ENCYCLOPEDIA OF

BIOPROCESS TECHNOLOGY: FERMENTATION, BIOCATALYSIS, AND BIOSEPARATION

VOLUMES 1 - 5

Michael C. Flickinger
University of Minnesota
St. Paul, Minnesota

Stephen W. Drew
Merck and Co., Inc.
Rahway, New Jersey

A Wiley-Interscience Publication
John Wiley & Sons, Inc.
New York / Chichester / Weinheim / Brisbane / Singapore / Toronto

This book is printed on acid-free paper. ♾

Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Flickinger, Michael C.
The encyclopedia of bioprocess technology : fermentation, biocatalysis, and bioseparation / Michael C. Flickinger, Stephen W. Drew.
p. cm.
Includes index.
ISBN 0-471-13822-3 (alk. paper)
1. Biochemical engineering--Encyclopedias. I. Drew, Stephen W., 1945- . II. Title.
TP248.3.F57 1999
660.6′03--dc21 99-11576
CIP

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

WILEY BIOTECHNOLOGY ENCYCLOPEDIAS

Encyclopedia of Bioprocess Technology: Fermentation, Biocatalysis, and Bioseparation
Edited by Michael C. Flickinger and Stephen W. Drew

Encyclopedia of Molecular Biology
Edited by Thomas E. Creighton

Encyclopedia of Cell Technology
Edited by Raymond E. Spier

Encyclopedia of Ethical, Legal, and Policy Issues in Biotechnology
Edited by Thomas J. Murray and Maxwell J. Mehlman

ENCYCLOPEDIA OF BIOPROCESS TECHNOLOGY: FERMENTATION, BIOCATALYSIS, AND BIOSEPARATION

EDITORIAL BOARD

PREFACE

The Wiley Biotechnology Encyclopedias, composed of the *Encyclopedia of Molecular Biology*; the *Encyclopedia of Bioprocess Technology: Fermentation, Biocatalysis, and Bioseparation*; the *Encyclopedia of Cell Technology*; and the *Encyclopedia of Ethical, Legal, and Policy Issues in Biotechnology* cover very broadly four major contemporary themes in biotechnology. The series comes at a fascinating time in that, as we move into the twenty-first century, the discipline of biotechnology is undergoing striking paradigm changes.

Biotechnology is now beginning to be viewed as an informational science. In a simplistic sense there are three types of biological information. First, there is the digital or linear information of our chromosomes and genes with the four-letter alphabet composed of G, C, A, and T (the bases guanine, cytosine, adenine, and thymine). Variation in the order of these letters in the digital strings of our chromosomes or our expressed genes (or mRNAs) generates information of several distinct types: genes, regulatory machinery, and information that enables chromosomes to carry out their tasks as informational organelles (e.g., centromeric and telomeric sequences).

Second, there is the three-dimensional information of proteins, the molecular machines of life. Proteins are strings of amino acids employing a 20-letter alphabet. Proteins pose four technical challenges: (*1*) Proteins are synthesized as linear strings and fold into precise three-dimensional structures as dictated by the order of amino acid residues in the string. Can we formulate the rules for protein folding to predict three-dimensional structure from primary amino acid sequence? The identification and comparative analysis of all human and model organism (bacteria, yeast, nematode, fly, mouse, etc.) genes and proteins will eventually lead to a lexicon of motifs that are the building block components of genes and proteins. These motifs will greatly constrain the shape space that computational algorithms must search to successfully correlate primary amino acid sequence with the correct three-dimensional shapes. The protein-folding problem will probably be solved within the next 10–15 years. (*2*) Can we predict protein function from knowledge of the three-dimensional structure? Once again the lexicon of motifs with their functional as well as structural correlations will play a critical role in solving this problem. (*3*) How do the myriad of chemical modifications of proteins (e.g., phosphorylation, acetylation, etc.) alter their structures and modify their functions? The mass spectrometer will play a key role in identifying secondary modifications. (*4*) How do proteins interact with one another and/or with other macromolecules to form complex molecular machines (e.g., the ribosomal subunits)? If these functional complexes can be isolated, the mass spectrometer, coupled with a knowledge of all protein sequences that can be derived from the complete genomic sequence of the organism, will serve as a powerful tool for identifying all the components of complex molecular machines.

The third type of biological information arises from complex biological systems and networks. Systems information is four dimensional because it varies with time. For example, the human brain has 1,012 neurons making approximately 1,015 connections. From this network arise systems properties such as memory, consciousness, and the ability to learn. The important point is that systems properties cannot be understood from studying the network elements (e.g., neurons) one at a time; rather the collective behavior of the elements needs to be studied. To study most biological systems, three issues need to be stressed. First, most biological systems are too complex to study directly, therefore they must be divided into tractable subsystems whose properties in part reflect those of the system. These subsystems must be sufficiently small to analyze all their elements and connections. Second, high-throughput analytic or global tools are required for studying many systems elements at one time (see later). Finally, the systems information needs to be modeled mathematically before systems properties can be predicted and ultimately understood. This will require recruiting computer scientists and applied mathematicians into biology—just as the attempts to decipher the information of complete genomes and the protein folding and structure/function problems have required the recruitment of computational scientists.

I would be remiss not to point out that there are many other molecules that generate biological information: amino acids, carbohydrates, lipids, and so forth. These too must be studied in the context of their specific structures and specific functions.

The deciphering and manipulation of these various types of biological information represent an enormous technical challenge for biotechnology. Yet major new and powerful tools for doing so are emerging.

One class of tools for deciphering biological information is termed high-throughput analytic or global tools. These tools can be used to study many genes or chromosome features (genomics), many proteins (proteomics), or many cells rapidly: large-scale DNA sequencing, genomewide genetic mapping, cDNA or oligonucleotide arrays, two-dimensional gel electrophoresis and other global protein separation technologies, mass spectrometric analysis of proteins and protein fragments, multiparameter, high-throughput cell and chromosome sorting, and high-throughput phenotypic assays.

A second approach to the deciphering and manipulation of biological information centers around combinatorial strategies. The basic idea is to synthesize an informational string (DNA fragments, RNA fragments, protein fragments, antibody combining sites, etc.) using all combinations of the basic letters of the corresponding alphabet, thus creating many different shapes that can be used to activate, inhibit, or complement the biological functions of designated three-dimensional shapes (e.g., a molecule in a signal transduction pathway). The power of combinational chemistry is just beginning to be appreciated.

A critical approach to deciphering biological information will ultimately be the ability to visualize the functioning of genes, proteins, cells, and other informational elements within living organisms (in vivo informational imaging).

Finally, there are the computational tools required to collect, store, analyze, model, and ultimately distribute the various types of biological information. The creation presents a challenge comparable to that of developing new instrumentation and new chemistries. Once again this means recruiting computer scientists and applied mathematicians to biology. The biggest challenge in this regard is the language barriers that separate different scientific disciplines. Teaching biology as an informational science has been a very effective means for breeching these barriers.

The challenge is, of course, to decipher various types of biological information and then be able to use this information to manipulate genes, proteins, cells, and informational pathways in living organisms to eliminate or prevent disease, produce higher-yield crops, or increase the productivity of animals for meat and other foods.

Biotechnology and its applications raise a host of social, ethical, and legal questions, for example, genetic privacy, germline genetic engineering, cloning of animals, genes that influence behavior, cost of therapeutic drugs generated by biotechnology, animal rights, and the nature and control of intellectual property.

Clearly, the challenge is to educate society so that each citizen can thoughtfully and rationally deal with these issues, for ultimately society dictates the resources and regulations that circumscribe the development and practice of biotechnology. Ultimately, I feel enormous responsibility rests with scientists to inform and educate society about the challenges as well as the opportunities arising from biotechnology. These are critical issues for biotechnology that are developed in detail in the *Encyclopedia of Ethical, Legal, and Policy Issues in Biotechnology*.

The view that biotechnology is an informational science pervades virtually every aspect of this science, including discovery, reduction to practice, and societal concerns. These Encyclopedias of Biotechnology reinforce the emerging informational paradigm change that is powerfully positioning science as we move into the twenty-first century to more effectively decipher and manipulate for humankind's benefit the biological information of relevant living organisms.

Leroy Hood
University of Washington

CONTRIBUTORS

Nicholas R. Abu-Absi, *University of Minnesota, St. Paul, Minnesota,* Cell Cycle, Eukaryotes

Luis A. Actis, *Miami University, Oxford, Ohio,* Plasmid DNA Replication

Michael W.W. Adams, *University of Georgia, Athens, Georgia,* Enzymes, Extremely Thermostable

P.S. Adams, *Trudeau Institute, Saranac Lake, New York,* Professional Societies, Association of Biomolecular Resource Facilities

William R. Adams, *Merck Manufacturing Division, West Point, Pennsylvania,* Adsorption, Protein, Batch

Eduardo Agosin, *Pontificia Universidad Católica de Chile, Santiago, Chile,* Solid Substrate Fermentation, Automation

Eric Ailor, *Johns Hopkins University, Baltimore, Maryland,* Insect Cells and Larvae, Gene Expression Systems

Kazuo Aisaka, *Kyowa Hakko Kogyo Co., Ltd., Tokyo, Japan,* Cholesterol Oxidase

Hiroyuki Akatsuka, *Tanabe Seiyaku Co., Ltd., Osaka, Japan,* Diltiazem Synthesis

Mohamed Al-Rubeai, *University of Birmingham, Birmingham, England,* Apoptosis; Cell Cycle

Seigo Amachi, *Hokkaido University, Sapporo, Japan,* Pyruvate, Production Using Defective ATPase Activity

Teruo Amachi, *Kyoto University, Kyoto, Japan,* β-Galactosidase, Enzymology and Applications

Graham Andrews, *Idaho National Engineering and Environmental Laboratory, Idaho Falls, Idaho,* Bioreactors, Gas Treatment

Ruth Hogue Angeletti, *Albert Einstein College of Medicine, Bronx, New York,* Professional Societies, Association of Biomolecular Resource Facilities

William Apel, *Idaho National Engineering and Environmental Laboratory, Idaho Falls, Idaho,* Bioreactors, Gas Treatment

Frances H. Arnold, *California Institute of Technology, Pasadena, California,* Enzymes, Directed Evolution

David L. Aronson, *Consultant, Bethesda, Maryland,* Coagulation Factors, Therapeutic

Yasuhisa Asano, *Toyama Prefectural University, Toyama, Japan,* D-Aminopeptidase, Alkaline D-Peptidase; Opine Dehydrogenase, Secondary Amine Dicarboxylic Acids; Phenylalanine Dehydrogenase; *Mitsubishi Chemical Co., Yokohama, Japan,* Malate, D-Malate

Yoshiro Ashina, *Nitto Chemical Industry Co., Ltd., Tokyo, Japan,* Nitrile Hydratase

John Aunins, *Merck and Company, West Point, Pennsylvania,* Roller Bottle Culture, Mixing

Hans Axelsson, *Alfa Laval Separation AB, Tumba, Sweden,* Cell Separation, Centrifugation

Irina Bagyan, *University of Connecticut Health Center, Farmington, Connecticut,* Gene Transfer, Gram-Positive Bacteria

Frederick S. Baker, *Westvaco Corporation, Charleston, South Carolina,* Activated Carbon, Decoloration of Pharmaceutical Products

Kym N. Baker, *University of Kent, Canterbury, U.K.,* Glycosylation of Recombinant Proteins

Richard H. Baltz, *CognoGen Enterprises, Indianapolis, Indiana,* Mutagenesis

A. Baradarajan, *Indian Institute of Technology–Madras, Chennai, India,* Dextran, Microbial Production Methods

Ana Paulina Barba de la Rosa, *University of California, Irvine, Irvine, California,* Crystallization, Bulk, Macromolecules

Claudia Bardouille, *ECACC CAMR Porton Down, Salisbury, Wiltshire, United Kingdom,* Animal Cells Used in Manufacturing

Ann-Kristin Barnfield Frej, *Amersham Pharmacia Biotech, Uppsala, Sweden,* Adsorption, Expanded Bed

Jonathan Basch, *Bristol-Myers Squibb Company, Syracuse, New York,* Cephalosporins

J.W. Bennett, *Tulane University, New Orleans, Louisiana,* Aspergillus

I.S. Bentley, *ABM Brewing and Enzymes Group / Rhone-Poulenc, Stockport, United Kingdom,* Enzymes, Starch Conversion

A.R. Bernard, *Serono Pharmaceutical Research Institute, Geneva, Switzerland,* Transient Expression Systems

Michael J. Betenbaugh, *Johns Hopkins University, Baltimore, Maryland,* Insect Cells and Larvae, Gene Expression Systems

Kami Beyzavi, *Bioprocessing Limited, Consett, County Durham, United Kingdom,* Adsorbents, Inorganic

John R. Birch, *Lonza Biologics PLC, Berkshire, United Kingdom,* Suspension Culture of Animal Cells

H.D. Blasey, *Serono Pharmaceutical Research Institute, Geneva, Switzerland,* Transient Expression Systems

Thomas Bley, *Dresden University of Technology, Dresden, Germany,* Flow Cytometry

Lynda F. Bonewald, *University of Texas Health Science Center, San Antonio, Texas,* Professional Societies, Association of Biomolecular Resource Facilities

C.J.-P. Boonaert, *Université Catholique de Louvain, Louvain-la-Neuve, Belgium,* Cell Separation, Flocculation

Michelle K. Bothwell, *Oregon State University, Corvallis, Oregon,* Adsorption, Proteins with Synthetic Materials

Joye Bramble, *Merck and Company, West Point, Pennsylvania,* Roller Bottle Culture, Mixing

R. Bruttini, *Criofarma Freeze-drying Equipment, Turin, Italy,* Freeze-drying, Pharmaceuticals

James D. Bryers, *University of Connecticut, Farmington, Connecticut,* Biofilms, Microbial

R. Buccholz, *Technical University of Berlin, Berlin, Germany,* Microencapsulation

Michael Butler, *University of Manitoba, Winnipeg, Canada,* Energy Metabolism, Microbial and Animal Cells

Eleanor Canova-Davis, *Genentech, San Francisco, California,* Professional Societies, Association of Biomolecular Resource Facilities

Ningjun Cao, *Purdue University, West Lafayette, Indiana,* Hemicellulose Conversion; Organic Compounds, Cellulose Conversion

Christine L. Case, *Skyline College, San Bruno, California,* Professional Societies, Society for Industrial Microbiology (SIM)

S.A. Casnocha, *Monsanto Co., St. Louis, Missouri,* Media, Animal Cell Culture

Francisco J. Castillo, *Berlex Biosciences, Richmond, California,* Hybridoma, Antibody Production

Marvin Charles, *Lehigh University, Bethlehem, Pennsylvania,* Fermenter Design

Peter S. J. Cheetham, *Zylepsis, Ltd., Ashford, United Kingdom,* Enzymes, for Flavor Production

Shu-Jen D. Chiang, *Bristol-Myers Squibb Company, Syracuse, New York,* Cephalosporins

Yusuf Chisti, *University of Almería, Almería, Spain,* Mass Transfer; Shear Sensitivity; Solid Substrate Fermentations, Enzyme Production, Food Enrichment

Jeffrey L. Cleland, *Genentech, Inc., South San Francisco, California,* Formulation and Delivery, Protein Pharmaceuticals

O. Colagrande, *Università Cattolica Sacro Cuore, Piacenza, Italy,* Wine Production

Attilio Converti, *Genoa University, Genoa, Italy,* Biofilters

Charles L. Cooney, *Massachusetts Institute of Technology, Cambridge, Massachusetts,* Bioreactors, Continuous Stirred-Tank Reactors

Athel Cornish-Bowden, *National Center for Scientific Research, Marseilles, France,* Kinetics, Enzymes

Rosalie J. Cote, *Becton Dickinson Microbiology Systems, Sparks, Maryland,* Media Composition, Microbial, Laboratory Scale; Medium Formulation and Design, *E. coli* and *Bacillus* spp.

John W. Crabb, *Cleveland Clinic Foundation, Cleveland, Ohio,* Professional Societies, Association of Biomolecular Resource Facilities

Steven M. Cramer, *Rensselaer Polytechnic Institute, Troy, New York,* Chromatography, Ion Exchange

Ronald L. Crawford, *University of Idaho, Moscow, Idaho,* Bioremediation

Jorge H. Crosa, *Oregon Health Sciences University, Portland, Oregon,* Plasmid DNA Replication

John Cullum, *LB Genetik, University of Kaiserslautern, Kaiserslautern, Germany,* Genetic Instability

Simon Cutting, *Royal Holloway University of London, Egham, Surrey, United Kingdom,* Gene Transfer, Gram-Positive Bacteria

Hans von Döhren, *Technical University Berlin, Berlin, Germany,* Peptide

Alessandro D'Aprano, *University of Rome, La Sapienza, Rome, Italy,* Conductivity

W.-D. Deckwer, *GBF-Gesellschaft für Biotechnologische Forschung GmbH, Braunschweig, Germany,* Xanthan Gum

Michael L. Dekleva, *Merck and Co., Inc., West Point, Pennsylvania,* Vaccine Technology

Arnold L. Demain, *Massachusetts Institute of Technology, Cambridge, Massachusetts,* Metabolites, Primary and Secondary

P.B. Dengis, *Université Catholique de Louvain, Louvain-la-Neuve, Belgium,* Cell Separation, Flocculation

C.P. Dillon, *C.P. Dillon & Associates, St. Albans, West Virginia,* Stainless Steels

Dennis Dobie, *Fluor Daniel, Marlton, New Jersey,* Heating, Ventilating, and Air Conditioning

D. Dochain, *Université Catholique de Louvain, Louvain-La-Neuve, Belgium,* Process Control, Strategy and Optimization

Ed Domanico, *Tri-Clover, Valencia, California,* Pumps, Industrial

John Dougherty, *Lilly Research Laboratories, Indianapolis, Indiana,* Professional Societies, Association of Biomolecular Resource Facilities

Alan Doyle, *Centre for Applied Microbiology and Research, Salisbury, Wilts, United Kingdom,* Culture Collections

Y.F. Dufrêne, *Université Catholique de Louvain, Louvain-la-Neuve, Belgium,* Cell Separation, Flocculation

C.C. Dupont-Gillain, *Université Catholique de Louvain, Louvain-la-Neuve, Belgium,* Cell Separation, Flocculation

H. Dziallas, *Technische Universität Braunschweig, Braunschweig, Germany,* Scale-Up, Stirred-Tank Reactors

Heinrich Ebner, *Linz, Austria,* Vinegar, Acetic Acid Production

G. Larry Eitel, *Jacobs Engineering Group, Denver, Colorado,* ASTM Standards for Biotechnology

Lynda B.M. Ellis, *University of Minnesota, Minneapolis, Minnesota,* Biocatalysis Databases

Takakazu Endo, *Nitto Chemical Industry Co., Ltd., Tokyo, Japan,* Nitrile Hydratase

Larry E. Erickson, *Kansas State University, Manhattan, Kansas,* Anaerobes

Bruno Fabiano, *University of Genoa, Genoa, Italy,* Corrosion, Microbial

Pierre Fauquenberque, *Gist-brocades, Seclin Cedex, France,* Enzymes, Fruit Juice Processing

Joseph Fernandez, *Rockefeller University, New York, New York,* Professional Societies, Association of Biomolecular Resource Facilities

J.-L. Flores Candia, *GBF-Gesellschaft für Biotechnologische Forschung GmbH, Braunschweig, Germany,* Xanthan Gum

Ian G. Fotheringham, *NSC Technologies, Mount Prospect, Illinois,* Phenylalanine

Beth Fowler, *Autoimmune, Inc., Lexington, Massachusetts,* Professional Societies, Association of Biomolecular Resource Facilities

Cornelius G. Friedrich, *University of Dortmund, Dortmund, Germany,* Waste Gas Cleaning, Biological

Roger Fuentes-Granados, *Iowa State University, Ames, Iowa,* Soybean (Fermentation, Meal Oil)

Daniel Y.C. Fung, *Kansas State University, Manhattan, Kansas,* Anaerobes

Sean R. Gallagher, *Motorola Phoenix Corporate Research Laboratories, Tempe, Arizona,* Electrophoresis of Proteins and Nucleic Acids

Subinay Ganguly, *Smith Kline Beecham Pharmaceuticals, King of Prussia, Pennsylvania,* Expression Systems, Mammalian Cells

F.A.P. Garcia, *University of Coimbra, Coimbra, Portugal,* Cell Disruption and Lysis

Alan R. Gardner, *Smith Kline Beecham, King of Prussia, Pennsylvania,* Process Validation

Maria Gavrilescu, *Research Centre for Antibiotics, Iaşi, Romania,* Static Mixing, in Fermentation Process

Craig J.L. Gershater, *S. B. Pharmaceuticals, Harlow, Essex, England,* Inoculum Preparation

Robert L. Gherna, *American Type Culture Collection, Rockville, Maryland,* Culture Preservation, Bacteria, Fungi, Yeast, and Cell Lines; Medium Formulation and Design, *E. coli* and *Bacillus* spp.

Siddartha Ghose, *Aston University, Birmingham, United Kingdom,* Protein Adsorption, Expanded Bed

Barbara Ghrist, *Genentech, San Francisco, California,* Professional Societies, Association of Biomolecular Resource Facilities

M. Gluz, *Ben-Gurion University of the Negev, Beer-Sheva, Israel,* Bioreactors, Airlift Reactors

Francesc Gòdia, *Universitat Autònoma de Barcelona, Barcelona, Spain,* Bioreactors, Fluidized-Bed

Victor Goetz, *Merck Research Labs, West Point, Pennsylvania,* Adsorption, Protein, Batch

C.S. Gong, *Purdue University, West Lafayette, Indiana,* Hemicellulose Conversion; Organic Compounds, Cellulose Conversion

Catherine Grassin, *Gist-brocades, Seclin Cedex, France,* Enzymes, Fruit Juice Processing

Thomas P. Graycar, *Genencor International Inc., Palo Alto, California,* Proteolytic Cleavage, Reaction Mechanisms

Lasse Greiner, *Forschungszentrum Jülich GmbH, Jülich, Germany,* Enzymes, Immobilized, Reactors

J. Bryan Griffiths, *Scientific Consultancy & Publishing, Salisbury, United Kingdom,* Mammalian Cell Culture Reactors, Scale-Up

E. Molina Grima, *University of Almería, Almería, Spain,* Microalgae, Mass Culture Methods

Tingyue Gu, *Ohio University, Athens, Ohio,* Chromatography, Radial Flow

Maria J. Guardia, *University of Minnesota, Minneapolis, Minnesota,* Mammalian Cell Bioreactors

S. Guillouet, *Massachusetts Institute of Technology, Cambridge, Massachusetts,* Corynebacteria, Brevibacteria

H. Håkanson, *Lund University, Lund, Sweden,* Sampling Methods (Reactors, Contamination)

Daniel D. Hanle, *Kendro Laboratory Products, Newtown, Connecticut,* Centrifuges, Animal Cells

Satoshi Hanzawa, *TOSOH Corp., Ayase-shi, Japan,* Aspartame; *Tokyo Research Laboratory, Hayakawa, Japan,* Thermolysin

Colin R. Harwood, *University of Newcastle upon Tyne, Newcastle upon Tyne, United Kingdom, Bacillus*

Junzo Hasegawa, *Kaneka Corporation, Hyogo, Japan,* Optically Active 1,2-Diols, Microbial Production by Stereoinversion

Gary Hathaway, *California Institute of Technology, Pasadena, California,* Professional Societies, Association of Biomolecular Resource Facilities

Holly Haughney, *Pall Corporation, Long Island, New York,* Filtration, Air; Filtration, Cartridge

J.J. Heijnen, *Delft University of Technology, Delft, The Netherlands,* Bioenergetics of Microbial Growth

Martin Held, *Swiss Federal Institute of Technology, Zürich, Switzerland, Pseudomonas,* Process Applications

Ulf Hellman, *Ludwig Institute for Cancer Research, Uppsala, Sweden,* Professional Societies, Association of Biomolecular Resource Facilities

D.C. Hempel, *Universität Gesamthochschule Paderborn, Paderborn, Germany,* Scale-Up, Stirred-Tank Reactors

Diane L. Hevehan, *Northwestern University, Evanston, Illinois,* Hypoxia, Effects on Animal Cells

S.M. Heydarian, *University College London, London, United Kingdom,* Rheology of Filamentous Microorganisms, Submerged Culture

Paula M. Hicks, *North Carolina State University, Raleigh, North Carolina,* Enzymes, Extremely Thermostable; Thermophilic Microorganisms

Sophia Hober, *Royal Institute of Technology, Stockholm, Sweden,* Affinity Fusions, Gene Expression

Ann C. Horan, *Schering Plough Research Institute, Kenilworth, New Jersey,* Secondary Metabolite Production, Actinomycetes, Other Than Streptomyces

Daslav Hranueli, *PLIVA d.d., Research Institute, Zagreb, Croatia,* Genetic Instability

Tsu-An Hsu, *Vaccine Pharmaceutical R&D, Merck & Co., Inc., West Point, Pennsylvania,* Insect Cells and Larvae, Gene Expression Systems

Wei-Shou Hu, *University of Minnesota, Minneapolis, Minnesota,* Mammalian Cell Bioreactors; Microcarrier Culture

Holger Huebner, *Technical University of Berlin, Berlin, Germany,* Microencapsulation

Tony Hunt, *Advanced Minerals Corporation, Santa Barbara, California,* Filter Aids

Takamitsu Iida, *Niigata University, Niigata, Japan,* Wastewater Treatment, Immobilized Cells

Yasuhiro Ikenaka, *Kaneka Corporation, Hyogo, Japan,* Aminohydrolases, for Production of D-Amino Acids

Masayuki Inui, *Mitsubishi Chemical Corporation, Ibaraki, Japan,* L-Isoleucine

A.P. Ison, *University College London, London, United Kingdom,* Rheology of Filamentous Microorganisms, Submerged Culture

Kathryn M. Ivanetich, *University of California–San Francisco, San Francisco, California,* Professional Societies, Association of Biomolecular Resource Facilities

Cornelius F. Ivory, *Washington State University, Pullman, Washington,* Electrophoresis, Proteins, Batch and Continuous

Mahendra K. Jain, *MBI International, Lansing, Michigan,* Anaerobes, Industrial Uses

David C. James, *University of Kent, Canterbury, United Kingdom,* Glycosylation of Recombinant Proteins

Nigel Jenkins, *Eli Lilly, Inc., Indianapolis, Indiana,* Protein Glycosylation

Steen Weber Jensen, *Novo Nordisk A/S, Bagsvaerd, Denmark,* Insulin, Purification

Michael R. Johns, *University of Queensland, Brisbane, Australia,* Crystallization, Proteins, Kinetics

Alois Jungbauer, *Institute of Applied Microbiology, Vienna, Austria,* Chromatography, Computer-Aided Design

Beth H. Junker, *Merck Research Laboratories, Rahway, New Jersey,* Good Manufacturing Practice (GMP) and Good Industrial Large Scale Practice (GLSP); Pilot Plants, Design and Operation

Oliver Kaltenbrunner, *Institute of Applied Microbiology, Vienna, Austria,* Chromatography, Computer-Aided Design

Manohar Kalyanpur, *Consultant, Bioseparations and Pharmaceutical Validation, Plaisir, France,* Membrane Separations

Per Kårsnäs, *Percreative, Göteborg, Sweden,* Chromatography, Hydrophobic Interaction

R.S. Karthikeyan, *Indian Institute of Technology–Madras, Chennai, India,* Dextran, Microbial Production Methods

Michihiko Kataoka, *Kyoto University, Kyoto, Japan,* Aldehyde Reductase; Lactonohydrolase; Pantothenic Acid and Related Compounds

Yasuo Kato, *Toyama Prefectural University, Toyama, Japan,* Opine Dehydrogenase, Secondary Amine Dicarboxylic Acids

Randal J. Kaufman, *University of Michigan Medical School, Ann Arbor, Michigan,* Secretion from Animal Cells

Tetsuya Kawakita, *Ajinomoto Co., Inc., Kanazawa, Japan,* Amino Acids, Glutamate

Takuo Kawamoto, *Kyoto University, Kyoto, Japan,* Cell Immobilization; Organosilicon Compounds

Robert M. Kelly, *North Carolina State University, Raleigh, North Carolina,* Enzymes, Extremely Thermostable; Thermophilic Microorganisms

Birgit Kessler, *Institute of Biotechnology, Zürich, Switzerland,* Poly(3-Hydroxyalkanoates)

Shun-ichi Kidokoro, *Sagami Chemical Research Center, Kanagawa, Japan,* Thermolysin

Shukuo Kinoshita, *Kyowa Hakko Kogyo Co., Ltd., Tokyo, Japan,* Glutamic Acid Producing Microorganisms; Production of L-Glutamic Acid

Horst Kleinkauf, *Technical University Berlin, Berlin, Germany,* Peptide

M.A. Klich, *U.S. Department of Agriculture, New Orleans, Louisiana, Aspergillus*

Miki Kobayashi, *Mitsubishi Chemical Corporation, Ibaraki, Japan,* L-Aspartic Acid

Arthur L. Koch, *Indiana University, Bloomington, Indiana,* Microbial Growth Measurement, Methods

Fragiskos N. Kolisis, *National Technical University of Athens, Athens, Greece,* Reverse Micelles, Enzymes

Hidenobu Komeda, *Toyama Prefectural University, Toyama, Japan,* D-Aminopeptidase, Alkaline D-Peptidase

Tadashi Kometani, *Toyama College of Technology, Toyama-shi, Japan,* Bioreduction; Yeast, Baker's

Udo Kragl, *Forschungszentrum Jülich GmbH, Jülich, Germany,* Biotransformation, Engineering Aspects; Enzymes, Immobilized, Reactors

M.R.V. Krishnan, *Anna University, Chennai, India,* Citric Acid, Processes

Maria-Regina Kula, *Heinrich Heine University Düsseldorf, Jülich, Germany,* Protein Purification, Aqueous Liquid Extraction

Hidehiko Kumagai, *Kyoto University, Kyoto, Japan,* Dihydroxyphenylalanine (4), Produced by Microorganisms; Tyrosine Phenol-lyase

Hidetoshi Kutsuki, *Kaneka Corporation, Hyogo, Japan,* Optically Active 1,2-Diols, Microbial Production by Stereoinversion

Andreas Lübbert, *Martin Luther University, Wittenberg, Germany,* Gas Hold-Up

David C. LaPorte, *University of Minnesota, Minneapolis, Minnesota,* Glyoxylate Bypass, Regulation

Xanthe M. Lam, *Genentech, Inc., South San Francisco, California,* Formulation and Delivery, Protein Pharmaceuticals

Per Larsen, *Novo Nordisk A/S, Bagsvaerd, Denmark,* Insulin, Purification

Michael E. Laska, *Massachusetts Institute of Technology, Cambridge, Massachusetts,* Bioreactors, Continuous Stirred-Tank Reactors

Anh LeDuy, *Laval University, Sainte-Foy, Canada,* Pullulan, Microbial Production Methods

David Leak, *Imperial College of Science, Technology, and Medicine, London, United Kingdom,* Methylotrophs, Industrial Applications

Ann L. Lee, *Merck Research Labs, West Point, Pennsylvania,* Adsorption, Protein, Batch

Gyun Min Lee, *Korea Advanced Institute of Science and Technology, Taejon, Korea,* Osmotic Stress, Secretion Rate

P.A. Lessard, *Massachusetts Institute of Technology, Cambridge, Massachusetts,* Corynebacteria, Brevibacteria

Andreas Liese, *Forschungszentrum Jülich, Jülich, Germany,* Biotransformations, Engineering Aspects

Carsten Lindemann, *University of Hannover, Hannover, Germany,* Fluorescence Techniques for Bioprocess Monitoring

M. Liu, *Rutgers University, Piscataway, New Jersey,* Roller Bottle Culture, Mixing

David R. Lloyd, *University of Birmingham, Birmingham, England,* Cell Cycle

Deryk T. Loo, *Bristol-Myers Squibb Pharmaceutical Research Institute, Seattle, Washington,* Attachment Factors

Shun-ichi Maemoto, *Kaneka Corporation, Hyogo, Japan,* Optically Active 1,2-Diols, Microbial Production by Stereoinversion

Savvas C. Makrides, *PRAECIS Pharmaceuticals, Inc., Cambridge, Massachusetts,* Expression Systems, *E. coli*

Stefan Marose, *University of Hannover, Hannover, Germany,* Fluorescence Techniques for Bioprocess Monitoring

R.E. Marquis, *University of Rochester, Rochester, New York,* Transport, Microbial Solute Uptake

Seiji Masuda, *Kyoto University, Kyoto, Japan,* Erythropoietin

Jennie P. Mather, *Rauen Biotechnologies, Inc., Mountain View, California,* Culture Media, Animal Cells, Large Scale Production

Hiroaki Matsumae, *Tanabe Seiyaku Co., Ltd., Osaka, Japan,* Diltiazem Synthesis

Ryuichi Matsumo, *Kyoto University, Kyoto, Japan,* Glucosidases

Tadashi Matsunaga, *Tokyo University of Agriculture and Technology, Tokyo, Japan,* Algal Culture

Ryuichi Matsuno, *Kyoto University, Kyoto, Japan,* Yeast, Baker's

B. Mattiasson, *Lund University, Lund, Sweden,* Sampling Methods (Reactors, Contamination)

Ian R. McDonald, *University of Warwick, Coventry, United Kingdom,* Methanotrophs

Joseph McGuire, *Oregon State University, Corvallis, Oregon,* Adsorption, Proteins with Synthetic Materials

Alexander McPherson, *University of California, Irvine, Irvine, California,* Crystallization, Bulk, Macromolecules

J.C. Merchuk, *Ben-Gurion University of the Negev, Beer-Sheva, Israel,* Bioreactors, Air-lift Reactors

Charles E. Miller, *Westvaco Corporation, Charleston, South Carolina,* Activated Carbon, Decoloration of Pharmaceutical Products

Stephen P. Miller, *University of Minnesota, Minneapolis, Minnesota,* Glyoxylate Bypass, Regulation

William M. Miller, *Northwestern University, Evanston, Illinois,* Hypoxia, Effects on Animal Cells

Sheenah Mische, *Rockefeller University, New York, New York,* Professional Societies, Association of Biomolecular Resource Facilities

David A. Mitchell, *Universidade Federal do Parana, Curitiba, Brazil,* Solid-State Fermentation, Microbial Growth Kinetics

Inger Mollerup, *Novo Nordisk A/S, Bagsvaerd, Denmark,* Insulin, Purification

Alison Moore, *Amgen, Inc., Newbury Park, California,* Culture Media, Animal Cells, Large Scale Production

Vadim V. Mozhaev, *University of Iowa, Iowa City, Iowa,* Denaturation, Proteins, Solvent Mediated

Susann Müller, *University of Leipzig, Leipzig, Germany,* Flow Cytometry

David W. Murhammer, *University of Iowa, Iowa City, Iowa,* Pluronic Polyols, Cell Protection

J. Colin Murrell, *University of Warwick, Coventry, United Kingdom,* Methanotrophs

F.J. Muzzio, *Rutgers University, Piscataway, New Jersey,* Roller Bottle Culture, Mixing

Clayton Naeve, *St. Jude Children's Research Hospital, Memphis, Tennessee,* Professional Societies, Association of Biomolecular Resource Facilities

Kazuhiro Nakanishi, *Okayama University, Okayama, Japan,* Membrane Surface Liquid Culture, Microorganisms, Fungi

Toru Nakayama, *Tohoku University, Sendai, Japan,* β-Galactosidase, Enzymology and Applications

M.P. Nandakumar, *Lund University, Lund, Sweden,* Sampling Methods (Reactors, Contamination)

Venkatesh Natarajan, *Rensselaer Polytechnic Institute, Troy, New York,* Chromatography, Ion Exchange

Robert Newburgh, *The Protein Society, Bethesda, Maryland,* Professional Societies, The Protein Society

Ronald L. Niece, *University of California–Irvine, Irvine, California,* Professional Societies, Association of Biomolecular Resource Facilities

Hans Ole Nielsen, *DELTA Light&Optics, Lyngby, Denmark,* Fluorescence Techniques for Bioprocess Monitoring

Jens Nielsen, *Technical University of Denmark, Lyngby, Denmark,* Fermentation Monitoring, Design and Optimization

Zivko L. Nikolov, *Iowa State University, Ames, Iowa,* Soybean (Fermentation, Meal Oil)

Joakim Nilsson, *Royal Institute of Technology, Stockholm, Sweden,* Affinity Fusions, Gene Expression

Yasuo Ninomiya, *Nitto Denko Company, Osaka, Japan,* Insecticides, Microbial Production; Optical Resolution, Biocatalysis

Takuo Nishida, *Tanabe Seiyaku Co., Ltd., Osaka, Japan,* Diltiazem Synthesis, Microbial Asymmetric Reduction

P.T. Noble, *Fluor Daniel GmbH, Wiesbaden, Germany,* Sterilization-in-Place

Per-Åke Nygren, *Royal Institute of Technology, Stockholm, Sweden,* Affinity Fusions, Gene Expression

Jun Ogawa, *Kyoto University, Kyoto, Japan,* Oils, Microbial Production

Masahiro Ogura, *Kaneka Corporation, Hyogo, Japan,* Optically Active 1,2-Diols, Microbial Production by Stereoinversion

Takehisa Ohashi, *Kaneka Corporation, Hyogo, Japan,* Optically Active 1,2-Diols, Microbial Production by Stereoinversion

Hiromichi Ohta, *Keio University, Yokohama, Japan,* Lactones, Biocatalytic Synthesis

Tetsuo Oka, *Kyowa Hakko Kogyo, Co., Ltd., Tokyo, Japan,* Amino Acids, Production Processes

Hans Sejr Olsen, *Novo Nordisk, Bagsvaerd, Denmark,* Enzymes, Protein Hydrolysis

Gustaf Olsson, *Lund Institute of Technology, Lund, Sweden,* Waste Treatment, Activated Sludge, Control Strategies

Tetsuo Omata, *Nitto Denko Company, Osaka, Japan,* Insecticides, Microbial Production; Optical Resolution, Biocatalysis

Sadettin S. Ozturk, *Bayer Corporation, Berkeley, California,* Ammonia Toxicity, Animal Cells

J. Ricardo Pérez-Correa, *Pontificia Universidad Católica de Chile, Santiago, Chile,* Solid Substrate Fermentation, Automation

Philip Packer, *Centre for Applied Microbiology and Research, Salisbury, Wilts, United Kingdom,* Culture Collections

Nicolai S. Panikov, *Institute of Microbiology, Russian Academy of Sciences, Moscow, Russian Federation,* Kinetics, Microbial Growth

Patrizia Perego, *University of Genoa, Genoa, Italy,* Corrosion, Microbial

Madhusudan V. Peshwa, *Dendreon Corp., Mountain View, California,* Microcarrier Culture

Joerg Peters, *Bayer AG, Wuppertal, Germany,* Cofactor Regeneration, Nicotinamide Coenzymes

David Pollard, *Merck and Co., Inc., Rahway, New Jersey,* Vent Gas Analysis

Kaisa Poutanen, *VTT Biotechnology and Food Research, Espoo, Finland,* Hemicellulases

Joan Qi Si, *Novo Nordisk Ferment Ltd., Dittingen, Switzerland,* Enzymes, Baking, Bread Making

Bruce Ramsay, *Queen's University, Kingston, Ontario, Canada,* Polyhydroxyalkanoates, Separation, Purification, and Manufacturing Methods

Juliana Ramsay, *Queen's University, Kingston, Ontario, Canada,* Polyhydroxyalkanoates, Separation, Purification, and Manufacturing Methods

Douglas W. Rea, *Cytogen Corp., Princeton, New Jersey,* Antibody Purification

Kenneth F. Reardon, *Colorado State University, Fort Collins, Colorado,* Fluorescence Techniques for Bioprocess Monitoring

Matthias Redenbach, *LB Genetik, University of Kaiserslautern, Kaiserslautern, Germany,* Genetic Instability

Hans Reichenbach, *Gesellschaft für Biotechnologische Forschung, Braunschweig, Germany,* Myxobacteria

Harold Reisman, *Biotechnology Results, Weston, Connecticut,* Economics

Albert J. Repik, *Westvaco Corporation, Charleston, South Carolina,* Activated Carbon, Decoloration of Pharmaceutical Products

Nadine Ritter, *Abbott Laboratories, Abbott Park, Illinois,* Professional Societies, Association of Biomolecular Resource Facilities

Palmer Rogers, *University of Minnesota, Minneapolis, Minnesota,* Clostridia, Solvent Formation

Rodica-Viorica Roman, *Chemical Pharmaceutical Research Institute, Iaşi, Romania,* Transfer Phenomena in Multiphase Systems in Mixing Vessels

Charles M. Roth, *Massachusetts General Hospital, Shriners Burns Hospital, and Harvard Medical School, Boston, Massachusetts,* Chromatography, Size Exclusion

P.G. Rouxhet, *Université Catholique de Louvain, Louvain-la-Neuve, Belgium,* Cell Separation, Flocculation

Joon Soo Ryu, *Korea Advanced Institute of Science and Technology, Taejon, Korea,* Osmotic Stress, Secretion Rate

Peter Salmon, *Merck and Co., Inc., Rahway, New Jersey,* Vent Gas Analysis

Gautam Sanyal, *Astra Research Center Boston, Cambridge, Massachusetts,* Thermal Unfolding, Proteins

Ryuzo Sasaki, *Kyoto University, Kyoto, Japan,* Erythropoietin

Tadashi Sato, *Tanabe Seiyaku Co., Ltd., Osaka, Japan,* Enzymes, Immobilization Methods; Malic Acid, Production by Fumarase; Production of L-Amino Acids by Aminoacylase

Catherine H. Schein, *University of Texas Medical Branch, Galveston, Texas,* Protein Aggregation, Denaturation; Protein Expression, Soluble

Thomas Scheper, *University of Hannover, Hannover, Germany,* Fluorescence Techniques for Bioprocess Monitoring

S. Schiewer, *McGill University, Montreal, Canada,* Biosorption, Metals

Andrew Schmid, *Swiss Federal Institute of Technology, Zurich, Switzerland, Pseudomonas,* Process Applications

Ole Schou, *Novo Nordisk A/S, Bagsvaerd, Denmark,* Insulin, Purification

Robert Schwartz, *Abbott Laboratories, North Chicago, Illinois,* Professional Societies, Society for Industrial Microbiology (SIM)

Henry Schwartzberg, *University of Massachusetts, Amherst, Massachusetts,* Food Process Engineering

James Searles, *Merck and Company, West Point, Pennsylvania,* Roller Bottle Culture, Mixing

Katsuya Seguro, *Food Research and Development Lab, Kawaski, Japan,* Transglutaminase

Klaus Selber, *Heinrich Heine University Düsseldorf, Jülich, Germany,* Protein Purification, Aqueous Liquid Extraction

Sylvia Sellmer-Wilsberg, *Bad Honnef, Germany,* Vinegar, Acetic Acid Production

Shuji Senda, *Nitto Denko Company, Osaka, Japan,* Insecticides, Microbial Production; Optical Resolution, Biocatalysis

Pir M. Shah, *De Montfort University, Leicester, United Kingdom,* Protein Glycosylation

P. Ayazi Shamlou, *University College London, London, United Kingdom,* Rheology of Filamentous Microorganisms, Submerged Culture

Bhav P. Sharma, *CV Therapeutics, Palo Alto, California,* Cell Separation, Sedimentation

Allan R. Shatzman, *SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania,* Expression Systems, Mammalian Cells

Li Shi, *Merck Research Laboratories, West Point, Pennsylvania,* Thermal Unfolding, Proteins

Takeji Shibatani, *Tanabe Seiyaku Co., Ltd., Osaka, Japan,* Diltiazem Synthesis; Diltiazem Synthesis, Microbial Asymmetric Reduction

Sakayu Shimizu, *Kyoto University, Kyoto, Japan,* Aldehyde Reductase; Lactonohydrolase; Oils, Microbial Production; Pantothenic Acid and Related Compounds

Suteaki Shioya, *Osaka University, Osaka, Japan,* Mixed Culture

M. S. Showell, *Procter and Gamble, Cincinnati, Ohio,* Enzymes, Detergent

Rabinder P. Singh, *University of Birmingham, Birmingham, England,* Apoptosis

Satinder K. Singh, *University of Minnesota, Minneapolis, Minnesota,* Glyoxylate Bypass, Regulation

A.J. Sinskey, *Massachusetts Institute of Technology, Cambridge, Massachusetts,* Corynebacteria, Brevibacteria

Augustine Smith, *Abbott Laboratories, Abbott Park, Illinois,* Professional Societies, Association of Biomolecular Resource Facilities

Thomas M. Smith, *Smith Kline Beecham, King of Prussia, Pennsylvania,* Process Validation

Leo Snel, *Novo Nordisk A/S, Bagsvaerd, Denmark,* Insulin, Purification

Gail Sofer, *BioReliance, Rockville, Maryland,* Cleaning, Cleaning Validation

Carles Solà, *Universitat Autònoma de Barcelona, Barcelona, Spain,* Bioreactors, Fluidized-Bed

Richard Sparling, *University of Manitoba, Winnipeg, Canada,* Energy Metabolism, Microbial and Animal Cells

Dave Speicher, *Wistar Institute, Philadelphia, Pennsylvania,* Professional Societies, Association of Biomolecular Resource Facilities

Koti Sreekrishna, *Procter and Gamble Co., Ross, Ohio, Pichia,* Optimization of Protein Expression

Friedrich Srienc, *University of Minnesota, St. Paul, Minnesota,* Cell Cycle, Eukaryotes

Stefan Ståhl, *Royal Institute of Technology, Stockholm, Sweden,* Affinity Fusions, Gene Expression

Keith Stephenson, *University of Newcastle upon Tyne, Newcastle upon Tyne, United Kingdom, Bacillus*

Kathryn Stone, *Yale University, New Haven, Connecticut,* Professional Societies, Association of Biomolecular Resource Facilities

Bob Stover, *Tri-Clover, Valencia, California,* Pumps, Industrial

William R. Strohl, *Merck Research Laboratories, Rahway, New Jersey,* Secondary Metabolites, Antibiotics

Daniel J. Strydom, *Bio Nebraska, Inc., Lincoln, Nebraska,* Professional Societies, Association of Biomolecular Resource Facilities

Deidre M. Stuart, *Queensland University of Technology, Brisbane, Australia,* Solid-State Fermentation, Microbial Growth Kinetics

John Stults, *Genentech, San Francisco, California,* Professional Societies, Association of Biomolecular Resource Facilities

Takeshi Sugai, *Keio University, Yokohama, Japan,* Lactones, Biocatalytic Synthesis

Masaru Suto, *Nitto Chemical Industry Co., Ltd., Tokyo, Japan,* Nitrile Hydratase

T. Swaminathan, *Indian Institute of Technology–Madras, Chennai, India,* Dextran, Microbial Production Methods

P.D. Swanson, *Rutgers University, Piscataway, New Jersey,* Roller Bottle Culture, Mixing

Kristine M. Swiderek, *ZymoGenetics, Seattle, Washington,* Professional Societies, Association of Biomolecular Resource Facilities

Satomi Takahashi, *Kaneka Corporation, Hyogo, Japan,* Aminohydrolases, for Production of D-Amino Acids; Phenylglycines, D-Phenylglycines

Hiroyuki Takano, *Chichibu Onoda Cement Corporation, Chiba-prefecture, Japan,* Algal Culture

Haruko Takeyama, *Tokyo University of Agriculture and Technology, Tokyo, Japan,* Algal Culture

Atsuo Tanaka, *Kyoto University, Kyoto, Japan,* Cell Immobilization

Robert D. Tanner, *Vanderbilt University, Nashville, Tennessee,* Solid-State Fermentation, Microbial Growth Kinetics

Maija Tenkanen, *VTT Biotechnology and Food Research, Espoo, Finland,* Hemicellulases

Masato Terasawa, *Mitsubishi Chemical Corporation, Ibaraki, Japan,* L-Aspartic Acid; L-Isoleucine

Jörg Thömmes, *Heinrich-Heine Universität Düsseldorf, Jülich, Germany,* Membrane Chromatography

Jules Thibault, *Laval University, Sainte-Foy, Canada,* Pullulan, Microbial Production Methods

J.S. Tolan, *Iogen Corporation, Ottawa, Canada,* Enzymes, Pulp and Paper Processing

E. Donald Tolles, *Westvaco Corporation, Charleston, South Carolina,* Activated Carbon, Decoloration of Pharmaceutical Products

Marcelo E. Tolmasky, *California State University, Fullerton, Fullerton, California,* Plasmid DNA Replication

Fusao Tomita, *Hokkaido University, Sapporo, Japan,* Pyruvate, Production Using Defective ATPase Activity

Tetsuya Tosa, *Tanabe Seiyaku Co. Ltd., Osaka, Japan,* Enzymes, Immobilization Methods; Malic Acid, Production by Fumarase; Production of L-Amino Acids by Aminoacylase

R. Reid Townsend, *University of California–San Francisco, San Francisco, California,* Professional Societies, Association of Biomolecular Resource Facilities

Mario R. Tredici, *University of Florence, Florence, Italy,* Bioreactors, Photo

G.T. Tsao, *Purdue University, West Lafayette, Indiana,* Hemicellulose Conversion; Organic Compounds, Cellulose Conversion

Satoru Tsuda, *Kaneka Corporation, Hyogo, Japan,* Optically Active 1,2-Diols, Microbial Production by Stereoinversion

Radu Z. Tudose, *Technical University Gh.Asachi Iaşi, Iaşi, Romania,* Static Mixing, in Fermentation Process

A.P.F. Turner, *Cranfield Biotechnology Centre, Bedfordshire, United Kingdom,* Process Monitoring

Makoto Ueda, *Mitsubishi Chemical Co., Yokohama, Japan,* Malate, D-Malate

Mathias Uhlén, *Royal Institute of Technology, Stockholm, Sweden,* Affinity Fusions, Gene Expression

Michiel E. Ultee, *Cytogen Corp., Princeton, New Jersey,* Antibody Purification

D.R. Unger, *Rutgers University, Piscataway, New Jersey,* Roller Bottle Culture, Mixing

Mahesh Uttamlal, *Glasgow Caledonian University, Glasgow, Scotland,* Optical Sensors

Robert van Reis, *Genentech, Inc., South San Francisco, California,* Protein Ultrafiltration

James Vaughn, *ARS, USDA, Beltsville, Maryland,* Insect Cell Culture, Protein Expression

Liisa Viikari, *VTT Biotechnology and Food Research, Espoo, Finland,* Hemicellulases

B. Volesky, *McGill University, Montreal, Canada,* Biosorption, Metals

David B. Volkin, *Merck Research Laboratories, West Point, Pennsylvania,* Thermal Unfolding, Proteins

Lawrence P. Wackett, *University of Minnesota, Minneapolis, Minnesota,* Biocatalysis Databases

David R. Walt, *Tufts University, Medford, Massachusetts,* Optical Sensors

Christian Wandrey, *Forschungszentrum Jülich GmbH, Jülich, Germany,* Enzymes, Immobilized, Reactors

Sally Warburton, *ECACC, Wiltshire, United Kingdom,* Human and Primate Cell Lines

Udo Werner, *University of Dortmund, Dortmund, Germany,* Waste Gas Cleaning, Biological

David White, *Indiana University, Bloomington, Indiana,* Electron Transport

S.F. White, *Cranfield Biotechnology Centre, Bedfordshire, United Kingdom,* Process Monitoring

Erik M. Whiteley, *Johns Hopkins University, Baltimore, Maryland,* Insect Cells and Larvae, Gene Expression Systems

Ken Williams, *Yale University, New Haven, Connecticut,* Professional Societies, Association of Biomolecular Resource Facilities

L.B. Willis, *Massachusetts Institute of Technology, Cambridge, Massachusetts,* Corynebacteria, Brevibacteria

Jack Wilson, *ABEC, Inc., Allentown, Pennsylvania,* Fermenter Design

Patrick L. Wintrode, *California Institute of Technology, Pasadena, California,* Enzymes, Directed Evolution

Bernard Witholt, *Swiss Federal Institute of Technology, Zürich, Switzerland,* Poly(3-Hydroxyalkanoates); *Pseudomonas,* Process Applications

K.D. Wittrup, *Kyowa Hakko Kogyo Co., Ltd., Tokyo, Japan,* Protein Secretion, *Saccharomyces cerevisiae*

R.A. Wolfe, *Monsanto Co., St. Louis, Missouri,* Media, Animal Cell Culture

Kathy Wong, *University of Minnesota, Minneapolis, Minnesota,* Microcarrier Culture

Marcel Wubbolts, *Swiss Federal Institute of Technology, Zürich, Switzerland, Pseudomonas,* Process Applications

Florian M. Wurm, *Swiss Federal Institute of Technology, Lausanne, Switzerland,* Chinese Hamster Ovary Cells, Recombinant Protein Production

Xinfa Xiao, *Apotex Fermentation, Inc., Winnipeg, Canada,* Energy Metabolism, Microbial and Animal Cells

Feng Xu, *Novo Nordisk Biotech, Davis, California,* Laccase

Hideaki Yamada, *Mitsubishi Chemical Co., Yokohama, Japan,* Malate, D-Malate

Tsuneo Yamane, *Nagoya University, Nagoya, Japan,* Monoacylglycerols

Martin L. Yarmush, *Massachusetts General Hospital, Shriners Burns Hospital, and Harvard Medical School, Boston, Massachusetts,* Chromatography, Size Exclusion

Atsushi Yokota, *Hokkaido University, Sapporo, Japan,* Pyruvate, Production Using Defective ATPase Activity

Hideaki Yukawa, *Mitsubishi Chemical Corporation, Ibaraki, Japan,* L-Aspartic Acid; L-Isoleucine

Dane W. Zabriskie, *Smith Kline Beecham Pharmaceuticals, King of Prussia, Pennsylvania,* Process Validation

J. Gregory Zeikus, *MBI International, Lansing, Michigan,* Anaerobes, Industrial Uses

Mario Zilli, *Genoa University, Genoa, Italy,* Biofilters

Marko Zlokarnik, *Graz, Austria,* Dimensional Analysis, Scale-Up

Andrew L. Zydney, *University of Delaware, Newark, Delaware,* Protein Ultrafiltration

ACTIVATED CARBON, DECOLORATION OF PHARMACEUTICAL PRODUCTS

FREDERICK S. BAKER
CHARLES E. MILLER
ALBERT J. REPIK
E. DONALD TOLLES
Westvaco Corporation
Charleston, South Carolina

KEY WORDS

Activated carbon
Adsorption isotherms
Adsorption of radionuclides
BET surface area
Emissions, motor vehicles
Food, beverage, and cooking oil
Pharmaceuticals
Pore volume
Powdered carbon regeneration
Solvent recovery

OUTLINE

INTRODUCTION

Activated carbon is a predominantly amorphous solid that has an extraordinarily large internal surface area and pore volume. These unique characteristics are responsible for its adsorptive properties, which are exploited in many different liquid- and gas-phase applications. Activated carbon is an exceptionally versatile adsorbent because the size and distribution of the pores within the carbon matrix can be controlled to meet the needs of current and emerging markets (1). Engineering requirements of specific applications are satisfied by producing activated carbons in the form of powders, granules, and shaped products. Through choice of precursor, method of activation, and control of processing conditions, the adsorptive properties of products are tailored for applications as diverse as the purification of potable water and the control of emissions from bioproduct recovery processes.

In 1900, two very significant processes in the development and manufacture of activated carbon products were patented (2). The first commercial products were produced in Europe under these patents: Eponite, from wood in 1909, and Norit, from peat in 1911. Activated carbon was first produced in the United States in 1913 by Westvaco Corp. under the name Filtchar, using a by-product of the papermaking process (3). Further milestones in development were reached as a result of World War I. In response to the need for protective gas masks, a hard, granular activated carbon was produced from coconut shell in 1915. Following the war, large-scale commercial use of activated carbon was extended to refining of beet sugar and corn syrup and to purification of municipal water supplies (4). The termination of the supply of coconut char from the Philippines and India during World War II forced the domestic development of granular activated carbon products from coal in 1940 (5). More recent innovations in the manufacture and use of activated carbon products have been driven by the need to recycle resources and to prevent environmental pollution.

PHYSICAL AND CHEMICAL PROPERTIES

The structure of activated carbon is best described as a twisted network of defective carbon layer planes, cross-linked by aliphatic bridging groups (6). X-ray diffraction

patterns of activated carbon reveal that it is nongraphitic, remaining amorphous because the randomly cross-linked network inhibits reordering of the structure even when heated to 3,000 °C (7). This property of activated carbon contributes to its most unique feature, namely, the highly developed and accessible internal pore structure. The surface area, dimensions, and distribution of the pores depend on the precursor and on the conditions of carbonization and activation. Pore sizes are classified (8) by the International Union of Pure and Applied Chemistry (IUPAC) as micropores (pore width <2 nm), mesopores (pore width 2–50 nm), and macropores (pore width >50 nm).

The surface area of activated carbon is usually determined by application of the Brunauer-Emmett-Teller (BET) model of physical adsorption (9,10) using nitrogen as the adsorptive (8). Typical commercial products have specific surface areas in the range 500–2,000 m^2/g, but values as high as 3,500–5,000 m^2/g have been reported for some activated carbons (11,12). In general, however, the effective surface area of a microporous activated carbon is far smaller because the adsorption of nitrogen in micropores does not occur according to the process assumed in the BET model, which results in unrealistically high values for surface area (10,13). Adsorption isotherms are usually determined for the appropriate adsorptives to assess the effective surface area of a product in a specific application. Adsorption capacity and rate of adsorption depend on the internal surface area and distribution of pore size and shape but are also influenced by the surface chemistry of the activated carbon (14). The macroporosity of the carbon is important for the transfer of adsorbate molecules to adsorption sites within the particle.

Functional groups are formed during activation by interaction of free radicals on the carbon surface with atoms such as oxygen and nitrogen, both from within the precursor and from the atmosphere (15). The functional groups render the surface of activated carbon chemically reactive and influence its adsorptive properties (6). Activated carbon is generally considered to exhibit a low affinity for water, which is an important property with respect to the adsorption of gases in the presence of moisture (16). However, the functional groups on the carbon surface can interact with water, rendering the carbon surface more hydrophilic (15). Surface oxidation, which is an inherent feature of activated carbon production, results in hydroxyl, carbonyl, and carboxylic groups that impart an amphoteric character to the carbon, so that it can be either acidic or basic. The electrokinetic properties of an activated carbon product are, therefore, important with respect to its use as a catalyst support (17). As well as influencing the adsorption of many molecules, surface oxide groups contribute to the reactivity of activated carbons toward certain solvents in solvent recovery applications (18).

In addition to surface area, pore size distribution, and surface chemistry, other important properties of commercial activated carbon products include pore volume, particle size distribution, apparent or bulk density, particle density, abrasion resistance, hardness, and ash content. The range of these and other properties is illustrated in Table 1 together with specific values for selected commercial grades of powdered, granular, and shaped activated carbon products used in liquid- or gas-phase applications (19).

MANUFACTURE AND PROCESSING

Commercial activated carbon products are produced from organic materials that are rich in carbon, particularly coal, lignite, wood, nut shells, peat, pitches, and cokes. The choice of precursor is largely dependent on its availability, cost, and purity, but the manufacturing process and intended application of the product are also important considerations. Manufacturing processes fall into two categories, thermal activation and chemical activation. The effective porosity of activated carbon produced by thermal activation is the result of gasification of the carbon at relatively high temperatures (20), but the porosity of chemically activated products is generally created by chemical dehydration reactions occurring at significantly lower temperatures (1,21).

Thermal Activation Processes

Thermal activation occurs in two stages: thermal decomposition or carbonization of the precursor and controlled gasification or activation of the crude char. During carbonization, elements such as hydrogen and oxygen are eliminated from the precursor to produce a carbon skeleton possessing a latent pore structure. During gasification, the char is exposed to an oxidizing atmosphere that greatly increases the pore volume and surface area of the product through elimination of volatile pyrolysis products and from carbon burn-off. Carbonization and activation of the char are generally carried out in direct-fired rotary kilns or multiple-hearth furnaces, but fluidized-bed reactors have also been used (22). Materials of construction, notably steel and refractories, are designed to withstand the high-temperature conditions (i.e., >1,000 °C) inherent in activation processes. The thermal activation process is illustrated in Figure 1 for the production of activated carbon from bituminous coal (23,24).

Bituminous coal is pulverized and passed to a briquette press. Binders may be added at this stage before compression of the coal into briquettes. The briquetted coal is then crushed and passed through a screen, from which the on-size material passes to an oxidizing kiln. Here, the coking properties of the coal particles are destroyed by oxidation at moderate temperatures in air. The oxidized coal is then devolatilized in a second rotary kiln at higher temperatures under steam. To comply with environmental pollution regulations, the kiln off-gases containing dust and volatile matter pass through an incinerator before discharge to the atmosphere.

The devolatilized coal particles are transported to a direct-fired multihearth furnace where they are activated by holding the temperature of the furnace at about 1,000 °C. Product quality is maintained by controlling coal feed rate and bed temperature. As before, dust particles in the furnace off-gas are combusted in an afterburner before discharge of the gas to the atmosphere. Finally, the granular product is screened to provide the desired particle size. A

Table 1. Properties of Selected U.S. Activated Carbon Products

Property	Manufacturer Precursor Product grade Product form Typical range	Gas-phase carbons Calgon coal BPL granular	 Norit peat B4 extruded	 Westvaco wood WV-A 1100 granular	Liquid-phase carbons Calgon coal SGL granular	 Norit peat SA 3 powdered	 Westvaco wood SA-20 powdered
Particle size, U.S. mesh[a,b]	<4	12 × 30	3.8 mm	10 × 25	8 × 30	60% <325	65–85% <325
Apparent density, g/cm^3	0.2–0.6	>0.48	0.43	0.27	0.52	0.46	0.34–0.37
Particle density, g/cm^3	0.4–0.9	0.80		0.50	0.80		
Hardness number	50–100	>90	99				
Abrasion number					>75		
Ash, wt %	1–20	<8	6		<10	6	3–5
BET surface area, N_2 m^2/g	500–2,500	1,050–1,150	1,100–1,200	1750	900–1,000	750	1,400–1,800
Total pore volume, cm^3/g	0.5–2.5	0.8	0.9	1.2	0.85		2.2–2.5
CCl_4 activity, wt%	35–125	>60					
Butane working capacity, g/100 cm^3	4–14			>11.0			
Iodine number	500–1,200	>1,050			>900	800	>1,000
Decolorizing index Westvaco	15–25						>20
Molasses number Calgon	50–250				>200		
Norit	300–1,500					440	
Heat Capacity at 100 °C, J/(g K)[c]	0.84–1.3	1.05			1.05		
Thermal conductivity, W/(m·K)	0.05–0.10						

Note: Specific values shown are those cited in manufacturers' product literature (19). Typical ranges shown are based on values reported in the open literature.
[a]Unless otherwise noted.
[b]Approximate mm corresponding to cited meshes are (mesh:mm 4:4.76; 8:2.38; 10:2; 12:1.68; 25:0.72; 30:0.59; 325:0.04.
[c]To convert J to cal, divide by 4.184.

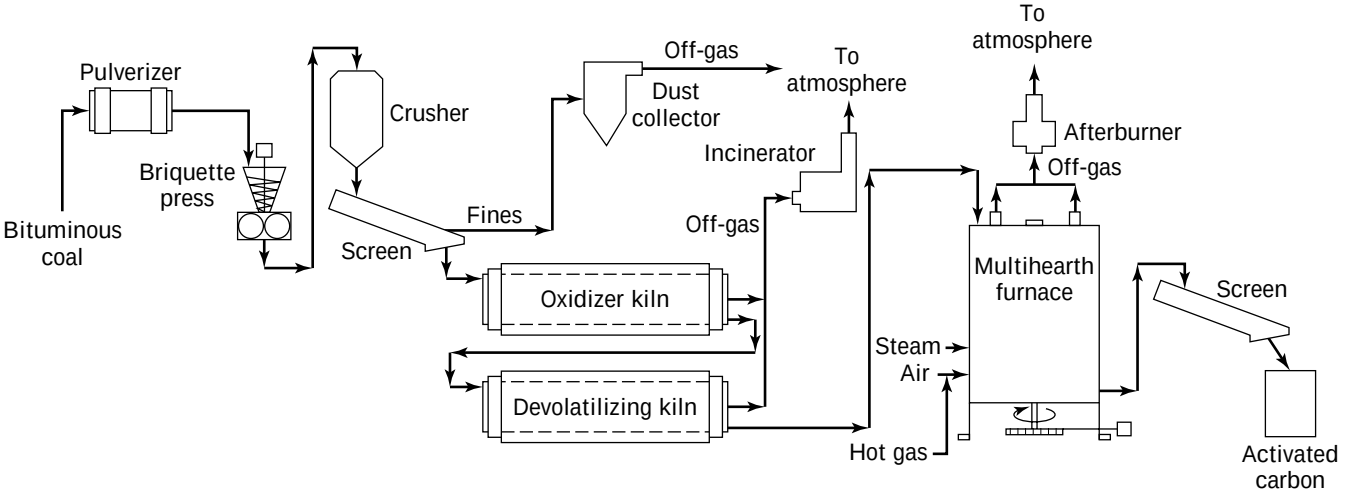


Figure 1. Thermal activation of bituminous coal.

typical yield of activated carbon is about 30–35% by weight based on the raw coal.

The process for the thermal activation of other carbonaceous materials is modified according to the precursor. For example, the production of activated carbon from coconut shell does not require the stages involving briquetting, oxidation, and devolatilization. To obtain a high-activity product, however, it is important that the coconut shell is charred slowly prior to activation of the char. In some processes, the precursor or product is acid-washed to obtain a final product with a low ash content (23,25).

Chemical Activation Processes

In contrast to the thermal activation of coal, chemical activation is generally carried out commercially in a single

kiln. The precursor, usually wood, is impregnated with a chemical activation agent, typically phosphoric acid, and the blend is heated to a temperature of 450–700 °C (26). Chemical activation agents reduce the formation of tar and other by-products, thereby increasing carbon yield. The chemical activation process is illustrated in Figure 2 for the production of granular activated carbon from wood (23,27).

Sawdust is impregnated with concentrated phosphoric acid and fed to a rotary kiln, where it is dried, carbonized, and activated at a moderate temperature. To comply with environmental pollution regulations, the kiln off-gases are treated before discharge to the atmosphere. The char is washed with water to remove the acid from the carbon, and the carbon is separated from the slurry. The filtrate is then passed to an acid-recovery unit. Some manufacturing plants do not recycle all the acid but use a part of it to manufacture fertilizer in an allied plant. If necessary, the pH of the activated carbon is adjusted, and the product is dried. The dry product is screened and classified into the size range required for specific granular carbon applications. Carbon yields as high as 50% by weight of the wood precursor have been reported (26).

Novel Manufacturing Processes. Different chemical activation processes have been used to produce carbons with enhanced adsorption characteristics. Activated carbons of exceptionally high surface area (>3,000 m^2/g) have been produced by the chemical activation of carbonaceous materials with potassium hydroxide (28,29). Activated carbons are also produced commercially in the form of cloths (30), fibers (31), and foams (32) generally by chemical activation of the precursor with a Lewis acid such as aluminum chloride, ferric chloride, or zinc chloride.

Forms of Activated Carbon Products

To meet the engineering requirements of specific applications, activated carbons are produced and classified as granular, powdered, or shaped products. Granular activated carbons are produced directly from granular precursors, such as sawdust and crushed and sized coconut char or coal. The granular product is screened and sized for specific applications. Powdered activated carbons are obtained by grinding granular products. Shaped activated carbon products are generally produced as cylindrical pellets by extrusion of the precursor with a suitable binder before activation of the precursor.

Shipping and Storage

Activated carbon products are shipped in bags, drums, and boxes in weights ranging from about 10 to 35 kg. Containers can be lined or covered with plastic and should be stored in a protected area both to prevent weather damage and to minimize contact with organic vapors that could reduce the adsorption performance of the product. Bulk quantities of activated carbon products are shipped in metal bins and bulk bags, typically 1–2 m^3 in volume, and in railcars and tank trucks. Bulk carbon shipments are generally transferred by pneumatic conveyors and stored in tanks. However, in applications such as water treatment where water adsorption does not impact product performance, bulk carbon may be transferred and stored as a slurry in water.

Specifications

Activated carbon producers furnish product bulletins that list specifications, usually expressed as a maximum or minimum value, and typical properties for each grade produced. Standards helpful in setting purchasing specifications for granular and powdered activated carbon products have been published (33,34).

ECONOMIC ASPECTS

Excluding eastern European countries and China, where production figures have not been published, the world pro-

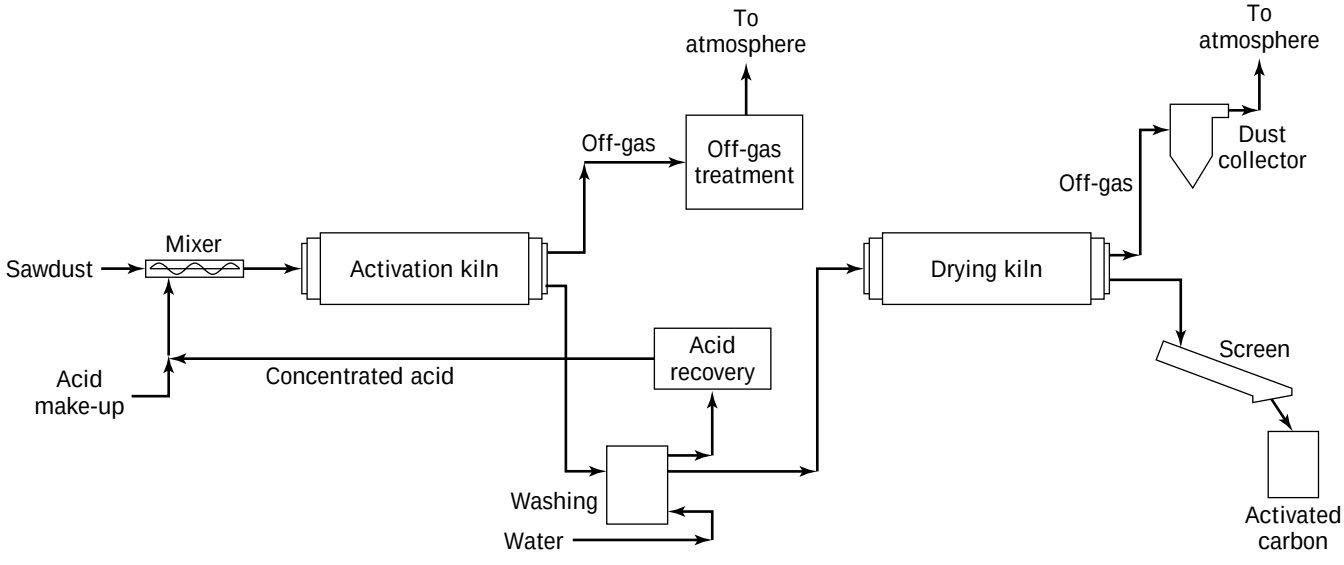


Figure 2. Chemical activation of wood.

duction capacity of activated carbon was estimated to be 375,000 metric tons in 1990 (35). The price of most products was 0.70 to 5.50 $/kg, but some specialty carbons were more expensive (36, pp. 731.2001P, 731.2001Q). Forty percent of the production capacity was in the United States, 30% in western Europe, 20% in Japan, and 10% in other Pacific Rim countries (Table 2).

Production capacity was almost equally split between powdered and nonpowdered activated carbon products. Powdered activated carbon, a less expensive form used in liquid-phase applications, is generally used once and then disposed of. In some cases, however, granular and shaped products are regenerated and reused (35). In 1990 production capacity for granular and shaped products was split, with about two-thirds for liquid-phase and one-third for gas-phase applications (37).

Over the last decade, production capacity in the United States remained essentially unchanged, but minor fluctuations occurred in response to changes in environmental regulations (36pp.731.2000S–731.2000Y). A similar reaction was noted worldwide (35). The current demand for activated carbon is estimated at 93% of production capacity. The near-term growth in demand is projected to be approximately 5.5%/year (37pp.4,5).

In 1970 the U.S. Congress enacted the Clean Air Act, the Clean Water Act, and the Safe Drinking Water Act. Because activated carbon can often be used to help meet Environmental Protection Agency (EPA) regulations, the U.S. activated carbon industry reacted by increasing its production capacity. A proposed amendment to the Safe Drinking Water Act in 1979 required the use of granular activated carbon systems, but the amendment was not enacted. In response to the projected increase in demand for activated carbon, production capacity remained high until the late 1980s, but when the anticipated need did not materialize, some production facilities were shut down. Currently, because of stricter EPA regulations implementing all three acts in 1990, the industry will increase production capacity by 25% during the next several years (35,38).

The estimated production capacity of activated carbon in the United States is shown in Table 3 for seven manufacturers (35pp.54–65). The principal producers are Calgon Carbon (37%), American-Norit (26%), Westvaco (19%), and Atochem (10%). Several other companies purchase activated carbon for resale but do not manufacture products.

Western Europe has seven manufacturers of activated carbon. The two largest, Norit and Chemviron (a subsidiary of Calgon), account for 70% of western European production capacity, and Ceca accounts for 13% (35p.13), Japan is the third-largest producer of activated carbon, having 18 manufacturers, but 4 companies share over 50% of the total Japanese capacity (35p.25–32). Six Pacific Rim countries account for the balance of the world production capacity of activated carbon, 90% of which is in the Philippines and Sri Lanka (35p.13). As is the case with other businesses, regional markets for activated carbon products have become international, leading to consolidation of manufacturers. Calgon, Norit, Ceca, and Sutcliffe-Speakman are examples of multinational companies.

Table 2. World Production Capacity, Estimated 1990

Country	Capacity (10^3 t)
United States	146
Western Europe	108
Japan	72
Pacific Rim, other	49
Total	375

Note: Excluding eastern Europe and China.

Activated carbon is a recyclable material that can be regenerated. Thus the economics, especially the market growth of activated carbon, particularly granular and shaped products, is affected by regeneration and industry regeneration capacity. The decision to regenerate an activated carbon product is dependent on the cost, size of the carbon system, type of adsorbate, and the environmental issue involved. Large carbon systems, such as those used in potable and wastewater treatment, generally require a high-temperature treatment, which is typically carried out in rotary or multihearth furnaces. During regeneration, carbon losses of 1 to 15% typically occur from the treatment and movement of the carbon (39). However, material loss is compensated for by the addition of new carbon to the adsorber system. In general, regeneration of spent carbon is considerably less expensive than the purchase of new activated carbon. For example, fluidized-bed furnace regeneration of activated carbon used in a 94,600-m^3/day water treatment system cost only 35% of new material (40). For this system, regeneration using either infrared or multihearth furnaces was estimated to be more expensive but still significantly less so than the cost of new carbon.

Because powdered activated carbon is generally used in relatively small quantities, the spent carbon has often been disposed of in landfills. However, landfill disposal is becoming more restrictive environmentally and more costly. Thus large consumers of powdered carbon find that regeneration is an attractive alternative. Examples of regeneration systems for powdered activated carbon include the Zimpro/Passavant wet air oxidation process (41), the multihearth furnace as used in the DuPont PACT process (41pp.389–447,42), and the Shirco infrared furnace (40p.51,43).

Other types of regenerators designed for specific adsorption systems may use solvents and chemicals to remove susceptible adsorbates (44), steam or heated inert gas to recover volatile organic solvents (45), and biological systems in which organics adsorbed on the activated carbon during water treatment are continuously degraded (46).

ANALYTICAL TEST PROCEDURES AND STANDARDS

Source references for frequently used test procedures for determining properties of activated carbon are shown in Table 4. A primary source is the *Annual Book of American Society for Testing and Materials (ASTM) Standards* (54). Other useful sources of standards and test procedures include manufacturers of activated carbon products, the American Water Works Association (AWWA) (33,34), and the U.S. Department of Defense (54).

Table 3. Production Capacity in the United States, Estimated 1990

Company	Location	Capacity (10^3 t)
Acticarb Division, Royal Oak Enterprises	Romeo, Fla.	6.8
American Norit Co.	Marshall, Tex.	38.6
Barneby and Sutcliffe	Columbus, Ohio	3.0
Calgon Carbon Corp.	Catlettsburg, Ky. and Pittsburgh, Pa.	53.5
Ceca Division, Atochem NA	Pryor, Okla.	15.0
Trans-Pacific Carbon	Blue Lake, Calif.	2.3
Westvaco Corp.	Covington, Va.	27.2
Total		*146.4*

Table 4. Source References for Activated Carbon Test Procedures and Standards

Title of procedure or standard	Source
Standard Definitions of Terms Relating to Activated Carbon	ASTM D2652
Apparent Density of Activated Carbon	ASTM D2854
Particle Size Distribution of Granular Activated Carbon	ASTM D2862
Total Ash Content of Activated Carbon	ASTM D2866
Moisture in Activated Carbon	ASTM D2867
Ignition Temperature of Granular Activated Carbon	ASTM D3466
Carbon Tetrachloride Activity of Activated Carbon	ASTM D3467
Ball-Pan Hardness of Activated Carbon	ASTM D3802
Radioiodine Testing of Nuclear-Grade Gas-Phase Adsorbents	ASTM D3803
pH of Activated Carbon	ASTM D3838
Determination of Adsorptive Capacity of Carbon by Isotherm Technique	ASTM D3860
Determining Operating Performance of Granular Activated Carbon	ASTM D3922
Impregnated Activated Carbon Used to Remove Gaseous RadioIodines from Gas Streams	ASTM D4069
Determination of Iodine Number of Activated Carbon	ASTM D4607
Military Specification, Charcoal, Activated, Impregnated	Ref. 47
Military Specification, Charcoal, Activated, Unimpregnated	Ref. 47
AWWA Standard for Granular Activated Carbon	Ref. 33
AWWA Standard for Powdered Activated Carbon	Ref. 34
BET Surface Area by Nitrogen Adsorption	Refs. 6, 8, 9, 48
Pore Volume by Nitrogen Adsorption or Mercury Penetration	Refs. 10, 49–52
Particle Density	Ref. 53

HEALTH AND SAFETY

Activated carbon generally presents no particular health hazard as defined by NIOSH (55). However, it is a nuisance and mild irritant with respect to inhalation, skin contact, eye exposure, and ingestion. On the other hand, special consideration must be given to the handling of spent carbon that may contain a concentration of toxic compounds.

Activated carbon products used for decolorizing food products in liquid form must meet the requirements of the *Food Chemical Codex* as prepared by the Food and Nutrition Board of the National Research Council (56).

According to the National Board of Fire Underwriters, activated carbons normally used for water treatment pose no dust explosion hazard and are not subject to spontaneous combustion when confined to bags, drums, or storage bins (67). However, activated carbon burns when sufficient heat is applied; the ignition point varies between about 300 and 600 °C (58p.353).

Dust-tight electrical systems should be used in areas where activated carbon is present, particularly powdered products (58pp.84,85). When partially wet activated carbon comes into contact with unprotected metal, galvanic currents can be set up; these result in metal corrosion (59).

Manufacturer material safety data sheets (MSDS) indicate that the oxygen concentration in bulk storage bins or other enclosed vessels can be reduced by wet activated carbon to a level that will not support life. Therefore, self-contained air packs should be used by personnel entering enclosed vessels where activated carbon is present (60).

LIQUID-PHASE APPLICATIONS

Activated carbons for use in liquid-phase applications differ from gas-phase carbons primarily in pore size distribution. Liquid-phase carbons have significantly more pore volume in the macropore range, which permits liquids to diffuse more rapidly into the mesopores and micropores (61). The larger pores also promote greater adsorption of large molecules, either impurities or products, in many liquid-phase applications. Specific-grade choice is based on the isotherm (62,63) and, in some cases, bench or pilot-scale evaluations of candidate carbons.

Liquid-phase activated carbon can be applied either as a powder, granular, or shaped form. The average size of powdered carbon particles is 15–25 μm (62). Granular or shaped carbon particle size is usually 0.3–3.0 mm. A significant factor in choosing between powdered and nonpowdered carbon is the degree of purification required in the adsorption application. Granular and shaped carbons are usually used in continuous flow through deep beds to remove essentially all contaminants from the liquid being treated. Granular and shaped carbon systems are preferred when a large carbon buffer is needed to withstand significant variations in adsorption conditions, such as in cases where large contaminant spikes may occur. A wider range of impurity removal can be attained by batch application of powdered carbon, and the powdered carbon dose per batch can be controlled to achieve the degree of purification desired (61).

Batch-stirred vessels are most often used in treating material with powdered activated carbon (64). The type of carbon, contact time, and amount of carbon vary with the desired degree of purification. The efficiency of activated carbon may be improved by applying continuous, countercurrent carbon-liquid flow with multiple stages (Fig. 3). Carbon is separated from the liquid at each stage by settling or filtration. Filter aids such as diatomaceous earth are sometimes used to improve filtration.

Granular and shaped carbons are generally used in continuous systems where the liquid to be treated is passed through a fixed bed (41pp.8–19,64). Compounds are adsorbed by the carbon bed in the adsorption zone (Fig. 4). As carbon in the bed becomes saturated with adsorbates, the adsorption zone moves in the direction of flow, and breakthrough occurs when the leading edge of the adsorption zone reaches the end of the column. Normally at least two columns in series are on line at any given time. When the first column becomes saturated, it is removed from service, and a column containing fresh carbon is added at the discharge end of the series. An alternative approach is the moving bed column (41pp.8–19). In this design the adsorption zone is contained within a single column by passing liquid upward while continuously or intermittently withdrawing spent carbon at the bottom and adding fresh carbon at the top.

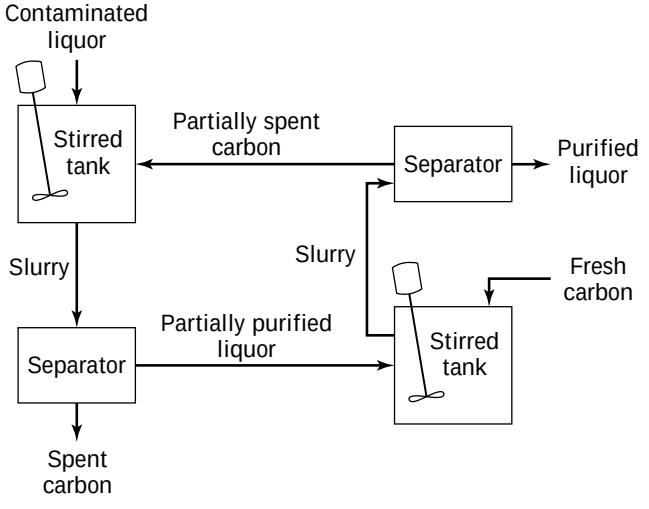


Figure 3. Multistage countercurrent application of powdered activated carbon.

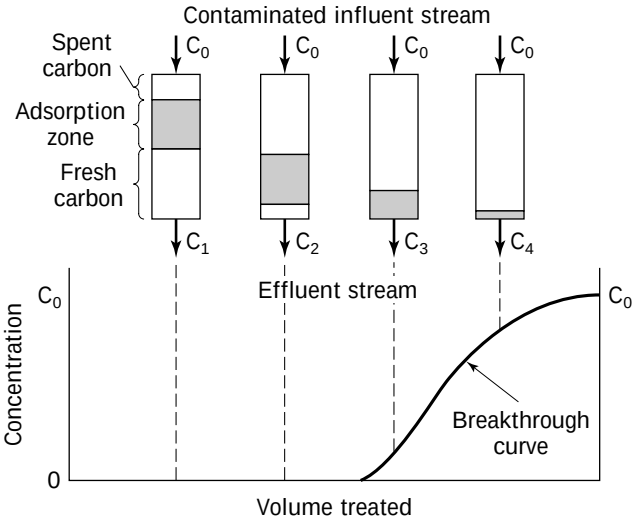


Figure 4. Adsorption and breakthrough curve for fixed bed of granular or shaped activated carbon.

The total activated carbon consumption for liquid-phase applications in the United States in 1987 was estimated to be about 76,700 t, which accounted for nearly 80% of the total activated carbon use. The consumption by application is summarized in Table 5 (36pp.731.2000V–731.2001L)

Potable Water Treatment

Treatment of drinking water accounts for about 24% of the total activated carbon used in liquid-phase applications (36pp.731.2000V–731.2001L). Rivers, lakes, and groundwater from wells, the most common drinking water sources, are often contaminated with bacteria, viruses, natural vegetation decay products, halogenated materials,

Table 5. Liquid-Phase Activated Carbon Consumption (10^3 t) in the United States in 1987.

	Granular-shaped	Powdered	Total
Potable water	4.5	13.6	18.1
Wastewater			
Industrial	6.4	6.6	13.0
Municipal	0.9	2.0	2.9
Sweetener decolorization	6.8	9.1	15.9
Chemical processing and misc.	4.1	2.3	6.4
Food, beverage, and oils	0.9	3.9	4.8
Pharmaceuticals	2.0	2.3	4.3
Mining	1.6	2.5	4.1
Groundwater	0.9	2.3	3.2
Household uses	1.4	0.9	2.3
Dry cleaning	0.7	0.4	1.1
Electroplating	0.2	0.4	0.6
Total	*30.4*	*46.3*	*76.7*

and volatile organic compounds. Normal water disinfection and filtration treatment steps remove or destroy the bulk of these materials (65). However, treatment by activated carbon is an important additional step in many plants to remove toxic and other organic materials (66–68) for safety and palatability.

Groundwater Remediation

Concern over contaminated groundwater sources increased in the 1980s, and in 1984 an Office of Groundwater Protection was created by the EPA (36pp.731.2000V–731.2001L). This led to an increase in activated carbon consumption in 1987 for groundwater treatment to about 4% of the total liquid-phase usage, and further growth is expected in the 1990s. There are two ways to apply carbon in groundwater cleanup. One is the conventional method of applying powdered, granular, or shaped carbon to adsorb contaminants directly from the water. The other method utilizes air stripping to transfer the volatile compounds from water to air. The compounds are then recovered by passing the contaminated air through a bed of carbon (69,70).

Industrial and Municipal Wastewater Treatment

Wastewater treatment consumes about 21% of the total U.S. liquid-phase activated carbon (36pp.731.2000V–731.2001L), and governmental regulations are expected to increase demand over the next several years. Wastewater may contain suspended solids, hazardous microorganisms, and toxic organic and inorganic contaminants that must be removed or destroyed before discharge to the environment. In tertiary treatment systems, powdered, granular, or shaped carbon can be used to remove residual toxic and other organic compounds after the primary filtration and secondary biological treatment (71). Powdered carbon is also used in the PACT process by direct addition of the carbon to the secondary biological step (41pp.389–447).

Sweetener Decolorization

About 21% of the liquid-phase activated carbon is used for purification of sugar (qv) and corn syrup (36pp.731.2000V–731.2001L). White sucrose sugar is made from raw juice squeezed from sugarcane or sugar beets. The clarified liquor is decolorized using activated carbon or ion exchange resins (58pp.87–125,274–292). High-fructose corn sweeteners (HFCS) are produced by hydrolysis of corn starch and are then treated with activated carbon to remove undesirable taste and odor compounds and to improve storage life. The demand for HFCS rose sharply in the 1980s primarily because of the switch by soft drink producers away from sucrose (83).

Chemical Processing

Activated carbon consumption in a variety of chemical processing applications is about 8% of the total (36pp.731.2000V–731.2001L). The activated carbon removes impurities to achieve high quality. For example, organic contaminants are removed from solution in the production of alum, soda ash, and potassium hydroxide (58pp.87–125,274–292). Other applications include the manufacture of dyestuffs, glycols, amines, organic acids, urea, hydrochloric acid, and phosphoric acid (35pp.92–135).

Food, Beverage, and Cooking Oil

Approximately 6% of the liquid-phase activated carbon is used in food, beverage, and cooking oil production (36pp.731.2000V–731.2000L). Before being incorporated into edible products, vegetable oils and animal fats are refined to remove particulates, inorganics, and organic contaminants. Activated carbon is one of several agents used in food purification processes. In the production of alcoholic beverages, activated carbon removes haze-causing compounds from beer, taste and odor from vodka, and fusel oil from whiskey (58pp.87–125,274–292). The feed water for soft drink production is often treated with carbon to capture undesirable taste and odor compounds and to remove free chlorine remaining from disinfection treatment. Caffeine is removed from coffee beans by extraction with organic solvents, water, or supercritical carbon dioxide prior to roasting. Activated carbon is used to remove the caffeine from the recovered solvents (35pp.92–135).

Pharmaceuticals

Pharmaceuticals account for 6% of the liquid-phase activated carbon consumption (36pp.731.2000V–731.2000L). Many antibiotics, vitamins, and steroids are isolated from fermentation broths by adsorption onto carbon followed by solvent extraction and distillation (58pp.87–125,274–292). Other uses in pharmaceutical production include removal of chromogenic impurities, process water purification, and removal of impurities from intravenous solutions prior to packaging (35pp.92–135).

Mining

The mining industry accounts for only 4% of liquid-phase activated carbon use, but this figure may grow as low-grade ores become more common (36pp.731.2000V–731.2000L). Gold, for example, is recovered on activated carbon as a cyanide complex in the carbon-in-pulp extraction process (58pp.87–125, 274–292). Activated carbon serves as a catalyst in the detoxification of cyanides contained in wastewater from cyanide stripping operations (46pp.8–19). Problems caused by excess flotation agent concentrations in flotation baths are commonly cured by adding powdered activated carbon (58pp.87–125,274–292).

Miscellaneous Uses

Several relatively low-volume activated carbon uses comprise the remaining 6% of liquid-phase carbon consumption (36pp.731.2000V–731.2000L). Small carbon filters are used in households for purification of tap water. Oils, dyes, and other organics are adsorbed on activated carbon in dry cleaning recovery and recycling systems. Electroplating solutions are treated with carbon to remove organics that can produce imperfections when the thin metal layer is deposited on the substrate (58pp.87–127,274–292). Medical ap-

plications include removal of toxins from the blood of patients with artificial kidneys (35pp.92–135), and oral ingestion into the stomach to recover poisons or toxic materials (58pp.87–125,274–292;72). Activated carbon also is used as a support for metal catalysts in low-volume production of high-value specialty products such as pharmaceuticals, fragrance chemicals, and pesticides (73).

GAS-PHASE APPLICATIONS

Gas-phase applications of activated carbon include separation, gas storage, and catalysis. Although only 20% of activated carbon production is used for gas-phase applications, these products are generally more expensive than liquid-phase carbons and account for about 40% of the total dollar value of shipments. Most of the activated carbon used in gas-phase applications is granular or shaped. Activated carbon use by application is shown in Table 6 (74).

Separation processes comprise the main gas-phase applications of activated carbon. These usually exploit the differences in the adsorptive behavior of gases and vapors on activated carbon on the basis of molecular weight and size. For example, organic molecules with a molecular weight greater than about 40 are readily removed from air by activated carbon.

Solvent Recovery

Most of the activated carbon used in gas-phase applications is employed to prevent the release of volatile organic compounds into the atmosphere. Much of this use has been in response to environmental regulations, but recovery and recycling of solvents from a range of industrial processes such as printing, coating, and extrusion of fibers also provides substantial economic benefits.

The structure of activated carbons used for solvent recovery has been predominantly microporous. Micropores provide the strong adsorption forces needed to capture small vapor molecules such as acetone at low concentrations in process air (75). In recent years, however, more mesoporous carbons, specifically made for solvent recovery, have become available and are giving good service, especially for the adsorption of heavier vapors such as cumene and cyclohexanone that are difficult to remove from micropores during regeneration (75). Regeneration of the carbon is performed on a cyclic basis by purging it with steam or heated nitrogen.

Table 6. Gas-Phase Activated Carbon Uses in the United States in 1987

Application	Consumption (10^3 t)
Solvent recovery	4.5
Automotive/gasoline recovery	4.1
Industrial off-gas control	3.2
Catalysis	2.7
Pressure swing separation	1.1
Air conditioning	0.5
Gas mask	0.5
Cigarette filters	0.5
Nuclear	0.3
Total	*17.4*

Gasoline Emission Control

A principal application of activated carbon is in the capture of gasoline vapors that escape from vents in automotive fuel systems (76). Under EPA regulations, all U.S. motor vehicles produced since the early 1970s have been equipped with evaporative emission control systems. Most other auto-producing countries now have similar controls. Fuel vapors vented when the fuel tank or carburetor are heated are captured in a canister containing 0.5–2 L of activated carbon. Regeneration of the carbon is then accomplished by using intake manifold vacuum to draw air through the canister. The air carries desorbed vapor into the engine where it is burned during normal operation. Activated carbon systems have also been proposed for capturing vapors emitted during vehicle refueling, and activated carbon is used at many gasoline terminals to capture vapor displaced when tank trucks are filled (77). Typically, the adsorption vessels contain around 15 m^3 of activated carbon and are regenerated by application of a vacuum. The vapor that is pumped off is recovered in an absorber by contact with liquid gasoline. Similar equipment is used in the transfer of fuel from barges (78). The type of carbon pore structure required for these applications is substantially different from that used in solvent recovery. Because the regeneration conditions are very mild, only the weaker adsorption forces can be overcome, and therefore the most effective pores are in the mesopore size range (79). A large adsorption capacity in these pores is possible because vapor concentrations are high, typically 10–60%.

Adsorption of Radionuclides

Other applications that depend on physical adsorption include the control of krypton and xenon radionuclides from nuclear power plants (80). The gases are not captured entirely, but their passage is delayed long enough to allow radioactive decay of the short-lived species. Highly microporous coconut-based activated carbon is used for this service.

Control by Chemical Reaction

Pick-up of gases to prevent emissions can also depend on the chemical properties of activated carbon or of impregnants. Emergency protection against radioiodine emissions from nuclear power reactors is provided by isotope exchange over activated carbon impregnated with potassium iodide (81). Oxidation reactions catalyzed by the carbon surface are the basis for several emission control strategies. Sulfur dioxide can be removed from industrial off-gases and power plant flue gas because it is oxidized to sulfur trioxide, which reacts with water to form nonvolatile sulfuric acid (82,83). Hydrogen sulfide can be removed from such sources as Claus plant tail gas because it is converted to sulfur in the presence of oxygen (84). Nitric oxide can be removed from flue gas because it is oxidized to nitrogen dioxide. Ammonia is added and reacts catalytically on the carbon surface with the nitrogen dioxide to form nitrogen (85).

Protection against Atmospheric Contaminants

Activated carbon is widely used to filter breathing air to protect against a variety of toxic or noxious vapors, including war gases, industrial chemicals, solvents, and odorous compounds. Activated carbons for this purpose are highly microporous and thus maximize the adsorption forces that hold adsorbate molecules on the surface. Although activated carbon can give protection against most organic gases, it is especially effective against high molecular weight vapors, including chemical warfare agents such as mustard gas or the nerve agents that are toxic at parts per million concentrations. The activated carbon is employed in individual canisters or pads, as in gas masks, or in large filters in forced air ventilation systems. In air-conditioning systems, adsorption on activated carbon can be used to control the buildup of odors or toxic gases such as radon in recirculated air (86).

Inorganic vapors are usually not strongly adsorbed on activated carbon by physical forces, but protection against many toxic agents is achieved by using activated carbon impregnated with specific reactants or decomposition catalysts. For example, a combination of chromium and copper impregnants is used against hydrogen cyanide, cyanogen, and cyanogen chloride, whereas silver assists in the removal of arsine. All of these are potential chemical warfare agents; the Whetlerite carbon, which was developed in the early 1940s and is still used in military protective filters, contains these impregnants (87). Recent work has shown that chromium, which loses effectiveness with age and is itself toxic, can be replaced with a combination of molybdenum and triethylenediamine (88). Oxides of iron and zinc on activated carbon have been used in cigarette filters to absorb hydrogen cyanide and hydrogen sulfide (89). Mercury vapor in air can be removed by activated carbon impregnated with sulfur (90). Activated carbon impregnated with sodium or potassium hydroxide has long been used to control odors of hydrogen sulfide and organic mercaptans in sewage treatment plants (91). Alkali-impregnated carbon is also effective against sulfur dioxide, hydrogen sulfide, and chlorine at low concentrations. Such impregnated carbon is used extensively to protect sensitive electronic equipment against corrosion by these gases in industrial environments (92).

Process Stream Separations

Differences in adsorptivity between gases provides a means for separating components in industrial process gas streams. Activated carbon in fixed beds has been used to separate aromatic compounds from lighter vapors in petroleum refining process streams (93) and to recover gasoline components from natural and manufactured gas (94,95).

Molecular sieve activated carbons are specially made with restricted openings leading to micropores. These adsorbents are finding increasing use in separations utilizing pressure swing adsorption, in which adsorption is enhanced by operation at high pressure and desorption occurs upon depressurization (96). Larger molecules are restricted from entrance into the pores of these carbons and, therefore, are not retained as strongly as smaller molecules. The target product can be either the adsorbed or unadsorbed gases. Examples include separation of oxygen from air and recovery of methane from inorganic gases in biogas production. Hydrogen can be removed from gases produced in the catalytic cracking of gasoline, and carbon monoxide can be separated from fuel gases. Use of pressure swing techniques for gas separation is an area of growing interest in engineering research.

The hypersorption process developed in the late 1940s used a bed of activated carbon moving countercurrent to gas flow to separate light hydrocarbons from each other and from hydrogen in refinery operations. The application is of interest because of its scale, treating up to 20,000 m^3/h of gas, but the plants were shut down within a few years, probably because of problems related to attrition of the rapidly circulating activated carbon (97). It should be noted, however, that in recent years moving-bed and fluid-bed adsorption equipment using activated carbon has been successfully employed for solvent recovery (98).

Gas Storage

Adsorption forces acting on gas molecules held in micropores significantly densify the adsorbed material. As a result, activated carbon has long been considered a medium for lowering the pressure required to store weakly adsorbed compressed gases (99). Recent work with modern high-capacity carbons has been directed toward fueling passenger cars with natural gas, but storage volume targets have not yet been attained (100). Natural gas storage on activated carbon is now used commercially in portable welding cylinders (101). These can be refilled easily at about 2,000 kPa and hold as much gas as a conventional cylinder pressurized to 6,000 kPa (59 atm).

Catalysis

Catalytic properties of the activated carbon surface are useful in both inorganic and organic synthesis. For example, the fumigant sulfuryl fluoride is made by reaction of sulfur dioxide with hydrogen fluoride and fluorine over activated carbon (102). Activated carbon also catalyzes the addition of halogens across a carbon-carbon double bond in the production of a variety of organic halides (73) and is used in the production of phosgene from carbon monoxide and chlorine (103,104).

BIBLIOGRAPHY

1. H. Jüntgen, *Carbon*, **15**, 273–283 (1977).
2. Brit. Pat. 14,224 (1900), R. von Ostrejko; Fr. Pat. 304,867 (1900); Ger. Pat. 136,792 (1901); U.S. Pat. 739,104 (1903).
3. J.W. Hassler, *Forest Products J.* **8**, 25A–27A (1958).
4. J.W. Hassler, *Activated Carbon*, Chemical Publishing Co., Inc., New York, 1963, pp. 1–14. A comprehensive account of the development and use of activated carbon products to about 1960.
5. R.V. Carrubba, J.E. Urbanic, N.J. Wagner, and R.H. Zanitsch, *AIChE Symp. Ser.* **80**, 76–83 (1984).
6. B. McEnaney and T.J. Mays, in H. Marsh ed., *Introduction to Carbon Science*, Butterworths, London, 1989, pp. 153–196. A good introduction to carbon science in general.

7. H. Marsh and J. Butler, in K.K. Unger, J. Rouquerol, K.S.W. Sing, and H. Kral eds., *Characterization of Porous Solids, Proceedings of the IUPAC Symposium (COPS I)*, Bad Soden a.Ts., FRG, Apr. 26–29, 1987, Elsevier, Amsterdam, The Netherlands, 1988, pp. 139–149.
8. K.S.W. Sing et al., *Pure Appl. Chem.* **57**, 603–619 (1985).
9. S. Brunauer, P.H. Emmett, and E. Teller, *J. Am. Chem. Soc.* **60**, 309–319 (1938).
10. S.J. Gregg and K.S.W. Sing, *Adsorption, Surface Area, and Porosity*, 2nd ed., Academic Press, London, 1982. An indispensable text on the interpretation and significance of adsorption data.
11. H. Marsh, D. Crawford, T.M. O'Grady, and A. Wennerberg, *Carbon* **20**, 419–426 (1982).
12. *Jpn. Chem. Week* **30**, 5 (Mar. 16, 1989).
13. M.M. Dubinin, *J. Colloid Interface Sci.* **46**, 351–356 (1974).
14. K.S.W. Sing, *Carbon* **27**, 5–11 (1989).
15. J. Zawadzki, in P.A. Thrower ed., *Chemistry and Physics of Carbon*, vol. 21, Marcel Dekker, New York, 1989, pp. 147–380. *Chemistry and Physics of Carbon*, published in 23 volumes through 1991, is a primary source of excellent review articles on carbon, many relevant to activated carbon.
16. D. Atkinson, A.I. McLeod, K.S.W. Sing, and A. Capon, *Carbon* **20**, 339–343 (1982).
17. J.M. Solar, C.A. Leon y Leon, K. Osseo-Asare, and L.R. Radovic, *Carbon* **28**, 369–375 (1990).
18. K.D. Henning, W. Bongartz, and J. Degel, *19th Biennial Conference on Carbon*, Penn State University, Pa., June 25–30, 1990, extended abstracts, pp. 94–95.
19. Product data bulletins from activated carbon manufacturers, Calgon Carbon Corp., 1990, American Norit Co., 1990, and Westvaco Corp., 1988.
20. T. Wigmans, *Carbon* **27**, 13–22 (1989).
21. F. Derbyshire and M. Thwaites, *Proc. 4th Australian Coal Science Conf.*, Brisbane, Australia, Dec. 3–5, 1990, pp. 372–379.
22. U.S. Pat. 3,976,597 (Aug. 24, 1976), A.J. Repik, C.E. Miller, and H.R. Johnson (to Westvaco Corp.).
23. W. Gerhartz, Y.S. Yamamoto, and F. Thomas Campbell eds., *Ullmann's Encyclopedia of Industrial Chemistry*, 5th ed., vol. A5, VCH, New York, 1986, pp. 124–140. Good descriptions of activation processes.
24. Product literature on Pittsburgh activated carbon, Pittsburgh Coke & Chemical Co. (now Calgon Carbon Corp.), Pittsburgh, Pa., ca 1960.
25. U.S. Pat. 4,014,817 (Mar. 29, 1977), B.C. Johnson, R.K. Sinha, and J.E. Urbanic (to Calgon Corp.).
26. A. Cameron and J.D. MacDowall, in J.M. Haynes and P. Rossi-Doria eds., *Principles and Applications of Pore Structural Characterization, Proceedings of the RILEM/CNR International Symposium*, Milan, Italy, Apr. 26–29, 1983, J.W. Arrowsmith, Ltd., Bristol, U.K., 1985, pp. 251–275.
27. R.C. Bansal, J.-B. Donnet, and F. Stoeckli, *Active Carbon*, Marcel Dekker, New York, 1988, p. 8. A modern treatise on activated carbon based on a comprehensive review of the literature.
28. U.S. Pat. 4,082,694 (Apr. 4, 1978), A.N. Wennerberg and T.M. O'Grady (to Standard Oil Co.).
29. T. Kasuh, D.A. Scott, and M. Mori, *Proceedings of an International Conference on Carbon*, The University of Newcastle upon Tyne, U.K., Sept. 18–23, 1988, pp. 146–148.
30. Product literature on activated carbon cloth, Charcoal Cloth Ltd., U.K., 1985, and on C-tex products, Siebe Gorman & Co., Ltd., U.K., 1985.
31. Product literature on KYNOL activated carbon fibers and cloths, GUN EI Chemical Industry Co., Ltd., Japan, 1987; product literature on AD'ALL activated carbon fibers, Unitika, Ltd., Japan, 1989.
32. Product literature on KURASHEET activated carbon foam sheets, Kuraray Chemical Co., Ltd., Japan, 1987.
33. *AWWA Standard for Granular Activated Carbon*, ANSI/AWWA B604, American Water Works Association, Denver, Colo., 1991.
34. *AWWA Standard for Powdered Activated Carbon*, ANSI/AWWA B600, American Water Works Association, Denver, Colo., 1990.
35. *The Economics of Activated Carbon*, 3rd ed., Roskill Information Services, London, 1990.
36. J. Goin, V. von Schuller-Goetzburg, and Y. Sakuma, *Chemical Economics Handbook—SRI International*, Menlo Park, Calif., 1989.
37. Wertheim Schroder, *Calgon Carbon Corp.*, Company Report No. 955346, Mar. 19, 1990.
38. S. Irving-Monshaw, *Chem. Eng.* **97(2)**, 43–46 (1990).
39. W.G.P. Schuliger, *Waterworld News* **4(1)**, 15–17 (1988).
40. R.M. Clark and B.W. Lykins, Jr., *Granular Activated Carbon—Design, Operation, and Cost*, Lewis Publishers, Chelsea, Mich., 1989, pp. 295–338.
41. P.N. Cheremisinoff and F. Ellerbusch, *Carbon Adsorption Handbook*, Ann Arbor Science Publishers, Ann Arbor, Mich., 1978, pp. 539–626. An excellent reference book on activated carbon, ranging from theoretical to applied aspects.
42. Product literature on PACT systems, Zimpro/Passavant, Inc., Rothschild, Wis., 1990.
43. W.E. Koffskey and B.W. Lykins, Jr., *J. Am. Water Works Assoc.* **82(1)**, 48–56 (1990).
44. A. Yehaskel, *Activated Carbon—Manufacture and Regeneration*, Noyes Data Corporation, Park Ridge, N.J., 1978, pp. 202–217. A dated, but still useful summary of key patent literature.
45. P.N. Cheremisinoff, *Pollut. Eng.* **17(3)**, 29–38 (1985).
46. R.G. Rice and C.M. Robson, *Biological Activated Carbon—Enhanced Aerobic Biological Activity in GAC Systems*, Ann Arbor Science Publishers, Ann Arbor, Mich., 1982, 611 pp.
47. *Department of Defense Military Specifications*, MIL-C-0013724D(EA), Sept. 22, 1983; MIL-C-0013724D(EA) Amendment 1, Mar. 5, 1986; and MIL-C-17605C(SH), Mar. 22, 1989.
48. S.J. Gregg and K.S.W. Sing, *Adsorption, Surface Area, and Porosity*, 1st ed., Academic Press (London) London, 1967, pp. 308–355.
49. H.M. Rootare, *Advanced Experimental Techniques in Powder Metallurgy*, Plenum, New York, 1970, pp. 225–252. A comprehensive review of the use of mercury penetration to measure porosity.
50. G. Horvath and K. Kawazoe, *J. Chem. Eng. Jpn.* **16**, 470–475 (1983).
51. D. Dollimore and G.R. Heal, *J. Appl. Chem.* **14**, 109–114 (1964).
52. M.M. Dubinin and H.F. Stoeckli, *J. Colloid Interface Sci.* **56**, 34–42 (1980).
53. C. Orr, Jr., *Powder Technol.* **3**, 117–123 (1970).

54. *Annual Book of ASTM Standards*, 15.01, Section 15, American Society for Testing and Materials, Philadelphia, Pa., 1989.
55. *1985–1986 Registry of Toxic Effects of Chemical Substances*, vol. 2, National Institute for Occupational Safety and Health, U.S. Department of Health and Human Services, Washington, D.C., 1987, p. 1475.
56. National Research Council, Assembly of Life Sciences, Division of Biological Sciences, Food and Nutrition Board, and Committee on Codex Specifications, *Food Chemicals Codex*, 3rd ed., National Academy Press, Washington, D.C., 1981, pp. 70, 71.
57. American Society of Civil Engineers, American Water Works Association and Conference of State Sanitary Engineers, *Water Treatment Plant Design*, American Water Works Association, New York, 1969, p. 297.
58. J.W. Hassler, *Purification with Activated Carbon*, 3rd ed., Chemical Publishing, New York, 1974. Contains much of the information given in Ref. 4 but with more emphasis on the commercial uses of activated carbon.
59. U.S. Environmental Protection Agency, *Process Design Manual for Carbon Adsorption*, Swindell-Dressler, Pittsburgh, Pa., 1971, pp. 3–68.
60. Material safety data sheets on activated carbon products, available from the manufacturers, 1991.
61. R.A. Hutchins, *Chem. Eng.* **87(2)**, 101–110 (1980). A particularly useful paper on liquid-phase adsorption.
62. M. Suzuki, *Adsorption Engineering*, Kodansha Ltd., Tokyo and Elsevier Science Publishers B.V., Amsterdam, The Netherlands, 1990, pp. 11, 35–62.
63. T.F. Speth and R.J. Miltner, *J. Am. Water Works Assoc.* **82(2)**, 72–75 (1990).
64. F.L. Slejko ed., *Adsorption Technology*, Marcel Dekker, New York, 1985, pp. 23–32. A good account of the theory, design, and application of adsorption systems.
65. American Water Works Association, *Water Quality and Treatment*, 3rd ed., McGraw-Hill, New York, 1971, pp. 1–216.
66. W.J. Weber, Jr. and B.M. Van Vliet, in I.H. Suffet and M.J. McGuire eds., *Activated Carbon Adsorption of Organics from the Aqueous Phase*, vol. 1, Ann Arbor Science Publishers, Ann Arbor, Mich., 1980, pp. 15–41. A comprehensive, two-volume treatise with many key references.
67. J.L. Oxenford and B.W. Lykins, Jr., *J. Am. Water Works Assoc.* **83(1)**, 58–64 (1991).
68. I.N. Najm and co-workers, *J. Am. Water Works Assoc.* **83(1)**, 65–76 (1991).
69. L.W. Canter and R.C. Knox, *Ground Water Pollution Control*, Lewis Publishers, Chelsea, Mich., 1985, pp. 89–125.
70. Environmental Science and Engineering, Inc., *Removal of Volatile Organic Chemicals from Potable Water—Technologies and Costs*, Noyes Data Corp., Park Ridge, N.J., 1986, pp. 23–40.
71. G. Culp, G. Wesner, R. Williams, and M.V. Hughes, *Wastewater Reuse and Recycling Technology*, Noyes Data Corp., Park Ridge, N.J., 1980, pp. 343–432. A useful review of wastewater treatment with activated carbon.
72. M. Smisek and S. Cerny, *Active Carbon—Manufacture, Properties, and Applications*, Elsevier Publishing, New York, 1970, pp. 290–294.
73. A.J. Bird, in A.B. Stiles ed., *Catalyst Supports and Supported Catalysts*, Butterworths, Stoneham, Mass., 1987, pp. 107–137.
74. pp. 731.2000W, 731.2001M-731.2001P.
75. P.J. Luft and P.C. Speers, Paper 52c, *AIChE Summer National Mtg.*, Aug. 19–22, 1990.
76. P.J. Clarke and co-workers, *SAE Trans.* **76**, 824–837 (1967).
77. Product literature on hydrocarbon vapor recovery systems, John Zink Co., Tulsa, Okla., 1990.
78. J. Hill, *Chem. Eng.* **97**, 133–143 (1990).
79. H.R. Johnson and R.S. Williams, S.A.E. Technical Paper No. 902119, International Fuels and Lubricants Exposition, Tulsa, Okla., Oct. 23, 1990.
80. D.W. Moeller and D.W. Underhill, *Nucl. Saf.* **22**, 599–611 (1981).
81. M.L. Hyder, *Comm. Eur. Communities [Rep.] EUR 1986, EUR 10580, Gaseous Effluent Treat. Nucl. Install.*, pp. 451–462 (1986).
82. F.J. Ball, S.L. Torrence, and A.J. Repik, *APCA J.* **22**, 20–26 (1972).
83. P. Ellwood, *Chem. Eng.* **76**, 62–64 (1969).
84. J. Klein and K.-D. Henning, *Fuel* **63**, 1064–1067 (1984).
85. E. Richter, *Catal. Today* **7**, 93–112 (1990).
86. M.A. Brisk and A. Turk, *Proc. APCA Ann. Meetg., 77th* **2**, 84–93 (1984).
87. U.S. Pat. 2,920,050 (Jan. 5, 1960), R.J. Grabenstetter and F.E. Blacet (to U.S. Dept. of Army).
88. U.S. Pat. 4,801,311 (Jan. 31, 1989), E.-D. Tolles (to Westvaco Corp.).
89. U.S. Pat. 3,460,543 (Aug. 12, 1969), C.H. Kieth, V. Norman, and W.W. Bates, Jr. (to Ligget & Meyers Corp.).
90. R.K. Sinah and P.L. Walker, *Carbon* **10**, 754–756 (1972).
91. W.D. Lovett and R.L. Poltorak, *Water and Sewage Works* **121**, 74–75 (1974).
92. G.N. Brown, M.A. Lunn, C.E. Miller, and C.D. Shelor, *Tappi J.* **66**, 33–36 (1983).
93. S. Dunlop and R. Banks, *Hydrocarbon Process.* **56**, 147–152 (1977).
94. G.F. Russell, *Petrol. Refiner.* **40**, 103–106 (1961).
95. T. Scott, *Gas. J.* **303**, 300–307 (1960).
96. E. Richter, *Erdol Kohle, Erdgas, Petrochem.* **40**, 432–438 (1987).
97. C. Berg, *Chem. Eng. Prog.* **47**, 585–590 (1951).
98. *Gastak Solvent Recovery System*, product literature, Kureha Chemical Industry Co., Ltd., New York, 1990.
99. H. Briggs and W. Cooper, *Proc. R. Soc. Edinburgh* **41**, 119–127 (1920–1921).
100. J. Braslaw, J. Nasea, and A. Golovoy, *Alternative Energy Sources: Proceedings of the Miami Int. Conf. on Alternative Energy Sources*, 4th ed., Ann Arbor Science Publishers, Ann Arbor, Mich., pp. 261–270, 1980.
101. U.S. Pat. 4,817,684 (Apr. 4, 1989), J.W. Turko and K.S. Czerwinski (to Michigan Consolidated Gas Co.).
102. U.S. Pat. 4,102,987 (July 25, 1978), D.M. Cook and D.C. Gustafson (to The Dow Chemical Company).
103. H. Jüntgen, *Fuel* **65**, 1436–1446 (1986).
104. H. Jüntgen, *Erdol Kohle, Erdgas, Petrochem.* **39(12)**, 546–551 (1986).

ADDITIONAL READING

"Active Carbon" under "Carbon" in *Encyclopedia of Chemical Technology (ECT)* 1st ed., Vol. 2, pp. 881–899, by J.W. Hassler, Nuchar Active Carbon Division, West Virginia Pulp and Paper Co., and J.W. Goetz, Carbide and Carbon Chemicals Corp.

"Activated Carbon" under "Carbon" in *ECT* 2nd ed., Vol. 4, pp. 149–158, by E.G. Doying, Union Carbide Corp., Carbon Products Division.
"Activated Carbon" under "Carbon (Carbon and Artificial Graphite)" in *ECT* 3rd ed., Vol. 4, pp. 561–570, by R.W. Soffel, Union Carbide Corp.

ADSORBENTS, INORGANIC

KAMI BEYZAVI
Bioprocessing Limited
Consett, County Durham, United Kingdom

KEY WORDS

Alumina
Chromatography
CPG (controlled pore glass)
Hydroxyapatite
Rigid matrices
Separation
Silica
Titania
Zirconia

OUTLINE

INTRODUCTION

The design of a process for the large-scale purification of biological molecules is one of the several complex tasks that must be successfully completed to bring a product to market. Most downstream processes include some chromatography steps as well as other conventional separation methods. Chromatography is generally recognized as the technique that enables the highest degree of purification of biomolecules. For example, the U.S. Food and Drug Administration (FDA) recommends that at least two chromatography steps be used in the purification of therapeutic-grade monoclonal antibodies.

Chromatography separations are sophisticated and can be expensive processes, which explains the continuous development of effective approaches by scientists and engineers to improve the productivity of chromatography technologies. Fluidized bed adsorption, various column technologies, computerized optimization strategies, and continuous and semicontinuous processes are some of the recent approaches that have been proposed. However, one of the most important aspects of chromatography technology is the choice of matrix.

In this article, we examine the use of inorganic rigid matrices in chromatography separation processes of biological molecules, with emphasis on the preparation, properties, and applications of controlled pore glass (CPG).

CHOICE OF MATRICES FOR BIOPROCESS SEPARATIONS

The subject of matrices for process chromatography and the requirements of an ideal matrix have been reviewed by many authors (1–3). In process-scale purification, the important characteristics of a matrix are chemical and physical stability, rigidity of the matrix to withstand high flow rates and differing hydrodynamic pressure, uniform porosity, and high surface area. In addition, for most chromatography applications, surface modification of the matrix will be necessary to facilitate ligand immobilization, which requires the presence of suitable functional groups on the matrix. Finally, the cost and commercial availability of the matrix are also important considerations in selecting a matrix.

The majority of matrices used in chromatography processes can be divided into two categories: natural and synthetic. Natural matrices include glass, silica, ceramic, alumina, agarose, chitosan, and cellulose. Synthetic matrices are prepared by polymerization of functional monomers. The incorporation of monomers with suitable functional groups can provide activation sites on these supports for ligand attachments. There are also the composite matrices, which are mixed networks constituted of at least two components. Generally, one component acts as a rigid skeleton, and the other component has active groups that interact with the biological molecule to be separated. Dextran-coated silica and polyacrylamide-agarose are examples of composite matrices that have been applied to the separation of biological molecules.

INORGANIC MATRICES

A large selection of inorganic matrices are commercially available for incorporating homogenous catalysts, and some have been applied to chromatography separations. These matrices include silica, alumina, glass, zeolite, clay, celite, zerconia, titania, magnesium oxide, and carbon. Only a limited number of these matrices with suitable pore

structure, particle size, and mechanical stability can be used for chromatography separation of biomolecules. Silica, alumina, and glass are available with a large variety of surface areas, pore structure, and particle sizes, and some have been used extensively in both analytical and preparative chromatography of biological molecules.

Controlled Pore Glass

Controlled pore glass (CPG) has excellent mechanical stability and can be manufactured with a wide range of pore diameters, a high surface area, and a range of particle sizes. The material is almost pure silica and has a very narrow pore size distribution. Surface-modified CPG with the trade name PROSEP manufactured by Bioprocessing Limited has been extensively used in large-scale purification of biomolecules. (See later sections on CPG for a more detailed discussion of this material.)

Silica

Silica is extensively used in the analytical field (HPLC) and to a much lesser extent for preparative chromatography of biomolecules. Silica matrices for chromatography applications are generally produced by a wet process, that is, by acidification of sodium silicate in water (4,5), and are available in particle sizes of up to 40 μm and pore sizes of up to 30 nm. In most cases, these materials are unsuitable for large-scale purification processes. The small particle size causes excessive backpressure in preparative chromatography columns, and the pore size of 30 nm is not large enough for the separation of most proteins, especially in the field of affinity chromatography, where both the ligand and the ligate could be large proteins.

In our laboratories, we have used Matrex™ silica (Amicon) for the large-scale purification of biomolecules. This material is available in a range of pore diameters and particle sizes (Table 1) and is manufactured by first forming a hydrosol from organosilicic acid. Hydrosol is then allowed to set to a jelly-like mass called hydrogel. After setting, the hydrogel is thoroughly washed to remove salts resulting from the reaction. Temperature and pH of the washing medium and the washing time influence the specific surface area, pore volume, and purity of the hydrogel. The hydrogel is subsequently dried to give a structurally stable xerogel. This material is 99.5% silica, with impurities including Na^+, Ca^{2+}, Fe^{3+}, SO_4^{2-}, and Cl^- ranging from 600 to 60 ppm.

Alumina

Alumina-based materials have a much wider pH stability than silica matrices and therefore are of great interest in chemical separations and industrial-scale chromatography. Commercially available alumina matrices are all aluminium oxides prepared by thermal dehydration of aluminum hydroxide (6,7). The origin of the starting material together with other factors, such as the heating rate, drying time, and temperature, influence the properties of the final product.

Alumina matrices can be obtained in a wide range of particles and pore sizes (Table 2). The physical stability of alumina particles in chromatography columns, even under moderate flow rates and pressure, has been a major problem in our laboratories. Generally, the particles break up and cause a substantial increase in the backpressure generated by the column. Furthermore, any risk of contamination of the purified product with aluminum may not be acceptable to the pharmaceutical and biotechnology industries, particularly when the reports relating to Alzheimer's disease to aluminum are considered (8).

Zirconia and Titania

Recently, there has been an increasing interest in zirconia- and titania-based matrices in chromatography applications, mainly because of their chemical stability over a relatively wide pH range and their physical characteristics. Zirconia or titania supports with properties suitable for preparative chromatography or HPLC are not commercially available, although zirconia and titania powders can be obtained commercially from a number of sources and have been used by researchers to prepare matrices with a small particle size (1–10 μm) and pore sizes of up to 100 nm, suitable for HPLC applications (9). More recently, larger porous zirconia particles (30–150 μm) have been prepared and used in packed and fluidized beds for protein purification (10–13).

Zirconia and titania as base stable matrices not only enable separation of biomolecules at high pH values, but they can also be sanitized with alkaline media, which is the standard sanitization procedure in downstream processing of biological molecules. Various adsorbents such as affinity, ion-exchange, and reversed phase can be prepared from these matrices for protein purification by either polymer coating or reaction with organosilanes (14–17).

Table 1. Properties of Matrex™ Silica

Nominal pore diameter (nm)	Surface area (m^2/g)	Pore volume (ml/g)	Particle size range (μm)
6	540	0.93	20–500
10	320	1.20	45–300
25	160	1.10	45–70
50	90	1.10	50
100	45	1.10	50–130
150	25	1.10	50

Table 2. Properties of Alumina Matrix (Selecto Scientific)

Nominal pore diameter (nm)	Surface area (m^2/g)	Particle size range (μm)
6	120–150	60–120
10	100	60–120
15	75	60–120
30	80	40
50	50	40
100	8	40
180	6–8	20

Hydroxyapatite

Hydroxyapatite is a special crystalline form of calcium phosphate that has been used in the separation of biological molecules. The adsorbent consists of a mosaic of crystals that result from a combination of phosphate and calcium salts after a number of steps of wet alkaline treatment and heat aging. The application of hydroxyapatite to the field of protein separation was first introduced by Tiselius et al. (18). The standard procedures for the operation of hydroxyapatite columns were developed by Bernardi and Gorbunoff (19,20) using a number of different proteins in several solvent systems, Bernardi also proposed a mechanism for protein adsorption to, and desorption from, hydroxyapatite.

Proteins bind reversibly to hydroxyapatite by a complex ionic mechanism different from the principle of ion-exchange chromatography (19,20). Both the amino and carboxyl groups on the protein molecules act in the adsorption mechanism. The adsorption of proteins to hydroxyapatite via the amino groups is the result of their positive charges and the negative charge on the adsorbent when the column is equilibrated in phosphate buffer. The surface of hydroxyapatite crystals consists of sites with both positive (calcium) and negative (phosphate) charges. When the hydroxyapatite columns are equilibrated in phosphate buffers, normally at pH 6.8, the surface of the adsorbent can be regarded as negative because of partial neutralization of the positive calcium loci by phosphate ions. The carboxyl groups act in two ways. First, they are repelled electrostatically from the negative charge of the adsorbent; second, they bind specifically by complex formation with the calcium sites.

The elution of basic proteins from hydroxyapatite is generally carried out by specific displacement with Ca^{2+} and Mg^{2+} ions that neutralize the negative charges of the hydroxyapatite phosphate sites by complex formation. Acidic proteins are eluted by displacement of their carboxyl groups from hydroxyapatite calcium sites by ions, such as fluoride or phosphate, which form stronger complexes.

The difficulty with using hydroxyapatite in medium- and large-scale purification of biological molecules is the physical nature of the material. The crystals are highly fragile, generating small particles resulting in column clogging and precluding any long-term usage of the column. An improved form of hydroxyapatite crystals has recently been prepared and is commercially available from a number of sources, including Bio-Rad. The new ceramic material from Bio-Rad is spherical and has been reported to be stable both chemically and physically.

CONTROLLED PORE GLASS

The method used for the preparation of CPG was originally invented by Hood and Nordberg (21). They discovered that by heat treating a certain sodium borosilicate glass composition, the glass separated into two intermingled and continuous glassy phases. One phase, rich in boric oxide, was soluble in acids; the other phase was high in silica and stable toward acid solutions. This process was used by the inventors to prepare nonporous silica particles by first removing the boron phase with an acid solution to obtain porous particles and then revitrifying this material by heating to 900 °C or above to yield transparent homogenous nonporous particles having a composition of greater than 95% silica.

The porous glass material obtained by this method typically had an average pore diameter of about 4 to 5 nm. A procedure for enlarging the pore size to approximately 20 nm was invented by Chapman and Elmer (22). This was carried out by impregnating a porous glass body with an aqueous solution of a weakly reactive fluorine-containing compound, reacting the compound in situ with a mineral acid to release hydrofluoric acid at a temperature sufficient to dissolve a portion of the glass body and washing the body to remove the soluble constituents. Ammonium bifluoride and nitric acid at 95 °C were applied to a porous glass disc to enlarge the pore size from 4 to 24 nm.

The process for making porous glass was modified by Haller (23) to produce the material with a range of controllable pore sizes. This was achieved by varying the temperature and duration of the heat treatment to obtain porous glass with pore sizes from 12 to 250 nm. Larger pore sizes were also obtained by the addition of a sodium hydroxide leaching process.

Manufacturing Process

CPG can be commercially obtained from a number of sources, and the manufacturing processes are all based on the process invented by Haller (23). The composition of the starting metal oxides mixture and the exact conditions used for the heat treatment and leaching of the glass, which influence the quality and physical properties of the porous glass, are the propriety technology of the manufacturing companies. An outline of the general procedure used for preparing CPG is shown in Figure 1.

CPG is manufactured in a range of particle sizes, 40 to 200 μm, and different pore sizes, 7 to 300 nm (Table 3). The manufacturing batch sizes are generally much smaller than silica particles, possibly because of the heat treatment process. The largest batch size commercially available is in the region of 50 L (20 kg).

Physical Characteristics

Most commercially available CPG granules are irregularly shaped and highly porous (Figs. 2 and 3). CPG has a narrow pore size distribution, with 80 to 90% of the pores showing a deviation less than $\pm 10\%$ from the nominal pore diameter (Fig. 4). CPG has a very high surface area that varies according to the pore diameter and pore volume. The surface area increases with increasing pore volume and decreases with increasing pore diameter, as is shown by the data in Table 3. CPG is a rigid insoluble matrix, compatible with most reagents used in the purification of biological molecules. Because CPG is essentially pure silica, it generally exhibits a high degree of solubility in alkali solutions. The solubility of CPG is a function of temperature, time, and pH of the solution.

The surface of native CPG is slightly hydrophobic and contains acidic hydroxyl groups that could denature bio-

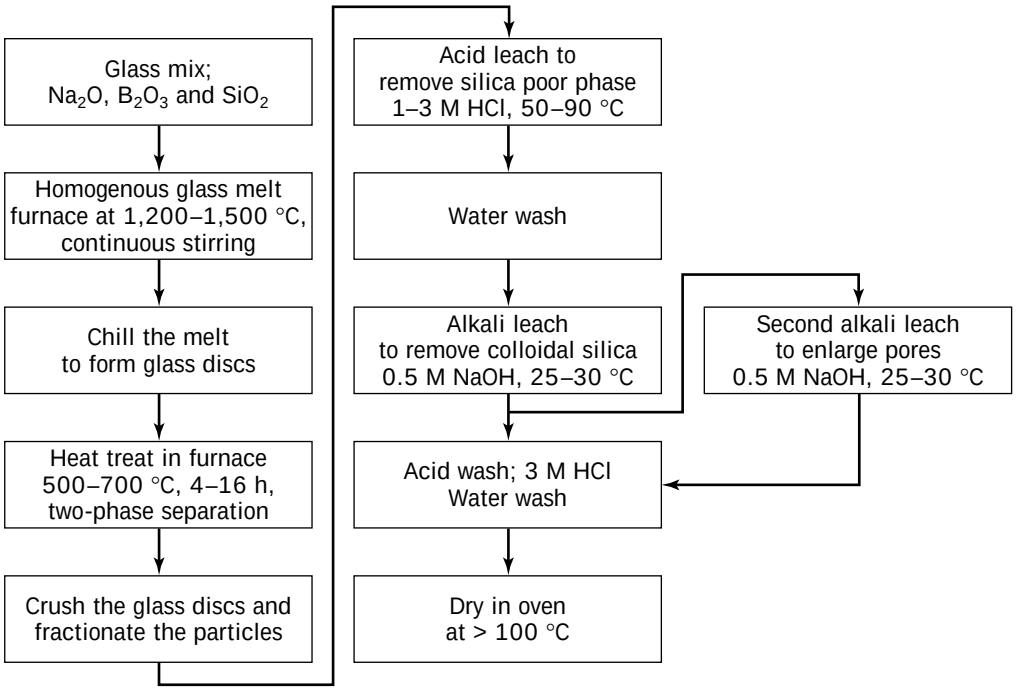


Figure 1. Process for the preparation of porous glass.

Table 3. Properties of Controlled Pore Glass (CPG, Inc.)

Nominal pore diameter (nm)	Surface area (m^2/g); pore volume = 1.0 cm^3/g	Surface area (m^2/g); pore volume = 0.9 cm^3/g
7.5	340	240
12	210	150
17	150	100
24	110	75
35	75	50
50	50	35
70	36	25
100	25	18
140	18	13
200	13	9
300	9	7

logical molecules and cause strong, often irreversible binding of proteins. For most separation applications, surface modification of CPG with organosilanes is necessary, although native CPG has been used for the purification of a number of biological molecules, such as factor VIII (24). Surface modification also enhances the stability and durability of CPG.

Surface Modification

Siliceous materials such as CPG can be reacted with the organosilane compounds in a number of different ways to obtain the desired bonding through hydrolyzable groups of the organosilane. Depending on the type of organosilane, the reaction can be performed in organic, aqueous, or vapour phase, resulting in matrices suitable for different applications. The majority of the work cited on this subject relates to the surface modification of silica material (25), which is not always applicable to CPG because of the difference in the surface characteristics of the two matrices.

Surface modification of CPG is carried out by treating the material with an organosilane containing an organic functional group at one end and a silylalkoxy group at the other. Attachment of the silane compound to the matrix is through the surface silanol or oxide groups to the silylakoxy groups with polymerization between the adjacent silanes. The result is an inorganic matrix with available organic functional groups that can be used in the preparation of chromatography adsorbents. A large number of different organosilanes are commercially available with functional groups such as epoxy, amine, and sulfhydryl (Table 4).

The first step in the surface modification of porous glass is to clean the surface by washing the matrix in nitric acid at elevated temperature. The matrix is then washed with water and dried in the oven at temperatures not exceeding 180 °C. The dry matrix is reacted with the silane compound dissolved in water or an organic solvent at temperatures ranging from 40 °C to 120 °C, depending on the solvent. After removing the excess silane by washing the matrix with suitable solvents or water, the functional groups on the matrix can be further modified to obtain various adsorbents. The exact conditions used in the surface modification of CPG will depend on the nature of the silane compound and the application of the resulting adsorbent.

Extensive optimization of the surface modification procedure is usually required to obtain a suitable adsorbent for a specific application. At Bioprocessing, special tech-

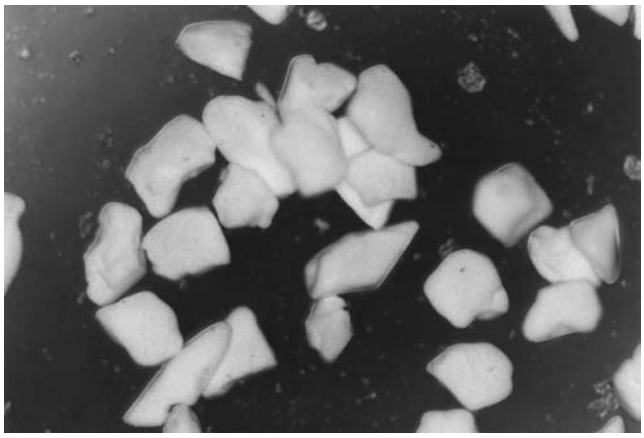


Figure 2. The irregularly shaped granules of CPG matrix.

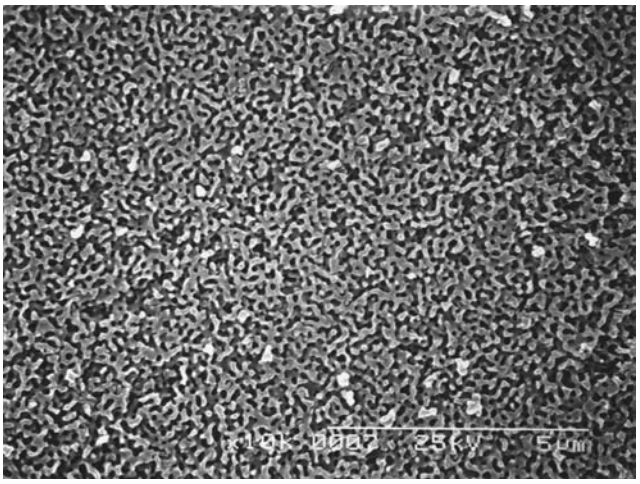


Figure 3. Scanning electron micrograph of CPG matrix.

niques for preparing surface modified CPG (PROSEP®) have been developed for applications in affinity and other types of separation processes. PROSEP® has been designed for effective immobilization of both large and small ligands, resulting in affinity adsorbents with a high capacity for the target molecules, minimum nonspecific binding of contaminating proteins, and low ligand leakage (26).

Applications

CPG has been effective in the separation of a range of biological molecules, using a variety of chromatography techniques that include adsorption, size exclusion, affinity, and ion exchange. It has also been used extensively in the field of enzyme immobilization for the modification of biological molecules. The rigidity and incompressibility of the matrix allows it to be used in large columns at very high flow rates generating low backpressure (Fig. 5), which makes this matrix especially useful for processing large volumes of feedstock in short cycle times.

Application of native CPG in adsorption chromatography has been generally limited because of its denaturing surface characteristics and irreversible binding of proteins. A number of proteins, such as factor VIII (24), have been purified from plasma by first blocking the surface of CPG with albumin. The use of CPG in ion exchange chromatography applications has also been very limited because of the instability of the matrix in sodium hydroxide solution, which is the standard cleaning and sanitization

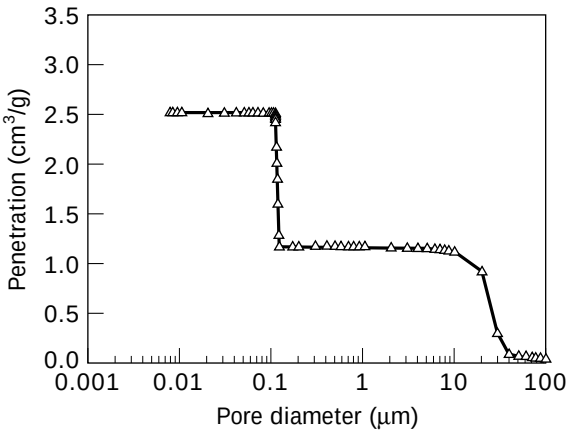


Figure 4. Pore size distribution of CPG matrix as determined by mercury porosimetery method. Average pore diameter was determined to be 0.1125 μ.

solution for ion exchange adsorbents. Furthermore, ion exchange adsorbents prepared with CPG as the matrix have generally shown low functional capacity when compared to the polymetric matrices. Nevertheless, some applications of surface-modified CPG in ion exchange chromatography have been reported (27–29).

The use of CPG in size-exclusion chromatography has generally been for the separation of macromolecules such

Table 4. Surface Modification of CPG

Organosilane	Reactant	Activated CPG
3-Aminopropyltriethoxysilane	—	CPG-amine
3-Aminopropyltriethoxysilane	Glutaraldehyde	CPG-aldehyde
3-Aminopropyltriethoxysilane	Succinic anhydride	CPG-carboxyl
3-Glycidoxypropyltrimethoxysilane	—	CPG-epoxy
3-Glycidoxypropyltrimethoxysilane	Acid solution	CPG-diol
3-Glycidoxypropyltrimethoxysilane	Sodium periodate	CPG-aldehyde
Mercaptopropyltrimethoxysilane	—	CPG-sulfyhyryl

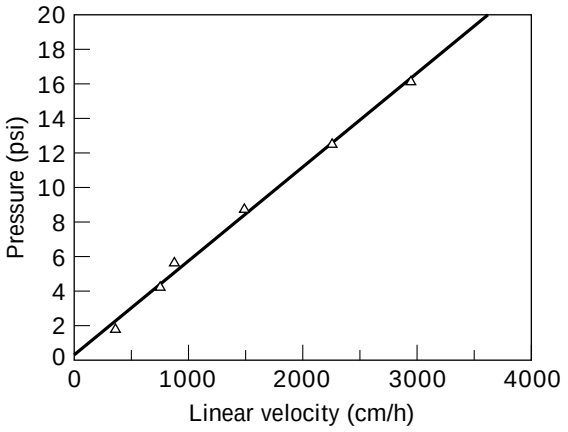


Figure 5. Backpressure generated from a column (11.3 × 10.0 cm) packed with PROSEP®.

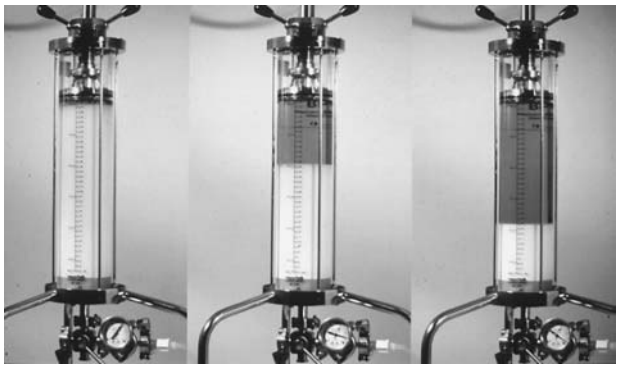


Figure 6. Principle of fluidized-bed adsorption. From left to right: PROSEP® adsorbent is packed into a column specially adopted for fluidized-bed operation. The bed is gradually expanded by an upward flow of solution through the column until full expansion is reached.

as viruses, cell components, and polymers (30,31). Separation of these species from low molecular weight contaminants has been achieved by taking advantage of the very sharp exclusion limits and narrow pore size distribution of CPG. Generally, surface-modified CPG has been used for size-exclusion chromatography to eliminate the problem of nonspecific and irreversible binding of biological molecules to native CPG. The surface of the matrix is modified with glycidoxypropyltrimethoxysilane, converting the solid phase to diol-CPG (Table 4). This is a nonionic, hydrophilic coating compatible with aqueous and most solvent systems (32).

Surface-modified CPG has been widely applied in the area of affinity chromatography for the purification of biological molecules. Biologically active molecules will often bind to other molecules in a reversible and highly specific manner. The formation of these stable, specific, and dissociable complexes forms the basis of this powerful separation technique.

Affinity adsorbents prepared using PROSEP® and other surface-modified CPG have been used in many applications for the purification of biological molecules, such as antibodies (33,34) and enzymes (35). PROSEP® adsorbents have been used in packed-bed chromatography columns and in fluidized or expanded bed mode (Fig. 6) for the direct capture of biological molecules from feedstocks containing particulate or colloidal solids (36,37). Table 5 shows the application of PROSEP® adsorbents in the field of affinity purification.

FUTURE TRENDS

The progress of inorganic matrices such as silica and CPG as adsorbents in the downstream processing of biological molecules is hampered by their instability toward alkaline solutions. Sodium hydroxide solution up to 0.5 M concentration is the standard cleaning and sanitizing reagent for chromatography adsorbents used in the separation processes of biomolecules. Furthermore, certain chemistries used in the surface modification of matrices and some purification protocols require alkaline conditions.

Alkaline-resistant inorganic matrices are under development using different approaches. Development of zirconia and titania matrices by sol-gel techniques have already been mentioned. Surface coating of silica and CPG with an inert metal oxide such as zirconium oxide and preparation of glass–ceramic matrices may result in suitable products that satisfy all the requirements of an ideal matrix for the purification of biomolecules.

Methods for coating silica matrices with zirconium oxide were studied by Meijers et al. (38) and applied to CPG at Bioprocessing. Great improvement in the stability of the

Table 5. Application of Surface-Modified CPG (PROSEP®) in Separation Processes

Product	Trade name	Application
Protein A adsorbent	PROSEP®-A High Capacity	Purification of IgG from all species and subclasses except rat IgG
Protein G adsorbent	PROSEP®-G	Purification of IgG from all species and subclasses
Heparin adsorbent	PROSEP®-Heparin	Purification of antithrombin III and blood clothing factors
Lysine adsorbent	PROSEP®-Lysine	Purification of plasminogen and tPA
Metal-chelating adsorbent	PROSEP®-Chelating	A metal-binding adsorbent used in purification of metal-binding proteins and metal removals from dilute solutions
Thiophilic adsorbent	PROSEP®-Thiosorb	Purification of antibodies
Immobilized antibodies	—	Purification of specific antigens and second antibodies

Note: IgG, immunoglobulin G; tPA, tissue-type plasminogen activator.

coated CPG to alkaline solutions were observed; however, total stability has not yet been achieved.

A preferred approach for making alkaline-stable inorganic matrices is the glass–ceramic route. Porous glass–ceramic with a skeleton of $CaTi_4(PO_4)_6$ crystals has been prepared by controlled crystallization of glasses in the CaO-TiO_2-P_2O_5-Na_2O system and subsequent acid leaching of the resulting dense glass–ceramic (39). The medium pore diameter has been reported to be controllable in the range of 40 to 150 nm by changing the temperature of the heat treatment. Such an approach is being examined at Bioprocessing in collaboration with a glass–ceramic research center for developing alkaline stable inorganic matrices suitable for downstream processing of biological molecules.

LIST OF SUPPLIERS

Bioprocessing Limited, Medomsley Road, Consett, Co Durham, DH8 6SZ, UK

Amicon Products–Millipore Corporation, 80 Ashby Road, Bedford, MA 01730-2271, USA

Selecto Scientific, 5933 Peachtree Industrial Blvd., Bldg. B, Norcross, GA 30092, USA

Bio-Rad Laboratories, 2200 Right Avenue, Richmond, CA 94804, USA

CPG Inc., 3 Borinski Road, Lincoln Park, NJ 07035, USA

BIBLIOGRAPHY

1. E. Boschetti, *J. Chromatogr. A* **658**, 207–236 (1994).
2. S.R. Narayanan, *J. Chromatogr. A* **658**, 237–258 (1994).
3. G.B. Wisdom, *Peptide Antigens: A Practical Approach*, IRL Press at Oxford Univ. Press, London, 1993, pp. 117–136.
4. R.K. Iler, *The Chemistry of Silica*, John Wiley, New York, 1979.
5. K.K. Unger, *Porous Silica*, Elsevier, Amsterdam, 1979.
6. M.A. Coletti-Previero and A. Previero, *Anal. Biochem.* **180**, 1–10 (1989).
7. L.K. Huston, C. Misra, and K. Weters, *Ulmann Encyclopedia of Industrial Chemistry*, Volume A1, VCH, Weinheim, 1985, p. 557.
8. *Lancet* **343**, 993 (1994).
9. J. Nawrocki, M.P. Rigney, A. McCormick, and P.W. Carr, *J. Chromatogr. A* **657**, 229–282 (1993).
10. M.J. Robichaud, A.N. Sathyagal, P.W. Carr, A.V. McCormick, and M.C. Flickinger, *Sep. Sci. Technol.* **32**, 2547–2559 (1997).
11. C.M. Griffith, J. Morris, M. Robichaud, M.J. Annen, A.V. McCormick, and M.C. Flickinger, *J. Chromatogr. A* **776**, 179–195 (1997).
12. A. Mullick, C.M. Griffith, and M.C. Flickinger, *Biotech. Bioeng.* (1998), **63**, 333–340.
13. A. Mullick and M.C. Flickinger, *Prep. Biochem. Biotechnol.* **28**, 1–21 (1998).
14. A.J. Alpert and F.E. Regnier, *J. Chromatogr.* **185**, 375–392 (1979).
15. J.J. Pesek and V.H. Tang, *Chromatographia*, **39**, 649–654 (1994).
16. J. Yu and Z. El Rassi, *J. High Resolut. Chromatogr.* **17**, 705–712 (1994).
17. H.J. Wirth and M.T.W. Hearn, *J. Chromatogr.* **646**, 143–151 (1993).
18. A. Tiselius, S. Hjerten, and O. Levin, *Arch. Biochem. Biophys.* **65**, 132–137 (1956).
19. G. Bernardi, *Methods Enzymol.* **22**, 325–330 (1971).
20. M.J. Gorbunoff, *Methods Enzymol.* **182**, 329–339 (1990).
21. U.S. Pat. 2,106,744 (February 1, 1938), H.P. Hood and M.E. Nordberg (to Corning Glass Works, Corning, N.Y.).
22. U.S. Pat. 3,485,687 (December 23, 1969), I.D. Chapman and T.H. Elmer.
23. U.S. Pat. 3,549,524 (December 22, 1970), W. Haller.
24. J. Margolis and P.H. Rhoades, *Lancet* 446–448 (August 29, 1981).
25. P. Van Der Voot and E.F. Vansant, *J. Liq. Chromatogr. Relat. Tech.* **19**, 2723–2752 (1996).
26. M.A.J. Godfrey, P. Kwasowski, R. Clift, and V. Marks, *J. Immunol. Methods* **160**, 97–105 (1993).
27. S.H. Chang, K.M. Gooding, and F.E. Regnier, *J. Chromatogr.* **120**, 321–333 (1976).
28. A.J. Albert and F.E. Regnier, *J. Chromatogr.* **185**, 375–392 (1979).
29. S.H. Chang, R. Noel, and F.E. Regnier, *Anal. Chem.* **48**, 1833–1845 (1976).
30. T. Darling, J. Albert, P. Russel, D.M. Albert, and T.W. Reid, *J. Chromatogr.* **131**, 383–390 (1977).
31. R.C. Collins and W. Haller, *Anal. Biochem.* **54**, 47–53 (1973).
32. F.E. Regnier and R. Noel, *J. Chromatogr. Sci.* **14**, 316–320 (1976).
33. K. Beyzavi and H.C. Wood, *Separation for Biotechnology*, Elsevier Applied Science, London, 1990, pp. 452–461.
34. M.A.J. Godfrey, *Affinity Separations: A Practical Approach*, IRL Press at Oxford Univ. Press, London, 1997, pp. 141–195.
35. P.J. Robinson, P. Dunnill, and M.D. Lilly, *Biochim. Biophys. Acta* **242**, 659–661 (1971).
36. K. Beyzavi and R. Clift, *Genetic Engineering News* **14** (March 1, 1994).
37. C. Spence, C.A. Schaffer, S. Kessler, and P. Bailon, *Biomed. Chromatogr.* **8**, 236–241 (1994).
38. A.C.Q.M. Meijers, A.M. de Jong, L.M.P. van Gruijthuijsen, and J.W. Niemantsverdriet, *Appl. Catal.* **70**, 53–71 (1991).
39. H. Hosono, Y. Sakai, and Y. Abe, *J. Non-Cryst. Solids* **139**, 90–92 (1992).

See also ADSORPTION, EXPANDED BED; ADSORPTION, PROTEIN, BATCH; ADSORPTION, PROTEINS WITH SYNTHETIC MATERIALS.

ADSORPTION, EXPANDED BED

ANN-KRISTIN BARNFIELD FREJ
Amersham Pharmacia Biotech
Uppsala, Sweden

KEY WORDS

Capture
Expanded bed
Chromatography
Downstream process
Fluidized bed
Protein adsorption
Protein purification
Recovery

OUTLINE

INTRODUCTION

Expanded-bed adsorption is a technique based on fluidization for recovery of proteins directly from unclarified feedstocks such as fermentation broths, cell homogenates, and other solutions containing suspended particulates. An expanded bed is a chromatography bed with an increased void volume between the adsorbent particles, as compared to a packed chromatography bed. An expanded bed is formed when high-density adsorbent particles are lifted by an upward liquid flow through a column specially designed for expanded bed adsorption. When the bed has expanded and is stable, the particulate-containing feedstock is applied, and the protein of interest can bind to the adsorbent particles while the particulates in the feedstock can pass unhindered through the void of the bed. Expanded-bed adsorption is a one-passage technique designed for large-scale recovery processes. Traditionally, when conventional packed-bed chromatography is used to recover proteins, the feedstock has to be clarified by centrifugation or filtration prior to application or the particulates can block the bed. Expanded-bed adsorption can handle particulate-containing feedstocks; as a result, clarification, concentration, and initial purification are accomplished in one unit operation. Reducing the number of unit operations in a downstream processing scheme decreases the risk of product loss, decreases the processing time, and lowers the overall production costs. The basic principles of expanded-bed adsorption and some applications are described in this article.

BACKGROUND

The biotechnology industry produces proteins for therapeutic and diagnostic use by using biological systems such as bacteria, yeast, and mammalian cells. A common bottleneck in the production of these proteins is the initial purification, during which the proteins are recovered from large volumes of cell- and debris-containing feedstocks, such as fermentation broths and homogenates. The purification processes are usually based on packed-bed chromatography techniques in which the protein is bound to an adsorbent packed in a column. A packed bed (fixed bed) cannot handle particulates very well and would eventually become totally blocked if an unclarified feedstock was passed through it. Another concern with these types of feedstocks is that because of the sometimes low production level of the protein of interest, the feedstock volume is large. A large volume takes a long time to pass through a chromatography column, and a long processing time increases the risk of proteolytic breakdown of the protein. Unit operations to clarify and reduce the volume of these feedstocks are therefore vital for a successful purification process. It is not uncommon for a purification process to consist of a sequence of unit operations just to achieve initial recovery. The techniques traditionally used for feedstock clarification are centrifugation and filtration (1).

Centrifugation is well suited to industrial processes because the throughput is high and it can be performed in a contained environment (a requirement when handling recombinant host organisms). The efficiency of the centrifugation process will depend on the properties of the feedstock, but it is usually difficult to remove all particulates by centrifugation, and the resulting fluid may therefore not be sufficiently free of particulates for direct application to a packed chromatography bed (2). Whole cells are relatively easy to remove by centrifugation but disrupted cells are not, and cells (especially mammalian) may even be damaged by the centrifugation process (3) because of shear stress and centrifugal forces. New and improved centrifuges that are more gentle can be used for mammalian cells (4), but often the clarification problem remains because at harvest, culture fluids usually contain relatively large amounts of subcellular particulates, such as debris. When the feedstock is secreting bacteria or a bacterial ho-

mogenate, for example, clarification is even more difficult and centrifugation is not very efficient.

When filtration is used for clarification of feedstocks, it is usually performed in cross-flow mode (5,6). The fluxes that can be used will depend on the properties of the feedstock, but it is often difficult to reach an acceptable economic limit (100 $L/m^2/h$) (7) for a cross-flow filtration process because fluxes are typically less than 50 $L/m^2/h$. Cross-flow filtration can easily be operated as a completely contained system, and it yields a particulate-free solution. The fouling of the filtration membranes is a major concern, and the resulting decrease in flux per filter area often leads to long processing times. In addition, the cost of membranes must be taken into account in a process where membrane fouling is severe because the membranes have to be replaced frequently. Because of the drawbacks associated with both the centrifugation and the filtration techniques, they are often used in combination in a purification process to achieve the best clarification of the feedstock.

The volume of the feedstock is not significantly reduced by either centrifugation or filtration, which is why product concentration does not usually increase until the protein has been adsorbed to a chromatography adsorbent in a packed-bed column. Because the feedstock volume is not reduced before application to the packed bed, processing time will be long. Thus, for a number of reasons, the best option for successful initial recovery of a protein from a large-volume cell- or debris-containing feedstock is direct adsorption of the protein.

A variety of methods for direct recovery have been tried. One such method is to use a batch procedure in which an adsorbent is added directly to the unclarified feedstock in a stirred tank (8,9). The advantage of using a batch procedure is that the feedstock does not have to be clarified. The disadvantage is that adsorption of protein to the chromatography adsorbent is inefficient, because the stirred tank acts as one single theoretical plate in a separation process (10). Several attempts, however, have been made to improve batch adsorption, for example, continuous-batch adsorption processes (11).

Another alternative for direct recovery is fluidized bed adsorption. This technique uses high-density adsorbent particles in an upflow fluidized bed column (12,13). As with a batch procedure, a fluidized bed behaves like a column with a low number of theoretical plates. The adsorbent particles and the feedstock are constantly "back-mixed," which usually results in poor adsorption of the protein by the adsorbent particles, making it necessary to recirculate the feedstock through the column, leading to long processing times. A number of solutions to improve the efficiency of fluidized beds have been suggested, such as columns fitted with perforated horizontal plates (14). This creates a column with a number of smaller, each completely back-mixed, fluidized beds. The number of fluidized beds depends on the number of plates inserted. The end result is a column with more than one theoretical plate and consequently a more efficient process. Another way to improve the efficiency of a fluidized bed is to use magnetically stabilized fluidized beds (15). The adsorbent particles are magnetically susceptible and are fluidized while being stabilized by a magnetic field. Stabilizing the particles in this way prevents back-mixing and gives the column several theoretical plates. There are, however, drawbacks that have prevented this technique from being used at industrial scale, such as the heat generated by the magnetic field.

The concept of expanded-bed adsorption (or expanded-bed chromatography) was introduced after another attempt to stabilize and prevent back-mixing in a fluidized bed. An expanded bed is a chromatography bed in which the bed voidage is larger than in a settled or packed bed. Furthermore, it displays no or very low back-mixing of adsorbent and liquid. An expanded bed can be also be described as a stable fluidized bed because the liquid flow shows a constant velocity profile over the whole cross-section of the bed (i.e., plug flow) (Worth mentioning is that the terms *fluidized bed* and *expanded bed* are used differently in the literature, and sometimes authors do not differ between the expressions.) The early expanded bed experiments were performed using Sepharose™ Fast Flow (Amersham Pharmacia Biotech, Uppsala, Sweden) as the adsorbent and a sintered glass filter as the liquid flow distributor at the bottom of the column (16). The results showed that if the bed expansion was limited to approximately twice the settled bed height, the liquid flow through the column was close to plug flow and the protein adsorption efficiency was similar to that of a packed bed. These promising results showed that in an expanded bed it is possible to combine the fluidized bed's ability to handle unclarified feedstocks with a packed bed's characteristic of efficient protein adsorption. The adsorbent used in these early experiments is intended for use in packed beds. The density of the particles is relatively low, which meant that the flow through the expanded bed column also had to be low, otherwise the bed would expand to an extent that adsorbent would start to pack at the top of the column. For this technique to be attractive at industrial scale, the flow velocities must be significantly increased. The sintered glass filter that was used as a flow distributor is not appropriate for large-scale operations for a number of reasons: particulates may get trapped in the filter, the filter is difficult to clean, and it is difficult to produce large homogenous sinters. Hence, it soon became apparent that for expanded-bed adsorption to be successful at industrial scale, new types of adsorbent particles and new column designs were needed.

Many different types of adsorbents have been tested: agarose (17,18), zirconia, silica (19), glass (20,21), hydrophilized perfluorocarbon (22), and composites such as kieselguhr/agarose (23), cellulose/titanium dioxide (24), and dextran/silica (25). These adsorbents are suitable for certain applications, but they lack the combination of properties needed for an efficient expanded-bed adsorption process. Agarose adsorbents are well suited for protein recovery but they are not dense enough to allow high flow velocities. Silica cannot withstand the harsh cleaning and sanitization conditions that are used in early recovery processes (26). Zirconia particles are the most dense, are stable at high pH, and are sterilizable. Porous glass and the described perfluorocarbons have promising sedimentation properties, but porous glass sometimes shows quite low protein-binding capacities (21), and the perfluorocarbons

have low protein capacities because of the low particle porosity (22).

With the recent development of STREAMLINE™ (Amersham Pharmacia Biotech) adsorbents (agarose/crystalline quartz composite, agarose-dextran/crystalline quartz composite, and agarose/metal alloy composite) and STREAMLINE™ columns, it was possible to secure a recovery system that has the properties needed for an efficient expanded-bed process. The columns have a flow distributor in the bottom that generates an even flow distribution over the whole cross-section of the bed. The density of the adsorbent permits process flow velocities (300–500 cm/h), and the size distribution creates a classified bed that enhances the stability of the expanded bed. The result is a bed with protein adsorption characteristics similar to those of a packed chromatography bed. Expanded-bed adsorption is a scalable, one-passage technique that is intended for industrial-scale operations (27).

There have been a number of applications reported that use expanded bed adsorption with various feedstocks: secreting bacteria (28), periplasmic preparations of bacteria (29), bacterial homogenates (30,31), secreting yeast (32–34), yeast homogenates (35,36), mammalian cells (37–45), milk/whey (46–48), and inclusion body preparations (49).

THEORY

As mentioned earlier, an expanded bed is a fluidized bed that is devoid of or shows low back-mixing of liquid and adsorbent. The stability of the expanded bed makes it resemble a packed bed in that it shows an almost constant flow velocity profile and yields a bed with an increased number of theoretical plates (relative to a fully mixed fluidized bed). These characteristics are achieved by a combination of column and adsorbent design.

Column

The liquid flow in a column used for expanded-bed adsorption must be uniform over the entire cross-section of the bed. A conventional packed-bed chromatography column with one or a few inlets cannot meet this criterion. In a packed bed, the bed itself creates a pressure drop that achieves the desired plug flow. In an expanded bed, there is a very low pressure drop over the column, thus requiring a specially designed flow distributor at the bottom of the column (50). This distributor must ensure that the flow is directed in a vertical direction only, because radial flow inside the column will cause disturbances in the expanded bed. It must also prevent the adsorbent particles from leaving the column while allowing cells and debris to pass through. Other requirements are that it must not create high shear stresses, or cells and shear sensitive molecules may break, and it should be designed to facilitate cleaning. There should be no stagnant zones in the column where cells and other contaminants can accumulate and, in the worst case, escape cleaning.

Adsorbent

The adsorbents used for expanded-bed adsorption must exhibit additional properties to those required of an adsorbent used in packed-bed chromatography. The density of the particles must be relatively high to enable high liquid flow velocities through the expanded bed without the adsorbent particles packing against the adaptor. The adsorbent particles in a fluidized/expanded bed are in constant motion; however, an efficient protein adsorption process requires that the solid dispersion (the movement of the adsorbent particles) is reduced to a minimum. This can be achieved if the adsorbent is polydispersed with respect to size, that is, the particles exhibit a size distribution. When such a polydispersed adsorbent with a small variation in particle density is expanded, each particle will find its equilibrium position in the upward flow. The result is called a classified bed (51), and at a closer look the larger particles will be found at the bottom of the bed and the smaller ones at the top. A distribution of density with size (a large particle exhibiting a larger density than a small particle) will further enhance the classification of the bed. One way of describing the behavior of such an adsorbent in an expanded bed is the Richardson-Zaki (52) model for monodispersed particles:

$$u_0/u_t = \epsilon^n$$

where u_0 is the superficial velocity of the fluid, u_t is the terminal falling velocity of a particle in infinite dilution, ϵ is bed voidage, and n is an index. The model describes expansion for a bed that has reached equilibrium (the expansion height is constant). The terminal falling velocity u_t is described by Stoke's law:

$$u_t = ((d_p)^2 g(\rho_p - \rho))/18\mu$$

where d_p is particle diameter, g is acceleration due to gravity, ρ_p is particle density, ρ is fluid density, and μ is fluid viscosity. Stoke's law can only be used for particles with a Reynold number lower than 0.2. The Reynold number Re_p for a particle in a fluid is defined by:

$$Re_p = u d_p \rho/\epsilon\mu$$

The Richardson-Zaki model is corrected for the particle size distribution of the adsorbent by using the perfectly classified bed model (51):

$$\Delta H(\epsilon) = H_0(1 - \epsilon_0)\Delta F_v/(1 - \epsilon)$$

where H_0 is sedimented bed height, ϵ_0 is sedimented bed voidage, ΔF_v is the proportion of the total volume of particles having the diameter within the interval $d - \Delta d/2 < d < d + \Delta d/2$, and ϵ is bed voidage for the particles within that interval. This means that the particle size group between d and $(d - \Delta d)$ is found within the expanded bed between heights h and $(h + \Delta h)$. Thus, each segment of height is considered a bed with a specific voidage that can be calculated using the Richardson-Zaki equation. The total bed height is obtained by adding up the heights of each segment.

EQUIPMENT AND METHODOLOGY

Column

The major difference between a standard packed-bed column and a column used for expanded-bed adsorption is the liquid distribution system. STREAMLINE™ columns are specially designed for expanded bed adsorption and have a fixed bottom distributor and a movable top distributor (adaptor). The bottom and top distributors have essentially the same design: a perforated plate with a single-weave screen or net facing the adsorbent. The distributor plate in the adaptor has a lower pressure drop than the bottom distributor. The purpose of the top distributor plate is to create a more even flow (compared with a column with one or a few inlets in the adaptor) when the column is run in downward mode during elution in settled bed mode. The steel screens help to stabilize the flow and confine the adsorbent particles to the column. These columns do not cause cell damage, and it has been shown that shear-sensitive cells pass through the column without breaking (40,42).

Expanded-bed adsorption handles unclarified feedstocks, and for this reason, harsh cleaning procedures are necessary. The columns are constructed of materials that withstand the chemicals used in such procedures and are also designed to have the minimum of stagnant zones where cells could accumulate and risk contaminating the next run. It has been shown that the columns can be efficiently cleaned and sanitized after passage of unclarified feedstocks such as whole cultures of bacteria and yeast (27).

Columns range in size from small lab-scale columns with a 25-mm inner diameter to industrial-scale columns, presently the largest with a 1,200-mm inner diameter. They all have essentially the same design, but differ in the material used for the column tube: the smaller columns have tubes made of glass and industrial columns have tubes made of stainless steel. The larger columns also have a greater number of inlets and a higher pressure drop over the distributors than the smaller columns. One method of verifying successful scale-up and comparing the performance of different sizes of expanded-bed columns is to determine the axial dispersion of a tracer substance (the method is described in "Operation of an Expanded Bed"). Another method is to determine protein breakthrough capacity using a model protein in a buffer system. The results obtained from these methods give a good measure of bed stability, a factor that is crucial for a successful expanded-bed process. The hydrodynamic properties of the columns are tested using the axial-dispersion method, and the "function"—the protein adsorption properties—is evaluated using the protein breakthrough method. The columns have been tested using both methods, and the results demonstrate that the performance of an industrial-sized column is very similar to that of a small lab-scale column (27). Table 1 shows that the number of theoretical plates, and the axial dispersion varies little between the different column sizes. Figure 1 compares protein breakthrough curves (bovine serum albumin [BSA] at 300 cm/h using 15-cm settled bed height), illustrating the maintained protein adsorption characteristics from lab scale to industrial scale.

Table 1. Number of Theoretical Plates (N and Axial Dispersion (D_a) for Expanded-Bed Columns with Inner Diameters from 25 to 600 mm

Column	N	D_a (m^2/s)
STREAMLINE™ 25	50–70	$3–5\ 10^{-6}$
STREAMLINE™ 50	40–60	$4–6\ 10^{-6}$
STREAMLINE™ 200	35–50	$5–7\ 10^{-6}$
STREAMLINE™ 600	35–50	$5–7\ 10^{-6}$

Note: The experiments were performed at 300 cm/h flow velocity and 15 cm settled bed height adsorbent using acetone as tracer substance.

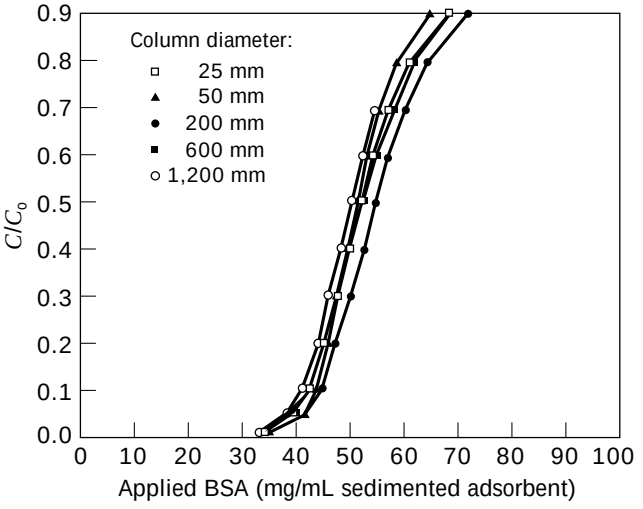


Figure 1. Protein breakthrough capacity. Breakthrough curves for bovine serum albumin (BSA) on STREAMLINE™ DEAE (15 cm settled bed height at 300 cm/h flow velocity) using STREAMLINE™ expanded-bed columns with inner diameters ranging from 25 to 1,200 mm. C_0 is the concentration of BSA in the applied feedstock, and C is the concentration in the flowthrough. *Source*: Copyright Amersham Pharmacia Biotech, Uppsala, Sweden, modified and reproduced with permission.

Because industrial-scale columns have to be automated, these columns are equipped with an adsorbent sensor that locates the position of the expanded bed surface. In response to the signal from the sensor, the adaptor is automatically adjusted to its correct position. Figure 2 shows different scales of expanded-bed columns.

UpFront™ Columns (UpFront Chromatography A/S, Denmark) offer another approach to expanded-bed column design (53). The columns are without screens and without distribution plates, and the feedstock is introduced at the bottom of the column, right below a stirring device that distributes the feedstock over the cross-section of the column. The stirrer causes a mixed zone, but gradually above that zone plug flow is obtained. Columns of 20-, 50-, and 200-mm inner diameter are available.

BioProcessing Ltd. (England) uses modified standard chromatography columns for expanded bed adsorption (54). The standard column bottom support net or screen is replaced with a 300-μm nylon mesh, and then the column

Figure 2. Different sizes of STREAMLINE™ expanded bed columns: 25-mm, 600-mm, and 1,200-mm inner diameters. *Source*: Courtesy of Amersham Pharmacia Biotech, Uppsala, Sweden.

is packed with a layer (approximately 3 cm in height) of surface-modified solid glass beads (0.4–0.52 or 0.85–1.23 mm in diameter). This bed of beads serves as a distributor of fluid to the bed. The adsorbent particles are then added directly on top of the distributor beads. During operation of the column, the bed of the relatively heavy distributor beads remains static, and the bed of adsorbent particles expands.

Adsorbent

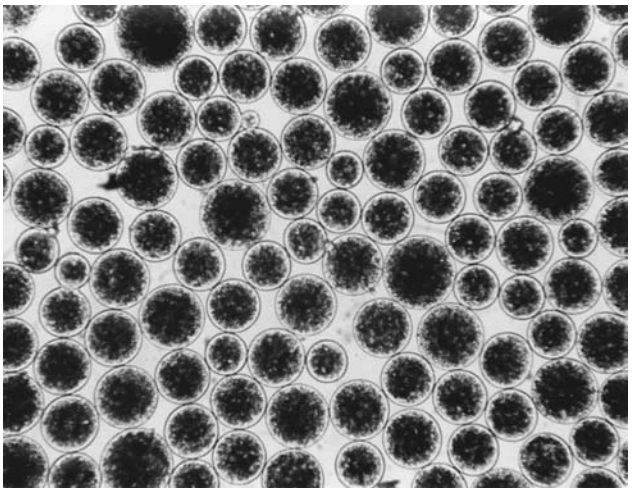
STREAMLINE™ adsorbents are specially designed for expanded bed adsorption. To ensure that the bed does not expand too much at industrial chromatography flow velocities, a dense core material has been incorporated into the agarose particles to increase the apparent density. The adsorbent has a defined Gaussian-like particle size distribution, which yields a stable classified bed when the adsorbent is expanded. The influence of particle size distribution on protein adsorption performance has been studied using three different particle fractions (120–160 μm, 250–300 μm, and 120–300 μm) (55). The results showed that a wide particle size distribution gives an expanded bed with low axial dispersion and that small particles give high dynamic protein capacity (due to reduced diffusion length and large surface area). The results also showed that with increasing (settled) bed heights, the dynamic binding capacity increases over a wide range of flow velocities (due to reduced axial dispersion and increased numbers of transfer units for pore and film diffusion).

The STREAMLINE™ DEAE, Q XL, SP, and SP XL, chelating and heparin adsorbents are spherical, macroporous, cross-linked 6% agarose particles with crystalline quartz as core material. The Q XL and SP XL adsorbents have dextran chains coupled to the agarose, and the Q and SP charged groups are attached to the dextran through chemically stable ether bonds. This increases the exposure of the Q and SP groups and results in very high protein binding capacities. The adsorbents can be cleaned with 1 M NaOH. The matrix used for the protein A adsorbent contains highly cross-linked 4% agarose with a metal alloy core. The alloy is stable and composed of 77% Ni, 15.5% Cr, and 7.5% Fe. No leakage of metal was detected in 0.1 M glycine, pH 3.0, a solution commonly used in protein A chromatography. The metal leakage was below the detection limit of the analysis method (0.02 μg/mL eluent using induced coupled plasma atomic-emission spectroscopy). The protein A adsorbent is stable in all aqueous buffers commonly used in protein A chromatography and cleaning. No significant change was observed in chromatographic performance after either 1 week's storage or 100 cycles, normal use, at room temperature with 10 mM HCl (pH 2), 0.1 M sodium citrate (pH 3), 1 mM NaOH (pH 11), 6 M GuHCl, or 20% ethanol. Figure 3 shows micrographs of the adsorbent particles, and Table 2 summarizes the properties of the various STREAMLINE™ adsorbents.

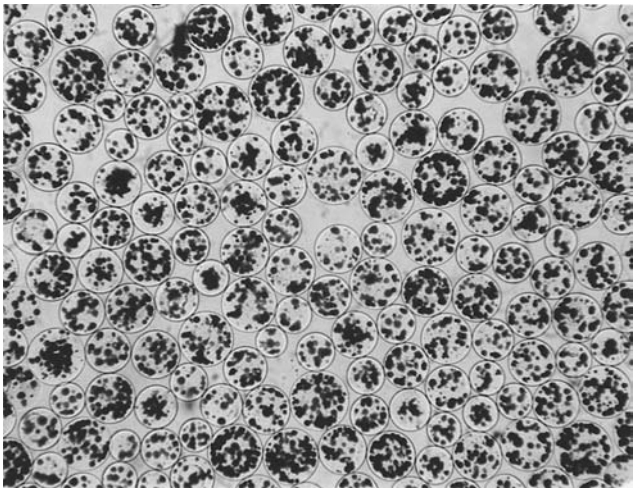
Apart from the purpose-designed column and adsorbent, the equipment and the experimental set-up used for expanded bed adsorption is very similar to that used for packed bed chromatography. An example of an industrial-scale set-up with a 600-mm-diameter column is seen in Figure 4.

Method Development

The strategy for developing an expanded-bed process is in general analogous to that of a conventional packed-bed process. In order to save time, it is advisable to do the first experiments with clarified feedstock, using the expanded-bed adsorbent in a small, packed-bed column, for example, 15 mm in diameter and with 15-cm bed/height. (A bed height below 10 cm of these relatively large particles usually results in a decreased dynamic binding capacity. The increase in dynamic binding capacity with increased bed height is an effect of longer residence time for the sample.) From the packed-bed experiments, the most suitable adsorbent is selected (e.g., ion exchanger or chelating adsorbent), and the conditions for protein binding and elution are determined. It is important to establish that the conditions selected for protein binding are compatible with the properties of the unclarified, cell- or debris-containing feedstock (e.g., low pH might cause aggregation). Such problems are not apparent with clarified feedstock and



(a)



(b)

Figure 3. STREAMLINE™ adsorbents. Micrographs of STREAMLINE™ DEAE (**a**) and STREAMLINE™ rProtein A (**b**). *Source*: From *Bioseparation and Bioprocessing, Vol. 1 Biochromatography: Expanded Bed Adsorption Chromatography*, Wiley-VCH Verlagsgesellschaft, Weinheim, Germany, 1998.

should be investigated before the process has progressed too far. Once the most promising method has been selected from the packed-bed experiments, it is then optimized in a lab-scale expanded-bed column using unclarified feedstock. At this stage, the effect of the now present cells or debris on protein binding is evaluated. In theory, if the net surface-charge of the cells and debris is negative, the protein and the cells or debris may compete for binding sites on an anion exchanger. A consequence of this may be that the feedstock load per milliliter adsorbent has to be decreased. When STREAMLINE™ DEAE was used to recover insulinase from whole yeast culture, the presence of cells did not cause any reduction in protein capacity (33). However, it has been reported that when the same adsorbent was used to recover glucose-6-phosphate dehydrogenase from a yeast homogenate, the presence of cell debris did cause a slight reduction in capacity of the adsorbent (36). Feedstock properties and their effects on the expanded-bed process are discussed further in "Feedstocks." Before the process is transferred into larger columns, appropriate cleaning procedures should be determined. The whole process is then verified at pilot scale before scaling up to the final production scale.

Operation of an Expanded Bed

Figure 5 is a schematic representation of how an expanded bed is operated—essentially as a packed bed. Avoid getting air into the column; air trapped in the bottom distributor may disturb the bed. If air does enter the column, it can usually be removed with high-flow velocity pulses combined with back-flushes (reversal of flow direction). Before start-up always check that the column tube is positioned vertically; otherwise, flow will be turbulent when the liquid hits the inside walls of the column tube. The height of the settled bed (H_0) is noted before start-up.

Bed Expansion and Equilibration. The bed is expanded and equilibrated with buffer. It usually takes 20–30 min at a flow velocity of 300 cm/h before the bed is stable and has stopped expanding. The height of the expanded bed (H) is then noted. At this stage, the bed can be checked empirically, before application of the feedstock, by determining the axial dispersion of a tracer substance. Figure 6 shows the UV signal recording from such a test procedure. Before applying the tracer substance, the adaptor is lowered to a position where the adaptor screen is 0.5–1.0 cm above the expanded-bed surface. When the signal from the UV monitor is stable, the solution is changed to buffer-tracer substance mixture (e.g., 0.25% acetone v/v), that is, positive step input signal. The solution is changed to buffer when the UV signal is stable at maximum absorbance (100%), that is, negative step input signal. The change is marked on the chart recorder paper, and the UV signal is allowed to stabilize at the baseline level (0%). The number of theoretical plates (N) is calculated from the negative step input signal (57):

$$N = t^2/\sigma^2$$

where t is the mean residence time, the distance from the mark on the chart recorder paper to 50% of the maximum absorbance; and σ is the standard deviation, half the distance between the points 15.85% and 84.15% of maximum absorbance.

The procedure gives the number of theoretical plates not only in the column but also the contribution from pumps, valves, tubing, and so forth. Because different experimental set-ups may vary in their contribution to N, the results should only be compared between runs performed using the same set-up. The negative input signal is recommended for evaluation (at least when acetone is used as tracer) because the reproducibility of the results is higher than that obtained from the positive step input signal (from 0% to 100%).

The relation between N and the axial dispersion coefficient (D_a) is

Table 2. Properties of STREAMLINE™ Adsorbents Used in Expanded-Bed Adsorption

Adsorbent name and base matrix	Ligand and functionality	Particle size distribution (μm)	Mean density (g/mL)	Ligand density	Breakthrough capacity (mg/mL)	Total binding capacity (mg/mL)
STREAMLINE™ DEAE Agarose/quarts	DEAE Anion exchange	100–300	1.15	0.13–0.21	35	55
STREAMLINE™ Q XL Agarose-dextran/quartz	Q Anion exchange	100–300	1.2	0.23–0.33	>110	Not done
STREAMLINE™ SP Agarose/quartz	SP Cation exchange	100–300	1.15	0.17–0.24	65	75
STREAMLINE™ SP XL Agarose-dextran/quartz	SP Cation exchange	100–300	1.2	0.18–0.24	>140	Not done
STREAMLINE™ Chelating Agarose/quartz	IDA Metal affinity	100–300	1.15	40	8	40
STREAMLINE™ Heparin Agarose/quartz	Heparin Affinity	100–300	1.15	4	Not done	Not done
STREAMLINE™ rProtein A Agarose/metal alloy	rProtein A Affinity	80–165	1.3	6	20	50

Note: Adsorbent name: XL stands for extreme load and consists adsorbents with very high loading capacities. *Ligand*: DEAE, diethyl aminoethyl; Q, quatenary amine; SP, sulphopropyl; IDA, iminodiacetic acid; rProtein A, recombinant protein A. *Functionality*: Functionality of STREAMLINE Heparin was evaluated by its ability to bind antithrombin from bovine plasma. *Ligand density*: Ligand densities for STREAMLINE DEAE, Q XL, SP, and SP XL were determined as millimole charged groups per milliliter of adsorbent. For STREAMLINE Chelating, ligand density is expressed as Cu^{2+} binding capacity in μmole per milliliter adsorbent. Ligand density for STREAMLINE Heparin is milligram of heparin bound per milliliter adsorbent. For STREAMLINE rProtein A, ligand density is expressed as milligram protein A per milliliter adsorbent. *Breakthrough capacity*: For STREAMLINE DEAE, SP, and rProtein A, breakthrough capacity was determined at 300 cm/h linear flow velocity in expanded-bed mode (15 cm settled bed height) at 1% breakthrough of 2 mg/mL bovine serum albumin (DEAE) or lysozyme (SP) or human immunoglobulin G (IgG) (rProtein A). For STREAMLINE Q XL AND SP XL, the breakthrough capacity was determined at 300 cm/h in packed bed mode (10 cm bed height) at 10% breakthrough of 2 mg/mL bovine serum albumin (Q XL) or lysozyme (SP XL). Breakthrough capacity for STREAMLINE Chelating was determined in packed bed mode at 300 cm/h flow velocity (5 cm sedimented bed height) at 1% breakthrough of 2 mg/mL bovine serum albumin. *Total binding capacity*: Total binding capacity was determined in packed bed mode at 50 cm/h flow velocity (15 cm bed height) at 100% breakthrough of 2 mg/mL bovine serum albumin (DEAE, Chelating), lysozyme (SP), or human IgG (rProtein A).

$$D_a = (u_0 \cdot H)/(2 \cdot N \cdot \epsilon)$$

where u_0 is the superficial velocity, H is the expanded bed height, and ϵ is the bed voidage in expanded mode. ϵ is calculated from $H/H_0 = (1 - \epsilon_0)/(1 - \epsilon)$. In a settled bed, the bed voidage (ϵ_0) is approximately 0.4. The higher the value for N, the lower the value for D_a, indicating a low degree of back-mixing (a flow profile approaching plug flow). The relation between N and D_a is derived from the coupling between the tanks in series and the axial dispersion models by comparing the variances between the models' step-response curves. It is not completely accurate to make this comparison because the responses of the two different models are never identical (56). But for a system with a low degree of back-mixing, this way of calculating the axial dispersion is accurate.

Feedstock Application. The feedstock is applied in expanded bed mode, usually at a constant flow velocity (e.g., 300 cm/h). The adaptor is kept at its uppermost position, or at a position 5–10 cm above the expanded-bed surface. The uppermost position is usually the choice when a new (unknown properties) feedstock is applied or when there are feedstock batch variations, making it difficult to predict how the feedstock will behave in the expanded bed from run to run. If feedstock properties, such as viscosity or biomass or others that affect expansion, are known to be consistent and do not vary, then the adaptor can be left at a position closer to the bed surface. The feedstock is always stirred during application to prevent cells and debris from settling. If there is a build-up of cells and debris on the adaptor, the column is intermittently back-flushed. The duration of a back-flush is about 5–10 seconds, and the interval between back-flushes is usually 10 to 30 minutes, depending on the properties of the feedstock. Another method of applying the feedstock is to change the flow velocity (normally lowering it) during the run, so that the expansion height is kept constant. It has been reported that this approach, under certain conditions, may give a higher productivity than the constant flow velocity method does (58).

Wash. The wash is performed in expanded-bed mode at the highest possible flow velocity. This means that the flow velocity at the beginning of the wash is often the same as during feedstock application, but as the bulk of the cells and debris are washed out, the flow velocity is increased. At the beginning of the wash, the column is usually back-flushed as described in the previous section. During the wash, when the bulk of the particulates has been washed out, the adaptor is lowered closer to the bed surface. Lowering the adaptor significantly reduces the volume of wash buffer and shortens the time required for the wash step. If, instead, the wash is performed with a wide gap between the bed surface and the adaptor, mixing of the liquids is usually severe (a low density wash buffer being introduced after the high density/viscosity feedstock), resulting in a large wash volume. After the wash, when the UV curve has levelled out, the column is back-flushed to remove any trapped cells or debris from the distribution system. The bed is allowed to settle and the adaptor is lowered to the bed surface. The wash is continued in settled-bed mode, in

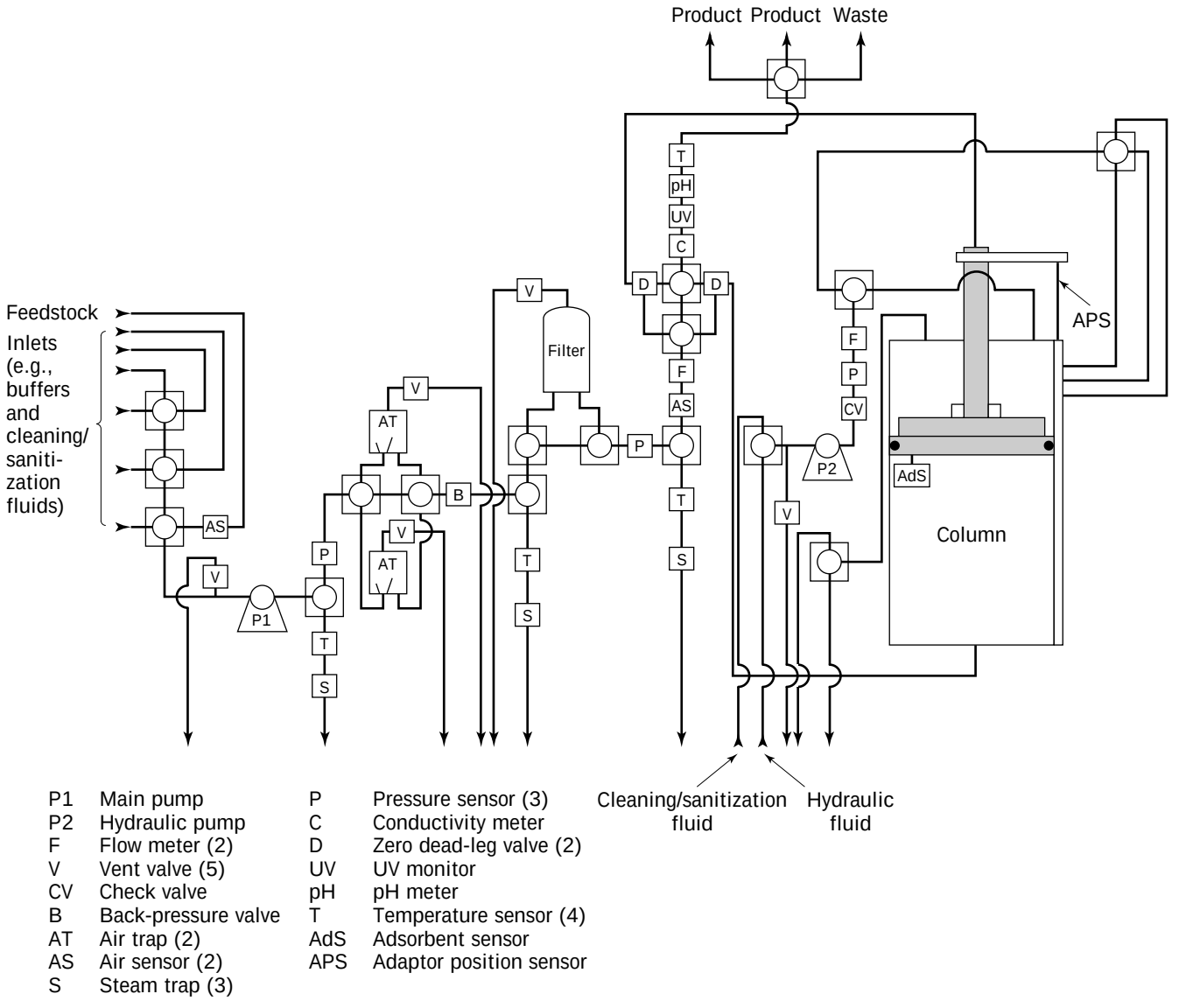


Figure 4. Experimental set-up for production-scale expanded-bed adsorption using a 600-mm inner diameter column. *Source*: From Ref. 27.

a downward flow direction at 300 cm/h, for at least one settled bed volume or until the UV curve levels out. When the wash step has been optimized approximately, it then requires 15 settled bed volumes of wash buffer to flush out particulates and unbound proteins (59).

Elution. Elution is usually performed in settled bed mode with downward flow. The flow velocity is lower than during the previous steps, usually 100 cm/h, in order to keep the elution volume as small as possible. A high flow velocity during elution will increase the elution volume. It has been shown that elution in expanded-bed mode may increase the volume by approximately 40% (58).

Cleaning. Cleaning is performed in expanded bed mode to allow any trapped particulates in the bed or column to be flushed out. The adaptor is typically placed at a position corresponding to twice the settled bed height, and the flow velocity is about 100 cm/h. Cleaning is facilitated if it is performed directly after each purification cycle. As with any protocol for cleaning, it will depend on the adsorbent and on the nature of the feedstock, etc. A general cleaning protocol suggested by the manufacturer for the different types of adsorbents (e.g., ion exchanger) is always a good starting point, as is a cleaning protocol used for a corresponding packed-bed process (if such exists). As previously mentioned, the cleaning is facilitated if as many of the particulates as possible have been washed out during the previous wash step. Back-flushes disturb the build-up of particulates, and it has been shown that back-flushes at a high flow velocity were effective in removing the build-up of sticky yeast from the distributor plate in an industrial-scale expanded-bed column (27).

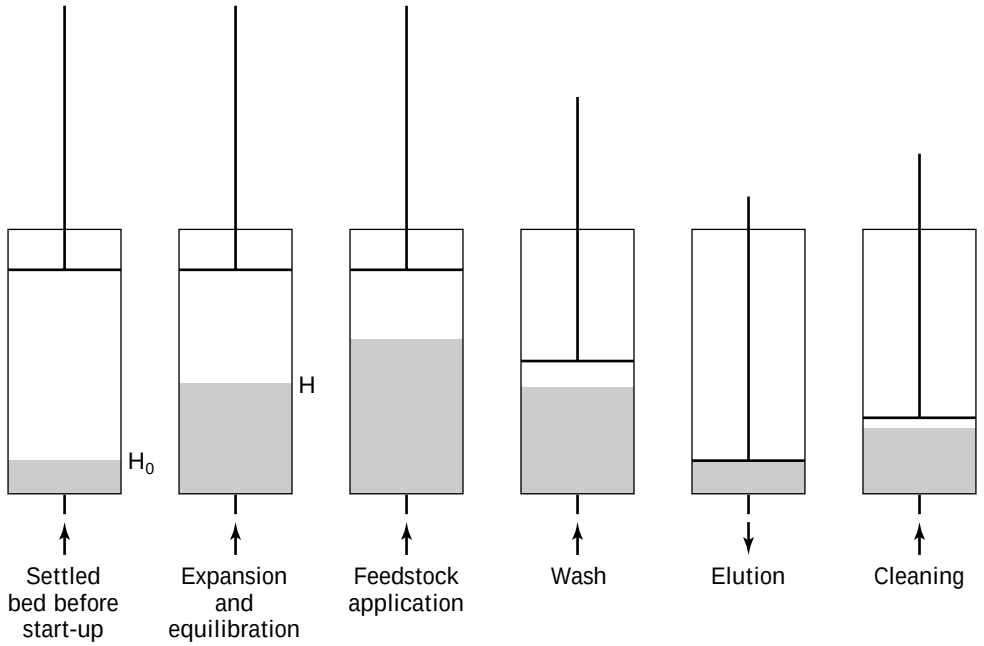


Figure 5. Operational stages in an expanded-bed adsorption process. The arrows indicate the liquid flow direction. H_0 is the settled bed height and H is the expanded bed height.

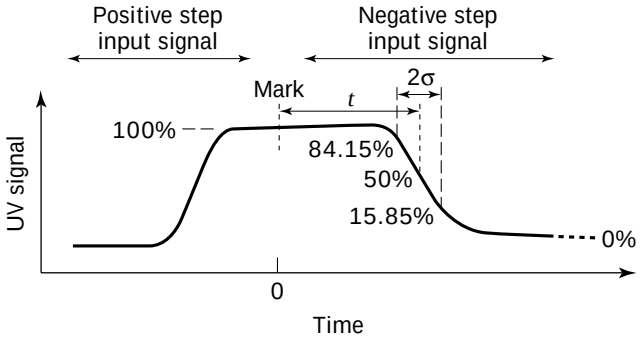


Figure 6. UV signal recording from the test procedure for determination of the number of theoretical plates (N). $N = t^2/\sigma^2$, where t is the mean residence time and σ is the standard deviation. *Source*: Courtesy of Amersham Pharmacia Biotech, Uppsala, Sweden.

After cleaning, the adsorbent can be stored in the column in an appropriate storage solution.

Feedstocks

Particulate-containing feedstocks can be applied to an expanded bed without prior clarification steps such as centrifugation and filtration; however, there are certain physical properties a feedstock must have to make it suitable for an expanded-bed process. The size and shape of the particulates must be such that they can pass through the screen or net in the column, that is, the particulates must be smaller than approximately 60 μm in diameter. Feedstocks containing small, round single cells or fine precipitates (e.g., renatured inclusion body preparations) will easily pass through the screens, but large aggregates of cells will not. If the particulates are too large, they will build up under the bottom screen and eventually block the column. When the feedstock contains particulates that are too large, they can be removed by centrifugation or by filtering through a simple mesh screen. It is of course desirable for these steps to be avoided completely, but they are relatively easy to perform and should not be compared to the complete clarification procedures needed for a packed-bed process. It is important to stir all feedstocks during application, otherwise the particulates may settle and disturb the expanded bed. Other parameters that will influence the performance of an expanded bed are the viscosity and biomass content of the feedstock. High values for these parameters will cause the bed to overexpand (relative to a buffer); if the bed expands up to the adaptor, the adsorbent particles will start to pack against the adaptor. This build-up will act as a filter and eventually cause the column to block. Furthermore, if the viscosity or the biomass content is too high, flow may become turbulent and channels may appear in the bed. If the turbulence is severe, the feedstock will pass through the channels in the bed, resulting in poor adsorption efficiency. The upper limits for feedstock viscosity and biomass content will vary from case to case, but for STREAMLINE™ with homogenates of *Escherichia coli*, there are a few guidelines (30). A dry weight of 3 to 4% gives the best result, 5% gives an acceptable result, and 7 to 8% is the upper limit. Viscosities up to 10 mPa s (measured at a shear rate of 1/s) give the best results; the upper limit is about 50 mPa s (at 1/s). In cases where the viscosity or biomass content of the feedstock does not cause the bed to expand too much, build-up on the adaptor can be prevented if the column is intermittently back-flushed. The only way to lower the biomass content of the feedstock is

to dilute it. The disadvantage is the resulting increase in feedstock volume. Dilution is only one of way of lowering viscosity. When the viscosity is caused by nucleic acids, the feedstock may be treated with a nuclease (30). When such a feedstock is a homogenate, further homogenization can be used to shear the nucleic acids and thereby lower the viscosity. The viscosity of a feedstock decreases with increased temperature, making it advantageous to apply the feedstock directly from the fermentor without prior cooling. Since temperature also has a positive effect on protein adsorption (59), expanded-bed adsorption should always be performed at room temperature (or without feedstock cooling), unless the protein of interest is, for example, sensitive to proteases.

To keep aggregation of cells and debris to a minimum, any pH adjustment of the feedstock should be performed immediately before application to the expanded-bed column. Aggregation of cells usually occurs at low pH and is therefore a problem associated with cation exchange, where pH has to be below the isoelectric point of the protein for it to bind to the adsorbent. If aggregation of cells and debris occurs within the expanded bed, when the feedstock is in contact with the adsorbent, and blocks the screen in the adaptor, the expanded bed may be operated without the adaptor screen (provided that the bed has not expanded all the way up to the adaptor) (49). As mentioned previously, cells or debris may, under certain conditions, interact with the adsorbent particles. If this occurs, it is usually with anion exchangers because they are positively charged and may attract the negatively charged cells or debris. A slight change in the binding conditions can help, for example, the addition of some salt and/or lowering the pH.

Alternatively, the feedstock load per milliliter adsorbent may have to be decreased, or it might be necessary to change to an adsorbent with another type of ligand, such as a cation exchanger.

INDUSTRIAL-SCALE APPLICATIONS

Expanded-bed adsorption is a technique with great potential for initial purification. The number of published industrial scale applications is limited because until recently columns and adsorbents suitable for use in expanded beds were not available. In addition to the three industrial-scale applications described below, an IND (Investigational New Drug) application has been filed with the Food and Drug Administration for a therapeutic recombinant protein in which expanded-bed adsorption is central to the purification process (60).

Purification of Recombinant Human Serum Albumin from Whole Yeast Culture

The number of unit operations in the process for purification of recombinant human serum albumin (rHSA) from yeast was reduced by three when expanded-bed adsorption replaced the conventional packed-bed method (32). In the conventional packed-bed method, three filtration steps were necessary to prepare the yeast culture for the first cation exchange column; when using expanded-bed adsorption, the unclarified culture could be applied directly to the expanded cation exchanger. By reducing the number of unit operations, the total processing time was deceased and the overall yield was increased. In contrast to the conventional packed-bed process, the expanded bed process could be performed in a totally closed system, and the quality (determined as degree of coloring) of the rHSA was reported to be higher using the expanded-bed process. The process was successfully scaled up (determined by comparing yield of rHSA) from a 50-mm diameter column with 300 mL adsorbent to a 1,000-mm diameter process scale column with 150 L adsorbent. Approximately 2,000 L unclarified feedstock (1,000 L culture diluted with 1,000 L distilled water) was applied to the process scale column, and the average yield of rHSA from four different runs was 87% (ranging from 82 to 91%). Using a specially made column with a 4-mm inner diameter, the lifetime of the resin was investigated. The results showed that the adsorbent could be reused up to 1,000 times without compromising performance (61).

Purification of Monoclonal Antibodies from Whole Mammalian Cell Culture

A process for recovery of recombinant monoclonal antibodies was directly scaled up from a small lab-scale expanded-bed column containing 70 mL cation exchange adsorbent to a process-scale expanded-bed column to which 12,000 L of cell culture was applied (45). In this single process step, all the cells were removed, the monoclonal antibody was concentrated fivefold, and the yield of antibody obtained was 95%.

Purification of Bovine Serum Albumin from Whole Yeast Culture

In a mimicked purification process, a 4.5% dry weight yeast culture was spiked with bovine serum albumin (2 mg/mL) and used as feedstock. The bovine serum albumin was then recovered from the feedstock using expanded-bed adsorption with an anion exchange adsorbent (27). The process was scaled up from a 25-mm diameter column to a pilot scale column with a 200-mm diameter and finally run in a fully automated system with a production-size column with a 600-mm diameter. The settled bed height was 15 cm throughout all scales, and the feedstock volume was approximately 750 mL in the 25-mm column and 430 L in the 600 mm column. The results from the different scales showed good agreement. The yield of bovine serum albumin was approximately 89% (ranging from 87 to 92%). The wash volume was approximately 11 settled bed volumes, and the volume of the elution peak was approximately 2 settled bed volumes.

Potential Industrial-Scale Applications

The successful use of expanded beds at laboratory and pilot scales has increased significantly during the past few years, giving rise to a number of potential industrial-scale applications. One such example is the purification of the protease inhibitor aprotinin (34), which has good potential as a therapeutic and diagnostic compound. Aprotinin was

expressed and secreted from the methylotrophic yeast *Hanensula polymorpha*, and the initial recovery was performed using 300 mL cation exchange adsorbent in a 50-mm diameter expanded-bed column. The only pretreatment of the yeast culture was a 1 + 1 dilution with water and an adjustment of pH. The final volume of the feedstock was 6.4 L. The dilution achieved two goals: it lowered the conductivity to favor binding to the ion exchanger, and it reduced the biomass and viscosity of the feedstock so that the bed did not expand too much. The yield of aprotinin was 76%, the purification factor was 4, and the concentration factor was 7.

Another example is the recovery of a recombinant mitotoxin fusion protein, fibroblast growth factor–saporin (rFGF-2-SAP) (31), which is a possible therapeutic agent for the treatment of diseases characterized by cellular proliferation (e.g., cancer). The fusion protein was expressed in *E. coli*. After homogenization, it was recovered from the diluted homogenate using 300 mL cation exchanger in a 50-mm diameter expanded-bed column. Homogenate from 300 to 400 g (wet weight) cells was applied to the expanded-bed column. The yield of fusion protein was 65%, and the purification factor was 20. It was also reported that a somewhat modified process was scaled up to handle 8,000 g (wet weight) cells. This process was carried out under GLP (good laboratory practice) conditions, and one such process yielded an average of 1.6 g of protein. The yield and quality of the fusion protein were comparable between the different scales.

Recombinant human nerve growth factor (rhuNGF) produced in Chinese hamster ovary cells is another example of a potential industrial scale application (38). Here, the rhuNGF was recovered using a 25-mm diameter expanded-bed column with cation exchange adsorbent using flow velocities up to 375 cm/h. The yield of rhuNGF was approximately 95% and concentrated approximately 45 to 50 times. During method development, it was found that the dynamic binding capacity increased when the feedstock was applied at 37 °C (normally at 25 °C). The optimal conditions for recovery were determined in this small scale before the process was scaled up 65-fold.

In another example, *E. coli* was used to express modified exotoxin A from *Pseudomonas aeruginosa* (29). The exotoxin lacks the enzymatic activity but has retained binding activity. It accumulates in the periplasm of *E. coli*. After release of the exotoxin by osmotic shock, it was recovered using expanded-bed adsorption with an anion exchange adsorbent. To reduce the viscosity of the extract, it was treated with a DNase prior to application to a pilot-scale expanded-bed column with a 200-mm diameter. Cell extract from 4.5 kg of *E. coli* was applied to 4.7 L adsorbent at 400 cm/h. Recovery by expanded-bed adsorption gave an exotoxin with a concentration three times greater than that obtained using the conventional packed-bed process and with a slightly higher yield. Furthermore, the entire process was completed in one third of the time.

Humanized immunoglobulin G_4 was affinity purified from myeloma cell culture using expanded-bed adsorption at pilot scale (41). A 200-mm diameter expanded-bed column containing 4.7 L protein A adsorbent was used. The cell reactor was connected directly to the expanded bed, and approximately 100 L of cell culture was applied to the column at 37 °C. The purification process (application, wash, and elution) took 2.5 h and 40 g of the antibody was quantitatively recovered. The purity of the antibody was comparable to that obtained by the corresponding packed-bed process.

Other potential large-scale processes include the purification of proteins from milk or whey (46–48).

BIBLIOGRAPHY

1. S.-M. Lee, *J. Biotechnol.* **11**, 103–118 (1989).
2. W. Berthold and R. Kempken, *Cytotechnology* **15**, 229–242 (1994).
3. R. Kempken, A. Preißmann, and W. Berthold, *J. Indust. Microbiol.* **14**, 52–57 (1995).
4. H. Tebbe, D. Lütkemeyer, F. Gudermann, R. Heidemann, and J. Lehmann, *Cytotechnology* **22**, 119–127 (1996).
5. F.J. Bailey, R.T. Warf, and R.Z. Maigetter, *Enzyme Microb. Technol.* **12**, 647–652 (1990).
6. A.G. Fane, and J.M. Radovich, in J.A. Asenjo ed. *Separation Processes in Biotechnology*, Marcel Dekker, New York, 1990, pp. 209–262.
7. K.H. Kroner, H. Schütte, H. Hustedt, and M.-R. Kula, *Process Biochem.* **19**, 67–74 (1984).
8. H.G.J. Brummelhuis, in J. Curling ed. *Methods in Plasma Protein Fractionation*, Academic Press, London, 1980, pp. 117–128.
9. M. Better, R. Richieri, A. Tran, and W.C. McGregor, *Abstract from Recovery of Biological Products VIII*, Tucson, Ariz., Oct. 1996, p. 11.
10. M.J. Slater, *The Principles of Ion Exchange Technology*, Butterworth-Heinemann, Oxford, 1991, pp. 41–49.
11. E. Pungor Jr., N.B. Afeyan, N.F. Gordon, and C.L. Cooney, *Biotechnology* **5**, 604–608 (1987).
12. C.R. Bartels, G. Kleiman, J.N. Korzun, and D.B. Irish, *Chem. Eng. Prog.* **54**, 49–51 (1958).
13. F.P. Gailliot, C. Gleason, J.J. Wilson, and J. Zwarick, *Biotechnol. Prog.* **6**, 370–375, (1990).
14. A. Buijs and J.A. Wesselingh, *J. Chromatogr.* **201**, 319–327 (1980).
15. M.A. Burns and D.J. Graves, *Biotechnol. Prog.* **1**, 95–103 (1985).
16. N.M. Draeger and H.A. Chase, *Inst. Chem. Eng. Symposium Series* **118**, 161–172 (1990).
17. H.A. Chase and N.M. Draeger, *J. Chromatogr.* **597**, 129–145 (1992).
18. N.M. Draeger and H.A. Chase, *Bioseparation* **2**, 67–80 (1991).
19. G. Dasari, I. Prince, and M.T.W. Hearn, *J. Chromatogr.* **631**, 115–124 (1993).
20. C. Spence, C.A. Schaffer, S. Kessler, and P. Bailon, *Biomed. Chromatogr.* **8**, 236–241 (1994).
21. J. Thömmes, M. Weiher, A. Karau, and M.-R. Kula, *Biotechnol. Bioeng.* **48**, 367–374 (1995).
22. G.E. McCreath, H.A. Chase, R.O. Owen, and C.R. Lowe, *Biotechnol. Bioeng.* **48**, 341–354 (1995).
23. C.M. Wells, A. Lyddiatt, and K. Patel, in M.S. Verrall and M.J. Hudson eds., *Separations for Biotechnology*, Ellis-Horwood, Chichester, UK, 1987, pp. 217–224.

24. G.R. Gilchrist, M.T. Burns, and A. Lyddiatt, in D.L. Pyle ed., *Separations for Biotechnology 3*, The Royal Society of Chemistry, London, 1994, pp. 184–190.
25. P.H. Morton and A. Lyddiatt, in M.J. Slater ed., *Ion Exchange Advances*, Elsevier, London, 1992, pp. 237–244.
26. K.K. Unger, *Journal of Chromatography*, vol. 16, Elsevier, Amsterdam, 1979, pp. 12–14.
27. A.-K. Barnfield Frej, H.J. Johansson, S. Johansson, and P. Leijon, *Bioprocess Eng.* **16**, 57–63 (1997).
28. M. Hansson, S. Ståhl, R. Hjorth, M. Uhlén, and T. Moks, *Biotechnology* **12**, 285–288 (1994).
29. H.J. Johansson, C. Jägersten, and J. Shiloach, *J. Biotechnol.* **48**, 9–14 (1996).
30. A.-K. Barnfield Frej, R. Hjorth, and Å. Hammarström, *Biotechnol. Bioeng.* **44**, 922–929 (1994).
31. J.R. McDonald, M. Ong, C. Shen, Z. Parandoosh, B. Sosnowski, S. Bussell, and L.L. Houston, *Protein Express. Pur.* **8**, 97–108 (1996).
32. European Patent Application EP 0 699 687 A2 (Published 1996), M. Noda, A. Sumi, T. Ohmura, and K. Yokoyama (to The Green Cross Corporation, Osaka, Japan).
33. A. Pessoa Jr., R. Hartmann, M. Vitolo, and H. Hustedt, *J. Biotechnol.* **51**, 89–95 (1996).
34. C. Zurek, E. Kubis, P. Keup, D. Hörlein, J. Beunink, J. Thömmes, M.-R. Kula, C.P. Hollenberg, and G. Gellissen, *Process Biochem.* **31**, 679–689 (1996).
35. Y.K. Chang, G.E. McCreath, and H.A. Chase, *Biotechnol. Bioeng.* **48**, 355–366 (1995).
36. Y.K. Chang and H.A. Chase, *Biotechnol. Bioeng.* **49**, 204–216 (1996).
37. B.C. Batt, V.M. Yabannavar, and V. Singh, *Bioseparation* **5**, 41–52 (1995).
38. J.T. Beck, A. Liten, S. Viswanathan, A. Lindgren, J.C. Emery, and S.E. Builder, *Abstract from Recovery of Biological Products VIII*, Tucson, Ari., Oct. 1996, p. 74.
39. C. Born, J. Thömmes, M. Biselli, C. Wandrey, and M.-R. Kula, *Bioprocess Eng.* **15**, 21–29 (1996).
40. J.C. Erickson, J.D. Finch, and D.C. Greene, in B. Griffiths, R.E. Spier, and W. Berthold eds., *Animal Cell Technology: Products for Today, Perspectives for Tomorrow*, Butterworth-Heinemann, Oxford, 1994, pp. 557–560.
41. C. Jägersten, S. Johansson, R. Pardon, and J. Bonnerjea, *Abstract from Recovery of Biological Products VIII*, Tucson, Ariz., Oct. 1996, p. 46.
42. D. Lütkemeyer, N. Ameskamp, H. Tebbe, J. Wittler, and J. Lehmann, *Abstract from Recovery of Biological Products VIII*, Tucson, Ariz., Oct. 1996, p. 78.
43. J. Thömmes, M. Halfar, S. Lenz, and M.-R. Kula, *Biotechnol. Bioeng.* **45**, 205–211 (1995).
44. J. Thömmes, A. Bader, M. Halfar, A. Karau, and M.-R. Kula, *J. Chromatogr., A* **752**, 111–122 (1996).
45. G. Zapata, A. Lindgren, A.-K. Barnfield Frej, P. Leijon, A.D. Liten, T.L. Mayes, G. Blank, D. Narindray, B. Wagner, W. Galan, J. Beck, M. Press, and S. Builder, *Abstract from Recovery of Biological Products VIII*, Tucson, Ariz., Oct. 1996, p. 16.
46. A. Degener, M. Belew, and W.H. Velander, *Abstract from Recovery of Biological Products VIII*, Tucson, Ariz., Oct. 1996, p. 15.
47. W. Noppe, I. Hanssens, and M. De Cuyper, *J. Chromatogr., A.* **719**, 327–331 (1996).
48. W. Noppe, I. Hanssens, F. Geeraerts, and M. De Cuyper, *Abstract from First International Conference on Expanded Bed Adsorption*, Cambridge, U.K., Dec. 1996, p. 15.
49. A.-K. Barnfield Frej, *Bioseparation* **6**, 265–271 (1996).
50. M. Asif, N. Kalogerakis, and L.A. Behie, *AIChE J.* **37**, 1825–1832 (1991).
51. M.R. Al-Dibouni and J. Garside, *Trans. Inst. Chem. Eng.* **57**, 94–103 (1979).
52. J.F. Richardson and W.N. Zaki, *Trans. Inst. Chem. Eng.* **32**, 35–53 (1954).
53. E. Zafirakos and A. Lihme, *Abstract from Second International Conference on Expanded Bed Adsorption*, Napa Valley, Calif., June 21–23, 1998, p. 15.
54. K. Beyzavi and V. Sarantschin, *Abstract from Second International Conference on Expanded Bed Adsorption*, Napa Valley, Calif., June 21–23, 1998, p. 42.
55. A. Karau, C. Benken, J. Thömmes, and M.-R. Kula, *Biotechnol. Bioeng.* **55**, 54–64 (1997).
56. *Bioseparation and Bioprocessing, Vol. 1 Biochromatography: Expanded Bed Adsorption Chromatography*, Wiley-VCH Verlagsgesellschaft, Weinheim, Germany, 1998.
57. O. Levenspiel, *Chemical Reaction Engineering*, 2nd ed., Wiley, New York, 1972, pp. 253–295.
58. Y.K. Chang and H.A. Chase, *Biotechnol. Bioeng.* **49**, 512–526 (1996).
59. R. Hjorth, S. Kämpe, and M. Carlsson, *Bioseparation* **5**, 217–223 (1995).
60. J.A. Purvis, A. Binieda, C.A. Lewis, M. Pearce-Higgins, and P.G. Varley, *Abstract from First International Conference on Expanded Bed Adsorption*, Cambridge, U.K., Dec. 1996, p. 27.
61. A. Sumi, K. Okuyama, K. Kobayashi, W. Ohtani, T. Ohmura, and K. Yokoyama, *Abstract from Second International Conference on Expanded Bed Adsorption*, Napa Valley, Calif., June 21–23, 1998, p. 12.

ADDITIONAL READING

A. Mullick, C.M. Griffith, and M.C. Flickinger, *Biotechnol. Bioeng.* **60**, 333–340 (1998).
M.C. Flickinger, M.J. Robichaud, J.E. Morris, C.M. Griffith, M.J. Annen, P.W. Carr, and C. Dunlap, *U.S. Patent 5,837,826*, November 17, 1998.

See also ADSORBENTS, INORGANIC; ADSORPTION, PROTEIN, BATCH; ADSORPTION, PROTEINS WITH SYNTHETIC MATERIALS.

ADSORPTION, PROTEIN, BATCH

WILLIAM R. ADAMS
Merck Manufacturing Division
West Point, Pennsylvania
VICTOR GOETZ
ANN L. LEE
Merck Research Labs
West Point, Pennsylvania

KEY WORDS

Adsorption
Batch

Modeling
On-off chromatography
Protein
Separation

OUTLINE

INTRODUCTION

Batch adsorption technology has been successfully applied in a number of important protein separation processes. The principles of batch adsorption are used in a wide range of diverse applications, including several commercial-scale processes for the purification of protein products, a large number of biochemical assays (e.g., enzyme-linked immunosorbent assays), and design of biosensors and biomedical devices, where the aim could be to promote or retard protein adsorption. The focus of this article is on methods for recovery and purification of proteins, although the principles and considerations can be readily applied to the full range of applications.

Batch adsorption is considered here as a separation technique in which an entire protein-containing solution is contacted simultaneously with a fixed quantity of adsorbent. This is in contrast to adsorption in a fixed-bed column mode, where the feed stream flows through the column and is differentially contacted with the adsorbent. After adsorption, the wash and desorption procedures can be performed in a number of different ways. For example, elution may be performed in the original mixing vessel by changing the surrounding solution, in a different vessel after adsorbent recovery through an additional step such as centrifugation, or after transfer of the adsorbent into a column for ease of flowthrough operation. For desorption to be considered a batch operation, however, the elution must be performed in a manner consistent with single-stage solid–liquid contact. This implies that the eluting solution will have uniform product concentration, which is in contrast to the differential elution characteristic of chromatography.

Protein batch adsorption is often regarded as an outdated unit operation now superseded by modern column chromatography techniques incorporating a new generation of chromatographic supports. Before the introduction of rigid resins and high-performance liquid chromatography techniques, however, stirred tank batch contactors were the preferred configuration for scale-up of adsorptive separation steps, because the highly compressible resins could not even support their own weight when packed into process-scale chromatography columns. Nevertheless, an appropriate train of such stirred tanks could provide a moderate number of theoretical separation stages.

Although it is true that modern chromatographic techniques in many applications offer efficiencies and purification factors far superior to the stirred tank systems, there remain a number of applications where creative implementations of batch adsorption offer decided advantages. Because of the immediate and uniform contact of solution and adsorbent, batch adsorption can be performed faster than column-mode adsorption. Thus, it could offer an advantage in applications where a large volume of feed solution must be processed quickly to achieve efficient cycle times or where the limited stability or proteolytic susceptibility of protein products makes rapid processing essential. Batch adsorption may also be favored for handling crude feedstreams having high viscosity or debris that cannot be processed readily through column operations.

This review summarizes the physicochemical principles that define and control batch adsorption, describes a variety of classical and novel applications of these principles, and attempts to demonstrate that a widely overlooked unit operation still has significant utility in bioprocessing.

PRINCIPLES GOVERNING PROTEIN BATCH ADSORPTION

The overall effectiveness of batch adsorption processes will depend on both the equilibrium adsorbent capacities as well as the mass transfer and kinetic limitations. This section reviews the fundamental principles of protein adsorption phenomena, through which the suitability of batch adsorption as a separation or recovery process may be more readily understood or predicted.

Models for Protein Adsorption: Capacity and Competition

Single Component Adsorption. The specific mechanisms by which large macromolecular proteins contact and adsorb to a solid surface are varied and generally highly complex. Depending on the choice of protein, adsorbent, and solution properties, the surface attachment mechanism may be driven by several different general types of interactions, such as ion exchange, affinity (protein–ligand reaction), less-specific hydrophilic, hydrophobic, van der Waals, or hydrogen bonding. In principle, the thermodynamic treatment of protein adsorption from solution could be directly analogous to that for adsorption of small nonpolar molecules from the vapor phase for which there are several well-developed surface equations of state. For such

simple molecule adsorption, the adsorbent surface geometric and energetic irregularities along with specific sorbate–surface interactions are generally the chief complexities (1). For protein adsorption in practice, however, thermodynamic treatment of equilibria is severely limited by the additional complexities arising from specific solvent–protein interactions, protein conformational variations both in solution and on the adsorbent surface, and frequent multipoint attachment configurations.

As a consequence of the somewhat intractable experimental and mathematical complexities, most models used for protein adsorption equilibria are simple empirical or semitheoretical models that are useful according to their goodness-of-fit to the experimental data. The semitheoretically based Langmuir adsorption isotherm given by

$$q = q_m C/(K_D + C) \quad (1)$$

where q is quantity adsorbed per unit of adsorbent, C is concentration of protein in solution, q_m is maximum quantity adsorbed at high C, and K_D is the disassociation or binding constant has been frequently fitted to protein adsorption data with adequate correlation. Although the physical assumptions underlying the development of this model are not followed by protein adsorption via ion exchange (2), the model nonetheless has provided a good fit if salt concentration-dependent parameters are used. For example, Weaver and Carta (3) fitted data for lysozyme on POROS 50 HS (a macroporous resin based on styrene-divinyl benzene copolymer) and on S-Hyper D-M (polystyrene-coated silica filled with functionalized polyacrylamide hydrogel), both of which possess strong cationic functionality. In their studies, the fitted values for q_m ranged from 109 to 262 mg/cm^3, and disassociation constants ranged from 5.3×10^{-3} to 2.6×10^{-1} mg/cm^3. The isotherms displayed classical protein adsorption characteristics; the nearly rectangular-shaped isotherms indicative of highly favorable adsorption at low salt concentration moderated to an almost linear-shaped isotherm indicative of weak adsorption at high salt concentration. Numerous other ion exchange protein adsorption studies have been performed with similar relative measures of adherence to the Langmuirian behavior.

Affinity adsorption may be highly selective, depending on the specific ligand–protein interaction, which may often be driven by a combination of electrostatic, hydrophobic, or hydrogen bonding type forces. Some affinity adsorption data has been successfully correlated with simple equilibrium models, based on the assumed simple reaction

$$[P] + [L] \Leftrightarrow [P - L] \quad (2)$$

where $[P]$ and $[L]$ denote concentration of protein and ligand, respectively, and from which an equilibrium constant may be defined such as

$$K_{eq} = [P][L]/[P - L_{complex}] \quad (3)$$

Generally, K_{eq} is small, often ranging between 10^{-4} M to 10^{-10} M, implying that binding is nearly irreversible or that the rate constants are related by $k_{adsorp} \gg k_{desorp}$. Adsorption data for many affinity systems can also be approximated through the fitted Langmuir model, although the range of fitted constants vary widely, depending on substrate–enzyme, antigen–antibody, carrier protein–hormone, or base sequence nucleic acid interactions. As an example, Chase describes the fit of data to equation 1 above for β-galactosidase adsorbing onto monoclonal antibodies immobilized on agarose gels, where it was found that K_D and q_m typically had values of 1.5×10^{-8} M and 6×10^{-9} gmol/mL, respectively (4).

Protein adsorption onto dye–ligand modified supports generally is controlled by a combination of electrostatic and hydrophobic protein–dye interactions. The Langmuir equation provides a reasonable fit to the data for many systems (5), although some data are better correlated through other models, such as the Freundlich isotherm. Typically, q_m for such systems is between those observed for high capacity ion exchange and the lower capacity affinity adsorbents. There are many other protein–adsorbent systems in which adsorption is driven predominantly by nonspecific mechanisms of interaction that could be used for either batch adsorption or chromatographic separations. Some of these systems are adequately fitted by simple models, whereas others, such as those displaying sigmoidal isotherms, clearly follow more elaborate adsorption mechanisms. A listing of representative adsorption capacities for several typical types of protein–adsorbent systems is shown in Table 1.

As indicated in Table 1, the choice of adsorbent and solution conditions provide a wide range of possible adsorbent capacities and relative strengths of adsorption. The suitability of any particular adsorbent–solution system to accomplish an effective separation process, however, will depend significantly on the adsorption capacity for the protein of interest relative to other competing proteins as well as on the relative rates of protein component adsorption and desorption.

Multicomponent Adsorption. Application of adsorption models to protein mixtures has been much more limited because of several factors:

1. Limitations of applicability of the empirical or semitheoretical models when extended to multiple adsorbing protein components
2. Limited adsorption database available for even model protein mixtures
3. The fact that industrially important separations, whether by batch adsorption or chromatographic adsorption processes, typically involve mixtures containing numerous protein components (many of which may be poorly characterized) or additional impurities that substantially confound the analysis

As an example model binary system, adsorption of mixtures of bovine serum albumin (BSA) and lysozyme on the strong cation exchanger S Sepharose FF has been studied (9). For this system, the single component isotherms were well correlated by the Langmuir equation. For the mixture data, the authors evaluated both competitive and

Table 1. Representative Capacities and Binding Constants for Protein Adsorption

Protein	Adsorbent	Mode	q_m (mg/mL)	K_D (mg/mL)	Ref. no.
BSA	DEAE Sepharose CL-6B	Ion exchange	82.1	6.0×10^{-2}	6
BSA	Q Sepharose (CL-4B) FF	Ion exchange	41.2	7.0×10^{-2}	6
α-Chymotrypsinogen	SynChroprep Propyl HIC	Hydrophobic	70.7	3.7×10^{-2}	7
β-galactosidase	Monoclonal antibody on agararose	Affinity (antibody)	3.2×10^{-3}	8.1×10^{-6}	4
IgG (human)	Avid AL gel	Affinity	22.7	5.8×10^{-3}	8

Note: BSA, bovine serum albumin; IgG, immunoglobulin G.

noncompetitive multicomponent extensions of the Langmuir equation. Neither model gave an accurate correlation to the data, although from this comparison and from additional chromatographic data it was concluded that competitive adsorption was apparent but that additional factors such as interprotein interactions were likely important as well. More recently, Lassen and Malmsten (10) reported on a study of multicomponent adsorption from ternary mixtures of human serum albumin (HSA), human immunoglobulin G (IgG) and human fibrinogen (Fgn) in 0.01 M phosphate buffer on three different polymer surfaces. The three surfaces were prepared to yield a spectrum of overall surface properties covering hydrophobic and hydrophilic with either positive or negative overall charge, respectively. For all surfaces, the single component isotherms were approximately Langmuirian in shape, with the relative maximum quantities adsorbed falling in the pattern Fgn > IgG > HSA. Adsorption from the multicomponent mixtures was notable in that (*1*) the total quantity adsorbed relative to the single component isotherms was reduced, and (*2*) the emergent dominant protein varied depending on the specific surface evaluated. Overall, the observed ternary behavior could not be readily predicted from the single component data.

Although not strictly of concern with respect to equilibrium loadings, the so-called Vroman effect, through which some protein mixtures exhibit sequential adsorption with smaller molecular weight proteins generally adsorbing early only to be later displaced by larger molecular weight proteins, has been noted for several protein–adsorbent systems (11). This rate-based competitive effect could occur within the time scale of practical application in separations, and thereby provide a contributing factor to adsorptive process performance.

The complex adsorption behavior observed even in relatively simple binary or ternary mixtures amply highlights the limitations of adsorption models and weakness of extensions thereof to multicomponent mixtures. It is therefore clear that the development of batch adsorption separation processes must be firmly based on experimental data.

Implications of Adsorption Performance to Separations. Equilibrium adsorption capacities (or q vs. C isotherms) define the first criteria for adsorption process feasibility. If capacity appears sufficient and if competitive adsorption is not a major problem, then the next concern would be fractional recovery. As discussed by Scopes (12), for practical applications that use small or modest volumes of adsorbent relative to solution (e.g., a ratio of volume adsorbent V_a to volume solution V_s of ~0.1), the separation factor α, defined by $\alpha = q/(C + q)$, should be greater than approximately 0.98. This in turn implies that K_{eq} values less than about 10^{-6} M are needed to achieve high fractional recoveries (e.g., a recovery of >90% in one adsorptive stage).

Assuming loading capacity and fractional recoveries from solution would be practical, the next major challenge in applying batch adsorption is to identify appropriate combinations of adsorbent and solution properties to provide a useful selectivity. Selectivity among a mixture of n components may be defined by

$$S = [(q_1/C_1)/(q_i/C_i)] \tag{4}$$

where subscript *1* denotes the desired product and i denotes some unwanted component, with i ranging from 2 to n. The required value for S depends of course on the objective of the separation. To achieve a 10-fold purification from an equimolar solution, for example, would require that S be greater than 10. High selectivity may be achievable through a particular combination of adsorption and desorption conditions; however, it is uncommon for most general adsorbents to provide high selectivity between many proteins. Because of the broad range of proteins generally present in bioseparation feedstreams (such as clarified broths or cell lysates originating from fermentation operations), the adsorbent candidates more likely to enable a successful application would employ specific interactions. Affinity adsorbents, for example, may be able to achieve such selectivity, and immunoadsorbents should be especially capable in light of the highly specific antigen–antibody complexation. If the selectivities of interest are low, highly effective separations could be more readily achieved through non–batch operations such as column chromatography, which can exploit even modest selectivity differences in equilibrium loadings or in relative rates of adsorption.

Dynamics of Protein Adsorption

Kinetic and Mass Transfer Considerations. Separations of practical industrial application will generally require substantial adsorbent capacity. The large surface areas consequently required will usually be provided through extensive macroporous structures. A general schematic of the paths available for a protein molecule in the bulk liquid phase to adsorb to the surface of such an adsorbent is indicated in Figure 1. The overall process will likely involve several diffusional steps as well as multiple interactions with the adsorbent surface moities.

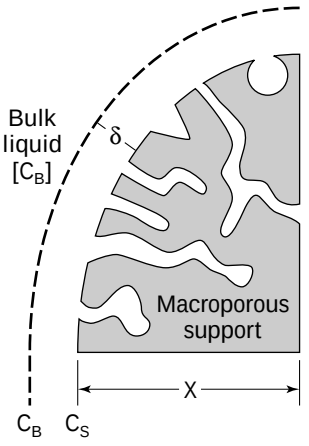


Figure 1. Schematic representation of available paths for protein adsorption on macroporous adsorbents.

The bulk phase in batch adsorption applications or experiments will normally be well mixed through appropriate design of mixing conditions in the contacting vessel. As such, the concentration of protein throughout the bulk phase will be uniform, as denoted by C_B. Surrounding each adsorbent particle will be a stagnant fluid layer, the thickness of which (δ) will depend on the hydrodynamic conditions. Diffusion through this film is normally approximated using a linear driving force in concentration

$$J = -D \times (C_B - C_S)/\delta \tag{5}$$

where J is the protein mass flux, D is the free solution diffusivity, and C_S is the concentration at the external surface of the adsorbent. Diffusion within the macroporous structure is generally approximated by an equation of form

$$J = -\epsilon_p \times D_e \times dC/dx \tag{6}$$

where ϵ_p is the adsorbent void fraction and D_e is the effective diffusivity, related to the free solution diffusivity by the approximation

$$D_e = D/\tau \tag{7}$$

where τ is the tortuosity factor. There are a number of means to measure or estimate values for D and δ (13,14) as well as τ (15). Although the approximation of pore diffusional mass transfer with such simplistic equations allows for practical modeling and data correlation, the mechanistic inaccuracies are quite apparent. For example, because of the irregular geometry of the pores, there will be a large distribution of actual mass transfer trajectories (or random walks). Furthermore, because the proteins may be of comparable size to some pore passageways in certain adsorbents, there may be frequent opportunities for pore blocking, diffusional hindrances, or substantial variation in the number of contacts with the surface.

The kinetics of interaction between the protein in solution and the adsorbent surface will depend on many factors. Because the transitions in three-dimensional conformational states of the protein during the adsorption process are not well known, simplistic mechanisms are usually postulated in order to construct a mathematical representation. Normally, the adsorption kinetics are modeled using first- or second-order reaction rate equations to describe rates of transition between a few different states of protein or protein–adsorbent complexes. The kinetic constants thereby serve as adjustable model parameters. The rationale for the use and the accuracy of such representations is discussed in the next section.

Kinetic Studies: Ideal Surfaces and Industrial Adsorbents. Recent progress has been made in the study of protein–adsorption kinetics using well-characterized surfaces. In contrast to macroporous industrial adsorbents, in which diffusional limitations will usually be significant, planar surfaces within small experimental reactor volumes often will not be diffusion limited and hence can allow intrinsic adsorption rate measurement. Adsorption onto planar optical waveguide surfaces as monitored through reflectance techniques can be measured to within approximately ± 10 molecules per μm^{-2} at high sample frequencies (14). Kurrat et al. (16), for example, evaluated the kinetics of HSA and BSA adsorption onto a hydroxylated silica-titania surface. Approach to half saturation was achieved within 5 to 10 min, although adsorption toward equilibrium loadings continued out beyond 50 min. In developing and fitting kinetic equations, based on an elementary mechanistic model, they found that the model required at least two states of protein (distinguished by reversibly and irreversibly bound, respectively) to provide an acceptable fit to the data. Within the context of their equations, the adsorption kinetic constants k_a were on the order of 10^{-6} cm/s.

Wahlgren and Eloffson (17) studied adsorption kinetics of β-lactoglobulins A and B to hydrophobic methylated silica waveguide surfaces. Adsorption of these proteins, which show complicating self-associating behavior in solution, proceeded to completion over the time course of approximately 50 min. In applying their kinetic model, which allowed for both a first-order surface conformational change as well as an exchange reaction between adsorbed and solubilized protein, a reasonable fit to the data was obtained. It appears that an increase in model accuracy, however, would have to come through an allowance of additional nonidealities, such as from the range of protein orientations and conformations, surface heterogeneity, and probability factors for interactions, steric hindrances, etc. Along these lines, Jin et al. (18) attempted to account for steric hindrance effects by applying random sequential adsorption (RSA) principles to the case where adsorption may proceed reversibly. The resulting formalism provided an explanation for why the apparent adsorption kinetic constants can show a functional dependence on surface coverage.

The state-of-the-art adsorption rate modeling exemplified in the works cited earlier are useful for correlation of data required for practical applications and for aiding the development of theoretical understanding of viable biophysical mechanisms. However, the data analysis and modeling efforts also highlight a fundamental limitation:

the inability to experimentally distinguish or measure protein conformational variations and the distribution thereof on adsorbent surfaces. Until further progress in technique is achieved, modeling and understanding will remain limited to semitheoretical models based on an indirectly assessed distribution of proteins between only a few postulated states of conformation and surface binding.

The kinetics of protein adsorption to industrial macroporous adsorbents has been evaluated for many systems, of course, since the definition of adsorption rates is a prerequisite to application of any batch adsorption process. Basic rate data have generally been obtained through classical methods using either batch- or column-based experiments. As mentioned above, the adsorption dynamics will often have some significant diffusional resistances, and hence the kinetic data will reflect combined mass transfer and intrinsic adsorption reaction rates.

Modeling of Adsorption Process Dynamics. The ultimate success of potential applications of batch adsorption separation depends on the adroit manipulation of the thermodynamic and kinetic properties of the adsorbent-solute system to achieve high process efficiency. This aim can best be realized through development and application of a mathematical model for the process dynamics, which necessarily would be based on the underpinning factors governing intrinsic adsorption kinetics, equilibrium phase distribution, and mass transfer effects. The model used to simulate the process dynamics would facilitate the following:

1. *Feasibility analysis*. To assess the performance, whether for batch or column mode, for approximate performance in terms of yield and purification achievable
2. *Process design*. To enable detailed design, including scale-up effects, as well as to perform operational analyses to ascertain cycle times, economics of operation, etc.
3. *Optimization*. To readily calculate effects resulting from changes to the myriad of variables, such as adsorbent properties, relative volumes, initial concentration, contact times, regeneration procedures, etc., all of which could have substantial impact on process viability

Modeling of adsorption process dynamics is straightforward in principle but requires numerous simplifications at the detail level in order to retain mathematical tractability as well as biophysical reality in the adjustable model parameters. The general approach follows directly from the differential equations for mass balance, mass transfer, and kinetics for adsorption and any additional reactions. The process model, which is the resulting system of differential equations, will vary in complexity depending on the simplifying assumptions used to approximate the collective molecular behavior, which is usually highly complex because of the distribution of interaction rates and pathways occurring throughout an irregular three-dimensional geometry.

A number of models for protein adsorption within a finite batch have been published. As an example, Mao et al. (19) have presented a model for batch protein adsorption applicable specifically for nonporous adsorbents along with a simplified extension for porous adsorbents. By combining variants of equations 1 and 2, along with the differential form of the overall mass balance and an assumed second-order adsorption kinetic equation, the resulting equations could then be integrated to provide C and q as functions of time. The parameters in their model were first fitted to adsorption data for single component proteins HSA, lysozyme, and ferritin, as adsorbed onto either dye-affinity or ion-exchange adsorbents. In assessing the model fits, it was observed that the sensitivity of the models with respect to the kinetic constants may not be significant enough relative to the accuracy of the adsorption data to accurately ascertain true values of parameters. This conclusion may generally be extended to more elaborate models as well, where the accuracy of the fitted model may be improved by including additional kinetic parameters to account for additional reactions (states of conformation, etc.), but where the resulting understanding and the accuracy of extrapolation using the models is not improved. Mao et al. further used the models to evaluate relationships between various operational parameters and system performance. For example, calculations showed that the initial concentration of HSA could significantly affect both the final concentrations as well as the time required to approach equilibrium (e.g., time to 90% of final C/C_o), where a C_o varied from 0.8 to 0.05 mg/mL increased the required adsorption time from about 10 to more than 30 minutes. Of particular note was the extreme sensitivity of both the rate of adsorption and the equilibrium loadings to the ratio of adsorbent volume to liquid volume, $R_v(V_a/V_s)$. This would clearly be an important parameter to routinely evaluate in process design and operational optimization.

These sample calculations highlight the real utility of such models, which is that once proven capable of accurately representing the mass transfer and kinetic reaction rates of adsorption, they can be readily applied to adsorptive system design and optimization for either batch or column modes of operation.

Implications and Applications to Batch Separations. Knowledge of the mass transfer and kinetic limitation principles provides several general guidelines. In terms of process timing, for many systems it can be appreciated that equilibrium loadings may take hours to attain, but practical process-type loadings may be achieved on the order of 10 minutes. Analogously, adsorbent regeneration will often require hours, depending on the conditions employed to induce complete desorption and the requirements for regenerated adsorbent site occupancy. Although some approximate predictions of capacity and adsorption rates could be made by considering the basic features of the adsorbent (particle diameter, average pore size, primary mode of interaction with the protein, etc.) and of the protein in solution (molecular weight, size and shape, average surface potential, etc.), it is clear that acquisition of adsorption rate data for the specific system under consideration will always be an important prerequisite to suc-

cessful process applications. Along these lines, relative to column-based experiments, batch experiments will generally be easier to perform and will provide more accurate information on the fundamentals of equilibrium capacity and limitations for mass transfer and kinetics of adsorption for particular protein–adsorbent systems.

The rationale for successful batch adsorption applications reviewed later in this article can be readily interpreted from the performance characteristics of relative adsorption rates and capacities. For example, it can be seen that if fast adsorption is required because of product lability, then a batch mode may be preferred to the column mode of operation. If speed and overall cycle time is a main driver, it has been shown (20) that a batch system may be more efficient than a column-based system under some conditions. Batch-contacting systems may also be suggested if there are other hydrodynamic factors, such as high solids/debris content or high viscosity which would negate other column operational advantages. Batch operations will also allow greater flexibility in choice of adsorbent, since even highly compressible gels can readily be contacted with the solution. On the other hand, if a high purification factor is desired from the adsorptive separation process, then a column mode of differential protein-adsorbent contact will generally be required other than for highly specific affinity adsorbents.

BATCH ADSORPTION AS A PRIMARY SEPARATION STEP

Except for highly specific affinity methods, batch adsorption techniques are not generally applicable to high-resolution separations. Instead, batch adsorption is ideal for so-called capture applications, where entire classes of impurities are to be eliminated in advance of high-resolution purification techniques, such as differential elution chromatography. Such bulk adsorptions also serve to reduce large volumes of dilute cell culture fluids or microbial lysates to more manageable volumes for further downstream processing.

Whereas solvent extraction is a popular first step in recovery of antibiotics from fermentation broths, adsorption is generally more gentle to proteins, which can be adsorbed without the denaturation often produced by organic extraction (21). Ion exchange adsorbents are preferred for highly polar molecules that, even in neutral form, exhibit low distribution coefficients for extraction.

Selectivity in batch ion exchange adsorption can be enhanced by judicious choice of adsorption buffer conditions (e.g., moderate salt concentration) to exploit differences in properties between the desired product and contaminants. This may be especially true in the case of recombinant proteins, which may be fundamentally different from the host organism's natural proteins. Adsorption equilibria that rely on weak physical interactions such as hydrophobic interaction, hydrogen bonding, and ion exchange are desirable for separations because these low-energy interactions can be readily reversed. Additional selectivity can be introduced during the desorption step so that adsorption processes are not without resolving power. Nonetheless, high selectivity is not the primary requirement for such an early process step.

With the introduction of the more specific nature of affinity interactions, well-thought-out batch adsorptions could accomplish clarification, concentration, and at least partial purification in one integrated step. Creative new configurations such as expanded-bed adsorption to approach this one-step purification ideal are discussed in more detail later ("Expanded Bed Chromatography").

CLASSICAL BATCH ADSORPTION PROCESSES

Antibiotic Purification

Much of the technology for batch adsorption of proteins and other biomolecules was originally developed for antibiotic purification from microbial fermentation sources. Both ion-exchange and hydrophobic adsorbents have been used for antibiotic purifications. Like other packed-bed applications, column adsorption is extremely susceptible to plugging by suspended insoluble materials in a feed-stream. Therefore, either efficient precolumn filtration or stirred tank adsorber configurations predominated.

In the late 1950s, Bartels et al. (22) developed the technology for sorption of the basic antibiotic streptomycin on cation exchange resins directly from fermentation broths in well-agitated tanks (actually in the form of upflow columns fitted with agitators to keep solids dispersed) without need for costly and loss-associated prefiltration. This breakthrough helped make adsorption processes competitive with traditional solvent extraction methods. Belter et al. published a similar anion exchange process to isolate the acidic antibiotic novobiocin in a series of specially designed, well-mixed columns with screens to pass mycelia but retain the resin (23). Periodically, the lead column was removed from the train, washed free of insolubles, and eluted in fixed-bed mode. Periodic countercurrent operation was obtained in the sorption sequence by advancing each of the uneluted columns one position and placing a freshly eluted column in the trail position of the train.

Modifications to surface chemistry of nonpolar adsorbent resins such as XAD-2 allowed specific capture of known fermentation by-products, such as removal of desacetylcephalosporin C from cephalosporin C, to produce a more potent antibiotic (24).

Blood Plasma Fractionation

The well-known Cohn fractionation process, originally developed in the 1940s to isolate albumin as a blood volume expander during World War II, relies primarily on a series of ethanol fractionation steps to separate the various proteins of human blood plasma from one another according to their solubility behavior in the presence of ethanol (25). As the role of other plasma components became recognized, such as factors VIII and IX for the treatment of clotting disorders and IgG for use in passive immunization, higher yield and higher purity large-scale purification methods for these components were sought, and the importance of adsorption techniques grew.

With the development of cellulose ion exchangers in the mid-1950s, followed by the introduction of agarose-based resins (Sephadex®) a few years later, new chromatographic techniques specifically designed for protein separations became available for use in purification of blood plasma components. Despite the advances made, large-scale application of these methods in the biological industry developed

slowly. One reason was the aseptic processing requirements often specified for biological raw materials. Chromatography using saline solutions and biodegradable supports could result in bacterial contamination if the appropriate precautions were not strictly followed. An inherent advantage of the Cohn fractionation process is that ethanol is bacteriostatic. But the main cause for the slow acceptance of chromatography at the large scale (i.e., fixed-bed operations rather than stirred-tank adsorption) was the poor mechanical qualities of the available macroporous supports designed for protein separations. However, these soft gels were used for some batch adsorption processes in stirred tank reactors for recovery of various blood products. Continual advances in resin technology through the 1970s provided supports that were mechanically strong and had surfaces and porosities well suited for protein chromatographic separations.

Although there are a number of advantages associated with column operations, packed beds of adsorptive media tended to act as depth filters for many of the biological feedstocks that were commonly used. This problem could be avoided by carrying out the adsorption in a stirred tank. This method has been used for many years for factor IX isolation from plasma, one of the first major applications of an adsorption method in plasma fractionation.

Factor IX complex, also known as prothrombin complex, can be prepared by a number of methods, but one of the most popular is to use batch adsorption with anion exchangers based on diethylaminoethyl (DEAE) groups attached to a stable support matrix such as cellulose, dextran, or agarose. Suomela et al. of the Finnish Red Cross Blood Transfusion Centre in Helsinki reported on a method developed to purify coagulation factor IX concentrate from human plasma (26). The production scale consisted of starting with 150 L of plasma, which was combined with pretreated and autoclaved DEAE-Sephadex A-50 and mixed for 30 minutes to complete adsorption. About 95% of the factor IX along with 5 to 6% of the other proteins were adsorbed from the plasma. The weakly bound proteins were removed by a low ionic strength wash. Adsorption was carried out in stainless steel tanks with mixing, followed by settling the gel to the bottom of the tank while the clear plasma was pumped off. The product containing DEAE-Sephadex gel was then pumped into a steel cylinder equipped with a bottom screen to retain the resin. Elution of the factor IX was achieved by increasing the buffer ionic strength. The resulting factor IX concentrate was purified about 100-fold, although it still contained relatively smaller amounts of factors II, VII, and X. The overall yield of the process was 60%.

A semicontinuous method for purification of factor IX complex from human plasma was described by Tharakan et al. (27). In the semicontinuous process, cryosupernatant plasma is pumped through a stirred column containing resin. Plasma was pumped out of the bottom of the column through the flow adapter screen at the same flow rate so that the volume in the column remained constant. After all the plasma had contacted the resin, the column was packed by gravity settling, and an upper flow adapter was installed. The column was washed with low ionic strength buffer until the absorbance of the effluent was negligible. The factor IX complex was then eluted with a higher ionic strength buffer. It was found that a residence time of 15 minutes was sufficient to capture 95% of the factor IX in the starting plasma. In the pilot plant, 550 L of plasma was passed through a 50-L column containing 8.5 L of resin, yielding a 68% recovery. Comparative studies involving batch, semicontinuous, and packed bed adsorption methods showed that although the amount of factor IX adsorbed from the plasma remained the same regardless of contacting mode, the mode of operation did affect the percentage of factor IX recovered, with increasing recoveries encountered as the resin density or ratio of resin to reactor volume increased (recoveries were 32%, 82%, and 100% batch, semicontinuous, and packed-bed operations, respectively). These results indicate that the adsorption and desorption process is not simple, and it was speculated that multipoint attachment resulting in denaturation or competitive binding of other proteins may have played a role. The semicontinuous method was developed to facilitate contacting plasma with inexpensive soft resins. It eliminates the need for handling a resin–plasma slurry, and it reduces total process time and labor requirements in comparison to batch adsorption. Furthermore, the resin is contained within one vessel (a modified column) throughout its use, thus minimizing handling and overall equipment requirements.

More recently, highly selective chromatography steps have been combined with the traditional batch adsorption step to separate factor IX from the other clotting factors (28) in human plasma. The factor IX concentrate from batch adsorption is passed through a more efficient anion exchanger, and factor VII is removed by stepwise elution. Separation of factor IX from factors II and X is carried out using a highly selective affinity column using heparin-derivatized agarose. Both factors II and X have a lower binding affinity for heparin than does factor IX. The resulting factor IX is purified more than 30-fold, resulting in an overall purification factor from plasma of about 10,000-fold.

The DEAE-ion exchangers have also been applied to the preparation of IgG, which, unlike other plasma proteins, is not bound by the DEAE-group at a neutral pH and low ionic strength. Several methods have been described that use ion-exchange materials alone or in combination with ethanol fractionation to purify IgG from plasma. A DEAE-Sephadex in combination with the ethanol fractionation technique was reported for the fractionation of IgG from blood plasma (29). The plasma proteins were first separated into three main fractions by increasing ethanol fractionations (8%, 25%, and 40% ethanol). The γ-globulin was recovered in the 25% ethanol fractionation precipitate. The paste was then dissolved and the pH adjusted to 6.5 in low ionic strength. DEAE-Sephadex (A50 coarse) was added at a level of about 1 g per gram of protein to be adsorbed, and the mixture was stirred for 30 minutes. The Sephadex gel was filtered off on a stainless steel Buchner-type funnel, and the γ-globulin, which was nonbound and in the filtrate, was subsequently dispensed for direct lyophilization.

Hepatitis B Purification

Fused silica, or Aerosil, was first used in a batch adsorption mode to adsorb lipoproteins from human serum (30).

The methodology was then applied to the adsorption of hepatitis B surface antigen (HBsAg) to remove hepatitis B virus from donated plasma (31). In the mid-70s, these methods were refined with respect to optimizing HBsAg adsorption selectivity and elution conditions for the purpose of purifying HBsAg from human plasma (32,33).

The classical methods for purification of HBsAg from sera were by isopycnic banding in a CsCl density gradient, followed by rate zonal centrifugation on a sucrose density gradient. It was cumbersome, however, to process large volumes of plasma or serum containing HBsAg by ultracentrifugation as the first step of purification. Duimel et al. (32) described a method whereby large quantities of antigen contained in serum could be partially purified by means of batch adsorption to colloidal silica (Aerosil) and then eluted with low ionic strength buffer (0.01 M borax, pH 9.3). Human positive serum was first extracted with Freon, which extracted part of the lipoids from serum, and then adsorbed with 2% Aerosil at 37 °C for 4 hours. After centrifugation for 15 min at 3,000 rpm, the Aerosil was washed four times with physiological saline. The Aerosil was eluted batchwise twice with borax, pH 9.3, for 30 min at 37 °C. Recoveries of HBsAg across this step were reported to be about 60%, providing a 15-fold purification factor. Pillot et al. (33) reported improving the recovery of the desorption step to 100% by eluting with 0.25% sodium deoxycholate in 0.01 M borax, pH 9.3, at 56 °C. Furthermore, elution at the higher temperature favored HBsAg desorption relative to some of the serum proteins, particularly albumin, which often represented the major contaminating protein eluted along with HBsAg.

The French vaccine manufacturer, Pasteur Merieux, used Aerosil adsorption with elution conditions similar to those described above for the production of a commercial plasma-derived HBsAg vaccine (34). In the mid-1980s, new second-generation hepatitis B vaccines based on recombinant DNA technology became available commercially. Many of the general purification principles that were used for the plasma-derived products were incorporated into purification schemes for HBsAg from recombinant sources (35). In the purification process for RECOMBIVAX HB® (Merck & Co., Inc.), batch adsorption of HBsAg onto colloidal silica is employed. The HBsAg is isolated from the clarified lysate by adsorption onto colloidal silica (Aerosil). The silica suspension is then collected by centrifugation, washed multiple times through resuspension and recentrifugation, and finally eluted from the silica by treatment with warm borate buffer. Sanford et al. reported that the antigen adsorbs to bare silica by a mixed mode and is specifically eluted by a chemical interaction between the silica surface and borate ion (36). More recently, this traditional batch operation has been evaluated for application in a fixed-bed column mode using a macroporous silica packing (37). The elution uses an extended warm borate buffer recycle through the column to achieve an equilibrium distribution between the entire bed volume before product collection. Although the adsorption, washing, and elution steps all take place within a column, it is still fundamentally a batch operation because the separation is achieved in a truly single stage contactor. The column mode of operation is able to be scaled and is more efficient in terms of overall cycle time as a result of the elimination of multiple centrifugation steps, despite the more lengthy column adsorption loading time.

NOVEL CONFIGURATIONS OF BATCH ADSORPTION

Fluidized Bed Contactors

In practice, batch adsorption generally requires large amounts of adsorbent because of relatively low product adsorption. Thus, a continuous or semicontinuous operation would be more efficient, but until recently, this has been severely constrained by the difficulties in processing viscous or debris-laden streams through fixed beds that act as depth filters and quickly clog. The use of a fluidized bed for batch adsorption of biomolecules is not entirely new, having been demonstrated in hybrid form for antibiotic recovery as early as 1958 (22). Because of the limitation of a well-mixed fluidized system to one theoretical plate, there was little inherent advantage of fluidized bed contactors over stirred tank systems. However, the introduction of segmented beds, magnetic stabilization, and, most recently, Pharmacia's expanded bed chromatography system has allowed multistage contacting in fluidized systems that exhibit comparable dispersion to packed beds of similarly sized particles (38).

Expanded-Bed Chromatography. In process-scale chromatography of proteins, most separations rely more on selectivity than on a high number of theoretical plates because achieving the latter necessarily requires high pressures (small diameter particles) and costly equipment. Given the accepted preference for low-pressure operations, a step from conventional packed-bed systems to recently developed expanded bed systems is often appropriate for unclarified feeds. These stable fluidized beds show axial dispersion and dynamic capacities comparable to those measured in packed bed.

An analogous technique for stabilizing fluidized beds incorporating the application of uniform magnetic fields to beds of magnetically susceptible adsorbent particles was extensively investigated by Rosensweig (39) and subsequently applied to biological separations by Burns and Graves (40). Although these systems have not been commercialized, they exhibit similar hydrodynamic characteristics to expanded-bed adsorption, facilitating processing of debris-laden streams while maintaining dispersion characteristics not dramatically greater than corresponding fixed beds. In this case, a uniform magnetic field prevents tumbling of the fluidized bed, creating a multistage contactor for potentially improved separation efficiency.

Whole Broth Adsorptive Extraction

Integrated fermentation and recovery may enable reduction of feedback inhibition resulting from product accumulation in the bioreactor, thus providing for an overall more productive system (41). Whereas integrated solvent

extraction requires intermittent phase separation and culture recycle, integrated adsorption can be accomplished either by inclusion of adsorbents within the bioreactor or by passing filter-clarified broth through an external contactor with continuous recycle to the bioreactor.

An improved system for direct-contact broth extraction was developed by immobilizing finely divided affinity adsorbent particles in a large hydrogel bead (42). The large bead size facilitates separation from broth components, and the reversible Ca^{2+} hydrogel facilitates efficient recovery of the costly affinity ligand for recycle and reuse. Hydrogels, by virtue of their extremely high water content (>90%), offer limited diffusional resistance to the desired product while protecting the affinity ligand from fouling components of the broth. Addition of adsorptive particles to a cell culture or fermentation broth can also remove those components from the liquid medium that are responsible for low filtration fluxes frequently encountered in clarification of untreated streams (43). Even in those cases where flocculants thus formed raise the viscosity of the medium, the resulting flux is still less significantly impaired than that decreased by the presence of a significant concentration polarization layer.

Incorporation of an external fluidized bed contactor minimizes attrition to both adsorbent particles and cells encountered in direct inclusion systems while allowing circulation of whole broth to avoid frequent membrane fouling associated with broth clarification. Single-stage recirculating fluidized beds are less susceptible to time-based changes in biomass and viscosity over the course of a fermentation cycle than expanded beds, which are best suited to single-pass batch treatments. The greatest limitation on implementation of such integrated fermentation and recovery systems, as in the case of fluidized-bed purification systems, has been the lack of available supports. Because of the strict sterility requirements in fermentation or cell culture applications, any such adsorbents must also withstand sterilization.

Continuous Affinity Recycle Extraction

A variant of batch adsorption referred to as continuous affinity-recycle extraction (CARE) was developed in the late 1980s. It combines adsorption purification with membrane filtration for the continuous separation of proteins (44). The method employs two continuous stirred-tank reactors (CSTRs), one for adsorption and the other for desorption, between which the adsorbent is recirculated. The feed solution is continuously added to the adsorption tank, where the desired product is contacted with the selectively adsorbing resin. Simultaneously, contaminants are diluted with the addition of wash buffer and removed by passage through a screen filter, which allows even crude lysates to be processed. A stream of the product-containing adsorbent is pumped into the second CSTR along with the desorbing buffer. The product is eluted in the tank and recovered through an ultrafiltration membrane in the permeate stream. The regenerated adsorbent is recycled to the adsorption tank. By controlling the relative rates of the feed, product, and recycle streams, which are dictated by the rates of adsorption in CSTR-1 and desorption in CSTR-2, the system can be operated continuously at a steady state.

This scheme was demonstrated for both ion exchange and affinity modes of purification for the recovery of β-galactosidase from a partially clarified *Escherichia coli* lysate. A recovery yield of 70% was achieved with a 35-fold purification factor, although the product was substantially diluted (45).

Ultrafiltration-Coupled Adsorption Systems

Mattiasson et al. (46) described a method of ultrafiltration affinity purification to combine purification based on affinity interactions with membrane separation. By selecting membranes with pore sizes large enough to freely pass the protein of interest but then introducing recirculating (membrane-rejected) macromolecular ligands with specific affinity for this protein, all molecules except the target protein are rapidly washed from the feed pool. When all proteins not bound to the ligand are removed, the affinity complex may be dissociated, and the liberated material can pass through the membrane for collection. In batch experiments capturing concanavalin A using yeast cells as ligands, yields up to 70% of electrophoretically homogeneous product have been obtained. As in column-affinity chromatography, the adsorption, wash, and desorption steps can be optimized independently.

Fletcher and Deley (47) applied a similar concept using recirculating fine adsorbent particles, such as Biocryl bioprocessing aids, to purification of a peptide from highly colored cell culture streams. After pH adjustment for optimal binding to the adsorbent resin and a short incubation, the bound peptide was concentrated twofold and then diafiltered to remove remaining color bodies. The permeate contained colored material but minimal peptide. The loaded particles were then diafiltered against elution buffer, and product peptide was collected in the permeate. This preliminary removal of load-limiting contaminant allowed a subsequent reverse phase chromatography step to be used for polishing purposes, with a 100-fold increase in column capacity.

OUTLOOK FOR PROTEIN BATCH ADSORPTION

As described in the section on protein adsorption principles and evidenced within the context of the examples discussed in the preceding text, the application of batch adsorption operations has been limited by the modest protein selectivities generally found for most commercially available adsorbent systems. Development of more sophisticated high-affinity adsorbents through combinatorial chemistry or phage display approaches, which use high throughout screening to tailor and design specific ligands, will certainly enable additional batch adsorption applications by providing high selectivity separations.

Important future applications of batch adsorption for biomolecule recovery and purification will likely be developed for several niche bioseparations. For example, adsorbent-based purification of very large biomolecules including plasmid DNA, viruses, and recombinant viruslike particles are often characterized by very slow diffusion through pores in even the most macroporous resins avail-

able. This can make chromatographic-based operations problematic, because the overall mass transfer can require long contact times or make the effective surface area appear quite small. The development of new adsorbent materials with megapores may provide improvements in separation performance. Moreover, limitations in column operation because of slow adsorption can be overcome in some cases if the separation is performed in a batch operation mode. Analogously, slow desorption may be handled optimally through batch operation using either a well-mixed contactor or a column with column effluent-recycle configuration. In such cases, the recycle mode may result in minimal volumes for product elution relative to the large volume if carried out in a single-pass flow mode.

Improvements in adsorbents and processing equipment, especially for product-laden adsorbent handling to be consistent with aseptic operations, along with greater understanding and modeling of protein transport and adsorptive reactions will allow invention of additional adsorbent-based process technology schemes. The coupling of purification technologies, such as through adsorptive membranes or more specific fluidized bed batch adsorbers, will likely be increasingly applied as enhanced variants of traditional batch adsorption.

BIBLIOGRAPHY

1. D.M. Ruthven, *Principles of Adsorption and Adsorption Processes*, Wiley, New York, 1984, pp. 29–121.
2. A. Velayudhan and C. Horvath, *J. Chromatogr.* **443**, 13–29 (1988).
3. L. Weaver and G. Carta, *Biotechnol. Prog.* **12**, 342–355 (1996).
4. H.A. Chase, *Chem. Eng. Sci.* **38**, 1099–1125 (1984).
5. P.M. Boyer and J.T. Hsu, *Adv. Biochem. Eng. Biotechnol.* **49**, 1–44 (1993).
6. G.L. Skidmore and H.A. Chase, in M. Streat ed. *Ion Exchange for Industry*, Ellis Horwood Limited, Chichester, UK, 1988, pp. 520–532.
7. A. Tongta, A.I. Liapis, and D.J. Siehr, *J. Chromatogr., A* **686**, 21–29 1994.
8. J.Y. Shi and R.A. Goffe, *J. Chromatogr. A* **686**, 61–71 1994.
9. G.L. Skidmore and H.A. Chase, in D.L. Pyle ed., *Separations for Biotechnology 2*, Elsevier, New York, 1990. pp. 112–128.
10. B. Lassen and M. Malmsten, *J. Colloid Interface Sci.* **186**, 9–16 (1997).
11. S.M. Slack and T.A. Horbett, in T.A. Horbett and J.L. Brash eds. *Proteins at Interfaces II*, ACS Symp. Series No. 602, American Chemical Society, Washington, D.C., 1995, pp. 112–128.
12. R.K. Scopes, *Protein Purification: Principles and Practice*, 3rd ed., Springer, New York, 1994, pp. 121–126.
13. M. Wahlgren, T. Amebrant, and I. Lundstrom, *J. Colloid Interface Sci.* **175**, 506–514 (1995).
14. J.J. Ramsden, *Chem. Soc. Rev.* **24**, 73–78 (1995).
15. D. Farnan, D.D. Frey, and C. Horvath, *Biotechnol. Prog.* **13**, 429–439 (1997).
16. R. Kurrat, J.E. Prenosil, and J.J. Ramsden, *J. Colloid Interface Sci.* **185**, 1–8 (1997).
17. M. Wahlgren and U. Elofsson, *J. Colloid Interface Sci.* **188**, 121–129 (1997).
18. X. Jin, J. Talbot, and N.-H.L. Wang, *AIChE J.* **40**, 1685–1696 (1994).
19. Q. Mao, R. Stockman, I. Prince, M. Hearn, *J. Chromatogr.* **646**, 67–80 (1993).
20. B. Yang, M. Goto, and S. Goto, *Sep. Sci. Technol.* **24**, 741–754 (1989).
21. P.A. Belter, E.L. Cussler, and W.-S. Hu, *Bioseparations*, Wiley, New York, 1988, pp. 145–177.
22. C.R. Bartels, G. Kleiman, J.N. Korzun, and D.B. Irish, *Chem. Eng. Prog.* **54**, 49–51 (1958).
23. P.A. Belter, F.L. Cunningham, and J.W. Chem, *Biotech. Bioeng.* **15**, 533–549 (1973).
24. J.L. Casillas, J.L. Garrido, J. Aracil, M. Martinez, F. Adde-Yobo, in D.L. Pyle ed. *Separations for Biotechnology*, Volume 2, Elsevier, New York, 1990, pp. 285–294.
25. E.J. Cohn, L.E. Strong, W.L. Hughes Jr., D.J. Mulford, J.N. Ashworth, M. Melin, and H.L. Taylor, *J. Am. Chem. Soc.* **68**, 459–475 (1946).
26. H. Suomela, G. Myllyla, and E. Raaska, *Vox Sang.* **33**, 37–50 (1977).
27. J.P. Tharakan, D.M. Gee, and D.B. Clark, *Vox Sang.* **57**, 233–239 (1989).
28. T. Burnouf, in C.V. Prowse ed. *Plasma and Recombinant Blood Products in Medical Therapy*, Wiley, New York, 1992, pp. 67–87.
29. L.-G. Falksveden and G. Lundblad, in J.M. Curling ed. *Methods of Plasma Protein Fractionation*, Academic Press, New York, 1980, pp. 93–103.
30. W. Stephan and L. Roka, *Z. Clin. Chem. Clin. Biochem.* **6**, 186–190 (1968).
31. J.C. Siebke, E. Kjeldsberg, and T. Traavik, *Acta Pathol. Microbiol. Scand., B*, **80**, 935–936 (1972).
32. W.J.M. Duimel, H.G.J. Brummelhuis, and H.W. Krijnen, *Vox Sang.* **23**, 249–255 (1972).
33. J. Pillot, S. Goueffon, and R.G. Keros, *J. Clin. Microbiol.* **4**, 205–207 (1976).
34. F. Barin, M. Andre, A. Goudeau, P. Coursager, and P. Maupas, *Ann. Microbiol.* **129B**, 87–100 (1978).
35. R.D. Sitrin, D.E. Wampler, and R.W. Ellis, in R. Ellis ed. *Hepatitis B Vaccines in Clinical Practice*, Marcel Dekker, New York, 1993, pp. 83–101.
36. W. Sanford, D. Kubek, and R.D. Sitrin, *Am. Inst. Chem. Eng.*, Fall Meeting, Abstract 2751, Los Angeles, Calif., November 17–22, 1991.
37. M. Kosinski, D. Krips, W. Adams, W. Sanford, D. Kubek, and R. Sitrin, *Recovery of Biological Products VIII Meeting*, cosponsored by ACS and Engineering Foundation, Abstract D.9, Tucson, AZ, October 20–25, 1996.
38. V. Goetz and D.J. Graves, *Powder Technol.* **64**, 81–92 (1991).
39. R.E. Rosensweig, *Science* **204**, 57–60 (1979).
40. M.A. Burns and D.J. Graves, *Biotechnol. Prog.* **1**, 95–103 (1985).
41. P. Morton and A. Lyddiatt, in D.L. Pyle ed. *Separations for Biotechnology*, Volume 3, Elsevier, New York, 1994, pp. 329–335.
42. S.C. Nigam and H.Y. Wang, in J.A. Asenjo and J. Hong eds. *Separation, Recovery, and Purification in Biotechnology*, ACS Symp. Series 314, American Chemical Society, Washington, D.C., 1986, pp. 153–168.
43. U.S. Pat. 4,830,753 (May 16, 1989), J.M.S. Cabral, E.M. Robinson, and C.L. Cooney (to Rohm and Haas Company, Philadelphia, Penn.).

44. N.F. Gordon and C.L. Cooney, in M.R. Ladisch ed., *Protein Purification: From Molecular Mechanisms to Large-Scale Processes*, ACS Symp. Series 427, American Chemical Society, Washington, D.C., 1990, pp. 118–138.
45. E. Pungor, N.B. Afeyan Jr., N.F. Gordon, C.L. Cooney, *Biotechnology* **5**, 604 (1987).
46. B. Mattiasson, T.G.I. Ling, and J.L. Nilsson, in I.M. Chaiken, M. Wilcheck, and I. Parikh eds. *Affinity Chromatography and Biological Recognition*, Academic Press, New York, 1983, pp. 223–227.
47. K. Fletcher and S. Deley, in D.L. Pyle ed. *Separations for Biotechnology*, Volume 2, Elsevier, New York, 1990, pp. 142–294.

See also ADSORBENTS, INORGANIC; ADSORPTION, EXPANDED BED; ADSORPTION, PROTEINS WITH SYNTHETIC MATERIALS.

ADSORPTION, PROTEINS WITH SYNTHETIC MATERIALS

JOSEPH MCGUIRE
MICHELLE K. BOTHWELL
Oregon State University
Corvallis, Oregon

KEY WORDS

Adhesion
Adsorption
Adsorption kinetics
Competitive adsorption
Conformational change
Exchange reaction
Interface
Kinetic modeling
Surface energetics
Synthetic materials

OUTLINE

INTRODUCTION

In general, the surfaces of materials of almost any type that come into contact with protein mixtures tend to become quickly occupied by proteins. These protein films can and often do lead to profound alterations in the properties of the material–fluid interface, affecting material performance in a number of bioprocess and biomedical applications. Performance of immobilized enzyme bioreactors, purification of therapeutic proteins in the biotech industry, drug formulation strategies, bacterial adhesion and contamination of industrial process equipment, and cell adhesion to natural and synthetic materials in the body are only a few examples of processes impacted by protein behavior at interfaces. The key impediment to needed progress in all these areas continues to be our inability to predict what the eventual make-up and nature of these films will be under a given set of conditions. Protein interactions with solid surfaces have been studied for decades, and several reviews are available (1–5). In this treatment, we describe some generally well-understood, important results that have provided a basis for meeting the challenges facing engineers and scientists in this area.

INTERFACES

Interfaces, as well as the interactions that take place in interfacial regions, can be complex. In fact, the interface has been described as a fourth state of matter (6). The properties of atoms or atomic groups at a material surface are different than those of the bulk material. The first layer of atoms, in contact with the fluid phase, is particularly unique. Chemical composition, molecular orientation, and properties relevant to crystallinity differ at the surface. In addition, surfaces have different electrical and optical properties and can be characterized by atomic- or molecular-level textures and roughnesses. Surfaces have wettabilities or hydrophobic/hydrophilic balances related to the factors named above. Further, surfaces are generally energetically heterogeneous. For example, although a surface may be assigned a particular wettability, it would most likely be the result of a distribution of surface regions of varying wettabilities.

In spite of this complexity, many researchers have met with success in describing some aspect of protein adsorption in terms of one or several surface properties. The effects of charge distribution, surface energy (i.e., whether it is high or low), and surface hydrophobicity, for example, have received much attention (1–5). From a purely thermodynamic standpoint, the extent of protein adsorption or biological adhesion in general could be determined purely by surface energetics, that is, the surface energies of the synthetic material, liquid medium, and adsorbates involved. Such an approach would imply that the free energy of adsorption is minimized at equilibrium. Adsorption would be favored if it caused the free energy function to decrease and would not be favored if it caused the function to increase. In the absence of electrostatic and specific receptor–ligand interactions, the change in free energy upon adsorption could be written

$$\Delta F_{ads} = \gamma_{AS} - \gamma_{AL} - \gamma_{SL} \quad (1)$$

where F_{ads}(J/m^2) is the free energy of adsorption per unit of surface area, and γ_{AS}, γ_{AL}, and γ_{SL} (J/m^2) are the

adsorbate-solid, adsorbate-liquid, and solid-liquid interfacial energies, respectively.

If all the required interfacial energies of equation 1 could be estimated, one could predict the relative extent of adsorption among different surfaces. This would lead to a distinction between two situations (7,8), depending on whether adsorbate surface energy is greater than or less than the surface energy of the suspending liquid. Concerning protein adsorption from aqueous media, equation 1 would predict increasing adsorption with decreasing surface energy. In other words, a given protein would be expected to adsorb with greater affinity to hydrophobic as opposed to hydrophilic surfaces.

The importance of hydrophobic/hydrophilic balance in protein adsorption has prompted numerous investigators to develop techniques for measurement of this property at solid surfaces. Contact angle methods have been prominent in this regard (9). Contact angle analysis is inexpensive, rapid, and fairly sensitive. However, contact angle data can be difficult to interpret, and the technique is subject to artifacts caused by macroscopic, energetic heterogeneities in the surface, hysteresis, and drop-volume effects, among others. Still, useful conclusions regarding biological interactions with surfaces have been based on the results of contact angle analysis in areas of red blood cell adhesion, platelet adhesion, bacterial adhesion, and protein adsorption (9,10). Surface properties have been correlated to biological responses using other methods as well, including electron spectroscopy for chemical analysis (ESCA), secondary ion mass spectroscopy (SIMS), infrared and vibrational methods, and scanning probe microscopies. These methods and their relevance to biomedical technology were reviewed by Ratner and Porter (9).

The properties of a synthetic material's surface play a large role in dictating any biological response the material may evoke. But although much is known about selected surface property effects on protein adsorption, in a quantitative sense we know very little about how the molecular properties of protein influence its adsorption. Interfacial behavior is a cumulative property of a protein, influenced by many factors; among these are its size, shape, charge, and thermodynamic (thermal, structural, or conformational) stability. Experimentally observed differences in interfacial behavior among different protein molecules have been very difficult to quantify in terms of these factors because proteins usually vary substantially from one another in each category. The following discussion is an attempt to summarize the salient results from a wide range of experiments, focused on study of surface, solution, and protein effects on adsorption. Note that many observations have been explained in terms of a protein's charge, its tendency to unfold, and contact surface hydrophobicity.

PROTEINS AT INTERFACES

General Features of Adsorption from Single-Protein Solutions

Solution Chemistry Effects on Adsorption. The net charge of a protein in solution is dependent on the difference between pH of the solution and the isoelectric point (pI) of the protein. If the pH of the solution is greater than the pI, the net charge of the protein would be negative, whereas if the pH is less than the pI, the net charge of the protein would be positive. It is generally accepted that maximum adsorption occurs at the isoelectric point. As the out-of-balance charge of a protein increases, it will be in a more extended form than when the net charge is zero (11).

Norde and Lyklema (12) suggested that the degree to which pH affects the adsorption of a protein is determined by its conformational stability. They found that plateau values of adsorbed mass were independent of pH for structurally stable proteins, whereas those of less stable proteins varied considerably, apparently because less stable proteins were able to change structure with solution conditions. The effect of pH on protein adsorption and desorption can depend on solution history as well (13). Kondo and Higashitani (14) studied the adsorption of model proteins with wide variation in molecular properties. They explained the pH dependence of adsorbed mass in terms of lateral interactions. In particular, they suggested that lateral interactions between large protein molecules are stronger than those between small molecules. Large proteins would thus be expected to show maximum adsorption around their pIs, whereas the effect of pH on smaller proteins would be less pronounced.

The degree to which ionic strength affects protein adsorption is a function of the role electrostatic plays in the adsorption driving force. At low ionic strength, protein surface charge fully contributes to the total electrostatic interaction (11). At high ionic strength, the surface charges of proteins are shielded, reducing electrostatic interactions between proteins, whether attractive or repulsive (13). Luey et al. (15) showed that ionic strength effects on adsorbed mass are very much related to solid surface properties. They observed that increased ionic strength reduced the electrostatic repulsion between negatively charged β-lactoglobulin molecules and the hydrophilic, negatively charged surface they studied, increasing adsorbed mass. By contrast, increased ionic strength resulted in little change in adsorbed mass at hydrophobic surfaces.

Surface-Induced Conformational Changes. It is well accepted that a given protein can exist in multiple adsorbed conformational states on a surface (16–20). These states can be distinguished by differences in occupied area, binding strength, propensity to undergo exchange events with other proteins, and catalytic activity or function. All these features of adsorbed protein are interrelated and can be time-dependent. For example, decreases in surfactant-mediated elution of proteins from an adsorbed layer (an indirect measure of binding strength) are observed as protein–surface contact time increases (21). This time-dependence is illustrated in Figure 1. As conformational change proceeds, the likelihood of desorption decreases.

It has been observed that the extent of conformational change experienced by adsorbed fibrinogen increases with contact surface hydrophobicity (22). This is consistent with findings of Elwing et al. (23), who used ellipsometry to make inferences regarding conformational changes experienced by complement factor III, a plasma protein, on hydrophilic and hydrophobic silica surfaces. The results of

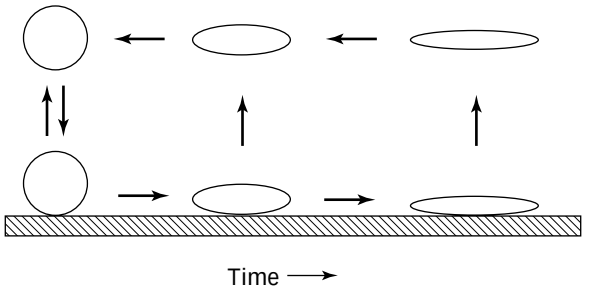


Figure 1. Surface-induced conformational changes undergone by adsorbed protein, resulting in multiple noncovalent bonds with the surface and coverage of greater interfacial area per molecule.

Elwing et al. also showed that greater values of adsorbed mass were found on hydrophobic as opposed to hydrophilic surfaces. Protein molecules are assumed, in general, to change conformation to a greater extent on hydrophobic surfaces. This is due to the effect of hydrophobic interactions between the solid surface and hydrophobic regions in the protein molecule. In fact, surface-induced unfolding is often characterized as entropically driven, because the hydrophobic protein interior associates with hydrophobic regions of the surface during unfolding. These interactions can give the molecule an extended structure, covering a relatively large area of the surface. If the repulsive force normally acting between native protein molecules is decreased for such structurally altered molecules, one would expect to measure a greater adsorbed mass on hydrophobic as opposed to hydrophilic surfaces. On the other hand, adsorption of positively charged protein to hydrophilic (negatively charged) silica can result in greater conformational change than adsorption of the same protein to hydrophobic silica, even with a greater extent of adsorption being observed at the hydrophobic silica surface (24). It is thus important to recognize that multiple factors affect the extents of protein adsorption and conformational change.

The concept that adsorbed proteins can exist in multiple states on a surface plays a role in interpretation of most if not all experiments in protein adsorption. Biophysicists rather routinely gain information relevant to protein structure in solution with circular dichroism (CD). It would be attractive to use CD in structural study of adsorbed protein as well. A recent innovation has made CD more applicable to study of structural changes during adsorption, and that is the use of colloidal silica particles, or nanoparticles (16–20). In these tests, particles have ranged from less than 10 to about 30 nm and are small enough not to interfere with the CD spectra. Individual molecules are allowed to adsorb to nanoparticles, resulting in a stable suspension of adsorbed protein. In this way, structural changes on adsorption have been unambiguously measured.

Work by Billsten et al. (19) and Tian et al. (20) have provided the most direct illustration of the effect of stability on structural rearrangements at a solid surface. Using site-directed mutants of bacteriophage T4 lysozyme, these investigators showed that both the rate and extent of secondary structure loss upon adsorption to colloidal silica were clearly related to protein thermal stability. With the same mutants, Fröberg et al. (25) used the interferometric surface force technique to study structural characteristics of adsorbed layers of T4 lysozyme. The results demonstrated that less-stable mutants lose their tertiary structure upon adsorption, whereas more stable mutants retain their globular shape.

Steady-State Adsorption Behavior. A great deal is known about how various conditions affect the steady-state adsorbed mass of protein. Numerous protein adsorption isotherms have been constructed and compared on the basis of temperature, pH, ionic strength, conformational stability of the protein in solution, and solid surface charge and hydrophobicity. The effects of protein conformational stability and solid surface properties are perhaps best revealed with reference to effects of pH and ionic strength.

In general, the effect of pH and ionic strength on protein adsorption is dependent on which type of interactions predominate (e.g., electrostatic, hydrophobic, or van der Waals interactions). At a negatively charged surface, if electrostatic interactions predominate, adsorbed mass should be greater at pH values below the isoelectric point relative to pH values above it. Below the isoelectric point, the protein and surface would be of opposite charge, whereas both the protein and surface would be negatively charged at pH values greater than the isoelectric point. As ionic strength increases, the electrostatic interaction would be reduced because of shielding of the protein by counterions; consequently, increasing the ionic strength should decrease adsorbed mass at pH values less than the isoelectric point and increase the adsorbed mass at greater values of pH. The relationship between adsorbed mass and changes in pH and ionic strength becomes inextricably linked to protein conformational stability. In general, pH and ionic strength conditions that lead to a less stable conformation for the protein in solution will lead to an increased adsorbed mass, assuming that the protein molecule would be more stable on the solid surface (15).

Another observation of importance is that protein adsorption is often an apparently irreversible process, at least in the sense that is often irreversible to dilution or buffer elution. The adsorbed mass remains constant or decreases very little when the solution in contact with the solid surface is depleted of protein. This irreversibility is more pronounced as protein–surface contact time increases. However, although spontaneous desorption is not generally observed, adsorbed protein can undergo exchange reactions with similar or dissimilar protein molecules adsorbing from solution (26). Such exchange reactions are shown schematically in Figure 2. Adsorbed protein exchange rates are likely state-dependent, being slower for more conformationally altered protein.

Some workers have reported that protein adsorbs onto a solid surface in more than one layer. Arnebrant et al. (27) studied adsorption of β-lactoglobulin and ovalbumin on hydrophilic and hydrophobic chromium surfaces using ellipsometry and electrical potential measurements. On hydrophilic surfaces, their results showed that a highly hydrated layer is obtained that can be partially removed by rinsing. They suggested that the protein adopts a bilayer formation

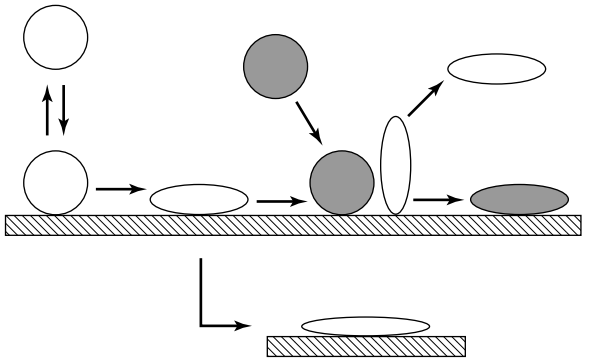


Figure 2. Exchange reaction between a conformationally altered, adsorbed protein and a dissimilar protein adsorbing from solution.

on the surface, with the layer in direct contact with the surface being unfolded and attached by strong polar bonds. Rinsing showed that the outer protein layer is loosely attached, which would imply that molecules in the outer layer have a structure closer to that of their native state. This adsorption behavior was described in terms of surface-induced conformational changes and charge interactions between the protein and surface. In particular, there are always polar amino acid side chains that can interact strongly with a surface, even if both the protein and surface are negatively charged. Such binding might be expected to result in unfolding of the protein. A consequence of this might be exposure of hydrophobic regions into aqueous solution; adsorption of a second protein layer would therefore reduce the interfacial free energy. In the case of protein in contact with a hydrophobic metal surface, values of adsorbed mass were found to be consistent with formation of a monolayer. Arnebrant and Nylander (28) reported possible bilayer formation upon adsorption of oligomeric units of insulin as well.

Adsorption Kinetics. In considering the kinetics of any interfacial process, the question of transport versus reaction control must be addressed. Protein adsorption at an interface is dependent not only on the intrinsic kinetic rate, which is a function of protein, solution, and surface properties, but also by the rate of protein transport from the bulk solution through the concentration boundary layer near the interface.

Proteins are macromolecules, and they can possess domains that differ chemically and physically. Diffusion coefficients may vary widely among proteins, depending on their concentration and the electrostatic condition of the solution (29). The initial adsorption rate of protein molecules at an interface can be transport-limited at either low or high concentration. The diffusion limitation exists as long as there is a significant concentration gradient near the solid surface. With careful design of an experimental system to minimize the transport-limited period, however, an intrinsic adsorption kinetic rate can be estimated. Still, relatively little is known about the nature of the adsorbed layer, and predictive models to describe any aspect of adsorption as a function of protein and interfacial properties are lacking. Protein adsorption is characterized by the likely presence of a time dependence in the development of bonds with the surface, a time dependence in the lateral mobility and exchangeability of the protein molecules, and time-dependent conformational changes; it is thus very difficult to describe mathematically.

Many experimental observations have indicated that a major portion of the final adsorbed amount had been adsorbed within the first few minutes of contact. Soderquist and Walton (30) proposed the existence of three distinct processes contributing to protein adsorption kinetics on polymeric surfaces. First, rapid and reversible adsorption of the proteins occurs during a short period of time. At up to 50 to 60% of surface coverage there is a random arrangement of adsorbed molecules, but then some form of surface transition occurs that is probably in the direction of surface ordering, thereby allowing further protein adsorption. Second, molecules on the surface undergo structural transitions as a function of time that occur in the direction of optimizing the protein-surface interaction. Third, as time increases the probability of desorption decreases, and the adsorption becomes irreversible to dilution.

Competitive Adsorption

Molecular Structure and Interfacial Behavior. Study of molecular influences on protein adsorption has received much attention because of its relevance to better understanding of adsorption competition in multiprotein mixtures. Important contributions to current understanding of molecular influences on protein adsorption have evolved from several comparative studies of protein interfacial behavior, in which similar or otherwise very well-characterized proteins (31–34), genetic variants (35,36), or site-directed mutants (19,20,24,25,37–39) of a single protein had been selected for study. A number of factors are known to affect protein adsorption, and these studies have stressed the importance of protein charge, hydrophobicity, and structural stability in interfacial behavior.

Shirahama et al. (33) studied hen lysozyme, ribonuclease A, and α-lactalbumin adsorption to hydrophilic and hydrophobic, polystyrene-coated silica (both negatively charged surfaces). At hydrophilic silica, they found that adsorbed mass increased with increasing charge contrast between the surface and protein. At the hydrophobic surface they found electrostatic interaction to have a lesser effect, in that the adsorbed mass was not clearly related to charge contrast between the surface and protein. Arai and Norde (31) described adsorption from single-protein solutions of hen lysozyme, ribonuclease A, myoglobin, and α-lactalbumin to synthetic materials varying in surface charge density and hydrophobicity. They concluded that at a given surface, adsorption of a globular protein is related to its structural stability. That is, proteins of high stability behave like hard particles at a surface, with the interactions governed by surface hydrophobicity and electrostatics, whereas adsorption of proteins of low stability (soft proteins) may be influenced by structural rearrangement, allowing adsorption to occur even under conditions of electrostatic repulsion.

Horsley et al. (35) compared isotherms constructed for hen and human lysozymes at silica derivatized to exhibit

negatively charged, positively charged, or hydrophobic surfaces. Differences in adsorptive behavior observed between the two lysozyme variants were largely explained with reference to the fact that human lysozyme contains one less disulfide bond and is less thermally stable than hen lysozyme. Xu and Damodaran (36) compared adsorption kinetic data measured for native and denatured hen, human, and bacteriophage T4 lysozymes at the air–water interface. Their results showed substantial differences in adsorption dynamics among the three variants, as influenced by their structural state and the physical and chemical nature of the protein and surface.

Kato and Yutani (37) evaluated the interfacial behavior of six site-directed mutants of tryptophan synthase α-subunits, produced by amino acid substitution in the protein's interior, using surface tension, foaming, and emulsifying property measurements. The stability of these subunits, as measured by their free energy of denaturation in water, varied from about 5 to 17 kcal/mol. They were able to attribute differences in interfacial behavior to protein stability with good success. In particular, they observed that less stable mutants were most surface active, that is, they more rapidly adsorbed or more readily unfolded at the hydrophobic interfaces studied in that work.

Multicomponent Systems. Shirahama et al. (33) used hen lysozyme, ribonuclease A, and α-lactalbumin to study sequential and competitive adsorption at hydrophilic and hydrophobic, polystyrene-coated silica. At hydrophilic silica, they found that once an adsorbed layer of a given protein was formed, almost total displacement of that protein would occur upon introduction of a second protein to solution if the second protein had a more favorable capacity for electrostatic attraction with the surface (otherwise, sequential adsorption was not observed). In addition, when adsorbed from a mixture, the protein capable of more favorable electrostatic attraction with the surface preferentially adsorbed, essentially to the exclusion of the other protein. At the hydrophobic surface, they found that once an adsorbed layer of a given protein was formed, only partial displacement of that protein would occur upon introduction of a second protein to solution, even if the second protein had a more favorable capacity for electrostatic attraction with the surface. Moreover, the eventual make-up of a film adsorbed from a mixture was not related to charge contrast between the protein and surface. Other experimental observations (40,41) indicated that adsorbed protein molecules undergo exchange with protein from solution more readily on hydrophilic than on hydrophobic regions of a surface. Arai and Norde (32) studied the sequential and competitive adsorption behavior exhibited by hen lysozyme, ribonuclease A, myoglobin, and α-lactalbumin and concluded that whether introduced in sequence or in a mixture, adsorption of a globular protein is related to its structural stability. In particular, interfacial behavior of proteins of high stability is governed by surface hydrophobicity and electrostatics, whereas that of proteins of low stability are more influenced by structural rearrangement.

The elution of adsorbed protein by surfactant has been used to provide a measure of protein binding strength (21,24,42–51). The essential steps of this type of experiment are illustrated in Figure 3. Adsorption is allowed to occur, followed by rinsing with protein-free buffer. A surfactant solution is then introduced, after which adsorbed protein is displaced or solubilized (45). This is followed by rinsing and comparison of the amounts of protein present before surfactant addition and after the final rinse. Elution by dissimilar protein has been used as a measure as binding strength as well. Slack and Horbett (52) evaluated the strength of attachment of fibrinogen to solid surfaces by measuring its time-dependent elution in plasma and modeled fibrinogen adsorption, with reference to its rate of conversion from a weakly bound (exchangeable) to a tightly bound (nonexchangeable) state. Wahlgren and Arnebrant (45,46) used in situ ellipsometry to continuously monitor the different effects of cationic and anionic surfactants on the elution of β-lactoglobulin and lysozyme from hydrophilic and hydrophobic surfaces as well as adsorption from protein and surfactant mixtures. The elution studies allowed postulation of four mechanisms for surfactant-mediated elution of adsorbed protein. With the aim of relating elutability to protein molecular properties, Wahlgren et al. (47) studied removal of well-characterized proteins from silica surfaces using dodecyltrimethylammonium bromide. Some general trends regarding molecular property effects on elutability emerged from that work, but clear correlations between molecular properties and elutability remained difficult to quantify. By contrast, similar tests conducted with synthetic mutants of bacteriophage T4 lysozyme showed a clear correlation between protein stability and elutability (24). In particular, less stable proteins are more resistant to elution, presumably because they are more able to alter their conformation at the surfaces.

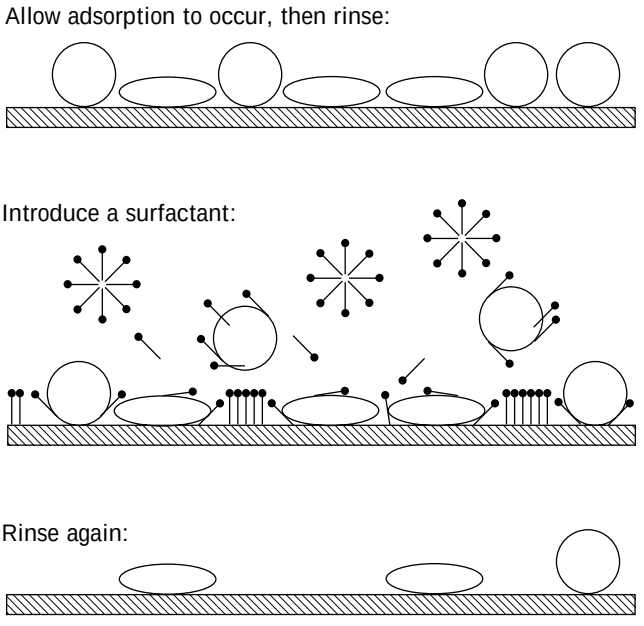


Figure 3. Experimental approach to evaluating adsorbed protein-binding strength using surfactant-mediated elution.

Modeling the Process. A number of mathematical models of protein adsorption at air–water and solid–water interfaces have been constructed (38,53–57). The problem is generally approached as an issue of molecular diffusion through a potential gradient, a reaction-diffusion problem involving interactions between diffusive-convective protein transport from the bulk solution and competitive adsorption and exchange kinetics on the surface, or as a kinetically controlled phenomenon, involving adsorption, unfolding, and exchange. Such models and the kinetic simulations they allow provide a framework with which the complexity of protein adsorption can be better understood and quantified. In addition, they can be used to provide direction for further experiments, particularly involving surface modification.

Lundström (58) presented an equilibrium model of protein adsorption on solid surfaces. The model described the fractional surface coverage of adsorbed molecules as a function of equilibrium concentration and allowed for reversible adsorption and conformation change. Later, Lundström and Elwing (26) described a model that allowed for bulk-surface exchange reactions among proteins in single-component and binary mixtures. The work featured manipulation of the equations describing the fractional surface coverage of protein in specific states and simulations of total surface coverage as a function of equilibrium concentration and as a function of time. Although no experimental data were presented, the shapes of the curves were in qualitative agreement with experimental observations. Currently, there is no adequate method to directly monitor changes in fractional surface coverages of protein in different adsorbed states. A less-complex model that can be statistically compared with the available data would be useful, because it would enable individual rate constants to be related to surface, solution, and protein properties.

Past work with synthetic mutants of bacteriophage T4 lysozyme have involved in situ ellipsometry and surfactant-mediated elution (24,51), ring tensiometry (38), the interferometric surface force technique (25), CD (19,20), and spectrophotometric assays of bound enzyme activity (39). These studies have shown that structural alterations definitely occur upon adsorption, with the extent and rate of structural change being related to thermal stability. In addition, mutants exhibited resistances to surfactant-mediated elution that were proportional to thermal stability and consequently related to extent of structural change. Finally, concerning a number of T4 lysozyme variants, results could be explained by modeling adsorption as occurring such that molecules adopt one of only two states, each varying in binding strength and occupied area, with differences in behavior among the molecules attributable to the relative populations in each state.

The simplest adsorption mechanism consistent with the fact that adsorbed proteins can exist in multiple states would include two adsorbed states. Figure 4 shows such a mechanism. Rate constants k_1 and k_2 govern adsorption into states 1 and 2, respectively. Although the mechanism is drawn to depict molecules adsorbing directly into states 1 and 2 from solution, a more accurate and detailed mechanism might include a multistep path to state 2. However,

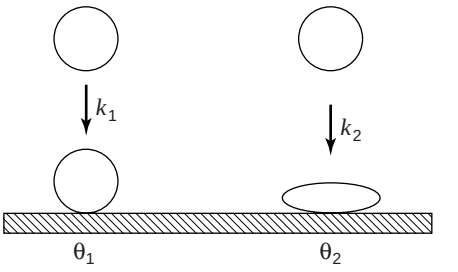


Figure 4. A simple mechanism for adsorption from a single-protein solution, allowing adsorption into one of two states exhibiting different resistances to elution and occupying different interfacial areas.

for modeling purposes, the actual path to state 2 is not consequential; we need only account for the different rates of generation of two functionally dissimilar forms of adsorbed protein. If adsorption of practically relevant proteins can be adequately described in this way, extension to the case of competitive protein adsorption would be straightforward.

Figure 5 shows a mechanism for competitive adsorption (between two proteins, A and B) based on Figure 4. In each case, all associated rate constants can be determined a priori. The protein-specific k_1C and k_2C of Figure 4 can be obtained from single-component kinetic data, and the various exchange constants can be determined through sequential adsorption experiments (59). Figure 5 can be easily redrawn to depict competitive adsorption of three proteins, A, B, and C or more. As long as sequential adsorption data are available for each pair permutation of A, B, and C, for example, an a priori estimate of all rate constants can be made, and the adsorption competition can be simulated. Such comparisons would provide a basis for design of further experiments to better resolve molecular ef-

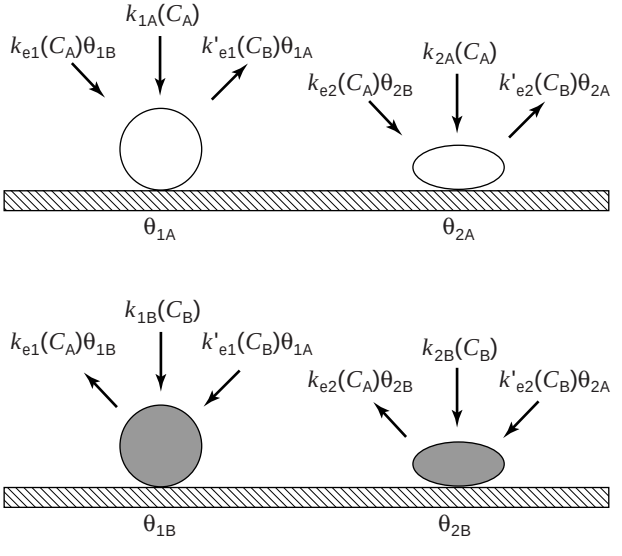


Figure 5. A mechanism for competitive adsorption between two proteins A and B, based on the single-component mechanism of Figure 4.

fects on competition in complex mixtures, subsequently enabling more quantitative prediction of adsorbed layer effects on practically relevant phenomena.

BIBLIOGRAPHY

1. J.D. Andrade, in J.D. Andrade ed. *Surface and Interfacial Aspects of Biomedical Polymers. Volume 2; Protein Adsorption*, Plenum Press, New York, 1985, pp. 1–80.
2. W. Norde, *Adv. Colloid Interface Sci.* **25**, 267–340 (1986).
3. T.A. Horbett and J.L. Brash, in J.L. Brash and T.A. Horbett eds. *Proteins at Interfaces: Physicochemical and Biochemical Studies*, ACS Symp. Series 343, American Chemical Society, Washington, D.C., 1987, pp. 1–35.
4. J.L. Brash and T.A. Horbett, in T.A. Horbett and J.L. Brash eds. *Proteins at Interfaces II: Fundamentals and Applications*, ACS Symp. Series 602, American Chemical Society, Washington, D.C., 1995, pp. 1–23.
5. J.D. Andrade, V. Hlady, L. Feng, and K. Tingey. in J.L. Brash and P.W. Wojciechowski eds. *Interfacial Phenomena and Bioproducts*, Dekker, New York, 1996, pp. 19–56.
6. C.B. Duke, *J. Vac. Sci. Technol., A* **2**, 139–143 (1984).
7. D. Absolom, F. Lamberti, Z. Policova, W. Zingg, C. Van Oss, and A. Neumann, *Appl. Environ. Microbiol.* **46**, 90–97 (1983).
8. D.R. Absolom, L.A. Hawthorne, G. Chang, *J. Biomed. Mater. Res.* **22**, 271–285 (1988).
9. B.D. Ratner, S.C. Porter, in J.L. Brash and P.W. Wojciechowski eds. *Interfacial Phenomena and Bioproducts*, Dekker, New York, 1996, pp. 57–84.
10. R.E. Baier and A.E. Meyer, in J.L. Brash and P.W. Wojciechowski eds. *Interfacial Phenomena and Bioproducts*, Dekker, New York, 1996, pp. 85–122.
11. S. Lee and E. Ruckenstein, *J. Colloid Interfac. Sci.* **125**, 365–379 (1988).
12. W. Norde and J. Lyklema, *J. Colloid Interfac. Sci.* **66**, 257–265 (1978).
13. P. Bagchi and S. Birmbaum, *J. Colloid Interfac. Sci.* **83**, 460–478 (1981).
14. A. Kondo and K. Higashitani, *J. Colloid Interfac. Sci.* **150**, 344–351 (1992).
15. J. Luey, J. McGuire, and R.D. Sproull, *J. Colloid Interface Sci.* **143**, 489–500 (1991).
16. A. Kondo, S. Oku, and K. Higashitani, *J. Colloid Interface Sci.* **143**, 214–221 (1991).
17. A. Kondo, F. Murakami, and K. Higashitani, *Biotechnol. Bioeng.* **40**, 889–894 (1992).
18. W. Norde and J.P. Favier, *Colloids Surf.* **64**, 87–93 (1992).
19. P. Billsten, M. Wahlgren, T. Arnebrant, J. McGuire, and H. Elwing, *J. Colloid Interface Sci.* **175**, 77–82 (1995).
20. M. Tian, W.-K. Lee, M.K. Bothwell, and J. McGuire, *J. Colloid Interface Sci.*, **200**, 146–154 (1998).
21. J.L. Bohnert and T.A. Horbett, *J. Colloid Interface Sci.* **111**, 363–377 (1986).
22. D.R. Lu and K. Park, *J. Colloid Interface Sci.* **144**, 271–281 (1991).
23. H. Elwing, S. Welin, A. Askendal, and I. Lundström, *J. Colloid Interface Sci.* **123**, 306–308 (1988).
24. J. McGuire, M.C. Wahlgren, and T. Arnebrant, *J. Colloid Interface Sci.* **170**, 182–192 (1995).
25. J.C. Fröberg, T. Arnebrant, J. McGuire, and P.M. Claesson, *Langmuir*, **14**, 456–462 (1998).
26. I. Lundström and H. Elwing, *J. Colloid Interface Sci.* **136**, 68–84 (1990).
27. T. Arnebrant, B. Ivarsson, K. Larsson, I. Lundström, and T. Nylander, *Prog. Colloid Polym. Sci.* **70**, 62–66 (1985).
28. T. Arnebrant and T. Nylander, *J. Colloid Interface Sci.* **122**, 557–565 (1988).
29. E.L. Cussler, *Diffusion: Mass Transfer in Fluid Systems*, Cambridge Univ. Press, Cambridge, Mass., 1989.
30. M.E. Soderquist and A.G. Walton, *J. Colloid Interface Sci.* **75**, 386–397 (1980).
31. T. Arai and W. Norde, *Colloids Surf.* **51**, 1–15 (1990).
32. T. Arai and W. Norde, *Colloids Surf.* **51**, 17–28 (1990).
33. H. Shirahama, J. Lyklema, and W. Norde, *J. Colloid Interface Sci.* **139**, 177–187 (1990).
34. A.-P. Wei, J.N. Herron, and J.D. Andrade, in D.J.A. Crommelin and H. Schellekens eds. *From Clone to Clinic*, Kluwer, Amsterdam, 1990, pp. 305–313.
35. D. Horsley, J. Herron, V. Hlady, and J.D. Andrade, in J.L. Brash and T.A. Horbett eds. *Proteins at Interfaces: Physicochemical and Biochemical Studies*, ACS Symp. Series 343, American Chemical Society, Washington, D.C., 1987, p. 290.
36. S. Xu and S. Damodaran, *J. Colloid Interface Sci.* **159**, 124–133 (1993).
37. A. Kato and K. Yutani, *Protein Eng.* **2**, 153–156 (1988).
38. J. Wang and J. McGuire, *J. Colloid Interface Sci.* **185**, 317–323 (1997).
39. C.K. Bower, Q. Xu, and J. McGuire, *Biotechnol. Bioeng.*, **58**, 658–662 (1998).
40. H. Elwing, S. Welin, A. Askendal, U. Nilsson, and I. Lundström, *J. Colloid Interface Sci.* **119**, 203–210 (1987).
41. H. Elwing, A. Askendal, and I. Lundström, *Prog. Colloid Polym. Sci.* **74**, 103–107 (1987).
42. R.J. Rapoza and T.A. Horbett, *J. Colloid Interface Sci.* **136**, 480–493 (1990).
43. R.J. Rapoza and T.A. Horbett, *J. Biomed. Mater. Res.* **24**, 1263–1282 (1990).
44. S.I. Ertel, B.D. Ratner, and T.A. Horbett, *J. Colloid Interface Sci.* **147**, 433–442 (1991).
45. M.C. Wahlgren and T. Arnebrant, *J. Colloid Interface Sci.* **142**, 503–511 (1991).
46. M.C. Wahlgren and T. Arnebrant, *J. Colloid Interface Sci.* **148**, 201–206 (1992).
47. M.C. Wahlgren, M.A. Paulsson, and T. Arnebrant, *Colloids Surf., A* **70**, 139–149 (1993).
48. M.C. Wahlgren, T. Arnebrant, A. Askendal, and S. Welin-Klinström, *Colloids Surf., A* **70**, 151–158 (1993).
49. V. Krisdhasima, P. Vinaraphong, and J. McGuire, *J. Colloid Interface Sci.* **161**, 325–334 (1993).
50. P. Vinaraphong, V. Krisdhasima, and J. McGuire, *J. Colloid Interface Sci.* **174**, 351–360 (1995).
51. J. McGuire, M.C. Wahlgren, and T. Arnebrant, *J. Colloid Interface Sci.* **170**, 193–202 (1995).
52. S.M. Slack and T.A. Horbett, *J. Colloid Interface Sci.* **133**, 148–165 (1989).
53. G. Narsimhan and F. Uraizee, *Biotechnol. Prog.* **8**, 187–196 (992).
54. V. Krisdhasima, J. McGuire, and R. Sproull, *J. Colloid Interface Sci.* **154**, 337–350 (1992).

55. C.F. Lu, A. Nadarajah, and K.K. Chitter, *J. Colloid Interface Sci.* **168**, 152–161 (1994).
56. M.C. Wahlgren, T. Arnebrant, and I. Lundström, *J. Colloid Interface Sci.* **175**, 506–514 (1995).
57. P. Dejardin, I. Cottin, *Colloids Surf., B* **4**, 111–120 (1995).
58. I. Lundström, *Prog. Colloid Polym. Sci.* **70**, 76–82 (1985).
59. M.A. Daeschel and J. McGuire, *Biotechnol. Genet. Eng. Rev.*, **15**, 413–438 (1998).

See also ADSORBENTS, INORGANIC; ADSORPTION, EXPANDED BED; ADSORPTION, PROTEIN, BATCH; CHROMATOGRAPHY, ION EXCHANGE; CHROMATOGRAPHY, SIZE EXCLUSION.

AFFINITY FUSIONS, GENE EXPRESSION

STEFAN STÅHL
JOAKIM NILSSON
SOPHIA HOBER
MATHIAS UHLÉN
PER-ÅKE NYGREN
Royal Institute of Technology
Stockholm, Sweden

KEY WORDS

Affibody
Affinity purification
Affinity tag
Downstream processing
Expression system
Functional cDNA analysis
IgG binding
Immunodetection
In vivo folding
Phage display
Renaturation
Serum albumin
Stabilization
Staphylococcal protein A
Streptococcal protein G
Vaccine development

OUTLINE

INTRODUCTION

The strong, specific, and well-characterized interaction between the Fc part of immunoglobulin G (IgG) and *Staphylococcus aureus* protein A (SpA) was the first biointeraction exploited for the creation of an affinity fusion system allowing affinity chromatography purification of expressed recombinant gene products (1). To date, hundreds of proteins have been produced as SpA fusions in a wide variety of host cells (2–5). During the past decade a large number of additional affinity fusion systems have been presented (for reviews, see 2,6–9). Affinity fusions are primarily used to simplify recovery of a fused target gene product but are also widely employed for detection and immobilization purposes (10). Most affinity tags are easy to use as detection moieties to monitor the production or distribution of a produced fusion protein, and certain affinity tags are particularly suited to obtain efficient and directed immobilization of a target protein.

This article describes the use of two affinity tags, the IgG-binding SpA domains and the serum albumin–binding region of streptococcal protein G (SpG), which both carry a number of features that make them highly suited as fusion partners in different applications. The wide spectrum of applications for affinity fusions in biotechnology is illustrated by selected examples in which SpA or SpG fusion proteins have been used to *(1)* improve the stability to proteolysis of produced recombinant proteins, *(2)* allow integrated downstream processing schemes by careful genetic design, *(3)* improve in vitro renaturation schemes for expressed gene products, *(4)* enable affinity-assisted folding in vivo of target proteins, *(5)* prolong the in vivo half-life of therapeutic proteins, *(6)* facilitate subunit vaccine development and functional cDNA analysis, and *(7)* select novel receptor variants with new specificities by the use of phage display technology.

COMMONLY USED AFFINITY TAGS

A large number of different affinity fusion systems have been described involving fusion partners that range in size from one amino acid to large multisubunit proteins capable of selective interaction with a ligand immobilized onto a chromatography matrix. Different types of interactions have been utilized, such as enzymes–substrates, bacterial receptors–serum proteins, polyhistidines–metal ions, and antibodies–antigens (2). The degree of specificity in the interaction and conditions for the affinity purification differ from system to system. The environment tolerable by the target protein is an important factor for deciding which affinity fusion system to choose. Also, other factors, including costs for the affinity matrix and buffers and the possibilities for column sanitation and reuse, are important to consider. Numerous gene fusion systems for affinity purification have been described in the literature, all with different characteristics (2,6–9). Some of the most commonly used systems are listed in Table 1, and some relevant references are given for each system. These different affinity fusion systems, with their advantages and disadvantages, were described in a recent review (10). The polyhistidyl tags, which suffer from a somewhat lower specificity, have the advantage of allowing affinity purification under denaturing conditions. The in vivo biotinylated "Bio-tag" is particularly suited for immobilization purposes because of the strong interaction (association constant; $K_A \approx 10^{15}$ M^{-1}) between biotin and streptavidin, but is inconvenient to use for affinity purification purposes because efficient elution is difficult to accomplish. One of the most extensively used affinity tags is the glutathione S-transferase (GST) from the parasitic helminth *Shistosoma japonicum* (17). GST fusion proteins can be purified using glutathione as immobilized ligand, also following renaturation from inclusion bodies after solubilization with 6 M guanidinium hydrochloride (18). A possible complication associated with the GST fusion system is the use of reduced glutathione (a reducing agent) for elution, which can affect target proteins containing disulfides (10).

Although a multitude of systems have been described, no single affinity fusion strategy is ideal for all expression or purification situations. For example, if secretion of the product into the periplasm or culture medium is desired, fusion to affinity tails derived from normally intracellular proteins (e.g., the β-galactosidase system) is not applicable, and if the purification must be performed under denaturating conditions, protein ligands such as monoclonal antibodies (mAbs; e.g., anti-FLAG M1 and M2 mAbs) are unsuitable. Therefore, to obtain general expression vectors for affinity gene fusion strategies applicable in several situations, a combination of affinity fusion domains could be introduced into a single fusion partner. Such composite fusion partners, consisting of several independent affinity domains, could potentially be used in different detection, purification, and immobilization situations, employing the affinity function that is most suitable for the situation. However, the included affinity domains should be carefully chosen not to interfere functionally with each other, and each moiety should be able to withstand the purification conditions dictated by the affinity domain requiring the harshest affinity chromatography conditions. A further advantage using combined affinity tags is that different methods for the detection and immobilization of the fusion protein also can be envisioned.

Recently, a new affinity fusion partner was described in which this strategy was utilized (26). An affinity-tag combination was designed consisting of the in vivo biotinylated peptide (Bio), a hexahistidyl peptide (His_6), and the SpG-derived albumin-binding protein (ABP) that could allow for flexible use in different detection, purification, and immobilization situations (26). Fusions to the Bio–His_6–ABP affinity tag could be used for easy detection using commercial streptavidin conjugates, for purification under both native and denaturating conditions, and for directed immobilization purposes with an option to choose among three strong affinity interactions. In addition, the ABP part of the affinity tag is highly soluble and effectively refolds after denaturation, which can be an advantage during refolding of the target protein fused to the Bio–His_6–ABP affinity tag. Such composite affinity tags have the potential of becoming widely used in the future because they offer flexibility.

This article does not focus on the description of different characteristics of various affinity-tag systems, as up-to-date reviews are available (4,9,10), but instead describes various application areas in which affinity fusions to the

Table 1. Commonly Used Affinity Fusion Systems

Fusion partner	Size	Ligand	Supplier	Refs.
Protein A	31 kDa	hIgG	Pharmacia Biotech	5,11
Z	7 kDa	hIgG	Pharmacia Biotech	12,13
ABP	5–25 kDa	HSA		5,14
Poly His	≈1 kDa	Me^{2+} chelator	Novagen, Qiagen	15,16
GST	25 kDa	Glutathione	Pharmacia Biotech	17,18
MBP	41 kDa	Amylose	New England Biolabs	19,20
FLAG	1 kDa	MAbs M1/M2	Kodak	21,22
PinPoint[a]	13 kDa	Streptavidin/avidin	Promega	23,24
Bio[b]	13 aa	Streptavidin/avidin		25,26

Note: aa, amino acids; ABP, albumin-binding protein; GST, glutathione S-transferase; hIgG, human immunoglobulin G; HSA, human serum albumin; MAb, monoclonal antibody; MBP, maltose-binding protein.

[a]A 13-kDA subunit of the transcarboxylase complex from *Propionibacterium shermanii*, which is biotinylated in vivo by the *E. coli* cells.

[b]A 13-amino acid peptide that is selected from a combinatorial library and found to be biotinylated in vivo.

IgG-binding domains of SpA and the serum albumin-binding regions of SpG have been used.

PROTEIN A AND PROTEIN G AFFINITY FUSION SYSTEMS

SpA (Fig. 1a) is an immunoglobulin-binding surface receptor found in the Gram-positive bacterium *Staphylococcus aureus*. SpA has found extensive use in immunological and biotechnological research (2–4,11,27,28). SpA binds to the constant (Fc) part of certain immunoglobulins, but the exact biological significance of this property is not clarified (29). Functional and structural analysis of SpA has revealed the presence of a signal peptide, S, which is processed during secretion, five highly homologous domains (E, D, A, B and C) all capable of binding to IgG (30), followed by a cell wall–attaching structure designated XM (31–35) (Fig. 1a). Here, X represents a charged and repetitive region, postulated to interact with the peptidoglycan cell wall (31), whereas M is a region common for Gram-positive cell surface-bound receptors containing an $LPX_{aa}TG$ motif, linked to a C-terminal hydrophobic region ending with a charged tail (32–35). It has been demonstrated that the complete M region is required for cell surface anchoring and that the cell wall sorting is accompanied by proteolytic cleavage at the SpA C-terminus and subsequent covalent linking of the surface receptor to the cell wall (32–35). SpA binds to IgG from most mammalian species, including humans. Of the four subclasses of human IgG, SpA binds to IgG1, IgG2, and IgG4 but shows only weak interaction with IgG3 (36).

Several characteristics of the IgG-binding domains of SpA have made them suitable as fusion partners for the production and purification of recombinant proteins: *(1)* SpA is proteolytically stable. *(2)* The structure of the IgG-binding domains, each being a three-helix bundle (37–39), appears to be favorable for independent folding of the fusion partner and the fused product because the N- and C-termini of each IgG-binding domain are solvent exposed (38,40). *(3)* It has been demonstrated to be feasible to introduce sequences accessible for site-specific cleavage of SpA fusion proteins so as to release the target gene product (41). *(4)* The high solubility of SpA enables the production

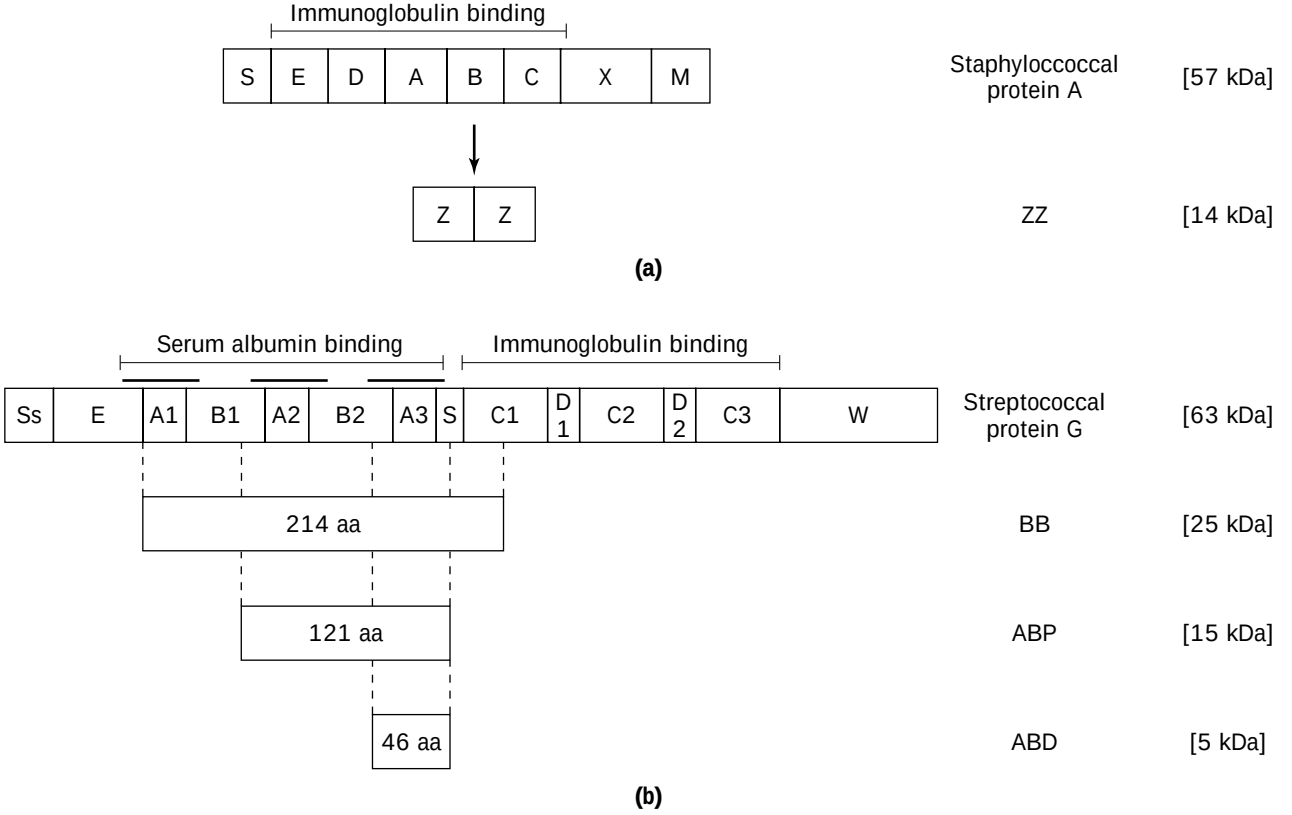


Figure 1. Schematic presentation of staphylococcal protein A (SpA) and streptococcal protein G (SpG). **(a)** The complete SpA gene product and the 7-kDa B-domain analog Z, represented in the most widely used divalent form. **(b)** The SpG gene product, with the separate albumin- and immunoglobulin-binding regions indicated, and a presentation of some of the most commonly used albumin-binding affinity tags. BB (25 kDa) was the first isolated serum albumin–binding affinity tag (14). More recently, the new tags ABP (15 kDa) and ABD (5 kDa), containing two or one serum albumin–binding motifs, respectively, were designed on the basis of postulated borders for the albumin-binding protein domains (13,47). The three postulated minimal albumin-binding motifs are indicated as *bars* above the SpG.

of soluble fusion proteins to very high concentrations within the *Escherichia coli* cell (42,43).

To create a "second-generation" affinity tag, an engineered IgG-binding domain, Z, was designed based on domain B of SpA (Fig. 1a) (12). This domain naturally lacks methionines, making it resistant to CNBr cleavage. In addition, an NG dipeptide sequence present in all native domains was changed to NA by altering the glycine codon to a codon for alanine, rendering the Z-domain also resistant to cleavage by hydroxylamine (12). Analysis of the interaction between IgG and Z domains, polymerized to different multiplicities, demonstrated that a dimeric form, ZZ (Fig. 1a), was the optimal fusion partner by its strong IgG-binding (44) and efficient secretion (12).

A different affinity-tag system, which has a number of features in common with the SpA system but also some additional advantageous properties, is based on the serum albumin-binding region of streptococcal protein G (SpG; Fig. 1b) (14,45). The regions of SpG mediating binding activities to serum albumin and IgG have been shown to be structurally separated (Fig. 1b) (14). Figure 1b illustrates SpG schematically and presents some existing variants of albumin-binding affinity tags. SpG binds to serum albumins of different mammalian species, including humans, mouse, and rat (45). The complete region comprises three serum albumin-binding motifs (46), each being approximately 46 amino acids in size. One of these postulated minimal motifs has been produced and tested for its serum albumin binding (47) and is considered to be a suitable affinity tag in a mono- or divalent fashion. However, tags comprising one, two, two-and-a-half, or three serum albumin-binding motifs (Fig. 1b) have all been successfully produced and used in different applications (4,13,27,47–50). The affinity tags, denoted BB, ABP (for albumin-binding protein), and ABD (for albumin-binding domain) (Fig. 1b), are proteolytically stable, highly soluble, and possible to produce at high yields (13,26,48,51). The 46–amino acid albumin-binding domain (ABD) was recently subjected to a heteronuclear nuclear magnetic resonance (NMR) analysis, concluding that it constitutes a three-helix bundle (52), remarkably similar to that of the IgG-binding domains of SpA.

Host Cell Systems, Affinity Resins, and Detection Systems

To date, a multitude of proteins have been produced as fusions to the IgG-binding domains of SpA (native domains or the engineered Z-domains; see following), in host cells from all five kingdoms of living organisms: *(1)* Gram-negative bacteria, such as *E. coli* (more than 100 published examples, including Refs. 13,28,53–55) and *Salmonella typhimurium* (56,57), as well as Gram-positive bacteria such as *Staphylococcus aureus* (58,59), *Streptococcus sanguis* (60), and *Bacillus subtilis* (61); *(2)* the yeast *Saccharomyces cerevisiae* (62–66); *(3)* plants (67); *(4)* insect cells (68,69); and *(5)* mammalian cells (70–75).

The SpG-derived expression systems have found several applications (as are described later), some completely new, and some as a complement to the SpA/ZZ systems. The albumin-binding tags have not yet been evaluated for production in all host organisms but have been successfully expressed for different purposes in Gram-negative bacteria such as *E. coli* (13,14,26,47,51), Gram-positive bacteria such as *Staphylococcus xylosus* and *Staphylococcus carnosus* (48,76–80), and mammalian cells such as Chinese hamster ovary (CHO) cells (81).

The SpA fusion proteins, which bind to the Fc region of IgG, are commonly affinity purified using IgG-Sepharose (50,82) and the serum albumin-binding SpG fusions are normally purified on human serum albumin (HSA) Sepharose (14,50). However, other matrices, such as protein-coated paramagnetic beads, have also been utilized (83). SpA (ZZ) and SpG (BB, ABP, or ABD) fusion proteins can be easily detected during expression, without the need of specific antisera, by blotting techniques after polyacrylamide gel electrophoresis (PAGE) by taking advantage of their IgG- (Fc) and HSA-binding properties, respectively. SpA fusions are efficiently stained using a commercially available complex of rabbit anti-horseradish peroxidase–IgG and horseradish peroxidase (13). SpG fusions are stained using biotinylated HSA and conjugated streptavidin–alkaline phosphatase (13).

Expression Vector Systems for *E. coli*

A number of expression vectors based on gene fusions to the SpA gene, often in the form of the divalent SpA analogue ZZ (12), or all five native IgG-binding domains, have been developed for recombinant protein production in *E. coli* (11,82,84). The promoter and secretion signals of SpA are functional also in the Gram-negative *E. coli* (1). Protein A fusions expressed in *E. coli* can thus be efficiently secreted to the periplasm of the bacteria, and in many cases also to the culture medium (50,53,55,85,86) from which they can be easily purified by IgG affinity chromatography (86,87). In the vectors designed for secretion, the expressed SpA-fusion products are in most examples transcribed from the SpA promoter, which is constitutive and efficiently recognized in *E. coli* (27,50).

Analogous expression vectors for production of secreted fusions also to the serum albumin–binding regions of SpG have been developed (14,50), taking advantage of the same SpA promoter and secretion signals. Fusion proteins produced in this way can be efficiently affinity purified using human serum albumin (HSA) columns (14).

There exists no general strategy applicable to all recombinant gene products to guarantee secretion. Instead, the inherent properties of the target protein largely dictate if a secretion route is possible to use. However, a large fraction of the proteins produced as SpA fusions are soluble and efficiently secreted to the periplasmic space and culture medium of *E. coli* (2,55,86). This may at least be partially caused by the high solubility of the ZZ domains, which most likely increases the overall solubility of the fusion proteins. There are a number of advantages connected with a secretion strategy. First, the gene product will be less exposed to cytoplasmic proteases, which might enable production of also labile proteins (88). Second, disulfide bond formation, which is occurring in the nonreducing environment outside the cytoplasm, could improve on the folding of certain gene products (86). Third, the recovery of the recombinant protein is simplified because a large de-

gree of the purification from host cell proteins has been achieved through the secretion (55,86). Hansson and coworkers (55) described the *E. coli* production of a fusion protein ZZ-M5, a malaria subunit vaccine candidate, using a secretion strategy. More than 65% of the recombinant gene product was found to be secreted to the culture medium, from which it could be recovered in a single step by expanded-bed anion exchange adsorption. As the ZZ-M5 protein was to be used in a preclinical malaria vaccine trial in *Aotus* monkeys (89), a polishing step was included in which the IgG-binding capacity of the fusion protein was employed. After affinity chromatography on IgG-Sepharose, contaminating DNA and endotoxin levels were well below the demands set by regulatory authorities. The overall yield of the process, performed in pilot scale, exceeded 90%, resulting in 550 mg product per liter of culture (55).

As mentioned earlier, there is no guarantee that a certain protein can be secreted even if an expression system with secretion signals is chosen for the production. Proteins with a strong tendency to precipitate intracellularly and proteins containing hydrophobic transmembrane regions may be extremely difficult to secrete (27,80). Intracellular expression of such proteins could be a more attractive alternative because of recent advances in in vitro renaturation of recombinant proteins from intracellular precipitates, that is, inclusion bodies (90). The mechanism for inclusion body formation is not fully elucidated, but expression driven from exceptionally strong promoters such as the T7 promoter (91), leading to very high expression levels, seems to increase the tendency for intracellular aggregation of the produced recombinant protein (90). Production by the inclusion body strategy has the main advantages that the recombinant product normally is protected from proteolysis and that it can be produced in large quantities. Levels up to 50% of total cell protein content have been reported (90).

Vectors for intracellular production of SpA fusions have thus been constructed in which the transcription is under the control of inducible promoters, such as *lac*UV5 (11), λ_{PR} (11), *trp* (92,93), *trc* (94), or T7 (13), enabling a more controlled expression. Also, for intracellular production of BB and ABP fusion proteins, inducible vector systems have been used, employing *trc* (94), *trp* (51,93), or T7 (13) promoters, respectively. The most robust and in addition the most tightly controlled of these systems seems to be the T7 RNA polymerase–regulated promoter system (91). Using this system, Larsson and coworkers (13) recently described the expression of five different mouse cDNAs as fusions to ZZ and ABP, respectively. Expression levels ranged from 4 to 500 mg gene product per liter of culture.

Intracellularly produced and still soluble fusion proteins can be released by sonication or high-pressure homogenization of cells before purification by IgG or HSA affinity. However, precipitated gene products need to be solubilized before the affinity purification procedure. It was recently demonstrated that affinity purification indeed can be useful also for the recovery of proteins with a strong tendency to aggregate during renaturation from inclusion bodies following standard protocols (93). Murby and coworkers (93) devised a recovery scheme for the production of such ZZ and BB fusions, applied on various fragments of the fusion glycoprotein (F) from the human respiratory syncytial virus (RSV). Earlier attempts to produce these labile and precipitation-prone polypeptides in *E. coli* had failed, but several different F fusion proteins could by this novel strategy be produced and recovered as full-length products with substantial yields (20–50 mg/L). Because it was demonstrated that the IgG and HSA affinity ligands both were resistant to 0.5 M guanidine hydrochloride, efficient recovery from inclusion bodies of the ZZ-F and BB-F fusions could be achieved by affinity chromatography in the presence of the chaotropic agent throughout the purification process (93). The described strategy has so far been successfully evaluated for a number of proteins of low solubility and should be of interest for efficient recovery of other heterologous proteins that form inclusion bodies when expressed in a bacterial host.

It is, of course, almost impossible to give general information concerning the production levels for the described expression systems because production levels are strongly dependent on inherent properties of the target gene products. It is however evident that both the secretion strategy and intracellular protein production in *E. coli* using the SpA and SpG expression systems can give high expression levels. For the secreted production, expression levels of 0.5 to 0.8 g product per liter of culture have been reported (55,95), and for intracellular production levels up to 3 g gene product per liter of culture have been obtained (96). The SpA and SpG expression systems are thus for many gene products a good choice because high expression levels can be obtained and efficient affinity purification of produced fusion proteins can be easily achieved.

Mild Elution of Protein A and Protein G Fusions by Competition

One possible drawback associated with the SpA and SpG affinity purification systems has been the need for elution by low pH from the affinity column. The elution of fusion proteins at pH 3, used routinely, has for some products been destructive, yielding biologically inactive products. By a recently presented concept, the low-pH elution can be circumvented by competitive elution based on an engineered competitor protein that can be efficiently removed from the eluate mix after elution. The principle for the competitive elution strategy was described by Nilsson and coworkers (47) (Fig. 2). The target protein in the presented examples was produced as a fusion to a monovalent Z-domain of SpA. A sample (e.g., crude cell lysate) containing the recombinant fusion Z-target is passed through an IgG-Sepharose column. The recombinant fusion protein is thus captured, enabling extensive washing to remove contaminants. A stoichiometric excess of an engineered bifunctional competitor fusion protein ZZ-BB (Fig. 2) was added in which BB constitutes a second affinity tail capable of selective binding (association constant; $K_A \approx 1.4 \times 10^9$ M^{-1}) to HSA. The competitor ZZ-BB, having an affinity to human IgG more than 10-fold higher than that to the monovalent Z, is used for competitive elution of the Z-fusion at neutral pH. Specific removal of ZZ-BB from the effluent mix is accomplished by a passage through a second affinity column (HSA) to which the ZZ-BB fusion is bound.

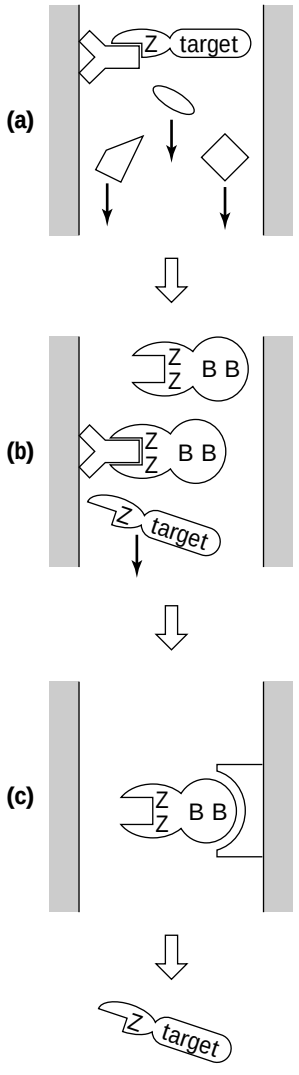


Figure 2. The concept of competitive elution, here used for elution of a monovalent Z-fusion by the competitor ZZ-BB. *Source*: Modified from Nisson and coworkers (47).

The described strategy has been used to produce the Klenow fragment of *E. coli* DNA polymerase I as a Z-Klenow fusion, and the recovered product retained full enzymatic activity (47).

Stabilization to Proteolysis of Labile Gene Products

When expressing mammalian proteins in a heterologous host such as *E. coli*, proteolytic degradation is one of the more common problems encountered. A review describing the different strategies available to minimize proteolytic degradation of recombinant gene products was recently presented (97). Fusion of a proteolytically labile target protein to stable fusion partners, such as the ZZ and ABP tags, has been shown to have a certain, but limited, stabilizing effect (88,97). However, a much more pronounced stabilizing effect has been obtained when employing a dual affinity fusion concept (Fig. 3) (98), where the protein of interest, X, is fused between the two different affinity tag

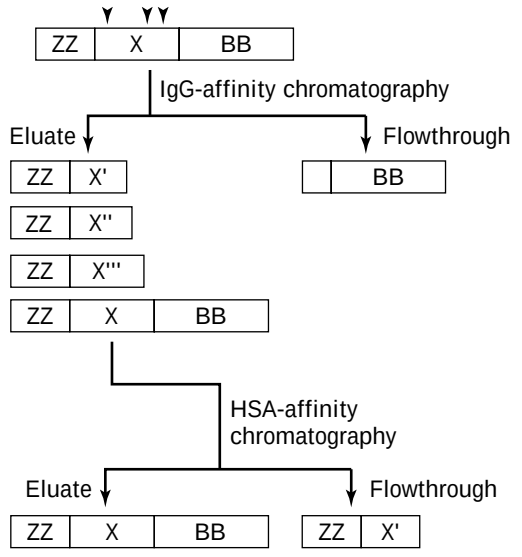


Figure 3. The concept of dual-affinity fusions for recovery of proteolytically labile target proteins. The gene encoding the target protein X is fused between the genes encoding the ZZ and BB tags, respectively. A first passage over an IgG column results in recovery of the full-length product as well as potential degradation products. A second passage over an HSA column results in recovery of only the full-length product.

proteins, ZZ and BB. The concept was applied to the production of proteolytically sensitive human insulin-like growth factor II (IGF-II) in *E. coli*. The concept first of all allows the use of two consecutive affinity chromatography steps, but in addition, dramatic stabilization effects can be obtained as compared with more conventional expression routes. The dual affinity fusion approach was found to have a stabilizing effect on three small and labile recombinant proteins, namely human proinsulin, a domain from rat protein disulfide isomerase, and a fragment from a human T-cell receptor (88), when expressed in *E. coli*. The concept of dual affinity fusions has thereafter been employed also with affinity tags other than the SpA and SpG tags (99–102).

Strategies for Integrated Downstream Processing

An important part of modern biotechnology is to develop simplified schemes for fermentation and downstream processing of recombinant proteins by the integration of unit operations. Genetic product design can be used in several ways to obtain such schemes, that is, by influencing the yield and localization of the product as well as to adapt the gene product by fusion technology to specific unit operations suitable for large-scale downstream processing.

The recently presented expanded-bed adsorption procedure (mentioned earlier) constitutes a nice example of process integration (55). Efficient recovery of a secreted recombinant fusion protein ZZ-M5 was achieved directly from a crude fermentor broth without prior cell removal. The fusion protein had a relatively low isoelectric point (pI), which allowed anionic exchange adsorption at pH 5.5 at which most *E. coli* host proteins are not adsorbed. This

strategy allowed an integration of the cell separation step with ion exchange adsorption of the gene product with simultaneous volume reduction, which resulted in a highly condensed but still efficient recovery process (55). The two-step purification process, ion exchange chromatography in an expanded-bed format followed by IgG affinity chromatography for polishing, demonstrated an overall yield of more than 90%.

An alternative strategy to simplify downstream processing was investigated by Köhler and coworkers (92), where recombinant technology was applied to alter the partitioning properties of a ZZ-fusion protein, which thereby could be efficiently recovered by aqueous two-phase extraction as the primary purification step. An attractive approach to simplify recovery of products, which may be produced with modified primary sequences, e.g., industrial enzymes, would be to alter the partitioning properties by substitution of one or more surface amino acid residues, which are not involved in the biological activity, for tryptophans, which are known to influence partitioning characteristics. This type of protein engineering will most probably be used in the future to reduce production costs for bulk proteins.

A third example of how unit operations can be integrated by careful design of the gene product and the recovery process was recently demonstrated for the production of native human insulin and its C-peptide in *E. coli* (96,103). Human proinsulin was genetically fused to ZZ, and recovered from inclusion bodies via solubilization and refolding of the intact fusion protein, followed by IgG affinity chromatography. Native insulin, consisting of cystein-bridged A- and B-chains, as well as the clinically relevant C-peptide (104), could be released by a single-step trypsin treatment, cleaving off the ZZ-tail simultaneously to the processing of the C-peptide (96,103). The careful genetic design of a fusion protein as an intermediate product, which could be efficiently converted to the final products, allowed the design of a pilot-scale process involving few unit operations (96) that thus might result in a cost-efficient bioprocess. A further improved bioprocess for production of the proinsulin C-peptide was presented (105) in which the C-peptide-encoding gene fragment was heptamerized to increase yields. A fusion protein containing seven copies of the C-peptide was thus produced, and after affinity purification, enzymatic digestion was employed to release native C-peptide (105).

Improved Renaturation Schemes for Recombinant Gene Products

The recovery of biologically active or native proteins by in vitro refolding from insoluble inclusion bodies can in some cases be hampered by the aggregation of the product during the procedure, leading to low overall yields. To make the procedure more efficient, several improved protocols have been developed, including the addition of different "folding enhancers" (4). Alternatively, the protein itself can be engineered to facilitate the refolding. The presence of hydrophilic peptide extensions during the refolding can dramatically improve the folding yield, probably by conferring a higher overall solubility of the protein (90). Samuelsson and coworkers (42) showed that the reshuffling of misfolded disulfides in recombinant insulin-like growth factor I (IGF-I) was greatly facilitated by fusion to the highly soluble ZZ fusion partner (42). Compared to unfused IGF-I, the fusion ZZ-IGF-I could be successfully refolded at a 100-fold higher concentration (1–2 mg/mL), without forming precipitates (43).

Affinity-Assisted In Vivo Folding

A completely new concept to improve the fraction of correctly folded recombinant IGF-I was recently presented (106). It was demonstrated that coexpression of a specific binding protein, IGF binding protein 1 (IGFBP-1), significantly increased the relative yield of IGF-I having native disulfide bridges when expressed in a secreted form in *E. coli*. In addition, a glutathione redox buffer was added to the growth medium to enhance formation and breakage of disulfide bonds in the periplasm of the bacteria. In the presented example, both IGF-I and IGFBP-1 were produced as affinity fusions, to Z and BB (Fig. 4a), respectively, which facilitated the purification of *in vivo*–assembled heterodimers by alternative purification methods (106). A further development of the strategy would be to express the target protein as a nonfused gene product and the specific binding protein in an affinity-tagged configuration (Fig. 4b). Correctly folded target protein would thus be affinity captured as a heterodimer via the tagged binding protein. This would employ the benefits of high specificity for affinity chromatography without introducing requirements of proteolytic processing (see following) to recover the native target protein.

Affinity-Tagged Proteases for Site-Specific Cleavage of Fusion Proteins

When gene fusion strategies are used for the production of native proteins, efficient means for site-specific cleavage of the fusion protein and subsequent removal of the affinity fusion partner are needed. In addition, if employing enzymatic strategies for cleavage, strategies for the removal of the proteolytic enzyme itself must be developed. Special considerations must be taken if the target protein itself also has to be further processed to give the desired final product (96,103). Careful upstream design of the fusion protein construct using genetic strategies can greatly facilitate the subsequent purification of the target protein and also allow for integrated systems involving coprocessing of the protein and efficient removal of the affinity fusion partner as well as the protease used for cleavage (96,102).

Enzymes for use in biotechnological applications should preferably be highly specific proteases, easy to produce by recombinant means in large scale. Interesting candidate enzymes can be found in human picorna viruses, whose maturation relies on the site-specific cleavage of a large polyprotein precursor to yield the viral components. Some of these proteases are functionally produced at high levels in bacterial expression systems (107,108), and this could facilitate the production of variants constructed by protein engineering. Recently, a new general strategy was described where the 3C protease of rhinovirus was fused to

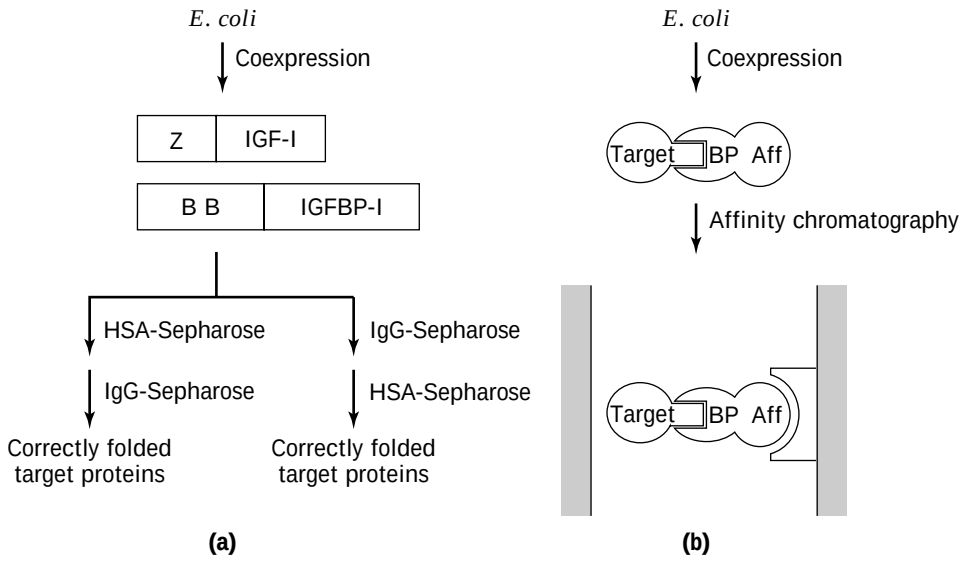


Figure 4. Affinity-assisted in vivo folding and its use for recovery of correctly folded gene products. **(a)** Schematic description of the example presented by Samuelsson and coworkers (106) where a fusion protein Z-IGF-I, with correctly folded IGF-I, could be recovered using alternative affinity chromatography strategies, taking advantage of the interaction between IGF-I and IGFBP-I. **(b)** The basic principle might be suitable for general use both to increase the yield of correctly folded target protein and to achieve recovery of only correctly folded target protein via the interaction of a specific binding protein (*BP*), which binds only to the correctly folded form of the target protein. If the binding protein is expressed as an affinity fusion (*Aff*), this fusion protein should be a suitable ligand for recovery of correctly folded target protein.

an affinity tag and subsequently used for specific cleavage of a target protein (108). This enabled removal of the affinity-tagged protease and the cleaved affinity fusion partner by simple affinity chromatography steps. It was in addition possible to add any desired amount of protease to ensure efficient cleavage because the protease could be easily removed after cleavage (108). Recently, a different affinity-tagged protease, consisting of coxsackievirus 3C ($3C^{pro}$) fused to the serum albumin–binding ABP at the N-terminus and His_6 at the C-terminus, was produced in *E. coli* and used for the production of a truncated *Taq* DNA polymerase (*ΔTaq*) (102). The heat-stable polymerase was produced as an ABP-*ΔTaq* fusion protein having a $3C^{pro}$ cleavage site introduced between the two protein moieties. After affinity purification of ABP-*ΔTaq*, the fusion was efficiently cleaved using the affinity-tagged protease ABP-$3C^{pro}$, which allowed for subsequent recovery of nonfused, fully active *ΔTaq* via passage of the cleavage mixture over an HSA-column (102). Note that the affinity-tagged protease, ABP-$3C^{pro}$, is removed together with eventual remains of unprocessed ABP-*ΔTaq* fusion protein.

The use of affinity fusions for recombinant production of therapeutic target gene products has probably been hampered by the difficulties with efficiently cleaving off the affinity tag and thereafter removing not only the tag but also the used protease. The increasing availability of affinity-tagged, highly specific proteases for cleavage of fusion proteins to enable straightforward methods for recovery of unfused authentic gene products (102,108) will probably influence researchers to further increase their use of fusion strategies for production of target proteins.

Increased In Vivo Half-Lives for Therapeutic Gene Products

Gene fusion technology might also prove to be useful in prolonging the in vivo half-life of pharmaceutical proteins by administering the therapeutic protein as fused to a protein with extended half-life. The principle was demonstrated to be efficacious by Capon and coworkers (109), who showed that by fusing a soluble portion of CD4 (the target receptor for the human immunodeficiency virus particle) to the Fc part of IgG, the serum half-life of CD4 in rabbits increased 200-fold. Nygren and coworkers (110) described an alternative strategy where in vivo stabilization was achieved by fusion of the unstable CD4 protein to BB, leading to the formation of complexes between the fusion protein and serum albumin. The serum half-lives of different fusion proteins containing the albumin-binding receptor were tested in mice. In these studies, the half-life of the CD4 fusion in vivo was shown to be comparable to the CD4-IgG molecule described earlier. Interestingly, the albumin-binding receptor in macaque monkeys was found to have a serum half-life similar to that of serum albumin itself (110). Recently, the same strategy was employed to increase the serum half-life of human soluble complement receptor type 1 (sCR1) in rats, by fusion to different serum albumin–binding fragments of SpG (81). These examples suggest that gene fusions to BB, ABP, or the minimal binding motif ABD of SpG might constitute useful strategies to improve the in vivo stability of pharmaceutical proteins. It may also be possible to combine these approaches with new methods for drug targeting or slow drug release.

Protein A Fusions in Immunodetection Techniques

Bifunctional SpA fusion proteins have found extensive use for various analytical purposes, for example, in enzyme immunoassays. Protein A fusion proteins can after expression be efficiently purified using the IgG-binding activity, and in most assays the protein A moiety binds antibodies to be quantified and the activity of the fused gene product is monitored. Thus, the fusion protein can replace more

traditional reagents, such as secondary antibodies conjugated to an enzyme. The concept of genetically engineered conjugates in immunoassays was recently reviewed (111). The fusion between catechol dioxygenase and protein A produced in *E. coli* was used in an enzyme immunoassay to quantify antibodies (112). The principle of a bioluminescent immunoassay using the fusion between protein A and bacterial luciferase has also been demonstrated (113), and recently a bifunctional fusion between domain D of SpA and luciferase from sea-firefly was used in a sandwich enzyme-linked immunosorbent assay (ELISA) (72). Furthermore, an SpA-streptokinase fusion was shown to be equally useful as a commercial SpA-horseradish peroxidase conjugate for use in immunoassays (60). A fusion between SpA and LacG, 6-phospho-β-galactosidase from *S. aureus*, was shown to be able to replace secondary antibodies in an ELISA (114). In a similar approach, an SpA–alkaline phosphatase fusion was used in ELISA (115) and a fusion between maltose-binding protein and protein A was used for a solid-phase immunoassay, in which the fusion protein was immobilized to determine IgG concentrations (116).

Recently, the use of SpA-LacI and SpA-streptavidin fusion proteins in various immunodetection techniques was reviewed (3). Using sandwich techniques, DNA fragments also can be involved in various purification and detection strategies (Fig. 5). Taking advantage of the lac repressor (LacI) from *E. coli*, binding to the *lac* operator sequence, a LacI-SpA fusion protein was produced (117) for simplified analysis of polymerase chain reaction (PCR) amplified DNA fragments generated for diagnosis of pathogens (118) (Fig. 5a). In addition, this LacI-SpA fusion protein has been used for mild reversible recovery of protein antigens or whole cells using paramagnetic beads with a coupled DNA fragment containing the *lacO* recognition sequence (119) (Fig. 5b).

An antigen detection system, termed immuno-PCR, has been described (120) in which the principle of immunodetection was combined with the sensitivity of PCR. A streptavidin-SpA chimera was used to attach a biotinylated DNA fragment to antigen–IgG complexes immobilized on microtiter plate wells (Fig. 5c). Although the method gives a sensitivity greater than any existing antigen–antibody detection system, the sensitivity of the PCR assay could have the drawback of causing false positives resulting from nonspecific binding of IgG, fusion protein, or DNA. Nevertheless, SpA fusion proteins have through the IgG Fc binding activity found extensive use in a vast number of immunodetection techniques.

Protein G Fusion Used to Accomplish "Hot-Start PCR"

Immobilization of enzymes has proven to be a useful procedure for many applications because of the altered characteristics of the immobilized enzyme as compared with the enzyme's soluble counterpart (121). One practical application utilizing such differential activities was recently described in which a heat-mediated release of an affinity-immobilized recombinant *Taq* DNA polymerase (TagTaq) was used to create a hot-start PCR procedure. *Taq* DNA polymerase was produced in *E. coli* as fused to the multifunctional Bio–His_6–ABP tag (see foregoing text; 122) for affinity purification and immobilization. HSA-affinity immobilization of the fusion employing the ABP moiety resulted in an apparent deactivation of the *Taq* DNA polymerase. However, the ABP–HSA interaction was shown to

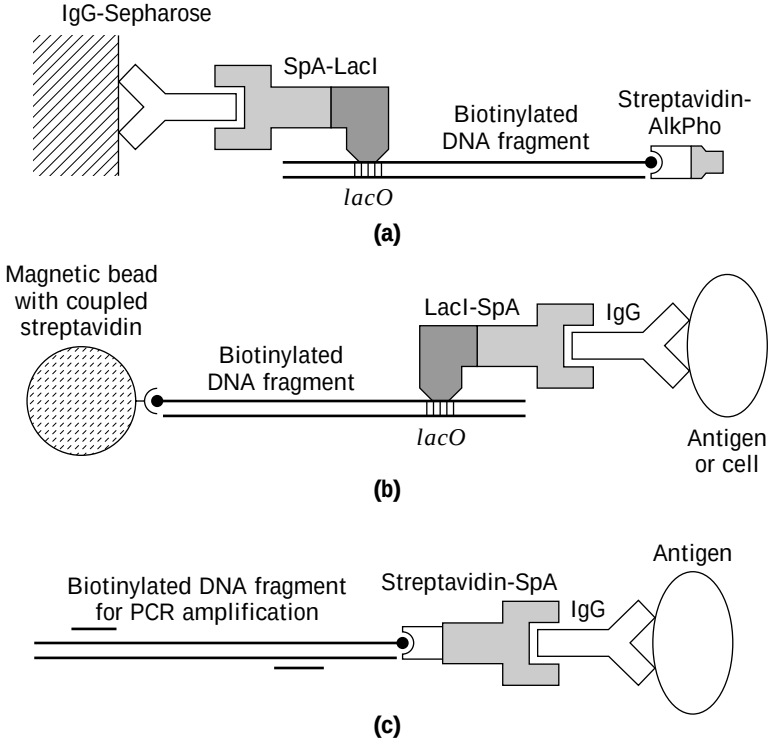


Figure 5. Schematic presentation of some sandwich concepts utilizing SpA fusion proteins and interacting DNA fragments for detection and purification purposes. **(a)** Colorimetric detection of DNA fragments containing the *lac* operator (*lacO*) sequence, using a lac repressor (LacI)-SpA fusion protein immobilized to IgG-Sepharose. The colorimetric signal was obtained using a streptavidin–alkaline phosphatate conjugate. The DNA fragments, generated by the PCR technique, introduce biotin and the *lac* operator by the PCR primers. This assay has been used to detect *Plasmodium falciparum* DNA in clinical samples (118). **(b)** This system takes advantage of the same interactions as described in **a**, with the only difference that paramagnetic beads with covalently coupled streptavidin are used as solid support. The technique has been used to purify different proteins and also whole cells using specific antibodies reversibly immobilized to the paramagnetic beads via the LacI-SpA fusion and a biotinylated DNA fragment. Mild recovery of the antigen–antibody complexes can be achieved using deoxyribonuclease (DNase) or restriction endonucleases (119). **(c)** In this system, the antigen, immobilized on a microtiter plate, is allowed to react with a monoclonal antibody, linked to a biotinylated DNA fragment via an SpA-streptavidin chimera. The detection is accomplished by PCR amplification using primers specific for the bound DNA. The generated fragment is analyzed by agarose gel electrophoresis (120).

be disrupted at elevated temperatures (>70 °C), resulting in a restored DNA polymerase activity of released *Taq* DNA polymerase fusion protein. This concept for controlling the activity of the *Taq* DNA polymerase was demonstrated to be of practical use to achieve a so-called hot-start PCR procedure, where it is desired to suppress the DNA polymerase activity until temperatures are reached at which nonspecific annealing of PCR primers is eliminated. The affinity immobilization concept for hot-start PCR was successfully used in full multiple-cycle PCR amplification and was shown to eliminate artefactual primer-dimer products in the amplification of the human oncogene K-*ras* gene in contrast to standard amplification protocols. Noteworthy, the ABP affinity fusion partner was still functional after 30 PCR cycles, allowing post-PCR reimmobilization of the *Taq* DNA polymerase fusion protein onto fresh HSA-Sepharose, facilitating the removal of the DNA polymerase from reaction mixtures after PCR (122). Further studies are needed to investigate if this concept for modulating biological activity can also be applied to other widely used enzymes, for example, reverse transcriptase.

Protein A and Protein G Fusion Proteins in Subunit Vaccine Development

Immunogenic proteins and peptides that are produced as fusion proteins have been frequently used for immunization purposes and represent a strategy to construct subunit vaccines against infectious diseases. An important aspect regarding production of antigens by recombinant DNA technology is the purification of the gene product from contaminating host components. The approach to express peptides or proteins as one part of a fusion protein, where the other part constitutes an affinity handle, allows rapid and efficient affinity purification of the gene product.

As a powerful strategy in subunit vaccine development, a dual expression system was devised, combining the two similar affinity-fusion expression systems (ZZ and BB) for the parallel expression of immunogenic peptides and proteins (Fig. 6) (50). After affinity purification, one fusion protein is used for immunization while the corresponding second fusion protein is used to analyze the induced antibody response to the fused peptide. This strategy eliminates any background originating from antibodies reactive with the affinity tag of the fusion used for immunization. In addition, the second fusion protein can, after immobilization, be used as a ligand for affinity purification of peptide-specific antibodies (Fig. 6). The dual expression system has been used to produce immunogenic structures derived from different *Plasmodium falciparum* malaria blood-stage antigens, and the expressed fusion proteins have been shown to be immunogenic in mice, rabbits, and monkeys and to induce antigen-specific antibody and T-cell responses (for review, see Ref. 27). The dual expression strategy has thus shown to be of significant value in subunit vaccine research.

Besides the use of BB and ABP fusion proteins in the dual expression concept for analysis of elicited immune responses, the serum albumin–binding region of SpG has been shown to have inherent immunopotentiating properties when used as a carrier protein genetically fused to

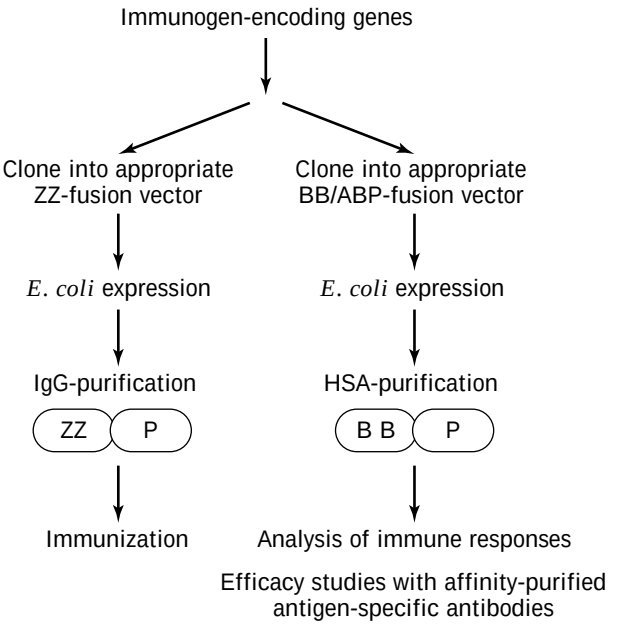


Figure 6. A flow chart representation of how the dual expression concept has been used in subunit vaccine research. Note that the BB-P fusion protein can be used both as coating antigen in various assays to analyze elicited immune responses and also in an immobilized form for enrichment of antigen-specific antibodies from the immune sera. The use of this concept has been thoroughly reviewed (27).

the immunogen used for immunization (27,49,123,124). Recently, it was demonstrated that a BB-fusion protein (BB-M3), containing a malaria peptide, induced significant antibody responses in mice strains that were nonresponders to the malaria peptide (M3) alone, suggesting that BB has the ability to provide T-cell help for antibody production (49). Furthermore, a fusion protein BB-G2N, comprising a 101–amino acid sequence from human RSV, was shown to induce protective immunity in mice to RSV challenge (123,124). It was shown that, by inclusion of the BB part, a more potent G2N-specific B-cell memory response was evoked (124). This indicates that the SpG-derived BB can function both as an affinity tag to facilitate purification and as a carrier protein with immunopotentiating properties. To date, it is not fully elucidated whether this capacity results from strong T-cell epitopes (49) or is related to the serum albumin–binding activity resulting in a prolonged exposure of the immunogen to the immune system, or is due to a combination of both.

Protein A and Protein G Fusion Proteins in Functional cDNA Analysis

The principle of dual expression systems was recently also implemented for functional analysis of cDNA-encoded proteins (Fig. 7) (13). Selected cDNAs were then expressed as ZZ fusion proteins and used for immunization to elicit antibodies, whereas the corresponding ABP fusion protein, having the cDNA-encoded portion in common, was used for purification of target protein–specific antibodies. Such

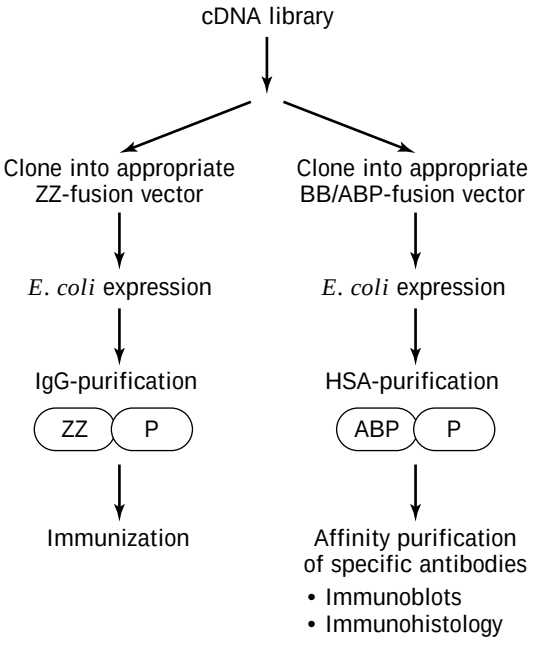


Figure 7. A flow chart representation showing how the dual expression concept has been used in functional cDNA research. The concept is thus used for the generation of highly specific antibodies employed in various analyses to assign function to cDNA-encoded proteins. Such sera have indeed shown to be useful to localize gene products to a certain cell type within a tissue section and also further down to subcellular levels (13).

highly enriched antisera should be suitable for functional analysis of the cDNA-encoded protein, for example, in Western blots and immunohistology on tissue sections or within cells. The principle was evaluated by expressing five cDNAs, isolated from prepubertal mouse testis by a differential cDNA library screening strategy. All five clones could be expressed intracellularly in *E. coli* as fusion proteins with high production levels, and affinity purification yielded essentially full-length products. Characterization of affinity-purified antibodies revealed that there existed no cross-reactivity between the two fusion systems, and that such antibodies indeed were valuable for various analyses to elucidate the functions of the cDNA-encoded proteins (13). For example, in the presented study a previously unknown cDNA, with no sequence homology to genes in the public databases, could be connected with a gene product found in nuclear bodies within spermatogonial cells (13).

To make it possible to use the dual expression system for functional analysis of cDNA, where robustness of the expression systems is crucial, completely new expression vectors were constructed. Intracellular expression was used and transcription was tightly controlled by the T7 promoter system (13). This method should ensure that a higher fraction of investigated cDNA-encoded proteins are successfully produced. The different ongoing genome projects have generated enormous amounts of partially sequenced cDNAs, and there exists a significant need for various alternative systems to take care of this information to elucidate the function of the corresponding proteins. By automation of some of the unit operations, the system described by Larsson and coworkers (13) might have the potential to make large-scale functional analysis of cDNA libraries possible.

Bacterial Surface Display

The display of heterologous proteins on the outer surface of bacteria has become an emerging topic in different fields of research within biotechnology, microbiology, and immunology (125,126). Genetic insertion of a target sequence into the genes for various outer membrane proteins has constituted the general strategy to enable secretion and subsequent surface anchoring of the recombinant target gene products. For Gram-negative bacteria, represented by *E. coli* and *Salmonella* spp., a number of different types of heterologous proteins have been surface displayed including antigenic determinants for the purpose of vaccine development, enzymes, metal-chelating histidyl peptides, and single-chain antibodies, as well as entire peptide libraries (125–128). Recently, surface display on Gram-positive bacteria has gained interest in the context of vaccine development. Approaches based on the attenuated mycobacteria (129), commensal streptococci (130), and nonpathogenic food-fermenting staphylococci have been presented (48,76–78).

To achieve surface display of various target proteins, the staphylococcal systems take advantage of the cell surface anchoring regions (XM) of SpA (see Fig. 1). The surface-exposed proteins are thus genetically fused between a functional signal peptide, responsible for the translocation from the cell, and the XM-regions of SpA. Heterologous antigenic determinants of various origin (bacteria, protozoa, and human virus) have been expressed as anchored on recombinant staphylococci, which are being investigated as a bacterial vaccine delivery system (48,76–78,80,131–133). Even active single-chain antibody fragments have been surface displayed on recombinant staphylococci (79). Such recombinant bacteria could be envisioned as inexpensive tools in diagnostic tests.

Protein A Domains and Phage Technology for Selection of "Affibodies"

An example of emerging interest in which gene fusions have made new technology possible is the display of libraries of peptides or proteins on filamentous phage surfaces. Phage particles decorated in such manner can be enriched in vitro from a background of irrelevant phages by biopanning techniques, employing an immobilized ligand capable of binding to the displayed protein. Moreover, inside its protein coat the captured phage carries the genetic information of the displayed protein, which allows later identification by DNA sequencing. Thus, the technology provides a link between genotype and phenotype and allows the selection of peptides or proteins of desired functions to be expressed on the surface of the filamentous phage. This technology, as pioneered by Smith (134), has emerged into a powerful technology to select, rather than screen, for rare phages capable of binding to a defined target molecule from libraries comprising as many as 10^{10}

phage species. Such phage libraries are obtained by the fusion to phage pIII or pVIII genes, of cassettes encoding randomized peptides (usually 6- to 10-mers), antibody repertoires, or multiple variants of a protein (135). Thus, it is possible to rapidly and efficiently select high-affinity ligands or effector molecules that can be used for analytical or preparative applications.

The Z-domain from SpA has several features including proteolytic stability, high solubility, small size, and compact and robust structure, which makes it ideal as a fusion partner for the production of recombinant proteins. The residues responsible for the IgG Fc–binding activity have been identified from the crystal structure of the complex between its parent B-domain and human IgG Fc, showing that these residues are situated on the outer surfaces of two of the helices making up the domain structure and do not contribute to the packing of the core (37). The Fc-binding surface covers an area of approximately 600 $Å^2$, similar in size to the surfaces involved in many antigen–antibody interactions. Taken together, these data suggest that if random mutagenesis of this binding surface was performed, novel domains could be obtained with the possibility of finding variants capable of binding molecules other than Fc of IgG. In addition, these domains would theoretically share the overall stability and α-helical structure of the wild-type domain.

Recently, an attempt to obtain such ligands with completely new affinities was initiated, employing phage display technology (136) (Fig. 8a). A genetic library was created in which 13 surface residues (involved in the IgG Fc binding) of the Z-domain were randomly and simultaneously substituted to any of the 20 possible amino acids. This Z-library has, after genetic fusion to the gene for coat protein 3 of filamentous phage M13, resulted in a phage library adapted for selection of novel specificities by biopanning (137). The library has been subjected to such affinity selection against different molecules for investigation as a source of novel binders (Fig. 8b). So far, novel Z-variants, so-called affibodies, have successfully been selected to diverse targets, such as *Taq* DNA polymerase, human insulin, and a human apolipoprotein variant (137). This concept suggests that the described strategy might prove to be beneficial in generating a "next generation" of ligands or artificial antibodies to various target molecules,

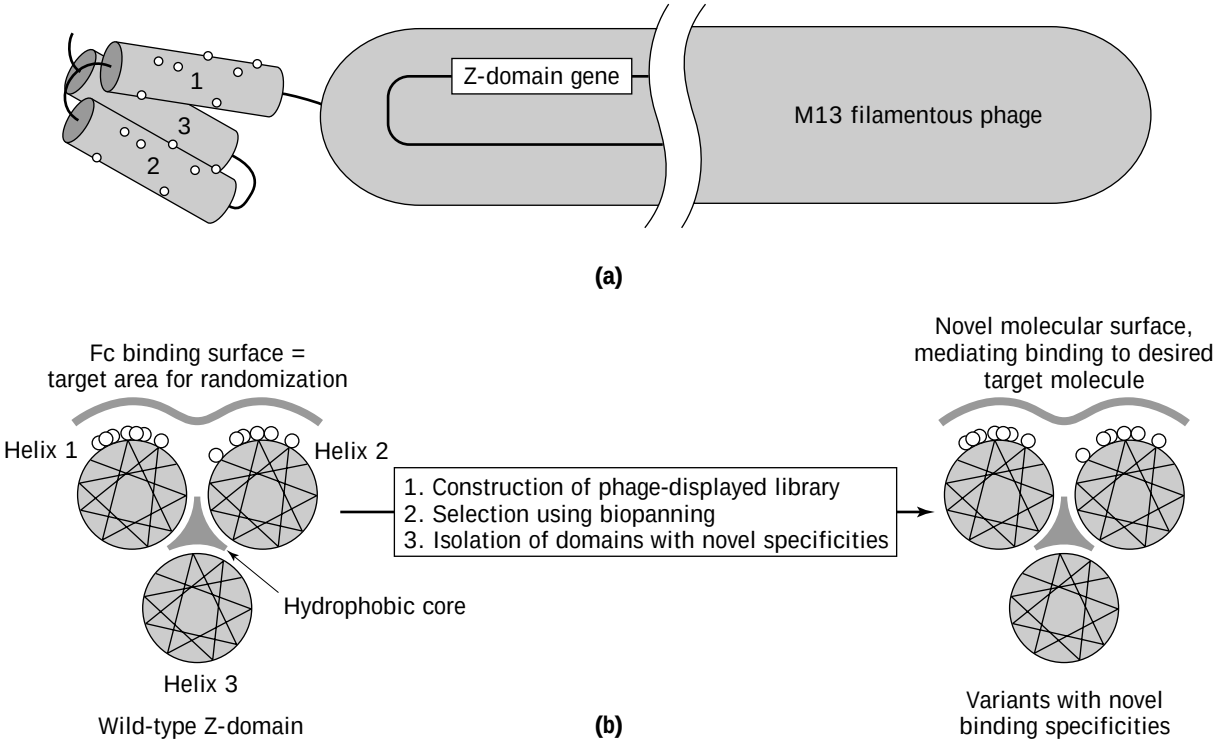


Figure 8. Schematic description showing how the Z-domain has been used as molecular scaffold in search of domain variants with novel binding specificities. **(a)** Using phage display technology, Z-variants can be expressed for presentation on the surface of filamentous phage particles. Phage libraries containing large numbers of phage clones, each displaying an individual Z-variant, can be subjected to rounds of biopannings in which Z-variants capable of binding to a desired target molecule can be selected. Identification of binding variants is accomplished through DNA sequencing of the packed phage DNA, containing the gene of the displayed Z-variant. **(b)** Top view of the Z-domain consisting of a triple α-helix. The library is constructed through randomization of amino acids present at the domain surface and involved in the Fc binding of the wild-type domain. Using the selection procedure (biopanning), variants with new molecular surfaces are isolated that are capable of binding the desired target.

to be used for purification, detection, immobilization, and perhaps even therapeutic purposes.

CONCLUSIONS

In this article, we have illustrated how fusions to the immunoglobulin-binding domains of SpA and the serum albumin-binding regions of SpG, respectively, have become important tools in many fields of research. Several strategies for improved heterologous expression and simplified recovery are based on such gene fusions. We focused on some current trends in certain fields of applications for SpA and SpG fusion proteins. It is apparent that SpA and SpG fusion proteins will soon find extensive use in novel applications in medicine, immunology, and biotechnology.

ACKNOWLEDGMENTS

All former and present members at the Department of Biochemistry and Biotechnology who have been involved in the development and evaluations of the SpA and SpG expression systems are gratefully acknowledged. We express our gratitude to Drs. Björn Nilsson, Tomas Moks, Lars Abrahmsén, and Maria Murby. In particular we thank Dr. Birger Jansson for help with literature searches concerning the use of SpA fusions in immunodetection techniques.

BIBLIOGRAPHY

1. M. Uhlén, B. Nilsson, B. Guss, M. Lindberg, S. Gatenbeck, and L. Philipson, *Gene* **23**, 369–378 (1983).
2. M. Uhlén, G. Forsberg, T. Moks, M. Hartmanis, and B. Nilsson, *Curr. Opin. Biotechnol.* **3**, 363–369 (1992).
3. S. Ståhl, P.-Å. Nygren, A. Sjölander, and M. Uhlén, *Curr. Opin. Immunol.* **5**, 272–277 (1993).
4. P.-Å. Nygren, S. Ståhl, and M. Uhlén, *Trends Biotechnol.* **12**, 184–188 (1994).
5. S. Ståhl, P.-Å. Nygren, and M. Uhlén, in R.S. Tuan ed., *Methods in Molecular Biology*, vol. 63, *Recombinant Proteins: Detection and Isolation Protocols*, Humana Press, Totowa, N.J., 1997, pp. 103–118.
6. E. Flaschel and K. Friehs, in M. Moo-Young and B.R. Glick, eds., *Biotechnology Advances*, Pergamon Press, Oxford, 1993, pp. 31–78.
7. C.F. Ford, I. Suominen, and C.E. Glatz, *Prot. Expr. Purif.* **2**, 95–107 (1991).
8. E.R. Lavallie and J.M. McCoy, *Curr. Opin. Biotechnol.* **6**, 501–505 (1995).
9. S.C. Makrides, *Microbiol. Rev.* **60**, 512–538 (1996).
10. J. Nilsson, S. Ståhl, J. Lundeberg, M. Uhlén, and P.-Å. Nygren, *Prot. Expr. Purif.* **11**, 1–16 (1997).
11. B. Nilsson and L. Abrahmsén, *Methods Enzymol.* **185**, 144–161 (1990).
12. B. Nilsson, T. Moks, B. Jansson, L. Abrahmsén, A. Elmblad, E. Holmgren, C. Henrichson, T.A. Jones, and M. Uhlén, *Prot. Eng.* **1**, 107–113 (1987).
13. M. Larsson, E. Brundell, L. Nordfors, C. Höög, M. Uhlén, and S. Ståhl, *Prot. Expr. Purif.* **7**, 447–457 (1996).
14. P.-Å. Nygren, M. Eliasson, E. Palmcrantz, L. Abrahmsén, and M. Uhlén, *J. Mol. Recognit.* **1**, 69–74 (1988).
15. J. Porath, J. Carlsson, I. Olsson, and G. Belfrage, *Nature* **258**, 598–599 (1975).
16. E. Hochuli, W. Bannwarth, H. Döbeli, R. Gentz, and D. Stüber, *Biotechnology* **6**, 1321–1325 (1988).
17. D.B. Smith and K.S. Johnson, *Gene* **67**, 31–40 (1988).
18. X. Zhao and C.H. Siu, *J. Biol. Chem.* **270**, 29413–29421 (1995).
19. C. di Guan, P. Li, P.D. Riggs, and H. Inouye, *Gene* **67**, 21–30 (1988).
20. R. Aitken, J. Gilchrist, and M.C. Sinclair, *Gene* **144**, 69–73 (1994).
21. T.H. Hopp, K.S. Prickett, V.L. Price, R.T. Libby, C.J. March, D.P. Cerretti, D.L. Urdal, and P.J. Conlon, *Biotechnology* **6**, 1204–1210 (1988).
22. B.L. Brizzard, R.G. Chubet, and D.L. Vizard, *Biotechniques* **16**, 730–734 (1994).
23. J.E. Cronan, *J. Biol. Chem.* **265**, 10327–10333 (1990).
24. J.K. Choi, F. Yu, E.S. Wurtele, and B.J. Nikolau, *Plant Physiol.* **109**, 619–625 (1995).
25. P.J. Schatz, *Biotechnology* **11**, 1138–1143 (1993).
26. J. Nilsson, M. Larsson, S. Ståhl, P.-Å. Nygren, and M. Uhlén, *J. Mol. Recognit.* **9**, 585–594 (1996).
27. A. Sjölander, S. Ståhl, and P. Perlmann, *ImmunoMethods* **2**, 79–92 (1993).
28. M. Uhlén and T. Moks, *Methods Enzymol.* **185**, 129–143 (1990).
29. J.J. Langone, *Adv. Immunol.* **32**, 157–252 (1982).
30. T. Moks, L. Abrahmsén, B. Nilsson, U. Hellman, J. Sjöquist, and M. Uhlén, *Eur. J. Biochem.* **156**, 637–643 (1986).
31. B. Guss, M. Uhlén, B. Nilsson, M. Lindberg, J. Sjöquist, and J. Sjödahl, *Eur. J. Biochem.* **138**, 413–420 (1984).
32. O. Schneewind, P. Model, and V.A. Fischetti, *Cell* **70**, 267–281 (1992).
33. O. Schneewind, D. Mihaylova-Petkov, and P. Model, *EMBO J.* **12**, 4803–4811 (1993).
34. W.W. Navarre and O. Schneewind, *Mol. Microbiol.* **14**, 115–121 (1994).
35. O. Schneewind, A. Fowler, and K.F. Faull, *Science* **268**, 103–106 (1995).
36. M. Eliasson, R. Andersson, A. Olsson, H. Wigzell, and M. Uhlén, *J. Immunol.* **142**, 575–581 (1989).
37. J. Deisenhofer, *Biochemistry* **20**, 2361–2370 (1981).
38. H. Gouda, H. Torigoe, A. Saito, M. Sato, Y. Arata, and I. Shimada, *Biochemistry* **31**, 9665–9672 (1992).
39. M. Tashiro and G.T. Montelione, *Curr. Opin. Struct. Biol.* **5**, 471–481 (1995).
40. B.A. Lyons, M. Tashiro, L. Cedergren, B. Nilsson, and G. Montelione, *Biochemistry* **32**, 7839–7845 (1993).
41. G. Forsberg, B. Baastrup, H. Rondahl, E. Holmgren, G. Pohl, M. Hartmanis, and M. Lake, *J. Prot. Chem.* **11**, 201–211 (1992).
42. E. Samuelsson, H. Wadensten, M. Hartmanis, T. Moks, and M. Uhlén, *Biotechnology* **9**, 363–366 (1991).
43. E. Samuelsson, T. Moks, B. Nilsson, and M. Uhlén, *Biochemistry* **33**, 4207–4211 (1994).
44. C. Ljungquist, B. Jansson, T. Moks, and M. Uhlén, *Eur. J. Biochem.* **186**, 557–561 (1989)
45. P.-Å. Nygren, C. Ljungquist, H. Trömborg, K. Nustad, and M. Uhlén, *Eur. J. Biochem.* **193**, 143–148 (1990).
46. C. Falkenberg, L. Björck, and B. Åkerström, *Biochemistry* **31**, 1451–1457 (1992).
47. J. Nilsson, P. Nilsson, Y. Williams, L. Pettersson, M. Uhlén, and P.-Å. Nyren, *Eur. J. Biochem.* **224**, 103–108 (1994).

48. P. Samuelson, M. Hansson, N. Ahlborg, C. Andréoni, F. Götz, T. Bächi, T.N. Nguyen, H. Binz, M. Uhlén, and S. Ståhl, *J. Bacteriol.* **177**, 1470–1476 (1995).
49. A. Sjölander, P.-Å. Nygren, S. Ståhl, K. Berzins, M. Uhlén, P. Perlmann, and R. Andersson, *J. Immunol. Methods* **201**, 115–123 (1997).
50. S. Ståhl, A. Sjölander, P.-Å. Nygren, K. Berzins, P. Perlmann, and M. Uhlén, *J. Immunol. Methods* **124**, 43–52 (1989).
51. M. Murby, E. Samuelsson, T.N. Nguyen, L. Mignard, U. Power, H. Binz, M. Uhlén, and S. Ståhl, *Eur. J. Biochem.* **230**, 38–44 (1995).
52. P.J. Kraulis, P. Jonasson, P.-Å. Nygren, M. Uhlén, L. Jendeberg, B. Nilsson, and J. Kördel, *FEBS Lett.* **378**, 190–194 (1996).
53. L. Abrahmsén, T. Moks, B. Nilsson, and M. Uhlén, *Nucleic Acids Res.* **14**, 7487–7500 (1986).
54. F. Ducancel, J.-C. Boulain, O. Trémeau, and A. Ménez, *Prot. Eng.* **3**, 139–143 (1990).
55. M. Hansson, S. Ståhl, R. Hjorth, M. Uhlén, and T. Moks, *Biotechnology* **12**, 285–288 (1994).
56. D. Haddad, S. Liljeqvist, S. Kumar, M. Hansson, S. Ståhl, H. Perlmann, P. Perlmann, and K. Berzins, *FEMS Immunol. Med. Microbiol.* **12**, 175–186 (1995).
57. S. Liljeqvist, D. Haddad, K. Berzins, M. Uhlén, and S. Ståhl, *Biochem. Biophys. Res. Commun.* **218**, 356–359 (1996).
58. X. Cao, T. Preiss, E.P. Slater, H.M. Westphal, and M. Beato, *J. Steroid Biochem. Mol. Biol.* **44**, 1–11 (1993).
59. B. Hammarberg, T. Moks, M. Tally, A. Elmblad, E. Holmgren, M. Murby, B. Nilsson, S. Josephson, and M. Uhlén, *J. Bacteriol.* **14**, 423–438 (1990).
60. K.H. Schmidt, C. Klessen, W. Kohler, and H. Malke, *J. Immunol. Methods* **143**, 111–117 (1991).
61. V. Nagarajan, H. Albertson, M. Chen, and J. Ribbe, *Gene* **114**, 121–126 (1992).
62. B. Antonsson, S. Montessuit, L. Friedli, M.A. Payton, and G. Paravicini, *J. Biol. Chem.* **269**, 16821–16828 (1994).
63. B. Jansson, Ph.D. Thesis, Royal Institute of Technology, Stockholm, Sweden, 1996.
64. R.L. McKown and G.J. Warren, *Cryobiology* **28**, 474–482 (1991).
65. D.A. Stirling, A. Petrie, D.J. Pulford, D.T.W. Paterson, and M.J.R. Stark, *Mol. Microbiol* **6**, 703–713 (1992).
66. J. Zueco and A. Boyd, *Anal. Biochem.* **207**, 348–349 (1992).
67. R. Hightower, C. Baden, E. Penzes, P. Lund, and P. Dunsmuir, *Plant Mol. Biol.* **17**, 1013–1021 (1991).
68. D. Andersons, A. Engström, S. Josephson, L. Hansson, and H. Steiner, *Biochem. J.* **280**, 219–224 (1991).
69. C. Oker-Blom, A.-M. Soumalainen, K. Åkerman, Z. Qi, C. Lindqvist, A. Kuusisto, and M. Karp, *Biotechniques* **14**, 800–809 (1993).
70. H. Dorai, J.E. McCartney, R.M. Hudziak, M.-S. Tai, A.A. Laminet, L.L. Houston, J.S. Huston, and H. Oppermann, *Biotechnology* **12**, 890–897 (1994).
71. F. Liu and B. Roizman, *J. Virol.* **65**, 5149–5156 (1991).
72. Y. Maeda, H. Ueda, T. Hara, J. Kazami, G. Kawano, E. Suzuki, and T. Nagamune, *Biotechniques* **20**, 116–121 (1996).
73. R. Sanchez-Lopez, R. Nicholson, M.-C. Gesnel, L.M. Matrisian, and R. Breathnach, *J. Biol. Chem.* **263**, 11892–11899 (1988).
74. T. de Vries, C. Srnka, M.M. Palcic, S.J. Swiedler, D.H. van den Eijnden, and B.A. Macher, *J. Biol. Chem.* **270**, 8712–8722 (1995).
75. L. Zhang and J.D. Esko, *J. Biol. Chem.* **269**, 19295–19299 (1994).
76. M. Hansson, S. Ståhl, T.N. Nguyen, T. Bächi, A. Robert, H. Binz, A. Sjölander, and M. Uhlén, *J. Bacteriol.* **174**, 4239–4245 (1992).
77. T.N. Nguyen, M. Hansson, S. Ståhl, T. Bächi, A. Robert, W. Domzig, H. Binz, and M. Uhlén, *Gene* **128**, 89–94 (1993).
78. T.N. Nguyen, M.-H. Gourdon, M. Hansson, A. Robert, P. Samuelson C. Libon, C. Andréoni, P.-Å. Nygren, H. Binz, M. Uhlén, and S. Ståhl, *J. Biotechnol.* **47**, 207–219 (1995).
79. E. Gunneriusson, P. Samuelson, M. Uhlén, P.-Å. Nygren, and S. Stähl, *J. Bacteriol.* **178**, 1341–1346 (1996).
80. A. Robert, P. Samuelson, C. Andréoni, T. Bächi, M. Uhlén, H. Binz, T.N. Nguyen, and S. Stähl, *FEBS Lett.* **390**, 327–333 (1996).
81. S.C. Makrides, P.-Å. Nygren, S.B. Andrews, P.J. Ford, K.S. Evans, E.G. Hayman, H. Adari, J. Levin, M. Uhlén, and C.A. Toth, *J. Pharmacol. Exp. Ther.* **277**, 534–542 (1996).
82. Amersham Pharmacia Biotech, Uppsala, Sweden.
83. A. Hedrum, J. Lundeberg, C. Påhlson, and M. Uhlén, *PCR Methods Appl.* **2**, 167–171 (1992).
84. B. Löwenadler, B. Jansson, S. Paleus, E. Holmgren, B. Nilsson, T. Moks, G. Palm, S. Josephson, L. Philipson, and M. Uhlén, *Gene* **58**, 87–97 (1987).
85. L. Abrahmsén, T. Moks, B. Nilsson, U. Hellman, and M. Uhlén, *EMBO J.* **4**, 3901–3906 (1985).
86. T. Moks, L. Abrahmsén, B. Österlöf, S. Josephson, M. Östling, S.-O. Enfors, I. Persson, B. Nilsson, and M. Uhlén, *Biotechnology* **5**, 379–382 (1987).
87. Nilsson, L. Abrahmsén, and M. Uhlén, *EMBO J.* **4**, 1075–1080 (1985).
88. M. Murby, L. Cedergren, J. Nilsson, P.-Å. Nygren, B. Hammarberg, B. Nilsson, S.-O. Enfors, and M. Uhlén, *Biotechnol. Appl. Biochem.* **14**, 336–346 (1991).
89. K. Berzins, S. Adams, J.R. Broderson, C. Chizzolini, M. Hansson, K. Lövgren, P. Millet, C.L. Morris, H. Perlmann, P. Perlmann, A. Sjölander, S. Ståhl, J.S. Sullivan, M. Troye-Blomberg, B. Wåhlin-Flyg, and W.E. Collins, *Vaccine Res.* **4**, 121–133 (1995).
90. R. Rudolph, in J.L. Cleland and C.S. Craik, eds., *Principles and Practice of Protein Folding*, Wiley, New York, 1995, pp. 283–298.
91. F.W. Studier, A.H. Rosenberg, J.J. Dunn, and J.W. Dubendorff, *Methods Enzymol.* **185**, 60–89 (1990).
92. K. Köhler, C. Ljungquist, A. Kondo, A. Veide, and B. Nilsson, *Biotechnology* **9**, 642–646 (1991).
93. M. Murby, T.N. Nguyen, H. Binz, M. Uhlén, and S. Ståhl, in D.L. Pyle, ed., *Separations for Biotechnology*, vol. 3, Bookcraft, Bath, UK, 1994, pp. 336–344.
94. A. Sjölander, S. Ståhl, K. Lövgren, M. Hansson, L. Cavelier, A. Walles, H. Helmby, B. Wåhlin, B. Morein, M. Uhlén, K. Berzins, P. Perlmann, and M. Wahlgren, *Exp. Parasitol.* **76**, 134–145 (1993).
95. S. Josephson and R. Bishop, *Trends Biotechnol.* **6**, 218–224 (1988).
96. J. Nilsson, P. Jonasson, E. Samuelsson, S. Ståhl, and M. Uhlén, *J. Biotechnol.* **248**, 241–250 (1996).
97. M. Murby, M. Uhlén, and S. Ståhl, *Prot. Expr. Purif.* **7**, 129–136 (1996).
98. Hammarberg, P.-Å. Nygren, E. Holmgren, A. Elmblad, M. Tally, U. Hellman, T. Moks, and M. Uhlén, *Proc. Natl. Acad. Sci. USA* **86**, 4367–4371 (1989).

99. B. Jansson, C. Palmcrantz, M. Uhlén, and B. Nilsson, *Prot. Eng.* **2**, 555–561 (1990).
100. J.S. Kim and R.T. Raines, *Anal. Biochem.* **219**, 165–166 (1994).
101. C.A. Panagiotidis and S.J. Silverstein, *Gene* **164**, 45–47 (1995).
102. T. Gräslund, J. Nilsson, M. Lindberg, M. Uhlén, and P.-Å. Nygren, *Prot. Expr. Purif.* **9**, 125–132 (1997).
103. P. Jonasson, J. Nilsson, E. Samuelsson, T. Moks, S. Ståhl, and M. Uhlén, *Eur. J. Biochem.* **236**, 656–661 (1996).
104. B.-L. Johansson, A. Kernell, S., Sjöberg, and J. Wahren, *J. Clin. Endocrinol. Metab.* **77**, 976–981 (1993).
105. P. Jonasson, P.-Å. Nygren, B.-L. Johansson, J. Wahren, M. Uhlén, and S. Ståhl, *Gene* **210**, 203–210 (1998).
106. E. Samuelsson, P. Jonasson, F. Viklund, B. Nilsson, and M. Uhlén, *Nat. Biotechnol.* **14**, 751–755 (1996).
107. K. Miyashita, M. Kusumi, R. Utsumi, T. Komano, and N. Satoh, *Biosci. Biotechnol. Biochem.* **56**, 746–750 (1992).
108. P.A. Walker, L.E.-C. Leong, P.W.P. Ng, S. Han Tan, S. Waller, D. Murphy, and A.G. Porter, *Biotechnology* **12**, 601–605 (1994).
109. D.J. Capon, S.M. Chamow, J. Mordenti, S.A. Marsters, T. Gregory, H. Mitsuya, R.A. Byrn, C. Lucas, F.M. Wurm, J.E. Groopman, S. Broder, and D.H. Smith, *Nature* **337**, 525–531 (1989).
110. P.-Å. Nygren, P. Flodby, R. Andersson, H. Wigzell, and M. Uhlén, in R.M. Chanock, H.S. Ginsberg, F. Brown, and R.A. Lerner eds., *Vaccines 91*, Cold Spring Harbor Laboratory Press, New York, 1991, pp. 363–368.
111. C. Lindbladh, K. Mosbach, and L. Bülow, *Trends Biochem. Sci.* **18**, 279–283 (1993).
112. E. Kobatake, Y. Ikariyama, M. Aizawa, K. Miwa, and S. Kato, *J. Biotechnol.* **16**, 87–96 (1990).
113. C. Lindbladh, K. Mosbach, and L. Bülow, *J. Immunol. Methods* **137**, 199–207 (1991).
114. E. Witt, G.C. Stewart, and W. Hengstenberg, *Prot. Eng.* **5**, 267–271 (1992).
115. P.S. Chowdhury, A. Kushwaha, S. Abrol, and V.K. Chaudhary, *Prot. Expr. Purif.* **5**, 89–95 (1994).
116. E. Kobatake, Y. Ikariyama, and M. Aizawa, *J. Biotechnol.* **38**, 263–268 (1995).
117. J. Lundeberg, J. Wahlberg, and M. Uhlén, *Gene Anal. Technol.* **7**, 47–53 (1990).
118. J. Lundeberg, J. Wahlberg, M. Holmberg, U. Pettersson, and M. Uhlén, *DNA Cell Biol.* **9**, 287–292 (1990).
119. C. Ljungquist, J. Lundeberg, A.M. Rasmussen, E. Hornes, and M. Uhlén, *DNA Cell Biol.* **12**, 191–197 (1993).
120. T. Sano, C. Smith, and C.R. Cantor, *Science* **258**, 120–122 (1992).
121. D.S. Clark, *Trends Biotechnol.* **12**, 439–443 (1994).
122. J. Nilsson, M. Bosnes, F. Larsen, P.-Å. Nygren, M. Uhlén, and J. Lundeberg, *Biotechniques* **22**, 744–751 (1997).
123. U.F. Power, H. Plotnicky-Gilquin, T. Huss, T.N. Nguyen, A. Robert, M. Trudel, S. Ståhl, M. Uhlén, and H. Binz, *Virology* **230**, 155–166 (1997).
124. C. Libon, N. Corvaïa, J.-F. Haeuw, T.N. Nguyen, S. Ståhl, J.-Y. Bonnefoy, and C. Andréoni, *Vaccine* **17**, 406–414 (1999).
125. G. Georgious, C. Stathopoulos, P.S. Daugherty, A.R. Nayak, B.L. Iverson, and R. Curtiss III, *Nat. Biotechnol.* **15**, 29–34 (1997).
126. S. Ståhl and M. Uhlén, *Trends Biotechnol.* **15**, 185–192 (1997).
127. Z. Lu, K.S. Murray, V. van Cleave, E.R. Lavallie, M.L. Stahl, and J.M. McCoy, *Biotechnology* **13**, 366–372 (1995).
128. C. Sousa, A. Cebolla, and V. de Lorenzo, *Nat. Biotechnol.* **14**, 1017–1020 (1996).
129. C.J. Stover, G.P. Bansal, M.S. Hanson, J.E. Burlein, S.R. Palaszynski, J.F. Young, S. Koenig, D.B. Young, A. Sadziene, and A.G. Barbour, *J. Exp. Med.* **178**, 197–209 (1993).
130. V. A. Fischetti, D. Medaglini, and G. Pozzi, *Curr. Opin. Biotechnol.* **7**, 659–666 (1996).
131. S. Ståhl, P. Samuelson, M. Hansson, C. Andréoni, L. Goetsch, C. Libon, S. Liljeqvist, E. Gunneriusson, H. Binz, T.N. Nguyen, and Uhlén M., in J. Wells and G. Pozzi, eds., *Recombinant Gram-Positive Bacteria as Vaccine Vehicles for Mucosal Immunization*, Landes, Springer-Verlag, New York, 1997, pp. 61–81.
132. S. Liljeqvist, P. Samuelson, M. Hansson, T.N. Nguyen, H. Binz, and S. Ståhl, *Appl. Environ. Microbiol.* **63**, 2481–2488 (1997).
133. C. Andréoni, L. Goetsch, C. Libon, P. Samuelson, A. Robert, T.N. Nguyen, M. Uhlén, H. Binz, and S. Ståhl, *Biotechniques* **23**, 696–704 (1997).
134. G.P. Smith, *Science* **228**, 1315–1317 (1985).
135. P.-Å. Nygren, and M. Uhlén, *Curr. Opin. Struct. Biol.* **7**, 463–469 (1997).
136. K. Nord, J. Nilsson, B. Nilsson, M. Uhlén, and P.-Å. Nygren, *Prot. Eng.* **6**, 601–608 (1995).
137. K. Nord, E. Gunneriusson, J. Ringdahl, S. Ståhl, M. Uhlén, and P.-Å. Nygren, *Nat. Biotechnol.* **15**, 772–777 (1997).

See also ANTIBODY PURIFICATION; HYBRIDOMA, ANTIBODY PRODUCTION.

AIR HANDLING. See FILTRATION, AIR; HEATING, VENTILATION AND AIR CONDITIONING.

ALANINE. See AMINO ACIDS, PRODUCTION PROCESSES.

ALDEHYDE REDUCTASE

SAKAYU SHIMIZU
MICHIHIKO KATAOKA
Kyoto University
Kyoto, Japan

KEY WORDS

Aldehyde reductase
Aldo-keto reductase superfamily
Chiral alcohol
4-Chloro-3-hydroxybutanoate ester
Glucose dehydrogenase
$NADP^+$
NADPH
Optically active alcohol

Oxidoreductase

Stereospecific reduction

OUTLINE

INTRODUCTION

There is a group of reductases that catalyze the reduction of various aldehydes and ketones (1,2). They occur widely in living organisms and are suggested to be involved in the metabolism of biogenic and xenobiotic carbonyl compounds. From the viewpoint of the practical application of these reductases, the stereospecific reduction of carbonyl compounds with enzymes exhibiting stereospecificity is an effective method for the production of various optically active alcohols (3–6).

Among these enzymes catalyzing the reduction of various carbonyl compounds, a class of monomeric NADPH-dependent oxidoreductases with molecular masses of about 35 kDa, called the aldo-keto reductase superfamily (7), has been well characterized. This superfamily contains many reductases, such as aldehyde reductase (EC 1.1.1.2), aldose reductase (EC 1.1.1.21), and so forth, and these proteins show high similarity in their substrate specificities and primary amino acid sequences (7). Here we describe aldehyde reductase and the application of microbial aldehyde reductase to the production of chiral alcohols.

ALDEHYDE REDUCTASE AND ALDO-KETO REDUCTASE SUPERFAMILY ENZYMES

The aldo-keto reductase superfamily is a class of monomeric NADPH-dependent oxidoreductases with molecular masses of about 35 kDa (7). This group includes aldehyde and aldose reductases (8–14), prostaglandin F synthase (EC 1.1.1.188) (15), 2,5-diketogluconate reductase (16,17), morphine dehydrogenase (EC 1.1.1.218) (18), chlordecone reductase (EC 1.1.1.225) (19), a number of hydroxysteroid dehydrogenases (EC 1.1.1.62, 1.1.1.149, 1.1.1.213) (20–25), plant chalcone reductases (26,27), and so on. From the strong similarity in their amino acid sequences, ρ-crystallin (28) and a yeast protein encoded by the *GCY* gene (29), which do not exhibit oxidoreductase activity, are also included in this superfamily. (A list of current members of the aldo-keto reductase superfamily can be seen on the AKR Web page at *http://pharme26.med.upenn.edu./*.) Among them, aldehyde reductase and aldose reductase catalyze the NADPH-dependent reduction of a wide range of aromatic and aliphatic aldehydes, such as *p*-nitrobenzaldehyde, pyridine-3-aldehyde, and glyceraldehyde, to the corresponding alcohols (8–14). These two enzymes are widely distributed in various kinds of mammals including humans, cows, pigs, rats, and mice.

Because aldose and aldehyde reductases cause complications in diabetes (30), the crystal structures of human and pig aldose reductases and aldehyde reductases have been determined by X-ray diffraction methods (31–34). As oxidoreductases, the enzymes are unique in that they have a β/α–triosephosphate isomerase (TIM) barrel structure, and are the first oxidoreductases known to possess such a structure instead of a dinucleotide, or Rossmann, binding fold. Elucidation of the tertiary structures and reaction mechanisms of these enzymes will facilitate the design of specific inhibitors that may be used as drugs for the treatment of diabetic complications.

NOVEL MICROBIAL ALDEHYDE REDUCTASE

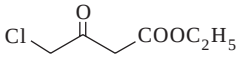
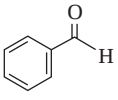
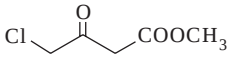
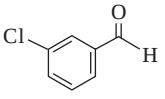
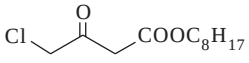
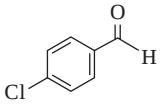
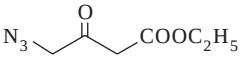
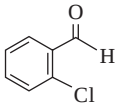
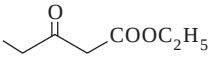
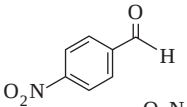
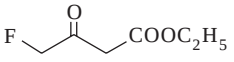
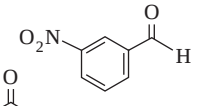
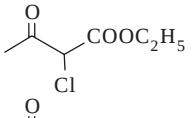
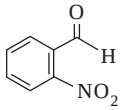
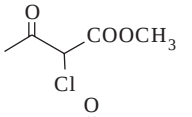
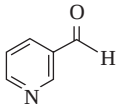
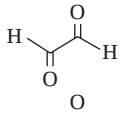
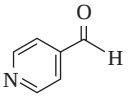
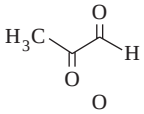
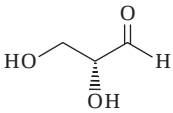
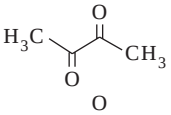
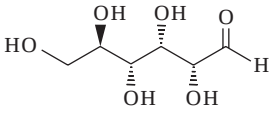
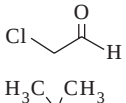
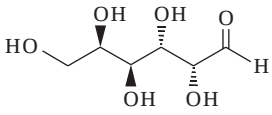
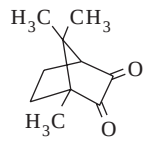
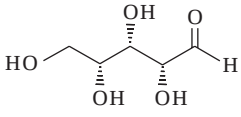
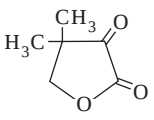
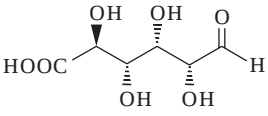
Yamada et al. found a novel aldehyde reductase in a red yeast, *Sporobolomyces salmonicolor* (35–37), and succeeded in cloning the aldehyde reductase gene from this strain (38).

Aldehyde reductase was isolated in a crystalline form from cells of *S. salmonicolor* and was characterized in some detail (35–37). It should be noted that the cellular content of the enzyme comprised more than 4.5% of the total extractable protein. The enzyme has a monomeric form of about 35 kDa. The enzyme absolutely requires NADPH as a cofactor. The substrate specificity of the enzyme is shown in Table 1. The enzyme did not catalyze the reduction of any keto acids or monoketones tested. Several aldehydes such as *p*-nitrobenzaldehyde, *p*-chlorobenzaldehyde, pyridine-3-aldehyde, or D-glyceraldehyde were readily reduced. These aldehydes are typical substrates for aldo-keto reductase family members, as already described. Quinones, such as *p*-benzoquinone, α-naphthoquinone, β-naphthoquinone, and menadione, which are good substrates for carbonyl reductases (14), were not reduced. Although no detectable oxidation of NADPH was observed with D-glucuronate, D-glucose, D-xylose, or D-galactose as the substrate at a low concentration (0.2 mM), the enzyme effectively catalyzed the reduction of these compounds when they were present at a high concentration (100 mM). The K_m values for these substrates are quite high in comparison with those for other substrates, which suggests that the affinity of the enzyme for these substrates is very low. Furthermore, the enzyme effectively catalyzed the reduction of 4-haloacetoacetate esters to the corresponding (*R*)-4-halo-3-hydroxybutanoate esters (100% ee) (35,36). Because the reduction products derived from 4-haloacetoacetate esters are optically active alcohols, the enzyme can recognize the stereopositions of these substrates.

2-Chloroacetoacetate esters were also effective substrates. Initial velocity and inhibition studies indicated that the enzyme reaction sequence is ordered, with NADPH binding to the free enzyme, and $NADP^+$ being the last product to be released. Quercetin, dicoumarol, and some SH-reagents inhibited the enzyme activity. These results of enzymological studies suggested that the enzyme of *S. salmonicolor* belongs to the aldo-keto reductase superfamily.

cDNA coding the aldehyde reductase was cloned from cells of *S. salmonicolor* and sequenced (38). The aldehyde reductase gene comprises 969 bp and encodes a polypeptide of 35,232 Da. The deduced amino acid sequence shows a high degree of similarity to that of other members of the aldo-keto reductase superfamily. The identities between *S. salmonicolor* aldehyde reductase and the bovine (39), rat (40), human (7), rabbit (41), and barley (42) aldose reductases, human aldehyde reductase (7), and xylose reductase from *Pichia stipitis* (43) are 45.1, 45.0, 43.7, 42.9, 39.5, 43.9, and 37.7%, respectively. In general, aldo-keto reductases exhibit greater similarity in their N-terminal regions than in their C-terminal regions (Fig. 1). Analysis of the genomic DNA sequence indicated that the aldehyde reductase gene was interrupted by six introns (two in the 5′ noncoding region and four in the coding region).

Table 1. Substrate Specificity of the Aldehyde Reductase from *S. salmonicolor*

Substrate	Relative activity (%)	Substrate	Relative activity (%)
	100		14
	25		56
	240		52
	65		58
	75		468
	153		63
	330		14
	74		228
	74		54
	219		64
	75		81
	17		24
	16		472
	0		173

Note: See Refs. 35 and 36 for details.

Figure 1. Comparison of the primary structures of the aldehyde reductase from *S. salmonicolor* and aldo-keto reductase superfamily proteins. The sequences of the aldehyde reductase (sALR), human aldehyde reductase (hALR), human aldose reductase (hADR), rat aldose reductase (rADR), bovine prostaglandin F synthase (bPGF), frog ρ-crystallin (fRHO), the yeast nuclear gene product (sGCY), and 2,5-diketogluconic acid reductase from *Corynebacterium* sp. (cDKG) are aligned. Gaps in the aligned sequences are indicated (–). Identical amino acid residues are enclosed in boxes. See Ref. 38 for details.

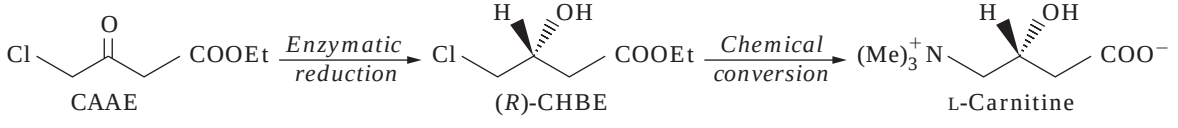


Figure 2. Chemicoenzymatic route for the synthesis of L-carnitine. CAAE, ethyl 4-chloroacetoacetate; CHBE, ethyl 4-chloro-3-hydroxybutanoate.

APPLICATION OF MICROBIAL ALDEHYDE REDUCTASE TO THE PRODUCTION OF OPTICALLY ACTIVE ALCOHOLS

The asymmetric reduction of carbonyl compounds with microorganisms or enzymes is a useful method for the synthesis of optically active compounds (3–6). The aldehyde reductase produced by *S. salmonicolor* effectively catalyzes the asymmetric reduction of 4-haloacetoacetate esters to the corresponding (*R*)-isomers, which are promising chiral building blocks for the chemical synthesis of L-carnitine (Fig. 2) and so forth (44–47).

Stereospecific Reduction of Ethyl 4-Chloroacetoacetate

Ethyl (*R*)-4-chloro-3-hydroxybutanoate formed through the reduction of ethyl 4-chloroacetoacetate with cells of *S. salmonicolor* showed low optical purity (8.5% ee) because the cells contained a significant amount of an enzyme(s) catalyzing the reduction of ethyl 4-chloroacetoacetate to ethyl (*S*)-4-chloro-3-hydroxybutanoate in addition to the aldehyde reductase forming ethyl (*R*)-4-chloro-3-hydroxybutanoate. Because heat treatment and acetone fractionation of a cell-free extract of *S. salmonicolor* remove the (*S*)-isomer-forming enzyme(s) from the cell extract, the cell extract after these treatments was used as the source of aldehyde reductase. Commercially available glucose dehydrogenase, $NADP^+$, and glucose were added as a cofactor regenerator. A practical preparation can be carried out in an organic solvent–water two-phase system, as shown in Fig. 3a, because the substrate is unstable in an aqueous solution, and both the substrate and product strongly inhibit the enzyme reaction. In this system, most of the substrate is present in the organic phase; consequently, the ethyl 4-chloroacetoacetate is quite stable and cannot in-

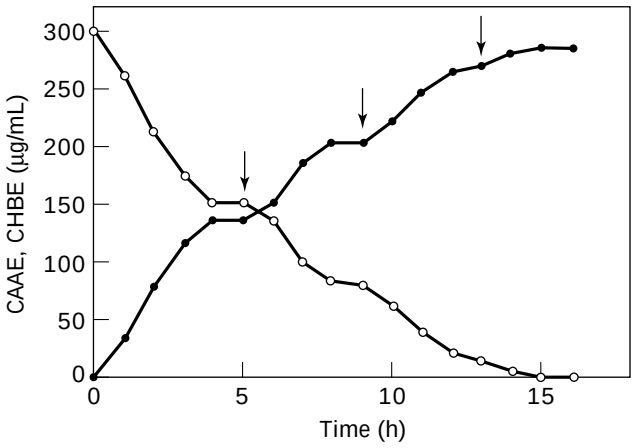


Figure 4. Conversion of ethyl 4-chloroacetoacetate to ethyl (*R*)-4-chloro-3-hydroxybutanoate in a two-phase reaction system by *E. coli* JM109 [pKAR, pKKGDH] cells. Glucose (100 mg/mL) was added to the aqueous phase periodically (arrows). Symbols: ○, ethyl 4-chloroacetoacetate (CAAE); ●, ethyl (*R*)-4-chloro-3-hydroxybutanoate (CHBE). See Refs. 49–52 for details.

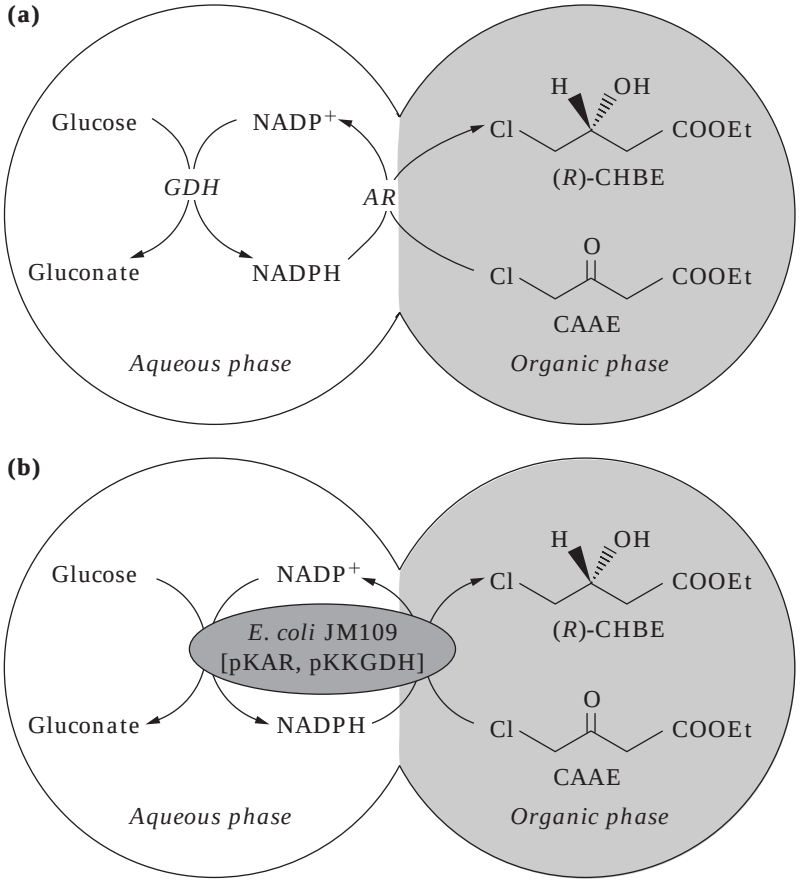


Figure 3. Outline of the stereospecific reduction of ethyl 4-chloroacetoacetate by aldehyde reductase (AR) and glucose dehydrogenase (GDH) (**a**), and washed cells of *E. coli* JM109 [pKAR, pKKGDH] (**b**), as catalysts, in an organic solvent–water two-phase system. CAAE, ethyl 4-chloroacetoacetate; CHBE, ethyl 4-chloro-3-hydroxybutanoate. See Refs. 48–52 for details.

hibit the enzyme reaction occurring in the aqueous phase. *n*-Butyl acetate is regarded as the most suitable organic solvent for such a two-phase system because it shows high partition efficiencies with regard to both the substrate and product, and both enzymes are stable in the presence of this organic solvent. In a bench-scale two-phase reaction, 83.8 g/L of ethyl (*R*)-4-chloro-3-hydroxybutanoate (86% ee) was produced from ethyl 4-chloroacetoacetate, with a molar yield of 95.4% (48,53).

Enzymatic Production of Ethyl (*R*)-4-Chloro-3-Hydroxybutanoate

The two-phase reaction was improved by using an *Escherichia coli* transformant. The aldehyde reductase gene from *S. salmonicolor* (38) and the glucose dehydrogenase gene from *Bacillus megaterium* (54) were transformed into the same *E. coli* strain; consequently, a transformant overproducing both aldehyde reductase and glucose dehydrogenase, *E. coli* JM109 [pKAR, pKKGDH], was obtained (49–52). The reduction reaction was carried out in an organic solvent–water two-phase system containing ethyl 4-chloroacetoacetate, glucose, and $NADP^+$, and *E. coli* JM109 [pKAR, pKKGDH] cells as the catalyst (Fig. 3b). When the *E. coli* cells were incubated in the *n*-butyl acetate–water two-phase system, 300 mg/mL of ethyl 4-chloroacetoacetate was almost stoichiometrically converted to ethyl (*R*)-4-chloro-3-hydroxybutanoate (92% ee) in 16 h (Fig. 4). Because the use of *E. coli* transformant cells as the catalyst is simple and does not require isolation of the aldehyde reductase, which is necessary for *S. salmonicolor* cells, it is highly advantageous for the practical synthesis of ethyl (*R*)-4-chloro-3-hydroxybutanoate.

BIBLIOGRAPHY

1. H. Weiner, D.W. Crabb, and T.G. Flynn eds., *Enzymology and Molecular Biology of Carbonyl Metabolism 4, Advances in Experimental Medicine and Biology*, Vol. 328, Plenum, New York, 1993.
2. H. Weiner, R.S. Holmes, and B. Wermuth eds., *Enzymology and Molecular Biology of Carbonyl Metabolism 5, Advances in Experimental Medicine and Biology*, Vol. 372, Plenum, New York, 1995.
3. W. Hummel and M.-R. Kula, *Eur. J. Biochem.* **184**, 1–13 (1989).
4. J.B. Jones, *Tetrahedron* **42**, 3351–3403 (1986).
5. G.M. Whitesides and C.-H. Wong, *Angew. Chem. Int. Ed. Engl.* **24**, 617–638, (1985).
6. H. Yamada and S. Shimizu, *Angew. Chem. Int. Ed. Engl.* **27**, 622–642 (1988).
7. K.M. Bohren, B. Bullock, B. Wermuth, and K.H. Gabbay, *J. Biol. Chem.* **264**, 9547–9551 (1989).
8. T.G. Flynn, *Biochem. Pharmacol.* **31**, 2705–2712 (1982).
9. B. Wermuth, J.D.B. Munch, and J.-P. von Wartburg, *J. Biol. Chem.* **252**, 3821–3828 (1977).
10. T.G. Flynn, J.A. Cromlish, and W.S. Davidson, *Methods Enzymol.* **89**, 501–506 (1982).
11. J.-P. von Wartburg and B. Wermuth, *Methods Enzymol.* **89**, 506–513 (1982).
12. B. Wermuth, H. Burgisser, K. Bohren, and J.-P. von Wartburg, *Eur. J. Biochem.* **127**, 279–284 (1982).
13. B. Wermuth and J.-P. von Wartburg, *Methods Enzymol.* **89**, 181–186 (1982).
14. B. Wermuth, *J. Biol. Chem.* **256**, 1206–1213 (1981).
15. K. Watanabe, Y. Fujii, K. Nakayama, H. Ohkubo, S. Kuramitsu, H. Kagamiyama, S. Nakanishi, and O. Hayaishi, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 11–15 (1988).
16. S. Anderson, C.B. Marks, R. Lazarus, J. Miller, K. Stafford, J. Seymour, D. Light, W. Rastetter, and D. Estell, *Science* **230**, 144–149 (1985).
17. J.F. Grindley, M.A. Payton, H. van de Pol, and K.G. Hardy, *Appl. Environ. Microbiol.* **54**, 1770–1775 (1988).
18. D.L. Willey, D.A. Caswell, C.R. Lowe, and N.C. Bruce, *Biochem. J.* **290**, 539–544 (1993).
19. C.J. Winters, D.T. Molowa, and P.S. Guzelian, *Biochemistry* **29**, 1080–1087, (1990).
20. Y. Deyashiki, K. Ohshima, M. Nakanishi, K. Sato, K. Matsuura, and A. Hara, *J. Biol. Chem.* **270**, 10461–10467 (1995).
21. A. Hara, K. Matsuura, Y. Tamada, K. Sato, Y. Miyabe, Y. Deyashiki, and N. Ishida, *Biochem. J.* **313**, 373–376 (1996).
22. R. Miura, K. Shiota, K. Noda, S. Yagi, T. Ogawa, and M. Takahashi, *Biochem. J.* **299**, 561–567 (1994).
23. W.R. Lacy, K.J. Washenick, R.G. Cook, and B.S. Dunbar, *Mol. Endocrinol.* **7**, 58–66 (1993).
24. J.E. Pawlowski, M. Huizinga, and T.M. Penning, *J. Biol. Chem.* **266**, 8820–8825 (1991).
25. M. Khanna, K.N. Qin, R.W. Wang, and K.C. Cheng, *J. Biol. Chem.* **270**, 20162–20168 (1995).
26. C. Sallaud, J. El-Turk, L. Bigarré, H. Sevin, R. Welle, and R. Esnault, *Plant Physiol.* **108**, 869–870 (1995).
27. R. Welle, G. Schröder, E. Schilts, H. Grisebach, and J. Schröder, *Eur. J. Biochem.* **196**, 423–430 (1991).
28. S.I. Tomarev, R.D. Zinovieva, S.M. Dolgilevich, S.V. Luchin, A.S. Krayev, K.G. Skryabin, and G. G. Gause, Jr., *FEBS Lett.* **171**, 297–302 (1984).
29. U. Oechsner, V. Magdolen, and W. Bandlow, *FEBS Lett.* **238**, 123–128 (1988).
30. J.H. Kinoshita and C. Nishimura, *Diabetes Metab. Rev.* **4**, 323–337 (1988).
31. O. El-Kabbani, N.C. Green, G. Lin, M. Carson, S.V.L. Narayana, K.M. Moore, T.G. Flynn, and L.J. DeLucas, *Acta Cryst.* **D50**, 859–868 (1994).
32. O. El-Kabbani, S.V.L. Narayana, Y.S. Babu, K.M. Moore, T.G. Flynn, J.M. Petrash, E.M. Westbrook, L.J. DeLucas, and C.E. Bugg, *J. Mol. Biol.* **218**, 695–698 (1991).
33. J.-M. Rondeau, F. Tête-Favier, A. Podjarny, J.-M. Reymann, P. Barth, J.-F. Biellmann, and D. Moras, *Nature* **355**, 469–472 (1992).
34. D.K. Wilson, K.M. Bohren, K.H. Gabbay, and F.A. Quiocho, *Science* **257**, 81–84 (1992).
35. H. Yamada, S. Shimizu, M. Kataoka, H. Sakai, and T. Miyoshi, *FEMS Microbiol. Lett.* **70**, 45–48 (1990).
36. M. Kataoka, H. Sakai, T. Morikawa, M. Katoh, T. Miyoshi, S. Shimizu, and H. Yamada, *Biochim. Biophys. Acta* **1122**, 57–62 (1992).
37. M. Kataoka, K. Yamamoto, S. Shimizu, K. Kita, T. Hashimoto, H. Yanase, K. Kita, and K. Miki, *Acta Cryst.* **D52**, 405–406 (1996).
38. K. Kita, K. Matsuzaki, T. Hashimoto, H. Yanase, N. Kato, M.C.M. Chung, M. Kataoka, and S. Shimizu, *Appl. Environ. Microbiol.* **62**, 2303–2310 (1996).

39. S.Z. Schade, S.L. Early, T.R. Williams, F.J. Kezdy, R.L. Heinrikson, C.E. Grimshaw, and C.C. Doughty, *J. Biol. Chem.* **265**, 3628–3635 (1990).
40. D. Carper, C. Nishimura, T. Shinihara, B. Dietzchold, G. Wistow, C. Craft, P. Kador, and J.H. Kinoshita, *FEBS Lett.* **220**, 209–213 (1987).
41. A. Garcia-Perez, B. Martin, H.R. Murphy, S. Uchida, H. Murer, B.D. Cowley, Jr., J.S. Handler, and M.B. Burg, *J. Biol. Chem.* **264**, 16815–16821 (1989).
42. D. Bartels, K. Engelhardt, R. Roncarati, K. Schneider, M. Rotter, and F. Salamani, *EMBO J.* **10**, 1037–1043 (1991).
43. R. Amore, P. Kotter, C. Kuster, M. Ciriacy, and C.P. Hollenberg, *Gene* **109**, 89–97 (1991).
44. J. Grunwald, B. Wirz, M.P. Scollar, and A.M. Klibanov, *J. Am. Chem. Soc.* **108**, 6732–6734 (1986).
45. W-R. Shieh, A.S. Gopalan, and C.J. Sih, *J. Am. Chem. Soc.* **107**, 2993–2994 (1985).
46. C-H. Wong, D.G. Drueckhammer, and H.M. Sweers, *J. Am. Chem. Soc.* **107**, 4028–4031 (1985).
47. B. Zhou, A.S. Gopalan, F. VanMiddlesworth, W-R. Shieh, and C.J. Sih, *J. Am. Chem. Soc.* **105**, 5925–5926 (1983).
48. S. Shimizu, M. Kataoka, M. Katoh, T. Morikawa, T. Miyoshi, and H. Yamada, *Appl. Environ. Microbiol.* **56**, 2374–2377 (1990).
49. M. Kataoka, L.P.S. Rohani, K. Yamamoto, M. Wada, H. Kawabata, K. Kita, H. Yanase, and S. Shimizu, *Appl. Microbiol. Biotechnol.* **48**, 699–703 (1997).
50. M. Kataoka, L.P.S. Rohani, M. Wada, K. Kita, H. Yanase, I. Urabe, and S. Shimizu, *Biosci. Biotechnol. Biochem.* **62**, 167–169 (1998).
51. S. Shimizu, M. Kataoka, and K. Kita, *J. Mol. Catal. B: Enzymatic* **5**, 321–325 (1998).
52. S. Shimizu, M. Kataoka, and K. Kita, in A. Laskin, G.-X. Li, and Y.-T. Yu eds., *Enzyme Engineering XIV, Annals of the New York Academy of Sciences*, Vol. 864, The New York Academy of Sciences, New York, 1998, in press.
53. S. Shimizu, M. Kataoka, A. Morishita, M. Katoh, T. Morikawa, T. Miyoshi, and H. Yamada, *Biotechnol. Lett.* **12**, 593–596 (1990).
54. Y. Makino, S. Negoro, I. Urabe, and H. Okada, *J. Biol. Chem.* **264**, 6381–6385 (1989).

See also BIOCATALYSIS DATABASES.

ALGAL CULTURE

TADASHI MATSUNAGA
HARUKO TAKEYAMA
Tokyo University of Agriculture and Technology
Tokyo, Japan

HIROYUKI TAKANO
Chichibu Onoda Cement Corporation
Chiba-prefecture, Japan

KEY WORDS

Algal biomass
CO_2 fixation
Cyanobacteria
Large-scale biomass culture systems
Microalgae
On-site CO_2 removal systems
Open culture systems
Photobioreactor
Useful material production

OUTLINE

INTRODUCTION

Photosynthetic microorganisms play an important role in the conversion of solar energy into chemical energy through their photosynthetic activities. Photosynthetic conversion can be expected to be an efficient and alternative process in several industrial fields; for example, microalgae have been studied because they produce a wide variety of metabolites.

Attempts to develop large-scale methodologies for the cultivation of microalgae have been summarized in a document produced by the Carnegie Institute of Washington (1), in which algae were predicted to provide an alternative to fermentation and agriculture products. Algal biomass has historically served as fertilizer, and as a food source for both humans and animals (2–4). They are also used for secondary wastewater treatment (5). This use of algal biomass is an important industrial application of microalgal cultures. With advances in processing technology, algal biomass has come to be seen as a possible source of fuel, as well as of fine chemicals and pharmaceuticals (6). Several microalgae that produce useful chemicals, such as amino acids, vitamins, carotenoids (β-carotene), fatty acids (DHA, EPA, γ-LA), polysaccharides, and antibiotics have been reported (Table 1). Several microalgal products have already been commercialized (7).

It is estimated that there are roughly over 10,000 species of microalgae in the natural environment. Unique algal strains may be found in highly diversified marine environments. We have discovered several marine microalgal strains that produce useful substances, including novel

Table 1. Commercial Products from Microalgae

Products	Uses	Value (dollars/kg)	Market (millions of dollars)
Phycobiliproteins	Research	10,000	1–10
	Food coloring	100	
β-Carotene	Provitamin A	500	1–10
	Food coloring	300	
Xanthophylls	Feed	200–500	10–100
Vitamin C	Vitamin	10	10–100
Vitamin E	Vitamin	50	10–100
Polysaccharides	Viscosifiers	5–10	10–1,000
Amino acids			
Proline		5–50	1–10
Arginine		50–100	1–10
aspartic acid	Gums	2–5	100–1,000
Single cell protein	Animal feeds	10–20	100–2,000

plant-growth regulators that promote redifferentiation, germination, and plantlet formation; tyrosinase inhibitors; a sulfated polysaccharide showing anti-HIV activity; novel antibiotics with light-regulated activity; and a UV-A absorbing compound (Table 2).

The development of an efficient culture system is necessary for algal mass production and the industrial application of microalgae. Hence, we describe several technologies for microalgal mass cultivation under open and closed systems. Furthermore, recent studies of applied microalgal culture investigating the possible use of biological CO_2 fixation are presented. Also discussed are examples of the technical designs for on-site CO_2 removal systems combined with algal cultures, which promise to play a role in developing technologies for environmental protection.

MICROALGAL MASS CULTIVATION TECHNOLOGIES

The growth rate and maximum biomass yield of microalgal strains are affected by culture parameters (light, temperature, and pH) and nutritional status (CO_2, nitrogen, and phosphate concentrations). On the other hand, increasing the density of cultures decreases photon availability to individual cells. Light penetration of microalgal cultures is poor, especially at high cell densities (Fig. 1), and such poor photon availability decreases specific growth rates. Higher biomass yields could be expected if sufficient photons were provided to high-density cultures of microalgae.

Large-scale culture systems have been constructed (classified as open and closed systems) with the greatest attention paid to the light supply (Fig. 2) (28). Strains such as *Chlorella*, *Scenedesmus*, *Dunaliella*, *Spirulina*, *Porphyridium*, and *Haematococcus* have been cultured using these systems to provide for several useful materials (6,29–32).

Open Systems

Several different types of open culture systems have been proposed (Fig. 2a–d). These open culture systems are the simplest method of algal cultivation and offer advantages in construction cost and ease of operation. Open culture systems require a large surface area and shallow depth (ca. 12–15 cm) to improve light utilization. Furthermore, mixing the culture prevents the cells from sinking to the bottom and distributes efficient cell growth with sunlight. The raceway pound proposed by Oswald (33) has been developed into various types, including raceway pounds employing a paddle wheel, which have most generally been used for outdoor production of microalgae (34,35). The raceway pounds for commercial production of microalgae require an area of 1,000–5,000 m^2.

Contamination by different algal species and other organisms is a serious problem in open culture systems. *Chlorella*, *Dunaliella*, and *Spirulina*, which are tolerant to extreme conditions (high nutrients, high salinity, and high pH), are adequate strains for open culture systems. Vonshak et al. (36) demonstrated that contamination by *Chlorella* in outdoor *Spirulina* culture was prevented by maintaining the culture medium at a high bicarbonate concentration (0.2 M). Grazers sometimes found in *Spirulina* culture were arrested by addition of ammonia (2 mM).

The open culture system is easily affected by weather conditions. Rain dilutes salinity, causing contamination.

Table 2. Production of Useful Materials by Marine Microalgae

Product	Strain	Amount (mg or unit*/g dry wt)	Ref.
Coccolith	*Emiliania huxleyi*	740	8–10
	Pleurochrysis carterae		
γ-Linolenic acid	*Chlorella* sp. NKG042401	9.5	11, 12
Palmitoleic acid	*Phormidium* sp. NKBG 041105	47	13
Docosahexaenoic acid (DHA)	*Isochrysis galbana* UTEX LB 2307	15.7	14
Polysaccharide	*Aphanocapsa halophytia*	45	15
Glutamate	*Synechococcus* sp. NKBG040607	15.4	16, 17
Phycocyanin	*Synechococcus* sp. NKBG042902	150	18
UV-A absorbing biopterin glucoside	*Oscillatoria* sp. NKBG 091600	0.2	19, 20
Antimicrobial compound	*Chlorella* sp. NKG 0111	—	21
Plant growth regulator	*Synechococcus* sp. NKBG042902	—	22–24
Lactic acid bacteria growth regulator	*Synechococcus* sp. NKBG040607	—	
Tyrosinase inhibitor	*Synechococcus* sp. NKBG15041c	107*	25, 26
SOD	*Synechococcus* sp. NKBG042902	—	27

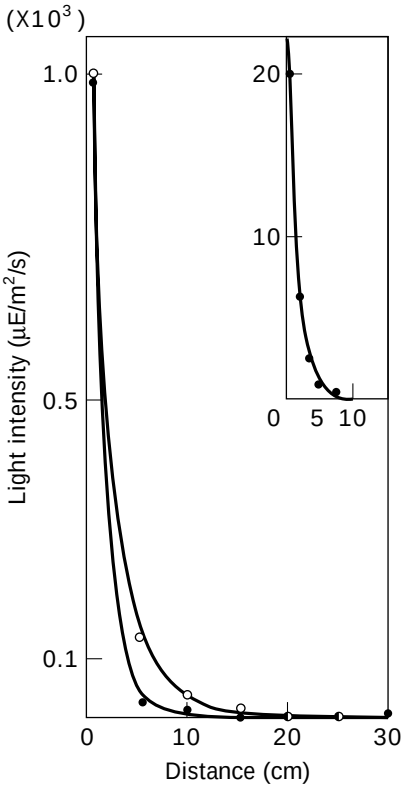


Figure 1. The distribution of light intensity as a function of distance from the surface at various cell concentrations; solid circles, 2.3×10^8 cells/ml; open circles, 7.9×10^7 cells/ml.

Furthermore, several algae that produce useful chemicals require more restricted conditions for efficient growth and for metabolite production.

Photobioreactor (Closed System)

Closed systems have been expected to overcome such open culture system disadvantages, and several photobioreactors have been devised. Figure 2e–i shows some large-scale closed photobioreactors. These photobioreactors offer several advantages: easier maintenance of monoalgal cultures because of protection from contamination; avoidance of water loss and subsequent increase of salinity in the culture medium; higher productivity with greater cell densities, reducing harvesting costs; and applicability to varying microalgal species under favorable culture conditions. However, for an efficient and reliable large-scale culture system, several criteria need to be considered (37), such as light-utilization efficiency, homogeneous mixing (turbulence), low-shear environment, temperature control, and efficient gas transfer.

Yield of algal biomass depends on the light path to each cell, and therefore, surface-to-volume ratio is an important factor for efficient light utilization in a photobioreactor. Tubular reactors have been well studied since their development by Tamiya et al. (38). The diameter of a tubular reactor is now less than 40 cm. Richmond et al. (39) demonstrated that reduction of the tube diameter from 5.0 cm to 2.8 cm enhanced the biomass yield 1.8 times. Narrower tube diameters may allow efficient light utilization as well as a faster flow rate, enhancing the algal productivity and reducing fouling on the inside walls of the tubes. In tubular reactors, flow rates of 30–50 cm/s are generally used, and air-lift is the most effective circulation method of culture, rather than a centrifugal pump, a rotary positive displacement pump, and a peristaltic pump. The main advantages of air-lift systems are their low shear and relative simplicity of construction. Furthermore, the advantage of the tubular system in biomass production per volume using *Anabaena siamensis* was shown where productivities of an open raceway (300-L capacity, height of 15 cm) and a tubular reactor with a 2.8-cm-diameter tube were 86 and 550 mg dry wt/L/day, respectively. Several modifications in tube arrangement have been carried out for optimizing light penetration. Most of the tubular reactor is laid on the ground, and the lower parts of the tubes receive less light than the upper parts. Torzillo et al. (40) have constructed a two-plane tubular bioreactor for optimizing light availability, in which each tube in the lower plane was placed in the vacant space between two tubes in the upper plane. They showed that this two-plane tubular bioreactor (145 L culture volume) has an effective surface area (ratio of surface area [m^2] to culture volume [m^3] = 49/m) and gave a net volumetric productivity of 1.5 g dry wt/L/day using *Spirulina platensis*. Furthermore, a pilot plant called the Biocoil has been set up (Fig. 2e) (41). This helical tubular reactor is constructed with a vertical tower coiled up with a tube, increasing the available efficient land surface area. When biomass productivity of various tubular reactors is compared on area basis (surface area), a similar production range of 15–30 g dry wt/m^2/day is obtained for all reactor types.

The effect of temperature on biomass yield is significant in algal culture. The culture in the tubular reactor is often maintained at higher temperature than that in the open raceway. Richmond et al. (39) demonstrated that the *Spirulina* culture using the tubular reactor could be warmed faster than in the open raceway. The culture could be kept at 35–37 °C, the optimal range for *Spirulina* growth during most of the light period. However, the closed reactors are sometimes overheated and thus are more suitable for the thermophilic or thermotolerant strains. Temperature control using a heat exchanger and/or evaporative cooling by spraying water onto the surface is sometimes required for the cultivation of general strains (40,42).

A problem in the closed system is photooxidative damage to the cells caused by accumulation of dissolved oxygen produced by photosynthesis during the light period. In the open system, evolved oxygen is diffused to the atmosphere. By contrast, because oxygen cannot escape from the closed reactors, degasser systems are sometimes required.

The culture part of closed photobioreactors has been constructed with materials such as glass, methyl polymetacrylate, polyethylene, polypropylene, polyvinylchloride, silicone, and stainless steel. These photobioreactors have been designed to utilize external illumination, such as sunlight. Fiber optics also have been employed as internal light sources (43). Photobioreactors employing fiber optics have the merit of controlling the illumination and duration of the light period. The authors have demonstrated growth

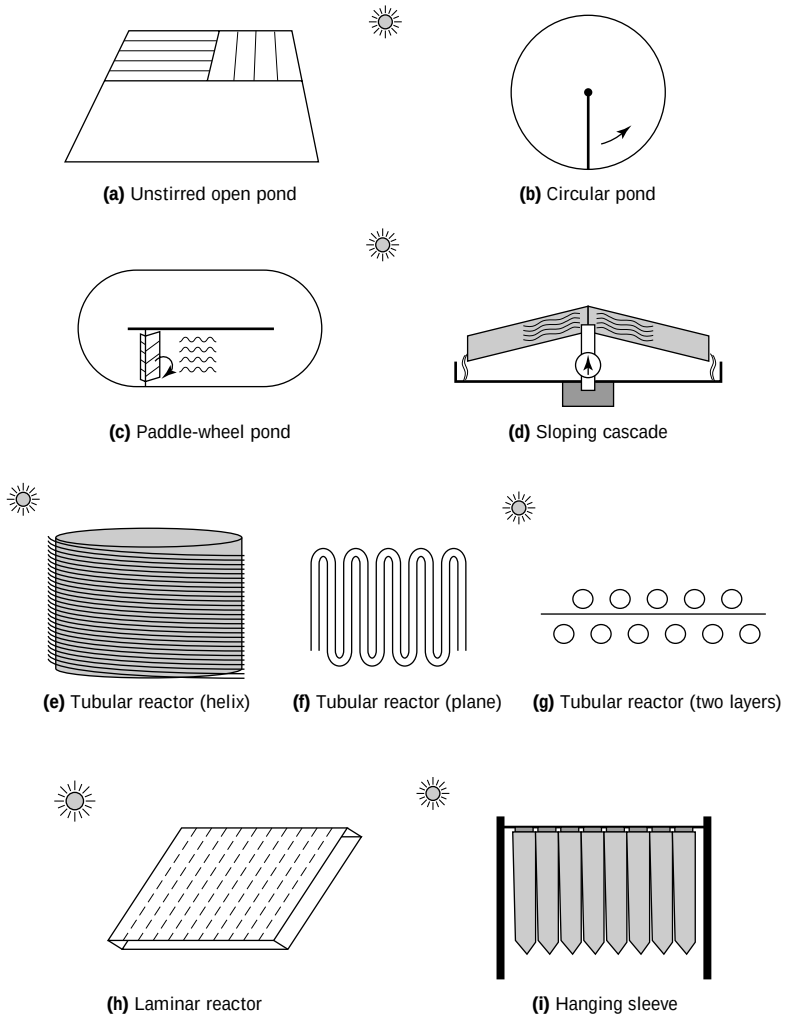


Figure 2. Illustration of algal mass culture systems.

of a high-density culture using a photobioreactor employing novel light-diffusing optical fibers (LDOF). Efficient light distribution was achieved as a result of a bundle of LDOFs passing through the cell culture (17). The ratio of surface area to volume was very high (692/m), and uniform illumination in the cell culture was obtained. Takano et al. (44) cultured a marine cyanobacterium *Synechococcus* sp. in the LDOF photobioreactor, achieving a biomass yield of 2g/L/day.

Photobioreactors are less economical but have many superior aspects to the open culture systems mentioned earlier. Thus, they are limited to high-value products. Suitable culture systems have to be adopted according to the target products and available environmental conditions.

PRODUCTION OF USEFUL MATERIALS BY RECOMBINANT CYANOBACTERIA

Higher productivity of valuable primary or secondary metabolites in microalgae will be possible by high-density cultivation and/or genetically engineered microalgae. Genetic engineering of cyanobacteria (blue–green algae) has been developed in which gene-transfer systems have been established in several cyanobacterial strains (45,46), and these have been used as hosts for production of foreign proteins (47).

Recent pharmaceutical interest in unsaturated fatty acids has triggered a search for sources of these valuable compounds. Monounsaturated fatty acids such as palmitoleic acid (C16:1), undecylenic acid (C11:1), and tridecenic acid (C13:1) have the potential to prevent several diseases.

Several eukaryotic microalgae are known to produce highly unsaturated fatty acids such as eicosapentaenoic acid (EPA; C20:5) and docosahexaenoic acid (DHA; C22:6), which are valuable dietary components (14,48). They may be involved in preventing several human diseases (49,50). However, wild-type cyanobacteria do not possess the biosynthetic pathway for these fatty acids. The EPA synthesis gene cluster (ca. 38 kbp) from the marine bacterium *Shewanella putrefaciens* SCRC-2738 was cloned into the marine cyanobacterium *Synechococcus* sp. using a broad-host cosmid vector, pJRD215 (10.2 kbp, Sm^r Km^r). The constructed plasmid (pJRDEPA; ca. 48 kbp) could be transferred by transconjugation to the cyanobacterial host at a frequency of 2.2×10^{-7}. Cyanobacterial transconjugants grown at 29 °C produced only 0.12 mg EPA/g dry wt, whereas those grown at 23 °C produced 0.56 mg/g dry wt.

Furthermore, the content of EPA in cells grown at 23 °C increased by 0.64 mg/g dry wt after preincubation for one day at 17 °C (51).

Spirulina spp. are commercially valuable cyanobacteria. Several *Spirulina* genes have been characterized (52). The genetic study of *Spirulina* has been focused on the construction of mutants and the cloning of genes. Gene-transfer systems have not been developed because *Spirulina* has high restriction endonuclease activity and is resistant to most antibiotics. Establishment of gene-transfer systems in *Spirulina* will contribute to their industrial application.

CO_2 FIXATION USING MICROALGAL CULTURES

Recently, the possible use of biological CO_2 fixation to reduce anthropogenic CO_2 emission has been investigated. For development of on-site CO_2 fixation systems using microalgae, efficient photobioreactors and strains that can fix large quantities of CO_2 are required. Recently, several applied studies aimed at the direct biological utilization of CO_2 in emission gas from power plants have been carried out. In Japan, coal-fired power plants and the steel and cement industries produce large quantities of CO_2. These industries also emit NO_x and SO_x. It is known that NO_x and SO_x reduce biological CO_2 fixation activity, primarily by inhibiting photosynthesis. Therefore, microalgae that can grow under such extreme conditions will be required for direct CO_2 fixation.

Identification of microalgal strains capable of growing rapidly in water sparged with emitted gas and under other extreme conditions, such as at high pH, has been carried out. Their ability to grow in saline water (e.g., seawater) and their resistance to other gas constituents (e.g., SO_2 and NO_x) are additional considerations. Desirable properties of microalgae for the direct biological utilization of CO_2 in gas emitted from a power plant are summarized in Table 3.Several strains have been isolated from environments that have some of the properties described above. Among those strains, the marine alga *Chlorococcum littorale* (53) is characterized by high CO_2 tolerance and high growth rate in the linear growth phase.

A Designed System for the Reduction of CO_2 Emission: Case I—Coal-Fired Power Plants

A designed system for the reduction of CO_2 emission from coal-fired power plants is presented (28). Microalgal cultivation using CO_2 in gas emitted from power plants has been described. In this system, CO_2-fixed product by microalgal culture is used as biomass fuel, which will substitute for fossil fuel (Fig. 3). CO_2 recycling in coal-fired power plants will achieve the reduction of CO_2 emission from power plants.

Cost of microalgae CO_2 mitigation using the designed system has been estimated based on several designed specifications:

1. Plant size: emission gas from a 500 MW power plant
2. Gas condition: 30.85 kg C/S (CO_2) concentration in dry flue gas: 14.2% (vol)
3. Biofixation conditions
   a. One hundred percent of the CO_2 produced by the power plant is fed to biological systems, and 90% of the CO_2 fed to the system is utilized during daylight summer hours.
   b. Direct biofixation of CO_2 is applied where CO_2 in emission gas from the power plant is directly fed to algal ponds for photosynthesis.
   c. Both 10% (case 1) and 20% (case 2) of photosynthetic efficiency based on visible light are studied as are 4.5% (case 1) and 9.0% (case 2) based on total light.
   d. The average solar radiation should be greater than 200 W/m^2.
4. Operation of the plant: The power plant is operated at 70% operation factor. That is to say, the plant is operated at 100% of rated capacity for 16.8 h (70% of 24 h) during the day and is shut down for 7.2 h (30% of 24 h) during the night.
5. CO_2 production rate
   a. Hourly CO_2 production: 407 tCO_2/h
   b. Annual CO_2 production: 2,495,724 tCO_2/year
6. Algal Strain: NANNP2 (*Nannochloropsis* sp. from the DOE/NREL/ASP culture collection) The expected productivities of algal biomass: 42 g/m^2/day (case 1) and 84 g/m^2/day (case 2)

Table 4 presents preliminary cost estimates for a large-scale (over 1,000 ha) microalgal system for biomass fuel production. Two different productivities, 42 and 84 g/m^2/day, were assumed to correspond to about 10 and 20% solar conversion efficiencies, based on visible light, and 4.5 and

Table 3. Desired Properties of Microalgae for CO_2 Fixation

1. High pCO_2 tolerance
2. Low pH tolerance
3. Stable and high growth rate in the linear growth phase
4. Capacity to grow at high cell densities
5. Acidic gas (NO_x, SO_x) tolerance
6. Thermotolerance

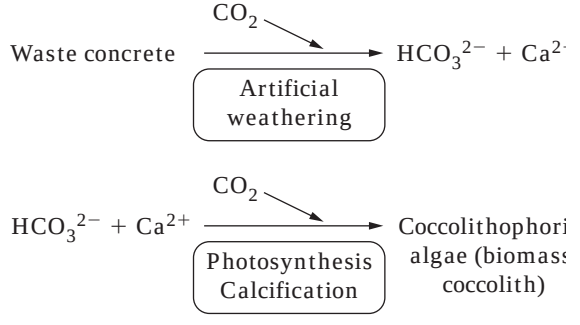


Figure 3. Design of CO_2 fixation by artificial weathering of waste concrete and culture of coccolithophorid algae.

Table 4. Capital and Operating Costs

Productivity assumption	Case 1	Case 2
Avg. ash-free dry weight (g/m^2/day) @20 Mj/m^2/day	42	84
Annual productivity (mt/ha/year)	101.9	203.8
Solar conversion efficiency (%)	10	20
CO_2 fixation (tC/ha/year)	48.4	96.8
Average annual utilization of CO_2 (%)	54	54
Growth pond area required (ha)	7,390	3,700
Land area required (500 MW power plant, growth pond × 1.2 ha)	8,900	4,500
Capital costs ($/ha)		
Ponds (earthworks, CO_2 sumps, mixing)	31,000	45,000
Harvesting (settling ponds, centrifuge)	10,500	15,000
Systemwide costs (water, CO_2 supply, etc.)	25,100	36,900
Drying	9,800	19,600
Engineering, contingency, fees	16,700	25,700
Land-related costs	3,800	5,800
Total capital cost ($/ha)	96,900	148,000
Capital costs ($tC/year)	2,002	1,529
Operation costs ($/ha/year)		
Power, nutrient, maintenance, labor	15,400	27,000
Annualized capital cost (10% DCF)	10,048	15,346
Credit for biomass fuel (82.4 $/mt)	−8,400	−16,793
Net operating cost	17,048	25,553
CO_2 mitigation cost ($/tC)	352	264

Note: Solar conversion efficiencies are for visible light.
The credit for biomass fuel is $25/barrel of oil.

9%, based on total light. The lower productivity reflects what could be expected based on present experience. The higher productivity is the theoretical limit. CO_2 mitigation costs greatly depend on the productivity of algae and solar radiation. Microalgal strains that can achieve higher photosynthetic efficiency at higher solar radiation are necessary. A photobioreactor, in which microalgae can convert from CO_2 to biomass at higher photosynthetic efficiency, is also required.

Further research into microalgae, photobioreactors, the downstream process of microalgal culture, and the burning process of microalgal biomass is still needed. If such research is successful and the economic projections shown in Table 4 are verified, this system will represent one of the most promising technologies for future CO_2 mitigation processes.

A Designed System for the Reduction of CO_2 Emission: Case II—Cement Plants

Coccolithophorids are unicellular planktonic marine algae that produce elaborate structures called coccoliths, comprising scales or plates of $CaCO_3$. In the oceans, huge blooms of coccolithophorid algae occur. These blooms have been recognized as contributing to ocean floor sediment formation, and algae play an important role in the global carbon cycle by recycling $CaCO_3$.

We have chosen coccolithophorid algae as model organism to investigate biomineralization and have focused on the ecological significance of $CaCO_3$ recycling. We have proposed CO_2 fixation by artificial weathering of waste concrete and coccolithophorid algae cultures (Fig. 4). During artificial weathering of waste concrete suspended in seawater, atmospheric CO_2 can be absorbed and dissolved as bicarbonate ions, which are a major source of coccolith particles. Coccolithophorid algae can use bicarbonate ions to form $CaCO_3$ particles. As a result, CO_2 absorbed by artificial weathering can be mineralized and fixed permanently. Artificial weathering of waste concrete is also a useful method to supply bicarbonate ions to cells of the coccolithophorid alga *Emiliania huxleyi*.

The CO_2 fixation method by artificial weathering of waste concrete and coccolithophorid algae cultures can be applied to the reduction of CO_2 emission in cement plants. A designed system for the reduction of CO_2 emission in cement plant is shown in Fig. 5. Coccoliths can be used as an alternative to limestone, which is a carbonate source for cement production. In the cement industry, CO_2 is mainly produced by the decomposition of limestone during the burning of cement clinker. If $CaCO_3$ recycling can be achieved by artificial weathering of waste concrete and coccolithophorid culture, CO_2 emission by the cement industry might be reduced (Fig. 5).

Moreover, it is proposed that microalgal biomass with fixed CO_2 products may be stored in concrete. If microalgal biomass can be stored in concrete without the decomposition of the biomass back to CO_2, removal of anthropogenic CO_2 may be achieved.

CO_2 absorption by artificial weathering of concrete has been determined. When concrete (calcium content was 15% wt/wt) pieces were weathered by CO_2 and seawater, over 310 mg CO_2/g concrete could be absorbed and dissolved in the form of HCO_3^- (54). The amount of CO_2 absorbed is directly related to calcium content in concrete, 310 mg CO_2/g concrete is equal to almost 1 g CO_2/g Portland cement (calcium content is almost 46%).

On the other hand, amount of emitted CO_2 during a cement production is 708 kg CO_2/ton cement. The amount of absorbed CO_2 by the weathering of waste concrete is greater than that of emitted CO_2 during a cement production. Moreover, CO_2 is absorbed by following coccolithophorid algae cultures.

CONCLUSION

The authors reviewed large-scale algal culture systems. Open pond systems have lower productivity of algal biomass, require larger land areas, and involve larger land costs. However, their operation cost is lower. On the other hand, closed systems can achieve high-density culture, and the volume of algal culture can be reduced. The reduction of culture volume decreases the land costs. However, a certain land area is still required for the collection of solar energy, and the operating cost is larger than that of open systems. Moreover, solar radiation, temperature, and other factors regulating algal productivity are significantly affected by location. Suitable culture systems have to be

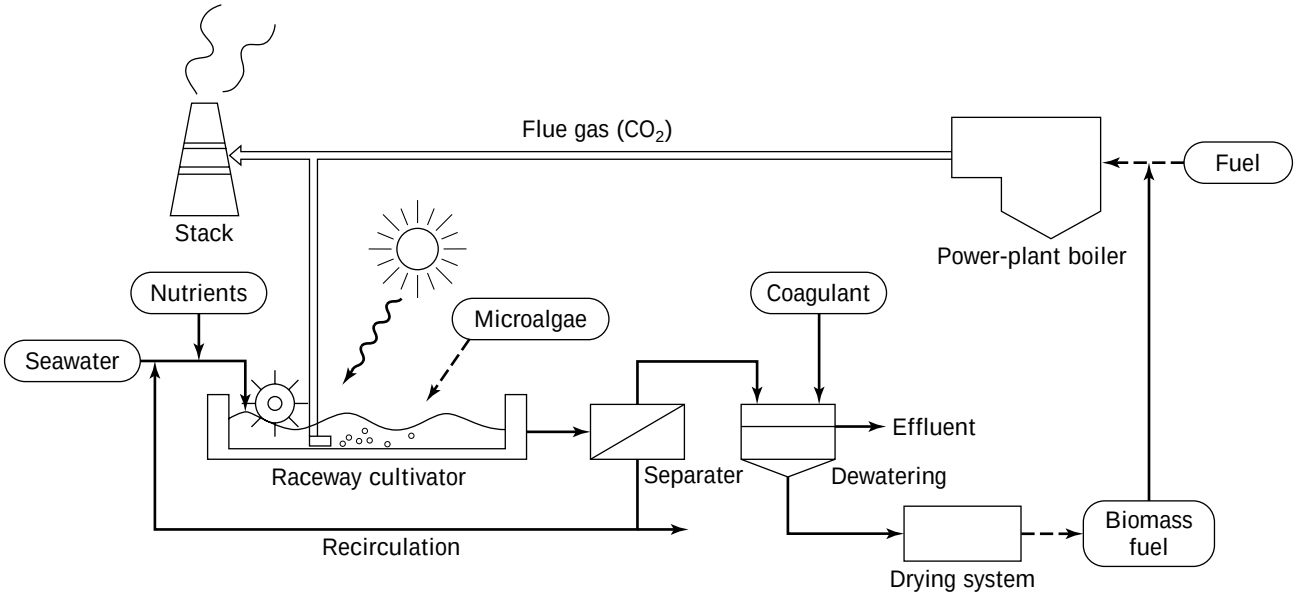


Figure 4. Conceptional system of CO_2 fixation by microalgae.

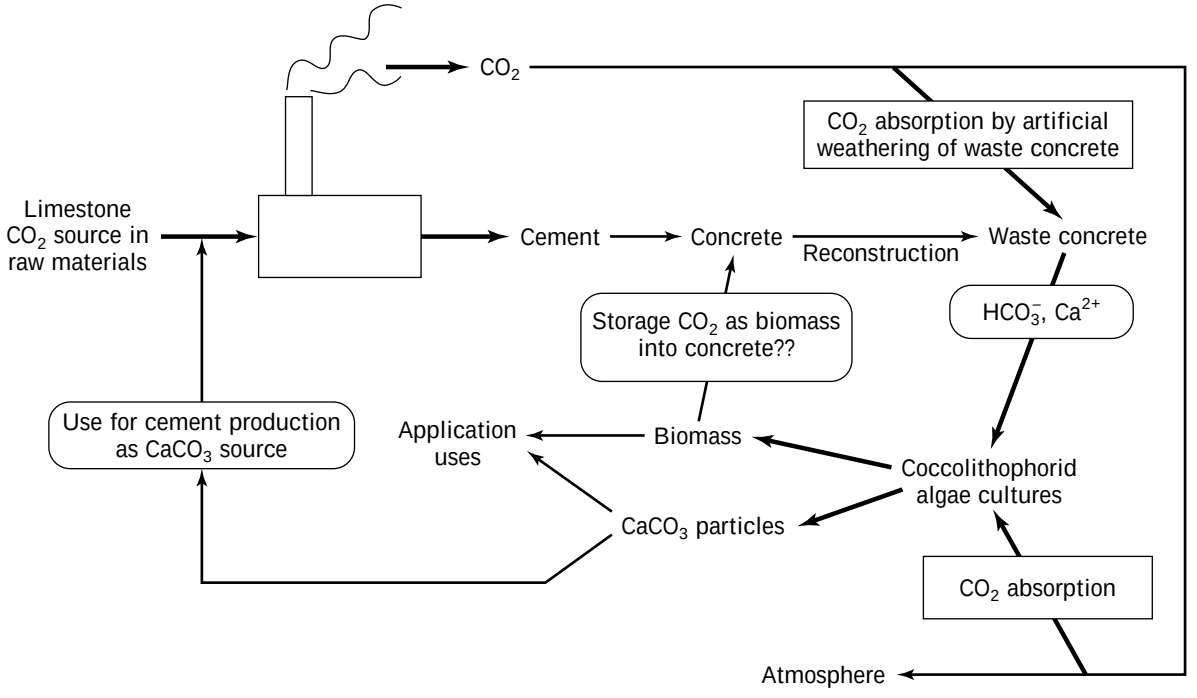


Figure 5. Design of CO_2 removal system in cement industry.

adopted according to the target products and available environmental conditions.

Extensive studies of biological CO_2 fixation using microalgal cultures have been pursued for the past 10 years. A primary goal is the complete removal of CO_2 in discharged gas emitted by such an on-site system. Because of the land-area requirements and the current CO_2 mitigation cost of $264/ton as carbon, it is to be difficult at present to apply microalgal cultures to CO_2 removal. Most nations are seriously considering the increase of atmospheric CO_2 concentration, and intensive efforts to reduce the anthropogenic CO_2 emission will be made. Microalgae culture may be one of the important processes used for such efforts (55).

Regarding technological use of microalgal biomass, several examples, including historical findings, have been pre-

sented. Microalgal production of useful chemicals and energy resources have been extensively investigated. Processes that utilize the majority of the resulting microalgal biomass as energy sources would be desired. Such processes may allow the recycling of evolved CO_2 from human energy consumption rather than a one-way emission, as is the case with fossil fuels. The following six products can be produced from microalgal biomass for use as fuels: hydrogen (through biophotolysis), methane (through anaerobic digestion), ethanol (through yeast or other alcohol fermentation), triglycerides (through extraction of lipids), methyl ester fuels (through transesterification of lipids), and liquid hydrocarbon (from *Botryococcus braunii*).

Increasing attention has been paid to resource sustainability in all industries. It must be considered whether resources used for manufacturing products can be sustainable. Developing technologies of microalgal culture will be expected to provide sustainable resources.

BIBLIOGRAPHY

1. J.S. Burlew ed., *Algal Culture from Laboratory to Pilot Plant*, Carnegie Institute of Washington. Washington, D.C., 1953.
2. B. Metting, *Micro-algal Biotechnology*, Cambridge Univ. Press, Cambridge, UK, 1988, pp. 288–304.
3. E.W. Backer, *Handbook of Microalgal Mass Culture*, CRC Press, Boca Raton, FL, 1986, pp. 339–420.
4. J.R. Benemann, *J. Appl. Phycol.* **4**, 233–245 (1992).
5. V.M. Kaya, J. de la Noüe, and G. Picard, *J. Appl. Phycol.* **7**, 85–95 (1995).
6. M.A. Borowitzka, *J. Appl. Phycol.* **7**, 3–15 (1995).
7. J.R. Benemann, D.M. Tillett, and J.C. Weissman, *Trends Biotechnol.* **5**, 47–53 (1987).
8. H. Takano, H. Furu-une, J.G. Burgess, E. Manabe, M. Hirano, M. Okazaki, and T. Matsunaga, *Appl. Biochem. Biotechnol.* **39/40**, 159–167 (1993).
9. H. Takano, J. Jeon, J.G. Burgess, E. Manabe, Y. Izumi, M. Okazaki, and T. Matsunaga, *Appl. Microbiol. Biotechnol.* **40**, 946–950 (1994).
10. H. Takano, R. Takei, E. Manabe, J.G. Burgess, M. Hirano, and T. Matsunaga, *Appl. Microbiol. Biotechnol.* **43**, 460–465 (1995).
11. M. Hirano, H. Mori, Y. Miura, N. Matsunaga, N. Nakamura, and T. Matsunaga, *Appl. Biochem. Biotechnol.* **24/25**, 183–191 (1990).
12. Y. Miura, K. Sode, Y. Narasaki, and T. Matsunaga, *FEMS Microbiol. Lett.* **107**, 163–168 (1993).
13. T. Matsunaga, H. Takeyama, Y. Miura, T. Yamazaki, H. Furuya, and K. Sode, *FEMS Microbiol. Lett.* **133**, 137–141 (1995).
14. J.G. Burgess, K. Iwamoto, Y. Miura, H. Takano, and T. Matsunaga, *Appl. Microbiol. Biotechnol.* **39**, 456–459 (1993).
15. H. Sudo, J.G. Burgess, H. Takemasa, N. Nakamura, and T. Matsunaga, *Curr. Microbiol.* **30**, 1–4 (1994).
16. T. Matsunaga, N. Nakamura, N. Tsuzaki, and N. Takeda, *Appl. Microbiol. Biotechnol.* **28**, 373–376 (1988).
17. T. Matsunaga, H. Takeyama, H. Sudo, N. Oyama, S. Ariura, H. Takano, M. Hirano, J.G. Burgess, K. Sode, and N. Nakamura, *Appl. Biochem. Biotechnol.* **28/29**, 157–167 (1991).
18. H. Takano, T. Arai, M. Hirano, and T. Matsunaga, *Appl. Microbiol. Biotechnol.* **43**, 1014–1018 (1995).
19. T. Matsunaga, J.G. Burgess, N. Yamada, K. Komatsu, S. Yoshida, and Y. Wachi, *Appl. Microbiol. Biotechnol.* **39**, 250–253 (1993).
20. Y. Wachi, J.G. Burgess, K. Iwamoto, N. Yamada, N. Nakamura, and T. Matsunaga, *Biochim. Biophys. Acta* **1244**, 165–168 (1995).
21. Y. Miura, K. Sode, Y. Narasaki, and T. Matsunaga, *J. Mar. Biotechnol.* **1**, 143–146 (1993).
22. H. Wake, H. Umetsu, K. Shimomura, and T. Matsunaga, *Plant Cell Rep.* **9**, 655–658 (1991).
23. H. Wake, A. Akasaka, H. Umetsu, Y. Ozeki, K. Shimomura, and T. Matsunaga, *Plant Cell Rep.* **11**, 62–65 (1992).
24. H. Wake, A. Akasaka, H. Umetsu, Y. Ozeki, K. Shimomura, and T. Matsunaga, *Appl. Microbiol. Biotechnol.* **36**, 684–688 (1992).
25. Y. Wachi, J.G. Burgess, J. Takahashi, N. Nakamura, and T. Matsunaga, *J. Mar. Biotechnol.* **2**, 210–213 (1995).
26. Y. Wachi, K. Sode, K. Horikoshi, H. Takeyama, and T. Matsunaga, *Biotechnol. Techniques* **9**, 633–636 (1995).
27. Y. Wachi, J.G. Burgess, J. Takahashi, T. Matsunaga, and N. Nakamura, *J. Mar. Biotechnol.* **3**, 258–261 (1995).
28. The IEA Greenhouse Gas R&D Programme, Report G/93/OE16B, Chemical Society of Japan, 1994.
29. A. Ben-Amotz, and M. Avron, *Trends Biotechnol.* **8**, 121–126 (1990).
30. A. Richmond, *Micro-algal Biotechnology*, Cambridge Univ. Press, Cambridge, UK, 1988, pp. 85–121.
31. A. Vonshak, *Micro-algal Biotechnology*, Cambridge Univ. Press, Cambridge, UK, 1988, pp. 122–134.
32. P. Bubrick, *Bioresource Technol.* **38**, 237–240 (1991).
33. W.J. Oswald, *Chem. Eng. Symp. Ser.* **65**, 87–92 (1969).
34. J.C. Dodd, *Handbook of Microalgal Mass Culture*, CRC Press, Boca Raton, FL, 1986, pp. 265–283.
35. A. Richmond, *J. Appl. Phycol.* **4**, 281–286 (1992).
36. A. Vonshak, S. Boussiba, A. Abeliovich, and A. Richmond, *Biotechnol. Bioeng.* **25**, 341–351 (1983).
37. M.A. Borowitzka, *J. Mar. Biotechnol.* **4**, 185–191 (1996).
38. H. Tamiya, E. Hase, K. Shibata, K. Mitsuya, T. Iwamura, T. Nihei, and T. Sasa, *Algal Culture from Laboratory to Pilot Plant*, Carnegie Institute of Washington, Washington, DC, 1953, pp. 204–232.
39. A. Richmond, S. Boussiba, A. Vonshak, and R. Kopel, *J. Appl. Phycol.* **5**, 327–32 (1993).
40. G. Torzillo, P. Carlozzi, B. Pushparaj, E. Montaini, and R. Materassi, *Biotechnol. Bioeng.* **42**, 891–898 (1993).
41. L.J. Borowitzka and M.A. Borowitzka, *Algal and Cyanobacterial Biotechnology*, Longman Scientific, London, 1989, pp. 294–316.
42. Y.K. Lee and C.S. Low, *Biotechnol. Bioeng.* **40**, 1119–1122 (1992).
43. K. Mori, *Biotechnol. Bioeng.* **15**, 331–345 (1985).
44. H. Takano, H. Takeyama, N. Nakamura, K. Sode, J.G. Burgess, E. Manabe, M. Hirano, and T. Matsunaga, *Appl. Biochem. Biotechnol.* **34/35**, 449–458 (1992).
45. T. Matsunaga and H. Takeyama, *J. Appl. Phycol.* **7**, 77–84 (1995).
46. J. Elhai, *J. Appl. Phycol.* **6**, 177–186 (1994).
47. R.C. Murphy and S.E. Stevens, Jr, *Appl. Environ. Microbiol.* **58**, 1650–1655 (1992).
48. H. Takeyama, K. Iwamoto, S. Hata, H. Takano, and T. Matsunaga, *J. Mar. Biotechnol.* **3**, 244–247 (1996).

49. T. Terano, J.A. Salmon, and S. Moncada, *Prostaglandins* **27**, 217–232 (1984).
50. M. Abbey, P. Clifton, M. Kestin, B. Belling, and P. Nestel, *Arteriosclerosis* **10**, 85–94 (1990).
51. H. Takeyama, D. Takeda, K. Yazawa, A. Yamada, and T. Matsunaga, *Microbiology* **143**, 2725–2731 (1997)
52. A.K. Vachhani and A. Vonshak, *Spirulina platensis (Arthrospira): Physiology, Cell-Biology and Biotechnology*, Taylor and Francis, London, 1997, pp. 67–77.
53. M. Kodama, H. Ikemoto, and S. Miyachi, *J. Mar. Biotechnol.* **1**, 21–25 (1993).
54. H. Takano and T. Matsunaga, *Energy Conversion Manage.* **36**, 697–700 (1994).
55. M. Murakami and M. Ikenouchi, *Energy Conversion Manage.* **38S**, 493–497 (1997).

AMINO ACIDS, GLUTAMATE

TETSUYA KAWAKITA
Ajinomoto Co., Inc.
Kanazawa, Japan

KEY WORDS

Acceptable daily intake (ADI)
Brevibacterium
Corynebacterium
Fermentation process
Glutamic acid
Polymorph of crystals of glutamic acid α
Pyrroridon carboxylic acid (PCA)
Synergy effect
Umami seasoning

OUTLINE

INTRODUCTION

Glutamic acid was isolated from wheat gluten hydralyzate by Ritthousen in 1866. Wolff carried out the first chemical synthesis of glutamic acid in 1890. In 1908, Ikeda (1) isolated the specific taste fraction from the kelplike seaweed *Laminariaceae*. Kelp, called konbu in Japan, is traditionally used as a soup-stock component with dried bonito, and has been identified as the salt of glutamate, used to enhance the flavor of foods. He also recognized it as "umami" taste, which does not belong to the four traditional taste categories; bitter, sour, salty, and sweet. Immediately after, Ikeda filed a patent application and began the commercial production of monosodium L-glutamate monohydrate (MSG) with his partner, Suzuki, in 1909.

With the increasing demand for MSG is use as a seasoning, biochemical procedures were developed for its production. For example, α-ketoglutaric acid was first fermented from glucose by using a species of *Pseudomonas* (2), followed by the conversion with amination to L-glutamate in a high yield. The vigorous study on the accumulation of L-glutamate in bacterial cultures by Kinoshita et al. (3) opened the door to the microbial production of L-glutamate and other amino acids.

The annual amount of production was estimated at more than 1,000,000 tons in 1997, an amount that increases by about 5% globally every year. Production is increased because of the demand due to changes in eating habits, and its use in commercially prepared, packaged foods; ready-made sours; and table seasonings in both Western and Eastern countries.

METABOLISM

L-glutamic acid is split into α-ketoglutaric acid and ammonia by glutamate dehydrogenase. By the reverse reac-

tion, L-glutamic acid is synthesized from α-ketoglutaric acid, a member of the tricarboxylic acid (TCA) cycle of glycolysis. Whereas glutamate transaminase is nonspecific for the pairs of keto acids and amino acids, L-glutamate is the only amino acid in mammalian tissues that undergoes oxidative deamination at an appropriable rate. The formation of ammonia from α-amino groups requires their conversion to the α-amino nitrogen of L-glutamate. Thus, L-glutamic acid is a key substance in the nitrogen metabolism of amino acids. Glutamate dehydrogenase is widely distributed in microorganisms and higher plants as a catalyst in the synthesis of L-glutamic acid from α-ketoglutaric acid and free ammonia. Transaminase is contained in a wide variety of microorganisms.

MICROBIAL PRODUCTION

Although various kinds of microorganisms accumulate amino acids in culture medium, only bacteria have sufficient productivity to warrant the commercial production of amino acids. Since amino acids are essential components of microbial cells, and their biosyntheses are teleologically regulated to maintain an optimal level, they are normally synthesized in limited quantities and subject to negative feedback regulation. It is necessary to overcome this regulation to achieve overproduction of amino acids. Microbial amino acid overproduction can be achieved using the following procedures and has been attained by introducing mutation techniques (4):

1. Stimulation of cellular uptake of the starting materials
2. Hindrance of the side reaction
3. Stimulation of the formation and the activity of the enzymes for biosynthesis
4. Inhibition or reduction of the enzyme concerned with the degradation of the amino acids produced
5. Stimulation of the excretion of the product into the extracellular space

L-Glutamate production by a wild-type strain of *Corynebacterium* was explained by the release of L-glutamate regulation and efflux of L-glutamate through the cell membrane by the limitation of biotin or the addition of penicillin ester of fatty acids to the culture medium (5).

Fermentation Method

Excretion of L-amino acids by *Escherichia coli* when an excess of ammonium salt was added was reported in 1950 (6). In 1957, a fermentation process was successfully commercialized (3) that used a strain of bacteria isolated from soil, later found to be *Corynebacterium glutamicum*, which is able to excrete considerable amounts of L-glutamic acid.

Numerous microorganisms have been isolated and found to be able to overproduce L-glutamic acid, including *Brevibacterium flavum, B. lactofermentum,* and *Microbacterium ammoniaphlum*. Because of minor differences in the character of these bacteria, which are all Gram positive, non–spore forming, nonmotile, and require biotin for growth, the name of genus *Corynebacterium* was suggested for these coryneform bacteria (7).

For production of L-glutamic acid, the key factors in controlling the fermentation are as follows:

1. The presence of biotin in the range of 5 to 10 μg/L, which is optimal for the excretion of L-glutamic acid through cell walls
2. A sufficient supply of oxygen to reduce the accumulation of lactic acid and succinic acid as by-products

When a biotin-rich medium, such as molasses, is used as the carbon source, the addition of penicillin or cephalosporin C favors the overproduction of L-glutamic acid, supposingly due to the repression of peptide glycan synthesis on the cell membrane. The supplementation of a fatty acid or surfactant also results in an increased permeability of the cell wall, thus enhancing L-glutamic acid excretion. Kramer (8) reports the existence of specific carrier mechanisms that are responsible for active L-glutamic acid secretion in *C. glutamicum*.

Generally, the intracellular accumulation of L-glutamic acid does not reach levels sufficient for feedback control in glutamate overproducers due to rapid excretion of glutamate. However, the regulatory mechanisms of L-glutamic acid biosynthesis have been studied intensively to obtain mutants with increased productivity, as shown in Figure 1.

Shiio (9) discussed two enzymes that have played key roles in the biosynthesis of L-glutamic acid. The first, phosphoenolpyruvate carboxylase, catalyzes carboxylation of phosphoenolpyruvate to yield oxaloacetate; it is inhibited by L-aspartic acid and repressed by both L-aspartic acid and L-glutamic acid. Second, α-ketoglutaric acid dehydrogenase converts α-ketoglutaric acid to succinyl-CoA; in L-glutamic acid–overproducing strains, α-ketoglutaric acid dehydrogenase limits further oxidation of α-ketoglutaric acid to carbon dioxide and succinic acid, thus favoring the formation of L-glutamic acid.

In L-glutamic acid–overproducing strains, the K_m value of α-ketoglutaric acid dehydrogenase was nearly two magnitudes lower than that of L-glutamic acid dehydrogenase, which catalyzes the last step to L-glutamic acid. Consequently, v_{max} of L-glutamic acid dehydrogenase proved to be about 150 times higher than that of α-ketoglutaric acid dehydrogenase. Borman (10) isolated and characterized the *C. glutamicum* glutamic acid dehydrogenase.

A strain of *Microbacterium ammoniaphlum* cultured under biotin-deficient conditions produced 58% of L-glutamic acid formed from glucose via phosphoenolpyrvate, citrate, and α-ketoglutaric acid; the other 42% was formed via the TCA cycle or glyoxylate cycle (11).

The mutants are reported as either having sensitivity in cell permeability (12), having the capability for increased carbon dioxide fixation (13), or having a too-low activity level of pyruvate dehydrogenase (14).

Another approach focused on the development of thermophilic mutants. A strain of *Corynebacterium thermoaminogenes* is reported to accumulate L-glutamic acid at temperatures above 43 °C (15).

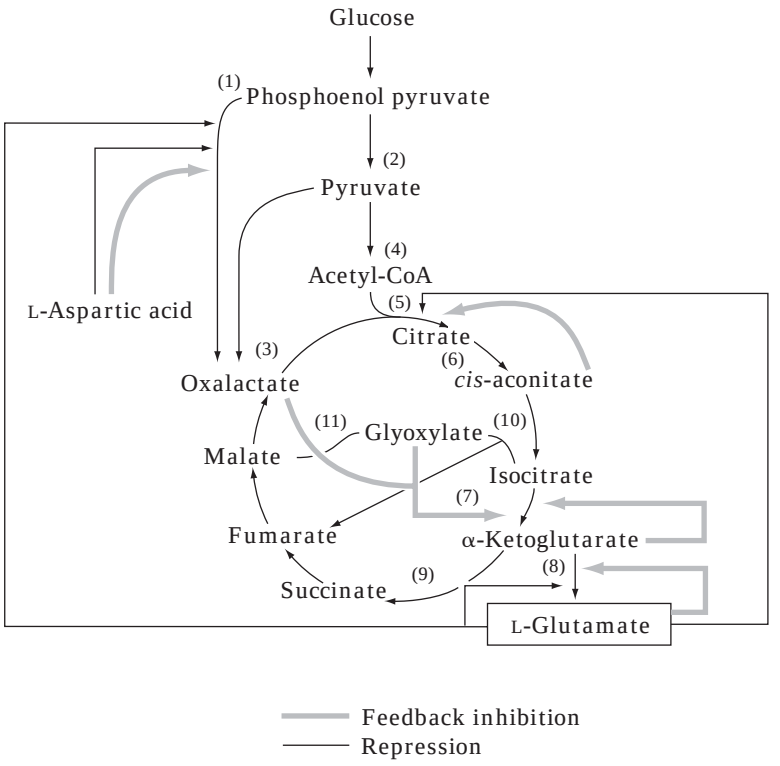


Figure 1. Regulation of L-glutamic acid biosynthesis in *Corynebacterium glutamicum* (5). Regulatory enzymes: 1, phosphoenolpyruvate carboxylase; 2, pyruvate kinase; 3, pyruvate carboxylase; 4, pyruvate dehydrogenase; 5, citrate synthetase; 6, aconitase; 7, isocitrate dehydrogenase; 8, L-glutamate dehydrogenase; 9, α-ketoglutarate dehydrogenase; 10, isocitrate lyase; 11, malate synthetase.

CHEMICAL PROPERTIES

Dissociation Constant

Amino acids exit aqueous solutions as dipolar ions (zwitterions), and glutamic acid may be represented as follows (16). Following Brönsted, an acid may be defined as a substance capable of combining with a proton. When an amino acid is titrated with HCl and NaOH, a curve exhibiting two points of inflection is obtained. Thus amino acid can react either as an acid or as a base, that is, as an ampholyte.

The dissociation constants for glutamic acid are pK_1 2.19 (COOH), pK_2 4.25 (COOH), and pK_3 9.67 (NH_3^+). (It is customary to designate the pK' values of amino acids in order of decreasing acidity.) The isoelectric point (pI) of glutamic acid is the pH 3.2 calculated from the relation $pI = pK_1 + pK2/_2$, at which there is no net charge on the amino acid molecule. At this pH, no migration will occur in an electric field.

Racemization

Racemization proceeds, in general, in alkali solutions as shown in Figure 2, such that α-H liberated from α-C by a base-catalyzed reaction yields carboion (C^-—C═O) on α-C, which equilibrates with α,β unsaturated from C═C—O^-. In the case of acidic solutions, racemization proceeds the same as in alkali, via carboion, C^-—C═OH^-, which equilibrates with α,β unsaturated from C═C—O^-.

Racemization in neutral solution occurs at 190 °C with the loss of taste characteristics after the formation of the lactam and pyroglutamic acid. This reaction is very slow, but the equilibrium shifts toward the lactam rather than

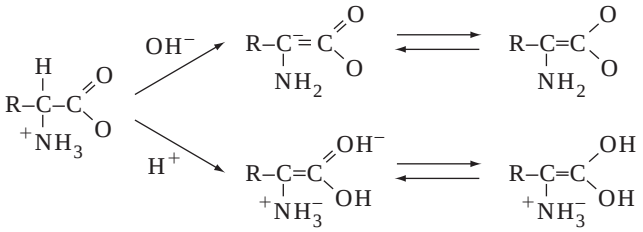


Figure 2. Mechanism of racemization of amino acid.

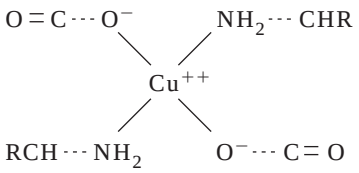
glutamic acid. Under strongly acidic or alkaline conditions, the ring-opening reaction requires a very short time; therefore, neutralization of L-glutamic acid should be performed cautiously to avoid intramolecular dehydration even below 190 °C (17).

Racemization is also brought about by a microbial racemase. As in other amino acids, the racemase of glutamic acid is found in *Lactobacillus fermenti*. This enzyme acts specifically on glutamic acid but cannot act on its simplest derivatives (18).

PROPERTIES OF L-GLUTAMIC ACID

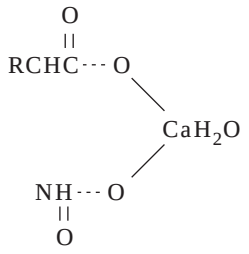
Salt Formation

An amino acid can react with a copper ion to form a copper complex of the following type:



Formation of the copper complex is the basis of a quantitative procedure for the determination of the amino acid.

In alkaline solution, amino acids react with carbon dioxide to form carbamino acids, which may be precipitated as salts:



Some of these are quite insoluble and have been used for isolation of amino acids. Decomposition of carbamino acid occurs by boiling in aqueous solution.

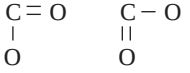
L-Glu is able to react with various kind of cations, especially with metallic ions to form their corresponding specific complexes. The complex constant (log K_s) for various complexes of L-Glu is 10.3 for Ni^{2+}, 8.5 for Zn^{2+}, 8.1 for Co^{2+}, 4.6 for Fe^{2+}, and 3.0 for Mn^{2+} (19). Glutamic acid reacted with sodium ion to form monosodium L-glutamate, on the other hand, reacts as a base with sulfuric acid to form diglutamate sulfate crystals.

Infrared Spectra

Comparison of the infrared (IR) spectra of L-Glu, L-GluHCl, and L-GluNa H_2O are shown in Figure 3 (20). The IR spectra of L-Glu shows strong bands at 1,620 and 1,420 cm^{-1}, which correspond to the COO; these are 1,660 cm^{-1} for COOH and 2,740 cm^{-1} for NH_3^+, respectively. The spectra of L-Glu HCl, on the other hand, does not show bands at 1,620 and 1,420 cm^{-1} for COO^- and that of L-Glu NaH_2O does not show the band at 1,670 cm^{-1} for COOH, but does show a strong band at 1,620 and 1,420 cm^{-1} for COO^-.

Ionic Forms

The four ionic forms of L-Glu exist depending on the pH of the solution. The dipolar nature of amino acids in solution has been shown by Raman spectroscopy. The Raman spectra of L-Glu shows a strong line at 1730 cm^{-1}, which corresponds to the C═O vibration. The spectrum of sodium L-Glu^-, on the other hand, does not show this frequency because the salt is completely dissociated and its structure represented is by the resonant forms



This does not have a C═O group, but instead intermediate

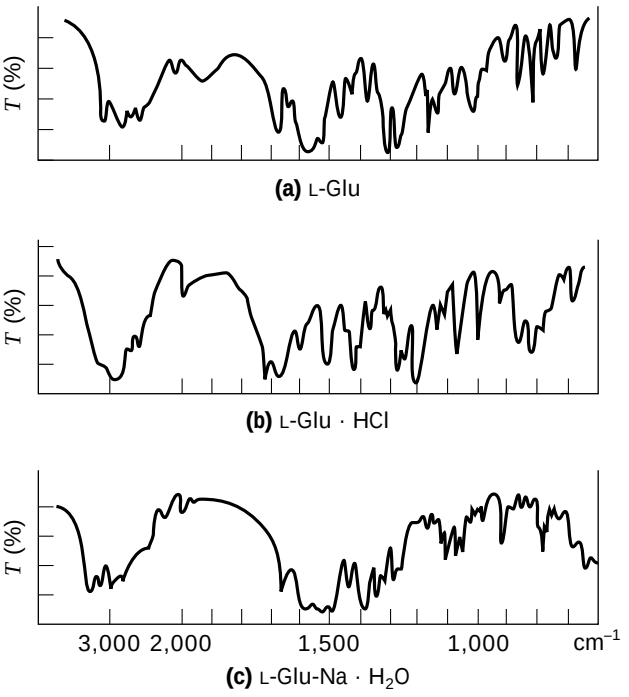


Figure 3. IR spectra of L-glutamic acid– related substances.

some between C—O and C═O. The spectra of the amino acids at the isoelectric point of pH do not show the existence of C═O, which indicates that the —COOH group is ionized. In a similar way, the ionization of the $—NH_2$ is indicated.

The optical rotation changes with changing pH in solution at 25 °C; glutamic acid $[\alpha]_D(G^{\pm}; H_2O)$ is +12.0, $[\alpha]_D(G^+; 5N\ HCl)$ is +31.8, $[\alpha]_D(G^-; NaOH)$ is −4.2, and $[\alpha]_D(G^=; NaOH)$ is +10.9, which corresponds to the ratio of ionic form in solution (21).

Solubility

For the production of the umami seasoning MSG, the characteristic solubility changes in L-Glu HCl, L-Glu, and L-GluNa are used to separate it from other impurities to meet the category of food additives. The change in solubility of L-Glu for α- and β-forms, respectively, with temperature are expressed by the following (22):

$$\log S = 0.0174t - 0.377 \text{ (0–30 °C) for } \alpha\text{-form}$$

$$\log S = 0.0153t - 0.328 \text{ (30–70 °C) for } \alpha\text{-form}$$

$$\log S = 0.0159t - 0.461 \quad \text{for } \beta\text{-form}$$

The solubility of the β-form is lower than that of the α-form throughout the temperature range measured. It follows that the β-form is a stable form from the aspect of thermodynamic kinetics. In Table 1, the solubility for various salts of L-Glu versus temperature are summarized (23).

The solubility changes with change in the concentration of both hydrochloric acid and sodium hydroxide in Figure 4, which corresponds to the existence of solid crystals, in

Table 1. Solubility of Various Salts of L-Glu (g/L)

Temperature (°C)	L-Glu	DL-Glu	L-GluNa	DL-GluNa	L-GluHCl	DL-GluHCl
0	3.41	8.55	514	158	298	471
25	8.64	20.5	627	243	479	698
50	21.9	49.3	765	372	769	1,030
75	55.3	119	933	570	1,240	1,540
100	140	285	1,140	875	1,990	2,280

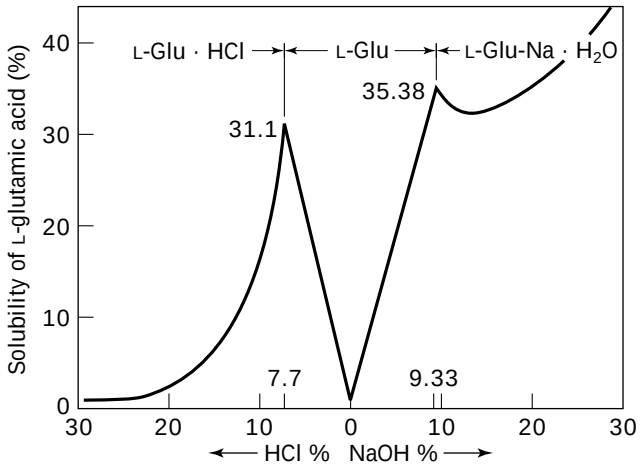


Figure 4. The change in the solubility of L-glutamic acid with change in the concentration of both hydrochloric acid and sodium hydroxide.

equilibrium solution, of the hydrochloride, free, and sodium salts of L-Glu, respectively (i.e., in the range of 0 to 30% for HCl and of 0 to 20% for NaOH). In the range of 0 to 7.7% of HCl, the solubility increases linearly so that L-Glu is soluble until the equivalent amount of HCl in solution in which L-Glu exists as solid crystals. Above the concentration of 7.7% HCl, with an increase in the HCl concentration, the solubility decreases steeply. It flattens in the vicinity 20% and then becomes a constant of about 1% above 25% HCl. Above 7.7% HCl, the hydrochloride is a solid crystal in solution. At the invariant point where equilibrium is attained between L-Glu and L-Glu HCl, the maximum solubility of 31.1% is obtained.

On the other hand, in the range of 0 to 9.33% of NaOH, L-Glu solvates an equal amount of NaOH in solution and exists as a solid crystal. At the invariant point (NaOH 9.33%), the solubility of L-Glu reaches the maximum of 35.38%. Above the concentration of 9.33% NaOH, L-GluNa H_2O exists as solid crystals. A slight decrease in solubility is observed in the range from 9.33 to about 13%, and then increases with increasing concentration of NaOH. The minimum value of the solubility of L-Glu is obtained from the range of pH 2 to 4, in the vicinity of the pI.

L-GLUTAMIC ACID (L-GLU)

$$\begin{array}{c} HOOCCH_2CH_2CHCOOH;\ \text{L-Glu} \\ | \\ NH_2 \end{array}$$

Chemical formula: $C_6H_9O_4N$ Molecular weight: 147.14

C = 40.82, H = 6.17, O = 43.50, N = 9.52 (%)

Crystal: L-Glu is a polymorph, having two crystal forms, α and β. Both are anhydrous.

Crystals of the α-form are obtained by crystallization from the saturated solution of L-Glu at 60 to 70 °C by cooling rapidly with agitation or from the acidic or alkalic solution of L-Glu to which alkali or acid was added to rapidly neutralize it until the isoelectric point of L-Glu was reached (pH 3.2).

The crystal shape is a column or pyramid. The α-form crystals transform gradually to β-form when kept in solution for a long time at room temperature. β-form crystals are obtained from either the saturated solution of L-Glu at 80 to 90 °C, which is cooled gradually with agitation, or the relatively higher concentration of L-Glu hydrochloride or sodium salt, to which alkali or acid was added to neutralize it slowly until pH 3.2 was reached. The crystal shape is a needle or thin plate (24).

Crystal Structure

The α- and β-forms of the crystals belong to the orthorhombic crystal group, and the space group is $P2_12_12_1$ with the crystal-lattice constant for both crystals as follows (25): a = 7.06, b = 10.3, and c = 8.75 Å for the α-form and a = 5.17, b = 17.34, and c = 6.95 Å for the β-form. The molecular arrangement obtained from an X-ray diffraction for two crystals are shown in Figure 5. From the crystal-structure analysis (25), L-glutamic acid in both forms exists as a zwitterion; a proton transfers from the α-carboxyl group to the amino group. The C(1)-O(1) and C(1)-O(2) distances of 1.24 and 1.27 Å, for both crystal forms, are the values of the ionized carboxyl group. This discrepancy of 0.03 Å found in both crystal forms can be attributed to differences in their intermolecular hydrogen bonds. O(2) makes a very short hydrogen bond with the OH group and a NH—O hydrogen bond with the NH_3^+ group, while O(1) makes one NH—O bond with NH_3^+.

In the γ-carboxyl group, the length of C(5)-O(4) and C(5)-O(3), 1.21 and 1.37 Å for the α-form, and 1.19, and 1.24 Å for the β-form, respectively, indicate that it is un-ionized in both crystal forms.

The bond angles around C(2), C(3), and C(4) in the α-form are significantly different from the corresponding ones in the β-form. This discrepancy may be a result of the steric repulsions that are due to the conformational differences. The OH—O bond seems to be most predominant of the four hydrogen bonds. It is reasonable to assume that considerable amounts of molecules are connected by OH—O hydrogen bonds to form chains in solution. The neigh-

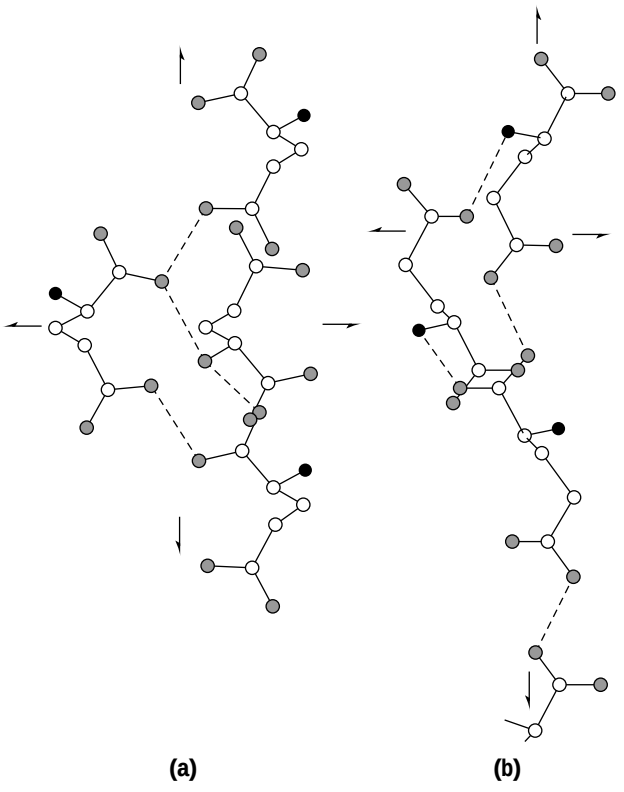


Figure 5. The molecular arrangement of L-glutamic acid crystals: a, α-form; b, β form; Dashed line shows the hydrogen bond systems in the crystals, ○, carbon; ◍, oxygen, ●, nitrogen.

boring chains can be connected by the NH—O hydrogen bonds, and the conformation and the arrangement of the molecules are stabilized by these hydrogen bonds.

Stability

When heated as a dried crystal to 150 °C, L-Glu undergoes cyclization and is converted to an interdehydrated compound of pyrrolidone carboxylic acid. Moreover, heated to 180–200 °C, the optically inactive DL-pyrrolidone carboxylic acid is obtained. At greater than 200 °C, glutamic acid begins to decompose (26). At pH 4 or 10 when the solution is heated at 100 °C for about 50 h, more than 98% of L-Glu undergoes cyclization. At pH 3 and when heated at 120 °C for 3 h, 100% cyclization is observed.

The melting point of L-pyrrolidone carboxylic acid is 160 to 161 °C, and the specific optical rotation in water is −11.35 at 20 °C.

This cyclization is a reversible reaction when treated with strong acid to yield glutamic acid, or when heated with 2N HCl or 0.5N NaOH at 100 °C for 1 to 2 h.

L-GLUTAMIC ACID HYDROCHLORIDE

Chemical formula: $C_5H_{10}O_4NCl$ Molecular weight: 183.61

C = 32.71, H = 5.49, O = 34.86, N = 7.53, Cl = 19.33 (%)

L-Glu HCl forms transparent crystals that crystallize from the high concentration of hydrochloric acid of L-Glu by cooling. The melting point is 160 °C and optical rotation calculated from the ionic form (G^+) is +25.65.

The relationship between the solubility of L-GluHCl and the temperature is (27):

$$\log S = 1.474 + 0.00825t$$

The solubility of L-glutamic acid hydrochloride crystals at 20 and 30% HCl with changing temperature is shown in Table 2.

As mentioned earlier, with an increase in the concentration of HCl above 17%, the solubility of hydrochloride decreases and at more than 30% becomes constant. Conversely, when the temperature is lowered, the solubility decreases sharply. For example, at 3 °C in 30% HCl, the solubility is only 0.38%.

It is profitable when crystallizing L-GluHCl crystals to allow the solution to cool to as low a temperature as possible. The crystal of DL-GluHCl is racemic compound, and it is possible to resolute by the seed method. The solubility of DL-Glu HCl is $\log S = 1.673 + 0.00685t$.

SODIUM SALTS OF L-GLUTAMIC ACID

Monosodium L-glutamate (MSG), L-GluNa

$$\begin{array}{l} HOOCCH_2CH_2CHCOONa\ H_2O \\ \qquad\qquad\quad | \\ \qquad\qquad NH_2 \end{array}$$

Chemical formula: $C_5H_8O_4N\ Na\ H_2O$

Molecular weight: 187.13

C = 32.09, H = 4.31, O = 34.20,
N = 7.49, Na = 12.29, H_2O = 9.63(%)

The monosodium salt of L-Glu crystallizes generally as a monohydrate. The melting point is 195 °C and dehydration temperature is 150 °C.

Crystal Structure

This crystal belongs to the orthorhombic crystal group $P2_12_12_1$. The crystal lattice constant is a = 15.30, b = 17.86, and c = 5.54 Å, and the crystal density 1.635 g/cm^3. The packed specific volume is 1.20 to 1.25 cm^3/g.

Table 2. The Solubility of L-GluHCl for 20% and 30% of HCl

	Hydrochloric acid	
Temperature (°C)	20%	30%
3	1.15	0.38
7	1.31	0.44
10	1.48	0.47
15	1.38	0.58
25	2.60	1.16
35	5.29	1.19
50	8.42	2.68

The crystal shape is an octahedral column and has bilateral planes at both termini where the atomic coordination is unstable, causing piezoelectricity. The crystal shows weak hygroscopic character with increasing temperature, as summarized in Table 3, which shows the relationship between the temperature and critical humidity, although the degree of hygroscopicity was relatively weak, when compared with that of NaCl. It is easily soluble in water, and the solubility S is designated by the following relation (28):

$$S = 35.30 + 0.098t + 0.012t^2$$

where S is g/100 g H_2O and t is in degrees centigrade.

L-Glutamic Acid Monosodium Salt Pentahydrate

$$\underset{\displaystyle \overset{|}{NH_2}}{HOOCH_2CH_2CHCOONa}\ 5H_2O$$

Chemical formula of pentahydrate: $C_5H_8O_4Na\ 5H_20$
Molecular weight: 259.5

The condensed solution of L-Glu Na was cooled between −0.8 and −8.5 °C, from which pentahydrate crystals were crystallized (29). The crystal structure belongs to the triclinic and space group P1, and lattice constant is (30)

$$a = 6.12,\ b = 16.06,\ \text{and}\ c = 6.01\ \text{Å}$$
$$\alpha = 99.3,\ \beta = 100.9,\ \gamma = 99.1$$

The crystal density is 1.480 g/cm^3.

As shown in Figure 6, the transition point form monohydrate to pentahydrate is −0.8 °C. Above −0.8 °C monohydrate crystals exist as the stable form, between −0.8 and −8.5 °C, pentahydrate exists as the stable form, and below −8.5 °C the solution itself turns into ice. The solubility of pentahydrate is changeable with temperature compared to monohydrate, and is designated as follows:

$$S = 35.72 + 0.630\ t + 0.0016\ t^2$$

Table 3. Critical Humidity of L-GluNa and NaCl

	Critical humidity (%)	
Temperature (°C)	MSG	NaCl
10	96.1	74.9
20	96.0	75.3
30	94.8	75.1
40	92.6	75.2
50	90.0	75.2

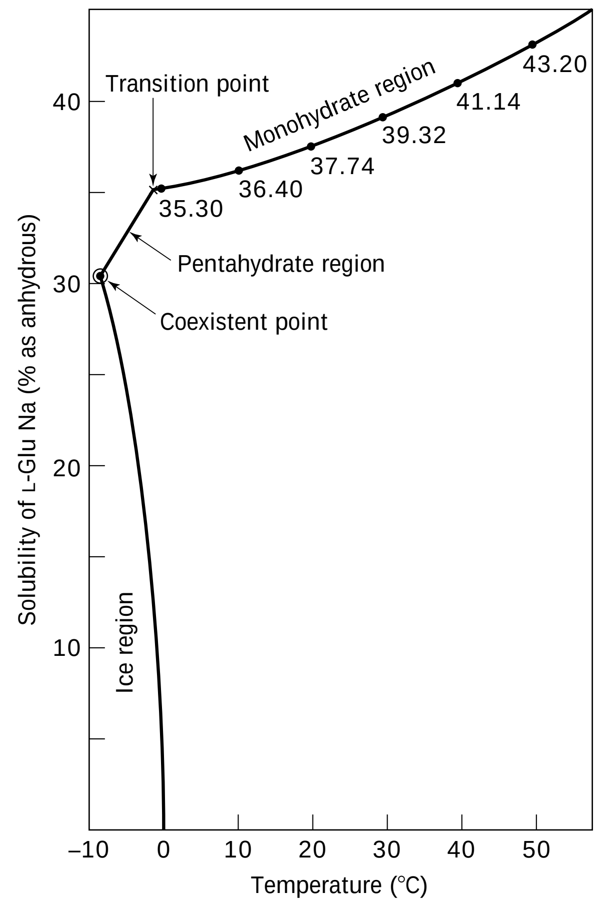


Figure 6. The solubility of sodium L-glutamate with change in temperature.

CALCIUM SALT OF L-GLUTAMIC ACID

Several forms of calcium salt are reported.

$(Glu)_2$-CaH_2O: An equimolar amount of $Ca(OH)_2$ was added to L-Glu solution to crystallize. This is soluble in water.

(Glu)-CaH_2O: L-Glu HCl in solution reacts with $Ca(OH)_2$ at 80 °C to form (Glu)-CaH_2O. This is sparely soluble in water.

$(Glu)_2$-$CaCaCl_2H_2O$: $Ca(OH)_2$ was added to L-Glu-HCl in solution, saturated with CO_2, and boiled. Concentration yielded the complex salt.

INDUSTRIAL PRODUCTION

Extraction Method

From 1909, when MSG production started, until 1965, when the extraction process ended, wheat gluten and defatted soybeans were used as the main raw materials, containing as much as 15 to 25% L-glutamaic acid. Crude raw materials were hydrolyzed by heating with hydrochloric acid. After concentration under reduced pressure, the hydrolyzate was cooled to crystallize L-glutamic acid hydrochloride. L-Glutamic acid can be easily separated from other amino acids in the form of its hydrochloride because

of its low solubility in concentrated hydrochloric acid. The crude hydrochloride is dissolved in hot water and filtered to separate the humic substances formed in the hydrolyzation process. The pH was adjusted with caustic soda to 3.2, the isoelectric point of L-glutamic acid needed to crystallize L-glutamic acid crystals at a yield of more than 90%. The crude L-glutamic acid crystals are suspended in water and neutralized with sodium hydroxide. The solution is decolonized with activated carbon and crystallized MSG (monosodium L-glutamate monohydrate) under reduced pressure. Commercially preferable crystals are grown with a small amount of amino acids (31). The crystals are separated by centrifuge and then dried for packaging.

Beet, sugarcane molasses or Stefan's waste (which contained 2-pyrrolidone carboxylic acid) has been used as raw material for MSG production in the United States and Europe. The pyrrolidone carboxylic acid was hydrolyzed at 85 °C at pH 10.5 to 11.5 for about 2 h to avoid racemization (32).

Industrial Fermentation Method

For industrial production of MSG, molasses or starch hydrolyzate is generally used as raw material. Sugar molasses usually contains excess amounts of biotin, so when used, it is necessary to repress the activity of biotin by adding penicillin or certain surface-active substances at an early stage of the production to retard microorganism propagation. Furthermore, sugar concentration in culture media at below about 10% is necessary so as not to retard propagation. Additional sugar was fed during fermentation to obtain a higher concentration of excreted L-glutamate. This batch-fed fermentation process made possible a concentration of more than 120 g/L of L-glutamic acid in media.

Microorganisms used for L-glutamate fermentation are usually preserved under lyophilization below -80 °C or, for a short period, by keeping the stock culture below 10 to 15 °C. To refresh the microorganisms, stocked in either form, they are inoculated on an agar medium composed of 1% yeast extract and polypeptone, 0.5% sodium chloride, and 2% agar, at an optimum temperature of microorganisms. The refreshed microorganisms are then cultivated in liquid medium, shaken vigorously, and transferred to a small fermenter to allow them to propagate to about several kiloliters for seed culture. Industrial-scale fermenters are pressure tight, stainless steel containers, built to hold up to several hundred kiloliters of cultivating medium. They are equipped with aeration and stirring devices, as well as other automatic controls. Fermentation takes from 30 to 45 h.

Gaseous ammonia or a solution of urea and ammonium salts are convenient nitrogen sources for fermentation, not only as the initial medium, but also to maintain the pH of the medium at 7 to 7.5 for microbial growth and product formation. The culture medium becomes acidic because of assimilation of ammonium ions and the formation of L-glutamate. Gaseous ammonia can be used advantageously to maintain neutral pH and avoid dilution of culture medium, resulting in the high accumulation of L-glutamate in the fermentation broth, because it does not contain OH^- ions or water.

Microorganisms required several minerals, such as ferrous and potassium ions, which play important roles in L-glutamate fermentation. Other important conditions include regulating the aeration stirring. The biosynthesis of glutamate is performed under regulated aerobic conditions. When oxygen is not sufficiently dissolved, lactic acid and succinic acids accumulate and reduce the accumulation of glutamate. On the other hand, an excess of dissolved oxygen results in the formation of α-ketoglutaric acid. The optimal oxygen transfer rate was determined by measuring the rate of consumption of sodium sulfite, which is considered to depend on the characteristics of microorganisms used (33). The pressure of dissolved oxygen was usually kept above 1 kPa (0.01 atm) by aeration and agitation in the fermenter. The sterilized air is fed through a filter. Carbon dioxide is formed during fermentation and causes the medium to become heavy. Mechanical and chemical defoaming systems are provided.

The optimum temperature for fermentation is dependant on the character of microorganisms used. Regulation of the temperature using heat exchangers installed inside of the fermenter is indispensable. Fermentation is an exothermic reaction, and the temperature critically affects not only the propagation of microorganisms, but also the formation of glutamate.

In large-scale production the formation of trehalose very often reduces the product yield of glutamate. Trehalose consists of two α-1,1-bond glucose molecules and is excreted by the bacteria as an osmoprotectant. A process has been developed and successfully industrialized in which trehalose formation is controlled and decreased by culturing the overproducing mutant in media with inverted sucrose from molasses (34).

When fermentation is complete and after sterilization of the broth, the glutamate-recovery process is performed.

Separation Process

The separation process consists mainly of the following two processes (35):

1. Recovery of L-glutamic acid crystals from the fermentation broth
2. The purification of L-monosodium glutamate monohydrate from crude L-glutamic acid crystals

The typical outline of a production flow sheet of L-MSG from the fermentation broth is shown in Figure 7.

Hydrochloric acid at 35%% is added to the fermentation broth with agitation to crystallize the L-glutamic acid crystals at pH 3.2, the isoelectric point of L-glutamic acid. The crude L-glutamic acid crystals, retaining some mother liquor when removed from the slurry, are dissolved in water with equimolar sodium hydroxide; then the solution is applied to the activated carbon column to obtain the decolonized L-MSG solution, on which most of the color-related and hydrophobic substances are adsorbed. This purified L-MSG solution is fed continuously to a vacuum crystallizer to crystallize L-MSG crystals. L-MSG crystals separated by centrifugation are dried, sieved, and finally packaged to deliver to the consumers.

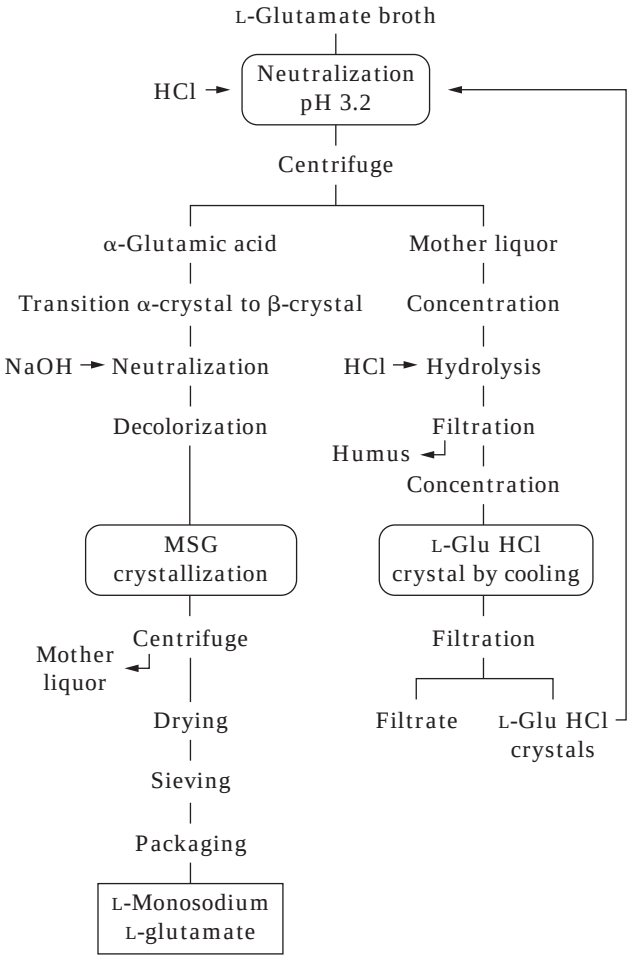


Figure 7. Flow diagram of production of MSG from the fermentation broth.

Crystallization of L-Glutamic Acid

In the fermentation broth, L-glutamic acid exists as an ammonium salt in the vicinity of pH 7. For obtaining L-MSG directly from the fermentation broth, it is necessary to exchange ammonium ion for sodium ions. This process is rather complicated, taking into account an environmental approach. To avoid this complexity, as a first step, crystallization of the α-form of L-glutamic acid crystals, which show lower solubility as compared to L-MSG and fairly good sedimentation as compared to β-form crystals, is employed in actual industrial production. The presence of salts can affect solubility of L-glutamic acid in solution. For example, the chloride-related inorganic substances NaCl, $CaCl_2$, NH_4Cl, and KCl increase the solubility of L-glutamic acid.

When crystallized under ordinary conditions, pyramid-type α-form crystals are obtained by the effect of the impurities contained in the fermentation broth.

Pyramid Type

In the growing of crystals, only those faces having the lowest translation velocities survive. The face of [001] of the α-form are fast-growing (high translation velocities) and this face tends to disappear, as overlapped by slower [111] faces, and finally results in the pyramid type.

If Ostwald's step theory is obeyed, as α-form crystals are unstable from the thermodynamic point of view, α-form crystals first appear and then gradually transform to β-form. But this theory is not based on the thermodynamics, but on the kinetics, so exceptions frequently occur.

The conditions to obtain α-form crystals are summarized as follows (36):

1. When neutralized, 35% hydrochloric acid is added as quickly as possible to the fermentation broth and then cooled to below 10 °C instantly with vigorous agitation.
2. A small amount of a third substance is added to affect the stability of α-form, such as a peptide, active surfactant, protein hydrolyzate, or amino acid, especially L-phenylalanine and L-cysteine.

Centrifugation

The α-form of L-glutamic acid crystals are separated by a centrifuge. Residual mother liquor after centrifugation amounts to 5 to 15% of the weight of the crystals, depending on the viscosity of the mother liquor and the degree of uniformity of crystal size. It is more effective to wash the crystals on the centrifuge with fresh water to reduce impurities than recrystallization. In this process, the bacterial cells remain effectively in the mother liquor, and without another bacterial cell separation process, L-glutamic acid crystals remain effectively separated from the bacterial cells.

Transformation of α-Crystals to β-Crystals

When the α-form crystals are poured into the water and the temperature is elevated to 60 to 80 °C, the transformation of α-form to β-form occurs within about 1 h. During this solvent-mediated transformation, a large amount of adhering impurities are rejected out of the crystals, and higher purified β-form miceous plate-shape crystals are obtained (37). This transformation process has so marked an efficiency for removing impurities, especially color-related substances, that when higher purity is required, a small amount of activated carbon is used to obtain a transparent MSG solution.

For another purification of L-glutamic acid solution, the application of UF membrane or NF membrane is investigated: When these membranes are applied to permeate the fermentation broth, the permeate solution obtained is neutralized followed by cooling. Purified β-form crystals are obtained directly, without α-form crystallization.

Crystallization of L-MSG

Neutralized L-glutamic acid solution (L-MSG solution) is applied to the activated carbon column to obtain decolonized solution, which is fed to the continuous evaporator to concentrate until the saturation of L-MSG (50% at 60 °C). To meet the requirements for L-MSG monohydrate crystals—crystal size, crystal shape (aspect ratio), and size distribution—it is necessary to strictly control the operational conditions of the crystallization process, in which nucleation and crystal growth occur simultaneously. The

result of the combined processes on the mentioned requirement for crystals depends on the relative rates of nucleation and growth. To obtain a more uniform product, nucleation is controlled by adding the desired number of crystals, usually crushed, to the crystallizer when the solution is either saturated or supersaturated. In a continuous crystallizer, the number of nuclei formed per unit time will be continuous and will equal the number of crystals withdrawn from the crystallizer, so that all nuclei can receive the same time of growth. Otherwise, there must be a classifying action in the crystallizer that will retain the small crystals under treatment until they have grown to the proper size, when they can be removed from the crystallizer.

For practical control, it is necessary to reduce as much as possible the local variation in the temperature or concentration of the solution that causes excess nucleation. Usually, the objective of the crystallizer operator is to achieve the maximum growth rate, consistent with low nucleation rate. More importantly, impurities in solution may inhibit the formation of new nuclei or retard crystal growth. The degree of the effect of a given impurity cannot be predicted; in general, higher molecular weight materials seem to be more effective inhibitors.

Crystal Shape (Aspect Ratio)

When no impurity is contained in the solution the aspect ratio (referred to as L/D in crystal) is 7 to 8 which is not convenient for umami seasoning for cooking. To overcome this problem, the appropriate amount of amino acids is added. The crystal shape is determined by the change in the difference in the relative crystal growth of constituted crystal faces. For L-MSG crystals, a small amount of amino acids changes the crystal shape, the amino acids such as L-alanine, L-lysine, and L-arginine are easily adsorbed on the [110] face of the L-MSG crystal and inhibit the growth of this face, those having the lowest translation velocities survive, and the aspect ratio is 4 to 5, which satisfies the demand of users (38).

Recovery of L-Glutamic Acid from Mother Liquor

In mother liquor–separated L-glutamic acid crystals other than L-glutamic acid, many impurities derived from the raw material used as the carbon source, as well as ammonium hydrochloride and bacterial cells, are found. For recovery of L-glutamic acid from this mother liquor, the following process is proposed: The mother liquor is condensed and ammonium chloride crystallized by cooling for a day and night followed by filtration to obtain the ammonium chloride crystals that are used as fertilizer. The separated mother liquor is hydrolyzed with excess hydrochloric acid to convert the L-glutamic acid to hydrochloride and the bacterial cells and other debris to insoluble matters (so-called humic acid). The humic acid is separated by filtration. The filtrate is condensed, then cooled to crystallize the hydrochloride crystals at lower than 10 °C for more than several days. This is recycled in the process of L-glutamic acid crystallization from the fermentation broth to curtail the consumption of hydrochloric acid. The residual mother liquor combined with humic acid is used as raw material for the organic fertilizer.

USES

Seasoning

MSG is widely used as a long-established seasoning or flavor enhancer to improve the palatability of foods. The effect of MSG is due to its characteristic taste, umami. The development of food science, including electrophysiology, psychophysics, and nutrition, have provided evidence that umami should be regarded as a basic taste, independent from the traditional four basic tastes—sweet, salty, sour, and bitter. Glutamate is ubiquitous in foods such as tomato, cheese, human milk, and seaweed. Other umami substances, that is, the 5′-ribonucleotides disodium 5′-inosinic acid and disodium 5′-guanylic acid, are found in meats, fish, vegetables, and mushrooms.

When glutamate and nucleotides coexist, a very strong synergistic effect occurs, and even a small amount of glutamate added to a food containing nucleotide causes umami to be dramatically enhanced—seven to eight times as much as the original umami of the food (39).

The taste threshold of MSG is about 0.03% in aqueous solution. The intensity of umami increases linearly with a logarithmic increase in the concentration of MSG. The synergistic effect of MSG with 5′-ribonucleotides is expressed by the following relation (41):

$$y = u + A\,u\,v$$

where y is the intensity equivalent to the concentration of MSG alone, and u and v are the concentrations of MSG and 5′-ribonucleotides, respectively. A is a constant that is 1,200 for disodium 5′-inosinate (IMP) and 2,800 for disodium 5′-guanylate (GMP). In some commercial umami seasonings, 5′-ribonucleotides are mixed in to take advantage of the synergistic effect with MSG.

Although MSG increases the palatability of food, it is not always palatable by itself. When MSG is tested in crystal form or in simple solutions, it does not cause a pleasant sensation. It gives a pleasant taste only in flavored solution or in actual foods. Another important property of MSG is the concentration dependence of its pleasantness rating. An excess of MSG causes an unpleasant sensation, and thus the amount of MSG that is added to foods should be limited (41).

The desirable usage level of MSG depends on the properties of the food, including the amount of original umami substances. Results of extensive taste studies indicate that the optimum MSG level in food ranges from 0.1 to 0.8 wt %, varying from food to food. In most cases, MSG is used in combination with salt. A relationship between the levels of MSG and salt, and palatability of foods, has been clarified. The lower the salt concentration, the higher the MSG concentration required to maintain the palatability, and vice versa. When umami substances are added to meals, the amount of salt can be reduced without a decrease in the palatability or degree of satisfaction.

Umami taste is considered to be a universal taste sensation for humankind. An ethnic comparison of taste sensitivity has shown that there is no difference in the thresholds of MSG and IMP between Japanese and European-American Caucasians (42).

Industrial Use

Depending on the pH, glutamic acid can be used to neutralize acidic or basic compounds, that is, arginine glutamate is used as a pharmaceutical and as a raw material for cosmetics. The reaction of glutamic acid with hydrochloric acid affords glutamic acid hydrochloride, which is used as a gastric acidifier. The property of glutamic acid as an amphoteric electrolyte is exploited in chelating agents, buffers, and builders for detergents.

The sodium salt of pyroglutamic acid (43), which is obtained by dehydration of glutamic acid followed by neutralization with sodium hydroxide, is very hydroscopic and used as a component of a natural moisturizing factor for human skin and as a humectant for cosmetics. The *N*-acylated compound (44) produced by reaction with a long-chain fatty acid is an anionic surfactant and is popularly used for skin and hair cleansers due to its mildness and high safety. The dibutylamine of the *N*-acyl-glutamate gelatinizes nonpolar oils and can be used as a recovering agent for marine-oil spills (45).

SAFETY

Monosodium L-glutamate is metabolized in the same way as glutamic acid from any digested protein. After oral ingestion, glutamate is absorbed from the intestine by an active-transport system. During this process, a large portion is metabolized to alanine and α-ketoglutarate, which enters the tricarboxylic acid cycle. Glutamate is further metabolized in the liver to give glucose, lactate, glutamine, and other amino acids. Consequently, blood glutamate levels do not rise significantly unless very large doses are administered. Blood glutamate levels transiently rise when large doses are ingested on their own, but the coingestion of foods that contain metabolizable carbohydrate increases the metabolism of monosodium L-glutamate and eliminates or greatly attenuates this rise. Human infants, including premature infants, metabolize glutamate similar to adults (46,47).

The acute toxicity of MSG is low (48), the oral LD_{50} for humans, calculated on the basis of doses administered in different ways to various animals, would represent a single dose greater than 1 kg for a person weighing 70 kg. In contrast, the oral LD_{50} for sodium chloride in rats is 3.75 kg/kg body weight (49).

The main use of MSG is as a umami seasoning or a food ingredient, and so its safety when used in the diet is the most important aspect of its use. Both short-term and chronic toxicity studies on MSG in the diet of several species at doses of up to 4% (approximately 6 to 8 g/kg body weight/day) in the diet showed no specific toxic effects and no evidence for carcinogenicity or mutagenecity. Reproduction studies of up to three generations did not reveal any adverse effects of dietary MSG ingestion. Fertility, gestation, viability, and lactation indexes, or the pre- and postweaning performance of offspring, were unaffected. In primates, it was found that the placenta serves as an effective barrier to the transfer of glutamate from the oral ingestion of large doses of MSG.

The concentration of glutamate is higher in the brain than in the blood. The demonstration that injected or force-fed neonatal rodents given extremely high doses of MSG showed evidence of brain leisons has led to much additional research to determine any possible link between neurotoxicity and human use of MSG (50). However, from animal tests on monkeys, guinea pigs, rat, and mice, glutamate levels in the brain remain unchanged after the administration of large oral doses of MSG (51). Numerous experiments on rodents, as well as dogs and monkeys, using dosage levels up to 43 g MSG/kg body weight have failed to show any link between dietary use of MSG and brain damage. In the case of dogs and monkeys, even experiments involving injection of MSG have not shown any effects on the brain.

An anecdotal report in 1968 triggered interest in Chinese restaurant syndrome (CRS), a reaction associated with transient subjective symptoms of burning, numbness, and a tight sensation in the upper part of the body. Possible association with food ingredients such as MSG was suggested. No objective changes in skin temperature, heart rate, ECG, or tone were observed, and no correlation was seen between plasma glutamate levels and symptoms. In 1979 a questionnaire survey of more than 3,000 people in the United States was conducted for CRS. Only 1.8% of those surveyed acknowledged that they experienced possible CRS feelings (52), and 0.19% of those with symptoms attributed them to eating in Chinese restaurants. These feelings were also found to occur after consumption of spicy tomato juice, orange juice, coffee, and tea. In 1986 self-identified MSG responders were challenged in a properly controlled double-blind study, and it was concluded that the link between MSG and CRS was not supportable (53). In 1987, the Joint FAO/WHO Expert Committee on Food Additives (JECFA), a scientific advisory body to the World Health Organization and Food and Agriculture Organization of the United Nations, concluded that properly conducted double-blind studies among individuals who claimed to suffer from the syndrome did not confirm MSG as the causal agent (54).

In 1981 Chinese restaurant asthma was reported following capsule administration of MSG to several asthmatics (55), but the researchers failed to account for other allergens to which the subjects could have been exposed and did not utilize the scientific practice of a "control" substance that would have helped to determine if glutamate triggered this response. In a double-blind crossover study in which chronic asthmatics were challenged with MSG or a placebo, no decrease in pulmonary function was observed (56).

JECFA reviewed the safety studies of glutamate and endorsed its safety by allocation of an acceptable daily intake (ADI) for L-glutamic acid and its monosodium salt, potassium salt, ammonium salt, calcium salt, and magnesium salt as being "not specified." This result is ascertained by the scientific committee for food of EC (57).

BIBLIOGRAPHY

1. K. Ikeda, *J. Tokyo Chem. Soc.* **30**, 820–836 (1908).
2. L.B. Lockwood and F.H. Stodola, *J. Biol. Chem.* **164**, 81–93 (1946).

3. S. Kinoshita, K. Nakayama, and S. Udaka, *J. Gen. Appl. Microbiol.* **3**, 193–205 (1957).
4. Y. Hirose and H. Okada, *Microbial Technology*, 2nd ed., vol 1, Academic Press, New York, 1979, p. 211.
5. I. Shiio, *Seikagaku* **50**, 1–16 (1978).
6. S. Dagley, E.A. Davis, and G.A. Morrison, *Nature* **165**, 437–438 (1950).
7. W. Liebl, U. Ehrman, W. Ludwig, and K.H. Schleifer, *Int. J. Syst. Bacteriol.* **41**, 225–260 (1991).
8. R. Kramer, *FEMS Microbiol. Rev.* **13**, 75–93 (1994).
9. I. Shiio and K. Ujikawa, *Agric. Biol. Chem.* **42**, 1897–1904 (1980).
10. E.R. Boermann, B.J. Eikmanns, and H. Sahm, *Mol. Microbiol.* **6**, 317–326 (1992).
11. T.E. Walker, C.H. Han, V.H. Kollman, R.E. London, and N.A. Matwiyoff, *J. Biol. Chem.* **257**, 1189 (1982).
12. H. Momose and T. Takagi, *Agric. Biol. Chem.* **42**, 1911–1917 (1978).
13. Japanese Pat. 50-92795 (1980), O. Tosaka, Y. Murakami, K. Akashi, and S. Ikeda, Ajinomoto Co.
14. Japanese Pat. 50-521762 (1980), E. Ono, O. Tasako, and K. Takinami, Ajinomoto Co.
15. French Pat. 2612937 (March 24, 1988), Y. Yamada and A. Seto, Ajinomoto Co.
16. H. Wilson and R.K. Cannan, *J. Biol. Chem.* **119**, 309–331 (1937).
17. C.B. Anfensen, J.T. Edsall, F.M. Richards, and M.L. Anson, *Advances in Protein Chemistry*, vol. 4, Academic Press, New York, 1948, p. 339.
18. W.F. Diven, *Biochim. Biophys. Acta* **191**, 702–706 (1969).
19. P. Pfeiffer and J. Wurgler *Z. Physiol Chem.* **97**, 128–147 (1916); P. Pfeiffer and O. Angern, *Z. Physiol. Chem.* **133**, 180–192 (1924).
20. T. Takenishi, *Journal of the Chemical Society of Japan* **81**, 1380–1382 (1960); **82**, 15 (1961).
21. M.C. Otey, J.P. Greenstein, M. Winitz, and S.M. Birnbaum, *J. Am. Chem. Soc.* **77**, 3112–3114 (1955).
22. Y. Sakata, *Agric. Biol. Chem.* **25**, 829–834 (1961).
23. T. Ogawa, *Journal of the Chemical Society of Japan, Ind. Chem. Section* **52**, 69 (1949).
24. Y. Sakata, *Agric. Biol. Chem.* **25**, 835–837 (1961); **26**, 355–361 (1962).
25. M.S. Lehman, T.F. Koetzle, and W.C. Hamilton, *J. Cryst. Mol. Struct.* **2**, 225–233 (1972); M.S. Lehman and A.C. Nenes, *Acta Crystallogr.*, **B36**, 1621–1625 (1980); N. Hiroyama, K. Shirahata, Y. Ohashi, and Y. Sasada, *Bull. Chem. Soc. Jpn.* **53**, 30–35 (1980).
26. S. Budavari ed., *Merck Index*, 12th ed., Merck Co., Inc. Rahway, N.J., 1996, p. 760.
27. A. Albert, *Biochem. J.* **50**, 690–698 (1952).
28. T. Ogawa, *Journal of the Chemical Society of Japan, Ind. Chem. Section* **52**, 102–105 (1949).
29. Japanese Pat. 23-3312 (Dec. 17, 1949), T. Ogawa, Ajinomoto Co.
30. T. Kashiwagi, N. Nagashima, C. Sano, and T. Kawakita, *Acta Crystallogr.*, **C51**, 1053–1056 (1995).
31. U.S. Pat. 2834805 (May 13, 1958) J.L. Purvis and V. Bassel.
32. H. Olbrich ed., *Die Melasse*, 4th ed., Nippon Seito Kogyokai, Tokyo, Japan, 1961, p. 227.
33. Y. Hirose, H. Sonoda, K. Kinoshita, and H. Otsuka, *Agric. Biol. Chem.* **30**, 585–593 (1966).
34. H. Yoshii, M. Yoshimura, S. Nakamura, and S. Inoue, *Nippon Nogei Kagakukaishi* **67**, 955–960 (1993).
35. T. Kawakita and M. Saeki, *Chem. Eng. Jpn.* **50**, 196–203 (1986).
36. S. Hiramatsu, Ph.D. Thesis, University of Tokyo, Tokyo, Japan (1976).
37. Japanese Pat. 45-4730 (Feb. 17, 1970), K. Ito, N. Mizoguchi, and M. Dazai, Ajinomoto Co.
38. H. Tasuke, *Nippon Nogei Kagakukaishi* **50**, 175–179 (1976).
39. A. Kuninaka, *Nippon Nogei Kagakukaishi* **34**, 489–492 (1960).
40. S. Yamaguchi, *J. Food. Sci.* **36**, 846–849 (1971).
41. S. Yamaguchi and C. Takahashi, *Agric. Biol. Chem.* **48**, 1077–1081 (1984).
42. S. Yamaguchi, H. Kiminami, and R. Ishii, *Proceedings of the 22nd Japanese Symposium on Taste and Smell, Japanese Association for Safety of Taste and Smell*, Gifu, 1988, p. 73.
43. K. Laden and R. Spitzer, *J. Soc. Cosmet. Chem.* **19**, 351–360 (1967).
44. M. Takehara, R. Yoshida, and M. Yoshikawa, *Cosmet. Toiletries* **94**, 31 (1979).
45. M. Honma and T. Ikeda, *Region of Chemistry* (in Japanese), **36**, 697–701 (1982).
46. L.J. Filer, Jr., S. Garattini, M.R. Kare, W.A. Reynolds, and R.J. Wurtman eds., *Glutamic Acid*, Raven Press, New York, 1979.
47. Joint FAO/WHO Expert Committee on Food Additives, Toxicological Evaluation of Certain Food Additives, World Health Organization, 1988.
48. H. Moriyuli and M. Ichimura, *Appl. Pharmaceuticals* (in Japanese) **15**, 433 (1978).
49. S. Budavari, M.J. O'Neill, A. Smith, P.E. Heckelman, and J.F. Kinneary eds., *Merck Index*, 12th ed., Merck Co., Rahway, N.J., 1996, p. 1474.
50. J.W. Olney, *Science* **164**, 719–721 (1969).
51. Y. Takasaki, *Toxicology* **9**, 293–305, 307–318 (1978).
52. G.R. Kerr, *J. Am. Diet. Assoc.* **75**, 29–33 (1979).
53. R.A. Kenny, *Food Chem. Toxicol.* **24**, 351–354 (1986).
54. D.H. Allen and G.J. Baker, *New Engl. J. Med.* **305**, 1154–1155 (1981).
55. R.E. Gosselin, R.P. Smith, H.C. Hodge, and J.E. Braddock, *Clinical Toxiocology of Commercial Products*; 5th ed., Williams & Wilkins, London, 1989, p. 407.
56. J.M. Drazen, *J. Asthma* **167**, 167–172 (1987).
57. Commission of the EC Reports of the Scientific Committee for Food, **25**, 16, 1990.

ADDITIONAL READING

T. Kawakita, in J.I. Kroschwitz ed., *Encyclopedia of Chemical Technology*, 4th ed. vol. 2, 1982, p. 410.

T. Kawakita, C. Sano, S. Shioya, M. Takehara, and S. Yamaguchi, in B. Elvers, S. Hawkins, and G. Schulz eds., *Ullman's Encyclopedia of Industrial Chemistry*, 5th ed., vol. A16, VCH, Weinheim, Germany, 1991, p. 711.

W. Leuchtenberger, in H.-J. Rehm and G. Reed eds., *Biochemistry*, 2nd ed., vol. 6, VCH, Weinheim, Germany, 1996, p. 466.

K. Soda, H. Tanaka, and N. Esaki, in H.-J. Rehm and G. Reed eds., *Biotechnology*, vol. 3, Verlag Chemie, Weinheim Germany 1983, p. 481.

T. Yoshida, in J.I. Kroschwitz ed., *Encyclopedia of Chemical Technology*, 3rd ed., vol. 2, 1982, p. 410.

AMINO ACIDS, PRODUCTION PROCESSES

TETSUO OKA
Kyowa Hakko Kogyo, Co., Ltd.
Tokyo, Japan

KEY WORDS

Amino acids
Corynebacterium
Corynebacterium glutamicum
Fermentation
Lysine
Mutant
Recombinant

OUTLINE

INTRODUCTION

Monosodium L-glutamate (MSG) was found by Ikeda in 1908 as a flavoring component of *konbu*, or sea tangle. This was the start of the later development of the amino acid industry. Initially, MSG was produced industrially by extraction from a protein hydrolysate, more specifically, first from wheat gluten hydrolysate with hydrochloric acid, and then from defatted soybean hydrolysate or by conversion from pyrrolidone-5-carboxylic acid in beet molasses, for use as a seasoning agent. This latter process was not sufficient to meet the increasing demand because it involved the problem of by-products disposal, requiring the development of new production processes. Patent applications filed around that time related to chemical synthesis of MSG from succinate semialdehyde or cyclopentane. In 1956, Kinoshita et al. discovered a microbial process for L-glutamic acid production by direct fermentation of a microorganism (*Corynebacterium glutamicum*) from sugar and ammonia (1). Consequently, industrial production of MSG as a seasoning was rapidly enlarged, and at the same time, a new industry of amino acid production by fermentation of microorganisms emerged. In 1958, Kinoshita, Nakayama, and Kitada found that an auxotrophic mutant of *C. glutamicum* requiring homoserine accumulated L-lysine in a medium (2), which enabled industrial production of L-lysine by fermentation. Establishing the basis for the development of the amino acid fermentation industry, this technology suggested the possibility of producing various other amino acids by auxotrophic mutants and also the importance of research in fermentative production of biological components with regulatory mutants. Since then, direct fermentation of various amino acids has been broadly studied.

Reduction of the cost of amino acids, owing to the development of fermentation techniques together with the establishment of their production processes, facilitated extended application of amino acids to uses other than as a seasoning. Particularly, L-lysine, an essential amino acid that is lacking in several food proteins, has been enjoying remarkably increased demand as a feed additive. Table 1 shows the production of the major amino acids. Most of the amino acids produced in large quantities are still manu-

Table 1. Estimated Worldwide Production of Amino Acids (1996)

Amino acids	Estimated production (ton/year)	Process[a]
Glycine	22,000	C
L-Aspartic acid	7,000	E
L-Arginine	1,200	F
L-Cysteine	1,500	E
Monosodium L-Glutamate	1,000,000	F
L-Glutamine	1,300	F
L-Histidine	400	F
L-Isoleucine	400	F
L-Leucine	500	Ex, F
L-Lysine HCl salt	250,000	F
DL-Methionine	350,000	C
L-Phenylalanine	8,000	F, C
L-Proline	350	F
L-Threonine	4,000	F
L-Tryptophan	500	F, E
L-Tyrosine	120	Ex

Source: Ref. 3.
[a]C, chemical synthesis; E, enzymatic method; Ex, extraction; F, fermentation.

factured by fermentation, except for glycine, which does not have optical isomers, and methionine, which has a similar effect as a feed additive in both L- and DL-forms.

In addition to being used as a seasoning agent and a feed additive, amino acids are currently used widely as a raw material for the sweetener Aspartame (*N*-L-α-aspartyl-L-phenylalanine 1-methyl ester), for pharmaceuticals and agrichemicals, for infusion and oral nutrition, for surfactants, and so forth, based on their reactivities and nutritional, pharmacological, and flavoring effects.

In connection with amino acid fermentation, many reports have been published on the physiological properties of amino acid–producing microorganisms. The following discussions stress problems from the standpoint of practically manufacturing amino acids on an industrial scale.

FERMENTATION PROCESSES

Amino acid fermentation may be defined in two ways. One is a definition in the narrow sense and refers to direct accumulation of amino acids in a medium containing a sugar, ammonia, and other nutrients, and the other is a broad definition that covers the fermentation process involving the addition of specific precursors and amino acid production by reaction using enzymatic functions of microorganisms. Here, mainly direct fermentation of the former type is discussed.

Fermentation processes generally have a number of advantages and benefits that accrue to the user of the process. On the other hand, they also have drawbacks. Both are summarized as follows:

Benefits

- Mild conditions are used both in fermentation and in product recovery; hence, little product degradation takes place.
- Fermentation requires relatively less complex operation.
- Once the plant is built and operation has begun, there are relatively low maintenance costs.
- Only L-form amino acids are obtained.

Problems

- Operations provide low product concentrations compared with chemical synthesis processes and require large volumes of water, large fermenter capacity, and comparatively high capital investment.
- The requirements for strict sterility add to capital costs and operation costs.
- Large amounts of energy for oxygen transfer and mixing are required.
- Product recovery may be complex, difficult, and expensive.
- The process time necessary to reach maximum concentrations of the desired product is usually comparatively long.

The aforementioned narrow-sense definition of direct fermentation of amino acids includes three types of fermentation: batch-type, fed-batch-type, and continuous.

Batch Fermentation

Industrial fermentation is mostly performed using batch processes. In a batch process, a large volume of a medium containing nutrients and substrate material is inoculated with a viable culture of one or more appropriate microorganisms. Microbial growth and biochemical synthesis are allowed to proceed until an optimum yield of metabolite or a desired biochemical transformation has been obtained. Although this process is the basic form of fermentation, it is not frequently practiced on an industrial scale because productivity is limited by the amount and nature of the nutrients present at the time of inoculation. In most amino acid fermentation, enhanced product concentration is important to improve production efficiency and requires more advanced fermentation processes.

Fed-Batch Fermentation

Fed-batch fermentation types are classified in Table 2. This fermentation method is aimed at efficiently carrying out fermentation and is characterized by a low concentration of components in the initial medium to minimize metabolic regulation. At the time of inoculation the medium promotes initial growth of microbes; subsequent supplies of more raw materials drive the desirable increase in metabolite biosynthesis. Feeding is effected either intermittently or continuously. Industrial fermentation of most amino acids is accomplished with this method. Fed-batch fermentation requires feeding equipment in addition to the equipment required for batch fermentation and therefore leads to higher fixed costs (Fig. 1). However, the process can provide improved productivity as a whole because of the enhanced yield and reduced fermentation time. Actual product cost depends on the cost of the carbon source or specific precursors that are used for feeding. The nitrogen source is supplied generally through pH control with ammonia. This process is of particular importance in industrial operations where the reaction of the microbial catalyst does not last long; hence, continuous fermentation cannot be practiced, unlike in the case of L-glutamic acid fermentation from cane molasses using penicillin, described next.

Table 2. Classification of Fed-Batch Microbial Reaction Processes

Case 1 (nonfeedback regulation)
Constant feeding-rate method
Exponential feeding method
Optimized feeding method
Others (intermittent feeding, etc.)
Case 2 (feedback regulation)
Indirect control
Direct control
Constant control
Program control
Others

Source: Ref. 4.

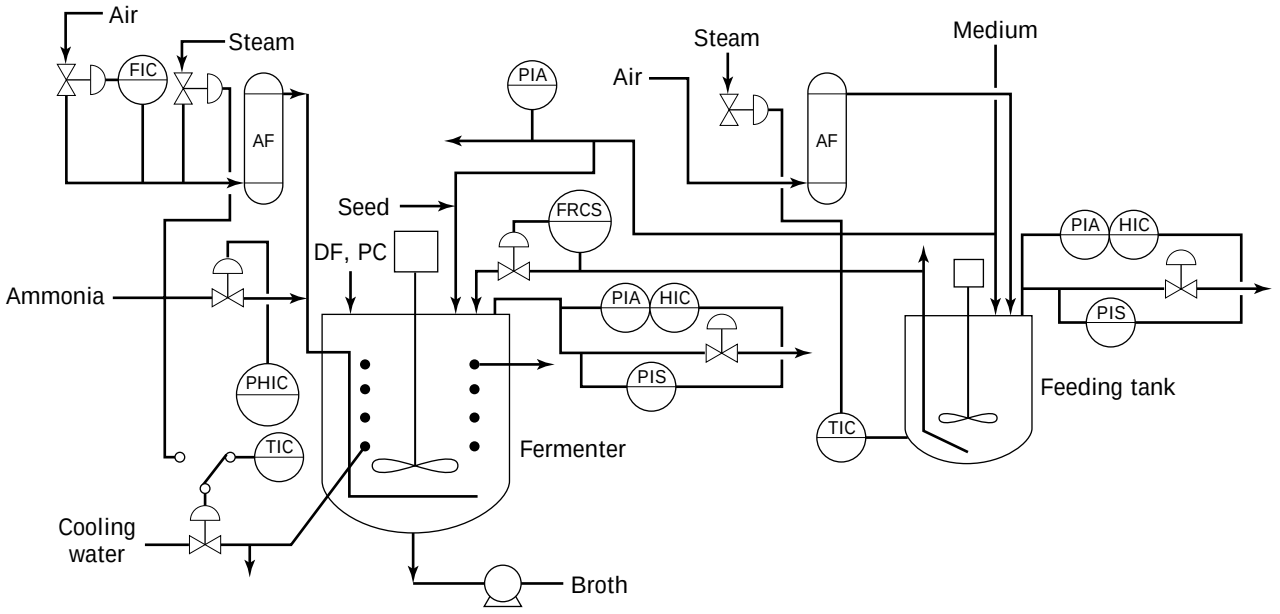


Figure 1. Typical system for fed-batch fermentation. AF, air filter; DF, defoamer; PC, penicillin; FIC, flow indication control; PHIC, pH indication control; TIC, temperature indication control; PIS, pressure indication sum; PIA, pressure indication alarm; HIC, highest indication control; FRCS, flow record control sum.

Continuous Fermentation

In continuous fermentation, a complete medium is fed to a fermenter after an appropriate period of batch fermentation, and the same quantity of broth is continuously taken from the fermenter to maintain the fermentation broth at a fixed volume. This may be performed either by the chemostat method using a substrate or limiting substance, or by the turbidostat method in which the cell level is adjusted to maintain constant cell mass. Because the continuous fermentation process allows improvement of productivity compared with the ordinary fermenter, the initial investment in equipment is small relative to the production volume, and operation cost is low. However, one drawback is that it is not suitable for small-scale production, and the challenges of sterile operation and equipment maintenance are more necessary than they are for batch and fed-batch fermentation.

This process shifts from batch fermentation to continuous fermentation when productivity per unit time in the former is relatively high. There are many reports analyzing the steady-state condition in continuous fermentation. Most of them relate to cell culture, but a few reports are available specifically on amino acid fermentation (5). One of the reasons may be that studies on amino acid fermentation have been mainly directed to the influence of metabolic regulation, as in the case of the penicillin addition method in glutamic acid fermentation, or because some amino acid processes have a distinct growth phase and production phase, and continuous culture cannot be used. Many of the microbial strains used in amino acid fermentation are released from metabolic regulation to a remarkable extent. This makes it easier to analyze continuous processes and thereby optimize them. It is necessary to study optimum conditions for each process and to optimize their industrial application.

Unlike processes of chemical synthesis, continuous fermentation processes have their own restrictions and duration. This is because microbes undergo spontaneous mutation within the system, and an increase in the fraction of microbes with decreased productivity may lead to rapid reduction in productivity (Fig. 2). Hence, it is necessary to breed a strain with high genetic stability.

Enzymatic Method

Of the amino acid production processes using direct enzymatic biotransformation, those for L-alanine (6), L-aspartic acid (7), L-lysine (8), and L-tryptophan (9,10) have been the most extensively studied, and some of the results have led to standard industrial processes. Although the enzymatic production process of L-lysine from DL-aminocaprolactam did not result in practical application, this technology is interesting because of its use of petrochemical products for fermentation raw materials. The outline of the process is shown in Figure 3.

MICROORGANISMS

In amino acid fermentation, screening/breeding of microorganisms to be used for the process is the most important step. These microorganisms are classified into (*1*) wild-type strains isolated from nature to meet a specific purpose, (*2*) mutants that provide the desired properties that are developed by spontaneous or artificial mutation of

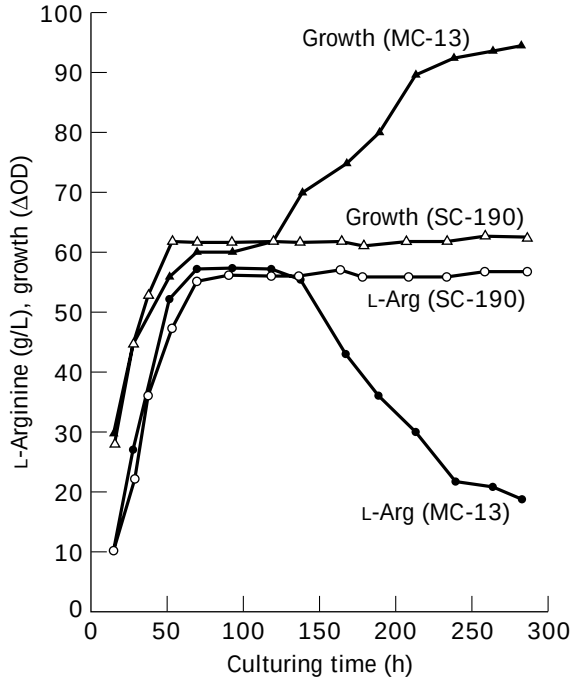


Figure 2. Comparison of L-arginine continuous culture profile of MC-13 and SC-190. *Source*: Ref. 5.

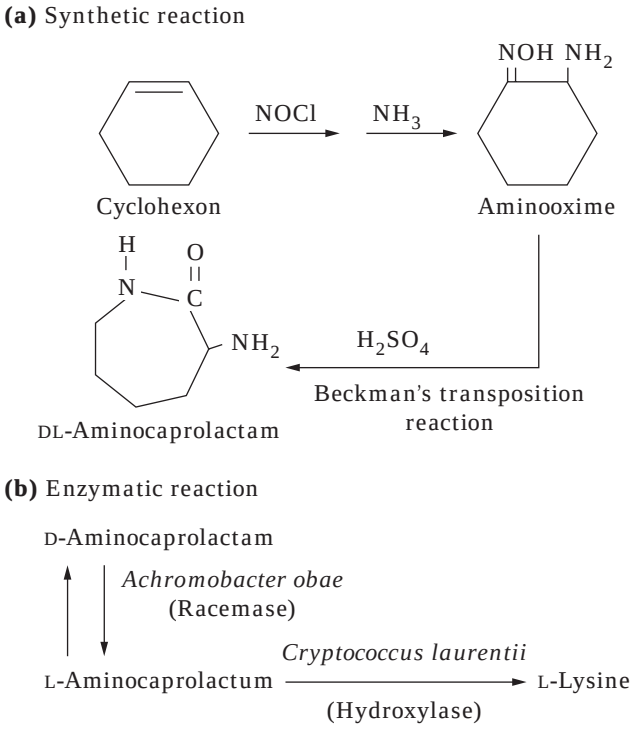


Figure 3. L-Lysine production by a combination of synthetic and enzymatic reactions. *Source*: Ref. 8.

wild-type strains, and (*3*) recombinant strains bred or constructed through genetic recombination technologies.

Wild-Type Strains

Microorganisms suitable for the purpose are selected from nature. They are cultured under adjusted fermentation conditions so as to produce excess amounts of the desired products. Typical examples are *C. glutamicum* and *Brevibacterium flavum* used for L-glutamic acid fermentation.

Mutant Strains

Auxotrophic strains and amino acid analog–resistant strains fall into the mutant strain category. Most amino acids can be produced by these types of microorganisms. Currently, most fermentative production of amino acids is accomplished in this fashion. Tables 3 and 4 and Figure 4 show L-lysine producers and L-tryptophan producers as typical examples of strain improvement.

Table 3. L-Lysine-Producing Microorganisms

Microorganisms[a]	L-Lysine HCl productivities: Conc. (g/L)	Yield (%)	Ref.
C. glutamicum (Hse^-)	13	13	2
B. flavum (Thr^-, Met^-)	34	34	11
B. flavum (AEC^r)	32	32	12
C. glutemicum (Hse^-, Leu^-, $Pant^-$, ACE^r)	42	42	13

Source: Refs. 2, 11–13.
[a]AEC, *S*-(2-aminoethyl)-L-cysteine; Pant, pantothenate.

Table 4. Isolation and Productivities of Tryptophan-Producing *C. glutamicum*

Strains	Phenotype[a]	Production of L-Trp (g/L)
KY 9456	Phe^-, Tyr^-	0.15
MT-11	Phe^-, Tyr^-, $5MT^r$, $TrpHx^r$, $6FT^r$, $4MT^r$	4.9
PFP-2-32	Phe^-, Tyr^-, $5MT^r$, $TrpHx^r$, $6FT^r$, $4MT^r$, PFP^r	5.7
PAP-136-50	Phe^-, Tyr^-, $5MT^r$, $TrpHx^r$, $6FT^r$, $4MT^r$, PFP^r, PAP^r	7.1
TX-49	Phe^-, Tyr^-, $5MT^r$, $TrpHx^r$, $6FT^r$, $4MT^r$, PFP^r, PAP^r, $TyrHx^r$	10.0
Px-115-L-67	Phe^-, Tyr^-, $5MT^r$, $TrpHx^r$, $6FT^r$, $4MT^r$, PFP^r, PAP^r, $TyrHx^r$, $PheHx^r$	12.0

Source: Ref. 14.
[a]5Mt, 5-methyltryptophan; TrpHx, tryptophanhydroxamate; 4MT; 4-methyltryptophan; 6FT, 6-fluorotryptophan; PFP, *p*-fluorophenylalanine; PAP, *p*-aminophenylalanine; TyrHx, tyrosinehydroxamate; PheHx, phenylalaninehydroxamate.

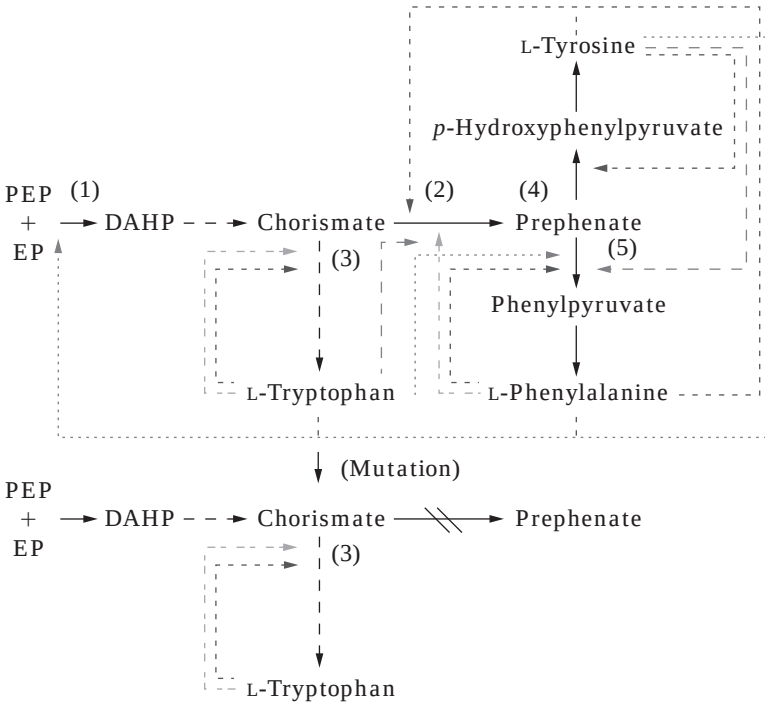


Figure 4. Metabolic regulation of aromatic amino acids in *C. glutamicum*. 1, DAHP synthetase; 2, chorismate mutase; 3, anthranilate synthetase; 4, prephenate dehydrogenase; 5, prephenate dehydratase. PEP, phosphoenolpyruvate; EP, erythrose 4-phosphate; DAHP, 3-deoxy-D-arabinoheptulosonic acid 7-phosphate; ·····➤, repression; ·····➤, inhibition; ·····➤, activation. *Source*: Ref. 14.

Recombinant Strains

There are many reports on the strains prepared by recombination of the various genes required for amino acid biosynthetic enzymes, as well as their possible applications to amino acid fermentation. Self-cloning recombinants are often reported in the literature because they are highly practical from the viewpoint of regulation. They will be put to use in the future as a primary technology for amino acid fermentation. Table 5 shows two L-threonine-producing strains. Whichever types of strains are constructed, conditions such as raw materials, temperature, process, and so on, as well as the by-products to be produced must be determined.

MEDIA

Components

Components of a medium are decided upon with regard to the microorganism, cost, purification process, waste treatment, and so forth. Except for particular cases, the cost of the carbon source, a main raw material, accounts for most of the cost of raw materials. Hence, its selection is of primary importance.

Cane molasses, beet molasses, glucose, methanol, *n*-paraffin, and other carbon source materials have been studied. In enzymatic processes and fermentation processes with addition of precursors, petrochemical products may be used as a precursor for the desired product. Carbon sources currently used for the industrial production of amino acids are mostly sugars, including cane molasses, beet molasses, and corn syrup (glucose), because of their relatively low cost and availability. When a natural substance such as molasses is used as a carbon source, composition of the medium must be properly adjusted and empirically optimized because natural substances contain additional components.

Typically, the concentration of the nitrogen source is relatively low in an initial medium to avoid overregulation of the biosynthetic pathways, and nitrogen is supplemented by way of pH control with ammonia. Because amino acid fermentation is a process of converting ammonia into amino groups, the importance of the nitrogen source is second only to that of the carbon source.

In addition, phosphate and other minerals are added in the required amounts. When an auxotrophic strain is used for fermentation, the required substance is added at an optimum concentration. Furthermore, an antifoam agent must be added to prevent foaming of the broth at the sterilization and fermentation stages.

Table 5. Accumulation of Amino Acids by *E. coli* Strain No. 29-4 and *β*-IM4

	Amino acids produced (g/L)						
Strain	Thr	Lys	Pro	Glu[a]	Ile	Asp	Phe
β-IM4[b]	3.82	0.31	0.25	0.20	0.14	0.09	0.06
No. 29-4[c]	13.40	0.03	0.00	0.13	0.12	0.00	0.02

Source: Ref. 15.
[a]Or homoserine.
[b]Threonine-overproducing mutant.
[c]Recombinant (self) or *β*-IM4 with pBR322-thr.

Sterilization

When the fermenter has a capacity of approximately 10,000 L or less, the preparation of raw materials and steam sterilization are often both carried out within the

fermenter. On the other hand, when a large fermenter (40,000-L capacity or more) is used, as in the case of amino acid fermentation, each piece of equipment is separately sterilized with steam and cooled with sterile air, and raw materials other than sugar are prepared in a make-up tank and subjected to steam sterilization and cooling. Sugar is sterilized separately and sent to the fermenter with other nutrients. The flow diagram is shown in Figure 5.

Similar procedures apply to preparation of the feed tank and the feeding solution in fed-batch fermentation. Feeding is carried out from the feed tank through a sterilized line while controlling the flow.

FACTORY FERMENTATION OPERATION

Figures 1 and 6, respectively, show a process flow diagram of amino acid fermentation. The process comprises the steps of laboratory seed culture, factory seed culture (Fig. 6), and fermentation (Fig. 1), and the size of culturing is scaled up in this order.

Laboratory Seed Culture and Factory Seed Culture

The most important feature in the steps of laboratory and factory seed culture is to ensure active microbes. This promotes initial growth of microbes at the fermentation step and stabilizes fermentation. Although the cost required for these steps accounts for a relatively large portion of that of the whole process, the steps are of much importance to properly operate the process; it is thus necessary to carry out the steps only under optimum conditions. Laboratory seed culture requires an exceedingly high level of sterility because all subsequent steps for all of the lots may be lost as a result of contamination at this step. Likewise, transfer from laboratory seed culture to factory seed culture must be performed under highly sterile conditions. Typically, all transfer steps after laboratory seed culture are closed, and the whole fermentation train of equipment must be sterilized carefully before the initiation of transfer. Because amino acid fermentation is generally carried out under aerobic conditions at a neutral pH, the process will easily support the growth of a contaminant.

Fermentation Equipment

Figure 7 illustrates a typical fermenter. Amino acid fermentation is performed aerobically, so an efficient oxygen supply is required. The effect of oxygen supply on amino acid fermentation is shown in Table 6, which indicates that the sufficiency rate of oxygen required (r_{ab}/KrM) by microbes is controlled by the oxygen transfer efficiency of the fermenter. Oxygen transfer efficiency correlates with aeration and agitation conditions. It strongly affects the results of fermentation and at the same time constitutes the main factor of the cost of utilities.

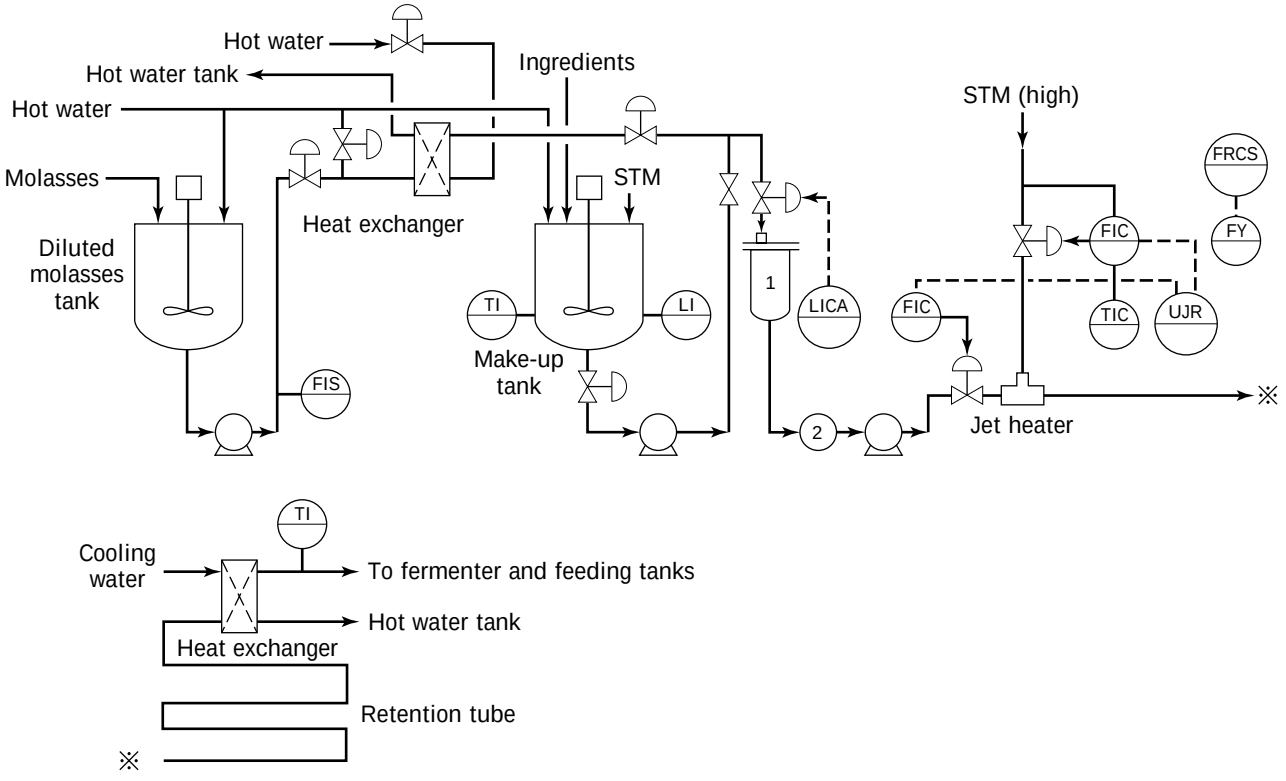


Figure 5. Piping and instrumentation drawing of the sterilizing section of a continuous sterilizer. 1, Cushion tank, 2, strainer; FIS, flow indication sum; TI, temperature indication; STM, steam LI, lowest indication; LICA, lowest indication control alarm; FIC, flow indication control; TIC, temperature indication control; FRCS, flow record control sum; FY, flow indication converter; UJR, point recorder. *Source*: Ref. 16.

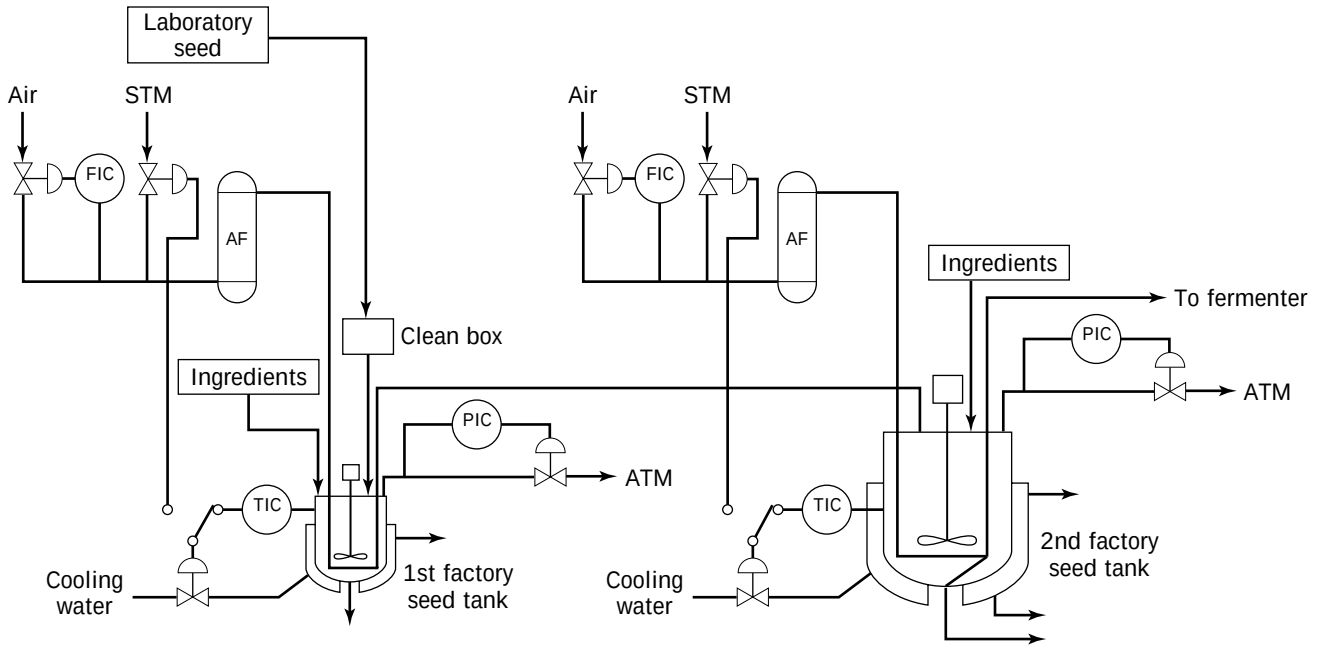


Figure 6. Factory seed culture equipment. AF, air filter; FIC, flow indication control; STM, steam; TIC, temperature indication control; PIC, pressure indication control; ATM, atmosphere.

Process Measurement and Control

In a fermentation process, aeration, pH, feed rate, dissolved oxygen, temperature, and foaming level must be measured and controlled. Parameters automatically measured during fermentation generally include fermenter temperature, volume and temperature of cooling water, volume of air supplied into the fermenter, concentrations of O_2 and CO_2 in the exhaust gas, pressure within the tank, feed volume, agitation rate, agitation energy, pH, dissolved oxygen, and degree of foaming. Sampled broth is analyzed mainly for optical density, sugar concentration, amino acids, and organic acids. Fermentation temperature, pH, aeration, feed volume, and foaming level are automatically controlled. In continuous fermentation, equipment capable of controlling the volume of both discharged and the liquid surface is provided.

Downstream Processing

Amino acids are typically recovered and purified either by chromatographic methods or by concentration–direct crystallization methods. Although most amino acids may be isolated and purified using an ion exchange resin, this tends to increase the cost of waste-liquor treatment because of the large volume of diluted waste liquor. On the other hand, with the concentration–direct crystallization method, the discharged mother liquors have a high biological oxygen demand and contain a large amount of plant-growth-promoting factors and may thus be used as fertilizer. This makes the cost of waste-liquor treatment less burdensome; however, this method cannot be applied to all amino acids because the chemical composition of fermentation broth strongly affects the purification steps.

Wastewater Treatment

In recent years, treatment of fermentation waste liquor has become more and more important in the planning process. When chromatographic purification is employed, the low concentration and large volume of waste liquor is burdensome, and the cost of treatment by the active sludge method is high, exerting much influence on the cost of the final product. In this respect, the direct crystallization method is advantageous, though it may not be applicable to many products due to characteristics peculiar to those products. It is necessary to construct a process that takes waste treatment into consideration as soon as the raw materials have been selected.

ECONOMY OF THE FERMENTATION PROCESS

As already mentioned, economy of the fermentation processes must consider the cost of waste-liquor treatment. It is difficult to accurately estimate the economic contribution of waste treatment because this factor is affected by many factors other than fermentation, such as environmental regulation around the location of the plant and so forth. In this discussion, process economics is limited to the steps that begin with the selection of raw materials and end with the desired products. A primary part of the cost of fermentation processes is the cost of the carbon source, which is decided by the *(1)* unit price of the carbon source, *(2)* fermentation yield, and *(3)* purification yield. In L-glutamic acid fermentation, because the carbon source to be used directly affects these three points, selection of the carbon source is particularly important. On the other hand, in the case of fermentation of other amino acids in

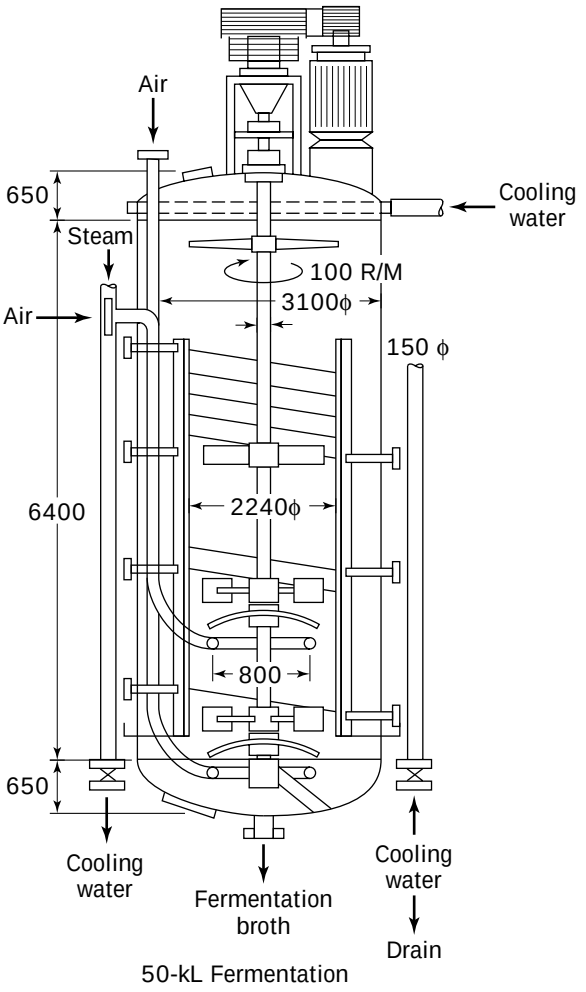


Figure 7. Example of an industrial fermenter. R/M, revolutions per minute; φ, diameter; all distances in millimeters. *Source*: From Y. Su and K. Yamada, *Bull. Agric. Chem.*, 1960, cited in Ref. 17.

Table 6. Rate of Oxygen Sufficiency for L-Glutamic Acid–Family Amino Acids

Amino acid	Rate of oxygen sufficiency[a]	Rate of productivity
Glutamic acid	0.2	0.1
	0.43	0.41
	1	1
Arginine	0.22	0.2
	0.4	0.44
	0.6	0.64
	1	1
Proline	0.22	0.2
	1	1
Glutamine	0.55	0.39
	0.25	0.1
	0.4	0.2
	0.9	0.8
	1	1

Source: Ref. 18.

[a] r_{ab}/KrM, where r_{ab} is the respiratory rate of the microbe (mol of O_2 mL^{-1} min^{-1}), and KrM is the maximum respiratory rate of the microbe (mol of 0_2 mL^{-1} min^{-1}).

which auxotrophic strains are used, the cost of the required nutrient also is a major factor. Thus, it is necessary to use a strain requiring a low concentration of such a nutrient. Production of pharmaceutical-grade neutral amino acids requires difficult separation of analogous amino acids by-produced to the level at which they contribute to increased cost. To overcome this problem, it is necessary to minimize by-produced amino acids to a level at which purification of the desired product is not affected, and to this end, breeding of suitable strains and improvement of fermentation processes that result in few by-products are important. The cost of utilities supporting fermentation is generally high because although the fermentation is carried out mostly at normal temperature and atmospheric pressure, the following factors create additional demands:

1. Sterility must be secured.
2. Aerobic fermentation is involved.
3. It is necessary to more intensely concentrate the fermentation broth because of low product concentration compared to other processes.

AMINO ACID PRODUCTION

Monosodium Glutamate

Corynebacterium glutamicum, which produces L-glutamate, requires biotin for growth but excessively forms L-glutamic acid at a biotin concentration suboptimal for growth. Under the condition of excess biotin, L-glutamic acid can be produced by *(1)* adding penicillin (19) or a surface-active agent (20), or *(2)* culturing an oleic acid–requiring strain at growth-limiting oleic acid concentrations (20,21). Cane molasses, beet molasses, and corn syrup are mainly used as the carbon source because of their low cost. By way of example, L-glutamic acid fermentation by the fed-batch method using cane molasses is described here. Although cane molasses contains most of the

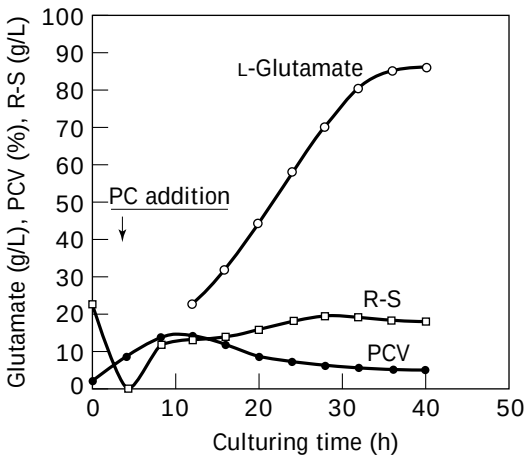


Figure 8. Typical fermentation profile (L-glutamate): *C. glutamicum*, 120-kL fermenter, fed-batch culture, cane molasses medium, penicillin method. R-S, residual sugar; PCV, packed cell volume.

carbon, nitrogen, phosphate, and minerals for microbial growth, its use as a carbon source is basic, and the growth medium is optimized by supplementing with other nutrient sources. Such medium contains an excess concentration of biotin. Potassium salt of penicillin G (8 units/L) is added to the culture broth late in the logarithmic growth phase. Upon addition of penicillin, the microorganism produces L-glutamic acid—mainly due to a change of the permeating system and/or the enzyme activities in TCA cycle (22)—and L-glutamic acid is excreted into the broth. Fermentation is continued while feeding molasses and controlling the pH to neutrality with liquid ammonia. Molasses is fed until L-glutamic acid is no longer produced in the broth. The final L-glutamic acid concentration is around 95 g/L. Figure 8 shows a typical fermentation profile in a commercial plant. The method of penicillin addition provides remarkably high productivity of this amino acid compared with fermentation under suboptimal biotin conditions.

For purification of L-glutamic acid, two methods are available: the direct crystallization method and the chromatographic ion exchange resin method. Because it is well known that the resin method can be used to purify all amino acids produced by fermentation, only the direct crystallization method is discussed here.

An outline of the recovery process (monosodium L-glutamate from L-glutamic acid fermentation broth) in a commercial plant is shown in Figure 9. The purification process may comprise an L-glutamic acid (GA) extraction process and an MSG process to obtain MSG from L-glutamic acid. In Figure 9 the GA extraction process discharges impurities out of the system as a GA mother liquor (containing microbial cells) to recover high-purity GA from the fermentation broth and an MSG mother liquor produced from the MSG process. High purity can be secured by efficiently washing crystals with water at a pH range near the isoelectric point. To this end, the crystal type (there are α- and β-types) must be in the α-form, which can be easily separated by centrifugation.

In the MSG process it is essential to maintain the purity of the product while securing a high yield. These may be achieved by enhancing the purity of L-glutamic acid in the GA process and efficiently using the carbon source to achieve minimum levels of residues. Because the MSG mother liquor contains a high concentration of dissolved MSG, it is sent back into the GA process to augment the total yield of L-glutamic acid.

The GA mother liquor (containing spent cells) and a decolorized active carbon cake are the main fractions dis-

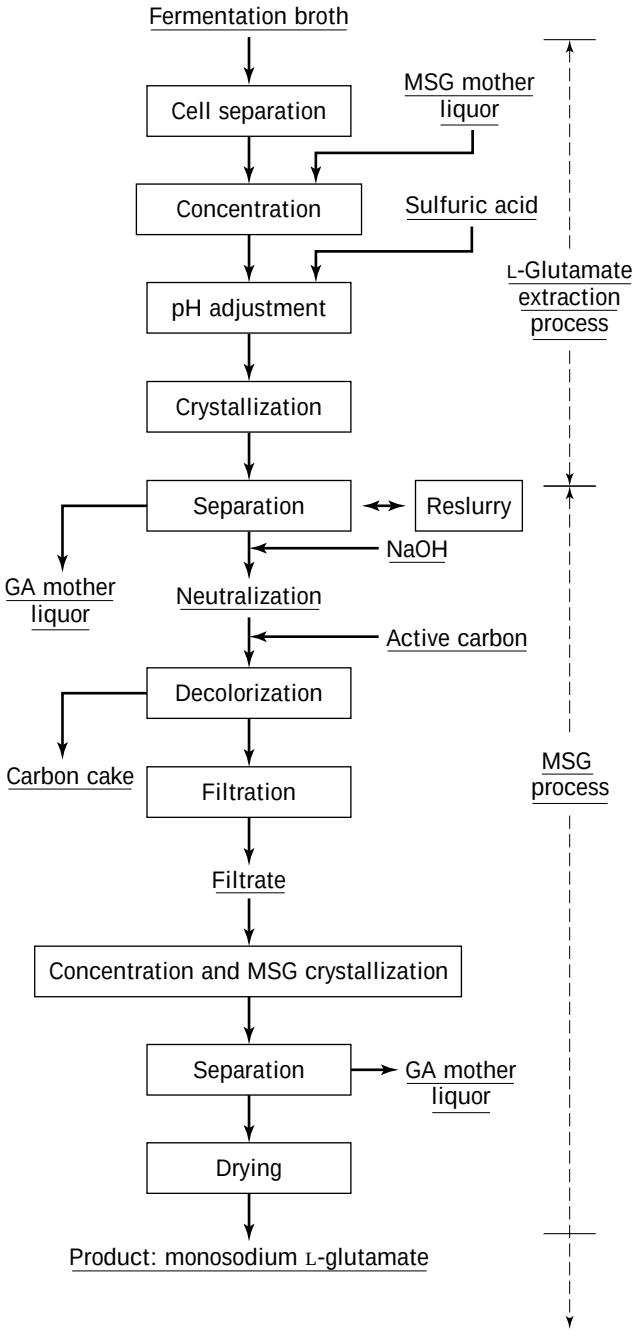


Figure 9. Recovery process flow of monosodium L-glutamate.

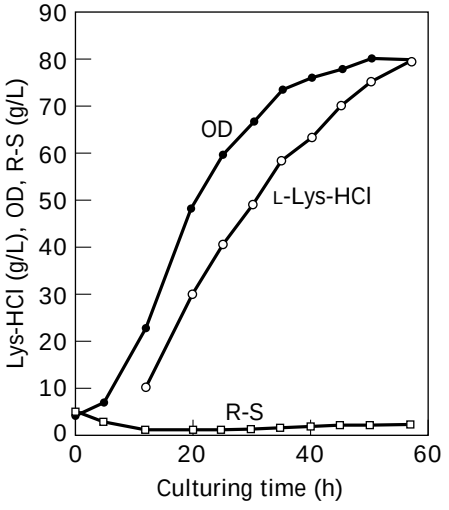


Figure 10. Typical fermentation profile (L-lysine): *C. glutamicum* (Homoser$^-$, Leu^L, AEC^r), 120-kL fermenter, fed-batch culture, molasses medium. OD, indicated cell concentration; R-S, residual sugar.

charged out of the system. Therefore, it is clear that the properties of the fermentation broth strongly affect the total yield of MSG. How the mother liquor of the GA process is used is also important. Hence, with the direct crystallization method of MSG production, there is unification of the whole process from raw materials to final product.

The yield of MSG from the fermentation broth after purification is 65–70 wt %. This is an excellent and effective production system with an attractive process economy if the mother liquor by-product can be used as a fertilizer. As already mentioned, yield is strongly affected by the ratio of L-glutamic acid concentration to total solids in the fermentation broth. This ratio varies, depending upon the purity of raw material sugar and the fermentation efficiency. Whereas the latter depends on the strain of bacteria used, purity of the sugar depends on the quality of the raw material.

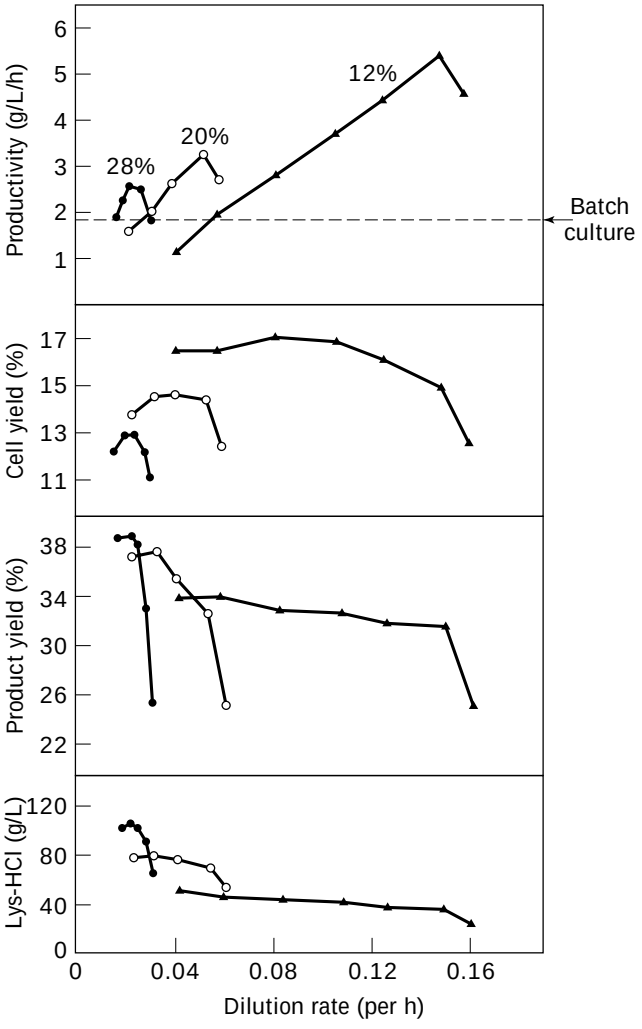


Figure 11. Effect of dilution rate on steady-state parameters for *C. glutamicum* B-6; *S*-(2-aminoethyl)-L-cysteine^r; rifampicin^r, streptomycin^r, 6-azauracil^r. Culture conditions: 5-L jar fermenter, molasses medium, pH 7.0, 32 °C, 600 rpm, air at 3 L/min, working volume of 2.5 L. Stable L-lysine production for over 300 h. Feeding sugar concentration: ▲, 12%; ○, 20%; ●, 28%. *Source*: Ref. 23.

L-Lysine HCl

L-Lysine HCl is demanded in large quantities for animal feed and must be available at low cost. Reduction of cost through the improvement of productivity is very important, as is the quality of the product.

Mutants of *C. glutamicum* (regulatory mutants such as homoserine and/or leucine leaky mutants, which are revertants from auxotrophic mutants and/or lysine analog–resistant mutants) are used. The main raw material for L-lysine fermentation is either molasses or corn syrup. Addition of a biotin source may be required, depending on the raw materials. Molasses contains a sufficient amount of biotin. Because the strains used for L-lysine fermentation are released from metabolic regulation to a remarkable extent, propagation and product formation proceed simultaneously. This is one of the characteristic features of L-lysine fermentation.

In L-lysine fermentation carried out by the fed-batch method, more nitrogen must be supplied compared to the amount required for L-glutamic acid production because L-lysine is a basic amino acid. It is necessary to add ammonium sulfate to the feed solution to supplement the NH_3 supplied for pH control and to neutralize the L-lysine formed. Furthermore, the feeding rate of the solution is controlled to maintain the sugar in the broth at a low concentration because a rapid increase causes catabolite repression. Fermentation is usually complete in about 60 h at a concentration of 95 g/L (as L-lysine HCl). Figure 10 shows a typical profile of the L-lysine fermentation process in a commercial plant. Continuous fermentation of L-lysine is interesting from the viewpoints of productivity and the price required by the market. There has been a report on laboratory-scale steady-state production of L-lysine HCl at a concentration of 80–100 g/L with a productivity of more than 3 g/L per h (Fig. 11). To be successful, however, the following conditions must be satisfied:

1. The equipment for industrial fermentation must ensure a pure culture of *C. glutamicum*.
2. The strain's production of L-lysine must not be decreased in the presence of a high concentration of L-lysine.
3. The strain must be genetically stable.

If these conditions are met, high productivity is secured by determining the optimal dilution rate (feeding volume/working volume) for stable operation (Fig. 11). Continuous fermentation is superior to batch fermentation in that the productivity per fermenter is higher. However, it is difficult to operate a continuous fermenter for a long period, unlike synthetic chemical processes, due to the three requirements just listed. All factors must be taken into consideration when deciding whether to employ this process.

Figure 12 shows a flow diagram of the purification process for feed-grade L-lysine HCl according to the adsorption-desorption method using a cation exchange resin. After separation of the cells, sulfuric acid–acidified broth is passed through the resin for adsorption, followed by washing with water and elution with ammonia. (Lysine is removed from the broth and spent water is sent to the

waste treatment step.) Fractions containing L-lysine are collected and concentrated by removing ammonia (ammonia recovery). The concentrate is neutralized, further concentrated in a vacuum pan, and subjected to crystallization. The resulting mother liquor is spent on the resin-adsorption step and recovered. The yield of the final product from the broth is 80–82 wt %. This process is readily operable and provides stable product quality. However, the relatively high cost of wastewater treatment is a problem because there is a large volume of wastewater with a low-concentration chemical oxygen demand.

L-Threonine

L-Threonine is used in pharmaceuticals and animal feed. In recent years, there has been a growing demand for L-threonine as an additive to low-protein feeds.

The strains used include mutants of *C. flavum* and *Escherichia coli* (mutants requiring diaminopimeric acid, methionine, or isoleucine, or their partial revertants) and recombinants of genes involved in the biosynthetic pathway of threonine. The fermentation process is basically similar to that of L-lysine. Because L-threonine is a neutral amino acid, less nitrogen is required compared with L-lysine fermentation. The purification process differs depending on whether the product is of feed grade or for pharmaceutical use. In the case of feed-grade L-threonine, the resin adsorption method or the concentration–direct crystallization methods may be selected depending upon the composition of the fermentation broth.

L-Tryptophan

L-Tryptophan is used in pharmaceuticals and animal feed. As with L-threonine, feed-grade L-tryptophan has been in increasing demand. L-Tryptophan is produced by direct fermentation, fermentation with a precursor, or an enzymatic method using indole and L-serine as raw materials. In the direct fermentation method, strains of the genus *Corynebacterium* requiring aromatic amino acids and having a resistance to aromatic amino acid analogues or re-

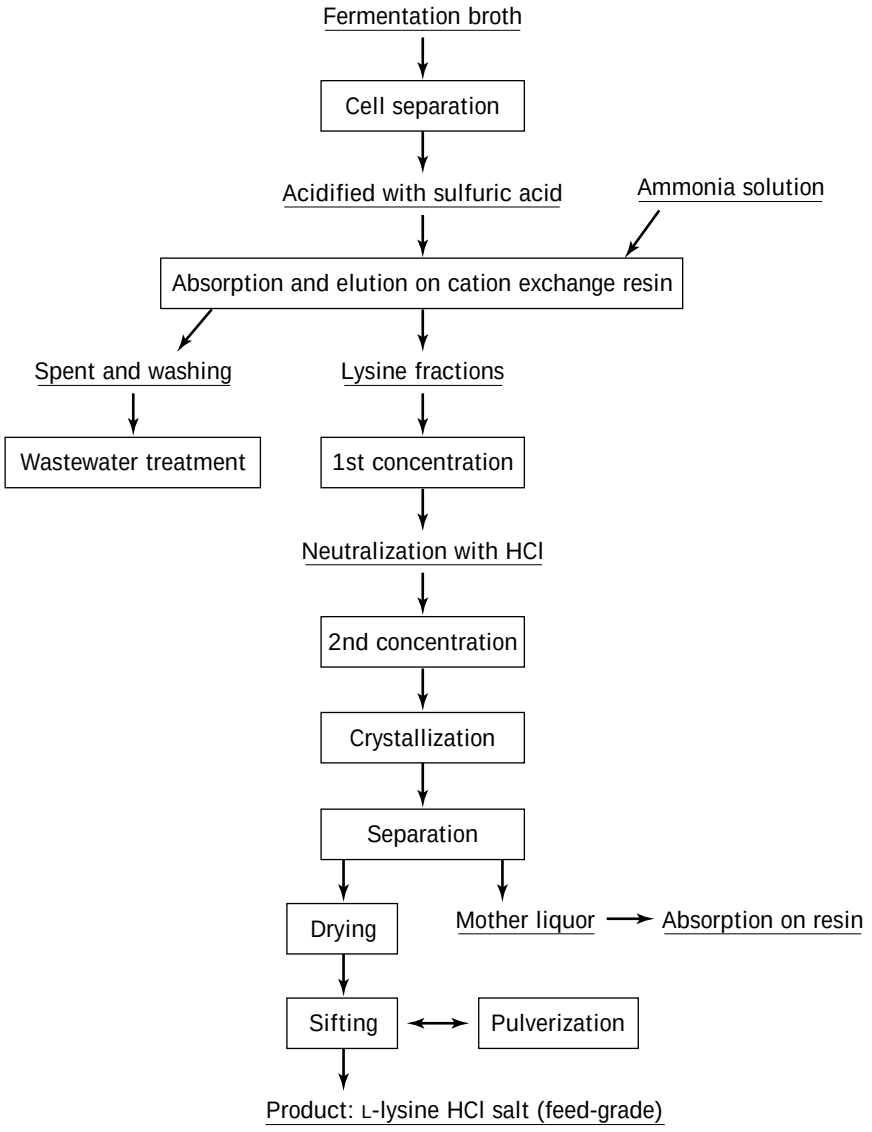


Figure 12. Purification process flow of L-lysine HCl salt (feed-grade).

combinants of aromatic amino acids biosynthetic genes in *E. coli* are used. The direct fermentation method using mutants and recombinants and the enzymatic method are currently practiced. Because L-tryptophan is not highly soluble in the aqueous media, the purification methods have been developed to take this into consideration.

CONCLUSION

Further technical improvement of amino acid fermentation is expected to greatly reduce the production cost of amino acids. Development of L-glutamic acid fermentation resulted in the growth of the seasoning market, and that of L-lysine promoted development of the market for amino acids used in animal feed. Likewise, it is expected that low-cost amino acid production will lead to the development of new uses of amino acids.

BIBLIOGRAPHY

1. S. Kinoshita, S. Udaka, and M. Shimono, *J. Gen. Appl. Microbiol.* **3**, 193–205 (1957).
2. K. Nakayama, S. Kitada, and S. Kinoshita, *J. Gen. Appl. Microbiol.*, **7**, 145–154 (1961).
3. Association for Essential Amino Acids (Japan), Tokyo, Japan.
4. T. Yamane, in T. Yamane ed., Bio-Reaction Engineering, Sangyo-tosho, Tokyo, 1979, pp 167–180.
5. T. Nakano, T. Hirao, T. Azuma, and T. Nakanishi, *Kagaku-kogakkai* (Yamaguchi) **42** (1990).
6. I. Chibata, T. Kakimoto, and J. Kato, *Appl. Microbiol.* **13**, 638–645 (1965).
7. M. Kisumi, Y. Ashikaga, and I. Chibata, *Bull. Agric. Chem. Soc. Jpn.* **24**, 296–305 (1966).
8. T. Hukumura, *Sekiyu-to-Biseibutsu* **15**, 10–17 (1976).
9. P. Zaffaroni, V. Vitobello, F. Cercere, E. Glacomozzi, and F. Morisi, *Agric. Biol. Chem.* **38**, 1335–1342 (1974).
10. H. Yoshida, H. Kumagai, and H. Yamad, *Agric. Biol. Chem.* **38**, 463–464 (1974).
11. K. Sano and I. Shilo, *J. Gen. Appl. Microbiol.* **13**, 349–358 (1967).
12. K. Sano and I. Shilo, *J. Gen. Appl. Microbiol.* **16**, 373–391 (1970).
13. Jpn. Pat. 1445888 (June 30, 1988), K. Nakayama and K. Araki to Kyowa Hakko Kogyo Co., Ltd.
14. H. Hagino, *Nogei-kagakukai-shi* **50**, R79–R81 (1976).
15. K. Miwa, T. Tsuchiya, O. Kurahashi, S. Nakamori, K. Sano, and H. Momose, *Agric. Biol. Chem.* **47**, 2329–2334 (1983).
16. H.C. Vogel, in H.C. Vogel ed., Fermentation and Biochemical Engineering Handbook, Noyes Publication, New Jersey, 1983.
17. S. Kinoshita, in S. Kinoshita ed., Fermentation Technology, Dainippon-tosho, Tokyo, 1975, pp 84–85.
18. K. Akashi, H. Shibai, and Hirose, *J. Ferment. Technol.* **57**, 321–327 (1979).
19. Jpn. Pat. 509952 (January 31, 1968), N.L. Somerson and T. Phillips (to Merck & Co., Inc.).
20. S. Okumura, R. Tsugawa, T. Tsunoda, and S. Kitai, *Nogei-kagakukai-shi* **36**, 197–211 (1962).
21. K. Takinami, H. Okada, and T. Tsunoda, *Agric. Biol. Chem.* **28**, 114–119 (1964).
22. Y. Kawahara, K. Takahashi, E. Shimizu, T. Nakamatsu, and S. Nakamori, *Biosci. Biotech. Biochem.* **61**, 1109–1112 (1997).
23. T. Hirao, T. Nakano, T. Azuma, and T. Nakanishi, *Int. Conf. Ferment. Technol.* Wellington, New Zealand, February 12–15, 1990.

See also AMINOHYDROLASES, FOR PRODUCTION OF D-AMINO ACIDS; FERMENTATION MONITORING, DESIGN AND OPTIMIZATION; GLUTAMIC ACID PRODUCING MICROORGANISMS; METABOLITES, PRIMARY AND SECONDARY; PHENYLALANINE; PRODUCTION OF L-AMINO ACIDS BY AMINOACYLASE.

AMINOHYDROLASES, FOR PRODUCTION OF D-AMINO ACIDS

YASUHIRO IKENAKA
SATOMI TAKAHASHI
Kaneka Corporation
Hyogo, Japan

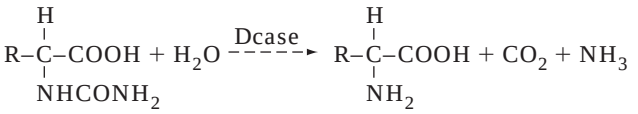
KEY WORDS

D-Amino acid
N-Carbamoyl-D-amino acid amidohydrolase
Thermotolerant enzyme

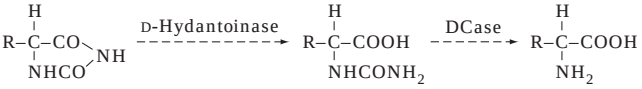
OUTLINE

INTRODUCTION

N-Carbamoyl-D-amino acid amidohydrolase (hereinafter abbreviated as DCase) is the enzyme that catalyzes the following reaction:



The DCase converts *N*-carbamoyl-D-amino acids to D-amino acids with strict D-form specificity and is expected to be applicable to the industrial D-amino acid production process instead of chemical diazonation, in which *N*-carbamoyl-D-amino acids are converted by DCase after hydrolysis of the corresponding 5-substituted hydantoins by the hydrolyzing enzyme, D-hydantoinase. These reactions can be performed separately as two serial reactions or as one batch reaction with microorganisms possessing both enzymes.



D-Amino acids are useful compounds for the preparation of physiologically active peptides and *β*-lactam antibiotics such as semisynthetic penicillins and cephalosporins. DCase activity was first found in rat liver (1) and microbial cells (2–5) about 40 years ago. These enzymes were examined as to their properties and substrate specificities, and their reaction mechanisms and biological functions were deduced. The enzymes from rat liver and *Clostridium uracilicum* (2) were found to be involved in the degradation of pyrimidine, and the enzyme from *Pseudo-*

Table 1. Substrate Specificities of Screened Enzymes

	Mesophile (30 °C)		Thermophile (45 °C)	
	KNK 712	KNK 1415	KNK 003A	KNK 505
c-D-Ala	+ + + +	+ + +	+ + + +	+ +
c-D-Val	+ + + +	+ + +	+ + +	+ +
c-D-Leu	+ + + +	+ + + +	+ + + +	+ + +
c-D-Phe	+ + + +	+ + +	+ + +	+ +
c-D-PEG	+ + +	+ + +	+ +	+
c-D-PG	+ + +	+ + +	+ −	+ −
c-D-HPG	+ + +	+ + +	+ +	+ −

Note: With 1 mL culture broth of each isolated strain, the resting cell reaction was performed under standard conditions at 30 °C (mesophile, 3-h reaction) or 45 °C (thermophile, 24-h reaction), followed by analysis by TLC. Seven *N*-carbamoyl-D-amino acids (c-D-amino acids) were examined, and the amount of D-amino acid produced in the reaction is represented approximately as follows: + −, 0.1 g/L; +, 0.2 g/L; + +, 0.5 g/L; + + +, 1–2 g/L; + + + +, 4 g/L. PEG and PG represent phenetylglycine and phenylglycine, respectively.

Table 2. Properties of DCases from *Agrobacterium* sp. KNK712 and *Pseudomonas* sp. KNK003A

	KNK712	KNK003A
Molecular weight (monomer)		
SDS-PAGE	37,000	38,000
DNA sequence	34,300	35,400
Thermal stability (60 °C, 20 min) (residual activity)	5%	95%
Optimal temperature	67 °C	
Optimal pH	6.8–7.3	7.0–7.5
Number of amino acids (deduced)		
Total	304	312
Hydrophobic	162	154
Neutral	50	51
Hydrophilic	92	107
Cystein	5	5
G + C Content	60.3%	62.9%
Homology		
DNA	62%	
Amino acid	60%	

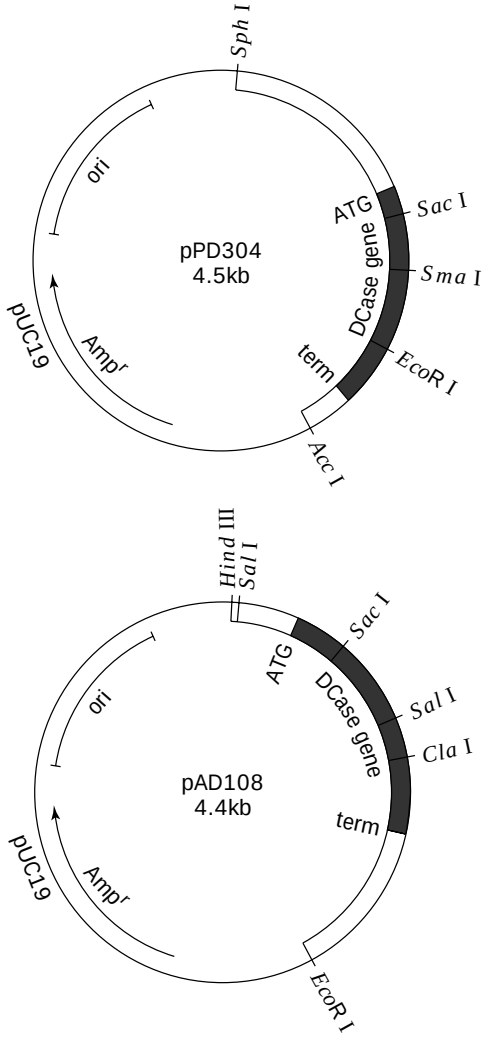


Figure 1. Restriction maps of pAD108 and pPD304. The *thin line* indicates the vector DNA and the *boxed line* the inserted DNA fragment. The open reading frame is represented from ATG to term (the *black-boxed* line).

(a)

```
GTCGACGGCGGGCTCGCGCGAGAGCTTGTCAAGCAGCGCAAATTCCGGTTCCGCTCCGGTTGACAGATCAAAAATTTTACGCCTGTTATTGTCGTGCTGC  100

ATGTAATATTTCGTACTTTATGTAGAATTTGCATTGCGCCGCGAGTCACAAAGCCGGTTTTCGGCGATGTGTTTCACAACGTTTTCCCGGCCGCTGGGCC  200

GGACATCACCTAGGAAGGAGCAGAGGTTCATGACACGTCAGATGATACTTGCAGTGGGACAACAAGGTCCGATCGCGCGCGCGGAGACACGCGAACAGGT  300
                                M T R Q M I L A V G Q Q G P I A R A E T R E Q V

CGTCGTTCGTCTTCTCGACATGCTGACGAAAGCCGCGAGCCGGGGCGCGAATTTCATTGTCTTCCCCGAACTCGCGCTTACGACCTTCTTCCCGCGCTGG  400
 V V R L L D M L T K A A S R G A N F I V F P E L A L T T F F P R W

CATTTCACCGACGAGGCCGAGCTCGATAGCTTCTATGAGACCGAAATGCCCGGCCCGGTGGTCCGTCCACTCTTTGAGAAGGCCGCGGAACTCGGGATCG  500
H F T D E A E L D S F Y E T E M P G P V V R P L F E K A A E L G I G

GCTTCAATCTGGGCTACGCTGAACTCGTCGTCGAAGGCGGCGTCAAGCGTCGCTTCAACACGTCCATTTTGGTGGATAAGTCAGGCAAGATCGTCGGCAA  600
  F N L G Y A E L V V E G G V K R R F N T S I L V D K S G K I V G K

GTATCGTAAGATCCATTTGCCGGGTCACAAGGAGTACGAGGCCTACCGGCCGTTCCAGCATCTTGAAAAGCGTTATTTCGAGCCGGGCGATCTCGGCTTC  700
 Y R K I H L P G H K E Y E A Y R P F Q H L E K R Y F E P G D L G F

CCGGTCTATGACGTCGACGCCGCGAAAATGGGGATGTTCATCTGCAACGATCGCCGCTGGCCTGAAGCCTGGCGGGTGATGGGCCTCAGGGGCGCCGAGA  800
P V Y D V D A A K M G M F I C N D R R W P E A W R V M G L R G A E I

TCATCTDCGGCGGCTACAACACGCCGACCCACAATCCCCCTGTTCCCCAGCACGACCACCTGACGTCCTTCCACCATCTCCTATCGATGCAGGCCGGGTC  900
  I C G G Y N T P T H N P P V P Q H D H L T S F H H L L S M Q A G S

TTATCAGAACGGGGCCTGGTCCGCGGCCGCGGGCAAGGTGGGCATGGAGGAGAACTGCATGCTGCTCGGCCACTCCTGCATCGTGGCGCCGACCGGGGAA 1000
 Y Q N G A W S A A A G K V G M E E N C M L L G H S C I V A P T G E

ATCGTCGCTCTCACTACGACGCTGGAAGACGAGGTGATCACCGCCGCCGTCGATCTCGATCGCTGCCGGGAACTGCGTGAACACATCTTCAACTTCAAGC 1100
I V A L T T T L E D E V I T A A V D L D R C R E L R E H I F N F K Q

AGCATCGTCAGCCCCAGCACTATGGTCTGATCGCGGAACTCTGAGGTTGCCGAAAAGCATGTGTGTCGTTGTTCTCGGCGCCTGGGTCACATCCAGGCGC 1200
  H R Q P Q H Y G L I A E L

GCCAGGGTGACGCTGGTGGAATAGTACCACGACCGCTTCAGGGCGATCCGCAAGGAGATGCGGGTCGCCGGAGCGGCAAAGCCCGACATTCGTTTCGCAC 1300

CGACGGCCGTCGTGAACTCGACAGTCCGCGAGAAGGGCGTATTGCGCGGCCTGGACCTGTACGTGGAACTGTAGCCCATATATAGATTTCCAAAGAGTTT 1400

CGGCGAGGCGCGGCGCGCCTAGCCCCATGTGAGCGAGAACCGTGCCCAGATCAAAGAATGAGACCGACGCGCCGGCCGCGGCAAAGGATGATCCTCAGGG 1500

TCGGATCTATCGCTCCGCCCTGAAGCAGGAGGGCGCACGCTGGCTGCTGACGGCGGAGGAAGGGTTGCTGGCAAAGCCCAAGCCGCCCGGCCTTGTTCCG 1600

GCACTTGAGAATGCGATCGCCATCGTCGATTACATCAACGGTACACCGCCCCATATCGCGTCCCTGGCGGAGCTTTCAACGACGCTCGGGATATCCAAGA 1700

GCCACTGTCACTCCATCCTCAAGACGCTGACGCATTTCGGCTGGCTGAAATTCGACAATCGCTCAAAGAGCTACGAGCTGAATTC
```

Figure 2. Nucleotide sequences of the insert DNAs of pAD108 (**a**) and pPD304 (**b**) and the deduced amino acid sequences of these DCases. The numbers at the *right* are nucleotide positions. Sequences similar to the −10 and −35 consensus sequences for the *E. coli* promoter are *double-underlined*, and the putative ribosome-binding site (Shine–Dalgarno) sequence and initiation and termination codons are also *double-underlined*. The open reading frame of the DCase is *underlined*.

monas putida 77 (3), which is *N*-carbamoyl-sarcosine amidohydrolase, is involved in creatinine metabolism. Olivieri et al. found DCase activity of *Agrobacterium radiobacter* NRRL B11291 in intact cells or a cell-free extract, together with D-hydantoinase activity (4). Yokozeki and Kubota attempted the partial purification of DCase from *Pseudomonas* sp. AJ11220 (5). These two enzymes were examined as to their substrate specificities and reaction profiles. They showed relatively broad substrate specificities such as for aliphatic and aromatic compounds, with strict D-form specificity, and almost the same optimal conditions, that is, pH 7 and 55–60 °C. Ogawa et al. screened enzyme producers by means of enrichment cultures with the *N*-carbamoyl-D-amino acids citrulline and ornithine as sole nitrogen or carbon sources. They isolated alkaliphilic and thermotolerant enzymes, which were produced by strains classified as *Comamonas* sp. and *Blastobacter* sp., respectively, purified them to homogeneity, and determined their properties (6,7). These enzymes also showed strict specificity toward *N*-carbamoyl-D-amino acids and hydrolyzed *N*-carbamoyl-amino acids having a hydrophobic side chain very efficiently; on the other hand, those having a polar group or short-chain alkyl group were only weakly hydrolyzed.

(b)

```
GCATGCGCGGGGAACTGAAGAACTTGCAAGACGAACTCGGCATTACCTTCGTGCATGTAACCCATACCCAGCCTGAGGCGATCGCGCTCGCCGACATGGT   100

GGTTGTGATGGATACGGGCCGCATAGAGCAGGCAGCGAGCGCCAACGAAATCTACAACCGGCCCGCGACGCCCTATGTGGCGCGCTTCATGGGCGGCCAA   200

AACGTGTTGACGGGGAGGGTGGAGAGCATCTCGCCCACCGGCATGGTGCTGAAAAGCGAAAAGGGCGAGATCTTCAATGCGCCTCTTACGGGTGCTGCGC   300

CGAAGCTGGGCGAACCCGTATCGATATCCATGCGCCGCGACCGCATCAGCATCAGCAAGCCGCAAAACGGCAAGGGCGCGCAGCAGGCTGACGCGGTAAC   400

GGGTGTGGTCGATTCCACGGAATACCAGGGCAGCTTCGTGAAGGTCAGCATAGTGCTCGACGGTGGCGAGACCTTCGTCGCAAACATGCCCGACCATGAA   500

TTTTTCGCGGAACCGGTGGATCACGGCGTCCCGGTGGTCGCCCGCTGGAAACCGGAGCATGTGCATGTCCTGTCCAAGTCTGACCGGGGCGCCGACCACA   600

CCGAAATCTACCGCTTCCCTGCAGGCGAAAATACCGTTTCAATGGGCAAGGGGCGGCAAACGGGGTTGAGACGACCCGGTTTATCGAGGAGGACGAGATG   700
                                                                                                    M

ACACGCATCGTCAATGCAGCCGCCGCGCAGATGGGGCCCATCAGCCGGTCCGAAACGCGCAAGGATACGGTCCGGCGCCTGATCGCGCTCATGCGCGAGG   800
T  R  I  V  N  A  A  A  A  Q  M  G  P  I  S  R  S  E  T  R  K  D  T  V  R  R  L  I  A  L  M  R  E  A

CGAAGGCCCGCGGTTCCGACCTTGTCGTCTTTACCGAACTCGCGCTCACCACCTTCTTTCCCCGCTGGGTGATCGAGGACGAAGCTGAGCTCGACAGCTT   900
   K  A  R  G  S  D  L  V  V  F  T  E  L  A  L  T  T  F  F  P  R  W  V  I  E  D  E  A  E  L  D  S  F

CTACGAGAAGGAGATGCCAGGGCCCGAAACCCAGCCGCTCTTCGATGAGGCGAAGCGCTTGGAGATCGGCTTCTATCTCGGTTATGCCGAGCTGGCGGAG  1000
   Y  E  K  E  M  P  G  P  E  T  Q  P  L  F  D  E  A  K  R  L  E  I  G  F  Y  L  G  Y  A  E  L  A  E

GAGGGCGGCAGGAAGCGGCGCTTCAACACCTCTATCCTTGTGGACCGCAGCGGCCGGATCGTCGGCAAGTACCGCAAGGTGCACCTGCCCGGGCACAAAG  1100
E  G  G  R  K  R  R  F  N  T  S  I  L  V  D  R  S  G  R  I  V  G  K  Y  R  K  V  H  L  P  G  H  K  E

AGCCGCAGCCCGGCAGGAAACACCAGCATCTCGAGAAACGCTATTTCGAGCCCGGCGATCTCGGCTTCGGTGTCTGGCGCGCCTTCGACGGCGTAATGGG  1200
   P  Q  P  G  R  K  H  Q  H  L  E  K  R  Y  F  E  P  G  D  L  G  F  G  V  W  R  A  F  D  G  V  M  G

CATGTGCATTTGCAACGACCGCCGCTGGCCGGAGACCTACCGGGTCATGGGCTTGCAGGGAGTGGAGATGGTCATGCTGGGCTACAACACGCCGTATGAC  1300
M  C  I  C  N  D  R  R  W  P  E  T  Y  R  V  M  G  L  Q  G  V  E  M  V  M  L  G  Y  N  T  P  Y  D

CATACCGGTCACGACGACATCGATTCACTCACCCAGTTTCACAATCATCTCTCCATGCAGGCGGGCGCCTACCAGAATTCGACCTGGGTGATCGGCACCG  1400
H  T  G  H  D  D  I  D  S  L  T  Q  F  H  N  H  L  S  M  Q  A  G  A  Y  Q  N  S  T  W  V  I  G  T  A

CCAAATGCGGCACCGAGGAGGGCTCCAAAATGGTGGGGCAGAGCGTGATCGTTGCGCCCTCCGGCGAGATCGTCGCTATGGCCTGCACGATCGAGGACGA  1500
   K  C  G  T  E  E  G  S  K  M  V  G  Q  S  V  I  V  A  P  S  G  E  I  V  A  M  A  C  T  I  E  D  E

GATCATCACCGCACGCTGCGATCTCGACATGGGCAAGCGCTACCGCGAGACCATCTTCGATTTCGCCCGCCATCGCGAGCCCGACGCCTATCGCCTGATC  1600
   I  I  T  A  R  C  D  L  D  M  G  K  R  Y  R  E  T  I  F  D  F  A  R  H  R  E  P  D  A  Y  R  L  I

GTCGAACGCAAAGGGGCTGTGCCGCCGCCGCAGTGATCGGAACCTGAAAACGAAATATCCCGCCGGACGGTGGGAAGGTGAAAGGAGGAGTCTCCATGAC  1700
V  E  R  K  G  A  V  P  P  P  Q

AACAGTTATCAAGGGTGGAACGATCGTCGCCGCCGATCGCAGCTATGAAGCCGATATCCTGATCGAAGGCGAAAAGATCGCCCAGATCGGCAGGGATCTG  1800

CAGGGCGACAAGATTGTCGAC
```

Figure 2. Continued.

However, these enzymes showed that the *N*-carbamoyl compounds involved in the metabolism of pyrimidine and purine are not hydrolyzed at all. On characterization of these enzymes, the reactive molecular weights of the native enzymes were found to be about 120,000 and those of the subunits to be about 40,000. Enzymochemical analysis was also performed. Recently, Louwrier and Knowles purified a DCase from *Agrobacterium* sp. which was composed of genetically engineered self-cloning cells, and characterized it (8). This enzyme was able to cleave a variety of *N*-carbamoyl substrates but was strictly D-form specific. The active enzyme was suggested to be present as a dimer with a subunit molecular weight of 38,000 Da, differing from the trimer enzymes of Ogawa et al. (6,7).

Concerning the enzyme genes, some DCase genes have been cloned and analyzed. The DCase gene from *Agrobacterium radiobacter* NRRL B11291 was cloned, analyzed, and expressed in *Escherichia coli*, and the recombinant enzyme was then characterized (9,10). By means of a site-directed mutagenesis experiment, the relationship between activity and amino acid substitutions was examined, and some mutations concerning enzyme stability were found. Neal et al. also isolated the DCase gene of *Agrobacterium* sp. and expressed the gene in *E. coli* and *Agrobacterium* sp. (11).

We also screened some strains producing the enzyme from mesophile (12) and thermotolerant strains (13) and cloned two enzyme genes. We tried to improve the native DCase to obtain a practical DCase that exhibits both high reactivity and sufficient stability for repeated use in a bioreactor system by means of amino acid substitutions using recombinant DNA technology (14,15). We succeeded in creating a practical DCase by the substitution of three amino acids (16), and applied it to an industrial production process as an immobilized enzyme (17,18).

In this section, we report our attempts to establish a new D-amino acid production process, such as screening of DCase-producing bacteria from soil, purification of the enzymes, cloning of the enzyme genes, mutagenesis to obtain thermostabilized enzymes, and immobilization of the enzymes for a bioreactor system to produce D-amino acids.

SCREENING OF *N*-CARBAMOYL-D-AMINO ACID AMIDOHYDROLASE

Screening of DCase-Producing Microorganisms

The screening and isolation methods used for soil microorganisms were as follows (12). For the isolation of microorganisms that can hydrolyze *N*-carbamoyl-D-amino acids to D-amino acids, soil samples were suspended in saline and the supernatants were inoculated into growth medium containing *N*-carbamoyl-D-amino acids as sole nitrogen sources for enrichment culture. After aerobic cultivation, the reduction of *N*-carbamoyl-D-amino acids and the production of D-amino acids were detected by silica gel thin-layer chromatography (TLC) with a solvent system of *n*-butanol:acetic acid:water (4:1:1). *N*-Carbamoyl-D-amino acids and D-amino acids were detected with *p*-dimethylaminobenzaldehyde in a 6 M HCl solution and ninhydrin, respectively. From the culture broth from which *N*-carbamoyl-D-amino acids disappeared or in which D-amino acids accumulated, microorganisms were isolated on agar plates.

The substrate specificities of these DCases were examined by means of the resting cell reaction with detection by the TLC method (Table 1). We screened two good mesophile strains producing a lot of enzymes that were classified as *Agrobacterium* sp. or *Rhizobium* sp. (12). A similar enzyme-producing strain, classified as *Agrobacterium radiobacter*, was also found by Olivieri et al. (4) and characterized. We also screened thermotolerant strains by means of enrichment culture at 45 °C and obtained some *Pseudomonas* sp. strains (Table 1) (13). The strain isolated by Yokozeki and Kubota was classified as *Pseudomonas* sp. (5) but was not a thermotolerant strain.

Purification and Characterization of the Enzymes

The DCases from *Agrobacterium* sp. KNK712 and *Pseudomonas* sp. KNK003A were purified after large-scale incubation by ammonium sulfate precipitation and chromatography such as ion exchange and gel filtration.

DCase activity was assayed by measurement of D-*p*-hydroxyphenyl glycine (D-HPG) production from *N*-carbamoyl-D-HPG (c-D-HPG) (12). The reaction was started by the addition of 100 μL of an enzyme solution diluted with 100 mM potassium phosphate buffer, pH 7.0 (hereinafter abbreviated as KP buffer), containing 5 mM dithiothreitol (DTT) to the assay mixture containing 47.6 μmol c-D-HPG and 100 μmol KP buffer in a total volume of 1 mL. After 20 min incubation at 40 °C, the reaction was stopped by the addition of 0.25 μL of 20% (w/v) trichloroacetic acid. The D-HPG produced was analyzed by means of the following high-performance liquid chromatography (HPLC) method. *N*-Carbamoyl-amino acids and amino acids were detected and quantified by HPLC at 210 nm on a reverse-phase HPLC column (e.g., Finepack SIL C18-5 column; Nipponbunko), using 36.7 mM KH_2PO_4 containing 15% methanol adjusted to pH 2.5 with phosphoric acid. One unit of the enzyme was defined as the amount of enzyme that catalyzed the formation of D-HPG at the rate of 1 μmol/min under the assay conditions already mentioned.

The reactivity of the DCase of *Agrobacterium* sp. KNK712 is about 20 times higher than that of *Pseudomonas* sp. KNK003A with c-D-HPG as the substrate, but as concerns heat stability, in contrast, the *Pseudomonas* sp. KNK003A enzyme is more stable, with an about 10 °C increase in view of its denaturation temperature.

The molecular weights of these DCases were found to be about 37,000 and 38,000 Da, respectively, on SDS-PAGE. The properties of these two enzymes are shown in Table 2.

ISOLATION OF ENZYME GENES

Cloning of Enzyme Genes

The DCase genes from *Agrobacterium* sp. KNK712 and *Pseudomonas* sp. KNK003A were isolated and their DNA sequences analyzed as follows (12). Chromosome DNA was partially digested with *Sau*3AI, and the fractionated DNA

fragment of 4–9 kb was inserted into pUC18 and then transformed into *E. coli* JM109. Recombinant colonies were collected, inoculated into enrichment culture medium containing *N*-carbamoyl-D-amino acids or 5-substituted hydantoins, and then incubated at 37 °C. After repeating the enrichment cultures, the recombinant clones were isolated. Plasmid DNAs were prepared from these clones and analyzed with several restriction endonucleases. The DCase gene of *Agrobacterium* sp. KNK712 was located in a 1.8-kb *Sal*I–*Eco*RI fragment, and the gene of *Pseudomonas* sp. KNK003A was found in a 1.8-kb *Sph*I–*Acc*I fragment. These fragments were inserted into pUC19, the resultant plasmids being designated as pAD108 and pPD304, respectively. The restriction maps are shown in Figure 1. The recombinant strains containing these plasmids showed increased DCase expression compared with the native strains.

Analysis of the Enzyme Genes

The DNA sequences of the genes in the 1.8-kb fragments were analyzed (Fig. 2), there being one open reading frame in each case, consisting of 912 and 936 bases with a starting triplet, ATG, and predicted to encode polypeptides of 304 and 312 amino acids with calculated molecular weights of 34,285 and 35,438, respectively. Upstream of the open reading frame, sequences similar to the −35 and −10 consensus sequences for *E. coli* promoters and a putative ribosomal-binding site (Shine–Dalgarno sequence) were found. The G + C contents of the open reading frames were 60.3% and 62.9%, respectively (see Table 2).

The homologies of the DNA sequences and the deduced amino acid sequences between these two DCases were 62% and 60%, respectively. An 8–amino acid length of the C-terminal region was different in these two enzymes. The homologies among those of *Agrobacterium* sp. KNK712 and those of two other reported *Agrobacterium* strains (9,10) which had completely the same sequences, were 97.0% (amino acid sequences) and 93.0% (DNA sequences). The homologies of the N-terminal regions of 30 amino acids in five DCases (6,7,9,12,13) were about 50%, the homology between the enzyme of *Pseudomonas* sp. KNK003A and the *Comamonas* sp. E222c DCase being the highest (56.7%), besides the homology among *Agrobacterium* strains. Multiple alignment of the amino acid sequences of these enzymes is shown in Figure 3.

Similarity research using a database (PIR release, 45.0) showed that the aliphatic amidase of *Pseudomonas aeruginosa* (19) showed the highest similarity (30%) to DCase, but a closely related enzyme was not found. Multiple alignment with Clustal W (1.4) software of the DCase and 10 known related enzymes (PIR accession numbers [PIR Ac. No.] are given in the legend to Fig. 4) showed that 6 amino acids, that is the 119th amino acid of DCase, glycine (hereinafter abbreviated as 119th Gly), 126th Arg, 127th Lys, 146th Glu, 153rd Gly, and 172nd Cys, were all at corresponding positions.

IMPROVEMENT OF THE ENZYME IN STABILITY

For its use in a bioreactor, DCase was immobilized on a resin and used repeatedly. Although a DCase requires both high reactivity and high enzyme stability for such use, the known DCases do not have both properties. Thus, the DCase gene of *Agrobacterium* sp. KNK712, which exhibits high reactivity, was mutagenized to increase the enzyme's thermostability as an index of enzyme stability.

Mutagenesis

The DCase gene of *Agrobacterium* sp. KNK 712 was separated from recombinant plasmid pAD108 and then inserted into M13mp18. Random mutation was performed using two mutagens, hydroxylamine hydrochloride and $NaNO_2$ (14,15). For the mutation with hydroxylamine, recombinant phage particles containing the DCase gene were treated with 0.25 M mutagen (pH 6.0) for 1–8 h at 37 °C. For the mutation with $NaNO_2$, single-strand recombinant phage DNA was treated with 0.9 M mutagen (pH 4.3) for 30 min at 25 °C. After preparation of double-strand phage DNA, the 1.8-kb *Eco*RI–*Hind*III fragment corresponding to the DCase gene was prepared, inserted into pUC19, and then transformed into *E. coli* JM109.

The recombinant clones were screened by means of a newly developed colorimetric enzyme assay to select thermostabilized enzyme-producing colonies. The method is as follows. For first screening for colony assaying, recombinant *E. coli* colonies on the plate medium (the plates being called the master plates) were transferred to sterilized filter paper and then lysed for 30 min at 37 °C by adding a lysis solution containing lysozyme and Triton X-100. After inactivation of the heat-sensitive enzymes (65 °C, 5 min), a color reagent (1 mL), a mixture of two solutions (Sol. A, 30 mM potassium phosphate buffer, pH 7.4, 3 mg/mL *N*-carbamoyl-D-phenylglycine, 10 mM phenol, and 0.8 u/mL D-amino acid oxidase; Sol. B, 70 u/mL horseradish peroxidase and 50 mM 4-aminoantipyrine: Sol. A:Sol. B = 100:1, mixed just before use) was added to the filter paper, followed by incubation at 37 °C. The candidate colonies, which were colored red, were separated from the master plates. These clones were then subjected to the following second screening for cell-free assaying to confirm the mutant clones.

Recombinant *E. coli* cells were disrupted by sonication and the heat-sensitive enzymes were then inactivated (65 °C, 10 min). The crude enzyme sample was dispensed in 150-μL aliquots onto a 96-well microtiter plate; these aliquots were serially diluted twofold and 50 μL of the modified color reagent (Sol. A:Sol. B = 19:1) was added. The plate was incubated at 37 °C, and then the enzyme activity was measured as absorbance at 505 nm.

The candidate clones that exhibited high absorbance were separated and the enzyme genes were analyzed.

Nucleotide Sequence Analysis of the Mutagenized DCase Genes

The DNA sequences of the 1.8-kb *Eco*RI–*Hind*III fragments that contained the mutagenized DCase genes from these representative clones were analyzed (Table 3). One or some nucleotides of the DCase gene were changed. The amino acid changes related to the thermostability increase proved that the 57th amino acid of DCase, histidine, was changed to tyrosine, the 203rd proline was changed to ser-

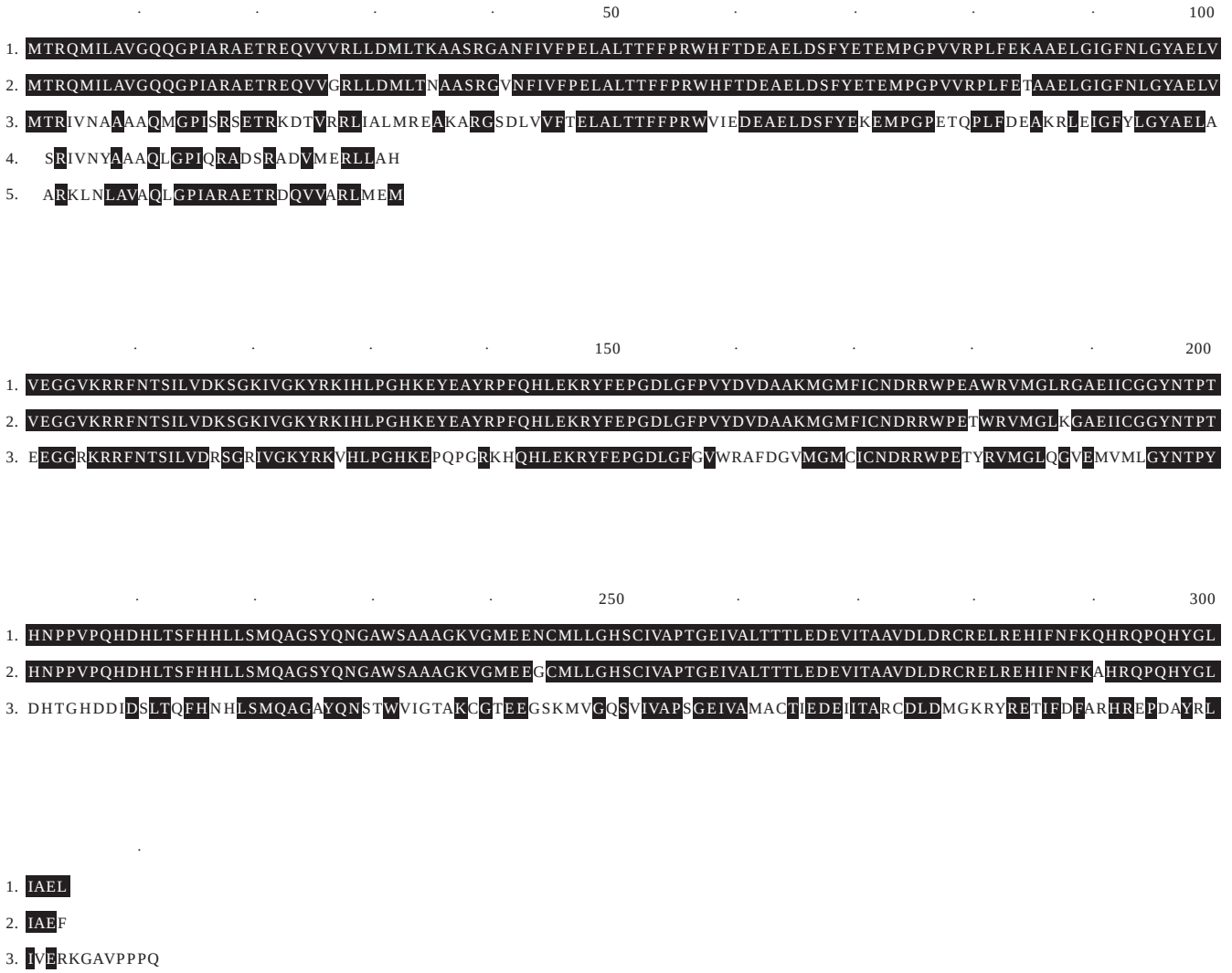


Figure 3. Multiple alignment of the deduced amino acid sequences of DCases. The amino acid sequences of the DCases of *Agrobacterium* sp. KNK712 and homologous amino acid residues in other DCases are shown in *black boxes. 1, Agrobacterium* sp. KNK712; *2, Agrobacterium radiobacter* NRRL B11291; *3, Pseudomonas* sp. KNK003A; *4, Comamonas* sp. E222C; *5, Blastobacter* sp. A17p-4.

ine or leucine, and the 236th valine was changed to alanine.

Amino Acid Substitution of Thermostability-Related Sites by PCR

The three thermostability-related sites of the DCase were substituted (15) through site-directed random amino acid changes by use of the PCR technique according to Ito et al. (20), and then thermostabilized DCase-producing strains were screened by means of the colorimetric colony assay method. As a result of DNA analysis of the mutant thermostabilized DCases, it was proved that the following amino acid changes at the sites increased the thermostability in addition to the known amino acid changes; the 57th amino acid of DCase was changed to leucine; the 203rd amino acid to alanine, asparagine, glutamate, histidine, isoleucine, or threonine; and the 236th amino acid to serine or threonine. These results are shown in Table 4.

Combination of Thermostability-Related Mutations

To produce multiple mutants with combinations of two or three thermostability-related amino acid mutations, restriction fragments containing the mutations were used to replace the corresponding mutation-free DNA fragments (16). In the DCase of *Agrobacterium* sp. KNK712, the mutation of the 57th amino acid is located in a *Nde*I–*Sac*I DNA fragment of about 190 bp; the 203rd amino acid mutation is in a *Sal*I–*Cla*I DNA fragment of about 170 bp (hereinafter this fragment is referred to as fragment A); and the 236th amino acid mutation is in a *Cla*I–*Sph*I DNA fragment of about 75 bp. These fragments were replaced and the thermostabilities of the enzymes were then mea-

Figure 4. Optimal alignment with the CLUSTAL W program of the high-scoring segments in the predicted amino acid sequences of DCases and related enzymes. Identical amino acid residues in all enzymes are indicated by *asterisks*. *1, Agrobacterium* sp. KNK712 DCase (104–192); *2, Agrobacterium radiobacter* NRRL B11291 DCase (104–192); *3, Pseudomonas* sp. KNK003A DCase (104–192); *4, Brevibacterium* sp. R312 aliphatic amidase, PIR Ac. NO. JC1174 (104–186); *5, Pseudomonas aeruginosa* aliphatic amidase, PIR Ac. No. A26741 (115–186); *6*, rat beta-alanine synthetase, PIR Ac. No. S27881 (176–252); *7, Arabidopsis thaliana* nitrilase, PIR Ac. No. S31969 (126–199); *8, Arabidopsis thaliana* nitrilase, PIR Ac. No. S22398 (133–206); *9, Rhodococcus rhodochrous* aliphatic nitrilase, PIR Ac. No. A43470 (117–190); *10, Rhodococcus rhodochrous* nitrilase, PIR Ac. No. A45070 (112–185); *11, Klebsiela ozaenae* nitrilase, PIR Ac. No. A28658 (108–181); *12, Alcaligenes faecalis* nitrilase, PIR Ac. No. A47181 (110–183); *13, Gloeocercospora sorghi* cyanide hydratase, PIR Ac. No. JQ1613 (109–183).

Table 3. Nucleotide Analysis of Mutagenized DCase Genes

Mutant	Locations of nucleotide changes	Amino acid substitutions
401M	840 C → T	203 Pro → Leu
402M	401 C → T	57 His → Tyr
	715 C → T	(161 Val → Val)
403M	177 C → T	(noncoding)
	499 C → T	(89 Ile → Ile)
	839 C → T	203 Pro → Ser
404M	840 C → T	203 Pro → Leu
406M	839 C → T	203 Pro → Ser
414M	939 T → C	236 Val → Ala

Note: The mutagenized DCase genes of thermostable enzyme-producing strains were analyzed and the locations of the nucleotide changes were determined. The amino acid substitutions resulting from the mutations were deduced.

sured (Table 5). The thermostabilities of the enzymes increased cumulatively with the accumulation of the individual mutations. Among the mutant enzymes with two amino acid changes, a mutant with mutations of the 203rd amino acid, proline, to glutamate, and the 236th, valine, to alanine, showed a thermostability increase of about 17 °C. Among the mutant enzymes with three amino acid changes, a mutant enzyme from KNK455M that had mutations of the 57th amino acid to tyrosine, the 203rd to glutamate, and the 236th to alanine, showed about 19 °C increase in thermostability.

CHARACTERIZATION OF THE THERMOSTABILIZED ENZYMES

Effects of Temperature and pH on Enzyme Stability

To evaluate the thermostability of the DCases, a cell-free extract (100 μL) was incubated at various temperatures (55–85 °C) for 10 min, the denatured protein was removed by centrifugation, and then the residual activity was measured by the standard HPLC method. The heat denaturation profiles of the four mutant enzymes from KNK402M, KNK404M, KNK416M, and KNK455M and the native enzyme were investigated. The relative activities in comparison with the activities of the non-heat-treated enzymes were plotted against temperature (Fig. 5). The thermostabilities of the mutant enzymes increased about 5–20 °C under these conditions compared to that of the native enzyme.

To examine the effect of pH on DCase stability, a cell-free extract (200 μL) was mixed with 800 μL of a pH-adjusted buffer (pH 6–10, 0.1 M each), followed by incu-

Table 4. Thermostability Analysis of Mutant DCases

Location of mutation	Substituted amino acid	Thermostable temperature (°C)
57 His	Tyr*	67.3
	Leu	67.5
203 Pro	Leu*	68.0
	Ser*	66.5
	Ala	67.7
	Asn	67.0
	Glu	70.0
	His	65.2
	Ile	67.2
	Thr	67.5
236 Val	Ala*	71.4
	Ser	72.0
	Thr	69.5
Native type	—	61.8

Note: The mutagenized DCase genes were analyzed and the amino acid substitutions resulting from the mutations were deduced. To compare the thermostability of the mutant enzymes easily, the index of thermostability, thermostable temperature, was defined as the temperature on heat treatment for 10 min that caused a decrease in DCase activity of 50%. The substituted amino acids, on random mutagenesis, in these mutants are indicated by asterisks.

Table 5. Thermostability of Mutant DCases Having Two or Three Mutations

Strain	Thermostable temperature[a] (°C)	Amino acid mutation		
		57 His	203 Pro	236 Val
KNK712	61.8	—	—	—
KNK421M	71.5	Tyr	Leu	—
KNK422M	70.0	Tyr	Ser	—
KNK423M	75.4	Tyr	—	Ala
KNK424M	77.5	—	Leu	Ala
KNK425M	75.9	—	Ser	Ala
KNK426M	78.8	Tyr	Leu	Ala
KNK427M	77.8	Tyr	Ser	Ala
KNK451M	74.0	Tyr	Glu	—
KNK452M	75.3	Tyr	—	Ser
KNK453M	76.0	—	Glu	Ser
KNK461M	79.0	—	Glu	Ala
KNK454M	80.4	Tyr	Glu	Ser
KNK455M	80.8	Tyr	Glu	Ala
KNK456M	78.5	Tyr	Leu	Ser

[a]Thermostable temperature is defined as the temperature on heat treatment for 10 min that caused a decrease in the DCase activity of 50%.

bation at 40 °C for 4 h, and then the residual activity was measured under standard conditions. The relative activity compared with the activity of each nontreated sample was plotted against the incubation pH (Fig. 6). The stability of the mutant enzyme from KNK455M was increased in the lower- and higher-pH regions in comparison with those of the native enzyme.

Substrate Specificities of the DCases

The substrate specificity of the enzyme from KNK455M was investigated using a broad range of *N*-carbamoyl-D-

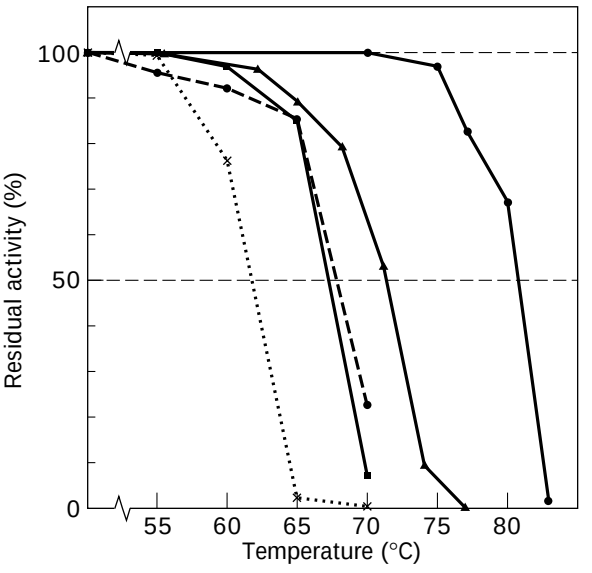


Figure 5. Effect of temperature on the stability of the mutagenized enzymes. A cell-free extract of each mutant DCase-producing strain was kept at various temperatures and the remaining activity was assayed under standard conditions. The remaining activities of the mutant enzymes from KNK402M (*solid line, squares*), KNK404M (*dotted line, circles*), KNK416M (*solid line, triangles*), and KNK455M (*solid line, circles*), and the native enzyme (*asterisks*), are expressed as a percentage of the activity of each nonheated enzyme.

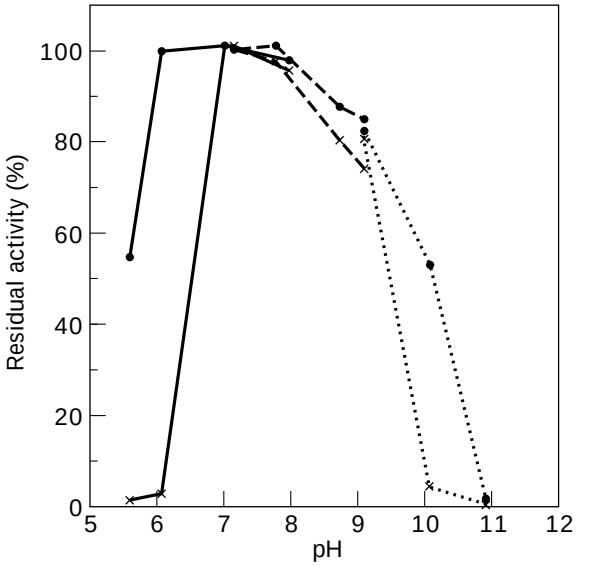


Figure 6. Effect of pH on the enzyme stability of the mutagenized DCases. The mutant enzyme from KNK455M was incubated at 40 °C for 4 h at the indicated pH in 0.1 M K_2HPO_4-KH_2PO_4 (*solid line*), 0.1 M Tris-HCl (*dashed line*), or Na_2CO_3-$NaHCO_3$ (*dotted line*) buffer. The remaining activities were then assayed under standard conditions and expressed as a percentage of the activity of each nontreated enzyme. As a control, the native enzyme (*asterisk*) was treated in the same manner.

Table 6. Comparison of Substrate Specificities of Native and Mutant DCases

Substrate	Relative activity (%)[a]	
	Native DCase	DCase of KNK455M
c-D-HPG	100	100
c-D-Phenylglycine	110	130
c-D-Ala	86	75
c-D-Val	42	47
c-D-Leu	120	110
c-D-Met	180	160
c-DL-Ile	79	76
c-DL-Phe	55	46
c-DL-Ser	58	39
c-DL-Asp	7.2	3.4
c-DL-Norvaline	82	69
c-DL-Norleucine	120	99
c-DL-α-Amino butyrate	100	87

[a]Relative activity is expressed as a percentage of the DCase activity in comparison with the activity with c-D-HPG as a substrate.

Table 7. Comparison of the Properties of Native and Mutant DCases

	Native DCase	DCase of KNK455M
Specific activity (U/mg)	6.8	6.6
Optimum pH	7.0	6.4
pH stability	6.5–7.5	6.0–8.0
Optimum temperature (°C)	65	75
Thermostable temperature (°C)	61.8	80.8
Substrate specificity	c-D-amino acid	c-D-amino acid
K_m (mM, pH 7.0 c-D-HPG)	0.89	1.3
V_{max} (μmol/min/mg, c-D-HPG)	9.6	7.6

amino acids and related compounds. Like the native enzyme, the enzyme showed strict specificity toward *N*-carbamoyl-D-amino acids, *N*-carbamoyl-L-amino acids not being hydrolyzed. *N*-Carbamoyl-D-amino acids having a hydrophobic side chain, as well as *N*-carbamoyl-amino acids having a polar group or short-chain alkyl group, were hydrolyzed. The substrate specificity of the mutant enzyme was almost the same as that of the native enzyme (Table 6). It seemed that these three thermostability-related amino acid changes little affected the substrate specificity. The other known DCases from *Agrobacterium* showed relatively similar tendencies as to substrate specificity.

Kinetic Properties of the Enzymes

The kinetic properties of the enzymes from *Agrobacterium* sp. KNK712 and mutant KNK455M were measured using c-D-HPG as a substrate. The K_m and V_{max} values obtained on Lineweaver–Burk plotting are shown in Table 7.

As the affinity toward substrates and the reactivity were slightly decreased by the mutations, the mutant enzyme was thought to be utilized under industrial conditions.

PRODUCTION OF AN IMPROVED ENZYME

Construction of an Expression Vector

To express the DCase gene effectively, the mutant DCase gene was inserted into a newly constructed vehicle, pUCNT. This vehicle was constructed as follows (21).

To insert the improved DCase gene just after the lac promoter of pUC19, a *Nde*I cleavage site was generated in the initiation codon of the *lacZ* gene, through PCR amplification of the *Hin*dIII–*Cfr*10I 1.3-kb fragment. The amplified fragment was used to replace the corresponding fragment of pUC19, and the plasmid constructed was designated as pUCNde. After generating a *Nde*I site instead of the *Nco*I site of pTrc99A in the same manner, a 0.6-kb *Nde*I–*Ssp*I fragment was ligated into the 2.0-kb *Nde*I-*Ssp*I fragment, pUCNT being constructed.

The 455M mutant DCase gene was inserted into pUCNT between the *Nde*I and *Pst*I cleavage sites, the recombinant plasmid pNT4553 being constructed (Fig. 7).

Expression of the Thermostabilized Enzyme

The expression plasmid was transformed into *E. coli* HB101, and then the recombinant clone was cultivated overnight in 2 × YT medium at 37 °C. In the recombinant *E. coli*, the plasmid was stably maintained without the addition of ampicillin and was expressed efficiently without IPTG, which revealed that the DCase accounted for nearly 50% of all the soluble protein in a cell.

PREPARATION OF AN IMMOBILIZED ENZYME

Immobilization of the Improved DCase

Escherichia coli JM109 (pAD108) and JM109 (pAD404) cells were harvested by centrifugation, suspended in 50 mL of 0.1 M KP buffer, pH 7.0, containing 5 mM DTT, and then disrupted by sonication. The cell debris was removed and the supernatant was obtained, as was the enzyme solution. Duolite A-568 (Rohm and Haas) was washed with 1 M NaCl, water, and then 0.1 M KP buffer, pH 7.0, and then equilibrated with the same buffer for 18 h at room temperature. The wet resin was subjected to filtration, and then DTT (final concentration of 5 mM) was added to the enzyme solution, followed by stirring at 4 °C for 20 h under nitrogen sealing. For adsorption, the weight of the wet resin and that of protein in the enzyme solution were in the ratio of 1 to 0.04. After adsorption, the resin was washed three times with a fivefold volume of 0.1 M KP buffer, pH 7.0, containing 10 mM DTT, cross-linked with a fivefold volume of 0.2% (w/w) of glutaraldehyde at 4 °C for 10 min, washed three times with a fivefold volume of the same buffer at 4 °C, and then filtered (17,18).

Stability of the Immobilized Enzyme

For repeated batch reactions with the immobilized DCase, a substrate solution (3%) of c-D-HPG (pH 7.0) was prepared, with nitrogen gas bubbling at 40 °C. The DCase immobilized on Duolite A-568 and the prepared substrate

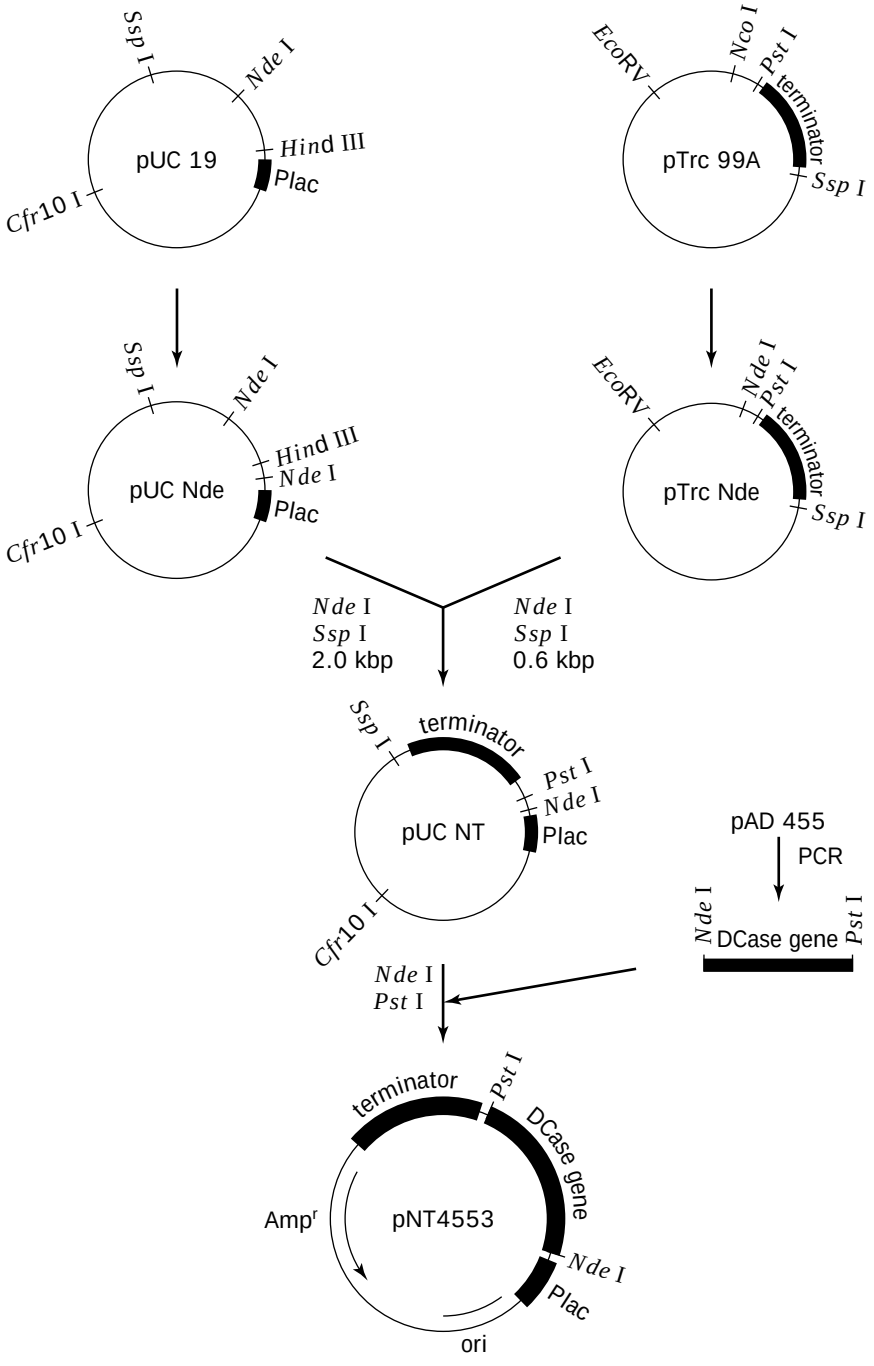


Figure 7. Construction profile of pNT4553.

solution were stirred at 40 °C under a stream of nitrogen gas while the pH was controlled at pH 7.0 with 2 M HCl. Samples were taken for activity measurement at 10 and 60 min, the reaction being continued for 23.5 h in total. The reaction mixture was removed by suction, after which a fresh substrate solution was introduced and allowed to react in the same manner as already described. The residual activity was measured as the relative activity compared with the initial activity and plotted against the reaction times (Fig. 8). The stability of the mutant enzyme increased in comparison with that of the native enzyme.

UTILIZATION OF THE IMPROVED DCASE FOR INDUSTRIAL PRODUCTION

An enzymatic conversion process involving the improved DCase in an immobilized form was introduced for D-HPG production instead of chemical diazonation. As a result of the introduction of the new process, the efficiency of the

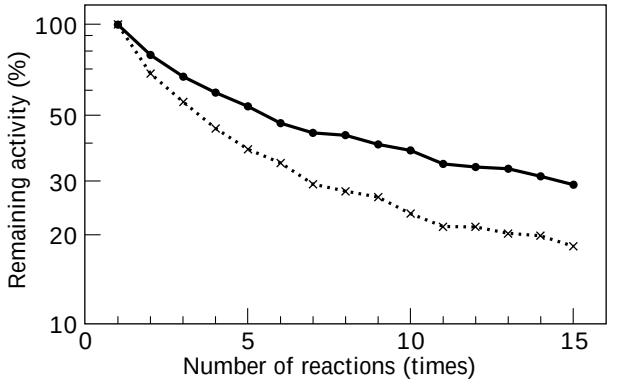


Figure 8. Stability of the immobilized DCase on repeated batch reactions. Repeated batch reactions were carried out using the immobilized mutant DCase from KNK404M (*circles*) and the native enzyme (*asterisks*).

reaction was greatly increased, the purification process became simpler because of the decrease in the undesirable color formation of the product, and the energy required for production was reduced. In addition to these advantages, by-products such as inorganic salts were greatly reduced by the change in the reaction mechanism. By means of these advantages, the new process decreases the burden on the environment.

BIBLIOGRAPHY

1. J. Caravaca and S. Grisolia, *J. Biol. Chem.* **231**, 357–365 (1958).
2. L.L. Campbell, *J. Biol. Chem.* **235**, 2375–2378 (1960).
3. J.M. Kim, S. Shimizu, and H. Yamada, *J. Biol. Chem.* **261**, 11832–11839 (1986).
4. R. Olivieri, E. Fascetti, L. Angelini, and L. Degan, *Biotechnol. Bioeng.* **23**, 2173–2183 (1981).
5. K. Yokozeki and K. Kubota, *Agric. Biol. Chem.* **51**, 721–728 (1987).
6. J. Ogawa, S. Shimizu, and H. Yamada, *Eur. J. Biochem.* **212**, 685–691 (1993).
7. J. Ogawa, M.C.M. Chung, S. Hida, H. Yamada, and S. Shimizu, *J. Biotechnol.* **38**, 11–19 (1994).
8. A. Louwrier and C.J. Knowles, *Enzyme Microb. Technol.* **19**, 562–571 (1996).
9. R. Grifantini, C. Pratesi, G. Galli, and G. Grandi, *J. Biol. Chem.* **271**, 9326–9331 (1996).
10. A. Buson, A. Negro, L. Grassato, M. Tagliaro, M. Basaglia, C. Grandi, A. Fontana, and M.P. Nuti, "*FEMS Microbiol. Lett.* **145**, 55–62 (1996).
11. World Pat. WO94/00577 (January 6, 1994), R.J. Neal, A.M. Griffin, M.O. Scott, A.R. Shatzman, and H.C. Gorham (to Smith-Kline Beecham P.L.C.).
12. H. Nanba, Y. Ikenaka, Y. Yamada, K. Yajima, M. Takano, and S. Takahashi, *Biosci. Biotechnol. Biochem.* **62**, 875–881 (1998).
13. Y. Ikenaka, H. Nanba, Y. Yamada, K. Yajima, M. Takano, and S. Takahashi, *Biosci. Biotechnol. Biochem.* **62**, 882–886 (1998).
14. Y. Ikenaka, H. Nanba, K. Yajima, Y. Yamada, M. Takano, and S. Takahashi, *Biosci. Biotechnol. Biochem.* **62**, 1668–1671 (1998).
15. Y. Ikenaka, H. Nanba, K. Yajima, Y. Yamada, M. Takano, and S. Takahashi, *Biosci. Biotechnol. Biochem.* **62**, 1672–1675 (1998).
16. Y. Ikenaka, H. Nanba, K. Yajima, Y. Yamada, M. Takano, and S. Takahashi, *Biosci. Biotechnol. Biochem.* **63**, in press.
17. H. Nanba, Y. Ikenaka, Y. Yamada, K. Yajima, M. Takano, K. Ohkubo, Y. Hiraishi, K. Yamada, and S. Takahashi, *J. Mol. Catal. B: Enzymatic*, in press.
18. H. Nanba, Y. Ikenaka, Y. Yamada, K. Yajima, M. Takano, and S. Takahashi, *Biosci. Biotechnol. Biochem.* **62**, 1839–1844 (1998).
19. R.P. Ambler, A.D. Auffret, and P.H. Clarke, *FEBS Lett.* **215**, 285–290 (1987).
20. W. Ito, H. Ishiguro, and Y. Kurosawa, *Gene* **102**, 67–70 (1991).
21. H. Nanba, Y. Ikenaka, Y. Yamada, H. Yajima, M. Takano, and S. Takahashi, *J. Ferment. Bioeng.*, in press.

See also AMINO ACIDS, PRODUCTION PROCESSES.

D-AMINOPEPTIDASE, ALKALINE D-PEPTIDASE

YASUHISA ASANO
HIDENOBU KOMEDA
Toyama Prefectural University
Toyama, Japan

KEY WORDS

Alkaline D-peptidase
D-Amino acid–containing peptide
D-Aminopeptidase
Bacillus cereus
Ochrobactrum anthropi
Organic solvent
Penicillin-recognizing enzyme

OUTLINE

INTRODUCTION

Enzymatic reaction has increasingly been applied to organic synthesis (1–3), and there are several reports of the attempted in vitro synthesis of the D-amino acid–containing peptides and amides (4–10). However, because these enzymatic reactions are not stereospecific for substrates with D-configurations, they are at an innate disadvantage. Some peptidases and amidases act on peptides and amides containing D-amino acids. Soluble *Streptomyces* carboxypeptidase DD catalyzes not only the transpeptidation reaction on the peptide intermediate in peptidoglycan biosynthesis but also the hydrolysis of N_{α},N_{ϵ}-diacetyl-L-lysyl-(D-Ala)$_2$ in water (11). A D-peptidase has been purified and characterized from an actinomycete, although it is not strictly specific toward peptides containing D-amino acids (12). In *Enterococcus*, the *vanX* gene product (D-Ala)$_2$ hydrolase plays a role in vancomycin resistance (13). The chemically synthesized D-enzyme of an HIV-1, in which all the amino acids were replaced with the corresponding D-amino acids, displays D-stereospecificity (14). D-Stereospecific amino acid amidases (amidohydrolases) were isolated from some microorganisms (15–17). Of the enzymes, a D-alaninamide–specific amidohydrolase from *Arthrobacter* has been used in the industrial manufacture of D- and L-alanine (17,18). Production of D-amino acids by use of D- or L-specific enzymes such as D-amino acylase, D-hydantoin hydrolase, *N*-carbamoyl-D-amino acid hydrolase, D-amidase, D-transaminase, and amino acid racemase was recently reviewed (19,20).

A new enzyme, D-aminopeptidase, was isolated from *Ochrobactrum anthropi* and characterized. We describe its

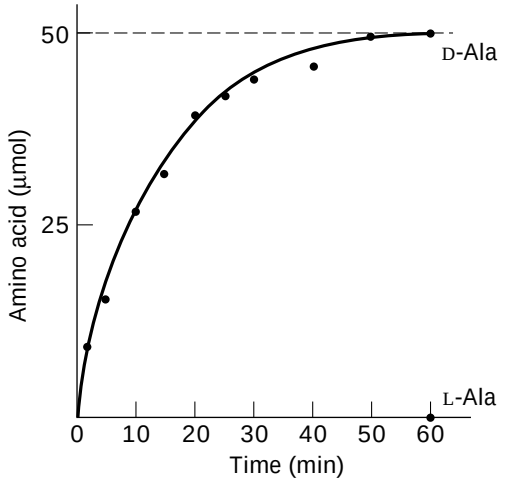


Figure 1. Kinetic resolution of DL-alanine amide by *O. anthropi* D-aminopeptidase. A reaction mixture containing 0.5 mmol tris(hydrochloride), pH 8.0, 0.1 mmol DL-alanine amide hydrochloride, and 5 units of the purified enzyme in a total volume of 5.0 mL was incubated at 30 °C. The content of L-alanine was measured by L-alanine dehydrogenase.

Table 1. Substrate Specificity of D-Aminopeptidase from *O. anthropi* SCRC C1-38

Substrate	Relative activity (%)	K_m (mM)	V_{max} (U/mg)
D-Alanine amide	100	0.65	600
Glycine amide	44	22.3	365
D-α-Aminobutyric acid amide	30	18.3	576
D-Serine amide	29	27.0	22.0
D-Threonine amide	9	100	60.3
D-Methionine amide	2		
D-Norvaline amide	1.8		
D-Norleucine amide	0.8		
D-Phenylglycine amide	0.7		
D-Alanylglycine	95	0.98	1,000
D-Alanylglycylglycine	45	0.37	799
D-Alanyl-D-alanine	21	10.2	326
D-Alanyl-D-alanyl-D-alanine	92	0.57	866
D-Alanyl-D-alanyl-D-alanyl-D-alanine	89	0.32	702
D-Alanyl-L-alanine	46	1.03	312
D-Alanyl-L-alanyl-L-alanine	100	0.65	730
DL-Alanyl-DL-serine	27		
DL-Alanyl-DL-methionine	20		
DL-Alanyl-DL-phenylalanine	9		
DL-Alanyl-DL-asparagine	7		
DL-Alanyl-DL-leucine	1		
DL-Alanyl-DL-valine	0.5		
Glycine methyl ester	229		
D-Alanine methyl ester	75		
D-Alanine-β-naphthylamide	32		
D-Alanine benzylamide	72	0.51	768
D-Alanine anilide	73		
D-Alanine-*p*-nitroanilide	96	0.51	696
D-Alanine *n*-butylamide	66	0.73	670
D-Alanine-3-aminopentane amide	32	2.27	288
D-Alanine *n*-laurylamide	19		
D-Threonine benzyl ester	3.2		

Table 2. D-Aminopeptidase-Catalyzed Synthesis of D-Amino Acid Derivatives in Organic Solvents

Acyl donor	Acyl acceptor	Yield (%)
In 1,1,1-trichloroethane		
D-Alanine methyl ester HCl	3-Aminopentane	99
D-Alanine methyl ester HCl	Neopentylamine	98
D-Alanine methyl ester HCl	*n*-Butylamine	66
D-Alanine methyl ester HCl	Benzylamine	15
In butyl acetate		
D-Alanine methyl ester HCl	3-Aminopentane	99
Glycine methyl ester HCl	3-Aminopentane	80
D-α-Aminobutyric acid methyl ester HCl	3-Aminopentane	44
In toluene		
D-Alanine methyl ester HCl	D-Alanine methyl ester HCl	58[a]

[a]Product: D-alanine dimer (plus D-alanine trimer, yield 6%).

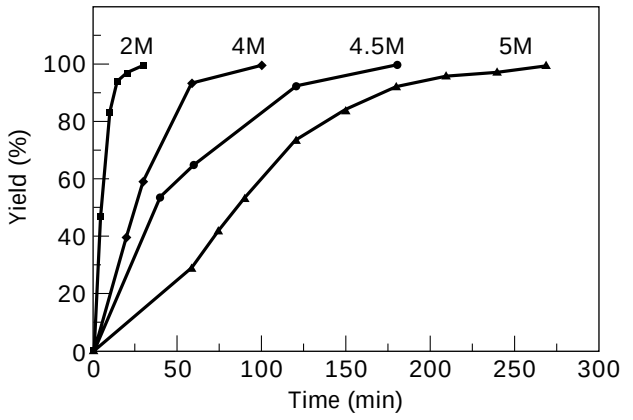


Figure 2. Synthesis of D-alanine from alanine amide HCl by the cells of *E. coli* JM109/pC138DP. The reaction mixture contained various amounts of racemic alanine amide HCl, washed *E. coli* cells harvested from 2.0 mL of culture broth, and 50 μmol of tris(HCl), pH 8.0, in a total volume of 2.0 mL. The enzyme activity shown by the *E. coli* transformant was 288,000 Units per liter of culture. The mixtures were incubated at 30 °C. Formation of D-alanine from D-alanine amide is shown as yield (%).

function, structure, and application to the D-stereospecific hydrolysis of racemic amino acid amides and the formation of peptide bonds. A new enzyme, alkaline D-peptidase, acting on a synthetic peptide, $(D\text{-Phe})_4$, was also isolated from *Bacillus cereus*. We describe its structure and function, as well as its application to the synthesis of D-phenylalanine oligomers. We propose that these two enzymes are new members of the group of penicillin-recognizing enzymes.

D-AMINOPEPTIDASE

During the course of studies on the enzyme-catalyzed organic synthesis of D-amino acid derivatives, a D-stereospecific aminopeptidase was needed. It was speculated that microorganisms might be a likely source for such an enzyme, given their important role in the environment as synthesizers and degraders of a wide variety of substances.

Isolation and Properties of the D-Aminopeptidase

An enrichment culture in a medium containing a synthetic substrate ($D\text{-Ala-NH}_2$) as the sole nitrogen source led to isolation of a bacterial strain, *Ochrobactrum anthropi* SCRC C1-38. The enzyme hydrolyzing $D\text{-Ala-NH}_2$, named D-aminopeptidase (EC 3.4.11.19), was purified to homogeneity from the cell-free extract of the strain and characterized in detail (21). The molecular weight of the native enzyme was estimated to be approximately 122,000, with two identical subunits of molecular weight of about 59,000.

A typical time course for the stereoselective hydrolysis of D-alanine amide is shown in Figure 1, with racemic alanine amide as a substrate. D-Alanine amide was completely hydrolyzed by the action of the enzyme, whereas the L-enantiomer remained unhydrolyzed in the reaction mixture. The rate of the hydrolysis of L-alanine amide was less than 0.01% that of D-alanine amide. The enzyme showed strict chemo- and stereospecificities toward D-amino acid amides, peptides, and esters, as shown in Table 1. Each substrate (100 mM) was incubated under the standard enzyme assay condition (21). The substrates include D-alanine amide (relative velocity: 100%, K_m value: 0.65 mM); glycine amide (44%, 22.3 mM); D-α-aminobutyramide (30%, 18.3 mM); D-serine amide (29%, 27.0 mM); D-alanine 3-aminopentane amide (32%, 2.27 mM); D-alanine-*p*-nitroanilide (96%, 0.51 mM); D-alanine methyl ester (75%), dimer (21%, 10.2 mM), trimer (92%, 0.57 mM), and tetramer (89%, 0.32 mM) of D-alanine; D-alanylglycine (95%, 0.98 mM), D-alanylglycylglycine (45%, 0.37 mM); and D-alanyl-L-alanyl-L-alanine (100%, 0.65 mM). These results show that the enzyme has higher affinity toward peptide substrates than amino acid amides. The enzyme showed neither endopeptidase nor carboxypeptidase activity. The mode of action of the enzyme toward a peptide substrate was studied with D-alanylglycylglycine. The reaction was followed over time, and it was observed that alanine and glycylglycine were released until the substrate tripeptide was nearly completely consumed, whereupon glycine release began, and finally alanine and glycine were produced. This result shows that the enzyme catalyzes the hydrolysis of a single amino acid from the N-terminus of a peptide. Because the mode of action is typical of an aminopeptidase, we named the enzyme D-aminopeptidase.

Application to Organic Synthesis

The D-aminopeptidase from *O. anthropi* SCRC C1-38 was utilized for a stereoselective synthesis of D-alanine *N*-alkylamide from an amine and D-amino acid amide or D-amino acid methyl ester (22). In water, D-alanine *N*-alkylamide once formed by the enzyme was successively hydrolyzed to yield D-alanine. The enzyme was immobilized by urethane prepolymer PU-6. When the aminolysis reaction was performed in water-saturated organic solvents such as butylacetate, benzene, and 1,1,1-trichloroethane with the immobilized enzyme as a catalyst, it progressed well quantitatively in a highly D-stereoselective manner, to give optically pure D-alanine *N*-alkylamides. The acyl donor specificity of this reaction was rather limited, with D-alanine, glycine, and D-α-amino butyric acid as substrates, while the acyl acceptor specificity was relatively wide, with bulky, aromatic-containing, and straight-chain amines (Table 2). The catalytic center activity (k_{cat}) of this reaction, 7,700 (min^{-1}), was much higher than in the case of the nonstereoselective synthesis of D-amino acid–containing peptides (4,5).

Alkaline conditions are usually essential for peptide bond formation in kinetically controlled systems (23,24). Several attempts to oligomerize D-alanine methyl ester with D-aminopeptidase in an alkaline aqueous medium were unsuccessful, probably because the substrate is unstable under alkaline conditions. The immobilized D-aminopeptidase catalyzed the synthesis of D-alanine oligomers from D-alanine methylester in organic solvents (25). Dimer and trimer of D-alanine were obtained in 58% and 6% yield, respectively, when urethane prepolymer PU-6 immobilized D-aminopeptidase (1.5 U/mL) was incubated

```
dd      1   MVSGTVGRGTALGAVLLALLAVPAQAGTAAAADLPAPDDTGLQAVLHTALSQGAPGAMVRVDDNGTIHQL
dap     1                                                   MSKFDTSALEAFVRHIPQNYKGPGGVVAVVKDGEVVLQH
bla     1                                 MFKTTLCALLITASCSTFAAPQQINDIVHRTITPLIEQQKIPGMAVAVIYQGKPYY-F

dd     71   SEGVADRATGRAITTTDRFRVGSVTKSFSAVVLLQLVDEGKLDLDASVNTYLPGLLPDDRITVRQVMSHR
dap    40   AWGFADLRTRTPMTLDTRMPICSVSKQFTCAVLLDAVGEPEL-LDDALEAYL-DKFEDERPAVRDLCNNQ
bla    58   TWGYADIAKKQPVTQQTLFELGSVSKTFTGVLGGDAIARGEIKLSDPTTKYW-PELTAKQWNGITL----

dd    141   SGLYDYTNDMFAQTVPGFESVRNKVFSYQDLITLSLKHGVTNAPGAAYSYSNTNFVVAGMLIEKLTGHSV
dap   107   SGLRDY----WALSVLCGADPEGVFLPAQAQSLLRRLKTTHFEPGSHYSYCNGNFRILADLIEAHTGRTL
bla   123   LHLATY-----TAGGLPLQVPDEVKSSSDLLRFYQNWQPA-WAPGTQRLYANSSIGLFGALAVKPSGLSF

dd    211   ATEYQNRIFTPLNLTDTFYVHPDTVIPGTHANGYLTPDEAGGALVDSTEQTVSW-AQSAGA-VISSTQDL
dap   174   VDILSERIFAPAGMKRAELI-SDTALFD-ECTGY--EGDTVRGFLPATNR-IQWMGDAGICASLNDMIAW
bla   187   EQAMQTRVFQPLKLNHTWINVPPAEEKN-YAWGY--REGKAVHVSPGALDAEAYGVKSTIEDMARWVQSN

dd    279   DTFFSALMSGQLMSAAQLAQMQQWTTVNSTQGYGLGLRRRDLSCGISVYGHTGTVQGYYTYAFASKDGKR
dap   239   EQFIDATRDDESGLYRRLSGPQTFKD-GVAAPYGFGLNLHETG-GKRLTGHGGALRGWRCQRWHCADERL
bla   254   LKPLDINEKTLQQGIQLAQSRYWQTGDMYQGLGWEMLDWPVNPDSIINGSDNKIALAARPVKAITPPTRA

dd    349   SVTALANTSNNVNVLNTMARTLESAFCGKPTTAKLRSATSSATTVERHEDIAPGIARD
dap   307   STIAMFNFEGGASEVAFKLMNIALGVSSSEVSRVEADSAWFGSWLDDETGLVLSLEDAGHGRMKARFGTS
bla   324   VRASWVHKTGATGGFGSYVAFIPEKELGIVMLANKNYPNPARVDAAWQILNALQ

dap   377   PEMMDVVSANEARSAVTTIRRDGETIELVRASENLRLSMKRVKGEAKHDIIGRYHSDELDADLLLVSEGG
dap   447   AIYGAFEGFLGKSDMYPLYSVGSDVWLLPVQRSMDAPSPGEWKLVFRRDDKGEITGLSVGCWLARGVEYR
dap   517   RVQP
```

Figure 3. Comparison of the amino acid sequences of the *dap* gene from *O. anthropi* SCRC C1-38, carboxypeptidase DD from *Streptomyces* R61, and class C β-lactamase from *E. coli*: asterisk (*), identical residue; single dot (·) similarity of functional group.The first amino acid sequence (dd) is the sequence of carboxypeptidase DD from *Streptomyces* R61 (sequence size: 406) (30). The second amino acid sequence (dap) is that of D-aminopeptidase from *O. anthropi* (sequence size: 521) (28). The third amino acid sequence (bla) is that of class C β-lactamase from *E. coli* K12 (sequence size: 377) (32).

in water-saturated toluene with 250 mM of D-alanine methyl ester HCl and 750 mM of triethylamine (Table 2). The k_{cat} of the reaction was calculated to be 19,500 (min^{-1}), which is several tens of thousands times greater than that of the known enzymatic syntheses of amino acid oligomers (26,27).

The gene for the D-aminopeptidase was cloned in *Escherichia coli* JM109 to overproduce the enzyme (28). An expression plasmid pC138DP was constructed. The amount of the enzyme in a cell-free extract of *E. coli* JM109/pC138DP was elevated to 288,000 units per liter of culture, which is about 3,600-fold over that of the wild strain (29). The intact cells of *E. coli* transformant were used as a catalyst for the D-stereospecific hydrolysis of several racemic amino acid amides HCl. Complete hydrolysis of D-alanine amide was achieved in a short time (4.5 h) from 5.0 M of racemic alanine amide HCl using the cells of *E. coli* transformant (Fig. 2). The concentration of D-alanine reached up to 220 g/L. The cells or the cell-free extract catalyzed the synthesis of D-α-aminobutyric acid, D-methionine, D-norvaline, and D-norleucine from their amides in a similar manner.

			Reference
DAP	*O. anthropi*	D L R T R T P M T L D T R M P I C **S** V S **K** Q F T C A V L L D A V G E P E L L	28
BLA Class A	*E. coli* RTEM	S G K I L E S F R P E E R F P M M **S** T F **K** V L L C G A V L S R V D A G Q E Q	33
	B. licheniformis	G T N R T V A Y R P D E R F A F A **S** T I **K** A L T V G V L L Q Q K S I E D L N	34
Class C	*C. freundii*	D I A N N H P V T Q Q T L F E L G **S** V S **K** T F N G V L G G D R I A R G E I K	31
	E. coli K12	L I A K K Q P V T Q Q T L F E L G **S** V S **K** T F T G V L G G D A I A R G E I K	32
	E. cloacae	D I A A N K P V T P Q T L F E L G **S** I S **K** T F T G V L G G D A I A R G E I S	52
	S. marcescens	Q T G K P I T E Q T L F E V G **S** L S **K**	51
Class D	*Staphylococcus*	A M L V F D P V R S K K R Y S P A **S** T F **K** I P H T L F A L D A G A V R D E F	41
	S. marcescens	S C A T N D L A R A S K E Y L P A **S** T F **K** I P N A I I G L E T G V I K N E H	42
PBP 1A	*E. coli*	F N Q S K F N R A T Q A L R Q V G **S** N I **K** P F L Y T A A M D K G L T L A S M	36
1B		P Q F A G Y N R A M Q A R R S I G **S** L A **K** P A T Y L T A L S Q P K I Y R L N	36
2		K S G K E V K F N S D K R F A Y A **S** T S **K** A I N S A I L L E Q V P Y N K L N	37
3		P K E A M R N R T I T D V F E P G **S** T V **K** P M V V M T A L Q R G V V R E N S	35
5		S G K V L A E Q N A D V R R D P A **S** L T **K** M M T S Y V I G Q A M K A G K F K	36
6		S G K V L A E G N A D E K L D P A **S** L T **K** I M T S Y V V G Q A L K A D K I K	36
PBP	*S. aureus*	D K K E P L L N K F Q I T T S P G **S** T Q **K** I L T A M I G L N N K T L D D K T	38
PBP 2	*N. gonorrhoeae*	D S E Q R R N R A V T D M I E P G **S** A I **K** P F V I A K A L D A G K T D L N E	40
CPase	*Streptomyces*	D R A T G R A I T T T D R F R V G **S** V T **K** S F S A V V L L Q L V D E G K L D	30

Figure 4. Partial alignments of the amino acid sequences of D-aminopeptidase and active sites of the penicillin-recognizing enzymes: BLA, β-lactamase; Cpase, carboxypeptidase; DAP, D-aminopeptidase; PBP, penicillin-binding protein.

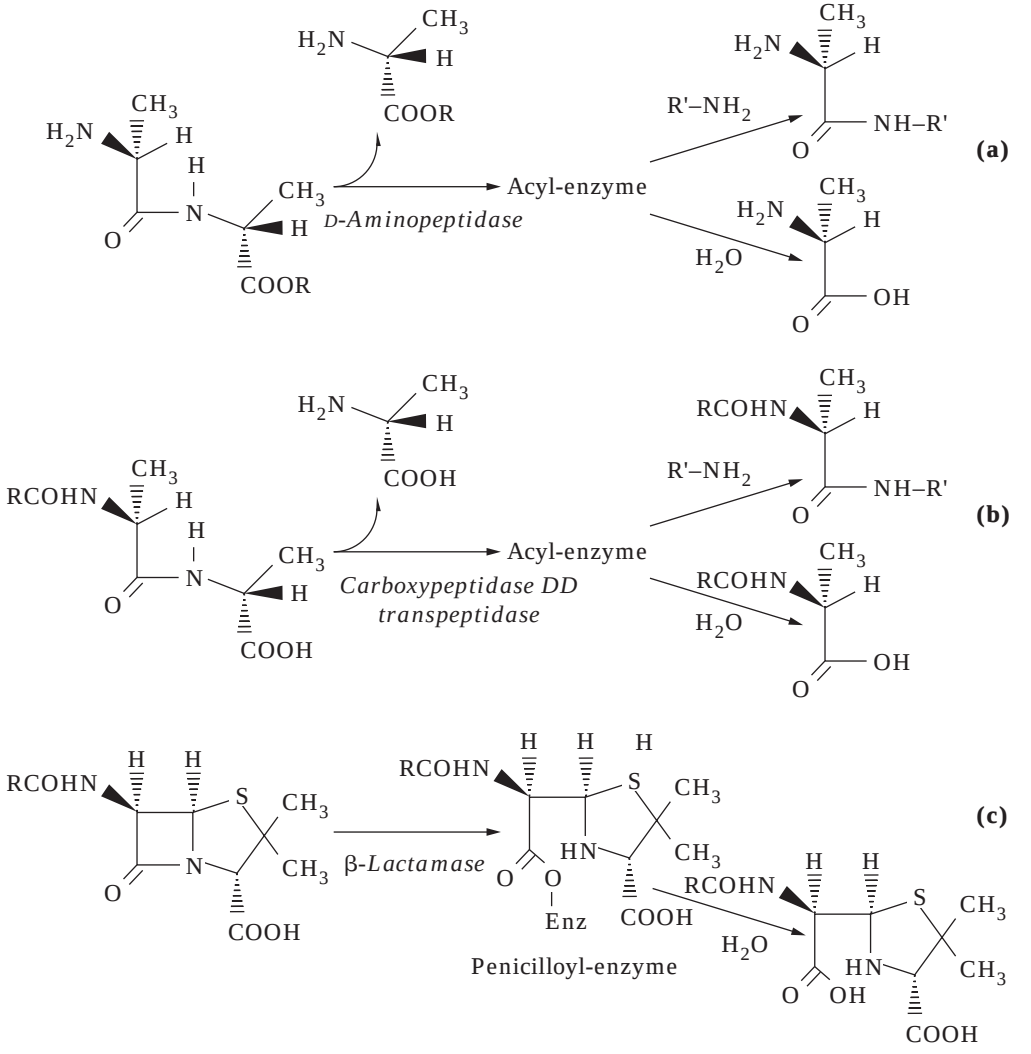


Figure 5. Reactions catalyzed by (**a**) D-aminopeptidase, (**b**) carboxypeptidase DD, and (**c**) β-lactamase.

Table 3. Substrate Specificity of Alkaline D-Peptidase from *Bacillus cereus* DF4-B

Substrate[a]	Relative activity (%)[b]	K_m (mM)	V_{max} (U/mg)	V_{max}/K_m (U/mg/mM)
(D-Phe)$_6$	1.8[c]			
(D-Phe)$_4$	100[c]	0.398	199	500
(D-Phe)$_3$	90[c]	0.127	130	1020
(D-Phe)$_2$	0.2[d]	50.1	13.7	0.270
D-Phe-L-Phe	<0.1[d]			
(D-Phe)$_2$-L-Phe	14.9[c]	0.522	30.6	59.0
L-Phe-(D-Phe)$_2$	119[e]	0.455	154	346
L-Phe-D-Phe-L-Phe	28.1[e]	1.63	66.0	41.0
D-Tyr(D-Phe)$_2$	83.6[d]			
(D-Phe)$_2$-D-Tyr	83.6[c]			
D-Phe-OMe	15[c] 1.8[d]			
D-Phe-NH$_2$	0.1[c] 0.1[d]			
D-Phe-*p*-Nitroanilide	4.2[d]			
Boc-(D-Phe)$_4$	1.8[f] 0.8[g]			
Boc-(D-Phe)$_3$	3.2[f] 1.1[g]			
Boc-(D-Phe)$_2$	7.0[f]			
Boc-(D-Phe)$_3$-O^{t}Bu	1.2[f] 0.3[g]			
Boc-(D-Phe)$_4$-OMe	0.5[f] 0.2[g]			
Boc-(D-Phe)$_3$-OMe	0.7[f] 0.3[g]			
Boc-(D-Phe)$_2$-OMe	1.4[f]			
Ampicillin	8.9[h]	73.1	262	3.58
Penicillin G	9.7[h]	48.9	250	5.11

[a]O^{t}Bu, *tert*-butyl ester; OMe, methyl ester.
[b]Formation of five D-Phe compounds and consumption of a β-lactam compound were measured.
[c]Formation of (D-Phe)$_2$.
[d]Formation of D-Phe.
[e]Formation of L-Phe-D-Phe.
[f]Formation of Boc-D-Phe.
[g]Formation of Boc-(D-Phe)$_2$.
[h]Consumption of a β-lactam compound.

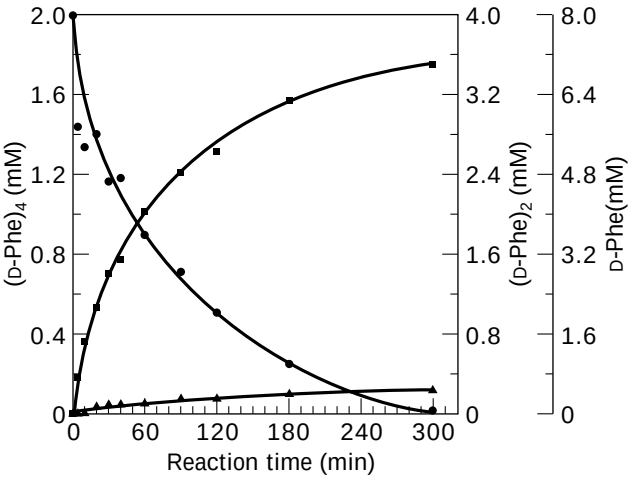


Figure 6. Time course of (D-Phe)$_4$ hydrolysis by D-aminopeptidase. The reaction mixture consisted of 2 mM of (D-Phe)$_4$, 100 mM of tris (HCl), pH 9.0, 2 mM of $MgSO_4$, 2% (v/v) of DMSO, and 0.038 unit of the enzyme solution in a total volume of 500 μL. The reaction was carried out at 30 °C and terminated with $HClO_4$. The amounts of (D-Phe)$_4$ (circles), (D-Phe)$_2$ (squares), and D-Phe (triangles) in the supernatant were measured by a Waters HPLC equipped with a Cosmosil 5C18-MS at a flow rate of 1.0 mL/min with a mixture of methanol (55%) and 5 mmol KH_2PO_4/H_3PO_4, pH 2.9 (45%).

Structural Similarities of D-Aminopeptidase to Carboxypeptidase DD and β-Lactamases

The nucleotide sequence of the D-aminopeptidase gene was determined and analyzed to study the structural relationship to other proteins. The gene consisted of an open reading frame of 1,560 nucleotides, which specifies a protein of M_r of 57,257 (28). The deduced primary structure is considerably similar to that of *Streptomyces* R61 carboxypeptidase DD (30) (27% identical over 287 amino acids) and class C β-lactamase from *E. coli* (22% identical over 234 amino acids) (31,32) (Fig. 3). The enzyme is also structurally related to class A β-lactamases (33,34), penicillin-binding proteins (35–40), and class D β-lactamases (41,42). The sequence Ser-X_{aa}-X_{aa}-Lys is perfectly conserved among this class of enzymes (Fig. 4). With carboxypeptidase DD and class C β-lactamase, the sequences at the highly similar area, Ser-Val-X_{aa}-Lys-X_{aa}-Phe-X_{aa}-Ala-X_{aa}-Val-Leu-Leu (30) and Ser-Val-Ser-Lys-X_{aa}-Phe-Tyr (31), respectively, are well conserved. Furthermore, some of the residues scattered around the conserved region are similar among the enzymes. The Met-57 located four residues upstream from the Ser-61 is found similarly in the carboxypeptidase DD and β-lactamases as a hydrophobic residue Phe. A common observation regarding D-aminopeptidase and the β-lactamase is that the consensus sequence is located around 60 from the N-terminus of the enzymes (43).

The inhibition by *p*-chloromercuribenzoate (PCMB) and other sulfhydryl reagents suggested that the enzyme would be a thiol peptidase (44). However, the findings of the site-specific mutagenesis study showed that the amino acid sequences Ser-X_{aa}-X_{aa}-Lys, which had been conserved in the penicillin-recognizing enzymes, are essential in exerting the D-aminopeptidase activity (28). The sites Ser-61 and Lys-64 are essential in the catalysis, because the V_{max}/K_m value measured with the mutants modified to Cys-61 and Asn-64 have been reduced to 0.26% and 0.008%, respectively, of those of the natural enzyme, although these K_m values were hardly lowered. The mutant with the inert Gly residue in place of the likely active center Ser-61 had a lower activity (10^{-5}-fold that of the native enzyme). On the other hand, the mutations at other sites (Met-57 and Cys-68) did not greatly affect the enzyme activity. The mutations at Cys-60, which is adjacent to the likely active center Ser-61, gave notable effects on the kinetic profiles, indicating that the residue is important in the binding of the substrate. The substitutions at Cys-60 to Ser and Gly produced mutants with slightly altered V_{max}/K_m values. They were tolerant to inhibition by PCMB, which suggests that the inhibition of the native enzyme by PCMB would have been due to the steric hindrance by a mercaptide bond formation between Cys-60 of the enzyme and PCMB.

A mutant of D-aminopeptidase with increased thermostability was obtained. The enhancement of the thermal stability of mutant enzyme was attributed to the substitution of Gly-155 to Ser (45).

Figure 5 shows similarities of the reactions catalyzed by D-aminopeptidase, carboxypeptidase DD, and β-lacta-

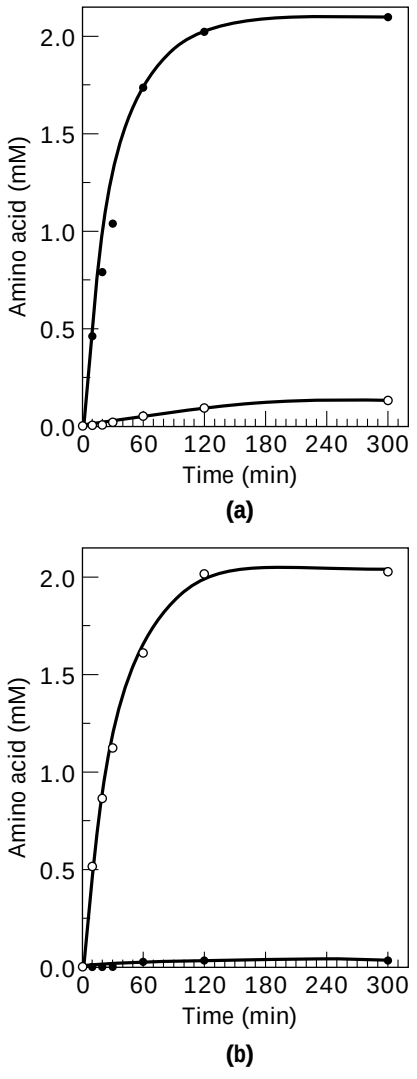


Figure 7. Mode of action toward (**a**) (D-Phe)$_2$-D-Tyr and (**b**) D-Tyr-(D-Phe)$_2$: solid circles, D-Tyr; open circles, D-Phe. The reaction mixture (500 μL) containing 2 mM of the substrate (D-Phe)$_2$-D-Tyr or D-Tyr-(D-Phe)$_2$, 50 mM of tris (HCl), pH 9.0, 2 mM of $MgSO_4$, 2% (v/v) of DMSO, and 0.038 unit of the enzyme was incubated at 30 °C. After termination of the reaction, the reaction mixture was analyzed with a Hitachi L-8500 amino acid analyzer.

mase. All three enzymes catalyze the hydrolytic cleavage of the amide bond between D-amino acids of the same configuration. The former two catalyze the transpeptidation reaction; carboxypeptidase DD catalyzes the cross-linkage of the peptideglycan, and D-aminopeptidase catalyzes the transpeptidation and aminolysis of D-amino acid ester and amides, respectively, in water and organic solvents. The aminolysis reaction catalyzed by D-aminopeptidase in butylacetate is efficient (k_{cat}: 7,700/min) (22). The transpeptidation reaction can also be catalyzed in water, although the product D-alanine-3-aminopentane amide was easily hydrolyzed in water (22). Inheritance of the carboxypeptidase-like tertiary structure in the D-aminopeptidase would have made the transpeptidation reaction in organic solvents possible. The enzyme is actually inhibited by β-lactam compounds, such as 6-APA, 7-ACA, benzylpenicillin, and ampicillin, although none of these is the substrate for the enzyme.

The D-aminopeptidase is a new member of the penicillin-recognizing enzymes by virtue of the following characteristics: similarities in the primary structure by gene sequencing, similarities in the reactions catalyzed in water and organic solvents, and the findings obtained by kinetic studies of the mutants generated by site-directed mutagenesis and the inhibition by β-lactam compounds. The mutants generated by inhibition of β-lactam compounds may constitute evidence to rebut the contention that the β-lactamases evolve from the penicillin-binding proteins (46). The existence of a third enzyme with a similar structure, which does not appear to be a selective target of the β-lactam compounds, was shown. To our knowledge, this enzyme is the first example of an aminopeptidase with Ser at the active site.

ALKALINE D-PEPTIDASE

D-Aminopeptidase (EC 3.4.11.19) from *Ochrobactrum anthropi* was discovered, and its primary structure has been found to have similarity to the β-lactamases and penicillin-binding proteins (21,28). The enzyme acts mostly on peptides with D-Ala at the N-terminus to yield D-amino acids; it does not act on D-amino acid derivatives with bulkier substituents. We proposed that D-aminopeptidase is a new penicillin-recognizing enzyme (11) based on its primary structure, inhibition by β-lactam compounds, and the ability to catalyze peptide bond formation in organic solvents, although the enzyme does not show β-lactamase activity (21,47).

Here, we describe the screening of soil microorganisms for D-stereospecific endopeptidases using a synthetic peptide (D-Phe)$_4$, characterization of the new enzyme alkaline D-peptidase (ADP), as well as cloning and sequencing of the *adp* gene from *B. cereus* strain DF4-B.

Isolation and Properties of the Alkaline D-Peptidase

An enrichment culture in LB medium containing a synthetic substrate (D-Phe)$_4$ led to the isolation of a bacterial strain *Bacillus cereus* DF4-B. The extracellular enzyme hydrolyzing (D-Phe)$_4$ was purified and characterized (48). The M_r of the subunit calculated was about 36,000 by gel electrophoresis (SDS-PAGE), 37,000 by HPLC, and 37,952 by mass spectrophotometry. The absorption of the purified enzyme in 0.01 M potassium phosphate buffer, pH 7.0, was maximal at 281 nm. The enzyme is an endopeptidase that acts D-stereospecifically on peptides composed of aromatic D-amino acids, recognizing the D-configuration of the amino acid whose carboxy-terminal peptide bond is hydrolyzed. The enzyme had an optimum pH at around 10.3. Thus, the enzyme was named alkaline D-peptidase (D-stereospecific peptide hydrolase [EC 3.4.11.–]). The substrate specificity of the enzyme was examined as shown in Table 3. Each substrate (2 mM) was incubated under standard enzyme assay conditions (48). The enzyme was active

```
                    9                  18                  27                  36                  45                  54                  63                  72                  81
5'  GAA TTC AAT GTT GAC TTT TTT TTT TGG TGG AGA TGA TAG CTG TGT GGT TAA TAC GAA ATA GCA AGA AAC CAC CTA TAC AAT

                   90                  99                 108                 117                 126                 135                 144                 153                 162
    ATG AAT AGA GGA GAG AAA TGC CCT ATG AAA ACA CGT AGT CAA ATT ACA TGT GCA AGT CTT GCC CTT TTA ATA GCT GGA AGT
                                    Met Lys Thr Arg Ser Gln Ile Thr Cys Ala Ser Leu Ala Leu Leu Ile Ala GGA AGT

                  171                 180                 189                 198                 207                 216                 225                 234                 243
    TCC CTG TTA TAC GAG ACG CAA ACC TTA ATT GTA AAA GCA GAA CCT ACT CAA AGT GTA TCT AGT TCG GTA CAA ACA AGT ACT
    Ser Leu Leu Tyr Thr Thr Gln Thr Leu Ile Val Lys Ala Glu Pro Thr Gln Ser Val Leu Ser Ser Val Gln Thr Ser Thr

                  252                 261                 270                 279                 288                 297                 306                 315                 324
    CAA CGA GAT CGT AAT TCT GTC AAG CAA GCA GTG CGG GAT ACA TTG CAA CTT GGA TTC CCG GGG ATA CTT GCT AAA ACT TCT
    Gln Arg Asp Arg Asn Ser Val Lys Gln Ala Val Arg Asp Thr Leu Gln Leu Gly Phe Pro Gly Ile Leu Ala Lys Thr Ser

                  333                 342                 351                 360                 369                 378                 387                 396                 405
    GAG GGT GGA AAA ACA TGG AGT TAT GCC GCT GGG GTA GCG AAT CTG AGC AGC AAG AAA CCC ATG AAA ACA GAT TTT CGC TTT
    Glu Gly Gly Lys Thr Trp Ser Tyr Ala Ala Gly Val Ala Asn Leu Ser Ser Lys Lys Pro Met Lys Thr Asp Phe Arg Phe

                  414                 423                 432                 441                 450                 459                 468                 477                 486
    CGC ATT GGT AGC GTG ACG AAG ACG TTC ACG GCA ACA GTT GTA CTT CAa TTA GCC GAA GAG AAC CGC TTG AAT CTA GAC GAC
    Arg Ile Gly Ser Val Thr Lys Thr Phe Thr Ala Thr Val Val Leu Gln Leu Ala Glu GLu Asn Arg Leu Asn Leu Asp Asp

                  495                 504                 513                 522                 531                 540                 549                 558                 567
    TCT ATT GAA AAA TCG TTG CCT GGT GTC ATT CAA GGA AAT GGG TAT GAT GAT AAA CAG ATT ACT ATC CGG CAA TTA TTG AAC
    Ser Ile Glu Lys Trp Leu Pro Gly Val Ile Gln Gly Asn Gly Tyr Asp Asp Lys Gln Ile Thr Ile Arg Gln Leu Leu Asn

                  576                 585                 594                 603                 612                 621                 630                 639                 648
    CAC ACA AGT GGT ATC GCT GAA TAC ACA AGG TCA AAA AGT TTT GAT CTT ATG GAT ACT AAA AAA TCG TAT AGG GCT GAA GAA
    His Thr Ser Gly Ile Ala Glu Tyr Thr Arg Ser Lys Ser Phe Asp Leu Met Asp Thr Lys Lys Ser Tyr Arg Ala Glu Glu

                  657                 666                 675                 684                 693                 702                 711                 720                 729
    TTA GTA AAG ATG GGG ATT TCG ATG CCC CCA GAT TTT GCT CCA GGA AAG AGC TGG TCT TAT TCA AAT ACA GGA TAC GTA TTA
    Leu Val Lys Met Gly Ile Ser Met Pro Pro Asp Phe Ala Pro Gly Lys Ser Trp Ser Tyr Ser Asn Thr Gly Tyr Val Leu

                  738                 747                 756                 765                 774                 783                 792                 801                 810
    CTT GGG ATC CTT ATT GAA ACA GTA ACC GGG AAC AGC TAT GCG GAA GAG ATT GAA AAT CGG ATT ATT GAA CCG CTT GAA TTA
    Leu Gly Ile Lys Ile Glu Thr Val THr Gly Asn Ser Tyr Ala Glu Glu Ile Glu Asn Arg Ile Ile Glu Pro Leu Glu Leu

                  819                 828                 837                 846                 855                 864                 873                 882                 891
    TCG AAT ACA TTC TTA CCT GGC AAT TCA AGC GTG ATT CCA GGA ACC AAG CAT GCC CGT GGA TAT ATC CAA CTA GAC GGA GCA
    Ser Asn Thr Phe Leu Pro Gly Asn Ser Ser Val Ile Pro Gly Thr Lys His Ala Arg Gly Tyr Ile Gln Leu Asp Gly Ala

                  900                 909                 918                 927                 936                 945                 954                 963                 972
    AGT GAG CCA AAA GAT GTT ACT TAT TAT AAC CCA AGT ATG GGG AGC TCG GCT GGA GAT ATG ATT TCT ACT GCT GAT GAT TTA
    Ser Glu Pro Lys Asp Val Thr Tyr Tyr Asn Pro Ser Met Gly Ser Ser Ala Gly Asp Met Ile Ser Thr Ala Asp Asp Leu

                  981                 990                 999                1008                1017                1026                1035                1044                1053
    AAC AAA TTC TTC TCT TAC TTA CTT GGG GGT AAA TTA CTA AAA GAA CAG CAA CTA AAA CAG ATG CTT ACT ACA GTT CCA ACA
    Asn Lys Phe Phe Ser Tyr Leu Leu Gly Gly Lys Leu Leu Lys Glu Gln Gln Leu Lys Gln Met Leu Thr Thr Val Pro Thr

                 1062                1071                1080                1089                1098                1107                1116                1125                1134
    GGA GAA GCT GCA CTT GGC AGA TAT GGT CTT GGA ATC TAT GAA ACT AAG CTT CCA AAC GGT GTC TCA ATA TGG GGA CAC GGA
    Gly Thr Ala Ala Leu Gly Arg Tyr Gly Leu Gly Ile Tyr Glu Thr Lys Leu Pro Asn Gly Val Ser Ile Trp Gly His Gly

                 1143                1152                1161                1170                1179                1188                1197                1206                1215
    GGT AGC ATT CCA GGG TTT GTT ACT TTT GCT GGA GGC ACA CTT GGA GGC AAG CAT ACA TTA GCT GTC AAT TTG AAC AGC CTT
    Gly Ser Ile Pro Gly Phe Val Thr Phe Ala Gly Gly Thr Leu Gly Gly Lys His Thr Leu Ala Val Asn Leu Asn Ser Leu

                 1224                1233                1242                1251                1260                1269                1278                1287                1296
    AAT GCT GAG AGT CCT GAT CCT TTT AAA AAT ATT TTA CTT GCT GAA TTT AGC AAG TAG GCC AAA AGG AAG AAC GGA TAA ATT
    Asn Ala Glu Ser Pro Asp Pro Phe Lys Asn Ile Leu Leu Ala Glu Phe Ser Lys ***

                 1305                1314                1323                1332                1341                1350                1359                1368                1377
    TTA TCA TAT TTT CAT AGT TAA ACA AAT AAA TGA TGA AAT TGA TCA CTT ATC TAA GGA GGC AAT TCC CTC ATG AAA AAT AAA


                 1386                1395                1404                1413                1422                1431                1440                1449                1458
    CTA ACT GGT GTA TTA GCC GCT ACA TTA GCG CTA ACT ATG ATT CTA CCA ACA GGG GCG AAA GCA TTT TCT GAC GGT AAG ACT

                 1467                1476                1485                1494                1503                1512                1521                1530                1539
    ACA GTT GTA TCT AAC GCA GAA GTA GCT AGC CAA GAA CTG AAG AAG ATT GCA GCA GAA AAG GCT GCG CTA CTC ACG AAG TCC

                 1548                1557                1566                1575                1584                1593                1602                1611                1620
    CAT GGG ACG AGT AGC GTG CAG TAT GCG CTA ATT GAT AAT GGC AAG CTT ACT TTG TCG GGG CAG GCT GGA AAG AAC GAT ATG

                 1629                1638                1647                1656                1665                1674                1683                1692                1701
    GAA GGG GAA CAG CCG CTA ACC AAA GAT ACT CTA TAC GGG ATT GGT TCA GTC AGT AAA ATG TAT GCA ACA GCG GCT GTA ATG

                 1710
    AAA TTG GTC GAC 3'
```

Figure 8. Nucleotide sequence of the *adp* gene and the predicted amino acid sequence. In this open reading frame of 1,164 bp (388 amino acids), shown with the deduced amino acid sequence, sequences confirmed by the protein sequencer are underlined.

Bacillus cereus ADP (48)	55	M	K	T	D	F	R	F	R	I	G	**S**	V	T	**K**	T	F	T	A	T	V	V	75
Streptomyces Cpase DD (30)	83	I	T	T	T	D	R	F	R	V	G	**S**	V	T	**K**	S	F	S	A	V	V	L	103
Streptomyces lactamdurans PBP (49)	50	V	T	T	D	S	V	F	Q	S	G	**S**	V	A	**K**	V	Y	T	A	T	L	V	70
Bacillus subtilis PBP 4 (50)	50	L	K	T	N	S	L	F	E	L	A	**S**	L	S	**K**	P	F	T	A	L	G	I	70
Serratia marcescens BLA (51)	69	I	T	E	Q	T	L	F	E	V	G	**S**	L	S	**K**	T	F	T	A	T	L	A	89
Enterobacter cloacae BLA (52)	74	V	T	P	Q	T	L	F	E	L	G	**S**	I	S	**K**	T	F	T	G	V	L	G	94
Dichelobacter nodosus Protein D (53)	103	M	R	V	N	S	R	F	R	Y	A	**S**	V	T	**K**	V	L	T	S	A	L	V	123
Ochrobactrum anthropi DAP (28)	55	M	T	L	D	T	R	M	P	I	C	**S**	V	S	**K**	Q	F	T	C	A	V	L	75
Pseudomonas sp. Esterase (31)	65	W	H	S	D	T	I	V	N	L	F	**S**	C	T	**K**	T	F	T	A	V	T	A	85

Figure 9. Partial alignments of the amino acid sequences of ADP and other similar enzymes. Met[55] found in the partial amino acid sequence of ADP corresponds to [387]ATG in Figure 8; abbreviations as in Figure 4.

Table 4. Comparison of Properties of the D-Amino Acid–Containing Peptide Hydrolases

Property	D-Aminopeptidase	Alkaline D-peptidase
Origin	*Ochrobactrum anthropi*	*Bacillus cereus*
Molecular weight	114,514 (57,252 × 2)	37,952
Number of subunits	2	1
Optimum pH	8.5	10.5–11
Substrates (V_{max}/K_m)		
(D-Ala)$_4$	2,190	
(D-Phe)$_4$		364
Ampicillin	0.04	3.58
Penicillin G	0.05	5.11
Mode of action	D-Stereospecific aminopeptidase	D-Stereospecific endopeptidase
Similarity	Cpase DD (27%)	Cpase DD (35%)
	Class C BLA (22%)	Class C BLA (25%)
	PBP 4 (29%), DAP (27%)	

Note: BLA, β-lactamase; C Base, carboxy peptides; DAP, D-aminopeptidase; PBP, penicillin-binding protein.

toward (D-Phe)$_3$ and (D-Phe)$_4$, forming (D-Phe)$_2$ and D-Phe. The enzyme is also active toward tripeptides with D-Tyr at the C- or N-terminus and on Boc-(D-Phe)*n* (*n* = 2–4), forming Boc-D-Phe, (D-Phe)$_2$, and D-Phe. The enzyme had esterase activity toward D-Phe methyl ester and (D-Phe)$_2$ methyl ester. The products from Boc-(D-Phe)$_3$ *tert*-butyl ester were Boc-D-Phe, D-Phe, and D-Phe *tert*-butyl ester. The enzyme was not active toward L-Phe methyl ester, (L-Phe)$_2$ methyl ester, (L-Phe)$_4$, Boc-(L-Phe)$_4$, Boc-(L-Phe)$_4$ methyl ester, (D-Val)$_3$, (D-Leu)$_2$, and (D-Ala)*n* (*n* = 2–5). These properties indicated that the enzyme is an endopeptidase that acts D-stereospecifically on peptides composed of aromatic D-amino acids. On the other hand, a dimer was formed when D-Phe methyl ester and D-Phe amide were the substrates. Since the enzyme is found to be a serine peptidase as described it is anticipated that later, the enzyme will be useful in the kinetically controlled synthesis of peptides.

Eight stereoisomers of phenylalanine trimer were synthesized, and their effectiveness as substrates for the enzyme was tested. The enzyme recognized the configuration of the second D-Phe of tripeptides and catalyzes the hydrolysis of the second peptide bond from the N-terminus. The calculated V_{max}/K_m values for the peptides containing L-Phe were lower than that for (D-Phe)$_3$, affected by the neighboring L-Phe. The enzyme also showed β-lactamase activity toward ampicillin and penicillin G. The calculated V_{max} values of the enzyme for β-lactam compounds were about the same as those for (D-Phe)$_3$ and (D-Phe)$_4$, whereas the K_m values were several hundred times larger. On the other hand, carboxypeptidase DD (30) and D-aminopeptidase (28) activities were undetectable.

The time course of the (D-Phe)$_4$ degradation was measured. As shown in Figure 6, (D-Phe)$_4$ was hydrolyzed to (D-Phe)$_2$ and D-Phe. No (D-Phe)$_3$ was detected. These results coincide with the kinetic properties of the enzyme described earlier. The mode of action of the enzyme was examined with the synthetic substrates D-Tyr-(D-Phe)$_2$ and (D-Phe)$_2$-D-Tyr, as shown in Figure 7. When D-Tyr-(D-Phe)$_2$ was the substrate, D-Phe was released first, then D-Tyr was slowly formed. When (D-Phe)$_2$-D-Tyr was used as a substrate, D-Tyr was released first, then D-Phe was slowly formed. In both reactions, the second peptide bond from the N-terminus of the substrate was hydrolyzed first. These results showed that the enzyme acts as a D-stereospecific dipeptidylendopeptidase.

The enzyme activity was maximal at 45 °C. About 60% activity remained after an incubation at 43 °C in 0.1 M potassium phosphate buffer, pH 8.0, for 10 min. No activity was lost between pH 5.0 and 10.0 after an incubation at 30 °C for 1 h in 0.05 M buffers at various pH values (48).

The enzyme activity was enhanced by Mg^{2+} (138%), Mo^{3+} (130%), and Ba^{2+} (123%). When measured after incubating at 30 °C for 30 min, the activity was inhibited to 94% by phenylmethylsulfonyl fluoride (PMSF), 76% by Ag^{+}, 74% by Fe^{2+}, and 32% by Hg^{2+} at 5 mM. These results, together with the information about its primary

structure as described later, indicated that the enzyme is a serine peptidase (48).

Cloning of the Gene for the Alkaline D-Peptidase

The gene coding for alkaline D-peptidase (*adp*) was cloned into plasmid pUC118, and a 1,164 bp open reading frame consisting of 388 codons with an M_r of 42,033 was identified as the *adp* gene (Fig. 8) (48). The enzyme would be synthesized with a signal peptide.

The deduced primary structure of the enzyme is similar to carboxypeptidase DD from *Streptomyces* R61 (35.0% identical over 346 amino acids [a.a.]) (14), penicillin-binding proteins from *Streptomyces (Nocardia) lactamdurans* (28.1% over 263 a.a.) (49) and that of *B. subtilis* (28.5% over 309 a.a.) (50), and class C β-lactamases of *Serratia marcescens* (24.9% over 217 a.a.) (51), class C β-lactamases of *Enterobacter cloacae* (25.1% over 191 a.a.) (52), fimbrial protein D from *Dichelobacter nodosus* (24.1% over 261 a.a.) (53), D-aminopeptidase from *O. anthropi* (27.5% over 182 a.a.) (28), and esterase from *Pseudomonas* sp. (30.5% over 154 a.a.) (54). As shown in Figure 9, the sequence of Ser-X_{aa}-X_{aa}-Lys is perfectly conserved among this class of enzymes and the consensus sequence is located around 60 residues from the N-termini of most of the enzymes. Thus, we propose that alkaline D-peptidase from *B. cereus* be categorized as a new member of the penicillin-recognizing enzymes, which include penicillin-binding proteins, β-lactamases, and D-aminopeptidase. Table 4 summarizes the properties of the D-amino acid–containing peptide hydrolases, D-aminopeptidase, and alkaline D-peptidase.

Synthesis of D-Phe Oligopeptides by the Alkaline D-Peptidase (55)

The alkaline D-peptidase not only hydrolyzed $(D\text{-Phe})_3$ and $(D\text{-Phe})_4$ to form $(D\text{-Phe})_2$ and D-phenylalanine but also acted on D-Phe methyl ester and D-Phe amide to form $(D\text{-Phe})_2$. This finding gave us the opportunity to use the enzyme to further investigate the synthesis of D-Phe oligopeptides from D-Phe methyl ester. Because the alkaline D-peptidase was found to be a serine peptidase (48), we attempted to use it for kinetically controlled peptide synthesis. An expression plasmid pKADP was constructed by placing the alkaline D-peptidase gene (*adp*), amplified by means of the polymerase chain reaction, under the *tac* promoter of pKK223-3. Oligomerization of D-phenylalanine methyl ester by means of the purified enzyme from the transformant *E. coli* was investigated under several conditions. D-Phe dimer, $(D\text{-Phe})_2$, and trimer, $(D\text{-Phe})_3$, were produced in 25.4% and 8.6% yield, respectively, when 50 mM of the substrate was incubated for 8 h with ADP (2.0 and 0.4 U/mL, respectively) in 100 mM triethylamine HCl (pH 11.5). Addition of dimethyl sulfoxide to the reaction mixture resulted in the production of tetramer $(D\text{-Phe})_4$ in 6.7% yield with the decrease of the $(D\text{-Phe})_2$ and $(D\text{-Phe})_3$ production. This is the first example of the synthesis of D-phenylalanine oligomers by means of a D-stereospecific endopeptidase.

BIBLIOGRAPHY

1. G.M. Whitesides and C.-H. Wong, *Angew. Chem. Int. Ed. Engl.* **24**, 617–638 (1985).
2. C. Laane, J. Tramper, and M.D. Lilly, eds., *Biocatalysts in Organic Media*, Elsevier, Amsterdam, 1987.
3. H. Yamada and S. Shimizu, *Angew. Chem. Int. Ed. Engl.* **27**, 622–642 (1988).
4. J.B. West and C.-H. Wong, *J. Org. Chem.* **51**, 2728–2735 (1986).
5. A.L. Margolin, D.-F. Tai, and A.M. Klibanov, *J. Am. Chem. Soc.* **109**, 7885–7887 (1987).
6. K. Kato, K. Kawahara, T. Takahashi, and S. Igarashi, *Agric. Biol. Chem.* **44**, 821–825 (1980).
7. I.B. Stoineva and D.D. Petkov, *FEBS Lett.* **183**, 103–106 (1985).
8. H. Nakajima, S. Kitabatake, R. Tsurutani, K. Yamamoto, I. Tomioka, and K. Imahori, *Int. J. Pept. Protein Res.* **28**, 179–185 (1986).
9. A.L. Margolin and A.M. Klibanov, *J. Am. Chem. Soc.* **109**, 3802–3804 (1987).
10. C.F. Barbas, III and C.-H. Wong, *J. Chem. Soc., Chem. Commun.* **1987**, 533–534.
11. J.-M. Frère and B. Joris, *CRC Crit. Rev. Microbiol.* **11**, 299–396 (1985).
12. M. Sugie and H. Suzuki, *Agric. Biol. Chem.* **50**, 1397–1402 (1986).
13. P.E. Reynolds, F. Depardieu, S. Dutka-Malen, M. Arthur, and P. Courvalin, *Mol. Microbiol.* **13**, 1065–1070 (1994). GenBank database accession number M97297.
14. R.C. Del Milton, S.C.F. Milton, and S.B.H. Kent, *Science* **256**, 1445–1448 (1987).
15. Y. Asano, T. Mori, S. Hanamoto, Y. Kato, and A. Nakazawa, *Biochem. Biophys. Res. Commun.* **162**, 470–474 (1989).
16. Japanese Pat. 88-87998 (1988), M. Dotani, H. Igarashi, and U. Uragami.
17. A. Ozaki, H. Kawasaki, M. Yagasaki, and Y. Hashimoto, *Biosci. Biotechnol. Biochem.* **56**, 1980–1984 (1992).
18. A. Ozaki, H. Kawasaki, M. Yagasaki, and Y. Hashimoto, *Biosci. Biotechnol. Biochem.* **57**, 520–521 (1993).
19. J. Ogawa and S. Shimizu, *J. Mol. Catal.* **B2**, 163–176 (1997).
20. M. Yagasaki and A. Ozaki, *J. Mol. Catal.* **B4**, 1–11 (1998).
21. Y. Asano, A. Nakazawa, Y. Kato, and K. Kondo, *J. Biol. Chem.* **264**, 14233–14239 (1989).
22. Y. Kato, Y. Asano, A. Nakazawa, and K. Kondo, *Tetrahedron* **45**, 5743–5754 (1989).
23. I.M. Chaiken, A. Komoriya, M. Ohno, and F. Widmer, *Appl. Biochem. Biotechnol.* **7**, 385–399 (1982).
24. K. Morihara, *Trends Biotechnol.* **5**, 164–170 (1987).
25. Y. Kato, Y. Asano, A. Nakazawa, and K. Kondo, *Biocatalysis* **3**, 207–215 (1990).
26. L. Sluyterman and J. Wijdenes, *Biochim. Biophys. Acta* **289**, 194–202 (1972).
27. K. Aso, T. Uemura, and Y. Shiokawa, *Agric. Biol. Chem.* **52**, 2443–2449 (1988).
28. Y. Asano, Y. Kato, A. Yamada, and K. Kondo, *Biochemistry* **31**, 2316–2328 (1992). GenBank database accession number M84523.
29. Y. Asano, K. Kishino, A. Yamada, S. Hanamoto, and K. Kondo, *Recl. Trav. Chim. Pays-Bas* **110**, 206–208 (1991).

30. C. Duez, C. Piron-Fraipont, B. Joris, J. Dusart, M.S. Urdea, J.A. Martial, J.-M. Frère, and J.-M. Ghuysen, *Eur. J. Biochem.* **162**, 509–518 (1987). GenBank database accession number X05109.
31. F. Lindberg and S. Normark, *Eur. J. Biochem.* **156**, 441–445 (1986). GenBank database accession number X03866.
32. B. Jaurin and T. Grundström, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 4897–4901 (1981).
33. R.P. Ambler and G.K. Scott, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3732–3736 (1978).
34. K. Neugebauer, R. Sprengel, and H. Schaller, *Nucleic Acids Res.* **9**, 2577–2588 (1981). GenBank database accession number V00093.
35. M. Nakamura, I.N. Maruyama, M. Soma, J. Kato, H. Suzuki, and Y. Horota, *Mol. Gen. Genet.* **191**, 1–9 (1983). GenBank database accession number K00137.
36. J.K. Broome-Smith, A. Edelman, S. Yousif, and B.G. Spratt, *Eur. J. Biochem.* **147**, 437–446 (1985). GenBank database accession number X02163.
37. S. Asoh, H. Matsuzawa, F. Ishino, J.L. Strominger, M. Matsuhashi, and T. Ohta, *Eur. J. Biochem.* **160**, 231–238 (1986). GenBank database accession number X04516.
38. M.D. Song, M. Wachi, M. Doi, F. Ishino, and M. Matsuhashi, *FEBS Lett.* **221**, 167–171 (1987). GenBank database accession number Y00688.
39. J.K. Broom-Smith, I. Ioannidis, A. Edelman, and B.G. Spratt, *Nucleic Acids Res.* **16**, 1617 (1988). GenBank database accession number X06479.
40. B.G. Spratt, *Nature* **332**, 173–176 (1988). GenBank database accession number X59632.
41. J.W. Dale, D. Godwin, D. Mossakowska, P. Stephenson, and S. Wall, *FEBS Lett.* **191**, 39–44 (1985). GenBank database accession number X03037.
42. P. Huovinen, S. Huovinen, and G.A. Jacoby, *Antimicrob. Agents Chemother.* **32**, 134–136 (1988). GenBank database accession number J03427
43. B. Joris, F. De Meester, M. Galleni, S. Masson, J. Dusart, J.-M. Frère, J. Van Beeumen, K. Bush, and R. Sykes, *Biochem. J.* **239**, 581–586 (1986).
44. Y. Asano, A. Nakazawa, Y. Kato, and K. Kondo, *Angew. Chem.* **101**, 511–512 (1989).
45. Y. Asano and K. Yamaguchi, *J. Ferment. Bioeng.* **79**, 614–616 (1995).
46. R.A. Nicholas and J.L. Strominger, *Rev. Infect. Dis.* **10**, 733–738 (1988).
47. Y. Asano, in K. Takai ed., *Frontiers and New Horizons in Amino Acid Research*, Elsevier, Amsterdam, 1992, pp. 333–336.
48. Y. Asano, H. Ito, T. Dairi, and Y. Kato, *J. Biol. Chem.* **271**, 30256–30262 (1996). GenBank database accession number D86380.
49. J.J.R. Coque, P. Liras, and J.F. Martin, *EMBO J.* **12**, 631–639 (1993). EMBL database accession number Z13971.
50. D.L. Popham and P. Setlow, *J. Bacteriol.* **175**, 2917–2925 (1993). GenBank database accession number L10629.
51. B. Joris, F. De Meester, M. Galleni, G. Reckinger, J. Coyette, and J.-M. Frère, *Biochem. J.* **228**, 241–248 (1985).
52. M. Galleni, F. Lindberg, S. Normark, S. Cole, N. Honore, B. Joris, and J.-M. Frère, *Biochem. J.* **250**, 753–760 (1988).
53. M. Hobbs, B.P. Dalrymple, P.T. Cox, S.P. Livingstone, S.F. Delaney, and J.S. Mattick, *Mol. Microbiol.* **5**, 543–560 (1991). GenBank database accession number M32230.
54. D.B. McKay, M.P. Jennings, E.A. Godfrey, I.C. MacRae, P.J. Rogers, and I.R. Beacham, *J. Gen. Microbiol.* **138**, 701–708 (1992). GenBank database accession number M68491.
55. H. Komeda and Y. Asano, *J. Mol. Catal. B.* In press.

AMMONIA TOXICITY, ANIMAL CELLS

SADETTIN S. OZTURK
Bayer Corporation
Berkeley, California

KEY WORDS

Ammonia
Ammonium
Cell culture
Fermentation
Inhibition
Ion exchange
Toxicity

OUTLINE

INTRODUCTION

Mammalian cells are used for the production of several important biologicals including interferons, growth factors, vaccines, hormones, and monoclonal antibodies (1–4).

The process efficiency of cell culture fermentation should be increased in most cases for economically viable production (4). Maintaining high numbers of viable cells in the bioreactor for extended periods of time results in higher product concentration and higher reactor productivity, thus increasing process efficiency.

The cell growth rate and the number of cells attained in the bioreactor are determined, in part, by the concentration of nutrients available to the cells. Even though fed-batch or perfusion systems can eliminate nutrient depletion, they are still prone to inhibitory metabolite accumulation. Several metabolic by-products have been reported to inhibit cell growth. Ammonia* and lactate are major metabolic by-products of glucose and glutamine metabolism (5,6), and their presence has been demonstrated to diminish growth of mammalian cells and to influence protein production. Lactate is generated mainly from incomplete oxidation of glucose, and it affects the cells through alteration of culture pH and osmolarity (6). The effects of ammonia on cells are more direct and occur at much lower concentrations. Thus, ammonia is generally more important as an inhibitory metabolic by-product in cell culture.

Ammonia is produced mainly from glutamine, an essential component of cell culture media. Glutamine gives rise to ammonia through chemical degradation and cellular metabolism. Concentration of ammonia in cell culture is influenced by many factors, including the mode of reactor operation, cell concentration, glutamine concentration, and cellular metabolism. Typically, ammonia concentrations reach 2 to 5 mM in batch cultures. This concentration can be still higher in fed-batch reactors where glutamine is the limiting substrate and therefore supplemented continually, leading to increased ammonia accumulation.

Although different cells exhibit different tolerance to ammonia, ammonia concentrations as low as 2 mM have been shown to adversely affect mammalian cells. Numerous studies have reported the effect of ammonia on mammalian cells in different cell culture systems. It has been established that ammonia not only affects cell growth but also alters cell metabolism, product expression, and product quality (7). The action mechanism for ammonia is complicated and not well characterized. Ammonia was shown to alter intracellular pH, membrane potential, and intracellular composition and to result in futile cycles.

Minimization of ammonia inhibition should yield better cell growth and increase the longevity of the cultures. Several techniques were used to minimize ammonia accumulation in cell culture systems. Media modification involves replacement or elimination of glutamine from cell culture medium. The success of this method depends on the cell line, and in most cases the cells should be adapted or genetically altered. Alternatively, the ammonia generated in the media can be removed by physical and chemical means.

In this article, we review the current knowledge on ammonia inhibition in cell culture. The sources of ammonia accumulation; the effects of ammonia on cell growth, metabolism, product secretion, and product quality; the mechanism of ammonia inhibition; and the methods of reducing ammonia concentration in cell culture are discussed.

*In this text we use ammonia in reference to both gaseous (ammonia) and ammonium ion. The ammonia and ammonium ion are present simultaneously in the culture. The relative concentration of ammonium ion is determined by the pH-dependent equilibria.

SOURCES OF AMMONIA ACCUMULATION

There are two major sources for ammonia accumulation, and in both cases glutamine is involved. Glutamine is one of the important energy sources for the mammalian cells in vitro (8,9). Addition of glutamine has been shown to stimulate cell growth and antibody production (10–15). Because almost all conventional cell culture media contain glutamine, the problem of ammonia accumulation is of universal importance (16,17). However, depending on the cell line and media composition, the effects can be different from one case to another.

Generation of Ammonia Caused by Glutamine Degradation

Glutamine is not stable in aqueous solutions and is degraded to yield a five-member ring structure, pyrrolidone carboxylic acid, and ammonia (18–20):

$$\text{Glutamine} \rightarrow \text{Pyrrolidone carboxylic acid} + NH_4^+ \quad (1)$$

This reaction is irreversible, and the extent of the reaction is affected by the chemical environment of the medium. Glutamine degradation follows first-order kinetics, the half-life of glutamine is reported to be between 6 and 20 days, depending on the conditions (16,21–23).

Glutamine degradation is pH (18,21,22,24) and temperature (20) dependent. Alkaline pH increases the degradation rate, and altering pH from 7.2 to 7.6 was reported to elevate the degradation rate by 300% (21,22). The presence of serum, especially if not heat inactivated, can elevate degradation of glutamine and subsequent ammonia accumulation caused by glutaminase and asparaginase activity (25). The media composition, especially the phosphate levels in the media, affects glutamine degradation significantly (22).

Glutamine degradation is an important problem for cell culture. The degradation not only generates inhibitory ammonia, but also decreases the level of glutamine available to the cells. The cell performance is affected negatively because of these two factors.

Generation of Ammonia Caused by Cell Metabolism

Although ammonia can be produced from other amino acids, glutamine metabolism is established to be the most important pathway. The pathways involving ammonia are summarized in Figure 1 and have been studied extensively in the literature (9,26–46).

Glutamine metabolism primarily takes place in mitochondria. Glutamine is initially deaminated by glutaminase, with ammonia as a byproduct (47,48). Ammonia can also be removed from glutamine via biosynthetic pathways, where the ammonia group is used in the formation of a biosynthetic product, particularly for pyrimidine and purine base synthesis (Fig. 1). The product formed by ei-

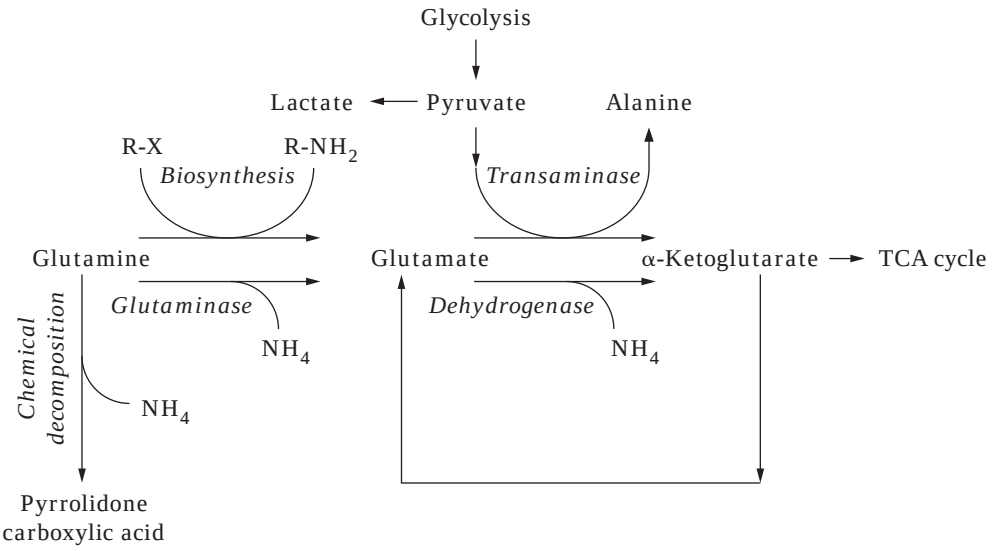


Figure 1. Ammonia generation by glutamine degradation and cell metabolism.

ther route is glutamate. The second step of glutaminolysis is the conversion of glutamate to α-ketoglutarate. Again, this conversion may be accomplished by more than one enzyme. An active transaminase can remove the ammonia group from glutamate. Less-active transaminases appear to deliver the ammonia group to the glycolytic intermediate 3-phosphoglycerate to form serine, which subsequently can be transformed to glycine. Glutamate can also be transformed by glutamate dehydrogenase to form α-ketoglutarate, which liberates a second ammonium ion. α-Ketoglutarate is then metabolized via the tricarboxylic acid (TCA) cycle.

The amount of ammonia generated is dependent on the metabolic pathway for glutamine. Depending on the pathway and the fate of glutamine, the ammonia yield on glutamine varies. Based on stoichiometry, Glacken (49) reported possible scenarios for glutamine metabolism and ammonia yield, between 1 and 2 mol of ammonia are generated per mole of glutamine consumed (Table 1). These yield coefficients, however, are very high and are not typically observed in cell culture.* For hybridoma cells, the ammonia yield was reported to be between 0.5 to 1 mol of ammonia generated per mole of glutamine consumed (6,33,50–52). Glutamine concentrations in typical cell culture media range between 1 and 7 mM (17). Thus, the ammonia concentrations in cell culture can reach 5 mM (6,50–53). Consequently, the higher glutamine concentrations found in special media formulations yield much more ammonia levels. Glutamine concentrations used in fed-batch cultures, for instance, can vary between 8 and 40 mM, and very high ammonia concentrations are obtained (14,54,55).

*It should be noted that the determination of ammonia yield from glutamine requires inclusion of the glutamine degradation rate in the analysis. When the degradation rate is omitted, the analysis results in erroneously high "apparent yields" (22).

Table 1. Amount of ATP Produced and Ammonia Generated in Different Metabolic Pathways

End product	ATP produced (mol/mol Gln)	NH_4 produced (mol/mol Gln)	NH_4/ATP
Alanine	9	1	0.11
Aspartate	9	1	0.11
Acetyl-CoA	15	2	0.13
Pyruvate	12	2	0.17
Lactate	9	2	0.22

Source: Glacken (49).

EFFECTS OF AMMONIA ON CELL PHYSIOLOGY

The influence of ammonia concentration on mammalian cell growth and product expression has been reported for a variety of cell lines and culture conditions (56). Most of these studies emphasized the influence of ammonia on cell growth. The effects of elevated ammonia concentrations on metabolic rates and cell productivity have been studied less extensively. Although it is not discussed in detail in this article, the effect of ammonia on virus replication and virus production from cell culture was demonstrated clearly in the literature (57–63).

Effects of Ammonia on Cell Growth

Although ammonia inhibited cell growth in all these studies, the sensitivity of the growth rate to ammonia concentration varied among the cell lines used. The extent of ammonia inhibition seems to be more severe for primary cells, whereas transformed cells tolerate higher ammonia levels. For 3T3 cells, ammonia levels as low as 0.6 mM reduced cell growth by 60%. However, when the cells are transformed, the reduction was reduced to 15% (64). Addition of 2 mM ammonia reduced growth by 30% for BSC-1 monkey epithelial cells (65). For established cell lines, inhibitory effects of ammonia were observed at higher ammonia concentrations. Hassell et al. (66) reported reduction of cell

growth at 2 mM for a variety of cell lines. Several reports indicate ammonia inhibition for hybridoma cells with varying severity (6,41,50,52,67–82). Inhibitory concentrations of ammonia were observed to be in the range of 2 to 10 mM in these studies. For BHK cells, even lower concentrations of ammonia resulted in growth inhibition. Wentz and Schügerl (83) reported a 80% reduction at 1 mM ammonia, and Butler and Spier (84) reported a 75% decrease in growth at 3 mM of ammonia in BHK cells. For CHO cells, the inhibitory ammonia levels seems to be higher. Kurano et al. (85) obtained a 50% reduction in growth at 8 mM ammonia. Some cell lines are reported to be more resistant to ammonia. For instance, Schneider (81) observed only a 20% reduction in growth at 6 mM ammonia for HeLa cells, and Hansen and Emborg (86) did not observe any inhibition at up to 8 mM for a CHO clone. Other factors that can influence the inhibitory effects of ammonia on growth are the serum level and cell adaptation. Holley et al. (65) for instance, observed that increasing the level of serum from 0.1 to 10% reduced the ammonia inhibition. Miller et al. (52) observed adaptation of hybridoma cells to 8 mM ammonia. Maiorella et al. (54) used serial passaging in the presence of 5 mM ammonia and demonstrated that the resulting cells survived better in 10 mM ammonia. Adaptation to high ammonia concentrations can be used for selection of ammonia-resistant cell lines.

The effects of ammonia on cell growth in most of these studies reported were quantified as the reduction in cell density, and only a few studies evaluated the specific growth rates as a function of ammonia concentration. Ammonia inhibition on the specific growth rate can be described by a second-order inhibition model:

$$\mu = \frac{\mu_0}{1 + [NH_4^+]^2/K_a} \qquad (2)$$

where μ and μ_0 are the growth rates in the presence and absence of ammonia, respectively; $[NH_4^+]]$ is the initial ammonia concentration; and K_a is the inhibition constant. The square root of K_a corresponds to the ammonia concentration at which the specific growth rate decreases by 50%. Figure 2 presents the effect of ammonia on cell growth rate and cell density for 163.4.G5.3 murine hybridoma cells. Equation 2 describes the cell growth rate accurately with parameters $\mu_0 = 0.037$/h and $K_a = 24$ mM^2. For comparison, K_a value of 26 mM^2 was reported for the CRL-1606 hybridoma cell line (87), however, the value reported for the SB-4082 hybridoma cell line was almost one order of magnitude lower ($K_a = 3.2$ mM^2) (65).

Effects of Ammonia on Cell Death

Even though ammonia affects the cell growth rate and maximum cell density, effect on cell death is not well documented. Ozturk et al. (6) reported no effect of ammonia on the cell-specific death rate in the range of 0 to 5 mM; their data on the death rate are presented in Figure 2. When the ammonia concentration was increased from 0 to 3.75 mM, the specific death rate did not change. Ozturk and Palsson (88) also observed that elevated ammonia concentrations did not significantly influence the cell viability.

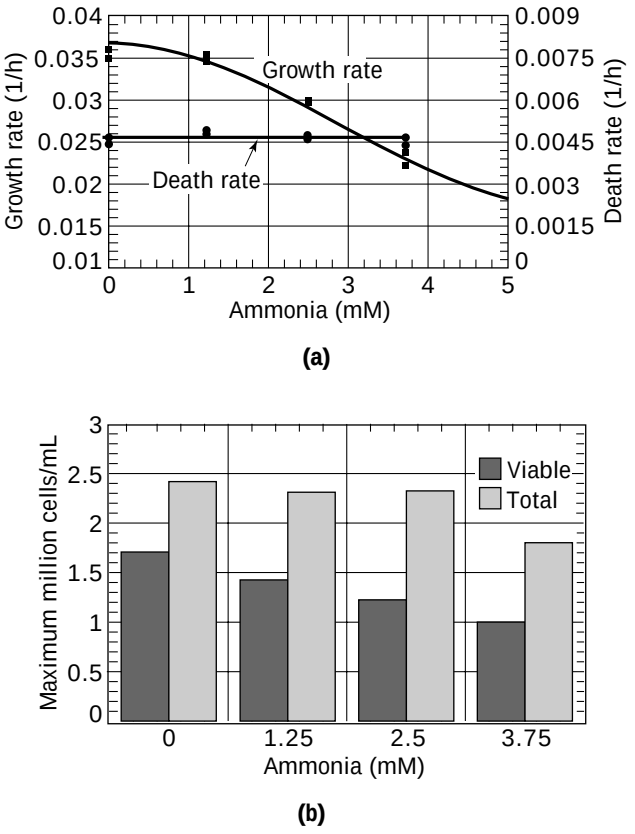


Figure 2. Effect of ammonia on murine hybridoma culture. (**a**) Cell growth and death rates, and (**b**) maximum viable and total cell concentrations in batch culture. *Source*: Data obtained from Ozturk and Palsson (88).

Thus, the reduction in cell density at higher ammonia concentrations was simply the result of lower cell specific growth rates. On the other hand, Goergen et al. (89) used much higher concentrations of ammonia (up to 17 mM) and observed an increase in the death rate. Newland et al. (80) also reported the influence on death rates by ammonia at high concentrations. More quantitative studies are needed to verify the effects of ammonia on specific death rates and the mechanism(s) involved.

Effects of Ammonia on Cell Metabolism

A number of studies reported the effects of ammonia on cell metabolism. For a murine hybridoma cell line, Ozturk et al. (6) observed a roughly twofold increase in glucose and glutamine consumption rates in cultures exposed to 3.75 mM of ammonia, compared to control conditions (Fig. 3). The production rates of lactate and ammonia were also enhanced by the presence of ammonia, indicating an elevated metabolic state of the cells. The yield coefficients of ammonia from glutamine and lactate yield from glucose decreased about 15% with 3.75 mM ammonia (Fig. 3). A similar elevation in metabolic rates was reported by Miller et al. (52) for AB2-143.2 hybridoma cells and by Alex et al. (90) for C127 cells. Glacken (87) has also shown an increase in the glutamine consumption rate by added ammonia.

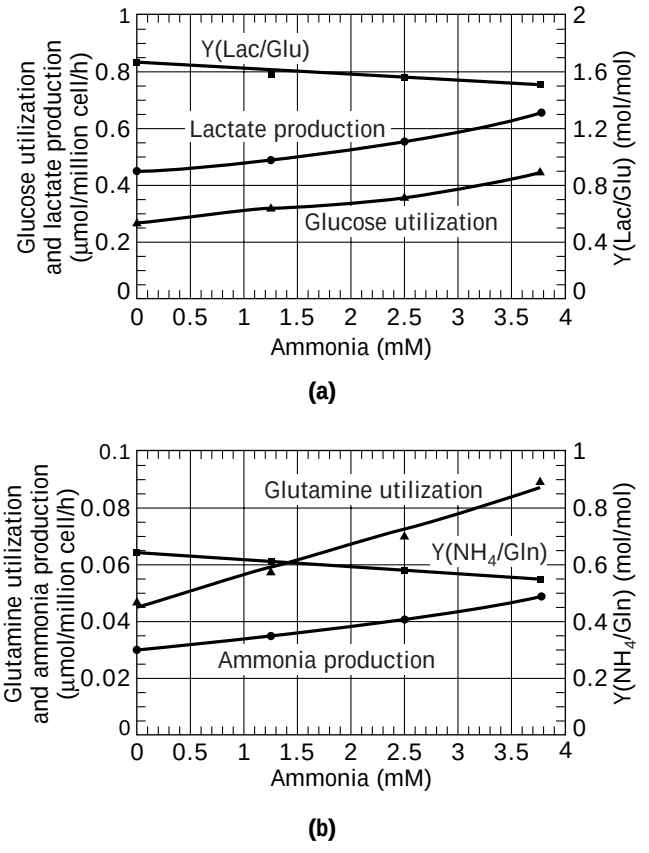


Figure 3. Effect of ammonia on murine hybridoma culture. (**a**) Glucose consumption, lactate production, and lactate yield on glucose; (**b**) Glutamine consumption, ammonia production, and ammonia yield on glutamine. Y, yield; Lac, lactate; Gln, glutamine; Glu, glucose. *Source*: Data obtained from Ozturk and Palsson (88).

McQueen and Bailey (91) indicated an increase in hybridoma cell yield from glucose and glutamine that resulted from a decrease in growth rate and an increase in metabolic rate.

Miller et al. (52) and Ozturk et al. (6) studied the effects of ammonia on amino acid metabolism. Table 2 summarizes the effects of ammonia on amino acid metabolic rates for murine hybridoma 163.4G5.3 (6). For this cell line, glutamate, serine, glycine, and alanine were produced and all other amino acids were consumed. The consumption or production rates of these amino acids were doubled at 3.75 mM ammonia. The increase in the amino acid consumption or production rates were parallel to the glutamine consumption rate because they increased at higher ammonia concentrations by the same magnitude. Only serine, arginine, alanine, and valine showed a relatively different response to ammonia concentration. The consumption rate for serine decreased. The rate of increase for arganine (1.5-fold increase at 3.75 mM ammonia) was lower compared to glutamine (1.8-fold increase). On the other hand, the increase in alanine production (2.1-fold) and valine consumption (2.6-fold) was higher (6).

The yield coefficients of lactate from glucose and ammonia from glutamine decreased, whereas alanine yield

Table 2. Amino Acid Consumption and Production Rates (Parentheses) as a Function of the Ammonia Concentrations

	$[NH_4^+]$			
Amino acid	0 mM	1.25 mM	2.5 mM	3.75 mM
Aspartate	1.74	2.21	2.58	3.75
Glutamate	(2.92)	(1.69)	(2.03)	(3.25)
Asparagine	2.13	2.26	2.67	4.18
Serine	(2.45)	(2.07)	(2.03)	(0.47)
Glutamine	48.50	58.00	67.50	89.00
Histidine	0.12	0.23	0.35	0.20
Glycine	(0.58)	(0.32)	(0.66)	(0.66)
Threonine	3.10	3.01	3.86	4.14
Arginine	3.18	3.24	3.61	4.89
Alanine	(34.92)	(42.78)	(55.33)	(72.14)
Tyrosine	3.75	4.51	5.08	6.84
Methionine	2.14	2.98	2.54	3.51
Valine	5.68	8.46	9.61	14.70
Phenylalanine	2.21	3.48	3.86	4.57
Isoleucine	7.06	10.84	11.26	14.16
Leucine	8.64	11.14	13.35	18.57
Lysine	5.39	6.94	8.17	11.73

Source: Ozturk et al. (6).
Note: Rates are in nmol/106/cell/h.

from glutamine increased as a result of ammonia addition. Miller et al. (52) observed similar data for ammonia yield from glutamine. Hansen and Emborg (86) also observed a decrease in ammonia yield from glutamine at high ammonia concentrations for CHO cells. A decrease in the lactate yield coefficient has been reported at elevated ammonia concentrations for other mammalian cells (90). Similarly, McQueen and Bailey (71) reported a decrease in lactate yield from glucose and in ammonia yield from glutamine. An increase in the alanine yield from glutamine was reported by Miller et al. (52). An increase in amino acid metabolic rates was observed for the C127 cell line (90).

The effects of ammonia on amino acid metabolism can be interpreted in the view of Figure 1. Ozturk et al. (6) showed a 12% increase in the alanine yield from glutamine at elevated ammonia concentrations, whereas the ammonia yield decreases by the same amount. These data indicate that at elevated ammonia concentrations, relatively more of the glutaminolytic flux is via the alanine transaminase pathway than that of glutamate dehydrogenase. This shift could be due to the fact that ammonia is a by-product of the dehydrogenase reaction, which is believed to operate nearly at equilibrium; thus, elevated ammonia concentrations would shift the equilibrium toward glutamate (Fig. 1). Both Miller et al. (52) and Ozturk et al. (6) showed an increase in the specific glutamine consumption rate at elevated ammonia concentrations. This increase in the consumption rate suggests that inhibition of glutaminase by ammonia is not operative in the hybridoma cell line investigated.

Effects of Ammonia on Energy Metabolism

Mammalian cells use various pathways to produce energy in the form of adenosine triphosphate (ATP). The produc-

tion rate of ATP can be estimated using the lactate production and oxygen consumption rates (87,92,93):

$$q_{ATP} = q_{Lac} + 2(P/O) \cdot q_{O_2} = q_{Lac} + 6 \cdot q_{O_2} \quad (3)$$

where q_{ATP}, q_{Lac}, and q_{O_2} are the production rates of ATP, lactate, and oxygen, respectively, and P/O is the phosphorylation ratio. The term q_{Lac} corresponds to the contribution of glycolysis, and the term $6 \cdot q_{O_2}$ corresponds to the contribution of oxidative phosphorylation.

Miller et al. (52) observed a decrease in oxygen uptake at high ammonia concentrations in a continuous reactor operation. On the other hand, Kimura et al. (94) and Ozturk et al. (6) reported no significant effect of ammonia on oxygen consumption rates for a human leukemia line and a hybridoma line, respectively. The data of Ozturk et al. (6) on the effect of ammonia on oxygen consumption rates are presented in Figure 4A.

Miller et al. (52) and Ozturk et al. (6) reported an increase in the ATP production rate at high ammonia concentrations. The dependence of the ATP production rate on ammonia is presented in Figure 4a. Although cells generated more ATP at higher ammonia levels, this increased energy production was not, however, used for growth, because the growth rate was inhibited (6). It appears that the cells produce more energy under stressful conditions and use it for maintenance. The relative contribution of oxidative phosphorylation was observed to decrease, and the contribution of glycolysis increased with elevated ammonia levels (6,52).

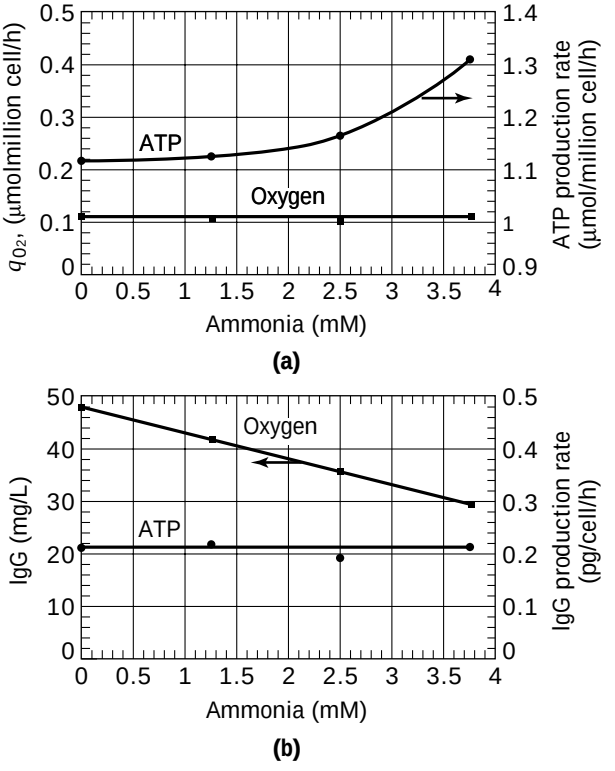


Figure 4. Effect of ammonia on murine hybridoma culture. (**a**) Oxygen consumption and ATP production rate, (**b**) monoclonal antibody concentration and antibody production rate. *Source*: Data obtained from Ozturk and Palsson (88).

Effects of Ammonia on Product Expression

Because ammonia affects cell growth and metabolism, it is conceivable that it also affects the product expression. Either directly or indirectly, ammonia can affect the product concentration obtained in the culture and thus the process efficiency.

The indirect effect of ammonia on production results from the reduced cell growth rate and subsequent cell density achieved in the culture. Even though the specific productivity may not have been altered, the cell density and longevity of the culture can be restricted by high ammonia levels. Thus, cultures with high ammonia levels result in lower product concentrations. Several studies have reported a decrease in monoclonal antibody production at high ammonia concentrations, mainly because of lower cell densities obtained (6,70,71). Figure 4b presents the effect of ammonia on antibody titer and specific productivity for a murine hybridoma line in batch culture. Although the titer was reduced, the specific antibody production rate remained constant under these conditions.

The direct effect of ammonia on specific productivity varies and the mechanism(s) are poorly understood. Dyken and Sambanis (95) observed a slight increase in protein secretion from AtT-20 cells. Ammonia was reported to alter the secretion rate of interferon (INF) γ (96) and synthesis of INF-β; (97); however, the specific production rates were not provided in these studies. Hansen and Emborg (86) reported a decrease in tissue-type plasminogen activator (t-PA) production from CHO cells at high ammonia concentrations. Specific antibody productivity was evaluated for a number of hybridoma cell lines in the literature. Glacken (87) reported a decrease in antibody productivity at elevated ammonia levels for the C-1606 hybridoma line. Other investigators, however, reported an unaltered specific antibody production rate (6,70,71).

Effect of Ammonia on Product Quality

Ammonia was reported to affect the conditions in intracellular compartments and enzymatic reactions for protein processing. High ammonia concentrations were reported to alter protein after translations, glycosylation, and secretion. Ammonia can alter the intracellular or intracompartmental pH, alter membrane potentials, and directly interact with enzymes. Thorens and Vassalli (98) reported inhibition of sialic acid transferase for Immunoglobulin (Ig) M at 10 mM ammonia. The results indicate a pH increase in Golgi because of ammonia load. Similarly, Andersen and Goochee (99) reported significant reduction in the terminal sialylation of O-linked glycosylation of recombinant granulocyte colony-stimulating factor. Borys et al. (100) studied the N-linked glycosylation of recombinant mouse placental lactogen-I from CHO cells and observed inhibition of glycosylation by ammonia. The effect of ammonia was also dependent on extracellular pH. Gawlitzek et al. (101,102) studied the glycosylation pattern of recombinant proteins expressed by BHK-21 cells. For the production of HuIL-2 variant in BHK cells, ammonia was

observed to increase the intracellular uridine diphosphate-N-actylglucosamine (UDP-GlcNAc) pool. UDP-GlcNAc is a precursor substrate for the glycosylation process in the cytosol and Golgi. High ammonia levels led to a decrease in terminal sialyation and to an increase in branching. To test the hypothesis that UDP-GlcNAc is involved, the authors used glucosamine, which is the precursor for UDP-GlcNAc. Kopp et al. (103) studied product consistency and glycosylation patterns in recombinant CHO-expressed glycoproteins. High ammonia levels influenced glycosylation patterns significantly for INF-ω and t-PA. The alteration of product glycosylation by ammonia was described in monoclonal antibody production (104). To minimize the effects of ammonia, a multilevel pH control was proposed for commercial production. Genetic and process engineering strategies for control of ammonia in cell cultures were recently described to yield enhanced glycosylation (104).

MECHANISM FOR AMMONIA INHIBITION

Although a significant number of articles have been published on the affect of ammonia, the mechanism of action is not fully understood. Several mechanisms were hypothesized and tested experimentally for ammonia inhibition. Alteration of intracompartmental pH and membrane potential, futile cycles, metabolic inhibition, and alteration of critical ribonucleotides are identified as potential mechanisms for ammonia inhibition.

Intracellular pH

A possible explanation for the effect of ammonia is simply the alteration of intracellular and intracompartmental pH (6,72,106,107). Ammonia in cell culture is present in both gaseous and ionic form, and these species are in equilibrium at a ratio determined by the medium pH. At normal pH values, almost all the ammonia present is in the form of ammonium ion, and the concentration of gaseous ammonia is very low ($pK_a = 9.2$).

The regulation of intracellular (cytosolic) and intracompartmental pH in response to added ammonia is complex (Fig. 5). Both gaseous ammonia and ammonia ion can permeate the membranes. This permeation involves both passive (diffusion) and active transport.

The permeability of gaseous ammonia is much greater than that of the ammonium ion (6,49,72,107,108). Thus, initially, gaseous ammonia permeates rapidly, raising the intracellular pH. Then, the slower penetration of ammonium ion decreases the intracellular pH. McQueen and Bailey (72,107) and Ozturk et al. (6) studied the response of hybridoma cells to added ammonia and verified this mechanism. Figure 6 shows the response of intracellular pH to ammonia addition, followed by fluorescent dye. After the addition, there is an immediate increase in intracellular pH. This is due to rapid gaseous ammonium diffusion into the cells. The intracellular pH decays to steady-state intracellular pH values that are lower than those observed prior to the exposure to ammonia because of the transport of ammonium ion. The steady-state intracellular pH values for this experiment are presented in Table 3.

The rate of diffusion for ammonium ion through the cell membrane is four to five orders of magnitude lower than the rate of diffusion for gaseous ammonia (109). However, ammonium ion can be actively transported across the cell membrane via [Na,K]ATPase, [Na,K,Cl]-cotransporter, and [Na,H]-exchanger, as indicated in Figure 5 (109–114). The dynamics of ammonia transport for mammalian cells were analyzed by mathematical models (107).

The transport of ammonia and ammonium ion to the cells were studied in detail by Martinelle and Haggsrom (115–117) for the murine myeloma cell (Sp2/O-Ag14). In addition to diffusional transport, these authors identified the active transport of ammonium ion via [Na,K]ATPase and [Na,K,C]-cotransporter in hybridoma cells. The presence of K^+ could inhibit the [Na,K]ATPase active transport and alter the transport of ammonium ion. When K^+ (10 mM) was added, the change in intracellular pH due to ammonia addition was observed to be negligible. The authors postulate that one of the reasons for ammonia inhibition is the increased energy demand resulting from energy wasted because NH_4^+ is actively transported to the cells via [Na,K]ATPase, and they proposed the use of K^+ for minimization of ammonia inhibition. Use of KOH instead of NaOH for pH control is proposed to increase K^+ concentration in the culture.

A change in intracellular or intracompartmental pH can alter the activity of numerous enzymes and, depending on the location, different results can be obtained. The *net* result of ammonium ion transport to the cells is a decrease in intracellular pH. However, the ammonia transported to the cells further diffuses to compartments in the cells. Gaseous ammonia easily permeates intracellular membranes and the pH in mitochondria, Golgi, endoplasmic reticulum, and lysosymes are elevated by its presence. The pH in vesicles is normally lower than in cytosol (118,119), and the pH can be altered by ammonia. The rise in pH by the transport of ammonia gas into the lysosomes has been demonstrated (120,121).

The generation of ammonia inside the cell caused by cellular metabolism alters intracompartmental pH. Ammonia is generated mainly in mitochondria because of glutamine metabolism. Ammonia gas readily diffuses out of the mitochondria, and the pH decreases. Ammonia diffuses not only to the cytoplasm of cells but also to the compartments. In all cases, ammonia that diffuses from mitochondria increases the pH in other intracellular compartments (i.e., Golgi, endoplasmic reticulum, and lysosymes).

McQueen and Bailey (72,107) have shown that the net result of ammonia addition is a decrease in intracellular pH for the hybridoma line ATCC TIB 131. When internal pH was altered by external pH, McQueen and Bailey associated ammonia effects to the variations in intracellular pH, because both ammonia addition and low external pH resulted in lower cell yields on glucose and glutamine. On the other hand, Ozturk et al. (6) individually evaluated the cell growth rate and the metabolic rates and could not relate the effects of ammonia to the effects of lowering external pH. There is a decrease in glucose consumption and an increase in glutamine uptake rates when pH was controlled below 7.2 (93). However, a decrease in intracellular pH as a result of ammonia addition increased both rates.

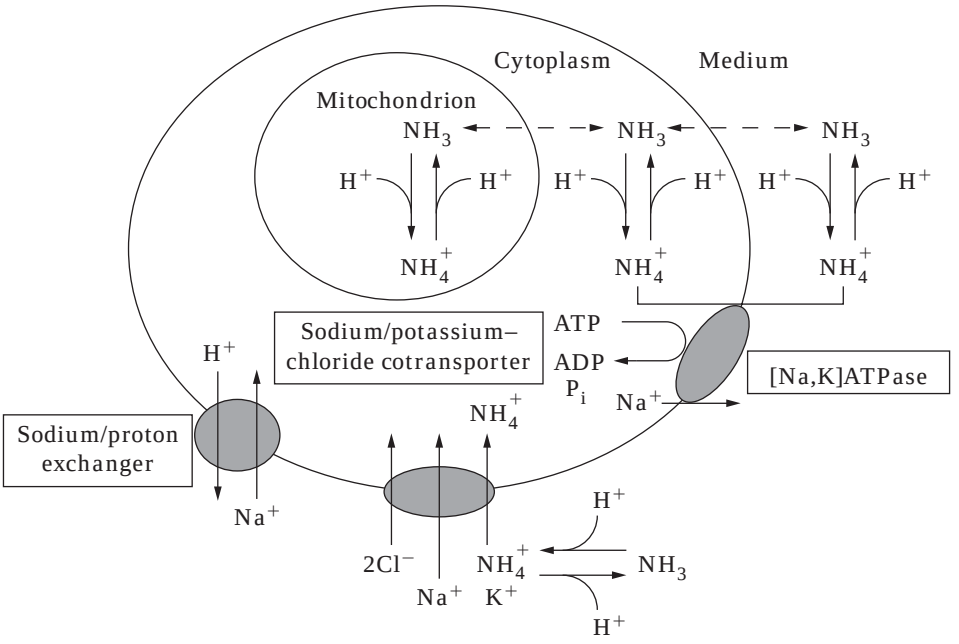


Figure 5. Ammonia transport and alteration of intracellular pH.

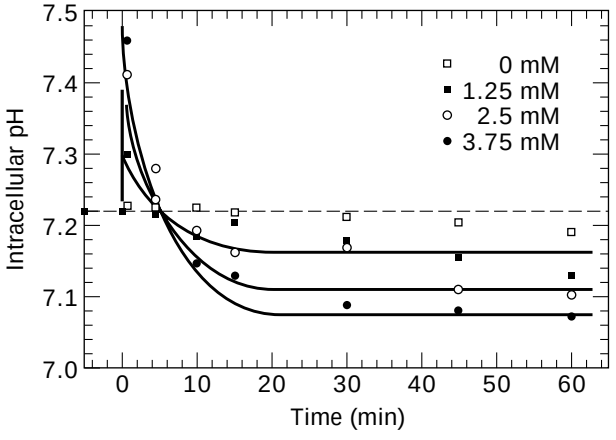


Figure 6. Alteration of intracellular (cytosolic) pH in response to ammonia addition. *Source*: Data obtained from Ozturk and Palsson (88).

Table 3. Effect of Ammonia on Intracellular pH at External pH of 7.20

$[NH_4^+]$ mM	Peak pH_i	Steady-state pH_i	Delta pH
0	7.23	7.19	0
1.25	7.3	7.14	−0.05
2.5	7.42	7.1	−0.09
3.75	7.43	7.07	−0.13

Source: Ozturk et al. (6).

ATP production was also influenced differently from ammonia addition (increasing ATP production) and from lowering external pH (decreasing ATP production). Miller et al. (52) observed an increase in glucose consumption rate as a result of ammonia addition, although the decrease in extracellular pH led to a decrease in glucose consumption rate. Thus, it can be concluded that the mechanisms of the ammonia effects may not be as simple as they were originally thought to be. The alteration of intracellular pH by ammonia cannot alone explain the observations.

Membrane Potential

As previously mentioned, ammonium ions can be transported across cellular membranes via [Na,K]ATPase and [Na,K,Cl]-cotransporters. Ammonium ions compete with K^+ ions and disturb the membrane potential. The [Na,K]ATPase transport system has a high energy demand (55,122,123), and the competition of ammonium ions with K^+ ions decreases the efficiency. Martinella and Hagstrom (116) reported futile cycles originating from ammonia diffusion and ammonium ion transport. These futile cycles increase the maintenance energy required by the cells. When ammonia is generated in the mitochondria, it diffuses out. However, ammonia is transported back in the form of ammonium ion, and cells use [Na,K]ATPase for this transfer. Thus, the alteration of membrane potential and elevation of maintenance energy are possible mechanisms for ammonia inhibition.

Alteration of Metabolic Pathways

Ammonia can also interact with the enzymes and alter the metabolic pathways. Phosphofructokinase (PFK) can be

activated by ammonia (124,125). PFK is a highly regulated key enzyme for glucose metabolism, and its activation by ammonia can explain the increase in glucose consumption at high ammonia levels. Ammonia can also affect the metabolic pathways in mitochondria. The glutamine metabolism pathway, for instance, involves glutaminase, transaminase, and glutamate dehydrogenase reactions. The glutaminase enzyme is known to be inhibited by ammonia. However, Ozturk and Palsson (88) did not observe a decrease in the specific glutamine consumption rate at elevated ammonia concentrations. The glutamate dehydrogenase reaction is known to operate near equilibrium, and elevated ammonia levels would shift the equilibrium toward glutamate. At high ammonia levels, Ozturk and Palsson (88) observed a shift to the glutaminolytic flux via the alanine transaminase reaction from glutamate dehydrogenase, confirming this change. Thus, at high ammonia concentrations, the ammonia production by the cells decreased and alanine yield increased. This finding was also supported by other investigators (52,87,126).

The involvement of ammonia in metabolic pathways can explain the effects on cell metabolism. Growth inhibition by ammonia, on the other hand, can be explained by the presence of futile cycles (49,127). Futile cycles are cyclic reactions in which the hydrolysis of ATP is the only result (49); thus, cells have to increase ATP production to maintain the energy required for their survival.

Tagler et al. (127) studied the glutamate dehydrogenase and transhydrogenase enzymes in isolated mitochondria and liver cells. This enzyme system catalyzes the following reactions:

$$\text{Glutamate} + \text{NAD}^+ \leftrightarrow \alpha\text{-Ketoglutarate} + \text{NH}_4^+ + \text{NADH} \quad (4)$$

$$\text{NADH} + \text{NADP}^+ + \text{ATP} \leftrightarrow \text{NADPH} + \text{NAD}^+ + \text{P}_i \quad (5)$$

$$\alpha\text{-Ketoglutarate} + \text{NADPH} + \text{NH}_4^+ \leftrightarrow \text{Glutamate} + \text{NADP}^+ \quad (6)$$

Ammonia can also drive the glutaminase–glutamine synthetase enzyme system in a futile cycle mode. Glutaminase enzyme catalyzes the deamidation of glutamine to glutamate and ammonia. Glutamine synthetase, on the other hand, uses ATP to convert the glutamate to glutamine:

$$\text{Glutamine} \leftrightarrow \text{Glutamate} + \text{NH}_4^+ \quad \text{(Glutaminase)} \quad (7)$$

$$\text{Glutamate} + \text{NH}_4^+ + \text{ATP} \leftrightarrow \text{Glutamine} + \text{ADP} \quad \text{(Glutamine synthetase)} \quad (8)$$

Thus, for both glutamate dehydrogenase–transhydrogenase and glutaminase–glutamine synthetase cycles, the ATP is dissipated and the reactions are driven by ammonia. This increases the maintenance energy of the cells and can explain the inhibitory effects of ammonia (49). Although these futile cycles have been studied for isolated mitochondria, their functionality has not been demonstrated for mammalian cells in culture.

Alteration of Ribonucleotide Pools

Another explanation for ammonia inhibition comes from detailed studies on various ribonucleotides by Ryll and Wagner (128) and Ryll et al. (129). The intracellular pools of UDP-*N*-actylglucosamine (UDP-GlcNAc) and UDP-*N*-actylgalactosamine (UDP-GalNAc) have been shown to be responsible for inhibition of protein, DNA, and RNA synthesis (129–132). The pool of UDP-GNAc refers to both UDP-GlcNAc and UDP-GalNAc collectively, and it was demonstrated to be elevated in response to increased ammonia levels for a variety of cell lines. Ammonia is incorporated into the glycolytic pathway in cytosol (Fig. 7). Fructose-6-phosphate and ammonia combine to form glucosamine-6-phosphate, a direct precursor for the UDP-GNAc pool in the cytosol. In the mitochondria, ammonium leads to formation of carbamoylphosphate. When the ammonia concentration is high in the mitochondria, then subsequently up-regulated carbamoylphosphate can enter the cytoplasm. Carbamoylphosphate in the cytosol stimulates de novo synthesis of UMP and UTP. Finally, UTP and glucosamine-6-phosphate are combined to supplement the UDP-GNAc pool.

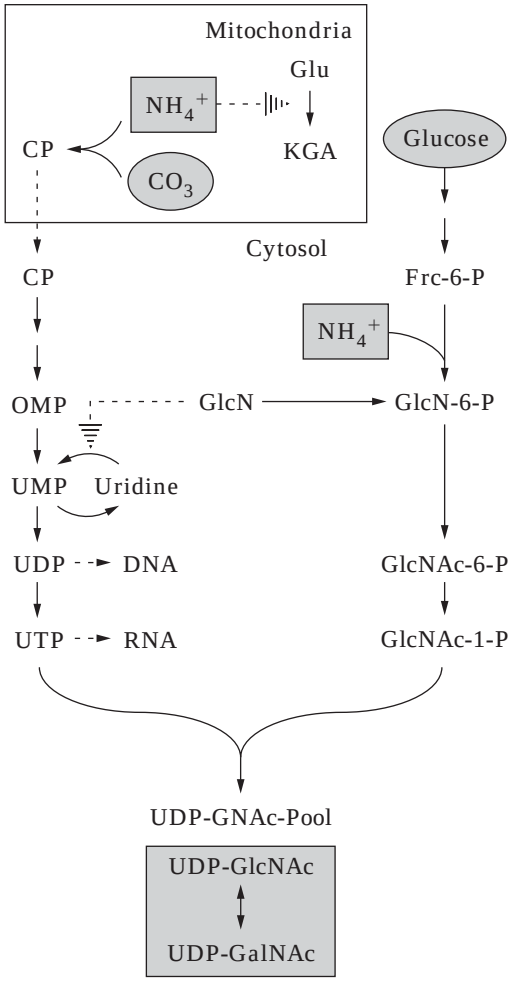


Figure 7. Biosynthesis of UDP-GNAc pool after ammonia application. *Source*: Adapted from Ryll et al. (129).

Ryll et al. (129) studied the biochemistry of the UDP-GNAc pool formation and regulation of the enzymes involved. The formation of the UDP-GNAc pool was observed to be dependent on both glucose and ammonia. Ammonia increases the size of the UDP-GNAc pool, which has been correlated with growth inhibition. It is also important to note that the UDP-GlcNAc is also known as a precursor substrate for glycosylation processes in the endoplasmic reticulum and Golgi. Thus, the effects of ammonia on glycosylation could also be explained by this mechanism (101,102).

CONTROL OF AMMONIA INHIBITION

Several techniques were used to minimize ammonia accumulation in cell culture. The techniques can be classified into two groups: minimization of ammonia generation and removal of ammonia.

Minimization of Ammonia Generation

As mentioned before, ammonia is generated mainly by the spontaneous degradation and cellular metabolism of glutamine. Ammonia generation can be minimized by controlling these processes.

Minimization of Glutamine Degradation. As previously mentioned, the glutamine degradation rate is dependent on several factors, and it can be minimized by several strategies. Temperature is the most obvious factor, and the media should be kept cold when stored. In the reactor, however, the temperature cannot be lowered much. Other factors, such as media pH, can be used to minimize the degradation. The pH optimum for cells varies between pH 7.0 and 7.4 (22,23). Cultivation at lower pH values should decrease the degradation rate and ammonia generation. Serum, if used, should be inactivated to minimize any glutaminase activity. Finally, glutamine degradation is dependent on several ions, such as phosphate. Various media should be evaluated for a given cell line, and glutamine degradation rates for each should be compared.

Attempts have been made to replace glutamine by glutamine-containing peptides (133,134). Peptides such as alanyl-glutamine or glycyl-glutamine are more stable, and the media prepared with these peptides can be autoclaved. These peptides were also shown to generate less ammonia. Although the cells exhibited a considerable lag phase, the cell yields from peptide-substituted media were comparable to those obtained from standard media. Holmlund et al. (135) used glycyl-glutamine and *N*-acetyl-glutamine for CHO cultures. Although cells did not perform well in the presence of *N*-acetyl-glutamine (low growth and t-PA production), glycyl-glutamine provided good culture performance.

Control of Cellular Metabolism. The specific ammonia production rate is affected by the cellular environment, and variables such as pH, dissolved oxygen (DO), temperature, and the level of metabolites can be manipulated to minimize the production rate.

There seems to be an optimal pH for reducing the ammonia production rate. For instance, Ozturk and Palsson (93) observed the ammonia production rates to be minimal at pH 7.2. The cell growth rates were optimal, and glutamine consumption was also observed to be minimal at this pH. The ammonia yield from glutamine, on the other hand, was reduced at low pH values (92,136). Cells seem to produce less ammonia when they are growing under optimal conditions. At very low (1% air saturation) and very high (100% air saturation) DO levels, the ammonia generation rates increased (93,136,137).

In most cases, the ammonia production follows closely with the glutamine consumption with a relatively constant yield coefficient. Reducing the glutamine concentration ([Gln]) decreases the glutamine consumption (q_{Gln}), and Monod-type saturation kinetics can be used to describe the data:

$$q_{Gln} = q^0_{Gln} \frac{[Gln]}{K_m + [Gln]} \tag{9}$$

where q^0_{Gln} and K_m are constants. The value of K_m is in the order of 0.2 to 1 mM (10,87).

Reducing glutamine not only decreases the glutamine consumption and ammonia production rates, but also decreases the yield of ammonia from glutamine. When glutamine was fed in a controlled fashion below 1 mM levels, Glacken et al. (50) observed a substantial decrease in ammonia production in Madin-Darby bovine kidney (MDCK) and human fibroblast cells. A similar strategy was successfully used to reduce ammonia production by 50% in hybridoma cells (36,38).

Although a glucose limitation alone did not cause a change in ammonia secretion, Ljunggren and Haggstrom (38) observed an enhancement in ammonia reduction when glucose *and* glutamine were limited. The replacement of glucose by fructose, mannose, and galactose virtually eliminated the ammonia-induced generation of UDP-GNAc.

Altering the amino acid composition of the media seems to affect ammonia generation. Hiller at al. (138) increased the concentrations of leucine, isoleucine, valine, and lysine and observed a decrease in ammonia secretion in a hybridoma line. In these experiments, alanine production rates increased, indicating an elevation in the activity of alanine transaminase.

Replacement of Glutamine. The replacement of glutamine in cell culture media can eliminate most of the ammonia generation in culture. Although the idea of replacement is a good one, it is difficult to implement in most cases because the cells seem to be strongly dependent on glutamine for energy and biomass production. Cells need to adapted or genetically altered to grow in the absence of supplemented glutamine.

Mammalian cells can use other amino acids as a substitute for glutamine. Studies were conducted to investigate the replacement of glutamine by glutamate, α-ketoglutarate, and asparagine. The efficiency of glutamate for supporting cell growth is very low, and high concentrations of glutamate (up to 20 mM) are required (139,140). The replacement of glutamine by glutamate was possible

for mouse LS cells (141) and for the McCoy cells (142). On the other hand, MDCK cells could not be adapted to glutamine-free media (143). Cells can convert glutamate to glutamine via the glutamine synthetase reaction, and the success of growing the cells on glutamate can depend on the concentration and activity of this enzyme. However, McDermott and Butler (143) observed that the key factor in cell adaptation to glutamine-free media is not the glutamine synthetase, but the uptake rate of glutamate by the cells.

Asparagine is another amino acid used to replace glutamine. Although asparagine is also unstable in media (25), the degradation rate is much lower (half-life, 87 days). Kurano et al. (85) observed the growth of CHO cells on asparagine after an initial lag phase. The asparagine-containing cultures generated 40% less ammonia than cultures in standard media. The use of α-ketoglutarate instead of glutamine was also effective in reducing ammonia generation (142).

The success of growing cells in the absence of glutamine can be increased by genetic engineering. Scientists at Celltech Ltd. developed a vector containing glutamine synthetase and infected NS0 myeloma and CHO cells with plasmid containing this vector (144–146). The construction of this vector is presented in Figure 8. The cells use glutamate as a substrate in glutamine-free media. The genes for protein expression were also integrated into the vector; thus, the glutamine synthetase gene was used as an amplifiable, selectable marker; in the glutamine-free media only the cells containing plasmid could grow. This resulted in the selection of high-producing clones. The system works best for NS0 myeloma cells because these cells lack any endogenous glutamine synthetase activity. CHO cells, on the other hand, contain endogenous glutamine synthetase genes, and the selective pressure induced by the absence of glutamine does not work effectively. The use of a specific inhibitor, methionine sulphoximine (MSX), increases the efficiency of selection and amplification.

Adaptation of Cells to High Ammonia. The adaptation of cells to grow at high ammonia concentrations can result in more ammonia-tolerant cultures. Adaptation is a complex process, and it is not clear whether the cells alter themselves or a particular clone is selected as a result of this process (147,148). Regardless of the mechanism, the cells can tolerate higher ammonia levels after the adaptation, and this method of ammonia adaptation can be used to minimize the ammonia inhibition. The adaptation of hybridoma cells to ammonia has been demonstrated by several investigators (52,54,149).

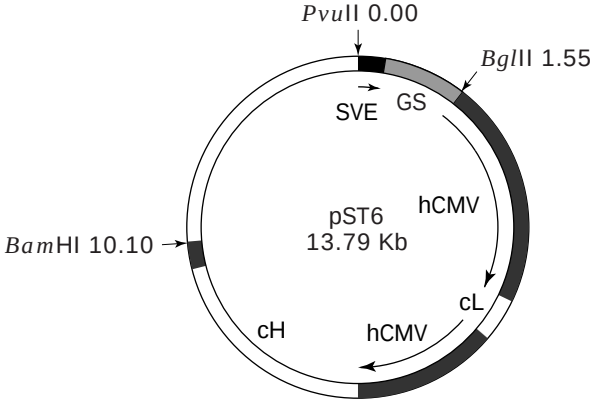


Figure 8. Expression plasmid for cB72.3 antibody containing glutamine synthetase. Expression vectors: GS, glutamine synthetase; cH, heavy chain; cL, light chain; hCMV, CMV promoter. *Source*: From Bebbington et al. (144).

Removal of Ammonia

Several techniques were investigated to remove ammonia during cell culture and to improve culture performance. These techniques involve the use of adsorbents, gas exchange and ion exchange membranes, and electrodialysis. Although the systems can be placed in the bioreactor, they are normally used as a loop system (Fig. 9). Cell culture fluid is recycled through a column, where ammonia is removed. In the case of adsorbents, the recycle is halted routinely to allow the regeneration of the column. The column is regenerated by stripping the ammonia from the adsorbents, using a stripping solution. The methods using gas and ion exchange membranes and electrodialysis, on the other hand, can be run continuously. In these cases, the ammonia fixing–stripping solution is continuously recycled on the other side of the membrane.

Use of Adsorbents. Several resins and natural adsorbents were investigated to selectively remove ammonia from cell culture media. Carbonell et al. (150) and Capiaumont et al. (151) used natural adsorbents such as Clinoptilolite to adsorb ammonia in a column where cells culture media is cycled through. The operation of the adsorbent had to be interrupted periodically to regenerate the column. Although these columns removed ammonia from the culture, the cell culture performance did not change significantly. Similar results were obtained when Zeolite (Phillipsite-Gismondine) and ion exchange resins were used as adsorbents (77). On the other hand, Nayve et al. (152,153) and Matsumura and Nayve (154) obtained good results in hybridoma cultures by combining the adsorption method and other removal systems. In fact, Matsumura and Nayve (154) obtained a viable cell concentration of 25 million cells/mL in a perfusion system with ammonia removal.

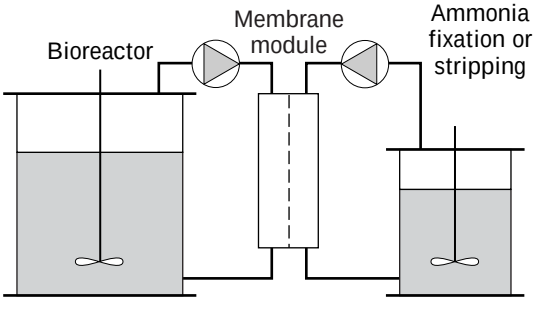


Figure 9. Ammonia removal from cell culture using adsorbents, gas or ion exchange membranes, or electrodialysis.

The use of adsorbents to remove ammonia requires complicated systems. The regeneration of the column prohibits continuous operation and requires complicated process control. These systems have to be proven to be reliable and effective for commercial scale operation.

Gas Exchange Membranes. Another method of ammonia removal from cell culture uses gas exchange membranes (such as polypropylene) where ammonia is stripped from the solution because of its gaseous properties (155–158). These systems were first developed for microbial systems and for chemical processes (159–162). The efficiency of ammonia removal in these systems depends on the concentration gradient for gaseous ammonia across the membrane (Fig. 10). The ammonia removal rate (R) is given by:

$$R = k \cdot A \cdot ([NH_3]_c - [NH_3]_s) \quad (10)$$

where k is the mass transfer coefficient, A is the surface area, and $[NH_3]_c$ and $[NH_3]_s$ are the concentrations of gaseous ammonia, in the cell culture and receiving side, respectively. The concentration gradient can be maximized by maintaining a very low gaseous ammonia concentration on the receiving side of the membrane ($[NH_3]_s \rightarrow 0$). In most cases, the transported ammonia is irreversibly converted by either enzymatic or chemical reactions (e.g., protonization). On the cell culture side, the concentration of gaseous ammonia ($[NH_3]_c$) is determined by the *total* ammonia concentration. Gaseous ammonia is at equilibrium with ammonium ion, and the culture pH determines the concentration of ammonia in gaseous form:

$$[NH_3]_c = [NH_4^+]_T/(1 + 10^{(pK_a - pH)}) \quad (11)$$

where $[NH_4^+]_T$ is the total ammonia concentration (ionized and gas) and pK_a is the ionization constant. At typical culture conditions (pH below 7.5), gaseous ammonia constitutes less than 1% of the total ammonia concentration (pK_a = 9.2 at 37 °C). Thus, only a minute fraction of ammonia is available for its removal. The efficiency of ammonia removal can be increased by increasing the pH on the cell culture side. Higher pH shifts the equilibrium to ammonia gas and increases the concentration gradient. However, this method cannot be utilized fully because the cells require a tight pH range for growth and maintenance. Increasing pH also affects the media components, and high pH can precipitate proteins, including the product.

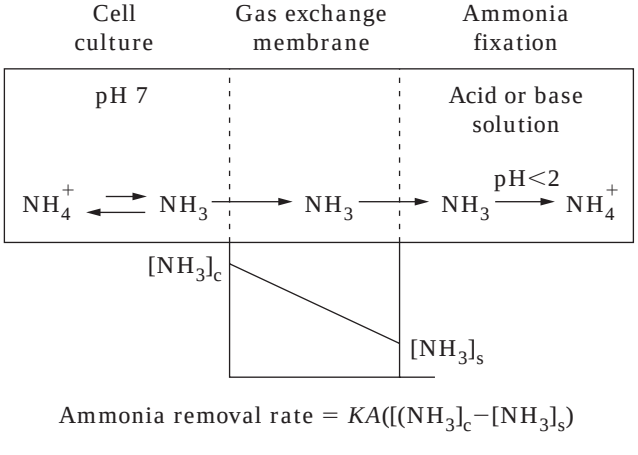


Figure 10. Ammonia removal from cell culture using gas exchange membranes. The rate of removal depends on the ammonia gas concentration difference.

The removal of ammonia using gas exchange membranes was tested in a variety of hybridoma lines (7,81,151,163) in an effort to increase antibody yield. Although the final cell density was reported to have increased, the effect on production was not significant. Ammonia removal, however, did result in an alteration in cell metabolism.

Ion Exchange Membranes. Ammonia can be selectively removed by ion (cation) exchange membranes. Contrary to gas exchange membranes, ion exchange membranes remove ammonium ion instead of gaseous ammonia. As mentioned before, ammonia in the cell culture is almost 99% in ammonium ion form, so the removal via ion exchange membranes should be more effective.

The use of autoclavable cation exchange membranes (DuPont, Wilmington, Del., Nafion 417) was described by Thommes et al. (164). Ammonium ions bind to the fixed anions in the membrane and pass into the strip–fixation solution, where they are deprotonized or stripped from the solution by pervaporation (Fig. 11). The elimination of ammonia on the strip–fixation side and a high concentration of ammonium ions on the cell culture side establishes an efficient concentration gradient so the removal of ammonia is very effective. Thommes et al. (164) reported an increase in cell density and antibody production in a murine hybridoma culture. The problem with the ion exchange membrane is the simultaneous removal of some other cations from the culture. Some cations such as calcium and magnesium can also precipitate as bicarbonates as a result of ion exchange (56).

Electrodialysis. Another technique for ammonia removal involves the use of an electrokinetic mechanism utilizing electrophoresis. Chang and Wang (82,165) applied a continuous DC electrical field to remove charged ammonia and lactate from the culture of ATCC CRL 1606 hybridoma cells. At a current density of 50 A/m^2, almost all ammonia in the culture could be removed. This ammonia removal system allowed the use of 4× concentrated media. The system allowed enhancement of cell density and antibody production significantly. The applied current did not cause any detrimental effect on the cells. Removal of ammonia increased the glutamine consumption rate. The electrodialysis system also removed lactate and minimized lactate related inhibition.

CONCLUSIONS

In this article, we reviewed several aspects of ammonia inhibition. Ammonia accumulation in cell cultures can be a serious problem for cell growth, productivity, and product quality.

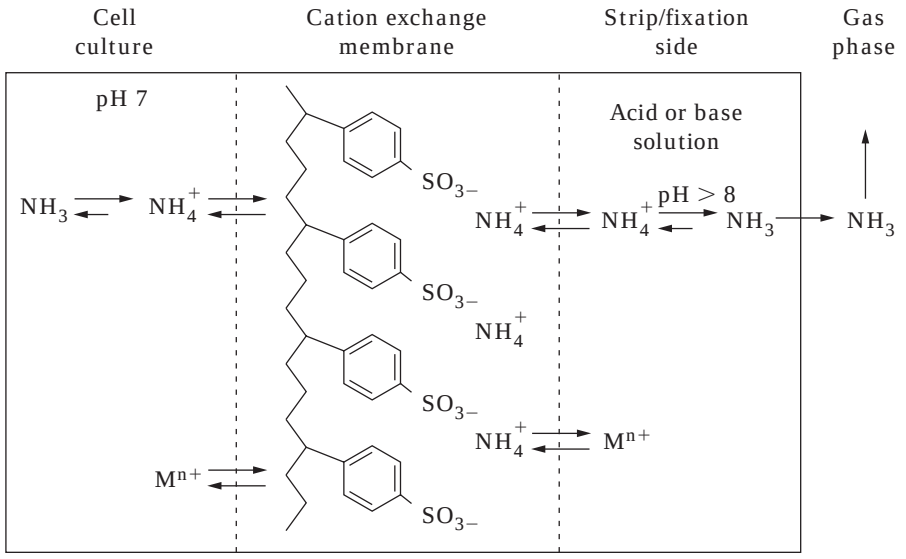


Figure 11. Ammonia removal from cell culture using cation exchange membranes. *Source*: Adapted from Thommes et al. (164).

Ammonia is generated mainly by glutamine degradation and glutamine metabolism. Cell physiology can be affected by ammonia concentrations between 2 and 10 mM. Different cell lines display varying tolerance to ammonia inhibition. Cell growth rate and final cell density in batch and fed-batch cultures decrease at high ammonia levels. Under normal conditions, ammonia does not *directly* influence the death rate or the decrease in viability of the cultures.

Ammonia is involved in several metabolic pathways and can influence the metabolic rates. Specific metabolic rates for glucose, glutamine, ammonia, lactate, amino acids are accelerated at higher ammonia concentrations. The ammonia yield from glutamine decreases, and more alanine is produced at higher ammonia levels. Although ammonia does not seem to alter the specific oxygen consumption rate, the ATP production rate increases at elevated ammonia concentrations. Cells become more active metabolically and generate more ATP for maintenance when their growth is suppressed by ammonia.

Specific rates for product secretion are not influenced by ammonia concentration. Cultures at higher ammonia levels result in lower product concentrations because of the low cell concentrations achieved. Ammonia can influence product quality, glycosylation, and sialyation; this was demonstrated for a number of cell lines and products.

Several mechanisms have been hypothesized and examined experimentally in an attempt to understand ammonia inhibition. Alteration of intracompartmental pH and membrane potential, futile cycles, metabolic inhibition, and alteration of critical ribonucleotides are identified as potential mechanisms for ammonia inhibition. Although the dynamics of ammonia transport to and from the cells were studied and intracellular pHs were measured, the data on metabolic rates could not be explained by internal pH hypothesis. Energy dissipation caused by futile cycles and disturbance of membrane potentials are viable explanations for ammonia inhibition. Ammonia was shown to elevate levels in several ribonucleotide pools. Ammonia increases the size of UDP-GNAc pool and results in growth inhibition.

Several techniques were investigated to minimize ammonia accumulation in cell culture. Generation of ammonia resulting from glutamine degradation and metabolism can be reduced by controlling the environment of the cells. By maintaining glutamine at low levels, the ammonia generation can be minimized. Glutamine can also be replaced by several other amino acids; however, the cells have to be adapted or genetically altered for this strategy to be successful. Cells can also be adapted to high ammonia concentrations and can thus tolerate higher ammonia levels.

Methods for ammonia removal in cell culture were investigated to improve culture performance. These techniques involve use of adsorbents, gas exchange and ion exchange membranes, and electrodialysis. These systems are fairly complicated, and more work has to be done to improve the effectiveness and reliability for commercial production.

BIBLIOGRAPHY

1. M.P. Backer, L.S. Metzger, P.L. Slaber, K.L. Nevitt, and G.B. Boder, *Biotechnol. Bioeng.* **32**, 993–1000 (1988).
2. C.H. Leist, H.-P. Meyer, and A. Fiechter, *J. Biotechnol.* **15**, 1–46 (1990).
3. A.S. Lubiniecki, *Large Scale Mammalian Cell Culture Technology*, Dekker, New York, 1990.
4. S.S. Ozturk, *Cytotechnology* **23**, 1–20 (1996).
5. K.W. Lanks and P.-W. Li, *J. Cell. Physiol.* **135**, 151–155 (1986).
6. S.S. Ozturk, M.R. Riley and B.O. Palsson, *Biotechnol. Bioeng.* **39**, 418–431 (1992).
7. M. Schneider, M. El Alaoui, U. Von Stockar, and I.W. Marison, *Enzyme Microb. Technol.* **20**, 268–276 (1997).
8. J. Crawford and H.J. Cohen, *J. Cell. Physiol.* **124**, 275–282 (1985).
9. H.R. Zielke, C.L. Zielke, and P.T. Ozand. *Fed. Proc.* **43**, 121–125 (1984).

10. M. Dalili, G.D. Sayles, and D.F. Ollis, *Biotechnol. Bioeng.* **36**, 74–82 (1990).
11. C. Dyring, H.A. Hansen, and C. Emborg, *Cytotechnology* **16**, 37–42 (1994).
12. D. Duval, C. Demangel, K. Munier-Jolain, S. Miossec, and I. Geahel, *Biotechnol. Bioeng.* **38**, 561–570 (1991).
13. O.T. Ramirez and R. Mutharasan, *Biotechnol. Bioeng.* **36**, 839–848 (1990).
14. M.C. Flickinger, N.K. Goebel, T. Bibila, and S. Boyce-Jacino, *J. Biotechnol.* **22**, 201–226 (1992).
15. T. Omasa, M. Ishimoto, K. Higashiyama, S. Shioya, and K. Suga, *Cytotechnology* **8**, 75–84 (1992).
16. S. Heeneman, N.E.P. Deutz, and W.A. Buurman, *J. Immunolog. Methods* **166**, 85–91 (1993).
17. R.I. Freshney, *Culture of Animal Cells: A Manual of Basic Technique*, 2nd ed., Liss, New York, 1987.
18. H.G. Bray, S.P. James, I.M. Raffan, and W.V. Thorpe, *Biochem. J.* **44**, 625–627 (1949).
19. J.B. Gilbert, V.E. Price, and J.P. Greenstein, *J. Biol. Chem.* **180**, 209–218 (1949).
20. G.L. Tritsch and G.E. Moore, *Exp. Cell. Res.* **28**, 360–364 (1962).
21. A.A. Lin, P. Agrawal, *Biotechnol. Lett.* **10**, 695–698 (1988).
22. S.S. Ozturk and B.O. Palsson, *Biotechnol. Prog.* **6**, 121–128 (1990).
23. S.S. Ozturk, M.E. Meyerhoff, and B.O. Palsson, *Biotechnol. Tech.* **3**, 217–222 (1989).
24. S.S. Seaver, J.L. Rudolph, and J.E. Gabriels Jr., *Biotechnology* **2**, 254–260 (1984).
25. J. Wein and I.E. Goetz, *In Vitro* **9**, 186–193 (1973).
26. M.S.M. Ardavi and E.A. Newsholme, *Biochem. J.* **212**, 835–842 (1993).
27. L.G. Baggetto, *Biochimie* **74**, 959–974 (1992).
28. K. Brand, J. von Hintzenstarn, K. Langer, and W. Fekl, *J. Cell Physiol.* **132**, 559–564 (1987).
29. K. Brand, W. Fekl, and J. von Hintzenstem, *Metabolism* **38**, 29–33 (1989).
30. M. Butler and H. Jenkins, *J. Biotechnol.* **12**, 97–110 (1989).
31. M. Donnely and I.E. Scheffler, *J. Cell Physiol.* **89**, 39–52 (1976).
32. L. Fitzpatrick, H.A. Jenkins, and M. Butler, *Appl. Biochem. Biotechnol.* **43**, 93–116 (1993).
33. H.A. Jenkins, M. Butler, and A.J. Dickson, *J. Biotechnol.* **23**, 167–182 (1992).
34. Z. Kovacevic and J.D. McGivan, *Physiol. Rev.* **63**, 547–605 (1983).
35. K.W. Lanks and P.-W. Li, *J. Cell Physiol.* **135**, 151–155 (1988).
36. J. Ljunggren and L. Haggstrom, *Biotechnol. Lett.* **12**, 705–710 (1990).
37. J. Ljunggren and L. Haggstrom, *Cytotechnology* **8**, 45–56 (1992).
38. J. Ljunggren and L. Haggstrom, *Biotechnol. Bioeng.* **44**, 808–818 (1994).
39. W.L. McKeehan, in M.J. Morgan ed., *Carbohydrate Metabolism in Cultured Animal Cells*. Plenum, New York, 1986, pp. 111–150.
40. R.W. Moreadith and A.L. Lehninger, *J. Biol. Chem.* **259**, 6215–6221 (1984).
41. M. Newland, P.F. Greenfield, and S. Reid, *Cytotechnology* **3**, 215–229 (1990).
42. L.J. Reitzer, PhD Thesis, Washington University, St. Louis, Mo., 1978.
43. L.J. Reitzer, B.M. Wice, D. Kennell, *J. Biol. Chem.* **254**, 2669–2676 (1979).
44. L.J. Reitzer, B.M. Wice, D. Kennell, *J. Biol. Chem.* **255**, 5616–5626 (1980).
45. S.T. Sharfstein, S.N. Tucker, A. Manusco, H.W. Blanch, and D.S. Clark, *Biotechnol. Bioeng.* **43**, 1059–1074 (1994).
46. M. Watford, *Biochim. Biophys. Acta* **1200**, 73–78 (1994).
47. E. Kvamme ed., *Glutamine and Glutamate in Mammals*, Vol. 1. CRC, Boca Raton, Fla., 1988.
48. J. Swierczynski, Z. Bereznowski, and W. Makarewicz, *Biochim. Biophys. Acta* **1157**, 55–62 (1993).
49. M.W. Glacken, *Biotechnology.* **6**, 941–1050 (1988).
50. M.W. Glacken, R.J. Fleischaker, and A.J. Sinskey, *Biotechnol. Bioeng.* **28**, 1376–1389 (1986).
51. P.M. Hayter, E.M.A. Curling, A.J. Baines, N. Jenkins, L. Salmon, P.G. Strange, and A.T. Bull, *Appl. Microbiol. Biotechnol.* **34**, 559–564 (1991).
52. W.M. Miller, C.R. Wilke, and H.W. Blanch, *Bioproc. Eng.* **3**, 113–122 (1988).
53. H. Buntemeyer, C. Wallerius, and J. Lehmann, *Cytotechnology* **9**, 59–67 (1992).
54. B.L. Maiorella, A. Shauger, M. Poulhazan, B. Howarth, and D. Inlow, *203th ACS Natl. Mtg.*, 1992.
55. L. Xie and D.I.C. Wang, *Biotechnol. Bioeng.* **43**, 1175–1189 (1994).
56. M. Schneider, I.W. Marison, and U. Von Stockar, *J. Biotechnol.* **46**, 161–181 (1996).
57. W.M. Canning and B.N. Fields, *Science* **219**, 987–988 (1983).
58. M.D. Eaton and A.R. Scala, *Virology* **13**, 300–307 (1961).
59. E. Furusawa and W. Cutting, *Proc. Soc. Exp. Biol. Med.* **111**, 71–75 (1962).
60. G. Farias, E. Navarrete, J. Kiss, and J. Kuznar, *Arch. Virol.* **98**, 155–162 (1988).
61. E.M. Jensen and O.C. Liu, *Proc. Soc. Exp. Biol. Med.* **107**, 834–838 (1961).
62. A.H. Koyama and T. Uchida, *Virus Res.* **13**, 271–282 (1989).
63. C. Scholtissek, R. Rott, and H.-D. Klenk, *Virology* **63**, 191–200 (1975).
64. W.J. Visek, G.M. Konoldny, and P.R. Gross, *J. Cell. Physiol.* **80**, 373–382 (1980).
65. R.W. Holley, R. Armour, and J.H. Baldwin, *Proc. Natl. Acad. Sci. USA* **75**, 1864–1866 (1978).
66. T. Hassell, S. Gleave, and M. Butler, *Appl. Biochem. Biotechnol.* **30**, 29–41 (1991).
67. G.A. Truskey, D.P. Nicolakis, D. Dimasi, A. Haberman, and R.W. Swartz, *Biotechnol. Bioeng.* **36**, 797–812 (1990).
68. T.C. Dodge, G.-Y. Ji, and W.-S. Hu, *Enzyme Microb. Technol.* **9**, 607–611 (1987).
69. J.G. Gaertner and P. Dhurjati, *Biotechnol. Prog.* **9**, 309–316 (1993).
70. S. Reuveny, D. Velez, L. Miller, and J.D. Macmillan, *J. Immunol. Methods* **86**, 53–59 (1986).
71. A. McQueen and E. Bailey, *Biotechnol. Bioeng.* **35**, 1067–1077 (1990).
72. A. McQueen and E. Bailey, *Biotechnol. Bioeng.* **35**, 897–906 (1990).
73. D. Doyle and M. Butler, *J. Biotechnol.* **15**, 91–100 (1990).
74. A.J. Racher, D. Looby, and J.B. Griffiths, *J. Biotechnol.* **15**, 129–146 (1990).
75. A.J. Racher, D. Looby, and J.B. Griffiths, *J. Biotechnol.*, **29**, 145–156 (1993).
76. O.W. Ronning, M. Schartum, A. Winsnes, and G. Lindberg, *Cytotechnology* **7**, 15–24 (1991).

77. Y-H. Jeong and S.S. Wang, *Biotechnol. Tech.* **6**, 341–346 (1992).
78. Y.-H. Jeong and S.S. Wang, *Enzyme Microb. Technol.* **17**, 47–55 (1995).
79. I. Ludemann, R. Portner, and H. Markl, *Cytotechnology* **14**, 11–20 (1994).
80. M. Newland, M.N. Kamal, P.F. Greenfield, and L.K. Neilsen, *Biotechnol. Bioeng.* **43**, 434–438 (1994).
81. M. Schneider, PhD Thesis No. 1412, Ecole Polytechnique Federal de Lausanne (EPFL), Switzerland, 1995.
82. Y.H.D. Chang, A.J. Grodzinsky, and D.I.C. Wang, *Biotech. Bioeng.* **47**, 308–318 (1995).
83. D. Wentz and K. Schügerl, *Enzyme Microb. Technol.* **14**, 68–75 (1992).
84. M. Butler and R.E. Spier, *J. Biotechnol.* **1**, 187–196 (1984).
85. N. Kurano, C. Leist, F. Messi, S. Kurano, and A. Fiechter, *J. Biotechnol.* **15**, 113–128 (1990).
86. H. Hansen and C. Emborg, *Biotechnol. Prog.* **10**, 121–124 (1994).
87. M.W. Glacken, Ph.D. Thesis, MIT, Cambridge, Mass., 1987.
88. S.S. Ozturk and B.O. Palsson, *Biotechnol. Bioeng.* **37**, 989–993 (1991).
89. J.L. Goergen, A. Marc, and J.M. Engasser, in R.E. Spier, J.B. Griffiths, W. Berthold eds. *Animal Cell Technology: Products of Today, Prospects of Tomorrow*, Butterworth-Heinemann, London, 1994, pp 161–164.
90. P.C. Alex, D. Barngrover, and R.W. Swartz, *ACS Natl. Mtg.*, Miami, Fla., 1989.
91. A. McQueen and J. Bailey, *Biotechnol. Bioeng.* **35**, 1067–1077 (1989).
92. S.S. Ozturk and B.O. Palsson, *Biotechnol. Prog.* **7**, 471–480 (1991).
93. S.S. Ozturk and B.O. Palsson, *Biotechnol. Prog.* **7**, 481–494 (1991).
94. T. Kimura, S. Iijima, and T. Kobayashi, *J. Ferment. Technol.* **65**, 341–344 (1987).
95. J.J. Dyken and A. Sambanis, *Enzyme Microb. Technol.* **16**, 90–98 (1994).
96. C.L. Atanassov, C.D. Muller, S. Sarhan, B. Knodgen, G. Rebel, and N. Seiler, *Res. Immunol.* **145**, 277–288 (1994).
97. M. Ito and W.F. McLimans, *Cell. Biol. Int. Rep.* **5**, 661–666 (1981).
98. B. Thorens and P. Vassalli, *Nature* **321**, 618–620 (1986).
99. D.C. Andersen and C.F. Goochee, *Biotech. Bioeng.* **47**, 96–105 (1995).
100. M.C. Borys, D.L.H. Linzer, and E.T. Papoutsakis, *Biotechnology* **11**, 720–724 (1993).
101. M. Gawlitzek, U. Valley, R. Wagner, and H.S. Conradt, *Cytotechnology* **14**, 3–5 (1994).
102. M. Gawlitzek, U. Valley, M. Nimtz, R. Wagner, and H.S. Conradt, in E.C. Beuvery, J.B. Griffiths, P.W. Zeijlemaker eds., *Animal Cell Technology: Developments Towards the 21st Century*. Kluwer, Norwell, Mass., 1995, pp. 379–384.
103. K. Kopp, W. Noe, M. Schlueter, W. Werner, and F. Goetz, in E.C. Beuvery, J.B. Griffiths, P.W. Zeijlemaker eds., *Animal Cell Technology: Developments Towards the 21st Century* Kluwer, Norwell, Mass., 1995, pp 403–407.
104. U.S. Pat. 5,156,964 (1992) Cetus Corp.
105. T.S. Stoll, M.B. Sliwkoeski, and Jacobson, *ACS Mtg.*, San Francisco, April 1977.
106. Y. Fukami, *Brain Res.* **463**, 140–143 (1988).
107. A. McQueen and J.E. Bailey, *Bioproc. Eng.* **6**, 49–61 (1991).
108. W.F. Boron and P. De Weer, *J. Gen. Physiol.* **67**, 91–112 (1976).
109. M.A. Knepper, R. Packer, and D.W. Good, *Physiol. Rev.* **69**, 179–249 (1989).
110. D. Kikeri, A. Sun, M.L. Zeidel, and S.C. Heben, *Nature*, **339**, 478–480 (1989).
111. R.M. Lynch and R.S. Balaban, *Am J. Physiol.* **252**, C225–C231 (1987).
112. R.M. Lynch and R.S. Balaban, *Am. J. Physiol.* **253**, C269–C276 (1987).
113. W.H. Moolenaar, R.Y. Tsien, P.T. van der Saag, and S.W. de Laat, *Nature* **304**, 645–648 (1983).
114. R.L. Post and P.C. Jolly, *Biochim. Biophys. Acta* **25**, 118–128 (1957).
115. M. Martinelle and L. Haggsrom, in J.B. Spier, C. Griffiths, C. MacDonald eds., *Animal Cell Technology: Developments, Processes, and Products*, Butterworth-Heinemann, London, 1992, p. 163.
116. K. Martinelle and L. Haggstrom, *J. Biotechnol.* **30**, 339–350 (1993).
117. K. Martinelle and L. Haggstrom, *Cytotechnology* **14**, 2–17 (1994).
118. R.G.W. Anderson and R.K. Pathak, *Cell*, **40**, 635–643 (1985).
119. D.D. Wagner, T. Mayadas, and V.J. Marder, *J. Cell Biol.* **102**, 1320–1324 (1986).
120. B. Poole and S. Ohkuma, *J. Cell. Biol.* **90**, 665–669 (1981).
121. F. Van Leuven, J.-J. Cassiman, and H. Van Den Berghe, *Cell* **20**, 37 (1980).
122. S.I. Harris, R.S. Balban, L. Barrett, and L.J. Mandel, *J. Biol. Chem.* **256**, 319–328 (1981).
123. S.M. Wall and L.M. Koger, *Am. J. Physiol. Ren. Fluid Electrol. Physiol.* **36**, F660–F670 (1994).
124. K. Uyeda and E. Racker, *J. Biol. Chem.* **240**, 4682–4688 (1965).
125. A. Parmeggiani, J.H. Luft, D.S. Love, and E.G. Krebs, *J. Biol. Chem.* **241**, 4625–4637 (1966).
126. J.C. Street, A.-M. Delort, P.S. Braddock, and K.M. Brindle, *Biochem. J.* **291**, 485–492 (1993).
127. J.M. Tagler, T.P.M. Akerboom, J.B. Hoek, A.J. Meijer, W. Vaartjes, L. Ernster, and J.R. Williamson, in F.A. Hommes, C.J. Van der Berg, eds., *Normal and Pathological Development of Energy Metabolism*. Academic Press, New York, 1975, pp. 63–75.
128. T. Ryll and R. Wagner, *Biotechnol. Bioeng.* **40**, 934–946 (1992).
129. T. Ryll, U. Valley, and R. Wagner, *Biotechnol. Bioeng.* **44**, 184–193 (1994).
130. I.G. Bekesi, E. Bekesi, and R.J. Einzler, *J. Biol. Chem.* **244**, 3766–3772 (1969).
131. J.G. Bekesi and R.J. Winzler, *Cancer Res.* **30**, 2905–2912 (1970).
132. I.-N. Chou, J. Zeigher, and E. Rapaport, *Proc. Natl. Acad. Sci. USA* **81**, 2401–2405 (1984).
133. E. Roth, G. Ollenschlager, A. Hamilton, K. Langer, W. Fekl, and R. Jakesz, *In Vitro Cell Dev. Biol.* **24**, 696–698 (1988).
134. A. Christie and M. Butler, in J.B. Spier, C. Griffiths, W. Berthold, eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, Butterworth-Heinemann, London, 1994, p. 164.
135. A.-C. Holmlund, S.L. Chatzisavido, S.L. Bell, and E. Lindner-Olsson, *Proc. of the 11th ESACT Mtg.*, Brighton, UK, 1992, pp. 176–179.

136. W. Miller, C. Wilke, and H. Blanch, *J. Cell Physiol.* **132**, 524–530 (1987).
137. S.S. Ozturk and B.O. Palsson, *Biotechnol. Prog.* **6**, 437–446 (1990).
138. G.W. Hiller, D.S. Clark, and H.W. Blanch, *Biotechnol. Bioeng.* **44**, 303–321 (1994).
139. J.E. Darnell Jr. and H. Eagle, *Virology* **6**, 556–566 (1958).
140. S.C. Nagle and B.L. Brown, *J. Cell. Physiol.* **77**, 259–264 (1971).
141. J.B. Griffiths and S.J. Pirt, *Proc. R. Soc. Lond., Ser. B* **168**, 421–438 (1967).
142. T. Hassell and M. Butler, *J. Cell Sci.* **96**, 501–508 (1990).
143. R.C. McDermott and M. Butler, *J. Cell Sci.* **104**, 51–58 (1993).
144. C.R. Bebbington, G. Renner, S. Thomson, D. King, D. Abrams, and G.T. Yarranton, *Biotechnology* **10**, 169–175 (1992).
145. S.L. Bell, M.E. Bushell, M.F. Scott, J.N. Wardell, R.E. Spier, and P.G. Sanders, *Proc. of the 11th ESACT Mtg.* Brighton, UK, 1992, pp. 180–182.
146. S.L. Bell, C. Bebbington, M.F. Scott, J.N. Wardell, R.E. Spier, M.E. Bushell, and P.G. Sanders, *Enzyme Microb. Technol.* **17**, 98–106 (1995).
147. S.S. Ozturk and B.O. Palsson, *Hybridoma* **9**, 167–175 (1990).
148. S.S. Ozturk and B.O. Palsson, *Biotechnology Bioeng.* **37**, 35–46 (1991).
149. M. Matsumura, M. Shimoda, T. Arii, and H. Kataoka, *Cytotechnology* **7**, 103–112 (1991).
150. D. Carbonell, B. Besnainou, J. Capiaumont, P. Lessart, and P. Nabet, in J.B. Spier, C. Griffiths, C. MacDonald, eds., *Animal Cell Technology: Developments, Processes, and Products*, Butterworth-Heinemann, London, 1992, p. 166.
151. J. Capiaumont, C. Legrand, D. Carbonell, B. Dousset, F. Belleville, and P. Nabet, *J. Biotechnol.* **39**, 49–58 (1995).
152. F.R.P. Nayve Jr., M. Motoki, M. Matsumura, and H. Kataoka, *Cytotechnology* **6**, 121–130 (1991).
153. F.R.P. Nayve Jr., T. Misato, M. Matsumura, and H. Kataoka, *J. Biotechnol.* **34**, 217–225 (1994).
154. M. Matsumura and F.R.P. Nayve Jr., *Cytotechnology* **18**, 35–50 (1995).
155. European Pat. Appl. No. 90301630.1 (1990), Bend Research Inc.
156. U.S. Pat. 5,071,561 (1991), Bend Research Inc.
157. V. Blet, M.-N. Pons, and J.L. Greffe, *Anal. Chim. Acta* **219**, 309–311 (1989).
158. D.J. Brose and P. Van Eikeren, *Appl. Biochem. Biotech.* **24/25**, 457–458 (1990).
159. V. Hecht, L. Bischoff, and K. Gerth, *Biotechnol. Bioeng.* **35**, 1042–1050 (1990).
160. M. Imai, S. Furusaki, and T. Miyauchi, *Ind. Eng. Chem. Proc. Des. Dev.* **21**, 421–426 (1982).
161. L. Qin and J.M.S. Cabral, *J. Chem. Technol. Biotechnol.* **67**, 323–328 (1996).
162. M.J. Semmens, D.M. Foster, and E.L. Cussler, *J. Membr. Sci.* **51**, 127–140 (1990).
163. M. Schneider, I.W. Marison, and U. Von Stockar, *Enzyme Microb. Technol.* **16**, 957–963 (1994).
164. J. Thommes, U. Garske, M. Biselli, and W. Wandrey, in J.B. Spier, C. Griffiths, MacDonald, eds., *Animal Cell Technology: Developments, Processes, and Products*, Butterworth-Heinemann, London, 1992, p. 171.
165. Y.-H.D. Chang, A.J. Grodzinsky, and D.I.C. Wang, *Biotechnol. Bioeng.* **47**, 319–326 (1995).

See also CULTURE MEDIA, ANIMAL CELL, LARGE SCALE PRODUCTION; ENERGY METABOLISM, MICROBIAL AND ANIMAL CELLS; FERMENTATION MONITORING, DESIGN AND OPTIMIZATION; INSECT CELL CULTURE, PROTEIN EXPRESSION; PROCESS CONTROL, STRATEGY AND OPTIMIZATION.

ANAEROBES

LARRY E. ERICKSON
DANIEL Y.C. FUNG
Kansas State University
Manhattan, Kansas

KEY WORDS

Anaerobes
Biodegradation
Bioenergetics
Cultivation
Mixed cultures
Products

OUTLINE

INTRODUCTION

There are both aerobic and anaerobic microorganisms that are important to the advancement of biotechnology. Anaerobic microorganisms include obligate anaerobes that require oxidation–reduction potentials of −150 to −420 mV and facultative anaerobes that can grow at oxidation–reduction potentials between +300 and −420 mV (1). Oxygen must be excluded in working with obligate anaerobes.

Many important fermentation products are produced under anaerobic conditions (1–69). These include food products such as bread, yogurt, cheeses, wine, beer, sufu, and sauerkraut. Acetone, butanol, and ethanol are examples of industrial chemicals produced by anaerobes. Specialty chemicals produced by anaerobes include vitamins and pharmaceutical products. Major references describing the microbial products and the microorganisms that produce them include those by Erickson and Fung (2), Reed (3), Steinkraus (4), and Zeikus and Johnson (20).

Anaerobic digestion and biodegradation processes are applied widely to treat process waste products and in environmental restoration. Environmental microbiology is significant because of the beneficial impact of natural and nurtured processes and because of the isolation and identification of useful microorganisms that have evolved in natural environments. Anaerobic processes occur in production agriculture; cattle and sheep are still the greatest commercial success in harvesting and utilizing cellulosic plant products because of the anaerobic fermentations that occur in the rumen of these animals.

The taxonomy of anaerobes has evolved as new information has become available (64–66). Based on 16/18 S rRNA sequence comparisons, a universal phylogenetic tree has been described that includes bacteria, archaea, and eukarya (64,65). Anaerobes are found in all three branches of this tree.

METHODS OF CULTIVATION

The microbial world consists of aerobic and anaerobic organisms, defined by the ability to tolerate molecular oxygen. Sonnewirth (63) traced the evolution of anaerobic methodology and recorded that Leeuwenhoek demonstrated that some life forms could exist in the presence of gases other than oxygen and that Pasteur discovered anaerobiosis in 1861, coining the terms *aerobes* and *anaerobes*. The separation between these groups is not clearly defined. A convenient way to categorize these organisms is by observing their ability to grow at different oxidation–reduction potentials (E_h). E_h can be measured and expressed + or − millivolts (mV). Thus, aerobes can grow between +300 and −50 mV; facultative anaerobes, +300 to −420 mV; aerotolerant anaerobes, +180 to −350 mV; and obligate anaerobes, −150 to −420 mV (28).

The microbiological procedures for studying obligate anaerobes are similar to those used for the more familiar aerobic organisms, with the exception of the need to exclude oxygen from the working environment during manipulation, inoculation, and incubation of cultures and samples. As a class of organisms, anaerobes are very diversified in their behavior in terms of oxygen tolerance and nutritional requirements for growth. There is no single set of procedures for isolation of all anaerobes. Clinically important anaerobes constitute the most studied class. Specific methods were developed for isolating anaerobes from various sites in patients (57).

A detailed treatment of clinically important anaerobic organisms is recorded in the *Manual for the Determination of the Clinical Role of Anaerobic Microbiology* by Gall and Riely (43). This manual contains a definition of clinically important anaerobic bacteria methods for the collection and transportation of anaerobic specimens, culturing of anaerobic specimens, and identification of anaerobes.

The book *Isolation of Anaerobes*, by Shapton and Board (56), contains chapters describing the isolation of clostridia from animal tissues, feces, soil, and foods. Some chapters deal with the isolation of anaerobes from rumen as well as with sulfate-reducing bacteria and photosynthetic bacteria. Zeikus (62) describes the biology of methanogenic bacteria and methods of studying these organisms. In the area of food microbiology, very little attention has been given to the anaerobic microbiology. The *Compendium of Methods for the Microbiological Examination of Foods* (59) contains some chapters on anaerobes, mainly on the spore formers and the detection of toxins in foods. Anderson and Fung (30) reviewed the status of anaerobic methods, techniques, and principles for food bacteriology in detail. The history of cultivation of anaerobes was detailed by Hall (44) and more recently summarized by Fung (1).

The general conclusion on isolation procedures is that samples must be free from contact with air during collection time. The best method of collecting samples is by aspiration with needle and syringe. Sample should be placed in media that are designed to handle anaerobes. A variety of liquid and solid media are marketed for anaerobic cultivation by such companies as DIFCO, BBL, UNIPATH/OXOID, and others (69).

Currently, anaerobic cultivation systems include anaerobic jars, anaerobic glove boxes, roll tubes, anaerobic tube systems, and biological membrane systems.

Anaerobic Jars

Anaerobic jars that use both H_2 and CO_2 were developed by Brewer and Allgeier (33). The chemical generator consisted of a sodium borohydride tablet and a citric acid–sodium bicarbonate tablet in a convenient package. These two tablets, when activated by the addition of water to the package, produce H_2 and CO_2 gas, respectively. The system required no vacuum or electric current because the H_2 + O_2 reaction was catalyzed by a cold catalyst (alumina pellets coated with 0.5% palladium). CO_2 is used in the system for the promotion of growth of some anaerobic bacteria. Some companies are marketing these anaerobic jars (e.g., BBL, B.T.L., and UNIPATH/OXOID). The advantages of the anaerobic jar include (*1*) ease of operation, especially for the small laboratory; (*2*) economy of space; (*3*) convenience of prepackaged materials for the generation of gases;

and (*4*) safety of the units. The disadvantages include (*1*) the culture plates are inoculated under aerobic conditions; (*2*) the jar remains aerobic for 15 to 20 minutes even while the Gas Paks are activated, which may affect some strict anaerobes; and (*3*) occasionally the jar fails to achieve anaerobiosis.

Anaerobic Glove Boxes

To eliminate the chance of exposure of anaerobes to the natural environment during bacteriological manipulation, the anaerobic chamber, or glove box, was developed. The basic principles for achieving anaerobiosis in the glove box and for promoting binding of residual O_2 and H_2 in the presence of a catalyst (Laidlaw principle) are to remove air first and then flush the chamber with inert gas. Reducing agents in the medium also help to reduce the E_h to a desired low level. Companies marketing anaerobic glove boxes include Format Scientific, Don Whitney Scientific, and Microflow MDH Ltd. The advantages of the glove box include (*1*) the bacteriological procedure can be operated under anaerobic conditions from the sampling stage to the incubation stage; (*2*) the area for manipulation is large; and (*3*) in combination with temperature control, the culture can be incubated in the chamber without further disturbance of the cultural environment. The disadvantages are (*1*) high initial cost; (*2*) high operating expenses (gas and electricity); (*3*) high maintenance cost; (*4*) requirement for operator training; (*5*) oxygen contamination through leakage, operator error, or mechanical malfunction; (*6*) occasional high labor requirements; and (*7*) moisture buildup in the box.

Roll Tubes

One of the most widely used anaerobic cultivation systems is the Hungate roll tube, developed by Hungate (46–48). Sterile tubes were flushed with CO_2 while the medium was aseptically added to the tubes to prevent air contamination. The tubes were then inoculated, sealed with rubber stoppers, cooled (48 °C), and hand spun so that the agar with cultures forms a thin film in the inner wall of the test tube. After incubation the colonies could be examined with ease. Isolation of colonies was also feasible with minimum effort by flushing CO_2 into the cavity to allow maintenance of reduced condition while performing microbiological operations. A popular version of the roll tube system is the VPI Anaerobic Culture System, developed by Moore (53). This apparatus has a platform so that the tubes can stand upright with cannulas positioned above each tube. The system also allows for streaking cultures into the roll tube anaerobically. The system is continuously flushed by CO_2 to exclude O_2 from the system.

The advantages of the Hungate roll tube system include (*1*) good anaerobiosis with minimum investment, (*2*) unit operation for convenience of small sample numbers; and (*3*) ease of observation of colonies and isolation of cultures for further studies. Disadvantages include (*1*) time requirement for large number of samples; (*2*) need for manual dexterity; and (*3*) occasional failure of some tubes to achieve anaerobiosis.

Anaerobiosis Tube Systems

Ogg et al. (54) designed a tube with double wells so that agar with cultures can be introduced into the tube and form a thin film. Anaerobiosis is initially achieved by boiling and autoclaving of the medium. By the action of reducing agent in the medium and minimum reabsorption of O_2 due to the design of the tube (commercially called the Lee tube), good anaerobiosis is maintained in the system. The glass cylinder forming the inner wall (the inverted smaller tube) makes it easy for enumeration and observation of colony morphologies. The Lee tube has the advantage of ease of operation and low operating cost (no gas or special equipment needed). The disadvantages include (*1*) difficulty in picking isolates from the agar; (*2*) difficulty in cleaning the tubes; and (*3*) fragility of the system.

Fung and Lee (41) developed a double-tube system for anaerobic cultivation of bacteria from foods. It consists of a glass tube (15 × 1.0 cm o.d.) placed into a larger glass tube (15 × 1.5 cm o.d.) with anaerobic agar (designed for specific organisms) and culture sandwiched between the two tubes. The screw cap of the larger tube when closed seals the system and makes it anaerobic. Anderson and Fung (29) were successful in cultivating strict anaerobic rumen microbes, such as *Megasphaera elsdenii, Butyrivibro fibrisolvens, Clostridium perfringens, Eubacterium limosum*, and anaerobes from beef and poultry. Ali et al. (31) successfully used this system to isolate *C. perfringens* from meat. Aramouni et al. (32) used the double-tube system to study *C. sporogenes* in home-style canned quick bread. The advantages of the system include (*1*) ease of operation; (*2*) inexpensive and ease for cleanup; (*3*) ease of enumeration and observation of cultures; (*4*) ease of obtaining colonies (by removing the inner tube and pick colonies); and (*5*) flexibility and ease of adaptation by other laboratories. Disadvantages include (*1*) operation of samples in the open environment and (*2*) occasional breakage of agar by gas-producing organisms.

Biological Membrane Fragments

A novel method using biological materials to achieve anaerobiosis was developed by Adler and Crow (26). Sterile suspensions of membrane fragments derived from *Escherichia coli* were able to remove oxygen rapidly and efficiently as a result of the presence of a cytochrome-based electron transport system that transfers hydrogen from suitable donors in the bacteriological media to oxygen and produces water as an end product. The membranes are stable, nontoxic, and active in both liquid and solid media. The history of the development of this concept and manufacturing of membrane fragments were described recently by Adler and Spady (27). The commercial product is named Oxyrase and marketed by Oxyrase Inc. (Mansfield, Ohio). Fung et al. (42) described the use of Oxyrase to stimulate the growth of *Listeria monocytogenes, Campylobacter jejuni, E. coli* O157:H7, etc. to high numbers (10^6 CFU/mL) so that secondary detection methods such as polymerase chain reaction (PCR), enzyme-linked immunosorbent assay, and DNA/RNA probes can detect the presence of these facultative anaerobic pathogens much faster for food safety concerns. Fung et al. (42) also detailed the use of Oxyrase

and membranes derived from *Acetobacter* and *Gluconobacter* to stimulate the growth of starter cultures to produce fermented food faster. Claerbout (34) described the applications for bacterial membrane fragments in sterile pharmaceutical quality control. Thurston and Gannon (58) made a detailed comparison between Oxyrase anaerobic agar plates and conventional anaerobic chambers for the isolation and identification of anaerobic bacteria from clinical infections and concluded that Oxyrase is a valuable compound for anaerobic cultivation of bacteria.

The future of membrane fragments for anaerobic cultivation of microorganisms from food, clinical, industrial, and environmental samples is very bright.

METHODS OF IDENTIFICATION

The identification of anaerobes is a concern for applied microbiologists. Classic methods for identification can be found in *Bergey's Manual of Systematic Bacteriology*, Volume 1 (50). A variety of rapid and automated methods in microbiology have been developed for the isolation, enumeration, and identification of microorganisms (37), and many of these new methods can be adapted to anaerobic bacteriology. Identification of anaerobes can be achieved by the following methods and procedures: conventional methods, miniaturized methods, commercial diagnostic kits, physical and biochemical methods, immunological methods, and PCR and related methods.

Conventional Methods

Identification of microorganisms using conventional procedures is summarized in Table 1 (38). For anaerobes, the most important aspect is to have all the liquid and solid media incubated in anaerobic environments, as described in the previous section. After the reactions are obtained, microbiologists will determine the identities of the cultures by a variety of schemes, charts, and flow diagrams. The conventional method is time consuming and highly dependent on the knowledge and skill of the microbiologist involved.

The morphology is one of the important considerations in identifying microorganisms by conventional methods. Under the microscope, the morphology of anaerobic microorganisms are similar to the aerobic counterparts. The morphology of bacteria are rod-shaped, coccal-shaped, and spiral-shaped organisms. An example of a sporulating anaerobic rod-shaped bacterium, *C. perfringens*, is presented in Figure 1. There are also other bacteria that are pleomorphic (variable shapes depending on growth conditions). The morphology of yeast and mold are similar to the aerobic counterparts. Morphology of an anaerobic protozoa, *Isotricha prostoma*, is presented in Figure 2.

Table 1. Needed Information for Identification of Microorganisms

Morphology under magnification and on agar plates
Gram reaction and special staining properties
Biochemical activity profile and special enzyme systems
Pigment production, bioluminescence, chemiluminescence, and fluorescent compound production
Nutritional and growth-factor requirements
Temperature and pH requirements and tolerance
Fermentation products, metabolites, and toxin production
Antibiotic sensitivity pattern (antibiogram)
Gas requirements and tolerance
Genetic profile; DNA/RNA sequences and fingerprinting
Pathogenicity to animals and humans
Serology and phage typing
Cell wall, cell membrane, and cellular components
Growth rate and generation time
Ecological niche and survival ability
Motility and spore formation
Extracellular and intracellular products
Response to electromagnetic fields, light, sound, and radiation
Resistance to organic dyes and special compounds
Impedence, conductance, and capacitance characteristics

Source: Adapted from Fung (38).

Miniaturized Methods

Fung and Hartman (40) developed a variety of miniaturized microbiological methods for the identification and enumeration of aerobic bacteria. The entire concept and system were adapted for miniaturized anaerobic bacteriology simply by operating the procedures in an anaerobic chamber (60,61). Guy Miller and T.G. Nagaraja at Kansas State University (personal communication, 1987) found that miniaturized techniques could be used for studying rumen bacteria effectively. The advantages of miniaturized techniques are savings of material, space, time of incubation and operation, and labor. Another advantage is the flexibility of the system to be able to adapt to many different laboratories, especially for environmental and industrial settings where commercial diagnostic kits are not available for specialized organisms.

Commercial Diagnostic Kits

Around the beginning of 1970s with the advancement of miniaturized microbiological techniques and concepts de-

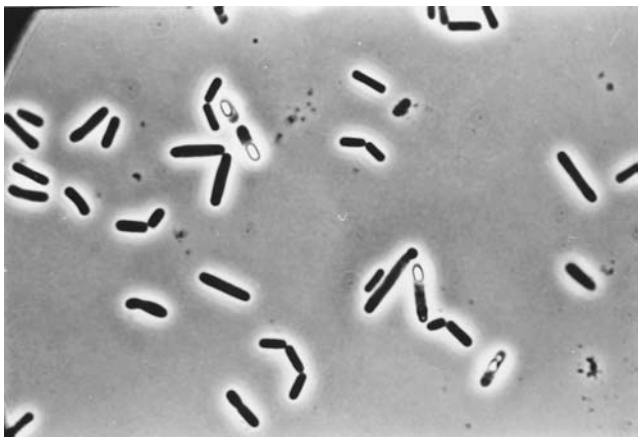


Figure 1. Morphology of *Clostridium perfringens* under 1,000X light-phase-contrast microscopy. This is a gram-positive, rod-shaped, anaerobic, spore-forming bacterium that causes large numbers of foodborne outbreaks and cases in the world. The dark rod-shaped cells are vegetative bacteria, and the bright refractile oval objects are spores developed in the vegetative cells.

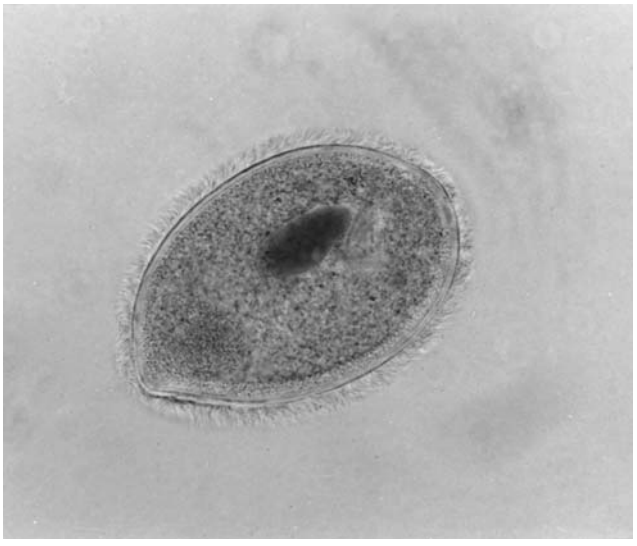


Figure 2. Morphology of *Isotricha prostoma*, under 100X light-phase-contrast microscopy. This is an anaerobic protozoan an important member of the ruminal flora in cattle. *Source*: Courtesy of T. G. Nagaraja of Kansas State University.

veloped by Fung and colleagues and other scientists, many diagnostic kits appeared on the market mainly for clinical microbiology. These methods gradually found their ways into food microbiology laboratories and sometime into industrial and environmental laboratories. Currently commercially successful diagnostic kits are marketed toward clinical settings. This may change if the need arises in the near and far future. There are two major types of diagnostic kits. One type depends on *growth* of the pure culture and biochemical reactions obtained in the media. This type of test is typically slow, requiring overnight incubation and anaerobic incubation of cultures. Another type is by monitoring *enzyme systems* of the pure culture. In this system, anaerobiosis is not necessary and the results can be obtained in about 4 hours. Kelly and Schipp (49) made a detailed analysis of the approaches and designs of anaerobic diagnostic kits, including discussions on substrate concentration, organism concentration, pH of reaction, buffering capacity, additives, incubation time, and growth medium for the identification of anaerobes. In that article, important variables affecting system performance and issues concerning comparative analysis of data reported in journal articles were also studied. To compare data reported in the literature, one needs to address the issues such as source of strains, database version, the "gold standard," number of species/strains tested, need of supplementary tests, and accuracy. In general, a diagnostic kit providing 90 to 95% accuracy compared with the conventional method is considered good. Any system with less than 85% agreement with the conventional method is considered unacceptable.

The major growth-dependent commercial kits for anaerobes are API 20A, marketed by bioMerieux Vitek Inc., Hazelwood, Mo., and Minitek Anaerobe system, marketed by BBL Microbiology Systems, Cockeysville, Md. The API 20A kit consists of 20 microtubes with dehydrated substrates. A liquid suspension of a pure culture is made and then introduced into each microtube. The kit is then incubated anaerobically, and after 1 to 2 days the reactions are read and data recorded in a convenient chart. The results are converted into a number system. From a code book the identity of the culture can be ascertained. These types of results provide a statistical probability of the unknown culture by comparing the attributes in the computerized database. Thus, an organism may be identified as *C. perfringens* at 0.9500 level. Sometimes the printout may include requests for additional tests, and in some cases the organism may be given two identifications with probability of each listed. This general format of identification is used by all commercial diagnostic kits.

Minitek Anaerobe II uses paper disks saturated with substrates. Thirty-five disks are available. After a suspension of the pure culture is made, the liquid is dispensed into wells of the units containing the appropriate disks. The assemblage is incubated anaerobically for 2 days. Reactions are read by color changes and recorded. The identity of the unknown can be obtained by following the procedure outlined for the API system.

Use of enzyme test systems include Rapid ID 32A (bioMerieux), AN-Ident (bioMerieux), RapID ANA II (IDS Inc., Norcross, Ga.), Vitek ANI Card (bioMerieux), and MicroScan Rapid Anaerobe Panel (Baxter Diagnostic, Sacramento, Calif.). All these methods are incubated aerobically because enzyme activities will not be influenced by the presence of air. Rapid ID 32A and RapID ANA II are similar in that the cultures are made into a liquid suspension, and the liquid is inoculated into microwells. After 4 hours, the color reactions are read and the identities are ascertained from a code book. Vitek ANI Card involves injection of liquid sample into a plastic card that is the size of a credit card with 30 wells. After 4 hours of incubation, the results are read and data are entered into the ANI program on the VITEK computer. Identification is made by the computer. In the MicroScan Rapid Anaerobe Panel system, the culture is inoculated into 24 substrates and two color interpretation controls. After 4 hours of aerobic incubation and addition of reagents, the panels are read either manually or with the aid of a Touchscan light box. Data are then interpreted by computer for identification.

It should be emphasized that the aforementioned systems are designed for clinical bacteriology. Identifications of unknowns from environmental, food, and industrial sources may not be as accurate because the databases of these systems are from clinical isolates.

Physical and Biochemical Methods

Physical and biochemical analysis of components of cells and metabolites of cells have also been used for identification of anaerobes. Fingerprinting techniques have been explored to identify anaerobes by matching profiles from unknown cultures with known profiles in the data base. One such approach is the gas–liquid chromatography method popularized by the VPI Anaerobic Laboratory (45). Mitruka (51) provided a detailed discussion on this method. The Hewlett-Packard company developed a procedure to fingerprint the fatty acid profiles of microorgan-

isms, including anaerobes, using a high-performance liquid chromatography technique. The AMBIS system involves computer-integrated fingerprinting of electrophoretically separated labeled proteins of bacteria to match profiles of unknowns versus knowns. In all these systems, these cultures must be pure and grown anaerobically before accurate identification can be made. Again, the database must be from the appropriate sources for the identification to be accurate.

Immunological Methods

Microorganisms have specific antigen properties, and many of these antigens can elicit the production of antibodies in animals. With the appropriate antibodies, one can then use the antibodies to identify the unknown by antigen–antibody reactions. Historically, polyclonal antibodies have been used for detection and characterization of bacteria. More recently, monoclonal antibodies have been used for the reactions. Both types of antibodies are useful for anaerobic microbiology. The reactions can be agglutination of cells directly with antibodies or indirect agglutination of antigens reacting with antibodies fixed on a variety of support systems, including blood cells, latex beads, plastic steel balls, etc. The most popular format of antigen–antibody reactions used is ELISA. In this test, the antibody is fixed on a support surface and the antigen is captured. A second antibody then reacts with another site of the antigen. The second antibody is conjugated with an enzyme. A substrate is then applied, and if the enzyme is present a color reaction will occur. The intensity of the color reaction will indicate the presence of the antigen. A cutoff color intensity is established to indicate a positive or negative test of the presence of the antigen. At first, this test was done manually, but recently, it has been completely automated. The analyst presents the instrument with a suspect sample, and the instrument will automatically complete the test with a computer printout.

An exciting new development in immunological analysis is the field of immunomagnetic capture technology. In this technology, beads are magnetized and then antibodies or other capture particles are attached on the surface of the beads. Antibodies against whole cells, antigens, and other target particles are fixed on the beads. The charged beads are then introduced to a broth or food sample to interact with target bacteria or other molecules. In this phase, the sample is mixed thoroughly with the charged beads tumbling in the liquid food to interact with target organisms. After the interaction, a powerful magnet is applied to the side of the reaction tube. The magnet holds all the beads with or without captured organisms. The rest of the debris is discarded, and the beads are released and washed. The captured organisms can be subject to growth in a special broth or agar, or these may be used for further immunological reaction, PCR, DNA/RNA reactions, or other tests.

PCR and Related Methods

DNA and RNA hybridization tests have been developed to detect organisms whose specific gene sequence of interest is known. At first, radioactive compounds were used to report the hybridization, but recently, the reaction has been reported by enzyme and color reactions. The truly revolutionary development in genetic type of identification of microorganisms is in the field of PCR technology. In this technology, the gene sequence of the organism of interest (an anaerobe, for example) is known at the DNA level. A piece of the DNA can be unfolded by heat, and after cooling of the unfolded DNA, two primers will interact with specific regions of the single DNA strands. A heat-stable polymerase will complete the complementary strand in the presence of nucleotides from the 3′ end and the 5′ end. Thus, in one cycle, one DNA molecule will become two, and after another cycle, two will become four, etc. In about 2 hours, one molecule of DNA can be amplified to 10^6. To use this technology, the genetic sequence of the target organism must be known, and the appropriate primers must be developed for successful use of this powerful tool. There are many new variations of this technology, such as the search for a cold amplification method to bypass the heating cycle and the development of an ELISA-type detection scheme to bypass the need to do electrophoresis of PCR products.

Another technology is called ribotyping. In this technology, the DNA of a target organism is extracted from the cells, an appropriate enzyme or a collection of enzymes is then used to cut the DNA. The DNA fragments are then separated by electrophoresis, and the pattern of the bands are photographed and quantified, resulting in special fingerprints.

The fingerprints can be matched with fingerprints of known cultures for identification of unknowns. One added feature is that for certain organisms (e.g., *E. coli*) there are different fingerprints, even for the same organism. This becomes very important when one wants to trace the occurrence of an organism in the case of an outbreak of foodborne disease. For example, finding *E. coli* O157:H7 in several foods and the environment in an outbreak is not good enough to trace the etiology of the outbreak. With ribotyping, the culprit *E. coli* O157:H7 can be traced to a particular food by matching the ribotyping pattern of the strain that caused the outbreak with the origin of the strain in certain foods or in the environment. This will help pinpoint the source of the infection. Dupont's Qualicon Division is leading the way for the development and commercialization of this technology.

There are, of course, many other methods for identification of unknowns, such as impedance, conductance, capacitance, pyrolysis pattern, microcalorimetry, flow cytometry, etc. *Automated Microbial Identification and Quantitation: Technologies for the 2000s*, an excellent recent reference book edited by Olson (55), highlights many of these developments. There will be an explosion of technologies in applied microbiology in the near future that will directly and indirectly influence the enumeration, characterization, isolation, and identification of anaerobes.

INDUSTRIALLY IMPORTANT STRAINS AND PRODUCTS

Many products are or have been produced by fermentation in commercial quantities (2–4). Food products, industrial bulk chemicals, solvents, enzymes, and amino acids are

produced with the help of anaerobic microorganisms. In this work, the emphasis is on those processes carried out by anaerobes, which generally means that important pathways yield useful products under conditions where oxygen is not required. However, oxygen is present in modest quantities to enhance microbial growth in some cases. Alternative manufacturing processes have been developed for some products; they can be produced by fermentation and by chemical processes that do not require microorganisms. Often these products are made by a microbial process in some countries and a chemical process in other countries; both processes may be used in the same country as well. The choice may depend on local raw materials, taxes, and scale of operation; often local economics determines which process is used.

Table 2 lists several compounds produced by anaerobes and some of the microorganisms that can produce significant quantities of the product. The pathways by which these products are produced have been investigated; many of the well-known industrially important pathways such as the Embden-Meyerhof-Parnas pathway are described in detail elsewhere (2,6,7). Ethanol is produced in large quantities for use as a motor fuel in several countries (2,68). Methane is often produced by mixed cultures as a product of anaerobic digestion of waste materials. Lactic acid is widely produced in food fermentations where milk, meat, or vegetables are fermented. Table 3 lists some of the many food fermentations and microorganisms that are present and active in these fermentations. Detailed descriptions of food fermentations are given by Reed (3) and Steinkraus (4). Flavors are introduced into fermented foods as a result of the fermentation process (8); however, preservation is often as important as flavor in food fermentations.

Table 2. Anaerobic Fermentation Products and Microorganisms That Can Produce Significant Quantities of the Product

Product	Microorganism
Methane	*Methanobacterium*
	Methanococcus
	Methanosarcina
	Methanobrevibacter
	Methanospirillum
Acetate	*Clostridium thermoaceticum*
Ethanol	*Saccharomyces cerevisiae*
	Saccharomyces uvarum
	Saccharomyces carlsbergensis
	Zymomonas mobilis
	Clostridium thermocellum
Lactic acid	*Lactobacillus*
	Streptococcus (Lactococcus)
	Leuconostoc
	Pediococcus
Propionate	*Clostridium propionicum*
	Clostridium tetanomorphum
Acetone	*Clostridium acetobutylicum*
Isopropanol	*Clostridium beijerisckii*
Acetoin	*Clostridium acetobutylicum*
Butanediol	*Enterobacter aerogenes*
	Bacillus
	Serratia
	Erwinia
	Aerobacter
Butyrate	*Clostridium butyricum*
	Clostridium tyrobutyricum
Butanol	*Clostridium acetobutylicum*
	Clostridium beijerinskii
	Clostridium tetranomorphum

Source: Linden et al. (5).

Knowledge of the biochemical pathways and the enzymes that facilitate the biochemical transformations is needed to commercialize and optimize many anaerobic fermentations. The control systems are often altered to overproduce desired products. In the case of biodegradation, it may be necessary to add plasmids that provide genetic material necessary for a biodegradation pathway to function. As new products such as gasoline oxygenates enter the soil and water environment, microorganisms capable of biodegradation of methyl-*tert*-butyl ether (MTBE) have evolved (67).

GROWTH AND PRODUCT FORMATION

Anaerobic growth and product formation are often investigated together when the product is growth associated. In anaerobic fermentations, adenosine triphosphate (ATP) needed for growth is often formed as a result of substrate phosphorylation. Products such as ethanol are produced because the organism is growing and needs ATP for synthesis of microbial cell mass. Mass balances and yield parameters that describe growth and product formation can be written in terms of the mass of substrate, biomass, and product or in terms of the equivalents of available electrons of each of these.

Consider the anaerobic production of a simple product, such as ethanol, by anaerobic fermentation. The chemical balance equation is as follows:

$$CH_mO_l + aNH_3 = y_cCH_pO_nN_q + zCH_rO_sN_t + cH_2O + dCO_2 \quad (1)$$

where CH_mO_l, $CH_pO_nN_q$, and $CH_rO_sN_t$ give the elemental compositions of the carbon, hydrogen, oxygen, and nitrogen in the substrate, biomass, and extracellular product, respectively. For glucose as substrate, $m = 2$ and $l = 1$; for ethanol as product, $r = 3$, $s = 0.5$, and $t = 0$. For the valences C = 4, H = 1, O = −2, and N = −3, the available electron balance (2,9) is

$$y_c \frac{\gamma_b}{\gamma_s} + z \frac{\gamma_p}{\gamma_s} = 1.0 \quad (2)$$

or

$$\eta + \xi_p = 1.0 \quad (3)$$

where γ is the reductance degree (2,9) and the subscripts s, b, and p refer to substrate, biomass, and product, respectively.

Luedeking and Piret (10,11) reported that product formation kinetics and growth kinetics were related; product

Table 3. Fermented Food Products and Associated Microorganisms

Product	Microorganisms	Reference
Yogurt	*Streptococus thermophilus* *Lactobacillus bulgaricus*	2
Sauerkraut	*Enterobacter cloacae* *Erwinia herbicola* *Leuconostoc mesenteroides* *Lactobacillus brevis* *Lactobacillus plantarum* *Pediococcus cerevisiae*	2,3
Cucumber fermentation (pickles)	*Lactobacillus brevis* *Lactobacillus plantarum* *Pediococcus cerevisiae* *Leuconostoc mesenteroides*	3
Sausage	*Lactobacillus* *Leuconostoc* *Pediococcus*	3
Bread and sourdough bread	*Saccharomyces cerevisiae* *Lactobacillus* *Candida milleri* *Leuconostoc mesenteroides* *Saccharomyces cerevisiae*	21
Cheddar cheese	*Streptococcus (Lactococcus) cremoris* *Streptococcus (Lactococcus) lactis*	3
Swiss cheese	*Lactobacillus bulgaricus* *Streptococcus thermophilus* *Propionibacterium shermanii*	3
Blue cheese	*Streptococcus (Lactococcus) lactis* *Streptococcus (Lactococcus) cremoris* *Penicillium roquiforti*	3
Camembert cheese	*Streptococcus (Lactococcus) lactis* *Streptococcus (Lactococcus) cremoris* *Penicillium camemberti*	3
Cottage cheese	*Streptococcus (Lactococcus) lactis* *Streptococcus (Lactococcus) cremoris*	3
Acidophilus milk	*Lactobacillus acidophilus*	3
Kefir	*Streptococcus (Lactococcus) lactic* *Streptococcus (Lactococcus) cremoris* *Torula* or *Candida yeast*	3
Buttermilk and sour cream	*Streptococcus (Lactococcus) lactis*, (including subspecies diacetylactis) *Streptococcus (Lactococcus) cremoris* *Leuconostoc cremoris*	3
Wine	*Saccharomyces cerevisiae* *Saccharomyces uvarum* Other *Saccharomyces*	3
Beer	*Saccharomyces cerevisiae* *Succharomyces uvarum*	3
Soy sauce	*Asperigillus oryzae* *Saccharomyces rouxii* *Torulopsis versatilis* *Pediococcus soyae*	3
Miso	*Pediococcus halophilus* *Saccaromyces rouxii* *Torulopsis* *Streptococcus faecalis*	3
Tempeh	*Rhizopus oligosporus*	3

formation was modeled using a growth-associated term and a maintenance-associated term; that is,

$$\frac{dP}{dt} = \alpha \frac{dX}{dt} + \beta X \tag{4}$$

where X is biomass concentration and P is product concentration. Growth is dependent upon the availability of substrate according to the Monod model

$$\frac{1}{X}\frac{dX}{dt} = \mu = \frac{\mu_{max} S}{K_s + S} \tag{5}$$

Here S is substrate concentration, μ is specific growth rate, μ_{max} is maximum specific growth rate, and K_s is the saturation constant. The organism uses substrate for growth and maintenance as follows:

$$-\frac{dS}{dt} = \frac{1}{Y_{x/s}^{max}}\frac{dX}{dt} + m_s X \tag{6}$$

where $Y_{x/s}^{max}$ is the true growth yield corrected for maintenance and m_s is the maintenance coefficient. In this process, the product formation kinetic parameters and the bioenergetic parameters are related. Equation 6 may be written as (2)

$$\frac{\mu}{\eta} = \frac{\mu}{\eta_{max}} + m_e \tag{7}$$

and equation 4 becomes

$$\frac{\mu}{\eta}\xi_p = \frac{\mu}{\eta}(1 - \eta) = \alpha_e \mu + \beta_e \tag{8}$$

if the following relationships are used (2)

$$m_e = \beta_e = m_s \frac{\sigma_s \gamma_s}{\sigma_b \gamma_b} = \beta \frac{\sigma_p \gamma_p}{\sigma_b \gamma_b} \tag{9}$$

$$\frac{1}{\eta_{max}} = \alpha_e + 1 = \frac{\sigma_s \gamma_s}{\sigma_b \gamma_b Y_{x/s}^{max}} = \alpha \frac{\sigma_p \gamma_p}{\sigma_b \gamma_b} + 1 \tag{10}$$

The true growth yield in equation 7 must satisfy the relationship (2)

$$\eta_{max} = \frac{(\sigma_b \gamma_b/12) Y_{ATP}^{max}}{(\sigma_b \gamma_b/12) Y_{ATP}^{max} + \delta} \tag{11}$$

where Y_{ATP}^{max} is the true growth yield based on ATP, δ is the equivalents of available electrons transferred to products to produce 1 mol of ATP from ADP, and σ_b is the weight fraction carbon in biomass. Information on the theoretical maximum yield with respect to ATP and the value of δ for a particular fermentation allows one to estimate the range of values of the growth-associated kinetic term in the product formation model. For $\delta = 12$ equivalents of available electrons per mole of ATP generated, $Y_{ATP}^{max} = 28.8$ g cells per mole ATP, $\sigma_b = 0.462$, and $\gamma_b = 4.291$; $\eta_{max} = 0.283$ and $\alpha_e \geq 2.52$.

Product formation is directly linked to the specific growth rate, true growth yield corrected for maintenance, and the maintenance coefficient as shown in equations 1 to 11. Equation 3 shows that the product yield is related to the biomass yield; the largest product yields are expected when the cells are using carbon and energy primarily for maintenance and nearly all the available electrons are being converted to product. When oxygen is present, some of the available electrons in the carbon substrate are transferred to oxygen, and the product yield is reduced accordingly.

Growth kinetics, product formation rates, and product yields are affected by temperature, pressure, water activity, osmolality, oxidation–reduction potential, bioenergetics, and the concentrations of organic substrate, inorganic nutrients, biomass, hydrogen ions (pH), carbon dioxide, products, and electron acceptors such as oxygen, nitrate, and sulfate (2).

BIOENERGETICS AND PRODUCT YIELDS

The product yields under anaerobic conditions depend on the operating conditions and the fermentation pathway for the selected strain of microorganisms. Equations 2 and 7 can be rearranged as follows:

$$\xi_p = 1 - \eta \tag{12}$$

$$\eta = \frac{\mu}{\mu/\eta_{max} + m_e} \tag{13}$$

If equation 13 is substituted into equation 12, one obtains

$$\xi_p = 1 - \frac{\mu}{\mu/\eta_{max} + m_e} \tag{14}$$

Equation 14 indicates that the product yield is a function of the biomass true growth yield, which depends on the growth yield with respect to ATP and the efficiency with which ATP is formed through substrate phosphorylation. The product yield also depends on the values of the maintenance coefficient and specific growth rate. Table 4 shows

Table 4. Effect of ATP Yield, True Growth Yield, Maintenance Coefficient, and Specific Growth Rate on Biomass Yield and Product Yield

Y_{ATP}^{max}	δ	η_{max} [a]	m_e	μ	η	ξ_p [b]
28.8	12	0.283	0.1	0	0	1.0
28.8	12	0.283	0.1	0.1	0.221	0.779
28.8	12	0.283	0.1	0.2	0.248	0.752
28.8	12	0.283	1.0	0.1	0.0739	0.926
10.5	12	0.126	0.1	0.1	0.112	0.888
10.5	12	0.126	1.0	0.1	0.056	0.944
10.5	24	0.0674	0.1	0.1	0.063	0.937
10.5	24	0.0674	1.0	0.1	0.040	0.960

[a] Values estimated from equation 11.
[b] Values estimated from equation 14.

how the true growth yield varies with the growth yield with respect to ATP and the number of moles of available electrons of substrate required to produce 1 mol of ATP according to equation 11. It also illustrates how the product yield varies as a function of specific growth rate and maintenance coefficient. The product yield decreases as the specific growth rate increases. As values of the maintenance coefficient increase, the product yield increases also.

Estimated values of the true growth yield and maintenance coefficient are presented in the book by Erickson and Fung (2) for several sets of experimental data reported by a variety of investigators. Selected values are presented in Table 5. These values are within the expected range of values based on theoretical yields and known pathways; however, not all of the estimated values reported by Erickson and Fung (2) are below the estimated theoretical yield. It is common to have experimental error because of the difficulty of making measurements. When data consistency is checked by making carbon and available electron balances and yields are compared with theoretical yields, it is possible to identify experiments where errors are present; however, it is not always possible to determine which particular measurements have the largest errors associated with them. When errors are present in the data, the covariate adjustment method (2) should be used in data analysis.

The results in Tables 4 and 5 may be related to metabolic pathways. The Embden-Meyerhof-Parnas pathway commonly produces 2 mol of ATP per mole of glucose (δ = 12 equivalents of available electrons per mole of ATP generated), whereas the Entner-Doudoroff pathway produces 1 mol of ATP per mole of glucose (δ = 24). *Zymomonas mobilis* has the Entner-Doudoroff pathway; thus, the lower values of true growth yield in Table 5 are expected. As shown in Table 4, the product yield is predicted to be larger for *Z. mobilis*.

ANAEROBIC DIGESTION AND BIODEGRADATION

Anaerobic digestion is widely used in wastewater treatment to degrade the solids from primary and secondary treatment. Anaerobic processes are also applied in the treatment of industrial effluents and in bioremediation of contaminated soil and ground water. Although environmental microorganisms have been functioning and serving the needs of humans for a very long time, the commercialization and engineering of this aspect of bioprocess technology was started after the discovery of microorganisms. Wastewater treatment is the second oldest bioprocess technology after food processing.

Anaerobic degradation processes depend on the electron acceptors that are present. Stouthamer (13) points out that when several electron acceptors are present, the microorganism selects the electron acceptor that yields the largest amount of energy by repressing the formation of reductase enzymes for the other electron acceptors. Table 6 gives the standard free energy changes for oxidation of acetate with oxygen, nitrate, sulfate, and water as electron acceptors, respectively. When oxygen is present, aerobic processes are observed and anaerobic processes are generally absent unless there are spatial regions or time periods where oxygen is absent. Nitrate is the electron acceptor of choice when oxygen is absent, and when nitrate and oxygen are both absent, sulfate is the favored electron acceptor. Methane production, which occurs when water is the electron acceptor, is generally inhibited if sulfate is present. Stouthamer points out that other electron acceptors may participate, including other oxidized forms of nitrogen and fumarate.

Many bacteria have been found that can use nitrate as the terminal electron acceptor. The first product of nitrate respiration or nitrate reduction is formation of nitrite, which can be further used as an electron acceptor by many microorganisms. In denitrification, the end product of further reduction is nitrogen gas; however, in other cases the end product is ammonia. Stouthamer (13) reviewed the enzymology and bioenergetics of nitrate reduction.

Sulfate-reducing bacteria include *Desulfovibrio, Desulfobacterium, Desulfotomaculum, Desulfococcus, Desulfosarcina, Desulfomonile, Desulfonema, Desulfoarculus, Desulfobulbus, Desulfobotulus,* and *Desulfobacula* (19). For these genera, many species have been described in the literature (13,14,19). These organisms have great diversity in biochemical properties, and they can completely oxidize a variety of organic substrates.

Methanogenic bacteria are widely found under anaerobic conditions in nature. Methane is produced in soil, landfills, anaerobic digesters, rumens of animals, and many other places. In anaerobic digestion, methanogenic bacteria are found together with hydrolytic fermentative microorganisms and syntrophic acetogenic bacteria. The hydrolytic fermentative organisms, which include eubac-

Table 5. Estimated Values of m_e and η_{max} for Selected Experiments

m_e	η_{max}	Organism	Substrate	Product
0.896	0.164	*Saccharomyces cerevisiae*	Glucose	Ethanol
1.791	0.0835	*Zymomonas mobilis*	Glucose	Ethanol
3.264	0.078	*Zymomonas mobilis*	Glucose	Ethanol
0.078	0.281	*Lactobacillus delbruekii*	Glucose and yeast extract	Lactic acid
0.112	0.308	*Lactobacillus delbruekii*	Glucose and yeast extract	Lactic acid
0.491	0.267	*Lactobacillus delbruekii*	Glucose	Lactic acid
0.396	0.206	*Lactobacillus bulgaricus*	3% Nonfat dry milk	Lactic acid
0.286	0.230	*Clostridium acetobutylicum*	Glucose	Acetone, butanol, ethanol, butyrate

Source: Data reported by Oner et al. (12).

Table 6. Standard Free Energy Changes per Equivalent of Available Electrons for Oxidation of Acetate with Various Electron Acceptors

Electron acceptor	Free energy change (kJ/equiv)
Oxygen	−106
Nitrate	−99
Sulfate	−6
Water	−4

Source: Stouthamer (13).

teria, fungi, and often protozoa, produce fermentation products such as ethanol, lactate, butyrate, succinate, and acetate. The acetogenic bacteria (*Pellobacter, Desulfovibrio, Syntrophobacter, Clostridium, Syntrophomonas*, and others) (16) produce acetate, formate, bicarbonate and hydrogen from the ethanol, lactate, butyrate, and other organic acids and alcohols with more than two carbons. The methanogenic bacteria use the acetate, formate, and hydrogen as substrates. In anaerobic digesters, about 75% of the available electrons in the complex organics are first transformed to fermentation products other than acetate, formate, and hydrogen, which make up the remaining 25%. More than 60% and as much as 72% of the available electrons is converted to acetate, which is then used by methanogens to form methane. More than 25% of the available electrons are converted to hydrogen before being incorporated into methane (2). Thiele and Zeikus (15) and Thiele (16) reviewed mixed culture interactions in methanogenesis and provided information on selected anaerobic microorganisms that are active in anaerobic digestion.

Many microorganisms that are active in environmental microbiology have been studied and described; however, each year many new organisms are described for the first time (17). New anaerobic treatment processes for specific applications, such as nitroaromatic munitions compounds, are being developed (18). Although anaerobic processes are often slower than aerobic processes, they have been found to be economically attractive in a wide variety of applications.

MIXED CULTURES

Mixed cultures are used extensively in food fermentations and in biodegradation of complex substances. Mutualism, commensalism, amensalism, competition, and prey–predator relationships are the main interactions observed in mixed cultures (16). Mutualism occurs when both species benefit from the interaction, whereas commensalism describes processes where only one species benefits and the other is not affected. Amensalism occurs when one species restricts another, whereas competition affects both species because of nutrient limitations. The predator consumes the prey in prey–predator interactions. In actual situations, more than two species are often present, and several types of interaction may be taking place.

In the recent book on mixed cultures by Zeikus and Johnson (20), most of the contributions are related to either food fermentations or environmental microbiology. There are food fermentation chapters on bread, milk, oriental foods, wine, and vegetables and environmental microbiology chapters related to degradation of polysaccharides, methanogenesis, detoxification of hazardous waste, corrosion, and leaching processes in mineral biotechnology.

In San Francisco sourdough bread, *Lactobacillus sanfrancisco* ferments only the maltose, whereas *Candida milleri* uses all the other free sugars but not maltose. Lactic acid, acetic acid, and carbon dioxide are the main products of the fermentation; however, small amounts of propionic, isobutyric, butyric, α-methyl-*n*-butyric, isovaleric, and valeric acids have been found in commercial San Francisco sourdough breads (21).

Anaerobic cellulose degradation is accomplished by a mixed culture of *Clostridium* cellulolytic bacteria that produce cellobiose, glucose, and cellodextrins; fermentative bacteria that produce propionate, butyrate and other fermentation products from glucose and other intermediates; syntrophic acetogenic bacteria that convert the fermentation products to acetate, hydrogen, and carbon dioxide; homoacetogens that convert hydrogen and carbon dioxide to acetate; and methanogens that produce methane from acetate, formate, and hydrogen (22). Leschine (22) reports that *Clostridium papyrosolvens* grow mutualistically in coculture with a noncellulytic *Klebsiella; C. papyrosolvens* hydrolyzes cellulose for the *Klebsiella,* whereas *Klebsiella* excretes vitamins required by the *C. papyrosolvens*. The soluble sugar products of cellulose hydrolysis provide substrates for many noncelluloytic commensal organisms that depend on the cellulolytic bacteria, but do not provide known compounds in return.

In the natural environment, the human environment, and even in the industrial environment, mixed cultures of bacteria, yeast, mold, viruses, protozoa, nematodes, and higher organisms live in antagonistic, cooperative, or inert coexistence. Pure culture systems are the result of understanding of the fundamental role of microbes in a particular process and the isolation and cultivation of such cultures for specific reactions. This section deals with single-culture processes mixed pure-culture processes, and mixed natural-culture processes, using food systems as examples. This information certainly can be applied to other industrial and fermentation processes.

Other publications deal adequately with the subject of single-cell process. The key success of single-culture process is to provide the culture with a sterile substrate and environment with no contamination during the process. Single-cell process is a manmade situation classified as a controlled process because the substrate is prepared and processed in such a way as to minimize contamination. Examples of this type of process are wine making, beer making, bread making, single-culture dairy product fermentation, and vinegar production. The kinetics of growth and product formation are easier to control and monitor.

Mixed Pure-Culture Process

For some processes, it is desirable to have more than one pure culture for proper product development. The mixed pure cultures can be a controlled mixture of bacterium

with bacterium or with a combination of yeast or mold, or both. The relationships among these pure cultures during growth become very complex.

Mixed Natural-Culture Process

In many situations, such as in nature and noncontrolled fermentation, the flora in the process are mixed natural cultures. The interactions of these microbes are even more complex than those of mixed pure-culture process. The microbial successions in these processes follow a special sequence. If the products are treated properly, the desirable cultures provide the characteristic end results. On the other hand, if the conditions are not treated properly, the process will fail.

Mixed Bacteria Interactions

An example of mixed pure bacterial culture fermentation is yogurt fermentation, which typically uses *Streptococcus thermophilus* and *Lactobacillus bulgaricus*. The interactions are commensal. Figure 3 illustrates the development of acidity of pure and mixed cultures of *S. thermophilus* and *L. bulgaricus* in liquid. In mixed cultures, *S. thermophilus* grows faster at the beginning, reduces the pH by lactic acid production, and also produces formic acid, which is a stimulatory compound for *L. bulgaricus*. *L. bulgaricus* produces amino acids that stimulate the growth of *S. thermophilus*. Higher acid production with mixed cultures was obtained and compared to pure cultures, as can be seen in Figure 3 (39,52). This is an example of commensalism in mixed cultures. Details of the kinetics and mathematical relationships of interactions between *S. thermophilus* and *L. bulgaricus* are presented in an article by Fung et al. (39).

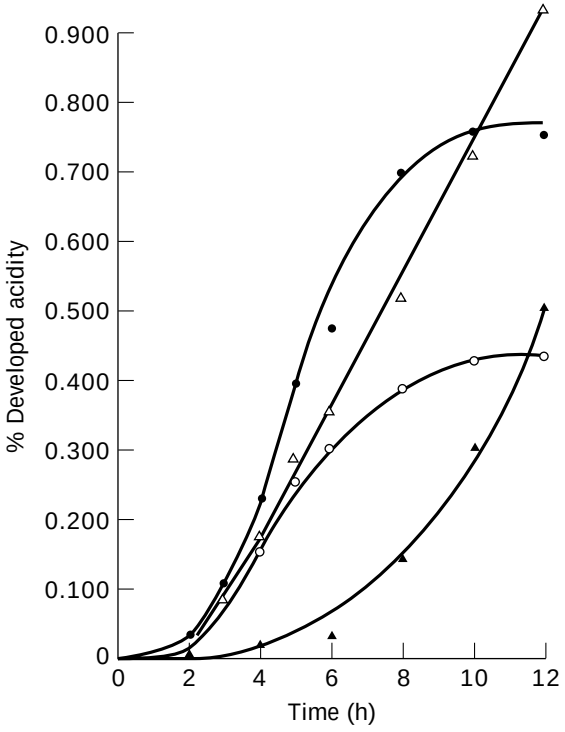


Figure 3. Developed acidity of pure and mixed cultures of *L. bulgaricus* and *S. thermophilus*. *Open circles*, *S. thermophilus*; *closed triangles*, *L. bulgaricus*; *open triangles*, sum of developed acidity of pure cultures; *closed circles*, mixed culture. *Source*: Reprinted with permission from *Journal of Food Protection* and Marcel Dekker, Inc. Reprinted from Reference 39, page 511, courtesy of Marcel Dekker, Inc.

Mixed Bacterium and Yeast Interaction

Yeast is one of the most important organisms for industrial use in alcoholic fermentation. Bacterial contamination may occasionally occur during yeast fermentation, which may pose serious problems for the fermentation industry. This section deals with the interaction of *Saccharomyces cerevisiae* with *Acetobacter suboxydans* or *E. coli* in mixed-culture growth. *S. cerevisiae* was grown in mixed cultures with either *A. suboxydans* or *E. coli* in a culture medium and glucose at 28 °C. Growth was monitored by viable cell counts of yeast and bacteria using selective agars, direct count by microscopic reading of cells in a Petroff-Hauser counting chamber, and electronic counts by the Coulter counter. By selecting the appropriate threshold windows, the electronic instrument can differentiate yeast count and bacterial count in the same liquid. During the growth period, pH values were monitored by a conventional pH meter and glucose was determined by the Glucostat, a commercial kit designed to detect glucose in solution. A vinometer was used to measure alcohol level of the solution.

Figure 4 illustrates the interaction between *S. cerevisiae* and *A. suboxydans* as reported by Fung et al. (39). For the yeast, both direct count and electronic count increased with time. The direct count registered about 10 times more cells than the electronic count. Because *A. suboxydans* does not grow well in solid agar medium, viable cell count data were not obtained. For *A. suboxydans*, the direct count increased with time, but the electronic count increased to about log 7.8 mL^{-1} and then started to decrease. The decrease in number is due to the adhesion of many of the bacterial cells to yeast cells. The electronic count cannot differentiate these two populations in this condition. However, microscopic observations can ascertain the different counts. The concentration of glucose completely disappeared in the first 10 h of yeast and bacterial interactions. The pH of the medium first decreased and then stabilized at pH 3. This is because of oxidation of alcohol (produced by *S. cerevisiae*) by *A. suboxydans* to acetic acid, making the reaction mixture acidic.

The interactions of yeast and *E. coli* (39) in Figure 5 show interesting contrasts compared with the yeast–*Acetobacter* interaction. Growth of yeast as monitored by viable cell count, direct count, and electronic counts all increased with time. After reaching a stationary phase, viable yeast count decreased, but direct count and electronic count registered no reduction in numbers. A likely explanation is that the former two biomass measurements counted both live and dead cells, but the viable cell count only registered living cells. After the stationary phase, some of the yeast cells went through the death cycle. The growth curves of *E. coli* showed an increase in the viable

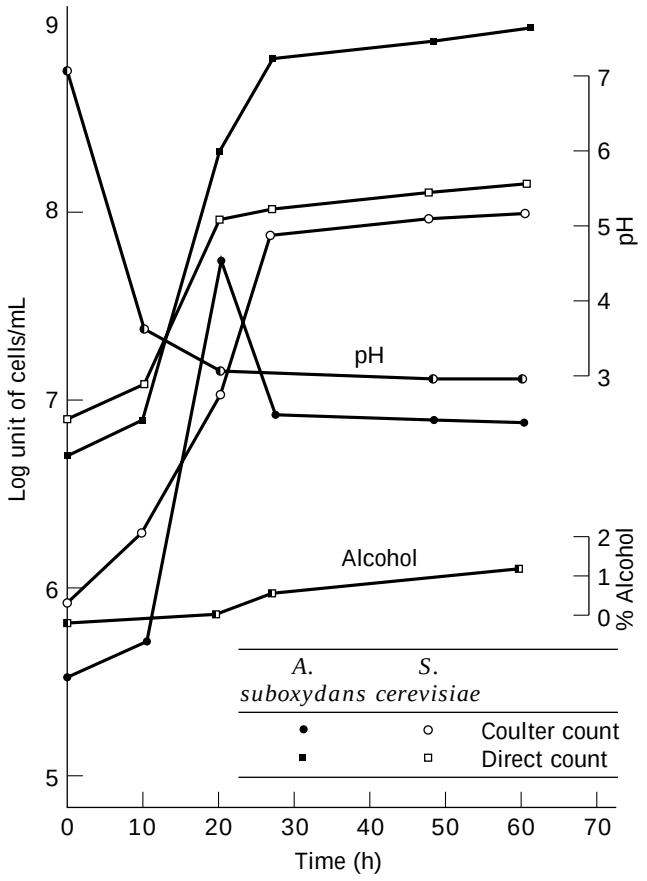


Figure 4. Interaction of mixed cultures of *S. cerevisiae* and *A. suboxydans* in terms of cell numbers monitored by direct count and electronic count as well as product formation (alcohol and acid production. *Source*: Reprinted with permission from reference 39, page 527, by courtesy of Marcel Dekker, Inc.

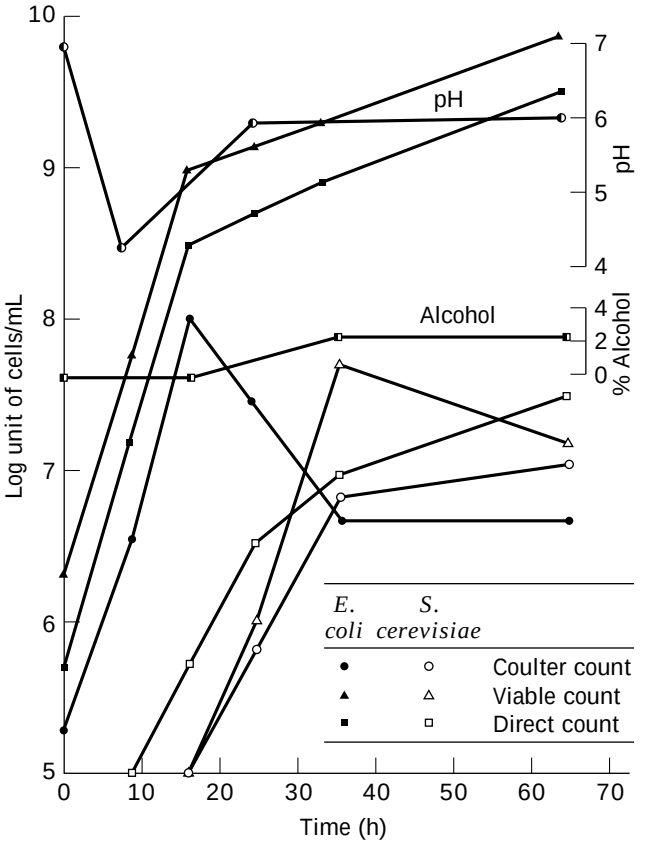


Figure 5. Interaction of mixed culture of *S. cerevisiae* and *E. coli* in terms of cell numbers monitored by direct count, electronic count, and viable cell count as well as product formation (alcohol and acid production). *Source*: Reprinted with permission from reference 39, page 528, by courtesy of Marcel Dekker, Inc.

cell count and direct cell count. The electronic count reached log 8 mL^{-1} and then declined, exhibiting a trend similar to that observed in the yeast–*Acetobacter* interactions.

Alcohol contents increased to about 2%, and the pH first dropped to around 4 and then rose to 6. This pattern differs from that obtained from the interaction for yeast and *Acetobacter*. *E. coli* cannot oxidize alcohol and thus did not create large amounts of acid to counteract the basic metabolites in the reaction vessel. This resulted in a medium that reverted to a more alkaline state.

Mixed Bacteria and Mold Interaction

Another example of mixed culture interaction is ensilage of crops. Dalamacio and Fung (35) and Dalmacio et al. (36) studied the influence of ammonia on bacteria and mold during ensilage of high-moisture corn. Bacteria were not affected by ammonia treatment as much as mold at the onset of the treatment. Bacterial populations remained detectable after treatment of 1.0, 1.5, and 2% ammonia. However, no mold was detected for corn treated with 1.0, 1.5, 2.0% ammonia. Corn treated with 0.5% ammonia had high bacterial and mold counts. After 2 months of storage, the bacterial population increased to more than 10^7 g^{-1}. Mold count increased to around 10^4 g^{-1} for the corn treated with 0.5% ammonia. After 4 months of storage, the bacterial count reached saturation level (higher than 10^{10} g^{-1}) but the mold population remained at around 10^4 or 10^5 g^{-1}. Not only did the bacteria and mold counts change in the fermentation, but the genera of bacteria and mold also changed concomitant to the fermentation stages. At the beginning, most of the bacteria isolated were *Bacillus*, but as the fermentation progressed most of the bacterial isolates belonged to *Lactobacillus*.

At the beginning of the storage of ammonia-treated corn (in the fall season), the predominant mold was *Mucor*, which is considered a field mold. As the storage continued into the cold winter, *Penicillium* isolates increased, and at the end of the storage in spring, *Scopulariopsis* predominated. These successions occurred for a variety of reasons. For example, the winter months favored the development of *Penicillium*, which is cold tolerant. As temperature increased in the spring, *Penicillium* population was overtaken by *Scopulariopsis*, which can use complex nitrogenous foods better than *Penicillium*.

The aforementioned examples of microbial interactions are just some of the typical interaction patterns in food and in nature. Countless other interactions can occur. The

main point to emphasize is that mixed-culture interactions are very complex, but scientists can use this knowledge to optimize process control in obtaining desirable products in food and industrial microbiological processes.

ACKNOWLEDGMENTS

This research was partially supported by the U.S. Environmental Protection Agency (EPA) under assistance agreement R-819653 to the Great Plains–Rocky Mountain Hazardous Substance Research Center. It has not been submitted to the EPA for peer review and, therefore, may not necessarily reflect the views of the agency; no official endorsement should be inferred. The Center for Hazardous Substance Research and the Agricultural Experiment Station also provided partial funding.

BIBLIOGRAPHY

1. D.Y.C. Fung, in L.E. Erickson and D.Y.C. Fung eds., *Handbook on Anaerobic Fermentations*, Dekker, New York, 1988, pp. 3–25.
2. L.E. Erickson and D.Y.C. Fung eds., *Handbook on Anaerobic Fermentations*, Dekker, New York, 1988.
3. Gerald Reed ed. *Prescott and Dunn's Industrial Microbiology*, 4th ed, AVI Publishing, Westport, Conn., 1982.
4. K.H. Steinkraus, *Handbook of Indigenous Fermented Foods*, Dekker, New York, 1983.
5. J.C. Linden, in L.E. Erickson and D.Y.C. Fung eds., *Handbook on Anaerobic Fermentations*, Dekker, New York, 1988, pp. 59–80.
6. J.E. Bailey and D.F. Ollis, *Biochemical Engineering Fundamentals*, 2nd ed. McGraw-Hill, New York, 1986.
7. C.K. Matthews and K.E. van Holde, *Biochemistry*, Benjamin/Cummings, New York, 1990.
8. E.R. Vedamuthu, in L.E. Erickson and D.Y.C. Fung eds., *Handbook on Anaerobic Fermentations*, Dekker, New York, 1988, pp. 641–692.
9. L.E. Erickson, I.G. Minkevich, and V.K. Eroshin, *Biotechnol. Bioeng.* **20**, 1595–1621 (1978).
10. R. Luedeking and E.L. Piret, *Biotechnol. Bioeng.* **1**, 393–412 (1959).
11. R. Luedeking and E.L. Piret, *Biotechnol. Bioeng.* **1**, 431–459 (1959).
12. M.D. Oner, L.E. Erickson, and S.S. Yang, in L.E. Erickson and D.Y.C. Fung eds., *Handbook on Anaerobic Fermentations*, Dekker, New York, 1988, pp. 293–323.
13. A.H. Stouthamer, in L.E. Erickson and D.Y.C. Fund eds., *Handbook on Anaerobic Fermentations*, Dekker, New York, 1988, pp. 345–437.
14. A.J.B. Zehnder, *Biology of Anaerobic Microorganisms*, Wiley, New York, 1988.
15. J.H. Thiele and J.G. Zeikus, in L.E. Erickson and D.Y.C. Fung eds., *Handbook on Anaerobic Fermentations*, Dekker, New York, 1988, pp. 537–595.
16. J.H. Thiele, in J.G. Zeikus and E.A. Johnson, eds., *Mixed Culture Biotechnology*, McGraw-Hill, New York, 1991, pp. 261–292.
17. S.G. Pavlostathis, G. Misra, M. Prytula, and D. Yeh, *Water Environ. Res.* **68**, 479–497 (1996).
18. J.C. Spain ed. *Biodegradation of Nitroaromatic Compounds*, Plenum, New York, 1995.
19. L.L. Barton ed. *Sulfate-Reducing Bacteria*, Plenum, New York, 1995.
20. J.G. Zeikus and E.A. Johnson eds., *Mixed Culture Biotechnology*, McGraw-Hill, New York, 1991.
21. H.J. Peppler, in J.G. Zeikus and E.A. Johnson eds., *Mixed Culture Biotechnology*, McGraw-Hill, New York, 1991, pp. 17–36.
22. S.B. Leschine, in L.N. Ornston, A. Balows, and E.P. Greenbert eds., *Ann. Rev. Microbiol.* **49**, 399–426 (1995).
23. J.C. Spain, in L.N. Ornston, A. Balows, and E.P. Greenberg eds., *Ann. Rev. Microbiol.* **49**, 523–555 (1995).
24. G.S. Omenn ed. *Environmental Biotechnology*, Plenum, New York, 1988.
25. M.L. Rochkind-Dubinsky, G.S. Sayler, and J.W. Blackburn, *Microbial Decomposition of Chlorinated Aromatic Compounds*, Dekker, New York, 1987.
26. H.I. Adler and W.D. Crow, *Biotechnol. Bioeng. Symp.* **11**, 533–540 (1981).
27. H.I. Adler and G. Spady, *J. Rapid Methods Automat. Microbiol.* **5**, 1–12 (1997).
28. American Society for Microbiology, *Fundamentals of Anaerobic Bacteriology as Related to the Clinical Laboratory*, American Society for Microbiology, Washington, D.C., 1980.
29. K.L. Anderson and D.Y.C. Fung, *Proc. of the 42nd Ann. Mtg.*, Inst. of Food Technologists, Abstract 163, Las Vegas, June 22–25, 1982.
30. K.L. Anderson and D.Y.C. Fung, *J. Food Protect.* **40**, 811–827 (1983).
31. M.S. Ali, D.Y.C. Fung, and C.L. Kastner, *J. Food Sci.* **56**, 367–370 (1991).
32. F.M. Aramouni, K. Kone, J.A. Craig, and D.Y.C. Fung, *J. Food Protect.* **57**, 882–886 (1994).
33. J.H. Brewer and D.L. Allgeier, *Appl. Microbiol.* **14**, 985–988 (1966).
34. M.E. Claerbout, *J. Rapid Methods Automat. Microbiol.* **5**, 21–28 (1997).
35. I.F. Dalmacio and D.Y.C. Fung, Kalikasan, Philippines. *Journal of Biology* **6**, 203–216 (1977).
36. I.F. Dalmacio, D.Y.C. Fung, and J.H. McNeil, *Philippine Phytopathol.* **18**, 102–107 (1982).
37. D.Y.C. Fung, in C.M. Bourgeois and J.V. Leveau eds., *Microbiology Control for Food and Agricultural Products*, VCH Publishers, New York, 1995, pp. 1–25.
38. D.Y.C. Fung, *Food Technol.* **49**, 64–67 (1995).
39. D.Y.C. Fung, M.A. Buono, and L.E. Erickson, in L.E. Erickson and D.Y.C. Fung eds., *Handbook on Anaerobic Fermentations*, Dekker, New York, 1988, pp. 501–536.
40. D.Y.C. Fung and P.A. Hartman, in C.G. Heden and T. Illeni eds., *New Approaches to the Identification of Microorganisms*, Wiley, New York, 1975.
41. D.Y.C. Fung and C.M. Lee, *Food Sci. (Republic of China)* **7**, 209–213 (1981).
42. D.Y.C. Fung, L.S.L. Yu, F. Niroomand, and K. Tuitemwong, in R.C. Spencer, E.P. Wright, and S.W.B. Newsom eds., *Rapid Methods and Automation in Microbiology*, Inercept Limited, Andover, UK, 1994, pp. 313–326.

43. L.S. Gall and P.E. Riely, *Manual for the Determination of the Clinical Role of Anaerobic Microbiology*, CRC, Boca Raton, Fla., 1981.

44. I.C. Hall, *J. Bacteriol.* **17**, 255–301 (1929).

45. L.V. Holdeman, E.P. Cato, and W.E.C. Moore, *Anaerobe Laboratory Manual*, 4th ed., Virginia Polytechnic Institute and State University, Blacksburg, Vir., 1977.

46. R.E. Hungate, *J. Bacteriol.* **53**, 631–645 (1947).

47. R.L. Hungate, *Bacteriol. Rev.* **14**, 1–49 (1950).

48. R.E. Hungate, in J.R. Norris and D.W. Ribbons eds., *Methods in Microbiology*, Vol. 3B, Academic Press, New York, 1969, pp. 117–132.

49. R.W. Kelly and T.C. Schipp, in W.P. Olson ed., *Automated Microbial Identification and Quantitation: Technologies for the 2000s*, Interpharm Press, Buffalo Grove, Ill., 1996.

50. N.R. Krieg and J.G. Holt, *Bergey's Manual of Systematic Bacteriology*, Vol. 1, Williams and Wilkins, Baltimore, 1984.

51. B.M. Mitruka, *Methods of Detection and Identification of Bacteria*, CTC, Cleveland, Ohio, 1976.

52. N.J. Moon and G.W. Reinbold, *J. Milk Food Technol.* **39**, 337–341 (1976).

53. W.E.C. Moore, *Inst. J. Syst. Bacteriol.* **16**, 173–190 (1966).

54. J.E. Ogg, S.Y. Lee, and B.J. Ogg, *Can. J. Microbiolol.* **25**, 987–990 (1979).

55. W.P. Olson, *Automated Microbial Identification and Quantitation: Technologies for the 2000s*. Interpharm Press, Buffalo Grove, Ill., 1996.

56. D.A. Shapton and R.G. Board, *Isolation of Anaerobes*, Academic Press, New York, 1971.

57. V.L. Sutter, D.M. Citron, and S.M. Finegold, *Anaerobic Bacteriology Manual*, C.V. Mosby, St. Louis, Mo., 1980.

58. M. Thurston and C. Gannon, *J. Rapid Methods Automat. Microbiol.* **5**, 13–20 (1997).

59. C. Venderzant and D. Splittoesser, *Compendium of Methods for the Microbiological Examination of Foods*, Amer. Public Health Assoc., Washington, D.C., 1992.

60. T.D. Wilkins and C.B. Walker, *Appl. Microbiol.* **30**, 825–830 (1975).

61. T.D. Wilkins, C.B. Walker, and W.E.C. Moore, *Appl. Microbiol.* **30**, 831–837 (1975).

62. J.C. Zeikus, *Bacteriol. Rev.* **41**, 514–541 (1977).

63. A.C. Sonnewirth, *Am. J. Clin. Nutr.* **25**, 1295–1298 (1972).

64. K. Horikoshi and W.D. Grant eds., *Extremophiles: Microbial Life in Extreme Environments*, Wiley, New York, 1998.

65. C.R. Woese, O. Kandler, and M.L. Wheelis, *Proc. Natl. Acad. Sci. USA* **87**, 457–466 (1990).

66. C.R. Woese, *Microbial. Rev.* **51**, 221–271 (1987).

67. M.R. Mormile, S.H. Lin, and J.M. Sulfita, *Environ. Sci. Technol.* **28**, 1727–1732 (1994).

68. R.J. Bothast and B.C. Saha, *Adv. Appl. Microbiol.* **44**, 261–286 (1997).

69. R.M. Atlas, *Handbook of Media for Environmental Microbiology*, CRC, Boca Raton, Fla., 1995.

See also BIOENERGETICS OF MICROBIAL GROWTH; CLOSTRIDIA, SOLVENT FERMENTATION; PROTEIN EXPRESSION, SOLUBLE; YEAST, BAKER'S.

ANAEROBES, INDUSTRIAL USES

MAHENDRA K. JAIN
J. GREGORY ZEIKUS
MBI International
Lansing, Michigan

KEY WORDS

Anaerobes
Enzymes
Extremophiles
Metabolic diversity
Methanogenesis
Organic acids
Solvents
Waste treatment

OUTLINE

INTRODUCTION

Anaerobes, in the simplest form, are microorganisms that are unable to grow at the surface of a medium exposed to air, where the oxygen may either be bacteriocidal or bacteriostatic. These organisms do not use oxygen for catabolic reactions, and oxygen is not required for their survival. Anaerobes, with the exception of autotrophic methanogenic and acetogenic bacteria, sulfur-metabolizing archaea, and sulfate-reducing bacteria, obtain their energy by fermentation, a redox process in which organic substances serve as electron donors and terminal electron acceptors. Obligate anaerobes share two common characters: an extreme sensitivity to molecular oxygen, and a capacity to produce energy and perform essential biosyntheses without molecular oxygen. The sensitivity to O_2 is extremely variable from one group to another, and evaluation of O_2 sensitivity requires that all organisms be cultivated under identical conditions, both for medium composition (including reducing substances) and gaseous environment. The physiological, ecological, and genetic diversity of anaerobic bacteria makes them very interesting organisms.

Unlike aerobic microorganisms whose anabolic features are largely exploited for production of single-cell proteins, antibiotics, and amino acids, anaerobes offer the potential for utilizing catabolic features in the production of specialty and bulk chemicals, fuels, and unique enzymes that are stable and active under extreme environmental conditions. These properties of anaerobes make them highly desirable for certain industrial uses. Today, a large body of information is available on the metabolic diversity of anaerobes. Anaerobes seem to have adapted from the primordial atmospheric conditions to inhabit both moderate and extreme environments with metabolic machinery that is active from about 0 °C to over 100 °C, in very little salt to saturated salt conditions, in highly acidic (pH < 2) to alkaline (pH > 9) conditions, and even in the presence of highly toxic chemicals. In this chapter, we will only cover procaryotes and archaea, not anaerobic eucaryotes (i.e., yeasts and fungi). We will include examples of pure cultures used to produce specific products and mixed cultures or consortia used in environmental processes.

PHYSIOLOGY, BIOCHEMISTRY, AND MOLECULAR BIOLOGY

Species Diversity

Knowledge about microbial diversity is essential in understanding the relationship between environmental factors and metabolic functions. Such knowledge can be used to assess the effects of environmental conditions on production of microbial products. However, relatively little is known about the entire spectrum of anaerobic bacteria that may be of importance to various industries. What is known about anaerobic species diversity is presented in *Bergey's Manual* (1), where over 170 genera and over 800 species of anaerobic bacteria and archaea are described.

Cultivation of microorganisms greatly helps in determining the full complement of genetic and physiological diversity of newly isolated species. Strains available in pure and mixed culture can be examined to determine what novel features they have, including those that might be of special interest to industry. Thus, this aspect of microbial diversity is particularly important. Pure cultures are suitable for determination of morphological, physiological, and metabolic characteristics and commercial potential. The 16S rRNA sequence can provide a powerful means to organize the phylogenetic relationships among different microbial species in a meaningful way, but this analysis is often considered too conservative to provide species-level distinctions because any two different species can show 97–100% 16S rRNA sequence homology.

Initially, only mesophilic anaerobic species were isolated and characterized from moderate environments. A good example is the rumen, which provides a great variety of anaerobic species with different metabolic properties. However, in recent years, studies on microbial diversity in extreme ecosystems, such as thermal springs and deep-sea vents, have offered an unparalleled diversity in extremophile communities. The Yellowstone National Park ecosystem holds the world's greatest diversity of accessible, extreme microbial habitats. Not only are there temperature and pH extremes, but also areas with high levels of sulfur, heavy metals, radionuclides, and UV radiation, as well as a range of organic nutrients. Since oxygen solubility in water decreases as the temperature increases, anaerobes are common at the highest temperatures. Interest in novel microorganisms thriving under environmental extremes reflects a growing awareness of their value in biotechnological and industrial processes. Not only are the enzymes and metabolites often more stable, but in some cases, novel enzymes, not known in organisms from moderate habitats, have been found in extremophiles (2,3).

The range of anaerobe diversity has proven to be extraordinary. For example, following the discovery of the deep-sea hydrothermal vents, hyperthermophilic anaerobes from hot springs and volcanic fields, as well as from shallow and abyssal marine hydrothermal vents were identified (4,5). Numerous *Archaea* have been identified and classified (6), and some species produce enzymes that have, or are expected to have, commercial utility (2,7).

Table 1 provides some examples of representative anaerobic bacteria and archaea that illustrate the spectrum of anaerobic bacteria that live under extreme environmental conditions. The different ecophysiologies of these species explain, in part, the great diversity among anaerobic microorganisms, as it is the function of the organism or its enzyme(s) and gene(s) that is practiced by industry.

Metabolic Diversity

Microbial diversity at the biochemical level is more important from the industrial perspective. During the past decade, molecular biological and genetic approaches have come to dominate the study of microorganisms. Although fruitful, exclusive reliance on such approaches poses both practical and philosophical problems for microbiology. After an early phase of discounting the importance of physiology, people in industry are learning from the "biotechnology revolution" that gene cloning and biochemical manipulations do not solve all the product-development problems that may arise. Some of the more valued intellectual benefits anticipated in the field will emerge from understanding the physiology of peculiar and diverse microorganisms from unusual environments. As long as oxygen is present as the electron acceptor, denitrification, sulfate reduction, and methanogenesis are inhibited. Denitrification begins after consumption of oxygen, and sulfate reduction begins after consumption of nitrate. Finally, methanogenesis occurs, which is partly a dissimilatory carbon dioxide reduction. The sequence of these reactions agrees with the sequence of their free-energy changes, which decrease from respiration to methanogenesis.

Table 2 lists some representative anaerobic bacteria that display different metabolic types based on their catabolic pathways. These different biochemistries explain, in part, the great diversity of anaerobic microorganisms.

Methanogenesis. Methanogenic bacteria are strictly anaerobic archaea with a unique form of energy metabolism involving the generation of methane. Biological methanogenesis is an important component of the carbon cycle in a variety of anaerobic habitats. It represents the terminal

Table 1. Representative Bacterial and Archaeal Species of Anaerobic Extremophiles

Extremophiles	Subgroup	Optimal growth condition	Representative species
Thermophile	Moderate	$60\ ^{\circ}C < T_{opt} > 80\ ^{\circ}C$	*Clostridium thermocellum*
		$60\ ^{\circ}C < T_{opt} > 70\ ^{\circ}C$	*Methanobacterium thermoautotrophicum*
	Hypher	$T_{opt} \geq 80\ ^{\circ}C$	*Thermotoga maritima*
			Pyrococcus furiosus
Halophile	Moderate salinity	$[NaCl]_{opt} \geq 1.8$ M	*Desulfohalobium retbaense*
	High salinity	$[NaCl]_{opt} = 3.0$ M	*Halomethanococcus doii*
Acidophile		$pH_{opt} \leq 4.0$	*Sarcina ventriculi*
Alkaliphile		$pH_{opt} \geq 9.2$	*Methanohalphilus zhilinae*
Syntrophile		Grows with a metabolic partner (e.g., a methanogen or a sulfate reducer)	*Syntrophospora bryantii*
Toxophile		Carbon monoxide as a substrate	*Clostridium ljungdahlii*

Table 2. Potential Industrial Fermentation Products from Anaerobic Bacteria

Fermentation product	Substrate	Typical species	Concentration (g/L)	Status	Reference
Ethanol	Glucose	*Zymomonas mobilis*	84.5	Bench-scale	Kesava et al. (8)
	Paddystraw	*Clostridium thermocellum*	23.6	Bench-scale	Sudha Rani et al. (9)
	CO	*Clostridium ljungdahlii*	48	Bench-scale	Klasson et al. (10)
n-Butanol	Starch/glucose	*Clostridium acetobutylicum*	20	Bench-scale	Glassner et al. (11)
	CO	*Butyribacterium methylotrophicum*	2.7	Bench-scale	Grethlein et al. (12)
1,3-Propanediol	Glycerol	*Clostridium butyricum*	58	Pilot-scale	Günzel et al. (13)
Lactic acid	Glucose	*Lactobacillus acidophilus*	82	Industrial	Datta (14)
Acetic acid	Glucose	*Clostridium thermoaceticum*	83	Bench-scale	Parekh and Cheryan (15)
Butyric acid	Glucose	*Clostridium tyrobutyricum*	62.8	Bench-scale	Fayolle et al. (16)
Propionic acid	Glucose	*Propionibacterium acidipropionici*	57	Bench-scale	Paik and Glatz (17)
	Glycerol	*Propionibacterium acidipropionici*	42	Bench-scale	Barbirato et al. (18)
Succinic acid	Glucose	*Actinobacillus succinogenes*	105	Pilot-scale	Guettler et al. (19)
Caproic acid	Cellulose & ethanol	*Clostridium kluyveri* plus *Fibrobacter succinogenes*	4.6	Bench-scale	Kenealy et al. (20)

step in the anaerobic breakdown of organic matter under sulfate-limiting conditions. The substrate range for methanogenesis is limited to carbon dioxide, formate, carbon monoxide, methanol, methylamines, and acetate, but no single isolate is able to utilize all these carbon sources (21). The metabolic pathways of methane formation are unique and involve a number of enzymes and coenzymes that occur only in methanogens. A primary Na^+ gradient and H^+ gradients are formed by interesting and unique membrane-bound enzymes, and the reactions by which these ion-motive forces are generated are also novel (22). The organisms and the methanogenic reactions (21,23), and the pathways of energy conservation, have been extensively reviewed (22).

In recent years, bioconversion of agricultural and industrial wastes to methane has been considered important from two perspectives, as a waste-disposal method for residual organic matter and as a source of energy. Anaerobic digestors for treatment of sewage, agricultural and animal waste, and industrial effluent are being used worldwide. More recently, methanogenic microbial communities have been found to be very efficient in treating toxic organic wastes. For example, transformation of C_1 and C_2 halocarbons was observed in the early 1980s and, as one of the mechanisms responsible for these transformations, was proposed as a biologically mediated reductive dechlorination (24–26). Anaerobic dechlorination and degradation of chlorinated aromatic compounds such as chlorophenols, chlorobenzenes, and polychlorinated biphenyls (PCBs) occurs in a wide variety of environments. Methanogenic microbial consortia that dechlorinate tetrachloroethylenes (PCE) and trichloroethylenes (TCE), and dechlorinate, degrade, and mineralize chlorophenols and PCBs have been developed (27–31). The potential of these consortia in bioremediating the contaminated soil and sediment systems is being examined. DDE [1,1-dichloro-2,2-bis(*p*-chlorophenyl)ethane], a metabolite of the pesticide DDT, has been observed to reductively dechlorinate to DDMU at a relatively high rate under methanogenic conditions as compared to sulfidogenic conditions (32).

The catabolic reactions carried out by methanogens yield very little energy compared to those of aerobes. For example, use of H_2 to reduce CO_2 to CH_4 has a $\Delta G^{0'}$ of -34 kJ/mol H_2, whereas the corresponding energy value for an aerobe oxidizing H_2 using O_2 as an electron acceptors is -237 kJ/mol H_2. Most of the known species of methanogens can use H_2 to reduce CO_2 to CH_4. Many of the hydrogenotrophic methanogens also utilize formate. In contrast, only species belonging to the genera *Methanosarcina* and *Methanosaeta* use acetate (acetotrophic) and convert it to methane. Acetate is an important end product of many fermentative anaerobes, and is a primary methanogenic sub-

strate in anaerobic digestors. The methylotrophic methanogens can utilize several simple methylated compounds. These methanogens include *Methanosarcina*, which can also use acetate and usually H_2—CO_2, and *Methanolobus* and *Methanococcoides*, which are only known to grow on methylated compounds. Carbon monoxide is also converted to methane by methanogenic species. Methanogenesis has been observed at low, moderate, and high temperatures, and in neutral to mildly acidic or alkaline conditions. In summary, methane is produced under a variety of environmental conditions and is considered a byproduct of the anaerobic treatment of industrial wastes.

Solventogenesis. Ethanol production is a very well-known fermentation process since it has been an outgrowth of the alcoholic beverages industry. During World War I, shortages of acetone in the manufacture of munitions led to the development of an acetone–butanol–ethanol process involving the fermentation of starch by *Clostridium acetobutylicum*. This fermentation gave approximately 2 parts butanol to 1 part acetone and 1 part ethanol. During the 1940s and 1950s, lower costs of petrochemical processes stopped butanol production by fermentation routes in the U.S. and Europe. The process was also stopped in South Africa, but continued in China into the early 1990s.

The present market for ethanol is met by yeast fermentation, but the industrial market for isopropanol, acetone, and butanol could be met by other anaerobic fermentations. Biomass-derived solvents produced by fermentation can enter into the current petrochemical synthetic pathways through a number of reactions. The most important of these is the dehydration of alkanols to alkene to form ethylene, propylene, butylene, and butadiene. Therefore, production of chemical feedstocks from biomass via fermentation is becoming increasingly attractive because biomass production costs are not as tightly bound to energy costs.

A great diversity exists among the solvent-producing microorganisms. These organisms span several groups of yeasts and a broad range of both mesophilic and thermophilic, and aerobic and anaerobic, bacteria. By far, the ethanol-producing organisms are the most abundant solventogens. However, industrial use is mainly limited to strains of yeasts, but great potential exists for recombinant *Zymomonas mobilis*. Several clostridia are capable of producing butanol and either acetone or isopropanol with ethanol as a byproduct. The *C. acetobutylicum* and *Clostridium beijerinckii* strains are of prime importance for these industrial fermentations.

Clostridium thermocellum, *Thermoanaerobacter ethanolicus*, *Clostridium thermosaccharolyticum*, *Thermoanaerobacterium thermosulfurigenes*, *Thermoanaerobacter brockii*, and *Thermobacteroides acetoethylicus*, use a wide range of substrates, from polymeric carbohydrates such as cellulose, pectin, xylan, and starch, to mono- and disaccharides such as glucose, cellobiose, xylose, and xylobiose. The primary fermentation product is ethanol, with the production of acetate, lactate, carbon dioxide, and hydrogen in various ratios. *T. acetoethylicus* has not been shown to produce lactic acid. At present, one of the major limitations in the use of thermoanaerobes is the variability of end-product ratios, yield, and low ethanol concentration. These are affected by species, enzyme complement, and environmental conditions. Certain strains of *T. ethanolicus* have the best conversion of carbohydrates to ethanol, forming 1.6–1.9 mol/mol of glucose fermented (33). High ethanol and hydrogen concentrations also reduce the yield of ethanol (33) owing to the flexibility of the carbon and electron pathways, which may possess many reversible enzyme systems (33,34).

In addition to ethanol, several saccharolytic clostridia produce butanol and acetone besides volatile fatty acids and gaseous products from carbohydrate fermentation. In some cases acetone is further reduced to isopropanol. In most species, the production of solvents only occurs late in the fermentation cycle, following a shift from the pathways leading to acetate and butyrate production. The production of butanol and ethanol is usually associated with the uptake and reutilization of acids, and the production of acetone or butanol. The ability to produce solvents is influenced by the type and concentration of substrate, the pH and the buffering capacity of the culture medium, and the environmental conditions. Several strains that have the ability to produce solvents undergo degenerative changes, resulting in the loss of their ability to produce solvents. The solvent-producing clostridia are found in a wide variety of natural habitats (35); however, the strains of *C. acetobutylicum* and *Clostridium beijerinckii* are the most frequently studied species. *C. acetobutylicum*, which characteristically produces butanol, acetone, and ethanol in the ratio of 6:3:1, has been used extensively for the industrial production of solvents. Most butanol producers are mesophiles with a temperature optima for fermentation between 30 and 37 °C. Recently, solvent yield was shown to increase at lower temperatures and acetone but not butanol yield to decrease at elevated temperatures (36; B.K. Soni and M.K. Jain, unpublished data 1991).

At present, many ethanol- and butanol-producing strains of *Clostridium* remain poorly classified, and there is still no accepted standard classification for the clostridia group as a whole. Strains of *C. beijerinckii* (formerly classified as *C. butylicum*) constitute a second group of solvent producers that do not show DNA homology with the *C. acetobutylicum* group. This group contains both high- and low-solvent-producing strains that produce either acetone or isopropanol in addition to butanol. The strains of *Clostridium aurantibutyricum* group also include butanol-producing strains that produce both acetone and isopropanol (37). *Clostridium tetanomorphum* produces butanol and ethanol, but not acetone or isopropanol. As obligate anaerobes, butanol producers require anaerobic conditions. However, vegetative cells of *C. acetobutylicum*, for example, survived several hours exposure to oxygen and formed fruiting-body-like structures on agar plates when exposed to oxygen (38).

Acetogenesis. The term acetogen is commonly applied to a bacterium that forms acetate, whether the acetate is produced by a catabolic process such as fermentation or by an autotrophic-type synthesis. Acetogenic bacteria are ubiquitous in anaerobic ecological systems. Homoaceto-

gens have the ability to grow well on H_2 plus CO_2 and poorly on CO, and to form acetyl CoA by a CO-dependent pathway involving CO dehydrogenase (39). *Clostridium thermoaceticum* and *Clostridium thermoautotrophicum* are homoacetogens and synthesize acetate from C_1 compounds by the recently established Wood pathway of acetyl CoA synthesis (40). This pathway was established by studies with *C. thermoaceticum* that can grow on CO or CO_2—H_2, utilizing them as carbon and energy sources (41).

In industrial fermentations, the formation of a single product is advantageous because its recovery would be simplified. In this respect, homoacetogenic bacteria essentially form only acetate in the fermentations of hexoses and pentoses. They appear to be ideal for the microbial production of acetate. Energy for cell growth is obtained by the reduction of CO_2 to acetate via the Wood pathway. The homoacetogens also grow on other one-carbon compounds, including CO, formate, and methanol with acetate as the product. In the acetyl CoA pathway for synthesis of acetate, acetyl CoA is the first two-carbon intermediate of the autotrophic fixation of CO_2, and is either used for the synthesis of cell carbon or converted to acetate.

Butyribacterium methylotrophicum is an acetogenic anaerobe that can grow on multicarbon compounds as well as on one-carbon compounds (e.g., CO or methanol). It produces acetate or butyrate using the Wood pathway, but acetyl CoA can be condensed and reduced to butyrate. Higher levels of NADH were found in butyrate-producing than in acetate-producing cells. In *B. methylotrophicum*, butyrate production is regulated by the carbon source and is dependent on cellular NADH/NAD ratios, and the levels and direction of ferredoxin- and NAD-linked oxidoreductases (42). Also, the growth pH regulates both hydrogenase and FD-NAD oxidoreductase activities such that, at acid pH, more intermediary electron flow was directed towards butyrate synthesis than H_2 production.

Acidogenesis. The acidogenic bacteria include those bacteria that only form acetate by fermentation. These bacteria invariably produce other acids in addition to acetate. The acidogenic bacteria grow heterotrophically on a variety of carbohydrates, producing a mixture of acids such as lactic, propionic, butyric, succinic, and caproic. The pathways for production of these acids have been reviewed (43). A generalized scheme for anaerobic production of fermentative organic acids and alcohols is presented in Figure 1. The acidogens differ from the acetogens in that they do not synthesize acetate from CO_2 or other C_1 compounds by the Wood autotrophic pathway. Many of the species can utilize hexoses such as glucose, fructose, galactose, and mannose; disaccharides including cellobiose, lactose, and maltose; pentoses represented by xylose; and other substrates such as glycerol, mannitol, sorbitol, inositol, and polymeric carbohydrates. For example, *Clostridium butyricum*, the type strain of the genus *Clostridium*, is saccharolytic in nature and metabolizes glucose, producing butyrate, acetate, CO_2, and H_2 as fermentation products. Glucose is converted to pyruvate by the Embden–Meyerhof–Parnas glycolytic pathway. Pyruvate is then simultaneously decarboxylated and oxidized by the enzyme complex pyruvate–ferredoxin oxidoreductase to yield acetyl CoA, CO_2, and reduced ferredoxin. The reduced ferredoxin is reoxidized in several reactions, the most important of which involves H_2 evolution catalyzed by hydrogenase. Acetyl CoA is condensed and reduced to butyrate.

In some other acidogens, phosphoenolpyruvate (PEP) serves as a branching point in the pathway producing pyruvate and oxaloacetate. The PEP–oxaloacetate step involves CO_2 fixation mediated by PEP carboxykinase. Depending upon the species, products such as succinate and propionate are produced as fermentation products with acetate, formate, or H_2. For example, oxaloacetate is converted to propionate in species like *Propionispira arboris* (44), whereas, oxaloacetate is further converted to succinate in species like *Anaerobiospirillum succiniciproducens* (45). Lactic acid bacteria such as *Lactobacillus acidophilus* convert pyruvate to lactic acid, whereas *Clostridium kluyvri* produces caproic acid from ethanol and acetate (43).

Sulfidogenesis. Sulfate, a chemically inert, nonvolatile, and nontoxic compound, is widespread in rocks, soil, and water. In contrast, hydrogen sulfide, because of its chemical properties and physiological effects, is a far more conspicuous substance and is chemically reactive. The presence of black-colored sediments due to the formation of ferrous sulfide from iron-containing materials, accompanied by a distinctive smell, is indicative of H_2S production. Hydrogen sulfide acts as a reductant, and is toxic to plants, animals, and humans. It is produced by reduction of sulfate by sulfate-reducing bacteria, which are present in ecological niches where oxygen has no access. In the absence of oxygen, the oxidized form of sulfur (SO_4^{2-}) is used as an electron acceptor by sulfate-reducing bacteria (sulfate reducers). Since reduction of the inorganic compound serves for energy conservation, the process is distinguished as dissimilatory reduction from the assimilatory reduction of sulfate in plants and bacteria.

Under anaerobic, reduced conditions, hydrogen sulfide is the energetically stable form of sulfur, as sulfate is the stable form under aerobic conditions. The most important reducers of elemental sulfur in nature are probably special anaerobes that may be designated as true sulfur-reducing bacteria; these carry out the dissimilatory reduction of sulfur as their primary or even obligate metabolic reaction during oxidation of organic substrates. The true sulfur-reducing bacteria do not reduce sulfate, and only some reduce other oxygen-containing sulfur compounds. The majority of the eubacterial sulfur reducers are mesophilic, as are most sulfate reducers. In contrast, all described archaebacterial sulfur reducers are hyperthermophiles with temperature optima (up to 110 °C) the highest known in the living world.

Sulfate-reducing bacteria include a rather heterogenous assemblage of microorganisms having in common merely dissimilatory sulfate metabolism and obligate anaerobiosis. The genus *Desulfovibrio* includes species that are still the best-studied sulfate reducers. These are nonsporing, having curved motile cells and growing on a relatively limited range of organic substrates, preferably lactate or pyruvate, which are incompletely oxidized to acetate and H_2S. Spore-forming sulfate-reducing bacteria with a similar metabolism were classified within the genus

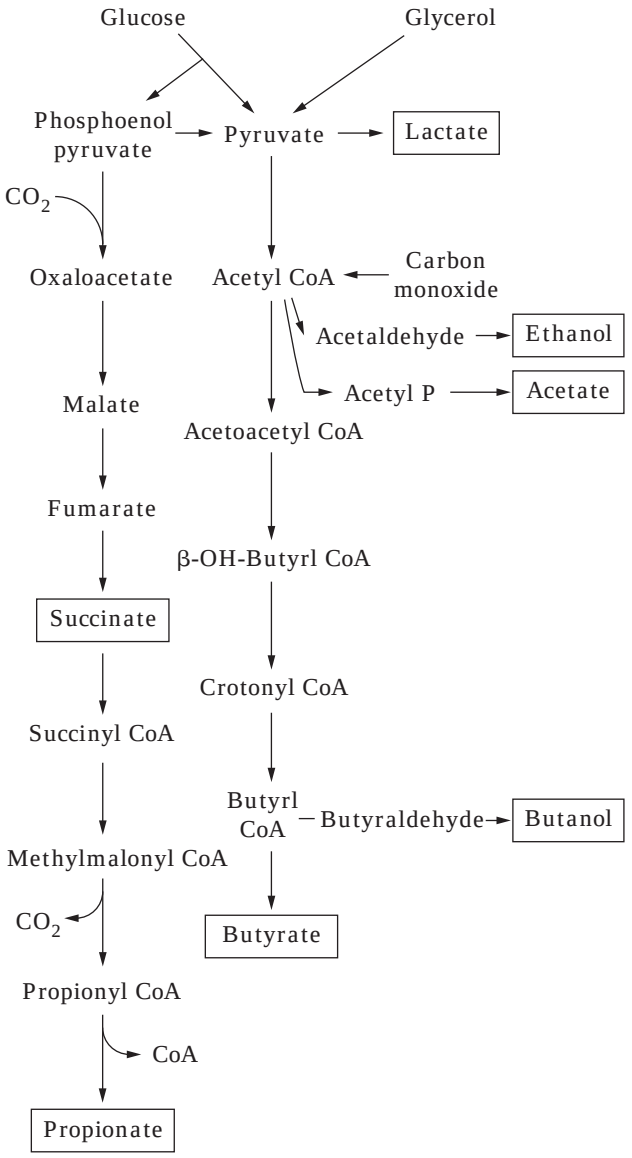


Figure 1. Generalized scheme for anaerobic fermentative production of common organic acids and alcohols.

Desulfotomaculum. Other types of sulfate-reducing bacteria differ markedly physiologically and morphologically from the known *Desulfovibrio* and *Desulfotomaculum* species. The reactions by which sulfate-reducing bacteria are involved in anaerobic degradation are indicated by their metabolic capacities. As far as tested, these bacteria use low molecular weight compounds as electron donors and therefore depend on fermentative bacteria that degrade the original polymers from biomass. Thus sulfate reducers are terminal degraders and their role is analogous to that of methanogenic bacteria that form methane and carbon dioxide as final anaerobic products. The recognition of dissimilatory sulfate reduction and methanogenesis as two alternative terminal carbon degradation processes has contributed a great deal to our understanding of anaerobic mineralization.

Genes and Genetic System

The field of genetics has been restricted for many years to *Escherichia coli* and few other genera of aerobic or facultative anaerobic bacteria such as *Pseudomonas*, *Bacillus*, and *Salmonella*. Anaerobic bacteria were known and studied since before 1900, but work on the genetics and molecular biology of anaerobic bacteria began to emerge only in the 1970s. The volume by Sebald is the most recent comprehensive review on genetics of anaerobes (46). Improvements in basic techniques for culturing anaerobes and recent advances in molecular biology techniques have helped in making rapid progress in understanding genes and genetic systems in anaerobic bacteria. Thermoanaerobic bacteria produce many thermostable enzymes but the yields are low. However, their enzymes can be overproduced by cloning the genes into mesophilic, industrially important aerobic bacteria. The genes can be overexpressed in both gram-positive and gram-negative hosts such as *Bacillus subtilis* and *E. coli* (47). One major advantage of cloning genes for thermostable enzymes in mesophiles is in the recovery of enzymes in greater than 95% purity by a single-step high-temperature treatment that inactivates host

proteins, including protease enzyme (47). Aerobic bacteria produce about 10-fold higher biomass yields than the anaerobic bacteria. Therefore, there is no advantage, except for biotransformation, in producing the enzymes in an anaerobic system.

Several species of the genus *Clostridium* are of biotechnological interest because of the end products of their fermentative metabolism, the stereospecific reductions that they undertake, and the potentially important enzymes that they produce. Other species, such as *C. botulinum*, produce some of the most powerful toxins known today, which have been commercialized for medical treatments such as the correction of crossed eyes.

Various elements of gene-transfer technology have been developed in several species. Two gene-transfer procedures should prove widely applicable throughout the genus—electroporation and conjugal plasmid mobilization. In clostridia, mutants defective in a variety of functions, including purine, pyrimidine, vitamin, and amino acid biosynthesis, pathways of fermentative metabolism, and sporulation, have been isolated with ethyl methane sulfonate, *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine, and ultraviolet light, and have been characterized (48–52). These mutant strains provide valuable information about gene regulation and function. Conjugal transfer of plasmids and transposons as well as phase transfection/plasmid transformation has now been documented in several species. In addition, recombinant DNA technology has also been employed to gain insights into the gene-transfer systems in *Clostridium*. Protoplasts of *C. acetobutylicum* can be transfected with phage DNA (53). The whole cells of *C. acetobutylicum* can be transformed with plasmid DNA using an electroporation procedure. Permealized cells of *Thermoanaerobacter thermohydrosulfuricus* have been transformed with plasmid pUB110 (encoding resistance to kanamycin, Km^R) and a derivative, pGS13, carrying a chloramphenicol resistance (Cm^R) marker (54).

A variety of cloning vectors have been developed for use in *C. acetobutylicum*. To date only pMTL500E has been used to introduce cloned heterologous genes into *C. acetobutylicum*. These genes included the *Clostridium pasteurianum leu*B gene (55) and the *C. thermocellum cel*A gene (see Ref. 56). Acquisition of the *cel*A gene was demonstrated by an in situ plate assay using Congo red. In most cases, screening of clostridial gene banks has relied on expression of the heterologous genes in *E. coli*. Appropriate *E. coli* mutants have been employed for the isolation of *C. acetobutylicum* genes involved in the acetone–butanol fermentation. The butyraldeyde dehydrogenase gene of *C. acetobutylicum* has been isolated by complementation of an *E. coli* aldehyde dehydrogenase–deficient mutant. Endo-*β*-glucanases and xylanse of *C. thermocellum* and other cellulolytic clostridia are easily detectable on plates containing the appropriate substrate by the Congo red assay. Upon staining with Congo red, positive clones are surrounded by a yellow hydrolysis zone on a red background.

BIOTECHNOLOGICAL APPLICATIONS

Industrial biotechnology has developed into a more than $25 billion business of enzymes and of chemically defined compounds produced by fermentation, microbial transformation, and enzymatic conversion. These include organic acids, amino acids, alcohols, vitamins, microbial polymers, antibiotics, industrial enzymes, biopesticides, and pharmaceutical specialities. Of these, ethanol, citric acid, L-lactic acid, gluconic acid, monosodium L-glutamate (MSG), and L-lysine are examples of bulk fermentation products.

Solvents

Solvents are compounds that are essentially neutral in character and comprise organic alcohols and ketones. The major solvent products of microbial carbohydrate catabolism include ethanol, acetone, isopropanol, butanol, acetoin, 1,3-propanediol, and butanediol. These are the most useful oxychemicals produced by fermentation. Solvents have various uses in the chemical industry, including the direct use as organic chemicals and indirect use as fuel additives, extenders, or feedstocks for further chemical synthesis. Butanol is primarily used as a feedstock chemical in the manufacture of lacquers, rayon, plasticizers, coatings, detergents, and brake fluids. It can also be used as a solvent for fats, waxes, resins, shellac, and varnish. In addition, butanol may also be used as an extractant and solvent in the food industry.

Butanol. The acetone–butanol fermentation currently has potential because butanol has many characteristics that makes it better than the currently used liquid-fuel extender, ethanol. Higher alcohols have several characteristics that are favorable for motor fuel use, either in gasoline blends or, in some cases, when used directly as fuel. They are miscible in gasoline; their heats of combustion per gallon are greater than those of methanol, and they have good octane-enhancing properties. Butanol, for example, has a heat of combustion 54% higher than methanol and 83% of that for gasoline (57,58). Although the octane numbers are less than that for methanol (RON 100 vs. 110), the research octane number (RON) value of 100 is still above that of gasoline (RON 92), and therefore useful for fuel blends. In addition, butanol has low miscibility with water and exhibits high miscibility with both diesel and gasoline. Owing to its high heat of combustion butanol solutions containing as much as 20% (v/v) water have the same combustion value as anhydrous ethanol.

In the early 1900s, butanol was used to produce butadiene, the most desirable raw material for synthetic rubber. The annual production of fermentation-derived butanol was over 45 million pounds during 1945. But the fermentation-derived butanol process declined after World War II in the U.S. due to both changes in availability of renewable feedstocks (molasses, sugarcane) and the increase in availability of inexpensive petrochemical feedstocks. Butanol fermentation of beet molasses continued through the 1970s in the Soviet Union, and fermentation of sugarcane molasses continued through the 1980s in South Africa. In China, butanol was still produced fermentatively until recently, but the last viable industrial acetone–butanol–ethanol (ABE) fermentation in the Western world was carried out by National Chemical Products

in Germiston, South Africa, using *C. acetobutylicum*.

Butanol is now synthesized chemically from petroleum-derived ethylene, propylene, and triethylaluminium, or carbon monoxide and hydrogen. The major domestic producers of butanol and its derivatives are BASF, Chem Service Inc., Dow Chemical, Eastman Chemical, Hoechst Celanese, Shell, Union Carbide, and Vista. The current U.S. production of butanol is more than 1.2 billion pounds per year and is experiencing a growth of 3–4% annually.

Butanol pricing has been rising steadily in recent years due to an increase in demand. The bulk butanol pricing in the United States is $0.40–$0.50 per pound. The price of butanol in Europe and the Far East is higher than $0.50 per pound. If the production cost of fermentation-derived butanol could be in the range of $0.25–$0.30 per pound, the market for butanol will likely expand: The market penetration of fermentation-based butanol will drive the utilization of corn from 7.2 million bushels in 1995 to 153 million bushels in 2010. In addition, the supply of other alternative feedstock, such as biomass, is more available than corn. However, biomass requires effective pretreatment technology that has not yet been developed. Biomass-derived solvents produced by fermentation can enter into the current petrochemical synthetic pathways through a number of reactions, the most important of which is the dehydration of alkanols to alkenes to form ethylene, propylene, butylene, and butadiene.

The ABE fermentation in batch culture is a sequential process characterized by a primary acid-producing or acidogenic phase that is coupled to growth, and a secondary solvent-producing or solventogenic phase. Thus, solventogenesis appears as a secondary metabolic process caused by uncoupling of growth. Acetate and butyrate are produced as soluble fermentation products during the acidogenic phase. As the acids accumulate, the growth rate and the culture pH decreases, and solventogenesis begins. The acids and H_2 are consumed and are reduced to solvents in the solventogenic phase. The "switch" from acidogenesis to solventogenesis is accompanied by a large number of physiological and biochemical changes, and is clearly a multifactorial process. The exact mechanism that regulates this switch has not been clearly elucidated, but the concentration of undissociated butyric acid within the cytoplasm functions as a "bioregulator." The end products of the ABE fermentation alter and inhibit growth and metabolism. Acids are more toxic than solvents, and the more hydrophobic a product, the higher is its toxicity. It should be noted that toxicity of acids depends upon the pH. At neutral pH, high concentrations of organic anions (i.e., butyrate, acetate) are not toxic because they are dissociated. In normal batch-fermentation conditions, organic acids are produced, and then consumed during solventogenesis to prevent dissipation of the cellular proton-motive force. Butanol inhibits the fermentation at a relatively low concentration.

Butanol producers use a wide variety of carbohydrates, including hexoses, pentoses, oligosaccharides, and polysaccharides as fermentation substrates, and including starch, molasses, whey, wood hydrolysates, pentosans, inulin, and sulfite liquor. Cellulosic materials are one of the most attractive substrates for chemicals, as they are the most abundant and least expensive raw material. Few attempts have been made to use cellulosic substrates for the production of butanol. A simultaneous saccharification fermentation process was studied to produce butanol from alkali-pretreated wheat straw using *Trichoderma reesi* cellulase and *C. acetobutylicum* (59). The final butanol concentration was 10.7 g/L from about 140 g of straw. Pentoses did not accumulate during the fermentation, suggesting simultaneous use of both hemicellulose and cellulose components. Different substrates alter fermentation parameters of *C. acetobutylicum*, such as growth rate and the solvent production ratio. Higher growth and substrate consumption rates were obtained when glucose, cellobiose, or mannose were used than when xylose, arabinose, or galactose were used. The solvent production ratio with the first sugar group was 1:4:10 (ethanol:acetone:butanol), whereas ratios of 1:2:5 were obtained with the latter.

Starch is at present the most useful raw material for fermentation because it is abundant, has high density, and requires limited pretreatments to be made from biomass. Starch is produced from corn, cereals (wheat, oats, rice), and starchy roots (sweet potatoes, cassava). Typically, *C. acetobutylicum* produces 12–13 g/L butanol with about 20–22 g/L total solvents. In recent years, the majority of investigators have focused their interests on the type culture strain *C. acetobutylicum* ATCC 824. Other cultures included DSM 792 strain and NCIMB 8052. The *C. acetobutylicum* ATCC 4259 strain and the equivalent DSM 1731, NCIMB 619, and NRRL B-530 strains are all derived from the original Weizmann industrial strain. An asporogenous mutant ATCC 39236 derived from this strain has been patented by scientists at the Moffett Technical Center CPC International (60).

Recently, a limited sporulating *C. acetobutylicum* E604 mutant strain has been developed by treating the parent *C. acetobutylicum* strain (ATCC 4259) with ethane methane sulfonate (61). This mutant strain has been used to ferment a high-carbohydrate substrate concentration in a multistage, continuous-temperature programmed-fermentation process followed by batch fermentation (11), to yield an extraordinarily high concentration of butanol (>20 g/L) and total solvent (>30 g/L) production (61). This mutant strain has higher butyrate uptake rate (0.33 g/L) in comparison to the parent strain (0.26 g/L) at a lower temperature of 30 °C (62). Butyrate up to an externally added concentration of 11.4 g/L did not inhibit butyrate uptake. Optimization studies for butyrate uptake by *C. acetobutylicum* suggested a direct correlation between minimum pH and butyrate concentration or temperature (63). These developments make it possible to commercially produce butanol fermentatively with no residual butyrate.

The saccharolytic strain *C. acetobutylicum* P262, one of the NCP production strains, has been studied by groups working in South Africa and New Zealand. The various solvent-producing strains now tend to be regarded as varieties of *C. acetobutylicum* or *C. beijerinckii*. For batch ABE fermentation to be industrially viable, and to compete with the chemical process, concentrations of between 22 and 28 g/L must be obtained in 40–60 h (64). A large number of publications on butanol-producing microbial cultures and the process demonstrate the extensive research work that has been done in past years (33,60,64–68).

Simultaneous recovery of butanol from the broth during fermentation can eliminate product inhibition. This will allow complete utilization of sugar and continual fermentation of the substrate at relatively higher rates. In contrast to the traditional distillation processes, removal of butanol can be accomplished by integrating a pervaporation process with the multistage fermentation process. In this process, the solvents will pass through the membranes preferentially and the organic acids and cells will remain on the liquid side and return to the fermentor. The permeate stream that is enriched in solvent relative to the feed will separate into a low-butanol aqueous phase and a butanol-rich solvent phase. This aqueous phase could be recycled back to the feed tank. Pervaporation also provides for the retention of cells in the fermentor, which improves volumetric productivity.

Ethanol. In the United States, use of ethanol as a fuel is almost wholly confined to its use as an octane enhancer in the higher grade of gasoline. Currently, this fuel ethanol is obtained from both fermentation and chemical synthesis. All of the fermentation ethanol produced in the United States is made by yeast fermentation.

Interestingly, the substrate range for both yeasts and *Zymomonas* for ethanol production is similar. Traditionally, yeasts have been used to produce alcoholic beverages and, more recently, to produce a considerable portion of industrial ethanol all over the world. Among yeasts, *Saccharomyces* is the most important and most commonly used organism. *Saccharomyces cerevisiae* converts glucose to ethanol and carbon dioxide. Other yeasts include *Pachysolen tannophilus* and *Kluyveromyces marxianus*. Many facultative and anaerobic bacteria produce varying amounts of ethanol as one of the metabolic end products. Bacterial species such as *Zymomonas mobilis*, *Sarcina ventriculi*, and *Erwinia amylovorans* utilize pyruvate decarboxylase to convert pyruvate to CO_2 and acetaldehyde, which is then reduced to ethanol. *Zymomonas* grows on glucose, fructose, and sucrose, but is unable to use starch, maltose, or pentose sugars. *Z. mobilis* is a promising bacterium with some potential to replace yeast as the major organism for production of industrial ethanol.

Z. mobilis was metabolically engineered to broaden its range of fermentable substrates to include the pentose sugar xylose (69). Two operons encoding xylose assimilation and pentose-phosphate-pathway enzymes were constructed and transformed into *Z. mobilis* in order to generate a strain that grew on xylose and efficiently fermented both glucose and xylose to ethanol. Subsequent metabolic-pathway engineering further expanded the fermentation substrate range of the ethanologenic bacterium *Z. mobilis* to include the pentose sugar L-arabinose (70). Direct fermentation of cassava starch to ethanol by *Z. mobilis* was obtained in the medium containing amylase-rich culture filtrate of *Endomycopsis fibuligera* (71). The ethanol concentration of 105 g/L could be increased to 132 g/L by further addition of glucoamylase enzyme at 0.01%.

A number of different groups of bacteria, including clostridia, produce ethanol via the reduction of acetyl CoA, which is generated by the cleavage of pyruvate produced during glycolysis. At least 30 species of *Clostridium* have been reported to produce ethanol in amounts varying from trace to close to the theoretical maximum of 2 mol ethanol/mol glucose fermented. In most cases, the extent of ethanol production is dependent upon the nature and concentration of fermentation substrate and the fermentation conditions, such as pH and temperature. For a number of mesophilic clostridial species, ethanol yields ranging from 1.7–1.9 mol/mol hexose fermented have been observed. In a single-step biomass to ethanol conversion process, *C. thermocellum* has been shown to produce 23.6 g/L ethanol from an alkali-treated paddy straw (9). The bacterium *C. ljungdahlii*, a fast-growing bacteria, has been shown to produce ethanol from carbon monoxide in concentrations up to 48 g/L from synthesis gas in a CSTR with cell recycling (10).

In recent years, thermophilic bacteria have been shown to produce solvents and are considered potential candidates for use in process development because of certain advantages over mesophiles. Thermophiles have the ability to use complex plant polymers such as cellulose or starch, and offer high growth rates, fast fermentation, reduced contamination, increased process stability, and eliminate high energy demand in cooling the media and product recovery. The major disadvantage is production of end products in low concentrations and yield due to lower end-product tolerance. However, this can be partially overcome if the solvents are removed simultaneously.

There are increasing number of patented processes concerning the use of thermophilic microorganisms, including ethanol fermentation (72–75). The thermophilic anaerobic bacteria have provided a great deal of information on ethanol-fermentation pathways via acetyl-CoA (34). There are two different ethanol pathways in thermophiles (76). The type I pathway uses NADH-linked alcohol dehydrogenase and produces ethanol from acetaldehyde. The Type II pathway uses NADPH-linked alcohol dehydrogenase to produce ethanol from both acetyl CoA directly or acetaldehyde. In Type II ethanologens, the NAD-linked alcohol dehydrogenase functions to consume ethanol. The Type I pathway is applicable in anaerobic bacteria such as *C. thermocellum*. The Type II pathway is present in *Thermoanaerobacter thermohydrosulfuricus*, *Thermoanaerobacter brockii*, and *Thermoanaerobacter ethanolicus*. Ethanol yields of heterofermentations vary considerably with the specific growth conditions used. The biochemical basis for different reduced end-product ratios of thermophilic ethanol producers that contain the same glycolytic pathways is related to subtle differences in the specific activities and regulatory properties of the enzymes that control electron flow during fermentation. Similarly, specific changes in culture conditions such as temperature and pH influence the rate and direction of the enzymatic machinery responsible for end-product formation.

T. ethanolicus 39E has a low tolerance for ethanol, with growth inhibition occurring at 2% (wt/vol) ethanol. An ethanol-tolerant strain (39EA) that was tolerant to 4% (wt/vol) ethanol at 60 °C, produced ethanol under these conditions (77), and lacked the NAD-linked alcohol dehydrogenase was selected. Another strain (H8) can grow at 8% ethanol but produces lactic acid at high solvent concentra-

tions. The mechanism of high (i.e., 8%) ethanol tolerance in *T. ethanolicus* is related to its ability to produce unique transmembrane lipids (C_{30} to C_{34} fatty acids) that provide solvent tolerance (78). Alcohol increases membrane fluidity, and these transmembrane lipids may serve to reduce the fluidity and maintain membrane integrity.

1,3-Propanediol. 1,3-Propanediol (PD) is a versatile intermediate compound for the synthesis of heterocycles and as a monomer for the production of polymers such as polyethers, polyesters, and polyurethanes. PD can also be used as a solvent and an additive for lubricants. The microbial conversion of glycerol to PD is simple in comparison to the chemical conversion of acrolein. It has been shown that some *Clostridium* species are able to convert glycerol to PD with an appreciable yield (79). 1,3-Propanediol from glycerol was produced by *C. butyricum* DSM 5431 at 50–58 g/L with productivities of 2.3–2.9 g/L per hour (13). The fermentations were conducted in 300- and 3000-L fermentors with almost similar results, indicating that the scale-up of this microbial fermentation process should not cause a major problem. The substrate glycerol is relatively cheap and its conversion to PD could help reduce glycerol surpluses in the market.

Organic Acids. Currently, organic acids such as acetic, propionic, butyric, fumaric, succinic, malic, and lactic acids can be produced by anaerobic fermentation technology. The key technical problems blocking the rapid advances in developing the bioprocess technology for organic-acid fermentations have been low product concentration, low specificity to desired product(s), low productivity or rate of fermentation, and inefficient or energy-intensive recovery processes. However, recently, significant technological advances have occurred in anaerobic fermentations for production of organic acids by fermentation of carbohydrates (80). Acidogenic fermentation of carbohydrates to volatile organic acids is well known (43). In general, anaerobic bacterial fermentations are of interest because of the wide range of products formed and substrate fermented, the high substrate-to-product conversion yields and rates, and the potential for enhanced process stability and product recovery due to physiological diversity of species (e.g., extreme thermophily, acidophily, halophily). For acetate fermentation, homacetogenic fermentation producing 3 mol of acetate from 1 mol of dextrose (e.g., by using *C. thermoaceticum*) has been described by many authors. Currently, the available organisms produce salts of organic acids in low concentrations (400–800 mM for acetate, 200–400 mM for propionate and 200–300 mM for butyrate) with low productivity of 0.2–2 g/L per hour (81). New fermentation processes with better yield and productivities have been developed, making it feasible to produce volatile as well as nonvolatile organic acids economically.

Lactic Acid. Lactic acid is a natural organic acid with a long history of use in the food industry. Lactic acid, because of its flavoring and preservation properties, is used as an acidulant, particularly in dairy products, confectionery, beverages, pickles, bread, and meat products. It is also used in the pharmaceutical and cosmetic industries. The cyclic dimers of lactic acid are used as raw materials in the synthesis of biodegradable polymers for a variety of uses, including absorbable surgical sutures, slow-release drugs, and prostheses. Ethyl lactate is the active ingredient in many anti-acne preparations. Crude grades of lactic acid are used for the deliming of hides in the leather industry, and it is used for fabric treatment in the textile and laundry industries. Its ability to form polymeric polylactic acids finds application in production of various resins. Because of its structure and its two functionally reactive groups, hydroxyl and carboxyl, lactic acid can be used to make numerous value-added functional chemicals (82). These include biodegradable polymers and copolymers, propylene glycol, propylene oxide, and acrylates. Lactic acid–based polymers and copolymers have potentially large-volume uses as biodegradable thermoplastics, pesticide formulation, and environmentally benign polymers, with a potential market exceeding several hundred million dollars per year.

The market for fermentation lactic acid is growing every year because it is now feasible to produce large-volume chemicals from lactic acid. Homofermentative lactic-acid bacteria produce none or only trace amounts of end products other than lactic acid, and are used for industrial processes. Numerous organisms are known to produce lactic acid with high (>90%) yield and productivity (>2 g/L per hour) from carbohydrate fermentation. Some of the common homolactic acid bacteria are: *Lactobacillus casei, L. pentosus, L. leichmannii, L. acidophilus, L. delbrueckii, L. bulgaricus, Streptococcus cremoris, S. lactis, S. diacetylactis, Sporolactobacillus* sp. and *Pediococcus* sp. Homolactic acid–producing bacteria ferment starch, hexoses, pentoses, and cellobiose. Depending upon the cultural conditions, the yield of lactic acid by the homolactic acid bacterium *L. casei* and heterolactic acid bacterium *T. brockii* are nearly equivalent. In non-energy-limited batch culture, *L. casei* produces lactic acid as the sole end product, whereas in energy-limited, continuous culture, nearly equivalent amounts of lactic, acetic, and formic acids, and as well as ethanol, are produced (43). Lactate is the major fermentation product of *T. brockii* grown in high-yeast extract, batch fermentation. It should be noted that lactate yields in homolactic-acid bacteria depend upon specific growth conditions and that additional products (e.g., formic acid, acetic acid, and ethanol) can also be formed. Lactic acid yields are generally highest during glycolysis via the homolactic-acid-fermentation pathway. Theoretically, 2 mol of lactate and 2 mol of ATP are formed per 1 mol of glucose fermented.

Anaerobic fermentations for production of organic acids operate optimally at pHs where salts of organic acids, rather than free acids, are produced. The free acids and their derivatives are required for the manufacture of functional chemicals. The development of new technologies to recover and purify the fermentation acids from broth have reduced the costs associated with product recovery and purification, thus making the commodity production of organic acids such as lactic acid practical and economically attractive. One of the approaches that has been traditionally applied is simultaneous fermentation and neutralization with calcium salt to obtain calcium lactate precipi-

tate. The precipitate can be separated easily, and free lactic acid can be obtained after treating the salt with acids such as sulfuric acid. The disadvantage of such a system is in handling the solids and slurry. Another emerging technology for separation and recovery of lactic acid is electrodialysis (ED). ED is a rapid process for altering the composition and/or concentration of electrolytes in aqueous solutions by transferring ions across ion-exchange membranes under the influence of a direct electric current. The recovery and purification of lactic acid using ED has been successfully demonstrated (14). In this process, sodium lactate is first selectively separated from proteins, amino acids, and cells, then, using water-splitting membrane, the lactic acid salt is converted to free lactic acid and base. The traditional process using lime for neutralization followed by cell filtration, carbon adsorption, evaporation, acidulation, gypsum filtration, carbon adsorption, evaporation, filtration, esterification, and final purification is uneconomical (83). However, separation and recovery of lactic acid by ED seems a cost-effective and attractive approach.

Succinic Acid. Succinic acid is a relatively new nonhygroscopic acidulant. Succinic acid is listed by the FDA as a GRAS additive for miscellaneous and/or general-purpose uses. It readily combines with proteins in modifying the plasticity of bread doughs. It is also an dibasic acid for producing edible synthetic fats with desirable thermal properties. Succinic acid is manufactured by the catalytic hydrogenation of maleic or fumaric acid. It has also been produced commercially by aqueous alkali or acid hydrolysis of succinonitrile, which is derived from ethylene bromide and potassium cyanide.

Succinic acid is a common intermediate in the metabolic pathway of several anaerobic microorganisms. For example, succinate is a key intermediate for anaerobic fermentations by propionate-producing bacteria, but it is only produced in low yields and in low concentrations. Succinate is also produced by anaerobic rumen bacteria. These bacteria include *Ruminobacter amylophilus* and *Prevotella ruminicola*. As such, accumulation of succinic acid in fermentation broth has been observed in a number of rumen microorganisms. Although the rumen bacteria give higher yields of succinate than do the propionate-producing bacteria, the reported fermentations were generally run in very dilute solutions and gave a variety of products in generally low yields. Moreover, the rumen organisms tend to lyse after a comparatively short fermentation time, thereby leading to unstable fermentations.

Anaerobiospirillum succiniciproducens, an anaerobic non-rumen bacterium, has been demonstrated to produce succinic acid in higher concentrations and yield (84,85). The pathway for succinic acid production in *A. succiniciproducens* has been established (45). Two processes to produce succinic acid as sodium succinate (86) and calcium succinate (87) were developed using *A. suciniciproducens*, as this organism produced succinate in a relatively higher concentration and yield with fewer cofermentation products, had a faster growth rate, and could grow in an industrial grade medium.

In order to develop a commercially attractive process to produce succinic acid by fermentation, the product should be produced in a high concentration and in a high yield (wt%) using inexpensive raw materials and nutrients. *A. succiniciproducens* produces succinate and acetate in a ratio of about 4:1, thus a good amount of carbon is lost as acetate and the separation of succinate and acetate becomes very critical. Therefore, a fluoroacetate-resistant variant FA-10 of *A. succiniciproducens* was developed that produced very minute amounts, if any, of acetic acid (88). The organism produces enough succinate to allow simultaneous fermentation and calcium succinate precipitation under strictly controlled conditions and when the correct calcium salt is used for neutralization. However, the process seems to have a major disadvantage in handling gypsum at a commercial scale. The succinic acid fermentation in *A. succiniciproducens* is regulated by fermentation pH and the amount of carbon dioxide (45). For example, lactate is the major fermentation product at pH 6.5 and higher with limiting levels of CO_2, whereas succinate is the major product at a pH of 6.0–6.2 with an abundance of CO_2. Phosphoenolpyruvate carboxykinase has been purified and described for *A. succiniciproducens* (89), a nonruminal anaerobic bacterium, as well as from a ruminal anaerobic bacterium, *Ruminococcus flavefaciens* (90). In *E. coli*, overexpression of PEP carboxikinase had no effect on succinic acid production, but succinic acid was produced as the major fermentation product by weight (3.5-fold increase in the concentration) by overexpression of PEP carboxylase (91).

Recently, a rumen-facultative anaerobic bacteria, *Actinobacillus* sp. strain 130Z, has been shown to produce succinic acid in much higher concentration (60–80 g/L) than any of the previously known succinic acid–producing microorganisms (92). Acetate and small amounts of pyruvate and formate are produced as cofermentation products. The variants of *Actinobacillus* sp. strain 130Z have been developed to produce succinic acid in a concentration of >105 g/L and yield of 85–95 wt % (19) under neutralization with magnesium. The organism is reported to be robust with a wide pH range and tolerance to high substrate and product concentrations. The levels of key succinic acid–producing pathway enzymes, such as phosphoenolpyruvate (PEP) carboxykinase, a key CO_2-fixing enzyme, malate dehydrogenase, and fumarase were significantly higher in *Actinobacillus* sp. strain 130Z than in *E. coli* K-12 (93). The key enzymes in end-product formation in *Actinobacillus* sp. 130Z were regulated by the energy substrates.

Butyric Acid. Butyric acid, butyrate esters, and other butyrate derivatives are important flavor ingredients in many natural and processed foods. For example, butyric acid is present in butter as an ester to the extent of 4–5%. Some of its esters serve as bases of artificial flavoring ingredients of certain liqueurs, soda water syrups, candies, and so on. Development of a bioprocess for production of natural butyric acid by anaerobic fermentation thus is likely to satisfy the potential market of natural fermentation products. In addition, ethyl and methyl esters of butyric acid can also be used as octane enhancers when added to gasoline (94). Some common anaerobic bacterial genera producing butyrate are *Butyribacterium*, *Butyrivibrio*,

Clostridium, Eubacterium, Fusobacterium, Haloanaerobium, and *Sarcina.* Most butyric-acid producing bacteria ferment starch, hexose, pentose, and cellobiose, and form acetic acid, in addition to butyric acid, as the major (i.e., percentage of substrate weight converted to product) fermentation product. Butyric acid production is also related to cultural conditions of the specific species. For example, *C. thermosaccharolyticum* produces butyric acid as the fermentation product during exponential growth (95). *B. methylotrophicum* can produce either acetate or butyrate as the sole end product, or it can form a mixture of butyrate and acetate. *B. methylotrophicum* ferments glucose to acetate with small amounts of butyrate at near neutral pH; it produces high levels of butyrate at acidic pH (42,96). In addition, it will produce butyrate from methanol in the presence of carbon dioxide and acetate under the same pH conditions. Products of CO fermentation by this organism are acetate and butyrate in a ratio of approximately 30:1. However, the fermentation can be shifted towards butyrate by decreasing the pH at the onset of the stationary phase (97). Theoretically, yields of 1 mol of butyrate and 2 mol of hydrogen and CO_2 are obtainable from 1 mol of hexose via the butyrate fermentation path (43). *B. methylotrophicum* or *C. butyricum* under cell-recycle can achieve this scenario.

In general, the most studied butyrate-producing microorganisms belonged to the genus *Clostridium*. A large number of fermentation substrates, including hydrolysates of waste cellulosic material, lactose from whey, molasses, and cellulosic materials can be utilized to produce butyric acid. However, the low concentration in the fermentation broth did not make the process commercially attractive as a commodity product. In addition, a major characteristic of this fermentation is the concomitant production of acetate, which is observed in butyrate-producing *Clostridium* and other bacterial species. Recently, fed-batch fermentation of glucose *with Clostridium tyrobutyricum* have yielded butyrate in a concentration of 42.5 g/L with a selectivity of 0.90, a productivity of 0.82 g/L per hour and a weight yield of 36% (98). In glucose-limited, fed-batch cultures, initially produced acetate was reutilized, resulting in exclusive production of butyrate. Because the ratio of butyrate to total acids was strongly influenced by the growth rate of bacteria, acetate being produced along with butyrate at higher growth rates (16) achieved increases in butyrate concentrations, productivity, selectivity, and yield by controlling the substrate feeding by the rate of gas production. In a fermentation of wheat flour hydrolysate (380 g/L of glucose) with *C. tyrobutyricum*, a butyrate concentration of 62.8 g/L was obtained with a productivity of 1.25 g/L per hour, a selectivity of 91.5%, and a weight yield of 45% (16). Thus, production of butyrate for specialty uses by fermentation is feasible.

Propionic Acid. Propionic acid is a value-added specialty chemical used in various chemical and food-processing industries. It has wide-ranging applications, such as an antifungal agent in foods and feeds, and as an ingredient in thermoplastics, antiarthritic drugs, perfumes, flavors, and solvents. Currently, propionic acid is produced by chemical synthesis from petroleum feedstocks, but small amounts of natural propionic acid are produced by fermentation. Propionic acid is a major end product of fermentations carried out by a variety of anaerobic bacteria, many of which ferment glucose to propionate, acetate, and CO_2. Lactate can also be fermented to propionate either via the acrylate pathway in a stepwise reduction to propionate, or via the succinate pathway, where lactate is converted to propionate via pyruvate and succinate.

Although several microorganisms can produce propionic acid, fermentations using propionibacteria have been studied most extensively. The slow growth rate of propionibacteria usually results in propionate production at a slow rate. *Propionispira arboris* ferments glucose to propionate, acetate, and CO_2 via the succinate pathway. The ratio of propionate to acetate is low, resulting in a carbon loss to cofermentation products. However, the ratio of propionate to acetate can be changed from 2 to 16:1, approaching a homofermentation yield from glucose, by use of H_2 as a cosubstrate because this species consumes H_2 (44). Various fermentation systems have been examined to increase the propionate fermentation rate, including batch, fed-batch, continuous, continuous with cell recycle, extractive, and immobilized cell fermentations. Also, various substrates such as glucose, xylose, lactate, glycerol, food-processing waste, and whey have been examined to produce propionate not only with increased productivity, but also economically. Using *Propionibacterium acidipropionici*, propionate has been produced at 42 g/L from glycerol (18) and at 57 g/L from glucose (17) in batch and fed-batch fermentation systems, respectively. It seems fermentative production of propionate is feasible if an appropriate fermentation system and recovery technology can be integrated to obtain the acid in high concentration, yield, and productivity.

Enzymes

In recent years, enzymes have gained wider applications in the biotechnology industry. The global industrial market has recently been valued at $1.4 billion, with an expected market increase of 4–5% (99). Enzymes are grouped into six major classes: oxidoreductases, transferases, hydrolases, lyases, ligases, and isomerases. The majority of commercialized enzymes are hydrolases such as amylase, cellulase, xylanases, pectinase, proteases, lipase, and collagenase. Other important enzymes are glucose isomerase and glucose oxidase. Anaerobic bacteria are very diverse and thus produce a wide variety of enzymes, excluding oxygenases. Enzymes are widely used in pulp and paper, textile, detergent, and food- and feed-additive industries. In addition, enzymes also find applications in pharmaceutical synthesis, therapeutic contexts, and clinical and chemical analysis. The rumen anaerobic microbial population presents a rich and, until recently, underutilized source of novel enzymes with tremendous potential for industrial applications. The enzymes from these microorganisms include cellulases, xylanases, β-glucanases, pectinases, amylases, proteases, phytases, and tannases. High-molecular-mass complexes containing numerous cellulases have been identified in a number of rumen bacteria, including *Butyrivibrio fibrisolvens, Ruminococcus albus,* and *Fibrobacter succinogenes.*

The most common enzyme-producing anaerobic bacteria are mesophilic and thermophilic clostridia, and moderate- and hyper-thermophilic nonclostridial species (100,101). Among the amylolytic enzymes, α-amylase hydrolyzes internal α-1,4 linkages of starch at random in an endo fashion, producing oligosaccharides of varying chain lengths. Generally, it cannot act on α-1,6 linkages of starch. *C. butyricum, C. acetobutylicum, T. ethanolicus, C. thermoamylolyticum* have been reported to produce this amylase enzyme. β-Amylase hydrolyzes alternate α-1,4-glycosidic linkages of starch in an exo fashion from the nonreducing end, producing β-maltose. The β-amylase has been produced in high yield as a primary product during growth of *Thermoanaerobacterium thermosulfurigenes* (102). A hyperproductive mutant was isolated that produced eightfold more β-amylase than the wild type. Synthesis of the enzyme was both constitutive and resistant to catabolite repression. The β-amylase has also displayed industrial potential for the production of high-maltose syrups from raw or soluble starch at 75 °C. Various maltose-containing syrups are used in the brewing, baking, canning, and confectionery industries.

Glucoamylase is an exoacting carbohydrase that cleaves glucose units consecutively from the nonreducing end of starch molecules. High levels of a thermostable glucoamylase activity has been reported in crude extracts of *T. ethanolicus* 39E (103), although it was purified and later described as an α-glucosidase activity. Glucoamylase is widely used in alcoholic fermentation of starchy materials and in the commercial production of glucose and high-glucose corn syrups. However, the current source of this enzyme is fungal. Pullulanase is a debranching enzyme that specifically cleaves α-1,6 linkages in starch, amylopectin, pullulan, and related oligosaccharides. It is generally used in combination with saccharifying amylases such as glucoamylase, fungal α-amylase, or fungal β-amylase for the production of various sugar syrups because it improves saccharification and yield (104). Hyun and Zeikus (103) found that *T. ethanolicus* produces highly thermoactive and thermostable cell-bound pullulanase. They also developed a hyperproductive mutant that displayed improved starch metabolism features. Pullulanase has also been reported to be produced by *C. thermosaccharolyticum*. Taking advantage of the high thermoactivity and acidoactivity of these pullulanases, it may be assumed that these enzymes might effectively replace α-amylase and pullulanase in both starch liquefaction and saccharification processes. α-Glucosidase hydrolyzes terminal nonreducing α-1,4-linked glucose residues of various substrates, releasing α-D-glucose. It is generally considered to be maltase, but has a wide specificity, being able to cleave glucosides of nonsugars in addition to maltose, maltotriose, and other maltooligosaccharides, and to transfer α-D-glycosyl residues of maltose and α-D-glucosides to suitable acceptors.

α-Amylase, which is used in starch liquefaction, solubilizes α-1-4 linkages in starch-forming maltodextrin syrups. α-Amylase has been recently described in thermoanaerobes. For example, the *Pyrococcus furiosus* α-amylase gene has been cloned and expressed in *E. coli* (105). The *P. furiosus* α-amylase is twice as active as *Bacillus litchenformis* commercial α-amylase, but does not require Ca^{2+} for stability at high temperature (i.e., >100 °C).

The cellulase system in both bacteria and fungi comprises three different classes of enzymes: (*1*) endo-1,4-β-glucanases; (*2*) exo-1,4-β-D-glucanases, including both 1,4-β-D-glucan cellobiohydrolases and 1,4-β-D-glucan glucohydrolases; and (*3*) 1,4-β-D-glucosidases, also referred as cellobiases. Several species of cellulolytic clostridia have been described in the literature. These include *Clostridium cellobioparum, C. acetobutylicum, Clostridium cellulovorans, Clostridium stercorarium*, and *Clostridium thermocellum*. Most of the work on industrial cellulases has been accomplished using aerobic fungal systems. However, due to the high specific-enzyme activity on the one hand and the high thermostability on the other, the cellulases from *C. thermocellum* have been considered for potential industrial utilization in direct alcohol fermentations, but not for saccharification *per se* (i.e., glucose production). Recently, endoglucanases have been used in laundry detergents.

Collagenases are endopeptidases that hydrolyze native, insoluble fibrous collagen. One of the anaerobic bacteria producing collagenase that has been extensively studied is *C. histolyticum*. The collagenase of *C. histolyticum* is available commercially. Other collagenase-producing anaerobic bacterial species include *C. collagenovorans* and *C. proteolyticum* (106). Pure collagenase can be applied as a sensitive probe for biosynthetic studies and sequence determinations. It is useful in prevention or cure of keloids. Collagenase is useful for the dispersal and dissociation of animal tissues in the laboratory. It is routinely used to separate cells from their parent tissues. *C. histolyticum* also produces an extracellular sulfhydryl proteolytic enzyme called clostripain (clostridiopeptidase). This enzyme possesses amidase, esterase, and proteolytic activity, which is directed toward the carboxyl peptide linkage of arginine (107).

Pectin-degrading enzymes are produced by a variety of microorganisms, including clostridia. These have been produced by *C. aurantibutyricum, C. felsineum, C. multifermentas, C. roseum*, and *T. thermosulfurigenes*. An active thermostable polygalactunonate hydrolase and pectin methylesterase have been produced by *T. thermosulfurigenes* (108). These thermostable pectinolytic activities may have application in fruit juice clarification and for processing food or agricultural/forestry products.

Several xylose (glucose) isomerases have been isolated from thermoanaerobes including from *Thermotoga neapolitana* (109). This enzyme is more stable and active than commercial enzymes. An alkaline phosphatase that is more stable and active than commercial calf enzyme has been purified and characterized from *T. neapolitana* (110). Table 3 provides examples of some selected enzymes of industrial importance from anaerobic extremophiles. This serves to illustrate the diversity of enzymes from anaerobes.

Specialty Chemicals and Biotransformations

Diverse species of acetogenic and methanogenic bacteria grow on CO, H_2—CO_2, or methanol. These bacteria have been studied for conversion of synthesis gas (i.e., CO + H_2) or CO to methane by pure cultures and consortia, or into organic acids and alcohols by pure culture. Some of these

Table 3. Enzymes of Potential Industrial Importance from Anaerobic Extremophiles

Enzyme	Optimal temperature (°C)	Localization	Organism	Reference
Endoglucanase	65	extracellular	*Clostridium thermocellum*	Ng et al. (111)
α-Amylase	100	extracellular	*Pyrococcus furiosus*	Dong et al. (105)
β-Amylase	75	extracellular	*Thermoanaerobacterium thermosulfurigenes*	Hyun and Zeikus (102)
Xylose/glucose isomerase	95	intracellular	*Thermotoga neopolitana*	Vielle et al. (109)
Amylopullulunase	85	extracellular	*Thermoanaerobacter ethanolicus* 39E	Mathupala and Zeikus (112)
	105	extracellular	*Pyrococcus furiosus*	Dong et al. (113)
Endo-xylanase	70	extracellular	*Thermoanaerobacterium saccharolyticum* B6A-RI	Lee et al. (114)
1° Alcohol dehydrogenase	80	intracellular	*Thermococcus litoralis*	Ma et al. (115)
2° Alcohol dehydrogenase	90	intracellular	*Thermoanaerobacter ehtanolicus* 39E	Burdette et al. (116)
β-Xylosidase	70	intracellular	*Thermoanaerobacterium saccharolyticum* B6A-RI	Lee and Zeikus (117)
Alkaline phosphatase	85	intracellular	*Thermotoga neopolitana*	Dong and Zeikus (110)

bacteria include *B. methylotrophicum*, *Eubacterium limosum*, *Peptostreptococcus productus*, *C. thermoaceticum*, *C. ljungdahlii*, and *Methanosarcina barkeri*. Synthesis gas represents a cheap feedstock for microbial conversion to higher-value commodity products. Recent work has focused at syngas fermentation to liquid-fuel additives. The feasibility of ethanol production from syngas fermentation by *C. ljungdahlii* (118) and ethanol plus butanol production by *B. methylotrophicum* (119,120) has been established. However, these processes need to be developed further for their commercialization.

Recently, anaerobes have also been examined for a wide range of specialty fermentation products including antimicrobials, bioflavors, biopigments, biopesticides, and anticancer agents. Antimicrobial compounds such as antibiotics and bacteriocins are produced by some of the anaerobic species. The most common examples are that of nisin and pediocin production by *Lactococcus lactis* and *Pediococcus acidilactici*, respectively (121). Some strains of lactobacilli catalyze the decarboxylation of glutamate to γ-aminobutyrate (122).

Many chiral compounds can be synthesized by microbial hydrogenation using hydrogen (or formate) and hydrogenase-containing microorganisms. Oxidoreductase enzymes are involved in electron-transfer reactions and can be applied in a stereoselective catalysis. The best known alcohol dehydrogenase is that present in both *T. ethanolicus* and other thermoanaerobes, the most notable being *T. brockii*. This NADP-linked secondary alcohol (aldehyde/ketone) dehydrogenase, found in thermoanaerobes, exhibits a wide substrate specificity toward linear and cyclic secondary alcohols and thioesters (116,123,124). Table 4 summarizes novel oxidoreductases present in some clostridial species that are of commercial interest (125). The biotransformation reactions can be carried out either by whole-cell fermentation (which eliminates coenzyme regeneration but may be subject to interference by competing enzymes), or by crude or purified enzyme in batch or continuous-flow systems using an immobilized enzyme column. *C. sporogenes* performs the stickland reaction, in which pairs of amino acids are fermented, one amino acid acting as an electron donor (e.g., valine, leucine, isoleucine), and the other acting as an electron acceptor (e.g., proline, glycine). Synthesis of pyruvate and other 2-oxacids from acyl phosphate derivatives have been shown with permeabilized cells of *C. sporogenes* (125). In amino acid–fermenting anaerobic bacteria, a set of unusual dehydratases is found that use 2-hydroxyacetyl-CoA, 4-hydroxybutyryl-CoA, or 5-hydroxyvaleryl-CoA as substrates. These anaerobic bacteria include *C. propionicum*, *C. aminobutyricum*, and *C. aminovalericum* (126).

Waste Treatment

Residual Organics. Microorganisms excel in using organic substances as sources of nutrients and energy. The challenge of wastewater treatment is to remove (*1*) compounds with a high biochemical oxygen demand, (*2*) pathogenic organisms and viruses, and (*3*) a multitude of human-made chemicals. Anaerobic digestion (AD) is commonly used to treat materials with a high content of insoluble organic matter, such as cellulose, and to degrade concentrated industrial wastes, such as those from the food-processing industry. AD is a complex biological process that utilizes a consortium of anaerobic microorganisms to act in concert to hydrolyze complex organics to simple monomers and then to volatile fatty acids (VFAs). These VFAs are ultimately converted to methane and carbon dioxide in the final step in the anaerobic food chain. The degradative and fermentative reactions in the anaerobic treatment processes can be divided into two stages: acid-forming and methane-forming, which involve at least three groups of anaerobic bacteria—acidogenic, syntrophic, and methanogenic. In the acid-forming stage, complex organic polymers, including carbohydrates, fats, and proteins, are hydrolyzed and fermented to VFAs, alcohols, and ketones by selective anaerobic bacteria. Organic acids and alcohols such as lactic acid, propionic acid, butyric acid, and ethanol are converted to acetic acid and carbon dioxide by acetogenic bacteria in the syntrophic association of, generally, methanogenic bacteria. The acetate and CO_2 are finally converted to CH_4 by methanogenic bacteria that

Table 4. Novel Oxidoreductase in *Clostridium* Species

Reduction	Electron donor	Organism
Aldehyde/ketone dehydrogenase		
Steroids	Unknown	*C. paraputrificum*
		C. bifermentans
Methyl ketones	NADPH	*C. thermohydrosulfuricum*
Ketones	Unknown	*C. pasteurianum*
	Unknown	*C. tyrobutyricum*
2-Oxoacid synthase		
Fatty acids	Ferredoxin	*C. sporogenes*
Acetate	Ferredoxin	*C. kluyveri*
Linoleic reductase		
Linoleic acid	Unknown	*C. sporogenes*
Enoate reductase		
Cinnamic acid	NADH	*C. sporogenes*
Crotonic acid	NADH	*C. tyrobutyricum*
2-Oxoacid reductase		
Phenylpyruvic acid	NADH	*C. sporogenes*
Nitroaryl reductase		
Chloramphenicol	Ferredoxin/flavodoxin	*C. acetobutylicum*
Metronidazole	Ferredoxin/flavodoxin	
Paro-nitrobenzoate	Ferredoxin/flavodoxin	
2-Nitrobenzene	Ferredoxin/flavodoxin	
Lipoamide hydrogenase		
NAD/lipoamide	Lipoamide/NADH	*C. kluyveri*

Source: From Lovitt et al. (125).

may include acetate- and H_2-CO_2-consuming methanogens. Anaerobic microorganisms in general show a high degree of metabolic specialization. The success of the anaerobic digestion process therefore depends upon cooperative interactions between microorganisms with different metabolic capabilities. Figure 2 illustrates the anaerobic food chain that converts complex organic matter into methane. Since the rate of VFA production can be significantly higher than that of VFA conversion to methane, an imbalance between these two rates can occur in a single-stage digestor, resulting in VFA accumulation, a concomitant pH drop, cessation of methane fermentation, and ultimate process failure. A two-stage AD process that separates the acid-forming and methane-forming stages has been practiced with a variety of wastes. Alternatively, a high-rate syntrophic methanogenic microbial consortium can be developed (127,128). The leach-bed two-phase digestion process, consisting of an acid-phase, solid-bed reactor operated in tandem with a separate, packed-bed, methane-phase digestor (anaerobic filter), for high solids waste has been described (129). Low-strength wastewaters can now be treated, even under psychrophilic conditions, by using specific rheologic conditions in the expanded granular sludge bed (EGSB) reactor (130). Solid wastes can be treated anaerobically using the thermophilic "high-solids" fermentation technology (131). New reactor designs permit S° recovery from SO_2-rich waste gases (132).

Many human-made organic compounds are degraded during sewage treatment. Since 1986, full-scale UASB reactors have been used to treat municipal wastewater throughout the world. Treatment costs are halved when anaerobic treatment (e.g., UASB reactor) is applied instead of just aerobic processes (133). In warm climates, a simple upflow sludge blanket (USB) septic tank with an aerobic posttreatment (e.g., a trickling filter) can be combined to remove the bulk of its suspended solids (134). AD is expected to reduce most of the pathogenic bacteria. Thus, the treated water may be considered reclaimed for use in crop production in the next decade. Organic slurries containing particulate organic matter, such as animal manures and primary or secondary sewage sludges, are normally digested in completely mixed reactors. Because the hydrolysis rates of certain solids is slow, separation of hydraulic retention time (HRT) and solids retention time (SRT) may improve performance of digestors treating slurries. AD does not remove NH_4^+. In a new NH_4^+-removal pathway observed in the methanogenic reactors, NH_4^+ was oxidized anaerobically to N_2 in the presence of NO_2^-, with a lab-scale reactor achieving a removal rate of 0.7 kg NH_4^+-N/m^3 per day (135).

The UASB technology is now widely used to anaerobically digest industrial wastewaters with concentrations in the range 2–20 g COD/L. UASB reactors are usually implemented when wastewater is rich in carbohydrates and relatively poor in other contaminants. Biomass retention through adequate granulation is of utmost importance in UASB technology, first to obtain a good effluent quality and, second, in order to ensure a minimal cell residence time of 7–12 day that is required to avoid the wash-out of the slowest-growing anaerobic bacteria. The onset of sludge granulation after start-up of an anaerobic-sequencing batch reactor was enhanced by adding a cationic polymer, divalent cations, and granulation nuclei such as clay minerals (136). Consideration of surface thermodynamics is important in granulation (137). Some key methanogenic species have been identified and implicated in the granu-

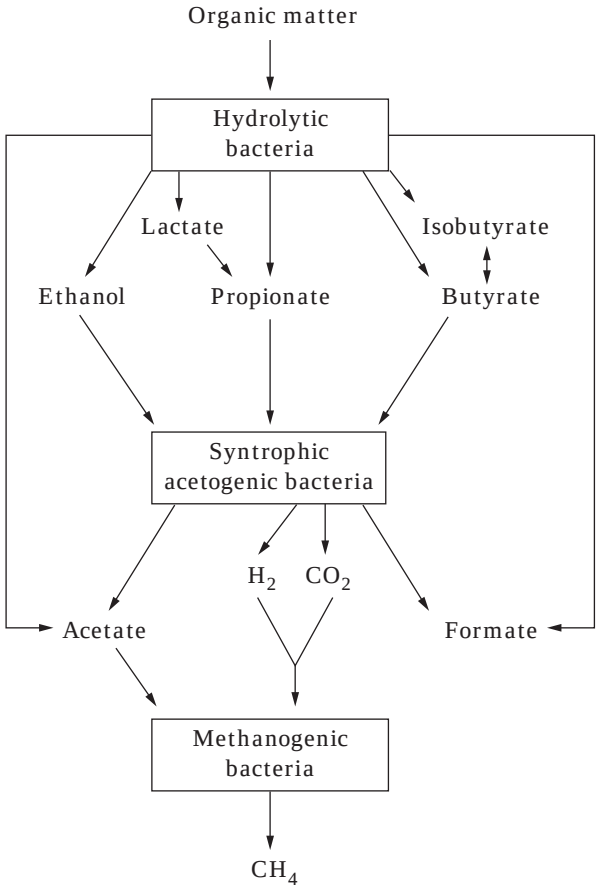


Figure 2. Anaerobic degradation of organic matter to methane in relation to microbial trophic groups.

lation process (138), and exopolysaccharides produced by these methanogens, specifically *Methanobacterium formicicum*, *M. mazei*, and *Methanosaeta* sp. seem to be responsible for the stability of granular structure (139). Industrial wastes are often unreliable in terms of their composition. Inhibition of methanogens caused by toxic compounds present in industrial wastewaters can be alleviated by adding activated carbon directly in the anaerobic reactors (140). Addition of GAC to UASB reactors treating textile wastewater prevented the chronic intoxication of anaerobic sludge that takes place in the absence of the GAC (134).

Anaerobic treatment systems are gaining in popularity and finding new applications. As revealed in papers presented at the 8th International Conference on Anaerobic Digestion, Sendai, Japan, in May 1997, the databank of anaerobic reactors presently includes 1066 operating anaerobic reactors worldwide. Of these reactors, 956 (89.7%) are for the treatment of industrial effluents, while 78 (7.3%) treat domestic sewage, and only 32 (3.0%) of the total are anaerobic-treatment systems for organic solid waste (excluding the vast number of biogas plants for manure treatment that have been installed worldwide). It is estimated that there are about two million biogas plants for manure treatment just in India alone. The largest number of municipal anaerobic treatment plants presently exist on a moderate, but remarkably growing, scale for sewage treatment in countries such as Brazil, Colombia, Mexico, India, and China. In industrial countries, only sewage sludge is treated anaerobically on a large scale due to climatic factors. For industrial wastewater treatment, countries such as India, China, Thailand, Brazil, and Mexico have a growing interest in anaerobic treatment for the local industry; in industrialized regions the leading countries, especially The Netherlands and Belgium, have almost already saturated their home markets. For anaerobic solid-waste treatment (and cofermentation), Germany and Denmark have significant numbers of plants so far. In the near future, anaerobic treatment of solid wastes will become a viable option provided that appropriate collection systems are introduced.

Toxic Organics. In recent years, AD has been used to degrade xenobiotics present in wastewater, soils, and river sediments. Complex xenobiotics often require more than one species to be completely mineralized. Microorganisms need to adapt, maybe for several months, before the maximal conversion rate of a new xenobiotic compound can be achieved. Anaerobic tunnel reactors are used to treat soils and sediments polluted with chloroethene, BTEX, or TNT (141). Efficiency of removal of adsorbable organic halogens by a UASB treating kraft-mill bleach wastewater varied between 27–65%, depending on the residence time (142). *Desulfomonile tiedjei*, which can rapidly transform 3-chlorobenzoate (3-CB), has been shown to become established in an UASB reactor. This indicates that a specific strain or desired species can be incorporated in microbial granular culture of an UASB reactor. Incorporation of an adapted microbial sludge culture into syntrophic biomethanation granules has been demonstrated to treat PCP (14). These studies were further validated by developing PCE/TCE-dechlorinating methanogenic granules (143) using a similar protocol. In this, the anaerobic granules, using methanol as an electron donor for reductive dehalogenation, completely converts highly oxidized PCE to ethylene. In contrast, PCBs biodegrade very slowly and persist in the environment for decades. An effective microbial enrichment that was shown to remove *ortho*-, *meta*-, and *para*-chlorine was incorporated into the PCP-degrading methanogenic microbial granules to produce PCB-dechlorinating anaerobic microbial granules in a manner similar to achieving PCP-granules in the first place (31). The PCB-dechlorinating anaerobic methanogenic granules have shown PCB-dechlorination of defined PCB-congeners (29) and Aroclor-contaminated river sediments (30). These granules are stable, culturable, and can be produced on a mass scale for bioaugmentation for application in both in situ and ex situ processes. These granules have also demonstrated degradation and mineralization of biphenyl, a non-chlorinated end-product of PCB-dechlorination process reductively (Natarajan, Wu, and Jain, unpublished data). Since PCBs are a complex mixture of various PCB compounds, and each PCB compound varies in its degree of dechlorination, it is expected that a mixture of bacteria capable of removing chlorines from different positions on the PCB molecule must be involved in achieving a com-

plete dechlorination of PCB. A similar assumption, however, is no longer true for dechlorination of perchloroethylene (PCE). Recently, Zinder and his group isolated a pure culture of an anaerobic bacteria, tentatively named *Dehalococcoides* strain 195, that dechlorinates PCE and TCE in a stepwise manner to ethene (144). This novel eubacterium is not methanogenic, acetogenic, or sulfidogenic.

2,4,6-Trinitrotoluene (TNT), one of the most widely used explosives, occurs as a pollutant of soil and ground water, especially at sites of ammunition factories. There have been several reports of degradation or transformation of nitro and aminoaromatic compounds under anaerobic conditions. A nonspecific reduction of the nitro group to the corresponding amine was assigned to a variety of methanogenic bacteria, sulfate-reducing bacteria, and clostridia (145). In contrast, it was reported that 2,4-dinitrophenol, 2,4- and 2,6-dinitrotoluene (146), and TNT (147) were used as sources of nitrogen by sulfate-reducing bacteria. Thus, it is clear that many nitroaromatic compounds can serve as growth substrates for anaerobic bacteria. Anaerobic treatment systems have been proposed as a means of avoiding the accumulation of partially reduced intermediates during degradation of TNT. Under strictly anaerobic conditions, TNT can be completely reduced to triaminotoluene (147). *C. acetobutylicum* transformed 2,4,6-TNT to undetermined end products via monohydroxylamino derivatives (148). In contrast to solventogenic cells, acidogenic cultures showed rapid transformation rates and the ability to transform TNT and its primary reduction products to below detection limits. Anaerobic treatment of explosive (primarily TNT)-contaminated soil was demonstrated in open bulk containers containing soil, phosphate buffer, and potato starch (149). The potato starch served as a rapidly degradable carbon source that allowed the rapid establishment of anaerobic conditions. Cresols and small organic acids were observed as end products. Similarly, Dinoseb, a very persistent herbicide that is highly toxic to virtually all living systems, has also been shown to be degraded by anaerobic microbial consortia using potato starch as a carbon source (150).

Inorganics. Microorganisms degrade certain toxic constituents used in mineral processing, and concentrate and immobilize soluble heavy metals released as a result of mining and mineral-processing activities. The initial mechanism of metal binding by microorganisms is electrostatic attraction between charged metal ions in solution and charged functional groups on microbial cell walls. The cell walls are composed of macromolecules with functional groups (principally carboxylate, amine, imidazole, phosphate, sulfhydryl, and sulfate) that contribute a net negative charge to the surface of the microorganism. These functional groups remain active even when the microorganism is not viable. Sulfate-reducing bacteria (SRBs) are used in highly controlled reactor systems and in constructed anaerobic wetlands for removal of sulfate and heavy metals from acid-rock drainage and other aqueous, metal-contaminated, streams. SRBs such as *Desulfovibrio* and *Desulfatomaculum* oxidize organic matter or H_2 by using sulfate as an electron acceptor to produce hydrogen sulfide and bicarbonates. The sulfide immediately reacts with soluble heavy metal ions to form highly insoluble metal sulfides. A system to evaluate anaerobic and aerobic treatment of mine effluents containing excessive sulfate and heavy metals has been described (151). In the continuous system, the H_2S is produced by the SRB biofilm on the dolomite pebbles in the anaerobic stage, which precipitates metal sulfides from a waste stream amended with an organic energy source. In the aerobic stage, a completely mixed reactor and a settling tank facilitate oxidation of residual H_2S and biodegradation of residual organics from the primary anaerobic column. Sulfate reduction has been considered as a method for permanent stabilization of sulfidic mine tailings.

High sulfate concentrations are common in wastewaters from paperboard industries, molasses-based fermentation industries, edible oil refineries, and in acidic leachates of pyritic waste rock and tailings (152). Sulfidogenic UASB reactors have attracted some attention for treating such polluted streams, but the biological processes to remove high SO_4^{2-} contents have not yet been optimized. A new bioprocess in which sulfate and heavy metals are removed simultaneously from groundwater has been developed (153). In this two-step process, SO_4^{2-} bacteria produce sulfides that precipitate with the heavy metals after ethanol, an electron donor, is added to the first reactor (BIOPAQ UASB). In the second reactor (THIOPAQ-submerged fixed-film reactor), sulfur bacteria reoxidize the excess sulfide selectively to solid S° under controlled dosage of O_2. With HRT of 4 h in the BIOPAQ reactor and 30 min in THIOPAQ reactor, removal efficiencies of 99% for heavy metals and of 85% for sulfate have been achieved. This technology has also been adapted for desulfurization of flue gas from power plants (132). In the modified process for flue-gas desulfurization, H_2 can be used as an electron source in lieu of ethanol.

Minewaters and industrial effluents containing high sulfate concentrations create a disposal problem that requires an urgent solution in order to avoid excess mineralization of surface waters. Sulfate can be converted quantitatively to hydrogen sulfide by *D. desulfuricans* (154). Further conversion to elemental sulfur can be effected by the photosynthetic bacteria *Chlorobium limicola* and *Chromatium cinosum*. Two separate reactors were used for hydrogen sulfide and sulfur production. A possible way to increase the sulfate reduction rate is by making use of a packed-bed, instead of a completely mixed, reactor.

CONCLUSIONS AND FUTURE RESEARCH

The use of anaerobes in industrial processes has grown dramatically over the past 20 years. Traditionally, anaerobes were used routinely in the processing of fermented foods such as lactic acid and propionic acid bacteria in cheeses, yogurts, sausages, and pickles. In waste treatment, mixed methanogenic cultures were used in both industrial and municipal anaerobic digestors for removal of residual organic matter.

Now, *C. botulinum* toxins are used as medicines. Numerous organic alcohols and acids are produced by industry as natural fermentation chemicals. These natural

chemicals are used in flavors, fragrances, preservatives, and other specialty products. Fermentation-derived lactic acid is now produced by industry as a commodity chemical and is used as an acidulant, disinfectant, green solvent, and precursor for polylactide-based biodegradable plastics. Anaerobic composting systems are used to treat food and municipal solid wastes because they produce less end product that must be used or land spread. Sulfite is also removed from coal-bearing flue gases in bioreactor systems comprising mixed sulfidogenic cultures.

With the rapid growth in isolation and characterization of new anaerobic species from normal and extreme environments, the diversity of anaerobic species with unique biochemical attributes far surpasses that of aerobic microbes. This great diversity of anaerobic microbes should be expected since anaerobes were the first organisms to evolve on earth, yet they were the last large natural group to be studied in detail by biologists. The vast array of biochemical diversity shown by anaerobes is related in part to the fact that they are not limited to use of O_2 as an electron acceptor, they can use fermentative metabolism or CO_2, SO_4, or other electron acceptors for anaerobic respiration. This biochemical diversity will undoubtably be exploited in the future since many of the fermentation products and enzymes of anaerobes are of interest to industry (i.e., food-feed, chemical, pharmaceutical, energy, and environmental companies).

Several research areas on anaerobes showing special industrial promise include succinate fermentation, thermozymes, and dechlorinating methanogenic granules. Ethanol fermentations are limited in part because two moles of CO_2 are lost from the product per mole of glucose fermented. On the other hand, one can derive more than one pound of succinate per pound of glucose fermented because the theoretical chemical yield is 1 glucose + 2 CO_2 + 2 $H_2 \rightarrow$ 2 succinate. Succinate has a wide variety of uses as both a specialty chemical and commodity intermediate chemical. Perhaps the largest markets for succinate includes their use as a feedstock to produce stronger-than-steel engineered plastics and polyesters, and as the chelator EDDS (ethylene diamine disuccinate) to replace nonbiodegradable EDTA.

The saccharolytic enzymes of thermoanaerobes are very active and stable. These thermozymes could be used to develop the next generation of enzymes used in the starch-processing industry (i.e., α-amylase, glucose isomerase, glucoamylase, and pullulanase) or to initiate a cellulose-processing industry for enhanced biomass utilization based on very active and stable cellulases and hemicellulases. Anaerobic sediments and soils are contaminated with a wide variety of chlorinated compounds. The use of dechlorinating methanogenic granules offers an alternative to expensive dredging and landfilling. Dechlorinating methanogenic granules can degrade a wide variety of toxicants (DDT, DIOXINS, PCE/TCE, PCBs, PCP, etc.) by in situ bioagumentation technology.

BIBLIOGRAPHY

1. N.R. Krieg and J.G. Holt eds., *Bergey's Manual*, Williams and Wilkins, Baltimore, Md., 1984.
2. M.W.W. Adams, *Annu. Rev. Microbiol.* **47**, 627–658 (1993).
3. J.G. Zeikus, in K.H. Park, J.F. Robyt, and Y.-D. Choi, eds., *Enzymes for Carbohydrate Engineering*, Elsevier Science, Amsterdam, 1996, pp. 145–161.
4. K.O. Stetter, A. Hoffman, and R. Huber, in R. Guerrero and C. Pedros-Alio eds., *Trends in Microbial Ecology*, Spanish Society for Microbiology, Barcelona, Spain, 1993, pp. 25–28.
5. L.D. Mermelstein and J.G. Zeikus. *Extremophiles: Microbial Life in Extreme Environments*, Wiley-Liss, New York, 1998, pp. 255–286.
6. C.R. Woese, *PNAS* **91**, 1601–1603 (1994).
7. M. Watanabe, *Scientist* **8**, 14–15 (1994).
8. S.S. Kesava, T. Panda, and S.K. Rakshit, *Proc. Biochem.* **31**, 449–456 (1996).
9. K.S. Rani, M.V. Swamy, and G. Seenayya, *Proc. Biochem.* **33**, 435–440 (1998).
10. K.T.A. Klasson, M.D. Ackerson, E.C. Clausen, and J.L. Gaddy, *Fuel* **72**, 1673 (1993).
11. U.S. Patent 5,063,156 (November 5, 1991), D.A. Glassner, M.K. Jain, and R. Datta (to Michigan Biotechnology Institute, Lansing, Mich.).
12. A.J. Grethlein, R.M. Worden, M.K. Jain, and R. Datta, *J. Fermen. Bioeng.* **72**, 58–60 (1991).
13. B. Günzel, S. Yonsel, and W.-D. Deckwer, *Appl. Microbiol. Biotechnol.* **36**, 288–294 (1991).
14. U.S. Patent 4,885,247 (December 5, 1989), R. Datta (to Michigan Biotechnology Institute, Lansing, Mich.).
15. S.R. Parekh and M. Cheryan, *Biotech. Lett.* **16**, 139–142 (1994).
16. F. Fayolle, R. Marchal, and D. Ballerini, *J. Ind. Microbiol.* **6**, 179–183 (1990).
17. H.-D. Paik and B.A. Glatz, *Appl. Microbiol. Biotechnol.* **42**, 22–47 (1994).
18. F. Barbirato, D. Chedaille, and A. Bories, *Appl. Microbiol. Biotechnol.* **47**, 441–446 (1997).
19. U.S. Patent 5,573,931 (November 12, 1996), M.V. Guettler, M.K. Jain, and D. Rumler (to Michigan Biotechnology Institute, Lansing, Mich.).
20. W.R. Kenealy, Y. Cao, and P.J. Weimer, *Appl. Microbiol. Biotechnol.* **44**, 507–513 (1995).
21. L. Bhatnagar, M.K. Jain, and J.G. Zeikus, in J.M. Shively and L.L. Barton eds., *Variations in Autotrophic Life*, Academic, London, 1991, pp. 251–270.
22. U. Deppenmeier, V. Müller, and G. Gottschalk, *Arch. Microbiol.* **165**, 149–163 (1996).
23. M. Blaut, *Antonie van Leewenhoek* **66**, 187–208 (1994).
24. E.J. Bouwer, B.E. Rittman, and P.L. McCarty, *Environ. Sci. Technol.* **15**, 596–599 (1981).
25. E.J. Bouwer and P.L. McCarty, *Appl. Environ. Microbiol.* **45**, 1286–1294 (1983).
26. E.J. Bouwer and P.L. McCarty, *Appl. Environ. Microbiol.* **45**, 1295–1299 (1983).
27. L. Bhatnagar, W.-M. Wu, M.K. Jain, and J.G. Zeikus, in *Proceedings of the International Symposium on Environmental Biotechnology*, Royal Flemish Society of Engineers, Ostend, Belgium, 1991, pp. 1–10.
28. W.-M. Wu, L. Bhatnagar, and J.G. Zeikus, *Applied Environ. Microbiol.* **59**, 389–397 (1993).
29. M.R. Natarajan, W.-M. Wu, J. Nye, H.-Y. Wang, L. Bhatnagar, and M.K. Jain, *Appl. Microbiol. Biotechnol.* **46**, 673–677 (1996).

30. M.R. Natarajan, J. Nye, W.-M. Wu, H.-Y. Wang, and M.K. Jain, *Biotech. Bioeng.* **55**, 182–190 (1997).
31. U.S. Patent 5,635,393 (June 3, 1997), L. Bhatnagar, W.-M. Wu, M.R. Natarajan, H.Y. Wang, and M.K. Jain (to Michigan Biotechnology Institute, Lansing, Mich.).
32. J.F. Quensen, III, S.A. Mueller, M.K. Jain, and J.M. Tiedje, *Science* **280**, 435–440 (1998).
33. R.W. Lovitt, B.-H. Kim, and J.G. Zeikus, *CRC Crit. Rev. Biotechnol.* **7**, 107–186 (1988).
34. R. Lamed and J.G. Zeikus, *J. Bacteriol.* **144**, 569–578 (1980).
35. H. Hippe, J.R. Andreesen, and G. Gottschall, in A. Balows, H.G. Truper, M. Dworkin, W. Harder, and K.H. Schleifer eds., *The Prokaryotes, a Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Applications*, Vol. 2, Springer-Verlag, New York, 1992, pp. 1800–1866.
36. B. McNeil and B. Kristensen, *Biotechnol. Lett.* **7**, 449 (1985).
37. H.A. George, J.L., Johnson, W.E.C. Moore, L.V. Holdeman, and J.S. Chen, *Appl. Environ. Microbiol.* **45**(3), 1160–1163 (1983).
38. D.T. Jones, J.R. Webster, and D.R. Woods, *J. Gen. Microbiol.* **116**, 195 (1980).
39. H.G. Wood and L.G. Ljungdahl, in J.M. Shively and L.L. Barton eds., *Variations in Autotrophic Life*, Academic Press, London, 1991, pp. 201–205.
40. H.G. Wood, *Annu. Rev. Biochem.* **54**, 1–41 (1985).
41. R. Kerby and J.G. Zeikus, *Curr. Microbiol.* **8**, 27–30 (1983).
42. G.J. Shen, B.A. Annous, R.W. Lovitt, M.K. Jain, and J.G. Zeikus, *Appl. Microbiol. Biotechnol.* **45**, 355–362 (1996).
43. J.G. Zeikus, *Ann. Rev. Microbiol.* **34**, 423–464 (1980).
44. T.E. Thompson, R. Conrad, and J.G. Zeikus, *FEMS Microbiol. Lett.* **22**, 265–271 (1984).
45. N. Samuelov, R. Lamed, S. Lowe, and J.G. Zeikus, *Appl. Environ. Microbiol.* **57**, 3013–3019 (1991).
46. M. Sebald, *Genetics and Molecular Biology of Anaerobic Bacteria*, Springer-Verlag, New York, 1993, pp. 1–703.
47. C. Lee, L. Bhatnagar, B.C. Saha, Y.-E. Lee, M. Takagi, T. Imanaka, M. Bagdasarian, and J.G. Zeikus, *Appl. Environ. Microbiol.* **56**, 2638–2643 (1990).
48. R.L. Robson, R.M. Robson, and J.G. Morris, *Biochem J.* **144**, 503–511 (1974).
49. M. Sebald and R.N. Costilow, *Appl. Microbiol.* **29**, 1–16 (1975).
50. I.R. Booth and J.G. Morris, *Biosci. Rep.* **2**, 47–53 (1982).
51. B.S. Mendez and R.F. Gomez, *Appl. Environ. Microbiol.* **43**, 495–496 (1982).
52. S.N. Bowring and J.G. Morris, *J. Appl. Bacteriol.* **58**, 577–584 (1985).
53. S.J. Reid, E.R. Allcock, D.T. Jones, and D.R. Woods, *Appl. Environ. Microbiol.* **45**, 305–307 (1983).
54. E. Stoutschek-Bauer, L. Hartl, and W.L. Staudenbauer, *Biotechnol. Lett.* **7**, 705–710 (1985).
55. J.D. Oultram, M. Loughlin, T.J. Swinfield, J.K. Brehm, D.E. Thompson, and N. Minton, *FEMS Microbiol. Lett.* **56**, 83–88 (1988).
56. M. Young, W.L. Staudenbauer, and N.P. Minton, in N.P. Minton and D.J. Clarke eds., Clostridia, Plenum Press, New York, 1989, pp. 63–103.
57. P.F. Levy, J.E. Sanderson, E. Ashare, and S.R. de Riel, *Enzyme Microbiol. Technol.* **3**, 207–215 (1981).
58. P.E. Liley, R.C. Reid, and E. Buck, in R.H. Perry, D.W. Green, and J.O. Maloney eds., *Perry's Chemical Engineers' Handbook*, 6th ed. McGraw-Hill, New York, 1950, pp. 3-1–3-291.
59. R. Marchal, M. Rebeller, and J.P. Vandecasteele, *Biotechnol. Lett.* **6**, 523–528 (1984).
60. J.A. Marlatt and R. Datta, *Biotech. Prog.* **2**, 23–28 (1986).
61. U. S. Patent 5,192,673 (March 9, 1993), M.K. Jain, D. Beacom, and R. Datta (to Michigan Biotechnology Institute, Lansing, Mich.).
62. B.K. Soni and M.K. Jain, *Bioprocess Eng.* **17**, 261–267 (1997).
63. B.K. Soni and M.K. Jain, *Bioprocess Eng.* **17**, 329–334 (1997).
64. D.R. Woods, *Trends Biotechnol.* **13**, 259–264 (1995).
65. D.T. Jones and D.R. Woods, *Microbiol. Rev.* **50**, 484–524 (1986).
66. H.P. Blaschek, *Food Technol.* **(October)**, 84–87 (1986).
67. M.R. Laddisch, *Enzyme Microbiol. Technol.* **13**, 280–283 (1991).
68. I.S. Maddox, N. Qureshi, and K. Roberts-Thomson, *Proc. Biochem.* **30**, 209–215 (1995).
69. M. Zhang, C. Eddy, K. Deanda, M. Finkelstein, and S. Picataggio, *Science* **267**, 240–243 (1995).
70. K. Deanda, M. Zhang, C. Eddy, and S. Picataggio, *Appl. Environ. Microbiol.* **62**, 4465–4470 (1996).
71. O.V.S. Reddy and S.C. Bassapa, *Biotechnol. Lett.* **18**, 1315–1318 (1996).
72. U.S. Patent 4,292,406 (September 29, 1981), L.G. Ljungdahl and J.K.W. Wiegel (to U.S. Department of Energy, Washington, D.C.).
73. U.S. Patent 4,292,407 (September 29, 1981), L.G. Ljungdahl and J.K.W. Wiegel (to U.S. Department of Energy, Washington, D.C.).
74. U.S. Patent 4,400,470 (August 23, 1983), J.G. Zeikus, T.K. Ng, A. Ben-Bassat, and R.J. Lamed (to Wisconsin Alumni Research Foundation, Madison, Wisc.).
75. U.S. Patent 4,385,117 (May 24, 1983), L.G. Ljungdahl and L.H. Carriera (to University of Georgia Research Foundation, Inc., Athens, Ga.).
76. D. Burdette and J.G. Zeikus, *Biochem. J.* **302**, 163–170 (1994).
77. R.W. Lovitt, R. Longin, and J.G. Zeikus, *Appl. Environ. Microbiol.* **48**, 764–770 (1984).
78. S. Jung, J.G. Zeikus, and R. Hollingsworth, *J. Lipid Research* **35**, 1057–1065 (1994).
79. H. Biebl, S. Marten, H. Hippe, and W.-D. Deckwer, *Appl. Microbiol. Biotechnol.* **36**, 592–597 (1992).
80. M.K. Jain, R. Datta, and J.G. Zeikus, in T.K. Ghose ed., *Bioprocess Engineering: The First Generation*, Ellis Harwood, Chichester, England, 1989, pp. 366–398.
81. L.G. Ljungdahl, L.H. Carreira, R.J. Garrison, N.E. Rabek, and J. Wiegel, *Biotech. Bioeng. Symp.* **15**, 207–223 (1985).
82. E.S. Lipinsky and R.G. Sinclair, *Chem. Eng. Prog.* **82**, 26–32 (1986).
83. Anonymous, *Biomass Process Handbook*, Technical Insights Inc., Fort Lee, N.J., 1981, pp. 72–110.
84. U.S. Patent 5,143,833 (September 1, 1992), R. Datta (to Michigan Biotechnology Institute, Lansing, Mich.).
85. N.P. Nghiem, B.H. Davison, B.E. Suttle, and G.R. Richardson, *Appl. Biochem. Biotechnol.*, **63–65**, 565–576 (1997).
86. U.S. Patent 5,143,834 (September 1, 1992), D. Glassner and R. Datta (to Michigan Biotechnology, Lansing, Mich.).
87. U.S. Patent 5,168,055 (December 1, 1992), R. Datta, D.A. Glassner, M.K. Jain and J.R. Vick Roy (to Michigan Biotechnology Institute, Lansing, Mich.).

88. U.S. Patent 5,521,075 (May 28, 1996), M.V. Guettler and M.K. Jain (to Michigan Biotechnology Institute, Lansing, Mich.).
89. S. Podkovyrov and J.G. Zeikus, *J. Gen. Microbiol.* **139**, 223–228 (1993).
90. L. Schöcke and P.J. Weimer, *Arch. Microbiol.* **167**, 289–294 (1997).
91. C.S. Millard, Y.P. Chao, J.C. Liao, and M.I. Donnelly, *Appl. Environ. Microbiol.* **62**, 1808–1810 (1996).
92. U.S. Patent 5,504,004 (April 2, 1996), M.V. Guettler, M.K. Jain, and B.K. Soni (to Michigan Biotechnology Institute, Lansing, Mich.).
93. M.J. van der Werf, M.V. Guettler, M.K. Jain, and J.G. Zeikus, *Arch. Microbiol.* **167**, 332–342 (1997).
94. R. Datta, *Biotech Bioeng. Symp.* **11**, 521–532 (1981).
95. E. Hsu and J. Ordal, *J. Bacteriol.* **102**, 369–376 (1970).
96. B.A. Annous, J.-S. Shieh, G.-J. Shen, M.K. Jain, and J.G. Zeikus, *Appl. Microbiol. Biotechnol.* **45**, 804–810 (1996).
97. R.M. Worden, A.J. Grethlein, J.G. Zeikus, and R. Datta, *App. Biochem. Biotech.* **20**, 687–698 (1989).
98. D. Michel-Savin, R. Marchal, and J.P. Vandecasteele, *Appl. Microbiol. Biotechnol.* **32**, 387–392 (1990).
99. D. Cowan, *Trends Biotechnol.* **14**, 177–178 (1996).
100. B.C. Saha, R. Lamed, and J.G. Zeikus, in N.P. Minton and D.J. Clarke eds., Clostridia, Plenum Publishing, New York, 1989, pp. 227–263.
101. C. Vieille, D.S. Burdette, and J.G. Zeikus, *Ann. Rev. Biotech.* **2**, 1–83 (1996).
102. H.H. Hyun and J.G. Zeikus, *Appl. Environ. Microbiol.* **49**, 1162–1167 (1985).
103. H.H. Hyun and J.G. Zeikus, *Appl. Environ. Microbiol.* **49**, 1168–1173 (1985).
104. B.E. Norman, *Starch / Starke* **34**, 340–346 (1982).
105. G. Dong, C. Vieille, A. Savchenko, and J.G. Zeikus, *Appl. Environ. Microbiol.* **63**, 3569–3576 (1997).
106. M.K. Jain and J.G. Zeikus, *Sys. Appl. Microbiol.* **10**, 134–141 (1988).
107. W.M. Mitchell and W.F. Harrington, in P.D. Boyer ed., *The Enzymes 3*, Academic, New York, 1971, pp. 699–719.
108. B. Schink and J.G. Zeikus, *FEMS Microbiol. Lett.* **17**, 295–298 (1983).
109. C. Vieille, J.M. Hess, R.M. Kelly, and J.G. Zeikus, *Appl. Environ. Microbiol.* **61**, 1867–1875 (1995).
110. G. Dong and J.G. Zeikus, *Enzyme Microbial Technol.* **21**, 335–340 (1997).
111. T.K. Ng, P.J. Weimer, and J.G. Zeikus, *Arch. Microbiol.* **114**, 1–7 (1977).
112. S.P. Mathupala and J.G. Zeikus, *Appl. Microbiol. Biotechnol.* **39**, 487–493 (1993).
113. G. Dong, C. Vieille, and J.G. Zeikus, *Appl. Environ. Microbiol.* **63**, 3577–3584 (1997).
114. Y.-E. Lee, S.E. Lowe, and J.G. Zeikus, *Appl. Environ. Microbiol.* **59**, 763–771 (1993).
115. K. Ma, F.T. Robb, and M.W.W. Adams, *Appl. Environ. Microbiol.* **60**, 562–568 (1994).
116. D.S. Burdette, C. Vieille, and J.G. Zeikus, *Biochem. J.* **316**, 115–122 (1996).
117. Y.E. Lee and J.G. Zeikus, *J. Gen. Microbiol.* **139**, 1235–1243 (1993).
118. S. Barik, S. Prieto, S.B. Harrison, E.C. Clausen, and J.L. Gaddy, *Appl. Biochem. Biotechnol.* **18**, 363–378 (1988).
119. A.J. Grethlein, R.M. Worden, M.K. Jain, and R. Datta, *J. Appl. Biochem.* **24/25**, 875–884, (1990).
120. A.J. Grethlein and M.K. Jain, *Trends Biotechnol.* **10**, 418–423 (1992).
121. R.W. Jack, J.R. Tagg, and B. Ray, *Microbiol. Rev.* **59**, 171–200 (1995).
122. Y. Hanaoka, *Hakkokogaku* **45**, 312–319 (1967).
123. R.J. Lamed and J.G. Zeikus, *Biochem. J.* **195**, 183–190 (1981).
124. R.J. Lamed, E. Keinan, and J.G. Zeikus, *Enzyme Microbiol. Technol.* **3**, 144–148 (1981).
125. R.W. Lovitt, E.W. James, D.B. Kell, and J.G. Morris, in G.W. Moody and P.B. Baker eds., *Bioreactors and Biotransformations*, Elsevier Applied Science, London, 1987, pp. 263–276.
126. W. Buckel, *FEMS Microbiol. Rev.* **88**, 211–232 (1992).
127. J. Thiele, M. Chartrain, and J.G. Zeikus, *Appl. Environ. Microbiol.* **54**, 10–19 (1988).
128. J.H. Thiele and J.G. Ziekus, in L.E. Erickson and D.Y.-C. Fung eds., *Handbook on Anaerobic Fermentations*, Marcel Dekker, New York, 1988, pp. 537–595.
129. S. Ghosh, in T. Noike, A. Tilche, and K. Hanaki eds., *Proceedings of the 8th International Conference on Anaerobic Digestion*, vol. 1, Sendai, Japan, 1997, pp. 9–16.
130. S. Rebac, J. Ruskova, S. Gerbens, J.B. van Lier, A.J.M. Stams, and G. Lettinga, *J. Ferm. Bioeng.* **80**, 499–506 (1995).
131. V. Gellens, J. Boelens, W. Verstraete, *Antonie van Leewenhoek* **67**, 79–89 (1995).
132. C.J.N. Buisman, in *Wider Application and Diffusion of Bioremediation Technologies: The Amsterdam '95 Workshop*, Organization for Economic Cooperation and Development, Paris, 1996, pp. 103–114.
133. P.N. Lens and W. Verstraete, in T.G. Villa and J. Abalde eds., *Profiles on Biotechnology*, Universidade de Santiago, Spain, 1992, pp. 333–356.
134. W. Verstraete and P. Vandevivere, in T. Noike, A. Tilche, and K. Hanaki eds., *Proceedings of the 8th International Conference on Anaerobic Digestion*, vol. 1, Sendai, Japan, 1997, pp. 67–74.
135. M. Strous, E. VanGerven, P. Zheng, J.G. Kuenen, and M.S.M. Jetten, *Water Res.* **31**, 1955–1962 (1997).
136. R.A. Wirtz and R.R. Dague, *Water Environ. Res.* **68**, 883–892 (1996).
137. J. Thaveesri, D. Daffonchio, B. Liessens, P. Vandermeren, and W. Verstraete, *Appl. Environ. Microbiol.* **61**, 3681–3686 (1995).
138. W.-M. Wu, M.K. Jain, and J.G. Zeikus, *Appl. Environ. Microbiol.* **62**, 2037–2044 (1996).
139. M.C. Veiga, M.K. Jain, W.-M. Wu, R.I. Hollingsworth, and J.G. Zeikus, *Appl. Environ. Microbiol.* **63**, 403–407 (1997).
140. S.R. Berchtold, S.L. Vanderloop, M.T. Suidan, and S.W. Maloney, *Water Environ. Res.* **67**, 1081–1091 (1995).
141. O. Meyer, R.I. Refae, J. Warrelmann, and H. von Reis, *Microb. Releases* **2**, 11–22 (1993).
142. W.H. Parker, E.R. Hall, and G.J. Farquhar, *Water Environ. Res.* **65**, 264–270 (1993).
143. W.-M. Wu, J. Nye, R.F. Hickey, M.K. Jain, and J.G. Zeikus, in R.E. Hinchee, A. Leeson and L. Semprini eds., *Bioremediation of Chlorinated Solvents*, Battelle Press, Columbus, Ohio, 1995, pp. 45–52.
144. X. Maym-Gatell, Y.T. Chien, J.M. Gossett, and S.H. Zinder, *Science* **276**, 1568–1571 (1997).
145. T. Gorontzy, J. Küver, and K.-H. Blotevogel, *J. Gen. Microbiol.* **139**, 1331–1336 (1993).

146. R. Boopathy and C.F. Kulpa, *Can. J. Microbiol.* **39**, 430–433 (1993).

147. A. Preuss, J. Fimpel, and G. Diekert, *Arch. Microbiol.* **159**, 345–353 (1993).

148. T.A. Khan, R. Bhadra, and J. Hughes, *J. Ind. Microbiol. Biotechnol.* **18**, 198–203 (1997).

149. S.B. Funk, D.J. Roberts, D.L. Crawford, and R.L. Crawford, *Appl. Environ. Microbiol.* **59**, 2171–2177 (1993).

150. R.H. Kaake, D.L. Crawford, and R.L. Crawford, in R.E. Hinchee, D.B. Richardson, F.B. Metting, Jr., and G. Sayles eds., *Applied Biotechnology for Site Remediation*, Lewis Publishers, Boca Raton, Fla., 1995, pp. 337–341.

151. J.P. Maree, A. Gerber, and E. Hill, *J. Water Pollut. Control Fed.* **59**, 1069–1074 (1987).

152. E. Colleran, S. Finnegan, and P. Lens, *Antonie van Leewenhoek*, **67**, 29–46 (1995).

153. P. Yspeert and C. Buisman, in Y. Ando and S. Wald eds., *Bioremediation: The Tokyo '94 Workshop*, Organization for Economic Cooperation and Development, Paris, 1995, pp. 335–344.

154. D.J. Cork and M.A. Cusanovich, in L.E. Murr, A.E. Torma, and J.A. Brierly eds., *Metallurgical Applications of Bacterial Leaching and Related Microbiological Phenomena, Proceedings of International Symposium*, Academic Press, New York, 1978, pp. 207–221.

ANIMAL CELLS USED IN MANUFACTURING

CLAUDIA BARDOUILLE
ECACC CAMR Porton Down
Salisbury, Wiltshire
United Kingdom

KEY WORDS

Animal cell line
Biopharmaceutical production
Posttranslational modifications
Protein
Recombinant
Three-dimensional culture

OUTLINE

INTRODUCTION

Animal cell lines are a resource in biotechnological processes. They are used increasingly in molecular biology developments, fermentation technology, the production of diverse health care products, and cell-based screening systems in toxicology and pharmacology. Large-scale technology began with the use of primary monkey kidney cells for the production of poliomyelitis vaccines in the 1950s. After intense debates on safety issues, human diploid cell strains (e.g., WI-38) were accepted for the production of mumps, measles, and rubella vaccines in the 1960s. The next step was taken in 1964 with the commercial production of foot-and-mouth disease virus (FMDV) from baby hamster kidney (BHK 21) cells for veterinary purposes. After the discovery of interferons and their clinical importance, the growing demand for these compounds led to the search for alternative, easily accessible sources. Interferon was the first licensed product derived from Namalwa cells, a human lymphoblastoid heteroploid cell line, in the late 1970s. Soon after this, hybridoma and monoclonal antibodies (mAb) took over the fast-developing market with their diverse applications in diagnostic and therapy.

Although strict safety standards have to be met, the value of animal cell lines is undeniable, and they are indispensable for the industrial production of eukaryotic proteins.

Since the acceptance of cell lines for the production of vaccines, cytokines, mAb, and therapeutic proteins or even the cells as product themselves, constant progress in a broad range of areas has contributed to the improvement of products. Research fields having major impact are analysis and adaptation of suitable cell lines, high-level expression systems, genetically engineered cell lines and products (e.g., protein modifications), scale-up methodology, bioreactor development, and downstream techniques (e.g., purification, quality control).

This chapter presents an overview of cell lines used for various applications and assays and gives an insight into different technologies and products.

ANIMAL CELL TYPES

The purification of therapeutic products from natural sources was and still is restricted because of limited supply, and contamination of known or unknown origin presents a serious danger (e.g., HIV in blood products). In addition, bacterial systems are not capable of the posttranslational modifications eukaryotic proteins undergo. Animal cell technology has been able to fill this gap and provide proteins of high quality on an industrial scale. Cell lines used most frequently for different production systems are described in the following sections.

Chinese Hamster Ovary

The anchorage-dependent cell line from Chinese hamster ovary (CHO) was derived from a biopsy of an ovary of an adult Chinese hamster *(Cricetulus griseus)* in 1957. Various derivatives were developed, including nutritionally deficient mutants and transfected cells. The most popular variants are CHO-K1, a proline-requiring mutant, and CHO/dhFr_, which is deficient in dihydrofolate reductase and therefore useful for the selection of transfected cells. Other variants have been adapted to grow in suspension and in serum-free media.

Advantages of CHO are their high capacity of amplification and expression of recombinant genes and proteins, the ability to synthesize proteins with similar glycosylation patterns as observed in native proteins, and the ability to grow in large-scale bioreactors. Therefore, they are very suitable for the expression of heterologous proteins by recombinant DNA technology. Table 1 presents examples of proteins produced by transfected CHO cells in a suspension or attached culture system. Because CHO cells are the most important substrate for the production of therapeutic proteins, new products are developed and licensed continuously (1). Product status, either in research or as industrial process, is indicated in the tables under Application.

Generally, CHO cells are kept in attached culture on tissue culture material or on microcarriers (see "Three-Dimensional Culture Systems"), but they can readily be adapted to grow as single cells or aggregates in suspension. This allows the usage of highly efficient bioreactors such as fluidized bed reactors (16). Because aggregates can lead to nonhomogenous product formation and affect system monitoring, single-cell suspension cultures are often preferred. To achieve this status, media formulations have been developed maintaining single-cell suspension during the production process (17). Cell lines that have been adapted to grow in suspension do not maintain all characteristics seen in attached cultivation. In a study by Brand et al. (18), it was shown that specific production rates of cells in attached cultures are not necessarily connected to similar rates after adaptation to suspension. This observation illustrates the difficulties of scale-up procedures.

Serum-containing media impede the purification of pharmaceutical products to a great extent, and therefore efforts are directed to establish cultures in serum-reduced, serum-free, or even protein-free media (19). To facilitate the scale-up of transfected cells, CHO cells preadapted to serum-free and suspension culture have been developed by Sinacore et al. (13). Other aspects important for efficient production are the study and regulation of inhibitory substances such as ammonia (20) that can exert an influence on the glycosylation pattern (21).

A reproducible product consistency is of course a major aim in all animal cell systems used for the manufacture of proteins to be administered as therapeutic agents. The effect of genetic and environmental factors on purity and posttranslational modifications are discussed in "Post-translational Modifications of Recombinant Proteins in Animal Cell Lines."

Baby Hamster Kidney Cells

The attached BHK 21 cell line was derived from five unsexed 1-day-old Syrian hamster (*Mesocricetus auratus*) kidneys in 1961. BHK 21 clone 13 was obtained after single-cell isolation and is now widely used, very often only addressed as BHK or BHK 21. Several variants have been developed that are being adapted to suspension culture or suitable for genetic engineering techniques.

Because of their susceptibility to various viruses, they have been extensively used for the production of vaccines. Since the 1960s, BHK cells have played the major role in the production of FMDV vaccine in industrial-scale fermenters (with a capacity of several thousand liters) for veterinary applications. The experience and know-how gained with this manufacturing system led to the increased use of animal cells and the improvement of process parameters. Like CHO, BHK cells show spontaneous aggregation that can be advantageous for cell retention and recycling systems (22). Efforts have also been directed toward high-cell-density culture and serum- or protein-free media formulations (23). The first product approved by the Food and Drug Administration (FDA) from BHK cells cultured in perfusion reactors was the recombinant human factor VIII for the therapy of coagulation disorders. Since then, many other proteins have been produced on laboratory- and industrial-scales using this versatile cell line, ranging from human erythropoietin and antithrombin III to enzymes and a range of vaccines for veterinary or human health applications (Table 2).

Hybridoma and Myeloma Cell Lines

Since Köhler and Milstein (30) discovered the technique to produce immortalized, antibody-secreting cells—hybridomas—the development in this area has had a major impact on the production and application of the secreted proteins (mAb). Therefore, it is not surprising that this tool occupies an important and still growing place in industrial biotechnology because the range of applications is constantly expanding. The advantages of hybridomas and myelomas (i.e., their natural secretory ability and, as lymphocytes, growth in suspension) renders these cell lines ideal for exploitation in biotechnical processes.

To optimize the production of mAb, efforts have been undertaken to develop efficient and economic media, including serum- and protein-free formulations (31), generating high cell densities by using perfusion reactors instead of batch processes (32), and selection of suitable clones with higher resistance against shear stress in agitated bioreactors (14). In addition to the ready-to-use serum-free media that are already commercially available, cultivation of hybridomas is constantly improved by research activities related to specific needs. Examples are cholesterol-free media or the use of genetically engineered cell lines that are glutamine-independent, which reduces the accumulation of inhibitory substances such as ammonia (33).

The ability to produce recombinant mAb contributed to the development of chimeric mAb containing several therapeutic advantages. It is now possible to construct heterohybridoma cell lines where different species can contrib-

Table 1. Recombinant Proteins Expressed in CHO cells

Product	Application[a]	Attached/suspension	Serum-free	Ref.
γ-Glutamyl transferase (hu)	Reference enzyme for clinical tests (diagnosis of hepatic cancer, alcohol abuse) (RA)	Attached	No	2
Humanized IgG1 binding CD52 (CAMPATH-IH®)	Transplant rejection, rheumatoid arthritis, non-Hodgkin's lymphoma (IP)	Suspension	Yes	3
Chimeric human/murine mAb reacting with tumor-associated glycoprotein TAG72	Tumor imaging and therapy (IP)	Suspension	Yes	4
Thyroid peroxidase (hu)	Testing of antibodies in autoimmune thyroid diseases (RA)	Attached	Yes	5
Interferon-γ	Immunostimulation; tumor therapy; antiviral therapeutic agent (IP)	Suspension	No	6
HIV 1 gp120	HIV research (RA)	Attached	No	7
Tissue plasminogen activator (t-PA) (hu); trade name Activase	Thrombolytic agent, acute myocardial infarction (IP)	Suspension	Yes	8,9
Antithrombin III (hu)	Antithrombin III deficiencies (IP)	n.k.	Yes	10
von Willebrand's factor (hu)	Hemostatic factor (IP)	Attached	Yes	11
Erythropoietin receptor (mu)	Study of physiological function detection assays (RA)	Attached	No	12
Erythropoietin; trade name Epogen/Procrit	Myocardial infarction (IP)	n.k.	n.k.	1
Macrophage colony-stimulating factor (hu)	Cytokine study (IP)	Suspension	Yes	13
Factor IX (hu)	Blood clotting factor (IP)	Suspension	Yes	13
Nerve growth factor (hu)	Bioactivity study (IP)	Suspension	Yes	14
Peptidylglycine α-amidating enzyme (rat)	Catalytic reagent for calcitonin production (IP)	Attached/Suspension	Yes	15

Note: hu, human; n.k., not known; mu, murine.
[a]RA, research application; IP, industrial process.

ute to the variant (e.g., mouse) and constant (e.g., human) region of the secreted mAb. These antibodies are less immunogenic and more effective in therapeutical applications. The constant region can also be replaced with enzymes or toxins, which is useful in tumor targeting. Bispecific mAb can react with two antigens (e.g., tumor markers and specific toxic drugs) and thereby bind and destroy specific cells only (34). Details of techniques for the construction of these new generations of hybridomas, bioreactor systems, purification, and applications are reviewed in James (35), Larrick and Balint (36), and Merten (37).

Because of the rapid development of new cell lines in this area, only a few examples were selected for inclusion in Table 3. Table 4 overviews the various applications for mAb and the potential for future developments in animal cell biotechnology.

Myeloma cell lines such as J558 L, NS0, SP2/0, or P3X63Ag.653 are not only malignant fusion partners used for the generation of hybridomas, but are also suitable for the production of recombinant proteins (40,41). Once again, it can be attributed to the tremendous new techniques of genetic engineering that opened the door for myeloma-based expression systems leading to a new branch in animal cell technology. These naturally secretory cells can be grown to high cell densities and are robust and relatively easy to transfect (42). With the use of recombinant myeloma cell lines, raising of antigen-specific B cells as prerequisite for the production of hybridomas and finally mAb can be avoided, and time- as well as cost-consuming optimization procedures can be limited.

Insect Cell Lines

In addition to mammalian cells, insect cell lines for the production of recombinant proteins have been established in the past decade. The techniques used most frequently can be distinguished by the application of different expression systems: the baculovirus expression vector system (BEVS) and the *Drosophila* metallothionein (Mt) promotor system.

Baculovirus Expression Systems. Baculoviruses replicate exclusively in invertebrates and are often highly species-specific. The BEVS most commonly used for the expression of heterologous proteins is the *Autographa californica* (alfalfa looper) nuclear polyhedrosis virus (AcNPV), where the gene of choice is inserted into the baculovirus genome under the control of a strong late viral transcription promoter. The recombinant virus is then used to infect lepidopteran insect cells, mainly from *Spodoptera frugiperda* (fall armyworm). These cells express the recombinant protein during the late stages of the infection process. Several human, animal, and plant proteins have been expressed using this system (43).

The insect cell line used most frequently in combination with BEVS is SF 9; the parental cell line, SF 21, was obtained from pupal ovarian tissue of *S. frugiperda*. SF 9 cells are highly susceptible to infection with various baculoviruses. A main advantage of the baculovirus/insect cell system is its potential to yield milligram quantities of recombinant product. To obtain mammalian cell lines producing reasonable amounts of heterologous proteins, it is

Table 2. Recombinant Proteins Expressed in BHK Cells

Product	Application[a]	Attached/suspension	Serum-free	Ref.
FMDV	Vaccine (cattle and pigs)	Suspension	No	24
Rabies	Vaccine	Suspension	No	25
Antithrombin III (hu)	Antithrombin III deficiencies	Suspension/attached	Yes	26
Erythropoietin (hu)	Regulation of red blood cell formation	Attached	No	27
Interleukin-2 (hu)	Immunodeficiencies, leukemia, cancer	Attached	Protein-free	28
Transferrin (hu)	Iron-transporting protein	Attached	No	29
Placental-secreted alkaline phosphatase	Aggregation studies	Suspension	No	22

Note: hu, human.
[a]All industrial processes.

Table 3. Hybridoma Cell Lines

Cell Line	MAb Specificity	Application[a]	Ref.
Orthoclone OKT®3 (mouse)	Anti-CD3	Suppression of rejection of transplanted kidneys	1
N12-16.63 (mouse-hu-hu, heterohybridoma)	Anti-tetanus toxoid	Tetanus therapy, diagnosis	14
HBW-4.16 (hu-hu; heterohybridoma)	Anti-hepatitis B virus	Hepatitis B	14
CB-STL-1 (hu-hu-mouse; heterohybridoma)	Anti-α-hemolysin	Sepsis, septic shock	38
2F5 (heterohybridoma)	HIV-1 neutralizing	HIV-1 therapy	39

[a]All industrial processes.

Table 4. Applications for mAb

Diagnostic
Clinical: virus, bacteria, parasite detection, tumor marker, allergy
Self-testing: pregnancy and fertility tests
Therapeutic (human and veterinary)
Vaccines
Growth factors, hormones
Disease treatment, e.g., autoimmune disorders, cancer, AIDS, infections, toxifications
Industrial purifications
Quality control laboratories

not only necessary to invest valuable time on transfection, selection, and optimization, but the resulting cell line might also show a basic constitutive expression activity and is constantly under selective pressure.

***Drosophila* Metallothionein Expression System.** In the *Drosophila* Mt. system, the inducible Mt. promotor is used to control the expression of recombinant proteins. This system offers the benefit of a well-known host and genomic structure with a tightly regulated promotor. In contrast to the viral system BEVS, where the production peak is found when the cells already die from the viral infection, *Drosophila* cells and the Mt. expression system can be used continuously.

Despite the different expression systems for insect cell lines, development of efficient fermentation techniques and process optimization are similar to the ones already described for mammalian cell types. High-cell-density reactors to achieve a maximum yield of desired product have been developed by various groups. Cavegn et al. established a system monitoring the production of insulin-receptor tyrosine kinase domain and the soluble part of endothelial leukocyte adhesion molecule as model proteins (44) and achieved up to 3×10^7 cells/mL in a serum-free perfusion bioreactor. Other developments are immobilized systems, such as cultivation on porous microcarrier (45), or the use of simulated microgravity to reduce hydrodynamic forces, as described by O'Connor et al. applying the rotating-wall vessel (46).

Examples of heterologous proteins expressed in insect cell lines can be found in Table 5. Investigations at the laboratory level indicate interesting areas still to be exploited for biotechnological purposes. The high level of expression of measles virus protein in SF 9 facilitates research into structure, function, and immunogenicity (53). Expression of recombinant bovine β-lactoglobulin for the processing of milk proteins to inhibit allergic reactions is another potential application (54). Because of the ability to express two or more proteins simultaneously in the baculovirus system, analysis of multisubunit formation (e.g., for bluetongue viruslike particles) is possible (49).

Research activities are also concentrating on other insect host cells to improve the production process and, as an important requirement for clinically administered compounds, obtain a correctly glycosylated and biological active protein. SF 9 cells display only a very simple glycosylation capacity, whereas *Estigmena acrea* cells are able to produce complex glycoforms very similar to their authentic counterparts (55). Other insect cell lines investigated for the manufacture of recombinant proteins are Schneider 2 *(Drosophila melanogaster)* and Tr 5 *(Trichoplusia ni)*. Apart from using insect cell lines to facilitate production, they can also be applied in screening or cytotoxicity tests.

Other Cell Lines

Although the cell lines described in the previous section probably contribute to the majority of industrial processes,

Table 5. Recombinant Proteins Expressed in Insect Cell Lines

Cell Line	Product	Application[a]	Expression system	Ref.
SF 9	Soluble human TNF receptor	Biochemical studies, structure-function analysis	BEVS	47
SF 9	Soluble IFN-γ receptor	Structure determination, identification of antagonists, rational drug design	BEVS	48
SF 9	Bluetongue viruslike particles	Vaccine in sheep and cattle	BEVS	49
SF 9/Schneider 2	VCAM	Model system	BEVS/Mt	50
SF 9/Tn 5	Osteoblast-specific factor 2	Model system	BEVS	51
SF 9	Prostate specific antigen	Tumor marker	BEVS	52

Note: TNF, tumor necrosis factor; IFN, interferon; VCAM, vascular cell adhesion molecule.
[a]All research applications

various other cell types are used as substrate for the production of biologicals.

On the one hand, this might be of historical reason. Certain cell lines are well established; the production of vaccines and systems for manufacture have been optimized already. On the other hand, cell lines synthesizing proteins naturally without transformation procedures have clear advantages considering the production process (where a selective pressure to maintain expression is hence not necessary) and acceptance of the final product. Other cell lines have been shown to possess valuable properties such as glycosylation characteristics. This section describes some of these cell lines and their biotechnological usage; although not great in number, they are still significant in the pharmaceutical industry.

VERO cells, derived from the kidney of a normal adult African green monkey, are susceptible to a wide range of viruses, including measles, polioviruses, and rubella. These properties led to the early establishment of vaccine production in VERO cells (for polio) (Table 6). This now rather classical cell line is still used as capable substrate for virus production. VERO cells replaced primary monkey kidney cells for the production of oral polio vaccine and are now cultured on microcarriers in large-scale bioreactors (68). They were also successfully applied for the study of infection with Aujeszky virus in the VERAX system, a perfused fluidized bed with macroporous collagen carriers as microspheres to evaluate the possibility of vaccine production in a high-cell-density system (67). Baijot et al. (65) established the manufacture of an Aujeszky vaccine using swine testicular cells (NLST) on microcarriers. Although VERO cells are mostly grown as attached cells, they can be adapted to grow in suspension in serum-free medium (70). The inducible expression of HIV-1 viruslike particles in transfected VERO and COS 7 cells (a SV 40 transformed monkey kidney cell line) was described by Haynes et al. (69). This important development offers an exciting alternative to the use of live virus vectors for the production and evaluation of AIDS vaccines based on noninfectious particles.

The human lymphoblastoid cell line Namalwa has been used as substrate for biologicals since the 1970s. The secretion of INF-α being naturally induced in Namalwa cells by treatment with sodium butyrate and addition of sendai virus is well established for industrial manufacturing (61). Process optimization and increase of productivity are still continuing and achieved by various modifications, such as treatment with tetramethyl urea (62) or sequential induction with sendai virus (63). Namalwa KJM 1 was derived from Namalwa and adapted to grow in serum-free medium. Miyaji et al. (71) describe the efficient expression of recombinant human INF-β in these cells in high-cell-density perfusion culture. The comparison of recombinant pro-urokinase expression in Namalwa KJM 1 and CHO cells revealed a partly cleaved and thereby inactive form secreted in CHO cells due to proteases present in the supernatant (64). Because the protein secreted by Namalwa KJM 1 showed a single-chain form, this indicates that

Table 6. Examples of Cell lines and Their Use in Biotechnological Processes

Cell Line	Product	Application[a]	Ref.
293 (human embryonal kidney, Ad5 transformed)	Acetylcholinesterase	Structural and functional analysis (RA)	56
293S (Adapted to suspension)	Tyrosine phosphatase 1C	Model system for scale-up (RA)	57
HuCCl (Human colon adenocarcinoma)	Carcinoembryonic antigen	Tumor-associated marker (IP)	58
HUH-6 (Human hepatoblastoma)	Fibronectin	Attachment factor (RA)	59
MDBK (Madin-Darby bovine kidney)	BVD virus (bovine viral diarrhea)	Vaccine (IP)	60
Namalwa (human lymphoblastoid)	IFN-α	Pharmaceutical agent (IP)	61,62,63
Namalwa KJM 1	Pro-urokinase	Model system (RA)	64
NLST (swine testicular)	Aujeszky virus	Vaccine (IP)	65
RPMI 8866 (human B cells)	CD 23 fragment	Structural and functional analysis (RA)	66
VERO (monkey kidney)	Aujeszky virus	Vaccine (IP)	67
VERO	Oral polio virus	Vaccine (IP)	68
VERO and COS-7 (monkey kidney)	HIV-1 viruslike particles	Vaccine (IP)	69

[a]RA, research applications; IP, industrial process.

these cells are more suitable for the production of protease-susceptible heterologous proteins.

The carcinoembryonic antigen (CEA) is a high molecular weight glycoprotein identified as tumor-associated marker. For the purpose of early clinical diagnosis, therapy response, and recurrences, it is important to be able to detect CEA in monitored patients. Several cell lines synthesize CEA and have been used in various large-scale production systems (58). mAb raised against this antigen can improve clinical detection assays and cancer specificity. Other examples of cell lines used for the production of naturally secreted proteins are HUH 6 (fibronectin) and RPMI 8866 (CD 23) (see Table 6 for details).

A list of all cell lines currently used in biotechnology is clearly beyond the scope of this article. The constant improvements in this field, such as transformation of cell types without losing specific properties or differentiation characteristics (e.g., hepatocytes, neurons, keratinocytes) and optimization of cell lines for fermentation processes, will certainly lead to a variety of options for animal cell biotechnologists.

THREE-DIMENSIONAL CULTURE SYSTEMS

Animal cell lines have to be classified according to their ability to grow in suspension or be anchorage-dependent. For the large-scale production of biologicals, suspension cells are clearly the substrate of choice. With the development of microcarriers initiated by van Wezel 1967 (72), it is now possible to culture attached cells in bioreactors very similar to suspension cells. Growth of animal cells on microcarriers or immobilized systems facilitates production processes and allows high yield of cells, heterologous proteins, or viruses (73).

Microcarrier techniques are especially important for the handling of primary cell types or cell strains because these are, with only a few exceptions, mostly anchorage-dependent. The microcarrier technology offers several advantages: providing a high surface-to-volume ratio; allowing manipulation during culture without interrupting the fermentation process, and facilitating scale-up possibilities (74). A further advantage is the stabilization of cells in culture, which applies to both attached and suspension cells, thereby permitting prolonged culture periods. Macroporous microcarriers have proved to be suitable for the protection of cells from shear forces and have been applied successfully for the immobilization of adherent and suspension cells (75–77).

The ability of some cell lines to form aggregates is another feature used for increased productivity in perfusion bioreactors (16). However, aggregation has to be evaluated in a case-to-case study because it can result in inhomogeneous expression of proteins and might impair monitoring (17).

Finally, three-dimensional culture conditions can simulate the natural environment of cells, including cell–cell interaction and differentiation processes that are not always possible to achieve in two-dimensional systems. An example is the establishment of a bioreactor that provides high, almost tissuelike cell densities and thereby simulates cell–cell interactions (78).

POSTTRANSLATIONAL MODIFICATIONS OF RECOMBINANT PROTEINS IN ANIMAL CELL LINES

Recombinant DNA technology opened the field for the production of an unlimited variety of clinically relevant therapeutics in various host cell lines. The majority of proteins applied undergo more or less extensive posttranslational modifications in their natural counterparts, including mainly glycosylation, but also phosporylation, carboxylation, and signal-peptide processing. These modifications affect the biological activity of the heterologous protein by being involved in solubility, pharmacokinetics, antigeneicity, circulating half-life, secretion, protein folding, oligomer assembly, and susceptibility to proteolytic attack (79,80).

Choosing a host cell line able to fulfill these criteria is, therefore, important. Regulatory authorities such as the FDA demand a comprehensive analysis of carbohydrate structures and consistency in the production process. Because prokaryotes are not capable of performing posttranslational modifications observed in eukaryotic cells, they cannot be utilized for the production of recombinant proteins, where bioactivity relies heavily on correct glycosylation, for example. These unique advantages of animal cell technology inspired research directed toward overcoming process or regulatory difficulties.

Research has increasingly revealed already substantial differences in the glycosylation of recombinant proteins using several host cell lines or even transgenic animals. Variations can also stem from different environmental factors, such as media or bioreactor configuration, or the internal cell status at the time of production (80).

A few examples illustrate the importance of choosing a host system able to produce heterologous proteins with appropriate characteristics for clinical applications. Although CHO and BHK lack certain functional glycosyltransferases found in human cells (80), they still represent the most favorable host cell lines for large-scale manufacture of protein-based pharmaceuticals such as erythropoietin or tissue-type plasminogen activator (t-PA). Studies on product consistency revealed an independence of glycosylation pattern by applying varied process parameters (8) and showed that glycosylation was identical to natural sources (81). Constant improvement in the methods available for the analysis of carbohydrate structures allow a detailed pattern identification (79) and have been used to detect species-specific oligosaccharide variations of recombinant human IFN-γ produced in CHO and SF cells and transgenic mice (82).

The glycosylation status of recombinant proteins synthesized in baculovirus-infected insect cells can vary quite considerably, as shown by Ogonah et al. (55). *Estigmena* cells are capable of producing complex oligosaccharides, whereas the still more frequently used SF 9 cell line is restricted to the performance of only simple glycosylation reactions.

These findings demonstrate the importance of investigating and identifying an animal cell system best suited for the production of therapeutics. Even shortcomings, such as the lack of glycosyltransferases in CHO cells, can

be overcome by transfection of appropriate enzymes (80), indicating the potential for future processes.

TRENDS

Several examples given in the preceding sections already point toward activities undertaken to optimize animal cell technology and bioproducts of interest.

Cell lines other than the well-established hamster lines CHO and BHK present valuable alternatives, for example, the efficient production of recombinant proteins. Concomitant with an increasing understanding of media formulations, bioreactor design, and processes, efficiency and quality are still subject to improvement. Tailored cell lines can be selected by appropriate cloning strategies, using either naturally occurring variants or conferring specific characteristics by recombinant DNA methods (33).

The development of vectors with a predictable integration site for foreign DNA into the host genomic structure, including efficient expression systems, can overcome low-productivity problems. An example for enhanced productivity is the oncogene-activated system applied by Ternya et al. (83), where cells are cotransfected with the *ras* oncogene.

Recombinant DNA technology gives new directions for the design of therapeutics with improved or even novel characteristics compared to their natural counterparts (e.g., pharmacokinetic properties). Ahern et al. (84) investigated recombinant variants of tPA produced by site-directed mutagenesis and identified proteins with increased fibrinolytic activity in vitro and decreased clearance in vivo.

New technologies, such as the encapsulation of cells, genetically engineered or as primary isolates, and transplantation into patients with various disorders or illnesses, could provide a cellular therapy tool for long-term treatment or even cure (85).

Other approaches attempt to simulate the three-dimensional microenvironment for constructing organotypical systems, such as artificial liver support devices (86) or drug metabolism studies (87). Investigations into the involvement of cytokines, extracellular matrixes, and cell–cell interactions are required to optimize organ models and lead to successful applications.

CONCLUSION

Animal cell lines are an established tool in biotechnology and will be used extensively in a variety of emerging areas, such as tissue engineering, organ replacement, or drug discovery. A vast amount of knowledge and experience has accumulated concerning the cell lines described in this article, and research into improved industrial large-scale use is still on going. Safety concerns using continuous cell lines with a tumorigenic potential are addressed frequently. As a result of numerous efforts made in response to recommendations by regulatory bodies and the World Health Organization, cell lines and derived products are now a prosperous factor in biotechnology.

Selecting a cell line for the manufacture of pharmaceuticals is dependent on the protein that has to be produced. Post-translational modifications, secretion, and purification are aspects that have to be considered, but species (e.g., human or rodent origin) and regulatory issues are equally important. The examples described in this article give some indications on cell line suitability, but research as well as industrial production is improving rapidly in this area. New sources (i.e., cell lines and methods) are being established continuously and should therefore be closely studied before choosing an appropriate cell substrate.

BIBLIOGRAPHY

1. A.S. Lubiniecki and S.A. Vargo, in A.S. Lubiniecki and S.A. Vargo eds., *Regulatory Practice for Biopharmaceutical Production*, Wiley-Liss, New York, 1994, pp. 1–12.
2. J.L. Goergen, I. Chevalot, A. Visvikis, T. Oster, K. Hess, J.M. Engasser, and A. Marc, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1992, pp. 593–595.
3. M.J. Keen and N.T. Rapson, *Cytotechnology* **17**, 153–163 (1995).
4. R.P. Field, H. Brand, G.L. Renner, H.A. Robertson, and R. Boraston, in R.E. Spier, J.B. Griffiths, and B. Meignier eds., *Production of Biologicals from Animal Cells in Culture*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1991, pp. 742–744.
5. T. Ii, M. Murakami, M. Mizuguchi, K. Matsumoto, and K. Onodera, in E.C. Beuvery, J.B. Griffiths, and W.P. Zeijlemaker eds., *Animal Cell Technology: Developments towards the 21st Century*, Proc. of the Veldhoven Mtg., Kluwer, Dordrecht, The Netherlands, 1995, pp. 385–389.
6. V. Leelavatcharamas, A.N. Emery, and M. Al-Rubeai, *Cytotechnology* **15**, 65–71 (1994).
7. S.J. Froud, G.J. Clements, M.E. Doyle, E.L.V. Harris, C. Lloyd, P. Murray, A. Preneta, P.E. Stephens, S. Thompson, and G.T. Yarranton, in R.E. Spier, J.B. Griffiths, and B. Meignier eds., *Production of Biologicals from Animal Cells in Culture*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1991, pp. 110–115.
8. K. Kopp, W. Noé, M. Schlüter, F. Walz, and R. Werner, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 661–666.
9. D. Beebe and G. Murano, in A.S. Lubiniccki and S.A. Vargo eds., *Regulatory Practice for Biopharmaceutical Production*, Wiley-Liss, New York, 1994, pp. 281–293.
10. G. Zettlmeissl, H.S. Conradt, M. Nimtz, and H.E. Karges, *J. Biol. Chem.* **264**, 21153– 21159 (1989).
11. G. Mignot, T. Faure, V. Ganne, B. Arbeille, A. Pavirani, and J.L. Romet-Lemonne, *Cytotechnology* **4**, 163–171 (1990).
12. M. Nagao, S. Masuda, S. Abe, M. Ueda, and R. Sasaki, in S. Kaminogawa, A. Ametani, and S. Hachimura eds., *Animal Cell Technology: Basic and Applied Aspects*, Proc. of the Fifth Int. Mtg. of the Japanese Assoc. for Animal Cell Technology, Kluwer, Dordrecht, The Netherlands, 1993, pp. 71–77.
13. M.S. Sinacore, T.S. Charlebois, S. Harrison, S. Brennan, T. Richards, M. Hamilton, S. Scott, S. Brodeur, P. Oakes, M.

Leonard, M. Switzer, A. Anagnostopoulos, B. Foster, A. Harris, M. Jankowski, M. Bond, S. Martin, and S.R. Adamson, *Biotechnol. Bioeng.* **52**, 518–528 (1996).

14. K. Kitano, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 35–42.
15. D.E. Matthews, K.E. Piparo, V.H. Burkett, and C.R. Pray, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 315–319.
16. M. Reiter, G. Blüml, T. Gaida, N. Zach, C. Schmatz, N. Borth, O. Hohenwarter, and H. Katinger, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1992, pp. 421–423.
17. R. Boraston, C. Marshall, P. Norman, G. Renner, and J. Warner, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1992, pp. 424–426.
18. H.N. Brand, S.J. Froud, H.K. Metcalfe, A.O. Onadipe, A. Shaw, and A.J. Westlake, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 55–59.
19. M. Zang, H. Trautmann, C. Gandor, F. Messi, F. Asselbergs, C. Leist, A. Fiechter, and J. Reiser, *Biotechnology* **13**, 389–392 (1995).
20. C. Dyring, H.A. Hansen, and C. Emborg, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 155–157.
21. E.T. Papoutsakis, D.I.H. Linzer, and M.C. Borys, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, 1994, U.K., pp. 658–660.
22. J.L. Moreira, P.M. Alves, J.G. Aunins, and M.J.T. Carrondo, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1992, pp. 411–413.
23. M. Lucki-Lange and R. Wagner, in R.E. Spier, J.B. Griffiths, and B. Meignier eds., *Production of Biologicals from Animal Cells in Culture*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1991, pp. 180–185.
24. P.J. Radlett, *Adv. Biochem. Eng. Biotechnol.* **34**, 129–145 (1987).
25. T.W.F. Pay, A. Boge, F.J.R.R. Menard, and P.J. Radlett, *Develop. Biol. Standard.* **60**, 171–174 (1985).
26. S.-Y. Li, B. Röder, and M. Wirth, in R.E. Spier, J.B. Griffiths, and B. Meignier eds., *Production of Biologicals from Animal Cells in Culture*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1991, pp. 445–450.
27. Y. Shirai, K. Hashimoto, H. Kawahara, R. Sasaki, K. Hitomi, and H. Chiba, *Cytotechnology* **2**, 141–145 (1989).
28. R. Kratje and R. Wagner, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1992, pp. 488–493.
29. D.W. Lee, D. Gregory, D.J. Haddow, J.M. Piret, and D.G. Kilburn, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1992, pp. 480–486.
30. G. Köhler and C. Milstein, *Nature* **256**, 495–497 (1975).
31. E.-J. Schlaeger, T. Klier, and K. Christensen, R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 130–133.
32. H. Pinton, A. Lourenco da Silva, J.L. Goergen, A. Marc, J.M. Engasser, J.N. Rabaud, and G. Pierry, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 470–475.
33. J.R. Birch, R.C. Boraston, H. Metcalfe, M.E. Brown, C.R. Bebbington, and R.P. Field, *Cytotechnology* **15**, 11–16 (1994).
34. S.A. Clark, *Animal Cell Biotechnol.* **4**, 318–341 (1990).
35. K. James, *Animal Cell Biotechnol.* **4**, 206–255 (1990).
36. J.W. Larrick and R. Balint, in S. Kaminogawa, A. Ametani, and S. Hachimura eds., *Animal Cell Technology: Basic and Applied Aspects*, Proc. of the Fifth Int. Mtg. of the Japanese Association for Animal Cell Technology, Kluwer, Dordrecht, The Netherlands, 1993, pp. 625–639.
37. O.-W. Merten, *Animal Cell Biotechnol.* **4**, 258–315 (1990).
38. R. Waterstrat, R.W. Glaser, H. Tanzmann, C. Riese, and U. Marx, in E.C. Beuvery, J.B. Griffiths, and W.P. Zeijlemaker eds., *Animal Cell Technology: Developments towards the 21st century*, Proc. of the Veldhoven Mtg., Kluwer, Dordrecht, The Netherlands, 1995, pp. 647–651.
39. M. Reiter, A. Buchacher, G. Blüml, N. Zach, W. Steinfellner, C. Schmatz, T. Gaida, A. Assadian, and H. Katinger, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 333–335.
40. C.A. Mitchell and P.P. Gray, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1992, pp. 583–585.
41. C.W. Buser, R. Beaudet, N. Soohoo, and G.G. Pugh, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 121–123.
42. A. Traunecker, F. Oliveri, and K. Karjalainen, *TibTech* **9**, 109–113 (1991).
43. V.A. Luckow and M.D. Summers, Biotechnology **6**, 47–55 (1988).
44. C. Cavegn, H.D. Blasey, M.A. Payton, B. Allet, J. Li, and A.R. Bernard, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1992, pp. 569–576.
45. E.-J. Schlaeger, H. Loetscher, and R. Gentz, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: De-*

velopments, Processes and Products, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1992, pp. 562–567.

46. K.C. O'Connor, T.L. Prewett, T.J. Goodwin, K.M. Francis, A.D. Andrews, and G. F. Spaulding, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 293–295.
47. G.R. Pettman and C.J. Mannix, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 264–266.
48. G. Schmid, N. Wild, M. Fountoulakis, H. Gallati, R. Gentz, L. Ozmen, and G. Garotta, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 625–631.
49. K. Baker, Y.-Z. Zheng, S. Reid, and P.F. Greenfield, in S. Kaminogawa, A. Ametani, and S. Hachimura eds., *Animal Cell Technology: Basic and Applied Aspects*, Proc. of the Fifth Int. Mtg. of the Japanese Assoc. for Animal Cell Technology, Kluwer, Dordrecht, The Netherlands, 1993, pp. 521–528.
50. A.R. Bernard, T.A. Kost, L. Overton, C. Cavegn, J. Young, M. Bertrand, Z. Yahia-Cherif, C. Chabert, and A. Mills, *Cytotechnology* **15**, 139–144 (1994).
51. T. Sugiura and E. Amann, *Biotechnol. Bioeng.* **51**, 494–499 (1996).
52. R. Kurkela, A. Herrala, P. Henttu, H. Nal, and P. Vihko, *Biotechnology* **13**, 1230–1234 (1995).
53. A.R. Fooks, J.R. Stephenson, A. Warnes, B. Dowsett, B.K. Rima, and G.W. Wilkinson, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 605–609.
54. K. Mizumachi, J. Kurisaki, and N.M. Tsuji, in S. Kaminogawa, A. Ametani, and S. Hachimura eds., *Animal Cell Technology: Basic and Applied Aspects*, Proc. of the Fifth Int. Mtg. of the Japanese Assoc. for Animal Cell Technology, Kluwer, Dordrecht, The Netherlands, 1993, pp. 115–121.
55. O.W. Ogonah, R.B. Freedman, N. Jenkins, and B.C. Rooney, in E.C. Beuvery, J.B. Griffiths, and W.P. Zeijlemaker eds., *Animal Cell Technology: Developments towards the 21st century*, Proc. of the Veldhoven Mtg., Kluwer, Dordrecht, The Netherlands, 1995, pp. 427–430.
56. A. Lazar, S. Reuveny, C. Kronman, B. Velan, and A. Shafferman, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 324–328.
57. A. Garnier, J. Côté, I. Nadeau, A. Kamen, and B. Massie, *Cytotechnology* **15**, 145–155 (1994).
58. N. Epstein, S. Reuveny, J. Friedman, and N. Ariel, *Animal Cell Biotechnol.* **4**, 445–473 (1990).
59. K. Nagamine, M. Shiraishi, Z. Kong, K. Shinohara, and H. Murakami, in R.E. Spier, J.B. Griffiths, and B. Meignier eds., *Production of Biologicals from Animal Cells in Culture*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1991, pp. 167–169.
60. W. Noe, R. Bux, W. Berthold, and W. Werz, *Cytotechnology* **15**, 169–176 (1994).
61. S.C. Musgrave, Y. Douglas, G. Layton, J. Merrett, M.F. Scott, and C.A. Caulcott, S. Kaminogawa, A. Ametani, and S. Hachimura eds., *Animal Cell Technology: Basic and Applied Aspects*, Proc. of the Fifth Int. Mtg. of the Japanese Assoc. for Animal Cell Technology, Kluwer, Dordrecht, The Netherlands, 1993, pp. 223–230.
62. S.C. Musgrave, N.J. Bangerter, C.M. Bentley, C.D. Brown, Y.E. Douglas, N.J. Furminger, S.T. McGowan, and M.F. Scott, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1992, pp. 89–91.
63. S.C. Musgrave, C.M. Bentley, C.D. Brown, G.G. Feeney, N.J. Furminger, and T.W. Steward, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1992, pp. 92–94.
64. S. Hosoi, M. Satoh, H. Miyaji, T. Nishi, T. Mizukami, M. Hasegawa, S. Itoh, and T. Tamaoki, in S. Kaminogawa, A. Ametani, and S. Hachimura eds., *Animal Cell Technology: Basic and Applied Aspects*, Proc. of the Fifth Int. Mtg. of the Japanese Assoc. for Animal Cell Technology, Kluwer, Dordrecht, The Netherlands, 1993, pp. 63–70.
65. B. Baijot, M. Duchene, and J. Stephenne, *Develop. Biol. Standard.* **66**, 523–530 (1987).
66. C. Isch, A. Bernard, J.Y. Bonnefoy, C. Cavegn, P. Graber, T. Jamotte, and H.D. Blasey, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1992, pp. 530–533.
67. T. Marique, D. Malarme, P. Stragier, and J. Wérenne, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 369–371.
68. B. Montagnon, B. Fanget, R. Vinas, L. Peyron, J.-C. Vincent-Falquet, and P. Caudrelier, in R.E. Spier, J.B. Griffiths, and B. Meignier eds., *Production of Biologicals from Animal Cells in Culture*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, U.K., 1991, pp. 695–704.
69. J.R. Haynes, S.X. Cao, B. Rovinski, C. Sia, O. James, G. A. Dekaban, and M. H. Klein, *AIDS Res. Hum. Retroviruses* **7**, 17–27 (1991).
70. J. Litwin, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, European Society for Animal Cell Technology, Butterworth-Heinemann, Oxford, 1992, pp. 414–417.
71. H. Miyaji, N. Harada, T. Mizukami, S. Sato, N. Fujiyoshi, and S. Itoh, *Cytotechnology* **4**, 173–180 (1990).
72. A.L. van Wezel, *Nature* **216**, 64–65 (1967).
73. S. Reuveny, in A.S. Lubiniecki ed., *Large Scale Mammalian Cell Culture Technology*, Dekker, New York, 1990, pp. 271–341.
74. C.A.M. van der Velden-de Groot, in E.C. Beuvery, J.B. Griffiths, and W.P. Zeijlemaker eds., *Animal Cell Technology: Developments towards the 21st century*, Proc. of the Veldhoven Mtg., Kluwer, Dordrecht, The Netherlands, 1995, pp. 899–905.
75. B. Griffiths, *Animal Cell Biotechnol.* **4**, 149–166 (1990).
76. P.M. Alves, J.L. Moreira, J.M. Rodrigues, J.G. Aunins, and M.J.T. Carrondo, *Biotechnol. Bioeng.* **52**, 429–432 (1996).
77. F. Cahn, *Tib Tech* **8**, 131–136 (1990).

78. M. Schläfke, W. Döcke, and U. Marx, in E.C. Beuvery, J.B. Griffiths, and W.P. Zeijlemaker eds., *Animal Cell Technology: Developments towards the 21st century*, Proc. of the Veldhoven Mtg., Kluwer, Dordrecht, The Netherlands, 1995, pp. 489–493.

79. N. Jenkins and E.M.A. Curling, *Enzyme Microb. Technol.* **16**, 354–364 (1994).

80. N. Jenkins, R.B. Parekh, and D.C. James, *Nat. Biotechnol.* **14**, 975–981 (1996).

81. H.S. Conradt, M. Nimtz, K.E.J. Dittmar, W. Lindenmaier, J. Hoppe, and H. Hauser, *J. Biol. Chem.* **264**, 17368–17373 (1989).

82. D.C. James, R.B. Freedman, M. Hoare, O.W. Ogonah, B.C. Rooney, O.A. Larionov, V.N. Dobrovolsky, O.V. Lagutin, and N. Jenkins, *Biotechnology* **13**, 592–596 (1995).

83. K. Teruya, S. Shirahata, T. Yano, J. Watanabe, K. Seki, K. Osada, H. Tachibana, H. Ohashi, and H. Murakami, in E.C. Beuvery, J.B. Griffiths, and W.P. Zeijlemaker eds., *Animal Cell Technology: Developments towards the 21st century*, Proc. of the Veldhoven Mtg., Kluwer, Dordrecht, The Netherlands, 1995, pp. 91–95.

84. T.J. Ahern, G.E. Morris, K.M. Barone, P.G. Horgan, G.A. Timony, L.B. Angus, K.S. Henson, J.B. Stoudemire, P.R. Langer-Safer, and G.R. Larsen, *J. Biol. Chem.* **265**, 5540–5545 (1990).

85. R.P. Lanza, J.L. Hayes, W.L. Chick, *Nat. Biotechnol.* 14, 1107–1111, (1996).

86. A. Bader, K. Bork, G. Schumann, and K.F. Sewing, in E.C. Beuvery, J.B. Griffiths, and W.P. Zeijlemaker eds., *Animal Cell Technology: Developments towards the 21st century*, Proc. of the Veldhoven Mtg., Kluwer, Dordrecht, The Netherlands, 1995, pp. 939–943.

87. A. Lazar, F.J. Wu, H.J. Mann, R.P. Remmel, F.B. Cerra, and W.-S. Hu, in E.C. Beuvery, J.B. Griffiths, and W.P. Zeijlemaker eds., *Animal Cell Technology: Developments towards the 21st century*, Proc. of the Veldhoven Mtg. Kluwer, Dordrecht, The Netherlands, 1995, pp. 1017–1023.

ANTIBIOTICS. See SECONDARY METABOLITES, ANTIBIOTICS.

ANTIBODY PURIFICATION

MICHIEL E. ULTEE
DOUGLAS W. REA
Cytogen Corp.
Princeton, New Jersey

KEY WORDS

Antibody
Chromatography
Isolation
Monoclonal antibody
Polyclonal antibody
Purification
Separations
Ultrafiltration

OUTLINE

INTRODUCTION

Scope

This review focuses on the fullest extent of purification: that needed for an injectable or parenteral product for human use. This application requires the use of multistep procedures, with steps of high resolution (1,2). Many uses of antibodies do not require such stringent purity. Our objective is to present methods that when combined are capable of achieving parts-per-million levels of contaminants, with the understanding that such criteria are not needed for many uses of antibodies. In such cases, single steps of many of the methods presented will prove to be adequate.

Our focus is also primarily on monoclonal antibodies, which now dominate the antibody field, but we include a special section on polyclonal antibodies. Similarly, we describe purification processes for the IgG and IgM classes only, because these account for almost all uses of antibodies. Occasional reference is made to purification of the IgA subclass, whose use is rare.

Our discussion is limited to methods of purification. The choice of equipment needed to carry out these methods is a very broad topic. It depends on whether process development or production is being performed. It also depends on the scale, on whether the method is good manufacturing process (GMP) or not, and on the nature of the antibody being purified. Thus, details of this selection process are beyond the scope of this article.

Properties of Antibodies Compared to Typical Feedstock Contaminants

Typical feedstock contains serum proteins such as albumin and transferrin as well as cell-degradation products such as DNA and cytosol proteins. In addition, if the antibody

is in ascites fluid, the host animal's (typically mouse) antibodies would be present as well. By a combination of chromatographic or precipitation methods, one seeks to purify the desired antibody away from these contaminants. More recently, cell-culture media have become available that are described as "protein-free" and usually contain only low molecular weight digests of proteins or amino acids. If the antibody-secreting cells can grow and produce well in such media, the purification process becomes much easier, because the IgG molecules at a molecular weight (MW) of 150 kDa are readily separated by size from the components of such media.

As described in other articles, IgG antibody molecules are multichain glycoproteins that have been divided into various subclasses based on structure and properties. Their carbohydrate content is only 2% and is invariably located in the CH_2 domain of the Fc portion. The carbohydrate is sequestered, in that the carbohydrate chains are directed inward and are the actual contact residues between the heavy chains in this region (3), with the two heavy chains bowed outward to accommodate the carbohydrate. Thus, lectins generally do not bind to IgG, and purification must be directed toward the polypeptide portion. It should be noted that there are occasional reports of a carbohydrate site in the variable region of the IgG (3), in which case a purification method could be directed toward the carbohydrate moiety.

IgM molecules, on the other hand, contain about 9 to 12% carbohydrate, attached at five locations throughout the constant domains (4,5). Therefore, carbohydrate-directed methods have been developed for IgM and will be discussed later.

Specific affinity interactions for IgG and to some extent for IgM have been exploited in their purification. These affinity methods fall into three categories. First are those that rely on the key functional activity of the antibody, that is, its binding to antigen. Thus, a classic method of purification of antibody is binding on a solid-phase absorbent to which antigen has been immobilized. The second category relies on certain proteins that specifically bind to IgG. In this category are the well-known bacterial proteins A and G as well as second antibodies raised to bind against the desired antibody. Third, less-specific but often considered affinity chromatographic techniques include those relying on the interaction of the antibody with immobilized metals (IMAC), dye-ligands, hydroxyapatite, and thiophilic adsorbents.

The isoelectric points vary widely among different antibodies, mostly between about pH 6–9, but within different subclasses the range is generally much tighter. The isoelectric point directly affects the ability to use ion exchange, hydroxyapatite, chromatofocusing, and preparative electrofocusing techniques to purify antibodies. For example, the isoelectric point of serum albumin at 4.5 to 5.0 (6,7) indicates that its removal by ion exchange should be feasible for most antibody-purification schemes. Transferrin, at pI = 5.5–6.0 (8), can be more challenging.

The hydrophobicity and solubility of antibodies are important in their purification by such methods as hydrophobic interaction chromatography and precipitation methods. In general, antibodies are more hydrophobic and less soluble in the presence of precipitating salts than feedstock contaminants.

Finally, the size of antibodies, 150 kDa for IgG and 900 kDa for IgM, makes this a key property to exploit in their purification. Size-exclusion chromatography is widely used, as are ultrafiltration methods. Most contaminants are of lower molecular weight than antibodies. The nucleic acids DNA and RNA, which vary widely in size, can be digested using nucleases (9–11) to convert them to small oligonucleotides that are readily removed by sizing methods.

PURIFICATION METHODS

Introduction

The choice of a purification method is based on a number of different factors: the antibody, the feedstock, scale, economics, timing, and desired purity. As mentioned above, this article will concentrate on the most demanding type of antibody purification, that of multigram lots of monoclonal antibodies produced under GMP and for use in parenteral formulations.

The purification process can be divided into four parts. Figure 1 presents a flow chart of these steps, along with the most common techniques used for each step. The purpose of the first part, *sample preparation*, is to ensure the readiness of the sample for the first purification step. Sample preparation is sometimes not necessary, depending on the nature of the crude feedstock and on the choice of the first purification step.

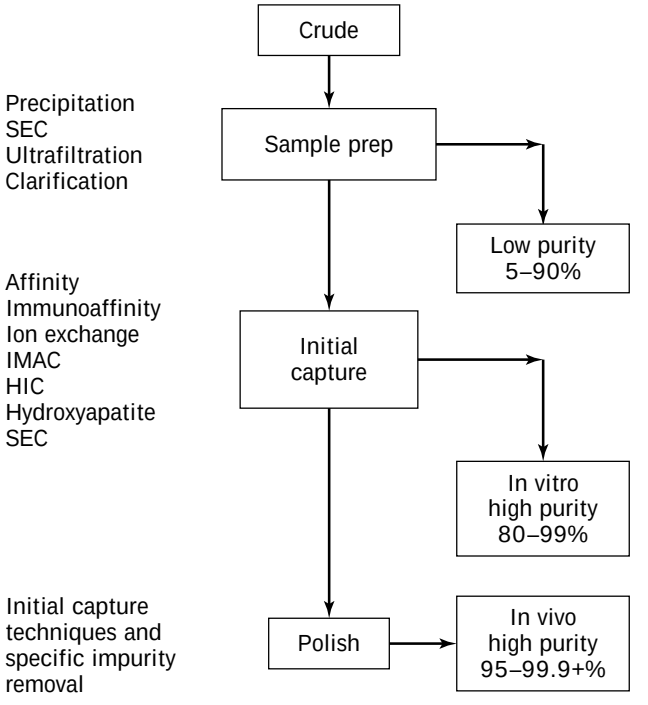


Figure 1. A general flow path for antibody purification, showing (on left) techniques typically used at each step.

Sample preparation is followed by the first purification step, often known as *initial product capture*. Most typically, the methodology for this step is chosen to give a high purification factor. Indeed, for antibody products with less stringent purification requirements (e.g., for in vitro use), this may be the only purification step required, if properly chosen.

The third part of the procedure is any *additional purification step(s)* necessary to reach the purity goal. These steps are sometimes referred to as polishing steps.

The final part of the procedure is to put the product into its *final formulation*, that is, the desired final concentration and physical state. This may sometimes be accomplished by the final purification step itself, eliminating the need for separate additional steps.

Sample Preparation

The purpose of sample preparation is to modify the chemical and physical state of the crude product into a state compatible with the initial capture step. From a chemical standpoint, this could involve changing the ionic strength, pH, or type of ion in the product solution. In the simplest cases, this may involve dilution or the addition of salt as a solid or concentrated solution. In a current GMP (cGMP) manufacturing setting, such methods are not particularly desirable. Addition of volume means increased costs. Solids that may be added will generally not be sterile, thus requiring an additional sterile filtration of the modified sample.

A more typical process at manufacturing scale would be a buffer exchange by size-exclusion chromatography (SEC) or by ultrafiltration or diafiltration. Because of its time- and volume-intensive nature, the related technique—dialysis—is typically used only on a laboratory scale.

From a physical standpoint, there are two typical processes carried out. One is concentration by ultrafiltration. The second is removal of particulate matter, by microfiltration or centrifugation. A very dirty feedstock (e.g., a batch fermentation) would clog a filter rapidly, making centrifugation the choice. However, for a relatively clean feedstock, the simplicity of filtration makes it the economic choice.

On a small scale, precipitative techniques such as ammonium sulfate, capyrilic acid, and polyethylene glycol (PEG) are also in common use. These techniques actually have at least modest purification power and generally have a concentrating effect as well (12–15). These techniques are less common at large scale, especially sterile operations, due to engineering considerations. Also, recovery from these operations can be low (15–17).

In addition to the above very common sample preparations, there are a number of other possibilities. It is sometimes desirable to remove certain contaminants early in the process, even though their concentration is low. Their removal is needed because their presence could affect subsequent purifications step(s). Some examples of these contaminants are DNA and RNA, endotoxins, and cell culture media dyes such as phenol red, which binds tightly to certain chromatography resins and could diminish their capacity. Bound dyes may also change the selectivity of a chromatographic resin (18,19). Many of these are removed when antibody is prepared by the precipitative methods noted earlier. Other techniques are possible, but these are generally used as secondary purification steps and will be described as such later. Another example of a troublesome contaminant is lipid, which can clog columns, affecting their performance and lifetime. Bulk lipids, especially as found in small volumes of ascites, may be partially removed by high-speed centrifugation. As before, the precipitative methods for antibody preparation will generally leave lipids in solution, thus effecting their removal. In addition, there are a number of extractive and chromatographic techniques that can remove lipids (16,20–21).

Initial Capture

The first major purification step in a process typically involves binding the antibody to some chromatography matrix, while impurities either flow through or are differentially eluted from the antibody. Alternately, conditions can be chosen where contaminants bind, and the antibody does not.

General Affinity. The most common matrix used to bind antibodies is immobilized protein A. This *Staphylococcus aureus* cell surface protein binds to many, although not all, antibodies and is widely available (17,22–31). Another bacterial cell surface protein, known as protein G, has been introduced more recently. Protein G binds more types of antibodies than does protein A, and it binds them more tightly (28,29,32,33). Although more versatile than protein A due to its wider specificity, the tighter binding of protein G can cause problems, as will be described later. Both of these are widely used because of their broad specificity, wide availability, and the high degree of purification. Indeed, purities of more than 95% are readily obtained, and antibodies of this purity are usable in many applications.

Historically, there have been three problems related to the use of protein A. First, it is expensive. However, given proper care, it is possible to reuse protein A resins many times, decreasing the long-term cost (34–37). Second, early protein A resins tended to leach protein A into the product (35). This is no longer a common problem, although leaching should be tested for, especially for products intended for in vivo use (38). Protein A assays are now readily available for such testing (39). The third problem is that many antibodies require a fairly low (pH 3–4) pH to elute from protein A. There are some antibodies that cannot withstand exposure to these conditions, even if the condition is neutralized as soon as possible. Needless to say, this is an even greater problem with the tighter binding of protein G (15,17,40).

In those instances where an antibody is acid sensitive, but it is still desirable to use protein A or G, a number of potential alternative elution conditions have been reported. These include the use of basic pH conditions (41), chaotropic ions or basic pH followed by ethylene glycol (42), and peptides that mimic the protein A or G binding site (43–45).

In addition to purifying bulk antibodies, proteins A and G have other applications in the field. For instance, they

can be used to remove residual whole antibody and Fc fragments in preparations of Fab and $F(ab')_2$ fragments (17), since both proteins A and G bind to the Fc region of antibodies. Another use is subclass separation from a polyclonal antibody feedstock (e.g., serum). Although this is not always the case, often different subclasses of antibodies have different affinities for proteins A and G, which allows chromatographic fractionation of the subclasses (15,17,22–23,40,46–48). Table 1 illustrates this effect. The differential binding of protein A for different subclasses is not only a blessing, but also a curse and a need for caution. One subclass that does not generally bind well to protein A under physiological conditions is murine IgG1. This is currently the most common monoclonal antibody type seen. However, this problem has been overcome, in that a number of different regimens, in particular the use of slightly alkaline binding buffers of high ionic strength, have been developed to efficiently purify murine IgG1 on protein A (49–52). The need for caution lies in that not even antibodies of the same subclass will behave exactly the same (Table 1; 40,46). Thus, although general guidelines for the binding and elution of different classes of antibodies are helpful, the protein A or G purification of each antibody must be individually optimized.

Similar natural affinity ligands exist with other specificities. Some of these are based on the fact that antibodies are glycoproteins. Examples are mannan binding protein, which targets the mannan sugars of IgM, and the lectin jacalin, which binds human IgA1 (26,49). Many others have been described in the literature (26).

Immunoaffinity. The most selective technique is that of immunoaffinity, which can take a number of forms. In the most specific, the antigen for the desired antibody, or an anti-antibody against it, is used to make a chromatographic resin. In a somewhat less-specific fashion, antibodies against specific types of light or heavy chains can be used. However, immunoaffinity suffers from a number of problems. First, the antigen or anti-antibody must be produced and purified. Second, binding is often so strong that very vigorous and potentially harmful elution conditions must be used (12,53–56). On a production scale, the economics of producing the antigen or anti-antibody often make this technique undesirable. Finally, for those antibodies intended for in vivo use, assay techniques must be available to show that the product is free of antigen or anti-antibody. Furthermore, unless process steps downstream of the immunoaffinity column provide adequate viral clearance, as shown by validation studies, the ligand used to make the immunoaffinity support will need to have been purified from a process validated to remove viruses (1). Also, it must be demonstrated that either the product is free of any leached ligand (as noted above) or the ligand does not affect product safety (1). These additional requirements would add significantly to the cost of immunoaffinity chromatography.

A renaissance of this technique may come from the advent of chemically and genetically engineered pseudoantigens. Developed for specificity and reasonable binding affinity and produced by recombinant or chemical methods, these compounds (often small peptides) could overcome most, if not all, of the earlier objections (57). A similar case is the development of mimics of protein A and of generic anti-antibodies (58,59). Generic anti-antibodies are those targeted against constant regions of the antibody, making

Table 1. Percentage of Given Antibody Eluted at Given pH

Clonality	Subclass	pH 8.5	pH 8.0	pH 6.0	pH 5.8	pH 5.5	pH 5.0	pH 4.5	pH 4.0	pH 3.5	pH 3.0	Ref.
Poly	G1			>90			2					23
Mono	G1	24.3			68							47
Mono	G1	73.2			3.1			1				47
Mono	G1			66				33.3				46
Mono	G2	1			44			25		0.6		47
Poly	G2a						90		10			23
Mono	G2a							100				46
Mono	G2a							56		44		46
Mono	G2a			45				55				46
Mono	G2a			61				39				46
Mono	G2a			38				62				46
Mono	G2a			60				40				46
Mono	G2a			7.5				92.5				46
Mono	G2a			29				71				46
Mono	G2a			12				88				46
Poly	G2b										>90	23
Mono	G2b			20.5				71.5				46
Mono	G2b			6				28		66		46
Mono	G2b			23				77				46
Mono	G2b							85.2		14.8		46
Mono	G3					6		103			5	48
Mono	G3			26				74				46
Poly	M		>80									23
Poly	A		>90									23

Note: mono, monoclonal; poly, polyclonal.

their usage more general. A number of these are available commercially.

Ion Exchange. Another very widely used technique is ion exchange chromatography (IEC). Although less specific than proteins A and G, it can often provide purities of more than 80% in one step, which is adequate for many purposes (18,60–64). Both cationic and anionic IEC can be used, in either binding or flowthrough modes. The choice of method and its utility depend on the relative ionic characteristics of the antibody and its impurities. Care must be taken in selecting the resin and conditions to be used, not only from the standpoint of resolution, but also of capacity. If conditions are chosen where both the antibody and impurities bind, the capacity of the resin for antibody could be dramatically reduced.

The pI of a protein is not an exact predictor of a protein's behavior on IEC. The charge distribution on the surface of the protein and a number of other minor effects also play a role. Nonetheless, the pI is a good first approximation. In one study of 15 murine monoclonal antibodies, pIs of 4.9 to 8.3 were reported (65). Bovine albumin has been reported to have a pI of about 4.5 to 5.0 (4,5). The pI of another common impurity, transferrin, ranges from 5.5 to 6.0, depending on the species and degree of iron saturation (66). Thus, it is quite possible to have an antibody that is acidic enough to be inseparable from transferrin and even albumin by IEC. One of the first steps in developing an antibody purification procedure is to determine the pI of the antibody and the contaminants. This may be done quickly and easily and will give a good indication of how useful IEC might be.

Immobilized-Metal Affinity Chromatography. In addition to the widely used techniques described in the previous section, a number of others may be used for the initial purification of antibody. One technique that is growing in use is immobilized metal affinity chromatography (IMAC). Metal ions, usually Ni^{2+}, are held to a chromatography resin by ligands. In turn, these metal ions can interact with proteins, primarily with histidine groups. Many engineered antibodies and fragments are being produced with a polyhistidine tail, which gives this technique a high specificity. IMAC can sometimes also be used for nonengineered antibodies that have naturally occurring surface histidines, although histidine is an uncommon amino acid (67–69).

Hydrophobic Interaction Chromatography. Hydrophobic interaction chromatography (HIC) is another technique that can be used for initial capture. Antibodies tend to be among the most hydrophobic of the proteins in the crude feedstock (70,71). Thus, this technique can be quite powerful (6,8,72,73). An example is given in Figure 2 (72), where bovine serum proteins are separated from antibody. On the down side, experience shows that the loading buffer must be carefully chosen (67). As with IEC, conditions where most of the contaminating proteins bind along with the antibody should be avoided for the initial capture. In such a situation, capacity is decreased and selectivity diminished. Also, selection of an appropriate media can be somewhat more tedious than for IEC, because the choice of potential binding strengths is very broad. Not only is there supplier-to-supplier variation in ligand density and surface chemistry, variabilities also seen in IEC, but there are a large number of binding groups, such as ethyl, propyl, butyl, phenyl, ether, phenylether, and others. Even reversed-phase media, such as octyl, can be used in the aqueous HIC mode with weakly hydrophobic proteins.

Because of the variation in manufacturing, very little can be said in a general fashion about the strengths of the various resins. Within an individual manufacturer's line,

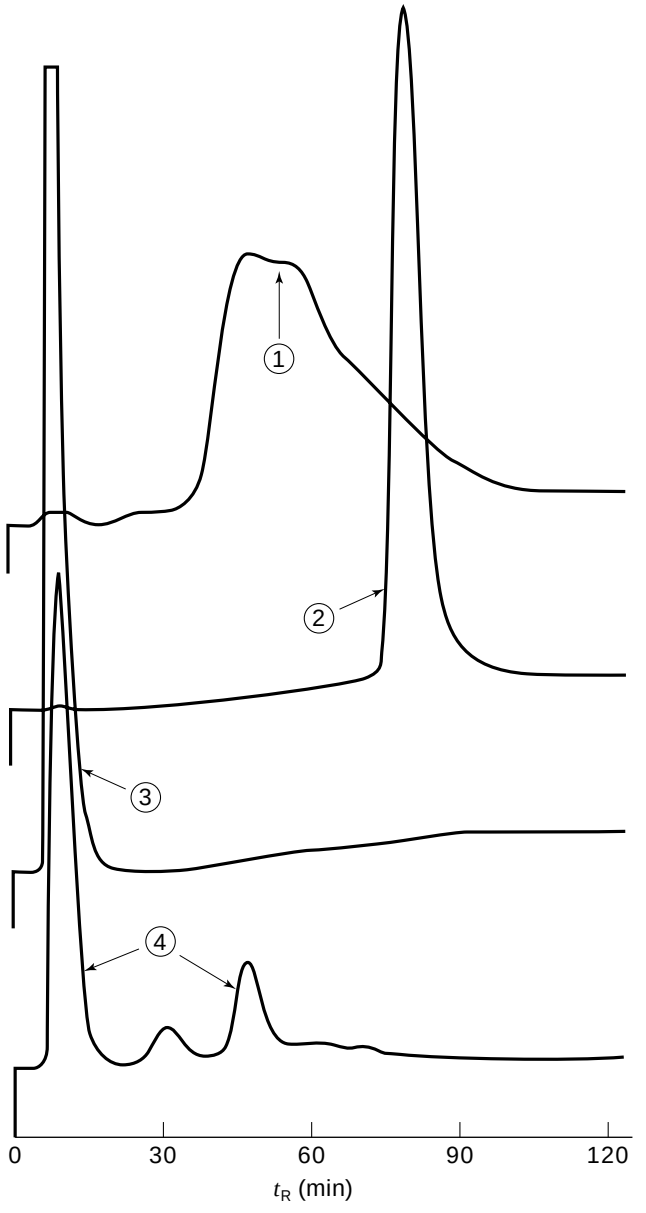


Figure 2. HIC separation of mouse polyclonal IgG from fetal bovine serum proteins and bovine polyclonal IgG. Chromatogram 1, bovine polyclonal IgG; 2, mouse polyclonal IgG; 3, partially purified (on ABx™, J.T. Baker) fetal bovine serum; 4, neat fetal bovine serum. *Source*: Reprinted from Ref. 72 with kind permission of Eaton Publishing.

usually the longer the alkyl chain length, the stronger the binding. In practice, however, tests should be made. It is important to make these tests, because the selection of a resin is an important way to manipulate selectivity.

Another means of altering selectivity is through the choice of buffer. In general, the stronger the lyotropic effect, the stronger the induction of hydrophobic binding by a salt (Fig. 3) (71). An example of this is seen in Figure 4 (74). However, strong salt solutions can have a detrimental effect on the integrity of proteins, so the choice should be limited, if possible, to those with strong salting out effects. Ammonium sulfate is the most commonly used salt, but it has the disadvantage of outgassing ammonia above pH 7.5 (71,75). Sodium sulfate and potassium phosphate are also good choices, but are somewhat limited in solubility (71,75). Because of the limited choice of buffers, the proper choice of resin becomes important.

In addition to salt, the selection of pH can have an effect on retention. In general, above pH 9 and below pH 4, retention increases, probably a result of denaturation of the protein. Between these pHs some proteins show an effect, but others do not. It is not possible to make generalizations; rather, the effect must be tested (71,75). Solution modifiers such as sugars, alcohols, PEG, and urea can also affect both binding and elution and can play a role in special situations (71,75).

HIC is perhaps deserving of more attention than it has received. Not only is it potentially a powerful technique itself, but for multistep procedures, it can be an ideal complement to IEC. IEC uses a low-ionic-strength load and a high-ionic-strength elution. HIC used a high-ionic-strength load and a low-ionic-strength elution. Thus, it is quite possible that conditions can be developed to allow one column to flow directly to the next, without intermediate sample handling.

Hydroxyapatite. Hydroxyapatite chromatography (HAP) has seen considerable use on a small basis, but little use on the large scale. The main reason for this is that until very recently, all versions of this material were mechanically fragile. Low flow rates were typical, and reuse of the columns was limited (53,76,77).

The mechanism of separation is not fully understood, but interaction of the protein with the calcium and phosphate in the resin matrix plays an important role, because a phosphate buffer gradient is most typically used for elution (53,76,77). As with other nonspecific techniques such as HIC and IEC, antibodies exhibit a wide variety of behavior on HAP, and generalizations are somewhat unsafe. However, HAP has been used not only for purifying antibodies (Fig. 5) (78), but also for separating different antibody subclasses (Fig. 6) (77). It is well known for its ability

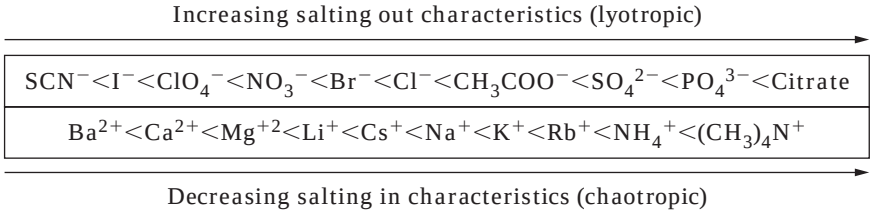


Figure 3. Lyotropic/chaotropic characteristics of salts.

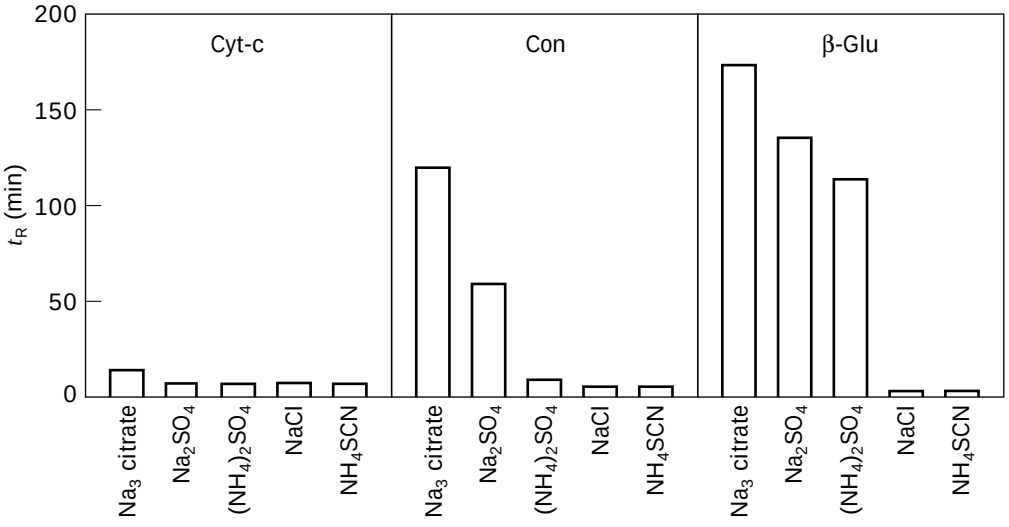


Figure 4. The effect of salt type on protein retention on a butyrate HIC column. The proteins were separated in a 20-min linear gradient from 1.0 M salt in 10 mM potassium phosphate buffer (pH 7.0) to 10 mM potassium phosphate buffer (pH 7.0). Cyt-c, cytochrome c; Con, conalbumin; *β*-Glu, *β*-glucosidase; t_R, retention time. *Source*: Reprinted from Ref. 74 with kind permission of Elsevier Science-NL, Sara Burgerharstraat 25, 1055 KV Amsterdam, The Netherlands.

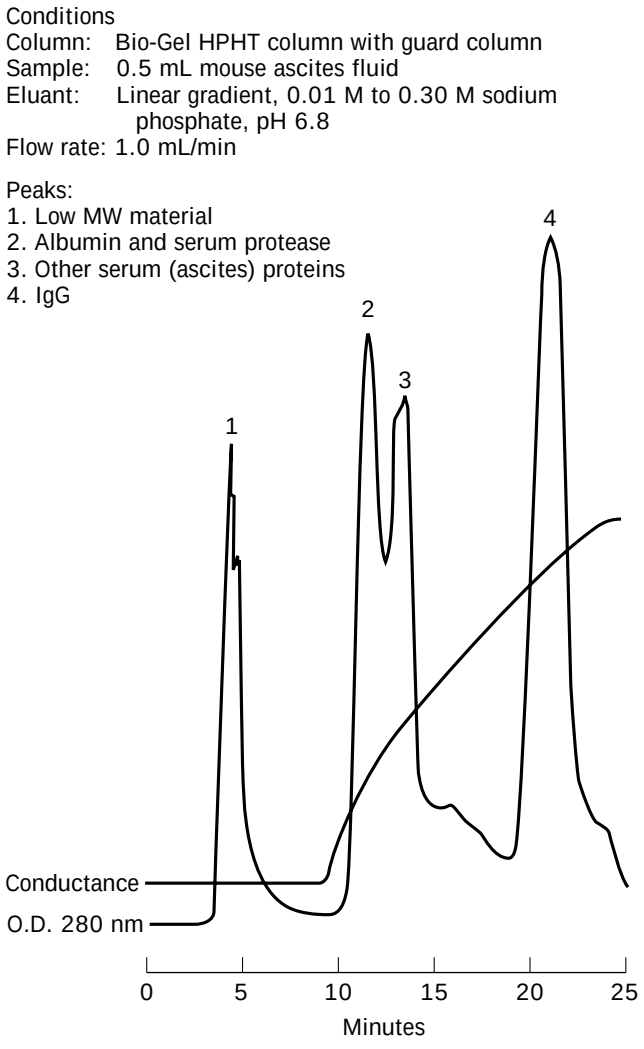


Figure 5. HPHT chromatography of unpurified ascites fluid containing monoclonal antibody. *Source*: Reprinted from reference 78 with kind permission of Eaton Publishing.

to bind DNA and to separate idiotypes, which will be discussed in "Secondary Purification Steps." With the recent development by several companies of more robust versions of this resin, its use for large-scale production of antibodies is bound to expand.

Size-Exclusion Chromatography. SEC is typically a low-resolution, low-throughput technique. Thus, in general, it is more often used for buffer exchange or polishing steps. However, there are instances when it can be used for initial purification. This is the case when dealing with very large antibodies such as IgM and IgA. IgM is typically a pentamer, with a weight nearing 1 million Da. IgA is usually found as a dimer, with a weight of about 300,000 Da. Thus, these are much larger than most contaminating proteins and can be purified by SEC.

Other Techniques. Another technique that has seen some use is thiophilic chromatography. In this technique, the sulfur ligand of the resin is thought to interact with the aromatic groups of the proteins. It has been shown that this technique can be used to effectively purify proteins (22,79). Like HIC, this technique generally loads proteins in high-salt buffers. However, unlike HIC, high salt concentrations also promote elution. Needless to say, different salts have different effects (79).

In addition to chromatographic techniques, there are a number of other techniques that can substantially purify antibodies. The precipitative techniques were mentioned earlier in "Sample Preparation." Ultrafiltration and diafiltration may also be used. Certainly, low molecular weight impurities may be removed by this technique. However, in addition, by careful selection of the membrane and buffer used, some contaminant proteins may be removed as well, such as albumin from IgG (80). The technique of preparative electrophoresis has gradually been advancing during the past few years. Currently it is can be considered useful for small-scale (milligram) bench preparations, especially non-GMP. Techniques for true production-scale electrophoresis that can be performed aseptically are being developed (81). Given the resolving power of analytical electrophoresis, success at developing production-scale techniques could add a valuable new tool to the development of purification methods.

In purification of antibodies on a large scale, process time is very important for economic reasons. One way to increase the efficiency of a process is to combine steps. In the past few years, a number of approaches to combining the concentration and clarification aspects of the sample prep step and the initial capture step have been commercialized. These approaches are generally either a fluidized bed, large-diameter bead resins, or wide-channel tangential flow filters. They are all designed so that cells and cell debris will flow through without clogging the setup, thus achieving clarification. The antibody is to bind to the resin or membrane, and after clarification is complete, it can then be eluted, providing both concentration and purification. Although simple in theory, a number of technical challenges had to be overcome before these approaches became practical.

Secondary Purification Steps

Depending on the feedstock, the technique selected, and the optimization, the initial capture step can often give purities in the range of 80 to 95%. This is often adequate for in vitro usages. However, for in vivo use, purities in excess of 99% are usually required. In addition to protein contaminants, other impurities such as DNA, endotoxins, viruses, and aggregates need to be removed. In such cases, a multistep procedure is almost inevitable.

All the same techniques used for initial capture can be used for secondary purification. Indeed, many of the techniques, such as HIC and HAP, are used more often as polishing steps than as initial capture steps.

The orientation of these secondary steps tends to be different from the initial capture step. With initial capture, the goal is to remove as many and as much contamination as possible. The secondary purification steps tend to be planned to remove individual contaminants. For instance,

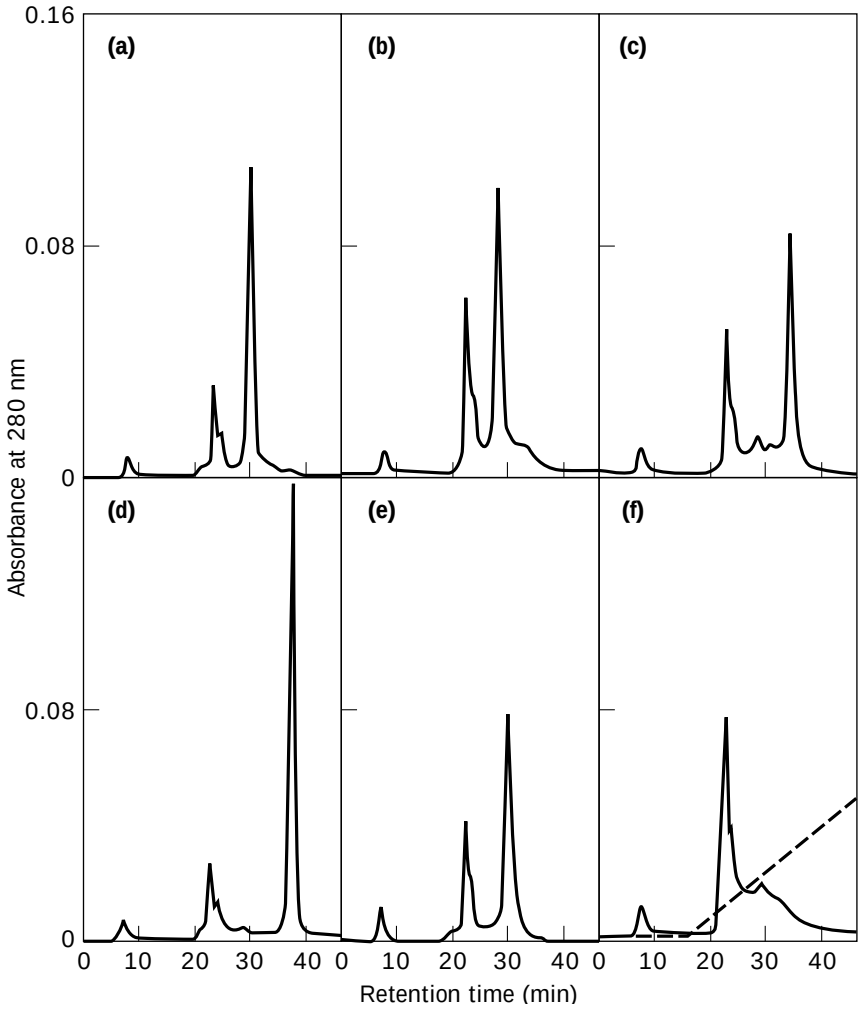


Figure 6. HPLC of mouse monoclonal antibodies in ascites on hydroxyapatite gel. (**a**), IgG1; (**b**), IgG2a; (**c**), IgG2b; (**d**), IgG3; (**e**), IgM; (**f**), antibody-free ascites. The *dashed line* on **f** indicates the gradient profile of phosphate ion concentration. *Source*: Reprinted from Ref. 77, p. 670, by courtesy of Marcel Dekker, Inc.

DNA and RNA tend to be highly negatively charged at just about any physiological pH. Thus, anion exchange resins are often used to remove these nucleic acids from antibodies, with conditions usually selected so that the nucleic acids bind while the antibody flows through.

In addition to the techniques mentioned in the initial capture step, there are a number of other techniques aimed at removing specific contaminants. For example, materials are available commercially that are intended specifically for the removal of impurities such as endotoxins, viruses, DNA and RNA, and albumin. In one way or another, these products take advantage of specific attributes of the impurity to remove them.

Some general examples of secondary purifications are discussed in this section. These include both techniques covered in the initial capture section and special polishing techniques.

The removal or inactivation of viruses must be validated for any animal-sourced or animal cell–sourced products, including antibodies, intended for in vivo use (1–2,82,83). All the techniques mentioned to date have the potential for removing virus. A specific attribute that is applicable to all viruses is their size. Being significantly larger than antibodies, they can be removed by a size basis. Although this can be done by SEC, a higher degree of virus removal is typically seen by ultrafiltration, with a membrane specifically selected for a molecular weight cutoff appropriate to retain virus and pass antibody (82,84,85).

In addition to virus clearance, virus inactivation is often performed. The most common techniques used are detergent treatment, low pH, or high temperature (82). Needless to say, the antibody must be tested to be resistant to these conditions before their routine use.

As mentioned earlier, one characteristic of DNA and RNA is the high negative charge. As well as anion IEC, HAP is well known to be effective at binding nucleic acids because of this high charge (76,86,87). HIC also may remove nucleic acids, because nucleic acids do not bind, whereas antibodies do bind to HIC resins (72).

Other properties of nucleic acids are the long, thin nature of the polymer, and, of course, the nucleotide sequence itself. Various precipitation techniques take advantage of this to remove DNA and RNA. Also, complementary nucleotides bound to chromatography supports can be used as an affinity adsorbent. Finally, as alluded to in the "Introduction," DNA and RNA can be enzymatically digested using nuclease enzymes (e.g., Benzonase®) to generate

very small fragments that can be readily removed from antibodies by sizing techniques (9,10).

Endotoxins are lipopolysaccharides. Many can be removed by anion exchange chromatography (6,53,88). As with the nucleic acids, HIC has been reported to be able to remove endotoxins for antibodies (72). In addition, both precipitative and chromatography techniques focusing in on the lipid portion of the endotoxins have been developed.

Albumin is probably the most typical and highest concentration contaminant in most feedstocks, whether serum, ascites, or cell culture media. As with the other contaminants, the techniques already mentioned have the potential for removing albumin. However, if additional specific removal is required, antialbumin affinity chromatography is possible. Also, many serum proteins, including albumin, can be removed using dye-ligand chromatography. A well-known technique in laboratory chromatography, it has not been much used in preparative-scale work. The dyes used were originally taken from the textile industry and were not particularly pure. In addition, there was concern about ligand leakage. However, modern dye-ligand resins are made with high-purity dyes prepared especially for this purpose; coupling procedures have improved, minimizing leakage; and assays are often available to track what leakage there may be (89–93).

Perhaps the greatest challenge for the polishing steps is the removal of form variants. The easiest form variant to remove is aggregate, which can be removed by SEC, a step that may be needed for final formulation (see next section) anyway. In addition, there may be light or heavy chain variants if a cell-line producing antibody is not stable and specific. There is the possibility of host antibody from serum, ascites, or serum-supplemented cell culture. And there can be intramolecule variants caused by variations in glycosylation, deamidation, proteolytic digestion, and a number of other possible causes. It is difficult to make a blanket statement in regard to how or even if such variants need be removed. If these variants prove to be extremely difficult to remove, and the product can be proved to be safe and effective with their presence, the economics of the situation may dictate that they not be removed.

Nonetheless, variants may need to be removed, and some examples of possible techniques follow. For chain variants and other source antibodies, if the physical characteristics are significantly different compared to the target antibody, they probably can be removed by IEC, HIC, etc. But, if not, one possibility is HAP. This technique has long been known to sometimes be able to separate different idiotypes (76,78) (Fig. 7). Dye-ligand chromatography has also been shown to be capable of separating subclasses (93). In addition, the different affinity of different antibodies for protein A and protein G allows these resins to be used for variant removal at times. Immunoaffinity directed against the undesirable variant heavy or light chain(s) is a powerful technique. Glycosylation variants may sometimes be removed by lectin or immunoaffinity chromatography directed against specific sugars. If physical changes such as deamidation do not change the antibody's charge enough for separation by IEC, preparative electrophoresis or chromatofocusing may sometimes be successful (94). But, basically, this is a struggle that must be developed on a case-by-case basis and often requires a unique solution.

In general, the difference between initial capture and additional purification steps may be summarized in one word: resolution. In initial capture, the feedstock is a moderately to highly complex mixture. The feedstock may be dirty, that is, contain particulates, if sample preparation is not performed. On the production scale, large volumes are the norm. The goals in initial capture are to remove most of the impurities, with good yield, and to concentrate the product. With secondary purification, the feedstock is relatively clean, most of the protein is the target antibody, and some residual impurities may be physically very similar to the target antibody. Thus, the goal in secondary purification is to achieve high-resolution separation.

Final Formulation

Final formulation can be considered as a part of purification in that it removes conditions that would impair the stability or utility of the antibody in its intended use. One goal of a purification strategy should be to limit the additional processing required to put the antibody into the final formulation. For example, ideally the antibody should be recovered from the final purification step in the same buffer needed for its final formulation. Often this can be done if this goal is kept in mind when the purification process is being developed.

Final formulation may be as simple as a straightforward sterilization by membrane filtration through a sterile filter with pores 0.2 μm or less. Another relatively simple formulation step is adjusting the antibody concentration, either by dilution with buffer or by concentration on ultrafiltration. In other circumstances, the buffer composition may need to be changed to achieve optimal stability of the antibody. For this purpose, SEC or diafiltration is widely used. A more complex formulation step would be the addition of excipients to confer stability (95). Finally, the antibody solution may need to be lyophilized to confer stability, and the liquid formulation may be changed to be compatible with the lyophilization process. Such complex formulation strategies are beyond the scope of this article, so the reader is referred to reviews on this subject (95–97).

PURIFICATION OF POLYCLONAL AND PROBLEMATIC ANTIBODIES

Polyclonal Antibodies

Purification of polyclonal antibodies is actually more of a challenge than the purification of monoclonal antibodies. This is the case for two reasons.

The first is related to the method of production. Polyclonal antibodies are produced naturally, by the immune response of an immunized animal. The antibody is purified from serum, with the necessity of removing all the serum impurities, including host antibodies. Monoclonal antibodies, however, are produced in incubators. Historically, for small batches, the incubator has been a mouse, with the resulting ascites fluid containing the desired antibody. This ascites fluid is related to serum and can be an equally

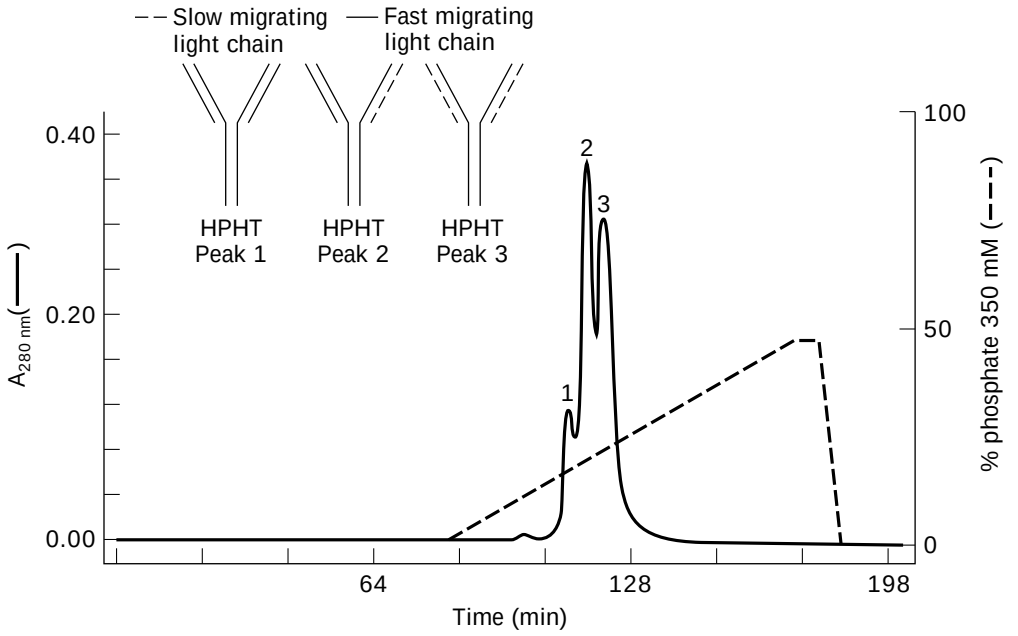


Figure 7. Preparative HPHT chromatography of IgG. *Source*: Reprinted from Ref. 76 with kind permission of Eaton Publishing.

complicated mixture. For large lots, and increasingly for small lots, monoclonal antibodies are produced in an artificial reactor, using at least a semidefined culture media. Again, historically, culture media could be considered only semidefined because it was supplemented with serum. However, even in this case, the concentration and number of macromolecular contaminants to be removed are lower than those in serum. Also, the current trend for monoclonal antibody production is toward use of serum-free or even protein-free media.

The second reason that polyclonal antibodies may be more challenging to purify is intrinsic to the antibody itself. Monoclonal antibodies are not necessarily absolutely homogenous. Often, there are differences in glycosylation. And there will be aging differences, such as in degree of deamidation. They will, however, all be of the same isotype, directed against the same epitope. This is not the case for polyclonal antibodies. Polyclonal antibodies tend to have a broader pI range than monoclonal antibodies, making IEC and related techniques such as HAP more complicated (77,86,94). They also tend to give broader peaks on hydrophobic interaction (72).

The same techniques used for purification of monoclonal antibodies are used for the purification of polyclonal antibodies. Indeed, most of the techniques were developed before the advent of monoclonal antibodies. As with monoclonal antibodies, the key to selecting the right purification techniques lies in knowing the desired yield, the desired purity, and the nature of the contaminants. Success at producing high-purity polyclonal antibodies centers around two points: extensive preparation and the use of affinity techniques.

In regard to preparation, characterizing all the individual impurities in a crude polyclonal mixture would be outlandishly time consuming. However, a valuable approach is to run standard IEC, HIC, and HAP conditions, where most if not all the proteins initially bind. Analysis of the elution fractions by an analytical technique such as two-dimensional electrophoresis will indicate which column technique can remove which impurities. The limitation to this approach is the well-known fact that the binding conditions used in chromatography affect the resolution of elution.

The affinity and pseudoaffinity techniques such as dye-ligand, IMAC, and thiophilic chromatography can be important in several roles. These techniques can be selected not only to bind to the target antibody, but also to remove specific impurities. For instance, one of the most vexing problems in antibody purification is removal of host antibody. If a specific isotype is desired, affinity techniques that bind to the other isotypes can be used to remove these unwanted host antibodies. Generic affinity techniques such as protein A and G are still very powerful tools. And, if practical to use, immunoaffinity has excellent selectivity.

In summary, there is nothing unique in purifying polyclonal antibodies compared to monoclonal antibodies. To reach a given degree of purity, the polyclonal antibody will probably take more planning; more preparation; and for high purity, use of more powerful techniques, especially affinity techniques.

Problematic Antibodies

As has been seen, antibodies can be remarkably diverse. One problem seen is when an antibody does not behave in a fashion similar to others of the same isotype. The solution of this type of problem is just good old-fashioned process optimization.

A much more serious problem is the purification of an antibody with limited solubility. A wide variety of additives can potentially be used for solubilization of antibodies, including urea, guanidine HCl, detergents, and sugars. Needless to say, there is the potential for these agents to interfere with purification steps as well. Thus, solution of this problem tends to require rather extensive development (48).

A similar problem is lack of stability. This may be intrinsic to the antibody, that is, it occurs even under physiological conditions, or it may be caused by a harsh purification condition, such as a low pH protein A elution or a virus-killing step. In the second case, generally alternate conditions are possible. Examples of alternates for both the examples are given earlier in this article. The solution of an intrinsic stability problem is similar to that of a solubility problem. These problems can be quite complicated and are beyond the scope of this article, but good references are available (98,99).

PURITY CONSIDERATIONS FOR USE OF ANTIBODIES AS PHARMACEUTICALS

Although purities of 95% or greater are often suitable for antibodies intended for nonpharmacological use, such as in vitro diagnostic reagents, much greater levels of purity are required for those intended for use as pharmaceuticals. In general, there are three types of contaminants of concern to regulatory agencies: microbes, viruses, and their products, such as endotoxins, proteins, and DNA (1,2). These contaminants must be removed to trace levels.

The need for sterility of a biological product intended for use in humans is a given. This is primarily assured by strict adherence to both aseptic technique and cGMP (100). Less apparent is the need for a validated process to remove any *potential* viruses that may be present in the crude material, including murine retroviruses, which are present in almost all hybridoma cell lines. These issues have been addressed in some detail by the regulatory authorities (1,2,83,101,102).

These regulatory authorities, such as the Center for Biologics Evaluation and Research division of the FDA, have relied on a three-tiered approach to controlling any infectious agents that may be present in the crude product or acquired during the downstream processing.

1. Careful selection of raw materials, including media components, for the absence of biological contaminants
2. Validating that the purification process can remove or inactivate viruses
3. Full testing of the product at appropriate stages in the process for the presence of viruses and infectious agents

Each of these three tiers of control must be adhered to by those engaged in the purification of antibodies for human use. A full treatment of these regulatory considerations is beyond the scope of this article. The reader is encouraged to refer to the cited references and keep abreast of current regulations issued by the various governmental agencies.

SUMMARY

The antibody molecule is one of the most versatile biological molecules created. Advances in immunology, biochemistry, and molecular biology have yielded a wealth of information about the structure and function of this class of proteins. The huge diversity of their binding sites allows one to acquire exquisitely specific binding reagents for almost any molecular target. Yet, the conserved nature of their constant domains as well as their characteristic three-dimensional structure allows the purification specialist to approach their recovery with much understanding of the behavior of these biomolecules. In spite of this, the adage that each monoclonal antibody is unique must be remembered, because even monoclonal antibodies from the same species, class, and subclass may behave very differently in purification procedures. With this caveat in mind, it is our hope that this review will provide a helpful guide to the purification of these fascinating biomolecules.

BIBLIOGRAPHY

1. Center for Biologics Evaluation and Research, *Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products for Human Use*, 3rd ed., U.S. Food & Drug Administration, 1997.
2. European Regulatory Guidance Document, *Production and Quality Control of Monoclonal Antibodies*, vol. 3, addendum 3, Committee on Proprietary Medical Products, 1995.
3. T.W. Rademacher, S.W. Homans, R.B. Parekh, and R.A. Dwek, *Biochemical Society Symposium* **51**, 131–148 (1986).
4. D.R. Anderson, P.H. Atkinson, and W.J. Grimes, *Arch. Biochem. Biophys.* **243**, 605–618 (1985).
5. A. Nisonoff, J.E. Hopper, and S.B. Spring eds., *The Antibody Molecule*, Academic Press, New York, 1975, pp. 87, 314.
6. L.S. Hanna, P. Pine, G. Reuzinsky, S. Nigam, and D.R. Omstead, *BioPharm* 33–37 (Oct. 1991).
7. P.G. Righetti and T. Caravaggio, *J. Chromatogr.* **127**, 1–28, (1976).
8. K. Morimoto and K. Inouye, *J. Biochem. Biophys. Methods* **24**, 107–117 (1992).
9. A.J. Hagen, R.A. Aboud, P.A. DePhillips, C.N. Oliver, C.J. Orella, and R.D. Sitrin, *Biotechnol. Appl. Biochem.* **23**, 209–215 (1996).
10. A-K.B. Frej, R. Hjorth, and Å. Hammarström, *Biotechnol. Bioeng.* **44**, 922–929 (1994)
11. G.B. Dove, G. Mitra, M.A. Shearer, and M.S. Cho, in M.R. Ladisch, R.C. Willson, and C.-D.C. Painton eds., *Protein Purification: From Molecular Mechanisms to Large-Scale Processes*, American Chemical Society, Symposium 427, Washington, D.C., 1989, pp 194–209.
12. E. Harlow and D. Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1988, pp. 283–318.
13. A. Polson, M.B. von Wechmar, and M.H. V. van Regenmortel, *Immunol. Commun.* **9**, 475–493 (1980).
14. S.H. Neoh, C. Gordon, A. Potter, and H. Zola, *J. Immunol. Methods* **91**, 231–235 (1986).

15. A.P. Phillips, K.L. Martin, and W.H. Horton, *J. Immunol. Methods* **74**, 385–393 (1984).
16. P.W. Thompson, A.C. Kenney, P. Moulding, and D. Wormald, *Ann. N.Y. Acad. Sci.* **529**, 529–539 (1990).
17. C.M. Fraser and J. Lindstrom in J. C. Venter and L.C. Harrison eds., *Receptor Biochemistry and Methodology*, Liss, New York, 1984, pp. 1–30.
18. M.J. Gemski, B.P. Doctor, M.K. Gentry, M.G. Pluskal, and M.P. Strickler, *Biotechniques* **3**, 378–384 (1985).
19. F. Chen, G.S. Naeve, and A.L. Epstein, *J. Chromatogr.* **444**, 153–164 (1988).
20. T.J. Smith, in D.J. Asai ed., *Methods in Cell Biology*, Vol. 37, Academic Press, San Diego, 1993, pp. 75–93.
21. U.S. Pat. 4,565,652 (Jan. 21, 1986), R. Schmidtberger (to Behringwerke Aktiengesellschaft).
22. L. Jacob, E. Schmitt, and W. Brümmer, *Am. Biotechnol. Lab.* 44–45 (Nov. 1994).
23. P.L. Ey, S.J. Prowse, and C.R. Jenkin, *Immunochemistry* **15**, 429–436 (1978).
24. J. Groudswaard, J.A. van der Donk, A. Noordzij, R.H. van Dam, and J.P. Vaerman, *Scand. J. Immunol.* **8**, 21–28 (1978).
25. J. Goding, *J. Immunol. Methods* **20**, 241–253 (1978).
26. M.D. P. Boyle, in M.D. P. Boyle ed., *Bacterial Immunoglobulin Binding Proteins*, Vol. 2, Academic Press, San Diego, 1990, pp. 1–21.
27. L. Schwartz, in M.D. P. Boyle ed., *Bacterial Immunoglobulin Binding Proteins*, Vol. 2, Academic Press, San Diego, 1990, pp. 309–340.
28. L. Björck and G. Kronvall, *J. Immunol.* **133**, 969–974 (1984).
29. A. Surolia, D. Pain, and M.I. Khan, *TIBS* 74–76 (Feb. 1982).
30. R. Lindmark, K. Thorén-Tolling, and J. Sjöquist, *J. Immunol. Methods* **62**, 1–13 (1983).
31. J.J. Langone, in F.J. Dixon ed., *Advances in Immunology*, Vol. 32, Academic Press, San Diego, 1982, pp. 157–252.
32. B.W. Walker, in M.D.P. Boyle ed., *Bacterial Immunoglobulin Binding Proteins*, Vol. 2, Academic Press, San Diego, 1990, pp. 255–368.
33. B. Äkerstrom and L. Bjorck, *J. Biol. Chem.* **261**, 10240–10247 (1966).
34. A.C. Kenney and H.A. Chase, *J. Chem. Technol.* **39**, 173–182 (1987)
35. P. Füglistaller, *J. Immunol. Methods* **124**, 171–177 (1989).
36. G. Hale, A. Drumm, P. Harrison, and J. Phillips, *J. Immunol. Methods* **171**, 15–21 (1994).
37. R. Francis, J. Bonnerjea, and C.R. Hill, in D.L. Pyle ed., *Separations for Biotechnology 2*, Elsevier, New York, 1990, pp. 491–498.
38. M.A. J. Godfrey, P. Kwasowski, R. Clift, and V. Marks, *J. Immunol. Methods* **160**, 97–105 (1993).
39. M.T. Dertzbaugh, M.C. Flickinger, and W.B. Lebherz III., *J. Immunol. Methods* **83**, 169–177 (1985).
40. C.L. Villemez, M.A. Russell, and P.L. Carlo, *Mol. Immunol.* **21**, 993–998 (1984).
41. Eur. Pat. Appl. 453,767 (Oct. 30, 1991), E.M. Croze (to E.R. Squibb and Sons, now Bristol-Myers Squibb).
42. N. Fornstedt, *FEBS Lett.* **177**, 195–199 (1984).
43. R. Bywater, G. Eriksson, and T. Ottosson, *J. Immunol. Methods* **64**, 1–6 (1983).
44. R. Bywater, *Chromatogr. Synth. Biol. Polym.* **2**, 337–340 (1978).
45. M. Yarnell and M.D.P. Boyle, *Biochem. Biophys. Res. Commun.* **135**, 1105–1111 (1986).
46. J.R. Stephenson, J.M. Lee, and P.D. Wilton-Smith, *Analyt. Biochem.* **142**, 189–195 (1984).
47. L. Manil, P. Motté, P. Pernas, F. Troalen, C. Bohuon, and D. Bellet, *J. Immunol Methods* **90**, 25–37 (1986).
48. W. Jiskoot, A-M.V. Hoven, A.A.M. DeKonig, M.F. Leerling, C.H.K. Reubsaet, D.J.A. Crommelin, and E.C. Beuvery, *J. Immunol. Methods* **138**, 181–189 (1991).
49. G.T. Hermanson, G.R. Mattson, and R.I. Krohn, *J. Chromatogr., A* **691**, 113–122 (1995).
50. G. Schuler and M. Reinacher, *J. Chromatogr.* **587**, 61–70 (1991).
51. A.P.G. van Sommeren, P.A.G.M. Machielsen, and T.C.J. Gribnau, *Prep. Biochem.* **22**, 135–149 (1992).
52. U.K. Pat. Appl. 2,160,530A (June 22, 1984), H. Juarez-Salinas and G.S. Ott (to Bio-Rad Laboratories).
53. D.R. Nau, *Biochromatography* **4**, 4–18 (1989).
54. Pharmacia, in Pharmacia ed., *Monoclonal Antibody Purification*, Pharmacia Biotech., Piscataway, N.J., 1993, pp. 21–25.
55. J.W. Eveleigh and D.E. Levy, *J. Solid Phase Biochem.* **2**, 45–78 (1977).
56. L.O. Narhi, D.J. Caughey, T. Horan, Y. Kita, D. Chang, and T. Arakawa, *Analyt. Biochem.* **253**, 236–245 (1997).
57. G.A. Baubach and D.J. Hammond, *Biopharm* 24–35 (May 1992).
58. N. Khatter, R.S. Matson, and T. Ngo, *Am. Lab.* 32LL– RR (May 1991).
59. J.Y. Shi and R.A. Goffe, *J. Chromatogr., A* **686**, 61–71 (1994).
60. J.R. Deschamps, J.E.K. Hildreth, D. Derr, and J.T. August, *Analyt. Biochem.* **147**, 451–454 (1985).
61. P. Clezardin, J.L. McGregor, M. Manach, H. Boukerche, and M. Dechanvanne, *J. Chromatogr.* **319**, 67–77 (1985).
62. B. Pavlu, U. Johansson, C. Nyhlén, and A. Wichman, *J. Chromatogr.* **359**, 449–460 (1986).
63. Y. Yang and K. Harrison, *J. Chromatogr., A* **743**, 171–180 (1996).
64. E. McCarthy, G. Vella, R. Mhatre, and Y. Lim, *J. Chromatogr., A* **743**, 163–170 (1996).
65. M. Carlsson, A. Hedin, M. Inganäs, B. Härfast, and F. Blomberg, *J. Immunol. Methods* **79**, 89–98 (1985).
66. E.N. Baker, *Perspect. Bioinorg. Chem.* **2**, 161–205 (1993).
67. V. Boden, J.J. Winzerling, M. Vijayalakshmi, and J. Porath, *J. Immunol. Methods* **181**, 225–232 (1995).
68. F.H. Arnold, *Biotechnology* **9**, 151–156 (1991).
69. J.E. Hale and D.E. Beidler, *Analyt. Biochem.* **222**, 29–33 (1994).
70. P. Gagnon, P.G. Cartier, J.J. Maikner, R. Eksteen, and M. Kraus, *LC-GC* **11**, 26–34 (1993).
71. P. Gagnon and E. Grund, *Biopharm*, 54–64 (May 1996).
72. D.R. Nau, *Biochromatography* **5**, 62–73 (1990).
73. O. Manzke, H. Tesch, V. Diehl, and H. Bohler, *J. Immunol. Methods* **208**, 65–73 (1997).
74. J.L. Fausnaugh, L.A. Kennedy, and F.E. Regnier, *J. Chromatogr.* **317**, 141–155 (1984).
75. T. Arakawa and L.O. Narhi, *Biotechnol. Appl. Biochem.* **13**, 151–172 (1991).
76. R. Vola, A. Lombardi, and M. Mariani, *Biotechniques* **14**, 650–655 (1993).

77. Y. Yamakawa and J. Chiba, *J. Liq. Chromatogr.* **11**, 665–681 (1988).
78. H. Juarez-Salina, S.C. Engelhorn, W.L. Bigbee, M.A. Lowry, and L.H. Stanker, *Biotechniques* 164–169 (May 1984).
79. T.W. Hutchens and J. Porath, *Analyt. Biochem.* **159**, 217–226 (1986).
80. S. Gadam, S. Orlando, S. Mundt, R. Kuriyel, S. Pearl, and M.E. Ultee, in S. Minden ed., *Monoclonal Antibody Purification*, International Business Communications, Southborough, Mass., 1996, pp. 1.3.1–1.3.4
81. C.W. Wrigley and H.P. Manusu, *Am. Lab.* **28**, 20K–20O (1996).
82. J.B. Grun, E.M. White, and A.F. Sito, *BioPharm* 22–30 (Nov.-Dec. 1992).
83. F. Brown and A.S. Lukiniecki eds., *Developments in Biological Standardization*, vol. 88, Karger, New York, 1996.
84. H. Maerz, S.O. Hahn, A. Maassen, H. Meisel, D. Roggenbuck, T. Sato, H. Tanzmann, F. Emmrich, and U. Marx, *Nat. Biotechnol.* **14**, 651–652 (1996).
85. L. Hoffer, H. Schwinn, L. Biesert, and D. Josic, *J. Chromatogr., B* **669**, 187–196 (1995).
86. T.L. Brooks and A. Stevens, *Am. Lab.* 54–64 (Oct. 1985).
87. R.J. von Wedel, in S.S. Seaver ed., *The Commercial Production of Monoclonal Antibodies*, Dekker, New York, 1987, pp. 159–173.
88. E.A. Neidhardt, M.A. Luther, and M.A. Recny, *J. Chromatogr.* **590**, 255–261 (1992).
89. C.R. Lowe, S.J. Burton, J.C. Pearson, and Y.D. Clonis, *J. Chromatogr.* **376**, 121–130 (1986).
90. C.R. Lowe, N. Burton, S. Dimaghanian, S. McLoughlin, J. Pearson, D. Stewart, and Y.D. Clonis, in M.A. Vijayalakshmi and O. Bertrand eds., *Protein-Dye Interactions: Developments and Applications*, Elsevier Applied Science, London, 1989, pp. 11–20.
91. G.E. McCreath, H.A. Chase, D.R. Purvis, and C.R. Lowe, *J. Chromatogr.* **629**, 201–213 (1993).
92. D.J. Stewart, D.R. Purvis, J.M. Pitts, and C.R. Lowe, *J. Chromatogr.* **623**, 1–14 (1992).
93. M. Hasnaoui, M. Debbia, S. Cochet, J.P. Cantron, P. Lambin, and O. Bertrand, *J. Chromatogr., A* **766**, 49–60 (1997).
94. J.C. Waldrep and J.R. Schulte, *J. Immunol. Methods* **118**, 273–277 (1988).
95. A. Wade and P.J. Weller eds., *Handbook of Pharmaceutical Excipients*, 2nd ed., American Pharmaceutical Assoc., Washington, D.C., and the Pharmaceutical Press, London, 1994.
96. J.L. Cleland and R. Langer, eds., *Formulation and Delivery of Proteins and Peptides*, American Chemical Society, Symposium 567, Washington, D.C., 1994.
97. A.P. MacKenzie, in D. Marshak and D. Liu eds., *Therapeutic Peptides and Proteins: Formulation, Delivery and Targeting*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989, pp. 17–21.
98. S. Li, C. Schöneich, and R.T. Borchardt, *Biotech. Bioeng.* **48**, 490–500 (1995).
99. T.J. Arern and M.C. Manning eds., *Stability of Protein Pharmaceuticals*, Plenum Press, New York, 1992.
100. Code of Federal Regulations 21:CFR 211, 600, 610.
101. Center for Biologics Evaluation and Research, *Q5 Viral Safety Evaluation of Biotechnology Products Derived from Cell Lines of Human or Animal Origin*, ICH Guidance Document, Washington, D.C., U.S. Food and Drug Administration, 1998.
102. F. Horaud and F. Brown eds., *Developments in Biological Standardization*, vol. 75, International Symposium on Virological Aspects of the Safety of Biological Products, Karger, New York, 1990.

See also ADSORPTION, PROTEINS WITH SYNTHETIC MATERIALS; AFFINITY FUSIONS, GENE EXPRESSION; CENTRIFUGES, ANIMAL CELLS; GOOD MANUFACTURING PRACTICE (GMP) AND GOOD INDUSTRIAL LARGE SCALE PRACTICE (GLSP); HYBRIDOMA, ANTIBODY PRODUCTION.

APOPTOSIS

RABINDER P. SINGH
MOHAMED AL-RUBEAI
University of Birmingham
Birmingham, England

KEY WORDS

Apoptosis
bcl-2
Bioreactor
Cell culture
Cell death
Flow cytometry
Necrosis

OUTLINE

INTRODUCTION

In the past it has generally been assumed that certain cell-culture conditions, such as interactions with specific hormones and deprivation of basic nutrients and growth factors, result in physical damage or metabolic collapse of the cell, leading to a passive, or "necrotic," death. It was thought that the cell was at the complete mercy of its environment and had no control over its fate. However, it is now clear that the cell does not always respond in such a simple and passive manner. Instead, a range of factors can trigger a highly complex and genetically regulated cellular response during which specific "death proteins" are activated, a phenomenon that has been named apoptosis (1). It is these proteins that are responsible, ultimately, for the death and destruction of the cell. Indeed, in many cases, the cell may have sustained only low levels of damage. Clearly, under these conditions apoptotic death will be premature and is often referred to as *cellular suicide*.

Apoptosis is now acknowledged as a fundamentally important process that plays an essential role in embryogenesis and in the maintenance and functionality of the highly ordered cell populations that constitute higher organisms. Indeed, there are now few aspects of biomedical research upon which apoptosis has not had an impact. It is as a consequence of its fundamental nature that any failure in its regulatory mechanisms leads to many of the diseases that pose the greatest challenges to medical science. This has resulted in an explosion of research into the genetic basis of apoptosis, the objective of which is to develop novel therapeutics for a wide range of disorders, including cancer, AIDS, and ischemic injury. Consequently a growing number of proteins have been identified that are involved in the induction suppression and execution of the apoptotic program. It is now clear that many of the cell lines used in the biotechnology industry for the production of therapeutics undergo apoptotic death. Obviously, the advances made in the characterization of the apoptotic pathway may be applied to the suppression of the apoptotic response in industrial culture processes. This should provide the biotechnologist with a new route to the optimization of culture performance.

We begin by describing the biochemical basis of apoptosis, the morphological changes that accompany it, and some of the techniques that have been used to identify apoptotic cells. This is followed by a general discussion of the regulation and induction of apoptosis. We then describe studies that have investigated this phenomenon from a biotechnological perspective. Particular reference is made to the conditions that elicit an apoptotic response, the cell lines that are susceptible, and the effect of *bcl-2* overexpression on cell survival and productivity in the bioreactor environment.

BACKGROUND

Morphology of Apoptosis

Apoptosis is defined by its highly characteristic morphology, which is illustrated in Figure 1. Early changes include a reduction in cell volume (2,3) and loss of surface microvilli. Cytoskeletal changes result in the formation of protrusions on the surface of the cell, which are referred to as blebbs. These may break away as intact vesicular structures, giving rise to "apoptotic bodies" (4).

One of the most striking changes during apoptosis occurs within the nucleus. Fluorescence microscopy of a typical viable cell following staining with a DNA stain such as acridine orange, reveals a large spherical nucleus that may constitute most of the cellular volume. Often, the chromatin is highly diffuse, although occasionally it may be possible to see condensed chromosomes in mitotic cells. During apoptosis, the nucleus undergoes major changes. Initially, the chromatin condenses and marginates to the nuclear membrane, forming crescent or ring-shaped structures that exhibit intense fluorescence. Eventually, the chromatin collapses into two or more particles, which are often highly spherical. Again, the remainder of the cell will be devoid of chromatin and will be almost transparent in appearance. The shape of the cell also undergoes a highly characteristic change. While viable cells tend to be highly irregular in shape, on entry into apoptosis they become smooth-surfaced and in many cases almost spherical.

The changes in nuclear morphology coincide with the activation of a nuclease enzyme that cleaves chromatin first into 300 and/or 50-kbp fragments (5), and then ultimately into multiples of 200 bp (6). The first stage is believed to be responsible for the morphological changes described above. The second stage represents the cleavage of chromatin at the internucleosomal linker regions. This generates a striking ladder like pattern when DNA samples from apoptotic cells are subjected to DNA gel electrophoresis. Together with condensation of chromatin, this so-called DNA ladder has become one of the hallmarks of apoptotic death.

Many cell types express a further enzymatic activity—that of the transglutaminase enzyme (7). This crosslinks proteins within the cell, generating a protein scaffold that is believed to hold the dead cell together, thus explaining the relative robustness of dead apoptotic cells as compared with cells that have undergone necrotic death.

These highly controlled changes appear to have a specific task: to provide a very clean and rapid method of eliminating dead cells, which is a vital consideration when one considers the extent of apoptotic cell death during, for example, embryogenesis. The cleavage of chromatin does not appear to be responsible, in itself, for the death of the cell. Instead, it has been argued that this ensures the complete destruction of the genetic material. It is suggested that this reduces the possibility of malignant transformation of surrounding cells by DNA from dying cells. The stabilization of the dead cell by the transglutaminase enzyme minimizes the probability of leakage of its contents onto surrounding cells. In vivo, surrounding cells and phagocytes engulf the dead cell before it can cause damage to surrounding tissue. Thus, inflammation of tissue as usually results from necrotic death is avoided. In vitro, in the absence of phagocytes, the apoptotic cell will eventually enter a degenerative phase called secondary necrosis.

Identification of Apoptosis

The search for simple techniques that allow for the identification of apoptotic cells has attracted a considerable

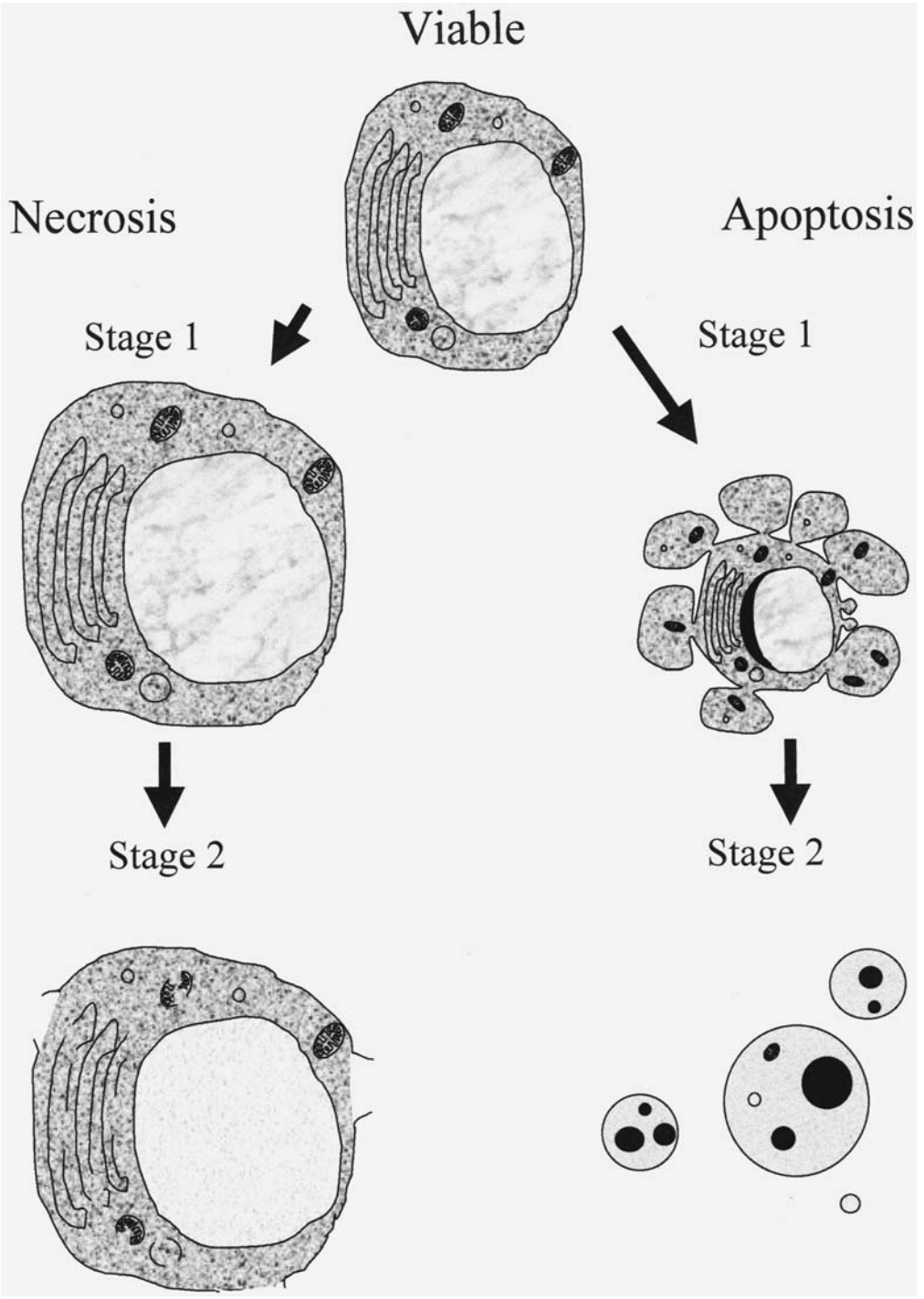


Figure 1. Schematic diagram showing the morphological features of apoptotic and necrotic cell death. The cell shown at stage 1 of necrosis exhibits clear swelling and dilation of mitochondria and Golgi. At this stage, nuclear morphology remains unchanged. By stage 2 of necrosis, disruption of organelles and plasma membrane is apparent, nuclear morphology is altered, and complete destruction of the cell quickly follows. During stage 1 of apoptosis, on the other hand, the cell exhibits extensive boiling or blebbing of the plasma membrane and a reduction in cellular volume. Condensation of the nuclear morphology may also be apparent. By stage 2, the chromatin has condensed and fragmented into several spherical particles and the cell is smooth surfaced in appearance. Apoptotic bodies containing chromatin and organelles are also shown. Eventually, plasma-membrane integrity is lost and degraded chromatin escapes from the dead cells, leaving nuclear-free ghosts. These structures may persist in the culture for many days.

amount of interest. As already stated, visualization of nuclease-mediated cleavage of DNA has been widely used to identify the presence of apoptotic cells. However, the technique can produce variable quality results that are of a qualitative, rather than a quantitative, nature.

Perhaps the simplest techniques for the study of apoptosis are based on the identification of the morphological features of cell death. For example, fluorescence microscopic analysis of nuclear morphology is a highly effective method for identification and quantification of apoptosis. There are two ways in which this may be done. If samples cannot be analyzed immediately, cells may be fixed in formaldehyde and stored at 4 °C. Analysis involves staining with acridine orange, which reveals the condensation of chromatin in apoptotic cells (7). However, this technique has one major drawback—an inexperienced operator may confuse early necrotic cells with viable cells, which have a very similar nuclear morphology. In order to avoid this difficulty, cells may be analyzed immediately while still in their culture medium by using the acridine orange–propidium iodide dual-staining technique (8). All cells are permeable to acridine orange, which stains chromatin green. Only membrane-damaged cells take up propidium iodide, and as a result exhibit red fluorescence. This technique therefore provides information regarding plasma membrane integrity, as well as nuclear morphology, and can consequently be used to simplify identification of necrotic cells. Additionally, the classification of apoptotic cells into early apoptotic and membrane-damaged apoptotic (sometimes referred to as secondary necrotic) cells gives an indication of the cell growth and death kinetics under the influence of a variety of environmental conditions during the cultivation process. The early-membrane-intact phase of death is relatively brief and, therefore, the presence of a large proportion of cells at this stage indicates that the rate of cell death is very high.

Microscopic techniques have two major drawbacks—they are subjective and time consuming. In theory, the development of flow cytometric methods should overcome these difficulties and provide a powerful tool for the study of the biochemical features of apoptosis in heterogenous cell populations. At present, there are a number of techniques that are available, and some of the most commonly used are discussed further.

Light-scattering Properties. The simplest flow cytometric method is based on the changes in light scattering properties that accompany cell death. Cells can be studied without the need for pretreatment, with the decrease in cell size producing a decrease in forward-scattered light. The increased granularity caused by nuclear condensation produces an increase in orthogonal light scatter (9,10). However, the latter is a transient stage, and eventually, a reduction in orthoganol light scatter is observed. The technique can also be used for the identification of necrotic cells (at least in their early stages). When necrosis is induced by a permeablizing agent such as saponin, a reduction in forward scatter is observed, but the increase in side scatter that occurs during apoptosis is not seen.

Changes in DNA Content. When stained with a DNA-specific stain such as propidium iodide (PI), apoptotic cells exhibit a characteristically low DNA content that appears as a sub-G_1 peak (i.e., it appears below the position of the G_1 peak of the cell cycle of viable cells). This is believed to result from the leakage of cleaved DNA from apoptotic cells (11–14). The technique provides a rather good correlation with the fluorescence microscopic technique already described. However, necrotic cells undergoing degradation may also exhibit a reduced DNA content, although the passive and asynchronous nature of this process means that a clear "peak" is not usually observed.

Annexin V. In viable cells, phosphatidylserine (PS) is located on the inner leaflet of the plasma membrane (15). During apoptosis, one of the earliest changes is the loss of this asymmetrical distribution (15,16). Annexin V has a very high affinity for PS. Conjugation of annexin V to a fluorescent tag, such as FITC, enables the use of this interaction to identify cells that have lost PS asymmetry. By combining this method with PI staining, it is also possible to classify apoptotic cells into two subpopulations: early (membrane intact and therefore PI negative) and late (membrane damaged and therefore PI-positive). The technique has been found to give a good correlation with levels of apoptosis during hybridoma batch cultures, during which apoptosis accounts for around 90% of cell deaths, as revealed by the fluorescence microscopic method described earlier. However, as with other flow cytometric techniques, necrotic cells can also give a false-positive result. This is because damage to the plasma membrane allows the annexin to enter the dead cell and bind to PS residues on the inner surface of the plasma membrane. Thus, the technique is reliable only when used to analyze early apoptotic cells, which can be seen as Annexin V positive and PI negative (16,17).

General Comments Regarding Identification of Apoptosis. Each of the techniques we described has advantages and drawbacks. When designing experiments to identify and quantify apoptosis, a number of points should be taken into consideration. First, there can be significant variations in the morphology of apoptosis from one cell type to the next, including, for example, absence of chromatin condensation. It is therefore recommended that several methods be used simultaneously to identify the mechanism of cell death. However, quantification of actual levels of apoptosis can vary significantly depending upon the technique used. Thus, in order to draw valid conclusions, it is essential that comparison of quantitative data collected using one technique is not made with data collected using a second technique.

Inducers of Apoptosis: Physiological Factors

The list of factors that induce apoptosis has grown steadily over the last few years. In the present section factors that have been the center of purely biological studies will be considered, although where necessary, implications for

animal cell technology will be highlighted. Factors that are specifically of interest to the process biotechnologist will be explored under "Apoptosis and Animal Cell Biotechnology."

Presence/Absence of Receptor–Ligand Interactions. The absence of certain hormones can result in the induction of apoptosis in certain cell types. This has led to the suggestion that apoptosis may play a pivotal role in the maintenance of tissue organisation in vivo. It is thought that all cells are primed to undergo apoptosis and are prevented from doing so through constant stimulation by parakrine survival factors (18). If a cell is removed from its physiologically correct location, the absence of the appropriate survival signal will lead to the induction of apoptosis. This role of survival factors as regulators of cellular distribution in vivo may also have an impact on the development of serum-free media preparations for industrial-scale cell-culture processes, as described later.

The presence of receptor–ligand interactions can also lead to the induction of apoptosis. For instance, in the Fas–FasL system, binding of the Fas ligand to the Fas receptor can trigger apoptosis (19). This mechanism is responsible for the regulation of the immune system. Autoreactive B cells undergoing maturation (20) and autoreactive mature T cells (21) are eliminated by the Fas-mediated induction of apoptosis. Fas also acts as the "off" switch for the immune system by inducing apoptosis in antigen-activated B and T cells (22). Molecular dissection of the Fas–FasL system has provided important insights into the early stages of the signaling cascade that leads to the expression of the death pathway (see "Genetics of Apoptosis").

Induction by Viruses. A number of viruses have been shown to interact with the cellular apoptotic machinery. An apoptotic response to viral infection would appear to act as a protective mechanism that prevents viral replication by triggering the suicide of the infected cell. However, a number of virally encoded antiapoptotic genes have now been identified that suppress the expression of the death program, thereby providing the virus with the opportunity to propagate itself. Perhaps one of the most interesting examples of this anti-death mechanism from a biotechnology perspective is that of baculovirus, which has been synthesized at production scale by infection of insect cell lines, and has applications as a biological pesticide and, more recently, for the expression of recombinant proteins. A mutant was identified that induced high levels of apoptosis during infection of Sf21 insect cells. This was attributed to a mutation in the p35 viral gene that, in its wild-type form, acts as an antiapoptosis gene (23).

Arguably, the most important example of virus-induced apoptosis is that mediated by HIV (24–27). A number of reports have indicated that the binding of the viral gp 120 protein to the $CD4^+$ receptor of T cells triggers the induction of apoptosis, thus leading to the depletion of this class of cells during HIV infection (28–30). Interestingly, HIV infection of $CD4^+$ cells appears to provide protection from apoptosis. The viral Nef protein downregulates the expression of $CD4^+$, thus preventing the induction of apoptosis. As a result, virus propagation in the infected cell is not prevented (31).

There have also been reports of virus-induced suicide of bacterial cells. Until recently, it was assumed that altruistic cell death was not possible in single-celled organisms, simply because the genes involved in such a phenomenon would be lost when the cell concerned dies. However, bacteriophage infection in bacterial colonies has been reported to induce a suicide response, thus preventing the spread of the infection to the remainder of the colony (32). The genes that mediate such a response are propagated by other clones in the colony, thus preserving the altruistic nature of cellular suicide. Bacteria and viral expression systems have been used for the production of recombinant proteins. Clearly, the possibility that bacterial cells in such systems exhibit an apoptosis-like response needs to be investigated.

Free Radicals and Apoptosis. Free-radical-mediated cellular damage has become an important area of study. In recent years there have been numerous reports that cell death following oxidative stress occur by apoptosis. Moreover, generation of free radicals has been postulated as being a universal triggering event in the induction of apoptosis (33–38). Indeed, for a period in the early 1990s, it was suggested that the widely studied antiapoptosis gene *bcl-2* functioned as an antioxidant (39). However, this theory has been challenged by the demonstration that anoxia-induced apoptosis can also be suppressed by Bcl-2 in the absence of measurable levels of free radicals (8,40,41).

Induction by Toxins and Therapeutic Agents. Exposure to high levels of toxic chemicals results in necrotic death of the cell. However, long-term exposure at a low level can trigger an apoptotic response. Many of the agents used in chemotherapy exert their affect by inducing apoptosis in tumor and normal cells. Overexpression of genes such as *bcl-2* has been linked to resistance to chemotherapy. Clearly, establishing the mechanism of induction of apoptosis in response to such agents and providing strategies that minimize the affect of antiapoptotic genes should provide novel and more effective chemotherapeutic strategies (42).

Exposure of cells to ionizing radiation induces high levels of apoptosis in many normal tissues. Particularly susceptible are those cells from tissues that undergo rapid proliferation, such as spermatogonia (43) and lymphocytes (44). Such tissue would be especially prone to malignant transformation, and consequently the induction of apoptosis following DNA damage minimizes the likelihood of such an event. Irradiation of tumors can also lead to the induction of apoptosis (45–47). Notably, tumors that respond least favorably to irradiation exhibit the lowest level of apoptosis under such conditions (45). The induction of apoptosis following DNA damage is regulated by the product of the p53 gene, which has been referred to as the gaurdian of the genome because of its central role in preventing the propagation of cells that may have sustained genetic damage and thus, potentially, malignant transformation (48).

Genetics of Apoptosis

The study of the genetic basis of apoptosis has become a highly complex and fast-moving area of research. Conse-

quently, this section will only deal with some of the most important aspects of this subject in order to provide a basis for the later discussion on genetic manipulation of the apoptotic pathway in commercially important cell lines.

The genes involved in apoptotic death may be classified into three groups. The first group consists of the modulators of the apoptotic pathway that suppress or induce death. The second group comprises the components of the cell death pathway that mediate the cell death signal. The final component is the group of effector enzymes that is responsible for the death and destruction of the cell. Some examples of these genes are given in Table 1.

Of the modulators of the apoptotic pathway, one group of closely related proteins, the Bcl-2 family, has attracted particular attention. The *bcl-2* gene, the first and best characterized member, was identified at the t(14;18)(q32⇗1) breakpoint found in human follicular lymphoma (54). It encodes a 24-kDa protein that is located on the outer mitochondrial membrane, the cytosolic face of the nuclear membrane, and endoplasmic reticulum. Numerous studies have demonstrated the ability of this protein to suppress apoptosis in response to a wide variety of inducers. Table 1 lists other members of this family, some of which are functionally similar to Bcl-2, whereas others act as antagonists of Bcl-2 and thereby trigger apoptosis.

Until recently, studies of the molecular basis of cancer were directed at the identification and characterization of genes involved in the regulation of cellular proliferation. However, it is now evident that the failure of cells to undergo apoptosis at the correct time and location is also an important step in malignant transformation. Indeed, studies of *bcl-2* in this context have been instrumental in establishing the role of apoptosis in cancer. Mutations that result in the overexpression of *bcl-2* lead to the accumulation of cells due to life span extension. These cells then undergo further mutation, most notably involving the *c-myc* gene, which leads to the formation of high-grade tumors that combine the characteristics of high cellular survival with a high rate of proliferation (61–63). Clearly, these are characteristics that would be desirable in the ideal candidate host for the expression of recombinant proteins.

One of the most common mutations identified in tumors is that of the p53 gene, mentioned earlier. In response to DNA damage, p53 causes cell cycle arrest, allowing the cell to repair damaged DNA, thus ensuring that any potentially carcinogenic mutation is not propagated. However, in some cell types, p53 triggers the induction of apoptosis, again in order to minimize the risk of transformation (48,64).

The well-characterized pattern of development of the nematode *Caenorhabtis elegans*, which includes the induction of apoptosis in specific cell types, has provided an important insight into the genetic basis of apoptosis. Most notable among the genes identified are *ced 9* and *ced 3*. The former is a *bcl-2* homologue that blocks the ability of *ced 3* to induce cell death (65). The *ced 3* gene is a cysteine protease that shares extensive homology with the mammalian protein called interleukin 1β converting enzyme (ICE) (66,67). As the name suggests, this enzyme cleaves the active precursor pro–interleukin 1β to generate the active molecule interleukin 1β, and a number of studies have now implicated it in the induction of apoptosis. However, apoptosis can also be induced in macrophages and thymocytes of ICE-negative mice, indicating that ICE is not a universal mediator of apoptosis (68). Indeed, several ICE-related enzymes have been identified in recent years (see Table 2).

The molecular mechanism of apoptosis induced by the interaction of the Fas-ligand with its receptor centers on the activity of ICE-related proteases. Indeed, important progress has now been made in deciphering the earliest events in the signal cascade that transduces the initial death stimulus to the death machinery of the cell. Activation of the Fas receptor by its ligand or agonist antibodies leads to the binding of the adapter protein MORT1/FADD (Fas-associating protein with death domain) (75,76). This, in turn, binds to the ICE homologue FLICE (FADD-like ICE) or MACH (MORT1-associated ced-3 homologue) (77,78). The targets of this enzyme are still under investigation.

APOPTOSIS AND ANIMAL CELL BIOTECHNOLOGY

At present, there are around 100 new biopharmaceuticals in phase I, II, and III clinical trials (79). Production of a biologically active therapeutic will, in many cases, necessitate expression of the protein in mammalian cell lines. As a result of intense commercial pressure, tried and tested production processes are often adopted in order to ensure that the product reaches the market in the shortest possible time. Initially, such production processes were relatively inefficient and there was tremendous scope for process optimization. Advances in cell biology, in addition

Table 1. Examples of Genes Involved in Apoptosis

Inducers/ Promoters	Ref.	Suppressors	Refs.	Mediators	Refs.
bax	49	*bcl-2*	54	ICE family	50
bak	50,51	*bcl-xL*	53	Trans-glutaminase	7
bad	52	*mcl-l*	55	NUC-18	58
bcl-xS	53	*ced-9*	56	DNase-I	59
FasL/Fas	19	*p35*	23	DNase-II	60
		bhfrI	57		

Table 2. Proteases Involved in the Induction of Apoptosis

Protease	Refs.
Ced-3	55
Interleukin 1β-converting enzyme (ICE)	69
Nedd2	70
CPP32	71
$ICE_{rel\ II}$/TX	72,73
$ICE_{rel\ III}$	73
mch2	74

to improvements in process monitoring, control, and optimization, and the development of novel bioreactor designs, are making the transition from the research laboratory to industry standard practice. Clearly, this will lead to a reduction in the cost and complexity of the process of recombination-protein production using mammalian cell lines by placing the technology on a firmer scientific footing.

In order to optimize the viable cell number and protein productivity, a detailed understanding of the factors that lead to cell death in the bioreactor is required. These factors may be classified into three groups:

1. The hydrodynamic environment of the cell
2. The accumulation of toxic metabolites
3. The exhaustion or local limitation of nutrients and oxygen

Although each of these areas has now been investigated to varying degrees, recent studies have demonstrated that at least some of the cell lines used in bioreactors undergo apoptotic death rather than necrosis, indicating that a reassessment of the subject is required. As described later, the greater understanding of the mechanism of cell death in commercial cultures should provide new routes to culture optimization. The resultant enhancement in culture efficiency would be expected to manifest itself in three forms:

1. First, the nutrients and culture time invested in generating a viable cell will be wasted if that cell should die prematurely. If the survival time of the cell can be enhanced, the proportion of culture resources utilized for production of the biopharmaceutical of interest can be increased by eliminating the need for regeneration of cellular biomass.
2. Second, as a cell dies, it releases proteolytic enzymes into the culture medium, which can lead to degradation of the product. Thus, product stability should be enhanced by minimization of cell death.
3. Finally, high levels of cellular debris in the culture medium can complicate the recovery of the target protein. This will add to the cost of downstream processing and lead to a reduction in the efficiency of target protein recovery.

A Brief History

The first suggestion that apoptosis may account for cell death during the cultivation of hybridoma cells came from an electron microscopic study conducted by Al-Rubeai et al. (73). Further studies by Franek and Dolnikova (74) demonstrated the accumulation of nucleosomal DNA fragments in culture medium during the death phase of batch hybridoma cultures. Based on this observation, they estimated that around 30% of cells had undergone apoptotic death. Further evidence of apoptosis during hybridoma cultures was provided by the studies of Mercille and Massie (80) and Singh et al. (81). Upon DNA gel electrophoresis, both studies revealed the laddering pattern that, as mentioned earlier, is characteristic of apoptosis. Morphological analysis of the nuclei of the cells indicated that apoptosis accounted for 90% of the dead cells. Both groups also found high levels of apoptosis during the cultivation of murine plasmacytoma cell lines (sometimes incorrectly referred to as myeloma cells). Furthermore, Singh et al. (82) reported an absence of apoptosis during the death phase of CHO and Sf-9 batch cultures. However, studies by Moore et al. (77) indicated significant levels of apoptosis during the death phase of serum-free batch cultures of CHO cells. Recent studies in our laboratory on a CHO 320 cell line overexpressing interferon indicate that this cell line may also be susceptible to apoptosis, although the morphology was not typical, and the frequency was much lower than that seen during hybridoma cultures under comparable conditions.

Studies are now required to give an indication of variability in susceptibility to apoptosis between clones of the same cell type, and between different unrelated cell lines. The objective will be to produce a correlation between susceptibility to apoptosis and general cell robustness. Clearly, such a correlation would provide a simple and easily identifiable predictor of robustness following exposure of cells to a range of stresses, thus simplifying the process of cell-line selection.

Induction and Suppression of Apoptosis in the Bioreactor Environment

The features of the bioreactor environment that may result in cell death were outlined in the introduction. In the present section, studies that have considered these factors in terms of their ability to induce apoptosis are described and the implications of suppression of apoptosis are discussed.

Death during Batch Cultivation. The nutrient limitations encountered by cells in the bioreactor may be classed into two groups: cycling or terminal limitation. The former may be encountered at all stages of large-scale or intensive culture systems. For example, in large-scale stirred tank reactors, the cells may be exposed to fluctuating nutrient levels because of inhomogeneity due to poor mixing. In intensive culture systems, the high cell densities reached will result in low local-nutrient concentrations. As a result, the level of cell death in such systems is relatively high. Terminal nutrient limitations will occur at the end of batch cultivation and, as a result, will become more extreme with time, invariably leading to cell death. In the case of hybridoma batch cultures, the first nutrient to become limiting is glutamine, and its exhaustion coincides with the onset of the death phase of the culture. As stated above, cell death under these conditions is almost exclusively by apoptosis. Two studies have reported on the effect of *bcl-2* overexpression on cell survival during the death phase of hybridoma cultures. Itoh et al. (78) found that bcl-2 overexpression significantly extended the duration of the culture by reducing the rate of cell death. Moreover, they reported a fourfold increase in the Mab (monoclonal antibody) titre in the culture medium. Simpson et al. (8) have also reported an extension in culture duration, although there was only a 40% improvement in Mab titre. Necrosis became the predominant mechanism of cell death in the *bcl-2*-transfected cell line, indicating near-complete suppression of apoptosis under batch culture conditions.

More recently, Suzuki et al. (79) have reported that transfection of COS-1 cells with *bcl-2* and then with the vector pcDNA-λ carrying the immunoglobulin λ gene for transient expression of λ protein have resulted in higher expression of the protein when compared to the control transfectant (i.e., *bcl-2* negative). In the same study, the mouse plasmacytoma p3-X63-Ag.8.653, which is used as a fusion partner in the generation of hybridomas, and the hybridoma cell line 2E3 were transfected with the human *bcl-2* gene. In both cases an extension of batch culture duration was reported. The *bcl-2*-transfected 2E3 cells survived 2 to 4 days longer in culture, producing a 1.5- to 4-fold larger amount of antibody in comparison with the control vector transfectants. A further enhancement in survival and antibody production in hybridoma 2E3 cultures was observed when cells were cotransfected with *bcl-2* and *bag-1*.

In contrast with these promising studies, Murray et al. (83) found that *bcl-2* transfection of the murine plasmacytoma NS0 failed to provide any protection from apoptosis. Although there was no endogenous Bcl-2 expression, they did report expression of Bcl-xL, a functional homologue of bcl-2. Thus, they suggested that Bcl-2 may be functionally redundant in this cell line.

Nutrient Limitation

Amino Acids and Glucose. The link between the onset of apoptosis and exhaustion of glutamine during batch cultures of hybridoma cells has prompted systematic studies of the role of the various nutrients used in culture medium. Initial studies indicated that deprivation of glucose, serum, glutamine, cysteine, and methionine (84–86) could all individually induce high levels of apoptosis. Moreover, recent studies in our laboratory (87) indicate that this is not a feature of these particular nutrients alone. Deprivation of each amino acid individually from the commonly used RPMI 1640 culture medium was found to result in the induction of apoptosis, with particularly high levels observed following deprivation of essential amino acids.

How might the deprivation of nutrients trigger an apoptotic response? Perreault and Lemieux (88) have found that hybridoma cells undergo apoptosis when their protein biosynthetic machinery is compromised. It may be that deprivation of nutrients such as amino acids has the same effect, possibly resulting in the failure of the synthesis of a critical regulatory protein required to keep the apoptotic pathway in check. Vaux and Strasser (89) suggest that some agents or treatments lead to a reduction in the cellular ATP pool, and that this may be a trigger of apoptosis. They propose that such changes may be interpreted by the cell as being a consequence of viral infection, and the cell responds by inducing the apoptotic pathway. Overexpression of Bcl-2 was found to offer a high degree of protection following deprivation of each individual amino acid, with two exceptions, glutamine and threonine, which exhibited relatively less protection (87). This may indicate that these two amino acids either play a particularly important role in cellular metabolism and biosynthesis, or that they are essential components of the mechanism by which Bcl-2 protects the cell.

The survival of *bcl-2*-transfected cells, even in the absence of supposedly essential amino acids, suggests a reduction in amino acid utilization due to downregulation of nonessential cellular functions. Indeed, metabolic arrest has been reported following interleukin 3 (IL-3) withdrawal from the IL-3-dependent murine cell line Bo, which consequently underwent apoptosis. This state was stabilized by *bcl-2* transfection of the cells, thus extending survival time by 300% (90). Presumably, Bcl-2 also stabilizes the metabolic arrest caused by amino acid starvation in murine hybridoma cells, possibly by maintaining the ATP pool above a threshold level. Moreover, studies by Simpson et al. (8) would suggest that this state may be reversible by feeding the cells with fresh medium. Clearly, such a characteristic is far more desirable than a rapid, and obviously nonreversible, entry into apoptosis that occurs in apoptosis-susceptible cell lines.

Serum. The role of serum in the suppression of apoptosis has been well documented, and consequently, it was of no surprise when it was reported that commercially important hybridoma and plasmacytoma cell lines also undergo apoptosis on withdrawal of serum (85). This would be expected to have important consequences for the development of new serum-free media formulations. Previously, the rational behind the design of such media was not particularly scientific and often involved the inclusion of chemicals that were identified by empirical studies. However, demonstration of the role of serum in the regulation of apoptosis may provide a new avenue of research for the development of novel, and perhaps cheaper, serum-free media. The Bcl-2-mediated suppression of apoptosis following serum withdrawal was the earliest demonstration of the antiapoptosis activity of this gene. A Burkitt's lymphoma cell line transfected with *bcl-2* has been reported to grow better than control vector–transfected cells in commercially available serum-free media without the need for adaptation (91). Similar results were obtained using a *bcl-2* murine hybridoma cell line.

Oxygen Limitation

In large-scale and intensive culture systems, effective aeration of the culture is a major difficulty. Thus, oxygen limitation is often the major limiting factor determining maximum cell number. Studies by Mercille and Massie (92) demonstrated that deprivation of oxygen can induce apoptosis. Subsequently, it has been shown that Bcl-2 overexpression protects hybridoma (8) and Burkitt's lymphoma cells (80) from apoptosis. Clearly this will provide the cells with a considerable advantage in oxygen-limited intensive culture systems.

Hydrodynamic Stress

During the very early days of mammalian cell culture, it was generally assumed that cell lines would be far more sensitive to shear damage due to the absence of a cell wall. Considerable progress has been made in our understanding of the exact interactions that result in the greatest damage to the cells. It is now clear that it is not the shear forces generated at the impeller tip that are responsible for cell death induced by the hydrodynamic environment of the reactor. Far more damage is caused by the events

that take place at the gas head space–liquid interface during bubble disengagement (95).

Despite the importance of the hydrodynamic environment, the biochemical response of the cell to this component of the bioreactor environment has been rather neglected. To address this issue Al-Rubeai et al. (91) investigated the mechanism of cell death following exposure of cells to very high agitation levels. Flow cytometric and morphological analysis indicated that cell death occurred mostly by apoptosis, although levels of necrosis were also significant. Singh et al. (82) found that a Burkitt's lymphoma cell line that had been routinely passaged in stationary cultures underwent apoptosis when attempts were made to grow the cells in suspension. *Bcl-2* transfection of this cell line was found to allow much better cell growth in suspension without the need for adaptation. Simpson et al. (8) reported similar behavior of a hybridoma cell line.

Physiological studies have also investigated the relationship between apoptosis and shear stress. Dimmeler et al. (89) have found that shear stress actually prevents induction of apoptosis in endothelial cells in the presence of the inducer tumor necrosis factor α or following growth factor withdrawal. Clearly, the possibility that a sublethal level of shear stress in the bioreactor protects the cells from apoptosis-inducing agents needs to be investigated.

Growth Arrest, Specific Productivity, and Apoptosis

Studies of p53 and c-*myc* have revealed a close relationship between cell proliferation and apoptosis. These studies may have important implications for process optimization strategies that have centered on the control of cellular proliferation. Such an approach allows for improvements in specific recombinant protein productivity during cultures that exhibit a negative correlation between growth rate and productivity. Furthermore, by controlling maximum cell number at an optimal level, cell death due to limitation of nutrients and oxygen should be minimized, thus simplifying medium clarification during downstream processing. However, such studies have had one major drawback: when attempts were made to control cellular proliferation, the cultures rapidly lost viability (93,94), and it would appear that, at least in the case of hybridoma cultures, this was due to the induction of apoptosis (85). Thus strategies designed to control cellular proliferation must also incorporate methods that minimize apoptosis. Initial results would appear to suggest that such an approach does work. Simpson et al. (8) found that Bcl-2 overexpression delays hybridoma cell death following cell cycle arrest induced by thymidine treatment. Similar results have been reported during Burkitt's lymphoma cultures (81).

FUTURE PROSPECTS

The study of apoptosis from an animal-cell-technology standpoint is still in its infancy. However, even at this early stage, significant progress has been made, and it would appear that suppression of apoptosis can lead to improvements in culture viability and, indeed, productivity. It is now necessary to study the prevalence of apoptosis in other culture systems, and then to examine the effect of the suppression of bcl-2 overexpression on cell survival. Studies of a variety of high-density perfusion systems are currently underway. It is also necessary to establish the incidence of apoptosis in other industrially important mammalian cell lines, and to investigate the level of protection afforded by antiapoptosis genes. If the results presented here are found to be reproducible in many different industrial culture systems and cell lines, the study of apoptosis will provide the process biotechnologist with a further strategy in the quest for the optimal production process.

BIBLIOGRAPHY

1. J.F.R. Kerr, A.H. Wyllie, and A.R. Currie, *Br. J. Cancer* **26**, 239–257 (1972).
2. H. Ohyama, T. Yamada, and I. Watanabe, *Radiat. Res.* **85**, 333–339 (1981).
3. N. Thomas and P.A. Bell, *Mol. Cell. Endocrinol.* **22**, 71–84 (1981).
4. C.J. Sanderson, in W.R. Clark and P. Golstein eds., *Mechanisms of Cell Mediated Cytotoxity*, Plenum, New York, 1982, pp. 3–21.
5. F. Oberhammer, J.W. Wilson, C. Dive, I.D. Morris, J.A. Hickman, A.E. Wakeling, P.R. Walker, and M. Sikorska, *EMBO J.* **12**, 3679–3684 (1983).
6. A.H. Wyllie, *Nature* **284**, 555–556 (1980).
7. M. Piacentini, L. Fesus, M.G. Farrace, L. Ghibelli, L. Piredda, and G. Melino, *Eur. J. Cell Biol.* **54**, 246–254 (1991).
8. N. Simpson, A.N. Milner, and M. Al-Rubeai, *Biotechnol. Bioeng.* **54**, 1–16 (1997).
9. K.D. Steck, T.J. Mcdonnell, and A.K. Elnaggar, *Cytometry* **20**, 54–161 (1995).
10. M.G. Ormerod, X.M. Sun, R.T. Snowden, R. Davies, H. Fearnhead, and G.M. Cohen, *Cytometry* **14**, 595–602 (1993).
11. K.H. Elstein and R.M. Zucker, *Exp. Cell Res.* **211**, 322–331 (1994).
12. K.H. Elstein, D.J. Thomas, and R.M. Zucker, *Cytometry* **21**, 170–176 (1995).
13. I. Nicoletti, G. Migliorati, M.C. Pagliacci, F. Grignani, and C. Riccardi, *J. Immunol. Methods* **139**, 271–279 (1991).
14. Y. Tsujimoto, N. Ikegaki, And C.M. Croce, *Oncogene* **2**, 3–7 (1987).
15. V.A. Fadok, D.R. Voelker, P.A. Campbell, J.J. Cohen, D.L. Bratton, and P.M. Henson, *J. Immunol.* **148**, 2207–2216 (1992).
16. G. Koopman, C.P.M. Reutelingsperger, G.A.M. Kuijten, R.M.J. Keehnen, S.T. Pals, and M.H.J. Vanoers, *Blood* **8**, 1415–1420 (1994)
17. M. Castedo, T. Hirsch, S.A. Susin, N. Zamzami, P. Marchetti, A. Macho, and G. Kroemer, *J. Immunol.* **157**, 512–521 (1996).
18. M.C. Raff, *Nature* **356**, 397–400 (1992).
19. S. Nagata, *Adv. Immunol.* **57**, 129–144 (1994).
20. J.C. Rathwell and C.C. Goodnow, *J. Immunol.* **153**, 2831–2842 (1994).
21. G.G. Singer and A.K. Abbas, *Immunity* **1**, 365–371 (1994).
22. P.T. Daniel and P.H. Krammer, *Immunology* **152**, 5624–5632 (1994).
23. R.J. Clem, M. Fechheimer, and L.K. Miller, *Science* **254**, 1388–1390 (1991).

24. J.C. Ameisen and A. Capron, *Immunol. Today* **12**, 102–391 (1991).
25. C. Terai, R.S. Kornbluth, C.D. Pauza, D.D. Richman, and D.A. Carson, *J. Clin. Invest.* **87**, 1710–1715 (1991).
26. H. Groux, G. Torpier, D. Monte, Y. Mouton, J.M. Bourez, A. Capron, and J.C. Ameisen, *J. Exp. Med.* **175**, 331–340 (1992).
27. L. Meyaard, S.A. Otto, R.R. Jonker, M.J. Mijinster, R.P.M. Keet, and F. Miedema, *Science* **257**, 217–219 (1992).
28. D.I. Cohen, Y. Tani, H. Tian, E. Boone, L. Samelson, and H.C. Lane, *Science* **256**, 542–545 (1992).
29. A.G. Laurent-Crawford, B. Krust, Y. Riviere, C. Desgranges, S. Muller, M.P. Kieny, C. Dauguet, and A.G. Hovanessian, *AIDS Res. Human Retroviruses* **9**, 761–773 (1993).
30. N.K. Banda, J. Bernier, D.K. Kurahara, R. Kurrle, N. Haigwood, R.P. Sekaly, and T.H. Finkel, *J. Exp. Med.* **174**, 1099–1106 (1992).
31. T.H. Finkel, G. Tudor-Williams, N.K. Banda, M.F. Cotton, T. Curiel, C. Monks, T.W. Baba, R.M. Ruprecht, and A. Kupfer, *Nature Med.* **1**, 129–134 (1995).
32. D.A. Shub, *Curr. Biol.* **4**, 55–556 (1994).
33. S.V. Lennon, S.J. Martin, and T.G. Cotter, *Cell Prolif.* **24**, 203–214 (1991).
34. L.T. Zhong, T. Sarafian, D.J. Kane, A.C. Charles, S.P. Mah, R.H. Edwards, and D.E. Bredesen, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4533–4537 (1993).
35. J.W. Larrick and S.C. Wright, *FASEB J.* **4**, 3215–3223.
36. M. Iwata, M. Mukai, Y. Nakia, and R. Iseki, *J. Immunol.* **148**, 3302–3308 (1992).
37. N. Ramakrishnan and G. Catravas *J. Immunol.* **148**, 1817–1821 (1992).
38. B. Brune, P. Hartzell, P. Nicotera, and S. Orrenius, *Exp. Cell Res.* **195**, 323–329 (1991).
39. D.M. Hockenbery, Z.N. Oltvai, X.M. Yin, C.L. Milliman, and S.J. Korsmyer, *Cell* **75**, 241–251 (1993).
40. S. Shimizu, Y. Eguchi, H. Kosaka, W. Kamiike, H. Matsuda, and Y. Tsujimoto, *Nature* **374**, 811–813, (1995).
41. M.D. Jacobson, and M.C. Raff, *Nature* **374**, 814–816 (1995).
42. J.F.R. Kerr, C.M. Winterford, and B.V. Harmon, *Cancer* **73**, 2013–2026 (1994).
43. D.J. Allen, B.V. Harmon, and J.F.R. Kerr, in C.S. Potten ed., *Perspectives on Mammalian Cell Death*, Oxford University Press, Oxford, 1987, pp. 229–258.
44. K.S. Sellins, and J.J. Cohen, *Immunol* **139**, 3199–3206 (1987).
45. L.C. Stephens, K.K. Ang, T.E. Schutheiss, L. Milas, and R.E. Meyn, *Rad. Res.* **127**, 308–316 (1991).
46. K.H. Falkvoll, *Int. J. Radiat. Oncol. Biol. Phys.* **21**, 989–94, 1991.
47. T.H. Forster, D.J. Allan, G.C. Gobe, B.V. Harmon, T.P. Walsh, and J.F.R. Kerr, *Int. J. Radiat. Biol.* **61**, 365–367 (1992).
48. D.P. Lane, *Nature* **362**, 786–787 (1993).
49. M.C. Kiefer, M.J. Brauer, V.C. Powers, J.J. Wu, S.R. Umansky, L.D. Tomei, and P.J. Barr, *Nature* **374**, 736–739 (1995).
50. M.L. Gaido and J.A. Cidlowski, *J. Biol. Chem.* **266**, 580–585 (1991).
51. M.C. Peitsch, B. Polzar, H. Stephan, T. Crompton, H.R. Macdonald, H.G. Mannherz, J. Tschopp, *EMBO J.* **12**, 371–377 (1993).
52. M.A. Barry and A. Eastman, *Arch. Biochem. Biophys.* **300**, 440–450 (1993).
53. Y. Tsujimoto, J. Cossman, E. Jaff, and C.M. Croce, *Science* **228**, 1440–1443 (1985).
54. T.J. McDonnell, N. Deane, F.M. Platt, G. Nunez, U. Jaeger, J.P. Mckearn, and S.J. Korsemeyer, *Cell* **57**, 79–88 (1989).
55. T.J. McDonnell, G. Nunez, F.M. Platt, D.M. Hockenbery, L. London, J.P. Mckearn, and S.J. Korsemeyer, *Mol. Cell. Biol.* **10**, 1901–1907 (1990).
56. A. Strasser, A.W. Harris, M.L. Bath, and S. Cory, *Nature* **348**, 331–333 (1990).
57. S.W. Lowe, E.M. Schmitt, S.W. Smith, B.A. Osborne, and T. Jacks, *Nature* **362**, 847–849 (1993).
58. M.O. Hengartner, and H.R. Horvitz, *Cell* **76**, 665–676 (1994).
59. J. Yuan, S. Shaham, S. Ledoux, H.M. Ellis, and H.R. Horvitz, *Cell* **75**, 641–652 (1993).
60. M. Miura, H. Zhu, R. Rotello, E. Hartwig, and J. Yuan, *Cell* **75**, 653–660 (1993).
61. K. Kiuda, J.A. Lippke, G. Ku, M.W. Harding, D.S. Livingstone, M.S.S. Su, and R.A. Flavell, *Science* **267**, 2000–2003 (1995).
62. M. Miura, H. Zhu, R. Rotello, E.A. Hartwieg, and J.Y. Yuan, *Cell* **75**, 653–660 (1993).
63. S. Kumar, M. Kinoshita, M. Noda, N.G. Copeland, and N.A. Jenkins, *Genes Dev.* **8**, 1613–1626 (1994).
64. A.T. Fernandes, G. Litwack, and E.S. Alnemri, *J. Biol. Chem.* **269**, 30761–30764 (1994).
65. C. Faucheu, A. Diu, A.W.E. Chan, A.M. Blanchet, C. Miossec, F. Herve, V. Collarddutilleul, Y. Gu, R.A. Aldape, J.A. Lippke, C. Rocher, M.S.S. Su, D.J. Livingston, T. Hercend, and J.L. Lalanne, *EMBO J.* **14**, 1914–1922 (1995).
66. N.A. Munday, J.P. Vaillancourt, A. Ali, F.J. Casano, D.K. Miller, S.M. Molineaux, T.T. Yamin, V.L. Yu, and D.W. Nicholson, *J. Biol. Chem.* **270**, 15870–15876 (1995).
67. A.T. Fernandes, G. Litwack, and E.S. Alnemri, *Cancer Res.* **55**, 2737–2742 (1995).
68. M.P. Boldin, E.E. Varfolomeev, Z. Pancer, I.L. Mett, J.H. Camonis, and D. Wallach, *J. Biol. Chem.* **270**, 7795–7798 (1995).
69. A.M. Chinnaiyen, K. O'Rourke, M. Tewari, and V.M. Dixit, *Cell* **81**, 505–512 (1995).
70. M.P. Boldin, T.M. Goncharov, Y.V. Goltsev, and D. Wallach, *Cell* **85**, 803–815 (1996).
71. M. Muzio, A.M. Chinnaiyen, F.C. Kischkel, K. O'Rourke, A. Shevchenko, J. Ni, C. Scaffidi, J.D. Bretz, M. Zhang, R. Gentz, P.H. Krammer, M.E. Peter, and V.M. Dixit, *Cell* **85**, 817–827 (1996).
72. L.C. Cooney, in E.C. Beuvery, J.B. Giffiths, and W.P. Zeijlemaker, eds., *Animal Cell Technology: Developments towards the 21st Century*, Kluwer Academic, Netherlands, 1995, pp. xxv–xxx.
73. M. Al-Rubeai, D. Mills, and A.N. Emery, *Cytotechnology* **4**, 13–28 (1990).
74. F. Franek and J. Dolnikova *FEBS Lett.* **248**, 285–287 (1991).
75. S. Mercille and B. Massie, Induction of apoptosis in nutrient-limited cultures of hybridoma cells. *Biotechnol. Bioeng.* **44**, 1140–1154 (1994).
76. R.P. Singh, M. Al-Rubeai, C.D. Gregory, and A.N. Emery, *Biotechnol. Bioeng.* **44**, 720–726 (1994).
77. A. Moore, C.J. Donahue, J. Hooley, D.L. Stocks, K.D. Bauer, and J.P. Mather, *Cytotechnology* **17**, 1–11 (1995).
78. Y. Itoh, H. Ueda, and E. Suzuki, *Biotechnol. Bioeng.* **48**, 118–122 (1995).
79. E. Suzuki, S. Terada, H. Ueda, T. Fujita, T. Komatsu, S. Takayama, and J.C. Reed, *Cytotechnology* **23**, 55–59 (1997).

80. S. Mercille and B. Massie, *Cytotechnology* **15**, 117–128 (1994).
81. R.P. Singh, A.N. Emery, and M. Al-Rubeai, *Biotechnol. Bioeng.* **52**, 166–175 (1996).
82. R.P. Singh, G. Finka, A.N. Emery, and M. Al-Rubeai, *Cytotechnology* **23**, 87–93 (1997).
83. K. Murray, C.E. Ang, K. Gull, J.A. Hickman, and A.J. Dickson, *Biotechnol. Bioeng.* **55**, 298–304 (1996).
84. N.H. Simpson, R.P. Singh, A. Perani, C. Goldenzon, and M. Al-Rubeai, *Biotechnol. Bioeng.* **59**, 90–98 (1997).
85. D.L. Vaux and A. Strasser, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2239–2244 (1996).
86. A. Handa-Corrigan, A.N. Emery, and R.E. Spier, *Enzyme Microb. Technol.* **11**, 230–235 (1989).
87. M. Al-Rubeai, R.P. Singh, M.H. Goldman, and A.N. Emery, *Biotechnol. Bioeng.* **45**, 463–472 (1994).
88. J. Perreault and R. Lemieux, *Cytotechnology* **13**, 99–105 (1994).
89. S. Dimmeler, J. Haendeler, V. Rippmann, M. Nehls, and A.M. Zieher, *FEBS Lett.* **399**, 71–74 (1996).
90. M. Al-Rubeai and A.N. Emery, *J. Biotechnol.* **16**, 67–86 (1990).
91. M. Al-Rubeai, A.N. Emery, S. Chalder, and D.C. Jan, *Cytotechnology* **9**, 85–97 (1992).
92. Z. Oltvai, C. Milliman, and S.J. Korsmeyer, *Cell* **74**, 609–619 (1993).
93. M.O. Hengartner, R.E. Ellis, and H.R. Horvitz, *Nature* **356**, 494–499 (1992).
94. S. Henderson, D. Huen, M. Rowe, C. Dawson, G. Johnson, and A.B. Rickinson, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8479–8483 (1993).

ADDITIONAL READING

E. Yang, J.P. Zha, J. Jockel, L.H. Boise, C.B. Thompson, and S.J. Korsmeyer, *Cell* **80**, 285–291 (1995).
S.N. Farrow, J.H.M. White, I. Martinou, T. Raven, K.T. Pun, C.J. Grinham, J.C. Martinou, and R. Brown, *Nature* **374**, 731–733 (1995).
L.H. Boise, M. Gonzalezgarcia, C.E. Postema, L.Y. Ding, T. Lindsten, L.A. Turka, X.H. Mao, G. Nunez, and C. Thompson, *Cell* **74**, 597–608 (1993).
K.M. Kozopas, T. Yang, H.I. Buchan, P. Zhou, and R.W. Craig, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 3516–3520 (1993).

ASPARTAME

SATOSHI HANZAWA
TOSOH Corp.
Ayase-shi, Japan

KEY WORDS

Aspartame
Enzymatic synthesis
Protease
Protein engineering
Sweetener

OUTLINE

INTRODUCTION

Aspartame, α-L-aspartyl-L-phenylalanine methyl ester (APM) (Fig. 1), is an artificial sweetener whose sweetness is about 200 times stronger than sucrose by weight. Its sweetness was found accidentally by a researcher of Searle Research Laboratories, where APM was produced as an intermediate of a C-terminal tetrapeptide of gastrin (1). Its strong sweetness allows reduction of the amount of calorie sweetener added to foods to achieve a sweetness equivalent to that from sucrose. Although other intense sweeteners often exhibit a slightly bitter aftertaste, APM has none (2). Its sweet taste is close to that of sucrose, a traditional sweetener and the standard of sweetness. Therefore, the demand for APM has grown rapidly since its approval by the Food and Drug Administration (FDA) in 1981. Worldwide APM consumption is currently over 10,000 tons per year, and it is the most common intense sweetener for beverages, tabletop sweetener, ice cream, chewing gum, and so forth. APM is produced by two methods: a chemical method, used by Nutrasweet in the United States and Ajinomoto in Japan, and an enzymatic method, used by Holland Sweetener Company, a joint venture of TOSOH (formerly Toyo Soda Manufacturing) in Japan and DSM in the Netherlands. In the enzymatic method, an APM precursor, *N*-benzyloxycarbonyl-APM (Z-APM) is produced by a protease-catalyzed condensation reaction of *N*-benzyloxycarbonyl-L-aspartic acid (Z-L-Asp) and L-phenylalanine methyl ester (L-PheOMe). The advantages of the enzymatic method are that a racemic substrate can be used and there is no production of β-aspartame (β-APM) (Fig. 1). The reduction of enzyme consumption was, however, desirable for further cost reduction. In the present article, the APM production process with the protease reaction and various efforts to reduce the enzyme consumption therein are described.

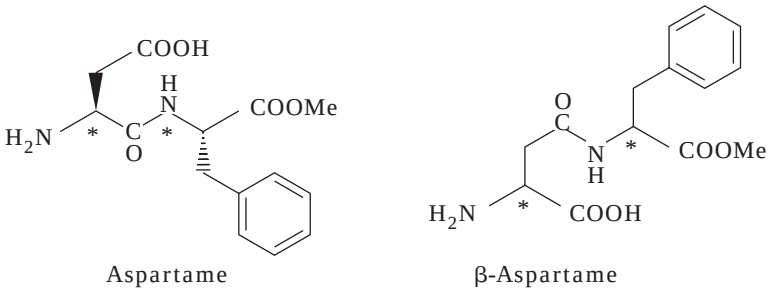


Figure 1. Structures of aspartame and its regioisomer. Asymmetric carbon atoms are indicated by an asterisk. The configuration of aspartame is L-α-L.

ASPARTAME PRODUCTION BY PROTEASE

Advantages of Aspartame Production by Protease

The structures of APM and its regioisomer are shown in Figure 1. In APM production by conventional organic synthesis, the side-chain carboxylic group (β-carboxylic group) of aspartic acid must be protected to avoid the formation of unwanted β-APM; additionally, optically pure substrates are required to avoid the production of stereoisomers of APM. These APM-related compounds, such as β-APM and stereoisomers, exhibit a bitter taste (1). Use of a protease as catalyst does not give formation of these undesirable compounds, even if racemic substrates are used as starting materials and without protection of the β-carboxylic group of aspartic acid. The use of racemic phenylalanine in the enzyme process is very advantageous, because L-phenylalanine is more expensive than racemic phenylalanine.

Proteases such as thermolysin, subtilisin, or papain were found to be useful tools for the peptide synthesis because of their high stereo and regio selectivity, and they have been applied in the synthesis of various biologically active peptides, such as peptide hormones, neuropeptides, and insulin (3). APM synthesis catalyzed by a protease was also investigated by Isowa et al. (4). They reported that N-protected aspartic acid and Phe-alkyl ester are condensed to N-protected L-α-Asp-L-Phe-alkyl ester by thermolysin in high yield (83–96%) as a precipitate consisting of the addition compound of the condensation product and one molecule of Phe-alkyl ester, as shown in Scheme 1.

In this scheme, R indicates N-protective group, such as Z or PMZ, and R1 indicates lower alkyl ester, such as methyl or ethyl ester. Proteases usually catalyze the hydrolysis of peptides in water as solvent; the condensation reaction is, in other words, the reverse reaction. Therefore, it is usually difficult to obtain products in high yield because of the occurrence of hydrolysis of the products by the proteases. The synthesis of APM-related peptides can, however, be performed in high yield when the product is removed from the reaction system as a precipitate and is not hydrolyzed back by the protease, for example, by thermolysin.

Although this method requires an expensive catalyst, such as thermolysin, the high stereo and regio selectivity is a great advantage in APM production. The industrial APM production process based on the proposal of Isowa et al. was developed by Toyo Soda Manufacturing and was adopted for APM production by Holland Sweetener Company.

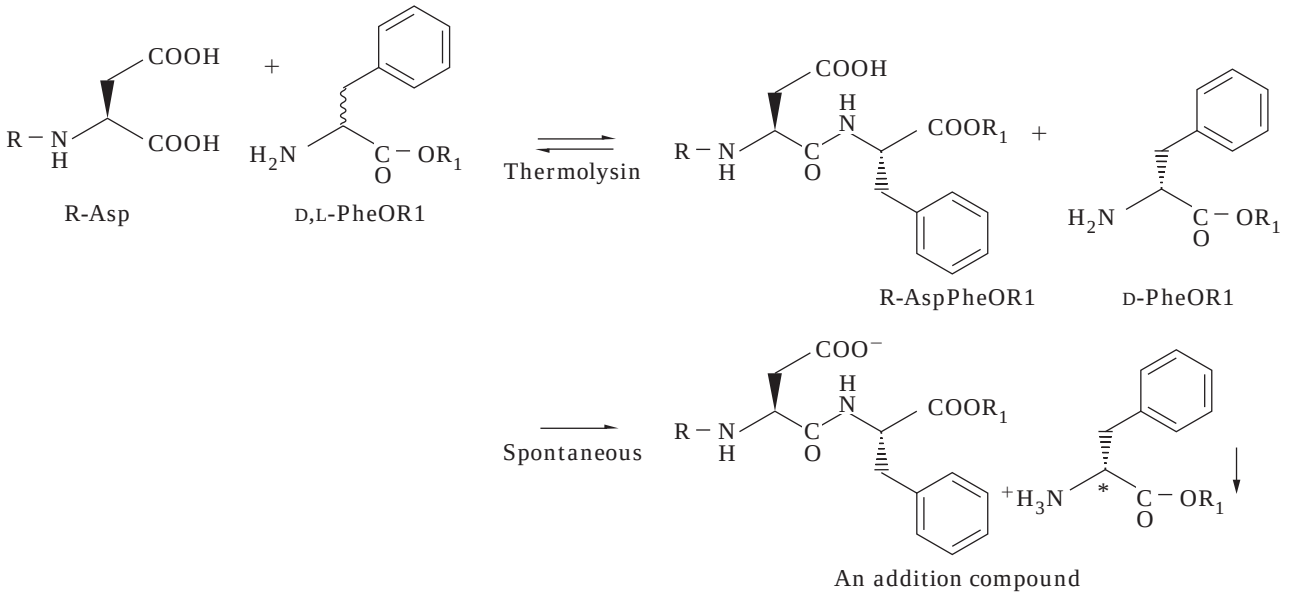
In contrast to the enzymatic method, the chemical method produces an APM precursor where formyl-APM (f-APM) is derived from *N*-f-Asp anhydride and L-PheOMe, (5) as is shown in Scheme 2.

Although β-APM is produced as by-product in this method, α-selectivity of the reaction can be enhanced by selection of the optimal solvent for the reaction, such as alkaline aqueous medium (6) or acetic acid partially replaced with an alkyl ester with a secondary or tertiary alcohol (7). It was reported that the production of β-APM was reduced to 20% of total peptide by selection of a proper solvent. The β-APM reportedly can be converted to APM by alkaline treatment in alcohol in the presence of metal ions such as zinc or copper as catalysts (8). Productivity of APM seems to be enhanced by these methods in the chemical procedure.

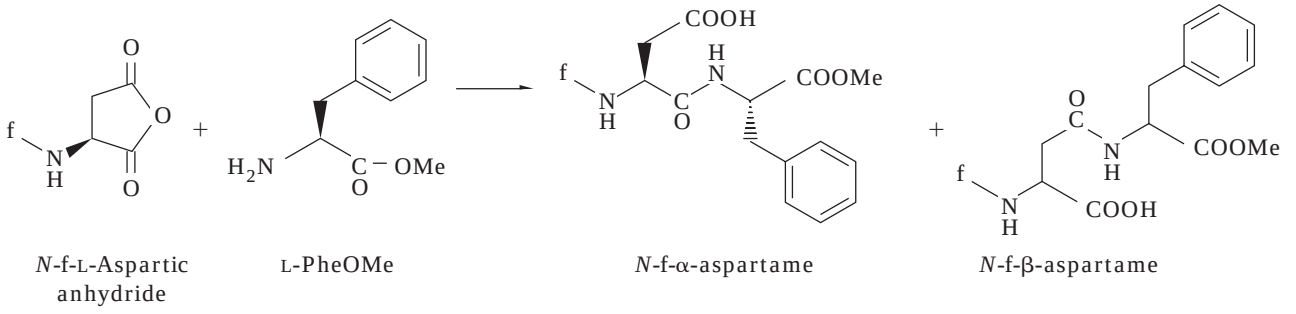
As described earlier, the enzymatic and chemical methods both have individual advantages. An advantage of the enzymatic method is the high selectivity of the reaction, where APM is produced from cheap raw materials by a simple procedure in high yield. Not requiring expensive catalysts, such as enzymes, in the condensation step is an advantage of the chemical method; however, expensive starting materials must be used and by-product is produced. Attempting cost reduction in APM production and improving the condensation step are of great importance for APM production by Holland Sweetener Company, and many efforts have been made by both TOSOH and DSM. Hereafter, I describe the enzymatic APM production process and the efforts to improve the condensation reaction.

Aspartame Production by Thermolysin

The principles of the thermolysin process were introduced by Oyama et al. (9) in 1987. At first, L-PheOMe from racemic PheOMe is coupled with Z-Asp stereo and regio selectively by thermolysin. The reaction product, Z-APM, forms an insoluble addition compound with one molecule of remaining D-PheOMe immediately and precipitates. When the reaction is completed, the Z-APM·D-PheOMe addition compound is collected as a solid product, and thermolysin is recovered from the solution. D-PheOMe is separated from Z-APM as its hydrochloride by acid washing (with hydrochloric acid, for example). The Z-group is removed from the Z-APM by hydrogenolysis. Finally, APM is obtained by crystallization and drying. D-PheOMe is recycled after racemization by alkaline treatment and reesterification in methanol. The process is summarized in Figure 2.



Scheme 1.



Scheme 2.

Although the β-carboxylic group of aspartic acid does not need to be protected, the α-amino group of aspartic acid and the carboxylic group of phenylalanine must be protected. This plays an important role not only for the selective reaction of the α-carboxylic group of L-aspartic acid and the amino group of L-phenylalanine, but also for being suitable to be used by thermolysin as substrates. Thermolysin can not catalyze the condensation reaction without such protections (9,10). The methyl ester group is used preferably for the protection of the carboxyl group of phenylalanine because this is an essential group expressing the sweetness of APM. The amino group of aspartic acid is protected by the Z-group because it is easily removed by hydrogenolysis without removal of the methyl ester group at the C-terminal of APM. Additionally, a higher reaction rate is obtained when a hydrophobic and bulky group exists at this position. Although formyl or acetyl groups can also be used for the protection, the Z-group, which is hydrophobic and bulky, is more suitable for the reaction rate of the condensation.

Thermolysin and similar microbial metalloproteases (thermolysin-like metalloproteases) were selected from among a number of industrially available proteases. Although other serine or cysteine proteases, such as subtilisin or papain, have been more widely used in various industrial fields, they are not very suitable for APM production because they tend to hydrolyze lower alkyl esters at the C-terminal of peptides (11–13) and also the methyl ester group of L-PheOMe and Z-APM. The cleaving sites of proteases in Z-APM are summarized in Table 1, where residues of Z-APM are numbered as P1, P2, and P3, respectively, for those toward the N-terminal from the cleavage site, and as P′1 and P′2, respectively, for those toward the C-terminal from the cleavage site. Specificity of thermolysin and thermolysin-like metalloproteases is restricted to hydrophobic and bulky amino acids, such as phenylalanine or leucine (14) at the P′1 position. The metalloproteases also have esterase activity, but this activity is restricted in peptide mimic esters like Bz-Gly-OPhe-Ala or furylacryloyl-Gly-OLeu-NH_2 (15). The methyl ester of L-PheOMe or Z-APM is not hydrolyzed by them at all, because the methyl ester group is too small at P′1 position. Z-APM is, therefore, produced by thermolysin or thermolysin-like metalloproteases without by-products.

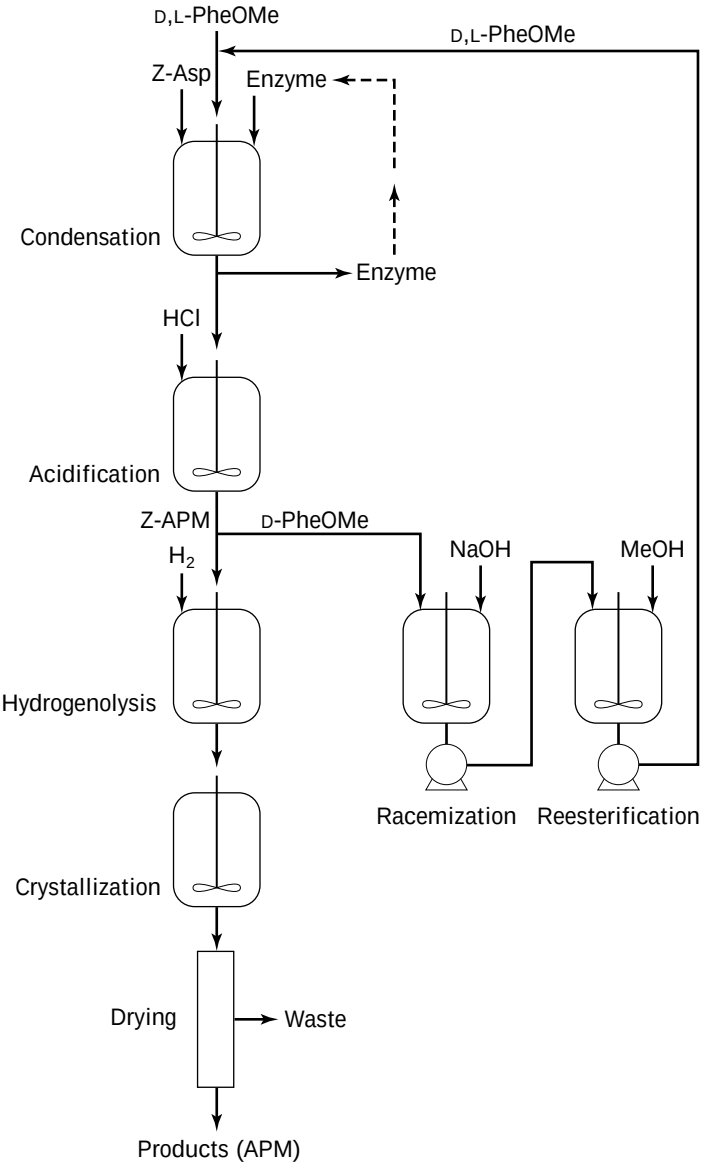


Figure 2. Process of aspartame production via thermolysin catalyzed reaction.

Table 1. Cleavage Sites on Z-APM

Cleavage site	P3 - P2 - P1 - P′1 - P′2 ↑
Thermolysin	Z - Asp - Phe- OMe
Subtilisin	Z - Asp-Phe - OMe
Papain	Z - Asp - Phe- OMe
	Z - Asp-Phe - OMe
V8 protease	Z - Asp - Phe- OMe

IMPROVEMENT OF THE CONDENSATION REACTION

Stability of Enzymes

To improve the condensation reaction in the enzymatic process, reduction of enzyme consumption may be of great importance. Although thermolysin is currently the most stable of the proteases industrially available (16), some portion of it is inactivated during the condensation reaction. This enzyme inactivation, of course, influences production costs. Stabilization of the enzyme during the condensation reaction thus was one of the most important aspects of cost reduction.

It is well known that calcium ion stabilizes many microbial proteases, and thermolysin is also stabilized by it (14). If the concentration of calcium ions in the reaction mixture is increased, it will diminish the inactivation of thermolysin. However, this would not be suitable to improve this process, because high concentration of calcium will frequently cause troubles caused by scaling.

If thermolysin can be replaced by a more stable protease, the enzyme loss can be reduced further. Extremely stable proteases such as archaelysin (17) and a serine protease from *Desulfurococcus* strain SY (18) were recently found in cultures of hyperthermophilic archaea (formerly called archaeobacteria). The archaea, such as *Desulfurococcus*, *Pyrococcus*, or *Thermococcus*, showed an optimum

temperature for growth around 90 °C and even can grow above 100 °C. Their enzymes exhibited extremely high stability against extreme conditions; archaelysin and the protease from *D.* strain SY retained 50% of their activity when they were heated at 95 °C for 80 min and 450 min, respectively (17,18). These proteases are expected to have applicability in various industrial fields (19). Unfortunately, no metalloprotease has yet been found from hyperthermophilic archaea. The proteases separated from the archaea up to now belong to serine proteases (17,18) or cysteine proteases (20). Because their specificity toward Z-APM is similar to that of subtilisin or papain, they are not appropriate for APM production. They will hydrolyze the ester bond of PheOMe or APM in the same way as subtilisn or papain. Although finding a metalloprotease from such archaea would be desirable, thermolysin remains the most stable protease currently available for APM production. Recently a metalloprotease was found from a culture broth of a hyperthermophilic aerobic archaeon "Aeropyrum pernix K1" that also exhibited extremely high stability against thermal denaturation with a half-life of 70 min at 125 °C (21), although its specificity for APM relating peptides has not been solved.

Immobilized Thermolysin

Immobilization of thermolysin was also expected to reduce enzyme inactivation because it is protected from autolysis. However, an application of immobilized thermolysin in the current production process seems difficult, because the reaction for Z-APM production takes place in the presence of a dense solid precipitate, that is, an insoluble addition compound of Z-APM and PheOMe. If both the immobilized enzyme and the reaction product occur in the reaction mixture as solids, they cannot be separated by a simple procedure. Additionally, if the reaction would be performed in a continuous process equipped with an immobilized thermolysin-packed column, the precipitate would block the flow of the reaction solution in the column. Z-APM forms its insoluble addition compound with PheOMe quite easily in water.

To overcome this problem, Oyama et al. (9,22) attempted to perform the reaction in ethyl acetate. Because ethyl acetate dissolves Z-APM without formation of the insoluble addition compound, it was expected that Z-APM could be recovered in solution and that the immobilized thermolysin could be recovered as a solid, thus enabling complete recycling of enzyme (9). Moreover, because the equilibrium state of the synthesis and hydrolysis of Z-APM shifts toward the synthesis side in organic solvents, it was expected that Z-APM could be obtained in high yield in ethyl acetate even though the product is not removed from reaction system in situ. Unfortunately, significant inactivation of immobilized thermolysin was observed in both batchwise operation and continuous operation. Additionally, Z-APM synthesis was disturbed by channeling of organic solvent and water in the immobilized thermolysin column when it was operated continuously (9); thus, high production yield could not be obtained by this system. Only if these problems can be solved, immobilized thermolysin can be applied in the APM production. Studies for APM production by immobilized thermolysin are being continued by various researchers (23,24).

Enzyme Improvement by Site-Directed Mutagenesis

Improvement of enzymes by site-directed amino acid substitutions at selected sites has recently become possible by the progress in genetic engineering and of computer simulation of the three-dimensional structure of the protein molecules. These so-called protein engineering techniques are becoming an important tool for the improvement of enzyme character. A first example of improved thermolysin-like metalloproteases was shown by Imanaka et al. (24) using a metalloprotease from *Bacillus stearothermophilus* CU21. This protease exhibited 85% homology in amino acid sequence with thermolysin. They substituted the 141st glycine of this protease to alanine and expected to find enhancement of the internal hydrophobicity and stabilization of the internal α-helix. The resulting mutant protease indeed exhibited higher thermostability than did wild-type protease (25). Another example of stabilization of this metalloprotease was reported by Hardy et al. (26), who introduced proline at the 65th or 69th site. Although the stability of this protease was lower than that of thermolysin even with introduction of these mutations, it may be expected that thermolysin can be improved by a similar procedure. Thus, studies for the improvement of thermolysin by protein engineering were started by a collaboration of TOSOH and Sagami Chemical Research Center.

In this research, attention was focused to enhancing the activity of thermolysin rather than its stability. Even when the stability of thermolysin is enhanced until the enzyme inactivation during the condensation reaction is at a negligible level, the enzyme has to be recycled for reduction of enzyme costs. This requires additional equipment or procedures to recover the enzyme from the process solution, such as concentration of enzyme by ultrafiltration, salting out, or immobilization of enzyme (9). This will increase the fixed costs at an APM plant. On the other hand, if the activity is enhanced adequately and if the enzyme amount in the APM production can be reduced to a negligible level by application of highly active mutant thermolysins, it would be possible to use the enzyme without recycling. A highly active mutant thermolysin would be more effective than a highly stable one. Additionally, useful knowledge would be gained about the mechanism determining the relationships among the proteins' structures, properties, and functions from the improvement of thermolysin activity.

Thus, studies for improvements of thermolysin were started, and highly active mutant thermolysins were successfully constructed. One of these mutants exhibited a five times higher activity toward Z-APM synthesis when it is evaluated by the initial reaction rates (27). Details of the improvements of thermolysin are described in the article entitled THERMOLYSIN. The stability of these mutant thermolysins was the same as that of wild-type thermolysins when it was determined by calorimetry (28). It is therefore expected that mutant thermolysin can be used for APM production in the same way as wild-type thermolysin, except that the enzyme amount to obtain the same reaction rate is lower with mutant thermolysin.

Evaluation of Mutant Thermolysins in APM Production

The highly active thermolysin exhibiting a five times higher activity than wild-type thermolysin was called TZ-1. TZ-1 has substituted amino acid residues at three

sites, namely the 144th leucine to serine (L144S), the 150th aspartic acid to histidine (D150H) and the 227th asparagine to histidine (N227H). Although this mutant exhibits five times higher activity toward Z-APM synthesis, it should be noticed that its activity was evaluated at conditions different from real APM production conditions as described in the article on thermolysin referred to earlier. The evaluation conditions were chosen to reduce the viscosity of the reaction solution; the concentration of substrates was lower than that under conventional production conditions, and the ratio of Z-Asp and PheOMe was set at about 1:1. These conditions enabled to obtain exact data in a milliliter scale of reaction mixture. These conditions are referred to as "model conditions." The model conditions enabled evaluating several candidates during molecular engineering of mutants. Because TZ-1 showed very high activity under the model conditions, its efficiency in small-scale APM production under the conventional production conditions was also examined.

The reactions were performed in 200 mL of a mixture containing 0.2 M of Z-L-Asp, 0.5 M of D,L-PheOMe, 5.1 mM $CaCl_2$, 0.17 M NaCl and 50–150 mg of the thermolysin at 40 °C and in a range of pH between 5 and 7. Mutant thermolysins used here were produced by fermentation of *Bacillus subtilis* MT-2, harboring recombinant plasmid pUBTZ2 in a jar fermentor equipped with a 16-L vessel (New Brunswick Scientific Co. Ltd., U.S.). Other conditions for the fermentation were the same as described in the article on thermolysin. Enzyme amounts in the reaction mixture were determined by casein hydrolytic activity according to the method of Endo (29) as are also described in the article on thermolysin. The condensation reactions were monitored by HPLC equipped with TSK gel G2000SW (TOSOH, Japan) with 5% acetonitrile in 0.05% trifluoroacetic acid as solvent for determining the amount of the product and the remaining substrates. Before HPLC, the samples were diluted by appropriate solution, such as 10 mM calcium acetate, and the insoluble addition compound of Z-APM and PheOMe was removed by centrifugation. The conversion rates were calculated from remaining Z-Asp using the following formula:

$$CV(\%) = ([\text{Z-Asp}]_0 - [\text{Z-Asp}]_t)/[\text{Z-Asp}]_0$$

Herein, $[\text{Z-Asp}]_0$ and $[\text{Z-Asp}]_t$ refer to the concentration of Z-Asp at the initial and at an appropriate later time, respectively.

Effect of pH on the initial reaction rates by 50 mg of TZ-1 or wild type is shown in Figure 3. The optimum pH of TZ-1 was 6.0, and that of wild type was 6.5. At pH 6.0, the activity of the TZ-1 was about fivefold that of wild type. The optimum pH shift of TZ-1 seems to result from the introduction of a plus charge of histidine instead of the 150th aspartic acid and the 227th asparagine. Because the 150th residue and the 227th residue exist at vicinity of the Zn ion and the 231st histidine residue in the active center of thermolysin, respectively, the plus charge of histidine may influence the electrostatic interactions in this center. Both Zn ion and the 231st histidine are regarded as important to express thermolysin activity (30,31). The mutation N227H especially seems to influence the pH profile

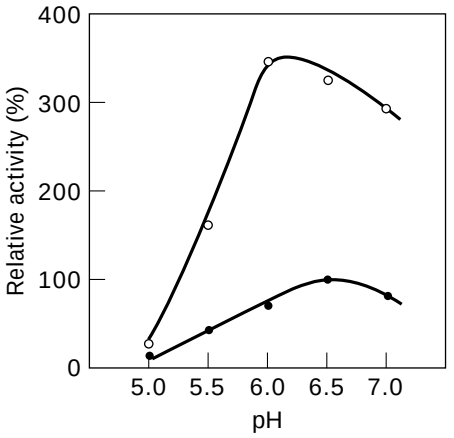


Figure 3. Effect of pH on the initial reaction rate in thermolysin catalyzed condensation reaction. Curves are shown for TZ-1 (open circle) and wild type (closed circle), respectively, for 0.2 M of Z-L-Asp and 0.5 M of D,L-PheOMe and containing 5.1 mM $CaCl_2$ and 0.17 M NaCl at 40 °C.

of thermolysin activity, because the pK_a value of the 231st histidine is likely to be shifted toward the acidic side by introduction of a plus charge of histidine at the 227th site. This shift in optimum pH is not observed, however, when the activity is evaluated under the model conditions where the optimum pH of both wild type and TZ-1 are pH 7.0, as is mentioned in THERMOLYSIN. An explanation for the difference of the optimum pH between the model conditions described in there and the production conditions described herein have not yet been found, although important factors may be *(1)* concentration of substrates, *(2)* ratio of the two substrates, and *(3)* optical purity of PheOMe.

Typical time courses of the Z-APM synthesis by TZ-1 and wild type at pH 6.0 are shown in Figure 4. Reaction rate and production yield by 50 mg of TZ-1 are the same as those by 150 mg of wild type. When the reaction is performed with 150 mg of TZ-1, the initial reaction rate is three times higher than that by the same amount of wild type. It seems that activity of TZ-1 is three times higher than that of wild type under this condition. Remaining activities of both wild type and TZ-1 after the condensation reaction were 70%. This indicates that the stability of TZ-1 is the same as that of wild type. Moreover, there are no differences in thermostability between TZ-1 and wild type as has been demonstrated by heat capacity analysis with the calorimeter as described earlier. From these results, it follows that mutant thermolysins can be useful for cost reduction in APM production.

Thus, studies for improvement of thermolysin were continued. Another important amino acid substitution, the 150th aspartic acid to tryptophan (D150W) (32) was recently found. The activity of the single site mutant D150W was five times higher than that of wild type. This is the same as the effect of the combination of three mutations, L144S, D150H, and N227H in TZ-1. Combination of D150W with L144S and N227H resulted in a new mutant, L144S-D150W-N227H, which exhibits about 10 times higher activity than wild type under the model conditions

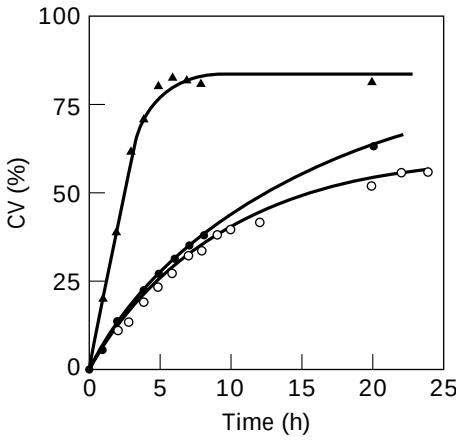


Figure 4. Time course of the condensation reaction of Z-L-Asp and L-PheOMe catalyzed by 150 mg of wild type (closed circle) and 150 mg (closed triangle) and 50 mg (open circle) of TZ-1, respectively, in 200 mL of a reaction mixture containing 0.2 M of Z-L-Asp 0.5 M of D,L-PheOMe, 5.1 mM $CaCl_2$, and 0.17 M NaCl at pH 6.0 at 40 °C.

(32). This mutant is called TZ-5 and will be useful for further reducing of costs in APM production.

Stabilization of Thermolysin by Change of Reaction Conditions

Studies on stabilization of thermolysin by controlling the reaction conditions were also continued separately by a collaboration of TOSOH and DSM. Harada et al. reported such stabilization by an N-protected amino acid, such as Z-Asp, Z-Glu or Z-Gly, during storage of thermolysin solution (33). Z-amino acids seem to act as inhibitors against the autolysis of thermolysin herein. The inhibitory effect of Z-Ala was reported earlier by Morihara and Tsuzuki (10) after they had found that Z-Ala inhibited hydrolysis of Z-Ala-Leu-Ala catalyzed by thermolysin at 30 mM of Ki value. Although these results for Z-Ala are not conclusive for determining whether inhibition by protected amino acids is competitive, it now appears that Z-amino acids are likely to bind to the active site of thermolysin instead of substrates.

Moreover, Harada et al. demonstrated that thermolysin in the reaction mixture was kept stable when Z-Asp concentration is kept at a high level during the condensation reaction. They studied effects of the ratio of Z-Asp and D,L-PheOMe at the start of the condensation reaction. An example of the results is summarized in Table 2. When the reaction was started at a 1:2 ratio of Z-Asp to PheOMe, the enzyme recovery was 92%, whereas it was 44% at a ratio of 1:2.5. Remaining Z-Asp after the reactions were 55.9 and 10.5 mmol/kg, respectively. At the ratio of 1:2, much more enzyme and Z-Asp were remaining than at the ratio of 1:2.5. The remaining Z-Asp apparently protects the enzyme from inactivation. Productivity of Z-APM at the ratio of 1:2 was 0.225 mol/kg; at 1:2.5 it was 0.238. Thus, significant decrease of productivity was not observed by higher concentration of Z-Asp. This indicates that the synthesis is not inhibited by Z-Asp, whereas Z-Asp is an inhibitor against autolysis or peptide hydrolysis. Nagayasu

Table 2. Condensation Reaction at Different Z-Aspartic Acid-to-PheOMe Ratio

Z-Asp/ PheOMe ratio	Initial [Z-Asp] (mol/kg)	Conversion rate (%)	Produced Z-APM (mol/kg)	Remained [Z-Asp] (mmol/kg)	Enzyme recovery (%)
1:2	0.28	80.5	0.225	55.9	92
1:2.5	0.25	95.2	0.238	10.5	44

et al., however, reported that Z-Asp inhibited the condensation reaction by immobilized thermolysin (24). The inhibition effect toward the condensation reaction can be avoided by choosing the reaction conditions, as suggested by Harada et al. According to Harada et al. (33), enzyme inactivation may be caused by adsorption of thermolysin on the surface of crystals of the addition compound of Z-APM and PheOMe. Apparently, Z-Asp is likely to prevent such adsorption.

In contrast to the protection by Z-Asp, thermolysin is not protected from inactivation by L-PheOMe, although the latter is another substrate of the condensation reaction; this may be because of much higher K_m value of thermolysin toward L-PheOMe. According to results of earlier kinetic studies of the condensation reaction by Oyama et al. (34), a double reciprocal plot toward L-PheOMe and D,L-PheOMe showed no apparent K_m value, whereas that toward Z-Asp is 10.3 mM. These authors presented a reaction model where thermolysin at first forms an enzyme-Z-Asp complex, and then Z-APM is produced by the attack of L-PheOMe on the complex, as shown in Scheme 3. In this scheme, TLN indicates thermolysin. This model suggests that thermolysin cannot form an enzyme–L-PheOMe complex to prevent autolysis and adsorption on crystals of the addition compound. Therefore, thermolysin does not seem to be stabilized by L-PheOMe.

The stabilization by Z-Asp is quite effective for thermolysin stabilization during APM production because the effect can be obtained quite simply by adjusting the concentration of the substrates. Moreover, Harada et al. mentioned (33) that Z-Asp stabilized thermolysin without presence of calcium ions. This is another advantage of this stabilization method because various problems caused by calcium scaling can be avoided by this method. As described earlier, enzyme inactivation during the condensation reaction can be prevented by the adjusting the ratio of the two substrates. Recently another method to stabilize thermolysin during condensation was revealed by Harada et al. (35): an addition of toluene into the reaction mixture at about 30% (w/w) of low materials. It prevented adsorption of thermolysin onto the crystal surface of the addition compound of Z-APM and PheOMe and resulted in an increase of thermolysin recovery after the reaction.

$$\text{Z-Asp} + \text{PheOMe} + \text{TLN} \rightleftharpoons \text{Z-Asp} - \text{TLN} + \text{PheOMe}$$

$$\text{Z-Asp} - \text{TLN} \longrightarrow \text{Z-APM} + \text{TLN}$$

Scheme 3.

ALTERNATIVE PROCESSES

Condensation without N-Protection

Various strategies have been used in the attempt to improve the APM production process. The possibility of reducing enzyme consumption was indicated by mutant thermolysins and by adjusting of the concentration of substrates. On the other hand, an alternative production system is desired to take long strides in reducing costs. Much of this research has been focused on avoiding N-protection of aspartic acid. If APM can be produced without N-protection, it will not only reduce raw materials costs but will also simplify the production process by eliminating the protection and de-protection steps. Although low yields have prevented industry adoption of a new method, the production costs of APM would be further reduced by N-protection-free methods if productivity could be improved. Hereafter, I summarize such attempts as appeared in patent applications.

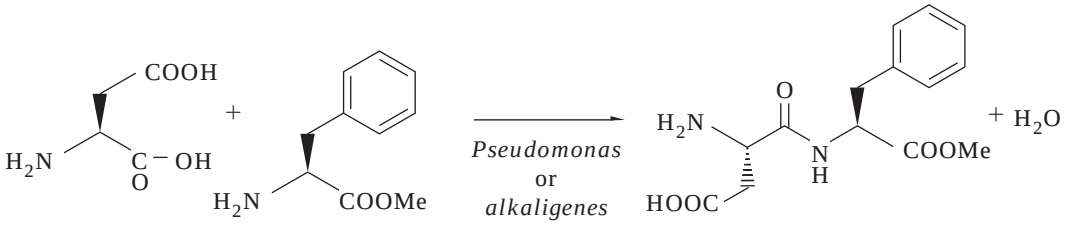
Direct production of APM by condensation of aspartic acid and PheOMe was proposed by TOSOH and Ajinomoto, respectively, where both starting materials are condensed by microbial cells belonging to *Pseudomonas, Alkaligenes*, and so forth (36,37), as shown in Scheme 4. In this process, the reaction seems to be catalyzed by aminopeptidases in the bacterial cells. Although this is a very simple production method, its yield was very low. According to the TOSOH patent application (36), only 0.06 g of APM was produced from 1.0 g of aspartic acid after 16 h of reaction. Because APM is not removed from the reaction mixture as a precipitate under this condition, it seems to be hydrolyzed back by the aminopeptidases. If the product could be separated from the reaction mixture in situ, the production yield will be increased. It is, however, difficult to remove APM from the reaction mixture as a precipitate, because APM is not only highly soluble (similar to the substrates) but also does not form an insoluble addition compound with remaining substrate. No other additive that would form an insoluble addition compound with APM has yet been found. Although separation of APM by electrodialysis was attempted by Snedecor and Hsu (38), sufficient separation yields do not seem to be obtained. Development of an effective method to separate APM from its starting materials and enzymes will be of importance in establishing a feasible production process based on those patent applications.

Another APM production method without N-protection is the coupling of α-AspOR (R indicates a lower alkyl group) and PheOMe by an ester–amino exchange reaction catalyzed by a protease from *Staphylococcus aureus* strain V8 (V8 protease) (39) as shown in Scheme 5. Because this reaction is a dealcoholation reaction, the reaction rate is higher than that of a dehydration reaction, such as the reaction shown in Scheme 4. Additionally, by-products due to misconfiguration, such as PheAspOR or PhePheOMe, are not produced because V8 protease is specific for a peptide bond or ester bond that involves the carbonyl group of amino acids with an acidic side chain (40), for example, aspartic acid at the P1 position (Table 1). However, removal of APM from the reaction system encounters the same difficulties as in the first example of the APM direct production; therefore, a higher production yield cannot be expected. It will also be difficult to produce α-AspOR industrially without formation of β-AspOR and $Asp(OR)_2$, because the reactivities of the two carboxyl group of aspartic acid are not so different. Thus, various barriers exist to establishing a production process based on this method.

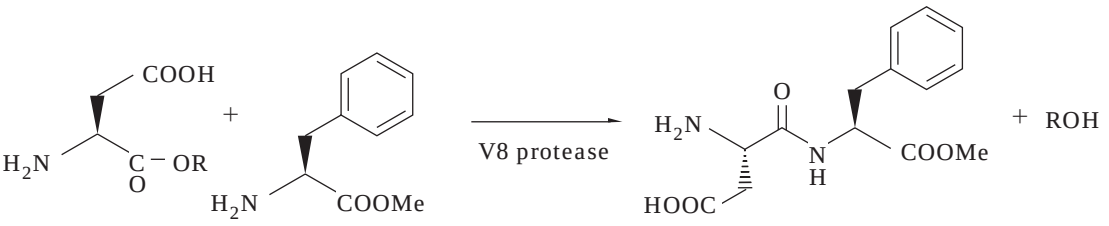
Production from Other Raw Materials

The methods described in the previous section are aimed at coupling of aspartic acid and PheOMe without protection; the methods described in this section do not depend on aspartic acid as the starting material. Because aspartic acid can be produced by ammonia addition to fumaric acid by an ammonia lyase, APM synthesis from fumaric acid, PheOMe, and ammonia as starting materials was proposed by TOSOH (41). *Pseudomonas putida* TS-15001 was found to be catalyzing the coupling of the three substrates simultaneously. Although all details of this reaction system have not been studied, this system seems to consist of two reactions in the cells. At first, aspartic acid is produced from fumarate by the ammonia addition reaction by an ammonia lyase; second, aspartic acid is coupled with PheOMe by an aminopeptidase. The low production yield seems to result from the hydrolysis of APM by the aminopeptidase in the same way as in the first example of the direct production.

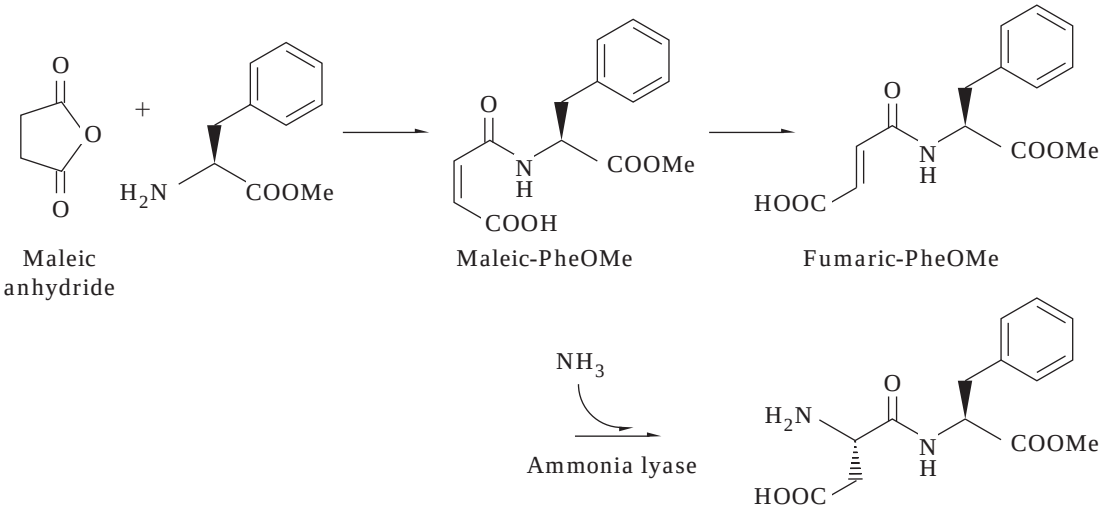
Production of APM from similar substrates was proposed by Showa Denko Corporation (42), where APM was derived from fumaric-PheOMe (FPM) by ammonia addition reaction by certain bacterial cells that produced an ammonia lyase, as shown in Scheme 6. In this proposal, FPM was derived from its *cis-trans* isomer, maleic-PheOMe (MPM) which is more easily produced by the coupling reaction of maleic anhydride and PheOMe than by the direct production of FPM by the coupling reaction of fumaric acid and PheOMe. When fumaric acid and PheOMe are coupled directly, use of a coupling reagent is required, such as oxalyl chloride, thionyl chloride, or carbodiimide. However, the coupling of maleic anhydride and



Scheme 4.



Scheme 5.



Scheme 6.

PheOMe proceeds spontaneously. Fumaric-PheOMe was reportedly produced in good yield by isomerization of MPM (40). Unfortunately, the yield at the step of ammonia addition is too low to be applicable in an industrial APM production process. A new microbial strain exhibiting higher lyase activity has recently been reported by Xu et al. (43), but even with this strain, productivity seems to be insufficient. Further screening of bacteria and improvement of ammonia lyase thus appears to be important.

CONCLUSION

Because APM is an excellent high-intensity sweetener, especially in its taste, demand for it will grow in the future, and APM will gradually become a commodity chemical. The drive for cost reductions will therefore increase. Although alternative production methods have been proposed, their yields are lower than the current, conventional production methods. The current processes will therefore continue to be used while attempts to improve them are ongoing. I focused on the reduction of enzyme costs to advance enzymatic APM production because the condensation by thermolysin is a key step of the production process used by Holland Sweetener Company. Enzymatic APM production processes will be further improved in many ways.

ACKNOWLEDGMENTS

Improvement of thermolysin was supported by Holland Sweetener Company. We are grateful for analysis of mutant thermolysin in production condition by Shin-ichiro Nakamura, Atsuo Aoyama, Yoshikazu Tanaka, and Akira Tokuda. We would like to give special thanks to Kiyotaka Oyama and Ichiro Hiraki for the organization of this work.

BIBLIOGRAPHY

1. R.H. Mazur, in L.D. Steging and L.J. Filer, Jr. ed. *Aspartame Physiology and Biochemistry*, Dekker, New York, 1984, pp. 3–9.
2. M. Tajima, in S. Tsumura and T. Itano ed. *Foods Bio Technology*, Sangyo Chosakai, Tokyo, 1988, pp. 132–134.
3. V. Kashe, in R.J. Beynone and J.S. Bond ed. *Proteolytic Enzymes*, Information Press, Oxford, U.K., 1989, pp. 125–143.
4. Y. Isowa, M. Ohmori, T. Ichikawa, K. Mori, Y. Nonaka, K. Kihara, K. Oyama, H. Saitoh, and S. Nishimura, *Tetrahedron Lett.* **28**, 2611–2612 (1979).
5. U.S. Pat. 3,786,039 (April 15, 1970), Y. Ariyoshi and M. Nagao (to the Ajinomoto Co. Inc.).
6. JP Pat. 2,513,159 (April 30, 1996), T. Takemoto, K. Hisamitsu, and T. Yukawa (to the Ajinomoto Co. Inc.).

7. EP Pat. 0468063B1 (June 5, 1996), J.B. Hill, R.A. Erickson, Y. Gelman, K. Kuang, H.L. Dryden Jr., and M.R. Hohnson (to the Nutrasweet Company).
8. JP Pat. 2,503,592 (June 21, 1988), S. Takahashi and T. Takemoto (to the Ajinomoto Co. Inc.).
9. K. Oyama, S. Irino, and N. Hagi, *Methods Enzymol.* **136**, 503–516 (1987).
10. K. Morihara and H. Tsuzuki, *Eur. J. Biochem.* **15**, 374–380 (1970).
11. M. Ottesen and I. Svendsen, *Methods Enzymol.* **19**, 199–215 (1970).
12. R. Arnon, *Methods Enzymol.* **19**, 226–244 (1970).
13. E.E. Ricks, M.C. Estrada-Valdes, T.L. McLean, and G.A. Iacobucci, *Biotechnol. Prog.* **8**, 197–203 (1992).
14. H. Matsubara, *Methods Enzymol.* **19**, 642–649 (1970).
15. B. Holmquist and B.L. Vallee, *Biochemistry* **15**, 101–107 (1976).
16. K. Aunstrup, in A.H. Rose ed. *Microbial Enzymes and Bioconversions*, Academic Press, New York, 1980, pp. 50–114.
17. D.A. Cowan, K.A. Smolenski, R.M. Daniel, and H.W. Morgan, *Biochem. J.* **247**, 121–133 (1987).
18. S. Hanzawa, T. Hoaki, H.W. Jannasch, and T. Maruyama, *J. Marine Biotechnol.* **4**, 121–126 (1996).
19. T. Coolbear, R.M. Daniel, and H.W. Morgan, *Adv. Biochem. Eng. Biotechnol.* **45**, 58–98 (1995).
20. M. Morikawa, Y. Izawa, N. Rashid, T. Hoaki, and T. Imanala, *Appl. Environ. Microbiol.* **60**, 4559–4566 (1994).
21. Y. Sako, P.C. Croocker, and Y. Ishida, *FEBS Lett.* **415**, 329–334 (1997).
22. K. Oyama, S. Nishimura, Y. Nonaka, K. Kihara, and T. Hashimono, *J. Org. Chem.* **46**, 5341–5242 (1981).
23. K. Nakanishi, A. Takeuchi, and R. Matsuno, *Appl. Microbiol. Biotechnol.* **32**, 633–636 (1990).
24. T. Nagayasu, M. Miyanaga, T. Tanaka, T. Sakiyama, and K. Nakanishi, *Biotechnol. Bioeng.* **43**, 1118–1123 (1994).
25. T. Imanaka, M. Shibazaki, and M. Takagi, *Nature* **324**, 695–697 (1986).
26. F. Hardy, G. Vriend, O.R. Veltman, B. van der Vinne, G. Venema and V.G.H. Eijsink, *FEBS Lett.* **317**, 89–92 (1993).
27. EP Pat. 0616033A1 (May 17, 1993), H. Nagao, T. Yoneya, T. Miyake, A. Aoyama, K. Kai, S. Kidokoro, Y. Miki, K. Endo, and A. Wada (to the Holland Sweetener Company and Sagami Chemical Research Center).
28. S. Kidokoro, K. Endo, Y. Miki, A. Wada, T. Yoneya, A. Aoyama, K. Kai, H. Nagao, T. Miyake, S. Hanzawa, K. Misawa, S. Ooe, T. Hanya, and H. Yagame, *Seibutu butsuri* **35**, supplement S19 (1995).
29. S. Endo, *J. Ferment. Technol.* **40**, 346–353 (1962).
30. S. Toma, S. Campagnol, E.D. Gregoriis, R. Gianna, I. Margrit, M. Zamai, and G. Grand, *Protein Eng.* **2**, 359–364 (1989).
31. W.R. Kester and B.W. Matthews, *J. Biol. Chem.* **252**, 7704–7710 (1977).
32. JP Pat. 9,402,050 (June 15, 1995), Y. Tanaka, T. Miyake, S. Hanzawa, S. Ohe, S. Kidokoro, Y. Miki, K. Endo, and A. Wada (to the Holland Sweetener Company and Sagami Chemical Research Center).
33. JP Pat. 0,8131,170 (May 28, 1996), T. Harada, Y. Kunizawa, K. Oyama, C.M.Y. Smith, S. Takasuga, and W.J.J. van den Tweel (to the Holland Sweetener Company).
34. K. Oyama, K. Kihara, and Y. Nonaka, *J. Chem. Soc., Perkins Trans.* 2 356–360 (1981).
35. EP Pat. 0768384A1 (September 23, 1996), T. Harada, S. Irino, Y. Kunisawa, and K. Oyama (to Holland Sweetener Company).
36. JP Pat. 9,014,040 (April 5, 1990), T. Harada, K. Takemoto, and T. Igarashi (to the TOSOH Corp.).
37. JP Pat. 9,015,196 (April 11, 1990), K. Yokoseki and K. Kubota (to the Ajinomoto Co. Ltd.).
38. U.S. Pat. 5,055,399 (Oct. 8, 1991), B.R. Snedecor and C.C. Hsu (to the Genentech Inc.).
39. EP Pat. 0269390B1 (Nov. 21, 1986), M.G. Boston and A.J. Poulose (to the Genencor Int. Inc.).
40. G.R. Drapeau, *Methods Enzymol.* **45**, 469–475 (1976).
41. JP Pat. 912,520 (Jan. 16, 1991), T. Harada, H. Takemoto, and I. Igarashi (to the TOSOH Corp.).
42. JP Pat. 63,185,395 (July 30, 1988), A. Nakayama and Y. Torigoe (to the Showa Denko Corp.).
43. J. Xu, W. Wu, S. Jin, and W. Yao, *J. China Pharm. Univ.* **25**, 53–55 (1994).

See also AMINO ACIDS, PRODUCTION PROCESSES.

ASPARTIC ACID. See AMINO ACIDS, PRODUCTION PROCESSES; L-ASPARTIC ACID.

L-ASPARTIC ACID

MIKI KOBAYASHI
MASATO TERASAWA
HIDEAKI YUKAWA
Mitsubishi Chemical Corporation
Ibaraki, Japan

KEY WORDS

Aspartase
L-Aspartic Acid
Coryneform bacteria
Isomerase
Ultrafiltration

OUTLINE

INTRODUCTION

L-Aspartic acid is widely used as a food additive as well as an ingredient of infusion solutions and medicines. In addition, a dipeptide, L-aspartyl-L-phenylalanyl methyl ester, is used as a synthetic low-calorie sweetener in many countries. Moreover, studies for production and application of L-aspartic acid derivatives have been carried out by many companies in Europe, America, and Japan because L-aspartic acid is also chemically reactive. L-Aspartic acid derivatives are hoped to function as a kind of biodegradable precursor for environment friendly products (Fig. 1). It is presumed that each of these would be produced at the 1,000-kiloton level in the near future. However, the high price of L-aspartic acid produced using the existing process has been a high hurdle in expansion of its use. Therefore, we tried to develop a novel process to produce L-aspartic acid economically.

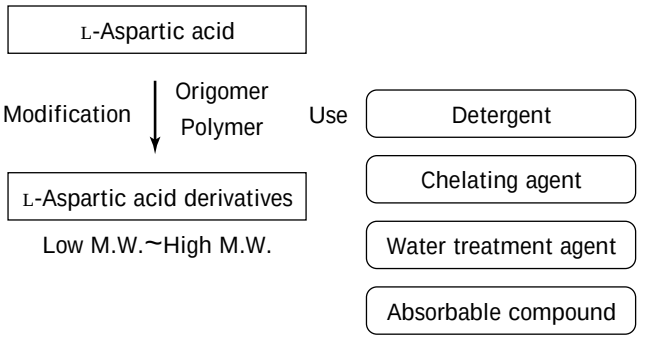


Figure 1. Uses of aspartic acid.

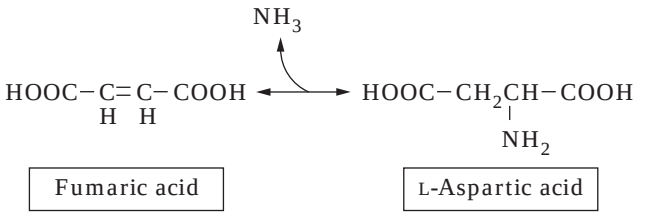


Figure 2. Pathway of L-aspartic acid formation from fumaric acid.

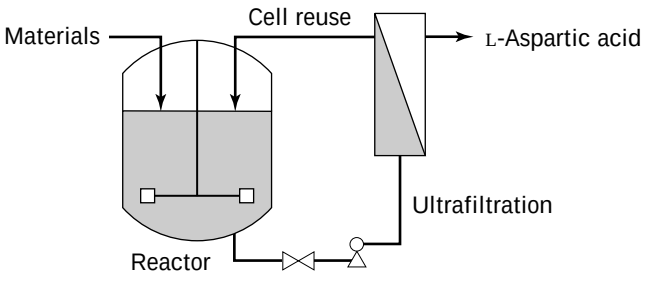


Figure 3. Membrane reactor system

CONVENTIONAL METHODS

Since Quastel and Woolf (1) observed the reversible formation of L-aspartic acid from fumaric acid and ammonium ion by a cell suspension of *Escherichia coli*, many reports have been published on aspartase (E.C.4.3.1.1), which catalyzes this reaction. This enzyme degrades L-aspartic acid into fumaric acid and ammonium ion under physiological conditions, but the reverse reaction proceeds in high concentrations of ammonia (Fig. 2).

High-yield conversion of L-aspartic acid from fumaric acid and ammonia was first reported by Sumiki et al. (2) in 1928. This method employed resting yeast cells to produce L-aspartic acid. Fermentation methods using medium containing fumaric acid, and enzymatic methods using bacterial aspartase activity have also been reported (3–5).

A process for industrial production of L-aspartic acid was first established in 1973; bacterial cells having aspartase activity were immobilized in a gel matrix to stabilize the enzyme activity. A continuous operation was also reported (6–9). In the immobilized-enzyme process, the enzyme cost was reduced compared with that of a batch process.

However, in the immobilized-cell process, cells had to be immobilized and packed into reactor columns under aseptic conditions. Contamination-free operations were needed, which resulted in additional costs. Also, it did not seem to be possible to use the reactor columns for conventional purposes because different specifications were required according to each individual reaction mode or the chemical properties of substrate and product.

PROPERTIES OF A NOVEL BIOREACTOR SYSTEM

As already mentioned, there are several biological and engineering disadvantages that need to be solved in conventional production methods. To overcome these disadvantages, we have developed a novel bioreactor process, the membrane reactor system, using microbial cells without artificial immobilization for the enzymatic reaction. Coryneform bacterium MJ-233, which is employed in this process, is characterized by nonlytic properties under nongrowing conditions. The cost of the enzyme is low because the bacterial cells can be reused repeatedly for enzymatic reactions without immobilization. Intracellular components such as protein and nucleic acid do not leak out, so the purity of the products is very high. From the engineering stand point, the ultrafiltration membrane system is employed for recycling bacterial cells (Fig. 3). The advantages of the ultrafiltration process are as follows:

1. Damage of bacterial cells by compression is reduced compared with centrifugation.
2. Capital outlay and running costs are relatively low.
3. Operation and maintenance are easier.

On the basis of these properties, we established an economical new bioprocess for the production of L-aspartic acid (10–13).

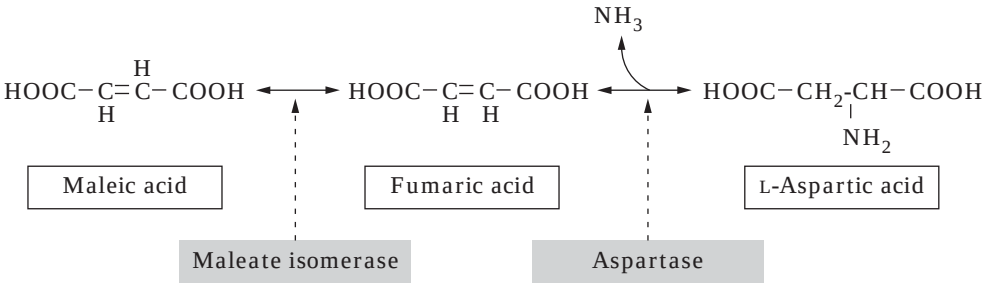


Figure 4. L-Aspartic acid formation from maleic acid

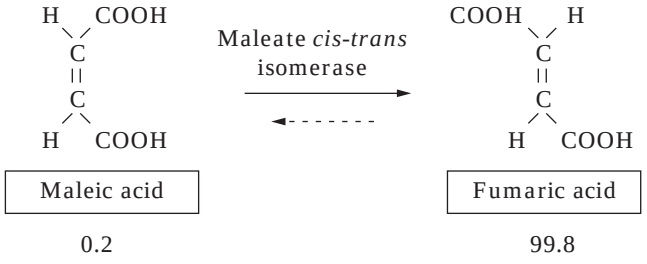


Figure 5. Equilibrium of the maleate isomerase reaction

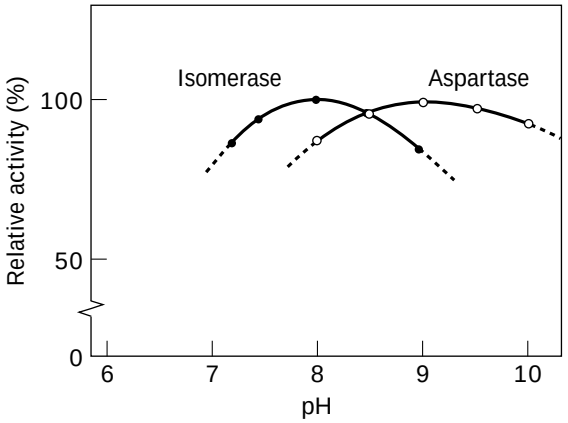


Figure 6. The relationship between pH and relative enzyme activity.

TECHNICAL COMPONENTS

Enzymes for L-Aspartic Acid Production

We tried to develop a new bioprocess for L-aspartic acid production using maleic acid as a starting material. Maleate isomerase (E.C.5.2.1.1), which isomerizes maleic acid to fumaric acid, and aspartase (E.C.4.3.1.1), which converts fumalic acid and ammonium ion to L-aspartic acid, were used for L-aspartic acid production from maleic acid (Fig. 4). The equilibrium of the isomerase reaction strongly favors fumaric acid formation (Fig. 5). This results in almost complete conversion of maleic acid to fumaric acid, so it is very useful for L-aspartic acid formation. On the other hand, the equilibrium of the aspartase reaction also favors L-aspartic acid formation under high concentrations of ammonium ion. Therefore, the conversion yield from maleic acid to L-aspartic acid can reach 99%.

Conditions for L-Aspartic Acid Production

In order to establish a one-pot reactor system using two enzymes, maleate isomerase and aspartase, we analyzed the properties of each enzyme, including pH profile (Fig. 6). Consequently, we could determine the potential for coexistence of two enzymes in a one-pot reactor even in the high concentrations of ammonium ion that are an essential condition for L-aspartic acid formation by aspartase.

Improvement of L-Aspartic Acid Productivity

In order to improve the productivity of the process, genetic engineering techniques were applied to coryneform bacterial strain MJ-233. An efficient transformation system in

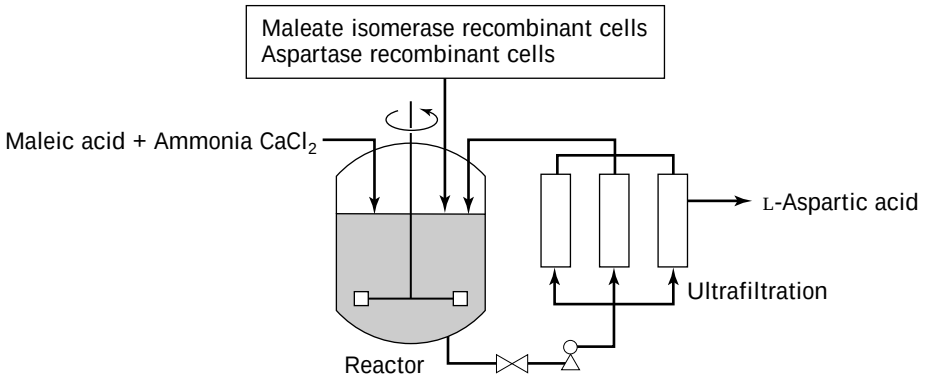


Figure 7. The L-Aspartic acid production process

coryneform bacteria using electroporation (14–16) and other techniques has also been developed (17,18). Both maleate isomerase and aspartase can be overproduced in coryneform bacterial strain MJ-233. These recombinant cells can therefore be employed for L-aspartic acid production from maleic acid as a starting raw material.

Development of a Stable Host–Vector System. One of the important factors in the application of genetic engineering techniques for industrial use is the development of a stable host–vector system. For this purpose, we isolated various plasmids from coryneform bacteria and identified a genetic element that is important for stable plasmid maintenance (19). We then constructed a highly stable host–vector system.

Cloning and Expression of the Maleate Isomerase and Aspartase Genes. Isolation of both the maleate isomerase gene and the aspartase gene from a genomic library was achieved by hybridization with DNA probes synthesized on the basis of the N-terminal amino acid sequences of each purified enzyme. Using this methodology, we cloned maleate isomerase (20) and aspartase (21) genes from the *Alcaligenas faecalis* and coryneform bacterial strain MJ-233, respectively, and achieved high-level expression in recombinant cells of coryneform bacterial strain MJ-233.

Membrane Reactor System

Coryneform bacterial strain MJ-233 exhibited nonlytic properties under nongrowing conditions and, therefore, showed no leakage of intracellular macromolecules, as already mentioned. Taking advantage of these physiological characteristics, an ultrafiltration system was adopted for the separation and recycling of cells from the reaction mixture. Employment of the ultrafiltration system greatly simplified the purification step (Fig. 7). In this ultrafiltration system, the filtration capacity can be easily modified by varying the number of modules. A hollow-fiber tubular-type membrane made of polysulfon, which can withstand high temperature and alkaline pH, was chosen for our process. We can operate this system at high temperature and alkaline pH, allowing us to carry out a contamination-free process.

L-Aspartic acid can be produced at high purity using this membrane reactor system and, therefore, more economically compared with conventional fermentation processes.

CONCLUSION

The host–vector system in coryneform bacterial strain MJ-233 was developed and L-aspartic acid can now be economically produced at high yield using recombinant intact cells in a membrane reactor system. We hope that our method can provide the quantities of L-aspartic acid needed to meet the increasing demands for biodegradable chemicals in an expanding market.

BIBLIOGRAPHY

1. J.H. Quastel and B. Woolf, *Biochem. J.* **20**, 545–555 (1926).
2. Y. Sumiki, *Bull. Jpn. Soc. Ferment.* **23**, 33–37 (1928).
3. M. Kisumi, Y. Ashikaga, and I. Chibata, *Bull. Agric. Chem. Society. Soc. Jpn.* **24**, 296–305 (1960).
4. S. Kinoshita, K. Nakayama, and S. Kitada. *Bull. Jpn. Soc. Ferment.* **16**, 517–520 (1958).
5. K. Kitahara, S. Fukui, and M. Misawa, *Bull. Agric. Chem. Soc. Jpn.* **34**, 44–48 (1960).
6. T. Tosa, T. Sato, T. Mori, Y. Matsuo, and I. Chibata, *Biotechnol. Bioeng.* **15**, 69–84 (1973).
7. I. Chibata, T. Tosa, and T. Sato, *Appl. Microbiol.* **27**, 878–885 (1974).
8. T. Sato, T. Mori, T. Tosa, I. Chibata, M. Furui, Y. Yamashita, and A. Sumi, *Biotechnol. Bioeng.* **17**, 1797–1804 (1975).
9. Y. Nishida, T. Sato, T. Tosa, and I. Chibata, *Enzyme Microb. Technol.* **1**, 95–99 (1979).
10. H. Yukawa, S. Yamada, T. Nara, M. Terasawa, and Y. Takayama, *Nippon Nogeikagaku Kaishi* **59**, 31–37 (1985).
11. H. Yukawa, T. Nara, M. Terasawa, and Y. Takayama, *Nippon Nogeikagaku Kaishi* **59**, 279–285 (1985).
12. H. Yukawa and M. Terasawa, *Nippon Nogeikagaku Kaishi*, **61**, 1279–1284 (1987).
13. M. Terasawa, H. Yukawa, and Y. Takayama, *Process Biochem.* **20**, 124–128 (1985).
14. Y. Satoh, K. Hatakeyama, K. Kohama, M. Kobayashi, Y. Kurusu, and H. Yukawa, *J. Ind. Microbiol.* **5**, 159–166 (1990).
15. Y. Kurusu, M. Kainuma, M. Inui, Y. Satoh, and H. Yukawa, *Agric. Biol. Chem.* **54**, 443–447 (1990).
16. A.A. Vertes, M. Inui, M. Kobayashi, Y. Kurusu, and H. Yukawa, *Res. Microbiol.* **144**, 181–185 (1993).
17. A.A. Vertes, M. Inui, M. Kobayashi, Y. Kurusu, and H. Yukawa, *Mol. Microbiol.* **11**, 739–746 (1994).
18. A.A. Vertes, K. Hatakeyama, M. Inui, M. Kobayashi, Y. Kurusu, and H. Yukawa, *Biosci. Biotechnol. Biochem.* **57**, 2036–2038 (1993).
19. Y. Kurusu, Y. Satoh, M. Inui, K. Kohama, M. Kobayashi, M. Terasawa, and H. Yukawa, *Appl. Environ. Microbiol.* **57**, 759–764 (1991).
20. K. Hatakeyama, Y. Asai, Y. Uchida, M. Kobayashi, M. Terasawa, and H. Yukawa, *Biochem. Biophys. Res. Commun.* **239**, 74–79 (1997).
21. Y. Asai, M. Inui, A.A. Vertes, M. Kobayashi, and H. Yukawa, *Gene* **158**, 87–90 (1995).

ASPERGILLUS

J.W. BENNETT
Tulane University
New Orleans, Louisiana

M.A. KLICH
U.S. Department of Agriculture
New Orleans, Louisiana

KEY WORDS

Aflatoxin
Amylases
Expression host
Filamentous fungi
Genetic model
Koji
Mycoses
Mycotoxins
Organic acids
Secondary metabolites

OUTLINE

INTRODUCTION

Biodiversity experts estimate that there may be 1.5 million species of fungi on earth, of which only approximately 100,000 have been described. Of these described species, the approximately 200 taxa that comprise the genus *Aspergillus* are among the most commonly encountered molds. Adapted to many habitats, exhibiting a wide range of metabolic activities, and capable of the production of astronomical numbers of spores, *Aspergillus* species are found throughout the biosphere. A few species are economically important as biotechnological resources, whereas others have gained notoriety for their negative impact on human health and commerce. For example, some aspergilli are used in industrial fermentations to produce enzymes, pharmaceuticals, and bulk chemicals; others are important components of Asian food fermentations; still others are major agents of biodeterioration; and a few common species cause allergenic, pathogenic, and toxigenic diseases in humans and other animals.

TAXONOMY

The filamentous fungi that comprise the genus *Aspergillus* reproduce mitotically through the production of spores borne on a specialized structure called the conidiophore. The morphology of this spore-bearing structure, with its distinctive foot cell, long stalk, swollen apex, and protuberant phialides, is the defining feature of the genus. It was first described in 1729 by Micheli, an Italian priest, who was reminded of a Roman Catholic device used to sprinkle holy water, the aspergillum.

All members of the genus *Aspergillus* produce an aspergillum-like conidiophore (Fig. 1). (The special vocabulary regularly encountered in *Aspergillus* taxonomy is summarized in Table 1.) The conidiophore wall is usually thicker and more hydrophobic than that of the other hyphal cells. The stipe of the conidiophore attaches to the vegetative mycelium in a T- or L-shaped base that has been traditionally called a foot cell, even though it is not a separate cell. The presence of the foot cell is an important diagnostic criterion of the genus. The apex of the stipe swells into the vesicle. Spore-producing phialides are borne directly (uniseriate) or indirectly on metulae (biseriate) on the surface of the vesicle, another important characteristic for distinguishing species within the genus. The phialides are sporogenous. Repeated mitotic divisions in the phialide nucleus lead to the formation of a chain of uninucleate or multinucleate conidiospores, also called conidia. The phialides also extrude primary wall layers and pigments that form part of the mature spore. Conidiospores are spherical or elliptical, smooth or echinulate, extremely hydrophobic, and easily airborne when mature.

Charles Thom (1872–1956), most famous for his work on penicillin and the other great mold genus, *Penicillium*, was also a major figure in the development of *Aspergillus* taxonomy. With Margaret Church in 1926 (1) and Kenneth Raper in 1945 (2) he produced the first two complete monographic treatments. Kenneth Raper went on to expand this work and with Dorothy Fennell published *The Genus Aspergillus* in 1965 (3), still considered by many to be the most authoritative source of *Aspergillus* taxonomy. Highly sensitive to substrate, temperature, moisture level, and other environmental variables, potential morphological plasticity is controlled in modern taxonomic schemes by the use of standard media and culture conditions. Pure cultures are grown on standardized agar media for specific lengths of time with defined temperature, light, and other environmental parameters. Species within *Aspergillus* are distinguished by cultural characters such as colony color, growth rate, and texture and by microscopic traits such as phialides and metulae, shape of the vesicle, spore size, and spore ornamentation. The ability to produce certain pigments, sclerotia, or cleistothecia is highly dependent on culture conditions, but can also be useful in distinguishing species. Modern taxonomists place increasing emphasis on secondary metabolite profiles, isoenzyme electrophoresis, and molecular genetic characters such as ribosomal DNA sequences to supplement the traditional characters.

After adequate characterization, identification of species depends on careful comparison with published descriptions. The last monograph of the entire genus was Raper and Fennell's *The Genus Aspergillus* (3), but several more recent books provide identification systems for the more common species (for example, see, Klich and Pitt [4]) Medically important species are described in Kwon-Chung and Bennett (5).

Some aspergilli possess a sexual stage. All the sexual stages result in the formation of ascospores inside of cleistothecia and are classified formally in the Ascomycete family Trichocomaceae. According to the rules of botanical nomenclature, which govern the naming of all fungi, the sexual (teleomorphic or perfect) state requires its own name. Because *Aspergillus* by definition refers only to the asexual (anamorphic or imperfect) state, many species within the genus have two names. For example, the asexual species *Aspergillus tetrazonus* produces a teleomorphic state called *Emericella quadrilineata*, and the teleomorph of the well-known genetic model *Aspergillus nidulans* is called *Emericella nidulans*. This dual nomenclature mystifies workers not trained in the rules of botanical nomen-

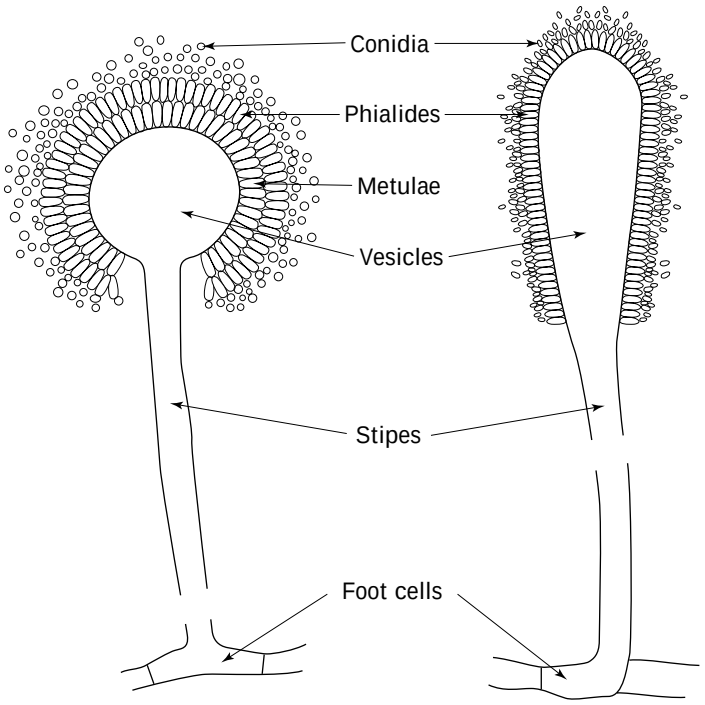


Figure 1. Conidiophores of *A. flavus* (left) and *A. clavatus* (right).

Table 1. Glossary of *Aspergillus* Taxonomy

Term	Definition
Anamorph	The nonsexual form of a fungus, loosely synonymous with *imperfect*
Ascus	A baglike sexual structure containing ascospores
Cleistothecium	A closed ascus-containing structure; the type of ascocarp (ascoma) found in the sexual stages of *Aspergillus*
Conidium (*pl.* conidia)	A thin-walled asexual spore borne on a conidiophore (also called *conidiospore*)
Foot cell	Lower portion of conidiophore, not a separate cell
Heterokaryon	A mycelium containing genetically dissimilar nuclei
Metullae	Sterile cells subtending the phialides in biseriate species; also called *secondary sterigmata*
Phialides	Conidiogenous cells; also called *primary sterigmata*
Sclerotium	Multicellular structure composed of hardened, thick-walled hyphae containing no spores inside or outside
Stipe	Stalk of the conidiophore
Teleomorph	The sexual stage of the fungal life cycle, loosely synonymous with *perfect stage*
Vesicle	The swollen apex of the conidiophore

clature. Moreover, Thom, Church, Raper, and Fennell all resisted these rules and applied the generic name *Aspergillus* to both the conidial and ascosporic stage of these organisms. Much common practice is also blind to the rules of botanical nomenclature.

Revisions of the Botanical Code in 1981 have added another layer of complexity by endangering many well-known specific epithets. Strict constructionists renamed the asexual state of *A. nidulans* as *Aspergillus nidulellus* (6). Rules of priority also endanger the name *Aspergillus niger*, one of the most widely used industrial species. Although the change to *A. nidulellus* is not followed, and although most workers continue to use the name *A. niger*, it demonstrates the potential for nomenclatural confusion in both the traditional literature and with on-line data retrieval. For a more detailed description of the nomenclatural conundrum, see Bennett (7); the contemporary rules of taxonomy are outlined in Samson and Pitt (6,8). A list of the Ascomycete genera (e.g., *Emericella, Eurotium*) with *Aspergillus* anamorphs are presented in Samson (9). Additionally, a list of *Aspergillus* species names in current use has been published in an attempt to help stabilize nomenclature in the genus (10).

DISTRIBUTION AND PHYSIOLOGY

Members of the genus *Aspergillus* are found almost everywhere. They are truly ubiquitous and have been isolated from penguin dung in the Antarctic, binocular lenses in the tropics, and spoiled foodstuffs all over the planet. Aspergilli can utilize an incredible variety of substrates, including plant debris, foods and feeds, fabrics, feathers, leather, and dung. Aspergilli have been associated with the biodeterioration of kerosene, paper, pesticides, plasticisers, and rubber and have been isolated from unlikely sources that range from biblical scrolls to sauna bath boards. They often grow on the surface of humid walls in cellars and stables, and their spores are found in high concentrations in attics,

crawl spaces, air conditioning ducts, house dust, and insulation. Although the majority of species are saprophytic, the genus also includes opportunistic pathogens of mammals, insects, and plants.

The major natural habitats for aspergilli are soil and decaying vegetation. Water and temperature are the major factors influencing mycelial growth. Within the soil environment, aspergilli occur most frequently in subtropical/warm temperate zones; however, certain species such as *Aspergillus niveus*, *A. nidulans*, and *Aspergillus tamarii* are more common in tropical zones. Most aspergilli prefer moist habitats, although several species are xerophilic. Whatever the optimal conditions for growth, all species of *Aspergillus* form abundant conidiospores that are resistant to harsh environments. These propagules bridge survival until hospitable conditions for germination prevail.

Aspergillus colonies are visible macroscopically as black, green, yellow, or brown mold. The colony color reflects the spore color of the particular species. Pigment production is also influenced by trace elements. In fact, the color of *A. niger* spores was once used as a bioassay to detect copper.

As with all filamentous fungi, the hyphae grow by apical extension, allowing for rapid exploitation of new habitats and facilitating the colonization of solid and semisolid substrates. The ramifying hyphal filaments are profusely branched and have a high surface-to-volume ratio in intimate contact with their substrate. As mycelia colonize, they extrude enzymes that break down complex carbon sources such as cellulose, hemicellulose, pectin, and starch, not only changing their substrate, but literally becoming part of it. Simultaneously, mycelia influence the growth of other organisms, especially bacteria, and create complex degradative consortia.

Because of their robust physiology, aspergilli are easy to grow in the laboratory and can be cultured with simple carbon and nitrogen sources, mineral salts, water, and oxygen. Mutations blocked in various catabolic and anabolic pathways are also easily isolated. Biochemical and molecular genetic analyses in *A. nidulans* have revealed considerable detail about the physiology and genetics of these processes. In addition, the weedy nature of these molds means they often cause trouble as common contaminants in tissue culture and bacteriology laboratories. Techniques for handling aspergilli as well as common media for culturing species are given by Raper and Fennell (3) and Onions et al. (11).

GENETICS

Aspergillus nidulans is an excellent model organism for genetics. If the work by Raper and Fennell (3) is the bible of *Aspergillus* taxonomy, then the paper of Pontecorvo et al. (12) is often referred to as the old testament. This paper "contained not only the foundations of *Aspergillus* genetics, but most of the superstructure as well" (13, p. 3).

The parasexual cycle was first elucidated in *A. nidulans*. The first step of the parasexual cycle is the production of heterokaryons, a mycelium with genetically dissimilar nuclei, usually forced between different auxotrophic mutants. Somatic diploids are formed that are identified by their ability to grow on minimal media. Parasexual recombination occurs through mitotic crossing over and vegetative haploidization. After the parasexual cycle was elucidated in *A. nidulans*, it was found in a large number of other fungi, allowing formal genetic analysis in anamorphic species. Because most species of *Aspergillus* do not form a sexual stage, these artificially produced diploids provide "an alternative to sex." For example, early genetic studies on the aflatoxin-producing species *Aspergillus flavus* and *Aspergillus parasiticus*, both of which lack a sexual stage, were conducted by means of the parasexual cycle. Even more importantly, parasexual cycle genetics was adopted to mammalian tissue culture, where it was usually referred to as somatic cell genetics. For several decades before the recombinant DNA revolution, somatic cell genetics was the major method for mapping human genes.

Aspergillus nidulans remains an important genetic model in its own right. The steps of conidiation display morphological phenotypes amenable to genetic dissection and have been exploited as a model for developmental biology through molecular analysis of successive stages of the defined developmental sequence. The ability of the species to degrade all kinds of natural materials and to utilize a wide range of inorganic nitrogen sources has also been exploited for the genetic dissection of carbon metabolism, alcohol metabolism, purine degradation, proline utilization, sulfur metabolism, and many other metabolic pathways. The genetic regulation of nitrogen metabolite repression and carbon catabolite repression, both tightly controlled by induction of appropriate enzymes, has also received considerable attention. In addition, *A. nidulans* has been a valuable model for the study of mitochondrial and population genetics, for penicillin biosynthesis, and for the molecular biology of gene expression. The formal genetics of *Aspergillus* has been reviewed by Smith and Pateman (14) and Martinelli and Kinghorn (15).

The small size of fungal chromosomes, long a disadvantage for cytogenetic analysis, has become an advantage for doing electrophoretic karyotyping by means of pulsed gel electrophoresis. Electrophoretic karyotypes allow workers to bypass conventional genetic and cytological maps. Any DNA sequence for which a probe is available can be mapped to a chromosome by Southern analysis. The fungal genomics community recently selected *A. nidulans* as a premier target for genomic analysis (16).

The genetic versatility and well-elucidated genetic systems have made *Aspergillus* a suitable genus for industrial application of the latest recombinant DNA techniques to large-scale fermentation processes. These industrial processes are discussed below.

ECONOMICALLY IMPORTANT PRODUCTS

Aspergilli have been used since prehistoric times as components of several Asian food fermentations. Members of the genus have exceptional protein- and metabolite-secreting capacity, and these physiological activities have been adapted to human enterprise. From a contemporary perspective, the most important economic impact is the

manufacture of fungal products such as enzymes, organic acids, and bioactive secondary metabolites.

Modern fermenters are sterilizable and operate under aseptic conditions. They are frequently controlled by computers that monitor temperature, agitation, pH, dissolved oxygen, and other parameters. In general, wild-type strains do not work well in industrial fermentations. Yield of the product is low, and the strain usually is not adapted to the conditions of large-scale fermentation. Genetic strain improvement has traditionally been conducted by means of brute strength mutagenesis and screening. Although repeated rounds of mutagenesis and screening are generally successful, they are also laborious. Recombinant DNA techniques have revolutionized strain improvement programs. *Aspergillus* gene expression systems out perform and are cheaper and simpler than analogous animal or insect systems.

Organic Acids

Citric acid is the most important acid produced by fungi and the only organic acid produced almost exclusively through fermentation. Widely used in the food and beverage industry, its pleasant acid taste enhances products such as soft drinks, fruit juices, jams, jellies, candies, prepared desserts, and frozen fruits. Citrates are also efficient buffering and chelating agents and are used by the cosmetics industry and in blood transfusion products, effervescent tablets, detergent manufacturing, electroplating, printing, inks, leather tanning, and a host of other applications. With increasing emphasis on nonpolluting chemical products, the market for nature's acidulant is increasing.

Many molds and yeasts accumulate large amounts of citric acid from glucose. The biochemical basis of this metabolic overproduction has been studied for a century. Names associated with early research include C. Wehmer in Germany, and J.N. Currie and C. Thom in the United States. The first commercial fermentation was perfected by Currie at Pfizer and Company in Brooklyn, New York, during the 1920s and involved surface fermentations of *A. niger*. This species is still used in many industrial processes worldwide, although submerged fermentations have generally replaced surface and solid-state systems. Production strains are selected not only for their high yield of citric acid, but also for their ability to use cheap raw materials as substrates. Optimal yields of citric acid are associated with low pH (below 2.5), high carbohydrate, and manganese deficiency.

Although a great deal is known about the environmental cues involved in maximizing citric acid production, the exact biochemical basis and the crucial regulatory control points are the subject of continuing research (17,18).

Mycelium is removed by filtering or centrifugation; citric acid is then precipitated. Subsequent recovery steps can include treatment with activated carbon, cation and anion exchanges, and final crystallization. The whole process is highly optimized, with *A. niger* capable of producing about 90% of the theoretical yield from a carbon source.

Other organic acids produced by *Aspergillus* fermentation are listed in Table 2. None of these is as economically important as citric acid.

Secondary Metabolites

The families of low molecular weight compounds that have no obvious role in the life cycle of producing species, often produced after active growth has ceased, are called secondary metabolites. Both biologically and chemically, secondary metabolites are extremely diverse; however, they arise from just a few biosynthetic pathways based in primary metabolism. Acetate-derived polyketides and amino acid–derived compounds are the most important biosynthetic routes in filamentous fungi, including aspergilli. Members of the genus *Aspergillus* are prolific producers of these natural products, and the structures of a large number of secondary metabolites from *Aspergillus* have been elucidated by chemists (19,20).

In general, the most famous natural products are bioactive and are classified on the basis of this activity as antibiotics, chemotherapeutic agents, toxins, and so forth (21). The single best-known fungal secondary metabolite is the antibiotic penicillin. Although several species of *Aspergillus* make penicillin, and although *A. nidulans* has been developed as a model system for studying its biosynthesis, high-yielding industrial fermentations all use strains of *Penicillium*.

Economically speaking, toxic secondary metabolites of *Aspergillus* are more important than the medically useful ones. Aflatoxins and sterigmatocystins, for example, cause millions of dollars of damage in agriculture each year (see "Mycotoxins"). Currently, the most important pharmacological agent is lovastatin (mevinolin) a potent inhibitor of cholesterol synthesis, derived from *Aspergillus terreus*. Lovastatin is a polyketide. Several amino acid–derived metabolites also show pharmacological activity. For example, echinocandin B (cilofungin), produced by *A. nidulans* and *Aspergillus rugulovalvus* (also called *E. rugulosis*), is under development as an antifungal agent (Table 2).

Food Production

Foods fermented with microorganisms have interesting flavors and improved digestibility. In Western culture, the majority of food fermentations involve yeasts (e.g., bread and wine) or bacteria (e.g., cheese and pickles). Aspergilli play only minor roles in Western cuisine (e.g., ripening salamis and country-cured hams). This is in sharp contrast to Asian food traditions, where filamentous fungi have been used for millennia, especially in complex two-stage fermentations. The best known of these processes include shoyu (soy sauce), sake (rice wine), and miso (fermented soybean paste). The initial step in each case is the production of a *koji*, a Japanese word that roughly translates as "bloom of mold." To produce a koji, *Aspergillus oryzae* or the closely related species *A. sojae* is grown on steamed rice or other cereal and incubated in warm, humid conditions. The hyphae infiltrate the grain, secreting hydrolytic enzymes that partially degrade the substrate, producing the koji, a fragrant, crumbly mass of mold and substrate. Perhaps because of the paucity of Western fungal fermentations, the English language does not have a word that similarly denotes a substance thus composed of living mycelia, extracellular enzymes, and partially degraded substrate. This mold–substrate mixture has different proper-

Table 2. Economically Useful Low Molecular Weight Products of *Aspergillus*

Product	Producing species	Use
Organic Acids		
Citric acid	*A. niger*	Flavoring agency, antioxidant, chelating and cleaning agent, component of effervescent powders
Gluconic acid	*A. niger*	Manufacturing of toothphaste
Itaconic acid	*A. terreus*	Copolymer in resin manufacture
Pharmacologically active secondary metabolites		
Asperillomarasmine	*A. oryzae*	Angiotensin-converting enzyme inhibitor
Asperlicins	*A. alliaceus*	Cholecystokinin antagonists
Lovastatin	*A. terreus*	Cholesterol inhibitor
Echinocandin	*A. nidulans* / *A. rugulovalvus*	Antifungal agent

ties than either alone and constitutes a form of matter not captured by the term solid-state fermentation.

Kojis are classified according to the degree of mycelial development and sporulation, their intended use, and the cereal substrate. Kojis are usually used as inocula for further fermentations. For soy sauce, the koji is mixed with soybeans and wheat. After fungal amylases and proteases have partially degraded the mixture, a salt solution is added. An anaerobic yeast–lactobacilli fermentation ensues; the final pressed liquid is clarified as soy sauce. For sake, a koji inoculum is mixed with steamed rice and water and incubated for 3 to 4 days in the moto stage. As the fungal enzymes partially saccharify the rice, a predictable change in the microbial flora takes place, with yeasts becoming the predominant microorganisms, to yield the moroni stage. After appropriate incubation, yeast metabolic activities yield alcohol. Miso is a fermented paste with the consistency of peanut butter, often made in Japanese homes. Widely used as a soup base, miso varies according to substrate (rice, barley, soybeans, mixtures), length of fermentation, amount of salt, etc. Similar products, based solely on fermented soybeans, are known as *chiang* in China, *jang* in Korea, *tao-jo* in Indonesian and Thailand, and *tao-tsi* in the Philippines (22,23).

Aspergillus oryzae and *A. sojae* are close relatives of the toxigenic species *A. flavus* and *A. parasiticus*. These domesticated species do not produce toxins; *A. oryzae* is on the GRAS list (Generally Regarded as Safe) by the U.S. Food and Drug Administration.

The first industrial enzyme patented in the United States was an amylase mixture purified from a koji by the early Japanese-American biotechnologist Jokichi Takamine in 1894 (7). Contemporary scientists have subsequently identified more than 50 different enzymes from koji.

Enzymes

Enzymes are organic catalysts. Long before they were purified by scientists, their properties had been used in baking, brewing, tanning, and other processes. Because of their high specificity, their biodegradability, and their ability to work in watery solutions at moderate temperatures, enzymes are becoming increasingly important in today's environmentally conscious industries. In many processes, enzymes have replaced the strong chemicals and high temperatures formerly used in commercial catalysis.

The major industrial enzymes from fungi are hydrolytic, and the major producing species are *A. oryzae* and *A. niger*. Amylases and glucoamylases are used for turning starches into sugars and oligosaccharides. Pectinases are applied in fruit juice clarification. *Aspergillus* proteases are used in bread making and in chill-proofing beer, and a thermostable phytase is used widely in the animal feed industry as an additive. Representative commercial enzymes produced by *Aspergillus* are listed in Table 3.

With the advent of modern molecular techniques, the genes for many native *Aspergillus* enzymes have been cloned, sequenced, amplified, and engineered to improve yields, stability, and other properties (24,25). Genetic manipulation also has allowed *Aspergillus* to become the host for the production of various nonnative (foreign) proteins, the best known of which is chymosin. Rennet is a substance extracted from the stomachs of slaughtered calves that is used in cheese making. The principle active ingredient of rennet is chymosin, a protease that coagulates milk. Structural genes for the calf stomach enzyme can be transferred to fungi. The mammalian enzyme is then expressed and secreted by the fungus in fermentation broths. The yeast *Kluyveromyces lactis* and *A. niger* have both been developed as production hosts for chymosin production. Given the economic advantages of fungal fermentations, it can be

Table 3. Commercially Important *Aspergillus* Enzymes

Enzyme	Applications
Amylases	Hydrolysis of starch; bread and beer production; removal of sizing from fabrics; high-fructose syrups
Invertase	Confections; soft center in chocolates
Lactase	Hydrolysis of lactose; production of syrups for sweetening agents
Pectinases	Pretreatment of fruit juices to remove turbidity, reduction cloudiness in wines
Phytase	Animal feed additive for liberation of phosphate from plant material
Proteases	Meat tenderizing; removal of bitter flavors and chill proofing of beer; reduction elasticity of gluten proteins in bread

predicted that *Aspergillus* will be the production host of choice for many other nonfungal enzymes that traditionally have been extracted from higher plants and animals.

Descriptions of the use of enzymes in fermentation, diagnostics, baking, brewing, cheese manufacturing, chemical biotransformation, detergents, effluent and waste treatment, leather, textile, starch conversion, and the like are presented in Godfrey and West's *Industrial Enzymology* (26). This work also gives detailed information on the legislation and regulation of commercial enzymes, toxicological and regulatory considerations, and the scientific and regulatory aspects surrounding the use of genetically engineered production strains (26). The long history of safe use and the GRAS status of *A. oryzae* and *A. niger* will ensure that they will remain the species of choice for industrial enzyme production, especially in the food industry.

DISEASE

Although *Aspergillus* species are very common in the human environment, only a limited number of taxa (approximately 17) are capable of initiating human disease. The usual reservoir for an *Aspergillus* infection is not another infected person but, rather, a nonliving site where the fungus is growing saprophytically. Air conditioning vents, attic insulation, cellar walls, stables, and even houseplants in hospitals have been implicated in various disease states. Because *Aspergillus* conidia are among the most common of airborne molds, the diseases they cause are usually pulmonary. The single species most commonly associated with human disease is *Aspergillus fumigatus*. In addition, *A. flavus*, *A. nidulans*, *A. niger*, and *A. terreus* are regularly implicated.

The term *aspergillosis* has been broadly used to describe any disease caused by an *Aspergillus* species, excluding mycotoxicoses. Because this term encompasses so many different clinical entities, with different researchers applying different emphases, and because some conditions defy simple categorization, the medical mycology literature can be confusing. The classification of Kwon-Chung and Bennett (5) and Dixon and Walsh (27) recognizes three broad categories of aspergillosis: *(1)* allergy, *(2)* saprophytic colonization, and *(3)* invasive mycosis.

Allergenic Responses

Like other fungal spores, *Aspergillus* conidia are potent allergens, especially associated with respiratory allergies. Spore inhalation may cause hypersensitivity reactions in susceptible individuals. For example, allergic bronchopulmonary aspergillosis occurs in patients with preexisting asthma who develop bronchial sensitivity reactions to inhaled conidia. Occupational exposure is a risk factor. Industrial-scale fermentations can cause sensitization of workers to high concentrations of airborne spores or to culture fluid materials, resulting in asthma or other adverse health effects. Malt worker's lung is a hypersensitivity pneumonitis caused by *Aspergillus clavatus*. Some medical mycologists classify respiratory allergies as extrinsic (inhalation of conidia) or intrinsic (*Aspergillus* growing in airways). Intrinsic conditions are an example of saprophytic growth.

Saprophytic Growth

In contrast to extrinsic respiratory allergies caused by spore inhalation, intrinsic aspergillosis is a localized hyphal growth associated with colonization of air spaces. Aspergilloma, or fungus ball, is formed when *Aspergillus* grows in a preexisting lung cavity. Fungus balls have also been reported to occur in the paranasal sinus.

Invasive Aspergillosis

Sometimes *Aspergillus* species are opportunistic pathogens causing frank mycosis. These mycoses are collectively called invasive aspergillosis and are usually classified by the portal of entry. Invasive pulmonary aspergillosis is typically manifested as acute pneumonia. Nonpulmonary portals of entry include the eye, paranasal sinus, or gastrointestinal tract or may be associated with vascular surgery and drug addiction. Invasive aspergillosis is largely a disease of a compromised host. Those at high risk include patients with diabetes, drug addicts, chronically ill patients, transplant patients receiving immunosuppressive drugs, and cancer patients receiving chemotherapy. Occupational exposure is also a factor. Historically, the earliest reports of aspergillosis were among wig cleaners and pigeon fanciers.

Aspergillus fumigatus and *A. flavus* are the most common species associated with invasive and noninvasive aspergillosis. It is hypothesized that these species possess special virulence factors that facilitate their role as opportunistic pathogens.

Invasive disease is treated with antifungal chemotherapy, but existing antimycotic drugs have toxic side effects, and the therapeutic outcome is often poor. As the use of immunosuppressive drugs increases in modern medicine, invasive aspergillosis is becoming more common. There is an important need for effective systemic antifungal drugs. Several pharmaceutical companies have active drug discovery programs in this area.

Mycotoxins

Filamentous fungal metabolites that cause human or veterinary diseases are called mycotoxins. The mycotoxins that have received the most attention are the aflatoxins, a family of highly toxic, mutagenic, and carcinogenic compounds produced by certain strains of *A. flavus*, *Aspergillus nomius*, and *A. parasiticus*. These species regularly contaminate peanuts, corn, cottonseed, rice, tree nuts, and other agricultural commodities (28). A large body of epidemiological and molecular biological data implicates aflatoxin with human liver cancer (29). Acute human aflatoxicosis, however, is rare and usually only occurs when famine or poverty forces people to subsist on moldy foodstuffs. On the other hand, veterinary aflatoxicosis is a major problem, with poultry and trout being particularly susceptible.

Other mycotoxins produced by members of the genus *Aspergillus* tend to be produced at toxicologically less im-

portant levels. These include cyclopiazonic acid (*A. flavus*), patulin (*A. clavatus*), and sterigmatocystin (*A. nidulans* and *Aspergillus versicolor*). The toxigenic potential of *Aspergillus ochraceus* is particularly varied, with different strains reported to make ochratoxin, citrinin, xanthomegnin, viomellein, and penicillic acid. Like aflatoxin, patulin and sterigmatocystin are carcinogenic; citrinin and ochratoxin A are nephroxic; cyclopiazonic acid is a neurotoxin; and xanthomegnin and viomellein have been associated with photosensitization and liver damage. Many of the compounds classified as cytotoxic or neurotoxic in vertebrates also show toxicity against insects. Sclerotia have been a particularly rich source of metabolites with insecticidal properties (30).

Mycotoxin contamination varies from year to year based on weather, agricultural practices, storage conditions, and other factors. Aflatoxin is usually worst in drought years. International agencies have set levels for permissible standards of contamination (31).

SUMMARY

Among the most commonplace and versatile of all fungi, the aspergilli play an important role in natural degradative cycles. Although a few species have gained notoriety as animal pathogens and mycotoxin producers, the vast majority are benign. In processes ranging from traditional Asian food preparation to massive contemporary industrial fermentations, the capacity of aspergilli to secrete enzymes and metabolites is exploited for human benefit. In basic science, *A. nidulans* is a model system for genomics and molecular biology as well as classical genetics. The increasing societal emphasis on environmentally safe processes and commodities guarantees that the metabolic capabilities of this genus will continue to be used in modern industry.

BIBLIOGRAPHY

1. C. Thom and M.B. Church, *The Aspergilli*, Williams & Wilkins, Baltimore, Md., 1926.
2. C. Thom and K.B. Raper, *A Manual of the Aspergilli*, Williams & Wilkins, Baltimore, Md., 1945.
3. K.B. Raper and D.I. Fennell, *The Genus Aspergillus*, Williams & Wilkins, Baltimore, Md., 1965.
4. M.A. Klich and J.I. Pitt, *A Laboratory Guide to the Common Aspergillus Species and Their Teleomorphs*, CSIRO Div. Food Processing, North Ryde, Australia, 1988.
5. K.J. Kwon-Chung and J.E. Bennett, in *Medical Mycology*, Lea & Febiger, Philadelphia, 1992, pp. 201–247.
6. R.S. Samson and J.I. Pitt, *Advances in Penicillium and Aspergillus Systematics*, Plenum Press, New York, 1985.
7. J.W. Bennett, in A.L. Demain and N.A. Soloman eds., *Biology of Industrial Microorganisms*, Benjamin/Cummings, Menlo Park, Calif., 1985, pp. 359–406.
8. R.A. Samson and J.I. Pitt, *Modern Concepts in Penicillium and Aspergillus Classification*, Plenum Press, New York, 1990.
9. R.A. Samson, in J.W. Bennett and M.A. Klich eds., *Aspergillus: Biology and Industrial Applications*, Butterworth-Heinemann, Boston, 1992, pp. 355–390.
10. J.I. Pitt and R.A. Samson, in W. Grauter ed. *NCU-2, Names in Current Use in the Families Trichocomaceae, Cladoniaceae, Pinaceae and Lemnaceae*, Koeltz Scientific Books, Konigstein, Germany, 1993, pp. 13–57.
11. A.H.S. Onions, D. Allsopp, and H.O.W. Eggins, *Smith's Introduction to Industrial Mycology*, E. Arnold, London, 1981.
12. G. Pontecorvo, J.A. Roper, L.M. Hemmons, K.D. MacDonald, and A.W.J. Bufton, *Adv. Genet.* **5**, 141–239 (1953).
13. O.A. Clutterbuck, in J.W. Bennett and M.A. Kich eds., *Aspergillus Biology and Industrial Applications*, Butterworth-Heinemann, Boston, 1992, pp. 3–18.
14. J.E. Smith and H.A. Pateman eds., *Genetics and Physiology of Aspergillus*, Academic Press, London, 1977.
15. S.D. Martinelli and J.R. Kinghorn eds., *Aspergillus: 50 Years On*, Elsevier, Amsterdam, 1994.
16. J.W. Bennett, *Fungal Gen. Biol.* **21**, 3–7 (1997).
17. M. Roehr, C.P. Kubicek, and J. Kominek, in J.W. Bennett and M.A. Klich eds., *Aspergillus: Biology and Industrial Applications*, Butterworth-Heinemann, Boston, 1992, pp. 91–131.
18. C.P. Kubicek, C.F.B. Witteveen, and J. Visser, in K.A. Powell, A. Renwick, and J.F. Peberdy eds., *The Genus Aspergillus: From Taxonomy and Genetics to Industrial Application*, Plenum, New York, 1994, pp. 135–145.
19. W.B. Turner, *Fungal Metabolites*, Academic Press, London, 1971.
20. W.B. Turner, *Fungal Metabolites II*, Academic Press, London, 1983.
21. G. Lancini and R. Lorenzetti, *Biotechnology of Antibiotics and Other Bioactive Microbial Metabolites*, Plenum, New York, 1993.
22. N.R. Reddy, M.D. Pierson, and D.K. Salunkhe eds., *Legume-Based Fermented Foods*, CRC, Boca Raton, Fla., 1986.
23. K.H. Steinkraus ed., *Handbook of Indigenous Fermented Foods*, Dekker, New York, 1983.
24. T. Christensen, in K.A. Powell, A. Renwick, and J.F. Peberdy eds., *The Genus Aspergillus: From Taxonomy and Genetics to Industrial Application*, Plenum, New York, 1994, pp. 251–259.
25. R.M. Berka, N. Dunn-Coleman, M. Ward, in J.W. Bennett and M.A. Klich eds., *Aspergillus: Biology and Industrial Applications*, Butterworth-Heinemann, Boston, 1992, pp. 155–202.
26. T. Godfrey and S. West eds., *Industrial Enzymology*, Stockton, New York, 1996.
27. D.M. Dixon and T.J. Walsh, in J.W. Bennett and M.A. Klich eds., *Aspergillus: Biology and Industrial Applications*, Butterworth-Heinemann, Boston, 1992, pp. 249–267.
28. L.A. Goldblatt, *Aflatoxin*, Academic Press, New York, 1969.
29. D.L. Eaton and J.D. Groopman eds., *The Toxicology of Aflatoxins*, Academic Press, New York, 1994.
30. V. Betina, *Mycotoxins—Production, Isolation, Separation, and Purification*, Elsevier, Amsterdam, 1984.
31. Food and Agricultural Organization of the United Nations, *World Wide Regulations for Mycotoxins: A Compendium*, FAO, Rome, 1995.

See also METABOLITES, PRIMARY AND SECONDARY; RHEOLOGY OF FILAMENTOUS MICROORGANISMS, SUBMERGED CULTURE; SECONDARY METABOLITE PRODUCTION, ACTINOMYCETES, OTHER THAN STREPTOMYCES.

ASTM STANDARDS FOR BIOTECHNOLOGY

G. LARRY EITEL
Jacobs Engineering Group
Denver, Colorado

KEY WORDS

ASTM standards
Bioprocess conversion
Consensus standards
Cryogenic
Preservation
Standards
Validation
Voluntary standards

OUTLINE

INTRODUCTION

As the biotechnology products and processes industry has grown during the past several decades, so has the realization grown that more standard test methods are needed to keep pace with regulatory requirements to assure the public that these products and processes are safe and effective methods of producing pure products. In 1996, the federal government passed a law directing federal agencies to participate in and use technical standards developed by voluntary consensus standards organizations, with few exceptions, to carry out policy objectives and activities. The American Society for Testing and Materials (ASTM) was specifically mentioned as an example of a voluntary consensus standards organization.

The ASTM E-48 Committee for Biotechnology was organized in 1985. This was in response to urging from the biotechnology community to lead the development of voluntary consensus standards in biosafety, characterization and identification of biotech organisms, preservation and storage, bioprocess controls and processing equipment, environmental issues, process validation, equipment qualification, and terminology. Biomass conversion standards responsibility was transferred to E-48 in 1987. Although many of these topics have analytical procedures currently in use by individual companies or trade organizations, there is a well-known need to standardize these procedures for fully acceptable application in industry and governmental regulatory organizations.

By 1997, the E-48 committee had published and is still responsible for more than 30 standards. More standards, based on the ASTM structured manner using technically qualified volunteers from international resources, are currently being developed.

This article explains the characteristics of voluntary consensus standards, how ASTM develops these standards, what standards are currently available, and which ASTM E-48 subcommittees are responsible for these and future standards development.

ASTM BACKGROUND

A Century of Consensus Standards

The ASTM has developed during the past century into one of the largest voluntary standards development organizations in the world. Organized in 1898, ASTM is a not-for-profit organization that enables all interests to participate in a common effort to write, review, and approve standards. These standards are used by producers, ultimate consumers, users, and general interest groups such as government, academia, and others in domestic and international settings.

What Is the Purpose of ASTM?

As stated in the *ASTM Technical Committee Officer Handbook* (1), the charter purpose is "the development of standards on characteristics and performance of materials, products, systems and services; and the promotion of related knowledge."

Currently, ASTM has 132 standards-writing technical committees, publishes more than 9,000 standards annually, and has more than 35,000 technically qualified members who voluntarily develop and maintain current standards in their committee specialty (2). The more familiar standards developed by ASTM include the concrete, iron, and steel products standards used in the construction

industry today. These standards have been developed for public safety. Other categories of ASTM standards include nonferrous metal products; metals; construction products (cement, road paving, roofing, masonry, insulation, sealants, soil, wood); petroleum fuels; lubricants; paints; coatings; textiles; plastics; rubber, electrical insulation; electronics; water; environmental technology; nuclear, solar, and geothermal energy; medical devices; emergency services; general laboratory analytical methods; hazardous substances; biotechnology, and others (2). These published standards are available for purchase from ASTM (100 Barr Harbor Drive, West Conshohocken, Penn., 19428-2959; phone (610) 832-9585).

What Are Consensus Standards?

Because ASTM standards are developed voluntarily and used voluntarily, the basic principle in their development is that a rigorous due process procedure is followed, resulting in a full consensus of all concerned parties prior to approval and publication of the standard by ASTM. ASTM technical committees are dedicated to this consensus process because these standards are often adopted by regulatory agencies in government regulations or referred to in business contracts between client's and contractors. This use of ASTM standards goes beyond the voluntary use and becomes legally binding.

Governmental agencies encourage the use of consensus standards, as demonstrated on March 7, 1996, when President Clinton signed Public Law 104-113, Technology Transfer Improvements Act of 1995 (3). This new law codified the existing Office of Management and Budget (OMB) Circular A-119 titled *Federal Participation in the Development and Use of Voluntary Standards*. This law directs all federal agencies and departments to use technical standards developed or adopted by voluntary consensus standards organizations, such as ASTM, to carry out policy objectives or activities. These federal agencies and departments consult with voluntary, private sector, and consensus standards bodies, when possible within budgetary resources and priorities, they should also participate in the development of these technical standards.

Senator John D. Rockefeller (D-W.Va.) proposed this legislation (4) and stated that

> Consensus standards are standards which are developed by voluntary, private sector, consensus standards bodies. These organizations are established explicitly for the purpose of developing such standards through a process having three characteristics. First, openness, defined as meaning that participation in the standards development process shall be open to all persons who are directly and materially affected by the activity in question. Second, balance of interest, which means that the consensus body responsible for the development of a standard shall be comprised of representatives of all categories of interest that relate to the subject. For example, manufacturer, user, regulatory, insurance, inspection, employee and union interest. Third, due process, which means a procedure by which any individual or organization who believes that an action or inaction of a third party causes unreasonable hardship or potential harm is provided the opportunity to have a fair hearing of their concerns. In short, a legitimate consensus standards organization provides open process in which all parties and experts have ample opportunity to participate in developing the consensus. Examples include traditional standards organizations such as the American Society for Testing and Materials.

How Is ASTM Organized to Produce Standards?

The basic organization of ASTM is illustrated in Figure 1. The board of directors is the governing body elected by the ASTM membership of more than 35,000. The standing committees maintain and enforce the procedures specified by the charter for developing standards. The Committee on Technical Committee Operations oversees the application of the regulations governing technical committee operations. It recommends changes to the *Regulations*, resolves jurisdictional disputes between committees, develops guidelines for long-range planning, and ensures that committee bylaws are consistent with ASTM regulations (5). The Committee on Standards maintains procedures for preparing standards and reviews all technical committee standards actions for consistency with the regulations before approval and publication by the society. The Committee on Publications advises the board on publication policy matters and oversees publication of all ASTM materials. The technical committees exist as semiautonomous groups to develop the standards. A more detailed description of the ASTM E-48 Committee for Biotechnology is presented later. Each technical committee is assigned a staff manager who assists the volunteer leadership in management functions and coordinates administrative services.

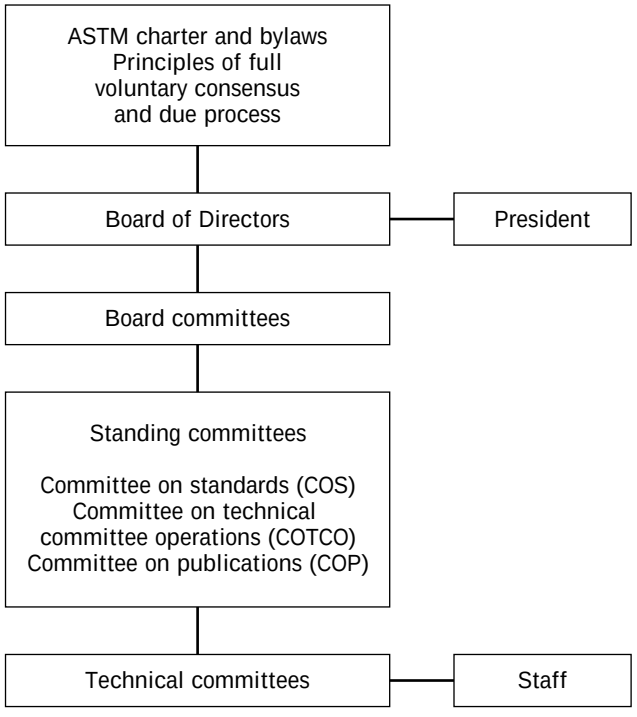


Figure 1. ASTM Organization. The president is a nonvoting member of the board of directors. For full descriptions of these committees, consult *Regulations Governing ASTM Technical Committees*.

TYPES OF ASTM STANDARDS

Categories of Standards

First, the word *standard* in ASTM is used primarily as an adjective with a specific document, such as a practice or test method, to denote consensus and approval (6). The technical committees, such as ASTM E-48 Committee on Biotechnology, determine the type of standard document to be developed and approved. There are two categories of documents within ASTM, as shown in the following (5):

- *Standard*. This document has been developed and approved by consensus within the principles of ASTM and meets the approval requirements of ASTM procedures and regulations.
- *Provisional Standard*. This document has been approved and published by ASTM for a limited period of time to meet urgent needs, such as an emergency situation, regulatory requirements, or other special circumstances. Provisional standards are not full consensus documents because they are approved by technical subcommittees only and not by full society consensus.

Types of Standards

The two categories of ASTM documents are found in the following seven types and forms, based on the technical content and intended use, not on the degree of consensus achieved (5). These different types of ASTM documents provide a flexibility of form, communication, and usage for the technical committees and the wide variety of users of these documents.

1. *Classification*. This is a systematic arrangement or division of products, systems, materials, or services into groups that have similar characteristics, such as composition, origin, properties, or use. An example is classifications of types of coal having similar specifications.
2. *Guide*. A guide is a documentation of general options or instructions that do not recommend a specific course of action. Normally, a guide suggests an approach based on consensus of viewpoints but does not establish a fixed procedure. A guide can be used to create increased awareness of the available techniques on a given subject and to provide information for subsequent evaluation and standardization. Note the difference between a guide and a practice: a guide suggests an approach, whereas a practice dictates a general usage principle.
3. *Practice*. This document is a definitive procedure for performing one or more specific operations or functions where a definitive test result is *not* produced (compare with test method). Examples of practices include procedures for conducting interlaboratory testing programs, statistical procedures, sampling procedures, testing equipment selection, operation, maintenance, installation, cleaning, training, application, assessment, decontamination, inspection, screening, and disposal.
4. *Specification*. This precise statement of requirements to be satisfied by a material, product, system, or service identifies the procedures for determining whether each of the requirements is satisfied. Examples include purity, composition, pH, or minimum or maximum ranges of acceptable values of the properties for a biochemical product. The specification could include physical, mechanical, or chemical properties of the material. Also applicable are quality, safety, or performance criteria with identification of the test methods to be used in evaluating these issues.
5. *Terminology*. This document contains definitions, descriptions of terms, and explanations of symbols, abbreviations, or acronyms used in the subject area. Each standard can have terminology specific to that standard, and then a consolidated standard of terminology can be approved for the general subject or committee area of responsibility. For example, ASTM Committee E-48 on Biotechnology has developed more than 30 individual standards with their own terminology sections. In addition, Standard E 1705-95 has been approved; it summarizes appropriate terminology specific to biotechnology and related ASTM standards.
6. *Test method*. A test method is a definitive procedure that produces a test result that can include identification, measurement, and evaluation of qualities, characteristics, or properties of the material, product, system, or service.
7. *Other types of ASTM documents*. Other documents include those less frequently used, such as charts, tables, matrixes, and reference photographs recommended by the sponsoring technical committees with approval from the ASTM Committee on Standards.

Who Uses ASTM Standards?

ASTM standards have a far-reaching international and domestic impact. Manufacturers of goods for the public use many ASTM procedures to specify raw materials used in the production of other products and services. The specifications for these produced goods are often tied into ASTM standards, where the producer and consumer are both testing the products for acceptability, identity, safety, and quality using the same test methods and specifications. Research and development interests use ASTM standards that are tried and proved to be accurate and reliable. In production of drugs by biotechnical or pharmaceutical routes, for example, it is critical that testing and analytical characterization of the chemical intermediates and final products be accurate, repeatable, and able to be validated. This enables regulatory agencies, such as the U.S. Food and Drug Administration (FDA) to monitor the quality of drug products using established, proved test methods and production protocols. Governmental agencies use consensus standards in concert with regulatory standards to assure the public that products being purchased are safe and proper. For example, in the mid-1980s the U.S. Department of Agriculture (USDA) asked ASTM to help prepare a checklist for USDA officials. This checklist would be used

to screen applicants requesting financial support to install small ethanol fermentation plants on their farms. Some of these plants did not work effectively because of incomplete designs. The checklists were prepared and used by the USDA to improve the effectiveness of evaluating the technical success potential of the new ethanol projects being proposed. These checklists were developed by volunteers in the ASTM Biomass Subcommittee with participation from USDA employees.

Other users of these consensus standards include insurance and inspection organizations that use standard guides, practices, and test methods to monitor and check the quality of products being used by the public. These standards may also find their way into the court system, where successful litigation may depend on using test methods that are credible and used as standards within the industry.

ASTM standards are international; the Biotechnology Committee E-48 includes members from Japan, Europe, and Canada, among others, often representing international standard-setting organizations with common interests. An example of cooperative efforts includes development of the polymerase chain reaction (PCR) test methods. ASTM E-48 members cooperated with the Deutsches Institut fur Normung (DIN) standards committee in Germany to develop PCR standards.

ASTM COMMITTEE E-48 FOR BIOTECHNOLOGY

Technical Committee Organization

The structure of the workhorse organization to produce ASTM standards in biotechnology is shown in Figure 2.

In 1985, the ASTM Board of Directors approved establishing the E-48 Committee on Biotechnology to answer requests from industry and government and provide voluntary consensus standards that would meet the needs in this rapidly emerging new field. E-48 founding members developed new bylaws and organization structure to represent a balanced interest from producers, general interest groups, and consumers. Because E-48 develops standards for products, services, systems, and materials that are offered for sale, the committee is required to be classified and balanced; the number of voting members representing producers does not exceed the total membership in the remaining classifications of voting members (5). Currently, there are about 100 members of E-48 from international and domestic countries. Officers are elected by the E-48 membership, and those in key positions serve for 2 years.

The chairperson and two vice chairpersons direct and manage the activities of the E-48 administrative and technical subcommittees (Fig. 2). The Executive Subcommittee provides direction for the main committee, including long-range planning, symposium planning, and coordination of the technical and technical service subcommittees. The Main Committee meets semiannually to conduct business, review balloting, approve new members, and plan for the short- and long-range interests of developing standards for biotechnology. E-48 has sponsored several symposia (7) for biotechnology and has published papers through ASTM. Papers resulting from the symposia are subject to peer review and approval by the Committee on Publications before publication.

E-48 technical subcommittees have been organized to represent specific subjects within the scope of the E-48 charter. Each technical subcommittee is the primary resource to develop and approve ASTM standards for biotechnology as described in "Standards Development". Each technical subcommittee is normally subdivided into task groups that focus on specific standards development. The

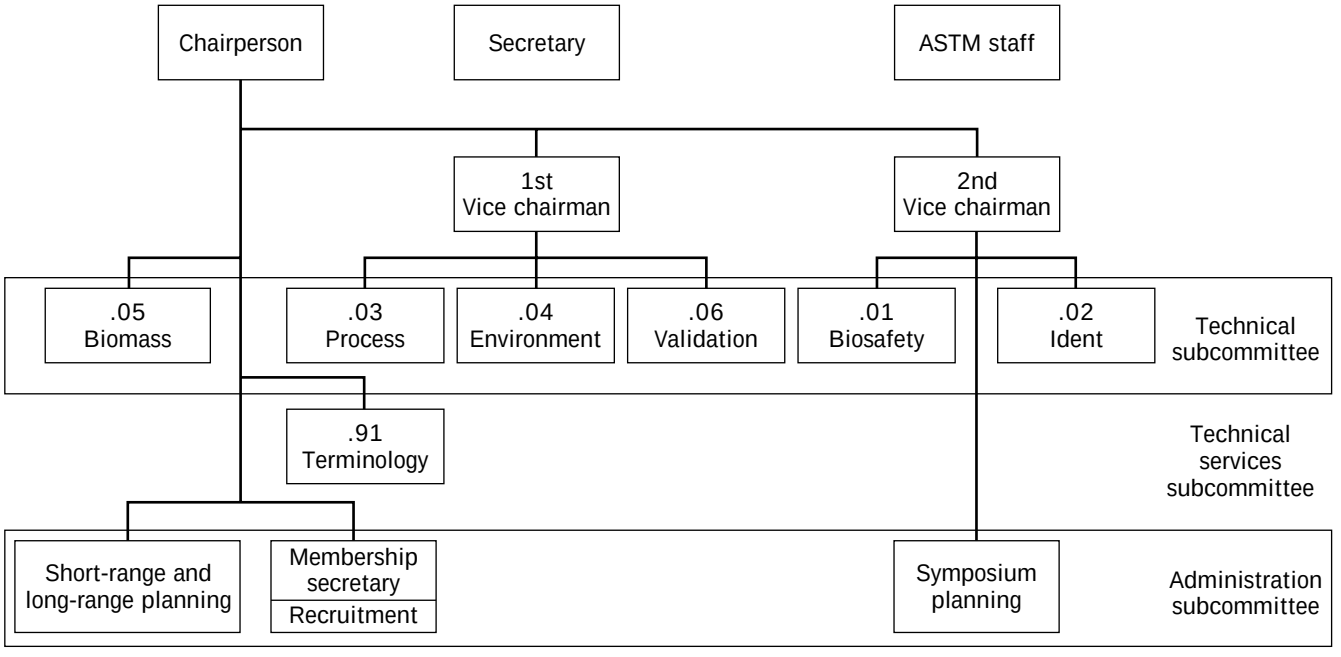


Figure 2. Main committee E-84 on biotechnology.

task group is composed of ASTM, affiliated, and non-ASTM members who are knowledgeable about the subject matter. The numbers of task groups vary, depending on the subjects being considered for standards. Minutes of all formal meetings are published for distribution to ASTM members to communicate the status and to correlate the overall committee work. For information on the E-48 subcommittee standards developed by specific subcommittees, refer to "E-48 Standards Summary."

Liaison assignments to committee members are made to coordinate E-48 activities and interests with other ASTM committees or other standards organizations (8). For example, the charter of Committee E-47 on Biological Effects and Environmental Fate includes standards on (*1*) the effects of physical and chemical stress on aquatic and terrestrial plants and animals (including humans); and (*2*) the properties of materials that affect and determine their fate, distribution, and persistence when introduced into the environment. The Committee D-22 on Sampling and Analysis of Atmospheres covers biological aerosols, mainly from sampling and analytical aspects. These and other ASTM main committees may have overlapping scopes of interest with E-48, and close coordination is needed to properly develop a standard of common interest. Sometimes, a joint task group is formed, and representatives from various committees develop the common standard. This is necessary because the charter of E-48 covers interests that include characterization of biological species, biosafety, processes using living organisms, environmental issues, biomass conversion, process validation, and terminology. Members of E-48 also serve as liaisons with other organizations, such as the American Society of Mechanical Engineers and professional, pharmaceutical, and drug manufacturing organizations.

ASTM STANDARDS DEVELOPMENT

Overview of Consensus Standard Development by E-48

The life cycle of an ASTM standard begins when a need is recognized and deemed appropriate for development by the E-48 technical subcommittee and executive subcommittee (1). Task groups are assigned by the subcommittee chair, and experts are assigned to develop a standard using proper format (6), which is then voted on by the subcommittee members.

- *Subcommittee ballots* require an affirmative vote of at least two-thirds of the votes cast by voting members, at least 60% of the voting members must return the letter ballots. If a negative vote is received with explanation of the reasons for the negative, then the subcommittee must resolve the negative by a meeting of the subcommittee or by another letter ballot. In a consensus standard, it is essential to properly resolve any negative ballot issues so that a full consensus is obtained.
- *Main committee ballots* are used to vote on the standard after completion of a subcommittee letter ballot. ASTM headquarters issues all letter ballots for the main committee. An affirmative vote of at least 90% of the voting members is needed, at least 60% of the voting members must return the letter ballots. The task of resolving all negative votes not previously considered belongs to the original subcommittee. If substantive changes to the standard are required, then the standard is reballoted normally at the subcommittee level.
- *Society review* of the standard is conducted simultaneously with the main committee ballot through publication in the monthly ASTM magazine, *Standardization News*. Each member of the society can then vote on the standard. Negatives again must be fully addressed by the appropriate subcommittee that developed the standard.
- The *Committee on Standards* then reviews the balloting and approves publication of the standard in the *Annual Book of ASTM Standards*.
- *Updating* the approved standard can be achieved whenever a significant change is needed by using the entire balloting sequence. Every 4 years, each standard is reviewed and balloted for either revision, reapproval, or withdrawal.

The E-48 Committee also seeks comments on new draft standards from other reviewers, even if they are not members of ASTM, to be sure that the new standard receives full consensus from biotechnical interests. Often, E-48 is requested by other biotechnology associations to discuss biotech standards and activities at ASTM.

E-48 STANDARDS SUMMARY

Since 1985 when E-48 was organized, the number of approved standards has increased to more than 40. Three symposia and seminars have been held on biotechnology subjects such as biosafety, validation, and general biotech interests (7,9,10). The abstracts from these documents are described in the next section, organized by subcommittee.

E-48.01 Subcommittee on Biosafety

Although the E-48.01 Subcommittee on Biosafety was established in 1996, biosafety was a main topic when E-48 was organized in 1985, and a symposium was conducted in 1987. This section summarizes the work of this subcommittee.

Scope and Background. Although multiple agencies regulate and provide guidelines for safety and transport of biomaterials within manufacturing facilities, a need for clarification of these standards was indicated in 1985. However, in the past 14 years, significant progress has been made in the industry to produce safety guidelines, especially at the laboratory-scale level. Possible augmentation of safety standards may be an area for ASTM to investigate. For example, the D-22 Committee on Biological Aerosols is considering standards to measure biocontaminants in indoor air. This is a very difficult area to develop effective test methods, but the industry is highly

interested in this topic. E-48 has interest in developing test methods that can be applied to large-scale operations relative to employee protection. Standardization of guidelines for safety action plans that can be customized for biotech operations could help regulatory agencies in their monitoring of industry manufacturing.

Available ASTM Standards. No biosafety standards have been produced to date. However, the BioSafety Symposia in 1987–1988 included papers on the following topics, some of which may provide basic information and guidance.

1. The section on safety in the bioprocessing industry section discussed the NIH Recombinant DNA Advisory Committee, which is an example of the self-regulatory process. Concepts and concerns of biosafety during the past decade were presented, such as risk taking, opportunities, and roadblocks to minimizing hazards in bioprocessing.
2. The section on design criteria for safety discussed biological production facility containment issues, regulations that govern, and examples of design criteria for safety in the isolation and purification of antibiotics and biologically active compounds. The emphasis was on protecting the worker and community, and significant strides have been made by the industry without additional standards from ASTM.
3. The personnel safety planning section reviewed the general considerations for work practices and personal protective equipment for biological operations. Significant standards are in place today from industry associations, regulatory agencies, and specialty safety equipment suppliers.
4. Safe waste disposal issues are as fundamental today as they were 10 years ago, and the need to minimize community environmental impact is just as essential. Community awareness and emergency response issues were presented. International, national, and local regulatory aspects of bioprocessing were presented. Ten years ago, the fear of overregulation was evident. Today, concern and self-regulation by the industry members have encouraged regulatory agencies to continue monitoring this progress without dictating additional guidelines. However, if these industry members or regulatory agencies see a need for additional standard guides, test methods, or practices, then ASTM E-48 will assist in developing a consensus procedure.

New Standards Being Developed ASTM E-48 is currently looking to the industry membership and regulatory agencies to identify standards that are justified to help improve biosafety.

Future Trends. Surveys and discussions with the industry have been and will continue to be used by the E-48.01 subcommittee to identify new standards for development. Participation with the overall E-48 effort to identify and develop needed, effective, consensus standards is encouraged from industry organizations, manufacturing companies, research institutions, and regulatory agencies.

E-48.02 Subcommittee on Characterization and Identification of Biological Systems

The E-48.02 subcommittee, organized in 1985, is responsible for standard test methods and guides that identify, characterize, and preserve biological species. At this time, 13 standards are published for use.

Scope and Background. The primary charter of Subcommittee E-48.02 is to develop test methods and guides for identifying plasmids, cells, viruses, enzymes, bacteria, and fungi. This includes methods for determining contamination in biological products, detecting mycoplasma, detecting nucleic acids by PCR, and methods for preserving biological products. Guides to meet FDA in vitro diagnostic requirements are also being considered.

Available ASTM Standards. The following standards are published by ASTM (2):

1. Identification of viruses
   a. E 1285, *Guide for Identification of Bacteriophage Lambda or Its DNA*, describes laboratory characterization procedures that are sufficient to verify that a biological preparation believed to contain λ or λ-DNA for use in any step of a biotechnology process actually does contain this material. This procedure should be used if the bacteriophage λ is to be used to construct a recombinant molecule.
   b. E 1286, *Standard Guide for Identification of Herpes Simplex Virus or Its DNA*, describes laboratory characterization procedures that would be sufficient to verify that a biological preparation believed to contain primarily HIV types 1 and 2 or their DNA for use in any step of a biotechnology process actually contains this virus or its DNA. This guide does not cover the identification of HSV in HSV-infected host cells without first isolating the virus from the sample using standard techniques of HSV purification.
   c. E 1493, *Standard Guide for Identification of Bacteriophage M13 or Its DNA*, includes procedures to verify that a biological preparation believed to contain M13 or M13 DNA for use in any step of a biotechnology process actually does contain this bacteriophage or its DNA. This procedure is applicable where M13 is to be used to construct a recombinant molecule and characterization of double-stranded RF DNA is made by restriction enzyme analysis or characterization of single-stranded phage DNA is made by DNA sequencing.
2. Detection of contamination or impurities
   a. E 1298, *Standard Guide for Determination of Purity, Impurities, and Contaminants in Biological Drug Products*, provides definitions of types and levels of impurities, purity of drug product, and

contaminants. It suggests analytical methods generally applied within the pharmaceutical industry to identify and quantitate the level of impurities in a drug product developed through biotechnology. The general effects of these contaminants on these drugs are also described.

b. E 1531, *Standard Practice for Direct Detection of Mycoplasma in Cell Culture by Broth Enrichment and Agar Growth*, describes the procedure that can be used for detection of mycoplasma contamination in cell cultures by direct microbiological culture. Preparation of the broth medium is described. This detection technique can be used in conjunction with the indirect method that is included in E 1532 93.

c. E 1532, *Standard Practice for Indirect Detection of Mycoplasma in Cell Culture by DNA Binding with Bisbenzamide Fluorochrome Stain*, describes the procedure that can be used for mycoplasma contamination detection, based on DNA staining and intended for use in examining cultured mammalian cells. It is not intended for use in detection of mycoplasma in serum, culture media, or systems other than cultures of mammalian cells. Preparation of the bisbenzamide stain concentrate is included with the cell staining procedures.

d. E 1533, *Standard Practice for Indirect Detection of Mycoplasma in Cell Culture by 4′-6-Diamidino-2-2 Phenylindole (DAPI) Staining*, is intended for use in examining only cultured animal cells for mycoplasma contamination using DAPI fluorochrome staining. The procedure includes preparation of DAPI stain concentrate.

e. E 1536, *Standard Practice for Large Volume Testing of Serum for Mycoplasma Contamination*, describes the procedure used for detection of mycoplasma contamination in serum, based on direct microbiological culture of a large volume (50 to 100 mL) of serum as an inoculum in a broth medium. The procedure includes preparation of the broth medium.

f. E 1759, *Standard Test Method for Isoaspartic Acid in Proteins: Method for the Determination of Asparagine Deamidation Products*, measures the amount of isoaspartic acid residues in a protein or peptide solution by using the enzyme protein isoaspartyl methyl transferase and radioactive *S*-adenosyl-L-methionine. The storage of proteins in aqueous solutions often results in the formation of isoaspartic acid linkages within the polypeptide chain as a result of the deamidation of aspargine residues and rearrangement of aspartic acid linkages. The concentration range of isoaspartic acid residues in the sample should be 2.5 to 50 μmol/L. The basis of the procedure is the production of radioactive methanol equal to the amount of isoaspartic acid residues in the sample by action of the enzyme protein isoaspartyl methyl transferase and radiolabeled *S*-adenosyl-L-methionine.

3. Preservation of biological materials

a. E 1342, *Standard Practice for Preservation by Freezing, Freeze-Drying, and Low Temperature Maintenance of Bacteria, Fungi, Protista, Viruses, Genetic Elements, and Animal and Plant Tissues*, outlines procedures to minimize the adverse effects of handling biological materials during low-temperature preservation and maintenance. Although a limited number of methods of preservation can be used on biological systems, the low-temperature methods provide the only real assurance of genetic stability. Cryogenic temperatures are those at or below $-70°C$. Suggested cooling rates and use of cryoprotectants are presented, and freeze-drying optimum temperatures are discussed. Safety precautions are included for handling cryogenic materials.

b. E 1564, *Standard Guide for Design and Maintenance of Low-Temperature Storage Facilities for Maintaining Cryopreserved Biological Materials*, covers recommended procedures for developing and maintaining low-temperature storage facilities for freezers with mechanical refrigeration and freezers cooled with liquid nitrogen. Suggestions are presented for selecting the proper equipment for the intended service.

c. E 1565, *Standard Guide for Inventory Control and Handling of Biological Material Maintained at Low Temperatures*, suggests safety and security procedures to be used for materials stored at low temperatures to avoid problems during storage. Properly designed inventory control systems ensure the maximum use of freezer space and that all material is located and retrieved easily, without compromising the stability of other items in the freezer (mechanical or liquid nitrogen design). Safety precautions are included for handling vials of materials at cryogenic temperatures.

d. E 1566, *Standard Guide for Handling Hazardous Biological Materials in Liquid Nitrogen*, lists special handling and storage procedures that avoid penetration of vials by liquid nitrogen during storage, which could contaminate the liquid nitrogen bath or contaminate personnel if improperly sealed vials were to explode. Safe handling techniques are discussed for retrieval of cultures from liquid nitrogen, it is recommended that vials be filled to only two-thirds of capacity. A leak test for sealed vials is also described.

4. Collaborative development of DIN standards on PCR. Subcommittee E-48.02 is actively participating in the collaborative development of DIN Committee E9 for the Polymerase Chain Reaction.

a. DIN 58967-10, *Serodiagnosis of Infectious Diseases and Diseases of the Immune System—Polymerase Chain Reaction* (PCR). *Part 60: Terminology, General Method-Specific Requirements*, is published in English and German.

b. DIN 58969-61, *Serodiagnosis of Infectious Diseases. Part 61: Polymerase Chain Reaction*

(PCR)—Special Requirements for the Detection of Nucleic Acid Sequences of the Human Immunodeficiency Viruses HIV-1 and HIV-2, is published in English and German.

5. Collaborative standards with Japan Bioindustry Association (JBA). Currently, JBA and E-48.02 are actively participating in a collaborative development of a standard on the detection of mycoplasma by PCR (*Mycoplasma Detection Method; Nested PCR Assay*)

New Standards Being Planned and Developed

1. Detection of nucleic acids by PCR
2. Identification of cell lines
3. Characterization of industrial bacteria
4. Guidance to help meet FDA requirements
5. Identification of plasmids and vectors
6. Characterization of fungi

Future Trends. Focus will be on completing the test methods and guides needed to safely and economically meet the rapidly growing application of biotechnology to produce drugs, chemicals, services, and products. Subcommittee E-48.02 continues to actively participate in the collaborative development of DIN (German Institute for Standardization) Committee E-9 standards for the PCR.

E-48.03 Subcommittee on Unit Processes and Their Control

The E-48.03 subcommittee was organized in 1985 for the purpose of developing guides and test methods for applying engineered equipment to biological processes. New instrumentation to detect process variables, such as pH, requires uniform procedures to ensure consistent performance. Separation techniques such as liquid chromatography provide opportunities for developing standard testing methods. The E-48.03 subcommittee was chartered to develop standards for these and other bioprocesses.

Scope and Background. The initial focus of this subcommittee was to develop standards for the emerging biosensor instrumentation and the specialized processing equipment used to properly handle living organisms to meet a specific objective. Unit processes being evaluated include bioreactor design, fermenter performance, membrane ultrafilter characterization, biosensors, and bio-specific equipment. Supporting processes such as aseptic sampling, sterile cleaning, clean-in-place, sterilize-in-place, and issues such as solution compatibility and materials of construction to meet aseptic requirements are also in the E-48.03 charter.

Available ASTM Standards. More than six standards have been developed or are in the process of development at this time. These standards are highlighted in the following, based on ASTM published resources (2).

1. Ultrafiltration membrane, chromatographic media, and protein characterization
   a. E 1294, *Test Method for Pore Size Characteristics of Membrane Filters Using Automated Liquid Porosimeter*, describes a procedure for measuring the pore size characteristics of membrane filters in the range of 0.05 to 300 μm. This method uses the automated bubble point method described in ASTM Test Method F 316. This test uses the capillary rise created by surface tension and uses the Washburn equation to calculate the pore diameter. This method determines maximum pore size, minimum pore size, and mean flow pore size of the membrane.
   b. E 1343, *Standard Test Method for Molecular Weight Cutoff Evaluation of Flat Sheet Ultrafiltration Membranes*, evaluates the molecular weight cutoff of flat sheet membranes between 4,500 and 1 million Da. A nonadsorbing penetrant liquid (dextran T-fractions) is used and can be applied to a wide variety of membrane substrates, excluding those that strongly adsorb dextran, from highly hydrophilic to highly hydrophobic.
   c. E 1772, *Standard Test Method for Particle Size Distribution of Chromatography Media by Electric Sensing Zone Technique*, is valuable for measuring particle size and distribution of a dilute suspension of chromatography media in the overall size range of 1 to 450 μm, using the electric sensing zone apparatus. This apparatus uses an electrical current path of small dimensions that modulates individual particles passing through an aperture to produce individual pulses of amplitude proportional to the particle volume.
   c. E 1470, *Standard Test Method for Characterization of Proteins by Electrophoretic Mobility*, is for determining the electrophoretic mobility of proteins of molecular weight greater than 10,000 Da using automatic electrophoretic light-scattering principles to determine its mobility. The instrument measures the Doppler shifts of scattered light at four different angles to determine the mobility distribution of particles in a dispersed, dilute suspension of the protein particles. The test results are expressed as either electrophoretic mobility or ζ potential distribution of the protein particles.
2. Bioprocesses
   a. E 1287, *Standard Practice for Aseptic Sampling of Biological Materials*, summarizes principles, state-of-the-art concepts, and generally accepted methods for aseptic sampling of materials in biotechnical process where contamination of either the sample or the source is not acceptable. Suggested designs include the advantages and disadvantages being discussed. Fabrication and maintenance considerations are presented.
   b. E 1357, *Standard Test Method for Determining the Rate of Bioleaching of Iron from Pyrite by* Thiobacillus ferrooxidans, describes how cells of *T. ferrooxidans* grown on ferrous iron are added to finely ground iron pyrite in an inorganic salt

medium of 2% pulp density. The culture is incubated with agitation, and samples are periodically withdrawn for soluble iron analysis. The rate of pyrite leaching is determined from the linear portion of a curve-plotting soluble iron produced over time. A high leaching rate is evidence for potential degradability of the mineral in mining operations using biologically active cultures.

New Standards Being Developed

1. Guide for evaluating the performance of stirred tank and bioreactors
2. Upgrading ultrafiltration membrane performance measurements

Future Trends. Equipment used to process cell cultures, viruses, bacteria, fungi, tissue, and other biological materials can be quite different from the equipment used in pharmaceutical processes. Unit processes such as reaction, separation, concentration, and purification have industry support to establish evaluation test methods, practices, and guides for proper use, cleaning, sterilization, and application. These needs are to receive focus by E-48.03 in the near future as experts in this field come forward to assist in identifying and preparing these standards.

E-48.04 Subcommittee on Environmental Issues

Although the E-48.04 subcommittee was established in 1985, its primary activity has been to review environmental issues in standards prepared by other E-48 subcommittees. This subcommittee has participated in symposiums, workshops, and surveys to define standards needed within the biotech industry.

Scope and Background. Two categories and specific topics of interest have been identified by industry, government, and research.

1. Biomanufacturing processes
   a. Containment of processes and local manufacturing areas
   b. Biotechnology risk assessment guides that are not currently available from National Institutes of Health, Centers for Disease Control, and other existing organizations
   c. Standardized methods for determining occupational exposure
   d. Decontamination of waste streams. Significant challenges still exist to effectively sample waste streams before discharge to the local waste treatment facilities.
   e. Standardizing construction of and protocols for model ecosystem mycocosm tests. Significant technical problems exist and need resolution before writing standards.
2. Deliberate release or uncontained applications could encourage development of guides for bioremediation microorganisms and their efficacy and efficiency.

Available ASTM Standards

1. E 1287, *Standard Practice for Aseptic Sampling of Biological Materials*, discusses containment options to avoid environmental releases. See E-48.03 "Subcommittee" for more details.
2. E 1344, *Evaluation of Fuel Ethanol Manufacturing Facilities*, presents containment and sterility issues for environmental, safety, and process yield considerations. See "E-48.05 Subcommittee" for other information.

New Standards Being Developed. Surveys and discussions with biotech industry members continue today, and volunteers are needed to identify new standards that are needed and have value to the industry.

Future Trends. Because the objectives and charters of E-48.04 include topics that are of interest to E-48.01 Subcommittee on BioSafety, a close cooperation between the two is planned to identify and develop the standards that are needed by the industry. Volunteers are needed to participate in this consensus standard development activity.

E-48.05 Subcommittee on Biomass Conversion

Biomass conversion standards were first developed in the biomass subcommittee of E-44 Committee on Solar Energy in the early 1980s. In 1982, standards were published for the measurement of moisture, volatiles, ash, and bulk density of particulate wood fuels. Initially, the focus of these standards was to characterize wood for fuel as an energy source and not to characterize biomass materials for conversion to other products using biotechnology. In 1982, the USDA requested ASTM to develop standards for evaluating the design and performance of small grain-fermentation plants that produce ethanol for oxygenate-enhanced fuel blending. This was the entry into biotechnology leading to biomass conversion using living organisms.

In 1985, biotechnology interests in government and industry encouraged ASTM to establish the E-48 Committee on Biotechnology. In 1987, the E-48 Committee requested and received approval for the transfer of the biomass subcommittee from E-44 because the biotechnical similarities would be more efficiently handled in E-48.

Scope and Background. The fuel crises starting in 1973 spurred interest in alternative fuels manufacturing, including fuel-grade ethanol from fermenting grains and other renewable feedstocks. Because grain-based feedstocks have high value for animal and human feeds, interest grew in converting other biomass materials to fuels using emerging bioconversion technology. Although wood, grasses, bagasse, and agricultural residues all contain carbohydrates, as do grains, their exposure for bioconversions to sugars and fermentation to alcohols is more difficult because insoluble lignin surrounds and protects the carbohydrates. This protection limits economical conversion by microorganisms and enzymes. However, with genetic engineering, a new, more efficient source of biological materials was emerging that could be commercially viable for

biomass conversions. With the intensified research by industry and U.S. Department of Energy–sponsored grants to organizations such as the National Renewable Energy Laboratory to conduct biomass conversion research, new test methods had to be developed to properly characterize biomass and measure the efficiency of its conversion to usable products. Existing standards from the wood products industry (TAPPI) describe characteristics of wood for use in the pulp and paper industry, but they have limited value for biotechnical use. Accordingly, the E-48.05 subcommittee has developed the following charter for developing standards for the biomass conversion industry.

1. Maintain current standards for wood materials used directly for fuel and energy generation
2. Develop standards for biomass feedstock characterization
3. Develop standards for biomass conversion to usable products through enzymatic saccharification, fermentation, and thermochemical processes
4. Develop terminology standards specific to the biomass feedstock characterization and conversion

Available ASTM Standards. The following standards are available for the wood and biomass conversion industries.

1. Wood fuels
   a. E 870, *Standard Test Methods for Analysis of Wood Fuels*, includes proximate and ultimate analysis of wood fuels. Lists of existing ASTM standards are referenced and include D 1102 (ash in wood), E 711 (gross calorific value of refuse-derived fuels [RDF] by bomb calorimeter), E 775 (total sulfur of RDF), E 777 (carbon and hydrogen in RDF), E 778 (nitrogen and oxygen in RDF), and other standards developed by E-44 and E-48.
   b. E 871, *Standard Method for Moisture Analysis of Particulate Wood Fuels*, determines total weight basis moisture in particulate wood fuel having a maximum particle volume of 16.39 cm^2 or 1 cu in. The method includes sample reduction techniques and oven-drying procedures for calculating moisture content, which can be used for establishing burning characteristics and process yields and as a basis for purchasing or selling wood fuels.
   c. E 872, *Standard Test Method for Volatile Matter in the Analysis of Particulate Wood Fuels*, determines the percentage of gaseous products, exclusive of moisture vapor, for wood particle sizes up to 16.39 cm^3. The technique includes sample preparation and volatilization in a vertical electric tube furnace at 950 °C for 7 minutes.
   d. E 873, *Standard Test Method for Bulk Density of Densified Particulate Biomass Fuels*, uses a fixed-volume open box with prescribed pouring height of cellulosic materials having up to 16.39-cm^3 particulate size. Repeated test values are averaged for final result.
   e. E 1288, *Standard Test Method for the Durability of Biomass Pellets*, describes a tumbling and sieve analysis test that evaluates the relative durability of a 45-kg sample of biomass fuel pellets. This sample is reduced to 1 kg, by a riffle or by coning and dividing. A 0.25-kg sample tumbles for 5 minutes at 50 rpm. A sieve test determines percentage fines passing through a no. 200 sieve.
   f. E 1358, *Standard Test Method for Determination of Moisture Content of Particulate Wood Fuels Using a Microwave Oven*, is a rapid test to determine moisture in wood fuels in several minutes using a microwave oven. The sample is stirred and reheated every 4 minutes or less, depending on values in the tables provided until a constant weight is reached and the percentage of moisture can be calculated.
   g. E 1534, *Standard Test Method for Determination of Ash Content of Particulate Wood Fuels*, discusses the ash content measurement of the noncombustible part of a particulate wood fuel, which uses an electric muffle furnace with an indicating pyrometer to oxidize a 2-g sample to ash.
2. Grain ethanol manufacturing plants
   a. E 869, *Standard Test Method for Performance Evaluation of Fuel Ethanol Manufacturing Facilities*, evaluates four main performance factors for a grain ethanol fermentation plant: conversion efficiency, energy for conversion, production rate, and mass balance. Procedures are contained to measure, interpret, and assign values for these performance parameters in a standard format. Data collection forms are also provided, and standard calculation equations are listed with examples.
   b. E 1117, *Standard Practice for Design of Fuel Alcohol Manufacturing Facilities*, presents minimum good engineering practices for use in designing, constructing, operating, and modifying fuel alcohol plants. Categories of equipment discussed include vessels, towers, heat exchangers, rotating equipment, electrical, instruments, safety, environmental, utilities, and piping.
   c. E 1344, *Standard Guide for Evaluation of Fuel Ethanol Manufacturing Facilities*, provides evaluation criteria to enable a prospective purchaser or lender to effectively review the plans, specifications, and plant operating concept of a fuel ethanol manufacturing facility and to determine whether its design meets the requirements of ASTM design practice standards. This guide can be used in conjunction with E 1117 practice. A design review checklist is included. The guide provides an explanation of the ethanol production process, followed by a discussion of design parameters and potential problem areas within the process.
3. Anaerobic digestion systems
   a. E 1535, *Standard Test Method for Performance Evaluation of Anaerobic Digestion Systems* pro-

vides for measuring the concentration and mass of the influent and effluent wastes and operational parameters, such as input energy, output gas production, and waste biomass, to provide a methodology for evaluation of the anaerobic digester operation. Data collection sheets, calculation procedures, and test report formats are presented.

4. Biomass and biotechnology terminology
   a. E 1126, *Standard Terminology Relating to Biomass Fuels*, presents and defines more than 70 terms used frequently in biomass fuels industries. These terms are directed to the wood fuels industry.
   b. E 1705, *Standard Terminology Relating to Biotechnology*, presents more than 160 terms used in biotechnology, with references to ASTM standards where they are used. As new E-48 standards are developed, new terms will be included.
5. Biomass feedstock characterization
   a. E 1690, *Standard Test Method for Determination of Ethanol Extractives in Biomass*, describes the determination of ethanol-soluble extractives as a percentage of the oven-dried biomass, which includes hard and soft woods, herbaceous materials (such as switchgrass and sericea), agricultural residues (corn stover, wheat straw, bagasse), and wastepaper (office waste, boxboard, newsprint). The extractive content of biomass is not considered to be part of the structural components of biomass and needs to be removed before any chemical analysis of the sample. Some of these extractives include waxes, fats, resins, tannins, gums, sugars, starches, and pigments. The procedure involves using a Soxhlet extraction apparatus and 190-proof ethanol to extract the sample for 24 hours. The sample slurry is filtered, and the filtrate is evaporated in a rotary vacuum evaporator. The resulting extractive weights are reported as a percentage of the original samples oven-dried weight.
   b. E 1721, *Standard Test Method for Determination of Acid-Insoluble Residue in Biomass*, can be used to determine the acid-insoluble residue of hard and soft woods, herbaceous materials, agricultural wastes, wastepaper, acid and alkaline pretreated biomass, and the solid fraction of fermentation residues. This acid-insoluble residue contains the acid-insoluble lignin and condensed proteins from the original oven-dried sample of biomass. Samples for this test can come from several sources and have separate calculation procedures; for example, prepared biomass samples are treated in accordance with practice E 1757, and extractives-free material is prepared in accordance with test method E 1690. The samples are reacted with 72% sulfuric acid (reagent grade) for hydrolysis and then diluted to 4% acid and autoclaved for an hour at 121 °C. The filtrate is removed by vacuum filtering, and the remaining insolubles are dried, weighed, and then ignited in a muffle furnace to determine the ash weight. The difference in weight of the dried insolubles minus the ash weight as a percentage of the original dry weight of the sample is defined as the acid-insoluble content of the biomass.
   c. E 1755, *Standard Test Method for Ash in Biomass*, expresses the ash content as the mass percentage of residue remaining after dry oxidation (575 °C in a muffle furnace) on the original sample dry weight. Samples of biomass can include hard and soft woods, herbaceous materials, agricultural residues, wastepaper, acid and alkaline pretreated biomass, and the solid fraction of fermentation residues.
   d. E 1756, *Standard Test Method for Determination of Total Solids in Biomass*, describes two test methods that depend on the nature of the sample. Because moisture is not considered a structural component of biomass and can change with storage and handling, it is important to analyze biomass samples and report the results on a moisture-free basis. Test method A uses a drying oven with controlled temperatures at 105 °C. Test method B uses an automated moisture analyzer with direct infrared heating elements; it should not be used for samples containing large amounts of free sugars or proteins that can degrade or caramelize.
   e. E 1757, *Standard Practice for Preparation of Biomass for Compositional Analysis*, describes a reproducible way to convert the many types of biomass feedstocks into a uniform material suitable for compositional analysis. Three methods are presented. Test method A uses field-collected samples and natural air drying to a constant weight. Test methods B and C produce dried feedstocks using convection oven and lyophilization, respectively. The samples are milled and sieved for additional analyses. Test methods B and C are used for very wet biomass samples that are at risk for mold growth during drying or that could degrade after standing for prolonged times.
   f. E 1758, *Standard Test Method for Determination of Carbohydrates in Biomass by High Performance Liquid Chromatography*, discusses the carbohydrates making up the major portion of dried biomass samples—polysaccharides constructed primarily of glucose, xylose, arabinose, galactose, and mannose subunits. These polysaccharides can be hydrolyzed to their component sugar monomers with 72% sulfuric acid in a two-stage hydrolysis process. These monosaccharides can then be quantified by ion-moderated partition HPLC using calibration sugar standards. These results can then be used with other biomass assays to determine the total composition of the biomass sample.

New Standards Being Developed. Now that the feedstock characterizations are almost complete, a new category of

standards is being developed including biomass conversion standards. Just as standards were developed to evaluate the yields, efficiency, and quality of alcohols produced by cooking and fermenting grain feedstocks, a similar opportunity applies to biomass conversion by the following methods that are currently in the research stage of development:

1. Chemical pretreatment using acid prehydrolysis
2. Biochemical conversion of biomass using liquefaction, saccharification, fermentation, simultaneous saccharification and fermentation, enzymatic hydrolysis, and bacterial conversion
3. Thermochemical conversion processes using pyrolysis and gasification

Future Trends. As research continues to uncover new biological and engineered organisms that improve the efficiency and economics to convert renewable biomass feedstocks to fuels, chemicals, drugs, and other usable products, additional standards may be justified in this rapidly expanding industry. A focused, concerted effort from the biomass research community is needed to develop these new test methods and to standardize biomass conversion processes and product characterization techniques. The ASTM E-48.05 Subcommittee on Biomass Conversion will continue to be a significant force to support this progress.

E-48.06 Subcommittee on Biotechnical Equipment Qualification and Process Validation

For biological processes producing drugs and chemicals that are under the regulatory responsibility of the FDA, there is a need to validate the biological processes and qualify the equipment used in the process. Because many processes use common equipment, the protocols used to qualify this equipment can be standardized to the point where their specific application may not be confidential or specific to the process owner. Process validation protocols can also be standardized for uniform application and acceptance by regulatory agencies. The current *Good Manufacturing Practices* (cGMPs) published by FDA call for process validation, equipment qualification, and utility qualification (12). The work of E-48.06 will assist the biotech industry to use standard test methods comparable to quality standards.

Scope and Background. Subcommittee E-48.06 was organized in October 1988 to develop standards and protocols for equipment qualification and process validation. Its charter is to develop standards required for or used in the qualification and validation of facilities, for process-related utilities, production equipment, computer controls, sampling methods, and processes that include at least one step involving living organisms.

Available ASTM Standards

1. E 1567, *Standard Guide for Biopharmaceutical Facilities Architectural Design Considerations*, covers architectural design considerations for buildings and facilities that use biological processing to produce drugs, chemicals, and other products, especially when cGMP criteria apply under Code of Federal Regulations Title 21 (2). The guide is intended for use in designing, constructing, and operating pilot plant and commercial production facilities to meet federal regulations. References are provided to existing federal regulations that apply to all general phases of a bioprocess activity.
2. The symposium *Validation Practices for Biotechnology Products* (9) brought together representatives from industry, federal agencies, research, and academia to identify more than 100 standards for development.

New Standards Being Developed. Task groups have been established to develop standards for the following issues:

1. Multiuse cleaning of process equipment
2. Computer validation protocols
3. Master plan for validation
4. Standards for installation qualification, operational qualification, performance qualification, purchase, or construction specifications
5. Qualification of utilities, filtered air, plant steam, and clean steam

Future Trends. By considering industry needs obtained from workshops, symposia, surveys, and technical committee meetings, E-48.06 will focus on protocols for fermenters, cell recovery systems, purification systems, sterile filling operations, environmental systems, sterilization equipment, and other production equipment. Computer controls validation is a high-interest topic in industry and regulatory agencies. In recent workshops (9,10) more than 100 potential standards have been identified for prioritization and development into voluntary consensus standards.

E-48.91 Subcommittee on Terminology

The charter of this subcommittee is to maintain E 1705, *Terminology Relating to Biotechnology* (2), which contains more than 130 definitions pertaining to the E-48 standards and other words relating specifically to biotechnology. As new standards are developed and approved, any new biotechnical words are added automatically to E 1705. As discussed previously, biomass standards currently include a biomass terminology standard, E 1126, *Terminology Relating to Biomass Fuels*.

BIBLIOGRAPHY

1. *ASTM Technical Committee Officer Handbook*, ASTM, June 1996, pp. 1–5.
2. *Annual Book of ASTM Standards*, vol. 11.05, ASTM, Philadelphia, Pa., 1996, pp. iii–v.

3. Technology Transfer Improvements Act of 1995, H.R. 2196, Provision 12(d), Public Law 104-113.
4. *Congressional Record—Senate*, February 7, 1996, p. 1081.
5. *Regulations Governing ASTM Technical Committees*, ASTM, July 1995, pp. 5–13.
6. *Form and Style for ASTM Standards*, ASTM, West Conshohocken, Pa., January 1996, p. vi.
7. *Symposium on Bioprocessing Safety Standard Guidelines and Practices to Insure Work Safety in the Bioprocessing of Industrial Chemicals, Foods, and Waste Products*, ASTM E-48 Symposium, October 1987, STP no. 1051.
8. *ASTM 1996 Directory*, ASTM, Philadelphia, Pa., pp. 40, 114.
9. *Validation Practices for Biotechnology Products*, ASTM E-48 Symposium, April 1995, STP no. 1260.
10. Conference Report, *J. Res. Natl. Inst. Standards Technol.* **99** (Jan.–Feb. 1994).
11. N.A. Roscioli, C.A. Renshaw, A.A. Gilbert, C.F. Kerry, and P.G. Probst, *BioPharm* 32–40 (Sept. 1996).
12. *ASTM Standardization News*, August 1989, pp. 44–49.

See also GOOD MANUFACTURING PRACTICE (GMP) AND GOOD INDUSTRIAL LARGE SCALE PRACTICE (GLSP); PROCESS VALIDATION.

ATTACHMENT FACTORS

DERYK T. LOO
Bristol-Myers Squibb Pharmaceutical Research Institute
Seattle, Washington

KEY WORDS

Cadherins
Cell adhesion molecules
Collagen
Fibronectin
Glycosaminoglycan
Laminin
Selectins
Tenascin
Thrombospondin
Vitronectin

OUTLINE

INTRODUCTION

The development and maintenance of tissues in vivo is acutely dependent on the ability of cells to form connections with substrates and neighboring cells. The formation of intimate contacts with various substrates provides sites of attachment and the scaffolding necessary for cells to communicate and develop cooperative function. These substrates, or attachment factors, include extracellular matrix (ECM) and basement membrane molecules such as collagen, laminin, fibronectin, and vitronectin. Additional substrates include proteins expressed on the surface of cells that function as attachment factors and mediate cell–cell interactions. Cell culture provides a powerful tool with which to identify attachment factors involved in cellular adhesion and to define the molecular mechanisms regulating cell adhesion. Additionally, in vitro approaches play a key role in the identification and cloning of specific attachment factor receptors. The cloning and characterization of attachment factor receptors has forced us to redefine the role of attachment factors in cellular functions. Growing in vitro evidence, combined with in vivo studies, make it clear that attachment factors serve not only as molecular scaffolds to support cell attachment and movement, but also as signaling molecules that play pivotal roles in regulating numerous cellular functions including growth, activation, differentiation, and death.

I review here the ECM molecules that are widely used as attachment factors in culture and illustrate ways in which attachment factors regulate cellular function. Additionally, cell surface molecules that serve as attachment factors and facilitate cell–cell interactions in vivo are also reviewed.

EXTRACELLULAR MATRIX MOLECULES

Fibronectin

Biochemical Properties. Fibronectins are glycoproteins consisting of two disulfide-linked subunits with a molecular weight of 220,000–250,000 kDa. Fibronectin was originally purified to homogeneity in 1970 by Mosesson and Umfleet from plasma cryoprecipitate and termed cold-insoluble globulin (CIg) (1). Later, fibronectin was detected as a cell surface protein on fibroblasts whose expression was lost upon transformation (2,3). In addition to existing as a disulfide-linked dimer, cell surface–associated fibronectin also exists as higher molecular weight disulfide-linked aggregates that are secreted by cells and accumu-

late as a major component of the basement membrane and ECM in vivo.

The primary structure of fibronectin consists of a series of three different types of modular units that combine to generate the functional domains of the protein. The two subunits of fibronectin are connected by a pair of interchain disulfide linkages at the carboxy-terminal end. The intact dimer appears as an extended, flexible molecule consisting of two linear arrays of globular domains. The identification and characterization of the individual globular domains has resulted largely from examination of proteolytic fragments of fibronectin, many of which retain biological function. The N-terminal domain of fibronectin binds fibrin (4–6), heparin (6–8), *Staphylococcus aureus* (8–11), thrombospondin (12), factor XIIIa transglutaminase (4,11), and IgG (13). Adjacent to the amino-terminal domain of fibronectin is a domain that binds gelatin and collagen (7,14). This interaction is important for localization of fibronectin to basement membranes and ECMs. The collagen-binding domain also binds to the C1q component of complement via the collagenous tail of C1q (15). Adjacent to the gelatin-binding domain of fibronectin is the cell-binding domain. Pierschbacher et al. (16) showed that the cell-binding activity of fibronectin resided within a 12-kDa proteolytic fragment that contains the tetrapeptide sequence Arg-Gly-Asp-Ser (RGDS) (17). Peptides containing soluble synthetic Arg-Gly-Asp (RGD) were shown to detach cultured cells and platelets from fibronectin substrates (18,19), implicating the RGDS sequence in cell binding.

In addition to binding collagen, fibronectin has an affinity for several other components of the ECM, including the glycosaminoglycans heparin, heparan sulfate and chondroitin sulfate proteoglycan, and hyaluronic acid (6,20–25). The interaction of fibronectin with these various components of the ECM is thought to play a role in the formation and maintenance of the ECM.

The high-affinity cell surface receptors for fibronectin are the integrins, which consist of heterodimeric glycoprotein α- and β-subunits of approximately 140 and 120 kDa that bind to the RGD sequence of fibronectin (26). The extracellular domains mediate cell adhesion via binding to fibronectin and other ECM proteins. The cytoplasmic domains are responsible for transmitting signals to the cell interior following ligand engagement (outside-in signaling), and for modulating the affinity of the integrin for its ligand by interacting with intracellular signaling molecules (inside-out signaling). The cytoplasmic domains also interact with the actin-binding proteins filamin (27,28) and α-actinin (29), thereby connecting the cell membrane with the cytoskeleton. These interactions play a role in the formation of focal adhesions and in phagocytosis.

Fibronectin plays a central role in vertebrate development. Fibronectin matrices are utilized by migrating cells during embryogenesis (30–32). The migration of avian, amphibian, and mammalian neural crest cells is blocked by RGD-containing peptides and by antibodies to fibronectin or its cell surface receptor (33,34). Furthermore, inactivation of the fibronectin gene in mice leads to early embryonic lethality (35). Its role in development and in maintenance of the adult organism is due to its diverse influences on cell growth, adhesion, migration, and differentiation.

Role in Cell Adhesion. Fibronectin promotes the attachment of multiple cell types, including epithelial cells, follicular cells, ovary cells, neuronal cells, fibroblasts, keratinocytes, and adult liver cells. As previously mentioned, fibronectin is not solely an attachment factor for cells. In many cases, cell attachment to fibronectin is required for cell survival and promotion of cell growth. Fibronectin promotes cell division of numerous cell types, including follicular cells (36), primary chick fibroblasts (37), murine and human fibroblasts (38), neuroblastoma cells (39), keratinocytes (40), rat granulosa cells (41), and lymphocytes (42). Fibronectin also promotes DNA synthesis by serum-deprived quiescent normal hamster fibroblasts (43).

In addition to regulating cell division, fibronectin has been shown to modulate the differentiation of cells in culture. Fibronectin coating of culture surfaces inhibits morphological changes and biosynthesis of lipogenic enzymes associated with the differentiation of preadipocyte fibroblasts (44,45). Fibronectin also attenuates hormone-induced expression of cytoskeletal proteins associated with granulosa cell differentiation (46). Additionally, exogenously added fibronectin inhibits both spontaneous and chemically induced differentiation of cultured normal human keratinocytes in vitro (47,48). Alternatively, fibronectin induces the differentiation of neural crest cells, neuroepithelial embryonal carcinoma cells (F9), and neuroblastoma cells in vitro (49–51), consistent with observations showing that receptors for fibronectin are found on neural crest cell derivitives colocalized with fibronectin (52) in vivo. Fibronectin also promotes the differentiation of primary human hepatocytes, as measured by the expression of albumin and the liver-specific enzyme cytochrome P450 (53). It has been suggested that modulation of cell growth and differentiation by fibronectin is due in part to its role in regulating the cytoskeletal architecture of cells (36,50). The recent cloning of fibronectin receptors and identification of signaling pathways coupled to these receptors has provided evidence that specific signaling events play an important role in controlling cell growth and differentiation.

Collagen

Biochemical Properties. Collagen is the most abundant component of the ECM and has been used extensively as an attachment factor for cultured cells. Collagen is composed of three helical α-chains that are wrapped around each other in a triple helical arrangement. In some forms of collagen the three chains are identical, whereas in other forms the chains differ from one another. The α-chains are composed of a series of triplet Gly-X-Y residues, where X and Y may be any amino acid. Frequently, X is a proline and Y is a hydroxyproline. Alternatively, hydroxylysine may also occupy the X or Y position and contribute to interchain cross-linking of the collagen α-chains. Collagen is secreted from cells as procollagen, with each of the α-chains containing large amino- and carboxy-terminal nonhelical domains. Although some collagens exist in vivo as

intact procollagen, many forms of collagen undergo proteolytic removal of the amino- and carboxy-terminal nonhelical domains.

To date, more than 14 collagen types have been described, which are generally subgrouped into six classes defined by the structural forms in which they assemble within the ECM. Several forms of collagen display ordered fibrous structures. For example types I, II, III, V, and XI collagen form fibrillar structures that contribute to connective tissues such as cartilage, skin, bone, and tendon, whereas type IX and XII collagens are fibril associated and interact with type I and II collagen fibrils. Type IV collagen is a nonfibrillar collagen and is the principle collagen type found in basement membrane. Type IV collagen differs from the fibrillar collagens by its existence in tissues in an unprocessed procollagen form (54,55). Other collagen types include filamentous collagen (type VI), short-chain collagens (types VIII and X), and long-chain collagen (type VII).

Collagen matrices are associated with many cell attachment proteins in vivo, including fibronectin and laminin. Thus the role of collagen as an attachment factor may be divided into two types: (*1*) a direct mechanism whereby collagen serves as the substrate for cell attachment (VLA-2; collagen receptor); and (*2*) an indirect mechanism whereby a secondary attachment factor, through its interaction with collagen, serves as the substrate for cell attachment.

Role in Cell Adhesion. The principle collagens used as attachment factors in cell culture are the fibrillar collagens and the nonfibrillar type IV collagen. Fibrillar collagen isolated from connective tissues was the first collagen used in cell culture. A mouse tumor source of type IV collagen was subsequently identified, which produced large amounts of this collagen type (56,57). Collagen substrates serve as an attachment and growth-promoting factor for a variety of cells in culture. Collagen promotes the sustained growth of epithelial cells from various tissues including mammary epithelial cells (58–61), endometrial epithelial cells (62), vaginal epithelial cells (63), esophageal epithelial cells (64), and liver epithelial cells (65). Collagen substrates also promote the extended viability of primary chondrocytes (66). Chondrocytes exhibit three-dimensional growth within collagen gels, deposit ECM, and incorporate or rearrange exogenous collagen into newly formed matrix, suggesting the formation of tissuelike structures.

In addition to providing attachment and growth-promoting activity, collagen also modulates the differentiation state of cells in culture. Collagen induces morphological and biochemical characteristics associated with the differentiated state of various epithelial cell types including human endometrial cells (62), trachial epithelial cells (67), uterine epithelial cells (68), mammary epithelial cells (69), and intestinal epithelial cells (70). Collagen also promotes the differentiation of adrenocortical cells (71,72), osteoblasts (73), keratinocytes (74), hepatocytes (75–79), liver dendritic cell progenitors (80), thyroid follicle cells (81), smooth muscle cells (82), and Sertoli cells (83). Regulation of growth and differentiation, which is accompanied by alterations in gene expression, implicate collagen as an active signal-transducing molecule. Candidate receptors for collagen that may play a role in cell signaling include the integrins VLA-1 (84,85), VLA-2 (86–88), and VLA-3 (86–88).

Laminin

Biochemical Properties Laminin is the first extracellular protein expressed during development and plays a critical role in cellular development and tissue organization through modulation of cellular attachment, motility, growth, and differentiation. Laminin is a large glycoprotein that is a major component of basement membranes. The prototypical laminin molecule is composed of three subunits: a 440-kDa A-chain, a 200-kDa B1-chain, and a 220-kDa B2-chain. These are joined together by disulfide linkages to form a cruciform-shaped structure possessing multiple globular and rodlike domains that are responsible for its cell-binding activity and its interaction with a variety of ECM-associated molecules. The terminal end of the long arm, which contains a large, multilobed globular domain and forms the base of the cruciform, is responsible for receptor-mediated cell attachment, promotion of neurite outgrowth, and heparin binding (89). Regions within the short B-chain arms provide a second cell attachment site and also play a role in cell signaling and binding of the glycoprotein nidogen/entactin. Similar to fibronectin, laminin also binds to bacteria (90) and glycosaminoglycans (91).

Laminin promotes cell attachment through interaction with specific cell surface receptors. The most extensively studied laminin receptors are the integrins. Laminin binds to at least four VLA integrins: $\alpha 1\beta 1$ (92,93), $\alpha 2\beta 1$ (87,94), $\alpha 3\beta 1$ (95), and $\alpha 6\beta 1$ (92,96). Laminin also binds to the $\alpha V\beta 3$ integrin (97) and $\alpha 6\beta 4$ integrin (98,99). The interaction of laminin with integrins is believed to occur through a non-RGD-mediated mechanism, although RGD-containing peptides can inhibit cell adhesion to the second cell adhesion site located within the central portion of the laminin cruciform (96). Other recently identified receptors for laminin include the 67-kDa laminin receptor, which has been implicated in tumor metastasis (100,101), and a 110-kDa laminin-binding protein from brain, which appears to be immunologically and biologically related to the β-amyloid precursor protein (APP) family (102). The interaction of the 110-kDa laminin-binding protein with the Ile-Lys-Val-Ala-Val (IKVAV) residues of the laminin A-chain plays a role in neurite outgrowth (102). A distinct 110-kDa protein, termed nucleolin, was copurified with the APP-related 110-kDa protein and has also been identified as a laminin-binding protein (103) that may play a role in ECM signaling. Laminin also binds to the actin-binding cell surface protein connectin (104) and the 67-kDa elastin/laminin receptor (105).

Role in Adhesion. Laminin plays an important role in the attachment, growth, and differentiation of many cell types in culture, including neural and epithelial cells. When added to the culture medium, laminin promotes Schwann cell attachment, growth, and the formation of a stellate, process-bearing morphology (106) and regulates glycolipid synthesis (107). Laminin promotes attachment

and neurite outgrowth of sensory neurons (108) and induces the differentiation of neuroblastoma cells (51). Antibodies directed against the heparin-binding domain of laminin reduce neuronal viability and inhibit neurite outgrowth in vitro, indicating that the heparin-binding domain is responsible for the effects of laminin on neurite outgrowth and neuronal survival (109). Laminin also modulates the proliferation of cultured astrocytes (110) and oligodendrocytes (111). Laminin promotes the differentiation of cultured endothelial cells to form capillary-like structures (112). Laminin also induces the differentiation of various epithelial cells, as measured by β-casein production by cultured mammary epithelial cells (113,114), the formation of cordlike structures by Sertoli cells (115), and the expression of alkaline phosphatase and lactase activity by intestinal epithelial cells (116). Laminin promotes the attachment and survival of F9 embryonal carcinoma cells (117), while it potentiates the differentiation of PCC4uva embryonal carcinoma cells to neurons following treatment with retinoic acid and dibutyryl cyclic AMP (118). Laminin also promotes the differentiation of granulosa cells (119) and hepatocytes (120,121).

Vitronectin

Biochemical Properties. Vitronectin, or serum-spreading factor, is a glycoprotein that is present in the blood and in other tissues. Vitronectin was initially described as a cell-spreading and growth-promoting α1-glycoprotein enriched in a fraction of human serum isolated by glass bead chromatography (122). The spreading–promoting activity of this fraction was further purified from this fraction by Barnes et al. (58) and was shown to be biochemically and immunologically distinct from fibronectin and laminin (123,124). Vitronectin exists in two distinct forms in human serum: a 75-kDa single-chain form and a 65-kDa two-chain form (125,126). The 65-kDa form is the product of proteolytic processing and lacks the carboxy-terminal 10-kDa fragment (127). Sensitivity to proteolytic processing may result from allelic differences in the vitronectin gene (128). The gene encoding vitronectin was identified by expression cloning (127), and sequence analysis revealed a 1,545-bp open reading frame corresponding to the full-length plasma vitronectin. Sequence analysis of an independently isolated serum protein associated with complement, termed S-protein, revealed it to be identical to vitronectin (129).

The major source of circulating vitronectin is the liver (130), although it is also produced by various hematopoetic cell types including platelets, monocytes, and macrophages (131,132). Vitronectin is also associated with tissues (124,133). In situ hybridization analysis indicates that vitronectin mRNA is expressed early in mammalian development primarily in the liver and central nervous system (134), suggesting that it plays an important role in mammalian development. In the mouse, vitronectin mRNA is detected by day 10 in the liver and central nervous system. Expression of vitronectin mRNA in the central nervous system is first detected in the floor plate; later, vitronectin mRNA can be observed in the meninges of the cortex and the spinal cord. Vitronectin mRNA is associated with the vasculature of the central nervous system but not with blood vessels of peripheral tissues, suggesting that vitronectin may play a specific role in vascular function in the central nervous system. Surprisingly, mice deficient in vitronectin gene expression demonstrate normal development and survival, indicating that the functional role of vitronectin in development may be compensated for by other molecules (135).

The vitronectin gene is composed of eight exons and seven introns (136), which define the domain structure of the mature protein. The amino-terminal cysteine-rich 44-residue domain is identical to somatomedin B (137). The second domain contains an RGD sequence that, analogous to fibronectin, is responsible for cell attachment via a number of integrins (137,138) and also contributes to collagen binding (139,140). The second domain is also implicated in binding to plasminogen-activator inhibitor-1, and consistent with its localization to platelets, this interaction plays an important role in blood clotting. Vitronectin also binds to heparin (125), and a domain located toward the carboxy-terminal region of vitronectin is responsible for the heparin-binding activity of vitronectin (137,141), which is dependent on the conformational state of vitronectin. Although native plasma vitronectin binds heparin weakly, vitronectin denatured by 8 M urea or other denaturing agents has an increased affinity for heparin (125,141). Denaturation is hypothesized to expose a cryptic heparin-binding domain and suggests that the biological functions of vitronectin are dependent on its conformation. Recently a form of vitronectin distinct from plasma vitronectin has been identified within platelet α-granules, which is able to bind heparin (142) and may play a role in platelet function at sites of vascular injury. Additionally, similar to collagen and fibronectin, vitronectin binds to bacteria, including *S. aureus*, *Escherichia coli*, *S. pneumoniae*, and *Neisseria gonorrhoeae* (143–147), implicating vitronectin in the adherence and phagocytosis of bacteria by host immune defense cells.

Role in Adhesion. Based on the tissue distribution of vitronectin, it is not surprising that vitronectin promotes the attachment and growth of a wide variety of cell types in vitro, including endothelial cells, lymphoid cells,and neural cells (58). Human endothelial cells attach to vitronectin, adopt a flattened morphology, and exhibit increased viability (148). Vitronectin promotes the attachment and proliferation of IL-3-dependent mast cells (149). However, vitronectin is unable to replace IL-3 as a survival factor, indicating that vitronectin is not itself a survival factor for bone marrow–derived mast cells but is able to augment the IL-3-dependent signal in mast cells. Vitronectin promotes the attachment of myeloblast cells through the αv-containing integrins (150) and stimulates megakaryocyte proplatelet formation through interaction with αv/β3 integrin (151). Vitronectin also promotes the attachment and long-term survival of cultured glioma cells (152), and this interaction may be mediated by gangliosides (153). In addition, vitronectin mediates attachment and nerve growth factor–dependent neurite outgrowth of rat peochromocytoma cells via an RGD-dependent mechanism (154).

Other Attachment Factors

In addition to collagen, fibronectin, laminin, and vitronectin, other components of the ECM and basement membrane also contribute to cell attachment. Thrombospondin is a large trimeric glycoprotein consisting of three identical 140-kDa subunits joined together through interchain disulfide linkages (155). Thrombospodin is released by activated platelets and is found associated with the ECM proteins heparin, collagen, laminin, and fibrinogen. Thrombospondin is synthesized by human long-term bone marrow cells in vitro, promotes the adhesion of hematopoietic progenitor cells (156), and promotes the attachment and migration of monocytes and neutrophils (157). Receptors for thrombospondin that play a role in its function as an attachment factor include membrane-bound heparan sulphate (158), platelet glycoprotein IV (CD36; GPIIIb) (159), and αv/β3 integrin (160).

Tenascin is a very large, star-shaped ECM glycoprotein composed of six related subunits joined through interchain disulfide linkages. The individual subunits contain 13 epidermal growth factor–like repeats, 8 or more fibronectin type III repeats, and a globular carboxy-terminal fibronectin-like domain. Tenascin is associated with mesenchymal–epithelial interactions during development, tissue remodeling, and wound healing. Tenascin promotes the attachment of many cell types in vitro, including tumor cells, fibroblasts, and endothelial cells, in an RGD-dependent manner (161) through interaction with multiple integrins (162). Janusin, an ECM protein with structural homology to tenascin, promotes attachment and neurite outgrowth of hippocampal neurons in vitro (163).

Glycosaminoglycans are a major component of the ECM and, as already mentioned, bind to fibronectin (164), collagen (165), laminin (166), and vitronectin (127). Glycosaminoglycans consist of long polysaccharide chains made up of disaccharide subunits and, with the exception of hyaluronan (HA), are sulfated and covalently attach to core proteins through serine residues. In addition to playing a role in tissue formation during development and remodeling, glycosaminoglycans also contribute to the process of inflammation and tumorigenesis (167). One glycosaminoglycan that has been widely studied is HA. The role of HA in various cellular functions including cell proliferation, cell activation, and cell migration is dependent on its interaction with specific cell surface glycoprotein CD44 (168–170). Molecular cloning of the CD44 gene revealed the presence of multiple isoforms, which result from alternate splicing. Analysis of the expression pattern of the different isoforms suggests that regulation of expression of specific isoforms plays a critical role in cellular interactions during early development (171). The CD44–HA interaction is involved in endothelial cell adhesion and proliferation (172) and T-cell adhesion and activation (173). CD44 has also been implicated in the regulation of tumor growth and metastasis, and this activity is dependent on its interaction with HA (174).

CELL SURFACE ATTACHMENT MOLECULES

In addition to molecules associated with the ECM and the basement membrane, cell surface receptors also serve as attachment factors to promote cell–cell interactions. Cell surface receptors implicated in cell attachment include the cell adhesion molecule family, the selectin family, and the cadherin family.

Cell Adhesion Molecules

The cell adhesion molecules, or CAMs, comprise a family of cell surface receptors that are members of the immunoglobulin gene superfamily. Members of this gene superfamily are characterized by the presence of multiple tandem immunoglobulin fold domains. Vascular cell adhesion molecule 1 (VCAM-1) is present on most peripheral blood leukocytes, activated endothelial cells, macrophages, dendritic cells, bone marrow fibroblasts, and myoblasts (175–177) and plays an important role in the recruitment of leukocytes to sites of inflammation. At sites of inflammation, activated endothelial cells express VCAM-1, which facilitates the adhesion and subsequent transmigration of circulating leukocytes. The β1 integrin VLA-4 (CDα4/β1) (178) and the β7 integrin (CDα4/β7) (179) are ligands for VCAM-1. Similar to VCAM-1, intercellular adhesion molecules (ICAM-1, -2, -3) are expressed on most circulating blood leukocytes, activated endothelial cells, fibroblasts, keratinocytes, chondrocytes, and epithelial cells (180,181) and play a role in immune and inflammatory responses by mediating cell–cell interactions. The ligands for ICAM-1 and ICAM-2 include the integrins LFA-1 (CD18/CD11a) (182) and MAC-1 (CD18/CD11b) (183,184), and CD43 (185). Neural cell adhesion molecule (NCAM) is expressed on neurons, astrocytes, Schwann cells, myoblasts, natural killer (NK) cells, and activated T lymphocytes. NCAM is thought to play an important role in vertebrate development through modulation of cell–cell interations between neurons, astrocytes, oligodendrocytes, and myoblasts (186). In addition to homotypic interactions, NCAM binds to heparan sulfate and heparin glycosaminoglycans (187). Two related members, L1-CAM and Ng-CAM, are expressed on postmitotic, premigratory neurons and play a role in neuronal migration and neurite outgrowth. Ligands for L1-CAM include axonin 1 (188) and αVβ3 integrin (189). Ng-CAM also binds to axonin 1 (190,191). Neurite outgrowth is stimulated by the interactions of L1-CAM and Ng-CAM with axonin I (188) and requires calcium influx through L- and N-type calcium channels (192). The condroitin sulfate proteoglycan neurocan also binds to NCAM, Ng-CAM, and L1-CAM and inhibits neuronal adhesion and neurite outgrowth (193). Platelet endothelial cell adhesion molecule (PECAM-1, CD31) is expressed on platelets, endothelial cells, monocytes, granulocytes, and a subset of T cells (194,195). Exogenous expression of PECAM-1 on mouse L cells induces cell aggregation both in a PECAM-1-dependent, homophilic manner (196), and a calcium-dependent heterophilic manner through interaction with heparin and chondroitin sulfate (197). Recently, the αv/β3 integrin has been demonstrated to be a ligand for PECAM-1 (198).

Selectins

The selectin family of cell surface adhesion molecules function specifically in leukocyte–endothelial cell adhesion. Se-

lectins have a multidomain structure consisting of an amino-terminal carbohydrate-binding lectin domain, an epidermal growth factor–homologous domain, and a number of cysteine-rich domains homologous to the consensus repeats contained within complement regulatory proteins. Three members of the selectin family have been cloned and characterized, to date. L-selectin (CD62L) is expressed by leukocytes, E-selectin (CD62E) is found on endothelial cells activated by inflammatory mediators, and P-selectin (CD62P) is found on platelets and activated endothelial cells (199). The E- and P-selectins are implicated in the initial adhesion and rolling of leukocytes on activated vascular endothelial cells that line blood vessels (200,201). Studies showing that monoclonal antibodies directed against the lectin domain of selectins blocked cell adhesion have implicated carbohydrates as ligands for selectins. Currently, a large number of sialylated, sulfated, and/or fucosylated carbohydrates have been identified as ligands for selectins (199,202). In addition to their adhesive properties, selectins activate leukocytes, causing intracellular calcium release, induction of IL-8 and TNF-α mRNA transcription, and activation of β2 integrin–dependent strong adhesion (203).

Cadherins

The cadherin family of cell adhesion molecules are calcium-dependent membrane glycoproteins that are grouped into nine types based on structural and functional differences (204). Type I, or classical cadherins (E-, P-, N-, B-, R-, EP-cadherin), are the only cadherins that have been conclusively shown to promote cell adhesion. Type I cadherins contain five extracellular domains of approximately 110 amino acids each, a single transmembrane domain, and two cytoplasmic domains. The extracellular domains contain Asp-X-Gln-Asp-Gln-Asp and Asp-X-Asp amino acid motifs, which are responsible for calcium binding. The amino-terminal extracellular domain contains the cell adhesion recognition sequence His-Ala-Val (HAV), which mediates calcium-dependent homotypic and heterotypic interactions between cadherins. The cytoplasmic domains of type I cadherins interact with a family of intracellular proteins termed catenins (105). The catenins in turn bind to microfilaments and serve to connect the cadherins to the cytoskeleton. The interaction of catenins with the cytoplasmic domain of cadherins is thought to play an important role in cadherin adhesive function.

Cadherin interactions regulate the formation of adherens junctions and tight junctions between cells. In addition to mediating cell–cell contact, cadherins also regulate cell morphology and differentiation. E-cadherin, which has been classically associated with the formation, differentiation, and polarization of developing epithelia, has recently been shown to regulate erythropoietin-mediated differentiation of bone marrow mononuclear cells (205). Neural cadherin, or N-cadherin, promotes neurite outgrowth, which can be blocked by HAV-containing peptides (206). N-cadherin also regulates the adhesion of O-2A-lineage glial progenitor cells in culture (207).

FUTURE DIRECTIONS

It has become clear that cell–cell and cell–substrate interactions mediated through adhesion molecules and their corresponding receptors play critical roles in cellular functions including growth, activation, differentiation, and death. One of the challenges facing researchers will be to define the molecular mechanisms triggered by the interaction between attachment factors and cell surface receptors that regulate cellular function. Recent advances in molecular biology have provided powerful tools by which to identify both extracellular and intracellular components of these complex signaling systems. Soluble, chimeric proteins containing the extracellular regions of attachment factors and cell surface receptors have been used to identify and clone novel interacting partners, to define domains and individual residues that participate in binding, and to examine the biological role of the attachment factor–cell surface receptor interaction. In addition to being a tool for studying the biochemistry and biology of attachment factor–cell surface interactions, soluble chimeric proteins are also being pursued as novel drug candidates. The ability of soluble, chimeric proteins to mimic their native counterparts may allow for the generation of effective biological-based drugs that can modulate protein–protein, or cell–cell interactions and their downstream signaling events.

In addition to defining extracellular interations, the development of powerful molecular biology approaches has provided novel methods by which to identify protein–protein interactions that occur inside cells. The yeast two-hybrid system has been used to identify signaling proteins that interact with the cytoplasmic domains of various attachment factor cell surface receptors. This system has been used to identify several proteins that play important roles in integrin function, including the cytoskeletal protein filamin (27,28), cytohesin 1 (208), integrin-linked kinase (209), and β3 endonexin (210). This system will continue to provide a means of identifying molecules that interact with the cytoplasmic domain of attachment factor cell surface receptors and participate in attachment factor–cell surface receptor interactions. These approaches, together with the current explosion in the identification and cloning of novel attachment factor receptors through use of genomic-based bioinformatics, will provide an exciting opportunity for interface between attachment factor biologists, molecular and cell biologists investigating cell signaling, and bioinformaticians in the development of the biology of cell attachment.

BIBLIOGRAPHY

1. M.W. Mosesson and R.A. Umfleet, *J. Biol. Chem.* **245**, 5728–5736 (1970).
2. R.O. Hynes, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3170–3174 (1973).
3. C.G. Gahmberg and S. Hakomori, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3329–3333 (1973).
4. H. Hormann and M. Seidl, *Hoppe Seyler's Z. Physiol. Chem.* **361**, 1449–1452 (1980).

5. M. Seidl and H. Hormann, *Hoppe Seyler's Z. Physiol. Chem.* **364**, 83–92 (1983).
6. K. Sekiguchi and S. Hakomori, *Biochemistry* **22**, 1415–1422 (1983).
7. K. Sekiguchi and S. Hakomori, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2661–2665 (1980).
8. M. Hayashi and K.M. Yamada, *J. Biol. Chem.* **257**, 5263–5267 (1982).
9. P. Kuusela, *Nature* **276**, 718–720 (1978).
10. P. Kuusela, T. Vartio, M. Vuento, and E.B. Myhre, *Infect. Immun.* **50**, 77–81 (1985).
11. D.F. Mosher and R.A. Proctor, *Science* **209**, 927–929 (1980).
12. G.A. Homandberg and J. Kramer-Bjerke, *Thromb. Res.* **48**, 329–335 (1987).
13. A.A. Rostagno, B. Frangione, and L.I. Gold, *J. Immunol.* **143**, 3277–3282 (1989).
14. D.D. Wagner and R.O. Hynes, *J. Biol. Chem.* **254**, 6746–6754 (1979).
15. D.H. Bing, S. Almeda, H. Isliker, J. Lahav, and R.O. Hynes, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 4198–4201 (1982).
16. M.D. Pierschbacher, E. Ruoslahti, J. Sundelin, P. Lind, and P.A. Peterson, *J. Biol. Chem.* **257**, 9593–9597 (1982).
17. M.D. Pierschbacher and E. Ruoslahti, *Nature* **309**, 30–33 (1984).
18. E.F. Plow, M.D. Pierschbacher, E. Ruoslahti, G.A. Marguerie, and M.H. Ginsberg, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 8057–8061 (1985).
19. E.G. Hayman, M.D. Pierschbacher, and E. Ruoslahti, *J. Cell Biol.* **100**, 1948–1954 (1985).
20. M.E. Perkins, T.H. Ji, and R.O. Hynes, *Cell* **16**, 941–952 (1979).
21. K.M. Yamada, D.W. Kennedy, K. Kimata, and R.M. Pratt, *J. Biol. Chem.* **255**, 6055–6063 (1980).
22. M. Isemura, Z. Yosizawa, T. Kiode, and T. Ono, *J. Biochem.* **91**, 731–734 (1982).
23. A. Oldberg and E. Ruoslahti, *J. Biol. Chem.* **257**, 4859–4863 (1982).
24. D.E. Smith and L.T. Furcht, *J. Biol. Chem.* **257**, 6518–6523 (1982).
25. J. Calaycay et al., *J. Biol. Chem.* **260**, 12136–12141 (1985).
26. M.A. Schwartz, M.D. Schaller, and M.H. Ginsberg, *Annu. Rev. Cell Dev. Biol.* **11**, 549–599 (1995).
27. D.T. Loo, S.B. Kanner, and A. Aruffo, *J. Cell Biol.* **273**, 23304–23312 (1998).
28. C.P. Sharma, R.M. Ezzell, and M.A. Arnaout, *J. Immunol.* **154**, 3461–3470 (1995).
29. C.A. Otey, F.M. Pavalko, and K. Burridge, *J. Cell Biol.* **111**, 721–729 (1990).
30. D. Newgreen and J.P. Thiery, *Cell Tissues Res.* **211**, 269–291 (1980).
31. J.L. Duband, S. Rocher, W.T. Chen, K.M. Yamada, and J.P. Thiery, *J. Cell Biol.* **102**, 160–178 (1986).
32. B.W. Mayer, E.D. Hay, and R.O. Hynes, *Dev. Biol.* **82**, 267–286 (1981).
33. J.C. Boucaut, T. Darribere, H. Boulekbache, and J.P. Thiery, *Nature* **307**, 364–367 (1984).
34. M. Bronner-Fraser, *Dev. Biol.* **117**, 528–536 (1986).
35. E.L. George, E.N. Georges-Labouesse, R.S. Patel-King, H. Rayburn, and R.O. Hynes, *Development* **119**, 1079–1091 (1993).
36. J. Orly and G. Sato, *Cell* **17**, 295–305 (1979).
37. J.R. Couchman, D.A. Rees, M.R. Green, and C.G. Smith, *J. Cell Biol.* **93**, 402–410 (1982).
38. P.B. Bitterman, S.I. Rennard, S. Adelberg, and R.G. Crystal, *J. Cell Biol.* **97**, 1925–1932 (1983).
39. J.E. Bottenstein and G.H. Sato, *Exp. Cell Res.* **129**, 361–366 (1980).
40. B.A. Gilchrest, R.E. Nemore, and T. Maciag, *Cell. Biol. Int. Rep.* **4**, 1009–1016 (1982).
41. P. Morley, D.T. Armstrong, and R.E. Gore-Langton, *J. Cell Physiol.* **132**, 226–236 (1987).
42. H.G. Klingemann and F.R. Kohn, *J. Leukoc. Biol.* **50**, 464–470 (1991).
43. M.J. Humphries and S.R. Ayad, *Nature* **305**, 811–813 (1983).
44. T.E. Broad and R.G. Ham, *Eur. J. Biochem.* **135**, 33–39 (1983).
45. B.M. Spiegelman and C.A. Ginty, *Cell* **35**, 657–666 (1983).
46. A. Ben-Ze'ev and A. Amsterdam, *J. Biol. Chem.* **262**, 5366–5376 (1987).
47. L. Staiano-Coico and P.J. Higgins, *Exp. Cell Res.* **201**, 126–136 (1992).
48. F.M. Watt, M.D. Kubler, N.A. Hotchin, L.J. Nicholson, and J.C. Adams, *J. Cell Sci.* **106**, 175–182 (1993).
49. M. Sieber-Blum, F. Sieber, and K.M. Yamada, *Exp. Cell Res.* **133**, 285–295 (1981).
50. A. Rizzino, *Dev. Biol.* **95**, 126–136 (1983).
51. T. Kidowaki, C.J. Thiele, H.K. Kleinman, and M.A. Israel, *Pathobiology* **59**, 316–323 (1991).
52. A.M. Sheppard, M.D. Onken, G.D. Rosen, P.G. Noakes, and D.C. Dean, *Cell Adhes. Commun.* **2**, 27–43 (1994).
53. C. Trautwein, M. Davies, E. Elias, A. Strain, and M.P. Manns, *J. Hepatol.* **22**, 50–56 (1995).
54. J.G. Heathcote, C.H. Sear, and M.E. Grant, *Biochem. J.* **176**, 283–294 (1978).
55. J.G. Heathcote, A.J. Bailey, and M.E. Grant, *Biochem. J.* **190**, 229–237 (1980).
56. R. Timpl, H. Wiedemann, V.v. Delden, H. Furthmayer, and K. Kuhn, *Eur. J. Biochem.* **120**, 203–211 (1981).
57. I. Oberbaumer, H. Wiedemann, R. Timpl, and K. Kuhn, *EMBO J.* **1**, 805–810 (1982).
58. D. Barnes, R. Wolfe, G. Serrero, D. McClure, and G. Sato, *J. Supramol. Struct.* **14**, 47–63 (1980).
59. J. Yang et al., *Proc. Natl. Acad. Sci. U.S.A.* **76**, 3401–3405 (1979).
60. J. Yang, R. Guzman, J. Richards, and S. Nandi, *In Vitro* **16**, 502–506 (1980).
61. J. Richards, D. Pasco, J. Yang, R. Guzman, and S. Nandi, *Exp. Cell Res.*. **146**, 1–14 (1983).
62. A. Akoum, et al., *J. Reprod. Med.* **41**, 555–561 (1996).
63. T. Iguchi, F.D. Uchima, P.L. Ostrander, and H.A. Bern, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3743–3747 (1983).
64. M.S. Babcock, M.R. Marino, W.T. Gunning, and G.D. Stoner, *In Vitro* **19**, 403–415 (1983).
65. L. Yang, R.A. Faris, and D.C. Hixson, *Gastroenterology* **104**, 840–852 (1993).
66. N. Yasui, S. Osawa, T. Ochi, H. Nakashima, and K. Ono, *Exp. Cell Biol.* **50**, 92–100 (1982).
67. E.A. Davenport and P. Nettesheim, *Exp. Cell Res.* **223**, 155–162 (1996).
68. D. Ghosh, K.G. Danielson, J.T. Alston, and S. Heyner, *In Vitro Cell Dev. Biol.* **27A**, 713–719 (1991).
69. H. Satoh et al., *Cell Struct. Funct.* **18**, 315–321 (1993).

70. M.D. Basson, G. Turowski, and N.J. Emenaker, *Exp. Cell Res.* **225**, 301–305 (1996).
71. C. Fujiyama, Z. Masaki, and H. Sugihara, *Pathol. Res. Pract.* **189**, 1205–1214 (1993).
72. J.A. East, S.P. Langdon, K.M. Townsend, and J.A. Hickman, *Differentiation* **50**, 179–188 (1992).
73. J. Green, S. Schotland, D.J. Stauber, C.R. Kleeman, and T.L. Clemens, *Am. J. Physiol.* **268**, 1090–1103 (1995).
74. S.T. Boyce and J.F. Hansbrough, *Surgery* **103**, 421–431 (1988).
75. A. Ben-Ze'ev, G.S. Robinson, N.L. Bucher, and S.R. Farmer, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2161–2165 (1988).
76. B. Rana, D. Mischoulon, Y. Xie, N.L. Bucher, and S.R. Farmer, *Mol. Cell Biol.* **14**, 5858–5869 (1994).
77. Z. Gatmaitan et al., *J. Cell Biol.* **97**, 1179–1190 (1983).
78. M. Haramaki, *Acta Pathol. Jpn.* **43**, 490–499 (1993).
79. N. Sawada et al., *In Vitro Cell Dev. Biol.* **23**, 267–273 (1987).
80. A.W. Thomson et al., *Adv. Exp. Med. Biol.* **378**, 511–518 (1995).
81. S. Toda et al., *Cell Struct. Funct.* **20**, 345–354 (1995).
82. M. Yamamoto, K. Yamamoto, and T. Noumura, *Exp. Cell Res.* **204**, 121–129 (1993).
83. M.A. Hadley, S.W. Byers, C.A. Suarez-Quian, H.K. Kleinman, and M. Dym, *J. Cell Biol.* **101**, 1511–1522 (1985).
84. R.H. Kramer and N. Marks, *J. Biol. Chem.* **264**, 4684–4688 (1989).
85. V.M. Belkin, A.M. Belkin, and V.E. Koteliansky, *J. Cell Biol.* **111**, 2159–2170 (1990).
86. W.D. Staatz et al., *J. Biol. Chem.* **266**, 7363–7367 (1991).
87. M.J. Elices and M.E. Hemler, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9906–9910 (1989).
88. Y. Takada and M.E. Hemler, *J. Cell Biol.* **109**, 397–407 (1989).
89. S. Sakashita, E. Engvall, and E. Ruoslahti, *FEBS Lett.* **116**, 243–246 (1980).
90. P. Speziale, M. Hook, T. Wadstrom, and R. Timpl, *FEBS Lett.* **146**, 55–58 (1982).
91. M.D. Rosso, R. Cappelletti, M. Viti, S. Vannucchi, and V. Chiarugi, *Biochem. J.* **199**, 699–704 (1981).
92. D.E. Hall et al., *J. Cell Biol.* **110**, 2175–2184 (1990).
93. E. Forsberg, M. Paulsson, R. Timpl, and S. Johansson, *J. Biol. Chem.* **265**, 6376–6381 (1990).
94. D. Kirchhofer, L.R. Languino, E. Ruoslahti, and M.D. Pierschbacher, *J. Biol. Chem.* **265**, 615–618 (1990).
95. K.R. Gehlsen, K. Dickerson, W.S. Argraves, E. Engvall, and E. Ruoslahti, *J. Biol. Chem.* **264**, 19034–19038 (1989).
96. M. Aumailley, R. Timpl, and A. Sonnenberg, *Exp. Cell Res.* **188**, 55–60 (1990).
97. R.H. Kramer, Y.-G. Cheng, and R. Clyman, *J. Cell Biol.* **111**, 1233–1243 (1990).
98. A. Sonnenberg et al., *J. Cell Biol.* **110**, 2145–2155 (1990).
99. M.M. Lotz, C.A. Korzelius, and A.M. Mercurio, *Cell Regul.* **1**, 249–257 (1990).
100. N.N. Tandon et al., *Biochem. J.* **274**, 535–542 (1991).
101. K. Satoh et al., *Cancer Lett.* **62**, 199–203 (1992).
102. M.C. Kibbey et al., *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10150–10153 (1993).
103. M.C. Kibbey, B. Johnson, R. Petryshyn, M. Jucker, and H.K. Kleinman, *J. Neurosci. Res.* **42**, 314–322 (1995).
104. S.S. Brown, H.L. Malinoff, and M.S. Wicha, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5927–5930 (1983).
105. L. Hinck, I.S. Nathke, J. Papkoff, and W.J. Nelson, *J. Cell Biol.* **125**, 1327–1340 (1994).
106. M.L. McGarvey, A.B.-V. Evercooren, H.K. Kleinman, and M. Bubois-Dalcq, *Dev. Biol.* **105**, 18–28 (1984).
107. R.G. Farrer and R.H. Quarles, *J. Neurosci. Res.* **45**, 248–257 (1996).
108. A.B.-V. Evercooren et al., *J. Neurosci. Res.* **8**, 179–193 (1982).
109. D. Edgar, R. Timpl, and H. Thoenen, *EMBO J.* **3**, 1463–1468 (1984).
110. N. Nagato, M. Aoyagi, and K. Hirakawa, *Glia* **8**, 71–76 (1993).
111. H. Ovadia et al., *Brain Res.* **322**, 93–100 (1984).
112. Y. Kubota, H.K. Kleinman, G.R. Martin, and T.J. Lawley, *J. Cell Biol.* **107**, 1589–1598 (1988).
113. D. Medina, M.L. Li, C.J. Oborn, and M.J. Bissell, *Exp. Cell Res.* **172**, 192–203 (1987).
114. C.H. Streuli et al., *J. Cell Biol.* **129**, 591–603 (1995).
115. M.A. Hadley, B.S. Weeks, H.K. Kleinman, and M. Dym, *Dev. Biol.* **140**, 318–327 (1990).
116. U. Hahn, A. Stallmach, E.G. Hahn, and E.O. Riecken, *Gastroenterology* **98**, 322–335 (1990).
117. A. Rizzino, V. Terranova, D. Rohrbach, C. Crowley, and H. Rizzino, *J. Supramol. Struct.* **13**, 243–253 (1980).
118. T.M. Sweeney, R.C. Ogle, and C.D. Little, *J. Cell Sci.* **97**, 23–31 (1990).
119. R.F. Aten, T.R. Kolodecik, and H.R. Behrman, *Endocrinology* **136**, 1753–1758 (1995).
120. K. Hirata, Y. Yoshida, K. Shiramatsu, A.E. Freeman, and H. Hayasaka, *Exp. Cell Biol.* **51**, 121–129 (1983).
121. D. Mooney et al., *J. Cell Physiol.* **151**, 497–505 (1992).
122. R. Holmes, G. Mercer, and N. Mohamed, *In Vitro* **15**, 522–531 (1979).
123. E.G. Hayman et al., *J. Cell Biol.* **95**, 20–23 (1982).
124. D.W. Barnes, J. Silnutzer, C. See, and M. Shaffer, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1362–1366 (1983).
125. D.W. Barnes, J.E. Reing, and B. Amos, *J. Biol. Chem.* **260**, 9117–9122 (1985).
126. M. Izumi, K.M. Yamada, and M. Hayashi, *Biochim. Biophys. Acta* **990**, 101–108 (1989).
127. S. Suzuki, A. Oldberg, E.G. Hayman, M.D. Pierschbacher, and E. Ruoslahti, *EMBO J.* **4**, 2519–2524 (1985).
128. M.G. Conlan, B.R. Tomasini, R.L. Schultz, and D.F. Mosher, *Blood* **72**, 185–190 (1988).
129. D. Jenne, and K.K. Stanley, *EMBO J.* **4**, 3153–3157 (1985).
130. D.W. Barnes, and J. Reing, *J. Cell Physiol.* **125**, 207–214 (1985).
131. C.J. Parker, O.L. Stone, V.F. White, and N.J. Bernshaw, *Br. J. Haematol.* **71**, 245–252 (1989).
132. G. Hetland, H.B. Pettersen, T.E. Mollnes, and E. Johnson, *Scand. J. Immunol.* **29**, 15–21 (1989).
133. E.G. Hayman, M.D. Pierschbacher, Y. Ohgren, and E. Ruoslahti, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4003–4007 (1983).
134. D. Seiffert, M.L. Iruela-Arispe, E.H. Sage, and D.J. Loskutoff, *Dev. Dyn.* **203**, 71–79 (1995).
135. X. Zheng, T.L. Saunders, S.A. Camper, L.C. Samuelson, and D. Ginsburg, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 12426–12430 (1995).
136. D. Jenne and K.K. Stanley, *Biochemistry* **26**, 6735–6742 (1987).

137. S. Suzuki et al., *J. Biol. Chem.* **259**, 15307–15314 (1984).
138. E. Ruoslahti and M.D. Pierschbacher, *Cell* **44**, 517–518 (1986).
139. M. Izumi, T. Shimo-Oka, N. Morishita, I. Ii, and M. Hayashi, *Cell Struct. Funct.* **13**, 217–225 (1988).
140. C. Gebb, E.G. Hayman, E. Engvall, and E. Ruoslahti, *J. Biol. Chem.* **261**, 16698–16703 (1986).
141. K.T. Preissner and G. Muller-Berghaus, *J. Biol. Chem.* **262**, 12247–12253 (1987).
142. D. Seiffert and R.R. Schleef, *Blood* **88**, 552–560 (1996).
143. J.I. Fuquay, D.T. Loo, and D.W. Barnes, *Infect. Immun.* **52**, 714–717 (1986).
144. G.S. Chhatwal, K.T. Preissner, G. Muller-Berghaus, and H. Blobel, *Infect. Immun.* **55**, 1878–1883 (1987).
145. M. Paulsson and T. Wadstrom, *FEMS Microbiol. Immunol.* **2**, 55–62 (1990).
146. M. Kostrzynska and T. Wadstrom, *Int. J. Med. Microbiol. Virol. Parasitol. Infect. Dis.* **277**, 80–83 (1992).
147. R.J. Arko et al., *J. Clin. Microbiol.* **29**, 70–75 (1991).
148. F. Re et al., *J. Cell Biol.* **127**, 537–546 (1994).
149. P.J. Bianchine, P.R. Burd, and D.D. Metcalfe, *J. Immunol.* **149**, 3665–3671 (1992).
150. K.M. Neugebauer, K.A. Venstrom, and L.F. Reichardt, *J. Cell Biol.* **116**, 809–815 (1992).
151. R.M. Leven and F. Tablin, *Exp. Hematol.* **20**, 1316–1322 (1992).
152. R.A. Wolfe, G.H. Sato, and D.B. McClure, *J. Cell Biol.* **87**, 434–441 (1980).
153. A. Merzak, S. Koochekpour, and G.J. Pilkington, *Cell Adhes. Commun.* **3**, 27–43 (1995).
154. P.W. Grabham, P.H. Gallimore, and R.J. Grand, *Exp. Cell Res.* **202**, 337–344 (1992).
155. D.F. Mosher, *Annu. Rev. Med.* **41**, 85–97 (1990).
156. M.W. Long and V.M. Dixit, *Blood* **75**, 2311–2318 (1990).
157. P.J. Mansfield and S.J. Suchard, *J. Immunol.* **153**, 4219–4229 (1994).
158. D.D. Roberts and V. Ginsburg, *Arch. Biochem. Biophys.* **267**, 405–415 (1988).
159. A.S. Asch, S. Silbiger, E. Heimer, and R.L. Nachman, *Biochem. Biophys. Res. Commun.* **182**, 1208–1217 (1992).
160. R.O. Hynes, *Cell* **69**, 11–25 (1992).
161. M.A. Bourdon and E. Ruoslahti, *J. Cell Biol.* **108**, 1149–1155 (1989).
162. A.L. Prieto, G.M. Edelman, and K.L. Crossin, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10154–10158 (1993).
163. A. Lochter, J. Taylor, B. Fuss, and M. Schachner, *Eur. J. Neurosci.* **6**, 597–606 (1994).
164. E. Ruoslahti, *Annu. Rev. Cell Biol.* **4**, 229–255 (1988).
165. J.E. Koda and M. Bernfield, *J. Biol. Chem.* **259**, 11763–11770 (1984).
166. G.R. Martin and R. Timpl, *Annu. Rev. Cell Biol.* **3**, 57–85 (1987).
167. L. Sherman, J. Sleeman, P. Herrlich, and H. Ponta, *Curr. Opin. Cell Biol.* **6**, 726–733 (1994).
168. I. Stamenkovic, M. Amiot, J.M. Pesando, and B. Seed, *Cell* **56**, 1057–1062 (1989).
169. L.A. Goldstein et al., *Cell* **56**, 1063–1072 (1989).
170. A. Aruffo, I. Stamenkovic, M. Melnick, C.B. Underhill, and B. Seed, *Cell* **61**, 1303–1313 (1990).
171. S.C. Wheatley, C.M. Isacke, and P.H. Crossley, *Development* **119**, 295–306 (1993).
172. V.B. Lokeshwar, N. Iida, and L.Y. Bourguignon, *J. Biol. Chem.* **271**, 23853–23864 (1996).
173. B.F. Haynes, M.J. Telen, L.P. Hale, and S.M. Denning, *Immunol. Today* **10**, 423–428 (1989).
174. A. Bartolazzi, R. Peach, A. Aruffo, and I. Stamenkovic, *J. Exp. Med.* **180**, 53–66 (1994).
175. M.E. Hemler, *Annu. Rev. Immunol.* **8**, 365–400 (1990).
176. P.-Y. Chan and A. Aruffo, *J. Biol. Chem.* **268**, 24655–24664 (1993).
177. A.E. Koch et al., *Lab. Invest.* **64**, 313–320 (1991).
178. M.J. Elices et al., *Cell* **60**, 577–584 (1990).
179. C. Ruegg et al., *J. Cell Biol.* **117**, 179–189 (1992).
180. D. Simmons, M.W. Makgoba, and B. Seed, *Nature* **331**, 624–627 (1988).
181. J. Fawcett et al., *Nature* **360**, 481–484 (1992).
182. D.E. Staunton, M.L. Dustin, and T.A. Springer, *Nature* **339**, 61–64 (1989).
183. S.D. Marlin and T.A. Springer, *Cell* **51**, 813–819 (1987).
184. M.S. Diamond, D.E. Staunton, S.D. Marlin, and T.A. Springer, *Cell* **65**, 961–971 (1991).
185. Y. Rosenstein et al., *Nature* **354**, 233–235 (1991).
186. G.M. Edelman, *Annu. Rev. Cell Biol.* **2**, 81–116 (1986).
187. G.J. Cole and R. Akeson, *Neuron* **2**, 1157–1165 (1989).
188. T.B. Kuhn, E.T. Stoeckli, M.A. Condrau, F.G. Rathjen, and P. Sonderegger, *J. Cell Biol.* **115**, 1113–1126 (1991).
189. A.M. Montgomery et al., *J. Cell Biol.* **132**, 475–485 (1996).
190. M. Grumet, *J. Neurosci. Res.* **31**, 1–13 (1992).
191. C. Rader et al., *Eur. J. Biochem.* **215**, 133–141 (1993).
192. S.J. Harper, S.R. Bolsover, F.S. Walsh, and P. Doherty, *Cell Adhes. Commun.* **2**, 441–453 (1994).
193. D.R. Friedlander et al., *J. Cell Biol.* **125**, 669–680 (1994).
194. H. Stockinger et al., *J. Immunol.* **145**, 3889–3897 (1990).
195. P.J. Newman et al., *Science* **247**, 1219–1222 (1990).
196. Y. Xie and W.A. Muller, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5569–5573 (1993).
197. H.M. DeLisser et al., *J. Biol. Chem.* **268**, 16037–16046 (1993).
198. L. Piali et al., *J. Cell Biol.* **130**, 451–460 (1995).
199. L.A. Lasky, *Annu. Rev. Biochem.* **64**, 113–139 (1995).
200. G.S. Kansas, K. Ley, J.M. Munro, and T.F. Tedder, *J. Exp. Med.* **177**, 833–838 (1993).
201. O. Abbassi, T.K. Kishimoto, L.V. McIntire, D.C. Anderson, and C.W. Smith, *J. Clin. Invest.* **92**, 2719–2730 (1993).
202. A. Varki, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7390–7397 (1994).
203. S.I. Simon et al., *J. Immunol.* **155**, 1502–1514 (1995).
204. S.B. Munro and O.W. Blaschuk, in P. Brodt ed., *Cell Adhesion and Invasion in Cancer Metastasis*, Chapman & Hall, New York, 1996, pp 17–34.
205. S. Armeanu, H.J. Buhring, M. Reuss-Borst, C.A. Muller, and G. Klein, *J. Cell Biol.* **131**, 243–249 (1995).
206. P. Doherty, L.H. Rowett, S.E. Moore, D.A. Mann, and F.S. Walsh, *Neuron* **6**, 247–258 (1991).
207. H.R. Payne and V. Lemmon, *Dev. Biol.* **159**, 595–607 (1993).
208. W. Kolanus et al., *Cell* **86**, 233–242 (1996).
209. G.E. Hannigan et al., *Nature* **379**, 91–96 (1996).
210. S.J. Shattil et al., *J. Cell Biol.* **131**, 807–816 (1995).

BACILLUS

COLIN R. HARWOOD
KEITH STEPHENSON
University of Newcastle upon Tyne
Newcastle upon Tyne, UK

KEY WORDS

Batch culture
Continuous culture
Gene regulation
Genetic manipulation
Genetic transformation
Industrial enzymes
Insecticidal toxins
Large-scale fermentation
Protein secretion
Taxonomy of *Bacillus*

OUTLINE

INTRODUCTION

Members of the genus *Bacillus* are gram-positive, endospore-forming bacteria that include both mesophiles and extremophiles. They are widely distributed in nature but are most commonly found in soil and associated water sources. Their extraordinary metabolic diversity has been exploited by old and new biotechnologies—from the production of Natto, a traditional Japanese food obtained by fermenting soybeans with *Bacillus subtilis* var. *natto*, to engineered industrial enzymes used in the food and detergents industries.

DISTRIBUTION AND TAXONOMY

The genus *Bacillus* comprises gram-positive, endospore-forming, aerobic or facultatively anaerobic bacteria. Members of the genus show extraordinary metabolic diversity and lifestyles and include thermophiles, psychrophiles, alkalophiles, and acidophiles (1,2). They are metabolically chemoorganotrophs, being dependent on organic compounds as sources of carbon and energy. *B. subtilis*, the most extensively characterized species, is a prototroph capable of growing at mesophilic temperatures on chemically defined salts media with glucose or other simple carbon sources. In contrast, some insect pathogenic species are nutritionally fastidious and require highly specialized growth media (3).

Bacillus species are among the most widely distributed microorganisms, found commonly in soil and associated environments, including plants and water sources such as rivers, estuaries, and coastal waters (4). Most species are harmless to humans and animals, and only a few pathogens have been identified. The latter include *B. anthracis*, the causative agent of anthrax, *B. cereus*, which causes food poisoning, and several insect pathogens.

The primary reservoir for most *Bacillus* species is the soil, where they secrete a variety of carbohydrate- and protein-degrading enzymes that allow them to grow at the expense of plant material and other nutrients (4). These macromolecular hydrolases are the source of much of the commercial interest in these species (5). Their ability to form spores means that they can often be detected in an environment long after periods of metabolic activity. The major soil types are members of the *B. subtilis* and *B. sphaericus* groups. However, representatives of the more fastidious species in the *B. polymyxa* group tend to accumulate in soils with a high organic content.

Some species form a close relationship with plant roots, particularly representatives of the *B. polymyxa* group. Several species (e.g., *B. azotofixans*, *B. macerans*, and *B. polymyxa*) fix nitrogen under anaerobic conditions. Certain strains of *B. subtilis* are categorized as plant growth–promoting rhizobacteria (PGPR). In return for stimulating plant growth by increasing the availability of inorganic nutrients, or secreting plant growth hormones or compounds

(often peptide antibiotics) that inhibit the growth of plant pathogens, PGPR receive nutrients from the plant in the form of plant exudates (6).

The metabolic diversity of the genus is matched by significant genetic diversity, and the %GC content of their genomic DNA varies from about 33 to 67 mol%. This variation is much larger than would be expected for a well-defined genus, and the 60 to 70 currently recognized species of *Bacillus* are likely to be reassigned to an increased number of more clearly defined genera. This process has already started with the reassignment of *B. acidocaldarius* and some other acidophilic thermophiles to the genus *Alicyclobacillus* (7). Numerical taxonomic and 16*S* ribosomal (r)RNA sequence analyses show good but not perfect congruence (8,9). These studies have resulted in *Bacillus* species being assigned to five groups. The evolutionary distance between these groups is large enough to affect the outcome of interspecies gene transfer and cloning experiments, particularly when there are large discrepancies between the %GC contents of the species involved.

The *B. subtilis* group includes many of the better-known bacilli. These include facultative anaerobes such as *B. licheniformis*, which grow fermentatively in the absence of exogenous electron acceptors, and aerobes such as *B. subtilis* that grow weakly in the absence of oxygen except in the presence of nitrate, which they can use as an alternative electron acceptor (10,11). Members of this group produce acid from several sugars, and several (e.g., *B. lentus*, *B. firmus*) are halotolerant. The group includes toxigenic species such as *B. anthracis*, *B. cereus*, and *B. thuringiensis*.

The *B. sphaericus* group is strictly oxidative, using acetate or amino acids rather than sugars as carbon and energy sources (12). Members of the *B. sphaericus* group are virtually unique among bacilli in having the *meso*-diaminopimelic acid that is usually present in cell walls replaced by lysine or ornithine.

The *B. polymyxa* group (renamed *Paenibacillus*) is made up of facultative anaerobes that use sugars and polysaccharides. *B. macerans* grows better in the absence of oxygen with an acidogenic type of fermentation that changes to an ethanolic fermentation at acidic pHs. *B. polymyxa* exhibits a butanediol type of fermentation. Members of the *B. brevis* group (renamed *Brevi bacillus*) are taxonomically heterologous. They are strict aerobes that generally do not produce acid from sugars. Some species, notably *B. azotoformans*, can grow anaerobically by using nitrate as an electron acceptor.

The thermophilic bacilli, which grow optimally at temperatures above 50 °C, represent a heterologous group with at least three lines of descent that range from anaerobes to strict aerobes. Some are able to grow fermentatively on lactic acid (*B. coagulans*) whereas others are chemolithotrophic autotrophs. The group includes *B. acidocaldarius* and the other acidophilic thermophiles that have been reassigned to the genus *Alicyclobacillus* (7).

Preservation of Strains

Although most strains of *Bacillus* survive well on agar plates, either at room temperature or at 4 °C, it is recommended that they are subcultured on a weekly basis. Because *Bacillus* species are able to sporulate, viable cells may even be recovered from severely dehydrated agar plates, particularly from minimal agar plates that encourage sporulation. Long-term stocks of *Bacillus* may be preserved as glycerol or lyophilized cultures, in the form of spores or vegetative cells (13). Strains that sporulate well may be preserved as spore suspensions in sterile water (14). Spore suspensions are generally stable for many years. Most *Bacillus* strains can be transported on freshly inoculated nutrient agar slopes or as spore suspensions spotted on sterilized filter paper disks encased in sterilized aluminium foil.

Culture Collections

The *Bacillus* Genetic Stock Center (BGSC, Department of Biochemistry, Ohio State University, 484 West 12th Avenue, Columbus, OH 43210-1292) has an extensive collection of mutant strains, bacteriophages, and plasmids. Although mainly devoted to *B. subtilis*, the collection also includes strains of *B. cereus*, *B. licheniformis*, *B. megaterium*, *B. pumilus*, *B. stearothermophilus*, *B. thuringiensis*, and a few others. BGSC produces an extensive catalog that can be requested by e-mail from *dzeigler@magnus.acs.ohio-state.edu*. Strains of *Bacillus* are also available from a variety of international culture collections and a comprehensive list has been published (15). The American Type Culture Collection (ATCC, 12301 Parklawn Drive, Rockville, MD 20852) has a World Wide Web (WWW) site at *http://www.atcc.org/*, and the Japan Collection of Micro-organisms (JCM; Riken, Wako-shi, Saitama 351, Japan) is available at *http://www.wdcm.riken.go.jp/wdcm/JCM*. More general links to culture collections are available at the World Data Collection for Micro-organisms at *http://www.wdcm.riken.go.jp/*.

The publication of the genome of *B. subtilis* 168 (16) has provided the opportunity to investigate the function of open reading frames (ORFs) with no known function (unknown reading frames (URFs) (17). A collection of isogenic mutants has been established in which the URFs of *B. subtilis* 168 have been systematically inactivated with a purpose-designed integrational vector, pMUTin. These mutants also permit the expression of the target genes to be monitored and any gene(s) downstream in the same transcriptional unit to be induced. The depository for the collection is accessible to account holders via the Micado WWW site at *http://locus.jouy.inra.fr/cgi-bin/genmic/madbase/progs/madbase.operl*.

GENETIC MANIPULATION

A wide range of genetic techniques has been developed for *B. subtilis*, making it among the most amenable bacteria to genetic manipulation. These techniques include classical in vitro and in vivo mutagenic methods and extremely efficient in vitro genome manipulation methods that are in advance of those available for other bacteria, including *Escherichia coli* (18). *Bacillus* vectors have been designed to allow recombinant or mutant genes to be maintained autonomously in the host or integrated into the chromo-

some by double-crossover (Fig. 1) or single-crossover (Fig. 2) recombinations. In the case of a single-crossover recombination, tandem copies of the gene, one wild type and one mutated, can be engineered into the chromosome, if required, each under the control of independently controlled promoters (Fig. 2). The techniques for the manipulation of other *Bacillus* species are limited by their transformability.

Mutagenesis

A number of DNA repair systems have been detected and studied in *B. subtilis*, including error-prone SOS-like translesion DNA synthesis but not photoreactivation (19). Consequently, classical in vivo mutagenic methods generate mutants efficiently. Limitations in whole-cell mutagenic techniques, including lack of specificity and the generation of multiple mutations that may influence the cell's phenotype in an unpredictable manner, can be overcome by use of an alternative approach in which cloned copies of the target DNA are subjected to in vitro mutagenesis and subsequently reintegrated into the chromosome (20). This approach is aided by the publication of the genome sequence of *B. subtilis* strain 168 (16).

Specific changes are currently introduced into DNA sequences by PCR techniques, with more extensive changes being engineered by splicing PCR methodologies (e.g., splicing by overlap extension [21]). The modified DNA can then be introduced into the host bacterium either on replicating or integration plasmids. The ability to generate insertion mutants is one of the great technical strengths of *B. subtilis*. The techniques rely on the ease with which *Bacillus* DNA can be cloned into plasmids that replicate in *E. coli* but not in *B. subtilis* and the high frequency with which homologous recombination occurs within *B. subtilis*.

The identification and analysis of *B. subtilis* genes involved in specific functions or responding to particular stimuli has been facilitated by the use of transposons and transpositional mutagenesis. Although the value of transposons as a tool for the analysis of *B. subtilis* has diminished with the publication of the genome sequence (16), nevertheless the technology remains important for the analysis of other *Bacillus* species. Because of the lack of suitable native *Bacillus* transposons, Tn*917* from *Enterococcus faecalis* and Tn*10* from *E. coli* have been used (22–24). In addition to their value for mutagenesis, transposons have been used to clone DNA adjacent to the site of integration, to generate transcriptional fusions to reporter genes, to control the expression of adjacent genes, and as phenotypic tags for mapping and cloning studies.

Transposon Tn*917* is a 5.3-kb Tn*3*-like transposon conferring inducible erythromycin resistance (MLS-type). This transposon has a number of properties that make it a valuable genetic tool, including *(1)* the ability to insert relatively randomly into bacterial chromosomes, *(2)* a relatively high transposition frequency, *(3)* the ability to accommodate at least 8 kb of DNA without markedly affecting transposition frequency or insertional stability, and *(4)* a host range that includes gram-positive and gram-negative bacteria. In addition to *B. subtilis*, Tn*917* has been used in *B. amyloliquefaciens*, *B. licheniformis*, and *B. megaterium* (23). Insertion of Tn*917* is not entirely random, and a limited preference for specific loci has been observed in the case of *B. subtilis* and *B. licheniformis* (25,26).

An alternative to Tn*917* is an engineered miniderivative of Tn*10* (24). This transposon consists of a *Staphylococcus aureus*–derived *cat* (chloramphenicol resistance) gene flanked by 307-bp fragments derived from IS*10*. The gene encoding the transposase is incorporated into a delivery vector rather than the minitransposon itself. Tn*10*-derived transposons have two potential advantages over Tn*917*-based transposons. They appear to transpose at a significantly higher frequency and insert randomly without bias to specific targets. A series of special-purpose delivery systems have been developed for gene inactivation, the recovery of adjacent chromosomal sequences, and transcription fusions.

Transformation

Bacillus subtilis was the first nonpathogenic bacterium to be successfully transformed with extracted DNA (27–29). The availability of methods for introducing DNA into cells is essential for recombinant DNA technology, and several transformation techniques have been developed for *B. subtilis*. The most frequently used method is transformation of naturally competent cells, although protoplasts of *B. subtilis* (30) and several other *Bacillus* species can be efficiently transformed. Current methods for electrotransformation of *Bacillus* species usually result in low yields of transformants.

Natural competence is one of several postexponential phase phenomena that are characteristic of *B. subtilis* strain 168. In glucose-minimal media, maximal competence develops shortly after the transition from exponential to stationary phase, and only about 10% of the cells are able to take up DNA. Natural competence has been observed with a limited number of *B. subtilis* strains, and few natural competence systems are known for *Bacillus* species. Unlike *Haemophilus influenzae*, competent *B. subtilis* cells do not discriminate between native and foreign DNA. In recombination-proficient strains, homologous chromosomal DNA is integrated in the host chromosome at the site of homology. If saturating amounts of DNA (>1 μg/mL of culture) are used, the cotransfer of unlinked genetic markers occurs. This phenomenon, called congression, can be used to introduce genes for which no selectable phenotype is known (31).

The frequency of plasmid-mediated transformation is usually between 0.001% and 0.1%, and this frequency is reduced still further with ligation mixtures. A major reason for these low efficiencies is that donor plasmid DNA becomes randomly fragmented and converted into single-stranded DNA before entry into the cell. Recirculation of the plasmid DNA requires a region of homology provided on fragments of opposite polarity, and consequently only plasmid multimers, present in most plasmid preparations, or monomers containing internal repeats are effective in transformation (32).

Bacillus protoplasts can incorporate DNA if treated with chaotropic agents such as polyethylene glycol. The

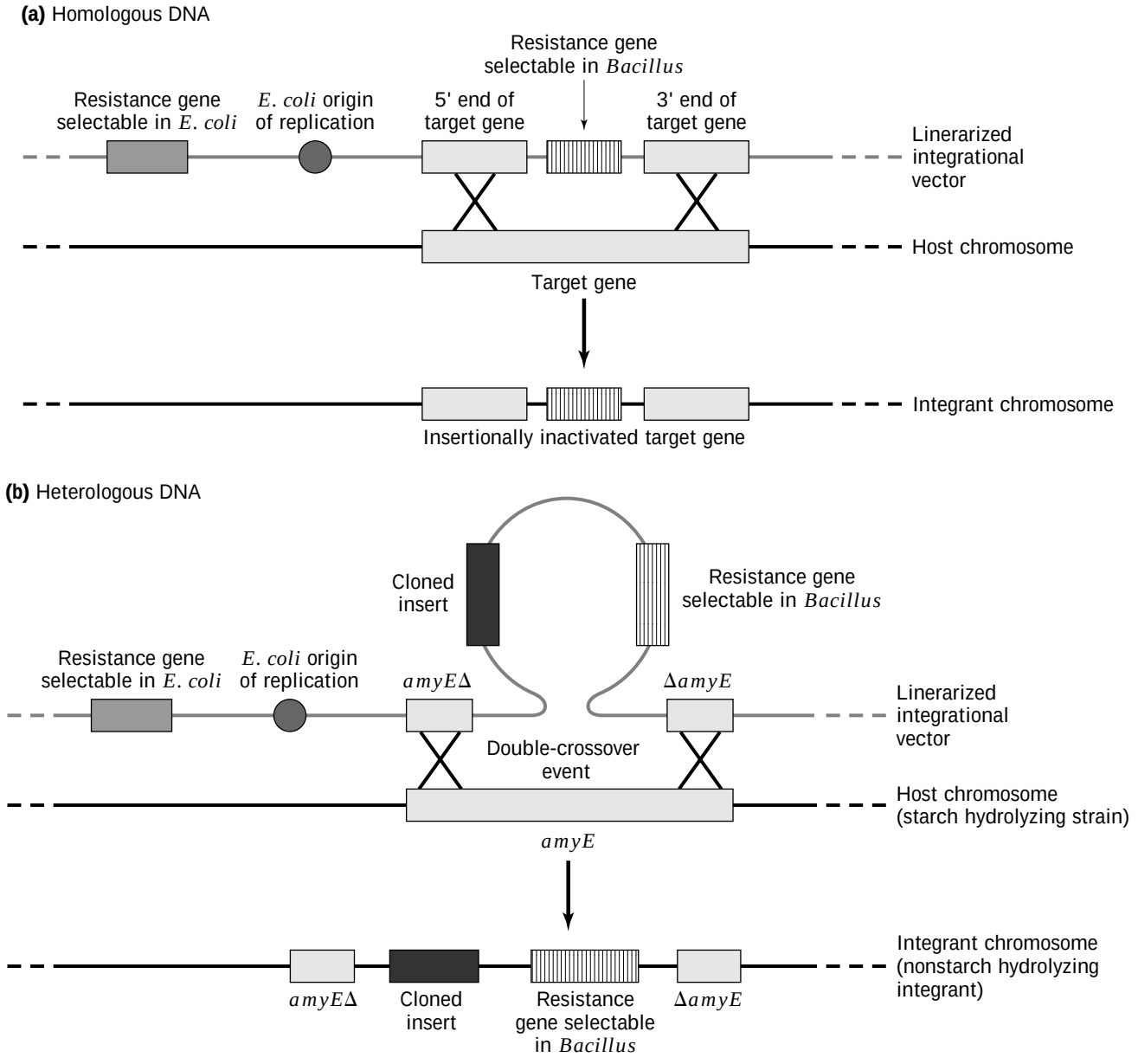


Figure 1. Double-crossover recombination shows the integration of a fragment of DNA from a linearized vector that is unable to replicate in *Bacillus* species. (**a**) Recombination event involving sequences homologous to the target DNA. (**b**) Recombination event involving heterologous DNA, in this case inserted between 5′ (*amyEΔ*) and 3′ (*ΔamyE*) fragments of the nonessential *amyE* gene encoding α-amylase. Integrants lose the ability to hydrolyze starch.

fragility of the wall-less protoplasts means that transformation needs to be carried out in an isosmotic medium; the cell walls are subsequently regenerated on complex regeneration media. Protoplast transformation was originally developed for *B. subtilis* by Chang and Cohen (30) but has subsequently been modified in many laboratories. A major advantage of protoplast transformation is that plasmid DNA remains double stranded throughout the procedure and consequently is less likely to undergo rearrangements. Despite its advantages, protoplast transformation methods are laborious and difficult to reproduce; transformation frequencies up to 10% per protoplast are claimed but in practice are usually a great deal lower. In addition, the complexity of the regeneration media precludes the selection for prototrophic markers. Protoplast transformation protocols have been developed for several other *Bacillus* species including *B. amyloliquefaciens*, *B. licheniformis*, *B. stearothermophilus*, *B. anthracis*, and *B. firmus*.

Genome Analysis and Manipulation

The genome of *B. subtilis* strain 168 is 4.2 Mbp in length, has a G+C content of 43.5%, and encodes slightly more

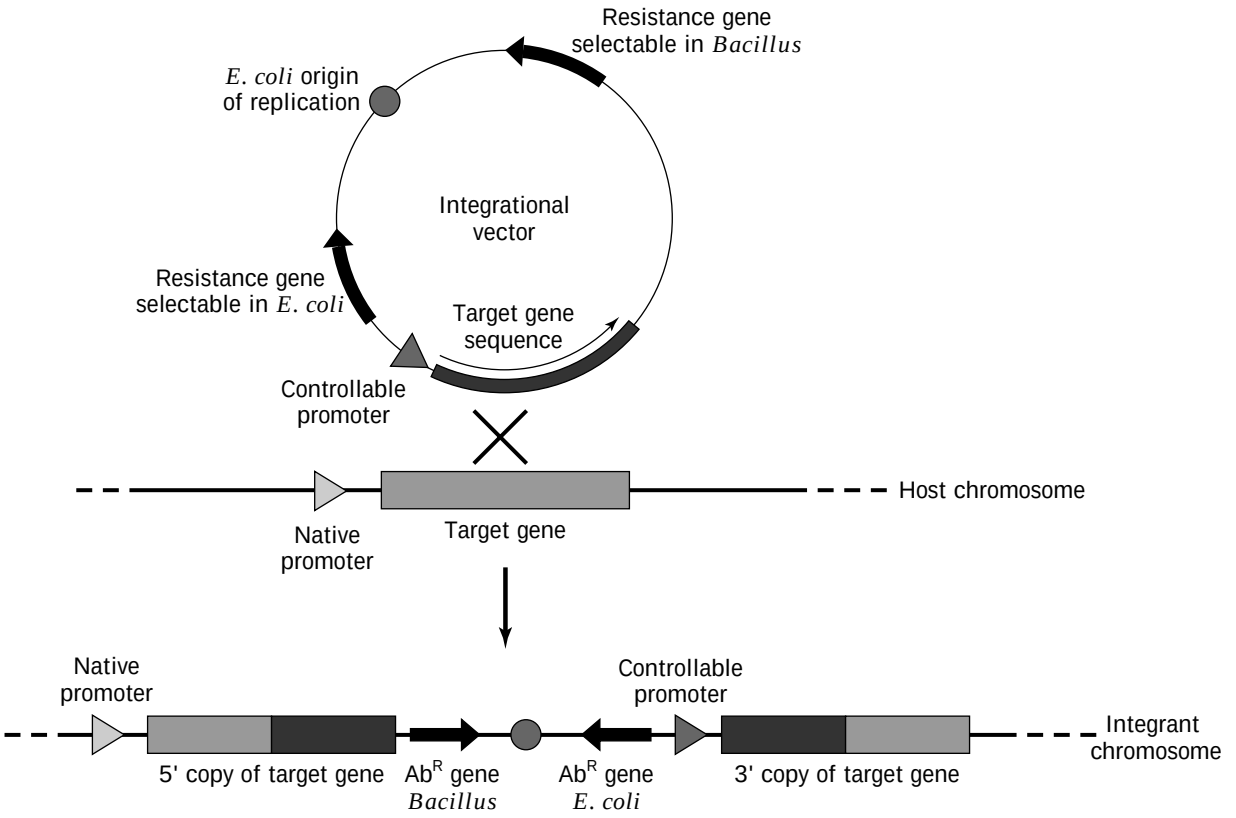


Figure 2. Single-crossover recombination showing integration of an intact integration vector into the host chromosome at the site of the target gene. Part or all of the target gene is duplicated as a result of the integration event. The inclusion of a controllable promoter upstream of the target gene fragment on the vector allows for controllable gene expression of the 3′ copy of the gene or any genes downstream in the same operon. *Ab*^{R}, selectable antibiotic resistance gene.

than 4,100 protein-coding genes (16). The genome shows evidence of 10 prophages or remnants of prophages, which may account for its susceptibility to lysis at the end of exponential growth in the absence of an energized membrane (33). Nearly half the protein-coding genes have paralogues in the genome, and several expanded gene families were detected, particularly for two-component signal transducers (34 genes) and ATP-binding transporters (77 genes).

Although analysis of a complete microbial genome sequence reveals many new ORFs, it is not possible to ascribe functions to all the detected ORFs on the basis of currently available sequence databases and bioinformatics tools. Even in well-characterized species such as *E. coli* and *B. subtilis*, URFs represent about 40% of the detected ORFs. Given the intensive in vitro studies on bacteria, it is likely that the function of a significant proportion of the URFs will relate to growth and survival in natural habitats.

The inability to ascribe function to more than a third of the ORFs in *B. subtilis* has required the development of new strategies for analyzing gene function. In the case of *B. subtilis*, a highly coordinated systematic functional analysis program involving groups with a wide range of technical expertise has been established. Approximately 2,000 unknown genes have been systematically disrupted with an integrational vector to generate a collection of isogenic mutants strains (34). The mutant construction strategy includes the introduction of a *lacZ* transcriptional reporter for monitoring the expression of the target gene and a controllable promoter that facilitates the expression of any genes downstream in the same operon (Fig. 2).

The availability of the *B. subtilis* genome sequence has revolutionized the methods of analysis and manipulation of this bacterium, and consequently traditional methods of genetic linkage mapping involving generalized transducing for long-range mapping and transformation for fine structure mapping have largely been rendered redundant (31). This is not the case for species such as *B. megaterium* for which generalized transducing phages such as MP13 and protoplast transformation remain important analytical tools (35).

Increasingly, integration systems are used for the analysis and manipulation of the *B. subtilis* genome. Transformation of cells with homologous DNA fragments is very efficient, and if heterologous sequences are flanked by sequences homologous to chromosomal DNA they can be inserted efficiently at the sites of homology via a double-crossover recombination event (Fig. 1). In the case of circular molecules, a single site of homology is all that is required and integration is via a single-crossover recombination event (Fig. 2).

Single or Campbell-type integration events generally use *E. coli*–based plasmid vectors that are not able to replicate in gram-positive bacteria but that include an antibiotic resistance gene that can be selected in *B. subtilis*. A fragment of *B. subtilis* DNA (≥0.4 kb) is cloned via *E. coli* into the integrational vector, which is then transformed into *B. subtilis*. Selection for an antibiotic-resistant gene results in transformants in which the vector has integrated into the host chromosome via a single-crossover recombination between homologous regions on the vector and bacterial chromosomes. Integration results in a duplication of the cloned fragment at the flanking ends of the integrated vector. These tandem repeats represent an amplifiable unit; increasing the amounts of antibiotic used for selection can result in the amplification of the intervening sequences up to 70 copies (36).

The integrational vector system can be used to generate different mutational outcomes. If the cloned fragment carries an entire operon, including promoter and transcription terminator, two fully functional copies of the gene are present in the integrant. If 5′- or 3′-end fragments of an operon are used, only one functional copy of the target gene is generated. If the 5′-end fragment includes a ribosome-binding site and start codon, and is placed downstream of a controllable promoter, then the expression of the single functional copy of the target gene is under the control of this promoter (Fig. 2). Finally, if regions internal to the target gene are used, no functional copies are formed after integration.

With double-crossover integrations, chromosomal DNA is replaced by either mutationally altered homologous DNA (Fig. 1a) or heterologous DNA (Fig. 1b). Only a single copy of the target gene is introduced. The incoming fragment, which must encode a selectable phenotype, is flanked by DNA that is homologous to sequences at the chromosomal target site. The *amyE* gene, encoding a nonessential α-amylase, is often used as a target of double-crossover integrations because its disruption provides a positive selection phenotype on starch plates (Fig. 1b).

Cloning and Expression Systems

The development of recombinant DNA techniques for *B. subtilis* is now well advanced, and previously encountered problems of vector stability have largely been overcome.

Plasmid Vectors. *B. subtilis* 168, the strain that has been sequenced and that is used for most genetic studies, does not contain endogenous plasmids, whereas most plasmids present in other *Bacillus* strains are cryptic. Consequently, plasmid vectors designed for use in *B. subtilis* have been taken from other gram-positive bacteria, notably *Enterococcus faecalis*, *Lactococcus lactis*, and *Staphylococcus aureus* (18,37). Staphylococcal plasmids such as pUB110 and pE194, which formed the basis of the first generation of *Bacillus* vectors, are still in common use for *B. subtilis* (38). The kanamycin resistance plasmid pUB110 is the most widely used replicon but suffers from problems of structural and segregational instability that result from its single-stranded (rolling circle) mode of replication (39). The second generation of *Bacillus* vectors involved the construction of chimeras between gram-positive and gram-negative replicons (40). These shuttle or bifunctional vectors enabled initial cloning experiments to be carried out in *E. coli*, exploiting this organism's highly efficient transformation system. In recent years, stable single-stranded replicating plasmids and theta-replicating plasmids have been identified and are being developed as a new generation of *Bacillus* vectors (18). The former includes derivatives of plasmids such as pWVO1 derived from *Lactococcus lactis* (41) and certain endogenous *Bacillus* plasmids, the latter plasmids based on pAMβ1, originally isolated from *Enterococcus faecalis* (42,43).

An alternative to using autonomously replicating plasmid vectors is to use plasmid integration vectors. Such integrants are usually stable; reversal of the process, by recombination in the duplicated sequences, occurs at a frequency of about 10^{-4} to 10^{-5} per cell generation (44). Integration vectors have been reviewed by Albertini and Galizzi (45) and Pérego (46).

Various special-purpose plasmid vectors have also been developed, including expression and secretion vectors. Although the expression of many genes has been studied in *B. subtilis*, relatively few promoter systems have been developed for controlled, high-level expression. The majority of extracellular enzymes synthesized for industrial purposes are expressed from native promoters induced at the end of exponential growth and for extended periods in the stationary phase (47). During this time they can be subjected to a variety of control pathways, including transition-state regulators and catabolite repression (48,49). Industry has developed promoter–expression systems (including growth regimes) that can direct the synthesis of extracellular proteins to concentrations of about 20 g/L, although for commercial reasons the details of these systems are not generally available. Three promoter systems that are widely used in research laboratories are discussed next.

P_{spac} Promoter. The P_{spac} promoter was constructed by fusing the 5′ sequences of a promoter from SPO1, a lytic *B. subtilis* bacteriophage, and the 3′ sequences of the *lac* promoter, including the entire *lac* operator. The point of fusion is within the −10 region of the hybrid promoter (50). The controllability of P_{spac} is dependent on the presence of the *lac* repressor, expressed in *B. subtilis* by placing it downstream of a promoter and ribosome-binding site (RBS) derived from a *B. licheniformis* penicillinase gene. The P_{spac} promoter can be induced 50 fold with the lactose analog IPTG.

XylR-Controlled Promoters. The genes for the utilization of xylose polymers are controlled in *B. subtilis* at the level of transcription (51) through the activity of a repressor, the product of the *xylR* gene. The xylose-inducible promoter–operator elements have been used without modification to control expression from chromosomal and plasmid locations. A copy of the *xylR* gene is usually included on high-copy-number expression vectors to maintain a balance between the number of repressor molecules and operator sites. Although genes in the xylose regulon are subject to catabolite repression, the catabolite-responsive element is located well downstream of the promoter and operator, and is therefore not included in the expression vectors. In con-

trast to IPTG, xylose is a cheap and readily available substrate and can be used for the induction of large-scale fermentations.

SacR-Controlled Promoters. The inducible expression of *sacB*, the gene encoding extracellular levansucrase by sucrose, involves a number of regulatory mechanisms, not all of which are fully understood. The *sacB* gene is controlled positively by sucrose, the *sacY* antiterminator, and the products of *degQ* and *sacU* and negatively by *sacX*, a putative PTS enzyme $II^{sucrose}$. The *sacR* regulatory region located upstream of the *sacB* promoter contains a target sequence for DegQ and SacU, which together enhance transcription. Downstream of this promoter, but before the *sacB* RBS, a termination-like stem-loop structure is present, which is acted on by the SacY antiterminator to mediate sucrose-induced expression (52). Expression vectors based on the *sacB* promoter have the advantage that this promoter can be activated during exponential growth when the expression of most *Bacillus* extracellular proteases is repressed and is not subject to catabolite repression. The level of induction is also geared to the level of sucrose in the medium; the promoter is only weakly induced in the presence of 1 mM sucrose and is fully induced at 30 mM.

Phage Vectors. The development of phage vectors for *B. subtilis* has lagged well behind that of *E. coli*, and no phage vectors with the range and versatility of those of lambda exist for this organism. Libraries have been generated in derivatives of *Bacillus* phages φ105 and SPβ, although the former has been used more extensively (53). Attempts to develop a cosmid system for *B. subtilis* have failed because of the lack of an efficient in vitro packaging system for any of the candidate phages.

Bacteriophages φ105 and PBSX (a defective phage of *B. subtilis* strain 168) have both been developed for the production of proteins in *B. subtilis* (53). In both cases the phages have been modified to make them temperature inducible by the introduction of mutations in the genes encoding the immunity repressor proteins. Target genes are introduced downstream of a strong prophage promoter. In the case of φ105MU333, temperature induction can lead to the production of the target gene product to 0.5 mg/mL culture supernatant (54). Bacteriophage-based expression systems have the advantage of being relatively stable because they are maintained as a single copy during growth. Copy number increases at the time of induction due to the replication of the phage. The promoter is tightly controlled but is induced cheaply and in a manner favored in industrial fermentations by increasing temperature. However, vector development is still required to maximize the full potential of this system for high-level protein production.

PROTEIN SECRETION

Bacillus subtilis and its close relatives secrete proteins directly into the growth medium efficiently and to high concentrations. The productivity of *Bacillus* species for the production of extracellular proteins has been improved by a variety of techniques including random and directed mutagenesis, recombinant techniques, and growth regimes. Yields of tens of gram per liter are expected from industrial strains secreting native *Bacillus* proteins (55–58).

The secretion process is highly specific, involving functional sequences on the secretory protein and recognition between the secretory proteins and the host's protein secretion machinery (secretory preprotein translocase). For convenience, the process in *B. subtilis* can be divided into three stages: intracellular processes, involving the targeting of secretory proteins to the cytoplasmic membrane–secretory translocase; translocation across the cytoplasmic membrane; and posttranslocational stages, involving the processing of the secreted protein, its folding into the native conformation, and passage through the cell wall into the growth medium.

A major limiting step in the production of native secretory proteins is the level of transcription of their genes. The transcription of genes encoding native secretory proteins is tightly controlled and involves a number of complex, often interacting, regulatory networks, including transcriptional activators (e.g., DegQ, DegR, DegU, and SenS) and repressors (e.g., AbrB, Hpr, and Sin). For example, the nucleotide sequence upstream of *apr*, encoding the major alkaline protease of *B. subtilis*, includes binding sites for Hpr, Sin, SacU/Q, and AbrB (48).

Most *Bacillus* secretory proteins are synthesized as precursors with amino-terminal extensions known as signal peptides, or pre-peptides. The signal peptide is required to target the precursor to the cytoplasmic membrane and for initiating the interaction with the secretory translocase. During or shortly after translocation across the membrane, the signal peptide is removed through the action of signal peptidases. *Bacillus* signal peptides vary in length from 18 to 35 amino acid residues; their average length of about 30 residues is 5 to 7 residues longer than their counterparts in gram-negative bacteria and eukaryotes. The C-terminal end of the signal peptide contains the signal peptidase recognition sequence. Lipoproteins, which remain attached at the *trans* side of the membrane after translocation, are cleaved by a prolipoprotein-specific signal peptidase that has a different recognition sequence (47).

Many exported *Bacillus* proteins are synthesised as pre-pro-proteins. The pro-peptide, which is located between the signal peptide (pre-) and the mature protein, can vary in length from about 8 (e.g., *B. subtilis* α-amylase) to more than 200 amino acids (various neutral proteases). Pro-peptides have been implicated in the folding of the mature protein into its native conformation following translocation (47,59). Cleavage of the pro-peptide occurs after translocation, autocatalytically in the case of proteases or via extracellular proteases in the case of other secretory proteins (60). Long (8–20 residues) hydrophobic domains, flanked by positively charged residues, are usually absent from secretory proteins because such domains act as "stop transfer" or "membrane anchor" sequences. Similarly, few positively charged residues are found in the amino terminus of the mature protein as these reduce export efficiency.

Surprisingly little is known about early events in the *B. subtilis* secretory pathway. Secretion-specific molecular chaperones, required to maintain secretory precursor proteins in an unfolded conformation, to prevent their aggregation, and to target them to the secretory machinery in

the membrane, have not been detected in *B. subtilis*. This could indicate that secretion is cotranslational, and indeed elements of the signal recognition particle–like (SRP-like) pathway have been found in this bacterium (47). The secretory translocase complex appears to be similar to that of *E. coli*. The SecAEY complex is primarily responsible for the ATP-dependent transport of proteins across the membrane, and homologs of other components of the *E. coli* translocase, such as SecD, F, and G, also appear to be present, in the former two cases as a SecDF fusion. A unique feature of *B. subtilis* strain 168 is the presence of five type I signal peptidases and a single type II lipoprotein signal peptidase; some industrial strains contain plasmids specifying additional type I signal peptidases.

The release of the translocated mature protein from the membrane is usually accompanied by its folding to the native conformation. For many *B. subtilis* secretory proteins, this requires folding factors such as Fe^{3+} and Ca^{2+} and/or the extracytoplasmic lipoprotein PrsA. Efficient folding of the mature protein on the *trans* side of the membrane appears to be important for avoiding or reducing electrostatic interactions with negatively charged components of the cell wall and degradation by extracellular proteases (61,62).

Although there are limits to the amounts of native proteins secreted by bacilli, such proteins are usually capable of being produced at commercially significant levels. In contrast, the extracellular production of heterologous proteins, especially from eukaryotic sources, is frequently inefficient. The reasons for this are not fully understood but are likely to include that the inherent incompatibilities with the secretory apparatus and their susceptibility to the many extracellular proteases, cell associated and released, produced by these bacteria (61,63).

GENE EXPRESSION

Most of the extracellular hydrolase enzymes produced by *Bacillus* species are induced during the transition from vegetative growth to stationary phase, where they appear to function to increase the range of substrates that may be used to sustain growth. The expression of the genes encoding some of these enzymes appears to be linked to, but not necessary for, sporulation itself (e.g., alkaline and neutral proteases), while others (e.g., α-amylases and alkaline phosphatases) are linked to the exhaustion of specific substrates.

Sporulation

Sporulation is a developmental process in which bacilli develop into heat-, chemical-, solvent-, radiation-, and desiccation-resistant spores (64,65). Their transition from dormancy to vegetative cells involves germination and outgrowth. Under laboratory conditions at 37 °C, sporulation takes approximately 8 h. The process has been divided into a series of stages that are defined primarily on the basis of their morphology. These stages, labeled 0, II, III, . . . through VII (stage I is no longer recognized), broadly correlate with the time (in hours) after the induction of sporulation (Fig. 3).

Sporulation is induced in response to nutrient deprivation and is only initiated after the failure of transition phase processes (e.g., motility and chemotaxis, production of macromolecular hydrolases, development of competence) to restore vegetative growth. Sporulation is induced by activated SpoOA (SpoOA–P) via a phosphorelay that facilitates multifactorial signal processing. The first morphologically distinct stage is stage II, which is defined by the formation of an asymmetric cell division event that divides the cell into two compartments, the larger sporangium (or mother cell) and the prespore, or forespore. During stage II, the prespore becomes completely engulfed by the sporangium so that, at stage III, the two compartments are separated by a double membrane system. During stage IV, peptidoglycan is laid down between the two membrane layers to form the spore cortex and the germ cell wall. During this stage the prespore becomes refractile to light. In stages V and VI, spore coat proteins are laid down on the surface of the inner and outer membrane of the prespore. As these mature the spore becomes increasingly resistant to chemicals and heat. In the final stage, stage VII, the by-now-dehydrated spore is released from the sporangium by cell lysis. The fully mature spore is metabolically inactive and capable of surviving as such for certainly tens, and possibly hundreds, of years.

Sporulation is a tightly controlled process requiring differential expression of genes on the chromosomes in each of the two compartments, as well as developmental checkpoints to ensure synchrony between their patterns of development. This is achieved through the differential activation of sporulation-specific sigma (σ) factors at or about stages II to III. σ^F is active in the prespore, and σ^E is active in the sporangium. These sigma factors are, in turn, responsible for the induction of sporangium-specific (σ^K) and prespore-specific (σ^G) sigma factors. Synchrony between the two compartments is maintained by signal communication processes.

Much less is known about germination and outgrowth. In the laboratory at 37 °C, the time between induction of germination and the first cell division event is about 90 min. Germination is triggered by a variety of chemical and physical stimuli; spores of *B. subtilis* germinate in response to L-alanine or to a mixture of asparagine, glucose, fructose, and KCl (66). Germination is accompanied by a series of rapid chemical and physiological changes including the degradation of the spore cortex, disruption of the spore coat, rehydration, and swelling. These changes are accompanied by a loss of heat and chemical resistance. It takes some 30 min before de novo RNA synthesis takes place. During outgrowth the cell begins to increase in mass and to replicate its DNA, forming an elongated cylindrical cell that is only subsequently reduced to a normal-sized cell by a delayed cell division.

Phosphate

Phosphorus is a major growth-limiting substrate in natural environments such as soil (67); consequently, under phosphate limitation, *B. subtilis* induces a programmed series of responses to reduce the cell's requirement for phosphorus, increase its affinity for phosphate uptake, and pro-

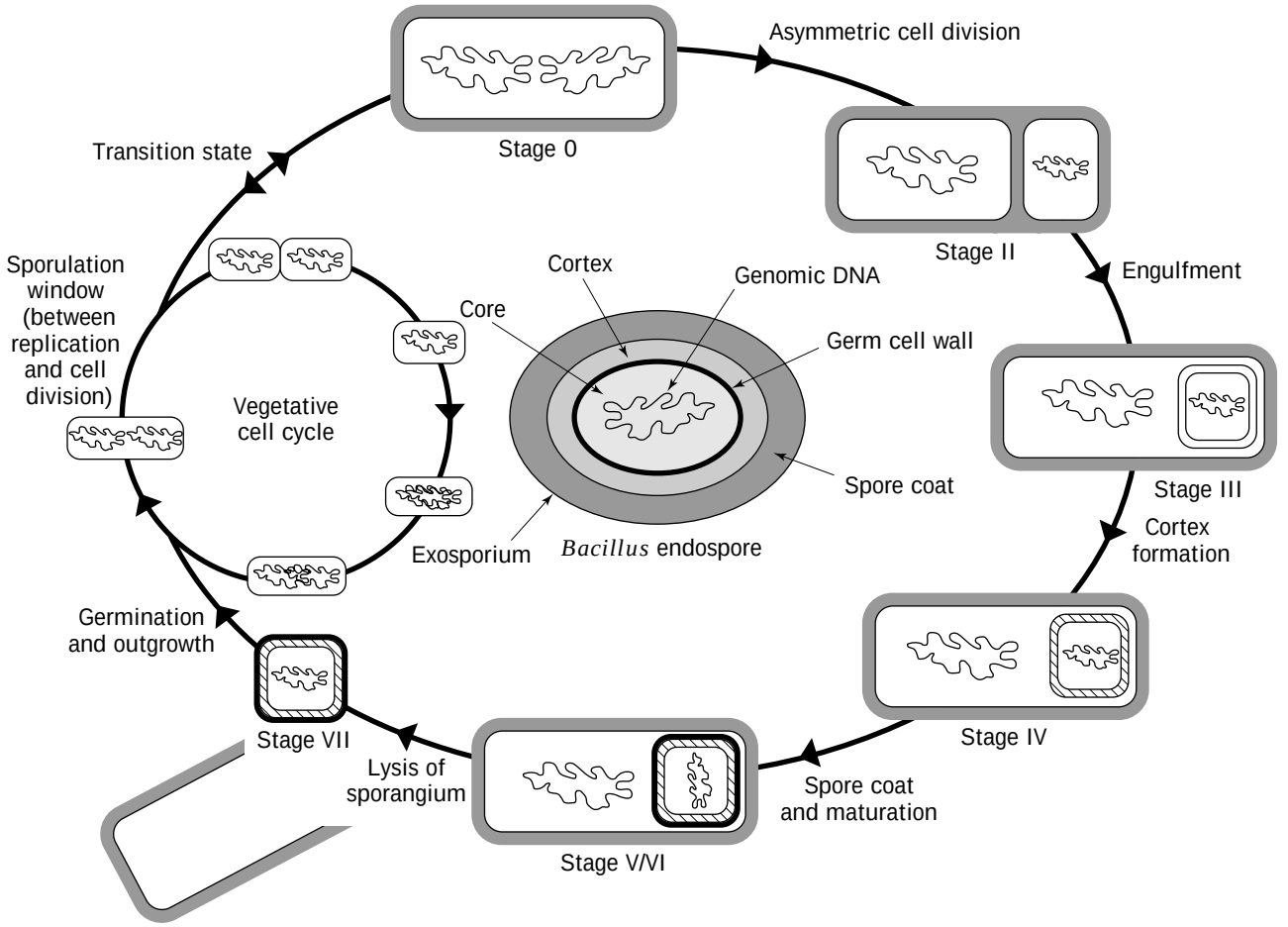


Figure 3. The sporulation pathway in *Bacillus subtilis* and its relationship to the vegetative cell cycle. Sporulation is only initiated toward the end of the vegetative cell cycle when the cell has two complete copies of the chromosome.

duce extracellular hydrolytic enzymes to recover inorganic phosphate from organic sources (68,69).

B. subtilis responds to phosphate stress by the coordinated induction of approximately 40 genes (69). Prominent among these are genes of the Pho regulon (67,70–72), whose products include alkaline phosphatases (APase), alkaline phosphodiesterase (APDase), a high-affinity phosphate transporter, and enzymes required for teichuronic acid synthesis.

The Pho regulon of *B. subtilis* is controlled by two *trans*-acting regulators, PhoP and PhoR, that together form the components of a two-component, environment-sensing, signal transduction system (70,73). PhoR is a membrane-spanning sensor protein with histidine protein kinase (HPKase) activity that is activated at low phosphate concentrations to PhoR–P. This in turn acts as a substrate for the activation of PhoP, a response regulator required for the induction of the Pho regulon. *B. subtilis* genes activated by PhoP–P, which include the bicistronic operon encoding PhoP and PhoR, constitute the Pho regulon. The response regulators of at least two other signal transduction systems, SpoOA–P and ResD, indirectly affect the expression of the Pho regulon by repressing the *phoP / R* operon (71,74,75).

The promoters of the Pho regulon genes/operons that have been studied to date have a conserved sequence, TTAACA (74), located around position −22. This sequence may be equivalent to the Pho box observed in *E. coli* (76), which also tends to be located toward the 5′ end of the promoter region of Pho regulon genes/operons.

Carbon

Bacillus species can grow on a wide range of simple carbon sources, including glucose, sucrose, and amino acids as well as a range of complex substrates such as proteins and polysaccharides, particularly plant-derived substrates such as starch and arabinans. In the case of more complex substrates, extracellular enzymes are required to reduce the size of these substrates before uptake and utilization. Carbon substrates are transported into the cell either in an unmodified form or via a PTS system (77).

In *B. subtilis* many genes/operons encoding enzymes involved in the catabolism of substrates as diverse as starch and histidine are repressed by readily metabolizable carbon sources such as fructose, glucose, glycerol, or mannitol—so-called PTS sugars. Catabolite repression functions at the level of transcription, but differs substantially from

the cyclic AMP-mediated system observed in *E. coli*. Three components of the *B. subtilis* system have been identified: a *cis*-acting catabolite-responsive element (CRE) and two *trans*-acting factors, CcpA and Hpr. The 14-bp palindomic CRE sequence (78) is essential for catabolite repression. CREs are found at diverse locations in relation to the transcriptional units of catabolite-sensitive genes, in some cases overlapping the promoter (*amyE*), in others being located within the structural gene itself (*xylA*). The latter positioning makes it extremely difficult to engineer catabolite insensitivity in such genes.

CREs are likely to be binding sites for a catabolite repressor protein. Although it has not been directly shown to be the case, CcpA, which is related to *E. coli* LacI-type repressors, is the most likely candidate for this protein. Mutants in the *ccpA* gene abolish catabolite repression, including glucose repression of sporulation, without affecting substrate-specific induction of the individual genes/operons concerned. They also severely affect growth on substrates such as glucose, fructose, glucitol, and glycerol that are metabolized via glycolysis or the TCA cycle, but not on substrates such as xylose which are metabolized via the pentose phosphate pathway. In the former case, growth is restored if TCA cycle intermediates are added to the growth medium. Interestingly, a CRE appears to be responsible for glucose-mediated activation of the acetate kinase gene (*ackA*) and in genes encoding enzymes of central metabolism.

HPr is a component of the phosphotransferase system (PTS) which, when activated by phosphorylation to HPr(Ser–P), binds to CcpA. The key signaling component of the catabolite repression pathway appears to be the Embden–Meyerhof pathway (EMP) intermediate fructose-1,6-diphosphate (FDP). High concentrations of this compound activate the ATP-dependent HPr serine kinase and possibly stimulates an interaction between CcpA and HPr(Ser–P). This complex may then bind to the CREs of catabolite-responsive genes/operons (49).

Nitrogen

Bacillus spp. can use ammonium and a number of other nitrogen-containing compounds (notably certain amino acids) as sole sources of nitrogen (79). Ammonium is assimilated via glutamine synthetase (GS), glutamine:2-oxoglutarate amidotransferase (GOGAT), and glutamate dehydrogenase (GDH) (80). Most members of the genus are able to use all three enzymes, while a smaller group, including many (but possibly not all) strains of *B. subtilis*, use GS and GOGAT, while nitrogen-fixing *Bacillus* species use GDH. In *B. subtilis*, the relative levels of GS and GOGAT are dependent on the available sources of nitrogen. The levels of GS synthesis are dependent on the need for glutamine synthesis; they are lowest in the presence of glutamine, a nitrogen source favored by this bacterium, and highest in the presence of nonfavored nitrogen sources such as ammonium. No global regulatory network analogous to that of Ntr of *E. coli* (81) has been detected in *B. subtilis*, although GS synthesis is autoregulated via the GlnR repressor and has also been implicated as a key signal for the nitrogen status of the cell modulating the activity of a variety of other metabolic pathways.

GOGAT levels are low in the presence of glutamate and amino acids (e.g., glutamine, arginine, and histidine) that can be catabolized to glutamate, and high in the presence of ammonium and amino acids that are not broken down to glutamate and other sources of nitrogen such as nitrate.

In addition to its use as a nitrogen source, nitrate (and nitrite) can be used by *B. subtilis* in place of oxygen as an electron acceptor, allowing growth under anaerobic conditions. Nitrate respiration requires the products of the *narGHJI* operon (10,11).

PRODUCTS

Bacillus species are an important source of industrial enzymes, fine biochemicals, antibiotics, and insecticides (5). Additionally, the ability of *B. subtilis* and close relatives to secrete grams per liter quantities of protein directly into the growth medium, their ease of growth, and their well-proven safety have also made them prime candidates for the production of heterologous proteins.

Many commercially important enzymes have traditionally been produced in their native hosts. However, when organisms with novel enzymic activities are newly isolated, the costs and uncertain outcome associated with process optimization are often prohibitive. This even applies to well-established strains of *B. amyloliquefaciens* and *B. licheniformis* that have been used extensively in commercial enzyme fermentations. Both these organisms are generally refactory to genetic engineering, and relatively little information is available concerning their biochemistry and physiology. Instead, companies are increasingly seeking to develop well-characterized strains such as *B. subtilis* strain 168 as universal hosts for commercial applications.

Enzymes

The world annual sales for industrial enzymes was recently valued at $1 billion, with strong growth in the paper, textiles, and waste treatment markets. Three-quarters of the market is for enzymes involved in the hydrolysis of natural polymers, including proteolytic enzymes used in the detergents, dairy, and leather industries and carbohydrases used in the baking, brewing, distilling, starch, and textile industries. About two-thirds of these enzymes are produced by fermentation from *Bacillus* species, the main producers, together with their products and activities, are listed in Table 1.

Two *Bacillus* enzymes dominate the industrial enzymes markets alkaline (serine) protease and α-amylase. Although many species of *Bacillus* secrete enzymes of these types, their catalytic properties vary from one producer strain to the next. Moreover, commercially significant enzymes may be modified to improve their performance or stability in particular industrial processes.

Alkaline proteases are the single largest enzyme market and are used extensively as additives to detergents. They are relatively nonspecific endoproteases that convert substrate proteins into small, readily soluble fragments. The two main proprietary products, subtilisin Novo (BPN′) from *B. amyloliquefaciens* and subtilisin Carlsberg from *B. licheniformis*, have been the subject of extensive develop-

Table 1. Commercially Significant Products Produced by *Bacillus* species

Product	Producer organism	Properties/use
Glucose isomerase	*B. coagulans*	Conversion of glucose to fructose in manufacture of low calorie sweeteners
α-Amylase	*B. subtilis*	Limited hydrolysis of starch
	B. amyloliquefaciens	Liquefication of starch
	B. stearothermophilus	Thermophilic liquefaction of starch
	B. licheniformis	Thermostable liquefaction of starch
β-Amylase	*B. polymyxa*	Formation of low-dextrose-equivalent syrups from starch
β-Glucanase	*B. amyloliquefaciens*	Hydrolysis of barley β-glucans in brewing and animal feeds
	B. circulans	
	B. subtilis	
Pullulanase	*Bacillus* species	1,6-Bond starch debranching enzyme
Alkaline protease	*B. alkalophilus*	Detergents and leather industries
	B. licheniformis	
Neutral protease	*B. thermoproteolyticus*	Temperature-stable enzyme
Thermolysin	*B. thermoproteolyticus*	Condensation of L-aspartic acid and D,L-phenylalanine to produce aspartame
β-1,4-Galactosidase	*B. stearothermophilus*	Dairy industry
	B. subtilis	
Chymosin (rennin/pepsin)	*B. mesenterius*	Rennet substitutes
	B. cereus	
	B. polymyxa	
	B. megaterium	
	B. licheniformis	
Glucose dehydrogenase	*B. megaterium*	Glucose assay
Glycerol kinase	*B. stearothermophilus*	Glycerol/trigycerides assay
Phospholipase C	*B. cereus*	Phosphatidylcholine assay
Restriction endonucleases	*B. aneurinolyticus*	*Ban*I
	B. amyloliquefaciens	*Bam*HI
	B. subtilis	*Bsu*
	B. caldolyticus	*Bcl*
	B. globiii	*Bgl*
	B. stearothermophilus	*Bst*
	B. sphaericus	*Bsp*
	B. brevis	*Bbv*
δ-Endotoxin	*B. thuringiensis*	Biocontrol agent
Peptide antibiotics	*B. licheniformis*	Bacitracin
	B brevis	Edeines, gramicidin, tyrocidine
	B. subtilis	Fengycin, subtilin, surfactin
	B. polymyxa	Polymyxins

ment programs. Their three-dimensional structures have been solved (82,83), and functionality testing has resulted in the identification of the catalytic triad, an oxyanion-binding site, and substrate-binding determinants. Subtilisins have been engineered to improve their characteristics in washing detergents by increasing their stability in the presence of oxidizing chemicals, at temperatures above 60 °C and at pH values up to 12. They have pH optima of about 11.0 and half-lives of approximately 50 min at 40 °C. They are used at 0.4–1% in formulations that includes anionic and nonanionic surfactants, oxidants, soap, and optical brighteners. Although these were originally used simply in the form of a powder, hypersensitivity reactions among producers and users meant that the enzymes had to be added to formulations in a microencapsulated form. This process not only reduces dust formation but also protects the enzymes from other components of the formulation during storage.

α-Amylases are used extensively in the starch industry where, because of the physical properties of starch, they need to be used at high temperatures and good thermostability is an important property of these enzymes. Industry has sought to obtain thermostable amylases by screening for new sources and by improving the thermostability of existing enzymes by protein engineering. For example, the α-amylases from *B. amyloliquefaciens*, *B. stearothermophilus*, and *B. licheniformis* are closely related (3) but have very different thermostabilities of 2, 50, and 270 min, respectively, at 90 °C and pH 6.5 (84,85). Extra salt bridges caused by specific lysine residues in the more-thermostable *B. licheniformis* amylase are responsible for its increased stability, and the equivalent residues in the less-stable enzymes have been the target for protein engineering. Additionally, the extensive homology between the genes encoding these amylases has also been used for in vivo recombination experiments in which hybrid enzymes combining beneficial characteristics of these enzymes have been screened. The three-dimensional structure of the *B. licheniformis* α-amylase has recently been published (86), and the three amino acid residues that constitute the calcium-binding site (i.e., asparagine-104, aspartate-200, and histidine-235), have been identified.

The α-amylase from *B. acidocaldarius* strain A2 shows greater thermostability under acid conditions (87). Additionally, *B. subtilis* strain KP1064 (88) produces an enzyme that is active against starch and pullulan.

The industrial production of α-amylase was developed in 1917 using a strain subsequently reclassified as *B. amyloliquefaciens*. This α-amylase could be used to a maximum temperature of about 85 °C. Because starch needs to be cooked to 100 °C to rupture the starch granules, the introduction of a thermostable α-amylase from *B. licheniformis* (89) that can operate at about 105 °C was a distinct improvement.

B. coagulans is an important source of glucose isomerases, required for the conversion of glucose (0.75 times as sweet as sucrose) to fructose (twice as sweet as sucrose) in the production of high-fructose corn syrup. In contrast to the proteases and amylases already discussed, glucose isomerase is an intracellular enzyme used either in the form of immobilized cells or extracted and immobilized on a solid matrix.

Metabolites

Bacillus species are used for the production of a number of primary metabolites for the food and health care industries. Extensive knowledge of their biochemistry and regulation has allowed growth and rational mutant isolation strategies to be developed to the extent that this organism has been considered for the production of a wide range of metabolites at commercially viable levels.

B. subtilis has been used for the production of the nucleotides xanthanylic acid (XMP), inosinic acid (IMP), and guanylic acid (GMP), which are of commercial importance as flavor enhancers. The primary fermentation products are the nucleosides hypoxanthine, inosine, and guanosine because their greater membrane permeability allows them to accumulate to higher concentrations in the culture medium. These are then converted to their respective nucleotides by chemical phosphorylation (90).

Attempts have been made to develop strains of *Bacillus* for the production of amino acids such as tryptophan, histidine, and phenylalanine and of vitamins such as biotin, folic acid, and riboflavin. The use of analogs has led to the isolation of resistance mutants overproducing a number of amino acids, nucleosides, and vitamins (91). Knowledge of *B. subtilis* is such that rational approaches are now possible and attempts have been made to engineer this bacterium for the production of folic acid. Promising strains have been obtained, although their production levels cannot yet compete with those of *Brevibacterium*, *Corynebacterium*, or *Serratia* species.

Peptide Antibiotics

Under conditions of nutritional stress *Bacillus* species produce a variety of special metabolites that enhance their survival in natural environments (92). Prominent among these are peptide antibiotics, generally short peptides (2–20 amino acid residues) that are synthesized by large multienzyme complexes (peptide synthetases) rather than ribosomes. Individual amino acid residues are often extensively modified and may be linked to each other by peptide bonds or through esters or lactones. Gramicidin-S is a cyclic decapeptide from *B. brevis* with antibacterial and surfactant properties. Bacitracin is a branched cyclic dodecapeptide produced by *B. licheniformis* and *B. subtilis* (93) that is used as a topical antibiotic. Its antibacterial activity is directed against bacterial cell wall synthesis where it inhibits the recycling of lipid-P carriers. Surfactin, produced by most strains of *B. subtilis*, has both antibacterial activity and powerful surfactant properties. Polymyxins are membrane-active branched cyclic acylpeptides produced by *B. polymyxa*.

A minority of the peptide antibiotics produced in *Bacillus* species are synthesized on ribosomes and subsequently modified extensively by posttranslational processing (92). The resulting products are usually larger, but only marginally so, than those produced by the peptide synthetases. The best studied are the lantibiotics, so-called because they contain the unusual amino acids lanthionine and methyl-lanthionine. They include subtilin, a 32-residue lantibiotic produced by *B. subtilis* that shows antibacterial and antitumor activity.

Heterologous Proteins

The efficient large-scale production of heterologous enzymes and proteins by *B. subtilis* has been the focus of considerable research effort. The ability of this bacterium to secrete proteins directly into the culture medium in quantities of grams per liter offers considerable process advantages, in terms of product yield, structural integrity, and downstream processing costs, over production systems in which proteins are accumulated in the cytoplasm or periplasm, often as insoluble aggregates.

Nevertheless, attempts to use *B. subtilis* for the production of heterologous proteins has met with only limited success, and there are no reports of the use of this host in an industrial fermentation. Although extracellular proteins from close relatives can be produced in this bacterium at high concentrations, the yield of proteins from unrelated species remains disappointingly low. This is likely to be due, at least in part, to the production of at least seven extracellular proteases, which have been shown to cause extensive degradation of heterologous proteins, and to incompatibilities with the *Bacillus* protein secretion pathway. The isolation of strains defective in the identified proteases has helped in some but by no means all cases (61,63).

Other species of *Bacillus* may prove useful for development as hosts for the production of heterologous proteins, particularly those naturally producing low levels of extracellular proteases. Of interest are species such as *B. brevis* that produce a crystalline surface (S) layer at the outer surface of a thin cell wall. S-layer proteins have been observed at concentrations up to 35 g/L in the culture medium and, consequently, attempts have been made to develop the expression and signal sequences of their genes into components of a new generation of secretion vectors (94). However, it may ultimately be necessary to develop strains that are currently used in industrial processes as production hosts by improving our understanding of their genetics and physiology.

B. subtilis has also proved useful for the intracellular production of outer membrane proteins of gram-negative pathogens (95). These proteins, which have potential for use for immunodiagnostics and as vaccines, are produced in *B. subtilis* in preference to their native hosts to avoid contamination with endotoxins with which they often form tight associations.

Insecticides

The use of chemical insecticides, with a world market worth $5 billion, is increasingly seen as problematical because of the development of resistance in target insect populations, lack of specificity against the insect pest, persistence in the environment, and toxicity to man and other animals. Alternative biocontrol methods have been developed involving the use of viruses, fungi, and bacteria. *B. thuringiensis* (96) is the most successful insecticidal biocontrol agent and accounts for about 90% of the worldwide market for bioinsecticides. Although *B. thuringiensis* was originally isolated in 1901 from an infection of a Japanese silkworm farm and has been used as an insecticidal agent since 1920, it still only represents about 1–2% of the total insecticide market.

Strains of *B. thuringiensis* have been identified that are pathogenic for each of the main groups of insect pests, namely Lepidotera (moths, butterflies), Diptera (flies and mosquitoes), and Coleoptera (beetles). More recently strains of *B. thuringiensis* have been isolated that are active against noninsect pests, notably nematodes, mites, and protozoa (97). The toxicity of *B. thuringiensis* results from the production of proteinaceous δ-endotoxins during sporulation. The toxins, which may represent as much as 30% of the cell's weight, form crystals within the mother cell (98). The toxins are synthesized from plasmid-encoded *cry* genes, and a single strain may produce several toxins encoded by more than one plasmid. They are synthesized initially in the form of an inactive pro-toxin (70–140 kDa) that is considerably larger than the active form (60 kDa). The δ-endotoxins, which have the properties of chemical agents rather than infectious agents, are approximately 300 times more potent than pyrethroids and 8,000 times more potent than organophosphates. *B. thuringiensis* toxins therefore combine high toxicity and specificity for their target pests with little or no toxicity for nontarget insects (e.g., pollinators, predators) and other animals.

A large number of *cry* genes have been cloned, and the regions responsible for the pro- and mature forms, host specificity, and toxicity have been mapped. As different toxins exhibit a range of host specificities and levels of toxicity, chimeric toxins that combine the specificities and toxic regions of different natural toxins are currently being developed.

B. thuringiensis strains have been used extensively for the control of crop and forest insect pests and of insect disease vectors. However, as an alternative means of application, genes encoding the δ-endotoxin have been used directly to transform a variety of plant crops, including tomatoes, tobacco, potatoes, and cotton, in which field trials have confirmed their activity against target pests (99).

SAFETY

Bacillus species have been used extensively in the fermentation industry for the production of a range of products including enzymes, fine biochemicals, antibiotics, and insecticides. In addition, a strain of *B. subtilis* is used for the fermentation of soybeans to Natto, an extensively consumed ($\sim 10^8$ kg/annum) traditional Japanese food product. In a review of the safety of *B. subtilis* and *B. amyloliquefaciens* (these strains were considered to be synonymous up to 1973 [100]), no reports were found of infection or toxicity relating to *B. amyloliquefaciens*. In the case of *B. subtilis*, the situation was complicated by the fact that before about 1970 diagnostic laboratories did not distinguish between *B. cereus* and *B. subtilis*. Virtually all reports of putative *B. subtilis* infections since that time either were associated with drug abuse or occurred in patients whose immune system had been compromised by treatment with immunosuppressants or chemotherapeutic agents. This low incidence of pathogenicity and the widespread use of their products in the food, beverage, and detergents industries have resulted in GRAS (generally regarded as safe) status being granted to these species by the U.S. Food and Drug Administration (FDA).

GROWTH

Laboratory Culture Conditions

The majority of *Bacillus* species will grow at mesophilic temperatures on commercially prepared nutrient media, although in some cases it is necessary to modify the pH or salt concentration. Obligate thermophilic species such as *B. stearothermophilus* that do not grow satisfactorily at 37 °C are usually grown at 60 °C. Moderately thermophilic species, such as *B. coagulans*, are grown between 45 °C and 50 °C. The more fastidious insect pathogens, *B. larvae and B. popilliae*, require the addition of thiamine for growth and are usually grown at 25 °C to 30 °C. *B. stearothermophilus* requires additional calcium and iron, and *B. pasteuri* requires 0.5–1% urea.

B. subtilis and many other species are able to grow in simple salts media containing ammonium or amino acids as nitrogen sources and glucose or other simple sugars as carbon sources. The most commonly used general-purpose chemically defined medium is Spizizen's minimal medium (27). The growth rate in this medium can be improved by the replacement of ammonium, a poor nitrogen source for *B. subtilis*, with glutamine.

B. subtilis is able to use a number of amino acids as nitrogen sources (e.g., arginine, glutamine, glutamate, asparagine, and aspartate), and the respective catabolic pathways are induced in the presence of these compounds. Consequently, the concentration of many amino acids recommended to overcome auxotrophy are actually growth limiting. It should also be borne in mind that derivatives of the well-studied *B. subtilis* strain 168 often show tryptophan auxotrophy, and consequently complex media in which acid-hydrolyzed casein provides the main source of nitrogen need to be supplemented with this amino acid.

Although many *Bacillus* species sporulate readily, special media and growth protocols are required for efficient sporulation (14). Sporulation is induced in response to nutrient deprivation and occurs after the end of exponential growth. In *B. subtilis*, sporulation is repressed in the presence of glucose and at low cell densities. A widely used sporulation medium was described by Schaeffer and colleagues (101). In some species sporulation can also be induced by decoyinine, an analog of GTP (14).

Growth in Batch Culture. Although many species of *Bacillus* grow under anaerobic conditions, they are normally grown aerobically. Some strains (e.g., *B. subtilis*) have a tendency to lyse in the absence of an energized membrane (33). Growth is normally carried out in conical flasks in an orbital (reciprocating) incubator at >200 rpm, although most species can be grown in batch fermentors. Ideally, the culture volume should not exceed 20% of the flask volume. The media should be prewarmed as temperature shock can also induce autolysis.

Nutrients are usually supplied in excess and growth continues until a particular substrate is exhausted, the oxygen tension or pH falls to an inhibitory level, or an inhibitor accumulates to an inhibitory level. The various phases in the growth cycle have been described in detail elsewhere (102).

Growth in Continuous Culture. *Bacillus* species undergo marked metabolic, physiological, and morphological changes during batch culture. Many of these changes occur during transition to and during stationary phase and include morphologically distinct spores and physiologically distinct competent cells. To some extent these population variations can be reduced by growth in a chemostat under continuous culture conditions. This method allows the effect of individual growth parameters (e.g., pH, oxygen tension, growth rate, limiting substrate) on particular cellular characteristics to be identified. Chemostat cultures can be used to study cells under "steady-state" conditions, or during transitions from one steady state to another (103).

In principle any liquid culture medium that supports growth in batch culture can be used for continuous culture. However, because the pH is actively controlled during growth, it is not necessary to include components required only for buffering pH. The concentration of the limiting nutrient determines the culture density, and all other essential nutrients are supplied so that they are in excess. For aerobes, the concentration of the limiting nutrient is set so that the culture density is not so high as to impose an excessive oxygen demand. For *B. subtilis* this may mean limiting the culture density to about 1–2 g dry weight per liter. Complex media can be used if a defined limitation can be achieved (104). Several defined media have been described using, for example, phosphate, magnesium, potassium, glucose, and sulfate as limiting substrates.

Commercial Culture Conditions

Most of the products traditionally produced from *Bacillus* species, including enzymes, antibiotics, and insecticides, are synthesized during stationary phase using a traditional batch fermentor. However, insecticides require both a high growth yield and efficient sporulation because the δ-endotoxin accumulates in the sporangium. The production of α-amylase and of *B. thuringiensis* δ-endotoxin are given as examples of the factors that need to be taken into account during commercial *Bacillus* fermentations.

Most species of *Bacillus* grow well on commercial formulations of nutrient media, although in some cases the medium needs to be modified by adjustment of the pH or ionic strength or by the addition of specific nutrient supplements. This means that media can, to some extent, be based on cheap, readily available substrates. Despite the extensive expertise in the fermentation of *Bacillus* species, commercial considerations have limited the published literature on large-scale fermentations. Culture conditions and media formulations are highly geared to the particular strains and processes used and these have been selectively reviewed elsewhere (105–107).

Media components have been developed empirically over a number of years to avoid catabolite repression that prevents the synthesis of a number of the commercially important products from *Bacillus* species. For example, α-amylases are produced in media containing starch in various forms including corn, wheat malt, and potato that is noncatabolite repressing. When carbohydrates are used as a carbon source, phosphate is required at a concentration of about 100 mM. Peptones and amino acids are a better source of nitrogen than ammonium for *B. subtilis*, probably because of the efficiency with which nitrogen is assimilated into amido groups. In the case of *B. apiaries*, however, ammonium is the preferred nitrogen source. Glutamine is the nitrogen source that sustains the fastest growth rate. The use of amino acids as a nitrogen source represses the synthesis of proteases and the onset of sporulation, allowing the productive stationary phase to be extended.

The concentrations of divalent metal cations Mn^{2+}, Ca^{2+}, and Zn^{2+} are important to the final yield of α-amylase, even when they have little or no effect on the growth of the bacterium. For example, Ca^{2+} is required for both the enzymatic activity and structural integrity of the liquefying α-amylases such as those from *B. licheniformis*, *B. amyloliquefaciens*, and *B. stearothermophilus*. We have evidence that this results, at least in part, from the rate of folding of the enzyme on the *trans* side of the cytoplasmic membrane (62). Other metal ions, including Fe^{2+}, Fe^{3+}, Na^{+}, and Mg^{2+}, enhance enzyme production presumably in part as a result of their effect in increasing growth yield.

The growth kinetics in *Bacillus* cultures is complicated by the operation of the stringent response and sporulation, both of which are similarly induced by nutrient limitations and other negative growth factors. Commercial processes often use fully or partially (oligosporogenous) sporulation-deficient strains. The stringent response, modulated by variations in the levels of ppGpp(p), appears to be a cause of linear growth. Substrates with a low carbon to energy ratio gives rise to carbon-limited growth and consequential "energy spilling" and the production of overflow metabolites. Glucose excess chemostat cultures of *B. stearothermophilus* accumulate acetate up to 34% of the carbon consumed. Under these conditions, energy efficiency is sacrificed to obtain the ATP turnover required to sustain

the high growth rates. Similar metabolite overflows have been observed for *B. licheniformis* (48%) and *B. subtilis* (22%).

Catabolite repression is one of the major factors limiting the production of many commercially important products. Although mutants (e.g., *ccpA*) have been isolated that lack elements of the catabolite repression pathway, these often grow poorly. An alternative approach in which CRE elements are deleted or inactivated may be frustrated by their occurrence within DNA sequences encoding the protein itself. Often, the most effective method for overcoming catabolite repression involves a combination of appropriate media formulations and growth regimes.

Many commercially significant products are produced after the end of exponential growth (transition phase) and during the early stages of sporulation. It is therefore essential that, when sporulation-deficient strains are used, they still retain the ability to synthesize the required product. For example, mutants deficient in Spo0A (the sporulation-specific response regulator) are unable to induce various extracellular proteases (108).

Although batch cultures are traditionally used for α-amylase production, because the main production phase is during stationary phase, Emanuilova and Toda (109) compared the production of α-amylase by *B. caldolyticus* in batch and continuous culture. Casitone and starch were required for optimal α-amylase production. The concentration of amylase was 10- to 20-fold higher in continuous rather than batch culture, although there was an inverse relationship between growth rate and amylase production.

Industrial fermentations are carried out at 30–40 °C and at pH 7.0 with $CaCO_3$ as the calcium source to stabilize the enzyme. The complex medium (110,111) allows growth to 10^{10} cell/ml within 20 h, after which time production continues for about 6 days (111). The fermentations also produce proteases that are important for the utilization of protein substrates, but which have to be separated from the α-amylase during subsequent downstream processing.

α-Amylase production can be increased by cloning the structural gene onto a multicopy plasmid, and production, possibly up to the limitation of the native secretory pathway, is achievable (112). However, because many of the plasmids used in *Bacillus* species are based on staphylococcal plasmids (e.g., pUB110), both segregational and structural instabilities are frequently observed.

Because δ-endotoxin production of *B. thuringiensis* occurs during sporulation, growth and process conditions need to be optimized to give cell densities $>5 \times 10^9$ per milliliter and sporulation rates greater than 90%. Although significant strain-to-strain variations are observed, certain basic requirements can be defined: these include good oxygenation, a growth temperature between 26 and 30 °C (higher temperatures can lead to plasmid curing and consequential loss of δ-endotoxin production [113]), and pH 6.5–7.5. The production of amylases and proteases by *B. thuringiensis* allows a wide range of natural raw materials to be used as growth substrates. Glucose is catabolized with the production of acetate, lactate, pyruvate, and acetoin (114,115), and pH control is required for some strains to prevent the acidic conditions restricting growth. The production of poly-β-hydroxybutyrate during early stationary phase under carbon excess can be mistaken for δ-endotoxin but disappears during sporulation.

Cultures are usually grown in submerged liquid cultures as large as 200 L using fermentors agitated with paddle blades and with static baffles on the side wall of the vessel (99). Occasionally vessels with impellers or airlift fermentors are used. Because the industrial media contain significant amounts of insoluble particles, sterilization is accompanied by agitation. Toward the end of the fermentation, the released spores attach to and stabilize the foam, and silicon- or polypropylene-based antifoam agents are required as foam suppressers. Antifoam agents can form aggregates that interfere with downstream processing, and consequently their use is kept to a minimum. During stationary phase, oxygen consumption is markedly reduced and, provided saturation levels do not fall below 20%, agitation can be reduced, with a concomitant reduction in foaming. At the end of fermentation, the fermentation vessel will be heavily contaminated with spores that need to be inactivated if different strains are used in the next fermentation run.

Bacillus thuringiensis will grow on glucose minimal medium only when supplemented with sources of amino acids (e.g., peptones, casamino acids) or proteins. Because *B. thuringiensis* produces amylases, glucose can be replaced with starch as the carbon source. Growth is stimulated by yeast extract and optimal concentrations of Fe^{3+}, Co^{+}, and Zn^{+}, and sporulation is improved by metal ions such as Ca^{2+} and Mn^{2+}. Phosphate is added both as a source of phosphorus and as a buffer.

Preparations of *B. thuringiensis* are applied at a rate of about 0.5 to 2 kg per hectare. Spores can be applied directly as a pest control agent, but are often processed to improve their physical characteristics. For example, granular or bait formulations are applied to the plants or soil; in other cases, concentrates are diluted in water and applied by spraying. Stickers, spreaders, or even chemical insecticides can be added to improve dispersion and activity (99).

BIBLIOGRAPHY

1. F.G. Priest, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., *Bacillus subtilis and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 1–16.
2. D. Claus and R.C.W. Berkeley, in P.H.A. Sneath ed., *Bergey's Manual of Systematic Bacteriology*, Williams & Wilkins, Baltimore, Md., 1986, pp. 1105–1141.
3. F.G. Priest and C.R. Harwood, in Y.H. Hui and G.G. Khachatourians eds., *Food Biotechnology Micro-organisms*, VCH, New York, 1994, pp. 377–421.
4. F.G. Priest, in C.R. Harwood ed., *Biotechnology Handbook 2: Bacillus*, Plenum Press, New York, 1989, pp. 27–56.
5. C.R. Harwood, *Trends Biotechnol.* **10**, 247–256 (1992).
6. J.W. Kloepper, R. Lifshitz, and R.M. Zablotowicz, *Trends Biotechnol.* **7**, 39–44 (1989).
7. J.D. Wisotzkey, P.J. Jurtshuck, G.E. Fox, G. Reinhard, and K. Poralla, *Int. J. Syst. Bacteriol.* **42**, 236–269 (1992).
8. C.H. Ash, A.E. Farrow, S. Wallbanks, and M.D. Collins, *Lett. Appl. Microbiol.* **13**, 202–206 (1991).

9. D. Rössler, W. Ludwig, K.H. Schleiffer, C. Lin, T.J. McGill, J. Wisotzkey, P.J. Jurtshuk, and G.E. Fox, *Syst. Appl. Microbiol.* **14**, 266–269 (1991).
10. T. Hoffman, B. Troup, A. Szabo, C. Hungerer and D. Jahn, *FEMS Microbiol. Lett.* **131**, 219–225 (1995).
11. M.M. Nakano, Y.P. Dailly, P. Zuber, and D.P. Clark, *J. Bacteriol.* **179**, 6749–6755 (1997).
12. B.L. Russell, S.A. Jelly, and A.A. Yousten, *Appl. Environ. Microbiol.* **55**, 294–297 (1989).
13. C.R. Harwood and A.R. Archibald, in C.R. Harwood and S.M. Cutting eds., *Molecular Biological Methods for* Bacillus, Wiley, Chichester, UK, 1990, pp. 1–26.
14. W.L. Nicholson and P. Setlow, in C.R. Harwood and S.M. Cutting eds., *Molecular Biological Methods for* Bacillus, Wiley, Chichester, UK, 1990, pp. 391–450.
15. D. Claus and D. Fritze, in C.R. Harwood ed., *Biotechnology Handbook 2:* Bacillus, Plenum Press, New York, 1989, pp. 5–26.
16. F. Kunst, N. Ogaswara, I. Moszer, et al. *Nature* **390**, 249–256 (1997).
17. R.F. Doolittle, Nature **392**, 339–342 (1998).
18. S. Bron, R. Meima, J.-M. van Dijl, A. Wipat, and C.R. Harwood, in G. Cohen ed., *Molecular Biology of Industrial Microorganisms*, American Society for Microbiology, Washington, D.C., 1998, in press.
19. R.E. Yasbin, W. Firshein, J. Laffan, and R.G. Wake, in C.R. Harwood and S.M. Cutting eds., *Molecular Biological Methods for* Bacillus, Wiley, Chichester, UK, 1990, pp. 296–320.
20. J.P. Müller, Z. An, T. Merad, I.C. Hancock, and C.R. Harwood, *Microbiology* **143**, 947–956 (1997).
21. R.M. Horton, H.D. Hunt, S.N. Ho, J.K. Pullen, and L.R. Pease, *Gene* **77**, 61–68 (1989).
22. P. Youngman, in C.R. Harwood and S.M. Cutting eds., *Molecular Biological Methods for* Bacillus, Wiley, Chichester, UK, 1990, pp. 221–266.
23. P. Youngman, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 585–596.
24. M.-A. Petit, C. Bruand, L. Jannière, and S.D. Ehrlich, *J. Bacteriol.* **172**, 6736–6740 (1990).
25. J. Perkins and P. Youngman, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 140–144 (1986).
26. M.R. Wati, F.G. Priest, and W.J. Mitchell, *FEMS Microbiol. Lett.* **71**, 211–214 (1990).
27. J. Spizizen, *Proc. Natl. Acad. Sci. U.S.A.* **44**, 1072–1078 (1958).
28. D. Dubnau, *Microbiol. Rev.* **55**, 395–424 (1991).
29. D. Dubnau, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 555–584.
30. S. Chang and S.N. Cohen, *Mol. Gen. Genet.* **168**, 111–115 (1979).
31. S.M. Cutting, and P.B. Vander Horn, in C.R. Harwood and S.M. Cutting eds., *Molecular Biological Methods for* Bacillus, Wiley, Chichester, UK, 1990, pp. 27–74.
32. B. Michel, B. Niaudet, and S.D. Ehrlich, *EMBO J.* **12**, 1565–1571 (1982).
33. L.K. Jolliffe, R.J. Doyle, and U.N. Streips, *Cell* **25**, 753–763 (1981).
34. C.R. Harwood and A. Wipat, *FEBS Lett.* **389**, 84–87 (1996).
35. P.S. Vary, *Microbiology* **140**, 1001–1013 (1994).
36. R. Meima, G. Venema, and S. Bron, *Plasmid* **35**, 14–30 (1996).
37. S. Bron, in C.R. Harwood and S.M. Cutting eds., *Molecular Biological Methods for* Bacillus, Wiley, Chichester, UK, 1990, pp. 75–175.
38. S.D. Ehrlich, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 1680–1682 (1977).
39. A.D. Gruss and S.D. Ehrlich, *Microbiol. Rev.* **53**, 231–241 (1989).
40. S.D. Ehrlich, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 1433–1436 (1978).
41. K.J. Leenhouts and G. Venema, in K.G. Hardy ed., *Plasmids, a Practical Approach*, Oxford Univ. Press, New York, 1993, pp. 65–94.
42. C. Bruand, S.D. Ehrlich, and L. Jannière, *EMBO J.* **10**, 2171–2177 (1991).
43. L. Jannière, A. Gruss, and S.D. Ehrlich, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 625–644.
44. V. Vagner and S.D. Ehrlich, *J. Bacteriol.* **170**, 3978–3982 (1988).
45. A.M. Albertini and A. Galizzi, *J. Bacteriol.* **162**, 1203–1211 (1985).
46. M. Perego, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 615–624.
47. M. Simonen and I. Palva, *Microbiol. Rev.* **57**, 109–137 (1993).
48. J. Pero and A. Sloma, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 939–952.
49. G.H. Chamblis, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 213–219.
50. D.G. Yansura and D.J. Henner, in A.T. Ganesan and J.A. Hoch eds., *Genetics and Biochemistry of* Bacilli, Academic Press, Orlando, Fla., 1984, pp. 249–263.
51. D.M. Gartner, M. Geissendorfer, and W. Hillen, *J. Bacteriol.* **170**, 3102–3109 (1988).
52. A.-M. Crutz, M. Steinmetz, S. Aymerich, R. Richter, and D. LeCoq, *J. Bacteriol.* **172**, 1043–1050 (1990).
53. J. Errington, in C.R. Harwood and S.M. Cutting eds., *Molecular Biological Methods for* Bacillus, Wiley, Chichester, UK, 1990, pp. 175–220.
54. S.J. Thornewell, A.K. East, and J. Errington, *Gene* **133**, 47–53 (1993).
55. D.E. Eveleigh and B.S. Montenencourt, *Adv. Appl. Microbiol.* **25**, 57–74 (1979).
56. V.P. Malik, *Adv. Genet.* **20**, 38–126 (1979).
57. K. Yamane and B. Maruo, in K. Sakaguchi and M. Ouishi eds., *Molecular Breeding and Genetics of Applied Microorganisms*, Academic Press, New York, 1980, pp. 117–123.
58. E.A. Ferrari, A.S. Jarnagin, and B.F. Schmidt, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and*

Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics, American Society for Microbiology, Washington, D.C., 1993, pp. 917–937.

59. U.P. Shinde, J.J. Liu, and M. Inouye, *Nature* **389**, 520–522 (1997).
60. V. Nagarajan, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 713–726.
61. K. Stephenson and C.R. Harwood, *Appl. Environ. Microbiol.* **64**, 2875–2881 (1998).
62. K. Stephenson, C.R. Harwood, N.M. Carter, M.-F. Petit-Glatron, and R. Chambert, *FEBS Lett.* **432**, 385–389 (1998).
63. X. Wu, W. Lee, L. Tran, and S.L. Wong, *J. Bacteriol.* **173**, 4952–4958 (1991).
64. J. Errington, *Microbiol. Rev.* **57**, 1–33 (1993).
65. P. Stragier and R. Losick, *Annu. Rev. Genet.* **30**, 297–341 (1996).
66. A. Moir, I.M. Feavers, A.R. Zuberi, R. Sammons, I.S. Roberts, J.R. Yon, E.A. Wolff, and D. Smith, in J.A. Hock and P. Setlow eds., *Molecular Biology of Differentiation*, American Society for Microbiology, Washington, D.C., 1985, pp. 36–46.
67. F.M. Hulett, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 229–235.
68. W.D. Grant, *J. Bacteriol.* **137**, 35–43 (1979).
69. C. Eymann, H. Mach, C.R. Harwood, and M. Hecker, *Microbiology* **142**, 3163–3170 (1996).
70. T. Seki, H. Yoshikawa, H. Takahashi, and H. Saito, *J. Bacteriol.* **169**, 2913–2916 (1987).
71. F.M. Hulett, J. Lee, L. Shi, G. Sun, G.R. Chestnut, E. Sharkova, M.F. Duggan, and N.V. Kapp, *J. Bacteriol.* **176**, 1348–1358 (1994).
72. F.M. Hulett, G. Sun, and W. Liu, in A. Torriani-Gorini, E. Yagil, and S. Silver eds., *Phosphate in Microorganisms: Cellular and Molecular Biology*, American Society for Microbiology, Washington, D.C., 1994, pp. 50–54.
73. T. Seki, H. Yoshikawa, H. Takahashi, and H. Saito, *J. Bacteriol.* **170**, 5935–5938 (1988).
74. F.M. Hulett, *Mol. Microbiol.* **19**, 933–939 (1996).
75. G.F. Sun, S.M. Birkey, and M.F. Hulett, *Mol. Microbiol.* **19**, 941–948 (1996).
76. B.L. Wanner, in F.C. Neidhardt ed., Escherichia coli *and* Salmonella: *Cellular and Molecular Biology*, American Society for Microbiology, Washington, D.C., 1996, pp. 1357–1381.
77. M.H. Saier, M.J. Fagan, C. Hoischen, and J. Reizer, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 133–156.
78. C.J. Hueck, W. Hillen, and M.H. Saier, *Res. Microbiol.* **145**, 503–518 (1994).
79. S.H. Fisher, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 221–228.
80. H.J. Schreier, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 281–298.
81. L. Reitzer, in F.C. Neidhardt ed., Escherichia coli *and* Salmonella: *Cellular and Molecular Biology*, American Society for Microbiology, Washington, D.C., 1996, pp. 391–407.
82. C.S. Wright, R.A. Alden, and J. Kraut, *Nature* **221**, 235–242 (1969).
83. D.J. Neidhart and G.A. Petsko, *Prot. Eng.* **2**, 271–276 (1988).
84. S.J. Tomazic and A.M. Klibanov, *J. Biol. Chem.* **263**, 3086–3091 (1988).
85. S.J. Tomazic and A.M. Klibanov, *J. Biol. Chem.* **263**, 3092–3096 (1988).
86. M. Machius, G. Weigand, and R. Huber, *J. Mol. Biol.* **246**, 545–559 (1995).
87. M. Kanno, *Agric. Biol. Chem.* **50**, 23–31 (1986).
88. Y. Suzuki and T. Imai, *Appl. Microbiol. Biotechnol.* **21**, 20–31 (1986).
89. G.B. Madsen, B.E. Norman, and S. Slott, *Die Starkr.* **25**, 304–308 (1973).
90. F.G. Priest, in C.R. Harwood ed., *Biotechnology Handbook 2:* Bacillus, Plenum Press, New York, 1989, pp. 293–320.
91. S.W. Queener and D.H. Lively, in A.L. Domain and N.A. Solomon eds., *Manual of Industrial Microbiology and Biotechnology*, American Society for Microbiology, Washington, D.C., 1989, pp. 155–169.
92. P. Zuber, M.M. Nakano, and M.A. Marahiel, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 897–916.
93. D. Konz, A. Klens, K. Schörgendorfer, and M.A. Marahiel, *Chem. Biol.* **4**, 927–937 (1997).
94. N. Tsukagoshi, S. Iritani, T. Sasaki, T. Takamura, H. Ihara, Y. Idota, H. Yamagata, and S. Udaka, *J. Bacteriol.* **164**, 1182–1187 (1985).
95. R. Puohiniemi, S. Butcher, E. Tarkka, and M. Sarvas, *FEMS Microbiol. Lett.* **83**, 29–34 (1991).
96. A.L. Aronson, in A.L. Sonenshein, J.A. Hoch, and R. Losick, eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 953–963.
97. J.S. Feitelson, J. Payne, and L. Kim, *Bio/technology* **10**, 271–275 (1992).
98. J. Li, J. Carroll, and D.J. Ellar, *Nature* **353**, 815–821 (1991).
99. S. Ely, in P.F. Entwistle, J.S. Cory, M.J. Bailey, and S. Higgs eds., Bacillus thuringiensis, *an Environmental Biopesticide: Theory and Practice*, Wiley, Chichester, UK, 1993, pp. 105–124.
100. A. Sietske de Boar and B. Diderichsen, *Appl. Microbiol. Biotechnol.* **36**, 1–4 (1991).
101. P. Schaeffer, J. Millet, and P.J. Aubert, *Proc. Natl. Acad. Sci. U.S.A.* **54**, 704–711 (1965).
102. U. Wanner and T. Egli, *FEMS Microbiol. Rev.* **75**, 19–44 (1990).
103. T. Merad, A.R. Archibald, I.C. Hancock, C.R. Harwood, and J.A. Hobot, *J. Gen. Microbiol.* **135**, 645–655 (1989).
104. A.R. Archibald, in I.C. Hancock and I.R. Poxton eds., *Bacterial Cell Wall Techniques*, Wiley, Chichester, UK, 1988, pp. 37–54.

105. R.J. Sharp, M.D. Scawen, and T. Atkinson, in C.R. Harwood ed., *Biotechnology Handbook 2*: Bacillus, Plenum Press, New York, 1989, pp. 255–292.

106. F.G. Priest and R.J. Sharp, in J.O. Neway ed., *Fermentation Process Development of Industrial Organisms*, Dekker, New York, 1989, pp. 73–132.

107. M.V. Arbige, B.A. Bulthuis, J. Schultz, and D. Crabb, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 871–895.

108. E. Ferrari, D.J. Henner, M. Perego, and J.A. Hoch, *J. Bacteriol.* **170**, 289–295 (1988).

109. E.I. Emanuilova and K. Toda, *Appl. Microbiol. Biotechnol.* **19**, 301–305 (1984).

110. L.A. Underkofler, in B.M. Miller and W. Litsky eds., *Industrial Microbiology*, McGraw-Hill, New York, 1976, pp. 128–146.

111. K. Aunstrup, *Appl. Biochem. Bioeng.* **2**, 27–69 (1976).

112. J. Vehmaanpera, P.N.A. Nybergh, R. Tanner, E. Pohjonen, R. Bergelin, and M. Korhola, *Enzyme Microb. Technol.* **9**, 546–549 (1987).

113. J.M. Gonzalez and B.C. Carlton, *Plasmid* **11**, 28–38 (1984).

114. K.W. Nickerson, G. St. Julian, and L.A. Bulla, *Appl. Environ. Microbiol.* **28**, 129–132 (1974).

115. T.G. Benoit, G.R. Wilson, and C.L. Baugh, *Lett. Appl. Microbiol.* **10**, 15–18 (1990).

BIOCATALYSIS DATABASES

LAWRENCE P. WACKETT
LYNDA B.M. ELLIS
University of Minnesota
Minneapolis, Minnesota

KEY WORDS

Biocatalysis
Biodegradation
Bioremediation
Microbial catabolism
World Wide Web

OUTLINE

INTRODUCTION TO MICROBIAL BIOCATALYSIS

The microbial transformation of organic compounds has been studied at the molecular level only during the last century of human history, driven by the development of techniques for determining the structure of organic compounds and cultivating microorganisms in pure culture. As information on these transformation reactions accrued, it became clear that microorganisms were, via their diverse metabolic pathways, principally responsible for the recycling of complex organic compounds on Earth. Almost simultaneously, it was recognized that the enormous diversity of microbial enzymes represents an invaluable resource for biosynthesizing desirable molecules. Historically, applications predate an understanding of the process, for example, the fermentation of sugars to yield ethanol. Although alcoholic beverage production was recorded in early written records, its biocatalytic basis was only proposed in the nineteenth century. Its enzymatic basis was first revealed in 1897 using a microbial cell-free system (1).

Now we appreciate that microorganisms collectively have evolved the greatest enzymatic diversity found on Earth to metabolically degrade disparate organic compounds, and thus obtain chemical energy for adenosine triphosphate (ATP) synthesis. This metabolism, called catabolism or biodegradation, is a major contribution to the Earth's carbon cycle (2). Microorganisms are increasingly used in engineered systems to biodegrade hazardous, xenobiotic compounds; this application is called bioremediation.

In the twentieth century, work on microbial biodegradation of organic molecules and on biocatalysis to yield desirable chemicals has gone hand in hand. The versatile enzymes found in biodegradation pathways are sometimes effective biocatalysts for specialty chemical synthesis. For example, recombinant *Escherichia coli* strains expressing naphthalene dioxygenase, the first enzyme in the biodegradation pathway of naphthalene and other polycyclic aromatic hydrocarbons, are used in an emerging commercial process for the environmentally benign production of the blue jean dye indigo (3). The process is based on the regiospecificity of naphthalene dioxygenase that dioxygenates indole at carbons 2 and 3 to yield an unstable dihydrodiol that spontaneously eliminates water. The resultant compound, indoxyl, spontaneously dimerizes to yield indigo. In a commercial biocatalytic process, indole is not supplied exogenously, because of its expense, but is generated in situ via overexpression of the enzymes functioning in the tryptophan biosynthetic pathway (4). A major impetus for bacterial indigo production is to preclude the formation of hazardous side products that are formed in the chemical process for indigo manufacture. Thus, bioca-

talysis and pollution prevention by biodegradation are intimately intertwined.

THE INFORMATION EXPLOSION AND THE WORLD WIDE WEB

New developments in commercial biosynthesis and environmental bioremediation are certain to emanate from a wider availability of information on microbial biocatalysis and biodegradation. Because the number of chemical substances subjected to microbial catabolism is large (10^6–10^7), a comparably large number of distinct enzymes and pathways must occur in nature. The vast majority have not yet been studied, in contrast to the better-characterized enzymes and pathways in common intermediary metabolism. Information on these reactions or pathways is not typically available in biochemistry or microbiology textbooks.

The increasing desire for this information has led to a virtual explosion of information published in journals, specialty books, and symposium proceedings over the past 2 years. There are also sources that contain extensive information but that are less readily available, for example, technical reports of government agencies such as the U.S. Environmental Protection Agency (EPA). The information retrieval problem is exacerbated by the different clientele for this information, including microbiologists, chemists, chemical engineers, and civil engineers. Most of these groups do not routinely survey the same publications and are accustomed to somewhat different presentations of the relevant data.

The rapid growth of the World Wide Web (WWW or Web) offers a way to overcome availability, outdatedness, and inconsistency problems because it can provide timely access to large amounts of appropriate, consistently formatted information. The Web builds on the success of the Internet, a collection of thousands of linked computer networks. Since the Internet's beginnings in the early 1970s, many scientists have used electronic mail as a very effective way to convey information. In January 1998, the Internet had approximately 30 million hosts (5), and its use is still growing rapidly.

In 1991, a client/server software system called Gopher was developed to allow individual client computers to access textual information on the computer networks of the Internet. Gopher, developed at the University of Minnesota, appears to be a hierarchical text database that allows the user to "go fer" (look for) Internet resources on various topics and display them on even a modest computer. The World Wide Web, developed shortly thereafter at the European Laboratory for Particle Physics, combines graphics with text and allows for links between related Internet resources. For example, clicking on a highlighted or underlined term transports the user to the designated Internet resource, which can be on a server housed anywhere in the world.

BIOCATALYSIS INFORMATION ON THE WORLD WIDE WEB

The Web information of most value to people involved in biocatalysis deals with chemical compounds, microorganisms, metabolic pathways, enzymes, and genes. Users can access each field of information separately, although in some cases there are hypertext links that facilitate finding related information in independent databases. There are commercial databases that cover biocatalysis, for example, focusing on the use of enzymes in organic synthesis. These databases are not freely available and thus are not linked to public Web databases and are not discussed here.

For each of the fields of chemical compounds, microorganisms, metabolic pathways, enzymes, and genes, Internet resources are listed with brief annotations in Table 1, along with their uniform resource locators (URLs). The enzyme and metabolism databases focus most directly on the biocatalysts that might be employed in organic synthesis or metabolic engineering. Most metabolism and enzyme databases, however, deal largely with the reactions common to most living things, often denoted as intermediary metabolism. Emerging applications in biocatalysis are less likely to involve such reactions, being more likely to be derived from metabolism that is carried out by a more limited range of organisms, typically bacteria. As an example, nitrile hydratase from *Rhodococcus* sp. is now used in the commercial synthesis of acrylamide from acrylonitrile. This is not a reaction carried out by animals, or even *E. coli*, but it nonetheless provides for clean catalysis with high product yield and purity. These more exotic enzymatic reactions are described in the first of the Web resources listed in Table 1, the University of Minnesota Biocatalysis/Biodegradation Database. It is described next in more detail.

THE UNIVERSITY OF MINNESOTA BIOCATALYSIS/BIODEGRADATION DATABASE

Introduction

The University of Minnesota Biocatalysis/Biodegradation Database (UM-BBD) first appeared on the World Wide Web (*http://www.labmed.umn.edu/umbbd/index.html*) on February 25, 1995, to meet the information needs of this distinct, diverse scientific community (6,7). It was designed to provide information on specialized nonintermediary microbial metabolism that is at the heart of scientific efforts in biocatalysis and biodegradation, in an easily accessible format, and to interface readily with Web-based databases of intermediary metabolism, such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) (8). Given that there are 10^6–10^7 organic compounds subject to catabolism, and that their catabolism is largely nonintermediary and unique to selected microorganisms, the UM-BBD has a much larger information domain to cover than intermediary metabolism databases. At the same time, information contained in the UM-BBD has potentially the greatest impact. Hazardous waste treatment is a $200 billion industry worldwide, and bioremediation comprises 5–10% of that total. There is the potential for an expansion of this market with the development of new methods for metabolic and protein engineering. For example, researchers working on biodegradative pathways have begun to use metabolic engineering to develop new metabolic routes to handle particularly problematic pollutants (9). One example de-

Table 1. Useful Internet Resources for Microbial Biocatalysis

University of Minnesota Biocatalysis / Biodegradation Database
http: / / www.labmed.umn.edu / umbbd / index.html
The University of Minnesota Biocatalysis/Biodegradation Database (UM-BBD) provides well-organized information on microbial metabolism of primarily xenobiotic nonintermediary compounds. Its metabolic pathways include appropriate hypertext links to compound, reaction and pathway information in complementary Web databases. The UM-BBD also provides a list of useful Internet resources for microbial biotechnology, a few of which are described below.
Intermediary Metabolism
KEGG: Kyoto Encyclopedia of Genes and Genomes
http: / / www.genome.ad.jp / kegg / kegg.html
KEGG contains current knowledge of molecular and cellular biology in terms of information pathways of interacting genes or molecules and links individual components of the pathways with the gene catalogs produced by genome projects.
WIT
http: / / www.cme.msu.edu / WIT
WIT helps users reconstruct metabolism from complete bacterial genomic sequences.
Chemical Compounds
ChemFinder
http: / / chemfinder.camsoft.com /
ChemFinder is a database containing thousands of chemical compounds, including, for each compound, physical property data, chemical structures, and links to external information.
National Toxicology Program
http: / / ntp-server.niehs.nih.gov /
The NTP website contains information about the Program, Chemical Health & Safety Data, 2D and 3D chemical structures, and abstracts of Short-Term Toxicity and Long-Term Carcinogenesis studies.
Enzymes and Genes
Ligand Chemical Database for Enzyme Reactions
http: / / www.genome.ad.jp / dbget / ligand.html
LIGAND contains entries for enzymes and their substrates, products, and cofactors.
GenBank
http: / / www.ncbi.nlm.nih.gov / Web / Genbank / index.html
GenBank is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences.
Microorganisms
World Data Center for Microorganisms
http: / / wdcm.nig.ac.jp /
The WDCM provides a comprehensive directory of culture collections, databases on microbes and cell lines, and a getaway to biodiversity, molecular biology, and genome projects.
Bacterial Nomenclature Up-to-Date
http: / / www.gbf-braunschweig.de / DSMZ / bactnom / bactname.htm
This site, compiled by the DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkuturen GmbH, includes all bacterial names that have been validly published since January 1, 1980, and nomenclatural changes which have been validly published since then.

Source: From Ref. 34.

scribes using a knowledge of existing biodegradative enzymes to design new pathways to metabolize polyhalogenated compounds (10). The UM-BBD could clearly stimulate similar projects by compiling and disseminating the crucial information needed to design such pathways de novo. We briefly describe the present status of the UM-BBD, including its format, use and users, peer acceptance, scientific advisory board, and future plans.

Format

In its first year of operation, the UM-BBD has evolved a standard format for reaction and compound pages. Both can link to many other internal and external information sources. One reaction page, for the first step in the biodegradation pathway for 1,2-dichloroethane, and a few of its many links, are shown in Figure 1. The Entrez Medline link (Fig. 1e) is especially interesting. The link back to the UM-BBD is prominent at the top. Also, when Entrez added a protein structure database, at our request they inserted the “Structure” link that is seen at the top of Figure 1e, so that our users would be only a click or two away from viewing and manipulating any available tertiary structure of the enzymes contained in the UM-BBD. Such a display for haloalkane dehalogenase is shown in Figure 1b. The

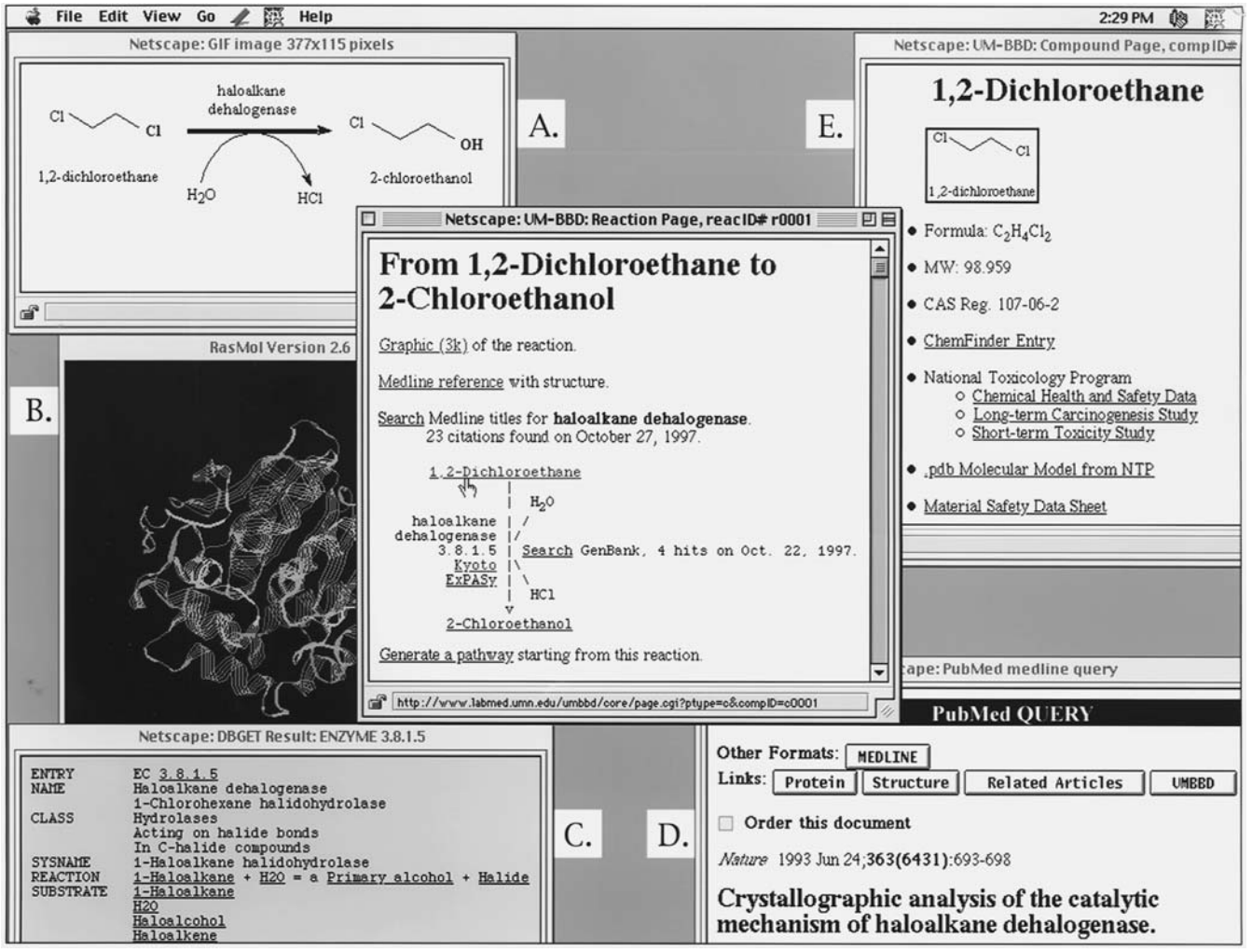


Figure 1. Selected windows of information from the UM-BBD, as viewed on a computer screen, representing the range of information available using the first reaction of 1,2-dichloroethane catabolism as an example. The respective windows are organized with the reaction page at the center and links from **a** to **e** at the periphery, starting with **a** in the upper-left corner and going counterclockwise. The six windows are (*center*) the haloalkane dehalogenase reaction page, (**a**) a reaction GIF showing the substrate and product structures, (**b**) a RasMol depiction of the haloalkane dehalogenase structure, (**c**) the enzyme page from the Kyoto Ligand Chemical Database, (**d**) the Entrez/Medline report on a selected paper on haloalkane dehalogenase, and (**e**) the 1,2-dichloroethane compound page.

UM-BBD now includes links to graphics of enzyme reaction mechanisms on selected reaction pages. An index to graphics helps users find them (11). The usefulness of a database is a function both of the information it contains and the access users have to this information.

Initially, the only way to access UM-BBD information was through browsing lists of pathways, reactions, and compounds. As the UM-BBD grew, browsing became an increasingly less convenient mode of access. The ability to search the database was then added (12). At present one can search for compounds by full or partial name or synonym (e.g., toluene or methyl benzene), CAS registry number (e.g., 108-88-3), or formula (e.g., C7H8), and for enzymes by full or partial name (e.g., urease) or full or partial EC number (e.g., 3.5.1.5 or 3.5.1). This search option will continue to be developed with input from the UM-BBD user community and may be expanded in the future to include chemical substructure searches.

Use and Users

Two months after the UM-BBD initially appeared on the Web, we began collecting use statistics. Usage has grown continuously; use statistics by Internet domain for 1997 are shown in Table 2. Currently, the UM-BBD has approximately 80,000 accesses per month (Fig. 2). U.S. educational (30%) and industrial (20%) users together make up half of all present use (Table 2). UM-BBD users are invited to "sign our guestbook" (i.e., voluntarily complete a seven-item electronic form). Those who sign are given the option

Table 2. UM-BBD Use Statistics by Domain from January 1 to December 31, 1997

#reqs	%bytes	Domain
187790	29.54%	.edu (USA Educational)
166052	19.66%	.com (USA Commercial)
87740	18.86%	[unresolved numerical addresses]
39050	8.13%	.net (Network)
12560	2.70%	.ca (Canada)
10008	2.40%	.de (Germany)
9527	2.23%	.uk (United Kingdom)
5944	1.56%	.gov (USA Government)
13285	1.45%	[Unknown]
4193	1.14%	.jp (Japan)
4960	1.07%	.au (Australia)
3950	0.97%	.fr (France)
3494	0.94%	.nl (Netherlands)
3718	0.68%	.se (Sweden)
2212	0.60%	.es (Spain)
2352	0.51%	.dk (Denmark)
1954	0.49%	.ch (Switzerland)
1948	0.46%	.us (United States)
2097	0.45%	.it (Italy)
1907	0.42%	.kr (South Korea)
1621	0.36%	.org (Non-Profit Making Organisations)
1473	0.34%	.mx (Mexico)
2210	0.33%	.il (Israel)
1332	0.31%	.be (Belgium)
1224	0.28%	.br (Brazil)
1014	0.26%	.fi (Finland)
1388	0.24%	.tw (Taiwan)
949	0.23%	.cz (Czech Republic)
925	0.23%	.no (Norway)
1022	0.22%	.mil (USA Military)
761	0.17%	.at (Austria)
672	0.14%	.sg (Singapore)
507	0.14%	.za (South Africa)
768	0.14%	.nz (New Zealand)
526	0.14%	.gr (Greece)
596	0.13%	.ar (Argentina)
467	0.11%	.pt (Portugal)
422	0.11%	.ru (Russian Federation)
468	0.11%	.my (Malaysia)
390	0.11%	.ie (Ireland)
524	0.11%	.co (Colombia)
406	0.10%	.th (Thailand)
436	0.09%	.pl (Poland)
381	0.09%	.su (Former USSR)
317	0.09%	.id (Indonesia)
284	0.08%	.cl (Chile)
331	0.08%	.hu (Hungary)
259	0.05%	.ee (Estonia)
131	0.05%	.hr (Croatia)
215	0.04%	.si (Slovenia)
112	0.04%	.sk (Slovak Republic)
258	0.04%	.tr (Turkey)
172	0.03%	.hk (Hong Kong)
185	0.03%	.in (India)
125	0.03%	.cn (China)
131	0.02%	.cr (Costa Rica)
78	0.02%	.ro (Romania)
124	0.02%	.lt (Lithuania)
70	0.02%	.bh (Bahrain)
64	0.02%	.pe (Peru)
104	0.01%	.ve (Venezuela)
71	0.01%	.uy (Uruguay)
69	0.01%	.eg (Egypt)
77	0.01%	.arpa (Old style Arpanet)
29	0.01%	.jo (Jordan)
57	0.01%	.ae (United Arab Emirates)
63	0.01%	.ph (Philippines)
31		.is (Iceland)
36		.yu (Yugoslavia)
18		.jm (Jamaica)
25		.pa (Panama)
23		.lb (Lebanon)
10		.do (Dominican Republic)
11		.lu (Luxembourg)
16		.pr (Puerto Rico)
11		.bm (Bermuda)
20		.ec (Ecuador)
10		.ua (Ukraine)
5		.zw (Zimbabwe)
5		.az (Azerbaidjan)
4		.md (Moldavia)
8		.gy (Guyana)
3		.am (Armenia)
8		.by (Belarus)
5		.mk (Macedonia)
9		.kw (Kuwait)
6		.bg (Bulgaria)
7		.gb (Great Britain)
9		.mt (Malta)
1		.dm (Dominica)
5		.gt (Guatemala)
4		.dz (Algeria)
9		.et (Ethiopia)
6		.cy (Cyprus)
7		.ge (Georgia)
6		.pg (Papua New Guinea)
2		.bo (Bolivia)
2		.pk (Pakistan)
2		.mu (Mauritius)
2		.lv (Latvia)
1		.om (Oman)

UM-BBD, University of Minnesota Biocatalysis/Biodegradation Database.

to regularly receive e-mail progress reports, and most choose to do so. This e-mail user list presently has about 400 members, who are surveyed on an annual basis. Users have recommended pathways and features to us, and we have acted on some of their suggestions: A pathway for degradation of a chlorinated aliphatic compound and graphical pathway maps were added in part because of user requests.

Connections

When the UM-BBD was first installed on a server in February 1995, it linked to several Internet resources. Links to others have been added since then. Most of the Web databases the UM-BBD links to have been contacted and collaborate on cross-linkage where possible. For example, the UM-BBD now has reciprocal links to Entrez PubMed (13), Entrez Nucleotide (14), Entrez Protein (15), ChemFinder (16), and Ligand (17), and others are under discussion. Each such interconnection is usually the result of individualized programming. The UM-BBD links to others

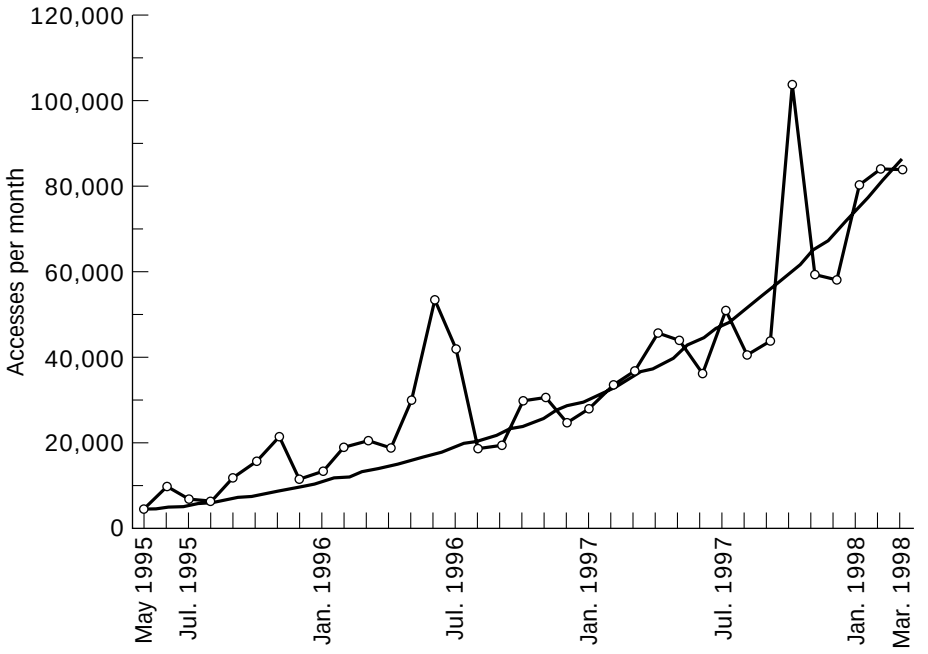


Figure 2. UM-BBD access statistics from May 1995 to March 1998.

in separate text lines, or, especially for reaction information, as links inserted in the text-formatted reaction (see Fig. 1, center).

Unique among present Web-based metabolic databases such as KEGG (8), the UM-BBD specializes in a more broadly based set of metabolic reactions. Because 10 million organic compounds are known and most are enzymically biodegraded, the potential range of the UM-BBD is enormous. These reactions are under active study now and will be in the foreseeable future, and the ability to search for current information in remote databases, as opposed to more static links to individual documents, is very important to UM-BBD users.

At present, the UM-BBD provides dynamic search links to EXTOXNET (18), Entrez/MEDLINE (13), and Entrez/GenBank (14). Although more difficult to construct and maintain than links to individual documents (in the first 6 months of 1996, each of these links was reconstructed at least once because the Web sites changed search format), they allow UM-BBD users to obtain the very latest information installed on these servers. Thus, a search is installed in preference to an individual link wherever possible, even if it returns only one document at present. Tomorrow or next week additional items may appear, and when they do UM-BBD users will have access to them.

UM-BBD search links include information on the number of hits that were returned when the search was last tested, the date of that search, and, if possible, the exact records being searched and terms that are being searched for. For example, Entrez/MEDLINE searches may be carried out for MEDLINE titles and abstracts, or if this produces too many hits, they may be restricted to titles alone. Entrez/MEDLINE and GenBank searches require custom formulation to return the best results because of the ambiguities and complexities of chemical and biochemical nomenclature and database search syntax.

Peer Acceptance

As mentioned earlier, the UM-BBD links to other metabolic databases. It also provides users with links to relevant information in several other biological and chemical Internet databases. The five prestigious international Web databases that it links to most closely—ChemFinder (16), Ligand (17), Entrez/Medline (13), Entrez/Protein (15), and Entrez/GenBank (14)—all link back to it. The UM-BBD has also been selected to appear on many peer-reviewed scientific Web referral lists, including Internet Chemical Resources (19), the American Society of Microbiology Biology Resources List (20), and BioTech: Biochemistry, Biophysics, and Molecular Biology (21). In addition, it was the subject of an article in the American Chemical Society's publication, *Chemical & Engineering News* (22), and in *Science* (23), and articles on it were solicited by the *Society of Industrial Microbiologists News* (6) and the *Journal of Microbial Methods* (7).

Scientific Advisory Board

During the first year of operation, a world-renowned scientific advisory board was established, including members in three countries from educational institutions, commercial companies, and government agencies: Peter Chapman (EPA, Florida); I. C. Gunsalus (EPA, Florida); Andreas Kiener (Lonza AG, Switzerland); Thomas Leisinger (ETH Zurich, Switzerland); and Jack T. Trevors (Guelph University, Canada). Most recently, Alfred Spormann (Stanford University) has joined the UM-BBD advisory board.

Growth and Future Developments

The UM-BBD began with only 4 pathways in February 25, 1995 and grew to 13 pathways by September 15, 1996, 25 pathways by February 20, 1997, and 59 pathways by April

21, 1998. However, with 10 million compounds in the biosphere, it will never, and is not meant to, contain information on the biodegradation of more than a representative fraction of these compounds. It was developed to contain reactions and pathways that span the diversity of organic functional groups and microbial metabolism. The next step in its development is to use this representative information to develop ways to predict microbial catabolic metabolism for compounds that the UM-BBD does not contain.

CONCLUSIONS

Biocatalysis will expand in scope as a knowledge of microbial catabolism increases. Most of the more than 10 million organic compounds are thought to be biodegradable, providing a virtually limitless potential for expansion of known microbial enzymatic reaction types. Additionally, information is rapidly emerging from the more than 60 bacterial genome sequencing projects completed or ongoing. This will lead to a new era of microbial metabolism studies, building on the extensive data on intermediary metabolism. These studies will increasingly reveal the vast catabolic capabilities of bacteria from diverse environments and taxonomic classes.

Research advances will create a need for more people trained in microbial genomics, metabolic reconstruction, and enzymology. The UM-BBD, and associated courses, can contribute to training the next generation of microbial biocatalysis specialists. The UM-BBD has been used instructionally at several colleges and universities, including the University of Minnesota (24–27), Iowa State University (28), Pomona College (29), the University of Waterloo (30), the University of Basel (31), and Hebrew University in Rehovot, Israel (32). A University of Minnesota course on biocatalysis and biodegradation is being offered completely over the Internet, using the UM-BBD as its focal point (24). The biocatalysis and biodegradation course instructs students on the general principles of microbial environmental biotechnology and microbial catabolic reactions. Using these fundamentals, students peruse the UM-BBD and related World Wide Web databases to research specific microbial transformation reactions. Student assignments are transferred to the instructors for assessment via electronic mail and the Web. This is a venue for advanced college-level instruction that is increasingly becoming available to students worldwide. It reflects a fundamental change in university education: for 6,000 years students were required to come to the scrolls, books, and personal contact with scholars that allowed learning to advance (33). Now, people can be at remote sites and increasingly access the accumulated knowledge of humanity via electronic means. The UM-BBD is one small part of this profound change in human information transfer and learning.

BIBLIOGRAPHY

1. E. Buchner, *Ber. Dtsch. Chem. Ges.* **30**, 117–124 (1897).
2. D.T. Gibson, *J. Ind. Microbiol.* **12**, 1–12 (1993).
3. B.D. Ensley, B.J. Ratzkin, T.D. Osslund, M.J. Simon, L.P. Wackett, and D.T. Gibson, *Science* **222**, 167–169 (1983).
4. H. Bialy, *Nat. Biotechnol.* **15**, 110 (1997).
5. R.H. Zakon, Robert H'obbes' "Hobbes' Internet Timeline" (May 1998): *http://info.isoc.org/guest/zakon/Internet/History/HIT.html*
6. L.B.M. Ellis and L.P. Wackett, *Soc. Ind. Microb. News* **45**, 167–173 (1995).
7. L.P. Wackett and L.B.M. Ellis, *J. Microbiol. Methods* **25**, 91–93 (1996).
8. KEGG: Kyoto Encyclopedia of Genes and Genomes: *http://www.genome.ad.jp/kegg/kegg.html*
9. K.N. Timmis, R.J. Steffan, and R. Unterman, *Annu. Rev. Microbiol.* **48**, 525–557 (1994).
10. L.P. Wackett, M.J. Sadowsky, L.M. Newman, H.-G. Hur, and S. Li, *Nature* **368**, 627–629 (1994).
11. S. Stephens, "UM-BBD Graphic Index" (May, 1998): *http://www.labmed.umn.edu/umbbd/search/graph_BBD.html*
12. R. McLeish and D. Hershberger, "UM-BBD Search Page" (May 1998): *http://www.labmed.umn.edu/umbbd/search/index.html*
13. National Center for Biotechnology Information (May 1998) Entrez/PubMed: *http://www.ncbi.nlm.nih.gov/PubMed/index.html*
14. National Center for Biotechnology Information (May 1998) Entrez/Nucleotide: *http://www.ncbi.nlm.nih.gov/Entrez/nucleotide.html*
15. National Center for Biotechnology Information (May 1998) Entrez/Protein: *http://www.ncbi.nlm.nih.gov/Entrez/protein.html*
16. CambridgeSoft (May 1998) ChemFinder: *http://chemfinder.camsoft.com/*
17. Kyoto University Ligand Chemical Database (May 1998): *http://www.genome.ad.jp/dbget-bin/www_bfind?ligand*
18. EXTOXNET: Extension Toxicology Network: *http://ace.ace.orst.edu/info/extoxnet/*
19. Internet Chemical Resources: *http://www.rpi.edu:80/dept/chem/cheminfo/chemres.html*
20. American Society for Microbiology—other biology sources: *http://www.asmusa.org/others.htm*
21. BioTech: Biochemistry, Biophysics, and Molecular Biology: *http://golgi.harvard.edu:80/biopages/biochem.html*
22. J. Kreiger, *Chem. & Eng. News*, p. 36–37, September 11 (1995).
23. D. Voss, *Science* **277**, 53 (1997).
24. Biocatalysis and Biodegradation, University of Minnesota: *http://biosci.cbs.umn.edu/class/bioc/5309/*
25. Advanced Biochemistry I, University of Minnesota: *http://biosci.cbs.umn.edu/class/bioc/8001/*
26. Ecological Biochemistry, University of Minnesota: *http://biosci.cbs.umn.edu/class/bioc/5301/*
27. M.C. Flickinger, *Applied Microbial Biochemistry*, Univ. of Minnesota, St. Paul, Minn. (1997).
28. Biotechnology Instructional Component, Iowa State University (1998): *http://www.lib.iastate.edu/biotech/biotech2/instr_bt.html*
29. Environmental Chemistry Sites, Pomona College, California (May 1998): *http://pages.pomona.edu/~wsteinmetz/envchem.htm*
30. Environmental Microbiology, University of Waterloo (May 1998): *http://bordeaux.uwaterloo.ca/biol447new/wwwsitesbiodegradation.htm*

31. Chemistry Sites, University of Basel, Switzerland (May 1998): *http://www.chemie.unibas.ch/Sites/Chemistry.html*
32. R. Mandelbaum, *Environmental Microbiology*, Hebrew University, Rehovot, Israel (1998).
33. E. Noam, *Science* **270**, 247–249 (1995).
34. L.B.M. Ellis and L.P. Wackett, *Nat. Biotechnol.* **15**, 1406 (1997).

See also ENZYMES, DIRECTED EVOLUTION; ENZYMES, EXTREMELY THERMOSTABLE; LACTONES, BIOCATALYTIC SYNTHESIS; LACTONOHYDROLASE; PHENYLALANINE DEHYDROGENASE.

BIOCATALYSIS, OPTICAL RESOLUTION.

See OPTICALLY ACTIVE 1,2-DIOLS, MICROBIAL PRODUCTION BY STEREOINVERSION; OPTICAL RESOLUTION, BIOCATALYSIS.

BIOENERGETICS OF MICROBIAL GROWTH

J.J. HEIJNEN
Delft University of Technology
DelftThe Netherlands

KEY WORDS

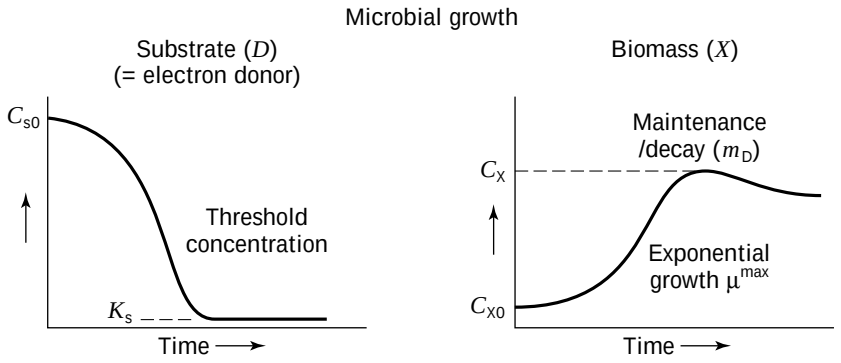
Biomass yield
Conservation constraints
Heat production
Maintenance
Maximal growth rate
Thermodynamics

OUTLINE

INTRODUCTION

The growth of microorganisms occurs within a wide range of pHs and temperatures, and on a wide variety of nutrients. Figure 1 shows a typical batch experiment where a substrate (starting at concentration C_{so}) is converted by a microorganism (C_{xo} at $t = 0$). The microorganism grows exponentially at a specific growth rate μ^{max} and with yield Y^m_{DX}. After depletion of the substrate, which is characterized by the substrate affinity constant K_s and a threshold concentration, the biomass concentration reaches $C_x = C_{xo} + Y^m_{DX}C_{so}$. Subsequently the biomass concentration decreases due to maintenance and/or biomass decay, which is characterized by the maintenance coefficient m_D (or the decay coefficient k_d). The relevant substrate always acts as electron donor, and therefore it is proper to define the biomass yield and maintenance on donor (D).

In the design of processes with growing microorganisms (fermentation processes and biological waste-treatment processes) the key parameters that need to be considered are the maximal biomass yield on substrate (Y^m_{DX}), the substrate maintenance coefficient m_D, the maximal growth rate (μ_{max}), and the substrate affinity constant (K_s). These four key parameters are sufficient to describe growth of



Four basic parameters

	Units	Range
Stoichiometric		
Y_{DX}^{m}	[C-mol X/(C)-mol D]	0.010–0.7
m_D	[(C)-mol D/C-mol X/h]	0.01–3
Kinetic		
μ^{max}	(h^{-1})	0.005–2
K_s	(mol D/L)	10^{-6}–10^{-3}

Figure 1. A typical microbial growth batch profile of biomass concentration and substrate concentration.

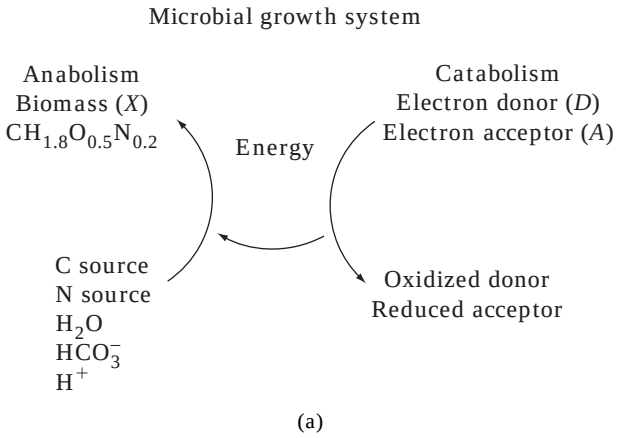
microorganisms in a standard mathematical model (1–4). A practical problem is, however, that the values of these parameters can vary by more than two orders of magnitude for different electron donors or acceptors used by different microorganisms, as indicated in Figure 1. It should be realized that conventionally stoichiometric parameters are expressed with C-mol of biomass, C-mol of electron donor for organic, and mol of donor for inorganic donors (C-mol X/(C)-mol D). It is therefore of interest to provide a general method to estimate values of these parameters for any chemotrophic growth system. Such methods have been provided by, for example, Battley (5), Roels (1), and Westerhoff (6). Recently (2,9) these methods have been critically evaluated with respect to general applicability and internal consistency. It was concluded (2,7,8) that none of these methods was satisfying. However a new method was proposed that is generally applicable and lacks the mentioned problems (2,8). Further, it should be recognized that in growth processes not only biomass production (r_x in C-mol biomass per m^3 reactor/h) and electron donor (substrate) consumption r_D in C-mol substrate (for carbon compounds), or mol substrate (for noncarbon compounds) per m^3 reactor/h are important. Also, the other conversions, such as O_2 consumption, N source consumption, heat production, and CO_2 production are highly relevant for the process design to calculate, for example, the required O_2 and heat transfer. Clearly, the full stoichiometry of the growth process should also be calculated and methods to achieve this are of major interest.

A STANDARD DESCRIPTION OF MICROBIAL GROWTH STOICHIOMETRY

The stoichiometry of microbial growth is most easily understood from Figure 2. Figure 2a introduces the biomass composition of 1 C-mol biomass (the ash-free organic fraction). The composition shown is fairly typical and is taken from Roels (1). One C-mol ash-free organic biomass is the amount of organic dry biomass that contains 12 g of carbon. The indicated biomass organic fraction corresponds to an elemental composition of 48.8% carbon, 7.3% hydrogen, 32.5% oxygen, and 11.4% nitrogen (w/w).

In practice, total dry biomass, which includes the organic fraction and the ash fraction (S, P, K, Mg, etc.), is measured. In general, the organic and ash fraction are obtained by combusting the organic biomass at 500 to 600 °C and weighing the ashes. Recently Battley (9) has indicated that this simple procedure underestimates the real organic biomass weight by 5 to 6%. This is due to the formation of P, S, and metal oxides in the ash during combustion, whereas such oxides are not present in the dry biomass. The composition formula follows directly from the elemental analysis of the biomass. In Figure 2a only the four major elements (C, H, O, N) are shown; however, it is straightforward to include P, S, and metals such as K or Mg in this composition formula, and also in the stoichiometric/energetic calculation. Figure 2a also shows that in the formation of biomass for all chemotrophic growth systems a C source, N source, H_2O, CO_2, and H^+ are always involved. These five compounds provide the building elements for making biomass. This is also called anabolism. For heterotrophic organisms the C source is organic; for autotrophic organisms the C source is CO_2.

Although it is possible to establish a stoichiometrically correct description to make biomass from these five building compounds, it is easily shown that this is not acceptable from the point of view of the second law of thermodynamics. It has been calculated that the Gibbs energy of such a hypothetical reaction, depending on the C source used, is often positive (7), although sometimes small negative values can also be calculated (5). In addition, it is



(a)

Relevant stoichiometry

$$-\frac{1}{Y_{DX}}\text{ (C-)mol electron donor} - (\ldots)\text{ N-source} - \frac{1}{Y_{AX}}\text{ mol electron acceptor} + 1\text{ C-mol biomass} + \frac{1}{Y_{QX}}\text{ (heat)} + \frac{1}{Y_{GX}}\text{ Gibbs energy} + (\ldots)\,H_2O + (\ldots)\,HCO_3^+ + (\ldots)\,H^+$$

Many different possibilities

Donor	Acceptor	C-source
Organic	Organic	Organic
Inorganic	Inorganic	Inorganic

$$Y_{DX}0.01 - 0.7\text{ C-mol}X\text{/(C)mol donor}$$

(b)

Figure 2. (**a**) System definition of microbial growth. (**b**) Macrochemical reaction equation of microbial growth.

known that to convert the five compounds into biomass, microorganisms use a large amount of biochemical energy in the form of ATP (10). Clearly the production of biomass from the five building compounds requires input of large quantities of Gibbs energy. The amount of energy needed to make biomass depends on the type of C source used. Intuitively, one expects that making 1 C-mol biomass from CO_2 requires more Gibbs energy than making 1 C-mol biomass from an organic compound. A quantitative relation for this energy need is presented later (equations 2, 3a, and 3b). The required energy, which must be taken as Gibbs energy and not as enthalpy, is delivered by a redox reaction between an electron donor and an electron acceptor. This redox reaction is called catabolism (Fig. 2a). Examples are the aerobic combustion of glucose ($C_6H_{12}O_6 + 6O_2 \rightarrow 6HCO_3^- + 6H^+$) and the anaerobic formation of ethanol from glucose ($C_6H_{12}O_6 + 2H_2O \rightarrow 2HCO_3^- + 2H^+ + 2C_2H_5OH$). Obtaining the required Gibbs energy is as essential for micro-organisms as it is for higher organisms, and even for human society. Therefore it should not be surprising that during evolution a wide diversity of microorganisms developed that are mainly different in the applied redox reaction for catabolism to obtain Gibbs energy (Fig. 2b). Electron donor or acceptor couples can be organic and inorganic compounds. This microbial variety in catabolic possibilities for generating Gibbs energy has led to the use of a classification system for naming microorganisms (Table 1). This system is understandably based on the source of Gibbs energy (light or chemical energy), the source of electron donor (inorganic or organic), and the source of biomass carbon (CO_2 or organic).

In addition, microorganisms may employ a wide variety of electron acceptors, as reflected in their class names. These class names are related to the electron acceptor used in catabolism (O_2, aerobic; NO_3^-, denitrification; SO_4^{2-}, sulphate reduction) fermentation (absence of external electron acceptor), or to the product of the catabolic reaction (CH_4, methanogenic; acetate, acetogenic; H_2S, sulphidogenic, etc.). The C source also functions often as electron donor, except in autotrophic microorganisms, where the C source is CO_2. For example, a microorganism growing aerobically in the dark on H_2S as the electron donor (inorganic compound) using CO_2 as the C source is called an aerobic chemolithoautotrophic organism. In summary, in each realistic chemotrophic microbial growth system there must be present the five compounds of anabolism and an electron donor/acceptor combination for catabolism.

These considerations bring us then to Figure 2b, which shows the macrochemical reaction equation containing all the stoichiometric information of the growth process. The macrochemical equation of Figure 2b should not be considered a mathematical equation but is a chemical reaction where substrates and products have negative and positive stoichiometric coefficients, respectively. Therefore for $a = 0$, sign is absent. In addition, the stoichiometric involvement of enthalpy and Gibbs energy is expressed in their respective stoichiometric coefficients (Y_{QX}, and Y_{GX}, which have units of C-mol X/kJ). The macrochemical reaction equation is therefore a compact, but exact, form of notation of the relevant stoichiometry of growth. This macrochemical equation shows that for the formation of $+1$ C-mol of biomass an amount of $-1/Y_{DX}$ of electron donor is required. The minus sign shows that the electron donor is consumed. Y_{DX} is in C-mol biomass per C-mol electron donor (in case of an organic donor) or per mol donor (in case of an inorganic donor). Its units are written as C-mol X/(C) mol D. An amount of $-1/Y_{AX}$ mol electron acceptor is consumed (minus sign) per 1 C-mol biomass produced and, in addition, $1/Y_{QX}$ kJ of heat and $1/Y_{GX}$ kJ of Gibbs energy are involved in the production of 1 C-mol of biomass. $1/Y_{QX}$ and $1/Y_{GX}$ are found as the conventionally calculated enthalpy of reaction and Gibbs energy of reaction of the macrochemical reaction equation, which produces 1 C-mol of biomass. Finally, certain amounts of H_2O, CO_2, (or HCO_3^-), H^+, and N source are involved. It is important that in each macrochemical reaction equation in which biomass is grown, HCO_3^-, H_2O, H^+, and N source be present. The differences between different organisms occur mostly in the electron acceptor/donor combinations used. The N source is often NH_4^+ and sometimes NO_3^-, N_2, or something else. The most important point in stoichiometry is to recognize that it is nearly always sufficient to measure one stoichio-

Table 1. Microbial Classification System

Source of Gibbs energy	Source of electron donor	C-source
Light (phototrophic)	Inorganic (lithotrophic)	CO_2 (autotrophic)
Chemical (chemotrophic)	Organic (organotrophic)	Organic (heterotrophic)

metric coefficient, that is, Y_{DX}, which is the traditional biomass yield on substrate (equal to carbon source and electron donor). All the other stoichiometric coefficients then follow from the so-called conservation equations (elements, electric charge, and enthalpy) (Example 1a) and the Gibbs energy balance (Example 1b).

EXAMPLE 1a

Calculation of stoichiometric coefficients in the macrochemical equation

Consider the aerobic growth of *Pseudomonas oxalaticus* on oxalate using NH_4^+ as the N source. The relevant chemical compounds in this growth system are the five compounds [biomass ($CH_{1.8}O_{0.5}N_{0.2}$), NH_4^+, HCO_3^-, H^+, H_2O], the electron donor oxalate ($C_2O_4^{2-}$), and the electron acceptor O_2. In total there are seven compounds and four elements (C, H, O, N). The conversion rates of these compounds are mathematically related by the conservation relations of C, H, O, N, and electric charge. In total there are five independent relations. This means that seven conversion rates are related by five conservation equations, and that the measurement of two rates (e.g., biomass production r_x and consumption of the electron donor oxalate r_D), which is equivalent to the measurement of $Y_{DX} = r_X/-r_D$, allows the calculation of all other yields.

Suppose that from measurement the biomass yield Y_{DX} is found to be +0.086 C-mol biomass produced per C-mol oxalate consumed. The proper macrochemical reaction equation can be written in a general form, without knowing all the stoichiometric coefficients but one (+1 for biomass), as

$$f C_2O_4^{2-} + a NH_4^+ + b H^+ + c O_2 + d H_2O + 1 CH_{1.8}O_{0.5}N_{0.2} + e HCO_3^-$$

The following conservation equations can now be written:

C conservation	$2f + 1 + e = 0$
H conservation	$4a + b + 2d + 1.8 + e = 0$
O conservation	$4f + 2c + d + 0.5 + 3e = 0$
N conservation	$a + 0.2 = 0$
Charge conservation	$-2f + a + b - e = 0$

Clearly there are six unknown stoichiometric coefficients (a–f) that are related by five conservation equations. (Biomass has been assigned a convenient, yet arbitrary coefficient +1.) Having one measured coefficient allows the calculation of all other coefficients. Y_{DX} was measured as 0.086. This means that 1/0.086 = 11.63 C-mol oxalate are consumed to produce 1 C-mol biomass. The previously defined macrochemical equation contains f mol of oxalate, which was two carbon atoms. The stoichiometric coefficient f therefore has the value $-11{,}63/2 = -5.815$ (remember the minus sign). Using this f value and the five conservation equations, one can calculate the whole chemical growth stoichiometry. The result is

$$-5.815 C_2O_4^{2-} - 0.2 NH_4^+ - 0.8 H^+ - 1.857 O_2 - 5.42 H_2O + 1 CH_{1.8}O_{0.5}N_{0.2} + 10.63 HCO_3^-$$

All the different biomass yields can be read from this reaction equation; thus, $Y_{AX} = 1/1.857 = 0.538$ C-mol biomass/mol O_2 or $Y_{CX} = 1/10.63 = 0.094$ C-mol biomass per mol CO_2.

In the Example 1a, only the chemical stoichiometry was calculated. However, there are two additional biomass yields of interest that relate the heat production and Gibbs energy dissipation occurring during the growth process to biomass production. These yields can be simply calculated if the full chemical stoichiometry is known by using tabulated ΔH_f^0 and ΔG_{f-}^{01} values (at pH = 7 and standard conditions) and calculating the enthalpy and Gibbs energy of reaction (Example 1b).

Table 2 contains all the required thermodynamic information as taken from Thauer et al. (11). The values for biomass are taken from Roels (1). Although there is some discussion about the value of ΔG_f^{01} for biomass, its value is not very important in thermodynamic calculations, as shown by Heijnen (3).

EXAMPLE 1b

Calculation of the yield of biomass on enthalpy and Gibbs energy (Y_{QX} and Y_{GX})

The chemical stoichiometry from Example 1a and the appropriate ΔH_f^0 and ΔG_f^{01} values from Table 2 can be used to obtain the heat (enthalpy) and Gibbs energy of reaction.

The enthalpy of reaction, using ΔH_f^0 from Table 2, is calculated as

$$(10.63)(-692) + 1(-91) - (5.42)(-286) - (1.857)(0) - (0.8)(0) - (0.2)(-133) - (5.815)(-824) = -1078.7 \text{ kJ}$$

For the Gibbs energy of reaction using ΔG_f^{01} values there follows a value of −1052.4 kJ. Because in the macrochemical reaction 1 C-mol of biomass is produced, this means that for each 1 C-mol biomass produced there is a heat production of 1078.7 kJ and a Gibbs energy dissipation of 1052.4 kJ, showing that $Y_{QX} = 1/1078.7 = 0.00093$ C-mol biomass produced per kilojoule heat produced and that $Y_{GX} = 1/1052.4 = 0.0095$ C-mol biomass produced per kilojoule of Gibbs energy dissipated. The complete chemical and energetic stoichiometry now can be written as

$$-5.815 C_2O_4^{2-} - 0.2 NH_4^+ - 0.8 H^+ - 1.857 O_2 - 5.42 H_2O + 1 CH_{1.8}O_{0.5}N_{0.2} + 10.63 HCO_3^- + 1078.7 \text{ kJ heat} + 1052.4 \text{ kJ Gibbs energy}$$

Example 1 shows that the complete chemical and energetic stoichiometry of microbial growth can be calculated from one measured yield using conservation equations and the Gibbs energy and enthalpy balance (elements, charge,

Table 2. Standard Gibbs Energy and Ethalpy of Formation

Compound name	Composition	ΔG_f^{01} (kJ/mol)	ΔH_f (kJ/mol)
Biomass	$CH_{1.8}O_{0.5}N_{0.2}$	−67	−91
Water	H_2O	−237.18	−286
Bicarbonate	HCO_3^-	−586.85	−692
CO_2 (g)	CO_2	−394.359	−394.1
Ammonium	NH_4^+	−79.37	−133
Proton	H^+	−39.87	0
O_2 (g)	O_2	0	0
Oxalate^{2-}	$C_2O_4^{2-}$	−674.04	−824
Carbon monoxide	CO	−137.15	−111
Formate	CHO_2^-	−335	−410
Glyoxylate$^-$	$C_2O_3H^-$	−468.6	—
Tartrate^{2-}	$C_4H_4O_6^{2-}$	−1,010	—
Malonate^{2-}	$C_3H_2O_4^{2-}$	−700	—
Fumarate^{2-}	$C_4H_2O_4^{2-}$	−604.21	−777
Malate^{2-}	$C_4H_4O_5^{2-}$	−845.08	−843
Citrate^{3-}	$C_6H_5O_7^{3-}$	−1,168.34	−1,515
Pyruvate$^-$	$C_3H_3O_3^-$	−474.63	−596
Succinate^{2-}	$C_4H_4O_4^{2-}$	−690.23	−909
Gluconate$^-$	$C_6H_{11}O_7^-$	−1,154	—
Formaldehyde	CH_2O	−130.54	—
Acetate	$C_2H_3O_2^-$	−369.41	−486
Dihydroxyacetone	$C_3H_6O_3$	−445.18	—
Lactate	$C_3H_5O_3^-$	−517.18	−687
Glucose	$C_6H_{12}O_6$	−917.22	−1,264
Mannitol	$C_6H_{14}O_6$	−942.61	—
Glycerol	$C_3H_8O_3$	−488.52	−676
Propionate$^-$	$C_3H_5O_{2+}$	−361.08	—
Ethylene glycol	$C_2H_6O_2$	−330.50	—
Acetoine	$C_4H_8O_2$	−280	—
Butyrate	$C_4H_7O_2^-$	−352.63	−535
Propanediol	$C_3H_8O_2$	−327	—
Butanediol	$C_4H_{10}O_2$	−322	—
Methanol	CH_4O	−175.39	−246
Ethanol	C_2H_5O	−181.75	−288
Propanol	C_3H_8O	−175.81	−331
n-Alkane	$C_{15}H_{32}$	+60	−439
Propane	C_3H_8	−24	−104
Ethane	C_2H_6	−32.89	−85
Methane	CH_4	−50.75	−75
H_2 (g)	H_2	0	0
N_2 (g)	N_2	0	0
Nitrite ion	NO_2^-	−37.2	−107
Nitrate ion	NO_3^-	−111.34	−173
Iron II	Fe^{2+}	−78.87	−87
Iron III	Fe^{3-}	−4.6	−4
Hydrogen sulfide (g)	H_2S	−33.56	−20
Sulfide ion	HS^-	+12.05	−17
Sulfate ion	SO_4^{2-}	−744.63	−909
Thiosulfate ion	$S_2O_3^{2-}$	−513.2	−608

Note: pH = 7, 1 atm, 1 mol/L, 298 K.

enthalpy, and the Gibbs energy balance). This means also that there must exist mathematical relations between Y_{DX}, Y_{AX}, Y_{CX}, Y_{QX}, and Y_{GX} (see Fig. 2b). These relations are addressed in a later section (see equations 9a–9e). It is obvious that this knowledge of the complete growth stoichiometry provides essential engineering information with respect to reactor design on the amount of O_2 that must be transferred (aeration capacity), the amount of carbon dioxide that must be removed (ventilation), the amount of heat to be removed (cooling capacity), or the amount of fermentation products (in anaerobic growth). The amounts of the required N source and HCO_3^- (autotrophic growth) also follow from these stoichiometric calculations.

MEASUREMENT OF GROWTH STOICHIOMETRY

As shown earlier, the measurement of one stoichiometric coefficient suffices, in general, to calculate all the other stoichiometric coefficients using the conservation relations. This measured stoichiometric coefficient requires the measurement of two conversion rates because, by definition, a stoichiometric coefficient is the ratio of two conversion rates. For example, $Y_{DX} = r_X/-r_D$. The most simple growth system contains eight conversion rates (biomass, N source, H^+, H_2O, CO_2, electron donor, electron acceptor, heat production) and six conservation equations (C, H, O, N, enthalpy, charge). Measurement of two conversion rates is then sufficient to calculate all other rates and, hence, the complete growth stoichiometry. Currently, the most common measurements are biomass production and substrate (equal to electron donor) consumption. For aerobic growth the on-line measurement of O_2 consumption and CO_2 production by the analysis of O_2 and CO_2 in the off gas in air-sparged fermentors is becoming more and more routine. Especially for autotrophic growth, the on-line measurement of CO_2 consumption by off-gas analysis gives direct and highly accurate information on microbial growth (because all consumed CO_2 appears as biomass). This method was very successfully applied to study the growth stoichiometry and kinetics of solid pyrite oxidation by Fe^{2+}-oxidizing bacteria (12,13) and of *Methanobacterium thermoautotrophicum* on H_2/CO_2 (14).

Most recently, it was also shown that on-line measurement of heat production during microbial growth can be used to explore growth stoichiometry and kinetics (15–17).

However, such a simple approach of measuring only two conversion rates often makes certain assumptions:

- Each chosen pair of measured conversion rates will allow the complete calculation of all other conversion rates.
- All measurements are reliable within a certain statistical error but without a systematic deviation.
- The assumed description of the growth system is correct, which means that by-products or additional substrates are assumed to be absent.

All these assumptions are subject to critical considerations, which are dealt with extensively in a recent series of publications (18–21). Here, simple examples are provided to illustrate the points of interest. The reader is referred to Refs. 18–21 for a more elaborate introduction, including the full mathematical and statistical aspects.

Noncalculability of Stoichiometry

Suppose that in Example 1a the chosen two conversion rates to be measured are biomass production (r_X) and

NH_4^+ consumption (r_N). Measurement of these two rates would not lead to a calculation of the other rates because r_X and r_N occur in the nitrogen-conservation equation in such a way that r_X uniquely determines r_N, and vice versa. It is then said that r_X and r_N are redundant. The N balance gives a constraint for these two measured conversion rates that can be used to calculate the statistically best estimate of r_X and r_N, which also exactly satisfies the N balance.

Clearly the choice of the two measured rates must be such that calculability of all other conversion rates is assured. In example 1a, suitable combinations would be the oxygen consumption rate (r_O) and biomass production rate (r_X), r_O and the carbondioxide production rate (r_C), or r_X and the heat production rate (r_Q).

Ill-Conditioned Calculability of Stoichiometry (Error Propagation)

It is well known that measured conversion rates have a certain measurement error. The subsequently calculated conversion rates, from combining the conservation relations and the two measurements, have an error due to error propagation. It is obviously of great practical importance to choose two measured conversion rates where this error propagation is minimal. A simple example to illustrate this problem is the aerobic growth of biomass on the donor glucose. If oxygen consumption ($-r_O$) and carbon dioxide production (r_C) are the measured rates, then the following relations (using conservation relations and the standard biomass composition) to calculate r_X and $-r_D$ (in C-mol glucose/m^3 h) from the measured r_C and ($-r_O$) can be derived:

$$r_X = 20r_C - 20(-r_O)$$
$$(-r_D) = 20(-r_O) - 21r_C$$

Due to the large multiplication factors of 20 and 21 in these equations, the propagation of the measurement errors in r_C and r_O into r_X and $-r_D$ is enormous.

If the donor conversion rate ($-r_D$) and the carbon dioxide production rate (r_C) were chosen as the measured rates r_X and ($-r_O$) would be calculated as

$$r_X = (-r_D) - r_C$$
$$(-r_O) = -0.05(-r_D) + 1.05r_C$$

The error propagation now is much lower and, therefore, from the measured ($-r_D$) and (r_C), r_X and ($-r_O$) can be calculated, as can be the other conversion rates involved. Clearly the aspect of error propagation is of major importance, and this propagation can be significantly decreased by a proper choice of the conversion rates to be measured.

Redundancy of Measurements

As stated earlier, in general two well-chosen measured conversion rates are usually sufficient to reliably calculate the complete stoichiometry. However, it is advantageous (Example 2) to measure more conversion rates than the minimum requirement of two. This leads to so-called redundant measurements, which can be used for two purposes error diagnosis and data reconciliation.

Error Diagnosis

- To check the validity of the defined growth systems with respect to the absence of by-products or possible second substrates
- To check the measured conversion rates for systematic errors

Data Reconciliation

- To decrease the measurement error in the calculated and measured conversion rates, provided that the statistically based checks (error diagnosis) on the validity of the growth system and the systematic errors in the measured conversion rates are passed

---

EXAMPLE 2

Use of redundant measurements to establish the presence of errors in the definition of the growth system or in the measurements

Consider the microbial growth system of Example 1a, where the following four conversion rates have been measured. The biomass has the standard elemental composition.

Biomass production	$r_X = +1$ C-mol/h
O_2 consumption	$-r_O = 1.2$ mol/h
HCO_3^- production	$r_C = 10.5$ mol/h
Oxalic acid ($C_2O_4^{2-}$ consumption)	$-r_D = 5.8$ mol/h

We know that a minimum of two rate measurements are needed to calculate the full stoichiometry. Therefore there are two redundant measurements. We can now establish the conservation equations (with r_W, r_H, and r_N as the water, proton, and NH_4^+ conversion rates, respectively) based on conversion rates as

C conservation	$2r_D + r_X + r_C = 0$
H conservation	$4r_N + r_H + 2r_W + 1.8r_X + r_C = 0$
O conservation	$4r_D + 2r_O + r_W + 0.5r_X + 3r_C = 0$
N conservation	$r_N + 0.2r_X = 0$
Charge conservation	$-2r_D + r_N + r_H - r_C = 0$

By eliminating the three nonmeasured rates (r_W, r_H, r_N) from these five conservation equations, one obtains $5 - 3 = 2$ equations, which relate the measured conversion rates only. The result is as follows:

$$2r_D + r_X + r_C = 0$$
$$2r_D - 4r_O + 4.2r_X = 0$$

The first relation can be recognized as the carbon balance, and the second is so-called electron balance or the balance of degree of reduction (1) (see also a later section). With respect to the C balance, one finds from the measurements:

$$\text{C-in} = 2 \times 5.8(\text{oxalate}) = 11.6 \text{ C-mol/h}$$
$$\text{C-out} = 1(\text{biomass}) + 10.5(CO_2) = 11.5 \text{ C-mol/h}$$

Clearly the C balance seems satisfying (0.86% gap).

For the balance of degree of reduction one obtains

$$\text{electrons in} = 2 \times 5.8 = 11.6 \text{ mol electrons/h}$$

$$\text{electrons out} = -1.2(-4) + 1(4.2) = 9 \text{ mol electrons/h}$$

Clearly there is a large gap of 2.6 mol electrons/h.

Because the C balance fits, it is reasonable to assume that the measured values of r_D, r_X, and r_C are reliable. The balance of degree of reduction can therefore be wrong for two reasons:

1. A very inaccurate measurement of r_O.
2. If the measurement of r_O is found to be correct, then the only other possibility is the presence of an additional electron acceptor (e.g., NO_3^-). This would be an error in the defined growth system.

A Mathematically Complete Analysis of Calculability, Analysis of Redundancy, Error Diagnoses, and Data Reconciliation

In the preceding section simple examples were provided to highlight the problems in accurately establishing the full growth stoichiometry from measurements. Because all these calculations are based on linear conservation relations, it is highly appropriate to use matrix algebra. Basic to these calculations is the "elemental" matrix, which specifies the element, charge, and enthalpy information for each compound in the growth system. Recently, an extensive and coherent mathematical description has been provided for calculability, redundancy analysis, error diagnosis, statistical aspects, and data reconciliation using involved matrix algebra (18–21). The developed mathematical theory has been put in a user-friendly computer program called Macrobal (22).

THE EFFECT OF GROWTH RATE ON GROWTH STOICHIOMETRY

Maintenance Energy Concept

In his pioneering work, Monod (23) found that in exponential growth the amount of biomass formed increased in proportion to the amount of substrate consumed. This led to the definition of growth yield Y_{DX} [amount of biomass produced per amount of electron donor (substrate) consumed]. We have seen that Y_{DX} usually determines the complete growth stoichiometry. With the introduction of the chemostat in the early 1950s, microbial growth could be studied at a range of growth rates, and it became clear that Y_{DX} decreased at lower growth rates μ, as shown in Figure 3 (24). This phenomenon was explained by two different concepts (24–26):

- Endogenous respiration or microbial decay, determined by the parameter k_d
- Electron donor (substrate) requirements for maintenance, determined by the parameter m_D

The basic idea is, however, similar in recognizing that a microorganism is a complex structure where the polymers (proteins, etc.) are subject to slow thermal denaturation and where there are numerous small leaks associated with the many transmembrane gradients (e.g., Na^+ leaking into the microorganism). These leaking substances must be pumped out, and the degraded polymers must be rebuilt at the expense of Gibbs energy. This results in a small, but finite, need of Gibbs energy to maintain the biomass structure and the transmembrane gradients (maintenance Gibbs energy). In the concept of endogeneous respiration or microbial decay, this energy is produced by catabolism of biomass itself. In the concept of maintenance this energy is produced by catabolism of a part of the substrate (electron donor).

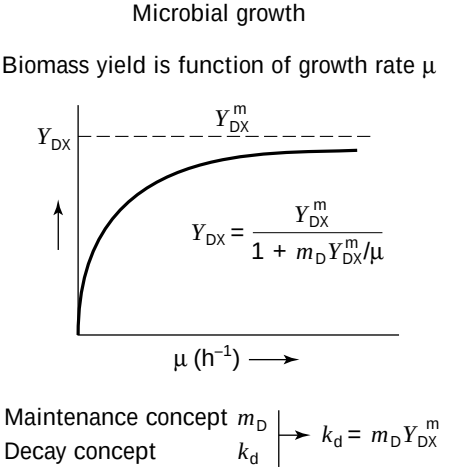


Figure 3. Dependence of Y_{DX} on growth rate μ.

Mathematically, the dependence of Y_{DX} on growth rate μ is described by equation 1a and shown in Figure 3.

$$1/Y_{DX} = 1/Y^m_{DX} + m_D/\mu \qquad (1a)$$

This equation contains two model parameters, Y^m_{DX} and m_D, where m_D is the rate of consumption of electron donor (substrate) that is catabolized to generate the necessary Gibbs energy flow for maintenance in C-mol electron donor per C-mol biomass per hour. Y^m_{DX} is the maximal biomass yield. Figure 3 shows that Y_{DX}, using equation 1a, decreases with decreasing growth rate. Clearly, at higher growth rates Y_{DX} comes close to Y^m_{DX}. Using typical values for m_D it can be shown that only for $\mu < 0.01$ to 0.05 h^{-1}, Y_{DX} starts dropping significantly below Y^m_{DX}. This means that in exponential growth, as occurs in batch fermentation where μ is high, the stoichiometry is properly covered by Y^m_{DX}. However, in many industrial-fed batch-production processes, maintenance is extremely important due to the low growth rates applied. For example, in penicillin fermentation $\mu \approx 0.01$ h^{-1} and about 70% of all consumed glucose is spent for maintenance (27). Similarly, in wastewater-treatment processes, where low growth rates are also applied, the maintenance effects are very relevant. However, in this area one often uses the biomass decay coefficient k_d. This coefficient is however related to m_D according to $k_d = m_D Y^m_{DX}$. In general, it can be shown that all biomass yields, Y_{iX} as defined in Figure 2a, decrease with decreasing growth rate μ.

Measuring m_D

Equation 1a shows that Y^m_{DX} and m_D can be obtained directly by measuring Y_{DX} at different growth rates μ. Using equation 1a to plot $1/Y_{DX}$ versus $1/\mu$ as a straight line to obtain $1/Y^m_{DX}$ and m_D is, however, not desirable, because the error distribution of the measurements Y_{DX} and μ is completely distorted due to the use of $1/Y_{DX}$ and $1/\mu$.

It is more proper to directly use the nonlinear equation 1a in combination with the measured Y_{DX} and μ, and to use an algorithm for nonlinear parameter estimation. It is stressed that for accurate m_D values one should measure Y_{DX} at low growth rates (0.005–0.03 h^{-1}).

EXAMPLE 3a

Calculating Y^m_{DX} and m_D from measured Y_{DX} as function of growth rate μ

Consider aerobic growth on glucose and that there are two measurements available. At $\mu = 0.5\ h^{-1}$, $Y_{DX} = 0.49$ C-mol biomass per C-mol glucose, and at $\mu = 0.02\ h^{-1}$, $Y_{DX} = 0.33$.

Applying equation 1a will show that $Y^m_{DX} = 0.50$ and $m_D = 0.02$ C-mol glucose/C-mol biomass/h.

Other Maintenance Quantities

Microorganisms require Gibbs energy for maintenance. This is obtained by catabolizing the required amount of electron donor m_D. It is then obvious that other quantities, such as electron acceptor, heat, Gibbs energy, oxidized electron donor, and reduced electron acceptor, are also involved in maintenance, to catabolize the m_D electron donor. These maintenance-related quantities are directly obtained from the stoichiometry of the catabolic reaction (Example 3b).

EXAMPLE 3b

Calculating other maintenance rates using the catabolic reaction

In Example 3a it was found that $m_D = 0.02$ C-mol glucose/C-mol biomass/h.

In the growth system being considered, the catabolic reaction is the aerobic oxidation of glucose according to

$$C_6H_{12}O_6 - 6O_2 + 6HCO_3^- + 6H^+ + 2{,}814\ \text{kJ heat} + 2{,}843\ \text{kJ Gibbs energy}$$

Using the stoichiometry of this catabolic reaction and the known m_D, it is now easy to calculate the other maintenance rates:

$$\text{maintenance glucose } m_D = 0.02 \text{ C-mol glucose/C-mol Xh} = 0.02/6 \text{ mol glucose/C-mol biomass/h}$$

$$\text{maintenance oxygen } m_A = \frac{0.02}{6} \times 6 \text{ mol } O_2\text{/C-mol biomass/h}$$

$$\text{maintenance } HCO_3^- \ m_C = \frac{0.02}{6} \times 6 \text{ mol } CO_2\text{/C-mol biomass/h}$$

$$\text{maintenance heat } m_q = \frac{0.02}{6} \times 2{,}814 \text{ kJ/C-mol biomass/h}$$

$$\text{maintenance Gibbs energy } m_G = \frac{0.02}{6} \times 2{,}843 \text{ kJ/C-mol biomass/h}$$

Complete Growth Stoichiometry as a Function of Growth Rate

Equation 1a shows how $1/Y_{DX}$ depends on the growth rate μ and the two parameters Y^m_{DX} and m_D. Completely similar equations can be derived for growth yields on acceptor (A), carbon dioxide (C), heat (Q), and Gibbs energy (G) according to equation 1b:

$$1/Y_{iX} = 1/Y^m_{iX} + \frac{m_i}{\mu} \qquad (1b)$$

Here i can be A, C, Q, or G. The maintenance coefficients for the different compounds are related to m_D according to Example 3b. The maximal yields Y^m_{AX}, Y^m_{CX}, Y^m_{QX} and Y^m_{GX} are related to Y^m_{DX}, and can be found by solving the macrochemical equation, using the available Y^m_{DX} value, according to Example 1.

A THERMODYNAMICALLY BASED METHOD TO ESTIMATE GROWTH STOICHIOMETRY

In the previous paragraphs the methods for accurate measurement of a growth stoichiometric coefficient, as, for example, the biomass yield on electron donor Y_{DX} and the subsequent calculation of all the nonmeasured stoichiometric coefficients of the macrochemical equation (using the conservation principles) have been provided. In past decades, the value of Y_{DX} for many different microorganisms, different electron donors, C sources, and electron acceptors has been measured under C- and energy-limited growth conditions. Many methods have been proposed to predict Y_{DX} because of its obvious importance. Recently, a critical evaluation of these methods has been performed (2). The following criteria were used for the evaluation:

- The method should be generally applicable to all chemotrophic growth systems.
- The method should relate directly to the second law of thermodynamics.
- No detailed knowledge of metabolism is required; only the identity of the electron donor, C source, and electron acceptor is known.
- Methodological problems are absent.

The conclusion of this evaluation was that none of the published methods satisfied these simple criteria. Therefore, an alternative method that satisfies the mentioned criteria has been proposed (2). This method is based on $1/Y_{GX}$, which is the amount of Gibbs energy (in kilojoules) that must be dissipated for the production of 1 C-mol biomass.

The Gibbs energy stoichiometric parameter $1/Y_{GX}$ has already been introduced as one of the stoichiometric coefficients in the macrochemical reaction equation (Fig. 2b). Therefore, it is obvious that this energetic parameter can be calculated directly if only one of the chemical stoichio-

metric coefficients has been measured and if the electron donor, electron acceptor, and C source are known (see Example 1b, where $1/Y_{GX} = 1{,}052$ kJ/C-mol biomass).

Furthermore it is well known that the value of growth yields depends on the growth rate (μ) due to the Gibbs energy that must be used for maintenance (see earlier section). This means that the Gibbs energy needed to produce biomass should be divided into two parts:

1. A growth-related part
2. A maintenance-related part

Mathematically this can be expressed as

$$\underset{\text{Total needed Gibbs energy kJ/C-mol biomass}}{1/Y_{GX}} = \underset{\text{Gibbs energy for new biomass}}{\frac{1}{Y^{m}_{GX}}} + \underset{\text{Maintenance Gibbs energy for existing biomass}}{\frac{m_G}{\mu}} \quad (1c)$$

where $1/Y^{m}_{GX}$ is the Gibbs energy needed to make 1 C-mol of biomass (kJ/C-mol X) and m_G is the Gibbs energy needed for biomass maintenance (kJ/C-mol biomass h). The biomass specific growth rate (h^{-1}) is μ.

Clearly, at high growth rate μ, the m_G/μ term becomes negligible and $1/Y_{GX}$ becomes practically equal to Y^{m}_{DX}. At low growth rates Y_{GX} becomes much lower than Y^{m}_{GX}.

Equation 1c shows that in order to calculate Y_{GX} as a function of growth rate μ we need information about Y^{m}_{GX} and m_G. In the past years two simple correlations have been found with which to estimate Y^{m}_{GX} and m_G (2,4). These correlations were established using a very large body of experimental growth yields, which covered carbon- and energy-limited growth for the following:

- Many different microorganisms (bacteria, fungi, plant cells)
- Many different C sources, including CO_2 and a wide variety of organic substrates
- Different electron acceptors (aerobic, anaerobic, denitrifying)
- Electron donors that need reversed electron transport (RET)

The resulting correlations are given in equations 2 and 3. Figure 4 shows the m_G data used to establish equation 2; Figure 5 shows the $1/Y^{m}_{GX}$ data used to establish the correlations 3a and 3b.

Maintenance Gibbs Energy Need m_G

The data for m_G as a function of temperature shown in Figure 4 can be correlated with an Arrhenius type of relation:

$$m_G = 4.5 \exp\left[\frac{-69{,}000}{R}\left(\frac{1}{T} - \frac{1}{298}\right)\right] \quad (2)$$

This correlation was found to hold (with ± 40% accuracy) for a very wide variety of organisms, for different electron donors, for aerobic and anaerobic conditions, and for a temperature range of 5–75 °C (4). Obviously, the main influencing factor is the temperature, which behaves as an Arrhenius function with an activation energy of 69,000 J/mol. The type of electron donor (organic or inorganic), the microorganism, and the electron acceptor are of minor importance. This seems logical, because maintenance is a biomass-linked Gibbs-energy-requiring process that counteracts the biomass-deteriorating processes (protein degradation, leakage over cell membranes, etc.).

Gibbs Energy for Growth

The data for Y^{m}_{GX} shown in Figure 5a (heterotrophic growth) and Figure 5b (autotrophic growth) can be correlated by equations 3a and 3b as shown in Ref. 2. For autotrophic growth it was found to be important to distinguish electron donors for which reversed electron transport (RET) was necessary. Such electron donors (e.g., Fe^{2+}/Fe^{3+}, NH_4^+/NO_2^-) provide electrons that have insufficient Gibbs energy to reduce the C source CO_2 to biomass. Microorganisms using such electron donors first have to increase the Gibbs energy level of the donor electrons by the biochemical process RET.

Heterotrophic growth/autotrophic growth (−RET)

$$1/Y^{m}_{GX} = 200 + 18(6 - C)^{1.8} + \exp[((3.8 - \gamma)^2)^{0.16}(3.6 + 0.4C)] \quad (3a)$$

Autotrophic growth (+RET)

$$1/Y^{m}_{GX} = 3{,}500 \quad (3b)$$

It was found (4) that the Gibbs energy dissipation required for the production of 1 C-mol biomass mainly depends on the C source used (equations 3a and 3b) The type of microorganism and the type of electron acceptor have only minor effects, as shown in Figures 5a and 5b. The influence of the C source on Y^{m}_{GX} can be characterized by the following:

- Its number C of carbon atoms (e.g., for CO_2 C = 1 and for glucose C = 6) as shown in Figure 5a
- Its degree of reduction γ (1)

γ is a stoichiometric number of a chemical compound that represents the number of electrons in the compound. For organic compounds γ is per C-mol, for inorganic compounds γ is per mol. For example, for CO_2 $\gamma = 0$, for CH_4 $\gamma = 8$, and for glucose $\gamma = 4$. The concept of degree of reduction will be further elucidated extensively later in this article. For organic compounds (Fig. 5a), γ has a value between 0 and 8. For inorganic compounds (Fig. 5b), only a lower value of 0 holds; a maximal value does not exist because there is no normalization per atom. It is relevant to know that biomass has a degree of reduction of about 4.2. Equation 3a and Figure 5a show that, in the situation that RET is not required for both hetero- and autotrophic growth, the Gibbs energy needed to produce biomass

- Increases if the number of C atoms in the carbon source (the parameter C in equation 3a) decreases

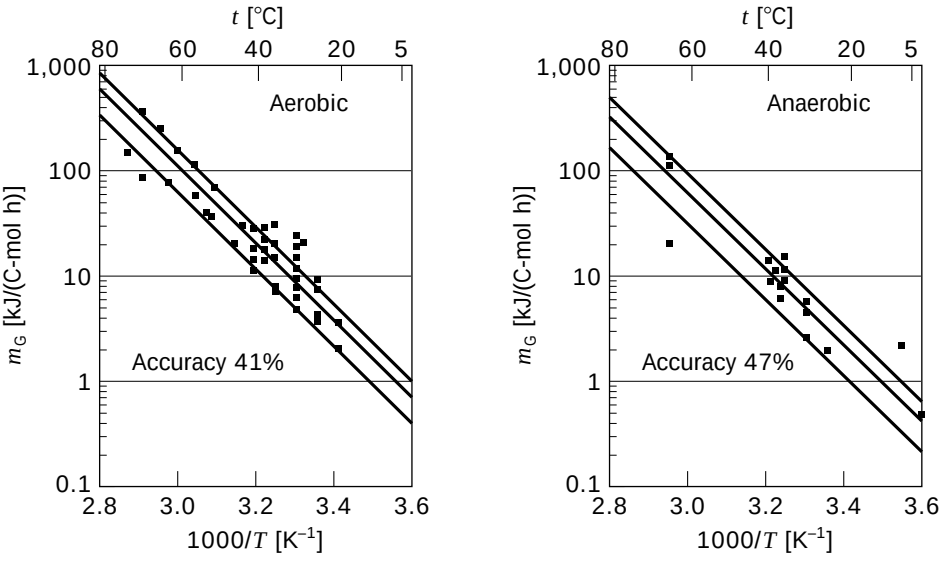


Figure 4. Maintenance Gibbs energy m_G (in kJ/C-mol biomass h) for aerobic (**a**) and anaerobic (**b**) growth, shown as an Arrhenius function of temperature. The lines represent equation 2.

- Increases if the degree of reduction of the carbon source (the parameter γ in equation 3a) is smaller or larger than about 3.8

Equation 3a further shows that for heterotrophic growth $1/Y^m_{GX}$ ranges between about 200 and 1,000 kJ/C-mol biomass, for the C sources explored, for which:

- The number of carbon atoms in the carbon source ranges between C = 1 (e.g., CO_2, formate, methane) and C = 6 (e.g., glucose, citrate).
- The degree of reduction of the C source γ ranges between 0 (for CO_2) and 8 (for CH_4).

The effect of the number of C atoms (C) and degree of reduction (γ) of the C source can be simply understood as follows:

- Biomass contains many polymers that contain monomers of four to six C atoms. If the C source contains fewer than four to six C atoms, the microorganism must perform extra biochemical reactions to achieve C–C couplings. This requires extra Gibbs energy, compared to a C source that has six C-atoms, Hence $1/Y^m_{GX}$ increases for C-sources with less carbon atoms.
- Biomass has $\gamma = 4.2$. If the C source is more reduced ($\gamma > 4.2$) or more oxidized ($\gamma < 4.2$), there is a need for additional oxidation reactions or reduction reactions, respectively, as compared to a carbon source (like glucose) with $\gamma = 4$. These additional reactions lead to extra Gibbs energy dissipation, leading to a higher value of $1/Y^m_{GX}$.

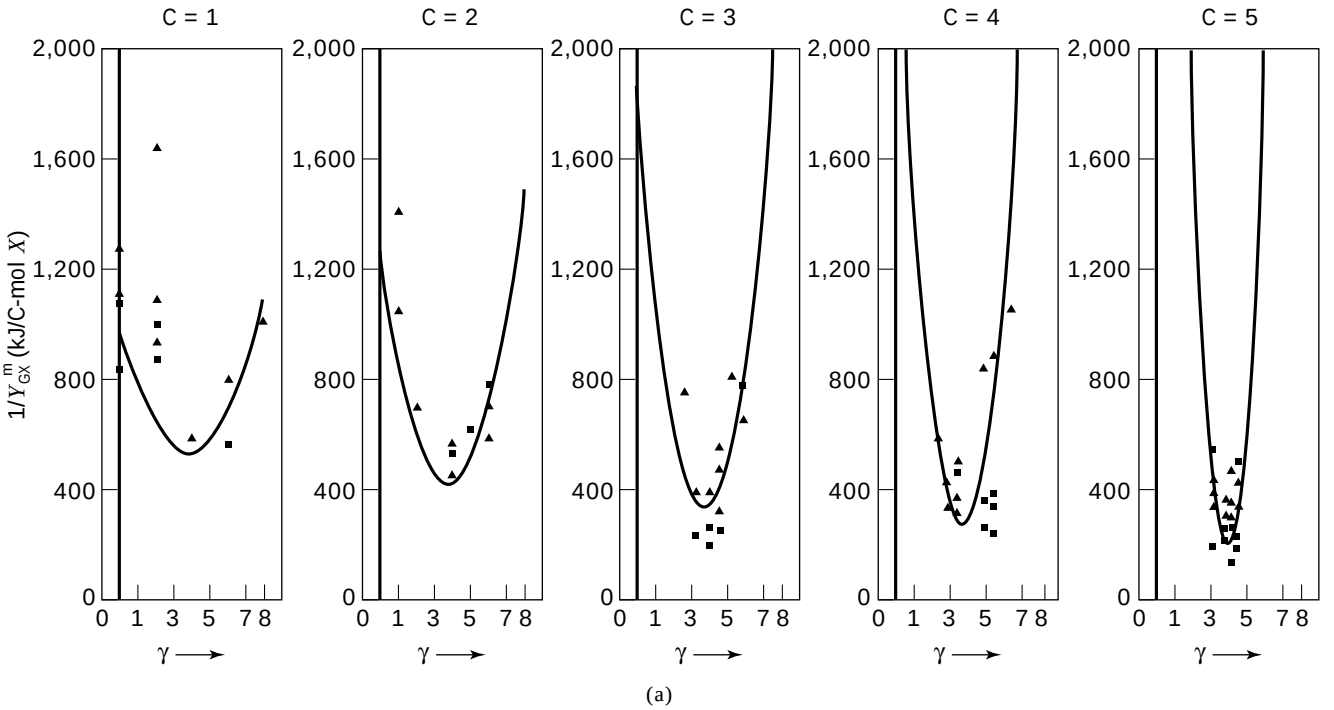
Simply stated, the more biochemical tinkering is needed to convert an organic C source into biomass, the more Gibbs energy is dissipated and the higher $1/Y^m_{GX}$ becomes. Obviously glucose (C = 6, $\gamma = 4$) is a nearly ideal C-source because it requires the least Gibbs energy dissipation for biomass production. According to equation 3a, for glucose $1/Y^m_{GX} = 200 + 0 + 36 = 236$ kJ/C-mol biomass. In contrast, CO_2 is a very poor C source, because it requires about four times as much Gibbs energy ($1/Y^m_{GX} = 200 + 236 + 460 = 986$ kJ/C-mol according to equation 3a). Equation (3b) shows that for autotrophic growth, in the situation where RET is needed (which occurs for many inorganic electron donors), $1/Y^m_{GX}$ has a very high value of 3,500 kJ/C-mol biomass. This value should be compared to autotrophic growth without RET as occurs with, for example, H_2 or CO as electron donor (for which $1/Y^m_{GX} \approx 1{,}000$ kJ/C-mol according to equation 3a).

Obviously, the use of RET increases the Gibbs energy dissipation needed for biomass production tremendously. The explanation is that, using the RET process, the electrons of the electron donor are increased in energy level, up to the energy level of electrons in NADH in order to make CO_2 reduction to biomass thermodynamically feasible. This "energy-pumping" process (RET) apparently requires a large amount of Gibbs energy, of about 3,500 − 1,000 = 2,500 kJ/C-mol biomass produced.

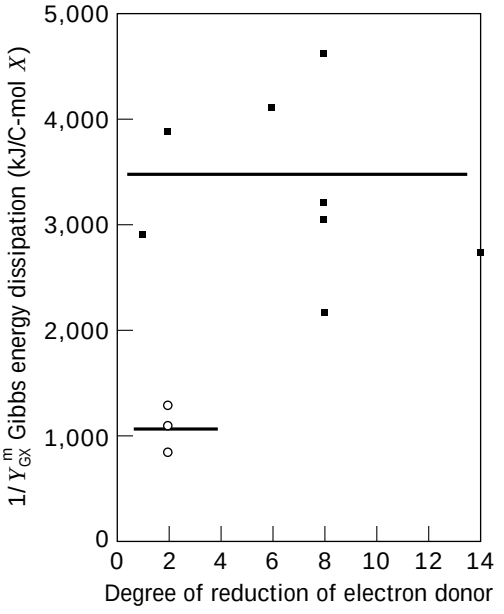
The effect that the type of the available C source has on the Gibbs energy needed for biomass synthesis is well known in biochemistry. Biochemists express the energy need in ATP. Figure 6 compares the calculated Gibbs energy dissipation needed for biomass synthesis ($1/Y^m_{GX}$, in kJ/C-mol*X*) with the theoretically calculated amount of ATP expenditure for biomass synthesis in mol ATP/C-mol *X*. The points shown are for different C sources, ranging from glucose (28.8) to CO_2 (2.5). The parenthetical numbers are the published (10) biomass yields on ATP in gram-*X*/mol ATP. It is clear that there is a close correspondence, which is logical. Equations 3a and 3b provide the energy needed for biomass synthesis in kilojoules, whereas the biochemists use mol ATP as the energy measure.

In conclusion, it should be realized that equations 2, 3a and 3b are completely sufficient to estimate a biomass yield and the full macrochemical equation for any arbitrary chemotrophic growth system (Example 4).

The predictive accuracy of this correlation for chemotrophic growth has been shown (2,4) to be ±10 to 20% rela-



(a)



(b)

Figure 5. Reciprocal maximal yield of biomass on Gibbs energy, $1/Y^{m}_{GX}$, (kJ/C-mol X); (**a**) Heterotrophic growth (triangles, aerobic; squares, fermentation, X's denitrifying systems); C is the number of carbon atoms in the carbon source; γ is the degree of reduction of the carbon source. (**b**) Autotrophic growth (squares, electron donors where reversed electron transport [RET] is needed; circles, donors without RET). The lines represent equations 3a and 3b.

tive error in a yield range of nearly two orders of magnitudes of 0.01–0.70 C-mol biomass per C-mol organic electron donor or per mol inorganic donor while covering aerobic, anaerobic, denitrifying, autotrophic microbial systems with and without RET (Fig. 7). The measured yield data used were taken from Refs. 2 and 4.

EXAMPLE 4

Calculation of the full macrochemical reaction equation using the correlations of equations 2, 3a, and 3b

It is assumed that a microorganism grows anaerobically on methanol as C source and electron donor with NH_4^+ as the N source

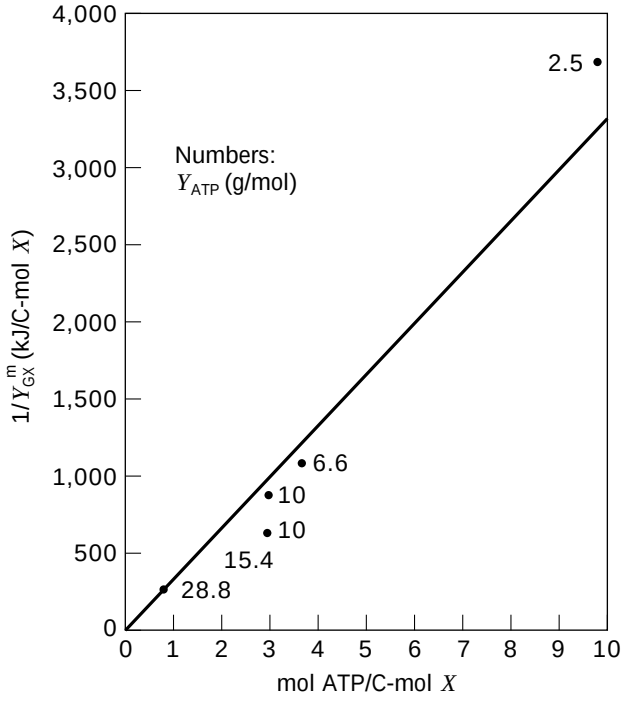


Figure 6. Comparison of energy needed for biomass synthesis on different carbon sources in mol ATP/C-mol biomass and in kJ/C-mol biomass ($1/Y^{m}_{GX}$). The numbers refer to the conventional biomass yield on ATP in gram biomass/mol ATP for different carbon sources (10).

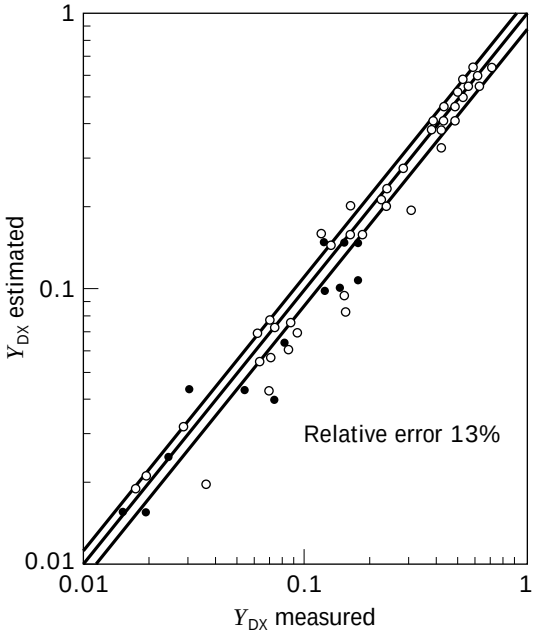


Figure 7. Comparison of measured and predicted biomass yield Y_{DX} (solid circles, fermentative; open circles, aerobic growth systems).

and acetate is produced. The growth system contains biomass, NH_4^+, H^+, HCO_3^-, H_2O, methanol, and acetate as the seven compounds. The general macrochemical reaction equation for the production of 1 C-mol biomass can be written as follows:

$$f\,CH_3OH + a\,NH_4^+ + b\,H^+ + c\,C_2H_3O_2^- + d\,H_2O + 1\,CH_{1.8}O_{0.5}N_{0.2} + e\,HCO_3^-$$

Clearly there are six unknown stoichiometric coefficients (a–f). However, using equation 3a we can calculate that (C = 1, γ = 6 for methanol, maintenance has been neglected) $1/Y^{m}_{GX}$ = 200 + 326 + 172 = 698 kJ/C-molX. This means that we know that the Gibbs energy of reaction of the macrochemical reaction equation equals −698 kJ.

We can now write the conservation equations for C, H, O, N, electric charge, and the Gibbs energy balance (taking values of ΔG_f^{01} from Table 2):

C balance	$f + 2c + 1 + e = 0$
H balance	$4f + 4a + b + 3c + 2d + 1.8 + e = 0$
O balance	$f + 2c + d + 0.5 + 3e = 0$
N balance	$a + 0.2 = 0$
charge balance	$a + b - c - e = 0$

$$\text{Gibbs energy balance } (-175.39)f + (-79.37)a + (-39.87)b + (-369.41)c + (-237.18)d + (-67)1 + (-586.85)e + 698 = 0$$

Solving these six equations gives, for a to f,

$$a = -0{,}2;\ b = 2.866;\ c = 8.898;\ d = 12.964;\ e = -6.232;\ f = -12.564$$

This gives a biomass yield on methanol of 1/12.564 = 0.08 C-mol biomass/C-mol methanol. The stoichiometric result also shows that the acetate production is 8.898/12.564 = 0.70 mol acetate/mol methanol, showing a C yield of 1.4 acetate carbon/methanol-carbon. This is, of course, due to the CO_2 fixation that occurs (6.232 mol HCO_3^- per 12.564 mol methanol).

If maintenance is not allowed to be neglected m_G must be taken into account. For example, the temperature is assumed to be 50 °C. Equation 2 then shows that m_G = 38.8 kJ/C-mol biomass h. If the growth rate μ = 0.03 h^{-1}, then we can calculate, using equation 1c that $1/Y_{GX}$ = 698 + 38.8/(0.03) = 1,991 kJ/C-mol biomass. Using this number one can solve the six equations to obtain the complete stoichiometry, which holds under these conditions.

Before ending this section, a final warning is relevant. The described thermodynamic method of predicting growth stoichiometry is based on a very wide database of experimentally measured growth systems. No detailed biochemical information is required, because intrinsically a kind of average biochemistry used by most organisms is assumed. This is an attractive feature, but in the end we should consider that, of course, the biochemistry used by microorganisms does have a significant influence. For example, for the anaerobic ethanol fermentation on glucose the mentioned method will give Y_{DX} = 0.15 C-mol biomass/C-mol glucose. This is indeed found for *Saccharomyces cerevisae*. However, another organism, *Zymomonas mobilis*, does the same glucose/ethanol process, but with Y_{DX} = 0.07. The explanation is that *Z. mobilis* uses a completely different biochemical pathway for glucose catabolism than *S. cerevisae*. From this example we can also

learn that if the predicted biomass yield differs very substantially from the actually measured yield, it might be possible that the microorganism being studied uses a novel pathway for catabolism or anabolism.

A USEFUL REFERENCE SYSTEM TO SIMPLIFY GROWTH STOICHIOMETRIC AND ENERGETIC CALCULATIONS AND TO GAIN INSIGHT

The Growth Reference System

In the preceding sections, the stoichiometric coefficients for the macrochemical reaction equation of biomass formation have been solved by setting up the proper conservation equations (C, H, O, N, charge, enthalpy) and the Gibbs energy balance. Although this is a sufficient and straightforward method, solving these linear equations remains unattractive and does not provide insight. To simplify these calculations and to gain insight, a special reference system has been designed—the growth reference system. This reference system is based on the observation that, in all chemotrophic growth systems, H_2O, HCO_3^-, H^+, and N source (mostly NH_4^+) occur as chemical compounds (see earlier section on growth system definition). In this special reference system each chemical compound is assigned three new numbers.

- γ the degree of reduction, which represents the electron content per C-mol (for organic compounds) or per mol (for inorganic compounds)
- ΔG_e the Gibbs energy per electron present in the compound
- ΔH_e the enthalpy per electron present in the compound

Clearly γ is a stoichiometric quantity and ΔG_e and ΔH_e are energetic parameters.

The reference system is designed such that for H_2O, HCO_3^-, H^+ (pH = 7), N source for growth, HPO_4^{2-}, NO_3^-, SO_4^{2-}, and Fe^{3+}, the values of γ, ΔG_e, and ΔH_e are zero. For ΔG_e, the biochemical standard conditions (1 mol/L, 1 bar, pH = 7,298 K) are assumed, ΔH_e is calculated for CO_2 (gas) because of the large heat effect of HCO_3^- (liq) $\Leftrightarrow$ CO_2 (gas) transfer.

The calculation of γ, ΔG_e, and ΔH_e follows from the reference redox half reaction where 1 C-mol of organic or 1 mol of inorganic compound is converted into the reference chemicals and a number of electrons. The number of electrons is by definition equal to γ (Example 5). From the Gibbs energy and enthalpy of this reference reaction, called ΔG_{ref} and ΔH_{ref} (calculated with the usual thermodynamic ΔG_f^{01} and ΔH_f^{01} values, see Table 2), the values of ΔG_e and ΔH_e follow from equations 4a and 4b.

$$\Delta G_e = \frac{-\Delta G_{ref}}{\gamma} \tag{4a}$$

$$\Delta H_e = \frac{-\Delta H_{ref}}{\gamma} \tag{4b}$$

EXAMPLE 5

The reference redox half reaction and calculation of γ and ΔG_e for chemical compounds

For methanol the following reference redox half reaction can be set up according to the preceding definition by converting methanol to the reference compounds HCO_3^-, H_2O, and H^+

$$-1CH_4O - 2H_2O + HCO_3^- + 7H^+ + 6e^-$$

In this reference redox half reaction, 1 C-mol methanol is converted and six electrons are produced, hence $\gamma = +6$ for methanol. Using the ΔG_{f-}^{01} values from Table 2, the ΔG_{ref} for the methanol-reference redox half reaction follows as (standard conditions)

$$\Delta G_{ref}^{01} = (7)(-39.87) + 1(-586.85) - (2)(-237.18) - (1)(-181.75) = -216.192 \text{ kJ}$$

This gives for the ΔG_e^{01} value of methanol by equation 4a

$$\Delta G_e^{01} = -\left(\frac{-216.192}{6}\right) = +36.032 \text{ kJ/e-mol}$$

Obviously ΔH_e can be calculated in a similar way by calculation of ΔH_{ref}.

For biomass the following redox half reaction can be set up, assuming that NH_4^+ is the N source:

$$-1CH_{1.8}O_{0.8}N_{0.2} - 2.5H_2O + HCO_3^- + 0.2NH_4^+ + 5H^+ + 4.2e^-$$

Obviously, the degree of reduction for biomass is 4.2. The ΔG_{ref}^{01} value is obtained similarly as earlier for methanol. ΔG_{ref}^{01} can be calculated to be -142.128 kJ, giving

$$\Delta G_e = -(-142.128)/(4.2) = +33.840 \text{ kJ/e-mol}$$

In a similar way as shown in Example 5 for each chemical compound, the values of γ, ΔG_e, and ΔH_e can be calculated for a large number of relevant compounds. Table 3 contains all relevant stoichiometric and energetic information for growth systems, clearly shown in the following. A point of attention is the finding (Table 3) that for biomass the degree of reduction depends on the N source used in the growth system. For example $\gamma = 4.2$ for NH_4^+ and 5.8 for NO_3^- as N source. This is a consequence of the reference definition. The advantage is that the N source disappears from the stoichiometric calculations using γ, ΔG_e, and ΔH_e (Examples 7a, 7b, and 8b). The defined reference system is closely related to the generalized degree of reduction as defined by Roels (1) and Erickson et al. (28). It can be seen that for reduced organic compounds γ is between 0 and 8 (per C-mol). For inorganic compounds, such an upper limit does not exist (because there is not a normalization per atom). For O_2, γ is negative (-4), which is logical for an acceptor. ΔG_e is related to the conventional redox potential of redox half reactions ($\Delta G_e^{01} = -FE_0^1$). ΔG_e is calculated using HCO_3^- (the most abundant form of carbon dioxide at pH = 7); ΔH_e has been calculated using CO_2 (gas) as reference, to take the large heat effect of $HCO_3^- \rightarrow CO_2$ (gas) into account.

Table 3. Calculated γ, ΔG_e^{01}, and ΔH_e^0 Values for Chemical Compounds under Standard Conditions

	γ Degree of reduction per C-mole for organic and per mole for inorganic compounds in electrons/(C)-mole	ΔG_e^{01} (kJ/e-mol)	ΔH_e^0 (kJ/e-mol)
biomass/NH_4^+ − N source	+4.2	+33.840	−26.1
Biomass/NO_3 − N source	+5.8	+14.820	−44.2
Biomass/N_2 − N source	+4.8	+32.948	−26.3
N source for growth	0	0	0
HCO_3^-	0	0	0
Oxalate	+1	+52.522	−20
Formate	+2	+39.186	−15.50
Glyoxylate	+2	+48.229	—
Tartrate	+2.5	+39.577	—
Malonate	+2.67	+28.976	—
Fumarate	+3	+33.662	−31.60
Malate	+3	+33.354	−32.20
Citrate	+3	+32.282	−33.90
Pyruvate	+3.33	+34.129	−23.60
Succinate	+3.50	+28.405	−36.30
Gluconate	+3.67	+39.106	—
Formaldehyde	+4	+45.326	−0.10
Acetate	+4	+26.801	−33.50
Lactate	+4	+31.488	−28.90
Glucose	+4	+39.744	−25.75
Mannitol	+4.33	+38.777	—
Glycerol	+4.67	+37.625	−24.30
Propionate	+4.67	+26.939	−33.80
Ethylene glycol	+5	+37.292	—
Acetoin	+5	+32.625	—
Butyrate	+5	+27.000	−33.30
Propanediol	+5.33	+33.177	—
Acetone	+5.33	+28.718	−30.90
Butanediol	+5.50	+31.374	—
Methanol	+6	+36.032	−23
Ethanol	+6	+30.353	−28.90
Propanol	+6	+29.144	−32.50
n-Alkane	+6.13	+26.694	—
Propane	+6.66	+25.948	−31.90
Ethane	+7	+25.404	−31.40
Methane	+8	+22.925	−31.50
CO	+2	+47.477	−1.5
H_2	+2	+39.870	0
SO_4^{2-}	0	0	0
SO_3^{2-}	+2	+50.296	—
S^0	+6	+19.146	−55.2
$S_2O_3^{2-}$	+8	+23.584	−27.5
HS^-	+8	+20.850	−43.9
NO_3^-	0	0	0
NO_2^-	+2	−41.650	−108.5
NO(g)	+3	−96.701	—
N_2O(g)	+8	−57.540	−124.55
NH_4^+	+8	−35.109	−101.9
N_2	+10	−72.194	−136.4
Fe^{3+}	0	0	0
Fe^{2+}	+1	−74.270	−46.8
H_2O	0	0	0
O_2	−4	−78.719	−143

Note: pH = 7, 1 mol/L, 1 atm, 298 K.

Balance of Degree of Reduction, Atomic Degrees of Reduction, and the COD Balance

In the previously defined "growth reference" system, γ_i was introduced as the degree of reduction of compound *i*. This parameter is important in stoichiometric calculations, because due to the principle of electron conservation, an electron balance can be defined. This is the so-called balance of degree of reduction. This is not an additional conservation principle (in addition to C, H, O, N, and charge conservation). The balance of degree of reduction can be obtained from the usual C, H, O, N, and charge balances by eliminating (by suitable substitutions) H^+, H_2O, HCO_3^-, and the N source. Hence, the balance of degree of reduction is a suitable linear combination of already available conservation equations. The importance of the balance of degree of reduction is that, by definition, in this balance, only biomass formation, consumption of electron donor, and consumption of electron acceptor are related. Based on the previous definition of the reference set of compounds in the growth reference system it is also possible to calculate the degree of reduction of atoms and of electric charge (Table 4).

It should be noted that the atomic degree of reduction for the N atom in biomass depends on the applied N source as a consequence of the defined growth reference system as explained earlier. Using the γ values of atoms and electric charge in Table 4, it is straightforward to calculate the γ values for any chemical compound for which the elemental composition is known (Example 6a). This is an equivalent alternative to writing the reference redox half reaction to obtain γ (Example 5).

EXAMPLE 6a

Direct calculation of γ from elemental composition

Using the atomic degrees of reduction (Table 4) it can easily be checked that indeed for the reference chemicals $\gamma = 0$:

H_2O	$\gamma = 2 \times 1 + 1(-2) = 0$
CO_2	$\gamma = 1 \times 4 + 2(-2) = 0$
HCO_3^-	$\gamma = 1 \times 1 + 1 \times 4 + 3(-2) + 1 = 0$
H^+	$\gamma = 1 \times 1 + 1(-1) = 0$

Table 4. Degree of Reduction of Atoms and Electric Charge According to the Definition of the Growth Reference System

Atom or charge	Degree of reduction of atoms
H	+1
O	−2
C	+4
Charge +1	−1
Charge −1	+1
S	+6
P	+5
N	+5
N in N index in biomass	+3 for NH_4^+ or NH_3 as N source 0 for N_2 as N source +5 for NO_3^- or HNO_3 as N source

For the degree of reduction of biomass (γ_X) it is easy to show that this is a function of the N source used. Using the standard elemental biomass composition $CH_{1.8}O_{0.5}N_{0.2}$ and using the N degree of reduction for the different N sources (Table 4) one obtains

$$-NH_4^+ \text{ as N-source } \gamma_X = 1 \times 4 + 1.8 \times 1 + 0.5(-2) + 0.2(-3) = 4.2$$

$$-NO_3^- \text{ as N-source } \gamma_X = 1 \times 4 + 1.8 \times 1 + 0.5 \times (-2) + 0.2(5) = 5.8$$

For the electron content of an organic substrate (e.g., acetate ion, $C_2H_3O_2^-$) the amount of electrons is $2 \times 4 + 3 \times 1 + 2(-2) + 1(+1) = 8$. Because for organic compounds γ is defined as the number of electrons per C atom, we obtain for acetate with 2 carbon atoms $\gamma = 8/2 = 4$.

The degree of reduction balance is also called chemical oxygen demand (COD) balance in wastewater engineering (26). The COD balance is equivalent to the balance of degree of reduction. COD is a number assigned to each chemical and represents the consumed O_2 on total oxidation in g O_2/g compound. There is a direct link with degree of reduction. Each mole of electrons represents 8 g COD. This is easily understood, because the consumption of 1 mol O_2 represents the acceptance of 4 mol electrons ($\gamma = -4$; see Table 3). One mol O_2 represents −32 g COD and therefore 1 mol electrons ≡ 8 g COD.

EXAMPLE 6b

Calculation of COD values

Consider glucose, in which 1 mol (≡ 180 gram) represents (according to Table 3) a total of 6×4 electrons ≡ $6 \times 4 \times 8 = 192$ g O_2. Clearly glucose has a COD value of 192/180 = 1.0667 g COD/g glucose.

The NO_3^-/N_2 acceptor couple has $\gamma_A = -5$ electrons. The COD value of NO_3^- − N is then $-5 \times 8/14 = -2.857$ g COD per gram nitrate-nitrogen.

The values of γ, ΔG_e, and ΔH_e from Table 3 can be used for very easy stoichiometric and energetic calculations as shown in Examples 7a and 7b.

EXAMPLE 7a

Calculation of the stoichiometry of example 1a using γ values

In Example 1a, five equations were solved to calculate the full macrochemical equation. Using the γ values of Table 3 we can first make the balance of degree of reduction. For the electron donor oxalate $\gamma_D = 1$ per carbon or 2 per mole oxalate; for biomass $\gamma_X = 4.2$ and for the electron acceptor $O_2 \gamma_A = -4$ (Table 3). For all the other chemicals (N source, H^+, H_2O, HCO_3^-) $\gamma = 0$ by definition. The γ balance is now

$$2f - 4c + 4.2 = 0$$

Because $f = -5.815$ we obtain $c = -1.857$ directly.

From the C balance we then obtain $e = +10.63$. From the N balance $a = -0.20$, from the charge balance $b = -0.8$, and from

the O or H balance we finally find $d = -5.42$. This is, as expected, the same result as before in Example 1a.

EXAMPLE 7b

Calculation of the Gibbs energy of reaction in example 1b using ΔG_e values

Using the now-available full macrochemical stoichiometry, it is possible to calculate the Gibbs energy of reaction using the Gibbs energy balance.

For each chemical compound the Gibbs energy contribution follows from the product of its number of electrons and its ΔG_e number. For example (using Table 3), the Gibbs energy contribution for oxalate, O_2, and biomass in the growth reference system follows as

$$\text{oxalate} = 2 \times 1 \times 52.522 = +105.04 \text{ kJ}$$

$$O_2 = -4 \times (-78.719) = 314.876 \text{ kJ}$$

$$\text{biomass} = 4.2 \times 33.840 = 142.128 \text{ kJ}$$

For all other reactants the Gibbs energy contribution in the growth reference system is zero. For the Gibbs energy of the macrochemical-reaction equation we obtain then from the available full stoichiometry:

$$-5.815(105.04) - 1.857(314.876) + 142.128 = -1053 \text{ kJ}$$

This is the same as obtained before, but now the calculation has only three terms.

Energetics of Redox Couples, Catabolic Redox Reactions, RET, and Energetic Regularities

It was pointed out earlier that for each microbial growth system a catabolic redox reaction is needed where an electron donor couple reacts with an electron acceptor couple. For the generation of maintenance energy the catabolic reaction is also required. For example, in aerobic growth on glucose the electron donor couple is glucose/HCO_3^- (glucose is oxidized to HCO_3^-) and the electron acceptor couple is O_2/H_2O (O_2 is reduced to H_2O). However, for anaerobic growth on glucose, where the catabolic reaction is the conversion of glucose into ethanol, the electron donor couple is glucose/HCO_3^- and the electron acceptor is the HCO_3^-/ethanol couple. To be able to quickly calculate the catabolic energy production, it is relevant to define the ΔG_e, ΔH_e, and γ values of redox couples i/j. These can be calculated from γ, ΔG_e, and ΔH_e values (Table 3) using equations 5a–5c.

$$\gamma_{couple} = \gamma_i - \gamma_j \tag{5a}$$

$$(\Delta G_e)_{couple} = \frac{(\gamma \Delta G_e)_i - (\gamma \Delta G_e)_j}{\gamma_i - \gamma_j} \tag{5b}$$

$$(\Delta H_e)_{couple} = \frac{(\gamma \Delta H_e)_i - (\gamma \Delta H_e)_j}{\gamma_i - \gamma_j} \tag{5c}$$

Equations 5a–5c have the following properties:

- If we invert the redox couple, for example, i/j into j/i, we are in fact inverting the redox half reaction of the redox couple. The value of γ changes sign, which is logical because produced electrons become consumed electrons. The value of ΔG_e and ΔH_e does not change, which is logical because the energetics of the reaction do not change.
- If the redox couple i/j contains a biological reference compound (e.g., $j = HCO_3^-$, H_2O, H^+, NO_3^-, SO_4^{2-}) then for compound j the value of γ, ΔG_e, and ΔH_e is zero. Equations 5a–5c then show that the value of γ, ΔG_e, and ΔH_e of the redox couple i/j becomes equal to the tabulated values of the i component in Table 3. For example, if the redox couple is an organic compound/HCO_3^-, then the γ, ΔG_e-, and ΔH_e value of the organic compound follow directly from Table 3.
- If the redox couple does not contain a reference chemical then equations 5a–5c must be used to calculate γ, ΔG_e, and ΔH_e. Table 5 shows some examples.

It can be seen from Table 5 that electron donor couples are characterized by positive γ values and electron-acceptor couples by negative γ values (which is logical). As stated earlier, in each microbial growth system there functions a catabolic reaction between an electron donor and an electron acceptor, which generates the required (for anabolism and maintenance) Gibbs energy. It is noted (equation 6c) that in a full catabolic redox reaction, the redox couple with the highest ΔG_e value must be the electron donor; the redox couple with the lowest ΔG_e value is the acceptor (Example 8a).

EXAMPLE 8a

Recognizing electron donor and acceptor

Consider the following catabolic reaction: $C_6H_{12}O_6 + 6O_2 \rightarrow 6HCO_3^- + 6H^+$. The two redox couples are $C_6H_{12}O_6/HCO_3^-$ and O_2/H_2O. According to Table 3, the glucose couple has $\Delta G_e = +39.744$ kJ/e-mol and the O_2 couple has $\Delta G_e = -78.719$ kJ/e-mol. Clearly glucose is the electron donor and O_2 is the acceptor. Consider now the catabolic reaction $C_6H_{12}O_6 + 2H_2O \rightarrow 2HCO_3^- + 2H^+ + 2C_2H_5OH$. This is the catabolic reaction in the ethanol fermentation. The redox couples are $C_6H_{12}O_6/HCO_3^-$ and

Table 5. ΔG_e, and ΔH_e Values of Redox Couples under Standard Conditions

Redox couple	γ (per (C-)mol)	ΔG_e (kJ/e-mol)	ΔH_e (kJ/e-mol)
$NH_4^+/½N_2$	3	+26.703	−44.3
NH_4^+/NO_2^-	6	−32.928	−99.665
NO_2^-/NO_3^-	2	−41.647	−108.5
Lactate/pyruvate	2	+18.283	−55.4
Fumarate/succinate	−2	−3.137	−64.5
$NO_3^-/½N_2$	−5	−72.194	−136.4
$NO_2^-/½N_2$	−3	−92.559	−155
$NO(g)/½N_2$	−2	−35.434	—
$N_2O(g)/N_2$	−2	−130.809	−183.8
Glucose/HCO_3^-	4	+39.744	−25.75

HCO_3^-/C_2H_5OH. From Table 3 we read for the glucose couple that $\Delta G_e = +39.744$ and for the ethanol couple $\Delta G_e = +30.353$. Now glucose is the donor and HCO_3^-/ethanol is the acceptor.

Because the catabolic reaction liberates the Gibbs energy required for the anabolism, it is important to calculate this amount of energy. Using the ΔG_e approach, we can then write directly equations 6a and 6b to calculate the Gibbs energy of reaction (ΔG_{CAT}) of the catabolic reaction. ΔG_{ED} and ΔG_{EA} are the electron Gibbs energy of the acceptor and donor couples:

$$-\Delta G_{CAT} = \gamma_D(\Delta G_{ED} - \Delta G_{EA}) \quad (6a)$$

The enthalpy of reaction ΔH_{CAT} of the catabolic reaction can be calculated similarly:

$$-\Delta H_{CAT} = \gamma_D(\Delta H_{ED} - \Delta H_{EA}) \quad (6b)$$

ΔG_{CAT} and ΔH_{CAT} represent the Gibbs energy and enthalpy of reaction of the catabolic reaction consuming 1 C-mol organic or 1 mol inorganic compound of electron donor. Dimensions are kJ/(C)-mol donor. γ_D is the degree of reduction of the donor couple in mol electrons/(C)-mol donor, which is always positive. According to the second law of thermodynamics ΔG_{CAT} must be negative. Therefore equation 6c, the second law of thermodynamics, holds:

$$\Delta G_{ED} > \Delta G_{EA} \quad (6c)$$

This shows that indeed the electron donor couple always has the highest ΔG_e value.

EXAMPLE 8b

Calculation of catabolic Gibbs energy production using ΔG_e

Consider the catabolic reactions in Example 8a. For aerobic glucose oxidation we can calculate, using equation 6a and Table 3

$$(-\Delta G_{CAT}) = 4 \times [39.644 - (-78.719)] = 473.852 \text{ kJ, so that } \Delta G_{CAT} = -473.852 \text{ kJ}$$

This is the Gibbs energy released for the aerobic combustion of 1 C-mol glucose. For 1 mol of glucose (6 C-atoms) the Gibbs energy of the catabolic reaction is $6 \times (-473.852) = -2{,}843$ kJ.

For anaerobic ethanol fermentation of glucose we can calculate for the catabolic reaction $(-\Delta G_{CAT}) = 4 \times (39.744 - 30.353) = 37.564$ kJ per C-mol glucose. For 1 mol glucose the catabolic Gibbs energy of the catabolic reaction becomes $6 \times (-37.564) = -225.34$ kJ. This is the same as calculated in Example 3b.

The use of ΔG_e and ΔH_e now also reveals some interesting energetic regularities (4). Table 3 shows that for many organic donor compounds, the ΔH_e and ΔG_e values are rather close, with an average $\Delta G_e = +32 \pm 8$ kJ/e-mol and average $\Delta H_E = -28 \pm 5$ kJ/e-mol. These values are also close to the ΔG_E and ΔH_E value of biomass. Hence we can write for organic electron donors the important regularity

$$\Delta G_{ED} \approx \Delta G_{EX} \quad (7a)$$

$$\Delta H_{ED} \approx \Delta H_{EX} \quad (7b)$$

For autotrophic growth using inorganic electron donors, the ΔG_{ED} values are, in general, much lower than ΔG_{EX} (consider NH_4^+/NO_2^-, F_e^{2+}/F_e^{3+}, etc., in Table 4). For autotrophic growth (CO_2 as C source), this means that for these electron donors there is a need for Gibbs energy input in order to realize CO_2 reduction to biomass. This is achieved by RET. Knowing this we can write equation 7c to recognize RET.

$$\Delta G_{EX} > \Delta G_{ED} \quad (7c)$$

Finally, it is now easy to calculate the heat production and Gibbs energy dissipation in oxidative catabolism of organic compounds. As stated earlier, for organic compounds the average $\Delta G_{ED} = 32$ kJ/e-mol and for O_2 as acceptor $\Delta G_{EA} = -78.719$ kJ/e-mol. Hence, per mole of electron transferred between donor and acceptor, the available Gibbs energy is $32 - (-78.719) = 110.72$ kJ. Per mole of consumed O_2 (which accepts four electrons, $\gamma_A = -4$) in the combustion of any organic compound, the Gibbs energy made available by combustion of the organic compound is then $4 \times 110.72 = 443$ kJ per consumed mol O_2. Analogously, one can find for the produced heat per mole of O_2 in the combustion of organic compounds a value of 460 kJ per mole O_2. It is also obvious that the mentioned inaccuracy in the average ΔG_E or ΔH_E values for organic compounds only results in a minor error of 5 to 8% in the calculated Gibbs energy dissipation and heat production. These are very important rules of thumb for the fermentation industry (1).

MATHEMATICAL EQUATIONS TO CALCULATE THE GROWTH STOICHIOMETRY FROM KNOWN GIBBS ENERGY DISSIPATION

Deriving the Equations

In Example 1, it was shown that one suitably measured stoichiometric coefficient, for example, Y_{DX}, allows the calculation of all other stoichiometric coefficients, including the dissipated Gibbs energy $1/Y_{GX}$. This means that knowledge of $1/Y_{GX}$ should enable the calculation of all stoichiometric coefficients, as shown in Example 4. It has also been shown that Y_{GX} can be estimated for arbitrary growth systems under different growth rates and temperatures using the correlations (equations 2 and 3). Here we show that particularly simple equations to calculate all yields from Y_{GX} can be obtained by using the γ, ΔG_e, and ΔH_e parameters introduced in the previous section.

The general macrochemical reaction equation can be written as shown in Figure 2b. In this macrochemical equation, the electron donor and acceptor are written in C-mol (for organic compounds) or in mol (for inorganic compounds). The (. . .) stoichiometric coefficients are not given separate symbols. They follow easily from the charge balance (for H^+), N balance (for N source), from O or H balance (for H_2O), and the carbon balance (for HCO_3^-). For

autotrophic growth, the HCO_3^- stoichiometric coefficient is -1. Three balances can be written based on the conservation principles of electrons and enthalpy and the Gibbs energy balance.

Balance of degree of reduction

$$-\gamma_D/Y_{DX} - \gamma_A/Y_{AX} + \gamma_X = 0 \tag{8a}$$

Enthalpy balance

$$-\gamma_D \Delta H_{ED}/Y_{DX} - \gamma_A \Delta H_{EA}/Y_{AX} + \gamma_X \Delta H_{EX} + 1/Y_{QX} = 0 \tag{8b}$$

Gibbs energy balance

$$-\gamma_D \Delta G_{ED}/Y_{DX} - \gamma_A \Delta G_{EA}/Y_{AX} + \gamma_X \Delta G_{EX} + 1/Y_{GX} = 0 \tag{8c}$$

From equation 8a it follows that Y_{AX} is directly related to Y_{DX} (equation 9a)

$$\frac{1}{Y_{AX}} = \frac{\gamma_D}{(-\gamma_A)Y_{DX}} - \frac{\gamma_X}{(-\gamma_A)} \tag{9a}$$

The degree of reduction of the acceptor γ_A is (by definition) a negative number. Using equations 8a–8c, it is also possible to calculate directly Y_{DX}, Y_{AX}, and Y_{QX} as function of Y_{GX}. If we further use the found energetic regularity that $\Delta G_{EX} \approx \Delta G_{ED}$ (equation 7a) and replace $(\Delta G_{ED} - \Delta G_{EA})$ by $(-\Delta G_{CAT})/\gamma_D$ using equation 6a, it is possible to derive the simple equations 9b–9e. It is noted that ΔG_{CAT} is the Gibbs energy of the catabolic reaction of 1 C-mol organic or 1 mol inorganic compound, and that therefore $(-\Delta G_{CAT})$ is the Gibbs energy released in the catabolic reaction of 1 C-mol of organic or 1 mol of inorganic electron donor. $(-\Delta G_{CAT})$ is then by definition >0 and its units are kJ/(C)-mol donor. ΔG_{CAT} and ΔH_{CAT} follow from equations 6a and 6b using ΔG_{EA} and ΔG_{ED} values (Tables 3 and 5).

$$Y_{DX} = \frac{(-\Delta G_{CAT})}{1/Y_{GX} + \gamma_X/\gamma_D(-\Delta G_{CAT})} \tag{9b}$$

$$Y_{AX} = \frac{(-\gamma_A/\gamma_D)(-\Delta G_{CAT})}{1/Y_{GX}} \tag{9c}$$

$$1/Y_{QX} = \frac{(-\Delta H_{CAT})}{(-\Delta G_{CAT})}\, 1/Y_{GX} \tag{9d}$$

Furthermore it is often interesting to study Y_{DA}, which is the amount of electron acceptor couple consumed relative to the amount of electron donor consumed. For microorganisms growing aerobically on organic matter, this would be the mole of O_2 consumed per C-mole of organic compound consumed. For anaerobic growth this would be the amount of anaerobic products per amount of organic substrate in C-mole product per C-mole substrate. Because $Y_{DA} = Y_{DX}/Y_{AX}$ we obtain

$$Y_{DA} = [\gamma_D/(-\gamma_A)]\, \frac{1/Y_{GX}}{1/Y_{GX} + \gamma_X/\gamma_D(-\Delta G_{CAT})} \tag{9e}$$

Application of the Mathematical Stoichiometry Relations

The obtained stoichiometric relations (equations 9a–9e) can now easily be applied. In this section their use is demonstrated with the following subjects:

- Calculation of the complete growth stoichiometry
- Calculation of maintenance coefficients and maximal growth yields
- Calculation of the limit to growth yield posed by the second law
- Calculation of COD-based growth yields
- Calculation of the relation between heat production and Gibbs energy dissipation
- Calculation of maximal product yields in anaerobic metabolism

Calculation of the Complete Growth Stoichiometry. If for a given growth system the C source, the electron donor (ΔG_{ED} to decide on RET using equation 7c), the temperature, and the growth rate are known, then equations 2 and 3 allow a direct calculation of the required Gibbs energy $1/Y^m_{GX}$ to produce 1 C-mol of biomass at high growth rates μ. Knowing the electron donor couple and acceptor couple, and using Table 3 and equations 5a–5c, the values of ΔG_{ED}, ΔG_{EA}, ΔH_{ED}, ΔH_{EA}, γ_D, and γ_A can be calculated and from this $(-\Delta G_{CAT})$ and $(-\Delta H_{CAT})$ using equations 7a and b. Using equations 9b–9e subsequently allows the complete stoichiometric calculation where H^+, N source, H_2O, and HCO_3^- must be calculated using the conservation equations of electric charge, N, O or H, and carbon.

EXAMPLE 9a

Calculation of stoichiometry using equations 9b–9e

Consider Example 4, where a microorganism is grown anaerobically on methanol, producing acetate. Assume first that the growth rate is high, such that maintenance can be neglected. Equation 3a then shows that $1/Y^m_{GX} = 698$ kJ Gibbs energy per C-mol biomass. Methanol is the C-source, methanol/HCO_3^- is the electron donor, and HCO_3^-/acetate is the electron acceptor. From Table 3 and using equations 5a–5c, we can then find that $\Delta G_{ED} = 36.032$ kJ/e-mol, $\Delta H_{ED} = -23$ kJ/e-mol; $\Delta G_{EA} = 26.801$ kJ/e-mol; $\Delta H_{EA} = -33.5$ kJ/e-mol. Also, $\gamma_D = 6$, $\gamma_A = -4$, and $\gamma_X = 4.2$.

This provides that $-\Delta G_{CAT} = 6(36.032 - 26.801) = 55.386$ and $-\Delta H_{CAT} = -23 - (-33.5) = 63$ kJ/C-mol methanol. Using equations 9b–9e, we obtain the maximal growth yields (maintenance neglected).

$$1/Y^m_{DX} = 13.3 \text{ mol methanol/C-mol biomass}$$

$$1/Y^m_{AX} = 18.9 \text{ C-mol acetate/C-mol biomass} = 9.45 \text{ mol acetate/C-mol biomass}$$

$$1/Y^m_{QX} = 794 \text{ kJ heat/C-mol biomass}$$

$$Y^m_{DA} = 1.42 \text{ C-mol acetate/C-mol methanol} = 0.71 \text{ mol acetate/mol methanol}$$

Using the C- balance $1/Y^m_{CX}$ is calculated as 6.6 mol HCO_3^- consumed/C-mol biomass produced. This overall stoichiometric result is very close to the exact solution obtained in Example 4.

The small deviation arises from the assumption that $\Delta G_{EX} \approx \Delta G_{ED}$ (as discussed before).

In general it can be shown that the simple set of equations 9b–9e seldom deviates more than 5% from the exact solution.

Maintenance Coefficients and Maximal Yield Coefficients. As indicated previously, relations between the maintenance coefficients follow from the catabolic reaction. The following relations can now be written to link the various maintenance coefficients to the maintenance Gibbs energy m_G using $-\Delta G_{CAT}$ (see also Example 9b). It is noted that m_G follows from the correlation (equation 2) and that ΔG_{CAT} is calculated for the catabolism of 1 (C)-mol of electron donor.

$$m_D = m_G/(-\Delta G_{CAT}) \tag{10a}$$

$$m_A = (\gamma_D/-\gamma_A)m_G/(-\Delta G_{CAT}) \tag{10b}$$

$$m_Q = m_G \frac{(-\Delta H_{CAT})}{(-\Delta G_{CAT})} \tag{10c}$$

$$m_C = m_G/(-\Delta G_{CAT}) \tag{10d}$$

Using equations 10a–10d, the value of the maintenance Gibbs energy requirement m_G can be calculated from either measured maintenance coefficients (for electron donor m_D, electron acceptor m_A, heat production m_Q, or carbon dioxide production m_C). Furthermore, it can easily be understood that the maximal biomass yields for electron donor, acceptor, and heat are found from equations 9b–9e by substitution of $1/Y^m_{GX}$ (instead of $1/Y_{GX}$) because the maintenance contribution is then neglected.

EXAMPLE 9b

Effect of maintenance on stoichiometry

In Example 9a the maintenance contribution was neglected. Assume that the microorganism is growing at 37 °C. Equation 2 then leads to $m_G = 13$ kJ/C-mol biomass h. Using equations 10a–10c and using ΔG_{CAT} and ΔH_{CAT} (Example 9) we obtain

$$m_D = 0.2347 \text{ mol methanol/C-mol biomass/h}$$

$$m_A = 0.3521 \text{ C-mol acetate/C-mol biomass/h}$$

$$m_Q = 14.787 \text{ kJ/C-mol biomass/h}$$

Further assume that the growth rate $\mu = 0.02\ h^{-1}$. Using equation 1b, the Y^m_{iX} values obtained in Example 9a, and the m_i values obtained here, one obtains for the stoichiometry:

$$\text{electron donor}(D)\,1/Y_{DX} = 13.3 + \frac{0.2347}{0.02} = 25.03 \text{ mol methanol/C-mol X}$$

$$\text{electron donor}(A)\,1/Y_{AX} = 18.9 + \frac{0.3521}{0.02} = 36.50 \text{ C-mol acetate/C-mol X}$$

$$\text{heat } 1/Y_{QX} = 794 + \frac{14.787}{0.02} = 1533 \text{ kJ heat/C-mol X}$$

$$Y_{DA} = 36.50/25.03 = 1.46 \text{ C-mol acetate/mol methanol}$$

These values can also be obtained directly from equations 9b–9e by substituting the complete Gibbs energy of growth and maintenance according to equation 1c:

$$1/Y_{GX} = 1/Y^m_{GX} + \frac{m_G}{\mu} = 698 + \frac{13}{0.02} = 1348 \text{ kJ Gibbs energy/C-mol X}$$

Clearly, comparing Examples 9a and 9b, one observes that the yield of biomass Y_{DX} drops from 0.077 to 0.04 due to maintenance, but the acetate/methanol yield Y_{DA} increases from 1.42 to 1.46 C-mol acetate/mol methanol.

Second Law Limit of Growth Yield. As for any chemical reaction, the microbial growth yield is also limited by the second law of thermodynamics. This limit is achieved if $1/Y_{GX} = 0$, because this defines equilibrium. Equations 9c–9f then show that for the thermodynamic limits we can write [see also Ref. 4] the following:

Thermodynamic limits for growth yields

$$Y_{DX} = \gamma_D/\gamma_X$$

$$1/Y_{AX} = 0$$

$$1/Y_{HX} = 0$$

Clearly, the more reduced electron donors (γ_D higher) have a higher Y_{DX} limit. This limit has already been determined (1).

COD-Based Yields. In wastewater treatment, the biomass yield is calculated on COD basis. Y_{COD} is the gram biomass COD over gram-substrate COD. Based on the COD definition we can write

$$Y_{COD} = \frac{\gamma_X}{\gamma_D} Y_{DX} \tag{11a}$$

This allows the following relation for Y_{COD} from equation 9b

$$Y_{COD} = \frac{(-\Delta G_{CAT})}{(-\Delta G_{CAT}) + (\gamma_D/\gamma_X)1/Y_{GX}} \tag{11b}$$

For aerobic growth on organic substrate $-\Delta G_{CAT} = \gamma_D[32-(-78.719)] \approx \gamma_D \times 110$ kJ/C-mol. Here the average value of $\Delta G_{ED} = +32$ kJ/e-mol for organic matter was used as shown before. Substitution of $-\Delta G_{CAT}$ gives Y_{COD} for aerobic growth on organic substrate

$$Y_{COD} = \frac{110}{110 + (1/Y_{GX})/\gamma_x} \tag{11c}$$

This equation shows that Y_{COD} for aerobic growth is not a constant, as often assumed (26) with $Y_{COD} \approx 0.50$–0.67, but it depends also on the type of C source, because this determines $1/Y_{GX}$ (equations 3a and 3b). Also, to decrease the Y_{COD} leading to lower surplus-sludge production, one must, according to equations 11b and 11c:

- Decrease ($-\Delta G_{CAT}$), by, for example, using anaerobic metabolism producing CH_4
- Increase $1/Y_{GX}$ by increasing temperature or decreasing the growth rate according to equation 1

These predicted phenomena are well known and applied in waste-treatment processes.

Relation between Heat Production and Gibbs Energy Dissipation. According to equation 9d, the heat production ($1/Y_{QX}$) is related to Gibbs energy need ($1/Y_{GX}$) by the enthalpy and Gibbs energy of the catabolic reaction (ΔG_{CAT} and ΔH_{CAT}). Table 6 shows some examples of growth systems to illustrate the relation between heat production and dissipated Gibbs energy for growth (29). From Table 6 we can conclude the following rules of thumb:

- For aerobic (or denitrifying) growth systems on organic substrate, the Gibbs energy dissipation and heat production are nearly equal. The entropy contribution in the catabolic reaction is minimal (see glucose and acetate aerobic growth).
- For anaerobic growth, heat production and Gibbs energy dissipation can be substantially different, due to entropic effects.

Obviously, if in the catabolic reaction there is a net decrease of molecules or a consumption of gaseous molecules, then there is a strong negative entropy contribution (see H_2/CO_2 aerobic and anaerobic) and there is a much higher heat production than Gibbs energy dissipation. If, however, in the catabolic reaction there is a net production of the amount of molecules and/or production of gaseous molecules (e.g., the glucose/ethanol fermentation or the methane production from acetate), then there is a very large positive entropy contribution, leading to a much lower heat production than the Gibbs energy dissipation. The entropic effect can even be so large that there is a calculated heat uptake during growth (e.g., methanation of acetate). This is obviously endothermic growth. So, contrary to a common belief, growth of microorganisms is not necessarily related to heat production; there can be heat uptake as well. Experimental proof is, however, still lacking.

Maximal Product Yields in Anaerobic Metabolism. In many microbial processes the valuable product (e.g., ethanol or lactic acid) is related to catabolism. The relevant stoichiometric coefficient is then the yield of the electron acceptor couple to electron donor Y_{DA}.

Equation 9e shows how this coefficient is determined by various factors and it appears that Y_{DA} is maximized.

- For high Gibbs energy dissipation $1/Y_{GX}$. This means that high catabolic product yields are achieved for poor carbon sources, low growth rate, and high temperature, because $1/Y_{GX}$ is then maximized.
- For catabolic reactions with low ΔG_{CAT}. This is understandable, because then the growth yield is minimized, which leads directly to higher product yield.
- For highly reduced electron donors (γ_D high) and highly oxidized products (γ_A low). It is then even possible to achieve C yields larger than 1. An excellent example is the anaerobic production of acetate from methanol, where $Y_{DA} \approx 1.4$ C-mol acetate/C-mol methanol (Example 4).

KINETICS OF MICROBIAL GROWTH FROM A THERMODYNAMIC POINT OF VIEW

In the previous sections, the full stoichiometric description of growth has been given from a unifying thermodynamic point of view. However, for a complete description, growth kinetics are also required. In this section these kinetics are also presented from a thermodynamic point of view.

A Basic Kinetic Description of Microbial Growth

In Figure 1, the typically observed batch-growth curves of a microorganism growing on one substrate (electron donor) are shown. In this section we use the subscript S to denote substrate (electron donor, D). Usually these curves are described (for constant batch volume) with two mass balances for substrate (donor) and biomass, giving two differential equations (DE):

Table 6. Relation between Heat Production and Gibbs Energy Need

Microorganism	Growth condition	Y_{DX} (C-mol/(C)-mol)	$1/Y_{GX}$, Gibbs energy (kJ/C-mol biomass)	$1/Y_{QX}$, heat (kJ/C-mol biomass)	Entropy contribution (kJ/C-mol biomass)
Saccharomyces cerevisae	Glucose aerobic	0.57	332	+339	−7
Saccharomyces cerevisae	Glucose anaerobic	0.14	270	+95	+175
Hydrogenotroph	$H_2 + CO_2$ aerobic	0.13	1,265	+1686	−421
Methanobacterium arborophilus	$H_2 + CO_2$ anaerobic	0.015	1,035	+3,923	−2,888
Pseudomonas oxalaticus	Acetate aerobic	0.406	562	+593	−31
Methanobacterium soehngenii	Acetate anaerobic	0.024	597	−90	+687

Note: Heat production and Gibbs energy dissipation for (an)aerobic growth on glucose, H_2, and acetate; relative contribution of heat- and entropy-related dissipation.

$$\frac{dC_s}{dt} = q_s C_x \tag{12a}$$

$$\frac{dC_x}{dt} = \mu C_x \tag{12b}$$

where q_s and μ are the biomass-specific electron-donor (substrate) uptake and growth rate. To solve these DEs, we need kinetic expressions for q_s and μ. However, q_s and μ are stoichiometrically coupled to each other according to the Herbert–Pirt equation (equation 1a). This equation for substrate can be written as (realizing that $Y_{DX} \equiv Y_{SX} = \mu/-q_s$, and $m_S \equiv m_D$)

$$-q_s = \frac{1}{Y^m_{SX}}\mu + m_S \tag{13}$$

It is now only necessary to specify one kinetic equation, which can be for q_s of μ. In practice the Monod kinetic equation for μ ($\mu = \mu^m C_s/(C_s + K_s)$ is often chosen. This choice leads, however, to a very nasty inconsistency. According to Monod, $\mu = 0$ if the concentration of substrate C_s becomes 0. Substituting this result in the Herbert–Pirt equation shows, however, that (for $\mu = 0$) $q_s = m_s$. The inconsistency is then that in the absence of substrate ($C_s = 0$, $\mu = 0$) there is still consumption of substrate for maintenance ($v = -q_s = m_s$); this is very strange indeed. These problems can be avoided in a most simple way by introducing the kinetics of the substrate consumption rate in the most simplistic way (a Michaelis–Menten type of relation, see Figure 8):

$$-q_s = q_s^{max}\frac{C_s}{K_s + C_s} \tag{14a}$$

Combining equation 14a with equation 13 by eliminating q_s, leads to the kinetic relation of μ

$$\mu = Y^m_{SX}\, q_S^{max}\frac{C_s}{K_s + C_s} - Y^m_{SX}\, m_S \tag{14b}$$

The kinetic relations 14a and 14b contain four model parameters: Y^m_{SX}, K_s, m_s, and q_s^{max}. Figure 8 shows how μ and q_s depend on C_s.

Clearly, there is a maximal substrate uptake rate q_s^{max}, but also for μ a maximal value (μ^{max}) is seen at high C_s. The value of μ^{max} follows from equation 14b by taking the limit $C_s \gg K_s$, and the result is

$$\mu^{max} = Y^m_{SX}(q_s^{max} - m_s) \tag{15a}$$

In addition, it is clear from Figure 8 that there is a minimal substrate concentration ($C_{s,min}$) at which $\mu = 0$. For $C_s < C_{s,min}$, the growth rate μ becomes negative. An expression for $C_{s,min}$ can be found from equation 14b by putting $\mu = 0$, which results in

$$C_{s,min} = K_s\left(\frac{m_S}{q_s^{max} - m_s}\right) \tag{15b}$$

The occurrence of a minimal substrate concentration has indeed been observed (30,31). It can also be seen (Fig. 8) that at $C_s = 0$ the growth rate is negative.

$$(\mu)_{C_s=0} = -Y^m_{SX}m_S = -k_d \tag{15c}$$

Clearly, in the absence of substrate, the microorganism decreases its biomass; this phenomena is often called endogenous respiration, or biomass decay, where the kinetic parameter is k_d. Equation 14b can be rewritten by eliminating q_s^m and m_S using equations 15a and 15b leading to

$$\mu = \mu^{max}\frac{(C_s - C_{s,min})}{K_s + C_s} \tag{14c}$$

Alternatively equation 14b can be rewritten using the parameter k_d

$$\mu = \mu^{max}\frac{C_s}{K_s + C_s} - k_d\frac{K_s}{K_s + C_s} \tag{14d}$$

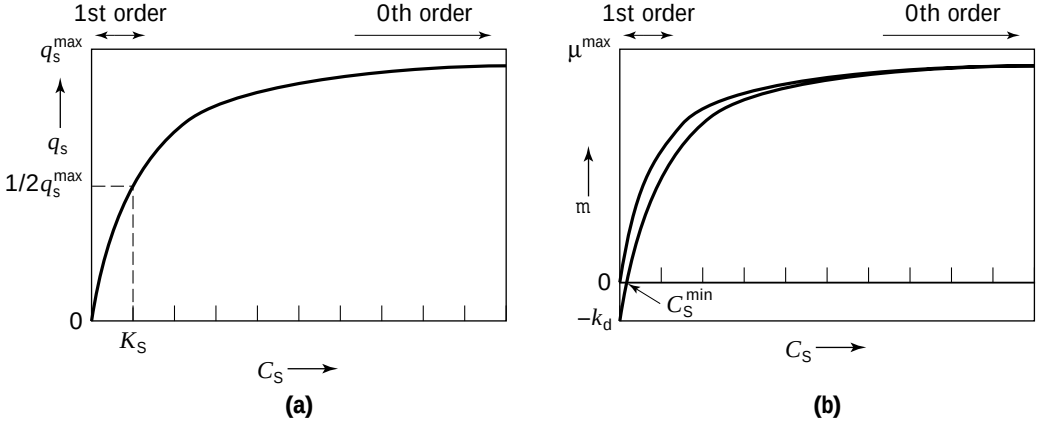


Figure 8. Dependence of q_s and μ as function of C_s for the Monod equation for μ (a) or the Michaelis–Menten equation for q_s (b).

It is clear that the obtained formulations for μ (equations 14b, 14c, and 14d) all become the familiar Monod equation if $m_S = 0$ (meaning $C_{s,min} = 0$ and $k_d = 0$; see equations 15b and 15c). The Monod equation and equations 14c and 14d are very close for $C_s \gg K_s$, but substantial differences arise for $C_s < K_s$ (see Fig. 8).

A final point to address is the implicit assumption (in equation 14a) that in the substrate-uptake process the substrate concentration can become zero. This is, in principle, not possible, because there always exists a substrate concentration where the catabolic reaction reaches a point where energy production in the form of ATP becomes impossible. This concentration is called the threshold concentration of substrate $C_{s,thresh}$.

From published data it is known that μ^{max} can vary in a wide range (0.001–1 h^{-1}) and that K_s is also covering an even wider range (10^{-6} to 10^{-3} mol/L). In the preceding sections m_S and Y^m_{SX} have been studied from a thermodynamic point of view. In the following, we present correlations and points of view that allow the estimation k_d, μ^{max}, K_s, and $C_{s,thresh}$, all based on a thermodynamic point of view. Such relations are very relevant in, for example, considerations of waste-treatment processes where the threshold concentration determines the effluent quality, where k_d determines the surplus sludge production, and where μ^{max} directly determines the reactor size or the type of reactor (suspended organisms or biomass retention systems). For batch-fermentation processes, μ^{max} is important because it determines the duration of a batch-growth process.

A Thermodynamic Relation for the Endogeneous/Decay Parameter k_d

According to equation 15c, k_d can be calculated from the relations for Y^m_{SX} and m_G found previously (equations 9c and 10a). Substitution gives

$$k_d = \frac{m_G}{1/Y^m_{GX} + (\gamma_X/\gamma_D)(-\Delta G_{CAT})} \qquad (16)$$

The values of m_G and Y^m_{GX} are obtained from the thermodynamic correlations equations 2 and 3. This equation shows that for aerobic heterotrophic growth (where $1/Y^m_{GX} \approx$ 300–600 kJ/C-mol biomass) the expected k_d will be a factor three times larger than for aerobic autotrophic growth (such as nitrification, for which $1/Y^m_{GX} \approx$ 3500 kJ/C-mol x). This has indeed shown to be the case (4), indicating that equation 16 is useful.

A Thermodynamic Correlation for μ^{max}

The value μ^{max} is the result of a limiting factor in metabolism. This metabolism can be represented very schematically as shown in Figure 9. Three possible bottlenecks are identified (Fig. 9):

1. The uptake rate of substrate
2. The rate of synthesis of biomass, as related to the ribosomal capacity to synthesize biomass protein

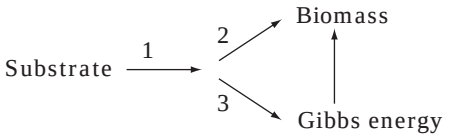


Figure 9. Possible bottlenecks in metabolism.

3. The rate of making Gibbs energy available to enable growth and maintenance

Here we assume as a simple working hypothesis that the rate of making Gibbs energy available for microorganisms is limited by a maximum rate of electron transport in the catabolic energy production. This hypothesis is supported by studies with *Escherichia coli*, where μ^{max} was measured aerobically (at the same temperature) for different growth substrates (32). It was observed that for all substrates the same maximal biomass specific O_2 uptake rate was achieved by *E. coli*. This supports the hypothesis of a constant maximal electron-transport capacity in microorganisms. A similar hypothesis was put forward by McCarthy (26). The following correlation is now proposed for the maximal electron-transport capacity in the microbial electron transport chain (in mol electrons/C-mol biomass/h) as a function of temperature:

$$\text{Maximal electron-transport capacity} = 3 \exp\left[\frac{-69000}{R}\left(\frac{1}{T} - \frac{1}{298}\right)\right]$$

This correlation is an Arrhenius type of relation with an energy of activation of 69,000 J/mol. R is the gas constant (8.314 J/mol K). The value of 69,000 J/mol is an activation energy that follows from observations of the effect of temperature on μ^{max} of microorganisms (1).

For a given electron donor/acceptor combination, the maximal rate of Gibbs energy made available per unit biomass (q_G^{max} in kJ Gibbs energy per C-mol biomass per hour) follows then as

$$q_G^{max} = 3[(-\Delta G_{CAT})/\gamma_D] \exp\left[\frac{-69{,}000}{R}\left(\frac{1}{T} - \frac{1}{298}\right)\right]$$

It should be noted that $(-\Delta G_{CAT})/\gamma_D$ is the Gibbs energy made available per mole electron transported between donor and acceptor by the electron transport chain.

Now, by definition, the Gibbs energy is spent for growth and maintenance and we can write under maximum growth rate condition

$$q_G^{max} = \frac{1}{Y^m_{GX}}\mu^{max} + m_G$$

Eliminating q_G^{max} leads to the final equation (equation 17), which gives μ^{max} as a function of temperature (after replacing m_G using equation 2), Y^m_{GX}, γ_D, and ΔG_{CAT}.

Table 7. Estimated μ^{max} Values at 25 °C Using Equation 17

Microbial system	$-\Delta G_{cat}/\gamma_D$ (kJ/mol electron)	$1/Y^m_{GX}$ (kJ/C-mol biomass)	μ^{max} (h^{-1}, 25 °C)
Aerobic/glucose	118.5	236	1.5
Aerobic/acetate	105.5	432	0.7
Anaerobic/CH_4 from acetate	3.87	432	0.015
Anaerobic/ethanol from glucose	9.39	236	0.10
Aerobic/Fe^{2+} oxidation (pH = 1.5)	38.6	3,500	0.03
Aerobic/sulfide oxidation to SO_4^{2-}	99.6	3,500	0.08
Aerobic/nitrification	45.8	3,500	0.04

$$\mu^{max} = \frac{[3(-\Delta G_{CAT})/\gamma_D - 4.5]}{1/Y^m_{GX}} \exp\left[\frac{-69{,}000}{R}\left(\frac{1}{T} - \frac{1}{298}\right)\right] \quad (17)$$

Table 7 shows the estimated μ^{max} values (25 °C = 298 K) for well-known growth systems. These estimated maximal growth rates are reasonable in the range of reported values. For the only prediction, which is rather wrong (system 4), it is indeed known that the Gibbs energy conversion is not due to electron transport, but due to substrate phosphorylation.

Affinity Constant of Electron Donor, K_s

The constant K_s is an affinity constant in the substrate-uptake kinetics. It is known that reported K_s values can vary over a wide range (10^{-3} to 10^{-6} mol/L). Even for the same organism, for example, *E. coli*, many different K_s values are found (30). Much of this wide spread is probably due to systematic errors in the determination of the K_s parameter. These errors can be due to ill-defined conditions of the microorganisms (adaptation), statistically unsound procedures to evaluate K_s from experimental data, neglect of mass-transfer limitation in the case of microorganisms growing as flocs, and analysis/sampling problems in determining the substrate concentration (see Ref. 30 for an extensive discussion). Moreover, K_s is a typical kinetic parameter, determined by the kinetic properties of the first step in substrate transport into the microorganism. For example, for the anaerobic acetate conversion into CH_4, the *Methanosarcina* type of organism has a high K_s value of 3 mM acetate and its substrate uptake is passive (non-energy linked). However, for the *Methanotrix* type of organism, where acetate is transported into the cell utilizing energy, $K_s \approx 0.5$mM (33).

Due to these factors, it is not possible to give, from a thermodynamic point of view, a generalization about the value of K_s for different microbial growth systems.

Threshold Concentration

In the previous section the substrate consumption kinetics q_s was assumed as an irreversible Michaelis–Menten kinetics equation (equation 14a). From this it follows that q_s becomes zero at substrate concentration $C_s = 0$. However, it is known that microbial metabolism stops at a certain concentration of the substrate. This is called the threshold concentration $C_{s,thresh}$.

This threshold concentration is thus the substrate concentration where substrate consumption q_s becomes zero. This should be distinguished from the minimal substrate concentration $C_{s,min}$, where the growth rate μ becomes zero.

This threshold situation is achieved when the coupled system of catabolic reaction and the energy generating system of the organism [most probably the proton motive force (pmf) process] are in equilibrium. Because the pmf requires at least 15–20 kJ for each proton (6), and assuming that a full catabolic reaction is minimally coupled to the extrusion of one proton, it appears that there must be a minimal catabolic Gibbs energy release of about 15–20 kJ. The following examples seem to support this idea where threshold behavior has been reported for a number for growth systems (Example 10):

- Anaerobic metabolism of CH_4 production from acetate (33)
- Aerobic metabolism of ferrous into ferric iron (12,13)
- Anaerobic production of H_2 converting 1 ethanol into 2 H_2 and 1 acetate and consumption of 4H_2 to produce acetate, CH_4, or H_2S from $2HCO_3^-$, HCO_3^- or SO_4^{2-} (34,35)

In all these situations of observed threshold concentration the actual ΔG_{CAT} could be calculated to be in the range of -10 to -30 kJ per catabolic reaction. Example 10 shows how the actual $(\Delta G)_{CAT}$ follows from textbook thermodynamic calculations.

EXAMPLE 10

Threshold concentrations and minimally required catabolic Gibbs energy for aerobic Fe^{2+}-oxidizing bacteria (12,13)

Fe^{2+} can be aerobically oxidized by specific bacteria. The catabolic reaction is $Fe^{2+} + 1/4O_2 + H^+ \rightarrow Fe^{3+} + 1/2H_2O$. The ΔG^0_{CAT} can be calculated to be -44.32 kJ (standard conditions). This aerobic iron oxidation can be performed by two different microorganisms, *Thiobacillus ferro oxidans* (T.f.) and *Leptospirillum ferro oxidans* (L.f.). Recently (12,13) the Fe^{2+} concentration was observed where the O_2 consumption stopped (threshold concentration). The following results were found (30 °C):

T.f.: $P_{02} = 0.12$ bar, pH = 1.85, $Fe^{3+} = 0.21$ M,
$Fe^{2+} = 3 \times 10^{-4}$ M (= threshold concentration)

L.f.: $P_{O2} = 0.12$ bar, pH = 1.55, $Fe^{3+} = 0.21$ M,
$Fe^{2+} = 7 \times 10^{-6}$ M (= threshold concentration)

Using these concentrations the following ΔG_{CAT} can be calculated for *T. ferro oxidans*

$$\Delta G_{CAT} = -44.32 + 5.8 \log\left[\frac{0.21}{3 \times 10^{-4} \times 10^{-1.85}(0.12)^{0.25}}\right] = -15.75 \text{ kJ}$$

Similarly we can calculate for *L. ferro oxidans* $\Delta G_{CAT} = -8.03$ kJ.

Threshold concentrations and minimally required Gibbs energy for anaerobic acetate consuming methanogens

CH_4 can be produced anaerobically from acetate according to the catabolic reaction Acetate + $H_2O \rightarrow CH_4 + HCO_3^-$. For this reaction ΔG^0_{CAT} and ΔH^0_{CAT} at 298 K follow as

$$\Delta G^{01}_{CAT} = -31.0 \text{ kJ}$$
$$\Delta H^0_{CAT} = +5 \text{ kJ}$$

This provides, using the Van 't Hoff equation, for $\Delta S^{01}_{CAT} = \Delta H_r - \Delta G_r/T = 0.121$ kJ/mol K.

There are two organisms known to perform this reaction at 60 °C, both of which have threshold acetate concentrations. According to Ref. 32 *Methano sarcina* has a threshold concentration of 0.3×10^{-3} M and *Methano trix* of 16×10^{-5} M of acetate.

Using a CH_4 pressure of 0.6 bar, T = 333 K, and a bicarbonate concentration of 0.03 M, we can calculate the following ΔG_{CAT} of the catabolic reaction at the threshold acetate concentration (using the Van 't Hoff relation $\Delta G = \Delta H - T\Delta S$ to take the temperature effect into account):

Methanosarcina $\Delta G_{CAT} = -23$ kJ

Methanotrixe $\Delta G_{CAT} = -15.7$ kJ

It appears indeed that at the threshold concentration the coupled system of catabolic reaction and proton translocation may be in equilibrium. This leads to the statement that a threshold concentration can be estimated from the requirement that

$$(\Delta G_{CAT})_{thresh} \approx -20 \text{ kJ} \qquad (18)$$

This explanation is further supported by the observation that for anaerobic H_2-consuming systems (36), the observed effect of temperature on H_2 threshold values can be explained from correlation equation 18. Example 11 shows how the expected threshold concentration can be estimated for such systems.

EXAMPLE 11

Calculation of the threshold concentration as a function of temperature

Consider the anaerobic catabolic system $4H_2 + HCO_3^- + H^+ \Leftrightarrow CH_4 + 3H_2O$. One can calculate that $\Delta G^{01}_{CAT} = -135.57$ kJ, and $\Delta H^0_{CAT} = -241$ kJ. Using the Van 't Hoff relation ($\Delta G = \Delta H - T\Delta S$) one calculates $\Delta S^0_{CAT} = -0.354$ kJ/molK.

Assuming $P_{CH_4} = 1$ bar, $HCO_3^- = 0.01$ M, $H^+ = 10^{-8}$ M, $T = 303$ K, and ΔG_{CAT} equals the threshold minimum of -20 kJ, one can calculate at threshold conditions:

$$-20 = -241 - (-0.354) \times 303 + 0.008314 \times 303 \ln \frac{1}{(10^{-8}/10^{-7})(0.01)(P_{H_2})^4}$$

From this follows that the threshold hydrogen pressure is $P_{H_2} = 7 \times 10^{-5}$ bar at a temperature of 30 °C (303 K).

For a temperature of 75 °C (=348 K), one finds that $(P_{H_2})_{threshold} = 120 \times 10^{-5}$ bar. Such an increase of hydrogen threshold partial pressure with increasing temperature has indeed been observed (36). The calculated threshold H_2 pressures are also in the correct range.

From the previous findings it is clear that, especially for systems with a low ΔG_{CAT} (e.g., anaerobic, inorganic electron donors), threshold values can be found. For aerobic growth, ΔG_{CAT} is so large that a measurable threshold value is not expected because the thermodynamically calculated $C_{s,thresh}$ would be extremely low.

The existence of threshold concentrations makes it desirable to change the kinetics of substrate uptake (q_s, equation 14a) from the irreversible to reversible form

$$q_s = \frac{q_s^m(C_s - C_{s,thresh})}{K_m + (C_s - C_{s,thresh})} \qquad (19)$$

In conclusion it appears that threshold concentrations of electron donor (or substrate) do exist and their value can be estimated from the catabolic reaction using the value of $(\Delta G_{cat})_{thresh}$ in equation 18. In addition a more proper kinetic expression for q_s is then given by equation 19.

NOMENCLATURE

Symbols

C_i	Concentration of compound *i*, mol/m^3
Y_{ij}	Yield of compound *j* on compound *i*, mol$_j$/mol$_i$
μ^{max}	maximal specific growth rate of biomass, h^{-1}
k_d	Decay coefficient of biomass, h^{-1}
m_i	Biomass specific maintenance requirement of compound i, mol$_i$/C-mol biomass h
K_s	Affinity constant for substrate uptake, mol/m^3
r_i	Rate of conversion of compound i per reactor volume, (C)mol/m^{3}h
q_i	Biomass specific rate of uptake or secretion of compound i, (C) mol i/C-mol biomass h
H	Enthalpy, kJ
G	Gibbs energy, kJ
Q	Heat, kJ
γ	Degree of reduction: for organic compounds it is defined per C-mol, electron/C-mol; for inorganic compounds it is defined per mol, electron/mol

Subscripts

D	Electron donor
A	Electron acceptor
X	Biomass
e	Calculated per mol electron
E	Calculated per mol electron
O	Oxygen
S	Substrate
C	Carbon dioxide
N	N-source
Q	Heat
H	Proton
W	Water
CAT	Catabolic reaction

Superscripts

m	Value at infinite substrate concentration
max	Value under maximal conditions
o	Reference conditions (1 mol/L, 1 atm, 298 K)
01	Reference conditions, but with pH = 7 ($H^+ = 10^{-7}$ mol/L)

ACKNOWLEDGMENTS

I would like to thank J. A. Roels, H. V. Westerhoff, U. von Stockar, and E. H. Battley for fruitful and pleasant discussions on thermodynamics in the past years.

BIBLIOGRAPHY

1. J.A. Roels, *Energetics and Kinetics in Biotechnology*, Elsevier, New York, 1983.
2. J.J. Heijnen and J.P. van Dijken, *Biotechnol. Bioeng.* **39**, 833–858 (1992).
3. J.J. Heijnen, M.C.M. van Loosdrecht, and L. Tijhuis, *Biotechnol. Bioeng.* **40**, 1139–1154 (1992).
4. L. Tijhuis, M.C.M. van Loosdrecht, and J.J. Heijnen, *Biotechnol. Bioeng.* **42**, 509–519 (1993).
5. E.H. Battley, *Energetics of Microbial Growth*, Wiley New York, 1987.
6. H.V. Westerhoff and K. van Dam, *Mosaic Non-Equilibrium Thermodynamics and the Control of Biological Free Energy Transduction*, Elsevier, Amsterdam, 1987.
7. J.J. Heijnen, *Antonie van Leeuwenhoek* **60**, 235–256 (1991).
8. J.J. Heijnen and J.P. van Dijken, *Biotechnol. Bioeng.* **42**, 1127–1130 (1993).
9. E.H. Battley, *Appl. Environ. Microbiol.* **61**, 1655–1657 (1995).
10. A.H. Stouthamer, *Antonie van Leeuwenhoek* **29**, 545–565 (1973).
11. R.K. Thauer, K. Jungermann, and K. Decker, *Bacterial. Rev.* **41**, 100–180 (1977).
12. M. Boon, J.J. Heijnen, and G.S. Hansford, in D.S. Holmes and R.W. Smith eds., *Minerals Bioprocessing II. Proceedings of the Engineering Foundations Conference: Minerals Processing II*, Snowbird, Utah, July 10–15, 1995; TMS, Warrendale, Pa., 1995, pp. 41–62.
13. M. Boon, J.J. Heijnen, and G.S. Hansford, in D.S. Holmes and R.W. Smith eds., *Minerals Bioprocessing II. Proceedings of the Engineering Foundations Conference: Minerals Processing II*, Snowbird, Utah, July 10–15 1995, TMS, Warrendale, Pa., 1995, pp. 63–82.
14. N. Schill, W.M. van Gulik, D. Voisard, and U. von Stockar, *Biotechnol. Bioeng.* **51**, 645–658 (1996).
15. B.H.A. van Kleeff, J.G. Kuenen, and J.J. Heijnen, Biotechnol. Bioeng. **12**, 510–518 (1996).
16. U. von Stockar, L. Gustafsson, C. Larsson, J. Marison, P. Tissot, and E. Gnaiger, *Biochim. Biophys. Acta* **1183**, 221–240 (1993).
17. U. von Stockar and I.W. Marison, *Adv. Biochem. Eng. Biotechnol.* **40**, 93–136 (1989).
18. H.J. Noorman, B. Romein, K.Ch.A.M. Luyben, and J.J. Heijnen, Biotechnol. Bioeng. **49**, 364–376 (1996).
19. R.T.J.M. van der Heijden, J.J. Heijnen, C. Hellinga, B. Romein, and K.Ch.A.M. Luyben, *Biotechnol. Bioeng.* **43**, 3–10 (1994).
20. R.T.J.M. van der Heijden, B. Romein, J.J. Heijnen, C. Hellinga, and K.Ch.A.M. Luyben, *Biotechnol. Bioeng.* **43**, 11–20 (1994).
21. R.T.J.M. van der Heijden, B. Romein, J.J. Heijnen, C. Hellinga, and K.Ch.A.M. Luyben, *Biotechnol. Bioeng.* **44**, 781–791 (1994).
22. C. Hellinga and B. Romein, in M.N. Karim and G. Stephanopoulos eds., *Modeling and Control of Biotechnical Processes 1992*, IFAC Symposia Series 10, 1992.
23. J. Monod, *Ann. Rev. Microbiology* **3**, 371–394 (1949).
24. S.J. Pirt, *Proc. R. Soc. London, Ser. B* **163**, 224–231 (1965).
25. O.M. Neijssel and D.W. Tempest, *Arch. Microbiol.* **107**, 215–230 (1976).
26. P.L. McCarty, in S.D. Faust and J.V. Hunter eds., *Organic Compounds in Aquatic Environments*, Marcel Dekker, New York, 1971, pp. 495–531.
27. J.J. Heijnen, J.A. Roels, and A.H. Stouthamer, *Biotechnol. Bioeng.* **21**, 2175–2201 (1979).
28. L.E. Erickson, I.G. Minkevich, and V.K. Eroshin, *Biotechnol. Bioeng.* **20**, 1595–1605 (1978).
29. J.J. Heijnen, *TIBTECH* **12**, 483–492 (1994).
30. K. Kovárová, A.J.B. Zehnder, and T. Egli, *J. Bacteriol.* **178**, 4530–4539 (1996).
31. M.E. Tros, G. Schraa, and A.J.B. Zehnder, *Appl. Environ. Microbiol.* **62**, 437–442 (1996).
32. K.B. Andersen and K. von Meyenburg, J. Bacteriol. **144**, 114–123 (1980).
33. S.H. Zinder, *FEMS Microbiol. Rev.* **75**, 125–138 (1990).
34. H.J. Seitz, B. Schink, N. Pfennig, and R. Conrad, *Arch. Microbiol.* **155**, 82–88 (1990).
35. H.J. Seitz, B. Schink, N. Pfennig, and R. Conrad, *Arch. Microbiol.* **155**, 89–93 (1990).
36. R. Conrad and B. Wetter, *Arch. Microbiol.* **155**, 94–98 (1990).

See also ANAEROBES; BIOREMEDIATION; CLOSTRIDIA, SOLVENT FORMATION; ELECTRON TRANSPORT; ENERGY METABOLISM, MICROBIAL AND ANIMAL CELLS; YEAST, BAKER'S.

BIOFILMS, MICROBIAL

JAMES D. BRYERS
University of Connecticut
Farmington, Connecticut

KEY WORDS

Bacterial adhesion
Biofilm detection
Biofilm eradication
Biofilm formation
Biofilms
Biofouling
Cell attachment
Sloughing

OUTLINE

THE IMPACT OF BIOFILM FORMATION

Biofilms are collections of microorganisms, predominantly bacteria, enmeshed within a three-dimensional gelatinous matrix of extracellular polymers secreted by the microorganisms (Fig. 1). In an aqueous environment, a support, termed a substratum, will be immediately biased by dissolved organic molecules and macromolecules that adsorb rapidly from the liquid phase. Bacterial cells present in the fluid contact the substratum by a variety of transport mechanisms. Once at the substratum, the cells can adsorb either reversibly or irreversibly. Provided the cells remain at the surface for a sufficient time, they will secrete extracellular polymers that serve to attach the cells tenaciously to the substratum. Attached cells metabolize prevailing energy and carbon sources (either dissolved within the surrounding fluid, adsorbed to the substratum surface, or existing as a constituent of the substratum itself), reduce terminal electron acceptors, grow, replicate, and produce insoluble extracellular polysaccharides, thus accumulating an initial viable biofilm community. Inert particles, bacterial cells of the same or different species, and higher life forms (e.g., algae, amoeba, protozoa) continue to be recruited from the fluid and incorporated into the biofilm community.

As the biofilm bacterial communities mature, the adherent populations will oxidize electron donors and reduce existing terminal electron acceptors in an order of decreasing redox potential. As the biofilm community becomes denser and thicker, the penetration depth of one electron acceptor will overlap that of another; thus, a succession of different microbial populations can be established. Biofilm layers near oxygen-saturated water will harbor aerobic heterotrophic or autotrophic activity, reducing the available oxygen. As the distance from the aerobic interface increases and oxygen is depleted, a sequence of terminal electron acceptors is utilized by specific microbial populations: facultative denitrifiers using nitrate and nitrite; anaerobic sulfate reducers using sulfate; fermentative microbes partially reducing organic carbon compounds; and finally, anaerobic methane producers. Depending upon the balance between mass transfer rates and microbial reaction rates, biofilms can develop stratified ecosystems, on a microscopic scale, similar to those observed in lake or marine sediments. In engineered systems, highly active aerobic activity within a biofilm can completely deplete oxygen concentrations over distances of 10–20 μm, creating anaerobic layers capable of supporting obligate, anaerobic sulfate-reducing bacteria. Consequently, it should come as no surprise to discover anaerobic sulfate reduction and hydrogen sulfide corrosion of the metallic surfaces in heat exchange systems that utilize highly aerated cooling water to condense steam.

As a result of hydrodynamic forces and stresses exerted by replication, there can be a continual erosion of cells and extracellular material from the biofilm back to the bulk fluid. A more random stochastic process known as "sloughing" can also occur where either large sections or the entire biofilm becomes displaced from the substratum and enters the liquid.

Biofilms are as versatile as they are ubiquitous. Intentional use of biofilms can serve many benefits, for example, in the water and waste treatment industry, in bioremediation applications, and in industrial biotechnology. Unintentionally formed biofilms can create such detriments as biofouling of heat exchange systems and marine structures; microbial-induced corrosion of metal surfaces or the deterioration of dental surfaces; contamination of household products, food preparations, and pharmaceuticals; and the infection of short- and long-term indwelling biomedical implants and devices. Such detriments can range in severity from being a mere nuisance to being life threatening. This chapter will first summarize the impact of biofilms, both beneficial and detrimental, on engineered and natural processes. A detail discussion of the fundamental rate processes governing biofilm formation and persistence is then given, followed by a summary of recent advances in biofilm research.

Detrimental Biofilms

Figure 2 illustrates, for an arbitrary analytical measure, the typical accumulation of biofilm at a surface as a func-

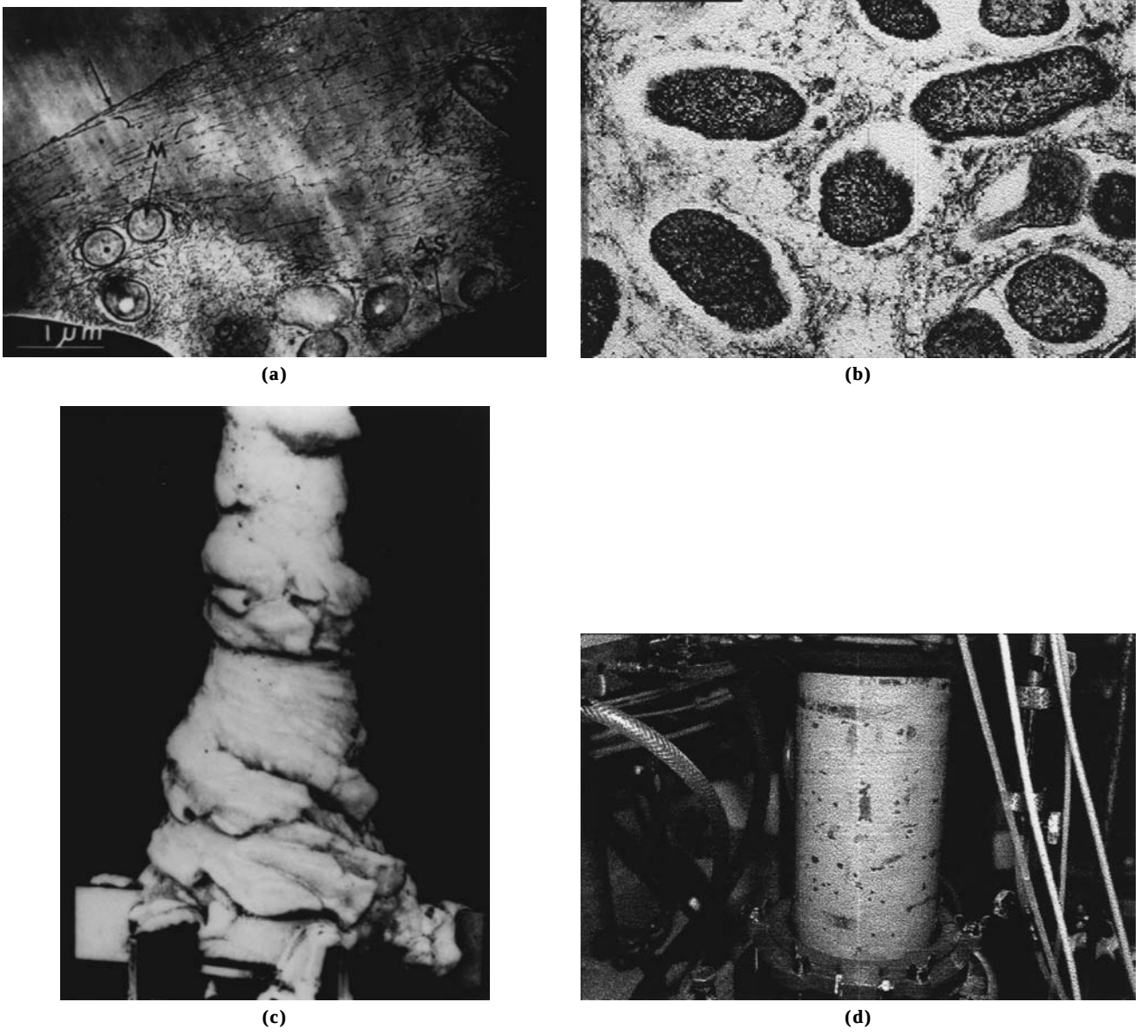


Figure 1. (**a**) Transmission electron micrograph of a sagital Section of a mixed-culture biofilm accumulating upon the surface of a rock in an alpine stream (1). (**b**) Transmission electron micrograph of a *Pseudomonas aeruginosa* biofilm isolated from a cystic fibrosis patient (2). (**c**) *Hyphomicrobium* spp. biofilm accumulating upon surfaces of a laboratory bioreactor vessel (3). (**d**) Accumulation of a yeast biofilm upon the impeller of a 200-L bioreactor (4).

tion of time. Also shown are the changes in frictional and heat transfer resistances observed in engineered systems experiencing biofilm formation.

Uncontrolled biofilm formation within natural, engineered, and biomedical systems can create numerous detriments as detailed in several definitive reviews (5–7). Detriments arise by way of a biofilm's influence on the transport of mass, momentum, and energy. Detriments attributed directly to bacterial biofilms include (*1*) material deterioration and corrosion (8–16), (*2*) increases in both frictional and heat transfer resistances (14,17), (*3*) attachment to and infection of biomedical implant devices (18–25), and (*4*) operational problems that plague lab-scale and full-scale bioreactors (6,7,26).

Beneficial Biofilms

Benefits afforded by biofilms in a continuous reactor situation arise chiefly because the cell population is immobilized, and thus the residence time of the cells in the reactor is independent of the fluid-phase residence time. In continuous suspended culture bioreactors (i.e., chemostats), the

mean residence time of the system cannot be less than the generation time of the bacterial species, otherwise cells do not have sufficient time to replicate within the reactor and are eventually diluted from the system.

Immobilized and biofilm-bound cells remain in a continuous reactor system independent of the fluid phase, thus the mass loading of limiting substrate (or influent pollutant in the case of a wastewater treatment reactor) can be increased well beyond the growth rate limit imposed on suspended cultures. Consequently, immobilized-cell or biofilm reactors can provide (*1*) added volumetric reactivity, (*2*) more stable operating performance, (*3*) an inherent ease in biomass–fluid separation, and (*4*) the prospect of staging different bioconversion processes in sequential reactors. Due to these inherent advantages, the use of biofilm reactors is not confined to just bacterial cells, but also comprises plant and animal cell applications.

Bacterial biofilm reactors are employed either in commodities production or in wastewater treatment applications. Biofilm reactors have been reportedly used to produce acetic acid (27), L-lysine (28), gluconic acid (29), kojic acid (30,31), and ethanol (32,33), and in the epoxidation of propene (34). Such biofilm reactors are operated either as packed- or fluidized-bed reactor systems, with cells either attached to an inert support particle or artificially immobilized within a gel matrix.

One major application that relies on a microbial culture's ability to form biofilms is wastewater treatment. Biofilm reactor configurations, applied in both pilot- and full-scale wastewater treatment, include packed-bed trickling filters, high-rate plastic media filters, rotating biological contactors, fluidized-bed biofilm reactors, and membrane-immobilized cell reactors. Examples of biofilm reactors employed for wastewater treatment reported in the literature include polychlorinated hydrocarbon degradation (35), toluene degradation (36), denitrification (37,38), cadmium removal (39), anaerobic butyrate degradation (40), nitrification (41–43), glyphosphate degradation (44), anaerobic propionate degradation (45), phenolic wastewaters (46), uranium removal (47), and anaerobic carbon removal (48,49).

PROCESSES GOVERNING THE FORMATION AND PERSISTENCE OF BIOFILMS

Figure 2 illustrates, for an arbitrary analytical measure, the typical accumulation of biofilm at a surface as a function of time. Initially, the substratum is conditioned and cells attach reversibly, then irreversibly. Next, attached cells grow, reproduce, and secrete insoluble extracellular polysaccharidic material. As the biofilm matures, biofilm detachment and growth processes come into balance, such that the total amount of biomass on the surface remains approximately constant in time. Multiple physical, chemical, and biological processes operate in each phase (Fig. 3).

Processes governing biofilm formation and persistence include the following:

1. Biasing or preconditioning of the substratum either by macromolecules present in the bulk liquid or, intentionally, as in the case of precoating biomedical materials with adhesion molecules (e.g., fibronectin, vitronectin, fibrinogen, von Willebrand's factor).
2. Transport of planktonic cells from the bulk liquid to the substratum.
3. Adsorption of cells at the substratum for a finite time followed by desorption (release) of reversibly adsorbed cells.
4. Irreversible adsorption of bacterial cells at a surface.
5. Transport of substrates into the biofilm, substrate metabolism by the biofilm-bound cells, and transport of products out of the biofilm. These processes are accompanied by cellular growth, replication, and extracellular polymer production.
6. Biofilm removal can be either continuous, as in the case of biofilm detachment brought on by forces associated with either fluid flow or cell replication, or discrete, as in the case of biofilm sloughing.

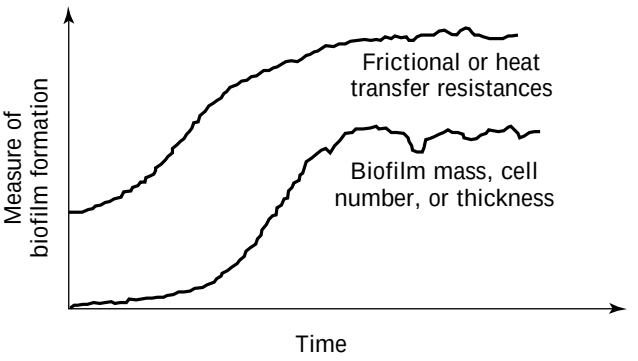


Figure 2. Net accumulation of biofilm and corresponding increases in frictional and heat transfer resistances.

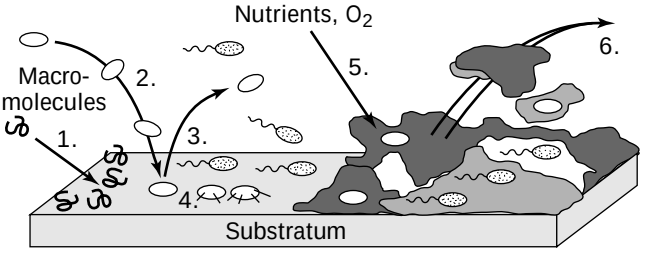


Figure 3. Schematic illustration of processes governing biofilm formation and persistence. (Processes are defined within the text.)

Each of these processes is reviewed in some detail.

Initial Events

Processes 1–4 are often grouped in mathematical models as a single process termed deposition.

Preconditioning of the Substratum. In the case of intentional pretreatment of a surface (i.e., adhesion molecules for biomaterials), there is no rate to be considered; the substratum essentially enters the system with a biased sur-

face chemistry. However, in cases where a clean substratum material is exposed to an aqueous environment, transport of dissolved organic molecules or macromolecules in laminar flow is basically by molecular diffusion; in turbulent flow, transport must also consider convective transport effects. Once at the surface, adsorption of macromolecules occurs almost instantaneously, immediately changing the surface chemistry of the exposed material. Loeb and Neihof (50) and Depalma et al. (51) measured adsorption rates of organic molecules to various solid substrata in seawater, and Bryers (52) reported analogous adsorption rates in a freshwater laboratory system. The net rate of adsorption in these studies can be described by equation 1.

$$R_{\text{adsorption}} = k_a S_i[1 - (S_i/k_s)] \quad (1)$$

where $R_{\text{adsorption}}$ is the rate of net adsorption ($ML^{-2}t^{-1}$); k_a is the adsorption rate constant (Lt^{-1}); k_S is the surface saturation coefficient (ML^{-2}); and S_i is the real concentration of adsorbed species (ML^{-2}). Although experiments indicate the maximum amount of adsorbed material may not exceed 0.10 μm in thickness, the surface properties resulting from adsorption of an organic film can significantly bias subsequent microbial events.

Cellular Adsorption and Desorption. The net deposition rate of cells upon a "conditioned" substratum is itself a composite of the individual rates of cell transport to the substratum, cell adsorption, and adherent cell desorption.

Many models exist to describe bacterial transport from the fluid phase to the target surface. Bowen et al. (53) and Beal (54) derive continuity equations describing the transport of suspended particles from a flowing fluid to the surfaces of a surrounding conduit that accounts for both convective and molecular transport mechanisms. From their work, the maximum flux of particles transported to the surface, L distance from the inlet, of a rectangular flow cell, with gap height $2h$, can be written as

$$J(\text{cells-}L^{-2}t^{-1}) = k_{tr} X \mathcal{D}_x X/h \quad (2)$$

where k_{tr} is the particle transport coefficient $((2/9)K^{0.33}/\Gamma(4/3))$; $K = (1/Pe)(8L/3h)$; Pe is the Peclet number $= 4vh/\mathcal{D}_X$; v is the average velocity of the fluid; Γ is the gamma function ($\Gamma(4/3) = 0.89338$); and $\mathcal{D}_X$ is the diffusivity of cells in the liquid. For nonmotile cells, the Brownian diffusion coefficient can be estimated from the Stokes-Einstein equation,

$$\mathcal{D}_x X = k_b T/3\pi\mu d_c \quad (3)$$

where k_b is the Boltzmann constant (1.38×10^{-23} J/K); T is absolute temperature; μ is absolute viscosity; d_c is cell diameter. To compensate for cell motility, Jang and Yen (55) propose the following equation,

$$\mathcal{D}_x X = v_m d_r/3(1 - \cos\sigma) \quad (4)$$

where v_m is the velocity of motility, d_r is the free length of a random run, and σ is the angle turned by the motile cell.

Once at the substratum interface, cells stick. Observations have indicated that cells can adsorb to a surface for a period of time and then may desorb from the substratum, returning to the bulk liquid. If cells spend sufficient time at the surface, they become permanently bound, initiate within minutes the secretion of extracellular polymer production, and can be removed only by rather aggressive physical or chemical means. One can conceive of this overall process as much like two reaction steps in series, with the reversible adsorption process followed by an irreversible step pertaining to permanent adsorption,

$$X \leftrightarrow X_{rev} \rightarrow X_{irr} \quad (5)$$

where X is the suspended cells, X_{rev} is the reversibly adsorbed cells, and X_{irr} is the irreversibly adsorbed cells. Assuming no cell replication, the total amount of cells observed at a substratum, X_t, at any one time comprises both reversibly and permanently adsorbed cells, that is, $X_t = X_{rev} + X_{irr}$.

Cell Attachment. Attachment is the deposition of cells from the fluid phase onto an existing biofilm. This process differs from adsorption in that the target surface is now a developed biofilm. The biofilm itself poses a different type of substratum in that biofilm surface morphology is generally more irregular (perhaps filamentous), very porous, compliant, and gelatinous. Experimental results indicate that the increased attachment of cells to a biofilm-covered substratum may depend on far more subtle mechanisms than mere changes in surface morphology. The local chemistry of a biofilm surface will reflect the composition of the bacterial cell and that of the insoluble polymer matrix, both of which could affect subsequent chemical interactions with incoming cells.

Banks (56) and Banks and Bryers (57) report that the rates of cell deposition increased when cells were exposed to biofilm surfaces versus clean glass (Fig. 4). Two species were investigated, *Pseudomonas putida* and *Hyphomicrobium ZV620*, as to their deposition rates, under laminar flow conditions, onto both clean glass and biofilm-covered surfaces. *Pseudomonas putida* cells attached to a *P. putida* biofilm at a rate (zero order in attached cell concentration) that was 3.6 times higher than the cell attachment rate to clean glass. *Pseudomonas putida* cells attached to a *Hyphomicrobium* biofilm (with zero-order kinetics in attached cell concentration) that was 2.4 times higher than to clean glass. *Hyphomicrobium* illustrated an attachment rate to a *Hyphomicrobium* biofilm (again a zero dependency on attached cells) that was 3.3 times higher than the *Hyphomicrobium* attachment rate to clean glass. Attachment rate of *Hyphomicrobium* spp. cells to a *P. putida* biofilm was only 1.3 times that to clean glass. The attachment rate of *Hyphomicrobium* cells to a *Hyphomicrobium* biofilm was about half of the rate of *Pseudomonas* cells attaching to a *Pseudomonas* biofilm, even though suspended *Pseudomonas* cell concentrations were three times higher. Results insinuate that species-dependent enhancement of cell attachment may occur by specific rather than nonspecific adhesion mechanisms.

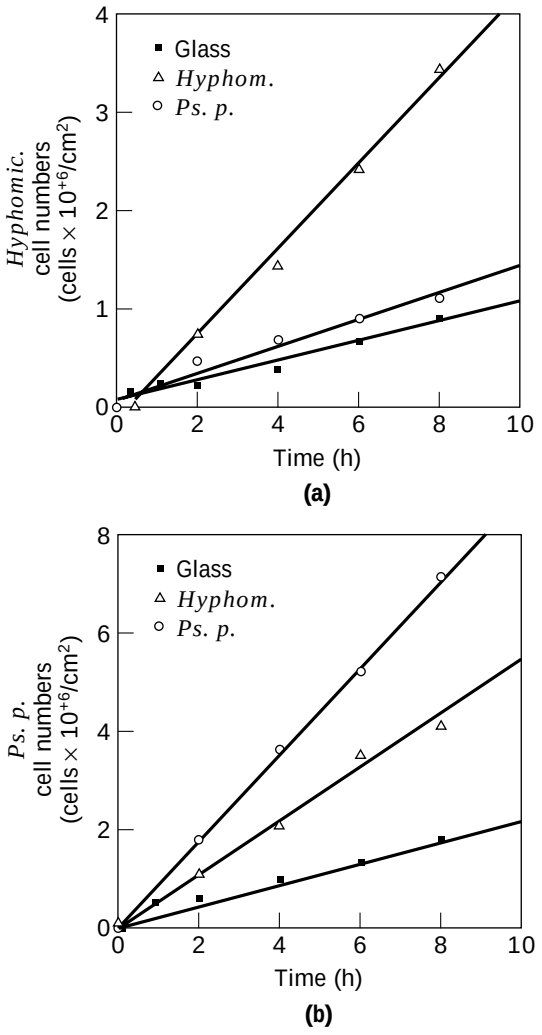


Figure 4. (**a**) Attachment of *Hyphomicrobium* spp. to various surfaces. Suspended cell concentration is 4.5E + 07 cells/mL. (**b**) Attachment of *Pseudomonas putida* to various surfaces. Suspended cell concentration is 1.5E + 08 cells/mL (57).

Biofilm Cell Metabolic Processes

Transformation collectively refers to those processes in which molecular rearrangements occur, that is, reactions. Once a cell adsorbs to a surface or becomes embedded within an existing biofilm, it will continue its metabolic processes in response to its immediate environment. Three fundamental rate processes can be identified: *(1)* alginate or exopolymer secretion, *(2)* cellular growth and replication, and *(3)* cell death and lysis. Prior to the mid-1990s; it was often easier to measure "per reactor area" performance using observed rates determined from bulk liquid parameters (e.g., electron donor removal rate or electron acceptor uptake rate) than to directly measure growth, replication, and death of cells in a biofilm. This experimental limitation led to the development of a number of unstructured mathematical models that were useful in estimating the flux of growth-rate-limiting substrate into a biofilm of fixed thickness, density, and reactivity, taking into account both external and internal molecular diffusion and biological reaction. It was not until the advent of various molecular methods and advances in computer-aided microscopy that it was possible to determine various biofilm processes directly within the biofilm.

Adhesion-Induced Alginate Synthesis. Reporter gene technology has been used to observe the regulation of the alginate biosynthesis gene, *algC*, in a mucoid strain of *P. aeruginosa* in developing and mature biofilms (58). In vivo detection of *algC* up-expression in developing biofilms was performed with a fluorogenic substrate for the plasmid-borne *lacZ* gene product (*β*-galactosidase) using microscopy coupled with image analysis. By this technique, cells were tracked over time and analyzed for *algC* activity. During the initial stages of biofilm development, cells attached to a glass surface for at least 15 min exhibited up-expression of *algC*, detectable as the development of whole-cell fluorescence. However, initial cell attachment to the substratum appeared to be independent of *algC* promoter activity. Furthermore, cells not exhibiting *algC* up-expression were shown to be less capable of remaining at a glass surface under flowing conditions than were cells in which *algC* up-expression was detected.

Cell Growth and Replication. The major transformation carried out by cells in the biofilm is the metabolism of both an electron donor and a terminal electron acceptor to produce soluble by-products, extracellular polymers, carbon dioxide, and water. Depending on the microbial population in question and the ambient concentration of electron donor and acceptor, a biofilm can be either aerobic, anoxic (denitrifying), anaerobic (sulfate-reducing bacteria, methane formers), or fermentative. Analysis of biofilm bacterial metabolic rates are frequently complicated by the effects of significant mass transfer resistances in both the liquid phase and within the developing biofilm. Steady-state mathematical models (59,60) were derived in the late 1970s to estimate the observed flux of growth-rate-limiting substrate into a biofilm of fixed thickness, density, and reactivity. These steady-state biofilm models are based on assumptions of a constant biofilm concentration (tacitly implying a spatially uniform reactivity) and a constant diffusion path (biofilm thickness). Such models allow one to predict *(1)* the concentration profile of limiting substrate (and, by stoichiometry, all other nutrients) with biofilm depth, and *(2)* the maximum substrate uptake or flux to the biofilm.

Biofilm cellular growth rates are typically simulated using a classical microbial growth rate expression, R_g, that is related by a stoichiometric coefficient, Y (in units of cell mass per substrate mass), to the rate of utilization of growth-limiting substrate, R_S ($M_{substrate}\, L^{-2} t^{-1}$), that is,

$$R_g(M_{cells} L^{-2} t^{-1}) = -R_S Y \qquad (6)$$

where the rate of substrate utilization is typically a Monod-like dependency on the concentration of one or more growth-rate-limiting substrates. For the example of a single substrate limitation,

$$R_S = \mu^* SB/(K_S + S) \quad (7)$$

where μ^* is the species maximum growth rate constant (t^{-1}); S is the local dissolved substrate concentration ($M_{substrate}L^{-3}$); K_S is the saturation kinetics half-saturation constant ($M_{substrate}L^{-3}$); and B is the local biofilm cell concentration ($M_{cell}L^{-2}$).

Some differences in biofilm kinetics versus typical suspended cell microbial kinetics include the following:

- The concentration of active bacterial cell mass metabolizing the substrate, B, is a per area concentration and may also be a function of time and space, $B = B(t, z)$. In a mixed culture biofilm, each bacterial species could exhibit a different dependency on time and space.
- The concentration of substrate S in the biofilm will be affected by mass transfer resistances and biological reaction rates, such that S may also be a function of time and space.
- Endogenous decay, maintenance energy, or death terms for cells not receiving adequate nutrients are treated as simple first-order functions of the biofilm population, for example,

$$R_{death} = k_d B \quad (8)$$

where k_d is the cell death specific rate constant (t^{-1}). Thus the local, instantaneous net production of a species biomass in the biofilm can be modeled as

$$R_g(t, z) = [\mu^* S(t, z)B(t, z)/Y(K_s + S(t, z))] - k_d B(t, z) \quad (9)$$

Biofilm Removal Processes

Biofilm removal processes involve the actual loss of material from the biofilm—either cells or cell–biofilm matrix debris. Desorption (see "Deposition Processes") is the release of cells from the substratum and occurs during the early stages of substratum colonization. Biofilm removal represents loss of biomass from a developing or mature biofilm that arises by way of two distinctively different mechanisms: detachment and sloughing.

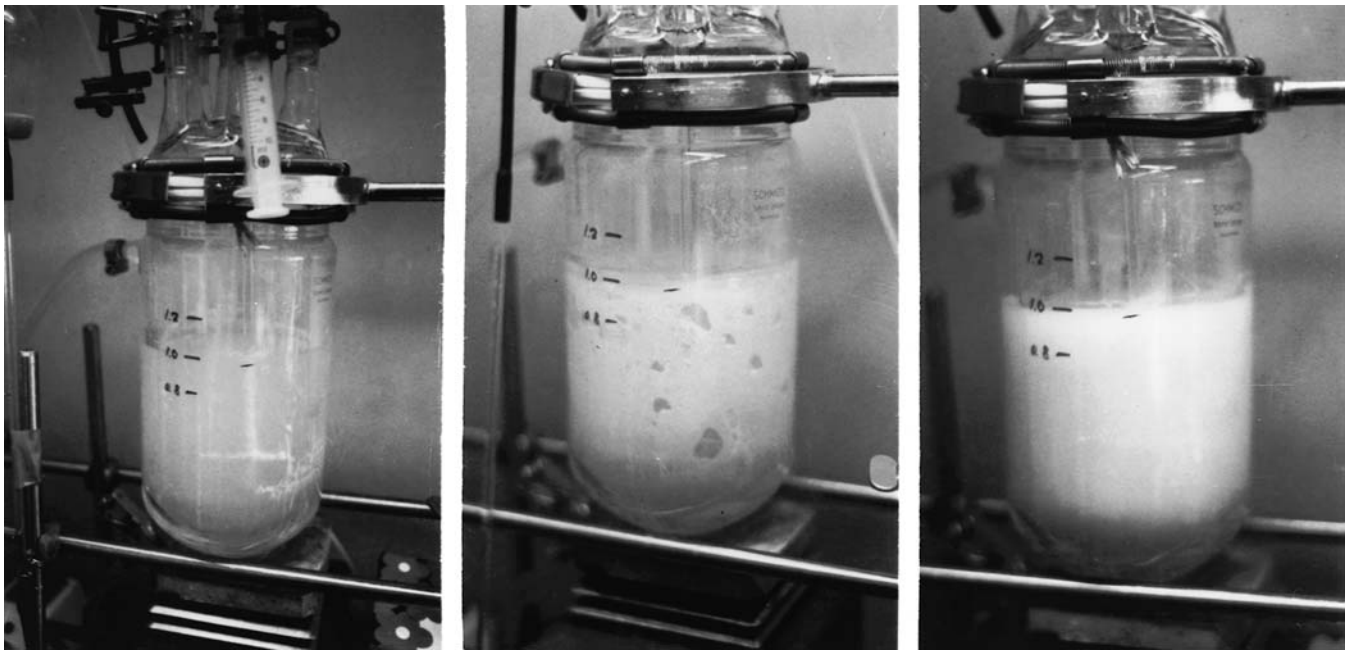
Detachment. Detachment is the continuous loss of biomass or individual cells from a biofilm. The rate of biofilm detachment has been mathematically modeled in a number of ways (61). In about the only definitive experimental study on biofilm detachment, Applegate and Bryers (62) reported that the growth conditions of the biofilm strongly affect biofilm removal processes. Carbon-substrate-limited growth conditions fostered biofilms that *(1)* contained less extracellular polymer per cell and per area, *(2)* bound less calcium per unit amount of biofilm, and *(3)* exhibited a higher detachment rate than oxygen-substrate-limited biofilms that accumulated more extracellular polymer per cell, bound more calcium per biofilm amount, and developed a rigid compact morphology that was very resistant to detachment prior to predictable sloughing. Carbon-limited biofilms did not exhibit sloughing for extended periods of cultivation or after prolonged (300 h) of carbon-nutrient deprivation.

Perhaps biofilm detachment processes may be controlled, not by fluid hydrodynamics as once imagined, but rather by cell physiology. Boyd and Chakrabarty (63) suggest biofilm detachment could be genetically controlled. Their work showed that the exopolysaccharide alginate of *P. aeruginosa* was important in determining the degree of cell detachment from an agar surface. Nonmucold strain 8822 gave rise to 50-fold more sloughed cells than mucoid strains 8821 and 8830. Alginate anchors the bacteria to the surface, thereby influencing the extent of detachment. The role of the *P. aeruginosa* alginate lyase in the process of cell detachment was investigated. Increased expression of the alginate lyase in mucoid strain 8830 led to alginate degradation and increased cell detachment. Similar effects were seen both when alginate lyase synthesis was induced at the initial stage of cell inoculation and when it was induced at a later stage of growth. Apparently, high molecular weight alginate polymers are required to efficiently retain the bacteria within the developing biofilm. When expressed from a regulated promoter, the alginate lyase did enhance detachment of cells through the degradation of the alginate. This suggests a possible role for the lyase in the removal process of bacterial growth films.

Sloughing. The catastrophic, apparently random, loss of large pieces or entire sections of biofilm is known as sloughing (Fig. 5). At the industrial scale, sloughing creates excursions of intolerably high biomass in bioreactor effluents, and periodic "blooms" of pathogenic bacteria in drinking water distribution systems. At the laboratory scale, all too often, random sloughing of a biofilm will terminate an experiment, resulting in very little quantitative data regarding the fundamental causes of sloughing. A number of situation-specific causes of sloughing have been identified: bubble formation in either anaerobic methane-producing biofilms or denitrifying biofilms, and artificially induced sloughing by way of EGTA addition, which chelates the bound calcium ion used in biofilm alginate cross-linking. Applegate and Bryers (62) reports that the growth conditions of the biofilm strongly affected the biofilm removal processes of detachment and sloughing. "Fluffier" biofilms, produced under carbon-limited growth, exhibited high detachment rates but never sloughed, even when subjected to over 300 h of nutrient starvation. Conversely, rigid biofilms cultivated under oxygen limitations showed little detachment, but the onset of catastrophic sloughing was repeatable between replicate experiments.

Net Accumulation of Biofilm

As shown in Figure 1, development of a bacterial biofilm comprises the following fundamental processes: *(1)* substratum preconditioning by adsorption of fluid-phase organic molecules; *(2)* bacterial cell deposition to the conditioned substratum (cell transport to the surface, cell adhesion to the substratum, cell desorption); *(3)* bacterial metabolism (cell substrate conversion; cell growth and replication; extracellular exopolymer production; cell starvation, death, lysis); *(4)* biofilm removal (cell and biofilm de-



(a) **(b)** **(c)**

Figure 5. Sequence of biofilm formation (**a**), onset of partial biofilm sloughing (**b**), and complete biofilm sloughing (**c**) of an anaerobic biofilm. *Source*: Ref. 3.

tachment, biofilm sloughing). Naturally, the relative influence of each process to the overall rate of accumulation is dependent upon the specific system, the prevailing environmental conditions, and biological changes throughout the lifetime of the biofilm.

Simple unstructured models of biofilm formation can be written. Biofilm formation is assumed to occur on surfaces exposed to a well-mixed bulk fluid phase, thus tacitly neglecting any concentration gradients in the bulk liquid. Any mass transfer limitations in the bulk liquid will only complicate the already difficult task of determining the kinetics of a heterogeneous reaction. Suspended biomass in the bulk liquid arises due to either cell growth and replication or to detachment of biofilm material. Planktonic cells leave the liquid phase by either the effluent liquid leaving the reactor volume or by cell deposition onto the reactor surfaces. A single, sterile growth-limiting substrate, S, enters the reaction volume, where it is consumed by either the suspended or biofilm-bound cells. Any reasonable measure of both biofilm and suspended cell concentrations is acceptable (cell number, biomass dry weight, biomass organic carbon). However, to complete material balances over the entire system, one most know the stoichiometric relationship between changes in the suspended and attached cell concentrations due to growth and the limiting substrate utilized.

Based on these assumptions, material balances for both the suspended cell biomass and single growth-limiting substrate can be written as equations 10 and 11:

Suspended biomass balance: $VdX/dt =$

$$-FX + (R_g V)_{\text{suspended}} + R_{\text{det}}A - R_{\text{dep}}A \quad (10)$$

Limiting substrate balance: $VdS/dt = F(S_{\text{in}} - S)$

$$- (R_g V/Y)_{\text{suspended}} - (R_g A/Y)_{\text{biofilm}} \quad (11)$$

where X is the suspended cell biomass (ML^{-3}); S is the limiting substrate concentration (ML^{-3}); V is the reaction volume (L^3); $(R_g)_{\text{suspended}}$ is the rate of suspended cell growth ($M_{\text{biomass}}L^{-3}t^{-1}$); R_{det} is the rate of biofilm detachment ($M_{\text{biomass}}L^{-2}t^{-1}$); R_{dep} is the net rate of bacterial cell deposition at a substratum ($M_{\text{biomass}}L^{-2}t^{-1}$); $(R_g)_{\text{biofilm}}$ is the rate of biofilm cell growth ($M_{\text{biomass}}L^{-2}t^{-1}$); Y is the yield coefficient ($M_{\text{biomass}}/M_{\text{substrate}}$); and A is the reaction area (L^2). Rate expressions for the process rates (R_g, R_{det}, R_{dep}) may depend directly on such changing parameters as biofilm amount, limiting substrate, and suspended biomass concentration.

Biofilm net accumulation can be described by equation 12:

Biofilm net accumulation:

$$A(dB/dt = R_{\text{dep}} + R_{\text{gb}} - R_{\text{det}}) \quad (12)$$

Because rate expressions for both planktonic and biofilm growth are nonlinear saturation kinetic functions of the instantaneous substrate concentration and first-order functions of X and B, respectively, equations 10–12 are coupled, nonlinear ordinary differential equations that require either simultaneous numerical integration (for the transient situation) or simultaneous algebraic solution under steady-state conditions (i.e., dS/dt, dX/dt, and $dB/dt = 0$). Numerous biofilm accumulation studies have been carried out that can be mathematically described by this set of equations.

Should the reactor be operated at a hydraulic residence time shorter than the generation time of the cells, then one can assume the term ($R_{gs}V$) is negligible; however, the corresponding term in equation 11, representing suspended cell substrate metabolism ($R_{gs}V/Y_{X/S}$), cannot be disregarded unless in the specific system this term is determined to be small compared to the other terms, or substrate is not supplied.

In a simple bacterial adhesion study, experiments can be carried out without exogenous substrate, thus the R_g terms in equations 10–12 could be ignored. Thus, equation 12 becomes

$$dB/dt = R_{dep} - R_{det} \quad (13)$$

and is valid provided the adherent bacterial replication is zero.

RECENT ADVANCES IN BIOFILM RESEARCH

This section highlights recent research advances in understanding biofilm phenomena, which include advances in biofilm experimental techniques, new findings regarding mass transport mechanisms in biofilm, and genetic processes within biofilm communities.

Advances in Biofilm Experimental Techniques

All biofilm experimental systems should contain certain critical components: *(1)* the biofilm study reactor itself (where the surface accumulation of cells and biofilm will be studied) and *(2)* the environmental support system that controls the experimental conditions (e.g., temperature, flow velocity, nutrient concentrations, pH) of the bulk fluid. Under quiescent conditions, the rate of transport of dissolved nutrients and particulates (i.e., the bacteria) to the substratum may be the slowest in a series of rate processes. For example, particle transport may limit the rate of cell deposition in quiescent conditions. Thus, it is recommended that only experimental systems that create a gradient-less, well-mixed bulk fluid phase be employed. Otherwise, instead of determining the true kinetics of whatever biofilm process is of interest, the resultant data will erroneously reflect mass transfer limitations.

Invasive Biofilm Diagnosis. Detection of biofilms often requires direct sampling and removal of a finite quantity of biofilm from a reactor surface to be used in a number of destructive analytical procedures that may measure *(1)* overall biofilm amount (cell mass or biofilm thickness), *(2)* a component of the biofilm (biofilm carbon, cell number, biofilm nitrogen content), or *(3)* a biofilm cell component or cellular activity (ATP, cell protein, dehydrogenase activity, active DNA synthesis, species-specific 16S rRNA). Reactor designs that provide access for destructive sampling involve removable sections of substratum and are quite numerous. The reader is directed to a number of excellent reviews for further details (5,7). One drawback to destructive analyses is that the parameter determined represents an average over the entire sample area, which ignores any spatial variations in that parameter that may exist.

Almost any parameter determined by a molecular or chemical assay (ATP, DNA, total protein, biofilm polysaccharide content, 16S rRNA, cellular phospholipid signatures, specific enzyme levels, and oligonucleotide probes), cytochemical (acridine orange, DAPI stain, and Hoechst DNA stains) or immunofluorescent staining, or cellular activity (CTC stains, INT stains, and plate counting) techniques that has been applied to planktonic (freely suspended) cells can be used to quantify changes in biofilm on a substratum. Rodriguez et al. (64) reports on the potential of a fluorescent redox probe, 5-cyano-2,3-ditolyltetrazolium chloride (CTC), for direct observation of actively respiring bacteria. Oxidized CTC is colorless, but upon reduction by electron transport activity, the insoluble CTC-formazan fluoresces at 365 nm upon excitation. Respiring marine bacteria within thin (10–50 μm) biofilms attached to optically nontransparent polysulfone substrata were rapidly and easily enumerated. Respiring cell counts via CTC were also useful in directly assessing the efficacy of various biocidal control agents.

The advent of modern molecular biology has provided many useful diagnostic methods with which to analyze biofilm ecology. For example, Amann et al. (65) characterize the population structure of sulfidogenic biofilms, established in anaerobic bioreactors, by selective polymerase chain reaction (PCR) amplification and fluorescent microscopy. 16S rRNA common to the sulfate-reducing bacteria in the biofilm was selectively amplified by PCR and used to design both general and specific fluorescent hybridization probes. Biofilms, (5–10 μm thick) on glass cover slips immersed within laboratory anaerobic biofilm reactors, were fixed in formaldehyde and dehydrated prior to hybridization with the RNA probes. Hybridized biofilm samples were then viewed under epifluorescent microscopy.

Rogers and Keevil (66) report on the destructive sampling of a multispecies biofilm, intentionally inoculated with *Legionella pneumophilia*, followed by both immunogold and fluorescein immunolabeling for this species. Episcopic differential interference contrast microscopy was employed to simultaneously visualize the total biofilm community and the labeled *Legionella* species. This techniques was able to provide observations of variations in the biofilm community from one position to another (parallel to the substratum) but was unable to discern such variations with depth (perpendicular to the substratum).

McCarter et al. (67) report a very elegant series of genetic techniques to investigate bacterial responses to various stimuli, one being the adhesion to a surface. Using a transposon (mini-Mu*lux*), the authors were able to incorporate recombinant reporter gene insertions adjacent to target gene promoter sites that encode for light production as a function of target gene expression. Light production is conveniently measured by exposure of X-ray or photographic film, visual examination, chemiluminescence, or photometry. Such techniques also obviate the need to disrupt biofilms prior to quantification. McCarter et al. employed a *lux* reporter gene adjacent to the chromosomal gene encoding for swarmer cell differentiation that *Vibria parahaemolyticus* experiences upon association with a solid surface. Resultant mutants thus emit light upon growth at a surface but not in suspension. Dagostino et al.

(68) also employed transposon mutagenesis to insert into an appropriate recipient bacteria a marker gene that lacks its own promoter. The premise is that if a suitable target gene is "on" at the surface, then expression of the marker will be observed only in the presence of the surface. Dagostino et al. employed *Escherichia coli* C600 (pRK2013: mini-MuTetr *lacZ*) as donor and two marine bacteria as recipients. The authors were able to successfully isolate a transposon-generated mutant in which the *lacZ* gene was not expressed in either liquid or agar, but was expressed when grown on a polystyrene substratum.

Noninvasive Biofilm Diagnosis. The "holy grail" of biofilm research is a diagnostic method that creates as little disturbance to the biofilm and the reactor operation as possible. Regrettably, most parameters of interest in cell adhesion and biofilm accumulation require destructive sampling and analyses.

True noninvasive diagnostics of biofilm accumulation, amount, and biofilm-cell reactivity are limited to relatively few sophisticated techniques. The simplest of these techniques is on-line microscopy, which places the biofilm surface within the viewing field of a microscope. This technique requires a reactor dimension compatible with the microscope, and a flat reactor surface on which to focus. Microscopic observation of adherent cells can be quantified and preserved by use of video recorders or image analyzer systems that digitize observed images and save those images on a computer for future numerical interpretations.

For about 14 years, confocal scanning laser microscopy (CSLM) has been a tool in the biological sciences. CSLM uses confocal apertures, or pinholes, to create a thin (0.4 μm) depth of field, which eliminates out-of-focus light emitted from the laser-excited fluorophores. Laser light sources provide the intense, highly coherent, collimated light necessary to penetrate thick specimens. The laser light is used to excite fluorophores—either those intrinsic (e.g., chlorophyll) to the sample, or selected chemical or immunological stains intentionally applied to the sample. Resultant fluorescence that passes through the aperture is detected by a photomultiplier, and a digital image is collected. Using a computer-manipulated stage, optical sections can be collected from a specimen in three dimensions. This allows, for the first time, the ability to digitally locate cells on a substratum or within a biofilm, noninvasively (69). Numerous fluorochromes are available that can be used to localize and measure intracellular and extracellular conditions in three dimensions within living biofilms.

Camper et al. (70) detail image analysis software that can be used to collect images of bacterial adhesion patterns from the CSLM, digitize those images, then superimpose those images onto digitized images of the substratum itself, collected from any one of a number of other analytical instruments (e.g., X-ray photoelectron spectroscopy, atomic force microscopy, time-of-flight secondary ion mass spectroscopy); all images are collected at the same location on the substratum. Thus, one can now correlate the adhesion of a single cell to the local topographical or chemical properties of the substratum directly underneath the cell.

Lawrence et al. (71) illustrates the ability to optically section fully hydrated biofilms, both horizontally and sagitally, producing optical sections of undisturbed biofilms with a spatial thickness of 2 μm. Any fluorescent probe (cellular DNA stain; immunofluorescent stain; pH-, ion-, and redox-sensitive stains; viability redox stains) or any combination of probes can be simultaneously recorded by the CLSM with great clarity and little background interference, due to the confocal exclusion of any fluorescence originating from excited fluorochrome above and below the focal point (Fig. 6). CLSM is a critical tool in population analysis of biofilms consisting of mixed strains, mixed species, and mixed cell lines and will allow detailed examination of the relationships between biofilm structure, adaptation, reactivity, and response to external stress (72).

Another noninvasive technique that can be used to collect information on the molecular chemistry of cell adhesion at the substratum–fluid interface is Fourier transform infrared spectroscopy (FT-IR) coupled with attenuated total reflectance (ATR) waveguides integrated within flow cells. Atoms and groups of atoms within a molecule vibrate with a characteristic frequency and will adsorb light at those frequencies. Light that contains infrared frequencies can be focused on a molecule, and the amount of light adsorbed is then measured as a function of frequency. Specific IR adsorptions can thus be assigned to particular bonds, and alterations in these bonds due to changes in local environment can be assessed from the resultant spectral details. Recent increases in the capability of FT-IRs and the focusing of the IR wave within specific crystal wave guides (germanium or zinc selenide crystals) allows one to establish a standing IR evanescent wave at the surface of the crystal. Fourier transform signal processing and multiple scanning allow aqueous samples to be processed where the IR spectra of molecules directly adjacent to the waveguide can be collected. In biological systems, the IR absorbance of water must be subtracted via computer manipulations to provide the IR spectra of molecules accumulating at the crystal surface. The effective depth of penetration for the evanescent IR wave is a function of crystal material and the wavenumber but ranges around 0.3–0.7 μm, thus providing spectra that reflect those chemical species directly adjacent to the surface.

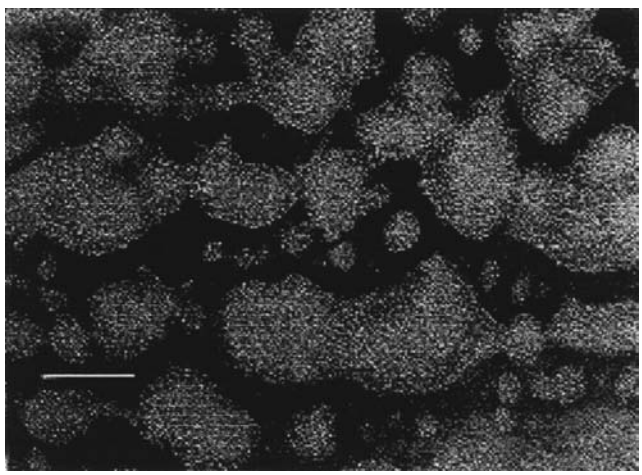


Figure 6. Confocal laser microscope optical section of a biofilm.

ATR/FT-IR has been used to study the effects of protein-conditioning films on alginate adsorption to the germanium surface (73). Alginate adsorbed in greater quantities to protein-coated ATR crystals than to uncoated ones. For the development of biofilms on ATR crystals, the observed spectrum is the integral of all biomolecules on the surface and over the entire surface. Bremer and Geesey (74) report the chemical changes as detected by FT-IR that occur upon inoculating with a microbial culture, at a germanium ATR crystal situated within a flow cell. Spectral intensities representing various bonds within proteins (amide I at 1,645/cm; amide II at 1,550/cm) and polysaccharide (C-O stretches at 1,058/cm) are seen to accumulate with time, indicating the feasibility of the ATR/FT-IR system to detect biofilm formation. Because the IR wave can only penetrate less than 1 μm into the biofilm, reported increases in absorbance intensities most likely represent increases in cell surface coverage of the crystal and not increases in biofilm thickness. A plot of attached cells per area versus the increases in area under the amide I peak was generated to estimate a lower limit of detection by FT-IR of $\sim 5 \times 10^5$ cells/cm^2. In a medical application, ATR/FT-IR has been used to follow the penetration of an antibiotic (100 μg/mL ciprofloxacin) into the IR wave region underneath a *P. aeruginosa* biofilm (75).

Several reports exist of the use of invasive but nondestructive analysis of solute concentration profiles, by microsensor chemical probes, as a function of spatial dimension within the developing biofilm. Microsensors (tip size ≤ 15 μm) exist that can detect various dissolved solutes including glucose, oxygen, pH, sulfide, and ammonia. DeBeer et al. (76) report using microoxygen sensors (tip size = 15 μm) to estimate oxygen profiles, local oxygen uptake rates, and oxygen diffusion coefficients in a mixed-culture biofilm community cultivated within a laboratory biofilm reactor. Schramm et al. (77) determined microprofiles of oxygen and nitrate in nitrifying biofilms from a trickling filter and correlated these solute profiles with bacterial population profiles for *Nitrosomonas* and *Nitrobacter* spp., as determined using fluorescence in situ hybridization of cells fixed with 16S rRNA oligonucleotide probes.

Transport Phenomena in Biofilms

Bacterial biofilms are heterogeneous reaction systems wherein limiting substrates, essential nutrients, cellular antagonists, and cellular by-products are exchanged between the surrounding fluid and the site of the reaction (i.e., the cell). In numerous circumstances, the performance of biofilm processes can be limited by both internal and external mass transfer rates (e.g., antibiotic treatment of a biofilm infection, biocide challenge to eliminate biofilm formation in heat exchange systems, biofilm reactor treatment of wastewaters or waste gases, and in situ hazardous waste decontamination). Recent advances in characterizing biofilm architecture have led to the need to reformulate concepts of the mass transport mechanism in biofilms.

For about 25 years, estimates of substrate consumption for biofilms of known thickness have been based on the steady-state solution of equation 14,

$$\mathcal{D}_{\text{eff}} d^2C/dx^2 = \upsilon R_c \tag{14}$$

where C is the solute concentration, R_c is all reactions transforming solute C, υ is the stoichiometric coefficient for the transformation, x is the coordinate of molecular transport, and $\mathcal{D}_{\text{eff}}$ is the effective diffusivity of solute in the biofilm. These various models are based on the assumption that transport of reactants (i.e., substrates) into the biofilm and export of products is by molecular diffusion alone and that bulk transport (convection) does not occur within the biofilm. These models also do not account for *(1)* any heterogeneities in the biofilm gel structure, *(2)* any locally dependent bacterial concentrations, or *(3)* any nonlinearity in diffusion coefficients.

Traditionally, either $\mathcal{D}_{\text{eff}}$ is estimated as the solute diffusivity in pure water, reduced by an arbitrary fraction, or $\mathcal{D}_{\text{eff}}$ is determined for the particular biofilm system in question from diffusion experiments. For small ions and uncharged small organic compounds, both Libicki et al. (78) and Siegrist and Gujer (79) indicate that transport in thin microbial layers is indeed by molecular diffusion. Both reports indicate that the decrease in observed diffusivity values relative to those in pure water is a function of the microbial aggregate cell-to-polymer composition. Siegrist and Gujer (79) provided the first report to insinuate that molecular diffusion may not be the only mass transport process acting within a biofilm and that convective mass transport may predominate in the upper layers of thick, filamentous biofilms.

Results of Drummond (80) and Bryers and Drummond (81) suggest these dense alginate clusters are themselves porous, with a wide distribution of pore sizes ranging from angstroms to tenths of micrometers. Their results have shown the transport of larger solutes (i.e., macromolecules) through the hydrated biofilm polysaccharide gel matrix is controlled by transport mechanisms far more complex than molecular diffusion. Overall effective diffusion coefficients for various molecules, increasing in molecular weight and differing in molecular size and stereochemistry, were determined using half-cell diffusion chambers. Transient diffusion studies were carried out at constant temperature and pH using bacterial biofilms of increasing thickness. Half-cell diffusion results were considered erroneous due to the "perforated" nature of the biofilms.

The existence of water channels in these biofilms was documented and found to significantly affected mass transfer coefficient estimates obtained from half-cell studies. Consequently, a local microtransport technique was developed based on the fluorescence recovery after photobleaching (FRAP) technique of Poitevin and Wahl (82). FRAP is a local, instantaneous, and noninvasive method that can be repeatedly carried out over a wide expanse of biofilm. Corrections of the FRAP "tracer" diffusion coefficients must be made in order to estimate "Fickian" diffusivities. Diffusion coefficients determined for large macromolecule solutes (dextrans, proteins, DNA fragments) moving in biofilms did not correlate simply with increasing molecular weight except within the homologous chemical groups. Transport of macromolecules through the biofilm was affected not only by molecular weight but by the size or length of the molecule in solution. Using a computer-

actuated laser and motorized microscope stage, FRAP analysis across a large section of biofilm illustrates that local mass transfer coefficients can vary significantly in a biofilm over very short distances (Fig. 7).

Thus, research reported by Drummond (80) and Bryers and Drummond (81) and independently corroborated by Stoodley et al. (84) has shown biofilms are not architecturally spatially continuous. Rather, biofilms are structurally nonuniform, consisting of dense microbial clusters surrounded by tortuous water channels that perforate the biofilm matrix. Thus, theory and experiments addressing overall mass transport within a real biofilm system must account for both convective transport and molecular diffusion.

Molecular Processes in Biofilm Communities

Plasmid Expression and Retention in Biofilm Cultures. Any exposure of plasmid-recombinant microorganisms to an open system environment, either inadvertently or intentionally, mandates research into those fundamental organism–plasmid processes that influence plasmid retention, transfer, and expression. Use of plasmid-recombinant bacteria to impart a desired phenotypic ability to an existing ecosystem has been desired for the bioremediation of hazardous wastes in both controlled reactor systems and in waste-contaminated open environments. In open environmental systems, a majority of the microbial activity occurs associated with an interface, within biofilms. In wastewater or waste gas treatment, efficient reactor design often dictates the use of immobilized, biofilm-bound bacterial communities. Consequently, the fate and ability to manifest a plasmid-borne phenotype within a biofilm community requires basic information on the biofilm's genetic processes.

A plasmid is deemed unstable if it either undergoes molecular rearrangement (structural instability) or is not absolutely inherited by progeny (segregational instability).

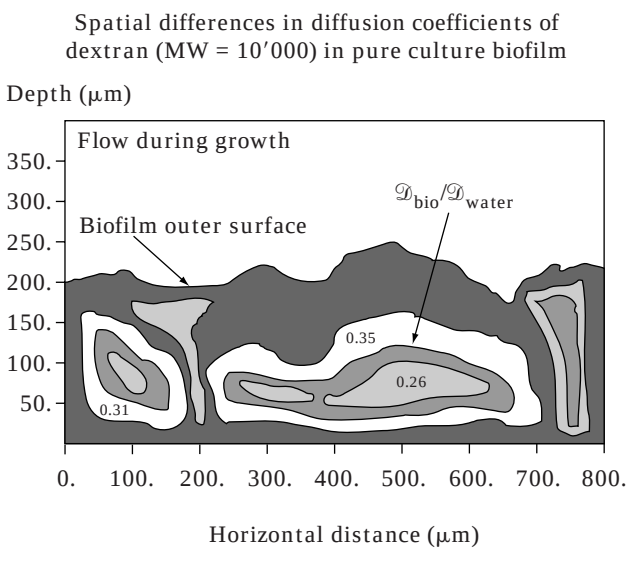


Figure 7. Spatial differences in diffusion coefficients of dextran (MW = 10,000) in pure culture biofilms. *Source*: Ref. 83.

Such phenomena can have significant effects on the outcome of cell cultivation processes. It has been established, both experimentally and theoretically, that plasmid maintenance and cloned gene expression can reduce the overall growth rate of the plasmid-bearing cell relative to the plasmid-free cell. Reduction of copy number and loss of plasmids from populations under continuous suspended culture have been reported in many cases, even in the presence of selective pressure.

That immobilization might stabilize a plasmid-bearing population can be shown mathematically (85,86). Plasmid persistence in suspended cultures has been observed in cases where the plasmid-bearing cell was at a growth rate disadvantage; this observation was directly attributed to biofilm formation and cell detachment from the biofilm (87). Improved plasmid stability upon artificial cell immobilization in κ-carrageenan gel beads has been reported by de Taxis du Poët et al. (88), Nasri et al. (89), and Sayadi et al. (90).

Huang et al. (91,92) cultivated biofilm cultures under controlled hydrodynamic conditions in a parallel-plate flow cell reactor using *E. coli* DH5α containing a recombinant plasmid either with (pTKW106) or without (pMJR1750) a segregational stability factor, *hok/sok*. Both plasmid constructs provided for the inducible expression of β-galactosidase. Results in these studies suggest, for these two plasmids, that the expression of the chromosomal gene for synthesis of the polysaccharide matrix of the biofilm commands a higher priority on cellular intermediates than the foreign genes. Consequently, a series of experiments were carried out to indirectly manipulate bacterial polysaccharide production in the biofilm by varying the nutrient carbon-to-nitrogen (C:N) ratio. Plasmid loss probability, total polysaccharide production, and ratios of total polysaccharide to total protein in biofilm cultures of *E. coli* DH5α (pMJR1750) all increased with increasing C:N ratios. Synthesis rates of total RNA, rRNA, and β-galactosidase-specific mRNA in *E. coli* DH5α (pTKW106) all increased after induction by IPTG, but levels of each parameter seen in biofilm cultures were 10–20 times lower than the same values observed in suspended cell cultures.

Cell–Cell Communication. Cell–cell communication involves a chemically unique molecule, produced biologically by one cell, that reacts with a set of specific receptors located at another cell's surface to subsequently alter cellular behavior. Two examples of cell–cell communication in higher life forms would be *(1)* the thousands of lectins that control embryonic development in plants and animals, and *(2)* signaling molecules, such as the pheromones, that control the behavior of adult insects.

Cell–cell communication in bacterial systems, both sessile and planktonic, has been recognized for years (93), but our understanding of the control mechanisms has entered a new era with the advent of various molecular diagnostic tools. Recent research has identified the specific chemical structures of various signaling molecules that manifest certain cellular responses. The majority of cell–cell communication studies have dealt with suspended bacteria (e.g., quorum sensing), but recent advances suggest that

certain structural architecture in a biofilm arises due to molecular signaling.

Bacterial swarming, floc formation, and biofilm formation or aggregation are multicellular, cell-concentration-dependent behavior that can be induced in response to the recognition of surfaces of certain viscosity. Cells differentiate into multinucleated, elongated, and hyperflagellated form and orient themselves lengthwise across the surface. For a number of species, the ability to differentiate into swarmer cells is critical to virulence, surface adhesion, and colonization.

Acylated homoserine lactones serve as one class of signal molecule in bacterial communication. These molecules and their derivatives have been linked to control of bioluminescence, Ti plasmid transfer, production of virulence factors, antibiotic resistance, and swarming motility (94). Higher life forms have evolved mechanisms to interfere with such signaling processes. For example, certain seaweeds and sea grasses produce halogenated furanones, structurally similar to acylated homoserine lactones, that interfere with the swarming process and have been shown effective as antifoulants and antimicrobials (95). Givskov et al. (96) illustrated that two different furanones derived from the seaweed *Delisea pulchra* could progressively inhibit and eliminate the swarming behavior of *Serratia liquefaciens* as concentrations were increased from 0 μg/mL to 100 μg/mL (concentrations far too low to affect the growth of the bacteria).

Davies et al. (unpublished results) employed homoserine lactone mutants of the classic PAO-1 strain of *P. aeruginosa* to determine if cell–cell communication plays any role in bacterial biofilm polysaccharide matrix formation. Wild strains and mutants were separately used to inoculate flow cells to observe cell adhesion to surfaces and subsequent formation of biofilms. Cells of the wild-type strain (PAO-1), retaining the ability to secrete homoserine lactone, adhered to surfaces avidly and formed complex biofilms comprising discrete microcolonies and well-developed water channels.

Cells of the double homoserine lactone mutant (JP2), which lacked the ability to synthesize either the *Pseudomomas* autoinducer, PAI, or the factor U homoserine lactone, did adhere to surfaces avidly but failed to form the complexly structured biofilms. These mutant cells did not form detectable amounts of exopolysaccharide matrix material, did not aggregate into discrete microcolonies, and did not develop the water channels characteristic of the wild-type biofilms. Cells of the double homoserine lactone mutant simply formed unstructured masses of cells that were tightly packed together. These unstructured adherent masses of cells could be easily removed from the colonized surfaces by simple washing with surfactants. Chemical analyses proved that these biofilm contained no detectable biofilm exopolysaccharide matrix material. When a cell-free supernatant from a wild-type biofilm (PAO-1), which contained both types of homoserine lactone molecule, was added to the bulk fluid delivered to the flow cell, adherent JP2 mutant cells gradually began to produce matrix material and form complex biofilms; 12 h after the addition of the homoserine lactone molecules from the wild-type supernatant, the JP2 mutant strain biofilm was observed to be identical to the wild-type biofilm strain. Such research could potentially lead to nontoxic control strategies for disrupting detrimental biofilms, which would be economically and environmentally significant.

Novel Biofilm Removal Strategies

One traditional approach to eliminate biofilm formation within an engineered system is to prevent cell adhesion. One such anti–cell adhesion approach used in the biomedical materials sector focuses on modification of the surface chemical properties of the substratum by various methods including photochemical coupling of benzephenone derivatives of polyethylene glycol, polyacrylamide, and poly vinyl pyrrolidone (97); passive adsorption of pluronic surfactants (copolymers of polyethylene oxide and polypropylene oxide) to polystyrene (98); incorporation of polyethylene oxide into the upper layers of polyethylene terephthalate by solvent swelling (99); and a series of neutral, anionic, and cationic surfactants adsorbed onto stainless steel or glass (100). Rather than prevent adhesion, one alternative control approach for biomedical materials is to retard surface microbial activity by incorporating within the substratum a slowly released antibiotic agent (20,101,106).

Remedial approaches to eliminate or eradicate fouling biofilms existing in nonmedical engineered systems consist of mechanical cleaning, materials or unit replacement, or chemical biocide challenges (107). One novel alternative to chemical antagonism of a biofilm has been to eliminate an nutrient essential for microbial growth (108). In the petrochemical, refining, and power industries, either oxidants (chlorine, ozone, bromides) or biocides are employed regularly to suppress biofouling of water-cooled heat exchangers; biocides prove mostly ineffective in that they inactivate the bacterial cells but, unlike oxidants, are incapable of eliminating the source of the system inefficiencies, the biofilm matrix.

One of the most innovative technologies to control detrimental biofilm formation was reported by Wood et al. (109). Substrata containing layers of copper and cobalt phthalocyanine catalysts were constructed for bacterial adhesion trials. These immobilized catalysts serve to promote the formation of active oxygen species from various oxidizing agents such as peroxide or persulfates. Biofilms on the treated catalytic surfaces were exposed to various concentrations of hydrogen peroxide. Generation of antibacterial oxidant concentrations right at the interface of the substrata aided in the almost complete removal of biofilm.

BIBLIOGRAPHY

1. H.C. Jones, I.L. Roth, and W.M. Sanders, III, *J. Bacteriol.* **99**, 316–325 (1969).
2. J.W. Costerton, in L.D. Sabath ed., *Int. Symp.* Pseudomonas aeruginosa, Huber, Berne, 1979, pp. 1–3.
3. J.D. Bryers, *Biotechnol. Progr.* **3**, 57–68 (1987).
4. B. Atkinson, H.W. Fowler, in T.K. Ghose, A. Fiechter, and N. Blakebrough eds., *Advances in Biochemical Engineering*, vol. 3, Springer-Verlag, Berlin, 1974, pp. 221–277.
5. W.G. Characklis and K.C. Marshall, *Biofilms*, Wiley-Interscience, New York, 1990.

6. S.A. Blenkinsopp and J.W. Costerton, *Trends Biotechnol.* **9**, 138–144 (1991).
7. J.D. Bryers, *Trends Biotechnol.* **9**, 422–426 (1991).
8. S. Deckena and K.H. Blotevogel, *Biofouling* **5**, 287–294 (1992).
9. N.J.E. Dowling, J. Guezennec, J. Bullen, B. Little, and D.C. White, *Biofouling* **5**, 315–322 (1992).
10. R.D. Bryant, W. Jansen, J. Boivin, E.J. Laishley, and J.W. Costerton, *Appl. Environ. Microbiol.* **57**, 2804–2809 (1991).
11. P.J. Bremer and G.G. Geesey, *Appl. Environ. Microbiol.* **57**, 1956–1962 (1991).
12. I.B. Beech, C.C. Gaylarde, J.J. Smith, and G.G. Geesey, *Appl. Environ. Microbiol.* **35**, 65–71 (1991).
13. M.J. Franklin, and D.C. White, *Curr. Opinion Biotechnol.* **2**, 450–456 (1991).
14. N.D. Benbouzid-Rollet, M. Conte, J. Guezennec, and D. Prieur, *J. Applied Bacteriol.* **71**, 244–251 (1991).
15. R. Boopathy and L. Daniels, *Appl. Environ. Microbiol.* **57**, 2104–2108 (1991).
16. M.B. Deshmukh, I. Akhtar, R.B. Srivastava, and A.A. Karande, *Biofouling* **6**, 13–32 (1992).
17. P. Vandevivere and P. Baveye, *Appl. Environ. Microbiol.* **58**, 1690–1698 (1992).
18. G. Reid, H.S. Beg, C.A.K. Preston, and L.A. Hawthorn, *Biofouling* **4**, 171–176 (1991).
19. I.H. Pratt-Terpstra, J. Mulder, A.H. Weerkamp, J. Feijen, and H.J. Busscher, *J. Biomater. Sci. Polym. Ed.* **2**: 239–253 (1991).
20. C.C. Chang and K. Merritt, *J. Biomed. Mater. Res.* **26**, 197–207 (1992).
21. K. Merritt and C.C. Chang, *J. Biomat. Appl.* **5**, 185–203 (1991).
22. L. Passerini, K. Lam, J.W. Costerton, E.G. King, *Crit. Care Med.* **20**, 665–673 (1992).
23. C.P. Virden, M.K. Dobke, P. Stein, C.L. Parsons, and D.H. Frank, *Aesthetic Plant. Surg.* **16**, 173–179 (1992).
24. D.A. Jennings, M.J. Morykwas, W.W. Burns, M.E. Crook, W.P. Hudson, and L.C. Argenta, *Ann. Plastic Surg.* **27**, 216–220 (1991).
25. J.C. Nickel, J. Downey, and J.W. Costerton, *Urology* **39**, 93–96 (1992).
26. R.L.S. Whitehouse, E. Peters, J. Lizotte, and C. Lilge, *J. Dent.* **19**, 290–295 (1991).
27. Y.S. Park and K. Toda, *Biotechnol. Lett.* **14**, 609–612 (1992).
28. S.G. Velizarov, E.I. Rainina, A.P. Sinitsyn, S.D. Varfolomeyev, V.I. Lozinsky, and A.L. Zubov, *Biotechnol. Lett.* **14**, 291–296 (1992).
29. M. Moresi, E. Parente, and A. Mazzatura, *Appl. Microbiol. Biotechnol.* **36**, 320–323 (1991).
30. M.Y. Kwak and J.S. Rhee, *Appl. Microbiol. Biotechnol.* **36**, 578–583 (1992).
31. M.Y. Kwak and J.S. Rhee, *Biotechnol. Bioeng.* **39**, 903–906 (1992).
32. J.W. Tzeng, L.S. Fan, Y.R. Gan, and T.T. Hu, *Biotechnol. Bioeng.* **38**, 1253–1258 (1991).
33. B. Demuyakor and Y. Ohta, *Appl. Microbiol. Biotechnol.* **36**, 717–721 (1992).
34. G.A. Kovalenko and V.D. Sokolovskii, *Biotechnol. Bioeng.* **39**, 522–528 (1992).
35. B.Z. Eathepure and T.M. Vogel, *Appl. Environ. Microbiol.* **57**, 3418–3422 (1991).
36. J.P. Arcangeli and E. Arvin, *Appl. Microbiol. Biotechnol.* **37**, 510–517 (1992).
37. I. Coelhoso, R. Boaventura, and A. Rodrigues, *Biotechnol. Bioengrg.* **40**, 625–633 (1992).
38. D. Lemoine, T. Jouenne, and G.A. Junter, *Appl. Microbiol. Biotechnol.* **36**, 257–264 (1992).
39. J.A. Scott and A.M. Karanjkar, *Biotechnol. Lett.* **14**, 737–740 (1992).
40. G. Zellner, M. Gevcke, E.C. Demacario, and H. Dickmann, *Appl. Environ. Microbiol.* **36**, 404–409 (1991).
41. C.D. Gooijer, R.H. Wijffels, J. Tramper, *Biotechnol. Bioeng.* **38**, 224–231 (1991).
42. R.H. Wijffels, C.D. Gooijer, S. Kortekaas, J. Tramper. *Biotechnol. Bioeng.* **38**, 232–240 (1991).
43. T. Sumino, H. Nakamura, N. Mori, Y. Kawaguchi, and M. Tada, *Appl. Microbiol. Biotechnol.* **36**, 556–560 (1992).
44. L.E. Hallas, W.J. Adams, and M.A. Heitkamp, *Appl. Environ. Microbiol.* **58**, 1215–1219 (1992).
45. B. Heppner, G. Zellner, and H. Dickmann, *Appl. Environ. Microbiol.* **36**, 810–816 (1992).
46. A.M. Anselmo and J.M. Novais, *Biotechnol. Lett.* **14**, 239–244 (1992).
47. L.E. Macaskie, P.J. Clark, J.D. Gilbert, and M.R. Tolley, *Biotechnol. Lett.* **14**, 525–530 (1992).
48. G. Gonzalez, F. Ramirez, and O. Monroy, *Biotechnol. Lett.* **14**, 149–154 (1992).
49. D.E. Fennel, S.E. Underhill, and W.J. Jewell, *Biotechnol. Bioeng.* **40**, 1218–1232 (1992).
50. G.I. Loeb and R.A. Neihof, in R.E. Baier ed., *Applied Chemistry at Protein Surfaces*, Vol. 145, American Chemical Society, Washington, D.C. 1975, pp. 314–335.
51. V.A. DePalma, D.W. Goupil, and C.K. Akers, 6th Offshore Thermal Energy Conversion Conf., Washington, D.C., May 15–18, 1979.
52. J.D. Bryers, Ph.D. Dissertation, Wm. M. Rice University, Houston, Tex., 1980.
53. B.D. Bowen, S. Levine, and N. Epstein, *J. Coll. Interfac. Sci.* **54**, 375–390 (1978).
54. S.K. Beal, *Nucl. Sci. Eng.* **40**, 1–11 (1972).
55. L.K. Jang and T.F. Yen, in Zajic and Donaldson, eds., *Microbes and Oil Recovery*, Vol. 1, International Bioresources J., 1985, pp. 226–246.
56. M.K. Banks, Ph.D. Dissertation, Duke University, Durham, N.C., 1989.
57. M.K. Banks and J.D. Bryers, *Appl. Microbiol. Biotechnol.* **33**, 596–601 (1990).
58. D. Davies and G.G. Gessey, *Appl. Environ. Microbiol.* **61**, 860–867 (1995).
59. P. Harremoës, *Vatten* **33**, 122–143 (1977).
60. H. Tanaka and I.J. Dunn, *Biotechnol. Bioeng.* **24**, 669–689 (1982).
61. P.S. Stewart, *Biotechnol. Bioeng.* **41**, 111–117 (1993).
62. D.H. Applegate, and J.D. Bryers, *Biotechnol. Bioeng.* **37**, 17–25 (1991).
63. A. Boyd and A. Chakrabarty, *Appl. Environ. Microbiol.* **60**, 2355–2359 (1994).
64. G.G. Rodriguez, D. Phipps, K. Ishiguro, and H.F. Ridgway, *Appl. Environ. Microbiol.* **58**, 1801–1808 (1992).
65. R.I. Amann, J. Stromley, R. Devereux, R. Key, and D.A. Stahl, *Appl. Environ. Microbiol.* **58**, 614–623 (1992).
66. J. Rogers and C.W. Keevil, *Appl. Environ. Microbiol.* **58**, 2326–2330 (1992).

67. L.L. McCarter, R.E. Showalter, and M.R. Silverman, *Biofouling* **5**, 163–176 (1992).
68. L. Dagostino, A.E. Goodman, K.C. Marshall, *Biofouling* **4**, 113–120 (1991).
69. T.C. Brelje, M.W. Wessendorf, and R.L., Sorenson, *Methods Cell Biol.* **38**, 97–181 (1993).
70. A.C. Camper, M.A. Hamilton, K.R. Johnson, P. Stoodley, G.J. Harkin, and D.S. Daly, *Ultrapure Water* **11**, 26–35 (1994).
71. J.R. Lawrence, D.R. Korber, B.D. Hoyle, J.W. Costerton, and D.E. Caldwell, *J. Bacteriol.* **173**, 6558–6567 (1991).
72. S. Moller, A.R. Pedersen, L.K. Poulsen, E. Arvin, and S. Molin, *Appl. Environ. Microbiol.* **62**, 4632–4640 (1996).
73. K.P. Ishida and P.R. Griffiths, *J. Colloid Interface Sci.* **160**, 190–200 (1993).
74. P.J. Bremer and G.G. Geesey, *Biofouling* **3**, 89–100 (1991).
75. P.A. Suci, M.W. Mittleman, F.P. Yu, and G.G. Geesey, *Antimicrob. Agents Chemother.* **38**, 2127–2133 (1994).
76. Z. Lewandowski, G. Walser, W.G. Characklis, *Biotechnol. Bioeng.* **38**, 877–882 (1991).
77. A. Schramm, L.H. Larsen, N.P. Revsbech, N.B. Ramsing, R. Amann, and K.-H. Schleifer, *Appl. Environ. Microbiol.* **62**, 4641–4647 (1996).
78. S.B. Libicki, P.M. Salmon, and C.R. Robertson, *Biotechnol. Bioeng.* **32**, 68–80 (1988).
79. H. Siegrist and W. Gujer, *Water Res.* **19**, 1369–1378 (1985).
80. F.E. Drummond, M.S. Thesis, Duke University, Durham, N.C. (1993).
81. J.D. Bryers and F.E. Drummond, in R.H. Wijffels, R.M. Buitelaar, C. Bucke, and J. Tramper eds., *Progress in Biotechnology 11: Immobilized Cells*, Elsevier Science BV, Amsterdam, The Netherlands, 1996, pp. 196–204.
82. E. Poitevin and P. Wahl, *Biophys. Chem.* **31**, 247–258 (1988).
83. J.D. Bryers and F. Drummond, *Biotechnol. Bioeng.* **60(4)** (1998).
84. P. Stoodley, D. deBeer, and Z. Lewandowski, *Appl. Environ. Microbiol.* **60**, 2711–2716 (1994).
85. J.E. Bailey and D.F. Ollis, *Biochemical Engineering Fundamentals*, 2nd ed., McGraw-Hill, New York, 1986.
86. D.F. Ollis, *Phil. Trans. R. Soc. London* B, **297**, 617–629 (1982).
87. D.E. Dykhuizen and D.L. Hartl, *Microbiol. Rev.* **47**, 150–168 (1983).
88. P. de Taxis du Poët, P. Dhuster, J.-N. Barbotin, and D. Thomas, *J. Bacteriol.* **165**, 871–877 (1986).
89. M. Nasri, S. Sayadi, J.-N. Barbotin, P. Dhuster, and D. Thomas, *Appl. Environ. Microbiol.* **53**, 740–744 (1987).
90. S. Sayadi, M. Nasri, J.N. Barbotin, and D. Thomas, *Biotechnol. Bioeng.* **33**, 801–808 (1989).
91. C.-T. Huang, S.W. Peretti, and J.D. Bryers, *Biotechnol. Bioeng.* **41**, 211–220 (1993).
92. C.-T. Huang, S.W. Peretti, and J.D. Bryers, *Biotechnol. Bioeng.* **44**, 329–336 (1994).
93. W.C. Fuqua, S.C. Winans, and E.P. Greenberg, *J. Bacteriol.* **176**, 269–275 (1994).
94. G.P.C. Salmon, B.W. Bycroft, G.S.A.B. Stewart, and P. Williams, *Mol. Microbiol.* **16**, 615–624 (1995).
95. J.S. Todd, R.C. Zimmerman, P. Crews, and R.S. Alberte, *Phytochemistry* **34**, 401–404 (1993).
96. M. Givskov, R. de Nys, M. Manefield, L. Gram, R. Maximillien, L. Eberl, S. Molin, P.D. Steinberg, S. Kjelleberg, *J. Bacteriol.* **178**, 6618–6622 (1996).
97. S.G. Dunkirk, S.L. Gregg, L.W. Duran, J.D. Monfils, J.E. Haapala, J.A. Marcy, D.L. Clapper, R.A. Amos, and P.E. Guire, *J. Biomaterials. Appl.* **6**, 131–155 (1991).
98. M.J. Bridgett, M.C. Davies, and S.P. Denyer, *Biomaterials* **13**, 411–416 (1992).
99. N.P. Desai, S.F.A. Hossainy, and J.A. Hubbell, *Biomaterials* **13**, 417–420 (1992).
100. W.K. Whitekettle, *J. Industrial Microbiol.* **7**, 105–116 (1992).
101. C.C. Chang and K. Merritt, *J. Orthop. Res.* **9**, 284–288 (1991).
102. G. Golomb and A. Shpigelman, *J. Biomed. Mater. Res.* **25**, 937–952 (1991).
103. G.D. Kamal, M.A. Pfaller, L.E. Rempe, and P.J. Jebson, *JAMA* **265**, 2364–2368 (1991).
104. J.W. Leung, G.T. Lau, J.J. Sung, and J.W. Costertoni, *Gastrointest. Endosc.* **38**, 338–340 (1992).
105. D. Stickler and P. Hewett, *Eur. J. Clin. Microbiol. Infect. Dis.* **10**, 416–421 (1991).
106. D. Stickler and P. Hewett, *Eur. J. Clin. Microbiol. Infect. Dis.* **10**, 157–162 (1991).
107. C.L. Wiatr and O.X. Fedyniak, *J. Industrial Microbiol.* **7**, 7–13 (1991).
108. R. Bakke, B. Rivedal, and S. Mehan, *Biofouling* **6**, 53–60 (1992).
109. P. Wood, M. Jones, M. Bhakoo, and P. Gilbert, *Appl. Environ. Microbiol.* **62**, 2598–2602 (1996).

See also BIOREMEDIATION; DEXTRAN, MICROBIAL PRODUCTION METHODS; ENZYMES, IMMOBILIZATION METHODS.

BIOFILTERS

MARIO ZILLI
ATTILIO CONVERTI
Genoa University
Genoa, Italy

KEY WORDS

Air pollution control biotechnologies
Biofilter
Biofiltration
Biological methods
Degradation pathways
Microorganisms

OUTLINE

INTRODUCTION

Biofiltration is the biological removal of air contaminants from waste gases by means of aerobic microorganisms that are immobilized on a porous solid support. In a biofilter, the waste-air stream is passed through a biologically active packing material; organic or inorganic air pollutants are degraded by microorganisms into harmless end products such as water, carbon dioxide, mineral salts, and new microbial mass. Clearly, biofiltration technology is limited to non-toxic gases with good biodegradability and water solubility, because these gases are used by microorganisms as their sole source of energy and carbon.

The idea of employing microorganisms to eliminate contaminants from waste gases is very old. Initially, biofiltration was used mainly for the treatment of odorous gases (such as hydrogen sulfide and ammonia) emanating from sewage-treatment plants. The first reports on actual applications of this technology, dating back to the late 1950s and the early 1960s, dealt with soil-bed installations in West Germany and the United States, where it was demonstrated that biodegradation, rather than sorption, was responsible for most of the odor abatement (1). After the first publication of Bohn's extensive investigation on soil-filter beds (2,3), several additional full-scale applications of biofilters in West Germany and the Netherlands followed. These biofilters were employed to control the odors from wastewater- and thermal-sludge-treatment plants, composting facilities, rendering plants, carcass incinerators, and others (4–6). Jäger and Jager, who compared several methods for off-gas purification from the composting plant of Heidelberg, first demonstrated the economic convenience of biofiltration (7).

Process and technology development was attempted by Thistlethwayte et al., who utilized a trickling filter column packed with river gravel or glass balls and seeded with activated sludge. Waste air contaminated by H_2S, C_2H_5SH, $(C_2H_5)_2NH$, and C_4H_9CHO was purified, in countercurrent with a nutrient solution, with removal yields ranging from 40% to almost 100% (8). In addition, applications in which organic impurities were eliminated from waste gas, or air in beds inoculated with biologically active material, have been reported together with the first criteria for correct process design and operation (9).

A further development of this technology was marked by a biofilter installation with continuous regeneration of the bed, which was obtained by removing active material at the bottom of the reactor and recharging it after regeneration at the top of the reactor (9). A regeneration process by microbial, thermal, and chemical methods was also reported by Helmer (10), who used dry- and wet-soil filter beds to extract products of anaerobic fermentation of organic materials (e.g., NH_3, H_2S, sulfides, alcohols, and aldehydes). The importance of a sufficient moisture content in the filter bed and of a good distribution of the gas supplied to the biofilter was demonstrated for the first time (11).

Leson and Winer estimated that in 1991 the total number of biofilter and soil-bed installations used in the United States and Canada mainly for odor control was less than 50 (1). There was greater confidence in biofiltration in Germany and the Netherlands, where over 500 biofilters were operating in the same period. In the last two decades, the number of installations has significantly increased in Japan, growing from about 40 in the 1980s to 90 in the 1990s (12). On the contrary, in other European countries, including Italy, Switzerland, and Austria, the number of biofilter installations has been quite limited (1).

Although biofiltration has developed mainly with the primary aim of eliminating odorous compounds from exhaust gases, since the early 1980s this technology has been of increasing importance in air pollution control, and the number of applications of biological filters has strongly increased due to an increasing concern about environmental quality and more stringent air-emission quality standards.

The application field of biofiltration has widened thanks to better knowledge of the biodegradation process, in particular the control of the process conditions, as well as to improved biofilter construction and packing material composition (13,14). In particular, special packing materials having optimal structural properties prevent aging phenomena, decrease the pressure drop over the packed bed, and increase the long-term operational stability of the filter, as well as its microbial activity.

Furthermore, the considerable research carried out on the isolation, selection, and construction of strains or mixed cultures of microorganisms (mainly bacteria) has led to the extension of biodegradation to both anthropogenic (i.e., manmade) and xenobiotic compounds (chemi-

cals that do not occur naturally). Compounds that sufficiently resemble structures of biological origin are eliminated rather quickly, whereas biodegradation of xenobiotics is very slow (recalcitrant compounds) or even impossible (persistent compounds). Nevertheless, the continuous adaptation of microorganisms to new substrates is at the basis of the increasing number of microbes capable of degrading xenobiotics.

As a consequence, biofiltration is now a well-recognized technology that is successfully applied to a wide range of industrial facilities, waste-disposal and food-processing activities to control odors, volatile organic compounds (VOCs), and air toxics (both inorganic and organic), especially those that are readily biodegradable.

Biofiltration development is the result of the implementation of health and safety guidelines, enforced in many countries to protect people and the environment, combined with several advantages that this technique has demonstrated over traditional chemical and physicochemical treatment methods, such as water washing, chemical scrubbing (ozonation and chlorination), activated carbon adsorption, incineration, and catalytic oxidation.

Besides the overall low energy requirements (moderate temperature and pressure conditions), easy maintenance, simple control, and low capital and operating costs, it must be stressed that biological filters do not transfer the pollution problem to another environmental compartment, as occurs with many other alternative control technologies. Because the pollutants are converted into harmless oxidation products (which are part of a natural cycle), they do not create any secondary pollution and thus can be considered environmentally safe.

BIOLOGICAL TECHNOLOGIES FOR WASTE-GAS TREATMENT

Biological Systems

There are three basic types of biological reactor systems used to treat waste gases: bioscrubbers, biotrickling filters, and biofilters. These can be grouped into those employing microorganisms dispersed freely throughout the liquid phase (bioscrubbers), and those using microorganisms immobilized on a packing or carrier material (biotrickling filters and biofilters). Moreover, in bioscrubbers and biotrickling filters the water phase is continuously moving, whereas in biofilters it is stationary.

A bioscrubber consists of a scrubber unit and a regeneration unit. In the scrubber (absorption column), water-soluble gaseous pollutants are absorbed and partially oxidized in the liquid phase (the culture medium containing the microorganisms), which is distributed from the top of the unit. The contaminated water is subsequently transferred into an aerated stirred-tank reactor (regeneration unit), similar to an activated-sludge unit, where the contaminants are fully biodegraded. The regenerated suspension is continuously recirculated to the top of the scrubber section, thereby enhancing process efficiency (Fig. 1).

Biotrickling filters (Fig. 2) and biofilters (Fig. 3) are different from bioscrubbers in that gaseous pollutant absorption and biological degradation occur simultaneously in the same compartment, resulting in more compact systems. The polluted air flows through a biologically active bed, where microorganisms are attached in the form of a biofilm. As the gas diffuses through the packed bed, the pollutants are transferred to the biolayer and degraded.

As illustrated in Figure 3, in order to ensure optimal operation of a biofilter, the inlet gas usually requires pretreatment processes such as: (*1*) particulate removal in order to prevent possible clogging and/or sludge build up, (*2*) load equalization in case the waste-gas concentration is subject to strong fluctuations, (*3*) temperature control, and (*4*) humidification.

In biological trickling filters the packed bed consists only of inert materials (glass, ceramics, and plastics), while the liquid phase, containing inorganic nutrients, flows downward over the packaging material in countercurrent with the contaminated gaseous stream and is continuously recirculated through the bioreactor.

Bioscrubbers and biotrickling filters are applicable mainly to the treatment of waste gases containing good or moderately water-soluble compounds, whereas biofilters, due to the large surface area available for mass transfer, are also suited to treat poorly water-soluble compounds. Moreover, due to their high reaction selectivity, biofilters are particularly suitable for treating large volumes of air containing easily biodegradable pollutants with relatively low concentrations, typically below 1,000 ppm.

Compared with the other biological systems, biofilters have the widest application because they are easy to operate, simply structured, and imply low installation and operating/maintenance costs. Also, the reliability of biofilter operation is higher than that of bioscrubbers, where the risk exists of washing away the active microorganisms. Moreover, the presence of a large amount of packing material with a buffering capacity diminishes the sensitivity of biofilters to different kinds of fluctuations (15). Because the major disadvantage is the difficult control of parameters such as pH, temperature, and nutrient supply, biofilters may be unsuitable for degrading halogenated compounds (as acid metabolites are produced) and treating gas streams containing high concentrations of VOCs, unless long residence times or large bed volumes are applied (3). Biotrickling filters and biofilters are currently utilized mainly in compost-production plants, sewage-treatment plants, and agriculture, whereas biofilters and bioscrubbers are preferred in industrial applications.

Different Types of Biofilters

Among the different types of biofilters described in the literature for a variety of applications (16), open biofilters, consisting of single beds of compost or porous soils, commonly 1-m deep, are used mainly for odor and VOC abatement. Because they are in direct contact with the open air, their performance is strongly influenced by weather conditions (rain, frost, temperature fluctuations, etc.). The simple design and low cost of these systems are counterbalanced by difficult monitoring and control, as well as by large space requirements.

All the other real-scale configurations are closed biofilters containing mixtures of organic materials and bulking

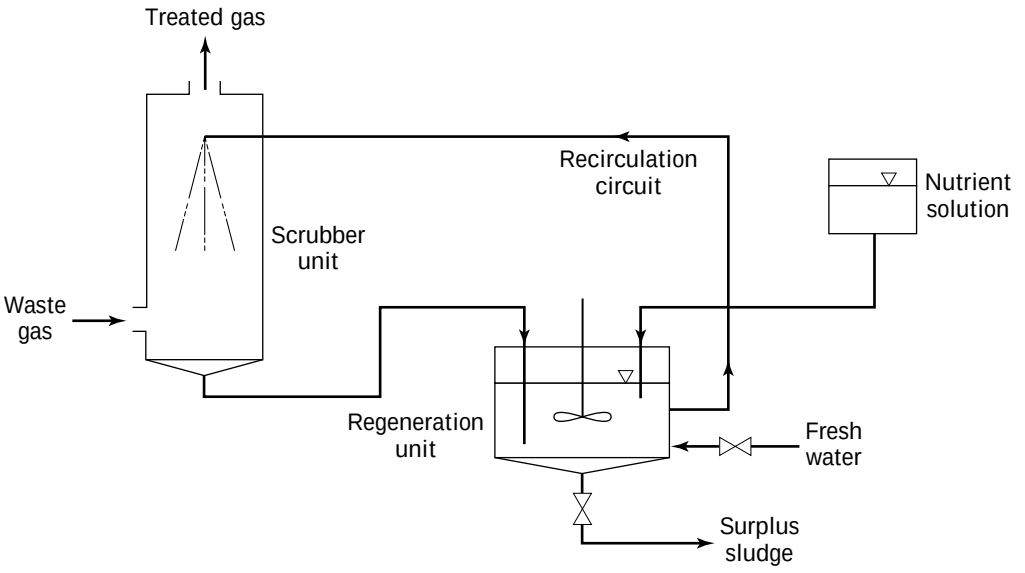


Figure 1. Scheme of a bioscrubber system.

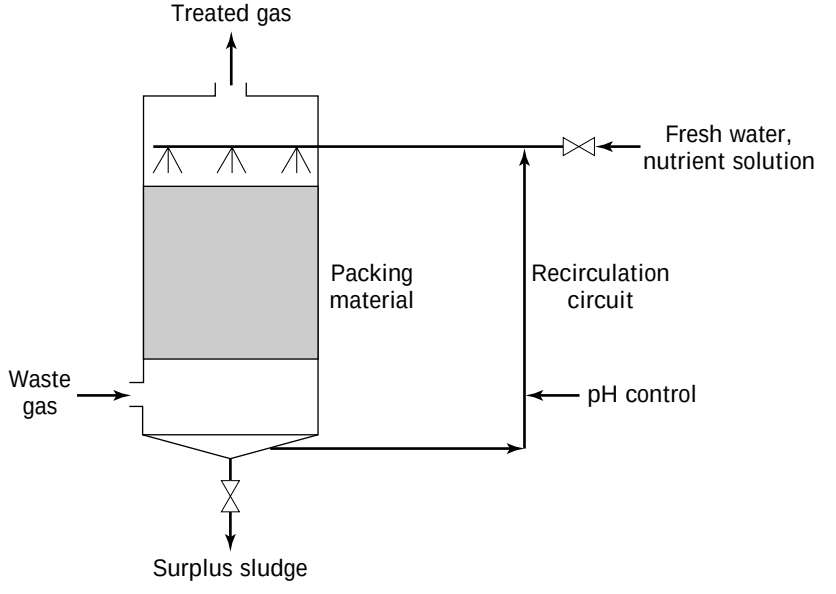


Figure 2. Scheme of a biotrickling filter system.

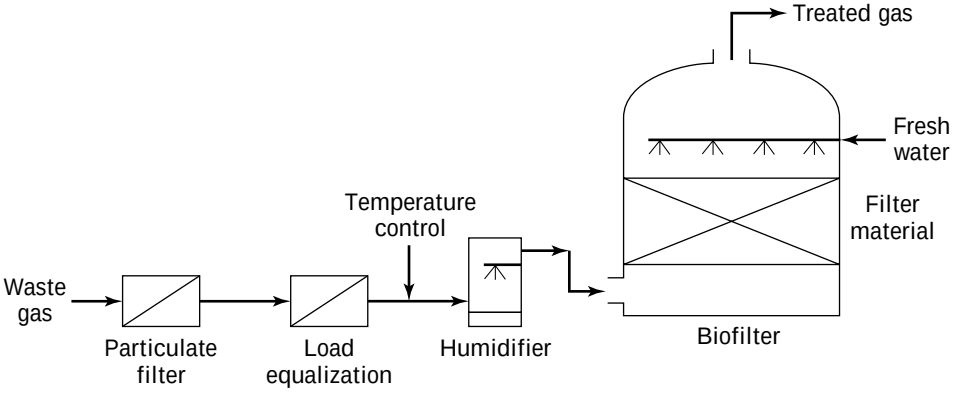


Figure 3. Scheme of a closed single-layer biofilter system.

agents. Among these, simple single-layer systems, often accurately controlled and monitored, but scarcely flexible and requiring large space, can be used for VOC treatment. Multiple-layer systems, characterized by increased complexity and cost but less space requirements, consist of separately supported layers that allow optimal growth conditions for different microbial populations in each stage; a better control of the bed environment during the treatment of pollutant mixtures is also achieved.

More sophisticated and very promising configurations have been successfully tested on bench- or pilot-scale projects for the treatment of VOC mixtures, BTEX (benzene, toluene, ethyl benzene and xylene), and reduced-sulfur compounds, among which there are multistage, modular, and step-feed systems, that in general are characterized by better control and flexibility, but higher costs and operative complexity.

PROCESS PARAMETERS

Biofiltration is a very complex process that involves mass transfer and reaction and is also influenced by the fluid dynamic of the gaseous stream through the reactor. Furthermore, it is necessary to maintain an environment as close as possible to the optimal conditions for the microorganisms. Therefore, several important process parameters, such as filter material structure, moisture content, temperature, and pH, must be kept under strict control and within an optimum range.

Filter Material

The choice of filter material is of great importance in order to maintain the efficiency of a biofilter, in that it must (*1*) guarantee optimum environmental living conditions for the microorganisms; (*2*) constitute at the same time a nutritious reserve, a humidity reservoir, and a mechanical odorless support; and (*3*) provide the structural stability of the bed.

The choice of the material is strongly influenced by the need to minimize the overall volume necessary for the reactor, optimize the removal efficiency, keep energy consumption to a minimum, and minimize maintenance. In addition, the characteristics of the carrier material impact directly on the microbial growth and activity, thus in turn affecting biofilter performance. Biofilter beds have the advantage of immobilizing the microflora on the packing material, as a result of which these organisms, forming a biolayer, are not drained from the system, as is often the case in freely dispersed systems.

Small particles of natural organic materials, such as compost, peat, soil, or mixtures of these materials with bark, leaves, wood chips, heather branches, humus earths, or brushwood (less than 10 mm in diameter), are widely used as packing media in biofilters because they provide a high specific surface area (from 300 to 1,000 m^{-1}), favorable living conditions for the resident microbial population (ensured by high retention capacities of water and nutrients), and favorable immobilization for the microflora involved. In practice these packing materials have shown the common disadvantage of being strongly subject to aging phenomena, resulting in bed shrinkage. This phenomenon disturbs the homogeneous flow distribution of the gas and may cause a considerable increase in the pressure drop, a decrease in the specific area, and shorter filter lifetime. Moreover, the natural unhomogeneity of the structure of these natural materials prevents the uniform distribution of the gas flow, provoking short-circuiting channel development in the filter bed, with an increase in the flow resistance and a decrease in the biological degradation capacity of the bed.

In order to prevent these problems, inert materials, such as polystyrene spheres, lava particles, glass beads, porous clay, and ceramic, are usually added to the natural filling materials. This combination, improving the uniform distribution of the gas flow, lowers the head losses up to 100–150 mm of water gauge (6), thus contributing to reducing the power requirement necessary to convey the gas through the filter and, consequently, the operating costs. Moreover, the addition of porous materials (e.g., granular activated carbon) with high internal porosity and hydrophilic properties, increases the buffering capacity of the filter, (which is very favorable, particularly in the treatment of gas streams with strongly fluctuating pollutant concentrations). The selection of the most suitable material for a specific application depends on the nature of the pollutants and on the magnitude of the concentration fluctuations in the gaseous stream. In order to buffer the fluctuations in the concentration so that a constant supply of contaminants to the biofilter can be achieved, the adsorbent should have both good adsorbing power and reasonable desorbing properties. This combination provides the adsorbent with the capacity of adsorbing at high concentrations and desorbing at low concentrations and allows a significant reduction in the required filter volume (15).

Compared to soil, compost has the advantage of providing lower resistance to the gas flow and consequently contained pressure drop, which should not, in general, exceed 250 mm of water gauge (17). Peat is the material with highest water-retention capacity and constitutes the optimal substrate for the microorganisms; nevertheless, as it implies a higher pressure drop, it is often mixed with other materials in order to improve the structural stability.

Moisture Content

The good performance of a biofilter depends greatly on an optimal moisture content of the filter bed. Moisture, the most critical operational parameter, is, in fact, essential for the survival and activity of the microbial flora (which can absorb and degrade substances only in the aqueous phase) and contributes to the filter buffer capacity.

A lack of moisture causes cracking of the bed mainly at the gas inlet side, where the microbial activity is stronger due to the higher air-pollutant concentration. This results in (*1*) a considerable decrease of the microbial activity, (*2*) development of short-circuiting channels that cause bypass of the gas flow, and (*3*) contraction of the filter bed with consequent breakthroughs. On the contrary, an excess of water content promotes the development of anaerobic zones in the bed, due to occlusion of the pores, and provokes the formation of odorous volatile metabolic products, which are released in these zones and transported by

the gas flow. Moreover, due to the higher gas-flow resistance, these zones are bypassed by the gas flow, resulting in a decrease in the mean residence time of the gas and, hence, in the efficiency of the process, and causes a 10% increase in the power requirement (2). A moisture excess causes problems also in the oxygen transfer due to a reduction of the gas–water interface per unit biofilm volume, and in the drainage of the filter components, with production of low-pH and high-load leachate that requires disposal. The optimal moisture content of the filter bed should then range between 40 and 60% by weight, depending on the material. In order to preserve the microbial activity, it should exceed at least 40% (18).

Water consumption by a biofilter is generally low, depending on the temperature and the relative humidity of the input gas and, for open filters, on precipitation. Due to evaporation of the aqueous phase, freshwater must be continuously supplied to the system in order to ensure a degree of saturation of more than 95%, which is necessary for optimal living conditions. This is normally realized by humidifying the inlet gas stream in a water-spray scrubber upstream of the filter unit and by periodically spraying water onto the surface of the packing material by means of surface sprinklers, spray nozzles, and so on. This periodical water supply is necessary not only because it is not always possible to humidify the inlet waste gas up to the required level, but also because heat is produced by the microbial oxidation of the pollutants.

Moistening equipment must be operated in such a way that the moisture content stays within the indicated limits at any point of the medium. This operation must be realized in a very controlled way, otherwise the structure of the bed may be disturbed, which results in compaction of the filter bed in the long run and in increased gas-phase pressure drop.

pH

Maintaining an optimal pH in the filter material is a very important operational requirement for maximizing the biofiltration process because the microbial activity is strongly dependent on the pH. The effect of pH on biological systems is the indirect result of its action on the enzymatic reaction rate, which usually falls outside an optimal value. In biofiltration, as well as in most aerobic biological processes, the range of pH within which the biological systems can operate is typically between 5 and 9, with an optimum range of 6–8.

Because in biofilters the aqueous phase is stationary, problems of package acidification may arise during the degradation of pollutants such as sulfur- or nitrogen-containing compounds and halogenated hydrocarbons, which can generate acidic metabolites such as sulfuric, nitric, or hydrochloric acid. As a consequence of their accumulation, the microbial activity may decrease or even stop, seriously compromising the overall elimination capacity of the biofilter. These problems can be prevented by addition of buffering agents, such as lime, dolomite, limestone, marl, phosphates, or other water-insoluble alkaline materials, even if the neutralization products can sometimes reach inhibitory levels (13).

Besides the above adverse effect on the process, acidification of the biofilter, as well as the acidic nature of the waste gas, filter material, or leachate, can corrode the air distribution and leachate collection pipes. In such cases, the use of materials resistant to corrosion, together with a continuous pH monitoring, are suggested.

Temperature

Temperature remarkably influences the microbial growth through its effects on both bacterial enzyme activity and on various metabolic processes. Because the microorganisms generally applied in biofilters are mesophilic, temperature should be maintained between 20 and 40 °C, with an optimum of 35 °C for most aerobic microorganisms. Suboptimal temperatures can slow down the degradation capacity. The bed temperature should not rise above 45 °C, although the resident microflora can gradually adapt itself to temperatures as high as 50 °C (19). Microbial activity slows dramatically above 65 °C. The effect of low temperature on biofilters depends largely on the inlet gas temperature. In fact, although low bed temperatures, especially below 10 °C, decelerate biological oxidation, this effect can be partially counterbalanced by increased gas sorption by solids (3). However, the microbial activity will recover again upon preheating the influent exhaust gas. Fortunately, microorganisms have shown a tremendous ability to rapidly adapt themselves to drastic variations of temperature.

Reaction and diffusion rates typically increase with increasing temperature; nevertheless this effect is offset by a decrease of the solubility of the compounds to be removed, as well as a decrease in the physical adsorption capacity of the packing material at higher temperatures. Operation of the biofilter at high temperature shifts the microbial population towards thermophilic microorganisms.

In order to prevent the death of the microorganisms, it is often necessary to control the influent gas temperature by precooling through heat recovery, water injection, or wetting of the packing material, thus offering additional economic advantage and the possibility of raising the gas relative humidity. Moreover, the heat liberated by aerobic microbiological activity causes a significant increase in temperature within the filter material, which must be taken into consideration at the design phase, particularly in biofilters that treat waste gases containing high concentrations of volatile organic compounds. For good operation of the biofilter, it is obviously necessary to continuously monitor the filter-bed temperature.

PROCESS PRINCIPLES

Physical Model

In a biofilter, the microorganisms adhere to the surface of the porous solid material used as the support and grow, thus forming a biofilm. All models commonly used to describe the phenomena taking place in a biofilter under steady-state conditions are based on the principles of heterogeneous catalysis. According to the macrokinetic model

proposed by Ottengraf (20), the support particle is surrounded by a biofilm; both the pollutant and oxygen molecules dispersed in the gaseous phase continuously transfer to the active biofilm region. Due to the metabolic activity of aerobic microorganisms, a concentration gradient forms that is responsible for the continuous mass transfer of the pollutant from the gaseous to the liquid phase. At the same time, the products of the aerobic degradation (carbon dioxide, water, inorganic compounds, etc.) continuously counterdiffuse toward the gaseous phase, where they are removed by the upward flow and are finally released from the biofilter.

The metabolic activity is made possible by the continuous diffusion of inorganic nutrilites (mainly nitrogenous and phosphoric compounds) from the support to the biofilm. Nutrilite consumption is only partially counterbalanced by the mineralization of dead cells, therefore periodic addition or support renewal is necessary.

Because the small particle size and the low solubility of the organic pollutants in water make the mass-transfer resistance in the gas phase negligible, pollutant concentrations at the biolayer interface and in the gas-phase bulk can reasonably be assumed in equilibrium and related by Henry's law.

Degradation Kinetics

The degradation macrokinetics in biofilters were reviewed by Ottengraf (20). Assuming gas flow through the biofilter to be plug flow, and that degradation follows typical Monod kinetics law (21), two different situations may occur: first-order or zero-order kinetics, respectively. Several experimental works (13) demonstrated that zero-order kinetics are typically observed in biofilters for most of the volatile compounds, even at very low concentration levels (13,20,22).

Reaction Limitation. The process is reaction-limited above a critical pollutant concentration (C_{crit} in the off-gas (Fig. 4a). Due to there being no diffusion limitation, the wet biolayer is fully active (e.g., the biofilter works at its maximum elimination capacity), and pollutant elimination is limited by the reaction (Fig. 5a). Under these conditions, the degree of conversion (η) is given by the ratio between

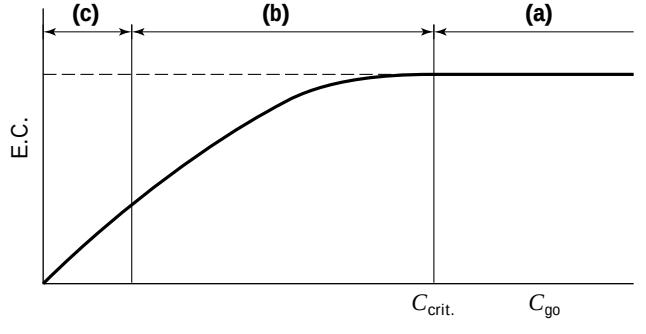


Figure 4. Dependence of the elimination capacity of a biofilter on the inlet off-gas concentration. (**a**) Reaction limitation; (**b**) Diffusion limitation; (**c**) 100% conversion.

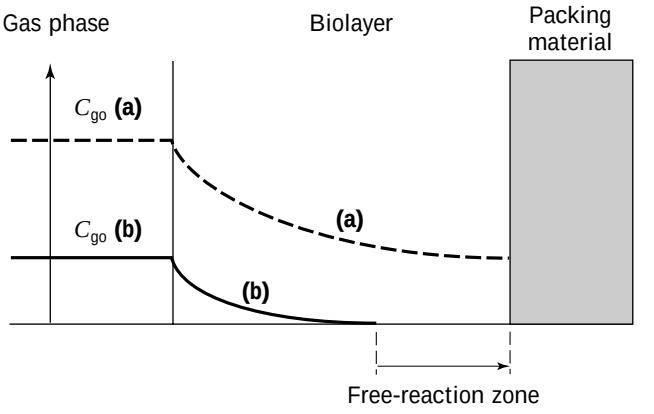


Figure 5. Biophysical substrate-penetration model.

the maximum elimination capacity, K_o which corresponds to an apparent zero-order kinetic constant, and the bed load per unit volume ($U_g C_{go}/H$) (20):

$$\eta = 1 - (C_{ge}/C_{go}) = HK_o/U_g C_{go}$$

So, at the macroscopic level, the pollutant concentration in the filter bed linearly decreases with the height, and the compound is completely consumed when $H\ K_o$ equals $U_g\ C_{go}$.

Diffusion Limitation. Below the critical concentration, diffusion limitation occurs and the rate of reaction tends to exceed that of diffusion (Fig. 4b). As a consequence, the depth of penetration becomes smaller than the layer thickness and the biolayer is not fully active (Fig. 5b). The pollutant is nearly completely consumed by the biofilm before having entirely crossed it, and the conversion rate, which decreases with decreasing pollutant concentration in the waste gas, is controlled by diffusion. Under these conditions, the elimination capacity of the biofilter is influenced by both the reaction and the diffusion rates and the compound is completely removed when $H\ K_o \geq 2\ U_g\ C_{go}$ (20).

The applicability of this model to multicomponent waste gases is limited by the theoretical complexity of the system because the various compounds are not biodegraded independently. Industrial application regarding mixture of pollutants requires pilot testing for accurate sizing of a full-scale system.

DESIGN AND MANAGEMENT CRITERIA

The main criteria for biofilter design and management are the degree of conversion or removal efficiency, the residence time, the superficial gas flow rate, the mass loading, and the elimination capacity (16).

Removal Efficiency

The removal efficiency (η), which provides information on the biofilter effectiveness, is given by the fraction of removed pollutant mass ($m_o - m_e$) with respect to total mass in the influent gas (m_o). The most common way of design-

ing and managing biofilters consists of selecting the desired removal efficiency for one or more pollutants and deriving from this all the other criteria. The best removal efficiencies (95–99%) are usually observed for aromatic compounds, such as benzene, toluene, and benzoic acid (23).

Residence Time

The residence time of a gaseous pollutant in a biofilter (τ), given by the ratio of the empty bed volume of the filter (V_e) to the waste-gas flow rate (Q_g), is a fundamental parameter that cannot assume values smaller than a critical level if satisfactory removal efficiencies are desired. However, due to the difficulty in estimating empty-bed volume, residence time is often referred to total-bed volume. According to the pollutant degradability, this parameter typically ranges between 15 and 60 s. For liquid pollutants it is necessary to take into account their partition coefficients between liquid and gaseous phases.

Superficial Gas Flow Rate

The superficial gas-flow rate (U_g), defined as the ratio of the gas flow rate (Q_g) to the total cross-sectional area (A), is a parameter increasing with flow rate and decreasing with residence time; it usually ranges, depending on pollutant type, from 50 to 200 $m^3\ m^{-2}\ h^{-1}$. Since a superficial gas-flow rate increase obviously implies a reduction in removal efficiency, a maximum threshold of about 200 $m^3\ m^{-2}\ h^{-1}$ should not be exceeded (24).

Mass Loading

The mass loading (L_v) is defined as the ratio of the gas-flow rate (Q_g) to the bed volume (V) multiplied by the inlet pollutant concentration in the off gas (C_{ge}); for VOCs it should not exceed values of 3,000 ÷ 5,000 mg m^{-3}. In a biofilter, load increase could be ensured by either increasing waste-gas concentration or decreasing residence time. Although increasing gas concentration should, in general, accelerate biodegradation, excess concentrations above a certain threshold can inhibit microbial activity, so that decreasing residence time should be preferred. In addition, since the removal efficiency is affected by an increase in mass loading, and because excess organic load may obstruct the biofilter with the possible formation of toxic and/or acidic intermediates (25), a value between 10 and 160 g $m^{-3}\ h^{-1}$ should be employed to achieve satisfactory removal yields.

Elimination Capacity

The elimination capacity (EC) is defined as the ratio of gas flow rate (Q_g) to the bed volume (V), multiplied by the difference between the pollutant concentration in the influent and the effluent waste gases ($C_{go} - C_{go}$), it indicates a measure of the biofilter ability in removing the pollutants. From this definition it is evident that, because removal efficiency is expected to be close to 100%, the elimination capacity also ranges between 10 and 160 g $m^{-3}\ h^{-1}$ (1). In addition, this parameter also is higher close to the inlet side, as happens for mass loading. The elimination capacity increases with mass loading and inlet pollutant concentration and decreases with residence time, the limiting value depending not only on the pollutant degradability, but also on the support material and the operating conditions. Beyond this value it is nearly constant (20).

APPLICATIONS OF BIOFILTERS

Bench-Scale Results

Although many experimental works on biofiltration, both at laboratory and full-scale size, have been published during the last two decades, a critical comparison of results is difficult due to various reasons, among which are the use of German and Dutch as the most common languages in this research field, the use of different criteria to express the performances, the utilization of different volatile organic and inorganic compounds, the different types of biofilters used, and so on. Nevertheless, the work of Ottengraf (20) can be taken as a useful reference basis to provide a complete overview on biofilter performance, as well as on the various effects influencing the process macrokinetics.

For most organic pollutants, the steady-state continuous elimination follows zero-order kinetics, so that pollutant concentration decreases with the column height. In the presence of mixtures of pollutants characterized by different biodegradability, a nearly complete removal of the most easily degradable compounds takes place at the bottom of the filter, whereas the most recalcitrant are mainly removed at the top. This suggests that different microbial populations could be responsible in the filter bed for the degradation of different organic components.

The path of Figure 4 explains another typical behavior of biofilter performances besides the one illustrated earlier. At low organic loads (corresponding to low concentrations for a given residence time), the elimination capacity linearly increases with this parameter, which means that nearly complete pollutant conversion is achieved. Over a critical load, the elimination capacity reaches a maximum value (corresponding to the apparent zero-order kinetic rate constant, K_o) and stays nearly constant. Maximum thresholds of 21, 27, 32, and 32 g $m^{-3}\ h^{-1}$, calculated for a mixture of toluene, ethyl acetate, butyl acetate, and butanol in a five-stage system simulating a biofiltration column with variable height (20), seem to reflect a direct dependence of this parameter on the pollutant degradability.

Pilot-Scale and Real-Scale Applications

Biofilters of different sizes (areas ranging from 10 to 2,000 m^2) are currently used in many industrial sectors, mainly in Europe, to remove a variety of mixtures of volatile organic and inorganic compounds from off-gases characterized by variable flow rates (1,000–150,000 $m^3\ h^{-1}$) and low pollutant concentrations (usually less than 1,000 ppm). As one can see from Table 1, where the main pollutants removable by biofiltration are listed (3,13,16,26,27), a relatively variable degradability characterizes not only the different classes of pollutants, but also the various pollutants within the same class. While for easily biodegradable substances the microorganisms suited for their degradation are usually present in the natural packing materials, in-

Table 1. Classification of the Main Pollutants that can be Removed by Biofiltration, according to their Biodegradability

Excellent	Good	Minimum	No	Uncertain
Aliphatics	*Aliphatics*	*Alicyclics*	*Halogenated*	*Aliphatics*
Butadiene	Hexane	Cyclohexane	1,1,1-Trichloroethane	Acetylene
Aromatics	*Aromatics*	*Aliphatics*		*Nitrogenous*
Ethyl benzene, xylene Cresols	Benzene, styrene, toluene Phenols	Methane, pentane		Isocyanates
Nitrogenous	*Halogenated*	*Aromatics*		*Oxygenated*
Trimethylamine	Chlorophenols	PAH		Methyl methacrylate
Oxygenated	*Nitrogenous*	*Halogenated*		*Sulfuric*
Alcohols: butanol, ethanol, methanol Aldehydes: acetaldehyde, formaldehyde Esters: Ethyl acetate Ethers: Tetrahydrofurane Ketones: Acetone Organic acids: Butyric acid	Amides Heterocycles: Pyridine Isonitriles Nitriles: Acetonitrile	Carbon tetrachloride Dicloroethane Dichloromethane Pentachlorophenol Perchloroethylene Trichloroethane Trichloroethylene		Isothiocyanates
Sulfuric	*Oxygenated*	*Nitrogenous*		
Methyl mercaptan	Methylisobutylketone	Nitrocompounds		
Inorganics	*Sulfuric*	*Oxygenated*		
HCl, HF, H_2S, NH_3, NO_x (except NO_2), PH_3, SiH_4, SO_2	Heterocycles: Thiophene Thiocyanates Thioethers: Dimethylsulphide	Dioxane		
		Sulfuric		
		Carbon disulfide		

Source: From Refs. 3, 13, 16, 26, and 27.

oculation and enrichment with cultivated organisms is often necessary for the slowly biodegradable ones.

On the other hand, Table 2 provides a picture of the most significant industrial sectors where biofiltration is applied with success, with indication of the related pollutants (26).

Finally, from Table 3, where the results of the principal pilot and full-scale applications of biofilters are summarized, it is evident that conversion yields often higher than 90% are not uncommon in industrial practice (16,17,20,27–32).

MICROORGANISMS

Another factor strongly influencing biofiltration performance is the composition of the heterogeneous microflora, which primarily depends on the composition of the polluted gaseous stream. As a consequence of the progressive adaptation of microorganisms to the organic pollutants contained in the off-gas, which typically takes about 10 days (20), the population distribution shifts towards strains that naturally catabolize these pollutants.

Microbial Populations of Biofilters

Microbial population distribution along biofilters has recently been reviewed by Benedusi et al. (33). The purification of waste air containing a limited number of pollutants is effectively achieved by a population restricted to a few microbial species and generally requires, to shorten the start-up operation time, inoculation of the biofilter with pure cultures or activated sludges previously adapted to each pollutant. This practice is particularly suited to the degradation of very recalcitrant or complex organic pollutants, such as halogenated and aromatic hydrocarbons, which often require cometabolism with more easily degradable substances to accelerate the process (34,35). On the other hand, gaseous emissions simultaneously polluted by several compounds, such as those coming from water works, compost plants, and so on, require as inoculum a more heterogeneous population possessing a wider variety of metabolic pathways, which is available in the microorganisms naturally present in activated sludges and compost. Microbial loads of about 10^5, 10^4, and 10^3 colony-forming units per gram of support (CFU/g) have been reported for oligoheterotrophic, copioheterotrophic, and autotrophic metabolic types, respectively in heather-peat biofilters treating septic-tank emissions (33). Although Bohn (3) demonstrated that enriching compost-based biofilters with microorganisms is not advantageous in terms of removal yields, probably because of the unfavorable competition with the indigenous population, the addition of activated sludges to medium devoid of microorganisms

Table 2. Typical Pollutants Removed by Biofiltration, in Addition to Odor Control, in Different Industrial Applications

	Organics						Inorganics	
Type of Industry	Aliphatic	Aromatic	Oxygenated compounds	S-containing compounds	N-containing compounds	Halogenated compounds	H_2S	NH_3
Food and Agriculture Industry								
Aroma extraction		X						
Yeast drying			X		X			
Blood and fish-meal production			X	X	X		X	X
Bristle and feather drying			X	X	X			X
Fish roasting			X	X	X		X	
Fat processing			X	X	X		X	X
Gelatin production			X	X				
Coffee and cacao roasting		X	X	X	X			
Bone processing			X	X	X		X	X
Meat and fish curing	X	X	X		X			X
Animal feed production			X	X	X		X	X
Slaughterhouses			X	X	X		X	X
Tobacco processing		X	X		X			X
Livestock farming			X	X	X		X	X
Chemical Industry								
Foundries	X	X	X					X
Plastics processing	X	X	X		X			
Lacquer production and processing	X	X	X					
Adhesive production		X	X					
Oil and fat production	X	X		X		X		
Polyester manufacture				X	X			
Friction-linings production			X	X	X			X
Rendering plants			X	X	X	X	X	X
Glue production		X	X					
Treatment Plants								
Used oil processing	X	X		X		X		
Dump gas removal			X	X			X	
Sewage treatment and sludge drying	X	X	X	X	X	X	X	X
Composting and waste processing	X		X	X	X	X	X	X
Manure drying	X		X	X	X		X	X

Source: Elaboration from Ref. 26.

is a common practice (1). Finally, the growth and activity of the microorganisms are strongly influenced by the availability of oxygen and nutrients, the eventual presence of toxic substances in the exit gas, the degree of moisture, temperature, pH, and so on.

A great advantage of the heterogeneous population present in the biofilter is the excellent ability to survive long periods (up to two months) without activity, provided that periodic aeration of the bed is ensured, even if microbial activity seems to be hardly affected (36).

As revealed by plate counting in nutrient agar, most of the microorganisms growing in biofilters treating organic pollutants are heterotrophic eubacteria, actinomycetes, and fungi (1,37), which utilize the organic compounds contained in the gaseous streams as carbon and energy sources. Even when the presence in the off-gas of inorganic substances, like thiosulfate or hydrogen sulfide, should favor the growth of chemio-litoautotrophic microorganisms, such as *Thiobacillus* sp., by utilizing these substances as the energy source and carbon dioxide as the carbon source, the heterotrophic population becomes prevalent and is probably responsible for most of the degradation activity (33,38,39).

The bacteria more frequently detected in biofilters are soil bacteria, such as *Bacillus cereus* var. *mycoides* and several strains and species of actinomycetes belonging to the genus *Streptomyces*. Helmer (40), who did a microbiological characterization of compost-filter population, observed abundant growth of eubacteria (10^9 CFU/g), actinomycetes (2×10^6 CFU/g), and fungi (10^6 CFU/g) along the whole filter height, with higher density of fungi at the bottom. The highest bacterial densities are usually detectable near the inlet side, where substrates are more concentrated and more easily degradable, while different strains well-adapted to low substrate levels, as well as to more recalcitrant compounds, can be found at higher depth (37). The growth of several bacteria and fungi, whose cell densities decreased according to the sequence: *Actinomyces*

Table 3. Management Data of Representative Pilot and Industrial-Scale Applications of Biofilters

Applications	U_g (m h^{-1})	Q_g (m^3/h)	Pollutants	C_{go} (mg/m^3)	τ (s)	N of stages	η (%)	Reference
			Food industry					
Gelatin production		35,000	Odor		12–21	0.6–1	70–93	27
Cocoa and chocolate processing		10,000	Odor		22	2	99	27
	60		Odor	$4{\cdot}10^7$ o.u.		1	90	28
Fishmeal factory		40,000	Odor	230 (C)	20	1	50–90	27
Flavor and fragrance industry		25,400	Odor	10^5 o.u.	22	2	98	27
	550		Odor	$5{\cdot}10^4$ o.u.		3	98	28
Food processing industry		9,000	Odor (oil)	10^5 o.u.	20	2	93	27
	400		Odor	$2{\cdot}10^5$ o.u.		2	95	28
Meat rendering		80,000–120,000	Odor	$2{\cdot}10^4$ o.u.	60–90	1	99	16
Commercial bakery		90–200	VOC		25–60	2		
			Ethanol	1,000–4,000			80–100	
			Methane	700			10	16
			Aliphatics	200			10–30	
			Ethyl acetate				80–90	
Coffee industry	500		Odor	10^8 o.u.		2	90	28
Yeast production	100		Odor	7,500 o.u.		2	99	28
			Ethanol	500			99	
			Petrochemical industry					
Gasoline tank		222	BTEX	3–70		2	75–90	16
Oil storage	100		H_2S	240		2	99	28
Tank cleaning			Benzene	30			50	
	100		Acrylonitrile	16–300		2	50–98	28
			Alcohols	60			80	
Petroleum processing			BTEX				>95	
		8.5–70	C2–C5		45–360	1	≤50	17
			Methane				≤5	
			>C5				40–90	
			Other industries					
Tobacco industry		30,000	Odor		14	2	95	
			NH_3	1.5				27
			Nicotine	3.5				
Paint production		11,700	VOC	1,800	38	2	90	27
Pharmaceutical plant		75,000	VOC		108	3	80	27
Photo film production		140,000	VOC	400	30	2	75	27
Ceramics production		30,000	Ethanol		8	1	98	27
Metal foundry		40,000	Benzene	9	30	1	80	27
Industrial emission		17,000	Ethanol	1,000–4,000	60	2	80–90	16
Resin synthesis	65		Organics	25,000		2	40	28
			odor	10^6 o.u.			45	
Chemical industry		0.7–0.8	VOC	2,000–10,000		1	82	29
Printing industry	200		Esters, alcohols	1200		2	90	28
			Treatment plants					
Wastewater treatment		10,000	Odor		29	2	90–95	
			H_2S	10				27
			Acetone	8				
		30,000	Odor			2		
			H_2S	4.7			96	30
			NH_3	0.15			45–60	
	200		H_2S	300		1	80	28
Petrochemical sewage treatment	100		Odor	$1.3–10^5$ o.u.		2	75	28
Composting		154,000–240,000	Odor		30–47	4	>90	31
		16,000–22,000	Odor	45–230 (C)	40–60	1	93–96	20
	100		Odor	$1.3{\cdot}10^4$ o.u.		2	95	28
			Handicraft activities					
Wood paint and varnish	200		VOC	250		1	50	28
Metal paint and varnish	70		H_2S	4000 o.u.		2	90	28
Paint spraying			Styrene	60–600	15–60	1	95	32

globisporus, Penicillium sp., *Cephalosporium* sp., *Mucor* sp., *Micromonospora albus, Micrococcus albus,* and *Ovularia* sp., was also reported for similar supports used for mixtures of halogenated and aromatic hydrocarbons (9). An opposite bacteria distribution has been reported by Benedusi et al. (33) along a pilot heather-peat filter used to remove hydrogen sulfide and ethyl mercaptan from septic-tank emissions. They observed maximum bacterial densities just near the biofilter outlet side, with oligo-heterotrophic population (2.4×10^8 CFU/g) predominating over all the others: copiochemioheterotrophs (5.7×10^6 CFU/g), sulfur-oxidizing chemioautotrophs (2.6×10^5 CFU/g), and nitrifying chemioautotrophs (7.2×10^5 CFU/g). Bacterial densities at biofilter bottom vary between 1% and 3% of these values, probably due to the hard pH decrease (up to 2.5–3.5) consequent to hydrogen sulfide oxidation (to H_2SO_4), as well as to the inhibiting effect of high concentrations of this pollutant on microbial activity. Other bacteria detected in biofilters are *Micromonospora vulgaris* and *Proteus vulgaris* (41).

Among the main soil fungal taxa able to degrade polysaccharides naturally present as symbionts or saprophites in heather or peat (*Geomyces pannorum* var. *pannorum, Penicillium citrinum, Oidiodendron griseum, Penicillium frequentans, Pestalotiopsis guepinii,* and *Trichoderma koningii*), only *G. pannorum* var. *pannorum* and *T. koningii* seem to be able to develop in the bed under regular and irregular moisture conditions, respectively (33). Other taxa, initially absent in the support material but normally living in polluted waters and soil, appear as a consequence of the air flowing and progressive acclimation, under conditions either of irregular (*Mucor* sp.) or regular moisture (*Mortierella bainieri, Drechslera australiensis,* and *Trichotecium roseum*). Other fungi detected in filter beds are *Circinella* sp., *Cephalotecium* sp., and *Stemphilium* sp. (41). Although the contribution of fungi in biofiltration is not so clear, it is likely that an important role is played in the degradation of organic nitrogen, sulfur, and potassium, such as takes place in the soil.

As far as the influence of the support structure on microbial density inside the bed is concerned, it was demonstrated that compost-based filters usually contain much higher population densities of these microorganisms than does soil (43). On the other hand, it seems that concentrations of bacteria cells and fungi spores in the treated emissions are only slightly higher than those observed in open air (43,44). Nevertheless, Liebe and Herbig advised against the risk of relevant fungi-spore emissions during biofilter maintenance and support-replacement operations (45).

A great number of other bacterial strains have been inoculated and tested in bench-scale biofilters to treat waste gases polluted by specific compounds, most of which appear to be particularly promising for future development of this technology (see "New Developments and Potentials").

Metabolic Pathways for Organic Pollutant Degradation

Different metabolic pathways are employed by microorganisms to degrade a variety of organic pollutants.

Aliphatic Hydrocarbons. Several bacteria (mainly belonging to the genera *Nocardia*, *Pseudomonas*, and *Mycobacterium*), and some yeast and fungi, which are capable of utilizing saturated aliphatic hydrocarbons for their growth, are the main candidates for new applications in bioscrubbers and biofilters for the removal of such air pollutants. The oxidation of the terminal carbon to an alcohol, catalyzed by a monoxygenase, is responsible for the introduction in the hydrocarbon structure of only one of the atoms of molecular oxygen, which acts as a cosubstrate, while the other is reduced to water:

$$\text{R-CH}_3 + \text{NADH} + \text{H}^+ + \text{O:O} \xrightarrow{\text{monoxygenase}} \text{R-CH}_2\text{-OH} + \text{NAD}^+ + \text{H}_2\text{O}$$

Further oxidations by NAD^+-dependent alcohol and aldehyde dehydrogenases lead to a carboxylic acid that can be decomposed to acetyl-CoA by the β-oxidation route (46):

$$\text{R-CH}_2\text{-OH} + \text{NAD}^+ \xrightarrow{\text{alcohol dehydrogenase}} \text{R-CH=O} + \text{NADH} + \text{H}^+$$

$$\text{R-CH=O} + \text{NAD}^+ \xrightarrow{\text{aldehyde dehydrogenase}} \text{R-COOH} + \text{NADH} + \text{H}^+$$

$$\text{R-COOH} + \text{ATP} + \text{CoA} \xrightarrow{\beta\text{-oxidation}} \text{CH}_3\text{-CO-CoA} + \text{AMP} + \text{PP}_i$$

On the other hand, polychlorinated hydrocarbons can be cometabolized in the presence of isoprene-degrading bacteria (47).

Aromatic Compounds. The aerobic degradation of aromatics is only possible if the aromatic ring undergoes an enzymatic cleavage. Benzene can be degraded by several microorganisms by two divergent pathways after preliminary dioxygenation to catechol: the so-called *ortho*-pathway, utilizing the intradiol cleavage catalyzed by the catechol-1,2-dioxygenase, and the so-called *meta*-pathway, following an extradiol cleavage. Mixtures of benzene, toluene, and xylene are preliminarily dioxygenated, decomposed mainly via the *meta*-pathway, and finally metabolized through the citric acid cycle (47).

The presence of an alkyl group in the benzene ring gives the microbe the opportunity to attack the compounds either on the side chain or the ring. In arenes with short side chains (up to seven carbons), both alternative routes are possible, mainly due to the widespread presence of transformable plasmids. In particular, the plasmid TOL (48), containing catabolic operons for both alkyl-group oxidation and ring cleavage via *meta*-pathway, gives several strains of *Pseudomonas* the ability to degrade toluene, *m*- and *p*-xylenes, and other monoalkylbenzenes (49,50). On the contrary, if the side chain is sufficiently long (more than seven carbons), its metabolization by initial side-chain attack via ω- and β-oxidation (51), provides the cells of different species with enough energy for growth, and then the ring cleavage is not necessary.

Of the dialkylbenzenes, *m*- and *p*-xylenes can be biode-

graded by certain strains of *Pseudomonas* containing the TOL plasmid by preliminary oxidation of one of the methyl groups to a carboxylic group and methylcatechol (52), followed by *meta*-cleavage. Members of *Nocardia* are able to cometabolize *m*- and *p*-xylenes via the *ortho*-pathway, and *o*-xylene via the *meta*-pathway in the presence of alkanes as carbon sources (53). More recently, it has also been demonstrated that *o*-xylene can be metabolized as the sole carbon and energy source, via 3,4-dimethylcatechol and subsequent *meta*-cleavage, by some strains of *Pseudomonas stutzeri* (54) and *Corynebacterium* (55).

There are reports of different bacteria (*Pseudomonas* sp., *Xanthobacter* sp., etc.) able to grow on styrene (56) and methylstyrenes (57) as sole carbon and energy sources, during which oxidation of the aromatic nucleus and subsequent *meta*-cleavage probably take place.

In several bacteria, after a preliminary monoxygenation to a catecholic structure, phenol is degraded (58) through the *meta*-pathway to directly give intermediates of the citric acid cycle (46). While for 4-chlorophenol degradation the chlorinated lactonic intermediate is directly dechlorinated before its introduction into the cycle, for 4-methylphenol a specific enzyme catalyzes the isomerization of the related intermediate into a more easily metabolizable lactone (47).

Biphenyl and Fused-Ring Hydrocarbons. Although the biodegradation of biphenyl, naphthalene, and polycyclic aromatic hydrocarbons (PAH) is a subject of less relevance for biofiltration applications given their solid state at ambient temperature, their high toxicity has brought lively interest in the actual possibility of purifying low contaminated off-gases by utilizing biofilters inoculated with liquid suspensions of strains specifically adapted.

Apart some subtle differences, biphenyl is biodegraded through a pathway common to many different bacteria belonging to the genera *Pseudomonas* (59) and *Nocardia* (60), including a preliminary dioxygenation of the ring in the C2 and C3 positions to give a catechol-type intermediate, which is subsequently *meta*-cleaved and catabolyzed via 2-oxopenta-4-enoate and benzoate (60).

Naphthalene is degraded mainly by members of the genus *Pseudomonas* via dioxygenation to 1,2-dihydroxynaphthalene, followed by extradiol cleavage of the ring between C1 and C9, whereas PAHS, which are potential carcinogens, mutagens, and tetragens, are dioxygenated and cleaved in two different ways, according to their phenanthrenic or anthracenic structures (53). Less detail is available for higher molecular weight PAHs.

ECONOMIC CONSIDERATIONS

Comparing cost data quoted in the literature for biofiltration systems is very difficult because of the use of different currencies, the fluctuations of their reciprocal rates, the cost variability with time, the different units employed, and so on. Nevertheless, there is no doubt of biofilter economic convenience when compared with other traditional abatement techniques, due to the following advantages (6): minimum operational input, reduced production and disposal of by-products, easy start-up procedures, operational stability at steady state, low operative temperature, low cost of materials, simple building technologies, simple control systems, and low power requirements.

From the comparison of different biofilters in use in Europe (1,37,61–63), total operating and maintenance costs ranging from \$0.2 to 0.5 per 1,000 m^3 of waste gas can be estimated, which includes depreciation and interest, water consumption, replacement of filter materials and personnel, wastewater treatment, and energy costs for compressors. Costs reductions showing a half of these amounts for United States installations do not include expenses for support replacement and reflect the lower cost of electricity (1).

Capital costs are often negatively influenced by several factors, among which high pollutant concentration in the waste gas, recalcitrant nature of the components, necessity of gas pretreatment, reduced waste-gas availability, lack of space near the polluting source, and transport. Since the transport costs are strongly dependent on the country and the nature of the site, a cost comparison is only possible if this factor is not considered. On these bases, capital costs for open single-bed filters are in the range between \$300 and \$1,000 per m^2 of filter material in Europe (64,65) and between \$600 and \$1,000 per m^2 in the United States (1). These costs would increase up to \$3,800–5,700 per m^2 if lack of space would force the build-up of the filter on the plant cover, because of the additional costs of personnel, piping, and compression (66), and up to \$1,000–5,000 per m^2 if particular meteorological situations (snow, frequent atmospheric precipitation, etc.) require an upper covering or enclosed systems (1). Multiple-bed systems seem to be twice as expensive as the single-bed ones.

The filling material may consist, according to the pollutants, of heather, peat, and earth, often mixed with other inert materials necessary to minimize head losses. Because of these variables and the market situation, costs may fluctuate remarkably from \$60 to \$700 per m^3 (65,66).

Taking odor abatement as a basis for comparison of different techniques, biofiltration is by far the most convenient process in terms of space availability, resulting in, as a function of waste-gas composition, total cost savings of 15–30%, 45–70%, 50–75%, and 80% with respect to bioscrubbing, chemical scrubbing, adsorption, and ozonation, respectively (17,65,67). The success of biofiltration appears to be the result of the very low incidence of running costs to total costs as compared to other techniques. Extending the above comparison also to incineration, it has been estimated that biofiltration allows total cost savings of 60–75%, and investment and operational cost savings of 40–75% and 70–80%, respectively (17,63).

NEW DEVELOPMENTS AND POTENTIALS

Pollutants' degradability often depends on their origin and source: although xenobiotics, which are characterized by unnatural structures, cannot be degraded (persistent compounds), or can be degraded only with difficulty (recalcitrant compounds), the biogenic ones are easily biodegradable. For this reason, much research effort is being spent on the isolation, selection, and construction of different mi-

croorganisms (27,69,70) able to grow on difficult carbon sources in pure cultures, at rates comparable with those usually encountered on common substrates (Table 4).

For these reasons, it is evident that future developments and new applications of this technology strongly depend on the actual possibility of utilizing these microbial strains in pure cultures or selected consortia under the peculiar conditions, as present in biofilters.

Table 4. Xenobiotic Compounds and Microorganisms Able to Degrade Them in Pure Cultures

Pollutant	Microorganism
Chlorinated hydrocarbons (27,68)	
Methyl chloride	*Hyphomicrobium* sp.
Dichloromethane	*Pseudomonas* DM1
	Methylobacter DM111
	Hyphomicrobium sp.
1,2-Dichloroethane	*Xanthobacter* GJ10
Vinylchloride	*Mycobacterium* L1
Epichlorohydrine	*Pseudomonas* AD1
Chlorobenzene	*Pseudomonas* WR 1306
1,2-Dichlorobenzene	*Pseudomonas* GJ 60
1,3-Dichlorobenzene	*Pseudomonas* sp.
1,4-Dichlorobenzene	*Alcaligenes* A175
Arenes (69)	
Benzene	*Pseudomonas* sp.
	Acinetobacter calcoaceticus RJE 74
Toluene	*Pseudomonas* sp.
	Pseudomonas putida
	Bacillus sp.
Monoalkylbenzenes	*Pseudomonas* sp. NCIB 10643
	Pseudomonas fluorescens
	Pseudomonas putida RE204
	Acinetobacter Iwoffi
m-,*p*-Xylene	*Pseudomonas* sp.
	Nocardia sp.
o-Xylene	*Pseudomonas stutzeri*
	Corynebacterium sp. C125
Styrene	*Xanthobacter* sp 124X
	Pseudomonas putida
α-Methylstyrene	*Pseudomonas* sp.
Biphenyl (69)	
	Pseudomonas sp. NCIB 10643
	Nocardia sp. NCIB 10503
	Pseudomonas cruciviae
	Pseudomonas putida
	Pseudomonas pseudoalcaligenes
	Pseudomonas aeruginosa
	Escherichia coli
(Fused-ring compounds) (69)	
Naphthalene	*Pseudomonas* sp.
	Pseudomonas fluorescens
	Pseudomonas putida
PAH	*Pseudomonas* sp.
	Pseudomonas paucimobilis
	Pseudomonas vesicularis
	Algaligenes denitrificans

CONCLUDING REMARKS

As pointed out here, biofiltration has become by now a well-recognized technology for the control of odors and the elimination of air contaminants from waste gases, and represents a new generation of air pollution control technologies.

Over the past two decades, thanks to the considerable progress made in the microbiological and process-engineering fields, biofiltration has gained the confidence of many of industries and has found increasingly widespread application. The suitability of this technique for the treatment of a wide range of inorganic and organic pollutants has been proven. Removal efficiencies as high as 90% or more are usually achieved for common pollutants, such as alcohols, ethers, aldehydes, ketones, and monoaromatic volatile chemicals.

The operating cost savings, the technological simplicity of the plants, the reduced maintenance and energy requirements, and the minimal generation of by-products that must be disposed of, make biofiltration, with respect to other air pollution control technologies, an attractive cost-effective alternative, particularly if applied to large volumes of gaseous-waste streams containing readily biodegradable contaminants in relatively low concentrations.

Although its application to different compounds has been demonstrated, considerable research is being undertaken to find appropriate strains able to degrade recalcitrant compounds, in particular, most of which are xenobiotics. In addition, new strategies should be developed to improve the control of essential operating parameters affecting biofilter performance and to optimize reactor design and performance, in order to widen the application field of biofilters.

BIBLIOGRAPHY

1. G. Leson and A.M. Winer, *J. Air Waste Manage. Assoc.* **41**, 1045–1054 (1991).
2. H.L. Bohn, *J. Air Pollut. Control. Assoc.* **25**, 953–955 (1975).
3. H.L. Bohn, *Chem. Eng. Progr.* **88**, 34–40 (1992).
4. W.H. Prokop and H.L. Bohn, *J. Air Pollut. Control. Assoc.* **35**, 1332–1338 (1985).
5. D.H. Kampbell, J.T. Wilson, H.W. Read, and T.T. Stocksdale, *J. Air Pollut. Control. Assoc.* **37**, 1236–1240 (1987).
6. F. Alfani, L. Cantarella, A. Gallifuoco, and M. Cantarella, *Acqua-Aria* **10**, 877–884 (1990).
7. B. Jäger and J. Jager, *Müll und Abfall* **2**, 48–54 (1978).
8. D.K.B. Thistlethwayte, B. Hardwick, and E.E. Goleb, *Chimie Ind.* **106**, 795–801 (1973).
9. M. Gust, H. Grochowski, and S. Schirz, *Staub-Reinh. Luft* **39**, 397–402 (1979).
10. R. Helmer, *Ges. Ing.* **94**, 21–30, (1974).
11. H. Hartmann, *Stuttg. Ber. Siedlungswasserwirtsch* **59**, 3–19 (1977).
12. S. Cernuschi and V. Torretta, *Ingegneria Ambientale* **25**, 248–264 (1996).
13. S.P.P. Ottengraf, J.J.P. Meesters, A.H.C. van den Oever, and H.R. Rozema. *Bioprocess Eng.* **1**, 61–69 (1986).
14. S.P.P. Ottengraf, A.H.C. van den Oever, and F.J.C.M. Kempenaars, in E.H. Houwink and R.R. van der Meer eds., *Inno-*

vations in Biotechnology, Elsevier, Amsterdam, 1984, pp. 157–167.

15. F.J. Weber and S. Hartmans, in A.J. Dragt and J. van Ham eds., *Biotechniques for Air Pollution Abatement and Odour Control Policies*, Elsevier, Amsterdam, 1992, pp. 125–130.
16. W.J. Swanson and R.C. Loeher, *J. Environ. Eng.* **123**, 538–546 (1997).
17. G. Leson and B.J. Smith, *J. Environ. Eng.* **123**, 556–563 (1997).
18. E. Eitner, in Verein Deutscher Ingenieure ed., *VDI Berichte 735: Biologische Abgasreinigung*, VDI Verlag, Düsseldorf, 1989, pp. 191–214.
19. H.L. Bohn, *86th Mtg. Air Waste Management Assoc.*, Denver, Col., June 14–18, 1993.
20. S.P.P. Ottengraf, in H.J. Rehm and G. Reed eds., *Biotechnology*, vol. 8, VCH, Weinheim, 1986, pp. 425–452.
21. J. Monod, *Ann. Rev. Microbiol.* **3**, 371–394 (1949).
22. S.P.P. Ottengraf and A. van den Oever, *Biotechnol. Bioeng.* **19**, 1411–1417 (1986).
23. R. Marsh, *Inst. Chem. Eng.* **3**, 13.1–13.14 (1992).
24. F. Sabo, U. Motz, and K. Fischer, *86th Mtg. Air Waste Management Assoc.*, Denver, Col., June 14–18, 1993.
25. J.S. Devinny and D.S. Hodge, *J. Air Waste Manage. Assoc.* **45**, 125–131 (1995).
26. Kommission Reinhaltung der Luft, in Verein Deutscher Ingenieure ed., *VDI 3477-Handbuch Reinhaltung der Luft: Biological Waste Gas / Waste Air Purification, Biofilters*, VDI Verlag, Düsseldorf, 1991, Band 6, pp. 1–29.
27. S.P.P. Ottengraf and R.M.M. Diks, in A.J. Dragt and J. van Ham eds., *Biotechniques for Air Pollution Abatement and Odour Control Policies*, Elsevier, Amsterdam, 1992, pp. 17–31.
28. A.J. Dragt and S.P.P. Ottengraf, *Proc. Conf. on Advanced Technologies and Biotechnology for the Pollution Prevention and Environmental Protection*, SIBA-ISER, Genoa, Italy, 1987, pp. 243–258.
29. M.A. Deshusses, *J. Environ. Eng.* **123**, 563–568 (1997).
30. M. Pescari and V. Pizzirani, *Microsymposium on Air Pollution Control*, Pavia, Italy, May 23, 1997.
31. I. Trentini and R. Vismara, *Biologi Italiani*, **28**, 16–26 (1998).
32. A.P. Togna and S. Frisch, *86th Mtg. Air Waste Management Assoc.*, Denver, Col., June 14–18, 1993.
33. L. Benedusi, G. Caretta, C. Collivignarelli, C. Dacarro, G. Del Frate, and V. Riganti, *Ingegneria Ambientale* **22**, 271–284 (1993).
34. S.J. Ergas, K. Kinney, M.E. Fuller, and K.M. Scow, *Biotechnol. Bioeng.* **44**, 1048–1054 (1994).
35. D.B. Janssen, R. Oldenhuis, and A.J. van den Wijngaard, in Verein Deutscher Ingenieure ed., *VDI Berichte 735: Biologische Abgasreinigung*, VDI Verlag, Düsseldorf, 1989, p. 25.
36. S.P.P. Ottengraf and A.H.C. van den Oever, *Biotechnol. Bioeng.* **25**, 3089–3102 (1983).
37. K.F. Kosky and C.R. Neff, *NWWA / API Conf. on Petroleum Hydrocarbons and Organic Chemicals in Ground Water*, Houston, Tex., November 9–11, 1988.
38. H. van Langenhove, E. Wuyts, and N. Schamp, *Water Res.* **20**, 1471–1476 (1986).
39. C. Kyeoung-Suk, H. Mitsuyo, and S. Makoto, *Appl. Environ. Microbiol.* **58**, 1183–1189 (1992).
40. R. Helmer, *Stuttgarter Berichte zur Siedlungswasserwirtschaft*, Band 49, Oldenbourg, München, 1972.
41. W. Ludwig, O. Fischer, and F.X. Kneer, *Chem. Ing. Tech.* **58**, 742 (1986).
42. M. Wainwright, *Trans. Br. Mycol. Soc.* **90**, 159–170 (1988).
43. S.P.P. Ottengraf and J.H.G. Konings, *Bioprocess Eng.* **7**, 89–96 (1991).
44. R.M.M. Diks and S.P.P. Ottengraf, in Verein Deutscher Ingenieure ed., *VDI Berichte 735: Biologische Abgasreinigung*, VDI Verlag, Düsseldorf, 1989, p. 7.
45. H.G. Liebe and V. Herbig, *Bericht II*, 1.1-50441-5/132, Umweltbundesamt, Berlin, May, 1986.
46. T.D. Brock, M.D. Madigan, J.M. Martinko, and J. Parker, *Microbiologia*, Città Studi Edizioni, Milan, 1995.
47. K.H. Engesser, in A.J. Dragt and J. Van Ham eds., *Biotechniques for Air Pollution Abatement and Odour Control Policies*, Elsevier Amsterdam, 1992, pp. 33–40.
48. R.C. Bayly and M.G. Barbour, in D.T. Gibson ed., *Microbial Degradation of Organic Compounds*, Marcel Dekker, New York, 1984, pp. 253–294.
49. S.F. Vecht, M.W. Platt, Z. Er-El, and I. Goldberg, *Appl. Microbiol. Biotechnol.* **27**, 587–592 (1988).
50. R.W. Eaton and K.N. Timmis, *J. Bacteriol.* **168**, 123–131 (1986).
51. O.O. Amund and I.J. Higgins, *Antonie van Leewenhoek* **51**, 45–56 (1985).
52. J.F. Davey and D.T. Gibson, *J. Bacteriol.* **119**, 923–929 (1974).
53. D.T. Gibson and V. Subramanian, in D.T. Gibson ed., *Microbial Degradation of Organic Compounds*, Marcel Dekker, New York, 1984, pp. 361–369.
54. G. Baggi, P. Barbieri, E. Galli, and S. Tollari, *Appl. Environ. Microbiol.* **53**, 2129–2132 (1987).
55. G. Schraa, B.M. Bethe, A.R.W. van Neerven, W.J.J. van den Tweels, E. van der Wende, and A.J.B. Zehnder, *Antonie van Leewenhoek* **53**, 159–170 (1987).
56. S. Hartmans, M.J. van der Werf, and J.A.M. de Bont, *Appl. Environ. Microbiol.* **56**, 1347–1351 (1990).
57. D.B. Dzhusupova, B.P. Baskunov, L.A. Golovleva, R.M. Alieva, and A.N. Ilyaletdinov, *Mikrobiologiya* **54**, 136–140 (1985).
58. M. Zilli, A. Converti, A. Lodi, M. Del Borghi, and G. Ferraiolo, *Biotechnol. Bioeng.* **41**, 693–699, (1993).
59. T. Omori, H. Ishigooka, and Y. Minoda, *Agric. Biol. Chem.* **50**, 1513–1518 (1986).
60. M.R. Smith and C. Ratledge, *Appl. Microbiol. Biotechnol.* **30**, 395–401 (1989).
61. J.A. Don, in Verein Deutscher Ingenieure ed., *VDI Berichte 561*, VDI Verlag, Düsseldorf, 1986, pp. 63–74.
62. B. Hippchen, *Mikrobiologische Untersuchungen zur Eliminierung Organischer Lösungmittel im Biofilter*, Stuttgarter Berichte zur Siedlungswasserwirtschaft, Band 94, Oldenbourg, München, 1985.
63. P.G. Maurer, *BMFT Forschungsbericht T*, 79–114 (1979).
64. L. Rodhe, L. Thyselius, and U. Berglung, Report N. 76 Jordbrukstekniska Intitutet, Uppsala, Sweden, (Joint JTI/AFRC Engineering, Silsoe, translation N. 1, New Series), 1988.
65. C.C. Pearson, V.R. Phillips, G. Green, and I.M. Scotford, in A.J. Dragt and J. van Ham eds., *Biotechniques for Air Pollution Abatement and Odour Control Policies*, Elsevier, Amsterdam, 1992, pp. 245–254.
66. V. Torretta, *Ingegneria Ambientale* **25**, 379–389 (1996).
67. J.L. Blitz and A.F. Elvidge, *Environ. Protect. Bull.* **16**, 3–6 (1992).
68. G. Stucki, R. Galli, H. Ebersold, and T. Leisinger, *Arch. Microbiol.* **130**, 366–371 (1981).
69. M.R. Smith, *Biodegradation* **1**, 191–206 (1990).

BIOREACTORS, AIR-LIFT REACTORS

J.C. MERCHUK
M. GLUZ
Ben-Gurion University of the Negev
Beer-Sheva, Israel

KEY WORDS

Bubble column
Fluid dynamics
Gas hold-up
Heat transfer
Liquid flow
Mass transfer
Scale-up
Three-phase airlift reactors

OUTLINE

INTRODUCTION

General

The term *airlift reactor* (ALR) covers a wide range of gas–liquid or gas–liquid–solid pneumatic contacting devices that are characterized by fluid circulation in a defined cyclic pattern through channels built specifically for this purpose. In ALRs, the content is pneumatically agitated by a stream of air or sometimes by other gases. In those cases, the name *gas lift reactors* has been used. In addition to agitation, the gas stream has the important function of facilitating exchange of material between the gas phase and the medium; oxygen is usually transferred to the liquid, and in some cases reaction products are removed through exchange with the gas phase.

The main difference between ALRs and bubble columns (which are also pneumatically agitated) lies in the type of fluid flow, which depends on the geometry of the system. The bubble column is a simple vessel into which gas is injected, usually at the bottom, and random mixing is produced by the ascending bubbles. In the ALR, the major patterns of fluid circulation are determined by the design of the reactor, which has a channel for gas–liquid upflow—the riser—and a separate channel for the downflow (Fig. 1). The two channels are linked at the bottom and at the top to form a closed loop. The gas is usually injected near the bottom of the riser. The extent to which the gas disengages at the top, in the section termed the *gas separator*, is determined by the design of this section and the operating conditions. The fraction of the gas that does not disengage, but is entrapped by the descending liquid and taken into the downcomer, has a significant influence on the fluid dynamics in the reactor and hence on the overall reactor performance.

Airlift Reactor Morphology

Airlift reactors can be divided into two main types of reactors on the basis of their structure (Fig. 1): (1) external-loop vessels, in which circulation takes place through separate and distinct conduits; and (2) baffled (or internal-loop) vessels, in which baffles placed strategically in a single vessel create the channels required for the circulation. The designs of both types can be modified further, leading to variations in the fluid dynamics, in the extent of bubble disengagement from the fluid, and in the flow rates of the various phases.

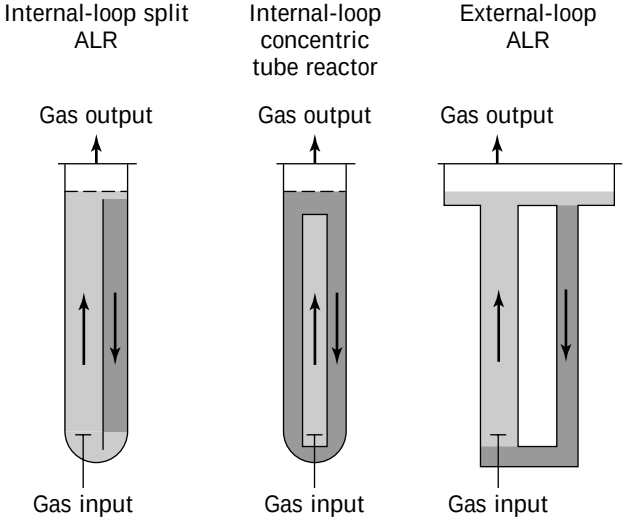


Figure 1. Different types of ALRs.

All ALRs, regardless of the basic configuration (external loop or baffled vessel), comprise four distinct sections with different flow characteristics:

- *Riser*. The gas is injected at the bottom of this section, and the flow of gas and liquid is predominantly upward.
- *Downcomer*. This section, which is parallel to the riser, is connected to the riser at the bottom and at the top. The flow of gas and liquid is predominantly downward. The driving force for recirculation is the difference in mean density between the downcomer and the riser; this difference generates the pressure gradient necessary for liquid recirculation.
- *Base*. In the vast majority of airlift designs, the bottom connection zone between the riser and downcomer is very simple. It is usually believed that the base does not significantly affect the overall behavior of the reactor, but the design of this section can influence gas holdup, liquid velocity, and solid phase flow (1,2).
- *Gas separator*. This section at the top of the reactor connects the riser to the downcomer, facilitating liquid recirculation and gas disengagement. Designs that allow for a gas residence time in the separator that is substantially longer than the time required for the bubbles to disengage will minimize the fraction of gas recirculating through the downcomer (Fig. 2).

Momentum, mass transfer, and heat transfer will be different in each section, but the design of each section may influence the performance and characteristics of each of the other sections, since the four regions are interconnected.

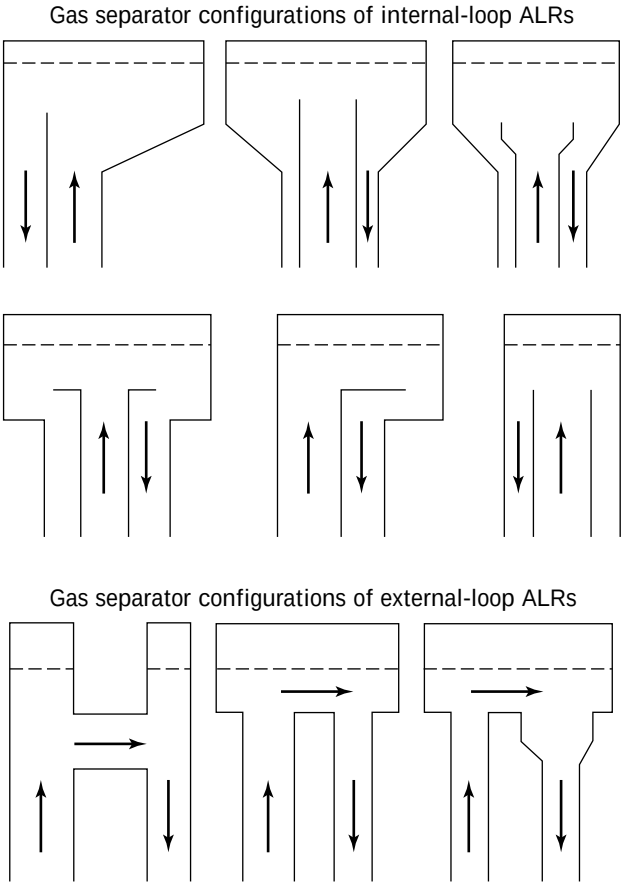


Figure 2. Different types of gas separators.

Advantages of Airlift Bioreactors

For the growth of microorganisms, ALRs are considered to be superior to traditional stirred-tank fermenters despite the fact that the conventional fermenters provide the major requirements for culturing microorganisms: gas–medium interface for the supply of oxygen and the removal of waste gases; means of agitation to ensure proper nutrient distribution and to minimize damage resulting from addition of concentrated acid or base (for pH control); means of heat transfer (for temperature control); and a contamination-free environment. Therefore, the reason for the more successful growth reported in ALRs (3,4) appears to lie in the difference in the fluid dynamics between ALRs and the more conventional fermenters. In conventional stirred tanks or bubble columns, the energy required for the movement of the fluids is introduced focally, at a single point in the reactor, via a stirrer or a sparger, respectively. Consequently, energy dissipation is very high in the immediate surroundings of the stirrer and decreases away from it toward the walls. Similarly, shear will be greatest near the stirrer (5), since the momentum is transferred directly to the fluid in that region (6), which, in turn, transfers this energy to the slower-moving, more distant elements of the fluid. This results in a wide variation of shear forces; for example, the maximum shear gradient in a stirred tank with a flat-blade turbine has been reported to be approximately 14 times the mean shear gradient (7).

Cells in culture may thus be exposed to contrasting environments in a mechanically stirred vessel, either to minimal shear forces that may generate potentially undesirable gradients in temperature and in substrate, metabolite, and electrolyte concentrations or, alternatively, to highly turbulent zones, with no problems of heat or mass transfer, but with very high shear gradients that may endanger cell integrity or exert some influence on cell morphology and metabolism (6). Changes in the morphology of microorganisms associated with high shear forces in the medium have frequently been observed (8–10). The nature of the relationship between such morphological changes and the rates of growth and metabolite production is still not properly understood, although it may be of great importance in the design and scale-up of bioreactors.

In ALRs, as in bubble columns, the gas is injected at a single point, but in ALRs the direct contribution of gas injection to the dynamics of the system is small; circulation of liquid and gas is facilitated by the difference in gas holdup between the riser and the downcomer, which creates a pressure difference at the bottom of the equipment:

$$\Delta P_b = \rho_L g(\varphi_r - \varphi_d) \quad (1)$$

where ΔP_b is the pressure difference, ρ_L is the density of

the liquid (the density of the gas is considered to be negligible), g is the gravitational constant, and φ_r and φ_d are the fractional gas holdup of the riser and downcomer, respectively. The pressure difference forces the fluid from the bottom of the downcomer toward the riser, generating fluid circulation in the ALR. Since φ_r and φ_d are both average values integrated along the height of the reactor, it follows that there are no focal points of energy dissipation and that shear distribution is homogeneous throughout the ALR. There is thus a relatively constant environment, with minimization of sharp changes in the mechanical forces acting on suspended particles. Because good mixing is required, shear forces cannot be avoided completely. One of the most critical points is the bottom, where there is a sharp 180° turn.

Shear-sensitive mammalian and plant cells in culture should benefit from such an environment. Currently, the research and development of new bioreactors for mammalian cells is indeed focusing on the issue of shear-related damage to suspended cells (8,11–24).

Mammalian and plant cells in culture are more susceptible than microorganisms to the reactor conditions. Mammalian cells, which lack the rigid cell wall of microorganisms, have a larger size (one order of magnitude) than microorganisms and are very sensitive to mechanical stress. Plant cells have a rigid cellulose wall, but they are also much larger than microorganisms (usually by about an order of magnitude) and are therefore also sensitive to reactor conditions. Kolmogoroff's model of isotropic turbulence (25) indicates that serious damage may occur at relatively large values of the length scale. The last length is a parameter of the model and indicates the size of the eddy where energy starts to be dissipated by viscous resistance. Indeed, it has been observed that plant cells, in spite of their rigid wall, are shear-sensitive, and difficulties have been found in stirred-tank cultures. This is especially true when large-scale systems are considered. Although high agitation rates may be detrimental to cell growth, low agitation rates lead to an increase in the number and size of cell aggregates—also an undesirable phenomenon. The aggregates are formed as a result of daughter cells failing to separate after division and as a consequence of the stickiness of the polysaccharides excreted by the cells, especially at the end of the growth phase. An optimal shear rate between these two extremes must be found for each culture.

It has recently been shown experimentally that velocity fluctuations related to turbulent shear are relatively homogeneously distributed in an ALR (26,27). The measurements of fluctuating velocity made by Tan et al. (26) show that the liquid turbulence in ALRs is homogeneously distributed in both the riser and the downcomer. It thus seems reasonable to assume that the homogeneity of the stress forces is the main advantage offered by ALRs and that this homogeneity is responsible for the success of shear-sensitive cultures in the ALR type of fermenter (3,28–31).

Another advantage of the ALR is the mechanical simplicity of the device. The absence of a shaft and of the associated sealing, which is always a weak element from the point of view of sterility, confers on the ALR an obvious advantage over agitated tanks. This consideration is especially important in processes involving slow-growing cultures, such as animal and plant cells, for which the risk of contamination is large.

All the points mentioned above are particularly relevant for sophisticated processes in which the product is usually of high value. But ALRs may be used also for processes involving low-value products, in which case efficiency of energy use may well become the key point for design, as in the use of ALRs for wastewater treatment (32). The superiority of ALRs over mechanically agitated contactors in terms of mass transfer rates for a given energy input has been demonstrated by Legrys (33). Comparison of the efficiency of oxygen transfer, that is, the mass of oxygen absorbed per unit energy invested and unit time, showed that the efficiency of the ALR is among the highest in agitated systems (32). The ALRs are particularly suited to processes with changing oxygen requirements because aeration efficiency and performance are relatively insensitive to changes in operating conditions. Performance decreases markedly in mechanically stirred systems as the energy input (or oxygen transfer rate) increases, but it is quite constant in ALRs (34) (Fig. 3).

The efficiency of ALRs decreases relatively slowly as the energy input per unit volume of reactor is increased, as is shown in Figure 4 (32). In contrast, in the operation of stirred tanks, the mass transfer rate can be easily increased by increasing the power input, but this improvement is achieved at the cost of a considerable decrease in the efficiency of oxygen transfer. This decrease may constitute a crucial disadvantage in a process like wastewater treatment, where the energy input is an important element in the cost of the final product and flexibility of operating conditions is required because of the constant change of feed composition and flow rate.

Energy economy in the ALR may be improved by placing a second sparger in the upper part of the downcomer (32,35,36). If the liquid velocity is greater than the free rising velocity of the bubbles generated, the gas is carried down, resulting in a longer contact time between the bubble and the liquid. This diminishes the energy requirements, since part of the gas is injected against a lower hydrostatic pressure.

The advantages described above counterbalance the obvious disadvantage of ALRs, which is the requirement for a minimum liquid volume for proper operation. Indeed, the changes in liquid volume in these reactors are limited to the region of the gas separator, since the liquid height must always be sufficient to allow liquid recirculation in the reactor and must therefore be above the separation between the riser and the downcomer.

FLUID DYNAMICS

The interconnections between the design variables, the operating variables, and the observable hydrodynamic variables in an ALR are presented diagramatically in Figure 5 (37). The design variables are the reactor height, the riser-to-downcomer area ratio, the geometrical design of the gas separator, and the bottom clearance (C_b, the distance between the bottom of the reactor and the lower end of the draft tube, which is proportional to the free area for

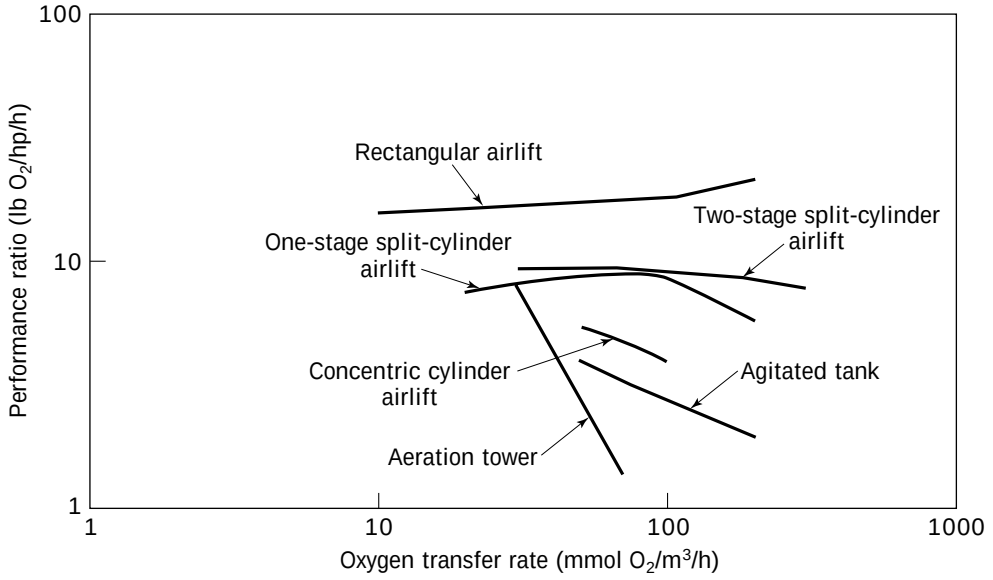


Figure 3. Performance ratio as a function of oxygen-transfer rate, showing that aeration efficiency and performance are relatively insensitive to changes in operating conditions in different types of ALRs (1–5) versus an agitation tank (6) and an aeration tower (7). Adapted from Orazem and Erickson (34).

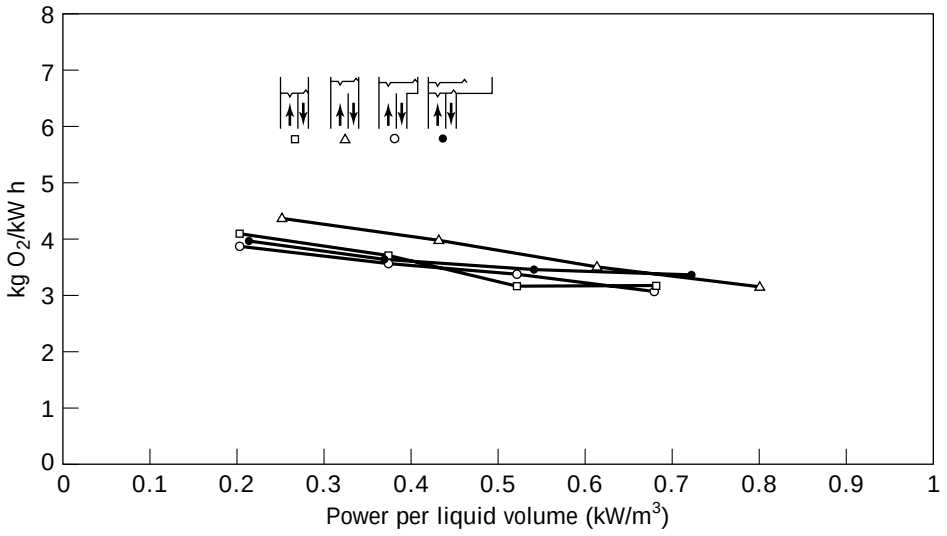


Figure 4. Aeration efficiency as a function of pneumatic power of gas input per unit volume in a straight-baffle ALR. The level indicated corresponds to no-aeration conditions. Adapted from Siegel et al. (32).

flow in the bottom and represents the resistance to flow in this part of the reactor). The main operating variables are primarily the gas input rate and, to a lesser extent, the top clearance (C_t, the distance between the upper part of the draft tube and the surface of the nonaerated liquid). These two independent variables set the conditions that determine the liquid velocity in the ALR via the mutual influences of pressure drops and holdups, as shown in Figure 5 (37). Viscosity is not shown in Figure 5 as an independent variable because in the case of gas–liquid mixtures, it is a function of the gas holdup (and of liquid velocity in the case of non-Newtonian liquids), and because in a real process, it will change with time due the changes in the composition of the liquid.

Flow Configuration

Riser. In the riser, the gas and liquid flow upward, and the gas velocity is usually larger than that of the liquid. The only exception is homogeneous flow, in which case both phases flow at the same velocity. This can happen only with very small bubbles, in which case the free-rising velocity of the bubbles is negligible with respect to the liquid velocity. Although about a dozen different gas–liquid flow configurations have been developed (38), only two of them are of interest in ALRs (39,40):

1. Homogeneous bubbly flow regime, in which the bubbles are relatively small and uniform in diameter and turbulence is low

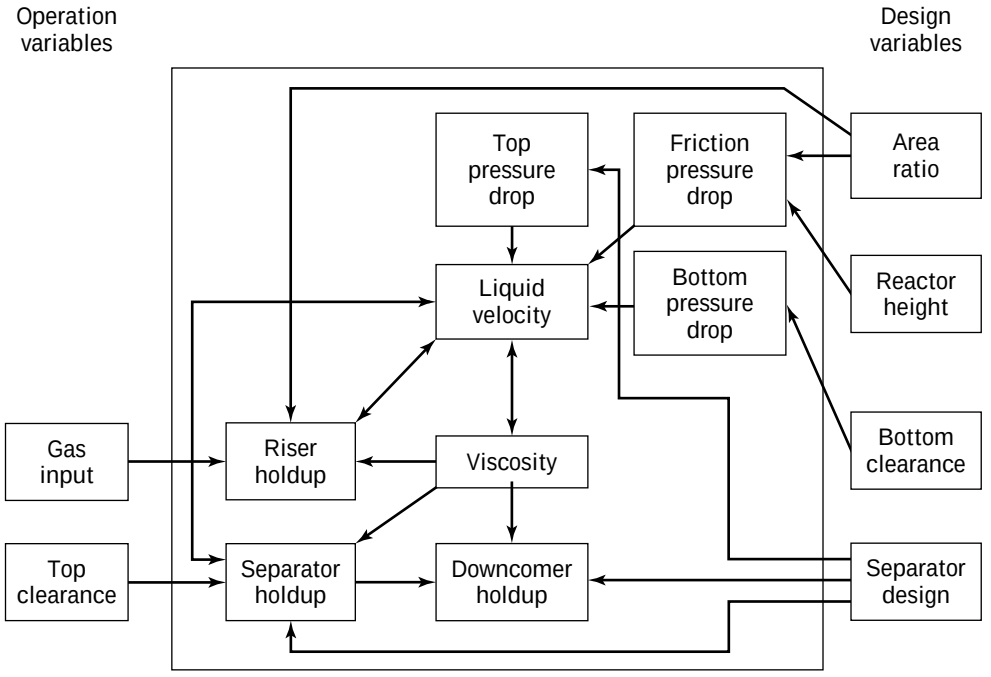


Figure 5. Interaction between geometric and fluid dynamic variables in an ALR. Adapted from Merchuk et al. (37).

2. Churn-turbulent regime, in which a wide range of bubble sizes coexist within a very turbulent liquid

The churn-turbulent regime can be produced from homogeneous bubbly flow by increasing the gas flow rate. Another way of obtaining a churn-turbulent flow zone is by starting from slug flow and increasing the liquid turbulence, by increasing either the flow rate or the diameter of the reactor, as can be seen in Figure 6 (41). The slug-flow configuration is important only as a situation to be avoided at all costs, because large bubbles bridging the entire tower cross-section offer very poor capacity for mass transfer.

Downcomer. In the downcomer, the liquid flows downward and may carry bubbles down with it. For bubbles to be entrapped and flow downward, the liquid velocity must be greater than the free-rise velocity of the bubbles. At very low gas flow input, the liquid superficial velocity is low, practically all the bubbles disengage, and clear liquid circulates in the downcomer. As the gas input is increased, the liquid velocity becomes sufficiently high to entrap the smallest bubbles. Upon a further increase in liquid velocity larger bubbles are also entrapped. Under these conditions the presence of bubbles reduces the cross-section available for liquid flow, and the liquid velocity increases in this section. Bubbles are thus entrapped and carried downward, until the number of bubbles in the cross-section decreases, the liquid velocity diminishes, and the drag forces are not sufficient to overcome the buoyancy. This feedback loop in the downcomer causes stratification of the bubbles, which is evident as a front of static bubbles, from which smaller bubbles occasionally escape downward and larger bubbles, produced by coalescence, escape upward. The bubble front descends, as the gas input to the system is increased, until the bubbles eventually reach the bottom and recirculate to the riser. When this point is reached, the bubble distribution in the downcomer becomes much more uniform. This is the most desirable flow configuration in the downcomer, unless a single pass of gas is required. The correct choice of cross-sectional area ratio of the riser to the downcomer will determine the type of flow.

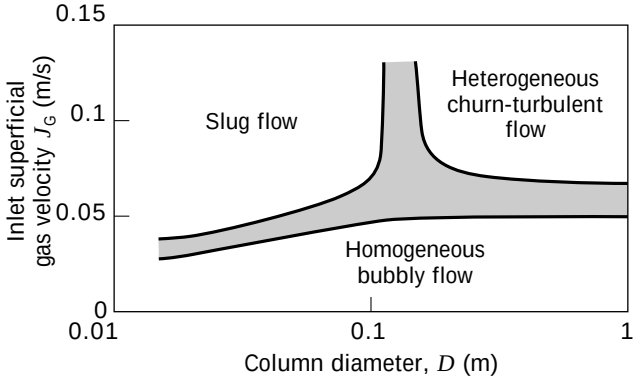


Figure 6. Map of flow configurations for gas–liquid concurrent flow in a vertical tube. Adapted with permission from Wiswanathan (40).

Gas Separator. The gas separator is often overlooked in descriptions of experimental ALR devices, although it has considerable influence on the fluid dynamics of the reactors. The geometric design of the gas separator will determine the extent of disengagement of the bubbles entering from the riser. In the case of complete disengagement, clear liquid will be the only phase entering the downcomer.

In the general case, a certain fraction of the gas will be entrapped and recirculated. Fresh gas may also be entrapped from the headspace if the fluid is very turbulent near the interface. The extent of this entrapment influences strongly gas holdup and liquid velocity in the whole reactor.

It is quite common to enlarge the separator section to reduce the liquid velocity and to facilitate better disengagement of spent bubbles. Experiments have been reported in which the liquid level in the gas separator was high enough to be represented as two mixed vessels in series (42,43). This point will be analyzed further in the section devoted to mixing.

Gas Holdup

Gas holdup is the volumetric fraction of the gas in the total volume of a gas–liquid–solid dispersion:

$$\varphi_i = \frac{V_G}{V_L + V_G + V_S} \tag{2}$$

where the subindexes L, G, and S indicate liquid, gas, and solid, and *i* indicates the region in which the holdup is considered, that is, gas separator (s) the riser (r), the downcomer (d), or the total reactor (T).

The importance of the holdup is twofold: (1) the value of the holdup gives an indication of the potential for mass transfer, since for a given system a larger gas holdup indicates a larger gas–liquid interfacial area; and (2) the difference in holdup between the riser and the downcomer generates the driving force for liquid circulation. It should be stressed, however, that when referring to gas holdup as the driving force for liquid circulation, only the total volume of the gas is relevant. This is not the case for mass-transfer phenomena, in this case, the interfacial area is of paramount importance, and therefore some information on bubble size distribution is required for a complete understanding of the process.

Because gas holdup values vary within a reactor, average values, referring to the whole volume of the reactor, are usually reported. Values referring to a particular section, such as the riser or the downcomer, are much more valuable, since they provide a basis for determining liquid velocity and mixing. However, such values are less frequently reported.

The geometric design of the ALR has a significant influence on the gas holdup. Changes in the ratio A_d/A_r, the cross-sectional areas of the downcomer and the riser, respectively, will change the liquid and gas residence time in each part of the reactor and hence their contributions to the overall holdup. Gas holdup increases with decreasing A_d/A_r (44–47).

Gas Holdup in Internal Airlift Reactors. Correlations presented for internal-loop ALRs are shown in Table 1. These take into account liquid properties and geometric differences within a particular design. Most of the correlations take the form:

$$\varphi_r = a(J_G)^{\alpha}\left(\frac{A_d}{A_r}\right)^{\beta}(\mu_{ap})^{\gamma} \tag{3}$$

where φ_r is the gas holdup in the riser, J_G is the superficial gas velocity (gas volumetric flow rate per unit of cross-sectional area), μ_{ap} is the effective viscosity of the liquid, and α, β, γ, and a are constants that depend on the geometry of the reactor and the properties of the liquid. The correlation can be used to predict the holdup in a system that is being designed or simulated as a function of the operating variables, the geometry of the system, or the liquid properties. Such correlations are effective for fitting data for the same type of reactor (e.g., a split-vessel reactor) with different area ratios or even different liquid viscosities, but they are mostly reactor-type specific.

The cyclic flow in the ALR complicates the analysis of the system. The riser gas holdup depends strongly on the geometric configuration of the gas–liquid separator and the water level in the gas separator. This has been shown experimentally in a split-vessel rectangular ALR (60), but the premise can essentially be extended to any internal-loop ALR. Analysis of the system revealed that these factors influence the gas disengagement and hence the gas recirculation in the downcomer. When this influence is taken into account and the holdup is plotted against the true gas superficial velocity, $J_{G,true}$, which is defined as the sum of the gas superficial velocity due to the freshly injected gas, Q_{in}, and to the recirculated gas, Q_d, that is,

$$J_{G,true} = \left(\frac{Q_{in} + Q_d}{A_r}\right) \tag{4}$$

then all the data for the different gas separators may be represented by a single relationship, such as equation 3. In other words, if the actual gas flow is known, the influence of gas recirculation (which depends on A_d/A_r and the design of the gas separator) has been already taken into account and does not need to be considered again. Nevertheless, this simple approach has a drawback in that the true gas superficial velocity is difficult to measure because the gas recirculation rate is usually not known. A method for evaluation of the extent of the maximum gas recirculation has recently been developed and will be discussed later in this article.

Thus, correlations that take into account all the variables, which may be easily measured, remain the option of choice. Table 1 shows most of the correlations of this type that have been proposed for the riser holdup in internal-loop ALRs. Comparison of a number of these correlations shows that there is reasonable agreement between the predictions of the different sources (Fig. 7).

Figure 7 can be used as an example of the actual state-of-the-art in ALR design. A number of correlations have been proposed, and three variables (A_d/A_r, μ_{ap}, and J_G) have been tested by most researchers. The ranges in which these variables were studied varies from source to source. In addition, some other variables (such as bottom clearance, top clearance or gas separator design, and surface tension) have been used by some authors but ignored by others. One example is the disengagement ratio defined by Siegel and Merchuk (64), which represents the mean horizontal path of a recirculating bubble relative to the external diameter and is equivalent to the parameter obtained by dimensional analysis (1) as:

$$M = \frac{D_S}{4D}$$

where D is the diameter of column and D_s the diameter of

Table 1. Gas Hold-Up in Internal-Loop ALR

No.	Formula	Ref.
1	$\varphi_r = 0.441 J_{Gr}^{0.841} \mu_{ap}^{-0.135}$ $\varphi_d = 0.297 J_{Gr}^{0.935} \mu_{ap}^{-0.107}$	48
2	$\varphi_r = 2.47 J_{Gr}^{0.97}$	49
3	$\varphi_r = 0.465 J^{0.65}\left(1 + \frac{A_d}{A_r}\right)^{-1.06} \mu_{ap}^{-0.103}$	50
4	$\varphi_r = 0.65 J_{Gr}^{(0.603 + 0.078 C_0)}\left(1 + \frac{A_d}{A_r}\right)^{-0.258}$ $\varphi_d = 0.46\varphi_r - 0.0244$	51
5	$\varphi_r = (0.491 - 0.498) J_{Gr}^{0.706}\left(\frac{A_d}{A_r}\right)^{-0.254} D_u \mu_{ap}^{-0.0684}$	52
6	$\varphi_r = 0.16\left(\frac{J_{Gr}}{J_{1r}}\right)^{0.57}\left(1 + \frac{A_d}{A_r}\right)$ $\varphi_d = 0.79\varphi_r - 0.057$	45
7	$\varphi_r = 0.364 J_{Gr}$	53
8	$\frac{\varphi_r}{1 - \varphi_r} = \frac{J_{Gr}^{n+2/2(n+1)}}{2^{3n+1/n+1} n^{n+2/2(n+1)}\left(\frac{K}{\rho_1}\right)^{1/2(n+1)} g^{n/2(n+1)}\left(1 + \frac{A_d}{A_r}\right)^{3(n+2)/4(n+1)}}$	54
9	$\frac{\varphi}{(1 - \varphi)^4} = \frac{0.124\left(\frac{J_G \mu_1}{\sigma_1}\right)^{0.996}\left(\frac{\rho_1 \sigma_1^3}{g\mu_1^4}\right)^{0.294}\left(\frac{D_r}{D}\right)^{0.114}}{1 - 0.276(1 - e^{-0.0368 Ma})}$	55
10	$\varphi_r = \frac{Fr}{0.415 + 4.27\left(\frac{J_{Gr} + J_{1r}}{\sqrt{gD_r}}\right)\left(\frac{g\rho_1 D^2}{\sigma_1}\right)^{-0.188} + 1.13 Fr^{1.22} Mo^{0.0386}\left(\frac{\Delta\rho}{\rho_1}\right)^{0.0386}}$	56
11	$\frac{\varphi}{(1 - \varphi)^4} = 0.16\left(\frac{J_G \mu_1}{\sigma_1}\right) Mo^{-0.283}\left(\frac{D_r}{D}\right)^{-0.222}\left(\frac{\rho_1}{\Delta\rho}\right)^{0.283} * (1 - 1.61(1 - e^{-0.00565 Ma}))^{-1}$	2
12	$\varphi_d = 4.51 \cdot 10^6 Mo^{0.115}\left(\frac{A_r}{A_d}\right)^{4.2} \varphi_r$ when $\varphi_r < 0.0133\left(\frac{A_d}{A_r}\right)^{-1.32}$ and $\varphi_d = 0.05 Mo^{-0.22}\left[\left(\frac{A_r}{A_d}\right)^{0.6} \varphi_r\right]^{0.31 Mo^{-0.0273}}$ when $\varphi_r > 0.0133\left(\frac{A_d}{A_r}\right)^{-1.32}$	57
13	$\varphi_r = 0.0057\left[(\mu_1 - \mu_w)^{2.75} - 161\frac{73.3 - \sigma}{79.3 - \sigma}\right] \cdot J_{Gr}^{0.88}$	58
14	$\varphi_r = \frac{0.4Fr}{1 + 0.4Fr\left(1 + \frac{J_1}{J_{Gr}}\right)}$	57
15	$\varphi = 0.24 n^{-0.6} Fr^{0.84 - 0.14n} Ga$	59

gas separator. If this parameter is not taken into account, then studies of the influence of the top clearance (42,65) are incomplete and difficult to extrapolate to other designs. The same can be said about the filling factor (66) given by the ratio of the gas separator volume to the total volume.

The foregoing discussion thus explains why all the correlations coincide for some ranges of these secondary variables while in other ranges they may diverge. In addition, in some cases the number of experiments may not have been sufficient to provide correlations or they may have been ill-balanced from the statistical point of view. The obvious solution to this problem lies in the collection of a large and detailed bank of reliable data that will constitute the basis for correlations with greater accuracy and validity.

The safest procedure for the prediction of the gas holdup in an ALR under design is to take data provided by researchers who have made the measurements in that par-

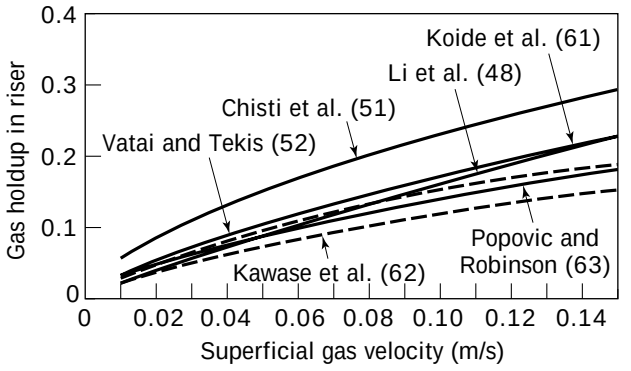


Figure 7. Some correlations proposed for prediction of gas holdup in the riser of internal-loop ALRs (Table 1). Gas holdup (φ_r) is presented as a function of superficial gas velocity (J_G). Other parameters related to geometry and physicochemical properties that were used in the calculations are shown on the figure.

ticular type of reactor with the same physicochemical properties of the system. If this option is not available, then correlation 9 in Table 1 (55) is recommended for prediction of the gas holdup in the riser.

Gas holdup in the downcomer is lower than that in the riser. The extent of this difference depends mainly on the design of the gas separator (67). The downcomer gas holdup is linearly dependent on the riser holdup, as a consequence of the continuity of liquid flow in the reactor. Many expressions of this type have been published (68). At low gas flow rates, φ_d is usually negligible, since most of the bubbles have enough time to disengage from the liquid in the gas separator. This usually happens at the low gas flow rates frequently used for animal cell cultures.

The gas holdup in the separator is very close to the mean gas holdup in the whole reactor (1) as long as the top clearance C_t is relatively small (one or two diameters). For larger top clearances, the behavior of the gas separator begins to resemble that of a bubble column, and the overall performance of the reactor is influenced by this change.

External-Loop Airlift Reactors. From the point of view of fluid dynamics, neither the external configuration (shape and architecture) nor the fact that both riser and downcomer are easily accessible is the most important difference between external- and internal-loop reactors. The most important point is that the gas separator of the external-loop ALR is built in such way that gas disengagement is usually much more effective in this type of reactor. This can be easily seen in Figure 2. In concentric tubes or split vessels, the shortest path that a bubble has to cover from the riser to the downcomer is a straight line across the baffle that separates the two sections. In the case of external-loop ALRs, there is usually a minimum horizontal distance to be covered, which increases the chances of disengagement of the bubbles. In this case, it is worth pointing out that if gas does appear in the downcomer, then most of it will be fresh air entrained in the reactor because of interfacial turbulence or vortices that appear in the gas separator above the entrance to the downcomer. In many of the studies reported in the literature on holdup in external-loop ALRs, total disengagement is attained. No such data are available for the concentric tubes of split-vessel ALRs, since total disengagement is possible only at very low gas flow rates.

Several authors (37,69–73) have presented their results of gas holdup as the gas velocity versus the superficial mixture velocity, based on the drift flux model of Zuber and Findlay (74). These authors derived general expressions for prediction of the gas holdup and for interpretation of experimental data applicable to nonuniform radial distributions of liquid velocity and gas fraction. The drift velocity is defined as the difference between the velocity of the particular phase (U) and the volumetric flux density of the mixture (J) where:

$$J = J_G + J_L \tag{5}$$

The drift velocities of the gas and liquid phases may thus be expressed as:

$$J_G = U_G - J \tag{6}$$

$$J_L = U_L - J \tag{7}$$

Zuber and Findlay (74) derived the relationship [8], which has been shown to be more than adequate to provide a correlation of gas holdup measurements in tower reactors with high liquid velocities, such as ALRs (71):

$$U_G = \frac{J_G}{\varphi} = C_0 J + \frac{\frac{1}{A}\int \varphi(U_G - J)dA}{\frac{1}{A}\int \varphi \cdot dA} \tag{8}$$

where A is cross-sectional area, C_0 is distribution parameter, J is superficial velocity, J_G is superficial gas velocity, U_G linear gas velocity, and φ is gas holdup.

Equation 7 describes the relationship between the gas velocity in a two-phase flow and the volumetric flow density of the mixture, J.

As stressed by Zuber and Findlay (74), J has the advantage of being independent on space coordinates for both one-dimensional flow and multidimensional irrotational flows. The distribution parameter C_0 is given by (75):

$$C_0 = \frac{\frac{1}{A}\int_A \varphi J \cdot dA}{\left[\frac{1}{A}\int_A J \cdot dA\right]\left[\frac{1}{A}\int_A \varphi \cdot dA\right]} \tag{9}$$

The value of C_0 depends mainly on the radial profile of the gas holdup. Zuber and Findlay (74) calculated $C_0 = 1$ for a flat profile and $C_0 = 1.5$ for a parabolic profile. Experimental values have been reported in the range of 1.03–1.2 for upflow (69–73,76) and 1.0–1.16 for downflow (70,73).

Equation 9 shows that this parameter is a function of the profiles of velocities and holdup. The last term of the right-hand side of equation 8 is the weighted mean value of the drift velocity:

$$U_{GI} = U_G - J \tag{10}$$

The drift velocity of a swarm of bubbles can be evaluated by using the expression given by Zuber and Findlay (74).

$$U_{2J} = 1.53 \cdot \left[\frac{\sigma g \Delta\rho}{\rho_L^2}\right]^{0.25} (1 - \varphi)^{1.5} \tag{11}$$

where U_{2J} is the velocity of the swarm of bubbles, g is gravitational acceleration, ρ_L is the density of liquid, $\Delta\rho$ is the density difference, σ is the surface tension, and φ is the gas holdup. This equation is valid for bubble diameters of the order of 0.1 to 2 cm, which covers the population usually observed in ALRs.

It has, however, been shown (71) that a plot of U_G versus J gives a straight line, suggesting that a constant value of the drift velocity satisfactorily represents the two-phase flow in the riser of an external-loop ALR. In this plot, the distribution parameter was $C_0 = 1.03$, and U_{GS}, the value of the slip velocity of a bubble, was taken as the mean drift velocity. Siegel et al. (35) applied the same model for the study of gas recirculation in a split-vessel ALR and obtained the values of $C_0 = 1.11$. The slip velocity that they obtained fitting their data to equation 8 was 0.238 m/s. It has been suggested (71) that this simplification holds as long as coalescence is not a predominant factor in the process.

It is very important to stress the difference between holdup, φ, and the flowing volumetric concentration (β), which is defined as:

$$\beta = \frac{Q_G}{Q_G + Q_L} = \frac{J_G}{J} \tag{12}$$

Zuber–Findlay's drift flux model allows us to derive the following equation, which establishes a connection between the gas holdup and β.

$$\frac{\beta}{\varphi} = C_0 + \frac{U_{b\infty}}{J} \tag{13}$$

where C_0 is the distribution parameter, J is the superficial velocity, $U_{b\infty}$ is the terminal gas velocity, β is the flowing volumetric concentration, and φ is the gas holdup.

Figure 8 gives a representation of the $\varphi - \beta$ plane. The 45° line indicates that $\varphi = \beta$, an equality that is true only for nonslip flow, where the velocity of the gas is equal to the velocity of the liquid. Such a situation can be visualized for the case of very small bubbles in a relatively fast liquid. In this case, there is no influence of one phase on the motion of the other. As indicated in Figure 8, all the points below the 45° line indicate operation situations in which the liquid is driven by the gas:

$$U_G > U_L;\ \varphi < \beta \tag{14}$$

This happens in the riser of ALRs. For all points above the line the opposite is true:

$$U_G < U_L;\ \varphi > \beta \tag{15}$$

This latter condition reflects the operation of the downcomer.

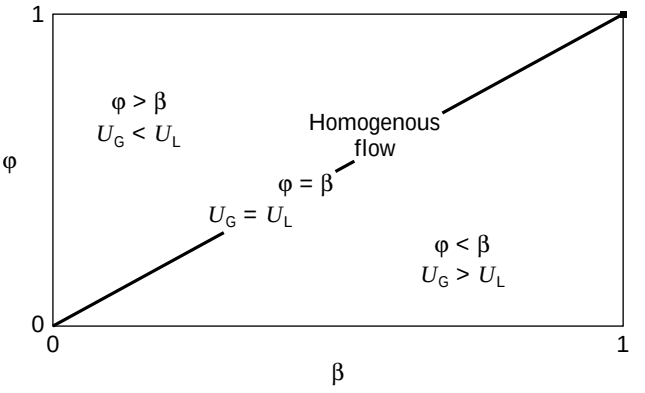


Figure 8. Gas flow holdup (φ) vs. flowing volumetric concentration (β). The different zones in the plane φ–β identify the two-phase flow. Adapted from Merchuk and Berzin (77).

A number of authors (71,76,78,79) have measured the local holdup profile along the riser of an external-loop ALR. In general, it was found that the holdup increases with height. This finding concurs with the expected expansion of gas bubbles as regions of lower pressure are reached. Common sense indicates that this situation must be limited to a certain range; an increase in bubble size will enhance turbulence and result in an increase in bubble encounters, leading eventually to bubble coalescence. The larger bubbles will rise much faster, resulting in a decrease in holdup. Such a scenario was indeed observed by Merchuk and Stein (71), as is illustrated in Figure 9. Merchuk and Stein (71) reported a maximum in the holdup profile for the case of a single-orifice gas distributor. For a multiple-orifice sparger, producing a more homogeneous bubble size distribution, a maximum was not observed within the studied length of the riser, which was 4 m.

Literature data from different sources for gas holdup in the riser under conditions of little or no carryover of gas from the separator into the downcomer for different A_d/A_r and top clearance C_t may be represented by the simple exponential:

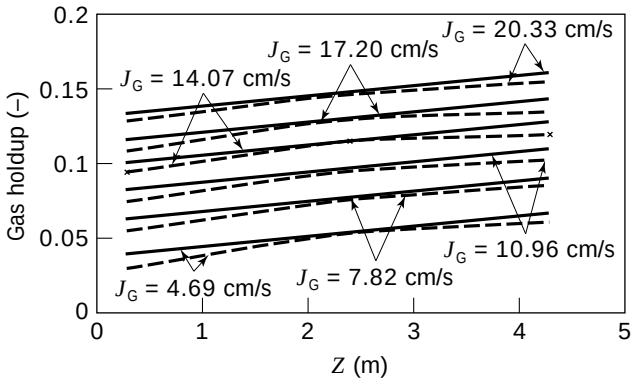


Figure 9. Dependence of the riser gas holdup in a 4-m high external-loop ALR with a multiple-orifice sparger (*solid lines*) and a single-orifice sparger (*broken lines*). Adapted from Merchuk and Stein (71).

$$\varphi_r = \alpha J_G^{\beta} \tag{16}$$

where the constant α depends on the friction losses in the loop, and β is usually a value between 0.6 and 0.7, as is illustrated in Figure 10 (65). The fact that neither the area ratio nor the top clearance affects the gas holdup demonstrates the role of the gas–liquid separator in determining the performance of the reactor in general. In the absence of gas recirculation, there is no effect on these variables. Moreover, this means that under conditions of no gas entrainment from the separator to the downcomer, it is possible to predict the riser gas holdup as a function of the riser superficial gas velocity alone, which is of great importance for design purposes.

It is accepted that liquid velocity has a mild negative effect on gas holdup in the riser. This effect is usually studied by reducing the liquid flow; this is achieved by adding resistance to the liquid loop by means of a valve or other controlled obstruction (71,78,82) under conditions of low or nil gas recirculation. Such experiments, which are relatively simple in external-loop ALRs, indicate that the holdup decreases as the liquid velocity is increased from zero (bubble column) to 0.3 m/s (which is close to the bubble free-rise velocity). For higher velocities, the effect of U_L is small. These findings add to our understanding of the fluid dynamics in the column. At liquid velocities that are smaller than the bubble free-rising velocity, the liquid transported in the wake of the bubbles, which must return downward to balance the mass flux, is the cause of the meandering and loops that typically appear in bubble column operation (83). As the overall liquid flux increases, the patterns straighten out, the bubbles begin to ascend in a straight pattern, and the holdup goes down. When the liquid velocity is higher than the free-rise velocity of the bubbles, piston flow of bubbles ensues in the tube, and the decrease in holdup for further increases in liquid velocity is due solely to the change in the ratio of gas–liquid volumetric flow rates.

When there is gas recirculation, the area ratio A_d/A_r becomes an important variable affecting gas holdup. The effect of A_d/A_r starts in the region in which gas entrainment from the separator to the downcomer occurs. If it is assumed that the riser cross-sectional area A_r remains unchanged and the downcomer cross-sectional area A_d is increased, then it can be expected that the liquid velocity in the riser will increase as a result of the smaller resistance to flow in the loop, which in turn leads to a decrease in the riser gas holdup. An increase in A_d/A_r will result in a decrease in the liquid velocity in the downcomer, which leads to a decrease in the gas recirculation, since fewer bubbles are entrapped in the downcomer. The final outcome of increasing A_d/A_r is thus a decrease is the riser gas holdup. A similar argument can be applied in the discussion of the effect of reactor height on the riser gas holdup, that is, an increase in the height of the downcomer will result in a higher liquid velocity, which will in turn lead to a decrease, as in the former case, in the holdup in the riser. In contrast, an increase in A_d/A_r will lead to an increase in the extent of bubble entrapment in the downcomer, which will serve to inject some additional gas into the riser. On the other hand, an increase of gas holdup in the downcomer diminishes the driving force for recirculation, as shown in equation 1, and this will moderate the increase of liquid velocity generated by the larger height. This feedback control of the liquid velocity is one of the characteristics particular to ALRs.

Table 2 shows most of the expressions published for the correlation of experimental data obtained in external-loop ALRs. Some of these expressions are presented in Figure 11. The differences between the predictions obtained with the different correlations are probably due to the design of the gas separator. The equation given by Popovic and Robinson (63) seems to give an average of the proposed correlations.

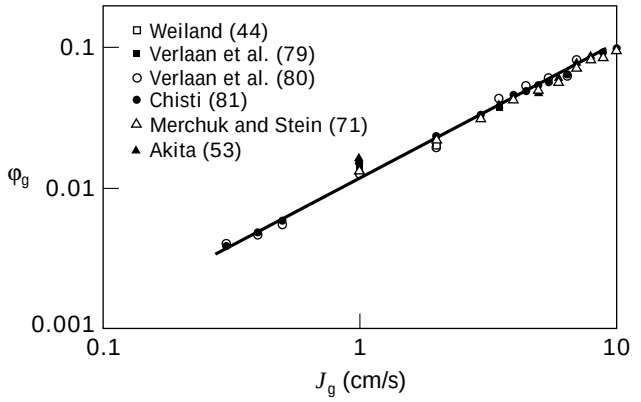


Figure 10. Gas holdup reported by various sources for the riser of airlift reactors under conditions of little or no gas recirculation. The data correspond to different A_d/A_r ratios.

Table 2. Gas Hold-up in External-Loop ALR

No.	Formula	Ref
1	$\varphi_r = \dfrac{0.6\rho_G^{0.062}\rho_l^{0.069}\mu_G^{0.107}}{\mu_l^{0.053}s_l^{0.185}} \cdot \dfrac{J_{Gr}^{0.936}}{(J_{Gr} + J_{lr})^{0.474}}$	84
2	$\varphi_r = 0.16\left(\dfrac{J_G^2}{J_{lr}}\right)^{0.56}\left(1 + \dfrac{A_d}{A_r}\right)$	85
	$\varphi_d = 0.89\varphi r$	
	$\varphi_r = 1.07Fr^{0.333}$	
3	$Fr = \dfrac{J_G^2}{gD_r}$	62
4	$\varphi = 0.55J_{Gr}^{0.78}F^{0.2}D_r^{0.42}$	66
	$F = \dfrac{V_{ls}}{V_l}$	
5	$\varphi_r = 0.203\dfrac{Fr_*^{0.31}}{Mo^{0.012}}\left(\dfrac{J_{Gr}}{J_{lr}} \cdot \dfrac{A_r}{A_d}\right)^{0.74}$	86
	$Mo = \dfrac{g(\rho_l - \rho_G)}{\sigma_l\rho_l^2} \cdot K^4\left(\dfrac{8J_{lr}}{D_r}\right)^{4(n-1)}\left(\dfrac{3n+1}{4n}\right)^{4n}$	
	$Fr_* = \dfrac{(J_{lr} + J_{Gr})^2}{gD_r}$	
6	$\varphi_d = 0.997\varphi_r$	81
7	$\varphi_r = 0.16\left(\dfrac{J_{Gr}}{J_l}\right)^{0.56}\left(1 + \dfrac{A_d}{A_r}\right)$	45

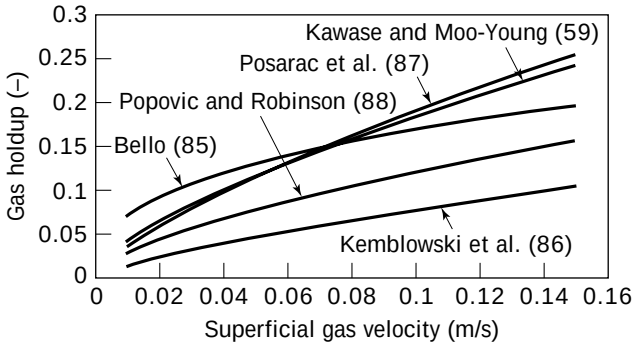


Figure 11. Some correlations proposed for prediction of gas holdup in the riser of external ALRs (Table 2). The gas holdup is presented as a function of the superficial gas velocity.

Effects of Liquid Rheology. The effect of rheology on the reactor behavior and performance is of great interest because in most biotechnological processes an increase in biomass provokes changes in the rheology of the fluid, especially in the case of mycelial growth. This effect is enhanced when in addition to the biomass growth, a product of the process is released into the medium in appreciable amounts. A good example of this scenario is the biosynthesis of polysaccharides, which cause an increase in the liquid viscosity.

The effect of viscosity on gas holdup in bubble columns has been studied by a number of authors. The main problem to be overcome is that of non-Newtonian flow. If the viscosity is not constant, but changes with changes in the shear rate, then the evaluation of shear rates becomes particularly relevant for the identification of the system. Several authors have confronted this issue. Nishikawa et al. (89,90) analyzed the problem of heat transfer in a bubble column with non-Newtonian liquids. They found a direct proportionality between the superficial gas velocity and the global shear rate:

$$\gamma = 5000 \cdot J_G \quad (J_G > 0.04\text{m/s}) \tag{17}$$

This global shear rate was then used to calculate a global viscosity. In shear-sensitive cultures, the definition of a global shear rate in itself is of great importance.

A number of researchers, Henzler (91), Kawase and Moo-Young (59), Schumpe and Deckwer (92,93) have followed the approach of Nishikawa et al. (89) but have suggested different proportionality constants relating the global shear rate to the superficial gas velocity. This approach is questionable from the rheological point of view because it will predict the same shear rate for a certain superficial gas velocity, no matter which liquid is used. El-Tamtamy et al. (94) introduced an improvement by calculating the shear rate from the bubble velocity divided by the bubble diameter. However, accurate evaluation of the latter two parameters is difficult. Henzler and Kauling (95) suggested relating the shear rate to power input based on dimensional analysis by expressing the shear rate as a function of the power input per unit volume, $(P/[V\rho\nu])^{1/2}$. Their analysis gives different shear rates for liquids that are rheologically different.

The above-described relationships predict different shear rates that vary in up to three orders of magnitude. It is thus generally agreed that the correct solution is still to be found. Recently, a more general approach, known as a global approach, has been proposed by Merchuk and Ben-Zvi (Yona) (96). The shear stress in a bubble column was defined as being equal to the acting force, which can be calculated from the power input divided by the sum of the areas of all the bubbles:

$$\tau = \frac{P}{L_R} S_{ab} \tag{18}$$

where L_R is an effective length that represents the mean circulation path of a bubble in the system considered, P is the power input, S_{ab} is the total surface of all of bubbles, and τ is the shear stress.

Assuming ideal gas isothermal expansion, the power input P can be calculated. The interfacial area can be evaluated from correlations or can be obtained by direct measurement if available. A correlation taking into account other variables, like sparger configuration, surface tension, etc., will broaden the range of applications of this approach.

If a constitutive equation describing the rheology of the system is available (such as the power law, which has been reported to correspond to many biological systems), equation 17 facilitates the calculation of a global shear force acting on the liquid. The shear rate can be in this case expressed as:

$$\gamma = \left[\frac{\tau}{k}\right]^{1/n}$$

where γ is shear rate and κ is behavior coefficient, and equation 17 can be now used to express γ as:

$$\gamma = \left[\frac{p_1 J_G \ln\left(\frac{p_1}{p_2}\right)}{aL_R^2\kappa}\right]^{1/n} \tag{19}$$

where the subindexes 1 and 2 represent the two extremes of the section considered.

Equation 19 thus gives a global shear rate that is a function of both fluid dynamics and rheology. This approach has been found to be useful for the presentation of results on mass transfer rates in bubble columns (96).

In contrast to the marked influence of rheology on gas holdup in bubble columns, the data available for ALRs show clearly that the effect of liquid viscosity is less dramatic, but not simpler. Figure 12 (65) illustrates the effect of the addition of glycerol to water in an internal-loop ALR. At low concentrations of glycerol, a moderate increase of the gas holdup is evident, particularly in the downcomer but also in the riser. These increases are caused by the lower free rise velocity of the bubbles, which increase the gas retention due to the longer residence time. In addition, the entrapment of the bubbles is increased, and this is reflected mainly in φ_d. When the concentration of glycerol becomes too high, a strong decrease of the gas holdup is

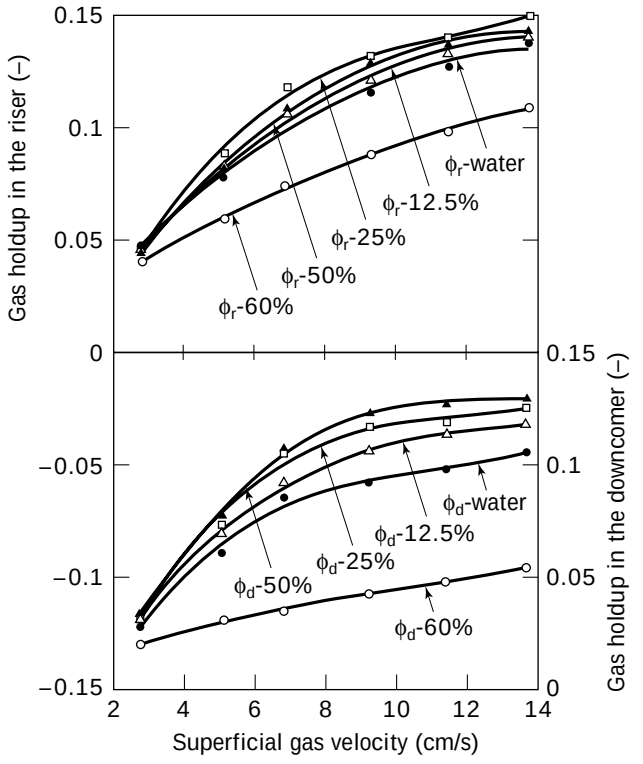


Figure 12. Effect of liquid viscosity on the gas holdup in the riser and in the downcomer of an internal-loop ALR. The viscosity corresponding to the solutions used was 4–14 mPa s. Percentages refer to percent glycerol in water (65).

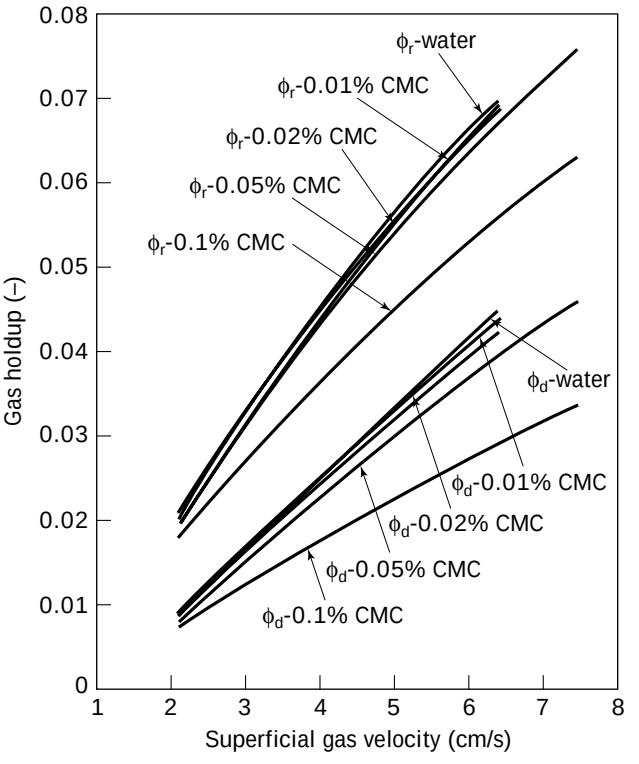


Figure 13. Effect of liquid rheology on the gas holdup in the riser and in the downcomer of an internal-loop ALR with a non-Newtonian liquid. The apparent viscosity corresponding to the solutions used, calculated for $\gamma = 50\ s^{-1}$, was 5–56 mPa s (65).

seen. This decrease is probably due to the onset of coalescence, which produces larger bubbles that ascend faster in the liquid and easily disengage in the gas separator. The viscosities corresponding to these solutions ranged from 4 mPa s to 14 mPa s. In Figure 13, the addition of carboxymethyl cellulose (CMC) to water is shown (65). The change in CMC concentration had only a slight effect on the gas holdup for additions in the range 0.01 to 0.05% CMC. Only for solutions with concentrations higher than 0.5% CMC was an appreciable decrease in holdup seen.

Effect of Liquid Level. The influence of the liquid level C_t on the gas holdup is exerted as a consequence of changes in the extent of disengagement of the bubbles in the gas separator. This influence is therefore dependent on the geometric design of this section. Whether a bubble will disengage or will be entrapped into the downcomer depends on the interrelationship of several parameters—the free rising velocity of the bubble $U_{b\infty}$ (a function of size and viscosity), the liquid velocity in the downcomer U_{Ld} (a function of the difference in gas holdup between the riser and the downcomer and frictional losses), and the residence time of the bubble in the gas separator (a function of geometric design and liquid height). For a given bubble size, if U_{Ld} is smaller than the $U_{b\infty}$ corresponding to the smallest bubble, then there is no carryover. For smaller bubbles, the balance between the time required to cover the path from the end of the riser to the zone near the entrance of the downcomer and the time needed for disengagement will give the fraction of bubbles recirculated. It should nevertheless be kept in mind that this is a feedback process. A higher bubble disengagement rate leads to a lower gas holdup in the downcomer, which in turn increases the liquid velocity, enabling larger bubbles to be trapped, until the system eventually reaches a steady state.

Due to the above-explained influence of the geometric design, the influence of the liquid height is completely different in internal- and external-loop ALRs. In internal-loop reactors, an increase in C_t increases the zone of the separator in which the gas holdup is higher, and as a result, the gas holdup increases. The extent of this increase depends, as said earlier, on the free-rise velocity of the bubbles. Figure 14 (1) shows the gas holdup in the riser and downcomer of a 30-liter ALR for two values of the top clearance, $C_t = 0.178$ m and $C_t = 0.308$ m, for two different liquids, water and a 0.5% CMC solution. It can be seen that although for water C_t has a small effect, this is not so for the more viscous solution. In the latter, the lower rising velocity of the bubbles causes a greater number of bubbles to be entrained and carried down by the liquid. Thus, the residence time in the disengagement section becomes very important in determining the fraction of bubbles that recirculate. A lower C_t will give a shorter residence time in the separator, a larger bubble recirculation, and, hence, a larger gas holdup.

In external-loop ALRs, the opposite effect is obtained, as may be seen in Figure 15 (65). The data in Figure 15,

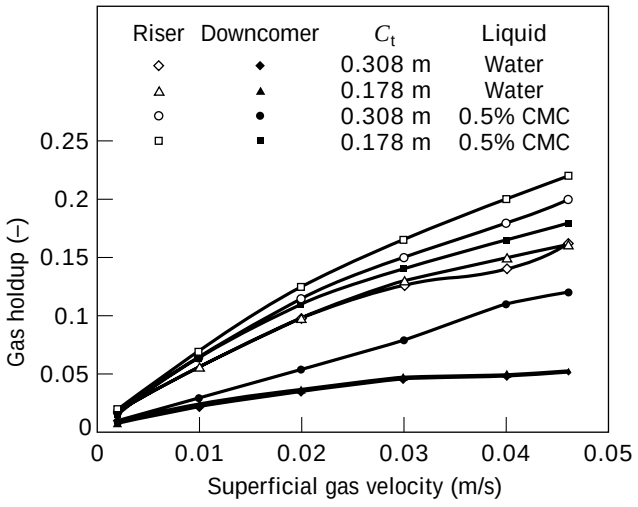


Figure 14. Riser and downcomer gas holdup in an internal-loop ALR for two different top clearances and two liquids. Adapted from Merchuk et al. (1).

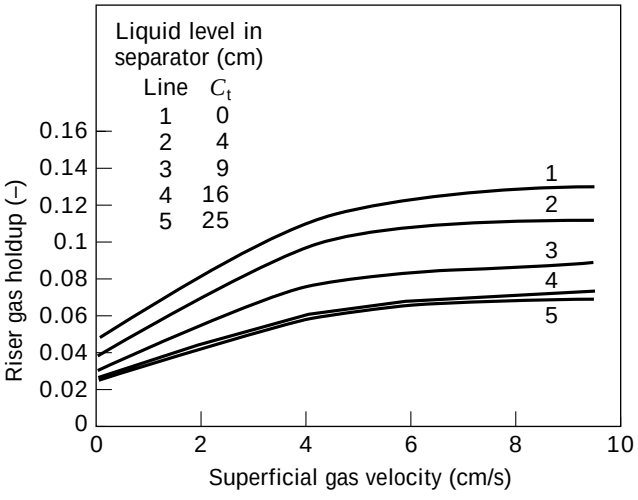


Figure 15. Gas holdup in the riser of an external-loop ALR for several top clearances. Adapted from Hallaile (65).

obtained for a 4-m high external-loop ALR, show that the holdup in the riser decreases as C_t increases. This is due to the construction of these reactors, in which much of the gas that enters into the downcomer is trapped from the headspace due to the turbulence in this zone. An increase in the liquid height serves to reduce the amount of gas trapped, so that less gas circulates in the downcomer and the liquid velocity increases. The final result is a reduction in the gas holdup, both in the riser and the downcomer (65).

Gas Recirculation

The degree to which gas flowing in the riser is entrapped and recirculated through the downcomer is an important variable, since it influences not only the flow configuration in the downcomer, but also the overall performance of the ALR. The liquid velocity depends mainly on the difference in holdup between the riser and the downcomer, and it in turn influences the gas holdup in the riser. Despite the importance of recirculation, very little quantitative data are available on this phenomenon. Siegel et al. (35) evaluated the gas recirculation in a split-vessel ALR by an indirect method based on holdup measurements. From their results, shown in Figure 16, it may be seen that the recirculation rate remains fairly constant for changing gas flow rates in the riser for high values of the last variable. Thus, the recirculation rate is determined largely by the geometric configuration of the gas–liquid separator and the liquid level in the separator.

Three zones are evident in Figure 16; they represent operating conditions giving oscillating, borderline, and straight bubble flow in the downcomer. Oscillating flow patterns produce much larger fractions of gas retained in the downcomer, but they are much more sensitive to J_G. At low superficial gas velocity, the recirculation increases very sharply with J_G. The bubbles exhibit an oscillating swirling flow pattern, with some larger bubbles escaping toward the top. The borderline condition is defined as oscillatory bubble flow at a low gas flow rate that shifts to straight bubble flow with increasing input gas flow rate. The straight flow operation zone is distinguished by bubble flow in a straight, well-defined flow pattern for all the input gas flow rates studied. If a straight bubble flow pattern is desired, the reactor should be operated at high riser gas flow rates, at which the reactor will shift toward stable operation.

Lubbert et al. (97) attempted to evaluate the recirculation of gas during the cultivation of yeast (*Saccharomyces cerevisiae*) on waste from a starch factory in a 4-m^3 pilot plant. They used microprocessor-aided pseudostochastic tracer input and cross-correlation techniques, which facilitated very reduced tracer feeds due to a high signal-to-noise ratio. The response to a pulse of helium was measured directly at the surface of the liquid in the separator by a quadruple mass spectrometer. The peak obtained showed pronounced shoulders (Fig. 17) which could be interpreted as superimposition of a second peak that represents the helium tracer one loop after. The fitting of such

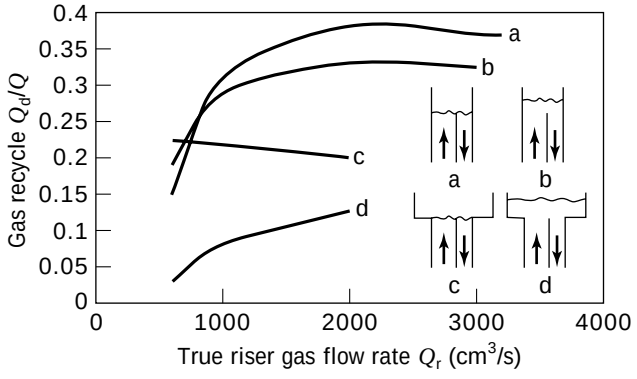


Figure 16. Gas recirculation in a split-cylinder ALR. The level indicated corresponds to no-aeration conditions. Adapted from Siegel et al. (35).

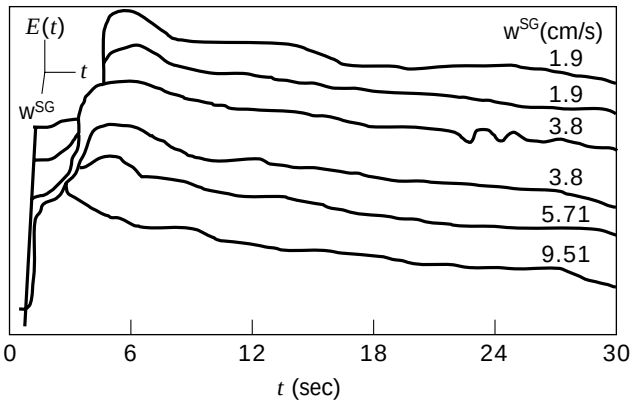


Figure 17. Response to a pulse of helium in an airlift reactor. The pronounced shoulders could be interpreted as the result of superimposing a second peak, which represents the He tracer one loop after. Reprinted with permission from Lubbert et al. (98).

a model to the experimental data suggested recirculation of 25% of the gas. This figure is in the range of the recirculation rates presented by Siegel et al. (35), considering the differences in the coalescing properties of the liquids used in these two works and the corresponding differences in the downcomer holdup.

Recently, Merchuk and Berzin (77) developed a mathematical model based on the application of the first law of the thermodynamics to each of the regions of an ALR. This model facilitates the evaluation of the maximum liquid recirculation possible in the system. The calculation is based on the premise that the gas recirculated must be compressed from the pressure at the top of the downcomer to the pressure at its bottom. The mathematical expression that gives this maximal gas recirculation is:

$$Q_d = \frac{Q_L(P_2 - P_3) + Q_L\rho_L gh - \frac{1}{2}\rho_L C_d A_d(1 - \varphi_d)U_L^3}{P_4 \ln\left(\frac{P_3}{P_2}\right)} \quad (20)$$

where Q_d is the gas flow rate in the downcomer, Q_L is the liquid circulation flow rate, P_i is pressure at point i of the reactor (1 is top of the riser, 2 is top of the downcomer, 3 is bottom of the downcomer, 4 is bottom of the riser), C_d is the hydraulic resistance coefficient, A_d is the downcomer cross-sectional area, U_L is the linear liquid velocity, g is the gravitational acceleration, ρ_L is the liquid density, and φ_d is the downcomer gas holdup.

The calculation of Q_d thus requires knowledge of the liquid flow rate, the pressures, and the geometry of the reactor. This equation represents the maximum recycling of gas in the downcomer, which will take place only if all the energy dissipated in the downcomer is invested in gas compression.

Liquid Velocity

The liquid velocity is one of the most important parameters in the design of ALRs. It affects the gas holdup in the riser and downcomer, the mixing time, the mean residence time of the gas phase, the interfacial area, and the mass and heat transfer coefficients.

Circulation in ALRs is induced by the difference in hydrostatic pressure between the riser and the downcomer as a consequence of a difference in gas holdup. Liquid velocity—like gas holdup—is not an independent variable, because (see Fig. 5) the gas flow rate is the only variable that can be manipulated. As shown in Figure 5, the geometric design of the reactor will also influence the liquid velocity, but this remains constant during operation. Experiments have been carried out in devices specially designed to artificially change the resistance to flow, with the aim of studying the effect of the velocity at a fixed rate of aeration (71,79). The information emerging from these experiments indicates that an increase in the liquid velocity leads to a decrease in the mean residence time of bubbles in the riser and hence of the gas holdup in the riser. In practice, when the gas flow rate is increased, the higher liquid velocity increases the carryover of bubbles from the gas separator into the downcomer; the carryover dampens the liquid flow by reducing the hydrostatic driving force. As a result, the overall change in liquid velocity is tempered.

Liquid Velocity Measurement. Several different methods can be used for measuring the liquid velocity. The most reliable ones are based on the use of tracers in the liquid. If a tracer is injected and two probes are installed in a section of the tube, the velocity of the liquid traveling the distance between probes can be taken directly from the recorded peaks, as the quotient of the distance between the two electrodes and the time required by the tracer to travel from the one to the other. The latter is obtained as the difference of between the first moments of the two peaks. A second method is to calculate the liquid velocity (U_L) from the circulation time (t_c) and holdup (φ) as:

$$U_L = \frac{\text{liquid volume}}{t_c \times A \times (1 - \varphi)} \quad (21)$$

where A is cross-sectional area.

In this case, only one electrode is necessary, φ is the holdup at the point at which the electrode is installed, and the circulation time is obtained from two successive peaks recorded by the electrode.

Modeling of Liquid Flow. A number of expressions are available for the estimation of the liquid velocity. Two main methods have been used for the modeling of two-phase flow in ALRs—energy balances and momentum balances. Chakravarty et al. (58) used the energy balance approach to obtain a relationship between superficial gas velocity, holdup, and liquid velocity. Lee et al. (99) calculated U_L by a similar type of model for a series of published data for concentric and external-loop ALRs and from their own results for split vessels. In both the above-mentioned models, constants accounting for friction losses were obtained by adjusting the models to the experimental data. Jones (100), on the other hand, managed to express the results of his energy balance (based on previous work of Niklin

[101] and Freedman and Davidson [73]) in a relationship free of empirical constants. His results, however, fit the experimental data only qualitatively, and the fit is satisfactory only for very small diameters. An improvement of this method was suggested by Clark and Jones (102), who took into account the radial distribution of the gas holdup through the drift flux model. However, the values of the distribution coefficient C_0 needed for satisfactory fitting of the experimental data for lager diameters is far from the range usual in this type of flow.

Chisti and Moo-Young (103) extended a model originally proposed by Bello (85), based on an energy balance over the airlift loop. Their expression for the average superficial liquid velocity is:

$$U_{Lr} = \left(\frac{2gH_d(\varphi_r - \varphi_d)}{\frac{K_t}{(1-\varphi_r)^2} + K_B\left(\frac{A_r^2}{A_d}\right)\frac{1}{(1-\varphi_d)^2}} \right)^{0.5} \qquad (22)$$

where U_L is the superficial liquid velocity, A_r is the riser cross-sectional area, A_d is the downcomer cross-sectional area, H_d is the downcomer height, K_b and K_t are the hydraulic pressure loss coefficients, φ_r is the riser gas holdup, and φ_d is the downcomer gas holdup.

By choosing suitable values for the friction coefficients in each case, the authors showed that much of the published data on liquid velocity for the different types of ALRs could be satisfactorily correlated by equation 22. Only one coefficient has to be adjusted, since the authors assume that K_t, the friction coefficient at the top of the loop, is negligible in concentric-tube type reactors and that in external-loop reactors K_t can be taken as equal to K_b, the friction coefficient for the bottom of the loop. Equation 22 has thus been adopted by many scientists. Wachi et al. (104) claimed that their derivation of the same equation gives a clearer physical meaning to the adjustable parameters. Equation 22 can also be also derived from a simple momentum balance (77).

Chisti et al. (51) presented an empirical correlation for K_b obtained by comparison of results obtained from several sources:

$$K_b = 11.402 \cdot \left(\frac{A_d}{A_b}\right)^{0.789} \qquad (23)$$

where A_b is the minimal cross section at the bottom of the airlift reactor and A_d is the downcomer cross-sectional area.

Equation 22 has the particularly that the gas flow rate, which is the main, and often the only, manipulable variable in the operation, is not present directly, but exerts its influence through the gas holdup. Therefore, either experimental data or a valid mathematical expression for the gas holdup in both the riser and the downcomer are required.

Chisti and Moo-Young (103) extended this model further in order to facilitate the prediction of liquid circulation in ALRs operating with pseudoplastic fluids, such as mold suspensions. This improvement is very important, since many commercial fermentation processes involve such non-Newtonian liquids. Kemblowski et al. (86) presented a method for the prediction of gas holdup and liquid circulation in external-loop ALRs. In their experiments there was almost no gas recirculation, because of the large size of the gas separators used.

Garcia Calvo (105) presented an ingenious model based on energy balances and on an idea originally proposed by Richardson and Higson (106), and Garcia Calvo and Leton (107) extended the model to bubble columns. The model is based on the assumption that the superficial gas velocity (J_G) in any region can be considered to be the sum of two streams (J' and J'') as follows. The J' stream has a velocity equal to that of the liquid and can therefore be treated by the laws of homogeneous two-phase flow (no slip between the bubbles and the liquid). The second stream (J'') is considered to be responsible for all the energy loses at the gas–liquid interface. The concept in itself is simple and elegant, and it is possible to envisage its application even to the flow in the downcomer, where $U_G < U_L$. In such a case, we would divide the gas flow rate into two parts as follows: One part would be larger than the actual flow rate, i.e., it would have the same velocity as the liquid. In order to arrive at the actual gas flow rate, the second flow rate must have the reverse direction. This type of gas flow can actually be seen under certain conditions, such as when there is coalescence of bubbles and larger bubbles ascend along the walls of the downcomer.

Another technique used by several researchers to predict liquid velocity is the momentum balance of the ALR. This method has been used by Blenke (108) in jet-loop reactors and by Hsu and Dudukovic (109), Kubota et al. (36), Bello (85), Koide et al. (47), and Merchuk and Stein (71). The latter authors presented a simple model for the prediction of the liquid velocity as a function of the gas input in an ALR. They assumed that the pressure drop between the bottom and the top of their external-loop reactor could be expressed as a continuation of the downcomer, using an equivalent length L_E. This length was set as an adjustable parameter describing the pressure loss in the loop. Kubota et al. (36) used a similar approach for the analysis of Imperial Chemical Industries' deep-shaft reactor. They were able to simulate the operation of the reactor and to predict the minimum air supply required to prevent flow reversal.

Verlaan et al. (76) used a similar model, in combination with the expression of Zuber and Findlay (74), to calculate the friction coefficients from experimental data reported by several authors for a wide range of reactor volumes. Koide et al. (47) presented an analysis of the liquid flow in a concentric-tube ARL that was also based on a momentum balance. The main difference between this model and that used by Merchuk and Stein (71) was that Koide et al. used a convergence–divergence flow model for the bottom and the top of the loop. At the bottom, the effect of flow reversal on the pressure drop was included in the effective width of the gas-liquid flow path under the lower end of the draft tube h_l, which was smaller than the actual gap. Miyahara et al. (57), who studied both the bubble size distribution in an internal-loop ALR and the pressure drop at the top and the bottom of the draft tube, also presented a model facilitating the prediction of the liquid velocity.

Other models use the drift-flux model (59) presented in equations 5–7, as:

$$U_{Lr} = \frac{J_{Gr}\left(\frac{1}{\varphi_r} - C_0\right) - U_{GJ}}{C_0(1 - \varphi_r)} \quad (24)$$

U_{GJ} can be taken from equation 10. The range of variation of C_0 is rather narrow, as shown in the previous section, and therefore it is not difficult to make a judicious guess as to the value of C_0 in an unknown system. The drift flux model has also been used together with energy balances (110) or with the momentum balance (111). Some studies on liquid measurement present the results in the form of empirical correlations (42,52,112). The usefulness of these correlations depends on the amount of data and the number of parameters taken into account.

Most of those correlations are shown on Table 3, and some of them are presented in Figure 18. In general, the superficial liquid velocity increases with an increase in the superficial gas velocity, but its rate of increase is much lower at high superficial gas velocities. From Table 3, it can be seen that the riser-to-downcomer cross-sectional area ratio and the reactor height are the main parameters that affect the superficial liquid velocity at constant superficial gas velocity. The superficial liquid velocity increases with an increase in A_d/A_r.

The effect of the properties of the liquid, such as viscosity, surface tension, and ionic strength, on the superficial liquid velocity are much milder in ALR than in bubble columns (113). It is expected that increasing the liquid viscosity will decrease the liquid velocity because of frictional losses, but this, in turn, will increase the gas holdup in the riser and consequently increase the driving force for liquid recirculation. Hence, it seems that these two effects balance each other partially and result in a milder effect on the superficial liquid velocity. The effects of the surface tension and the ionic strength are also exerted via their influence on the gas holdup, as analyzed above.

As a rule, it can be recommended that the Zuber-Findlay expression (equation 20) be used when the holdup is known and the liquid circulation velocity is high. For low liquid velocity, a correlation obtained in a piece of equipment as similar as possible to the one under design should be used.

Table 3. Liquid Circulation Velocity in ALRs

No.		Formula	Ref.
1	ELR	$J_{Lr} = 1.55 J_{Gr}^{0.33}\left(\frac{A_d}{A_r}\right)^{0.74}$	45
	ILR	$J_{Lr} = 0.66 J_{Gr}^{0.33}\left(\frac{A_d}{A_r}\right)^{0.78}$	
2	Bubbly	$J_{Lr} = 0.024 J_{Gr}^{0.322}\left(\frac{A_d}{A_r}\right)^{0.794} \mu_{ap}^{-0.395}$	50
	Slug flow	$J_{Lr} = 0.052 J_{Gr}^{0.322}\left(\frac{A_d}{A_r}\right)^{0.794} \mu_{ap}^{-0.395}$	
3	Slug flow	$J_{Lr} = 0.23 J_{Gr}^{0.322}\left(\frac{A_d}{A_r}\right)^{0.97} \mu_{ap}^{-0.39}$	63
4		$J_{Lr} = 2.858 J_{Gr}^{0.482}\left(\frac{A_d}{A_r}\right)^{0.97} 416\mu_{ap}^{-0.0105}$	52
5		$J_{Lr} = \left[\frac{2gH_d(\varphi_r - \varphi_d)}{K_b\left(\frac{1}{(1-\varphi_r)^2} + \left(\frac{A_d}{A_r}\right)^2 \frac{1}{(1-\varphi_d)^2}\right)}\right]^{0.5}$	103
		$K_b = 11.4\left(\frac{A_d}{A_r}\right)^{0.79}$	

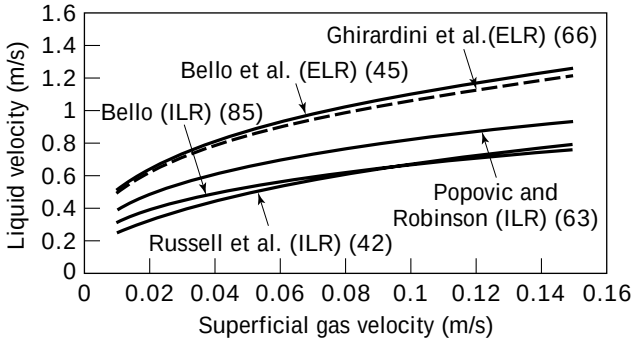


Figure 18. Liquid velocity predicted by some of the proposed correlations from Table 3. ILR, internal-loop ALR; ELR, external-loop ALR.

Liquid Mixing

For the design, modeling, and operation of ALRs, a thorough knowledge of mixing behavior is necessary. This is of particular importance during the process of scale-up from laboratory-scale to industrial-scale reactors. The optimum growth rate of a microorganism or the optimum production rate of a specific secondary metabolite usually relates to well-defined environmental conditions, such as pH range, temperature, substrate level, limiting factors, dissolved oxygen, and inhibitor concentration in a specific well-mixed laboratory-scale vessel. Because of the compromises made during scale-up, it is difficult to keep, at different scales of operation, the same hydrodynamic conditions established in the laboratory; mixing on an industrial scale may not be as good as mixing on a laboratory scale (5). In smaller-scale reactors it is easier to maintain the optimal conditions of pH, temperature, and substrate concentration required for maximum productivity of metabolites in a fermenter.

Furthermore, in fermentation systems efficient mixing is required to keep the pH within the limited range, giving maximum growth rates or maximum production of the microorganism during addition of acid or alkali for pH control. Mixing time—or the degree of homogeneity—is also very important in fed-batch fermentation, where a required component, supplied either continuously or intermittently, inhibits the microorganisms or must be kept within a particular concentration range (114,115). A large number of commercially important biological systems are operated in batch or fed-batch mode. In this operation mode, fast distribution of the incoming fluid is required, and the necessity for understanding the dynamics of mixing behavior in these vessels is obvious. Even for batch systems, good control of the operating conditions, such as pH, temperature, and dissolved oxygen, require prior es-

timation of mixing so that the addition rates can be suitably adjusted. Deviation of the pH or temperature from the permitted range may cause a damage to the microorganism, in addition to its effect on the growth and production rates. Moreover, a knowledge of the mixing characteristics is required for modeling and interpreting mass and heat transfer data.

A parameter used frequently to represent mixing in reactors is the mixing time (t_m). It has the disadvantage that it is specific to the reactor design and scale, but it is easy to measure and understand. Mixing time is defined as the time required to achieve the desired degree of homogeneity (usually 90–95%) after the injection of an inert tracer pulse into the reactor. The so-called degree of homogeneity (I), is given by:

$$I = \frac{C - C_m}{C_m} \tag{25}$$

where C is the maximum local concentration and C_m is the mean concentration of tracer at complete mixing.

A more comprehensive way of analyzing mixing, applicable to continuous systems, is a study of the residence time distribution (RTD). Although ALRs are usually operated in a batch-wise manner, at least in the laboratory, advantage is taken of the fact that the liquid circulates on a definite path to characterize the mixing in the reactor. Hence, a single-pass RTD through the whole reactor or through a specific section is usually measured. Based on the observed RTD, several models have been proposed. These models have the advantage of reducing the information of the RTD to a small number of parameters, which can later be used in design and scale-up.

The axial dispersion model, which has the advantage of having a single parameter, is widely accepted for the representation of tower reactors. This model is based on visualization of the mixing process in the tower reactor as a random, diffusion-like eddy movement superimposed on a plug flow. The axial dispersion coefficient D_z is the only parameter in the formulation:

$$\frac{\partial C}{\partial \tau} = D_z \frac{\partial^2 C}{\partial z^2} + U_L \frac{\partial C}{\partial z} \tag{26}$$

where C is the concentration of a tracer. The boundary conditions depend on the specific type of tower reactor. This model is attractive, since it has a single parameter, the Bodenstein number (Bo), which is used to describe the mixing in the reactor:

$$Bo = \frac{U_L L}{D_z} \tag{27}$$

where L is the characteristic length. When the Bo number tends to infinity, the mixing conditions are similar to those of a plug-flow reactor, and the reactor can be considered as well-mixed for low Bo numbers.

The alternative approach of Buffham and Mason states that the mixing characteristics of a piece of equipment should be expressed as the variance σ^2 of the distribution obtained by injection of a pulse of tracer without adopting any mechanistic model (116). The relationship between Bo and σ^2 depends on the reactor configuration (117). The approach of Buffham and Mason facilitates the presentation of mixing characteristics free of any modeling assumptions. The variance σ^2 is the second moment of the distribution and carries information on the spread of the distribution around the mean value (first moment). Nevertheless, most of the data on mixing in bioreactors are presented either as t_m or as overall Bo numbers, which can be obtained by relatively simple experiments of pulse injection.

Single-pass mixing in the ALR is due to mixing in the individual and interrelated sections of the reactor—riser, separator, downcomer, and bottom. Repeated passage mixing is the sum of the mixing in the subsequent passages. The latter is usually reported as the mixing time (t_m), the former as Bo or σ^2. Indeed, these parameters are interrelated, and knowledge of Bo or t_m is sufficient for calculating, theoretically, the mixing time (108,118) based on the deviation of the envelope of the maxima in the response curve to a pulse, which is a measure of the degree of inhomogeneity. Verlaan et al. (80) and Lin et al. (119) correlated their results as follows:

$$t_m = MBo \tag{28}$$

where M is a constant equal to 0.093 (80) or to 0.089 (119). The coefficient M given by Verlaan et al. is in exact agreement with the theoretical relationship derived by Murakami et al. (118) for $Bo > 50$ and a degree of inhomogeneity, $I = 0.05$.

Equation 28 shows that the circulation path, which enters in the definition of Bo, has a linear effect on the mixing time. If the mean circulation time and the axial dispersion coefficients are known, it is possible to theoretically estimate the mixing time using equation 28. Experimental details must, however, be carefully planned to avoid complications. Note that in order to simplify data processing it is important to inject the signal and to measure the response at exactly the same point (120) (often the position of the injection point is not specified despite its effect on the mixing time). In a study of the effect of the injection point on the dynamics of the mixing time, Schügerl et al. (12) concluded that the gas–liquid separator is the best choice for tracer injection for short mixing times. Fields and Slater (114) reported a marked dependence of the respiratory quotients upon the injection point of methanol during unlimited fed-batch growth of *Methylophilus methylotropus* in a concentric-tube ALR. The lines in Figure 19 show experimental data for Bo (as overall values) reported for different types of ALR (in which the reactor is considered as a single unit). The dimensions of the reactors are given in Table 4. Because of the definition of the overall Bo, the values are specific to the reactor for which they were obtained and can be used only as indication of trends and orders of magnitude.

As explained above, the ALR is, in fact, a combination of several regions having quite different fluid dynamic characteristics. The overall mixing is the result of the contributions of each of them, and the overall Bo represents

Table 4. Bodenstein Number as a Function of the Superficial Gas Velocity (Key for Fig. 19)

Curve number	Authors	Type	A_d/A_r	H/D_r	Liquid	Ref.
1	Weiland	External loop	0.25	85	Water	44
2	Weiland	External loop	0.25	85	2-propanol 1.65%	44
3	Weiland	External loop	0.25	85	CMC 50%	44
4	Lin et al.	External loop, unbaffled	0.11	20	Water	119
5	Lin et al.	External loop, baffled	0.11	20	Water	119
6	Lin et al.	External loop, unbaffled	0.11	40	Water	119
7	Bello et al.	External loop	0.69	12	Water	45
8	Bello et al.	External loop	0.69	12	Water	45
9	Bello et al.	External loop	0.11	20	Water	45
10	Moor Nagar	External loop	1.0	30	Water	122
11	Fields and Slater	Concentric tube	1.56	10.5	Water	114
12	Fields and Slater	Concentric tube	1.56	21	Water	114
13	Fields and Slater	Concentric tube	1.56	10.5	Water, antifoam	114
14	Fields and Slater	Concentric tube	1.56	10.5	1% Ethanol	114
15	Bello et al.	Concentric tube	0.13	39.65	Water	112
16	Bello et al.	Concentric tube	0.56	55.6	Water	112
17	Verlaan et al.	External loop, total *Bo* number	0.25	16.5	50 mM KCl aqueous solution	79
18	Verlaan et al.	*Bo* number in the riser	0.25	16.5	50 mM KCl aqueous solution	79
19	Verlaan et al.	*Bo* number in the downcomer	0.25	16.5	50 mM KCl aqueous solution	79
20	Verlaan et al.	External loop, total *Bo* number	0.25	16.5	50 mM KCl aqueous solution	79

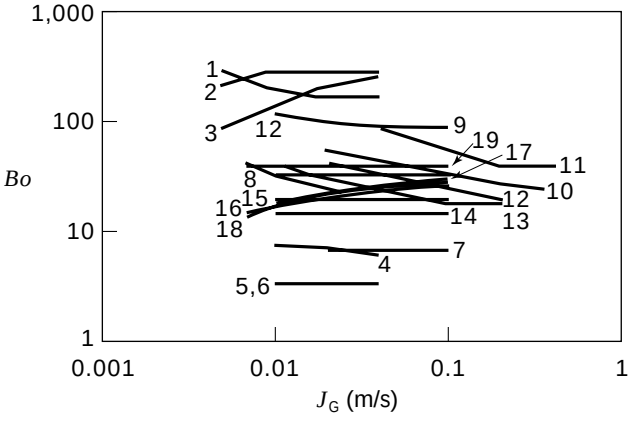


Figure 19. Global Bodenstein numbers reported for ALRs as a whole unit. See references in Table 4.

this combination for a particular reactor. This value thus has some limitations for extrapolation to other configurations or scales. Because each section has different mixing characteristics (it may have a different cross-section, flow configuration, etc.), from a strictly engineering point of view, the mixing in each of the sections of the ALR should be defined and considered separately.

One way of obtaining information on the mixing characteristics of each of the regions of the ALR is the simultaneous measurement of the response in the ALR at several points, so that after one single pulse injection the response of each section in the loop can be obtained (37,80). This method of measurement has the advantage that multiple measurements are made for the same tracer injection experiment. This enables us to check the consistency of the liquid velocity results obtained, since independent measurements can obtained in the same run, as can be seen in Figure 20.

The results for *Bo* obtained by Verlaan (80) in an external-loop ALR are shown in Figure 21 (the much higher *Bo* in the downcomer was explained by the fact that the data were obtained under conditions of complete gas disengagement so that only liquid flowed). The results indicated that most of the mixing took place in the gas separator, and both riser and downcomer could be considered as plug-flow conduits. The same conclusions were drawn for internal-loop ALRs by Merchuk et al. (37) by analysis of mixing times: the shift in results for mixing times obtained with different probes in response to the same impulse perturbation indicated that the riser and the downcomer were introducing a pure delay in the response, that is, that they behaved as plug-flow sections.

Indeed, the simplest model of an ALR, from the point of view of fluid dynamics, is a combination of two plug-flow reactors, representing the riser and the downcomer, and a perfectly stirred tank, representing the gas separator (Fig. 22) (67,123). Merchuk and Yunger (67) showed that this simple model could take into account the mixing in the ALR. The validity of using this simple model as a first approximation is supported by experimental evidence that shows that the mixing time decreases when the separator volume increases (31,124). When the volume of the separator is increased substantially without changing the reactor diameter, the gas separator becomes tall and slender and may depart from total mixing behavior. It has been reported that under such conditions the circulation time measured from two successive peaks of a probe is not influenced by changes in the top clearance (42,43). This indicates that the gas separator acts as two interconnected well-mixed regions, one of them (the lower one in Figure 23) being the link between riser and downcomer, as shown schematically in Figure 23.

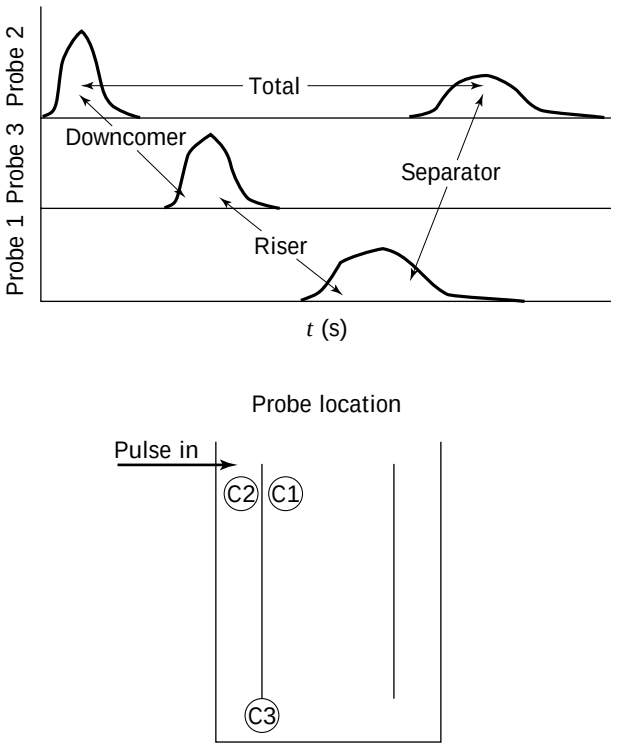


Figure 20. Response of three electrodes in the riser, separator, and downcomer to a pulse of electrolyte. The information that can be obtained from a single experiment is indicated by the *arrows*. From Merchuk et al. (37).

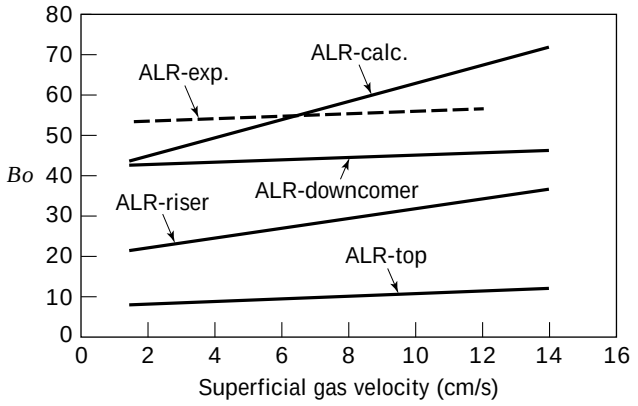


Figure 21. Bodenstein numbers as a function of the superficial gas velocity in an external-loop ALR. ALR-exp, experimental values for the whole reactor; ALR-calc, calculated values for the whole reactor. Adapted from Vorlaan et al. (80).

The extent of mixing in the gas separator of internal-loop ALRs can probably be compared to that in a bubble column of the same dimensions. In the case of external-loop ALRs, the extent of mixing depends on the geometric design. In special designs, like the channel-loop reactor (125), a different approach must be taken. However, in general, a rule of thumb may be recommended: consider that all the mixing takes place in the gas separator, and calculate the degree of mixing in it using published data for bubble columns. Correlations available for bubble columns can thus be used. Godboole and Shah (126) recommend the use of the correlation proposed by Deckwer et al. (127):

$$D_Z = 0.678 \cdot D^{1.4} \cdot J_G^{0.3} \qquad (29)$$

where D_Z is the dispersion coefficient and D is the column diameter, or the relationship proposed by Joshi (128):

$$D_Z = 0.33(U_c + J_L)D \qquad (30)$$

where J_L is the superficial liquid velocity and U_c is the cell circulation velocity given by:

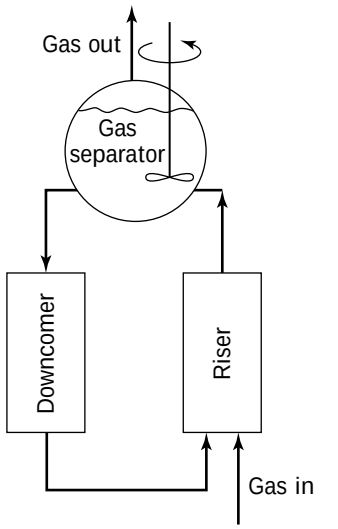


Figure 22. Simple model of an ALR representing the riser and the downcomer as plug-flow reactors and the gas separator as a perfectly mixed region. From Merchuk and Siegel (67).

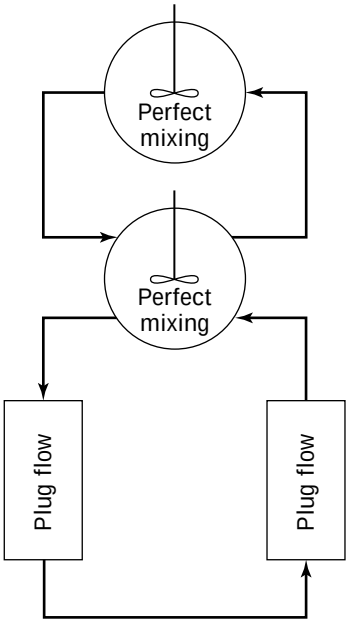


Figure 23. Schematic representation of the effect of a tall slender gas separation region in a concentric-tube ALR.

$$U_C = 1.31\left[gD\left(J_G - \frac{\varphi_G}{1 - \varphi_G} J_L - \varphi_G U_{b\infty}\right)\right] \quad (31)$$

where $U_{b\infty}$ is the terminal bubble velocity.

The liquid superficial velocity (J_L) can be easily calculated for external-loop reactors, and for internal loop reactors, J_L may be taken as half the liquid velocity in the riser. The viscosity may also play a role in the rate of mixing. An increase in liquid viscosity will increase the energy dissipation in the loop and result in an increase in mixing time and *Bo*. In non-Newtonian fluids, however, the behavior may be different, as shown in Figure 24. It has been reported that in the case of the addition of polymers that confer pseudoplastic behavior to the liquid, low concentrations produce a decrease of the mixing time (65,129). This can be explained in terms of the drag reduction due to the presence of polymers in the boundary layer near the walls. Fields et al. (129) found that the mixing time increased for concentrations of the natural polymer, xanthan gum, above the critical concentration. (The critical concentration is a theoretically calculated value at which appreciable overlapping of polymer molecules occurs and which marks the onset of a rapid increase of the apparent viscosity [130]).

Mixing in the Gas Phase

For all practical purposes, the gas in the riser of an ALR exhibits plug flow behavior. Only for extremely high J_G or hindered liquid circulation will the axial dispersion of the bubbles have some effect on the gas RTD. In the downcomer, the gas flow is almost plug-flow when the bubble recirculation is fully developed. But at the stage at which a stationary phase of suspended bubbles appears at the top of the downcomer, appreciable dispersion will occur. This zone has a large degree of mixing due to coalescence and consequent rise of larger bubbles amid smaller ones, with repeated events of breakup and coalescence. However, this type of operation has no relevance to practical applications. It is an operation mode to be avoided at all costs. Indeed, no data on mixing under these conditions have been reported. The main question related to the mixing of the gas phase is, in fact, related to gas recirculation. When a particular gas flow has developed in the downcomer, part of the gas is being recirculated (see "Liquid Velocity"). A pulse of gas tracer at the inlet would produce, as a response, a series of pulses, separated one from the other by the gas circulation time. In practice, not many of these pulses would be detected, due to dilution and disengagement of the tracer in each pass through the separator.

The only reported study on gas phase mixing in an ALR is that of Frehlich et al. (131). The distribution of the gas residence times in two reactors, one of 0.08 m^3 and the other of 4 m^3, was measured using pseudostochastic tracer signals and a mass spectrometer. The values of *Bo* were calculated from the first and second peaks, indicating the main gas stream and the recirculation. Figure 25 shows the *Bo* obtained for a laboratory-scale ALR. The values obtained in a pilot plant were of the same order.

Energy Dissipation and Shear Rate in Airlift Reactors

ALRs are being increasingly used in processes involving shear-sensitive cells (mammalian, insect, and plant cell cultures) (3,11,14,18,21,28,29). This situation has created the need for considering shear stress as one of the parameters relevant in the design of such reactors. Although the a priori evaluation of shear rates is a matter that has been studied for many years in stirred tanks, information on this subject is scanty for pneumatically stirred reactors. The first approach made in this direction in pneumatically agitated vessels was that of Nishikawa et al. (70), who were interested in the problem of heat transfer to a non-Newtonian liquid in a bubble column. This study was extended to mass transfer by Nakanoh and Yoshida (132), who proposed the expression:

$$\gamma = 5000\ J_G \quad (32)$$

This expression has been widely accepted despite the criticism sometimes leveled at it (62,68,123,133). Some modifications have thus been proposed (92,95,134), which, like the original approach, were based on data for bubble

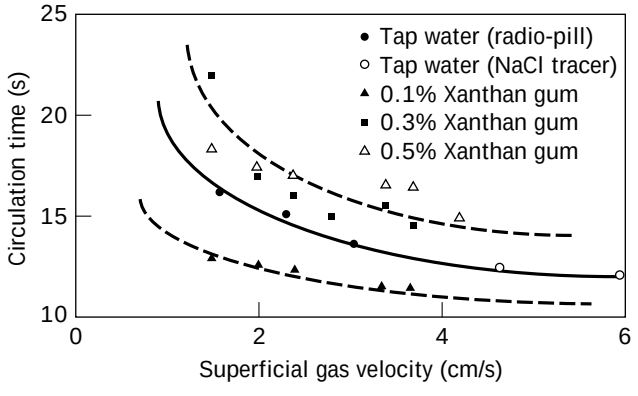


Figure 24. Variation of mean circulation times with J_G for several solutions. Adapted with permission from Fields and Slater (129).

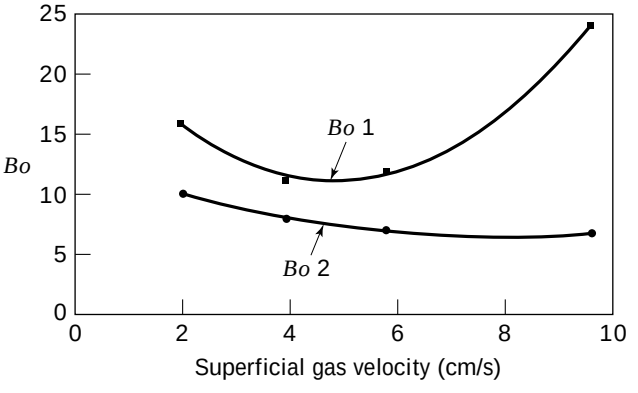


Figure 25. Bodenstein numbers of the first and second peaks of the gas residence time distribution in an 80-L internal-loop ALR, as a function of the gas superficial velocity. From Frehlich et al. (131).

column performance and therefore have limited application for ALRs. Recently, a method facilitating the prediction of the distribution of the energy dissipated in an ALR, based on a simple thermodynamic approach, has been developed (77). Energy dissipation was considered to occur in the ALR by two main mechanisms, wall friction and bubble-associated dissipation (ideal gas behavior was assumed). The work done by the gas on the liquid (and vice versa) was expressed assuming isothermal expansion of the bubbles. The energy dissipation inside the gas phase was considered negligible. The general energy balance (135) was written as:

$$\Delta(PQ) + \Delta E_p + E_D = W_S \tag{33}$$

In this equation, the first term represents the flow work lost by the system under consideration, E_D is the energy dissipated per unit of time, and W_S is the shaft work done by the surroundings on the system under consideration. The schematic representation of the concentric-tube ALR in Figure 26 indicates the different points in the reactor considered in the mathematical expressions. The expressions found for the energy dissipated in each zone were the following:

Riser

$$(E_d)_R = Q_L(P_4 - P_5) - Q_L\rho_L gh - Q_r P_4 \ln\left(\frac{P_5}{P_4}\right) \tag{34}$$

Gas separator

$$(E_d)_S = Q_L(P_5 - P_2) - Q_{in}P_4 \ln\left(\frac{P_1}{P_5}\right) - Q_d P_4 \ln\left(\frac{P_2}{P_5}\right) \tag{35}$$

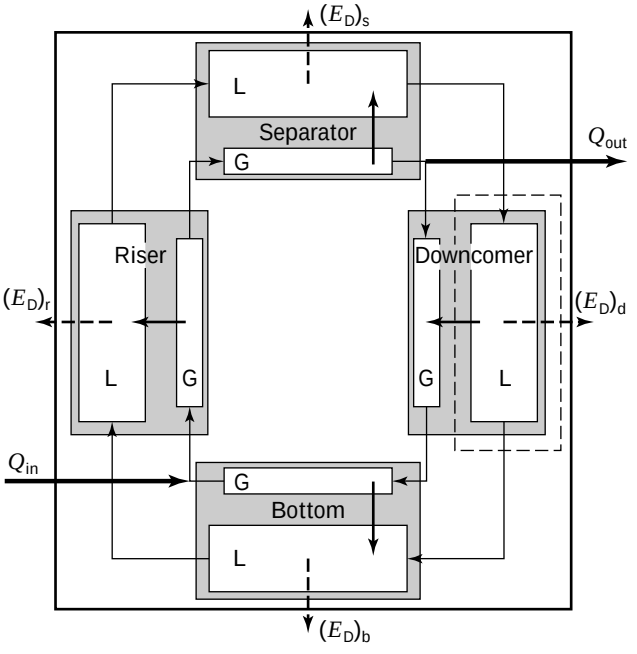


Figure 26. Schematic description of the variables in the thermodynamic model for energy dissipation distribution in an ALR. From Merchuk and Berzin (77).

Downcomer

$$(E_d)_d = Q_L(P_2 - P_3) - Q_L\rho_L gh - Q_d P_4 \ln\left(\frac{P_3}{P_2}\right) \tag{36}$$

Bottom

$$(E_d)_b = Q_L(P_3 - P_4) - Q_d P_4 \ln\left(\frac{P_4}{P_3}\right) \tag{37}$$

Results of this model of an ALR can also be used to estimate the global shear rate in each region of the reactor, according to the global approach presented Merchuk and Ben-Zvi (Yona) (96). The shear stress in the liquid of each region of the reactor can be defined as the energy dissipated divided by the mean path of circulation in the region and by the sum of the areas of all the bubbles. For the region i in the ALR

$$\tau_i = \frac{(E_d)_{bulk,i}t_i}{a_i h_i^2 A_i} \tag{38}$$

where t_i is the residence time of the liquid, h_i is the effective length, and a_i is the specific interfacial area, in the region i.

A global shear rate γ_i can be calculated for each region i as

$$\gamma_i = \frac{\tau_i}{\mu} \tag{39}$$

where μ is the effective viscosity of the fluid.

For liquids exhibiting different types of rheological behavior, the corresponding constitutive equation must be used. Such an approach has been used for the interpretation of shear effects on mammalian cells (136) and algal growth (43).

MASS TRANSFER

The volumetric mass transfer coefficient ($k_L a$) is the rate of gas transfer across the gas-liquid interface per unit of driving force (the driving force is the gas concentration gradient between the liquid and the gas). The mass transfer coefficient $k_L a$ can be seen as the product of two terms: the mass transfer coefficient k_L and the specific interfacial area a. Both terms depend on a series of variables that can be grouped into three categories: (*1*) static properties of the liquid, such a density, diffusivity, and surface tension; (*2*) dynamic properties of the liquid (related to liquid flow), such as rheological parameters; and (*3*) liquid dynamics. In general, the variables in group 1 do not change very drastically. The variables in groups 2 and 3, however, may span wide ranges.

Mass Transfer Rate Measurements

Methods for the determination of $k_L a$ in a reactor can be grossly classified as steady-state and nonsteady-state methods. In the steady-state methods, the rates of oxygen uptake in steady-state operation are evaluated, either by measurement of inlet and outlet rates of oxygen or by di-

rect analysis of a compound that reacts with the oxygen, as in the case of the sulfite method. One of the problems associated with these procedures is that the changes in oxygen concentration in the gas streams are usually small, and the errors of measurement thus have a substantial influence. When a chemical is added to the system, there is the question of whether the addition has provoked changes in the physicochemical properties, which thus become different from the properties of the original system.

Transient methods may be applied to follow the response of the dissolved oxygen concentration in the system after a step-change of oxygen concentration in the inlet gas stream. These methods have the advantage that addition of an alien material is not required and that a single concentration is measured. The correct use of this method has been analyzed in depth by Linek et al. (137).

One important point to take into account is the dynamics of the oxygen electrode. The lag in the response of the oxygen electrode makes it necessary to discern between the electrode response and the real oxygen concentration, especially when close to a sharp change in concentration. A correct analysis should also include the model of the dynamic behavior of the electrode. In order to simplify these procedures, approximations based on truncation of parts of the response curve have been proposed (133,138). These methods are based on truncating the first part of the electrode response obtained in a transient experiment. Once the error included in the value of k_La is set, the extent of truncation is fixed, allowing simplification of the analysis of the remainder of the curve. It should be kept in mind that this simplification implies the loss of part of the information, and due care should be given to statistical analysis of the results.

Variations of the method have been proposed to minimize disturbances in the system by introduction of step variations of agitation or pressure (137,139). In this way, the method can be applied to bioreactors during real operation of the system.

One problem that may appear in the measurement of mass transfer rates, especially when viscous liquids are used, is related to the presence of very small bubbles that are depleted of oxygen very rapidly but do not disengage in the gas separator, thus constituting an inert volume of gas in the reactor (134). Kawase and Moo-Young (140) analyzed the use of transient absorption of CO_2 for the determination of k_La and concluded that the error due to small bubble retention was much smaller than that in the case of O_2.

Whatever the method used for the determination of k_La, the process of data elaboration is basically the same as is shown schematically in Figure 27 (141). The measured variable (usually oxygen concentration) is compared to the value predicted by a mathematical model of the process. The model includes k_La as a parameter, and the value of k_La that gives the best fit is chosen. It is thus obvious that the choice of the model is very important, and poor assumptions on flow characteristics of gas or liquid phases may lead to errors and deviations from the true values (142). All the models are a simplified (amenable to mathematical treatment) description of the system. It should be kept in mind, therefore, that k_La is not a property of the

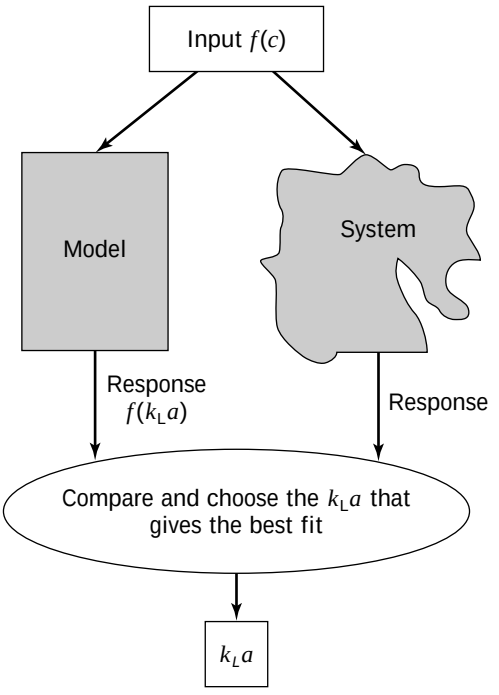


Figure 27. Steps in the determination of the mass transfer coefficient (k_La). From Merchuk et al. (141).

system, but a parameter of the model adopted. If total mixing is assumed in the model adopted, the mass transfer coefficient obtained will consequently be limited. The assumption of complete mixing is such a common practice that many reports do not even specify explicitly that this has been done, and in many texts the mass transfer coefficient k_La is defined by the equation:

$$\text{OTR} = k_1a(C^* - C) \tag{40}$$

It is often forgotten to state that this equation is valid only for perfectly mixed systems.

Strictly, the different sections of ALRs (riser, downcomer, and gas separator) have different flow characteristics, and the mass transfer coefficient may be expected to differ from one region to another. Some researchers (108) have assumed that the contribution of the downcomer to the overall mass transfer is negligible and have reported values of k_La that are, in fact, the entire mass transfer divided by the volume of the riser. The fact that the values of the mass transfer coefficient are calculated and reported in different ways introduces some ambiguity into the published data; special care should therefore be taken when comparing data or using published mass transfer coefficients for design purposes.

The assumption of perfect mixing in the liquid phase may be questionable in the case of tall reactors. One very simple method to verify this assumption is the simultaneous measurement of the response of the concentration to a step-change in the inlet. In a perfectly mixed system, the location of the probe should be irrelevant. Figure 28 (143) shows the response of three probes, one located at

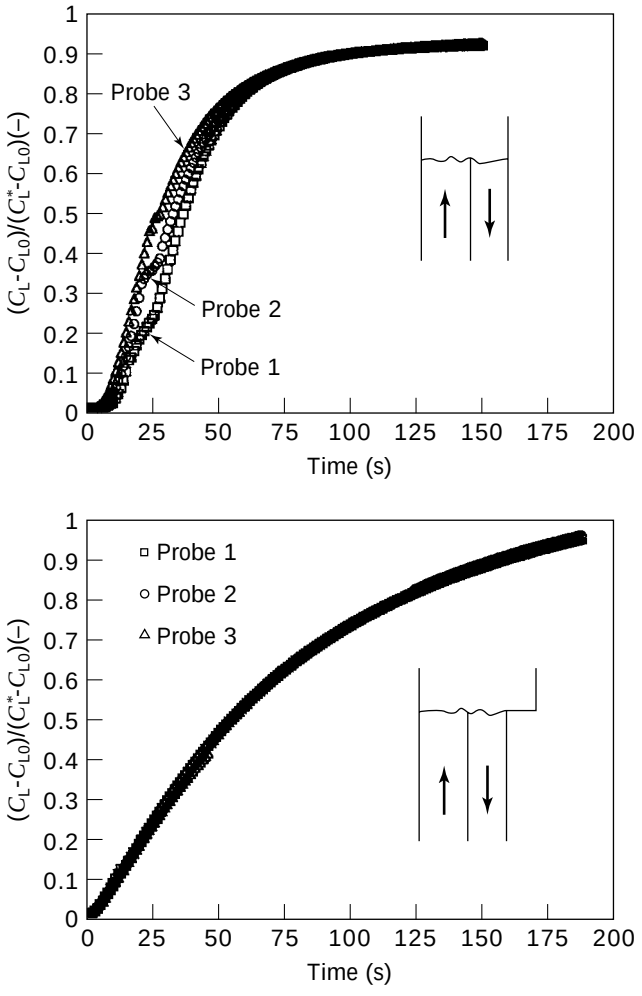


Figure 28. Response of three probes, one located at the end of the riser, one at the inlet of the downcomer, and one at the bottom of a split-cylinder ALR. From Siegel et al. (143).

the end of the riser, one at the inlet of the downcomer, and one at the bottom of a split-cylinder ALR. The difference between the figures stems from a different gas separator section, which changes the fluid dynamics in the system. In one case the system behaved as perfectly mixed, because of the large volume of the gas separator and the faster liquid circulation. In the other (closed system in the figure), the responses of the three electrodes were clearly different, calling for a different analysis. The validity of the criterion, originally proposed by Andre et al. (144), that compares the circulation time in the reactor (t_c) with the characteristic time for mass transfer, which is the inverse of the mass transfer coefficient, was thus confirmed (143):

$$t_c \cdot k_L a < 0.5 \tag{41}$$

If this condition is not fulfilled, the bioreactor cannot be considered as a perfectly mixed volume, and more sophisticated ways of analysis are required. In this case, the mathematical model, and, consequently, the analysis of the experimental data become much more complicated. Nevertheless, this approach has been used (48,122,145,146), and the mass transfer experiments render values of $k_L a$ for each of the regions of the ALR. Figure 29 shows the results reported by Hwang and Lu (146) in an internal-loop ALR. The graph shows that the mass transfer takes place at the highest rate in the riser. Values for the downcomer are 50% lower, and those for the gas separator are intermediate between the riser and the downcomer. The overall mass transfer rate is the result of the balance between the volumes and rates in the three sections.

Bubble Size and Interfacial Area

As said earlier, the interfacial area per unit volume is an important component of the volumetric mass transfer coefficient. In fact, it is the part of $k_L a$ that is most susceptible to changes in operation variables and fluid properties. The mass transfer coefficient k_L varies only within a limited range (147), but the interfacial area is the main component responsible for the changes in mass transfer rate due to variations in turbulence, initial bubble size, and liquid properties.

The methods for measurement of the interfacial area are based either on rapid chemical reactions or on direct measurement of bubble size. If a mean bubble size can be defined, then the interfacial area can be evaluated with the aid of the holdup measurement, since, in a population of homogeneous bubble size, it can be applied:

$$a = \frac{6\varphi}{d_s} \tag{42}$$

where the Sauter mean diameter (d_S) is given by

$$d_S = \frac{\sum n_i d_i^3}{\sum n_i d_i^2} \tag{43}$$

and n_i is the number of bubbles of diameter d_i

There are very few published data on bubble size applicable to ALRs. For the riser, the correlation presented by Miyahara et al. (57) for the volumetric mean diameter

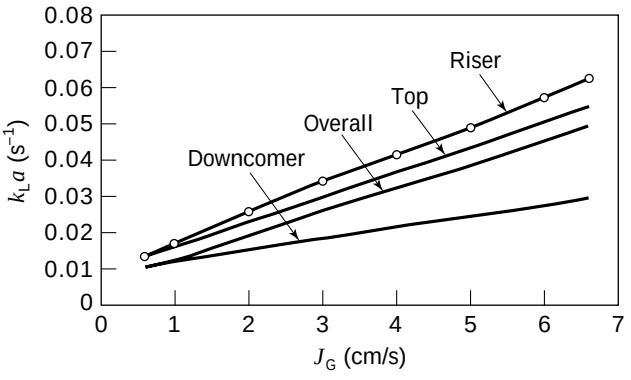


Figure 29. Influence of the superficial gas velocity on overall $k_L a$ and on the $k_L a$ in each of the regions of an ALR. Adapted from Hwang and Lu (146).

of the bubbles in the riser of a concentric-tube ALR can be used:

$$\bar{d}\left(\frac{g\rho_L}{\sigma d_0}\right)^{1/3} = f(N_W) \tag{44}$$

where d_0 is the diameter of the sparger orifice and σ the surface tension,

$$N_W = \frac{We}{Fr^{0.5}} \tag{45}$$

and the function $f(N_W)$ is different for each range of N_W:

$$\begin{aligned} f(N_W) &= 2.9 && N_W < 1 \\ f(N_W) &= 2.9N_W^{-0.188} && 1 < N_W < 2 \\ f(N_W) &= 2.9N_W^{-0.5} && 2 < N_W < 4 \\ f(N_W) &= 3.6 && 4 < N_W \end{aligned} \tag{46}$$

Data on bubble size in the downcomer of an ALR were published by Popovic and Robinson (63). No data on interfacial area measured by chemical methods have been published.

Data Correlations for Mass Transfer Rate

There are two ways of correlating experimental data from ALRs. First, the hydrodynamic point of view suggests that the movement of the fluid in the reactor determines its overall behavior; the gas superficial velocity is therefore the more appropriate independent variable. Second, the thermodynamic point of view is based on a consideration of energy balance, a more global approach to the system. This will lead to correlation of the phenomena in the system as a function of the energy input. Indeed, it is easier to compare mass transfer coefficients in ALRs with those in conventional reactors when the data are presented as a function of the total power input (both mechanical and pneumatic) per unit volume of the medium (34) (Fig. 30).

Figure 31 shows data for mass transfer coefficients as a function of specific power input (149). References and geometric characteristics can be seen in Table 5.

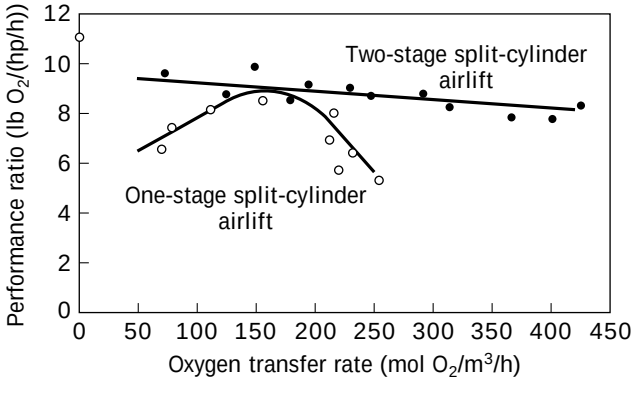


Figure 30. Performance of ALRs. Adapted from Orazem and Erickson (148).

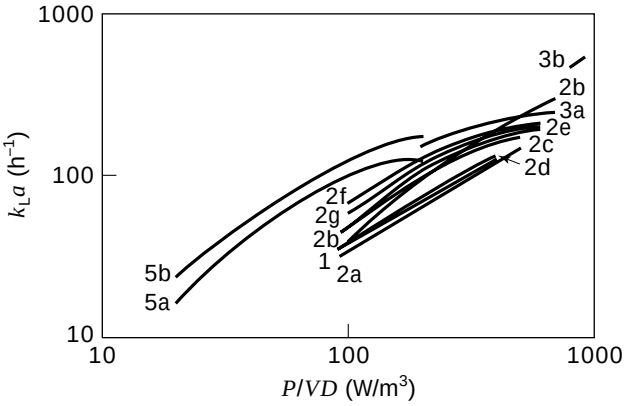


Figure 31. Mass transfer coefficients as a function of specific power input. With permission from Siegel et al. (123). See Table 5.

Selected of correlations proposed for the prediction of mass transfer coefficients are shown in Tables 6 and 7 for internal- and external-loop ALRs, respectively, and the predictions are compared in Figures 32 and 33. Among the internal-loop reactors, two correlations presented for reactors with a rectangular cross-section are shown (32,154). For external-loop ALRs, the correlation by Popovic and Robinson (88) is recommended. In the case of internal-loop ALRs, most of the correlations predict similar values. The correlation of Merchuk et al. (37) can be recommended on the basis that more geometric parameters have been taken into account, and this gives greater generality. Nevertheless, the general considerations already expressed when analyzing other correlations in this chapter are valid here as well.

HEAT TRANSFER

Because of the relatively low reaction rates of processes involving microorganisms and cells, it may—in a very general way—be said that heat-effect problems related to local variations of temperature are not common in bioreactors. Even in the case in which polymeric products are released into the medium and very high viscosity is reached, heat transfer is not the controlling step, because such viscous media will hinder the mass transfer, and heat generation will consequently be limited. In such cases the main point of focus is thus, mass, rather than heat, transfer. Reactions catalyzed by immobilized enzymes, however, may require different considerations, because of higher reaction rates.

There is a far greater body of published data on heat transfer in bubble columns than in ALRs, and some of the basic observations are valid for both types of reactor. The heat transfer rate in bubble columns is much larger than that expected from single-phase flow (155). This is a result of the bubble-driven turbulence and liquid recirculation, which are characteristic of the flow in pneumatically agitated reactors.

Several correlations have been proposed for the prediction of the heat transfer coefficient in these reactors. Recently, Kawase and Moo-Young (62) presented an expres-

Table 5. Mass Transfer in ALRs

Curve no.	Reactor type	Height (cm)	*D* Riser (cm) draught tube	*D* Downcomer (cm) reactor tube	A_d/A	Ref.
1	Concentric tube	143	21	30	1.04	150
2a	External loop	180	15.2	76 + 10.2	0.69	45
2b	External loop	180	15.2	7.6	0.25	45
2c	External loop	180	15.2	5.1	0.11	45
2d	Concentric tube	180	8.9	15.2	0.56	45
2e	Concentric tube	180	7.6	15.2	0.35	45
2f	Concentric tube	180	5.1	15.2	0.13	45
2g	Bubble column	180	0	15.2	0.00	45
3a	Rec. split vessel; all experiments, except high recirculation and two spargers	435–450	9 × 250	7 × 25	0.78	32
3b	Rec. split vessel; high recirculation and two spargers	435–450	9 × 25	7 × 25	0.78	32
4	External loop	850	10	5	0.25	78
5a	Concentric tube	170	11.7	20	1.96	44
5b	Concentric tube	170	17.6	20	0.26	44

Note: Curve numbers correspond to Figure 31.

Table 6. Mass Transfer in Internal-Loop ALRs

No.	Formula	Ref.
1	$Sh = 3 \cdot 10^4 Fr^{0.97} M^{-5.4} Ga^{0.045} \left(1 + \frac{A_d}{A_r}\right)^{-1}$	1
2	$Sh = 2.66 Sc^{0.5} Bo^{0.715} Ga^{0.25} \left(\frac{D_r}{D}\right)^{-0.429} \varphi^{1.34}$	151
3	$k_L a = 0.0343 J_{Gr}^{0.524} \mu_{ap}^{-0.255}$ $\gamma = 5000 J_{Gr}$ for $J_{Gr} \geq 0.04$ ms^{-1} $\gamma = 5000 J_{Gr}^{0.5}$ for $J_{Gr} \leq 0.04$ ms^{-1}	48
4	$Sh = 0.68 n^{-0.72} Fr^{0.38n+0.52} Sc^{0.38-0.14n}$	59

Table 7. Mass Transfer in External-Loop ALRs

No.	Formula	Ref.
1	$K_L a = k_1 J_{Gr}^{0.8} \left(1 + \frac{A_d}{A_r}\right)^{-2}$ Water $K_L = 0.75 \mu^{-0.8} \sigma^{-0.2}$ NaCl $K_L = 0.79 \mu^{-0.8} \sigma^{-0.2}$	45
2	$k_L a = 0.5 \cdot 10^{-2} J_{Gr}^{0.52} \left(1 + \frac{A_d}{A_r}\right)^{-0.85} D_1^{0.5} \rho_1^{1.03} \sigma_1^{-0.25} \mu_{ap}^{-0.89}$ or in simplified form: $k_L a = 1.911 \cdot 10^{-4} J_{Gr}^{0.52} \left(1 + \frac{A_d}{A_r}\right)^{-0.85} \mu_{ap}^{-0.85}$ $k_L a = 0.24 J_{Gr}^{0.837} \left(1 + \frac{A_d}{A_r}\right)^{-1}$	88
3	in case of suspended solids: $k_L a = (0.349 - 0.102 C_s) J_{Gr}^{0.837} \left(1 + \frac{A_d}{A_r}\right)^{-1}$	152
4	$k_L a = 913 \left(\frac{P}{VD}\right)^{1.04} (U_L)^{-0.15}$	32
5	$k_L a = (0.349 - 0.102 C_s \left(1 + \frac{A_d}{A_r}\right)^{-1} J_{Gr}^{0.837}$	152

sion that satisfactorily fits most of the published data. The model is based on Levich's (156) three-zone concept and Kolmogoroff's (157) isotropic turbulence theory and has no empirical adjustable parameters. The model can take into consideration non-Newtonian behavior of the liquid and predicts the enhancement of the heat transfer due to the shear-thinning effect of the fluid. The dimensionless expression may be written as:

$$Nu = 0.075(10.3n^{-0.63})^{\beta} n^{1/3} (Pr^*) \frac{1}{3} Fr^{-\beta} (Re^*)^{\beta + 3(n+1)} \tag{47}$$

with

$$\beta = \frac{4 - n}{6(n + 1)}; \quad Pr^* = \frac{\kappa D^{1-n}}{\left(\frac{\kappa}{C_p}\right) J_G^{1-n}}; \quad Re^* = \frac{D^n J_G^{2-n}}{\left(\frac{k}{\rho}\right)} \tag{48}$$

where the symbols defined in "Nomenclature" have been used.

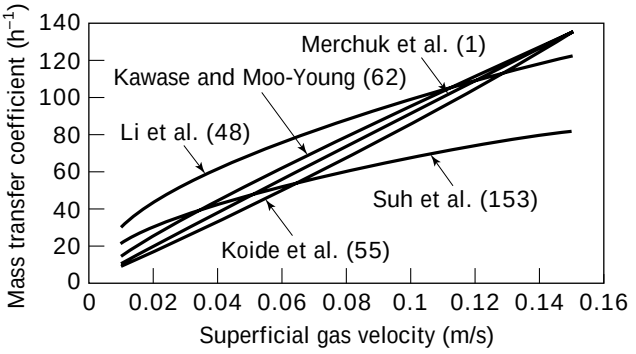


Figure 32. Mass transfer coefficient $k_L a$ as a function of gas superficial velocity for internal-loop ALRs. Correlations are presented in Table 6.

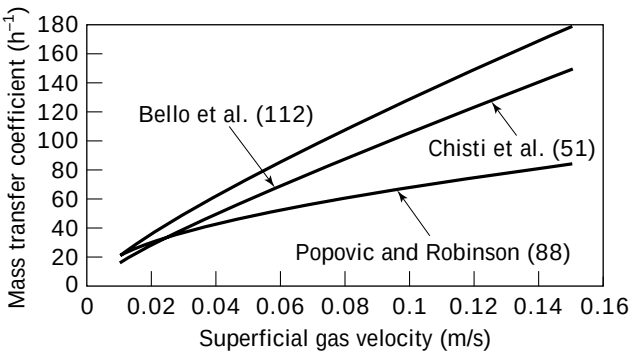


Figure 33. Mass transfer coefficient k_La as a function of gas superficial velocity for external-loop ALRs. Correlations are presented in Table 7.

For Newtonian liquids, this equation reduces to:

$$Nu = 0.134^{1/3}(Pr)^{1/3}Fr^{-1/4}(Re)^{3/4} \qquad (49)$$

Kawase and Moo-Young (140) compared their model with data and correlations published by various investigators and found satisfactory agreement for both Newtonian and non-Newtonian liquids.

For ALRs (64), the flow in the reactor may be similar to that in bubble columns if the internal recirculation is high, or it may be even closer to net two-phase flow in pipes, in which case the equations for heat transfer in pipes can be used (126).

Blakebrough et al. (158) studied heat transfer rates in an external-loop airlift fermentor used for culturing *Aspergillus niger*. They concluded that the enhancement of the heat transfer rate could be explained by disturbances in the liquid layer near the surface caused by the presence of the microorganisms.

THREE-PHASE AIRLIFT REACTORS

The special qualities of the ALR stem, as stated before, from its fluid dynamic characteristics. One such characteristic is the directionality of the liquid flow. Independently of superimposed fluctuations, a clear net flow is present in the reactor, with exception of the gas separator in internal-loop designs. Therefore, it is to be expected that the fluidization capacity of the ALR will be markedly superior to that of a bubble column. Several studies have been conducted on the suspension of solids in ALRs, particularly on the use of this type of device for catalytic processes in the chemical industry, where the solid support is usually heavy (61,159–163). In this regard, a very important point is the minimum gas superficial velocity that leads to complete solid fluidization (61,87,159,162,164–166). Hysteresis has been observed in some cases; once total fluidization has been attained, the superficial velocity can be reduced to values lower than that required to reach this state. This is due to the high pressure drop related to passage of liquid through a bed of solids, before fluidization, as compared to the drag forces required to maintain the solid in suspension after all the solids are suspended. Contradictory data on the effect of the suspended solids on the reactor performance have been reported. Fan et al. (167) claimed that the overall gas holdup increased due to the presence of the solids, whereas Koide et al. (168,169) showed the opposite effect on the gas holdup and reported a small decrease in k_La as well. It is possible that these discrepancies are due to the use of different solids. One of the properties of solids that is often overlooked is wettability. Small bubbles may adhere on wettable solids, leading to a change in the apparent density of the particle and thus changing their solid circulation velocity.

In the case of suspended solids that take an active part in the process, the mass transfer rate from the liquid to the solid may become the limiting step. The dependence of particle size on the mass transfer to the suspended solids has been studied by several authors (19,170,171).

All the comments made above relate to heavy solids. This situation is not very frequent in biological processes, with the exception of biolixiviation (172) or the special case of microbial desulfuration of coal (173). In most of the biological processes that may take place in an ALR, the solids are either cells, clumps of cells, or supports that are not much more dense than the medium. Therefore, neither fluidization nor the distribution of solids in the reactor constitutes a problem. Assa and Bar (174) found very small variations in the axial distribution of animal and plant cells suspended in an internal-loop ALR. This is due to the small free-falling velocity of the solids, which is the reason for the difference in loading between the riser and the downcomer when heavy particles are used (160–162). Because of the small difference between solid and medium density, the movement of the particles usually present in biological processes is not as dependent on gravity forces as on liquid and bubble movement. In this case, the transport of liquid in the wake of the bubbles may be considered to be the prevailing transport mechanism (175). Snape and Thomas (176) proposed a Monte Carlo algorithm for modeling particle movement by this mechanism in bubbly flow to predict distribution of circulation times in the reactor. Koide et al. (169) conducted a broad study of gas holdup and mass transfer rates in an internal-loop ALR containing Ca-alginate beads, which are often used for cell immobilization. The authors found that solid loading had a negative effect on both parameters, but the particle diameter had no influence in the range studied (1.8 mm $< d_p <$ 3.98 mm). Chang et al. (31) studied the influence of suspended pellets (cylindrical pellets of immobilized penicillin acylase) on the mixing time in an internal-loop ALR. They found mixing time increased when the solid volumetric concentration was raised up to 15%. For higher concentrations the trend inverted, and the mixing time decreased. No such effects were observed with heavier solids. Increases in gas superficial velocity and in top clearance both lead to decreases in mixing time in all cases.

AIRLIFT REACTOR—SELECTION AND DESIGN

Scale-up of Airlift Bioreactors

The problems encountered in the scale-up of bioreactors can be concentrated into two groups. The first includes the cases in which a high power input per unit volume is used

on a laboratory scale, but cannot be maintained on an industrial scale due to economic or mechanical limitations. This is not the case for plant or animal cultures, for which a very high specific power input cannot be used because of cell fragility. The second group of problems can be generalized as lack of knowledge about the hydrodynamics of large-volume vessels.

The methods available for scale-up of bioreactors have been reviewed by Oosterhuis (177), Kossen and Oosterhuis (178), Sweere et al. (179), and Sola and Godia (180), among others. Because, in general, design from first principles cannot be undertaken because of the lack of basic knowledge about the hydrodynamics of the bioreactors, one possible solution is the semifundamental method, which comprises using approximate simple models for fluid dynamics and integrating them with basic known kinetics and heat and mass transfer rates.

It does not happen very often that all this information is available with a degree of certainty that allows safe design of a large system. Thus, the designer must usually resort to dimensional analysis. This method requires a knowledge of all the variables affecting the process, which can be obtained from a qualitative, but realistic, model (181).

A simplified version of this method is to limit the number of variables to one or two and to follow rules of thumb (182), which, depending on the specific case, may be constant P/V, constant k_La, etc. The literature shows, however, many cases of inconsistency for this method. For example, design of a scaled-up bioreactor in which the oxygen transfer rate is kept constant can lead to a better performance than expected, as in the example reported by Taguchi et al. (183) for glucoamylase production by *Endomyces* sp., or worse than expected, as in the case of protease production by *Streptomyces* sp. (184). These two examples are shown in Figure 34.

The method of regime analysis and scale down, proposed by Kossen and Oosterhuis (178), combines two tools to overcome the problem posed by the complexity of biochemical reactors. Their method is based on considering the regime of the full-scale process as the objective and planning the strategy of process development from this point on. This method is therefore applicable only to conventional and well-studied bioreactors that are to be used in a new bioprocess.

Regime analysis is based on the consideration that, generally, biochemical processes involve a series of steps, some being mass or heat transfer by convection, some being diffusive mechanisms (activated or not), and others being chemical reaction steps. In the latter case, a mass-transfer mechanism is superimposed, since molecules must encounter one another in order to react, and usually a heat effect will accompany the reaction. Depending on whether these steps take place in parallel or in series and on the relaxation time for each step (185), the rate of the total process is often given as the rate of one single step. But the equilibrium between all the individual rates can be (and usually is) upset by a change in scale. This is to be expected, because a change in scale will not bring a change in the physicochemical or kinetic parameters (scale-insensitive variables), but will affect the overall convective mass and

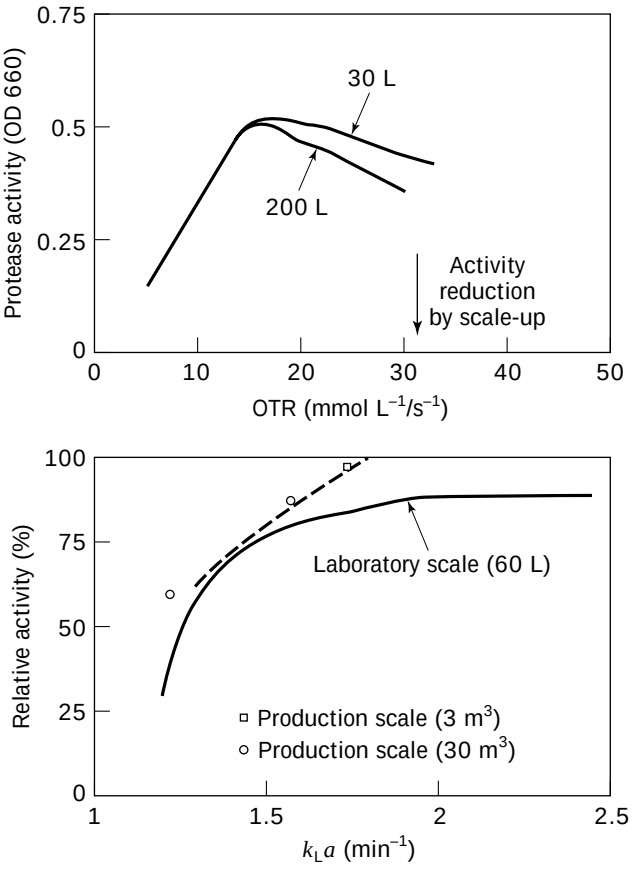


Figure 34. Results for a scaled-up bioreactor with a constant oxygen transfer rate (183,184). OD 660 is the optical density measured at wavelength of 660 nm.

heat rates (scale-sensitive variables). A new equilibrium will be established, and the interplay of all the parameters of the system may lead to a regime in which a different step becomes the step-controlling the process rate.

The method of Kossen and Oosterhuis (178) starts with an analysis of the operation of the large-scale system. Once the regime is clarified, a small-scale system is designed in such a way that it simulates the operation regime of the larger one. Optimization studies can be done on the smaller model, and conclusions will then be extrapolated to the full-scale process. This concept is depicted in Figure 35. An example of this method is given by Oosterhuis (177), in which a large-scale stirred tank reactor is simulated on a laboratory scale by two interconnected vessels, one with a small liquid volume, high agitation rate, and high rate of oxygen supply (representing the zone near the impeller-sparger in the reactor), and a second vessel, with a much larger liquid volume, oxygen consumption (simulated by nitrogen sparging), and gentle agitation. The success of this model suggests that large vessels must be carefully analyzed and their internal structure studied for proper modeling and design to be performed. The analysis starts with the definition of the characteristic time constants of the system.

In the case of low gas recirculation, the relationship between the total circulation time and the residence time in

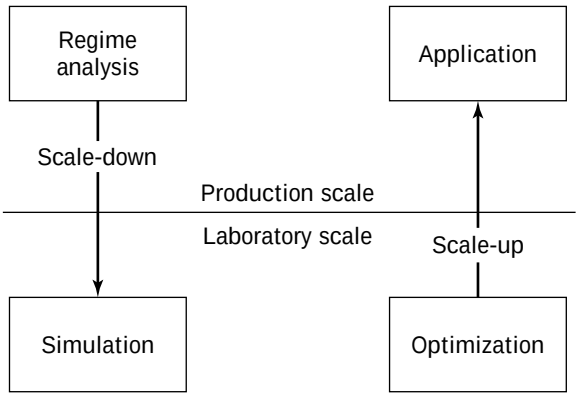


Figure 35. Schematic representation of the scale-down method. Adapted from Oosterhuis (177).

the downcomer may become very relevant. This was shown by MacNeil and Kristiansen (180,186) for the scale-up of an external-loop ALR for the production of *Aureobasidium pullulans*. The poor oxygen supply in the downcomer was reflected on the level of product synthesis, as shown in Figure 36. The smaller residence times lead to higher *A. pullulans* production.

An analysis of the time constants required for the design of an ALR for plant cell culture has recently been presented (187). Figure 37 shows the reported values of the time constants calculated for an external-loop ALR for mixing, mass transfer, and oxygen consumption at two different cell concentrations: 5 and 30 kg/m^3. These graphs exemplify the changes in the controlling mechanism that may take place as the physical properties of the broth change in the course of the bioprocess.

Details on bioreactor scale-up may be found in the recent review of Sola (180). It should be remembered, however, that the reliability of a scale-up will always be limited by the quality of the predictions of gas holdup, liquid velocity, and mass transfer rate.

Design Improvements

In the design of ALR, several modifications aiming at the improvement of some of the characteristics of the equipment have been proposed. One of the earliest modifications was the two-staged ALR, proposed by Orazem and Erickson (148). Their design was inspired by the improvement observed by them in the performance obtained with multistage sieve trays over single-stage bubble columns. They claimed that a substantially higher mass transfer coefficient was obtained, as was a better performance in terms of oxygen transferred per unit of energy invested.

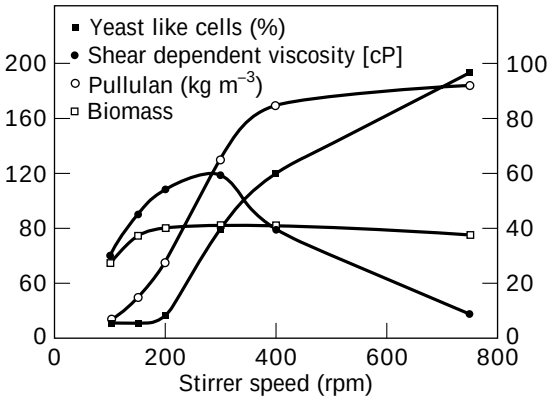


Figure 36. Influence of residence time in the poorly aerated downcomer on the production of *A. pullulans*. Adapted from McNeil and Kristiansen (186).

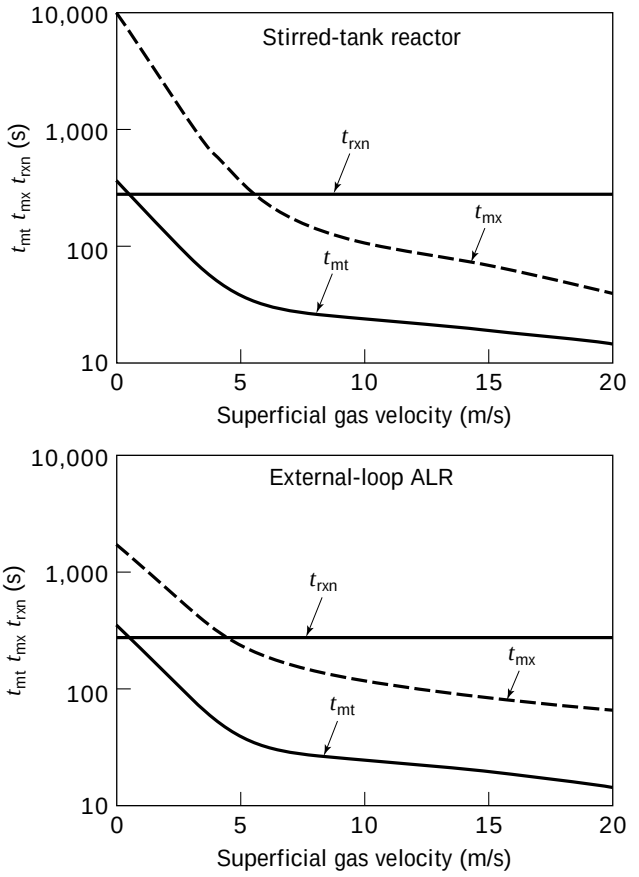


Figure 37. Time constants calculated in an external-loop ALR as a function of the superficial gas velocity (187).

The combination of a concentric-tube ALR with a marine propeller (1000 rpm) was studied by Pollard et al. (188) (Fig. 38). This modification, in which the stirrer was located near the bottom of the draft-tube, served to increase the circulation rate, which may be low for viscous liquids. This configuration showed enhanced oxygen transfer and more uniform distribution of the dissolved oxygen concentration throughout the reactor in comparison with that of the unmodified ALR. The liquid circulation was also intensified. However, the improvement was achieved at the cost of the introduction of a focus of energy input. As has been mentioned before, the absence of such a high shear region is one of the advantages of the ALR. This advantage is, in part at least, lost. It is interesting to note that a bakers' yeast culture (188) grown in the same type of modified reactor did not reveal any improvement. This may be related to the strong shear rate in the vicinity of the propeller.

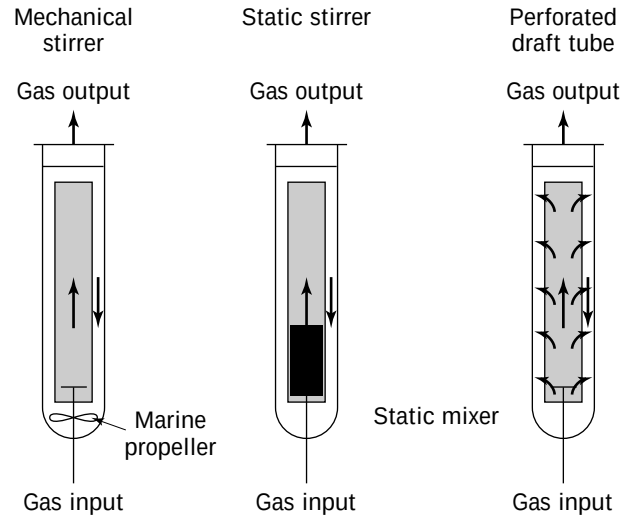


Figure 38. Modifications proposed in ALRs.

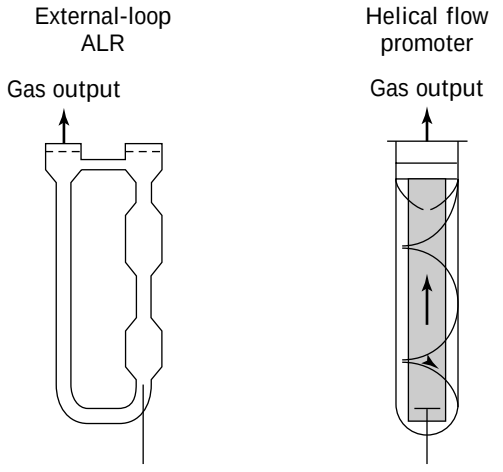


Figure 39. Comparison between a bubble column an ALR and an ALR with helical flow promotors. From Schlotelburg et al. (196).

Different types of static mixers, usually located in the riser, have also been used for enhancing the performance of ALRs (Fig. 38). Potucek (189) studied the influence of static mixers on the gas holdup and interfacial gas-liquid area in a concentric-tube ALR and found an improvement in interfacial area. Enhancement of $k_L a$ has been also obtained, especially in viscous liquids, as reported by Gaspillo and Goto (190) and Chisti et al. (191). Zhou et al. (192) showed better growth of *Cephalosporum acremonium* in an ALR with static mixers.

As mentioned above, it has been shown repeatedly that the mass transfer rate in a bubble column is usually higher than that in the conventional ALR. It therefore makes sense to try to bring into the ALR some of the characteristics of bubble columns in a controlled fashion. This was done by Bando et al. (193,194), who tested a perforated draft tube in a concentric-tube ALR. The perforations in the draft tube facilitated communication between the less-well aerated liquid in the downcomer and the better aerated riser (Fig. 38). The reported improvements in mass transfer rates were undoubtedly obtained at the cost of a reduction of circulation velocity.

Another variant tested is the converging-diverging tube ALR (195), which can be seen in Figure 38. The authors report that the changes in cross-section in the riser produced an increase in the gas holdup.

One of the advantages of the directionality of flow in ALRs is the improved fluidization capacity. The strengthening of this advantage was the aim of another modification, the helical flow promoter, proposed by Gluz and Merchuk (170). The helical flow promoter causes the fluid to flow down in the downcomer (Fig. 39) in a helical pattern. The device comprises several fins or baffles, which have the effect of modifying the flow paths; instead of going in straight lines along the axis, the flow paths move along an helix. The baffles may be installed in a small section at almost any place along the riser or the downcomer, and the effect is perceived throughout the reactor. One of the best positions for the helical flow promoter is the top of the riser. A helical flow is then generated in the downcomer to produce a swirl at the bottom and a corkscrew-like path in the riser. This has a strong potential for the culture of photosynthetic microorganisms. The helical movement causes secondary flow, which leads to enhanced radial mixing, and therefore more homogeneous distribution of light and heat among liquid elements and suspended particles. With the helical flow promoter, it is thus more likely that all the elements of the fluid get the same exposure to light and heat exchange.

One of the most important characteristics of the helical flow promoter is the enhanced capacity for fluidizing solid particles. This is due to the swirls that develop at the bottom of the reactor. Thus, this modification is especially suited to processes operating with cells immobilized on a solid. The minimal gas flow rate for complete fluidization of solids in an ALR may be reduced drastically by the use of a helical flow promoter (170). In addition, the mass transfer rate to suspended solids may be enhanced up to 50%, because of the higher relative velocity between the particles and solids.

Another interesting aspect of the performance of helical flow promoters can be seen in Figure 39 (196). Mass transfer coefficients obtained in a bubble column, an ALR, and an ALR with HFP were compared, for water and two solutions of CMC: 0.2% and 0.04%. The corresponding effective viscosities for $\gamma = 50\ s^{-1}$ are 16 and 38 mPa s respectively. It is clearly seen that although with water there is a consistent difference in favor of the bubble column, the differences diminish as the CMC concentration increases, and for the highest concentration tested the difference is negligible. This implies that under these conditions the disadvantage of the HFP in mass transfer vanishes, whereas all the advantages mentioned before subsist.

SUMMARY AND CONCLUSIONS

ALRs are popular in modern bioprocess research and development. These reactors are particularly suitable for biological processes in which a high mass transfer rate is re-

quired, but excessive power input may lead to damage of the cells due to shear effects. ALRs also have very appealing characteristics for bioprocesses for low-value products in which efficiency of energy utilization may become the key point for design. Such is the case for ALRs for wastewater treatment. The ALR is particularly effective for solid fluidization, which is important in many biological processes in which the biocatalyst is available in the form of pellets or is immobilized on a solid support.

The distinctive characteristics of ALRs are conferred by the fluid dynamics of the liquid–gas or liquid–gas–solid mixtures in it. These characteristics are expressed as gas holdup, liquid velocity, and mixing in each of the zones of the ALR. It is important to recognize the differences in the fluid dynamic characteristics of these zones: the riser, the gas separator, and the downcomer. Only a correct understanding of behavior and interconnection of these regions can make possible the correct design of a new reactor or the scale-up of a laboratory device up to pilot or industrial size.

The purpose of scale-up is to conserve and repeat on a larger scale the fluid dynamics of the reactor. Therefore, one of the most important factors in the design and scale-up of reactors is the influence of the geometric characteristics of the system on the flow of the different phases present.

The variables affecting the performance of the reactor are geometric design, operation variables, and fluid properties. Several correlations are available in the technical literature for the prediction of the fluid dynamic characteristics of the reactor (gas holdup, liquid velocity, mixing time of axial dispersion coefficient) and of the mass and heat transfer coefficients. Although attempts have been made to study all the aspects described, no single research group has managed to cover all the variables over a wide range. It is therefore extremely important that the engineer confronting scale-up or de novo design of an ALR analyzes in depth the range of validity of the correlations used for the calculations.

NOMENCLATURE

a	Interfacial area (m^{-1})
A	Cross-sectional area (m^{-2})
B	Heat transfer area (m^2)
Bo	Bond number or Bodenstein number (−)
C	Concentration (mg/L)
Co	Distribution parameter (−)
C	Clearance or fiction losses coefficient (−)
d	Mean diameter (mm)
D	Diameter (m), diffusivity, dispersion coefficient
D_L	Diffusivity, dispersion coefficient (m^2 sec^{-1})
E	Energy (J)
Fr	Froude number (−)
g	Gravitational constant (m sec^{-2})
Ga	Galileo number (−)
h	Height (m)
h_b	Overall heat transfer coefficient ($W\ m^{-2}sec^{-1}K^{-1}$)
I	Degree of homogeneity (−)
J	Superficial velocity (cm sec^{-1})
K	Saturation constant in Monod's equation mgL^{-1}
k_La	Volumetric mass transfer coefficient (hr^{-1})
K_b	Hydraulic coefficient bottom (−)
K_t	Hydraulic coefficient top (−)
L	Length (m)
M	Wetted surface parameter (−)
Ma	Bubble coalescence parameter (−)
Mo	Morton number (−)
n	Flow index (−)
N	Agitation speed (sec^{-1})
Nu	Nusselt number (−)
N_W	Parameter (−)
OTR	oxygen transfer rate
P	Pressure, power input (Pa)
Pe	Peclet number (−)
Pr	Prandtl number (−)
r	Reaction rate ($mol\ m^{-3}sec^{-1}$)
Re	Reynolds number (−)
S	Surface area (m^2)
Sc	Schmidt number (−)
Sh	Sherwood number (−)
t	Time (sec)
U	Velocity (m sec^{-1})
V	Volume (m^3)
Q	Flow rate (m^3sec^{-1})
We	Weber number (−)
Wi	Weissenberg number (−)

Greek Letters

α	Coefficient (−)
β	Flowing volumetric concentration (m^3/m^3)
γ	Shear rate (sec^{-1})
Δ	Difference
φ	Gas holdup (−)
κ	Behavior coefficient (Pa s^{n+2})
λ	Heat conductivity coefficient ($W\ m^{-1}K^{-1}$)
μ	Dynamic viscosity (Pa s), specific growth coefficient (h^{-1})
ρ	Density ($kg\ m^{-3}$)
σ	Surface tension ($N\ m^{-1}$)
τ	Shear stress (Pa)
ν	Kinematic viscosity (m^2 sec^{-1})

Indices

ab	All bubbles
ap	Apparent
b	Bottom
$b\infty$	Terminal
c	Circulation

d	Downcomer, dissipation
D	Dissipation
d	Downcomer
g	Growth
G	Gas
ht	Heat transfer
HM	Heat production by the microorganisms
HS	Heat production by stirring
i	Region of reactor
in	Input
L	Liquid
m	Mixing
max	Maximal
mt	Mass transfer
oc	Oxygen consumption
p	Potential
r	Riser
s	Gas separator
sc	Sugar consumption
sh	Shaft
sol	Solid
t	Top
T	Total
w	Water

BIBLIOGRAPHY

1. J.C. Merchuk, N. Ladwa, A. Cameron, M. Bulmer, and A. Pickett, *AIChE J.* **40**, 1105–1117 (1994).
2. K. Koide, K. Kurematsu, S. Iwamoto, Y. Iwata, and K. Horibe, *J. Chem. Eng. Jpn.* **16**, 413–418 (1983).
3. H. Tanaka, *Proc. Biochem.* **22**, 106–112 (1987).
4. W.R. Kessler, M.K. Popovic, C.W. Robinson, *Can. J. Chem. Eng.* **71**, 195–202 (1993).
5. J.Y. Oldshue, *Fluid Mixing Technology*, McGraw-Hill, New York, 1983.
6. J.C. Merchuk, *Adv. Biochem. Eng.* **44**, 65–95 (1991).
7. R.J. Weetman and J.Y. Oldshue, *Power, Flow and Shear Characteristics of Mixing Impellers, Proceedings of the 6th European Conference on Mixing*, Cranfield, 1988, p. 43.
8. H. Markl, R. Bronnenmeier, B. Wittek, *Int. Chem. Eng.* **31**, 185–197 (1991).
9. JD.a.P. Wase, Y.R., *J. Gen. Microbiol.* **131**, 725–731 (1985).
10. N. Edwards, S. Beeton, A.T. Bull, and J.C. Merchuk, *Appl. Microbiol. Biotechnol.* **30**, 190–196 (1989).
11. J. Tramper, D. Smit, J. Straatman, and J.M. Vlak, *Bioprocess Eng.* **3**, 37–41 (1988).
12. L. Yerushalmi and B. Volesky, *Biotechnol. Bioeng.* **27**, 1297–1305 (1985).
13. F.J. Petersen, V.L. McIntire and E.T. Papoutsakis, *J. Biotechnol.* **7**, 229–246 (1988).
14. D.E. Martens, C.D. de Gooijer, C.A.M. van der Velden-de Groot, E.C. Beuvery, and J. Tramper, *Biotechnol. Bioeng.* **41**, 429 (1993).
15. H.R. Millward, B.J. Bellhouse, A.M. Nicholson, S. Beeton, N. Jenkins, and C.J. Knowles, *Biotechnol. Bioeng.* **43**, 899–906 (1994).
16. H. N.a.U. Shiragami, *Bioprocess Eng.* **10**, 43–45 (1994).
17. N.A. Stathopoulos and J.D. Hellums, *Biotechnol. Bioeng.* **27**, 1021–1026 (1984).
18. H.J. Silva, T. Cortinas, and J.R. Ertola, *J. Chem. Technol. Biotechnol.* **40**, 41–49 (1987).
19. H. Tanaka, H. Semba, T. Jitsufuchi, and H. Harada, *Biotechnol. Lett.* **10**, 485–490 (1988).
20. C.G. Smith, P.F. Greenfield, and D.H. Randerson, *Biotechnol. Tech.* **1**, 39–44 (1987).
21. Z. Zhang, Y. Christi, and M. Moo-Young, *J. Biotechnol.* **43**, 33–40 (1995).
22. J.J. Meijer, H.J.G. ten Hoopen, Y.M. van Gameren, and K.Ch.A.M. Luyben, *Enzyme Microb. Technol.* **16**, 467–477 (1994).
23. G.Z. Lu, B.G. Thompson, and M.R. Gray, *Biotechnol. Bioeng.* **40**, 1277–1281 (1992).
24. H.E. Dunlop, K.N. Pradyumna, and Z.M. Rosenberg, *Chem. Eng. Sci.* **49**, 2263–2276 (1994).
25. L.D.a.L. Landau, E.M., *Fluid Mechanics*, Volume 6, Pergamon, Oxford, U.K., 1975, pp. 115–123.
26. W.S. Tan, G.C., Dai, W. Ye, and J.P. Shen, *Chem. Eng. J. Biochem. Eng. J.* **57**, B31–B36 (1995).
27. H.F. Svendsen, H.A. Jakobsen, and R. Torvik, *Chem. Eng. Sci.* **47**, 1297–1304 (1992).
28. A. Stafford, L. Smith, and M.W. Fowler, *Plant Cell Tissue Organ Culture* **4**, 83–94 (1985).
29. P.J. Wilkinson, in G.W. Moody and P.B. Baker eds., *Biotransformations*, Elsevier, London, 1992.
30. A. Gebauer, T. Scheper, K. Schugerl, *Bioprocess Eng.* **2**, 13–23 (1987).
31. C.M. Chang, W.J. Lu, K.S. Own, and S.J. Hwang, *Biotechnol. Tech.* **7**, 317–322 (1993).
32. M.H. Siegel and J.C. Merchuk, *Biotechnol. Bioeng.* **32**, 1128–1137 (1988).
33. G.A. Legrys, *Chem. Eng. Sci.* **33**, 83–86 (1978).
34. M.E. Orazem and L.E. Erickson, *Biotechnol. Bioeng.* **21**, 69–88 (1979).
35. M.H. Siegel, J.C. Merchuk, and K. Schugerl, *AIChE J.* **32**, 1585–1596 (1986).
36. H. Kubota, Y. Hosono, and K. Fujie, *J. Chem. Eng. Jpn.*, **11**, 319–325 (1978).
37. J.C. Merchuk, N. Ladwa, A. Cameron, M. Bulmer, A. Pickett, and I. Berzin, *Chem. Technol. Biol.* **66**, 174–182 (1996).
38. D. Barnea, Y. Taitel, in N.P. Cheremisinoff, ed. *Encyclopedia of Fluid Mechanics*, Gulf, Houston, Tex., 1986, pp. 403–491.
39. G.B. Wallis, *One Dimensional Two-Phase Flow*, McGraw-Hill, New York, 1969.
40. K. Wiswanathan, *Flow Patterns in Bubble Columns*, Gulf, Houston, Tex., 1986, pp. 291–308.
41. J.C. Merchuk, in K. Schügerl ed. *Biotechnology*, Springer, New York, 1990.
42. A.B. Russell, C.R. Thomas, and M.D. Lilly, *Biotechnol. Bioeng.* **43**, 69–76 (1994).
43. A.C. Gomez, Univ. of Almeria, Spain, 1996.
44. P. Weiland, *Ger. Chem. Eng.* **7**, 374–385 (1984).
45. R.A. Bello, C.W. Robinson, and M. Moo-Young, *Chem. Eng. Sci.* **40**, 53–58 (1985).
46. M.Y. Chisti and M. Moo-Young, *Chem. Eng. Commun.* **60**, 195–242 (1987).

47. K. Koide, S. Iwamoto, Y. Takasaka, S. Matsura, E. Takasashi, M. Kimura, and H. Kubota, *J. Chem. Eng. Jpn.* **17**, 611–618 (1984).
48. G.Q. Li, S.Z. Yang, Z.L. Cai, and J.Y. Chen, *Chem. Eng. J.* **56**, B101–B107 (1995).
49. J. Cai, T.J. Niewstad, and J.H. Kop, *Water Sci. Technol.* **26**, 2481–2484 (1992).
50. M. Popovic and C.W. Robinson, *Proc. 34th Can. Chem. Eng. Conf.* Quebec, Oct., 1984, pp. 258–264.
51. M.Y. Chisti, B. Halard, and M. Moo-Young, *Chem. Eng. Sci.* **43**, 451–457 (1988).
52. G. Vatai and M.N. Tekis, *Effect of Pseudoplasticity on Hydrodynamic Characteristics of Airlift Loop Contactors*, Prague, 1986.
53. K. Akita and M. Kawasaki, *Proc. 48th Meet. of Chem. Engrs. Engrs. Jpn.* Kyoto, 1983, pp. 122–126.
54. Y. Kawase, M. Tsujimura, and T. Yamaguchi, *Bioproc. Eng.* **12**, 21–27 (1995).
55. K. Koide, K. Horibe, H. Kawabata, and I. Shigetaka, *J. Chem. Eng. Jpn.* **18**, 248–254 (1995).
56. K. Koide, M. Kimura, H. Nitta, and H. Kawabata, *J. Chem. Eng. Jpn.* **21**, 393–399 (1988).
57. T. Miyahara, M. Hamaguchi, Y. Sukeda, and T. Takahashi, *Can. J. Chem. Eng.* **64**, 718–725 (1986).
58. M. Chakravarty, H.D. Singh, J.N. Baruah, and M.S. Lyengar, *Indian Chem. Eng.* **16**, 17–22 (1974).
59. Y. Kawase, M. Moo-Young, *Chem. Eng. Commun.* **40**, 67–83 (1986).
60. J.C. Merchuk, M. Bulmer, N. Ladwa, A.M. Pickett, and A. Cameron, in R. King, ed., *Bioreactor Fluid Dynamics* Elsevier, London, 1988, pp. 391–414.
61. K. Koide, K. Horibe, H. Kawabata, and S. Ito, *J. Chem. Eng. Jpn.* **17**, 368–374 (1984).
62. Y. Kawase, M. Moo-Young, *Chem. Eng. Res. Des.* **65**, 121–126 (1987).
63. M. Popovic and C.W. Robinson, *Biotechnol. Bioeng.* **32**, 301–312 (1988).
64. M. Siegel and J.C. Merchuk, *Can. J. Chem Eng.* **69**, 465–471 (1991).
65. M. Hallaile, *Biotechnology*, Ben-Gurion Univ. Beer Sheva, Israel, 1993.
66. M. Ghirardini, G. Donati, and F. Rivetti, *Chem. Eng. Sci.* **47**, 2209–2214 (1992).
67. J.C. Merchuk and R. Yunger, *Chem. Eng. Sci.* **45**, 2973–2976 (1990).
68. Y. Chisti and M. Moo-Young, *Biotechnol. Bioeng.* **34**, 1391–1392 (1989).
69. J.H. Hills, *Ind. Chem. Eng.* **52**, 1–9 (1974).
70. N.N. Clark, C.M. Atkinson, and R.L.C. Flemmer, *AIChE J.* **33**, 515–521 (1987).
71. J.C. Merchuk and Y. Stein, *AIChE J.* **27**, 377–381 (1981).
72. B. Glennon, W. Al-Masry, P.F. MacLoughlin, and D.M. Malone, *Chem. Eng. Commun.* **121**, 181–192 (1993).
73. W. Freedman and J.F. Davidson, *Trans. Inst. Chem. Eng.* **47**, 251–262 (1969).
74. N. Zuber and J.A. Findlay, *Trans. ASME J. Heat Transfer* **87**, 453–468 (1965).
75. S.G. Bankoff, *J. Heat Transfer* **82**, 252–258 (1960).
76. P. Verlaan, J. Tramper, K. van't Riet, and K. Luyben, *Chem. Eng. J.* **33**, B43–B48 (1986).
77. J.C. Merchuk and I. Berzin, *Chem. Eng. Sci.* **50**, 2225–2233 (1995).
78. P.a.O.U. Weiland, *Ger. Chem. Eng.* **4**, 174–181 (1981).
79. P. Verlaan, J. Tramper, K. van't Riet, and K.Ch.A.M. Luyben, *Hydrodynamics and Axial Dispersion in an Airlift Loop Reactor with Two and Three Phase Flow*, Cambridge, UK, 1986.
80. P. Verlaan, A.M.M. Van Eijs, J. Tramper, K. van't Riet, and K. Ch. A.M. Luyben, *Chem. Eng. Sci.* **44**, 1139–1144 (1989).
81. Y. Chisti, *Airlift Bioreactors*, Elsevier, London, 1989.
82. P. Verlaan, J.C. Vos, and K. van't Riet, in R. King ed., *Bioreactor Fluid Dynamics II*, Elsevier, London, 1988.
83. J.C. Merchuk, *Chem. Eng. Sci.* **41**, 11–16 (1986).
84. D. Posarac and D. Petrovic, *Chem. Eng. Sci.* **43**, 116–119 (1991).
85. R.A. Bello, Ph.D. Thesis, Univ. of Waterloo, Ontario, Canada, 1981.
86. Z. Kemblowski, J. Prywarski, and A. Diab, *Chem. Eng. Sci.* **48**, 4023–4035 (1993).
87. D. Posarac, D. Petrovic, A. Dudukovic, and D. Skala, *J. Serb. Chem. Soc.* **56**, 227–240 (1991).
88. M.K. Popovic and C.W. Robinson, *AIChE J.* **35**, 393–405 (1989).
89. M. Nishikawa, H. Kato, and K. Hashimoto, *Ind. Eng. Chem. Process Des. Dev*, **16**, 133–137 (1977).
90. M. Nishikawa, *Biotechnol. Bioeng.* **37**, 691–692 (1991).
91. H.J. Henzler, *Chem.-Ing.-Tech.* **52**, 643–652 (1980).
92. A. Schumpe and W.D. Deckwer, *Bioproc. Eng.* **2**, 79–94 (1987).
93. A. Zaidi, H. Bourziza, and L. Echihabi, *Chem.-Ing.-Tech.* **59**, 748–749 (1987).
94. S.A. El-Tamtamy, S.A., Khalil, A.A. Nour-El-Din, and A. Gaber, *Appl. Microbiol. Biotech.* **19**, 376–381 (1984).
95. H.J. Henzler, J. Kauling, *Scale-Up of Mass Transfer in Highly Viscous Liquids*, Wurzburg, Germany, 1985, pp. 303–312.
96. J.C. Merchuk and S. Ben-Zvi (Yona), *Chem. Eng. Sci.* **47**, 3517–3523 (1992).
97. A. Lubbert, S. Frohlich, B. Larson, and K. Schugerl, *Fluid Dynamics in Airlift Loop Reactors as Measured During Real Cultivation Processes*, London, 1988, pp. 379–394.
98. A. Lubbert, B. Larson, L.W. Wan, and S. Broring, *I. Chem. E. Symp. Ser.* **121**, 203–213 (1990).
99. C.H. Lee, L.A. Glasgow, L.E. Erickson, and S.A. Patel, *Liquid Circulation in Airlift Fermentors*, Miami Beach, 1986.
100. A.G. Jones, *Chem. Eng. Sci.* **40**, 449–462 (1985).
101. D.J. Niklin, *Chem. Eng. Sci.* **17**, 693–702 (1962).
102. M. Clark and A.G. Jones, *Chem. Eng. Sci.* **42**, 378–384 (1987).
103. Y. Chisti and M. Moo-Young, *J. Chem. Technol. Biotechnol.* **42**, 211–219 (1988).
104. S. Wachi, A.G. Jones, and T.P. Elson, *Chem. Eng. Sci.* **46**, 657–661 (1991).
105. E. Garcia Calvo, *Chem. Eng. Sci.* **44**, 321–323 (1989).
106. R.S. Richardson and D.J. Higson, *Trans. Inst. Chem. Eng.* **40**, 169–182 (1962).
107. E. Garcia Calvo and P. Leton, *Chem. Eng. Sci.* **46**, 2947–2951 (1989).
108. H. Blenke, *Adv. Biochem. Eng.* **13**, 121–127 (1979).
109. Y.C. Hsu and M.P. Dudukovic, *Chem. Eng. Sci.* **35**, 135–141 (1980).

110. P. Ayazi Shamlou, D.J. Pollard, A.P. Ison, and M.D. Lilly, *Chem. Eng. Sci.* **49**, 303–312 (1994).
111. W.J. Lu, S.J. Hwang, and C.M. Chang, *Chem. Eng. Sci.* **30**, 1301–1310 (1995).
112. R.A. Bello, C.W. Robinson, and M. Moo-Young, *Biotech. Bioeng.* **27**, 369–381 (1985).
113. J. Zahradnik, M. Fialova, F. Kastanek, K.D. Green, and N.H. Thomas, *Trans. IChE* **73A**, 341–346 (1995).
114. P.R. Fields and N.K.H. Slater, *Chem. Eng. Sci.* **38**, 647–651 (1983).
115. T. Yamane, in J.A. Asenjo and J.C. Merchuk eds., *Bioreactor System Design*, Dekker, New York, 1994.
116. G. B.A.a.M. Buffham and G. Mason, *Chem. Eng. Sci.* **48**, 3879–3887 (1993).
117. O. Levenspiel and K.B. Bishoff, *Adv. Chem. Eng.* **4**, 95–126 (1963).
118. Y. Murakami, T. Hiroshe, S. Ono, and T. Nishijima, *J. Chem. Eng. Jpn.* **15**, 121–126 (1982).
119. C.H. Lin, B.S. Fang, H.Y. Wu, T.F. Fang, T.F. Kuo, and C.Y. Hu, *Biotechnol. Bioeng.* **18**, 1557–1562 (1976).
120. J.C. Merchuk and J.A. Asenjo, *Bioreactor System Design*, Dekker, New York, 1994.
121. K. Schügerl, *Adv. Biochem. Eng.* **19**, 71–94 (1981).
122. Moor Nagar, Graduate Thesis, Ben Gurion University, Dept. of Chemical Engineering, 1992.
123. M.H. Siegel, M. Hallaile, and J.C. Merchuk, in A. Mizrahi ed., *Advances in Biotechnological Processes*, Liss, New York, 1988, pp. 79–124.
124. W.J. Lu, S.J. Hwang, and C.M. Chang, *Ind. Eng. Chem. Res.* **33**, 2180–2186 (1994).
125. M. Siegel and J.C. Merchuk, *Influence of the Gas-Liquid Separator on Air-Lift Reactor Design and Operation*, 1990.
126. S.P. Godboole and Y.T. Shah, *Design and Operation of Bubble Column Reactors*, Gulf, Houston, Tex., 1986.
127. W.D. Deckwer, R. Burckhart, and G. Zoll, *Chem. Eng. Sci.* **29**, 2177–2187 (1974).
128. J.B. Joshi, *Trans Inst. Chem. Eng.* **58**, 155–161 (1980).
129. P.R. Fields and N.K.H. Slater, *Biotechnol. Bioeng.* **26**, 719–723 (1984).
130. M.R. Makley and A. Keller, *Phylos. Trans. Roy. Soc. London*, **278**, 1276–1281 (1975).
131. S. Frehlich, M. Lotz, T. Korte, A. Lübbert, K. Schügerl, and M. Seekamp, *Biotech. Bioeng.* **38**, 43–48 (1991).
132. M. Nakanoh, F. Yoshida, *Ind. Eng. Chem. Process Des. Dev.* **19**, 190–195 (1980).
133. J.C. Merchuk, S. Yona, M.H. Siegel, and A. Ben-Zvi, *Biotech. Bioeng.* **35**, 1161–1163 (1990).
134. S.P. Godbole, A. Schumpe, Y.T. Shah, and N.L. Carr, *AIChE J.* **30**, 213–220 (1984).
135. R.B. Bird, *Chem. Eng.* **27**, 102–109 (1993).
136. E. Molina-Grima, Y. Chisty, and M. Moo-Young, *J. Biotechnol.* **52**, 195–201 (1997).
137. V. Linek, V. Vacek, and P. Benes, *Chem. Eng. J.* **34**, 11–34 (1987).
138. T. Philichi, L. Stenstrom, K. Michael, *J. Water Pollut. Control Fed.* **4**, 83–86 (1989).
139. C.F. Mignone and J. Ertola Rodolfo, *J. Chem. Technol. Biotechnol.* **34**, 121–126 (1984).
140. Y. Kawase and M. Moo-Young, *Biotechnol. Bioeng.* **30**, 345–347 (1987).
141. J.C. Merchuk, S. Ben-Zvi, and K. Niranjan, *TibTech* **12**, 501–511 (1994).
142. W.D. Deckwer, *Bubble Column Reactors*, Wiley, New York, 1991.
143. M. Siegel, S. Ben-Zvi, and J.C. Merchuk, *Measurement and Interpretation of Mass Transfer Data in Air-Lift Bioreactors*, Kuyngju, Korea, 1990, pp. 449–453.
144. G. Andre, C.W. Robinson, and M. Moo-Young, *Chem. Eng. Sci.* **38**, 1845–1854 (1983).
145. J.C. Merchuk, G. Osemberg, M. Siegel, and M. Shacham, *Chem. Eng. Sci.* **47**, 2221–2226 (1992).
146. S.-J. Hwang, W.-J. Lu, *Chem. Eng. Sci.* **52**, 853–857 (1997).
147. J.C. Merchuk, *TibTech*, **8**, 66–71 (1990).
148. M.E. Orazem and L.E. Erickson, *Biotechnol. Bioeng.* **21**, 69–88 (1979).
149. M. Siegel, M. Hallaile, and J.C. Merchuk, in A. Mizrahi ed. *Advances in Biotechnological Processes*, Liss, New York, 1988, pp. 51–75.
150. T.W. Barker and J.T. Worgan, *Eur. J. Appl. Microbiol. Biotechnol.* **13**, 77–82 (1981).
151. K. Koide, K. Horibe, H. Kawabata, and S. Ito, *J. Chem. Eng. Jpn.* **18**, 248–254 (1985).
152. Y. Chisti, K. Fujimoto, and M. Moo-Young, *Biotechnol. Prog.* **5**, 72–76 (1988).
153. I.S. Suh., A. Schumpe, and W.D. Deckwer, *Can. J. Chem. Eng.* **69**, 506–512 (1991).
154. M.Y. Chisti and M. Moo-Young, *Biotechnol. Bioeng.* **31**, 487–494 (1988).
155. Y.T. Shah, B.G. Kelkar, S.P. Godbole, and W.D. Deckwer, *AIChE J.* **38**, 353–359 (1982).
156. V.G. Levich, *Physicochemical Hydrodynamics*, Prentice Hall, Englewood Cliffs, N.J., 1962.
157. A.N. Kolmogoroff, *Dokl. Akad. Nauk SSSR* **66**, 1949.
158. N. Blakebrough, I.A. Fatile, W.J. McManamey, and G. Walker, *Chem. Eng. Res. Des.* **61**, 383–387 (1983).
159. K. Koide, M. Terasawa, and H. Takekawa, *J. Chem. Eng. Jpn.* **19**, 341–344 (1986).
160. R.S. Douek, G.F. Hewitt, and A.G. Livingston, *Trans. Inst. Chem. Eng.* **73A**, 336–340 (1995).
161. A.G. Livingstone and S.F. Zhang, *Chem. Eng. Sci.* **48**, 1641–1654 (1993).
162. J.C. Merchuk and M. Herskowitz, *Can. J. Chem. Eng.* **64**, 57–62 (1986).
163. L.S. Fan, S.J. Hwang, and A. Matsuura, *Chem. Eng. Sci.* **39**, 1677–1681 (1984).
164. D. Petrovic, D. Posarac, and D. Skala, *Chem. Eng. Sci.* **44**, 997–998 (1989).
165. D. Petrovic, D. Posarac, A. Dudukovic, and D. Skala, *Chem. Eng. Sci.* **45**, 2967–2970 (1990).
166. M. Immich and U. Onken, *Chem. Eng. Sci.* **47**, 3379–3386 (1992).
167. L.S. Fan, K. Fujie, and T.R. Long, *AIChE Symp. Ser.* **80C24T**, 102–107 (1984).
168. K. Koide, K. Horibe, H. Kawabata, and S. Ito, *J. Chem. Eng. Jpn.* **18**, 248–254 (1985).
169. K. Koide, K. Shibata, H. Ito, S.Y. Kim, and K. Ohtaguchi, *J. Chem. Eng. Jpn.* **25**, 11–16 (1992).
170. M.D. Gluz and J.C. Merchuk, *Chem. Eng. Sci.* **51**, 1920–1925 (1996).
171. K.B. Kushalkar and V.G. Pangarkar, *Chem. Eng. Sci.* **49**, 139–144 (1994).

172. G. Rossi, *Biohidrometallurgy*, McGraw-Hill, Hamburg, Germany, 1990.
173. B.C. Smith and D.R. Skidmore, *Biotechnol. Bioeng.* **35**, 483–491 (1990).
174. A. Assa and R. Bar, *Biotechnol. Bioeng.* **38**, 1325–1330 (1991).
175. J. Schmidt, R. Nassar, and A. Lubbert, *Chem. Eng. Sci.* **47**, 2295–2300 (1992).
176. J.B. Snape and N.H. Thomas, *Biotechnol. Bioeng.* **40**, 337–345 (1992).
177. N.M.G. Oosterhuis, Ph.D. Thesis, Dept. of Chemical Technology, Univ. of Technology, Delft, The Netherlands, 1984.
178. N.W.F. Kossen and N.M.G. Oosterhuis, in N.J. Rhem and J. Reed eds. *Biotechnology* VCH, Weinheim, Germany, 1985.
179. A.P.J. Sweere, K.C.A.M. Luyben, and N.W.F. Kossen, *Enzyme Microb. Technol.* **9**, 386–392 (1987).
180. C. Sola and F. Godia, in J.A. Asenjo, J.C. Merchuk eds., *Bioreactor System Design*, Dekker, New York, 1994.
181. F.H. Johnson, H. Eyring, and M.J. Polissar, *The Kinetic Basis of Molecular Biology*, Wiley, New York, 1954.
182. D.G. Jordan, *Chemical Process Development*, Volume 6, Wiley, New York, 1968.
183. H. Taguchi, T. Imanaka, S. Teramoto, M. Takatsu, and M. Sato, *J. Ferment. Technol.* **46**, 823–830 (1986).
184. H. Takei, K. Misusawa, and F. Yoshida, *Ferment. Technol.* **53**, 151–156 (1975).
185. J.A. Roels, *Energetics and Kinetics in Biotechnology*, Elsevier, Amsterdam, 1983.
186. B. McNeil and B. Kristiansen, *Biotechnol. Lett.* **9**, 101–107 (1987).
187. P.M. Doran, *Adv. Biochem. Eng. Biotechnol.* **48**, 115–168 (1993).
188. D.J. Pollard, A.P. Ison, P. Ayazi Shamlou, and M.D. Lilly, *The Examination of Bioreactor Heterogeneity with Rheological Different Fermentation Broths*, Kluwer, Dordrecht, The Netherlands, 1994, pp. 163–170.
189. F. Potucek, *Collect. Czech. Chem. Commun.* **55**, 981–986 (1990).
190. P.D. Gaspillo and S. Goto, *J. Chem. Eng. Jpn.* **24**, 680–682 (1991).
191. Y. Chisti, M. Kasper, and M. Moo-Young, *Can. J. Chem. Eng.* **68**, 45 (1990).
192. S.Q. Zhou, L.M. Tang, and K. Schugerl, *J. Biotechnol.* **28**, 165–177 (1993).
193. Y. Bando, M. Nishimura, H. Sota, S. Suzuki, and N. Kawase, *Chem. Eng. Sci.* **47**, 3371–3378 (1992).
194. Y. Bando, M. Nishimura, A. Hayaashi, S. Hiura, S. Indo, and A. Idota, *J. Chem. Eng. Jpn.* **28**, 225–227 (1995).
195. T.K. Ghosh, B.R. Maiti, and B.C. Bhattacharyya, *Biotechnol. Tech.* **7**, 301–307 (1993).
196. C. Schlotelburg, M. Gluz, M. Popovic, and J.C. Merchuk, *Int. Symp. on Bubble Columns*, Kyoto, Nov. 1997.

See also BIOREACTORS, CONTINUOUS STIRRED-TANK REACTORS; BIOREACTORS, FLUIDIZED-BED; BIOREMEDIATION; ENZYMES, IMMOBILIZATION METHODS; FERMENTATION MONITORING, DESIGN AND OPTIMIZATION; MASS TRANSFER; RHEOLOGY OF FILAMENTOUS MICROORGANISMS, SUBMERGED CULTURE; SCALE-UP, STIRRED TANK REACTORS.

BIOREACTORS, CONTINUOUS STIRRED-TANK REACTORS

MICHAEL E. LASKA
CHARLES L. COONEY
Massachusetts Institute of Technology
Cambridge, Massachusetts

KEY WORDS

Chemostat
Continuous culture
Continuous stirred-tank reactor (CSTR)

OUTLINE

INTRODUCTION

Definitions

A continuous stirred-tank reactor (CSTR) is defined as an agitated vessel with continuous addition and removal of material and energy. The CSTR is one of the basic continuous reactor types widely used in the chemical process industries because of its amenability to process control and scale-up, although in biotechnology applications the CSTR is used more often as a research tool than as a production technology. An idealized, well-mixed CSTR can be modeled as having no spatial variations in temperature, concentration, fluid properties, or reaction rates. Thus, the properties of the exit stream may be considered the same as those throughout the vessel. Although such ideal mixing is never observed, the vessel is designed to provide good mixing through selection of operating conditions and vessel, baffle, and impeller geometries. The stirred tank used in continuous bioprocesses is similar to that used in batch bioprocesses, with the exception that the CSTR likely has an overflow or other level control device. Oxygen can be introduced into the vessel by sparging through inlets at the base of the vessel, where impellers then disperse the bubbles. Vessel jacketing or internal cooling coils provide a means for heat transfer. Continuous systems that are not agitated vessels often are modeled as CSTRs when their behavior

approximates that of the ideal CSTR. CSTRs are also known as backmix reactors, continuous-flow stirred-tank reactors (CFSTRs), or chemostats, when used for cell growth.

Strategy for CSTR Analysis

The performance and analysis of a CSTR is based on the material and energy conservation balances and the underlying processes governing the reaction kinetics. Because of the large variety of CSTR applications and limited space for discussion in this article, a brief outline of the steps in systems analysis will be beneficial in understanding any CSTR-based process.

The strategy for analyzing CSTR performance first requires defining the problem statement and goals. The second step is system identification, which includes defining the system boundaries and the interactions between the system and its environment across the system boundaries. The system could be a cell, the fluid in the reactor, or an entire bioprocessing plant. The third step is to identify the state variables that characterize the system. During the course of the analysis, new state variables may be identified and added to the original list. The fourth step is to characterize the state of the system using material and energy balances that account for the accumulation of mass and energy. In a general balance for a particular quantity, the rate of accumulation of that quantity in the system is equal to the net influx of the quantity across the system boundaries plus the rate at which the quantity is generated. Separate material balances are written for the reactants, products, and the catalyst (e.g., cells or enzymes). The final step is to calculate performance metrics and revisit the assumptions to determine the conditions under which they are valid.

REACTION KINETICS

This section discusses the theory governing ideal CSTR performance in two important bioprocess applications (cell growth and enzyme reactions), the underlying assumptions of the theory, and some practical aspects associated with CSTR operation.

Cell Growth

Introduction. Cell growth occurs in response to the environment. It is useful to classify growth in four mechanistic categories and identify those features relevant to the analysis of CSTRs. These classes are fission, budding, mycelial, and viral growth. The modes of growth serve to highlight the impact of morphology on bioprocessing considerations.

Bacteria grow by a process of binary fission yielding two identical daughter cells with doubling times typically between 0.5 and 3.0 h. The high specific growth rates of bacteria make them especially suitable for many CSTR applications; however, high O_2 requirements and metabolic heat generation become important considerations in bioprocess scale-up. Animal cells (5 to 20 μm) are much larger than bacteria ($\sim$1 μm) and yeast (5 to 10 μm) cells but grow more slowly, with typical doubling times between 18 and 48 h. Because of their larger size and lack of a protective cell wall, mammalian cells are particularly sensitive to the fluid shear in the vessel as well as to the osmotic pressure of the medium. Mammalian cells grow over a narrow range of osmotic/pressures and pH and typically have more complex nutritional requirements than bacteria and yeasts do.

Cell division in yeast occurs primarily by budding, with typical doubling times between 1 and 3 h. The budding process leads to a mother and a daughter cell, each having different growth rates and cell surface characteristics. Unlike bacterial cultures, yeast cell populations have a broad and time-varying distribution of ages and properties that may influence the formation of a desired product.

Mycelial growth occurs in molds, actinomycetes, and some yeasts by a process of hypha chain elongation and branching. Mycelial cultures also are characterized by a distribution of ages, with younger cells located at the hyphal tips. The hyphae form intertwined cellular strands, or mycelia, that increase the broth viscosity and lead to nonideal fluid mixing. High broth viscosity can be problematic for process monitoring and control, cell separation and recycle, and oxygen and heat transfer. Furthermore, fluid shear in the vessel can cause hyphal breakage and the formation of denser, more highly branched pellets or flocs. Cells in these pellets may be exposed to different microenvironments because of mass transfer limitations, which can vary with culture conditions and influence important cell properties.

Viral growth initially requires the infection of a host cell, which occurs by attachment of the virus to the cell surface and injection of viral nucleic acids into the cell interior. New viruses are constructed from biological molecules synthesized by the host cell under the direction of the viral genome. Viral nucleic acid is replicated many times (e.g., $>$500) and encapsulated in coat proteins to form a large number of new viral particles. Viral growth can proceed to one of two phases. In the lytic cycle, the host cell will lyse or break open and release infectious viral particles, whereas in the lysogenic cycle, the viral DNA will be integrated into the host cell DNA and the host cell will continue to reproduce normally.

CSTRs used for cell growth are commonly referred to as chemostats or turbidostats, depending on the strategy used to control the vessel environment. The most common arrangement is the chemostat (1–3), in which the medium fed to the vessel is designed so that all but a single nutrient essential for growth are present in excess of the cells' requirements. Any nutrient necessary for growth can be used to control the size of the cell population in the vessel, making the chemostat a flexible tool to study cellular behavior under different nutrient limitations. In a turbidostat, the cell concentration in the vessel is maintained constant by monitoring the optical density of the culture and modulating the medium feed rate to achieve a set point optical density. When the optical density rises above the set point, the feed rate is increased and, because the fluid volume is maintained constant by an overflow device, the well-mixed culture is diluted and the optical density approaches the set point value. The turbidostat is less commonly used because of difficulties in continuously monitoring the cell con-

centration. Its main utility is to control the growth rate near the maximum growth rate, an operating region in which the chemostat is less stable.

Material Balances. Cell mass is most often used to quantify microbial growth and usually is proportional to cell number under conditions of balanced growth, in which cellular chemical and physical properties are preserved in subsequent generations. Material balances based on cell number may have particular utility for some applications, such as mammalian cell culture, where number rather than mass is the conventional method of analysis. One may need to be wary of variations in cell size and morphology that may not be apparent from measurements of cell mass or number. The material balance for a uniform cell population in a CSTR can be written as shown in equation 1, in which $\mu[h^{-1}]$ represents the specific growth rate and $\alpha[h^{-1}]$ represents the specific rate of cell lysis and/or endogenous metabolism (i.e., resulting in a decrease in cell mass). The specific growth and death rates differ among organisms and are functions of the cell environment (e.g., pH, temperature, nutrients).

$$\underbrace{F \cdot X_0}_{\text{CELLS IN}} - \underbrace{F \cdot X}_{\text{CELLS OUT}} + \underbrace{V \cdot \mu \cdot X}_{\text{CELL GROWTH}} - \underbrace{V \cdot \alpha X}_{\text{CELL LYSIS}} = \underbrace{\frac{d(V \cdot X)}{dt}}_{\text{ACCUMULATION}} \tag{1}$$

It may be necessary to reformulate the cell balance if the cell population is significantly differentiated; examples include mixed cultures, mammalian cell culture, and recombinant fermentations with plasmid instability (see "Selection/Mutation and Contamination"). If we assume the system is operating at steady state and there is no accumulation of fluid or cells in the vessel, then the time derivative can be set to zero. Under normal bioprocessing conditions, cell death is assumed to be negligible (i.e., $\mu \gg \alpha$). Bacterial and yeast cells maintain approximately complete viability except in suboptimal environments and at very low dilution rates, whereas in mammalian cell culture, viability and cell lysis may vary significantly with time.

Assuming that the feed to the reactor is sterile ($X_0 = O$), the material balance reduces to equation 2, in which the specific growth rate is equal to the liquid flow rate from the vessel divided by the liquid volume. This quantity is called the dilution rate D (the inverse residence time for the CSTR) and has units of h^{-1}.

$$\mu = \frac{F}{V} = D \tag{2}$$

Equation 2 illustrates one of the most important attributes of the chemostat: that the specific growth rate can be controlled by manipulating the dilution rate. Control of the specific growth rate, combined with the ability to maintain a constant, defined cell environment, makes the chemostat a powerful experimental tool with which to investigate the many factors that influence cell growth, metabolism, and product formation. It is as an investigative tool, rather than as a production technology, that continuous culture has found its widest and most successful application.

A second material balance can be written for the growth-limiting substrate (S) using an *allocation model* for substrate utilization, in which substrate uptake is divided into cell growth, cell maintenance, and product formation components.

$$\underbrace{F \cdot S_0}_{\text{SUBSTRATE IN}} - \underbrace{F \cdot S}_{\text{SUBSTRATE OUT}} - V \cdot \left(\underbrace{\frac{\mu \cdot X}{Y_{X/S}}}_{\text{GROWTH}} + \underbrace{m \cdot X}_{\text{MAINTENANCE}} + \underbrace{\frac{q_P \cdot X}{Y_{P/S}}}_{\text{PRODUCT FORMATION}} \right) = \underbrace{\frac{d(V \cdot S)}{dt}}_{\text{ACCUMULATION}} \tag{3}$$

In this equation, $Y_{X/S}$ is the cell yield (or dry cell weight, DCW) on substrate (g DCW/g substrate), m is the cell maintenance coefficient (g substrate/g DCW $\times$ h), q_P is the specific product formation rate (g product/g DCW $\times$ h), and $Y_{P/S}$ is the mass yield of product on substrate (g product/g substrate). Each specific yield coefficient describes the allocation of substrate to cells, product, or maintenance. Depending on whether the fermentation goal is to produce cells (biomass) or a metabolic product, it may be possible to simplify the substrate balance by assuming that some uptake terms are dominant. The following sections describe the analysis of biomass production and product formation.

Biomass Production. In the case of biomass production, in which the goal of the fermentation is to produce cells, the large majority of nutrient uptake goes toward cell growth. The rate of substrate uptake for growth is assumed to be much greater than that for maintenance (i.e., $\mu/Y_{X/S} \gg m$) and product formation (i.e., $\mu/Y_{X/S} \gg q_P/Y_{P/S}$). Assuming the system is at steady state, the substrate balance (equation 3) can be rewritten and solved for the cell concentration using equation 2.

$$X = Y_{X/S} \cdot (S_0 - S) \tag{4}$$

To determine the relationship between the specific growth rate and the cell environment, a suitable growth model must be adopted. The simplest and most common relationship used is the unstructured Monod growth model, in which cell growth is a function of a single limiting substrate, usually the carbon source. Alternative unstructured growth models are given in Table 1. These unstructured growth models are empirically derived from observations of chemostat behavior, and their applicability to dynamic-batch or fed-batch processes should not be assumed. Some unstructured models (e.g., substrate inhibition) can have more than one solution, making different steady states possible depending on the starting conditions. Structured models, which typically account for either changes in cell composition, intracellular concentra-

Table 1. Common Unstructured Growth Models

Monod
$\mu = \dfrac{\mu_{max} \cdot S}{K_S + S}$
Modified monod
$\mu = \dfrac{\mu_{max} \cdot S^{\lambda}}{K_S + S^{\lambda}}$
where λ is an adjustable parameter

Inhibition models	Noncompetitive	Competitive
Substrate inhibition	$\mu = \dfrac{\mu_{max} \cdot S}{(K_S + S) \cdot \left(1 + \dfrac{S}{K_I}\right)}$	$\mu = \dfrac{\mu_{max} \cdot S}{K_S + S + \left(\dfrac{K_S}{K_I}\right) \cdot S}$
Product inhibition	$\mu = \dfrac{\mu_{max} \cdot S}{(K_S + S) \cdot \left(1 + \dfrac{P}{K_P}\right)}$	$\mu = \dfrac{\mu_{max} \cdot S}{K_S + S + \left(\dfrac{K_S}{K_P}\right) \cdot P}$

tions (4), or cell morphology (5), have been proposed in varying degrees of complexity (6,7). Structured models can have utility when such properties significantly influence the kinetics and are required to accurately describe behavior (such as in dynamic process modeling).

$$\mu = \frac{\mu_{max} \cdot S}{K_S + S} = D \tag{5}$$

In the Monod model, μ_{max} is the maximum specific growth rate of the organism, and K_S, called the saturation constant, is inversely proportional to the cell's affinity for the substrate. The value of K_S is typically quite low (1 to 5 mg/L for *Escherichia coli* on glucose), which means that $\mu \approx \mu_{max}$ when $S > 10K_S$ and that μ only becomes a strong function of the substrate concentration when $S < 10K_S$.

Note that the maximum specific growth rate of the organism limits the extent to which the dilution rate can be increased.

$$\mu_{max} = D_C \tag{6}$$

This threshold dilution rate is called the critical dilution rate (D_C), and increasing the dilution rate beyond D_C results in "washout"; cells are removed from the vessel at a rate faster than their growth rate. This can limit the productivity of the simple CSTR and motivates cell retention or recycle strategies that enable operation at higher throughputs.

Rearranging the growth model to solve for the substrate concentration in the vessel gives

$$S = \frac{D \cdot K_S}{\mu_{max} - D} \tag{7}$$

At dilution rates above D_C, cells have been washed out of the vessel, and the substrate concentration equals the inlet concentration (S_O). By substituting equation 7 into equation 4, the cell concentration becomes an explicit function of D and S_O.

$$X = Y_{X/S} \cdot \left(S_0 - \frac{D \cdot K_S}{\mu_{max} - D}\right) \tag{8}$$

The biomass productivity of the CSTR (R_{CSTR}), defined as the cell output per reactor volume, is calculated as

$$R_{CSTR} = \frac{F \cdot X}{V} = (D \cdot X) = D \cdot Y_{X/S} \cdot \left(S_0 - \frac{D \cdot K_S}{\mu_{max} - D}\right) \tag{9}$$

Figure 1 shows the steady-state cell and substrate concentrations and biomass productivity for the ideal chemostat. Note that the substrate is almost completely utilized

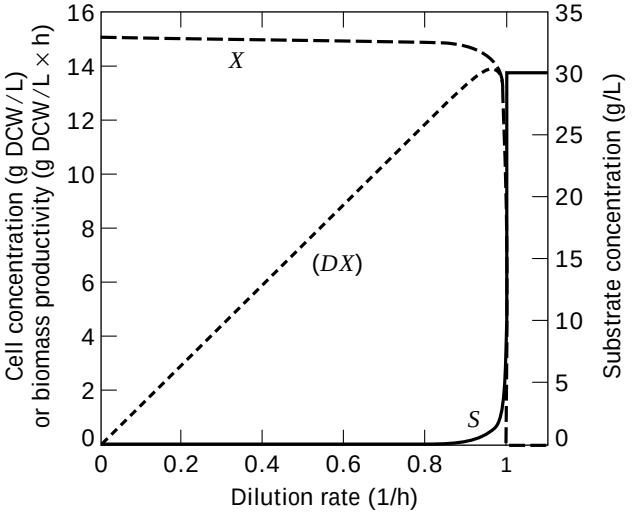


Figure 1. Ideal chemostat performance: $\mu_{max} = 1.0\ h^{-1}$, $K_S = 0.05$ g/L, $S_0 = 30$ g/L, $Y_{X/S} = 0.5$ g DCW/g substrate.

over most of the operating dilution rates. This high conversion of substrate is a key attribute of the ideal CSTR system, improving the economics of processes that depend on efficient substrate utilization and minimizing the effects of substrate inhibition. Also shown in Figure 1 is the high concentration of cells in the vessel until washout at the critical dilution rate. The maximum biomass productivity is located close to the washout point, making operation at the maximum productivity point very sensitive to deviations in dilution rate.

The dilution rate associated with maximum biomass productivity for a fixed limiting nutrient concentration in the feed (D_M) can be calculated by setting the derivative of (DX) with respect to D to zero and solving for the dilution rate. D_M is then given as

$$D_M = \mu_{max}\left(1 - \sqrt{\frac{K_S}{K_S + S_0}}\right) \tag{10}$$

By substituting D_M into equation 8, we can solve for X_M, the cell concentration corresponding to D_M.

$$X_M = Y_{X/S}\left[S_0 + K_S - \sqrt{K_S(S_0 + K_S)}\right] \tag{11}$$

Considering the limiting nutrient concentration (S_0) as an independent design variable, equation 11 suggests that the substrate concentration in the feed can be increased arbitrarily to achieve extraordinary cell densities and productivities. In reality, the productivity of an aerobic reactor system is ultimately limited by the rates of heat and/or mass transfer when the reaction kinetics are fast (i.e., high X and D) (see "Mass Transfer"). Oxygen transfer, and not the carbon substrate, is often growth limiting because oxygen is an essential nutrient for aerobic metabolism. It is poorly soluble in the medium (typically around 7 mg/L for air at 1 atm), and its transfer rate is restricted by the physical capabilities of the oxygenation system. Oxygen-limited growth may be expressed as a steady-state balance between the oxygen uptake rate (OUR) and the oxygen transfer rate (OTR). Oxygen transfer is usually limited by transfer from the gas to liquid phases, leading to the steady-state balance:

$$\underbrace{\left(\frac{\mu \cdot X}{Y_{X/O_2}} + m_{O_2} \cdot X\right)}_{\text{OUR}} = \underbrace{k_L a(C^* - C_L)}_{\text{OTR}} \tag{12}$$

in which Y_{X/O_2} is the cell yield on oxygen (g DCW/g O_2), m_{O_2} is the maintenance coefficient for oxygen (g O_2/g DCW $\times$ h), k_L is the liquid phase mass transfer coefficient (cm/h) a is the specific interfacial area for mass transfer (cm^2/cm^3), $k_L a$ is the mass transfer coefficient (h^{-1}), and ($C^* - C_L$) is the driving force for mass transfer where C^* is the equilibrium oxygen concentration (mmol/L) and C_L is the dissolved oxygen concentration (mmol/L). Typical values for Y_{X/O_2} are located in Table 2. The ability of the heat transfer system to remove heat generated during microbial growth can also limit R_{CSTR}, as discussed in "Energy Balance". These limitations will be important in evaluating and comparing CSTR performance.

Table 2. Yield Coefficients for Bacteria on Different Carbon Substrates

Substrate (i)	$Y_{X/S}$ [g DCW/g substrate]	Y_{X/O_2} [g DCW/g O_2]	Y_{HI} [g DCW/kcal]
Acetate	0.36	0.70	0.21
Glucose	0.51	1.5	0.42
Methanol	0.40	0.44	0.12
Ethanol	0.68	0.61	0.18
n-Paraffins	1.0	0.50	0.16
Methane	0.62	0.20	0.061

Source: From Ref. 16.

The biomass productivity of the CSTR can be compared to the productivity of a batch fermentor (R_{BATCH}) by defining a relevant batch productivity. The batch fermentation cycle consists of a lag phase, an exponential growth phase, cell harvest, and a batch turnaround time associated with cleaning, sterilizing, and filling the vessel. The lag, harvest, and turnaround activities can be grouped into a term $t_{TURNAROUND}$ in order to determine the batch cycle time (t_{CYCLE}) using equation 13, in which X_i is the concentration of cells in the vessel following inoculation (typically $X_i \sim 0.1 \times X$).

$$t_{CYCLE} = \frac{1}{\mu_{max}} \ln\left(\frac{X}{X_i}\right) + t_{TURNAROUND} \tag{13}$$

Cell growth can be calculated from the cell yield on the growth-limiting substrate and the initial concentration of substrate, again assuming cell maintenance and product formation are negligible.

$$X - X_i = Y_{Y/S} S_0 \tag{14}$$

Subsequently, the ratio of biomass productivities in the CSTR at D_M, $(R_{CSTR})_M$, and in the batch fermentor is given by equation 15, in which it is assumed that $S_0 \gg K_S$ (as it often is).

$$\frac{(R_{CSTR})_M}{R_{BATCH}} = \frac{(D_M \cdot X_M)}{\left(\frac{X - X_i}{t_{CYCLE}}\right)} = \ln\left(\frac{X}{X_i}\right) + \mu_{max} \cdot t_{TURNAROUD} \tag{15}$$

Equation 15 often appears as a measure of relative biomass productivity, in which the CSTR is favored over batch operation at high growth rates and long turnaround times. However, equations 10 and 11, associated with maximum productivity, do not reflect the ultimate limitation posed by heat and mass transfer in industrial aerobic processes. Given that productivity has a limit dictated by the system, the independent parameters, D and S_0, can be adjusted so that the maximum productivity is attained during operation. Recognizing that (R_{CSTR}) is ultimately limited by the maximum oxygen transfer rate, equation 12 can be rearranged to give the biomass productivity.

$$(D \cdot X) = \frac{D \cdot Y_{X/O_2} \cdot OTR_{MAX}}{(D + m_{O_2} \cdot Y_{X/O_2})} \tag{16}$$

Notice that the cell concentration is fixed once an operating

dilution rate is specified. Solving equation 16 for X and substituting into equation 4 yields

$$S_0 = \frac{Y_{X/O_2} \cdot OTR_{MAX}}{Y_{X/S} \cdot \delta_S \cdot (D + m_{O_2} \cdot Y_{X/O_2})} \quad (17)$$

where δ_S is the fractional substrate conversion. High substrate conversion leads to lower raw materials costs and reduces the burden of residual substrate on downstream purification and waste treatment operations. When operating near the maximum reactor productivity (i.e., maximum OTR), the dilution rate and the inlet substrate concentration can be adjusted independently to achieve target levels of cell concentration and/or substrate conversion, as shown in Figure 2.

For biomass production when O_2 transfer is not limiting productivity, the CSTR is favored over batch operation when the specific growth rate is high and the batch turnaround time is long. However, most industrial processes will be operating at or near the oxygen or heat transfer limitation. The desirability of enhanced O_2 transfer has motivated the development of novel bioreactor designs (e.g., bubble columns, loop or airlift reactors). For therapeutic products, however, the overwhelming majority of fermentations are batch or fed-batch processes. In these applications, the choice of operating mode is not based on biomass productivity. Performance metrics such as volumetric productivity of the product, product yield, and product concentration become more important in evaluating potential operating strategies than simply the biomass concentration.

Product Formation. When the goal of the fermentation is to produce a product other than biomass, the criteria used to evaluate alternative operational modes are less straightforward than biomass productivity. In order to evaluate process alternatives, a proper set of performance metrics must be identified that relate to overall process economics. When process economics are dominated by fermentation costs (e.g., fuel alcohol or gluconic acid production), volumetric productivity and conversion yields are especially relevant criteria as they relate to the size and cost of the reactor system and the cost of raw materials. In processes where recovery costs dominate (e.g., antibiotics), however, the size and operational costs of the recovery system are proportional to the fluid volume processed and inversely proportional to the product concentration (8). As a result, the final product concentration, or titer, is more important than biomass productivity.

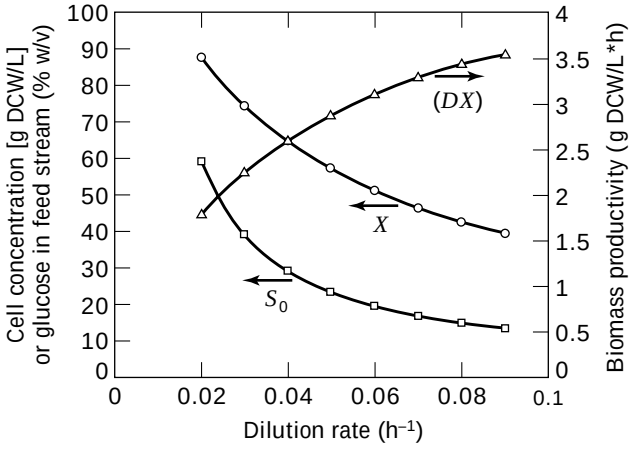


Figure 2. Chemostat operating at maximum oxygen transfer rate: OTR^{max} = 100 mmol O_2/L × h, μ_{max} = 0.09 h^{-1}, Y_{X/O_2} = 1.56 g DCW/g O_2, m_{O_2} = 0.024 g O_2/g DCW × h, $Y_{X/S}$ = 0.45 g DCW/g substrate, 95% substrate conversion.

The material balance on the product can be written as shown below, in which product formation is expressed using a specific product formation rate q_P (g product/g DCW × h), and product degradation by a specific rate constant $k_P(h^{-1})$.

$$\underbrace{q_P \cdot X}_{\text{PRODUCT FORMATION}} - \underbrace{k_P \cdot P \cdot D}_{\text{PRODUCT DEGRADATION}} - \underbrace{D \cdot P}_{\text{PRODUCT OUT}} = \underbrace{\frac{dP}{dt}}_{\text{PRODUCT ACCUMULATION}} \quad (18)$$

Product formation can be characterized in relation to growth, being growth (primary metabolites) or nongrowth (secondary metabolites) associated. Examples of growth-associated products are direct catabolic products of the carbon substrate, such as ethanol and citric or acetic acid. Nongrowth-associated products, comprising many antibiotics, are metabolites that are not necessary for cell growth and typically are only produced during slow or stationary growth phases. Some products, such as xanthan gum and lactic acid, are mixed growth associated in that they are produced during slow and stationary growth phases. Numerous models of product formation have been proposed, taking into account variables such as hyphal morphology, cell age, surface area, metabolic carbon flux, and plasmid copy number. A simple model expresses the growth dependence of q_P as (9)

$$q_P = \underbrace{\alpha \cdot \mu}_{\text{GROWTH ASSOCIATED}} + \underbrace{\beta}_{\text{NONGROWTH ASSOCIATED}} \quad (19)$$

The steady-state product balance can be rewritten and solved for P.

$$P = \frac{q_P \cdot X}{D} = \frac{(\alpha \cdot \mu + \beta)}{D} \cdot X = \left(\alpha + \frac{\beta}{D}\right) \cdot X \quad (20)$$

For growth-associated products (i.e., $\alpha \gg \beta$) product concentration is proportional to biomass and is independent of the dilution rate when X is approximately constant. Increasing the dilution rate results in increased product formation up to the region near D_C. The growth dependence of product formation and intracellular metabolic fluxes can be determined using a chemostat (10). An example of growth-associated product formation is shown in Figure 3 (11), in which product concentration is proportional to biomass and the growth dependence of q_P is evident.

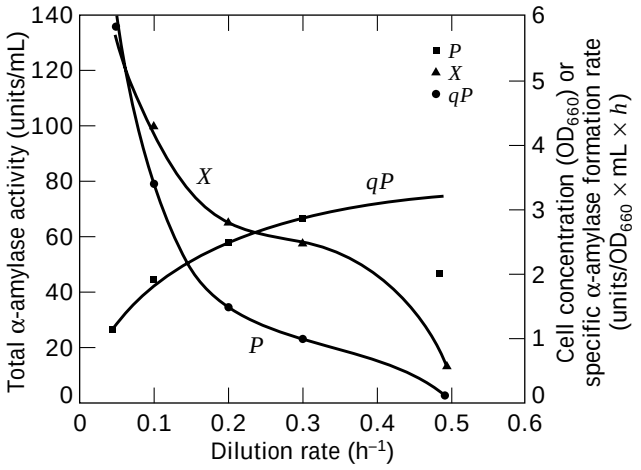


Figure 3. Production of α-amylase in a chemostat by recombinant *Escherichia coli*. The authors (11) used a modified Leudeking–Piret model, $q_P = (\alpha\mu + \beta)(1 + k\mu)^{-1}$, to describe α-amylase kinetics. The term $(1 + k\mu)^{-1}$ accounts for an observed increase in plasmid copy number with decreasing growth rate. Model parameters were regressed from data: $\alpha = 34.12$ units/mL $\times$ OD_{660}, $\beta = 4.2 \times 10^{-10}$ U/mL $\times$ OD_{660} $\times$ h, and $k = 8.63$ h. Experimental data are depicted as points, model predictions as solid lines, and trend lines as dotted lines.

In the case of secondary metabolites (i.e., $\alpha \ll \beta$), product concentration is inversely proportional to the dilution rate, and the productivity ($D \times P$) is independent of dilution rate. The low dilution rates favorable for secondary metabolite production approach batch operation, which is generally favored over the CSTR in such instances.

Other issues can impact the decision between batch and continuous culture. The ability of the CSTR to maintain an ideal environment for product formation may offer a competitive advantage over the batch fermentation, with its time-varying environment and prolonged lag and stationary growth phases. Regulatory and market factors also play an important role in deciding on the operating mode. The CSTR is a dedicated manufacturing system used to produce a single product. Such a system may not be well suited for the production of specialty chemicals and pharmaceuticals because it can neither adapt to variable market demand nor satisfy demand for multiple products.

Energy Balance. Heat transfer is an important consideration in fermentor design, scale-up, operation, and sterilization. The energy balance is used to determine the time–temperature profile of the fermentation broth by accounting for the transfer and accumulation of energy. The heat transfer rate limits the ability to reduce the cycle time for sterilization (12). More importantly, the rate of heat removal from the broth during cell growth can constrain volumetric productivity, an issue in very large reactors with reduced area-to-volume ratios. Because considerable heat generation accompanies rapid cell growth, the high specific growth rates favored in industrial CSTR applications will exacerbate the problem of heat transfer in large reactors. The steady-state energy balance for the fluid in the CSTR is written as

$$Q_{AGIT} + Q_{MET} + \Delta Q_{SENS} - Q_{LOSS} - Q_{EVAP} - Q_{EXCH} = 0 \quad (21)$$

in which Q_{AGIT} is the mechanical energy imparted to the fluid through impeller agitation (equal to the gassed power input), Q_{MET} is the metabolic heat generated by cell growth, ΔQ_{SENS} is the net sensible heat added to the system by streams entering and leaving the system, Q_{LOSS} is the sum of the heat losses from the system to the surroundings, Q_{EVAP} is the latent heat removed by evaporation, and Q_{EXCH} is the heat removed from the system by an appropriate heat exchanger system (13). In some cases, the heats of solution and mixing must be accounted for, but in most cases they are negligible. The terms ΔQ_{SENS}, Q_{LOSS}, and Q_{EVAP} are comparatively small, leading to the simplified energy balance

$$Q_{EXCH} = Q_{MET} + Q_{AGIT} \quad (22)$$

For fast-growing microorganisms, the heat exchanger duty can be as high as 7.7 to 23.2 kW/m^3, of which Q_{MET} and Q_{AGIT} typically represent about 75% and 25% of the total, respectively (14).

Approximately 40 to 50% of the energy contained by a substrate is converted into useful chemical energy, whereas the balance is released as heat. If this metabolic heat is not removed from the fermentation broth, the temperature will rise and possibly hinder performance. Metabolic heat generation is a function of the growth rate of the organism, the cell concentration, the fluid volume, and the efficiency of cell growth on a particular substrate (i), which can be expressed as the metabolic heat released per gram of cell produced ($1/Y_{Hi}$) (kcal/g DCW).

$$Q_{MET} = V\frac{\mu \cdot X}{Y_{Hi}} \quad (23)$$

In aerobic fermentations, oxygen is the final electron acceptor in substrate metabolism, enabling a correlation between the rates of oxygen uptake and heat generation. The following empirical correlation (15) gives Q_{MET} (kcal/h) as a function of specific oxygen consumption, q_{O_2} (mmol O_2/L $\times$ h), and V(L):

$$Q_{MET} = 0.12 \cdot V \cdot q_{O_2} \quad (24)$$

where the oxygen demand during exponential growth can be expressed using the cellular yield on oxygen, Y_{X/O_2} (g DCW/g O_2).

$$Q_{O_2} = \frac{\mu \cdot X}{Y_{X/O_2}} \quad (25)$$

Table 2 gives values of Y_{Hi} and Y_{X/O_2} for bacterial growth that can be used to estimate oxygen demand and metabolic heat generation. Notice that the more reduced substrates result in greater O_2 demand and heat generation and subsequently a larger burden on the O_2 transfer and heat exchange system.

Heat exchange systems are chosen based on the expected heat exchanger duty, influence on fluid mixing, im-

pact on cleaning and sterilization, utility economy, and maintenance, operating, and capital costs. Heat exchangers for fermentors commonly consist of a jacket or shell around the vessel, internal coolant coils, or occasionally an external heat exchanger. The rate of heat removal by a heat exchanger system can be described using a Fourier's law expression:

$$Q_{\mathrm{EXCH}} = U \cdot A \cdot (T - Tc) \tag{26}$$

in which A is the surface area available for heat transfer, U is an overall heat transfer coefficient accounting for all heat transfer resistances, and $(T - Tc)$ is the driving force for heat transfer, where T is the bulk temperature of the fermentation broth and T_C is the temperature of the cooling fluid used in the heat exchanger. Typical values of the heat transfer coefficient are 50 to 150 BTU/ft^2 × h × °F (280 to 850 W/m^2 × K). An important consequence of fermentor scale-up is a decreasing surface-area-to-volume ratio (A/V): at increasing scale the capacity for heat removal relative to heat generation diminishes and often becomes limiting at larger volumes.

The rate of heat removal can be increased by increasing the temperature driving force, increasing the heat transfer surface area, or reducing resistance to heat transfer (increasing U). Water is primarily used as the coolant fluid because of its availability and low cost relative to a refrigerated coolant system. The temperature of the cooling water increases as it passes through the heat exchanger, so an arithmetic or logarithmic mean Tc may give a more representative measure of the coolant temperature. Increasing the coolant flow rate can decrease the mean coolant temperature but with a subsequent increase in utility consumption. Furthermore, the pressure drop of the exchanger and pump capacity limit the extent to which coolant flow rate can be increased. The heat exchanger area is dictated by the size and type of heat exchange system chosen during process scale-up. The overall heat transfer coefficient and constituent resistances are thoroughly discussed elsewhere. In general, poor mixing contributes to decreased heat transfer (see "Nonideal Mixing"). It is sufficient to focus on the convective heat transfer coefficient on the fermentation broth side, h[W/m^2 ∗ K], which is often the dominant heat transfer resistance. The convective heat transfer coefficient is a function of the Reynolds number (Re), Prandtl number (Pr), viscosity ratio (Vi), and the system geometry (F_{GEOMETRY}). In many cases $Vi \approx 1$, and the exponents a and b have typical values of 0.8 and 0.3, respectively.

$$Nu = C_1 \cdot Re^{\mathrm{a}} \cdot Pr^{\mathrm{b}} \cdot Vi^{\mathrm{c}} \cdot F_{\mathrm{GEOMETRY}}$$

$$\frac{h \cdot D_{\mathrm{T}}}{k} = C_1 \cdot \left(\frac{D_{\mathrm{i}}^2 N \rho}{\mu}\right)^{\mathrm{a}} \left(\frac{Cp\mu}{k}\right)^{\mathrm{b}} \left(\frac{\mu}{\mu_{\mathrm{w}}}\right)^{\mathrm{c}} F_{\mathrm{GEOMETRY}} \tag{27}$$

Economic Considerations. Operating costs are those costs associated with maintaining a given production level dictated by the scale and scheduling of the process, typically grouped as raw materials, direct expenses, and indirect expenses. Although the focus here will be on raw materials and direct expenses (e.g., utilities, maintenance, operating supplies, and labor) associated with the CSTR, it is important to view the reactor system in the context of the entire process. The process flow diagram will form the basis for determining not only the reactor operating costs but also the capital and operating costs of air compressors, pumps, holding tanks, and various downstream units. It is important to remember that the relative costs associated with the reactor system, batch or continuous, may not have a significant impact on the total process economics if product recovery costs are dominant (see "Product Formation"). Raw materials costs include the material and handling costs of components added to the system to satisfy metabolic (e.g., C, N, O sources) or process (e.g., acids or bases and antifoam agents) requirements. Direct expenses include the cost of utilities, maintenance, operating supplies, operating labor, direct supervision, laboratory charges, and patent royalties. Continuous systems are easier to automate and offer the potential of lower labor costs than batch production systems, with their labor intensive start-up and shut-down operations. Indirect expenses include taxes and depreciation, usually expressed as a percentage of the plant cost.

Media. The selection of bioprocess media typically involves a trade-off between media cost and the product yield and titer. It is an important stage of development that can influence the design and performance of the entire process. Selecting an apparently cheap media, for example, may result in more expensive downstream recovery and waste treatment operations. In general, media selection includes a number of technical and economic considerations: yield or titer of desired and undesired products; cost; variability in composition and price; availability; effect on downstream processes; need for pretreatment or supplements; shipment, storage, and handling, and need for testing and validation.

Media can be classified as defined or undefined with respect to chemical composition. Laboratory-scale chemostat investigations commonly use defined media to allow precise control over the growth-limiting nutrient and medium composition. Small bioreactors (1 to 4 L) are preferred for continuous operation in the laboratory because media preparation and storage are less of a burden. Industrial microbial fermentations predominantly use undefined media because they generally are less costly and perform better than defined media. Undefined microbial media often contain agricultural by-products (e.g., molasses or corn steep liquor) and thus are subject to source market fluctuations in quality and price. Processes that have been validated with a variety of media may change the production media to take advantage of market changes. The costs of raw materials as a percentage of operating costs for primary metabolites can range from 40% for citric acid to 70% for ethanol from sugar cane (17), whereas for secondary metabolites media costs can be around 10 to 20%. At small scale (less than 10 m^3), mammalian cells are often grown in undefined serum, an expensive and sometimes scarce media derived from mammalian plasma. Viral contamination and the presence of serum proteins can complicate cultivation, product recovery, and quality control and quality assurance (QA–QC) in industrial processes. These complications, combined with regulatory pressure to safeguard

against viral contamination, are a driving force toward the development of defined, serum-free media in order to avoid use of animal-derived media components.

Utility Expenses. Steam, cooling water, and power requirements comprise the majority of utility expenses. Process demand for water-for-injection (WFI) must also be determined for clean-in-place (CIP) systems. Steam usage occurs predominantly during media and equipment sterilization and can be calculated using knowledge of the sterilization cycle. Continuous reactors readily lend themselves to the use of continuous, as opposed to batch, media sterilization systems, offering advantages in reduced thermal degradation of heat-sensitive media, reduced sterilization time, more efficient fermentor use, and greater steam economy (about 20 to 25% of the steam used by batch sterilization) (12). The cooling water requirement, w_{H_2O} (kg/h), can be calculated knowing the heat exchanger duty from the energy balance (equation 21):

$$w_{H_2O} = \frac{Q_{EXCH}}{C_p \cdot (T_{out} - T_{in})} \tag{28}$$

where C_p is the heat capacity of water [$\sim$1 kcal/kg $*$ K] and T_{out} and T_{in} are the outlet and inlet cooling water temperatures, respectively. The availability of abundant low-temperature water can reduce cooling water requirements and possibly allow the use of larger reactors because of the improved ability to remove metabolic heat.

Oxygen demand and mixing requirements drive power consumption by agitators and compressors, whereas larger volumes, higher cell densities, higher specific O_2 uptake rates, and higher broth viscosities result in increased power requirements. Total power consumption for agitated vessels is typically in the range of 2 to 10 kW/m^3. Centrifugal pumps are predominantly used in bioprocesses for their relatively low cost and ability to handle suspended solids. Although pumps represent only a small fraction of power consumption, they can be a significant percentage of overall maintenance costs. In addition, the performance of continuous processes is particularly susceptible to pump failure and may warrant additional capital investment toward the installation of backup pumps in parallel.

Actual vs. Ideal Behavior. In this section, the assumptions used in the development of the ideal CSTR theory are revisited in order to determine when they are invalid and to gauge the impact of nonidealities on performance.

Nonideal Mixing. The major assumption of the ideal CSTR is that there are no spatial variations of properties inside the vessel. Such ideal mixing is never observed in actual systems and even deteriorates on scale-up, although it can be a good approximation of behavior. The challenge is to determine when nonidealities can be expected and how they will influence performance.

Equipment design, operating conditions, and broth properties all influence the quality of fluid mixing in the vessel. Agitated tanks are designed to provide good mixing through selection of tank geometry, baffle placement, and impeller design, although mixing quality decreases with increasing scale. Agitator power input and gas sparging rates, although typically associated with oxygen transfer, are critical to mixing. High broth viscosities, typical of mycelial fermentations, contribute to poor mixing. Poor mixing can affect oxygen transfer (18), product formation (19), heat transfer, process monitoring and control, and the distribution of components added to the system (20,21).

Nonideal mixing in continuous systems is typically characterized by either the residence time distribution, the distribution of fluid residence times around the ideal residence time (τ) (22), or the mixing time (t_M), the time required for a system to respond to a feed disturbance. The ideal mixing time is equal to zero (instantaneous), and the mean value of the residence time distribution is τ (or D^{-1}). Residence time distributions are typically measured by pulse or step addition of a tracer into the reactor feed and tracking the appearance of the tracer in the exit stream as a function of time. Correlations obtained from dimensional analysis can be used to predict mixing times, which increase with scale and broth viscosity (23). Small vessels ($<$500 L) are generally well mixed ($t_M \sim$ s), but large fermentors ($>$5,000 L) typically have poor mixing ($t_M \sim$ min). Compartmental mixing models, in which the bulk fluid is modeled as discrete CSTRs and plug flow reactors (PFRs) with fluid interchange, have been used to describe heterogeneity in large vessels (24).

Wall growth is a special case of nonideal mixing in which cells adhere and proliferate on vessel surfaces (25), where they are hidden from cell mass measurements obtained from samples of the bulk fluid. The system, in addition to being heterogeneous, is no longer at steady state because cells are accumulating in the vessel. This can pose serious problems if the accumulating organism is a contaminant or otherwise undesirable organism. The metabolism and growth of wall-bound cells can be quite different from the suspended population because of mass transfer limitations. Wall growth can reduce heat transfer, create sterilization and cleaning problems, and corrupt measurements in experimental systems. It can be a significant factor when the surface-area-to-volume ratio (A/V) is high (e.g., laboratory-scale systems and vessels with internal cooling coils) and may require modifications to chemostat design or operation (26). For example, the glass walls of laboratory vessels are sometimes treated with organosilane compounds to minimize wall growth. Yeast and mycelial cells with a propensity to form pellets are prone to wall growth under the same conditions that favor flocculation. A discrete washout point does not exist with wall growth, because cells are effectively immobilized in the vessel even past D_C.

Substrate Assumptions. Deviations from ideal chemostat behavior may arise when assumptions regarding the magnitudes of nutrient uptake, the consistency of cell composition, or the identity of the growth-limiting nutrient become invalidated. Substrate uptake for growth in equation 3 was assumed to be much larger than that for maintenance and product formation. At low dilution rates (i.e., low growth rates) the maintenance term becomes significant ($m \sim \mu/Y_{X/S}$) and less substrate goes toward cell growth, causing the actual X to be less than predicted by the ideal theory at low D. Similarly, at high dilution rates, the production and accumulation of growth-related prod-

ucts and intermediates can become significant, leading to a reduced cell yield at higher D.

The ideal chemostat derivation assumes that cell composition does not change over the operating region. In actuality, cell composition varies with pH, temperature, growth rate (27), and medium composition. As cell composition changes, the demand for essential nutrients will change in ways that were not accounted for in the derived equations. Proper evaluation of experimental data from a chemostat may require consideration of the variation of nutrient uptake and cell composition with environment and growth rate.

When using complex and undefined media or an organism with complex nutritional requirements (e.g., mammalian cells), it is often difficult to identify the growth-limiting nutrient. In addition, the limiting nutrient itself may change because the nutrient demand, and subsequently the media composition, may change with operating conditions. The cell concentration profile in these situations likely would be constantly decreasing with dilution rate, unlike the ideal chemostat where X is approximately constant over the majority of $0 < D < D_C$. More complicated, structured growth models would be required to account for such behavior.

Non-Steady-State Behavior. The potential sources of process variability and non-steady-state behavior are perhaps too numerous to mention. However, typical instances include chemostat start-up, execution of control actions, induced disturbances (e.g., pulse and shift methods [28]), variations in feed composition, culture degeneration (e.g., plasmid loss or apoptosis), wall growth, or equipment failures. Steady state often is declared when the measurable process states are maintained constant for 3 to 5 residence times. Sustained oscillations are sometimes observed in continuous culture and often are the result of growth inhibition resulting from either an accumulated product (29) or the burden of product formation (30).

Selection/Mutation and Contamination. By controlling the culture conditions in the CSTR, a highly selective environment for the selection and proliferation of certain microorganisms can be created. Cells in this selective environment with growth rates less than the dilution rate will be washed out of the reactor, leaving only those cells with the properties that have been selected for. In this way, continuous culture can be used as a strain improvement tool to select organisms that possess a desirable trait, such as yeast with higher ethanol tolerance (31).

Because of the metabolic burden imposed by high levels of product formation, the production strain has a growth disadvantage relative to unproductive strains that are present in the reactor. Without selection pressure in favor of the production strain, a gradual decline in productivity will be observed over time as nonproductive cells dominate the culture. Examples include the reversion of specially selected antibiotic strains to low productivity mutants or the domination of recombinant protein processes by plasmid-free cells. The configuration of CSTRs in series can be used to circumvent this problem by providing separate environments for growth and product formation (see "CSTRs in Series"). Selective pressure in favor of the production strain (and against contaminants) may be exerted through the application of elevated temperatures, extremes of pH, the use of narrowly defined or modified media, and the use of specially selected cultures (e.g., antibiotic-resistant strains).

The prolonged operating periods of continuous culture increase the probability of contamination by a foreign organism. The threat posed by contamination depends on the ability of the undesirable microorganism to complete and thrive in the CSTR environment. Consider the case of two types of microorganisms with concentrations X (the desired strain) and Z (the contaminant) competing for the same limiting substrate in a CSTR. The material balances on cell mass, neglecting cell lysis, can be written as

$$\frac{dX}{dt} = \mu \cdot X - D \cdot X \tag{29}$$

$$\frac{dZ}{dt} = \mu_z \cdot Z - D \cdot Z \tag{30}$$

Subtracting equation 30 from equation 29 and rearranging leads to equation 31:

$$\frac{d \ln[X/Z]}{dt} = \mu - \mu_z \tag{31}$$

which shows that the growth rates and their dependence on the limiting substrate will determine the fate of the culture (32). The contaminant (Z) could be washed out ($\mu > \mu_z$), remain at a stable level ($\mu = \mu_z$), or dominate ($\mu < \mu_z$) the culture. This simple analysis of selection can be complicated if the contaminant has properties that prevent it from being washed out (e.g., adhesion to reactor surfaces), if the contaminant competes for a different substrate than the production strain, or if there are interactions from inhibitory cellular products. For example, lactobacilli are often a persistent contaminant of continuous ethanol production processes because of their intimate association with flocculant yeast aggregates and ability to adapt to high ethanol concentrations (33). Selective recycle of desirable organisms back to the vessel has been used to prevent domination of the culture by undesirable strains (34–36).

Enzymes

Introduction. Enzymes are biological catalysts with high selectivity toward reactants and products, making them attractive for use in a number of industrial applications. Enzyme activity is strongly influenced by the environment (e.g., pH, temperature, metal ions). Loss of activity or denaturation can be reversible or irreversible, depending on the type, strength, and duration of an unfavorable interaction. A benefit of using the CSTR for enzyme reactions is that the constant, controlled reactor environment can be designed for maximum enzyme activity and life.

Material Balances. Assuming that the inlet and outlet flow rates are approximately equal (i.e., solutions are dilute), the steady-state material balance on the substrate can be written as

$$F \cdot (S_0 - S) = v \cdot V \tag{32}$$

where F is the volumetric flow rate, v is the rate of substrate consumption by reaction, V is the fluid volume in the reactor, and S_0 and S are the substrate concentrations in the feed and vessel, respectively. Rewriting equation 32 in terms of the fluid residence time (τ) and the fractional substrate conversion (δ_S) yields the CSTR design equation

$$\tau = \frac{V}{F} = \frac{S_0 \cdot \delta_S}{v} \tag{33}$$

Using a valid rate expression for v, the design equation can be used to determine the reactor volume required to yield a given conversion rate ($S_0 \times \delta_S \times F$). Fast reaction kinetics are obviously favorable because reactor cost scales with reactor size. Enzyme loading in the reactor can be increased beyond the solubility limit by immobilization on inert support particles, which increases v, reduces the necessary reactor volume, facilitates enzyme retention and recycle, and may improve enzyme stability.

Enzyme Reaction Kinetics. Numerous mechanistic models have been developed to describe enzyme reaction rates as a function of enzyme and substrate concentrations. Some of the more common models appear in Table 3 accompanied by the corresponding solution to the design equation (equation 33).

Unlike a plug flow reactor (PFR), in which the substrate enters at a high concentration and leaves at a lower concentration, the substrate concentration in a CSTR is at a uniform, low concentration. The reduced substrate concentration leads to a slower reaction rate, so that the CSTR requires more of the active enzyme than the PFR to attain the same substrate conversion rate. Substrate inhibition is less problematic in a CSTR than in a PFR because of the lower substrate concentration in the bulk fluid, whereas product inhibition is generally more of a problem in CSTRs than PFRs. Arranging CSTRs in series can reduce the effects of product inhibition (approaching PFR behavior) while taking advantage of the good mixing characteristics of the CSTR to provide optimal pH control (37).

Temperature Effects. Like many chemical reactions, increasing the temperature enhances the rate of enzyme reactions. Higher temperatures also result in increased rates of thermal denaturation and loss of the active biocatalyst. Process economics often depend on optimal temperature control to maintain high substrate conversion and long catalyst life (38–40). The effects of temperature on the catalytic rate constant (k_2) can be described using an Arrhenius expression

$$k_2 = A \cdot e^{-Ea/RT} \tag{34}$$

where A is the Arrhenius constant, E_a is the activation energy, R is the gas constant, and T is the absolute temperature. The activation energy of enzyme-catalyzed reactions ranges from 4 to 20 kcal/mol, with most reactions near 11 kcal/mol.

Thermal denaturation usually can be described as a first-order decay reaction:

$$\frac{dE}{dt} = -k_d \cdot E \text{ or } E = E_0 e^{-k_d t} \tag{35}$$

where k_d is the thermal denaturation constant, which also follows an Arrhenius temperature dependence. For thermal denaturation, E_a varies from 40 to 130 kcal/mol, with most in the vicinity of 70 kcal/mol. Increasing temperature has a greater effect on the rate of denaturation than catalysis. For a typical enzyme (i.e., 11 and 70 kcal/mol), an increase in temperature from 30 to 40 °C results in a 1.8-fold increase in the rate of catalysis, but a 41-fold increase in the denaturation rate.

Energy Balance. An optimal temperature control strategy requires good heat removal because most industrial enzyme reactions are exothermic. Although heat transfer is generally good for soluble enzymes in agitated tanks, the high enzyme concentrations attained with immobilization can result in fast reaction rates and appreciable heat generation. Heat transfer resistance within the catalyst pellet can reduce heat removal rates, resulting in higher pellet temperatures and shorter catalyst life. Catalyst degradation from insufficient heat removal is more of a concern in packed beds with high enzyme loading, where heat transfer resistances can be significant. The steady-state energy

Table 3. Some Common Enzyme Kinetic Expression

	Rate expression $v =$	Design equation $\tau =$
Michaelis–Menten	$\frac{v_{max} \cdot S}{K_m + S}$	$\frac{1}{v_{max}}\left[S_0 \cdot \delta_S + K_m \cdot \frac{\delta_S}{1-\delta_S}\right]$
Substrate inhibition	$\frac{v_{max} \cdot S}{S + K'_m + \frac{S^2}{K'_S}}$	$\frac{1}{v_{max}}\left[S_0 \cdot \delta_S + K'_m \cdot \frac{\delta_S}{1-\delta_S} + \frac{S_0^2}{K'_S} \cdot (\delta_S - \delta_S^2)\right]$
Competitive product inhibition	$\frac{v_{max} \cdot S}{K'_m\left(1 + \frac{P}{K_P}\right) + S}$	$\frac{1}{v_{max}}\left[S_0 \cdot \delta_S + K'_m \cdot \frac{\delta_S}{1-\delta_S} + \frac{K'_m}{K_P} \cdot \left(\frac{S_0 \cdot \delta_S^2}{1-\delta_S}\right)\right]$
	Where $v_{max} = k_2 \times E_0$	Where the product concentration (P) is related to converted substrate; $P = \delta_S \times S_0$

balance for an exothermic ($\Delta H_{RXN} < 0$) enzyme reaction in a well-mixed CSTR is

$$F\rho c_P(T_F - T) + v' \cdot E \cdot V(-\Delta H_{RXN}) - Q_{EXCH} = 0 \quad (36)$$

where F is the volumetric flowrate (m^3/h), ρ is the fluid density (kg/m^3), c_P is the fluid heat capacity (kJ/kg × K), v' is the specific reaction rate (kmol substrate/kg enzyme × h), E is the enzyme concentration (kg enzyme/m^3), ΔH_{RXN} is the heat of reaction (kJ/kmol substrate), Q_{EXCH} is the heat removed by the heat exchanger (kJ/h) from equation 26, and T_F and T are the feed and bulk fluid temperatures (K), respectively.

Economic Considerations. For enzyme CSTRs, the primary operating costs are associated with enzyme replacement. Prolonged catalyst activity and marked reductions in raw materials costs can be achieved by maintaining an optimal environment for the enzyme during operation. Preserving enzyme activity reduces the number and frequency of labor-intensive cleaning and changeovers, facilitating downstream operations by consistently providing a constant-quality product stream. With multiple reactors installed in parallel, changeovers can be scheduled to minimize production variations and downtime. Industrial enzymes are often sold as a crude mixture containing only a fraction of active enzyme. Selecting enzymes among different vendors may involve a trade-off between cost and purity (percent active enzyme) and a consideration of how the impurities may affect the process.

MASS TRANSFER

Introduction

The rate of mass transfer ultimately will limit the maximum aerobic reactor performance. Oxygen transfer to the fermentor broth, for example, can limit both the extent and rate of cell growth. Mass transfer limitations to microbial flocculants and immobilized catalyst pellets can result in reduced reaction rates and inefficient conversion. Liquid–liquid mass transfer rates from hydrocarbon substrates to suspended cells may limit productivity in two-phase systems (41). To determine the rate-controlling regime, it is useful to characterize the relative rates of mass transfer and reaction using the dimensionless Damkohler number (Da):

$$Da = \frac{\text{Maximum rate of reaction}}{\text{Maximum rate of diffusion}} \quad (37)$$

The observed reaction rate may be limited by the rate of diffusion depending on the value of the Damkohler number: if $Da \gg 1$ the diffusion rate is limiting, if $Da \ll 1$ the reaction rate is limiting, and if $Da \sim 1$ then the reaction and diffusion rates are comparable. As with all dimensionless numbers, the Damkohler number is only meaningful if it is calculated using the proper time and length scales for a given system. Consider spherical pellets ($r = 400\,\mu$m) of *Penicillium chrysogenum*, assuming a pellet cell density of 0.1 g/cm^3, an effective oxygen diffusivity ($D_{O_2}^{eff}$) of 1×10^{-6} cm^2/s, a particle size of 400 μm, an oxygen concentration (c_{O_2}) of 7 mg/L, and the following cell parameters (42): $\mu = 0.075\ h^{-1}$, $Y_{X/O_2} = 1.56$ g DCW/g O_2, and $m_{O_2} = 0.024$ g O_2/(g DCW × h). The Damkholer number shows that internal mass resistance is considerable and that cell growth in the pellet is likely to be limited by oxygen transfer.

$$Da = \frac{\overbrace{\left(\frac{\mu}{Y_{X/O_2}} + m_{O_2}\right)}^{O_2\ DEMAND} \cdot \underset{\sim}{\overset{PELLET\ CELL\ DENSITY}{X}} \cdot \overbrace{\left(\frac{4\pi r^3}{3}\right)}^{PELLET\ VOLUME} \cdot}{\underbrace{\left(D_{O_2} \cdot \frac{c_{O_2}}{r}\right)}_{DIFFUSIVE\ FLUX} \cdot \underbrace{(4 \cdot \pi \cdot r^2)}_{PELLET\ SURFACE\ AREA}} =$$

$$\frac{\left(\frac{0.075\ h^{-1}}{1.56\ g\ DCW/g\ O_2} + \frac{0.024\ g\ O_2}{g\ DCW \cdot h}\right) \cdot \frac{0.1\ g\ DCW}{cm^3} \cdot (4 \times 10^{-2}\ cm)^2}{3 \cdot (1 \times 10^{-6}\ cm^2/s) \cdot (7 \times 10^{-6}\ g/cm^3)}$$

$$\cong 5.5 \times 10^5 \quad (38)$$

Gas–liquid mass transfer is often rate-limiting for gases that are sparingly soluble in the broth, such as oxygen and methane. Although highly soluble, carbon dioxide exhibits pH-dependent partitioning between gaseous and dissolved forms (CO_2, H_2CO_3, HCO_3^-, CO_3^{2-}) that is influenced by the rates of both reaction and mass transfer. Proper interpretation of the respiratory coefficient (RQ) in fermentations operated at neutral pH requires consideration of CO_2 dynamics. The OTR has already been used to determine the productivity limit of a CSTR used for biomass production in equation 12. In general, the rate of mass transfer from the gas to the liquid phase is given as

$$N_A = k_L a(C^* - C_L) \quad (39)$$

where N_A is the rate of gas transfer (mmol/l × h) and the remaining terms have the same definitions as in equation 12. For sparged, agitated tanks, $k_L a$ has typical values in the range from 50 to 1,400 h^{-1}. Several correlations have been developed for $k_L a$ as a function of the gassed power input per unit volume and the superficial gas velocity for Newtonian broths in a variety of fermentors (43). The correlations can offer wide variability in mass transfer estimates and should be used in conjunction with knowledge from past experience or empirical measurements of $k_L a$ (e.g., dynamic or sulfite oxidation methods). Oxygen transfer to shear-sensitive mammalian cells requires gentle agitation combined with surface or membrane aeration, or light sparging, as opposed to the large power inputs and high rates of gas sparging in microbial fermentations. This limitation is somewhat offset by the fact that mammalian cells have lower O_2 requirements (0.05 to 0.5 mmol/10^9 cells × h) (44) and grow to lower cell densities (10^6 to 10^7 cells/mL) than microbial cultures.

VARIATIONS ON THE SINGLE CSTR

Single CSTR with Recycle

Volumetric productivity is related to the concentration of active catalyst. Cell or enzyme concentrations greater than

the steady state obtained from the simple CSTR can be achieved by separating cells from the effluent stream and recycling them to the vessel (27,45) or by retaining them within reactor. Higher catalyst concentrations enhance substrate conversion and reduce the reactor size necessary to attain a given conversion. Recycle operation improves system stability in the face of feed disturbances by retaining cells in the vessel even under conditions that would cause washout in the simple CSTR. Recycle systems can be operated at dilution rates, or throughputs, greater than the specific growth rate of the organism. Productivity improvements achieved with cell recycle are demonstrated in Table 4 for *Saccharomyces cerevisiae* ATCC 4126 and *Zymomonas mobilis* ATCC 10988 at 100 g/L glucose feed.

Cell Recycle Methods. Cell recycle is implemented through a cell separation step, often by a unit operation commonly used in the initial stages of downstream processing. Typical methods of continuous cell separation include centrifugation, filtration, and sedimentation. Cell separation can be viewed as having two often equally important purposes: (*1*) the recovery or retention of cells for reuse and (*2*) the removal of potentially inhibitory byproducts or products from the culture environment. The separation step often has to satisfy additional performance requirements such as handling of shear- or temperature-sensitive materials, selectivity in rejection or recovery, containment, maintenance of asepsis, corrosion resistance, brief retention time, and ease of cleaning, sterilization, maintenance, and validation. Recycle operation is standard for reactors using stable enzymes, because discarding expensive active catalyst is economically unfeasible. Cell separation operations are discussed elsewhere in the context of downstream processing, although a brief description is presented here in relation to cell recycle.

Sedimentation. Sedimentation is the settling of particles in a gravitational field. With low energy requirements and simple equipment, sedimentation is a relatively inexpensive way of separating a dilute cell phase. Waste treatment is by far the largest application of sedimentation-based cell recycle, in which cells are typically separated in large sedimentation tanks using lime or clay to enhance flocculant formation. The settling velocity (u_0) for an isolated spherical particle can be described using Stoke's law:

$$u_0 = \frac{d_P^2(\rho_P - \rho_F)g}{18\eta} \tag{40}$$

in which d_P is the particle diameter, ρ_P is the particle density (the specific gravity of a typical cell is 1.05), ρ_F is the fluid density, g is the gravitational acceleration constant, and η is the fluid viscosity. The functional dependence of Stoke's law suggests ways of increasing the settling velocity. The easiest and most common method is to increase the effective cell size by promoting flocculation (cell aggregation) through physiological, chemical, and physical factors: selection of flocculant strains; modification of cell wall structure or surface charge; changing the pH, temperature, or shear stress; addition of inorganic salts (e.g., Ca^{2+} and Mg^{2+} salts) or clays; controlling the concentration of certain nutrients or products (e.g., extracellular polysaccharides); and controlling the cell age or growth phase. Selective cell recycle has been implemented using the differential sedimentation properties of a desired and unwanted microorganism (34–36). The properties of a particular broth are generally unchangeable and will probably only impede particle settling.

Although equation 40 holds for dilute suspensions of cells, the interactions among settling particles in concentrated slurries results in hindered settling. The hindered particle velocity (u_h) is influenced by the particle concentration and can be expressed with the following correlation (49):

$$\frac{u_h}{u_0} = \frac{1}{1 + \lambda \cdot \epsilon_P^{1/3}} \tag{41}$$

in which ϵ_P is the volume fraction of particles and λ is an empirical function of ϵ_P. For dilute suspensions, $\epsilon_P < 0.15$, whereas in slurries $0.15 < \epsilon_P < 0.50$.

The limiting settling velocity for a system has a strong influence on equipment design and operation. Consider the case of a continuous sedimentation tank with volumetric throughput (F) and constant cross-sectional area (A). The sedimentation tank performance can then be described by equation 42, in which throughput is directly proportional to A and independent of tank depth:

$$F = u_{lim} \cdot A \tag{42}$$

The throughput and limiting settling time will thus dictate equipment size and costs. Another design consideration is the residence time of cells in the settling tank, which must be considered in the context of nutrient depletion (particularly for oxygen) and its potential effects on performance.

Sedimentation at laboratory scale may be implemented with an external settling column (47,50,51). Similar devices may be used at bioprocessing scales, whereas large open-air tanks must be used in high-volume wastewater treatment. Internal sedimentation has been implemented

Table 4. Ethanol Productivity Enhancements for *S. cerevisiae* and *Z. mobilis*

		Dilution rate (h^{-1})	Ethanol productivity g/(L × h)	Cell density (g DCW/L)	Reference
S. cerevisiae	CSTR	0.17	7	12	46
	CSTR with recycle	0.68	29	50	47
Z. mobilis	CSTR	0.175	8	2.5	29
	CSTR with recycle	2.7	120	38	48

in tower fermenters, in which immobilized cells and enzymes or microbial flocculants are retained in the vessel by a sedimentation zone within the vessel. Unlike the ideal CSTR, tower fermenters may exhibit spatial variations in nutrient concentrations and broth properties along the height of the tower that can significantly influence reactor performance. In addition, the productivity in these reactors may be limited by the need to maintain low upward velocities (e.g., low aeration or CO_2 evolution) to allow adequate cell sedimentation.

Centrifugation. The operating principle behind centrifugation is the same as that of sedimentation; however, much higher settling velocities than in sedimentation may be obtained in the centrifugal field. Centrifugal separators enable high-volume continuous processing of fluids containing many particles, with short retention times and small space requirements. To determine the unhindered particle velocity in a centrifugal field (u_{0C}), equation 40 is multiplied by the centrifugal coefficient (C), also known as the G-value, which describes the increase in sedimentation rate due to centrifugation relative to gravitational settling:

$$u_{0C} = \frac{d_P^2 g(\rho_P - \rho_F)}{18\eta} \underbrace{\left(\frac{r\omega^2}{g}\right)}_{C} \tag{43}$$

where r is the radial distance from the axis of rotation and ω is the angular velocity.

Industrial centrifuges are most often classified by internal structure (e.g., disk stack, tubular bowl) and mode of operation (e.g., solids retaining, continuous or intermittent solids ejecting). The selection of sturdier construction and materials will enable higher rotation speeds for separation of smaller particles. The equation describing throughput in a centrifuge is analogous to equation 42, except that the centrifuge area is expressed using the Σ value, which is the area equivalent for a given centrifuge and rotation speed. Centrifuge manufacturers will often provide machine-specific Σ values, although the Σ value for simple disk-stack and tubular-bowl centrifuges can be calculated directly.

Filtration. Filtration is separation based on size, allowing retention of molecules larger than the pore size of the filter and passage of smaller molecules. Membrane filtration thus offers the twin benefits of cell retention and inhibitory by-product removal. In cell recycle systems, the most common arrangements are internal filters for cell and enzyme retention (52,53) or external membrane filters (54,55) in plate and frame, spiral cartridge, and hollow fiber configurations. In all these configurations, flow patterns tangential to the membrane surface can reduce fouling and improve the filtrate flux across the membrane. Compared to internal filters, external filters have higher surface-area-to-volume ratios and may be easier to maintain; however, they may be less easily sterilized (particularly for some polymer membranes) and could introduce problems of nutrient depletion in the external recycle loop. Membrane selection depends primarily on the critical particle size, with other criteria being cost, mechanical stability, and susceptibility to plugging and fouling. Because membranes have the potential for complete cell recycle, a purge or bleed stream is typically split from the recycle stream to prevent accumulation of inert particles and debris in the vessel.

Material Balances. A schematic of a CSTR with recycle of cells is shown in Figure 4. A material balance on cell mass for the CSTR with recycle system, neglecting cell death, may be written

$$\underbrace{F \cdot X_0}_{\text{CELLS IN FEED}} + \underbrace{\alpha \cdot F \cdot C \cdot X_1}_{\text{CELLS IN RECYCLE STREAM}} - \underbrace{(1 + \alpha) \cdot F \cdot X_1}_{\text{CELLS OUT}} + \underbrace{V \cdot \mu \cdot X_1}_{\text{CELL GROWTH}} = \underbrace{\frac{d(V \cdot X_1)}{dt}}_{\text{CELL ACCUMULATION}} \tag{44}$$

where α is the recycle ratio equal to the recycle volume divided by the feed volume, C is the concentration factor (cell concentration in the recycle divided by the effluent concentration) related to the efficiency of the separation step, and X_0, X_1, and X_2 are the cell concentrations in the feed, recycle, and separator effluent streams, respectively. Note that the low substrate concentrations in waste treatment create suboptimal growth environments in which cell death cannot be neglected.

Assuming the system is at steady state ($dX_1/dt = 0$) and that the feed is sterile ($X_0 = 0$), equation 44 yields

$$\mu = (1 + \alpha - \alpha \cdot C) \cdot D \tag{45}$$

The dilution rate is no longer equal to the specific growth rate; in fact, because C > 1 and $\alpha < 1$, the dilution rate is greater than the specific growth rate.

A material balance on the limiting substrate, again neglecting maintenance and product formation, may be written

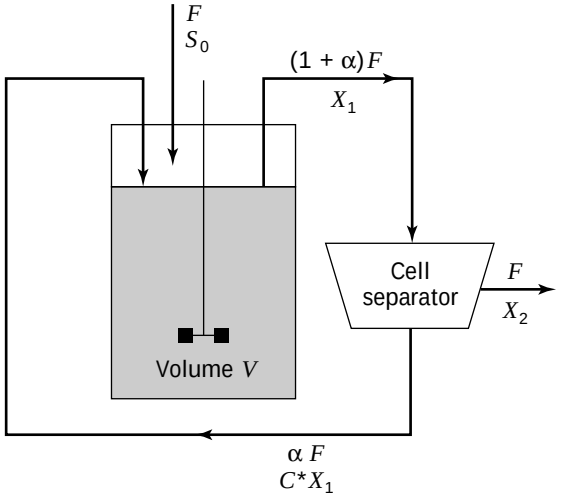


Figure 4. CSTR with recycle. Cell separation can be achieved through centrifugation, sedimentation, or filtration.

$$\underbrace{F \cdot S_0}_{\text{SUBSTRATE IN FEED}} + \underbrace{\alpha \cdot F \cdot S}_{\text{SUBSTRATE IN RECYCLE}} - \underbrace{(1 + \alpha) \cdot F \cdot S}_{\text{SUBSTRATE OUT}} - \underbrace{\frac{\mu \cdot X}{Y_{X/S}} \cdot V}_{\text{SUBSTRATE CONSUMED}} = \underbrace{V \cdot \frac{dS}{dt}}_{\text{SUBSTRATE ACCUMULATION}} \quad (46)$$

Solving equation 46 for the cell concentration, assuming steady-state ($dS/dt = 0$) operation, yields

$$X_1 = \frac{Y_{X/S} \cdot (S_0 - S)}{[1 + \alpha(1 - C)]} \quad (47)$$

in which the steady-state cell concentration with recycle is greater than that in the simple CSTR by a factor of $1/[1 + \alpha(1 - C)]$. By adopting a suitable expression for cell growth, the substrate concentration can be determined. Using the Monod expression, as before, and solving for the substrate concentration gives

$$S = K_S \frac{\mu}{\mu_{max} - \mu} = K_S \frac{D(1 + \alpha - \alpha C)}{\mu_{max} - D(1 + \alpha - \alpha C)} \quad (48)$$

Substituting equation 48 into equation 47 yields

$$X_1 = \frac{Y_{X/S}}{(1 + \alpha - \alpha C)} \left[S_0 - K_S \frac{D(1 + \alpha - \alpha C)}{\mu_{max} - D(1 + \alpha - \alpha C)} \right] \quad (49)$$

A material balance on cell mass around the separator gives the cell concentration in the outlet:

$$X_2 = (1 + \alpha - \alpha \cdot C) \cdot X_1 \quad (50)$$

Figure 5 shows the cell mass and biomass productivity of a CSTR with recycle compared to a simple CSTR. The higher biomass productivity of the recycle system results from a dilution rate higher than the specific growth rate and the increased cell concentration in the vessel.

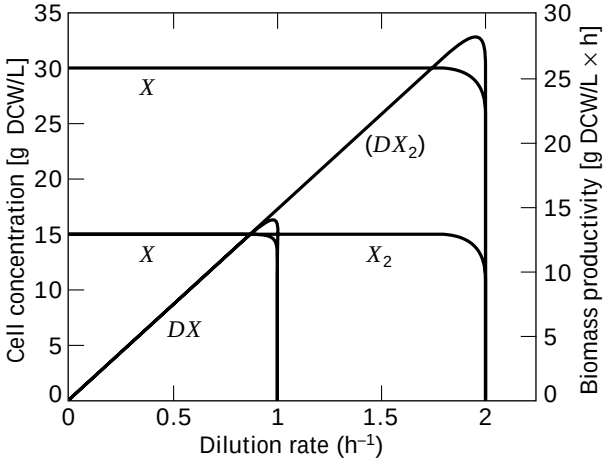


Figure 5. Comparison of steady-state behavior of a chemostat (*solid lines*) and a chemostat with recycle (*dotted lines*) using the following parameters: $Y_{X/S} = 0.5$ g DCW/g substrate, $\mu_{max} = 1.0$ h^{-1}, $K_S = 0.02$ g/L, $S_0 = 30$ g/L, $C = 2$, and $\alpha = 0.5$.

Implementation of a recycle system is often critical to the economic viability of processes using expensive biocatalysts (e.g., enzymes). Typically, this is accomplished by immobilizing the enzymes on inert support particles to facilitate either internal or external recycle. The potential use of cell recycle in an industrial process involves weighing the effectiveness and economics associated with the cell separation step against the marginal improvement in process performance. It should also be noted that higher cell densities exacerbate the oxygen transfer and heat removal burden of the system.

CSTRs in Series

In the single CSTR, the constant, controlled environment gives the advantage of being able to control the cellular–enzyme environment for maximum utility. Sometimes, however, a particular cell system will exhibit multiple properties of interest that can only be realized in different environments. The optimal environments for cell growth and product formation, for example, may be characterized by different temperatures, pH, and limiting nutrients. The configuration of CSTRs in series lends itself to those applications in which multiple environments are required.

Cell Growth. Bacterial growth in the presence of multiple carbon substrates often results in diauxic growth, in which cells preferentially metabolize a single substrate over all others. In a waste treatment application, the preferred substrate would be consumed by the microorganisms and the remaining substrates would pass through the system untreated. Configuring CSTRs in series provides a partitioning of cell metabolism so that less-favored substrates are consumed in subsequent stages.

CSTRs in series have been used to improve recombinant protein fermentations in which performance is threatened by plasmid instability (56,57) and lethal protein overproduction (30). Cells are grown to high density in the first stage without inducer so that plasmid-free cells have little growth advantage over plasmid-containing cells. Induction in the second stage results in higher productivity than the simple CSTR because the continuous introduction of plasmid-containing cells from the first stage reduces the ability of nonproductive cells to dominate the culture.

Consider the two-stage system for biomass production in Figure 6, in which a separate feedstream can be added to the second stage. The steady-state material balances for cells and substrate in the first reactor are identical to the single CSTR case (equations 1 and 3), with steady-state solutions as equations 5 and 8 for Monod growth. Considering the case without the second feedstream, the material balances on cell mass and growth-limiting substrate in the second stage can be written as

$$\underbrace{\frac{F \cdot X_1}{V_2}}_{\text{CELLS IN}} - \underbrace{\frac{F \cdot X_2}{V_2}}_{\text{CELLS OUT}} + \underbrace{\mu_2 \cdot X_2}_{\text{CELL GROWTH}} = \underbrace{\frac{dX_2}{dt}}_{\text{ACCUMULATION}} \quad (51)$$

and

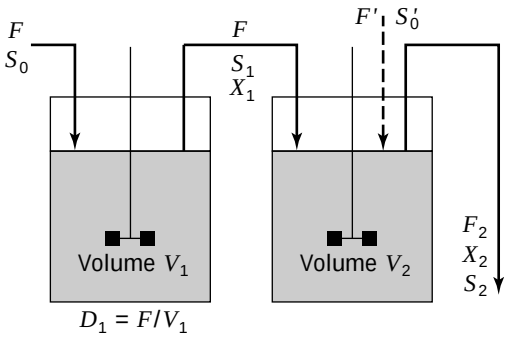


Figure 6. Two-stage chemostat system with possibility of a separate feed to second reactor.

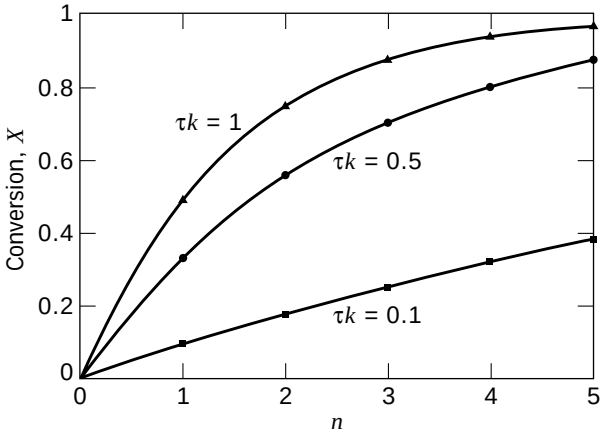


Figure 7. Substrate conversion for first order reaction in n CSTRs in series.

$$\underbrace{\frac{F \cdot S_1}{V_2}}_{\text{SUBSTRATE IN}} - \underbrace{\frac{F \cdot S_2}{V_2}}_{\text{SUBSTRATE OUT}} - \underbrace{\frac{\mu_2 \cdot X_2}{Y_{X/S}}}_{\text{CONSUMPTION FOR GROWTH}} = \underbrace{\frac{dS_2}{dt}}_{\text{ACCUMULATION}} \tag{52}$$

with steady-state solutions shown in Table 5. A growth model on the limiting substrate must be adopted to further complete the system description. Minimal cell growth will occur in the second stage if no additional substrate is added, because the majority of substrate is consumed in the first stage. Thus, non-growth-related cellular behavior and product formation could be studied in the second CSTR.

Adding an additional feedstream to the second stage (Fig. 6) provides the opportunity to introduce more of the limiting nutrient, other nutrients required for growth or product formation, inducers, or inhibitors. The material balances on cell mass and substrate on the second stage can be written as equations 53 and 54, respectively, with steady-state solutions given in Table 5. In equation 53, F' is the volumetric flow of the second feedstream, which is assumed to be sterile ($X' = 0$).

$$\underbrace{\frac{F_1}{V_2} X_1}_{\substack{\text{CELLS IN}\\ \text{FROM}\\ \text{STAGE 1}}} - \underbrace{\frac{F_1 + F'}{V_2} X_2}_{\substack{\text{CELLS}\\ \text{OUT}}} + \underbrace{\mu_2 \cdot X_2}_{\substack{\text{CELL}\\ \text{GROWTH}}} = \underbrace{\frac{dX_2}{dt}}_{\text{ACCUMULATION}} \tag{53}$$

The dilution rate for the second stage is given by $D_2 = (F_1 + F')/V_2$, and the concentration of the limiting nutrient in the second feed is S'_0.

$$\underbrace{\frac{F_1}{V_2} S_1}_{\substack{\text{SUBSTRATE}\\ \text{IN FROM}\\ \text{STAGE 1}}} + \underbrace{\frac{F'}{V_2} S'_0}_{\substack{\text{SUBSTRATE}\\ \text{IN FROM}\\ \text{SECOND FEED}}} - \underbrace{\frac{F_1 + F'}{V_2} S_2}_{\substack{\text{SUBSTRATE}\\ \text{OUT}}} - \underbrace{\frac{\mu_2 \cdot X_2}{Y_{X/S}}}_{\substack{\text{CONSUMPTION}\\ \text{FOR GROWTH}}} = \underbrace{\frac{dS_2}{dt}}_{\text{ACCUMULATION}} \tag{54}$$

Feeding additional substrate to the second stage allows for more growth to occur. In addition, the dilution rate in the second stage is larger than the maximum specific growth rate of the organism because the second stage has a continuous feed of cells.

Enzyme Reaction. Enzyme reactions may also be carried out in multiple CSTRs. The performance of CSTRs in series approaches that of a single PFR while maintaining the good mixing characteristics of the stirred-tank reactor (37). Considering a first-order enzymatic reaction for substrate conversion ($v = k \times S$), the reactor design equation for a single CSTR can be written as

Table 5. Steady-State Solutions to Material Balances for a Two-Stage Chemostat

	Cell mass	Substrate
First stage	$\mu_1 = D_1$	$X_1 = Y_{X/S}(S_0 - S_1)$
Second stage	$\mu_2 = D_2\left(1 - \frac{X_1}{X_2}\right)$	$X_2 = \frac{D_2}{\mu_2} Y_{X/S}(S_1 - S_2)$
Second stage with additional feedstream	$\mu_2 = D_2 - \frac{F_1 \cdot X_1}{V_2 \cdot X_2}$	$X_2 = \frac{Y_{X/S}}{\mu_2}\left(\frac{F_1}{V_2} S_1 + \frac{F'}{V_2} S' - D_2 S_2\right)$
	where $D_2 = D_1 = F/V_1$	where $D_2 = (F_1 + F')/V_2$

$$\tau = \frac{S_0 - S}{k \cdot S} \tag{55}$$

Assuming that there is no volume change upon reaction, the conversion in the single CSTR is given by

$$X = \frac{\tau \cdot k}{1 + \tau \cdot k} \tag{56}$$

Then for a system of n CSTRs in series with equal volumes and reactor conditions (constant k) the conversion in the nth CSTR is given by equation 57 and depicted in Figure 7.

$$X_n = 1 - \frac{1}{(1 + \tau \cdot k)^n} \tag{57}$$

NOMENCLATURE

Abbreviations

CER	Carbon dioxide evolution rate
CIP	Clean-in-place
CSTR	Continuous stirred-tank reactor
OTR	Oxygen transfer rate
OUR	Oxygen uptake rate
PFR	Plug flow reactor
RQ	Respiratory quotient, RQ = CER/OUR
WFI	Water-for-injection

Symbols

ΔH_{RXN}	Heat of reaction, >0 endothermic, <0 exothermic, kJ/mol
ΔQ_{SENS}	Net sensible heat input, W
$(R_{CSTR})_M$	Biomass productivity at X_M and D_M, g DCW/m^3 × h
A	Area, m^2
C	Concentration factor in cell recycle system
C^*	Equilibrium dissolved oxygen concentration, g/m^3
C_L	Dissolved oxygen concentration, g/m^3
Cp	Heat capacity, J/kg × K
D	Dilution rate, h^{-1}
Da	Dimensionless Damkohler number
D_C	Critical dilution rate, h^{-1}
D_M	Dilution rate associated with maximum R_{CSTR} at fixed S_0, h^{-1}
d_P	Particle diameter, m
E	Enzyme concentration
E_0	Initial enzyme concentration
Ea	Activation energy, kJ/mol
F	Volumetric flow rate, m^3/h
g	Gravitational acceleration, 9.8 m/s^2
k_2	Catalytic rate constant, h^{-1}
k_d	Thermal denaturation constant, h^{-1}
k_La	Overall mass transfer coefficient, h^{-1}
k_P	Specific product degradation constant, h$_{-1}$
K_S, K_I, K_P	Model parameters for cell growth
K_m, K_S, K'_m, K'_S	Model parameters for enzyme kinetics
m	Maintenance coefficient g/g DCW × h
N_{discs}	Number of discs in a disk-stack centrifuge
OD_{660}	Optical density at 660 nm; measure of cell concentration
P	Product concentration, g/m^3
Q_{AGIT}	Agitation heat input, W
Q_{EVAP}	Heat loss by evaporation, W
Q_{EXCH}	Heat removal by heat exchanger, W
Q_{LOSS}	Heat loss to surroundings, W
Q_{MET}	Metabolic heat generation, W
q_{O_2}	Specific oxygen uptake rate, g O_2/g DCW ∗ h
q_P	Specific product formation rate, g/g DCW ∗ h
R	Ideal gas constant, J/mol ∗ K
r_2, r_1	Outer and inner radii for centrifuge, m
R_{BATCH}	Biomass productivity for batch fermentation, g DCW/m^3 × h
R_{CSTR}	Biomass productivity for CSTR, g DCW/m^3 × h
S	Substrate concentration, g/m^3
S_0	Inlet substrate concentration, g/m^3
T	Bulk fluid temperature, K
T_C	Coolant temperature, K
t_{CYCLE}	Batch cycle time, h
t_M	Mixing time, h
$t_{TURNAROUND}$	Lumped batch turnaround time, h
U	Overall heat transfer coefficient, W/m^2 × K
u	Particle velocity, m/h
u_{lim}	Limiting particle velocity for a separator, m/h
V	Fluid volume, m^3
v	Reaction rate,
v_{max}	Maximum reaction rate
w_{H_2O}	Mass flow rate of water, kg/h
X	Cell concentration, g DCW/m^3
X_0	Inlet cell concentration, g DCW/m^3
Xi	Inoculum cell concentration, g DCW/m^3
X_M	Cell concentration associated with D_M, g DCW/m^3
Y_{Hi}	Cell mass produced per heat evolved, g DCW/kcal
$Y_{P/S}$	Product yield on substrate, g/g
Y_{X/O_2}	Cell yield on oxygen, g DCW/g O_2

$Y_{X/S}$	Cell yield on substrate, g DCW/g
Z	Concentration of contaminant microorganism, g DCW/m^3

Greek Symbols

μ	Specific growth rate, h^{-1}
α	Specific rate of cell lysis or endogenous metabolism, h^{-1}
λ	Adjustable parameter
τ	Residence time, h
Σ	Σ Factor, area equivalent for a centrifuge
ω	Angular velocity
α	Recycle ratio
ρ	Density, kg/m^3
η	Viscosity, kg/m $\times$ s
λ	Empirical function of ϵ_P
α, β	Product formation parameters, g product/g DCW and g product/g DCW $\times$ h, respectively
μ_{max}	Maximum specific growth rate, h^{-1}
ϵ_P	Volume fraction of particles
δ_S	Fractional substrate conversion

BIBLIOGRAPHY

1. J. Monod, *Ann. Rev. Microbiol.* **3**, 371 (1949).
2. A. Novick and L. Szilard, *Proc. Nat. Acad. Sci. Wash.* **36**, 708 (1950).
3. D. Herbert, R. Elsworth, and R.C. Telling, *J. Gen. Microbiol.* **14**, 601–622 (1956).
4. M.L. Shuler and M.M. Domach, in H.W. Blanch, E.T. Papoutsakis, G. Stephanopoulos eds., *Foundations in Biochemical Engineering*, American Chemical Society, Washington, D.C., 1983, p. 93.
5. J. Nielsen, *Biotechnol. Bioeng.* **41**, 715–727 (1993).
6. A. Harder and J.A. Roels, *Adv. Biochem. Eng.* **21**, 55–107 (1982).
7. J. Nielsen and J. Villadsen, *Chem. Eng. Sci.* **47**, 4225–4270 (1992).
8. N.M. Fish and M.D. Lilly, *Biotechnology* **2**, 623–627 (1984).
9. R. Luedeking and E.L. Piret, *J. Biochem. Microbiol. Tech. Eng.* **1**, 431–459 (1959).
10. R.D. Kiss and G. Stephanopoulos, *Biotechnol. Bioeng.* **39**, 565–574 (1992).
11. P. Yu and K.-Y. San, *Biotechnol. Prog.* **9**, 587–593 (1993).
12. G.K. Raju and C.L. Cooney, in H.-J. Rehm, G. Reed, A. Puhler, P. Stadler eds., *Biotechnology*, 2nd ed., vol. 3, VCH, Weinheim, Germany, 1993, pp. 159–184.
13. M.F. Edwards and W.L. Wilkinson, *Chem. Eng.* **265**, 310 (1972).
14. M. Charles, *Trends Biotechnol.* **3**, 134–139 (1985).
15. C.L. Cooney, D.I.C. Wang, and R.I. Matales, *Biotechnol. Bioeng.* **11**, 269–281 (1968).
16. B.J. Abbott and A. Clamen, *Biotechnol. Bioeng.* **15**, 117–127 (1973).
17. B.L. Maiorella, H.W. Blanch, and C.R. Wilke, *Biotechnol. Bioeng.* **26**, 1003–1025 (1984).
18. A.P.J. Sweere, L. Mesters, L. Janse, K.Ch.A.M. Luyben, and N.W.F. Kossen, *Biotechnol. Bioeng.* **10**, 567–578 (1988).
19. F. Vardar and M. Lilly, *Eur. J. Appl. Microbiol. Biotechnol.* **14**, 203–211 (1982).
20. C.G. Sinclair and D.E. Brown, *Biotechnol. Bioeng.* **12**, 1001–1017 (1970).
21. P.J. Senior and J. Windass, *Biotechnol. Lett.* **2**, 205–210 (1980).
22. O. Levenspiel, *Chemical Reaction Engineering*, 2nd ed., Wiley, New York, 1972, pp. 253–325.
23. M. Charles, in T.K. Ghose, A. Fiechter, N. Blakebrough eds., *Advances in Biochemical Engineering*, vol. 8, Springer-Verlag, New York, 1978, pp. 1–62.
24. N.M.G. Oosterhuis and N.W.F. Kossen, *Biotechnol. Bioeng.* **26**, 546–550 (1984).
25. J.A. Howell, C.T. Chi, and U. Pawlowsky, *Biotechnol. Bioeng.* **14**, 253–265 (1972).
26. R.C. Righelato and S.J. Pirt, *J. Appl. Bacteriol.* **30**, 246–250 (1967).
27. D. Herbert, in *Continuous Culture of Microorganisms*, Monograph no. 12, Society of Chemistry and Industry, London, 1961, pp. 21–53.
28. H. Kuhn, U. Friederich, and A. Fiechter, *Eur. J. Appl. Microbiol.* **6**, 341–349 (1979).
29. K.J. Lee, D.E. Tribe, and P.L. Rogers, *Biotechnol. Lett.* **1**, 421–426 (1979).
30. J. Fu, D.B. Wilson, and M.L. Shuler, *Biotechnol. Bioeng.* **41**, 937–946 (1993).
31. S.W. Brown and S.G. Oliver, *Eur. J. Microbiol. Biotechnol.* **16**, 119–122 (1982).
32. D.E. Dykhuizen and D.L. Hartl, *Microbiol. Rev.* **47**, 150–168 (1983).
33. M. Nagashima, in H. Verachtert, and R.D. Mot eds., *Yeast: Biotechnology and Biocatalysis*, Dekker, New York, 1990, pp. 57–84.
34. R.H. Davis and C.S. Parnham, *Biotechnol. Bioeng.* **33**, 767–776 (1989).
35. K.L. Henry and R.H. Davis, *Biotechnol. Prog.* **6**, 7–12 (1990).
36. K.L. Ogden and R.H. Davis, *Biotechnol. Bioeng.* **37**, 325–333 (1991).
37. S.W. Carleysmith and M.D. Lilly, *Biotechnol. Bioeng.* **21**, 1057–1073 (1979).
38. L.Y. Ho and A.E. Humphrey, *Biotechnol. Bioeng.* **12**, 291–311 (1970).
39. S.H. Park, S.B. Lee, and D.Y. Ryu, *Biotechnol. Bioeng.* **23**, 1237–1254 (1981).
40. C. Kim, H.S. Kim, and D.D. Ryu, *Biotechnol. Bioeng.* **24**, 1889–1896 (1982).
41. M. Moo-Young, *Can. J. Chem. Eng.* **53**, 113–118 (1975).
42. R.C. Righelato, A.P.J. Trinci, S.J. Pirt, and A. Peat, *J. Gen. Microbiol.* **50**, 399–412 (1968).
43. M. Moo-Young and H.W. Blanch, *Adv. Biochem. Eng.* **19**, 1–69 (1981).
44. R.J. Fleischaker and A.J. Sinskey, *Eur. J. Appl. Microbiol. Biotechnol.* **12**, 193–197 (1981).
45. S.J. Pirt and W.M. Kurowski, *J. Gen. Microbiol.* **63**, 357–366 (1970).

46. G.R. Cysewski and C.R. Wilke, *Biotechnol. Bioeng.* **18**, 1297–1313 (1976).
47. G.R. Cysewski and C.R. Wilke, *Biotechnol. Bioeng.* **19**, 1125–1143 (1977).
48. P.L. Rogers, K.J. Lee, M.L. Skotnicki, and D.E. Tribe, *Adv. Biochem. Eng.* **23**, 37–84 (1982).
49. J.E. Bailey and D.F. Ollis, *Biochemical Engineering Fundamentals*, 2nd ed., McGraw-Hill, New York, 1986, pp. 733–734.
50. T.K. Ghose and R.D. Tyagi, *Biotechnol. Bioeng.* **21**, 1387–1400 (1979).
51. D. Gold, A. Mohagheghi, C.L. Cooney, and D.I.C. Wang, *Biotechnol. Bioeng.* **23**, 2105–2116 (1981).
52. A. Margaritis and C. Wilke, *Biotechnol. Bioeng.* **20**, 727–754 (1978).
53. H.N. Chang, W.G. Lee, and B.S. Kim, *Biotechnol. Bioeng.* **41**, 677–681 (1993).
54. Y.L. Lee and H.N. Chang, *Biotechnol. Bioeng.* **36**, 330–337 (1990).
55. H.N. Chang, I.-K. Yoo, and B.S. Kim, *Biotechnol. Adv.* **12**, 467–487 (1994).
56. R. Siegel and D.D.Y. Ryu, *Biotechnol. Bioeng.* **27**, 28–33 (1985).
57. S.B. Lee, D.D.Y. Ryu, R. Seigel, and S.H. Park, *Biotechnol. Bioeng.* **31**, 805–820 (1988).

See also BIOREACTORS, AIR-LIFT REACTORS; BIOREACTORS, FLUIDIZED-BED; FERMENTATION MONITORING, DESIGN AND OPTIMIZATION; MASS TRANSFER; SCALE-UP, STIRRED TANK REACTORS.

BIOREACTORS, FLUIDIZED-BED

FRANCESC GÒDIA
CARLES SOLÀ
Universitat Autònoma de Barcelona
Barcelona, Spain

OUTLINE

KEY WORDS

Bioreactors
Fluidized-bed
Hydrodynamics
Immobilized biocatalysts

INTRODUCTION

Fluidized-bed bioreactors are directly linked to the use of biocatalysts (cells or enzymes) for transformations in an immobilized form. The solid particles of the immobilized biocatalyst are maintained in fluidization by means of the circulation of a fluid phase (either liquid, gas, or a mixture of both) that compensates their weight. In this way, good liquid mixing and mass transfer between the solid and the liquid phases can be obtained with low attrition. Also, fluidized-bed bioreactors can accommodate a gas phase and can be used to feed solids in suspension. High productivities can be achieved in these systems, but their hydrodynamic complexity and operational stability have to be well defined for a proper operation.

THE FLUIDIZATION CONCEPT: GENERAL CONSIDERATIONS

The term *fluidized-bed* is used to define those physical systems composed of a solid phase in the form of individual particles that move within a fluid phase and are not in continuous contact with each other. Fluidization of the solid particles is reached when the flow of fluid through the bed is high enough to compensate their weight. On the other hand, in order to be kept in the fluidized-bed reactor and not be washed out (elutriated), the superficial velocity of the fluid in the bed (that is, the ratio between the flow rate and the bed cross-sectional area) has to be lower than the settling velocity of the particles. These two extreme situations are outlined in Figure 1. When the flow rate of a fluid through a packed bed of solid particles steadily increases, the pressure drop increases proportionally to the flow rate, as long as the bed height remains constant. When the drag force of the fluid equilibrates the weight of

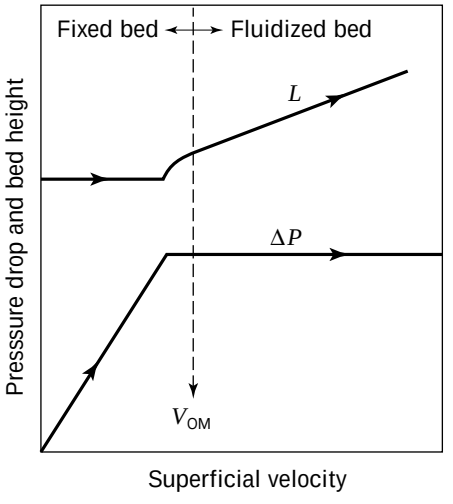


Figure 1. Illustration of the basic concept of fluidization. Variation of the pressure drop (ΔP) and the height of the solid particles in a bed (L), with increasing fluid superficial velocity. v_{OM}, minimum fluidization velocity.

the particles, the bed starts to expand and, after a transition period, reaches fully developed fluidization. At this point, further increments in the flow rate do not produce an increase in pressure drop, but instead lead to an increase of the height occupied by the solid particles in the reactor. If the flow rate is increased significantly, the elutriation of the solid particles occurs when the fluid's superficial velocity is higher than the solid's settling velocity. The fundamentals of the fluidization phenomena are discussed comprehensively in the chemical engineering literature (1,2).

Figure 2 represents the basic scheme of a fluidized-bed bioreactor. Although various configurations are possible (3), the most extensively used is the gas–liquid cocurrent up-flow reactor. In it, liquid usually comprises the continuous phase and is fed from the reactor bottom. Its flow upward in the reactor promotes fluidization of the solid particles. Usually, the reactor will have two or three phases. In addition to the liquid and solid phases, the occurrence of a gas phase is quite common in those systems using cells as biocatalysts, either for aeration requirements (in which case, an air or oxygen stream is fed to the reactor, as shown in Fig. 2) or because cell metabolism produces a gas product (for example, CO_2, CH_4). In systems using enzymes as biocatalysts, the most common situation is two-phase fluidization, without any gas phase. Very often, due to the low reaction rates of most biological transformations, long liquid residence times are needed for the completion of the reaction, and therefore the drag force created by the low liquid flow rate in a single pass reactor is not enough to promote fluidization of the solid particles. Fluidization is obtained either by external liquid recirculation or by the gas loaded to the reactor, as depicted in Figure 2. In systems where a gas is produced by cell metabolism, the gas can also be an additional factor contributing to solid particle fluidization, although other effects are also observed in this case, such as internal liquid recirculation patterns. Fluidization at relatively low liquid flow rates is also favored in tapered fluidized-bed configurations; the liquid superficial velocity at the bottom of the reactor is higher due to the reduced cross-sectional area. In general, one can distinguish three main sections in fluidized-bed bioreactors: *(1)* the bottom section, where feed (liquid, gas, or both) and recirculation are provided; *(2)* the central main section, where most of the reaction takes place; *(3)* and the top section, with a wider diameter that serves to decelerate the movement of the particles by decreasing the superficial velocity of the liquid, thus enhancing the retention of the solid phase and at the same time allowing gas disengagement from the liquid phase. It is a common trend for fluidized-bed bioreactors to use biocatalysts, either cells or enzymes, in the form of immobilized preparations. In general, the particles can be of three different types: *(1)* inert cores on which a biofilm is created by cell attachment, or in the case of enzymes, by adsorption or covalent binding immobilization; *(2)* porous particles in which the biocatalysts are entrapped; *(3)* cell aggregates obtained by self-immobilization caused by the ability of some cell strains to form flocs, pellets, or aggregates. Fluidized-bed bioreactors are usually differentiated from air-lift bioreactors by the fact that the latter do not specifically require the use of immobilized biocatalysts. Indeed, they were developed for free cell suspensions. In addition, air-lift bioreactors have different compartments, created by physical internal divisions, with different degrees of aeration.

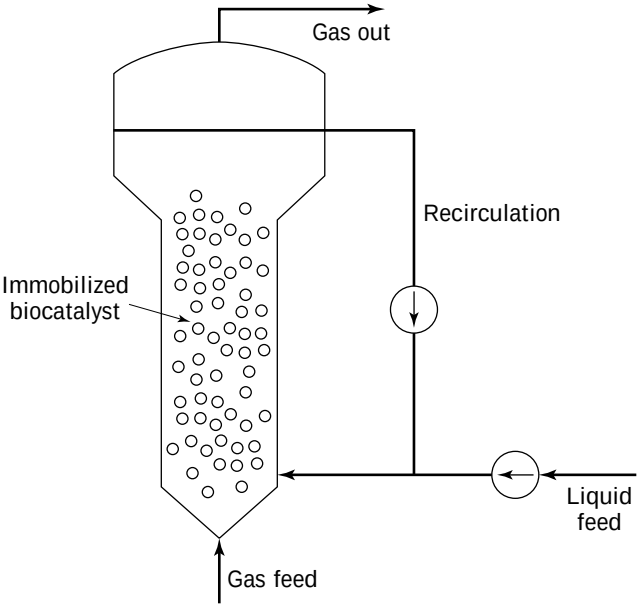


Figure 2. General scheme of a fluidized-bed reactor.

CHARACTERISTICS AND POTENTIAL OF FLUIDIZED-BED BIOREACTORS

The use of fluidized-bed bioreactors can provide a number of advantages that makes them an interesting alternative for bioprocesses, especially for continuous operation. In comparison with conventional mechanically stirred bioreactors, fluidized-bed bioreactors provide a much lower attrition of solid particles, and almost any kind of immobilized biocatalyst preparation can be used without physical disruption. Biocatalyst concentration can be significantly higher because of immobilization, and the typical wash-out limitations of continuous bioreactors operating with free cells are overcome because the solid particles are physically retained in the reactor vessel; operation at flow rates higher than the maximum growth rate of the cells can be achieved. As a consequence, the final productivity of the bioreactor can be increased substantially. Compared with packed-bed reactors, fluidized-bed bioreactors can be operated with smaller-size particles and without the drawbacks of clogging, creation of preferential flowpaths, or particle compression due to bed weight. Moreover, the smaller particle size minimizes the internal diffusional resistances, and the higher level of mixing enhances external mass and heat transfer from liquid to solid phase.

The degree of internal mixing in a fluidized bed can vary to a great extent, and it depends on various factors: density

and diameter of the solid particles, liquid and gas flow rates at the reactor inlet, endogenous gas production by cell metabolism, and recirculation rate. For some operational conditions, high gas, liquid, or recirculation flow rates, the internal mixing will be very high, and the reactor will approach the behavior of a complete mixed tank. On the other hand, operation with low gas, liquid, or recirculation rates will provide a flow pattern close to plug flow, with some degree of axial mixing. This implies that reaction kinetics is an important factor to consider in the analysis and design of fluidized-bed bioreactors. Configurations favoring liquid mixing will be more appropriate for substrate-inhibited reactions, and configurations approaching plug flow will be indicated for product-inhibited reactions. Another advantage of fluidized-bed bioreactors is the ease of separation of the gas produced in most transformations involving cells (i.e., CO_2), or the feed of a gas stream to the reactor, for example, for aeration purposes. Also, fluidized-bed reactors make biocatalyst replacement easy, without disruption of the operation, enabling good control of the overall activity of the reactor. For example, one may replace particles with deactivated enzyme or removing an excess of biomass created by biofilms. On the other hand, solids attrition is higher in fluidized-bed than in packed-bed bioreactors. From the productivity point of view, the advantages of the fluidized bed, especially with respect to mass transfer rates, make it possible to obtain higher levels of overall productivity than in packed-bed reactors, in spite of the fact that the fraction of immobilized biocatalyst particles is lower for a fluidized bed.

Fluidized-bed bioreactors are complex with regard to hydrodynamic aspects, especially taking into account that the properties of the biocatalyst particles may change considerably during the operation time and in the presence of three different phases (solid, liquid, gas) in many cases. In fact, the nature of the particles (for example, their density and size, or their evolution with time, which are especially important with respect to certain kinds of immobilized cells), the liquid and gas flow rates employed, the type of reaction kinetics, and the kinetics of cell growth or enzyme deactivation influence each other and have a direct effect on the reactor design and performance (4).

Using the biocatalyst in immobilized form also contributes to the complexity of a fluidized-bed bioreactor. The behavior of the immobilized biocatalyst, especially when cells are used, can be substantially different than that of free suspensions (5). The behavior must be determined at the kinetic level, the physiological level, and the genetic level, and the biocatalyst's relationship with the diffusional restrictions in the particles and the possible direct effects associated with the immobilization itself must be well understood and correctly described in order to build appropriate and reliable models of for reactor design, control, and scaling up.

The potential of fluidized-bed bioreactors can be further exploited by considering multistage units and using two solid particles with different properties. Figure 3 gives an example of the concept of multistage operation, in particular, a countercurrent multistage fluidized bed working with immobilized enzymes (6). The main characteristic of this bioreactor is the continuous transport of the solid particles of immobilized enzymes from one stage to another in a downward direction. The overall catalytic activity of the reactor remains constant as the exhausted enzyme is removed from the reactor bottom stage while fresh biocatalyst is added at the top stage. A second advantage of dividing the reactor into compartments is the very low degree of back-mixing of the biocatalyst, and the plug-flow regime attained in the liquid phase. The use of two solid particles with different properties, particularly with different densities, can be used in a fluidized bed to achieve the in situ separation of a product of the reaction; this is a clear advantage in systems with product inhibition or when unfavorable thermodynamic equilibria limit the conversion rate for a reaction. For example, Davison and Scott (7) have proposed a system based in two different types of particles with different densities. As one type of particle, containing the biocatalysts (in this particular example, cells of *Lactobacillus delbreuckii*), remains fluidized in the bioreactor, the second type, which is heavier and contains no cells, is introduced from the top of the bioreactor and collected at the bottom. This second type of particle is selected to selectively remove the inhibitory product of the fermentation, for example, lactic acid. Another possibility is to combine both approaches, that is, to design a multistage fluidized bed working with two types of particles to achieve a selective in situ removal of the product. Van der Wielen et al. (8) have used this approach to enhance the enzymatic deacylation of benzylpenicillin, providing the desired product, 6-aminopenicillanic acid, and as a byproduct, phenylacetic acid, by means of light particles of immobilized enzymes and heavy particles of an adsorbent of the acidic by-product. Figure 4 presents a scheme of two different possibilities for the design of a two-solid-phase fluidized-bed bioreactor: semicontinous multistage pulsed flow and continuous trickle flow. In the first, the movement of the heavier solid particles downward in the reactor is obtained by periodic pulsations; in the second, the denser particles move continuously from the top to the bottom of the reactor.

MAIN ASPECTS FOR THE DESIGN AND OPERATION OF FLUIDIZED-BED BIOREACTORS

As already described, the design and operation of a fluidized-bed bioreactor has to take into account a number of aspects regarding both physical characteristics and reaction performance. These aspects include the hydrodynamics of the bioreactor (usually with three phases), characterization of the type of flow, degree of mixing for each phase and the heat transfer, mass transfer between phases, mass balancing of the species taking part in the reaction, intrinsic kinetics of the immobilized biocatalyst, diffusion transport of the species inside the solid particles, enzyme deactivation, and cell growth. Accurate knowledge concerning these points will allow complete characterization of a fluidized-bed reactor and the development of reliable mathematical models describing its performance. Next, we discuss the conceptual aspects intrinsically related to fluidized beds. For more detailed information, including mathematical formulation of the models and their

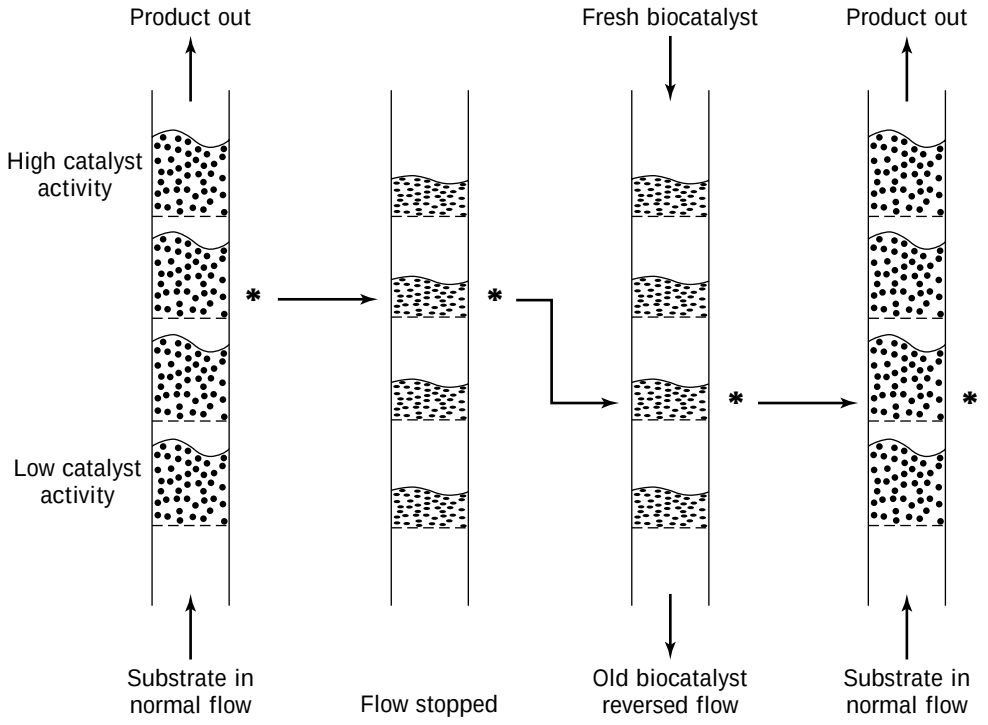


Figure 3. Operation phases of a multistage fluidized-bed reactor for a deactivating biocatalyst *Source*: From Ref. 6.

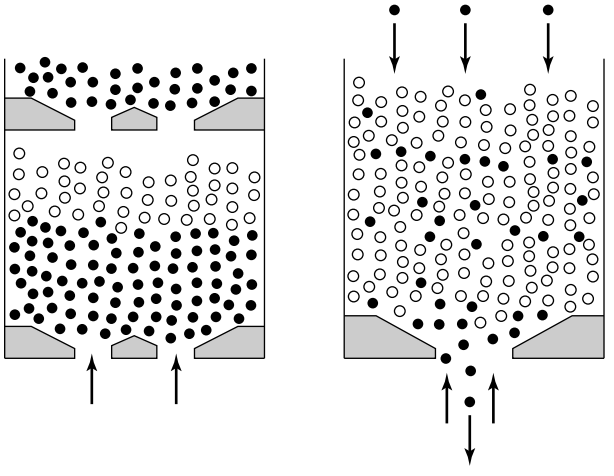


Figure 4. Schemes corresponding to two types of operation of a fluidized-bed as countercurrent adsorptive reactors, with two types of solid particles: light/small immobilized biocatalyst (○) and dense/large sorbent (●). Left: one compartment of the semicontinuous, multistage pulsed flow. Right: continuous trickle flow mode. *Source*: From Ref. 8.

application, the reader is referred to Refs. 6, 9–12 as well as general reviews in Refs. 13 and 14.

The first step in hydrodynamic characterization requires a determination of the fraction of the total volume to be occupied by the different phases (gas, liquid, solid). This is known as *phase hold-up* and can be determined using different techniques (3,15). Strictly speaking, phase hold-up may not be uniform for the whole reactor, especially with low axial dispersion. However, in systems with appreciable mixing, phase hold-up can often be considered uniform.

With respect to mixing behavior, most attention is generally paid to characterizing the liquid phase because this is, in most cases, the continuous phase in bioreactors, that is, the one in which substrates are fed and products of the reaction accumulate. For a three-phase fluidized-bed bioreactor with cocurrent up-flow circulation of liquid and gas, the type of flow pattern is very much dictated by the value and the ratio of the liquid and gas superficial velocities in the reactor (3,16). Three main regimes are possible: *(1)* the dispersed flow regime occurs at high ratios of liquid velocity to the gas velocity and is characterized by the homogeneous dispersion of small gas bubbles in the liquid; *(2)* the coalesced bubble flow regime occurs at increased gas flow rates and is characterized by the formation of bigger bubbles as a result of the coalescence of smaller ones; the coalesced bubbles have a nonuniform distribution in the liquid; *(3)* the slug-flow regime is the consequence of a further increase of gas flow rates, occurs at high ratios of gas velocity to liquid velocity, and is characterized by the formation of large gas bubbles that, in small-diameter bioreactors, tend to completely occupy their cross-sectional area. Slug formation breaks the bed continuity and causes great instability. Figure 5 gives an example of experimental data in a fluidization chart, showing these different regimes. If the gas flow rate is increased dramatically, it will eventually become the continuous phase in the bioreactor. A very detailed study on the characteristics of these dif-

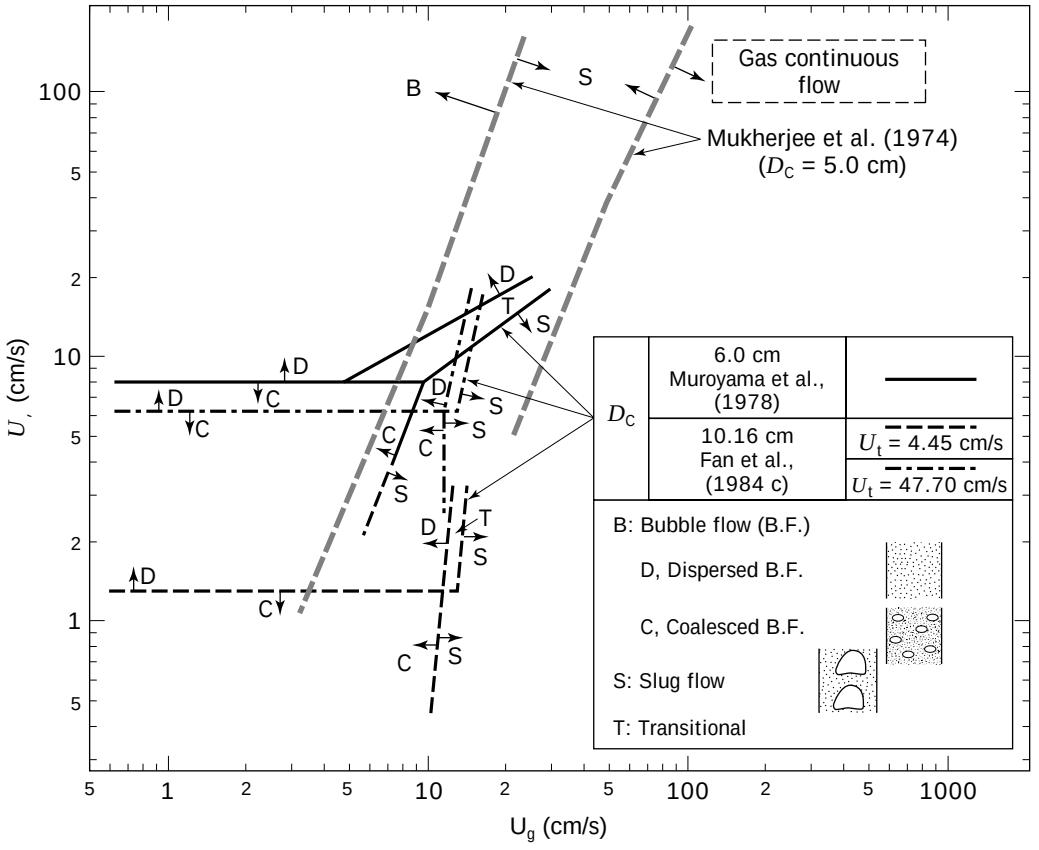


Figure 5. Flow regime diagram for the concurrent gas–liquid–solid fluidized-bed, U_g, gas superficial velocity. U_l, liquid superficial velocity. *Source*: From Ref. 3.

ferent regimes and the effect of the density and size of the solid particles on the regime transitions has been conducted by Zhang et al. (17). The liquid flow pattern in the bioreactor is directly influenced by the degree of mixing associated with these different regimes, as well as by other factors such as internal gas generation, bead size distribution, and external liquid recirculation. Basically, the flow will approach plug flow in systems with high velocity in the liquid phase. Quite often, bioreactors experience different degrees of mixing, and complete mixed flow can be observed in fluidized-bed bioreactors.

The influence of the gas phase on the liquid and solid mixing in a fluidized-bed bioreactor has been studied by Gommers et al. (18), considering two extreme situations: a reactor operating without gas and a reactor to which gas was introduced artificially from the bottom. The results showed that the gas phase greatly influences the degree of liquid mixing in the reactor, and that this effect increases sharply with the bed diameter. Another fact to consider is that in systems with gas generation associated with the reaction progress (for example, in most fermentations), the gas flow will change from the bottom to the top of the reactor, proportionally to the substrate consumption, and the axial dispersion will change with reactor height (19). An important aspect to consider in the study of the influence of the gas produced in a fluidized bed on the degree of liquid mixing is that the results will be affected by the experimental system used for gas injection, as discussed in detail by Buffière et al. (20). When liquid recirculation is used to promote fluidization because of the slow reaction rate and the long liquid-residence time required, plug flow is usually disrupted, and completely mixed flow is usually achieved.

Other factors influencing the hydrodynamic behavior of fluidized-bed bioreactors are the properties of the solid and liquid phases. In general terms, the weight and size of the solid particles will directly influence the liquid and gas flow rates required for bed fluidization. If the liquid residence time is fixed by criteria of substrate conversion, for example, then the heavier particles will require higher L/D ratios to increase liquid superficial velocity and enhance fluidization, or as an alternative, they will require high recirculation rates. In absence of recirculation, and when particles with a certain distribution in size and weight are used, solid-particle stratification is commonly found. Under stratification conditions, movement of the solid particles in the bed is very limited, and the particles are ordered by decreasing settling-velocities, from the bottom to the top of the bed. One of the consequences of this situation is that, in liquid plug-flow regimes, the particles at different reactor heights will experience different environments. For cells growing as biofilms around a solid particle, it is typical for the particles at the bottom of the bed to provide better conditions for cell growth (for example, substrate

availability), and as a consequence, they will decrease their overall density and therefore migrate to the upper part of the reactor. Removal of excess biofilm may require external treatment of the particles. The operation of a stratified fluidized bed is illustrated in Figure 6, where the stratification of a bed with activated carbon as a solid support for cell growth is combined with the capacity to absorb the product of the reaction and, therefore, remove it selectively (21). The crucial effect of particle size and density distribution, as well as gas and liquid superficial velocities, on the phenomena of solid-particle stratification is discussed in detail in various recent studies (22–24). In general, solids mixing has been studied less comprehensively than liquid mixing, and it is less well understood. The possibility of using new experimental techniques (25) has enabled researchers to propose new models for the interpretation of the solids mixing, for example, trajectory length distribution (26,27), in contrast to more classical models that assume that solids mixing can be defined by means of an axial dispersion model (28).

Another important aspect to be considered from the hydrodynamic point of view is that fluidization properties depend on the difference in densities of the solid and liquid phases in the fluidized bed, and these values change during the operation time. This situation is of particular relevance for those systems where the absolute values for the liquid and solid densities are relatively close, such as for cells immobilized in natural origin polymers (alginate, agarose, etc.) or self-aggregated cells (pellets, flocs, etc.). In this case, relatively low variations in the absolute value of the solid density (usually associated with cell growth or the accumulation of CO_2 gas in the particles) and the liquid density (usually associated with substrate consumption) may cause an important percentile variation in the density difference between both phases, producing a relevant impact on the reactor hydrodynamics, as the regime changes from dispersed to coalescing (29), and on reactor stability. This problem can be of particular relevance at the start-up of a bioreactor operation, when the density values may change to a greater extent.

A second aspect to be addressed in the characterization of a fluidized-bed reactor is the definition of the flux model. As mentioned previously, most of the attention is focused on the liquid phase, and the flux model is closely connected to the hydrodynamic conditions in the reactor that will generate a given degree of internal mixing, which is somewhere between the two extreme situations of perfectly mixed or plug flow. The determination of the real liquid flux model in the bioreactor is a necessary step for the application of the mass balance equations for the species taking part in the reaction. Stimulus–response techniques are commonly used for such a purpose and are based on the introduction of an inert tracer at the reactor inlet and the analysis of the response curve obtained at the outlet, which reflects the type of flux (30). The models that describe the liquid flux in a real reactor (31) can be *(1)* an axial dispersion model, in which axial dispersion is superimposed on the liquid convective flux; *(2)* a tank-in-series model, in which the bioreactor is considered as a series of CSTR reactors of the same volume; and *(3)* a compartmented model, in which the flux model in the bioreactor is described as the combination of different ideal compartments. For fluidized-bed bioreactors, especially when fermentation gas is produced, some interesting contributions have been proposed, such as the consideration of a variable dispersion coefficient, which increases its value in proportion to the fermentation gas generated in the reactor, up to a given point (10).

In addition to hydrodynamic and mixing characteristics, another relevant aspect of a fluidized-bed bioreactor is its biocatalytic activity. Because immobilized biocatalysts are always used, the definition of the problem must simultaneously include the reaction characteristics (kinetics, stoichiometry, equilibrium) and transport characteristics, including both the transport between the liquid and the solid phase and the transport within the solid phase. When the general situation of a solid biocatalyst particle immersed in a liquid phase, as shown in Figure 7, is studied, different phenomena have to be considered. First, substrate is transported from the liquid phase to the solid ex-

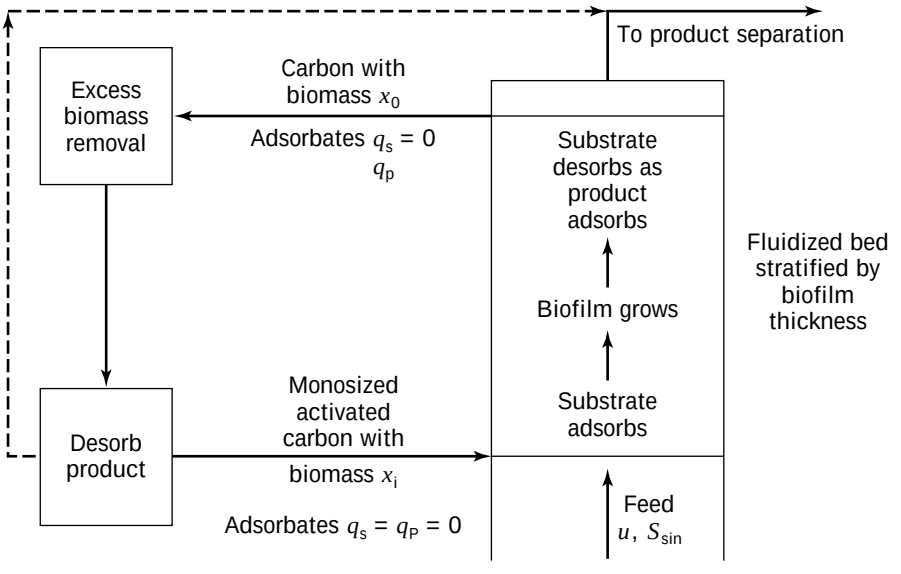


Figure 6. Operation diagram of a fluidized-bed bioreactor with simultaneous bioconversion and adsorption/desorption of substrate and product. *Source*: From Ref. 21.

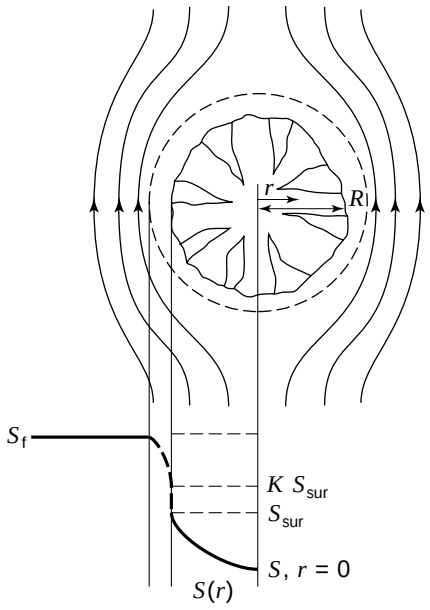


Figure 7. Inter and intraparticle mass transfer of a single porous spherical bead of radius R. Substrate concentration profiles across the stagnant liquid film and inside the solid particle, $S(r)$. S_f, concentration on the bulk liquid; S_{sur}, concentration on the solid surface; K, partition coefficient. *Source*: From Ref. 11.

ternal surface. This external mass transfer is usually modeled by means of a hypothetical liquid film that creates resistance to the transport, and it is characterized by the external mass transfer coefficient. The value of the coefficient, and thus of the external mass transfer velocity, depends on the physical properties of the liquid and the superficial velocity of the liquid around the solid; thus beds with more mixing will present better conditions for external mass transfer. When external mass transfer resistance is negligible, the concentration at the surface of the particle is the same as in the fluid phase ($S_f = S_{sur}$, according to Figure 7). An additional possibility is the occurrence of partitioning phenomena in the substrate, between the particles and the liquid, as a result of the material properties. In that case, the ratio between the substrate concentration at both sides of the solid–liquid interface will be given by the partition coefficient, K. Equally, partitioning phenomena can be determined for the product. Once in the solid particle, the substrate will diffuse inside the solid, following Fick's law, and simultaneously the reaction will take place. Therefore the corresponding equations of kinetic reaction and diffusion in the solid have to be solved simultaneously (10,11) in order to obtain the internal concentration profiles. To define this part correctly, the values for the intrinsic kinetic parameters of the reaction (that is, in absence of mass transfer limitations) and the effective diffusivities for the substrate and the product of the reaction in the solid have to be known. Moreover, the substrate conversion implies product generation, and this can affect the kinetics (for example, by product inhibition); usually the diffusion and reaction analysis is made simultaneously for both substrate and product. The overall activity of the biocatalytic particles will be dictated by the relative velocities of the two phenomena taking place inside them: reaction and diffusion. Systems with low diffusion rates with respect to the reaction rates will be diffusion controlled; on the other hand, systems with high diffusion rates with respect to the reaction rate will be controlled by the reaction. In the first case, the low diffusion will limit the efficiency of the particle because the reaction potential of the immobilized biocatalyst will not be fully used. Usually, this is reflected in terms of the efficiency factor, which is defined as the ratio between the actual reaction rate occurring in the system, and the reaction rate that would occur if no diffusional limitations existed, that is, when all solid particles would present a uniform concentration, equal to that of their surface. The relationship between the effectiveness factor and some moduli, such as the observable modulus, being proportional to the ratio between reaction and diffusion rates, is given in Figure 8. One of the interesting properties of such graphs is that they are very similar for different types of geometries and kinetics. Therefore they enable a direct analysis of the degree of diffusional limitations in a given type of particle in a fluidized bed and also suggest quantitative modifications to be performed to avoid such limitations, for example, changes in the particle size or diffusion conditions. In general terms, fluidized beds are interesting with regard to these aspects because since they require comparatively small-diameter particles for better fluidization, the potential diffusion limitations are reduced.

As mentioned at the beginning of this section, the combination of the different aspects we discussed allows the elaboration of mathematical models describing the behavior of fluidized-bed bioreactors. The reliability of these models depends greatly on the accuracy of the determination of the various parameters involved in the definition of the reactor and the reaction system (flux model, intrinsic kinetic parameters, external mass and heat transfer coefficients, internal effective diffusivities), and they will serve for various purposes such as conceptualization and understanding of the bioreactor itself (an important aspect required for building the model), design of similar bioreactors (especially in the case of scaling-up), simulation of the bioreactor operation at different conditions, and bioreactor control. As an example of how the output of a mathematical model can describe the internal concentration profiles in a fluidized-bed bioreactor, the data corresponding to a continuous fermenter for the production of ethanol from glucose by the bacteria *Zymomonas mobilis* immobilized in carrageenan beads (10) are given in Figure 9.

INDUSTRIAL APPLICATIONS OF FLUIDIZED-BED BIOREACTORS

Fluidized-bed bioreactors have been studied in a wide spectrum of applications as a consequence of the different advantages that they offer (14,33). Nevertheless, this interest has been translated with relatively low intensity into the industrial scale of operation. Possible reasons for this may be the greater complexity of the reactor compared with conventional types (especially regarding long-term

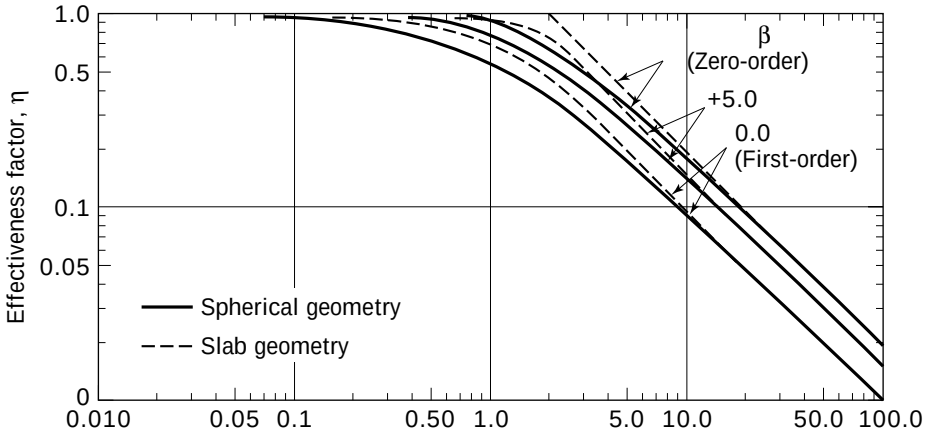


Figure 8. Variation of the effectiveness factor values for immobilized enzymes with Michaelis–Menten intrinsic kinetics with respect to the observable modulus, Φ. *Source*: From Ref. 32.

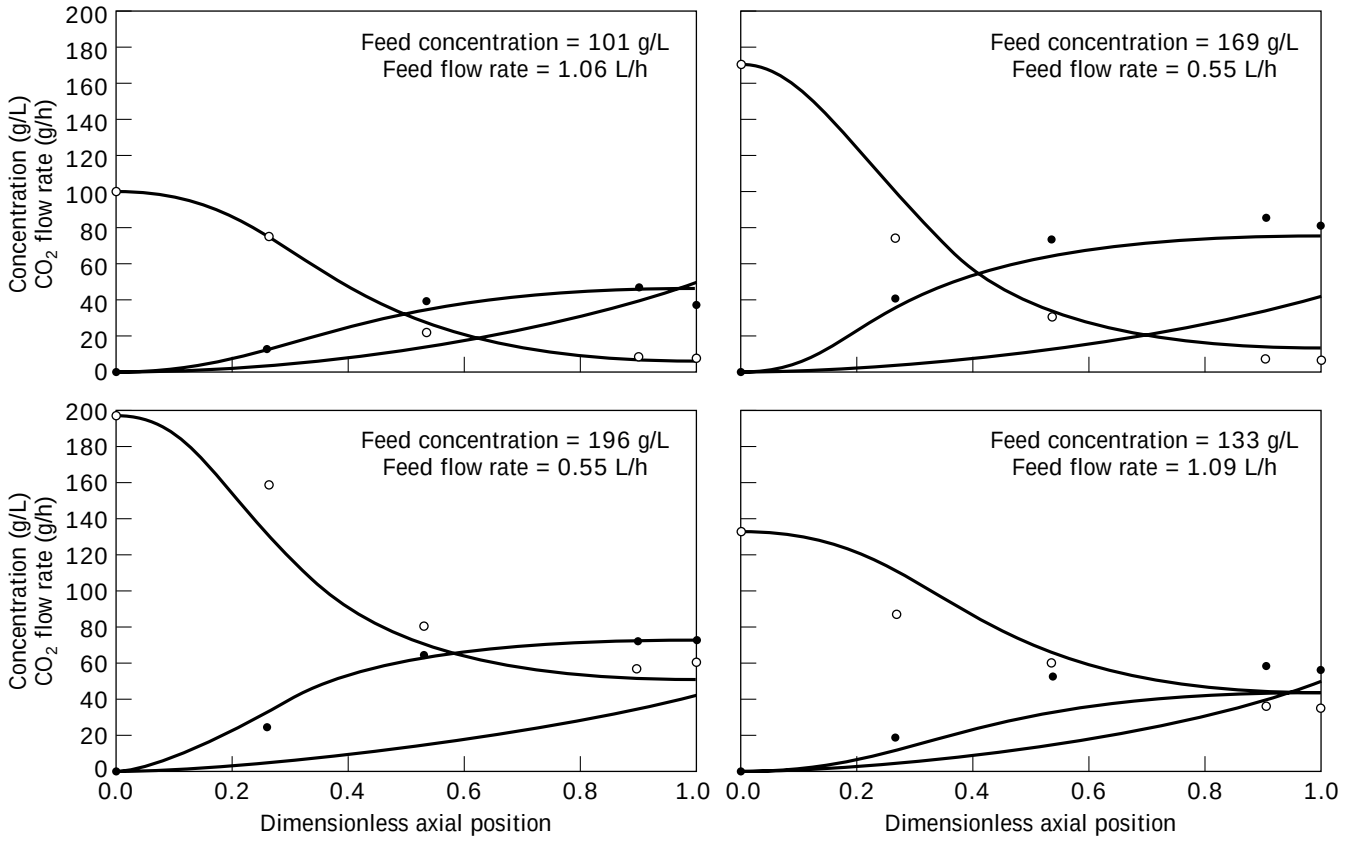


Figure 9. Results of the mathematical model for a tapered fluidized-bed producing ethanol from glucose with *Zymomonas mobilis* cells immobilized in carrageenan beads. Internal concentration profiles with the dimensionless fermenter height: experimental values (points) and values predicted by the model (solid lines). Dotted lines represent the calculated CO_2 flow rate produced by fermentation. *Source*: From Ref. 10.

operation, hydrodynamic behavior and evolution of the immobilized biocatalyst), and the time required in any industrial process to implement any change in technology. Some specific examples from different biotechnological processes illustrate the feasibility of the fluidized bed at a pilot plant and on the industrial scale. The area of wastewater treatment provides examples of high-volume operation and has made use of three-phase fluidized-bed bioreactors, as well as other related designs (such as the up-flow anaerobic sludge blanket reactor), to a significant extent. The wastewater from the baker's yeast industry, particularly from the Gist-brocades company in Delft,

Figure 10. View of an industrial anaerobic wastewater treatment plant based on the operation of fluidized-bed reactors, Gist-brocades, Delft, The Netherlands. *Source*: From Ref. 34.

Netherlands, has been treated with two anaerobic fluidized-bed reactors in series, each one 21 m high and with 390 m^3 total volume, operating with bacterial cells immobilized by attachment onto 0.25- to 0.5-mm-diameter sand particles, a support that is typical for other fluidized-bed reactors applied to anaerobic and aerobic wastewater treatment (34). The volume of the reactor employed for the disengagement of the CH_4 gas obtained was important, and the fluidized-bed volume was 215 m^3. The carbon–oxygen demand (COD) conversion of these reactors is 22 kg/(m^3 day), which is 65% of the COD in the load. A first set of these reactors started operation in 1984, and a second set in 1986. A picture of these industrial units is given in Figure 10. More recently, data from two full-scale an-

aerobic three-phase fluidized-bed plants have also been reported (33). A 700-m^3-total-volume reactor (500-m^3 working volume) has been installed for wastewater treatment at a sugar beet factory of United Northern Sugar Factories Co. in Clauen, Germany, using cells immobilized onto pumice and having COD conversion of 20 kg/(m^3 day). A 734-m^3-total-volume multistage reactor (650- to 700-m^3 working volume) has been installed for wastewater treatment at a yeast factory of Hamburg Co., in Hamburg, Germany, using free cells and a new principle that allows control, in each stage, of the ratio of the recirculated gas amount with respect to the rising gas flow. This enables optimization of the hydrodynamic conditions in each stage, according to the reactor loading. Nitrification of the effluent from a municipal wastewater treatment plant in order to eliminate ammonium salts has also been carried out in large-scale fluidized-bed reactors (35) with a cross-sectional area of 2.2 m^2 and a height of 5 m, using 5 tons of sand particles 0.4–0.6 mm in diameter as support particles for cell biofilm development. The reported results show the feasibility of the system, from both the technical and economic points of view.

Ethanol fermentation has been studied intensively in fluidized-bed reactors and brought up to the pilot-plant level in several cases. A pilot plant with a volume of 1 m^3, using *Saccharomyces cerevisiae* cells immobilized in carrageenan beads has been successfully operated for half a year (36). As an alternative to yeast, the bacteria *Zymomonas mobilis* can also be used. Very promising results have been reported on the operation of a series of two pilot reactors of 55 L each, under nonsterile conditions (37). The support for the immobilization in this case was a macroporous glass particle, Siran, that allowed cell attachment to the internal surface of the pores. After an initial period of sterile colonization of the beads in a conventional stirred tank, the particles, with a high *Z. mobilis* cell concentration, can be transferred to the fluidized bed, which then can be operated at high dilution rates under nonsterile conditions because all the possible contaminants in the feed (a solution of hydrolyzed B-starch) are washed out from the reactor. Complete conversion of 120 g/L sugar concentration with a 4.25-h residence time can be consistently obtained, with ethanol productivities of 18 g/(L h), calculated with the total reactor volume as a basis. A complete technical and economic evaluation of the use of fluidized-bed reactors for the production of ethanol with *Z. mobilis* has shown promising features (38).

Other examples of the use of fluidized-bed bioreactors on the pilot scale in fermentation processes include the production of alcohol-free beer with immobilized yeast in a 50-L pilot-plant fluidized-bed reactor (39) and the production of penicillin using *Penicillium chrysogenum* immobilized in urethane particles in a 160-L pilot-plant reactor (40). *Penicillium chrysogenum* cells have also been used in different studies with fluidized-bed reactors with celite as the support material (41). Also, the area of animal cell culture provides a number of examples of the application of fluidized-bed technology, taking into account that smaller production volumes are usually required in this case. The use of immobilized cells is particularly required for those cells that are anchorage dependent, and a variety of porous carriers and beads have been developed for this purpose. The production of a human anti-HIV-1 antibody using recombinant Chinese hamster ovary cells grown onto a macroporous polyethylene carrier has been carried out in an 84-L-volume pilot-scale fluidized-bed novel reactor equipped with a low shear stress internal impeller for the recirculation of the liquid (42).

BIBLIOGRAPHY

1. D. Kunii and O. Levenspiel, *Fluidization Engineering*, Robert E. Krieger, Huntington, New York, 1977.
2. L.S. Fan, *Gas–Liquid–Solid Fluidization Engineering*, Butterworths, Stoneham, 1989.
3. K. Muroyama and L.S. Fan, *AIChE J.* **31**, 1–34 (1985).
4. G. F. Andrews and J. Przezdziecki, *Biotechnol. Bioeng.* **28**, 802–810 (1986).
5. J.A.M. de Bont, J. Visser, B. Matiasson, and J. Tramper, *Physiology of Immobilized Cells*, Elsevier, Amsterdam, 1990.
6. H.J. Vos, D.J. Groen, J.M. Potters, and K.Ch.A.M. Luyben, *Biotechnol. Bioeng.* **36**, 367–376 (1990).
7. B.H. Davison and C.D. Scott, *Biotechnol. Bioeng.* **39**, 365–368 (1992).
8. L.A.M. Van der Wielen, P.J. Diepen, A.J.J. Straathof, and K.Ch.A.M. Luyben, *Ann. N.Y. Acad. Sci.* **750**, 482–490 (1995).
9. Y. Sun and S. Furusaki, *J. Ferment. Bioeng.* **69**, 102–110 (1990).
10. J.N. Petersen and B.H. Davison, *Appl. Biochem. Biotechnol.* **28/29**, 685–698 (1991).
11. T. Kiesser, G.A. Oertzen, and W. Bauer, *Chem. Eng. Technol.* **13**, 20–26 (1990).
12. N. Qureshi and I.S. Maddox, *J. Chem. Tech. Biotechnol.* **34/35**, 441–448 (1992).
13. G.F. Andrews, *Biotechnol. Gen. Eng. Rev.* **6**, 151–178 (1988).
14. F. Godia and C. Solà, *Biotechnol. Prog.* **11**, 479–497 (1995).
15. W.M. Ohkawa, A. Maezawa, and S. Uchida, *Chem. Eng. Sci.* **52**, 3941–3947 (1997).
16. C. Zheng, B. Rao, and Y. Feng, *Chem. Eng. Sci.* **43**, 2195–2200 (1988).
17. J.-P. Zhang, J.R. Grace, N. Epstein, and K.S. Lim, *Chem. Eng. Sci.* **52**, 3979–3992 (1997).
18. P.F.J. Gommers, L.P. Christoffels, J.G. Kuenen, and K.Ch.A.M. Luyben, *Appl. Microbiol. Biotechnol.* **25**, 1–7 (1986).
19. B.H. Davison, *Appl. Biochem. Biotechnol.* **20/21**, 449–460 (1989).
20. P. Buffière, Ch. Fonade, and R. Moletta, *Chem. Eng. Sci.* **53**, 617–627 (1998).
21. G.F. Andrews and J.P. Fonta, *Appl. Biochem. Biotechnol.* **20/21**, 375–390 (1989).
22. N. Funamizu and T. Takakuwa, *Chem. Eng. Sci.* **51**, 341–351 (1996).
23. L.A.M. Van der Wielen, M.H.H. Van Dam, and K.Ch.A.M. Luyben, *Chem. Eng. Sci.* **51**, 995–1008 (1996).
24. T. Matsumoto, N. Hidaka, Y. Takebayasi, and S. Morooka, *Chem. Eng. Sci.* **52**, 3961–3970 (1997).
25. J. Reese and L.S. Fan, in J. Chaouki, F. Larachi, and M.P. Dudukovuc eds., *Non-invasive Monitoring of Multiphase Flows*, Elsevier, Amsterdam, 1997.
26. J. Villermaux, *Chem. Eng. Sci.* **51**, 1939–1946 (1996).

27. K. Kiared, F. Larachi, Ch. Guy, and J. Chaouki, *Chem. Eng. Sci.* **52**, 3931–3939 (1997).
28. T.C. Bickle and M.G. Thomas, *Ind. Eng. Chem. Process Des. Dev.* **21**, 377–381 (1982).
29. P. Béjar, C. Casas, F. Gòdia, and C. Solà, *Appl. Biochem. Biotechnol.* **34/35**, 467–475 (1992).
30. O. Levenspiel, *The Chemical Reactor Omnibook*, OSU Book Stores, Corvallis, Ore., 1979.
31. D.E. Swaine and A.J. Dauglis, *Biotechnol. Prog.* **4**, 134–148 (1988).
32. J.E. Bailey and D.F. Ollis, *Biochemical Engineering Fundamentals*, McGraw-Hill, New York, 1986, pp. 213.
33. K. Schügerl, *Chem. Eng. Sci.* **52**, 3661–3668 (1997).
34. J.I. Heijnen, A. Mulder, R. Weltevrede, J. Hols, and H.L.I.M. van Leeuwen, *Water Sci. Technol.* **23**, 1427–1436 (1991).
35. P.F. Cooper and S.C. Williams, *Water Sci. Technol.* **22**, 431–442 (1990).
36. M. Nagashima, M. Azuma, S. Noguchi, K. Inuzuka, and H. Samejima, *Biotechnol. Bioeng.* **26**, 992–997 (1984).
37. D. Weuster-Botz, A. Aivasidis, and C. Wandrey, *Appl. Microb. Biotechnol.* **39**, 685–690 (1993).
38. R.M. Busche, C.D. Scott, B.H. Davison, and L.R. Lynd, *Appl. Biochem. Biotechnol.* **34/35**, 395–417 (1992).
39. A. Aivasidis, Ch. Wandrey, H.G. Eils, and M. Katzke, *Proc. 23rd. Congr. Eur. Brew. Conv.*, 569–576 (1991).
40. I. Endo, T. Nagamune, K. Tachi, and T. Kobayashi, in G. Durand, L. Bobichon, and Florent eds., *Int. Biotechnol. Symp.*, Societé Française de Microbiologie, Paris, 1988, pp. 433–441 (1988).
41. T. Keshavarz, R. Eglin, E. Walker, C. Bucke, G. Holt, A.T. Bull, and M.D. Lilly, *Biotechnol. Bioeng.* **36**, 763–770 (1990).
42. M. Reiter, G. Blüml, T. Gaida, N. Zach, F. Unterluggauer, O. Doblhoff-Dier, M. Noe, R. Plail, S. Huss, and H. Katinger, *Bio / Technology* **9**, 1100–1102 (1991).

See also BIOREACTORS, AIR-LIFT REACTORS; BIOREACTORS, CONTINUOUS STIRRED-TANK REACTORS; FERMENTATION MONITORING, DESIGN AND OPTIMIZATION; MASS TRANSFER; SCALE-UP, STIRRED TANK REACTORS.

BIOREACTORS, GAS TREATMENT

GRAHAM ANDREWS
MMBD Consulting
Gresham, Oregon

WILLIAM APEL
Idaho National Engineering and Environmental Laboratory
Idaho Falls, Idaho

KEY WORDS

Biofilters
Biophase
Bioscrubbers
Biotricking filters
Bubble-gasified bioreactors
Cometabolism
Electron acceptors
Electron acceptors
Electron donors
Nutrient addition

OUTLINE

INTRODUCTION

The most common type of bioprocess, at least in the United States, is the aerobic treatment of wastewater, in which contaminants in the aqueous phase are biodegraded with the help of oxygen transferred from the gas phase (air) to generate a metabolic product (carbon dioxide) that, being volatile, is then stripped out of the aqueous phase into the gas. The familiarity of this arrangement should not be allowed to limit our imaginations about the possibilities of bioprocessing. It is perfectly possible to construct a process in which a gas-phase contaminant (hydrogen sulfide) is biodegraded with the help of a nonvolatile oxidant (nitrate), generating a product (sulfuric acid) that remains dissolved in the aqueous phase. The bioreactor for such a process must share many of the features found in aerobic wastewater treatment, such as efficient gas–liquid contacting, cell recycling or immobilization to create a high concentration of biomass, control of temperature and pH, and addition of nutrients in order to create a comfortable environment for microbial metabolism. However, the different objective, the treatment of a gas rather than a liquid, must dictate differences in the details. This point can be illustrated by thinking of a conventional trickling filter not as a device for treating sewage, but as one for removing

oxygen from the air that flows through it. It would work, but terribly inefficiently, with a fractional removal on the order of 1%.

This chapter presents a three-step study of the use of bioprocesses for the removal of undesirable components from gas streams. The existing and potential applications are described first, followed by the types of bioreactors that may be employed. A simple process analysis is then presented to illustrate how the choice and design of the bioreactor is dictated by the particular application, the type of microbial metabolism involved, the concentration and solubility of the contaminant, and so on. It should be noted that there are also a number of potential bioprocesses in which the objective is not the removal of a contaminant from a gas stream, but the bioconversion of gaseous substrates into useful products. They include the partial oxidation of methane to methanol, and the production of ethanol and acetic acid from mixtures of carbon monoxide and hydrogen. Work on these processes is not described in detail, except when doing so illustrates the common problems of bioreactors with gaseous substrates.

MICROORGANISMS AND APPLICATIONS

Due to inherent difficulties in achieving the sterile conditions necessary for prolonged maintenance of pure cultures in field-scale bioreactors, most gas treatment bioreactors contain a mixed culture or consortium of microorganisms. These consortia may be derived from a number of different inocula, from common sources such as sewage-treatment facilities, biofilms from established bioreactors used for similar applications, or soils and waters from areas contaminated with the substrate of interest. In some instances, it is possible to enrich for the desired microbial consortium directly from the bioreactor bed medium, particularly when media components such as soil, compost, peat, or bark chips are used, because these materials naturally contain a mixture of microbes with wide-ranging physiological capacities. Regardless of the inoculum source, the bioreactor must be operated under conditions that select for and maintain, over the life of the bioreactor, microorganisms with the physiological capabilities necessary to catalyze the desired bioconversion. Typical selective factors include the electron donor (i.e., the metabolizable energy source), available terminal electron acceptors, supplemental nutrients, pH, Eh, and temperature.

Regardless of the selection strategy used to achieve a microbial consortium with the desired physiological properties, the use of a consortium versus the attempted use of a pure culture of microbes has a number of practical advantages. As mentioned above, it is difficult to achieve the sterility standards necessary to maintain a pure culture of microbes in a gas treatment bioreactor under field conditions. In addition, a pure culture of microbes is often metabolically incapable of fully degrading a contaminant, so hazardous intermediates may be created and build up within the bioreactor or in effluent streams from the bioreactor. Because mixed consortia inherently tend to be composed of a number of different microbial species with a gamut of physiological capabilities, undesired intermediates that may be created by pure cultures are often fully degraded by consortia. Similarly, in many applications, the concentrations and chemical compositions of gas streams to be treated may be transient. Consortia generally have a higher probability than pure cultures of adapting to this transience. The contaminant gases and vapors that can be successfully biodegraded can be classified by their role in microbial metabolism.

Electron Donors

Microorganisms require a constant supply of metabolic energy, which is normally derived from the oxidation of either organic or inorganic compounds, with these compounds serving as electron donors to the microorganisms. Many environmentally or industrially significant gases and vapors are metabolically oxidizable by microorganisms and can serve as electron donors. Thus, many different types of gas and vapor streams can be and have been, treated microbiologically (1). Microbially oxidizable gases and vapors that have been treated in bioreactors include hydrocarbons, ketones, ammonia, xylene, alcohol, terpene, and carbon disulfide vapors, as well as gases such as methane and hydrogen sulfide (2–10). In the case of carbon-containing substrates (hydrocarbons, carbon monoxide, etc.), the metabolism of these compounds provides heterotrophic microbes not only with an electron source, but also with a source of cellular carbon. Non-carbon-containing substrates such as ammonium, while providing a metabolic electron source, must be metabolized by autotrophic microorganisms capable of obtaining their cellular carbon via the fixation of atmospheric carbon dioxide.

Cometabolism

Some substrates cannot serve as sole energy sources for microorganisms, but are nevertheless potentially bioconvertable. Their biodegradation is achieved in the presence of another compound that serves as the microorganism metabolic electron donor, a process that is termed cometabolism. Trichloroethylene (TCE) provides a well-documented example of a substrate that is oxidized cometabolically by microorganisms. Typical of cometabolic TCE degraders are the methanotrophic bacteria, aerobes that oxidize methane as their sole carbon and energy source (11). This oxidation takes place in a sequential manner, with methane first being converted to methanol, which is in turn oxidized to formaldehyde, formate, and finally, carbon dioxide. Along this oxidative pathway, the methanotrophs generate cellular energy as well as fixing methane-derived carbon into additional biomass. The first step, in which methanol is formed from methane, is catalyzed by the enzyme methane monooxygenase (MMO) (12,13). Although MMO is selective for its natural substrate, it can, under certain conditions, catalyze the oxidation of a variety of other compounds (12–14), including partially chlorinated aliphatic solvents such as TCE. Two distinct forms of MMO have been identified, membrane MMO (mMMO) and soluble (sMMO) (15,16). sMMO, ex-

pressed by methanotrophs growing under copper-limited conditions, supports much higher rates of TCE oxidation than does mMMO, which is expressed under conditions of copper sufficiency (17).

Air streams contaminated with TCE and similar chlorinated solvents are generated by industrial operations and the remediation of contaminated ground and water by air stripping. They can be treated in a bioreactor that is started up with methane as the sole carbon and energy source, while supplying the necessary oxygen and supplemental nutrients under copper-limiting conditions (6). After the methanotrophic consortium is established, the methane load to the reactor can be decreased or fed on an intermittent basis concurrently with TCE vapors. The methanotrophic microbes then will be maintained by the lessened methane feed, while catalyzing the oxidation of the TCE via their sMMO. The reduced or intermittent methane feed not only reduces the amount of methane that must be supplied, but is essential in achieving optimal TCE removal rates since methane itself is an effective competitive inhibitor of TCE oxidation. An alternate, but more expensive, strategy used to avoid this type of competitive inhibition is to supply a catabolic intermediate such as formate in order to provide energy to the methanotrophs, because formate does not compete with TCE for access to the sMMO. It should be noted that a number of other microbial oxygenases have been shown to be capable of catalyzing cometabolic TCE oxidation. These include toluene oxygenases from a number of *Pseudomonas* spp., propane monooxygenases from propanotrophs, and ammonia monooxygenase from *Nitrosomonas europea* (18–21).

Electron Acceptors

Most gas treatment bioreactors remove contaminants from air streams and oxidize them to innocuous end products (CO_2, Cl^-, etc.), using the oxygen in the air as the electron acceptor. However, there are also gas streams containing little or no oxygen where the objective is to remove a compound that can act as an alternative physiological terminal electron acceptor. This approach has been shown to work for compounds like sulfur oxides (SO_x) and nitrogen oxides (NO_x) (8,24,25). NO_x biofilters have been developed using a consortia of denitrifying microbes that have the physiological capacity, as a part of their normal metabolism, to use a variety of NO_x compounds as terminal electron acceptors, reducing them to innocuous nitrogen gas (N_2). This process generally requires low-oxygen conditions because denitrifiers preferentially use oxygen, if present, as their terminal electron acceptor. As a result, selective conditions for denitrifying microbes include a suitable carbon and energy source with a NO_x compound as a terminal electron acceptor, a lack of oxygen, and the presence of supplemental nutrients necessary for microbial growth. The potential application to stack gases raises the question of the maximum temperature at which the microbes will function. Research (25) has shown that naturally occurring thermophilic microbes selected for from compost are capable of reducing nitric oxide (NO) to nitrogen gas (N_2) at temperatures up to 60 °C.

Other processes are far more sensitive than denitrification to trace amounts of oxygen. The sulfite-reducing bacteria responsible for the removal of SO_2 are obligate anaerobes, although they may be active in anoxic microniches, for example, deep in a biofilm. Chlorinated aliphatic solvents will also accept electrons under anaerobic conditions, a degradative mechanism that works best on the completely chlorinated compounds that are not touched by the oxidative enzymes already discussed (22,23). Carbon tetrachloride, for example, will be progressively reduced to chloroform, dichloromethane, and methyl chloride. This mechanism is often categorized as cometabolism, implying that the microbes derive no metabolic benefit from it, but this remains uncertain.

BIOREACTOR TYPES

Figure 1 shows a generic gas treatment bioprocess along with the nomenclature to be used in the subsequent analysis. Gas treatment bioreactors are sometimes described as *gas-phase bioreactors*, but this term is inexact in the sense that the actual reactions happen inside a microbial cell, which is necessarily an aqueous phase. Some reactions carried out in immobilized enzyme reactors do seem to involve direct interaction between the gas and the enzyme (33), although even here the humidity remains a critical variable and the enzyme may be covered with a layer of water a few molecules thick. For our purposes, the bioreactor is divided into a gas phase, a solid phase consisting of any solid packing or biofilm support particles, and a biophase, which contains the water and the microorganisms. The fraction of the bioreactor volume occupied by each phase is called its holdup, and the three values, e_g, e_s, and e_b, must add up to 1.

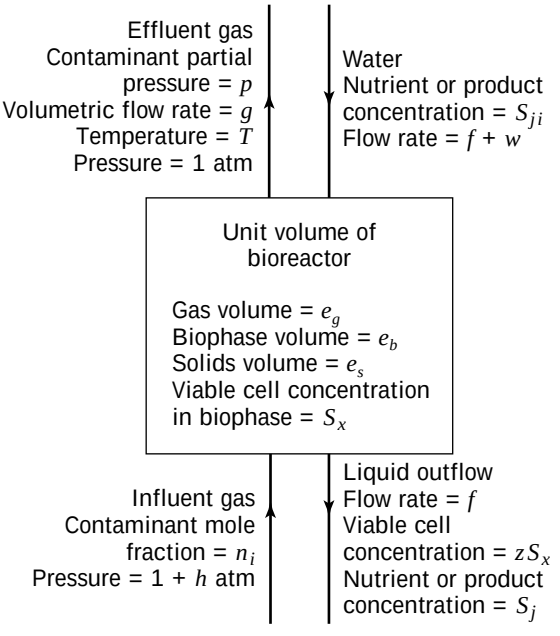


Figure 1. The generic gas-treatment bioreactor.

For convenience, we think in terms of a bioreactor of unit volume and express the gas flow rate as the volume per volume per minute (VVM) flow rate, g, at the effluent gas conditions, taken as temperature T and atmospheric pressure. Others choose to work in terms of the molar outlet gas flow rate (g/RT, where R is the gas constant) or a gas residence time, either the empty bed residence time, $1/g$, or the average time an element of gas actually spends in the reactor, e_g/g. The flow rate of the influent gas is not the same as g because its temperature may not be T, it is at a higher pressure, ($1 + h$ atmospheres, where h is the pressure drop through the reactor), and because various components (contaminant, oxygen, metabolic products, water vapor) are added to, or removed from, the gas stream as it flows through the reactor. When these effects are significant, it is essential to specify exactly what is meant by "gas flow rate."

The objective of some processes is to remove a gaseous contaminant by converting it into a soluble, nonvolatile metabolic product. Some such products, H^+ for example, can be neutralized in the bioreactor, but others, such as the Cl^- from the degradation of chlorinated organics, must be removed in the liquid if they are not to accumulate to concentrations that inhibit the microorganisms. There may therefore be a liquid outflow, f per unit bioreactor volume, and an inflow that must be greater than f to allow for the evaporation rate, w. This outflow tends to wash microorganisms out of the bioreactor, and it is usually beneficial to retain them either by cell recycle or immobilization, making the parameter z in Figure 1 less than 1.

The question is what type of bioreactor should be put in the "black box" in Figure 1. Almost every possible configuration by which a gas can be contacted with a biofilm or a suspension of microorganisms has been employed in some application, and these can be roughly divided into four categories.

Biofilters

The earliest gas-treatment bioreactors were used to control odors in air coming from slaughterhouses and rendering plants and consisted of porous pipes buried in the ground, the noxious compounds being removed by a combination of adsorption and biodegradation as the air flowed up through the soil (26). Applications of the idea have since expanded to include the removal of hazardous, volatile compounds from chemical plants, paint shops, foundries, and a variety of agricultural and food-processing operations. Although often effective, the operation of these simple soil beds depends critically on the local soil and weather conditions. Unless they are covered and/or periodically sprayed with water, they may stop working, either due to becoming waterlogged or to drying out. They have largely been replaced by custom-made biofilters, which are essentially boxes made of plastic, wood, sheet metal, concrete, or even stainless steel, containing a bed of media and a gas distribution system to ensure uniform flow of the gas through the bed.

In terms of the variables defined in Figure 1, biofilters are the small e_b, small f, and small z approach. The biophase consists of individual microorganisms attached to the media, perhaps enough to constitute a thin biofilm. The amount of moisture in this layer is critical, and is controlled by humidification of the inlet gas and/or by spraying water directly over the bed. Little or no water drips out of the bed, making it difficult to provide soluble nutrients and pH control chemicals, or to wash out nonvolatile metabolic products. Biofilters are thus best suited to applications that generate no such products and involve no pH swings, the treatment of air streams containing low levels of volatile organics being the obvious example.

The great advantage of biofilters over the early soil beds is that they allow the choice of the nature and particle size of the bed media. This media must support a high density of attached microorganisms suited to the particular application, perhaps even providing some of the nutrients needed for their growth. The ability to adsorb the contaminant is an advantage because adsorption can provide some contaminant removal during the start-up period, before a large population of well-acclimated microorganisms has developed (sometimes called "the bed-ripening period"), and during any sudden slugs of concentrated contaminant in the influent gas. Also, although the interactions between such adsorption and biodegradation are complex and poorly understood, they are generally favorable to biofilter performance. Finally, the media must be mechanically strong and resistant to disintegration, compacting, and the resulting channeling of gas flow. Soil, compost, peat, and wood-chip mixtures are all inexpensive media that have been used successfully. Materials such as activated carbon and limestone, although more expensive, may be mixed in to provide extra adsorption and pH buffering.

The particle size of the media is critical. Small particles give a huge gas–biophase interfacial area per unit bed volume (a quantity subsequently called "a"), thus eliminating concerns about mass-transfer limitations between the two phases. But beds of small particles have a higher resistance to gas flow, are more prone to waterlogging and plugging by excess biofilm, and if too light, may get blown out of the reactor altogether. A general guideline is to choose the smallest available particles that avoid the latter problems, and then ensure that the shape of the bed is optimized to give the desired contaminant removal with a reasonable pressure drop (see "Scale-Up").

If the life of the bed is found to be finite due to compaction, loss of pH buffering capacity, accumulation of salts from the evaporation of added water and so on, one solution is to employ two beds in series. Bed 1 can provide most of the contaminant removal, while bed 2 polishes the effluent, mainly by adsorption, while developing its own population of well-acclimated microorganisms. When bed 1 is exhausted it can be replaced by bed 2, which is itself replaced by a bed of fresh media. The arrangement of the media on a series of trays in the bioreactor facilitates this mode of operation.

Bubble-Gasified Reactors

Conventional bubble-aerated fermenters make poor gas treatment bioreactors because they are designed for a high rate of transfer of gaseous nutrients and products (O_2 in

and CO_2 out), rather than a high fractional removal of a component from the air stream. This latter objective requires taller bioreactors with longer gas residence times, but tall reactors mean higher pressure drops due to the hydrostatic head. Nevertheless, tall, mechanically agitated bioreactors with multiple impellers have been proposed for removing H_2S from industrial gas streams (29), and bubble columns with no mechanical agitation have been demonstrated for the removal of TCE from air using a toluene-oxidizing organism (30).

These bioreactors are in many ways at the opposite extreme from biofilters. The biophase occupies most of the reactor ($e_b \rightarrow 1$ in Fig. 1), and it is dilute and well-mixed rather than dense and immobilized. The large volume of water makes it easy to add soluble nutrients, control the pH, and remove water containing nonvolatile products. However, care must be taken to not wash out the biomass, because $z = 1$ (Fig. 1) unless there are extra surfaces in the bioreactor for the attachment of a biofilm, or if a cell separation and recycle loop is added. Another major difference is that the gas–biophase interfacial area, a, is not only much smaller than in a biofilter, but is not even a constant, varying with the gas flow rate. Bubble gasified bioreactors are therefore best suited to contaminants like NO_x that are very soluble in water (thus reducing concerns about the gas–liquid mass-transfer rate), whose dissolution in water cause pH swings, and whose metabolism requires large amounts of dissolved nutrients.

Biotrickling Filters

Trickling filters for gas treatment are similar to, but usually taller than, those in use for decades for the secondary treatment of municipal and industrial wastewaters. Like biofilters, they consist of a bed of media through which the gas flows either upward or downward, but unlike biofilters, the microbial culture is recirculated continuously over the bed from a reservoir beneath it. Stone aggregate or wood chips are the traditional media, but even seashells have been used (27). Most modern filters employ ceramic or plastic media, Pall rings, Raschig rings, and so on specifically designed to make a bed with a high surface area to maximize gas transfer, and a high porosity to minimize the pressure drop and the chance of clogging or flooding (28).

These bioreactors can be thought of as intermediate between biofilters and bubble-gasified systems and are best suited to applications between those already described. If the characteristics of the media and the microorganisms are such that most of the biomass stays attached, then a trickling filter is essentially a biofilter with a higher liquid flow and a smaller gas–biophase interfacial area, a, due to the larger media. If on the other hand, most of the biomass stays suspended and the recycle ratio (RR = flow of recycled liquid/f) is high enough to keep the biophase well mixed, then it is functionally similar to a bubble-gasified bioreactor. The main differences are that the pressure drop, h, is usually much smaller, and the interfacial area, a, is a constant that is independent of the gas flow rate.

Bioscrubbers

The systems just described essentially combine two functions, stripping the contaminant from the gas into the aqueous biophase and then subjecting it to biodegradation. But what if the physicochemical conditions in the biophase resulting from the stripping are not suitable for the microbial metabolism? The removal of sulfur dioxide from stack gas is a good example because stack gases contain a small percentage of oxygen that although relatively insoluble, inhibits the bacteria responsible for the reduction of the sulfite formed by the dissolution of SO_2. One solution is the bioscrubber, which consists of an aqueous stripper in which the contaminant is transferred from the gas to water, and a separate bioreactor, in which the liquid effluent from the stripper is essentially treated as wastewater. The bioreactor should be a completely mixed, immobilized-cell type so that a large population can grow up in a well-controlled environment without continuously being recirculated through the potentially toxic environment in the stripper. Several configurations are possible, the simplest being that in which the liquid reservoir at the bottom of a packed-bed stripper is used as the bioreactor. Bioscrubbers are not considered further here because building a separate bioreactor and stripper is necessarily more expensive than combining them in a single unit, and when it is necessary, the two unit operations can be designed by conventional methods.

BIOREACTOR DESIGN

A process variable is any quantity that is under the direct control of the designer or operator of the process. For gas treatment bioreactors, the variables to be fixed by the designer include the bioreactor type, shape, and size, whereas those fixed by the operator include the addition of water, nutrients, and chemicals for pH control. The objective in fixing these variables is to satisfy a set of process requirements that specify the flow rate and composition of the gas stream, the nature and concentration of the contaminant, and the fractional removal required. The difficulty is that many designs, that is, many combinations of the process variables, will meet the process requirements for a particular case. The proper goal of engineering design is not simply to pick one of these designs at random, but to find the single design that meets the requirements at the lowest possible cost. The design of gas treatment bioreactors, like that of any other process, is an economic optimization problem, but one that is so complex that the exact solution cannot be found within the constraints of time and money normally imposed on the design process. The best solution for a particular case will certainly depend on the scale of the problem, the solubility and possible inhibitory effects of the contaminant, the type of microbial metabolism (aerobic or anaerobic, growth associated or cometabolic, etc.) needed to degrade it, and the temperature and pressure of the gas stream. Even coming close to the best solution requires a careful scale-up program in which the insights available from education and experience are used to integrate experiments at different scales with mathematical models that can extrapolate the results from one scale to the design of the equipment at the next scale. The bioreactor analysis presented in this section is intended as a guide to this program, sug-

gesting which bioreactor types be chosen for experimental evaluation in each case, specifying the operating conditions around which the experiments should be designed, and showing how performance may be expected to vary with scale.

Because the design goals are essentially economic, the relationship of the variables shown in Figure 1 to the process costs must be kept in mind. The volume of the commercial-scale bioreactor equals the actual gas flow to be treated (a process requirement) divided by g; a large value for g translates directly into a smaller and, all other things being equal, cheaper bioreactor. The liquid flow, f should be kept as small as possible because providing clean water and disposing of any process wastewater both cost money, as do any nutrients and chemicals that must be added. The pressure drop, h, through the bioreactor may be significant, particularly on a large scale, and it determines the capital cost of the gas compressors and the energy costs for running them. Compressing the feed gas is not only expensive in itself, it heats the feed such that heat exchangers may be needed to cool it to the desired temperature.

Process Analysis

It will be assumed that the biophase is completely mixed and that the gas approximates plug flow through the bioreactor. This provides a realistic description of some bioreactors (bubble columns, some biotrickling filters), and a reasonable starting point for quantitative thinking about the others. More detailed mathematical models are available in the literature for specific types of bioreactors and metabolism (31,32). The equilibrium solubility of the contaminant in the biophase, S^*, will be described by Henry's law, $S^* = p/H$, where p = partial pressure of the contaminant in the gas and H = Henry's law constant. Although this is adequate for relatively insoluble and dilute contaminants, it is a considerable oversimplification for others. For example, the dissolution of oxides of nitrogen (NO_x: a mixture of NO and NO_2) from stack gases into water is an extremely complex process (34,35) that involves the gas-phase chemical oxidation of NO to NO_2 and a dissociation reaction that generates H^+, NO_3^-, and NO_2^- when the gas dissolves. In such cases, Henry's law can only be interpreted in the purely qualitative sense that low H means a soluble gas.

Any complete bioprocess model combines the mass conservation equations that describe the bioreactor with equations that try to approximate the complexities of microbial metabolism. Because the objective here is to compare different bioreactor configurations, it is essential to have consistent descriptions of the metabolism. In most gas treatment applications the contaminant is a major microbial nutrient, either the electron donor or the electron acceptor. Its consumption provides metabolic energy for growth and cell maintenance, so its specific consumption rate, q is given by the yield equation

$$q = \frac{\mu}{Y} + k \qquad (1)$$

where μ is the specific growth rate of cells, Y is the cell-yield coefficient, and k is the cell-maintenance coefficient.

In a well-operated bioreactor, the pH and temperature are carefully controlled and all required nutrients are provided in adequate amounts, making the specific consumption rate a function only of the concentrations of contaminant, S, and any inhibitory product, S_m, dissolved in the biophase. The exact form of this function, written $q(S, S_m)$, will not be specified because it varies among cases, but it must describe the effects on the metabolic rate of substrate limitation, product inhibition, and possibly (when n_i/H is large) substrate inhibition. It is more common to describe the specific cell growth rate, μ, by a Monod-type function of the concentration S, but this is clearly incorrect at the low concentrations found in gas treatment bioreactors, because it predicts that growth stops and substrate uptake continues when no substrate is available ($\mu = 0$ and $q = k$ when $S = 0$). In fact, when a major nutrient is exhausted, its consumption must stop, and the microorganisms go into an endogenous state in which the viable biomass declines, a phenomenon correctly approximated by requiring that $q(S, S_m) = 0$, and thus $\mu = kY$ (equation 1) when $S = 0$. Cell growth now stops at a nonzero "stationary phase" concentration, S_s, corresponding to the point where substrate uptake is just sufficient to satisfy the maintenance requirement, and defined by $q(S_s, 0) = k$.

For cometabolic processes, the contaminant is neither the main electron donor nor the acceptor; two rate equations are needed, and equation 1 contains an extra parameter whose value may be found from knowledge of the degradative pathway. See, for example, the analysis by Andrews of chloroform degradation by methanotrophic bacteria (36). The design of cometabolic bioreactors follows many of the principles described here but is too complex and case specific to be analyzed in detail.

At steady state the conservation of mass requires that the rate of growth of microorganisms in the bioreactor, $\mu e_b S_x$, equals the rate at which they flow out, fzS_x (see Fig. 1 for nomenclature). The loss of microorganisms in mist suspended in the effluent gas may be significant in some bioreactors without demisting devices but is not included explicitly in the analysis. The specific growth rate, μ, a measure of the physiological state of the microorganisms, thus equals zf/e_b, the reciprocal of the mean cell-residence time (the number of cells in the bioreactor divided by the cell outflow rate), which is a quantity under our direct control. This remarkable result, although well known for chemostats, has not been fully appreciated for gas treatment. It follows from equation 1 that the concentrations in the biophase, S and S_m, are constrained by

$$q(S, S_m) = k(1 + D) \qquad (2)$$

where $D = zf/Yke_b$ is the dimensionless dilution rate or reciprocal mean cell-residence time.

The steady-state conservation of mass for the contaminant requires that whatever is removed from the gas phase be transferred to the biophase, where it is either consumed by the microorganisms or flows out in the liquid. For a bioreactor, as opposed to a gas stripper, the latter term

should be negligible, but the flows of dissolved nutrients and nonvolatile metabolic products must be included

$$\frac{g}{RT}(n_i - p) = \frac{k_1 a(n_i - p)}{H \ln\dfrac{(n_i - HS)}{(p - HS)}} = e_b q S_x = \pm\frac{(fS_j - (f + w)S_{ji})}{y_j} \quad (3)$$

where y_j is the amount of component j produced (+ sign) or consumed (− sign) per unit of contaminant consumed (see "Nutrient Addition and Start-Up").

The description of the driving force for mass transfer in terms of the log–mean concentration difference in the second term of this equation follows from the assumption of plug flow of the gas. The expression is not strictly correct when the pressure drop through the bioreactor is large, as in a tall bubble column, because the contaminant partial pressure in the influent is then $(1 + h)n_i$; but this complication is rarely included in practice. The essential points here are that the mass-transfer factor, k_1a, is an empirically determined parameter, that it can be based on different forms of the driving force, and that the best choice depends on the reactor type and must be clearly understood by everyone involved in a project.

Equation 3 can be rearranged to give the fractional removal of the contaminant from the gas, only the product (subscript m) being considered in the last term:

$$x = 1 - \frac{p}{n_i} = (1 - e^{-G})\left(1 - \frac{HS}{n_i}\right) = \frac{S_x}{X}(1 + D) = FD\,\frac{S_m}{S'_m} \quad (4)$$

where $G = k_1aRT/gH$ is the dimensionless mass transfer coefficient; $F = YS'_m/zXy_m$ is the dimensionless product inhibition factor; $X = gn_i/ke_bRT$ is the maximum possible viable cell concentration in the biophase; and S'_m is the product concentration that completely inhibits cell growth.

Rather than solving equations 2 and 4 exactly, which would require the form of the inherent kinetic function $q(S, S_m)$, consider some limiting cases in order of decreasing liquid outflow rate. These cases are used to summarize the bioreactor performance in Figure 2, a graph that can be universal, at least within the limitations of the assumptions, because the parameters describing the contaminant (H, n_i), the reactor (f, k_1a), and the metabolism (k, S'_m) in individual cases have been incorporated into the dimensionless numbers x, D, F, and G.

- *Washout.* $D > (q(n_i/H,0)/k) - 1$: Because the dissolved concentration, S, must obviously be less than that in equilibrium with the feed gas, n_i/H, equation 2 now has no solution, meaning that all of the biomass is washed out of the reactor and there can be no contaminant removal.
- *Substrate limitation.* $D > (1 - e^{-G})/F$: Equation 4 shows that under this condition, $S \to 0$, while $S_m < S'_m$, meaning that microbial kinetics are limited by the availability of substrate rather than product inhibition and the exact value of the inhibition number, F, is irrelevant. If gas-biophase mass transfer is very rapid, the dissolved contaminant is in equilibrium with the outlet gas, giving the best possible bioreactor performance. The shape of this $G = \infty$ line in Figure 2 is simply a reflection of the relationship between $\mu(=zf/e_b)$ and $S(=p/H)$ from the inherent microbial kinetics. Mass-transfer limitation ($G < \infty$) leads to a progressive reduction in the contaminant removal, x.
- *Product inhibition.* $D < (1 - e^{-G})/F$: Now the concentration of metabolic products, S_m reaches inhibitory levels, S'_m, while dissolved contaminant is still available, meaning that metabolism is limited by product inhibition rather than substrate limitation, and the exact value of G is irrelevant. Increasingly stronger inhibition (smaller F) produces a series of curves in Figure 2 whose exact shape depends on the details of the inherent inhibition kinetics. All of these curves pass through the origin because, with no liquid outflow ($D = 0$), any soluble, nonvolatile metabolic products must eventually accumulate to inhibitory levels.

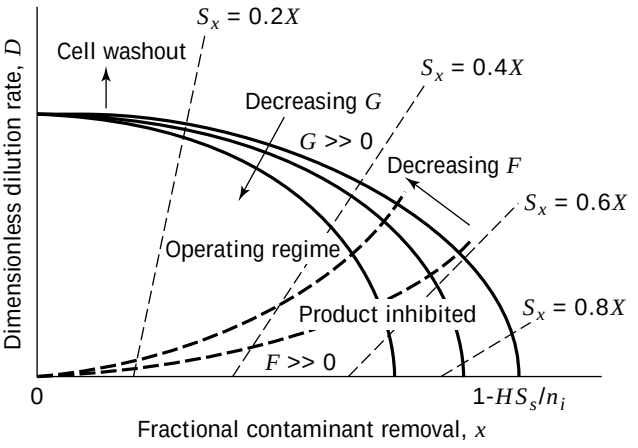


Figure 2. Operating regime for a completely mixed bioreactor.

Consequences for Process Design

Any pair of values of F and G define a regime in Figure 2 within which the bioreactor can operate. The most important result is that for any such pair of values there exists a value of D that maximizes the contaminant removal, x. If the form of the rate equation $q(S, S_m)$ were known, the pointed maximum seen in Figure 2 would be replaced by a smooth curve that defined this optimum design more accurately. In the absence of such detailed modeling the optimum must be found experimentally, the analysis providing a starting point and overall guidance. Because the gas flow rate, g, appears in the denominator of both F and G, the first step is to establish the maximum g for which the desired contaminant removal, x, falls safely within the operating regime. It is clear from equation 4 that a high re-

moval rate (e.g., $x = 0.95$) with reasonable values of $S(\gg Hn_i)$ and $S_m(<0.5\ S'_m)$ requires both $G \sim 3$ and $FD \sim 2$, and such values are reasonable, if approximate, general design goals. Much larger values represent inefficient design because they imply excessive consumption of energy and water.

For processes that produce large amounts of very inhibitory, nonvolatile, soluble metabolic products, the problem is to wash the products out of the bioreactor without washing out the biomass. Mathematically, $y_m/S'_m \to \infty$, making it difficult to keep $FD > 2$ without D exceeding the washout criterion. These are the processes that require cell retention ($z \to 0$), either by cell recycle or by immobilization as a biofilm. For processes with more moderate values of y_m/S'_m, the analysis shows that there exists an optimum set of values of the process variables g, f, and z. Some processes, the aerobic mineralization of hydrocarbons for example, generate only innocuous (H_2O) or gaseous (CO_2) products, so $y_m/S'_m \to 0$, and the analysis suggests operating with no liquid outflow. This not only gives the best effluent-gas quality (Fig. 2 with $F = \infty$) but, because the biomass is in a maintenance condition ($\mu = 0$ when $D = 0$), it also avoids the cost of adding nutrients, such as phosphates, that are needed for microbial growth (see "Nutrient Addition and Start-Up"). In practice there are reasons to avoid this condition. Even if y_m/S'_m is too small to be observed in short-term laboratory experiments, any nonvolatile product must eventually accumulate to inhibitory levels in a continuous bioreactor with no liquid outflow. Salts in the water added to compensate for evaporation or for pH control will also accumulate in the biophase. Also, it must not be forgotten that complete mixing of the biophase is an abstraction never achieved in practice and sometimes, as in the biofilter, not even attempted. The discussion prior to equation 2 shows that any bioreactor with $D = 0$ must have a cell growth rate $\mu = 0$ as an average, but it will contain microniches where cells are dying because conditions are worse than average, and others where conditions are slightly better and cells are growing. Although the growing cells may assimilate some of the material released by the lysis of the dying cells, there must be some long-term accumulation of dead biomass, with consequent reduction in bioreactor performance. Keeping D on the of order of 1, a very low flow rate in most cases can wash out the dead biomass and ensure stable operation for the cost of a few growth nutrients.

The analysis assumed the usual case in which the cell concentration in the bioreactor is kept as high as possible, with sufficient amounts of other nutrients provided to ensure that the contaminant is the limiting nutrient. The quantity X, defined under equation 4, is then the maximum possible viable-cell concentration in the biophase, corresponding to the situation where all of the contaminant flowing into the bioreactor, gn_i/RT, is being used to satisfy the cell maintenance requirement, ke_bX. The actual concentration at any given operating condition, S_x, can be calculated from the fourth term of the equation and is shown in Figure 2 as a function of D and x. There is, however, no guarantee that this concentration corresponds to a dilute, easily mixed cell suspension. If the maintenance requirement, k, is small, and the contaminant loading gn_i is large, it may be a paste of cells too dense to be, for example, pumped around a biotrickling filter. Although it is possible to reduce S_x by restricting the amounts of growth nutrients provided, such artificial reductions in the microbial "catalyst" concentration are generally poor bioprocess design. The alternative is to choose the right type of bioreactor, recognizing that, although the mass of microorganisms in the bioreactor is fixed by the process requirements and the microbial metabolism, the amount of water is a process variable that can be controlled. The concentration, X, is an order of magnitude higher in a trickle-bed type bioreactor where $e_b \approx 0.1$, than in a bubble column where $e_b \approx 1$.

Nutrient Addition and Start-Up

In some processes, the aerobic mineralization of organics in beds of compost for example, the air provides the electron acceptor and the media provides what little growth nutrients are needed. In other processes these substances, and sometimes the electron donor, must be provided in the liquid inflow. The amounts can be estimated from the "stoichiometry" of metabolism, a procedure best illustrated by example. Ongcharit et al. (37) have shown that hydrogen sulfide can be removed from industrial gas streams by the acidophilic, autotrophic, sulfur-oxidizing, denitrifying bacterium *Thiobacillus denitrificans*. Even such esoteric metabolism can be approximated by a pseudo-chemical reaction in which all of the main nutrients and products are written in the ionic form in which they enter or leave the reactor. After performing all of the element and charge balances the result is

$$\begin{aligned} H_2S + (1.6 - 0.2y\gamma)\,NO_3^- + yCO_2 + ypH_2PO_4^- \\ \rightarrow\!\!\rightarrow yCH_hN_nO_oS_sP_p + (0.8 - 0.1y(\gamma + 5n))N_2 \\ + (1 - ys)SO_4^{2-} + (0.4 + y(0.2\gamma - p - 2s))H^+ + wH_2O \end{aligned} \tag{5}$$

The first term on the right-hand side represents the dry weight of the biomass produced. $\gamma = 4 + h - 2o + 6s + 5p$ is a measure of its oxidation/reduction state and is remarkably consistent between species (36) (note that the base oxidation state for nitrogen here is N_2). The actual cell yield is y, the ratio of cells produced to H_2S consumed. For a completely mixed biophase this equals μ/q or, from equations 1 and 2,

$$y = \frac{DY}{(1 + D)} \tag{6}$$

The "stoichiometric coefficients" on the other components in equation 5 then provide the y_j values needed to calculate the concentrations of nutrients and products from the last term in equation 3. These simple calculations can provide a surprising amount of approximate information about process design and development. A feature of the example in equation 5 is the large amount of acid (H^+) produced by the metabolism. Even with the high liquid flow rate allowed by cell immobilization or recycle (29,37) its neutralization requires a high pH in the liquid inflow, so the biophase must be very well mixed if metabolism is

not to be inhibited by pH gradients, suggesting some type of bubble-gasified bioreactor. If all of the NO_3^- and OH^- are provided by their sodium salts, then the salinity in the biophase would be high, suggesting a need for halophilic strains of bacteria. A reasonable alternative would be to explore the use of calcium salts, which would precipitate some of the product sulfate as gypsum, reducing inhibitory effects.

Like all the previous analyses, equation 6 applies only to steady-state operation. Gas treatment bioreactors, like any other bioprocess, are generally started with a relatively small inoculum of acclimated microorganisms, the objective of the start-up period being to increase the population to the steady-state value, S_x. This requires additional growth nutrients, and the initial feed concentration, S_{ji}, can be calculated as above, but with setting the observed cell yield, y, equal to the growth yield coefficient, Y. These concentrations can then be decreased slowly as steady state is approached. Start-up of biofilters is sometimes called ripening of the bed, and it is not always necessary to add growth nutrients, because the microorganisms capable of growth on the contaminant may obtain nutrients from the support material (compost, soil, etc.), or may grow at the expense of cells that die and lyse under conditions in the biofilter. If these sources are not sufficient to produce the biofilm thickness needed for maximum contaminant removal, some nutrient addition after steady-state contaminant removal has been reached will further improve performance. However, excessive nutrient addition must be avoided because modeling (38) suggests that, under contaminant limited conditions, a biofilm will grow much thicker than is needed. The inlet end of the biofilter may plug with useless biofilm.

Just because a steady state exists in theory, it does not follow that it can be reached by any start-up procedure. Of particular concern are those contaminants that are so concentrated (high n_i) and soluble (low H) that they will cause substrate inhibition of the microbial metabolism if their dissolved concentration in the biophase ever reaches equilibrium with the feed (n_i/H). The steady state analyzed above still exists, but $S \ll n_i/H$ only as a consequence of the metabolism of the large numbers of microorganisms present. Starting with a small inoculum, the dissolved concentration in the biophase will be much higher, metabolism will be inhibited, and the biomass may wash out. For similar reasons, batch experiments may produce the erroneous conclusion that the gas cannot be treated biologically. A proper start-up procedure may involve dilution of the feed gas and careful control of the dilution rate and nutrient addition, until a dense, well-acclimated microbial culture has developed.

Gas-Biophase Mass Transfer

Figure 2 shows that high levels of contaminant removal require high values of the dimensionless mass-transfer factor, G, which is actually the product of two dimensionless numbers: H/RT, an inverse measure of the contaminant solubility, and k_1a/g, a measure of the efficiency with which the gas flow generates mass transfer. This is not usually a difficulty for biofilters with their thin biophase and very large a but how can it be achieved for other bioreactor types? Measurements of gas–liquid mass transfer factors (28) are often given as correlations of the form

$$k_1a = AU_g^c \tag{7}$$

The superficial gas velocity U_g is related to the flow rate g by $U_g = gL$, where L is the bioreactor height. Consider first an idealized bubble column, a hypothetical device in which all the bubbles are the same size, and doubling U_g simply doubles their number. The exponent c must then be 1, making $k_1a/g = AL$, a constant dependent upon scale but independent of gas-flow rate. In real bubble columns, with their imperfect bubble dispersion, c is closer to 0.7, but this still implies that G varies only slightly with g. The copious data available for oxygen transfer from air (28) shows that, at normal gas flows (g on the order of 1 VVM), k_1a/g is on the order of 1 at the laboratory scale, rising toward 10 in commercial-scale equipment. Assuming for the purposes of illustration that k_1 for a gas is proportional to its diffusivity in water raised to the 2/3 power (strictly valid only for relatively insoluble gases, for which most of the mass-transfer resistance is on the liquid side of the gas–biophase interface), this can be used to estimate G and the maximum possible fractional removal, $1 - e^{-G}$, (equation 4 with $S \to 0$) for contaminant gases of low, medium, and high solubility. The results shown in Table 1 suggest that bubble columns would be useless for insoluble contaminants like carbon monoxide, but worth consideration for moderately soluble contaminants like hydrogen sulfide. Most hydrocarbon vapors fall into the former category, but many alcohols and chlorinated aliphatic solvents are in the latter (30). For very soluble contaminants, such as sulfur dioxide, G tends to be unnecessarily large in bubble columns, and while acceptable removal with a moderate pressure drop could be achieved in very short, wide columns, this configuration has not been explored in practice.

The addition of a mechanical agitator can increase k_1a by roughly one order of magnitude, mainly by breaking up the gas flow into many small bubbles with a large interfacial area. Equation 7 still applies, but the A parameter becomes a function of the mechanical power input per unit reactor volume. Unfortunately, agitation also undermines the assumption that the gas moves though the reactor in plug flow. In a mechanically agitated tank, the gas bubbles are assumed to recirculate several times so that the gas phase is completely mixed; a sample of gas containing many bubbles taken from anywhere in the tank would have the same composition as the outlet gas. The effect on the driving force for mass transfer can be seen by comparing the concentration profiles shown in Figure 3a and b. Mathematically, the long-mean concentration difference used in equation 3 must be replaced by the more familiar effluent partial pressure driving force ($p/H - S$), which makes the maximum possible removal not ($1 - e^{-G}$), but $G/(1 + G)$. These two functions are virtually identical at small G, which explains why the exact definition of the driving force is not an issue for aeration (oxygen is an insoluble gas, $H/RT = 30$ at 20 °C, and its small fractional removal from the air stream in conventional fermenters is

Table 1. Performance of Bioreactors for which k_1a/g = 1–10 for O_2

Contaminant gas	H/RT at 20 °C	G	Maximum possible removal	
			Mixed gas phase $G/(1 + G)$	Plug flow gas $1 - \exp(-G)$
CO	37	0.023–0.23	0.022–0.19	0.023–0.20
H_2S	0.41	1.8–18	0.65–0.95	0.84–1
SO_2	0.0088	88–880	0.99–0.999	1

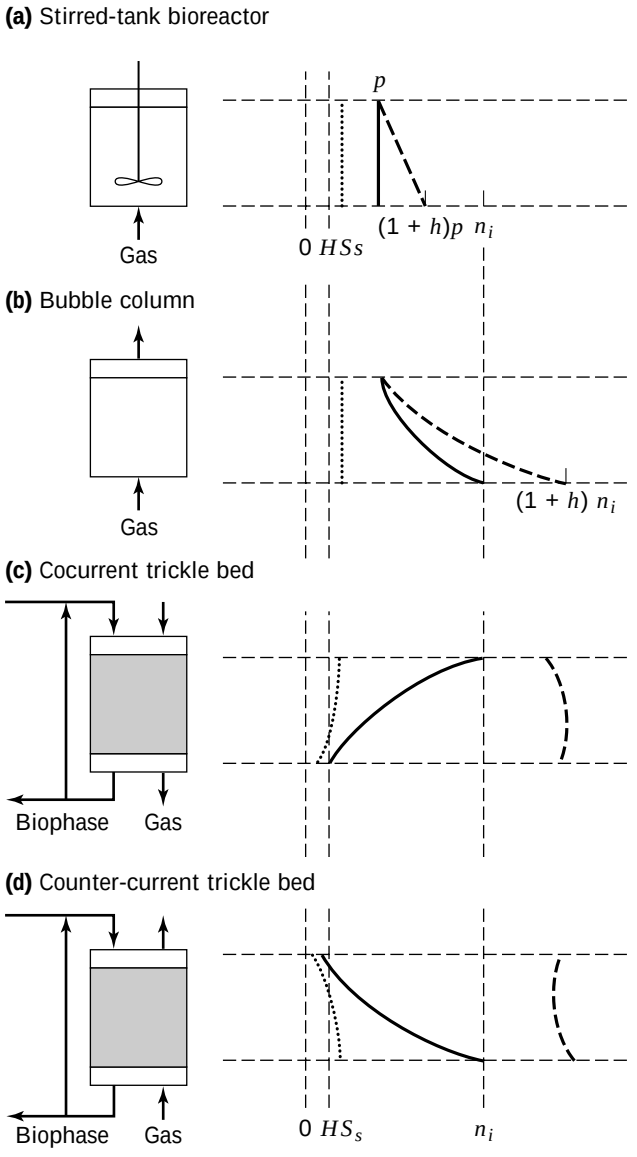


Figure 3. Concentration Profiles in Different Bioreactors. (**a**), (**b**) Partial pressure in gas phase, dot-dash line (**c**), (**d**) Viable biomass concentration, dot-dash line. Dotted line, partial pressure in biophase = HS; solid line, mole fraction in gas phase.

unimportant). However, a goal of 95% contaminant removal now requires not $G > 3$, but $G > 19$, which gives some idea of the lower effectiveness of a completely mixed gas phase. The numbers in Table 1 show that this lower effectiveness cancels out the increased k_1a leading to conclusions similar to those for bubble columns. The extra expense of mechanical agitation is rarely justified in gas treatment bioreactors, although reactors designed for a high k_1a/g with large aspect (height:diameter) ratios and multiple impellers to prevent bubble coalescence (29) have been proposed for removing hydrogen sulfide. Stirred tanks have been used successfully (39) for the microbial production of ethanol and acetic acid from carbon monoxide and hydrogen (another insoluble gas; H/RT = 51 at 20 °C), but the objective there was high volumetric productivity, rather than complete gas consumption. Several tanks were used in series, and they were pressurized to several atmospheres to improve gas transfer.

In contrast to bubble-gasified reactors, the gas–biophase interfacial area in biotrickling filters is almost constant, fixed by the choice of packing media and only slightly influenced by U_g and the superficial velocity at which the liquid is pumped over the bed U_1. This makes it possible to achieve the necessary value of G for contaminants of almost any solubility by appropriate choice of the packing media and the gas flow per unit volume, g. Many types of packing are available and the best choice in a particular situation is a complex trade-off between the following considerations:

- Less soluble contaminants will need the large a provided by smaller media and the small g implied by the lower permeability of such media to gas flow. For these contaminants, most of the gas–biophase mass-transfer resistance is on the biophase side of the interface, so in correlations of the form of equation 7, $c \to 0$, but A is a fairly strong function of U_1.
- More soluble contaminants can achieve the same G with the smaller a and larger g given by larger media. In this case most of the mass-transfer resistance is on the gas side of the interface, so $c > 0$ and A is less dependent on U_1.
- The conservation of mass shows that the change in concentration of a nonvolatile metabolic product in the biophase as it moves down the bed is $U_g x n_i y_m / U_1 RT$. If this number is small then the biophase approximates complete mixing. If it is large, then the concentration gradients, including that of pH, may be large enough to inhibit metabolism.

- Although a high liquid flow improves mixing and mass transfer, there is a definite upper limit of U_1 at which the bed floods. This limit is predictable (28) and is much smaller for small media.
- It is difficult to generalize about the attachment of the microorganisms to the media because this depends on both the nature of the media and the type of biomass involved. However, it is clear that a bed of small media with attached biomass may quickly become plugged if the contaminant is sufficiently concentrated and the biofilm is allowed to grow. It is reasonable to suppose that the larger the media and the higher the liquid velocity, the more cells will be suspended in the liquid rather than being attached. The z parameter in Figure 1 will be correspondingly larger.
- A biophase containing suspended and attached biomass may, if too thick or too dense, have its own mass-transfer resistance (38). While such resistance is usually to be avoided, it may be beneficial. For example, in trying to separate one electron acceptor, NO_x, from stack gases that contain another, O_2, which is preferred by the microorganisms, it should be possible to take advantage of the much higher solubility of the NO_x to create locally anaerobic environments in the biophase.

Somewhere among all of these considerations is a best choice of media for any situation. For dilute, less-soluble contaminants whose metabolism does not produce any nonvolatile metabolic products, the biotrickling filter will look more like a biofilter, with small media, little liquid flow, and attached biomass. For more soluble contaminants which do generate such products, it will tend toward the other extreme of high liquid flows and a completely mixed biophase. If there seems to be no sensible solution, the answer may be to adopt a bioscrubber. Although more expensive, it greatly simplifies the design problem by separating gas–liquid mass transfer from the microbiological considerations. It is certainly preferable to think in terms of this continuum of possibilities and to choose the one best suited to a particular application, rather than considering the different types of gas treatment bioreactors as completely separate technologies, each of which can be forced to fit any job.

Biophase Mixing and the Minimum Effluent Concentration

The single bioreactors with a completely mixed biophase analyzed here have an obvious drawback: all of the microorganisms are exposed to the same physicochemical environment, it is the worst possible environment for metabolism, with the lowest nutrient concentrations and the highest concentrations of metabolic products. One consequence appears in Figure 2. Even with no mass transfer resistance ($G = \infty$), no product inhibition ($F = \infty$), and no wastage of biomass ($D = 0$), the analysis suggests that complete contaminant removal is still impossible. With no biomass outflow there is no net cell growth, and all of the contaminant is used for maintenance metabolism (equation 2 with $\mu = D = 0$ gives $q = k$). Under these conditions, the dissolved contaminant concentration is not zero, but the stationary phase concentration, S_s, so the partial pressure in the effluent gas cannot be less than HS_s, and the fractional removal, x, can never exceed $(1 - HS_s/n_i)$, even with perfect mass transfer (equation 4 with $G = \infty$). This maximum removal also exists for completely mixed biological wastewater treatment systems, but in that case it does not matter because S_s is much less than common effluent standards. It may matter for some insoluble gas-phase contaminants because the corresponding minimum-achievable partial pressure, HS_s, may be significant, even when S_s is small. Bioreactors with a completely mixed biophase are a poor choice for such contaminants.

The simplest improvement would be two bioreactors through which the gas flows in series and between which the microbial culture is continuously recirculated. Reactor 1 would have a relatively high effluent partial pressure and dissolved contaminant concentration, providing a good environment for metabolism and cell growth. If the process variables are fixed correctly (a complete analysis is beyond the scope of this chapter) reactor 2 could run with a dissolved concentration $S < S_s$. The biomass in reactor 2 would then be in an endogenous state, but this would no longer matter because fresh, viable cells would be continuously supplied from reactor 1, and the dying cells would either flow back there to recover or be lost in the liquid effluent. The liquid inflow would go to reactor 1, where the growth nutrients are needed, and where it would dilute the concentrations of any metabolic products.

The more bioreactors in series, the better the average physicochemical environment for metabolism. Columns containing many sieve trays or bubble-cap trays have long been used for gas–liquid contacting in the chemical process industry and should be considered for the biological treatment of gases in cases where the biomass grows best in suspension and the contaminant is relatively dilute, insoluble, and a poor microbial substrate (i.e., HS_s/n_i is large). These columns are inherently countercurrent contacting devices with the gas flowing upward while the microbial culture falls down from tray to tray. An individual microorganism experiences a series of environments with ever-increasing dissolved contaminant concentration, S, until it is suddenly pumped back to the top tray where conditions are endogenous ($S < S_s$), but where it can continue metabolizing contaminant and survive long enough to start through the cycle again. Nutrient addition and pH control can be done on each tray as needed, and the liquid effluent can be taken from near the top of the column where cell viability is lowest.

A similar effect, with the gas closer to the theoretical optimum of plug flow, can be achieved in an upflow biotrickling filter with a small liquid recycle ratio. The only drawbacks are that pH control and nutrient addition can now only be done at the top of the bed, and the liquid can only be removed from the bottom, which is not necessarily the best arrangement. If all the microbial cells are suspended, the concentration profiles of viable biomass and contaminant will be as shown in Figure 3d. In real trickle beds some of the biomass remains attached, which is a great advantage over tray columns when a nonvolatile metabolic product is generated, because it allows a large

liquid outflow to remove the product without the danger of washing out the biomass. (Mathematically, D is small because z is small, even though f is large.) Biotrickling filters can also be operated in the gas-downflow, cocurrent mode, which gives concentration profiles (again assuming suspended biomass) more like those shown in Figure 3c. Now, the microorganisms moving down the bed experience an ever-worsening environment for metabolism, with lower S and higher product concentration S_m. They leave the bottom of the bed in an endogenous, low-viability state, which is very suitable for the liquid outflow stream, but which may cause a significant lag phase when cells are recycled to the better environment at the top of the bed. The question of whether the cocurrent mode works better can only be answered definitively by experiments on a particular gas stream.

In biofilters the biophase is far from being completely mixed. It consists of an immobilized mass of up to xXe_b viable cells per unit bioreactor volume (less if cell growth has been limited by some nutrient other than the contaminant), and one of the goals of media selection is to provide sufficient surface area so that these cells can be distributed as a biofilm thin enough to avoid mass transfer limitation due to diffusion through the film. However, simple models (38) suggest that biofilters should share the minimum effluent partial pressure, HS_s, predicted earlier for bioreactors with a completely mixed biophase, because biofilms exposed to any lower partial pressures near the biofilter effluent would be in an endogenous state and eventually die. Experience suggests otherwise, probably because the adsorption of the contaminants on the media increases their effective concentration in the microorganisms' immediate environment. It follows that highly adsorptive support particles, such as compost or activated carbon, will work best, and that biofilters can treat very insoluble, hydrophobic contaminants if they adsorb well.

Scale-Up

First-stage scale-up experiments can be done in a reactor of any size that is convenient or available, as long as it is large enough so that the flows of gas and liquid are not significantly altered by entrance effects, channeling, or wall effects. Ideally, several different bioreactor types or packing materials are studied, but this is not always possible. The objective is to vary the gas flow, g, the liquid flow, f, nutrient addition, Sji, and so on, to determine how the specified contaminant removal can be achieved most economically. The scale-up problem is then to extrapolate the results in order to calculate the necessary volume (volume = gas flow to be treated/g) and shape (specifically the height, L) of the commercial-scale bioreactor, which must satisfy the process requirements without further experimentation immediately after start up. The best combination of process variables found in the experiments must be close to the optimum identified by the point at the right end of the operating region in Figure 2. To preserve this solution at the commercial scale, the dimensionless numbers D, F, and G that define this point must be kept the same between scales. Scaling up a gas treatment bioreactor is thus an example of Reynold's similarity principle, which states that two physical situations are similar only if the all of the relevant dimensionless numbers are identical.

Consider first the case where there is no inhibitory metabolic product and the process needs little or no liquid outflow. $F(\rightarrow\infty)$ and $D(\rightarrow 0)$ both drop from the list of relevant dimensionless numbers, and the scale-up problem reduces to finding combinations of gas flow, g, and bioreactor height, L, that keep G constant between scales. Given a mass-transfer correlation of the form of equation 7, G can be written $ARTL^c/Hg^{(1-c)}$. The problem is illustrated graphically in Figure 4, where lines of constant G are drawn on a log–log plot of g versus L. If the small-scale experiments have found different combinations of g and L that produce the specified contaminant removal, x, then this data should fall on such a line, labeled *empirical* in Figure 4, because the theory suggests that "constant x" implies "constant G." Also shown in Figure 4 is an upper limit on the gas superficial velocity, $U_g = gL$, caused by physical phenomena such as slugging of bubble columns, fluidization of the media in upflow biofilters, or flooding of the impeller in stirred tanks. This limit may be slightly different at different scales. Two extreme cases of the scale-up problem can be identified.

For a biofilter or a biotrickling filter treating sparingly soluble contaminants, the mass-transfer factor is little affected by the gas velocity, implying that $c = 0$ and making the empirical line horizontal. Several important phenomena are lumped into the parameter A, but there is no reason to expect them to vary with scale because microorganisms cannot "know" if they are in a small or large bioreactor. Consequently, performance on the large scale is defined by the same horizontal line, meaning that scale-up is done at constant g, or equivalently at constant gas-residence time ($=1/g$). Because $g = U_g/L$, constant g can be achieved either with a deep bed and a high gas velocity, the right end of the $c = 0$ line in Figure 4, or with a shallow bed and a low velocity, the left end of the line. Although the former are easier and cheaper to build, they have a much higher pressure drop though the bed, meaning higher costs for gas compression and possibly subsequent

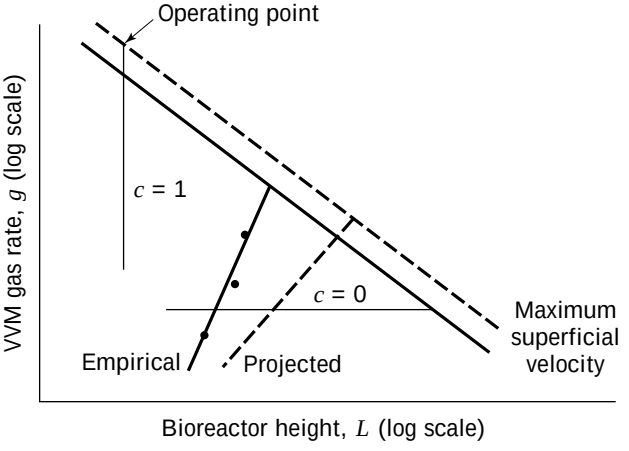


Figure 4. Dotted line, large scale; solid line, small scale.

cooling to return the gas to a temperature tolerable to the microorganisms. The latter are much cheaper to operate and work well as long as the beds are deep enough (at least several hundred times the particle size) to avoid problems of channeling and short-circuiting of the gas flow. Many biofilters consist of shallow beds of media on trays arranged in various configurations in a reactor designed to ensure an even flow of gas into the beds.

At the other extreme is the ideal bubble column described in the section "Gas Biophase Mass Transfer" for which $c = 1$. The value of G, and thus the contaminant removal, is now independent of gas flow but proportional to bioreactor height. Scale-up must therefore be done by keeping L constant but increasing the reactor width, the large-scale bioreactor operating at maximum g in order to minimize its volume (the *operating point* in Figure 4). Although this ideal bubble column is an abstraction, it illustrates a paradox that can arise in scaling up practical bubble columns (for which $c \cong 0.7$); if the desired removal can be achieved in a small-scale column, then a bubble column is probably not the best solution at the commercial scale. For very soluble contaminants, the L needed to obtain the desired fractional removal may be in the range of laboratory reactors, around 1 m. Scale-up would produce a very unusually shaped large-scale bioreactor, a similar height but several meters wide. There would be no point in making it any deeper, except to provide a safety factor in the design, and it could not be made any narrower and still meet the process requirements. For insoluble contaminants the necessary L may be several meters. The conclusion from laboratory experiments would then be that a bubble column is not a feasible bioreactor, even though the desired contaminant removal may be achievable (albeit with a high pressure drop) with the reactor heights used at the commercial scale.

All practical situations fall between these two extremes, so neither simple scale-up rule, constant g or constant L, is necessarily appropriate. The empirical line will have a positive slope, and a similar line for the same G at the larger scale can be projected from it if any variations of the parameters A and c with scale can be estimated from the published literature, experience, consultations with equipment manufacturers, and so on. The design of the large-scale bioreactor can be anywhere along this projected line in Figure 4, the exact design point again being constrained by the maximum U_g line, but essentially an economic trade-off between the large bioreactors demanded by a small g and the large pressure drops associated with a large L. The essential point is that the shape, and the size of the large-scale bioreactor are fixed by this procedure and should not be chosen based simply on guesswork or convenience.

Some bioreactor types allow design flexibility because they have variables that can be adjusted when drawing the projected line. Both A and c tend to decrease with scale in any bubble-gasified reactor due to poorer gas dispersion, but in mechanically agitated tanks, A is a function of the power input per unit volume, which can be varied to some extent. For trickle beds and biofilters, both A and the bed permeability are predictably related to packing size, allowing the packing to be chosen to give the large-scale reactor both a convenient shape and a reasonable pressure drop. Note that, whether or not such changes are made between scales, the estimates on which the projected line in Figure 4 is based should always be conservative in order to provide a factor of safety in the large-scale design.

When product inhibition is significant, the liquid outflow, f, becomes a critical variable, and all three dimensionless numbers D, F, and G must be kept constant if the physicochemical environment of the biomass (S, S_m) and the optimum bioreactor performance established in the small-scale experiments are to be extrapolated to the large scale. An exact solution is possible only when the retention of the biomass by cell immobilization or recycle (the z parameter) can be controlled, although when retention is almost complete ($z \to 0$) only the product DF (from which z cancels) is important, implying that the liquid flow, f, must be kept proportional to the gas flow, g. Biotrickling filters have an extra dimensionless number, the liquid recycle ratio, RR, which introduces the extra difficulty that increasing the bed height, while keeping f and RR constant increases the liquid superficial velocity [$U_l = Lf(1 + RR)$], thus changing the mass-transfer situation and possibly flooding the bed. The procedure just described will not work in these more complex situations (constant x no longer implies constant G), but the general considerations still apply, including the physical limits on the gas superficial velocity, the desirability of tall small-scale bioreactors to reduce uncertainties, the need for compromise between capital and energy costs, and the possibility of predicting a very inconvenient aspect ratio for the large-scale bioreactor. Judgment and calculation are needed in each case because the trivial solution to the scale-up problem—"build a bigger bioreactor with the same shape, f, and g as the small one"—is unlikely to be the best. Gas treatment bioreactors are one of the few devices whose performance tends to improve with scale, but taking advantage of this tendency in practice is not straightforward.

NOMENCLATURE

a	Gas–biophase interfacial area per unit bioreactor volume
A,c	Constants in mass-transfer correlation
D	$zf/e_b Yk$ = dimensionless liquid flow rate
e	Hold-up
f	Liquid outflow rate per unit bioreactor volume
F	$YS'_m/zy_m X$ = dimensionless product-inhibition number
g	Volumetric gas outflow rate per unit bioreactor volume (the VVM gas rate)
G	$k_1 aRT/Hg$ = dimensionless mass-transfer-effectiveness number
h	Gauge pressure (atm) at gas inlet
H	Henry's law constant
k	Cell maintenance coefficient
k_1	Gas-to-biophase mass-transfer coefficient
L	Bioreactor height
n_i	Mole fraction of contaminant in feed gas

p Partial pressure (atm) of contaminant in effluent
q Specific consumption rate of contaminant
R Gas constant
S Dissolved concentration in biophase
S_s Stationary phase contaminant concentration (concentration at which $\mu = 0$)
S'_m Product concentration that totally inhibits metabolism
T Outlet gas temperature
U_g Superficial gas velocity at bioreactor outlet
U_1 Superficial liquid velocity in a trickle bed
X gn_i/ke_bRT = maximum possible cell concentration in biophase
x $1 - p/n_i$ = fractional removal of contaminant
w Evaporation rate per unit bioreactor volume
y Mole of nutrient consumed or product formed per unit of contaminant metabolized
Y Cell-yield coefficient
z Cell concentration in liquid effluent/cell concentration in biophase
γ Available electrons per carbon equivalent of biomass
μ Specific growth rate of biomass

Subscripts

i Inlet conditions
j Any nutrient or product
m Metabolic product
x Biomass

BIBLIOGRAPHY

1. G. Leson and A.M. Winer, *J. Air & Waste Management Assoc.* **41**, 1045–1053 (1991).
2. D.S. Hodge, V.F. Medina, R.L. Islander, and J.S. Devinny, *Environ Technol.* **12**, 655–662 (1991).
3. W.A. Apel, W.D. Kant, F.S. Colwell, B. Singleton, B.D. Lee, G.F. Andrews, A.M. Espinosa, and E.G. Johnson, *Emerging Technologies in Hazardous Waste Management IV*, American Chemical Society, Washington, D.C., 1994, pp. 142–159.
4. S. Dharmavaram, J. Casey, and C. van Lith, *Proc. Air Waste Management Association*, 93-WP-52C.01, Denver, Colo., June 13–18, 1993.
5. I. Togashi, M. Suzuki, M. Hirai, M. Shoda, and H. Kubota, *J. Ferment. Technol.*, **64**, 425–432 (1986).
6. W.A. Apel, P.R. Dugan, M.R. Wiebe, E.G. Johnson, J.H. Wolfram, and R.D. Rogers, *Emerging Technologies in Hazardous Waste Management*, vol. 3, American Chemical Society, Washington, D.C., 1993, pp. 411–428.
7. W.A. Apel, B.D. Lee, M.R. Walton, L.L. Cook, and K.B. Dinerstein, *Proceedings of the Air & Waste Management Association*, 95-TP9C.07, San Antonio, Tex., June 18–23, 1995.
8. P. Plumb, K.H. Lee, and K.L. Sublette, *Appl. Biochem. Biotechnol.* **24/25**, 785–797 (1990).
9. J. Croonenberghs, F. Varani, and P. Le Fevre, *Proceedings of the Air & Waste Management Association*, 94-RP115B.06., Cincinnati, Ohio, June 19–24, 1994.
10. Y. Yang, A.P. Tongas, and J.R. Blunk, *Proceedings of the Air & Waste Management Association*, 94-RA115A.04., 1994.
11. R. Whittenbury, K.C. Phillips, and J.E. Wilkinson, *J. Gen. Microbiol.* **61**, 205–218 (1970).
12. C.L. Haber, L.N. Allen, S. Zhao, and R.S. Hanson, *Science.* **221**, 1147–1153 (1983).
13. J. Pilkington, and H. Dalton, *Methods in Enzymology*, vol. 88, Academic, New York, 1990, pp. 181–190.
14. J. Colby, D.I. Stirling, and H. Dalton, *Biochem. J.* **165**, 395–402 (1977).
15. D. Scott, J. Brannan, and I. Higgins, *J. Gen. Microbiol.* **125**, 63–72 (1981).
16. Stanley, S. Prior, D.J. Leak, and H. Dalton, *Biotechnol. Lett.* **5**, 487–492 (1983).
17. R. Oldenhuis, R.L.J.M. Vink, D.B. Janssen, and B. Witholt, *Appl. Environ. Microbiol.* **55**, 2819–2826 (1989).
18. V. Hecht, D. Brebbermann, P. Bremer, and W.D. Deckwer, *Biotechnol. Bioeng.* **47**, 461–469 (1995).
19. B.R. Folsom, P.J. Chapman, and P.H. Pritchard, *Appl. Environ. Microbiol.* **56**, 1279–1285 (1990).
20. D. Arciero, T. Vannelli, M. Logan, and A.B. Hooper, *Biochem. Biophys. Res. Commun.* **159**, 640–643 (1989).
21. L.P. Wackett, G.A. Brusseau, S.R. Householder, and R.S. Hansen, *Appl. Environ. Microbiol.* **55**, 2960–2964 (1989).
22. M. Kastner, *Appl. Environ. Microbiol.* **57**, 2039–2046 (1991).
23. J.N. Petersen, R.S. Skeen, K.M. Amos, and B.S. Hooker, *Biotechnol. Bioeng.* **43**, 521–528 (1994).
24. J.M. Barnes, W.A. Apel, and K.B. Barrett, *J. Hazard. Mat.* **41**, 315–326 (1995).
25. W.A. Apel, J.M. Barnes, and K.B. Barrett. *Proceedings of the Air & Waste Management Association*, San Antonio, Tex., June 18–23, 1995.
26. Bohn H.L., *J. Air Pollut. Control Assoc.* **25**, 953–95 (1975).
27. M. Shields, M. Reagin, R. Gerger, C. Somerville, R. Schaubhaut, R. Campbell, and J. Hu-Primmer, in R.E. Hinchee, A. Leeson, L. Senprini, and S.K. Ong eds., Bioremediation of Chlorinated and Polycyclic Aromatic Compounds, Lewis, Ann Arbor Mich. 1994, pp. 50–65.
28. R.H. Perry and C.H. Chilton, *Chemical Engineers Handbook*, 6th ed., McGraw Hill, New York, 1988.
29. H.M. Lizama and E.N. Kaufmann, in D.S. Holmes and R.W. Smith eds., Minerals Bioprocessing II, TMS, Warrendale, Penn, 1995, pp. 241–253.
30. B.D. Ensley, Dept. of Energy Report DOE/CH-9207, 1992.
31. Z. Shareefden and B.C. Baltzis, *Chem. Eng. Sci.* **49**, 4347–4360 (1994).
32. R.M. Diks and S.P. Ottengraf, *Bioprocess Eng.* **6**, 93–99 (1991).
33. S. Lamare and M. Legoy, *Trends in Biotechnology* **11**, 413–419 (1993).
34. R.M. Counce and J.J. Perona, in N. Cheremisenoff ed., *Handbook of Heat and Mass Transfer*, vol. 2, Gulf, Houston, Tex., 1986, pp. 953–966.
35. Y.N. Lee and S.E. Schwartz, *J. Phys. Chem.* **85**, 840–848 (1981).
36. G.F. Andrews, Appl. Biochem. Biotechnol. **51/52**, 329–338 (1994).
37. C. Oncharit, Y.T. Shah, and K.L. Sublette. Chem. Eng. Sci. **45**, 2383–2389 (1990).
38. G.F. Andrews and K.S. Noah, *Biotechnol. Prog.* **11**, 498–509 (1995).
39. K.T. Klasson, M.D. Ackerson, E.C. Classen, and J.L. Gaddy, *Fuel* **72**, 1673–1678 (1993).

BIOREACTORS, MAMMALIAN CELLS. See

HYPOXIA, EFFECTS ON ANIMAL CELLS; MAMMALIAN CELL BIOREACTORS.

BIOREACTORS, PHOTO

MARIO R. TREDICI
University of Florence
Florence, Italy

KEY WORDS

Microalgae mass cultivation
Photobioreactors
Phototrophic microorganisms

OUTLINE

INTRODUCTION

Photobioreactors are bioreactors in which phototrophic microorganisms or plant cells are grown or used to carry out a photobiological reaction. In a broad sense, the open shallow basins used for microalgae cultivation could also be viewed as photobioreactors (1). In this article, however, the term *photobioreactor* is used to indicate only closed systems, that is, reactors that do not allow direct exchange of gases (e.g., CO_2, O_2) or contaminants (dust, microorganisms, etc.) between the culture and the atmosphere.

At present, commercial production of phototrophic microbial biomass is limited to two microalgae (*Dunaliella* and *Chlorella*) and one cyanobacterium *(Arthrospira)*. Mass cultivation of these species is carried out in open ponds by means of a selective environment (high bicarbonate and high pH for *Arthrospira*, high salinity for *Dunaliella*), or by making use of large amounts of inoculum produced under strictly controlled conditions and taking advantage of a high growth rate *(Chlorella)*. Most microalgal species cannot be maintained long enough in outdoor open-culture systems because of contamination by heterotrophic microorganisms (fungi, bacteria, protozoa) and competition by faster-growing microalgae that tend to dominate, regardless of the species used as inoculum. With the few exceptions mentioned earlier, virtually all attempts at mass cultivation of specific species of phototrophs in outdoor open ponds for prolonged periods have failed. An alternative choice is to use photobioreactors, which can provide an environment that is protected from direct fallout, are relatively safe from invasion by competing microorganisms, and in which conditions (pH, pO_2, pCO_2, temperature) can be better controlled so that dominance of the desired species is established for a sufficient period of time.

The use of photobioreactors could extend the exploitation of microbial phototrophs from the few strains commercially grown today to the more than 50,000 known species, many of which are potential sources of high-value products. Examples of such products are specialty lipids (sterols and polyunsaturated fatty acids); antifungal, antibacterial, or antiviral compounds; radioactively or isotopically labeled molecules (such as ^{13}C-labeled glucose and amino acids); vitamins (e.g., vitamin E); pigments (such as β-carotene, astaxanthin, and phycobiliproteins); pharmaceuticals (e.g., antineoplastic agents); and aquaculture feeds. The latter represent a good example of a case that could benefit in a decisive manner from the introduction of new photobioreactor concepts.

Understanding photobioreactors requires background knowledge on the fundamentals of the growth of phototrophs. The first part of this article therefore describes the principal characteristics of phototrophic microorganisms and their growth kinetics. The second part deals with photobioreactor designs, advantages, limitations, and commercial applications.

The topic of photobioreactors has also been reviewed by Lee (2), Chaumont (3), Prokop and Erickson (1), and more recently by Torzillo (4), Tredici and Chini Zittelli (5), and Pulz and Scheibenbogen (6).

PRINCIPAL CHARACTERISTICS OF PHOTOTROPHIC MICROORGANISMS

Photosynthesis and the Photosynthetic Apparatus

All organisms require a source of energy for growth to take place. Phototrophic organisms, or phototrophs, use light as

an energy source. Light, or visible radiation, is the electromagnetic radiation in the 400- to 700-nm range, to which the human eye is sensitive. Because it is used by algae and plants for photosynthesis, visible radiation is also called photosynthetic active radiation (PAR).

Photosynthesis is the process by which phototrophs capture and convert the energy of photons (quanta of electromagnetic radiation, i.e., the smallest "packets" of light energy) to biochemically usable energy. It is through photosynthesis that phototrophs obtain the energy (ATP) and the reducing power (NADH or NADPH) necessary to incorporate CO_2 and to manufacture the organic molecules they require for growth.

Photosynthesis is generally associated with higher plants, yet about half the photosynthesis on earth is carried out by microorganisms. There are three major groups of phototrophic microorganisms: the photosynthetic purple and green bacteria, the cyanobacteria (also referred to by botanists as blue–green algae), and the microalgae. Photosynthetic bacteria, and cyanobacteria are prokaryotes; microalgae have a eukaryotic-cell-type organization. Photosynthetic bacteria differ from cyanobacteria and microalgae because of their photopigments and photosynthetic processes.

The ability to perform photosynthesis depends on the presence of light-absorbing pigments, the chlorophylls: large planar molecules composed of four substituted pyrrole rings with a magnesium atom coordinated to the four central nitrogen atoms. Chlorophylls absorb light primarily in the blue–violet and red regions of the spectrum and absorb weakly in the green region; this is why many phototrophs are green in color. Photosynthetic bacteria differ from cyanobacteria and microalgae in that their chlorophylls, called bacteriochlorophylls, have different molecular structure and absorption characteristics. For example, bacteriochlorophyll *b* of some purple bacteria has an in vivo maximum absorption in the infrared region between 1,020 and 1,040 nm. For purple bacteria, photosynthetic active radiation covers a much wider wavelength range than it does for microalgae and cyanobacteria, from the near ultraviolet to the long-wave infrared region (Fig. 1).

Besides the chlorophylls, other photosynthetic pigments, called accessory pigments, are also involved in the capture of light. Accessory pigments absorb light in the blue–green through yellow–orange range where chlorophylls absorb poorly and transfer the trapped energy to the chlorophylls, making photosynthesis more efficient over a broader range of wavelengths. The most widespread accessory pigments are the carotenoids, fat-soluble molecules that absorb light in the blue region of the spectrum, and the water-soluble phycobiliproteins of red algae and cyanobacteria that absorb in the green–yellow region. The accessory pigments also protect phototrophs from sunlight, which can damage the photosynthetic apparatus when it is too intense.

Cyanobacteria and microalgae also differ greatly from photosynthetic bacteria because of their photosynthetic process. Like higher plants, cyanobacteria and microalgae are oxygenic phototrophs; that is, they exhibit a type of photosynthesis that involves two photosystems (PSI and PSII), uses H_2O as electron donor, and evolves oxygen as

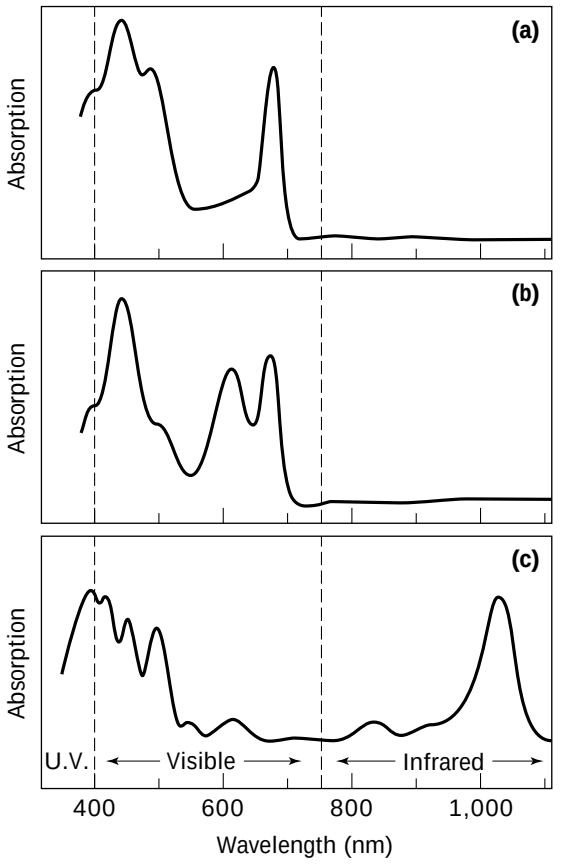


Figure 1. *In vivo* absorption spectra of (**a**) a microalga (*Scenedesmus acutus*), (**b**) a cyanobacterium (*Synechocystis* sp.), and (**c**) a purple bacterium (*Rhodopseudomonas viridis*).

a result of water photolysis. Green and purple photosynthetic bacteria lack photosystem II and cannot use water as an electron source; hence, they do not produce oxygen photosynthetically. For this reason they are called anoxygenic phototrophs. These definitions apply irrespectively of whether the organism obtains its cellular carbon principally from inorganic (autotrophically) or organic (heterotrophically) sources. Microalgae and cyanobacteria are mostly aerobic (they use oxygen as electron acceptor for respiratory metabolism) and autotrophic; purple bacteria are mainly anaerobic and heterotrophic; green bacteria are prevalently anaerobic and autotrophic. For a clear, comprehensive treatment of both plant and bacterial photosynthesis, see Hall and Rao (7).

Although this article deals exclusively with photobioreactors for the cultivation of oxygenic phototrophs (microalgae and cyanobacteria), some of the aspects covered and most of the conclusions can be extended to other phototrophic groups. For brevity, the term microalgae is used here in its botanical meaning, which includes cyanobacteria.

Response of Phototrophs to Light Quality and Quantity

The rate of photosynthesis achieved by a phototrophic cell depends on the rate of photon absorption and on the effi-

ciency with which the absorbed photons are used. The rate of absorption and the efficiency of conversion of absorbed photons are in turn determined by the characteristics of the photosynthetic apparatus and by the spectral quality and intensity* of the incident light.

Photosynthesis as a Function of Light Quality. The wavelengths absorbed by a phototroph depend on its primary and accessory pigment content as shown by its absorption spectrum (Fig. 1). Only absorbed wavelengths are used, but not all the wavelengths absorbed are equally efficient for photosynthesis. Therefore, the spectral suitability of a light source for a phototroph is better represented by the *action spectrum* of photosynthesis (i.e., the curve obtained by plotting the *specific rate of photosynthesis [P]*† against wavelength across the photosynthetic range for the cultured organism). Although action spectra give more information than absorption spectra, the latter are more commonly reported because they are easier to determine and interpret. The curve obtained when growth, rather than photosynthesis, is plotted against wavelength is called a *growth action spectrum*.

Photosynthesis as a Function of Light Intensity. The rate of photosynthesis is not simply proportional to the rate of photon absorption, since photons may be captured by the pigments much faster than the photosynthetic apparatus can make use of them. If photosynthesis is measured at different light intensities, the specific photosynthesis rate *(P)* can be plotted as a function of light intensity *(I)* in the well-known light response curve of photosynthesis, or the *P–I* curve (Fig. 2). In the dark, where there is, of course, no photosynthetic activity, O_2 is consumed and CO_2 is released due to cellular respiration. As the light intensity gradually increases, photosynthetic O_2 evolution (and CO_2 consumption) takes place and a light intensity value, called the light compensation point or compensation irradiance *(I_c)*, is eventually reached, at which value photosynthetic O_2 evolution balances respiratory O_2 consumption. Beyond this value, production exceeds consumption and net photosynthesis is achieved. From here on, *P* increases linearly with *I* up to a certain point, beyond which the rate of increase of *P* per increment in light intensity declines and eventually levels off. The slope of the linear portion of the curve, called alpha (α), represents the rate of photosynthesis per unit incident light or relative quantum yield (also called maximum photosynthetic efficiency or maximum quantum yield), and is a measure of the efficiency

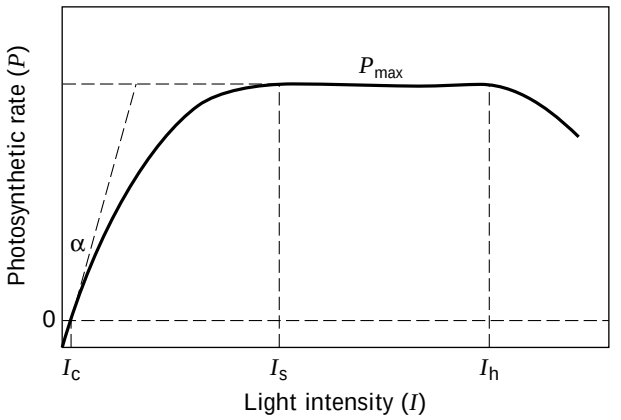


Figure 2. Light-response curve of photosynthesis (known as *P–I* curve). The intercept on the vertical axis is the measure of O_2 uptake due to dark respiration. I_c, light compensation point; I_s, light saturation intensity; I_h, light intensity value at which photoinhibition occurs. For details refer to the text.

with which the organism uses low light intensities. The relative quantum yield is expressed as moles of O_2 evolved per mole of photons of incident PAR. Plotting the absorbed light (incident light minus reflected and transmitted light) instead of incident light will give the absolute quantum yield. The reciprocal of the quantum yield (i.e., the number of photons required to evolve one molecule of oxygen) is called the quantum requirement of photosynthesis. In the range of intensities where *P* does not vary with *I*, photosynthesis is said to be light saturated and the light intensity is said to be saturating *(I_s)*; *P* reaches its maximum value (P_{max}) in this range. If light intensity further increases, *P* may decrease again. This phenomenon involves damage to some components of the photosynthetic apparatus and is referred to as photoinhibition. At low light intensities, *P* is limited by the light reactions of photosynthesis (thermodynamic constraint). At saturating light intensities, *P* (P_{max}) is limited by the so-called dark reactions, in which the products of light reactions are used to fix CO_2 (metabolic constraint). *P–I* curves are species specific and must be determined under conditions such that all the cells receive the same photon flux and are not mutually shaded.

Hill and Bendall's widely accepted Z-scheme of oxygenic photosynthesis incorporates two photosystems (PSII and PSI), and establishes that it is necessary to move four electrons through each photosystem to evolve one molecule of O_2 and fix one molecule of CO_2. Since the transfer of each electron requires two photons (one in PSII and one in PSI), it follows that the theoretical minimal quantum requirement cannot be lower than 8 (i.e., not less than eight photons are needed to move four electrons from water to NADP through the two photosystems) and the maximal quantum efficiency cannot be higher than 0.125 (1/8).

The curve relating growth to light intensity has a pattern similar to the *P–I* curve (with growth rate increasing proportionally to light intensity when intensity is limiting and being independent of intensity when saturating values are reached), since phototrophic growth is a direct consequence of the photosynthetic rate.

*The spectral quality of light depends on the wavelength of each single photon. The energy content of a photon varies with its wavelength according to $\epsilon = hc/\lambda$, where ϵ is energy content, h is Planck's constant, c is the speed of light, and λ is the light wavelength in nm. Light intensity is given by the number of photons impinging on unit area in unit time (expressed as μmol photons m^{-2} s^{-1}). More properly the incident photon flux on a unit area in unit time is defined as photon flux density (PFD).

†Photosynthesis can be measured in terms of either CO_2 uptake or O_2 release. Rates of photosynthesis can be expressed per unit biomass, or any quantity proportional to it, to obtain the specific photosynthetic rate *(P)*. Typical units for *P* are μmol evolved O_2 (or fixed CO_2) mg Chla^{-1} h^{-1}.

Growth of Microbial Phototrophs in Batch Culture

Phototrophic microorganisms can be grown in batch or continuous culture. A batch culture is initiated by the transfer of a small amount of cells (inoculum) into fresh growth medium; after the inoculum, nothing is added to (except light and carbon dioxide), or removed from (except oxygen), the system.

Growth Rate. Growth can be defined as the orderly increase of all the cell constituents. During this process, unicellular organisms grow to a maximum size, then divide to form two daughter cells. The daughter cells in turn repeat the process, and the number of cells doubles at regular intervals of time. The time required for a growing population to double is called the generation time, g. The rate of increase in biomass concentration is generally expressed as the specific growth rate, μ, which is related to the generation time by $\mu = \ln 2/g$. The specific growth rate, which can be defined as a proportionality constant for the autocatalytic growth reaction, is described by the following equation:

$$\mu = 1/X(dX/dt) \qquad (1)$$

where μ is the specific growth rate (h^{-1}), X is the biomass concentration (g L^{-1}), t is time (h), and dX/dt is the variation in biomass concentration with respect to time (g L^{-1} h^{-1}), that is, the culture productivity per unit volume. The volumetric productivity ($dX/dt = \boldsymbol{\mu}X$) of a culture in a pond or a flat reactor can be easily normalized on a per-unit-area basis by considering the culture volume *(V)* corresponding to the unit of illuminated surface area *(A)*. Therefore, the areal productivity can be calculated as $\mu XV/A$, or, since V/A is the depth d of the culture, as μXd.

Growth Curve. When a phototroph grows under suitable conditions in batch culture, its growth curve follows a sigmoidal pattern showing six principal phases (Fig. 3a). Each phase is characterized by a particular value of the specific growth rate (Fig. 3b) and reflects a particular metabolic state of the cell population. After the inoculum, growth does not necessarily start right away since most cells may be viable, but not in condition to divide, especially when the parent culture is old. The interval necessary for the transferred cells to metabolically adapt to the new situation and start to grow is called the lag phase. After the lag phase, the culture enters into the acceleration phase, during which μ increases continuously. When a constant growth rate is reached, the culture is said to be in the exponential growth phase. During exponential growth, all the cells are light saturated (if a saturating light intensity is provided at the culture surface) and growth proceeds according to equation 1. The maximum μ is achieved in this phase, and in this respect phototrophs do not behave differently from chemotrophs. (*Note*: Organisms that obtain energy for growth from oxyreduction reactions involving either organic or inorganic electron donors.) The exponential growth phase of phototrophs in batch culture normally lasts for a relatively short period because the cells start to shade each other as their number increases; growth becomes light limited and μ decreases (deceleration phase). Self-shading and the fact that the energy source is instantaneously provided at the culture surface and cannot be stored in the reactor are the two fundamental factors that differentiate photobioreactors from reactors for chemotrophs. During the deceleration phase, although μ decreases continuously, the culture productivity may remain stable at maximum value because of the continuous proportional increase in cell concentration *(X)*. Since the increase in cell biomass becomes almost linear, sometimes the initial part of this phase is called linear phase. Conditions of linear growth are typically sought in mass cultures of phototrophs, since they lead to the highest productivity. Following the linear growth phase, the cell population continues to increase, but μ decreases until it reaches zero, at which point the culture enters the stationary phase, during which the cell concentration remains constant at its maximum value. The stationary phase is followed by the death phase, in which the population of viable cells decreases.

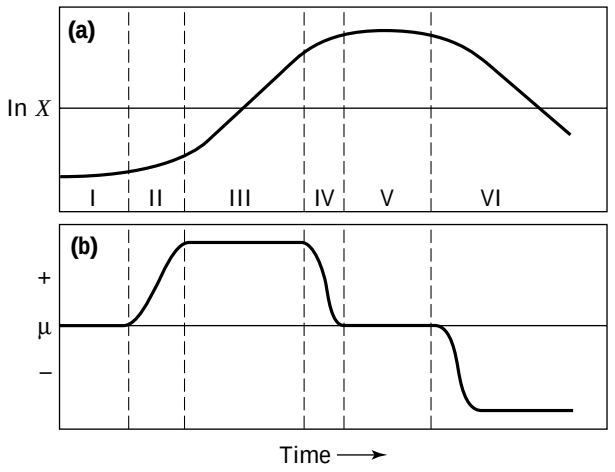


Figure 3. (**a**) Growth curve of a phototrophic microorganism in batch culture. The following six phases are shown: I, lag; II, acceleration; III, exponential; IV, deceleration; V, stationary; VI, death (after Ref. 8). (**b**) Changes in growth rate (μ) during the six typical growth phases.

To achieve continuous exponential light-saturated growth, the culture process has to be changed from batch to continuous; this is accomplished by continuously removing the newly formed cells. This condition is not adopted in photobiological process targeting production of biomass or related products, since it leads to low efficiency of conversion of the irradiance impinging on the culture surface and to low productivity (see later section).

Light distribution, Growth Rate, and Volumetric Productivity in Phototrophic Cultures

If nutrients are provided in suitable amounts and temperature is maintained at the optimum value, the productivity of phototrophic cultures becomes a function of light availability and cell concentration. In dilute cultures, self-shading is minimal and all the cells receive the same amount of light, independent of their position. It is self-

evident that, at these low densities, even if the light intensity provided would permit the maximum growth rate, the culture will achieve low efficiency of light conversion and low productivity, since most of the impinging photon flux will pass through the culture unabsorbed (Fig. 4, upper curve). Mass phototrophic cultures, in order to obtain maximum productivity, must be kept dense enough to absorb all the light impinging on their surface. Under these conditions, self-shading, due to light absorption by cell pigments and scattering by the cells causes light intensity to fall off exponentially with the culture depth (Fig. 4, lower curve). In dense cultures, the light availability for each single cell is much reduced; this brings about a significant diminution of the specific growth rate, but productivity is, in any case, higher than in dilute cultures because the decrease in growth rate is more than compensated by the increase in cell concentration. For achieving maximal productivity in mass cultures it is mandatory to operate at a proper combination of culture depth and cell concentration. This concept, known as operating at the optimal population density (OPD), has been investigated by many researchers, most recently by Richmond and coworkers (9–11). In general, maximal productivity is obtained at relatively high population densities and at specific growth rates that are about one-half of the maximum. Increasing the cell concentration above the OPD is generally associated with a reduction in productivity because of the presence of a large dark zone inside the reactor, which leads to high maintenance requirements in relation to the overall photosynthetic activity of the culture. Given the nonuniformity of the light environment that dominates in dense cultures, one requirement for achieving high photosynthetic efficiencies is that the individual cells be intermittently moved into and out of the illuminated zone. Although it has been shown that conventional mixing can do little to improve photosynthetic efficiency through what is known as the "flashing light effect,"* mixing is, in any case, required to provide intermittent exposure of the cells to light and darkness. A means to increase volumetric productivity in phototrophic cultures without increasing the light intensity at the culture surface is to shorten the light path by increasing the illuminated surface-to-volume ratio (S/V ratio) of the reactor.

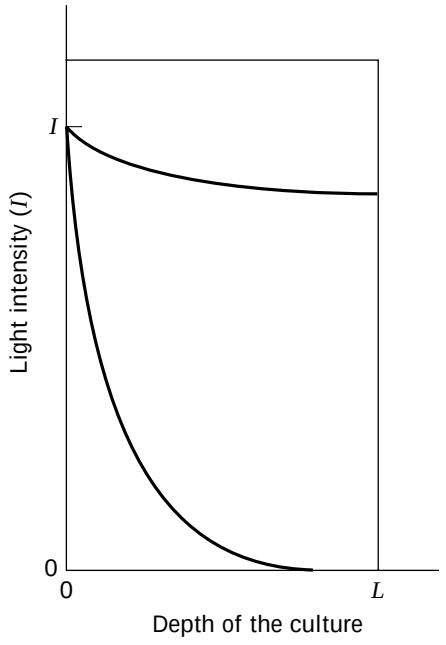


Figure 4. Light extinction in a phototrophic culture illuminated from one side with saturating light intensity. The upper curve represents a condition of low-cell concentration in which light absorption is negligible and all the cells receive almost the same light intensity. The lower curve describes a condition of high cell density in which the impinging photon flux is fully absorbed at two-thirds of the culture depth.

Modeling Photosynthetic Growth in Photobioreactors

Mass phototrophic cultures are normally light limited, even though the light intensity striking their surface may be several times above saturation. For example, in outdoor algal cultures at midday, cells in the top layers will receive light at intensities above I_s (they will be photosaturated), or may even be photoinhibited, but a considerable portion of the culture will be photolimited or even below the compensation point (bottom layers). The rate of photosynthesis will thus be different in different regions of the culture, since it depends on the local environment experienced by the single cells, but the photosynthetic rate of the culture as a whole will be an average of all the individual rates. In the same way, while for a single cell the instantaneous specific growth rate, which depends on the instantaneous rate of light absorption and of photosynthesis, varies depending on the position of the cell, the average specific growth rate of the culture depends on the average amount of light received by the cells. It is hence a function of the culture depth (or light path) and of the cell concentration, as well as of the intensity of the incident light.

The exponential decay of the intensity of light as it penetrates into dense cultures is generally described by Lambert–Beer's law. This equation, however, accounts for the sole absorption by pigments and neglects light scattering by cells. More sophisticated models, (e.g., the radiative transfer theory developed in astrophysics), have been proposed to obtain radiant energy profiles inside reactors with a reasonable accuracy (13). Models of photosynthetic growth are based on the assumption that growth rate is closely related to the light availability inside the culture. It is thus necessary to consider both the effect of the light intensity available at each point of the reactor on growth rate and how the light environment is modified by absorption and scattering. This creates a high degree of complexity for the modeling of phototrophic cultures. A complete treatment of the kinetics and energetics of photosynthetic growth in photobioreactors has been presented by Cornet et al. (13). Accurate models for calculating local and average irradiance inside photobioreactors as functions of the design and of the impinging irradiance, and for relating

*The phenomenon, first described by Kok (12), by which the photosynthetic efficiency is increased when cells are exposed to light flashes instead of continuous light.

growth rate and productivity to average irradiance have been developed by Molina Grima and coworkers (14,15).

Light Utilization and Productivity in Phototrophic Cultures

Photosynthetic Efficiency of Outdoor Cultures. The photosynthetic efficiency of phototrophic growth is defined as the energy stored in biomass per unit of light energy absorbed (9). This parameter is of fundamental importance in mass cultures, since high photosynthetic efficiency leads to high biomass production per unit of illuminated area. As mentioned earlier, according to the current model of photosynthesis, eight photons are required to transfer four electrons from water to NADP and bring about the reduction of one molecule of CO_2 to carbohydrate in the dark reactions of photosynthesis. Since one mole of photons in the visible range contains on average about 50 kcal of free energy, and since the energy stored in carbohydrate during photosynthesis is about 114 kcal/mol of reduced CO_2, the theoretical maximum photosynthetic efficiency is approximately 28.5% [114/(8 × 50)]. The biosynthesis of the other macromolecules making up the microbial cell (lipids, protein, nucleic acids) requires additional reducing power and ATP and hence additional photons per CO_2 incorporated; this brings the quantum requirement for growing cells to 10–12, rather than 8, and reduces the maximum efficiency to about 20%. In practice, this efficiency cannot be reached with sunlight because of several losses and limitations, which include the fact that PAR represents only 43% of the total solar radiation at the earth's surface, the loss of part of the incident radiation by reflection at the culture surface, and the fact that part of the absorbed energy is invested in cell maintenance. Another major limitation, often ignored, is represented by the light saturation effect, that is, the fact that the photosynthetic apparatus of most phototrophs saturates at only 10–20% of full sunlight (about 2,000 μmol photons m^{-2} s^{-1}). According to the equation derived by Bush ($f = I_s/I\,[\ln I/I_s + 1]$) (16), which expresses the fraction (f) of the thermodynamic efficiency that can be attained as a function of the incident light intensity (I) and the saturation intensity (I_s), the quantum efficiency comes very close to the theoretical at low light intensities (<100 μmol photons m^{-2} s^{-1}). When the light intensity approaches that of full sunlight, utilization drops to less than 20% and most of the light energy absorbed by the culture is dissipated as fluorescence and heat. There are not many options available to overcome the limitation imposed by the light saturation effect. In dense phototrophic cultures, mixing partially compensates for the limitation imposed by the light saturation effect, but does not overcome the substantial inability of phototrophs to use effectively high irradiances (17). Benemann (18) discussed the alternative concept of increasing the light intensity at which photosynthesis saturates by decreasing the amount of accessory pigments in the cells. This however, is not easy to achieve and may lead to another problem: Contaminating algae with higher pigment content, because of their ability to absorb more light in the light-limited zones of the culture, will overgrow the cultured alga. Based on studies that took this direction in the 1950s, Tredici and Chini Zittelli (19) have recently proposed spatial light dilution as a means to increase the photosynthetic efficiency of outdoor algal cultures (see later section). Considering all of these limitations, it is reasonable to assume that only about 4–5% of solar energy can be converted to the chemical energy of biomass, under even the best conditions (17). For locations characterized by an annual average insolation of 4,500 kcal m^{-2} d^{-1} (e.g., southwestern United States), and assuming a caloric content of the biomass of 5.4 kcal g dry wt^{-1}, this efficiency will result in a calculated average annual productivity between 33 and 42 g m^{-2} d^{-1}.

Productivity of Outdoor Cultures of Phototrophs. The productivity of outdoor cultures is one of most debated issues connected with mass cultivation of phototrophs. Extrapolations from efficiencies of light conversion obtained in the laboratory had led to predict annual productivities as high as 140 g m^{-2} d^{-1} (20), but actual productivities are only one-tenth or less these figures, with about 40 g m^{-2} d^{-1} being considered the highest value obtainable by outdoor algal cultures under favorable conditions (17,18). A common mistake made (more or less consciously) in this respect is that the results obtained on a small scale over short periods and under optimal conditions are extrapolated to all-year-round, full commercial scale. As a matter of fact, despite statements in the early years of algal biotechnology, large algal cultures are not much more productive than terrestrial systems, and generally are much less reliable. In the author's opinion, however, productivity should not be considered as being the decisive factor in determining the viability of algal biotechnology. It is true that productivity greatly influences the economy of the process, but other factors, like sustainability and reliability, are just as, or even more, important.

Several models have been proposed to predict the productivity of outdoor cultures of phototrophs both in open ponds (21,22) and in photobioreactors (13,15).

How Productivity of Phototrophic Cultures Is Expressed. In the case of algae cultivation in open ponds, it is common to evaluate productivity as areal yield (i.e., productivity per unit of occupied land area per unit time, expressed either as g m^{-2} d^{-1} or tons ha^{-1} $year^{-1}$). This is both practical, since it makes comparison with terrestrial systems easy and immediate, and correct, since areal productivity in ponds is proportional to the amount of light entering the system per unit illuminated area. In a pond this corresponds substantially to the land surface area occupied by the system.

With photobioreactors the situation is much more complex. Except in the cases of flat horizontal systems and tubular horizontal photobioreactors with no empty space between contiguous tubes (in both these cases the areal productivity referred to the occupied land is correct and easy to interpret), expressing productivity on a per land surface area basis is misleading. For example, in some horizontal serpentine bioreactors it is difficult to decide whether and how to compute the empty space between contiguous tubes and how to account for the fact that the reactor intercepts a different proportion of the radiation impinging on the horizontal, depending on the elevation of the sun and, hence, on the hour of the day. With elevated

systems, the inadequacy of the areal yield is even more evident, since the amount of radiation intercepted by a vertical or sharply inclined reactor is not related to the land surface area it occupies. Expressing the productivity of elevated systems on a per unit occupied land surface area basis (i.e., with reference to the horizontal projection of the reactor on the ground) may lead to high unrealistic figures (see later section). Productivity can also be referred to the illuminated surface area of the reactor (biomass produced per unit illuminated surface area per unit time, in g m^{-2} d^{-1}), but this parameter is also not adequate for all situations, since it can be difficult to define the correct illuminated surface area to which to refer productivity. In tubular reactors, for example, the illuminated surface area generally corresponds to half the total surface area of the reactor; using it as basis for calculation of productivity in most cases underestimates reactor performance. In some vertical reactors (e.g., large-diameter coils that receive a significant amount of radiation on their inner central part at midday, but not in the early morning nor in the evening), the illuminated surface area changes during the day with sun elevation. Another complication with vertical reactors is determined by the fact that they intercept a significant amount of diffuse radiation on the surface not directly illuminated by the sun, and this is not always taken into account.

For these reasons, photobioreactor performance, like that of fermenters, is frequently reported in terms of volumetric productivity (amount of biomass produced per unit reactor volume per unit time, commonly expressed in g L^{-1} d^{-1}). But volumetric productivity is a function of the number of photons that enter the unit reactor volume in the unit time and, as such, is dependent on the illuminated S/V ratio of the reactor. The higher this ratio, the higher the volumetric productivity. It is well known, for example, that small-diameter tubular reactors or thin flat systems achieve high volumetric productivities, even if they perform poorly. We should be aware of the fact that a volumetric productivity of 1 g L^{-1} d^{-1} assumes a completely different significance if obtained in a tubular reactor of 1 cm or of 5 cm diameter.

To overcome these limitations, Tredici and Chini Zittelli (19) suggested evaluating the performance of vertical photobioreactors on the basis of the photosynthetic efficiency achieved by the culture. Unfortunately this parameter is not always easy to calculate (e.g., in coiled reactors) and since it depends heavily on the irradiance (the lower the irradiance, the higher the efficiency) favors experiments carried out under low light intensities, such as in winter.

The productivity of a photobioreactor could be expressed on a per shaded surface area basis (the productivity of the reactor divided by the average ground surface area shaded by the reactor). This parameter, never suggested before, substantially reflects the efficiency of light utilization, but also takes the actual irradiance conditions into account and offers the possibility to compare very different culture systems. For example, the performance of an open pond and of a vertical coiled reactor could be compared on this basis, whereas it is rather difficult to envisage any other suitable comparison parameter for these two very different systems. The per shaded surface area parameter cannot be used alone, however, since in some cases (e.g., serpentine reactors) it does not give a correct evaluation of the efficiency of land utilization.

In conclusion, each method for expressing productivity has its own limitations, and complete evaluation of a reactor performance in terms of productivity will be achieved only by a combination of the different available methods. For example, expressing the productivity on the basis of unit shaded area, unit occupied area (including eventual empty spaces), and unit volume will allow evaluation (and comparison) of photobioreactors in terms of efficiency of light conversion, efficiency of land use, and efficiency with which the unit volume of reactor is used. This combined approach is highly recommended.

Artificial Light Sources. Solar energy is a relatively cheap energy source, but given the low photosynthetic efficiency attainable under full sunlight, and the fluctuations in supply determined by the weather, the time of the year, and the natural day/night cycle, the use of artificial light sources to grow phototrophs has been taken into serious consideration, at least for production of high-value compounds.

The use of artificial illumination offers the possibility of selecting light quality and intensity, as well as the length of the illuminated period, and thus of significantly enhancing the reliability of the production process. This, together with the higher photosynthetic efficiency attainable (a photosynthetic efficiency of 15% is not uncommon with artificial light of relatively low intensity), may compensate in many cases for the higher cost.

One of the fundamental factors guiding choice of the light source should be the efficiency of conversion of electric energy into PAR energy. This efficiency varies considerably with the type of lamp; for example, it is only about 5% in tungsten–halogen (incandescence) lamps, 10–15% for fluorescent tubes, and 20–25% for high-pressure sodium lamps. Only with xenon lamps and light-emitting diodes are very high efficiencies attained (50% and up to 80%, respectively). Considering both the efficiency of conversion of electrical energy into PAR and the efficiency of conversion of light energy into biomass, an overall efficiency of conversion of electrical energy into the energy of biomass between 0.7% and 2% has been attained by *Arthrospira* cultures grown under fluorescent lamps (13). Another aspect to consider is the spectral quality of the emitted light, which ought to match the action spectrum of the cultivated organism as closely as possible. For example, tungsten lamps have a low output in the blue region and are therefore not appropriate for oxygenic phototrophs. High-pressure sodium lamps are relatively efficient as regards electrical energy conversion but emit in a limited yellow–orange region. Practicality often prevails over efficiency considerations in the choice of an artificial light source. For example, fluorescent lamps are widely adopted, independently of the spectral quality of their emission, because they provide a relatively uniform irradiance over a large surface.

A SURVEY OF PHOTOBIOREACTORS

Classification of Photobioreactors

Photobioreactors can be classified on the basis of both design and mode of operation. In design terms, the main categories of reactors that can be envisaged are *(1)* flat or tubular; *(2)* horizontal, inclined, vertical, or spiral; and *(3)* manifold or serpentine. An operational classification of photobioreactors would include *(4)* air or pump mixed and *(5)* single-phase reactors (filled with media, with gas exchange taking place in a separate gas exchanger), or two-phase reactors (in which both gas and liquid are present and continuous gas mass transfer takes place in the reactor itself). Construction materials provide additional variation and subcategories—for example, *(6)* glass or plastic and *(7)* rigid or flexible photobioreactors.

Axenic photobioreactors are reactors operated under sterile conditions. Although a major characteristic of photobioreactors is their ability to limit contamination, it must be made clear that an effective barrier, and thus operation under truly sterile conditions, is not achieved, except in the few special designs developed expressly for that purpose. This is a major difference between photobioreactors and bioreactors for chemoheterotrophs, which typically operate under sterile conditions. Like fermenters, photobioreactors can be divided into laboratory photobioreactors used for research, pilot-plant photobioreactors used to scale-up and optimize the process, and production units. Unlike fermenters, there are only a few examples of commercial photobioreactors in operation. Some companies sell photobioreactors in a variety of sizes for use in teaching, research, and commercial applications. The emphasis in this article is on pilot-culture systems and the two commercial-scale plants realized to date. Laboratory systems, such as illuminated carboys and fermenters, immobilized cell systems, and optical-fiber reactors are not considered. Cell-immobilization techniques have been reviewed by Brouers et al. (23), whereas optical fiber reactors have been discussed by Pulz and Scheibenbogen (6).

Early Photobioreactors

The use of photobioreactors for microalgae mass culture dates back to the late 1940s and was a logical consequence of the investigations carried out with *Chlorella* on the fundamentals of photosynthesis. In those early days of algae culture, the emphasis was on continuous controlled cultivation; hence, closed reactors were given priority. Open systems were considered inappropriate to guarantee the necessary degree of control and optimization of the process.

One of the first closed systems was set up at Stanford University in California in the late 1940s and involved use of vertical tubular reactors (24). The units consisted of glass columns 1.8 m in height and 10 cm in diameter, constricted at the bottom to prevent algal settling. The system was operated turbidostatically, with medium addition being regulated by a photoelectric cell. Culture temperature was maintained constant by means of a cooling coil inserted into the reactor. Carbon dioxide–enriched air was injected at the bottom of the column to obtain mixing and pH/pCO_2 control. The system was experimented with in the cultivation of *Chlorella* under both artificial light and sunlight. Maximum volumetric productivity indoors was 0.48 g L^{-1} d^{-1}. Outdoors, productivity averaged 0.28 g L^{-1} d^{-1}, with a maximum of 0.35 g L^{-1} d^{-1}. The reason for the lower outdoor productivity was not clear.

In the early 1950s, the Arthur D. Little Company at the Massachusetts Institute of Technology, under contract with the Carnegie Institution of Washington, developed two basic types of photobioreactors: an array of 10 vertical rigid plastic tubes and a 50-m^2 horizontal flexible tubular unit (25). The vertical units, which were essentially the same as those operated at Stanford University, provided inoculum to the larger horizontal unit and were reported to operate satisfactorily, although no productivity data were actually presented. The horizontal tubular unit may be considered the first pilot plant for microalgae production. It consisted of 4-mm-thick polyethylene tubes that, once filled with algal culture up to a depth of about 8 cm, assumed a more or less elliptic shape of 1.2-m width. The tubes were joined at the ends to form a raceway 21 m in length, for a total surface area of about 56 m^2. The culture (*Chlorella*) was circulated by means of a centrifugal pump. The pumping system was not capable of generating circulation velocities over 7.5 cm s^{-1} and there was considerable settling of the culture. With the same pump and a tube width of 0.43 m, flow rates of 25 cm s^{-1} were achieved and problems of settling were eliminated. Leakage and contamination by *Chlorococcum* and protozoa were the major problems encountered in the operation of the system. Spray cooling proved inadequate for temperature control, and it was necessary to install a heat exchanger. Despite these difficulties, a 40-day run attained an average areal productivity of 9 g m^{-2} d^{-1}. The data collected with this pilot plant provided the material for an economic analysis of algae production in photobioreactors (26), which concluded that the cost of products obtained in photobioreactors was excessive in comparison to that of equivalent materials obtained using conventional methods. Cooling of the culture was found to be a major cost factor.

At about the same time in Japan, Tamiya and coworkers (27) developed a system consisting of a serpentine tubular reactor immersed in a water bath for temperature control and an external column for gas exchange. This system provided the basic design criteria for all subsequent, essentially identical, serpentine reactors (e.g., those adopted by Gudin in France and Molina Grima and coworkers in Spain) (see later section). The reactor was made up of 3-cm-diameter glass tubes joined to form a serpentine and had a total length of 33 m. The total culture volume was 40 L. The culture *(Chlorella ellipsoidea)* was circulated at a linear velocity of 15–30 cm s^{-1}. A separate column, in which a mixture of 5% CO_2 in air was injected at a rate of 1–1.5 L min^{-1}, provided mixing and gas exchange. Productivity was directly related to solar irradiance with the highest value, achieved at the maximum irradiance, of about 0.6 g L^{-1} d^{-1}. After several months of experimentation, the authors concluded that one of the main limitations was temperature control and that a significant increase of productivity could be attained through cultivation of algae exhibiting higher temperature optima

and through vigorous mixing aimed at increasing growth by providing intermittent illumination. These concepts still guide most of our research.

Following these initial studies in the 1950s, little work was carried out in this field for several decades. Two exceptions were the work of Setlik in Czechoslovakia (28) and Jüttner in Germany (29). The reasons for this almost total abandonment of research activity on photobioreactors are not clear, although the successful production of *Chlorella* in open ponds in Japan, the development of microalgae-based wastewater treatment in California, and later success with open-raceway systems for outdoor production of *Spirulina* (*Arthrospira*) and *Dunaliella* perhaps explain the relative lack of interest in closed systems. The difficulties experienced by the A.D. Little Co. and Fisher's negative conclusion in his economic analysis also contributed to discouraging further studies on the application of photobioreactors.

Recent Types of Photobioreactors

Several types of photobioreactors have been designed and experimented with since the work of Pirt and coworkers sparked renewed interest in closed systems in the early 1980s (20). Most of these are small-scale systems; some are at the pilot or prepilot stage of development. Only in a very few cases has scale-up to commercial size been attempted. For convenience, these systems are described herein by categories: *(1)* axenic photobioreactors, *(2)* vertical tubular photobioreactors, *(3)* serpentine photobioreactors, *(4)* manifold photobioreactors, *(5)* helical photobioreactors, and *(6)* flat photobioreactors. Small systems and reactors for which experimentation has been conducted prevalently indoors with artificial light are not considered.

Axenic Photobioreactors Setlik (28) and later Jüttner (29) and Krüger and Eloff (30) developed glass tubular reactors for the axenic cultivation of microalgae. Jüttner's system, which substantially reproduced Setlik's design, was an all-glass, 110-L vertical flat coiled photobioreactor that could be sterilized by steam. The reactor, made of 4-cm-ID tubes, had a total length of 79 m. The algal suspension was circulated by a glass centrifugal pump at a flow rate of 20 to 60 cm s^{-1}. The temperature was regulated through a 2-m water-jacketed section. Gas exchange took place in an external exchanger in which a CO_2–air mixture was injected at a rate of 300 L h^{-1}. The system attained a maximum productivity of about 0.5 g L^{-1} d^{-1} under artificial illumination. Cell damage caused by the centrifugal pump was a problem in the case of fragile cyanobacteria; other problems encountered were foam formation and cell adhesion to the glass walls. Jüttner concluded that cell adhesion was the most discouraging feature and suggested that cultivation ought to be limited to "nonsticky" algae. Wall growth and fouling are still today major limitations in photobioreactors, although the problem is not always recognized as such because experiments generally do not last long enough for these effects to become evident.

A similar 60-L system, composed of 12 5-cm-ID glass tubes each 2 m in length, coupled by U-bends to form a vertical flattened spiral, was operated by Krüger and Eloff (30), who reported a productivity of about 0.13 g L^{-1} d^{-1} with a nontoxic species of *Mycrocystis*.

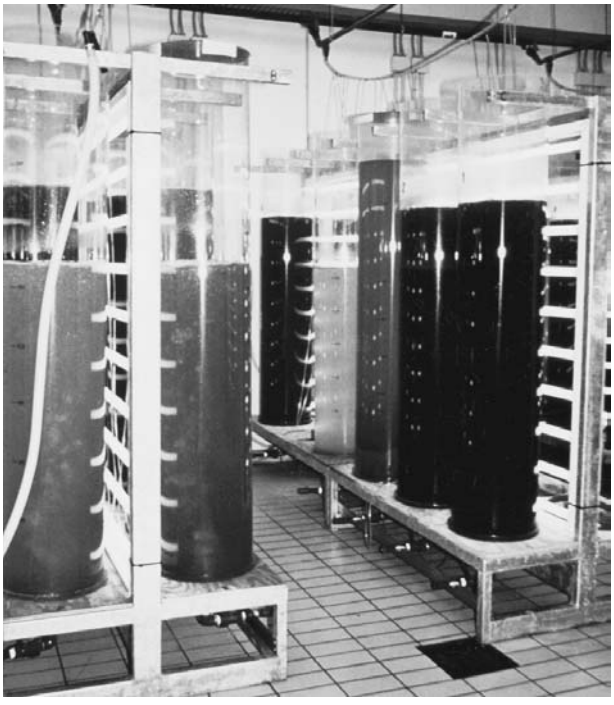
Soeder et al. (31) devised a rocking tray for the sterile cultivation of *Scenedesmus*. It consisted of a flat horizontal unit (inner dimensions 59.5 × 36 × 9 cm) made of stainless steel and covered with a glass lid. The tray was placed on a support connected to a rocking device. Temperature was controlled by circulating thermostated water through the double bottom of the unit.

Although these systems appear difficult to scale up and are rather expensive, they are mentioned here since with the exclusion of laboratory reactors they are the only systems that allow sterile cultivation of phototrophic microorganisms.

Vertical Tubular Reactors. Vertical tubular reactors (or column reactors) are simple systems in which mixing is achieved by injecting compressed air at the bottom of the column. Several vertical tubular photobioreactors were developed following Cook's first design. An example is the 30-L (2-m height, 15-cm ID) tower-type glass reactor devised by Jüttner (29). The main characteristic of this system was a wide, hollow glass finger inserted from above into the column to reduce dark space inside the reactor. Since the finger was open at the upper end, it could also be used to regulate the culture temperature. A thickness of the annular culture chamber of about 3.5 cm was found to be optimal for the growth of the light- and shear-sensitive *Microcystis aeruginosa*.

The rigid vertical columns, typically 2–2.5 m in height and 30–50 cm in diameter, that are extensively used in hatcheries to produce algal biomass for feeding the larval stages of marine bivalves and fishes must be included in this category (Fig. 5a). Most commonly, these systems are made of translucent fiberglass sheets formed into cylinders. Air is bubbled at the bottom for mixing, and either artificial illumination or natural light (or a combination of both) is provided. James and Al-Khars (32) cultivated *Chlorella* and *Nannochloropsis* in 200-L, 30-cm-diameter fiberglass cylinders of this type. Five such units were connected in series and operated continuously. Illumination was provided externally by a vertical bank of daylight fluorescent lamps. Although as a consequence of the low S/V ratio of the cylinders, mean volumetric productivities (0.17 g L^{-1} d^{-1} with *Chlorella* and 0.05 g L^{-1} d^{-1} with *Nannochloropsis*) were not high, exceptionally high productivities per unit occupied area were reported (up to 200 g m^{-2} land area d^{-1} for *Chlorella*). However, these values must not be taken as constituting a valid parameter for evaluation and comparison of this system, since they suffer from the limitation inherent in expressing the productivity of elevated reactors on a per occupied land area basis, as discussed earlier.

Cylinders illuminated from inside, which may be considered a variation of the annular reactor devised by Jüttner (29), have been proposed for the production of marine microalgae biomass in hatcheries (33). Although internally lit cylinders have high construction costs, they are less expensive to operate than completely filled columns. They also typically attain higher volumetric productivity and greater efficiency of light utilization since the photon flux



(a)

(b)

Figure 5. Vertical cylinders (**a**) and polyethylene sleeves (**b**) used in hatcheries to produce algal biomass for the larval stages of marine bivalves and fishes (photographs by G. Chini Zittelli).

provided to the culture is completely trapped by it. A patented internally illuminated annular airlift column is used by Martek Co. for specialty chemicals production from microalgae (34).

Miyamoto et al. (35) experimented with vertical tubular reactors (2.35-m height, 5-cm ID) made from the low-cost (about $2.50 per tube), mass-produced glass tubes used in the fluorescent lighting industry. Volumetric productivities of about 0.6 g L^{-1} d^{-1} were obtained with *Nostoc* and *Anacystis* in outdoor cultivation. Although these systems do not represent an innovation, they are materially useful because of their low cost and relatively high S/V ratio in comparison to previous reactors of this type. Using a similar system consisting of 32 glass tubes 1.5 m in height and with a 2.6-cm ID, Hu and Richmond (36) obtained a volumetric productivity of 1.6 g L^{-1} d^{-1} with *Isochrysis galbana* cultivated outdoors.

Transparent polyethylene tubing of various diameters and thicknesses is available in long rolls at very low cost. A disposable vertical reactor can be easily built by cutting a suitable length of tubing and heat-sealing one end. This bag or sleeve reactor can be suspended from a framework (Fig. 5b) or supported within a mesh frame. Such reactors are used indoors with artificial illumination (generally vertically mounted fluorescent lamps) or outdoors under sunlight. The temperature reached in the manufacturing of the polyethylene tubing is high enough to sterilize the material. These systems have a relatively short life because of biofouling but are easy and inexpensive to replace at the end of the culture run. Two-meter-long transparent polyethylene sleeves sealed at the bottom and hung on an iron structure have been used to grow *Porphyridium* and *Dunaliella* outdoors (in Israel), obtaining a significantly higher productivity compared to open ponds (37). The main drawbacks of sleeve reactors are the relatively low S/V ratio and the low efficiency of utilization of artificial light when it is provided from outside of the reactor.

Serpentine Photobioreactors. Serpentine photobioreactors are systems in which several straight culture tubes are connected in series by U-bends to form a flat loop that can be arranged either vertically or horizontally. Gas exchange and nutrient addition take place in a separate vessel. Circulation between the flat loop and the gas exchanger is achieved by the use of a pump or of an airlift. Several reactors of this type have been developed in the last two decades following the original design by Tamiya (27).

In the early 1980s, Gudin and coworkers (3,38), working at the Centre d'Etudes Nucléaires de Cadarache, France, devised (and since that time have further developed) a horizontal serpentine reactor the main characteristic of which is the use of a water pond to obtain temperature control (a method anticipated by Tamiya). After a first 1-m^2 version consisting of glass tubes of 3-cm diameter, a 10-m^2 and then a 100-m^2 plant were built using larger diameter flexible plastic tubes. The 100-m^2 system, experimented with from 1986 to 1989, was composed of five identical 20-m^2 units, each of which consisted of 20 tubes, 20 m in length and 6 cm in diameter, made of 0.25-mm-thickness low-density polyethylene. The total culture volume was about 7 m^3. Temperature control was attained by floating or immersing the reactor in a water bath by inflating or deflating a second layer of tubes placed below the reactor. Initially, the culture was circulated by means of a pump, but later airlift systems were adopted to limit damage to shear-sensitive cells and at the same time provide CO_2 supply and O_2 degassing. Productivities from 20 to 25 g m^{-2} d^{-1} were obtained with *Porphyridium cruentum* in a 2-month continuous run. The use of rigid rather than flex-

ible tubes in the more recent work by Gudin's group has permitted operating a self-cleaning system consisting of two plastic balls, one with higher density than the culture medium and one lighter, which are hydraulically pushed through the system. Although flotation immersion in a water basin can provide efficient thermoregulation, the cost of such a system is prohibitive for most applications.

In 1992, Dr. Gudin founded the Heliosynthese S.A. company for development of tubular photobioreactor technology in France. In 1996, Heliosynthese changed its name to Thallia Pharmaceuticals S.A. Thallia presently makes use of serpentine reactors to produce high-value algal compounds for the nutriceutical and pharmaceutical markets. The major products under development are polyunsaturated fatty acids, carotenoids (zeaxanthin from *Porphyridium cruentum*, astaxanthin from *Haematococcus pluvialis*) and antioxidants (superoxide dismutase).

Reactors of the type developed by Gudin are also being operated by Molina Grima and coworkers in Spain (Fig. 6a) (14,15,39,40). These systems consist of a horizontal loop, made of Plexiglas® tubes (about 100 m in length and 2.6 cm in diameter), which is connected to a 3- to 3.5-m-high airlift. The tubular solar receiver is immersed in a shallow pond for temperature regulation. Dissolved oxygen, pH, and temperature probes are connected to an online control unit and a computer for data acquisition. The Spanish group has made important improvements to the reactor originally purchased from Heliosynthese, and with it has carried out, and is still conducting, extensive research activity with the dual aim of shedding light on basic aspects of tubular photobioreactor technology (e.g., mass transfer and light distribution), and of defining guidelines for application of closed systems to production of specialty chemicals from microalgae (e.g., polyunsaturated fatty acids).

In the early 1980s, Pirt and coworkers developed a serpentine tubular reactor at the University of London's Queen Elizabeth College (20,41). The photobioreactor consisted of a photostage formed of 52 Pyrex glass tubes (each approximately 1 m long and with a 1-cm ID) horizontally stacked. The tubes were connected to Pyrex glass U-bends by silicone-rubber tubing to form a vertical loop. The loop outlet was connected through a vertical riser to a degasser. A second tube connected the degasser to the inlet of the loop. The reactor covered an area of about 0.5 m^2. The culture suspension (about 4.6 L) was circulated by either a peristaltic pump, a rotary positive-displacement pump, or an airlift; the latter method was preferred because of the adverse effects observed with the two pumps. With a defined consortium consisting of a chlorella-type green alga and three chemotrophic bacteria, cell concentrations in excess of 20 g L^{-1} and a very high productivity of 2.2 g m^{-2} h^{-1} were achieved at a light intensity of only 38 W m^{-2} (PAR). The work carried out by Pirt and coworkers was remarkable for three reasons: It provided the first detailed analysis of the fundamental engineering parameters of a closed photobioreactor, introduced the concept of high S/V ratio reactors (127 m^{-1} in this case), and attained extremely high light-conversion efficiencies. It is worth noting that the about 40% reported efficiency for *Chlorella* growth corresponds to a quantum yield of about 6, as compared to the theoretical value of 8 (see earlier section). Pirt's analysis, however, did not consider the issue of the accumulation of photosynthetically generated oxygen, which represents a key limiting factor in large-scale tubular reactors (see later section), and ignored the light saturation effect, assuming that the efficiencies observed at low light intensity could be achieved in full sunlight as well. Although Pirt's serpentine reactor did deal efficiently with, and solve, many of the problems commonly encountered in closed systems, the concept of very high S/V ratio reactors revealed all its drawbacks when the experimental setup was scaled up to industrial level (see later section). The vertical serpentine reactor designed by Pirt was patented in 1983 (41). The patent also covers the design of a full-scale, near-horizontal manifold system according to which the culture flows from a tank through a descending conduit to the higher ends of a set of 50-m-long transparent tubes supported by a slightly inclined framework. Each tube communicates at its lowest point with its own riser, which leads up to the common reservoir. The tubes are arranged close together so as to intercept a high portion of the solar radiation impinging on the area occupied by the reactor. Both this manifold and the vertical serpentine design were scaled up to commercial size in the late 1980s by PBL (see later section).

In the mid-1980s at the Centro di Studio dei Microrganismi Autotrofi (Florence, Italy), Florenzano and coworkers started their experiments with horizontal serpentine tubular photobioreactors for the cultivation of *Arthrospira* (42). One of the first pilot reactors they set up was made of 14-cm-diameter, 0.3-mm-thick flexible polyethylene tubes. Because of problems of leakage, these collapsible sleeves were later replaced with rigid Plexiglas® tubes. The standard unit consisted of straight tubes, with a 13 cm ID and 4 mm thick, laid parallel on a white sheet and joined by PVC U-bends to form a loop 500 m in length (Fig. 6b). As it exited the reactor, the culture suspension flowed into a receiving tank, from which a diaphragm pump raised the suspension to a feeding tank placed at 3 m above ground level. The feeding tank was provided with a siphon that allowed discharge of about 340 L of culture into the photobioreactor every 4 minutes, thus sustaining a culture flow of 26 cm s^{-1}. The photobioreactor had a total volume of 8 m^3 and covered an area of 80 m^2. The productivity of *Arthrospira* in this system was 50% higher than in open ponds, due both to the better diurnal temperature profile and to the extended cultivation period. More recently, the same group developed and experimented with a 145-L two-plane serpentine reactor (43) (Fig. 6c). The system consisted of a 245-m-long loop made of 2.6-cm ID Plexiglas® tubes arranged in two planes to maximize capture of the radiation impinging on the horizontal, and thus increase areal yield. The two planes were connected to form a single loop and the culture, circulated by means of two airlifts, flowed from one plane to the other. A maximum areal productivity of 27.8 g m^{-2} d^{-1} was achieved in July. It would appear that for such a high-cost system, the loss of volumetric productivity due to shading by the upper plane would exceed the benefits deriving from improved areal yield and land saving. It must be noted, however, that the volumetric productivity achieved with this

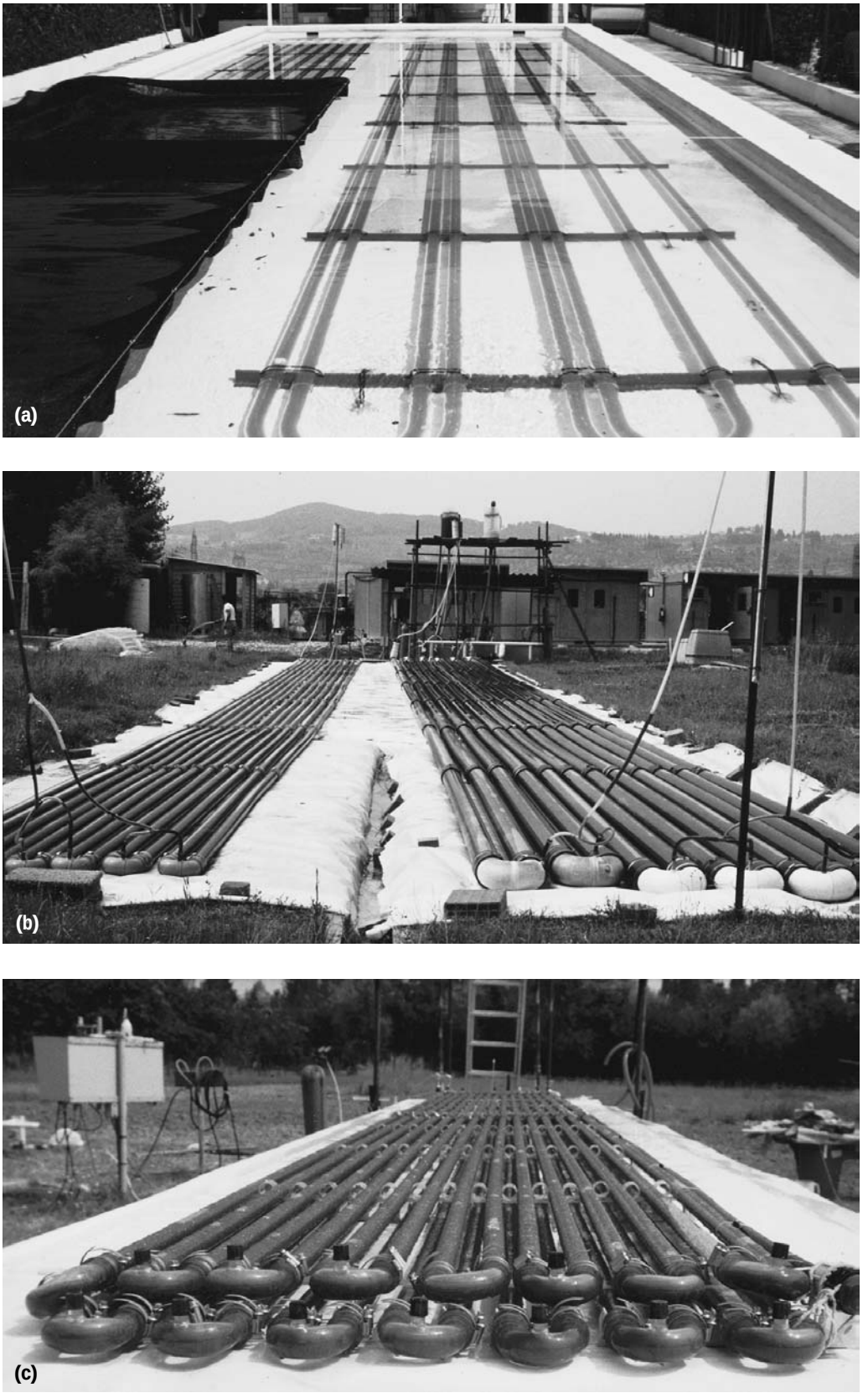


Figure 6. Horizontal serpentine photobioreactors set up at the University of Almeria, Spain (**a**) (photograph courtesy of E. Molina Grima) and at the Centro di Studio dei Microrganismi Autotrofi of Florence, Italy (**b, c**) (photographs courtesy of G. Torzillo).

two-plane photobioreactor (1.5 g L d^{-1}) was relatively high. The main merit of this study was that it addressed fundamental engineering parameters such as the effect of the rheological behavior of the algal culture on the mixing characteristics of the system, and the energy requirement for turbulent flow, which are often ignored.

Aquaserch (Kona, Hawaii) has recently devised and patented (44) a serpentine tubular reactor made of low-density polyethylene tubes. As in several previous serpentine designs, cooling is obtained by immersion of the reactor in a pond. The system, called the Aquaserch Growth Module (AGM), has been tested with different microalgae and is presently used to grow *Haematococcus* for astaxanthin production. The AGM is claimed to be the largest commercial-scale photobioreactor ever devised and one of the least expensive.

Manifold Photobioreactors. In manifold photobioreactors, a series of parallel tubes is connected at the ends by two manifolds, one for distribution and one for collection of the culture suspension. This type of photobioreactor represents one of the simplest configurations possible. As shown by Pirt et al. (20), about 15% of the energy consumed for recycling the culture in serpentine reactors is spent in moving the culture suspension around the bends; thus, in comparison to serpentine reactors, in which the culture inverts the direction of motion at each turn, manifold systems allow a significant saving of energy.

In 1991 in Javan (Tadshikistan), Sagdiana PGT built and used a large-scale glass reactor of the manifold type (Fig. 7a) (45) for cultivation of *Chlorella*. The system, very likely the largest glass reactor ever built, comprised 10 units consisting of two banks of 28 5.7-cm-ID tubes, each 73.5 m long, for a total of about 41,000 m of tubing. The tubes were connected in parallel to two metal headers each 20 cm in diameter. The culture suspension (13 m^3 in each unit) was circulated between the two banks of tubes at a speed of 1.1 m s^{-1} by a 290-m^3 h^{-1} capacity pump. Each unit was provided with one 3.6-m-high, 2.4-m-diameter gas exchanger. Temperature control was obtained through circulation of water from a single cooling tower serving all 10 units. The reactor was operated from August to October 1991, during which time it produced 0.5 ton of *Chlorella* biomass. Instability of the production process and contamination by *Scenedesmus* were reported to be the main problems.

Tredici and coworkers, working at the University of Florence (Italy), have developed several types of near-horizontal manifold-type photobioreactors (NHTR) (Fig. 7b). Units made of rigid or flexible tubes ranging from 6 to 85 m in length and from 4 to 5 cm in diameter have been built and used to grow *Arthrospira platensis*, *Anabaena siamensis*, and *Nannochloropsis* sp. under the climatic conditions of central Italy. Two eight-tube reactors of this type, 20 m in length, have been installed at the University of Hawaii as part of a joint project (Fig. 7c) (46). Typically, a NHTR consists of some 10–15 flexible tubes, each 45 m long and 4 cm in diameter, connected by two manifolds; the upper manifold, about 30 cm in diameter, acts as the degasser. The tubes are placed side by side on white corrugated plastic sheeting, facing south and inclined at a slight angle to the horizontal (4–6°). The corrugated sheeting keeps the tubes well aligned at a uniform inclination, reflects light onto the culture, and also provides drainage for rain and cooling water. Temperature control during the day is obtained by water spraying; the cooling water is collected and reused. The largest unit experimented with thus far occupies an area of 30 m^2 and contains about 600 L of culture. The main characteristic of the NHTR reactors developed in Florence is the use of air injection, at rates between 0.02 and 0.1 L L^{-1} min^{-1}, to attain mixing and gas exchange. The air is injected at the lower end of each tube by means of a perforated pipe inserted into the bottom manifold. In a different configuration used more recently, a number of tubes (typically one out of four) are not gassed, and are used as return flow tubes to increase the circulation speed of the culture and to obtain better mass transfer. Carbon dioxide addition to the air stream or in the return tubes is regulated through a pH-stat system. This reactor is shown in schematic form in Figure 8. Volumetric productivities of up to 1.4 g L^{-1} d^{-1} and areal productivities of more than 28 g m^{-2} d^{-1} have been obtained with *Arthrospira platensis* grown in these manifold systems. Besides simplicity of operation and low maintenance cost (due to the absence of moving parts), this internal gas-exchange reactor has other advantages: low shear stress compared to pump mixed reactors; no need for pumping the culture to a gas exchanger; and reduced fouling and wall growth due to the scouring effect of the gas bubbles. Moreover, this system appears to have potential for unit scale-up to about 100 m^2, and is relatively low cost in comparison to other designs. A preliminary estimate for a unit of 30 m^2, built and tested at Montepaldi (near Florence), sets the cost of this type of reactor at about \$60/m^2 (M.R. Tredici and G. Chini Zittelli, unpublished material, 1997), with the following cost per square meter: land, \$1.50; site clearing, grading, and leveling, \$6; liners, \$3.50; support framework, \$14.50; corrugated sheeting, \$5.50; tubing, \$13; manifolds, connectors, and piping, \$11.50; cooling system, \$1.50. A cost analysis for a large NHTR system for biological hydrogen production was recently presented (47).

A parallel-flow tubular photobioreactor was devised and experimented with in outdoor cultivation of cyanobacteria at the Microalgal Biotechnology Laboratory of Ben Gurion University (Israel) (48). The system consisted of eight 20-m-long, 3-cm-ID polycarbonate tubes laid parallel on the ground or on a horizontal platform and connected by manifolds made of 6.3-cm-diameter black polyethylene tubing (Fig. 7d). Circulation and gas exchange were obtained by means of a 2.2-m-high, 4.6-cm-ID airlift connected to a 20-L gas separator. A manifold received the culture from the separator and distributed it among a first series of culture tubes. At the far end of these tubes, the culture suspension was collected by a second manifold and distributed among a second series of returning tubes through which the culture ran back to the airlift riser and the gas separator. Small differences in flow rates were measured between the tubes. Temperature control was provided by water spraying. A productivity of 0.55 g L^{-1} d^{-1} was obtained with *Arthrospira platensis* and *Anabaena siamensis*. The main advantages of the system over serpentine reactors are the reduction of head losses and lower oxygen

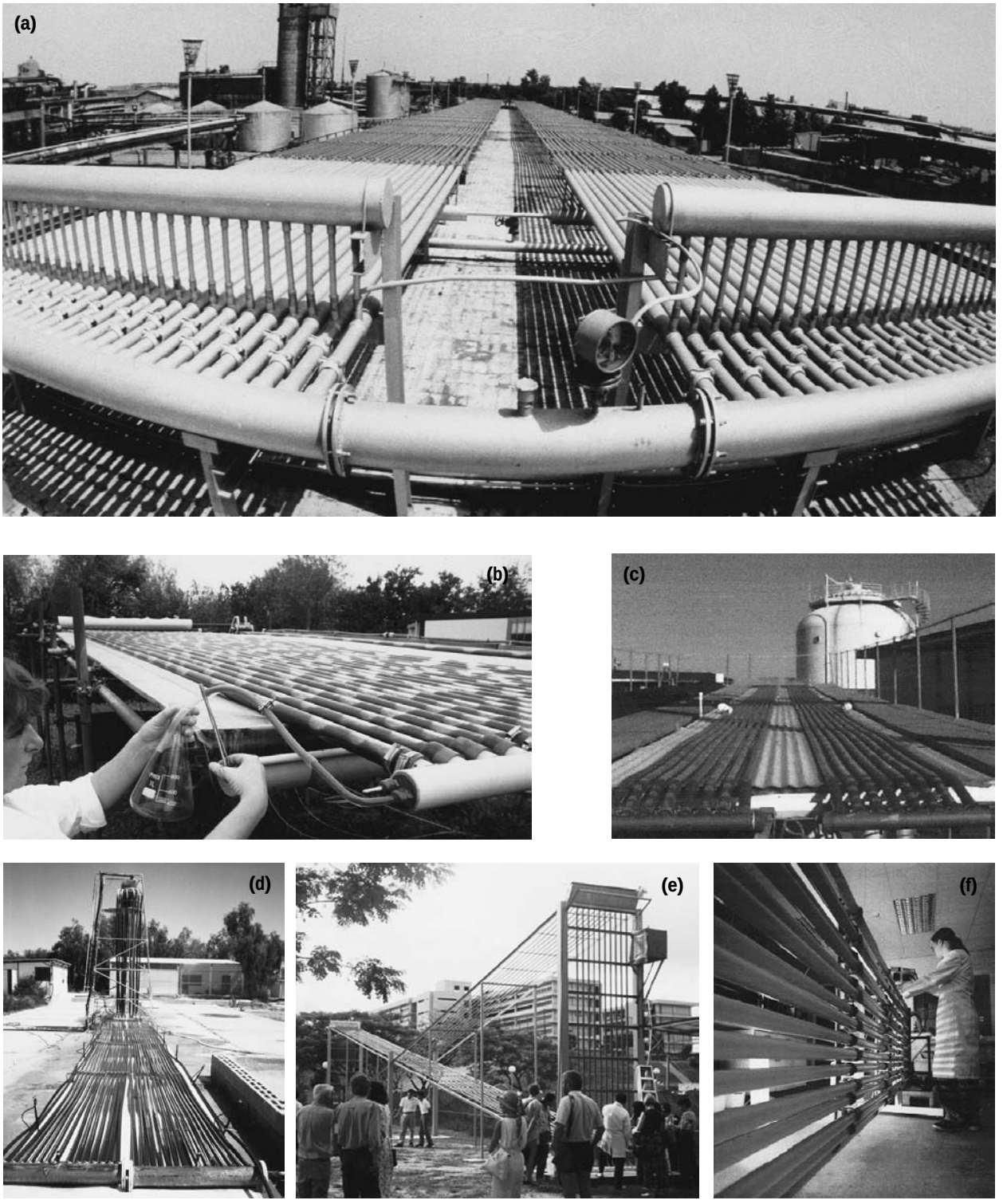


Figure 7. Manifold photobioreactors; (**a**) glass photobioreactor set up at Javan (Tadshikistan) by Sagdiana PGT (photograph by S. Markov); (**b**) NHTR photobioreactor developed at the Dipartimento di Scienze e Tecnologie Alimentari e Microbiologiche of the University of Florence, Italy (photograph by S. Pintus); (**c**) NHTR photobioreactor installed at the University of Hawaii, Honolulu, Hawaii, (photograph courtesy of O. Zaborsky); (**d**) parallel flow reactor designed at the Ben Gurion University of the Negev, Israel (photograph courtesy of A. Vonshak); (**e**) α-type photobioreactor set up at the Biotechnology Department of the National University of Singapore, Singapore (photograph courtesy of Y.K. Lee); (**f**) vertical manifold designed by Bio-Fence Ltd, Manchester, UK (photograph courtesy of Bio-Fence Ltd).

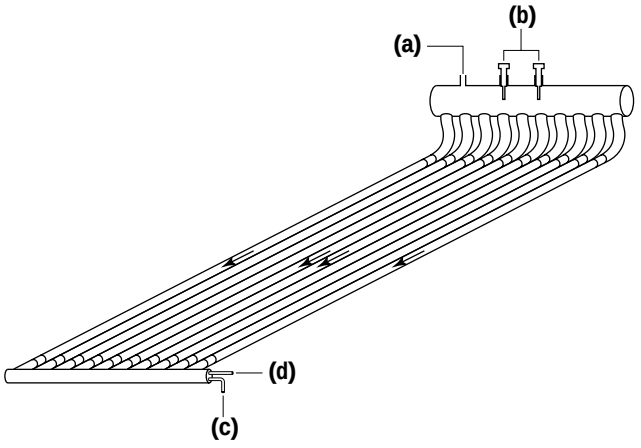


Figure 8. Scheme of the NHTR photobioreactor developed at University of Florence, Italy: (**a**), gas outlet; (**b**), electrodes and sensors; (**c**), culture outlet; (**d**), air and CO_2 injection point. Arrows indicate not-bubbled tubes used to circulate the culture suspension.

concentrations; these two factors facilitate scale up to industrial size.

A manifold elevated system called the α-type tubular photobioreactor was devised and experimented with in Singapore (Fig. 7e) (49). The reactor consisted of two sets of 10 parallel transparent PVC tubes (25 m long with a 2.5 cm ID), both placed at an angle of 25° to the horizontal but inclined in opposite directions. The algal suspension was lifted by air injection through a series of vertical PVC air-riser tubes (3.5 cm in diameter, 5 m in height) to a receiver tank from where it flowed down the first set of 25°-inclined tubes. These were connected to the base of a second series of air risers that lifted the culture to another receiver tank, from where the culture descended again through the second set of 25°-inclined tubes to the base of the first series of risers. The reactor thus looked like the Greek letter α. This elevated reactor, containing about 300 L of culture and covering a land area of about 12 m^2, was shown to be particularly well-suited to operation at low latitudes. *Chlorella pyrenoidosa* was grown for several months and reached very high cell densities (>10 g L^{-1}). The major problems were biofouling, solved by injecting air from the upper end of the inclined tubes, and foam formation at the higher cell densities. An areal productivity of 72 g m^{-2} occupied land d^{-1} was reported; as mentioned previously, the latter parameter can be misleading when referred to elevated systems.

At the University of Manchester (UK), Applied Photosynthetics Limited (APL) has recently developed a photobioreactor known as the Bio-Fence (Fig. 7f). The company offers several sizes of modular reactors of this type, designed specifically for the cultivation of the marine algal species used in aquaculture or for waste treatment with either microalgae or photosynthetic bacteria. The Bio-Fence consists of rigid transparent tubes racked together in banks and connected by manifolds in a fencelike structure. The culture suspension is circulated continuously between the photostage and a holding tank by a centrifugal pump or by an airlift. A patented scouring mechanism provides continuous cleaning of the inner walls of the tubes. The pH of the culture is controlled automatically by injection of CO_2 in the photostage. The most distinctive feature of the Bio-Fence is the stacked manifold configuration, which provides optimal gas exchange with elimination of pressure drops within the reactor. A 100-L unit requires a tube stack 5 m long by 0.75 m high. Units with a total volume of 500 L, with a tube length of 5 m, or 1,000 L volume, with a tube length of 10 m, can be assembled. The Bio-Fence is produced and marketed as units of from 50- to 1,000-L capacity that can be connected in parallel to any scale. Small steam-sterilizable systems for axenic operation are also available. In the case of systems connected in parallel, the units can be placed as close as 30 cm apart, as in the vertical system built by PBL in Spain and Pulz's alveolar plates (see later section). This arrangement will reduce volumetric productivity in case of outdoor utilization because of shading among the units but will attain two important results: high efficiency of land utilization (due to the very compact structure) and high photosynthetic efficiency (due to dilution of sunlight). The possibility of installing fluorescent tubes between the parallel units for additional lighting has also been considered. Typical cell concentrations achieved in the Bio-Fence are about 1 g L^{-1}. Productivity data are not available. The Bio-Fence is an accurately designed, well-engineered reactor. The only limitation seems to be the presence of a large dark tank that might reduce productivity in case of an excessively high ratio between darkstage and photostage volume.

Helical Photobioreactors. Helical photobioreactors consist of small-diameter, generally flexible tubing wound around an upright structure. This design is not new; it was used in the 1950s by Davis to grow *Chlorella* (50) and later adopted, in a flattened version made of glass tubes, by Setlik (28), Jüttner (29), and Krüger and Eloff (30). This coiled arrangement allows deployment of relatively long tubes on a small land area.

Helical photobioreactors called Biocoils were devised and patented in the late 1980s by Biotechna Ltd. (51). The Biotechna reactor consists of low-density polyethylene or clear food-grade PVC tubing, 2–3 cm in diameter, wound in a helix around an upright cylindrical core structure, typically 8 m in height with a core diameter of 2 m. Compressed air or peristaltic or centrifugal pumps may be used to induce culture circulation, with the culture being pumped downward through the coiled tube or upward through the tube to a header tank. Temperature control is achieved by a heat exchanger placed between the photostage and the pump. Pilot "biocoil" facilities have been installed at Luton (UK) and in Australia (52). Biotechna-Graesser A.P. Ltd. (based in the UK) marketed biocoils of different sizes (from a few liters capacity up to 10,000 L) for microalgae cultivation or for use with algae and heterotrophic bacteria in the treatment of wastewaters or toxic wastes. Small biocoil units are being operated for experimental and teaching purposes at King's College (London) (53) and at Murdoch University (Murdoch, Australia) (54).

A 120-L helical bubble reactor was devised and operated by Tredici and Chini Zittelli (19) to grow *Anabaena siamensis* and *Arthrospira platensis* outdoors (Fig. 9a). The reactor consisted of three 49-m-long transparent PVC tubes (3-cm ID, 0.5-cm wall thickness) wound around a vertical structure with an inclination of 2° to the horizontal in such a manner as to form a 1.65-m-high, 1.1-m-OD coil. The higher ends of the tubes were connected to a 20-L degasser. From the degasser, the culture flowed down through a single pipe connected to the lower ends of each tube. In this system, as in the NHTR devised by the same group, air was injected at the bottom of each tube and the bubbles flowing up the tubes provided mixing and gas exchange and prevented biofouling. Cooling was achieved by water spraying. With *A. platensis*, a mean volumetric productivity of 0.9 g L^{-1} d^{-1} and a photosynthetic efficiency of 6.6% (PAR) were achieved. The system operated with remarkable stability. The cost of a unit of this type was estimated at about $150.

A coiled, internal gas-exchange photobioreactor of about 3,000 L total volume was built in 1996 by Inalco S.p.A near Vicchio (Florence, Italy). The system consists of straight polycarbonate tubes, 5 cm in diameter, secured to an upright hexagonal frame with an inclination to the horizontal of 1.4°. The reactor is about 7 m high and occupies an area of about 40 m^2. Culture circulation is obtained by air injection at the base of the tubes. Temperature is controlled by a water-jacketed section. The reactor has been successfully experimented with in the cultivation of *Arthrospira platensis*.

Flat Photobioreactors. Flat photobioreactors and flat culture chambers have often been used to grow phototrophic microorganisms in the laboratory because they provide a simple geometry and greatly facilitate the measurement of irradiance at the culture surface. Despite their apparent simplicity, few such systems have been used for mass cultivation of algae.

The first flat culture unit devised for mass production of algae was the "rocking tray" used by Milner in 1953 (50) to grow dense *Chlorella* cultures in a thin turbulent layer. This system consisted of a 2.65-m^2, 9-cm-deep tray covered with transparent wire-reinforced plastic sheeting and supported on a shaker. At a mean culture depth of 1.7 cm, an average productivity of 8.2 g m^{-2} d^{-1} was obtained in a 20-day run. Following this experience, flat reactors were practically neglected until the 1980s.

Anderson and Eakin (55) cultivated *Porphyridium cruentum* outdoors for polysaccharide production using 3-m^2 flat plates 5–20 cm in depth (made of stainless steel and covered with glass or plastic) that were inclined to the horizontal to maximize interception of solar irradiance. About 20 g of polysaccharide m^{-2} d^{-1} were produced during the summer, when temperature control was provided.

In the mid-1980s, two groups working independently in France and in Italy introduced flat alveolar panels in algae cultivation. These systems are constructed from transparent PVC, polycarbonate, or polymethyl methacrylate sheets that are internally partitioned to form rectangular channels called alveoli. These sheets are commercially available in standard thicknesses from 4.5 to 40 mm.

Alveolar plates as culture systems for algae were first used by Ramos de Ortega and Roux (56) at the Centre d'Etudes Nucléaires de Grenoble (France). These authors experimented with 6-m-long, 0.25-m-wide, 4-cm-thick, double-layer panels made of transparent PVC for growing *Chlorella* in a greenhouse and outdoors. The plates were laid horizontally on the ground; the upper layer of channels was used for algal growth and the lower for thermoregulation. The culture suspension was circulated through a pump. A productivity of 24 g m^{-2} d^{-1} was achieved in the summer using units of 1.5-m^2 surface area. The authors stressed the fact that the double-layer plate permitted efficient thermoregulation of the culture and that significantly higher performances were obtained with the panels than with flexible sheets and semirigid tubes.

The first panel reactors built and experimented with in Florence by Tredici and coworkers were placed vertically, with the channels (and hence culture flow) running parallel to the ground, a design later adopted in Germany by Pulz. A pump was used to circulate the suspension. In 1988, the Consiglio Nazionale delle Ricerche (CNR) of Italy presented this reactor at the ITALIA 2000 exhibition, held in Moscow. The same year saw the awarding of a patent to a different design (57) in which the plate was placed vertically, but with the channels perpendicular to the ground; mixing was achieved by bubbling air at the bottom of the reactor. This latter version (Fig. 9b) has been used extensively in Florence for outdoor cultivation of cyanobacteria (5,19,58–61). The plates used in Florence are constructed from 16-mm-thick Plexiglas® alveolar sheets commonly available on the market at a cost of $30–40 m^{-2}. The main features that characterize these reactors are the high *S/V* ratio (80 m^{-1}), the vertical or tilted inclination from the horizontal of the channels, and the fact that no mechanical device is used to mix the cell suspension, circulation and degassing of the culture being attained through bubbling air at the base of each channel. Elevated "bubble column" plates of 2.2-m^2 surface area have been used to study the influence of some of the main cultural parameters, such as areal density, turbulence, oxygen tension, and orientation of the photobioreactor, on productivity and biomass composition of *Arthrospira platensis* (5). Alveolar plates, due to their high *S/V* ratio, attain high volumetric productivities (>2 g L^{-1} d^{-1}) and permit operation at high cell concentrations (4–6 g L^{-1}); these systems, however, have typically attained lower areal yields (about 24 g m^{-2} d^{-1}) compared to tubular reactors (about 28 g m^{-2} d^{-1}) (19). The lower performance achieved by cultures in plates has been attributed to the fact that these systems, unlike tubular reactors, do not achieve light dilution (unless they are placed at a high inclination with the horizontal) and thus cultures in plates suffer more strongly because of the light saturation effect and from photoinhibition (19). Alveolar plates, like other small flat systems, may be used profitably for research or in small production plants, however, due to the fact that the surface area of the cultivation unit is limited to a few square meters, they are not suitable for use in large commercial plants.

More recently, at the IGV Institut für Getreideverarbeitung (Bergholz-Rehbrücke, Germany), cultivation units based on 32-mm-thick alveolar panels were developed and

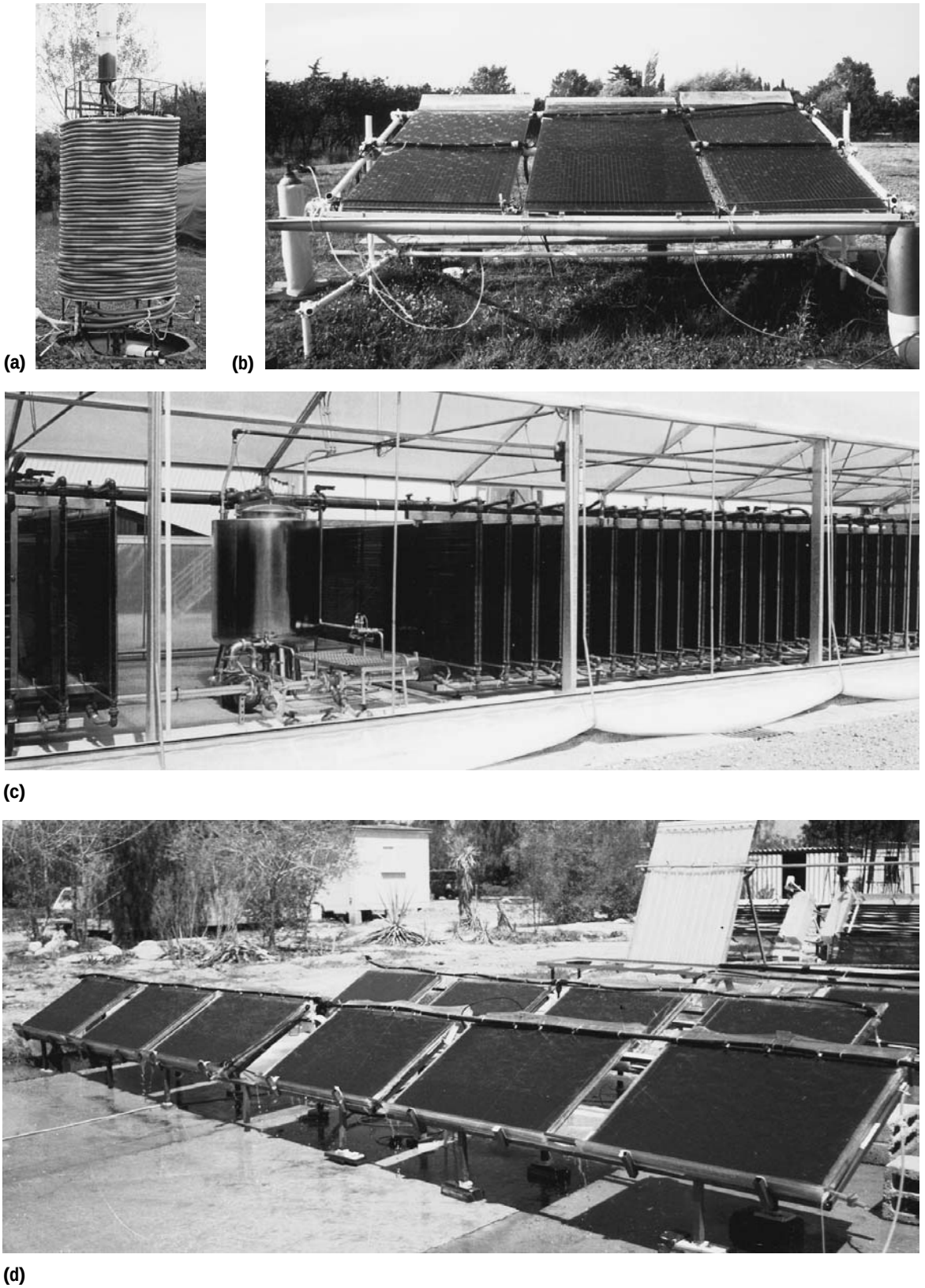


Figure 9. Biocoils and flat photobioreactors: (**a**) helical bubble-column reactor designed at the University of Florence, Italy (photograph by S. Pintus); (**b**) flat alveolar panels experimented with at the University of Florence, Italy (photograph by G. Chini Zittelli); (**c**) PBR system built by B. Braun Biotech Int., Germany (photograph courtesy of O. Pulz); (**d**) glass flat reactors experimented with at the Ben Gurion University of the Negev, Israel (photograph courtesy of A. Richmond).

operated by Pulz and coworkers (6). Typically, the plates are placed vertically with the channels arranged horizontally, just 20 cm apart, to form a compact structure called PBR (Fig. 9c). The culture, circulated by a piston pump, must reach a velocity adequate to maintain turbulent flow inside the channels. The distance to the degasser is critical in avoiding toxic concentrations of oxygen. Units from 100 to 10,000 L, made of closely packed plates, have been developed and tested both indoors and outdoors with different microalgae and cyanobacteria. Productivity levels of 1.3 g l^{-1} d^{-1}, corresponding to 28 g m^{-2} of illuminated surface d^{-1}, have been obtained with *Chlorella* under natural illumination. Among the problems that have emerged are buildup of oxygen (up to 30–40 mg L^{-1} at the outlet of the circuit) and fouling. The main characteristic of the PBR is the closely packing of the panels: 42 plates, for a total volume of 6 m^3, can be packed on 100 m^2 of land. Areal productivities of up to 130 g m^{-2} d^{-1} have been reported for such compact units, but these figures seem excessive and indeed confirm the incorrectness of referring the yield of vertical reactors to unit of occupied land. The main merit of the PBR, besides its high degree of technical development, is the achievement of significant light dilution. PBR reactors are sold commercially by B. Braun Biotech Int. in sizes from 10 to 6,000 L (6).

Richmond and coworkers (62) developed an inclined modular photobioreactor consisting of a series of flat glass chambers, 0.7 m high, 0.9 m long, and usually 2.6 cm in internal width, connected in cascade and tilted at the proper angle to maximize solar radiation capture (Fig. 9d). The culture volume varied from 6 to 50 L according to reactor width. Two mixing methods were tested: an airlift system consisting of five transparent PVC columns inserted vertically into the reactors with compressed air injected at the base of the columns, and air-bubble mixing by means of two perforated tubes running horizontally along the entire length of the reactor, one at the bottom and the other at midheight. The latter method yielded better performance. By adopting a reactor width of 2.6 cm and vigorous mixing (2.5 L air L^{-1} culture suspension min^{-1}), a record areal productivity of 49.4 g m^{-2} illuminated surface d^{-1} was achieved. When the back and side surfaces of the reactor were covered to avoid entry of diffuse light, productivity was reduced by about 20%. Nevertheless, this figure is the highest ever reported for a cyanobacterium cultivated outdoors. Using this flat modular system, the Israeli group has addressed and contributed to clarifying such fundamental aspects of photobioreactor design as optimal cell density with reference to the light path, the effect of reactor inclination, and the importance of the mode and intensity of mixing.

Commercial-Scale Photobioreactors

Only a few large commercial-scale photobioreactors have been built and operated: the PBL photobioreactors (based on the two different designs described in Pirt's patent) in southern Spain in the late 1980s and the Hidrobiologica SA manifold photobioreactor (probably inspired by previous similar designs devised in Italy by the Centro di Ricerca Agro-Industriale and Montedison) set up in the mid-1990s in northern Argentina. Despite the very different designs, both the plants suffered from similar mistakes with respect to some fundamental technical aspects and the scaling-up phase.

Photo Bioreactors Ltd. Plant. Although Pirt's measurements and extrapolations were not flawless, the quality of the basic work done at Queen Elizabeth College and the high projected productivities attracted investments and led to the creation of a company, Photo Bioreactors Ltd. (PBL UK), in 1986. A pilot plant was built and experimented with at the Reading University Research Farm (Sonning, UK). Three years later, PBL Spain (PBL SA) was founded with investments from private industries (among which the Danish Chr. Hansen Laboratory) and public Spanish sources (Instituto de Fomento de la Region de Murcia and Centro de Desarrollo Tecnico Industrial). A second pilot plant was set up in Santomera (Murcia, Spain), and finally a commercial plant for *Dunaliella* production was established in Santa Ana near Cartagena (Murcia, Spain). Full-scale activity started in May 1990, with a vertical system made of 50 rows, each of which contained 50 1.2-cm-diameter, 50-m-long polyethylene tubes (for a total tubing length of 125,000 m) (Fig. 10a). Circulation was provided by an airlift, and temperature was controlled by shading the reactors with nets or by water spraying. A near-horizontal system was built later according to the full-scale design patented by Pirt, using about 200,000 m of polyethylene tubing. Several major technical errors are apparent in these two systems: The tube diameter was too small for effective mixing considering the design (Fig. 10b), the tube material degraded quickly under sunlight, oxygen degassing was inadequate, wall growth was almost unavoidable, and control of temperature was problematic.

The very high S/V ratio in relation to the length of the tubes, together with the insufficient circulation and the improper management of the culture, led to poor growth of the alga, biofouling, and heavy contamination. But, as far as we know, the main reason for failure was inadequate piloting: although the pilot plants had been carefully planned and constructed with an appropriate 10-fold increase factor in the different stages, the problems encountered during the pilot stage were not taken into due consideration and had not been solved by the time commercial activity started. In September 1991, after an almost complete crash of the culture, the workers were dismissed and PBL closed without ever entering into production (63).

Hidrobiologica SA Plant. In 1996, Hidrobiologica SA, with an initial investment of about $2 million, built the largest known photobioreactor (Fig. 10c) about 15 km south of La Rioja (northern Argentina). The system consisted of 96 0.4-mm-thick polyethylene tubes, each 120 m long and 25.5 cm in diameter, laid parallel and horizontally on the ground and arranged like a manifold with feeding, connecting, and collecting channels made of concrete. The tubes, only partially filled, were maintained at very low positive pressure so that they assumed a roughly elliptical shape (about 35 cm wide and 9 cm high). The plant was divided into four modules, each of which comprised four six-tube submodules. An interspace of 0.5 m was allowed

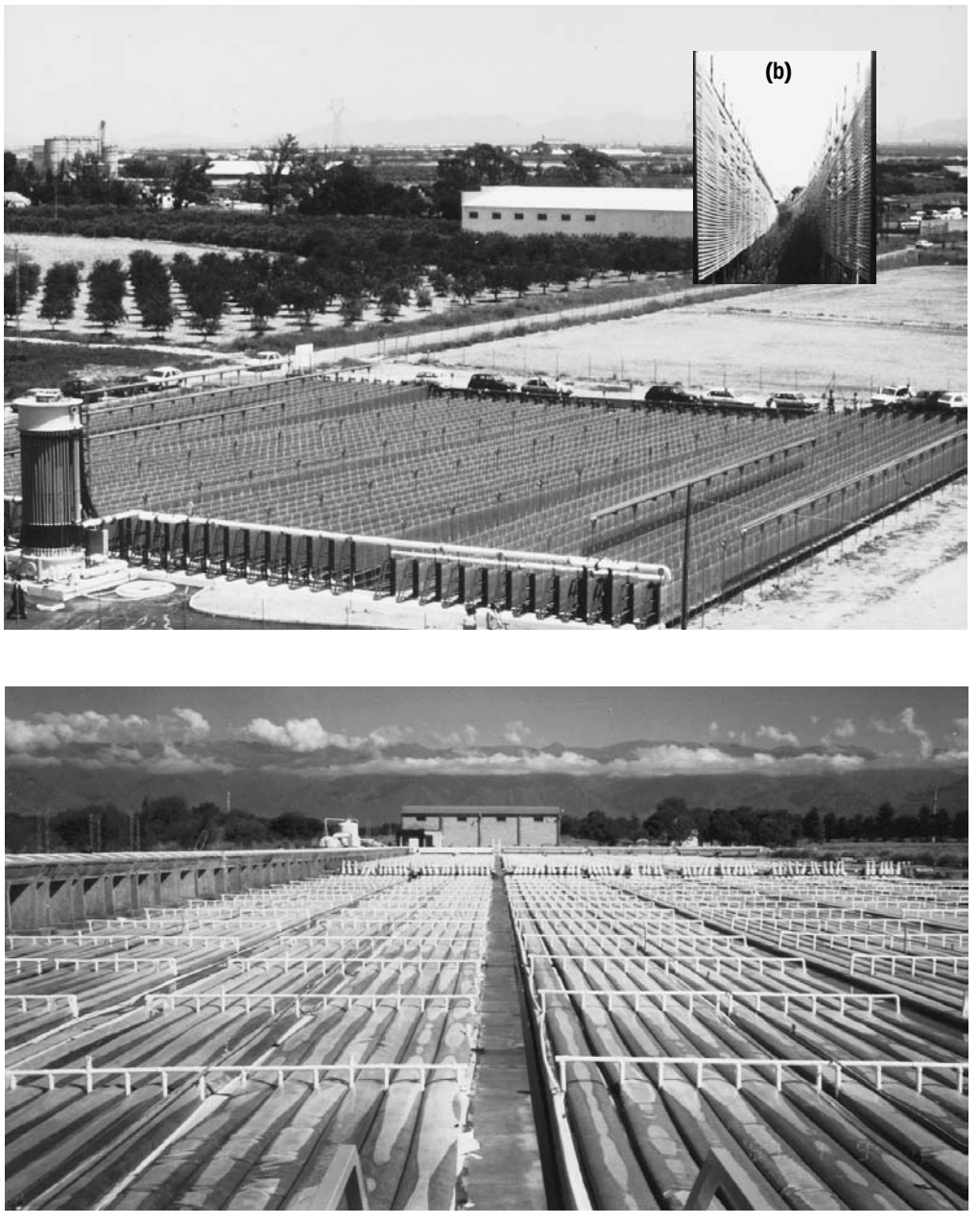


Figure 10. Commercial-scale photobioreactors: (**a**) vertical system set up by Photo Bioreactors Ltd (PBL) in Santa Ana, Murcia, Spain (photograph courtesy of D. Hall); (**b**) a particular of the PBL vertical system as it appeared in July 1995, four years after shutdown (photograph by the author); (**c**) the Hidrobiologica SA plant built at La Rioja, Argentina (photograph by G. Chini Zittelli).

between one submodule and the next, while about 1 m separated the modules. The four modules were connected to common feeding and collecting channels. The surface occupied by the whole plant, including the gaps between tubes, was about 5,000 m^2, and the total culture volume was 600 m^3. The culture (*Arthrospira platensis*) was circulated at a flow speed between 6 and 10 cm s^{-1} by a single axial flow pump with a maximum capacity of 900 $m^3\ h^{-1}$. The pump lifted the suspension from the collecting channel, into which it was fed by the culture tubes, up to a second raised, sloped channel. The suspension flowed by gravity through this second channel, which ran parallel to the culture tubes for 120 m, to the feeding channel at the opposite end of the reactor. The feeding channel, about 42 m long, distributed the culture suspension to the inlets of the single culture tubes. During the initial period of activity, the cell concentration was typically maintained at about 1 g L^{-1} and the productivity was calculated to be around 0.2 g $L^{-1}\ d^{-1}$; the biomass produced was of excellent quality. In the long run, however, several problems became evident; the most significant were poor mixing and biofouling deriving from the low flow speed and the un-

equal distribution of the culture among the tubes. The limited capacity to control temperature was also a major drawback in summer. It is likely that oxygen built up to dangerous levels despite the gas outlet provided halfway along the tubes, but no data are available. The plant is at present undergoing complete restructuring; the main objectives of this effort are to subdivide it into separate units, reduce the tube diameter, and improve mixing and culture circulation. The major errors, besides poor mixing and limited temperature control, were probably the absence of accurate piloting and the fact that only one large single unit was built and operated.

DESIGN CRITERIA FOR PHOTOBIOREACTORS

The basic function of a photobioreactor is to provide a controlled environment in which to achieve optimal growth and maximum, sustainable biomass and/or product synthesis with a specific phototroph (4). This means that a photobioreactor cannot be properly designed without adequate knowledge of the growth physiology of the organism to be cultivated. Besides, since phototrophic microorganisms are highly diverse as concerns their morphology, biology, nutritional and light requirements, and resistance to stresses, a single photobioreactor design cannot adapt to all organisms and all conditions.

Design criteria for photobioreactors should aim at achieving high efficiency of light conversion and at providing the necessary reliability and stability to the cultivation process by solving the main problems encountered in photobioreactor operation, such as overheating, oxygen buildup, biofouling, contamination, and photostage material deterioration. The fundamental design criteria for photobioreactors include reactor configuration, S/V ratio, mixing and degassing devices, temperature regulation, orientation, and the material used to build the photostage. These aspects are briefly discussed in the following sections. Reliability, ease of operation and scale-up, and construction and operating costs also take on particular relevance in relation to commercial application of photobioreactors.

Surface-to-Volume Ratio

The surface-to-volume ratio of the photobioreactor (i.e., the ratio between the illuminated surface of the reactor and its volume) determines the amount of light that enters the system per unit volume and the light regimen to which the cell population is subjected and is consequently one of the most important issues in photobioreactor design. The higher the S/V ratio, the higher the cell concentration at which the reactor is operated and the volumetric productivity of the culture. The cell concentration, in turn, greatly influences medium handling and harvesting costs and, to some extent, contamination level. The temperature profile and the hydrodynamic behavior of the culture suspension are also affected by this parameter (4,5,64). For example, in small-diameter tubular systems operated outdoors, the time required to reach the optimal temperature in the morning is reduced, and in the evening the relatively low culture volume cools down more quickly to ambient night temperatures, resulting in lower respiratory losses (4). For these reasons, in recent years we have witnessed a general trend toward a reduction of the diameter of tubular reactors and of the thickness of flat panels. Yet, it must be pointed out that although high S/V ratio reactors may work well in the laboratory, they may become extremely inefficient when scaled up. In high S/V ratio photobioreactors, all the volumetric activities that depend on the input of light energy per unit volume (such as O_2 evolution, CO_2 absorption, nutrient depletion, metabolite excretion, and heat production) change at a high rate, and this could have long-term negative effects on the stability of the process. These effects are rarely observed because experiments do not last long enough for them to become evident (63). The industrial systems set up by PBL (see earlier section) provide clear examples of the kinds of problems that may be absent at laboratory scale and yet arise when high S/V ratio photobioreactors are adopted in large-scale plants. It is not possible to fix an optimal S/V ratio for large-scale photobioreactors, since this parameter depends on the organism to be grown, on the type of mixing, and on other specific conditions. However, my laboratory experience suggests that a light path between 3 and 6 cm may represent a reasonable compromise.

Oxygen Accumulation

Accumulation of photosynthetically generated oxygen is one of the main factors that limit scale-up of photobioreactors. Dissolved oxygen concentrations over 35 mg L^{-1} (equivalent to about four times air saturation at 25 °C) may be easily reached in outdoor dense algal cultures at midday. Such high oxygen levels are toxic to many phototrophs and, if coupled with prolonged exposure to full sunlight, may lead to the photooxidative death of the culture (4,9). Since volumetric oxygen production is directly related to volumetric biomass productivity (about 1 g of oxygen is evolved per gram of biomass synthesized), oxygen buildup becomes an especially serious problem in high S/V ratio photobioreactors operated outdoors. At maximal rates of photosynthesis, a 1-cm-diameter tubular reactor would accumulate about 8–10 mg of oxygen L^{-1} min^{-1} (64). In such systems, the maintenance of dissolved oxygen below the toxic concentration requires adoption of very short loops and this fact makes the serpentine design difficult to scale up. The situation is less dramatic in tubes of greater diameter, since the volumetric rate of oxygen evolution falls in proportion to the decrease of the S/V ratio, but it has been shown that even in 14-cm-diameter tubular reactors dissolved oxygen may reach levels that significantly reduce productivity (42). In serpentine reactors, the time cycle (time required for the culture to reach the degassing point), the speed of circulation, and the length of the loop must be very carefully chosen in relation to the expected volumetric productivity to prevent buildup of toxic oxygen concentrations. For example, in a tubular reactor with a 2.6-cm ID and a culture flow speed of 0.3 m s^{-1}, the time cycle should not exceed 6 min, hence, the length of the reactor must be less than 100 m (4). To increase the length of the reactor it is necessary to increase the culture circulation speed or reduce the S/V ratio. Internal gas-exchange systems (bub-

ble columns) offer some advantage in this respect. Unlike serpentine reactors, in which oxygen accumulates during the loop cycle and is removed only at the degasser, bubble columns present the advantage that the whole culture is subjected simultaneously to the degassing action of air.

Mathematical models can be employed to predict the effect on oxygen concentration of various design variables, such as the length of the loop, the tube diameter, and the liquid circulation speed. A useful model for prediction of dissolved oxygen and carbon dioxide axial concentration profiles in serpentine tubular reactors has been developed and verified experimentally with *Porphyridium cruentum* in outdoor continuous culture by Camacho Rubio et al. (40).

Mixing

The type of device used to mix and circulate the culture suspension is essential in the design of a successful photobioreactor. Both the productivity of a photobioreactor and the cost of its construction and operation are influenced to a great extent by the type of device used for mixing (65). The beneficial effects that mixing has on growth and productivity of mass cultures of phototrophs are well known (9). Mixing is necessary to prevent cells from settling, to avoid thermal stratification, to distribute nutrients and break down diffusion gradients at the cell surface, to remove photosynthetically generated oxygen, and, above all, to ensure that all the cells experience alternating periods of light and darkness. It has been postulated that cells quickly brought into and out of the lighted portion of the culture could take advantage of the "flashing light effect" to increase photosynthetic efficiency. Conflicting results have been reported, however, for the light–dark fluctuations of medium frequency (i.e., in the order of seconds to minutes) that prevail in mixed-mass cultures as regards their stimulatory effect on productivity, and the true role of the mixing rate in determining algal productivity still remains to be clarified (9,64–66). With a few exceptions, significant positive effects of intense stirring on productivity of algal mass cultures have been observed only in paddle-wheel-mixed raceway ponds (9,67). Carlozzi and Torzillo (68) observed a 30–40% increase of the productivity of *A. platensis* grown in tubular reactors when the flow speed was increased from 0.18 to 0.75 m s^{-1}. Tredici and Chini Zittelli (5,60) studied the influence of the bubbling rate on productivity of *A. platensis* in inclined flat plates. At an air supply rate of 1.21 L L^{-1} min^{-1}, productivity increased by 35% over that at a rate of 0.36 L L^{-1} min^{-1}. The authors suggested that the higher bubbling rate enhanced productivity by providing a more adequate light regimen to the cells, as well as by reducing the oxygen level in the culture. Hu and Richmond (62) have recently attained astonishingly high productivities with *A. platensis* cultivated in flat plates by adopting ultrahigh cell densities (higher than 15 g L^{-1}) and providing mixing at the maximal rate permissible without causing cell damage. The choice of the intensity of mixing and of the mixing system must be dictated by the characteristics of the organism to be cultivated. Unlike *Chlorella* or the marine *Nannochloropsis*, which are endowed with strong cell walls, cultures of filamentous cyanobacteria and fragile dinoflagellates cannot be mixed by pumps that subject the cells to high shear stress and so lead to fragmentation of thricomes and to cell lysis. There is general agreement that bubble columns and airlift systems are low-shear-stress mixing systems and therefore are recommended for shear-sensitive organisms. Gudin and Chaumont (69) have reported that productivity may increase up to 75% when pumps are replaced by airlift systems. However, it is known that even air bubbling produces some stress on the cells at the level of bubble formation and breakup, bubble coalescence, and bubble-bursting at the culture surface (4). Other noticeable advantages of bubble columns and airlifts, from the point of view of industrial application, are their simplicity of operation and low maintenance costs due to the absence of moving parts. Mixing devices and power requirements for mixing have been dealt with by Weissman et al. (64) and by Torzillo (4). Torzillo has also analyzed the influence of the rheological properties of the culture on mixing and mass transfer (4). Gas–liquid mass transfer in a serpentine airlift photobioreactor has been described in detail by Camacho Rubio et al. (40).

Temperature Control

Although photobioreactors are said to provide a better process control than do ponds, the maintenance of optimal temperatures in closed reactors is not an easy task, especially when the systems are operated outdoors. Photobioreactors behave as solar collectors: Any light absorbed by the reactor walls or by the culture and not used in photosynthesis is converted into thermal energy, and may lead to an increase of the culture temperature of up to 30 °C or more above ambient temperature, causing the collapse of even rather tolerant species (59). With a few exceptions (e.g., cultivation of thermophilic species) photobioreactors operated outdoors require cooling. Among the solutions experimented with to control overheating in outdoor photobioreactors are shading, water-bath immersion, and water spraying. Shading, to be effective, requires that a large proportion of the reactor be covered for several hours a day, and this greatly reduces the irradiance at the culture surface and, consequently, the productivity (4). Cooling by water spraying has been shown to be reliable and cost-effective when used in dry climates (4,5). Immersion of the reactor in a cooling water basin is very efficient, but its cost-effectiveness is rather doubtful. It may, however, be the sole available option in humid climates, where water spraying does not provide evaporative cooling. Although it varies with the reactor design, the local climate, and the specific organism used, the need for cooling can significantly increase operating costs. In cold climates and in temperate regions during winter, heating may instead be required to maintain the optimal growth temperature.

Orientation of the Photobioreactor and Light Dilution

One of the main impediments to the efficient use of solar energy is the continuous variation of the available solar radiation throughout the day. Solar irradiance is very low in the early morning and in the evening hours and strong enough to inhibit photosynthesis at noon. Although the di-

urnal pattern of the solar irradiance incident on a horizontal reactor cannot be easily modified, some elevated systems can be oriented at various angles to the sun and thus offer the possibility to vary, continuously or discontinuously, the distribution of the light that strikes the reactor surface. The effect of the photobioreactor orientation on productivity of outdoor algal cultures was investigated by Lee and Low (70), who did not find the angle of inclination of the bioreactor to have any significant influence on the productivity of *Chlorella pyrenoidosa*. Because of the particular reactor design (two tubular panels were connected by hinges, erected at an angle to the ground, and placed in an east–west direction) and, above all, due to the location of the experimental facilities (about 1° latitude north), the findings of these researchers cannot be generalized. The influence of the orientation of flat photobioreactors in higher latitude regions has been studied by Tredici and Chini Zittelli (5) and by Hu et al. (11). These studies have shown that the reactor tilt angle exerts a significant effect on both productivity and photosynthetic efficiency. Systems that are sun oriented so as to intercept the maximum solar radiation on a daily basis (e.g., south-facing plates inclined at 10–30° in summer and at about 60° in winter) typically achieve the highest productivities, even if a culture placed perpendicularly to the rays of the sun at midday may suffer due to excessively high irradiance (19). Vertical plates or systems oriented so as to dilute the impinging radiation achieve lower productivities but higher photosynthetic efficiencies (19). A major objective of photobioreactor design should be to reduce the so-called light saturation effect of photosynthesis (see earlier section). One approach has been to install systems vertically (like Pulz's PBR series) or at a high inclination to the horizontal so that the reactor is at a large angle to the sun rays and solar irradiance is attenuated (diluted) at its surface (19). Another approach involves transmitting sunlight into the photobioreactor through light-diffusing prisms or optical fibers (6,71), but this option appears too complex to be effectively scaled up and applied commercially in the near future. By operating a near-horizontal flat panel and a near-horizontal tubular reactor in parallel, Tredici and Chini Zittelli (19) have shown that tubular reactors can achieve, per se, higher photosynthetic efficiencies and productivities than flat reactors because their curved surfaces achieve a certain degree of sunlight dilution and thus reduce the light saturation effect.

Materials

Materials for photobioreactors must have several properties: high transparency, high mechanical strength, lack of toxicity, resistance to weathering (durability), chemical stability, antifouling surface, and low cost. G. Chini Zittelli and M.R. Tredici (unpublished material, 1998) have carried out a preliminary investigation of the durability of some plastic materials. In Figure 11, the loss of transparency of six different tubing materials exposed to natural weathering in central Italy is shown. Polyethylene and polypropylene tubes are inexpensive, but both lose transparency very quickly and present serious limitations as regards biofouling and mechanical strength. Polycarbonate tubes are seriously weakened by alkaline solutions and become brittle and break easily after 2–3 years of exposure outdoors; they lose transparency mainly in the 300- to 500-nm range. Tubes of polymethyl methacrylate or Teflon have been shown to possess high resistance to weathering and good overall properties. PVC tubes lose transparency significantly after hydration. Some of the flexible plastic materials tested have shown limited mechanical strength that has led to trivial problems (such as leakage due to punctures inflicted by cats and wild animals) that could become rather serious in large-scale systems. Glass might be an excellent material for photobioreactors, due to its high transparency, chemical stability, and durability. Glass tubular systems, however, would require mostly field assembly and numerous connections, and this would greatly increase installation costs; furthermore, fragility may constitute a serious limitation to the use of glass tubes in large-scale installations. The tubes required for building the photostage represent a major cost component of the photobioreactor but not as high as commonly thought. Suitable long-lasting tubular material is available for about $1 per meter or about $15 per square meter of photobioreactor.

CONCLUDING REMARKS AND PROSPECTS

Photobioreactors are considered to have several major advantages over open ponds: They can *(1)* prevent or minimize contamination, *(2)* offer better control over cultural conditions (pH, pCO_2, pO_2, temperature, etc.), *(3)* prevent evaporation and reduce water use, *(4)* lower CO_2 losses due to outgassing, *(5)* permit higher cell concentrations that result in reduced operating costs, and *(6)* attain higher productivity. However, the ultimate and most important advantage is that photobioreactors permit cultivating algal species that cannot be grown in open ponds. On the other hand, certain requirements of photobioreactors, including the need for cooling, the need for strict control of oxygen accumulation and biofouling, and the need for frequent replacement of the material of the photostage, make these systems more expensive to build and operate than ponds. Thus, despite their advantages, the use of photobioreactors must be limited to production of very-high value compounds from phototrophs that cannot be cultivated in open ponds (18,52). Photobioreactors and open ponds should not be viewed as competing technologies, since it is very unlikely that photobioreactors will have a significant impact on any product or process that can be attained in large outdoor open ponds.

In recent years, significantly higher photosynthetic efficiencies and a higher degree of system reliability have been achieved due to better understanding of the growth requirements of phototrophic microorganisms in mass culture. Despite these advances, no photobioreactor has yet proven to be commercially successful, even for very-high-value products. This certainly does not depend on a lack of photobioreactor designs: There exist several valid basic concepts and myriad variations. The principle obstacle to

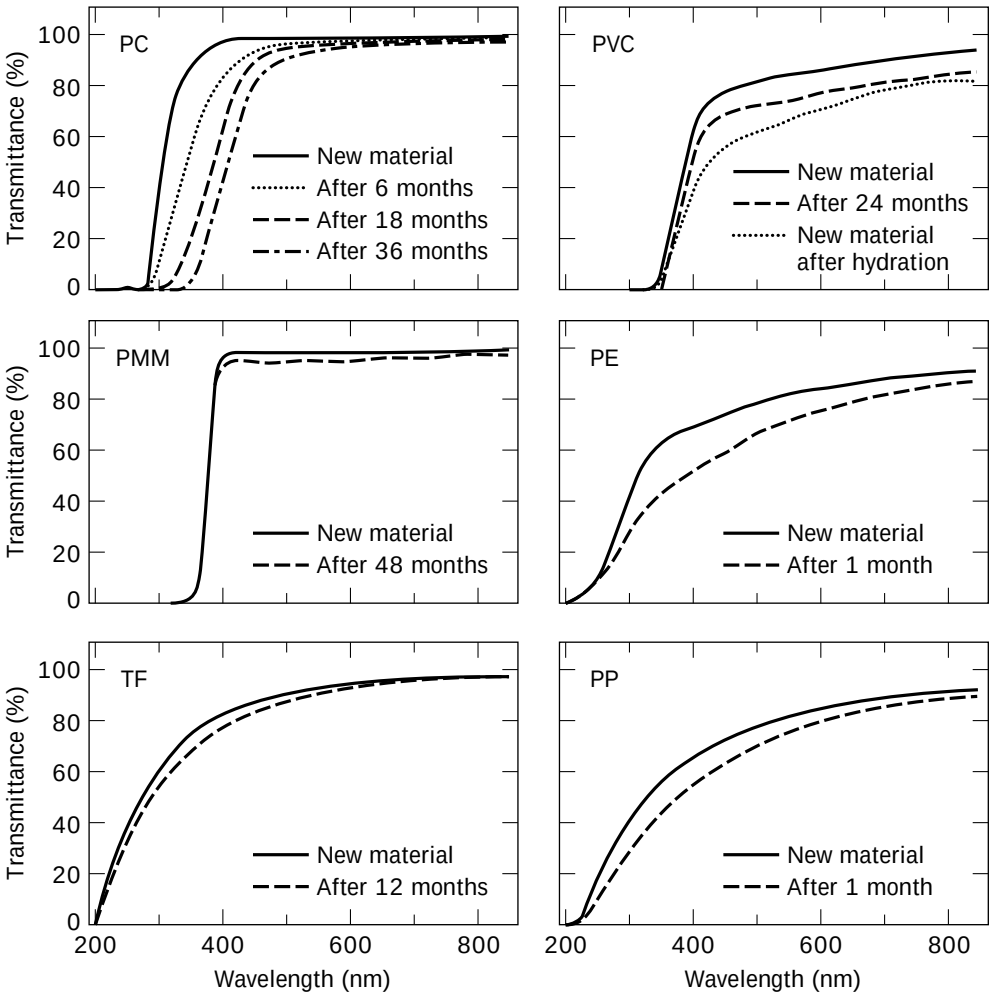


Figure 11. Loss of transmittance of different plastic tubing exposed outdoors to the climatic conditions of central Italy. PC, polycarbonate; PVC, polyvinyl chloride; PMM, polymethyl methacrylate; PE, polyethylene; TF, Teflon; PP, polypropylene. The tubing thickness was: TF, 0.15 mm; PE, 0.2 mm; PP, 0.4 mm; PC and PVC, 1 mm; PMM, 2 mm.

commercial application of this biotechnology seems instead to lie at the level of scaling-up. Transferring those processes developed at the laboratory scale to the industrial scale is never a simple proposition, and it is particularly difficult in the case of photobioreactors because of the peculiar problems of these systems. One of the major causes of the failures of photobioreactor technology at commercial scale is a lack of adequate piloting. The old rule of thumb used in the chemical industry, and still valid in fermentation technology, that each individual increase of scale should not surpass a factor of 10 (52,72) has been ignored in practice. This was especially true in the case of Hidrobiologica SA, where the full-scale plant was set up after only limited experimentation on a very small scale. At the small scale, the photobioreactor performance is usually evaluated in terms of productivity, whereas issues such as contamination, material degradation, and the reliability and sustainability of the production process, which more than the productivity determine the failure or success of the industrial-scale activity, are ignored. Only accurate piloting can shed light on the key problems and provide adequate solutions. It is not clear how much the failure of the PBL photobioreactors in Spain and of the 600-m^3 plant in Argentina depended on system design flaws or on management errors. However, in neither case had the basic technical and biological problems (mixing, oxygen buildup, temperature control, biofouling, contamination) that must be solved for successful full-scale operation been subjected to rigorous investigation at an adequately scaled pilot level. Another mistake, common to both plants, is that to keep operating costs low, the units built and operated were very large. This deprived the systems of necessary flexibility and negated the possibility of making improvements by trial and error. Under these conditions, any problem that might arise (e.g., contamination or toxic oxygen tensions) affects the whole culture and jeopardizes all of production. It is worth pointing out that both industrial plants failed in the cultivation of two phototrophs (*Dunaliella* and *Arthrospira*) that are commonly grown in the much less rigorously controlled open ponds.

The establishment of a commercial activity based on the cultivation of phototrophic microorganisms requires realizing that a photobioreactor design that is efficient at the laboratory level is not necessarily so at the level of industrial production—hence, the importance of accurate piloting. A few simple guidelines should be followed during piloting (63):

1. Experiments must last long enough to reveal the "aging" problems, including contamination, biofouling, and material deterioration, that become evident only after several months of continuous operation.
2. The units experimented with should be similar in size and type to the cultivation units envisaged for the full-scale plant.
3. The productivity of the system must be correctly evaluated and extrapolated.
4. Experiments must be carried out under the climatic and operating conditions that will prevail in the final commercial process.
5. Unless true axenic conditions are adopted, the possibility of contamination must not be underestimated.
6. Appropriate solutions to the main problems commonly encountered in photobioreactors (overheating, biofouling, oxygen accumulation, etc.) must be provided before beginning operation of the commercial plant.
7. The data obtained during the scaling-up stage must be subjected to rigorous analysis and used to reevaluate the economics of the process.

Today the efforts of many researchers, from both academia and industry, are directed toward development of new photobioreactor designs. In the majority of cases, such designs are anything but new and remain an academic exercise with little or no practical value. What we really need is a serious evaluation of the performance of photobioreactors at a relatively large scale. Through such analysis, and together with likely further improvements in engineering design, the development of new materials with suitable characteristics, and the improvement of cultivation techniques derived from a better knowledge of the physiology of phototrophs in mass culture, we can reasonably expect to achieve a degree of reliability in large-scale cultivation of phototrophic microorganisms such as to bridge the gap between laboratory experimentation and commercial application.

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to my colleague Dr. Graziella Chini Zittelli, with whom I have been working on microalgae mass culture and photobioreactor development since 1989. I also wish to thank Dr. Susan Blackburn and Dr. Laura Barsanti for critically discussing the manuscript. Parts of this article have been influenced by exchange of ideas over the years with Dr. John Benemann, to whom I feel particularly indebted.

BIBLIOGRAPHY

1. A. Prokop and L.E. Erickson, in J.A. Asenjo and J.C. Merchuk eds., *Bioreactor System Design*, Marcel Dekker, New York, 1995, pp. 441–477.
2. Y.K. Lee, *Trends Biotechnol.* **4**, 186–189 (1986).
3. D. Chaumont, *J. appl. Phycol.* **5**, 593–604 (1993).
4. G. Torzillo, in A. Vonshak ed., *Spirulina platensis (Arthrospira): Physiology, Cell Biology, and Biotechnology*, Taylor & Francis, London, 1997, pp. 101–115.
5. M.R. Tredici and G. Chini Zittelli, in A. Vonshak ed., *Spirulina platensis (Arthrospira): Physiology, Cell Biology and Biotechnology*, Taylor & Francis, London, 1997, pp. 117–130.
6. O. Pulz and K. Scheibenbogen, in T. Scheper ed., *Bioprocess and Algae Reactor Technology, Apoptosis*, Springer-Verlag, Berlin, 1998, pp. 123–152.
7. D.O. Hall and K.K. Rao, *Photosynthesis 5th ed.*, Cambridge University Press, Cambridge, UK, 1995, pp. 1–194.
8. G.E. Fogg, *Algal Cultures and Phytoplankton Ecology*, The University of Wisconsin Press, Madison, Wisconsin, 1965, p. 12.
9. A. Richmond, in F.E. Round and D.J. Chapman eds., *Progress in Phycological Research*, vol. 7, Biopress, Bristol, UK, 1990, pp. 269–330.
10. Q. Hu, H. Guterman, and A. Richmond, *Biotechnol. Bioeng.* **51**, 51–60 (1996).
11. Q. Hu, D. Faiman, and A. Richmond, *J. Ferment. Bioeng.* **85**, 230–236 (1998).
12. B. Kok, in J.S. Burlew ed., *Algal Culture from Laboratory to Pilot Plant*, Carnegie Institution of Washington Publication No. 600, Carnegie Institution, Washington, D.C., 1953, pp. 63–75.
13. J.F. Cornet, C.G. Dussap, and J.B. Gros, in T. Scheper ed., *Bioprocess and Algae Reactor Technology, Apoptosis*, Springer-Verlag, Berlin, 1998, pp. 153–224.
14. F.G. Acien Fernandez, F. Garcia Camacho, J.A. Sanchez Perez, J.M. Fernandez Sevilla, and E. Molina Grima, *Biotechnol. Bioeng.* **55**, 701–714 (1997).
15. F.G. Acien Fernandez, F. Garcia Camacho, J.A. Sanchez Perez, J.M. Fernandez Sevilla, and E. Molina Grima, *Biotechnol. Bioeng.* **58**, 605–616 (1998).
16. J.S. Burlew, in J.S. Burlew ed., *Algal Culture from Laboratory to Pilot Plant*, Carnegie Institution of Washington Publication No. 600, Carnegie Institution, Washington, D.C., 1953, pp. 3–23.
17. J.C. Goldman, *Water Res.* **11**, 119–136 (1979).
18. J.R. Benemann, in R.C. Cresswell, T.A.V. Rees, and N. Shah eds., *Algal and Cyanobacterial Biotechnology*, Longman, Harlow, UK, 1989, pp. 317–337.
19. M.R. Tredici and G. Chini Zittelli, *Biotechnol. Bioeng.* **57**, 187–197 (1998).
20. S.J. Pirt, Y.K. Lee, M.R. Walach, M.W. Pirt, H.H.M. Balyuzi, and M.J. Bazin, *J. Chem. Tech. Biotechnol.* **33B**, 35–58 (1983).
21. A. Sukenik, R.S. Levy, Y. Levy, P.G. Falkowsky, and Z. Dubinsky, *J. appl. Phycol.* **3**, 191–201 (1991).
22. H. Guterman, A. Vonshak, and S. Ben-Yaakov, *Biotechnol. Bioeng.* **35**, 809–819 (1990).
23. M. Brouers, H. deJong, D.J. Shi, and D.O. Hall, in R.C. Cresswell, T.A.V. Rees, and N. Shah eds., *Algal and Cyanobacterial Biotechnology*, Longman, Harlow, UK, 1989, pp. 272–293.

24. P.M. Cook, in J. Brunel, G.W. Prescott, and L.H. Tiffany eds., *The Culturing of Algae*, Charles F. Kettering Foundation, Yellow Springs, Ohio, 1950, pp. 53–75.
25. Anonymous, in J.S. Burlew ed., *Algal Culture from Laboratory to Pilot Plant*, Carnegie Institution of Washington Publication No. 600, Carnegie Institution, Washington, D.C., 1953, pp. 235–272.
26. A.W. Fisher, in L.W. Douglas and M.L. Kastens eds., *Proc. World Symp. Appl. Solar Energy*, Stanford Research Institute, Menlo Park, Calif., 1956, pp. 243–253.
27. H. Tamiya, E. Hase, K. Shibata, A. Mituya, T. Iwamura, T. Nihei, and T. Sasa, in J.S. Burlew ed., *Algal Culture from Laboratory to Pilot Plant*, Carnegie Institution of Washington Publication No. 600, Carnegie Institution, Washington, D.C., 1953, pp. 204–232.
28. I. Setlik, J. Komarek, and B. Prokes, in J. Necas and O. Lhotsky eds., *Annual Report of the Laboratory of Experimental Algology and Department of Applied Algology for the Year 1966*, Knihtisk, Praha, Czechoslovakia, 1967, pp. 5–36.
29. F. Jüttner, *Proc. Biochem.* **17**, 2–7 (1982).
30. G.H.J. Krüger and J.N. Eloff, in J.U. Grobbelaar, C.J. Soeder, and D.F. Toerien eds., *Wastewater for Aquaculture*. University of Orange Free State Publication, Series C (no. 3), Bloemfontein, South Africa, 1981, pp. 16–23.
31. C.J. Soeder, A. Bolze, and H.D. Payer, *Br. Phycol. J.* **16**, 1–7 (1981).
32. C.M. James and A.M. Al-Khars, *Aquaculture* **87**, 381–393 (1990).
33. M.M. Helm, I. Laing, and E. Jones, *MAFF Direct. Fish. Res.* **53**, 1–7 (1979).
34. R. Radmer, P. Behrens, and K. Arnett, *Biotechnol. Bioeng.* **29**, 488–492 (1987).
35. K. Miyamoto, O. Wable, and J.R. Benemann, *Biotechnology Lett.* **10**, 703–708 (1988).
36. Q. Hu and A. Richmond, *J. appl. Phycol.* **6**, 391–396 (1994).
37. E. Cohen and S. (Malis) Arad, *Biomass* **18**, 59–67 (1989).
38. D. Chaumont, C. Thepenier, and C. Gudin, in T. Stadler, J. Morillon, M.C. Verdus, W. Karamanos, H. Morvan, D. Christiaen eds., *Algal Biotechnology*, Elsevier Applied Science, London, 1988, pp. 199–208.
39. E. Molina Grima, J.A. Sanchez Perez, F. Garcia Camacho, J.M. Fernandez Sevilla, F.G. Acien Fernandez, and J. Urda Cardona, *Appl. Microbiol. Biotechnol.* **42**, 658–663 (1995).
40. F. Camacho Rubio, F.G. Acien Fernandez, J.A. Sanchez Perez, F. Garcia Camacho, and E. Molina Grima, *Biotechnol. Bioeng.* **62**, 71–86 (1999).
41. U.K. Pat. 2,118,572 (March 17, 1983), S.J. Pirt (to Queen Elizabeth College, University of London).
42. G. Torzillo, B. Pushparaj, F. Bocci, W. Balloni, R. Materassi, and G. Florenzano, *Biomass* **11**, 61–74 (1986).
43. G. Torzillo, P. Carlozzi, B. Pushparaj, E. Montaini, and R. Materassi, *Biotechnol. Bioeng.* **42**, 891–898 (1993).
44. U.S. Pat. 5,541,056 (July 22, 1994), M.R. Huntley, P.P. Nijler, and D. Rodalje (to Aquaserch Inc.).
45. S.A. Dvorin, in P. Kretschmer, O. Pulz, and C. Gudin eds., *Proc. 1st Eur. Workshop Microalgal Biotechnology "Algology,"* Print Express K. Beyer, Potsdam, Germany, 1992, pp. 91–92.
46. J.P. Szyper, A.Y. Brandon, J.R. Benemann, M.R. Tredici, and O.R. Zaborsky, *Int. Conf. Biological Hydrogen Production*, Kona, Hawaii, June 23–26, 1997.
47. M.R. Tredici, G. Chini Zittelli, and J.R. Benemann, *Int. Conf. Biological Hydrogen Production*, Kona, Hawaii, June 23–26, 1997.
48. A. Richmond, S. Boussiba, A. Vonshak, and R. Kopel, *J. appl. Phycol.* **5**, 327–332 (1993).
49. Y.K. Lee, S.Y. Ding, C.S. Low, Y.C. Chang, W.L. Forday, and P.C. Chew, *J. appl. Phycol.* **7**, 47–51 (1995).
50. E.A. Davis, J. Dedrick, C.S. French, H.W. Milner, J. Myers, J.H.C. Smith, and H.A. Spoehr, in J.S. Burlew ed., *Algal Culture from Laboratory to Pilot Plant*, Carnegie Institution of Washington Publication No. 600, Carnegie Institution, Washington, D.C., 1953, pp. 105–153.
51. European Pat. 0,239,272 (March 6, 1987), L.F. Robinson and A.W. Morrison (to Biotechna Limited).
52. L.J. Borowitzka and M.A. Borowitzka, in R.C. Cresswell, T.A.V. Rees, and N. Shah eds., *Algal and Cyanobacterial Biotechnology*, Longman, Harlow, UK, 1989, pp. 294–316.
53. Y. Watanabe, J. de la Noüe, and D.O. Hall, *Biotechnol. Bioeng.* **47**, 261–269 (1995).
54. T. Chrismadha and M.A. Borowitzka, in P.S. Moi, Y.K. Lee, M.A. Borowitzca, and B.A. Whitton eds., *Algal Biotechnology in the Asia Pacific Region*, University of Malaya, Kuala Lumpur, 1994, pp. 122–129.
55. D.B. Anderson and D.E. Eakin, *Biotechnol. Bioeng. Symp.* **15**, 533–547 (1985).
56. A. Ramos de Ortega and J.C. Roux, *Biomass* **10**, 141–156 (1986).
57. Italian Pat. 9357-A/88 (March 11, 1988), M.R. Tredici, D. Mannelli, and R. Materassi (to Consiglio Nazionale delle Ricerche).
58. M.R. Tredici, P. Carlozzi, G. Chini Zittelli, and R. Materassi, *Biores. Technol.* **38**, 153–159 (1991).
59. M.R. Tredici and R. Materassi, *J. appl. Phycol.*, **4**, 221–231 (1992).
60. M.R. Tredici, G. Chini Zittelli, and S. Biagiolini, in P. Kretschmer, O. Pulz, and C. Gudin eds., *Proc. 1st Eur. Workshop Microalgal Biotechnology "Algology,"* Print Express K. Beyer, Potsdam, Germany, 1992, pp. 58–60.
61. B. Pushparaj, E. Pelosi, M.R. Tredici, E. Pinzani, and R. Materassi, *J. appl. Phycol.* **9**, 113–119 (1997).
62. Q. Hu and A. Richmond, *J. appl. Phycol.* **8**, 139–145 (1996).
63. M.R. Tredici and G. Chini Zittelli, in P. Kretschmer, O. Pulz, C. Gudin, and V. Semenenko eds., *Proc. 2nd Eur. Workshop Biotechnology of Microalgae*, Institut fur Getreideverarbeitung, Bergholz-Rehbrucke, Germany, 1995, pp. 21–24.
64. J.C. Weissman, R.P. Goebel, and J.R. Benemann, *Biotechnol. Bioeng.* **31**, 336–344 (1988).
65. K.L. Terry and L.P. Raymond, *Enzyme Microbiol. Technol.* **7**, 474–487 (1985).
66. J.U. Grobbelaar, *J. appl. Phycol.* **6**, 331–335 (1994).
67. E.A. Laws, K.L. Terry, J. Wickman, and M.S. Chalup, *Biotechnol. Bioeng.* **25**, 2319–2335 (1983).
68. P. Carlozzi and G. Torzillo, *Appl. Microbiol. Biotechnol.* **45**, 18–23 (1996).
69. C. Gudin and D. Chaumont, *Bioresource Technol.* **38**, 145–151 (1991).
70. Y.K. Lee and C.S. Low, *Biotechnol. Bioeng.* **38**, 995–1000 (1991).
71. K. Mori, *Biotechnol. Bioeng. Symp.* **15**, 331–345 (1985).
72. C. Solà and F. Gòdia, in J.A. Asenjo and J.C. Merchuk eds., *Bioreactor System Design*, Marcel Dekker, New York, 1995, pp. 511–552.

BIOREDUCTION

TADASHI KOMETANI
Toyama College of Technology
Toyama-shi, Japan

KEY WORDS

Asymmetric reduction
Chiral alcohol
Coenzyme regeneration
NAD(P)H
Resin

OUTLINE

INTRODUCTION

Bioreduction is one of the important biological transformations catalyzed by whole cells. The reducing action of fermenting baker's yeast (*Saccharomyces cerevisiae*) was first reported during the reduction of furfural to furfuryl alcohol. Since the early twentieth century, numerous bioreductions using microorganisms other than baker's yeast have been studied by both biotechnologists and organic chemists. Though the ketone, C═C double bond, azide, and nitro compound are reduced, synthetic chemists have mainly applied the microbial reduction of prochiral ketones to the synthesis of chiral alcohols in an enantiomerically pure form. Until now, a large number of references to scientific studies in this area have been reported.

During the bioreduction of carbonyl compounds, carbonyl reductases in microorganisms catalyze the stereospecific reduction using NADH, NADPH, and, sometimes, the flavine mononucleotide as hydrogen donors. The microorganisms can reduce many types of ketones and aldehydes. However, this useful bioreduction has the disadvantage of needing regeneration system of cofactors, such as NAD(P)H, in addition to the reductases in whole cells. To overcome this problem, it is important to utilize the ability of the living microorganism to regenerate NAD(P)H from $NAD(P)^+$ by their natural process of respiration (Fig. 1).

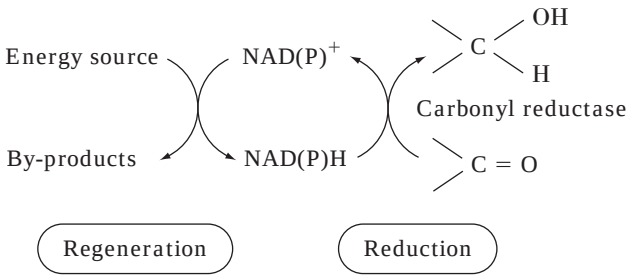


Figure 1. Outline of bioreduction.

This article describes the bioreduction of ketones with microbial cells, including reductases and nicotinamide coenzyme regeneration. An application of the bioreduction to the synthesis of an optically active alcohol is then described.

REDUCTASES OF KETONE

The most important microorganism catalyzing bioreduction is baker's yeast. The scientific studies devoted to bioreduction using baker's yeast have gradually increased since the 1950s. The transformations of various classes of synthetic organic molecules were examined. These transformations were reviewed from the viewpoint of their chemical reactions in three excellent articles published in 1990, 1991, and 1992 (1–3). In these reviews, the uses of baker's yeast in synthetic organic chemistry are fully covered.

Other microorganisms also catalyze the bioreduction of ketones. However, the reports on other microorganisms do not outnumber those on baker's yeast. The main difficulty in the widespread use of microorganisms among the organic chemists' community certainly depends on their familiarity with microbiological techniques. Therefore, good cooperation between synthetic organic chemists and microbiologists oriented toward such applications can be fruitful.

Baker's Yeast

Baker's yeast can be used by people not trained in microbiology and can be considered a useful "reagent," easily accessible in every laboratory. Fermenting baker's yeast using glucose in tap water as an energy source is generally used for highly enantioselective reactions. However, when the process is not completely enantioselective, simple modifications of the experimental conditions may influence the stereochemistry and the enantiomeric excess of the product. For instance, the use of organic media, the addition of other compounds, a change in energy source, and immobilization techniques are helpful.

Traditionally, baker's yeast can reduce variously substituted carbonyl groups to the corresponding hydroxyl compounds, the stereochemical outcome of these highly enantioselective reductions depends on the presence of dehydrogenases, which generally follow the so-called Prelog's rule. Several exceptions have been found, which can be due to the presence of competing dehydrogenases with different stereochemical requirements.

The baker's yeast–mediated reduction of acetol and ethyl acetoacetate afforded (*R*)-propylene glycol and (*S*)-ethyl 3-hydroxybutanoate, respectively. The large-scale preparations of these compounds were accomplished by using a bubble-column reactor and ethanol as the energy source in more than 98% enantiomeric excess (4). The usefulness of the yeast-mediated reduction was also shown in the diltiazem synthesis. The transformation of a racemic ketone (*RS*)-2-(4-methoxyphenyl)-1,5-benzothiazepin-3,4(2*H*,5*H*)-dione afforded only the desired (2*S*,3*S*)-2,3-dihydro-3-hydroxy-2-(4-methoxyphenyl)-1,5-benzothiazepin-4(5*H*)-one, a key intermediate in the diltiazem synthesis. The difficulty in its industrial application because of its low productivity, which is due to the low solubility of the substrate in water, was overcome by changing its physical properties. Thus, 100 g of the substrate in 1 L of reaction medium was reduced to produce the chiral intermediate with over 99% ee in 80% yield (5) (Fig. 2).

Other Microorganisms

Bioreduction using other microorganisms was systematically examined by Shimizu and Yamada using ketopantoyl lactone as the substrate (6). They found that about 10% of the microorganisms tested (yeasts, molds, bacteria, actinomycetes, and basidiomycete) reduced ketopantoyl lactone to pantoyl lactone in the presence of glucose. Their enantioselectivities varied at random and were independent of genus and species.

D-(−)-Pantoyl lactone, the γ-lactone of D-(−)-pantoic acid, is an important starting material for the synthesis of the vitamin D-(+)-pantothenic acid. The conventional synthesis involved the tedious chemical resolution of racemic pantoyl lactone. In this synthesis, the enantioselective bioreduction of ketopantoyl lactone to D-(−)-pantoyl lactone was examined. When *Candida parapsilosis* or *Rhodotorula minuta* cells were incubated with ketopantoyl lactone (more than 50 g/L) and glucose, D-(−)-pantoyl lactone was produced in excellent chemical and optical yields. In this reaction, the asymmetric reduction is catalyzed by a newly found carbonyl reductase specific to conjugated polyketones, which requires NADPH as a cofactor. Another approach is the enantioselective reduction of ketopantoic acid. *Agrobacterium* sp. was screened and incubated with ketopantoic acid and fructose to give D-(−)-pantoic acid. The acid was easily cyclized with hydrochloric acid to afford the D-(−)-pantoyl lactone (Fig. 3).

For the stereospecific reduction of 4-chloroacetoacetate, each stereoisomer of the product (*R*)- or (*S*)-4-chloro-3-hydroxybutanoate is a useful compound. Shimizu and Yamada also screened a microorganism for the asymmetric reduction of 4-chloroacetoacetate. Almost all microorganisms, such as *R. minuta* and *Paecilomyces marquandii*, reduce to (*S*)-4-chloro-3-hydroxybutanoate, which converts to (*S*)-1-benzyl-3-hydroxypyrrolidine. In the carnitine synthesis, 4-chloroacetoacetate should be reduced to (*R*)-4-chloro-3-hydroxybutanoate, which would be converted to carnitine by the chemical method. *Sporobolomyces salmonicolor, Micrococus luteus*, and *Cellulumonas* sp. predomi-

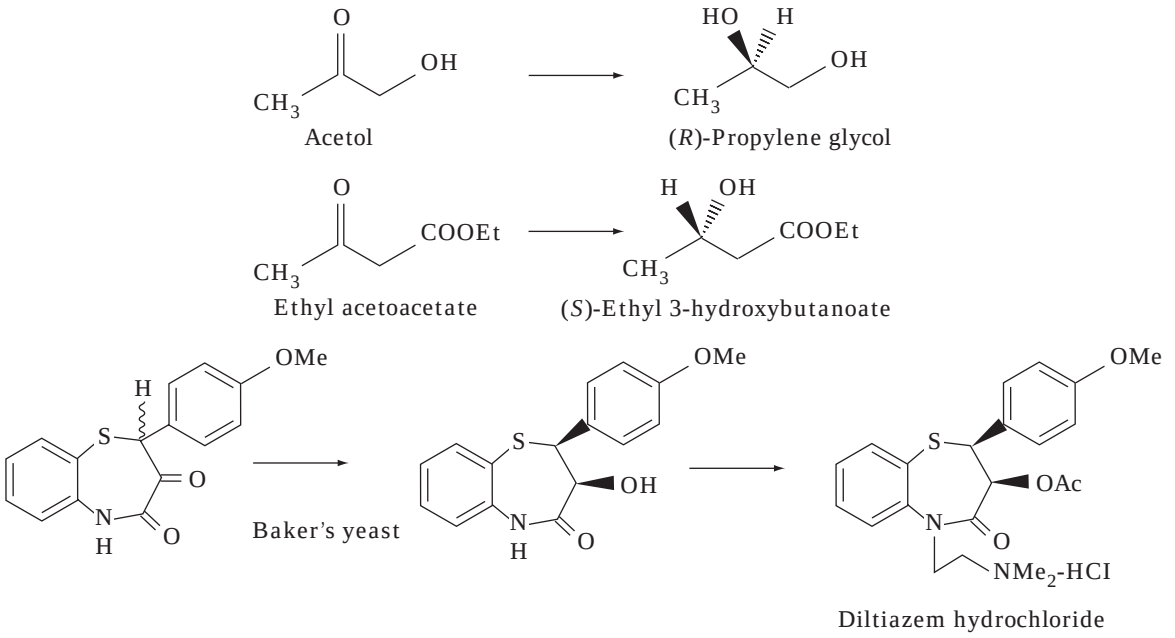


Figure 2. Production of optically pure alcohols through the baker's yeast–mediated reduction.

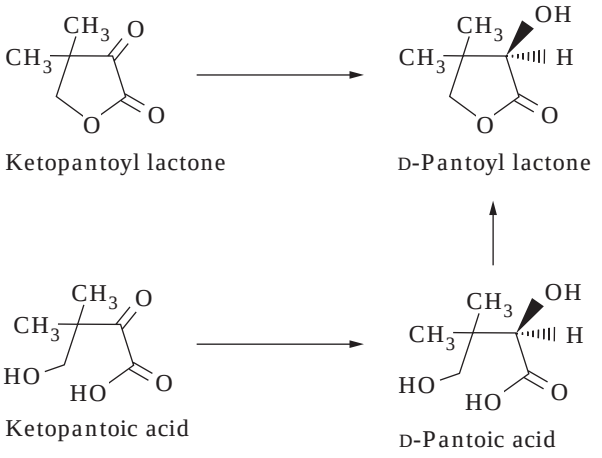


Figure 3. Production of D-pantoyl lactone.

nantly afforded the (*R*)-compound. However, they could not obtain a sufficient result for the enantiomeric excess. Therefore, *S. salmonicolor* was treated with acetone to remove (*S*)-reductase. The treated cells were used for the stereospecific reduction of ethyl 4-chloroacetoacetate to afford (*R*)-4-chloro-3-hydroxybutanoate in a water–organic solvent two-phase system. In this synthesis, NADPH must be regenerated from $NADP^+$. The regeneration was accomplished by using glucose and glucose dehydrogenase in the aqueous phase. This process can be used in the carnitine synthesis (Fig. 4).

NAD(P)H REGENERATION

During bioreduction, NAD(P)H coenzymes must be successively regenerated from the corresponding oxidized form in the cells. The system responsible for the coenzyme regeneration is coupled with the reduction of a substrate. A large amount of work has been devoted to developing and testing various types of systems for NAD(P)H regeneration.

D-Glucose Derivatives

Yamada et al. characterized the enzymes responsible for the NADPH regeneration system using glucose as the energy source (7). They examined the NADPH-forming ability of the cell-free extracts of *S. cerevisiae, Candida parapsilosis*, and *Mucor ambiguus*, with several sugars as the substrate. NADPH was not formed from D-fructose 1,6-diphosphate. This suggests that NADPH could not be produced after D-fructose 6-phosphate was phosphorylated to D-fructose 1,6-diphosphate. They found that NADPH is produced from the hexose monophosphates, D-glucose 1-phosphate, D-glucose 6-phosphate, or D-fructose 6-phosphate. Thus, with *S. cerevisiae*, *C. parapsilosis*, and *M. ambiguus*, the enzymes were shown to be D-glucose 6-phosphate dehydrogenase and 6-phospho-D-gluconate dehydrogenase in the hexose monophosphate pathway. By means of these systems, endogenous NADPH can be efficiently regenerated by the simple mixing of microbial cells with glucose. The reductions of various carbonyl compounds by microorganisms in synthetic organic chemistry will be mainly based on this system.

Ethanol

During the baker's yeast–mediated reduction using ethanol as the energy source, NAD(P)H is regenerated from $NAD(P)^+$ through the oxidative pathway of ethanol to carbon dioxide, and the system is also efficiently coupled with reduction of a ketone (4). The enzyme system responsible for NAD(P)H regeneration was investigated by studying the reduction of acetol and ethyl acetoacetate, catalyzed by NADH- and NADPH-dependent carbonyl reductases, respectively. NAD^+ formed during the reduction of acetol is reduced to NADH through the oxidative pathway of ethanol to acetate, catalyzed by alcohol dehydrogenase and aldehyde dehydrogenase. On the other hand, $NADP^+$ formed during the reduction of ethyl acetoacetate is reduced to NADPH through the oxidative pathway of acetate to carbon dioxide under aerobic conditions. The enzyme responsible for NADPH regeneration must be malate dehydrogenase (decarboxylating) in leaving the TCA cycle. By using the NAD(P)H regeneration system, the baker's yeast–mediated large-scale productions of (*R*)-propylene glycol from acetol and (*S*)-ethyl 3-hydroxybutanoate from ethyl acetoacetate were efficiently accomplished. Other yeasts also use ethanol as the energy source for the NAD(P)H regeneration system in the same manner (8).

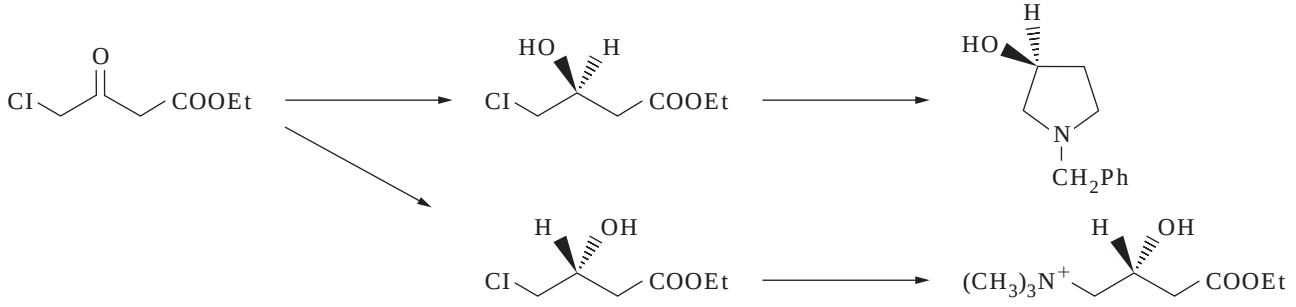


Figure 4. Production of (*S*)-1-benzyl-3-hydroxypyrrolidine and (*R*)-carnitine.

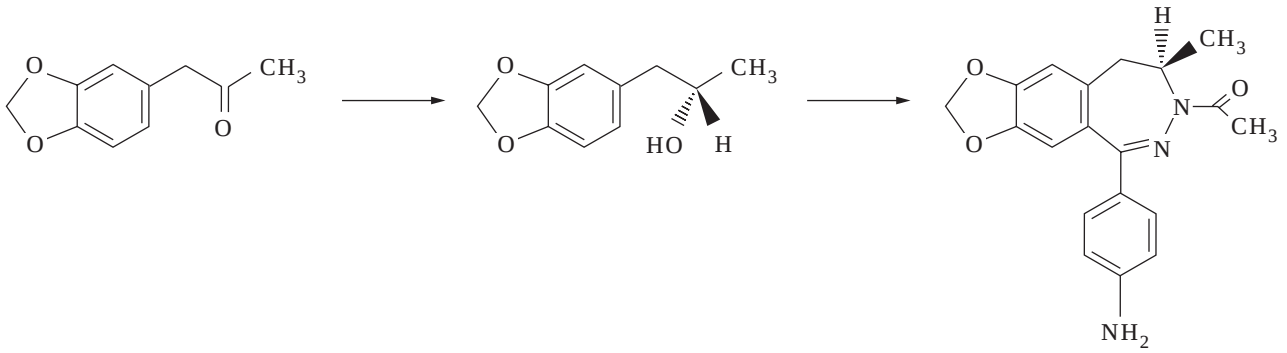


Figure 5. Production of (*S*)-3,4-methylenedioxyphenyl isopropanol for the synthesis of LY300164.

Methanol

Similarly, the reaction involved in the oxidation of methanol to CO_2 by methylotrophic microorganisms, methanol dehydrogenase, formaldehyde dehydrogenase, and formate dehydrogenase is considered to be a very useful system for NADH regeneration.

BIOREDUCTION OF 3,4-METHYLENEDIOXYPHENYLACETONE

The clinical potential for the (*R*)-isomer LY300164 has led to interest in developing syntheses that would not be limited by the inherent inefficiency of the resolution methods and would be amenable to large-scale execution. The retrosynthetic strategy required the synthesis of the benzodiazepine nucleus by intramolecular cyclization of a suitably functionalized hydrazone. This disconnection was chosen because it was anticipated that the required stereocenter could be accessed from the relatively simple alcohol. Asymmetric reduction of the corresponding ketone was considered the most desirable option for the preparation of the optically active alcohol (9,10) (Fig. 5).

Screening of Microorganisms

The chemical method for the enantioselective reduction of methyl benzyl ketones is typically nonselective or requires the stoichiometric use of costly reagents. Unlike closely related examples, however, selectivity and conversion with baker's yeast, as well as other common microorganisms, were unsatisfactory. After an extensive screening of microorganisms, the yeasts *Candida famata* and *Zygosaccharomyces rouxii* could catalyze the reduction of 3,4-methylenedioxyphenylacetone to 3,4-methylenedioxyphenyl-(*S*)-isopropanol with >99% enantiomeric excess in the presence of glucose. However, the productivity number using *C. famata* was too low to be practical and the microorganism was extremely sensitive to concentrations of both the ketone and the alcohol greater than 2 g/L. On the other hand, *Z. rouxii* was more tolerant to higher substrate concentrations and had a higher productivity number.

Practical Application

In this bioreduction, both the product and substrate were toxic to *Z. rouxii*, and this limited the maximum substrate concentration to only 6 g/L. It should be noted that toxicity to the substrate alone would not have been a problem, since this could be readily overcome by the slow addition of the substrate. However, the only way to overcome product toxicity was to devise a method to remove the product from the reaction as it was formed.

For in situ product removal, immiscible organic solvents are commonly used. This type of system is called a two-liquid-phase bioreactor. However, one of the major difficulties is that many organic solvents are highly toxic to microorganisms. In this reaction, polymeric hydrophobic resin (Amberlite XAD-7 from Rohm and Haas, Philadelphia, PA) was used to both supply substrate to, and remove the product from, the reaction mixture. The overall effect is that a high reaction concentration is achieved (40 g/L) while never exposing either the substrate or the product to more than 2 g/L. Thus, in many cases, a polymeric adsorbent resin will provide an attractive alternative to the use of organic solvents in whole cells.

BIBLIOGRAPHY

1. S. Servi, *Synthesis* **1990**, 1–25 (1990).
2. R. Csuk and B.I. Gläzer, *Chem. Rev.* **91**, 49–97 (1991).
3. E. Santaniello, P. Ferraboschi, P. Grisenti, and A. Manzocchi, *Chem. Rev.* **92**, 1071–1140 (1992).
4. T. Kometani, H. Yoshii, and R. Matsuno, *J. Mol. Cat. B: Enzymatic* **1**, 45–52 (1996).
5. T. Kometani, Y. Sakai, H. Matsumae, T. Shibatani, and R. Matsuno, *J. Ferment. Bioeng.* **84**, 195–199 (1997).
6. S. Shimizu and H. Yamada, *J. Synth. Org. Chem. Japan* **49**, 52–62 (1991).
7. M. Kataoka, Y. Nomura, S. Shimizu, and H. Yamada, *Biosci. Biotech. Biochem.* **56**, 820–821 (1992).
8. T. Kometani, Y. Sakai, H. Urai, H. Yoshii and R. Matsuno, *Biosci. Biotech. Biochem.* **61**, 1370–1372 (1997).
9. M.J. Zmijewski, J. Vicenzi, B.E. Landen, W. Muth, P. Marler, and B. Anderson, *Appl. Microbiol. Biotechnol.* **47**, 162–166 (1997).
10. J.F. Vicenzi, M.J. Zmijewski, M.R. Reinhard, B.E. Landen, W.L. Muth, and P.G. Marler, *Enzyme Microb. Technol.* **20**, 494–499 (1997).

BIOREMEDIATION

RONALD L. CRAWFORD
University of Idaho
Moscow, Idaho

KEY WORDS

Air sparging
Biofiltration
Biopiles
Bioreactors
Bioremediation
Bioventing
Composting
Intrinsic remediation
Landfarming
Phytoremediation
White-rot fungi

OUTLINE

INTRODUCTION

As a result of the manufacture, use, spillage, and improper disposal of anthropogenic chemicals and their waste products during past decades, toxic chemicals have become ubiquitous contaminants of soils and waters around the globe. These compounds include everything from petroleum and synthetic organic chemicals to heavy metals and radionuclides. Governments throughout the world now require that the worst of these contamination problems be cleaned up to eliminate, insofar as possible, their hazards to human health and the environment. Traditional methods such as excavation and incineration of contaminated soil or pumping and aboveground treatment of groundwater are proving to be either too expensive, politically unpopular, or both. Thus, innovative technologies are being sought to address the needs of the environmental restoration industry. *Bioremediation* is such a technology.

Bioremediation involves the use of microorganisms, plants, or products yielded by living organisms to degrade, detoxify, or permanently sequester toxic chemicals present in environmental systems. Bioremediation can be used either ex situ or in situ to treat soil, sediment, sludge, water, or air. In ex situ bioremediation procedures, contaminated materials are first removed from the polluted site and then treated, whereas in situ techniques are used to treat contaminants in place without relocating them to a treatment facility. Bioremediation can be promoted by augmenting natural systems with exogenous biological materials or by stimulating indigenous contaminant-transforming microbial populations. Natural systems can be augmented by the introduction of bacteria that degrade the targeted contaminants or plants that sequester pollutants. Native microbial populations are stimulated by adding nutrients or modifying environmental conditions to favor the growth and activity of appropriate organisms. Although intrinsic bioremediation, or microbial processes that occur without human intervention, may ultimately be able to clean up a contaminated site, a great variety of techniques are used by the bioremediation industry to enhance and accelerate degradative processes.

BIOREMEDIATION TECHNOLOGIES

Bioreactors

Bioreactors, or vessels of various configurations (1) containing degradative microbes, are a common means of ex situ treatment of contaminated materials, typically soil that has been excavated from a polluted site or groundwater that has been pumped from a polluted aquifer. A bioreactor may contain, for example, a slurry of soil and water to which nutrients have been added, such as sources of nitrogen, phosphorus, and/or carbon that support the growth and metabolism of degradative microorganisms. Sometimes the microorganisms themselves may have to be added (40). Slurry reactors can be run either aerobically, by stirring and mixing in air or pure oxygen (2), or anoxically, with little or no mixing (3,4). Slurry phase reactor treatment cells can be designed to hold a few thousand to millions of gallons. Aeration systems for aerobic slurry reactors must be able to maintain dissolved oxygen concentrations at 2.0 mg/L or higher. Aeration is a major expense for such systems. To increase oxygen transfer rates, such systems may be sparged with pure oxygen. Often reactor designs include a series of eductors to oxygenate the mixed liquor while minimizing air emissions. Anaerobic slurry systems usually require some mixing devices to evenly distribute soil or sediment materials throughout the aqueous phase, but they spare the need and cost of transferring large amounts of oxygen to the solution phase. Aerobic slurry reactors have been used to treat materials contaminated with chemicals such as polynuclear aromatic hydrocarbons (PAHs) and chlorinated phenols (35). Anaerobic slurry reactors have been used to treat materials polluted by compounds such as chlorinated solvents, polychlorinated biphenyls (PCBs), nitrated munitions like 2,4,6-trinitrotoluene (TNT) (Fig. 1) (4), and nitrated herbicides such as dinoseb (37).

Figure 1. Pilot-scale anaerobic slurry reactor for treating TNT-contaminated soil. *Source*: University of Idaho, with permission.

Water may be treated in simple batch reactors that hold a few hundred to many thousands of liters; nutrients and microorganisms may be added as required. Fluidized beds that contain biomass-bearing carriers are popular for water treatment (5) and can be run aerobically or anaerobically, depending on the contaminants being treated (6,7). Even normally recalcitrant compounds such as aniline and nitrobenzene can be treated in such systems. Alternatively, biomass can be grown on a supporting matrix that is packed into a vessel such as a column, and the water to be treated can be passed through the packed bed to remove contaminants by degradation or absorption (8). Packed beds can also be used to treat vapor-phase contaminants such as chlorinated solvents, which are sparged through the reactor as a separate gaseous stream (9). A novel, mechanically mixed packed-bed reactor that works as a biotrickling filter is a variation on packed-bed technology for treatment of vapor-phase contaminants (43). Another variation on reactor design for water treatment is the upflow sludge blanket reactor using granules containing anaerobic bacteria acclimated to the degradation of a particular waste stream (38). Bioreactors can be used in treatment trains that vary conditions, for example, by alternating anaerobic and aerobic conditions (36), so that mixtures of contaminants or highly chlorinated compounds can be treated successfully.

Biofiltration

Biofiltration is an ex situ treatment for vapor-phase contaminants (44,45). Typically, it involves passing contaminated air through a bed of soil, peat, or compost that has been preacclimated to develop a microbial population that completely degrades the targeted pollutants (10). This type of technology works well for volatile components of petroleum, and for many years it has been used in Europe for odor control (11). Biofilters have been developed for toxic airborne industrial contaminants as diverse as styrene (46), pentane and isobutane mixtures (47), toluene (48), chlorinated benzenes (49), dimethylsulfide (50), and volatile organic compounds (VOCs) (51). For example, a biofiltration process was developed for styrene-containing industrial offgases. This process used peat as filter material. Styrene reduction after 190 days of operation was 70% to 98%. The styrene elimination capacity was 12 g m^{-3} h^{-1} (maximum, 30 g m^{-3} h^{-1}). Degradation required addition of nutrients to the peat, adjustment of the pH, and control of humidity. The optimum operation temperature was about 23 °C, but the styrene removal was still satisfactory at 12 °C (46). Composting of leaves and alfalfa was used for the treatment of toluene-contaminated air in a biofilter. During the thermophilic phase (45–55 °C), toluene biodegradation rates reached 110 g toluene m^{-3} h^{-1} at an inlet concentration of about 5 g/m^3 and a gas residence time of 90 s (48). Removal rates by biofilters for chlorinated benzenes reached 300 g solvent vapor h^{-1} m^{-3} of filter volume (49).

Air Sparging

The biodegradation of many contaminants in groundwater is limited by the shortage of electron acceptors for oxidative processes. The most effective electron acceptor for the bioremediation of many contaminants is oxygen, which is present in groundwater at only 8–10 mg/L even under optimal conditions. Air sparging is an in situ technology for introducing oxygen into contaminated aquifers (12). Air or pure oxygen is injected into the aquifer through sparging wells or infiltration galleries (13). Potential problems associated with the introduction of oxygen into groundwater include biofouling, in which injection wells become plugged with biomass, and obstruction of well screens or subsurface geologic formations by iron complexes that precipitate as a result of the oxidation of Fe(II) to Fe(III). Hydrogen peroxide, which decomposes in the presence of iron or biomass to produce oxygen (14), is sometimes substituted for air or pure oxygen. Even when oxygen is supplied by sparging, it can be used so rapidly that its supply can still be rate limiting for bioremediation processes. Thus, the bioremediation industry has developed oxygen-releasing compounds (ORCs) composed of materials such as magnesium peroxide that allow controlled, long-term release of oxygen when introduced into boreholes as a slurry (15). Most successes with air sparging, as with many other bioremediation systems, have been reported in petroleum-contaminated aquifers. The reintroduction of indigenous microorganisms isolated from a contaminated site after culturing is one approach to in situ hydrocarbon bioremediation. It is effective especially when microorganism growth is supplemented by addition of oxygen and nitrogen (52). At the Moffett Field (Sunnyvale, Calif.) groundwater test site, in situ removal of *cis*-dichloroethene (*c*-DCE) and trichloroethene (TCE) coincided with biostimulation through phenol and oxygen injection and utilization, with c-DCE removed more rapidly than TCE. Greater TCE and c-DCE removal was observed when the phenol concentration was increased. More than 90% removal of *c*-DCE and TCE was observed in a 2-m biostimulated zone (53). Thus, introduction of oxygen into contaminated aquifers can be of use even for recalcitrant compounds such as VOCs.

Bioventing

Bioventing is related to air sparging in that the objective of this in situ technology (17) is to provide oxygen to the microbial populations in the subsurface. Bioventing systems use equipment similar to that used in soil vapor extraction (SVE) processes. Air injection or extraction wells or trenches are connected by pipes to blowers or vacuum pumps which either push or pull air through the vadose zone (unsaturated soil horizons) to stimulate microbial populations capable of degrading contaminants. The objective is to provide sufficient oxygen to the subsurface to promote in situ contaminant degradation, but not so much that volatile contaminants are released undegraded to the atmosphere. Thus, bioventing is best used with contaminants that are moderately volatile, such as diesel or jet fuels or residues of aged petroleum (13,21). For bioventing to succeed, the vadose zone must be sufficiently permeable to allow air exchange at least every 24 h (16), so this process is not suitable for all sites. It may be possible in some instances to open up impermeable soils by fracturing them with hydraulic or air pressure injections (26), allowing bioventing strategies to be employed. Bioventing has produced excellent results at hundreds of sites worldwide in recent years (13,18–20).

A typical bioventing system (Fig. 2) involves soil vapor extraction followed by bioventing via vent wells located within an aquifer in areas of high pollutant contamination. Monitoring wells in both the vadose zone (unsaturated soils above the aquifer) and within the aquifer are used to monitor bioventing progress. Neutron probes may sometimes be used for soil moisture monitoring, to assure that venting does not dry out the soil so that microbial populations are inhibited. A plastic liner is usually installed over the surface of part or all of a contaminated area to prevent infiltration of air that might bypass the desired path of air to the vent wells. A catalytic incinerator to destroy extracted vapor that was not degraded by microbes in the vadose zone is often installed as a safety device, especially at sites near populated areas. If bioventing proceeds effectively, there may be no need for treatment of extracted vapors because the concentration of contaminants in the vented gases may be very low. In bioventing, as with all biological processes in soil, soil moisture must be maintained within a range of 6% to 12%, and supplemental phosphorus or nitrogen may be needed (e.g., N as ammonia provided in the gas phase) to maintain a C:N:P ratio of about 100:10:1.

Not all sites are amenable to bioventing. The Battelle Corporation has used a special site evaluation protocol to

Figure 2. Petroleum bioventing operation. *Source*: University of Idaho, with permission.

determine the suitability of a site for remediation by this technique (41). Four steps are performed sequentially for characterization of a potential bioventing site:

1. A soil-gas survey is conducted to determine if the subsurface is oxygen limited. Soil gas is analyzed for concentrations of oxygen and carbon dioxide. Oxygen is, of course, required for successful bioventing of many contaminants, such as petroleum hydrocarbons.
2. A soil-gas permeability test is conducted to see if air can be injected at rates sufficient to aerate the vadose zone. This test requires injection and withdrawal of air for measuring changes in subsurface pressures at discrete distances from the point of injection. Soils must be sufficiently permeable to be candidates for bioventing. "Tight" soils (e.g., clay-rich soils) can usually be eliminated at this stage.
3. Oxygen-limited but sufficiently permeable soils are then subjected to an in situ respiration test to measure biological contaminant degradation rates for comparison to respiration rates in similar, but uncontaminated, soils. This involves aerating the soils for 24 h and then monitoring the oxygen and carbon dioxide concentrations in the soil-gas phase after the final air injection. An oxygen utilization rate is calculated and used to estimate a contaminant degradation rate, which can in turn be used to estimate the time that will be required to biovent the site.
4. Finally, a pilot- and/or full-scale field bioventing remediation demonstration is conducted at the site. There may or may not be a need for pilot-scale testing before a full-scale process is implemented. This is determined on a site-specific basis according to the prior testing.

Landfarming

Landfarming is an aboveground ex situ process in which contaminated soils are excavated, supplemented with nutrients (primarily sources of nitrogen or phosphorus), and spread in relatively thin layers on a prepared site so that biodegradation of contaminants is promoted. These soils are aerated by periodic tilling and kept moist by irrigation, thus explaining the name of this procedure (13). Sometimes soils are spread directly on the ground at a graded site, if there is no danger that the contaminants will leach through the soil to groundwater or vaporize to dangerous levels in the atmosphere. More commonly, soils are spread on sites with preinstalled plastic liners or clay barriers and leachate collection systems to prevent any possible leach-

ing of contaminants offsite or to groundwater, as well as vapor containment enclosures, especially if the site is located near populated areas. Petroleum is commonly treated in this manner (20). Highly toxic or biologically recalcitrant compounds generally should not be landfarmed, and the presence of heavy metals such as lead usually indicates that landfarming is not an appropriate technology for a particular soil. Some advantages and disadvantages of landfarming of petroleum-contaminated soil are shown in Table 1 (data from U.S. EPA).

The presence of vegetation can enhance the degradation of contaminants during landfarming of contaminated soil. Greenhouse experiments have shown the degradation of PAHs to be greater in the presence of plants than in their absence. PAHs were detectable in the plant tissue, but overall uptake was not significant. Enhanced microbial activity appeared to be responsible for the increased removal of PAHs. Thus, planting of vegetation is promising as an inexpensive means of speeding landfarm-based remediation of sites contaminated with petroleum (42).

Biopiles

Biopiles are basically soil-heaping operations, a variation on landfarming. The primary advantages of biopiles over landfarms are that biopiles save land space, and air emissions can be more inexpensively captured and treated (20). In this technique soil is excavated and piled into a mound within a lined treatment cell. The pile is typically covered by a plastic sheet, and an internal system of perforated pipes is usually installed to draw air in by means of vacuum pumps. Alternatively, air may be blown into the pile through the internal pumping system, or supplied from oxygen-releasing compounds (ORCs; see "Air Sparging") incorporated into the piles. Water, nutrients, and microbial inoculants are added as needed, and leachates and gaseous emissions are collected for separate treatment. In theory, biopiles can be used for most biodegradable contaminants, but the limitations of landfarming generally apply. In practice, most biopile operations have treated petroleum-contaminated soils (20).

Table 1. Advantages and Disadvantages of Landfarming for Treatment of Petroleum-Contaminated Soil

Advantages	Disadvantages
Relatively simple to design and implement	Concentration reductions >95% and constituent concentrations <0.1 ppm are very difficult to achieve
Short treatment times (usually 6 months to 2 years under optimal conditions)	May not be effective for high constituent concentrations (>50,000 ppm total petroleum hydrocarbons)
Cost competitive: $30–60/ton of contaminated soil	Presence of significant heavy metal concentrations (>2,500 ppm) may inhibit microbial growth
Effective on organic constituents with slow biodegradation rates	Volatile constituents tend to evaporate rather than biodegrade during treatment
	Requires a large land area for treatment
	Dust and vapor generation during aeration may pose air quality concerns
	May require bottom liner if leaching from the landfarm is a concern

Source: U.S. Environmental Protection Agency (*http://www.epa.gov/swerust1/cat/landfarm.htm*).

"Biomounding" is a low-technology, low-cost variation of biopiles. It has been used to treat petroleum-contaminated soil, with the addition of bulking materials like animal manure and wood chips to promote good conditions (sufficient moisture retention, good aeration, source of nitrogen, etc.) for building up microbial populations capable of hydrocarbon degradation. This simple process usually is well accepted by local governments and the general public. Biomounding can reach fairly stringent soil cleanup standards in the range of 10 ppm as total petroleum hydrocarbons (TPH). Generally, indigenous bacteria present in petroleum-contaminated soil proliferate under biomounding conditions, so inoculation with exogenous microbes is not normally needed. By engineering a favorable environment for microorganisms in a contaminated soil, it is possible to construct biomounds that are reasonably inexpensive to manage, yet allow fairly fast breakdown of petroleum compounds in soil, without the need for continually turning of the soil as in composting. Simple passive aeration systems normally are used in biomounds. For example, a grid of horizontal piping constructed of flexible perforated drain tile, along with vertical piping (risers) of PVC, can be used in the system. Multiple layers of piping can be employed, depending on the volume of soil to be treated. Moisture and heat loss from biomounds are usually controlled by covering the mounds with plastic sheeting.

Composting

In composting, excavated soil is mixed with bulking agents such as straw, horse manure, or other agricultural residues to form compost piles, which are then managed for good aeration by periodic turning with specially designed machines. Vigorous microbial activity generates considerable heat, so the piles are termed thermophilic, operating at temperatures around 50–60 °C. Composting has been examined as a treatment for a variety of contaminants, including petroleum sludges (22), chlorophenols (23,24), and munitions residues (25). This approach is still experimental, since some questions remain as to its overall efficacy in treating some contaminants. For example, nitroaromatic compounds such as 2,4,6-trinitrotoluene (TNT) tend to be polymerized or incorporated into soil humus fractions, with little or no conversion to carbon dioxide or products that might themselves be degradable to carbon dioxide (25). The long-term stability and potential for release of toxic compounds from such uncharacterized materials in soil is not known.

In one study, a reduction in direct lethality of TNT to earthworms reflected the removal of explosive compounds by composting, as indicated by chemical analysis. How-

ever, assays by the Mutatox procedure indicated increased mutagenicity, a result that would not have been shown by chemical analysis alone, which implied that mutagenic metabolites of explosives were formed during composting and were incompletely degraded (54). In another study, [^{14}C]TNT was subjected to composting during which much of the ^{14}C becomes incorporated into the "humic" fraction of the compost where it supposedly becomes unavailable. The composted TNT resulting from this process was then fed to rats. Although only 2.3% of the total ^{14}C dose appeared in the rats' urine during the first 3 days after feeding, low levels of ^{14}C continued to be excreted in urine from compost-treated rats for more than 6 months. The total amount of ^{14}C excreted in the urine was comparable to that excreted by animals treated with pure TNT (55). Thus, it is clear that composted TNT is not irreversibly immobilized, but that TNT or its biotransformation products in compost are bioavailable. Thus, composting may not be the treatment of choice for some wastes where there is little mineralization of the target pollutant.

Despite the experimental nature of composting, it is likely to become a useful tool for treatment of some chemically polluted soils. Variations on composting appear to be possible to overcome some of these problems. For example, TNT is degraded more effectively when the compost is allowed to first undergo a fairly long anaerobic treatment phase before traditional aerobic composting (39). The critical stage of the traditional composting process is the thermophilic (hot) phase. The fate and mechanism of removal of chemicals such as pesticides and other xenobiotic compounds in composting processes are largely unknown and in need of comprehensive analysis (57–59).

The hazards of composting stem not only from chemical agents but also from microbiological agents (61,62), whose presence depends heavily on composting methods. Further research clearly is needed on the risks related to human exposure to chemical residues and microorganisms in composted wastes (56). The fate of toxic heavy metals in compost applied to land is of particular concern (60), as is the danger to the public health from pathogenic microorganisms such as *Aspergillus fumigatus*. Spores of *A. fumigatus* (the causal agent of the lung disease aspergillosis) are abundant in the outdoor air at some composting sites (61,62), indicating the importance of site location among process and design criteria for composting operations. In summary, although composting is basically a simple, affordable process, its efficacy and safety for treatment of hazardous wastes are still open to question.

Intrinsic Bioremediation

Intrinsic remediation has been defined as the combined effect of natural destruction and other nondestructive processes to reduce the mobility and mass, and therefore the associated risks, of a contaminant (27). The mechanisms that mediate intrinsic remediation of a contaminant include biodegradation by indigenous microbial populations, sorption to, or trapping within matrixes that make the compound nonbioavailable, and loss of mass by dilution or volatilization (Fig. 3).

Intrinsic remediation of petroleum hydrocarbons is well documented (28,29), and it has also been confirmed for

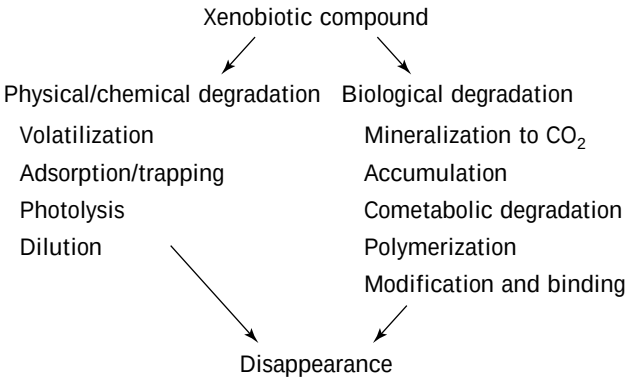


Figure 3. Intrinsic Bioremediation Processes.

chlorinated hydrocarbons, typically under anoxic environmental conditions that favor reductive dehalogenation processes (27,28). Before intrinsic processes can be accepted as a cleanup option for a particular site, regulators must agree to accept exposure and risk assessment analyses as the basis for determining cleanup end points. Therefore, in sites undergoing intrinsic remediation, the distributions of contaminants (e.g., the extent of contaminant plume migration in relation to sensitive receptors) must be thoroughly characterized, and it must be possible to confirm that pollutant remediation is proceeding at a rate sufficient to reduce risks to human health and the environment in an acceptable period. To confirm that intrinsic processes are effectively removing contaminants from soil or groundwater, several factors can be periodically measured, including the following:

- Contaminant mass and its environmental distribution
- Quantity of contaminant degraders or transformers, as indicated by microbial counts
- Temperature and pH, to ensure that they are conducive to microbial activity
- Types of electron donors and acceptors and changes in their concentrations; for example, nitrate, nitrite, metal species (especially iron), methane, ammonia, carbon dioxide, and sulfate
- Redox potential
- Other factors determined by site and contaminant

Ideally, the long-term efficacy of intrinsic remediation processes could be predicted at individual sites through modeling. It is obvious that intrinsic remediation requires specific, extensive, and long-term activities at a site and is far from a "do-nothing" option.

Phytoremediation

Phytoremediation, in which plants are used to remediate contaminated soils or waters, has been studied mainly for the removal of toxic metals but has also been proposed for remediation of some organic contaminants (30). Phytoremediation can take place through the phytoextraction of

contaminants from soil or water, where contaminants are taken up by a plant and concentrated in the plant tissues (31,32). The plant materials can then be harvested, removing the sequestered contaminant from the environment. Bacteria associated with the root zones of plants may in some instances degrade organic contaminants, although more study in this emerging area of research is required to confirm which plants promote degradation of which contaminants. Plants may remove contaminants from soil by transpiration of water to the atmosphere, carrying volatile contaminants along. Plants can help in the management and stabilization of polluted sites by reducing soil erosion caused by wind and water and by reducing the infiltration of water into soil, which will in turn slow down or eliminate the leaching of mobile contaminants into groundwater (30,32). An especially promising line of phytoremediation research concerns the use of plants to remove radionuclides from soil or water (33).

After plants or trees have concentrated metals or radioactive compounds, they can be harvested and then burned, which reduces the volume of the collected material by about 95%. Concentration and volume reduction makes disposal of the ash far less expensive than disposal of the original material itself. Alternatively, useful heavy metals might be extracted from the ash as industrial products, especially if the ash contains one or a few valuable metals. It has been suggested that if a plant contained 1% by weight of a metal, then its ash would contain 20% by weight of that metal, a higher metal content than is found in most ores. Any remaining residue after metal recovery would be disposed of as a toxic material.

Phytoremediation does have limitations. It can be slow, so it is not an option for sites that represent an immediate public health or environmental threat. Plant roots also may not penetrate the soil as deeply as necessary to access the targeted contaminants. The roots of most plants generally reach depths of only 30 cm to about 10 m, although a few plants, such as certain varieties of poplar trees, may send their roots somewhat deeper. Thus, deep contamination can be inaccessible to phytoremediation. Furthermore, as living organisms, plants have specific environmental and nutritional needs. They require a rather narrow range of growth temperatures, grow well only in the presence of favorable soil qualities (texture, water-holding capacity, pH, etc.), must obtain enough water to support growth and metabolism, and usually require the presence of oxygen. Plants may also be inefficient at taking up some pollutants, allowing contaminants to reach groundwater.

Some plants proposed for use in phytoremediation (*http://www.ecological-engineering.com/phytorem.html*) include alfalfa (symbiotic with hydrocarbon-degrading bacteria); *Arabidopsis* (genetically engineered to carry a bacterial gene that transforms mercury into a gaseous state); bladder campion (accumulates zinc and copper); *Brassica juncea* (Indian mustard greens; accumulates selenium, sulfur, lead, chromium, cadmium, nickel, zinc, and copper); Buxaceae (boxwood; accumulates nickel); the Compositae family (symbiotic with bacteria; accumulates cesium and strontium); Euphorbiaceae (succulent; accumulates nickel); tomato (accumulates lead, zinc, and copper); poplar (used in the absorption of the pesticide atrazine, and solvents such as carbon tetrachloride); and *Thlaspi caerulescens* (alpine pennycress; accumulates zinc and cadmium).

White-Rot Fungi

White-rot fungi (Fig. 4) produce enzymes that degrade the plant polymer lignin and oxidize a great variety of chemicals, including many environmental pollutants (34). Because these fungi degrade such a broad range of anthropogenic compounds, they have been examined extensively as potential bioremediation agents. The peroxidases of the basidiomycete *Phanerochaete chrysosporium* are the most thoroughly investigated of the xenobiotic compound–degrading fungal enzymes. These and other oxidative fungal enzymes such as laccase are thought to be the primary catalysts responsible for initiation of degradation of many xenobiotics. For example, enzyme measurements indicated that both high and relatively stable activities of Mn-dependent peroxidase, Mn-independent peroxidase, lignin peroxidase, and laccase characterized efficient PCB degraders among a selection of white-rot fungi (63). For application in the field, white-rot fungi grown on a substrate of wood chips or other specially formulated lignocellulosic materials are mixed into a polluted soil. The soil is tilled, watered, and generally managed to encourage the growth and lignin-degrading activity of the introduced organism. As the fungus degrades the lignocellulosic carrier, it simultaneously degrades the contaminants. This technology is considered experimental, because the fungi so far examined are often inhibited by the contaminant concentrations found in some polluted soils, and even when degradation does take place, it may be incomplete. Clearly, more development work needs to be done in contaminated soil systems.

The focus of one recent study was to determine whether *P. chrysosporium* could effectively operate in an actual field sample of oil tar–contaminated soil. The soil was supplemented with [^{14}C]phenanthrene to serve as a representative polynuclear aromatic compound (PAH). This study confirmed earlier work showing that the native soil microflora could acclimate to provide full mineralization of PAHs in some contaminated soils. However, this mineralization could be enhanced by supplementation with *P. chrysosporium* (64). In another soil study, investigators examined the role of fungal oxidative enzymes in the degradation of soil contaminants by looking at gene expression in the soil matrix. Among the genes implicated in pollutant degradation were two closely related lignin peroxidase genes, whose mRNA transcripts were detected in the soil under study (65). Thus, the use of white-rot fungi appears quite promising for certain pollutants such as the PAHs (64) and chlorinated phenols (34). However, more proof of efficacy is required before this becomes a commonly used bioremediation technique. Hundreds of species of white-rot fungi have not yet been examined as bioremediation agents, so the possibility of identifying strains more suitable for fungal bioremediation is good.

LIMITATIONS OF BIOREMEDIATION

As a biological process, bioremediation has the limitations inherent in any biological system, so it may not always be

Figure 4. The white-rot fungus *Ganoderma aplanatum* (common name, Artist's conk). (Photographed by University of Idaho Photo Services.)

an appropriate technology for treating a contaminated site. For a site to be treatable by bioremediation, environmental conditions must support biological activity because extremes of pH, temperature, radioactivity, or redox potential may not be tolerated by degradative organisms. Essential nutrients such as nitrogen, phosphorus, sulfur, and trace elements must be present and available. The soil or water to be treated by bioremediation must be free of extremely high concentrations of certain toxic chemicals (e.g., the heavy metals mercury, lead, and zinc) and other antimicrobial substances. The concentrations of contaminants may be too low to provide sufficient energy for microbial growth, and cometabolism may not be an option. Contaminants may not be bioavailable, or accessible to microorganisms, but they may still be subject to remediation standards required by regulatory agencies. In complex mixtures of contaminants, the feasible bioremediation processes may be unable to remove all the components. The site to be treated, a vast, deep aquifer, for example, may not be accessible to engineers. Finally, some xenobiotic compounds simply are not biodegradable. Thus, bioremediation must be viewed as only a single tool in the environmental restoration toolbox, one that will be used frequently but not for all sites or situations.

BIBLIOGRAPHY

1. W. Admassu and R.A. Korus, in *Bioremediation: Principles and Applications*, Cambridge Univ. Press, Cambridge, UK, 1997.
2. M.D. LaGrega, P.L. Bunkingham, and J.C. Evans, *Hazardous Waste Management*, McGraw-Hill, New York, 1994.
3. S.B. Funk, D.L. Crawford, D.J. Roberts, and R.L. Crawford, *Bioremediation of Pollutants in Soil and Water*, ASTM STP 1235, American Society for Testing and Materials, Philadelphia, Pa., 1995, pp. 177–189.
4. S.B. Funk, D.L. Crawford, R.L. Crawford, G. Mead, and W. Davis-Hoover, *Appl. Biochem. Biotechnol.* **51**, 625–633 (1995).
5. E.K. Nyer, *Groundwater Treatment Technology*, Van Nostrand Reinhold, New York, 1992.
6. T.C. Voice, X. Zhao, J. Shi, and R.F. Hickey, in R.E. Hinchee, G.D. Sayles, and R.S. Skeen eds., *Biological Unit Processes for Hazardous Waste Treatment*, Battelle Press, Columbus, Ohio, 1995, pp. 29–36.
7. G.D. Sayles and M.T. Suidan, in M.A. Levin, and M.A. Gealt eds., *Biotreatment of Industrial and Hazardous Waste*, McGraw-Hill, New York, 1993.
8. R.B. King, *Practical Environmental Bioremediation*, CRC Press, Boca Raton, Fla., 1992.
9. E.H. Marsman, J.M.M. Appleman, L.G.C.M. Urlings, and B.A. Bult, in R.E. Hinchee, D.B. Anderson, and F.B. Metting, Jr. eds., *Applied Biotechnology for Site Remediation*, Lewis Publishers, Boca Raton, Fla., 1994, pp. 391–399.
10. J.M. Yudelson and P.D. Tinari, in R.E. Hinchee, G.D. Sayles, and R.S. Skeen eds., *Biological Unit Processes for Hazardous Waste Treatment*, Battelle Press, Columbus, Ohio, 1995, pp. 205–209.
11. S.P.P. Ottengraf, *Biotechnology*, **8**, 301–332 (1986).
12. R.L. Raymond, V.W. Jamison, and J.O. Hudson, API Publication 4427, American Petroleum Institute, Washington, D.C., 1975.
13. G.F. Bowlen and D.S. Kosson, in L.Y. Young, and C.E. Cerniglia eds., *Microbial Transformation and Degradation of Toxic Organic Chemicals*, Wiley-Liss, New York, 1995, pp. 515–542.
14. W.T. Frankenberger, Jr., K.D. Emerson, and D.W. Turner, *Environ. Manage.* **13**, 325–332 (1989).
15. C.-M. Kao and R.C. Borden, in R.E. Hinchee, B.C. Alleman, R.E. Hopple, and R.N. Miller eds., *Hydrocarbon Bioremediation*, Lewis Publishers, Boca Raton, Fla., 1994, pp. 262–266.

16. R.E. Hinchee and S.K. Ong, *J. Air Waste Manage. Assoc.* **42**, 1305–1312 (1992).
17. D.F. Kidd, *In Situ Treatment Technology*, Lewis Publishers, Boca Raton, Fla., 1996, pp. 245–269.
18. R.R. Dupont, *Environ. Prog.* **12**, 45–53 (1993).
19. R.E. Hinchee, D.C. Downey, and R.R. Dupont, *J. Hazard. Mater.* **27**, 315–325 (1991).
20. I.D. Bossert and G.C. Compeau, in L.Y. Young and C.E. Cerniglia eds., *Microbial Transformation and Degradation of Toxic Organic Chemicals*, Wiley-Liss, New York, 1995, pp. 77–125.
21. W.T. Frankenberger, Jr. and U. Karlson, *On-Site Bioreclamation: Applications and Investigations for Hydrocarbon and Contaminated Site Remediation*, Butterworth-Heinemann, Stoneham, Mass., 1991, pp. 239–254.
22. S.J. McMillin, J.M. Kerr, and N.R. Gray, *Proceedings of the EPA Exploration and Production Environmental Conference*, U.S. Environmental Protection Agency, Washington, D.C., March 3–7, 1993, pp. 389–401.
23. R. Valo and M. Salkinoja-Salonen, *Appl. Microbiol. Biotechnol.* **25**, 68–75 (1986).
24. J.G. Mueller, S.E. Lanz, B.O. Baltmann, and P.J. Chapman, *Environ. Sci. Technol.* **25**, 1055–1061 (1991).
25. R.T. Williams, P.S. Ziegenfuss, and W.E. Sick, *J. Ind. Microbiol.* **9**, 137–144 (1992).
26. D.L. Kaplan, in D. Kamely, A. Chakrabarty, and G.S. Omenn eds., *Biotechnology and Biodegradation*, Portfolio Publishing, Houston, Tex., 1990, pp. 155–182.
27. R.E. Hinchee, J.T. Wilson, and D.C. Downey, *Intrinsic Bioremediation*, Battelle Press, Columbus, Ohio, 1995.
28. H.S. Rifai, R.C. Borden, J.T. Wilson, and C.W. Ward, in R.E. Hinchee, J.T. Wilson, and D.C. Downey eds., *Intrinsic Bioremediation*, Battelle Press, Columbus, Ohio, 1995, pp. 1–29.
29. T.H. Wiedemeier, M.A. Swanson, J.T. Wilson, D.H. Kampbell, R.N. Miller, and J.E. Hansen, in R.E. Hinchee, J.T. Wilson, and D.C. Downey eds., *Intrinsic Bioremediation*, Battelle Press, Columbus, Ohio, 1995, pp. 31–51.
30. H. Bolton, Jr. and Y.A. Gorby, in R.E. Hinchee, J.L. Means, and D.R. Burris eds., *Bioremediation of Inorganics*, Battelle Press, Columbus, Ohio, 1995, pp. 1–16.
31. R.T. Kelly and T.F. Guerin, in R.E. Hinchee, J.L. Means, and D.R. Burris eds., *Bioremediation of Inorganics*, Battelle Press, Columbus, Ohio, 1995, pp. 25–32.
32. S.D. Cunningham, W.R. Berti, and J.W. Huang, in R.E. Hinchee, J.L. Means, and D.R. Burris eds., *Bioremediation of Inorganics*, Battelle Press, Columbus, Ohio, 1995, pp. 33–54.
33. J.E. Cornish, W.C. Goldberg, R.S. Levine, and J.R. Benemann, in R.E. Hinchee, J.L. Means, and D.R. Burris eds., *Bioremediation of Inorganics*, Battelle Press, Columbus, Ohio, 1995, pp. 55–63.
34. A. Paszczynski and R.L. Crawford, *Biotechnol. Prog.* **11**, 368–379 (1995).
35. C. Barbeau, L. Deschenes, D. Karamanev, Y. Comeau, and R. Samson, *Appl. Microbiol. Biotechnol.* **48**, 745–752 (1997).
36. B.S. Evans, C.A. Dudley, and K.T. Klasson. *Appl. Biochem. Biotechnol.* **57–58**, 885–894 (1996).
37. R.H. Kaake, D.J. Roberts, T.O. Stevens, R.L. Crawford, and D.L. Crawford, *Appl. Environ. Microbiol.* **58**, 1683–1689 (1992).
38. W.M. Wu, L. Bhatnagar, and J.G. Zeikus, *Appl. Environ. Microbiol.* **59**, 389–397 (1993).
39. J. Breitung, D. Bruns-Nagel, K. Steinbach, L. Kaminski, D. Gemsa, and E. von Low, *Appl. Microbiol. Biotechnol.* **44**, 795–800 (1996).
40. F.R. Brunsbach and W. Reineke, *Appl. Microbiol. Biotechnol.* **42**, 415–420 (1994).
41. R.E. Hinchee and S.K. Ong, *J. Air Waste Manage. Assoc.* **42**, 1305–1312 (1992).
42. A.P. Schwab and M.K. Banks, in T.A. Anderson and J.R. Coats eds., *Bioremediation through Rhizosphere Technology*, ACS Symp. Ser. 563, American Chemical Society, Washington, D.C., 1994, pp. 132–141.
43. U.S. Pat. 5,494,574 (February 27, 1996). R. Unterman, B.R. Folsom, and A.P. Togna, (to Envirogen, Inc., Lawrenceville, N.J.).
44. M.A. Deshusses, *Curr. Opin. Biotechnol.* **8**, 335–339 (1997).
45. G. Leson and A.M. Winer, *J. Air Waste Manage. Assoc.* **41**, 1045–1054 (1991).
46. M. Arnold, A. Reittu, A. von Wright, P.J. Martikainen, and M.L. Suihko, *Appl. Microbiol. Biotechnol.* **48**, 738–744 (1997).
47. J.W. Barton, K.T. Klasson, L.J. Koran, Jr., and B.H. Davison, *Biotechnol. Prog.* **13**, 814–821 (1997).
48. Y. Matteau and B. Ramsay, *Biodegradation* **8**, 135–141 (1997).
49. Y.S. Oh and R. Bartha, *Appl. Environ. Microbiol.* **60**, 2717–2722 (1994).
50. A. Pol, H.J. Op den Camp, S.G. Mees, M.A. Kersten, and C. van der Drift, *Biodegradation* **5**, 105–112 (1994).
51. G. Leson and A.M. Winer, *J. Air Waste Manage. Assoc.* **41**, 1045–1054 (1991).
52. A. Korda, P. Santas, A. Tenente, and R. Santas, *Appl. Microbiol. Biotechnol.* **48**, 677–686 (1997).
53. G.D. Hopkins, L. Semprini, and P.L. McCarty, *Appl. Environ. Microbiol.* **59**, 2277–2285 (1993).
54. S. Jarvis, V.A. McFarland, and M.E. Honeycutt, *Ecotoxicol. Environ. Saf.* **39**, 131–135 (1998).
55. W.G. Palmer, J.R. Beaman, D.M. Walters, and D.A. Creasia, *J. Toxicol. Environ. Health* **51**, 97–108 (1997).
56. I. Deportes, J.L. Benoit-Guyod, and D. Zmirou, *Sci. Total Environ.* **172**, 197–122 (1995).
57. A.M. Fogarty and O.H. Tuovinen, *Microbiol. Rev.* **55**, 225–233 (1991).
58. F.C. Michel, Jr., C.A. Reddy, and L.J. Forney, *Appl. Environ. Microbiol.* **61**, 2566–2571 (1995).
59. M. Kastner, S. Lotter, J. Heerenklage, M. Breuer-Jammali, R. Stegmann, and B. Mahro, *Appl. Microbiol. Biotechnol.* **43**, 1128–1135 (1995).
60. L. Keller and P.H. Brunner, *Ecotoxicol. Environ. Saf.* **7**, 141–150 (1983).
61. P.B. Marsh, P.D. Millner, and J.M. Kla, *Mycopathologia* **69**, 67–81 (1979).
62. P.D. Millner, P.B. Marsh, R.B. Snowden, and J.F. Parr, *Appl. Environ. Microbiol.* **34**, 765–772 (1977).
63. C. Novotny, B.R. Vyas, P. Erbanova, A. Kubatova, and V. Sasek, *Folia Microbiol. (Praha)* **42**, 136–140 (1997).
64. T.S. Brodkorb and R.L. Legge, *Appl. Environ. Microbiol.* **58**, 3117–3121 (1992).
65. R.T. Lamar, B. Schoenike, A. Vanden Wymelenberg, P. Stewart, D.M. Dietrich, and D. Cullen, *Appl. Environ. Microbiol.* **61**, 2122–2126 (1995).

See also BIOFILMS, MICROBIAL; MIXED CULTURE.

BIOSORPTION, METALS

B. VOLESKY
S. SCHIEWER
McGill University
Montreal, Canada

KEY WORDS

Aqueous environment
Biomass utilization
Biosorption process
Detoxification
Metal biosorption
Metal concentration
Metal recovery
Metal removal
Metal uptake
Passive biosorption
Wastewater treatment

OUTLINE

BIOSORPTION OF METALS

It has been known for decades that some types of microbial biomass tend to accumulate metals rather effectively. This is the case even when the biomass is dead, that is, metabolically inactive, since it is the chemical composition that is responsible for sequestering metal ions and/or metal complexes from solutions. This fact obviates the necessity to maintain special growth-supporting conditions to make the biomass sequester metals. This metal-sorption phenomenon can find enormous applications wherever metals would need to be removed from solution and concentrated.

Rather than searching among thousands of microbial species for biomass with this particular metal-sequestering feature, feasibility dictates us to look first at biomass types that are readily available and in quantities to support the potential demand. Copious quantities of seaweeds, which could easily be collected, grow in the oceans and, indeed, some of them have been identified for their especially outstanding metal-sorbing properties. Other quantities of biomass are generated as by products of large-scale industrial fermentation processes. Often posing disposal problems, some of these readily available industrial-biomass types have been found to be excellent metal sorbers.

A whole new family of metal (bio)sorbing materials appears to be on the horizon, serving as a basis for the new technology of biosorption aimed specifically at metal removal and recovery from metal-bearing industrial solutions and/or effluents.

REMOVAL AND RECOVERY OF HEAVY METALS BY BIOSORPTION

Need for Heavy-Metal-Removal Processes

By far the largest demand for particularly cost-effective metal-removal technologies from dilute solutions is in ecoapplications. Metal ions in the environment are biomagnified in the food chain and accumulated in tissues. Cd, for example, has a half-life of 10–30 years in the human body (1). Therefore, their toxic effects are especially pronounced in animals of higher trophic levels, particularly humans. From the ecotoxicological point of view, the most dangerous metals are the so-called big three: Hg, Pb, and Cd. Following well-documented and publicized disasters and problems with Hg and Pb, the large-scale use of these metals by various industries has been either eliminated or very curtailed. Technological processes making use of these metals have recently been phased out, particularly in developed countries, and replaced by environmentally safer technologies. A major step in limiting the amount of lead discharge into the environment, for instance, has been made by the elimination of leaded gasoline.

On the other hand, the use of Cd has been steadily increasing in growing applications such as in electroplating, pigments, and batteries. Cd toxicity causes kidney damage, bone disease, and cancer (1). Other common heavy metals, such as Cr, Cu, Zn, Ni, and V are less toxic, but are among the most widely used metals (after Fe and A1) (1). Cu is mainly employed in the electrical industry (e.g., wires) and brass production. The major applications of Zn are galvanization and production of brass and other alloys (1). Mine tailings and effluents from nonferrous metal industry are the major discharge sources of different combinations of dangerous heavy metals.

Therefore, it would be beneficial to devise separate pretreatment facilities for the elimination of heavy metals from wastewater, ideally at the site of those plants that discharge significant amounts of heavy metal ions with

their wastewater. Such on-site treatment of waste streams contributing a load of heavy metals would be more efficient than treating large volumes of mixed wastewater in a general sewage plant. Additionally, from a policy viewpoint it seems reasonable that those who cause metal pollution should be responsible for its elimination. While industrial enterprises are presently obliged to meet certain discharge standards, drinking-water quality (Table 1) cannot be considered since such treatment would be prohibitively costly.

Natural-metal recycling in the environment makes metal recovery the only effective way of preventing heavy metals from poisoning the environment. Yet if a cost-effective way of completely eliminating heavy metals could be found, this goal could be attainable. Current research activity in the field attempts to evaluate whether biosorption may eventually provide such an effective and economical treatment-process alternative.

While biological treatment is reasonably effective in removing organic pollutants, heavy metals tend to accumulate in biological sludges (5), which are considered unsafe as fertilizer and require incineration. The accumulation of metals by the microbial biomass is a result and typical example of combined metabolism-dependent active bioaccumulation and passive biosorption of metals.

Possible Applications of Metal Biosorption

The term *biosorption* commonly refers to the passive binding of metal ions by biomass, which may even be dead. It must be distinguished from bioaccumulation, which is usually understood to be an active, metabolically mediated process occurring in living organisms.

It has been known for decades that different types of microbial biomass bind trace metal ions, achieving very high concentration factors. Black and Mitchell (6) reported a 1,000-fold increase of metal concentration in biomass of some brown marine algae as compared to the seawater. The biomass-metal-uptake phenomenon has been made use of in, for example, electron microscopy, to visualize cellular components. The focus in early studies of microbial-metal uptake has been almost exclusively on the nutritional and toxicological aspects of metal presence.

Recently, however, efforts have been made to develop the process of metal accumulation by biomass into a technique for the treatment of toxic metal-bearing industrial effluents, or for the recovery of precious elements from processing solutions, or even seawater. The process of gold sorption by *Sargassum* has been patented, and the immobilized biosorbents AlgaSORB™ and AMT-BIOCLAIM™ are being commercially developed (7).

Biosorption removal of toxic heavy metals is especially suited as a "polishing" water-treatment step because it is possible to reach drinking water quality of the treated water (initial metal concentrations, e.g., 1 to 100 mg/L, final concentration <0.01 to 0.1 mg/L), especially in packed-bed flow-through sorption column applications. To prevent unnecessarily rapid exhaustion of the sorption capacity when the metal concentrations in wastewater to be treated are high (>100 mg/L), it may be desirable to use a different pretreatment technique, such as precipitation or electrorecovery, for removal of the bulk of the metal content. However, generation of toxic sludges during pretreatment must be considered, since these represent another type of hazard and the eventual recovery of metals from them may not be feasible.

The metal-laden biosorbent can be regenerated, incinerated, or stored in landfills. The biosorption process basically serves to reduce the waste volume. Alternatively and preferably, regeneration of the biosorbent material or multiple (re)use is desirable to increase the process economy. It can be accomplished by metal desorption with, for example, acids or salt solutions. The resulting highly concentrated metal solution can be processed by other techniques such as precipitation or electro winning to remove or concentrate the metal, which could be the recovered and resold. The latter process, in particular, is aimed at recuperation of the metal. The overall achievement of the biosorption (complete adsorption + desorption cycle) process is to concentrate the metal solution, possibly by at least a factor of 100 or more. The flowchart of the biosorption process with in situ regeneration of the biosorbent is quite similar to the use of ion-exchange resins (Fig. 1).

Compared to conventional wastewater "polishing" techniques such as ion-exchange, activated carbon treatment,

Table 1. Drinking Water Standards and Discharge Limits of Heavy Metals

Metal	Drinking water standards WHO (2) (mg/L)	Discharge limits Canada (3) (mg/L)	Discharge limits United States (4) (mg/L)
Cd	0.003[a]	1.5	0.1
Cu	2[b]	1	3.3
Cr	0.05[b]	1	2.8
Ni	0.02[a]	2	4
Pb	0.01[a]	1.5	0.6
Zn	3[c]	2	2.6

[a]Maximum acceptable concentration for health reasons.
[b]Provisional value for health reasons.
[c]Limit for aesthetic or consumer oriented reasons.

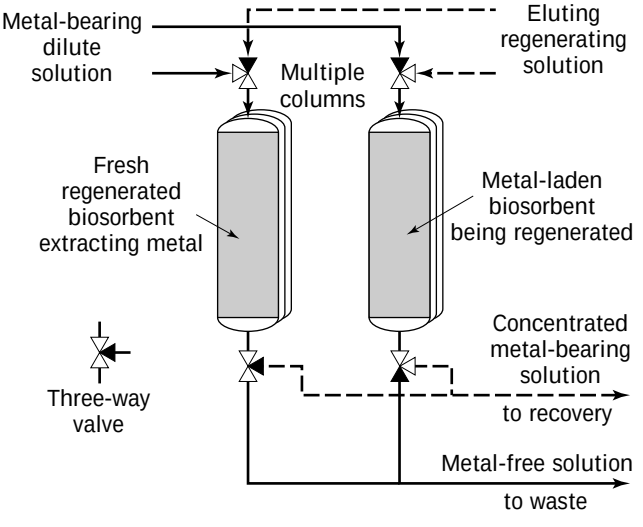


Figure 1. Schematic flowchart of a continuous-flow column biosorption process.

or membrane technologies (electrodialysis, reverse osmosis), the advantage of biosorption is not only that it can be operated under a broad range of conditions (pH, temperature), but especially that it may be economically attractive due to the cheap raw materials that can be used as biosorbents (7). These include waste products from other industries (e.g., fermentation by-products) or naturally abundant biomass (e.g., ubiquitous marine algae). Recent studies show that for certain types of seaweed biomass, for example, virtually no pretreatment may be necessary: Untreated *Sargassum* biomass was successfully applied in packed-bed column biosorption of Cu (8). This means that the sole costs of the seaweed-based biosorbent may be only those of collection and transport.

To evaluate the expected performance of biosorption for treating typical waste waters, it is desirable to reliably predict how much metal could be bound under different operating-process conditions. Computerized process simulations could effectively and economically provide this relevant information, which is essential for the engineering process design and scale-up.

Characteristics of Metal Biosorption

Biomass Types. A multitude of biomass types (9), comprising fungal biomass (10), bacterial biomass (11), algae (12–14), peat (15,16), and so on, have been studied for their biosorption of metals. It was observed that not only the species but also growth conditions such as culture medium, as well as physiological state and age of the organism, influence the biomass-sorption performance (17). In many cases, the sorption by dead material proved to be more effective than that by living organisms (18,19). Biomass of some marine brown algae has been found to be particularly effective in metal-ion binding (20,21), showing similar or better performance than industrial ion-exchange resins. Gold accumulation by the brown alga *Sargassum* can constitute up to 40% of the biomass dry weight (22). Further encouraging results with this sorbent material were obtained by Kuyucak and Volesky (22), Holan et al. (20), Leusch et al. (23), and Ramelow et al. (24).

The particular amount of metal bound in the biosorbent depends, however, not only on the chosen biosorbent but also on the type of metal ion, its concentration, and other physicochemical parameters of the solution (e.g., pH, ionic strength).

Temperature Effect. Simple physical sorption processes are generally exothermic (i.e., the equilibrium constant decreases with increasing temperature). Biosorption ion exchange among alkaline earth metals in binding to alginate was observed to be exothermic (25). However, when Cu was the replacement, the reaction exhibited a positive enthalpy change (i.e., it was endothermic—the equilibrium constant rose with temperature). The driving force in this case was attributed to a large positive entropy change, possibly caused by a larger ordering effect of Cu (than Ca) on the water molecules in the hydration sphere. This is illustrated by the following equation:

$$K = \exp\left(\frac{-\Delta G^0}{RT}\right) = \exp\left(\frac{-\Delta H^0 + T\Delta S^0}{RT}\right) = \exp\left(\frac{-\Delta H^0}{RT}\right)\exp\left(\frac{\Delta S^0}{R}\right) \quad (1)$$

A rise in K with increasing T means that $\Delta H^0 > 0$, that is, that the reaction is endothermic. But since a reaction can only proceed when $\Delta G^0 = \Delta H^0 - T\,\Delta S^0 < 0$, it follows that ΔS^0 must have been >0, that is, the system's entropy must have increased.

Similar results were obtained by Weppen and Hornburg (26) in calorimetric studies of the binding of divalent metal ions by potassium-saturated microbial biomass. The reactions were slightly endothermic and therefore driven by an increase in the entropy of the system. For most metals, the heat of reaction was constant, independent of the degree of site occupation. For Cu, however, the heat of reaction decreased with increasing degree of site occupation from 27 kJ/mol to 14 kJ/mol, which may indicate the involvement of different binding sites or the formation of different types of Cu complexes with the biomass. For other heavy metals, the heat of reaction was between about 7 and 11 kJ/mol, for light metals between about 2.5 and 6 kJ/mol (26).

Overall, however, the effect of temperature in biosorption is relatively small and perhaps only slightly more pronounced in the lower temperature range: the binding of Co by the brown alga *Ascophyllum nodosum* increased by 50–70% when the temperature was raised from 4 °C to 23 °C. With further temperature increase to 40 °C the binding increased only slightly, whereas temperatures of 60 °C or more caused a change in the texture of the sorbent and a loss in the sorption capacity. The distribution ratio (metal bound/metal in solution) for biosorption of Cd, Zn, Pb, Ni, and Cu to *Spirulina* algae increased by only about 20% when the temperature was raised from 4 °C to 55 °C. It is reasonable to expect that possible application of biosorption will be in a rather narrow temperature range (~5 to 50 °C) to avoid decomposition of biosorbent materials.

Presence of Anions (Ligands). Theoretically, the presence of ligands (at levels that do not cause precipitation) can lead to the following (27):

1. Formation of complexes that have a higher affinity to the sorbent than the free metal ions (i.e., an enhancement of sorption)
2. Formation of complexes that have a lower affinity to sorbent than the free metal ions (i.e., a reduction of sorption)
3. Interaction of anions with the biomass, changing the state of the active sites such that the binding is either enhanced or reduced

There are no indications of the third kind of interaction reported in the biosorption literature. Additionally, these are unlikely to occur in brown algal biomass since the probable binding sites (negatively charged carboxyl and sulfate groups) are not expected to interact with anions such as sulfate, nitrate, and so on.

It has been reported for sorption of Cu^{2+} and Ag^{+} to inorganic oxide surfaces that metal binding at low pH was enhanced in the presence of certain ligands. Since the ligand itself was able to bind to the oxide, and since a 1:1 stoichiometry of ligand and bound metal was noted, it was concluded that some metal complexes may bind more strongly than the respective free ions (28). Negatively charged metal–EDTA complexes were bound to oxide surfaces (probably releasing OH^{-}), with sorption decreasing with increasing pH (29).

Chen et al. (16) report that up to 50% of Cu binding by peat occurred as complexation of copper nitrate [$Cu(NO_3)_2$] with the binding sites. The remaining metal binding was due to ion exchange such that the overall charge neutrality was conserved. The data allow, however, another interpretation as well: Instead of up to half of the Cu being bound as $Cu(NO_3)_2$ and the other half as Cu^{2+}, up to 100% could have been bound as $Cu(NO_3)^{+}$. Glaus et al. (30) found that the binding of Co and uranyl (UO_2^{2+}) to humic substances could be modeled better when binding of metal–ligand (oxalate) complexes was included in the model.

In most cases of biosorption, however, metal binding tends to be reduced in the presence of ligands (exception: OH^{-}) (22,27,31). This means that the biomass apparently has less affinity for many metal–ligand complexes than for free hydrated metal ions. Therefore, the influence of ligands in solution can be understood as a competition with the biomass for binding of the metal ions.

This effect is, however, not very pronounced unless these anions show a strong complexing power for the metal ions of interest (e.g., EDTA) (24,27), which at only 1.5 times excess caused 100% inhibition of metal binding. The addition of 0.1% EDTA resulted in >90% reduction of binding for most cases of Cd, Cu, Zn, and Pb biosorption by seaweeds (also *Sargassum*) (24). A large excess (1,000 times) of the ligands glutamate, sulfate, phosphate, carbonate, or chloride resulted in the maximum reduction of, for example, Cd binding to *Rhizopus arrhizus* biomass of only 21% (27).

Since significant concentrations of strong complexing agents are rare in typical metal-bearing industrial effluents, the effect of the presence of ligands is of secondary importance.

Influence of pH. Of great importance in sorption is the pH value of the solution. It is commonly agreed that the sorption of metal cations (e.g., Cd, Cu, Zn, Pb, Ni, Mn, Al, Co) increases with increasing pH (14,20,24,32–35) as the metal ionic species become less stable in the solution. Only those metal ions that can occur as negatively charged complexes or that have a strong "b" character (i.e., tendency to form strong covalent bonds), such as Ag, Hg, or Au (e.g., as tetrachloroaureate), may show either a decrease in binding with increasing pH or may have no significant pH effect at all (24,31,33). There are three ways that the solution pH can influence metal biosorption, as described in the following.

First, the state of the active sites may be changed. When the binding groups are acidic, the availability of free sites depends on pH: At lower pH the active sites are protonated, therefore, competition between protons and metal ions for the sorption site occurs (36,37). At a low enough pH, virtually all sites become protonated and complete desorption of the bound metal ions is possible (38), which is why acid treatment is a method for metal elution and regeneration of the sorbent material. Decreasing the pH value by 2 units can in some cases result in a ~90% reduction of metal binding (14,24,31,32,38).

Second, extreme pH values, as they are employed in the regeneration (desorption) of the sorbent, may damage the structure of the biosorbent material. Microscopic observations have shown distorted cells; significant weight loss and decrease in the sorption capacity have been observed (39).

Third, the speciation of the metal in solution is pH dependent. Whereas metals in aqueous solutions occur as hydrated cations in solvation shells when the pH is low, hydroxides may form at a higher pH, especially for cations of high charge and small size (40–42). The formation of metal oxide and hydroxide complexes and precipitates is often called hydrolysis (i.e., decomposition or conversion by water) (43).

Adsorption depends not only on the attraction of the sorbate to the solid surface but also on its lyophobic behavior (44). This means that sorption increases with decreasing solubility. Since the solubility of many metal complexes in solution decreases with increasing pH, this provides an additional possible explanation as to why sorption increases with increasing pH. In the narrow pH range where the metal ions are hydrolyzed, sorption is especially enhanced, so that even a charge reversal of the surface to positive values can occur (44,45). Further possible explanations of increasing sorption with increasing pH are that hydrolyzed species have a lesser degree of hydration, that is, less energy is necessary for removal or reorientation of the hydrated water molecules upon binding. Furthermore, OH shows affinity for some sorbents (46).

With further increase of pH, the solubility of metal complexes decreases enough for precipitation to occur. Although precipitation may contribute to the overall removal of metals from solution (and therefore be desirable for metal-removal applications), it renders the study of the biosorptive binding more difficult. For scientific purposes it is therefore recommended to study biosorption at pH values where precipitation does not occur.

Presence of Other Cations. Other sorbable ions in the solution may compete with the metal ion of interest for sorption sites. The binding of this metal ion is then decreased. The amount of inhibition depends on the binding strength of the respective ions to the biomass.

For example, Mg and Na are effective in competing with Zn for binding sites of green algae [90% reduction of Zn binding at 200 mM NaCl or $Mg(NO_3)_2$], but they interfere less strongly with the binding of Pb or Cu [<50% reduction of Pb or Cu binding at 200 mM NaCl or $Mg(NO_3)_2$] (14).

The inhibition by alkali metals (K, Na) of heavy-metal-ion binding to microbial biomass was much less pronounced than the inhibition by heavy metals (Zn, Cu, Fe) of uranium or radium binding (19).

For *R. arrhizus* fungal biomass, the presence of uranium reduced the binding of Cd, Zn, and Ag more strongly

than vice versa; Ag in particular was a weak competitor, but nevertheless its binding was unaffected by the presence of Na (47).

In the sorption of Cd, Cu, and Zn on formaldehyde crosslinked biomass of the brown alga *A. nodosum*, the presence of each ion reduced the binding of each other ion, with Zn being the weakest binding, and Cu the most strongest binding, of the three cations (48).

From the different literature sources it can be generally concluded that the light metals (alkaline and alkaline earth metals) bind less strongly than the heavy-metal ions or radioactive elements. Therefore, the former do not strongly interfere with the binding of the latter. There are indications that among the heavy metals, Zn binds rather weakly and its binding is therefore more strongly affected by other ions.

In order to understand why some ions bind more strongly than others, it is useful to look at the properties of different metals, to be discussed later. Such knowledge can also serve to estimate how strongly an ion may bind that has not yet been investigated. This knowledge serves as a basis to better understand the metal-binding mechanism.

Mechanism of Biosorption

The effect of any of the previously mentioned influencing factors can be best quantitatively estimated when the mechanism of biosorption for different ions is known and when a mechanism-oriented mathematical model is used for predicting the effect of these factors.

Classification of Binding Mechanisms. The metal binding in biosorption has been attributed to a number of different sequestering mechanisms, such as complexation, coordination, chelation, adsorption, ion exchange, or microprecipitation (as metal or metal complex) (49). It is a recognized fact that a combination of several mechanisms, each functioning independently, can contribute to the overall metal uptake. In the studies of biosorption conducted so far, very little attention has been paid to examination of a well-defined metal uptake (by a specific mechanism) as opposed to the overall uptake where several types of sequestration may be taking place simultaneously. Systematic understanding of the metal-uptake mechanisms and their relationships may greatly clarify the otherwise confusing broad definition of biosorption mechanisms in the literature.

For example, some mechanisms can be submechanisms of other overall mechanisms. The claim of an ion-exchange metal-binding mechanism for a specific system may not necessarily be contradictory to reports of chelation being the relevant mechanism. The overall mechanism of ion exchange can be based on the submechanism of complexation. This means that one complexed ion exchanges against another one.

A systematic presentation of the relationships between different mechanisms is compiled in Figure 2. The classification of bond types (Fig. 2, bottom) as described by Myers (50) was used, as well as information on sorption processes from Westall (51) and Pagenkopf (44).

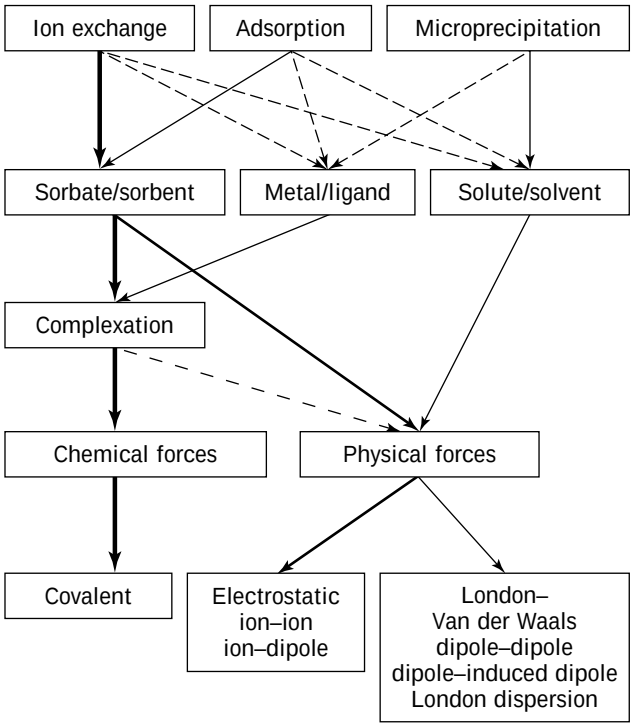


Figure 2. Metal biosorption mechanisms: bold lines, mechanisms probably important in biosorption by *Sargassum* biomass; dashed lines, biosorption binding relations of secondary importance.

In the context of this work, the term *sorption* refers to binding of a solute (usually a metal cation) to free sites that previously were not occupied by another cation. If the sites are initially occupied by another cation (regardless of the submechanism, i.e., whether covalent or electrostatic binding is involved), and if this second ion is released on the binding of the first ion, then the term *ion exchange* is used.

Microprecipitation is the deposition of electrically neutral material (metal or metal salt) at the surface of the biomass, and does not necessarily involve a bond between the biomass and the deposited layer. Microprecipitation may, however, be facilitated by initial binding of metal ions to reactive sites of the biomass, which serve as nucleation sites for further precipitation (52). This process is not limited to a monolayer (or saturation of sites): Cells can accumulate several times their dry weight in metal (53,54). Microprecipitation is based on interactions between the solute (dissolved solid) and the solvent, and occurs when the local solubility is exceeded.

Ion exchange and adsorption can be the result of three different interactions (which can act in combination): The main contribution for free metal ions (which are highly soluble in water) is usually attraction of the sorbate (metal ion) to the sorbent (biomass). In the case of binding of metal–ligand complexes, or in the case of ligand exchange (where the metal ion in solution is complexed with a ligand that has to be released when the metal ion is bound to the biosorbent [51]), interactions between ligands and metal cations must be taken into account. Additionally, hydrophobic (or, generally, lyophobic) expulsion (which is an interaction between sorbate and solution) may also play a

role (44,51), especially in the adsorption of organic dipoles and large organic ions because of their low affinity to the aqueous phase (46). Among inorganic ions, the less hydrated ones are more readily accumulated at the interface (46). The hydration effect is further discussed later.

Complexation plays an important role in both metal–ligand and sorbate–sorbent interactions (51). A complex (also referred to as a coordination compound) is a polyatomic molecule that consists of one or more (then called polynuclear complex) central atoms (usually metal cations) surrounded by ligands (other atoms or groups, usually of negative or neutral charge) that are attached to it. Complexes can be neutral, positively charged, or negatively charged. The number of coordinating atoms in the ligands that are directly attached to the central atom is called the coordination number, and can be larger than the valence of the central atom. (The most common coordination numbers are 4 and 6, but 2 and 8 also occur frequently.) If one ligand is attached to the central atom through two or more coordinating atoms, then the complex is called a chelate. The bonds between the central atom and the coordinated groups can be arising from principal or auxiliary valence forces (55,56).

Binding Forces. In general, the forces between atoms or molecules can be classified into chemical and physical ones. Chemical forces extend over very short distances (0.1–0.2 nm) (50). This type of bond is rather strong: Stumm and Morgan (46) report an energy of $\geqslant$40 kJ/mol, Myers (50) 150–900 kJ/mol, Smith (57) for chemisorption 20–400 kJ/mol and Pagenkopf (44) 40–400 kJ/mol. Covalent bonds are formed by merging electron clouds such that a nonionic molecule is formed. These bonds are directional (characteristic bond angles and lengths) and localized (50).

Physical forces can be subdivided into electrostatic and London–van der Waals forces (50). The energy of physisorption is reported as 2–20 kJ/mol and 20–40 kJ/mol by Smith (57) and Pagenkopf (44), respectively. In the resulting bonds, the electrons stay in their original systems. Electrostatic (or coulombic) forces between ions or between ions and dipoles extend over a long range and are the strongest among the physical bonds (50), with energies $\geqslant$40 kJ/mol (46). The interaction is repulsive for ion charges of the same sign and attractive for unlike charges. The magnitude of the force is proportional to the charge of each ion and inversely proportional to the square of the distance between the ions.

London–van der Waals forces can be divided into three categories: dipole–dipole interactions (creating orientational energy), dipole-induced dipole interactions, and the London dispersion force (50). The first two are closely related to coulombic forces, while the last is of quantum–mechanical nature and acts over a long range, of up to about 10 nm (50). The energy of the dispersion force (8–40 kJ/mol) is larger than the one of orientational or induced dipole (or polarization) (51) energy ($<$8 kJ/mol) (46). An example of untypical strong (almost ionic) dipole interactions is hydrogen bonding (which may even display some covalent bond character). It occurs between molecules in which H is bound to a very electronegative atom such as O (e.g., in H_2O) (58).

The mechanisms and forces that are probably (as pointed out later) the most important ones in the case of metal binding by, for example, brown algal biomass are indicated in bold type in Figure 2, that is, complexation by chemical covalent bonds based on ion exchange (for which the driving force is the affinity of the sorbate [metal ion] to the sorbent [biomass]) supported by physical electrostatic interactions between positively charged metal ions and negatively charged biomass.

Ion Exchange versus Sorption. There have been some indications that ion exchange plays an important role in metal-ion sorption by algal biomass. Kuyucak and Volesky (59) noted that the amount of ions (K, Na, Ca, Mg) released from the marine brown alga *A. nodosum* was much more pronounced in metal (Co) bearing than in metal-free solutions. A linear correlation between Ca release and Co uptake (2:3) was found when the biomass had been previously washed with CaCl and HCl. It was concluded that ion exchange was responsible for the sorption (59). If ion exchange was responsible, an exchange ratio of 1:1 should, however, be expected for two divalent ions such as Ca and Co. Since it can be expected that some sites would have been protonated due to the pretreatment with a CaCl and HCl mixture, and since some proton release was noticed, it would have been more appropriate to relate the Co binding to the sum of Ca and H released (in milliequivalent not mole). This would have led to an exchange ratio much closer to 1:1.

Crist et al. (60) included protons in the quantitative study of ions (Ca, Mg, H) released during metal (Cu) uptake by the freshwater alga *Vaucheria* and found that the amount of divalent metal ions taken up equaled the amount of other divalent ions plus protons released (in milliequivalent terms). Treen et al. (61) reported that in column application of *Rhizopus* fungal biosorbent 2 mol of protons were released for 1 mol of uranyl sequestered.

According to Chen et al. (16), the Cu binding to peat at low metal concentrations was predominantly ion exchange (with H, Ca, Mg), while at higher concentrations, binding of copper–nitrate complexes occurred. The latter mechanism was dominant above 6 mM of Cu.

It appears reasonable that in many cases ion exchange rather than sorption to free sites is the relevant overall mechanism for the binding of metal ions in biosorption. Since the overall charge of the biomass particle has to be neutral, any binding of one cation must be accompanied by either a stoichiometric release of other cations or by the binding of anions. Since the charge of algal biomass is negative (see "Surface Charge"), it appears unlikely that free anions will be bound by the biomass. If complexes in solution occur, double exchange of cations and anions or binding of the complex may take place. Generally, sorption of neutral (metal cation + anion) complexes would be possible. However, it is unlikely to be the case at the moderate pH and metal, as well as the ligand concentrations, relevant to the biosorption process.

Desorption/Ion Exchange of the Metal from Biosorbent. Due to the weak nature of metal-binding forces, it is possible to wash the bound metal from the sorbent in a small

volume of solution offering another cation (e.g., H^+, Ca^{2+}) in higher concentrations. From the process point of view, this enables the reuse of the biosorbent in multiple uptake/desorption cycles allowing for a greater process economy. Desorption can be achieved by acids (e.g., HCl, H_2SO_4) and/or salt solutions (e.g., $CaCl_2$) (38,39). The effectiveness of desorption depends in this case on the binding strength of the added cation to the biosorbent. A different approach to desorption is the addition of a complexing agent (e.g., EDTA, KSCN) (39) that binds free-metal ions and thereby lowers their concentration in solution to the extent that they are released from their binding sites on the solid. In this case the driving force is the complexation of the metal ion. The efficiency of the desorption depends on the binding strength of the desorbed ion to the complexing ligand. Nevertheless, a charge balance of the sorbent particle has to be maintained, that is, ion exchange still occurs. That means some cation (e.g., the one that was added along with the complexing ligand, often Na) has to replace the desorbed ion at the binding sites, even if its affinity toward the binding sites is lower.

Chemical (Covalent) versus Physical (Electrostatic) Binding. The proportion of protons among the released species varied with the metal taken up (25,62) and corresponded to the binding strength: More protons were released by the stronger binding ions. It could be concluded that in those cases where no protons were involved, as for most alkaline and alkaline earth metals, electrostatic attraction was the binding mechanism (63), whereas covalent binding occurred when protons were released (62).

It appears plausible that the number of protons released increases with the binding strength of the metal ion because it can more effectively compete with protons. Also, as a general rule, the force of electrostatic binding alone is weaker than binding with covalent contribution (especially since alkaline earth ions have similar radii to those of many divalent heavy metal ions, so that large differences in the binding strength are due to different amounts of covalent binding). However, it does not necessarily follow that proton release indicates covalent binding; part of the protons (depending on the relationship between pH and $p^{BH}K$) may also be bound by electrostatic attraction and therefore be easily replaceable by electrostatically bound ions.

Binding Sites. Biosorption has been attributed to different types of groups such as carboxylate, carbonyl, hydroxyl, amine, amide, imidazole, phosphate, thio, and thioether groups (33). Which binding mechanism applies in any specific case depends, of course, largely on the type of biosorbent used and on the types of binding sites it contains. Crist et al. proposed that, for example, carboxylate, sulfate, and amino groups may be responsible for metal binding by freshwater algae (62). Indications for carboxyl- and sulfate-group participation were found by titration of marine algal biomass (64). The chelating ability of polysaccharides from green freshwater alga *Chlorella* was connected to their content of uronic acids. Their carboxyl groups would be negatively charged and could bind metal ions (65). Tobin et al. (37) concluded that in *R. arrhizus* fungal biomass at pH 4, carboxylate and phosphate groups offered the primary metal-binding sites, supported by hydroxyl groups. Greene et al. (31) noticed a complete loss of available sulfhydryl groups (determined by polarimetric titration) after the binding of Au(III) to the freshwater alga *Chlorella*, thus giving evidence for the involvement of this group, but it could only account for a small amount of total gold binding. Cell modifications with succinic anhydride indicated the relevance of amine groups, since the sorption of tetrachloroaureate decreased by 50% after their modification (31).

The modification of carboxyl groups by esterification with acidic methanol resulted in a significant decrease in the binding of Cu and Al, but a slight increase in Au binding by five different algae (66). The decrease in sorption was more pronounced for Al, which as a hard metal is expected to bind preferably to hard sites, such as carboxyl groups, whereas Cu as a softer metal may bind to nitrogen- or sulfur-containing group instead. The binding of the negatively charged Au complex increases, probably due to a decrease in the negative surface charge.

For brown marine algae it is assumed that alginates play a key role in metal-ion binding (59). Fourest and Volesky (67) investigated the binding of H, Cd, and Pb by alginic acid and *Sargassum fluitans* biomass before and after modification of the carboxyl groups using acidic methanol or propylene oxide. A linear correlation between the binding capacity for Cd and the content of weak acidic groups (probably carboxylate) resulting from different degrees of active-site blocking was noted. The maximum metal-binding capacity of different algae was linearly correlated with their titer of weak acidic groups, with a stoichiometric relation between metal bound and acidic titer of 0.36 to 0.5 mol/mol. This means that about two acidic sites are necessary per bound metal ion (E. Fourest and B. Volesky, personal communication, 1997).

Surface Charge. The relevance of electrostatic attraction in biosorption depends on the types and amounts of sites present in the biomass and on whether they are ionized or occupied by protons or other ions. That, in turn, depends on the pH and on the pK_a of the respective group. Usually, a good agreement between the pK_a values of inorganic acids and their corresponding acidic groups in biomolecules exists, as evident from Table 2. Amine groups are positively charged in their protonated form and neutral

Table 2. K_a Values of Inorganic Acids and Acidic Groups in Biomolecules

Inorganic acid	pK_a	Acidic group in biomolecules	pK_a
Sulfuric acid[a]	1.9	Sulfate[b]	~1.5
Phosphoric acid[a]	2.1, 7.2	Phosphate[c]	2
Benzoic acid[a]	4.2	Carboxyl[c]	3–5
Acetic acid[a]	4.8	Carboxyl[c]	3–5
Ammonium[a]	9.2	Amine[c]	8–10
Phenol[a]	10.0	Phenolic OH[c]	10

[a]Bower and Bates (68).
[b]Crist et al. (64).
[c]Buffle (69).

when deprotonated. Carboxyl, sulfate, and phosphate groups are neutral when protonated and negatively charged when deprotonated.

If the freshwater alga *Chlorella* could be taken as an example, its isolelectric point (zero overall charge) is at pH 3 (33). At lower pH, the positive charge (possibly of amine groups) leads to an attraction of anions; at higher pH, the negative charge (probably of carboxyl groups) attracts cations and repels anions (33). Biomass of the marine brown alga *Ascophyllum nodosum* bears a negative charge whose magnitude sharply increases from pH 3 upward. It is probable that this surface charge originated from deprotonated carboxyl groups of the alginate in the cell wall (34).

Location of Bound Metal. Knowledge pertaining to the location of metal uptake in the cell can provide some indication about the relevance of binding groups if it is correlated with information about which biomolecules occur in the respective parts of the cell. Electron micrographs of bacterial (70,71), fungal (72,73), and algal (74) metal-bearing cells in the literature indicate that the cell wall plays an important role in metal-ion biosorption. While the basic biomass metal-sequestration process is usually very fast, some time-lapse electron micrography indicates slower metal penetration into cell interiors and/or building up of external metal deposits (18,31,59,63).

For a better understanding of the relevant mechanisms in metal binding by microbial biomass, more detailed and accurate knowledge about the biomolecules and binding sites offered by different types of cells is necessary.

PROPERTIES OF METAL IONS IN BIOSORPTION

Biosorption Affinity Sequence Trends

Studying metal-ion binding to alginic acid from *Laminaria digitata* Haug (75) reported that the amount of protons released into solution decreased in the order Pb > Cu > Cd > Ba > Sr > Ca > Co > Ni > Zn > Mn > Mg. This was explained by a better ability of strongly binding ions to compete with protons for organic binding sites. The affinity sequence for metal-ion binding to alginate from *L. digitata* followed a similar trend Cu > Ba > Ca > Co (76), and for binding of alkaline earth metals to polymannurate and polygulurate the binding strength decreased in the order Ba > Sr > Ca > Mg (25). According to Haug and Smidsrod, the preferential binding of the heavier ions could have been due to stereochemical effects, as larger ions might better fit a binding site with two distant active groups. The amount of proton displacement by metal ions decreased in the order Pb > Cu > Ba > Sr > Ca > Cd > Zn > Ni > Mg > Co, Mn for polyguluronic acid, and in the order Pb > Cu > Cd > Ba > Ni > Sr > Ca > Mg > Zn, Co, Mn for mannuronic acid (25). Again, the trend was concurrent with the one in binding strength.

Schweiger (77) found a different order of complex stability for binding to alginate as being Ba > Cd > Cu > Sr > Ni > Ca > Zn > Co > Mn > Mg, but still the binding of alkaline earth metals increased with their weight. For the elements Cu, Ni, Co, and Mn of the third period, an increase of binding strength with increasing atomic number was noticed. In binding to fucoidan, the affinity sequence Pb > Ba > Cd > Sr > Cu > Fe > Co > Zn > Mg > Cr > Ni > Hg > Ca was reported (78).

For *R. arrhizus* biomass, the maximum metal-binding capacity decreased in the order $UO^{2+} > Cr^{3+} > Pb \sim Ag^{+} > Ba > La^{3+} > Zn > Hg > Cd > Cu > Mn$ (37). The metal-binding capacity increased linearly with the crystal ionic radius, and no correlation to the charge was noticed, except that the binding of alkaline metal ions was negligible (37).

The observations of higher binding (both capacity and binding constant) with increasing ionic radii were interpreted as a consequence of the binding-group conformation: Assuming that binding groups are distant from each other, large ions could not only be more easily dehydrated, but could also bind to several groups simultaneously (79). This could explain the binding strength; however, it does not explain the higher binding capacity. To the contrary, if several groups are involved in the binding of one large ion, the binding capacity of small ions, each binding to one site, could be higher.

Avery and Tobin (80) found that the amount of proton displacement decreased in the order Cu > Cd > Zn > Mn > Sr. It was assumed by these authors that proton release occurred only when the metal was covalently bound, and that the release of Ca and Mg indicated ionic binding. The ratio of (Mg + Ca) released per H^{+} released was used as an index for the relative amount of ionic/covalent binding (80). It was, however, not considered that, for example, Cu might bind covalently to a site from which Ca, which had been bound electrostatically, was released. This index should therefore not be used in a quantitative way.

Evans (81) states that the affinity sequence in alkali-metal binding to soil (Cs > Rb > K > Na > Li) reflects an increase in the hydrated radii of these ions; for divalent ions the binding strength to soil organic matter at pH 5.8 decreased in the order Hg = Pb = Cu > Cd > Zn > Ni > Co > Mn.

For freshwater algae, Crist et al. observed a trend in binding strength Cu > Sr > Zn > Mg > Na being similar to the one for proton replacement Cu > Zn > Mg > Sr > Na, and concluded that a trend from covalent to ionic bonding was responsible (62). As pointed out earlier, however, electrostatic effects could have contributed to proton binding on the one hand and, on the other hand, covalent binding to sites previously occupied by a light metal could be possible.

For binding of light metal ions to freshwater algae, the affinity sequence was Na > K > Li > Cs and Mg > Ca > Ba > Sr (63). It was concluded that if the affinity between biomass and metals was high, the binding should increase with increasing charge density, that is, with decreasing radius. If binding was weak, it would be controlled by dehydration energy and the larger ions would be preferably bound because they could be dehydrated more easily. The observed sequence, which does not show a clear preference for large or small ions, was interpreted as resulting from these opposite trends, indicating medium to strong bonds (63). It is questionable, however, whether the use of crystal ionic radii for characterizing the ionic binding of the hydrated ion is appropriate.

Leusch et al. determined the affinity sequence for untreated *Sargassum fluitans* as Pb > Cu > Mi ~ Cd > Zn

(82). Fourest and Volesky (67) observed that the binding capacity for divalent metal binding to *Sargassum* decreased in the order Pb > Cu > Cd > Zn > Ca, and related this to decreasing electronegativity. However, the binding strength, rather than the maximum binding capacity, should be influenced by the electronegativity. Possibly, the maximum binding was not reached under the conditions employed by these authors.

Conclusion. Generally, these binding-strength trends, except for the work of Crist et al. (63), follow the pattern for monovalent ions—Cs > Rb > K > Na > Li—and for divalent ions—Ba > Sr > Ca > Mg; Cu > Ni > Co > Mn; Pb > Cd > Zn. While the preceding expresses basic trends, no fixed relation between these series can be established. Whether, for example, Cu or Cd happens to be more strongly bound depends on the sorbent material.

Factors Influencing the Binding Strength

There are three factors that can determine the affinity of a metal ion (because they contribute to the free energy of binding of hydrated metal ions to biosorbent binding groups).

1. *Hydration effects:* A change in the orientation of the hydrated water molecules occurs in conjunction with electrostatic binding (69), but in such "outer-sphere" complexes both partners retain their hydration spheres (69). Complete or partial dehydration takes place for "inner-sphere" complexes resulting from covalent binding (69).

   Both types of changes in the hydration state can occur more easily if the hydrated water molecules are not held strongly by the metal ion. Williams and Hale (83) showed that a good correlation (Born equation) between the ion hydration energy and the term (z^2/r) exists. For estimating the strength of hydration, the radius of the nonhydrated cation should be used, of course. This is usually taken as the ionic crystal radius (listed in Table 3).

   Taking into account the hydration sphere, the effective hydrated radius is larger than the crystal-ion radius. In general, the larger the hydration energy of an ion (i.e., the smaller the crystal radius) the larger the effective hydrated radius (58). Consequently, the hydrated radii follow a reverse trend as the ionic crystal radii, both for the groups of alkali and alkaline earth ions (see Table 3). It is difficult to quantify the radius of the hydrated ion, but some values have been compiled (84,85).

   According to Jain (86), hydration effects can be dominating when the electric field of the sorbent is weak (i.e., weak binding). In this case, larger ions (comparing crystal radii of ions of the same charge) that exert a smaller ordering influence on their hydrated water molecules (i.e., that are less strongly hydrated) are preferably accumulated at the interface.
2. *Ionic (electrostatic) binding:* If the binding groups are negatively charged, they can attract metal cations. The interaction is stronger the higher the charge density of both biosorbent and metal ion. If a strong electric field is present, electrostatic effects may become the dominant factor such that small ions that have a higher charge density are bound more strongly (86). In the past, ionic-crystal radii have been used in estimation of the ionic binding strength, as well as in aqueous solution (69,87). However, since the cation retains its hydrated water molecules, the hydrated cation radius is more characteristic for electrostatic attraction (as opposed to hydration). According to Marcus and Kertes (85), the selectivity increases with increasing charge and decreasing hydrated radius.
3. *Covalent binding:* This type of bond involves sharing of the electrons. Therefore, the more similar the electronegativities of the metal ion and the coordinating atom of the ligand, the higher the covalent character of the bond (88). The selectivity (or binding strength) increases with increasing polarizability of the ion (i.e., softness) (85).

Conclusion. Overall, the binding strength increases with increasing ionic radius and decreasing charge if binding is weak and largely due to hydration effects, decreasing hydrated radius and increasing charge if binding is intermediately strong and due to electrostatic effects, and decreasing electronegativity difference if binding is strong and covalent.

Since the first two criteria are almost opposed to each other, the affinity sequence can be reversed with changing electric field strength of the sorbent (86), except that the ionic and hydrated radii can exhibit opposite trends. Therefore, it is possible that for the alkaline and alkaline earth metals, for example, the affinity sequences Cs > Rb > K > Na > Li and Ba > Sr > Ca > Mg > Be observed in industrial ion exchangers are due both to decreasing crystal and increasing ionic radii (85).

The Concept of Hard and Soft Ions

In order to classify the behavior of metal ions, a division into so-called hard and soft ions was introduced by Pearson (89). Hard ions are characterized by predominantly electrostatic bonds where the entropy gain in the change of the orientation of the hydrated water molecules makes an important contribution to the free energy change ΔG (69). Hard ions have a weak polarizing power and retain their electrons strongly. Williams used the charge density (z^2/r) (with z being the charge and r the cation radius) as a measure for the strength of ionic binding (90,91). Table 3 shows that the ionic binding strength of most divalent metal ions is similar and that it is higher than the one for monovalent ions. This is also illustrated in Figures 3 and 4.

The characteristics of hard ions resemble the ones of class "a" cations in the "a"–"b" categorization of metal ions. Type "a" ions form their strongest bonds with the first element of each group, "b" ions with the second or higher (92). Type "a" ions tend to form strong bonds with highly electronegative donors that are difficult to polarize (i.e., hard) (93). The difference between the electronegativity (x) of the

Table 3. Parameters Characterizing Binding Strength

Metal	z (−)	r_{cryst}[a] (Å)	r_{hyd}[b] (Å)	x[c] (−)	z^2/r_{cryst}[d] (1/Å)	z^2/r_{hyd}[e] (1/Å)	"b" type[f] (Å)	Δx[g] (−)	*ionic* total[h] (−)	χ[i] (1/Å)
H	1	—	2.82	2.2	—	0.35	—	1.3	0.34	1.03
Li	1	0.76	3.82	1.0	1.32	0.26	1.61	2.5	0.79	0.33
Na	1	1.02	3.58	0.9	0.98	0.28	1.51	2.6	0.82	0.34
K	1	1.51	3.31	0.8	0.66	0.30	1.51	2.7	0.84	0.36
Rb	1	1.61	3.29	0.8	0.62	0.30	1.57	2.7	0.84	0.36
Cs	1	1.74	3.29	0.7	0.57	0.30	1.27	2.8	0.86	0.35
Mg	2	0.72	4.28	1.2	5.56	0.93	2.26	2.3	0.73	1.27
Ca	2	1.00	4.12	1.0	4.00	0.97	1.85	2.5	0.79	1.23
Sr	2	1.26	4.12	1.0	3.17	0.97	2.11	2.5	0.79	1.23
Ba	2	1.42	4.04	0.9	2.82	0.99	1.84	2.6	0.82	1.21
Mn	2	0.67	4.38	1.5	5.97	0.91	3.42	2.0	0.63	1.44
Co	2	0.65	4.23	1.8	6.15	0.95	4.86	1.7	0.51	1.84
Ni	2	0.69	4.04	1.8	5.80	0.99	4.99	1.7	0.51	1.92
Cu	2	0.73	4.19	2.0	5.48	0.95	6.32	1.5	0.43	2.22
Zn	2	0.74	4.30	1.6	5.41	0.93	4.07	1.9	0.59	1.56
Cd	2	0.95	4.26	1.7	4.21	0.94	5.20	1.8	0.56	1.69
Pb	2	1.19	4.01	1.8	3.36	1.00	6.61	1.7	0.51	1.94
Al	3	0.54	4.75	1.5	16.67	1.89	3.60	2.0	0.63	3.00
Ag	1	1.15	3.14	1.9	0.87	0.29	7.22	1.6	0.47	0.62

[a]Shannon crystal radii (95).
[b]Nightingale hydrated ion radii (84).
[c]Pauling electronegativity (88).
[d]Parameter for hydration strength (91).
[e]Parameter for ionic bonding strength (91).
[f]Parameter for covalent bond ("b") character x^2 (r_{cryst} + 0.85) (96).
[g]Parameter for ionic bond character (88).
[h]Fraction of ionic bond character 1-exp($-\Delta x^2/4$) (94).
[i]Parameter for total binding strength $\xi = z^2/r_{hyd}/1\text{-exp}(-\Delta x^2/4)$.

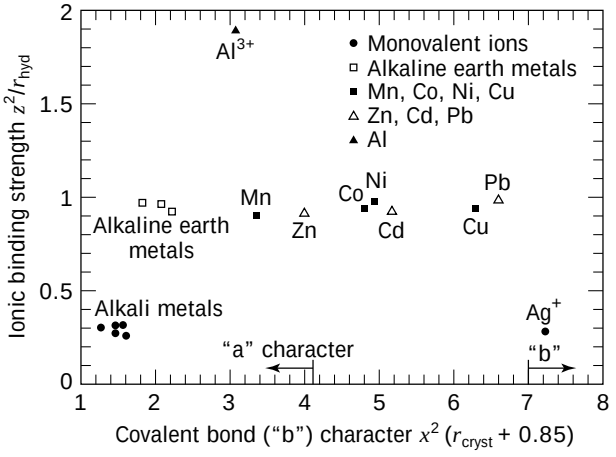


Figure 3. Ionic- and covalent-bond character for different metals.

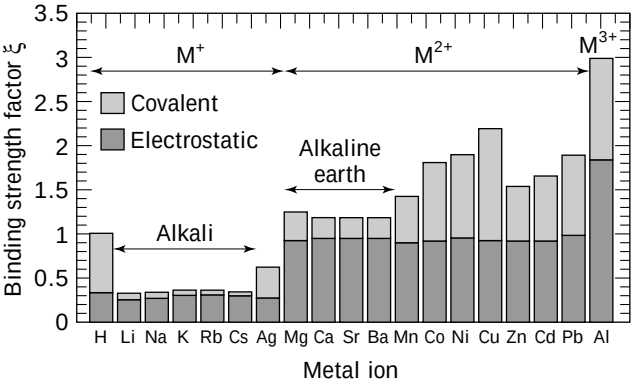


Figure 4. Binding-strength parameters for different metals.

metal ion and that of oxygen has been used as a measure for "a" (i.e., ionic bond) character, which increases with $\Delta x^2 = ({}^{M}x - {}^{0}x)^2$ (69). According to Pauling (94), the fraction of ionic bond character is $1 - \exp(-\Delta x^2/2)$, that is, equal contributions of covalent and ionic binding occur when $\Delta x = 1.7$. As seen in Table 3, the ionic bond character is larger than 70% for alkaline (earth) metals and about 50% for most heavy metals. Dean (88) describes a similar criterion: The ionic bond character increases with Δx and 100% ionic bonding occurs for $\Delta x > 2$ (but according to Pauling's criterion the ionic-bond character at $\Delta x = 2$ is 63%, not 100%). Table 3 shows that the Δx criterion is fulfilled for alkaline and alkaline earth ions.

Soft ions participate in more covalent bonding where the free energy is mostly enthalpic (69). The characteristics of soft ions are opposite those of hard ions, resulting in a reversed affinity sequence. The affinity sequence of soft, or class "b", ions for ligands is I > Br > Cl > F and S > N > O (97). A combined affinity sequence was described by Buffle: S > I > Br > Cl 39 N > O > F (69). Therefore, the difference between complexation constants for F and Cl, respectively, has been introduced by Turner et al. (87) to characterize softness, which increases with decreasing ΔK

$= \log({}^{F}{}^{M}K) - \log({}^{Cl}{}^{M}K)$. Type "a" ions are characterized by $\Delta K < 2$ and type "b" ions by $\Delta K > 2$; among the border elements, tendencies for "a" or "b" character are present for $\Delta K <$ or > 0, respectively (87). Typical hard, or class "a," ions are the alkaline and alkaline earth metals, transition elements up to class VI, and, for example, Zn (93). As opposed to the original classification of Pearson, who considered Cd a soft ion and H a typical hard ion, these two are now regarded as borderline ions (97). Table 3 and Figure 3 confirm that light metals and Mn can be classified as hard, and that most of the heavy metals may be listed as intermediate, except for Zn, which is just on the threshold to being hard, and Ag, which is soft.

For any donor (and especially for soft ones) the binding strength increases with cation softness (69). Soft cations form more stable complexes with soft donors (S, P, or As), while hard cations prefer hard donors (F or O) (89). N donors have intermediate characteristics. For O or N donor atoms, the Irving–Williams order of complex stability for divalent ions in the first series is Mn < Fe < Co < Ni < Cu > Zn (93). This can be explained by the electronegativity differences with respect to the donor atoms O or N.

Examples of hard ligands are carbonate, phosphate, sulfate, or hydroxide ions (1,94) as well as sulfate, carboxylate, alcohol, or hydroxyl groups (97). Soft ligands include sulfhydryl, thioether, and amino groups (97).

The relevant parameters for different metals are listed in Table 3. The hardness parameter (z^2/r_{hyd}) for different metal ions is plotted over the softness parameter $X^2(r_{cryst} + 0.85$ Å) in Figure 4. Since the parameter (z^2/r_{hyd}) expresses the ionic, binding strength, and the parameter $1 - \exp(-\Delta x^2/2)$ is supposed to characterize the relative contribution of ionic bonding to the total binding strength, one can obtain an indicator for the total binding strength by dividing (z^2/r_{hyd}) by $[1 - \exp(-\Delta x^2/2)]$. This parameter is called ξ; it is listed in Table 3 and plotted in Figure 4 as the total height of the bar. The lower part represents the ionic binding strength; the upper part represents the covalent binding strength. The graph should, however, be used only for illustrative purposes, rather than in a quantitative way, because the exact magnitude of the contribution of ionic bonding is disputable. According to the Δx criterion, for example, there is no covalent bonding for the light metals, while this plot, using Pauling's criterion, shows ~20% covalent binding.

The binding trends summarized in the section "Need for Heavy-Metal-Removal Processes" can now be explained: The tendency of stronger binding for the heavier elements in both the alkaline and alkaline earth metals can be explained as being due to lower charge density (z^2/r_{cryst}) (i.e., weaker hydration) of the ions themselves and higher charge density (z^2/r_{hyd}) of the hydrated ions (i.e., stronger electrostatic bonds) in the heavier elements of each series (see Table 3). Since these ions are hard, there is little contribution of covalent bonding; therefore, the electronegativity differences between them are irrelevant. The trend among the elements in the second period (Cu > Ni > Co > Mn) and the trend Pb > Cd > Zn are probably mainly due to a decrease of the covalent binding [characterized by $x^2(r_{cryst} + 0.85)$], but a decrease in the electrostatic binding (z^2/r_{hyd}) (except for the relation between Ni and Cu) and increasing hydration strength (z^2/r_{cryst}) (except for Mn) may also contribute.

Since the binding trends in biosorption by algal biomass or its constituent molecules indicate that there is significant electrostatic and covalent binding, the latest biosorption models described at the end of the next section ("Modeling") attempt to incorporate these stronger interactions and neglected hydration effects, which are weaker.

MODELING

In order to evaluate the feasibility and effectiveness of biosorption in wastewater treatment, it is essential to make predictions of the sorption performance (e.g., for facilitating process design). Therefore it is necessary to develop appropriate mathematical models of biosorption. Modeling the biosorption-binding equilibrium is a prerequisite for all further work involving batch kinetic studies and, in particular, for the continuous-flow sorption-column applications that represent the most effective configuration of the sorption-based process.

The amount of metal M (or in general, sorbate) bound per mass of sorbent is called the binding (or uptake), ${}^{M}q$. The binding is not only dependent on the sorbent material but also on the equilibrium concentration $[M]$ of the sorbate in solution and on other parameters, such as pH and equilibrium concentration of other ions in solution. The relationship between the equilibrium binding and the concentration of an ion (at constant temperature) is depicted in an isotherm plot of ${}^{M}{}_{q}$ versus [M]. With increasing metal concentration in solution its binding increases from zero to the maximum (Figure 5). It is desirable for a sorbent to possess a high sorption capacity (maximum binding obtained at high equilibrium concentrations) and a high affinity for the sorbate species, which is reflected in a steep slope of the isotherm curve at low equilibrium concentrations.

Langmuir and Freundlich Isotherms

Most mathematical biosorption models used in the literature describe simple Langmuir (13,14,20,59,63,64) or Freundlich (16,35,98) sorption isotherms, where the metal-ion binding is determined as a function of the equilibrium concentration of that metal ion in the solution, without reference to pH or other ions in the same solution system.

The Freundlich isotherm (99) is originally of empirical nature, but has later been interpreted as sorption to sites with an affinity distribution; that is, it is assumed that the stronger binding sites are occupied first and that the binding strength decreases with increasing degree of site occupation. More specifically, the Freundlich isotherm is obtained when a log-normal affinity distribution [i.e., a normal distribution of log (K)] is assumed (57p.319;100). The Freundlich isotherm is defined by the following expression:

$$ {}^{M}q = k[\mathrm{M}]^{1/n} \qquad (2) $$

where k and n are empirically determined constants, with

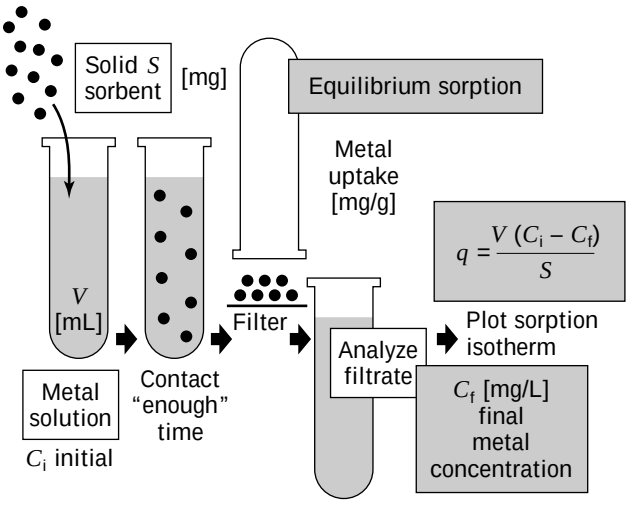


Figure 5. Experimental methodology for deriving the biosorption isotherm.

k being related to the maximum binding capacity, and n related to the affinity or binding strength (19,101).

The Langmuir isotherm (102) is based on considering sorption as a chemical phenomenon. It is assumed that the forces exerted by chemically unsaturated surface atoms (total number of binding sites B) do not extend further than the diameter of one sorbed molecule, and that therefore sorption is restricted to a monolayer. In the simplest case the following assumptions are made.

- All sorption sites are uniform (i.e., constant heat of adsorption).
- There is only one sorbate.
- One sorbate molecule reacts with one active site.
- There is no interaction between sorbed species.

The rate of adsorption is proportional to the rate constant of the forward reaction ${}^{BM}k_f$, the number of free sites ($B = {}^tB - {}^Mq$), and the number of sorbate molecules hitting the surface per unit time (i.e., proportional to their concentration [M]). Under the same conditions, the rate of desorption is proportional to the rate constant of the backward reaction ${}^{BM}k_b$ and the number of occupied sites ${}^Mq$. The equilibrium is attained when the rate of adsorption equals the rate of desorption:

$$ {}^{BM}k_f B[M] = {}^{BM}k_b {}^Mq \quad \text{(mol/s)} \tag{3} $$

With the equilibrium constant ${}^{BM}K$, indicating the affinity for the sorbate, defined as the ratio of adsorption and desorption rates ${}^{BM}K = {}^{BM}k_f / {}^{BM}k_b$, it follows:

$$ {}^{BM}q = \frac{{}^tB \cdot {}^{BM}K[M]}{1 + {}^{BM}K[M]} \quad \text{(mmol/g)} \tag{4} $$

Both Langmuir and Freundlich isotherms have been used to describe the equilibrium of metal biosorption. While it is possible to obtain good agreement between predictions of these models and the experimentally determined metal binding, their simple hyperbole-based equations have significant limitations: They cannot describe a more complex sorption behavior often reflected by nonhyperbolic relationships and they cannot predict the effect of pH or of other ions in the solution, among other shortcomings.

pH Effect. Although the crucial role of protons in biosorption is generally known, it is usually neglected in the mathematical description of the process. The complex nature of biosorbent materials that may contain a multiplicity of binding groups makes the modeling of the pH effect in biosorption difficult. As a consequence, it has been common practice to determine a separate isotherm and a new set of parameters, for each pH value (15,103,104) or for each initial biomass saturation state (e.g., protonated or loaded with ions from the seawater). This has been necessary because the most frequently used Langmuir or Freundlich sorption models do not take into account the fact that metal-ion biosorption is largely an ion-exchange phenomenon. They do not allow the prediction of the remaining binding of those ions (e.g., protons or sodium) that were initially loaded onto the biosorbent, and neither do they incorporate the concentration of the exchanged species (e.g., protons) as a parameter, normally not allowing reliable calculated predictions of the biosorbent performance.

Some attempt has been made, assuming a simplified linear relationship between the metal uptake and pH difference, to empirically model the biosorptive binding of metal ions by a Freundlich isotherm, which was modified by a factor proportional to the pH difference (105). This approach could not represent the change of proton binding in the presence of metal ions. Though protonation reactions for two acidic sites (alga) have been considered (106), the model simplification by lumping pH and proton- and metal-binding constants into one parameter for each resulted in a standard two-site Langmuir isotherm. A rare example of a Langmuir model including protons and one or two binding sites (104,107) assumed a 1:1 stoichiometry of metal binding and proton release, although it was recognized that metal binding with a 1:2 stoichiometry probably occurred. A modified Langmuir model for competitive Cu and H binding to two sites on alginate (108) required

iterative calculations of each cation uptake as a function of the binding of the other one. Another two-site model (109) based on a 2:1 proton–metal exchange constant allows direct calculation of the metal binding as a function of the concentration ratio $[H]^2/[M]$, but it does not account for free sites and it cannot easily be extended to include more species.

Competing Ions in Solution. Often, the prediction of metal-sorption equilibria is complicated by the presence of several sorbed ions, requiring the use of multicomponent isotherm equations. This is especially relevant when the sorption of one ion influences the binding of another when different ions compete for the same binding site. A variety of multicomponent sorption isotherm models is presented in a review by Yu and Neretnieks (110).

The use of multicomponent Langmuir isotherm models was demonstrated in describing the equilibrium metal biosorption (48,111,112). However, biosorption was not regarded as an ion-exchange process; the ions with which the biomass had originally been saturated were not considered.

Instead of using multicomponent isotherm models, more complex systems involving different cations and ligands can be modeled by incorporating binding constants for metal-ion–biomass complexes into a computer program for the calculation of the chemical equilibrium. Reviews of different computer programs for calculating the equilibrium speciation have been presented (44,113). These programs take acid/base reactions, coordination, and dissolution/precipitation into account. Their use is advantageous if complexation in solution is expected or if final concentrations are to be predicted from the initial state (27,47).

Ion-Exchange Constants

There have been indications that ion exchange is an important biosorption mechanism. However, this does not preclude an adequate use of simple sorption isotherms (16,59,64) instead of appropriate ion-exchange models with equilibrium constants taking into account the reversibility of the ion-exchange reactions: The ion that is displaced into the solution can compete with the sorbed metal ion for the sites (25,62,76,114).

The Langmuir equation and the ion-exchange constant for the binding of a metal ion M (for simplicity here a monovalent ion) replacing a proton H on a complexing site B are related as follows:

Langmuir: $B + M \rightleftharpoons BM$

$$^{BM}K^* = \frac{BM}{B[M]}\,[B]_t = [B] + [BM] \qquad (5)$$

Ion exchange: $BH + M \rightleftharpoons BM + H$

$$^{BM}K = \frac{BM[H]}{BH[M]}\,[B]_t = [BH] + [BM] \qquad (6)$$

therefore:

$$^{BM}K^* = {}^{BM}K/[H] \qquad (7)$$

The difference between the two approaches is that the first one assumes that all sites are initially free and does not consider any reverse reaction of a displaced ion, in this case a proton, with the site. The second model assumes that all sites to which metal ions are sorbed are initially occupied, that is, the number of free sites stays constant. Crist et al. (114) compared the fit of the Langmuir sorption isotherm model and the one using ion-exchange constants. The differences between the two models were especially pronounced at low metal-ion concentrations because the reverse reaction involving the displaced ion is most noticeable if the concentration of the bound cation is low.

The ion-exchange approach is probably somewhat closer to reality than the simple Langmuir model, but it is not completely satisfying either. Although the constant number of free sites [in other words, the number of sites occupied by the bound ion is equal to the number of sites from which the exchanged ion is released so that $^tB = (B^1M) + (B^2M)$ can be assumed] may be a reasonable working assumption for a constant pH system, it may not hold for systems with changing pH values. According to Stumm and Morgan (46), the cation-exchange capacity increases with increasing pH above the isoelectric point. Therefore, one cannot simply model the competitive binding of metals and protons by using only a metal–proton-ion-exchange constant. It is necessary to include at least one reaction where a cation reacts with a free site.

In the modeling of industrial ion exchangers, protons have been considered (115,116). However, these synthetic resins are chemically rather simple compounds that mostly contain only one active group. Biosorbents, on the other hand, can be heterogeneous and contain different acidic groups active in metal ion binding. The $p^{BH}K$ value, the quantity of binding sites, and the metal-ion binding constant must be known for each of these sites.

Models Including Ionic Strength Effects

Ubiquitous ions such as sodium may interfere with heavy metal biosorption through the effect of ionic strength. In biosorption, no modeling of the effects of ionic strength and electrostatic attraction has been done. Several models of ionic-strength effects, however, have been published concerning metal binding to humic and fulvic acids. The physical characteristics of these molecules, which are degradation products of organic matter, are quite different from microbial biosorbents. Humic and fulvic acids are small water-soluble molecules. However, since these molecules contain carboxyl groups as one of their main functional groups they may in some aspects be comparable in particular to those biosorbent materials based on seaweeds.

Buffle (117) systematically described possible models for complexation equilibria, including site protonation, multiple sites, and secondary effects such as the influence of the electric field in polyelectrolytes on the local concentration. The simplest approach to modeling the influence of Na in the biosorption system consists of the introduction of a binding constant for Na, as done by Westall et al. (118) for the binding of H, Co, and Na to humic acids. This approach enables derivation of a multicomponent sorption isotherm model so that the binding of each species can be directly calculated. This was done for biosorbents only re-

cently in the work of Schiewer and Volesky (119–121). Although the Na binding constant can easily be incorporated into computer programs for the calculation of chemical equilibria in aqueous solutions, such as MINEQL + (122), the disadvantage of this approach can be that the Na binding constant is purely empirical and does not reflect the electrostatic nature of the Na binding.

Another possibility is the use of the Donnan model (123). An application of this approach for different charged polymers has been pioneered by Marinsky (124) and coworkers, who established the use of Donnan models for interpreting protonation equilibria in organic polyelectrolytes including humic and fulvic substances and some metals (125–127) and alginic acid (128). The latter contribution also involves an estimation of the water uptake (swelling) using osmotic coefficients. Unfortunately, this modeling work did not include model predictions of proton or metal binding under different conditions. More Donnan-based model work done with alginate considered competition of Cu and H for the same sites (129,130).

Güntelberg activity coefficients have been used to account for ionic-strength effects in conjunction with fulvic acid and Cu–alkaline earth metal binding competition (131). Although this method may be suitable for rather small molecules, it would not be appropriate for polyelectrolytes, such as the alginic acid found in seaweed-based biosorbents. While the isotherm sorption model did not include a term for free sites, it was noticed that increasing ionic strength (Na) or presence of Ca or Mg can reduce Cu binding.

A model that is suitable for both oligo- and polyelectrolytes was used only for modeling pH titrations and Cu binding of humic acid molecules (132). Two of the sites were assumed to be diprotic and able to bind metal ions and one was assumed to be monoprotic and not to bind metals. In this approach, the humic acid molecules were assumed to be penetrable or impenetrable spheres of two size classes (assuming that the total charge is proportional to the volume of the sphere). The advantage of this approach is its generality concerning the size of the molecule. It is particularly suited for oligoelectrolytes that show a behavior intermediate between simple ions and polyelectrolyte gels. Most biosorbent materials could be considered true polyelectrolytes. However, the semiempirical correlation used in this approach required three fitting parameters for the electrostatic effects for each of the two particle sizes. Moreover, the fitting parameters had to be determined individually for each ionic strength. This means that the effect of ionic strength could not be predicted.

A Helmholtz metal-sorption model for the charged interface (133) requires only one fitting parameter, the constant capacitance, for modeling the electrostatic effects, but not the ionic strength effects. Another model considering eight sites of two general types for light-metal binding to humic acids used relatively few fitting parameters for the prediction of a wide range of experimental conditions (134,135). Unfortunately, it also relies on an empirical relationship involving the ionic strength. A Poisson–Boltzmann model for proton binding to humic substances (136) incorporates the effect of ionic strength on the concentration factor in a realistic manner and requires only one fitting parameter for electrostatic effects (the particle radius). The assumption of small impermeable spheres or cylinders, however, does not appear applicable to the large, penetrable particles as in the preferable case for biosorbents.

Recently, Schiewer and Volesky (119–121) presented a series of equilibrium biosorption models gradually incorporating the metal-ion concentration, pH, and ionic strength. The models proposed had only one fitting parameter, instead of up to six as in some of those already mentioned. The formation of $BM_{0.5}$ complexes was assumed, and a combined Donnan–biosorption–isotherm equation was derived that allowed for direct calculation of cation binding without iterations when swelling proportional to the number of free sites was assumed. Using the parameters (number of binding sites, proton binding constant, and specific particle volume) obtained from pH titrations, it was possible to predict the effect of pH, ionic strength, and Ca concentration on Cd binding for a seaweed biosorbent.

BIOSORPTION PROCESS DESIGN

Because biosorption is just another form of a basic sorption operation, the same principles apply for the design and sizing of the equipment. The approach is always a combination consisting of basic equilibrium sorption tests (batch) and dynamic continuous-flow apparatus studies, which can become very sophisticated if empiricism is to be minimized. Invariably, numerical computer methodologies provide a powerful tool for the design of a continuous-flow (bio)sorption process.

Equilibrium Sorption Process Basis

When contacting the solid-phase biosorbent with the sorbate species (i.e., metal ions) in solution, a sorption equilibrium will become established after an appropriate length of time. When equilibrium between the species sorbed on the solid and the portion remaining in solution is established, the solids are separated and the supernatant analyzed for the residual, final, equilibrium concentration of the remaining sorbate (Fig. 6). The metal uptake by the solid sorbent is calculated (according to the formula in Fig. 6) and used for the sorption isotherm plot, which serves to compare the sorption performance of different systems. This is easily done for one-sorbate (one-metal) systems. The situation becomes more complicated when there are more than one sorbed species in multisorbate systems.

Of interest is often the case of "interference" assessment when the uptake of one key metal (e.g., toxic Cd) may be affected by the presence of a ubiquitous metal of no major interest (e.g., Fe). This situation can be still depicted by three types of three-dimensional sorption isotherm surfaces, depending on what particular metal uptake appears in the plot: (*1*) Cd uptake, (*2*) Fe uptake, and (*3*) combined (Cd and Fe) uptake. The base two axes are always the equilibrium concentrations of the two metals, respectively.

Since the experimental equilibrium (final) concentrations are the product of the experiment, there is no control

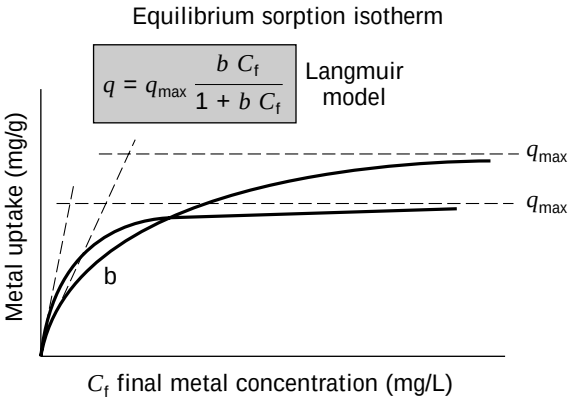


Figure 6. Biosorption-isotherm relationship curves; *b* is related to initial slope and indicates the sorbent "affinity."

over their values and they are generated more or less randomly. Consequently, the sorption isotherm surface needs to be fitted to the date points. Finding a suitable mathematical model capable of representing reasonably well the two-metal sorption behavior enables this and more; it can facilitate the summarizing of the effect of one metal on the uptake of the other for the desirable conclusion such as depicted in Figures 7 and 8 (137).

For many metal biosorption cases an equilibrium model considering only one type of binding site would be adequate assuming the following reaction (138):

$$M^{2+} + 2B = 2BM_{0.5} \quad H^+ + B = BH \tag{8}$$

The model provides explicit formulas for metal uptakes as functions of final concentrations of the species present in the system, even for the case of heterovalent exchange, which makes it especially suitable for use in numerical simulations of dynamic sorption systems. The model equation for the equilibrium uptake of a bivalent metal ion M1 present in a binary system with hydrogen protons can be written as follows:

$$q^*_{M1} = \frac{0.5M_{WM1}C_t(K_{M1}C_{fM1})^{1/2}}{1 + K_H C_{fH} + (K_{M1}C_{fM1})^{1/2}} \tag{9}$$

Similarly, for a binary system containing bivalent species M1 and M2, the expressions for uptakes of each of the species can be written as follows:

$$q^*_{M1} = \frac{0.5M_{WM1}C_t(K_{M1}C_{fM1})^{1/2}}{1 + (K_{M1}C_{fM1})^{1/2} + (K_{M2}C_{fM2})^{1/2}} \tag{10}$$

$$q^*_{M2} = \frac{0.5M_{WM2}C_t(K_{M2}C_{fM2})^{1/2}}{1 + (K_{M1}C_{fM1})^{1/2} + (K_{M2}C_{fM2})^{1/2}} \tag{11}$$

The values of the parameters K_H, K_{M1}, and $K_{M2}$are normally determined by fitting of the model to the experimental data using, for example, the method of weighed least squares. The constant $C_t$in equations 9–11 may be obtained from pH and/or conductometric titration of protonated biomass with NaOH (67). Metal biosorption can often be viewed as a simple ion exchange; hence, the classical ion-exchange concept based on exchange-equilibrium constants and separation factors (139,140) can be applied to this case. For the ion-exchange reaction:

$$bA^{+a} + aB^{+b} = bA^{+a} + aB^{+b} \tag{12}$$

the equilibrium constant K_{AB} and the separation factor r_{AB} are defined as

$$K_{AB} = \frac{q_A^b C_{Bf}^a}{C_{Af}^b q_B^a} = \frac{y_A^b x_B^a}{x_A^b y_B^a} \cdot \frac{C_0^{a-b}}{Q^{a-b}} \tag{13a}$$

$$r_{AB} = \frac{y_a x_B}{x_a y_b} \tag{13b}$$

for the case of ideal behavior of the exchanging species in both of the phases. As can be seen from equation 13a, the total normality C_0 of the solution affects the equilibrium of a heterovalent exchange (i.e., when $a \neq b$), but it does not affect the equilibrium when the valence of the exchanging species is the same (i.e., $a = b$). Consequently, the ion-exchange isotherms of a species in a binary heterovalent system vary in shape depending on the total normality of the solution (141).

Ion-Exchange Equilibrium Relationships. The term *equilibrium sorption isotherm* is commonly used in biosorption to refer to a line connecting equilibrium metal uptakes plotted against the final concentrations of the metal. Typically, biosorption equilibrium data are obtained under the same pH. However, this is not the case in ion exchange where the sorption isotherm represents data points obtained at a constant normality of the solution. Furthermore, in compliance with the established convention, ion-exchange isotherms for binary systems usually display dimensionless uptakes as functions of dimensionless concentrations for the species of interest (Fig. 9):

$$\frac{q_M}{Q} = f\left(\frac{C_{ML}}{C_0}\right) \tag{14}$$

Connection between Equilibrium and Dynamic Sorption in Columns. Unlike the biosorption isotherm, the ion-exchange isotherm depicts an equilibrium situation that is relevant to the case of the metal biosorption in a fixed-bed adsorber in which the concentration of the species released from the biosorbent increases monotonically in the direction of the flow. In fact, the ion-exchange isotherm represents the "equilibrium line" that is, together with the "operating line," sometimes used for assessing design and performance of fixed-bed adsorbers (142,143). According to the shape of the ion-exchange isotherm, the equilibrium is classified as (*1*) favorable when the isotherm is concave, (*2*) linear when the isotherm is linear, or (*3*) unfavorable when the isotherm is convex. It is known that the favorable isotherm leads to the development of a self-sharpening dynamic sorption zone (front) in the fixed-bed adsorber, whereas the unfavorable isotherm causes its broadening (144,145). As an alternative to ion-exchange isotherms, tables of values of separation factors for the compositions

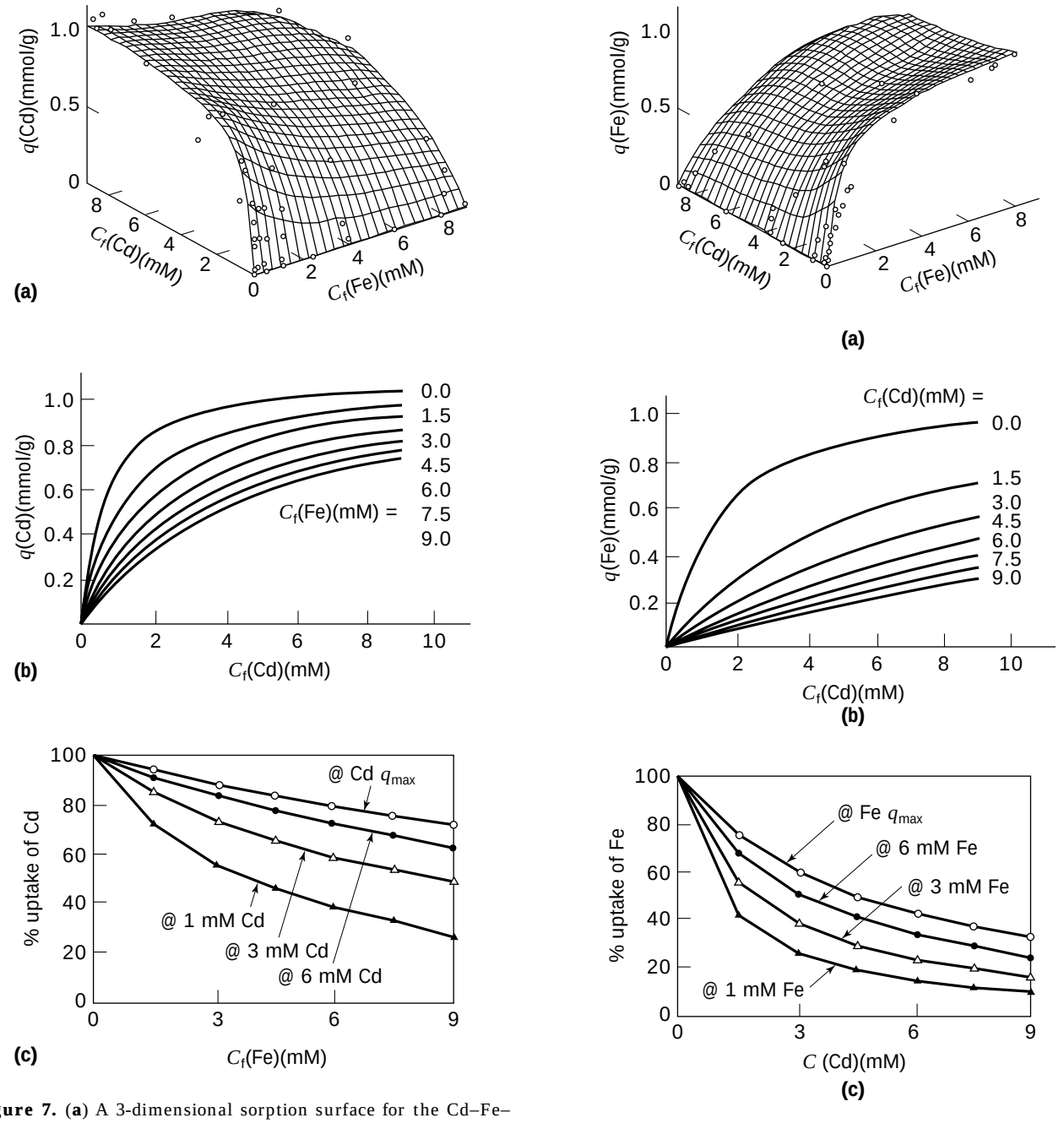


Figure 7. (**a**) A 3-dimensional sorption surface for the Cd–Fe–*Sargassum* biosorption system, cadmium uptake at pH 4.5; (**b**) the effect of Fe on the equilibrium uptake of Cd by *Sargassum* biomass at pH 4.5; (**c**) The summary of the effect of Fe presence on the Cd uptake by *Sargassum* biomass at pH 4.5. Equilibrium Cd concentrations were arbitrarily selected: ○, decreasing Cd uptake for different q_{max} at "high" C_f(Fe); ●, [Cd] = 6 mM; △, [Cd] = 3 mM; ▲, [Cd] = 1 mM.

Figure 8. (**a**) A 3-dimensional sorption surface for the Cd–Fe–*Sargassum* biosorption system, iron uptake at pH 4.5. (**b**) The effect of Cd on the equilibrium uptake of Fe by *Sargassum* biomass at pH 4.5. (**c**) The summary of the effect of Cd presence on the Fe uptake by *Sargassum* biomass at pH 4.5. Equilibrium Fe concentrations were arbitrarily selected: ○, decreasing Fe uptake for different q_{max}; ●, [Fe] = 6 mM; △, [Fe] = 3 mM; ▲, [Fe] = 1 mM.

existing in an adsorber can be used to represent the equilibrium for the purpose of computer simulation of the fixed bed (146).

Continuous-Flow Fixed-Bed Column Biosorption Process

In process applications, the most effective apparatus for cyclic sorption/desorption making the most effective use of the reactor volume is a fixed-bed column that also optimizes the concentration difference known to be the driving force in sorption processes. The process of metal biosorption in a fixed-bed column is influenced by three key regimes: the sorption equilibrium, the sorption particle mass transfer, and the flow pattern through the packed bed. Combination of these three determines the overall performance of the sorption column, which is judged by its ser-

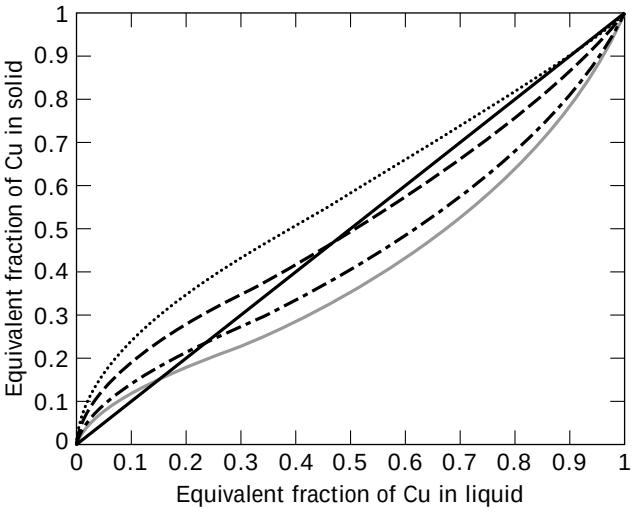


Figure 9. Calculated Cu^{2+} ion-exchange sorption isotherms for *Sargassum* biosorption system. C_o: —, 6 mequiv/L; —·—, 4 mequiv/L; — —, 2 mequiv/L; ..., 1 mequiv/L.

vice time, that is, the length of time until the sorbed species breaks through the bed to be detected at a given concentration in the column effluent (Fig. 10). At that point, the bed is considered for all practical purposes saturated, and the equipment has to be taken out of operation for some kind of regeneration (Fig. 1).

The lack of design information on biosorption in columns is due, in part, to the fact that the metal uptake by different biomaterials is still not fully understood, as well as because most of the biomasses have to be immobilized, granulated, or otherwise processed before their use in a column to provide good flow-through conditions and an acceptable pressure drop across the bed. This process condition must eventually be balanced and optimized against the maximum-sorption-performance requirement.

As opposed to activated carbon-sorption cases, metal biosorption is much closer to the heterovalent ion-exchange model because a considerable number of ions is being released from the biosorbent in exchange for the sorbed metal ion. The most universal approach seems to be that adopted by Tan and Spinner (146), who numerically solved a mixed system of partial differential, ordinary differential, and algebraic equations describing the dynamics of multicomponent ion exchange in a fixed bed. A term can be added to the differential molar balance equation to account for axial dispersion of the fluid in the packed bed. For ionic species M the molar balance is

$$\frac{C_M}{\partial z} - \frac{1}{Pe_c}\frac{\partial^2 C_M}{\partial z^2} + \frac{\partial C_M}{\partial t} + D_{gM}\frac{q_M}{t} = 0 \qquad (15)$$

The sorption rate equation can be written as

$$\frac{q_M}{t} = Sh_M(q_M^* - q_M) \qquad (16)$$

assuming a linear driving force for the sorption process and combined film and intraparticle mass transfer resistance: The equilibrium relationship for a metal species represented by equations (9) and (10) can be written schematically for a general counterion, either proton or metal, as

$$q_M^* = f(C_M, C_{CNT}) \qquad (17)$$

and the electroneutrality equation for binary systems is

$$\frac{C_{M0} - C_M}{M_{WM}} = \nu\frac{C_{CNT0} - C_{CNT}}{M_{WCNT}} \qquad (18)$$

To obtain equations 15–18, the original set of equations was transcribed with dimensionless variables using the following transformations of the variables:

$$C_M = \frac{C_{ML}}{C_0}\; q_M = \frac{q_{MS}}{Q}\; t = \frac{\tau v}{L_0}\; z = \frac{l}{L_0}$$

and yielding the following dimensionless groups:

$$D_{gM} = \frac{\rho_b Q}{C_0 \epsilon}\; Sh_M = \frac{K_{fM} L_0}{v}\; Pe_c = \frac{L_0 v}{D_z}$$

The initial and boundary conditions pertaining to the situations under question are specified as follows:

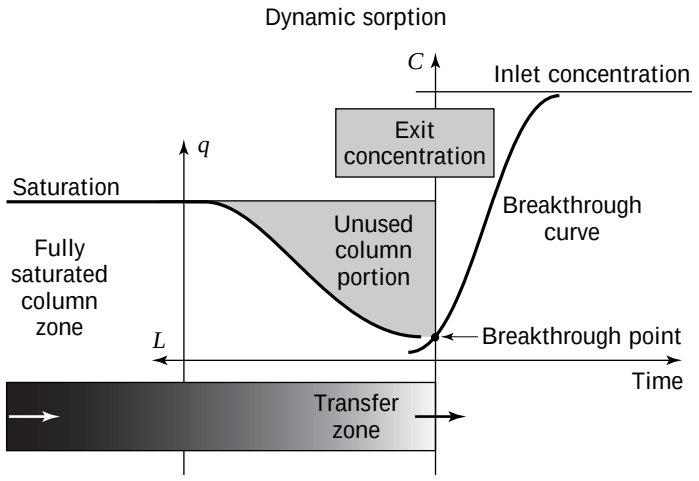


Figure 10. Biosorption column operation: the breakthrough point and the breakthrough curve for the column exit concentration.

$$t = 0 \quad l = (0, L_0) \quad C_M = 0$$

$$t > 0 \quad l = 0 \quad C_M = C_{M0} + \frac{1}{Pe_c}\frac{C_M}{z}$$

$$t > 0 \quad l = L_0 \quad \frac{C_M}{z} = 0$$

The system of equations 15–18 can be solved numerically using the Galerkin finite element method (147), combined with the implicit Euler method for time integration. Linear-basis functions would routinely be used for space discretization of both concentration C and uptake q, ensuring that the solution is mesh independent.

Fitting the Equilibrium Model and Calculation of Ion-Exchange Isotherms. Equilibrium data provide the experimental basis for the fitting of the biosorption model (148). First, the concentration of binding sites C_t can be determined from the uptake experiments; equilibrium constants for the three ions considered here (K_H, K_{Cu}, and K_{Ca}) can then be found by the least square fitting.

Once these parameters are known, the biosorption model can be used to calculate ion-exchange isotherms as follows: First, the normality of a binary mixture C_0 was fixed and the uptake Q, at the concentration of the sorbing species C_0, was calculated from equation 2 or 3. Second, a series of liquid-phase compositions was selected in such a way that the dimensionless concentration C/C_0 of one of the species covers the interval of interest. Finally, a series of dimensionless uptakes corresponding to the series of compositions was calculated using equations 2, 3, and 7.

Fitting and Use of the Flow-through Fixed-Bed Sorption Column Model. The dimensionless groups Sh, Dg, and Pe, together with the equilibrium parameters for a given binary system, represent the parameters of the fixed-bed model. However, because the equilibrium parameters can be determined from batch equilibrium studies, the values of C_0, v, L_0, and ρ_b are usually known, and Q can be calculated from equations 2 and 3, the only remaining unknown elements in the model are K_{fM}, D_z, and ϵ. The values of the three unknown elements K_{fM}, D_z, and ϵ must be determined simultaneously by fitting the fixed-bed model to an experimental breakthrough curve.

The value of an overall mass-transfer coefficient for a given species generally depends on the flow rate, the axial dispersion coefficient, and the values of the diffusivities of the species in the liquid and solid phases. The diffusion coefficients of metal ions in diluted aqueous solutions can often be obtained from the literature (149,150).

Conclusion. In preparation for the biosorption process design the sorption/desorption performance and characteristic properties (mass transfer, pressure and attrition resistance, chemical stability, etc.) of the biosorbent should be reasonably well known. Although equilibrium sorption tests yield some basic information, the sorption system behavior is invariably examined under continuous-flow conditions. The fixed-bed sorption column arrangement is usually the most effective process configuration. Its performance assessment and prediction based on both equilibrium and dynamic studies eventually leads to sizing of the equipment. Advanced methodologies used for this purpose are based on sophisticated computer-based modeling and optimization of the (bio)sorption process.

NOMENCLATURE

a,b	stoichiometric coefficients in the ion-exchange reaction equation 12
$\bar{b}$	Langmuir constant related to affinity
A,B	species in the resin phase
A,B	species in the liquid
B	biosorbent binding site
B_t	total number of binding sites on the biosorbent
BH	biosorbent–proton complex
BM	biosorbent–metal complex
C_i	initial metal concentration
C_0	normality of the solution [mequiv/L]
$C_M = x_M$	equivalent fraction of species M in liquid phase
C_{ML}	concentration of species M in liquid phase [mequiv/L]
$C_{Mf} = C_f$	equilibrium final concentration of species M in liquid phase [mequiv/L]
C_{M0}	equivalent fraction of species M in the feed to a column
C_{CNT}	equivalent fraction of counterion in liquid phase
C_{CNT0}	equivalent fraction of counterion in the feed to a column
C_t	concentration of binding sites in biomass [mequiv/L]
D_z	axial dispersion coefficient [cm^2/s]
G	free (Gibbs) energy
H or H^+	proton
H	enthalpy
${}^{BM}k_f$	rate of adsorption
${}^{BM}k_b$	rate of desorption
K_{AB}	ion-exchange equilibrium constant
${}^{BM}K$	biosorbent–metal ion-exchange equilibrium constant
K_H, K_{M1}, K_{M2}	equilibrium constants of biosorption equilibrium model
K_{fM}	overall mass-transfer coefficient of species M [min^{-1}]
l	vertical distance from the top of a column
L_0	length of the column [cm]
M	metal ion in solution
M_{WM}	molecular weight of the sorbing metal
M_{WCNT}	molecular weight of the counterion
n	Freundlich sorption isotherm model constant

$q_{MS} = {}^{M}q = q$	uptake of species M [mequiv/L]
q_{max}	maximum uptake (Langmuir parameter)
$q_M = y_M$	equivalent fraction of species M in solid phase
q_M^*	dimensionless equilibrium uptake of species M at C_M
Q	equilibrium uptake of M at $C_{ML} = C_{M0}$. C_0 [mequiv/L]
r	ionic radius (subscripts: hyd = hydrated; cryst = crystalline)
r_{AB}	ion-exchange separation factor
R	gas constant
S	weight of biosorbent solids (Fig. 5) or entropy (equation 1)
t	dimensionless time
T	absolute temperature (K)
v	interstitial velocity of the fluid [cm/min]
V	volume of solution
x	electronegativity
z	ionic charge

Greek Symbols

Δ	difference (equation 1)
ϵ	column void fraction
ν	stoichiometric coefficient of the exchange reaction
ρ_b	packing density [g/L]
τ	time [min]
ξ	parameter in Table 1 $\xi = (z^2/r_{hyd})/(1 - \exp(-\Delta x^2/2))$

Dimensionless Groups

$D_{gM} = \dfrac{\rho_b Q}{C_0 \epsilon}$	solute distribution parameter
$Pe_c = \dfrac{L_0 v}{D_Z}$	column Peclet number
$Sh_M = \dfrac{K_{fM} L_0}{v}$	modified Sherwood number

BIBLIOGRAPHY

1. J.W. Moore and S. Ramamoorthy, *Heavy Metals in Natural Waters*, Springer-Verlag, New York, 1984, pp. 28–57, 77–99, 182–204.
2. WHO, *Guidelines for Drinking Water Quality*, World Health Organization, Geneva, 1993, pp. 174–181.
3. EPS, *Metal Finishing Liquid Effluent Guidelines, Report EPS 1-WP-77-5*, Water Pollution Control Directorate, Environmental Protection Service, Fisheries and Environment Canada, 1977, pp. 5958–5960.
4. UNEP, *Environmental Aspects of the Metal Finishing Industry: A Technical Guide*, United Nations Environment Programme, Industry and Environment Office, Paris, 1989, pp. 19–39, 53–57.
5. C.T. Tien and C.P. Huang, in J.P. Vernet ed., *Heavy Metals in the Environment*, Elsevier, Amsterdam, 1991, pp. 295–311.
6. W.A.P. Black and R.L. Mitchell, *J. Marine Biol. Assoc.* **30**, 575–584 (1952).
7. N. Kuyucak, in B. Volesky ed., *Biosorption of Heavy Metals*, CRC Press, Boca Raton, Fla., 1990, pp. 371–378.
8. D. Kratochvil, E. Fourest, and B. Volesky, *Biotechnol. Lett.* **17**, 777–782 (1995).
9. B. Volesky and Z.R. Holan, *Biotechnol. Prog.* **11**, 235–250 (1995).
10. A. Kapoor and T. Viraraghavan, *Biores. Technol.* **53**, 195–206 (1995).
11. H. Mann, B. Volesky ed., *Biosorption of Heavy Metals*, CRC Press, Boca Raton, Fla., 1990, pp. 93–137.
12. N. Kuyucak and B. Volesky, in B. Volesky ed., *Biosorption of Heavy Metals*, CRC Press, Boca Raton, Fla., 1990, pp. 173–198.
13. Z.R. Holan and B. Volesky, *Biotechnol. Bioeng. 43*, 1001–1009 (1994).
14. J. Ferguson and B. Bubela, *Chem. Geol.* **13**, 163–186 (1974).
15. Y.S. Ho, D.A.J. Wase, and C.F. Forster, *Wat. Res.* **29**, 1327–1332 (1995).
16. X.H. Chen, T. Gosset, and D.R. Thevenot, *Wat. Res.* **24**, 1463–1471 (1990).
17. B. Volesky, in B. Volesky ed., *Biosorption of Heavy Metals*, CRC Press, Boca Raton, Fla., 1990, pp. 3–6.
18. B. Volesky and H.A. May-Phillips, *Appl. Microbiol. Biotechnol.* **42**, 797–806 (1995).
19. M. Tsezos, in H.L. Ehrlich and C.L. Brierley eds., *Microbial Mineral Recovery*, McGraw-Hill, New York, 1990, pp. 325–340.
20. Z.R. Holan, B. Volesky, and I. Prasetyo, *Biotechnol. Bioeng.* **41**, 819–825 (1993).
21. I. Prasetyo, M. Eng. Thesis, McGill University, Montreal, Canada, 1992.
22. N. Kuyucak and B. Volesky, *Biorecovery* **1**, 189–204 (1989).
23. A. Leusch, Z.R. Holan, and B. Volesky, *J. Chem. Tech. Biotechnol.* **62**, 279–288 (1995).
24. G.J. Ramelow, D. Fralick, and Y. Zhao, *Microbios* **72**, 81–93 (1992).
25. A. Haug and O. Smidsrod, *Acta Chem. Scand.* **24**, 843–854 (1970).
26. P. Weppen and A. Hornburg, *Thermochim. Acta* **269/270**, 393–404 (1995).
27. J.M. Tobin, D.G. Cooper and R.J. Neufeld, *Biotechnol. Bioeng.* **30**, 882–886 (1987).
28. J.A. Davis and J.O. Leckie, *Environ. Sci. Technol.* **12**, 1309–1315 (1978).
29. B. Nowack, J. Luetzenkirchen, P. Behra, and L. Sigg, *Environ. Sci. Technol.* **30**, 2397–2405 (1996).
30. M.A. Glaus, W. Hummel, and L.R. Van Loon, *Environ. Sci. Technol.* **29**, 2150–2153 (1995).
31. B. Greene, M. Hosea, R. McPherson, M. Henzl, M.D. Alexander, and D.W. Darnall, *Environ. Sci. Technol.* **20**, 627–632 (1986).
32. D.W. Darnall, B. Greene, M.T. Henzl, J.M. Hosea, R.A. McPherson, J. Sneddon, and M.D. Alexander, *Environ. Sci. Technol.* **20**, 206–208 (1986).

33. B. Greene, R. McPherson, and D. Darnall, in J.W. Patterson and R. Pasino eds., *Metals Speciation, Separation, and Recovery*, Lewis, Chelsea, Mich., 1987, pp. 315–338.
34. N. Kuyucak and B. Volesky, *Biotechnol. Bioeng.* **33**, 809–814 (1989).
35. M. Tsezos and B. Volesky, *Biotechnol. Bioeng.* **23**, 583–604 (1981).
36. B. Greene, M.T. Henzl, J.M. Hosea, and D.W. Darnall, *Biotechnol. Bioeng.* **28**, 764 (1986).
37. J.M. Tobin, D.G. Cooper, and R.J. Neufeld, *Appl. Envir. Microbiol.* **47**, 821–824 (1984).
38. I. Aldor, E. Fourest, and B. Volesky, *Can. J. Chem. Eng.* **73**, 516–522 (1995).
39. N. Kuyucak and B. Volesky, *Biotechnol. Bioeng.* **33**, 815–822 (1989).
40. C.F.J. Baes and R.E. Mesmer, *The Hydrolysis of Cations*, Wiley New York, 1976, pp. 397–419.
41. F.M.M. Morel, *Principles of Aquatic Chemistry*, Wiley, New York, 1983, pp. 237–266.
42. W. Stumm and J.J. Morgan, *Aquatic Chemistry*, Wiley, New York, 1970, pp. 238–299.
43. C.F.J. Baes and R.E. Mesmer, *The Hydrolysis of Cations*, Wiley, New York, 1976, pp. 1–7.
44. G.K. Pagenkopf, *Introduction to Natural Water Chemistry*, Marcel Dekker, New York, 1978, pp. 161–167, 214–216, 220–230.
45. Y.E. Collins and G. Stotzky, *Appl. Environ. Microbiol.* **58**, 1592–1600 (1992).
46. W. Stumm and J.J. Morgan, *Aquatic Chemistry*, Wiley, New York, 1970, pp. 445–513.
47. J.M. Tobin, D.G. Cooper, and R.J. Neufeld, *Biotechnol. Bioeng.* **31**, 282–286 (1988).
48. K.H. Chong and B. Volesky, *Biotechnol. Bioeng.* **47**, 451–460 (1995).
49. B. Volesky, in B. Volesky ed., *Biosorption of Heavy Metals*, CRC Press, Boca Raton, Fla., 1990, pp. 7–43.
50. D. Myers, *Surfaces, Interfaces, Colloids: Principles and Applications*, VCH, Weinheim, Germany, 1991, pp. 39–67.
51. J.C. Westall, in W. Stumm ed., *Aquatic Surface Chemistry*, Wiley, New York, 1987, pp. 3–32.
52. I.T. Mayers and T.J. Beveridge, *Can. J. Microbiol.* **35**, 764–770 (1989).
53. L.E. Macaskie, R.M. Empson, A.K. Cheetham, C.P. Grey, and A.J. Skarnulis, *Science* **257**, 782–784 (1992).
54. L.E. Macaskie, A.C.R. Dean, A.K. Cheetham, R.J.B. Jakeman, and A.J. Skarnulis, *J. Gen. Microbiol.* **133**, 539–544 (1987).
55. R.S. Cahn and O.L. Dermer, *Introduction to Chemical Nomenclature*, Butterworths, London, 1979, pp. 17–18.
56. W.C. Fernelius, in *Chemical Nomenclature*, American Chemical Society, Washington, D.C., 1953, pp. 9–15.
57. J.M. Smith, *Chemical Engineering Kinetics*, McGraw-Hill, New York, 1981, pp. 310–322.
58. J.B. Russell, *General Chemistry*, McGraw-Hill, New York, 1980, pp. 314–316, 340–341.
59. N. Kuyucak and B. Volesky, *Biotechnol. Bioeng.* **33**, 823–831 (1989).
60. R.H. Crist, J.R. Martin, P.W. Guptill, J.M. Eslinger, and D.R. Crist, *Environ. Sci. Technol.* **24**, 337–342 (1990).
61. M.E. Treen-Sears, B. Volesky, and R.J. Neufeld, *Biotechnol. Bioeng.* **26**, 1323–1329 (1984).
62. R.H. Crist, K. Oberholser, N. Shank, and M. Nguyen, *Environ. Sci. Technol.* **15**, 1212–1217 (1981).
63. R.H. Crist, K. Oberholser, D. Schwartz, J. Marzoff, D. Ryder, and D.R. Crist, *Environ. Sci. Technol.* **22**, 755–760 (1988).
64. R.H. Crist, K. Oberholser, J. McGarrity, D.R. Crist, J.K. Johnson, and J.M. Brittsan, *Environ. Sci. Technol.* **26**, 496–502 (1992).
65. D. Kaplan, D. Christiaen, and S. Arad, *Appl. Environ. Microbiol.* **53**, 2953–2956 (1987).
66. J.L. Gardea-Torresdey, M.K. Becker-Hapak, J.M. Hosea, and D.W. Darnall, *Environ. Sci. Technol.* **24**, 1372–1378 (1990).
67. E. Fourest and B. Volesky, *Environ. Sci. Technol.* **30**, 277–282 (1996).
68. V.E. Bower and R.G. Bates, in L. Meites ed., *Handbook of Analytical Chemistry*, McGraw-Hill, New York, 1963, pp. 1.20–1.27.
69. J. Buffle, *Complexation Reactions in Aquatic Systems: An Analytical Approach*, Ellis Horwood, Chichester, U.K., 1988, pp. 59–65, 74, 284–286.
70. T.J. Beveridge and R.G.E. Murray, *J. Bacteriol.* 876–887 (1980).
71. T.J. Beveridge, in H.L. Ehrlich and D.S. Holmes eds., *Biotechnology and Bioengineering Symposium No. 16: Biotechnology for the Mining, Metal-Refining, and Fossil Fuel Processing Industries*, Wiley, New York, 1986, pp. 127–140.
72. M. Tsezos and B. Volesky, *Biotechnol. Bioeng.* **24**, 955–969 (1982).
73. M. Tsezos and B. Volesky, *Bioeng. Biotechnol.* **24**, 385–401 (1982).
74. N. Kuyucak and B. Volesky, *Biorecovery* **1**, 219–235 (1989).
75. A. Haug, *Acta Chem. Scand.* **15**, 1794–1795 (1961).
76. A. Haug and O. Smidsrod, *Acta Chem. Scand.* **19**, 341–351 (1965).
77. R.G. Schweiger, *Kolloid Z.* **196**, 47–53 (1964).
78. A.J. Paskins-Hurlburt, Y. Tanaka, and S.C. Skoryna, *Bot. Mar.* **19**, 327–328 (1976).
79. J.M. Tobin, Ph.D. Thesis, McGill University, Montreal, Canada, 1986.
80. S.V. Avery and J.M. Tobin, *Appl. Environ. Microbiol.* **59**, 2851–2856 (1993).
81. L.J. Evans, *Environ. Sci. Technol.* **23**, 1046–1056 (1989).
82. A. Leusch, Z.R. Holan, and B. Volesky, *Appl. Biochem. Biotechnol.* **61**, 231–249 (1996).
83. R.J.P. Williams and J.D. Hale, *Structure and Bonding* **1**, 249–281 (1966).
84. Y. Marcus and A.S. Kertes, *Ion Exchange and Solvent Extraction of Metal Complexes*, Wiley London, 1969, pp. 10–33.
85. Y. Marcus and A.S. Kertes, *Ion Exchange and Solvent Extraction of Metal Complexes*, Wiley London, 1969, pp. 277–298.
86. M.K. Jain and R.C. Wagner, *Introduction to Biological Membranes*, Wiley, New York, 1980, pp. 196–201.
87. D.R. Turner, M. Whitfield, and A.G. Dickson, *Geochim. Cosmochim. Acta* **45**, 855–881 (1981).
88. J.A. Dean, *Lange's Handbook of Chemistry*, 13th ed., McGraw-Hill, New York, 1985, pp. 3.11–3.12.
89. R.G. Pearson, *J. Am. Chem. Soc.* **85**, 3533–3539 (1963).
90. R.J.P. Williams, *Proc. Chem. Soc.* 20–21 (1960).
91. C.S.G. Phillips and R.J.P. Williams, *Inorganic Chemistry: Principles and Non-Metals*, Oxford University Press, New York, 1965, pp. 142–164.

92. S. Ahrland, J. Chatt, and N.R. Davis, *Quart. Rev. Chem.* **12**, 265–276 (1958).
93. C.F. Bell, *Principles and Applications of Metal Chelation*, Clarendon, Oxford, UK 1977, pp. 30–50.
94. L. Pauling, *Nature of the Chemical Bond*, Cornell University Press, Ithaca, N.Y., 1967, pp. 55–73.
95. H.T. Evans, D.R. Lide ed., *CRC Handbook of Chemistry and Physics*, CRC Press, Boca Raton, Fla., 1993, pp. 12.8–12.9.
96. E. Nieboer and W.A.E. McBryde, *Can. J. Chem.* **51**, 2512–2524 (1973).
97. E. Nieboer and D.H.S. Richardson, *Environ. Poll.* **1B**, 11–13 (1980).
98. M. Tsezos and A.A. Deutschmann, *J. Chem. Tech. Biotechnol.* **48**, 29–39 (1990).
99. H. Freundlich, *Z. Physik. Chem.* **57**, 385–470 (1907).
100. W. Stumm, *Chemistry of the Solid–Water Interface.* Wiley, New York, 1992, pp. 87–97.
101. W.J.J. Weber, in F.L. Slejko ed., *Adsorption Technology: A Step by Step Approach to Process Evaluation and Application*, Marcel Dekker, New York, 1985, pp. 9–17.
102. I. Langmuir, *J. Am. Chem. Soc.* **40**, 1361–1403 (1918).
103. H.B. Xue and L. Sigg, *Wat. Res.* **24**, 1129–1136 (1990).
104. C. Huang, C.P. Huang, and A.L. Morehart, *Wat. Res.* **25**, 1365–1375 (1991).
105. J. Votruba, in *Proceedings of IUMS Congress*, Prague, 1994, p. 283.
106. M. Gonzales-Davila, J.M. Santana-Casiano, J. Perez-Pena, and F.J. Millero, *Environ. Sci. Technol.* **29**, 288–301 (1995).
107. J.-P. Huang, C.P. Huang, and A.L. Morehart, in J.P. Vernet ed., *Heavy Metals in the Environment*, Elsevier Science, Amsterdam, 1991, pp. 329–349.
108. L.K. Jang, D. Nguyen, and G.G. Geesey, *Wat. Res.* **29**, 315–321 (1995).
109. H. Seki, A. Suzuki, and I. Kashiki, *J. Coll. Int. Sci.* **134**, 59–65 (1990).
110. J.-W. Yu and I. Neretnieks, *Ind. Eng. Chem. Res.* **29**, 220–231 (1990).
111. E.M. Trujillo, T.H. Jeffers, C. Ferguson, and H.Q. Stevenson, *Environ. Sci. Technol.* **25**, 1559–1565 (1991).
112. L.K. Jang, D. Nguyen, and G.G. Geesey, *Wat. Res.* **29**, 307–313 (1995).
113. D.K. Nordstrom and J.W. Ball, in C.J.M. Kramer and J.C. Duinker eds., *Complexation of Trace Metals in Natural Waters*, Nijhof/Junk, The Hague, The Netherlands, 1984, pp. 149–164.
114. R.H. Crist, J.R. Martin, D. Carr, J.R. Watson, H.J. Clarke, and D.R. Crist, *Environ. Sci. Technol.* **28**, 1859–1866 (1994).
115. Y. Marcus and A.S. Kertes, *Ion Exchange and Solvent Extraction of Metal Complexes*, Wiley London, 1969, pp. 345–351.
116. F. Helfferich, *Ion Exchange*, McGraw-Hill, New York, 1962, pp. 72–94.
117. J. Buffle, *Complexation Reactions in Aquatic Systems: An Analytical Approach*, Ellis Horwood, Chichester, 1988, pp. 195–303.
118. J.C. Westall, J.D. Jones, G.D. Turner, and J.M. Zachara, *Environ. Sci. Technol.* **29**, 951–959 (1995).
119. S. Schiewer and B. Volesky, *Environ. Sci. Technol.* **31**, 1863–1871 (1997).
120. S. Schiewer and B. Volesky, *Environ. Sci. Technol.* **31**, 2478–2485 (1997).
121. S. Schiewer, Ph.D. Thesis, McGill University, Montreal, Canada, 1996.
122. W.D. Schecher, *MINEOL+: A Chemical Equilibrium Program for Personal Computers, Users Manual Version 2.22*. Environmental Research Software, Inc., Hallowell, Maine, 1991.
123. F.G. Donnan, *Z. Elektroch.* **17**, 572–581 (1911).
124. J.A. Marinsky, in W. Stumm ed., *Aquatic Surface Chemistry*, Wiley, New York, 1987, pp. 49–81.
125. J.A. Marinsky and J. Ephraim, *Environ. Sci. Technol.* **20**, 349–354 (1986).
126. J. Ephraim, S. Alegret, A. Mathuthu, M. Bicking, R.L. Malcolm, and J.A. Marinsky, *Environ. Sci. Technol.* **20**, 354–366 (1986).
127. J.A. Marinsky, S. Gupta, and P. Schindler, *J. Coll. Int. Sci.* **89**, 412–426 (1982).
128. F.G. Lin and J.A. Marinsky, *React. Polymers* **19**, 27–45 (1993).
129. L.K. Jang, N. Harpt, D. Grasmick, L.N. Vuong, and G. Geesey, *J. Phys. Chem.* **94**, 482–488 (1990).
130. A. Katchalsky, R.E. Cooper, J. Upadhyay, and A. Wassermann, *J. Am. Chem. Soc.* **83**, 5198–5204 (1961).
131. S.E. Cabaniss and M.S. Shuman, *Geochim. Cosmochim. Acta* **52**, 185–193 (1988).
132. B.M. Bartschat, S.E. Cabaniss, and F.M.M. Morel, *Environ. Sci. Technol.* **26**, 284–294 (1992).
133. H.-B. Xue, W. Stumm, and L. Sigg, *Wat. Res.* **22**, 917–926 (1988).
134. E. Tipping, C.A. Backes, and M.A. Hurley, *Wat. Res.* **22**, 597–611 (1988).
135. E. Tipping, *Environ. Sci. Technol.* **27**, 520–529 (1993).
136. J.C.M. de Wit, W.H. van Riemsdijk, and L.K. Koopal, *Environ. Sci. Technol.* **27**, 2005–2014 (1993).
137. M.M. Figueira, B. Volesky, and V.S.T. Ciminelli, *Biotechnol. Bioeng.* **54**, 344–350 (1997).
138. S. Schiewer and B. Volesky, *Environ. Sci. Technol.* **29**, 3049–3058 (1995).
139. T. Vermeulen, G. Klein, and N.K. Heister, in R.H. Perry and C.H. Chilton eds., *Chemical Engineers' Handbook*, 5th ed., McGraw-Hill, New York, 1973, Ch. 16.
140. M. Shallcross, C.C. Herrmann, and B.J. McCoy, *Chem. Eng. Sci.* **43**, 279–288 (1988).
141. W.A. Selke, in J. Schubert and F.C. Nachod eds., *Ion Exchange Technology*, Academic Press, New York, 1956, pp. 53–62.
142. A.S. Michaels, *Int. Eng. Chem.* **44**, 1922–1929 (1952).
143. L. Liberti, in L. Liberti and G. Helfferich eds., *Mass Transfer and Kinetics of Ion Exchange. NATO ASI Series E. Applied Sciences No. 71*, Martinus Nijhoff, The Hague, The Netherlands, 1983, pp. 181–206.
144. F. Helfferich, *Ion Exchange*, McGraw-Hill, New York, 1962, pp. 299–319.
145. J.M. Coulson and J.F. Richardson, *Chemical Engineering*, 4th ed., Pergamon Press, New York, 1991, pp. 774–775.
146. H.K.S. Tan and I.H. Spinner, *Can. J. Chem. Eng.* **72**, 330–341 (1994).
147. C.A.J. Fletcher, *Computational Galerkin Methods*, Springer-Verlag, New York, 1984, pp. 246–276.
148. D. Kratochvil, B. Volesky, and G. Demopoulos, *Wat. Res.* **31**, 2327–2339 (1997).

BIOTRANSFORMATIONS, ENGINEERING ASPECTS

UDO KRAGL
ANDREAS LIESE
Forschungszentrum Jülich
Jülich, Germany

KEY WORDS

Chiral diol
Enzyme
GDP-mannose
Kinetics
Product inhibition
Reactor

OUTLINE

INTRODUCTION

The use of biotransformations in food, pharmaceutical, and chemical industries has increased steadily since the early 1970s. Applications range from synthesis on the milligram scale for precious oligosaccharides, to the multiton scale for bulk chemicals using either whole cells or isolated enzymes. This process has been documented in numerous reviews and textbooks (1–5). When new processes are being developed, the special properties of the biocatalyst must be considered, as well as the basic elements of chemical engineering. There are strong interactions among the individual aspects, which may be considered for process optimization. These are summarized in Figure 1.

The aim of this article is to outline the basic strategy for the development of a biocatalytic step. For some selected examples the beneficial role of (bio)chemical engineering is demonstrated. A more detailed description can be found in textbooks of (bio)chemical engineering (1,3, 6–8).

WHOLE CELLS OR ISOLATED ENZYMES?

Whole microorganisms or their components, enzymes, may be used for biotransformations. Both approaches have specific advantages and requirements, as summarized in Table 1. Commonly, the word *biotransformation* is used to describe the transformation of a substrate or precursor into a certain product in only a few steps. Fermentation processes, for example, production of amino acids from glucose and ammonia or production of (recombinant) proteins, are normally not regarded as biotransformation. Litera-

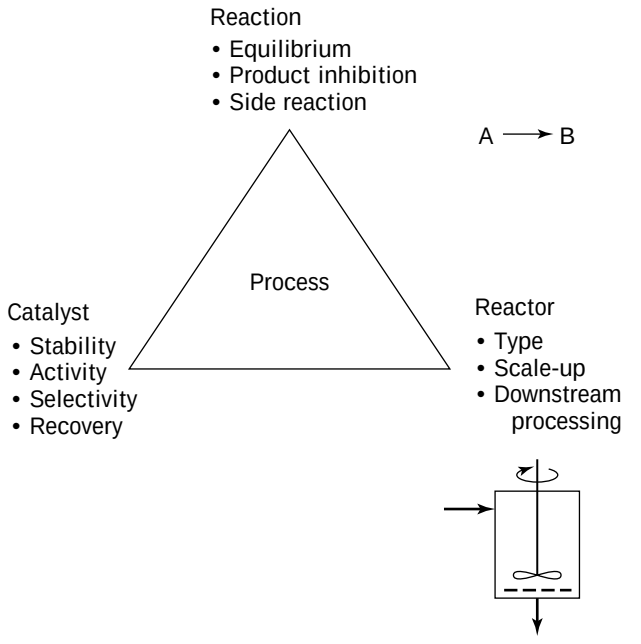


Figure 1. Elements of process development.

Table 1. Comparison of Biotransformations Using Whole Cells vs. Isolated Enzymes

Whole cells	Isolated enzymes
Large amounts of catalyst easy available	Additional step for obtaining the catalyst required
No stability problem when growing cells are used	Catalyst stability strongly influenced by reaction conditions
Limited volumetric catalyst activity	High volumetric catalyst activity possible
Multistep processes	Multistep processes difficult
Cofactor recycling solved	Cofactor recycling so far only for NADH and ATP
Side reactions may occur	Few to no side reactions
Toxicity of substrate/product may be complex	Enzyme stability may be influenced by substrate or product, but can be enhanced by choosing the appropriate reactor
Oxygen supply may control scale-up	Nonnatural substrates and reaction conditions possible
Normally no cosolvent, possibly leading to low concentrations of hydrophobic compounds	Cosolvents possible, but often reducing enzyme stability
Transport in and out of the cell may be limiting	Mass transport limitations only with immobilized enzymes

ture describing those processes can be found elsewhere in this volume or in external references (2,9–13).

Examples of industrial biotransformations using whole cells or isolated enzymes are given in Table 2. Some special information about process characteristics is given as well. When isolated enzymes are used for biotransformations, the purity must only be such that no side reactions occur. There is no need to purify the enzyme to homogeneity, as is required for biochemical characterization. On the contrary, many enzymes show reduced stability with increased purity. In some cases, the addition of inert proteins, such as bovine serum albumin, may be used as a simple way to increase the stability of the production enzyme. In others, enzyme stability may be increased by immobilization on a support. Covalent attachment fixes the conformation of the enzyme often limiting denaturation. However, the increase of stability must be balanced against the losses during the immobilization step, as well as the additional work and material costs. A method introduced quite recently for the stabilization of enzymes and their immobilization is the formation of cross-linked enzyme crystals (CLECs) (22). Reactors for immobilized enzymes are discussed elsewhere in this volume (see ENZYME, IMMOBILIZED, REACTOR). However, the engineering approach, as outlined in the following sections is the same for all forms of biocatalysts. Specific requirements and properties such as oxygen delivery, mass transport limitations, or particle size must be considered for individual cases.

CLASSIFICATION OF DIFFERENT REACTORS

There are three basic types of reactors, which have several variations (Fig. 2):

1. Stirred-tank reactors (STRs) or batch reactors
2. Continuously operated stirred-tank reactors (CSTRs)
3. Plug-flow reactors (PRFs)

Unlike the stirred-tank reactor, which is operated batchwise, the latter two are operated continuously. By knowing the main characteristics of these fundamental reactors and some of their variations, for example the cascade reactor of Figure 2, it is possible to choose the appropriate reactor for a specific application. This is especially important when dealing with a kinetically or thermodynamically limited system, as demonstrated later. In the following analyses, only the basic terms are explained. For more details the reader is referred to textbooks (6–8).

The stirred-tank reactor is operated in an instationary way (Fig. 2a). Assuming ideal mixing characteristics, the instantaneous concentration of every species is the same in every volume element. With advancing conversion the substrate concentration is decreased and the product concentration is increased. A steady state is reached only at equilibrium conversion.

In the plug-flow reactor the product concentration increases slowly over the length of the reactor (Fig. 2b). Therefore, the average reaction velocity is faster than in the continuously operated stirred-tank reactor. In each single-volume element in the reactor, the concentration is constant in the steady state. In other words, the dimension of time is exchanged with the dimension of place in comparison to the stirred-tank reactor.

The continuously operated stirred-tank reactor works under product outflow conditions, meaning that the concentrations in every volume element are the same as those at the outlet of the reactor (Fig. 2c). If the steady state is

Table 2. Examples of Industrial Biotransformations

Product	Substrate	Catalyst	Comments	References
Acrylamide	Acrylnitrile, H_2O	Nitril hydratase in whole cells of *Rhodococcus rhodochrous*	Very high substrate concentration	14
Chiral hydrophobic alcohol	Ketone	Whole cells of *Zygosaccharomyces rouxil*	XAD-resin as substrate reservoir	15
β-Tetralol derivative	*β*-Tetralone derivative	Whole cells of *Trichosporon capitatum*		16
Trimegestone	Corresponding ketone	Baker's yeast	In water	17
Ketodeoxynonulosonic acid	Mannose, pyruvate	Neuraminic acid aldolase	Continuously operated enzyme membrane reactor	18
Fructose	Glucose	Fructose glucose isomerase	Immobilized thermostable enzyme	4
Aspartame	L-phenylalanine-OMe,*Z*-L-asparaginic acid	Thermolysine	Precipitation of the product	19
(*S*)-Phenylethylamine	Racemic phenylethyl amine	Lipase from *Pseudomonas* spec.	Activation of enzyme by lyophilization with fatty acids	20
L-*tert*-Leucine	2-Oxo-3,3-dimethylbuturic acid, NH_4COOH	Leucine dehydrogenase, formate dehydrogenase	NADH regeneration	21

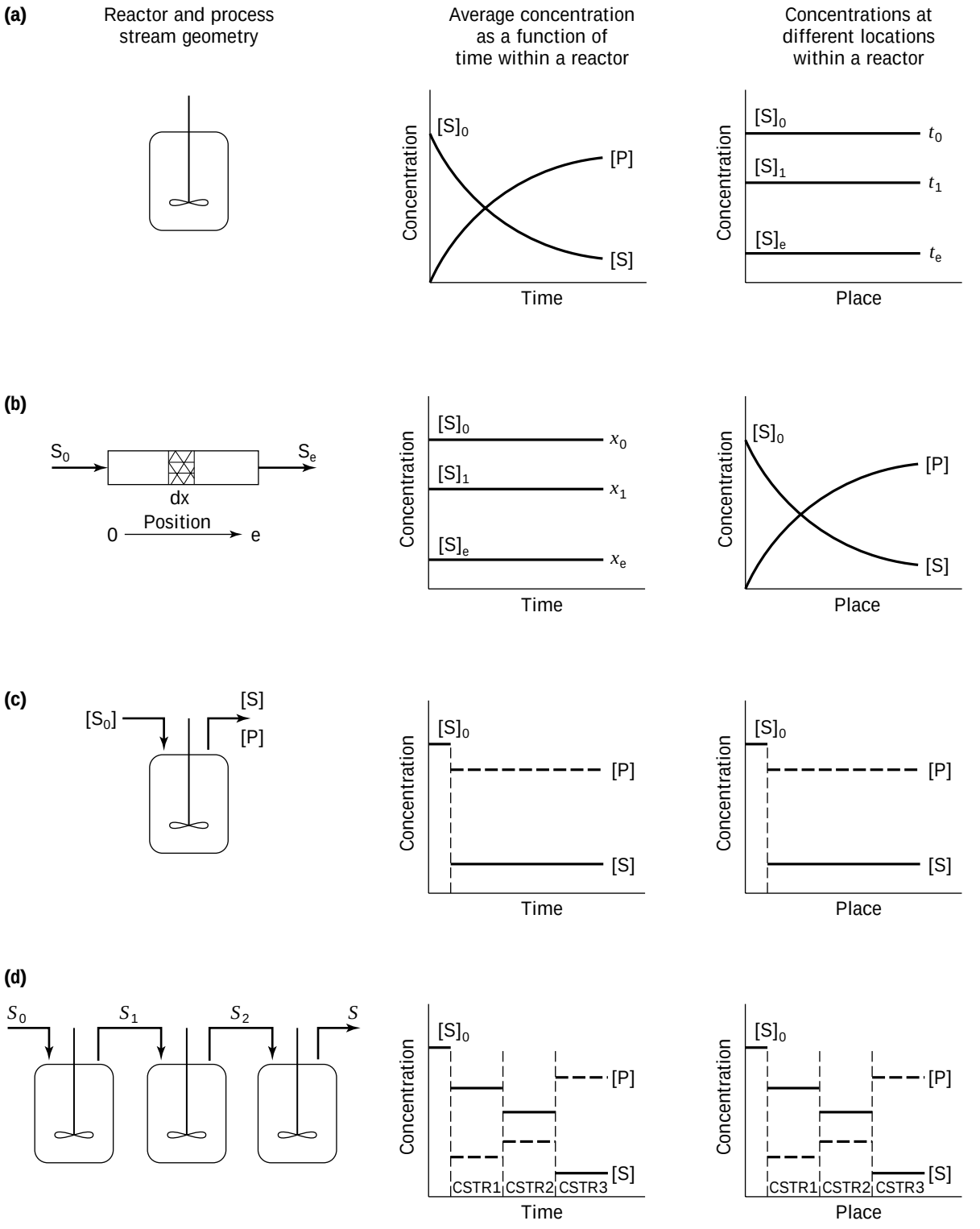


Figure 2. Characterization of different reactors. (**a**) Stirred-tank reactor. (**b**) plug-flow reactor. (**c**) continuously operated stirred-tank reactor. (**d**) Cascade.

reached, the concentrations are independent of time and place. The conversion is controlled by the catalyst concentration and the residence time, τ:

$$\tau = \frac{V}{F}$$

where $\tau[h]$ is the residence time, $V[L]$ is the total reactor volume, and F is the $[L/h]$ substrate feed rate.

One very common variation of the CSTR is the cascade of n CSTRs (Fig. 2d). With an increasing number, n, of reaction vessels, the cascade approximates the plug-flow reactor (23). The product concentration is increased stepwise from vessel to vessel.

Mass Balances

The performance of the different reactor types concerning the reaction can be simulated mathematically. This is also the verification of the kinetic model of the reaction because it should describe the course of the concentration for each compound with only a small error. The main part of the simulation model is the coupled system of differential equations of first order, which are the mass balances of all substrates and products. The change in the concentration of one compound in time and in a volume element (called the accumulation) is the sum of convection, reaction, and diffusion (equation 1).

$$\text{accumulation} = \text{convection} + \text{reaction} + \text{diffusion} \quad (1)$$

The convection term describes the change of the concentration of one compound in the reactor as the difference of the influx into the reactor and the efflux. The reaction term describes, by use of the kinetic model, the change of the concentration of one compound as result of the reaction. The reaction velocity, V_s, is the sum of the individual reaction velocities describing consumption of a substrate or formation of a product. Diffusion occurs only in the case where nonideal mixing is stated.

With regard to the chosen reactor type the mass balance can be simplified by assuming ideal mixing or ideal plug-flow behavior.

Stirred Tank Reactor

Diffusion = 0	At a defined time the concentrations are the same in every volume element.
Convection = 0	There is no influx or efflux of substrates or products to a single volume element in time.

The mass balance is simplified to

$$-\frac{d[\mathrm{S}]}{dt} = v_s \rightarrow t = [\mathrm{S}]_0 \int_0^{\text{desired conversion}} 1/v \, dx \quad (2)$$

Continuously Operated Stirred-Tank Reactor

Diffusion = 0	See previous.
Accumulation = 0	In the steady state the concentrations will not change with time.

The mass balance is simplified to:

$$0 = -\frac{d[\mathrm{S}]}{dt} = \frac{[\mathrm{S}]_0 - [\mathrm{S}]}{\tau} + v_s \quad (3)$$

Plug Flow Reactor

Diffusion = 0	Diffusion is neglected in an ideal plug-flow reactor.
Accumulation = 0	In the steady state the concentrations will not change with time.

The mass balance is the same as for the stirred-tank reactor. In the integrated form the time, t, has to be replaced by the residence time, τ.

Kinetics

The fundamental equation (equation 4) for the description of enzyme kinetics was derived by Michaelis and Menten (24).

$$v = [\mathrm{E}] \frac{A_{\max} \cdot [\mathrm{A}]}{K_m + [\mathrm{A}]} \quad (4)$$

The enzyme activity is expressed as a function of the substrate concentration for an irreversible one-substrate reaction without inhibition. As this is rarely found in nature, equation 5 gives the rate for a reversible reaction A ↔ B. In this form the product inhibition is considered as a competition of A and B for the active sites, as expressed by the two K_m values. This form might be extended to cover substrate surplus inhibition as well as multisubstrate reactions by adding additional terms to the numerator and denumerator (6,25–27).

$$v = [\mathrm{E}] \cdot \frac{\dfrac{A_{\max,A} \cdot [\mathrm{A}]}{K_{m,A}} - \dfrac{A_{\max,B} \cdot [\mathrm{B}]}{K_{m,B}}}{1 + \dfrac{[\mathrm{A}]}{K_{m,A}} + \dfrac{[\mathrm{B}]}{K_{m,B}}} \quad (5)$$

The determination of the kinetic parameters can be performed either by initial reaction velocity measurements or by batch experiments at different substrate concentration, such that substrate limitation is encountered. The kinetic parameters are determined by fitting the kinetic model to the experimental data by use of nonlinear regression; in the case of the batch experiments, numerical integration of the reaction velocity equations is required.

PROCESS DEVELOPMENT, STEP BY STEP

Figure 3 shows different phases and topics to be addressed during development and optimization of a biotransformation. The main focus is on the use of isolated enzymes, but for whole-cell processes a similar scheme must be followed. A more detailed description can be found in Ref. 6. In practice, some of the investigations listed here stepwise will be performed simultaneously. For example, first experiments should hint to possible problems from side reactions. Other factors may be considered as well.

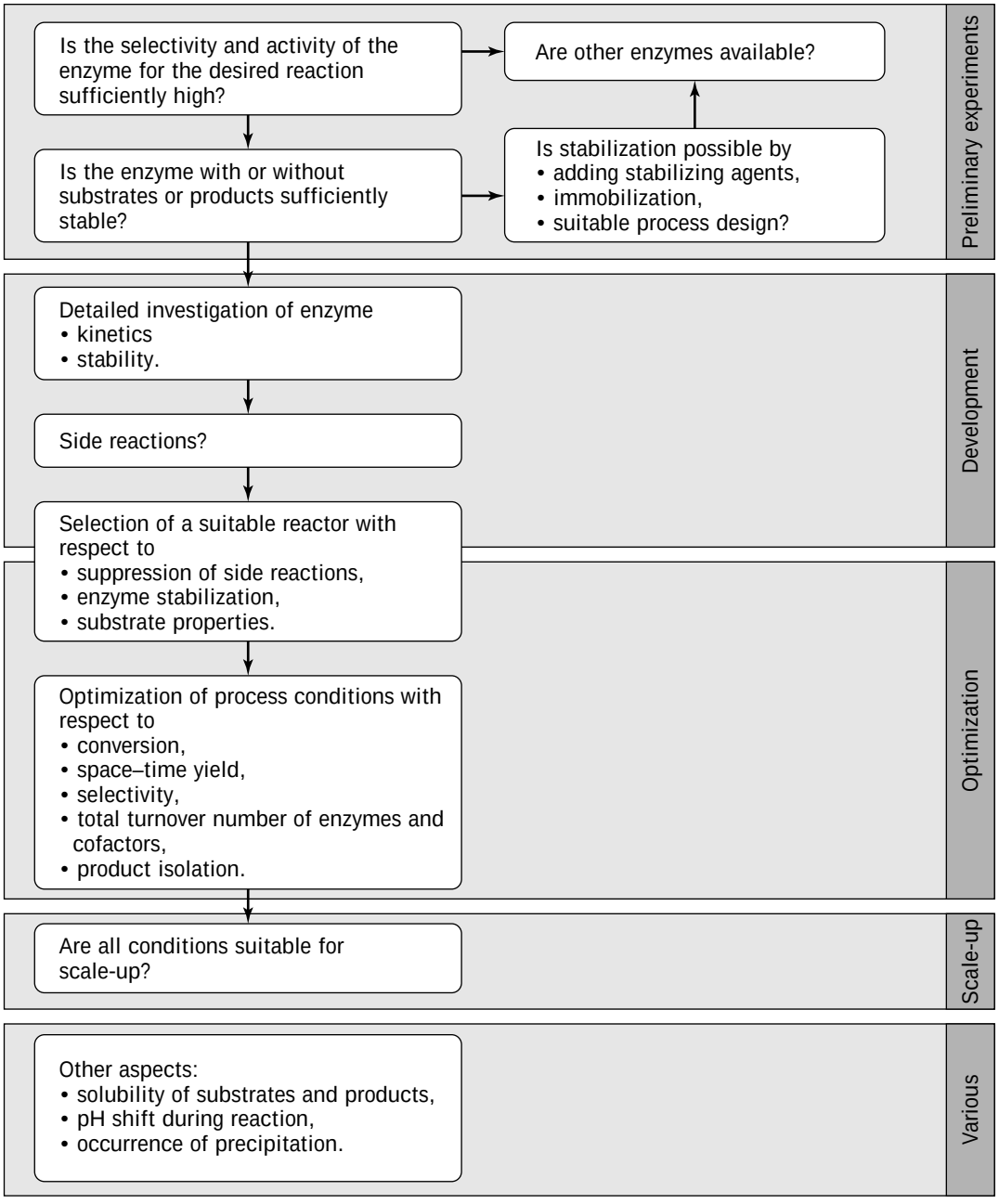


Figure 3. Process development, step by step.

- Buffer compounds may be important for enzyme stability, but must be removed during product isolation. Avoiding them may simplify product isolation and save costs.
- Any reaction exhibiting substrate surplus inhibition should not be carried out in a batch-reactor setup, since this results in longer reaction times due to the high substrate concentration at startup, which lowers the initial reaction velocity. In such a case, a continuously operated stirred-tank reactor that is operated at a high conversion in the steady state is preferred, resulting in a low substrate concentration. Another possibility is the use of a fed batch reactor, which means that the substrate is added in small portions as a function of time, also resulting in a small substrate concentration. One example of such a system is the enzymatic synthesis of L-phenylalanine by reductive amination, which can be found in the literature (25,29).
- If product inhibition occurs, either a stirred-tank reactor, a plug flow reactor, or a cascade of n continuously operated stirred-tank reactors should be chosen. In all these reactors the product concentration increases over time. Another possibility is the use of

a differential reactor with integrated product separation (30). These two possibilities will be discussed in more detail in the section titled "Examples".

- The biocatalyst may be recirculated in and between reactors by various means to enhance catalyst concentration. Whole cells are often concentrated by microfiltration membranes or centrifugation. Isolated enzymes are much smaller, but can be retained by an ultrafiltration membrane. This type of equipment is commercially available for the processing of volumes of a few microliters up to the cubicmeter scale. If the enzyme is immobilized on a support, its size determines the separation step. For example, penicillin amidase is immobilized on a support suitable for separation using a stainless steel mesh in a stirred tank reactor. Also, whole cells might be immobilized on a support, most often simply by adsorption, or by entrapment in alginate gel or similar material.
- The optimization process shown here and discussed later (see "Examples") is based on a kinetic model of the whole reaction system, including side reactions. This model, together with the mass balance for the reactor, can be used for simulation. By a variation of feed concentrations, catalyst concentration, and residence time, conditions can be found that result in high conversion and high space–time yield. When enzyme deactivation is known, the product-specific catalyst consumption may be calculated as well. Alternatively, an empirical optimization can be performed using a simplex method or a generic algorithm. The latter have been successfully applied to fermentation processes (31), but only tested in preliminary studies for optimization of multienzyme systems. These methods are especially powerful for optimization of processes where a direct connection between input and output values cannot be given.
- A discussion of the influence of organic solvents on the solubility of reactants and on enzyme stability can be found in literature (32–36).

Figure 4 summarizes some measures that can be taken to overcome or modulate against enzyme deactivation caused either by thermal condition or by the reaction partners. Whereas in the first case lowering of the temperature, application of stabilizers, or immobilization on a support may be helpful, the deactivation caused by the reaction partners, may be overcome with the help of the tools of reaction engineering. For the deactivation caused by substrates or products, different tools can be used. The different reactor types are discussed in "Classification of Different Reactors".

EXAMPLES

By choosing the appropriate fundamental reactor type on the basis of the individual reaction kinetics, the production of the fine chemicals as given in Table 3 is possible even in processes that exhibit a strong inhibition by one of the reaction compounds.

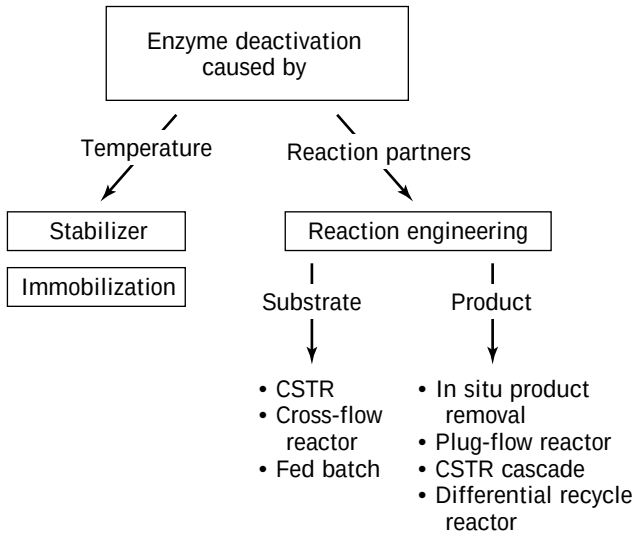


Figure 4. Tools used to overcome enzyme deactivation.

Product Inhibition

Product inhibition is the consequence of the competition of a substrate and a product for complexation in the active site of a catalyst. This kind of inhibition is not only observed with enzymes, but is also found for reactions that are catalyzed by homogeneous soluble and heterogeneous chemocatalysts (28). Whereas the reaction velocity is reduced by product bonding in the case of an irreversible reaction, the situation is more complex for reversible reactions ("principle of the microscopic reversibility"). In terms of the mathematics, product inhibition is expressed as an additional term in the denominator of the Michaelis–Menten equation (equation 6). This form of the equation describes competitive inhibition. An explicit discussion of other types of inhibition can be found in the textbooks of enzyme kinetics (6,25,26). By transformation of equation 6, equation 7 is can be obtained, which is similar in its structure to the common reaction velocity equation (equation 8), derived for chemical systems by Hougen and Watson in 1947 (28,41).

$$v = [\mathrm{E}] \cdot \frac{A_{\max} \cdot [\text{substrate}]}{K_{\mathrm{m}}^{\text{educt}} \cdot \left(1 + \dfrac{[\text{product}]}{K_i^{\text{product}}}\right) + [\text{substrate}]} \tag{6}$$

where v(kat/L) is reaction velocity (1 kat = 1 mol/s), [E](g/L) is enzyme concentration, [substrate](mol/L) is concentration of substrate, $A_{\max}$(kat/g, U/mg) is maximal mass specific activity of the reaction, $K_{\mathrm{m}}^{\text{substrate}}$(mol/L) is Michaelis–Menten constant of substrate, and K_i^{product}(mol/L) is competitive product inhibition.

$$v = [\mathrm{E}] \cdot \frac{\dfrac{A_{\max}}{K_{\mathrm{M}}^{\text{substrate}}} \cdot [\text{substrate}]}{1 + \dfrac{[\text{substrate}]}{K_{\mathrm{M}}^{\text{substrate}}} + \dfrac{[\text{product}]}{K_i^{\text{product}}}} \tag{7}$$

$$v = \frac{(\text{kinetic term}) \cdot (\text{potential term})}{(\text{adsorption term})^n} \tag{8}$$

Table 3. Examples for the Contribution of Biochemical Engineering

Product	Catalyst	Problems	Solution	Reference
(*S*)-1-Phenyl-1,2-ethanediol	Glycerol dehydrogenase	Product inhibition Chemical side reaction	Differential batch with continuous product extraction	30
GDP-mannose	GDP-Man pyrophosphorylase	Product inhibition Chemical complexation of reaction compounds	Cascade of continuously operated stirred-tank reactors	37
CMP-Neu5Ac	CMP-Neu5Ac synthetase	Substrate-surplus inhibition	Fed-batch reactor	38
Oxindol, chiral sulfoxide	Chloroperoxidase	Catalyst deactivation	H_2O_2-dosing with closed-loop control	39
Chiral lactone	Monooxygenase	Oxygen delivery	Bubble-free aeration using silicone tubing	40

where n is the number of centers, which are part of a velocity-limiting elementary reaction.

The decrease of the reaction velocity with increasing product concentration is described by the ratio of K_m to Ki. The relative reaction velocity is plotted as function of the conversion in Figure 5. Three different ratios of K_m to Ki are plotted, which relate to strong (curve A), moderate (curve B), and weak (curve C) product inhibition.

In 1985, Lee and Whitesides suggested, using the example of an oxidoreductase-catalyzed alcohol oxidation, that the feasibility of a process can be judged by the ratio of $K_m^{substrate}$ to $K_i^{product}$ (42):

- If $K_m^{substrate}$: $K^{product} \leq 1$, it is, in principle, possible that the reaction can proceed with practicable conversions (Fig. 5, curves B and C).
- If K_m^{educt}: $K_i^{product} > 1$, the reaction is unusable with regard to preparative syntheses (Figure 5: curve A).

This clearly is a simplified view. It is also important to note, that the kinetic parameters of a given reaction cannot be changed.

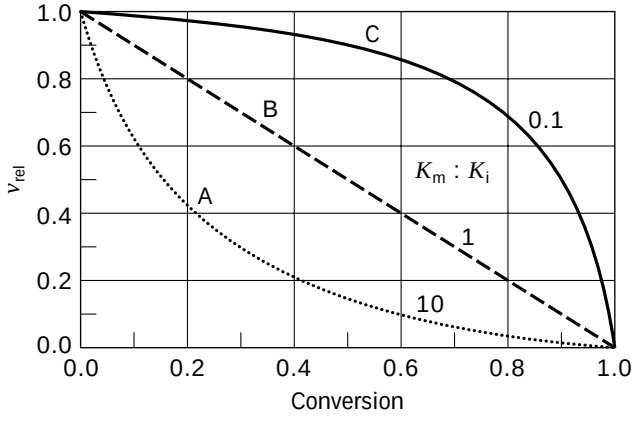


Figure 5. Influence of product inhibition on reaction velocity as function of conversion. Curve A, strong product inhibition with the ratio of Michaelis–Menten constant, K_m, and the product inhibition constant, K_i; K_m:K_i = 10. Curve B, moderate product inhibition with K_m:K_i = 1. Curve C, weak product inhibition with K_m:K_i = 0.1.

Overcoming Product Inhibition by Product Extraction

Chiral diols are important building blocks for the synthesis of chiral ligands, agrochemicals, and pharmaceutical intermediates (43–45). Using the example of the oxidative resolution of 1-phenyl-1,2-ethanediol, it is demonstrated that product inhibition can be overcome in the case of hydrophobic products. The enantioselective oxidation of the (*R*)-enantiomer to ω-hydroxyacetophenone (1-oxo-1-phenyl-2-hydroxyethane) is catalyzed by glycerol dehydrogenase (GDH) from *Enterobacter aerogenes* [EC 1.1.1.6] (46). Because GDH needs β-nicotinamide adenine dinucleotide (NAD^+) as a cofactor, a cofactor regeneration carried out by lactate dehydrogenase (LDH) [EC 1.1.1.27] was incorporated, as depicted in Figure 6. LDH reduces pyruvate to (L)-lactate enantioselectively. The cofactor regeneration must proceed faster than the racemic resolution in order not to become the limiting step. The remaining (*S*)-enantiomer of the diol is the desired product for further synthesis.

Detailed investigation of the reaction conditions and kinetics revealed strong product inhibition by the hydroxyketone, as well as by the NADH (30). This resulted in a ratio of $K_m^{substrate}$: $K_i^{product}$ = 43.2 for the ω-hydroxyacetophenone. The inhibition by NADH was not taken into account, because NADH is converted in situ to NAD^+ by LDH. Utilizing the above-determined criteria, the reaction as investigated here would never proceed efficiently. Additionally, it was observed that the hydroxyketone reacted in a chemical side reaction with NAD^+. Thus, ω-hydroxyacetophenone not only reduced the GDH activity due to its inhibition, it furthermore reduced the available NAD^+ concentration. Therefore, it is very important that the ω-hydroxyacetophenone concentration is kept as low as possible. One possibility is an integrated extraction of the inhibiting compound. Best results were obtained using hexane as the extracting solvent. In contrast to phenylethanediol, which was not extracted, up to 98% of the hydroxyketone was extracted after eight extraction steps.

The flow sheet of the two-phase membrane reactor is shown in Figure 7. The reactor consists of three different loops. The membrane reactor (loop I) integrates the stirred-reaction vessel and the membrane module (48,49). A hydrophilic ultrafiltration membrane (YM 10, cutoff 10,000 Da, Amicon GmbH, Witten, Germany) is used to retain the glycerol dehydrogenase and the lactate dehy-

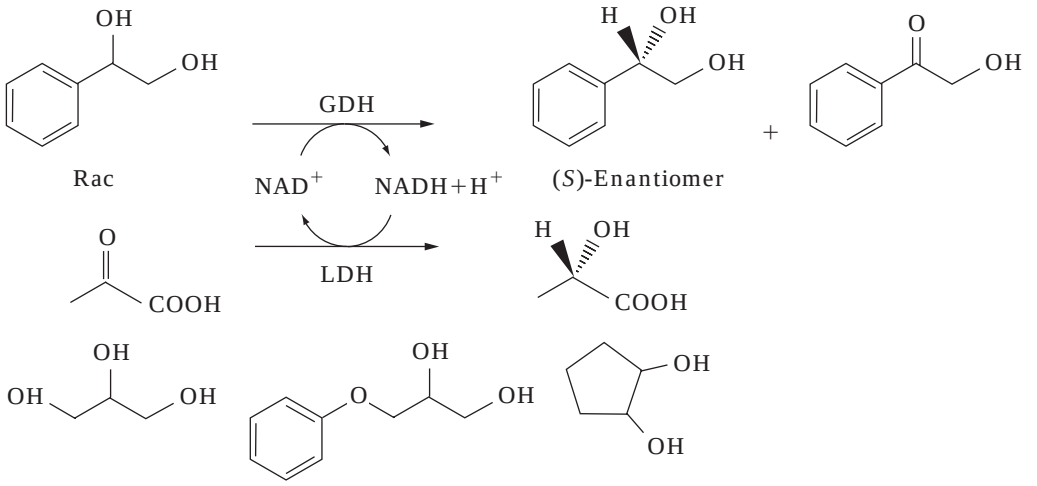


Figure 6. Enzymatic resolution of racemic 1-phenylethan-1,2-diol. GDH, glycerin dehydrogenase; LDH, lactate dehydrogenase. Relative activities: glycerol, 100%, 3-phenoxypropan-1,2-diol, 31%, 1-phenylethan-1,2-diol, 8%, *trans*-cyclopentan-1,2-diol, 3%.

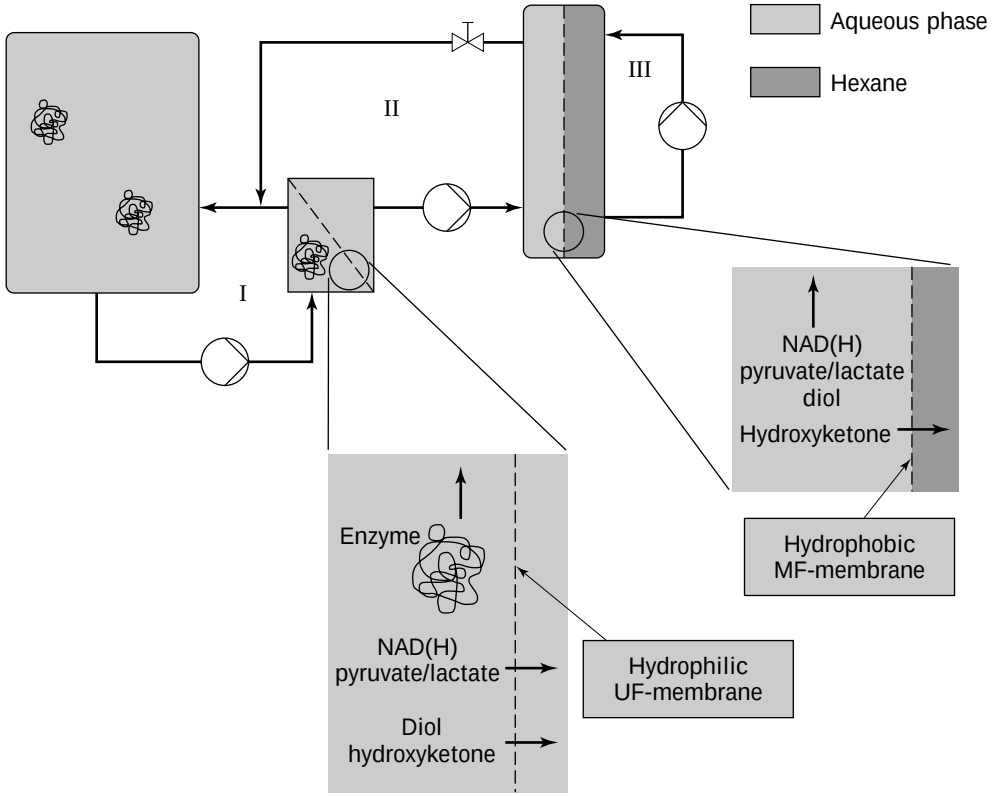


Figure 7. Flowsheet of the two-phase membrane reactor; I, membrane reactor loop; II, extraction loop; aqueous phase; III extraction loop, organic phase. For details refer to Refs 47 and 54.

drogenase. The consequence is the decoupling of the residence times of the catalysts and the reactants. Prior to the membrane reactor, an extraction module is positioned in a closed-loop arrangement (loop II). The aqueous phase is continuously extracted by means of liquid–liquid extraction using a hollow-fiber module (Liqui-Cel® 5PCM-106, Hoechst, Frankfurt, Germany) consisting of Celgard® microporous, hydrophobic hollow-fiber membranes (50). The module was circulated countercurrently by a hexane stream. The best ratio of the flow rates of the aqueous phase to the organic phase was determined to be 1:6. To reduce the amount of organic solvent used, a distillation

(not shown) is incorporated in the extraction loop III. The oxidation product ω-hydroxyacetophenone is collected in the bottom of the distillation. The entire reactor behaves as a batch reactor with recirculation, meaning that only a low conversion per cycle is reached, corresponding to a small amount of oxidation product, which is directly extracted (30). The consequence is a continuous low concentration of ω-hydroxyacetophenone in the membrane reactor. This reactor represents a modification of the continuously operated bimembrane reactor that was constructed for the conversion of sparingly soluble substrates (47,51,52). In Figure 8, two experiments are compared, one with the integrated extraction and the other without. When no extraction is incorporated, an enantiomeric excess of only 58% and a conversion of 37% was reached after 170 h. The reason for this unsatisfactory result was the accumulation of the reaction product, which lowers the reaction velocity of the GDH and also reduces the concentration of the NAD^+ by the chemical side reaction. With integrated product extraction a conversion of 50% and an enantiomeric excess >99% is reached after 64 h.

This process example demonstrates that by means of reaction engineering, kinetically unfavorable reaction can still be used on a preparative scale. By application of the integrated extraction, a transition from curve A to curve C in Figure 5 is reached.

Overcoming Product Inhibition by a CSTR Cascade

D-Mannose can be found as building block of numerous glycoproteins and glycolipids, for example, in the common core structure of N-glycans (53). For the enzymatic in vitro synthesis of oligosaccharides, guanosine-5′-diphospho-α-D-mannose (GDP-Man) is used as donor for mannosyltransferases. Recombinant mannosyltransferases have been used for the synthesis of the trisaccharide Manβ1-4GlcNAcβ1-4GlcNAc (54). The enzymatic synthesis of GDP-Man starting from mannose-1-phosphate (Man-1-P) and guanosine-5′-triphosphate (GTP) using a recombinant GDP-Man pyrophosphorylase [E.C. 2.7.7.13] has been recently described (55) (Fig. 9). Both reaction products show a very strong inhibition of GDP-Man pyrophosphorylase. The inhibition by pyrophosphate can be overcome by cleaving it to phosphate using an inorganic pyrophosphatase [E.C. 3.6.11]. The product inhibition by the product GDP-Man itself can be reduced in a CSTR-cascade, as shown below. Additionally, the cofactor Mg^{2+}, required by the GDP-Man pyrophosphorylase, is complexed by the substrate GTP. Its influence on enzyme activity and the preferable use of a continuously operated reactor to establish a constant Mg^{2+} to GTP ratio, together with a complete set of kinetic data for the pyrophosphorylase, has been recently published (43). For the ratio $K_m^{\text{Man-1-P}}/K_i^{\text{GDP-MAN}}$, a value of 4.4 is obtained. According to the classification outlined in the section "The Problem of Product Inhibition", this reaction would be of only limited use. Unfortunately, the substrates and products are quite similar with respect to size, hydrophilicity, or charge, prohibiting a selective removal of the inhibiting product. However, simulations on the basis of the kinetic model and various reactor configurations revealed that a cascade of two CSTR in series reduces the required amount of enzyme by >50% (23). A two-stage cascade is a good compromise: plug-flow behavior is approximated to a large extent, and the experimental setup does not require too much effort (7). Figure 10 compares the conversion of Man-1-P as function of the product of enzyme concentration and residence time for a single CSTR and a two-stage CSTR cascade. In a two-stage cascade 98% conversion is reached for 4.5 (U/mL)min, whereas a single CSTR requires 9.0 (U/mL)min. That means that at a constant enzyme concentration the residence time can be cut in half, resulting in a doubling of the space–time yield. When the enzyme deactivation is only a function of time, by an increase of the space–time yield at

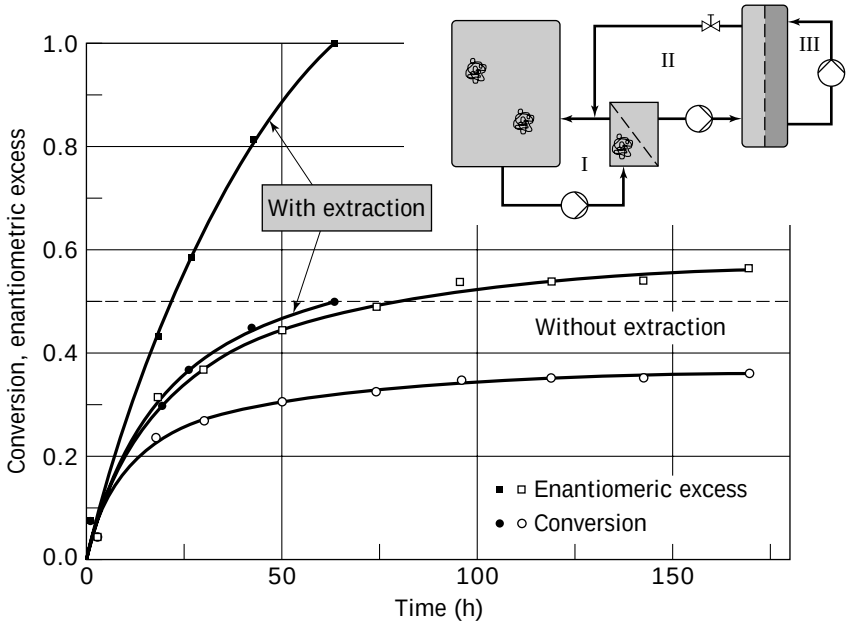


Figure 8. Increase of enantiomeric excess and conversion for the kinetic resolution of 1-phenyl-1,2-ethanediol due to product extraction. Residence time in the membrane reactor: 15 min; pressure difference: 2 bar; total volume 500 mL; 100 mM glycine/KOH buffer; pH 9; T = 25 °C; 0.054 U/mL GDH; 0.6 U/mL LDH; 200 mM ammonia; 40 mM (*R*,*S*) 1-phenyl-1,2-ethanediol; 2 mM NAD^+; 25 mM pyruvate.

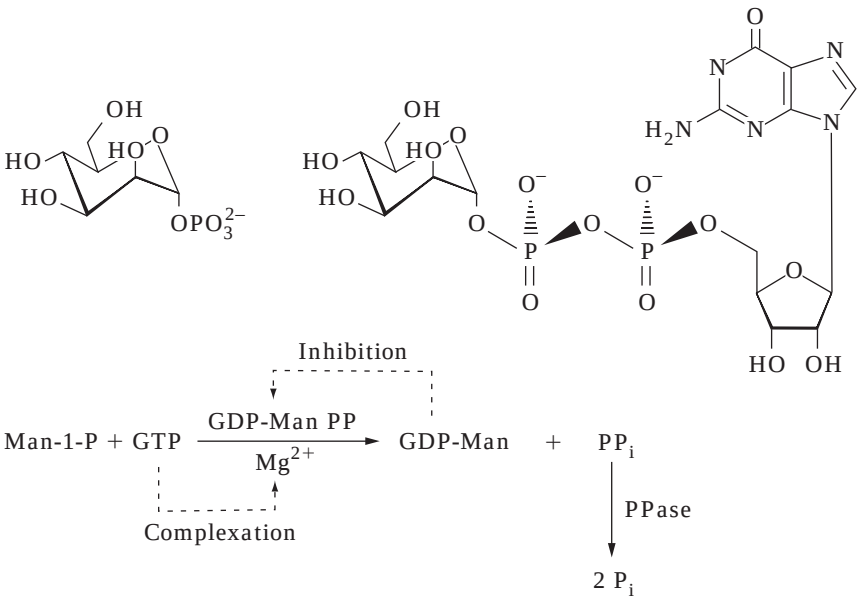


Figure 9. Enzymatic synthesis of GDP-Man starting from Man-1-phosphate and GTP; GDP-Man PP = GDP-Man pyrophosphorylase, PPase = inorganic pyrophosphate.

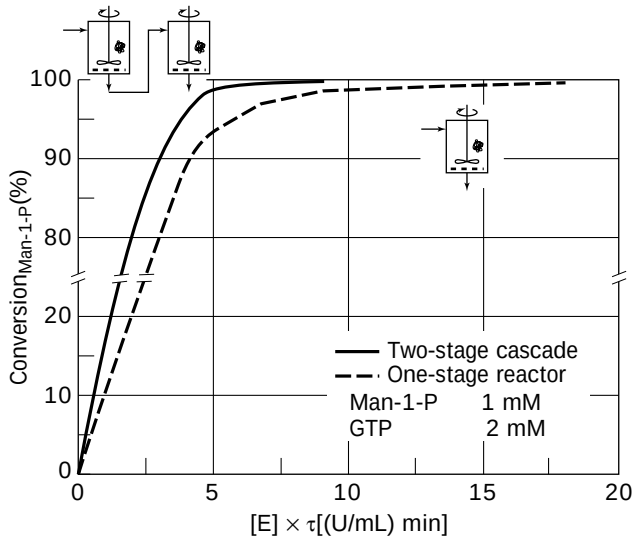


Figure 10. Simulation of GDP-Man synthesis in a one-stage CSTR (dotted line) and a two-stage CSTR cascade (solid line), Man-1-P, 1 mM; GTP, 2 mM; 25 °C; pH 8.0.

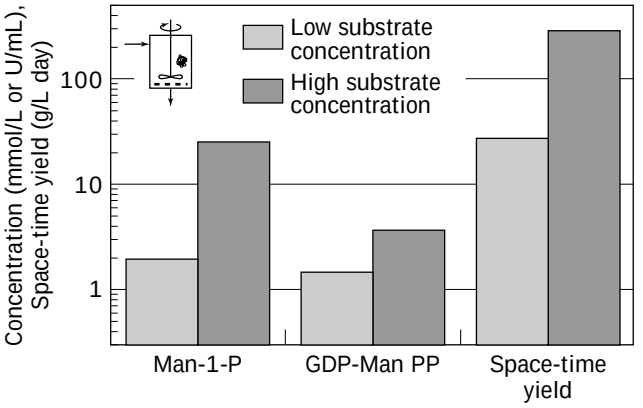


Figure 11. Improvement of space–time yield of GDP-Man by higher substrate and enzyme concentration in a one-stage CSTR; [GTP] 2.5 mmol/L or 25 mmol/L, [Mg^{2+}] 1 mmol/L or 9 mmol/L, residence time 1 h; conversion >98%. Black bars high substrate concentration, gray bars, low substrate concentration.

constant enzyme concentration, the product-specific enzyme consumption can be decreased. The experimental verification revealed good agreement between calculated and experimental values (37). At a conversion at 98%, a space–time yield of 28 g/(L day) and an enzyme consumption of 900 U/kg$_{product}$ was obtained. For a single CSTR, an enzyme consumption of 2,000 U/kg$_{product}$ was reached due to the higher enzyme concentration required to reach the desired conversion under otherwise identical conditions. The relatively low space–time yield is caused by the low substrate concentrations. Complex interactions between K_m values of the substrates and the K_i of GDP-Man require simulation to show that the space–time yield can be increased 10-fold by a corresponding increase of the substrate concentration (Fig. 11). In order to keep a high conversion at >98%, the enzyme concentration must be increased at the same time, but only 2.5-fold. This clearly demonstrates the potential of process optimization based on a model and a complete set of kinetic data.

BIBLIOGRAPHY

1. K. Faber, *Biotransformations in Organic Chemistry*, Springer-Verlag, Berlin, 1995.
2. A. Tanaka, T. Tosa, and T. Kobayashi eds., *Industrial Application of Immobilized Biocatalysts*, Marcel Decker, New York, 1993.
3. K. Drauz and H. Waldmann eds., *Enzyme Catalysis in Organic Synthesis*, VCH, Weinheim, 1995.

4. T. Godfrey and S. West eds., *Industrial Enzymology—Application of Enzymes in Industry*, 2nd ed., Macmillan, London, 1996.
5. R.A. Sheldon, *Chirotechnology: Industrial Synthesis of Optically Active Compounds*, Marcel Decker, New York, 1993.
6. M. Biselli, U. Kragl and C. Wandrey, in K. Drauz and H. Waldmann eds., *Enzyme Catalysis in Organic Synthesis*, VCH, Weinheim, 1995, pp. 89–155.
7. O. Levenspiel, *Chemical Reaction Engineering*, J Wiley, New York, 1972.
8. J.E. Bailey and D.F. Ollis, *Biochemical Engineering Fundamentals*, McGraw-Hill, New York, 1986.
9. D. Rozzell and F. Wagner, *Biocatalytic Production of Amino Acids and Derivatives*, Hanser, München, 1992.
10. W.S. Kelley, *Bio / Technology* **14**, 28–31 (1996).
11. A. Grobiollot, D.K. Boadi, D. Poncelet, and R. Neufeld, *Crit. Rev. Biotechnol.* **14**, 75–107 (1994).
12. S. Norton and J.C. Vuillemard, *Crit. Rev. Biotechnol.* **14**, 193–224 (1994).
13. J.A.M. de Bont, J. Visser, B. Mattiasson, and J. Tramper, eds., *Physiology of Immobilized Cells*, Elsevier, Amsterdam, 1990.
14. T. Nagasawa and H. Yamada, *Trends Biotechnol.* **7**, 153–158 (1985).
15. B.A. Anderson, M.M. Hansen, A.R. Harkness, C.L. Henry, J.T. Vicenzi, and M.J. Zmijewski, *J. Am. Chem. Soc.* **117**, 12358–12359 (1995).
16. J. Reddy, D. Tschaen, Y.-J. Shi, V. Pecore, L. Katz, R. Greasham, and M. Chartrain, *J. Ferment. Bioeng.* **81**, 304–309 (1986).
17. V. Crocq, C. Masson, J. Winter, C. Richard, G. Lemaitre, J. Lenay, M. Vivat, J. Buendia, and D. Prat, *Org. Process Res. Dev.* **1**, 2–13 (1982).
18. C. Salagnad, A. Gödde, B. Ernst, and U. Kragl, *Biotechnol. Progr.* **13**, 810–813 (1997).
19. V.H.M. Elferink in *Reprints of Chiral USA '96, Boston*, Spring Innovations, London, 1996, pp. 79–80.
20. F. Balkenhol, K. Ditrich, B. Hauer, and W. Ladner, *J. Prakt. Chem.* **339**, 381–384 (1997).
21. A.S. Bommarius, M. Schwarm, K. Stingl, M. Kottenhahn, K. Huthmacher, and K. Drauz, *Tetrahedron: Asymmetry* **6**, 2851–2888 (1995).
22. T. Zelinski and H. Waldmann, *Angew. Chem.* **109**, 746–748, 1997, and literature within.
23. U. Kragl, D. Vasic-Racki, and C. Wandrey, *Bioprocess Eng.* **14**, 291–297 (1996).
24. L. Michaelis and M.L. Menten, *Biochem. Z.* **49**, 333–369 (1913).
25. A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, Butterworth, London, 1995.
26. I.H. Segel, *Enzyme Kinetics*, J Wiley, New York, 1975.
27. H. Bisswanger, *Theorie und Methoden der Enzymkinetik*, Verlag Chemie, Weinheim, 1979.
28. M. Baerns, H. Hofmann, and A. Renken, *Chemische Reaktionstechnik*, Thieme Verlag, Stuttgart, 1992.
29. E. Schmidt, E. Fiolitakis, and C. Wandrey, *Ann. N. Y. Acad. Sci.* **501**, 434–437 (1988).
30. A. Liese, M. Karutz, J. Kamphuis, C. Wandrey, and U. Kragl, *Biotechnol. Bioeng.* **51**, 544–550 (1996).
31. D. Weuster-Botz and C. Wandrey, *Process Biochem.* **30**, 563–571 (1995).
32. A.M.P. Koskinen and A.M. Klibanov eds., *Enzymatic Reactions in Organic Media*, Chapman & Hall, London, 1996.
33. P. Adlerkreutz, *Biocatal. Biotrans.* **14**, 1–30 (1996).
34. L. Kvittingen, *Tetrahedron* **50**; 8253–8274 (1994).
35. G. Carrea, G. Ottolina, and S. Riva, *Trends Biotechnol.* **13**, 63–70 (1995).
36. A.L. Gutmann and M. Shapira, *Adv. Biochem. Eng. Biotechnol.* **52**, 87–128 (1995).
37. S. Fey, L. Elling, and U. Kragl, *Carbohydr. Res.* **305**, 475–481 (1997).
38. U. Kragl, T. Klein, D. Vasic-Racki, M. Kittelmann, O. Ghisalba, and C. Wandrey, *Ann. N. Y. Acad. Sci.* **709**, 577–583 (1996).
39. K. Seelbach, M.P.J. van Deurzen, F. van Rantwijk, R.A. Sheldon, and U. Kragl, *Biotechnol. Bioeng.* **55**, 283–288 (1997).
40. S. Rissom, U. Schwarz-Linek, M. Vogel, V.I. Tishkov, and U. Kragl, *Tetrahedron: Asymm.* **8**, 2523–2526 (1997).
41. O.A. Hougen and K.M. Watson, *Chemical Process Principles*, J Wiley, New York, 1947.
42. L.G. Lee and G.M. Whitesides, *J. Am. Chem. Soc.* **107**, 6999–7008 (1985).
43. J. Hasegawa, M. Ogura, S. Tsuda, S. Maemoto, H. Kutsuki, and T. Ohashi, *Agric. Biol. Chem.* **54**, 1819–1827 (1990).
44. A. Bosetti, D. Bianchi, P. Cest, P. Golini, and S. Spezia, *J. Chem. Soc. Perkin Trans.* **1**, 2395–2398 (1992).
45. P.E. Kündig, C. Dupre, B. Bourdin, A.C. Jr., and D. Pons, *Helv. Chim. Acta* **77**, 421–428 (1994).
46. E.C. Webb, *Enzyme Nomenclature*, Academic, New York, 1992.
47. W. Kruse, W. Hummel, and U. Kragl, *Recl. Trav. Chim. Pays-Bas* **115**, 239–243 (1996).
48. M.-R. Kula and C. Wandrey, *Methods Enzymol.* **136**, 9–21 (1987).
49. R. Wichmann, C. Wandrey, A.F. Bückmann, and M.-R. Kula, *Biotechnol. Bioeng.* **23**, 2789–2802 (1981).
50. U. Daiminger, P. Plucinski, and W. Nitsch, *Chem. Ing. Tech.* **67**, 217–220 (1995).
51. U. Kragl, W. Kruse, W. Hummel, and C. Wandrey, *Biotechnol. Bioeng.* **52**, 309–319 (1996).
52. S. Kise and M. Hayashida, *J. Biotechnol.* **14**, 221–228 (1990).
53. M. Fukuda and O. Hindsgaul eds., *Molecular Glycobiology*, Oxford University Press, London, 1994.
54. S.L. Flitsch, D.M. Goodridge, B. Builbert, L. Revers, M.C. Webberly, and I.B.H. Wilson, *Bioorg. Med. Chem.* **2**, 1243–1250 (1994).
55. L. Elling, J.E. Ritter, and S. Verseck, *Glycobiology* **6**, 591–597 (1996).

BREVIBACTERIA. See CORYNEBACTERIA, BREVIBACTERIA.

CELL CYCLE

DAVID R. LLOYD
MOHAMED AL-RUBEAI
University of Birmingham
Birmingham, England

KEY WORDS

Apoptosis
Cell cycle
Division
Flow cytometry
Growth
Productivity
Proliferation

OUTLINE

INTRODUCTION

The cell cycle is the name given to the process by which a cell "matures," synthesizes DNA, and divides to form daughter cells. Thus the cell cycle is a fundamental process with analogous mechanisms found in all cells, from the most primitive bacterium to higher animals and plants, from the unicellular to the most complex multicellular organism. Here we describe the mammalian cell cycle, highlight its importance in biotechnology, and describe some cell cycle applications in animal cell biotechnology.

The cell cycle model (G_1-event model) currently accepted by authors of the major cell and molecular biology books (1–4) describes the cell cycle as discrete phases controlled by cell cycle proteins interacting with each other and directing cellular events including DNA synthesis and cell division. However, the validity of this model and the interpretation of many experiments that underpin the current established view of the cell cycle and cell cycle regulation have been questioned (5,6). The proposed alternative, the continuum model, suggests that the cell cycle does not exist as currently recognized but takes the form of a continuum from one cell division to the next, rather than discrete phases with discrete groups of regulatory proteins. A hypothetical initiator of DNA replication is synthesized by the cell during all phases of the cell cycle, and its concentration reaches a critical level that initiates DNA synthesis. Thus the cell can control its rate of division by regulating the amount of initiator. Nevertheless, until discussion of these competing models is concluded, with perhaps more work to determine the correctness and applicability of the continuum model, it is only reasonable to accept the weight of evidence for the established view of the cell cycle and its regulation. Thus the established model is summarized here and is the basis for the process biotechnology applications that are described.

The pathways of cell cycle regulation are, unsurprisingly perhaps, shared with another important cell process: apoptosis, or "programmed cell death," which is, in effect, the opposite of the cell proliferation cycle. Because the cell cycle and apoptosis share some common pathways, regulatory imbalance can lead to cancer in multicellular organisms (7). Regulation of the cell cycle is also crucial in embryology and ontogeny, where features such as limb formation, tissue differentiation, and growth are a function of the balance between proliferation and apoptosis. Therefore, much of the cell cycle (and apoptosis) literature is from the field of biomedical science, particularly cancer and anticancer chemotherapy research. General texts, as well as much of what is written here, provide a consensus of information describing a hypothetical typical cell and thus, by implication at least, describe what is true of all eukaryotic cells.

THE CELL CYCLE

Overview

The cell cycle and its control mechanisms are highly conserved, so most of what is described in general terms for a "typical mammalian cell" is also true for other animal cells, such as insect and plant cells, although there is the obvious exception of the cell wall and its effect on cell division. The

cell cycle and its control in yeast are essentially similar to but much simpler than the analogous mechanisms in higher eukaryotes; where one protein pair exerts a particular control function in a yeast, several may be required for the same function in an animal cell. As a result of this less complex system and the fortuitous discovery of cell cycle control mutants (cdc mutants) in which the cell cycle is inhibited at specific stages, many early discoveries about cell cycle control were made using the budding yeast *Saccharomyces cerevisiae* and the fission yeast *Schizosaccharomyces pombe*. Subsequent research has shown homology between cell cycle control proteins in yeast and those in mammalian cells.

However, the bacterial cell cycle is markedly different from the eukaryotic cell cycle, and its control mechanisms are (to date) less well understood than the analogous processes in eukaryotes. Given a hypothetical typical bacterial cell, its structure and thus the function of its cell cycle are very different from those of a eukaryotic cell. Bacteria, like all cells, are enclosed in a cytoplasmic membrane, but in addition, they have a cell wall outside the cytoplasmic membrane. Within the cell there are no organelles, only free ribosomes and the nucleus, which is actually an area that contains DNA but is not segregated from the cytoplasm by a membrane. The entire genome is contained in a single, circular, double-stranded DNA chromosome; one strand is a closed circle and the other is open. There are no histone or similar DNA-associated proteins; instead there is a relatively high proportion of small, basic polyamines whose function may be to counteract the acidity of the DNA. The "joint" in the closed DNA strand, the origin of replication, is attached to a mesosome, an invagination of the cytoplasmic membrane. When DNA replication starts, the 5′ end of the open strand attaches to the cytoplasmic membrane and starts to form a second mesosome; meanwhile, the DNA unwinds in one direction from the origin of replication and splits into two single strands to form a replication fork. As the replication fork progresses along the chromosome, complementary strands of DNA are synthesized until two daughter chromosomes of double-stranded DNA are formed, each with an origin of replication attached to its mesosome. While the chromosome replicates, the cytoplasmic membrane and cell wall are synthesized in the region between the two mesosomes, enlarging the cell. After chromosome replication, a septum forms across the middle of this region, finally separating the two daughter cells. However, it takes about 40 minutes to duplicate the single chromosome; hence, to allow doubling times as low as 20 minutes, each daughter chromosome must start to replicate in exactly the same way before the parent has finished replication. Thus instead of one there will be three origins of replication operating at once. In addition, when the doubling time of a culture is less than about 40 minutes, DNA synthesis is continuous, so it may be that DNA replication and cell division cycles are independent in bacteria (8). However, despite the essentially continuous DNA synthesis and cell division shown in the fast-growing cultures likely to be used in process biotechnology, it has recently been shown (9) that slowly growing *Escherichia coli* cells (doubling time up to 5 hours) display cell cycle phases analogous to eukaryotic G_1, S, and G_2 phases, *vide infra*, followed by cell division. Furthermore it was shown that initiation of DNA replication was related to cell mass. Two genes, *seqA* and *dnaA*, were shown to be involved in the regulation of replication initiation; gene mutants exhibited decreased and increased mass, respectively, at initiation of replication. Recent work with *Caulobacter* (10) described a protein, CtrA, that has sequence homology with response regulatory proteins in other bacteria including *E. coli* and is involved in cell cycle control including cell division, and cell cycle–specific transcription. It was proposed that CtrA is part of a phosphorelay signal transduction system responsible for transcriptional-level control of bacterial cell cycle events.

The mammalian cell cycle is best described by illustration (Fig. 1a). Starting from a recently divided daughter cell, untransformed, or "normal," cells may enter gap phase 0, or G_0. This is an offshoot from the cell cycle in which cells may remain viable and nondividing but retain the potential to divide given the appropriate stimulus. A cell, depending upon its type, can remain out of the cell

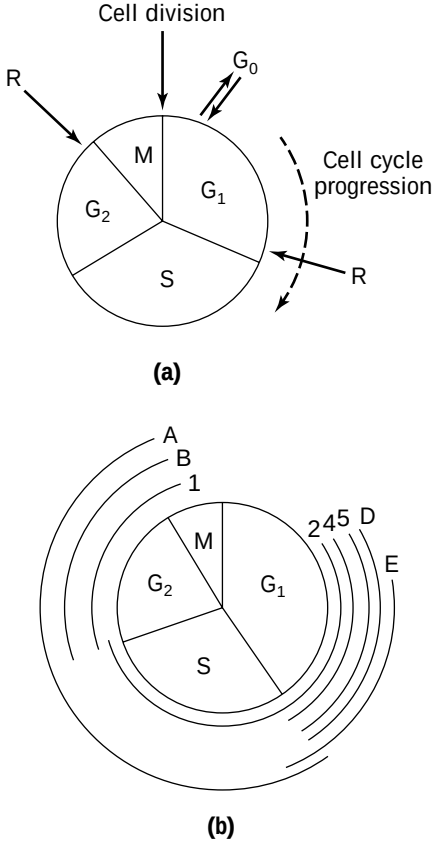


Figure 1. (**a**) A simple diagrammatic representation of the cell cycle. G_0, gap phase 0; G_1, gap phase 1; S, synthesis phase, in which the DNA and cell cytoplasmic components are duplicated; G_2, gap phase 2; M, cell division by mitosis (or meiosis); R, restriction point (check point). (**b**) A simple diagrammatic representation of cyclin and cyclin-dependent kinase activation during cell cycle traverse. A, cyclin A; B, cyclin B; D, cyclin D; E, cyclin E; 1, cyclin-dependent kinase 1 (homologous to cdc2); 2, cyclin-dependent kinase 2; 4, cyclin-dependent kinase 4; 5, cyclin-dependent kinase 5.

cycle, in G_0, for extended periods but can return to gap phase 1, or G_1, due to cell maturation after a certain time has elapsed, after a critical mass/size is reached, or as a response to an external stimulus. In contrast, cells from continuous, sometimes called *established* or *immortal*, cell lines and transformed cells (which includes most animal cells used in biotechnology) do not enter G_0 after division but restart their cell cycle after mitosis by progressing directly to G_1. During G_1, cells prepare to duplicate themselves. When a cell enters the synthesis, or S phase, it starts to precisely and accurately duplicate its genome by synthesizing DNA and the associated nuclear protein components, so that daughter cells carry a genome identical to each other and to the parent. During this time, ribosomes, mitochondria, endoplasmic reticulum, and all the other subcellular organelles are also replicated in preparation for cell division. Thus by the end of S phase, a normally diploid ($2n$) cell will be tetraploid ($4n$), containing enough DNA (and cellular components) for two cells. After S phase, the cycling cell enters the second gap, or G_2, phase. During G_2 the cell synthesizes the proteins required to direct chromosome segregation and cell division in the following mitosis, or M phase. After mitotic division the daughter cells return to G_1 at the beginning of a new cell cycle or exit the cycle to G_0. In a typical cell cycle of 24 hours' duration, G_1 would last 10 hours, S phase 9 hours, G_2 would last 4 hours, and mitosis 1 hour. However, the duration of the cell cycle varies both between cell lines and between individual cells within the same culture. Variation in cycle time is usually due either to cells exiting the cycle and arresting in G_0 or to a variation in the duration of G_1. It has been suggested that the intercellular variability of the postmitotic early G_1 period is smaller than that seen with initiation of DNA replication (presynthetic late G_1) (11). It is also postulated, by the same author, that cells that make the "yes or no" decision in the postmitotic early G_1 period about whether to proceed through the cell cycle also have the capacity to decide, in the presynthetic late G_1 period, when the cell will enter S phase. Heterogeneity in the duration of G_1 is the main reason why a synchronized culture loses synchrony with time to the extent that the culture is completely asynchronous within three or four cell cycles.

This apparently simple cell cycle is subject to complex control mechanisms. Some of the molecular pathways have been understood only in recent years and there may be others that have not yet been described. Mammalian cells periodically transcribe cell cycle genes that encode proteins directly involved in cell cycle regulation. These include the cyclins and the cyclin-dependent kinase (Cdk) families. The cyclin-dependent kinases are phosphorylating enzymes that are inactive when not associated with an appropriate cyclin; it is the associated cyclin that determines the substrate specificity of the complex. The activity of this heterodimeric complex is then regulated by phosphorylation/dephosphorylation of specific activation and/or inhibition sites, as well as by further binding with inhibitory proteins. Many of the cyclins, cyclin-dependent kinases, and related regulatory proteins are expressed or form active complexes only during particular cell cycle phases. Cdk4–cyclin D1 drives cells through mid G_1, Cdk2–cyclin E drives late G_1, Cdk2–cyclin A controls entry into S phase, and Cdk1–cyclin B drives the G_2–M transition; their phase-related activities are summarized in Figure 1b and Table 1. Others, including Cdk4 (13,14) are expressed throughout the cell cycle. The noncyclic proteins exert their cell cycle–controlling effects by association with phase-specific cyclins or cyclin-dependent kinases, association with Cdk-cyclin inhibitors, and specific amino acid phosphorylation/dephosphorylation to form active complexes in specific phases. A family of genes encoding proteins related to the retinoblastoma susceptibility gene, *RB*, and genes for transcription factors including *BYMB*, *E2F1*, and *JUN* are all induced in late G_1 or early S phase and have been shown to interact with cell cycle regulatory complexes. In addition, there are restriction points in the cell cycle that, once past, commit the cell to continuing the division cycle. There are also check points where, in effect, the cell ensures that the previous phase has been completed properly before proceeding or that there has been no DNA damage. To continue to mitosis before the genome is properly duplicated, for example, would be catastrophic for the cell; one if not both daughters would not be viable. It has been suggested that induction of the p53 tumor suppressor gene and its transcriptional activation of the Cdk inhibitor $p21^{WAF1/CIP1}$ with subsequent cell cycle arrest, following DNA damage, allows time for DNA repair before entering S phase (15). Activation of p53 in cells with unrepaired DNA damage may lead to apoptotic cell death.

Gap Phase 1, or G_1

G_1 was originally thought to be a gap in the cell cycle—a resting phase while the cell recovered from division before starting a new cell cycle by synthesizing more DNA—although this is now known not to be so. Rather, complex control processes are preparing the cell to synthesize the cellular machinery required for DNA synthesis in S phase. There is also a "point of no return" near the end of G_1 called the restriction point (R), which is analogous to START in yeast. Once cells pass this point they are committed to the next stages of the cell cycle. Signals for passing this point include adequate cell size and a favorable environment, including the presence of growth factors. The transition at the R point from growth-factor dependence to growth-factor independence in normal cells, as opposed to cancer or transformed cells, is likely to reflect a stringent growth regulatory mechanism that is defective in the latter cell types. Coordination between cell cycle commitment, START (or R), and cell size, involving the G_1 cyclins, has been suggested as a regulatory mechanism for cell growth in yeast. During early G_1, immediately after G_0 cells are stimulated to reenter the cell cycle, a critical step occurs involving the induction of cell cycle–specific genes including *myc, cyclin D, and fos*, and Cdk genes. It is now apparent that Cdk 2, Cdk 4, and Cdk 5, despite constitutive expression of Cdk4 thoughout the cell cycle, start to complex with cyclins D1, D2, and D3 and become active in early to mid-G_1 (possibly at a point when or soon after noncycling cells return from G_0). The level of activation of the G_1 Cdk-cyclin complexes peaks at about R; so it is possible that, by reaching a threshold activity, it is these complexes that allow cells to pass R. Cdk2 also forms active complexes

Table 1. Cyclin-Cdk Complexes, Stability, and Associated Proteins

Cyclin	Associated Cdk	Cell cycle activity	Associated proteins	Stability	Degradation
A	Cdk2 and cdc2 (Cdk1)	S and $G_2 \rightarrow M$	p107, E2F, p21, PCNA	Unstable in mitosis	Ubiquitin
B1	cdc2 (Cdk1)	$G_2 \rightarrow M$	(p21), (PCNA)	Unstable in mitosis	Ubiquitin
B2		$G_2 \rightarrow M$	(p21), (PCNA)	Unstable in mitosis	Ubiquitin
C-type	Cdk X	?	ND	ND	PEST
D1	Cdk4(2, 5, 6)	G_1	pRb, p21, p27, PCNA, p16, p15	Rapid turnover	PEST
D2	Cdk4(2, 5, 6)	G_1	pRb, p21, p27, PCNA, p16, p15	Rapid turnover	PEST
D3	Cdk4(2, 5, 6)	G_1	pRb, p21, p27, PCNA, p16, p15	Rapid turnover	PEST
E	Cdk2	G_1 and $G_1 \rightarrow S$	p107, E2F, p21, PCNA and p27 after TGFβ treatment	Rapid turnover	
F			ND	ND	
G			ND	ND	
H	Cdk7	All phases	ND	ND	ND

Source: From Ref. 12: T. Hunter and J. Pines, *Cell* **79,** 573–582 (1994), Copyright Cell Press, used with permission.
Note: ND, not determined; PCNA, proliferating cell nuclear antigen; TGFβ, transforming growth factor β; PEST, proteins that contain sequences rich in proline (P), glutamate (E), serine (S), and threonine (T).

with cyclin E starting from mid-G_1 and peaking later than the Cdk–cyclin D complexes, in other words, between R and the G_1–S phase transition.

The primary substrate of the cyclin D complexes is the retinoblastoma tumor suppresser gene product pRB. Active, unphosphorylated pRB complexes with and inhibits the transcription factor E2F. When pRB is phosphorylated and inactivated, the complex dissociates, releasing active E2F; the cell cycle is not stopped at the G_1 checkpoint, and cells are allowed to proceed to S phase (16). E2F was originally characterized as an activator of the adenovirus early region 2 transcription unit (16). It is now recognized that E2F promotes the transcription of enzymes associated with DNA synthesis (18,19). It has been proposed that the cyclin E complex also phosphorylates pRB, with similar effects (20).

Synthesis Phase, or S

The transition from G_1 to S phase is marked by the transcription (as already described) and activation of the many synthetic enzymes required for cellular duplication, including DNA polymerase α, thymidine kinase, and dihydrofolate reductase (DHFR) (21). This is followed by the exact duplication of the genetic material and replication of organelles and structural proteins. At the same time there is a marked change in regulatory proteins activated. At the onset of S phase, the activation of the G_1-phase Cdk-cyclin complexes decreases in the same temporal order as it increased, and the complexes dissociate. As Cdk2 disassociates from cyclins D and E, it complexes with cyclin A, which starts to be synthesized at about the G_1–S transition. The Cdk2–cyclin A complex retains its activity until the end of S phase. Disruption of cyclin A function inhibits DNA synthesis, suggesting that a threshold activity of the Cdk2–cyclin A complex is essential for the maintenance of S phase.

Gap Phase 2, or G_2

As with G_1, G_2 was originally thought to be a gap in the cell cycle—a resting phase while the cell recovered from DNA synthesis before dividing. This has now been shown not to be so, although G_2 remains analogous to G_1 in that, once the appropriate control signals are received, genes encoding for the proteins responsible for cell division are transcribed and ready for cell division. At the end of S phase the activity of the Cdk2–cyclin A complex decreases as the components dissociate. Cyclin A and cyclin B then form complexes with Cdk1 (also often referred to as cdc2 because of its homology with the product of *S. pombe* cell division cycle gene 2). It is these complexes, increasing in activity throughout G_2 and continuing into mitosis, that mediate activation of the cell division machinery. At the transition point between G_2 and M is a second check point that, in effect, ensures that the cells are big enough, the environment is favorable, and DNA synthesis has been fully and accurately completed before the cell continues to mitosis and division.

Mitosis, or M: Cell Division

Mitosis is the division of one normally diploid, or $2n$, somatic cell into two genetically identical (both to each other and the parent) cells. Mitosis, like the cell cycle itself, displays distinct phases that are best represented diagramatically (Fig. 2). When the cell leaves interphase at the end of G_2, it enters prophase. The nucleolus disappears (disperses), and nuclear DNA condenses into chromosomes. Following DNA synthesis in S phase, a normally diploid ($2n$) cell is now tetraploid ($4n$) and contains an identical pair of each chromosome, which in turn is organized as a pair of identical chromatids joined to each other at the centromere. During late prophase, the nuclear envelope is broken down. At the same time, the centrioles divide; one pair migrates to each cell pole with spindle fibers, which are microtubular structures, linking each pair. The transition from prophase to metaphase is marked by alignment of all the chromosome pairs on an equatorial plane, the metaphase plane, with their centromeres connected to each polar centriole pair by spindle fibers. The cell progresses into anaphase when the chromosomes separate and one chromatid from each homologous chromosome pair migrates along the spindle fibers toward each centriole. In the next phase, telophase, when the chromatids (one

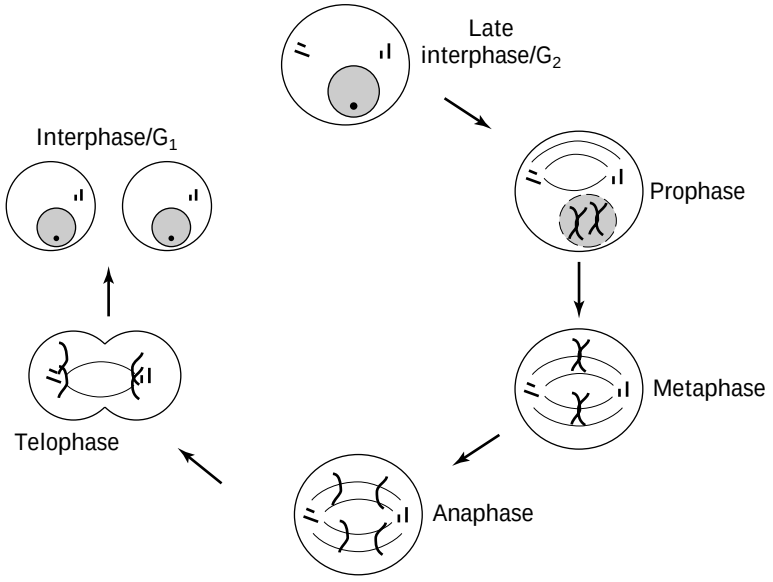


Figure 2. A simple diagrammatic representation of mitosis.

from each homologous chromosome pair) arrange at each pole, the spindle fibers depolymerize, and cytokinesis (division of cytoplasm) begins as a cytoplasmic cleavage furrow begins to form. Further changes follow: the cytoplasm cleaves completely, chromosomes uncoil, the nuclear envelope reforms around each nucleus, and the nucleoli reappear to give two genetically identical daughter cells in interphase. The final telophase changes, in which the cytoplasm divides and the cells separate, are markedly different in plant cells and in animal cells due to the constraints of the rigid cellulose cell wall. As with other cell cycle phases, Cdk-cyclin complexes perform a major regulatory function in mitosis. The activity of Cdk1–cyclin A and Cdk1–cyclin B, which started in G_2, continues to increase until the middle of mitosis, at the metaphase–anaphase boundary. At this point the activity of both complexes declines abruptly to zero, allowing the cell to complete mitosis. It is also thought that the rapid decrease in G_2 cyclin-Cdk complexes is necessary for cells to start a new cell cycle at G_1.

Meiosis. Although unlikely to be encountered in process biotechnology, it is useful to point out that germ cells may undergo an alternative form of division to form zygotes, which may also be considered to be a form of terminal differentiation, called *meiosis*. Assuming a normally diploid ($2n$) cell, initial cell division occurs exactly as in mitosis until the end of anaphase, or in this case the first meiotic anaphase. At telophase the cell may divide or partially divide and enter a brief meiotic interphase comprising two identical diploid ($2n$) daughter nuclei (if not cells). Otherwise the cell progresses directly to the second meiotic prophase. Now there is only a single copy of each chromosome, still comprising two chromatids. During the second meiotic anaphase, the chromatids separate again to each of four poles. The cell then progresses to (the second meiotic) telophase and on to interphase to give four haploid (n) daughter cells instead of the two diploid ($2n$) cells produced by mitosis.

Gap Phase 0, or G_0

The term G_0 was first used in the early 1960s when three types of cell were recognized: static cells, incapable of further division (for example, a terminally differentiated adult neurone); renewing, constantly dividing cells (for example, bone marrow cells); and conditionally renewing cells, which do not normally divide but can respond to an appropriate extraordinary stimulus (for example liver cells starting to proliferate after partial hepatectomy). G_0 is used to describe this last category of cells, which have "opted out" of the cell cycle after a mitosis because of, for example, growth factor depletion. In this circumstance, in contrast, transformed or tumor cells do not enter G_0 but continue through the cell cycle until they die through nutrient exhaustion, often by apoptosis. It has been argued that G_0 should not be used to describe cells, for example, in the stationary phase of a hybridoma, or CHO batch cultures, which are in a pseudo-steady-state of growth and death due to nutrient depletion and/or accumulation of toxic metabolites. Unlike the normal cells, they have not downregulated themselves; they have been closed down by their environment. Cells that lose their ability to respond to growth factor depletion or do not exit the cell cycle to G_0 may have a defective G_1 program, which is different from the cyclin-Cdk4-pRB phosphorylation mechanism involved in R regulation in normal cells. However, G_0 is commonly used today, some claim inappropriately, to describe any cell (containing a G_1 amount of DNA) that is not expected to proceed to S phase for any reason, including the result of any means of artificial interference with the cell cycle. Markers of G_0 include reduced RNA, reduced ribosome content, and an absence of cyclins D and E (22).

External (Environmental) Influence on the Cell Cycle

Extrinsic cell cycle regulation may be considered at two levels. At the more basic level is the effect of the physical environment and nutrient availability; the more complex level of extrinsic cell cycle control is through growth fac-

tors. The physical environment has a great effect on the cell cycle. When temperature is reduced from the optimum, the length of the cell cycle increases, and growth rates slow down. When temperature exceeds the optimum, even by a few degrees, profound changes in cell behavior occur, including the expression of heat shock proteins, and cells eventually die. Alternatively, if appropriate nutrients such as amino acids and glucose are not available in the medium, cell cycle duration and even progression will be influenced due to lack of energy and the precursors for protein and DNA synthesis.

Serum starvation, in the presence of appropriate nutrients and physical conditions, may also affect the cell cycle, but this is more likely to be caused by the absence of specific growth factors that serum provides. Most, if not all, cell lines were originally grown in medium containing animal serum, usually newborn or fetal bovine serum. The main contribution to the cell cycle and, hence, growth of a culture from these undefined and heterogeneous supplements is the supply of growth factors that promote or enhance the cell cycle. Thus, because G_0 cells reenter the cell cycle as a result of serum addition, serum is regarded as a source of positive regulatory growth factors. The presence of growth factors in the environment, their binding to specific transmembrane receptor proteins, and subsequent signal transduction effects changes in cellular processes, including the cell cycle. Expression of cyclin D, important in progression through G_1, is induced by growth factors (23). These effects are cell-type dependent because cells do not all express the same growth factor receptors. Furthermore, different cell types/lines may not only require different individual growth factors but may require a specific cocktail for optimal proliferation. Fibroblast growth, for example, requires growth factors including fibroblast growth factor (FGF), epidermal growth factor (EGF), and platelet-derived growth factor (PDGF). Other growth factors that promote cell division include insulin-like growth factor (IGF) I and II, macrophage–granulocyte colony stimulating factor (G-CSF), and endothelial cell growth factor. Each growth factor stimulates a specific receptor on the plasma membrane, which in turn activates protein kinases, many of which phosphorylate tyrosine residues on their substrates, so they are generically called tyrosine kinases. Subsequent signal transduction by phosphorylation/dephosphorylation cascades can lead to an increased rate of protein synthesis, influence the activity of DNA topoisomerase, or modify the glucose transport system—all with the similar effect of promoting cell division. Transformed cells are known to have a reduced (or no) requirement for external growth factors compared with normal cells. However, through adaptation of cell lines and extensive development of medium formulation, many cells can now be grown in protein-free medium, indicating that it is possible to bring about cell proliferation and support long-term culture in media free of (extrinsic) growth factors, hormones, and other contaminating proteins. This absence of extraneous protein is particularly important in large-scale process biotechnology for facilitating downstream processing and product recovery.

It is commonly accepted that the accumulation of toxic metabolites (ammonia and lactate, for example) inhibits culture proliferation. However, in addition to the positive growth-promoting (cell cycle–promoting) factors, there are also negative, growth-inhibiting factors, although these may be the same growth factors displaying opposite effects under different conditions. Transforming growth factor $\beta 1$ (TGF-$\beta 1$), in addition to its strong positive stimulatory effect, has been shown to produce a density-independent negative response that prolongs G_1 in the NIH-3T3 mouse embryo cell cycle and has a similar effect on mortal human fibroblasts (24).

METHODS FOR DESCRIBING THE CELL CYCLE

Mitotic Index

There are three main approaches to describing the cell cycle. The first and most historically established approach is to determine the percentage of cells in mitotic cell division. This method, called the *mitotic index*, is simple and can be applied to attached cells grown on a microscope slide or any other appropriate substratum, as well as cells in suspension. There are many fixation and staining methods described in any textbook of cytology methods, from which a protocol appropriate to animal cell biotechnology can be selected. A generic protocol is as follows. Take cells grown on a glass microscope slide (or coverslip) or prepare a smear of suspended cells on a microscope slide, using an attachment agent such as polylysine if necessary. Fix the cells with a fixative appropriate to the proposed staining method. Stain the cells with either a standard cytological or histological stain that differentiates DNA, such as Feulgen staining, or alternatively, stain with a fluorescent DNA stain such as ethidium bromide (EB), Hoechst 33258, or DAPI. Examine stained slides by conventional or fluorescence microscopy, as appropriate; count the number of nuclei in mitosis and express the result as a percentage of the total. A recent publication, investigating mitosis in a hydrodynamically stressed environment, described the use of the fluorescent DNA-intercalating stain propidium iodide (PI) with cells fixed with 3:1 ethanol:glacial acetic acid, followed by fluorescence microscopy (25). A major disadvantage to measuring the mitotic index is that, although it gives information about cell division in M phase, it can give no information about the other cell cycle phases or their associated processes. Measuring the mitotic index also exhibits disadvantages common to all manual microscopic methods, including sampling error because a small proportion of the total population is examined, and interobserver variation resulting from subjective discrimination between mitotic and interphase cells. Furthermore this slow and labor-intensive method is not appropriate for online automation or use in integrated process control.

Labeled Nucleotide Uptake in S Phase

An alternative approach for describing the cell cycle is to document the incorporation of labeled nucleotide into nuclear DNA during S phase. This method relies on [^{3}H]thymidine incorporation followed by autoradiography and can be used in one of two ways: continuous labelling, in which the cells are exposed to the label for a fixed time,

usually 24 or 48 h followed by analysis at that time; or pulse–chase labeling, in which the cells are exposed to the label for a short pulse (0.5–1 h), then analyzed at regular intervals for 24–48 h. Continuous labeling gives information about the growth fraction of a culture, which is the percentage of the cells that are in the cell cycle during the labeling period. However, this information is not likely to be of great use to the process biotechnologist because most cells under optimum conditions, when viability is close to 100%, will also have a growth fraction close to 100%. A simple generic protocol for continuous labeling is as follows: [^{3}H]thymidine is added to the culture medium, and the culture continues for a specified period; cells are fixed and subjected to autoradiography; when the autoradiograph has developed, the number of radiolabeled nuclei is counted and expressed as a percentage of the total.

In contrast, pulse–chase labeling is used to determine the average duration of each cell cycle phase. A generic protocol for [^{3}H]thymidine pulse–chase labeling is as follows: [^{3}H]thymidine is added to the culture; after a short period (1–2 h), cells are removed to fresh medium without [^{3}H]thymidine; the culture is sampled regularly at times greater than the length of one cell cycle (24–48 h); each sample is subjected to autoradiography; and the number of labeled mitoses is counted as a percentage of the total. The average duration of G_2, S, M, and the total cell cycle is measured from the plot of labeled mitoses against time, then G_1 is calculated. Details of using [^{3}H]thymidine incorporation for determining growth fraction and cell cycle phase duration are summarized elsewhere (26). As with the mitotic index, any autoradiographic technique exhibits inherent disadvantages including sampling error because a small proportion of the total cell population is examined; the technique is slow, labor intensive, and not appropriate for on-line automation or use in integrated process control; and the use of radioisotopes requires additional containment and waste disposal facilities.

In recent years the thymidine analogue bromodeoxyuridine (BrdU) has been used as a safer, nonisotopic alternative to [^{3}H]thymidine for DNA labeling. BrdU can be used in exactly the same way as [^{3}H]thymidine, as either a continuous or pulse–chase label. The only difference in the protocol is that BrdU incorporation is detected by immunostaining with fluorescent anti-BrdU antibodies followed by fluorescence microscopy instead of autoradiography, so the method still suffers the disadvantages of any manual microscopic method. Alternatively, BrdU incorporation can be measured by flow cytometry, *vide infra*.

Measuring DNA Content

The third and probably easiest approach to adopt for cell cycle analysis is to measure the DNA content of each cell in a population by flow cytometry and thus determine the proportion of cells in G_0/G_1, S, and G_2/M at the time of sampling. A major and obvious advantage to this approach is that heterogeneous nonsynchronized cultures can be analyzed easily and rapidly, with the useful bonus that sub-G_1 apoptotic populations can also be identified. Other advantages include minimizing sample error by analyzing a large number of cells, and minimal interobserver variation because of the objective discrimination between G_1, G_2, and S cells. Furthermore, the method is appropriate for near on-line automation and use in integrated process monitoring or control. Although this approach is not able to resolve G_0/G_1 or G_2/M because DNA content is the same, this is of little concern in most process biotechnology applications because the (transformed) cells commonly used do not enter G_0. Furthermore the G_2/M cell subpopulation can be resolved, if necessary, by measuring the mitotic index of the population (as already described) to determine the proportion of cells in M phase.

Flow Cytometry for Cell Cycle Analysis

Flow cytometric cell cycle analysis by measuring DNA content may follow any one of many protocols. A generic protocol is as follows: fix/permeablize the cells; stain with a fluorescent DNA stain; measure stained DNA content by flow cytometry; and analyze data by simple gating or using dedicated cell cycle analysis software. Variations include pretreatment of cells; altered fixation/permablization protocols; postfixation treatment of cells; different stains, although stains such as PI or EB, which intercalate into double-stranded nucleic acid, are commonly used; and methods for data analysis. A typical flow cytometric measurement of PI-stained DNA in ethanol-fixed recombinant CHO cells and the subsequent cell cycle analysis, using software based on a polynomial S-phase algorithm with an iterative nonlinear least squares fit, is shown in Figure 3. By fitting the G_1 and G_2 peaks as Gaussian distributions, and the S-phase distribution as a Gaussian broadened distribution, the cell cycle phase distribution can be deconvoluted.

Flow cytometry can also be used with BrdU incorporation to measure the growth fraction or to determine average cell cycle phase times. For flow cytometric analysis, BrdU incorporation can be detected by either fluorescent immunostaining or its quenching effect on Hoechst DNA stains. Thus BrdU incorporation can be analyzed simultaneously with DNA content. The advantage of this method is that it can recognize arrested but non-DNA-synthesizing cells during S phase. Cell size and/or protein content can also be analyzed simultaneously to give further related information. Again, there are many variations on a theme, details of which and other flow cytometric methods for cell cycle analysis have been summarized elsewhere (27). Flow cytometry is also very useful for analyzing specific cell cycle–related proteins, following appropriate fluorescent immunostaining. For example, the absence of cyclin D or E in G_0 cells, shown by flow cytometry, has been used to distinguish G_0 from G_1 cells (22), and flow cytometric measurement of the Ki67 nuclear antigen has been shown to correlate with [^{3}H]thymidine and BrdU incorporation (28). Multiparameter flow cytometric analysis of G_1 and G_2 cyclin expression and DNA content has also been shown to be a useful tool in cell cycle analysis (29) because G_1 cells can be differentiated from G_0 cells, and cells with unusual karyotypes (for example, G_1 tetraploid) can be distinguished from G_2 diploid cells. Flow cytometry can also be applied to investigating the relationship between the cell cycle and recombinant protein expression in

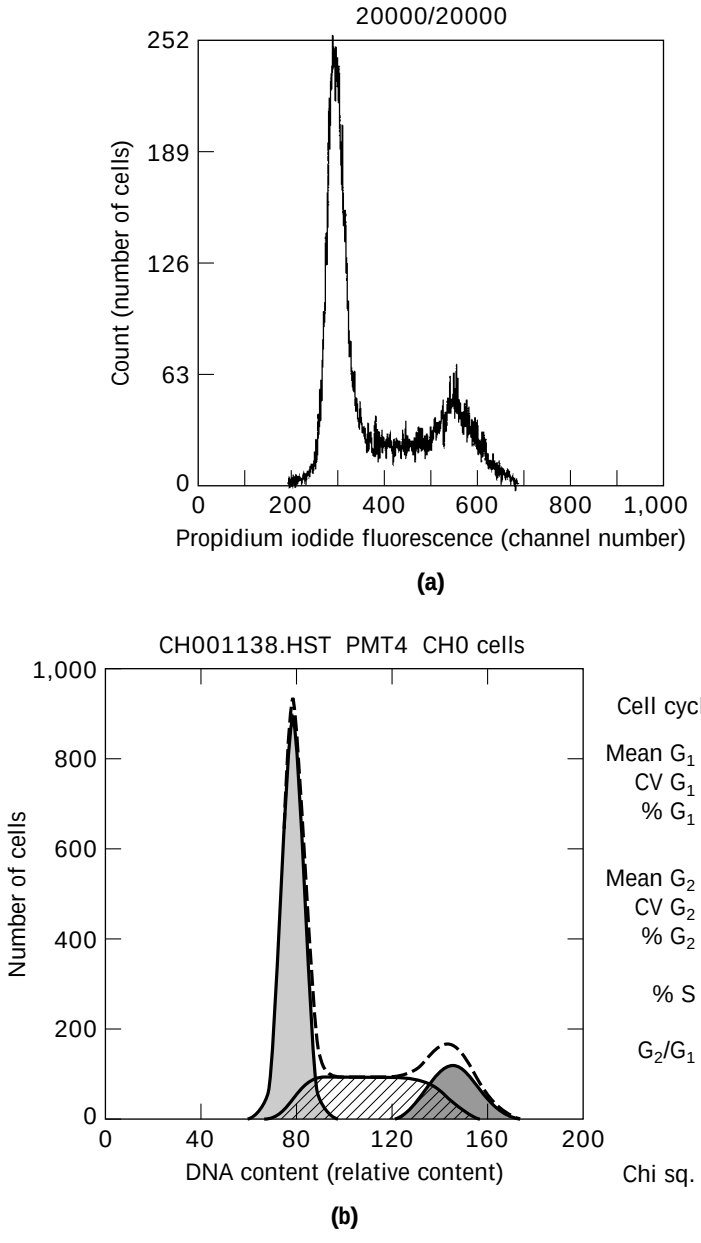


Figure 3. (**a**) Typical flow cytometric measurement of fluorescent (propidium iodide–stained) DNA content in recombinant CHO cells. (**b**) Resolution of cell cycle fractions from the fluorescent DNA content of recombinant CHO cells, using dedicated cell cycle analysis software (Multicycle®, Phoenix Flow Systems).

economically important cell lines as a means of optimizing medium and environmental conditions for successful commercial operation.

THE IMPORTANCE OF THE CELL CYCLE IN PROCESS BIOTECHNOLOGY

Culture Proliferation and Growth

The cell cycle is the means by which a culture proliferates. Cell division and cell death, both active programmed cell death (apoptosis) and passive necrosis, are responsible for the change in the number of cells present. In addition to the cell cycle control genes, the transcription of genes coding for proteins involved in nucleotide and DNA synthesis is also cyclic. However, the progress of the cell cycle is costly in nutrient utilization and energy expenditure, particularly during S phase, which is when the nucleic acid content of the cell is duplicated, but also during other phases because regulatory proteins are continuously synthesized and degraded. In theory it should be possible to arrest cells, in other words, stop them from cycling, thus preventing the "wastage" of energy and nutrients that could otherwise be directed into product synthesis. The maintenance energy model (30) states that energy (as adenosine triphosphate [ATP]) consumption is divided into that needed for cell growth (cell cycle progression, as determined by increase in cell number) and that needed for cell maintenance. Based on this model it has been determined that the maintenance energy requirement of a hybridoma is 14.4×10^{-9} mmol ATP/(cell day) at a specific growth rate (μ) of 0.66 per day (31), corresponding to 62%

of the energy (ATP) consumed. These values were confirmed by work that showed that the maintenance energy requirement of the same cells under slightly different nutrient and specific growth rate conditions was 11.4×10^{-9} and 12.0×10^{-9} mmol ATP/(cell day), again, corresponding to approximately 60% of the estimated ATP consumption rate (32). These observations agree with the value of 65%, which was determined as the proportion of energy consumption used for maintenance by mouse LS cells grown at $\mu = 0.60$ per day (33). Thus it is apparent that about two-thirds of the energy used by a culture is for its own maintenance, and this proportion is not likely to be readily altered. However the remaining third, which is used for growth (and product elaboration), may be exploited. If the growth (cell cycle progression) requirement was removed, all of the remaining resources could be directed toward product elaboration, or alternatively, cultures would require up to one-third less energy and nutrient resources for the same product yield.

Two main approaches have been used to prevent cells from cycling. The first is chemical blockade. For example, the addition of thymidine to culture medium causes cells to arrest at the G_1–S phase boundary, so cells accumulate in G_1, and specific productivity increases. However, this production advantage is short lived because the imbalanced growth state leads to culture death (usually by apoptosis) within a few days of the onset of the chemical blockade (D. C. Edwards and M. Al-Rubeai, unpublished data, 1997). The second approach is not to arrest the cell cycle but to substantially reduce its progression by reducing the culture growth rate. Using a custom-built total cell recycle bioreactor system, where the growth rate of a mouse hybridoma was substantially reduced ($\mu < 0.05\ \mu_{max}$), Flickinger et al. (34) showed that specific monoclonal antibody secretion remained constant as the specific growth rate decreased. This demonstrated that production could be uncoupled from growth (cell cycle progression), although there was no evidence that the energy saved by substantial (>95%) reduction of growth was diverted into production. However, using this method for production-scale processes may have the advantage of less medium (energy/nutrient) expenditure per unit product, although the total product yield per unit of production (from each bioreactor) would not be increased. Furthermore, questions of product fidelity, especially with respect to glycosylation, are raised when growth is limited by the stress of nutrient deprivation or limitation.

Until recently, these were the only means available for controlling the cell cycle (i.e., preventing cycle progression). In 1997, Fussenegger et al. (35) described a novel cytostatic process that enhances productivity. This was achieved by adopting a completely different approach compared with previous work. Metabolic engineering was used instead of the application of an external stress to the cells to prevent cell cycle progression. A series of transient CHO transfects was produced containing inducible dicistronic expression vectors, each of which contained a model product gene (human secretory alkaline phosphatase [hSEAP]) and a tumor-suppressor gene coding for an intracellular regulator of cell cycle progression (either p21, p27, or a mutant p53 [p53175P] that is said to have lost its proapoptotic properties while retaining its cell cycle regulation properties (36)). When the expression vector was induced, cells arrested and began to accumulate in G_1. The arrested cells showed a fourfold increase in hSEAP production compared with cells transfected with a control expression vector coding for the product only. Using stable dicistronic transfects, only p27 overexpression successfully induced cell cycle arrest (in G_1), but with it came a 10-fold increase in specific hSEAP productivity (35). Building on the success of these discistronic vectors, the tricistronic pBluescript®-based pTRIDENT family of vectors was developed (37), allowing the possibility of a single stable transfection coding not only for a product but also for cell cycle control (cell cycle arrest) and enhanced culture survival using an antiapoptosis oncogene such as *bcl-2* or *bcl-*x_L (38–40). Subsequent work with pTRIDENT vectors containing genes encoding for hSEAP with either p21 and the CCAAT/enhancer binding protein α (which induces and stabilizes p21 (41,42)) or p27 and Bcl-x_L showed even greater benefit, with specific productivities increased 15- and 30-fold, respectively, with respect to proliferation-competent controls (43).

Production in Relation to Cell Cycle Phase

It is important not to confuse the relationship between cell cycle phase and productivity with the relationship between cell cycle arrest and productivity. The substantial increases in productivity from the metabolically engineered transfects just described are most likely to exist because the arrested cells do not need to devote resources to biomass production (35). This section addresses the situation in "simple" recombinant cells that may be grown in standard industrial processes such as batch, continuous, or draw-and-fill semicontinuous culture modes. Determining the relationship between cell cycle phase and productivity is important for developing automated on-line (or near on-line) process control strategies. These strategies will be based on models that, in addition to the basic parameters of pH, dO_2, cell density, and so forth, must incorporate information concerning cell state, including cell cycle phase and its effect on productivity.

It has been shown that specific monoclonal antibody production by a synchronized mouse–mouse hybridoma culture is maximal during G_1; furthermore, the specific activity increases when the cell cycle is arrested and cells are maintained in late G_1 (44). However, apparent specific productivity also increases when the culture declines (viability decreases); this is the result of substantial passive release of intracellular product into the culture medium from dead and dying cells. Phase dependence was also found in recombinant CHO cells transfected with an expression vector containing *dhfr* and *lacZ* genes; the former was controlled by an SV40 promoter, and the latter was controlled by the constitutive hCMV promoter (45). In this study, the nonsecreted product, β-galactosidase, was produced almost exclusively during S phase. Furthermore, cell cycle arrest in S phase, by specific inhibition of DNA polymerase, did not inhibit transcription or translation and thus had no effect on β-galactosidase expression.

However, the literature on the relationship between cell cycle and protein expression, summarized in Table 2, is

Table 2. Variation in Phase Expression of Different Products with Cell Line and Promoter

Cell line	Promoter	Product	Phase expression	Ref.
Mouse leukaemia	Endogenous	DHFR	Aphasic	46
CHO	Endogenous	DHFR	Aphasic	47
Human osteosarcoma	Endogenous	DHFR	Aphasic	47
Mouse embryo fibroblast	Endogenous	DHFR	Aphasic	47
Human prostate epithelium	Endogenous	PSAP	All except early G_1	48
Hybridoma	Endogenous	MAb	G_1/G_0 (mainly)	49
CHO	AMLP	DHFR	G_1 (mainly)	50
Hybridoma	Endogenous	MAb	G_1/S	44
Mouse plasmacytoma	Endogenous	Immunoglobulin	Late G_1/early S	51
Human lymphoid cell	Endogenous	Immunoglobulin	Late G_1/S	52
CHO	Endogenous	DHFR	G_1/S[a]	22
Mouse leukemia	Endogenous	TMPS	S	46
CHO	CMV	β-Gal (*lacZ*)	S	45
CHO	SV40	tPA	S (mainly)	50
CHO	SV40	IFN-γ	S	53
Mouse L-cell	MMTV	β-Gal (*lacZ*)	S (predominantly)	54
CHO	Endogenous	DHFR	S	55
Rat adenocarcinoma	Endogenous	tPA	Shortly after S	56
Rat adenocarcinoma	Endogenous	uPA	Shortly after S	56
CHO	Endogenous	PA	G_2/M	57

Note: DHFR, dihydrofolate reductase; CHO, Chinese hamster ovary; TMPS, thymidine monophosphaste synthetase; MAb, monoclonal antibody; PSAP, prostate-specific acid phosphatase; AMLP, adenovirus major late promoter; CMV, cytomegalovirus; β-Gal (*lacZ*), β-galactosidase (product of the *Escherichia coli lacZ* gene); SV40, simian virus 40; tPA, tissuelike plasminogen activator; IFN-γ, interferon-γ; uPA, urokinase-like plasminogen activator; PA, plasminogen activator.

[a]Gene expression activates at G_1–S transition.

inconsistent. Different researchers have shown that product expression (or the maximum rate of expression) is related to different cell cycle phases. In addition, others have shown that product expression is aphasic. These observations do not necessarily contradict each other and may be a true reflection of the systems used, because a variety of cells, products, and constructs were studied. Cell cycle–dependent product expression may not be universal and could vary with cell type or cell line, the construction of the expression vector, the nature of the recombinant gene expressed, or the promoter/enhancer used to drive product expression. Other variables may include posttranscriptional regulation (mRNA level), product gene copy number, and even the culture mode, (e.g., whether the cells are attached to a substrate or are free in suspension) may be relevant. However, there are two potential flaws in studying cell cycle–related productivity: synchronizing cells using chemical means, such as thymidine block or nutrient starvation, and comparing cells of different phases from the same culture that are separated in time. The two problems often go together, for example, following a culture for a day after release from thymidine block. The difficulty with the former is that the chemical stress used (what ever its nature) may, in addition to its desired effect of cell synchrony, directly cause perturbations in productivity that could wrongfully be ascribed to the cell cycle. The difficulty with the latter is simply that if two cell populations from a culture are separated in time, they are also separated in environmental conditions; recent work (58) has shown that the effect of medium condition is dominant to any cell cycle effect on productivity. The effect of temporal separation is made more uncertain following stress synchronization; an early population may have only partially recovered from the stress, whereas a later population has fully recovered.

Cell Cycle Analysis as a Proliferation Predictor

Cells may be arrested in G_1 before the restriction point, but once past that point, they are committed to continue their cycle and divide unless prevented from doing so by catastrophe or artificial intervention. Therefore, one can presume that, once a cell enters S phase, under normal growth conditions, one single cell will in due course become two cells. This phenomenon can be used to predict culture growth. The growth-enhancing effect of peptone supplementation on hybridoma culture was predicted by cell cycle analysis 5 h after supplementation, although the effect on cell density was not apparent until 16 h later, 21 h after supplementation (59). A similar use of cell cycle analysis was made with CHO cells: the percentage of S-phase cells correlated with the specific growth rate 10.5 h later; the time difference was the average duration of S phase remaining plus the duration of G_2 and M for that population (60). Thus, the proportion of cells in S phase can be used to predict the subsequent proliferation rate of that culture.

Cell Cycle–Related Susceptibility to Hydrodynamic Damage

The effect of hydrodynamic stress, as found in stirred and sparged bioreactors, is at least in part related to cell cycle phase, and there may be a role for hydrodynamic cell damage in the induction of apoptosis (61). It has been reported that DNA synthesis (62) and the proportion of the cell population in S phase (63) were increased under conditions of hydrodynamic stress (high sparging and agitation rates). However, the cell number remained low even though the S-phase population correlated with an increase in cell number. The observed reduction in growth rate was

shown to be a consequence of a decline in the proportion of G_2- and M-phase cells, with a corresponding decrease in the cell size of the population, suggesting that the larger G_2 cells were particularly susceptible to hydrodynamic damage (61). Subsequently it was shown that mitotic cells, while also being susceptible to hydrodynamic damage due to their size, were even more sensitive to hydrodynamic damage than G_2 cells, perhaps due to cytoskeletal rearrangements, alteration of plasma membrane fluidity, or morphological changes during mitosis (25).

Cell Cycle and Apoptosis

Apoptosis (active or programmed cell death) is the means by which intrinsic or extrinsic stimuli activate a series of molecular cascades, triggering events that culminate in a particular morphologically, ultrastructurally, and biochemically defined process of cell death that is, in effect, the opposite of the cell cycle. Entry of a cell into apoptosis can be restricted to a particular cell cycle phase (54–66), although this cell cycle dependence does not appear to be universal (67). Cell cycle progression and apoptosis are not only opposite in effect, they may also be alternate outcomes from a single cell control mechanism; they are inextricably linked by common components in the signal transduction pathways for both processes. In addition, circumstantial evidence from the induction of apoptosis by DNA damage or cell cycle disruption and the existence of many oncogenes that both promote the cell cycle and induce apoptosis implies that there is an equilibrium between these two antagonistic processes in all eukaryotic cells.

BIBLIOGRAPHY

1. B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J.D. Watson, *The Molecular Biology of the Cell*, 3rd ed., Garland, New York, 1994.
2. J.D. Watson, N.H. Hopkins, J.W. Roberts, J.A. Steitz, and A.M. Weiner, *Molecular Biology of the Gene*, 4th ed., Benjamin/Cummings, Melno Park, Calif., 1987.
3. H. Lodish, D. Baltimore, A. Berk, S.L. Zipursky, P. Matsudaira, and J. Darnell, *Molecular Cell Biology*, 3rd ed., Scientific American Books, New York, 1995.
4. A. Murray and T. Hunt, *The Cell Cycle: An Introduction*, Freeman, New York, 1993.
5. S. Cooper, *Nature* **280**, 17–19 (1979).
6. S. Cooper, *J. Theor. Biol.* **135**, 393–400 (1988).
7. G.I. Evan, E. Harrington, N. McCarthy, C. Gilbert, M.A. Benedict, and G. Nuñez, in N.S.B. Thomas ed., *Apoptosis and Cell Cycle Control in Cancer*, Bios Scientific, Oxford, 1996, pp. 109–129.
8. I.B. Holland, *Cell* **48**, 361–362 (1987).
9. E. Boye, T. Stokke, N. Kleckner, and K. Skarstad, *Proc. Natl. Acad. Sci.* **96**, 12206–12211 (1996).
10. K.C. Quon, G.T. Marczynski, and L. Shapiro, *Cell* **84**, 83–93 (1996).
11. R. Zetterberg, in N.S.B. Thomas ed., *Apoptosis and Cell Cycle Control in Cancer*, Bios Scientific, Oxford, 1996, pp. 17–36.
12. T. Hunter and J. Pines, *Cell* **79**, 573–582.
13. M. Matsuoka, J.Y. Kato, R.P. Fisher, D.O. Morgan, and C.J. Sherr, *Mol. Cell. Biol.* **14**, 7265–7275 (1994).
14. J.Y. Kato, M. Matsuoka, D.K. Strom, and C.J. Sherr, *Mol. Cell. Biol.* **14**, 2713–2721 (1994).
15. W.S. El-Deiry, in N.S.B. Thomas ed., *Apoptosis and Cell Cycle Control in Cancer*, Bios Scientific, Oxford, 1996, pp. 55–75.
16. R.A. Weinberg, *Cell* **81**, 323–330 (1995).
17. Kovesdi, R. Reichek, and J.R. Nevins, *Cell* **45**, 219–228 (1986).
18. E. Ogris, H. Rotheneder, I. Mudrak, A. Pichler, and E. Wintersberger, *J. Virol.* **67**, 1765–1771 (1993).
19. B.E. Pearson, H.P. Nasheuer, and T.S.F. Wang, *Mol. Cell. Biol.* **11**, 2081–2095 (1991).
20. V. Dulic, E. Lees, and S.I. Reed, *Science* **257**, 1958–1961 (1992).
21. J. Wells, P. Held, S. Illenye, and N.H. Heintz, *Mol. Cell. Biol.* **16**, 634–647 (1996).
22. Z. Darzynkiewicz, J. Gong, G. Juan, B. Ardelt, and F. Traganos, *Cytometry* **25**, 1–13 (1996).
23. C.J. Sherr, *Cell* **73**, 1059–1065 (1993).
24. D. Zhang and J.W. Jacobberger, *Cell Proliferation* **29**, 289–307 (1996).
25. F.A. Abdul Majid and M. Al-Rubeai, in M.J.T. Carrondo, B. Griffiths, and J.L.P. Moreira eds., *Animal Cell Technology from Vaccines to Genetic Medicine*, Kluwer, Dordrecht, 1997, pp. 731–736.
26. R. Baserga, in G.P. Studzinski ed., *Cell Growth and Apoptosis: A Practical Approach*, IRL Press at Oxford Univ. Press, Oxford, 1995, pp. 1–19.
27. H.A. Crissman, in G.P. Studzinski ed., *Cell Growth and Apoptosis: A Practical Approach*, IRL Press at Oxford Univ. Press, Oxford, 1995, pp. 21–43.
28. S.D. Gore, L.J. Weng, and P.J. Burke, *Exp. Hematol.* **21**, 1702–1708 (1993).
29. J. Gong, X. Li, F. Traganos, and Z. Darzynkiewicz, *Cell Proliferation* **27**, 357–371 (1994).
30. S.J. Pirt, *Principles of Microbe and Cell Cultivation*, Blackwell Scientific, Oxford, U.K., 1975.
31. W.M. Miller, C.R. Wilke, and H.W. Blanch, *Biotechnol. Bioeng.* **33**, 477–486 (1989).
32. W.M. Miller, C.R. Wilke, and H.W. Blanch, *Biotechnol. Bioeng.* **33**, 477–486 (1989).
33. D.G. Kilburn, M.D. Lilly, and F.C. Webb, *J. Cell. Sci.* **4**, 645–654 (1969).
34. M.C. Flickinger, N.K. Goebel, and M.A. Bohn, *Bioprocess Eng.* **5**, 155–164 (1990).
35. M. Fussenegger, X. Mazur, and J.E. Bailey, *Biotechnol. Bioeng.* **55**, 927–939 (1997).
36. S. Rowan, R.L. Ludwig, Y. Haupt, S. Bates, X. Lu, M. Oren, and K.H. Vousden, *EMBO J.* **15**, 827–838 (1996).
37. M. Fussenegger, X. Mazur, and J.E. Bailey, *Biotechnol. Bioeng.* **57**, 1–10 (1998).
38. Y. Itoh, H. Ueda, and E. Suzuki, *Biotechnol. Bioeng.* **48**, 118–122 (1995).
39. R.P. Singh, A.N. Emery, and M. Al-Rubeai, *Biotechnol. Bioeng.* **52**, 166–175 (1996).
40. D.C. Huang, S. Cory, and A. Strasser, *Oncogene* **14**, 405–414 (1997).
41. A.L. Gartel, M.S. Serfas, and A.L. Tyner, *Proc. Exp. Biol. Med.* **213**, 138–149 (1996).

42. N.A. Timchenko, M. Wilde, M. Nakanishi, J.R. Smith, and G.J. Darlington, *Genes Dev.* **10**, 804–815 (1996).
43. M. Fussenegger, S. Schlatter, D. Dätwyler, X. Mazur, and J.E. Bailey, *Nature Biotechnology* **16**, 468–472 (1998).
44. M. Al-Rubeai and A.N. Emery, *J. Biotechnol.* **19**, 67–86 (1990).
45. M.B. Gu, P. Todd, and D.S. Kompala, *Biotechnol. Bioeng.* **42**, 1113–1123 (1993).
46. L.H. Matherly, J.D. Schuetz, E. Westin, and I.D. Goldman, *Anal. Biochem.* **182**, 338–345 (1989).
47. J.N. Feder, Y.G. Assaraf, L.C. Seamer, and R.T. Schimke, *J. Biol. Chem.* **264**, 20583–20590 (1989).
48. F. Herz, B. Czerniak, D. Deitch, R.P. Wersto, D.A. Simmons, and L.G. Koss, *Cell Proliferation* **24**, 321–330 (1991).
49. S.J. Kromenaker and F. Srienc, *Biotechnol. Bioeng.* **38**, 665–677 (1991).
50. M. Kubbies and H. Stockinger, *Exp. Cell Res.* **188**, 267–271 (1990).
51. O. Garatun-Tjeldstø, I.F. Pryme, J.K. Weltman, and R.M. Dowben, *J. Cell Biol.* **68**, 232–239 (1976).
52. D.N. Buell and J.L. Fahey, *Science* **164**, 1524–1525 (1969).
53. V. Leelavatcharamas, A.N. Emery, and M. Al-Rubeai, in K. Funatsu, Y. Shirai, and T. Matsushita eds., *Animal Cell Technology: Basic and Applied Aspects*, Vol. 8, Kluwer, Dordrecht, 1997, pp. 55–60.
54. M.B. Gu, P. Todd, and D.S. Kompala, *Biotechnol. Bioeng.* **50**, 229–237 (1996).
55. B.D. Mariani, D.L. Slate, and R.T. Schimke, *Proc. Natl. Acad. Sci.* **78**, 4985–4989 (1981).
56. F.M. Scott, V. Sator de Serrano, and F.J. Castellino, *Exp. Cell Res.* **169**, 39–46 (1987).
57. J. Aggeler, L.N. Kapp, S.C.G. Tseng, and Z. Werb, *Exp. Cell Res.* **139**, 275–283 (1982).
58. D.R. Lloyd, V. Leelavatcharamas, A.N. Emery, and M. Al-Rubeai, *Cytotechnology*, in press.
59. M. Al-Rubeai, S. Chalder, R. Bird, and A.N. Emery, *Cytotechnology* **7**, 179–186 (1991).
60. V. Leelavatcharamas, A.N. Emery, and M. Al-Rubeai, in M. Al-Rubeai and A.N. Emery eds., *Flow Cytometry Applications in Cell Culture*, Marcel Dekker, New York, 1996, pp. 1–15.
61. M. Al-Rubeai, R.P. Singh, A.N. Emery, and Z. Zhang, *Biotechnol. Bioeng.* **46**, 88–92 (1995).
62. Lakhotia, K.D. Bauer, and E.T. Papoutsakis, *Biotechnol. Bioeng.* **40**, 978–990 (1992).
63. S. Oh, M. Al-Rubeai, A.N. Emery, and A.W. Nienow, in R.E. Spier, J.B. Grifiths, and W. Berthold eds., *Animal Cell Technology Products for Today, Prospects for Tomorrow*, Butterworth Heinmann, Oxford, 1994, pp. 405–407.
64. J. Klucar and M. Al-Rubeai, *FEBS Lett.* **400**, 127–130 (1997).
65. K.S. Sellins and J.J. Cohen, *J. Immunol.* **139,** 3199–3206 (1987).
66. S.-C. Shih and O. Stutman, *Cancer Res.* **56**, 1591–1598 (1996).
67. Y.A. Lazebnik, S. Cole, C.A. Cook, W.G. Nelson, and W.C. Earnshaw, *J. Cell Biol.* **123**, 7–22 (1993).

See also BIOENERGETICS OF MICROBIAL GROWTH; CULTURE COLLECTIONS; ENERGY METABOLISM, MICROBIAL AND ANIMAL CELLS.

CELL CYCLE, EUKARYOTES

NICHOLAS R. ABU-ABSI
FRIEDRICH SRIENC
University of Minnesota
St. Paul, Minnesota

KEY WORDS

Animal cell culture
Cell cycle
Cell cycle controls
Cell cycle models
Cell cycle regulation
Cyclin-dependent kinases
Cyclins
Hybridomas
Protein secretion

OUTLINE

INTRODUCTION

The growth of eukaryotic cells is governed by a complex array of biochemical reactions that are collectively known as the cell cycle. The cell cycle controls are intimately coupled to the internal and external states of the cell and thus ultimately determine when a cell devotes the majority of its energy to growth, DNA replication, successful cell division, or cell maintenance. A recent explosion in the amount of research devoted to cell cycle controls has revealed that the activity of enzymes known as cyclin-dependent kinases (CDKs) regulates the order of cellular events that culminate in cell division. Furthermore, these CDKs are themselves subject to strict controls that govern their expression, activation, and degradation, giving a cell

the freedom to respond to a variety of situations. An overview of the findings from the numerous biological studies that have been conducted on cell cycle control is presented here. Additionally, because the cell cycle controls the most fundamental aspects of cell growth and division, it can be linked to the dynamics of many other cellular processes important to cell culture such as protein accumulation and secretion. Therefore, in order to achieve a more complete understanding of the cell cycle, a discussion of some of the investigations that have coupled the cell cycle to the dynamics of cellular process is presented. Indeed, the framework diagrammed here offers many opportunities for further research that can be used to engineer more efficient methods of cell culture.

THE BIOLOGY OF THE CELL CYCLE

The cell cycle comprises a group of complex interrelated events that exists to ensure that cells successfully divide to form two normal daughter cells. The eukaryotic cell cycle was originally divided into two stages, interphase and mitosis, as observed by light microscopy. Interphase was defined as the period between two successive rounds of cell division, whereas mitosis was identified as the period when the cell undergoes many structural changes, divides it chromosomes, and undergoes cell division. The advent of biochemical and genetic techniques later showed that interphase consists of many cellular processes necessary for the regulation of cell proliferation that can not be discerned under the light microscope. The cell cycle was then partitioned into four separate phases: G_1 (or gap 1), DNA synthesis (S phase), G_2 (gap 2), and mitosis (M phase), where G_1 refers to the gap between mitosis and DNA synthesis, and G_2 to the gap occurring after DNA synthesis and prior to mitosis.

Understanding the complex network of biochemical events that guide a cell through successful completion of cell division is of utmost importance in understanding how cells behave in response to different situations. There has been an explosion in the amount of cell cycle related research being done in the last 10 years. This is due, in part, to the fact that a wide variety cellular processes are thought to be somehow regulated by cell cycle events and proteins. Additionally, improper control of cell cycle processes in higher eukaryotes has been linked to the formation of cancerous tumors. Here we discuss the biochemical events that occur in the cell cycle and the control mechanism that maintains the temporal order of these events.

The G_1 and G_0 Phases

Upon initial inspection, it would seem that the G_1 phase of the cell cycle is a particularly uneventful time period (or gap) between the more spectacular events that occur during mitosis and DNA synthesis. Closer analysis of this interval, however, reveals that the cell is closely monitoring its environmental and internal state to determine whether the cell has sufficient resources and an intact genome to safely complete the rest of the cell cycle. In contrast to the other phases of the cell cycle, G_1 is highly variable between different cell lines and growth conditions. For example, if certain factors are not available, some cells do not progress toward S phase, but instead may enter a resting state called the G_0 phase, or quiescence. During this time, the cell undergoes a significant reduction in RNA and protein synthesis and will remain in this arrested state, with unduplicated DNA content, until the cells receive the proper stimuli.

The sequence of events that takes place during the G_1 phase has been deduced by observing how cells respond when cultured in different environments. For instance, when cultured cells are deprived of serum containing growth factors, the synthesis of new proteins is drastically reduced and a portion of newborn cells cease to advance through the cell cycle. Through observations made using time-lapse cinematography on a culture of Swiss 3T3 cells, it was determined that only those cells that exited from mitosis less than 4 h prior to growth factor deprivation were arrested in G_0, while all other cells continued to mitosis as normal (1). This observation led to the division of G_1 into two subphases. The first subphase consists of cells that are still dependent on growth factors for cell cycle progression and are referred to as G_1 postmitotic (G_{1pm}) cells, whereas cells that can proceed through the rest of the cycle are called G_1 presynthesis phase (G_{1ps}) cells. The transition point where cells pass from G_{1pm} to G_{1ps} is commonly called the restriction point in mammalian cells (2) and is analogous to "Start" in yeast. During G_{1pm}, progression through the cycle is highly dependent on growth factors and amino acids (3) that are necessary for the production of certain regulatory proteins (called cyclins) that are discussed later. It is thought that the purpose of the G_{1ps} period is to allow cells to grow, perhaps so that later in the cycle, the cell can focus its energy on other cellular processes such as DNA replication in S phase and reorganization of the cellular infrastructure during mitosis. This also explains the fact that the G_{1ps} portion of G_1 accounts for nearly all the temporal variability in the cell cycle both within a population and between cell lines (2,3). For example, after having passed the restriction point, larger cells may proceed almost immediately to the S phase, while smaller cells may linger in G_{1ps} for up to 10 h before beginning the transition to S phase (1). Additionally, although growth factors are not required to promote S phase entry once past the restriction point, the length of G_{1ps} is significantly increased when factors such as insulin and insulin-like growth factor (IGF) are lacking in the environment, presumably due to a decrease in the net protein synthesis. Under these conditions, cells undergo mitosis with a significantly reduced protein content (2), suggesting that different growth factors play an important role in growth and progression through the cell cycle.

A culture of 3T3 cells arrested in G_0 by serum deprivation can reenter the cell cycle and proceed through the restriction point only after stimulation by the addition of platelet-derived growth factor (PDGF) (1); this process is referred to as competence. Factors that initiate competence, however, vary between cell lines (4), and several other growth factor/cell-line combinations have been described (5). The biochemical effects of competency factors such as PDGF are discussed later. Time-lapse cinematog-

raphy also revealed that cell cultures that lack serum do not contain cells arrested all at one point, but at numerous times throughout G_{1pm}. Additionally, when stimulated to proliferate, arrested cells require an extra 8 h in G_1 before proceeding to S phase, presumably so the cells can recover from G_0 (2).

The DNA Synthesis Phase

The S phase is the period when the cell replicates its entire genome so that each daughter cell will receive a complete, intact set of chromosomes. Due to the immense size of mammalian genomes, multiple bidirectional replication sites are necessary for the synthesis of new DNA to be accomplished in such a short time (approximately 8 h in most cell lines). In addition to chromosome duplication, the cell must also synthesize proteins such as histones for the formation of nucleosomes and the chromosomal scaffolding. The replication of DNA has been shown to be a highly ordered process. This was observed by culturing cells with a thymidine analog called bromodeoxyuridine (BrdU). Cells in the process of synthesizing DNA take up BrdU and integrate it into newly formed DNA. Cells can then be separated by their position in S phase, based on the amount of BrdU in the DNA, to determine the exact time of replication of certain genes (6–8). These studies revealed several aspects of the replication of certain genes. First, individual genes within a cell line are replicated at defined periods in the S phase, with many of the more active genes replicated earlier than others. Furthermore, newly replicated DNA is packed into nucleosomes a short distance behind the replication fork (1).

It is also very important that the cells not replicate the same chromosome more than once; the mechanism that ensures that this does not happen is known as the rereplication block (reviewed in Ref. 2). Early cell fusion experiments on human HeLa cells showed that, prior to S phase, nuclei could be induced to undergo DNA synthesis early, whereas nuclei that have already undergone a round of DNA synthesis do not undergo another round of DNA synthesis (3). Later cell fusion experiments showed that nuclei with replicated DNA can undergo another round of DNA synthesis before mitosis after the nuclear membrane has been permeablized and then repaired (4). A model involving a replication-licensing factor (RLF) was then proposed to describe these phenomena. DNA would be replicated only if RLF was bound to the chromatin to be broken down by the replication machinery during the next S phase. During mitosis, the nuclear envelope breaks down, allowing new RLF to access the chromosomes and bind to them, thus licensing the chromosomes to undergo a single round of DNA replication. Support for this model was gained when two protein complexes, called RLF-M and RLF-B, were shown to be required to modify chromatin before replication is allowed to proceed (5). These protein complexes were shown to bind to chromatin in late mitosis and be gradually removed during S phase (6,7). The rereplication block model described earlier has recently been modified to accommodate a more detailed understanding of the proteins and mechanisms involved (2). Chromosomes in G_1 must first be "primed" by the establishment of a pre-replication complex (pre-RC) containing an origin-recognition complex (ORC), an RLF complex, and other proteins. After initiation of DNA synthesis, the pre-RC is disrupted to become the post-RC, which still contains the ORC. The ORC may be some sort of "docking" factor that allows other components of the pre-RC to assemble on the DNA.

The G2 Phase and Mitosis

Cells enter the G_2 phase upon completion of DNA synthesis, and this period lasts between 2 and 4 h, culminating with the onset of chromosome condensation. During this time, duplication of the centriole is completed, and cells produce many mRNAs and proteins that will be used during mitosis. Cells may also use this period of time before mitosis to ensure that the cell is sufficiently large and is ready to undergo cell division.

Mitosis is generally the shortest phase of the cell cycle, lasting only 15 min to 1 h, depending on the cell line. However, visual inspection of a cell culture reveals that this is a very physically active period for the cell in that it goes through many structural changes in preparation for cell division. Mitosis is traditionally divided into six stages that describe the mechanical events that occur during this period: prophase, prometaphase, metaphase, anaphase, telophase, and cytokinesis. The beginning of mitosis, known as prophase, is marked by condensation of the chromosomes into compact structures that allow sister chromatids to be more easily separated. The cytoskeletal microtubules begin to disassemble, and the mitotic spindle, a weblike array of microtubules that mediates the events in mitosis, begins to form. Also during this phase, the Golgi apparatus breaks up into tiny vesicles, and pinocytosis and RNA synthesis stop. Prometaphase begins with the disruption of the nuclear envelope, followed by the attachment of the chromosomes to the microtubules that make up the mitotic spindle. The microtubules attach to protein structures on the DNA called kinetochores, which as we shall see are very important in regulating progression through mitosis. During metaphase, the mitotic spindle arranges all the chromosomes in a single plane across the middle of the cell, called the metaphase plate. After alignment of the chromosomes, the cell enters anaphase, which is when sister chromatids are pulled to opposite sides of the cell by the microtubules so that there is a complete genome on either side of the cell. Microtubules not attached to kinetochores then elongate, pushing the spindle poles further apart. Telophase then begins, resulting in the formation of two nuclei, one around each set of chromosomes on opposite sides of the cell. The cell then begins the process of cell division, called cytokinesis, when a ring of actin and myosin that has formed around the center of the cell begins to contract. Upon completion of this process, two new daughter cells are formed, each containing a complete, intact genome and cytoplasmic organelles sufficient to support further growth and proliferation.

The temporal organization of the cell cycle phases is shown in Figure 1, where the numbers and arrows represent when the corresponding cyclin–CDK complexes listed in Table 1 are active.

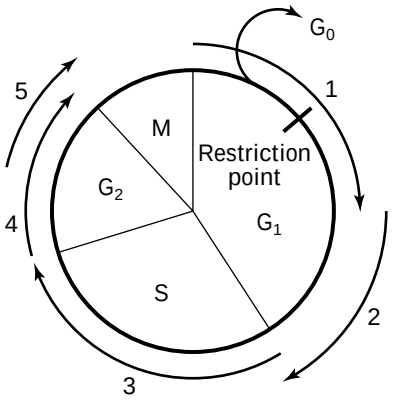


Figure 1. The temporal order of the cell cycle. Cells exit mitosis after cell division and are sensitive to the presence of growth factors until they pass the restriction point. The large block arrows depict the active period of the best-defined cyclin–CDK complexes. The numbers next to the arrows correspond with the numbers assigned to the various complexes as defined in Table 1.

CONTROL OF THE CELL CYCLE

Successful completion of cell division requires that the cell reside in an environment capable of supporting growth, and that the cell cycle proceed as normal. As discussed previously, cells require the presence of growth factors and essential amino acids to support the synthesis of new proteins. In addition, the immediate environment must not be overcrowded; daughter cells must have room to grow and subsequently divide. If any of these conditions is not met, cells enter the G_0 phase, which eventually results in cell death if environmental conditions are not improved. Similarly, if the events in the cell cycle do not occur in the proper sequence, daughter cells may not inherit the proper genetic material or organelles to properly function, and they will die. Therefore, cells must possess some means of sensing whether the environment is able to support growth. There must also be a control mechanism in place that regulates the sequence of events that make up the cell cycle. An understanding of the mechanisms that control cell growth is of interest to many scientists including oncologists and biochemical engineers. Oncologists may be able to use the knowledge of the events regulating the cell cycle to develop therapies against certain types of cancers, whereas the control network of cells in culture can be engineered to reduce the culture's dependence on serum.

Central to the mechanisms that control the events of the cell cycle are the checkpoints. Checkpoints are points in the cell cycle where certain criteria must be met before the cell cycle can proceed any further. For example, the restriction point in G_1 is a checkpoint because cells not cultured in the presence of growth factors are arrested in G_0 and not allowed to proceed through the cell cycle until growth factors are supplied. Indeed, several other checkpoints that control the order of the cell cycle phases and the events within each phase exist throughout the cell cycle. For instance, cells with unduplicated DNA do not enter mitosis, and chromosomes are replicated only once in each cycle. Additionally, cells with damaged DNA do not undergo mitosis, rather they become arrested in G_2 to allow the DNA repair machinery to correct any defects in the genetic material. Controls that mediate the order of events in mitosis also exist. For example, entry to metaphase is prevented until the chromosomes are lined up at the metaphase plate. As we will see, complexes made up of two classes of proteins are involved in regulating all these events throughout the cell cycle. These two classes of proteins are cyclins and cyclin-dependent kinases (CDKs).

Partners in Cell Cycle Regulation: Cyclins and CDKs

Cyclins were originally defined as molecules that accumulated throughout the cell cycle in marine invertebrate eggs and were then degraded during each mitosis (33). Several years later, the amino acid sequences of various cyclins were compared, and a consensus sequence of approximately 100 amino acids was found and dubbed the "cyclin box" (34). Cyclins are now defined on the basis of the presence of the cyclin box in the native sequence. Currently, 12 cyclins have been identified in mammalian cells (cyclins A–H, with two A-type, three B-type, and three D-type cyclins) (35). The cyclin box domain facilitates binding to CDKs, whose kinase activity is dependent on being bound

Table 1. Activity of the Major Cyclin–CDK Complexes

Complex (see Fig. 1)	Cyclin–CDK complex	Cell cycle phase	Function	Ref.
1	Cyclin D–CDK4	G_0, early to mid-G_1	Stimulates exit from G_0 and progression through the restriction point; phosphorylates pRb	8, 9, 10, 11–13
2	Cyclin E–CDK2	G_1–S transition	Triggers S phase entry; activates cdc25A	14, 15, 16, 17
3	Cyclin A–CDK2	S	Phosphorylates DNA synthesis machinery; inhibits E2F/DP1 transcription factor	18, 36, 37, 91
4	Cyclin A–CDK1	G_2, M	Triggers formation of mitotic spindle	22, 23, 24
5	Cyclin B–CDK1	G_2–M transition, M	Triggers entry into M; responsible for many structural changes in mitosis	18, 22, 25, 26–30
6	Cyclin H–CDK7	All cell cycle phases?	Cyclin activating kinase (CAK); phosphorylates other cyclin–CDK complexes	31, 32

Note: The numbers listed with each complex correspond with the numbers listed in Figure 1. See the text or the references for more detail on the substrates of the complexes.

to a cyclin molecule (36) and are thus defined as proteins that require binding to cyclins in order to attain activity (37). There are seven known CDKs in animal cells (named CDK1–CDK7). Different cyclins may interact with different CDKs at specific points in the cell cycle, and each of these complexes is specific for substrates with differing functions. Because several of the complexes have similar substrate specificity in vitro, the change in specificity is thought to be due not only to changes in the affinity toward different substrates, but also to localization of complexes in distinct parts of the cell by various cyclins (38). We first briefly discuss the different cyclin–CDK complexes that form throughout the cell cycle and then review the regulation of these complexes. The major cyclin–CDK complexes that are active throughout the cell cycle are listed in Table 1 along with their corresponding times of activity and their functions.

The major studies that resulted in the discovery of G_1 cyclins were performed in budding yeast (39). It was shown that three cyclin molecules, Cln1, Cln2, and Cln3, have important functions in advancing cells through the restriction point (called Start in yeast). In addition, inactivation of all three molecules is required to cause G_1 arrest of the cell cycle. By searching for factors that could complement the function of Clns in yeast, several mammalian cyclins were isolated: cyclin C, cyclin D, and cyclin E (40). Other studies revealed several other G_1 cyclins, cyclins G and F, however, neither the roles that these cyclins play nor their CDK partners are known (41). It has been shown that cyclin C mRNA synthesis reaches its peak in mid-G_1, and cyclin G mRNA synthesis starts soon after growth-factor stimulation of quiescent cells (42). Conversely, cyclins D and E are fairly well characterized and have been the focus of much attention because defects in control of these proteins have been implicated as a cause of cancer (18).

Although D-type cyclins have been shown to associate with CDK2, CDK5, and CDK6, they seem to interact primarily with CDK4 (43). Transcription of various combinations of D-type cyclins (cyclins D1, D2, and D3) is stimulated by different mitogens (growth factors), depending on the cell type, and their synthesis throughout the rest of the cell cycle is maintained by the presence of growth factors (44). Removal of growth factors halts the synthesis of D cyclins, resulting in a rapid decrease in their levels because they have very short half-lives (45). Furthermore, in cycling cells, the level of cyclin D hardly varies, and only a small peak showing its accumulation is observed. As noted previously, growth-factor deprivation has no effect on cells that have passed through the G_1 restriction point, and these cells continue to cycle until their progeny become arrested in early G_1. Additionally, microinjecting cells with cyclin D1–specific antibodies results in their cell cycle arrest only during early to mid-G_1 (8), whereas overexpression of cyclin D1 results in a reduction in the length of G_1, the size of the cell, and growth-factor dependency (46,47). These findings indicate that the primary function of D-type cyclins is to link the cell cycle machinery with the extracellular environment, which permits the cell to become committed to DNA synthesis when in a favorable environment or become arrested in G_0 under unfavorable conditions.

Cyclin E is expressed soon after cyclin D expression in G_0 cells stimulated with growth factors, and its level peaks late in G_1. It has been shown to interact with a single catalytic subunit, CDK2 (48). In contrast to D-type cyclins, cyclin E is synthesized periodically and is rapidly degraded soon after entry into S phase, thus adhering to the original definition of a cyclin (49). Similar to cyclin D, cyclin E function has been shown to be essential for progression through G_1, because microinjection of anti–cyclin E antibodies prevents cells from entering S phase, whereas its overexpression results in acceleration toward S phase (9). This acceleration is, however, balanced by an increase in the lengths of the S and G_2 phases so that the length of the entire cycle remains unchanged. Because an increased amount of cyclin E advanced progression to the S phase early, it is possible that the cyclin E–CDK2 complex is responsible for initiating the synthesis of proteins that are used during DNA synthesis. This hypothesis accounts for the increase in the length of the S and G_2 phases following increased expression of cyclin E. These findings, when taken together, suggest that the cyclin E–CDK2 complex is involved in regulating the passage from G_1 into S (41).

Upon entry into S phase, cyclin E is degraded and cyclin A takes its place as the partner of CDK2 (50). Transcription of cyclin A is begun immediately after entry into S phase, and its inhibition by microinjection of anti–cyclin A antibodies allows, at most, only 10% of the DNA to be replicated (51) and may also disrupt the temporal relationship between DNA synthesis and mitosis (10). As a cell progresses into the G_2 phase, cyclin A is no longer associated with CDK2, however, it forms a new complex with CDK1 (also known as cdc2 due to its homology with the budding yeast kinase of the same name) to regulate progression through G_2 and into mitosis. It has been shown that CDK1 forms complexes with both cyclin A and cyclin B near the G_2–M transition (52). These complexes are located in different parts of the cell because cyclin A accumulates in the nucleus, and cyclin B resides in the cytoplasm until the breakdown of the nuclear envelope during mitosis (53). Once they have completed their respective functions, both cyclins are then rapidly degraded later in mitosis.

Experiments done in fission yeast suggest that the cell monitors the presence of cyclin B, because cells undergo another round of DNA replication when cyclin B–CDK1 complexes are not present (14). This implies that the level of cyclin B–CDK1 complexes, both active and inactive, are constantly monitored, and the presence of these complexes maintains cells in G_2. Although this theory has not been investigated in other eukaryotes, it is not entirely unlikely given the high degree of conservation of cell cycle controls among eukaryotic cells.

Regulation of the Cyclin–CDK Complex Activity. Three different mechanisms allow tight control over the myriad of cyclin–CDK complexes: (*1*) cyclin transcription and degradation, (*2*) CDK phosphorylation and dephosphorylation, and (*3*) binding of the cyclin–CDK complexes to various inhibitory proteins. These three forms of cyclin–CDK regulation interact to form an intricate regulatory network that allows cells to temporally control events throughout the cell cycle and to ensure that these processes are suc-

cessfully completed. These three mechanism of cyclin–CDK complex regulation are depicted in Figure 2.

Intracellular concentrations of different cyclins, and thus the activity of cyclin–CDK complexes, are regulated by synthesis and degradation. There are two different pathways by which cyclins can be degraded, depending on the type of cyclin. Cyclins D and E, which function in early G_1 and the G_1–S transition, respectively, are inherently unstable and thus have short half-lives of approximately 30 min. The instability of these cyclins has been attributed to the presence of a PEST sequence near the C-terminus of the cyclin box (54). In contrast, cyclins A and B, which function in the S, G_2, and M phases, are stable throughout interphase, even when they are not bound with CDKs. However, these cyclins are degraded during mitosis in the proteosomes through a ubiquitin-mediated pathway (55). These proteins contain a sequence of amino acids near the amino terminus, called the destruction box, that marks them for destruction. Degradation of these proteins is also regulated so that cyclin–CDK complexes that are necessary for cell cycle progression are not inactivated before their specific tasks are successfully completed. For example, it has been shown that proteolysis of cyclin A and B does not occur in *Xenopus* egg extracts until purified cyclin B–CDK1 complexes are added (19). This suggests that the proteolysis pathway is initiated by active cyclin B–CDK1 complexes. Conversely, the addition of purified cyclinA–CDK1 complexes to the extracts had the opposite effect of delaying cyclin degradation triggered by cyclin B–CDK1 kinases (20). The fact that only cyclin B–CDK1 complexes can activate proteolysis may explain why cyclin degradation does not occur until correct assembly of the mitotic spindle is achieved in metaphase (56). The exact mechanism of regulation of cyclin degradation is unknown, but it has been hypothesized that cyclin B–CDK1 activates a cyclin-specific enzyme that catalyzes the conjugation of ubiquitin to cyclins (22).

Cyclin–CDK complexes are also regulated by phosphorylation and dephosphorylation. Specifically, phosphorylation on T160/161 (depending on the CDK) is necessary for activation of the complex, whereas phosphorylation of T14 and Y15 residues inhibits cyclin–CDK activity. The T160 residue is located on a T loop, which is conserved in all CDKs. When not phosphorylated,

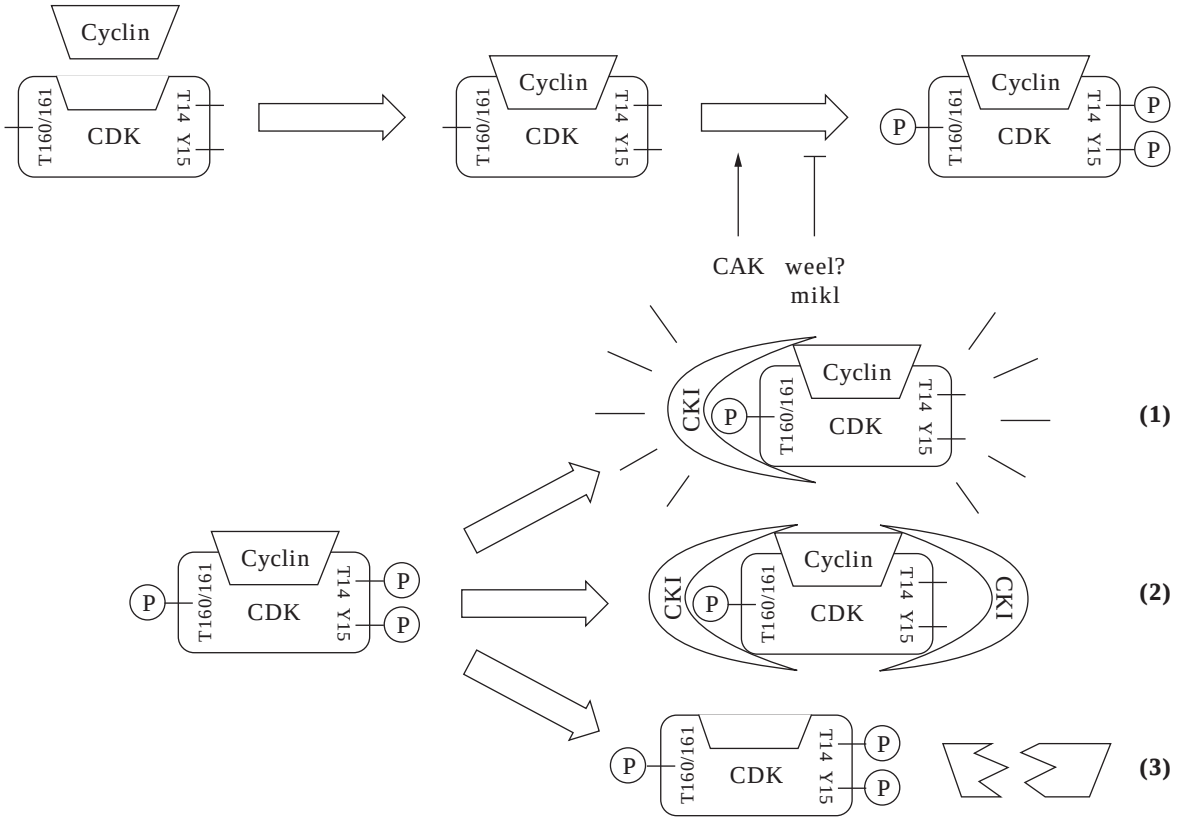


Figure 2. Control of cyclin–CDK complex activity. Cyclin–CDK activity is governed by three mechanisms: phosphorylation/dephosphorylation reactions, cyclin-dependent kinase inhibitors (CKIs), and cyclin expression and degradation. After assembly of the cyclin–CDK complex, the activity of CAK activates the complex while phosphorylation by Mik1 and Wee1 inhibits complex activity. The phosphorylated complex is then activated by the phosphatase activity of cdc25 and is denoted as complex 1 in this figure. Complex 2 has the correct residues phosphorylated, which is necessary for activation; however, it is bound by two CKIs and is inactivated. Complex 3 is shown inactivated by cyclin degradation.

this loop prevents kinase activity by covering the substrate binding domain. Phosphorylation at this site serves two purposes: it increases the affinity between the cyclin and the CDK and therefore stabilizes the complex, and it alters the conformation of the T loop, allowing access to a substrate (57).

The kinase that is responsible for phosphorylation of this residue on several CDKs is also a cyclin–CDK complex. This complex, commonly referred to as CDK-activating kinase (CAK), is made up of cyclin H, CDK 7, and a third 32-kDa protein (58). Although the level of CAK activity has not been shown to vary in a cell cycle–dependent manner, it is possible that cyclin H is regulated by its synthesis and degradation (56). The kinase responsible for activating CAK has not been identified, but it has been shown not to be an autophosphorylation process (58,59). Further studies done with this complex have shown that it may play a role in activating the RNA polymerase II C-terminus domain kinase of transcription factor IIH (60,61). This suggests that cyclin H and CDK7 may have a more diverse role than merely activating other CDKs (56).

Cyclin–CDK complexes are also subject to negative regulation through the phosphorylation of the T14 and Y15 residues. These residues are found within the ATP binding site of the enzyme. Phosphorylation of the Y15 residue inhibits kinase activity by interfering with the transfer of a phosphate group to a bound substrate (62), whereas phosphorylation of the T14 residue acts by interfering with binding of ATP (25). The regulation of phosphorylation on these residues is not yet well defined in animal cells. In addition, only the process of tyrosine phosphorylation is known in yeast because phosphorylation of the threonine residue is not necessary in this organism (35). In yeast, the tyrosine residue is phosphorylated by the Wee1 and Mik1 kinases (63). The Wee1 kinase is in turn regulated by the Nim1 kinase, which inhibits Wee1 activity. The kinase responsible for phosphorylation of the Y15 residue in animal cells has been isolated (64) and is closely related to Mik1 (35). The dephosphorylation of these residues, which is necessary for cyclin–CDK activation, is performed by the phosphatases Cdc25A, B, and C (31). Cdc25A has been shown to be activated by the cyclin E–CDK2 complex in early S phase, which then activates cyclin A–CDK2 complexes. The Cdc25C phosphatase activates cyclin B–CDK1 complexes at the transition from G_2 to mitosis (35); however, the function of cdc25B is as yet unknown.

The activity of cyclin–CDK complexes is also controlled by certain proteins known as cyclin-dependent kinase inhibitors (CKIs or CDIs) that bind to these complexes and repress their activation. Several different inhibitors have been isolated from mammalian cells that interact with either different cyclin–CDK complexes or CDKs themselves. These proteins are commonly named according to their molecular weights, such as the 21-kDa inhibitor known as p21. Other cyclin–CDK inhibitors are p27, p15, and p16. These inhibitors are intricately involved in regulating progression through the G_1 phase, and have thus been the subject of recent reviews (32,65,66).

Immunoprecipitation studies have revealed a very interesting fact: cyclin D1 is commonly found among a quaternary complex of proteins that includes either CDK4 or CDK6, p21, and proliferating-cell nuclear antigen (PCNA) (discussed later) (43). The p21 protein acts as an inhibitor by binding directly to cyclin B–CDK1, cyclin E–CDK2, cyclin A–CDK2, and cyclin D–CDK4 complexes. These cyclin–CDK complexes containing p21 are, oddly enough, still enzymatically active in vitro, which directly conflicts with studies showing that p21 is an inhibitor of CDK activity and is able to induce cell cycle arrest (67,68). It was later demonstrated that active cyclin–CDK–p21 complexes isolated from cycling cells could be inactivated by the addition of more p21 (69), suggesting that the stoichiometry of p21 binding to these complexes is very important. For example, in cycling cells, there is only enough p21 present so that each cyclin–CDK complex has a single p21 molecule and thus remains active. However, when the amount of p21 is increased or the cyclin–CDK pool depleted, the ratio of p21 to cyclin–CDK complexes increases so that more than one p21 molecule binds to each complex, resulting in inhibition of the complex's kinase activity. We explore the circumstances under which this sort of inhibition becomes important in a later section.

A second inhibitor that has been identified is called p27. p27 and p21 have been shown to have highly similar amino acid sequences near their amino terminus domains. In addition, they also share the same range of specificity for cyclin–CDK complexes. The cell cycle arrest of certain cell lines in response to some environmental signals such as contact inhibition and transforming growth factor β (TGF-β) has been shown to be due, in part, to inactivation of cyclin–CDK complexes by this inhibitor (70), and the exact mechanism of this type of arrest will be described shortly. A recent discovery has given some insight into one of the mechanisms by which this protein inhibits cyclin–CDK activity. Cyclin–CDK complexes that have not yet been activated by CAK phosphorylation can be bound by p27, which then blocks the CAK's access to the threonine residue found on the CDK subunit (71). Presumably, another mechanism exists for inhibition of active complexes.

The main function of the p27 inhibitor seems to be to prevent the G–S transition from occurring when the need arises by inhibiting the activity of cyclin E–CDK2 complexes (72). In normal cycling cells, the p27 inhibitor seems to be dispersed among the different cyclin–CDK complexes that are active during G_1, therefore it does not have a significant impact on any essential events. However, in response to stimulation by TGF-β, it has been suggested that the expression of CDK4 is downregulated, resulting in an increased amount of p27 bound to cyclin E–CDK2 complexes (73).

The p16 inhibitor was originally found associated with CDK4 in a transformed cell line (74). Unlike the p21 and p27 inhibitors, p16 was found to associate specifically with the cyclin D CDKs, specifically, CDK4 and CDK6. It competes with cyclin D for the CDK subunits and may even act to disrupt cyclin D–CDK complexes. This inhibitor is thought to play an important role in regulating growth because the p16 gene has been shown to be deleted in many tumor-cell lines (75).

Examination of TGF-β-arrested cells led to the discovery of another inhibitor, p15, because it was shown to be bound to CDK4 (76). Similar to p16, p15 binds to CDK4

and CDK6, competes with cyclin D, and may be involved in arresting cells treated with TGF-β in G_1. TGF-β stimulates the synthesis of p15 so that its level is increased approximately 30 times. This allows p15 to sequester the CDK4 complexes so that a much larger amount of the p27 inhibitor binds to and inactivates cyclin E–CDK2 complexes, resulting in G_1 arrest.

The Cyclin–CDK Complexes in Action

Now that we have discussed the events that occur in the cell cycle and defined the major regulatory proteins, we are in a position to examine the mechanisms that control cell cycle progression. Indeed, control of the activity of the cyclin–CDK complexes is at the heart of cell cycle regulation because these complexes either directly or indirectly act on cellular constituents that are capable of bringing about changes in intracellular events. This implies that the proper regulation of these cyclin–CDK complexes is essential for maintaining the temporal integrity of cell cycle events. This is accomplished by repressing the activity of the appropriate kinase through different combinations of cyclin degradation, phosphorylation/dephosphorylation of specific CDK residues, and action of CDK inhibitors. These controls not only maintain temporal stability, but also ensure that DNA synthesis is not carried out when DNA is damaged or that mitosis does not occur in the presence of damaged or misaligned chromosomes.

G_1 to S Phase Control. An important property in normal cells is their ability to sense the state of the extracellular environment, such as the presence of various growth or inhibitory factors or contact inhibition. This communication between the cell and its surroundings takes place in the G_0 and G_1 phases, and the cells respond by becoming quiescent, undergoing programmed cell death (known as apoptosis), or continuing with the cell cycle. The interaction with the environment ceases when the cell passes the restriction point, after which the cell is said to be committed to DNA synthesis. We now consider the machinery responsible for directing the path that the cell will traverse in response to signals from the environment.

Several lines of evidence indicate that the product of the retinoblastoma gene, pRb, is responsible for linking signals from the extracellular domain (77). Specifically, mutation or deletion of this gene is found in many cancerous cell lines that have lost their ability to regulate proliferation (78). Additionally, pRb accumulation was found to coincide with suppression of growth in TGF-β-induced arrest (79), whereas inhibition of pRb results in a resumption of cell cycle progression. Mammalian DNA viruses, whose main purpose is to activate quiescent cells so that they can use the cell's replication machinery to duplicate their own DNA, have been found to seek out the hypophosphorylated form of pRb (80). This led to the belief that pRb is in its active state when it is underphosphorylated, and phosphorylation of this protein results in its inactivation. This is supported by the fact that hypophosphorylated pRb is present in G_0 cells and is phosphorylated later during the G_1 phase. In its active form, pRb binds to several proteins including a heterodimer of proteins from the E2F and DP family of proteins. This heterodimer, called DRTF1/E2F (81), is a transcription factor that stimulates the transcription of numerous proteins necessary for cell cycle progression, such as CDK1, cyclin A, and a variety of proteins that are indispensable during DNA synthesis (82).

In order to further characterize G_1 events, it is necessary to find out what is responsible for inactivating pRb. A likely candidate seems to be the cyclin D–CDK4 complex. The presence of this active kinase is highly dependent on the presence of growth factors that stimulate the synthesis of cyclin D, and thus, the activity of that complex appears during the same period when pRb is inactivated. Moreover, cells that lack functional pRb do not require the presence of cyclin D–CDK4 to proceed through the cell cycle, implying that the primary function of this complex is to inactivate pRb (83). Indeed the belief that pRb is a substrate of this complex was confirmed because cyclin D–CDK4 is capable of binding to and phosphorylating pRb in vitro (84,85). The following feedback loop, involving the levels of cyclin D and the functional state of pRb, has been proposed (86). Cells that lack functional pRb have drastically reduced levels of cyclin D; however, addition of pRb to these cells results in the synthesis of cyclin D. Active kinase complexes are then allowed to form, which act by phosphorylating pRb. The hyperphosphorylated form of pRb downregulates the levels of cyclin D during late G_1, causing the inactivation of CDK4, which in turn allows pRb to be activated once again by the action of phosphatases, and the cycle begins anew upon completion of mitosis.

Because the CDK inhibitors p21 and p27 are constitutively expressed in cycling cells, these molecules are present throughout G_1. This indicates that in order for cells to advance past the restriction point, the level of cyclin D–CDK4 complexes must increase until it exceeds the number of inhibitor molecules present so that there is at most one CKI per cyclin–CDK complex to allow kinase activation (70). This model not only satisfies the timing of cyclin D–CDK4 activation, but suggests a dual role for this complex. Its primary function is to phosphorylate pRb to promote the transcription of important genes, while its secondary role is to sequester the CKIs so that the cyclin E–CDK2 complex, which is the next kinase to be active in the cell cycle, cannot be inhibited by these molecules, thus allowing the transition to S phase to occur (32). A schematic representation of the control mechanism that regulates progression through G_1 is shown in Figure 3.

As mentioned previously, the activity of cyclin E–CDK2 complexes during the cell cycle has been narrowed down to the transition from G_1 to S. Some of the possible substrates for this complex include the so-called pocket proteins, which include pRb, p107, and p130 (82). The term *pocket protein* refers to those proteins involved in binding to certain proteins, such as transcription factors, and releasing them after activation of the pocket protein by the action of a CDK kinase. The cyclin E–CDK4 complex is localized in the nucleus where it is bound to E2F, p107, and p130 (although it has not been shown to phosphorylate these proteins) (87), suggesting that it too plays a role in regulation of transcription. The E2F subunit that interacts with cyclin E–CDK4 has been shown to differ from the E2F found associated with pRb, which hints that they have dif-

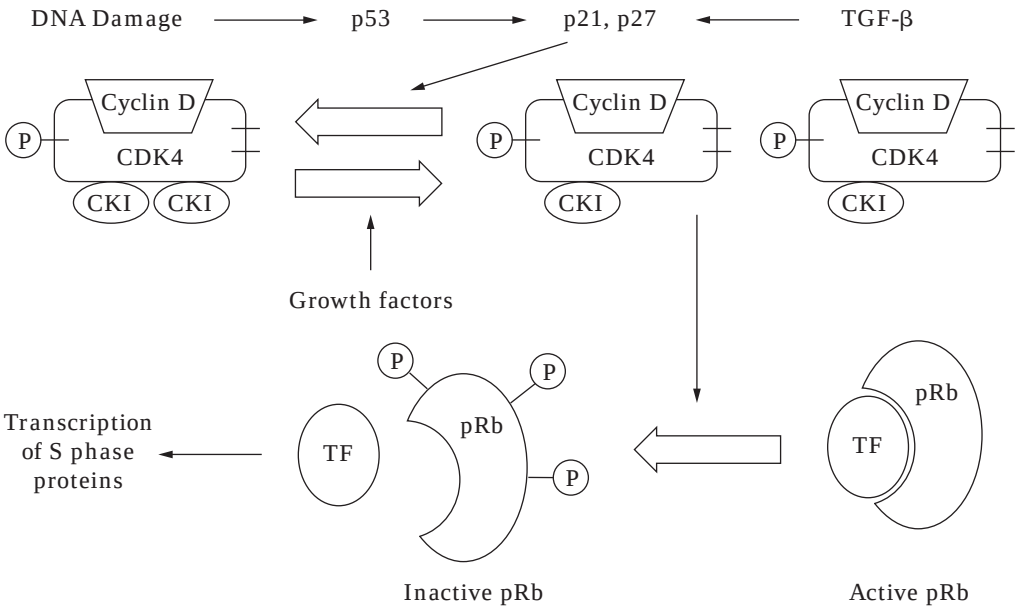


Figure 3. The G_1 control mechanism. Schematic of the control mechanism that governs the progression past the restriction point and through G_1. The presence of growth factors stimulates the expression of cyclin D to promote the formation of cyclin D–CDK4 complexes, increasing the amount of complex to a level above the inhibitory threshold of CKIs. This increase in cyclin D–CDK4 complexes results in kinase activity that then promotes the release of transcription factors. DNA damage results in the expression of p53, which increases the amount of CKIs present to arrest cells in G_1 and allow DNA repair to take place.

ferent properties (88). One confirmed substrate of cyclin E–CDK2 is the Cdc25A phosphatase that is involved in activating CDKs by cleaving the inhibitory phosphates on both a tyrosine and a threonine residue (89). The cdc25A phosphatase has been proposed to activate both cyclin E–CDK2 and cyclin A–CDK2 complexes, creating a positive feedback loop between cyclin E–CDK2 and Cdc25A (Fig. 4a). If this feedback loop does indeed exist, it has been hypothesized that the activation of Cdc25A occurs in response to the completion of some G_1 event such as the duplication of the centrosome or the accumulation of certain DNA synthesis proteins (15).

Experiments with yeast suggest that proteolysis may be involved in activating the cyclin E–CDK2 homologue that plays an important role in the G_1–S transition (11). It has been shown that a protein called Cdc34 is necessary for the G_1–S transition to take place, possibly because it is responsible for initiating the degradation of certain CKIs. Functional homologues of the Cdc34 protein have been isolated in vertebrate cells and are thought to stimulate the degradation of the p27 inhibitor (12). This, in addition to isolation of CKIs by the cyclin D–CDK4 complex, could possibly aid in beginning the onset of the synthesis phase.

In summary, a cell becomes arrested in G_0 in response to some extracellular signal that triggers a decrease in the level of cyclin D below a threshold level of CKIs. This allows p21 and p27 to bind to cyclin D–CDK4 complexes in a 2:1 ratio, resulting in an inhibition of kinase activity. Following stimulation of cyclin D synthesis due to the presence of growth factors, there is an increase in cyclin D–CDK4 kinase activity because there is no longer enough inhibitor present to inactivate all the complexes. This active CDK kinase then serves two purposes: (*1*) to inactivate pRb, allowing the release of transcription factors that induce the synthesis of other cyclins and DNA synthesis proteins, and (*2*) to isolate the CKIs to allow activation of the cyclin E–CDK2 complex. This complex then stimulates the release of additional transcription factors and initiates a positive feedback loop that triggers the onset of the DNA synthesis phase.

Once cells move beyond the restriction point, they are committed to entering the DNA synthesis phase. Cells do, however, have the power to delay the onset of later events to ensure the proper coordination of important processes. For example, failure to repair damaged DNA greatly increases the occurrence of genetic mutations, which can have disastrous consequences for the cell. Thus, an additional checkpoint exists in late G_1 phase, which ensures that the DNA is free from damage and ready to be replicated. We will now see that cells possess the means to sense when there is damage to the genetic material and then halt cell cycle progression to allow the DNA machinery time to repair any defects in the genetic material. It has been determined that the there is a single protein, p53, that is held accountable for inducing G_1 arrest in response to DNA damage (13) (see Fig. 3 for a schematic representation).

By inducing DNA damage in normal cells, it was determined that cells responded by increasing the level of p53 and the CKI p21 (90). The increase in p21 then leads to

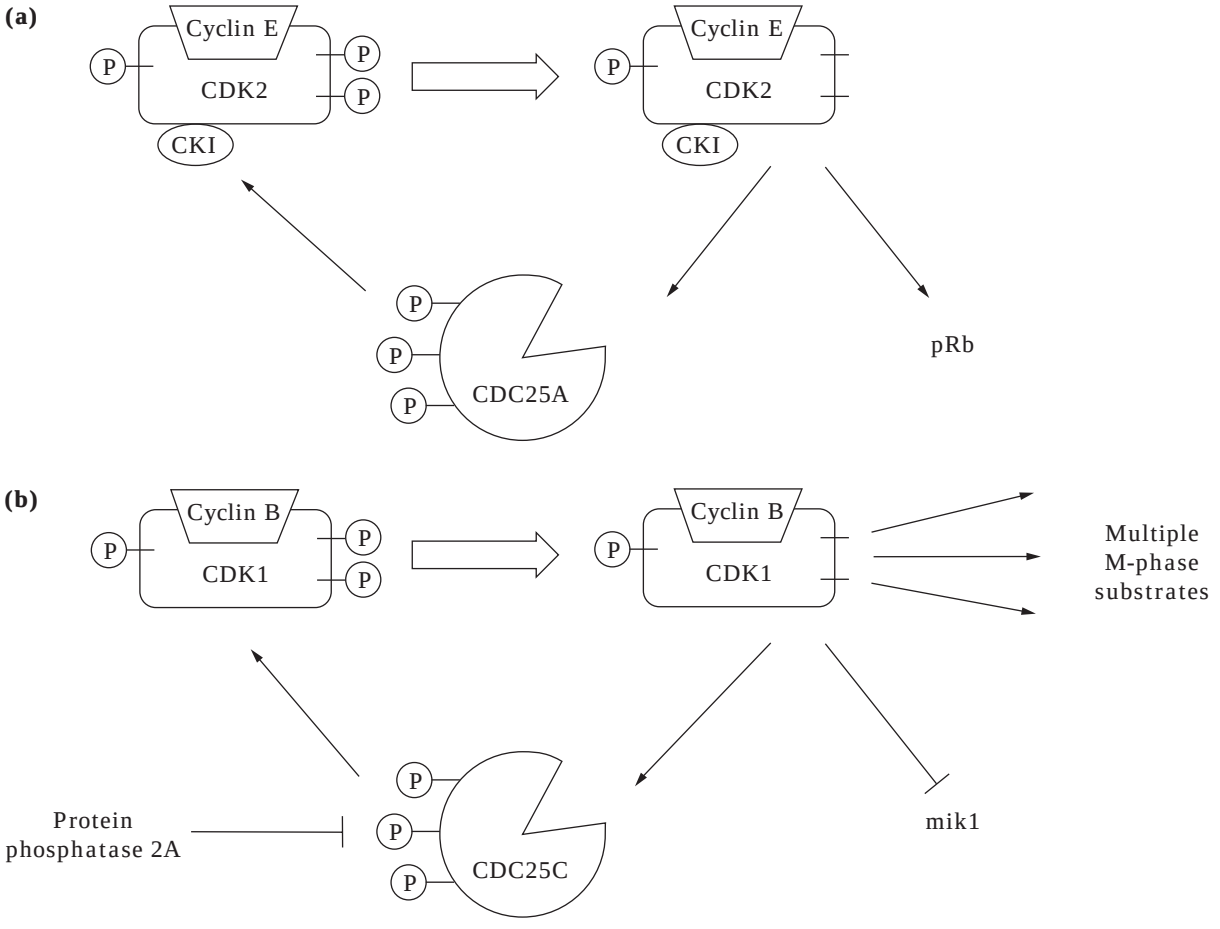


Figure 4. The S- and M-phase entry mechanisms. The positive feedback mechanisms that trigger the entry into S phase (**a**) and mitosis (**b**). The transition to S phase is triggered when an active cyclin E–CDK2 complex activates the Cdc25A phosphatase, which in turn activates more cyclin E–CDK2 complexes, causing a sharp rise in active complexes. These complexes then phosphorylate pRb to stimulate the release of transcription factors. The M-phase transition is triggered by a positive feedback mechanism between cyclin B–CDK1 complexes that is similar to that described in (**a**). In addition, the active cyclin B–CDK1 complex inactivates the mik1 kinase, which phosphorylates the complex on an inhibitory residue, thus contributing to the steep rise in cyclin B–CDK1 activity.

the inhibition of both cyclin D–CDK4 and cyclin E–CDK2 complexes to prevent entry into the S phase. Additionally, p21 also functions to oversee the proper repair of DNA by interaction with PCNA. PCNA is required for both DNA synthesis and repair. When bound with p21, PCNA is no longer able to interact with polymerase-δ, thus preventing DNA synthesis, however, the DNA repair function remains unaffected by the inhibitor (16). Therefore, p21 must inhibit both CDK activity and interaction between PCNA and polymerase-δ in order for DNA repair to be carried out before advancing to the S phase.

S Phase and G_2 Progression. Shortly after the cyclin E–CDK2 complex triggers entry into the S phase, cyclin E is degraded and CDK2 interacts with cyclin A. There is evidence that cyclin E may be degraded in a ubiquitin-mediated pathway, as in budding yeast, but this has not been confirmed (91). The cyclin A–CDK2 complex is localized in the nucleus where it associates with p107, p130, and E2F (17). Interacting with E2F allows the kinase to phosphorylate the DP1 subunit of the transcription factor complex, which inhibits its DNA binding capability (92). This suggests that when cyclin A synthesis begins in early S phase, it aids in halting the transcription of genes triggered by the E2F complex in the late G_1 phase. No other substrates of this cyclin–CDK complex have been identified except for helicase RF-A, which is one component of the DNA synthesis machinery (35).

Upon completion of DNA synthesis, CDK2 is degraded and cyclin A forms a new complex with CDK1 (93). It is not known if this complex has any function in the G_2 phase; however, its activity has been implicated in triggering initial events that occur in mitosis. During G_2, cyclin B begins associating with CDK1 in the cytoplasm and is transported to the nucleus just prior to mitosis. Mitosis begins when the cyclin B–CDK1 complex is activated, and it has been

suggested that a checkpoint mechanism exists that prevents activation of this complex until duplication of the centriole has been completed (94). One model that describes this transition involves the regulation of Cdc25C activity by downregulating protein phosphatase 2A (PP2A), which controls Cdc25C activity (15). Signals could be initiated from the DNA replication machinery or from the proper organization of chromatin, which would inhibit the action of PP2A. This allows Cdc25C to become activated and dephosphorylate the inhibitory residues on CDK1, causing its activation. A positive feedback loop is then initiated (Fig. 4b), whereby active cyclin B–CDK1 complexes activate more Cdc25C and, in budding yeast, inactivate the Wee1 kinase by phosphorylizing a residue that inhibits its action (95). This mechanism brings about a sharp transition into mitosis almost immediately after a small amount of CDK1 is activated.

Another mechanism exists that allows cells with damaged DNA to become arrested in G_2. In response to DNA damage, transcription of the CHK1 kinase is initiated. This kinase targets Cdc25C and phosphorylates the serine-216 residue, which allows formation of a binding site for the 14-3-3 protein that inhibits Cdc25C activity (96). This then allows the DNA repair machinery to work because the cyclin B–CDK1 complex can not be activated by cdc25C.

Supervising the Order of Events in Mitosis. The point when the cyclin A–CDK1 complex is most active occurs just after the cell enters into mitosis, which coincides with nucleation of the centrosomes and the formation of the long microtubular extensions of the mitotic spindle. Indeed it has been shown that cyclin A–CDK1 activity has a profound affect on the nucleating ability of the centrosomes in *Xenopus* extracts (97). Once spindle formation is complete, cyclin A begins to be degraded and the activity of cyclin B–CDK1 complexes takes over. This complex is either indirectly or directly involved in regulating almost all of the structural reorganization that takes place during mitosis.

Chromosome condensation begins in prophase when cyclin B–CDK1 phosphorylates histone H1, transcription factors, and several other proteins (98). Phosphorylation of these proteins is thought to weaken the interaction between these proteins and the DNA, allowing the chromosomes to pack more tightly (21). In addition, after formation of the mitotic spindle is complete, the cyclin B–CDK1 complex reduces the nucleating ability of the centrosomes, which causes the microtubules to shorten to a length suitable for the alignment and separation of the chromosomes (35). The breakdown of the nuclear envelope, which marks the beginning of prometaphase, and the reorganization of the cytoskeleton are also brought about by phosphorylation. The nuclear lamina are present in a hypophosphorylated form in an intact nucleus, and when phosphorylated by the cyclin B–CDK1 complex, the lamina network breaks down, contributing to nuclear disassembly (99). Lamina phosphorylation itself is not sufficient to achieve disruption of the nucleus, however, and these mechanisms remain unclear (56). Caldesmon has also been shown to be a substrate of the mitotic CDK complex, and microtubular reorganization ensues after its phosphorylation. Caldesmon binds to actin and calmodulin, and its phosphorylation weakens its affinity for actin, causing the microfilaments to dissociate (23). This process causes the cells to take on a round shape that is characteristic of anchorage-dependent cells undergoing mitosis. The activity of the cyclin B–CDK1 complex is also responsible for repressing the formation of the contractile bundle until the cell enters anaphase (100). The CDK1 complex phosphorylates myosin, which prevents its association with actin filaments so the contractile bundle is unable to form. This inhibitory function of the CDK complex stays in effect until cyclin B is destroyed in the ubiquitin degradation pathway. The activity of the cyclin B–CDK1 complex is, ironically, responsible for the initiation of its own destruction (101), which stays in effect until this pathway is shut down in late G_1, possibly by the action of the cyclin E–CDK2 complex (102).

The final regulatory mechanism that exists in the cell cycle is a checkpoint that prevents the onset of anaphase until several other structural events have taken place in the cell. Some of these events include the assembly of a bipolar spindle, attachment of the chromosomes to the spindles through the kinetochore, and the migration of the chromosomes to the metaphase plate. This suggests that proper completion of these processes is somehow monitored to ensure that anaphase does not begin prematurely (24). Several experiments (recently reviewed in Ref. 26) suggest that one of two events signals that the cell is ready to undergo the chromosome separation of cell division. These events are either the attachment of the chromosomes to the mitotic spindle (27) or the presence of tension on the kinetochores due to their attachment to a bipolar spindle (28). Because all the structural reorganization processes that occur in mitosis depend on the successful completion of the other events, failure to complete any one of these processes could result in the failure of the signaling event to occur. This would then stimulate a cascade of phosphorylation events, culminating in cell cycle arrest in mitosis.

Understanding Cell Cycle Regulation through Mathematical Modeling

Understanding the complex array of biochemical events we have outlined can be difficult at times. Moreover, our knowledge of the controls responsible for cell cycle regulation is rapidly becoming more and more complicated. This complexity makes it extremely difficult to understand the mechanisms involved through logical thinking alone. Mathematical modeling has been used for the purpose of explaining the various physiological events that are commonly observed in different situations, and the events observed in *Xenopus* oocytes and yeast cells. The cell cycle controls discussed in this review have been described for animal cells. However, the majority of initial experiments on the cell cycle were carried out in yeast cells and in *Xenopus* oocyte extracts. The components that make up the control systems in these organisms are largely identical to those of animal cells, and many findings resulting from studies on these organisms have been extrapolated to apply to animal cells. An initial review of the literature would suggest that the control systems in these organisms are

different; however, this is not the case—only the terminology used to represent the components differs. The cell cycle nomenclature used in these organisms has been reviewed elsewhere (35,56).

Modeling of the molecular controls that are thought to control the cell cycle in *Xenopus* oocytes (29) and yeast cells (30,103,104) has been done in order to determine the effect of changing the activity of various components of the regulatory network. We now believe that progression through the cell cycle is governed by the activity of cyclin–CDK complexes, in other words, a sharp rise in the complex's activity results in a transition to the next phase. The aim of the molecular modeling done by Novak and Tyson was to determine whether the control mechanisms presented here could be used to describe the observed behavior of cells in a variety of situations, or more simply stated, to compare theory with experiment. In order to do this, a set of equations (one for each component of the system) was generated, based on the mechanisms described, to describe the rate of change of each component as a function of time by translating the proposed control mechanisms into simple reaction rate equations. The resulting set of differential equations was then solved using numerical techniques; the activity of the cyclin–CDK complexes could be followed with time, and the sharp oscillations in complex activity could be varied. Because sets of equations commonly contain more than one solution (depending on rate constants and initial conditions used), phase-plane plots were constructed. Observation of limit-cycle-type behavior in these plots indicates an oscillatory solution that is characteristic of proposed cell cycle control (see Refs. 30 and 105 for details). This type of model is easily solvable by today's personal computers and has been successfully implemented to describe the observed behavior of a variety of yeast cell cycle mutants (103,104) and embryo and oocyte maturation (29).

IMPLICATIONS FOR ANIMAL CELL CULTURE

The use of large-scale cultures of animal cells is becoming increasingly important as the demand for their biological products grows. Animal cell cultures are used for the production of many different products of diagnostic and therapeutic value, including vaccines, monoclonal antibodies, and in some cases, the cells themselves (for example, to replace the function of damaged tissue). Determining the culture conditions that optimize the production and secretion of proteins is therefore of utmost importance in meeting these demands. Consequently, the key to designing such systems is not only determining the kinetics of growth, secretion, and proliferation, but also understanding how the cell cycle regulates certain cellular processes. The heterogeneous nature of cell populations has made it extremely difficult to obtain these data. Various mathematical models have been developed to describe cell behavior under different conditions; however, they are, for the most part, unable to account for the heterogeneous nature of cell populations and usually do not provide insight into variations introduced by the cell cycle. One model used to describe cell growth within a heterogeneous population of cells, known as population balances, has been developed (106–109). However, the implementation of this approach has been hampered by the difficulty in obtaining the single-cell parameters necessary for its use.

These problems are being tackled through the development and improvement of tools such as flow cytometry and fluorescent markers, which allow the examination of single-cell parameters within a heterogeneous population. Flow cytometry is a very sensitive tool that allows one to simultaneously obtain several single-cell parameters related to growth physiology from a large number of cells in the population. Collecting these data allows one to determine the size, DNA content (and thus cell cycle position), and protein content (when the protein is conjugated to a fluorescent marker) of individual cells in a population. Many flow cytometric studies relating the cell cycle to optimization of cell culture performance are available in the literature. Several of these analyses and their findings are outlined next.

Cell Cycle–Dependent Protein Accumulation

In the area of cell cycle–dependent protein accumulation, hybridoma cells have received the most attention because they are widely used in the production of monoclonal antibodies (MAbs). Numerous studies have suggested that lower rates of growth result in higher levels of protein production (110–112) (see also papers referenced in Ref. 113). It has been argued that because populations growing at slow rates have a larger fraction of G_1 cells, the G_1 phase must be a period of enhanced protein production and secretion. However, few studies have been done to give insight into the mechanism for this behavior. Additionally, other studies have been done that show a positive correlation between growth rate and protein production (briefly discussed in Ref. 113). These variations in response observed by different researchers may be explained by the fact that the overall growth rate of a culture depends on both cell growth and cell death. This means that the same overall growth can be observed in two cultures with different rates of cell loss and cell growth. The conclusion that can be made from the examination of these different investigations is that it is important to be aware of the different methods used to quantify growth and protein production in cell cultures because they may result in conflicting results.

Al-Rubeai and colleagues used flow cytometric results obtained from synchronous cultures and concluded that the rates of both immunoglobulin synthesis and secretion are at a maximum at the G_1–S transition (114). The synchronization of the culture was achieved by the addition of thymidine, which arrests cells prior to entry into the S phase. This type of analysis presents another problem, however, because factors used for cell cycle arrest have been shown to alter cell metabolism (115); therefore these results should be viewed with care. A more complete series of studies on rates of MAb production during the cell cycle was performed by Kromenaker and Srienc. The first study involved using flow cytometric data obtained in asynchronous cultures and applying these data to population balances to find the single-cell rates of protein accumulation

in various cell cycle phases for both antibody-producing and non-antibody-producing cell lines (116). It was found that during the early G_1, producer cells with low protein content have a specific rate of accumulation of protein approximately 3–5 times greater than the specific rate observed for the remainder of G_1 cells. This indicates that the rate of protein synthesis is much greater than the rates of protein degradation and secretion in G_1. The S phase of the cell cycle was shown to a have a low specific rate of protein accumulation that remained relatively steady throughout this period. Extending this analysis to the $G_2 + M$ phase yielded a negative specific rate of protein accumulation, indicating that protein is lost due to secretion and degradation during this period of the cell cycle. From these findings, it was concluded that the disruption of the secretion system that occurs during the division of the cell accounts for the extremely high rate of accumulation observed in early G_1 because the transport system probably has not had time to recover from its disruption.

A later report by Kromenaker and Srienc described the relation between growth rate and secretion rates during the cell cycle as determined using single-cell data and population balances (117). Lactic acid was used as a growth suppressor, and the concentrations were changed to vary the specific growth rate. This study brought several previously unconsidered factors to light. By analysis of the DNA distribution of the populations at different growth rates, it was found that the lengths of all the cell cycle phases were increased by the same factor that the specific growth rate was increased. This is in direct contrast to other studies that suggest that the increase in productivity is due to the increase in the length of the most productive phase of the cell cycle. Secondly, the application of population balances in this study revealed that the net specific rate of antibody synthesis is independent of growth rate. Additionally, the specific secretion rate of $G_2 + M$ cells is greater than the rate in the other two phases and remains constant with decreasing growth rates. In contrast, the specific secretion rates for G_1 and S increase with decreasing growth rate. These findings are consistent with the hypothesis that the disruption of the secretion pathway during cell division results in lower secretion throughout G_1, until the Golgi apparatus has had sufficient time to reassemble. During this time, antibodies accumulate in the cell until they can be secreted during $G_2 + M$, before the Golgi apparatus is disrupted once again. Thus, at lower growth rates, more time is spent in G_1, so an increased secretion rate is observed because the cells have had time to recover from cell division. A subsequent study tracked the amount of cell surface antibodies as a function of DNA content (118). Because prior experiments determined that the amount of cell surface antibodies observed on the membrane is a good measure of the rate of secretion, these data were used to determine the pattern of secretion during the cell cycle. The results obtained from this method also agreed with the hypothesized regulation of secretion already described.

Understanding the regulation of foreign genes in animal cells is also of great interest because recombinant DNA is commonly used to express proteins of interest in immortal cell lines such as Chinese hamster ovary (CHO) cells. This provided the motivation for studies in which the expression of foreign genes in animal cells was tracked throughout the cell cycle. One study tracked the secretion of tissuelike plasminogen activator (tPA) in recombinant CHO cultures synchronized by fluorescence-activated cell sorting (FACS) based upon DNA content (119). This study shows that secretion of tPA reaches its maximum near late S phase, while the increase in this protein is maximum in G_1. These results are fairly similar to those obtained by Kromenaker and Srienc using different methods. Gu and colleagues also measured the expression of foreign β-galactosidase during the cell cycles in recombinant CHO cells (120) and in recombinant mouse cells (121) and hypothesize that control of foreign gene expression is promoter dependent. These particular studies, however, showed that the use of either the cytomegalovirus (CMV) promoter or the mouse mammary tumor virus (MMTV) promoter causes the production of β-galactosidase mainly in S phase.

Cell Cycle–Dependent Behavior of Culture Systems

There is a steady increase in the number of reports in the literature that look to the cell cycle to explain various phenomena such as differing growth rates between cultures and cell death due to shear stress in bioreactors. For example, the effect of different inoculation sources was examined by Al-Rubeai et al. (122). Two cultures that were inoculated with cells from either the exponential or stationary growth phase were examined using flow cytometry. The culture inoculated with cells from the exponential phase grew at a much greater rate than the other culture and therefore reached the death phase more quickly due to nutrient exhaustion and buildup of toxins. Analysis of the cell cycle, cell size, and mitochondrial activity of the cultures revealed differences among the cultures. For instance, the slower-growing culture was found to have a larger proportion of its cells in G_1, while the culture with the greater growth rate had a larger fraction of its cells in S. Because G_1 cells are generally smaller than the S-phase cells, the culture with a larger fraction of cells synthesizing new DNA uses more nutrients than the one with smaller cells. Additionally, cultures with elevated mitochondrial activity were also shown to have a greater capacity for proliferation. These findings lead the researchers to conclude that examination of these parameters can provide a means of predicting how the culture will behave and a means of explaining variations in the growth patterns of cultures.

It is also of great interest to determine how cells behave in different culture systems. For cells in continuous culture (123), it has been found that the fraction of cells in the G_1 phase decreases with increasing growth rate, while the fraction of cells in the S phase increases. As expected, the G_1 phase accounts for most of the resulting variability in the lengths of the cell cycle phases; however, the S phase is affected slightly more than the G_2 and M phases. A more detailed analysis of the cell cycle was performed in cultures containing microcarriers, and the fraction of contact-inhibited cells was determined and compared to a model (124). Microcarriers are commonly employed in cultures of anchorage-dependent cells, such as CHO cells, because

these types of cells must be attached to a surface in order to proliferate. When the culture grows to the point where there is no longer any surface area available for newborn cells, contact inhibition occurs, and a portion of the cells cease to proliferate even when provided with an ample supply of nutrients. Analysis of this system yielded the distribution of cells in the cell cycle during the lag phase, exponential growth, and confluency. During the lag phase, the culture is predominantly in the G_1 and/or G_0 phases and rapidly enters the other cell cycle phases upon reaching exponential growth. As the culture reaches confluency, the culture gradually shifts back to predominantly the G_1 and/or G_0 phases; however, a portion of the cells continue cycling. These findings are similar to those of suspension cultures. Furthermore, a method was developed to determine the fraction of cells affected by contact inhibition, and the results showed good agreement with a model that had been developed.

Furthermore, when culturing animal cells in large-scale bioreactors, the cells can be injured due to the shear stress that results from agitation and sparging. There have been several studies that examined the mechanism of cell damage and sought to minimize injury to the cells through the use of various shear-protecting agents such as Pluronic F-68 and serum. However, there have been relatively few studies that examine the effect that these forces may have on cell metabolism and growth. It is evident that shear forces have an effect on the overall growth rate of the culture due to an increase in cell death. Lakhotia and coworkers (125) analyzed the rate of DNA synthesis and the cell cycle phase distributions of cultures grown in small-scale bioreactors with gas sparging and varying agitation intensities to determine the effect that these forces may have on growth kinetics of the surviving cells. Measurement of the kinetics was accomplished by measuring the uptake of BrdU and the DNA content distributions of the surviving cells with the aid of flow cytometry. They determined that, in cultures exposed to high agitation, the fraction of cells in the S phase increased by up to 45%, while the fraction of cells in G_1 decreased by up to 50%, compared to control cultures. The simultaneous analysis of the DNA synthesis rate suggests that the change in the cell cycle phase distribution is due to an increase in the DNA synthesis rate. Additionally, when cultured in the absence of shear stress, the apparent specific growth rate of the cells previously exposed to the shear forces was initially higher than and later the same as that of control cells. It was also shown that an increase in the agitation intensity resulted in a further increase of the DNA synthesis rate. By subjecting cells to extreme, short-term stresses induced by turbulent capillary flow, it was determined that larger cells are more susceptible to damage than smaller cells (126). This conclusion was made based on the preferential disappearance of S and G_2 cells when passed through the capillary. Another possible explanation for these results is that an apoptotic response triggered by the shear forces caused a decrease in the size of the cells, but verification of this hypothesis is difficult.

Engineering the Cell Cycle

Once a more complete picture of how the cell cycle regulates cell physiology and metabolism has been painted, it can possibly be manipulated in culture in such a way that cells will perform desired processes at an optimal level. For example, controlling the regulatory mechanisms of cell growth and proliferation, or controlling the rate at which cells advance through the cell cycle can lead to the enhancement of protein production, improved economic feasibility, and ease of downstream processing. Indeed there have been several studies published to date that address these issues and will undoubtedly lead to further examination of directing cellular operations through the engineering of the cell cycle in cell cultures.

Cells in culture are subjected to many different forms of stress that may ultimately lead to cell death. Cell death can be caused by external factors such as hydrodynamic forces that may cause the cell to rupture. Additionally, the response to certain stimuli such as limitation of glutamine, glucose, or serum or the buildup of toxins may result in programmed cell death, called apoptosis, which is regulated by some of the same pathways as is the cell cycle. When in culture, cells reach their maximum cell density in a few days, after which the process of programmed death is induced in the majority of the cells (127). The discovery that the product of the protooncogene *bcl-2* is able to suppress apoptosis lead to its use in cell culture to prolong the productive period of the culture (128–131). Several recent studies show that cells expressing this gene increased MAb production by 40% (129) and prolonged the viable culture time by up to 4 days by delaying the onset of apoptosis for 2 days (131). Additionally, it was determined that the increase in MAb production was equally due to both the prolonged survival of cells and the increase in protein production per cell (131). The expression of the *bcl-2* gene also allows cells to be cultured with a significant decrease in serum supplementation (130). Specifically, *bcl-2* overexpressing cells could maintain high viable cell concentrations for up to 4 days, while the control culture lasted only 1 day when supplemented with 0.5% serum. Moreover, cells cultures supplemented with 9% serum with additional 2% serum fed after 4 days sustained productivity for 10 days, while the control culture underwent apoptosis after 4 days (130). These experiments show that the cell cycle controls can be bypassed when expression of the appropriate gene is induced in culture.

A similar approach has been used in CHO cells to bypass the cell cycle controls that require the presence of growth factors for efficient proliferation. The observation that cells stimulated by growth factors possess elevated cyclin E levels led researchers to transfect the CHO cells with a cyclin E expression vector (132). The expression of this gene in cells cultured in protein-free media resulted in cell proliferation and levels of cyclin E, cell morphologies, and cell cycle phase distributions comparable to those obtained in growth-factor-stimulated cells. An approach related to the one just described employed expression of the transcription factor E2F-1 in protein-free media to bypass the growth-factor requirements for proliferation (133). Although the desired effect of eliminating growth-factor dependency of the cells was achieved in both cases, the two approaches had different effects on several cellular properties. For instance, the cells expressing cyclin E rapidly round up and move into suspension, whereas the

E2F-1 cells adhered more tightly to the surface. Additionally, the same doubling time was observed for both cell types; however, the cyclin E cells resided in G_1 for approximately 20 min longer, while the S phase was reduced by 2 h as compared with the E2F-1 cells. Finally, the expression of certain proteins varied between the two cell types. The expression of cyclin A was shown to be 2 times higher in the E2F-expressing cells, and these same cells expressed a greater number of proteins at higher levels than wild-type cells than did the cyclin E–expressing cells. These results present two different ways of circumventing growth-factor-dependent controls and show different effects on cellular processes. Additionally, these findings may give some insight into the roles that these factors play in the regulation of the cell cycle.

Another recent study involved increasing protein production in cultured cells by arresting cells in G_1 (134). This was done by transfecting CHO cells with a gene coding for either p21, p27, or p53175P controlled by a tetracycline-regulated promoter. The p53 protein used in this study is a mutant that has lost its apoptotic function, resulting in a culture that undergoes G_1 arrest rather than programmed cell death. The culture was allowed to proliferate until it reached the desired cell density, after which expression of the cell cycle arrest molecules was stimulated, thus blocking proliferation. This sort of cytostatic process has several advantages over other processes involving uncontrolled growth of cultures, including slower rates of nutrient exhaustion, higher levels of protein production, and low genetic drift.

Gene Therapy

A firm grasp of the events that control the cell cycle will also yield significant advances toward the treatment of various illnesses including cancer and genetic disorders. The connection between the cell cycle and some cancers has been reviewed elsewhere (18,65,135,136). Many genetic disorders that alter the way cells function result from either the absence of an enzyme or the presence of a flawed enzyme. This situation may be remedied by the introduction of the appropriate genes into a target cell, a procedure known as gene therapy. Although cells that have incorporated the desired gene into their genome have displayed the desired effects, this technique is limited by low rates of gene transfer (reviewed in Ref. 137). Numerous studies have attempted to uncover the factors that determine transfer efficiency, resulting in the hypothesis that gene transfer is cell cycle dependent (reviewed in Refs. 138 and 139). More specifically, the most widely used method for gene transfer is recombinant retroviruses, and genes introduced into cells in this manner are not integrated into the host's genome until mitosis. It has been postulated that nuclear envelope breakdown is necessary for the viral integration complexes to act. Based on these studies, Andreatis and Palsson suggested that transfer efficiency depends both on cell cycle and retroviral life cycle events and developed a mathematical model describing the interplay between the two (140). This model aided in the design of experiments that allow the direct measure of intracellular retroviral decay (138). Furthermore, they developed a method to monitor the cell cycle dependence of gene transfer by using the transduction of β-galactosidase as a marker. Thus it was determined that almost all transduced cells show up in the early G_1 phase, supporting the hypothesis that only mitotic cells are susceptible to gene transfer (139). These studies reveal the importance of the kinetics of both retroviral decay and cell cycle progression of the target cells in the efficiency of gene therapy.

CONCLUSIONS

An overview of the mechanisms by which cells respond to different environmental stimuli has been presented. It is evident that even after the commitment to synthesize new DNA is made, the cell possesses the ability to monitor and direct internal events. For instance, the transcription and activation of key enzymes are regulated by cell cycle controls to ensure that necessary cellular processes are carried out successfully. The existence of this type of control suggests that almost every cellular process is somehow linked to the cell cycle machinery to guarantee the proper coordination of events throughout the cell cycle. A failure to maintain the order of these events may result in cell death or a disruption of normal cell behavior. Thus, a more complete understanding of factors that have an effect on cell cycle regulation would allow one to gain insight into better directing the cellular processes bound to cell cycle control. This fact is characterized by the recent increase in the number of studies that have focused on better understanding the components involved in directing cell division and their influence on fundamental cellular processes. It seems that obtaining a better understanding of the cell cycle and understanding how the cell cycle directs fundamental components of cell growth, such as protein production and secretion, may provide the missing link that will allow researchers to not only improve cell culture, but develop treatments for many types of ailments and dysfunctions such as cancer, which are indubitably cell cycle related.

BIBLIOGRAPHY

1. R.A. Laskey, M.P. Fairman, and J.J. Blow, *Science* **246**, 609–613 (1989).
2. B. Stillman, *Science* **274**, 1659–1664 (1996).
3. P.N. Rao and R.T. Johnson, *Nature* **159**, 159–164 (1970).
4. G.H. Leno, C.S. Downes, and R.A. Laskey, *Cell* **69**, 151–158 (1992).
5. J.P.J. Chong, H.M. Mahbubani, C.Y. Khoo, and J.J. Blow, *Nature* **375**, 360–361 (1995).
6. P. Romanowski, M.A. Madine, and R.A. Laskey, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10189–10194 (1996).
7. Y. Kubota, S. Mimura, S.I. Nishimoto, H. Takisawa, and H. Nojima, *Cell* **81**, 601–609 (1995).
8. V. Baldin, J. Lukas, M.J. Marcote, M. Pagano, and G. Draetta, *Genes Dev.* **7**, 812–821 (1993).
9. M. Ohtsubo and J.M. Roberts, *Science* **259**, 1908–1912 (1993).
10. D.H. Walker and J.L. Maller, *Nature* **354**, 314–317 (1991).

11. E. Schwob, T. Bohm, M.D. Mendenhall, and K. Nasmyth, *Cell* **79**, 181–184 (1994).
12. M. Pagano, S.W. Tam, A.M. Theodoras, P. Beer-Romero, G.D. Sal, V. Chau, P.R. Yew, G.F. Draetta, and M. Rolfe, *Science* **269**, 682–685 (1993).
13. M.B. Kastan, O. Onyekwere, D. Sidransky, B. Voglestein, and R.W. Craig, *Cancer Res.* **51**, 587–597 (1991).
14. M.J. O'Conell and P. Nurse, *Curr. Opin. Cell Biol.* **6**, 867–871 (1994).
15. I. Hoffmann and E. Karsenti, *J. Cell Sci.* **Suppl. 18**, 75–79 (1994).
16. M.K.K. Shivji, S.J. Grey, U.P. Strausfeld, R.D. Wood, and J.J. Blow, *Curr. Biol.* **4**, 1062–1068 (1994).
17. M.E. Ewan, B. Faha, E. Harlow, and D.M. Livingston, *Science* **255**, 85–90 (1992).
18. C.C. Orlowski and R.W. Furlanetto, *Endocrinol. Metab. Clin. N. Am.* **25**, 491–502 (1996).
19. M.A. Felix, J.C. Labbe, M. Doree, T. Hunt, and E. Karsenti, *Nature* **346**, 379–382 (1990).
20. T. Lorca, J.C. Labbe, A. Devault, D. Fesquet, U. Strausfeld, J. Nilsson, P.A. Nygren, M. Uhlen, J.C. Cavadore, and M. Doree, *J. Cell Sci.* **102**, 55–62 (1992).
21. R. Reeves, *Curr. Opin. Cell Biol.* **4**, 413–423 (1992).
22. A. Hershko, D. Ganoth, V. Sudakin, A. Dahan, L.H. Cohen, F.C. Luka, J.V. Ruderman, and E. Eytan, *J. Biol. Chem.* **269**, 4940–4946 (1994).
23. S. Yamashiro, Y. Yamakita, R. Ishikawa, and F. Matsumura, *Nature* **344**, 675–678 (1990).
24. W. Wells, *Trends Cell Biol.* **6**, 228–234 (1996).
25. J.A. Endicott, P. Nurse, and L. Johnson, *Protein Eng.* **7**, 243–253 (1994).
26. S.J. Elledge, *Science* **274**, 664–1672 (1996).
27. X. Li and R.B. Nicklas, *Nature* **373**, 630–632 (1995).
28. C.L. Reider, R.W. Cole, A. Khodjakov, and G. Sluder, *J. Cell Biol.* **130**, 941–948 (1995).
29. B. Novak and J.J. Tyson, *J. Cell Sci.* **106**, 1153–1168 (1993).
30. B. Novak and J.J. Tyson, *J. Theor. Biol.* **165**, 101–134 (1993).
31. K. Galaktianov and D. Beach, *Cell* **67**, 1181–1194 (1991).
32. C.J. Sherr and J.M. Roberts, *Genes Dev.* **9**, 1149–1163 (1995).
33. T. Evans, E.T. Rosenthal, J. Youngblom, D. Distel, and T. Hunt, *Cell* **33**, 389–396 (1983).
34. T. Hunt, *Semin. Cell Biol.* **2**, 213–222 (1991).
35. J. Pines, *Adv. Cancer Res.* **66**, 181–212 (1995).
36. H. Kobyashi, E. Stewart, R. Poon, J.P. Ademczewski, J. Gannon, and T. Hunt, *Mol. Biol. Cell* **3**, 1279–1294 (1992).
37. J. Pines and T. Hunter, *Trends Cell Biol.* **1**, 117–121 (1991).
38. D.S. Peeper, L.S. Ewan, M.E. Ewan, M. Toebes, F.L. Hall, M. Xu, A. Zantema, A.J. van-der-Eb, and H. Piwnica-Worms, *EMBO J.* **12**, 1947–1954 (1993).
39. S.I. Reed, *Annu. Rev. Cell Biol.* **8**, 529–561 (1992).
40. C.J. Sherr, *Cell* **73**, 1059–1065 (1993).
41. X. Grana and E.P. Reddy, *Oncogene* **11**, 211–219 (1995).
42. G.F. Draetta, *Curr. Opin. Cell Biol.* **6**, 842–846 (1994).
43. Y. Xiong, H. Zhang, and D. Beach, *Cell* **71**, 504–514 (1992).
44. C.J. Sherr, *Trends Biochem. Sci.* **20**, 187–190 (1995).
45. H. Matsushime, M. Roussel, R. Ashmun, and C.J. Sherr, *Cell* **65**, 701–713 (1991).
46. D.E. Quelle, R.A. Ashmun, S.A. Shurtleff, J. Kato, D. Bar-Sagi, M.F. Roussel, and C.J. Sherr, *Genes Dev.* **7**, 1559–1571 (1993).
47. D. Resnitzky, M. Gossen, H. Bujard, and S.I. Reed, *Mol. Cell. Biol.* **14**, 1669–1679 (1994).
48. A. Koff, A. Giordano, D. Desai, K. Yamashita, J.W. Harper, S.J. Elledge, T. Nishimoto, D.O. Morgan, B.R. Franza, and J.M. Roberts, *Science* **257**, 1689–1694 (1992).
49. V. Dulic, E. Lees, and S.I. Reed, *Science* **257**, 1958–1961 (1992).
50. L.H. Tsai, E. Harlow, and M. Meyerson, *Nature* **353**, 174–177 (1991).
51. L.H. Tsai, E. Lees, B. Faha, E. Harlow, and K. Riabowol, *Oncogene* **8**, 1593–1602 (1993).
52. G. Draetta, F. Luka, J. Westendorf, L. Brizuela, J. Ruderman, and D. Beach, *Cell* **56**, 829–838 (1989).
53. J. Pines and T. Hunter, *J. Cell Biol.* **115**, 1–17 (1991).
54. S.I. Reed, C. Wittenberg, D.J. Lew, V. Dulic, and M. Henze, *Cold Spring Harbor Symp. Quant. Biol.* **56**, 61–67 (1991).
55. M. Glotzer, A.W. Murray, and M.W. Kirschner, *Nature* **349**, 132–138 (1991).
56. J. Pines, *Biochem. J.* **308**, 697–711 (1995).
57. J. Pines, *Cancer Biol.* **5**, 305–313 (1994).
58. R.P. Fisher and D.O. Morgan, *Cell* **78**, 713–724 (1994).
59. T.P. Makela, J.P. Tassan, E.A. Nigg, S. Frutiger, G.J. Hughes, and R.A. Weinberg, *Nature* **371**, 254–257 (1994).
60. W.J. Feaver, J.Q. Svejstrup, N.L. Henry, and R.D. Kornberg, *Cell* **79**, 1103–1109 (1994).
61. R. Roy, J.P. Adamczewski, T. Seroz, W. Vermeulen, J.P. Tassan, L. Schaeffer, E.A. Nigg, J.H.J. Hoeijmakers, and J.H. Egly, *Cell* **79**, 1093–1101 (1994).
62. S. Atherton-Fessler, L.L. Parker, R.L. Geahlen, and H. Piwnica-Worms, *Mol. Cell. Biol.* **13**, 1675–1685 (1993).
63. C. Featherstone and P. Russell, *Nature* **349**, 808–811 (1991).
64. R.N. Booher, R.J. Deshaies, and M.W. Kirschner, *EMBO J.* **12**, 3417–3426 (1993).
65. T.K. MacLachlan, N. Sang, and A. Giordano, *Crit. Rev. Euk. Gene Express.* **5**, 127–156 (1995).
66. S.J. Elledge and J.W. Harper, *Curr. Opin. Cell Biol.* **6**, 847–852 (1994).
67. Y. Xiong, G.J. Hannon, H. Zhang, D. Casso, R. Kobayashi, and D. Beach, *Nature* **366**, 701–704 (1993).
68. J.W. Harper, G.W. Adami, N. Wei, K. Keyomarsi, and S.J. Elledge, *Cell* **75**, 805–816 (1993).
69. H. Zhang, G.J. Hannon, and D. Beach, *Genes Dev.* **8**, 1750–1758 (1994).
70. K. Polyak, J.-Y. Kato, M.J. Solomon, C.J. Sherr, J. Massague, J.M. Roberts, and A. Koff, *Genes Dev.* **8**, 9–22 (1994).
71. K. Polyak, M.-H. Lee, H. Erdjument-Bromage, P. Tempst, and J. Massague, *Cell* **78**, 59–66 (1994).
72. E.J. Firpo, A. Koff, M.J. Solomon, and J.M. Roberts, *Mol. Cell. Biol.* **14**, 4889–4901 (1994).
73. M.E. Ewan, H.K. Sluss, L.L. Whitehouse, and D.M. Livingston, *Cell* **74**, 1009–1020 (1993).
74. M. Serrano, G.J. Hannon, and D. Beach, *Nature* **366**, 704–707 (1993).
75. T. Nobori, K. Miura, D.J. Wu, A. Lois, K. Takabayashi, and D.A. Carson, *Nature* **368**, 753–756 (1994).
76. G.J. Hannon and D. Beach, *Nature* **371**, 257–261 (1994).
77. D.S. Peeper and R. Bernards, *FEBS Lett.* **410**, 11–16 (1997).
78. R.A. Weinberg, *Science* **254**, 1138–1146 (1991).

79. M. Laiho, J.A. DeCaprio, J.W. Ludlow, D.M. Livingston, and J. Massague, *Cell* **62**, 175–185 (1990).
80. E. Moran, *Curr. Opin. Gen. Dev.* **3**, 63–70 (1993).
81. S.P. Chellappan, S. Hiebert, M. Mudryj, J.M. Horowitz, and, J.R. Nevins, *Cell* **65**, 1053–1061 (1991).
82. E.W.-F. Lam and N.B.L. Thangue, *Curr. Opin. Cell Biol.* **6**, 859–866 (1994).
83. J. Lukas, H. Muller, J. Bartkova, D. Spitkovsky, A.A. Kjerulff, D.P. Jansen, M. Strauss, and J. Bartek, *J. Cell Biol.* **125**, 625–638 (1994).
84. M.E. Ewan, H.K. Sluss, C.J. Sherr, H. Matsushime, J. Kato, and D.M. Livingston, *Cell* **73**, 487–497 (1993).
85. S.F. Dowdy, P.W. Hinds, K. Louie, S.I. Reed, A. Arnold, and R.A. Weinberg, *Cell* **73**, 499–511 (1993).
86. H. Muller, J. Lukas, A. Schneider, P. Warthoe, J. Bartek, M. Eilers, and M. Strauss, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2945–2949 (1994).
87. E. Lees, B. Faha, V. Dulic, S.I. Reed, and E. Harlow, *Genes Dev.* **6**, 1874–1885 (1992).
88. N. La-Thangue, *Curr. Opin. Cell Biol.* **6**, 443–450 (1994).
89. I. Hoffman, G. Draetta, and E. Karsenti, *EMBO J.* **13**, 4302–4310 (1994).
90. A. Di-Leonardo, S.P. Linke, K. Clarkin, and G.M. Wahl, *Genes Dev.* **8**, 2540–2551 (1994).
91. R.W. King, R.J. Deshaies, J.M. Peters, and M.W. Kirschner, *Science* **274**, 1652–1659 (1996).
92. W. Krek, M.E. Ewan, S. Shirodkar, Z. Arany, W.G. Kaelin, and D.M. Livingston, *Cell* **78**, 161–172 (1994).
93. M. Pagano, R. Pepperkok, F. Verde, and G. Draetta, *EMBO J.* **11**, 961–971 (1992).
94. A. Bailley and M. Bornens, *Nature* **355**, 300–301 (1992).
95. T.R. Coleman and W.G. Dunphy, *Curr. Opin. Cell Biol.* **6**, 877–882 (1994).
96. Y. Sanchez, C. Wong, R.S. Thoma, R. Richman, Z. Wu, H. Piwnica-Worms, and S.J. Elledge, *Science* **277**, 1497–1501 (1997).
97. F. Verde, M. Dogterom, E. Stelzer, E. Karsenti, and S. Leibler, *J. Cell Biol.* **118**, 1097–1108 (1992).
98. E.A. Nigg, *Curr. Opin. Cell Biol.* **5**, 187–193 (1993).
99. E.A. Nigg, *Curr. Opin. Cell Biol.* **4**, 105–109 (1992).
100. Y. Yamakita, S. Yamashiro, and F. Matsumura, *J. Cell Biol.* **124**, 129–137 (1994).
101. F.C. Luka, E.K. Shibuya, C.E. Dohrmann, and J.V. Ruderman, *EMBO J.* **10**, 4311–4320 (1990).
102. J.A. Knoblich, K. Sauer, L. Jones, H. Richardson, R. Saint, and C.F. Lehner, *Cell* **77**, 107–120 (1994).
103. B. Novak and J.J. Tyson, *J. Theor. Biol.* **173**, 283–305 (1995).
104. B. Novak and J.J. Tyson, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9147–9152 (1997).
105. J.J. Tyson, B. Novak, G.M. Odell, K. Chen, and C.D. Thron, *Trends Biochem. Sci.* **21**, 89–96 (1996).
106. J.F. Collins and M.H. Richmond, *J. Gen. Microbiol.* **28**, 15–33 (1962).
107. R.J. Harvey, A.G. Marr, and P.R. Painter, *J. Bacteriol.* **93**, 605–617 (1967).
108. A.G. Fredrickson, D. Ramkrishna, and H.M. Tsuchiya, *Math. Biosci.* **1**, 327–374 (1967).
109. D. Ramkrishna, A.G. Fredrickson, and H.M. Tsuchiya, *Bull. Math. Biophys.* **30**, 319–323 (1968).
110. S. Terada, E. Suzuki, H. Ueda, and F. Makishima, *Cytokine* **8**, 889–894 (1996).
111. K. Takahashi, S. Tereda, H. Ueda, F. Makishima, and E. Suzuki, *Cytotechnology* **15**, 57–64 (1994).
112. F. Makishima, S. Terada, T. Mikami, and E. Suzuki, *Cytotechnology* **10**, 15–23 (1992).
113. M. Al-Rubeai, A.N. Emery, S. Chalder, and D.C. Jan, *Cytotechnology* **9**, 85–97 (1992).
114. M. Al-Rubeai and A.N. Emery, *J. Biotech.* **16**, 67–86 (1990).
115. N.J. Cowan and C. Milstein, *Biochem. J.* **128**, 445–454 (1972).
116. S.J. Kromenaker and F. Srienc, *Biotech. Bioeng.* **38**, 665–677 (1991).
117. S.J. Kromenaker and F. Srienc, *J. Biotech.* **34**, 13–34 (1994).
118. M. Cherlet, S.J. Kromenaker, and F. Srienc, *Biotech. Bioeng.* **47**, 535–540 (1995).
119. M. Kubbies and H. Stockinger, *Exp. Cell Res.* **188**, 267–271 (1990).
120. M.B. Gu, P. Todd, and D.S. Kompala, *Biotech. Bioeng.* **42**, 1113–1123 (1993).
121. M.B. Gu, P. Todd, and D.S. Kompala, *Biotech. Bioeng.* **50**, 229–237 (1996).
122. M. Al-Rubeai, M.S. Chalder, R. Bird, and A.N. Emery, *Cytotechnology* **7**, 179–186 (1991).
123. D.E. Martens, C.D. d. Gooijer, C.A.M. v. d. V.-d. Groot, E.C. Beuvery, and J. Tramper, *Biotech. Bioeng.* **41**, 429–439 (1993).
124. K.A. Hawboldt, T.I. Linardos, N. Kalogerakis, and L.A. Behie, *J. Biotech.* **34**, 133–147 (1994).
125. S. Lakhotia, K.D. Bauer, and E.T. Papoutsakis, *Biotech. Bioeng.* **40**, 978–990 (1992).
126. M. Al-Rubeai, R.P. Singh, A.N. Emery, and Z. Zhang, *Biotech. Bioeng.* **46**, 88–92 (1995).
127. R.P. Singh, M. Al-Rubeai, C.D. Gregory, and A.N. Emery, *Biotech. Bioeng.* **44**, 720–726 (1994).
128. E. Suzuki, S. Terada, H. Ueda, T. Fujita, T. Komatsu, S. Takayama, and J.C. Reed, *Cytotechnology* **23**, 55–59 (1997).
129. N.H. Simpson, A.E. Milner, and M. Al-Rubeai, *Biotech. Bioeng.* **54**, 1–16 (1997).
130. S. Terada, Y. Itoh, H. Ueda, and E. Suzuki, *Cytotechnology* **24**, 135–141 (1997).
131. Y. Itoh, H. Ueda, and E. Suzuki, *Biotech. Bioeng.* **48**, 118–122 (1995).
132. W.A. Renner, K.H. Le, V. Hatzimanikatis, J.E. Bailey, and H.M. Eppenberger, *Biotech. Bioeng.* **47**, 476–482 (1995).
133. K.H. Lee, A. Sburlati, W.A. Renner, and J.E. Bailey, *Biotech. Bioeng.* **50**, 273–279 (1996).
134. M. Fussenegger, X. Mazur, and J.E. Bailey, *Biotech. Bioeng.* **55**, 927–939 (1997).
135. C. Cordon-Cardo, *Am. J. Pathol.* **147**, 545–559 (1995).
136. L.H. Hartwell and M.B. Kastan, *Science* **266**, 1821–1828 (1994).
137. B. Palsson and S. Andreadis, *Exp. Hematol.* **25**, 94–102 (1997).
138. S.T. Andreadis, D. Brott, A.O. Fuller, and B.O. Palsson, *J. Virol.* **71**, 7541–7548 (1997).
139. S. Andreadis, A.O. Fuller, and B.O. Palsson, *Biotech. Bioeng.* **58**, 272–281 (1998).
140. S. Andreadis, and B.O. Palsson, *J. Theor. Biol.* **182**, 1–20 (1996).

See also BIOENERGETICS OF MICROBIAL GROWTH; CULTURE COLLECTIONS; ENERGY METABOLISM, MICROBIAL AND ANIMAL CELLS.

CELL DEATH.

See APOPTOSIS; CELL CYCLE.

CELL DISRUPTION AND LYSIS

F. A. P. GARCIA
University of Coimbra
Coimbra, Portugal

KEY WORDS

Autolysis
Ball mill
Cell wall
Debris separation
Disintegration
Disruption
Homogenization
Lysis
Ultrasonics

OUTLINE

INTRODUCTION

Currently, biological products of commercial interest are excreted by producer cells into the surrounding medium, but the larger proportion of cellular products, many of which have a real or potential value, remain inside the cell (1). These intracellular products must be released into the medium before they can be recovered and purified. There are several strategies for accomplishing this: the cells can be genetically engineered to excrete the products, the cell wall can be made more permeable, cell lysis may be induced, or the cell wall can be disrupted. In preparative-scale applications the disruption of the cell wall is the most common method.

The cell wall is a rigid and complex structure that protects the cell from the environment, maintaining cell shape and providing osmotic stability, while performing biochemical functions of regulation and selective control of the exchange of nutrients and metabolites with the medium. However, microbial cell wall characteristics vary widely in composition and structure according to species; it is important to recognize the effects of these differences on the strength of the wall and their implications which disruption techniques should be adopted. Mammalian tissues, halophilic bacteria (which do not possess a true rigid cell wall), and the Mycoplasma bacteria (which lack a wall) have the lowest mechanical strength and can be easily lysed under osmotic stress in hypotonic media. The more structured and stronger cell walls in some bacteria and fungi are more difficult to disrupt, and more energy is required to overcome the toughness of the wall. Moreover, the cell age and physiological state are also important differences that contribute to the strength of a given species' cell wall.

THE CELL WALL

Chemical Composition and Structure

The diversity of cell wall composition and structure is well documented (2–5). Purified cell walls (6) can be hydrolyzed, and the different components can be identified.

Bacterial Cell Walls. Bacteria differ with respect to their ability to retain a crystal violet–iodine stain when treated with organic solvents (alcohol or acetone). Accordingly, they can be classified into two main groups: Gram positive if they retain the stain, and Gram negative if the stain is not retained. This difference reflects the diversity in cell wall composition and structure as well as differences in biochemical, physiological, or genetic features. Common to all cell walls is the element of rigidity, which is provided by a supramolecular structure known as a peptidoglycan, or murein (Fig. 1). This is a regular heteropolymer of alternating *N*-acetylglucosamine and *N*-acetylmuramic acid (the lactyl ether of *N*-acetylglucosamine) linked by β-(1–4) glycosidic bonds in linear chains of up to 65 disaccharide residues. A tetrapeptide side chain, consisting of at least three amino acids (some of the unusual D-amino acids), is attached to the carboxyl group of the *N*-acetylmuramic acid. The tetrapeptide chain of one polysaccharide chain is linked to a tetrapeptide chain of an adjacent polysaccharide chain, either directly (the terminal D-alanine of one to the *meso*-diaminopimelic acids in position 3 of the other, as in *Escherichia coli*), or indirectly through a pentaglycine peptide (the terminal D-alanine of one to the L-lysine in position 3 of the other, as in *Staphylococcus aureus*). Not all pairs of these tetrapeptide chains are involved in the cross-linking of the glycan chains; the degree of cross-linking, as well as the nature of the amino acids in positions 2 and 3 are dependent on the species.

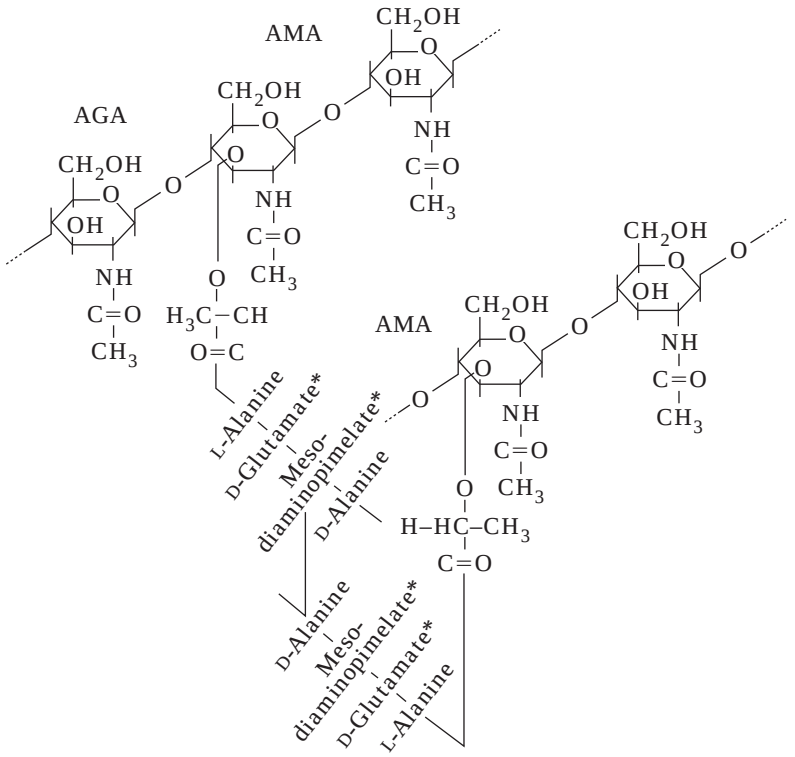


Figure 1. Structure of the building blocks of peptidoglycan in *E. coli*. The nature of the amino acids at positions 2 and 3, as well as the degree of cross-linking differ among species. AGA, *N*-acetylglucosamine; AMA, *N*-acetylmuramic acid.

Murein is the skeletal component of the cell wall in Gram-positive bacteria; it has a multilayered architecture that is 20–80 nm thick. In addition, teichoic acids (polymers of glycerol or ribitol phosphate) are associated with murein on the external side of the wall and are unique to Gram-positive cells. Teichoic acids do not contribute to wall strength but provide antigenic activity or serve as an attachment anchor. Gram-positive cell walls are devoid of significant amounts of lipids and proteins.

In Gram-negative bacteria the cell wall is thinner and more complex and stratified. Mechanical resistance is provided by an inner layer of murein, which is much thinner than in Gram-positive bacteria. This murein is covered externally by a bilayered lipid membrane, the internal leaflet being lipoprotein, containing phospholipids, and the outer leaflet being lipopolysaccharide (LPS), which has toxic properties and is able to resist the chemical aggressions of the environment.

Fungi and Yeasts. Under the electronic microscope, the walls of fungi and yeasts are seen to have a fibrillar structure that results from the presence of certain components such as cellulose or chitin. The main constituents are homo- or heteropolysaccharides, accounting for more than 80% of the wall, built up from different sugars (mainly D-glucose, D-mannose and *N*-acetylglucosamine, but also galactose, galactosamine, fucose, xylose, etc.). Usually, pairs of carbohydrates dominate the composition of the fungal walls, and a vast majority of fungi have a chitin–glucan combination, whereas the majority of yeasts have a glucan–mannan combination. Cellulose is also dominant in the walls of many lower fungi. Glucans are usually considered to be the main skeletal component in yeasts and consist of linear chains of β-(1–6)-linked D-glucose and side chains of D-glucose linked to the first ones through β-(1–3) bonds (Fig. 2a). Chitin is a polymer of *N*-acetylglucosamine and is always found in yeast cells, although in small amounts (Fig. 2c).

Lipids and proteins are also present in fungi and yeasts cell walls in different amounts. Proteins are now regarded as true structural wall components. In yeasts the wall appears stratified into two layers: the protein is associated with mannans in the form of mannoproteins that form an external layer covering an internal layer of glucans (Fig. 2b).

Algae. Cellulose is also the main skeletal component of the algal cell wall, but xylans and mannans (depending on which sugars are the most abundant) are also found. Algae differ considerably from the fungi with respect to cellulose content and crystallinity. The physical aspect is also fibrillar, as with other eukaryotes.

Effects of Upstream Conditions

The conditions that affect the synthesis of the cell wall constituents during the growth process will be reflected in the strength of the cell wall. For instance, cells of *E. coli* grown in a complex medium are less easily disrupted than cells grown in a simple medium (7). The ease of disintegration also depends on the batch growth of *E. coli*; the cells are weaker during the logarithmic phase and become stronger in the transition to the stationary phase (1,11). This corresponds to an increase in the murein thickness and cross-

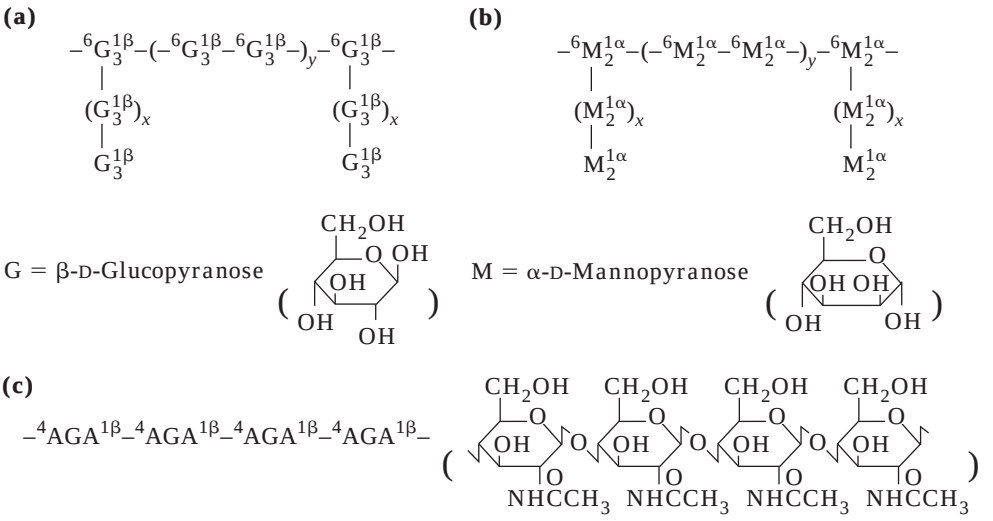


Figure 2. Structure of the building blocks of (**a**) glucan, (**b**) mannan, and (**c**) chitin in yeasts.

linking (8,10,11). A quantitative relationship has been established between mean cell strength and cell length and degree of cross-linkage (9). A higher specific rate of growth, as attained in continuous culture of *Candida utilis*, leads to a weaker cell wall, whereas a lower growth rate, as in a batch operation, allows the cells to build a stronger cell wall (12). Wall thickness and resistance to disruption in bakers' yeast increase with age, possibly due to exhaustion of the carbon source (13). Bud scars, much more frequent in a younger yeast population, may introduce local areas of lower wall strength (12,13). The compressive force required for the disruption of brewers' yeast cells was measured by using a micromanipulation technique and ranged from 40 μN for cells in the exponential phase to 90 μN for cells in the stationary phase (14).

ASSESSMENT OF THE DISRUPTION YIELD

Suitable analytical techniques are required to assess the extent of cell disruption in order to develop models for process selection, scale-up, and optimization. Different methods have been used to quantify the degree of disintegration at specified conditions; each one has its advantages and drawbacks. The number of disrupted cells must be estimated, which can be done by direct counting of nondisrupted cells or, indirectly, by determination of any intracellular metabolite whose quantitative release into the medium is proportional to the number of disrupted cells.

Direct counting of suitably diluted samples can be achieved by plating out techniques (14), microscopic counting (16,17), electronic counting (18), or using an analytical disk centrifuge (19).

Plating out techniques are time consuming, and only viable cells can be counted. However, nonviable intact cells can be present in a significant proportion and are uncountable. The situation is worsened if the cells have a tendency to agglomerate.

Microscopic counting is a faster and simpler alternative. Whole cells in suitably diluted samples of original and treated suspensions are counted in microscope hemocytometers (16,17). However, the counting process, though easier for yeasts, is tedious and difficult with very small bacterial cells, and a distinction between intact cells and cells that have been slightly damaged is generally impossible to make. Proportional counts of stained smears may be advantageous as, for instance, in the case of yeasts (20). By adopting the staining technique of Bianchi (21), it is possible to distinguish among whole intact cells, broken cells contaminated with cytoplasm, and empty debris, at 1,000× magnification. The main difficulty lies in finding a suitable staining technique for the cells concerned.

Electronic particle counters can be used to evaluate the number of whole yeast cells. The Coulter counter, for instance, was used to assess the concentration of yeast cells (18), but it is not a sensitive method for bacteria (22), and the machine tends to be fouled by cellular debris (19).

The analytical disk centrifuge was recently introduced for monitoring the disruption of *E. coli* containing recombinant inclusion bodies (19). This technique provides a size distribution in the homogenate and requires that the size distribution of the inclusion bodies be subtracted from the homogenate sample distribution by numerical or graphical deconvolution of both distributions.

Indirect counting is based on the determination of certain compounds (e.g., soluble protein, enzyme activities) that are released into the solution when the cell wall is broken: the yield of disintegration is the percentage ratio of the released material, R, to the maximum amount, R_{max}, that can theoretically be released into the solution. The most common material that has been used for this purpose is total soluble protein assayed in the sample supernatant (22–24). The "dilution method," using the Biuret method, was based on the comparison of the protein concentrations in the supernatants obtained by centrifugation of diluted and undiluted samples of homogenized suspensions (23). However, dilution interferes with protein solubility (22), because of partial denaturation. The method also fails when the protein is produced as an inclusion body (19).

Overheating, particularly during a prolonged operation to determine R_{max}, can cause thermal denaturation or enzyme deactivation.

Other bioactive products and, in particular, enzyme activities (25) released into the medium can be useful. However, the assessed value of the disruption yield under fixed conditions depends largely on the enzyme selected for analysis, its location within the cell, whether it is bound to membranous materials, and its heat or shearing lability (26), which can affect the values of both R and R_{max}. The most adequate material would be the product to be extracted and isolated, if it can be easily assayed.

Continuous monitoring of the disruption process and the performance of the equipment, particularly if wearing problems are severe, is advantageous. For instance, a relationship between the degree of cell disruption and the temperature increase of a baker's yeast suspension when passing through a mechanical homogenizer can be established. Therefore, by inserting a thermocouple in the homogenizer, continuous monitoring of its performance can be achieved (27).

METHODS OF CELL WALL DISRUPTION

Techniques for cell wall disruption have evolved over the years and are well documented in the literature (27–31). There are both mechanical and nonmechanical methods. Mechanical methods are based on shear forces that deform and rupture the cell wall. Nonmechanical methods induce the lysis of the cell by chemical, physical, or enzymatic means.

Many of these methods were developed for laboratory-scale use, with the objective of purifying biomolecules or organelles while maintaining as much integrity as possible for fundamental studies. For preparative or commercial applications, the strategy is to maximize product release, keeping the biological activity and minimizing the formation of fine cell debris, which can interfere with further downstream operations. The mechanical methods, based on solid or liquid shear, are best suited for this purpose.

High-Pressure Mechanical Homogenization

High-pressure mechanical homogenizers, typified by the APV Manton–Gaulin homogenizer (Fig. 3), have been one of the most widely used devices for the disruption of microbial cell walls in large-scale recovery of intracellular bioproducts. These devices consist of a positive displacement pump capable of generating high pressures to force the cell suspension through an adjustable discharge needle valve under high shear stresses, which disrupts the cells. The body of the valve is positioned against a valve seat by a spring-loaded rod, and the disrupting pressure is adjusted by the load that is applied, hydraulically or mechanically, to the spring. The treated suspension can be either collected or recycled for more passes through the valve.

In an earlier report (23) such a device was used to disrupt baker's yeast, and the degree of disruption was proportional to the number of passes, N, through the valve according to the model (first-order kinetics relative to N),

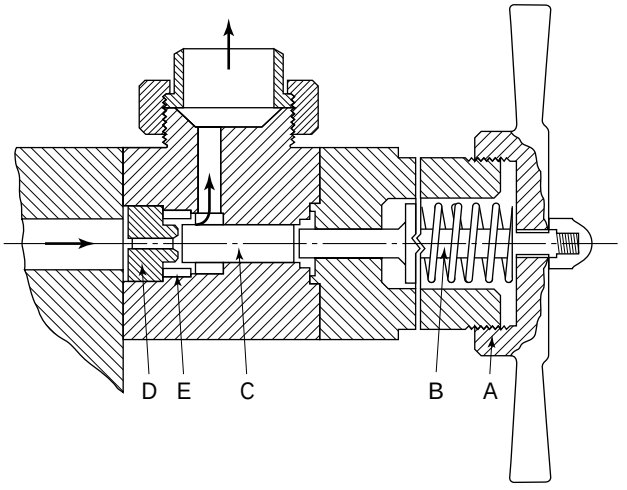


Figure 3. High-pressure homogenizer discharge valve unit. The discharge pressure is controlled by a handwheel assembly (A), which, through a spring-loaded valve rod (B) positions the valve (C) in relation to the valve seat (D). During discharge the material passes between the valve and its seat and impinges on an impact ring (F). *Source*: Reproduced by permission of the Institution of Chemical Engineers from Ref. 23.

$$\ln \frac{1}{1 - X} = kNP^a \tag{1}$$

where $X = R/R_{max}$ is the fraction of disrupted cells, k is a dimensional, machine-dependent (it depends on the geometry of the needle valve) specific rate constant $[(k) = (P)^{-a} = \text{MPa}^{-a}]$; the constant a is a measure of the cell resistance to disruption (the value $a = 2.9$ was reported for baker's yeast, and 2.2 for *E. coli*) and is dependent on the type of organism and the physiologic conditions and rate of growth (12).

Besides the pressure, the type of organism and the cell growth conditions, other variables such as the temperature and the cell concentration affect the specific rate of disintegration. The process has been found to be independent of cell concentration in the range 84–168 g dry matter/L and for pressures extending from 19.6 to 53.9 MPa, although lower protein recoveries for higher cell concentrations at higher pressures were obtained (23). Further studies extended the analysis of the influence of the cell concentration up to 240 g dry matter/L, and no significant effect was detected (32). However, a theoretical analysis based on the turbulence theory of Kolmogoroff (33) showed that the amount of soluble protein release in a Manton–Gaulin homogenizer was dependent on cell concentration according to

$$R = 1 - \exp\{-[(P - P_0)/z]^{\beta}\} \tag{2}$$

where z is a constant, P is the disrupting pressure, P_0 is a threshold pressure, and β is a parameter that depends on cell concentration. More recently, the disruption of *E. coli* in another type of homogenizer, the Microfluidizer (Microfluidics Corporation, Newton, Mass.), was correlated by (34)

$$\ln \frac{1}{1 - X} = kN^b P^a \quad (3)$$

where the exponent b was found to be a function of the cell concentration and the specific growth rate.

The rate of cell wall disruption has been shown to be highly temperature dependent, but a temperature increase when the cells are disrupted can lead to severe thermal denaturation of heat-labile products. An increase of about 2–3 °C (depending on cell concentration) can be expected for every 10 MPa increase in the backpressure of the needle valve (31,32). In fact, the suspension has to be precooled in order to keep the temperature under control, and cooled immediately after leaving the homogenizer. In most cases no losses of enzyme activity or protein solubility were detected if the temperature was not allowed to rise above 30 °C (12,23).

Increased productivity can be attained with higher pressures, allowing the recovery of more product with fewer passes. However, a higher disrupting pressure demands more energy and can cause severe wearing problems in the valve seat. Erosion of the valve seat is, in fact, of great concern in using this type of equipment.

Optimization of the equipment design and operation requires a deep understanding of the mechanisms of cell disruption. Breakage of the cell wall occurs as a result of normal stresses or shear forces developed during the passage of the suspension through the needle valve. One mechanism considers the magnitude and the rate of pressure drop to be responsible for cell breakage (32). A hydrodynamic mechanism in which turbulent eddies of smaller dimensions than the cells cause the intracellular liquid to oscillate with a kinetic energy large enough to rupture the cell (33). Impingement of a high-velocity jet against a flat surface, as occurs in the needle valve when the direction of the jet is suddenly and abruptly changed, has also been considered to be a cause of cell disruption (12,35). In the case of the Manton–Gaulin–type of homogenizers, this led to the redesign of the valve geometry (35); the so-called knife-edge configuration showed the best performance, whereas the flat valve configuration had the lowest yields in the disruption of baker's yeast (Fig. 4).

These homogenizers have proved to be efficient for all types of microorganisms, with high degrees of disruption even with only one pass. Blockage has, nevertheless, remained a major concern in the use of high-pressure homogenizers for the disruption of filamentous microorganisms (36). The use of a Manton–Gaulin homogenizer resulted in a severe degradation of plasmid recovered from *E. coli* (37).

Solid-Shear Methods

Ball mills, or bead mills, originally designed for nonbiotechnological applications, have also been used to disintegrate microorganisms, particularly yeasts. They are the most suitable for scaling up and can be continuously or discontinuously operated. They consist of agitator disks mounted concentrically or excentrically on a motor-driven central shaft inside a grinding chamber (Fig. 5). During operation, the chamber is filled with small glass or stainless steel beads, which are necessary to promote an abrasive action to grind the cells. In continuous operation, a bead retainer (a sieve plate, a vibrating slot, or a special rotating disk) prevents the beads from leaving the grinding chamber. Different models of the Netzsch bead mill (Netzsch Feinmahltechnik, Germany) or the Dyno-Mill (Willy A. Bachofen AG, Basel, Switzerland) have been used.

The mechanism of disintegration is very complex and results from both a liquid shearing and a solid shearing, resulting from the impacts of the beads in a field of intense turbulence. The disruption process can be properly described by a first-order kinetics law (24,38) and, for a batch operation or continuous plug flow system, the integrated rate equation is

$$\ln \frac{R_{\max}}{R_{\max} - R} = kt \quad (4)$$

where R is the amount of protein released in time t, $R_{\max}$ is the maximum amount that could be released if 100% disrupted had occurred, and k is the specific velocity of disintegration (protein released).

For a continuous operation there is a certain degree of backmixing, and n tanks (CSTR) in a series model is a better representation of the hydrodynamics in the chamber. In this case a mass balance to the released protein gives (38)

$$\frac{R_{\max}}{R_{\max} - R} = [1 + (k\theta/n)^n] \quad (5)$$

where $\theta = V/Q$ is the mean residence time, V is the total free volume of the grinding chamber, and Q is the flow rate of the microbial suspension.

Several process parameters affect the yield and the rate of cell wall disruption. Besides the design characteristics of the mill (volume and geometry of the grinding chamber, the number of disks per unit volume of the mill, and geometry and construction materials of the disks), the rate constant is influenced by a number of process parameters such as the agitator speed, cell concentration, bead size, bead load, temperature, and flow rate of the suspension (24,38,39).

The speed of the agitator disks determines the frequency of bead collisions (40) and the intensity of shear, and, within certain limits, the specific rate of disintegration is proportional to the disk tip speed, U (38):

$$k = KU \quad (6)$$

Practical limitations arising from high energy consumption, high heat generation, erosion of the beads, and inactivation of shear-labile products require that the disk tip speeds remain between 5 and 15 ms^{-1}.

The influence of cell concentration requires a more detailed investigation because contradictory results have been reported. The disruption of baker's yeast was reported to be independent of the cell concentration in the range of 4–20% w/v on a dry matter basis (25), as expected for a first-order process. However, others (38,39) have re-

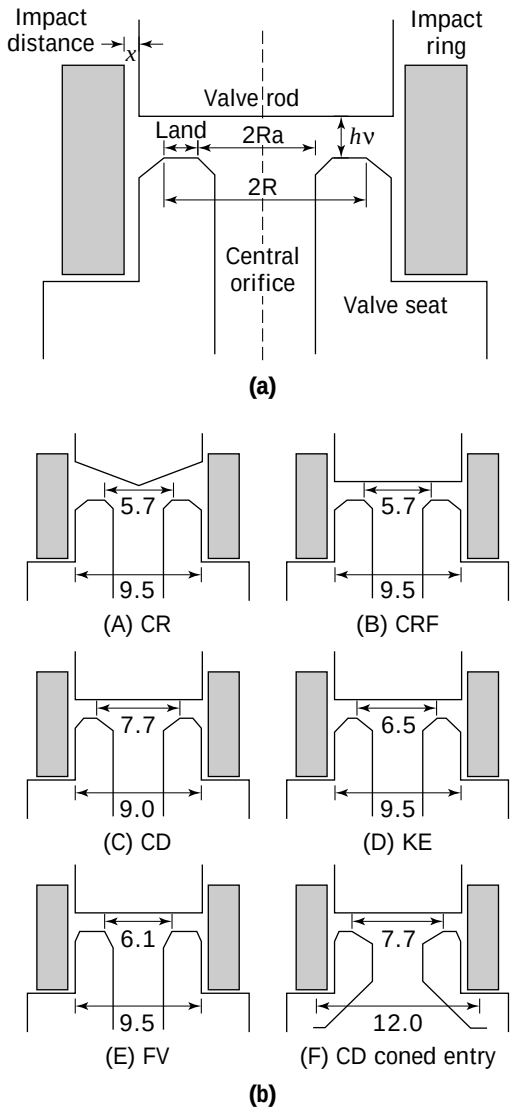


Figure 4. (**a**) Schematic diagram of a typical valve unit. (**b**) Schematic diagram of valve units used: (A) CR, cell rupture with coned valve rod; (B) CRF, cell rupture with flat valve rod; (C) CD, cell disruption; (D) KE, knife edge; (E) FV, flat valve unit; (F) CD-type unit with coned entry—dimensions in millimeters. Land distances ($R - R_o$) are for CRF, 0.7 mm; CD, 0.5 mm; KE, 0.1 mm; FV, 1.5 mm. *Source*: Reprinted by permission of Elsevier Science Inc. from Ref. 35.

ported a maximum protein release for cell concentrations of 35–40% (wet wt), and decreasing thereafter, depending on the disk materials (38), such effect being more pronounced when the agitator speed was low (39). The viscosity of the suspension increases with the cell concentration and is further increased with the progress of the disruption process, therefore increasing the rate at which the turbulence is damped out (38). However, insufficient data are generally generated or reported regarding the rheology of the suspension. The optimal cell concentration varies from case to case, and one must also consider that less heat is generated but the power consumption per unit weight of treated cells is higher for lower cell concentrations, and

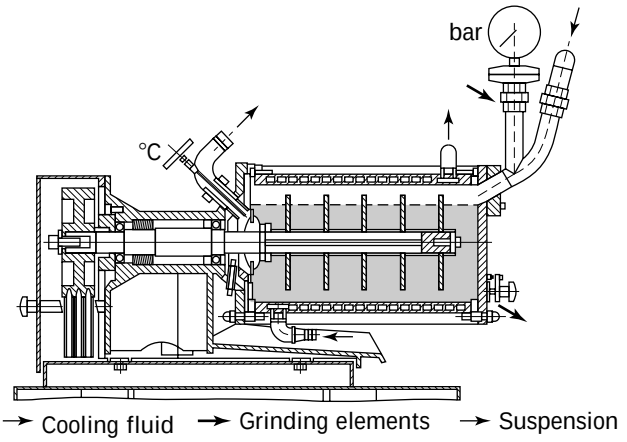


Figure 5. The Dyno-Mill. *Source*: Reprinted from a commercial brochure, by permission of Willy A. Bachofen AG, Basel/Switzerland.

that for low cell concentrations (<5% solids) increased wear is observed on the mill and beads. Hence, an optimum must be found by experiment and is generally in the range of 30–60% or, more usually, in the range of 40–50% (cell wet wt).

The bead diameters that have been reported range from 0.1 to 1.5 mm. Usually the disruption is faster with smaller beads, but the tendency of very small beads to float and difficulties in retaining them inside the grinding chamber impose practical limitations on the lower limits of size. In practice, the lower limit for batch operation of laboratory mills is 0.2 mm, whereas for production equipment working continuously the lower sizes are ≥0.4 mm. For the same bead load, larger particles correspond to fewer beads and reduced abrasive action or bead collisions. For the same bead load, there are also reports of an optimum bead size, which also has to be found by experiment and depends on the location (24,41) of the product (enzyme) inside the cells; larger beads are preferred for the recovery of enzymes located in the periplasmic region, and smaller ones for cytoplasmic material.

The rate of protein release is also enhanced by increasing the bead load (volume percentage of bead material relative to the total volume of the grinding chamber). Bead loads ranging from 30 to 90% are reported in the literature, but the packing density, depending on the bead diameter inside the chamber, is generally in the range of 80–90%. The upper limit is generally imposed by the high power consumption and difficulties in removing the heat released in the operation.

The rate constant is, obviously, dependent on the temperature that results from the balance between the heat generated during the milling process and the heat removed by a cooling liquid, brine, or a cooling liquid circulating through a cooling jacket of the grinding chamber. The heat generated in the process, if accumulated, would raise the temperature to unacceptable levels; in fact, the suspension has to be precooled and cooled immediately after leaving the chamber to avoid inactivation of heat-labile products.

The effects of the flow rate should mirror the effects of the residence time of the cells inside the grinding chamber,

and assuming a first-order process, the yield after one passage should decline as the feed rate increases. However, degree of caution should be exercised because the degree of longitudinal dispersion or backmixing increases as the flow rate decreases, and the yield of disruption is not a simple function of the flow rate. On the other hand, the power consumption per unit mass of processed cells strongly decreases as the feed rate increases, and for this reason, a high flow rate should be adopted and the decline in the yield of disruption could be alleviated by recycling part of the suspension. It should be emphasized that by increasing the residence time the viscosity may be decreased because prolonged shear may cause degradation of the shear-sensitive DNA polymer (51).

In one study (41), the disintegration rate was correlated with the specific energy as determined by the following combination of parameters: small grinding beads, moderate to high cell concentrations, and/or low to moderate agitator speeds. In this case the utilization of energy input reaches its maximum. In other words, high cell concentrations are energetically favorable. This result is considered to be the basis for scale-up. For cases of low cell concentrations in the suspension, high agitator speeds, and larger beads, the disintegration rate was no longer correlated with specific energy input but was a function of the stress frequency (which is proportional to the product of the agitator speed and disintegration time), which was experimentally proved.

Ultrasonics

Ultrasound has been used on the laboratory scale for analytical and preparative extraction and purification of cell constituents of bacteria or yeasts. The mechanism of cell disruption derives from the intensive shear induced by sonicating the suspension at sound frequencies above 20 kHz. A magnetorestrictive or a piezoelectric transducer converts the alternating current of an electric oscillator into mechanical waves that are transmitted to the suspension through a metallic probe (usually made from titanium) vibrating with the same frequency as the oscillator. The sound waves create many microbubbles at various nucleation sites in the suspension, which collapse implosively during the rarefaction period of the sound waves. This phenomenon of cavitation (formation, growth, and collapse of the vapor-filled bubbles) produces intense local shock waves, and intense local shear gradients are generated that cause the cells to deform beyond the limit of elasticity and rupture. The rate of microbial cell disintegration is usually low for preparative purposes, and the inclusion of small stainless steel or glass (Ballotini) beads will help in the nucleation for cavitation as well as with a grinding action (solid shear) that will increase the efficiency of disintegration.

The kinetics of disruption (protein release) has been reported to follow a first-order law (42,43), and for a well mixed pattern inside the disruption chamber the integrated equation is

$$1 - X = \exp(-kt) \tag{7}$$

where X is the fraction of released protein and k is the rate constant.

There are a number of variables that can markedly affect either the sample or the efficiency of the sonication process. One of these is the temperature rise. Most of the energy that is released when the bubbles collapse is absorbed by the liquid as heat. Although the value of the specific rate k increases with the temperature, the thermal energy has to be extracted to avoid temperatures that could destroy the activity of the useful bioproducts. To control the temperature within acceptable limits, the suspension should be precooled to 0–5 °C, and jacketed disruption vessels should be used with a continuous circulation of a cooling fluid. Alternatively or concomitantly, short periods of sonication can alternate with short periods of cooling. The ratio of the time of sonication to the time of cooling is known as the duty factor.

The integrity of the bioproduct can also be adversely affected if free radicals are allowed to accumulate during the intense cavitation. The addition of radical scavengers (e.g., cysteine or glutathione) or pregassing the suspension with hydrogen can alleviate this problem, particularly if a prolonged sonication is expected.

Other operating parameters affect the efficiency of the sonication process. One is the amplitude of vibrations of the sound waves (directly related to the acoustic power supplied to the suspension), which affects the rate constant according to the following relation (42):

$$k\alpha(W - W_0)^{0.9} \tag{8}$$

where W is the acoustic power and W_0 is the cavitation threshold power. However, a linear relationship between the rate constant and the acoustic power was recently reported (43).

The viscosity of the suspension affects the hydrodynamics and the rate of energy dissipation and inhibits the phenomenon of cavitation. When the viscosity is high, the operation should be started at a lower power, and the power should be gradually increased. The initial cell concentration that affects the viscosity of the suspension was reported to have no effects on the rate constant (43).

Problems of foaming can result from the presence of surface active agents (e.g., proteins released from the disrupted cells) that lower the surface tension of the liquid medium. An intense foaming, particularly when high power is applied, contributes to cavitation “unloading” (therefore decreasing the efficiency of the process) and protein denaturation at the gas–liquid interface.

High working volumes require high acoustic power, which results in violent turbulence and in the formation of large bubbles that cause cavitation unloading. Rate constants were found to decrease linearly with increasing working volumes (43). The difficulties associated with higher working volumes (100–200 mL) can be alleviated by good design of the disruption vessel. Commonly a jacketed glass beaker with a continuous flow of cooling fluid is used; inlet and outlet tubes can be inserted for continuous operation. In continuous operation, the flow rate determines the residence time of the cells in the vessel and affects the overall yield of the process. For tougher cells the

inclusion of Ballotini beads is necessary, and for further increasing the grinding action, the vessel should have internal tubing loops to circulate the suspension under the pumping energy of the sonication (27,31). The load and size of beads must in any case be optimized, but the value of the rate constant increases to a maximum and then decreases with increases in the bead diameter for the same bead load.

The shape of the probe is of particular concern. For a constant power level the amplitude of the sound waves at the probe tip is inversely proportional to the tip area. Therefore large tips dissipate the acoustic energy over a larger area, and the amplitude is smaller. However, if the tip diameter is too small the acoustic energy is dissipated over a smaller field and is less efficiently used, this is particularly important if the working volume is small because intense turbulence can lessen the power absorbed near the tip.

Manufacturing materials dictate the severity of the wearing problems that are also common with the probe tips. Titanium is a good material because of its good acoustic and mechanical properties and low toxicity to bioactive materials. However, titanium probes are expensive and difficult to machine. Pitting erosion produces deep cavities on the tip surface, particularly when Ballotini beads are included, which lessens the efficiency of sonication. The tip can usually be detached for servicing to produce a new flat surface, which is difficult with titanium, stainless or hard steels are better alternatives despite their poorer acoustic and mechanical properties.

Sonicating equipment can be found in many laboratories. In preparative applications the performance is generally not as good as with the high-pressure homogenizers or bead mills. In the case of plasmid recovery from *E. coli*, even when the disruption yield was good, sonication failed to produce intact plasmids (37). The application of this technique to large-scale purification of bioproducts is rather limited: it is difficult to transmit high acoustic power to large volumes, it is difficult to extract the large amounts of heat generated during the cavitation process, and the equipment has to be shielded to protect humans.

Other Methods

The mechanical methods are, generally, effective in disrupting tough microbial cell walls and much more efficient than the nonmechanical methods. Concomitantly, there are also several disadvantages: they are energy intensive, generate high temperatures and shear forces that can damage labile products, are nonspecific with respect to the organisms and the product to be released, and cause intense cell fragmentation. Although several attempts have been made to optimize operations and minimize these effects, process modifications that could result in a significant reduction in the overall energy requirements are still being pursued.

Nonmechanical methods are not efficient on their own, and very often they only increase the permeability of the cell envelope without breaking the wall. However, by weakening the cell wall, lytic methods can be combined with the more efficient mechanical methods, and therefore require less harsh conditions.

Among the physical methods of lysing cells, two have been reported more frequently, although they are not production oriented: osmotic shock and freeze–thawing. Osmotic stresses are generated when a cell is placed in either a hypertonic or a hypotonic medium; in the first case the cell tends to shrink (plasmolyze), and in the second case to expand. In both cases the cells can suffer potentially serious injuries. Usually the cells have osmoregulating mechanisms in addition to tough cell walls that protect them from lysis. However, if suddenly placed in a hypotonic medium (a medium of very low salt concentration), the cells suffer an osmotic shock, absorb water, and swell; eventually they reach a point of unsustainable internal pressure and burst. This bursting process is effective and upscaleable with weak animal cells. However, osmotic shocks are rather inefficient with tougher bacterial, fungal, or yeast cells, though some periplasmic enzymes and even some proteins/enzymes located near the inner side of the cytoplasmic membrane can be selectively released. Osmotic shock is frequently used to prepare membrane vesicles for studies of membrane proteins. The important operational parameter is the time to fully resuspend the cells in the hypotonic medium, which should be as short as possible to avoid the osmoadaptation of the organisms.

The release of intracellular components can also be induced by freeze–thawing the cells. The ice crystals that form and grow during the freezing step will mechanically disrupt the integrity of the inner cell membranes, which become permeable and allow several enzymes to be extracted into appropriate buffers during the thawing step. The precise mechanism for the release of proteins has not yet been elucidated; however, the formation of small transient pores (44) in the membranes is a logical explanation consistent with the facts that protein solubility is important in governing its release to the medium, the inclusion bodies remain unreleased, and large molecular weight proteins are excreted with very low yields (45). The efficiency of this method depends largely on the rates of thawing. With a large mass of frozen suspension, the rate of thawing is higher in the outside region, while the interior of the block remains frozen, allowing chemical or enzymatic (eventually, proteolytic) reactions to proceed to different degrees in different parts of the block. On the other hand, the lifetime of the transient pores (regions of discontinuity) depends on the temperature of thawing that affects the rate of membrane resealing (44) and is longer at low temperatures. This may explain the lack of reproducibility of the results in the ball mill disintegration of brewers' yeast samples that were kept frozen and separately thawed (24). The freeze–thawing process is not easily scaled up and find application mainly in laboratory-scale recovery of recombinant proteins. It can also be useful in disrupting pathogenic cells in a sealable package (7).

Chemical methods on lysis are milder and often anticipate an autolytic enzymatic activity. Organic solvents (e.g., toluene, ethanol, *n*-butanol, isobutanol, acetone, methylethyl ketone, etc.) are believed to extract, nonselectively, lipids from the cell wall or cytoplasmic membrane, therefore increasing the permeability to intracellular components. However, organic solvents might cause protein denaturation, and by diffusing into the cell, they also de-

stroy the membranes of intracellular organelles, thus allowing destructive enzymes, such as proteases, to be liberated. These harmful effects of the organic solvents were minimized in the extraction of β-D-galactosidase by mixing yeast cells (*Kluyveromyces fragilis*) with concentrated selected solvents and, following filtration, resuspending and agitating the cells in an adequate buffer for extraction (46); 95% of the enzyme was released at the end of a 20-h extraction period. The process was claimed to be suited for scale-up, and the organic solvents were beneficial in killing the viable yeast cells, therefore avoiding the need to separate them from the lysate and the addition of antimicrobial agents.

The use of concentrated organic solvents, however, imposes a fireproof environment. Other chemicals, such as detergents, chaotropic agents (e.g., guanidine HCl or urea), EDTA, phenethyl alcohol (PEA), or alkaline solutions have been used to increase the permeability of the cell walls. All chemical methods suffer from long holding times and, thus, have long been considered unattractive for larger-scale applications. Addition of glycine in low concentrations to the medium at an appropriate pH could also increase the permeability of recombinant *E. coli* immobilized in κ-carrageenan gel beads to α-amylase.

However, the recent progress in recombinant DNA technologies has renewed interest in the development of methods to increase the permeability of the cell envelope in order to avoid a detrimental accumulation of plasmid-encoded proteins in the cytoplasm. If these proteins are allowed to accumulate, further synthesis is reduced while proteolytic degradation may cause loss of protein activity, or the plasmids may loose stability. The benefits of whole viable cell immobilization may also be lost if the foreign proteins are not excreted and the cells would require a disruption step for the product to be released and further purified. Chemical methods to recover these active recombinant products in situ, with increased selectivity concerning the intracellular location, have been investigated, with great promise. *E. coli*, a Gram-negative bacterium, has typically been used as the host organism, and many high-value recombinant proteins are secreted into its periplasmic space. Treatment of *E. coli* with guanidine HCl (a chaotropic agent) increases the permeability of the outer membrane, and periplasmically expressed proteins such as β-lactamase (47) can thus be leaked into the surrounding medium with high yield (better than 80% in 5 h), free of many intracellular contaminants. After treatment of the same cells with Triton X-100, which attacks the inner membrane, the cytoplasmic proteins could be released but with low yield and only when the cells were treated with both the guanidine HCl and Triton X-100 was the yield of intracellular protein higher than 50%. The cost of Triton X-100 is high but, nevertheless, this process was claimed to be less expensive than enzymatic methods. The release of β-lactamase was also enhanced by treating recombinant *E. coli* with EDTA and PEA at appropriate concentrations; cell viability and synthesis of the plasmid encoded β-lactamase were retained (49).

Pretreatment with guanidine HCl and Triton X-100 was also used in conjunction with a high-pressure homogenization step to disrupt recombinant *E. coli* expressing human growth hormone inclusion bodies (48). A reduced number of passes through the homogenizer at lower disrupting pressure were required to recover the protein from the inclusion bodies; due to their synergistic effects, Triton X-100 was used to reduce the concentration of guanidine HCl required to weaken the cell wall and to wash the inclusion bodies from contaminant proteins.

Pretreatment of *Candida lipolytica* with an alkaline solution resulted in a weakened cell wall and increased protein solubility, and further mechanical homogenization was thus facilitated at lower pressures than with non-treated cells (20).

Enzymic methods of cell wall hydrolysis are also becoming increasingly more attractive. The advantages associated with enzymic lysis include an increased selectivity of product release (which is desirable for further product isolation and purification), an increased rate and yield of product release, the minimization of product damage, requirements for mild conditions of pH and temperature, and no debris being produced.

More than any other methodology, enzymatic lysis requires a good understanding of the details of the chemical and physical structure of the cell walls. No single enzyme is capable of fully degrading the cell walls. As already described, peptidoglycan is the structural element in bacterial cell walls. In Gram-positive bacteria a single enzyme could lyse the cell, but a mixture of several enzymes (glycosidases [endoacetylmuramidase, *endo*- and *exo*acetylglucosaminidase], *N*-acetylmuramyl L-alanine amidase or peptidases) acting synergistically would assist in the lysis of the bacterial peptidoglycan; in Gram-negative species a pretreatment with a surface-active agent is usually required to remove the outer membrane. For yeast cells a wall lytic protease and β-(1–3)-glucanase are essential, but

Table 1. Important Microbial Cell Wall Degrading Enzymes

Organisms	Enzymes	Type of hydrolyzed linkage
Bacteria	Glycosidases	β-(1,4)-linkages between AGA and AMA residues in a peptidoglycan
	Acetylmuramoyl-L-alanine amidases	Link between *N*-acetylmuramoyl residues and L-amino acid residues in certain glycopeptides
	Peptidases	Gly-gly, ala-gly, etc., peptide bonds
Fungi, yeasts	β-(1,3)-Glucanases	Random β(1,3)-linkages in glycans
	β-(1,6)-Glucanases	Random β(1,6)-linkages in glycans
	Mannanases	(1,2)- (1,3)-, (1,6)-β-D-Mannosidic linkages
	Chitinase	*N*-Acetyl-*b*-D-glucosaminide β(1,4)-B-linkages in chitin and chitodextrins
	Proteases	
Algae	Cellulases	α(1,4)-Linkages in cellulose

the inclusion of β-(1–6)-glucanase, mannase, and chitinase assist the lytic process (Table 1). Most lytic enzyme systems are active against pretreated nonliving cells (lyophilized or heat-killed cells) or against cell walls.

The main restrictions in using multienzyme systems are the availability and high costs associated with their preparation. Different sources of lytic enzymes have been investigated. Various species of *Arthrobacter* can be induced, by whole yeast cells or cell walls, to produce yeast lytic activity (β-(1–3)-glucanase, mannase, and protease) subject to catabolite repression by glucose. Yeast lytic activity was also identified in *Cytophaga* sp. and *Oerskovia xanthineolytica*. Different strains of *Streptomyces* sp., *Pseudomonas* sp., *Staphilococcus* sp., and *Cytophaga* sp. have a wide spectrum of bacteriolytic activities. Lysozyme from hen egg white is bacteriolytic against only a few bacteria.

For the extraction of a given product from an organism, the selection of the most appropriate lytic system depends on the nature of the wall to be lysed and the nature and location of the product inside the cell. The composition of a multienzyme system determines the efficiency and the selectivity of the lytic process (50). If the product is proteinaceous, a highly proteolytic system is undesirable. Thus, if the organism to be lysed is a yeast, a high proteolytic activity should be provided in only the early part of the lytic process, when the substrate is the mannan–protein complex of the wall; the protease should then be inactivated, or a protease inhibitor should be added. However, the composition of lytic systems can be manipulated in part in the production phase, using different inducers and growth conditions. For instance, in the case of the lytic enzyme system from *Cytophaga* sp., high β-(1–3)-glucanase activities and high glucanase: protease ratios were obtained when produced in continuous culture at low dilution rates under carbon-limited conditions. The quantitative data concerning the effects of pH, temperature, concentrations of lytic enzyme, and of cells on the lysis rate should also be obtained to optimize a lytic process (50).

As with all the nonmechanical methods, extensive enzymatic lysis is a very slow process, but in conjunction with others, the rate and the yield of product extraction can be much increased. Mutants that are defective in β-glucan can be produced by mutagenesis and are, therefore, more susceptible to β-(1–3)-glucanase (commercially known as zymolyase) attack (51). Enhanced disruption of *C. utilis* was obtained by high-pressure homogenization (using a microfluidizer) when the cells were pretreated with zymolyase (51). *Bacillus cereus* pretreated with 0.5 mg cellosyl (a lytic enzyme produced by *Streptomyces coelicolor*) per gram of cells (wet wt), at pH 5.5 and 30 °C, was disrupted with 98% yield in one single passage in a Manton–Gaulin homogenizer at 7×10^7 Pa, as compared with less than 40% yield when the cells were untreated. The same enzymatic treatment facilitated disruption in a bead mill (52).

Autolysis is the process by which the cells use their own lytic capabilities and produce lytic enzymes. Autolysis, as a self-digestion process, is very slow and often requires several days, which is very inconvenient. Because of that, it is not a common method of cell disruption; however, yeasts are autolyzed on an industrial scale for the production of yeast autolysates. The process can be speeded up by drying or freeze-thawing the cells. As already mentioned, autolysis can also be induced by plasmolysis with surfactants, thiol reagents, and organic solvents (e.g., toluene or butanol) or by aging or temperature. β-Lactam antibiotics are known to cause the activation of autolytic enzymes (autolysins), and this has also been suggested to be the mechanism by which surfactants induce the autolytic process. Autolysis enzymes are important for cell growth and reproduction, but once they are deregulated, they attack bacterial peptidoglycan. A key factor is the pH, which is generally optimum around 8.0 when yeast is to be plasmolyzed by organic solvents or detergents (56).

Microbial cells differ in their potential for autolysis; often they lack the genetic information to produce lytic enzymes, in which case the genetic coding must be introduced from outside. Lytic enzyme systems that can be cloned include not only the autolysins (several peptidoglycan hydrolases) but also colicin enzymes (a class of antibiotic proteins) and phages (57).

EFFECTS OF CELL DISRUPTION ON DOWNSTREAM OPERATIONS

One of the most challenging problems in solid–liquid separation is the removal of cell debris, which includes cell wall fragments, cell membranes, and protein precipitates. Disruption and lytic treatments cause a reduction in the sedimentation properties of the cells and fine debris (<0.1 μm), leading to inefficient centrifugation and/or filtration due to changes in the surface properties, high compressibility of the solids, high viscosity of the disrupted suspension, and a small difference between the densities of the solids and the suspending medium (53). *Centrifugation* has been used, but it is unsatisfactory when the particles are smaller than 0.5 μm, particularly in viscous homogenates, therefore requiring very low flow rates (on the order of 10 times less than for whole cells). A previous flocculation can improve the centrifugation efficiency. A highly selective *flocculation* of cell debris by borax (sodium tetraborate) was improved by further addition of poly(ethylene imine) to the clarified homogenate supernatant (54). *Rotary vacuum precoat filtration* has been used for the clarification of disrupted yeasts, but *crossflow microfiltration* has the advantage of preventing cake formation and avoiding the need for expensive filter aids. The separation of cell debris is, however, the most complex problem that crossflow microfiltration has been tested against. Membrane fouling and concentration polarization are some of the difficulties that contribute to a rapid flux decline and limit the concentration of debris that can be attained. Apart from the influence of the operational parameters (pressure, feed velocity, fouling factors) on the flux and transmission, which are well documented in the literature, we have to consider that strong adsorption of proteins to cell debris may occur in the lysate, depending on the concentration, temperature, pH, or ionic strength. *Aqueous two-phase extraction* is attractive for the recovery of whole cells and debris from protein solutions but involves high costs and difficult recovery of reagents. Phase

components include water-soluble polymers and electrolytes, the best-known systems being PEG-dextran and PEG-phosphate.

Separation techniques depend on many specific aspects, but the average particle size and the size distribution are crucial to all of them: the maximum amount of released protein increases with increasing micronization of the cell debris (36), but these become increasingly more difficult to separate.

Mechanical methods are generally, effective in disrupting tough microbial cell walls. The major difference among them lies in the different sizes of cell fragments that are generated; some methods disrupt the cell envelope with less overall destruction than others. This is significant because the size of the fragments has an important impact on downstream processing when they are removed from the lysate or separated from other subcellular species (37): the submicron particles are easily resuspended in the supernatant after centrifugation, hinder the fractional precipitation, and tend to foul chromatographic columns or membranes that follow (37,54).

The behavior of *E. coli* cell debris from a bead mill and high-pressure homogenizers was investigated (53). No significant differences in the yield and rate of release of protein were reported. However, the characteristics of the debris (particle size distribution, separability by centrifugation, and dead-end filtration) were quite dependent on the disruption method. The mean residence time in the bead mill did not significantly affect the particle size distribution, but the number of passes in the mechanical homogenizer broadened the debris size distribution.

The effect of operating conditions on the cell debris size distribution was investigated for bakers' yeast disrupted by high-pressure homogenization, and a two-parameter model was developed, incorporating both the operating pressure and the number of passes, to simulate the particle size distribution in the homogenate. A threshold pressure of 11.5 MPa was determined, below which no disruption of yeast cells will occur (55). Yeast cell debris particle size distribution was also reported to depend on the growth conditions (13): in batch cultivation the median diameter, d_{50}, was less in the exponential phase and greater in the stationary phase; in continuous culture the median diameter was less for higher dilution rates (younger cells).

High-pressure homogenization of of *E. coli* cells treated with a combination of guanidine HCl and Triton-X-100 produced cell debris of larger sizes than when the pretreatment was made with only guanidine HCl, which is highly desirable when the debris are to be separated by cross-flow microfiltration.

The cell suspension should also be conditioned to avoid unwanted degradation of labile products liberated from the disrupted cells. Protease inhibitors are often required and should be added to the cell slurry. Antioxidants are useful to avoid oxidation of S—S bonds. An appropriate buffer should be included to maintain the pH within the desired range. All the chemicals that are added to the system in all phases of the process (starting from the medium preparation and fermentation steps) may have effects on the product and subsequent operations; this should be taken into account.

Evolving from the recent developments in recombinant DNA technology, a large number of proteins of therapeutical importance, now in use or undergoing clinical evaluation, are produced as inclusion bodies (or "refractile" bodies) in the cytoplasm, mainly in *E. coli*, where they accumulate due to the absence of a secretory mechanism. These inclusion bodies are dense and amorphous aggregates of protein, consisting of misfolded (frequently denatured) polypeptides, often as large (0.8–1.5 μm) as a normal bacterium cell, with a relatively higher density (~1.2–1.35 kg/L) than the rest of the cell material. They require the use of chaotropic agents or detergents to be solubilized in an aqueous environment (58). The standard procedure for the purification of a protein expressed as an inclusion body begins with the isolation of the inclusion bodies, which itself begins with the disruption of the cell. The best cell disruption technique is high-pressure-homogenization. The aggregates are then separated from the soluble components and from the debris of the lysed cells, usually by centrifugation, typically using disc-stack centrifuges, at 5,000–10,000 *g* during 5–10 min; the lighter cell debris is removed along with the supernatant (58,59). Cross-flow membrane filtration has also been investigated (60). The pellets are then washed with either dilute detergents or chaotropic reagents (urea or guanidine HCl) to remove contaminants, and often reducing agents (e.g., dithiothreitol or β-mercaptoethanol) are added to disrupt incorrect disulfide bonds. The solubilized and reduced protein is then ready for subsequent purification and refolding.

BIBLIOGRAPHY

1. M.D. Lilly, *Appl. Biochem. Biotechnol.* **2**, 1–26 (1979).
2. F.C. Neidhardt, J.L. Ingraham, and M. Schaechter, *Physiology of the Bacterial Cell. A Molecular Approach*, Sinauer Associates Publishers, Sunderland, Massachussets, 1990.
3. S. Bartnicki-Garcia, *Annu. Rev. Microbiol.* **22**, 87–108 (1968).
4. R. Bonali, *Sci. Pharm. Biol. Lor.* **2**, 25–40 (1974).
5. J.L. Strominger, and J.-M. Ghuysen, *Science* **156**, 213–221 (1967).
6. E. Work, in J.R. Norris and D.W. Ribbons eds., *Methods in Microbiology*, vol 5A, Academic, London, 1971, pp. 361–418.
7. D.I.C. Wang, C.L. Cooney, A.L. Demain, P. Dunnill, A.E. Humphrey, and M.D. Lilly, *Fermentation Technology*, Wiley, New York, 1979, p. 241.
8. A.P.J. Middelberg and B.K. O'Neil, *Biotechnol. Prog.* **9**, 109–112 (1993).
9. A.P.J. Middelberg, B.K. O'Neil, and I.D.L. Bogle, *Trans. Inst. Chem. Eng.* **70**, Part C, 205–212 (1992).
10. A.P.J. Middelberg, B.K. O'Neil, I.D.L. Bogle, N.J. Gully, A.H. Rogers, and C.J. Thomas, *Trans. IChemE* **70**, Part C, 213–218 (1992).
11. T. Sauer, C.W. Robinson, and B.R. Glick, *Biotechnol. Bioeng.* **33**, 1330–1342 (1989).
12. C.R. Engler, and C.W. Robinson, *Biotechnol. Bioeng.* **23**, 765–780 (1981).
13. S.F. Siddiqi, M. Bulmer, P. Ayazi Shamlou, and N.J. Titchener-Hooker, *Bioprocess Eng.* **14**, 1–8 (1995).
14. D.A. Roberts, Z. Zhang, T.W. Young, and C.R. Thomas, *Proc. Inst. Chem. Eng. Ann. Res. Event*, University College, London, 1994, pp. 73–75.

15. J.R. Woodrow and A.V. Quirk, *Enzyme Microb. Technol.* **4**, 385–389 (1982).
16. G. Hedenskog, L. Enebo, J. Vendlova, and B. Prokes, *Biotechnol. Bioeng.* **11**, 37–51 (1969).
17. A. Heim and R. Czerski, *Bioprocess Eng.* **13**, 23–30 (1995).
18. K.E. Magnesson and L. Edebo, *Biotechnol. Bioeng.* **16**, 1273–1282 (1974).
19. A.P.J. Middelberg, B.K. O'Neil, I.D.L. Bogle, and M.A. Snoswell, *Biotechnol. Bioeng.* **38**, 363–370 (1991).
20. C.H. Lee, S.K. Tsang, R. Urakabe, and R.A. Rha, *Biotechnol. Bioeng.* **21**, 1–17 (1979).
21. D.E. Bianchi, *Stain Technol.* **40**, 79–82 (1965).
22. C.R. Englet and C.W. Robinson, *Biotechnol. Bioeng.* **21**, 1861–1869 (1979).
23. P.J. Hetherington, M. Follows, P. Dunnill, and M.D. Lilly, *Trans. Inst. Chem. Eng.* **49**, 142–148 (1971).
24. F. Marffy and M.-R. Kula, *Biotechnol. Bioeng.* **16**, 623–634 (1974).
25. H. Mogren, M. Lindblom, and G. Hedenskog, *Biotechnol. Bioeng.* **16**, 261–274 (1974).
26. M. Follows, P.J. Hetherington, P. Dunnill, and M.D. Lilly, *Biotechnol. Bioeng.* **13**, 549–560 (1971).
27. F.A.P. Garcia, in J.F. Kennedy and J.M.S. Cabral eds., *Recovery Processes for Biological Materials*, Wiley, Chichester, U.K., 1993.
28. D.E. Hughes, J.W.T. Wimpeny, and D. Lloyd, in J.R. Norris and D.W. Ribbons eds., *Methods in Microbiology*, vol 5B, Academic, London, 1971, pp. 1–54.
29. W.T. Coakley, A.J. Bater, and D. Lloyd, *Adv. Microb. Physiol.* **17**, 279–341 (1977).
30. Y. Chisti and M. Moo-young, *Enzyme Microb. Technol.* **8**, 194–204 (1986).
31. M.D. White and D. Marcus, in A. Mizrahi ed., *Downstream Processes: Equipment and Techniques*, Alan R. Liss, New York, 1988, pp. 51–96.
32. J.S.G. Brookman, *Biotechnol. Bioeng.* **16**, 371–383 (1974).
33. M.S. Doulah, T. Hammond, and J.S.G. Brookman, *Biotechnol. Bioeng.* **17**, 845–858 (1975).
34. T. Sauer, O.W. Robinson, and B.R. Glick, *Biotechnol. Bioeng.* **33**, 1330–1342 (1989).
35. E. Keshavarz Moore, M. Hoare, and P. Dunnill, *Enzyme Microb. Technol.* **12**, 764–770 (1990).
36. E. Keshavarz, J. Bonnerjea, M. Hoare, and P. Dunnill, *Enzyme Microb. Technol.* **12**, 494–498 (1990).
37. A. Carlson, M. Signs, L. Liermann, R. Boor, and K.J. Jem, *Biotechnol. Bioeng.* **48**, 303–315 (1995).
38. J. Limon-Lason, M. Hoare, C.B. Orsborn, D.J. Doyle, and P. Dunnill, *Biotechnol. Bioeng.* **21**, 745–774 (1979).
39. H. Schütte, K.H. Kroner, H. Hunstedt, and M.-R. Kula, *Microb. Technol.* **5**, 143–148 (1983).
40. Melendres, A.V., H. Unno, N. Shiragami, and H. Honda, *J. Chem. Eng. Jpn.* **25**, 354–356 (1992).
41. F. Bunge, M. Pietzsch, R. Müller, and C. Syldatk, *Chem. Eng. Sci.* **47**, 225–232 (1992).
42. M.S. Doulah, *Biotechnol. Bioeng.* **19**, 649–660 (1977).
43. R. Kuboi, H. Umakoshi, N. Takagi, and I. Komasawa, *J. Ferment. Bioeng.* **79**, 335–341 (1995).
44. B. Lee, K. McKenna, and J. Bramhall, *Biochim. Biophys. Acta* **815**, 128–134 (1985).
45. B.H. Johnson and M.H. Hecht, *Bio / Technology* **12**, 1357–1360 (1994).
46. D.M. Fenton, *Enzyme Microb. Technol.* **4**, 229–232 (1982).
47. T.J. Naglak and H.Y. Wang, *Enzyme Microb. Technol.* **12**, 603–611 (1990).
48. Bailey, S.M., P.H. Blum, and M.M. Meagher, *Biotechnol. Prog.* **11**, 533–539 (1995).
49. W. Ryan and S.J. Parulekar, *Biotechnol. Prog.* **7**, 99–110 (1991).
50. J.A. Asenjo, B.A. Andrews, J.B. Hunter, and S. LeCorre, *Process Biochem.* **20**, 158–164 (1985).
51. C.V. Baldwin and C.W. Robinson, *Biotechnol. Bioeng.* **43**, 46–56 (1994).
52. G. Vogels and M.-R. Kula, *Chem. Eng. Sci.* **47**, 123–131 (1992).
53. Agerkvist, I. and S.-O. Enfors, *Biotechnol. Bioeng.* **36**, 1083–1089 (1990).
54. A.I. Clarkson, P. Lefevre, and N.J. Titchener-Hooker, *Biotechnol. Prog.* **9**, 462–467 (1993).
55. S.F. Siddiqi, N.J. Titchener-Hooker, and P. Ayazi Shamlou, *Biotechnol. Bioeng.* **50**, 145–150 (1996).
56. K. Breddam and T. Beenfeldt, *Appl. Microbiol. Biotechnol.* **35**, 323–329 (1991).
57. R.L. Dabora and C.L. Cooney, *Adv. Biochem. Eng. Biotechnol.* **43**, 11–30 (1990).
58. A. Mukhopadhyay, *Adv. Biochem. Eng. Biotechnol.* **56**, 61–109 (1997).
59. H.H. Wong, B.K. O'Neilland, A.P.J. Middelberg, *Bioseparation* **6**, 361–372 (1997).
60. S.M. Bailey and M.M. Mragher, *Biotechnol. Bioeng.* **56**, 304–310 (1997).

See also CELL SEPARATION, CENTRIFUGATION; CELL SEPARATION, SEDIMENTATION; MEMBRANE SEPARATIONS; PROTEIN PURIFICATION, AQUEOUS LIQUID EXTRACTION.

CELL IMMOBILIZATION

ATSUO TANAKA
TAKUO KAWAMOTO
Kyoto University
Kyoto, Japan

KEY WORDS

Alginate gel method
Biocatalyst
κ-Carrageenan gel method
Carrier-binding method
Cell
Cross-linking method
Entrapment method
Immobilization

Polyacrylamide gel method
Synthetic resin prepolymer method

OUTLINE

INTRODUCTION

Immobilized cells are defined as "cells physically confined or localized in a certain defined region of space with retention of their catalytic activities, which can be used repeatedly and continuously" (1). This definition is applicable not only to cells, but also to single enzymes, multienzyme conjugating systems, and cellular organelles, that is, to all types of biocatalysts. Bioprocesses using biocatalysts have increasingly received worldwide attention because of their benefits of low environmental pollution and economical utilization of natural resources and energy. They have also contributed to a wide variety of fields. The use of immobilized biocatalysts increases the efficiency of such bioprocesses because they catalyze biochemical reactions under more stabilized states than their free counterparts; moreover, they can be used repeatedly or continuously. In particular, immobilized cells can catalyze complicated reactions and have many useful characteristics that other biocatalysts do not have, as described later. The European Federation of Biotechnology, Working Party on Applied Biocatalysis (formerly, the Working Party on Immobilized Biocatalysts) proposed that immobilized cells should be classified as viable or nonviable, where viable includes growing, nongrowing, and respiring (2). According to this proposal, terms such as treated, resting, growing, and living should not be used. However, in general, these terms are used for classifying the condition of immobilized cells. In the definition of "treated," cells are subjected to chemical or physical treatment before or after immobilization, and resting cells do not show growth during utilization, although they have the ability to grow.

HISTORY

In 1916, Nelson and Griffin fortuitously found that yeast invertase adsorbed on activated charcoal retained its catalytic ability in the hydrolysis of sucrose (3). This is said to be the first demonstration of immobilization of a biocatalyst. Although several reports were published on the immobilization of physiologically active proteins by covalent binding on various supports after this finding, immobilized biocatalysts were not used in practice until Grubhofer and Schleith immobilized several enzymes (carboxypeptidase, diastase, pepsin, and ribonuclease) on diazotized polyaminostyrene resin by covalent binding in 1953 (4). Thereafter, many studies on enzyme immobilization have been carried out, and studies on cell immobilization became to be popular in 1970s.

The first industrial application of immobilized biocatalysts was carried out by Chibata and coworkers of Tanabe Seiyaku Co., Japan, in 1969. Fungal aminoacylase immobilized on DEAE-Sephadex through ionic binding was used for stereoselective hydrolysis of *N*-acyl-D,L-amino acids to produce L-amino acids and *N*-acyl-D-amino acids (5). The first industrial application of immobilized cells was also performed by Chibata and his coworkers in 1973, producing L-aspartate from ammonium fumarate by polyacrylamide-gel-entrapped *Escherichia coli* cells containing a high level of aspartase activity.

Several other processes, such as production of L-alanine, L-malic acid, 6-aminopenicillanic acid, milk having a low lactose content, and high-fructose syrup also have been industrialized with the use of immobilized enzymes and immobilized cells, and a vast amount of literature on immobilization is available (6–15).

SIGNIFICANCE

In biological systems, such as microbial cells, plant cells, and animal cells, various chemical reactions are carried out. These chemical reactions are catalyzed by enzymes that are water-soluble globular proteins. *Compared to chemical catalysts, enzymes as catalysts have the following superior characteristics.*

1. Enzyme-catalyzed reactions take place under mild conditions, such as ordinary temperature, atmospheric pressure, and pH value around neutrality.
2. Enzymes have strict substrate specificity, stereospecificity, regiospecificity, and reaction specificity.

These facts suggest that energy-saving, resource-saving, and low-pollution processes could be designed by using biocatalysts because such processes could be carried out at relative low temperatures and at ambient pressure with little by-product formation. However, biocatalysts are, in general, liable to be deactivated under conditions such as high temperature, extremes of pH, presence of organic solvents, or even conditions suitable for their catalysis. The repeated or continuous use of biocatalysts that are required in order to construct efficient processes is another problem faced when free biocatalysts are used. Immobilization is one of the most useful methods for eliminating some of the disadvantages described above. In general, immobilization makes biocatalysts stable and easy to use repeatedly in a long-term series of batchwise reactions, or continuously in flow systems. Therefore, the

immobilization of biocatalysts is attracting attention worldwide.

At present, immobilized biocatalysts have contributed to a wide variety of fields including (1) the production of useful compounds by stereospecific and/or regiospecific reactions, *(2)* the production of energy by biological processes, *(3)* the selective degradation or removal of pollutants to solve environmental problems, *(4)* analyses of various compounds with high sensitivity and high specificity, and *(5)* utilization in new types of drugs such as slow-releasing medications, in making artificial organs, and so on.

These processes require the immobilization not only of single enzymes that catalyze simple reactions, but also of multienzyme conjugating systems that mediate more complex reactions. These systems often require the regeneration system of ATP and/or coenzymes. Because cells contain metabolic systems that mediate such complicated reactions, immobilization of cells serves as immobilization of such multienzyme conjugating systems. Furthermore, in the case of immobilization of cells, the procedure of extracting enzymes from the cells is no longer required. This avoids the inactivation of enzymes during these time-consuming and expensive procedures, making it possible to use the enzymes in a stable state.

However, immobilized cells have some disadvantages that must be overcome to realize their practical applications.

1. By-products are easily synthesized because the cells contain the various metabolic systems and the many enzymes that catalyze undesirable reactions. To defeat this problem, selection of a proper strain, mutation, or the proper treatment of cells in order to lower the levels of enzymes participating in side reactions and increase the activities of desirable enzymes, or genetic improvement of a cell line is essential.
2. Cell walls and cell membranes of intact cells often hamper the permeation of substrates, products, and other reaction components into and out of cells. In this case, appropriate treatment of cells before or after immobilization is required to destroy the barriers to permeability.

If immobilized cells are kept in a growing state by a proper supply of suitable nutrients, and if their biological functions are useful, as in the case of conventional fermentation procedures, these immobilized cells (immobilized growing cells) seem very advantageous. Immobilized growing cells serve as self-proliferating and self-regenerating biocatalysts, and catalytic activities of immobilized growing cells are expected to be maintained during long-term repeated or continuous use of cells.

However, immobilized growing cells also have some disadvantages as biocatalysts, and ways to overcome these disadvantages must be studied.

1. Various nutrients and energy sources are demanded by immobilized growing cells to maintain the cells growing state(s), and yield may be lowered by the consumption of substrates and/or products as nutrients and energy sources.
2. Contamination of product is liable to occur by cells leaked from carriers.

In spite of these faults, cell immobilization is a key technology used to construct an efficient bioprocess, as immobilized cells have the attractive characteristics described earlier.

METHODS

For the application of immobilized cells, screening or breeding of cells having the desired activity and characteristics is very important. In addition, selection of the appropriate immobilization method is necessary. There is, however, no ideal universal immobilization method fitted to various types of cells at present, although a vast amount of literature on immobilization is now available.

The immobilization methods can be classified into four categories: carrier-binding, cross-linking, entrapment, and combined methods, in which combined methods are a combination of former three. Because each method has its own merits and faults (to be described later), selection should be carried out in consideration of the purpose of immobilization, characteristics of cells, type of reaction, type of reactor, and so on.

Carrier-Binding Method

The carrier-binding method is based on the binding of biocatalysts to water-insoluble carriers through covalent bonds, ionic bonds, physical adsorption, or biospecific binding. Carrier materials should have adequate function to bind biocatalysts, have sufficient mechanical strength; physical, chemical, and biological stability; and nontoxicity. Ease of shaping is also important for applying immobilized biocatalysts to various types of reactors. Furthermore, the economic feasibility should be considered. As carriers, various types of insoluble materials, such as water-insoluble polysaccharides (cellulose, dextran, agarose derivatives, etc.), proteins (gelatin, albumin, etc.), synthetic polymers (polystyrene derivatives, ion exchange resins, polyurethane, etc.), and inorganic materials (brick, sand, glass, ceramics, magnetites, etc.) have been developed and utilized as they are, or after proper modification or activation. The development of new carrier materials having desired characteristics is also important.

Any reactive components of the cell surface, for example, amino, carboxyl, thiol, hydroxyl, imidazole, or phenol groups of proteins or saccharides, can be used for covalent bonding between cells and a modified or activated carrier. The covalent-binding method is one of the most frequently used techniques for the immobilization of enzymes because of the following advantages: *(1)* enzymes do not leak or detach from carriers during utilization because of the tight binding, *(2)* enzymes can easily contact with substrates because enzymes are localized on the surface of carriers, and *(3)* an increase in stability is often observed because of the strong interaction between enzymes and carriers. However, this method is not as widely used for cells because of

the toxicity of the agents needed for formation of covalent bonds, and the difficulty in finding the conditions for immobilization. In the limited number of reports on cell immobilization by this method, inorganic and organic carriers have been used as follows: Navarro and Durand (16) immobilized *Saccharomyces carlsbergensis* cells by covalent binding to porous silica beads that had been treated with aminopropyltriethoxysilane and activated with glutaraldehyde. *Serratia marcescens, Saccharomyces cerevisiae*, and *Saccharomyces amurcea* cells were also immobilized by covalent binding to borosilicate glass and zirconia ceramics with an isocyanate coupling agent (17). In the case of organic carriers, *Micrococcus luteus* cells were immobilized to the carboxyl group of agarose beads in a two-step process that avoided exposure of the cells to carbodiimide (18), and the uronic acid–forming activity of these immobilized cells was retained, even though cell viability was lost. A disadvantage of this method is that leakage of the enzymes may occur due to autolysis during the reaction.

Recently, Klein and Ziehr have summarized the immobilization of cells by adsorption to carriers (19). One of the first studies on immobilized cells (20,21) is on the adsorption of *Escherichia coli* and *Azotobacter agile* cells to Dowex-1 for the oxidation of succinic acid. This adsorption method is very simple and mild, and can be applied to the immobilization of various types of cells. However, adsorption of cells to carriers is sensitive to environmental factors such as pH, ionic strength, concentration of reactants, and temperature because the adsorption phenomenon is based primarily on van der Waals forces, ionic or hydrogen bonds, or hydrophobic interactions between the cell surface and the carrier. Therefore, cells are liable to leak from carriers.

Biospecific binding with lectins introduced on carriers was also used for immobilization of cells (22,23). Lectins that are glycoproteins bind tightly to specific carbohydrate residues on the cell surface. Concanavalin A, which can be obtained from jack beans, is one of the most useful lectins.

Cross-Linking Method

The cross-linking method utilizes bi- or multifunctional reagents, but carriers are not used. Cells are linked with each other by the reagents, and the crosslinked cells form pellets.

Glutaraldehyde is one of the most popular cross-linking reagents (24,25), and combines amino groups through a Shiff base linkage. Diisocyanates, such as toluene diisocyanate and hexamethylene diisocyanate, represent another type of cross-linking reagent, forming a urea linkage with amino groups.

E. coli having high aspertase activity was immobilized by this method with glutaraldehyde or toluene diisocyanate (26). A 34% retention of aspartase activity was obtained in the case of glutaraldehyde, but the immobilized cells were inactivated in the presence of toluene diisocyanate. So far there have been very few studies on this method, but *Bacillus coagulans* immobilized by this method is commercially available for immobilized glucose isomerase and is widely used for isomerization of glucose in the industrial field.

Entrapment Method

Entrapment methods can be classified as follows: lattice type, microcapsule type, liposome type, and membrane type.

In lattice type entrapment, biocatalysts are entrapped in gel matrices of various types of polymers. Microcapsule type is entrapment in microcapsules of semipermeable synthetic polymers. Liposome type is entrapment within a liquid membrane prepared from amphiphatic compounds. In membrane type, biocatalysts are separated from the environment as spent reaction solution by ultrafiltration membranes, microfiltration membranes, or hollow fibers.

Entrapment methods are very popular for immobilization of cells. Microbial cells, as well as plant and animal cells, can be immobilized by this method. However, this method also has disadvantages, as follows:

1. High molecular weight substrates can hardly access the cells entrapped.
2. Renewal of carriers is hardly possible.

Membrane-type entrapment can avoid these disadvantages, although it has no effects on stability of cells. The lattice-type method is most widely applied for preparing immobilized cells, and there is a great deal of literature on this method. Several representative techniques for the lattice-type method are described next.

Polyacrylamide Gel Method. Bernfeld and Wan (27) reported the entrapment of several enzymes in polyacrylamide gels in 1963, and this method was applied to the cells of the lichen *Umbilicaria pustuluta* by Mosbach and Mosbach in 1966 as reviewed by Chibata et al. (28). Various types of biocatalysts have so far been immobilized by this method. This method was used for not only the first industrial application of immobilized microbial cells as described earlier, but also for the first application of immobilized growing cells, that is, the production of L-glutamic acid by *Corynebacterium glutamicum* (29).

The procedure for the formation of gel to entrap cells is identical to that employed for the preparation of gel commonly used for electrophoresis. For the immobilization of cells by this method, acrylamide, and *N,N′*-methylenebisacrylamide (BIS) as the cross-linking reagent, are mixed with a suspension of cells and polymerized in the presence of an initiator, such as potassium persulfate, and a simulator, such as 3-dimethylaminopropionitrile (DMAPN) or *N,N,N′,N′*-tetramethylenediamine (Fig. 1).

Although this method can be used for various purposes, a major disadvantage of this method is the toxicity of the acrylamide monomer, the cross-linking reagent, the initiator, and the stimulator. In some cases, cells are damaged by free radicals during polymerization. Several analogues or derivatives of acrylamide are also useful for this method.

Table 1 summarizes some examples of cells immobilized with polyacrylamide.

Alginate Gel Method. Several natural polysaccharides, such as alginate, κ-carrageenan, and agar, are excellent gel

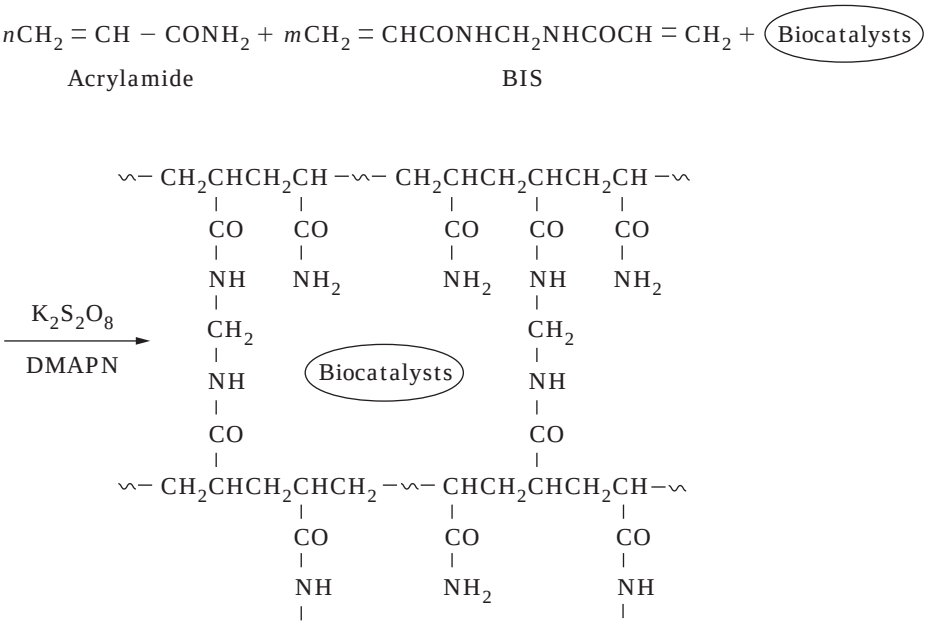


Figure 1. Immobilization with polyacrylamide.

Table 1. Examples and Applications of Cells Immobilized with Polyacrylamide

Cells	Products	Reference(s)
Acromobacter liquidum	Urocanic acid	41
Arthrobacter globiformis	Prednisolone	42
Aspergillus niger	Citric acid	43
Bacillus subtilis	α-Amylase	44
Brevibacterium ammoniagenes	L-Malic acid	45,46
Corynebacterium dismutans	L-Alanine	47
Corynebacterium glutamicum	L-Glutamic acid	29
Curvularia lunata and *Corynebacterium simplex*	Prednisolone	48
Escherichia coli	L-Aspartic acid	29,49,50
Humicola lutea	Acid protease	51
Penicillium chrysogenum	Penicillin G	52
Streptomyces clavuligerus	Cephalosporin C	53
Recombinant *Bacillus subtilis*	Proinsulin	54

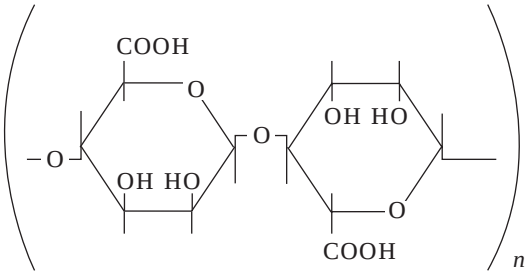


Figure 2. Structure of alginic acid (molecular weight, 50,000–200,000).

materials and are widely used for the entrapment of various biocatalysts.

Alginate, which is extracted from seaweed, is a linear copolymer of D-mannuronic and L-guluronic acids (Fig. 2) and can be gelled by multivalent ions such as calcium and aluminum. For immobilization of cells, sodium alginate, which is soluble in water, is mixed with a suspension of cells and then dropped into a calcium chloride solution to form water-insoluble beads of calcium alginate gels (30).

However, calcium alginate gels are apt to disrupt by mechanical stress and are gradually solubilized in the presence of calcium-ion-trapping reagents, such as phosphate ion and citrate. Leakage of cells is then liable to occur, but free cells can be recovered from the immobilized preparation by this solubilization, which allows the characterization of the cells in gels (e.g., viability). Several multivalent metal ions, such as aluminum ion or strontium ion, can also be used instead of calcium ion. The treatment of formed calcium alginate gels with cationic polymers can improve gel resistivity against the solubilization by calcium-ion-trapping reagents (31). It is also reported that addition of polyacryl acid to sodium alginate in the entrapment by the procedure described can increase the mechanical strength of gels formed and prevent them from dissolving because of bridging polyacrylate chains with calcium ions (32).

The method using alginate gel is widely applicable to immobilization of various useful biocatalysts, including very fragile ones such as plant cells and animal cells, as shown in Table 2, because of the simplicity of the procedure, easy availability of sodium alginate, and mildness of immobilization conditions. It is worthy of note that the first report on immobilization of plant cells employed this method (33).

Table 2. Examples and Applications of Cells Immobilized with Alginate

Cells	Products	Reference(s)
Arthrobacter simplex	Prednisolone	55
Aspergillus niger	Citric acid	56
	Gluconic acid	56
Curvularia lunata	Hydrocortisone	57
Corynebacterium glycinophilum	L-Serine	58
Lactobacillus lactis	Lactic acid	56
Penicillium chrysogenum	Penicillin G	59
Propionibacterium sp.	Vitamin B_{12}	60
Saccharomyces cerevisiae	Ethanol	30
Serratia marcescens	Protease	61
Streptomyces sp.	Tylosin	62
Streptomyces tendae	Nikkomycin	62
Trigonopsis variabilis	α-Keto acid	63
Zymononas mobilis	Ethanol	64
Recombinant *Bacillus subtilis*	Proinsulin	54
Recombinant *Escherichia coli*	β-Lactamase	65
Catharanthus roseus	Ajmalicine isomers	33
Digitalis lanata	Digoxin	33
Lavandula vera	Blue pigments	66
Morinda citrifolia	Anthraquinones	33
Hybridoma cell 4H11	Monoclonal IgA	67

κ-Carrageenan Gel Method. κ-Carrageenan is a readily available nontoxic natural polysaccharide, isolated from seaweed. It is composed of a unit structure of β-D-galactose sulfate and 3,6-anhydro-α-D-galactose (Fig. 3) and is widely used in the food and cosmetic industries as a gelling, thickening, and stabilizing agent.

κ-Carrageenan becomes a gel by cooling or in the presence of a gel-inducing reagent, such as ammonium ion, potassium ion, calcium ion, aluminum ion, magnesium ion, diamines, or water-miscible organic solvents (34). If the proper gel-inducing reagent is used to form a κ-carrageenan gel in the presence of cells, immobilized cells can be easily obtained.

κ-Carrageenan gels can be easily solubilized in saline or in water, and a suspension of free cells can be obtained to investigate their characteristics. On the other hand, leakage of cells easily occurs by dissolution of gel when a gel-inducing reagent is not added to the reaction mixture.

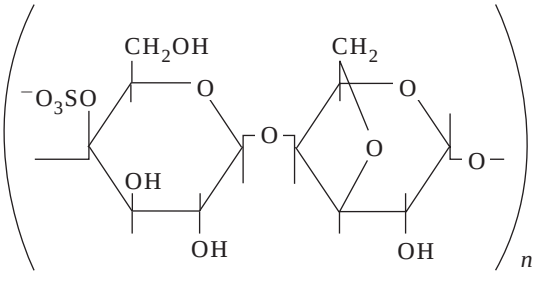


Figure 3. Structure of κ-carrageenan (molecular weight, 100,000–800,000).

In order to increase the stability of the κ-carrageenan-entrapped cells, treatment with glutaraldehyde and hexamethylenediamine after entrapment is often effective.

The conditions of immobilization by this method are also mild, and this method can be used for immobilization of various biocatalysts (Table 3). This method has replaced the polyacrylamide gel method in industrial production of L-aspartate (35) and L-malate (36) by means of immobilized microbial cells. Another advantage of this method is that various shapes of immobilized cells can be prepared (34). The typical procedures for preparing various shapes of κ-carrageenan gel-entrapped cells are as follows:

Cubic type. A suspension of cells and κ-carrageenan dissolved in physiological saline are mixed and cooled to about 10 °C to form a gel. After the gel is soaked in a cold potassium chloride solution in order to increase the gel strength, it is cut to give a cubic-type gel preparation.

Bead type. A mixture of κ-carrageenan and cells is dropped into a 0.3 M potassium chloride solution through a nozzle. Bead-type gels are obtained by this procedure.

Membrane type. A mixture of κ-carrageenan and cells is spread on a plate or sheet and soaked in a cold potassium chloride solution. Consequently, a membrane-type gel is obtained.

Synthetic Resin Prepolymer Method. With the application of a variety of bioreactions, including synthesis, transformation, degradation, or analysis of various compounds having different chemical properties under various reaction conditions, gels having controlled physicochemical properties come to be required for entrapment of biocatalysts. Selection of suitable gel materials among natural polymers for each purpose is usually difficult. For this reason, entrapment methods using synthetic resin prepoly-

Table 3. Examples and Applications of Cells Immobilized with κ-Carrageenan

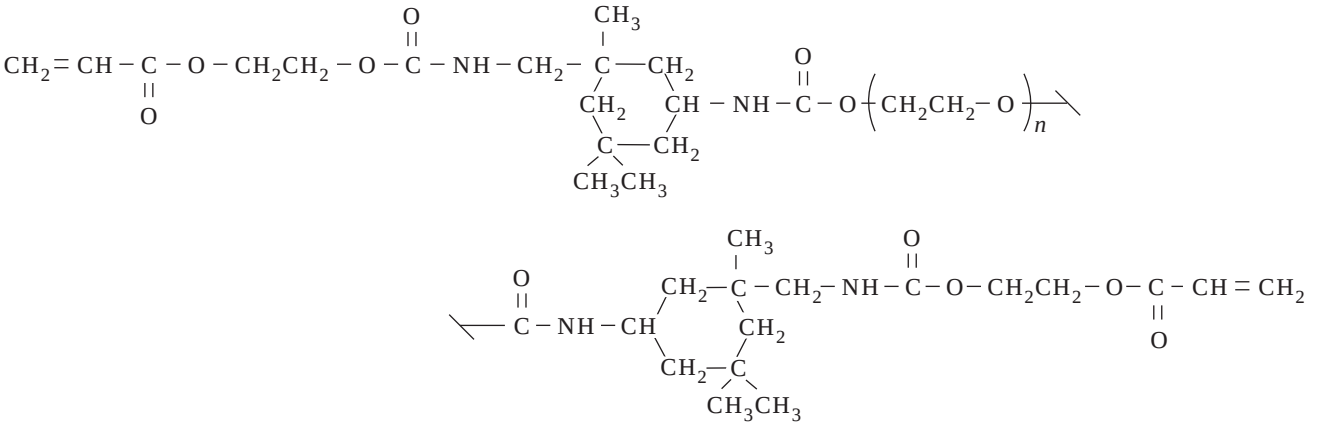
Cells	Products	Reference(s)
Bacillus subtilis	α-Amylase	68
Brevibacterium flavum	L-Malic acid	36,69
Corynebacterium dismutans	L-Alanine	47
Erwinia herbicola	L-DOPA	70
Escherichia coli	L-Aspartic acid	35,71
Penicillium urticae	Patulin	72
Propionibacterium sp.	Vitamin B_{12}	60
Pseudomonas dacunhae	L-Alanine	73
Saccharomyces bayanus	Ethanol	74
Serratia marcescens	L-Alginine	75
	L-Isoleucine	76
Zymomonas mobilis	Ethanol	64
Recombinant *Escherichia coli*	Catechol-2,3-dioxygenase	77,78
Recombinant *Saccharomyces cerevisiae*	Human chorionic gonadotropin	79
Catharanthus roseus	Ajmalicine	80

mers, such as photo-cross-linkable resin prepolymers and urethane prepolymers, are promising. Synthetic resin prepolymers having desired characteristics can be synthesized chemically in advance in the absence of biocatalysts. The methods are as follows:

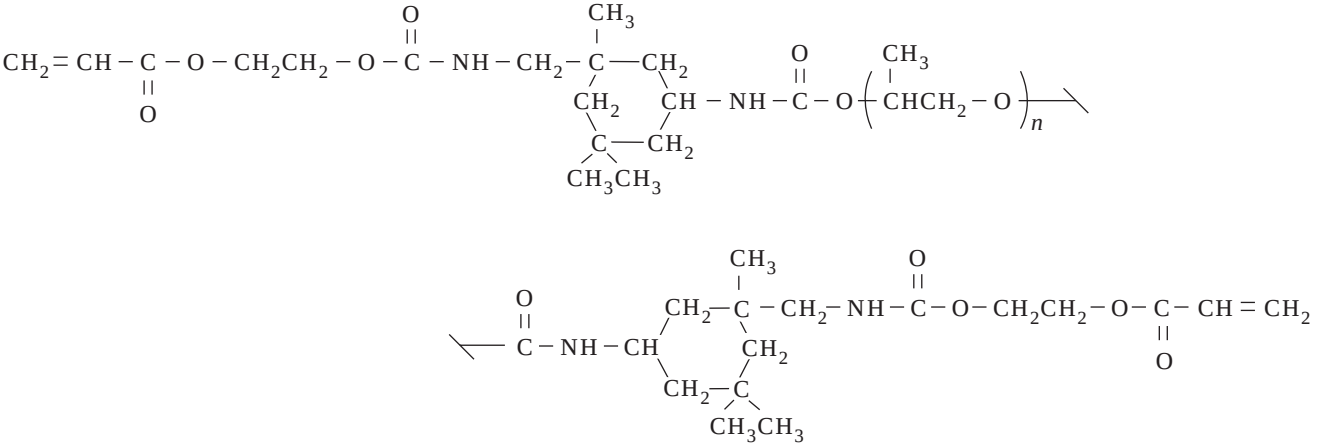
Photo-cross-linkable resin prepolymer method. Entrapment methods using photo-cross-linkable resin prepolymers (Fig. 4) with hydrophilic (ENT) or hydrophobic (ENTP) properties, cationic or anionic nature, and different chain lengths have been developed by Fukui and coworkers (37,38). ENT and ENTP can be prepared from hydroxyethylacrylate, isophorone diisocyanate, and poly(ethylene glycol) or poly(propylene glycol), respectively.

A mixture of the prepolymers and cells, layered (about 0.5-mm thickness) on a transparent polyester sheet is gelled by irradiation with near-ultraviolet light (wavelength rage 300–400 nm) for several minutes in the presence of a proper sensitizer, such as benzoin ethyl ether. This method has been applied for the pilot-scale continuous ethanol fermentation by immobilized growing yeast cells (39). Table 4 lists some examples of cells immobilized with photo-cross-linkable resin prepolymers and their applications. These prepolymers are available from Kansai Paint Co., Japan. Photo-cross-linkable resin prepolymers can be stored in the dark and sterilized at 120 °C.

Urethane prepolymer method. There are also various types of urethane prepolymers (PU) (Fig. 5), which have different hydrophilicities or hydrophobicities and chain length. Urethane prepolymers are synthesized from toluene diisocyanate and polyether diols composed of poly(ethylene glycol) and poly(propylene glycol) or poly(ethylene glycol) alone, although these are not commercially available. Hypol®, which has a similar structure, can be obtained from W. G. Grace Co., Boca Raton, Fla., USA.



ENT



ENTP

Figure 4. Structures of typical photo-cross-linkable resin prepolymers (ENT, hydrophilic; ENTP, hydrophobic).

Table 4. Examples and Applications of Cells Immobilized with Photo-Cross-Linkable Resin Prepolymers

Cells	Products	Reference(s)
Arthrobacter simplex	Prednisolone	81
Bakers' yeast	ATP	82
Brevibacterium ammoniagenes	Coenzyme A	83
Corynebacterium sp.	9α-Hydroxy-4-androstene-3,17-dione	84
Curvularia lunata	Hydrocortisone	85
Curvularia lunata and *Arthrobacter simplex*	Prednisolone	86
Enterobacter aerogenes	Adenine arabinoside	87
Nocardia rhodochrous	3-Keto-Δ^4-steroids	88
	1,4-Androstadiene-3,17-dione	89
Propionibacterium sp.	Vitamin B_{12}	60
Rhizopus stolonifer	11α-Hydroxyprosterone	90
Rhodotorula minuta var. *texensis*	*l*-Menthol	91
Saccharomyces sp.	Ethanol	39

Table 5. Examples and Applications of Cells Immobilized with Urethane Prepolymers

Cells	Products	Reference(s)
Arthrobacter simplex	Prednisolone	92
Enterobacter aerogenes	Adenine arabinoside	87
Escherichia coli	6-Aminopenicillanic acid	93
	L-Aspartate	94
Nocardia rhodochrous	3-Keto-Δ^4-steroids	88
	4-Androstene-3,17-dione	95
Propionibacterium sp.	Vitamin B_{12}	60
Rhodotorula minuta var. *texensis*	*l*-Menthol	91
Streptomyces rimosus	Oxytetracycline	96

Isocyanate functional groups at the terminals of urethane prepolymers react with each other in the presence of water to form urea bonds and liberate carbon dioxide. Therefore, gel is easily formed only by mixing the prepolymer with an aqueous suspension of cells, and entrapment of cells can be easily carried out (38,40). Some examples of cells immobilized with urethane prepolymers are listed in Table 5.

Urethane prepolymers can also be stored in a desiccator and sterilized at 120 °C, avoiding moisture.

In the cases of these synthetic prepolymer methods, the size of the network of gel matrices (which affects diffusion of substrates and products inside the gel, the mechanical strength of the gel formed, the cell-holding capacity of the gel, and the growth ability of cells inside gel matrices) can be regulated by the chain length of the prepolymers and content of the reactive functional groups.

Furthermore, a balance between hydrophilicity and hydrophobicity of gels can be optionally controlled by selecting suitable prepolymers or by mixing prepolymers that have different properties at a proper ratio. The hydrophilicity–hydrophobicity balance of the gels often has a critical effect on the bioconversion by strongly affecting the partition of reactants to the gels.

The Ionic character of the gel, which is also an important factor in bioconversion, can be introduced into the prepolymers by using ionic oligomers as the main chain for their synthesis.

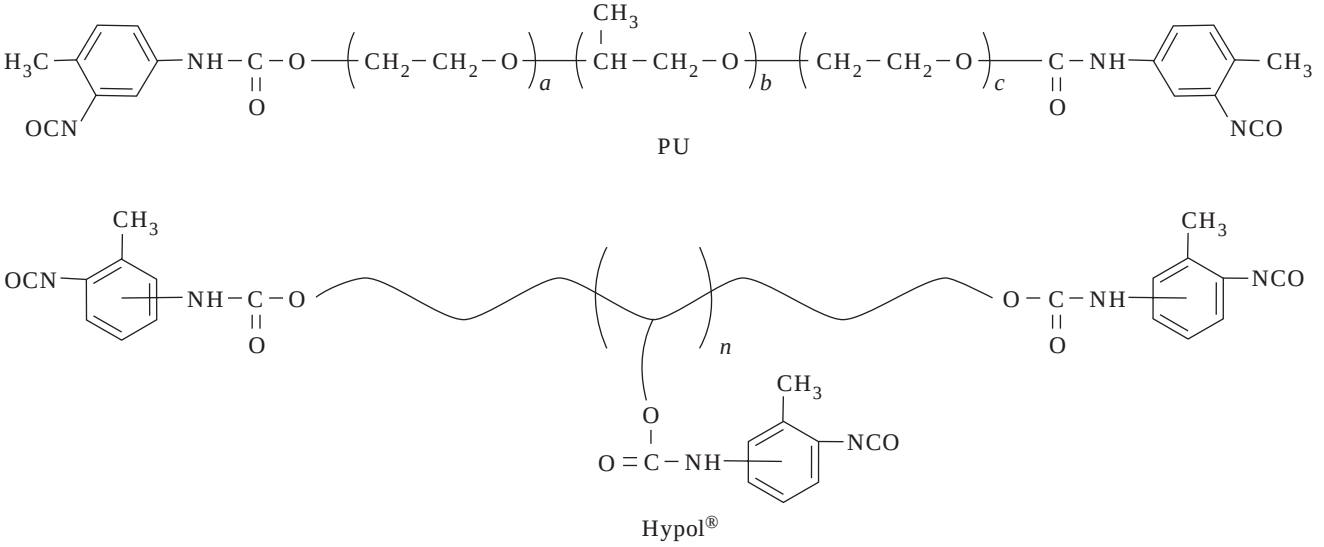


Figure 5. Structures of typical urethane prepolymers.

Other specific features and advantages of the synthetic prepolymer methods are as follows.

1. Procedures for entrapment are very simple and proceed under very mild conditions, and various types of biocatalysts can be immobilized by these methods.
2. Prepolymers do not contain monomers, which may have unfavorable effects on biocatalysts, such as cells. Therefore, the synthetic prepolymer method seems to be one of the most practically applicable immobilization methods for a wide variety of purposes.

BIBLIOGRAPHY

1. I. Chibata, *Immobilized Enzymes, Research and Development* (in Japanese), Kodansha, Tokyo, 1978.
2. M.D. Lilly, *Enzyme Microb. Technol.* **8**, 315 (1986).
3. J.M. Nelson and E.G. Griffin, *J. Am. Chem. Soc.* **38**, 1109–1115 (1916).
4. N. Grubhofer and L. Schleith, *Naturwissenschaft* **40**, 508 (1953).
5. I. Chibata, T. Tosa, T. Sato, T. Mori, and T. Matuo, in *Proceedings of the Fourth IFS. Fermentation Technology Today*, Society of Fermentation Technology, Osaka, Japan, 1972, pp. 383–389.
6. K. Mosbach, *Methods in Enzymology*, vol. 44, Academic, New York, 1976.
7. I. Chibata, *Immobilized Enzyme: Research and Development*, J Wiley, New York, 1978.
8. I. Chibata, T. Tosa, and T. Sato, *Adv. Biotechnol. Processes* **10**, 203–222 (1983).
9. I. Chibata and L.B. Wingard, Jr., *Appl. Biochem. Bioeng.* **4**, 1–349 (1983).
10. B. Mattiason, *Immobilized Cells and Organelles*, Chemical Rubber Co., Cleveland, Ohio, 1983.
11. K. Mosbach, *Methods in Enzymology*, vol. 135, Academic, New York, 1987.
12. I.A. Veliky and R.J.C. McLean, *Immobilized Biosystems, Theory and Practical Applications*, Blackie, London, 1994.
13. R.H. Wijffels, R.M. Buitelaar, C. Bucke, and J. Tramper, *Immobilized Cells: Basics and Applications*, Elsevier, Amsterdam, 1996.
14. R.G. Willaert, G.B. Baron, and L. De Backer, *Immobilized Living Cell Systems*, J Wiley, New York, 1996.
15. G.F. Bickerstaff, *Immobilization of Enzymes and Cells*, Humana Press, Totowa, N.J., 1997.
16. J.M. Navarro and G. Durand, *Eur. J. Appl. Microbiol.* **4**, 243–254 (1977).
17. R. Messing, R.A. Oppermann, and F.B. Kolot, *ACS Symp. Ser.* **106**, 13–28 (1979).
18. T.R. Jack and J.E. Zajic, *Biotechnol. Bioeng.* **19**, 631–648 (1977).
19. J. Klein and H. Ziehr, *J. Biotechnol.* **16**, 1–16 (1990).
20. T. Hattori and C. Furusaka, *J. Biochem.* **48**, 831–837 (1960).
21. T. Hattori and C. Furusaka, *J. Biochem.* **50**, 312–315 (1961).
22. C.H. Bornman and A. Zachrisson, *Plant Cell Rep.* **1**, 151–153 (1982).
23. G.S. Warren and R. Fallon, *Planta* **161**, 201 (1984).
24. F.A. Quiocho and F.M. Richards, *Proc. Natl. Acad. Sci.* **52** 833–839 (1964).
25. A. Schejter and A. Bar-Eli, *Arch. Biochem. Biophys.* **136**, 325–330 (1970).
26. I. Chibata, T. Tosa, and T. Sato, *Appl. Microbiol.* **27**, 878–885 (1974).
27. P. Bernfeld and J. Wan, *Science* **142**, 678–679 (1963).
28. I. Chibata, T. Tosa, and T. Sato, in A.L. Demain and N.A. Solomon eds., *Manual in Industrial Microbiology and Biotechnology*, ASM, Washington, D.C., 1986, pp. 217–229.
29. W. Slowinski and S.E. Charm, *Biotechnol. Bioeng.* **15**, 973–979 (1973).
30. M. Kierstan and C. Bucke, *Biotechnol. Bioeng.* **19**, 387–397 (1977).
31. I.A. Veliky and R.E. Williams, *Biotechnol. Lett.* **3**, 275–280 (1981).
32. T. Mano, S. Mitsuda, E. Kumazawa, and Y. Takeshita, *J. Ferment. Bioeng.* **73**, 486–489 (1992).
33. P. Brodelius, B. Deus, K. Mosbach, and M.H. Zenk, *FEBS Lett.* **103**, 93–97 (1979).
34. T. Tosa, T. Sato, T. Mori, K. Yamamoto, I. Takata, Y. Nishida, and I. Chibata, *Biotechnol. Bioeng.* **21**, 1697–1709 (1979).
35. T. Sato, Y. Nishida, T. Tosa, and I. Chibata, *Biochim. Biophys. Acta* **570**, 179–186 (1979).
36. I. Takata, K. Yamamoto, T. Tosa, and I. Chibata, *Enzyme Microb. Technol.* **2**, 431–437 (1980).
37. S. Fukui, A. Tanaka, T. Iida, and E. Hasegawa, *FEBS Lett.* **66**, 179–182 (1976).
38. S. Fukui, K. Sonomoto, and A. Tanaka, in K. Mosbach ed., *Methods in Enzymology*, vol. 135, Academic, New York, 1987, pp. 230–252.
39. T. Iida, in A. Tanaka, T. Tosa, and T. Kobayashi eds., *Industrial Application of Immobilized Biocatalysts*, Marcel Dekker, New York, 1993, pp. 163–182.
40. S. Fukushima, T. Nagai, K. Fujita, A. Tanaka, and S. Fukui, *Biotechnol. Bioeng.* **20**, 1465–1469 (1978).
41. K. Yamamoto, T. Sato, T. Tosa, and I. Chibata, *Biotechnol. Bioeng.* **16**, 1601–1610 (1974).
42. K.A. Koshcheenko, G.V. Sukhodolskaya, V.S. Tyurin, and G.K. Skryabin, *Eur. J. Appl. Microbiol. Biotechnol.* **12**, 161–169 (1981).
43. H. Horitsu, S. Adachi, Y. Takahashi, K. Kawai, and Y. Kawano, *Appl. Microbiol. Biotechnol.* **22**, 8–12 (1985).
44. T. Kokubu, I. Karube, and S. Suzuki, *Eur. J. Appl. Microbiol. Biotechnol.* **5**, 233–240 (1978).
45. K. Yamamoto, T. Tosa, K. Yamashita, and I. Chibata, *Eur. J. Appl. Microbiol. Biotechnol.* **3**, 169–183 (1976).
46. K. Yamamoto, T. Tosa, K. Yamashita, and I. Chibata, *Biotechnol. Bioeng.* **19**, 1101–1114 (1977).
47. J.M. Sarkar and J. Mayaudon, *Biotechnol. Lett.* **5**, 201–206 (1983).
48. K. Mosbach and P.-O. Larsson, *Biotechnol. Bioeng.* **12**, 19–27 (1970).
49. T. Tosa, T. Sato, T. Mori, and I. Chibata, *Appl. Microbiol. Biotechnol.* **27**, 886–889 (1974).
50. T. Sato, T. Mori, T. Tosa, I. Chibata, M. Furui, K. Yamashita, and A. Sumi, *Biotechnol. Bioeng.* **17**, 1797–1804 (1975).
51. P. Aleksieva, B. Tchorobanov, G. Suchodol'skaya, and K. Koschteenko, *Appl. Microbiol. Biotechnol.* **29**, 239–241 (1988).
52. Y. Morikawa, I. Karube, and S. Suzuki, *Biotechnol. Bioeng.* **21**, 261–270 (1979).

53. A. Freeman and Y. Aharonowitz, *Biotechnol. Bioeng.* **23**, 2747–2759 (1981).
54. K. Mosbach, S. Birnbaum, K. Hardy, J. Davies, and L. Bulow, *Nature*, **302**, 543–545 (1983).
55. S. Ohlson, P.-O. Larsson, and K. Mosbach, *Eur. J. Appl. Microbiol. Biotechnol.* **7**, 103–110 (1979).
56. P. Linko, in M. Moo-Young, C.W. Robinson, and C. Vezina eds., *Advances in Biotechnology*, Pergamon, Elmsford, N.Y., 1981, pp. 711–716.
57. S. Ohlson, S. Flygare, P.-O. Larsson, and K. Mosbach, *Eur. J. Appl. Microbiol. Biotechnol.* **10**, 1–9 (1980).
58. T. Tanaka, K. Yamamoto, S. Towprayoon, H. Nakajima, K. Sonomoto, K. Yokozeki, K. Kubota, and A. Tanaka, *Appl. Microbiol. Biotechnol.* **30**, 564–568 (1989).
59. W. Kurzatkowski, W. Kurytowicz, and A. Paszkiewicz, *Eur. J. Appl. Microbiol. Biotechnol.* **15**, 211–213 (1982).
60. B. Yongsmith, K. Sonomoto, A. Tanaka, and S. Fukui, *Eur. J. Appl. Microbiol. Biotechnol.* **16**, 70–74 (1982).
61. J.C. Vuillemard, S. Terre, S. Benoit, and J. Amiot, *Appl. Microbiol. Biotechnol.* **27**, 423–431 (1988).
62. M. Veelken and H. Pape, *Eur. J. Appl. Microbiol. Biotechnol.* **15**, 206–210 (1982).
63. P. Brodelius, B. Hagerdal, and K. Mosbach, *Enzyme Eng.* **5**, 383–387 (1980).
64. W. Grote, K.J. Lee, and P.L. Rogers, *Biotechnol. Lett.* **2**, 481–486 (1980).
65. G. Georgious, J.J. Chalmers, M.L. Shuler, and D.B. Wilson, *Biotechnol. Progr.* **1**, 75–79 (1985).
66. H. Nakajima, K. Sonomoto, N. Usui, F. Sato, Y. Yamada, A. Tanaka, and S. Fukui, *J. Biotechnol.* **2**, 107 (1985).
67. Y. Shirai, K. Hashimoto, H. Yamaji, and M. Takashiki, *Appl. Microbiol. Biotechnol.* **26**, 495–499 (1987).
68. P. Chevalier, and J. Noüe, *Enzyme Microb. Technol.* **9**, 53–56 (1987).
69. I. Takata, K. Yamamoto, T. Tosa, and I. Chibata, *Eur. J. Appl. Microbiol. Biotechnol.* **7**, 161–172 (1979).
70. G. Para, S. Rifai, and J. Baratti, *Biotechnol. Lett.* **6**, 703–708 (1984).
71. Y. Nishida, T. Sato, T. Tosa, and I. Chibata, *Enzyme Microb. Technol.* **6**, 85–90 (1979).
72. Y.M. Deo and G.M. Gaucher, *Biotechnol. Lett.* **5**, 125–130 (1983).
73. K. Yamamoto, T. Tosa, and I. Chibata, *Biotechnol. Bioeng.* **22**, 2045–2054 (1980).
74. M. Wada, J. Kato, and I. Chibata, *Eur. J. Appl. Microbiol. Biotechnol.* **8**, 241–247 (1979).
75. M. Fujimura, J. Kato, T. Tosa, and I. Chibata, *Appl. Microbiol. Biotechnol.* **19**, 136–139 (1984).
76. M. Wada, T. Uchida, J. Kato, and I. Chibata, *Biotechnol. Bioeng.* **22**, 1175–1188 (1980).
77. P. De Taxis du Poet, P. Dhulster, J.N. Barbotin, and D. Thomas, *J. Bacteriol.* **165**, 871–877 (1986).
78. J. Huang, P. Dhulster, J.N. Barbotin, and D. Thomas, *J. Chem. Tech. Biotechnol.* **45**, 259–269 (1989).
79. S.B. Karkare, D.H. Burke, R.C. Dean Jr., J. Lemontt, P. Souw, and K. Venkatasubramanian, *Ann. N. Y. Acad. Sci.* **469**, 91–103 (1986).
80. P. Brodelius and K. Nilsson, *FEBS Lett.* **122**, 312–316 (1980).
81. K. Sonomoto, A. Tanaka, T. Omata, T. Yamane, and S. Fukui, *Eur. J. Appl. Microbiol. Biotechnol.* **6**, 325–334 (1979).
82. M. Asada, K. Morimoto, K. Nakanishi, R. Matsuno, A. Tanaka, A. Kimura, and T. Kamikubo, *Agric. Biol. Chem.* **43**, 1773–1774 (1979).
83. M. Asada, K. Nakanishi, R. Matsuno, and T. Kamikubo, *Agric. Biol. Chem.* **46**, 1687–1688 (1982).
84. K. Sonomoto, N. Usui, A. Tanaka, and S. Fukui, *Eur. J. Appl. Microbiol. Biotechnol.* **17**, 203–210 (1983).
85. K. Sonomoto, M.M. Hoq, A. Tanaka, and S. Fukui, *J. Ferment. Technol.* **59**, 465–469 (1981).
86. T.K. Mazumder, K. Sonomoto, A. Tanaka, and S. Fukui, *Appl. Microbiol. Biotechnol.* **21**, 154–161 (1985).
87. K. Yokozeki, S. Yamanaka, T. Utagawa, K. Takinami, Y. Hirose, A. Tanaka, K. Sonomoto, and S. Fukui, *Eur. J. Appl. Microbiol. Biotechnol.* **14**, 225–231 (1982).
88. T. Omata, T. Iida, A. Tanaka, and S. Fukui, *Eur. J. Appl. Microbiol. Biotechnol.* **8**, 143–155 (1979).
89. T. Yamane, H. Nakatani, E. Sada, T. Omata, A. Tanaka, and S. Fukui, *Biotechnol. Bioeng.* **21**, 2133–2145 (1979).
90. K. Sonomoto, K. Nomura, A. Tanaka, and S. Fukui, *Eur. J. Appl. Microbiol. Biotechnol.* **16**, 57–62 (1982).
91. T. Omata, N. Iwamoto, T. Kimura, A. Tanaka, and S. Fukui, *Eur. J. Appl. Microbiol. Biotechnol.* **11**, 199–204 (1981).
92. K. Sonomoto, I.-N. Jin, A. Tanaka, and S. Fukui, *Agric. Biol. Chem.* **44**, 1119–1126 (1980).
93. J. Klein and F. Wagner, *Enzyme Eng.* **5**, 335–345 (1980).
94. M.C. Fusee, W.E. Swann, and G.J. Calton, *Appl. Environ. Microbiol.* **42**, 672–676 (1981).
95. S. Fukui, A. Ahmed, T. Omata, and A. Tanaka, *Eur. J. Appl. Microbiol. Biotechnol.* **10**, 289–301 (1980).
96. M. Ogaki, K. Sonomoto, H. Nakajima, and A. Tanaka, *Appl. Microbiol. Biotechnol.* **24**, 6–11 (1986).

CELL SEPARATION, CENTRIFUGATION

HANS AXELSSON
Alfa Laval Separation AB
Tumba, Sweden

KEY WORDS

Asepsis
Cell debris separation
Cell harvesting
Centrifugation
Containment
Scale-up
Separability
Solvent extraction
Spin test
Sterilization
Two-phase aqueous systems

OUTLINE

INTRODUCTION

The first application of centrifugal separation other than in milk processing was in the harvesting of cells in baker's yeast production. Centrifugation is still the only choice for cell separation and protein recovery in many fermentation processes.

CENTRIFUGAL SEPARATION

Centrifugal separation is a sedimentation operation accelerated by centrifugal force. Thus, a prerequisite for the separation is a difference in density between the phases. This applies to both solid–liquid and liquid–liquid separation.

The settling velocity (V_g) of a solids particle (or droplet) under the influence of gravity alone is given by Stokes' law (see "Nomenclature" for abbreviations and symbols):

$$V_g = \frac{\rho_p - \rho_f}{18 \cdot \eta} \cdot d_p^2 \cdot g \quad (1)$$

In the centrifuge field, the settling velocity becomes:

$$V_c = V_g \cdot Z \quad (2)$$

where

$$Z = \frac{\omega^2 \cdot r}{g} \quad (3)$$

is called relative centrifugal force (RCF) or G-number.

It is, however, not the RCF only that determines the performance of a given centrifugal separator. The residence time of the particle or droplet in the centrifugal force field is important, so a larger volume of the rotor will increase the possible flow rate. In the most common centrifugal separators, the disk bowl machines, a given bowl material can cope with a certain peripheral speed—the stronger the bowl material, the higher the speed. This leads to the fact that the maximal RCF is inversely proportional to the bowl diameter, but not necessarily that the lower RCF machine has a lower performance.

TYPES OF CENTRIFUGAL SEPARATORS

General

What determines the design of a centrifuge is the method by which the solids phase in the process material is handled. The first two types of centrifuges developed were applied to duties where the resulting products were in liquid form, such as cream and skim milk and yeast cream and cell-free wort. In the beginning of this century, centrifuges were applied to processes with solids, amorphous or crystalline. In the biotechnological industry, the solids phase is often recovered in a diluted suspension or is, even at high solids concentrations, a slurry with good flow properties.

Types of machines are described in the work by Rushton et al. (1). The most comprehensive description of centrifuges can be found in the book by Sokolov (2). Brochures and technical publications from sedimenting centrifuge manufacturers such as Alfa-Laval Separation AB also provide information on machine types and sizes.

In this article, centrifugal filtration will not be dealt with because of the limited use in this industry, with the exception of crystal recovery in antibiotic production. For a description of various centrifugal filters, see Rushton et al. (1).

Disk Bowl Machines

Solid Bowl. A solid bowl machine (Fig. 1a) has two liquid outlets, each with a paring disk. Other versions have

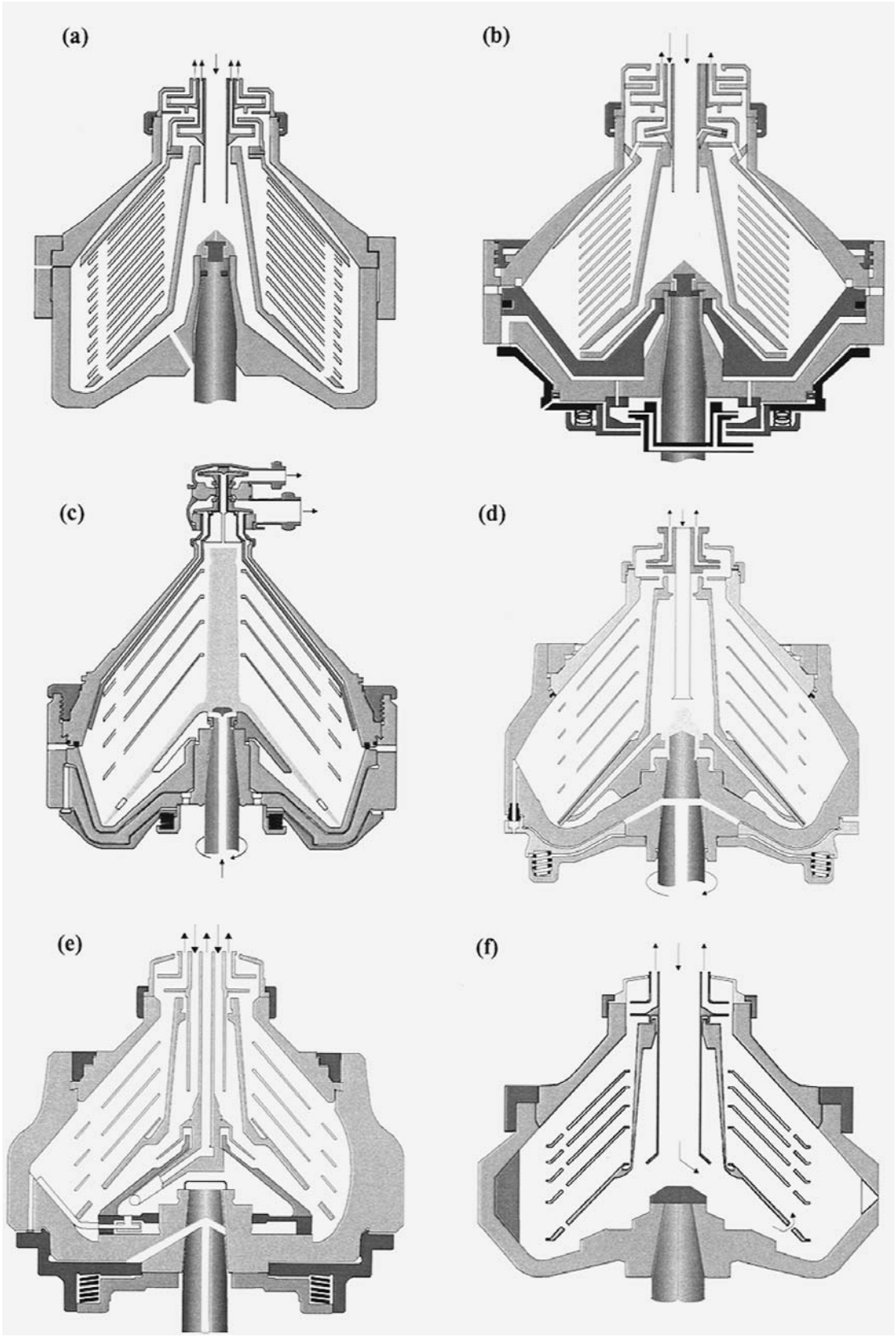


Figure 1. (**a**) Solids-retaining disk bowl, vertical cut. (**b**) Radial solids-ejecting disk bowl, top feed, vertical cut. (**c**) Radial solids-ejecting disk bowl, hermetic feed, vertical cut. (**d**) Axial solids-ejecting disk bowl, vertical cut. (**e**) Nozzle bowl with pressurized solids discharge, vertical cut. (**f**) Nozzle bowl with peripheral discharge nozzles, vertical cut.

none or one paring disk. The solids accumulate at the periphery and must be removed manually. In the separation of two nonmiscible liquids, a coalescence of the dispersed droplets takes place at an internal interface in the bowl. The positioning of this interface is very important. The nearer it is placed to the light liquid outlet, the more contaminated the light liquid will be with droplets of the heavy phase. The position is determined by a gravity disk, over which the heavy liquid flows. The gravity disk diameter is given by the relation:

$$D_h = \sqrt{D_i^2\left(1 - \frac{\rho_{fl}}{\rho_{fh}}\right) + D_l^2 \frac{\rho_{fl}}{\rho_{fh}}} \quad (4)$$

If properly designed and if the feed flow rate is high enough to prevent sedimentation in the bowl, this type of machine can be used to recover a dilute cell concentrate in the heavy liquid outlet. This is also valid for three-phase solids-ejecting machines. Solid-bowl machines are made in a large number of sizes, with bowl diameters ranging from 140 to 750 mm.

Solids-Ejecting Separators, Radial Discharge. Of all types of disk bowl machines, the solids-ejecting separators with radial discharge (Fig. 1b, c) are the most common. With an automated, periodic partial discharge of the sediment, it is possible to obtain a considerably higher concentration of the solids than in the peripheral nozzle machine (Fig. 1f). The machine is discharged in about 0.1 s by pressing a sliding bowl bottom in the bottom part of the bowl downward, thereby opening large discharge ports in the outer bowl wall. To make the discharge at the right moment, the discharge can be initiated by a self-triggering mechanism that hydraulically senses the solids level in the solids space (3). Three-phase separation units are available. The solids-ejecting machines can be equipped with hermetic seals and a hermetic feed from below in a hollow spindle (Fig. 1c). In these true hermetic machines, the split between the exiting light and heavy liquids can be made by controlling the back pressure in the outlet pipes. The heavy liquid can contain large amounts of solids (cells or proteins) if the machine is properly designed. The solids-ejecting separators are available in about 10 sizes, with bowl diameters between 180 and 900 mm. The smaller units are more often available in versions equipped for BL2-LS requirements than are larger units. A pilot-scale sterilizable separator module is seen in Figure 2. A large-scale unit is seen in Figure 3.

Solids-Ejecting Separators, Axial Discharge. When comparing axial and radial discharging separators (Fig. 1b, d), one finds that the axial machine has a different geometry. The lock ring can be placed at a smaller diameter, because no sliding bowl bottom needs to be fitted. This reduces the stresses in the winding of the lock ring. Also, the solids outlet takes place through a few small axial channels that do not induce such large stresses in the bowl shell. These two factors make it possible to increase the bowl speed considerably above that of the radially discharging machines, so that the RCF is at least doubled. The discharge system is operated by pressurized air, which forces a ring slide, equipped with valve seats, downward, opening the axial channels. To facilitate the transport of solids toward the channels, the bowl is machined to have a star-shaped inner contour. This type of machine exists at present in three sizes, with bowl diameters ranging from 500 to 900 mm. BL2-LS versions are available.

Nozzle Machines with Pressurized Discharge of Concentrate. The apparent viscosity of bacteria and yeast suspensions as well as some protein precipitates decreases with increasing velocity gradient. Therefore, they are able to flow even at high concentrations in channels or tubes from the periphery of the bowl in nozzle machines with pressurized discharge of concentrate (Fig. 1e). The tubes empty into a central chamber, where a centripetal pump, a paring tube, picks up the concentrate and pumps it out of the bowl. The nozzles are placed at the end of the concentrate tubes, just in front of the chamber. This machine type is suitable only for applications where the solids have the specific flow properties mentioned earlier and in the absence of other type of solids. In some machines, of this type of vortex chamber is placed in front of the nozzle (4). Liquid enters the chamber tangentially, creating a viscosity-dependent whirl, and leaves through the nozzle in the centre of the chamber. The viscosity effect is giving a self-regulating function so that varying solids flow rates (within limits) to the machine will result in a constant concentration of the concentrate, thereby reducing the risk of clogging the machine. One of the units of this type (Fig. 4) is available in a sterilizable, contained, and aseptic version, conforming to BL2-LS. To make this possible, the cleaning-in-place (CIP) has to be efficient. Therefore, the sterilizable machine uses the discharge system of the machine in Figure 1d during CIP. This system is also used in noncontained versions. Bowl diameters between 500 and 900 mm are available.

Nozzle Machines with Peripheral Nozzles. In nozzle machines with peripheral nozzles (Fig. 1f), nozzles with a diameter of 0.5 to 3 mm are situated at the periphery of the bowl, which has sloping walls toward the nozzles. The number of nozzles varies between 4 in small bowls and 20 in the largest bowls. It can be equipped with a recirculation device for the sediment, a separate central tube that is divided into tubes along the bowl bottom ending just in front of the nozzle. This makes it possible to increase the concentration of the solids phase without the need to decrease the nozzle size, leading to increased risk of nozzle clogging. The nozzle flow rate is proportional to number of nozzles, distance from the center of rotation, bowl speed, and square of the nozzle diameter. The peripheral nozzle machines are used for the largest flow rates; the bowl diameters are up to 1,000 mm.

Decanter Centrifuges

In the 1950s, the solid-bowl separator with scroll discharge (decanter centrifuge) (Fig. 5) was developed for large quantities of solids. It is equipped with a screw conveyor that rotates at a speed slightly higher or lower than the bowl speed.

The differential speed can be varied. Careful control of this parameter, torque control, and some newer design fea-

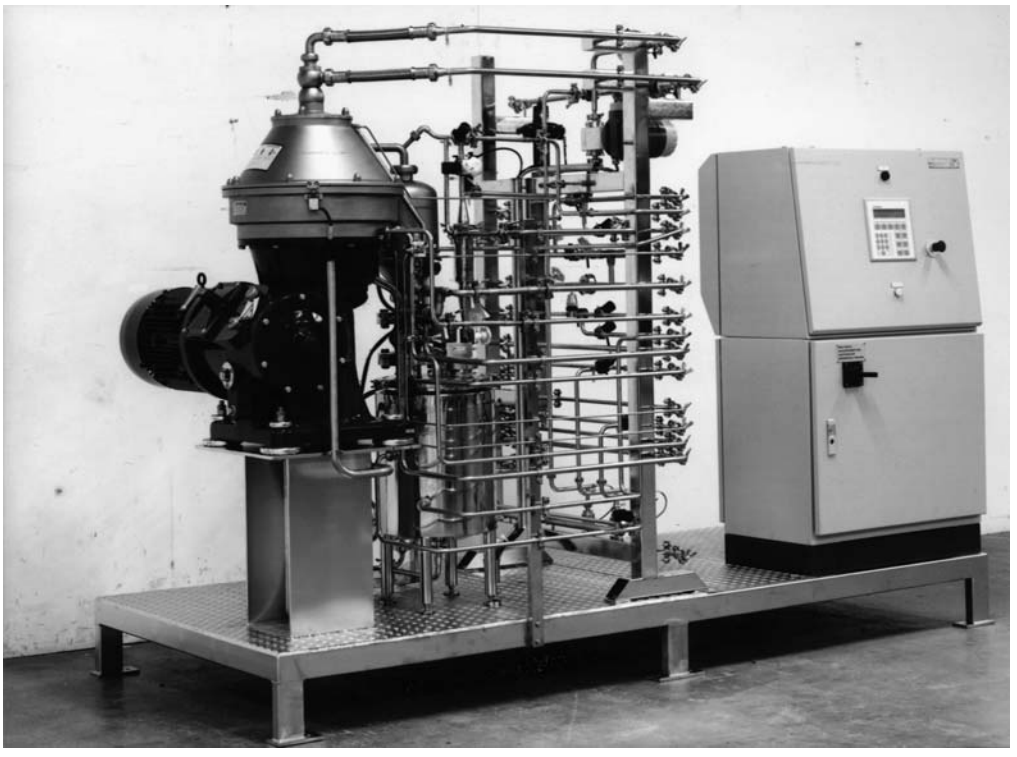


Figure 2. Solids-ejecting disk bowl centrifuge of pilot size for sterilizable contained installation. Type BTPX 205. The height of the separator only is 1.3 m. *Source*: Courtesy of Alfa-Laval Separation AB.

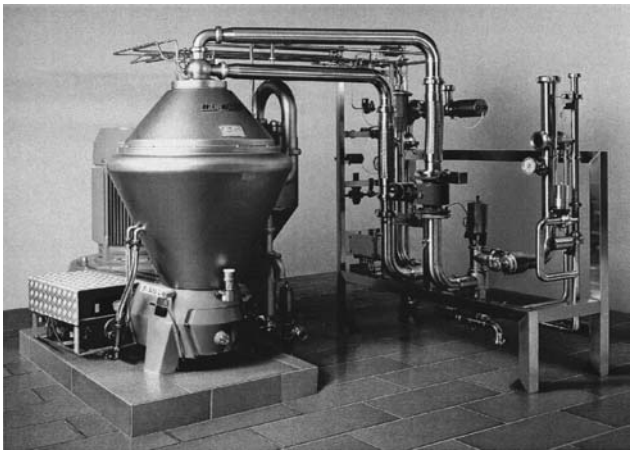


Figure 3. Solids-ejecting disk bowl centrifuge of production size. Capacity in breweries up to 75 m^3/h. Type BREW 2000. The height of the separator is 2.2 m. *Source*: Courtesy of Alfa-Laval Separation AB.

tures (e.g., the baffle disk), have recently made the decanter interesting to use also for biological sludges. However, the conventional mechanical design of these machines does not allow more than 5,000 G for a medium-diameter machine (5). To reach higher *g*-forces new design features are needed (6). With suspended main bearings, a separate gear box, and a swimming conveyor, held in place by the liquid, 10,000 G is feasible.

For three-phase decanters, the gravity disk dimensioning follows in principle equation 4. It is, however, very important to take the weir head into account, because a variation of the light-to-heavy phase ratio gives a very unstable interface position (7).

The range of bowl diameters is very wide, from 150 to 1,200 mm; the bowl length is often 4 to 5 times the diameter.

The newest addition to the decanter family is a vertical decanter equipped with a disk stack (8). The thought is to combine the best characteristics of a decanter (dry solids) and a disk bowl machine (good clarification).

Other Classical Bowl Designs

Tubular Bowl. In the tubular bowl (Fig. 6) the rotor, driven from above, consists of a long cylinder, into which liquid is pumped from below. The design shown in Figure 6 is equipped for separation of two liquid phases. Most tubular centrifuges are used for separating very valuable solids from a liquid. The liquid is pumped out by a paring disk, avoiding aerosol and foam formation. To remove the solids accumulated at the periphery, the machine must be stopped and dismantled. The bowl diameters are up to 130 mm with a solids volume of up to 6 L. In the industrial models, 20,000 G are generated. Sterilizable and contained units are available.

Multichamber Bowl. In the multichamber bowl machine (Fig. 7), the process liquid flows in the annuli between con-

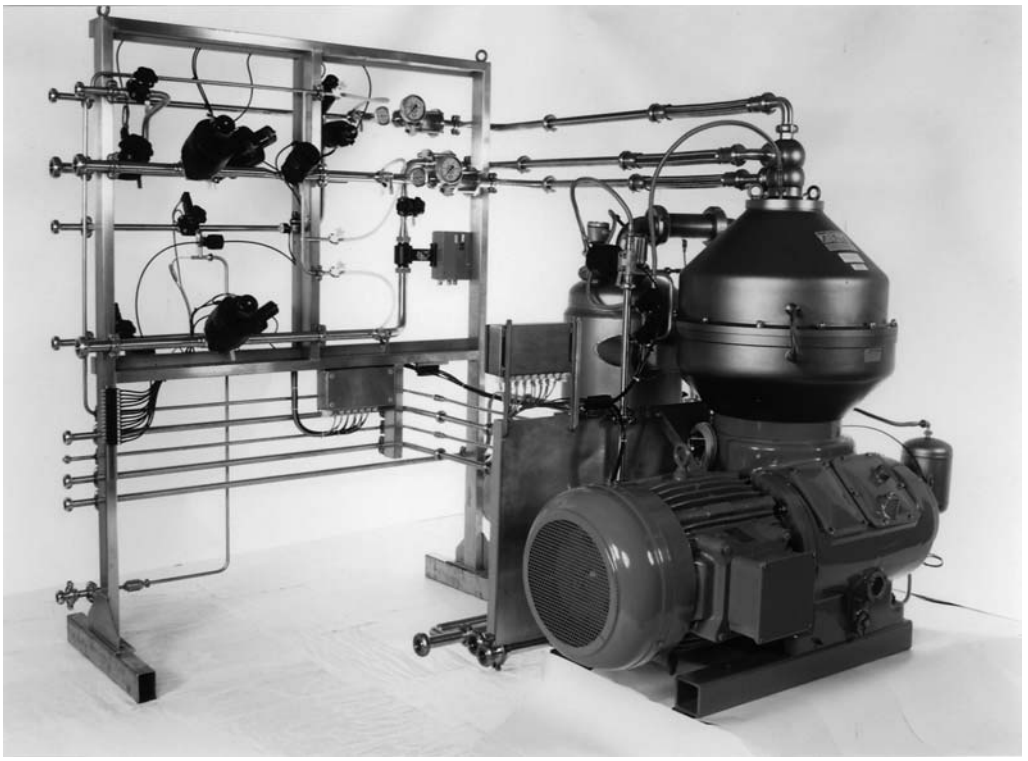


Figure 4. Nozzle machine with pressurized solids discharge for yeast and bacteria separation. Capacity on *E. coli* about 3,000 L/h. Type BTUX 510. The height of the separator is 1.8 m. Sterilizable and contained installation. *Source*: Courtesy of Alfa-Laval Separation AB.

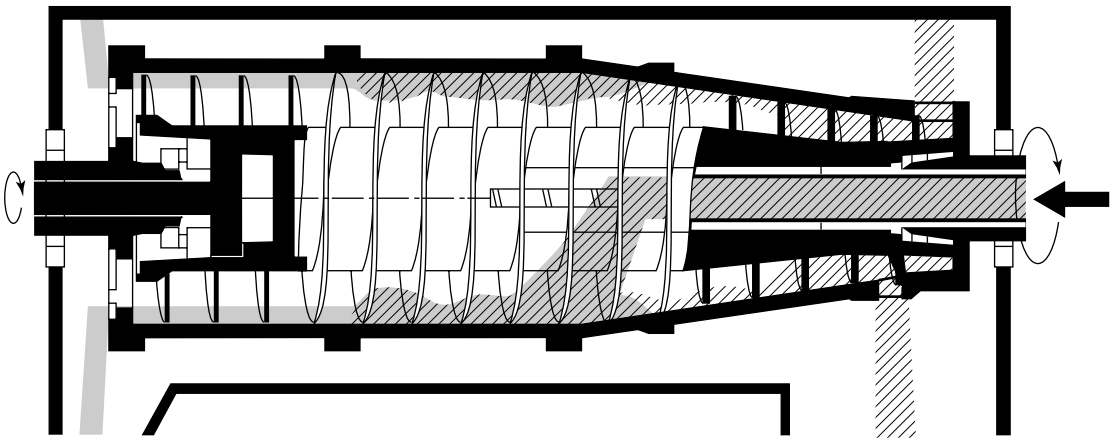


Figure 5. Decanter centrifuge, vertical cut.

centric cylinders. Having a relatively low separation capability, this type of machine is used for recovery of valuable solids, when the solids volume of the tubular bowl is too small. Some designs include cooling channels in the bowl wall, a cooling jacket around the bowl, and a cooled centripetal pump. The maximum bowl diameter is about 500 mm.

New Bowl Designs

Centritech® Machines. In the Centritech® machine (Fig. 8), a disposable separation insert in plastic is placed in a rotor. The feed liquid enters at the top of the insert, and the main liquid outflow is at the opposite end of the top of the insert. The heavy liquid phase (cell concentrate) is withdrawn intermittently through the bottom outlet. Through the principle of the inverted comma no seals between rotating and nonrotating parts are needed. The G-number is maximum 100. The mechanics are further described in Apelman et al. (9). Its only described application is separation of mammalian cells from culture fluid. Two sizes of the machine are available. The diameter of the insert is about 300 mm in the larger unit.

Inverted Chamber Bowl with Intermittently Operated Scraper. The inverted chamber bowl (Fig. 9) is a relatively new type of machine, sold under the name Powerfuge™ (10). Its bowl has its opening downward, where the centrate leaves under gravity and is collected in a chute. The feed is from the top. When the bowl is full of solids, it is decelerated. At low speed, the asymmetrically positioned scraper is engaged and removes the solids. Four sizes are available, with bowl diameters from 150 to 600 mm. They are designed for CIP and sterilization-in-place (SIP). The smallest unit can develop 20,000 G.

Summary

In Table 1, the characteristics of various separator types are found. It has to be remembered that the hydraulic capacities given by centrifuge manufacturers are the maximum possible feed flow rates that often are far above the flow rates that will give satisfactory separation results for difficult separations, such as bacteria and protein separations.

FLUID AND PARTICLE DYNAMICS

Introduction

A successful scale-up method should be based on a mathematical model that accurately describes the separation process, the hydrodynamics and particle dynamics inside the bowl. The phenomena going on inside the centrifuge bowl are, however, so complicated that no mathematical model is able to predict in absolute terms the separation efficiency as a function of machine and process media parameters. Axelsson (11) reviewed of some of the flow phenomena, which are briefly discussed in this section, along with examples of more recent findings.

Centrifuges with Conical Disks

Fluid and Particle Dynamics in the Disk Stack. The disks in the disk bowl split the flow in thin layers. The Reynold number for the flow between the disks is usually low, and the flow is laminar, resulting in an easy particle settling. The critical Reynolds number for transition to turbulent flow increases with decreasing disk spacing and increasing bowl speed (2), which is beneficial for the separation efficiency because the settling distance at a small disk spacing will be short, and the high bowl speed gives a large settling velocity. Also, a small disk spacing keeps the liquid rotating at or near the bowl speed, thus reducing the pressure drop in the disc stack.

In Figure 10 is a vertical cut of two disks in the disk stack, where the flow is driven toward the center by a radial pressure gradient. The character of the velocity profile between the disks depends on a dimensionless number, λ, which is defined as

$$\lambda = t \cdot \sqrt{\frac{\omega \cdot \sin \alpha}{\nu}} \qquad (5)$$

When λ is small, the velocity profile is parabolic, but for λ values above 5, normally encountered in industrial centrifuges, the radial velocity component splits up in two thin layers, Ekman layers (Fig. 10). In these layers, all the net transport of liquid toward the center takes place. Because separated particles (heavy phase) are collected at the underside of the disk, they are subjected to a strong shear force from the Ekman layer that drags the particle toward the liquid outlet. The velocity in these layers increases with increasing λ, such as with increasing bowl speed. Therefore, a higher bowl speed cannot always be fully used because of the counteracting shear force.

In between the Ekman layers is the geostrophic layer. In Figure 11 its streamlines have been visualized by dye

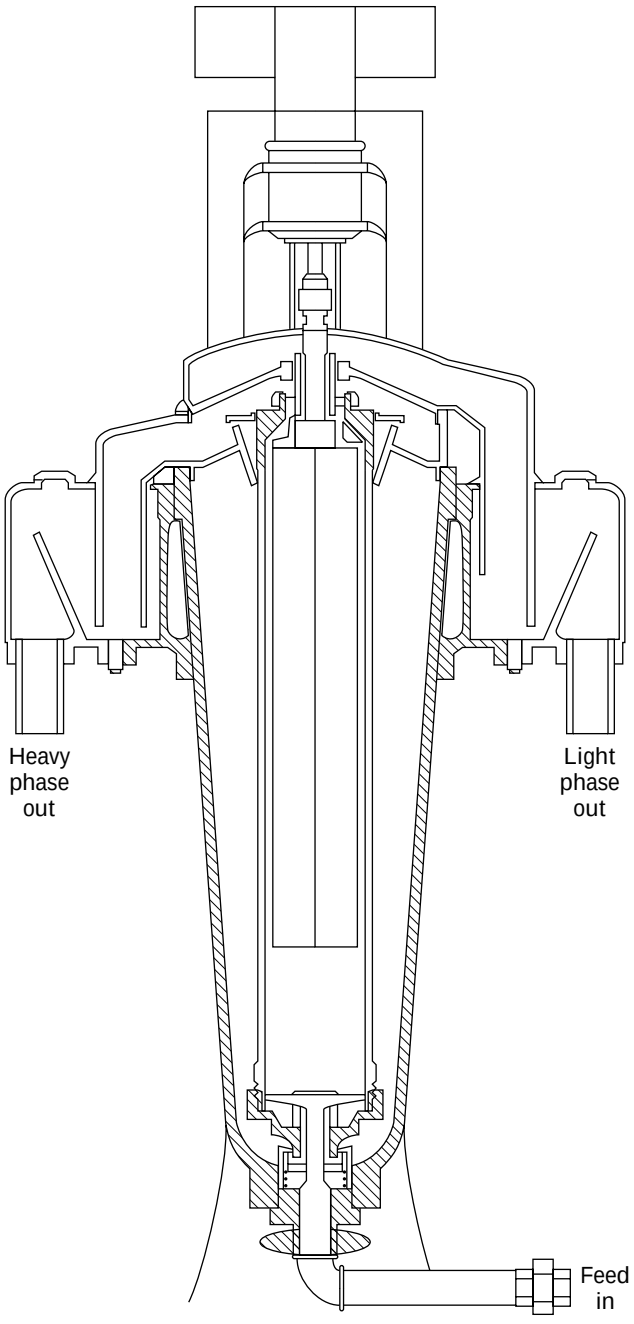


Figure 6. Tubular bowl centrifuge, vertical cut.

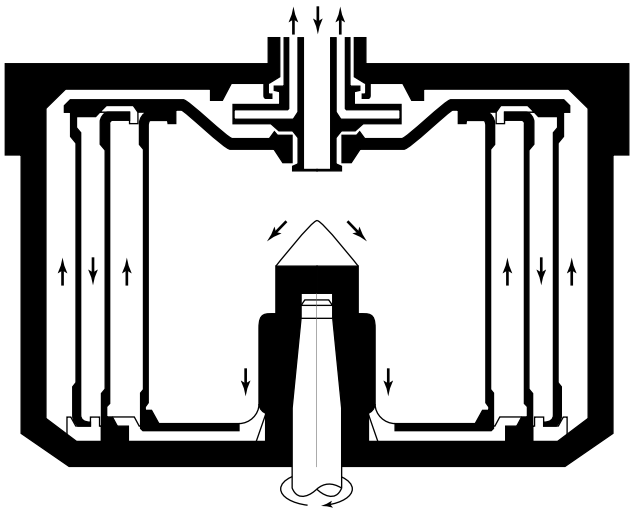


Figure 7. Multichamber bowl, vertical cut.

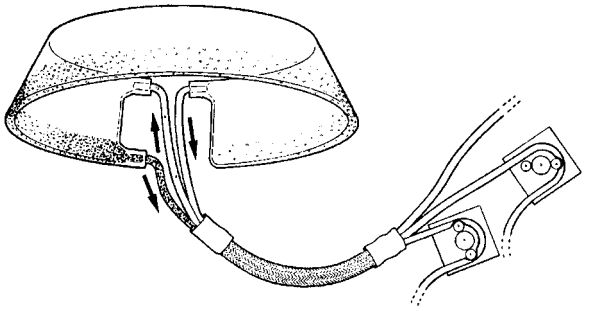


Figure 8. Centritech® Cell separation insert. *Source*: Courtesy of Sorvall (U.K.) Ltd.

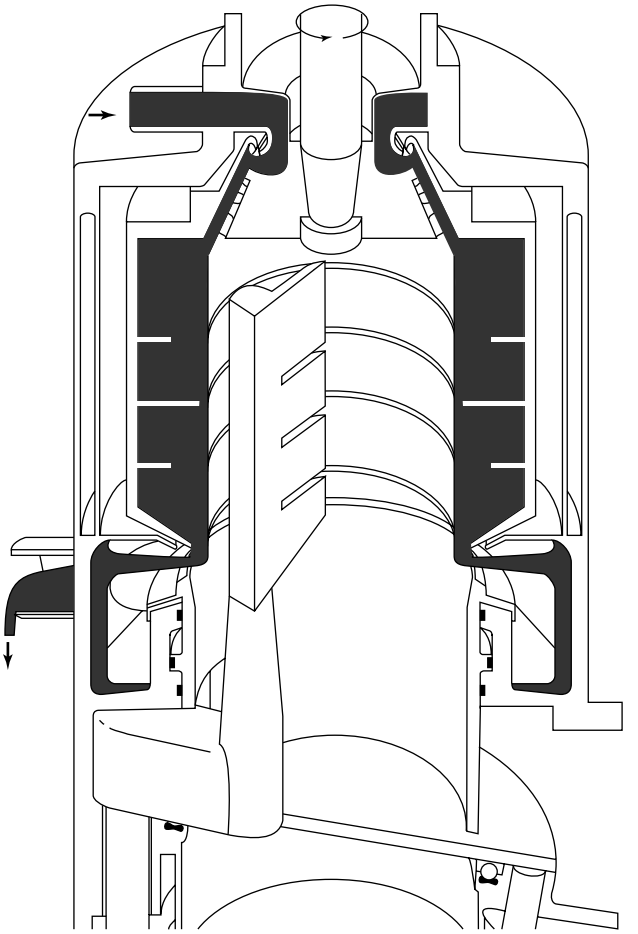


Figure 9. Vertical cut of a Powerfuge™ one-chamber bowl with scraper. Feed mode. *Source*: Courtesy of Carr Separations, Inc.

Table 1. Characteristics of Separator Types

Type	Mode of solids discharge	Feed flow rate (L/h)	Feed solids content (% by vol.)	Solids flow rate (L/h)	Σ value (m^2)	*G*-force developed	Consistency of solids
Disk solid bowl	Manual	20–100,000	<1	0	1,000–300,000	<10,000	Firm paste
Solids ejecting, radial	Intermittent	20–100,000	<25	<3,000	1,000–170,000	<14,500	Thick flowing slurry
Solids ejecting, axial	Intermittent	1,000–150,000	<15	<1,000	110,000–220,000	<15,000	Thick flowing slurry
Nozzle, pressurized discharge[a]	Continuous	1,000–180,000	4–30	>150 <40,000	69,000–180,000	<15,000	Thick flowing slurry
Peripheral nozzle	Continuous	300–250,000	2–30	>3000 <60,000	35,000–180,000	<11,000	Thin slurry
Decanter centrifuge	Continuous with scroll	20–120,000	5–50	<~50,000	400–25,000	<10,000	Thick slurry to semisolid
Tubular bowl	Manual	20–7,000	<1	0	1400–4500	<31,000	Firm paste
Multichamber bowl	Manual	100–20,000	<5	0		<9,000	Firm paste
Centritech® machines[b]	Intermittent	5–100	<1	<15	Not applicable	<100	Very thin slurry
Inverted solid bowl	Intermittent with scraper	10–6,000	1–30	<~1,000	800–~20,000	<20,000	Firm paste

[a]Only for slurry of single cells (see text).
[b]Only for mammalian cells.

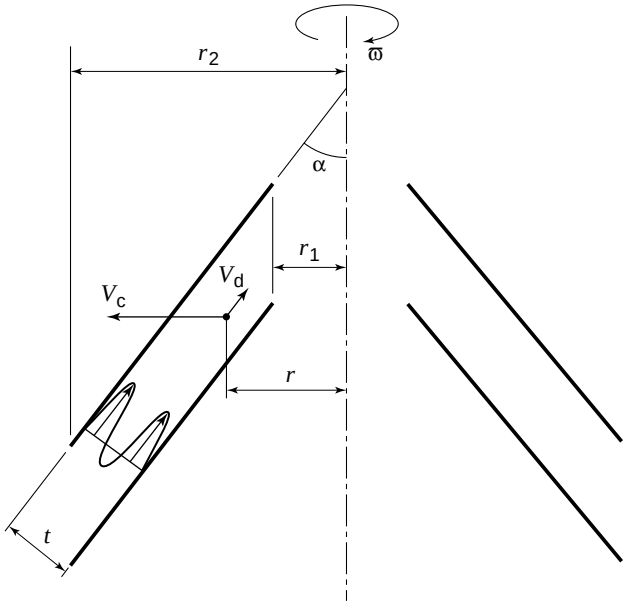
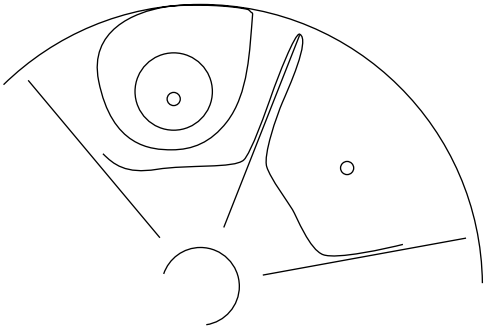


Figure 10. Flow profile between conical disks at high λ value. Particle separation.

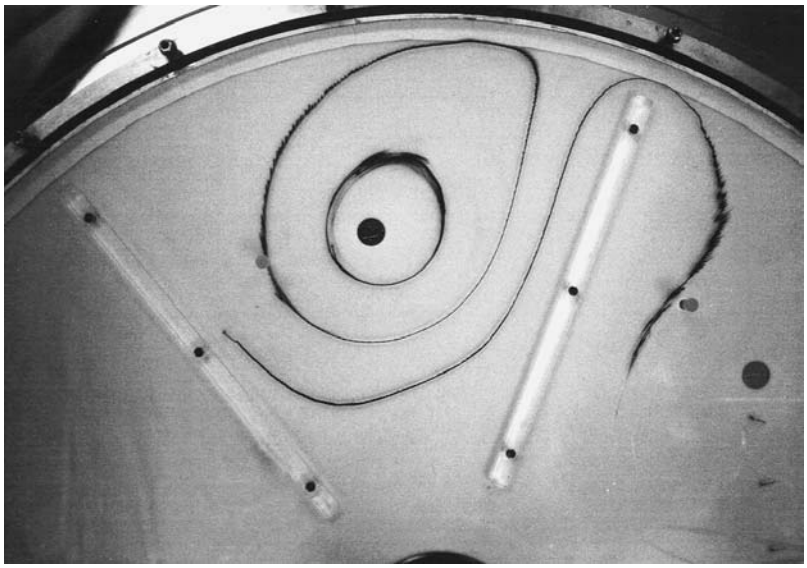
injection, and they can be compared with the calculated streamlines in a gap between two disks with radial spacers and with a distribution hole where the liquid enters the gap (12). Particle separation has also been studied in the same geometry (13). It has been found that the drag force on the particles collected at the underside of the disk depends on the position on the disk and that some particles are easily separated, whereas others are forced toward the center. The drag force can be computed (14), and the calculations of separation efficiency including the drag force verify these findings.

There is also an accumulation of particles at the radial spacer, along which the solids slide toward the periphery (13). At the end of the spacer, there is a wake that forces the particles back toward the center until the bulk weight of the solids becomes large enough to cause an intermittent discharge of particles into the solids space outside the perimeter of the disk.

Inlet. All bowls in the machines in Figures 1 to 9, except the hermetic machine (Fig. 1c) and the Centritech® machine (Fig. 8), have an open chamber in the center of the bowl, where the nonrotating feed liquid is received. In most inlets, the liquid forms a cylindrical interface with the gas (usually air) in the receiving chamber. The nonrotating in-



(a)



(b)

Figure 11. Comparison of calculated (**a**) and actual (**b**) flow patterns in a disk bowl.

coming liquid hitting the rotating interface results in a turbulence that entrains into the feed and splits particles that are susceptible to shear; the further out from the center the interface, the larger the turbulence. If it is not practical to use a hermetic inlet, the interface in the inlet must have as little an inner radius as possible. Also, the unavoidable energy dissipation must be distributed over a large volume, thereby diminishing the shear force. One design that allows smoother inlet conditions and a reduced air entrainment is the disk inlet (15).

A mathematical model for feed zone breakage of shear-sensitive particles has been verified for a conventional pilot-scale solids-ejecting separator (16). The extent of breakage was found to be independent of flow rate. This agrees with the finding that temperature increase in the separated liquid after the separator is also independent of flow rate.

Outlets. The liquid streams leaving the separator have picked up energy in the form of kinetic energy and heat caused by inner friction. When the liquid is decelerated outside the separator, the kinetic energy is converted to heat. The temperature increase is proportional to the square of the bowl speed and to the square of the radius at which the liquid leaves the separator. It is therefore not surprising that the discharge in a peripheral nozzle or a solids-ejecting machine causes breakage of cells because of the high kinetic energy of the fluid. A typical degree of bacterial cell breakage is 30%. The consequence of this is that if the particles to be separated are susceptible to shear and cell integrity is important, the cell concentrate should leave the centrifuge as near to the center as possible, such as in a machine according to Fig. 1c or e. It has been reported (17) that in such machines the breakage of *Escherichia coli* cells is low.

The outlet from the centrifuge is often of the paring disk type. For a good function, this outlet must be designed to avoid air entrainment and large shear forces. This is especially important for chemically sensitive substances such as proteins. A hermetic design (Fig. 1c) can be used in some applications.

Decanter Centrifuges

The objective of a decanter is usually to concentrate the solids present in concentrated suspensions. The hydrodynamics of a decanter are complicated. The conveyor adds turbulence and forces the liquid into a helical path (2). Below the top layer, vortices are formed. The transport of solids with low inner friction will influence the sedimentation process. The turbulence at the inlet zone, with possible resuspension of separated solids, and the lower solids concentration toward the liquid outlet contribute to the fact that a slender decanter will separate better than a short one with the same theoretical size (18).

THE THEORETICAL SIZE OF A CENTRIFUGAL SEPARATOR

Introduction

From the descriptions in the previous section it is evident that mathematical models describing the flow phenomena in centrifuges are very difficult to build. However, the attempts made are successful when machines of different sizes but with identical geometry are compared, for example, for scale-up purposes (see "Pilot Machine Testing").

The Σ Theory

The Generalized Σ Formula. The most-used quantity to characterize centrifuges, the Σ concept, was presented by Ambler (19,20). In its derivation he considered particles with a critical diameter (d_c) that were separated to 50%. Today, however, the most-used definition of the critical particle is the particle that is separated to 100%. This does not influence the formulas for the Σ value, but the value for the feed flow rate (Q) will be halved. Σ is defined as

$$Q_{\text{theor}} = V_g \cdot \Sigma \tag{6}$$

V_g is the Stokes settling velocity (equation 1). Σ has the general expression

$$\Sigma = \frac{V \cdot \omega^2 \cdot r_e}{g \cdot s_e} \tag{7}$$

The equation for the critical diameter becomes

$$d_c = \left(\frac{18 \cdot \eta \cdot Q_{\text{theor}}}{(\rho_p - \rho_f)} \cdot \frac{s_e}{V \cdot \omega^2 \cdot r_e}\right)^{1/2} = \left(\frac{18 \cdot \eta \cdot Q_{\text{theor}}}{\Sigma \cdot (\rho_p - \rho_f) \cdot g}\right)^{1/2} \tag{8}$$

The derivation in equation 8 is based on the following assumptions:

- Viscous drag is determining the particle movement.
- The flow in disk bowls between the disks is laminar and symmetrical.
- The liquid rotates at the same speed as the bowl.
- The particle concentration is low (no hindered settling).
- The particle always moves at its final settling velocity.
- This settling velocity (V_c) is proportional to the g-force.

Ambler uses the critical particle in his formula; therefore, his analysis is a special case of Svarovsky's grade efficiency function (21), where for each particle size the ratio of sedimented mass to the mass in the feed material is determined and integrated over all particle size to get a more realistic measure of the separation efficiency.

Σ for Disk Bowl Centrifuges. The flow system in between two conical disks is shown in Figure 10. The particle has a radial velocity component because of the centrifugal force (V_c), which increases with increasing r. It has also a drag-dependent velocity component (V_d), which can have any size or direction depending on the flow situation round the particle. In the Σ theory, it is assumed that V_d depends on

uniform radial plug flow parallel to the disks. The derivation of the equation for Σ is found in the paper by Ambler (20), and the result is shown in Table 2.

Summary. In Table 2, the resulting expressions of the Σ value for four different bowl geometries in continuous centrifuges are shown. These have been obtained by inserting actual values of V, r_e, and s_e for the various geometries.

Separation Efficiency

The assumptions in the Σ theory are not fulfilled in reality, as mentioned previously. Therefore, equation 6 should include an efficiency factor, so that in practice the flow rate is reduced at the requested separation performance. The efficiency factors have been estimated to have values according to Table 3 (11). This table was compiled in 1984 from data in literature. By careful measurement of the density difference and particle dimensions, it has been shown that in one case an efficiency factor of maximum 16% was found to fit the experimental data in a pilot-size disk bowl machine (16). However, this probably also included effects of hindered settling (see "Hindered Settling").

In the absence of hindered settling, the equation describing the relationship between the feed flow rate, the Σ value, and the settling rate, will thus be

$$Q_{act} = \mu \cdot V_g \cdot \Sigma \tag{9}$$

The KQ Formula for Disk Bowl Centrifuges

In 1961 a semi-empirical equation for the separation area was published (22). It is based on the finding that separation data fitted the relation

$$\frac{V_c}{V_g} = \left(\frac{\omega^2 \cdot r}{g}\right)^k \tag{10}$$

with $k = 0.75$ better than the classical analysis (equations 2, 3, and 6), where $k = 1$. This fact is a consequence of the increasing shear forces in the Ekman layers (see "Fluid and Particle Dynamics in the Disk Stack"). For practical purposes, the following relation is used:

$$KQ = 280 \cdot \left(\frac{n}{1000}\right)^{1.5} \cdot N \cdot \cot \alpha (r_2^{2.75} - r_1^{2.75}) \tag{11}$$

where n = bowl speed in rpm and r_2 and r_1 are the radii of the disk in centimeters. This form of the KQ formula is dimensionally incorrect, but it is mathematically possible to make it correct, expressing the KQ value as an entity with the dimension area. In practice, the difference between the KQ and the Σ for scale-up purposes is negligible, compared to other inaccuracies.

Hindered Settling

One of the few experimental investigations into hindered settling of cells (23) shows that complex phenomena occur at high concentrations. A model with several experimental constants was developed. It showed that the settling rates drop very fast with increasing solids content, by a factor of about 2 at 20% by volume, and 15 to 20 at 50% by volume. The classical model used in the work by Clarkson et al. (16) gave similar results. In the tests by Aiba et al. (23), it was also found that cells may settle flocculated at higher concentrations. Hindered settling is one further example of the difficulty to calculate theoretically the performance of centrifuges.

SELECTION OF TYPE AND SIZE OF CENTRIFUGE

Introduction

The choice of type and size of centrifuges is a combined theoretical and practical exercise. In most cases the final choice for a new centrifuge application cannot be made without a test of the type of centrifuge preliminary selected, at between 25 and 100% of the final plant flow rate. There are at least two reasons for this. One is the elimination of the technical risk. The other is most common in

Table 2. Σ Equations for Various Separator Types

Machine type	Equation for Σ	Symbols used	References
Disk bowl	$\frac{\pi \cdot \omega^2}{g} \cdot \frac{2}{3} \cdot N \cdot (r_2^3 - r_1^3) \cdot \cot \alpha$	r_2 = max radius of disk r_1 = min. radius of disk N = number of disk α = half cone angle of disk	19,20
Decanter	$\frac{\pi \cdot \omega^2}{g}\left[L_1\left(\frac{3}{2} \cdot r_2^2 + \frac{1}{2} \cdot r_1^2\right) + L_2\left(\frac{r_2^2 + 3r_2r_1 + 4r_1^2}{4}\right)\right]$	L_1 = length of cylindrical part L_2 = length of conical part r_1 = inner radius of liquid r_2 = inner radius of bowl	20,21
Tubular bowl	$\frac{\pi \cdot \omega^2}{g} L \frac{r_2^2 - r_1^2}{\ln\left(\frac{2r_2^2}{r_2^2 + r_1^2}\right)}$	r_1 = inner radius of liquid r_2 = inner radius of bowl L = inner length of bowl	20
Multichamber bowl	$\frac{\pi \cdot \omega^2}{g} \frac{L}{3} \sum_{i=0}^{i=n} \frac{r_{2i+1}^3 - r_{2i+2}^3}{r_{2i+1} - r_{2i+2}}$	Indexes with even numbers = inner radius of chamber Indexes with odd numbers = outer radius of chamber $n + 1$ = number of chambers L = height of chambers	2

Table 3. Efficiency Factors μ for Various Separator Types (%)

Disk bowl machines	45–73
Decanter centrifuges	54–67
Tubular bowls	90–98

the pharmaceutical industry. During the preapproval inspection, those in the inspecting and license-giving institutions are concerned with the comparability between batches of product from the development phase and from the production phase (24). Because the validation procedures are costly and time consuming, it is very important to select at least the right type of centrifuge already in pilot scale. It is important that alternative process and plant designs are discussed before the centrifuge type and size are selected (Table 4).

Main Process Parameters

It is important to define at least the characteristics of the duty listed in Table 5. The *feed flow rate* is often determined by the available or necessary time to process one batch. The *separability* can be estimated on a bench scale by test tube centrifugation (spin test) (see next section). The *flow of solids* is the feed flow rate multiplied by the percent by volume of suspended solids. This solids flow rather than the suspended solids content alone influences the choice of machine type. Different machine types differ in their ability to produce high *concentration of the separated solids* (Table 1). The *required degree of separation* has a strong influence on the possible flow rate in a given machine. To increase separation efficiency of *E. coli* from 90 to 99%, the flow rate must be decreased by 30 to 35% (25). To reach 99.9%, a further reduction of about 30% is necessary. Many cell suspensions, including yeast, have the same narrow particle size distribution and the numbers are the same. With the wider distribution of cell debris, for example, the flow rate reduction percentages are much higher. *Solids rheology*, which is the reaction of the solids to pressure and shear forces, is very important for the choice of machine type. Pseudoplastic (shear thinning) solids (such as unflocculated yeast, rod-shaped bacteria, and cocci as well as some protein precipitates) will flow out of the machines seen in Figure 1c and e at 80% by volume of suspended solids, which often has positive consequences (reduced cell breakage). Shear thinning solids may have an angle of repose of close to 0 degrees (see "Important Properties of the Process Streams"), but this is not always the case, such as with flocculated or mycelium-forming bacteria. On the other hand, dilatant solids (starch or calcium carbonate) will not be discharged in concentrated form without manual handling from any machines other than those seen in Figures 1f and 5. The rheology of inclusion body concentrates is not much studied, but they are preferably discharged from the machine seen in Figure 1b. Because of the content of a valuable dissolved substance in the entrained liquid, *washing of solids* may be necessary.

Another reason for washing solids is, of course, to get rid of impurities. This is usually done by resuspension and further separation steps.

Table 4. Process and Plant Considerations

Possibility to increase particle size
Possibility to reduce viscosity
Chemical stability of product
Containment requirements
Aseptic requirements
Sterilization requirements
Mode of operation (hours/day, days/year)
Integration with unit operations upstream and downstream
Scale of pilot- and full-scale operation
Handling of solids from the centrifuge

Table 5. Main Process Parameters

Feed flow rate
Flow of solids
Required concentration of separated solids
Solids washing requirements
Separability (V_g)
Required degree of separation
Solids rheology

Test Tube Centrifugation (Spin Test)

Determination of Separability. The most important task for a spin test is the determination of separability, that is, the settling velocity. In practice, it is difficult to calculate, because often the information needed in Stokes' law is not complete. The spin test is done by spinning a sample of the feedstream for different periods of time or different G-numbers or both. For each of these spins, the supernatant is analyzed with regard to the important process parameter, for example, residual solids, turbidity and cell count, or water content in the case of light solvent. Thereafter, the time of the spin with the acceptable supernatant quality is recorded, and the necessary Σ is calculated according to equation 16.

The Spin Test as an Analytical Tool. The spin test is also a valuable tool for measuring the phase-to-volume ratios. The solids content in percent by volume from the spin test will be a very important parameter for the next step, which is to make a crude mass balance around the centrifuge. The spin test performed under different time and *g*-force conditions gives also an indication of solids compressibility and rheology. It is also valuable to prepare a sample of the solids phase after spinning to get a good indication of its dry solids content after the centrifugation. In cases where the fermentation substrate contains suspended solids, residues of these may be present after fermentation. A spin test will make it possible to identify the nature and amount of these, which is very important because they may obstruct the solids phase flow in nozzle and solids-ejecting machines.

Theory of the Test Tube Centrifuge. The swing-out test tube (bottle) centrifuge is shown schematically in Figure 12. In order for a particle to be separated out with 100%

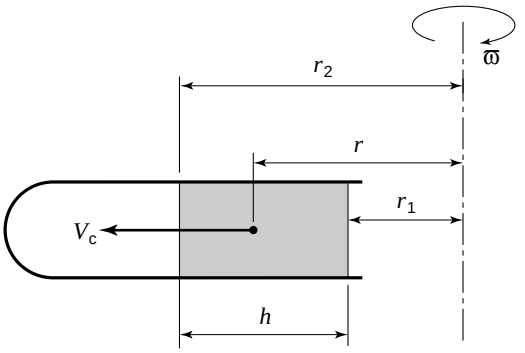


Figure 12. Separation of particles in a swing-out test tube centrifuge.

chance, it has to move from radius r_1 to radius r_2 in the spin time (T). The shaded volume with height (h) is the volume taken out for sampling of the result. The following equations describe the particle movement.

$$V_c = \frac{dr}{dt} \tag{12}$$

$$V_c = V_g \cdot \frac{r \cdot \omega^2}{g} \tag{13}$$

Insertion of equation 10 in equation 9 and rearranging and integration gives

$$\int_0^T dt = \frac{g}{V_g \cdot \omega^2} \int_{r1}^{r2} \frac{dr}{r} \tag{14}$$

$$T = \frac{g}{V_g \cdot \omega^2} \cdot {}^e\log \frac{r_2}{r_1} \tag{15}$$

In the test tube centrifuge, the efficiency factor μ in equation 9 is 100% because the settling process is considered ideal. Therefore, equation 6 is valid and

$$\Sigma = \frac{T \cdot Q_{theor} \cdot \omega^2}{g \cdot {}^e\log \frac{r_2}{r_1}} \tag{16}$$

and

$$\frac{Q_{theor}}{\Sigma} = \frac{g \cdot {}^e\log \frac{r_2}{r_1}}{T \cdot \omega^2} = V_g \tag{17}$$

For a test tube centrifuge with angle head, ${}^e\log(r_2/r_1)$ is exchanged for ${}^e\log(1 + D/r_1 \cdot \cos\beta)$.

Important Properties of the Process Streams

In Table 6, some solids phase data, such as *particle size* and *density difference*, have been collected. These data show that many important biological processes belong to the most demanding centrifuge applications regarding separability. The data given in Table 6 are very approximate. Cell size will vary with fermentation conditions. Cell debris size distribution is very wide and depends on the method of cell breakage. The density difference between cell and medium depends on the osmotic pressure and of the possible content of inclusion bodies. Apparent density and the density difference of inclusion bodies vary with size, so that larger ones have lower density (28).

As many as possible of the items in Table 7 should be defined before the selection of a centrifuge is made. Some of the items need further comments. Apart from Stokes' law, the *density* is important for the stress in the bowl material. *Low pH* and *high temperature* may, together with even a moderate chloride concentration, lead to difficulties with *corrosion*. In the biological process industries, steels 316 and 316L are often preferred (38). These steels have, however, such a low tensile strength that they are less suitable for high-speed centrifuges. Instead, duplex steel has been developed, it has high tensile strength coupled with a good corrosion resistance (39). The corrosive properties of process materials can be very serious, but it is in many cases possible to reduce the corrosion rate by intermittent washing with caustic solution every 4 to 5 hours of running to reestablish a passivating layer on the metal surfaces. *Erosion* may be a problem with silt-containing process water or other raw materials such as molasses. This may necessitate the use of hydrocyclones to remove these particles. The *angle of repose* (see "Main Process Parameters") is the angle a cone solid will form if allowed to settle, for example, between the nozzles in a nozzle bowl. If it is 0, or close to 0, degrees, the flow properties are extremely good. The *shear sensitivity* was mentioned in "Centrifuges with Conical Disks". The most shear-sensitive cells are mammalian cells, for which the shear forces must be minimized when the cells are used in perfusion systems.

In the Centritech® machines, the RCF is limited to 300, more than one order of magnitude lower than for other centrifuges. The cells may be so vulnerable that the residence time in the centrifugal force field must be minimized (40). Some protein precipitates are also sensitive to shear (16), including those formed in the maturation tanks in breweries. Many bacteria are sensitive, especially starter culture bacteria, where the activity of the cultures depends on intact cells. Yeast cells are the most robust cells, but they are also sheared to some extent at discharge in radial solids-ejecting centrifuges.

Improvement of Separability

According to Stokes' law, the settling velocity improves when the density difference and the particle density increase and when the viscosity decreases. Bowden (41) lists physical, chemical, and biological techniques for broth conditioning. The most-used techniques are coagulation and flocculation, but other methods that induce cell lysis can improve aggregation of cells, such as aging, pH, and change and addition of enzymes or other chemicals. In some cases it is possible to flocculate submicron suspensions, so that a decanter centrifuge can recover the precipitate (cell debris) with good efficiency and a dry cake (34).

Pilot Machine Testing

After maximizing the separability of the feedstream, the preliminary Q_{theor}/Σ is determined by spin tests. The ma-

Table 6. Size, Density Difference, and Relative Feed Flow Rate for Some Solid–Liquid Systems

Type of Solids	Particle size (μm)	Density difference (kg/m^3)	Relative throughput[a] (relative to baker's yeast)
Cell debris	0.2–0.5 (mean, bacteria)	—	0.2
	2 (median, yeast)	—	1
Streptococii	1 × 1	—	1–2
	1 (chain)	—	3–6
Clostridia	1 to 3	—	5
E. coli	1 × 2	70–85	2–6
Actinomyces spp.	10 × 10–143 × 143	1 (for large particles)	1–7
Inclusion bodies	0.7 × 0.7–1.3 × 1.3	29–154	5–8
Corynebacterium	—	—	6
Bacillus spp.	1 × 3	—	6–10
Aspergillus spp.	—	—	15
Mammalian cells	20 to 40	70	20
Flocculated material (e.g., debris)	—	—	20
Chlorella	4 × 4	—	30
Saccharomyces (alcohol yeast)	4 × 6	90	50–75
Saccharomyces (baker's yeast)	6 × 8	90	100

Source: Data from references 11, 23, 26–36.
[a]Proportional to Q_{act}/Σ.

Table 7. Important Process Liquid Properties

Chemical composition of liquid and solid phases
Densities of liquid and solid phases
Viscosities of liquid and solid phases
Particle size distribution
pH
Temperature
Vapor pressure
Solubility of one phase in the other
Corrosivity
Erosivity
Shear sensitivity
Angle of repose
Toxicity
Explosion hazard

Source: Adapted from reference 37.

chine type and size are tentatively chosen based on the preliminary mass balance and on the requested Q_{act}. With the separation efficiency μ, Q_{act}/Σ can be calculated from the equation

$$\frac{Q_{act}}{\Sigma} = \mu \cdot V_g \tag{18}$$

It is often advisable to make a test with the chosen type of machine. The purpose of such a test is to obtain a more reliable Q_{act}/Σ value, but also to assess solids flow properties, emulsification, and particle breakup tendencies and clogging risks.

Scale-Up

For all types of centrifugal separators, the Q_{act}/Σ value can be used for scale-up of the sedimentation performance if some general considerations are taken into account:

- The type of centrifuge is similar.
- The geometry is similar; for decanters, feed inlet positions, ratio of length to diameter, or beach angle, and for disk machines, cone angle, disk spacing, and number of feed holes.
- The bowl speed gives similar linear velocity of, for example, wings when treating shear-sensitive process feeds.
- The λ value (equation 5) is almost the same in the machines.
- The larger machine has no hydraulic limitations in relation to the calculated flows. The limitations could be inlet and outlet pressure drops, paring disk pumping capacities (upper and lower limit), and the upper and lower limit of solids flow.
- There are no mechanical limitations, such as power transmission and gear-box torque in decanters.

The equation to use in scale-up is

$$\left(\frac{Q_{act}}{\Sigma}\right)_1 = \left(\frac{Q_{act}}{\Sigma}\right)_2 (= \mu \cdot V_g) \tag{19}$$

where indexes 1 and 2 denote the two machines. Direct calculations of V_g and μ are thus avoided.

DESCRIPTION OF SOME APPLICATIONS

Introduction

The most important centrifuge application in biotechnology is cell harvesting, where either the liquid phase or the solid phase contains (or is) the product. Downstream of the harvesting stage the number of applications for centrifuges is growing. The operations performed are many: classification, dewatering, liquid–liquid extraction, and precipi-

tate recovery. The applications discussed in this section are examples only; there is an infinite variation in the downstream process design.

Cell Harvesting

Introduction. As can be seen from Table 5, the throughput capacity of a given centrifuge (i.e., Q_{act}/Σ) may vary by a factor of 100 depending on the microorganism used.

Yeasts. The second application of disk bowl centrifuges developed after milk skimming is concentration of *baker's yeast*. Today, nozzle machines with pressurized discharge (Fig. 1) are very common in the industry. The flow rates are the highest in the biotechnology industry, up to 150 m^3/h of fermented broth. Cell concentrations in the fermenters have increased over the years and have now reached above 25% by volume. The baker's yeast is washed, often twice and countercurrently, to increase shelf life. The pressurized discharge of the solids concentrate makes it easy to use an inline mixer to mix the wash water with the cells. The Q_{act}/Σ is about 2×10^{-7} m/s for a 99.9% reduction of the cell count.

In *distilleries*, nozzle bowl centrifuges have been used for a long time. The centrifuge has been the prerequisite for the development of new alcohol fermentation systems (11). In the continuous fermentation from spent sulfite liquor, centrifuges are used for recycling of the yeast. This is also the case in the batchwise Melle-Boinot process, where the recovered yeast inoculates the next batch. In the continuous Biostil process, in which the distillery slop is returned to the fermenter, the centrifuge recovers the yeast from the feed to the distillation column and returns it to the fermenter.

In *breweries*, centrifuges are used in several process steps. In the separation step after fermentation, *green beer separation*, the cell count can be as high as 80×10^6 cells per mL (2% by volume), but during the emptying of the fermentation tank the solids decrease. Radial solids-ejecting machines are used, equipped with self-triggering and special outlet devices to prevent air contact. At a cell reduction of 99% and 20×10^6 cells/mL, the Q_{act}/Σ is about 2×10^{-7} m/s, increasing somewhat with higher cell count. The flow rates are up to 75 m^3/h. In this step, the same machine type as for baker's yeast could be used if the solids flow to the machine could be controlled. New instruments (42–44) measuring the dielectric constant or the inductive permittivity may be used.

For the separation step after lagering and before final filtration, *preclarification*, the cell count is much smaller, and less solids handling capacity is needed. Therefore, the bowl of the solids-ejecting machine can have the shape of the bowl in Figure 1c. The hermetic inlet is an advantage because the separation of the shear-sensitive protein particles is helped by the gentle acceleration. The bowl has only one liquid outlet. The solids are ejected at the periphery. Because the emphasis is on cell reduction, the Q_{act}/Σ is about 5×10^{-8} m/s. Flow rates are up to 45 m^3/h.

The corresponding separation step is also present in winemaking. Because the same machine unit is used also in must separation before fermentation, radial solids-ejecting machines are used.

In rDNA yeast processes, the same type of machine as for baker's yeast can be used. Such machines are available for BL2-LS containment.

Bacteria. In modern biotechnology, solids-ejecting machines, nozzle machines with pressurized discharge, and decanters are used to harvest bacteria. For microorganisms with sizes about 1 μ, axial discharging machines (Fig. 1d) are used. If the solids concentration is so high that the solids discharge frequency has to be more than 1 per 60 s, a radial discharging machine (Fig. 1b) has to be used, but this is less cost effective. The vortex nozzle machine (4) (Fig. 1e) has the disadvantage of accepting only clear substrate media. Residues of substrate may clog the concentrate tubes. If a clear substrate is used, this type of machine is as cost effective as the axial discharging machine but without its solids-handling limitation. A further possibility for clear media is the hermetic machine in Figure 1c, where the cell concentrate leaves the machine continuously and under pressure. The intermittent discharge mechanism is used for CIP.

A typical Q_{act}/Σ value for 99% recovery of *E. coli* K12 in a disk stack machine is 10^{-8} m/s. But the variation between various types of *E. coli* can be very large indeed; between 30 and 300% of the above value has been encountered by the author.

Corynebacterium, somewhat larger cells used in amino acid production, are harvested in nozzle machines with pressurized discharge, often with washing of the bacteria (repulping) and a second separation to improve the yield. A typical Q_{act}/Σ is 5 times that of *E. coli*. *Actinomyces* mycelia can be harvested by radial discharging machines. As in nozzle machines, the cells are washed, often in two steps, to improve the yield.

There are cases where repulping and reseparation are not possible because the cells break up and may damage the rest of the process. To avoid losing yield, the cells must be recovered with maximum dryness. The obvious choice of machine is then the decanter. The low Σ value of this machine type makes it impossible to use it without flocculation of cells. Polymer addition gives a separation efficiency of 98 to 99% at a Q_{act}/Σ of 3×10^{-7} m/s for a spore-free *Bacillus subtilis* broth with a solids content of 12 to 15% by volume. The dry solids (DS) content of the concentrate is 27 to 28% by weight at 3,200 G. This can be compared to solids-ejecting or nozzle machines, which give around 15% by weight for bacteria. Longer residence time in the bowl or higher RCF gives still higher DS contents. At 10,000 G and similar residence time, a decanter can produce an *E. coli* concentrate of 32 to 36% by weight at a Q_{act}/Σ comparable to that in a disk bowl centrifuge.

The high DS capability of decanters is also used to concentrate solids from nozzle machines or from ultrafiltration units, working on fermenter broths in production of amino acids and antibiotics, respectively. The improvements are yield and less costly waste disposal (e.g., drying costs).

Single-cell protein production from hydrocarbons or their simple derivatives is today a not-so-profitable business, but the technology exists. Mayer and Woernle (45) presented a flow sheet for biomass downstream processing.

Bacteria are preconcentrated in large nozzle machines to about 6% by weight DS and then flocculated. This biomass is then dewatered to about 30% by weight DS in decanters and goes after granulating to drying. The liquid phase from the decanter is clarified in high-speed separators and sent to sterilization together with the clarified liquid from the preconcentration step. From sterilization it is recycled back to the fermenter.

Mammalian Cells. It is possible to use high-speed centrifuges in mammalian cell processes (46), in which the cells are discarded after removal. Usually, the inlet shear is not large enough to release DNA into the culture fluid. Care must be taken that shear and air contact in the centrate are avoided, because some protein substances may be destroyed. The centrifugation of intact cells for perfusion must be performed at very low pressure and shear forces. Special machines have been developed recently for this purpose (9,40,47,48).

Cell Debris

In those cases where the intracellular products do not form inclusion bodies, the cell debris often has to be removed after cell breakage. The separation is difficult because of the very fine particles produced (Table 6). Moreover, the viscosity may be high. Dilution of the suspension reduces the viscosity, but the Q_{act}/Σ may still be 5 to 10% of that of the whole cell. It is necessary to use high-speed solids-ejecting machines for this step. In cases where it is possible to flocculate the debris, the Q_{act}/Σ increases by a factor of at least 10 (17). It has been reported (34) that borax-flocculated yeast cell debris can be removed to about 80% in a pilot-size decanter centrifuge at a Q_{act}/Σ of 4×10^{-4} m/s.

Inclusion Body Recovery and Washing

In cases where the substance forms inclusion bodies, it is advantageous to remove cell debris and other cell constituents before the next processing step. The inclusion body has a higher settling velocity than the cell debris, so that a classifying centrifugation will recover the inclusion bodies while leaving the cell debris in the centrifugate. There is a strong interaction between, on the one hand the homogenization method (i.e., the number of passes) and the resulting particle size distribution and, on the other hand, the ease of washing the inclusion bodies (49). There exists an optimum number of passes. The classification split is not 100% sharp, making it necessary to repeat the separation a number of times. This is done either by resuspending the recovered inclusion bodies batchwise or by a continuous simultaneous feeding of the inclusion body concentrate and water into a well-mixed flowthrough feed tank, continuing until the required purity has been reached. Axial or preferably radial discharging solids-ejecting machines are applied at a Q_{act}/Σ of 1 to 2×10^{-8} m/s. To reduce the consumption of chemicals in the next stage, the final separation is often a dewatering step. This can be performed in a high-speed decanter at a Q_{act}/Σ of 2×10^{-6} m/s.

Extraction of Cell-Free Broth

Solvent Extraction. For 50 years, liquid–liquid extraction has been used to concentrate and recover penicillin from filtered broth. This was often performed in multistage centrifugal extractors. Today, the trend for filtered broth is to use a two-stage, countercurrent system with mixers in between disk bowl machines, which are of the three-phase solids-ejecting type. Many other antibiotics need only a one-stage extraction, for example, erythromycin and tetracycline.

Two-Phase Aqueous Systems. The polyethylene glycol (PEG)/dextrane or PEG/salt systems used in protein purification are characterized by a very small density difference between the phases, from 100 down to 14 kg/m^3. The viscosity of the polymer phase could be high, perhaps 100 mPa s, and a very low interfacial tension, leading to easy splitting of drops. Therefore, it is not surprising that the Q_{act}/Σ may be rather low, from 5×10^{-8} down to 1 to 2×10^{-9} m/s. The type of equipment used for the systems with low-density difference is radial solids-ejecting machines. Decanters may be used when the density difference is larger. It is important for optimum performance that the interface between the light and heavy phase in the bowl be in the right position. Equation 4 shows the relation between the three liquid levels in the bowl. The interface should usually be placed just outside of the disk stack in the disk bowl machines.

The two-phase aqueous systems can also be used for partitioning cell debris and other impurities from a dissolved protein product (50). Datar used a PEG 4,000/potassium phosphate system. No density difference data were given. The Q_{act}/Σ was 2×10^{-7} m/s.

Precipitation

Precipitation is often used to isolate and concentrate antibiotics (e.g., rifamycin, erythromycin, and tetracycline). The suitable machine is the radial solids-ejecting disk bowl type. The methods and parameters involved in protein precipitation have been reviewed (51).

Whole-Broth Extraction

Solvent Extraction. Whole-broth extraction of penicillin was the first development of this technique (52). The solids flow is so high that decanter centrifuges must be applied. Many of the constituents of the broth (e.g., macromolecules and the mycelium) have emulsion-stabilizing properties, so de-emulsifying agents must be added (53). After extraction, the rich solvent is separated to reduce the water content before further processing. The solvent-contaminated mycelium must be further treated before disposal.

Two-Phase Aqueous Systems. The two-phase aqueous systems can be applied for extraction also in the presence of cell debris. Hustedt et al. reported on three enzyme recovery processes that use cross-current extraction (54). The phase separation is made in small solids-ejecting or nozzle disk bowl machines. Provided that the density difference between phases is large enough, extraction decanter centrifuges can also be used. Among the commercial

uses of this technology on a large scale is the recovery of enzyme in a polymer–salt system, where the separation is performed in a three-phase radial solids-ejecting centrifuge.

INSTALLATION AND OPERATION

Plant Design

In modern biotechnology plants, the design philosophy must be documented as part of the validation procedure. These procedures are becoming common also in the food industry. Stringent guidelines for safe design of plants are being issued more frequently. In Europe, the draft standard for performance criteria for centrifuges in biotechnology deals with leak tightness, cleanability, and sterilizability. Centrifuge installations designed according to BL2-LS are seen in Figures 2 and 4.

Cleaning-in-Place

Automated CIP is today regarded as a quality guarantee. The validation of CIP procedures is one of the necessary protocols to obtain approval to manufacture bioproducts. An automated and logged CIP procedure is repeatable and traceable, less costly than manual cleaning, and safer for the plant personnel than manual cleaning is. Centrifuges and their peripherals are today designed with cleanability in mind. With use of modern programmable logic controllers (PLCs) for valve control, the CIP of a disk-stack centrifuge (or decanter) has been greatly simplified. A CIP cycle consists of several steps of rinsing with different cleaning liquids. The sequence could be cold water, warm (60–70 °C) 4% caustic, 3% nitric acid if necessary, hot water, and finally cold water. It is often necessary to use some type of detergent to improve the cleaning effect, mixed with the caustic or the water. In the radial solids-ejecting machines and the modern pressurized discharge nozzle machines, the discharge mechanism is activated especially during the initial stages.

Containment, Asepsis, and Explosion Protection

These three areas are related in the sense that the installation designer must define a space around the centrifuge, including tanks and piping with controlled interfaces with the environment. In the case of containment, the task is to protect process operators from harmful process substances. The National Institutes' of Health biosafety level is an attempt to classify the risks involved (55). The interface with the environment is sterile HEPA filters. This is a case of "no leaking out." The reverse case, "no leaking in," applies for asepsis and explosion protection. The barriers are formed by HEPA filters and liquid seal pots, respectively.

In both these cases, it is necessary to keep the centrifuge system under a slight, positive pressure relative to the environment, either by lowering the pressure in the room (asepsis) or increasing it in the centrifuge system. For explosion protection, this is done by filling the system with inert gas (e.g., nitrogen) and bubbling the inert gas through a water seal pot. Other measures taken with explosion hazards have been reviewed (56). Centrifuges in closed systems have to be modified in a number of ways: mechanical instead of labyrinth seals, cooling jacket around the bowl cover, and devices that can handle the pressure wave at discharge.

Sterilization

Sterilization with saturated steam at temperatures above 100 °C is by many considered to be the reliable way for avoiding contamination, and it is prescribed for BL2-LS. For the centrifuge manufacturer, this means that in addition to the modifications described in the previous section, the bowl cover must be designed as a pressure vessel. But the most complicated changes concern the installation, because the bowl internal and external surfaces, the cover, the service liquid lines, etc., must all be sterilized. To reach the desired temperature (above at least 121 °C), condensate must be drained away. Krook et al. (4) discuss the sterilizable system for a medium-sized machine with pressurized nozzle discharge in some detail. The sterilization efficiency for a radial solids-ejecting centrifuge of pilot scale was investigated (57). It was found that it was not possible to sterilize the whole system at once. By trying different combinations of open and closed valves, regimes were developed for heating different sections, and these were combined to sterilize the whole system. In the end, sterilization was validated with different chemical and biological indicators.

CENTRIFUGATION VERSUS MICROFILTRATION

After the development of sterilizable microfilters, there is available an alternative to centrifugation in separating solids and liquids. A general rule is that with increasing particle size and increasing scale of operation, centrifugation becomes more interesting than microfiltration and vice versa. It is conceivable that *E. coli* on a 200-L fermenter scale is a microfiltration application. Many factors must be considered in the choice. What is becoming more apparent is that the choice done in pilot scale becomes the most difficult one, because the scale-up and validation procedures make it easier to keep same technology on the larger scale, even if it is not the optimal choice. In many cases, a combination of a filter and a centrifuge is the best answer.

NOMENCLATURE

d_p	Particle diameter, m
D	Inner diameter of test tube, m
$D_{h,l}$	Diameter of heavy and light liquid outlet, m
D_i	Diameter of interface, m
g	Equal to 9.81 m/s^2
h	Distance between inner and outer radius in test tube, m
k	Constant in equation 10
KQ	Semi-empirical measure of centrifuge size, eq. 11

n	Bowl speed, rpm
N	Number of disks
$Q_{act,theor}$	Actual and theoretical feed flow rate, m^3/s
$r_{e,1,2}$	Equivalent, min. and max. radius, m
S_e	Equivalent settling distance, m
t	Distance between disks, m
T	Spin time, s
V	Bowl volume, m^3
$V_{g,c}$	Settling velocity at 1 g and in centrifugal field, m/s
V_d	Velocity component caused by drag forces, m/s
Z	Relative centrifugal force (RCF) or G-number
α	Half cone angle of disc
β	Angle of tube in test tube centrifuge
η	Dynamic viscosity of fluid, mPa s
λ	Dimensionless number, defined in equation 5
μ	Separation efficiency factor
ν	Kinematic viscosity of fluid, m^2/s
ρ_p	Particle density, kg/m^3
$\rho_{f,fh,fl}$	Fluid density, general, heavy liquid, light liquid, kg/m^3
ω	Velocity of rotation, rad/s
Σ	Area equivalent according to Ambler, m^2

BIBLIOGRAPHY

1. A. Rushton, A.S. Ward, R.G. Holdich, *Solid Liquid Filtration and Separation Technology*, VCH Verlagsgesellschaft mbH, Weinheim, Germany, 1996.
2. V.I. Sokolov, *Moderne Industriezentrifugen*, VEB Verlag Technik, Berlin, 1971.
3. T.A. Barker, *Filtr. Sep.* **12**, 33–36 (1975).
4. G. Krook, H. Axelsson, C. Thorsson, in A. Blazej and V. Privarova eds., *Environmental Biotechnology*, Elsevier, Amsterdam, 1991.
5. G. Bergjohann, *Chem. Tech.* **17**, 85–86 (1988).
6. N.F. Madsen, *Inst. Chem. Engrs., Symp. Ser.* No. 113, 281–299 (1989).
7. F. Litvine, *Inf. Chimie* **310**, 327–334 (1989).
8. A. Lee, *Adv. Filtr. Sep. Technol.* **10**, 248–253 (1996).
9. S. Apelman, T. Björling, *Biotech. For. Eur.* **8**, 356–358 (1991).
10. Carr Separations, Inc., brochure, Franklin, Mass., 1997.
11. H. Axelsson, in C.L. Cooney and A.E. Humphrey eds., *Comprehensive Biotechnology. vol. 2: The Principles of Biotechnology: Engineering Considerations*, Pergamon Press, Oxford, U.K., 1985, pp. 325–346.
12. T. Lagerstedt, O. Nåbo, in C. Veret ed., *Flow Visualization IV*, Hemisphere, Washington, D.C., 1986.
13. O. Nåbo, T. Lagerstedt, in R. Reznicek ed., *Flow Visualization V*, Hemisphere, Washington, D.C., 1990.
14. O. Nåbo, C.-G. Carlsson, *Proc. Filtech Conf.* **1**, 1–7 (1987).
15. L. Borgström, C.-G. Carlsson, C. Inge, T. Lagerstedt, H. Moberg, *Appl. Sci. Res.* **53**, 35–50 (1994).
16. A.I. Clarkson, M. Bulmer, N.J. Titchener-Hooker, *Bioproc. Eng.* **14**, 81–89 (1996).
17. B. Lindman, *Meded. Fac. Landbouwwet., Rijksuniv. Gent* **52**, 1383–1387 (1987).
18. B. Madsen, *Inst. Chem. Eng., Symp. Ser.* No. 113, 301–317 (1989).
19. C.M. Ambler, *Chem. Eng. Progr.* **48**, 150–158 (1952).
20. C.M. Ambler, *J. Biochem. Microbiol. Technol. Eng.* **1**, 185–205 (1959).
21. L. Svarovsky, in L. Svarovsky ed., *Solid–Liquid Separation*, Butterworth, London, 1977, pp. 124–147.
22. F.E. Sullivan, R.A. Erikson, *Ind. Eng. Chem.* **53**, 434–438 (1961).
23. S. Aiba, S. Kitai, H. Heima, *J. Gen. Appl. Microbiol.* 243–256 (1964).
24. H. Avallone, P. D'Eramo, *Pharm. Eng.* **12**, 36–39 (1992).
25. J.J. Higgins, D.J. Lewis, W.H. Daly, F.G. Mosqueira, P. Dunnill, M.D. Lilly, *Biotechnol. Bioeng.* **20**, 159–182 (1978).
26. R.V. Datar, C.-G. Rosén, in G. Stephanopoulos ed., *Biotechnology. Vol 3: Bioprocessing*, 2nd ed, VCH Verlagsgesellschaft mbH, Weinheim, Germany, 1993, pp. 469–503.
27. R. Datar, *Filtr. Sep.* **21**, 402–406 (1984).
28. G. Taylor, M. Hoare, D.R. Gray, F.A.O. Marston, *Biotechnology*, **4**, 553–557 (1986).
29. N.J. Titchener-Hooker, D. Gritsis, R. Olbrich, K. Mannweiler, S.A.M. Gardiner, N.M. Fish, M. Hoare, *Pharm. Technol. Int.* **3**, 42–48 (1991).
30. S.O. Hwang, *Biotechnol. Tech.* **10**, 157–160 (1996).
31. A.I. Clarkson, P. Lefevre, N.J. Titchener-Hooker, *Biotechnol. Prog.* **9**, 462–467 (1993).
32. S.F. Siddiqi, N.J. Titchener-Hooker, P. Ayazi Shamlou, *Biotechnol. Bioeng.* **50**, 145–150 (1996).
33. I. Agerkvist, S.-O. Enfors, *Biotechnol. Bioeng.* **36**, 1083–1089 (1990).
34. A.C. Bentham, J. Bonnerjea, C.B. Orsborn, P.N. Ward, and M. Hoare, *Biotechnol. Bioeng.* **36**, 397–401 (1990).
35. R. Datar, C.-G. Rosén, *Chem. Eng. J.* **34**, B49–B56 (1987).
36. R. Kempken, A. Preissman, W. Bertold, *Biotechnol. Bioeng.* **46**, 132–138 (1995).
37. R.J. Wakeman, *Filtr. Sep.* **32**, 337–341 (1995).
38. G.J. Curiel, G. Hauser, P. Peschel, D.A. Timperley, *Trends Food Sci. Technol.* **4**, 225–229 (1993).
39. I. Weibull, *Mater. Des.* **8**, 35–40, 82–88 (1987).
40. M. Johnson, S. Lanthier, B. Massie, G. Lefebvre, A.A. Kamen, *Biotechnol. Progr.* **12**, 855–864 (1996).
41. C. Bowden, *Chem. Eng. (Rugby)*, **415**, 50–54 (June 1985).
42. G.D. Austin, R.W.J. Watson, T. D'Amore, *Biotechnol. Bioeng.* **43**, 337–341 (1994).
43. S.A. Siano, *Biotechnol. Bioeng.* **55**, 289–304 (1997).
44. G. Siems, *Brauwelt Int.* **2**, 132–137 (1997).
45. M. Mayer, R. Woernle, *Chem.-Ing.-Tech.* **57**, 152–153 (1985).
46. M.P. Backer, L.S. Metzger, P.L. Slaber, K.K. Nevitt, and G.B. Boder, *Biotechnol. Bioeng.* **32**, 993–1000 (1988).
47. V. Jäger, in R.E. Spier, J.B. Griffiths, W. Berthold eds., *Animal Cell Technology: Developments, Processes and Products*, Butterworth-Heinemann, Oxford, U.K. 1992.
48. N. Chatzisavido, T. Björling, C. Fenge, S. Boork, E. Lindner-Olsson, S. Apelman, in T. Kobayashi, Y. Kitagawa, K. Okamura eds., *Animal Cell Technology: Basic and Applied Aspects*, Kluwer Academic Publishers, Dordrect, The Netherlands, 1993, pp. 463–468.

49. P.J. Middelberg, B.K. O'Neill, I.D.L. Bogle, *Trans. Inst. Chem. Eng.* **70**, 8–12 (1992).
50. R. Datar, C.-G. Rosén, *J. Biotech.* **3**, 207–219 (1986).
51. D.J. Bell, M. Hoare, and P. Dunnill, in A. Fiechter ed., *Advances in Biochemical Engineering/Biotechnology*, vol. 26, Springer-Verlag, Berlin, 1983, pp. 1–72.
52. H. Katinger, F. Wibbelt, and H. Scherfler, *Verfahrenstechnik* **15**, 179–182 (1981).
53. C. Szabo, *Biotechnol. Bioeng.* **40**, 247–251 (1992).
54. H. Hustedt, K-H. Kroner, and N. Papamichael, *Process. Biochem.* **23**, 129–137 (1988).
55. P. Hambleton, J. Melling, T.T. Salusbury eds., *Biosafety in Industrial Biotechnology*, Chapman & Hall, London, 1994.
56. J. Lindley, *Chem. Eng. (Rugby)* **411**, 41–44 (1985).
57. M.J. Frude, M.T. Simpson, *Proc. Biochem.* **28**, 297–303 (1993).

See also CELL DISRUPTION AND LYSIS; CELL SEPARATION, SEDIMENTATION; CENTRIFUGES, ANIMAL CELLS.

CELL SEPARATION, FLOCCULATION

C.J.-P. BOONAERT
C.C. DUPONT-GILLAIN
P.B. DENGIS
Y.F. DUFRÊNE
P.G. ROUXHET
Université Catholique de Louvain
Louvain-la-Neuve, Belgium

KEY WORDS

Activated sludge
Aggregation
Brewing yeast
Coagulation
Downstream processing
Flocculation
Hydrophobicity
Interfaces
Separation
Zeta potential

OUTLINE

INTRODUCTION

In colloid science, flocculation is defined as an aggregation of solid particles induced by addition of polymers; coagulation refers to an aggregation due to van der Waals or electrostatic attraction between particles (1). Cell aggregation was defined by Calleja (2) as the gathering together of cells to form stable, contiguous, multicellular associations under physiological conditions. The criteria of intimate cell–cell contact and of physiological conditions are not met in the case of floc formation by the addition of flocculating agents; therefore, this definition establishes a distinction between aggregation and flocculation. However, polymers formed at a cell surface or excreted into the medium may be responsible for cell–cell interactions of the same nature as those produced by polymers added to the system.

Calleja's definition of cell aggregation involves a physical movement, either active or stochastic, which is responsible for the change from a dispersed to an agglomerated state. The criterion of stability may refer to aggregation under dispersive conditions, typically in an agitated suspension (as opposed to formation of a colony on agar), of a centrifugation pellet or a sediment. The criterion of physical movement emphasizes that a bond is being formed between distinct bodies, as opposed to chains and mycelial structures. However, certain cell aggregates of great significance in biotechnological processes are not simply formed by a gathering of initially dispersed cells. In activated sludges and in cell granules formed in anaerobic digestion, a dynamic situation is obtained in which some newly formed cells remain more or less loosely attached to each other, either by cell–cell or cell–debris interactions, or because a matrix prevents their dispersion, while free cells may be captured by an existing aggregate.

From the biotechnologist's point of view, the distinction between coagulation, aggregation, flocculation, and agglomeration is not always clear. The words *aggregate, floc, agglomerate*, and *cluster* are often synonymous. They refer

to an assemblage of individual cells, formed by initially free cells or by the lack of cell dispersion and occurring under physiological conditions or after addition of various agents to the medium.

Aggregated microbial cells are of great interest in biotechnology. They offer advantages, with respect to fully suspended cells in downstream processing, by facilitating cell-broth separation by sedimentation or filtration. They also allow continuous fermentation processes by keeping the cells in the bioreactor with an upflow of the broth. Further details concerning cell flocculation can be found in several books (2–4). Biofilms (5–8) are not considered explicitly here; however, as concerns the mechanisms of association, they have many features in common with flocs.

PHYSICOCHEMICAL ASPECTS OF AGGREGATION

Interactions between Particles

The essential features of the interactions between colloidal particles may be understood by considering three different aspects (9,10): (*1*) DLVO (Derjaguin, Landau, Verwey, Overbeek) theory (see "DLVO Theory") accounts for van der Waals and electrostatic forces; (*2*) when surfaces are close to molecular contact, solvation interactions (see "Solvation Interactions") may play an important role; (*3*) when mobile macromolecules are present at the interface, they may exert a particular influence at long distances (see "Influence of Macromolecules at the Interface") due to their interactions with the solvent, with the surfaces, and with one another. In the section "Charge Neutralization by Cationic Polymers", a particular mention is made of the possibility to neutralize the surface charge and thereby provoke flocculation by oppositely charged macromolecules.

DLVO Theory. According to DLVO theory (9,10), the free enthalpy of interaction, G_t, between two particles approaching each other may be accounted for by the sum of two contributions, corresponding respectively to electrostatic and van der Waals interactions

$$G_t = G_e + G_v \tag{1}$$

The electrostatic contribution, G_e, is due to the overlap of electrical diffuse double layers of approaching particles. For two spheres of the same nature and same radius, R, suspended in a symmetrical electrolyte solution,

$$G_e = (64\pi kTRNC\kappa^{-2}T_d^2)\exp(-\kappa D) \tag{2}$$

where k is the Boltzmann constant, T the temperature, N the Avogadro number, C the electrolyte concentration (mol/L), and D the separation between particle surfaces.

$$T_d = \tanh[ze\psi_d/4kT] \tag{3}$$

where z is the ion charge, e is the electron charge, and ψ_d the electric potential at the inner limit of the diffuse part of the double layer. It differs from the surface potential, ψ_0, due to adsorption of ions in intimate contact with the surface, that is, in the Stern layer (chemically adsorbed or retained by localized electrostatic attraction) (10,11). For computing G_e, the value taken for ψ_d is usually the zeta potential ζ, that is, the electric potential at the plane of shear between the cell surface and the solution (11). κ, the reciprocal of the Debye length, determines the rate of the variation of the electric potential as a function of the distance to the surface, within the diffuse double layer.

$$\kappa = (1000e^2N/\epsilon kT)(2I)^{1/2} = 2.32\ 10^9(2I)^{1/2}\ (\text{at } 25\ ^\circ\text{C}) \tag{4}$$

where I is the ionic strength of the bulk solution and ϵ the electric permittivity. At an ionic strength of 10^{-1}, 10^{-3}, and 10^{-5} M, κ^{-1} is close to 1, 10, and 100 nm, respectively. Note that the ionic strength is commonly in the range of 10^{-2} to 10^{-3} M for natural water supplies and near 10^{-1} M for microbial culture media.

When the surface potential is low ($<25/z$ mV), G_e simplifies as follows:

$$G_e \cong 2\pi R\epsilon\psi_d^2 \exp(-\kappa D) = 2\pi R\sigma^2 \exp(-\kappa D)/\kappa^2\epsilon \tag{5}$$

where σ is the surface density of charge (C m^{-2}). This equation is also valid for nonsymmetrical electrolytes.

For two spheres with the same radius, R, the contribution of van der Waals interactions is given by

$$G_v = -\frac{AR}{12D} \tag{6}$$

where A is the Hamaker constant. This is mainly determined by dispersion forces (surface–surface, surface–liquid) and thus by the polarizability of the surface-constituting molecules. For organic particles in suspension in water, it is reasonable to consider A values in the range of 0.2 to 1.6 kT (kT = 4.1×10^{-21} J mol^{-1} at room temperature). Values are close to 5 kT for mica and in the range of 2 to 50 kT for oxides depending on the atomic number of the metal ion (9,12).

Figure 1 presents the computed G_t vs D curves for spherical particles of 1-μm radius, A = 1 kT, $\psi_d = -20$ mV, and κ^{-1} = 10 nm. The respective influence of variation of A, ψ_d, and κ^{-1} is also illustrated. The contribution of van der Waals forces is shown in Fig. 1b (ψ_d = 0 mV). The figure shows that, even at low values of the diffuse-layer potential (Fig. 1b), a potential energy barrier appears at a distance of a few nanometers. This barrier may be very high compared with the kinetic energy associated with Brownian motion, equal to 1.5 kT. At sufficiently high ionic strength (Fig. 1c), the variations of G_e and G_v as a function of the interdistance bring about a secondary minimum that holds the particles together, although not in close contact.

The efficiency of electrolytes to provoke coagulation of colloidal particles is usually expressed by the concentration required to achieve a rapid coagulation, called critical coagulation concentration. According to the valency rule of Schulze–Hardy, this is strongly dependent upon the valency of the ion with a charge opposite to that of the surface (counterion) and the critical coagulation concentration decreases as the counterion valency increases (10). Consider a negatively charged surface; the decrease of the critical

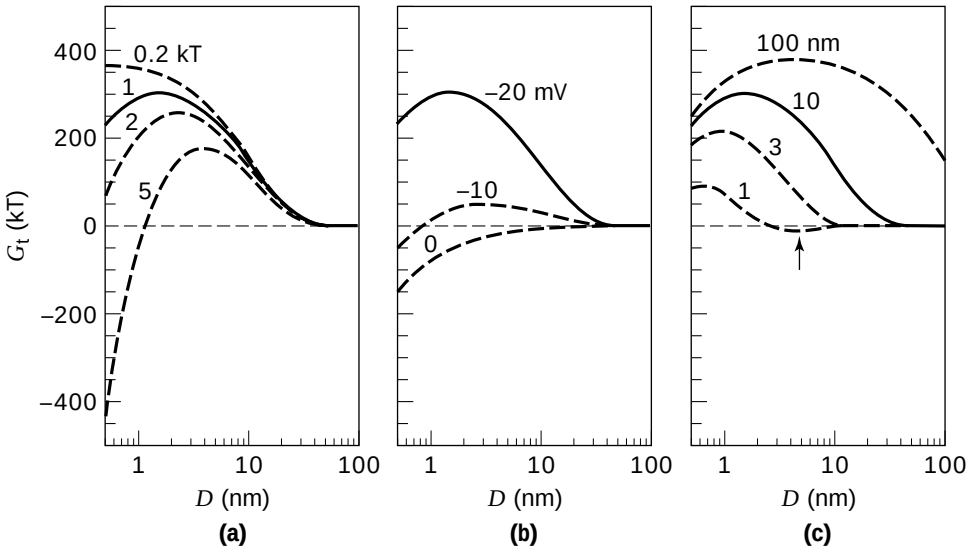


Figure 1. Variation of the free enthalpy of interaction as a function of the distance between two identical spherical particles of 1-μm radius, according to DLVO theory. Solid curve: Hamaker constant $A = 1$ kT, potential of the diffuse double layer $\psi_d = -20$ mV, $\kappa^{-1} = 10$ nm; (**a**) $\psi_d = -20$ mV; $\kappa^{-1} = 10$ nm, A varying from 0.2 to 5 kT; (**b**) $\kappa^{-1} = 10$ nm, $A = 1$ kT, ψ_d varying from 0 to -20 mV; (**c**) $A = 1$ kT, $\psi_d = -20$ mV, κ^{-1} varying from 1 to 100 nm (secondary minimum indicated by arrow).

coagulation concentration is explained by a decrease of the potential energy barrier as a result of two effects: (*1*) if the valency of the cation present in the solution is higher, its adsorption in the Stern layer rises, the potential drop in the Stern layer is larger, and ψ_d is less negative; and (*2*) as the valency of the cation increases at a given concentration, the ionic strength rises, the Debye length κ^{-1} decreases, and the potential barrier decreases (Fig. 1c). The critical coagulation concentration was found to increase in the order $Al^{3+} < Ca^{2+} < Na^+$ for *Escherichia coli* (13).

Equations 2 and 6 are restricted to situations where $D \ll R$; within this limit, the interaction energy is proportional to the particle radius. In DLVO theory, the interacting surfaces are considered as homogeneous, smooth, and rigid. This model is very rough for particles of biotechnological interest. Microbial cells may have complex shapes and bear surface appendages such as fimbriae, pili, fibrils, and flagella. As a consequence, the radius of curvature ruling the interactions may be much smaller than the overall cell radius. In particular, it is conceivable that cells become bound via appendages through a potential barrier resulting from average surface properties, as illustrated by Figure 2. It is also possible that certain zones of the surfaces of two particles are in molecular contact, and thus that solvation interactions have a significant influence, while other zones are still separated by an appreciable distance. The higher the ionic strength, the larger the possible impact of this mechanism.

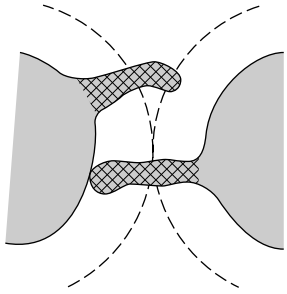


Figure 2. Schematic illustration of the penetration of a surface appendage through a potential energy barrier, the location of which is represented by dashed lines.

Solvation Interactions. Between hydrophilic surfaces in water, a short-range repulsive force arises, which tends to prevent the removal of water from the surfaces and is commonly referred to as hydration force. This has been attributed to the structuring of water molecules at the surfaces through H-bonding with hydrophilic surface groups. However, consideration of recent experimental and theoretical results has led to the suggestion of an alternative interpretation, in which the short-range force is an entropic repulsion arising from the confinement of mobile surface groups (14). The hydration repulsion G_{rh} between two hydrophilic surfaces appears to follow the empirical equation (9)

$$G_{rh} = +G_{rho} \exp(-D/\lambda_{rh}) \qquad (7)$$

where the decay length, λ_{rh}, is of the order of 1 nm for 1:1 electrolytes, and G_{rho} (3–30 mJ m^{-2}) depends on hydration of the surface, including hydration of the cations bound to a negatively charged surface, that is, held in the Stern layer.

The effect of electrolytes on the hydration forces may thus determine whether particles will coagulate or not. The effectiveness of cations as coagulants usually increases according to the so-called lyotropic series (9,10): $Li^+ < Na^+ < K^+ < Cs^+$ for monovalent cations, $Mg^{2+} < Ca^{2+}$ for divalent cations. The decrease of ion hydration is responsible for a decrease of hydration forces between the surfaces. For some surfaces, it is also accompanied by increased ion binding at the surface (11), and thus by a larger drop of the electric potential in the Stern layer.

A hydrophobic surface cannot bind water molecules via ionic or hydrogen bonds, therefore, an attractive hydrophobic interaction arises between two hydrophobic surfaces in water. This results from the strong tendency of water molecules to associate through hydrogen bonding (9,15). To a large extent, hydrophobic bonding is due to the fact that the surface reduces the degrees of freedom offered to adjacent water molecules for forming as many hydrogen bonds as possible; as a consequence, the water structure near the surface is more ordered. Hydrophobic bonding is thus entropically determined, and the resulting association of surfaces should be viewed as the result of a thermodynamic balance rather than the effect of a force, in the common sense. As a consequence of the entropic character, the effect of hydrophobic bonding increases markedly with temperature. For two surfaces in water, the hydrophobic interaction energy, G_{ah} (ignoring DLVO and other solvation forces), is given by (9)

$$G_{ah} = -2\gamma \exp(-D/\lambda_{ah}) \tag{8}$$

where the decay length λ_{ah} is equal to 1–2 nm and γ is the surface–water interfacial energy (10–50 mJ m^{-2}). The rather long-range nature of the hydrophobic interaction accounts for the rapid coagulation of hydrophobic particles in water and the folding of proteins.

Influence of Macromolecules at the Interface. A nonionic polymer dissolved in an ideal solvent (no interaction between polymer segments and the solvent) takes the shape of an "unperturbed" random coil, characterized by the radius of gyration

$$R_g = \frac{ln^{0.5}}{\sqrt{6}} \tag{9}$$

where l is the effective segment length and n is the number of segments. R_g may thus vary from a few to about 100 nm. In a real solvent, the size of the coil may be different from R_g and is given by the Flory radius

$$R_F = \alpha R_g \tag{10}$$

In a poor solvent, the segments attract each other, the coil is more compact, and α is smaller than unity. In a good solvent, the coil swells and its Flory radius is given by

$$R_F \cong ln^{0.6} \tag{11}$$

A poor solvent may become a good solvent by adding certain solutes or changing the temperature. The temperature at which $\alpha = 1$ is known as the theta (θ) temperature, and an ideal solvent is known as theta solvent (9).

Figures 3 and 4 illustrate the different situations that may be encountered when two surfaces approach each other while carrying a nonionic polymer. Once surface separation is a few times the radius of gyration, the segments overlap, which leads to a repulsive osmotic force, that is, to the decrease of entropy associated with a closer confinement of the polymer chains between the two surfaces. This is known as steric repulsion and is responsible for steric stabilization of colloids by polymer additives (9).

As a simple case, consider a low coverage of polymer molecules (no overlap or entanglement of neighboring chains), each being grafted at one end to the surface (Fig. 3 top left; Fig. 4 right, low density). This may be relevant to microbial surfaces. In a theta solvent, if the number of chains grafted per unit area is Γ, the repulsive energy per unit area is approximately given by

$$G_s \cong 36\Gamma kT \exp(-D/R_g) \tag{12}$$

In a good solvent, the coils swell and the range of repulsion is broader. When the surface density of grafted chains is high, so that they are forced to extend away from the surface much farther than R_g, a polymer brush is obtained (Fig. 3, top right). The thickness of the polymer layer, L, no longer varies according to the power 0.5 or 0.6 of the degree of polymerization but is proportional to the polymer length. Once two brush-bearing surfaces are closer than $2L$, a repulsive energy arises (Fig. 4 right, high density). For a good solvent and for $D/2L$ in the range of 0.2 to 0.9, if s is the average distance between the grafting points, the interaction energy per unit area is approximately given by

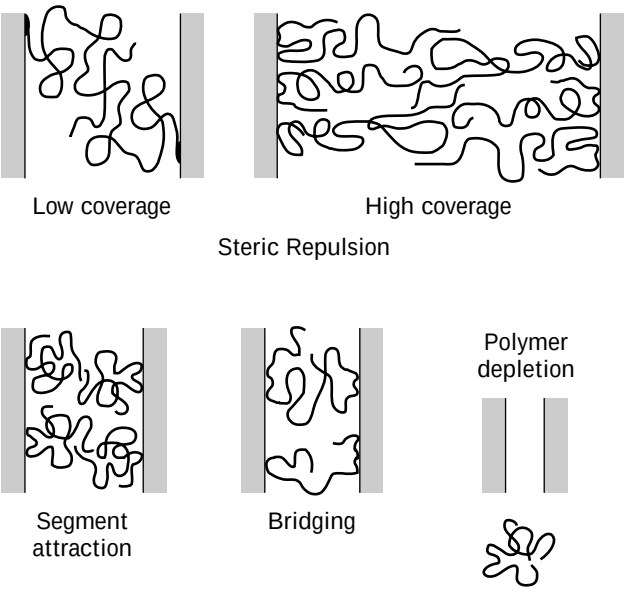


Figure 3. Illustration of steric forces generated by polymers between two surfaces (grafted or adsorbed). *Source*: Adapted from Ref. 9.

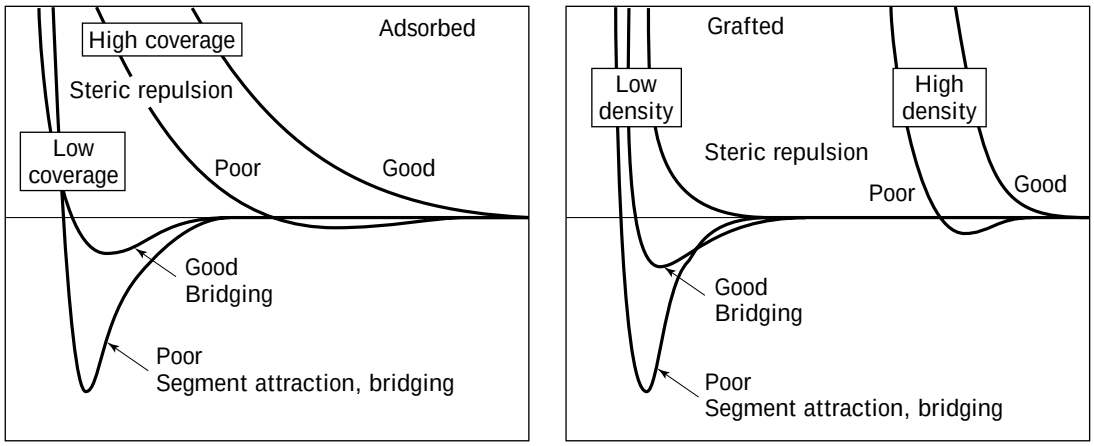


Figure 4. Variation of interaction free enthalpy as a function of the separation distance between two particles as a result of macromolecular interactions; influence of the mode of polymer retention at the surface (adsorption, grafting), polymer concentration at the surface (low, high), and solvent quality (good, poor). *Source*: Adapted from Ref. 9.

$$G_s \cong \frac{100L}{\pi s^3}\, kT\, \exp(-\pi D/L) \qquad (13)$$

In this case, G_s is the balance of an osmotic repulsion between the coils and the elastic energy that opposes chain stretching.

If the polymer is adsorbed (Fig. 3 left, bottom compared to top), a fraction of its segments (trains) are in contact with the surface; the other segments belong to loops and tails dangling in the solution. The polymer may adopt a conformation leading to an adsorbed thickness smaller than the radius of gyration. Things are more complicated than for grafted polymers due to the dynamics allowed by a multiplicity of weak bonds between the polymer and the surface, and to a possible variation of the adsorbed amount and the binding sites as the two surfaces approach each other. As a consequence, the variation of the interaction free energy as a function of distance is less sharp (Fig. 4, left compared with right). Moreover the force between two surfaces at a given separation distance may take a long time (hours) to reach a stationary value (16).

Attractive forces may also be mediated by adsorbed or grafted polymers (Fig. 3, bottom; Fig. 4) (9).

- In a poor solvent, segment attraction between molecules bound to opposite surfaces leads to overall attraction between the surfaces until a balance with steric repulsion is reached.
- When the degree of coverage is weak and the polymer is strongly adsorbed, the macromolecules may form bridges between the two surfaces. Under suitable conditions, bridging forces can extend over longer distances than van der Waals forces.
- When a polymer does not adsorb, bringing two surfaces at a distance smaller than R_g will push the coils out of the space between the surfaces. The polymer depletion is responsible for a difference of osmotic pressure between the bulk liquid phase and the liquid phase held between the surfaces. This effect is small compared to van der Waals and solvation interaction energies.

In a good solvent, the efficacy of a polymer to provoke flocculation is mainly determined by its molecular mass and its concentration, which determine the degree of coverage and the thickness of the adsorbed layer. An optimal flocculation concentration is found for a given molecular mass. It is influenced by pH and ionic strength, which determine the electrostatic repulsion between the surfaces. Flocs will form if the surface is partly covered and if the adsorbed molecules extend far enough from the surface to span the distance at which the electrical repulsion operates (16,17), as illustrated by Figure 5. On the left side of the figure, the molecular mass or the ionic strength are too low to allow bridging. On the right side, the ionic strength is higher (and the Debye length, lower) or the surface potential is weaker, the electric repulsion operates at short distances compared with the thickness of the adsorbed macromolecules, and bridging occurs.

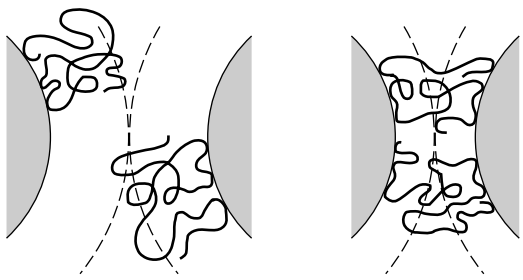


Figure 5. Illustration of the effect of ionic strength or surface potential on polymer bridging; the dashed lines represent the distance at which the electrical repulsion operates. The increase of ionic strength or decrease of surface potential from left to right allows polymer bridging. *Source*: Adapted from Ref. 18.

As polymer adsorption is slow and irreversible, flocculation may often proceed through nonequilibrium processes, which are strongly dependent on the procedure followed (manner and rate of mixing the particle suspension and the polymer solution; stirring). For instance, heterocoagulation may occur between those particles free of polymer and those that are saturated. In other situations, the surface may be partly covered by macromolecules that have not yet collapsed, thus allowing polymer bridging and producing transient flocs, the lifetime of which may vary from seconds to days, depending on the conditions (16,19).

Charge Neutralization by Cationic Polymers. As materials of biotechnological relevance are usually negatively charged, only cationic polyelectrolytes are considered here. Due to electrostatic repulsion between segments, polyelectrolytes display more or less extended conformations according to ionic strength and to their charge density, and thus to the fraction of charged segments τ. For a given molecular mass, the maximum adsorbed amount increases with τ because of the stronger polymer–surface interaction, passes through a maximum, and decreases due to increasing electrostatic repulsion between polymer segments (18,20,21). At low polymer ionicity (small τ), the maximum amount adsorbed (expressed in mass) increases as a function of molecular mass. At high ionicity, the maximum adsorbed amount and the adsorbed-layer thickness are practically independent of the molecular weight; the loop size is small and the polymer chain adopts a flat conformation on the surface. The flat conformation realizes a local neutralization of the particle surface charge, but reduces the possibility of bridging contacts with other particles.

With a cationic polyelectrolyte of high charge density, the influence of polymer molecular mass and of ionic strength can be explained by the electrostatic patch model (18,20,21) illustrated by Figure 6. The adsorbed polyelectrolyte represents patches of positive charges surrounded by the original particle surface. Particles with this mosaic surface–charge distribution may interact in such a way that positive and negative patches of different particles come into contact, causing a strong attachment. As the surface coverage increases, the particle becomes totally coated by the cationic polymer, electrostatic repulsion arises between particles, and the colloid remains in a dispersed state. The optimum flocculation concentration corresponds to a situation where the particle surface is partially covered by patches of cationic polymer. This is illustrated by Figure 7, which shows the influence of polymer concentration on flocculation of a clay, evaluated by the relative supernatant turbidity and the sedimentation time (22). The optimum flocculation concentration coincides with reversal of the zeta potential of the particles.

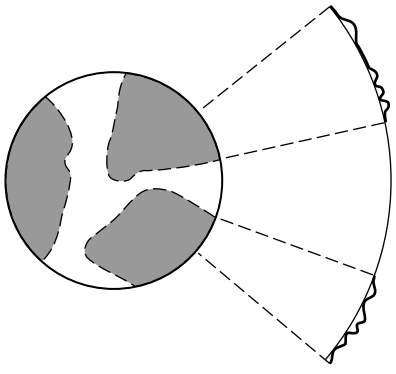


Figure 6. Patch model for a cationic polyelectrolyte (shaded area) adsorbed on a negatively charged surface. Left, view of the particle surface; right, section through a part of the surface. *Source*: Adapted from Refs. 20 and 21.

With a cationic polymer of low ionicity, the influence of charge density, concentration, and molecular mass is more complex, due to the combination of the charge-neutralization and bridging mechanisms. The influence of ionic strength is also complex because it affects the screening of segment/segment lateral repulsion (which decreases chain stiffness and increases loop size and adsorbed amount) and the screening of segment/surface attraction (which reduces the proportion of trains). In practice, increasing the ionic strength reduces the concentration at which the flocs initially appear and broadens the flocculation domain (18).

Rate of Aggregation

The rate of aggregation is determined by the probability of collision between particles, which depends on their concentration, and by the efficiency of a collision to make a bond between the particles. The collision may occur as a result of two types of transport processes: agitation and Brownian motion. If one considers the collision between two identical spherical particles of diameter D^*, the ratio between the probability of collision due to agitation P_o (orthokinetic conditions) and the probability of Brownian collision P_p (perikinetic conditions) is given by

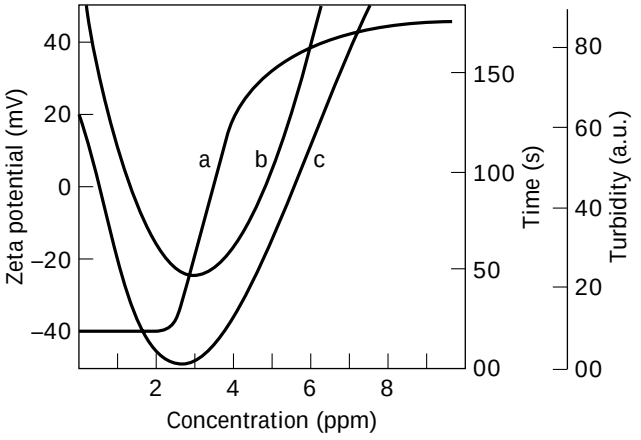


Figure 7. Flocculation of Na-montmorillonite by *N,N,N*-trimethyl aminoethyl chloride acrylate. Variation of (**a**) zeta potential, (**b**) sedimentation time, (**c**) relative supernatant turbidity, as a function of the amount of polymer introduced per unit volume. Polymer molecular weight 10^6; τ = 100%; clay concentration 0.05 g L^{-1}; NaCl concentration 1.2 mM. *Source*: Adapted from Ref. 22.

$$\frac{P_o}{P_p} = \frac{\eta D^{*3} G^*}{2kT} \quad (14)$$

where η is the dynamic viscosity, G^* the shear rate, k the Boltzmann constant, and T the temperature (23). In water at room temperature (η = 1 mPa s), with agitation typical of a magnetic stirrer (G^* = 100 s^{-1}), this ratio is about 100 for particles of 2 μm (24). The rate of aggregation of microbial cells thus depends strongly on agitation due to their relatively large size and slow Brownian motion, unlike small colloidal particles. On the other hand, agitation can also break the flocs so that a stationary state will be reached, with a characteristic distribution of aggregation numbers.

The chance that a collision gives rise to a bond between the particles is determined by the variation of the free energy of particle–particle interactions as a function of the distance. Four typical situations, illustrated by Figures 1 and 4, may be encountered: *(1)* the interaction energy is attractive at all distances, and a quick coagulation occurs (low ψ_d, high ionic strength, high Hamaker constant; Fig. 1); *(2)* the interaction is attractive at short distances but is repulsive at longer distances and passes through a maximum (Fig. 1); in this case the probability that a collision leads to bond formation is the probability of passing over the energy barrier; *(3)* the interaction is repulsive at all distances and no aggregation occurs; the colloid suspension is said to be stabilized (high ψ_d, low ionic strength, low Hamaker coefficient, Fig. 1; steric repulsion, Fig. 4); *(4)* the interaction energy curve presents a well at a certain separation distance (secondary minimum, Fig. 1; bridging, segment attraction, Fig. 4), and the aggregates will be constituted by particles maintained at a certain distance from each other.

Surface Properties of Particles

The nonspecific interactions that control particle aggregation (see "Interactions between Particles") depend on the physicochemical properties of the surfaces (electrical properties, hydrophobicity, presence of macromolecules) and consequently on their surface chemical composition. Physicochemical methods that are well-established or are being developed to investigate these properties are described in the following sections.

Electrical Properties. As discussed earlier, it appears that the electrical properties of particle surfaces strongly influence coagulation by addition of salts, or influence flocculation by polyelectrolytes. The zeta potential, ζ, is usually considered to be close to the potential relevant for DLVO computations (10,11,25,26). It can be deduced from the electrophoretic mobility, μ. For spherical particles of radius R,

$$\zeta = \frac{\eta}{\epsilon}\mu B = 12.85\,\mu B \quad (15)$$

at 298 K, if ζ is given in mV and μ in 10^{-8} m^2 V^{-1} s^{-1}. B = 1 (Helmholtz–Smoluchowski equation) when $\kappa R > 100$, which is relevant for microbial cells in suspensions of common ionic strength; B = 3/2 (Hückel–Onsager equation) when $\kappa R < 1$. In the intermediate situations, B is a function of κR.

Equation 15 implies that: *(1)* the solid plus stagnant layer together behave electrokinetically as a solid particle, *(2)* particles are nonconducting, *(3)* particle surfaces are smooth and homogeneous, and *(4)* the viscosity, dielectric constant, and ion conductivities in the diffuse part of the double layer are equal to their bulk values (11). When the surface conductance or the bulk (or cell wall) conductance are not negligible, the relationship between electrophoretic mobility and zeta potential is more complex (27). Therefore it is recommended to report experimentally determined electrophoretic-mobility values instead of ζ potentials.

Hydrophobicity. The surface hydrophobicity of microbial cells can be assessed by the water contact angle measured on particle lawns (28). A greater cell-surface hydrophobicity is of practical concern in separation: It favors cell coagulation and cell accumulation at the liquid–gas interface. The latter effect may be used for increasing cell concentration by flotation (29).

Contact angles of liquids have been used to deduce the surface energy of solids and to predict the stability of solid–solid associations. The model is based on the balance of interfacial free energies involved in the creation of a new solid–solid interface and in the destruction of the interface of each partner with water. This approach is subject to severe limitations. *(1)* There are different theoretical frames that lead to different values of interfacial energy (30), indicating that absolute values cannot be considered individually. Nevertheless, classification of different materials according to their surface free energy is the same, whatever the approach used. For polymer materials, this classification is the same as that given by the contact angle of water. *(2)* Because interfacial energies are deduced from contact angles of liquids, the energy balance does not incorporate the influence of electrostatic interactions, which act at long distances, or the influence of ionic bridges and steric interactions. *(3)* The system is supposed to be at equilibrium and thus reversible, which is not in accordance with experimental data. *(4)* A close contact is assumed to be achieved between the surfaces, which is not reasonable in view of the roughness of most surfaces.

Other methods of evaluation of microbial-cell hydrophobicity involve adhesion at hydrophobic surfaces or aggregation at high salt concentration (28).

Surface-Chemical Composition. The surface-chemical composition of solids, including microbial cells, can be determined by X-ray photoelectron spectroscopy (XPS). The technique involves irradiation of the sample by an X-ray beam, which induces ejection of electrons. The kinetic energy of the emitted electrons is analyzed, and their binding energy determined. Each peak is then characteristic of a given element. Due to inelastic scattering of electrons in the sample, the collected information concerns only the outermost molecular layers of the surface (2–10 nm). As the position of the photoelectron peak is slightly affected by the chemical environment of the element, peak-shape analysis provides further information on chemical func-

tions (31). The data can be converted into concentrations of model constituents: proteins, polysaccharides, and hydrocarbon-like compounds (32).

For microorganisms, correlations have been obtained between the surface-chemical composition obtained by XPS and surface properties, that is, hydrophobicity and electrical properties. Furthermore, the surface composition has been found to be related to the behavior of microorganisms at interfaces (aggregation, adhesion) (33).

Nanometer-Scale Surface Properties. The physicochemical methods just described have a poor lateral resolution, the information provided concerning only the overall surface properties of particles. Exciting new possibilities are now offered by atomic force microscopy (AFM) to investigate the molecular interactions and properties of surfaces on the nanometer scale. AFM, invented in 1986 (34), consists of measuring the forces acting between a sharp probe and the surface of a sample. It allows the nanoscale mapping of surface topography, as well as the direct measurement of mechanical properties, friction, and interaction forces (electrostatic, van der Waals, solvation, steric interactions). As opposed to more conventional surface analysis techniques, AFM can be operated in aqueous solutions, which makes it possible to study living cells in physiological conditions (35).

A very challenging goal for future research is to directly map the surface properties of living cells in aqueous conditions. The use of functionalized AFM probes, possessing different charges, surface energies, and solvation properties, should make it possible to measure: *(1)* the local surface hydrophobicity and electrical properties of individual cells, *(2)* the surface distribution of solvated macromolecules such as exopolysaccharides, and *(3)* steric interactions associated with these macromolecules. These data will provide new insight into the molecular mechanisms of cell aggregation.

BREWING YEAST: A MODEL FOR THE IMPLICATION OF SPECIFIC INTERACTIONS IN FLOCCULATION

Flocculation and Flocculence

Flocculation of brewing yeasts, thoroughly reviewed by Stratford (36), is the most extensively studied of all systems in which cell aggregates are of interest. It is illustrative of the diversity of approaches used to understand flocculation of microbial cells and the interplay between specific and nonspecific interactions. Spontaneous aggregation of yeast cells occurs near the end of fermentation and is of considerable importance in brewing. For instance, early sedimentation in lager beer production will provoke an incomplete or too slow fermentation; failure to sediment will complicate processing and may lead to alteration of taste.

In brewing technology, the flocculence of a strain means its ability or tendency to flocculate in industrial conditions. More broadly, flocculence is the ability to flocculate if all environmental conditions are favorable (37). The term *flocculability* has been proposed to designate the ability to flocculate in a given medium if hydrodynamic conditions and cell concentration are adequate (38). Beyond the terminology, it is essential to keep in mind that the occurrence of flocculation depends on *(1)* the nature of the biological material (strain; physiological state defined by growth conditions and growth phase), *(2)* the physicochemical environment (temperature, medium composition), and *(3)* the physical conditions (velocity gradient, cell concentration). Thus, the evaluation of the ability to flocculate must always be related to experimental conditions.

In order to test the ability of yeast to flocculate in a given medium, it is important to control the initial state of dispersion and to distinguish the conditions selected for the aggregation process from those allowing evaluation of the final state of dispersion. A screening test may conveniently be performed according to the following procedure (24,39). Introduce the washed cells into test tubes containing the desired medium and vigorously stir to ensure dispersion. Agitate gently for a given period to allow flocculation to take place, then let stand undisturbed to allow floc sedimentation. Finally, measure the absorbance of the supernatant, which reflects the concentration of cells remaining free. The volume of the sediment may also be a way to evaluate the final state of dispersion, as the bed volume of flocculated particles is larger than that of individual particles (40).

Yeast cell behavior in laboratory conditions does not always reproduce the behavior in industrial conditions. Among 45 strains noted as flocculent in brewing, only 20% did flocculate during a culture in laboratory conditions (shaken Erlenmeyer flasks) (41). This was attributed to agitation and aeration rather than to final pH or the requirement of particular wort constituents.

Several parameters were used to quantify yeast flocculation: initial rate of flocculation, rate of settling, number of free cells, floc size, and floc strength (38,42). It was estimated that the attraction force between flocculent cells is three orders of magnitude larger than the force calculated with the DLVO theory (43). The dynamic equilibrium between floc formation and floc breaking (i.e., splitting, which determines the average floc size, and erosion, which determines the number of single cells) was modeled. This showed the dependence of floc diameter and number of single cells on bond strength and agitation (44).

Yeast Behavior and Cell Surface Properties

There are two types of processes for beer brewing. Bottom fermentation, producing lager beer, is carried out at low temperature (near 10 °C) by *Saccharomyces carlsbergensis*; at the end of fermentation, the cells flocculate and settle. Top fermentation, typically producing ale, is carried out at higher temperature by *Saccharomyces cerevisiae*; when the cells flocculate, they do not settle but they tend to remain associated with CO_2 bubbles and rise to the top of fermented wort. The surface properties (electrophoretic mobility, water contact angle, and surface composition) of a collection of yeasts were investigated in order to examine their possible relationship with the type of strain, growth phase, flocculence, and occurrence of flocculation (45). There is no clear-cut distinction between surface properties of top and bottom strains; however, they differ on a

statistical basis. On the average, top strains have more protein at their surface and are more hydrophobic (Fig. 8), which may be related to their tendency to remain associated with gas bubbles and move to the top of the wort. Bottom strains contain more phosphate. The difference between the profiles of electrophoretic mobility vs. pH (Fig. 9) may be explained by modelling the surface charge according to the surface chemical composition, as given by XPS. For the bottom strains, the electrical properties are controlled mainly by phosphate, resulting in a low isoelectric point (pH 2 or below) and an electrophoretic mobility that does not change appreciably above pH 4. For the top strains, the electrical properties are determined mainly by the balance of protonated amino and carboxylate groups in proteins, which give a higher isoelectric point (pH 4) and an electrophoretic mobility changing regularly from pH 2 to 7.

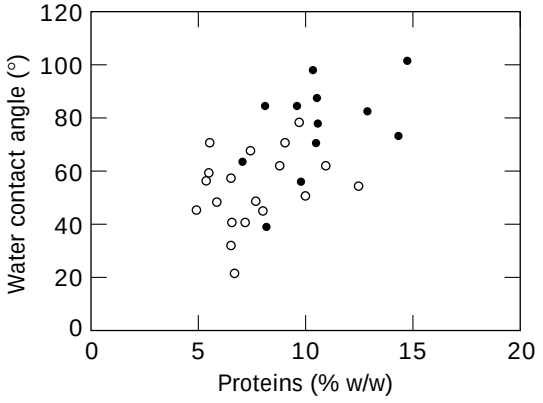


Figure 8. Water contact angles of bottom-fermenting strains (○) and top-fermenting strains (●) as a function of the cell-surface concentration (weight %) of proteins as obtained by modeling XPS data. *Source*: Adapted from Ref. 45.

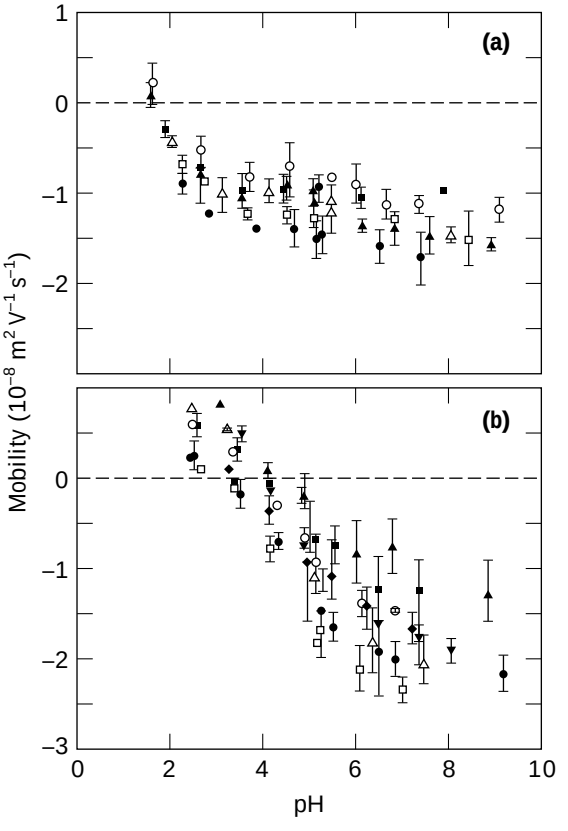


Figure 9. Electrophoretic mobility as a function of pH for bottom-fermenting strains (**a**) and top-fermenting strains (**b**) in 1 mM KNO_3 (45).

Although surface properties of brewing yeasts were statistically related to the top/bottom character, they did not show correlations with the flocculating character. No differences of surface properties were found between strains that did and did not flocculate in laboratory conditions, or between cells from the exponential and stationary growth phases, even for strains where flocculation occurred during the transition from one growth phase to the other. This is in contrast with reports (46,47) that ability to flocculate during growth coincides with an increase in cell-surface hydrophobicity correlated with the appearance of fimbriae-like structures. It was also reported (48) that flocculent cells possess a fimbriate or hairy surface, whereas nonflocculent cells are smooth.

Certain yeast strains produce aggregates with a particular geometric shape, such as cubes, lozenges, cylinders, and disks (36,49), the shape being strongly dependent on agitation. Cubic flocs, close to 1 cm in size, produced by *S. cerevisiae* used for continuous production of ethanol are shown in Figure 10. When agitation is stopped, the flocs collapse and, within a few tens of minutes give rise to a cell carpet on the bottom of the flask. This is tentatively attributed to the fact that the cells are coated by polymers that ensure a weak bridging, as schematized in Figure 11. The viscoelastic properties of the interface would be such that, upon agitation, binding prevails due to chain entanglement, with the hydrodynamic conditions controlling the floc shape; on the other hand, repulsion would prevail in resting conditions (49).

Genetic and Biochemical Aspects

According to the most widely considered mechanism of flocculation proposed by Miki et al. (50), lectinlike molecules (sugar-binding proteins that need Ca^{2+} to adopt their active conformation) recognize and bind mannans at the surface of adjacent cells (36,51), as illustrated by Figure 12(a).

Stratford and Assinder (52) distinguished two flocculation phenotypes, both based on a lectin-mediated mechanism: Flo1 phenotype strains, which included most of the laboratory strains, had *FLO1* genes and possessed mannose-specific lectins; NewFlo phenotype strains, which included most ale-brewing strains, possessed less-specific lectins, binding manno- and glucopyranoses. Masy et al. (53) found the same two phenotypes and called them MS (mannose sensitive) and GMS (glucose mannose sensitive). On the other hand, a different phenotype, called MI (mannose insensitive), representing strains that are not inhibited by sugars, including mannose and glucose, was reported (41,53).

In the last few years, much progress has been made as a result of advances in the fields of genetics and biochem-

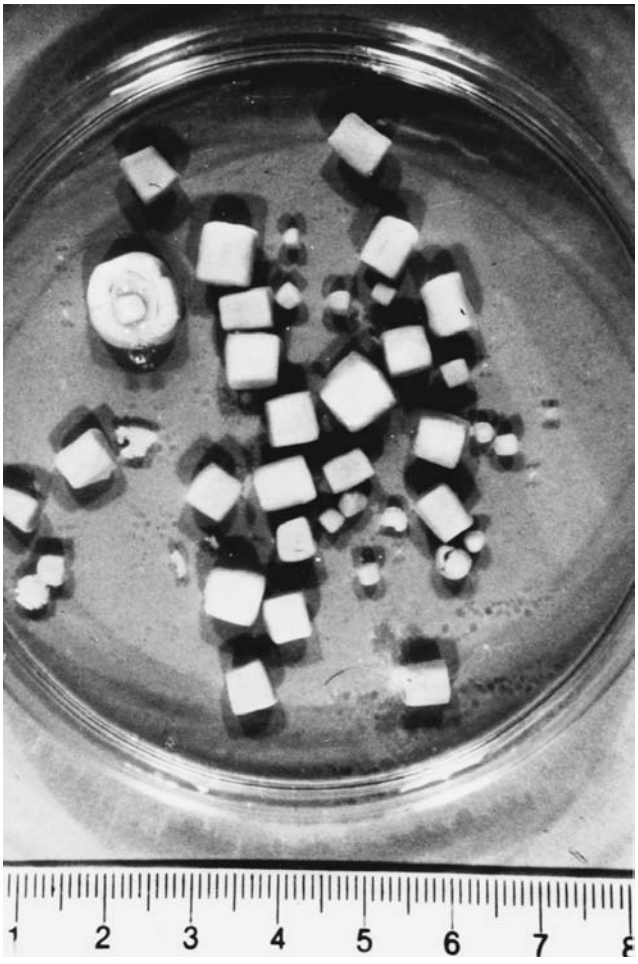


Figure 10. Flocs produced by a strain of *S. cerevisiae*; picture taken shortly after stopping agitation (scale in cm).

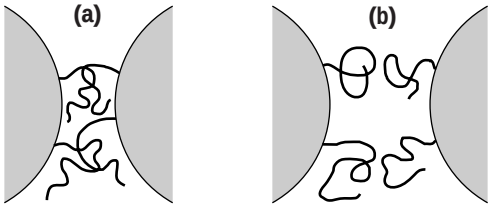


Figure 11. Schematic illustration. (**a**) Yeast surfaces weakly bound by entangled polymer chains upon agitation. (**b**) The same surfaces repelling each other in resting conditions.

istry. In a review (54), 33 genes involved in yeast aggregation have been reported; these include structural genes, regulatory genes, and genes involved in cell wall synthesis and mitochondrial functions. Four classes of strains were distinguished: *(1)* strains lacking an intact flocculation gene, *(2)* strains in which intact flocculation genes are constitutively repressed, *(3)* a strain with flocculation genes that are transcriptionally regulated, *(4)* strains with genes that encode for constitutive flocculence. For strains in which flocculation was regulated, induction was ascribed to nutrient limitation (46,55).

The best-studied dominant flocculation gene, *FLO1*, has been cloned and sequenced. Based on the genetic sequence, it was hypothesized that the Flo1 protein is present in the cell wall, and its N-terminus is exposed to the medium. The location of this protein in the cell wall was confirmed by the use of immunofluorescent probes. It was suggested that the exposed part of the Flo1 protein binds directly with neighboring cell wall mannans (54). The *FLO8* gene was also cloned (56); no homology was found with the *FLO1* gene. It was suggested that *FLO8* does not encode a cell wall protein, but is involved in the regulation of the transcription of the *FLO1* gene.

Two types of surface proteins involved in flocculation of bottom-fermenting strains were isolated from cell walls. A mannose-specific calcium-dependent agglutinin was isolated from a nonflocculent strain and from flocculent strains, both in nonflocculating and flocculating growth stages; it was released during the flocculation assay (57). A protein isolated from cell walls of flocculent cells only, called flocculin, showed homology to the protein coded by the flocculation gene *FLO1*; it was associated with fimbriae-like structures, did not show lectinlike activity, and was thought to act as a lectin ligand (58). The authors proposed a model (57) in which at least three factors determine the onset of flocculation: *(1)* release of agglutinin, *(2)* appearance or activation of flocculin, and *(3)* formation of fimbriae-like structures accompanied by an increase in cell-surface hydrophobicity.

For NewFlo phenotype *S. cerevisae* strains, there is evidence that lectins are synthesized continuously from an early stage of growth and rapidly inserted into the cell wall, but remain inactive until they are activated at flocculation onset (59).

Toward a Comprehensive View of the Role of Specific and Nonspecific Interactions

A comparative study of nine brewing strains flocculating in laboratory conditions (three bottom- and six top-fermenting strains) was performed (41,45). Flocculation most often occurred only in the stationary growth phase, and the onset of flocculation could not be related to a change of physicochemical properties (chemical composition by XPS, electrophoretic mobility, water contact angle).

Laboratory tests showed that flocculation of bottom strains requires the addition of calcium, is inhibited with mannose and sucrose, but not with galactose, and is not affected by ethanol. This is typical lectin-based behavior and the NewFlo/GMS phenotype. The following flocculation mechanism may be proposed: *(1)* there would basically be a repulsion between cell surfaces due to electrostatic or steric interactions (mobile polysaccharides); *(2)* in the stationary growth phase, lectins would appear at the cell surface or be made accessible; and *(3)* when the sugar concentration in the solution is sufficiently low, these lectins would bind sugar residues of adjacent cells across the repulsion barrier. The onset of flocculation thus seems to be triggered by the conjunction of the appearance of active lectin at the surface and of sugar attenuation, rather than by the calcium concentration (41,60). The occurrence of flocculation about 2 pH units above the isoelectric point in

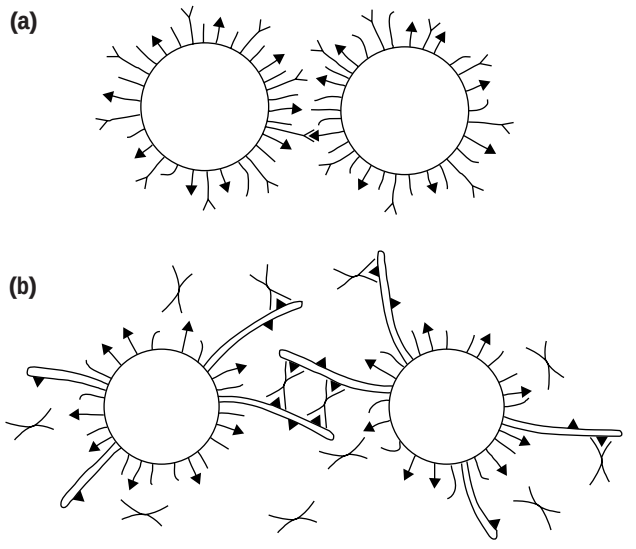


Figure 12. Schematic representation of possible cell–cell binding processes involving a lectinlike mechanism. (**a**) Direct lectin–ligand binding. (**b**) Implication of surface hairs and agglutinins (double forks).

laboratory tests, and not at the isoelectric point, indicates a major influence of pH on lectin conformation. The ionic strength has no effect on flocculation, which may be explained either by a minor role of electrostatic interactions or by a compensation of the influences of ionic strength on electrostatic interactions and lectin conformation. The penetration of the repulsion barrier will be more likely if the lectins or the lectin–ligands stick out of the surface, which is especially the case if they are located on surface hairs. Furthermore, flocculation might be favored by the release of agglutinins. The possible action of hairs and agglutinins is schematically illustrated by Figure 12b.

Flocculation of top-fermenting strains is induced by ethanol and other alcohols and does not require addition of calcium. For certain strains, there is no inhibition by mannose, sucrose, or galactose, whereas flocculation of other strains is inhibited by mannose only. The former would belong to the MI (mannose-insensitive)-flocculation phenotype, which has not been frequently reported. For this collection of top strains, there is no evidence of lectinlike interactions. Other organic solvents (methanol, *n*-propanol isopropanol, *n*-butanol, isopropanol, acetone, 1,2-propylene glycol) have the same effect as ethanol when added at a concentration provoking the same decrease of the surface tension of water. This indicates that the influence of ethanol is related to an adsorption process. The following flocculation mechanism may be proposed. There would exist a repulsion between the cells. As exponential cells, which have similar physicochemical properties as stationary cells, do not flocculate near the isoelectric point, the repulsion would be steric in nature. However, the influence of ionic strength on flocculation suggests that electrostatic interactions play a role as well. In the stationary growth phase, additional macromolecules or hairs would appear at the surface, providing binding sites for specific (lectinlike or other) or nonspecific binding. In the presence of sufficient ethanol concentration and, for the mannose-inhibited strains, sufficiently low sugar concentrations, these macromolecules or hairs would protrude from the surface and manage to bind adjacent cells across the repulsion barrier. The onset of flocculation thus seems to be triggered by the conjunction of a change of cell surface and an increase of ethanol concentration.

FLOCCULATION IN DOWNSTREAM PROCESSING

Production of fine chemicals by living cells involves a range of downstream processes, the first of which is usually the separation of cells or cell debris from the broth. This step is largely improved if flocculation can be carried out (61–63). Flocculation, as a pretreatment for the continuous centrifugal separation of whole cells, should produce flocs with high mechanical strength. In the production of microalgae, the best practical and economical way of harvesting was found to be chemical flocculation, followed by sedimentation, flotation, or filtration (64). The following sections review how flocculation can be produced in downstream processes and, to a certain extent, can be understood in the light of physicochemical theories described earlier.

Use of Colloidal Particles

S. cerevisiae forms flocs in the presence of positively charged submicrometer-sized aluminium hydroxide particles. The respective surface charges of yeast and aluminium hydroxide prevent homocoagulation and the process is a heterocoagulation of oppositely charged particles (65), which can easily be understood in the light of DLVO theory (discussed earlier).

Positively charged organic-polymer particles [poly(styrene-divinyl benzene) with quaternary ammonium surface groups; 0.27-μm diameter] were used to obtain flocs of *S. cerevisiae* and *E. coli* whole cells or cell debris. The settling time of the cells was considerably decreased in the presence of the particles, and the centrifugation energy and power required for the removal of *E. coli* cell debris were reduced by one order of magnitude. The optimal polymer particle concentration found for the clarification of suspensions of *E. coli* cells may be related to a partial coverage of the cell surface by the polymer particles, allowing the cells to be bound together by one layer of positively charged particles; complete coverage of every cell would lead to charge reversal and overall repulsion (see "Charge Neutralization by Cationic Polymers"). An optimal polymer-particle concentration was also found for the clarification of *E. coli* cell-debris suspensions. In the recovery of several enzymes by removal of cell debris following this method, complete enzyme activity was recovered in the supernatant, indicating that the polymer particles did not adsorb or inactivate the enzymes (61).

Use of Salts

Simple electrolytes favor particle coagulation by compressing the electric double layer, as explained in the section "DLVO Theory". Metal ions of higher charge are more efficient both because they have a higher tendency to adsorb

in the Stern layer and because they give a smaller Debye length. Moreover, other factors affect the efficacy of hydrolyzing metal ions, such as Al^{3+} and Fe^{3+}, which are used in water and sludge conditioning (66). Upon hydrolysis, these metal ions form positively charged polynuclear hydroxo complexes. The latter adsorb specifically and are able to reverse the surface charge. Therefore, depending on the pH and on the concentration of colloidal particles, the Al^{3+} and Fe^{3+} salts may act as coagulants only in a narrow concentration range. The concentration must be high enough to allow compression of the electric double layer or surface neutralization; however, overdosing may reverse the charge of the colloids and restabilize them. Above a certain concentration of metal ions, colloidal particles may again form flocs by enmeshing in aluminium or ferric hydroxide precipitates.

Jiang et al. (67) tested the ability of ferric and aluminium salts to coagulate algae. Hydrolized ferric species produced by oxidation of ferrous sulfate were more efficient than ferric sulfate, aluminium sulfate, or partially neutralized aluminium chloride. Floc size was larger and cell removal higher. Humic substances added to the algae suspension reduced the performances of the coagulants.

Use of Soluble Polymers

Nature and Efficiency of the Polymers. Both synthetic and biological polymers have been tested to induce cell flocculation, and results can be understood in the light of the principles presented earlier. Nonionic and anionic synthetic polymers were found to be ineffective for flocculation of *E. coli, S. cerevisiae*, and microalgae, while cationic polymers were good flocculating agents (61,64,68–70). Increasing the charge density of cationic polyacrylates and polyacrylamides enhanced flocculation (62,70). Cationic polyethyleneimine (PEI) was effective to flocculate *Chlorella ellipsoida* (68). With low molecular mass polymers, no reversal of the cell-surface charge occurred and the amount (mass) required to initiate flocculation decreased considerably as the molecular mass increased from 800 to 2,000 Da. Above 2,000 Da, the amount required to initiate flocculation did not vary much as a function of molecular mass; at high concentrations the polymer reversed the cell-surface charge; when the polymer concentration exceeded an optimal value, the flocs began to redisperse. Earlier work with silica and *E. coli* showed that an increase of pH from 4 to 6 or 9 increased the PEI concentration required for optimum flocculation, as expected from the variation of polymer ionicity. With *C. ellipsoida*, little effect of pH was observed between pH 4 and 7. The polymer concentration needed for flocculation was one order of magnitude larger for *E. coli*, compared to crystalline silica (68). With diethylaminoethylmetacrylate acrylic acid copolymers, washed cells of *E. coli* flocculated at polymer concentrations at which the isoelectric point was reached, indicating that the mechanism of flocculation was in agreement with the patch model (70).

Downstream processes are large consumers of polyacrylamide derivatives because they are economical and effective flocculating agents. However, neurotoxic and strong carcinogenic activities of acrylamide monomer were reported. Hence, there is a need to produce safe, biodegradable compounds that will minimize environmental and health risks. Several microorganisms have been reported to produce yeast-cell flocculating agents; they were effective between pH 3 and 5 and lost their effectiveness above pH 5 (71). Cultures of *Rhodococcus erythropolis* were found to flocculate *E. coli, Microcystis aeruginosa*, yeast, activated sludge, kaolin clay, and other particles. Flocculation was considerably improved by adding Ca^{2+} (72). However, the respective roles of *R. erythropolis* cells, possibly provoking heterocoagulation, and of polymers released into the suspension or at the cell surface, were not clarified. A *Paecilomyces* culture filtrate, containing a polysaccharidic compound, was shown to flocculate several microorganisms, including *Bacillus subtilis* and *E. coli*. Its effectiveness with *E. coli* did not change significantly between pH 4 and 8, and in an ionic strength range of 0.001 M to 0.1 M, except in the presence of CO_3^{2-} and Fe^{3+} ions. However, it decreased steeply below pH 4 and above pH 8, or at ionic strength of 1 M (73).

Compared to synthetic polymers, chitosan, also commercially available, possesses several advantages: monomers are not toxic, it can be heat-sterilized without degradation, and it is biodegradable. Chitosan, a partially *N*-acetylated poly (glucose-amine), is derived from chitin, the major constituent of crustacean exoskeleton (62). At acidic pH, chitosan is soluble and exists as well-extended linear chains due to repulsion of charged $-NH_3^+$ groups. At alkaline pH, the positive charges gradually disappear (the neutralization point is about pH 7.9), and chitosan tends to coil and precipitate (74).

Chitosan-induced flocculation of washed *E. coli* cells at concentrations much lower than those required to reach the isoelectric point, as opposed to the use of diethylaminoethylmetacrylate acrylic acid copolymers. This was attributed to the rigidity of chitosan, which undergoes weak deformation during adsorption and is able to form bridges between the cells (70). The optimal dose for flocculation of microalgae was about twice as low for chitosan as compared with polyacrylamides; this superiority was also attributed to the rigid backbone of chitosan (64). With washed cells of *B. subtilis* and *E. coli*, the effectiveness of chitosan and cationic polyacrylamides did not vary between pH 4 and 7. Above pH 7, the effectiveness of chitosan dropped due to a decrease of ionicity, whereas the change of polyacrylamide effectiveness was less marked. None of these compounds was effective in flocculating *Zymomonas mobilis* (62).

An optimal chitosan concentration was found near 0.2 $g\ L^{-1}$ for flocculation of *Euglena gracilis* cultures. Addition of chitosan above this concentration led to a partial cell redispersion due to repulsion of positively charged cells. With the optimal chitosan concentration, a sharp pH optimum was observed at pH 7.5, the decrease of flocculation at lower and higher pH being attributed to a lower negative charge of *Euglena* cells and to deprotonation of chitosan $-NH_3^+$ groups, respectively. At pH 7.5, flocculation did not decrease appreciably at high chitosan concentration, due to the low ionicity of the polymer at this pH (74).

Using chitosan, cell debris and nucleic acids could be separated selectively from β-D-galactosidase in *E. coli* dis-

integrates. For a highly charged chitosan, a pH decrease induced a reduction of the optimal flocculant dosage (63). Highly charged chitosan can achieve a more efficient removal of cell debris than less-charged chitosan. The required flocculation dosage of highly charged chitosan was relatively independent of the charge density, but increased with the molecular mass. At the required flocculation dosage, electrophoretic-mobility measurements showed that the flocs formed were not completely neutralized. It was therefore suggested that the main mechanism of flocculation was a "nonequilibrium" bridging process (75). This occurs when reconformation following polymer attachment to the surface is slow compared with polymer attachment or particle–particle collision (76). It is especially pronounced with high molecular mass polymers and leads to trapping of surface-negative sites between the bridges formed by the polycations. Highly charged and high molecular mass chitosan produced large and shear-resistant aggregates suitable for filtration, settling, and centrifugation, whereas highly charged low molecular mass chitosan gave smaller and more shear-sensitive flocs suitable only for centrifugation (75).

Influence of Medium Composition and Cell-Produced Polymers. Ionic strength can dramatically affect the flocculant effectiveness. Bilanovic et al. (64) studied the effect of salinity on microalgae flocculation with cationic polyacrylamides (1×10^7 and 2×10^7 Da) and chitosan (4×10^5 Da) at pH 7. For the two types of polycation, removal of more than 90% was obtained with freshwater cultures of *Chlorella*, whereas removal was very poor for cultures of marine species of the same genera. The efficiency of the cationic polymers to flocculate *Isochrysis galbana* dropped considerably as the ionic strength increased from 0.02 to 0.70 M, without appreciable alteration of the optimal dose. While the rise of ionic strength was expected to increase the screening of polymer–cell surface attraction, it also provoked a shrinkage of the macromolecule, as revealed by a decrease of the intrinsic viscosity.

The flocculation is also strongly influenced by the extracellular polymers (ECP) released by the cells in the solution and by the components of the culture media. The optimal dosage of cationic polymers was lower for washed *E. coli* cells than for the initial culture suspension. The interaction between anionic polymers in the medium and added cationic polymers led to an increase in flocculant requirement (69,70). On the other hand, it has been reported that the strongly acidic polymers released in solution by *Lactobacillus delbruekii* and *Lactobacillus fermentum* act as flocculating agents (strong anionic polyelectrolytes) that only need a small amount of cationic substances to be sensitized (69).

Polyacrylamides could flocculate *E. coli* and *B. subtilis* cells in defined media, but were ineffective in complex media because of precipitation. Chitosan did not precipitate below pH 7.5 in a complex medium and allowed flocculation of *E. coli* cells, while requiring twice the dose used in a defined medium; the flocculant efficiency decreased above pH 7 in both media and washed cells. On the other hand, flocculation of *B. subtilis* by chitosan in complex media occurred at pH 4 but dropped as the pH was increased from 4 to 7, and did not change with washed cells (62).

Flocculation of yeast by chitosan in a complex medium occurs simultaneously with the precipitation of chitosan with components of the medium. These can be regarded as competitive processes with respect to the availability of chitosan. The detrimental effect of precipitation (increase of dosage, contamination of biomass, removal of the fermentation product) can be minimized by a rapid addition and efficient mixing of the flocculant with fermentation media of high biomass concentration (77). Cells may also become entrapped in the precipitate of polyelectrolyte complexes, with the unexpected benefit that strong flocs are formed (78).

The effectiveness of flocculation is also dependent on the cell-surface constituents but is not simply related to the cell wall type (Gram positive or Gram negative). Polyacrylamide and chitosan were effective in flocculating *E. coli* (Gram negative) and *B. subtilis* (Gram positive) washed cells (62). Their inefficacy to flocculate washed cells of *Z. mobilis* (Gram negative) was attributed to the presence of nonionic extracellular polymers attached to the cell surface. It was reported that strongly acidic polymers adsorbed on *L. delbruekii* and *L. fermentum* surfaces induced an increase in the flocculant dosage, probably due to a high charge density and, therefore, a high polycation demand (69).

Combined Use of Flocculating Agents

The combined use of two flocculating agents has often proved to be an effective way to form flocs. Chitosan, polyethyleneimine, and Ca^{2+} were used simultaneously to perform flocculation of *E. coli* cell debris (79). This was applied in a process using a continuous centrifugal separator and allowed an at least 10-fold improvement in cell debris removal, compared with the same process without flocculation.

Sewage sludges were successfully flocculated by the addition of a cationic, followed by an anionic, polymer (80). The recovery of bacterial extracellular proteinases obtained in *B. subtilis* culture was investigated by adding bentonite (a clay with a negatively charged surface), followed by a cationic aminoamidic polyelectrolyte. The yield of recovery was more than 84% at the laboratory scale and up to 82% at the pilot-plant scale (81).

Polyelectrolytes used singly (62), or combined with an oppositely charged polymer (82), were inefficient in flocculating *Z. mobilis* in complex medium. However, effective flocculation could be performed by adding highly charged polyacrylamide or chitosan, followed by bentonite, providing high broth clarities between pH 5 and 8.5 or between pH 4 and 8, respectively. It was suggested that a combination of electrostatic interactions and bridging capability was responsible for the higher effectiveness of chitosan at low pH. Addition of bentonite after chitosan allowed avoidance of precipitation of the latter at high pH. Flocculation effectiveness was also influenced by the floc collision during settlement (82).

Use of Specific Interactions

Flocculation can be controlled by genetic factors, as discussed for brewing yeast earlier. Addition of starch to *E.*

coli suspension resulted in the formation of bacterial flocs, allowing a very efficient cell-broth separation by settling (83). Flocculation was dependent on the high expression of a cell-surface maltoporin receptor for starch: floc formation was evident with bacteria containing wild-type maltoporin, but was faster and occurred to a greater extent with strains expressing a high-affinity allele of the starch receptor or with strains grown at high levels of maltose. Because expression of the maltoporin receptor can be genetically controlled, its presence at the cell surface can be made largely independent of the variations in surface properties, thus avoiding the variability of classical flocculation. The optimal starch concentration was 1–2 mg mL^{-1} for a strain with a high-affinity allele. Decrease of flocculation at higher concentrations was attributed to saturation of binding sites. The optimal starch concentration was 5–10 mg mL^{-1} for a strain with a low-affinity allele. The optimal conditions for floc formation were pH 6–7 and a cell concentration of 2×10^9 cells mL^{-1}, which should be readily achievable in both laboratory and large-scale cultures. It was suggested that the presence of maltodextrin and that the lack of accessible α-1,4-linked chains in the starch preparation could decrease the starch effectiveness.

This starch-dependent aggregation is potentially applicable to other microorganisms if they comply with the following requirements: *(1)* the starch receptors are expressed, *(2)* the cell surface is not masked by a polysaccharide or a protein layer, and *(3)* the cell does not produce extracellular amylolytic enzymes. For harvesting bacteria, starch offers several advantages: *(1)* binding to cells is rapid; *(2)* it is nontoxic; *(3)* unlike electrically charged flocculating agents, it does not affect the pH, and its binding is effective in the pH range usually found in fermentation conditions; *(4)* it is a low-cost material; and *(5)* it is easily separable from bacteria, if necessary (83).

FLOCS AS A STATE OF PROLIFERATING BIOMASS

A biomass proliferating in the form of flocs is a system that is much more complex than flocs produced in a given downstream operation, due to the diversity of interfaces with respect to space and time. Brewing yeast, discussed earlier, illustrates the fact that a floc is the result of the integration of different processes and of their development on the time scale. In this section, even more complex systems of practical relevance are considered. Whenever possible, the presentation points out physicochemical features that provide a guide for understanding and controlling the biosystem behavior.

Continuous Ethanol Fermentation

The ability to flocculate is an important characteristic in the continuous production of ethanol, so that a high concentration of cells can be maintained within the fermentor (84). This has been used in upflow tower fermentors, which were successfully implemented at industrial scale (85). The yeast producing the particular flocs shown in Figure 10 was used in a gaslift laboratory fermentor (86). A more complete understanding of the phenomena controlling both the shape and size distribution of the flocs appeared to be a prerequisite for a full control of the highly promising fermentation process.

Continuous ethanol fermentation has also been performed with *Zymomonas mobilis*. In a laboratory upflow fermentor (volume of 0.92 L), flocs were obtained by starting-up (10 days) at a low dilution rate (0.065 h^{-1}) (87). After this period, the dilution rate was incrementally increased to 2.5 h^{-1}, allowing an ethanol productivity as high as 100 g $L^{-1}h^{-1}$. However, no attempt was made to evaluate long-term performance. A simultaneous addition of cationic polyacrylamide and titanium hydrous oxide was used to flocculate *Z. mobilis* (88). This provided an ethanol productivity of 16.4 g L^{-1} h^{-1} for at least 100 h in a gaslift tower fermentor (volume of 5 L). In similar conditions, productivity of unflocculated cells was evaluated to be about 2 g L^{-1} h^{-1}.

Mixed cultures of flocculent strains of *Z. mobilis* and *Saccharomyces* sp. gave higher ethanol production, yield, and volumetric productivity than pure cultures (84).

Activated Sludge

A common biological treatment of sewage waste is the activated sludge process, which is essentially a way to concentrate the organic matter of the waste. In its simplest form, it is a two-stage operation involving aeration followed by sedimentation. The aeration stage allows proliferation of different organisms and floc formation. An essential requirement for successful operation is the formation of settleable flocs (89). Among the physical characteristics of the flocs, the size distribution and structure are most important (90–94).

Influence of Physicochemical Factors. The floc behavior can be affected by hydrophobic interactions. Zita and Hermansson (95) isolated, by hydrophobic interaction chromatography, hydrophilic and hydrophobic bacteria from flocs coming from an aeration tank of a wastewater treatment plant. Attachment of these strains to activated sludge liquor containing flocs and wastewater was evaluated. The results showed a correlation between adhesion and surface hydrophobicity up to a threshold, above which bacterial attachment did not increase with surface hydrophobicity. This was attributed to an equilibrium between reversibly attached cells and free-living cells.

Urbain et al. (96) showed that a high internal hydrophobicity of the flocs is correlated with a decrease of the sludge volume index (SVI), which reveals an improvement of settleability. The internal hydrophobicity of the floc was determined by a salt aggregation test performed on a sonicated sludge sample mixed with increasing concentrations of $(NH_4)_2SO_4$ in phosphate buffer.

Bacterial strains of different hydrophobicity were separated by subjecting sonicated sludge to an octane–water mixture. The amount of exopolymers released by sonication of a hydrophilic strain cultured in the laboratory was greater than that of a more hydrophobic strain. This observation may suggest that the hydrophilic strains are not simply trapped in a matrix, but are also implicated in the floc organization (97).

No clear relationship has been established between the electrophoretic mobility of sludge particles and their settling properties. A strongly negative electrophoretic mobility of sludge particles was found to induce a high SVI (98,99). However, an inverse correlation (100) and the absence of correlation between electrophoretic mobility and SVI were also reported (101).

Floc stability was evaluated by gentle mixing in $CaCl_2$ and KCl solutions of increasing ionic strength (102). Floc dissociation decreased as the ionic strength rose from 50 μM to 5 mM; it decreased more slowly between 5 and 50 mM. This was explained by DLVO theory, the flocs being stronger as the diffuse double layer was more compressed. The reversibility of the process was tested by reflocculation of degraded activated sludge. Reflocculation increased between 0 and 5 mM ionic strength and remained constant up to 50 mM. Between 50 and 500 mM, floc dissociation increased and reflocculation decreased. According to the authors, at high ionic strength, the distance between particles is fairly constant and the floc structure appears to be influenced by mechanisms of unknown nature, unlike the mechanisms considered in DLVO theory. The influence of wastewater on floc stability was similar to that described for salt solutions.

Polyvalent ions can act by binding ECP and bacteria, or by binding bacteria with one another, through electrostatic interactions (90). The addition of Ca^{2+} to activated sludge induced a decrease of both the sludge volume and the residual turbidity in the supernatant after 30 min of sedimentation (103). Removal of Ca^{2+} from activated sludge by ion exchange (H^+, K^+, Na^+, Mg^{2+}) or by a chelating agent (EGTA) induced a sludge deterioration: small particles appeared in the supernatant, causing an increase of the specific resistance to filtration. On the contrary, exchange of Ca^{2+} by Cu^{2+} decreased the specific resistance to filtration and thus resulted in a more stable floc structure. This was attributed to the ability of Cu^{2+} ions to maintain the 3-dimensional structure of the exopolymers (90).

Influence of Biological Factors. Bioflocculation in activated sludge also involves extracellular polymers (ECP). Among the ECP, proteins (96,97,104,105), polysaccharides (96,97,104), lipids (104), humic acids (105), and DNA (96,97) were reported. Humic acids originate from wastewater itself, whereas other compounds arise from cell metabolism or from cell lysis. The influence of ECP on sludge settling (notably on SVI) has been widely studied. However, because there is no unified method of extraction (e.g., centrifugation with or without heating and/or sonication), care must be taken in comparing results from different studies. A general trend found in the literature is that an increasing concentration of ECP extracted from activated sludge is related to an increasing SVI (96,97,100,105); however, the absence of a relationship between ECP production and SVI has also been reported (101).

ECP were isolated after thermal extraction from sludge and their molecular mass was determined by gel filtration. A major fraction was found between 30 and 100 kDa, and an increase in its apparent molecular mass was correlated with an increase of the stirred specific volume index (SSVI). This could be due to steric interactions, allowing greater amounts of interstitial water to be retained in the sludge (106).

The presence of growing filamentous microorganisms is an important feature determining floc properties. It was suggested that they could form a backbone to which floc-forming bacteria could attach (107). This idea is supported by the presence of filamentous microorganisms in numerous activated-sludge samples (96). However, sludges have poorer settling characteristics when filaments grow out of the flocs. The outgrowth was quantified by the total extended filament length, that is, the sum of the lengths of filaments extending out of the flocs. An increase of SVI (94) or of SSVI (108) was observed as the total extended-filament length increased up to 10 m mg^{-1} of suspended solid. When the total extended-filament length exceeded this value, SVI increased dramatically and bulking phenomena were observed (101).

Microbial Aggregates in Anaerobic Wastewater Treatment

Anaerobic processes, resulting in the production of biogas, are used for the decomposition of organic compounds in wastewater; the main problem encountered is to efficiently retain the biomass in the reactor. Several reactor designs have been considered. Bacteria may be immobilized on a packing material or carrier particles; the ability of bacteria to aggregate can be used to produce granules with good settling properties. This latter is the case of the upflow anaerobic sludge blanket reactor (UASBR) (109). Bacterial aggregation ensures the conservation of the biomass inside the reactor, but also enhances the interactions between the various microorganisms needed for anaerobic digestion (110).

The formation of bacterial granules in UASBR is governed by the surface properties of the bacteria, which may themselves depend on the operational parameters. Hydrophobicity is the surface property that has been the most studied for bacteria originating from granules. Contact-angle measurements on cells found in the reactor effluent and on cells originating from granules led to the conclusion that more hydrophobic bacteria remain in the reactor in the aggregated state, whereas more hydrophilic bacteria are isolated and leave the reactor (111). It was also reported that the more hydrophobic bacteria found in UASBR had a better ability to flocculate (112). Moreover, the ECP produced by flocculent bacteria were more hydrophobic than those produced by nonflocculent bacteria. Different populations of bacteria found in UASBR are responsible for different steps of waste digestion and are classified as acidogens, acetogens, and methanogens. Most acidogens found in granules are more hydrophilic than the acetogens and methanogens; moreover, granules are formed of layers of different hydrophobicities, each layer being responsible for one step of the degradation process (113).

The surface charge of the bacteria is also recognized as a parameter governing the flocculation behavior in UASBR (109). The negative charge is attributed to carboxyl groups located on the ECP; it is indeed reduced when the cells are submitted to shear (112).

Divalent or trivalent metal ions, which can be added to the reactor, may influence granule properties by acting on cell–cell interactions, as described by DLVO theory, but also by affecting the conformation of ECP or by cross-linking the ECP to form a kind of matrix. It was shown that the size of the granules, and thereby their settleability, is significantly increased in the presence of Ca^{2+} (100 mg L^{-1}), which leads to shorter retention times in the reactor (111). When EDTA, a chelating agent, was added to the medium, the aggregation decreased but was not suppressed. It was thought that part of the Ca^{2+} was involved in irreversible bonds between the cells. The flocculation could also be induced in an anaerobic digestor by the addition of various divalent and trivalent cations, and the tendency to flocculate was related to a reduction in the negative charge of the cell surface. However, divalent and trivalent cations may also play a role through other mechanisms. Precipitates of calcium or iron may provide the bacteria with nucleation sites for granule formation. An alternative mechanism is by an increase in surface hydrophobicity provoked by complex formation between divalent cations and carboxyl or phosphate groups (112).

Other operational parameters may influence granule formation or the efficiency of the process. When the substrate is rich in protein, highly hydrophilic cells are favored, which leads to the formation of fluffy flocs rather than granules. On the contrary, a substrate rich in volatile fatty acids produces granules with a very hydrophobic surface; this may be detrimental to the process because such granules may be eliminated from the reactor by flotation (113). Lowering the surface tension of the medium by adding alkylbenzenesulfonate was shown to favor the formation of granules with a more hydrophilic surface, thus preventing their flotation (114).

CONCLUSION

Cell aggregation, which is of considerable importance in bioprocess engineering, can be understood and controlled to a certain extent in the light of established principles of colloid and surface chemistry. A major key is the interplay between electrostatic interactions and interactions involving macromolecules. In this respect, microbial cells represent a complex system that can be regarded as a solid surface, a surface grafted with macromolecules, or a surface with adsorbed macromolecules, depending on the strain and the approach followed. Furthermore, the extracellular polymers may constitute a matrix in which the cells are more or less loosely entrapped. For certain microorganisms, another important mechanism is provided by specific interactions between the cells, or between the cells and foreign substances.

In opposition with this complexity, most of the methods used so far to examine the physicochemical properties of microbial surfaces consider them to be, more or less implicitly, homogeneous and molecularly smooth, although key factors are clearly the spatial distribution of constituents and properties. Atomic-force microscopy is undoubtedly a very promising tool in this respect. In the near future, it may provide direct information, with molecular-scale spatial resolution, as regards chemical composition, physicochemical properties, and interactions of cell surfaces, in conditions relevant to practical interest. This constitutes an exciting field in which a strong synergy may develop between biotechnology and nanotechnology.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of the Foundation for Training in Industrial and Agricultural Research (FRIA), the National Foundation for Scientific Research (FNRS), the Department of Education and Scientific Research (Concerted Action, Physical Chemistry of Interfaces and Biotechnology), and the Interuniversity Poles of Attraction Programme (Federal Office for Scientific, Technical and Cultural Affairs, Supramolecular Chemistry and Supramolecular Catalysis).

BIBLIOGRAPHY

1. R.J. Hunter, *Foundations of Colloid Science*, vol. 1, Clarendon Press, Oxford, 1987.
2. G.B. Calleja, *Microbial Aggregation*, CRC Press, Boca Raton, Fla., 1984.
3. K. Marshall, *Microbial Adhesion and Aggregation*, Springer-Verlag, Berlin, 1984.
4. Y.A. Attia, *Flocculation in Biotechnology and Separation Systems*, Elsevier, New York, 1987.
5. W.G. Characklis and K.C. Marshall, *Biofilms*, Wiley, New York, 1990.
6. L.F. Melo, T.R. Bott, M. Fletcher, and B. Capdeville, *Biofilms—Science and Technology*, Kluwer Academic, Dordrecht, The Netherlands, 1992.
7. J. Wimpenny, W. Nichols, D. Stickler, and H. Lappin-Scott, *Bacterial Biofilms*, Bioline, Cardiff, U.K., 1994.
8. H.M. Lappin-Scott and J.W. Costerton, *Microbial Biofilms*, University Press, Cambridge, 1995.
9. J. Israelachvili, *Intermolecular and Surface Forces*, 2nd ed., Academic, London, 1992, pp. 176–311.
10. P.C. Hiemenz and R. Rajagopalan, *Principles of Colloid and Surface Chemistry*, 3rd ed., Marcel Dekker, New York, 1997.
11. J. Lyklema, *Fundamentals of Interface and Colloid Science*, vol. 2, Academic, London, 1995, pp. 4.1.–4.135.
12. S. Nir, *Progr. Surface Sci.* **8**, 1–58 (1976).
13. L. Eriksson and C. Axberg, *Water Res.* **15**, 421–431 (1981).
14. J. Israelachvili and H. Wennerström, *Nature* **379**, 219–225 (1996).
15. J. Lyklema, *Fundamentals of Interface and Colloid Science*, vol. 1, Academic Press, London, 1991, pp. 4.48–4.51.
16. F. Lafuma, K. Wong, and B. Cabane, *J. Colloid Interface Sci.* **143**, 9–21 (1991).
17. B. Vincent, in Th. F. Tadros ed., *Solid/Liquid Dispersions*, Academic, London, 1987, pp. 149–162.
18. F. Lafuma, in J.C. Roberts ed., *Paper Chemistry*, 2nd ed., Chapman and Hall, London, 1995, pp. 44–63.
19. V. Chaplain, M.L. Janex, F. Lafuma, C. Graillat, and R.E. Audebert, *Colloid Polymer Sci.* **273**, 984–993 (1995).
20. A. Guyot, R. Audebert, R. Botet, B. Cabane, F. Lafuma, R. Jullien, E. Pefferkorn, C. Pichot, A. Revillon, and R. Varoqui, *J. Chim. Phys.* **87**, 1859–1899 (1990).

21. J. Gregory, in Th. F. Tadros ed., *Solid/Liquid Dispersions*, Academic, London, 1987, pp. 163–181.
22. G. Durand-Piana, F. Lafuma, and R. Audebert, *J. Colloid Interface Sci.* **119**, 474–480 (1987).
23. J. Th. G. Overbeek, in H.R. Kruyt ed., *Colloid Science*, Elsevier, New York, 1952, pp. 58–89, 278–301.
24. J.C. Kihn, C.L. Masy, M.M. Mestdagh, and P.G. Rouxhet, *Can. J. Microbiol.* **34**, 779–791 (1988).
25. A.M. James, in N. Mozes, P.S. Handley, H.J. Busscher, and P.G. Rouxhet eds., *Microbial Cell Surface Analysis: Structural and Physicochemical Methods*, VCH Publishers, New York, 1991, pp. 221–262.
26. R.J. Hunter, *Zeta Potential in Colloid Science—Principles and Applications*, Academic Press, London, 1981, 386 p.
27. A. van der Wal, M. Minor, W. Norde, A.J.B. Zehnder, and J. Lyklema, *J. Colloid Interface Sci.* **186**, 71–79 (1997).
28. H.C. van der Mei, M. Rosenberg, and H.J. Busscher, in N. Mozes, P.S. Handley, H.J. Busscher, and P.G. Rouxhet eds., *Microbial Cell Surface Analysis: Structural and Physicochemical Methods*, VCH Publishers, New York, 1991, pp. 263–287.
29. R. Tybussek, F. Linz, K. Schügerl, N. Mozes, A.J. Leonard, and P.G. Rouxhet, *Appl. Microbiol. Biotechnol.* **41**, 13–22 (1994).
30. M.N. Bellon-Fontaine, N. Mozes, H.C. van der Mei, J. Sjollema, O. Cerf, P.G. Rouxhet, and H.J. Busscher, *Cell Biophysics* **17**, 93–106 (1990).
31. P.G. Rouxhet and M.J. Genet, in N. Mozes, P.S. Handley, H.J. Busscher, and P.G. Rouxhet eds., *Microbial Cell Surface Analysis: Structural and Physicochemical Methods*, VCH Publishers, New York, 1991, pp. 173–220.
32. Y.F. Dufrêne and P.G. Rouxhet, *Can. J. Microbiol.* **42**, 548–556 (1996).
33. P.G. Rouxhet, N. Mozes, P.B. Dengis, Y.F. Dufrêne, P.A. Gerin, and M.J. Genet, *Colloids Surfaces B: Biointerfaces* **2**, 347–369 (1994).
34. G. Binnig, C.F. Quate, and C. Gerber, *Phys. Rev. Lett.* **56**, 930–933 (1986).
35. M. Radmacher, R.W. Tillmann, M. Fritz, and H.E. Gaub, *Science* **257**, 1900–1905 (1992).
36. M. Stratford, *Adv. Microbiol. Physiology* **33**, 1–71 (1992).
37. D.E. Amory, J-P. Dufour, and P.G. Rouxhet, *J. Inst. Brew.* **94**, 79–84 (1988).
38. E.H. van Hamersveld, M.C.M. van Loosdrecht, R.G.J.M. van der Lans, and K.Ch.A.M. Luyben, *J. Inst. Brew.* **102**, 333–342 (1996).
39. P.B. Dengis, L.R. Nélissen, and P.G. Rouxhet, *Appl. Environ. Microbiol.* **61**, 718–728 (1995).
40. N. Mozes, F. Marchal, M.P. Hermesse, J.L. Van Haecht, L. Reuliaux, A.J. Leonard, and P.G. Rouxhet, *Biotechnol. Bioeng.* **30**, 439–450 (1987).
41. P.B. Dengis and P.G. Rouxhet, *J. Inst. Brew.* **103**, 257–261 (1997).
42. M. Stratford and M.H.J. Keenan, *Yeast* **4**, 107–115 (1988).
43. E.H. van Hamersveld, M.C.M. van Loosdrecht, and K.Ch.A.M. Luyben, *Colloids Surfaces B: Biointerfaces* **2**, 165–171 (1994).
44. E.H. van Hamersveld, R.G.J.M. van der Lans, P.J.C. Caulet, and K.Ch.A.M. Luyben, *Biotechnol. Bioeng.* **57**, 330–341 (1998).
45. P.B. Dengis and P.G. Rouxhet, *Yeast* **13**, 931–943 (1997).
46. M.H. Straver, P.C. v.d. Aar, G. Smit, and J.W. Kijne, *Yeast* **9**, 527–532 (1993).
47. M.H. Straver, G. Smit, and J.W. Kijne, *Colloids Surfaces B: Biointerfaces* **2**, 173–180 (1994).
48. G.G. Stewart, I.F. Garrison, T.E. Goring, M. Meleg, P. Pipasts, and I. Russell, *Kemia-Kemi* **10**, 465–479 (1976).
49. N. Mozes, L.L. Schinckus, C. Ghommidh, J.M. Navarro, and P.G. Rouxhet, *Colloids Surfaces B: Biointerfaces* **3**, 63–74 (1994).
50. B.L.A. Miki, N.H. Poon, A.P. James, and V.L. Seligy, *J. Bacteriol.* **150**, 878–889 (1982).
51. M. Stratford, *Colloids Surfaces B: Biointerfaces* **2**, 151–158 (1994).
52. M. Stratford and S. Assinder, *Yeast* **7**, 559–574 (1991).
53. C.L. Masy, A. Henquinet, and M.M. Mestdagh, *Can. J. Microbiol.* **38**, 1298–1306 (1992).
54. A.W.R.H. Teunissen and H.Y. Steensma, *Yeast* **11**, 1001–1013 (1995).
55. A.W.R.H. Teunissen, J.A. van den Berg, and H.Y. Steensma, *Yeast* **11**, 435–446 (1995).
56. O. Kobayashi, N. Hayashi, and H. Sone, *European Brewery Convention, Proc. 25th Congress, Brussels*, Oxford University Press, London, 1995, pp. 361–368.
57. M.H. Straver, V.M. Traas, G. Smit, and J.W. Kijne, *Yeast* **10**, 1183–1193 (1994).
58. M.H. Straver, G. Smit, and J.W. Kijne, *Appl. Environ. Microbiol.* **60**, 2754–2758 (1994).
59. M. Stratford and A.T. Carter, *Yeast* **9**, 371–378 (1993).
60. Ch.L. Masy, M. Kockerols, and M.M. Mestdagh, *Can. J. Microbiol.* **37**, 295–303 (1991).
61. C.W. Kim, S.-K. Kim, C. Rha, and E. Robinson, in Y.A. Attia ed., *Flocculation in Biotechnology and Separation Systems*, Elsevier, Amsterdam, 1987, pp. 429–439.
62. J. Hughes, D.K. Ramsden, and K.C. Symes, *Biotechnol. Techniques* **4**, 55–60 (1990).
63. I. Agerkvist, L. Eriksson, and S.-O. Enfors, *Enzyme Microb. Technol.* **12**, 584–590 (1990).
64. D. Bilanovic, G. Shelef, and A. Sukenik, *Biomass* **17**, 65–76 (1988).
65. G. Kayem and P. Rouxhet, *J. Chem. Soc. Faraday Trans.* **1. 79**, 561–569 (1983).
66. W. Stumm, in W. Stumm ed., *Chemistry of the Solid Water Interface*, Wiley, New York, 1992, pp. 253–278.
67. J.-Q. Jiang, N. Graham, and C. Harward, *Water Sci. Technol.* **27**, 221–230 (1993).
68. R. Tilton, J. Murphy, and J. Dixon, *Water Res.* **6**, 155–164 (1972).
69. L. Eriksson, and A.-M. Härdin, in Y.A. Attia ed., *Flocculation in Biotechnology and Separation Systems*, Elsevier, Amsterdam, 1987, pp. 441–455.
70. A. Baran, *Colloids Surfaces* **31**, 259–264 (1988).
71. J. Nakamura, S. Miyashiro, and Y. Hirose, *Agric. Biol. Chem.* **40**, 377–383 (1976).
72. R. Kurane, K. Takeda, and T. Suzuki, *Agric. Biol. Chem.* **50**, 2301–2307 (1986).
73. H. Takagi and K. Kadowaki, *Agric. Biol. Chem.* **49**, 3151–3157 (1985).
74. P. Gualtieri, L. Barsanti, and V. Passarelli, *J. Microbiol. Methods* **8**, 327–332 (1988).

75. I. Agerkvist, *Colloids Surfaces* **69**, 173–187 (1992).
76. E. Pelssers, M. Cohen Stuart, and G. Fleer, *Colloids Surfaces* **38**, 15–25 (1989).
77. S. Weir, D.K. Ramsden, J. Hughes, and F. Le Thomas, *Biotechnol. Techniques* **7**, 199–204 (1993).
78. S. Weir, D.K. Ramsden, J. Hughes, and F. Le Thomas, *Biotechnol. Techniques* **8**, 129–132 (1994).
79. I. Persson and B. Lindman, in Y.A. Attia ed., *Flocculation in Biotechnology and Separation Systems*, Elsevier, Amsterdam, 1987, pp. 457–466.
80. UK Pat. Appl, GB 2 082 163 A (August 7, 1981), M. Hashimoto, Y. Ishii and Y. Ohi (to Kuzita Water Industries, Ltd., Osaka, Japan).
81. V. Sitkey, E. Kuzmova, R. Horenitzky, P. Michalik, and M. Minarik, *Acta Biotechnol.* **12**, 299–302 (1992).
82. J. Hughes, D. Ramsden, K. Symes, and P. Williams, *Biotechnol. Techniques* **4**, 233–236 (1990).
83. T. Ferenci and K.-S. Lee, *Biotechnol. Bioeng.* **38**, 314–318 (1991).
84. C. Abate, D. Callieri, E. Rodriguez, and O. Garro, *Appl. Microbiol. Biotechnol.* **45**, 580–583 (1996).
85. K.H. Kwok and I.G. Prince, *Enzyme Microb. Technol.* **11**, 597–603 (1989).
86. A. Fontana, C. Ghommidh, J.P. Guiraud, and J.M. Navarro, *Biotechnol. Lett.* **14**, 505–510 (1992).
87. I.G. Prince and J.P. Barford, *Biotechnol. Lett.* **4**, 525–530 (1982).
88. C. Ghommidh and J.D. Bu'Lock, *Biotechnol. Techniques* **2**, 249–252 (1988).
89. J.A. Sodell and R.J. Sevior, *J. Appl. Bacteriol.* **69**, 145–176 (1990).
90. J.H. Bruus, P.H. Nielsen, and K. Keiding, *Water Res.* **26**, 1597–1604 (1992).
91. D. Li and J.J. Ganczarczyk, *Biotechnol. Bioeng.* **35**, 57–65 (1990).
92. F. Jorand, F. Zartarian, F. Thomas, J.C. Block, J.Y. Bottero, G. Villemin, V. Urbain, and J. Manem, *Water Res.* **29**, 1639–1647 (1995).
93. B.K. Mobarry, M. Wagner, V. Urbain, B.E. Rittmann, and D.A. Stahl, *Appl. Environ. Microbiol.* **62**, 2156–2162 (1996).
94. L. Eriksson, I. Steen, and M. Tendaj, *Water Sci. Technol.* **25**, 251–265 (1992).
95. A. Zitta and M. Hermansson, *Appl. Environ. Microbiol.* **63**, 1168–1170 (1997).
96. V. Urbain, J.C. Block, and J. Manem, *Water Res.* **27**, 829–838 (1993).
97. F. Jorand, P. Guicherd, V. Urbain, J. Manem, and J.C. Block, *Water Sci. Technol.* **30**, 211–218 (1994).
98. C.F. Forster, N.J.B. Knight, and D.A.J. Wase, *Adv Biotechnol. Processes* **4**, 211–240 (1985).
99. A.E. Steiner, D.A. McLaren, and C.F. Forster, *Water Res.* **10**, 25–30 (1976).
100. Y. Magara, S. Nambu, and K. Utosawa, *Water Res.* **10**, 71–77 (1976).
101. J.B. Barber and J.N. Veenstra, *J. Water Pollut. Control Fed.* **58**, 149–156 (1986).
102. A. Zitta and M. Hermansson, *Appl. Environ. Microbiol.* **60**, 3041–3048 (1994).
103. K. Kakii, S. Kitamura, T. Shirakashi, and M. Kuriyama, *J. Ferment. Technol.* **63**, 263–270 (1985).
104. J.A.S. Goodwin and C.F. Forster, *Water Res.* **19**, 527–533 (1985).
105. L. Eriksson and B. Alm, *Water Sci. Technol.* **24**, 21–28 (1991).
106. C.F. Forster, *Water Res.* **19**, 1259–1264 (1985).
107. M. Sezgin, D. Jenkin, and D.S. Parker, *J. Water Pollut. Control Fed.* **50**, 362–381 (1978).
108. L. Eriksson and A.-M. Härdin, *Water Sci. Technol.* **16**, 55–68 (1984).
109. N. Kosaric and R. Blaszczyk, *Adv. Biochem. Eng. Biotechnol.* **42**, 27–62 (1990).
110. W. De Zeeuw and G. Lettinga, *Antonie van Leeuwenhoek* **46**, 110–112 (1980).
111. E. Mahoney, L. Varangu, W. Cairns, N. Kosaric, and R. Murray, *Water Sci. Technol.* **19**, 249–260 (1987).
112. N. Kosaric, E. Mahoney, L. Varangu, and W. Cairns, *Water Pollut. Res. J. Can.* **22**, 289–297 (1987).
113. D. Daffonchio, J. Thaveesri, and W. Verstraete, *Appl. Environ. Microbiol.* **61**, 3676–3680 (1995).
114. J. Thaveesri, D. Daffonchio, B. Liessens, P. Vandermeren, and W. Verstraete, *Appl. Environ. Microbiol.* **61**, 3681–3686 (1995).

CELL SEPARATION, SEDIMENTATION

BHAV P. SHARMA
CV Therapeutics
Palo Alto, California

KEY WORDS

Cell density
Cell separation
Cell size
Gravity
pH
Reynolds number
Sedimentation
Sedimentation velocity
Stokes' law
Viscosity

OUTLINE

INTRODUCTION

By definition, the bioprocessing industry uses biological cells as little factories to make a variety of products. Once they are presumed to have finished their product synthesis step, the cells need to be separated from the liquid medium containing them. During the course of research involving biological cells or during the isolation and purification of products derived from cells, cell separation is like the midgame of a chess match. Correct moves at this stage provide an excellent chance to obtain a product of the desired characteristics.

At least five general approaches are available for cell separation: centrifugation, cyclone separation, sedimentation, filtration, and microfiltration. In addition, cell flocculation can often be used to enhance the efficiency of these separation techniques. Sedimentation as a tool for cell separation is the focus of this article. It is also noted how centrifugation and flocculation, too, draw on sedimentation theory.

SEDIMENTATION AS A TOOL IN PRACTICE

The starting point in selecting a cell separation strategy is to locate the product of interest in a given suspension of cells. Is it the cells themselves? Is it all the cells or a specific cell type? Is the interest actually in a product inside the cells (i.e., intracellular) or in a product outside the cells (i.e., extracellular)? If the product is intracellular, cell removal may need to be achieved without significant cell damage. If the product is extracellular, the liquid phase must be separated while the solid content of the liquid phase is minimized, and the cells along with the other solids are merely the solid waste to be removed and disposed of. If the cells *are* the product, they must be separated without damage and their functionality kept intact. Sedimentation may be a useful tool for any of these situations.

Sedimentation is relevant in a variety of biotechnology/bioprocessing situations:

1. Production of therapeutic or diagnostic proteins via mammalian cells
2. Achievement of high cell concentrations via cell recycle in continuous fermentors
3. Biological waste or water treatment
4. Production of beer
5. Study of different cell types involved in immunological response
6. Harvesting of cells in microcarrier or suspension cell culture

Table 1 gives specific examples of research in which sedimentation has been used as a cell separation tool (1–16). Sedimentation is also utilized extensively in other fields such as water or solids removal in mineral recovery and processing (17).

SEDIMENTATION EQUIPMENT

In processing applications, the sedimentation step is usually carried out in a typical processing tank. Such tanks are designed to have the capabilities of a feed line, an output or an overflow line, and a line to allow the removal of the sedimented material. However, equipment designs other than a processing tank have been shown to be very effective in special situations. One such design uses an inclined, parallel-plate sedimentation device (5,18,19) shown in Figure 1, which can be used for cell recycle in continuous fermentations or cell culture. The inclined sedimentor design not only allows the fermentation or the cell culture to achieve high cell densities, it also offers an efficient means of selective cell retention.

A CelSep apparatus meant primarily for mammalian cell separations at the 1-L scale was commercially available (Wescor, Inc., Logan, Utah). Although the unit is no longer on the market, the autoclavable unit worked well and allowed gravity separations under a continuous density gradient (20).

ASPECTS OF SEDIMENTATION AS A CELL SEPARATION TOOL

1. Sedimentation as a separation tool is a mild technique. This is important in cell biology, where the functional capabilities of the isolated subpopulation of cells must not be altered by the separation step.
2. Sedimentation lends itself to aseptic handling if needed. Although other techniques may allow cell separation without contamination, sedimentation devices for cell separation are typically easier to assemble, clean, and sterilize. Hence aseptic handling is more easily performed with sedimentation than with other options.
3. From a practical viewpoint, "unaided" sedimentation can be too slow a process if the sedimentation rates are low. In industrial-scale cell separations, where microbial growth or enzymatic degradation can be a problem, long processing times are generally not acceptable. Such situations call for centrifugation, which "enhances" the sedimentation rates via the centrifugal force, because the problems of microbial growth or enzymatic degradation are less dominant for short processing times.
4. In its simplest form, sedimentation may be the lowest-cost technique for cell separation.
5. Sedimentation is a simple and a highly reliable separation technique; unlike other options, such as centrifugation; it does not require extensive training of the user.

Overall, sedimentation principles have broad applicability in cell separations and extend beyond simple gravity sedimentation. Sedimentation principles also apply to centrifugation and ultracentrifugation.

SEDIMENTATION THEORY AND APPLICATION

To consider the practical issues in cell separation problems, it is useful to understand the principles behind grav-

Table 1. Some Examples of Sedimentation as a Cell Separation Tool in Research

Cell and cell culture type	Sedimentation application	Ref.
Mammalian, suspension	Sedimentation chamber as a separation device in a perfusion system for hybridoma cell culture for antibody production	1
Mammalian, microcarrier	Sedimentation separator pipe as a separation device for microcarrier cell culture	2
Mammalian, suspension	Gravitational sedimentation used to separate cells aggregated with the help of high-frequency, low-energy ultrasonic resonance fields	3
Bacteria	Sedimentation rates of cells of different shapes and sizes used to develop a method for enumeration and quantitation of bacteria	4
Yeast, *Saccharomyces cerevisiae*	Gravitation sedimentation used to recycle yeast cells for alcohol fermentation	5
Yeast, *S. cerevisiae* and yeast–mouse/mouse hybridoma cells	Sedimentation chamber used to separate cells from a chemostat culture	6
Starch endosperm of barley, *Hordeum vulgare L.*	Endosperm protoplast isolation method utilized gravity sedimentation to get good preparations	7
Spleen B-lymphocytes	Hybridoma-forming immune spleen B-lymphocyte cells separated based on their size and density using gravity sedimentation	8
Yeast, *Candida rugosa*	Sedimentation velocity measured to compare cells grown at different temperatures	9
Bacteria (*E. coli*) and yeast (*S. cerevisiae* and baker's)	Differences in flocculation and sedimentation properties utilized to achieve selective separation of cells carrying a certain plasmic	10,11
Sperm cells	A migration–sedimentation technique used to study whether it is a useful method to separate sperm cells leading to improved fertilization	12,13
Mammalian cells, erythrocytes	Utility of the erythrocyte sedimentation rate studied as a possible indicator of clinical disease or health	14,15
Sulfate-reducing bacteria, *Desulfovibrio desulfuricans*	Gravity sedimentation utilized for cell recycle in a system for microbial reduction of sulfur dioxide in an anaerobic digester	16

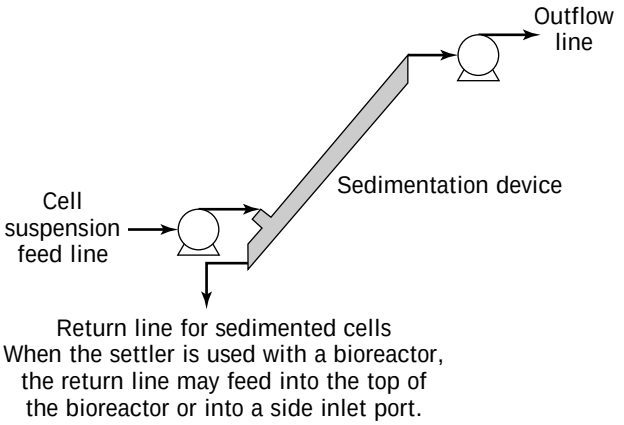


Figure 1. Conceptualized representation of an inclined, parallel-plate, cell sedimentation device used for cell recycle in a continuous fermentation or a perfusion bioreactor. When the settler is used with a bioreactor, the return line may feed into the top of the bioreactor or into a side inlet port (5,18,19).

itational sedimentation. Several factors affect the sedimentation of cells in a liquid medium:

- The *difference* between the density of a cell and that of the surrounding liquid medium
- The density of the liquid medium
- Cell size
- Cell shape
- Cell concentration
- The viscosity of the liquid medium

As a practical rule of thumb, the particle size needs to be greater than about 1 μm for gravitation sedimentation to be relevant as a separation tool. Other approaches such as centrifugation, ultracentrifugation, and membrane separation are usually needed for smaller particles.

The variables just listed are linked together by what is known as the modified *Stokes law*, named after the Ireland-born British scientist George Stokes (1819–1903), who is credited with having developed much of the general theory on the movement of a particle in a flowing fluid. Stokes was a gifted mathematician (21) who also explained the existence of fluorescence and the variation of gravity at the surface of the Earth—all during the brief span of 1845–1854. Although Stokes' law is not directly used in cell separations, it helps us understand how the various parameters affect the settling operation. For a simplified case of gravitational settling when the velocity relative to the surrounding medium is relatively modest (often expressed in chemical engineering literature via a dimensionless parameter called the Reynolds number being set at < 0.2), the Stokes' law modifications (22) suggest that

$$V_s = \frac{g(\rho_p - \rho)D_p^2}{18\mu}$$

where V_g is sedimentation rate (ft/s or m/s); g is gravitational acceleration constant (ft/s^2 or m/s^2), ρ_p is density of the particle (or cell) (s/ft^3 or kg/m^3), ρ is density of the liquid (s/ft^3 or kg/m^3), D_p is diameter of the particle (or cell) (ft or m), and μ is viscosity of the liquid [lb/(s ft) or Pa s&rsqb. Many of the factors in this equation may be beyond one's control in a given practical situation; it is not used very much for actual numerical calculation in cell separa-

tions. However, we can learn how cell sedimentation is affected under various conditions by examining this equation.

Effect of Density

In many large-scale industrial cell separation situations, the cell sedimentation velocity is too slow for sedimentation to be a practical tool. The sedimentation velocity is slow because biological cells and cell aggregates (if cell aggregates are present) tend to have a low density. The Stokes' law equation teaches us that the greater the *difference* between the densities of the cells and the liquid medium, the greater the cell sedimentation velocity. This property can be usefully exploited by a controlled increase of the liquid medium density. For illustration purposes, as shown in Figure 2, let us assume we need to separate a cell mixture containing two different cell densities in a given liquid medium. If the liquid density is increased via the addition of agents like Ficoll, fetal calf serum, or albumin, the *relative* ratio of sedimentation rates of the two cell types is increased, leading to a better separation. A refinement of this idea is to use a density gradient in the liquid medium through which the sedimentation is taking place. A technique based on this idea has been called isokinetic sedimentation (20,23,24).

Effect of Cell Size

Table 2 shows the distance traversed in 1 h by hypothetical spherical particles of specific gravity 2.0 and different cell sizes, and settling in a fluid with a specific gravity of 1.0 (25). In practice, the specific gravity of cells tends to be well below 1.2; the example merely illustrates the effect of cell size. Table 3 gives an idea of the typical sizes for the various cell types encountered in practice.

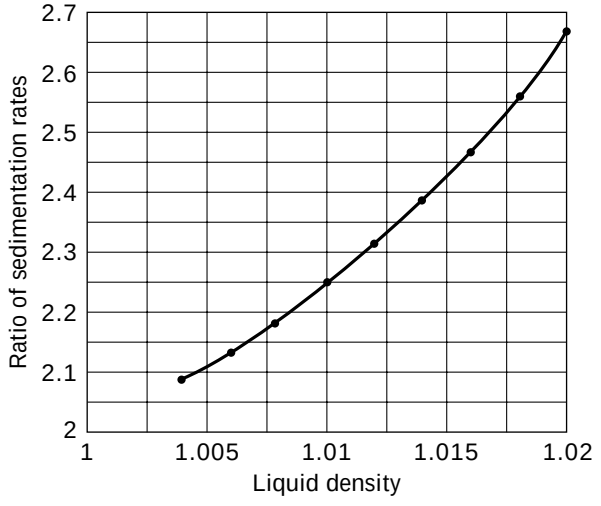


Figure 2. Effect of liquid density on relative sedimentation rate. This data assumes a hypothetical, idealized case where a mixture of two cell types, A and B (densities 1.05 and 1.10, respectively), is sedimenting in liquids of slightly differing densities.

Table 2. Effect of Spherical Particle Size on the Sedimentation Distance after 1 h in Water at 293 K

Particle radius (nm)	Distance settled
0.1	0.08 nm
1	8 nm
10	800 nm or 0.8 μm
100	80 μm
1,000	8 mm

Source: Ref. 25.

Table 3. Typical Sizes of Various Cell Types

Cell type	Size range (nm)
Bacterial virus	0.08
Bacteria	0.1–2.5
Yeast cells	2–10
Human red blood cells (erythrocytes)	6–8
Human white blood cells (leukocytes)	8–12
Filamentous fungi	10–40
Human liver cells	30
Plant cells	10–100
Paramecium, didinium (protozoa)	100–200

The sedimentation rate is proportional to the square of the diameter of the cell. Thus, any approach that leads to cell aggregation will help increase sedimentation substantially—albeit not as much as the "diameter square" relationship would seem to imply. The relative increase is less than that implied by the "square" relationship if the cell aggregate is porous and its density turns out to be less than that of the individual cells. In one study, the settling of larger-diameter activated sludge flocs was found to be proportional only to (floc diameter)$^{0.55}$ (26).

Bacterial cells are often difficult to separate owing to their small size and low density, and so flocculants such as sodium polyacrylate (27), polydimethyldiallyl ammonium chloride (28), and carboxymethyl cellulose (28) are added in some situations to achieve just this benefit. Ultrasound has been used to enhance the sedimentation of monoclonal antibody producing hybridoma cells (29) because the acoustic treatment aggregates the cells. Addition of ferromagnetic particles to a bacterial cell suspension (gluconic acid producing *Acetobacter methanolicus*) followed by an exposure to a magnetic field has been shown to enhance sedimentation (30) as a result of aggregation of cells with the magnetic particles. Two commercially products, Dynabeads® and DETACHaBEAD®, offer another way to achieve the immunomagnetic separation of cells (31). In this approach, sedimentation is enhanced by forming rosetted cells on the antibody-coated Dynabeads. The rosetted cells are washed and concentrated using a magnetic device. The positively isolated cells may then be detached for further study.

Because of the dramatic sensitivity of the sedimentation rate to the cell size, clinical researchers have attempted to determine whether the sedimentation rate of erythrocytes can be a useful indicator of patient health (19,15,32,33), results have been mixed, however.

Size differential between cells can thus also be a useful tool when one is trying to separate a specific cell type from a mixture of cells. If one were to take samples at a fixed point below the liquid surface from a container in which sedimentation of otherwise identical but differing cell sizes is under way, the initial samples would contain the largest cells. With time, smaller and smaller cells would be in the sample. This principle is useful in the separation of lymphocytes from erythrocytes where sedimentation through a dense medium in a centrifuge is one approach (32,33).

Effect of Viscosity

The sedimentation rate is inversely proportional to the viscosity of the liquid medium. This property can also lead to some practical options. Because the viscosity change with dilution is often nonlinear, one may dilute the liquid medium by 25–50% to achieve a greater decrease in viscosity, and hence, better sedimentation rate. Conversely, one may add something that increases the viscosity of the liquid medium to slow down the sedimentation of cells. An example of this is the use of 4% dextran (of average molecular weight 500,000 Da) as a viscosity-increasing agent in the study of antibody-binding rates by FACScans in immunology (34), where suspended cells must be prevented from settling.

Effect of Gravity

Cell separation via centrifugation (see CELL SEPARATION, CENTRIFUGATION) seeks to increase the sedimentation rates hundreds or thousands of times over simple sedimentation, by operating at a force much greater than the normal gravity. Because normal gravitational sedimentation is limited by the gravitational constant g, the relative efficiency of centrifugal separations has been examined by measuring the sedimentation velocity of disrupted yeast cells with and without a flocculent (35).

Effect of Chemical and Environmental Factors

It was pointed out earlier that polymers and other flocculating agents can aid sedimentation by increasing the effective cell size. In addition, the following factors can affect the behavior of particles in solution, hence can influence sedimentation rates:

- pH and temperature of the solution
- Presence of salts (e.g., of calcium, magnesium, manganese, sodium, potassium)
- Presence of sugars

Although in most cases the interaction of these parameters is complex, the question is one of how the chemical environment affects the surface charge of the cells and their tendency to form cell aggregates. The idea is that the surface interactions will control the degree of cell aggregation, and hence, the sedimentation rates. Most microbial cells usually have a negative surface charge (36,37). The system pH or temperature can affect this charge (37,38). Divalent cations, especially calcium, and certain pH ranges can promote cell aggregation, leading to better sedimentation rates (39–41). Monovalent cations and sugars can reduce the tendency of cells to aggregate, thus reducing the likely sedimentation rates (41,42). The overall effect depends on the specific cell type and the physiologic state of the cells involved.

BIBLIOGRAPHY

1. K.M.K. Lassen and C. Emborg, *FEMS Microbiol. Rev.* **14**(1), 89–91 (1994).
2. Japanese Pat 06209761 (Aug. 2, 1994), (to Toyobo Company Ltd).
3. F. Trampler, S.A. Sonderhoff, P.W.S. Pui, D.G. Kilburn, and J.M. Piret, *Bio/Technology* **12**(3), 281–284 (1994).
4. R.V. Sharma, R.T. Edwards, and R. Beckett, *Appl. Environ. Microbiol.* **59**(6), 1864–1875 (1993).
5. A.B.R.A. Maia and D.L. Nelson, *Biotechnol. Bioeng.* **41**(3), 361–369 (1993).
6. K.M.K. Lassen, P.R.E. Carcia de Cladeano, and C. Emborg, *Biotechnol. Tech.* **6**(2), 121–126 (1992).
7. I. Diaz and P. Carbonero, *Plant Cell Reprod.* **10**(12), 595–598 (1992).
8. K. Kuus-Reichel, A. Beebe, and C. Knott, *Hybridoma* **10**(4), 529–538 (1991).
9. K.Y. Lee and G. Baerwald, *Biotechnol. Lett.* **13**(8), 595–598 (1991).
10. K.L. Henry, R.H. Davis, and A.L. Taylor, *Dev. Ind. Microbiol.* **31**, 53–58 (1990).
11. K.L. Henry, C.S. Parnham, R.H. Davis, and A.L. Taylor, *Appl. Biochem. Biotechnol.* **24–25**, 651–662 (1990).
12. H. Yavetz, R. Hauser, Z.T. Homonnai, G.F. Paz, J.B. Lessing, A. Amit, and I. Yogev, *Andrologia (Germany)* **28**(1), 3–6 (1996).
13. J. Risopatron, R. Sanchez, N. Sepulveda, P. Pena, E. Villagran, and W. Miska, *Theriogenology* **46**(1), 65–73 (1996).
14. O.H. Iverson, M. Roger, H.E. Solberg, and P. Wetteland, *J. Intern. Med.* **240**(3), 133–141 (1996).
15. P. Wetteland, M. Roger, H.E. Solberg, and O.H. Iverson, *J. Intern. Med.* **240**(3), 125–131 (1996).
16. P.T. Selvaraj and K.L. Sublette, *Biotechnol. Prog.* **11**(2), 153–158 (1995).
17. D.R. Nagaraj, "Minerals Recovery and Processing, in J.I. Kroschwitz ed., *Kirk-Othmer Encyclopedia of Chemical Technology*, Wiley, New York, 1992, pp. 833–834.
18. R.H. Davis, in M.A. Hjortso and J.W. Roos eds., *Cell Adhesion*, Marcel Dekker, New York, 1995, pp. 135–185.
19. J.A. Searles, P. Todd, and D.S. Kompala, *Biotechnol. Prog.* **10**, 198–206 (1994).
20. J.R. Wells, in T.G. Pretlow II and T.P. Pretlow eds., *Cell Separation*, Academic, New York, 1982, pp. 169–189.
21. R. Porter ed., *The Biographical Dictionary of Scientists*, vol. 16, Oxford University Press, New York, 1994, pp. 647–648.
22. J.M. Coulson, J.F. Richardson, J.R. Backhurst, and J.H. Harker, *Chemical Engineering*, vol. 2, Pergamon Press, Oxford, 1978, pp. 174–175.
23. T.G. Pretlow II and T.P. Pretlow, in T.G. Pretlow II and T.P. Pretlow eds., *Cell Separation*, Academic, New York, 1982, pp. 41–60.

24. R.G. Miller, in R.H. Pain and B.J. Smith eds., *New Techniques in Biophysics and Cell Biology*, Wiley, New York, 1973, pp. 87–112.
25. A. Bleier, in J.I. Kroschwitz ed., *Kirk-Othmer Encyclopedia of Chemical Technology*, vol. 6, Wiley, New York, 1992, pp. 813–814.
26. D.-H. Li and J.J. Ganczarczyk, *Water Res.* **21**, 257–262 (1987).
27. Japanese Pat. 05003780 (Jan. 14, 1993), (to Mitsui Toatsu Chemicals, Inc.).
28. German Pat. DD298263 (Feb. 13, 1992), (to Forsch Biotechnology).
29. P.W.S. Pui, F. Trampler, S.A. Sonderhoff, M. Groeschl, D.G. Kilburn, and J.M. Piret, *Biotechnol. Prog.* **11**(2), 146–52 (1995).
30. German Pat. DD278598 (May 9, 1990), (to Karl Marx University).
31. DYNAL Dyna and DETACHaBEAD, Dynal A.S., PO Box 158, Skoyen N-0212, Oslo, Norway.
32. W. Emlen, J. Niebur, G. Flanders, and J. Rutledge, *J. Rheumatol.* **23**(6), 974–978 (1996).
32. S.L. Berger, in W.B. Jacoby and I.H. Pastan eds., *Methods in Enzymology*, vol. 58, Academic, New York, 1979, pp. 490–491.
33. M.N. Penev, P. Doukovapeneva, and K. Klinov, *Scand. J. Clin. Lab. Invest.* **56**(3), 285–288 (1996).
34. E. Martz, *J. Immunol.* **115**(1), 261–267 (1975).
35. R.L.P. Adams, in T.S. Work and R.H. Burdon eds., *Cell Culture for Biochemists*, Elsevier/North Holland, New York, 1980, pp. 70–73.
35. A.I. Clarkson, P. Lefevre, and N.J. Titchener-Hooker, *Biotechnol. Prog.* **9**(5), 462–467 (1993).
36. K. Thethi, P. Jurasz, A.J. MacDonald, A.D. Befus, S.F. Man, and M. Duszyk, *J. Biochem. Biophys. Methods* **34**(2), 137–145 (1997).
37. Y.E. Collins and G. Stotzky, *Can. J. Microbiol.* **42**(7), 621–627 (1996).
38. G.V. Sherbet and S. Jackson, *Anticancer Res.* (**1**), 129–134 (1986).
39. A.M. Porter and R.J. MacAulay, *J. Inst. Brew.* **71**, 175–179 (1965).
40. G.G. Stewart, I. Russell, and I.F. Garrison, *J. Inst. Brew.* **81**, 248–257 (1975).
41. M.A. Amri, R. Bonaly, B. Duteurtre, and M. Moll, *Eur. J. Appl. Microbiol. Biotechnol.* **7**, 235–240 (1979).
42. P.M. Jayatissa and A.H. Rose, *J. Gen. Microbiol.* **96**, 165–174 (1976).

See also CELL DISRUPTION AND LYSIS; CELL SEPARATION, CENTRIFUGATION; FILTER AIDS.

CELLULOSE CONVERSION.

See ORGANIC COMPOUNDS, CELLULOSE CONVERSION.

CENTRIFUGES, ANIMAL CELLS

DANIEL D. HANLE
Kendro Laboratory Products
Newtown, Connecticut

KEY WORDS

Centrifuge
Continuous flow
Decanter
Disc stack
Fed batch
Inclined separator
Perfusion
Spin filter
Tangential flow filtration
Totally closed system

OUTLINE

INTRODUCTION

With the emergence and growth of animal cell bioprocesses during the past 15 to 20 years, the need has arisen for unit operations that gently and effectively separate animal cells at the process scale. Centrifuge-based separation systems available today are especially well suited to this task. By its nature, centrifuge separation is inherently robust and delivers consistent separation performance. Centrifuge systems now incorporate a number of design changes, including some that are quite novel, aimed at tailoring their performance to the special requirements of animal cell bioprocesses. As a result, today we have a powerful array of options available for centrifuge-based unit operations for use in both the animal cell processes under development and the existing processes that are being improved for added efficiency and productivity.

ANIMAL CELL BIOPROCESSES

As biotechnology's commercial success has grown, it has created the need for highly regulated, effective, and, most recently, efficient production processes. Such biotechnology production processes (bioprocesses) have historically drawn heavily from other process industries, such as the chemical process industry for concepts, practices, and equipment, including centrifuges. Bioprocess has therefore become, in some cases, a hybrid field where established process techniques are intermingled with laboratory methods that have been multiplied manyfold for the required

process scale. Nowhere does this appear to be more the case than for the bioprocess of animal cells.

Although bacteria or yeast has been the more common choice for biological production to date, each can have such drawbacks as producing improperly formulated product or product that is unrecoverable as insoluble inclusion bodies (1). In such cases, animal cells have often been called on to overcome these limitations. Unfortunately, animal cells do not share some of the more desirable attributes of these other organisms. By comparison, animal cells are slower growing; exhibit lower protein expression levels; are susceptible to infection by a large number of adventitious agents such as viruses, prions, and so forth; and, because of their lack of a cell wall, are much less tolerant of many traditional process operations. Numerous advances in animal cell biotechnology and fermentation have enhanced the productivity of animal cells. But the bioprocess equipment has often lagged. Animal cell bioprocess has therefore become a "force fit" of techniques and equipment originally designed and developed for other purposes.

Nonetheless, the fact that animal cells are often the only practical way to produce what is required has generated a strong push by many in the industry to develop processes and equipment tailored to animal cell bioprocess. Centrifugation is no exception.

ANIMAL CELL CENTRIFUGATION

Process Applications

Centrifugation's role in animal cell bioprocess is primarily as a unit operation immediately adjacent to the bioreactor, as illustrated in Figure 1. In this role, the centrifuge is positioned to either (*1*) clarify the cell mass exiting the bioreactor, as in the case of a batch operation, or (*2*) retain the cells for further processing, as in the case of a continuous operation. These two operating modes, however, differ somewhat depending on whether the product is intracellular or extracellular. Intracellular processes, where the focus is on harvesting the product-laden cells, require that the centrifuge concentrate the intact cells from the fermentation broth. Care is taken to avoid damaging the cells during this concentration step to maximize product recovery. Any ruptured or lysed cells also release agents into the process streams that can be harmful to the desired product, especially when that product is a protein. Occasionally, this type of care is also needed when the cells are to be stored or are otherwise required to remain viable for further processing. In many of these cases, aseptic operation is essential.

Extracellular processes, on the other hand, require a different set of operating principles where the key focus is on keeping the cells viable. Any cell damage that the centrifuge causes in these processes will have a direct deleterious effect on the process productivity. Not only would the producing cells themselves be lost, but here, too, harmful intracellular agents would be released. The situation is further aggravated by the fact that the extracellular product is relatively vulnerable to any harmful agent because both are present in the bulk fluid. It is clear that centrifugation processes that avoid cellular damage are preferable in these cases, especially when operating in a continuous or perfusion process mode.

A further distinction in animal cell bioprocess that bears mentioning is that the producing cells can be either freely suspended or immobilized. The earlier discussion relates directly to suspension systems. For immobilized systems, where, for example, the cells require that they be tethered to induce growth or production, the need for centrifugation is either at the end of the process, whenever the tethered cells are released for further processing, or for clarification of debris released by the tethered cells, or both. When called for in this manner, the centrifugation considerations are similar to those for freely suspended cells.

Fed Batch versus Perfusion

Perhaps the most prevalent bioprocess mode in use today for freely suspended cells is a simple batch system using carefully constituted media. But in the past decade or so, two more involved processes, fed batch and perfusion, have gained much of the attention. The fed-batch process tends

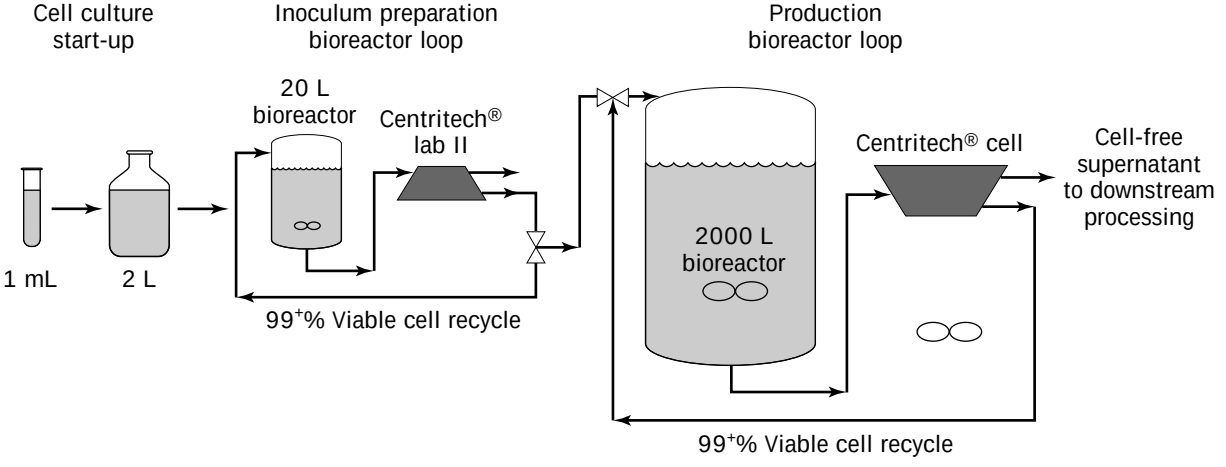


Figure 1. Centrifuge system as integral unit operation for animal cell bioprocess.

to make more efficient use of the media expense and other operating costs, whereas the perfusion process, being continuous, tends to make more efficient use of the producing cells and the physical plant investment.

The fed-batch process is currently the more prevalent of these two processes, partly because it is easier to operate in the laboratory and at the pilot and production scale. Because a batch culture is by its nature limited in its duration, process experience can be built up relatively quickly during the course of several runs. Here the need for processing the cells is confined to the end of the production run. Centrifugation may be used to concentrate the cells for intracellular product or clarify the supernatant for extracellular product. In either case, as mentioned earlier, avoiding cell damage would likely lead to important productivity gains.

Perfusion processes afford the benefits of continuous operation. To do so, however, perfusion requires an effective way to retain the cells throughout the production run with little or no cell viability loss. Centrifugation can deliver the needed performance provided it is continuous in both the liquid stream and the cell concentrate stream, can be operated aseptically, and does not damage the cells. The specific designs of centrifuges applicable to animal cell bioprocess are reviewed in the next section with reference to these important aspects. A continuous-flow centrifuge system installed in a perfusion animal cell process is shown in Figure 2.

The considerations of such processes are directly applicable to anchored cells when those cells are attached to microcarriers. Cells attached to microcarriers are, however, particularly sensitive to fluid shear forces (2–5). For example, cells that are not otherwise damaged by fluid shear are lost if the shear that is present separates them from the microcarriers and they do not reattach (6). In these cases, the bioprocess centrifuge needs to treat the process fluid especially gently.

The various roles of centrifugation in animal cell bioprocess are summarized in Table 1 for the cases discussed. Although generalizations tend to be difficult, these roles as delineated in the table are in many ways representative of the scope of possibilities.

The emergence of insect cell bioprocess has added yet another dimension to the need for gentle handling of the cells. Although insect cells have several advantages over animal cells, they tend to be even less robust than animal cells are during rough handling in the process (7–11). Coupled with the fact that insect cells typically generate intracellular product (specifically so with baculovirus systems), the centrifuge operation must generate especially low shear forces to avoid cell rupture and the resulting productivity loss.

Centrifuge Designs

Design Considerations. Centrifugation is a very robust and versatile separation technique. Because it relies primarily on density differences, centrifugation can readily account for changing process conditions by simple modulation of the intensity or duration of the *g*-field as needed during operation (12). At lower speeds and *g*-forces, the centrifuge can be used to gently remove the larger and more dense cells from the process stream; at higher speeds and *g*-forces, the centrifuge can be used to spin down ever smaller cell debris to increasingly clarify the process stream to the degree required.

The actual design of the centrifuge, however, depends on the various optimization trade-offs specific to the process application. Typical design trade-offs include high sample throughput versus instrument size, high centrifu-

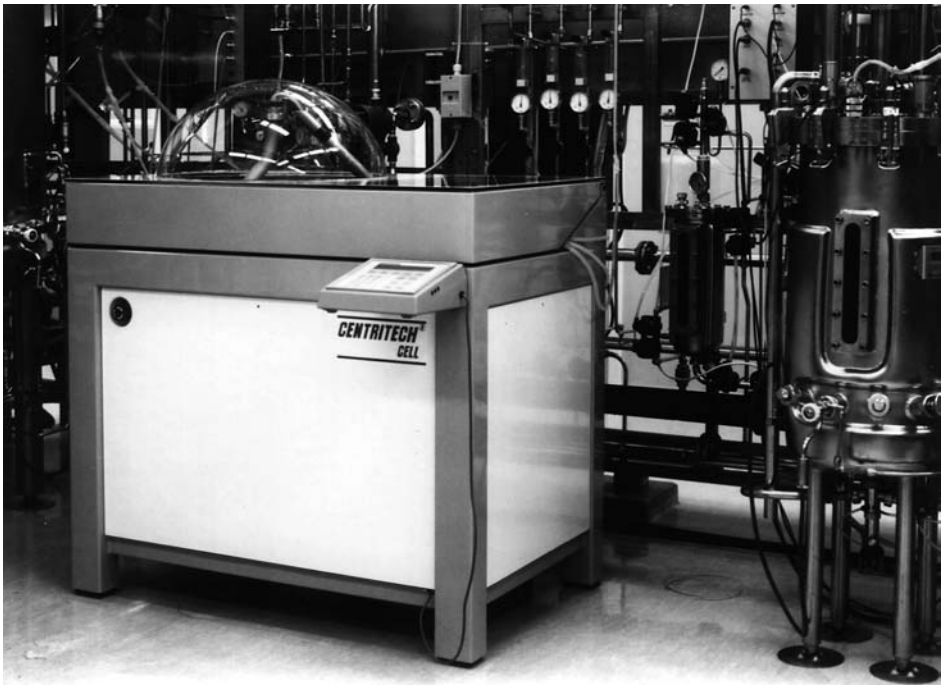


Figure 2. Continuous-flow centrifuge system installed in a production-scale perfusion process.

Table 1. Centrifugation Roles in Animal Cell Bioprocess

	Suspension cultures		Immobilized cultures	
	Fed batch	Perfusion	Anchored	Microcarriers
		Extracellular product		
During the process	N/A	Retain cells with minimal cell damage	Remove/recover cells released into the process stream	Retain the microcarriers with minimal cell loss or damage
At the end of the process	Collect supernatant with minimal cell debris	Recover the final volume of the supernatant	Recover cells with minimal damage for next inoculation	Collect supernatant with minimal cell debris
		Intracellular product		
During the process	N/A	Retain most of the cells with minimal cell damage	Remove/recover cells released into the process stream	Retain most of the microcarriers with minimal cell loss or damage
At the end of the process	Collect cells with minimal cell damage	Recover the final volume of the cells	Recover cells with minimal damage for next inoculation	Collect supernatant with minimal cell debris

Note: N/A, no centrifuge role directly applicable to this case.

gal speeds versus large load volumes, versatility versus dedicated instrument specialization, batch versus continuous operation, and, for continuous operation, solids-retaining versus solids-ejecting design. Laboratory-scale applications tend toward smaller, more versatile, batch or continuous solids-retaining systems, whereas pilot- to production-scale applications tend toward high throughput, specialized, continuous solids-ejecting systems.

When applied specifically to the separation of animal cells, centrifugation typically involves the added considerations of avoiding high fluid shears and maintaining high aseptic integrity. Additional design trade-offs are therefore involved, including (*1*) increased volume throughput versus maintaining the shear forces below critical levels that can result in cell death or damage, and (*2*) providing a totally closed aseptic system versus relying on rotating seals, a source of potential contamination and heat generation as well as generation of high shear stresses.

Design Approaches. To date, the design of animal cell centrifuges has been approached from three distinct directions. The first two directions have been through engineering extensions of other centrifuge equipment. The third is a centrifuge design specifically engineered for animal cell bioprocess.

The first direction has been to take centrifuge designs for larger process scale or bacteria and yeast and adapt them for animal cells. This approach has been more or less successful depending on whether cell damage was to be avoided or continuous operation was desired. Such centrifuges tend to spin the cells down into a paste or cake, which requires some type of cell removal process such as scraping or flushing out of the cells under high shear forces.

The second direction has been to take centrifuge designs for blood component manufacture and adapt them for animal cells. Blood component manufacturers have been making commercial use of continuous-flow centrifugation for many years now. Numerous such blood separation system designs have focused on cell concentration, supernatant clarity, throughput, and other issues that are also important to animal cell bioprocess. But there are several key differences between blood component manufacture and animal cell bioprocess. Blood cells, for example, flow much more readily than typical suspended animal cells. Animal cell centrifuges therefore require some type of low shear method to nudge the animal cells out of the centrifuge in a continuous mode. Also, formed elements of blood are more tolerant of shear stresses than are most animal cells in bioprocess. Flow channel designs in the centrifuge that work for blood therefore tend to be not at all suitable for animal cells. And the usual time frame for blood separation (on the order of a few hours) is much less than that necessary for animal cell bioprocess (on the order of days or weeks).

The third direction was taken by Alfa-Laval Separation AB (Tumba, Sweden) during the 1980s when a continuous-flow centrifuge design was developed specifically for animal cell bioprocess. Combining some of the features of larger process centrifuges and those of the smaller centrifuges, this unique design focused on maintaining a continuous flow of both the clarified supernatant and the cell concentrate. Key engineering issues that had to be overcome were (*1*) gently coaxing the concentrated animal cells to smoothly exit the system without significant cell damage; (*2*) sustaining continuous operation over days and weeks; and (*3*) creating a totally closed system for inherently safe aseptic operation.

These three approaches have had the common goal to overcome the limitations of laboratory-scale and even small production–scale centrifugation. At these smaller scales, research centrifuges have often been pressed into service. Several companies provide floor-model centrifuges that can process, for example, as much as 6 L at *g*-forces

more than adequate for cellular separations. Among these companies are Beckman Instruments, Inc. (Fullerton, Ca.), Kendro Laboratory Products L.P. (Norwalk, Conn.) (Sorvall® and Heraeus® brands), Hitachi Koki Co., Ltd. (Hitachi-naka City, Japan), Jouan S.A. (St. Herelain, France), and Hettich GmbH (Tüttlingen, Germany). Kendro, through its Sorvall® brand, has recently introduced a series of major software packages to their centrifuges to provide automatic process documentation for easier and more reliable process control and regulatory compliance. Other suppliers are now providing software as well. These centrifuges are designed for batch operation only. Some of these companies offer continuous-flow models, but they tend to have small throughput capacity, do not eject the cellular material, and are typically difficult to set up and clean.

Because one of the keys to successful animal cell centrifugation is the gentle handling of the cells, there have been over the years several research groups exploring the various aspects of flow-induced shear stress on different animal cell lines. Although it has been shown that cell damage in bioreactors can be attributed to bubble breakup and other gas–liquid interface phenomena (13–16), other findings indicate that cell damage can also occur at shear stress levels at lower levels of energy dissipation (17–19). Shear stress susceptibility, moreover, varies significantly among cell lines (8,20), thereby adding shear sensitivity to the list of key criteria for selecting a particular cell line for process scale-up.

Current Products. Several suppliers provide centrifuges for animal cell bioprocess that are based on proven designs for other applications. These designs typically provide high throughput and *g*-forces, clean-in-place (CIP) and steam-in-place (SIP) capability, and some method for removing the cellular material that accumulates in the centrifuge. Each design is discussed briefly in the following and then summarized in Table 2. More complete descriptions, including specific schematic depictions of how these systems operate, are available through the respective suppliers.

At the small, batch end of the spectrum, Alfa-Laval's AS-26 SP Super® Centrifuge is capable of high throughput rates in the range of 300 to 500 L/h. The AS-26, however, is strictly a batch centrifuge because it retains the solids that must later be removed by manual intervention. Also, the animal cells are removed as cell debris, so this centrifuge would be unsuitable for processes where cell viability after separation is desired.

At the pilot scale, Alfa-Laval's BTPX 205 system uses the classical disc-stack design to generate up to 12,800 × *g* and a maximum throughput of 1,200 L/h. The 205 has a solids-ejecting design that is tailored to produce very dry solids. It is therefore well suited for, but limited to, batch animal cell processes where cell viability is unimportant. Although its solids-ejecting capability would allow the 205 to be considered for use in a continuous animal cell process, its rough cell handling would result in unacceptable levels of cell damage.

At the production scale, Alfa-Laval's BTAX 510 system delivers up to 15,000 × *g* and a maximum throughput of 4,000 L/h. The 510, like the 205, uses a disc-stack, solids-ejecting design. So the 510 process applicability is, at its larger scale, similar to that for the 205.

Westfalia Separator AG's (Oelde, Germany) CFA-01 separator is also a continuous-flow system based on a disk-stack solids-ejecting design. Specific investigation (21) has demonstrated the effective use of this system for animal cell processing. This system, however, like the systems discussed earlier and the decanter designs discussed next, requires a rotating seal. To remove the frictional heat from its seal, the CFA-01 requires a sealing liquid that must be highly purified to maintain aseptic operation. These key considerations combine to significantly increase the complexity of the overall system. Later work with the West-

Table 2. Comparison of Centrifuges for Animal Cell Bioprocess

Supplier	Scale	AS 26-SP-System	Max *g*-force (×*g*)	Max throughput (L/h)	Solids ejecting	Low shear	Closed system	Comments
Alfa-Laval	Small	AS 26-SP	20,000	600	No	No	No	Not suitable for continuous processes
Alfa-Laval	Pilot	BTPX 205	13,000	1,200	Yes	No	No	
Alfa-Laval	Prod'n	BTAX 510	15,000	4,000	Yes	No	No	
Westfalia	Pilot/Prod'n	CFA-01	300	240	Yes	No	No	
Westfalia	Pilot/Prod'n	CSA-1	5,000	50	Yes	No	No	
Alfa-Laval/ Sharples	Prod'n	Super-D-Canter®	10,000	15,000	Yes	No	No	Difficult to CIP and SIP
Carr	Pilot/Prod'n	Powerfuge® P6	20,000	60	Yes	No	No	
Kendro Heraeus®	Small	Contifuge® 17 RS	19,000	54	No	No	No	Not suitable for continuous processes
Baxter	Small	Cell Harvester®	1,000	60	No	Yes	Yes	Not suitable for continuous processes
Kendro Sorvall®	Small	Centritech® LAB	250	10	Yes	Yes	Yes	Well suited to continuous processes
Kendro Sorvall®	Pilot/Prod'n	Centritech® CELL	300	100	Yes	Yes	Yes	Well suited to continuous processes

Note: Prod'n, production.

falia CSA-1 disc-stack centrifuge (22) came to mixed conclusions and reported heating issues.

Decanter centrifuges, such as Alfa-Laval's Sharples® Super-D-Canter® systems, use a solids-ejecting design that differs from that of the disc-stack systems. The decanter design displaces the solids by pushing or scraping them out of the centrifuge using a type of helical blade arrangement. Using this design, high animal cell viabilities may be obtainable in the ejected solids. But these solids are not well suited to continuous processes where the cells must be resuspended, for example, to be introduced back into the bioreactor.

Carr Separations, Inc. (Farmington, Mass.), as a relative newcomer (the company was founded in 1987), has developed a different design specifically for bioprocess as well as for other pharmaceutical and food applications. The Powerfuge® P6 Separation System uses a new solids separation design that can generate very dry solids. This new design also allows for easier CIP than traditional disc-stack designs, an issue that can be a critical drawback for the original design, especially in the larger-scale equipment. The Powerfuge® is capable of generating up to 20,000 $\times$ g and throughput rates of 60 L/h. Recently, Carr introduced a smaller, low-end bench-top version of the Powerfuge® called the Pilot Separator®. For animal cell process, the suitability of these systems is similar to that for decanter centrifuges.

Kendro, through its Heraeus® brand, offers a pair of bench-top centrifuges designed to operate with a continuous-flow rotor system that can generate 24,000 $\times$ g and a flow rate of 54 L/h. These units are compact and readily portable. They are continuous in the liquid phase but are not designed to eject the solids. Although popular for laboratory-scale work, these units can provide only qualitative support to any planned scale-up plans during process development.

Baxter International, Inc. (Deerfield, Ill.) has modified the design of one of their blood component systems to focus on animal cell bioprocess at the laboratory and smaller production scales. Called the Cell Harvester®, this system can be used to develop centrifuge-based cell separation protocols and to prove out process concepts. Because manual intervention is needed to remove the cells, the applicability of this system to continuous processes is extremely limited. One key advantage of the Baxter system, however, is its totally closed system design. This design allows for the use of a consumable to avoid the various issues that accompany the need for providing CIP/SIP.

The third, unique design approach mentioned earlier that was taken by Alfa-Laval in the 1980s, also uses a consumable insert. These disposable inserts are provided sterile so that, upon making sterile connections to the system, the user has created a totally closed, sterile flow path and separation zone for the process fluids (23). This design approach, called Centritech® by Alfa-Laval (see Figure 3), is now available through and supported by Kendro under its Sorvall® brand. Key trade-offs of the Centritech® design are very gentle cell handling and a totally closed, inherently aseptic system versus high throughput and g-force (24).

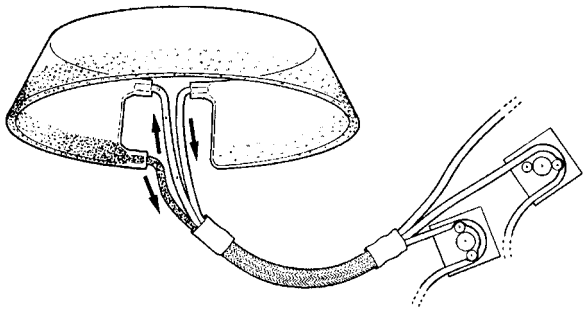


Figure 3. Sketch of the sterile, disposable insert for the Centritech® continuous-flow centrifuge system.

The Centritech® technology is embodied in both a smaller bench-top laboratory-scale system (25) and a larger floor-model pilot- to production-scale system (26). Depending on the application, however, the laboratory-scale unit has been and can be used for small production processes (27).

ALTERNATIVES TO ANIMAL CELL CENTRIFUGATION

Filtration has been the traditional choice as an alternate bioprocess unit operation to animal cell centrifugation. The most common form of filtration used to separate the cellular material from the fermentation broth is tangential flow filtration (TFF).

The foremost advantage of using TFF is that it is a technology that has been widely used to good effect (28,29). TFF does, however, have the disadvantage that, as for all filter-based systems, fouling and clogging can be a constant and often unpredictable problem if not sized properly. Filter systems also deliver performance that degrades over the course of the run, necessitating either that they be reconditioned or the run stopped to replace the filters. Even when reconditioned, filters do not usually return to their original performance, so operating them becomes a classic case of diminishing returns. Recently, issues such as documenting well-characterized products and other concerns have made the reuse of filters less desirable. In these cases, the only viable option is to stop the run, change the filters, resterilize the process, and start up again.

A further disadvantage of TFF is the inherent presence of shearing at the interface caused by the membrane. Material buildup can affect the way the system works, but a substantial shear remains across the boundary layer. Such shearing can run counter to the low shear requirements for animal cell processing and also requires that a recycle loop (with a pump) circulates the process fluid many times greater than that necessary for continuous (perfusion) operation.

Another alternative to animal cell centrifugation is the use of spin filters. Spin filters have the very attractive aspect that as a cell retention device for perfusion processes, they can be integral to the bioreactor. Thus, a single unit operation is created. Cell damage is a concern (30), but evidence suggests that with proper design and operation this issue can be minimized or even overcome (31).

A disadvantage of the spin-filter systems is that they too, like TFF, are susceptible to fouling and clogging. What can be worse, however, is that when a spin-filter system integral to the bioreactor clogs unexpectedly, the necessary maintenance typically requires that the bioreactor containment be breached. The risk of infection from such a breach can be significant. Another disadvantage is that the scaleability of spin-filter systems is limited (32). They may not be a practical unit operations choice if larger scales are anticipated during process development or once in full production.

Another device that has been used for cell retention in continuous processes is the inclined separator (33). The design of this device is mechanically simple, can be made to be relatively easy to CIP and SIP, and, because it relies on gravity (i.e., $1 \times g$), has many of the same process benefits as centrifugation. A drawback with this technology today appears to be its relative lack of throughput scaleability and, of course, that its separation efficiency is limited to that achievable at $1 \times g$ (vs. centrifugation at $100 \times g$ or more).

THE FUTURE OF ANIMAL CELL CENTRIFUGATION

Going forward, centrifugation is likely to be as commonly used for bioprocess as it is for the comparable chemical process industry. The inherent advantages of centrifugation (such as robustness and versatility) continue to fuel creative new concepts for harnessing these advantages for animal cell bioprocess. Novel centrifuge designs and the unique continuous methodologies for bioprocess will likely combine to make the centrifuge a necessary unit operation in a large number of the new animal cell bioprocesses being developed in the near future. Moreover, recent changes in the FDA regulations are conducive to the accelerated adoption of new and more efficient process methodologies. Although in the past there were often heavy disincentives to consider the type of process improvements that, for example, continuous-flow centrifugation allows, many of the regulatory-based concerns have been reduced or eliminated. As a result, we are likely to see adoption of new, higher productivity process improvements in existing processes as well as the development of new ones.

BIBLIOGRAPHY

1. G. MacMichael, *BioPharm.* **10**, 38–40 (1997).
2. M.S. Croughan and D.I.C. Wang, *Biotechnol. Bioeng.* **33**, 731–744 (1989).
3. M.S. Croughan and J.-F.P. Hamel, *Biotechnol. Bioeng.* **32**, 975–982 (1988).
4. E.T. Papoutsakis and R.S. Cherry, *Biotechnol. Bioeng.* **32**, 1001–1014 (1988).
5. R.V. Venkat and J.J. Chalmers, *Cytotechnology* **22**, 95–102 (1996).
6. M.S. Croughan, E.S. Sayre, and D.I.C. Wang, *Biotechnol. Bioeng.* **33**, 862–872 (1989).
7. J.J. Chalmers, in M.L. Shuler, H.A. Wood, R.R. Granados, and D.A. Hammer eds., *Baculovirus Expression Systems and Biopesticides*, Wiley-Liss, New York, 1995, pp. 175–204.
8. Technical Support Package, *NASA Tech. Brief* MSC-22336 (1993).
9. J. Tramper, J.B. Williams, and D. Joustra, *Enzyme Microb. Technol.* **8**, 33–36 (1986).
10. K. Trinh, M. Garcia-Briones, F. Hink, and J.J. Chalmers, *Biotechnol. Bioeng.* **43**, 37–45 (1994).
11. J.J. Chalmers, *Cytotechnology* **20**, 163–171 (1996).
12. T.C. Ford and J.M. Graham, *An Introduction to Centrifugation*, BIOS Scientific Publishers Limited, Oxford, U.K., 1991.
13. E.T. Papoutsakis, *Trends Biotechnol.* **9**, 427–437 (1991).
14. N. Kioukia, A.W. Nienow, A.N. Emery, and M. Al-Rubeai, *Trans. IChE* **70**, 143–148 (1992).
15. J.J. Chalmers, *Cytotechnology* **15**, 311–320 (1994).
16. D. Chattopadhyay, J.F. Rathman, and J.J. Chalmers, *Biotechnol. Bioeng.* **45**, 473–480 (1995).
17. M. Al-Rubeai, A.N. Emery, S. Chalder, and M.H. Goldman, *J. Biotechnol.* **31**, 161–177 (1993).
18. C. Born, Z. Zhang, M. Al-Rubeai, and C.R. Thomas, *Biotechnol. Bioeng.* **40**, 1004–1010 (1992).
19. K.T. Kunas and E.T. Papoutsakis, *Biotechnol. Bioeng.* **36**, 476–483 (1990).
20. Z. Zhang, M. Al-Rubeai, and C.R. Thomas, *Biotechnol. Tech.* **7**, 177–182 (1993).
21. V. Jager, in H. Murakami ed., *Proc. of the 4th JAACT Mtg.*, Kluwer, Dordrecht, The Netherlands, 1991.
22. R. Kempken, A. Preissmann, and W. Berthold, *Biotechnol. Bioeng.* **46**, 132–138 (1995).
23. D.D. Hanle, *Int. Biotech. Lab.* 12, (October 1996).
24. J. Hodgson, *Biotechnology* **9**, 628–629 (1991).
25. N. Chatzisavido, T. Bjorling, C. Fenge, S. Boork, E. Lindner-Olsson, and S. Apelman, in T. Kobayashi, Y. Kitagawa, and K. Okumura eds. *Animal Cell Technology: Basic and Applied Aspects*, Kluwer, Dordrecht, The Netherlands, 1993, pp. 463–468.
26. T. Bjorling, U. Malmstrom, in R.E. Spier, J.B. Griffiths, and B. Meignier eds., *Production of Biologicals from Animal Cells in Culture*, Butterworth-Heinemann, Oxford, U.K., 1991, pp. 394–399.
27. C.W. Kemp and A. Chan, *Waterside Monoclonal Antibody Mtg.*, Norfolk, Vir., 1997.
28. B. Maiorella, G. Dorin, A. Carion, and D. Harano, *Biotechnol. Bioeng.* **37**, 121–126 (1991).
29. R. vanReis, L.C. Leonard, C.C. Hsu, and S.E. Builder, *Biotechnol. Bioeng.* **38**, 413–422 (1991).
30. A.J. Brennan, J. Shevitz, and J.D. MacMillan, *Biotechnol. Tech.* **1**, 169–174 (1987).
31. D.C-H. Jan, A.N. Emery, and M. Al-Rubeai, *Biotechnol. Tech.* **7**, 351–356 (1993).
32. V.M. Yabannavar, V. Singh, and N.V. Connelly, *Biotechnol. Bioeng.* **43**, 159–164 (1994).
33. J.A. Searles, P.W. Todd, R.H. Davis, and D.S. Kompala, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products for Today, Prospects for Tomorrow*, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 240–242.

See also ANTIBODY PURIFICATION; CELL SEPARATION, CENTRIFUGATION; CELL SEPARATION, SEDIMENTATION.

CEPHALOSPORINS

SHU-JEN D. CHIANG
JONATHAN BASCH
Bristol-Myers Squibb Company
Syracuse, New York

KEY WORDS

7-ACA
7-ADCA
Biosynthetic genes
Biosynthetic pathways
Cephamycins
Extraction
Fermentation
β-Lactams
Semisynthetic cephalosporins
Strain improvement

OUTLINE

INTRODUCTION

Since the discovery of penicillins in 1928 and the discovery of cephalosporin C in 1945 (1), β-lactams have become the most important class of antibiotics because of their low toxicity and their effectiveness against bacterial infection. In cephalosporins, the four-membered β-lactam ring is fused to a six-membered dihydrothiazine ring rather than to the five-membered thiazolidine ring found in penicillins (Fig. 1). The natural product, cephalosporin C, has low activity against gram-negative bacteria, but it is resistant to penicillinase hydrolysis. Many chemical derivatives of cephalosporin C have a higher and broader spectrum of antibacterial activity than cephalosporin C. After 40 years of development, cephalosporins (i.e., chemical derivatives of cephalosporin C or semisynthetic cephalosporins) have become one of the most widely used antibiotics in the world; in 1995, the world market for cephalosporins was valued at $9.3 billion at the dosage form level (2). Of this total, $6.5 billion represented products made from cephalosporin C, with the remainder derived from penicillins. In this article, we discuss the development and production of cephalosporins and their chemical derivatives.

Historical Background

Penicillin G (benzylpenicillin) was considered a miracle drug, when discovered in the 1940s, for its effective treatment of the systemic infections caused by gram-positive bacteria such as *Staphylococcus aureus*. Unfortunately, penicillin G was generally ineffective against gram-negative bacteria, and within a few years of its introduction many common gram-positive bacteria had acquired resistance to penicillin G as a result of the expression of bacterial penicillinases. At about this time, a new antibiotic, cephalosporin C, was discovered from a species of mold, *Cephalosporium acremonium*. The *C. acremonium* strain was originally isolated from seawater near a sewage outflow in Cagliari, Sardinia, by Brotzu in 1945. The culture filtrates of this fungus exhibited a broad spectrum of antibacterial activity both in vitro and in preliminary clinical studies (3). In 1948, Brotzu published his findings, which failed to attract attention. At the British Medical Research Council's suggestion, Brotzu sent his fungal culture, *C. acremonium*, which was later reclassified as *Acremonium chrysogenum* in 1971 by Gams (4), to Florey at Oxford. The work of Newton and Abraham at Oxford led to the purification of penicillin N (Fig. 1) and cephalosporin C in 1955 (5). Both compounds had activity against gram-negative bacteria. Penicillin N was as susceptible to staphylococcal penicillinase hydrolysis as was penicillin G. Cephalosporin C, however, had a high degree of resistance to hydrolysis by staphylococcal penicillinase. Further study also indicated that cephalosporin C had low toxicity in mice, but unfortunately, cephalosporin C had relatively low in vivo antibacterial activity.

In 1961, the structure of cephalosporin C was proposed by Abraham and Newton (6) and confirmed by an X-ray crystallographic analysis by Hodgkin and Maslen (7). It was found that a small quantity of the nucleus of cephalosporin C, 7-aminocephalosporanic acid (7-ACA) (Fig. 1), could be produced by gentle acid hydrolysis of cephalosporin C. Acylation of 7-ACA yielded a cephalosporin with much higher antibacterial activity than cephalosporin C against gram-positive bacteria. In 1962, Morin et al. (8) in the Lilly Research Laboratories developed a chemical method for the production of 7-ACA in kilogram quantities, and this opened the way for the extensive exploration of the synthesis of semisynthetic cephalosporins (discussed

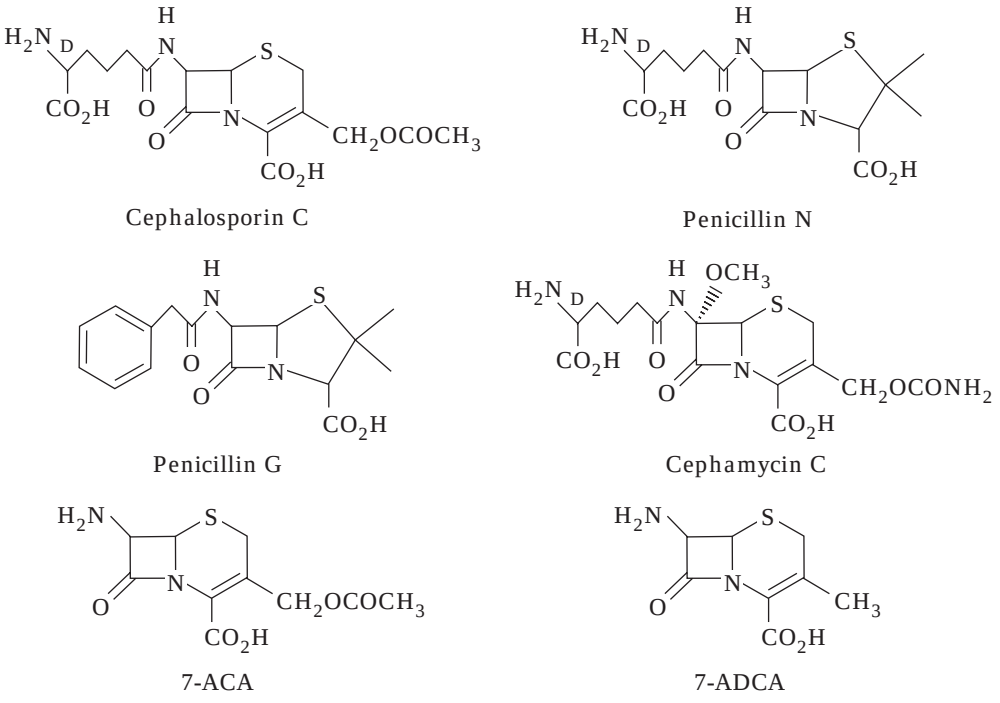


Figure 1. Chemical structure of β-lactams.

in a later section). By 1970, total synthesis of the cephalosporin double-ring system had been achieved by chemists working at Roussel Laboratories and E.R. Squibb & Sons (9).

The extensive use of semisynthetic cephalosporins resulted in the selection in nature of gram-negative bacteria containing R-plasmids that carried a gene encoding β-lactamase. This β-lactamase can hydrolyze most penicillins and all the semisynthetic cephalosporins that had been introduced in the mid 1960s. While searching for new β-lactam antibiotics that are stable in the presence of β-lactamases, culture filtrates of *Streptomyces* species yielded cephamycins (10) containing a methoxy group at the C-7α position of the cephalosporin nucleus (Fig. 1). These compounds were found to be completely resistant to the plasmid-encoded β-lactamase. Derivatives of cephamycin C will be described in "Semisynthetic Cephalosporins".

MODE OF ACTION OF β-LACTAMS

Cephalosporins and other β-lactams inhibit bacterial growth by interfering with bacterial cell wall synthesis. All bacterial cell walls have a protective layer of peptidoglycan at the outer surface of the cell membrane. The peptidoglycan layer consists of chains of alternating *N*-acetylglucosamine and *N*-acetylmuramic acid connected by peptides. These peptides cross-link adjacent peptidoglycan chains to form a net of peptidoglycan layers. In gram-positive bacteria, the cell walls are composed of peptidoglycan and teichoic acids (a polymer of glycerol phosphate or ribitol phosphate) or teichuronic acids (phosphate-free polysaccharides). In gram-negative bacteria, in addition to the peptidoglycan layer, there is a distinct outer membrane layer consisting of lipoproteins and lipopolysaccharides. During the course of bacterial cell wall synthesis, the monomers of *N*-acetylglucosamine and *N*-acetylmuramic acid peptide are linked to form the peptidoglycan chains, and three enzymes (transpeptidase, DD-carboxypeptidase, and endopeptidase) catalyze the cross-links between these peptidoglycan chains by forming peptide bridges (11). Binding of β-lactam antibiotics to the transpeptidase and the DD-carboxypeptidase enzymes, members of the penicillin-binding proteins (PBP) group, interferes with the formation of peptide bridges between peptidoglycan chains and thus inhibits the final stage of cell wall synthesis in bacteria (12).

BIOCHEMISTRY AND GENETICS OF CEPHALOSPORIN BIOSYNTHESIS

The biosynthetic pathways of cephalosporins have been well characterized (13). The current view of the overall biosynthetic scheme for cephalosporin C and cephamycin C is depicted in Figure 2. The nomenclature of genes encoding the enzymes involved in each step of the biosynthesis of cephalosporins follows the recommendation by Ingolia and Queener (14).

Biosynthetic Pathways for Cephalosporins

The biosynthesis of cephamycin C by *Streptomyces clavuligerus* and *Nocardia lactamdurans* and the biosynthesis of cephalosporin C by *A. chrysogenum* all begin with the

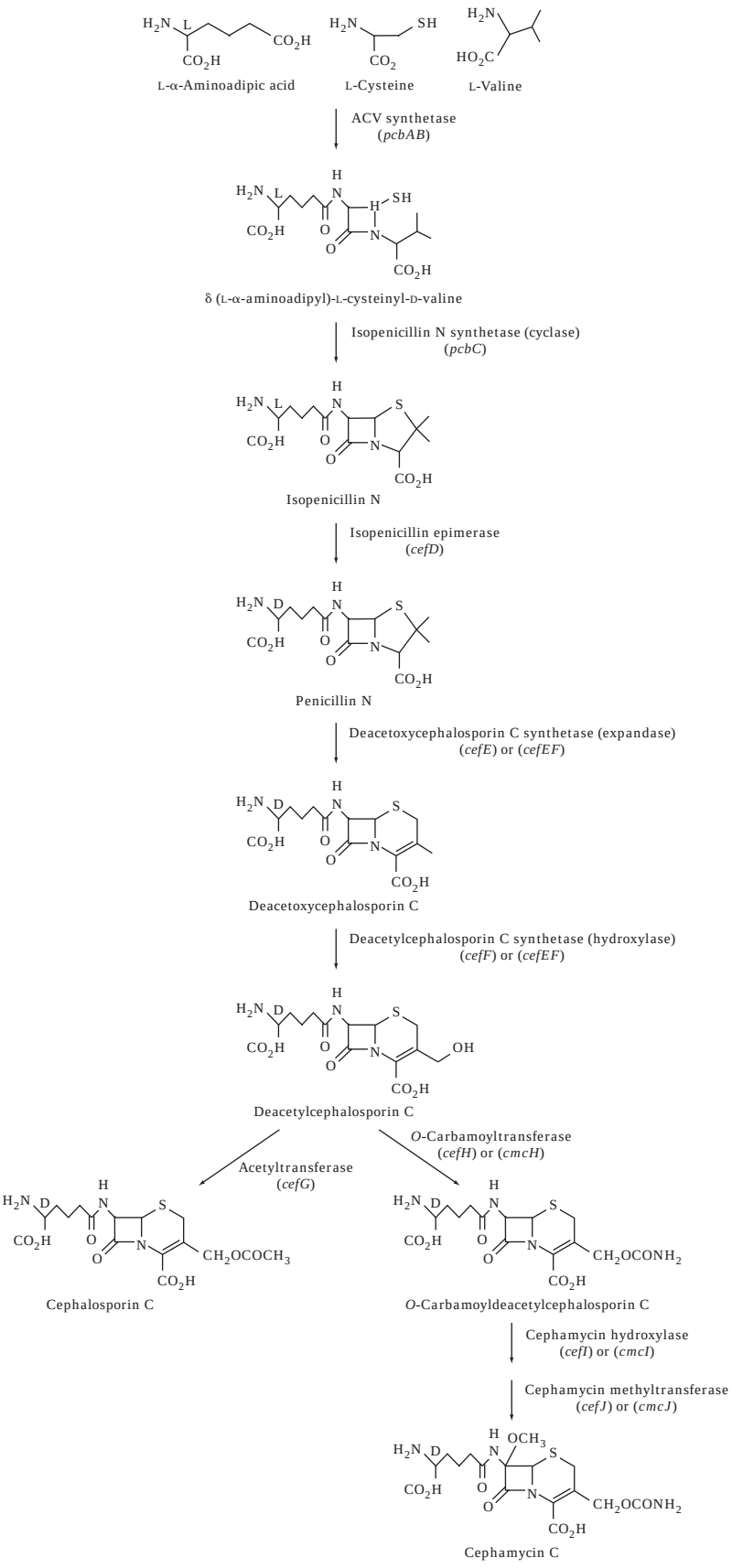


Figure 2. Biosynthetic pathways for cephalosporin C and cephamycin C.

condensation of three amino acids, L-α-aminoadipic acid, L-cysteine, and L-valine, to form the tripeptide δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (LLD-ACV) by the ACV synthetase enzyme. The racemization of valine from the L- to the D-configuration appears to take place during the synthesis of the tripeptide. The LLD-ACV tripeptide is cyclized to form the penam nucleus, isopenicillin N, by isopenicillin N synthetase (or cyclase). Both the LLD-ACV tripeptide and isopenicillin N also serve as intermediates in the biosynthesis of penicillins. The genes coding for enzymes involved in the biosynthesis of the common intermediates LLD-ACV and isopenicillin N are designated *pcb* (*p*enicillin/*c*ephalosporin *b*iosynthesis). The genes coding for the remaining enzymes involved in the biosynthesis of cephalosporins are designated *cef* (14).

The isopenicillin N epimerase (IPNE) catalyzes the conversion of the L-α-aminoadipyl side chain of isopenicillin N into the D-α-aminoadipyl side chain of penicillin N. The five-membered thiazolidine ring of penicillin N is then expanded to give the six-membered dihydrothiazine ring of deacetoxycephalosporin C by deacetoxycephalosporin C synthetase (expandase). The next enzyme in the pathway, deacetylcephalosporin synthetase (hydroxylase), catalyzes hydroxylation of deacetoxycephalosporin C at the C-3 methyl group to yield deacetylcephalosporin C. In *A. chrysogenum*, both the ring expansion and hydroxylation activities reside on the same protein, which is coded by one gene. Unlike fungi, the expandase and the hydroxylase are two distinct enzymes in *S. clavuligerus* and *N. lactamdurans* and encoded by two separate genes (13).

In *A. chrysogenum*, the last step in cephalosporin C biosynthesis, catalyzed by cephalosporin C synthetase (acetyltransferase), involves the transfer of an acetyl group from acetyl coenzyme A to the C-3 hydroxymethyl moiety of deacetylcephalosporin C.

In the cephamycin C producers *S. clavuligerus* and *N. lactamdurans*, the *O*-carbamoyltransferase catalyzes the transfer of a carbamoyl group from carbamoylphosphate to the C-3 hydroxymethyl moiety of deacetylcephalosporin C. The *O*-carbamoyldeacetylcephalosporin C is then hydroxylated at the C-7α position by cephamycin hydroxylase. The final biosynthetic step to cephamycin C (7α-methoxy-*O*-carbamoyldeacetylcephalosporin C) is the transfer of the methyl group from *S*-adenosylmethionine (SAM) to the C-7α hydroxyl moiety of 7α-hydroxyl-*O*-carbamoyldeacetylcephalosporin C by cephamycin methyltransferase. The biosynthetic pathways of other cephamycins, such as cephamycin A and cephamycin B, have not been fully investigated.

Cloning of the Cephalosporin Biosynthetic Genes

Because the DNA transformation system developed for *A. chrysogenum* is much less efficient than the bacterial systems, cloning of the cephalosporin C biosynthetic genes by complementation of blocked mutants is not feasible. However, the reverse genetics approach has been used successfully in the cloning of a number of cephalosporin C biosynthetic genes (14). This approach begins with the purification of the desired enzyme, determination of a partial amino acid sequence, and synthesis of oligonucleotide probes based on the amino acid sequence. A genomic DNA library of *A. chrysogenum* is then constructed in *Escherichia coli* and screened for clones that hybridize to the DNA probes. Sequencing of the cloned DNA will confirm that the DNA sequence matches the known amino acid sequence. The reverse genetics approach has also been used in the cloning of the cephamycin C biosynthetic genes despite the existence of highly efficient transformation systems developed for *S. clavuligerus* and *N. lactamdurans*.

In 1985, the first β-lactam biosynthetic gene, *pcbC* (encoding cyclase), was cloned from *A. chrysogenum* (15). The *A. chrysogenum pcbC* gene has no intron and encodes a polypeptide of 338 amino acids. When the open reading frame of the *A. chrysogenum pcbC* gene was expressed in an *E. coli* expression vector, the purified recombinant cyclase protein was able to convert LLD-ACV to isopencillin N with identical kinetics to the authentic protein, and it also converted ACV tripeptide analog substrates to novel β-lactam antibiotics. Recombinant cyclase is readily produced in the recombinant *E. coli* and can be used for in vitro synthesis of novel β-lactam antibiotics (16). Other *pcbC* genes have been cloned by reverse genetics or heterologous hybridization from *Aspergillus nidulans, Penicillium chrysogenum, S. clavuligerus, S. griseus, S. jumonjinensis, S. lipmanii, N. lactamdurans, Flavobacterium* sp., and *Lysobacter lactamgenus* (17). When these genes were compared at the DNA sequence and the predicted amino acid sequence levels, a high degree of similarity was found to exist among the *pcb*C genes from different sources.

The reverse genetics approach was also used to clone the *pcbAB* (encoding ACV synthetase), *cefD* (isopenicillin N epimerase), *cefE* (expandase), *cefF* (hydroxylase), *cefEF* (*A. chrysogenum* expandase–hydroxylase bifunctional enzyme), *cefG* (acetyltransferase), *cefH* (also known as *cmcH*, *O*-carbamoyltransferase), *cefI* (*cmcI*, cephamycin hydroxylase), and *cefJ* (*cmcJ*, cephamycin methyltransferase) genes (14,17–20). The only gene of the cephalosporin biosynthetic genes that has not been cloned is the *cefD* gene from *A. chrysogenum*. Most of the cephalosporin C biosynthetic genes cloned from *A. chrysogenum* do not have introns and have a high degree of similarity with their bacterial counterparts, suggesting a bacterial origin for these genes. The *cefG* gene is the exception with two introns. All the fungal genes and some of the bacterial genes encoding for the enzymes involved in cephalosporin biosynthesis have their own promoter for transcriptional control. However, the *Streptomyces cefD* and *cefE* genes are transcribed from the same promoter, with the *cefD* gene located upstream of the *cefE* gene and a 130 base pair (bp) intergenic region separating them (18). In *N. lactamdurans*, a 7-bp separation between the *cefI* and *cefJ* genes and the presence of a putative ribosome binding site upstream of the *cefI* gene suggests that the two genes are cotranscribed from the same promoter and also cotranslated (19).

Organization of the Cephalosporin Biosynthetic Genes

In 1988, Chen et al. demonstrated that a cloned 29.3-kb DNA fragment from *S. cattleya* carried the entire set of genes encoding enzymes involved in cephamycin C biosyn-

thesis (21). No further characterization of this cloned DNA fragment was reported. Subsequent work on other bacterial cephamycin C producers also indicated that the cephamycin C biosynthetic genes are clustered in one region of the bacterial chromosome (17–20) (Fig. 3a). However, in *A. chrysogenum*, the two early cephalosporin C biosynthetic genes, *pcbAB* and *pcbC*, reside on a different chromosome from the two late biosynthetic genes, *cefEF* and *cefG*. Eight chromosomes have been detected by pulsed-field electrophoresis, an electrophoretic technique that allows separation of chromosomes in an agrose gel system. The *pcbAB* and *pcbC* genes are located on chromosome VI and transcribed in opposite directions. The *cefEF* and *cefG* genes are located on chromosome II and are also transcribed in opposite directions (17) (Fig. 3b). Because the *cefD* gene of *A. chrysogenum* has not been cloned, it is not known whether it is linked to either of the two clusters.

STRAIN IMPROVEMENT

Mutagenesis and Selection

Most of the literature describing strain improvement reports the use of the cephalosporin C producer *A. chrysogenum*. However, similar strategies could be applied to the cephamycin C producers (i.e., *Streptomyces* spp.). Improvement in the yield of cephalosporin C has been achieved primarily by increasing the productivity of the *A. chrysogenum* culture by traditional strain mutation and selection techniques. The first significant titer improvement was achieved when in 1959 mutant 8650 was isolated. This strain enabled 100 g of cephalosporin C to be purified for structure determination and is the parent of all the industrial production strains (1).

Mutagenesis, followed by high-throughput screening of the culture, has been the traditional means of strain improvement. Conidia are prepared from the mycelial culture to increase the probability of isolating mutants derived from a single nucleus. Chemical mutagens, such as *N*-methyl-*N*-nitro-*N*-nitrosoguanidine and ethyl-methanesulfonate, or *UV* light are used to induce chromosomal mutations. Growth of the mutants on selective agents can be used to increase the probability of isolating improved clones. Toxic analogs of cysteine, valine, aminoadipate, lysine, and methionine are used to select mutants that overproduce these amino acids, providing a larger precursor pool for β-lactam biosynthesis. The toxicity to *A. chrysogenum* of heavy metallic ions, such as Cu^{2+}, Mn^{2+}, and Hg^{2+}, has also been exploited for selection purposes. Because the β-lactam antibiotics complex with these ions, one can theorize that mutants resistant to inhibitory concentrations of Cu^{2+}, Mn^{2+}, or Hg^{2+} can overproduce the β-lactam nucleus to detoxify these metallic ions. Elander et al. demonstrated the isolation of superior cephalosporin-C-producing strains from mutants resistant to mercuric chloride (22).

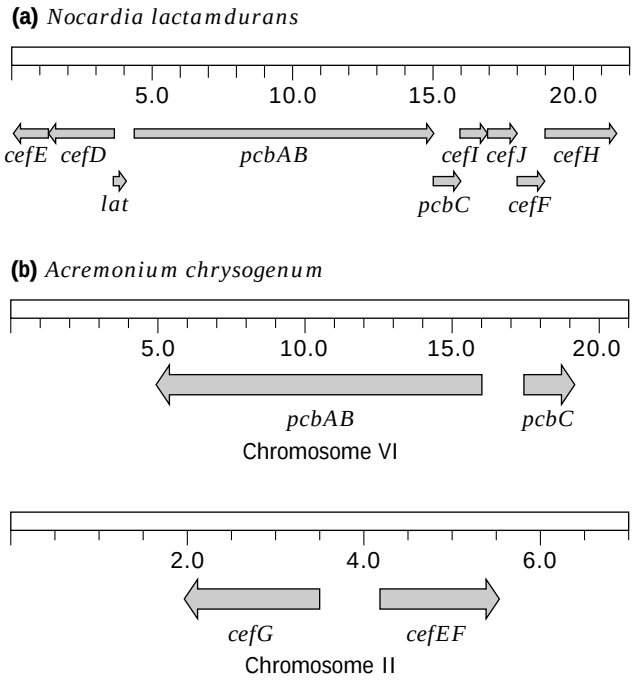


Figure 3. Organization of the cephalosporin biosynthetic genes. (**a**), Physical map of a 25-kb region of *N. lactamdurans* DNA carrying all genes related to cephamycin C biosynthesis. *lat* = gene encoding lysine-6-aminotransferase; *pcbAB* = ACV synthetase; *pcbC* = isopenicillin N synthetase (cyclase); *cefD* = isopenicillin N epimerase; *cefE* = deacetoxycephalosporin C synthetase (expandase); *cefF* = deacetylcephalosporin C synthetase (hydroxylase); *cefH* or *cmcH* = *O*-carbamoyltransferase; *cefI* or *cmcI* = cephamycin hydroxylase; *cefJ* or *cmcJ* = cephamycin methyltransferase (17–20). (**b**), Physical map of the clusters of the cephalosporin C biosynthetic genes of *A. chrysogenum*. *cefEF* = expandase/hydroxylase bifunctional enzyme; *cefG* = acetyltransferase (17). The direction of arrows indicates the transcriptional direction of a gene.

Protoplast Fusion

Protoplast fusion has been explored as a means of combining the advantageous characters of different culture lineages (e.g., high titer and fast growth). Mycelia are treated with dithiothreitol (DTT), followed by Novozym™ 234 (a multienzyme preparation from the fungus *Trichoderma harzianum* available from Novo Biolabs) in the presence of an osmotic stabilizer, usually 0.7 M NaCl. Polyethylene glycol (PEG) is then used to induce membrane fusions between protoplasts; the fusants are identified by regeneration on selective media. Auxotrophic markers have been used to select for fusants, but auxotrophic mutants often have reduced titers and are therefore not suitable for industrial strain improvement. Strains that have been transformed with positive selective markers such as phleomycin and hygromycin resistance are better candidates for protoplast fusion.

Genetic Engineering

In the past 10 years, the cloning of many of the genes involved in the biosynthetic pathway of cephalosporins has led to attempts to improve strains through genetic engineering. Enzymes that have been identified as rate-limiting in the biosynthesis of cephalosporins or their precursors are promising targets for strain improvement through genetic engineering (23). Transformation of industrial strains with additional copies of the genes coding

for these rate-limiting enzymes can result in alleviation of a metabolic bottleneck and increase antibiotic production. Transformation methods for *A. chrysogenum* have been developed using protoplasts and PEG treatment.

Based on the kinetic information available for the enzymes in the cephalosporin C biosynthetic pathway in *A. chrysogenum* and the cephamycin C pathway in *S. clavuligerus*, models have been developed to predict the flux of intermediates in these pathways. The model developed for *S. clavuligerus* predicted that the amino acid precursors, in particular α-aminoadipate, were limiting. Introduction of additional copies of the lysine ϵ-aminotransferase gene *(lat)*, required for conversion of lysine to α-aminoadipate in *Streptomyces*, resulted in a fivefold increase in cephamycin C production (24).

Skatrud et al. have transformed industrial strains of *A. chrysogenum* with a recombinant plasmid containing the hygromycin resistance gene and the *cefEF* gene encoding the expandase–hydroxylase bifunctional enzyme. At the shake flask scale, a 39% increase in cephalosporin C titer was observed, with a significant reduction in the levels of penicillin N and deacetoxycephalosporin C produced. The improvement in cephalosporin C titer at the 150-L pilot plant scale was 15%. Enzyme assays indicated that the expandase–hydroxylase activity in the transformant was twofold higher than in the host strain, and Southern blot analysis revealed a single chromosomal integration at a heterologous position (25). Interestingly, when the *cefG* gene (encoding acetyltransferase) was identified, it was found to be adjacent to the *cefEF* gene and had incidentally been included on the genomic fragment that Skatrud's group had used for transformation. Chiang and Basch also reported that recombinant *A. chrysogenum* strains transformed with a plasmid containing the *cefEF* gene had a 40% reduction of deacetoxycephalosporin C present in the cephalosporin C fermentation broth as compared to the parental strain (26). This process was scaled successfully to production tanks.

CEPHALOSPORIN FERMENTATION

Physiology of Cephalosporin C Fermentation

Cephalosporin C is produced by *A. chrysogenum* as a secondary metabolite, with production of the antibiotic associated with a distinct morphological change. When no nutrient stress is placed on the culture, the organism grows as branched mycelia by extension of the hyphal tips, and little or no cephalosporin C is produced. Under glucose-limiting conditions and in the presence of DL-methionine, the mycelial tips swell and form arthrospores. The mycelial cell walls weaken and fragment, releasing the arthrospores into the media. Throughout the formation of the arthrospores, cephalosporin C is produced, with the swollen hyphae being the most productive morphology.

DL-Methionine addition significantly increases cephalosporin C production. This effect is independent of the methionine role as the sulfur donor for the formation of cysteine. The presence of DL-methionine induces small increases in the enzyme activities of expandase and cyclase and a large increase in the ACV synthetase activity (27). In the presence of high concentrations of glucose or glycerol in fermentation media, the cephalosporin C productivity decreases dramatically. The high concentration of glucose represses the production of expandase, but not of cyclase or ACV synthetase (28).

Large-Scale Fermentation

Cephalosporin C is produced industrially by fed-batch fermentation of *A. chrysogenum*. The pH is maintained between 6.0 and 7.0 in a temperature range of 24 to 28 °C. The media must supply carbon and nitrogen for growth, but must also induce the differentiation of the culture required for antibiotic production. The carbon source may be supplied as separate simple and complex carbohydrates feeds, enabling better control of growth and glucose levels. Monosaccharides, particularly glucose, promote rapid growth of the culture but reduce antibiotic titer. Simple sugars may be batched into the media or fed at higher rates earlier in the fermentation. Galactose and sucrose support slower growth than glucose but a higher rate of specific productivity. As the fermentation progresses, this sugar feed is reduced and a larger proportion of carbon is supplied as soybean or peanut oil. This limits the availability of glucose and allows the formation of arthrospores and cephalosporin C. The oil may also serve as a surfactant to reduce potential foaming of the culture.

The organic nitrogen is usually supplied by some combination of agricultural by-products, such as soybean and cottonseed meals, and supplemented with an inorganic nitrogen source such as ammonium sulfate. Ammonia gas and ammonium hydroxide can also be used to supply inorganic nitrogen and to provide pH control. Corn steep liquor is commonly used in the medium as an inexpensive source of amino acids, trace elements, and vitamins.

DL-Methionine induces the formation of arthrospores in *A. chrysogenum*, which as mentioned earlier is correlated with cephalosporin C production. Swartz et al. examined the growth and productivity of fed-batch cultures of *A. chrysogenum*. Batch addition of methionine at 3 g/L interfered with sucrose utilization of these cultures but did not affect glucose utilization. Methionine was found to be a competitive inhibitor of the invertase enzyme required for the metabolism of sucrose. Optimal productivity was achieved when sucrose and DL-methionine were fed in an exponentially ramped feed. Surprisingly, feeding of DL-methionine increased the productivity of strains that were considered methionine independent, meaning batched methionine had no effect on titer (29).

As with all fungal fermentations, dissolved oxygen is an important physical parameter. It is particularly important in the fermentation of cephalosporin C, because the cyclase and particularly the expandase enzymes in the biosynthetic pathway require oxygen for effective catalysis. If the dissolved oxygen level drops below 20%, penicillin N will accumulate and the output of cephalosporin C will be reduced.

Chemical breakdown of cephalosporin C to 2-(D-4-amino-4-carboxybutyl)-thiazole-4-carboxylic acid (compound X) (Fig. 4) results in the loss of ~40% of the cephalosporin C produced, a significant loss of the biosynthetic

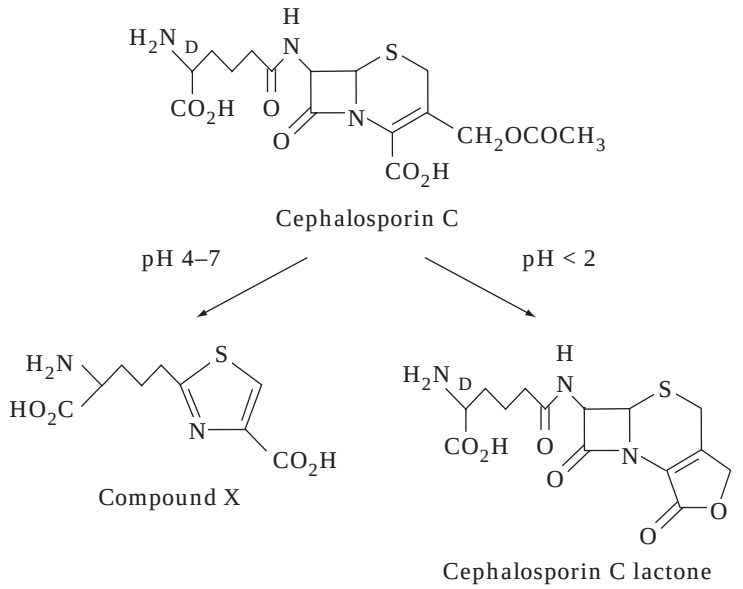


Figure 4. Chemical degradation of cephalosporin C.

potential of the organism (30). Deacetylcephalosporin C, however, is far more stable in the fermentation broth in comparison to cephalosporin C. Addition of an acetyl esterase produced by *Rhodosporidium toruloides* to the fermentation results in the accumulation of deacetylcephalosporin C and an increase in antibiotic yield of ~40% (31).

EXTRACTION OF CEPHALOSPORINS

Cephalosporin C

After the fermentation is complete, the mycelia and insoluble media components are usually removed by filtration or centrifugation. In the fermentation broth, besides cephalosporin C, there is also present small quantities of penicillin N, deacetoxycephalosporin C, and deacetylcephalosporin C. Cephalosporin C must be extracted from the fermentation broth. Under neutral and slightly acidic conditions, a temperature-dependent chemical conversion of cephalosporin C to compound X occurs in the fermentation broth (30). At a pH less than 2, cephalosporin C lactone is formed (32) (Fig. 4). To minimize this degradation, it is important that the cephalosporin C broth be processed rapidly, avoiding extreme pH conditions and keeping the temperature low.

Two main strategies are used for the purification of cephalosporin C. One strategy uses activated carbon column or nonionic resin column to adsorb cephalosporin C from fermentation broth. Because of the high selectivity of activated carbon and nonionic resin, cephalosporin C is preferentially adsorbed over penicillin N, deacetylcephalosporin C, and deacetoxycephalosporin C. A subsequent anion-exchange column gives high purity cephalosporin C product. Alternatively, a combination of anion- and cation-exchange columns arranged in series can be used to purify cephalosporin C directly from the fermentation broth (33).

Because cephalosporin C will be further derivatized to semisynthetic end products, modifications to the cephalosporin C molecule that do not affect the cephalosporin nucleus can be used to facilitate its purification. Therefore, an alternative strategy for the purification of cephalosporin C is substitution of the amine moiety on the α-aminoadipyl side chain of cephalosporin C. Some of the N-substituted derivatives, such as *N*-2,4-dichlorobenzoyl cephalosporin C and tetrabromocarboxybenzoyl cephalosporin C, can be crystallized in acidic aqueous solutions. Alternatively, salts can be formed between the N-substituted derivatives and organic base, such as dicyclohexylamine and dimethylbenzylamine, and they are solvent extractable. Several extraction steps are required to achieve an acceptable purity of the end product, although small amounts of penicillin N, deacetylcephalosporin C, and deacetoxycephalosporin C derivatives are still present (33). N-Substituted cephalosporin C derivatives can be converted to 7-ACA as effectively as cephalosporin C can.

Cephamycin C

The clarified cephamycin C fermentation broth is usually adjusted to pH 2.5 to inactivate most of the penicillin N present. Then, cephamycin C can be isolated with a combination of activated carbon adsorption and anion-exchange resin adsorption. A small amount of deacetoxycephalosporin C and the remaining penicillin N copurifies with cephamycin C. These can be removed by a subsequent silica gel column (34).

PRODUCTION OF THE CEPHALOSPORIN NUCLEUS

Cephalosporin C itself has a relatively low level of antibacterial activity. The semisynthetic cephalosporins with higher antibacterial activity, β-lactamase resistance, and desirable pharmacological properties are derived by modifications of 7-ACA or 7-aminodeacetoxycephalosporanic acid (7-ADCA) (Fig. 1).

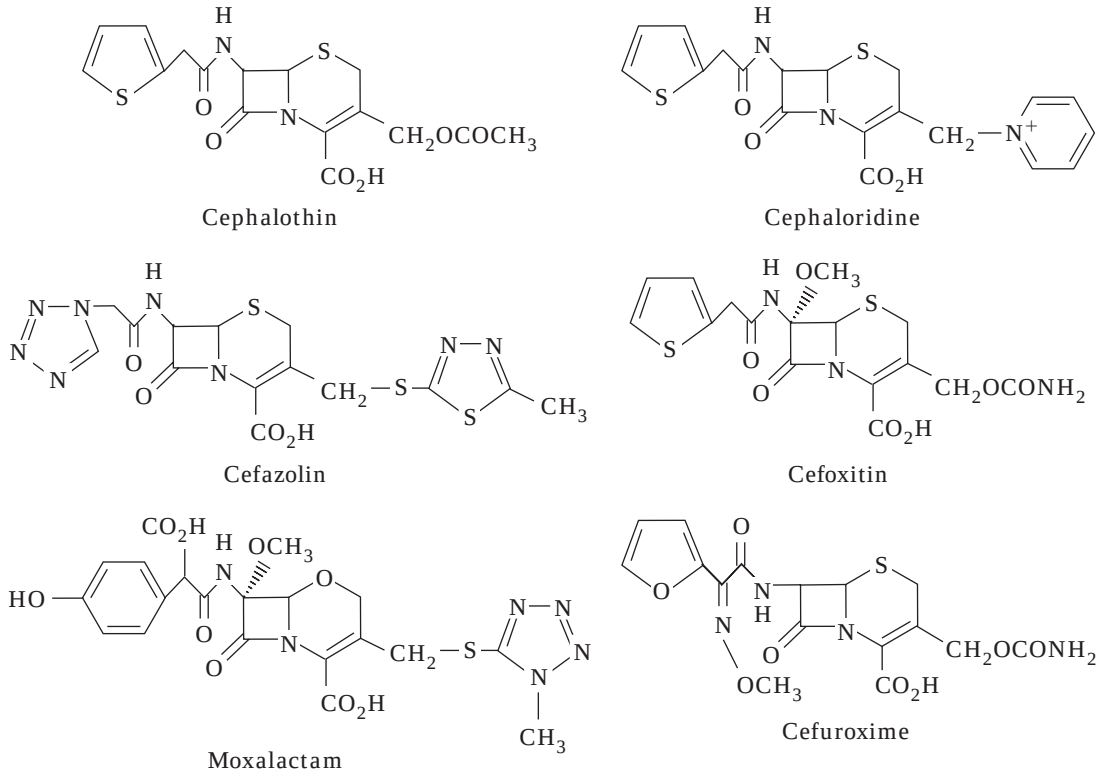


Figure 5. Chemical structure of first- and second-generation cephalosporins.

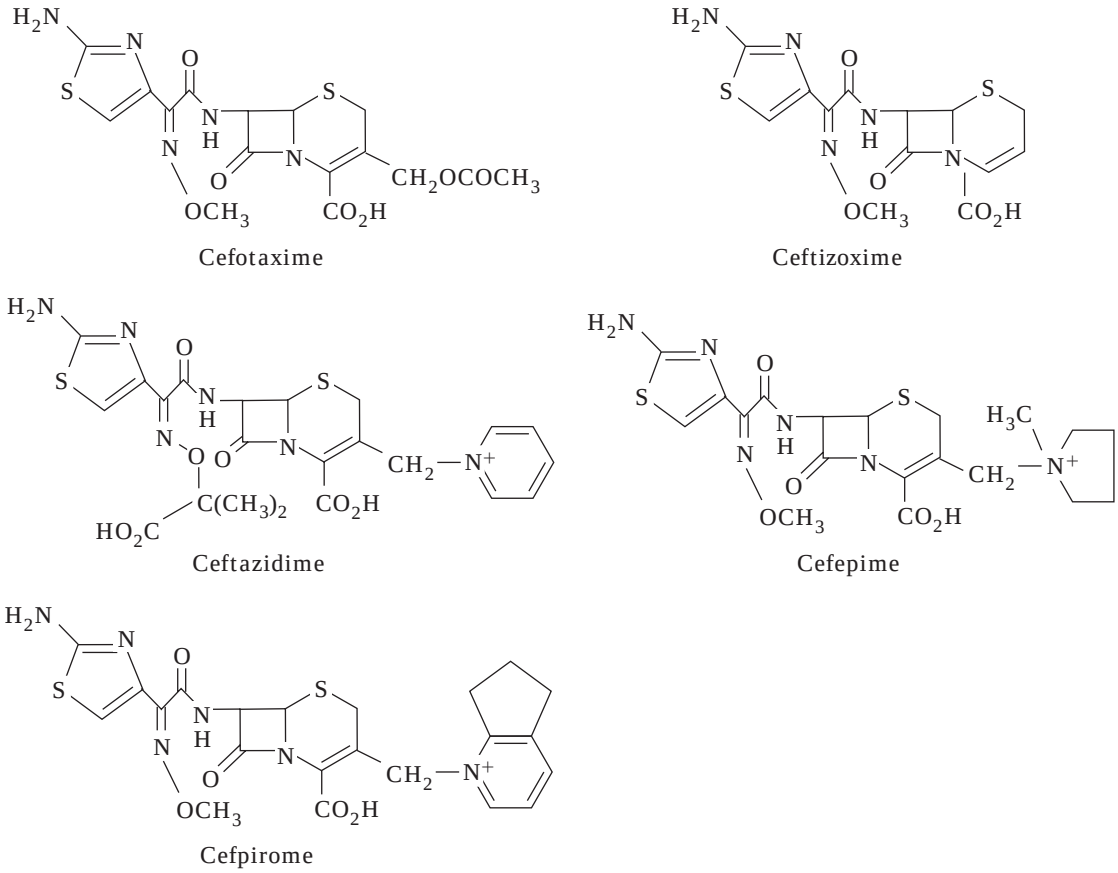


Figure 6. Chemical structure of third- and fourth-generation cephalosporins.

Almost one-third of commercial cephalosporins are derived from 7-ADCA. Because of the lower cost of fermenting penicillins, 7-ADCA is most economically produced from penicillin G by ring expansion of a penicillin sulfoxide ester to yield a cephalosporin ester. The ester group is removed, followed by removal of the phenylacetyl side chain to give 7-ADCA (35). Recently, Merck and Panlabs patented a process for producing 7-ADCA. In this process, *P. chrysogenum* was transformed with the *S. clavuligerus cefE* (expandase) gene. When fed with disodium adipate, this strain produced adipoyl 6-APA, which was then converted by the recombinant expandase to adipoyl 7-ADCA in vivo. The adipoyl side chain was subsequently cleaved in vitro by a *Pseudomonas* adipoyl acylase to produce 7-ADCA (36).

Two-thirds of the commercial cephalosporins are derived from 7-ACA. 7-ACA is produced from cephalosporin C by chemical or enzymatic deacylation. In the chemical process, after protection of the amino and carboxyl groups of cephalosporin C, reaction with PCl_5 in the presence of base yields an iminochloride derivative. Addition of alcohol is used to form the imino ether, which can then be hydrolyzed to form 7-ACA (37).

Alternatively, conversion of cephalosporin C to 7-ACA can be performed enzymatically in two steps. The first step is reaction of the α-aminoadipyl group with D-amino acid oxidase to produce glutaryl-7-ACA. This reaction proceeds through a keto-7-ACA intermediate that undergoes an oxidative decarboxylation in the presence of H_2O_2. Next, a glutaryl acylase is used to remove the glutaryl side chain to produce 7-ACA. An 85% yield of this reaction can be achieved using an immobilized enzyme system (38). Acai et al. coimmobilized the D-amino acid oxidase and the glutaryl acylase enzymes to perform a single-step conversion of cephalosporin C to 7-ACA (39).

SEMISYNTHETIC CEPHALOSPORINS

The outer membrane of gram-negative bacteria is impermeable to penicillins because of their hydrophobicity. Cephalosporins, however, are less hydrophobic, and therefore can pass through the outer membrane to reach their targets. The natural product, cephalosporin C, has low antibacterial activity. Chemical modification at the C-7 position of cephalosporin C yields compounds with a much improved antibacterial activity. Since the development of an efficient method for preparation of 7-ACA, a large number of semisynthetic cephalosporins have been made in the past 30 years. The development of these compounds can be divided into four stages (40).

The first-generation semisynthetic cephalosporins introduced in the mid-1960s include cephalothin, cephaloridine, and cefazoline (Fig. 5). These cephalosporins are modified at the C-3 and C-7 positions. The antibacterial activity of these compounds is still limited because indole-positive *Proteus* species and opportunistic pathogens, such as *Pseudomonas aeruginosa*, are resistant because of the presence of β-lactamases. Since the late 1960s, an enormous effort has been devoted to the development of cephalosporins resistant to β-lactamases.

Out of this research, the second-generation cephalosporins were generated. Among this group, two types of compounds, cephamycins and 7-oximinoacyl cephalosporins, were developed. In 1971, cephamycins were discovered from actinomycetes (10). Cephamycin C has a 3-carbamoyl and a 7α-methoxy group, which confers a remarkable resistance to β-lactamases. Chemical derivation of cephamycin C led to the development of cefoxitin and moxalactam (41) (Fig. 5). The compound developed from 7-oximinoacyl cephalosporin is cefuroxime (41) (Fig. 5). Cefuroxime carries the same 3-carbamoyl-oxymethyl group as cephamycin C does.

In the search for better cephalosporins with broader antibacterial activity and resistance to different β-lactamases, the addition of a hydrophilic aminothiazole group to oximinocephalosporins produced the third-generation cephalosporins, such as cefotaxime, ceftizoxime, and ceftazidime (41) (Fig. 6). Again, the third-gener-

Table 1. U.S. Manufacturers' Trade Names for Semisynthetic Cephalosporins

Semisynthetic cephalosporins	Trade name (manufacturer)
First generation	
Cefadroxil*	Duricef (Bristol-Myers Squibb)
Cefazolin	Ancef (SmithKline Beecham), Kefzol (Lilly)
Cephalexin*	Biocef (International Ethical), Cephalexin (Biocraft, Lederle Standard Products, Novopharm USA, Warner Chilcott), Keflex (Dista), Keftab (Dista)
Cephradine*	Cephradine (Biocraft, Lederle Standard Products, Warner Chilcott)
Second generation	
Cefaclor*	Ceclor (Lilly)
Cefamandole	Mandol (Lilly)
Cefmetazole	Zefazone (Upjohn)
Cefonicid	Monocid (SmithKline Beecham)
Cefotetan	Cefotan (Zeneca)
Cefoxitin	Mefoxin (Merck & Co., Inc.)
Cefprozil*	Cefzil (Bristol-Myers Squibb)
Cefuroxime	Kefurox (Lilly), Zinacef (Glaxo Wellcome)
Cefuroxime axetil*	Ceftin (Glaxo Wellcome)
Third generation	
Cefixime*	Suprax (Lederle)
Cefoperazone	Cefobid (Roerig)
Cefotaxime	Claforan (Hoechst-Roussel)
Cefpodoxime*	Vantin (Upjohn)
Ceftazidime	Ceptaz (Glaxo Wellcome), Fortaz (Glaxo Wellcome), Tazicef (SmithKline Beecham), Tazidime (Lilly)
Ceftrizoxime	Cefizox (Fujisawa)
Ceftriaxone	Rocephin (Roche Laboratories)
Loracarbef*	Lorabid (Lilly)
Four generation	
Cefepime	Maxipime (Bristol-Myers Squibb)

Note: Orally administered cephalosporins are marked with an asterisk. All others are administered by intramuscular or intravenous injection.

ation cephalosporins have been compromised by the rapid emergence of resistant bacterial strains with plasmid-encoded β-lactamases and by their ineffectiveness against strains of *Enterobacter cloacae* and *P. aeruginosa* with constitutively expressed chromosomal β-lactamase.

By the late 1980s, the fourth-generation cephalosporins, cefepime and cefpirome (Fig. 6), had been developed. The new compounds retain the aminothiazole moiety at C-7 with an oxyimino substitution and contain a positively charged quaternary nitrogen at the C-3 position (40). Table 1 summarizes the semisynthetic cephalosporins available in the U.S. market (40,42).

CONCLUSION

Since the discovery of cephalosporin C 50 years ago, rapid progress has been made in the strain improvement and fermentation process development for the production of cephalosporin C. Genetic manipulation of *A. chrysogenum* and *S. clavuligerus* allows the potential for engineering of cephalosporin biosynthetic enzymes with broader substrate specificity for in vivo or in vitro production of new β-lactams. Because of the continuous selection of resistant bacteria carrying newly formed β-lactamases, the development of new semisynthetic cephalosporins remains a constant effort to find new compounds effective against these pathogens.

BIBLIOGRAPHY

1. S. Selwyn, in S. Selwyn ed., *The Beta-Lactam Antibiotics: Penicillins and Cephalosporins in Perspective*, Hodder & Stoughton, London, 1980, pp. 1–55.
2. M. Barber, *The Cephalosporin Business 1995–2005*, Michael Barber Assoc. Surrey, U.K., 1995.
3. H.S. Burton and E.P. Abraham, *Biochem. J.* **50**, 168 (1951).
4. W. Gams, *Cephalosporium-artige Schimmelpilze (Hyphomycetes)*, Gustav Fischer Verlag, Stuttgart, Germany, 1971.
5. G.G.F. Newton and E.P. Abraham, *Nature* **175**, 548 (1955).
6. E.P. Abraham and G.G.F. Newton, *Biochem. J.* **79**, 377–393 (1961).
7. D.C. Hodgkin and E.N. Maslen, *Biochem. J.* **79**, 393–402 (1961).
8. R.B. Morin, B.G. Jackson, E.H. Flynn, and R.W. Roeske, *J. Am. Chem. Soc.* **84**, 3400–3401 (1962).
9. S.L. Neidleman, S.C. Pan, J.A. Last, and J.E. Dolfini, *J. Med. Chem.* **13**, 386–388 (1970).
10. R. Nagarajan, L.D. Boeck, M. Gorman, R.L. Hamill, C.E. Higgens, M.M. Hoehn, W.M. Stark, and J.G. Whitney, *J. Am. Chem. Soc.* **93**, 2308–2310 (1971).
11. J. Baddiley, in M. Salton and G.D. Shockman eds., **β**-*Lactam Antibiotics- Mode of Action, New Developments, and Future Prospects*, Academic Press, New York, 1982, pp. 13–30.
12. J.L. Strominger, *Harvey Lect.* **64**, 171–213 (1970).
13. S.E. Jensen, *Crit. Rev. Biotechnol.* **3**, 277–301 (1986).
14. T.D. Ingolia and S.W. Queener, *Med. Res. Rev.* **9**, 245–264 (1989).
15. S.M. Samson, R. Belagaje, D.T. Blankenship, J.L. Chapman, D. Perry, P.L. Skatrud, R.M. Van Frank, E.P. Abraham, J.E. Baldwin, S.W. Queener, and T.D. Ingolia, *Nature* **318**, 191–194 (1985).
16. U.S. Pat. 5,472,853 (December 5, 1995), J.E. Baldwin (to University of Oxford).
17. J.F. Martin, S. Gutierrez, F.J. Fernendez, J. Velasco, F. Fierro, A.T. Marcos, and K. Kosalkova, *Antonie Van Leeuwenhoek* **65**, 227–243 (1994).
18. S. Kovacevic, M.B. Tobin, and J.R. Miller, *J. Bacteriol.* **172**, 3952–3958 (1990).
19. J.J.R. Coque, F.J. Enguita, J.F. Martin, and P. Liras, *J. Bacteriol.* **177**, 2230–2235 (1995).
20. W.O. Pat. Appl. 95/29253 (November 2, 1995), J.J.R. Coque, F.J. Enguita, J.L. Fuente, F.J. Llarena, P. Liras, and J.F. Martin (to Merck & Co, Inc.).
21. C.W. Chen, H.-F. Lin, C.L. Kuo, H.-L. Tsai, and J.F.-Y. Tsai, *Biotechnology* **6**, 1222–1224 (1988).
22. R.P. Elander and S.J. Chiang, in A. Prokop, R.K. Bajpai, and C.S. Ho eds., *Recombinant DNA Technology and Applications*, McGraw-Hill, New York, 1991, pp. 153–170.
23. J.J. Usher, D.W. Hughes, M.A. Lewis, and S.-J.D. Chiang, *J. Ind. Microbiol.* **10**, 157–163 (1992).
24. L.H. Malmberg, W.-S. Hu, and D.H. Sherman, *J. Bacteriol.* **175**, 6916–6924 (1993).
25. P.L. Skatrud, A.J. Tietz, T.D. Ingolia, C.A. Cantwell, D.L. Fisher, J.L. Chapman, and S.W. Queener, *Biotechnology* **7**, 477–484 (1989).
26. J. Basch and S.-J.D. Chiang, *J. Ind. Microbiol. Biotechnol.* **20**, 344–353 (1998).
27. J.-Y. Zhang, G. Banko, S. Wolfe, and A.L. Demain, *J. Ind. Microbiol.* **2**, 251–255 (1987).
28. J.-Y. Zhang, S. Wolfe, and A.L. Demain, *Curr. Microbiol.* **18**, 361–367 (1989).
29. R.W. Swartz, S.M. Vicik, A.J. Fedor, *Biotechnol. Prog.* **6**, 333–340 (1990).
30. J.J. Usher, M.A. Lewis, D.W. Hughes, and B.J. Compton, *Biotechnol. Lett.* **10**, 543–548 (1988).
31. A. Smith, in H.W. Blanch, S. Drew, and D.I.C. Wang eds., *Comprehensive Biotechnology*, vol. 3. *The Practice of Biotechnology: Current Commodity Products*, Pergamon Press, New York, 1985, pp. 168–185.
32. J. Konecny, E. Felber, and J. Gruner, *J. Antibiot.* **26**, 135–141 (1973).
33. M.E. Wildfeuer, in D. LeRoith, J. Schiloach, and T.J. Leahy eds., *Purification of Fermentation Products*, American Chemical Society, Washington D.C., 1985, pp. 155–174.
34. U.S. Pat. 3,719,563 (March 3, 1973), R.L. Hamill, C.E. Higgens, and M.M. Hoehn (to Eli Lilly & Co.).
35. R.R. Chauvette, P.A. Pennington, C.W. Ryan, R.D.G. Cooper, F.L. Jose, I.G. Wright, E.M. Heyningen, and G.W. Huffman, *J. Org. Chem.* **36**, 1259–1267 (1971).
36. U.S. Pat. 5,318,896 (June 7, 1994), M.J. Conder, P.C. McAda, and J.A. Rambosek (to Merck & Co., Inc.).
37. L.D. Hatfield, W.H.W. Lunn, B.G. Jackson, L.R. Peters, L.C. Blaszczak, J.W. Fisher, J.P. Gardner, and J.M. Dunigan, in G.I. Gregory ed., *Recent Advances in the Chemistry of β-Lactam Antibiotics*, Royal Society of Chemistry, Burlington House, London, 1981, pp. 109–124.
38. H.D. Conlon, J. Baqai, K. Baker, Y.Q. Shen, B.L. Wong, R. Noiles, and C.W. Rausch, *Biotechnol. Bioeng.* **46**, 510–513 (1995).
39. P. Acai, E. Michalkova, V. Bales, *Bioproc. Eng.* **14**, 105–109 (1996).

40. R. Periti, *J. Chemother.* **8**, 3–6 (1996).
41. C.E. Newall, *Chem. Br.* 976–980 (October 1987).
42. *Physicians' Desk Reference*, 50th ed., Medical Economics, Montvale, N.J., 1996.

See also METABOLITES, PRIMARY AND SECONDARY; SECONDARY METABOLITE PRODUCTION, ACTINOMYCETES, OTHER THAN STREPTOMYCES.

CHINESE HAMSTER OVARY CELLS, RECOMBINANT PROTEIN PRODUCTION

FLORIAN M. WURM
Swiss Federal Institute of Technology
Lausanne, Switzerland

KEY WORDS

Cell culture process
Dihydrofolate reductase
DNA integration
Gene stability
Large scale
Manufacturing
Methotrexate
Suspension cells
Transfection

OUTLINE

INTRODUCTION

Continuously propagatable mammalian cells, in general, and Chinese hamster ovary (CHO) cells, in particular, have become preferred host systems for the production of complex proteins for clinical applications. Nowadays, in pharmaoriented biotech companies, more protein species are being produced in CHO cells than in *E. coli* and in *Saccharomyces*, which were production hosts of first choice during the early years of applied DNA technology. These microbial organisms are capable of producing large amounts of recombinant protein. However, the decision to employ somewhat more complex and costly in-vitro cultures of recombinant mammalian cells has been driven by the need to obtain proteins with superior biological activity and more complex biochemical structures. Biological activity and, last but not least, the pharmacokinetic characteristics of recombinant proteins, when administered in therapy, frequently depend on a number of complex protein modifications, that is, proper folding and disulfide bridge formation, oligomerization, proteolytic processing, phosphorylation, and the addition of specific and complex carbohydrate groups. Most of these and other protein processing steps can be executed with high efficiency by mammalian cells.

CHO cells have become dominant in industry, and not only for historic reasons. Of course, they were chosen as the host substrate for the first recombinant protein from mammalian cells, human tissue plasminogen activator (Activase® tPA) (1), achieving market approval in the United States and other global markets. CHO cells are unique in many ways and exhibit a number of advantages, ranging from ease of introduction of exogenous DNA to the capacity for growth and production at very large scales. This article attempts to summarize and highlight arguments, issues, advantages, questions, and ongoing research, in a cursory way, for the industrial production of high-value proteins derived from CHO cells. In doing so, the development path followed to establish a high-yield process will be followed, from the construction and use of appropriate plasmid vectors to the planning and design of the cell culture manufacturing process. Issues raised and

discussed herein are not meant to be comprehensive, yet it is hoped that the most critical points impacting development efforts will be addressed.

PRINCIPLE OF DNA TRANSFER AND AMPLIFICATION IN CHO CELLS

The following gives a brief introduction into the underlying principles for use of CHO cells as recipients of protein-encoding DNA vectors, and the most popular way to improve productivity with these cells: experimentally induced gene amplification.

A mutant cell line lacking enzyme activity for dihydrofolate reductase, developed by Urlaub and Chasin (2), can be cultivated in the presence of glycine, hypoxanthine, and thymidine (GHT). When these cells are transfected with a functional expression vector for dihydrofolate reductase (DHFR vector), cell lines that have acquired a functional DHFR gene can be selected and expanded in media lacking GHT. A second (or third or more) expression cassette for a product of interest (e.g., a protein with therapeutic value) can be included in the DHFR plasmid or cotransfected as part of one or more additional vectors. Screening a number of "clones," or more correctly, cell lines derived from colonies, will lead to the identification of cells coexpressing the product of interest.

In a second and often time-consuming phase of cell culture work, expression can be augmented by exposing recombinant cells to the drug methotrexate (MTX), a folate derivative that blocks DHFR activity completely and irreversibly. Usually, within a period of 2–3 weeks, from a population of cells a few colonies emerge (while the majority of cells die) that are resistant to the drug due to elevated expression of DHFR. It was found that these cells frequently contain chromosomally integrated plasmid sequences in a higher copy number than observed in cells prior to exposure to MTX. Treatment with stepwise elevated concentrations of MTX can be repeated several times and may result in the isolation of cells that contain dramatically increased copy numbers of the transferred genes.

The same principle for MTX-driven amplification applies to other immortalised cell lines as well (3); however, DHFR-mediated DNA transfer and methotrexate-driven amplification works best with cells that lack an endogenous DHFR gene (activity) prior to transfection.

REGULATORY ISSUES: PRODUCT SAFETY, CELL CHARACTERIZATION, IDENTITY, AND CONSISTENCY FOR PRODUCT FORMATION

CHO cells, originally derived from the ovary of a female Chinese hamster, emerged as a spontaneously immortalized line, more than 40 years ago in the hands of Kao and Puck (4,5). The cells have been the object of many studies since then and are very well characterized with respect to a variety of aspects, including karyotype, chromosome structures, gene mapping, general culture conditions, cell physiology, and media requirements. For the purpose of genetic and physiological studies on the activity of dihydrofolate reductase and its genomic alleles, in the late 1970s, Urlaub and Chasin (2) mutagenized the original cell line with the help of radiation and chemicals. The goal was to create a mutant cell line lacking functional activity of both alleles of the DHFR gene locus. The resulting cell line was called CHO-DUX (sometimes also referred to as CHO-DUKX-B11) and, although not initially intended for this purpose, proved to be an ideal substrate for a number of pioneering gene transfer experiments (6–8).

The "climate" for permission by regulatory agencies, in particular the FDA in the United States, was not favorable in the early 1980s for using "tumor" cell lines, such as immortalized CHO cells, for the production of recombinant proteins. Discussions about risks associated with the use of mammalian cells were controversial, and had been initiated more than two decades earlier (8) when a first generation of "classical" biological products (i.e., vaccines and, later, natural interferons) were developed on the basis of primary monkey kidney cells, human diploid cells, and later, transformed mammalian cells. Manufacturers of recombinant proteins for clinical applications and regulatory agencies were in agreement that it was very important to minimize eventual risks associated with the use of recombinant mammalian cell hosts. Risks were seen in tumor principles, carried by the DNA of the host and in adventitious agents (viruses, mycoplasma, etc.), which could infect the host cell lines. Also, consistency and quality of the recombinant proteins were discussed in the context of risk assessment and risk control. The result of a decade-long discussion was that very stringent manufacturing controls and monitoring procedures were enforced as a prerequisite for manufacture of proteins from such cells. A balance had to be found between the almost assured clinical benefit of some of those first recombinant products and the perceived (real or imaginary) risks associated with the unavoidable necessity to produce them in "tumor" cells.

The first product, recombinant tissue plasminogen activator (Activase® rtPA), proved to be a good candidate to achieve such a balance because the benefit, saving lives of heart-attack patients, outweighed by far anticipated risks, for example, the debated risk of exposing patients to DNA sequences that may transfer a tumor-inducing principle. However, approval was achieved only after a huge amount of data was provided to the regulatory agencies showing that (1) consistently, only a very small quantity of CHO DNA (less then 10 pg/dose recently related to less than 100 pg/dose) was present in the final product, (2) the product itself could be produced with a very high degree of reproducibility, and (3) it was produced with a purity not achieved before in any biological derivation from mammalian cell culture.

Adventitious Agents: Bacterial and Fungal Contamination

Prevention of airborne and other bacterial or fungal infections of cell culture and recovery systems can be assured, to a very high degree of confidence, by the use of a piping and vessel system that maintains absolute containment of the sterile medium fluids. The equipment used must be of a nature that allows cleaning and sterilization by clean-in-place (CIP) and sterilize-in-place (SIP) procedures (usually using high-quality stainless steel). Most cell culture pro-

cesses require complex media containing amino acids, vitamins, protein hormones, and fetal bovine serum. Some of these components cannot be autoclaved. To exclude, therefore, the introduction of bacterial and fungal contamination through these raw materials, prior testing and, in addition, filtration through membranes of 0.2-μm-or even 0.1-μm-pore-size into presterilized containers is employed.

Rigorous testing cell banks (Master Seed Cell Bank, MSCB, Manufacturers Working Cell Bank, MWCB) (for review see Ref. 10), done by analyzing cells of a number of representative cryovials, which represent the bank when frozen in liquid nitrogen, assures that the production cell line itself is not contaminated with viruses, bacteria, mycoplasma, and fungi. Sterility testing of the cell line must be done in appropriate media lacking antibiotic or antifungal compounds, for obvious reasons. Virus testing is done in suitable cell systems that are validated for each of the individual virus species.

Adventitious Agents: BSE

The use of sera or other products derived from bovine sources represents a potential risk of transfer of the causative agent for bovine spongiformous encephalopathy (BSE) to patients receiving recombinant proteins.

Regulatory agencies therefore request detailed information on the origin and processing by the vendor of bovine derived products. For example, sera obtained from countries in which BSE was diagnosed, even in a small number of animals, are considered unacceptable by regulatory agencies. Companies have, in most cases anticipating these regulations, assured their supply of bovine-derived process materials from countries like New Zealand or the United States where BSE has not yet been reported. Recently, due to the high degree of concern raised in Europe, but also in the United States, very stringent process regulations have been discussed. These regulations would restrict to a very large extent, or prevent altogether, the use of bovine-derived process additives, not taking into account the processing of the materials by the vendor or the receiving company. The high degree of concern regarding BSE is based on the possibility that *(1)* a very small proteinaceous compound could be responsible for the disease, *(2)* transfer to the human population as Creutzfeld–Jacob disease has been considered possible, and *(3)* detection of the causative agent is extremely difficult, even with the most sophisticated methods available today. It appears that the complexity of the involved issues will require further discussion in the future and must occur at an international level.

Adventitious Agents: Viral Contaminants

Why were hamster cells chosen as a host for making human recombinant proteins? A strong argument, favoring nonhuman cell lines over human cell lines for the production of proteins, is the fact that certain life-threatening human viruses cannot be propagated, or grow poorly, in nonhuman cell lines. CHO cells have been found to not support the replication of pathogenic viruses such as polio, herpes, hepatitis B, HIV, measles, adenoviruses, rubella, and influenza. Thus, the risk of a viral adventitious agent, belonging to this group of viruses, involuntarily carried along with the product of interest can be considered extremely low. Wiebe et al. tested a total of 44 human pathogenic viruses for replication in CHO cells and found only seven (Reo 1,2,3; mumps; and parainfluenza 1,2,3) that were able to infect these cells (11). Exclusion of these few virus species, as well as some other viruses that can be propagated on CHO cells, such as the minute virus of the mouse (MVM), can be assured to a very high degree of certainty through testing. Tests can be performed with all materials that enter the manufacturing process and would support the viability of the virus in question. Tests are also obligatory with fluids that contain the product of interest and that represent, for example, toward the end of the cell culture production cycle, a manufactured product unit, defined as a "batch."

Sterility filtration of fluids containing a variety of raw materials, some of which may have been exposed to viruses, does not prevent the introduction of viral contaminants into the process of cell culture–based manufacturing of proteins. Only recently have membranes with pore sizes small enough to exclude passage of virus particles become available for industrial-scale operations. However, these membranes cannot be introduced into existing processes without complex consequences on regulatory issues. Testing therefore still appears to be the most efficient method to exclude viruses that may reside within the host cell line itself, or that could be introduced by biologically derived raw materials needed for the cell culture process. Specialized service companies have developed, in close collaboration with the pharmaceutical client companies, batteries of validated test procedures. These utilize cell culture systems in which supernatants or lysates of the production cells are cocultivated with the corresponding virus-sensitive substrates. Because cells in master seed banks and in working cell banks are of the highest importance for the cell culture production process, samples from these are the first to be considered for the rather expensive and time-consuming testing exercise.

Another and complementary approach to virus safety is the design of virus kill and removal steps of the recovery process for the protein of interest. These include physical and chemical principles separating (theoretical) viral contaminants from the product or inactivating them. Again, appropriate testing procedures and the demonstration of inactivation and removal of model viruses, as discussed in the important paper by Wiebe et al. (11) was considered a major provision for the achieved safety of recombinant products from hamster cells.

Product Consistency, Quality, and Purity: DNA, Endotoxin, and Protein Characterization

A very rich collection of methods has become available for the analysis of purified proteins within the last two decades. In addition, most of these methods have been recently optimized and fine-tuned to very high sensitivities and resolution. When employed as routine analytical procedures during the manufacturing process, they can assure high quality and consistency of protein products (12–17). Still, the manufacturer of a pharmaceutical protein has

one major concern: will a product that has demonstrated its efficacy in extensive and expensive clinical trials remain the same in its essential characteristics when employing a defined manufacturing process for its production over many years? It seems surprising, but due to the large size and complexity of proteins under study for clinical use today, their structural features and their functional principles within the human body may not be fully understood by the manufacturer when completing successful clinical trials. This is especially true for the newer generation of pharmaceutical proteins that are larger in size and often contain multiple polypeptides and/or multiple specialized domains exhibiting secondary modifications such as glycosylaton, phosphorylation, and sialylation. Subtle changes, sometimes difficult to detect due to inherent heterogeneities of structures of protein populations, may result in loss or modification of activity and could pose risks to the patient. In order to reduce this possibility, batteries of in-process controls and tests are an inherent part of the production of clinical proteins. In the following section, sensitive and robust techniques and methods are outlined and discussed. The objective is to *(1)* prevent the occurrence of even small changes in the procedures employed to produce the product of interest and *(2)* enable the detection and exclusion from the final vial protein variants differing in a major way from the one tested in clinical trials.

Adherence to GMP Protocol for Manufacturing. The first level of protection against inadvertent changes of the product rests in effective management of the production process over time. The challenge is the same as that of any mature industry: to produce large amounts of material at competitive cost while ensuring that product consistency and quality are maintained. Obviously, manufacturing teams and their supervisors undertake every effort to reproduce the manufacturing process to the utmost detail in every production run. This is achieved by defining and describing each of the various steps in the form of detailed protocols. Almost every aspect of the procedure is documented. These documents establish, in the form of cGMP (current good manufacturing practice) protocols, the basis for the overall procedure. A "sign off" procedure by supervisors represents an integral part, assuring that the operating personnel for the manufacturing process are in fact controlling, assessing, and executing it according to the established protocols. At critical points in the overall process, the signatures are prerequisites allowing the progression of the process to the consecutive steps.

The time period of cell line and process development leading up to the establishment of the manufacturing process is very important for the definition of critical "checkpoints." During this period, knowledge about parameters and steps is acquired that can result in product changes. Once certain limits of variations of process conditions have been identified (within which no change was observed) the cGMP protocol is drafted and finalized. Specific events, defined in precise terms as part of the manufacturing protocols, can trigger a more elaborate investigation. Supervisors and managers can even order an interruption of the manufacturing process. In extreme cases, crude product batches are withheld from further processing and are discarded.

Quality Control for the Detection of Variant Product Forms. Quality control (QC) is an integral part of the manufacturing process of recombinant products. A comprehensive approach, utilizing independently validated techniques, is applied to assess the quality and identity of the product from various angles (for review see also Ref. 18). It is the goal of QC efforts to assure that products made over years of manufacturing will meet the stated specifications in terms of identity, quantity, activity, and purity.

In principle, it is no longer very difficult to produce large quantities of highly purified recombinant proteins. However, methods to produce these proteins are still considered part of a young technology, because its basis is the manipulation of genetic material of various sources in the laboratory. Those manipulations involve the creation of plasmid vectors, based on fragments of genetic information that are linked together as a first step. In a second step, the transfer of such DNA for transgene-dependent protein synthesis to a complex organism such as *E. coli*, yeast, or, as discussed here, into the chromosomes of mammalian cells cultivated in vitro occurs. A major concern has been the fidelity (amino acid sequence identity) of the final product, particularly in view of the high degree of ignorance with respect to gene transfer mechanisms in higher eucaryotes and its consequences. Also, it appears that transfected DNA may have, possibly during the transfer to the host cell, a somewhat elevated propensity for mutation (19). A discussion of issues concerning structure and fate of DNA transferred to mammalian cells will be discussed later.

Nevertheless, based on a history of experience with this technology for more than 10 years, it can be stated that this young technology is in fact a very reliable one. It has been suggested that rigorous and extensive nucleic acid based tests, most notably a complete sequence assessment of the genomically integrated DNA (or transcribed RNA from recombinant cells) should be performed (20,21). It appears that such rather complex, expensive, and time-consuming tests are not necessary. An entire session was devoted to nucleic acid considerations during a conference organized by the International Association of Biological Standardization, held in Annecy, France in 1994. Berthold specifically addressed genetic stability in CHO cells (22). As already pointed out, following the rapid development of techniques allowing manipulation and transfer of DNA molecules to various hosts, powerful and reliable new methods have emerged in analytical protein chemistry. These have increased the capacity to characterize purified protein preparations to an extremely high degree of sensitivity and resolution. These methods represent efficient tools to assess identity (including the amino acid sequence), quantity, potency, and purity of the product at the level where it is most needed—just prior to administration (delivery) to the patient. The results from QC assessments from thousands of lots of approved products are, as far as this author knows very encouraging with respect to (amino acid sequence) identity: no lot has been called into question

based on a QC finding that has indicated a newly emerged mutant gene.

Still, this statement does not negate the possibility that mutants may occasionally emerge. They will, over and over again, within the large population of cells in bioreactors. A 1,000-L reactor usually contains more than 10^{12} cells. However, there is no reason to believe that mutations would occur at a frequency significantly different from the frequency of mutations occurring within the human body (1 bp change in 10^9 bp for each generation) (23). Of course, such mutations will be "scattered," and as individual entities are extremely rare. Therefore, individual and specific variants are not likely to emerge in the population of protein molecules as a sizeable subclass of identical molecules.

A telling example of the power of in-process controls and the associated biochemical assays to identify variants in a population of molecules is given in an article by Sliwkowski and coauthors (24). It describes the detection of an amino acid exchange mutant of a recombinant monoclonal antibody (anti-Her2). This variant was detected early during process-development efforts when using a methotrexate-amplified clone of CHO cells. The variant represented about 10% of the total population of antibody molecules. The origin of this mutant remains mysterious, but it seems it was the result of an early event during the development of the cell line that was used in the process.

MOLECULAR BIOLOGY OF VECTOR DESIGN, TRANSFECTION, AND IDENTIFICATION OF HIGHLY PRODUCTIVE CELL POPULATIONS

DNA Transfer to CHO Cells Using Calcium Phosphate

Graham and van der Eb (25) showed more than 20 years ago that the transfer of DNA into mammalian cells in culture can be achieved by exposing cells to microprecipitates of DNA and calcium phosphate. This simple method (termed calcium phosphate transfection) has become very popular. Typically, a glycerol shock is part of the transfection procedure. The shock is thought to promote the transport of individual DNA–calcium phosphate complexes across the cell membrane, perhaps by inducing endocytosis or phagocytosis. When optimized transfection conditions are used, a high percentage of cells take up DNA. However, in addition to the movement across the outer membrane of the cells, the transport of DNA from the cytoplasm to the nucleus seems to be a significant barrier. In 1982 Loyter et al. reported that only 1–5% of transfected cells have detectable calcium phosphate complexes within the nucleus (26,27). The stable integration of transfected DNA into chromosomes of mammalian cells was first shown by Axel and collaborators (28,29). The number of emerging colonies upon transfection by the calcium phosphate technique has been reported to be low: usually, only between 0.01% and 0.5% of the cells gave rise to recombinant colonies (30). More recent investigations and appropriate improvements of crucial physicochemical parameters of the calcium phosphate transfection methodology in several laboratories have increased transfection efficiencies for stable recombinant cells to about 5–10% (and those of transient transfections to levels of 50% or higher) of the population of cells exposed to exogenous DNA (31–33). Calcium phosphate transfection still seems to be the most frequently used method for the generation of recombinant CHO cells.

Yet, it appears that the transfer of a sufficient quantity of DNA into the nucleus and the integration of at least a few functional plasmid molecules into transcriptionally active regions of the receiving genome is by no means an efficient process. Half a million or more plasmid molecules are usually provided per cell in a typical calcium phosphate transfection, the result of which is stable gene transfer in a small percentage of the exposed cells. It has been shown that the association of the DNA with calcium phosphate protects against nuclease attack during the transport to the nucleus (26). Eventually these complexes become dissociated and nuclear endo- and exonucleases will have access to the "naked" DNA. Moreover, it has been shown that supercoiled plasmid DNA molecules will be converted into relaxed (single-stand cut), open circular, and finally linear molecules (double-strand cut) within the nucleus after 1–2 h (34). This is, of course, an essential step for integration into the linear DNA backbone of a chromosome.

The mechanism by which the DNA is transported across the nuclear membrane (nuclear pores are not large enough for diffusive transport), and finally to the site of chromosomal integration, are not yet known. On the receiving side, at least one strand of the chromosomal DNA must be opened while a linear plasmid molecule is in sufficiently close proximity. Nuclear ligases could then mediate covalent linkages. Very little is known about the mechanism and about the factors affecting site specificity for integration.

Site specificity is a very relevant factor in this context. It is believed that it has influence on the transcription rate of the integrated DNA, even if the DNA carries a powerful (viral) promotor, such as the SV40 early promoter or a CMV promoter. The current assumption is that exogenous DNA will integrate at "random" within the genome, even if the transfection cocktails contain DNA with homology to sequence segments of the receiving genome. This is an observation from numerous gene-targeting experiments in mammalian cells. It was found that only 1 in 1,000 to 1 in 10,000 events would result in targeted integration (35,36). With respect to the actual sites of integration in CHO cells, there is only limited and sporadic data available at present. Only a small proportion of recombinant cell lines should contain the transferred DNA within transcriptionally active regions of chromosomes if integration would truly occur at random subsequent to transfection. In reality, we cannot know whether DNA will end up more frequently than not in transcriptionally inactive regions of the genome. We have no opportunity to analyze cells that did not overcome the selection step. The selection step is, as we know from the low transfection efficiencies in stable transfections, a very significant barrier.

Among the emerging cell lines there is usually extensive heterogeneity in productivity for recombinant proteins. The expression levels from primary CHO clones for proteins like rtPA, a secreted version of the CD4 receptor (rCD4-truncated), and a chimeric TNFR-FC fusion protein had a range of more than two orders of magnitude (37,38), observations which this author believes can be generalized

for all transfections in mammalian cells. As a consequence, the identification of high-producer cell lines is a tedious and labor-intensive exercise and requires the screening of hundreds of individual cell lines. In the author's experience, it is usually necessary to invest 2–4 months of laboratory work into this task. Only then has one enough confidence in having appropriately assessed the upper range of sustainable expression for a novel protein in CHO cells.

Nuclear enzymes such as endo- and exonucleases, but also ligases (39) and possibly recombinogenic enzymes (40,41), acting on the population of plasmid molecules within the nucleus are responsible for the modification of transfected DNA that becomes eventually integrated into the genome. This modification may not only degrade in part the plasmid DNA but is also responsible for creation of larger, that is, more extended DNA complexes. These DNA molecules seem to be created before the integration into the genome. They provide the physical basis for linking the genetic selection marker and the gene(s) of interest in those transfections that utilize separate vectors. It should be noted here that DNA transfer methods, other than the use of calcium phosphate, might deliver different quantities of DNA to the nucleus. This may have profound effects on the structure and copy number of integrated DNA molecules, as will be discussed later.

In discussing the integration loci for transfected DNA in the mammalian genome, it is interesting to note that integration sites for amplified DNA sequences, detectable by fluorescence in situ hybridization (FISH) after selection in methotrexate, did not show a preference for a particular chromosome or chromosomal region. However, in light of recent data on gene amplification mechanisms (42), it is not possible to draw conclusions from these findings toward the identification of the initial (primary) integration site(s).

Additional DNA Transfer Methods

In addition to calcium phosphate transfection (25,31), more recently electroporation (43,44), cationic lipids, and liposome mediated (45–47) transfection techniques have been developed for transfer of DNA into mammalian cells. Most of these techniques have been reported to mediate higher transfection efficiencies as compared with calcium phosphate–mediated DNA transfer. Such claims must be regarded with some caution because all DNA transfer techniques established so far suffer from high variability. Some of the variabilities observed are probably due to technical difficulties in producing reliable and consistently highly efficient carrier systems. The other factors that cause major variations in transfection efficiency are the type of cells used and the condition of the cells prior to transfection.

Because carrier principles for the transfer of DNA to cells can vary widely, one suspects that very different consequences may result within the cell, events over which the person conducting the experiment has no or little control. Just one of the possible consequences should be mentioned here: The technique applied for DNA transfer may affect the copy number average of individual clones, depending on the amount of DNA transferred to the nucleus. It has been proposed that electroporation may result in a smaller copy number of chromosomally integrated DNA than, for example, calcium phosphate–based transfection. In view of the applied selection procedures for the generation of recombinant cell "clones," very different chromosomal structures of integrated plasmid sequences may result.

Another factor to be taken into account for the genomic structure of integrated plasmid sequences is the physical form of the DNA for transfection. Because linearization of plasmid DNA is a prerequisite for integration into any chromosomal site, the structure of the plasmid prior to transfection will affect the final structure when integrated. When circular plasmids are used, as predominant in the past, linearization is dependent on cellular enzyme activity. This will happen most likely at random, and the integrated DNA structure cannot be predicted. Depending on the overall size of the plasmid molecule and the size of the DHFR gene cassette relative to the vector, a significant number of molecules would be linearized within the DHFR sequence, which jeopardizes the functionality of this plasmid. Obviously, molecules whose DHFR expression cassette are interrupted or deleted by the linearization process would be unable to transfer a functional marker gene to the host cells, and such cells would be eliminated through the selection procedure. Therefore, opening the circular plasmid molecule by restriction enzyme digestion with appropriate enzymes prior to transfection has been recommended for improved transfection efficiencies in protocols that aim at integration of plasmid DNA into chromosomal DNA.

We have seen in our laboratory that linearization is, in fact, advantageous in terms of transfection efficiency with stable transfections (32). Linear plasmid molecules are also used routinely in gene transfer experiments that attempt to transfer the desired DNA into specific (homologous) DNA sequences of the receiving genome (gene targeting). In these experiments, linearization is executed in a way that results in cutting the homologous DNA fragment approximately in half. The molecules for DNA transfer then carry DNA sequences that are homologous at their ends (48–50). We have used such an approach in our laboratory by employing (defective) retroviral DNA sequences derived from the CHO genome in expression vector cocktails. They have quite significantly improved the frequency of high-producer cell lines upon transfection (37,38).

CHO CELL ENGINEERING

Use of serum and other more or less undefined and complex media additives for growth of mammalian cells may result in problems of reproducibility in the process, and thus additional costs. To avoid a negative impact on cell growth, testing procedures have to be incorporated prior to use of these additives to identify possible variations from batch to batch. Therefore, the desire to replace some or all of these variable components in cell culture media with well-defined components having high purity and little or no variation in quality has been a priority in the industry for quite some time. Also, use of serum and animal-tissue-derived growth factors (insulin, transferrin) bears

the considerable risk of transmittal of undesired adventitious agents to the production process because these additives frequently cannot be heat sterilized.

Although serum-free and growth-factor-free media may support the growth of mammalian cells, the robustness of the cell culture process, in general, and maximally achievable cell densities in particular may be reduced when using media free of serum- and/or protein growth factors. Therefore, efforts have been undertaken to modify production host cell lines prior to the transfer of expression vector DNA. The hope was to generate hosts for production that would achieve superior growth rates, and so would be more useful for scale-up and higher in productivity. Various oncogenes, cell cycle genes, and/or hormone genes (51–53) have been transferred into the genome of cells, resulting in novel and possibly superior production hosts. In view of actual discussions regarding BSE and other, so far unknown, adventitious agents, which pose the risk of entering production through process additives, the interest in such novel hosts will be very high. It needs to be seen how soon and how frequent these "super CHO" cell lines will be used in the manufacturing of novel products.

METHOTREXATE-MEDIATED VECTOR AMPLIFICATION IN DNFR-TRANSFECTED RECOMBINANT CHO CELLS

Treatment of DHFR-positive CHO cells with incremental concentrations of methotrexate has been used to select for populations and clones of populations of CHO cells with an increased (i.e., amplified) copy number of the integrated DNA within the host genome. The phenomenon of methotrexate-mediated gene amplification had been observed before the use of recombinant DNA technology, most notably in cancer patients who had been treated for extended periods of time with methotrexate (54).

With CHO cells, cell lines containing several hundreds, and sometimes a few thousand, copies of transfected DNA fragments integrated within chromosomes have been established (55,56). Most amplified cells produce more product than unamplified host cells. However, the improvement of specific productivity is highly variable when studying individual clones (57) and also varies from product to product (Wurm, unpublished observation, 1994).

Cytogenetics of Amplified CHO Populations: Long-Term Genetic Stability and Maintenance of Productivity

Subpopulations and clones of methotrexate-treated CHO cells may contain the transfected DNA sequences at very high copy numbers (55,56,58). Studies concerning genetic features of these amplified DNA sequences within the CHO genome have been ambiguous. Whether continued presence of a selective agent (i.e., methotrexate) in long-term cultures of recombinant CHO cells is required for production stability is still controversial. Whereas some studies suggest MTX is required for stable production of recombinant proteins (59), others indicate that continued culture of clonal cell lines in MTX may not be necessary (60). Cytogenetic studies using fluorescence in situ hybridization (FISH) (61) were done in the author's laboratory with clonal and nonclonal recombinant CHO cell lines. They showed that continuous exposure to MTX at the same concentration under which the cell populations were initially established promotes genetic instability at the chromosomal level (62). Cell lines selected at micromolar concentrations frequently show elongated chromosomal structures that hybridize to probes representing the transfected DNA. Some of the chromosomes, containing a large number of more or less tightly arranged bands or regions, differ dramatically from normal CHO chromosomes. This raised questions about cytogenetic stability of such aberrant chromosomes during continuous subcultivation and, in consequence, the production of recombinant proteins based on genes residing within these structures.

The first of our FISH analysis studies was done with cell lines expressing the mouse c-*myc* gene in an inducible manner. The advantage of this system is that the phenomena observed in the genetic context can be attributed solely to the status (expression level, chromosomal location, degree of amplification) of the DHFR selector gene and the inhibitor MTX. The c-*myc* gene is truly silent during regular growth conditions and could not act as a metabolic burden nor result in any selective forces. When these cells are cultivated at 37 °C, the inducible *Drosophila melanogaster* heat shock protein 70 promoter is in "off" status. Even in a cell line containing 1,000–2,000 copies (selected at 320 μM methotrexate) of the transfected DNA, no trace of transcribed c-*myc* RNA and translated c-*myc* protein was observed at 37 °C. However, high-level expression of c-*myc* RNA was induced upon shifting cells to a culture temperature of 43 °C for 1 h.

The continuous exposure of the 320-μM-MTX-selected cell line (c-*myc*) to the drug resulted in about a twofold higher copy number of plasmid DNA as compared with a population of cells that had been cultivated in the absence of any selective pressure for at least 4 weeks. A similar observation, based on a cell line with a constitutive promoter driving tPA expression, was reported by Weidle et al. (59). These authors reported a loss of about 50% in copy number when cells were grown in the absence of methotrexate for about 50 days. This decline in gene number also resulted in a similar drop in productivity. No further copy number and productivity loss was found during the subsequent observation period of up to 100 days.

Following our initial observation of a drop in copy number in cells containing the mouse c-*myc* gene, we extended our studies with this line and also analyzed the chromosomal integration patterns of transfected sequences in two additional CHO cell lines. The three cell lines studied under this program expressed one of the three following proteins: *(1)* human rtPA, *(2)* a truncated version of the human CD4 receptor, and *(3)* the murine c-*myc* protein. The three cell lines had been established by transfecting CHO-DUKX cells using standard calcium phosphate precipitation techniques. Either individual clones or pools of clones were cultivated for extended periods in media containing stepwise increasing concentrations of MTX. In each case, a single individual clone expressing the desired level of recombinant protein was expanded for long-term growth and production. FISH was performed using biotinylated probes complementary to the recombinant sequences. Three dif-

ferent sets of typical integration patterns (site, degree) of amplified sequences were found for lines under study.

FISH signals in the rtPA-expressing cell lines were localized, in the majority of cells, in one of the larger submetacentric chromosomes (probably either chromosome 3 or 4, or the X chromosome). The signals (a bright band of fluorescence) covered most of the length of the short arm and showed weaker fluorescing spots over the center of the chromatids in the long arm. A number of other types of integration patterns (including an elevated percentage of quasi-tetraploid cells) were found in these cells (about 25%) when they were grown in the presence of methotrexate. When methotrexate was omitted, more than 95% of metaphases had the integration described.

In cells expressing c-*myc*, continuous cultivation in media containing high concentrations of MTX was found to be associated with rearrangement and variable amplification of transfected sequences within almost half of the cells (62). Several types of integration patterns were found in these cells: *(1)* chromosomes with highly extended regions of hybridizing sequences, *(2)* cells containing transfected DNA integrated into multiple sites on an individual chromosome, *(3)* cells containing DNA integrated into multiple chromosomes, *(4)* chromosomes joined at amplified regions, *(5)* circular chromosomes, *(6)* cells with small derivative chromosomes or fragments, and *(7)* cells containing a single "normal" integration, exhibiting a bright band of hybridization in the center of the long arm of what we assume is chromosome 2. The last type *(7)* was found in about 40% of cells when grown in the presence of methotrexate. This integration type increased in frequency to 95% when cells were grown in the absence of methotrexate for 3 weeks or more.

Similar results showing quite extensive heterogeneity were observed with yet another recombinant CHO cell line (the fourth studied), producing a chimeric CD4-immunoglobulin protein. When grown in the presence of MTX, CD4-IgG-producing cells showed heterogeneity in both chromosomal location and copy number of amplified sequences coding for CD4-immunoglobulin (CD4IgG) (63).

For the cell line expressing the truncated CD4 molecule, a strong signal was found over the telomeric end of a small acrocentric chromosome (possibly a derivative of chromosome 5, 6, or 7), which covered about one-third of the total length.

In summary, when cultivating these cell lines in the absence of methotrexate, unique and characteristic integrations were found in 95–99% of metaphases. Subsequent to the first analysis, we continued to cultivate these cell lines for extended periods in the absence of methotrexate, and we occasionally performed additional FISH analysis. We found the chromosomal structures described here to be stable within the observation period (a minimum of 60 days and in one case of 160 days). Cytogenetically, these observations indicate a high degree of genetic stability of chromosomally amplified sequences in the absence of methotrexate.

Assessment by Southern Hybridization and FISH of the Chromosomal Status of Amplified DNA Sequences.

Southern hybridization (64) of genomic DNA is a useful tool to acquire insights into the molecular structure of integrated plasmid sequences. Using suitable restriction enzymes it will provide information on the integrity of the expression fragments of interest. Estimates of copy numbers of the chromosomally integrated DNA can be established when, in the same experiment, an appropriate range of quantities of plasmid DNA, restricted with the same enzyme as the one used for restriction of genomic DNA, is used. Also, Southern hybridization may indicate whether one integration locus or more than one exists for the plasmid sequences. However, conclusions on multiplicity of integration must be taken with a degree of caution. Aberrations from the expected signals can be due to postintegration rearrangement in a fraction of the cell population or in the initial cointegration of a few copies of the plasmid sequences that had been subjected to nuclease attack, resulting in the deletion of the restriction enzyme site used for the analysis.

All Southern hybridizations are based on DNA extracted from thousands of individual cells. Even if the cell lines are based on a "cloning" step, one must be aware that none of the so-called clonal cell lines are clonal in the true sense. Mammalian cells, in particular immortalized cells, have an intrinsic genomic fluidity that is most evident when studying chromosome numbers in metaphase spreads (for a review see Ref. 65).

FISH can be used to obtain insights into the extent of chromosomal amplification (not to be confused with copy number estimates) and the location of recombinant DNA within the chromosomal set of cells. In order to provide some useful information, FISH studies must be supplemented by a statistical analysis of identified integration sites (and structures observed). The time point of analysis, with respect to the total period of time (since thaw) for use of these cells, should also be taken into consideration. A reasonable value (from the invested work) may be gained from a FISH study performed shortly after cells have been thawed from a bank of cells stored in liquid nitrogen. Depending on the culture conditions (with or without methotrexate, with or without serum, etc.) and the length of time during which cells have been exposed to in vitro culture, results of FISH analysis may vary considerably. In spite of variability, in the studies discussed we were able to separate without any doubt or difficulty the identity of one recombinant cell line from a second recombinant cell line by using FISH. This was possible because we were able to identify a chromosomal marker containing hybridizing DNA that was present in a large fraction of the individual cells of the populations studied. The identifying chromosomal markers containing recombinant sequences were termed *master integrations*, which we found to be the genetically stable entities in the system.

CHOOSING THE PRODUCTION SYSTEM: ADHERENT CELLS, MICROCARRIERS, AND FREE SUSPENSION ISSUES OF ECONOMY, EXPERTISE, AND TIME FRAME TO CLINIC AND MARKET

Process scientists and their managers must eventually decide which type of production system to choose for the product in question. The choice of a production system is driven

by the anticipated scale of operation for the manufacturing process, as well as by other issues. A number of very successful recombinant products from mammalian cells, such as EPO and GM-CSF, have been administered worldwide to a large number of patients. For these peptide hormones, the applied doses for treatment of patients are relatively small (about 100 μg/patient). Another successful protein therapeutic, recombinant tissue plasminogen activator (hu-rtPA, Activase®) requires a dose of 100 mg/patient, a 1,000-fold difference in dose as compared with erythropoeitin. The number of patients treated per annum with the two categories of products is about the same. It is obvious that such a difference in the annual product requirement would drive very different manufacturing decisions.

In the case of EPO and GM-CSF, a process based on roller bottles appears to be sufficient to supply the market with product. Using roller bottles, a "simple" and robust process can be established with minimal initial investment in hardware.

The Roller Bottle to Roller Bottle Process

Cells thawed from a cell bank can be expanded by subcultivation into a fixed number of roller bottles. This approach is based on cells growing attached to the inner surface of the roller bottle. Standard 2-L roller bottles are usually filled with 300–500 ml of medium. The remaining volume (air) provides the necessary oxygen while the bottles are slowly rolled (at about 1 rpm) in an incubator at 37 °C. The caps on the bottles are closed during the period of incubation. A sufficiently large number of roller bottles containing a healthy cell population represents the starting point for several production cycles. Part of the cell mass in the roller bottles can be immediately used for reproduction of cells in order to initiate the production cycle following the one already initiated. In other words, the bulk of the cell mass will serve as a seed culture to initiate the first production cycle. The remainder will provide, within one or two subcultivations, the basis for the subsequent production cycle.

Media for production are usually rich in nutrient content in order to maintain viability and productivity of the cells for a minimum of 5–7 days. Subsequently, the medium with the secreted product will be harvested, leaving the adherent cells inside the bottle. Sometimes, a refeeding with fresh medium for a second production cycle is possible, on the condition that the product of the first and the second harvest will be similar in composition and quality. Since a standard 2-L roller bottle will have about 300 ml of medium for harvest, 1,000 rollerbottles can provide product from 300 L or 600 L of supernatant, from one or two harvests, respectively. If such a campaign lasts all year long, a manufacturing process based on this principle will deliver 15,000 L (30,000 L) of cell-free culture medium containing the product of interest. Product concentrations in the 10- to 100-mg/L range are possible, resulting in 150 to 1,500 of unpurified product. Of course it must be realized that such a process is very labor intensive, requires the repeated use of trypsin for detachment of cells, and carries with it a considerable risk of contamination with adventitious agents through the handling personnel. However, because subcultivation in plastic cell-culture vessels represents a standard procedure in all laboratories, the scale-up from the lab-scale procedure to the pilot or manufacturing scale is a convenient approach for obtaining considerable (but not large) quantities of recombinant proteins.

The Bioreactor to Roller Bottle Process

The ability to cultivate CHO cells both in suspension and, subsequently, in an attached form allows the roller bottle production process to be streamlined. The seed culture for the roller bottle production phase can be generated in spinners or in bioreactors. The advantage of such a process would be that fewer interactions in terms of subcultivation are needed to generate sufficient cell mass. Being less labor intensive, it also reduces the risk of contamination by adventitious agents of the feed culture.

The Stirred Tank Bioreactor Process

Kilogram-scale production of recombinant proteins from CHO cells was first achieved in stirred bioreactors at very large scale (1,000 L and 10,000 L) with cells synthesizing while floating in free suspension (66). While such adaptation to suspension growth had been, in the early years of CHO technology, a somewhat tedious process, today multiple factors seem to have made this easier. For one, cell culture media have been developed that support the growth of cells in suspension better than those based on formulations such as DMEM and Ham's F12. Also, a number of purposeful genetic modifications (transfection of hormone genes or cell cycle genes, selection of CHO populations, or extended growth of CHO cells in media with little or no fetal calf serum and other modifications) appear to have generated CHO phenotypes that more readily support the transition from adherent growth to suspension growth.

In a bioreactor-based process, the scale-up to very large volumes can occur rather fast. It is usually executed by subcultivations, that is, by diluting the entire volume of one bioreactor into 5–20 volumes of fresh medium, maintained prewarmed in the next larger reactor. Within 10–15 days, therefore, a suspension culture at the 50-L scale could inoculate a 10,000-L reactor. It is major goal of process development work to carefully optimize media for the very last step—the production phase. Such a medium must support good growth in order to achieve the highest cell density possible, and it also must provide the nutritional basis and physiological balance needed to maintain viability and productivity for extended periods. While the termination (i.e., harvest) of such a culture is driven mainly by issues of plant capacity and volumetric productivity, the other important issue to consider is the quality of the derived product. The continuously changing composition of the culture medium during the production phase can affect the quality of earlier synthesized product through degradative activities, such as sialidase activity and proteolytic activity.

BATCH, EXTENDED-BATCH, AND PERFUSION PROCESSES

The batch process is considered a very simple, robust production process for stirred bioreactors. In the step preced-

ing production, cells in a (somewhat smaller) bioreactor are cultivated to maximal cell density and are then transferred, still suspended in the exhausted growth medium, to the production reactor. After addition of fresh medium and eventual adjustments in control of certain process parameters (O_2 supply, pH, temperature, etc.) the production process begins. It ends at a predetermined point in time when synthesis of recombinant protein seizes or slows down due to exhaustion of nutritional components in the medium and/or accumulation of toxic end products of cellular metabolism. For CHO cells the production phase will usually last, depending on the proteins' susceptibility to degradative enzymes and a number of other process-related factors, between 5 and 14 days. The advantage of such a process is obvious. Providing that the inoculating cell mass can be made reproducible and with little variation in biological quality, the resulting production process will show a high degree of reproducibility with respect to cell growth, viability, quantity, and quality of the product being synthesized.

The issue of reproducibility of process parameters and of achievable product quantity and quality is of highest significance, of course. It provides confidence-generating data for the long-term future. Another important point is the obligation to present such data to regulatory agencies that will ultimately give permission to use the process for the generation of the product. For investigational new drug applications (IND) and process licence applications (PLA), rather specific requirements must be met with respect to the minimal number of product batches analyzed. For INDs no less than three product runs are recommended to be presented. Shorter batch processes (5–7 days) have the advantage of generating more data within a given time frame. This can be a very important cost factor because the time period necessary to acquire and evaluate necessary data from the new process will eventually affect the overall time necessary for entering the market.

On the other hand, there are viable arguments that could sway process-development decisions in another direction. The option to prolong the synthetic activities of cells in the production vessel would capitalize on the process investment that allows generation of the necessary cell mass for the production vessel. Therefore, more and more extended-batch or perfused-batch cultures are used. With a longer-lasting production phase in cell culture, feeding additional medium components becomes necessary. Obviously, extending the process for a considerable period of time, (for example, from 7 days to 13 days) makes sense only when the return for this investment in labor and in occupation of the production facility results in a sufficiently high increase in product concentration within the vessel. There are various ways in which medium and medium components can be added to a culture that was initiated a few days earlier in the same tank. This could be done by feeding (batchwise) highly concentrated mixtures of essential amino acids and other medium components, thus not significantly affecting the volume in the tank. Alternatively, a culture may be started in the production vessel at half or so of the working volume, then standard concentration medium could be pumped slowly and continuously, or batchwise, into the tank until the final working volume has been reached. The choice between the modes, or combinations thereof, is in the hands of the process development scientist, who has to carefully evaluate advantages and disadvantages. No matter what principle will be used for extending the production phase, the overall result will always be a tank that contains, in one batch, the entire protein population for subsequent recovery and purification.

A very different philosophy for manufacturing is represented by continuously perfused production processes. Here, the goal is to achieve the highest cell concentrations possible within small tanks. In this process up-front investment for manufacturing equipment is reduced, and product quantities can be quite high. The much improved knowledge base in technology, and in physiology of CHO cells in culture, has made this more-complicated approach to manufacturing attractive. Perfused cultures can be maintained for many weeks and months, with product harvests occurring repeatedly throughout that period. A protein of high interest to the pharmaceutical industry for decades, Factor VIII, is reliably being manufactured using perfusion technology with baby hamster kidney cells, a "cousin" of Chinese hamster ovary cells. While the production of Cognate® Factor VIII (67) has pioneered the use of perfusion technology for recombinant proteins with mammalian cells in a non-CHO cells, other products from CHO cells have recently been approved as well using this technology (68).

Finally, due to the fact that many monoclonal antibodies have been produced at the laboratory scale in perfusion systems for many years by using hollow-fiber technology, the transfer of this technology to the production scale for pharmaceutical manufacturing has been an obvious conclusion for quite a while (69,70). Recently, a diagnostic monoclonal antibody for imaging in patients with prostate cancer (ProstaScint™) was the first FDA-approved product for injection into humans derived from the hollow-fiber perfusion process (71). The application of this and other perfusion technologies, using immobilization of CHO cells on macroporous carriers to produce recombinant antibodies or other proteins, is likely to be yet another option for future manufacturing processes (72).

CONCLUSIONS

The technology of using mammalian cells, and in particular CHO cells, for recombinant protein production is, in spite of significant success stories, still in its infancy. Much has to be done to establish production processes in a more straitforward way and to make them more productive. A data explosion is occurring with DNA sequences from microorganisms, plants, and animals, including the human genome. Subsequently, a tremendous need is arising for more detailed studies and, frequently, the urge to use the translation products of these sequences in many applications. Also, the possibility of designing new proteins, based on concepts provided by nature, will even increase the need for flexible and robust synthetic machinery. Mammalian cells have been used as production hosts in some of the first pioneering experiments for the design of novel pro-

teins (73,74). However, to deliver significant quantities of the desired proteins to analytical chemists, cell biologists, clinical researchers, and so on, takes far too long and is still a very demanding task. Some of the bottlenecks in the knowledge base have been addressed in this article. It is hoped that research efforts will solve some of them in the near future, resulting in more efficient, faster, and less expensive synthesis of recombinant proteins by mammalian cells. It appears that mammalian cells inherently have, due to their large size, the capacity to take up large quantities of foreign DNA without any negative effects, the responsiveness to perform a large number of synthetic tasks, and the capability to meet the challenges of the future.

BIBLIOGRAPHY

1. D. Pennica, W.E. Holmes, W.J. Kohr, R.N. Harkins, G.A. Vehar, C.A. Ward, W.F. Bennett, E. Yelverton, P.H. Seeburg, H.L. Heyneker, D.V. Goeddel, and D. Collen, *Nature* **301**, 214–221 (1983).
2. G. Urlaub and L.A. Chasin, *Proc. Natl. Acad. Sci. USA* **77**, 4216–4220 (1980).
3. M.B. Hendricks, C.A. Luchette, M.J. Banker, *Biotechnology* **7**, 1271–1274 (1989).
4. F.-T. Kao, and T.T. Puck, *Proc. Natl. Acad. Sci. USA* **60**, 1275–1281 (1968).
5. T.T. Puck, in M.M. Gottesman ed., *Molecular Cell Genetics*, Wiley, New York, 1985, pp. 37–64.
6. G. Ringold, B. Dieckmann, and F. Lee, *J. Mol. Appl. Genetics* **1**, 165–175 (1981).
7. R.J. Kaufman and P. Sharp, *J. Mol. Biol.* **159**, 601–621 (1982).
8. R.J. Kaufman, L.C. Wasley, A.J. Spiliotes, S.D. Gossels, S.A. Latt, G.R. Larsen, and R.M. Kay, *Mol. Cell. Biol.* **5**, 1750–1759 (1985).
9. L. Hayflick, *Science* **276**, 336–337 (1997).
10. M.E. Wiebe, and L.H. May, in A.S. Lubiniecki ed., *Large Scale Mammalian Cell Culture Technology*, Marcel Dekker, New York, 1990, pp. 147–160.
11. M.E. Wiebe, F. Becker, R. Lazar, L. May, B. Casto, M. Semense, C. Fautz, R. Garnick, C. Miller, G. Masover, D. Bergman, and A.S. Lubiniecki, in R.E. Spier, J.B. Griffiths, J. Stephenne, and P.J. Crooy eds., *Advances in Animal Cell Biology and Technology for Bioprocesses*. 1987, pp. 68–71.
12. J.V. O'Connor, R.G. Keck, R.J. Harris, and M.J. Field, in: F. Brown, and A.S. Lubiniecki eds., *Genetic Stability and Recombinant Product Consistency*, Karger, Basch, 1994, pp. 165–173.
13. R.L. Garnick, *Develop. Biol. Standard*, **76**, 117–130 (1992).
14. H.S. Lu, L.B. Tsai, W.C. Kenney, and P-H. Lai, *Biochem. Biophys. Res. Comm.* **156**, 807–813 (1988).
15. R.C. Chloupek, R.J. Harris, C.K. Leonard, R.G. Keck, B.A. Keyt, M.W. Spellman, A.J.S. Jones, and W.S. Hancock, *J. Chromatog.* **463**: 375–396 (1989).
16. J.V. O'Connor, *Biologicals* **21**, 111–117 (1993).
17. R.C. Stephenson, and S. Clarke, *J. Biol. Chem.* **264**, 6164–6170 (1989).
18. V. Anicetti, and W.S. Hancock, in R. Harrison ed., *Protein Purification Process Engineering*, Marcel Dekker, Inc. New York, 1994, pp. 25–56.
19. M. Calos, J.S. Lebkowski, and M.R. Botchan, *Proc. Natl. Acad. Sci. USA*, **80**, 3015–3019 (1983).
20. F. Galibert, *Biologicals* **18**, 221–224 (1990).
21. F. Galibert, in F. Brown and A.S. Lubiniecki eds., *Genetic Stability and Recombinant Product Consistency*, Karger, Basel, 1994, pp. 27–30.
22. W. Berthold, in F. Brown and A.S. Lubiniecki eds., *Genetic Stability and Recombinant Product Consistency*. Karger, Basel, 1994, pp 67–79.
23. J.W. Drake, *Annu. Rev. Genet.* **25**, 125–146 (1991).
24. R.J. Harris, A.A. Murnane, S.L. Utter, K.L. Wagner, E.T. Cox, G.D. Polastri, J.C. Helder, and M.B. Sliwskowski, *Bio/Technology* **11**, 1293–1297 (1993).
25. F.L. Graham, and A.J. van der Eb, *Virology* **52**, 456–467 (1973).
26. A. Loyter, G.A. Scangos, and F.H. Ruddle, *Proc. Natl. Acad. Sci. USA* **79**, 422–426 (1982).
27. A. Loyter, G. Scangos, D. Juricek, D. Keene, and F. Ruddle, *Exp. Cell Res.* **139**, 223–234 (1982).
28. M. Wigler, S. Silverstein, L.-S. Lee, A. Pellicer, Y.-C. Cheng, and R. Axel, *Cell* **11**, 223–232 (1977).
29. A. Pellicer, M. Wigler, and R. Axel, *Cell* **14**, 133–141 (1978).
30. D.M. Robins, S. Ripley, A. Henderson, and R. Axel, *Cell* **23**, 29–39 (1981).
31. C. Chen, and H. Okayama, *Mol. Cell. Biol.* **7** 2745–2752 (1987).
32. M. Jordan, A. Schallhorn, and F.M. Wurm, *Nucl. Acid. Res.* **24**, 596–601 (1996).
33. M. Jordan, C. Köhne, and F.M. Wurm in M.J. Carrondo, B. Griffiths and J.P. Moreira eds., *Animal Cell Technology—from Vaccines to Genetic Medicine*, Kluwer Academic, Dordrecht, 1997, pp. 47–50.
34. G.K. Finn, B.W. Kurz, R.Z. Cheng, and R.J. Shmookler *Mol. Cell. Biol.* **9**, 4009–4017 (1989).
35. S.L. Mansour, K.R. Thomas, and M.R. Capecchi, *Nature* **336**, 348–352 (1988).
36. H. Zheng, and J.H. Wilson, *Nature* **344**, 170–173 (1990).
37. F.M. Wurm, A. Johnson, T. Ryll, C. Koehne, H. Scherthan, F. Glaab, Y.S. Lie, C.J. Petropoulos, and W.R. Arathoon, *Ann. of the Acad. Sci.* **782**, 70–78 (1996).
38. F.M. Wurm, A. Johnson, Y.S. Lie, M.T. Etcheverry and K.P. Anderson, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, Butterworth-Heinemann, Oxford, 1992, pp. 35–41.
39. M. Perucho, D. Hanahan, and M. Wigler, *Cell*, **22**, 309–317 (1980).
40. G.K. Finn, B.W. Kurz, R.Z. Cheng, and R.J. Shmookler Reis, *Mol. Cell. Biol.* **9**, 4009–4017 (1989).
41. K.R. Folger, E.A. Wong, G. Wahl, M.R. Capecchi, *Mol. Cell. Biol.* **2**, 1372–1387 (1982).
42. J. Ruiz, and G.M. Wahl, *Mol. Cell. Biol.* **10**, 3056–3066 (1990).
43. G. Chu, H. Hayakawa, and P. Berg, *Nucl. Acids Res.* **15**, 1311–1326 (1987).
44. J. Barsoum, *DNA Cell Biol.* **9**, 293–300 (1990).
45. P.L. Felgner, and G.M. Ringold, *Nature* **337**, 387–388 (1989).
46. J.-P. Behr, B. Demeneix, J.-P. Loeffler, and J. Perez-Mutul, *Proc. Natl. Acad. Sci. USA* **86**, 6983–6986 (1989).
47. S.R. Muller, P.D. Sullivan, D.O. Clegg, and S.C. Feinstein, *DNA Cell Biol.* **9**, 221–229 (1990).
48. M. Jasin, and P. Berg, *Genes Develop.* **2**, 1353–1363 (1988).
49. O. Smithies, R.G. Gregg, S.S. Boggs, M.A. Koralewski, and R.S. Kucherlapati, *Nature* **317**, 230–234 (1985).
50. H. Zheng, and J.H. Wilson, *Nature* **344**, 170–173 (1990).

51. S. Shirahata, K. Teruya, K. Seki, T. Mori, H. Ohashi, H. Tachibana, and H. Murakami, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Technology: Developments, Processes and Products*, Butterworth-Heinemann, Oxford, 1992, pp. 54–59.
52. K.H. Lee, A. Guerini-Sburlati, W.A. Renner, and J.E. Bailey in M.J.T. Carrondo, et al. eds., *Animal Cell Technology from Vaccines to Genetic Medicine*, Kluwer Academic, Dordrecht, 1997, pp. 621–623.
53. P. Gray, S.C.O. Pak, S.M. Hunt, M.W. Bridges, N.A. Sunstrom, C. Bailey and M.J. Sleigh, *Proc. 10th Intern. Biotechnology Symposium*, Sydney, Australia, 1996.
54. J.L. Biedler and B.A. Spengler, *J. Natl. Cancer Inst.* **57**, 683–689 (1976).
55. R.J. Kaufman and P.A. Sharp, *J. Mol. Biol.* **159**, 601–621 (1982).
56. F.M. Wurm, K.A. Gwinn, and R.E. Kingston, *Proc. Natl. Acad. Sci. USA* **83**, 5414–5418 (1986).
57. G. Zettlmeissl, M. Wirth, H.J. Hauser, and H.A. Küpper, *Gene* **73**, 419–426 (1988).
58. J.H. Nunberg, R.J. Kaufman, R.T. Schimke, G. Urlaub, and L.A. Chasin, *Proc. Natl. Acad. Sci. USA* **75**, 5553–5556 (1978).
59. U.H. Weidle, P. Buckel, and J. Wienberg, *Gene* **66**, 193–203 (1988).
60. J.A. Vitek, *Neoplasma* **34**, 665–670 (1987).
61. D. Pinkel, T. Straume, and J.W. Gray, *Proc. Natl. Acad. Sci. USA* **83**: 2934–2938 (1986).
62. M.G. Pallavicini, P.S. DeTeresa, C. Rosette, J.W. Gray, and F.M. Wurm, *Mol. Cell. Biol.* **10.1**, 401–404 (1990).
63. F.M. Wurm, in H.J. Hauser and R. Wagner eds., *Genetic Manipulation of Mammalian Cells*, Walter de Gruyter, Berlin, pp. 87–120.
64. F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, *Current Protocols in Molecular Biology*, Wiley, New York, 1988.
65. T.C. Hsu, *Int. Rev. Cytol.* **12**, 69–121 (1961).
66. J.R. Birch and W.R. Arathoon, in A.S. Lubiniecki ed., *Large Scale Mammalian Cell Culture Technology*, Marcel Dekker, New York, 1990, pp. 251–270.
67. B.G.D. Bödecker, R. Newcomb, P. Yuan, A. Braufman, and W. Kelsey, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, Butterworth-Heinemann, Oxford, U.K., pp. 580–590.
68. Annual Report 1995, ARES-SERONO: Follicle Stimulating Hormone.
69. R.J. van Wedel, in S. Seaver, ed., *Commercial Production of Monoclonal Antibodies*, Marcel Dekker, New York, 1987, pp. 159–173.
70. J.M. Davies, C.M. Lavender, K.J. Bowes, J.A.J. Hanak, B.S. Combridge, and S.L. Kingsland, in E.C. Beuvery, J.B. Griffiths, and W.P. Zeijlemaker eds., *Animal Cell Technology: Developments towards the 21st Century*, Kluwer Academic, Dordrecht, 1995, pp. 149–153.
71. S. Heimbuch, Press Release of Cellex Biosciences: Cellex Biosciences perfusion equipment used for 1st injectable product produced in hollow fiber licensed by FDA, October 30, 1996.
72. A. Preissmann, R. Bux, P. Schorn, W. Noe, in E.C. Beuvery, J.B. Griffiths, and W.P. Zeijlemaker eds., *Animal Cell Technology: Developments towards the 21st Century*, Kluwer Academic, 1995, pp. 841–845.
73. R.A. Byrn, J. Mordenti, C. Lucas, D. Smith, S.A. Marsters, J.S. Johnson, S.M. Chamow, F.M. Wurm, T. Gregory, J.E. Groopman, and D.J. Capon, *Nature* **344**, 667–670 (1990).
74. W.F. Bennett, N. Paoni, D. Botstein, A.J.S. Jones, B. Keyt, L. Presta, F.M. Wurm, and M. Zoller, *J. Biol. Chem.* **266**, 5191–5201 (1991).

See also EXPRESSION SYSTEMS, *E. COLI*; EXPRESSION SYSTEMS, MAMMALIAN CELLS; INSECT CELL CULTURE, PROTEIN EXPRESSION; INSECT CELLS AND LARVAE, GENE EXPRESSION SYSTEMS; PICHIA, OPTIMIZATION OF PROTEIN EXPRESSION.

CHOLESTEROL OXIDASE

KAZUO AISAKA
Kyowa Hakko Kogyo Co., Ltd.
Tokyo, Japan

KEY WORDS

Brevibacterium sterolicum
4-Cholesten-3-one
Cholesterol
Cholesterol oxidase (EC 1.1.3.6)
3β-Hydroxysteroids
Steroid δ-isomerase (EC 5.3.3.1)

OUTLINE

INTRODUCTION

Cholesterol oxidase is the enzyme that catalyzes the oxidation of the 3β-hydroxy group of various kinds of steroids. The oxidation is the important step in the catabolism of the sterols. This catalytic property plays an important role in various bioprocesses

The catalytic and structural properties of cholesterol oxidases from microbial origins are described. Further, some applications are described, including cholesterol determination in medical diagnosis, reduction of cholesterol content in the food industry, and pest control (based on its insecticidal activity) in the agriculture industry.

CHOLESTEROL

Cholesterol (cholest-5-en-3β-ol) is principal sterol of higher animals. Cholesterol is practically insoluble in water, but it is soluble in aqueous solutions of bile salts. It is found in all body tissues, especially the brain and spinal cord, and in animal fats or oils.

Cholesterol is considered to be a key intermediate in the biosyntheses of fecal sterols, bile acids, and steroid hormones in animals. Turfitt first demonstrated the formation of 4-cholesten-3-one from cholesterol by *Proactiomyces erythropolis* (*Nocardia erythropolis*) and showed that this oxidation of the 3β-hydroxy group to the keto group of cholesterol was the initial step in the aerobic catabolism of the sterol (1). On the other hand, in the intestinal tract of humans, cholesterol can be anaerobically metabolized by bacterial microflora to coprostanol (65%) and coprostanone (10%) and to a lesser extent to cholestanone and epicoprostanol (2) (Fig. 1).

CHOLESTEROL OXIDASE

Cholesterol oxidase (EC 1.1.3.6) catalyzes the oxidation by dioxygen of the 3β-hydroxyl group of cholesterol and of related steroids (other 3β-hydroxy steroids) to a keto group and catalyzes the successive isomerization of the δ5 double bond to the conjugated position: δ4 (Fig. 2). Thus, in many cholesterol oxidases, both the oxidase and isomerase activities are catalyzed by the single enzyme protein, although the two kinds of enzymes, cholesterol dehydrogenase and δ5-3-ketosteroid isomerase (EC 5.3.3.1) are present in some microorganisms containing *Pseudomonas testosteroni* and *P. putida* (3).

The oxidation is estimated by the following methods.

1. Conjugated enone formation by the increase in UV absorption
2. Monitoring oxygen uptake
3. Hydrogen peroxide production by the oxidative coupling of 4-aminoantipyrine with phenol by peroxidase to give a quinoneimine dye absorbing at 500 nm

Generally, the enzyme is a flavin-dependent oxidase containing one mole of tightly bound flavin adenine dinucleotide (FAD) per mole of protein (4). However, cholesterol oxidases from *Nocardia rhodochrous*, *Streptomyces griseocarneus*, and *Streptomyces violascens* lack FAD or any other cofactor.

The enzyme has been isolated from a number of different soil bacteria that are capable of using cholesterol as the sole source of carbon and energy. They include Gram-positive bacteria such as *Arthrobacter simplex*, *Brevibacterium sterolicum*, *Corynebacterium sterolicum*, and *Rhodococcus equi*; Gram-negative bacteria such as *Pseudomonas* spp., actinomycetes such as *Actinomyces lavendulae*, *N. erythropolis*, *N. rhodochrous*, *Nocardia* spp., *Streptomyces griseocarneus*, and *Streptomyces violascens*; and fungi such as *Schizophyllum commune* (5). There are also two types of enzymes, secreted and membrane bound, according to their localization.

Commercial preparations of cholesterol oxidase can be obtained from the following suppliers: *Streptomyces* spp. (15 units/mg) from Toyobo Co.; *B. sterolicum* ATCC21387 (13 units/mg) from Kyowa Hakko Kogyo Co.; *Pseudomonas* spp. (3 units/mg) from Wako Pure Chemical Industries; *R. erythropolis* (25 units/mg) from Boehringer Mannheim; and *Cellulomonas* spp. (20–60 units/mg) and *Schizophyllum commune* (5 units/mg) from Sigma Chemicals.

The enzymes from *B. sterolicum* (6), *Schizophyllum commune* (7), and *Arthrobacter simplex* (8) are purified, and their enzymatic properties have been examined.

This enzyme can oxidize 3β-hydroxy groups in cholesterols, the biosynthetic precursors such as desmosterol, 17 ketosteroids such as dehydroepiandrosterone, and plant sterols such as stigmasterol, β-sitosterol, and campesterol. Of these, the length of the C17 side chain is especially an important factor in determining the reaction rate. On the other hand, 3α-hydroxysteroids, estrogens, and cholecalciferol do not serve as substrates (9). Further, the relative activities are also dependent on the origin of the enzymes.

Murooka et al. cloned the cholesterol oxidase gene from a *Streptomyces* sp. SA-COO and demonstrated high productivity of the extracellular cholesterol oxidase by using the *Streptomyces lividans* host-vector system (10). Ohta et al. cloned the cholesterol oxidase gene from *B. sterolicum* (11). The gene was also overexpressed into *Streptomyces lividans* (12). Both genes show identities of 64% in nucle-

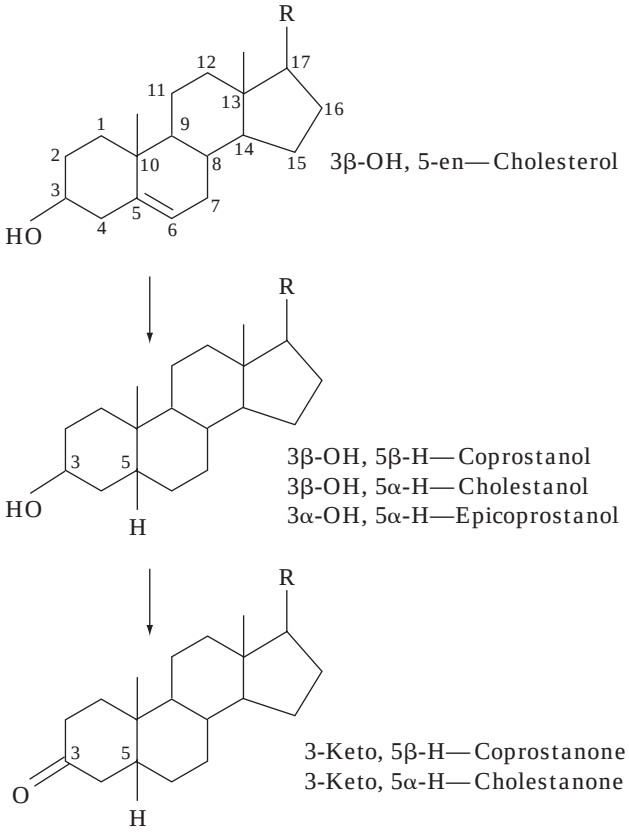


Figure 1. Anaerobic metabolism of cholesterol.

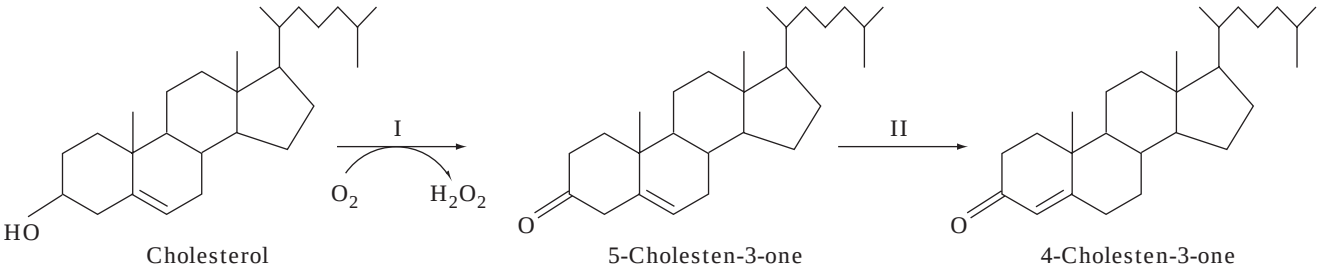


Figure 2. Scheme for the cholesterol oxidase reaction. I, oxidation step; II, isomerization step.

otide sequence level and 58% in amino acid sequence level, respectively. The gene analysis shows that the mature enzyme consists of 503 or 507 amino acids and is preceded by a 42 or 45 amino acid signal sequence for *Streptomyces* sp. or *B. sterolicum*. There are at least six highly conserved regions between the two enzymes, which may play structurally or functionally important roles such as substrate binding or catalysis. The FAD-binding domain in a number of FAD-dependent enzymes contains a pattern of three glycine residues, Gly-X-Gly-X-X-Gly, common to NAD and FAD-binding proteins. This pattern of residues is also found near the N-terminal of the sequences of cholesterol oxidases from *Streptomyces* sp. and *B. sterolicum*.

The crystal structure of the enzyme from *B. sterolicum* has been determined using the method of isomorphous replacement and refined to 1.8-Å resolution (13). The structure is made up of two domains: an FAD-binding domain and a steroid-binding domain. The FAD-binding domain consists of three noncontinuous segments of sequence, including the N-terminal and the C-terminal. The overall topology of this domain is very similar to other FAD-binding proteins. The steroid-binding domain consists of two noncontinuous segments of sequence and contains the roof of the active-site cavity. The active site lies at the interface of the two domains. The FAD lies on the floor of the cavity. The steroid substrate is buried within the internal cavity. The hydroxyl group of the steroid is hydrogen bonded to the flavin ring system of the FAD. The active-site cavity contains one charged residue, Glu361. This residue is appropriately positioned to act as the proton acceptor in the proposed mechanism for the isomerization step.

FAD-dependent cholesterol oxidase shares significant structural homology with another flavoenzyme, glucose oxidase, suggesting that it might also be a member of the glucose-methanol-choline (GMC) oxidoreductase family. A histidine residue is completely conserved in the known sequences of GMC oxidoreductases and may act as the general base in the oxidation reaction.

APPLICATION OF CHOLESTEROL OXIDASE

Diagnosis

Cholesterol oxidase is used in the determination of serum cholesterol concentration for the assessment of arteriosclerosis and other lipid disorders and of the risk of thrombosis.

Richmond (14) and Flegg (15) first showed the possibility of using cholesterol oxidase from microorganisms as a basis for a specific assay for serum cholesterol. In the procedure, 4-cholesten-3-one produced was assayed by its UV absorption after solvent extraction of the reaction mixture. However, their procedures were tedious and time consuming. Allain et al. developed the excellent method in which the hydrogen peroxide produced during the cholesterol oxidase reaction was oxidatively coupled with 4-aminoantipyrine and phenol by peroxidase to form a quinoneimine dye (16). Figure 3 shows the scheme for the enzymatic determination of total cholesterol. First, cholesterol esters are able to be converted to free cholesterol by the action of cholesterol esterase. Then, the liberated cholesterol can be determined by cholesterol oxidase and peroxidase. This enzymatic method is simple and specific under extremely mild conditions. The results with serum samples agreed very well with those obtained by chemical procedures based on the Liebermann-Burchard reaction. This assay system is marketed as a serum cholesterol determination kit; worldwide annual sales amount to $100 to 200 million.

Food Industry

Dietary cholesterol has been implicated as one of the causes of coronary artery disease. Because of this, reducing dietary cholesterol is said to cause a gradual reduction in serum cholesterol and regression of atherosclerotic plaque. The National Cholesterol Education Program coordinating board recommends a daily cholesterol intake in the American diet of 300 mg or less. As a result, reduction of cholesterol in food items relatively rich in cholesterol (dairy foods, eggs) has become a challenging problem.

Some organisms have been shown to degrade cholesterol, as described earlier, but only a few do so without accumulating any steroid intermediates. Watanabe et al. described the degradation of cholesterol in egg yolk and lard by crude cholesterol oxidase in *R. equi* no. 23 and the bacterial cell suspension (17). *R. equi* no. 23 degraded cholesterol with little or no accumulation of steroid intermediates such as 4-cholesten-3-one and 1,4-androstadiene-3,17-dione, which are recognized as key intermediates in a comprehensive pathway of cholesterol degradation.

A biotechnological approach to reducing the cholesterol content of milk involves the genetic engineering of starter culture bacteria (lactococci, lactobacilli, streptococci), leading to the expression of phenotypic traits that are related to cholesterol metabolism. These cultures are then used in

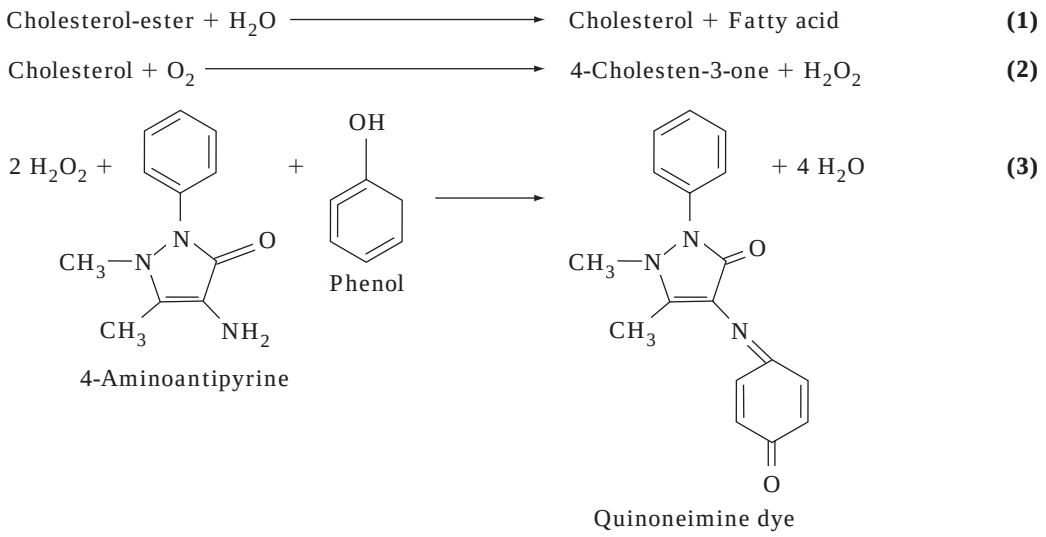


Figure 3. Scheme for the enzymatic determination of total cholesterol. Reactions 1, 2, and 3 are catalyzed by cholesterol esterase, cholesterol oxidase, and peroxidase, respectively.

the production of fermented dairy foods (yogurt and cheeses).

Somkuti et al. have investigated the possibility of introducing the cholesterol oxidase gene from *Streptomyces* sp. SA-COO into a common dairy starter culture bacteria such as *Streptococcus thermophilus* (18) and *Lactobacillus casai* (19). The biosynthesis of a functional cholesterol oxidase in *Streptococcus thermophilus* was confirmed by TLC analysis of reaction products after incubation of sonically disrupted cells with cholesterol. However, unlike *Streptomyces lividans*, *Streptococcus thermophilus* did not secrete cholesterol oxidase into the culture medium.

At present, the use of cholesterol-degrading enzymes in the food industry remains unrealistic because of the lack of toxicological data on enzyme extracts and the intermediate products of cholesterol degradation.

Agriculture

The discovery of novel insecticidal proteins is vital for pest control strategies using transgenic crops. Filtrates from several thousand microbial fermentations were assayed for insecticidal activity against crop-damaging insects. Two *Streptomyces*-culture filtrates killed boll weevil larvae in feeding studies, and the bioactive component was identified as cholesterol oxidase (20). The enzyme was active in a dose-dependent manner, and the LC50 against neonate boll weevil larvae was 20.9 μg/mL after 6 days in diet-incorporation bioassay. *Bacillus thuringensis* subsp. Kurstaki protein has an LC50 of 1 μg/mL against tobacco budworm and 37 μg/mL against European corn borer in similar diet-incorporation bioassay. Thus, cholesterol oxidase is active against the boll weevil in the same concentration range that *B. thuringiensis* protein is active against lepidopterans.

Morphological changes induced by ingestion of cholesterol oxidase suggest that the enzyme has a direct effect on the midgut tissue of boll weevil larvae. Thus, cholesterol oxidase disrupted the midgut epithelium at low doses and lysed the midgut cells at higher doses.

The cholesterol oxidase gene from *Streptomyces* spp. was expressed in tobacco and tomato, and successful control of boll weevil was achieved by the Monsanto company (21).

BIBLIOGRAPHY

1. G.E. Turfitt, *Biochem. J.* **38**, 492–496 (1942).
2. I.A. MacDonald, V.D. Bokkenheuser, J. Winter, A.M. McLernon, and E.H. Mosbach, *J. Lipid. Res.* **24**, 675–681 (1983).
3. P. Talalay, *Ann. Rev. Biochem.* **34**, 347–380 (1965).
4. T. Uwajima, H. Yagi, and O. Terada, *Agric. Biol. Chem.* **38**, 1149–1156 (1974).
5. *Enzyme Handbook*, Springer-Verlag, Berlin, 1995, EC 1.1.3.6.
6. T. Uwajima, H. Yagi, S. Nakamura, and O. Terada, *Agric. Biol. Chem.* **37**, 2345–2350 (1973).
7. M. Fukuyama and Y. Miyake, *J. Biochem.* **85**, 1183–1193 (1979).
8. W.-H. Liu, M.-H. Meng, and K.-S. Chen, *Agric. Biol. Chem.* **52**, 413–418 (1988).
9. A.G. Smith and C.J.W. Brooks, *J. Steroid Biochem.* **7**, 705–713 (1976).
10. Y. Murooka, T. Ishizaki, O. Nimi, and N. Maekawa, *Appl. Environ. Microbiol.* **52**, 1382–1385 (1986).
11. T. Ohta, K. Fujishiro, K. Yamaguchi, Y. Tamura, K. Aisaka, T. Uwajima, M. Hasegawa, *Gene* **103**, 93–96 (1991).
12. T. Ohta, K. Fujishiro, K. Yamaguchi, T. Uwajima, K. Aisaka, and M. Hasegawa, *Biosci. Biotech. Biochem.* **56**, 1786–1791 (1992).
13. A. Vrielink, L.F. Lloyd, and D.M. Blow, *J. Mol. Biol.* **219**, 533–554 (1991).
14. W. Richmond, *Scan. J. Clin. Lab. Invest.* **29**, suppl. 126, abs. no. 3.25 (1972).
15. H.M. Flegg, *Ann. Clin. Biochem.* **10**, 79–84 (1973).

16. C.C. Allain, L.S. Poon, C.S.G. Chan, W. Richmond, and P.C. Fu, *Clin. Chem.* **20**, 470–475 (1974).
17. K. Watanabe, H. Aihara, N. Tachi, and R. Nakamura, *J. Appl. Bacteriol.* **62**, 151–155 (1987).
18. G.A. Somkuti, D.K.Y. Solaiman, T.L. Johnson, and D.H. Steinberg, *Biotech. Appl. Biochem.* **13**, 238–245 (1991).
19. G.A. Somkuti, D.K.Y. Solaiman, and D.H. Steinberg, *Appl. Microbiol. Biotechnol.* **37**, 330–334 (1992).
20. J.P. Purcell, J.T. Greenplate, M.G. Jennings, J.S. Ryerse, J.C. Pershing, S.R. Sims, M.J. Prinsen, D.R. Corbin, M. Tran, R.D. Sammons, and R.J. Stonard, *Biochem. Biophys. Res. Commun.* **196**, 1406–1413 (1993).
21. D.R. Corbin, J.T. Greenplate, E.Y. Wong, and J.P. Purcell, *Appl. Environ. Microbiol.* **60**, 4239–4244 (1994).

CHROMATOGRAPHY, COMPUTER-AIDED DESIGN

ALOIS JUNGBAUER
OLIVER KALTENBRUNNER
Institute of Applied Microbiology
Vienna, Austria

KEY WORDS

Biochromatography
Computer simulation
Neural network
Optimization
Productivity
Simplex optimization
Throughput

OUTLINE

INTRODUCTION

Currently, chromatographic purification processes are optimized empirically. Separation media, running conditions, and sequence of purification steps are selected according to the expertise of the responsible scientists.

In biotechnology, computer-aided design is less far advanced than in certain areas of chemical engineering. The lack of expert systems is the reason for the traditional optimization, which is based on trial and error. Mathematical models for description of the chromatographic separation process have been around since the early 1940s (1–4). The common use of computers in the separation laboratory, the need of rapid process development, and the theoretical knowledge accelerated the development of expert systems and software tools. They are considered to be beneficial for design and optimization of purification steps.

The area of computer-aided chromatography design includes the design of experiments, the appropriate interpretation of experimental results (by the aid of mathematical algorithms), the prediction of separation conditions, and the rescaling of unit operations (5).

Four levels of complexity have to be considered for optimization of separation steps. The first level is the general process economics. This level will be influenced by the second level: flow sheeting. The third level is the optimization of particular unit operations. The fourth level is the sectioning of unit operations into their thermodynamical and molecular considerations by the aid of molecular simulation.

Simulated moving bed, countercurrent chromatography, and annular chromatography are not been considered in this chapter, because these modes of chromatography have not been used for bioprocessing to large extent.

BASIC CONSIDERATIONS FOR COMPUTER-AIDED DESIGN OF CHROMATOGRAPHY

Terminology

The word *model* denotes a system of algebraic or differential equations, which are held to be a representation of a physical system. From this general point of view, any relationship between quantities is a model, and the principles of modeling are the principles of chemical engineering science. *Process models*, and here particularly chromatography, can be counted to a chemical process. They are established by application of the three fundamental conservation principles: conservation of mass, momentum, and energy. Because liquid chromatography is carried out under isothermal conditions, the conservation of energy can be neglected in general. The contribution of liberated

or consumed thermal energy during adsorption and desorption is negligibly small. The conservation of mass and momentum and the virtual breakdown of a chromatography system to a small unit, where the adsorption process takes place, are the fundamentals for modeling of chromatography (6).

In principle, there is no difference between *computer models* and other models. Many models are so complex that they can be solved only numerically. In these cases, computers can be used for reasons of convenience, independent of the complexity of the model applied. Further considerations about modeling are given by Vajda and Valkó (7).

We usually assume that the input, initial conditions, and parameters of a model determine the response (result). The *parameters*, which are the different coefficients in mathematical equations, must be evaluated. This process is called parameter estimation. Once a model is established, the performance of a system can be predicted by varying the input and initial conditions. The prediction will yield a correct answer (response) as long as the predictions are carried out within a certain range of validity.

The words *prediction* and *simulation* describe, in this context of modeling, a similar process. The forecast of process conditions, using process parameters acquired at a small scale, is called prediction. Frequently parameters cannot be estimated because of the lack of a determination method or accessibility. Thus, a hypothetical value is used to solve the set of mathematical equations. This process is called simulation.

Depending on the task, some of the *operational variables* may be kept constant or may not be taken into consideration. The most important variables are listed in Table 1. The residence time of a product in a column is a critical factor, when the product undergoes a conformational change or a reaction in the column. Alteration of the residence time in a column can completely change the result of the purification. Reversed-phase chromatography (RPC) and hydrophobic interaction chromatography (HIC) are more susceptible to residence time changes than ion-exchange chromatography or affinity chromatography are.

The *response values* depend on the strategy and the stage where the chromatographic purification is placed. Frequently used response values are:

- *Productivity* (P) is defined as

$$P = \frac{Q_R \cdot C_0 \cdot V_F}{V_t \cdot t_C} \qquad (1)$$

where Q_R is the recovery ratio, C_0 is the sample concentration, V_F is the sample volume, V_t the column volume, and t_C the cycle time, which includes all time periods necessary to start the next cycle (process time). The purity has to be defined before optimization. The productivity in the aforementioned case does not consider the lifetime of a column.

- *Yield* or *recovery* is simply the percentage of recovered material and material present in the feed. Yield is not easily accessable by mathematical models. In most cases, product loss is caused by a very complex event. Unspecific (irreversible) adsorption to stationary phases, aggregation, and denaturation together with enzymatic degradation are responsible for incomplete recovery. These effects are also dependent on age and history of the column. These events are not covered by chromatography models.
- *Dynamic capacity* is the capacity at a certain flow in contrast to the capacity determined by batch experiments. BTC experiments at 5% breakthrough are usually used to determine the dynamic capacity. Originally, the analytical solution of an idealized model, known as the Thomas solution (8), has been used for approximation and prediction of breakthrough curve (BTC). Under conditions applied in preparative chromatography, the BTC deviates significantly from the ideal model. Recently, models have been published that take effects of kinetics under overloaded conditions into consideration (9).
- To minimize the number of purification steps, the highest *purity factor* should be achieved at each step. The regulations for production of pharmaceuticals by recombinant DNA technologies require very high purity standards. The standards can only be met when a purification power of each step is used to the optimal extent. Other factors, such as design of a plant, quality of the materials and buffers, operator training, etc., must also be obeyed. The optimization of the

Table 1. Important Variables for Optimization of Preparative Chromatography

Column size/shape
Height (Z)
Diameter
Height-to-diameter ratio
Height-to-particle diameter
Flow rate (u)
Flow rate during loading, during elution
During regeneration
Residence time (Z/u)
Physical conditions of packing
Packing quality
Packing density
Mobile phase
Modifier composition
Rate of change of mobile phase
Modifier composition (gradient shape)
Loading
Volume
Concentration
Flow rate
Stationary phase
Ligand type
Ligand density
Type of support
Porosity, tortuosity
Bead diameter
Beads size distribution (monodispers)
Temperature
Scarcely a process in biotechnology, which uses temperature gradients for elution

chromatographic process is only one piece of the puzzle.

- A high *concentration factor* leads to a smaller volume for the next step, which has to be handled and the product becomes more stable. Also, an upper limit in the concentration can be of consideration, because the solubility of the product is exceeded. The concentration factor (CF) can be defined as:

$$\mathrm{CF} = \frac{C_R \cdot V_R}{C_0 \cdot V_F} \tag{2}$$

where C_0 is the initial (feed) concentration, V_F is the feed volume, C_R is the recovered concentration, and V_R is the recovered volume.

- The *process time* may be limited by the stability of the product. Economical considerations also come into play, such as shift time and batches, which have to be produced by a certain time period. The process time includes the optimization of the gradient, flow rate, etc. Process time has to be optimized with regard to total process time. Cleaning, regeneration, and, if necessary, repacking have to be included.

General Criteria of Mathematical Models

Today, broadening of scientific and technological knowledge is based on three cornerstones: theory, experiments, and computer simulation. Reliable computer simulation needs an appropriate mathematical model. Without a valid mathematical model that reflects the real physical situation, the most sophisticated computer algorithm will fail. A model will never be as good as the physical reality, because the complexity has to be reduced to a certain formalism. A fictional model of the complexity of the reality would be as slow and laborious as the reality.

Vajada and Valkó (7) compiled a list of criteria for establishing a model in chemical engineering. These rules also apply for the chromatography of proteins.

- First, scientific curiosity is a strong motivation for establishing a model. A priori curious-appearing models may be of considerable value, when they are integrated with the appropriate objective. This requires the innovative selection of parameters and design of experiments to determine those parameters.
- The second reason for using models is the reduction information and indirect measurements. The most common information reduction in chromatography peak is the approximation of a peak by its first and second moment, instead of using the information from the entire peak curve. Using these information reduction processes, the appropriate model (10–14) must be selected to describe the peak area, width, and the asymmetric shape.
- The ability to design and optimize a separation scheme quickly and effectively is also a motivation for developing a model. To minimize process time and determine optimal column size and optimal gradient shape, loading and sequence of the different process steps (equilibration, loading, washing, elution, and regeneration) are necessary.
- The design of optimal particle properties is also an important goal for manufacturers of resins and for engineers selecting a resin. Modeling of packing provides a directive on the pressure drop and performance loss caused by irregularities (15).
- Control and simulation for validation purposes will become an important issue in modeling of chromatography. The ability to calculate acceptance criteria for scale-up or scale-down will allow rapid, accurate validation of tasks, such as virus clearance. The correct column dimensions, gradient shapes, limits of operation parameters, and rapid diagnosis of separation deficiencies are also added benefits. Scale-up of the production of a biologic is often considered a major change of the process, requiring additional clinical trials that are time consuming and costly. However, if the appropriate scale-up process can be performed with the generation of the same product exhibiting identical qualities, no additional clinical trial would be required. Computer-assisted optimization evaluates the influence of the process change. It would save significant manufacturing costs if the altered scheme were not considered a major change.

The existing models for simulating chromatography are based on different core assumptions. Depending on these assumptions, different levels of complexity are handled. At each level of complexity, different simplifications are made in order to make the mathematical treatment accessible.

According to van Deemter et al. (16) the following assumptions are made:

- The equilibrium concentrations in the two phases are proportional (linear chromatography).
- The exchange process is thermodynamically reversible, that is, the equilibrium between particle and fluid is immediate; the mass transfer coefficient is infinitely high; and longitudinal diffusion and other processes with similar effects can also be ignored. The latter processes are considered ideal chromatography.

From these assumptions one can categorize chromatography into four cases (Table 2). Although the cases 1, 2, and 3 are of great value for understanding chromatography, only case 4 reflects the real situation in preparative chromatography.

Table 2. Modes of Complexity in Chromatography Modeling

	Ideal chromatography	Nonideal chromatography
Linear isotherm	1	2
Nonlinear isotherm	3	4

Source: van Deetmer et al. (Ref. 16).

Differences in Analytical Chromatography

Besides the size and the shape of the separation column, the difference between analytical and preparative chromatography is the intended use (Table 3). In analytical chromatography, the components are separated and identified according to their retention times. In some cases, a component-specific detector is used for the identification. Generally, a small volume (mass) is applied onto the column. Therefore, the separation is carried out under linear isothermal conditions. Frequently, the chromatography can be described as case (Table 2). Preparative chromatography is intended for isolation of a compound. In contrast to analytical chromatography, the separation is carried out in the nonlinear regime. Peak-broadening effects resulting from the nonlinearity of the isotherm must be taken into consideration. The sample is applied as a high concentrated solution or in a large volume as a very dilute solution. The Dirac δ function can no longer be used for describing the sample volume and concentration. The Dirac δ function is a hypothetical approximation for an infinitesimal small volume containing one mass unit.

The whole system is designed so that extra-column effects cannot be neglected. When the performance of a chromatographic separation should be predicted from small scale to large scale, those effects must not be neglected. Special emphasis is put on this problem, because the successful application of computer-assisted optimization is strongly dependent on the extrapolation from small scale to large scale. The extra-column effects can dominate a preparative separation in small scale. In the mathematical description of an analytical column, various mass–transfer problems can be handled by using fine adsorbent particles at a high flow rate. However in a large-scale column, fine particles introduce hydrodynamical problems such as high pressure drop, channeling, and fingering of flow. Consequently, in large chromatography, larger particles are used and mass–transfer resistencies are more pronounced.

Objectives for Computer-Assisted Optimization

The general objective for computer-assisted optimization of chromatography is the reduction of the number of experiments required to identify an optimum. The optimization is accomplished by an iterative process. Frequently, optimal running conditions must be identified. Optimal running conditions can be defined as best buffer for loading, washing, and elution while maintaining highest resolution. The same can be valid for the best throughput (productivity) under highest resolution (17). The latter require chemical parameters (i.e., buffer composition) physical and engineering parameters (i.e., flow rate, height equivalent to theoretical plate). In other words, selectivity, efficiency, and productivity (throughput) must be optimized. Depending on the task, one must rate selectivity, efficiency, and throughput in different ways. Selectivity is mainly achieved by the chromatography sorbent and buffer composition and to some extent by engineering parameters. Efficiency depends on the physical characteristics of the sorbent, the overall setup of the chromatography system, and the chemistry of the mobile phases used. Throughput is influenced by the dynamic capacity of the sorbent, which depends on the composition of the sorbent and the buffers used. All three important characteristics of preparative chromatography are interconnected. Changing one parameter influences the others. Traditionally, after selection of a particular sorbent, one can identify the appropriate separation conditions by seeking the appropriate buffer conditions. Then the throughput is maximized by overloading and increasing the flow rate (18).

Another very important objective in preparative chromatography is the extrapolation from small-scale data to the performance of large scale (scale-up) and the extrapolation from a large scale to small scale (scale-down). When a process is translated from a laboratory scale to a pilot and industrial scale, very often the linear gradient elution must be changed to stepwise elution, because this elution mode is more simple (19,20).

After optimization, the single steps in the sequence of steps have to be optimized, and thus flow sheeting comes into play. In flow sheeting, the performance of the different steps is interconnected, and a strong reciprocal action is observed. However, multiple interconnected experiments must be carried out. The more steps and performance characteristics that are simulated, the fewer experiments needed and thus the faster an optimal separation process is identified.

What are the prerequisites for computer-assisted optimization of preparative chromatography? Because experiments have to be performed in order to have access to parameters for simulation, the experiments should be carried out at the smallest applicable scale. This is not only a heuristic for computer-assisted optimization, but is also valid for all kind of optimization. Whenever possible, the experiments should be carried out with equipment that allows a high degree of automation (21). Electronic data acquisition of a chromatogram is essential for mathematical treatment. The predicted data should be valid for a broad range of flow rates, column sizes, and loading conditions. The main parameters influencing the performance (including the extra-column parameters) should be known. Once the extra-column peak broadening can be discriminated from the parameters of the column, the simulation is fed with this information (22). During scale-up, the contribution of extra columns and column band broadening

Table 3. Differences between Analytical and Preparative Chromatography Other than the Intended Use of the Technique

Analytical	Preparative
Identification by retention or selective detection	Separation coupled with collection of separated proteins
High efficiency	High capacity
Small particles	Large particles
Small loading of columns	Overloading of columns
Linear adsorption isotherms	Nonlinear adsorption isotherms
Sample volume negligible	Sample volume not negligible, (Dirac δ function not applicable) Extracolumn space not negligible

changes. Observing these facts improves the accuracy of the prediction of chromatography from small scale to large scale.

Objectives for Computer-Aided Design

The objective in computer-aided design is the proper selection and order of chromatography steps in a purification process. Based on a primary knowledge on the properties (e.g., isoelectric point, molecular weight, surface hydrophobicity, biospecificity) of the protein of interest and its impurities, expert systems can assist in setting up a prospective purification design.

In parallel, these systems have to consider some basic rules for the separation sequence consisting of capture, purification, and polishing (23):

- Choose separation steps based on different properties.
- Separate the largest impurity first.
- Choose steps that use differences in the properties of the product and the impurities best.
- Use a high-resolution step as soon as possible.
- Set the most laborious step last.

After characterizing the chosen steps experimentally, economic considerations concerning productivity, capital, or production costs must be evaluated. Using computer-aided design routines, several characteristics of the product (amount required in a calculation period, purity required, recovery), resin (cost per liter, lifetime), labor constraints (labor costs, number of operators required), and production constraints (column dimensions available, costs of waste disposal, costs of overheads) can be evaluated to find the optimal scale of an individual purification step.

MATHEMATICAL DESCRIPTION OF THE SEPARATION PROCESS

Chromatographic theories have been applied for description of band broadening and retention of a solute while migrating through the column. In liquid chromatography, the column geometry can be described by the radius (r) and the length (z) of the column. Independently of the model applied, some basic assumptions are made in preparative chromatography of biomolecules:

- The column is considered to be homogeneous in radial direction. The applied protein is distributed homogeneously over the cross-sectional area of the column. As a consequence, the r dimension can be disregarded, and the effective parameter for the column geometry is z.
- The mobile phase is incompressible, and the mobile phase velocity is constant along the column.
- There is no influence of thermal effects. The adsorption process is considered to be isothermal, although this is not consistent with regard to rigorous thermodynamic treatment. The temperature is constant during the experiment, independent of position and time.
- Mass exchanges between mobile and stationary phase are infinitely fast, and the two phases are permanently in composition equilibrium. The need of this assumption is circumvented by kinetic models.
- The ratio between the mobile and stationary phase volume is constant during the experiment, independent of position and time. It is also independent of the mobile phase concentrations of protein and salt.

Depending on the type of chromatography and the intended use, these basic assumptions can heavily deviate from the real conditions.

Plate Models

Using these models, the column is divided into identical theoretical plates, and for each of these plates a mass balance is established. Originally, a linear isotherm was assumed, and in each plate mobile and stationary phase concentrations were in equilibrium. These models introduce the concept of the height equivalent of a theoretical plate (HETP). For a column of length (L) with N theoretical plates, HETP is

$$\mathrm{HETP} = \frac{L}{N} \tag{3}$$

Comparison of HETP of different columns or different conditions is a well-accepted measure for the performance of a column. HETP is not a constant value for a sorbent species, it depends on the column length (24), packing quality, flow rate, solute used for determination, and calculation method (25).

Martin and Synge Plate Model. This model (2) considers the column as a series of stages, and an infinitesimal fraction of the mobile phase volume passes from one stage to the next (Fig. 1). Because an infinitesimal small fraction of volume is transferred from one stage to the next, this model is denominated a continuous plate model. The ratios between mobile and stationary phase volume in each vessel are identical and constant. The mass balance for the jth vessel can be written as

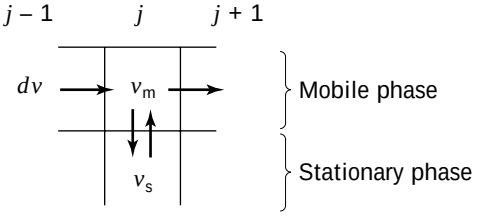


Figure 1. Schematic diagram of the distribution in a Martin and Synge plate model. The infinitesimal small fraction of the mobile phase volume (dv) enters stage j while the identical volume fraction leaves the stage. The distribution between the mobile and stationary phases occurs immediately and infinitely fast.

$$C_{\mathrm{m},j-1} \cdot dv = C_{\mathrm{m,j}} \cdot dv + v_{\mathrm{m}} \cdot dC_{\mathrm{m,j}} + v_{\mathrm{s}} \cdot dC_{\mathrm{s,j}} \quad (4)$$

when $\boldsymbol{dv}$ denotes the movement of the mobile phase volume, v_{m} and v_{s} are the mobile and stationary phase volume of a stage and C is the concentration of the species under consideration. Martin and Synge introduced one mass unit into the first stage. Solving equation 4 with respect to a linear isotherm and sample only in the first stage gives a Poisson distribution.

$$C_{\mathrm{N}} = C_1^0 \cdot \exp\left(-\frac{\mathrm{v}}{v_{\mathrm{m}} + K \cdot \mathrm{v_s}}\right) \cdot \frac{\left(\frac{v}{v_{\mathrm{m}} + K \cdot v_{\mathrm{s}}}\right)^{N-1}}{(N-1)!} \quad (5)$$

which can be approximated by a Gaussian distribution for high plate numbers. In equation 5, C_{N} is the concentration of the Nth stage, C_1^0 is the initial concentration of the first stage, and v is the volume of mobile phase that has left stage N from the start.

Craig Plate Model. Analogously to the Martin and Synge model, this model (26) divides the column into a series of discrete stages (plates), and each stage contains the same volume of stationary and mobile phase. But in contrast to the previous model, the flow is not continuous. It is assumed that the whole amount of mobile phase moves to the next stage after equilibrium is attained (Fig. 2). Then again, the distribution between stationary and mobile phase takes place immediately. Hence, the Craig model is a discontinuous plate model. The mass balance for the jth plate can be written as

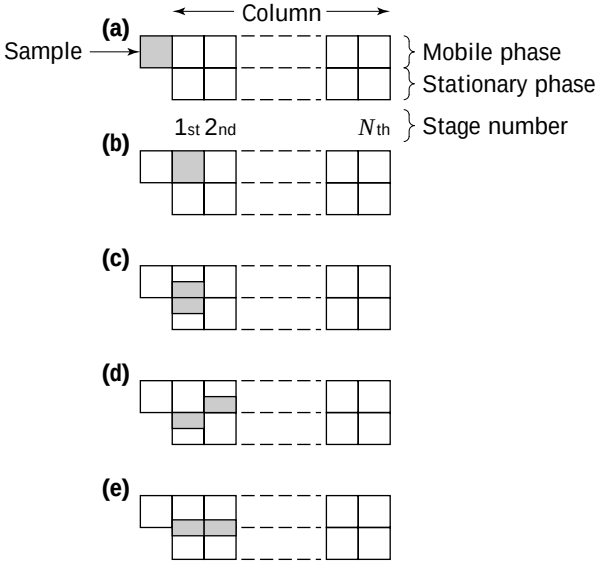


Figure 2. Schematic diagram of a Craig machine for N stages and a constant distribution coefficient. (**a**) A sample segment is arriving at the first stage. (**b**) The segment arrives in the mobile phase of the first stage. (**c**) The sample is distributed between mobile and stationary phases according to the distribution coefficient. (**d**) The mobile phase volumes are moving one stage further. (**e**) Mobile and stationary phases are distributed again.

$$v_{\mathrm{m}} \cdot C_{\mathrm{m},j,t} + v_{\mathrm{s}} \cdot C_{\mathrm{s},j,t} = v_{\mathrm{m}} \cdot C_{\mathrm{m},j-1,t-1} + v_{\mathrm{s}} \cdot C_{\mathrm{s},j-1} \quad (6)$$

The indexes t and $t-1$ denote for the actual and the previous time increment. For a linear isotherm and sample only in the first plate, this model gives a binomial distribution that also can be approximated by a Gaussian distribution for large plate numbers.

In the Craig model, peak dispersion is smaller than in the Martin and Synge model (27). This difference can be related by

$$N_{\mathrm{C}} = N_{\mathrm{M,S}} \cdot \frac{k_0'}{1 + k_0'} \quad (7)$$

when N_{C} is the number of plates of the Craig model, $N_{\mathrm{M,S}}$ is the number of plates of the Martin and Synge model, and k_0' is the retention factor

$$k_0' = \frac{t_{\mathrm{R}} - t_0}{t_0} \quad (8)$$

at infinite dilution. Here, t_{R} is the retention time at the conditions of interest and t_0 is the void time, that is, the retention time assuming nonbinding conditions.

Other Plate Models. The plate models described in this section are extended to kinetic models by introduction of mixing in the stages. Hence, the equilibrium in the stages is no longer instantaneous. The finite time for adsorption and transport kinetics between the phases are compensated by mixing kinetics within the stages. The sectional model is an extension of the Craig model (28). At the end of each time, increment the whole content of the liquid phase of each stage is transferred to the next stage. Therefore, the flow of the mobile phase is discontinuous.

The tanks-in-series model is an extension of the Martin and Synge model (29). The flow of the mobile phase is continuous, and the column is considered as an array of well-mixed cells.

Rate Theory

Although the plate theory was originally developed for linear chromatography with the intention to describe the nonideal effects, this theory was, in the early works, focused on ideal nonlinear chromatography.

The Ideal Model. The mass balance equation was derived independently by Wicke (30) and Wilson (1). In their mass balance equations, axial dispersion was disregarded. DeVault (4) introduced the influence of the pore volume, and Weiss (3) discussed Wilson's approach for several types of adsorption isotherms. The mass balance equation of this ideal model is

$$\frac{\partial C_{\mathrm{m},i}}{\partial t} + H \cdot \frac{\partial C_{\mathrm{s},i}}{\partial t} + u \cdot \frac{\partial C_{\mathrm{m},i}}{\partial z} = 0 \quad (9)$$

Here, C_{m} and C_{s} denote for the protein concentration of species i in the mobile and stationary phase, respectively. H is the phase ratio, t is time, z is column length and u is

the mobile phase velocity. This model is of interest only for nonlinear isotherms. In the case of a linear isotherm and a column of infinite efficiency, the sample pulse travels through the column without any change (31).

Equilibrium-Dispersive Model. According to Van Deemter et al. (16), Giddings (13), and Haarhoff and van der Linde (32), equation 9 can be replaced by

$$\frac{\partial C_{\mathrm{m},i}}{\partial t} + H \cdot \frac{\partial C\mathrm{s},_i}{{}_\partial t} + u \cdot \frac{\partial C_{\mathrm{m},i}}{\partial z} = D_{\mathrm{a},i} \cdot \frac{\partial^2 C_{\mathrm{m},i}}{\partial z^2} \qquad (10)$$

when the mass transfer is not infinitely fast. Here, $D_{\mathrm{a},t}$ is the apparent dispersion coefficient that lumps together axial dispersion and nonequilibrium from adsorption and desorption.

$$D_{\mathrm{a},i} = \frac{\mathrm{HETP} \cdot u}{2} \qquad (11)$$

It is assumed that the mobile and stationary phases are instantaneously in equilibrium. Assuming a column of infinite length, a Dirac $\delta(z)$ pulse as an injection profile, and a linear isotherm, the solution of this equation is a Gaussian distribution (33).

Lumped Kinetic Model. In this model, the dispersion coefficient D_{L} accounts only for the axial dispersion and is not lumped to the influence of nonequilibrium effects.

$$\frac{\partial C_{\mathrm{m},i}}{\partial t} + H \cdot \frac{\partial C\mathrm{s},_i}{{}_\partial t} + u \cdot \frac{\partial C_{\mathrm{m},i}}{\partial z} = D_{\mathrm{L},i} \cdot \frac{\partial^2 C_{\mathrm{m},i}}{\partial z^2} \qquad (12)$$

Depending on the assumptions, different models can be distinguished. If the influence of the mass transfer resistance is negligible and only the kinetics of adsorption and desorption are taken into account, the model is called reaction-dispersive. For a linear isotherm, the kinetic equation is

$$\frac{\partial C_{\mathrm{s}}}{\partial t} = k_{\mathrm{a}} \cdot C_{\mathrm{m}} - k_{\mathrm{d}} \cdot C_{\mathrm{s}} \qquad (13)$$

when k_{a} and k_{d} are the rate constants of adsorption and desorption, respectively. When a Langmuir-type isotherm is assumed, the Langmuir kinetic model is obtained (8).

$$\frac{\partial C_{\mathrm{s}}}{\partial t} = k_{\mathrm{a}} \cdot (C_{\mathrm{s,max}} - C_{\mathrm{s}}) C_{\mathrm{m}} - k_{\mathrm{d}} \cdot C_{\mathrm{s}} \qquad (14)$$

If we assume the adsorption-desorption kinetics as infinitely fast but not the mass transfer kinetics, the model is called transport-dispersive. Then, the kinetic equation for C_{s} is

$$\frac{\partial C_{\mathrm{s}}}{\partial t} = k_{\mathrm{m}} \cdot (C_{\mathrm{m}} - C_{\mathrm{s}}) \qquad (15)$$

when k_{m} is the apparent mass transfer coefficient.

Assuming a Dirac $\delta(z)$ pulse as an injection profile, a linear isotherm, and that mass transfer kinetics in the transport-dispersive model and adsorption-desorption kinetics in the reaction-dispersive model are fast, the solution of these equations is a Gaussian distribution (16). The general rate model consists of two partial differential equations for each component, one for the mobile phase in the interparticle area of the column and one for the fluid in the intraparticle area. It was used by Kucera (34) for linear isotherms and by Lee et al. (35) and Yu and Wang (36) for nonlinear isotherms.

For numerical solutions of these equations several different approaches can be used. The techniques generally applied are finite-difference and finite-element methods. The direct calculation of numerical solutions of the mass balance equations by finite-difference models leads to stability problems of the calculations (37). A method applied extensively replaces the axial dispersion by introducing a numerical dispersion. In this case, a solution for the ideal model is looked for that gives a first-order error equivalent to the axial dispersion. The numerical dispersion can be adjusted by changing the space increment (z) or time increment (τ) of the iteration. The basis for all calculations is a modified form of equation 10

$$\frac{\partial C_{\mathrm{m}}}{\partial z} + \frac{\partial \left[\frac{1}{u} \cdot (C_{\mathrm{m}} + H \cdot C_{\mathrm{s}})\right]}{\partial t} = \frac{D_{\mathrm{a}}}{u} \cdot \frac{\partial^2 C_{\mathrm{m}}}{\partial z^2} \qquad (16)$$

which is identical with

$$\frac{\partial C_{\mathrm{m}}}{\partial z} + \frac{\partial G(C_{\mathrm{m}})}{\partial t} = \frac{D_{\mathrm{a}}}{u} \cdot \frac{\partial^2 C_{\mathrm{m}}}{\partial z^2} \qquad (17)$$

For the ideal model this equation simplifies to

$$\frac{\partial C_{\mathrm{m}}}{\partial z} + \frac{\partial G(C_{\mathrm{m}})}{\partial t} = 0 \qquad (18)$$

Replacing the first term by the backward finite difference and the second by the forward finite difference (backward-forward differences), this equation for the $j + 1$ at the position n is

$$\frac{C^j_{\mathrm{m},\mathbf{n}} - C^{j}_{\mathrm{m},\,-\mathbf{1}}}{h} + \frac{G^{j+1}_n - G^j_n}{\tau} = 0 \qquad (19)$$

This calculation scheme is identical to the Craig model if the ratio of the calculation intervals h/τ equals u. This Craig model was widely used by several authors (38–42). In contrast to a forward-backward scheme (43), which is put to use for isocratic conditions (44,45), it can be applied to calculate gradient elution chromatography (Fig. 3).

More exact calculations with respect to the estimation of the axial dispersion effects are possible using finite-element methods. In chromatography, the method of orthogonal collocation on fixed elements is preferred to the method of moving elements (36,46,47). Arve and Liapis (48–50) call this approach the finite bath method instead of finite element. All these methods are limited to isocratic

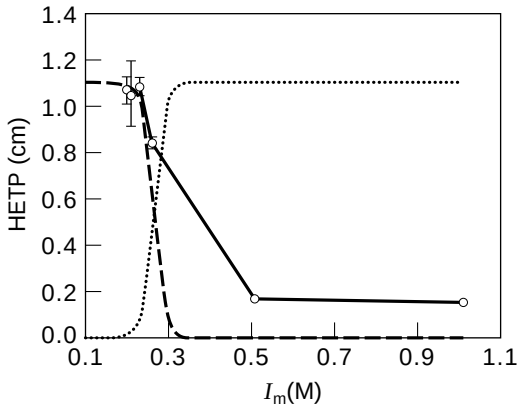


Figure 3. Dependence of HETP on I_m measured by BSA pulses applied to an anion exchanger (Q-HyperD). The column was equilibrated with elution buffers containing various amounts of NaCl. u = 3,006 cm/h; i.d. × h = 5 mm × 5.2 cm. The *dashed line* is a calculation of HETP values according to a Craig model. The *dotted line* occurs when using forward-backward calculation of finite differences.

conditions by means of numerical stability of the calculations (31).

Recently, a fast algorithm based on the assumption of elution as a collocation of point masses migrating with a velocity dependent on the local concentration and the flow rate was presented (51). This model gives comparable results to a Craig model, but the computation time is two orders of magnitude smaller.

Stochastic Theory

A statistical approach developed by Giddings and Eyring (52) has been used for a stochastic model to describe migration and broadening of a sample in a chromatographic column. They considered the migration of a solute in a column as a Poisson distribution process and ignored the axial dispersion. They obtained following probability distribution:

$$P = \frac{k_a \cdot k_d \cdot t_0}{2 \cdot \sqrt{\pi} \cdot (t - t_0)^{0.75}} \cdot \exp[-(\sqrt{k_d \cdot (t - t_0)} - \sqrt{k_a \cdot t_0})^2] \tag{20}$$

This model was extended to more complex situations accounting for axial dispersion and mass transfer resistances (53). For an increasing number of entries of a molecule into the stationary phase and a linear adsorption isotherm, this probability distribution tends to a Gaussian distribution.

Extra-Column Broadening

To be able to describe the behavior of a protein sample in different scales of chromatographic systems, the influence of the extra-column space on the band broadening must be distinguished from the broadening occurring in the column. A comprehensive description of the extra-column effects on gas chromatography is given by Sternberg (54). Most of the explanations are also valid for liquid chromatography (55).

Contributions to Extra-Column Broadening. The extra-column effects can be subdivided into dispersion in tubes, contributions of dead volumes, finite detector volume, and dynamic behavior of transducers and electronics. According to Taylor (56), tubes introduce a symmetrical Gaussian-type broadening (σ_t^2) to the sample. The band spreads under the combined action of molecular diffusion and the parabolic axial flow velocity profile. This broadening can be described by

$$\sigma_t^2 = \frac{t_R \cdot r^2}{24 \cdot D_M} \tag{21}$$

where t_R is the retention time, r is the radius of the tube, and D_M is the molecular diffusion coefficient. Dead volumes create a type of washout kinetic (57). Therefore, they introduce an exponential contribution to the band broadening. This exponential contribution (τ_{dead}) is a function of the dead volume (V_{dead}) and the flow rate (F)

$$\tau_{dead} = \frac{V_{dead}}{F} \tag{22}$$

It was also shown by Sternberg (54) that the finite sensing volume of a detector introduces an additional symmetrical, rectangular broadening. This broadening (σ_d) depends on the sensitive detector volume (V_d) and the flow rate. According to Cram and Glenn (57) it can be described by

$$\sigma_d^2 = \frac{V_d^2}{12 \cdot F^2} \tag{23}$$

An exponential contribution to the band broadening is introduced by the finite response rate of the electronical amplifier and recorder. This contribution (τ_{el}) is independent of the flow rate (57,58).

Mathematical Treatment. Assuming additivity of the variances, the total extra-column broadening can be written as

$$\sigma_{ex}^2 = \sigma_s^2 + \sigma_t^2 + \sigma_d^2 + \tau_{dead}^2 + \tau_{el}^2 \tag{24}$$

where σ_{ex}^2 is the lumped extra column dispersion and σ_s^2 is the variance of the initial theoretical injection profile. Assuming that in liquid chromatography the variances introduced by the electronics and the detector volume are negligibly small compared to the other variances, equation 24 simplifies to

$$\sigma_{ex}^2 = \sigma_s^2 + \sigma_t^2 + \tau_{dead}^2 \tag{25}$$

Lumping together the F-independent variances (σ_A^2) and the F-dependent variances (σ_B^2), the extra-column broadening (σ_{ex}^2) is (59)

$$\sigma_{ex}^2 = \sigma_A^2 + \frac{\sigma_B^2}{F^2} \tag{26}$$

An equation for the prediction of the extra-column contribution to the total broadening can be derived based on the definition of the plate number N of a system

$$N = \frac{V_P^2}{\sigma_{c.i.}^2} \tag{27}$$

If the total broadening (σ_{total}^2) is considered as the sum of the column internal broadening ($\sigma_{c.i.}^2$) and the extra-column broadening (σ_{ex}^2)

$$\sigma_{total}^2 = \sigma_{c.i.}^2 + \sigma_{ex}^2 \tag{28}$$

and N is replaced according to equation 3, the following relationship can be obtained

$$\sigma_{total}^2 = \frac{HETP}{L} \cdot V_R^2 + \sigma_{ex}^2 \tag{29}$$

This equation was derived for analytical chromatography by Huber and Rizzi (60) and has also been applied for preparative chromatography (22). Using this relationship it is possible to exhibit a quantitative picture of the contribution of extra-column broadening on the total broadening (Fig. 4).

ACADEMIC AND COMMERCIAL AVAILABLE COMPUTER PROGRAMS

Computer-aided tools have been developed to different levels of complexity. Programs have been written for prediction of peak profiles, breakthrough curves, gradient shapes, and optimization of resolution (61). On the other hand, expert systems have been created for development of chromatographic separations (62), by using the prediction of peak profiles as tools for optimization.

Expert Systems

An expert system applies heuristic rules for the solution of a problem by using a knowledge base and an interference engine. The knowledge base is a collection of rules and facts gathered from practical experience. The interference engine applies domain knowledge to draw conclusions for the problem under consideration (Fig. 5).

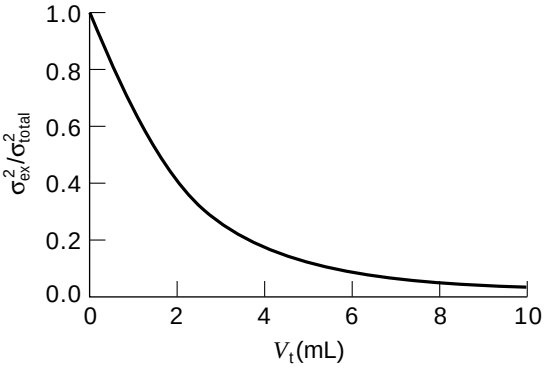


Figure 4. Contribution of extra-column band broadening to total broadening as a function of total column volume (V_t) (22). Original data from experiments using a 0.5-cm i.d. column.

Although the usefulness of expert systems was shown in the late 1960s for chemical applications with the expert system DENDRAL (64), expert systems for chromatography were not introduced until the early 1980s. The expert system ECAT (Expert Chromatographic Assistance Team) (65) provides information for column packing and geometry, and known analyte properties determine the type of chromatographic mode (reversed phase, ion exchange, etc.). Optimizations based on the linear solvent strength model (LSS), assuming a linear relationship between the log k' and the mobile phase modifier concentration, Φ, (66) are offered. A linear relationship between the log k' and the solvent strength of the mobile phase modifier is assumed

$$\log k' = \log k_0' - S\Phi \tag{30}$$

where k' is the capacity factor (normalized retention time in respect to the dead time) defined as

$$k' = \frac{t_R - t_0}{t_0} \tag{31}$$

and φ is the volume fraction of the mobile phase modifier. S reflects the dependency of k' on φ, and k_O' is the capacity factor without the mobile phase modifier.

The expert system expert system for chromatography (ESC) (67) was generated for method development in gas chromatography and HPLC. A set of about 500 chromatograms from the literature constitutes the basis for optimization. Peaks can be simulated, and mobile and stationary phases can be selected. For the peak identification, a curve-fitting tool based on the exponential modified Gaussian function is implemented. Fell et al. (68) developed an expert system to optimize the mobile phase composition in RPC. Therefore, they applied simplex optimization and iterative regression techniques based on experimentally determined retention times.

The European project ESCA (Expert Systems for Chemical Analysis) provided three complex expert systems for method development in chromatography. The first deals with method ruggedness (69); the second, with method (70), and the third, with method development from initial method selection to optimization (71).

The expert system HPLC Doctor (72) and a similar system developed by Tsuji and Jenkins (73) promise help when specific problems with HPLC arise.

Numerous expert systems were developed for RPC (74–76).

A comprehensive description of expert systems in chromatography is given by Hamoir and Massart (62). Presently, for preparative chromatography expert systems are not available. The starting material is very complex, and the purification is generally not completed in one cycle. Expert systems on downstream processing of fermentation products have been developed by Leser and Asenjo (63). These systems comprise optimal flow sheeting of the possible unit operation suited for recovery of bioproducts.

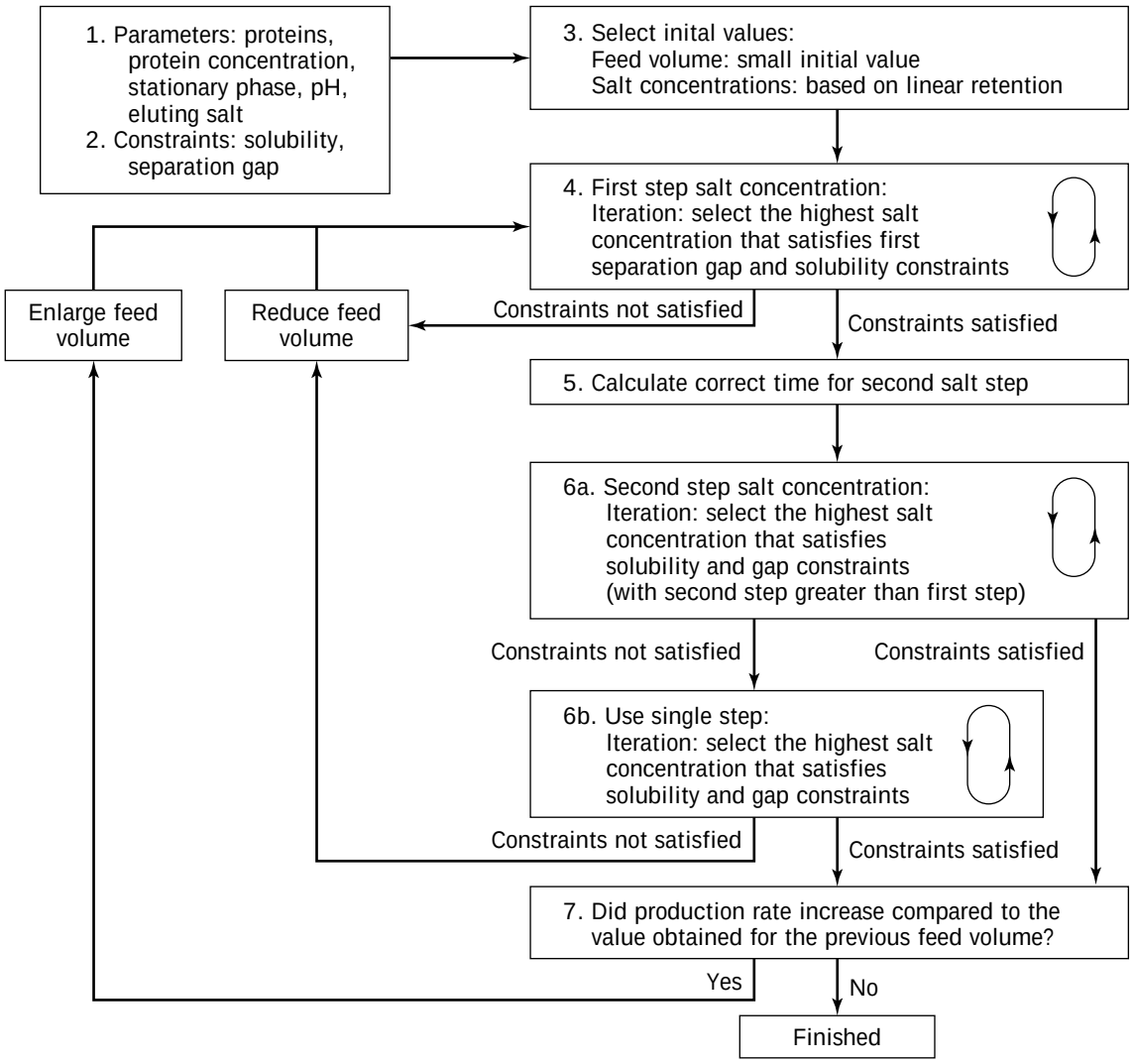


Figure 5. Architecture of the expert system and the links with external databases and algorithms. *Source*: Leser and Asenjo (63).

Leser et al. built an expert system for selecting the sequence of processes for downstream processes, using a commercially available shell (77). The knowledge base contains around 600 logical rules, methods, and structures that emulate the expert's reasoning in selecting downstream processes. The creation of the rationale for the selection of a sequence of high-resolution purification steps must be based on heuristic rules because the existing models for chromatographic separations are not adequate for solving questions of selecting processes, they relate more to the mechanisms of interaction between proteins and matrixes. The heuristic rules are similar to those used to guide the selection of processes for multicomponent mixtures in chemical engineering (Table 4). The rationale for selection of high-resolution purification operations uses the theoretical concept of separation coefficients (23,80).

Simulation Programs

Since the early 1980s, several computer programs for simulating chromatographic behavior have been developed. Most of the programs are only of academic applicability because the experimental accessibility of model parameters is very cumbersome.

Smit et al. (81,82) describe a discrete mathematical model for nonlinear, nonideal chromatography. The nonlinearity can be represented by several different isotherms (linear, Langmuir, Freundlich, sigmoidal). The nonideality is represented by a lumped parameter reflecting axial molecular diffusion, Eddy diffusion, and different sample shapes. No dispersion occurs in the stationary phase. The conservation of mass is used to set up a mass balance. The resulting differential equations are solved by a fourth order Runge-Kutta method.

Phillips et al. (83) use a probabilistic approach for the simulation of a chromatographic separation process, called discrete event simulation. All molecules are executed with the same adsorption–desorption mechanism. Each molecule undergoes a certain number of events (adsorption and desorption). The program can only simulate isocratic conditions and was originally developed for gas chromatog-

Table 4. Classification of Heuristics

Type	Applications and examples
Method heuristics	Specify choice between different unit operations. Describe conditions for preference of microfiltration over centrifugation.
Design heuristics	Specify order or sequence of process steps. Perform low-resolution step before high-resolution step. Use large particles at the beginning of a purification sequences.
Species heuristics	Specify based on the properties of the components. Choose separation method based on the greatest difference in properties of product and impurities. This rule tells us to use always affinity chromatography. There are certain other limitations.
Composition heuristics	Specify feed composition based on separation costs. This rule tells us that affinity chromatography may be too expensive.

Source: Wheelright (78), Nadgir and Liu (79).

raphy. Nonlinear isotherms are also taken into account by applying a molecular density map. The probability of adsorption is related to the accessible (free) binding sites.

Whitley et al. (84) established a very advanced computer program for simulation of peak profiles named VERSEC-LC. The calculation is based on an highly evolved rate model, where the column is virtually divided into different stages (vessels). This model treats the column as a series of N equal volume stages (equivalent to plates or series of vessels or transfer units). The mobile phase flows through these stages at a given rate. Each stage is divided into three fractions: the bulk solution, pore solution, and the solid phase. A gradient between the bulk solute concentration and pore solute concentration results from the combined effects of film and pore diffusion mass transfer resistances. Equilibrium assumed between the solute in the pore solution and the solute on the solid phase is described by nonlinear isotherms. The program can simulate gradients, different loading concentrations, and volumes using different pore structures and diffusion coefficients.

Several computer programs (e.g. Dry-Lab®) use the linear solvent strength theory (LSS) for the calculation of peak positions (27,85). A column is calibrated by two runs, and the optimal gradient shape can be calculated by using a resolution map. The program is mainly suited for RPC, because ion-exchange chromatography (IEC) does not follow the LSS theory (86). The program is intended for optimization of analytical chromatography.

Gallant et al. developed a computer model based on the stochastic mass action (SMA) formalism. SMA is a further improvement of the stoichiometric displacement model (87). The SMA formalism represents the chromatographic adsorption process as a stoichiometric exchange of mobile phase protein and bound counterions. Furthermore, a fraction of counterions is shielded by the adsorbed protein. The accessible sites are corrected by a so-called shielding factor, which is proportional to the stationary phase concentration of protein. For acquisition of the parameters, describing the protein adsorption at different salt concentrations, the LSS formalism is applied, although this formalism is not completely valid for IEC (86). Prediction of protein zone migration and broadening is carried out by a single lumped dispersion model (which is described by equation 10). Gallant et al. (88) propose an optimization routine using simulation of chromatography by the aforementioned models. The optimization runs through different stages. Initially, modeling parameters have to be acquired, and protein constraints (e.g., solubility, resolution) have to be defined. These constraints meet actual requirements of large-scale protein production (Fig. 6).

The program Simulus uses a similar approach as described by Whitley et al. (84), and diffusional resistances and a surface reaction resistance were included. The fundamentals for this program were published by Cowan et al. (29). Furthermore, they have integrated moving grids for calculation of the peak profiles of Wiblin et al. (89). The program is marketed by BIOSEP, AEA Technology (Harwell, U.K.).

ProSys™ Simulation, distributed by BioSepra Inc. (Marlborough, Mass.) is a computer package using the simple Craig simulation (5,90). The computer package is connected to hardware that allows automated determination of the parameters required for optimization of preparative chromatography (ProSys™ Workstation). Currently, it is the only commercially available integrated package of parameter acquisition and simulation software. For IEC, distribution coefficient values as functions of salt concentrations are predicted by linear gradient experiments as proposed by Yamamoto et al. (19). The calculation of simulation parameters can be carried out semiautomatically, assisted by the software. Validation of the simulated peak profiles can easily be accomplished because the software does not discriminate between experimental and simulated runs. So, comparison and superposition are simple. The software package implemented in the ProSys™ Workstation also contains a tool for economical evaluation of chromatographic unit operations (ProSys™ Scale-Up).

All the aforementioned computer programs do not take into account the extra-column band spreading. This mechanism may become very important when it comes to scale-down and parameter acquisition at very small scale (22).

The authors of this chapter have developed a computer program that is an extended version of the simulation program used in the commercial system ProSys. This extended version allows the prediction of peak profiles using a big variety of different adsorption isotherms (90). Furthermore, the influence of extra-column inlet and outlet can be simulated. The modifications of gradient profiles by the mixing device is also part of the simulation (91). Rescaling and influence of the setup can be virtually studied. The simulation program can also be used for validation of chromatographic conditions.

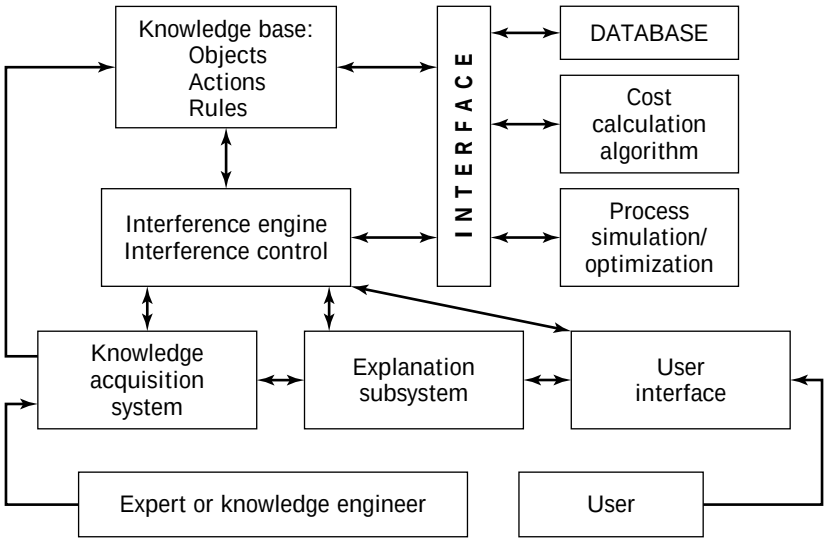


Figure 6. Iterative optimization scheme for step-gradient chromatography, according to Gallant et al. (20). *Source*: Reproduced by kind permission of the authors.

Artificial Neural Networks

Artificial neural networks have been recently applied for chromatographic peak classification (92). Neural networks are specialized computer systems having processing capabilities similar to the human brain. They are distinct from conventional computer systems, because they process entire patterns, rather than successive individual items of data. They are trained by example to carry out their tasks, instead of being explicitly programmed. The processing patterns and learning capabilities displayed by human brains are difficult to mirror using conventional computer programming. Nevertheless, a neural network is, ultimately, a particular kind of computer program, far removed in its theory or implementation from any biological system.

Modern digital computers built with traditional designs have a fundamental limitation, the so-called von Neumann bottleneck. Traditional computation requires a problem to be broken down to a set of operations that are performed in serial fashion, that is, one instruction at a time. Typically, each instruction must be completed before the next instruction is executed. Artificial neural networks represent a fundamentally different approach to computation. They are explicitly designed to mimic the basic organizational features of biological nervous systems: parallel, distributed processing. Artificial neural networks consist of a large number of simple interconnected processing elements, where the processing elements are simplified models of neurons and the interconnections between the processing elements are simplified models of the synapses. The processing of information in such networks therefore occurs in parallel and is distributed throughout each unit composing the network.

Neural networks are particularly appropriate for a defined class of tasks, which may be characterized as follows:

- Pattern recognition, rather than sequential data processing
- tasks demanding nonlinear capabilities in producing their classification, which can be a problem for other techniques
- technically ill-defined tasks, even though the decision may be obvious
- tasks with sufficient examples of the target data available to allow a learning system to be considered

Peak-shape classification is a pattern recognition task that is pertinent to chromatography. The automated determination of the peak profiles in a chromatogram gives an objective answer when two chromatograms are identical or different. It may be a valuable contribution to validation for a purification sequence or validation of scale-up.

APPLICATIONS AND LIMITATIONS OF MODELING

For a conventional software, exact definition of the conditions leading to the optimum is necessary. Mathematical algorithms (explicit equations or differential equations) must be available to calculate the optimum. The major obstacle in preparative chromatography is the complexity of the variables, the problem to estimate the parameters, and the response value. It may be also possible that an optimum is not found within the investigated range of variables.

In Figure 7a and b, a system with a distinct optimum is described. Independently from the starting point (initial pair of variables 1 and 2), an optimum will be found. Systems, where the variable and the response value are related by a peak function or where the variable consists of several variables lumped together, produce a surface with a unique maximum of the response value. The lumped variables or reciprocal action with other variables should be carefully investigated. Minor changes in the experimental conditions may cause a big variation in the optimal response value.

The optimum obtained in a system shown in Figure 7c and d depends on the starting point. Several optima and even local minima are present. Such a system can be characterized by carrying out an optimization strategy with starting points over the whole area in small increments of the considered variables and comparing the reached min-

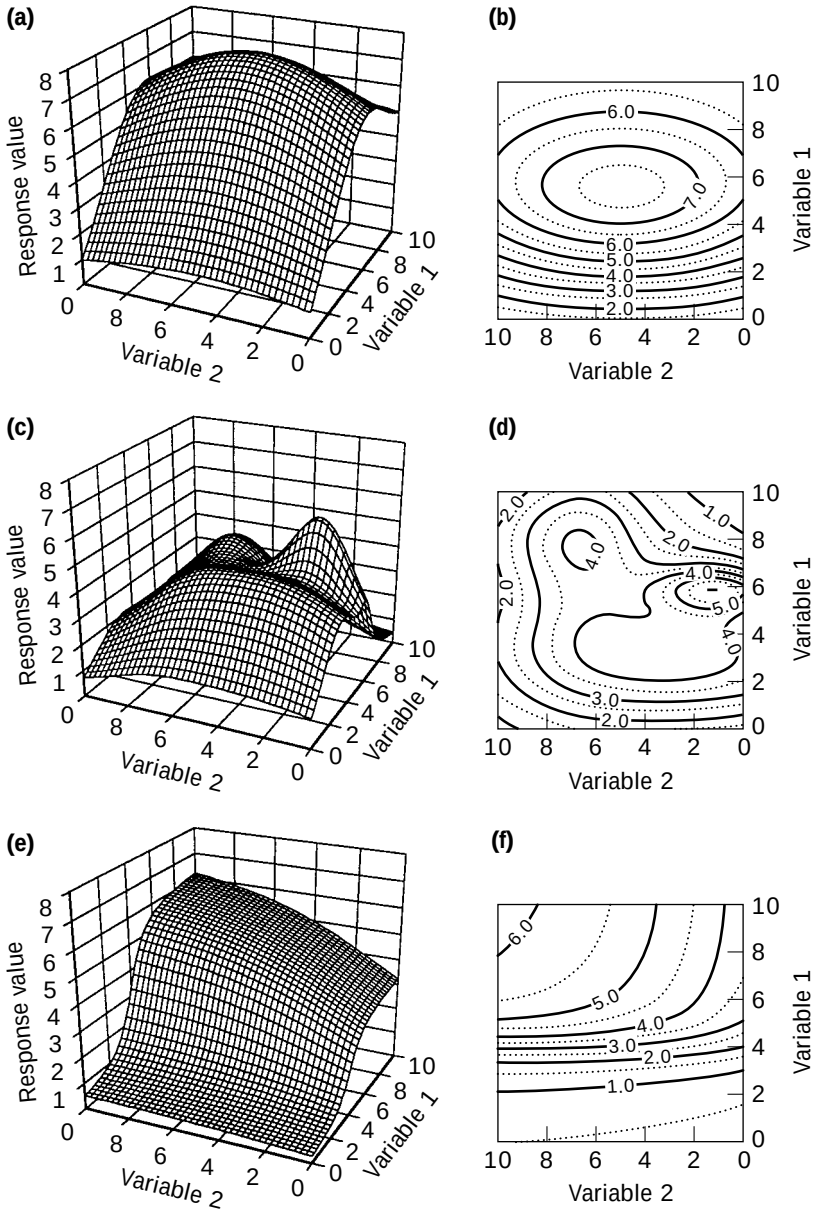


Figure 7. (**a** to **d**) Evaluation of optimum with two variables and one response value. (**e** and **f**) An optimal response value is not found in the considered range. In chromatography, the variable can be, for example, flow rate or loading, and the response value can be, for example, resolution, productivity, or purity.

ima and maxima. Figure 7e and f shows a system where the optimum cannot be reached within the allowed range of variables. This situation is often observed, when the productivity of a chromatographic separation is optimized in respect to flow rate and loading. When the separation is not sensitive to resolution, the system is hypothetically only constrained by the range of the variables. The performance of chromatography is influenced by far more variables than two, but it is very complicated to plot a function with more than two variables in a nonconfusing manner. One way is the overlay of contour plots, as shown in Figure 7b, d, and f and in Figure 8.

Let us assume that a chromatographic separation depends on two variables, such as flow rate and loading. The hypothetical task is to find the optimal response value, such as productivity. Using experimental and simulated results, the optimal response values are searched. A certain purity is presumed. Then the productivity and the purity are plotted versus flow rate and loading. Both contour plots are superimposed. The area, where the presumed purity is reached, indicates the optimal range of operational variables. This approach was used by Mao et al. (93).

Sometimes the conditions are constrained by pragmatic issues, such as availability of equipment (commercially available columns with a certain diameter, costs of a buffer) or by the physicochemical characteristics of the product. Many proteins have limited stability and solubility. There are conditions where the protein may be inactivated. These conditions are not easily transferred into mathematical equations. Sadana (94) summarized possible causes for inactivation of biological macromolecules during separation. During the adsorption and desorption process, proteins may undergo a conformational change. Proteins can associate or they can precipitate on the surface (95–

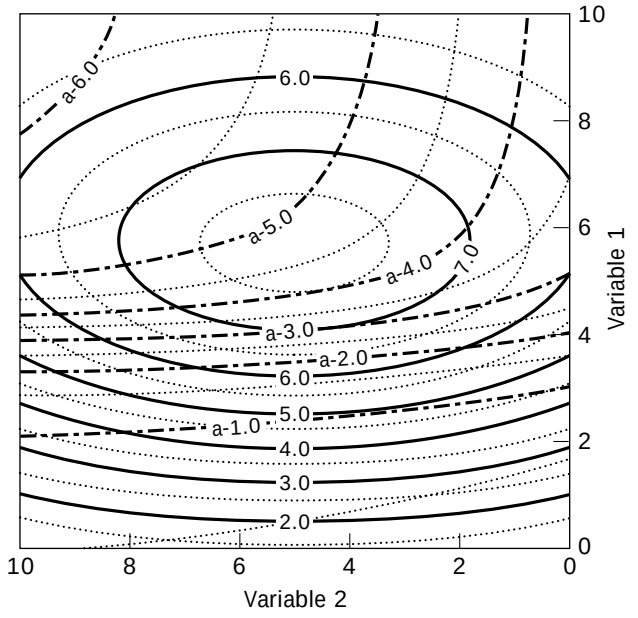


Figure 8. Evaluation of an optimal operational range, when two different response values are related to the two variables. A hypothetical system represented by the graphs **a** and **b** and **e** and **f** in Figure 7 are used to describe the strategy. The contour plots in Figure 7**b** and **f** are superimposed. The intersecting plane formed by the contours with the lower limits of the response values (e.g., a.5.5 and 7.0) yield the optimal range.

97). Self-association was modeled by Lemque et al. (98). They could derive quantitative information on cooperativity effects from the adsorption isotherms. From their model, they showed that protein–protein interaction exists in the adsorbed state and led to an increase in protein self-adsorption. Considering a crude protein sample from a clarified fermentation broth, many proteins and other biological macromolecules are present in this solution. Changing the concentration or pH will affect the self-association. On the other hand, one can also expect the formation of hetero di- and oligomeric protein complexes. Protein-DNA complexes can also occur. The capacity and the lifetime of a column are affected by the aforementioned aggregation phenomenon. This short discussion of protein interaction shows a very important aspect in optimization of chromatography, and complications arise when they have to be translated into a mathematical formulation. One possible way to circumvent these obstacles is the standardization of the sample as far as possible.

Conventional Optimization Methods

Conventionally, resolution has been optimized, because the major activities in computer-assisted optimizations have been carried out in analytical applications. Though some rules may also be applied for optimization of preparative chromatography, in analytical chromatography resolution (R_S) is defined as:

$$R_S = \frac{2 \cdot (t_2 - t_2)}{w_1 + w_2} \tag{32}$$

where t_2 and t_1 are the retention time of the separated substance, and w_1 and w_2 are the peak width at the base. A resolution map is the response value of the optimization. A resolution higher than 1 should be achieved throughout the whole chromatogram. A high resolution is not desired, because it increases the analysis time. The optimization of the running conditions (mobile-phase modifier composition, gradient shape, flow rate, etc.) can be carried out by factorial design, simplex optimization, and others.

Simplex Optimization. Although there are many variants of the simplex procedure (99), they are all based on the basic procedure of Spendley et al. (100). A simplex is defined as a geometric figure having one more point (vertex) than the number of variables being optimized. Thus, for two variables, a simplex is a triangle and for three variables the simplex is a tetrahedron. Although it is difficult to visualize a simplex for more than three variables, the mathematics does not become significantly more complex and the procedure is easily handled by manual or digital computation. Figure 9 shows a two-variable (dimension) simplex as it moves, with constant step sizes, across a response surface. The optimization proceeds by rejection of the vertex, which has the worst experimental response, and by reflecting its coordinates through the midpoint of the hyperplane.

Unfortunately, this original, fixed-step size procedure suffers from several severe limitations. In particular, progress across the response surface is made at a constant rate because the step size is fixed; this fixed size also means that the optimum may not be precisely located because the simplex will be forced (in two dimensions at

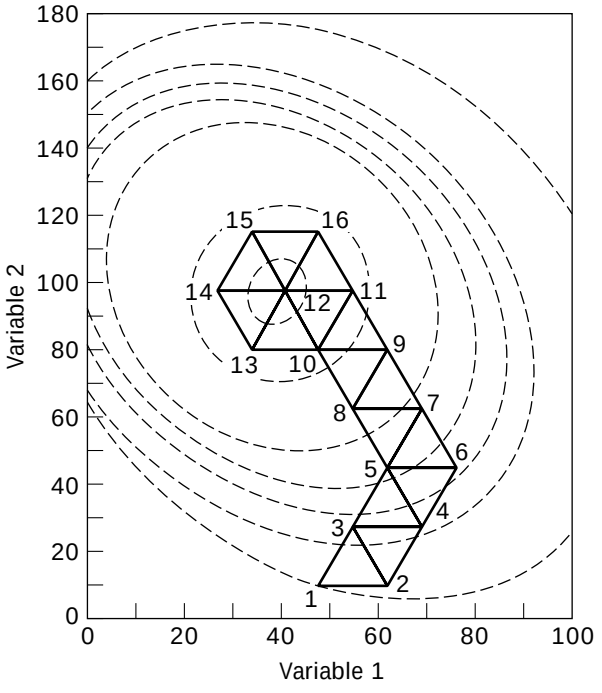


Figure 9. Fixed-step-size simplex optimization of two variables. Initial simplex is 123 and the optimum region is at vertex 10, 11, 12. *Source*: Reproduced with kind permission of Berridge (101).

least) to circle around it. Additionally, the fixed-step size simplex is prone to failure on a surface ridge, in that it will become stranded and not make progress toward the optimum. These limitations have been largely overcome in the modified procedure of Nelder and Mead (102). They introduced two new operations to the fixed-step size procedure, expansion and contraction. The latter operations allow the simplexes to expand and thus accelerate toward the optimum region. When they have approximately located the optimum region, the simplexes contract and reduce the search region until the optimum is precisely located. Figure 10 shows a variable-step size simplex on the same response surface as in Figure 9.

Berridge (101) reviewed the chromatography response functions (CRF) used in analytical chromatography. These functions can be also applied to preparative chromatography (103). The functions successfully used to date for simplex optimization are in general based on sums of terms, usually reflecting resolution and analysis time. The CRF are derived for the separation of a small number of components from a complex sample. A few response functions will be discussed as examples. A comprehensive treatment of the CRF used in analytical chromatography would exceed the scope of this chapter.

For isocratic separations, Berridge (104) developed a function to provide greater flexibility in meeting the needs of fully automated optimization:

$$\mathrm{CRF} = \sum R_i + n^a - |t_m - t_n| - c(t_0 - t_1) \quad (33)$$

where R_i is the actual resolution, n is the number of detected peaks, t_m is the desired retention time, t_n is the retention time of the last peak, t_0 is the dead time, t_1 is the retention time of the first elute peak, and a is a selectable weighting parameter.

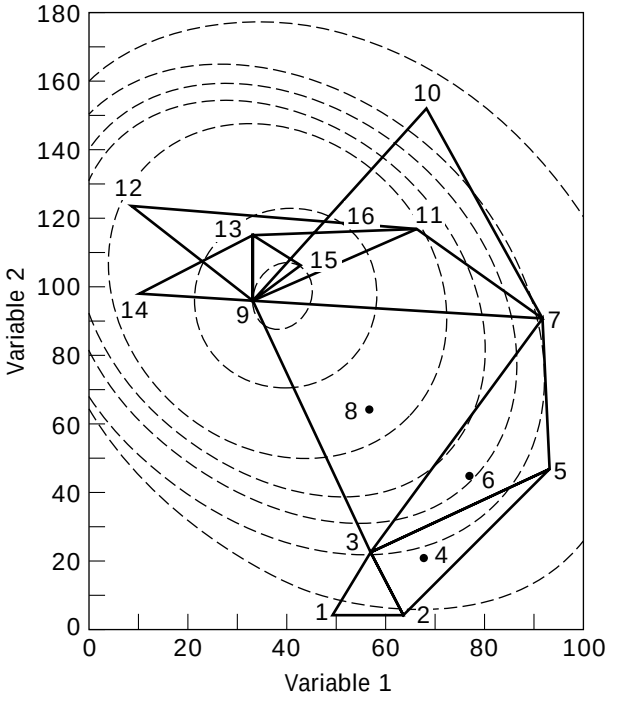


Figure 10. Variable-step simplex. *Source*: Reproduced by kind permission from Berridge (101).

To overcome some of the anomalies generated by using a multicriterion CRF in simplex optimization, Wright et al. (105) simplified the CRF (equation 33) to

$$\mathrm{CFR} = \sum R_i + n - (t_m - t_n) \quad (34)$$

with $t_m - t_n > 1$. The time constraint for the first eluted peak is deleted and that for the last peak is considered only if the difference between the retention time and the target retention time is greater than 1 min. The term for the number of peaks (n) is not raised to a power to ensure that when the term is included, it has a significant influence on the CRF value.

Isocronal Analysis. Atamna and Grushka (106) describe the concept of optimization by isochronal analysis, formerly known as the time normalization for HPLC. Briefly, in this mode of optimization, two experimental conditions are changed simultaneously. One is altered to improve resolution, whereas the other is changed to maintain the analysis time constant. The constant analysis time means:

$$t_{R1} = t_{R2} \quad (35)$$

where the subscripts 1 and 2 indicate two different experimental conditions. From equation 36, the following conditions can be derived:

$$\frac{Z_1(1 + k_1')}{u_1} = \frac{Z_2(1 + k_2')}{u_2} \quad (36)$$

where Z is the column length, u is the linear velocity, and k' is the capacity factor. The need to change two experimental conditions simultaneously is clearly seen from equation 36. The most practical pairs of experimental conditions that can be exploited in isochronal analysis are mobile phase velocity and composition, mobile phase velocity and column temperature, or mobile phase composition and column temperature. When the residence time of a product is of concern, this type of optimization may have some value also for preparative chromatography.

Flow Sheeting

Flow sheeting in bioprocessing is a very complex task. A process for purification of a protein consists of several steps. Usually there is a link between the performance of each step and the performances of the previous step and the subsequent steps. The complexity of the problem is obvious when considering a tree structure, assuming a three-step purification with only two possibilities at each step (Fig. 11). Even in this simplified case we end up with eight different purification schemes. Additionally, an economical evaluation of the individual arrangement is hard because the steps of one sequence are influencing each other.

The use of expert systems is suggested for the rational selection of purification processes starting from different

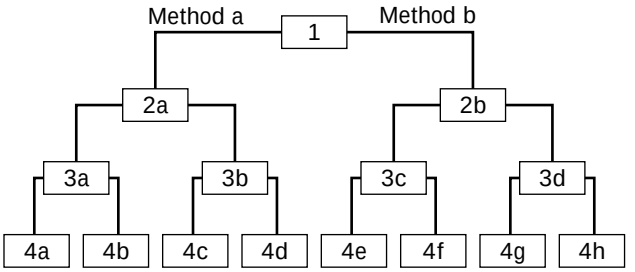


Figure 11. Tree structure of an idealized three-step purification. Each step offers only two possibilities.

microbial sources, such as harvesting, preconditioning, concentrating, purification, and polishing (77,107).

Artificial neural networks are suited to improve the quality of expert systems for flow sheeting. The numerical certainty for defined steps can be fine tuned by feedback with existing results. A successful application has not yet been shown today. However, the large number of software tools for artificial neural networks will promote this application.

The Biopro Designer is a computer program allowing the economical evaluation of each separation step. As in real life, chromatography is limited to other separation steps, such as ultrafiltration, precipitation, etc.

BIBLIOGRAPHY

1. J.N. Wilson, *J. Am. Chem. Soc.* **62**, 1583–1591 (1940).
2. A.J.P. Martin and R.L.M. Synge, *Biochem. J.* **35**, 1358–1368 (1941).
3. J. Weiss, *J. Chem. Soc.* 297–303 (1943).
4. D. DeVault, *J. Amer. Chem. Soc.* **65**, 532–540 (1943).
5. A. Jungbauer and O. Kaltenbrunner, *Biotechnol. Bioeng.* **52**, 223–239 (1996).
6. E.N. Lightfoot, R.J. Sanches-Palma, and D.O. Edwards, in H.M. Schoen ed., *New Chemical Engineering Separation Technique*, Interscience, New York, 1962, pp. 99–181.
7. S. Vajada and P. Valkó, in I. Pallai and G.E. Veress eds., *Studies in Computer-Aided Modeling, Design and Operation. Part B: Systems*, Elsevier, Amsterdam, 1992, pp. 28–95.
8. H.C. Thomas, *J. Am. Chem. Soc.* **66**, 1664–1666 (1944).
9. R.D. Whitley, R. Wachter, F. Liu, and N.-H.L. Wang, *J. Chromatogr.* **465**, 137–156 (1989).
10. E. Grushka, *Anal. Chem.* **44**, 1733–1738 (1972).
11. W.W. Yau, *Anal. Chem.* **49**, 395–398 (1977).
12. J.P. Foley and J.G. Dorsey, *Anal. Chem.* **55**, 730–737 (1983).
13. J.C. Giddings, *Dynamics in Chromatography. Part I. Principles and Theory*, Dekker, New York, 1965.
14. J.L. Wade, A.F. Bergold, and P.W. Carr, *Anal. Chem.* **59**, 1286–1295 (1987).
15. P.L. Spedding and R.M. Spencer, *Comput. Chem. Eng.* **19**, 43 (1995).
16. J.J. van Deemter, F.J. Zuiderweg, and A. Klinkenberg, *Chem. Eng. Sci.* **5**, 271–289 (1956).
17. J.C. Janson and P. Hedman, *Biotechnol. Prog.* **3**, 9 (1987).
18. J.H. Knox and H.M. Pyper, *J. Chromatogr.* **363**, 1 (1986).
19. S. Yamamoto, K. Nakanishi, and R. Matsuno, *Ion-Exchange Chromatography of Proteins*, Dekker, New York, 1988.
20. S.R. Gallant, A. Kundu, and S.M. Cramer, *Biotechnol. Bioeng.* **47**, 355–372 (1995).
21. D. Berger and D. Gillard, *J. Chromatogr.* **210**, 33 (1981).
22. O. Kaltenbrunner, A. Jungbauer, and S. Yamamoto, *J. Chromatogr., A* **760**, 47–53 (1997).
23. J.A. Asenjo, in JA. Asenjo ed., *Separation Processes in Biotechnology*, vol. 9, Dekker, New York, 1990, pp. 3–16.
24. R.M. Nicoud and M. Perrut, in G. Ganetsos and P.E. Barker eds., *Preparative and Production Scale Chromatography*, Dekker, New York, 1993, pp. 47–77.
25. H.P. Lettner, O. Kaltenbrunner, and A. Jungbauer, *J. Chromatogr. Sci.* **33**, 451–457 (1995).
26. L.C. Craig, *J. Biol. Chem.* **155**, 519–534 (1944).
27. L.R. Snyder, J.W. Dolan, and G.B. Cox, *J. Chromatogr.* **483**, 63–84 (1989).
28. Q.M. Mao and M.T.W. Hearn, *Biotechnol. Bioeng.* **52**, 204–222 (1996).
29. G.H. Cowan, in A.E. Rodrigues, M.D. LeVan, and D. Tondeur eds., *Adsorption: Science and Technology, NATO ASI Series E.* vol. 158, Kluwer, Dordrecht, The Netherlands, 1989, pp. 517–537.
30. E. Wicke, *Kolloid-Z.* **86**, 295–313 (1939).
31. G. Guiochon, S. Golshan-Shirazi, and A.M. Katti, *Fundamentals of Preparative and Nonlinear Chromatography*, Academic Press, Boston, 1994.
32. P.C. Haarhoff and H.J. van der Linde, *Anal. Chem.* **38**, 573–582 (1966).
33. L. Lapidus and N.R. Amundson, *J. Phys. Chem.* **56**, 984–988 (1952).
34. E. Kucera, *J. Chromatogr.* **19**, 237–248 (1965).
35. C.K. Lee, Q. Yu, S.U. Kim, and N.-H.L. Wang, *J. Chromatogr.* **484**, 29–59 (1989).
36. Q. Yu and N.-H.L. Wang, *Comput. Chem. Eng.* **13**, 915–926 (1989).
37. B. Lin, Z. Ma, S. Golshan-Shirazi, and G. Guiochon, *J. Chromatogr.* **475**, 1–11 (1989).
38. J.E. Eble, R.L. Grob, P.E. Antle, and L.R. Snyder, *J. Chromatogr.* **384**, 25–44 (1987).
39. M. Czok and G. Guiochon, *Comput. Chem. Eng.* **14**, 1435–1443 (1990).
40. M. Czok and G. Guiochon, *Anal. Chem.* **62**, 189–200 (1990).
41. M.Z. El Fallah and G. Guiochon, *Anal. Chem.* **63**, 859–867 (1991).
42. A. Felinger and G. Guiochon, *J. Chromatogr., A* **658**, 511–515 (1994).
43. M.Z. El Fallah and G. Guiochon, *Anal. Chem.* **63**, 2244–2252 (1991).
44. P. Rouchon, P. Valentin, M. Schoenauer, C. Vidal-Madjar, and G. Guiochon, *J. Phys. Chem.* **89**, 2076–2082 (1985).
45. P. Rouchon, M. Schoenauer, P. Valentin, and G. Guiochon, *Sep. Sci. Technol.* **22**, 1793–1833 (1987).
46. A.I. Liapis and D.W.T. Rippin, *Chem. Eng. Sci.* **33**, 593–600 (1978).
47. T. Gu, *Mathematical Modeling and Scale-Up of Liquid Chromatography*, Springer, Heidelberg, 1995.
48. B.H. Arve and A.I. Liapis, *AIChE J.* **33**, 179–193 (1987).
49. B.H. Arve and A.I. Liapis, *Biotechnol. Bioeng.* **30**, 638–649 (1987).

50. B.H. Arve and A.I. Liapis, *Biotechnol. Bioeng.* **31**, 240–249 (1988).
51. S.E. Asplund and R.K. Edvinson, *Comput. Chem. Eng.* **20**, 507–516 (1996).
52. J.C. Giddings and H. Eyring, *J. Phys. Chem.* **59**, 416–421 (1955).
53. F. Dondi and M. Remelli, *J. Phys. Chem.* **90**, 1885–1891 (1986).
54. J.C. Sternberg, in J.C. Giddings and R.A. Keller eds., *Advances in Chromatography*, vol. 2, Dekker, New York, 1966, pp. 205–270.
55. J.Å. Jönsson, in J.Å. Jönsson ed., *Chromatographic Theory and Basic Principles*, Dekker, New York, 1987, pp. 27–157.
56. G. Taylor, *Proc. R. Soc. London* **219**, 186–203 (1953).
57. S.P. Cram and T.H. Glenn, *J. Chromatogr.* **112**, 329–341 (1975).
58. H. Poppe, *Anal. Chim. Acta* **114**, 59–70 (1980).
59. J.L. Rocca, J.W. Higgins, and R.G. Brownlee, *J. Chromatogr. Sci.* **23**, 106–113 (1985).
60. J.F.K. Huber and A. Rizzi, *J. Chromatogr.* **384**, 337–348 (1987).
61. J.L. Glaich and L.R. Snyder, *Computer-Assisted Method Development for High-Performance Liquid Chromatography*, Elsevier, Amsterdam, 1990.
62. T. Hamoir and D.L. Massart, in J.C. Giddings and R.A. Keller eds., *Advances in Chromatography*, vol. 33, Dekker, New York, 1993, pp. 97–145.
63. E.W. Leser and J.A. Asenjo, *J. Chromatogr.* **584**, 43–57 (1992).
64. J. Lederberg, G.L. Sutherland, B.G. Buchanan, E.A. Feigenbaum, A.V. Robertson, A.M. Duffield, and C. Djerassi, *J. Am. Chem. Soc.* **91**, 2973–2976 (1969).
65. R. Bach, J.F. Karnicky, J. Excoffier, and S.R. Abbot, in R.T. Pierce and B.A. Hohne eds., *ACS Symp. Ser. No. 306*, American Chemical Society, Washington D.C., 1986, p. 278.
66. L.R. Snyder, *Principles of Adsorption Chromatography*, Dekker, New York, 1968.
67. L. Peichang and H. Hongxin, *J. Chromatogr.* **452**, 175–189 (1988).
68. A.F. Fell, T.P. Bridge, and M.H. Williams, *J. Pharm. Biomed. Anal.* **6**, 555 (1988).
69. P.J. Schoenmakers, N. Dunand, A. Cleland, G. Musch, and T. Blaffert, *Chromatographia* **26**, 37–44 (1988).
70. M. Mulholland, J.A. van Leeuwen, and B. Vandeginste, *Anal. Chim. Acta* **223**, 183–192 (1989).
71. T. Hamoir, M. de Smet, H. Pyrins, P. Conti, N. vanden Driessche, D.L. Massart, F. Maris, H. Hindriks, and P.J. Schoenmakers, *J. Chromatogr.* **589**, 31–43 (1992).
72. S.A. Borman, *Anal. Chem.* **58**, 1192A (1986).
73. K. Tsuji and K.M. Jenkins, *J. Chromatogr.* **485**, 297–309 (1989).
74. H. Yuzhu, A. Peeters, G. Musch, and D.L. Massart, *Anal. Chim. Acta* **223**, 1–17 (1989).
75. K. Valkó, G. Szabó, J. Röhricht, K. Jemnitz, and F. Darvas, *J. Chromatogr.* **485**, 349–363 (1989).
76. R.M. Smith and C.M. Burr, *J. Chromatogr.* **485**, 325–340 (1989).
77. E.W. Leser, M.E. Lienqueo, and J.A. Asenjo, *Ann. N.Y. Acad. Sci.* **782**, 441–455 (1996).
78. S. Wheelwright, *Protein Purification: Design and Scale Up of Downstream Processing*, Hanser Publishing, Munich, 1991.
79. V.M. Nadgir and Y.A. Liu, *AIChE J.* **29**, 926–934 (1993).
80. E.W. Leser and J.A. Asenjo, *Ann. N.Y. Acad. Sci.* **721**, 337–347 (1994).
81. J.C. Smit, H.C. Smit, and E.M. De Jager, *Anal. Chim. Acta* **122**, 1–26 (1980).
82. J.C. Smit, H.C. Smit, and E.M. De Jager, *Anal. Chim. Acta* **122**, 151–169 (1980).
83. J.B. Phillips, N.A. Wright, and M.F. Burke, *Sep. Sci. Technol.* **16**, 861–884 (1981).
84. R.D. Whitley, J.A. Berninger, N. Rouhana, and N.-H.L. Wang, *Biotechnol. Prog.* **7**, 544–553 (1991).
85. J.W. Dolan, D.C. Lommen, and L.R. Snyder, *J. Chromatogr.* **485**, 91–112 (1989).
86. M.A. Quarry, R.L. Grob, and L.R. Snyder, *Anal. Chem.* **58**, 907–917 (1986).
87. W. Kopaciewicz, M.A. Rounds, J. Fausnaugh, and F.E. Regnier, *J. Chromatogr.* **266**, 3–21 (1983).
88. S.R. Gallant, A. Kundu, and S.M. Cramer, *J. Chromatogr., A* **702**, 125–142 (1995).
89. D.J. Wiblin, S.D. Roe, and R.G. Myhill, *J. Chromatogr., A* **702**, 81–87 (1995).
90. O. Kaltenbrunner and A. Jungbauer, *J. Chromatogr., A* **734**, 183–194 (1996).
91. O. Kaltenbrunner and A. Jungbauer, *J. Chromatogr., A* **769**, 37–48 (1997).
92. R.C. Rowe, V.J. Mulley, J.C. Hughes, I.T. Nabney, and R.M. Debenham, *LC-GC Int.* **7**, 36–42 (1994).
93. Q.M. Mao, I.G. Prince, and M.T.W. Hearn, *J. Chromatogr.* **646**, 81–89 (1993).
94. A. Sadana, *BioSeparation* **3**, 145–165 (1992).
95. R.H. Ingraham, S.Y.M. Lau, A.K. Taneja, and R.S. Hodges, *J. Chromatogr.* **327**, 77–92 (1985).
96. S.-L. Wu, A. Figueroa, and B.L. Karger, *J. Chromatogr.* **371**, 3–27 (1986).
97. J. Fausnaugh and F.E. Regnier, *J. Chromatogr.* **359**, 131–146 (1986).
98. R. Lemque, A. Jaulmes, B. Sébille, V.-M.C., and P. Cysewski, *J. Chromatogr.* **599**, 255–265 (1992).
99. F.H. Walters, L.R. Parker Jr., S.L. Morgan, and S.N. Deming, *Sequential Simplex Optimization*, CRC Press, Boca Raton, Fla., 1991.
100. W. Spendley, G.R. Hext, and F.R. Himsworth, *Technometrics* **4**, 441–461 (1962).
101. J.C. Berridge, *Chemom. Intell. Lab. Sys.* **5**, 195–207 (1989).
102. N.A. Nelder and R. Mead, *Computer J.* **7**, 308–314 (1965).
103. L.S. Jennings, K.L. Teo, F.Y. Wang, and Q. Yu, *Comput. Chem. Eng.* **19**, 567–573 (1995).
104. J.C. Berridge, *J. Chromatogr.* **244**, 1–14 (1982).
105. A.G. Wright, A.F. Fell, and J.C. Berridge, *J. Chromatogr.* **458**, 335–353 (1988).
106. I. Atamna and E. Grushka, *J. Chromatogr.* **355**, 41 (1986).
107. G.J. Prokopakis and J.A. Asenjo, in J.A. Asenjo ed., *Separation Processes in Biotechnology*, vol. 9, Dekker, New York, 1990, pp. 571–601.

See also ADSORBENTS, INORGANIC; ADSORPTION, EXTENDED BED; ADSORPTION, PROTEIN, BATCH; CHROMATOGRAPHY, RADIAL FLOW; PILOT PLANTS: DESIGN AND OPERATION.

CHROMATOGRAPHY, HYDROPHOBIC INTERACTION

PER KÅRSNÄS
Percreative
Göteborg, Sweden

KEY WORDS

Chromatography
HIC
Hydrophobic interaction
Protein purification
Protein separation

OUTLINE

INTRODUCTION

Hydrophobic interaction chromatography (HIC) is an important chromatography technique used for the separation of biomolecules. The interaction is based on physicochemical characteristics of the molecules other than the charge, as in ion-exchange chromatography (IEC), or the size, as in gel filtration. Although its importance is increasing in industrial applications, the start has been slow. The technique is certainly very powerful when the optimal conditions for a certain separation are determined, but to achieve this has often been far from easy. One reason is that the mechanism of the interaction is complex, depending on several properties of the separation medium as well as of the separated substances, and not yet fully understood. There are a number of hypotheses that approach the phenomenon of hydrophobic interaction from different points of view, but the practitioner will still find that there is no simple theory that will always facilitate the optimization of a separation, as is the case when using other chromatography techniques. Hydrophobic interaction is favored by high salt concentration and is thus a logical purification step after salt precipitation or IEC, where the environment has a high ionic strength.

HIC has been used as a group separation method, rapidly separating what retards on the column from what is not interacting, and as a powerful high resolution technique with gradients of descending salt concentration.

In addition to what is usually designated as HIC, there are several related chromatography techniques based entirely or partly on hydrophobic interaction. Reversed-phase chromatography (RPC) is the most important, working with strong hydrophobic interaction that is often achieved by high substitution with long aliphatic chains. The principal difference between HIC and RPC is methodological in nature. In HIC, the interaction is increased by salt in order to retard the sample on the column and, basically, a return to native conditions will elute the compounds, whereas in RPC the sample is retarded on the column and eluted by addition of organic solvents. The methods based on what are described as salt-mediated and thiophilic interactions gain in importance, whether they should be considered as HIC techniques is perhaps a matter of semantics. Moreover, the binding in affinity chromatography and to dye–ligand matrixes often have significant elements of hydrophobic interaction.

HYDROPHOBIC INTERACTION

Basic Theory

There are several approaches to theoretically elucidate the nature of hydrophobic interaction, ranging from thermodynamics over surface tension to van der Waals forces (1,2). From different angles they all try to explain the observed influence on hydrophobic interaction by ions with salting-out effect, chaotropic ions, type of salt, temperature, pH, and additives. No single theory, however, manages to fully cover the mechanisms of interaction. In HIC the added complexity, contributed by matrixes, spacers, ligands, and the effects of buffer environment on different species of biomolecules, implicates that, apart from very general recommendations, the design of a HIC procedure still has to be founded on careful practical and empirical studies.

Salting Out and Hydrophobic Interaction

There is definitely a relationship between the salting out process used for the precipitation of proteins and hydrophobic interaction as used in chromatography. The Hofmeister series list anions and cations according to decreasing ability to precipitate and increasing ability to dissolve proteins, and the combinations (salts) from the beginning of the list are well known to promote hydrophobic interaction (3–6) (Fig. 1). Probably precipitation by salting-out and HIC are based on the same phenomenon, the principal difference being that in precipitation the hydrophobic interaction has to take place between molecules of the same kind or others that happen to be present in the solution, whereas in chromatography a well-defined counterpart is hydrophobic enough to bind some of the passing compounds, preferably at lower salt concentration than is needed for precipitation.

Thermodynamic Theory

The main and very elegant theory of hydrophobic interaction (7) refers directly to the Gibbs free energy equation (equation 1), which describes the change in energy when a chemical system is transferred from one state to another:

$$\Delta G = \Delta H - T\Delta S \qquad (1)$$

where ΔG is the change in Gibbs free energy, ΔH is the change in enthalpy or heat content, T is the temperature in Kelvin, and ΔS is the change in entropy or disorder of the system. If ΔG is negative, the second state has a lower energy and the system will adapt to it, provided that there are no energy barriers. In hydrophobic interaction, the entropy change is of greatest importance (8). If hydrophobic molecules such as aliphatic carbon chains are immersed in water, the monolayer of water molecules in contact with a carbon chain will be in higher order than those in the bulk of the water. If two or more hydrophobic structures come together, the surface that has to be covered by ordered water molecules decreases, some water molecules join the less-ordered bulk water, and the entire system gains in entropy and thus, according to equation 1, decreases its free energy (Fig. 2). This state will be favorable for energetic reasons and thus promoted. If some of the hydrophobic molecules are attached to a chromatographic matrix as ligands, the setup for HIC is obvious (Fig. 3).

The phenomenon of increased interaction in the presence of salting out ions is explained by a higher gain in entropy when water is transferred from the surface of a nonpolar molecule to the bulk of water. Whether this is achieved because the surface water in the presence of salt is initially more ordered or the bulk water less ordered or both is not quite clear.

Equation 1 reveals that at a higher temperature an increase in entropy will give a higher energy gain because the entropy factor of the equation will be multiplied by a higher number. Consequently, if the binding to a medium has been shown to increase with the temperature, this has been taken as an indication that the interaction is really hydrophobic.

Surface Tension, van der Waals Forces

Another theory (3) suggests that the surface tension of the water surrounding a nonpolar structure and the tendency to minimize it are responsible for the interaction. If salt is added to the water, it has been shown that the surface tension increases. As a consequence, the energy gain in minimizing the surface exposed to the hydrophobic molecules is larger and thus the interaction is stronger. Some salts that influence the surface tension are listed in order of decreasing effect in Fig. 1. Although the salts that increase the surface tension most are among those that should promote hydrophobic interaction according to the Hofmeister series, the order is not quite the same.

A third theory suggests that van der Waals forces (4) are responsible for the hydrophobic interaction in HIC. This is supported by the fact that these forces should be increased as the order of water increases in the presence of salt. Probably the theories mentioned here all describe one part of the complicated interaction, and a combination of them is more relevant to describe the phenomenon of hydrophobic interaction.

(a)

Anions: Increasing "salting-out" effect

$\longleftarrow$

PO_4^{3-}; SO_4^{2-}; CH_3COOH^-; CI^-; Br^-; NO_3^-; ClO_4^-; I^-; SCN^-

(b)

Cations: Increasing chaotropic effect

$\longrightarrow$

NH_4; Rb^+; K^+; Na^+; Cs^+; Li^+; Mg^{2+}; Ca^{2+}; Ba^{2+}

(c)

Salts promoting the molal surface tension of water:

$Na_2SO_4 > K_2SO_4 > (NH_4)_2SO_4 > Na_2HPO_4 > LiCI > KSCN$

Figure 1. The Hofmeister series of anions (**a**) and cations (**b**). The *arrows* indicate increasing salting-out effect and chaotropic effect, respectively. Salts with high ability to promote hydrophobic interaction are combinations of ions found to the left in the series. The same salts are found early in the series of salts listed in order of decreasing ability to promote the surface tension of water (**c**). This indicates the relationship between hydrophobic interaction and surface tension. Which salt to chose depends on the solubility and the production cost of the salt as well as the cost of preventing pollution of the environment by waste buffers.

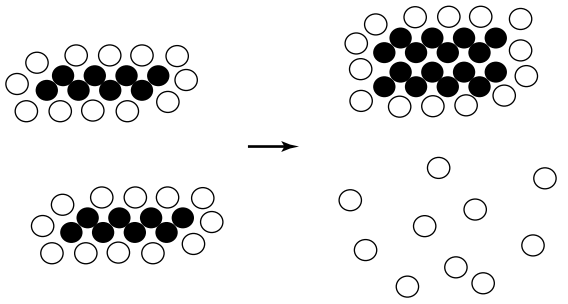


Figure 2. Hydrophobic interaction. Two aliphatic carbon chains (*black circles*) are immersed in water. Only the water molecules (*white circles*) with high order in the monolayer adjacent to the nonpolar molecules are indicated. When two chains are in contact, the surface area that is covered with ordered water molecules is decreased and some water is in a state of less order in the bulk of the water. The free energy of the system decreases, and thus the binding together of the two molecules is favored.

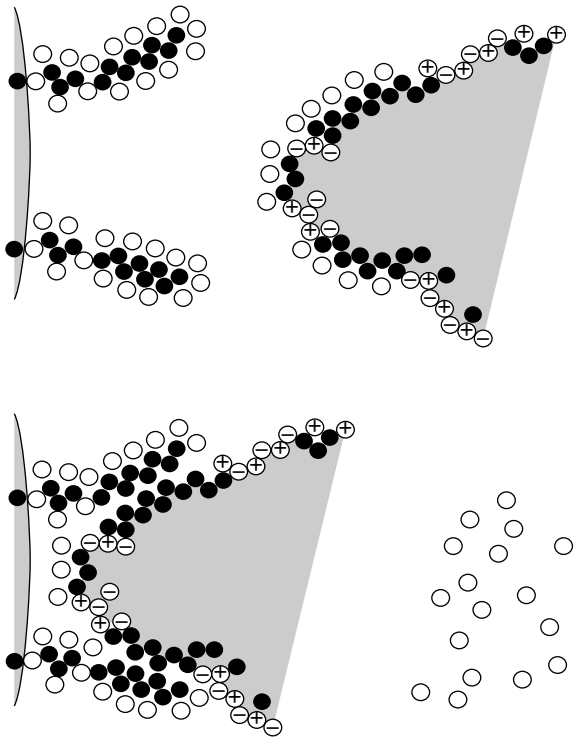


Figure 3. Hydrophobic interaction used for chromatography. Two octyl chains are coupled to a chromatographic medium. An energy gain, equivalent to that described in Figure 1, is achieved by the interaction of the octyl chains with nonpolar moieties on a protein. The carbon atoms of the hydrophobic parts of the protein are indicated with *black circles*, the oxygen atoms of the water with *white circles*, and charged atoms or groups of the protein with *white circles* containing plus or minus signs. The *hatched area* indicates the interior part of the protein.

HYDROPHOBIC INTERACTION CHROMATOGRAPHY

Development of the Technique

As early as 1948 Tiselius presented what he called Adsorption separation by salting out (9), but not until the beginning of the 1970s were reports on the synthesis of media published. Some researchers were working with a mixed mode because of their medias' content of charges (10–13), but charge-free hydrophobic media were also synthesized (5,14,15), and the first commercially available products were presented some years later (16). At present, a wide range of products, including HIC-media for HPLC, are available but phenyl, octyl, butyl, and alkyl ligands are still by far the most common (17,18).

Factors Influencing HIC

In addition to the environmental parameters, including ionic strength, temperature, and pH that influence the hydrophobic interaction as such, there are several factors that have to be controlled to use the hydrophobic interaction for chromatographic purposes. The chromatographic medium is constructed by coupling hydrophobic ligands to a chromatographic matrix, often by use of a short spacer to enhance the sterical possibilities of interaction with large molecules. The type of ligand and the ligand density affect the characteristics of the medium. The choice of buffer ions has been found to change the interaction pattern more profoundly than generally anticipated. The optional use of organic solvents and detergents have to be taken into consideration. The molecules themselves have certain characteristics and will often change chromatographic interaction behavior with the buffer salts in an individual manner (19). Finally, the choice of dimensions of the chromatographic column and the mode of elution will depend on the size range of biomolecules that are subject to processing and the desired throughput of the separation step. One approach, to describe the complex interaction of HIC from a practical point of view, has been to characterize a HIC medium by the interaction pattern of a significant set of protein molecules and to evaluate the result by the use of principal component analysis (20). As described in later sections, the media can be classified by this approach, and the importance of different characteristics of the proteins can be correlated to the interaction of different media.

The Matrix

Agarose-based matrixes with varying degrees of cross-linking are most commonly used for HIC, but organic resins and silica-based media are also used, especially in HPLC applications. All common chromatographic matrixes, such as cross-linked agarose (21,22), organic polymers (23), and silica (24), can be derivatized with hydrophobic groups to give HIC media. Normally, the influence of the by-design hydrophilic chromatographic matrix will just add slightly, if at all, to the performance of the medium, but it is a well-known fact that unsubstituted agaroses intended for gel filtration will expose what can only be characterized as hydrophobic interaction at extremely high salt concentrations (25–27). This allows for using HIC

even for rather hydrophobic proteins that would stick irreversibly to conventional hydrophobic gels. It has to be noted, though, that under normal salt concentrations (e.g., those used in gel filtration) no hydrophobic interaction will be taking place because the gels were successfully designed to be as hydrophilic as possible.

The Ligands

Alkyl, butyl, phenyl, and octyl groups are the most common ligands in commercially available HIC media (Fig. 4). They are listed in what is commonly considered as in order of increasing hydrophobicity, but as mentioned earlier there is no linear or single dimension relationship between their interaction with different species or sets of molecules. It is obvious that the aromatic part of the phenyl group can bind with π-to-π binding as well thus adding to the complexity of the interaction. Other ligands of interest include mercapto derivates, which have yet another spectrum of interaction. They have been claimed to work partly with a special thiophilic interaction, but there is some doubt whether their content of sulfur is really responsible for their interaction pattern or whether they are just another type of hydrophobic medium among the others and working by a combination of the same mechanisms as other hydrophobic media.

The Coupling Reactions

Glycidyl ether is the coupling agent used most often to immobilize hydrophobic ligands (18), but in principle all the common coupling procedures can be used to attach the ligands to the chromatographic matrix. An extensive review of coupling methods is presented in the work by Carlsson et al. (28). The coupling procedure will almost inevitably result in the ligands being attached via an often very short spacer, which is an advantage because conformational restrictions when large protein molecules bind are decreased. On the other hand, the choice of coupling chemistry will always influence the performance of the chromatographic medium by introducing hydrophobic or charged moieties, and the selectivity and the capacity may be very different even if the coupled ligand happens to be the same.

Binding Capacity

The binding capacity of an HIC medium depends partly on the ligand density (11,21,29,30), but it is important to remember that any figures given will always be related to the specific protein that was used to determine the capacity and to the conditions under which the determination was performed. The capacity can be determined as total capacity, with the protein exposed to the medium without flow, or as dynamic capacity, where the linear flow and the residence time of the molecules in the column will influence the result. There may well be a difference of magnitude of 10 to 20 between the capacities for different proteins, a fact that has to be considered when choosing a medium from a catalog where the dynamic capacity is often determined using serum proteins like albumin.

The capacity that can be used is dependent on the contamination picture as well and whether the main part of the impurities are more or less hydrophobic than the desired substance, so in practice it is necessary to determine the dynamic capacity in every specific application; this should be seen as a normal part of an optimization process.

Much work is being done to optimize the dynamic capacity of media, both by suppliers and by users at research sites, but the question is whether the most convenient and uncostly way of increasing the capacity of a column is to increase the amount of gel!

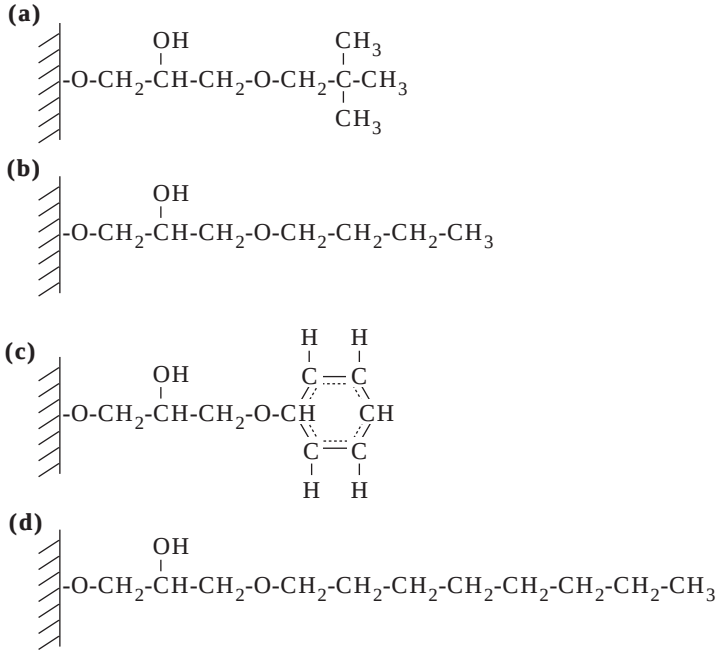


Figure 4. Common hydrophobic ligands. Alkyl (**a**), butyl (**b**) phenyl (**c**), and octyl (**d**) ligands coupled to a chromatographic matrix by glycidyl ether.

HIC and RPC and Column Design

HIC and RPC are both based on hydrophobic interaction, and the borderline between the techniques is more defined by difference in methodology and the type of sample than by the interaction as such.

In RPC the hydrophobicity of the media is higher because the ligands are more hydrophobic and the ligand density is higher. When the sample is run through an RPC column at conditions in the range of physiological pH and salt concentrations, most compounds are retarded on the column. Elution has to be achieved by using increasing gradients of organic solvents. Commonly used solvents are, in order of increasing elution ability, methanol, isopropyl alcohol, and acetonitrile.

In contrast, in HIC very little of the sample is retarded at physiological conditions. The ionic strength of the sample has to be increased to promote attraction, and the elution is normally done by decreasing salt gradients. If some components are still attached to the matrix in weak buffer, a mild organic solvent such as ethylene glycol will be used to wash them out.

As a consequence, HIC is more often used for protein separation where the task often is to retain the tertiary conformation and the biological activity of the separated molecules, whereas RPC is often used for smaller molecules, such as peptides, which lack tertiary structure and subsequently will withstand the elution conditions without denaturation.

Historically, RPC was developed as an HPLC technique to separate small molecules using mainly silica and organic polymer-based matrixes, and HIC was developed as a technique for the separation of proteins, as an alternative to IEC and gel filtration and using the common agarose matrixes as carriers of the hydrophobic ligands. The difference in molecular size range of the molecules to be separated leads to several important differences in the methodologies of the two techniques (31). The differences in behavior between small and large molecules are important in all attraction techniques, such as IEC, HIC, and RPC. But because of the restrictions to expose large molecules to organic solvents, the borderline between separation of small and large molecules happens to coincide more or less with the borderline between the two hydrophobic techniques. But there are other technical reasons behind the difference of methodology based on molecular size (31).

Small molecules interact with the matrix by single point attachment. A molecule is bound to the gel for a certain fraction of the time and is free to move with the buffer the rest of the time. Another way to express the same fact is that at a certain time one fraction of the molecules attach to the column while the rest are moving with the flow. Even if the equilibrium is shifted far toward binding, the macroscopic effect will be that small substances move down the column at some rate in a broad range of salt concentrations, although the speed may be rather low. In a mixture of substances, the rates of movement down the column will be different, as described by the individual equilibrium constants, and a separation will take place. To allow for the molecules to separate, a long path length is needed, and thus long columns and keeping the separating zones sharp are essential for high effectivity. This fact has led to a hunt for high plate numbers, a feature that can be achieved by using small matrix particles. This in turn will cause a high backpressure and the need to work under HPLC conditions. The full utilization of the selectivity of the column will be achieved by elution under isocratic conditions or by shallow gradients. The resulting extremely resolving techniques have been very successful and impressive and have been used for peptide fingerprinting and other types of qualified analysis of small organic molecules.

Proteins, on the other hand, interact with the matrix by multipoint attachment. Even if at some salt concentration, some of the interacting parts of the large molecules may be unbound, there are often several points of interaction left that keep the protein attached to the column. Not until a certain elution condition is reached, such as higher ionic strength in IEC or lower salt concentration in HIC, where all binding sites of the protein are simultaneously unbound, will the protein move down the column. The range between when the protein is not moving at all and moving at full speed is very narrow compared to when small molecules are separated. Isocratic elution is not possible because at constant buffer concentration some proteins will move without retardation and some will hardly move at all. A gradient is needed to elute the proteins in order. Because the resolution is not very dependent on the effectivity of the column, an HIC step where proteins are separated should be performed in short columns, and the linear flow should be kept low by increasing the diameter of the column. In large scale, the use of short and wide columns in protein separation based on IEC and HIC opens up the possibility to run with a high total flow and throughput and still keep the linear flow low enough to allow for the adsorption and desorption to take place without being too far from equilibrium.

CLASSIFYING HIC MEDIA BY MULTIVARIATE ANALYSIS

The performance of HIC media can be monitored by multivariate analysis, using a set or probe of protein molecules with different hydrophobic properties. The interactions of the molecule set with the HIC media could be correlated to molecular weight, isoelectric point, fraction of hydrophobic amino acids, surface charge, and several hydrophobicity indexes (20) without the need for detailed knowledge about the parameters or the mechanisms involved. Instead, the different combinations of media and chromatography conditions can be fingerprinted by their performance in practice and easily compared. As well as artificial mixtures of proteins, serums with 15 to 20 of the proteins identified have been used as probes to determine adsorption patterns (32).

Inversely, biomolecules can be classified by their hydrophobic characteristics as measured by their interactions with a standard set of different hydrophobic media. Molecules with similar adsorption patterns over the media should have similarities on a molecular level.

The results hint that in the tested buffer system, there are at least two main mechanisms that have an impact on HIC (20,33). One mode of interaction was typical on highly

hydrophobic media, for example, with an octyl chain as ligand, where the binding seems to be entirely correlated to the fraction of hydrophobic amino acids and to the hydrophobicity index of the protein molecule. An explanation may be that a higher content of hydrophobic amino acids automatically leads to more hydrophobic amino acids being exposed on the surface of the protein. But because the fraction of exposed hydrophobic areas of proteins is always in the vicinity of 40 to 50% (34,35), another possible explanation takes into consideration that proteins are not rigid bodies with rigid surfaces but that the conformation should be considered as a statistical mean around which the molecule must have the property to breath or change the microconformation, thus allowing a long hydrophobic chain to come in contact with hydrophobic structures that statistically are not always exposed at the surface.

An interesting finding is that the binding to the octyl medium is highly correlated with the molecular weight as well. This phenomenon is logical, though, taking into consideration that larger protein molecules generally have to have a higher fraction of hydrophobic amino acids to keep the tertiary structure.

The other mode of action does not correlate with the content of hydrophobic amino acids of the proteins at all, but instead correlates inversely with the surface charge of the molecules as found in biochemical handbooks. This interaction was displayed on weakly hydrophobic matrixes with short alkyl chains but also with some thiophilic media. An explanation is that charged groups will mask the type of hydrophobic interaction taking place on these gels under the tested conditions. Probably this interaction will be influenced significantly by altering pH.

Very commonly used HIC media, based on phenyl groups and butyl groups as ligands, were shown to work with a mixed mode, that is, combination of the two pure mechanisms. This fact adds to the explanation of the difficulty, especially using these HIC media, to find simple rules after which the separation should be optimized.

It should also be noted that the performance of a HIC medium will always be the result of the combination of ligand, spacer, and matrix. Phenyl gels from different manufacturers have quite different patterns of interaction if tested by a set of proteins as described earlier. This has to be remembered when, after the optimization in laboratory scale, alternative media are tested before scaling-up of a HIC procedure. This optimization is only valid for the medium used. There is thus a risk that competing media, which might be advantageous, will be tested under nonoptimal conditions and rejected under false premises.

THE CHROMATOGRAPHY CONDITIONS

Ionic Strength

A high ionic strength is needed by definition to attach biomolecules to the column. Very often, 1.5 to 1.7 M ammonium sulphate is used as initial buffer concentration without further trials to optimize. Another common approach is to equilibrate the column in weak tris-buffer to preserve the pH and increase the ionic strength by the addition of sodium chloride to a concentration of up to 3 M. It has to be noted, though, that ammonium sulphate, at a certain concentration, is between three and four times as effective in promoting the hydrophobic interaction compared to sodium chloride, so if the substances in question need a stronger interaction to be retarded on the column than can be achieved by almost saturated sodium chloride, ammonium sulphate is a natural choice. Using ammonium sulphate extends the range of HIC to be used for the separation of less hydrophobic molecules.

In production-scale operations, the use of ammonium sulphate has been questioned because of the effect the nitrogen of used buffers may have on the environment by increasing the overnutrition. The choice between overnutrition and possible toxic effects of substitute buffers such as sodium sulphate may not be easy, and all the factors, from the conveniency of the separation procedure to the cost of purifying the sewage, will have to be considered.

Buffer Ions

Recently, it has been shown that the choice of buffer ions greatly influences the function of an HIC medium (32,36). The overall behavior of a buffer-ligand-spacer-matrix system could be switched in a wide range, such as from typical phenyl-like to typical butyl-like or vice versa, by just changing the buffer ions. Optimization work by testing different buffers might then be more beneficial than usually anticipated but, as with pH, the task still includes much work based on trial and error.

pH

It has been shown that hydrophobic interaction between proteins and an octyl gel generally increases with lowered pH and most often decreases with high pH, but the pH dependence in this case is not very profound from pH 6 to pH 8.5. For some proteins, though, the interaction increases with an increase in pH (21). One reason for the different results might be that sodium sulphate was used above pH 8.5 and ammonium sulphate buffer was used below pH 8.5. According to Oscarsson and Kårsnäs (32), the adsorption behavior of serum proteins on octyl gel is quite different in the two environments. The reason for this complexity again has certainly to do with the multimechanism nature of HIC, with irregularities implied by different kinds of interactions between buffer ions and protein surfaces (18) and thus, secondarily, interaction with the HIC matrix.

In practice, changes in pH to get a desired separation can be very powerful because it may alter the entire adsorption spectrum, but, because of the trial-and-error nature of the task, pH is not the primary parameter to use for an optimization.

Temperature

High temperature promotes the binding, and HIC is preferably run at room temperature if the compounds to be separated can withstand this condition without losing activity (4,7,37,38). In room temperature, slightly lower buffer concentrations can be used than if the separation is made in a cold room. Elution of hardly bound substances

may also be facilitated by performing the elution step at cold room temperature. In practice, however, the increase of temperature at the reequilibration of the column in room temperature inevitably produces air bubbles that will destroy the packing of the column, so if a two-temperature scheme is chosen, repacking of the column will most certainly have to be included in the separation cycle.

Flow, Residence Time

Hydrophobic interaction shows signs of being a slower process compared to other interactive chromatography techniques. The desorption of compounds from the column is often run under nonequilibrium conditions, resulting in long tailing peaks that seriously damage the resolution.

It is also generally found that the adsorption of molecules sometimes seems to be a two-step reaction, where the initial binding is weaker and more readily broken under the elution conditions. After some time the bound molecules seem to adapt themselves to the column, possibly involving conformational changes (39). More harsh conditions may have to be introduced for the elution, and there is a risk that the recovery will decrease. On the other hand, there are reports showing that an elution pattern can be quite reproducible even when the bound substances have been immobilized on the column for a long time (21), and HIC is generally considered a high-recovery technique (40).

Elution

Generally, elution of substances bound to an HIC column can be achieved by altering one of the parameters that promote binding. It is possible to elute by increasing pH, if the adsorption was performed at low pH, or by lowering the temperature, although the effect of this may be too small. But most commonly, elution is performed by a descending salt gradient or, in large scale, by stepwise lowering of the salt concentration (5,14).

If substances remain on the column at very low buffer concentration, an organic additive such as ethylene glycol or a detergent can be used for the elution. The ethylene glycol will alter the water structure and weaken the hydrophobic interaction, and the detergent will act by competing with and displacing the protein on the chromatographic medium.

REGENERATION AND CLEANING-IN-PLACE

For process economy, the regeneration of the HIC column is essential. Fouling of the column and irreversible binding of contaminants to the column are common obstacles in HIC. There are three degrees of regeneration: removal of strongly bound or precipitated proteins, removal of hydrophobic substances like lipids and lipoproteins, and sanitation, which means removal or deactivation of microorganisms. Preferably, the regeneration is performed as a cleaning-in-place (CIP) procedure without costly and cumbersome repacking of the column. All washing steps should be performed with reversed flow to give the contaminants the shortest distance out of the column and avoid secondary interaction with the rest of the column. First, all conditions that will decrease possible hydrophobic interaction can be used for the regeneration. Detergents that will compete with the attached substances and ethanol and isopropyl alcohol, buffers normally used in RPC, are obvious choices NaOH, 0.5 to 1 M, has been shown to almost universally sanitize the column and has a good cleaning ability as well. It is important to note that washing with narrow zones of the washing solution is often more effective in removing the column contaminants than exposing the medium for a constant concentration of the washing agent for an extended time. It is not unusual when cleaning HIC media that material will be washed out both in the beginning and at the end of a washing zone where the concentration is increasing and decreasing, respectively. Probably, a lot of equilibria are present, including hydrophobic interaction and interactions involving charges, and if one type of binding is minimized another may be enhanced and keep the contaminants on the column. By wobbling the concentration, the differences of reaction velocities are probably utilized, with elution as the result.

OPTIMIZATION PROCEDURES

Although all the parameters mentioned above can be used to optimize a hydrophobic interaction chromatography step, for many of them the effect of a change is not easy to predict because it will also be dependent on the individual properties of the proteins subject to separation. Thus, a pH change may alter the entire elution pattern or just influence the binding of a single component. Other parameters that may change the entire interaction picture are the buffer ions and the chosen chromatography medium.

Choice of Medium

Ideally, it is preferable that the desired components attach to the medium under moderate salt concentrations (e.g., up to 1 M ammonium sulphate), and a screening of the most common commercially available media is always advisable. This screening is easily performed in small scale and should include a test of whether elution with a high recovery is possible with water or weak buffer. If a lower salt concentration is sufficient on a certain gel, this is a preferable choice for economical and environmental reasons provided that the resolution obtained is still acceptable.

Determination of the Starting Buffer Concentration

An HIC run is normally performed by equilibrating the column in a starting buffer, applying the sample, and eluting with a stepwise descending ionic strength or a descending continuous gradient.

Optimizing the starting concentration, or the concentration at which the proteins are adsorbed to the column, is important to really utilize the HIC step (Fig. 5). In laboratory scale and analytical experiments, where the resolution can be achieved by gradient elution, the ionic strength should be high enough to retard the desired component and to let it elute somewhere in the middle of the

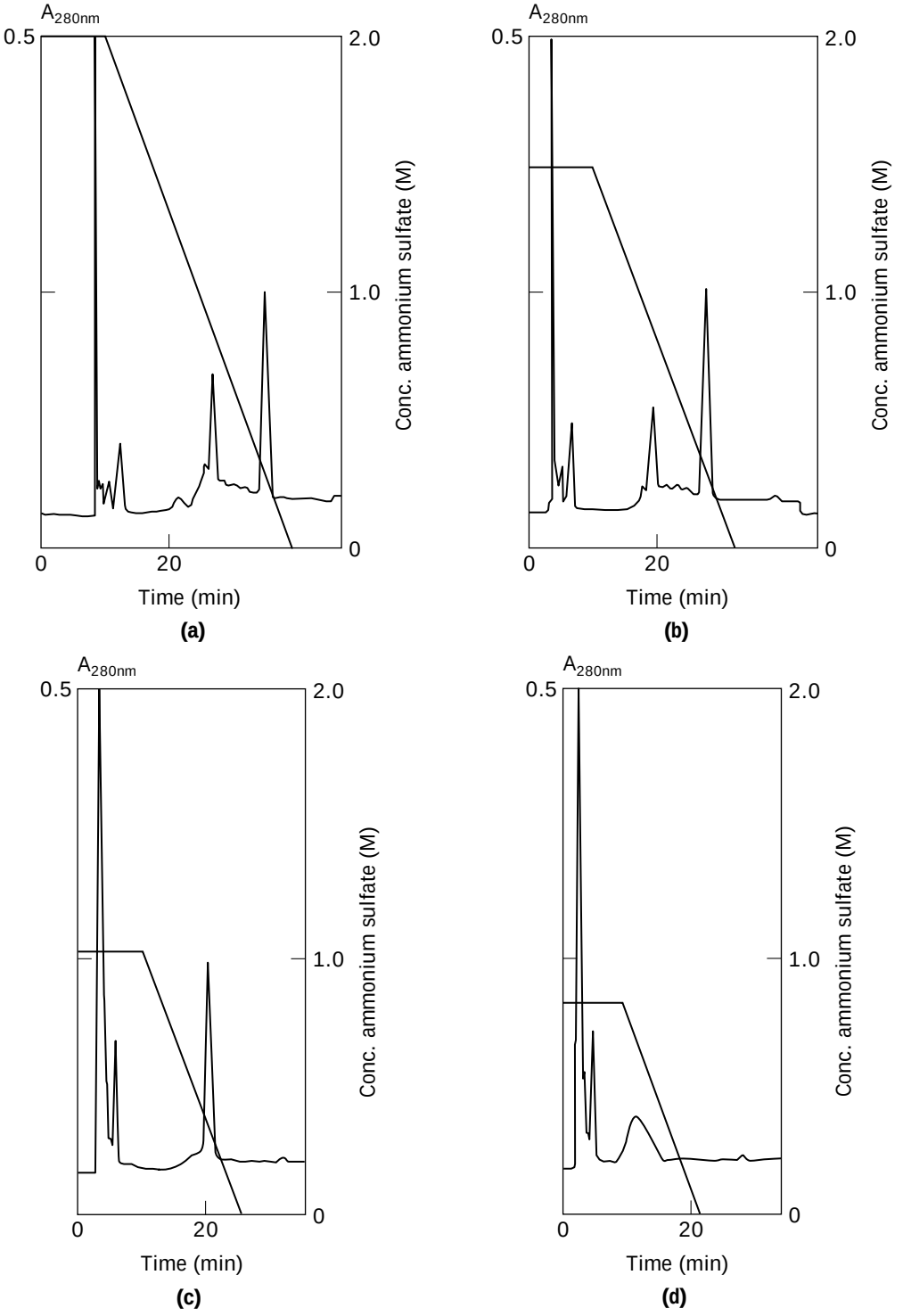


Figure 5. The effect of starting concentration. 100 μL anti-CEA MAB (−IgG1) from mouse ascites fluid, in 0.8 M $(NH_4)_2SO_4$ (AS), applied to an Alkyl Superose HR 5/5 column with a flow rate of 2.5 cm/min. The different concentrations of AS in the initial buffer were mixed in 0.1 M sodium phosphate, pH 7. (**a**) The sample is applied in 2 M AS. Both albumin and IgG are accumulated on the column. (**b**) The sample is applied in 1.5 M AS. Less albumin is bound, which indicates a weaker interaction and less binding capacity. (**c**) In 1 M AS only IgG is binding, and (**d**) in 0.8 M AS the IgG is still retarded but eluted in a broad peak. *Source*: Figure reprinted from ref. 17 by kind permission of Pharmacia Biotech, Uppsala, Sweden.

descending gradient. In this way, the combination of the capacity for the component of interest and the resolution will be favorable. In large-scale runs, well-defined and reproducible gradients are more difficult to achieve, so the elution is preferably done by stepwise decrease of the ionic strength. In some cases, it will work beautifully to chose an ionic strength that is just enough to retard the substance of interest, but in other cases, the peak containing the desired substance will broaden unacceptably at this condition, and the salt concentration will have to be increased (Fig. 5).

Gradient Volume

A gradient volume between 5- and 10-column volumes will give enough resolution in most cases. By increasing the gradient volume, the resolution will increase, but the components will be eluted in a larger volume.

Selectivity of the On and Off Reactions

It is often found that the reaction of binding to the column is rather rapid and selective, whereas the elution of a compound from the column is slower. As a result of this, tailing of zones at elution is very often experienced and the resolution of substances is jeopardized. A way of optimizing a method fully, utilizing the selectivity of the adsorption, is to use the following scheme:

1. Equilibrate the column with a starting buffer that has a lower concentration than is needed to retard the desired substance, adjust the sample accordingly, and run it through the column.
2. Elute everything that was retarded on the column with water or low salt. Check for activity.
3. Add some salt to the buffer and the sample and repeat steps 1 to 3 the desired substance is eluted from the column. To speed up the procedure, a prerun with a gradient starting at high salt concentration will hint at a proper initial concentration when step 1 is run the first time.

An optimized separation scheme includes as its first step running the column in a salt concentration just below the one where the desired substance was found to be retarded. All impurities that bind harder than the desired substance attach to the column and should be eluted. The next step is to bind the desired substance using a slightly higher salt concentration. If this is done carefully, very few contaminants will bind, and subsequently the desired substance can be eluted in high yield by a steep descending gradient, if any contaminating substances were retarded as well, or by a step of very diluted buffer. The described method is advantageous not only because the selectivity of the interaction is fully utilized but also because of the simplicity to recondition sample and buffers by just adding salt. Moreover, the capacity of the column is utilized to bind mostly the compound of interest and very little contaminants. A drawback may be that the capacity of the column for the substance of interest might be lower be-

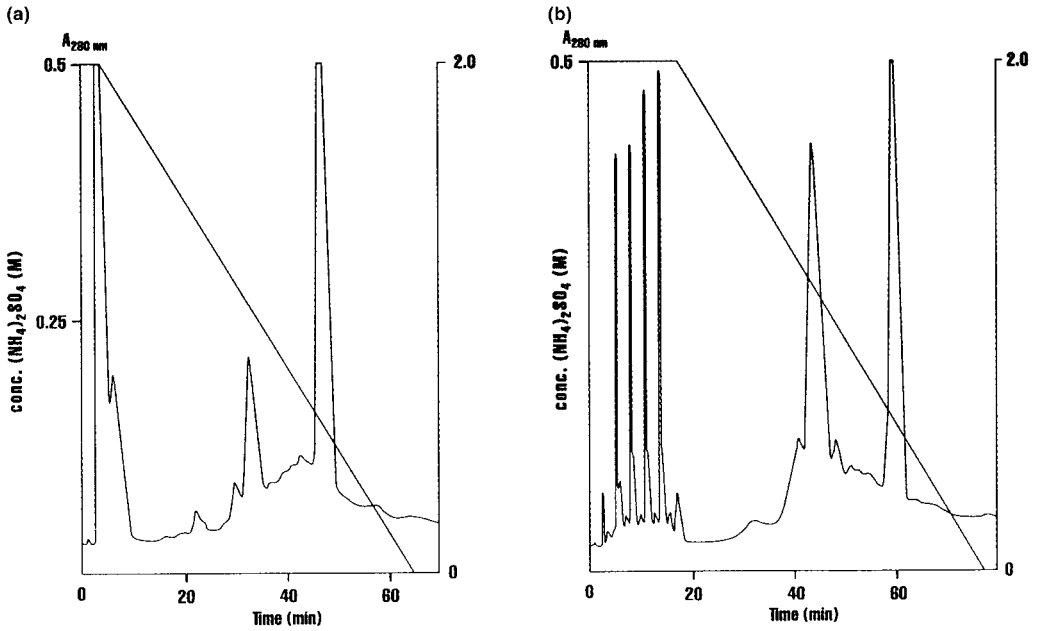


Figure 6. The effect of the suggested method of injecting the sample in diluted aliquots intermissioned by zones of concentrated buffer. In (**a**) 500 μL anti-CEA MAB (−IgG1) from mouse ascites fluid, in 0.9 M $(NH_4)_2SO_4$ (AS), was injected on a column equilibrated in 2 M AS in 0.1 M sodium phosphate, pH 7. In (**b**) the same sample was injected in 100 μL aliquots. 1.3 mL of 2 M AS were injected between every injection of sample. The capacity of the column was fully utilized without the risk of precipitating the sample in tubings and adapters. *Source*: Figure reprinted from ref. 17 by kind permission of Pharmacia Biotech, Uppsala, Sweden.

cause the binding is taking place near the elution concentration.

Equilibration of the Column and Sample Application

There is always a risk that the high salt concentration needed to attach the sample to the column will cause precipitation in tubings, pumps, columns, and detectors. This may be prevented by the following scheme: The column should be equilibrated with the proper buffer concentration but the sample with slightly lower salt concentration. This is easily accomplished by dilution of the sample. Instead of injecting the whole sample volume at one time, the injection should be done in aliquots divided by zones of the equilibration buffer. If the volume of the aliquots is kept less than 10% of the column volume and the volume of the intermediate zones with full buffer concentration exceeds 10% of the column volume, the sample condition will adapt to that of the equilibrated column. In this way, the precipitation in wrong places is avoided and the interaction with the column is still taking place in the correct environment (Fig. 6).

Scaling-Up

To get a similar result in large scale as in laboratory scale, it is important to keep some parameters constant. The column length, the linear flow, and the ratio between gradient and column volumes belong to this category. A constant column length leads to the fact that the desired capacity will be proportional to the cross-sectional area of the column and actually decide it, because the same proportion of the column length should be occupied by the sample in both scales.

Because of the relationship between the volumetric flow, the linear flow, and the cross-sectional area of the column on one side and the gradient-to-column-volume ratio on the other side, the gradient volume will be decided when the diameter of the column is set. In principle, a gradient elution in process scale will then take the same time as in laboratory scale. In practice, a decrease in resolution caused by a more ineffective solvent distribution system at the inlet of a process scale column will sometimes have to be compensated for by a lower flow.

APPLICATIONS

HIC has been used for the purification of serum proteins (41,42), nuclear proteins (43), membrane-bound proteins (44), and viruses and cells (45), and recombinant proteins (46) and insulin receptors (47) have been successfully prepared. Interesting applications include the use of HIC media as biochemical reactors by coupling enzymes by hydrophobic interaction (48) and a method for the exchange of detergents bound to membrane proteins (18).

For further applications and more discussion about large-scale processing using HIC, the works by Pharmacia Biotech and Eriksson (17,18) are recommended.

BIBLIOGRAPHY

1. C. Tanford, *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, Wiley, New York, 1973.
2. T.E. Crighton, *Proteins, Structures and Molecular properties*, W.H. Freeman, New York, 1984.
3. W. Melander, C. Horwath, *Arch. Biochem. Biophys.* **183**, 200–215 (1977).
4. R. Srinivasan, E. Ruckenstein, *Sep. Purif. Methods* **9**, 267–370 (1980).
5. J. Porath, L. Sundberg, N. Fornstedt, I. Olsson, *Nature* **245**, 465–466 (1973).
6. J. Porath, *J. Chromatogr.* **376**, 331–341 (1986).
7. S. Hjertén, *J. Chromatogr.* **87**, 325–331 (1973).
8. S. Lewin, *Displacement of Water and Its Control of Biochemical Reactions*, Academic Press, New York (1974).
9. A. Tiselius, *Ark. Kemi* **26B**, 1–5 (1948).
10. B.H.J. Hofstee, *Anal. Biochem.* **52**, 430–448 (1973).
11. S. Shaltiel, Z. Er-el, *Proc. Natl. Acad. Sci. USA* **70**, 778–781 (1973).
12. G. Halperin, M. Breitenbach, M. Tauber-Finkelstein, S. Shaltiel, *J. Chromatogr.* **215**, 211–228 (1981).
13. M. Wilchek, T. Miron, *Biochem. Biophys. Res. Commun.* **72**, 108–113 (1976).
14. S. Hjertén, J. Rosengren, and S. Påhlman, *J. Chromatogr.* **101**, 281–288 (1974).
15. S. Påhlman, J. Rosengren, and S. Hjertén, *J. Chromatogr.* **131**, 99–108 (1977).
16. J.-C. Jansson, T. Låås, in E. Roger ed., *Chromatography of Synthetic and Biological Macromolecules*, Ellis Horwood, Chichester, U.K., 1978.
17. Pharmacia Biotech, *Hydrophobic Interaction Chromatography. Principles and Methods*, Uppsala (1992).
18. K.-O. Eriksson, in J.-C. Jansson, L. Rydén ed., *Protein Purification: Principles, High Resolution Methods, and Applications*, VCH Publishers, Cambridge, U.K. 1989, pp. 207–226.
19. J.L. Fausnaugh, F.E. Reigner, *J. Chromatogr.* **359**, 131 (1986).
20. P. Kårsnäs, T. Lindblom, *J. Chromatogr.* **599**, 131–136 (1992).
21. S. Hjertén, K. Yao, K.-O. Eriksson, B. Johansson, *J. Chromatogr.* **359**, 99–109 (1974).
22. F. Maisano, M. Belew, J. Porath, *J. Chromatogr.* **321**, 305–317 (1985).
23. Y. Kato, T. Kitamura, T. Hashimoto, *J. Chromatogr.* **360**, 260 (1986).
24. J.L. Fausnaugh, E. Pfannkoch, S. Gupta, F.E. Reigner, *Anal. Biochem.* **137**, 464 (1984).
25. F. von der Haar, *Biochim. Biophys. Res. Commun.* **70**, 1009–1013 (1976).
26. K. Adachi, *Biochim. Biophys. Acta* **912**, 139 (1987).
27. J. Porath, *Nature* **196**, 47 (1962).
28. J. Carlsson, J.-C. Jansson, and M. Sparrman, in J.-C. Jansson, L. Rydén ed., *Protein Purification: Principles, High Resolution Methods, and Applications*, VCH Publishers, Cambridge U.K., 1989, pp. 275–329.
29. H.P. Jennissen, I.M.G. Heilmeyer, *Biochemistry* **14**, 754–760 (1975).
30. J. Rosengren, S. Påhlman, M. Glad, S. Hjertén, *Biochim. Biophys. Acta* **412**, 51–61 (1975).
31. B. Ekström, G. Jacobson, *Anal. Biochem* **142**, 134–139 (1984).
32. S. Oscarsson, P. Kårsnäs, *J. Chromatogr* **803**, 83–93 (1988).
33. W. Melander, Z. El Rassi, C. Horvath, *J. Chromatogr.* **496**, 3–27 (1989).
34. S. Miller, J. Janin, A.M. Lesk, C. Chotia, *J. Mol. Biol.* **196**, 641 (1987).
35. B. Lee, F.M. Richards, *J. Mol. Biol.* **55**, 379–400 (1971).
36. T. Arakawa, L.O. Narhi, *Biotechnol. Appl. Biochem.* **13**, 151–172 (1991).

37. V.A. Parsegian, B.W. Ninham, *Biophys. J.* **10**, 664–674 (1970).
38. H.P. Jennissen, *J. Chromatogr.* **159**, 71–83 (1978).
39. H.P. Jennissen, *J. Colloid. Interface Sci.* **111**, 570 (1986).
40. R.K. Scopes, *Protein Purification, Principles and Practice*, Springer, New York, 1994.
41. J.-C. Jansson, T. Låås, in E. Roger ed., *Chromatography of Synthetic and Biological Macromolecules*, Ellis Horwood, Chichester, U.K., 1978.
42. Z. Hrkal, J. Rejnkova, *J. Chromatogr.* **242**, 385–388 (1982).
43. D.E. Comings, A.G. Miguel, H.H. Lesser, *Biochim. Biophys. Acta* **563**, 253–260 (1979).
44. R.D. McNair, A.J. Kenny, *Biochem. J.* **179**, 379–395 (1979).
45. S. Hjertén, in D. Glick ed., *Methods of Biochemical Analysis*, Wiley, New York, 1981, pp. 89–108.
46. M. Belew, M. Yafang, L. Bin, J. Berglöf, J.-C. Jansson, *Bioseparation* **1**, 397–408.
47. L. Kuehn, H. Meyer, H. Reinauer, *Proc. 2nd. Int. Insulin Symp.* 243–250 (1980).
48. K.D. Caldwell, R. Axén, J. Porath, *Biotech. Bioeng.* **17**, 613–616 (1975).

CHROMATOGRAPHY, ION EXCHANGE

STEVEN M. CRAMER
VENKATESH NATARAJAN
Rensselaer Polytechnic Institute
Troy, New York

KEY WORDS

Ion exchange
Models
Nonlinear chromatography
Preparative chromatography

OUTLINE

INTRODUCTION

Preparative ion exchange chromatography is a technique for separating biomolecules based primarily on their electrostatic interactions with charged stationary phase materials. This technique is the most widely employed chromatographic separation technique in the biotechnology industry today. It is used for both capture and high-resolution separations. The last decade has seen dramatic advances in the state of the art of chromatographic bioprocessing with ion exchange materials. This article presents recent advances in resin materials, modes of operation, and modeling.

BACKGROUND

Ion exchange chromatography (IEC) is the most widely employed chromatographic separation technique in the biotechnology industry today. It is based on the principle of electrostatic interactions between charged surfaces. In IEC, one employs a charged stationary phase material to separate oppositely charged solutes. Cation exchange employs a negatively charged surface to separate positively charged solutes, and anion exchange employs a positively charged surface to separate negatively charged solutes. The binding of solutes to ion exchangers may be viewed as an exchange process in which the solutes compete for binding with salt counterions and other charged components in the mobile phase.

This article focus on preparative protein separations in ion exchange systems, beginning with a brief discussion of various factors that affect protein retention (e.g., pH, ionic strength, mobile phase modifiers). A detailed treatment of the state of the art of stationary phase materials is then presented. A description of different modes of preparative IEC is given, along with separation examples from the literature. Finally, the modeling of preparative ion exchange is addressed.

Factors Affecting the Ion Exchange of Proteins

pH of the Mobile Phase. Proteins are ampholytes, possessing both negative and positive surface charges. In fact, the zwitterionic nature of proteins is what makes ion exchange such a powerful tool for these separations. The charges on the proteins arise because of the presence of acidic and basic amino acid residues on the protein surface as well as the amino and carboxyl termini of the protein.

Thus, the net charge on a protein is a strong function of the pH of the solution. Under acidic conditions, the basic side chains are charged, under basic conditions, the acidic residues will be charged. At the isoelectric point (pI), the net charge on the protein will be zero. In general, when the pH of the mobile phase is greater than the pI of the protein, the protein will have a net negative charge and will be retained on an anion exchanger. On the other hand, when the pH of the mobile phase is less than the pI of the protein, the protein will have a net positive charge and will be retained on a cation exchanger. It turns out that the net charge of the protein is not the sole determinant of its retention (1). The distribution of charges on the surface of the protein is also important. The protein may have patches of positively or negatively charged residues that may interact with the stationary phase even when its net charge is zero. Furthermore, the microenvironment pH often can differ slightly in the resin and bulk solutions. Thus, a protein may be retained even at its pI. In addition, it is important to consider protein solubility constraints in the development of an ion exchange separation, since proteins are sparingly soluble near their isoelectric point. In practice, to prevent protein denaturation, the operating pH in IEC is limited to the range 2–9.

The choice of buffer to establish the operating pH is of paramount interest in ion-exchange chromatography. It is desirable to employ a buffering agent that has the same charge as the stationary phase material or is neutral. This precludes the binding of these ions onto the stationary phase and thus prevents pH shifts within the column. Hence, it is desirable to use cationic buffers (e.g., alkylamines, Tris, aminoethyl alcohol) in anion exchange chromatography and anionic buffers (e.g., phosphate, acetate, citrate, barbiturate) in cation exchange chromatography.

Ionic Strength. The chromatographic retention of proteins is modulated by the ionic strength of the mobile phase. In cation exchange systems, the relative strength of cations is as follows (2): $Ba^{2+} > Pb^{2+} > Sr^{2+} > Ca^{2+} > Ni^{2+} > Cd^{2+} > Cu^{2+} > Co^{2+} > Zn^{2+} > Mg^{2+} > Ti^{2+} > Ag^{+} > Cs^{+} > Rb^{+} > K^{+} > NH_4^{+} > Na^{+} > H^{+} > Li^{+}$. In anion exchange, the relative strength of anions is citrate > sulfate > oxalate > I^{-} > nitrate > CrO_4^{2-} > Br^{-} > SCN^{-} > Cl^{-} > formate > acetate > OH^{-}.

Thus, Br^{-} will bind stronger than Cl^{-} in an anion exchange column. In cation exchange, an equivalent solution of KCl would elute a protein faster than NaCl. Continuous and step gradients of ionic strength are often employed to effect protein separations. While protein retention is generally reduced as the ionic strength is increased, at very high ionic strength (>1 M), hydrophobic interactions can become pronounced.

Mobile Phase Modifiers. Addition of modifiers to the mobile phase can often influence the selectivity of a given separation. Recently, Gagnon *et al.* (3) reported an enhancement of the selectivity of proteins on the addition of poly(ethylene glycol) to the mobile phase. Other additives include chelating agents such as EDTA (to protect enzymes against metal impurities), reducing agents (e.g., dithiothreitol) to retard oxidation, and surfactants and chaotropic agents (e.g., urea) to solubilize membrane proteins (4).

Temperature. There are conflicting reports in the literature on the effect of temperature on resolution and band broadening. Increasing the temperature may reduce the viscosity of the mobile phase and increase the kinetics of the adsorption process. This may result in more efficient separations. However, changes in the temperature may also induce conformational changes in the protein. This may denature the protein or expose more hydrophobic patches to the stationary phase, enhancing nonspecific interactions and causing dramatic increases in protein retention (5). In general, the practice in industry is to employ either room temperature or low temperatures (4 °C) to minimize protease activity and to ensure protein stability.

STATIONARY-PHASE MATERIALS

The stationary phase is of paramount importance in IEC. Ideally, the stationary phase should be rigid enough to withstand high flow rates, provide sufficient selectivity and capacity for the problem at hand, and not denature the protein of interest. Furthermore, it should be stable over a wide range of operating conditions (pH, solvents, etc.) and amenable to rigorous cleaning-in-place with acid or alkali. The choice of the manufacturer is also important (6). The manufacturer should be able to ensure a regular supply with little lot-to-lot variation. Finally, the resin should not leach materials into the process stream.

Selectivity in ion exchange materials is affected by the ligand density, spacer arm chemistry, and the backbone resin material. Unfortunately, state-of-the-art chromatographic modeling cannot a priori predict the selectivity of different resin materials for various classes of separation problem. Thus, selection of an ion exchange resin is problem specific and requires the screening of a large variety of materials for a given separation. In bioprocessing applications, IEC is used for both capture and high-resolution separations. Capture applications demand that the resin possess sufficient dynamic capacity and be rigid enough to allow operations at high flow rates. On the other hand, selectivity is paramount for high-resolution separations.

Particle Size

While reversed-phase chromatography often employs materials of relatively small particle diameter for preparative applications, stationary-phase materials of 30- to 200-μm diameter are typically employed for preparative IEC. Higher resolutions can be achieved with smaller particles at the cost of increased backpressure. Finally, the size of the particle can also influence the mode of packing (7).

Ion Exchange Functional Groups

The stationary phase of an ion exchanger consists of a base matrix that is functionalized with basic (anion exchange) or acidic (cation exchange) groups. The charge on the stationary phase is pH dependent. This pH dependency allows one to categorize ion exchange stationary phases into

strong and weak exchangers. Strong ion exchangers retain their charge over a very wide range of pH, whereas weak ones do so over a narrower range of pH. Thus, the capacity of weak ion exchangers is a function of the pH. Typically, strong exchangers have functional groups with very low pK_a (sulfopropyl in cation exchange) or very high pK_a (quaternary amino–type anion exchangers). Table 1 lists some of the functional groups commonly employed in ion exchange stationary phases and their corresponding pK_a values (8,9).

Capacity and Porosity

The capacity of the ion exchanger can be measured as both the total number of charges per unit stationary-phase volume or the mass of protein adsorbed per unit volume. When one is determining protein-binding capacity, it is important to ensure that the activity of the adsorbed protein is not compromised. The capacity expressed in terms of adsorbed protein is the effective capacity of the column. It is dependent on both the specific protein and the operating conditions (pH, ionic strength). It is also dependent on the pore size and the porosity of the particles. If the particle is pellicular (i.e., not porous), the surface area is relatively low, resulting in reduced capacity. On the other hand, if the particle is highly porous, the capacity increases, but the resistance to mass transfer also increases. Pores that are too small may be inaccessible to the proteins. The pore size distribution is usually unavailable, and a nominal pore size is supplied by the manufacturer.

Conventional ion exchange media have nominal pore sizes between 100 and 1,500 Å. At times, the pore size is also specified as the molecular weight of the largest globular protein that can be accommodated in the material.

Novel and Traditional Matrices

Table 2 lists representative resins available in the market today along with their manufacturers and trade names. As seen in the table, the resins can be roughly divided into five categories. Polysaccharide gels have been in existence since the 1950s (10). These matrices are inert toward proteins, can be easily derivatized, have very high capacities, and are stable over a wide range of pH values. Polysaccharide gels based on weakly cross-linked agarose, dextran, or cellulose have weak mechanical properties and are limited to low-pressure regimes. In contrast, Sepharose® Fast Flow ion exchangers (Pharmacia) are composed of a higher degree of cross-linked agarose, which increases their mechanical stability. High-performance polysaccharide materials (Pharmacia) are also available in relatively small particle diameter (34 μm) and have been shown to be particularly useful for difficult high resolution separations. These polysaccharide-based resins are currently the most widely employed chromatographic materials in the biopharmaceutical industry.

Although silica-based ion exchangers have good mechanical properties, they are not as widely used for bioprocessing because their instability under alkaline conditions precludes alkaline cleaning. Furthermore, surface silanol and siloxane groups can have adverse effects on proteins (11). To overcome this problem, silica supports have been coated with a polymeric layer (12).

Resins based on porous particles of zirconium oxide have been synthesized and characterized for biochromatography (13). Zirconia possesses all the mechanical advantages offered by silica. In addition, it has good thermal and chemical stability. Unfortunately, this backbone material tends to bind proteins, which necessitates coating with a biocompatible layer. Clausen and Carr (14) recently employed a phosphonate analog of EDTA for the modification of zirconia supports to provide a biocompatible stationary phase. As described shortly (see "Dynamic Capacity"), a potential advantage of zirconium is its relatively high density, which has been shown to have potential for expanded-bed applications (15).

The polymer-based matrices (16) have good mechanical properties and are stable over a wide pH range. Furthermore, they have a relatively high loading capacity because of their macroporous structure. The hydrophobic polymer matrices are typically coated with a hydrophilic polymer to reduce nonspecific interactions. Polymeric supports are being used increasingly in the preparation of ion exchangers for bioprocessing (17). In addition, many of the novel materials described below are based on polymeric supports.

The penultimate category of particulate ion exchangers in Table 2 is entitled "Novel" media. Each of the stationary phases in this category possesses unique features that distinguish it from more traditional materials. The POROS® supports are synthesized by a suspension polymerization process. These chromatographic materials consist of a rigid polystyrene–divinylbenzene base matrix covered with a polyhydroxyl layer that can be further derivatized to produce various functional groups (18). This type of resin belongs to the class of particles termed "perfusive" (19,20). Such particles have a bimodal pore distribution containing macropores (6,000–8,000 Å) that transect the particle and interconnecting "diffusive" micropores (500–1,500 Å). The macropores facilitate mass transport, while the diffusive pores increase the available surface area. Their permeability has been shown to be independent of

Table 1. Functional Groups Used in IEC

Name	Abbreviation	pK_a
Anion exchangers		
Diethyl aminoethyl	DEAE	9–9.5
Quaternary aminoethyl	Q	9.5
Dimethyl aminoethyl	DMAE	9
Trimethyl aminoethyl	TMAE	>13
Cation exchangers		
Carboxymethyl	CM	3.5–4
Phosphate	Phospho	<2 & 6
Sulfonate	S	2
Sulfoethyl	SE	2
Sulfopropyl	SP, SO_3^-	<1

Source: Compiled from Refs. 8 and 9 and from manufacturers' data.

Table 2. Representative Commercial Ion Exchange Media

Trade name	Manufacturer	Function groups	Particle sizes (μm)	Matrix
		Polysaccharide gels		
Bio-Gel A	Bio-Rad	DEAE, CM	80–150	Cross-linked agarose
Sepharose CL-6B	Pharmacia LKB	DEAE, CM	45–165	Agarose
Sephacel	Pharmacia LKB	DEAE	40–160	Cross-linked cellulose
Express-Ion	Whatman	DEAE, Q, CM, SE	60–130	Microgranular cellulose
Sephadex	Pharmacia LKB	DEAE, Q CM, SP	40–125	Cross-linked dextran
Sepharose Fast Flow	Pharmacia LKB	Q, DEAE, CM, SP	45–165	Cross-linked agarose
Sepharose High Performance	Pharmacia LKB	Q, SP	24–44	Cross-linked agarose
Ultrogel	BioSepra	Hydroxy-apatite	60–180	Cross-linked agarose
		Inorganic matrix		
Accell	Waters	CM, QMA	40	Polymer-coated silica
TSKgel 3SW	TosoHaas	CM, DEAE	5, 10	Silica
Spherodex	BioSepra	DEAE, SP, CM	40–300	Silica–dextran
Spherosil	BioSepra	DEAE, QMA	40–300	Polymer-grafted silica
Zephyr	Sepracore	DEAE, S	20	Silica–dextran
		Synthetic polymers		
TSKgel 5PW	TosoHaas	SP, DEAE	30	Methacrylate copolymer
Mono Q/S	Pharmacia LKB	Q, S	10	Cross-linked polystyrene–divinylbenzene
Macro-Prep	Bio-Rad	S, Q, CM, DEAE	50	Methacrylate copolymer
Protein-Pak	Waters	DEAE, Q, SP, CM	8, 15, 40	Rigid polymeric gel
Trisacryl	BioSepra	Q, DEAE, CM, SP	40–160	Macroporous synthetic polymer
Trisacryl Plus	BioSepra	DEAE, SP	40–160	Base-resistant polymer
Toyopearl	TosoHaas	DEAE, CM, SP	35, 65, 100	Methacrylate
Super Q	TosoHaas	Q	35, 65, 100	Synthetic polymer
		Novel media		
POROS	Perseptive	Q, S	10, 20	Coated polystyerene
SOURCE	Pharmacia LKB	Q, S	15, 30	Cross-linked polystyrene–divinylbenzene
Hyper D	BioSepra	Q, S	10, 20, 35, 60	Polystyrene shell and hydrogel
Fractogel EMD (tentacle resins)	Merck	TMAE, DEAE, DMAE, SP, CM	20–40	Vinylic grafted polymer
STREAMLINE	Pharmacia	SP, DEAE	200	Macroporous cross-linked agarose with a quartz core
		Membrane ion exchangers		
MemSep	Millipore	DEAE, Q, CM		Cellulose
QuickDisk	Knauer	DEAE, Q		Poly(glycidylmethacrylate)
Sartobind	Sartorius	Q, S		Synthetic copolymer

the particle size (21). The solute breakthrough curves for the POROS media have been shown to be relatively unaffected by the flow rate in contrast to conventional ion exchangers (20). Gerstner et al. (22) have demonstrated that displacement chromatography of proteins can be carried out on these particles at very high linear velocities. The SOURCE™ materials have also been shown to exhibit chromatographic behavior that is independent of flow rate. In addition, the SOURCE™ resins have relatively high efficiency owing to their monodisperse particle size distribution. Recently, the development of agarose-based perfusive materials has also been reported (23).

Hyper-D® materials are composite stationary phases composed of a rigid macroporous support filled with a weakly cross-linked ion exchange gel. The rigid particle provides the necessary mechanical strength for high flow rate operations, while the soft gel provides high binding capacity. The equilibrium uptake of model proteins and the mass transfer rates in Q-Hyper-D® has been studied by Fernandez and Carta (24,25). These materials have a very high binding capacity even at elevated flow rates. In fact, it has recently been shown (26) that large-particle-diameter Hyper-D® materials (200 μm) also exhibit high binding capacities even at elevated linear velocities. It has been proposed that the mass transport mechanism in these materials is surface diffusion due to the high concentrations obtained in the gel (25).

The "tentacular" stationary phases comprise another novel sorbent (27). These chromatographic materials have polyelectrolyte chains (typically 5–50 monomers long)

grafted onto porous matrices. In contradistinction to the supports mentioned thus far, the charge moieties in tentacular phases are attached to the matrix via relatively long, flexible spacer arms. This results in reduced nonspecific interactions between the proteins and the base matrix. Furthermore, since the ligands are able to move relatively freely in space, they can explore the steric distribution of charges on a protein in a more efficient manner. This can potentially result in higher-resolution separations. Tentacular ion exchangers have been employed to differentiate between native and deamidated recombinant human DNAse which differ in a single surface charge (Fig. 1) (28).

Dynamic Capacity

The dynamic capacity of the column is the effective capacity determined from the breakthrough curve (i.e., when the effluent concentration reaches 10–50% of the inlet protein concentration). Table 3 lists the dynamic capacities of several anion exchange media obtained at low linear velocities; these capacities are seen to range from 30 to 165 mg/mL. The cellulosic anion exchangers such as Sepharose® Fast Flow have high dynamic capacities under these conditions. Importantly, dynamic capacities can be strongly affected by the flow rate. In Figure 2, which compares the flow rate dependence of the dynamic capacities of five Q exchangers, the dynamic capacities of the tentacular (Frac-

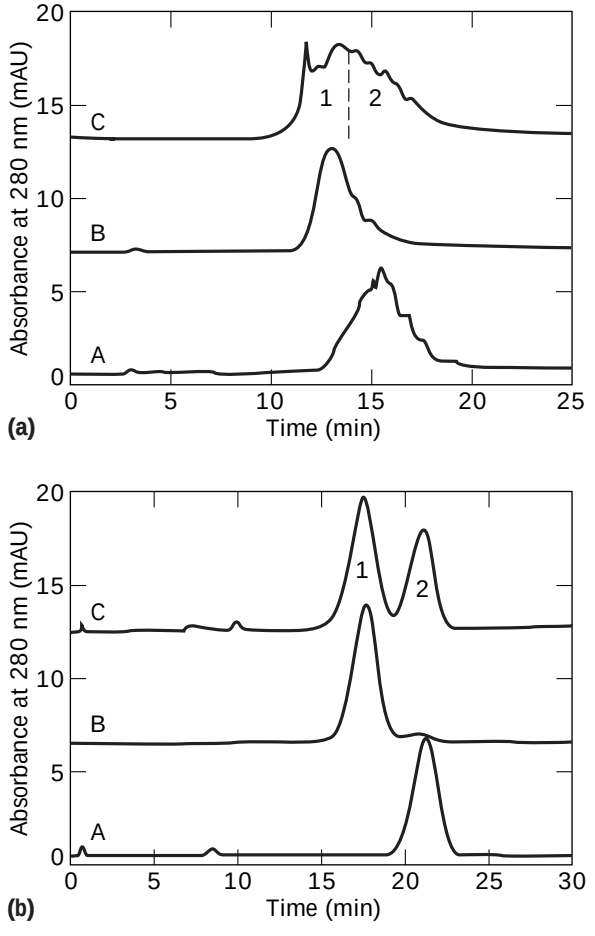


Figure 1. (**a**) Strong cation exchange chromatography on a sulfopropyl column (TSK SP 5PW, 75 mm × 7.5 mm i.d.) of three samples: A, rhDNAse; B, its deamidated variant; and C, a mixture of the two variants. (**b**) Cation exchange chromatography of three samples: A, rhDNAse; B, its deamidated variant; and C, a mixture of the two variants on a tentacle ion exchanger (LiChrosphere SO_3^- 50 mm × 4.6 mm i.d.). The two sets of chromatograms were obtained under the same chromatographic conditions. (*Source*: Reproduced with permission from Ref. 28.)

Table 3. Ionic Capacities of Commonly Used Anion Exchangers

Ion exchanger	BSA sorption capacity (mg/mL)[a]
QMA Accell	111
DEAE Bio Gel A	38
DEAE Fractogel 650	32
DEAE Sephacel	131
DEAE Sepharose CL 6B	164
DEAE Sepharose Fast Flow	88
DEAE Spherodex	98
DEAE Trisacryl	86
DEAE Fractogel EMD	100[b]
Q Fractogel EMD	55
Q POROS II	40
Q Sepharose Fast Flow	115
Q Super	127
Q Hyper D	125
Express-Ion Q	55

Source: From Ref. 8 and manufacturers' data.
[a]Determined for BSA in 50 mM Tris–HCl buffer (pH 8.6) except as noted.
[b]Determined for BSA in 50 mM Tris–HCl buffer (pH 8.3).

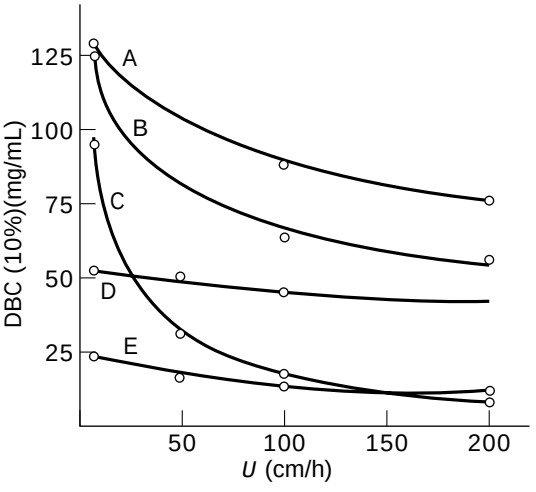


Figure 2. Variation of dynamic binding capacity (DBC) as a function of the flow rate (*U*) for Q anion exchangers. Determination of DBC were performed using breakthrough curves (10%) with a column of 1 cm i.d. containing 4 mL of sorbent. Protein solution was bovine serum albumin at 5 mg/mL in 50 mM Tris–HCl buffer (pH 8.6): A, Q HyperD F; B, Super Q; C, Q Sepharose Fast Flow; D, Q Fractogel EMD; E, Q POROS. *Source*: Reproduced with permission from Ref. 8.

togel EMD) and perfusive (POROS®) resins are relatively low at low flow rates, but their dependence on the flow rate is rather weak. On the other hand, the dynamic capacity of the Sepharose® FF resin exhibits a stronger dependence on the flow rate. The Hyper D® and the Super materials exhibit a moderate flow rate dependence. However, the Hyper D® material has a capacity substantially higher than those of the other materials at elevated flow rates. An important caveat to this discussion on dynamic capacity is that it is often desirable to use low operating pressures in a bioprocessing environment. This practice may affect the linear velocities that can be employed under actual operating conditions and may make performance at elevated velocities less important.

An important recent advance in the state of the art of chromatographic bioprocessing is the development of expanded-bed systems, in which the bed expands during the loading phase, enabling it to directly process unclarified feedstock. This can potentially obviate the need for membrane or centrifugal processing prior to the ion exchange capture step. A schematic of the basic principles of expanded bed adsorption is shown in Figure 3. The Streamline adsorbent, which has a mean particle size of approximately 200 μm, is a specially designed stationary phase for expanded-bed chromatography. It is stable over a wide pH range but unstable in solutions containing oxidizing agents. Accordingly, in situ sterilization may be a problem with these sorbents. Streamline forms a stable expanded bed without the need for external magnetic fields. Its capacities are lower than those of typical commercial adsorbents such as Sepharose FF and Spherodex (29,30). This may be due to the presence of the quartz core, which provides the necessary particle density but reduces the available surface area. The performance of this sorbent has been shown to be similar in packed-bed and expanded-bed modes (29), indicating that the expanded-bed mode behaves like a packed bed with increased voidage.

In addition to the commercially available Streamline material, several other expanded-bed supports are in development. Flickinger and Carr (15) are exploring the potential of zirconia stationary phases for expanded-bed applications. These materials have the desirable property of having sufficient density without further modification. Gilchrist et al. (30) synthesized a polysaccharide–ceramic composite sorbent for use in expanded-bed chromatography. Thommes et al. (31) derivatized Bioran (controlled pore glass) with sulfonic acid residues. These particles compared well with Streamline media in initial evaluations.

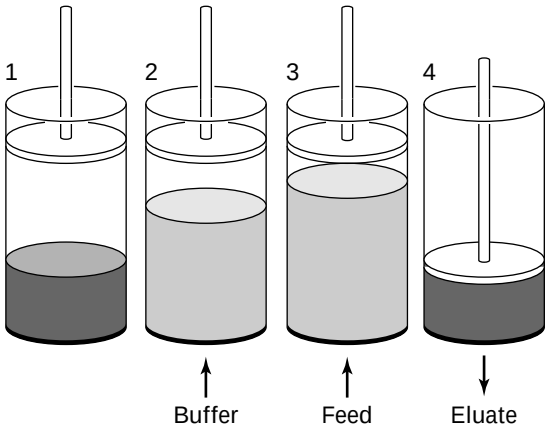


Figure 3. Principles of expanded-bed operation: **1**, sedimentation of adsorbent, **2**, upward flow of equilibration buffer resulting in a stable expanded bed, **3** application of unclarified feed by upward flow, and **4**, downward flow of the eluent to elute the bound protein of interest. *Source*: Adapted from Pharmacia catalog.

Membrane Adsorbers

In contrast to all the chromatographic supports discussed thus far, an alternative approach to ion exchange is the use of charged membrane systems. Mass transfer in membrane exchangers is typically dominated by convection, which enables these systems to be operated at relatively fast linear velocities resulting in high throughputs. The typical configuration of membrane ion exchangers is the "pancake" configuration, which has a large width-to-length aspect ratio that provides higher surface areas and lower backpressures. The principal disadvantage of these configurations is the difficulty of achieving a uniform inlet flow distribution and collection of the eluate with minimal backmixing.

Monoclonal antibodies and human recombinant antithrombin III were subjected to preparative purification on Sepharose® FF columns and Sartobind units (32). The membrane exchanger yielded higher throughputs (>13-fold in the case of monoclonal antibodies) and comparable or higher recoveries. However, higher concentration factors were achieved on the Sepharose columns. An extensive review of the state of the art of membrane chromatography can be found in Ref. 33.

Regeneration, Cleaning-in-Place, and Sanitization

A critical issue in ion exchange bioprocessing is the ability to clean the column. This includes column regeneration (between each separation cycle), cleaning-in-place operations (after every 1–10 cycles), and sanitization (between batches). Any commercial ion exchange material must be stable enough to withstand high pH as well as oxidizing conditions.

Column regeneration and cleaning-in-place operations are important for establishing long-term use of ion exchange columns. The conditions are usually determined empirically and are strongly affected by the feed composition. Typical conditions include high salt (1 M NaCl) and/or 1 M NaOH. A complete list of regeneration protocols can be found in Ref. 34. The monograph by Sofer and Nystrom (35) offers a more detailed discussion of column maintenance procedures.

In situ sanitization of the column is necessary to prevent microbial contamination between batches. Formalin, peracetic acid, and sodium hypochlorite are effective chemical disinfectants. Alkaline solutions of ethanol are effective for spore destruction. Clearly, the stability of the resin must be kept in mind when the sanitization conditions are chosen. A recommended validation protocol for the in situ sanitization of an ion exchanger is outlined in Ref. 34.

Relative Efficacy of Resin Materials

In process chromatography, it is important to distinguish between initial capture (recovery) and high-resolution or polishing steps. For product recovery, the capacity and production rate are paramount. In addition, it is often necessary to process large volumes of relatively dilute feedstock, which often contain particulates. It is precisely this challenge that has catalyzed the development of many of the "novel" materials described above. The advantage of perfusive-type materials (e.g., POROS®) and membrane chromatography is that they can be operated at high linear velocities. However, these materials often have relatively low capacities. While the problem can be overcome by carrying out many column cycles for a given batch, this approach has not been widely adopted by the industry. Composite stationary phases (e.g., Hyper-D®) possess high binding capacities and can be operated at reasonably high flow rates without a drastic loss of dynamic capacity or increased backpressure. This makes their use for capture operations potentially very attractive. Expanded-bed sorbents are particularly attractive for capture operations because they combine clarification, concentration, and capture into a single step. It turns out that the stability of expanded beds is quite sensitive to such factors as the solids content, pH, ionic strength, and viscosity of the feed (36). Since the effect of these factors has not been fully elucidated, the choice of the mobile phase conditions in expanded-bed applications will be highly problem specific. Nevertheless, expanded-bed chromatography has generated tremendous interest from the industry.

As mentioned in the beginning of this section, the effect of stationary phase chemistry on selectivity is still unpredictable, and the selection of an ion exchange resin is problem specific. Recently, Levison et al. (37) evaluated the functional and physical performance of a range of ion exchange media based on polysaccharide and polymeric or composite materials. Although the data suggest significant differences in the performance of various media for a common set of criteria, the selection of an ion exchanger remains problem specific. Thus, until the state of the art of chromatographic modeling is further improved, the biochromatographic engineer is left with the daunting task of screening a large variety of materials for a given separation. Fortunately, there are several automated chromatographic workstations that can readily screen a large number of resins in a relatively short period of time. Once the resin with the optimal selectivity has been identified, the task remains to develop an efficient, high-throughput chromatographic process. This requires the selection of the appropriate mode of operation.

MODES OF COLUMN OPERATION

Preparative ion exchange can be carried out by means of a variety of resin materials, column configurations, and modes of operation. While the selection and development of resin materials often receives significant attention, the choice of the proper mode of operation is often given less attention. In fact, the selection of the optimal mode of operation for a given bioseparation problem can have a profound effect on the production rate, yield, and purity for that separation. To date, linear gradient or step gradient modes have been used to carry out most industrial chromatographic bioprocessing. This section briefly describes these gradient modes and presents examples from the recent literature. In addition, the displacement and chromatofocusing modes of operation are discussed.

Step Gradient

Step gradient is often employed for initial capture or recovery operations where the primary objective is to recover and to concentrate the bioproduct of interest. This operation is usually carried out in short, large-diameter columns packed with relatively large-particle diameter stationary-phase materials (>90 μm). This mode of chromatography consists of first loading the column with a large volume of low ionic strength feed solution, followed by a wash step, and finally an elution step based on an elevated salt and/or modified pH eluent. This type of chromatography can be characterized as essentially an on/off operation. The most important stationary-phase characteristic in preparative step gradient chromatography is the dynamic capacity of the resin. Thus, these separations do not require very efficient columns and can be readily carried out in short "pancake"-type columns. These capture steps can be carried out by means of conventional packed-bed columns or with "expanded-bed"-type configurations. As described earlier, it is important to screen a wide variety of ion exchange resins to identify a material that has sufficient selectivity and capacity for a given separation. Several examples of step gradient ion exchange chromatography are now presented.

The use of an expanded bed in the pilot-scale recovery of the anticoagulant recombinant Annexin V from unclarified *E. coli* homogenate has been reported (36). The optimal binding and elution conditions for Annexin V were determined in a 1.6 cm × 20 cm packed bed of Streamline DEAE adsorbent using clarified homogenate. Then 1.7 L of unclarified homogenate was applied to the laboratory-scale Streamline 50 column (5 cm × 100 cm, Streamline DEAE adsorbent) under the mobile phase conditions established with the small packed bed. The protein was eluted using a 250-mM step gradient of NaCl. The elution was carried out in the sedimented mode at a mobile phase velocity of 100 cm/h. The results were very similar to those obtained with the small packed bed. The presence of cell debris did not seem to affect the process. The biomass dry weight and the viscosity of the homogenate were 3.6% and 10 mPa/s, respectively.

Finally, the process was scaled up to a pilot-scale expanded bed: Streamline 200 (20 cm × 95 cm, Streamline DEAE adsorbent). The adsorbent volume, feed volume, and flow rate were scaled up by a factor of 16 to account for the increased diameter of the column, resulting in a feed load of 26.5 L. The rest of the conditions were kept the same. Figure 4 shows the chromatogram from the pilot-scale recovery of the protein. Yields greater than 95% were obtained with both the Streamline 50 and Streamline 200 columns. The expanded bed was cleaned after each run using 0.5 M NaOH with 1 M NaCl followed by distilled

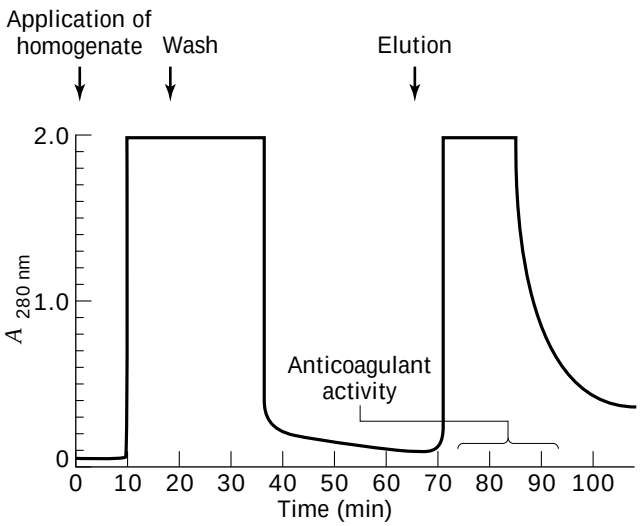


Figure 4. Chromatogram from pilot-scale, expanded-bed recovery of recombinant Annexin V from unclarified *E. coli* homogenate: feed, 26.5 L of unclarified *E. coli* homogenate in 30 mM ammonium acetate (pH 5.5); column, Streamline 200 with Streamline DEAE adsorbent; elution with 30 mM ammonium acetate (pH 5.5) containing 250 mM NaCl; flow rate: upward flow at 300 cm/h during application and wash and downward flow at 100 cm/h during elution. *Source*: Reproduced with permission from Ref. 36.

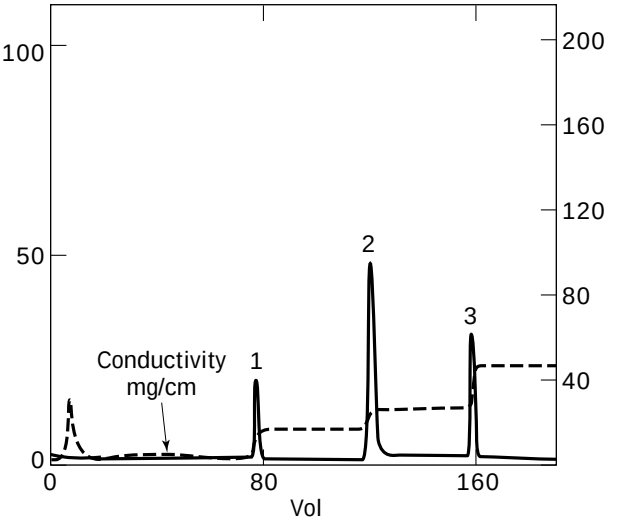


Figure 5. Purification of *Pseudomonas aeruginosa* exotoxin A on Q Sepharose HP (10 cm × 10 cm i.d.): buffer, 20-mM Tris, 1-mM EDTA (pH 7.4). Gradient program, first step salt concentration, 0.15 M NaCl (5 column volumes); second step salt concentration, 0.25 M NaCl (5 column volumes); final step salt concentration, 0.5 M NaCl (5 column volumes). *Source*: Reproduced with permission from Ref. 38.

water, 30% isopropanol, and 25% acetic acid. This example illustrates the potential of expanded-bed chromatography for ion exchange step gradient chromatography.

The use of step gradient in the production and purification of *Pseudomonas aeruginosa* exotoxin A (PE) from *E. coli* has been reported (38). The initial recovery of this protein from the fermentation broth involved cell lysis and centrifugation. Purification was achieved in three steps: weak anion exchange followed by hydrophobic interaction and a strong anion exchange step. The weak anion exchange step was carried out by means of a DEAE Sepharose® Fast Flow column in the flowthrough mode to remove DNA and pyrogens. In flowthrough column operations the impurities are retained on the column, but the product is not. Thus, this type of operation is quite simple to operate and can result in a rapid reduction of impurities from the feedstock. This flowthrough step resulted in a 10-fold reduction in DNA and a protein purity of approximately 25%. In fact, it is quite common to add a flowthrough anion exchange step in a process for DNA clearance.

The final anion exchange operation was carried out on a Q Sepharose high-performance (10 cm × 10 cm) column using multiple step gradients (Fig. 5). The protein of interest was eluted with the second step, resulting in the purification of 0.3 g of protein at a purity of greater than 99%. Since the column loading was relatively low, multiple cycles were necessary to produce sufficient amount of material. The ion exchange columns were cleaned and regenerated between runs with 1 M NaOH and 3 M NaCl. This example illustrates how step gradient chromatography can sometimes be employed for high-resolution separations for systems exhibiting sufficient separation factors. However, as shown in this example, these operations are often limited by relatively low column loading.

Step gradient has many advantages, namely, ease of operation, no need for complicated gradient pumping apparatus, and concentration of the bioproduct. When the separation problem is more difficult, however, a higher-resolution mode of operation is often required.

Linear Gradient

For high-resolution separations, linear gradient chromatography is often employed. This operation is usually carried out in large-diameter columns using a range of column lengths and particle diameters. It turns out that depending on the difficulty of the separation, it may be necessary to employ longer columns and/or materials of smaller particle diameter for these separations. This mode of chromatography is typically carried out by first loading the column with a relatively large volume of low-ionic-strength feed solution, followed by a continuous linear gradient of increasing ionic strength. This type of chromatography cannot be characterized as an on/off operation and is thus dependent on the length of the column and the column efficiency. The most important stationary-phase attribute in preparative linear gradient chromatography is the selectivity of the resin for the given separation. Accordingly, it is again essential that a wide variety of resin materials be screened to identify the material with the highest selectivity. In addition to the selectivity, the dynamic capacity of the stationary phase is also important for preparative linear gradient separations and must be considered when selecting a resin. The major advantage of linear gradient is its ability to resolve difficult separation problems. Several examples of linear gradient ion exchange chromatography are now presented.

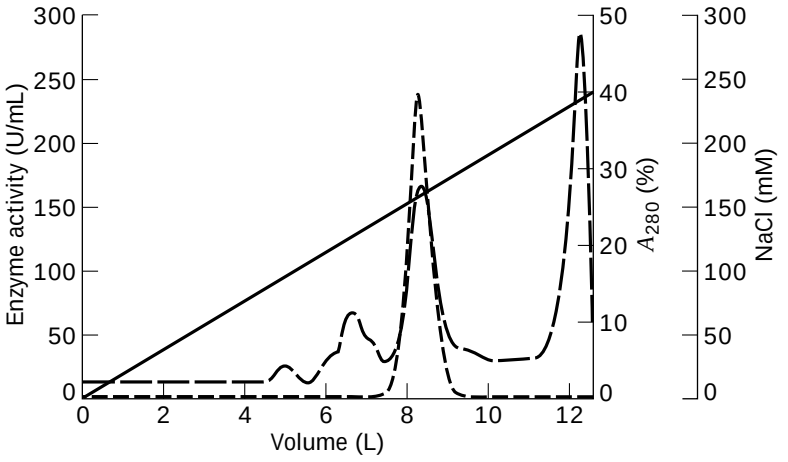


Figure 6. Cation exchange of L-asparaginase on a CM-Sepharose column (24 cm × 11.3 cm i.d.): buffer, 5 mM succinic acid–NaOH (pH 6); feed, dilute solution of cell extract with a conductivity of less than 2 mS; flow rate, 24 cm/h; gradient, 0–300 mM NaCl in 16 L of buffer. Curves: solid, NaCl; dotted, enzyme, dashed, protein. *Source*: Reproduced with permission from Ref. 39.

The use of linear gradients in the purification of L-asparaginase from *Vibrio (Wolinella) succinogenes* has been reported (39). The purification of this enzyme was achieved through a cation exchange step followed by an anion exchange polishing step. Approximately 13 g of protein (184,000 units) was loaded onto a 11.3 cm × 24 cm CM Sepharose cation exchanger at a mobile phase velocity of 24 cm/h. Upon elution of the enzyme by means of a linear gradient of 0–300 mM NaCl, the chromatogram shown in Figure 6 was obtained. This gradient resulted in the recovery of 91% of the enzyme and a 28-fold purification. An anion exchange polishing step followed. Approximately 0.4 g of protein (167,000 units) was loaded onto a 5 cm × 10.2 cm Q Sepharose® FF column at a mobile phase velocity of 15 cm/h. A linear gradient of 0–200 mM NaCl was used to elute the enzyme. This separation scheme resulted in an 87% recovery and a 47-fold purification of the enzyme. These separations required quite low mobile-phase velocities. In addition, the high-resolution step was limited to a column loading of 2g/L, which is typical of high-resolution linear gradient separations.

Often, the purification scheme for a protein involves both step and linear gradient operations. The purification of factor VIII (FVIII) and von Willebrand factor (vWF) from human plasma (40) illustrates the use of both step and linear gradients in the purification protocol. The FVIII–vWF complex is a complicated system in which the two proteins are highly glycosylated, have low pIs, and exhibit strong binding even under conditions of relatively high ionic strength. Furthermore, FVIII is prone to proteolytic degradation without its carrier protein vWF. The naturally adhesive vWF exhibits strong nonspecific interactions with organic, hydrophobic supports.

A wide variety of stationary-phase materials were evaluated for this separation. Silica-based supports could not be used in the purification scheme because they tended to activate FVIII. Supports such as cellulose and poly(styrene–divinylbenzene) were unable to effect the separation owing to the strong nonspecific interactions of vWF. Columns packed with traditional media such as Toyopearl® DEAE, MonoQ, MonoP, and TSK® DEAE 5PW were subject to fouling. On the other hand, tentacular supports gave good results for this separation, possibly because their extended tentacular arms prevent the protein from interacting with the base matrix.

The purification scheme for this complex involved three anion exchange steps and two virus inactivation steps. All the anion exchange steps involved the use of tentacular supports. Weak Fractogel® EMD-DEAE supports were used in the step gradient mode for the first two steps, and strong Fractogel® TMAE in the linear gradient mode was employed for the final purification. The FVIII–vWF complex was strongly retained on the DEAE support, enabling the use of step gradient chromatography to remove a variety of contaminating proteins, solvents, detergents, and stabilizers (employed during viral inactivation steps).

The FVIII–vWF complex was dissociated and separated in the final step by means of a linear gradient. Calcium ions were added to the buffer to cause the dissociation of the complex. Figure 7, which illustrates the separation achieved with the strong tentacular exchanger, indicates that this difficult separation was complicated by the heterogeneity of the protein glycosylation and incomplete dissociation. This example illustrates how linear gradient ion exchange with highly selective resins can be employed for

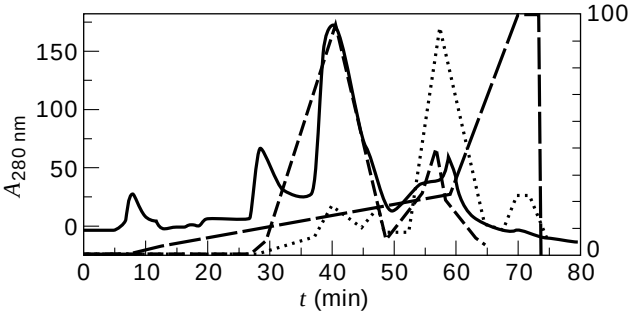


Figure 7. Dissociation and separation of FVIII and vWF from their complex in the presence of 50 mM calcium: buffer, 20 mM Tris–HCl (pH 7.4): feed, 300 units of FVIII in buffer; column, Fractogel® EMD TMAE (20–40 μm, 120 mm × 16 mm i.d.); flow rate, 3 mL/min; gradient, 0–200 mM NaCl. Curves: short dashes, vWF activity; dots, FVIII activity; long dashes, % B (IM NaCl) *Source*: Reproduced with permission from Ref. 40.

very difficult separations. This process resulted in a purification factor of 8,000 with an overall yield of approximately 20%.

According to the conventional wisdom, the more difficult the separation, the lower the column loading that can generally be processed by linear gradient. Thus, linear gradient is often limited to lower column loadings. In addition, linear gradient requires more complicated gradient pumping apparatus. In contrast, displacement chromatography is able to carry out high-resolution separations at very high column loadings, using the simple pumping systems employed for step gradient chromatography.

Displacement

The displacement mode of chromatography, first identified in the 1940s by Tiselius (41), combines the high throughput and concentrating power of step gradient chromatography with the high resolving power of linear gradient chromatography. Although displacement chromatography has been shown to hold significant promise for high-resolution preparative separations, a major impediment to the implementation of displacement chromatography has been the lack of suitable displacer compounds. Recent advances in this field have generated significant interest in displacement technology.

Operationally, displacement chromatography is performed in a manner similar to step gradient chromatography, where the column is subjected to sequential step changes in the inlet conditions. The column is initially equilibrated with a carrier buffer of relatively low ionic strength. The feed mixture is then introduced into the column, followed by a constant infusion of the displacer solution. The displacer is selected such that it has a higher affinity for the stationary phase than any of the feed components. Under appropriate conditions, the displacer induces the feed components to develop into adjacent "square-wave" zones of highly concentrated pure material. After the breakthrough of the displacer from the column effluent, the column is regenerated and reequilibrated with the carrier buffer.

The separation of oligonucleotides from their failure sequences presents a significant challenge to chromatographic engineers. The failure sequences are typically very closely related to the main product. Furthermore, there is considerable conformational heterogeneity of both the product and the failure sequences. Displacement chromatography has been successfully employed for resolving oligonucleotide mixtures characterized by low separation factors. For example, 1.2 g of a 24-mer phosphorothioate oligonucleotide was purified on a POROS® HQ/M column using dextran sulfate as the displacer (42) (Fig. 8). The separation yielded 70% of the product at a purity of about 96%. Methanol was added to the mobile phase to eliminate hydrophobic interactions between the oligonucleotide and the stationary phase and to improve the recovery. The column was regenerated with 3 M NaCl in 0.5 M NaOH. On a semipreparative scale, the entire separation (equilibration + loading + displacement + regeneration) could be accomplished in approximately 32 mins.

An important recent advance in the state of the art of displacement chromatography has been the discovery that low molecular weight displacers can be successfully employed for protein purification in ion exchange systems (43). Pentaerythritol-based dendritic polyelectrolytes ranging in molecular weight from 480 to 5,100 Da have been shown to be effective displacers for protein purification (44). In addition, protected amino acids (45), antibiotics (46), and a variety of anionic displacers have been developed (47–49). Low molecular weight displacers have been successfully employed for very difficult protein separations (50). These novel displacers have significant operational advantages over large polyelectrolyte displacers. First and foremost, if there is any overlap between the displacer and the protein of interest, the low molecular weight materials can be readily separated from the purified protein during postdisplacement downstream processing involving size-based purification methods (e.g., size exclusion chromatography, ultrafiltration).

The relatively low cost of synthesizing low molecular weight displacers can be expected to significantly improve the economics of displacement chromatography. Furthermore, the salt-dependent adsorption behavior of these low molecular weight displacers greatly facilitates column regeneration. In addition, these displacers can be employed in "selective" displacement chromatography, where the product(s) of interest can be selectively displaced while the low-affinity impurities can be desorbed in the induced salt gradient ahead of the displacement train and the high-affinity impurities either retained or desorbed in the displacer zone (51).

Figure 9 shows an example of a selective displacement separation: a displacement experiment was carried out using a feed mixture of ribonuclease B, bovine heart cytochrome *c*, horse heart cytochrome *c*, and lysozyme with *N*-α-benzoylarginine ethyl ester (BAEE) as the displacer. As seen in the figure, the low-affinity protein ribonuclease B eluted ahead of the displacement train in the induced salt gradient, whereas the two cytochrome *c*'s were well resolved in the displacement train. Thus, in spite of the extreme closeness in their adsorption behavior, displacement chromatography resulted in excellent separation of the two cytochrome-*c* proteins with both high yields and high purity. The highest-affinity protein, lysozyme, remained on the column under the conditions of this experiment. This experiment demonstrates that selective displacement chromatography can result in baseline separation of the product of interest from the low- and the high-affinity impurities, while resolving the target bioproduct from very similar affinity impurities in an isotachic displacement train. Recent results have indicated that displacement chromatography may have significant utility for extremely difficult bioseparations such as protein variants (52) and glycoforms.

Chromatofocusing

In chromatofocusing retained, internal pH gradients are employed to effect the separation of protein mixtures. These pH gradients are produced internally by running a buffer of a particular pH through an ion exchanger equilibrated at a different pH. The focusing effects produced by these pH gradients were originally studied by Sluyterman

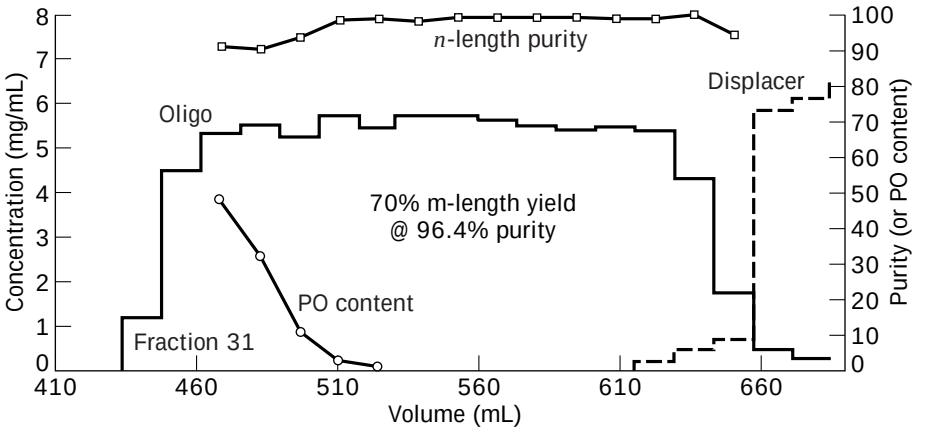


Figure 8. Preparative-scale purification of a 24-mer phosphorothioate oligonucleotide: feed, 1.2 g in 100 mM NaCl, 10 mM NaOH, 5% methanol; displacer, 800 mL of 7 mg/mL dextran sulfate; column, POROS HQ/M (20 μm, 100 mm $\times$ 50 mm i.d.); flow rate, 50 mL/min; fraction size, 14 mL. *Source*: Reproduced with permission from Ref. 42.

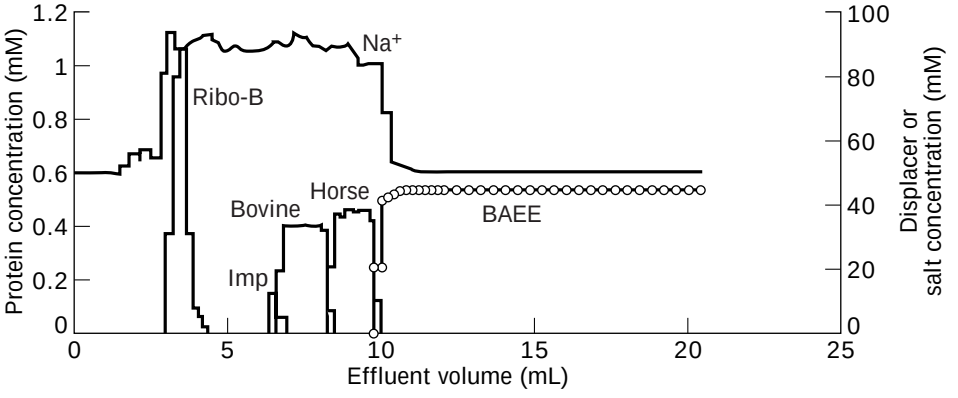


Figure 9. Displacement separation of a four-component protein mixture using BAEE as a displacer-column, 100 mm $\times$ 5 mm i.d. strong cation exchanger (8 μm); buffer, 50 mM sodium phosphate (pH 6.0); feed, 1.6 mL of 0.38 mM ribonuclease B, 0.45 mM horse heart cytochrome *c*, 0.37 mM bovine heart cytochrome *c*, and 0.34 mM lysozyme in buffer; displacer concentration, 45 mM; flow rate, 0.2 mL/min; fraction size, 200 μL. *Source*: Reproduced with permission from Ref. 51.

and coworkers (53,54). Typically, polyampholyte buffers are employed in this mode of chromatography. These buffers are not only expensive but also contain components that can associate with proteins (55). However, Hearn and Lyttle (56) and Frey et al. (57) demonstrated the efficacy of employing multicomponent mixtures of common buffer systems to produce internal pH gradients in ion exchange columns. Recently, Strong and Frey (58) demonstrated the chromatofocusing of model proteins under trace- and volume-overloaded feed conditions using mixtures of common buffers. The ability to carry out chromatofocusing with inexpensive, commonly employed buffers offers new opportunities for this high-resolution technology.

MODELS

Methods development for chromatographic bioprocessing typically involves an evaluation of the linear chromatographic behavior followed by empirical scale-up. Mathematical models are not widely employed for methods development and optimization of chromatography in the biopharmaceutical industry. Thus a significant amount of time is needed for methods development, as well as for the implementation of nonoptimal processes. As a result, there is increasing interest in the use of mathematical models for the development of efficient preparative ion exchange processes.

The classical view of protein retention in ion exchange columns involves the stoichiometric exchange of proteins and mobile-phase counterions (1,59–61). This is referred to as the stoichiometric displacement model (SDM). In this model, the retention of proteins on ion exchange stationary phases is described as follows (for a 1:1 salt, e.g., NaCl)

$$C_{pb} + \nu Q_m \Leftrightarrow Q_p + \nu C_m + \nu C_b \qquad (1)$$

where p is the protein molecule, ν is the ratio of the characteristic charge of the protein to the valence of the counterion, b is the co-ion (the ion with the same charge as the stationary phase) associated with the protein, m is the mobile-phase counterion (ion with the opposite charge of the stationary phase), C denotes the concentrations in the mobile phase, and Q denotes the concentrations on the stationary phase. The characteristic charge of the protein is a statistical average of the electrostatic interactions of the protein with the stationary phase as it travels down the column.

Some of the counterions remain bound to the protein on adsorption, while the rest make way for the interaction of the protein with the stationary phase. Drager and Regnier (60) argued that the equilibrium between the protein and the co-ion can be neglected. This results in the simplified equilibrium reaction

$$C_p + \nu Q_m \Leftrightarrow \nu C_m + Q_b \quad (2)$$

The SDM predicts that the retention of the protein under isocratic, linear conditions (low solute concentrations) is related to the counterion concentration as follows:

$$\log(k) = -\nu \log(C_m) + \text{const} \quad (3)$$

where k is the retention factor under isocratic linear conditions, and C_m is the concentration of the counterion in the mobile phase. This relationship has been experimentally demonstrated (60,62,63) under conditions usually employed in the IEC of proteins. Equation 3 forms the basis for the quantification of the effect of the ionic strength on the selectivity of linear separations.

Linear Chromatography

In linear chromatography, the various solutes migrate down the column independent of each other. The retention is characterized by a linear isotherm and an absence of intersolute competition for the adsorption sites. The plate model (64) is used extensively to characterize the efficiencies and resolutions obtained in these systems. In the plate model, all the nonidealities of the column are lumped into one parameter: N, the number of plates [or, alternatively, in height equivalent to one theoretical plate, HETP = (length of column)/N]. The HETP is widely used for the comparison of different sorbents and the validation of packing procedures. It is usually determined from the effluent profiles of isocratic pulse injections of a tracer. For a gaussian effluent peak, the HETP can be determined from the following equation

$$\text{HETP} = \left(\frac{w_h}{t_r}\right)^2 \frac{L}{5.545} \quad (4)$$

where w_h is the peak width at half-height, t_r is the retention time of the peak, and L is the length of the column. In ion exchange systems, the HETP measurement is affected by the salt in the background. It is recommended that pulse injections be made in a relatively high salt background to reduce Donnan exclusion (65).

The HETP is also a function of the flow rate. For conventional, macroporous sorbents, the plot of HETP versus flow rate (termed a Van Deemter or Knox plot) exhibits a minimum, indicating the presence of an optimum flow rate for operation. The HETP can be viewed as a sum of all the mass transfer, kinetic, and diffusional resistances in the column (66,67).

Yamamoto and coworkers have employed a continuous-flow plate model to characterize and scale up linear gradient and step gradient operations (68). Graphical techniques have been developed for scaling up linear gradient elution in linear chromatography. In this approach, the ionic strength at which the peak will elute in a large column is obtained directly from a plot of GH (gradient slope times the stationary phase volume) versus I_r (ionic strength at the elution time of the peak), which is established with the use of a relatively small column (69). The dimensionless retention time and the peak width in the large column can then be determined from analytical expressions (68). These methods have been applied to the scale-up of β-galactosidase purification in a 30-L column based on linear gradient data obtained with a 23-mL column (69).

Although the foregoing approach appears to work well for linear gradient scale-up of single solutes and multicomponent separations with baseline resolution, an alternative method (70) is required for general multicomponent linear separations. The resolution R_s is related to the operating conditions and the column dimensions as follows (70):

$$R_s \propto \left(\frac{L}{G\ \text{HETP}}\right)^{1/2} = \left(\frac{D_m I_a L}{GHud_p^2}\right)^{1/2} \quad (5)$$

where D_m is the molecular diffusivity, u is the mobile-phase velocity, d_p is the particle size, L is the length of the column, G is the gradient slope times the column dead volume, and I_a is a constant, introduced to dedimensionalize the right-hand side.

The HETP in equation 5 is the HETP in linear gradient elution. An algorithm to estimate the HETP from a set of linear gradient experiments at various flow rates has been developed (71). Thus, for scale-up,

$$\left(\frac{L}{G\ \text{HETP}}\right)_{\text{small column}} = \left(\frac{L}{G\ \text{HETP}}\right)_{\text{large column}} \quad (6)$$

While these approaches have proved useful for the scale-up of linear chromatography, the production rate in these ion exchange systems can be dramatically enhanced by operating these columns under nonlinear conditions. To model various modes of operation under nonlinear (high loading) conditions, an alternative approach must be employed.

Nonlinear Chromatography

To study the effects of nonlinear chromatography, one needs a suitable isotherm that accurately describes nonlinear protein adsorption. Guiochon et al. (72) have employed the Langmuir isotherm extensively to study the nonlinear chromatographic behavior of both small molecules and macromolecules. Berninger et al. (73) have developed the VERSE-LC model, which solves the complete mass transfer equations in conjunction with the Langmuir or modulator-modified Langmuir isotherms to describe nonlinear chromatography. While the Langmuir isotherm has many advantages, it is inadequate for describing protein adsorption in ion exchange systems. Protein adsorption is multipointed in nature (74) and is affected by the presence of mobile-phase modifiers (75–77). The Langmuir isotherm is unable to accurately describe these features.

The multivalent ion exchange model (75,78) incorporates both the multipointedness of the adsorption process and the effect of the mobile phase modifier on the process. It is derived from the SDM in conjunction with the law of mass action. Thus the multivalent ion exchange formalism predicts an isotherm of the form

$$Q_p = K_{eq} C_p \left(\frac{\bar{Q}_1}{C_1}\right)^{\nu_p} \quad (7)$$

where $\bar{Q}_1$ is the number of sites available for the adsorption of proteins, C_1 is the mobile-phase concentration of the counterions, K_{eq} is the equilibrium constant for the exchange process, and the subscript p denotes the protein.

The effect of the size of the protein molecules and lateral interactions on the adsorption process can be considerable (76,79–81). Furthermore, the adsorption of a protein molecule will result in the steric shielding of a number of stationary phase sites underneath the molecule. Brooks and Cramer (76) have defined a steric factor, σ, to account for such steric hindrance and have used the postulate that the steric hindrance will be proportional to the concentration of proteins on the surface (79). A steric mass action (SMA) model was developed that essentially employs equation 7 in concert with the following electroneutrality condition for single component equilibrium (76):

$$\Lambda = \bar{Q}_1 + (\nu_p + \sigma_p) Q_p \quad (8)$$

The parameters for the SMA model (K_{eq}, ν, and σ) can be readily obtained using standard chromatographic techniques. The equilibrium constant K_{eq} and the characteristic charge (for a monovalent counterion) ν are obtained from either isocratic or linear gradient elution chromatography (62). The steric factor can be obtained from nonlinear frontal chromatography. Knowledge of the bed capacity and the protein parameters allows the generation of nonlinear isotherms at any salt concentration.

The SMA formalism can be extended to multicomponent mixtures. In this case, the electroneutrality condition becomes

$$\Lambda = \bar{Q}_1 + \sum_{i=2}^{N} (\nu_i + \sigma_i) K_{eq,i} \left(\frac{\bar{Q}_1}{C_1}\right)^{\nu_i} C_i \quad (9)$$

where N is the total number of components (including the mobile phase counterion) in the system. Equations 7 and 9 then describe multicomponent nonlinear protein equilibria at various salt concentrations.

The SMA model accurately predicts a variety of nonlinear ion exchange behavior in different modes of chromatography. Two distinct approaches have been employed to date. The first involves the development of graphical techniques based on the method of characteristics for the prediction of single solute peaks in nonlinear isocratic and gradient chromatography (77,82). The second approach has employed the SMA formalism in concert with appropriate mass transport equations to describe multicomponent separations. As seen in Figures 10 and 11, the SMA model can accurately predict separations in multistep gradient and linear gradient separations, respectively. In addition, the model can predict sample displacement effects, which can have a profound effect on linear gradient separations under high loading conditions (Fig. 11). This model has also been shown to accurately predict the behavior of complex displacement systems (83). Furthermore, the SMA formalism has been used along with iterative optimization techniques to develop optimal step (84) and linear gradient (85) separations. Finally, the SMA model has been employed to describe the dynamic affinity of low molecular weight displacers in displacement systems, resulting in a simple approach to methods development in displacement chromatography (86).

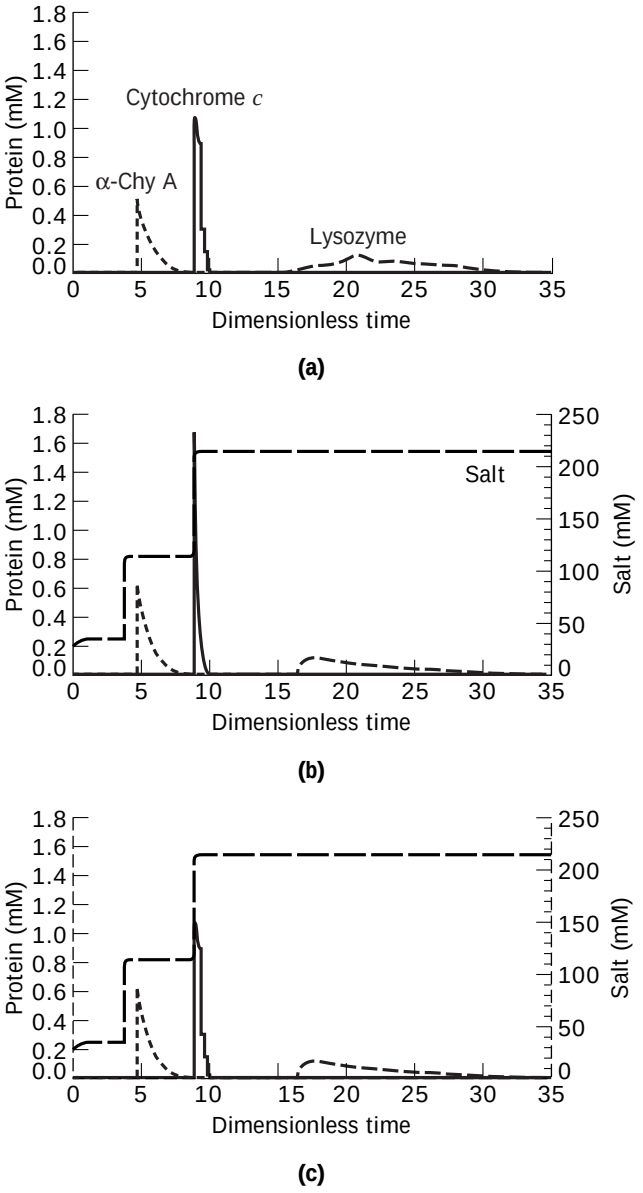


Figure 10. Comparison of step gradient chromatograms: (**a**) experimental, (**b**) simulated (using SMA), and (**c**) simulated (using SMA with fraction collection. Feed 2.8 column dead volumes of 0.19 mM α-chymotrypsinogen A, 0.19 mM cytochrome *c*, and 0.21 mM lysozyme in 30 mM sodium phosphate (pH 6). Gradient program: first step salt concentration, 112 mM, second step salt concentration, 213 mM; delay of 5.1 column dead volumes; flow rate, 0.4 mL/min; column, Waters SP (8 μm, 50 mm $\times$ 5 mm i.d.); fraction size, 147 μL (before 11 column dead volumes) and 413 μL (after 11 column dead volumes). *Source*: Reproduced with permission from Ref. 84.

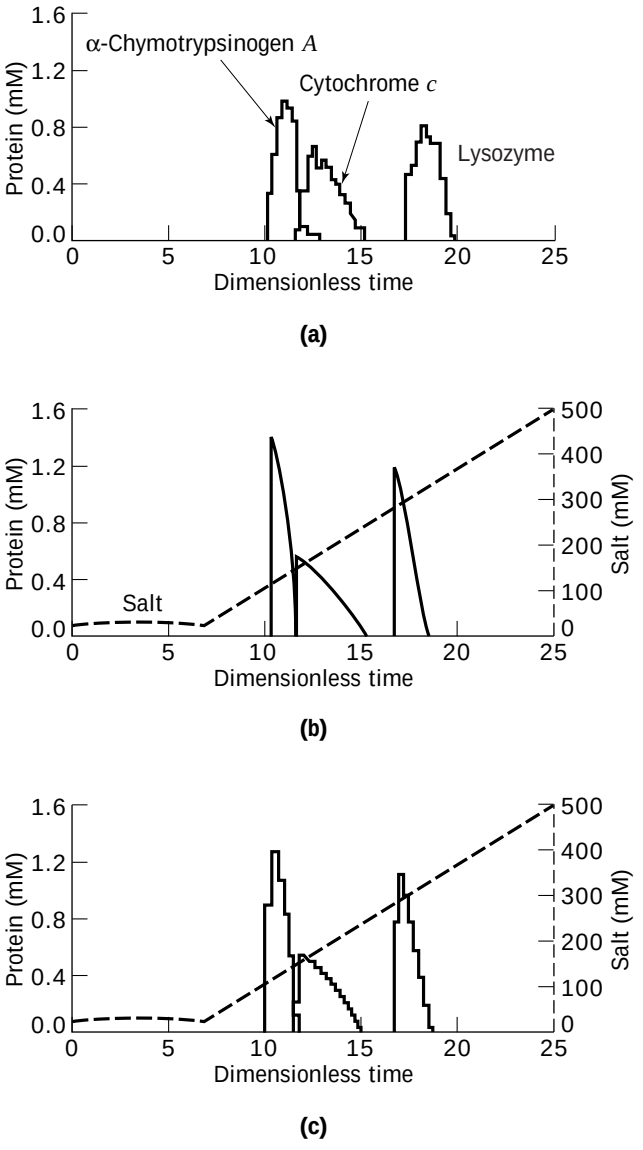


Figure 11. Comparison of linear gradient chromatograms: (**a**) experimental, (**b**) simulated (using SMA), and (**c**) simulated (using SMA with fraction collection). Feed, 6 column dead volumes of 0.2 mM α-chymotrypsinogen A, 0.2 mM cytochrome *c* and 0.2 mM lysozyme in 30 mM sodium phosphate (pH 6); gradient slope, 25 mM Na^+ per column dead volume; flow rate, 0.5 mL/min; column: strong cation exchanger (15 μm, 54 mm $\times$ 5 mm i.d.); fraction size, 200 μL. *Source*: Reproduced with permission from Ref. 85.

Nonstoichiometric Models

Though the SMA model is capable of accurately describing the behavior of nonlinear ion exchange chromatographic systems, it assumes that the individual charges on the protein molecule interact with the discrete charges on the ion exchange surface. In reality, the retention in ion exchange is more complex and is primarily due to the interaction of the *electrostatic fields* of the protein molecules and the chromatographic surface. Haggerty and Lenhoff (87) demonstrated a correlation between the retention of various proteins with their surface potential. Stahlberg and coworkers have developed the electrostatic interaction model (EIM), which treats the retention process as a coulombic interaction between two charged flat plates (88,89). Roth and Lenhoff (90) employ van der Waals forces in concert with electrostatic interactions to quantify the adsorption of proteins onto charged stationary phases. Recently, a mechanistic model has been developed that incorporates stationary-phase properties such as the surface charge density and short-range interaction energies to predict retention on various stationary phases (91). The drawback with the nonstoichiometric approaches reported thus far is that they are limited to describing retention in the linear mode of chromatography.

CONCLUSIONS

The last decade has seen dramatic advances in the state of the art of chromatographic bioprocessing with ion exchange materials. This article has presented recent advances in resin materials, modes of operation, and mathematical modeling. The need to process large volumes of relatively dilute feedstock has catalyzed the development of an array of "novel" chromatographic materials. While many novel materials have been developed, the effect of stationary-phase chemistry on selectivity is still essentially unknown, and thus the selection of an ion exchange resin requires significant column screening. Once a resin with sufficient selectivity has been identified, the task remains to develop an efficient, high-throughput chromatographic process. This requires the selection of an appropriate mode of operation. In fact, the selection of the optimal mode of operation for a given bioseparation problem can have a profound effect on the production rate, yield, and purity for that ion exchange step. Finally, the development of mathematical models capable of predicting scale-up and nonlinear protein chromatography now make it possible to employ modeling for bioprocess methods development.

NOMENCLATURE

C	mobile phase concentration (mM)
d_p	Particle size (cm)
D_m	Molecular diffusivity (cm^2/s)
G	Gradient slope times column dead volume (mM)
H	Column phase ratio (dimensionless)
I_r	Ionic strength at peak elution
k	Retention factor (dimensionless)
K_{eq}	Equilibrium constant (dimensionless)
L	Length of column (cm)
Q	Stationary phase concentration (mM)
$\bar{Q}_1$	Concentration of sites available for protein adsorption (mM)
R_s	Resolution (dimensionless)
t_r	Retention time (min)
u	linear velocity (cm/min)
w_h	Peak width at half-height (min)

Subscripts

i	solute index ($i = 1$ for salt counterion)
p	Protein
m	mobile-phase counterion

Greek

β	Phase ratio (dimensionless)
ν	Characteristic charge (dimensionless)
σ	Steric factor (dimensionless)
Λ	Bed capacity (mM)

BIBLIOGRAPHY

1. W. Kopaciewicz, M.A. Rounds, J. Fausnaugh, and F.E. Regnier, *J. Chromatogr.* **266**, 3–21 (1983).
2. Pharmacia, *FPLC Ion Exchange and Chromatofocusing*, Pharmacia, Uppsala, Sweden, 1985.
3. P. Gagnon, B. Godfrey, and D. Ladd, *J. Chromatogr. A* **743**, 51–55 (1996).
4. Dj. Josic, W. Hofmann, and W. Reutter, *J. Chromatogr.* **371**, 43–54 (1986).
5. F. Fang, M-I. Aguilar, and M.T.W. Hearn, *J. Chromatogr. A* **729**, 49–66 (1996).
6. W.F. Prouty, in Cs. Horvath and L.S. Ettre eds. *Chromatography in Biotechnology*, American Chemical Society, Washington D.C., 1993, pp. 43–58.
7. R. Henry and M.V. Piserchio, in C.T. Mant and R.S. Hodges eds., *High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis, and Conformation*, CRC Press, London, 1991, pp. 537–549.
8. E. Boschetti, *J. Chromatogr. A* **658**, 207–236 (1994).
9. E. Karlsson, L. Ryden, and J. Brewer, in J.-C. Janson and L. Ryden eds. *Protein Purification Principles, High Resolution Methods, and Applications*, VCH, New York, 1989, pp. 107–148.
10. H.A. Sober and E.A. Peterson, *J. Am. Chem. Soc.* **76**, 1711–1712 (1954).
11. S.H. Chang, R. Noel, and F.E. Regnier, *Anal. Chem.* **48**, 1839–1845 (1976).
12. A.J. Alpert and F.E. Regnier, *J. Chromatogr.* 185, 375–392 (1979).
13. P.W. Carr, J.A. Blackwell, T.P. Weber, W.A. Schafer, and M.P. Rigney, in Cs. Horvath and L.S. Ettre eds., *Chromatography in Biotechnology*, American Chemical Society, Washington D.C., 1993, pp. 146–164.
14. A.M. Clausen and P.W. Carr, *Anal. Chem.* **70**, 378–385 (1998).
15. M.C. Flickinger and P.W. Carr, presentation at Recovery of Biological Products VII, San Diego, CA 1994.
16. M.A. Rounds, W.D. Rounds, and F.E. Regnier, *J. Chromatogr.* **397**, 25–38 (1987).
17. C.K. Ratnayeke and F.E. Regnier, *J. Chromatogr. A* **743**, 15–23 (1996).
18. J.A. Gerstner, T. Londo, T. Hunt, J. Morris, P. Pedroso, and R. Hamilton, *CHEMTECH* **25**, 27–32 (1995).
19. N.B. Afeyan, N.F. Gordon, I. Mazsaroff, L. Varady, S.P. Fulton, Y.B. Yang, and F.E. Regnier, *J. Chromatogr.* **519**, 1–29 (1990).
20. N.B. Afeyan, S.P. Fulton, and F.E. Regnier, *J. Chromatogr.* **544**, 267–279 (1991).
21. J.F. Pfeiffer, J.C. Chen, and J.T. Hsu, *AIChE J.* **42**, 932–939 (1996).
22. J.A. Gerstner, J. Morris, T. Hunt, R. Hamilton, and N.B. Afeyan, *J. Chromatogr. A* **695**, 195–204 (1995).
23. P-E. Gustavsson and P-O. Larsson, *J. Chromatogr. A* **734**, 231–240 (1996).
24. M.A. Fernandez and G. Carta, *J. Chromatogr. A* **746**, 169–183 (1996).
25. M.A. Fernandez and G. Carta, *J. Chromatogr. A* **746**, 185–198 (1996).
26. J. Coffman, presentation at ISPPP 96, Boston, Mass., 1996.
27. W. Muller, *J. Chromatogr.* **510**, 133–140 (1990).
28. J. Cacia, C.P. Quan, M. Vasser, M.B. Sliwkowski, and J. Frenz, *J. Chromatogr.* **634**, 229–239 (1993).
29. Y.K. Chang and H.A. Chase, in D.L. Pyle ed. *Separations for Biotechnology*, Vol. 3, Royal Society of Chemistry, Cambridge, 1994, pp. 106–112.
30. G.R. Gilchrist, M.T. Burns, and A. Lyddiatt, in D.L. Pyle ed. *Separations for Biotechnology*, Vol. 3, Royal Society of Chemistry, Cambridge, 1994, pp. 186–192.
31. J. Thommes, M. Halfar, S. Lenz, and M-R. Kula, *Biotechnol. Bioeng.* **45**, 205–211 (1995).
32. D. Lutkemeyer, M. Bretschneider, H. Buntemeyer, and J. Lehmann, *J. Chromatogr.* 639, 57–66, (1993).
33. J. Thommes and M.-R. Kula, *Biotechnol. Prog.* **11**, 357–367 (1995).
34. A. Jungbauer and E. Boschetti, *J. Chromatogr. B* **662**, 143–179 (1994).
35. G.K. Sofer and L.-E. Nystrom, *Process Chromatography: A Guide to Validation*, Academic Press, London, 1991, pp. 5–30.
36. A.-K. Barnfield Frej, R. Hjorth, and A. Hammarstrom, *Biotechnol. Bioeng.* **44**, 922–929 (1994).
37. P.R. Levison, C. Mumford, M. Streater, A. Brandt-Nielsen, N.D. Pathirana, and S.E. Badger, *J. Chromatogr. A* **760**, 151–158 (1997).
38. A. Tsai, M. Gallo, T. Petterson, and J. Shiloach, *Bioprocess Eng.* **12**, 115–118, (1995).
39. R.J. Hinton, R.F. Sherwood, and J.R. Ramsay, in D.L. Pyle ed. *Separations for Biotechnology*, vol. 3, Royal Society of Chemistry, Cambridge, 1994, pp. 207–213.
40. Dj. Josic, H. Schwinn, M. Stadler, and A. Strancar, *J. Chromatogr. B* **662**, 181–190 (1994).
41. A. Tiselius, *Ark. Kemi Mineral. Geol.* **16A**, 1 (1943).
42. J.A. Gerstner, P. Pedroso, J. Morris, and B.J. Bergot, *Nucleic Acids Res.* **23**, 2292–2299 (1995).
43. U.S. Pat. 5,478,924 (December 26, 1995), S.M. Cramer, J.A. Moore, A. Kundu, Y. Li, and G. Jayaraman (to S.M. Cramer).
44. G. Jayaraman, Y.F. Li, J.A. Moore, and S.M. Cramer, *J. Chromatogr. A* **702**, 143–155 (1995).
45. A. Kundu, S. Vunnum, G. Jayaraman, and S.M. Cramer, *Biotechnol. Bioeng.* **48**, 452–460 (1995).
46. A. Kundu, S. Vunnum, and S.M. Cramer, *J. Chromatogr. A* **707**, 57–67 (1995).
47. S.D. Gadam and S.M. Cramer, *Chromatographia* **39**, 409–418 (1994).
48. J.A. Gerstner and S.M. Cramer, *Biopharm* **5**, 42–45 (1992).
49. S.C.D. Jen and N.G. Pinto, *J. Chromatogr.* 519, 87–98, (1990).
50. A. Kundu and S.M. Cramer, *Anal. Biochem.* **248**, 111–116 (1997).
51. A. Kundu, K.A. Barnthouse, and S.M. Cramer, *Biotechnol. Bioeng.* **56**, 119–129 (1997).

52. K.A. Barnthouse, W. Trompeter, R. Jones, P. Inampudi, R. Rupp, and S.M. Cramer, *J. Biotechnol.* (in press).
53. L.A. Sluyterman and O. Elgersma, *J. Chromatogr.* **150**, 17 (1978).
54. L.A. Sluyterman and J. Wijdenes, *J. Chromatogr.* **150**, 31 (1978).
55. J.H. Scott, K.L. Kelner, and H.P. Pollard, *Anal. Biochem.* **149**, 163 (1985).
56. M.T.W. Hearn and D.J. Lyttle, *J. Chromatogr.* **218**, 483–495 (1981).
57. D.D. Frey, A. Barnes, and J. Strong, *AIChE J.* **41**, 1171–1183 (1995).
58. J.C. Strong and D.D. Frey, *J. Chromatogr. A* **769**, 129–143 (1997).
59. N.K. Boardman and S.M. Partridge, *Biochem. J.* **59**, 543–552 (1955).
60. R.R. Drager and F.E. Regnier, *J. Chromatogr.* **359**, 147–155 (1986).
61. M.A. Rounds and F.E. Regnier, *J. Chromatogr.* **283**, 37–45 (1984).
62. S.D. Gadam, G. Jayaraman, and S.M. Cramer, *J. Chromatogr.* **630**, 37–52 (1993).
63. M.T.W. Hearn, in C.T. Mant and R.S. Hodges eds., *High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis, and Conformation*, CRC Press, London, 1991, pp. 105–122.
64. A.J.P. Martin and R.L.M. Synge, *Biochem. J.* **35**, 1358–1368 (1941).
65. H.P. Lettner, O. Kaltenbrunner, and A. Jungbauer, *J. Chromatogr. Sci.* **33**, 451–457 (1995).
66. J.J. van Deemter, F.J. Zuiderweg, and A. Klinkenberg, *Chem. Eng. Sci.* **5**, 271–289 (1956).
67. Cs. Horvath and H.J. Lin, *J. Chromatogr.* **149**, 43–70 (1978).
68. S. Yamamoto, K. Nakanishi, R. Matsuno, and T. Kamikubo, *Biotechnol. Bioeng.* **25**, 1465–1483 (1983).
69. S. Yamamoto, M. Nomura, and Y. Sano, *AIChE J.* **33**, 1426–1434 (1987).
70. S. Yamamoto, M. Nomura, and Y. Sano, *J. Chromatogr.* **409**, 101–110 (1987).
71. S. Yamamoto, *Biotechnol. Bioeng.* **48**, 444–451 (1995).
72. G. Guiochon, S.G. Shirazi, and A.M. Katti, *Fundamentals of Preparative and Nonlinear Chromatography*, Academic, New York, 1994, pp. 245–286.
73. J.A. Berninger, R.D. Whitley, X. Zhang and N.-H.L. Wang, *Comput. Chem. Eng.* **15**, 749–768 (1991).
74. H.P. Jennissen, *Biochemistry* **15**, 5683–5692 (1976).
75. A. Velayudhan and Cs. Horvath, *J. Chromatogr.* **443**, 13–29 (1988).
76. C.A. Brooks and S.M. Cramer, *AIChE J.* **38**, 1969–1978 (1992).
77. S.R. Gallant, A. Kundu, and S.M. Cramer, *J. Chromatogr. A* **702**, 125–142 (1995).
78. P. Cysewski, A. Jaulmes, R. Lemque, B. Sebille, C. Vidal-Madjar, and G. Jilge, *J. Chromatogr.* **548**, 61–79 (1991).
79. A. Velayudhan, *Studies in Nonlinear Chromatography*, Ph.D. dissertation, Yale University, New Haven, Conn., 1990.
80. Y.-L. Li, and N.G. Pinto, *J. Chromatogr. A* **658**, 445–457 (1994).
81. R.D. Whitley, R. Wachter, F. Liu, and N.-H.L. Wang, *J. Chromatogr.* **465**, 137–156 (1989).
82. S.R. Gallant, S. Vunnum, and S.M. Cramer, *AIChE J.* **42**, 2511–2520 (1996).
83. S.D. Gadam, S.R. Gallant, and S.M. Cramer, *AIChE J.* **41**, 1676–1686 (1995).
84. S.R. Gallant, A. Kundu, and S.M. Cramer, *Biotechnol. Bioeng.* **47**, 355–372 (1995).
85. S.R. Gallant, S. Vunnum, and S.M. Cramer, *J. Chromatogr. A* **725**, 295–314 (1996).
86. C.A. Brooks and S.M. Cramer, *Chem. Eng. Sci.* **51**, 3847–3860 (1996).
87. L. Haggerty and A.M. Lenhoff, *J. Phys. Chem.* **95**, 1472–1477 (1991).
88. J. Stahlberg, B. Jonsson, and Cs. Horvath, *Anal. Chem.* **63**, 1867–1874 (1991).
89. J. Stahlberg, B. Jonsson, and Cs. Horvath, *Anal. Chem.* **64**, 3118–3124 (1992).
90. C. M. Roth and A.M. Lenhoff, *Langmuir* **9**, 962–972 (1993).
91. C.M. Roth, K.K. Unger, and A.M. Lenhoff, *J. Chromatogr. A* **726**, 45–56 (1996).

See also ADSORBENTS, INORGANIC; ADSORPTION, EXPANDED BED; ADSORPTION, PROTEIN, BATCH; ADSORPTION, PROTEINS WITH SYNTHETIC MATERIALS; CHROMATOGRAPHY, SIZE EXCLUSION; MEMBRANE CHROMATOGRAPHY.

CHROMATOGRAPHY, RADIAL FLOW

TINGYUE GU
Ohio University
Athens, Ohio

KEY WORDS

Chromatography
Protein
Purification
Radial
Scale-up
Separation

OUTLINE

INTRODUCTION

The production of a modern biotechnology product, typically a recombinant protein, requires a multistage downstream process. Such a process usually centers on two or more liquid chromatography steps. As production scales escalate, the columns used become larger and more expensive. It is not uncommon to have bed volumes reaching hundreds of liters at industrial scales.

Radial flow columns were first used for gas–solid catalytic reactions in large packed beds. They were designed to increase gas flow rate and reduce pressure by increasing the cross-sectional flow area. Radial flow chromatography (RFC) columns first entered the commercial biotechnology market in the mid-1980s (1). RFC was intended as an alternative to the conventional axial flow chromatography (AFC) for preparative- and large-scale applications. In a radial flow chromatography column (Fig. 1), the mobile phase flows in the radial direction, not in the axial direction. The mobile enters from the outside tube and merges into the center tube (Fig. 2). In comparison to a slim AFC column, an RFC column provides a relatively larger flow area and a shorter flow path. It allows a higher volumetric flow rate with a lower bed pressure. The effect is equivalent to using a short pancakelike AFC column (Fig. 3). Pancakelike AFC columns are quite common in industrial applications. They are available from most major commercial column vendors. In scale-up to accommodate large feed loads, it is impractical to increase the column height of AFC columns by too much because excessive bed pressure drop will result. This is especially troublesome when soft gels are used. Thus, pancakelike AFC columns are used.

This article is devoted to the discussion of the applications of RFC columns in bioseparations and RFC modeling and scale-up issues. Applications examples will be provided based on existing literature. A general rate model will be presented for the modeling of RFC. Experimental and theoretical comparisons between AFC and RFC will be discussed.

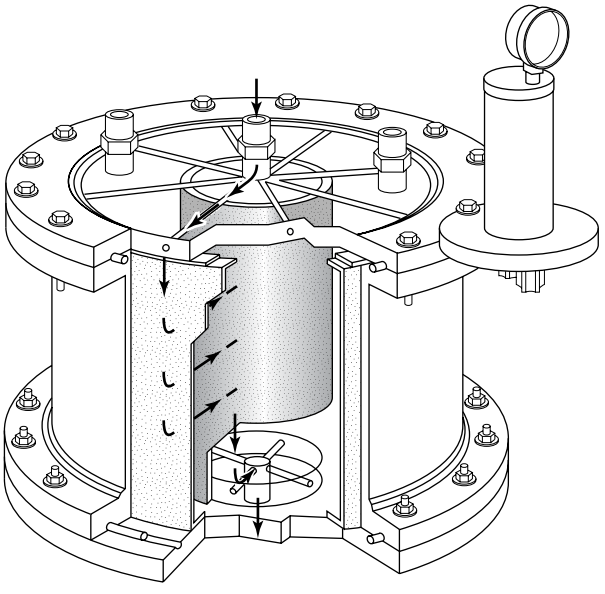


Figure 1. Diagram of an acrylic Superflo® column. *Source*: Courtesy of Sepragen Corp.

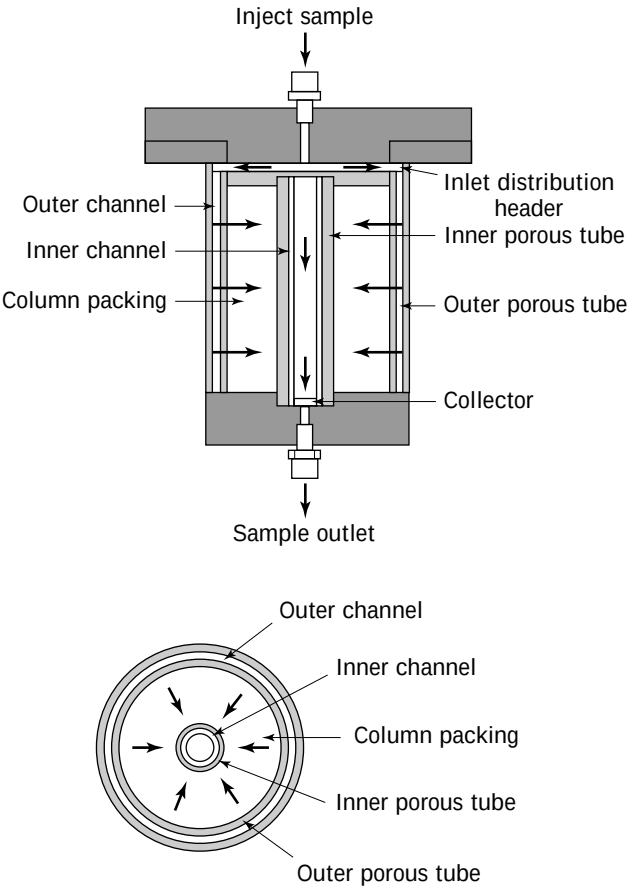


Figure 2. Anatomy of a Superflo® column. *Source*: Courtesy of Sepragen Corp.

RADIAL FLOW COLUMN CONFIGURATIONS

In 1947, Hopf (2) described a radial chromatography device. The device had a feed pipe in the center. The outward liquid flow in the radial direction was forced by the centrifugal force when the device was rotated. He called this device a chromatofuge. Such a device was obviously too complex and expensive for large-scale industrial application, and thus it was not adopted in biotechnology.

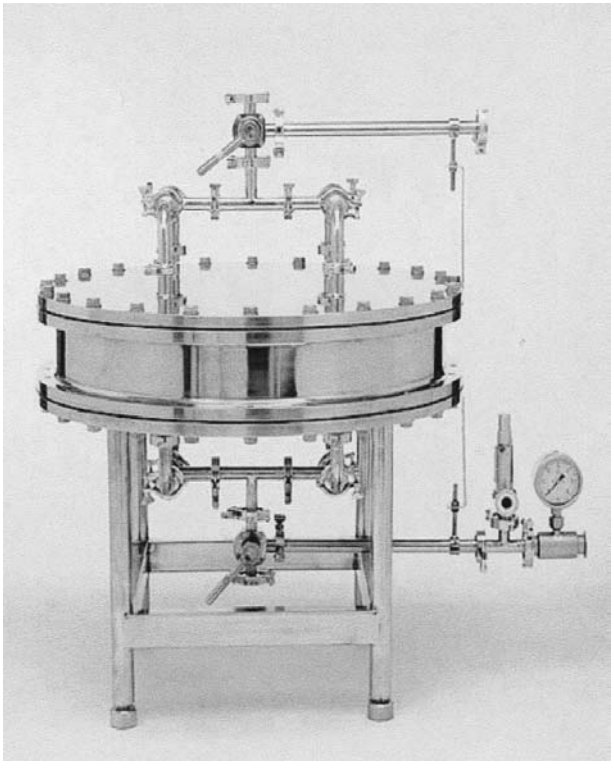


Figure 3. A large-scale pancakelike axial flow column. Courtesy of Pharmacia Biotech, Inc.

To date, two commercial companies have marketed RFC columns. The first one is CUNO, Inc., in Connecticut. They marketed the Zetaffinity series of preparative-scale radial flow cartridges, which looked like a spiral-wound filtration cartridge (Fig. 4). Such a design obviously was rooted in the fact the CUNO is a major manufacturer of industrial filtration systems. The packing for CUNO's RFC cartridges

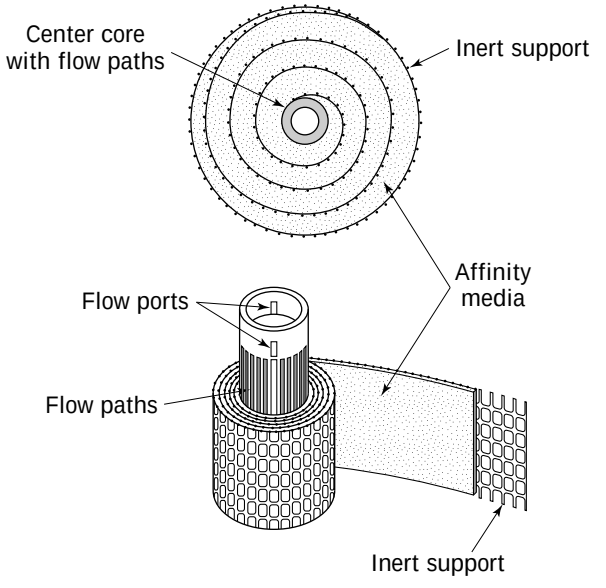


Figure 4. Structure of a Zetaffinity® cartridge.

was based on modified cellulose (3). These cartridges were used for affinity chromatography, hence the trade name Zetaffinity. CUNO discontinued its Zetaffinity product line in 1991.

The second vendor and manufacturer of RFC columns is Sepragen Corporation in Hayward, California. Their RFC product line carries the trade name Superflo®. Sepragen markets unpacked RFC columns ranging from 50 mL to 200 L. Figure 5 shows several Superflo columns with a similar diameter but different column heights. Superflo columns come in stainless steel, acrylic, polycarbonate, and polyethylene. They can sustain a pH range of 2 to 12 and a maximum pressure of 50 psi. They are used by major biopharmaceutical companies at production scales. Sepragen's president owns a U.S. patent (4) on the design of unpacked RFC columns. Figure 1 shows an acrylic RFC column with inward radial flow. The feedstream enters from a center input port at the top. It is then distributed through several flow channels to the outer shell and subsequently enters the packing media in the radial direction toward a center collection tube. The effluent exits the tube through an outlet port at the bottom. The two ports next to the center input port at the top in Figure 1 are packing ports used during media packing. A bubble trap with a pressure gauge sits on the edge of the column's top lid. Figure 2 is an anatomical view of a Superflo column. Superflo columns usually use inward flow instead of outward flow because it is more difficult to distribute outward flow without increasing flow distortion caused by gravity. The other reason for not using outward flow is because inward flow provides slightly sharper peaks than outward flow based on computer simulation (5). Outward flow is usually used for the packing or regeneration of a RFC column.

PACKING PROCEDURES FOR RFC COLUMNS

Because of their structure, CUNO's Zetaffinity cartridges do not require special packing procedures. Sepragen's Superflo columns are packed by first displacing air using outward flow with a buffer solution. Packing slurry is then pumped into the column through the two packing ports on the column top surface (Fig. 1). Excess buffer is squeezed out of the column through the input port on the column top surface. After that, the column is ready for use.

PRESSURE DROPS OF RFC COLUMNS

Because of a short flow path and a low pressure drop, both CUNO's and Sepragen's RFC devices exhibit a highly linear relationship between bed pressure drop versus flow rate. Figure 6 shows a linear relationship between bed pressure and flow rate for an 800-mL Zetaffinity cartridge studied by Huang et al. (3).

Figure 7 shows the pressure drops for a Superflo 20 L column and a Superflo 50 L column packed with several different soft-gel chromatographic media (6). It indicates that the pressure drops are quite low at relatively high flow rates. At lower flow rates, the pressure drop curve is linear. This behavior is similar to that of short pancakelike AFC columns.

Figure 5. Superflo columns. *Source*: Courtesy of Sepragen Corp.

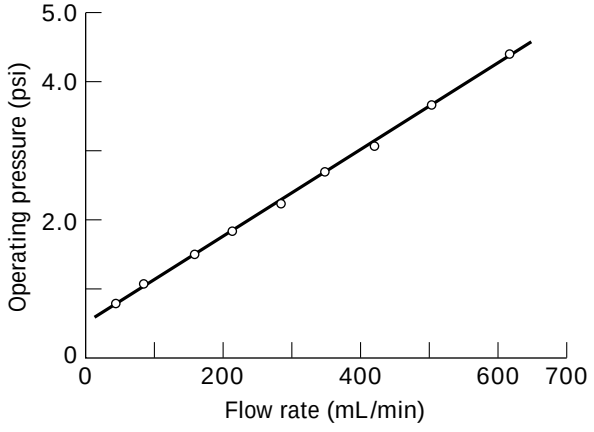


Figure 6. Pressure drop vs. flow rate for an 800-mL Zetaffinity cartridge.

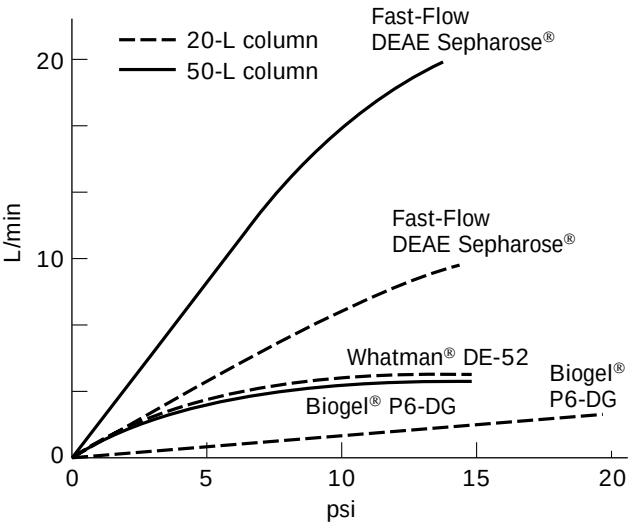


Figure 7. Pressure drop vs. flow rate for 20-L and 50-L Superflo® columns. *Source*: Courtesy of Sepragen Corp.

COMPARISON OF RADIAL FLOW AND AXIAL FLOW COLUMNS

Saxena and Weil (7) did an experimental case study of the comparison of RFC and AFC. They used a 100-mL axial flow glass column (Econocolumn®) with 2.5-cm i.d. from BioRad Laboratories (Richmond, Calif.) and a Superflo-100 RFC column from Sepragen with a bed volume of 100 mL. Both columns were packed with QAE cellulose with the following procedures. The Superflo column was packed using a 25% slurry (with 0.5 M NaCl) that was pumped in through the two packing ports at a flow rate of 30 mL/min. The final QAE cellulose density after packing was 6 mL per dry gram. The Superflo column was packed in about 20 min. The AFC column was packed using 50% slurry added from the top of the column. After the liquid was drained and the bed was settled, additional slurry was added. This process was repeated until the same amount of QAE cellulose was packed into the AFC column. The two columns were used to separate two 10-mL ascites fluid samples. The samples were pretreated by dialyzing the ascites fluid with 10 mM phosphate buffer of pH 8.0 for 2 days with three changes. Precipitates and debris were then removed by centrifugation. Stepwise salt gradients were used. Figure 8 shows the comparison between the 100-mL

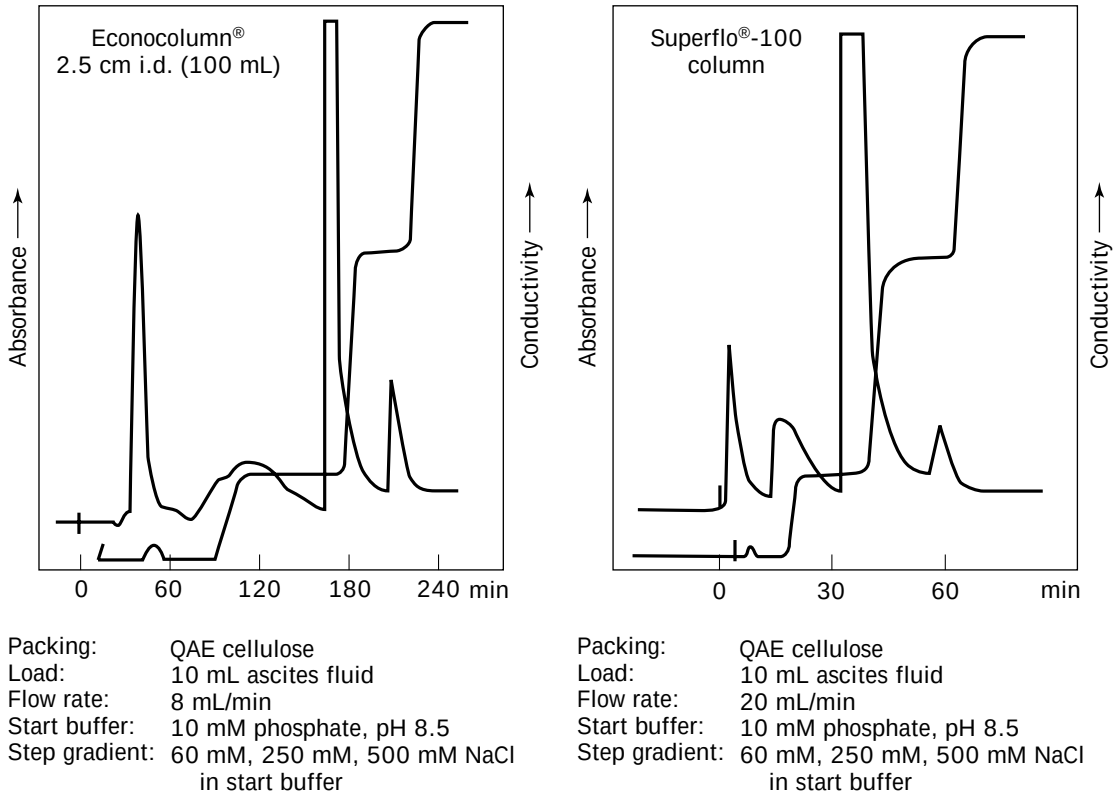


Figure 8. Comparison of AFC and RFC columns. *Source*: Courtesy of Sepragen Corp.

AFC column and the 100-mL Superflo-100 RFC column. From Figure 8, it can seen that the RFC column achieved similar separation results with much less time. In this case, the two columns were equivalent only in bed volume. A stricter comparison should use a pancakelike, short AFC column with its bed height about the same as the packing's thickness in radial direction for the RFC column, and both columns should have equal bed volume. The case study discussed next is close to satisfying these conditions.

Tharakan and Belizaire (8,9) used a 50-mL RFC column with a bed height of 0.95 cm and packing thickness of 3.0 cm in the radial direction and a 50-mL AFC column with a bed height of 2.8 cm. Both columns were packed with Sepharose CL2B resin-containing monoclonal antibody from Pharmacia (Piscataway, N.J.). They studied the purification of a protein called factor IX. Their experimental results indicated that the two columns gave similar purification results. This was expected because the AFC column was pancakelike. Tharakan and Belizaire (8) also packed the same column with S-200 Sephacryl size-exclusion gel from Sigma Chemical Co. (St. Louis). The protein band was more diffused using RFC than AFC.

Lane et al. (10) compared the performance of a Superflo-100 RFC column and a 6.6 $\times$ 4.4 cm i.d. AFC column for the separation of egg-white proteins. Both columns had a nominal bed volume of 100 mL. They were packed with the same media for comparison. Two anion-exchange cellulose media were tested, Whatman DE52 and QA52 from Whatman Specialty Products Division (Maidstone, U.K.). Egg whites were first separated from fresh hen eggs and then treated with a buffer. After treatment with a cell debris remover, egg-white suspensions were filtered using filter paper. The samples for chromatography had a protein concentration of 14 mg/mL. Sample load volume was 40 mL. After sample loading, the column was washed with a buffer. Elution was carried out using a linear gradient of 0 to 0.5 M NaCl in 0.025 M tris/HCl buffer at pH 7.5. Various flow rates were tested, ranging from 5 to 50 mL/min for the AFC column and 5 to 150 mL/min for the RFC column. Figures 9 and 10 represent typical results obtained by Lane et al. (10). They indicate that the AFC column gave slightly sharper peaks and faster elution times for both DE52 and QA52 media at a flow rate of 25 mL/min.

PROS AND CONS OF RADIAL FLOW COLUMNS

RFC columns provide a short flow path and a large cross-sectional area. This has the same effect as short pancakelike AFC columns. However, RFC columns occupy considerably less floor space. Both RFC and pancakelike AFC columns face flow distribution problems. According to Sepragen, its Superflo columns have better flow distribution than typical large pancakelike columns.

Compared to long AFC columns, RFC columns produce smaller pressure drops, thus enabling larger volumetric flow rates. If soft gels are used as separation media, the low pressure drop of RFC columns helps prevent bed compression (11,12). RFC is especially suitable for affinity chromatography using soft gels. In affinity chromato-

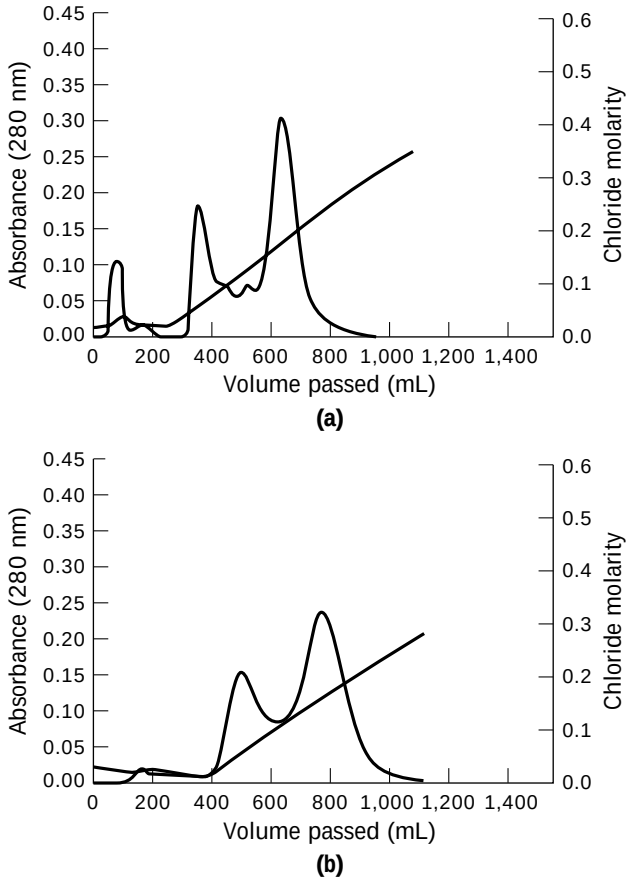


Figure 9. Comparison between a 100-mL AFC column (a) and a 100-mL RFC column (b) packed with DE52 in the separation of egg-white proteins.

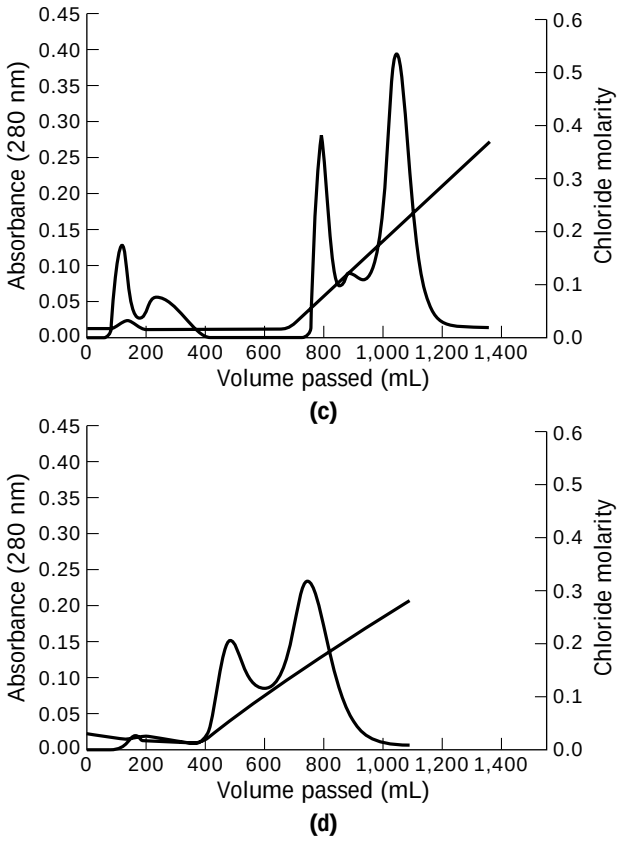


Figure 10. Same as Figure 9, but the columns were packed with QA52.

graphic operations, dilute feeds are typically used. Because of the extremely high affinity between the product and the gel matrix, very high flow rates are permitted without sacrificing column resolution. RFC is also a good choice for strong reversed phase, hydrophobic interaction and ion-exchange media.

Scale-up of RFC columns are relatively more straightforward because it is usually done by increasing the column height. To a certain extent, this does not seems to increase flow distortion problems in practice. If an AFC column is scaled up by increasing diameter, the flow distribution behaves quite differently from a small AFC column for which plug flow is easy to achieve. Thus, predicting actual performance of a large pancakelike AFC column from a small one is difficult.

The disadvantage of the RFC column is primarily its limited resolution caused by a short flow path. If the flow path is increased to a large extent, there will be flow distribution problems in the radial direction because of gravity. High resolution for more demanding separations can only come from using AFC columns with sufficiently large column length. This is precisely the reason why RFC has no use in analytical HPLC. RFC is not suitable for separations in which solute-stationary interactions are weak. For example, RFC is not a suitable choice for size-exclusion chromatography (SEC), because SEC depends strongly on the length of flow path for its resolution. RFC's short flow length cannot meet the demand. It is also expected that RFC is not suitable for other chromatography with weak solute-stationary phase interactions.

If mechanically strong packing materials, such as silica-based particles, are used, a longer column flow path can be used because the bed can sustain a much higher pressure. If a relatively high resolution is desired, RFC would be at a disadvantage compared to AFC since RFC is limited by its short flow path.

APPLICATION EXAMPLES

Applications Using Zetaffinity Cartridges

Huang et al. (3) studied several Zetaffinity cartridges from CUNO that contained modified cellulose-based affinity media. Table 1 lists the dimensions of three Zetaffinity cartridges tested by Huang et al. Figure 11 shows a chromatogram for the removal of proteases from human plasma using an 800-mL Zetaffinity cartridge with a flow rate of 100 mL/min (3). The cartridge contained modified cellulose with para-aminobenzamidine (PAB) as ligand. A 70% efficiency of protease removal was achieved in a single pass. The treated plasma was expected to have a three-fold increase in stability. The same type of 800-mL Zetaf-

Table 1. Dimensions of CUNO Radial Flow Cartridges Tested by Huang et al. (3)

Parameter	Small	Medium	Large
Nominal size	250 mL	800 mL	3200 mL
Bed volume	210 mL	810 mL	3020 mL
Outer diameter	7.0 cm	12.7 cm	12.7 cm
Inner diameter	0.6 cm	0.9 cm	0.9 cm
Height	6.4 cm	6.4 cm	23.8 cm

finity cartridge loaded with 1,260 mg PAB ligand was also used to purify crude trypsin purchased from Sigma Chemical. The data in Figure 12 were obtained by Huang et al. (3) using a flow rate of 295 mL/min.

Planques et al. (13) used a 250-mL Zetaffinity cartridge from CUNO packed with modified cellulose as chromatography media. They first chemically treated the cartridge to couple the media with L-lysine. This affinity chromatography media was able to bind with a human plasminogen protein. After centrifugation and microfiltration, human plasma was diluted with a buffer and then applied to the cartridge at a flow rate of 20 mL/min. After washing and elution, a 85% recovery yield and an increase of 110-fold in specific activity were achieved.

Applications Examples Using Sepragen's Superflo Columns

Akoum et al. (14) used a Superflo-400 RFC column packed with histidyl-Sepharose gel for the purification of myxalin, a glycopeptide with anticoagulant property. The feed for the column was obtained from a fermentation broth with the microorganism *Myxococcus xanthus*. Before the feed was applied to the column, it was clarified and concentrated using centrifugation, microfiltration, and reverse osmosis. A relative short processing time was achieved.

Strætkvern et al. (15) used a 60-mL, 2.2-cm i.d. AFC column, a Superflo-250 column (250 mL in bed volume), and a 2,500-mL AFC column for the separation of DNase from the extracts of cod pyloric cacca. The packing medium was Q-Sepharose Fast Flow anion exchange gel from Pharmacia. Column dimensions and operating conditions are listed in Table 2. Table 3 is a summary of their experimental results. Their results indicate that the RFC column required much less time and achieved higher productivity. The superior performance in this case should not be inter-

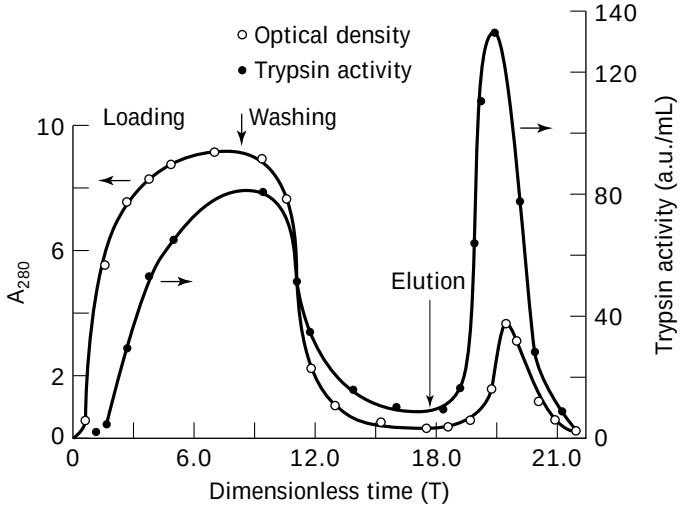


Figure 11. Protease removal from human plasma using a Zetaffinity® cartridge. a.u., affinity unit.

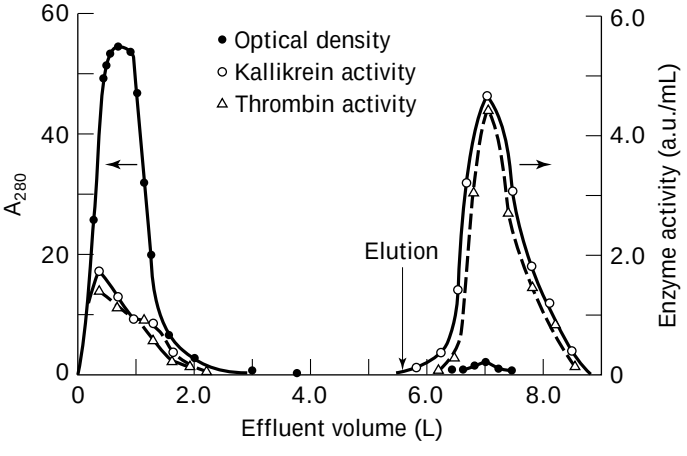


Figure 12. Trypsin purification using a Zetaffinity® cartridge. a.u., affinity unit.

Table 2. Columns Used by Strætkvern et al. (15)

Parameter	Small column	Medium column	Large column
Column type	Axial flow	Radial flow	Axial flow
Column volume (mL)	60	250	2,500
Cross-sectional flow area (cm^2)	3.8	120 (outer)	154
Volumetric flow rate (L/h)	0.6	17.4	20.9
Sample volume (L)	0.033	0.135	1.32
Protein concentration (mg/mL)	2.6	2.6	2.4
Scale-up factor	1	4	40

Table 3. Purification Results Obtained by Strætkvern et al. (15)

Parameter	Small column	Medium column	Large column
Elution volume (L)	0.20	0.87	3.57
Protein (mg)	10	44	350
Total activity (units $\times 10^{-6}$)	20	78	783
Specific activity (units $\times 10^{-6}$ mg^{-1})	2.00	1.77	1.35
Yield (%)	100	107	76
Purification (fold)	20	17	13
Cycle time (h)	1.3	0.25	1.4
Productivity (units $\times 10^{-6}$ $h^{-1} \cdot mL^{-1}$ gel)	0.256	1.25	0.135
Productivity ($mg \cdot h^{-1} \cdot mL^{-1}$ gel)	0.13	0.70	0.1

preted as a fixed rule here because the AFC columns they used were not equivalent pancakelike columns.

Weaver et al. (16) used a 10-L Superflo RFC column packed with Q-Sepharose Fast Flow medium from Pharmacia for the separation of uridine phosphorylase from total crude extracts of *Escherichia coli*. After fermentation, 375 L of broth was concentrated to 10 L. It was then washed with 50 L of 20 mM K_3PO_4 buffer containing 1 mM $MgCl_2$. After adding 20 μg lysozyme per milliliter of broth, 100 mg of DNase, and 100 mg of RNase, the cells were homogenized using a bead mill. The supernatant was diluted to 30 L and then applied to the 10-L RFC column. The feed was recirculated back to the column at a flow rate of 1.3 L/min for 3 h and then discharged. The column was washed with three different buffer solutions to remove bound lipid and hydrophobic proteins. Elution was then carried out using a buffer containing 0.225 mM NaCl. The collected effluent was 50 L. The final product after dialysis had a purity of 85% and a recovery yield of 82%.

McCartney (17) used two Superflo-100 RFC columns packed with S-Sepharose FF (Pharmacia) ion-exchange media in tandem for the purification of an undisclosed recombinant protein from *E. coli*. The system was able to process 4 L feed in less than 2 h with a concentration factor of 64 times.

Saxena et al. (12) applied a Superflo-1500 column (1,500 mL in bed volume) packed with immobilized Protein-A Sepharose in the purification of an antimelanoma IgG2A antibody from ascites fluids. Sample was loaded with a flow rate of 104 mL/min. After loading, the column was washed with a buffer using a flow rate of 170 mL/min. Elution was carried out with a flow rate of 92 mL/min. An actual recovery of 3.1 g of the antibody with a product purity greater than 97% was achieved in 3.5 h. A Superflo-1500 packed with riboflavin immobilized on a Sepharose-4B matrix via an epoxy linkage was used by the same researchers to obtain a crude preparation of a riboflavin-binding protein (12). The entire run was carried out using a flow rate of 350 mL/min. They also used a Superflo-200 column packed with antiricin B-chain antibody immobilized on cross-link agarose (12). The flow rate used was 45 mL/min. A 100% yield was achieved with a product amount of 2.1 g and a purity of 100%.

MATHEMATICAL MODELING OF RADIAL FLOW CHROMATOGRAPHY

In 1950, Lapidus and Amundson (18) proposed a simplified theoretical model for RFC. Their model ignores radial diffusion in the bulk-fluid phase and intraparticle diffusion. It is similar to that used by Rachinskii (19). Inchin and Rachinskii (20) subsequently included molecular diffusion in the bulk-fluid phase. Lee et al. (21) proposed several single component rate models for the comparison of statistical moments for RFC and AFC. They included radial dispersion, intraparticle diffusion, and external mass transfer effects. Kalinichev and Zolotarev (22) performed an analytical study on moments for single component RFC in which they treated the radial dispersion coefficient as a variable.

A rate model for nonlinear single component RFC was solved numerically by Lee (23) by using the finite difference and orthogonal collocation methods. His model considered radial dispersion, intraparticle diffusion, external mass transfer, and nonlinear isotherms. It used averaged radial dispersion and mass transfer coefficients instead of treating them as variables. A nonlinear model of this kind of complexity has no analytical solution and must be solved numerically.

Realistic modeling of RFC should treat the radial dispersion and external mass transfer coefficients as variables rather than constants because the linear flow velocity (v) in the RFC column changes continuously along the radial coordinate of the column. Without this distinctive important feature the curvature in the flow path is lost, and thus the column can be imaginatively cut and spread out to become exactly like a pancakelike AFC column.

General Rate Model for Multicomponent RFC

Gu et al. (24,25) presented a general rate model for RFC in which radial dispersion and external mass transfer coefficients are treated as variables rather than constants. The model was solved numerically using finite element and the collocation methods for discretizations of the bulk-fluid phase and the particle phase governing partial differential equations (PDEs), respectively. Figure 13 is an anatomy of a RFC column for the purpose of modeling. The following basic assumptions are made in order to formulate a general rate model for RFC.

1. The column is isothermal.
2. The porous particles in the bed can be treated as spherical and uniform in diameter.
3. The concentration gradients in the axial direction are negligible. This means that the maldistribution of radial flow is ignored.
4. The fluid inside particle macropores is stagnant, that is, there is no convective flow inside macropores.
5. An instantaneous local equilibrium exists between the macropore surfaces and the stagnant fluid in the macropores.
6. The film mass transfer theory can be used to describe the interfacial mass transfer between the bulk-fluid and particle phases.
7. The diffusional and mass transfer coefficients are constant and independent of the mixing effects of the components involved.

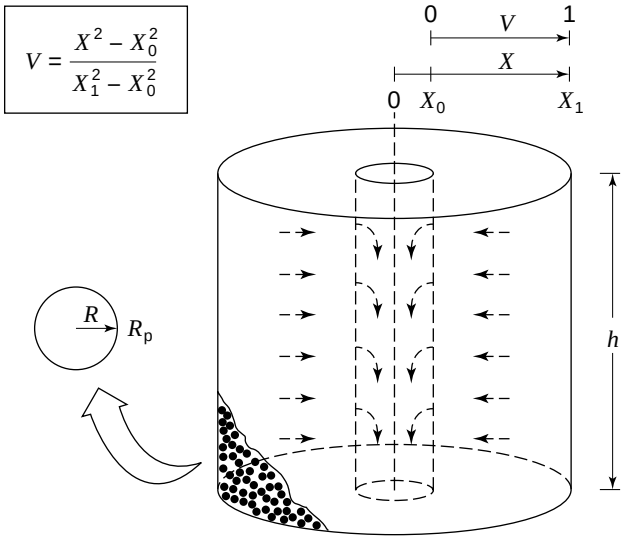


Figure 13. Anatomy of an inward-flow RFC column.

Based on these basic assumptions, equations 1 and 2 are formulated from the differential mass balance for each component in the bulk-fluid and particle phases. In equation 1, the plus sign represents outward flow, and −v is inward flow.

$$-\frac{\partial}{\partial X}\left(D_{bi}X\frac{\partial C_{bi}}{\partial X}\right) \pm v\frac{\partial C_{bi}}{\partial X} + \frac{\partial C_{bi}}{\partial t} + \frac{3k_i(1-\epsilon_b)}{\epsilon_b R_p}(C_{bi} - C_{pi,R=R_p}) = 0 \quad (1)$$

$$(1-\epsilon_p)\frac{\partial C_{pi}^*}{\partial t} + \epsilon_p\frac{\partial C_{pi}}{\partial t} - \epsilon_p D_{pi}\left[\frac{1}{R^2}\frac{\partial}{\partial R}\left(R^2\frac{\partial C_{pi}}{\partial R}\right)\right] = 0 \quad (2)$$

in which C_{pi}^* is related to C_{pi} via isotherm (24).

The initial conditions for the PDE system are at

$$t = 0,\ C_{bi} = C_{bi}(0, X) \quad (3)$$

and

$$C_{pi} = C_{pi} = C_{pi}(0, R, X) \quad (4)$$

The boundary conditions are at the inlet X position

$$\partial C_{bi}/\partial X = (v/D_{bi})[C_{bi} - C_{fi}(t)] \quad (5)$$

and at the outlet X position

$$\partial C_{bi}/\partial X = 0 \quad (6)$$

Equations 1 and 2 can be written in dimensionless forms as follows.

$$-\frac{\partial}{\partial V}\left(\frac{\alpha}{Pe_i}\frac{\partial c_{bi}}{\partial V}\right) \pm \frac{\partial c_{bi}}{\partial V} + \frac{\partial c_{bi}}{\partial \tau} + \chi_i(c_{bi} - c_{pi,r=1}) = 0 \quad (7)$$

$$\frac{\partial}{\partial \tau}[(1-\epsilon_p)c_{pi}^* + \epsilon_p c_{pi}] - \eta_i\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial c_{bi}}{\partial r}\right)\right] = 0 \quad (8)$$

In equations 1 to 7, the dimensionless variable, $V = (X^2 - X_0^2)/(X_1^2 - X_0^2) \in [0, 1]$, is based on the local volume averaging method [24]. $\alpha = 2\sqrt{V + V_0}(\sqrt{1 + V_0} - \sqrt{V_0})$ is a function of V.

The dimensionless initial conditions are at

$$\tau = 0,\ c_{bi} = c_{bi}(0, V) \quad (9)$$

and

$$c_{pi} = c_{pi}(0, r, V) \quad (10)$$

The dimensionless boundary conditions are

$$\partial c_{bi}/\partial V = Pe_i[c_{bi} - C_{fi}(\tau)/C_{0i}] \quad (11)$$

At the inlet V position, for frontal adsorption, $C_{fi}(\tau)/C_{0i} = 1$. For elution,

$$C_{\mathrm{fi}}(\tau)/C_{0\mathrm{i}} = \begin{cases} 1 & 0 < \tau < \tau_{\mathrm{imp}} \\ 0 & \text{otherwise} \end{cases}$$

After the introduction of a sample in the form of a rectangular pulse, it component I is displaced, $C_{\mathrm{fi}}(\tau)/C_{0\mathrm{i}} = 0$. If component I is a displacer, $C_{\mathrm{fi}}(\tau)/C_{0\mathrm{i}} = 1$. At the outlet V position, $\partial c_{\mathrm{bi}}/\partial V = 0$. For the particle phase governing equation, the boundary conditions are at r = 0

$$\partial c_{\mathrm{pi}}/\partial \mathrm{r} = 0 \tag{12}$$

and at r = 1

$$\partial c_{\mathrm{pi}}/\partial r = Bi_{\mathrm{i}}(c_{\mathrm{bi}} - c_{\mathrm{pi,r=1}}) \tag{13}$$

Note that all the dimensionless concentrations are based on $C_{0\mathrm{i}}$, that is, the maximum of the feed profile $C_{\mathrm{fi}}(\tau)$ for each component.

The radial dispersion coefficient (D_{bi}) depends on the linear velocity (v). In liquid chromatography, it can be assumed (5,22,25) that $D_{\mathrm{bi}} \propto v$. Thus, $Pe_{\mathrm{i}} = v\,(X_1 - X_0)/D_{\mathrm{bi}}$ can be considered constant in liquid RFC. The variation of Bi_{i} values observes the following relationship:

$$Bi_{\mathrm{i}} \propto k_{\mathrm{i}} \propto v^{1/3} \propto (1/X)^{1/3} \propto (V + V_0)^{-1/6} \tag{14}$$

If $Bi_{\mathrm{i,V=1}}$ values are known, Bi_{i} values anywhere else can be obtained from equation 15.

$$Bi_{\mathrm{i,V}} = [(1 + V_0)/(V + V_0)]^{1/6} Bi_{\mathrm{i,V=1}} \tag{15}$$

χ_{i} can be calculated from Bi_{i} using its definition $\chi_{\mathrm{i}} = 3Bi_{\mathrm{i}}\eta_{\mathrm{i}}(1 - \epsilon_{\mathrm{b}})/\epsilon_{\mathrm{b}}$.

Numerical Solution

The PDE system of the governing equations is first discretized. The finite element and orthogonal collocation methods are used to discretize the bulk-fluid phase and the particle phase governing equations, respectively. The resulting ordinary differential equation (ODE) system is then solved using the public domain ODE solver called DVODE written by Brown et al. (26). The compiled Fortran program is free to academic researchers. Both DOS and Windows 95 versions are available. Information on the software is available at *http://www.ent.ohiou.edu/~guting/CHROM/*.

A study of the effects of treating D_{bi} and k_{i} as variables compared to treating them as constants was carried out by Gu et al. (5,25). The comparison between RFC and AFC was also studied through computer simulation. Figure 14 shows that outward flow gives slightly sharper concentration profile (5). The figure also shows that RFC gives similar concentration profiles as AFC with equivalent physical parameters. These theoretical results support the experimental results obtained by Tharakan and Belizaire (8).

SCALE-UP OF RFC COLUMNS

One of the advantages of RFC columns is the relative ease for scale-up. In Sepragen's Superflo column series, increasing the column height is a rather safe way to accommodate an increase of sample size to a certain degree. However, one must keep in mind that maldistribution in radial flow may deteriorate slightly. If the bed thickness in the radial direction is increased, bed pressure usually increases proportionally. The performance prediction will be less reliable if the column diameter is increased. Figure 15 shows that a 15-fold increase of sample load and bed column produced very similar performances when DEAE cellulose is used to separate an ascites fluid (27,28). Figure 16 is an example with a 50-fold increase of sample load and bed column (27). These two examples are very successful examples.

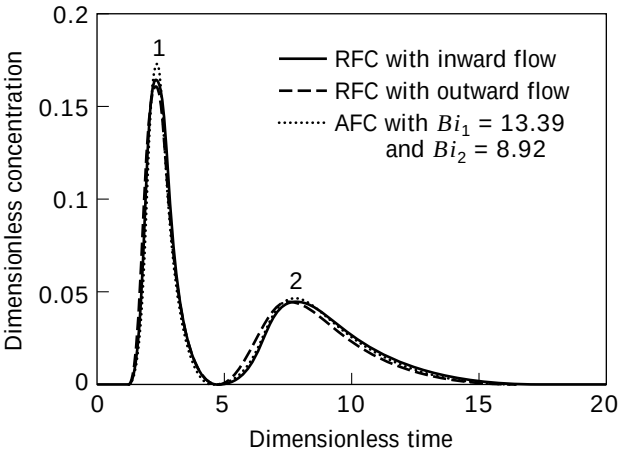
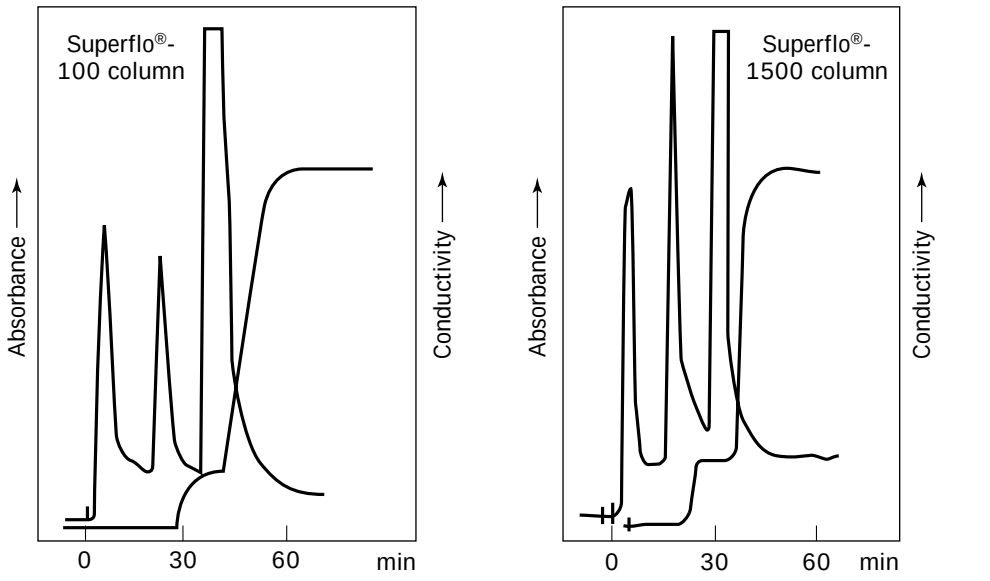


Figure 14. Simulated comparison for inward- and outward-flow RFC and AFC.

The mathematical model introduced above can be used to help the scale-up process. Before a column is bought, its performance can be predicted using computer simulation. Isotherm data must be obtained experimentally or be supplied by the vendor of the packing material. Mass transfer parameters can be estimated by using existing correlations (24).

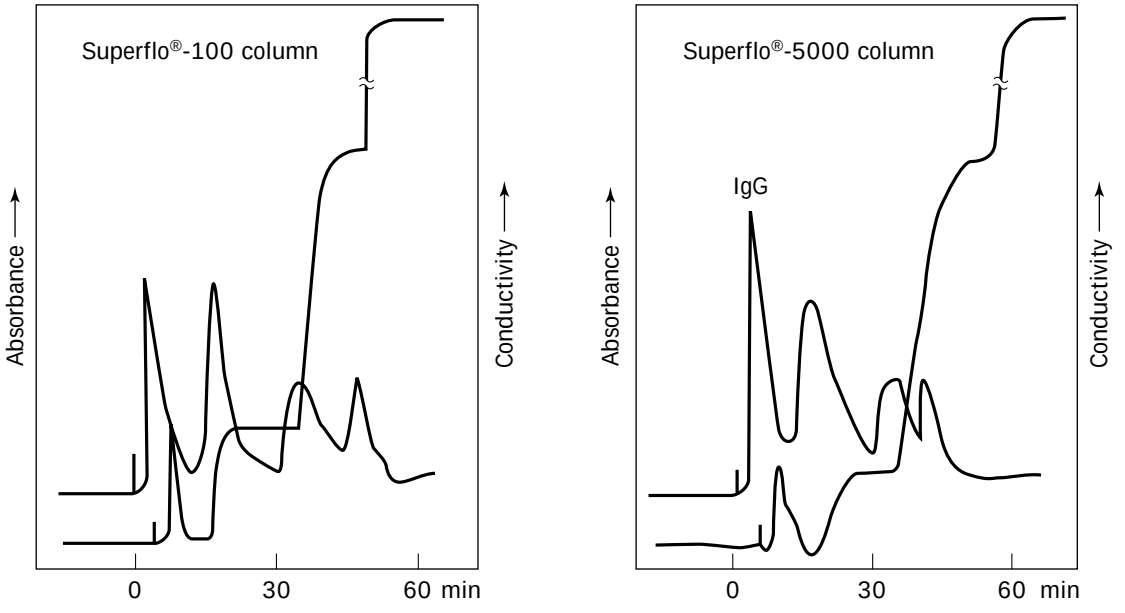
CONCLUSIONS

RFC columns have a short flow path and a large flow area, resulting in a small bed pressure. They are especially suitable when soft-gels are used. Because of its limited resolution, RFC should be used in chromatographic separations involving strong solute–stationary phase interactions, such as affinity chromatography, strong anion or cation exchange, strong reversed phase, and strong hydrophobic interaction chromatography. RFC is not suitable for SEC and other chromatography with weak solute–stationary phase interactions. RFC is an attractive alternative to AFC in preparative- and large-scale separations. Both experimental and theoretical modeling indicate that to a large extent an RFC column behaves much like a pancake-style AFC column with the same packing volume and with its bed height about the same as the packing's radial thickness in RFC column. However, RFC has a much



Packing:	DEAE cellulose	Packing:	DEAE cellulose
Load:	10 mL ascites fluid	Load:	150 mL ascites fluid
Flow rate:	10 mL/min	Flowrate:	150 mL/min
Start buffer:	10 mM phosphate, pH 8.5	Start buffer:	10 mM phosphate, pH 8.5
Step gradient:	60 mM, 500 mM NaCl in start buffer	Step gradient:	60 mM, 500 mM NaCl in start buffer

Figure 15. A 15-fold scale-up example using Superflo® columns.



Packing:	DEAE cellulose	Packing:	DEAE cellulose
Load:	10 mL cell culture fluid (murine IgG)	Load:	500 mL cell culture fluid (murine IgG)
Flow rate:	10 mL/min	Flow rate:	500 mL/min
Start buffer:	10 mM phosphate, pH 8.5	Start buffer:	10 mM phosphate, pH 8.5
Step gradient:	60 mM, 250 mM, 700 mM NaCl in start buffer	Step gradient:	60 mM, 250 mM, 700 mM NaCl in start buffer

Figure 16. A 15-fold scale-up example using Superflo® columns.

smaller footprint, and it seems to handle flow distribution better than very wide pancakelike columns.

NOMENCLATURE

Bi_i	Biot number of mass transfer for component i, $k_i R_p(\epsilon_p D_{pi})$
C_{0i}	Concentration used for nondimensionalization, $\max\{C_{fi}(t)\}$
C_{bi}	Bulk-fluid phase concentration of component i
C_{fi}	Feed concentration profile of component i, a time-dependent variable
C_{pi}	Concentration of component i in the stagnant fluid phase inside particle macropores
C^*_{pi}	Concentration of component i in the solid phase of particle (based on unit volume of particle skeleton)
c_{bi}	$= C_{bi}/C_{0i}$
c_{pi}	$= C_{pi}/C_{0i}$
c^*_{pi}	$= C^*_{pi}/C_{Oi}$
D_{bi}	Axial or radial dispersion coefficient of component i
D_{pi}	Effective diffusivity of component i, porosity not included
k_i	Film mass transfer coefficient of component i
Pe_i	Peclet number of radial dispersion for component i, $v(X_1 - X_0)/D_{bi}$
R	Radial coordinate for particle
R_p	Particle radius
r	$= R/R_p$
t	Dimensional time ($t = 0$ is the moment a sample enters a column)
v	Interstitial velocity
V_0	Dimensionless constant corresponding to X_0
V	Dimensionless volumetric coordinate
X	Coordinate in the radial direction for a RFC column.

Greek Letters

α	$= 2\sqrt{V + V_0}(\sqrt{1 + V_0} - \sqrt{V_0})$ for RFC
ϵ_b	Bed void volume fraction
ϵ_p	Particle porosity
η_i	Dimensionless constant, $\epsilon_p D_{pi} L/(R_p^2 v)$
χ_i	Dimensionless constant for component i, $3Bi_i\eta_i(1 - \epsilon_b)/\epsilon_b$
τ	Dimensionless time, vt/L
τ_{imp}	Dimensionless time duration for a rectangular pulse of the sample

ACKNOWLEDGMENT

The author wishes to thank Mr. Sanjeev Sexena of Sepragen Corp., Hayward, Calif., for providing information on Sepragen's Superflo® radial flow columns.

BIBLIOGRAPHY

1. D. McCormick, *Biotechnology* **6**, 158–160 (1988).
2. P. Hopf, *Ind. Eng. Chem.* **39**, 938–940 (1947).
3. S.H. Huang, S. Roy, K.C. Hou, and G.T. Tsao, *Biotechnol. Prog.* **4**, 159–165 (1988).
4. U.S. Pat. 4,627,918 (Dec. 9, 1986), V. Saxena (to Sepragen Corporation).
5. T. Gu, G.-J. Tsai, and G.T. Tsao, in A. Fiechter ed., *Advances in Biochemical Engineering / Biotechnology*, Springer-Verlag, Berlin, 1993, pp. 73–95.
6. V. Saxena and M. Dunn, *Biotechnology* **7**, 250–255 (1982).
7. V. Saxena and A.E. Weil, *BioChromatography* **2**, 90–97 (1987).
8. J.P. Tharakan and M. Belizaire, *J. Liq. Chromatogr.* **18**, 39–49 (1995).
9. J.P. Tharakan and M. Belizaire, *J. Chromatogr.* **702**, 191–196 (1995).
10. L. Lane, M.L. Koscielny, P.R. Levison, D.W. Toome, and E.T. Butts, *Bioseparation* **1**, 141–147 (1990).
11. P. Ernst, *Aust. J. Biotechnol.* **1**, 22–26 (1987).
12. V. Saxena, A.E. Weil, R.T. Kawahata, W.C. McGregor, and M. Chandler, *Am. Lab.* **19**, 112–120 (1987).
13. Y. Planques, H. Pora, and F.D. Menozzi, *J. Chromatogr.* **539**, 531–533 (1991).
14. A. Akoum, F. Devichi, M. Kalyanpur, J.P. Neff, M.A. Vijayalakshmi, and M. Sigot, *Process Biochem.* 55–59 (April 1989).
15. K.O. Strætkvern, A.J. Raae, K. Folkvord, B.A. Næss, I.M. Aasen, *Bioseparation* **2**, 81–93 (1991).
16. K. Weaver, D. Chen, L. Walton, L. Elwell, and P. Ray, *BioPharm* 25–29 (July/August 1990).
17. J.E. McCartney, *BioTechniques* **11**, 648–649 (1991).
18. L. Lapidus and N.R. Amundson, *J. Phys. Colloid. Chem.* **54**, 821–825 (1950).
19. V.V. Rachinskii, *J. Chromatogr.* **33**, 234–242 (1968).
20. P.A. Inchin and V.V. Rachinskii, *Russ. J. Phys. Chem.* **47**, 1331–1333 (1977).
21. W.-C. Lee, S.H. Huang, and G.T. Tsao, *AlChE J.* **34**, 2083–2087 (1988).
22. A.I. Kalinichev and P.P. Zolotarev, *Russ. J. Phys. Chem.* **51**, 871–873 (1977).
23. W.-C. Lee, Ph.D. Thesis, Purdue University, West Lafayette, Ind., 1989.
24. T. Gu, *Mathematical Modeling and Scale-Up of Liquid Chromatography*, Springer-Verlag, Berlin, 1995, pp. 102–106, 116.
25. T. Gu, G.-J. Tsai, and G.T. Tsao, *Chem. Eng. Sci.* **46**, 1279–1288 (1991).
26. P.N. Brown, G.D. Byrne, and A.C. Hindmarsh, *Sci. Stat. Comput.* **10**, 1038–1051 (1989).
27. V. Saxena and A.E. Weil, *BioChromatography* **2**, 90–97 (1987).
28. V. Saxena, K. Subramanian, S. Saxena, and M. Dunn, *BioPharm* 46–49 (March 1989).

See also CHROMATOGRAPHY, COMPUTER-AIDED DESIGN; CHROMATOGRAPHY, ION EXCHANGE; CHROMATOGRAPHY, SIZE EXCLUSION.

CHROMATOGRAPHY, SIZE EXCLUSION

CHARLES M. ROTH
MARTIN L. YARMUSH
Massachusetts General Hospital, Shriners Burns Hospital, and Harvard Medical School
Boston, Massachusetts

KEY WORDS

Desalting
Excluded volume
Large scale
Packing materials
Purification
Resolution

OUTLINE

INTRODUCTION

Size-exclusion chromatography (SEC) is a widely used means for separation of macromolecules of different sizes. Porath and Flodin in 1959 first demonstrated the use of SEC, or gel filtration chromatography, for separations by resolving glucose from two size fractions of dextrans (1). In addition, they applied the technique to desalting of serum proteins. Remarkably, these two applications—crude separations and desalting—remain the primary uses of SEC today. A wide range of materials, including sugars, proteins, and nucleic acids, have been analyzed or separated using this technique. In addition, SEC is used extensively in the characterization of polymers. However, in this article, the emphases are on analysis and on separations involving biological macromolecules.

Basic Principle

The essential principle of size exclusion chromatography is partitioning of species on the basis of preferential sieving into the pores of support particles. Sample is applied to the top of a bed packed with porous stationary phase particles. Elution results from flow of a fixed composition buffer through the bed (isocratic elution), in contrast with other modes of chromatography in which the buffer composition is varied (gradient elution). Molecules that are larger than the pores of the support access only the interstitial space between particles and therefore pass through the column quickly, whereas smaller particles access the volume within the pores and are retained for a longer time by the column (Fig. 1). Therefore, a support must be selected with a pore size distribution so that there exists a differential partitioning of the molecules to be separated.

Lexicon

In any chromatographic process, sample is injected into the column and eluted by continuous flow of a buffer through the column. The time between injection and the peak of any eluted solute is termed its retention time (t_R). An equivalent measure of retention is the retention volume (V_R), which, for the usual case of constant flow rate, is merely the product of the flow rate and retention time. The terms *elution time* and *elution volume* are used interchangeably with *retention time* and *retention volume*. For solute molecules larger than the size of the pores, the elution volume represents the volume of the column between the support particles plus that of extracolumn tubing between the injection loop and the column and from the column to the detector. This is known as the void volume (V_0), and it is frequently estimated using large, inert tracer solutes. The pore volume (V_p) is the volume within the particles occupied by fluid and, in principle, accessed entirely by very small particles, for which the retention volume is maximized:

$$V_{R,max} = V_0 + V_p \quad (1)$$

In general, the elution volume is intermediate between the void volume and the sum of the void and pore volumes and thus defines a distribution coefficient (K_d) through the relation

$$V_R = V_0 + K_d V_p \quad (2)$$

By definition, the distribution coefficient assumes values only between zero and unity. The other relevant volume is that of the support particles or chromatographic matrix (V_m). In principle, one can calculate the pore volume based on the amount and density of the support through the relation

$$V_p = \frac{W_r \rho_{gel}}{1 + W_r \rho_{H_2O}} (V_t - V_0) \quad (3)$$

where V_t is the total volume (sum of void, pore, and matrix volumes, the first corrected for the volume of extracolumn tubing), ρ represents mass density, and W_r is the water

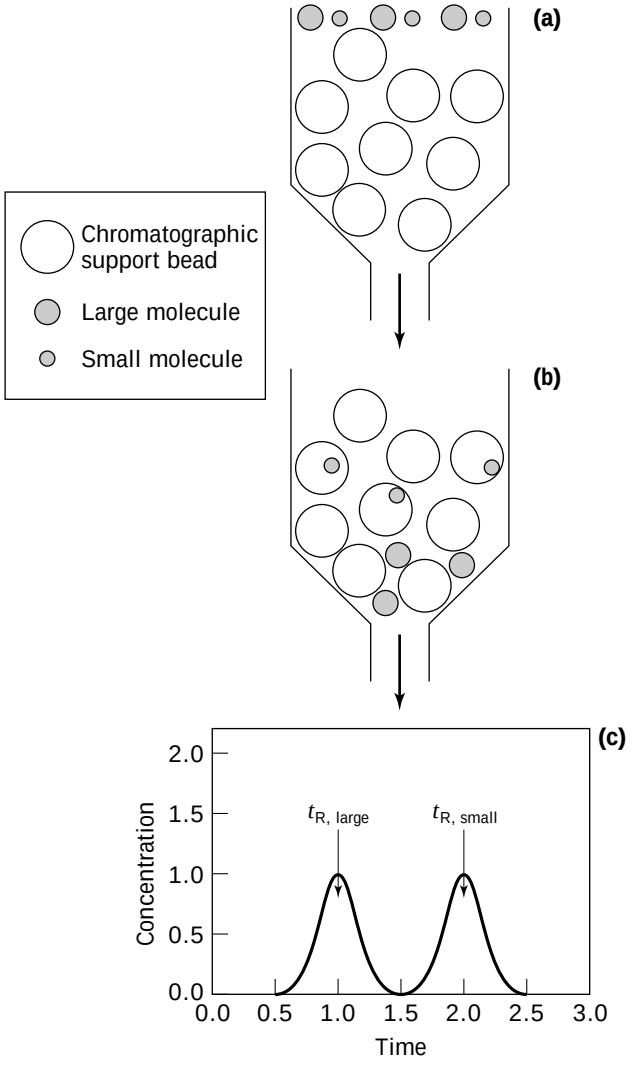


Figure 1. Principle of SEC. (**a**) A mixture of large and small particles is introduced together into the top of the column. (**b**) As the particles move down the column, the small particles are able to access the volume within the pores of the particles. Therefore, the larger particles are eluted from the column before the smaller ones. (**c**) Schematic chromatogram showing the elution profiles of the large and small particles.

regain of gel (grams of water per gram of dry gel divided by the density of water). Experimentally, V_p can be measured as the retention volume of a small molecule that does not interact with the support. With determination of V_0 and V_p, the distribution coefficient of any solute can be determined from its retention and equation 2; conversely, the retention time of a solute could be predicted from a model for its distribution coefficient (see "Process Variables") and from the column parameters.

The typical chromatographic peak is Gaussian in shape. The quality of separation of two (or more) species in SEC is characterized by differences in their retention times and by certain measures of the distribution of material in each peak about the mean. The resolution (R_s) is a measure of separation of two bands, defined by

$$R_s = \frac{t_{R2} - t_{R1}}{\frac{1}{2}(w_2 - w_1)} \quad (4)$$

where w_i is the baseline width of peak i, related to the standard deviation of the Gaussian peak through $w_i = 4\sigma_i$, and t_{Ri} is the retention time of the ith peak. Another measure of separation is the selectivity, α, between two bands. It is defined by

$$\alpha = \frac{k'_2}{k'_1} \quad (5)$$

where k′ of species i is its retention factor (or capacity factor). The retention factor is a relative retention time defined as

$$k' = \frac{t_R - t_0}{t_0} = \frac{V_R - V_0}{V_0} \quad (6)$$

For non-Gaussian bands, t_R and σ can be determined via statistical analysis of the chromatogram from the first and second moments, respectively, of a trace.

MATERIALS

Equipment

A typical laboratory setup for SEC is depicted in Figure 2. The equipment required includes solvent and reservoir, valves and tubing, pump, packed column, detector, fraction collector, and recording device. For the most part, this equipment is universal to the various modes of chromatography. The precolumn setup for SEC, especially at low pressure, is relatively straightforward. Solvent delivery is simpler in SEC than for other modes of chromatography because there is no need for gradient formation and the only pump requirement is for stable flow. Peristaltic pumps are adequate for low-pressure work, whereas piston or diaphragm pumps are used for high pressures (2).

The most important aspect of the equipment is the column packing. The gel must be packed uniformly throughout the column to avoid channeling or other heterogeneous effects that act to blur separation of the solutes. Frequently, the gel is supplied as a powder or as a suspension in a storage solvent. In the latter case, this solvent is filtered off, and in both cases the gel is then suspended in a packing solvent, which is usually an aqueous buffer with some NaCl added. A stable, uniform packing is attained by degassing the gel slurry before packing, pouring it at once, and then packing at a flow rate greater than used in normal operation.

Detection of proteins and nucleic acids is most commonly accomplished with UV absorbance detectors, although more sensitive techniques such as light-scattering and mass spectrometry are also possible (3). The detection limit (C^*) is dictated by the absorbance sensitivity (A^*) of the instrument, the extinction coefficient (ϵ) of the solute, and the path length (L) of the detector, according to

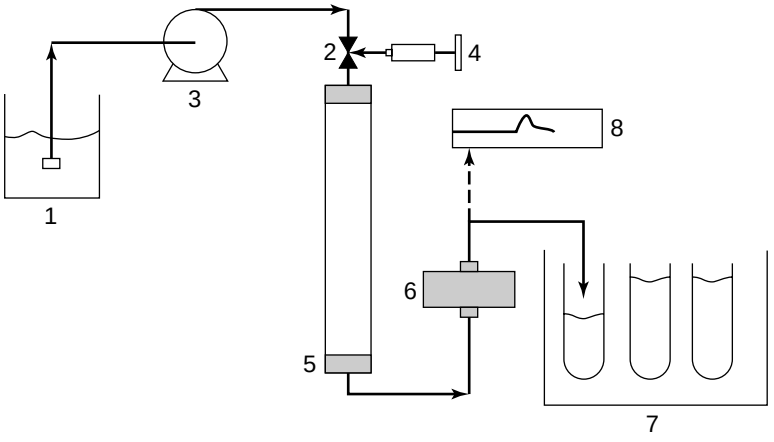


Figure 2. Laboratory SEC setup. Solvent (*1*) is pumped through a three-way valve (*2*) by a peristaltic pump (*3*). The desired volume of sample is introduced by syringe (*4*) through one port of the valve while the solvent flow is temporarily blocked. The valve is repositioned to allow elution of the sample though the packed column (*5*). The eluting solution is monitored by a detector (*6*) and collected in a fraction collector (*7*). A permanent record of the chromatogram is obtained from a recorder (*8*) or computer.

$$C^* > \frac{A^*}{\epsilon l} \tag{7}$$

The sensitivity of absorbance detectors is usually about 5×10^{-4}, and their path lengths can be up to 1 cm. Typical proteins have extinction coefficients on the order of 1 cm^2/mg; therefore, the detection limit is on the order of 1 μg/mL. Nucleic acids exhibit strong absorbance, peaking around 260 nm, with extinction coefficients on the order of 25 to 50 cm^2/mg; consequently, lower concentrations down to 20 ng/mL may be detected. Carbohydrates, on the other hand, exhibit no significant absorbance in UV or visible light; refractive index detectors are often used for their quantitation.

Packing Materials

A wide variety of packing materials are available for use in SEC. The choice of gel for a particular application depends on a number of factors, including materials compatibility, operating conditions, and particularly the sizes of the molecules to be separated. Manufacturers typically provide a good deal of information about their packing materials, including fractionation range, maximum flow rate or pressure drop, and stable operating ranges of pH and temperature.

Properties. The key feature of the packing material with respect to a particular separation is the pore size distribution, because this determines the range of particles that may be fractionated. Very small pores are required for desalting applications, so that all macromolecules are excluded and low molecular weight species are retained. Materials possessing much larger pores are used for analysis and purification of biological macromolecules. Because the pore sizes are normally inferred from the sieving of test solutes, the useful molecular weight fractionation range, rather than the actual pore size, is generally reported. It is important to note that different types of molecules exhibit different fractionation ranges; therefore, a fractionation range reported using protein standards will not be identical to that of, for example, nucleic acids. Some manufacturers report two fractionation ranges: one using protein standards and another using nucleic acids or dextrans. Usually, these two values will be sufficient to at least select an appropriate gel for the application of interest.

Once a fractionation range is determined, a material appropriate for the solutes of interest must be identified. Although the gel materials used in SEC have been developed to exhibit minimal adsorption, biological macromolecules inevitably interact to some extent with these materials. Proteins in particular will adsorb to a variety of surface functionalities. A general rule is that the gel material should be hydrophilic and uncharged. However, most hydrophilic materials exhibit at least a weak charge. Undesirable electrostatic interactions caused by this charge can be minimized by eluting with a buffer of moderate ionic strength, 0.3 to 0.5 M. In addition, the column must be operated under conditions at which the material is stable. Packing materials can deteriorate when subjected to extremes of pH, ionic strength, flow rate, and temperature. For some materials, enzymatic degradation is also possible in the event of bacterial contamination.

The size of the particles and their degree of compressibility determine the pressure drop across a bed and limit the flow rate that may be used. Particles ranging from 5 to 600 μm are available (4,5). Some SEC media, particularly cross-linked materials, are rigid over their normal operating range (Table 1). For these rigid materials, the pressure drop (Δp) across a gel of length (L) is inversely proportional to the square of the particle diameter (d_p) and proportional to velocity (v_0) (flow rate divided by cross-sectional column area), in accordance with Darcy's law:

$$\Delta p = \frac{\mu v_0 L}{k d_p^2} \tag{8}$$

where k is a constant and μ is the fluid viscosity. However, as the particles compress, the pressure drop increases dramatically, effectively setting an upper limit on the maximum attainable velocity. From the standpoint of materials selection, manufacturers supply information regarding particle size, fraction range, and maximum allowable flow.

Available Materials. Materials comprised of dextran, agarose, acrylamide, or combinations have traditionally been the workhorses of SEC. However, novel polymeric materials and polymer-coated silica media have been de-

Table 1. Packing Materials for SEC

Name	Manufacturer	Composition	MW fraction range (proteins)	DNA exclusion limit (bp)	Bead size (μm)	pH range	Maximum back pressure (MPa)	Maximum flow rate (cm/h)
Superose 12 prep grade	Pharmacia	Highly cross-linked agarose	1×10^3–3×10^5	150	20–40	3–12	0.7	30
Superose 6 prep grade	Pharmacia	Highly cross-linked agarose	5×10^3–5×10^6	450	20–40	3–12	0.4	30
Sephacryl S-100 HR	Pharmacia	Dextran/bisacrylamide	1×10^3–1×10^5	—	25–75	3–11	0.2	20–39
Sephacryl S-200 HR	Pharmacia	Dextran/bisacrylamide	5×10^3–2.5×10^5	118	25–75	3–11	0.2	20–39
Sephacryl S-300 HR	Pharmacia	Dextran/bisacrylamide	1×10^4–1.5×10^6	118	25–75	3–11	0.2	24–48
Sephacryl S-400 HR	Pharmacia	Dextran/bisacrylamide	2×10^4–8×10^6	271	25–75	3–11	0.2	33–63
Sephacryl S-500 HR	Pharmacia	Dextran/bisacrylamide	4×10^4–2×10^{7c}	1,078	25–75	3–11	0.2	24–48
Sephacryl S-1000 SF	Pharmacia	Dextran/bisacrylamide	5×10^5–1×10^{8e}	20,000	40–105	3–11	—	40
Sepharose 6B	Pharmacia	Agarose	1×10^4–1×10^6	194	45–165	4–9	2.8	14[a]
Sepharose CL-6B	Pharmacia	Cross-linked agarose	1×10^4–1×10^6	194	45–165	3–13	2.8	30[a]
Sepharose 4B	Pharmacia	Agarose	6×10^4–2×10^7	872	45–165	4–9	0.9	11.5[a]
Sepharose CL-4B	Pharmacia	Cross-linked agarose	6×10^4–2×10^7	872	45–165	3–13	1.7	26[a]
Sepharose 2B	Pharmacia	Agarose	7×10^4–4×10^7	1,353	60–200	4–9	0.6	10[a]
Sepharose CL-2B	Pharmacia	Cross-linked agarose	7×10^4–4×10^7	1,353	60–200	3–13	0.9	15[a]
Sephadex G-10	Pharmacia	Cross-linked dextran	<700	—	40–120	2–13	DL	2–5
Sephadex G-25	Pharmacia	Cross-linked Dextran	1×10^3–5×10^3	—	20–150	2–13	DL	2–5
Sephadex G-50	Pharmacia	Cross-linked dextran	1×10^3–3×10^4	—	20–80	2–10	DL	2–5
Sephadex G-75	Pharmacia	Cross-linked dextran	3×10^3–7×10^4	—	10–120	2–10	0.016[b]	18
Sephadex G-100	Pharmacia	Cross-linked dextran	4×10^3–1.5×10^5	—	10–120	2–10	0.0096[b]	12–50
Sephadex G-150	Pharmacia	Cross-linked dextran	5×10^3–3×10^5	—	10–120	2–10	0.0036[b]	6–23
Sephadex G-200	Pharmacia	Cross-linked dextran	5×10^3–6×10^5	—	10–120	2–10	0.0016[b]	3–12
Superdex 75 HR 10/30	Pharmacia	Dextran/Cross-linked agarose	3×10^3–7×10^4		13–15	3–12	1.8	110
Superdex 200 HR 10/30	Pharmacia	Dextran/cross-linked agarose	1×10^4–6×10^5	200	13–15	3–12	1.5	76
Bio-Gel P-2	Bio-Rad	Polyacrylamide	1×10^2–1.8×10^3	—	45–90	2–10	—	5–10[d]
Bio-Gel P-4	Bio-Rad	Polyacrylamide	8×10^2–4×10^3	—	45–180	2–10	—	10–20[d]
Bio-Gel P-6	Bio-Rad	Polyacrylamide	1×10^3–6×10^3	5	45–180	2–10	—	10–20[d]
Bio-Gel P-10	Bio-Rad	Polyacrylamide	1.5×10^3–2×10^4	—	45–180	2–10	—	10–20[d]
Bio-Gel P-30	Bio-Rad	Polyacrylamide	2.5×10^3–4×10^4	20	45–180	2–10	—	6–13[d]
Bio-Gel-P-60	Bio-Rad	Polyacrylamide	3×10^3–6×10^4	55	45–180	2–10	—	3–6[d]
Bio-Gel P-100	Bio-Rad	Polyacrylamide	5×10^3–1×10^5	—	45–180	2–10	—	3–6[d]
BioGel A 0.5m	Bio-Rad	Agarose	$<1 \times 10^4$–5×10^5	—	38–300	4–13	—	7–20[e]
BioGel A 1.5m	Bio-Rad	Agarose	$<1 \times 10^4$–$1.5^1 0^6$	—	38–300	4–13	—	7–20[e]
BioGel A 5m	Bio-Rad	Agarose	1×10^4–5×10^6	—	38–300	4–13	—	7–20[e]
BioGel A 15m	Bio-Rad	Agarose	4×10^4–1.5×10^7	—	38–300	4–13	—	7–20[e]
BioGel A 50 m	Bio-Rad	Agarose	1×10^5–5×10^7	350	75–300	4–13	—	5–25[e]
Macro-Prep SE 100/40	Bio-Rad	Cross-linked Agarose	5×10^3–1×10^5	—	40	1–14	4	240
Macro-Prep SE 1000/40	Bio-Rad	Cross-linked Agarose	1×10^4–1×10^6	—	40	1–14	3	240
Hydropore-5-SEC	Varian/Rainin	Polymer-coated silica	5×10^3–1×10^6	—	5	—	—	—
Toyopearl HW-40	TosoHaas	Poly(ethylene glycol/ methacrylate)	1×10^2–1×10^4	—	20–60	2–12	0.3	DL
Toyopearl HW-50	TosoHaas	Poly(ethylene glycol/ methacrylate)	5×10^2–8×10^4	—	20–60	2–12	0.3	DL
Toyopearl HW-55	TosoHaas	Poly(ethylene glycol/ methacrylate)	1×10^3–2×10^5	—	20–60	2–12	0.3	DL
Toyopearl HW-65	TosoHaas	Poly(ethylene glycol/ methacrylate)	4×10^4–5×10^6	—	20–60	2–12	0.3	DL
Toyopearl HW-75	TosoHaas	Poly(ethylene glycol/ methacrylate)	$5 \times 10^5 \times 10^7$	—	30–60	2–12	0.3	DL
TSK-GEL G2000SW	TosoHaas	Bonded silica	5×10^3–1.5×10^5	55	4–13	2.5–7.5	1–12	DL
TSK-GEL G3000SW	TosoHaas	Bonded silica	1×10^4–5×10^5	110	4–10	2.5–7.5	1–12	DL
TSK-GEL G4000SW	TosoHaas	Bonded silica	2×10^4–1×10^7	375	8–17	2.5–7.5	1–3.5	DL
TSK-GEL G1000PW	TosoHaas	Poly(ethylene glycol/ methacrylate)	$<2 \times 10^3$	—	10	2–12	2	DL
TSK-GEL G2000PW	TosoHaas	Poly(ethylene glycol methacrylate)	$<5 \times 10^3$	—	10–20	2–12	2–4	DL
TSK-GEL G2500PW	TosoHaas	Poly(ethylene glycol/ methacrylate)	$<8 \times 10^3$	—	6–20	2–12	2–4	DL

Table 1. Packing Materials for SEC (*continued*)

Name	Manufacturer	Composition	MW fraction range (proteins)	DNA exclusion limit (bp)	Bead size (μm)	pH range	Maximum back pressure (MPa)	Maximum flow rate (cm/h)
TSK-GEL G3000PW	TosoHaas	Poly(ethylene glycol/ methacrylate)	5×10^2–8×10^5	—	6–20	2–12	2–4	DL
TSK-GEL G4000PW	TosoHaas	Poly(ethylene glycol/ methacrylate)	1×10^4–1.5×10^6	—	10–22	2–12	1–2	DL
TSK-GEL G5000PW	TosoHaas	Poly(ethylene glycol/ methacrylate)	$<1 \times 10^7$	1,500	10–22	2–12	1–2	DL
TSK-GEL G6000PW	TosoHaas	Poly(ethylene glycol/ methacrylate)	$<2 \times 10^8$	—	13–25	2–12	1–2	DL
TSK-GEL GMPW	TosoHaas	Poly(ethylene glycol/ methacrylate)	$<2 \times 10^8$	—	13–17	2–12	1–2	DL
TSK-GEL G-Oligo-PW	TosoHaas	Poly(ethylene glycol/ methacrylate)	$<3 \times 10^3$	—	6	2–12	4	DL
TSK-GEL G-DNA-PW	TosoHaas	Poly(ethylene glycol/ methacrylate)	$<2 \times 10^8$	7,000	10	2–12	2	DL

Note: Properties described are as stated by the manufacturers. In many cases, several sizes of particles are available for a given material; these are indicated as a range. DL indicates that the material obeys Darcy's law (i.e., flow rate proportional to pressure drop per bed length) over the normal operating range.
[a]Determined using distilled water at room temperature on a 2.5- by 30-cm column.
[b]Determined using distilled water at room temperature on a 2.6- by 30-cm column.
[c]The indicated fractionation range was determined for dextrans rather than proteins.
[d]Typical flow rates determined using a 1.5- by 70-cm column.
[e]Typical flow rates determined using a 1.5- by 20-cm column.

veloped. Silica possesses excellent mechanical strength, but it is prone to adsorption of biomolecules. Materials are available from a number of manufacturers, and many are summarized in Table 1. Although this list is not exhaustive, it represents many of the commonly used materials, and it is notable that these have not changed significantly over the past 15 years (5–7).

Dextrans are used in a wide variety of protein and nucleic acid separations. They are polysaccharides that are made into a porous gel by cross-linking with epichlorohydrin (8). These materials are quite stable for pH $>$ 2 and for temperatures up to at least 120 °C and can be used with either aqueous or organic solvents. Dextran gels with small pore sizes exhibit low compressibility, but versions with larger pore sizes are quite compressible (9). As a result, some dextran gels are limited to flow rates that can be less than 1 mL/min. Because dextrans are polysaccharides, they are susceptible to bacterial degradation; however, because of their thermal stability, they can be autoclaved. A variation of the traditional dextran gel is provided by cross-linking dextran with *N,N′*-methylenebisacrylamide. This provides a material that is substantially more rigid but has a smaller stable pH range and is more susceptible to adsorption of proteins. The latter obstacle can often be overcome by eluting with a higher ionic strength solvent.

Acrylamide gels are formed by cross-linking acrylamide monomer with *N,N′*-methylenebisacrylamide. The resulting matrix is quite stable over the pH range from 2 to 10 and to high temperatures and is not prone to bacterial degradation. It is less adsorptive than the corresponding dextran material. Consequently, polyacrylamide gels are commonly used for protein separations. Their main disadvantage is that they are incompatible with organic solvents.

A copolymer of D-galactose and 3,6-anhydro-L-galactose is used to form an agarose gel. The composition and degree of cross-linking have a profound effect on the properties of agarose-based media. With less cross-linking, the gels are typically stable only in the pH range of 4 to 9, are compressible at flow rates as low as 1 mL/min, and must be sterilized chemically, but they exhibit little adsorption of macromolecules and broad fractionation ranges that make them much more suitable than acrylamide or dextran for the separation of large proteins and nucleic acids. With a higher degree of cross-linking, the material is stable over a wider range of pH and flow rate, and it can be autoclaved; however hydrophobic interactions may occur between proteins and the support, and the available fractionation range is reduced.

A number of variations on traditional materials have been established. Generally, these are produced to combine the desirable sieving properties of one material with the mechanical or chemical stability of another. The three most commonly used materials—dextran, acrylamide, and agarose—have been combined in all possible pairs. The dextran–bisacrylamide material (Sephacryl) mentioned earlier is one such example. Another is a composite of dextran and cross-linked agarose, covalently bound together to form a material (Superdex, [9]) that has the fractionation range of a dextran support with the rigidity of agarose. The resulting material is autoclavable and stable in the range 3 $<$ pH $<$ 12 and to a variety of strong solvents. Acrylamide has also been entrapped in an agarose gel to form the commercial Ultrogel matrix (7).

Other supports, based on materials such as silica and hydrophilic polymers, have been produced for SEC (5). Although porous silica is inherently rigid and can be manufactured in a variety of pore sizes, it exhibits marked ad-

sorption of proteins and is unstable at basic pH. Other polymer gels exist commercially, such as Toyopearl and Spheron, which are based on polyethers and poly(2-hydroxyethylmethacrylate), respectively. The latter is a material approved by the Food and Drug Administration and can thus be used to produce biocompatible supports.

APPLICATIONS

A number of applications exist for SEC in the research laboratory and in industrial processes. These can be divided into the two categories of sample analysis and purification. The former is most often performed on small columns with high resolution, and the latter is a bulk operation in which the separations are usually fairly crude. A summary of some examples of SEC applications is provided in Table 2.

Analysis

At the analytical level, SEC is used as a tool in characterizing the size, shape, and interactions of biological macromolecules. For example, SEC has been used to study the kinetics and equilibrium dimerization of proteins (10,11). Likewise, it can be used to study the hybridization of nucleic acids (12). Similar applications have been made to reversible interactions between macromolecules and low molecular weight species (13,14). It has also been used to identify and characterize molten globule states in studies of protein unfolding (15). In addition, it is an excellent tool for identification and separation of small molecules that are conjugated to biomolecules, such as probes attached to antibodies or drugs conjugated to targeting molecules (16–18).

SEC is often used for molecular weight determination, although this is an approximate method. Molecules that are similar to the solute of interest but with known molecular weight are used to create a calibration curve of retention time versus molecular weight. The unknown molecular weight of the solute is then estimated from its retention time. The degree of success of this approach is highly dependent on how similar the unknown solute is to the molecular weight standards. That is, a calibration curve prepared with globular proteins will give an inaccurate molecular weight for an oligonucleotide or even for a glycoprotein. Nonetheless, for similar molecules, molecular weights can be successfully estimated. For example, the retention times of a number of globular proteins correlate well with log(MW), as shown in Figure 3. Under the conditions of Figure 3, the retention time of α_1-acid glycoprotein is 8.70, from which an apparent molecular weight of 57 kDa is calculated, significantly greater than its actual molecular weight of 40 kDa.

For molecules of similar chemistry, such as globular proteins, the deviation of the observed elution volume from the calibration curve can be an indication of molecular shape. Deviation from spherical to ellipsoidal shape results in increased molecular friction, and thus the molecule sieves as if its apparent molecular weight is greater than its actual value. Although exact molecular weight determination is complicated by the shape and chemical nature of different protein molecules, changes in a particular structure can be discerned with high resolution. For instance, the extent of modification of a protein by a dye or other ligand can be monitored by the relatively small decrease in retention time if the same column, buffer, and operating conditions are used (16).

Separations

By design, SEC separates molecules based on their size, so it is not possible to achieve well-resolved separations of biomolecules from a mixture of similar species using this technique. Although fair discrimination of molecules according to molecular weight was seen in the previous section, small differences in retention time are insufficient for resolution in purification (see "Dependence of Separation Parameters on Operating Variables"). Nonetheless, crude fractionation of a particular size range is possible with

Table 2. Examples of SEC Applications

Analysis
Molecular weight estimation
Identification and analysis of protein dimers
Analysis of reversible protein–small molecule interactions
Studies of protein unfolding
Identification of products in molecular conjugates
Evaluation of the stability of macromolecular complexes
Separations
Desalting
Crude protein separations
Purification of molecular conjugates
Removal of endotoxins and viruses from pharmaceutical formulations

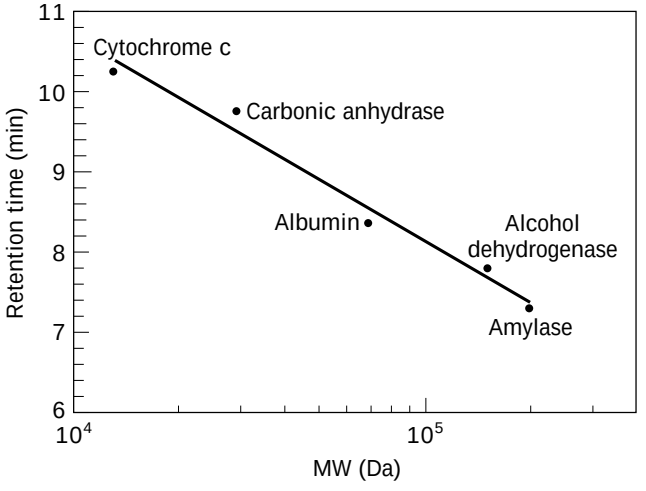


Figure 3. Calibration curve of an analytical SEC column. The elution volumes of several globular proteins were determined on a TSK-GEL G3000SW$_{XL}$ column, 7.8 mm i.d. and 30 cm length. Twenty microliter aliquots (concentration approx. 1 mg/mL) were injected and elution was performed in a buffer of 0.1 M KH_2PO_4 + 0.1 M Na_2SO_4, pH 7. The eluate was monitored by UV absorbance detection at 280 nm.

SEC. Thus, it is best suited either for purification of one molecule of interest from a small set of dissimilar contaminants in the laboratory or as an early or late step in conjunction with other steps in an industrial process.

SEC finds many laboratory applications in manipulations of biological molecules. For example, unreacted species in molecular conjugates are often removed via SEC (18). Cleanup of materials involved in the synthesis of peptides and oligonucleotides is also performed frequently via SEC (19). It is also used for removal of aggregates resulting from a variety of biochemical preparations. For example, an intermediate or final step in purification of a monoclonal antibody is often removal of dimers and aggregates that have been generated as a result of prior processing. In the production of recombinant proteins, a certain fraction may be clipped by proteolysis; these smaller molecules can be separated by SEC (20).

The most common application of SEC at both laboratory and process scales is desalting or buffer exchange. Many processing steps (e.g., extraction, crystallization, other modes of chromatography) involve the use of a solvent that is undesired for future processing or formulation steps. Very small pore SEC columns are used to replace the undesired buffer by the desired one. Usually, a wide disparity in molecular sizes exists between the biological solute and the buffer salts, so short columns can be used, limiting the dilution to twofold or less (21).

PROCESS VARIABLES

Selection of a support material for an application of interest is only the first step in performing an analysis or separation. The amount of sample to be analyzed or purified must be considered. The column length and total amount of material to be injected into the column must be determined. Finally, a flow rate must be found that will provide adequate throughput and resolution. In the process of designing a configuration, convective dispersion, transport of the solutes to and within the particles, and partitioning of the solutes between the mobile and gel phases must all be considered. For these reasons, a substantial body of literature has been devoted to the engineering issues associated with SEC. The most general and applicable models are discussed in this section.

Plate Theory

An operational description of chromatography can be obtained from plate theory, which is a phenomenological description of zone spreading in chromatography, first proposed by Martin and Synge in 1941 (22). The essential concept is that as the solute traverses the column, it undergoes repeated equilibration with the stationary phase, so that each equilibration defines one stage, or plate. The column length over which equilibration occurs is known as the height of an equivalent theoretical plate (HETP), or simply plate height. The development of a solute peak zone is shown schematically in Figure 4. In each stage, fluid in the mobile phase is mixed, and partitioning occurs between the stationary and mobile phases. A mass balance on a particular stage takes the form

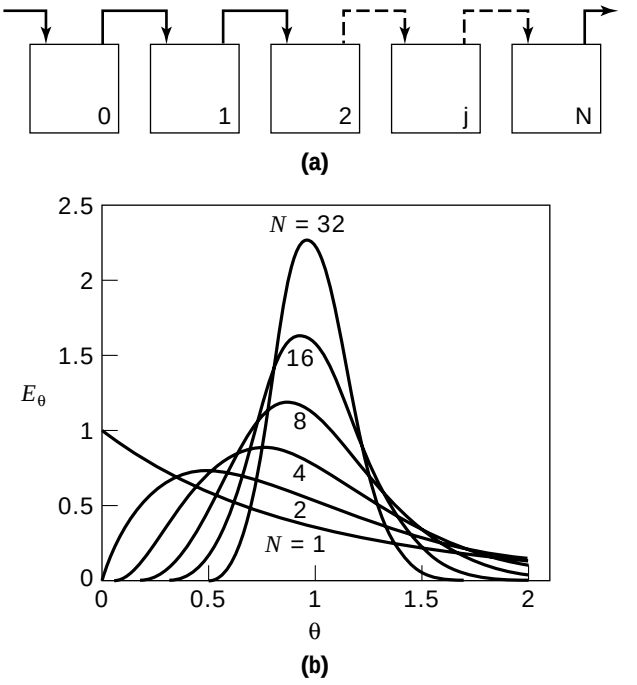


Figure 4. (**a**) Plate theory. The progression of solute through the column can be viewed as passing through a number of stages. In each stage, fluid in the mobile phase is mixed and partitioning occurs between the stationary and mobile phases. (**b**) Residence time distribution from N successive plates. The residence time distribution E, which is the output from a pulse input, in terms of the dimensionless variable $\theta = t/\tau$, is shown as a function of θ for various numbers of plates (N). It can be seen that as the number of plates exceeds approximately 10, the residence time distribution begins to assume a Gaussian shape, which is enhanced as N increases further.

$$Q(C_{j-1} - C_j) = v_m \frac{dC_j}{dt} + v_p \frac{d\bar{C}_{p,j}}{dt}, \qquad (9)$$

where Q is the flow rate, C_j is the concentration of solute in the jth stage, v_m is the mobile phase volume per stage, v_p is the pore volume per stage, and $\bar{C}_{p,j}$ is the average pore concentration in the jth stage. For linear chromatography, the average pore concentration is proportional to the bulk concentration, and equation 9 becomes directly analogous to the mass balance on stirred tanks in series (23).

The input to a chromatography column is essentially a pulse; consequently, the output from a series of N stages is analogous to the residence time distribution (E) from N tanks in series. The output from a single stage is an exponential distribution, but, as the number of plates increases, the residence time distribution takes on a Gaussian shape, as shown in Figure 4. It can be shown that the variance σ^2 of the output peak is dependent only on the retention time and number of stages through

$$\sigma^2 = t_R^2/N \qquad (10)$$

Therefore, the number of plates (N) in chromatography is defined as

$$N = t_R^2/\sigma^2 \quad (11)$$

and the plate height (H) as the length of the column (L) divided by the number of plates

$$H = L/N = L\sigma^2/t_R^2 \quad (12)$$

Relation of Size to Sieving

Because the purpose of SEC is to separate molecules on the basis of their size, interest developed from the time of its introduction to understand quantitatively the relationship between the size of a solute molecule and its retention volume. This is of importance in the use of SEC for molecular weight determination, in its interpretation for applications involving monitoring changes in molecular structure, and in the selection of SEC media and operating conditions to perform a separation.

Volume Exclusion. The primary mechanism of sieving is based on the excluded volume of particles in the pores. At equilibrium, the solute molecules will attain a concentration within the pore based on the volume accessible to the center of the molecule. When the pore size is of the same order as the solute molecular size, the accessible volume becomes significantly less than the pore volume (Fig. 5). As a result of the excluded volume effect, even a monodisperse pore size distribution will produce a continuum of distribution coefficients with increasing molecular size.

To develop this concept into a quantitative model, a geometry is usually assumed. An early result was that of Laurent and Killander (24), who used the expression of Ogston (25) for the available volume in a network of rigid rods to develop an expression for the distribution coefficient:

$$K_d = \exp[-\pi L(r_s + r_r)^2] \quad (13)$$

where L is the total length of rods per volume, r_s is the radius of the (assumed) spherical particles, and r_r is the radius of the rods. The parameters L and r_r are generally not known and must be fit to experimental data. Other expressions were derived for cylindrical, conical, and slit-like pores (26). The problem of ascribing a particular pore geometry was circumvented by Ackers (27), who represented a pore by the maximum characteristic radius particle that could pass through it and postulated that the distribution of pore sizes is distributed normally. For this model, the distribution coefficient takes the form

$$K_d = \text{erfc}\left[\frac{a - a_0}{b_0}\right] \quad (14)$$

where a is the radius of the solute particle, a_0 is effectively an average pore radius, and b_0 is a measure of the variance of the distribution. These parameters must still be determined for each gel from sieving experiments on at least two (preferably more) solutes.

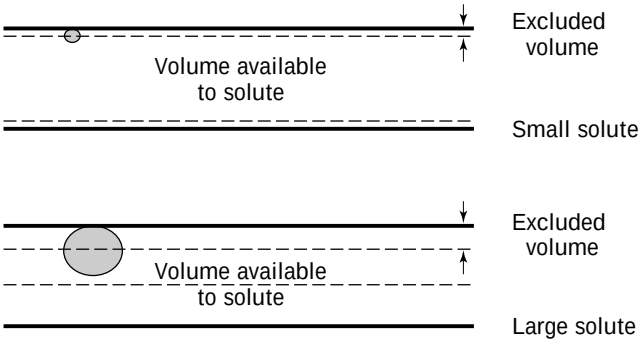


Figure 5. Volume exclusion of solutes in pores. Solutes assume concentrations within pores of varying sizes according to the volumes accessible by the centers of its molecules.

Rates of Pore Diffusion. The models based on volume exclusion (see previous section) are entirely thermodynamic; that is, they do not include any transport limitations and do not predict a dependence of distribution coefficient on flow rate. In practice, a slight dependence is observed, which led to the development of theories that account for transport of solute molecules into the pores. An approach to accounting for diffusion is to model the distribution of molecules between the mobile and gel phases using the one-dimensional (1-D) unsteady-state diffusion equation (28), which has the solution

$$C_g = C_m \text{ erfc}\left[\frac{x}{2(Dt_d)}^{1/2}\right] \quad (15)$$

where C_g and C_m are concentrations in the gel and mobile phases, respectively, x is distance into the 1-D gel phase, t_d is the diffusion time, and D is the diffusion coefficient of the solute. With t_d inversely proportional to the mobile phase velocity (v_0) and b representing the power to which diffusion coefficient is inversely proportional to molecular weight (M), the distribution coefficient can be determined to be (28)

$$K_d = \frac{k}{(\pi v_0 M^b)^{1/2}}(1 - e^{-v_0 M^b/k^2}) + \text{erfc}\left[\frac{(v_0 M^b)^{1/2}}{k}\right] \quad (16)$$

where k is a grouping of constants fit to experimental data.

Purely Empirical Correlations. The equations initially used to correlate SEC data were essentially empirical. The most commonly used forms are

$$V_R = A - B \log(M) \quad (17)$$

and

$$\log(V_R) = A' - B' \log(M) \quad (18)$$

where M is the molecular weight of the solute and A, B, A', and B' are empirical constants for the particular column configuration. As shown in Figure 3 and discussed in "Analysis", very good correlations of data for similar species are possible, but a priori prediction or extrapolation beyond the measured range of molecular weights or to mol-

ecules of different chemistry is problematic. Note that the more sophisticated theoretical models discussed earlier also make use of essentially empirical constants that require calibration of the column using proteins (or other molecules) of known molecular weight. Calibration of the column must be repeated if the pH, ionic strength, or temperature is changed significantly or if the column is repacked. Furthermore, the calibration standards must be well characterized and of a similar nature to the solutes of interest. Nonetheless, the use of theoretical models attaches a physical significance to the parameter's fit from experimental data and provides some insight into the relative contributions of the volume-exclusion versus pore-diffusion mechanisms (28).

Transport Issues

The previous section dealt with the retention time, which mostly results from the equilibrium partitioning of species into the gel support. For separations, it is necessary to consider that chromatography is conducted in a packed bed, under flow, sometimes at high pressures. The configurational aspects of the chromatography column lead to band broadening, or spreading of the eluted solute band. The main contributions to band broadening are eddy diffusion, axial dispersion, and mass transfer resistance at the surface of the particles and within the pores. Eddy diffusion results from the tortuosity of extraparticle paths that solute particles take from the top to the bottom of the column. Axial dispersion is essentially molecular diffusion along the length of the column. The mass transfer resistance at the surface arises from the boundary layer, or film, through which solute must pass to reach the gel particle, and the mass transfer resistance of the pores is due to diffusion of solute into them and is enhanced by their nonuniformity.

A chromatography column is a packed bed of gel particles and is subject to the same types of scaling analysis as for packed beds in other chemical processes. Most of the resistance to mass transfer is at the level of the particle; consequently, the particle diameter is the characteristic length for scaling. Hence, a reduced velocity or Péclet number, *(Pe)* can be defined in terms of the particle diameter as

$$Pe = \frac{d_p u}{D_0} \tag{19}$$

where u is the interstitial velocity of the fluid and D_0 is the bulk solute diffusion coefficient. Likewise, the reduced plate height (h) is

$$h = \frac{H}{d_p} \tag{20}$$

where H is the plate height defined in equation 12.

The relation of interest is that between plate height and flow rate (or velocity) or, equivalently, between reduced plate height and Péclet number. This is obtained from solution to the governing differential mass balance equations in lesser or greater complexity (29–31).

Dependence of Separation Parameters on Operating Variables

Analysis of mass transport in chromatography columns has made it possible to make quantitative predictions regarding plate heights as a function of flow rate in SEC (32). The original analysis of mass transport was conducted by van Deemter (29). His equation has served as the basis for all subsequent analyses:

$$h = 2\lambda + \frac{2\gamma}{Pe} + CPe \tag{21}$$

where λ, γ, and C are constants accounting for eddy dispersion, axial diffusion, and mass transfer resistance at the particle surface, respectively. This equation predicts a minimum plate height with respect to increasing fluid velocity. More detailed accounting of the mass transfer contributions, including that caused by intraparticle diffusion, have produced alternative expressions for the plate height, for example, the expanded form of Horváth and Lin (33):

$$h = \frac{2\gamma}{Pe} + \frac{2\lambda}{1 + \omega Pe^{-1/3}} + \frac{\kappa k_0^2}{(1 + k_0)^2} Pe^{2/3} + \frac{\theta k_0}{30(1 + k_0)^2} Pe \tag{22}$$

where θ is a geometric factor for a given particle, λ, γ, ω, and κ are parameters dependent on the details of the particular column packing, and k_0 is the ratio of intraparticle to interstitial volumes available to the solute of interest. The third and fourth terms represent mass transfer resistances at the particle surface and within the pores, respectively.

For liquid chromatography involving macromolecules, the contributions from axial dispersion and resistance at the particle surface are small, and the largest contribution is from pore diffusion. Consequently, the minimum observed by van Deemter for GLC (29) is typically not observed in SEC of biomolecules, and the plate height is roughly proportional to velocity (32,34,35). Plate height also depends on particle size:

$$H \sim \frac{u d_P^2}{D_0} \tag{23}$$

From this approximate equation, two practical points can be inferred. First, because plate height is a measure of zone spreading and should generally be minimized, smaller particles are preferred because of the strong dependence of plate height on diameter. However, pressure drop is inversely proportional to the square of particle diameter, so the mechanical properties of the material limit the extent to which particle size can be reduced. Second, lower velocities, that is, slower flow rates, will result in less zone spreading. This advantage may be offset where the speed of a separation is important. Furthermore, the flow rate cannot be reduced too drastically or the effects of axial diffusion will become limiting.

An important measure of the separation is the resolution between two or more solutes to be separated. The de-

pendence of the resolution on column variables can be determined from the equations already developed. Combining equations 4, 6, and 12 gives an expression for resolution in terms of capacity factors and plate heights (36):

$$R_s = \frac{L^{1/2}}{2} \frac{k_2' - k_1'}{H_2^{1/2}(1 + k_2) + H_1^{1/2}(1 + k_1')} \qquad (24)$$

Employing the approximation that the plate heights of the two solutes are equal, an expression for resolution as a function of retention, selectivity, and column efficiency can be derived by combining the definition of selectivity (equation 5) with equation 24:

$$R_s = \frac{1}{2}\left(\frac{\alpha - 1}{\alpha + 1}\right) \frac{\bar{k}'}{1 + \bar{k}'} \frac{L^{1/2}}{H^{1/2}} \qquad (25)$$

where $\bar{k}' = (k_1' + k_2')/2$ is the average retention time of the solutes.

Equation 25 expresses resolution as a function of three terms representing selectivity, retention, and zone spreading, respectively. Selectivity is the most important determinant of resolution, up to a point. An increase in α from 1.1 to 1.2 results in an 83% increase in resolution, all other factors being equal; however, an increase from 2.1 to 2.2 results in only a 6% further increase in resolution. Selection of a column best suited to fractionate the species of interest is thus critical to obtaining high resolution. Because the capacity factor is greater than or equal to unity, the middle term takes values only between 0.5 and 1; nonetheless, some improvement in resolution is seen to be possible by increasing the retention time. Because of the weak dependence of resolution on length, significant increases in column length are often required to effect a modest improvement in resolution. The inverse relationship between resolution and plate height suggests that minimizing zone spreading by the methods discussed earlier—reducing particle diameter or flow rate—will concomitantly improve resolution. This is true, subject to the same limitations on flow rate and pressure drop. Furthermore, if speed of separation is important, a trade-off exists.

In addition to the particle size and composition, column length, and operating flow rate, the SEC practitioner must choose column diameter, sample volume, and buffer composition. Column diameter does not appear explicitly in the equations describing resolution, but they are derived with the assumption of a uniform packing, which is increasingly difficult to achieve as the diameter-to-length ratio increases. Therefore, columns used for fractionation are prepared with a length-to-diameter ratio of about 10:1. But in industrial scale desalting applications, where the selectivity is great and resolution not an issue, much wider columns can be used. Sample volumes are usually chosen to be less than 5% of the total column volume, because values greater than a couple of percent tend to produce significant increases in zone spreading. In principle, the solute molecules should not interact with the stationary phase in SEC, but in practice biological molecules exhibit some interaction with a variety of materials, including gels used for SEC (37). Variation of buffer composition can be used to modulate adsorptive interactions. For example, adsorption is often mediated by charge interactions between solute and sorbent and can be reduced by increasing the ionic strength of the buffer, which provides double-layer screening of the electrostatic interactions.

CONSIDERATIONS FOR LARGE SCALE

SEC is used in large-scale industrial processes for the production of proteins for agricultural and pharmaceutical use. At this scale, economic considerations become imperative, and in particular it is important for the chromatographic operation to achieve a high level of productivity and throughput while maintaining high resolution.

General Strategies

For large-scale operation, the throughput required necessitates reuse of chromatography columns in a repeated-batch operation. Usually, the primary goal is to maximize productivity, which is defined as the amount of purified material per column area per time. This can be achieved by maximizing flow rate and minimizing the amount of time during which the column is not actively performing a separation. From a productivity standpoint, the time between elution of the valuable solute and all others is wasted time. One route to improving productivity is to reduce the resolution of the solutes to the minimum necessary (Fig. 6). Likewise, the time after elution of all solutes before the next injection can be used more efficiently by applying the next sample before the column volume of the present run is finished (21). These general guidelines govern both desalting and fractionation; further optimization depends on the particular application.

Another strategy to maximize throughput is to maximize the sample size. However, dispersion effects are enhanced for large sample volumes. Consequently, for fractionation applications, it is advisable to use sample sizes less than 3% of the total column volume (38). Although it is often desirable to have highly concentrated samples, higher concentrations result in higher viscosities, which result in slower diffusion into and equilibration with the gel. The requirement of sample equilibration with the gel also limits the flow rate that may be applied.

For desalting, on the other hand, larger sample volumes can be applied, because the sample elutes in the void volume and the salts are retained in the pore. Because of this large discrepancy in elution volumes, the increased dispersion resulting from large sample sizes is not much of an issue, and sample sizes as great as 25 to 30% of the column volume may be used. Furthermore, small molecules (e.g., ions) equilibrate quickly with the pores because of their small excluded volume and high diffusion coefficients, and macromolecules are completely excluded, so high flow rates may be used without compromising the separation.

Scale-Up Guidelines

The principle of scale-up is to determine the values of new process variables that will allow a desired change in one

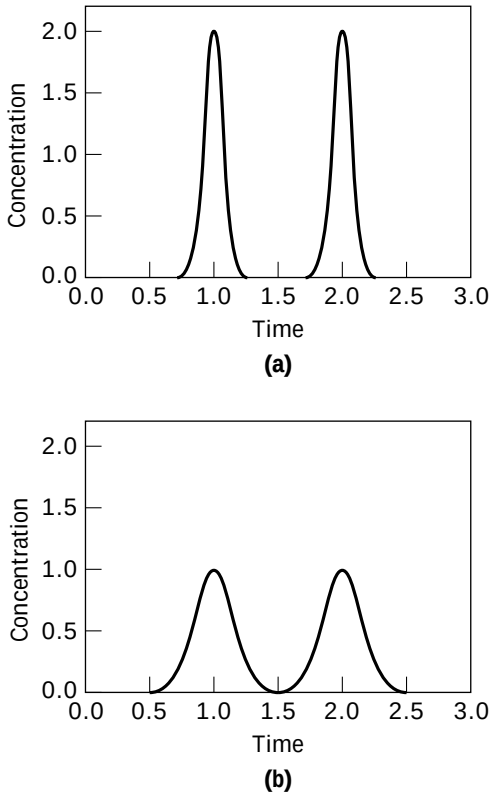


Figure 6. Optimization of resolution for large-scale SEC. (**a**) A small injection of solutes at low flow rate results in well-resolved bands. (**b**) The productivity is increased by increasing flow rate, increasing sample size, or decreasing column length to the point where the solutes are just resolved.

process variable while keeping others constant. Generally, it is desired to increase the throughput while maintaining the same resolution, but sometimes it is necessary to change resolution while keeping throughput constant. In any case, the key is to have a description of the relationship among process variables. It is here that the analysis of the preceding section is most useful.

Scaling rules for isocratic elution chromatography, which includes both SEC and adsorption chromatography, have been described by Wankat and Koo (36). The starting point is the definition of scaling parameters describing the ratios of process variables in the new and old designs, that is,

$$\delta = \frac{d_{p,new}}{d_{p,old}}$$
$$\lambda = \frac{L_{new}}{L_{old}}$$
$$\rho = \frac{D_{new}}{D_{old}}$$
$$\pi = \frac{\Delta p_{new}}{\Delta p_{old}}$$
$$\theta = \frac{Q_{new}}{Q_{old}} \qquad (26)$$

where D is the column diameter and Q is the buffer flow rate. The capacity factor in SEC is a thermodynamic quantity representing the partitioning of solutes in the pores of the support particles and therefore should not vary with column configuration or process variables. Therefore, the description of column resolution in equation 24 can be used for both the new and old designs to define a ratio of design resolutions, Ψ (36):

$$\Psi \equiv \frac{R_{s,new}}{R_{s,old}} = \left(\frac{L_{new}}{L_{old}}\right)^{1/2} \frac{[H_2^{1/2}(1 + k_2') + H_1^{1/2}(1 + k_1')]_{old}}{[H_2^{1/2}(1 + k_2') + H_1^{1/2}(1 + k_1')]_{new}} \qquad (27)$$

The difference between plate heights of the different solutes is usually ignored (see equation 25) and equation 27 is greatly simplified, that is,

$$\Psi = \lambda^{1/2}\left(\frac{H_{old}}{H_{new}}\right)^{1/2} \qquad (28)$$

As discussed in "Dependence of Separation Parameters on Operating Variables", transport resistances at the particle surface and within the pores usually dominate, with the result that the plate height can be represented by a scaling of the form equation 21. Substituting in equation 28 gives a simple design equation:

$$\Psi = \left(\frac{\lambda}{\theta}\right)^{1/2} \frac{\rho}{\delta} \qquad (29)$$

Note that more complete plate height expressions (derived from either theoretical or empirical correlations) may be substituted into equation 28 to develop more sophisticated design equations.

Equations such as equation 29 can be used to derive new sizes and operating conditions for a given change in any of the others. As a simple example, suppose that a design involving a reduction in particle size is desired, while maintaining constant column dimensions and throughput. In this case, λ, ρ, and θ are equal to unity, and the resolution is inversely proportional to the particle size

$$\Psi = \frac{1}{\delta} \qquad (30)$$

Alternatively, the objective of reducing particle size may be to allow a greater flow rate while maintaining constant resolution. In this case, Ψ, ρ, and λ are equal to unity, and the flow rate scales as the inverse square of particle diameter

$$\theta = \frac{1}{\delta^2} \qquad (31)$$

However, this is only valid in the regime where the particles are rigid, and it may not be feasible because of the large increase in pressure drop that would result.

The application of scaling rules to the interrelationship of process variables in equation 24 is an idealization that is subject to a number of constraints. Arbitrary column diameters and lengths are not possible, and certain length-

to-diameter ratios (on the order of 10:1) are desirable to achieve a uniform packing. Also, only certain particle sizes are available, and the range in which they are manufactured is dependent to a large degree on their resistance to flow (i.e., pressure drop). As discussed in "Packing Materials", many packings are rigid only within a certain range of pressure drops. If a new column design takes the pressure drop out of this range, the performance may be drastically altered. For rigid particles, the pressure drop is described by equation 8; therefore, its scaling is

$$\pi = \frac{\theta\lambda}{\rho^2\delta^2} \tag{32}$$

For compressible particles, the same procedure may be used with a correlation that describes the relationship between pressure drop and flow rate (39).

SUMMARY

SEC has been applied extensively to biological macromolecules in the laboratory and has found utility as a step in the production of proteins at large scale as well. Further growth of the biotechnology industry and the ease and applicability of the technique ensure that SEC will be used widely in the future as well. Many packing materials have been developed to enable the chromatographer to perform an analysis or separation quickly, efficiently, and unambiguously through improvements in the inertness and mechanical stability of the support. Engineering analysis has enabled quantitative descriptions to be made regarding the influence of mass transfer resistances to the efficiency (plate height) of the column. These can be incorporated into simple design equations for the development and optimization of column configurations and operating conditions for a number of industrially important separations.

BIBLIOGRAPHY

1. J. Porath and F. Flodin, *Nature* **183**, 1657–1659 (1959).
2. H. Colin, in G. Ganetsos and P.S. Barker eds., *Preparative and Production Scale Chromatography*, Dekker, New York, 1993, pp. 11–45.
3. H.G. Barth, B.E. Boyes, and C. Jackson, *Anal. Chem.* **68**, 445R–466R (1996).
4. C.V. Uglea, *Liquid Chromatography of Oligomers*, Dekker, New York, 1996, pp. 269–274.
5. R. Eksteen, in E.D. Katz ed., *High Performance Liquid Chromatography: Principles and Methods in Biotechnology*, Wiley, New York, 1996, pp. 144–145.
6. M.L. Yarmush, K. Antonsen, and D.M. Yarmush, in C.L. Cooney, A.E. Humphrey eds. *The Principles of Biotechnology: Engineering Considerations*, Pergamon Press, Oxford, UK, 1985, pp. 489–505.
7. J.-C. Janson, in G. Subramanian, ed., *Process Scale Liquid Chromatography*, VCH, Weinheim, Germany, 1995, pp. 81–98.
8. H. Determann, *Gel Chromatography*, Springer-Verlag, New York, 1969.
9. Pharmacia Biotech catalog, 1995, pp. 308–322.
10. C. De Felice, K. Hayakawa, T. Watanabe, and T. Tanaka, *J. Chromatogr.* **645**, 101–105 (1993).
11. T.W. Patapoff, R.J. Mrsny, and W.A. Lee, *Anal. Biochem.* **212**, 71–78 (1993).
12. M.K. Dewanjee, A.K. Ghafouripour, M. Kapadvanjwala, and A.T. Samy, *Biotechniques* **16**, 844–850 (1994).
13. J.P. Hummel and W.J. Dreyer, *Biochim. Biophys. Acta* **63**, 532–534 (1962).
14. S. Kadokura, T. Miyamoto, and H. Inagaki, *Biopolymers* **20**, 1113–1122 (1981).
15. V.N. Uversky, *Biochemistry* **32**, 13288–13298 (1993).
16. X.-M. Lu, A.J. Fischman, S.L. Jyawook, K. Hendricks, R.G. Tompkins, and M.L. Yarmush, *J. Nucl. Med.* **35**, 269–275 (1994).
17. M.S. Wadhwa, D.L. Knoell, A.P. Young, and K.G. Rice, *Bioconjugate Chem.* **6**, 283–291 (1995).
18. S.S. Wong, *Chemistry of Protein Conjugation and Cross-Linking*, CRC Press, Boca Raton, Fla., 1991, p. 213.
19. R.K. Mishra, C. Moreau, C. Ramazeilles, S. Moreau, J. Bonnet, and J.Y. Toulmé, *Biochim. Biophys. Acta* **1264**, 229–239 (1995).
20. V.B. Lawlis and H. Heinsohn, *LC-GC* **11**, 720–729 (1993).
21. J.-C. Janson and T. Pettersson, in G. Ganetsos and P.S. Barker eds., *Preparative and Production Scale Chromatography*, Dekker, New York, 1993, pp. 559–590.
22. A.J.P. Martin and R.L.M. Synge, *Biochem. J.* **35**, 1358–1368 (1941).
23. O. Levenspiel, *Chemical Reaction Engineering*, Wiley, New York, 1972, pp. 1–293.
24. T.C. Laurent and J. Killander, *J. Chromatogr.* **14**, 317–330 (1964).
25. A.G. Ogston, *Trans. Faraday Soc.* **54**, 1754 (1958).
26. P.G. Squire, *Arch. Biochem. Biophys.* **107**, 471–478 (1964).
27. G.K. Ackers, *J. Biol. Chem.* **242**, 3237 (1967).
28. W.W. Yau and C.P. Malone, *Polym. Lett.* **5**, 663–669 (1967).
29. J.J. van Deemter, F.J. Zuiderweg, and A. Klinkenberg, *Chem. Eng. Sci.* **5**, 271–289 (1956).
30. J.F.G. Reis, E.N. Lightfoot, P.T. Noble, and A.S. Chiang, *Sep. Sci. Technol.* **14**, 367–394 (1979).
31. A.M. Lenhoff, *J. Chromatogr.* **384**, 285–299 (1984).
32. A.M. Athalye, S.J. Gibbs, and E.N. Lightfoot, *J. Chromatogr.* **589**, 71–85 (1992).
33. C. Horváth and H.-J. Lin, *J. Chromatogr.* **149**, 43–70 (1978).
34. F.E. Regnier and K.M. Gooding, *Anal. Biochem.* **103**, 1–25 (1980).
35. H.W. Blanch and D.S. Clark, *Biochemical Engineering*, Dekker, New York, 1996, p. 584.
36. P.L. Wankat and Y.-M. Koo, *AIChE J.* **34**, 1006–1019 (1980).
37. D. Corradini, R. Filippetti, and C. Corradini, *J. Liq. Chromatogr.* **16**, 3393–3408 (1993).
38. R.K. Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag, New York, 1984, pp. 238–250.
39. A.W. Mohammed, D.G. Stevenson, and P.C. Wankat, *Ind. Eng. Chem. Res.* **31**, 549–561 (1992).

See also ADSORPTION, PROTEINS WITH SYNTHETIC MATERIALS; CHROMATOGRAPHY, ION EXCHANGE; MEMBRANE CHROMATOGRAPHY.

CITRIC ACID, PROCESSES

M.R.V. KRISHNAN
Anna University
Chennai, India

KEY WORDS

Aspergillus niger
Enzyme CoA
Fermentation
Molasses
Mold
pH
Solid substrate fermentation
Submerged culture
Surface culture
Yeast

OUTLINE

INTRODUCTION

Citric acid ($C_6H_8O_7$) is a white or translucent solid with a molecular weight of 192.12. It occurs as a natural constituent in citron, lemon, lime, pineapple, pear, peach, and similar fruits (Table 1). It is also found in animal tissues. The popular forms of citric acid are the anhydrous forms, the monohydrate and sodium salts of the acid.

Table 1. Citric Acid Contents of Various Fruits and Vegetables

Fruit/vegetable	Citric acid (wt %)
Lemon	4.0–8.0
Grape	1.2–2.1
Tangerine	0.9–1.2
Orange	0.6–1.0
Currant	
Black	1.5–3.0
Red	0.7–1.3
Raspberry	1.0–1.3
Strawberry	0.6–0.8
Apple	0.008
Tomato	0.25
Potato	0.35–0.5
Asparagus	0.08–0.2
Turnip	0.05–1.1
Pea	0.05
Corn	0.02
Lettuce	0.16
Eggplant	0.01

Citric acid was first isolated by Scheele in 1784 when he crystallized it from lemon juice. It was later synthesized (via symmetrical dicholoroacetone) from glycerol by Grimocex and Adam in 1880. It was first made in crystallized form from lemon juice. Calcium citrate was precipitated through the reaction of hot lemon juice with calcium carbonate, followed by decomposition of the product with sulphuric acid. Two hundred years later, this process is still being used. However, raw material is the principal limiting factor for production on a large scale. Thirty tons of lemon are needed to produce 1 ton of citric acid. The first commercial citric acid was prepared in 1860 in England from calcium citrate imported from Italy and Sicily. By 1880, France, Germany, and the United States had begun manufacturing citric acid using similar methods. In 1913, recovery of citric acid from calcium citrate began in Italy, and by 1922, Italy produced 90% of the world's supply.

MICROBIAL PRODUCTION

In 1893, Wehmer recognized citric acid as a microbial metabolite (1) (Table 2). Although his attempt to develop a commercial fermentation process with a species of *Penicillium* was unsuccessful, he established the basis for the subsequent complete reorganization of the citric acid industry. It was later confirmed that the use of *Aspergillus niger* (Fig. 1) favors sucrose fermentation to citric acid.

In 1923, a plant was started in New York to make citric acid by a fungal fermentation technique developed by Currie (2). This production effectively broke the power of the Italian cartel. In this process, *A. niger* was grown on the surface of a shallow pan of sugar medium. This surface fermentation process was subsequently used in England, The Netherlands, Belgium, Germany, Switzerland, Argentina, and the former Soviet Union.

A new fungal fermentation process was introduced successfully in the United States in 1952. In this process,

Table 2. Chronology of Events in Citric Acid Production

Authors (ref.)	Investigation/remarks
Wehmer 1893 (cited in 1)	Demonstrated citric acid to be a metabolic product of certain molds.
Currie 1917 (2)	First chose *Aspergillus niger* for deriving citric acid.
Doelger & Prescott 1934 (3)	Fermentation above 30 °C decreased citric acid yield and increased accumulation of oxalic acid.
Cahn 1935 (4)	First to describe solid-state fermentation for production of citric acid.
Mazzadroli 1938 (5)	Potassium ferrocyanide could be added to eliminate excess Fe^{2+}.
Szucs 1944 (6)	Citric acid synthesis took place only after total assimilation of phosphates.
Moyer 1953 (7)	Suggested use of methanol as stimulant for increasing acid yield.
Usami & Takatomi 1958 (1)	Spores produced under submerged fermentation were poor, and spores from surface culture were better acid formers.
Usami et al. 1960 (9)	Increase in aeration resulted in higher yields and reduction in cycle time.
Kovats 1960 (10)	Demonstrated lower concentration of sugar leads to lower yields of citric acid and accumulation of oxalic acid.
Schweiger 1961 (11)	Iron concentration above 0.2 ppm affected yield of acid. However, addition of copper 0.1 to 500 ppm at the time of inoculation or during the first 50 h of fermentation countered the deleterious effect of iron.
Naguchi & Bando 1960 (12)	Concentration of ammonium nitrate greater than 0.25% lead to accumulation of oxalic acid.
Sanchez-Marroquin et al. 1963 (13)	Higher yields of citric acid occurred in a simple medium rather than a complex medium.
Clark 1962 (14)	Ferrocyanide addition on the acid yield 10–200 μg/mL was tolerated during growth. But less than 20 μg/mL helped acid production.
Millis et al. 1963 (15)	Addition of vegetable oils, fatty acids, etc., increased acid yield.
Czech. Acad. Sci. 1964 (16)	Mycelial digests from *A. terreus* or *A. Niger* stimulated acid production by 60–70% when added to the medium.
Leopold 1965 (17)	Addition of pressed baker's yeasts to culture medium increased acid yield from 61.5 to 72.1%.
Bruchmann 1966 (18)	Addition of mild oxidizing agents stimulated acid production.
Sussman & Halvorson 1966 (19)	Spore viability varies with age.
Wendel 1967 (20)	In the initial stages, citric acid diffused into the medium, which leads to a stratification on the mycelium. This stratification inhibited fungal metabolism, leading to decreased acid yields. This stratification can be eliminated by stirring.
Tabuchi et al. 1969 (21)	Production of citric acid from *n*-paraffin using yeast, (*Candida*) medium composition. Hydrocarbon, 40–60 gs; NH_4Cl_2, 2 gs; KH_2PO_4, 0.5 g; $MgSO_4$, 0.5 g; cornsteep liquor, 1 g; $CaCo_3$, 30 g/L.
Khan et al. 1970 (22)	Phosphates promoted more growth at less acid yields.
Fedoseev et al. 1970 (23)	Copper sulfate at 4.7 mg/100 g molasses resulted in better conversion of sugar to citric acid.
Kyowa Fermentation Industry 1970 (24)	Citric acid production from dodecane (or) C_{12}–C_{14}. Using arthrobacterium in aqueous medium, inoculum 5% air at 3 vvm[a] at 28 °C, yielded 28 mg/mL.
Sanchez-Marroquin et al. 1970 (25)	The use of ion-exchange resin for reduction of metal content was better than chemical treatment.
Fukuda et al. 1970 (26)	Use of *Corynebacterium* on *n*-paraffins; yield 41.4 mg/mL; culture period 64 h, 32 °C.
Leopold 1971 (17)	Glycerol addition increased acid yield by 30%.
Dhankar et al. 1972 (27)	Acid yield improved by addition of peanut oil to molasses medium.
Kumamoto & Okamura 1972 (28)	Mn^{2+}, Ba^{2+}, Al^{3+} had an effect on fungal morphology.
Sardinas 1972 (29)	Patented a process for citric acid production involving *B. licheniformis*.
Chaudhary et al. 1972 (30)	A low pH in molasses medium was inhibitory to growth.
Ohmori & Ikeno 1973 (31)	Bacterial production of citric acid from media containing isocitric acid.
Dhankar et al. 1974 (32)	Sodium nitrate at a concentration of 0.4% was superior to ammonium nitrate.
Zhuravskii, 1974 (33)	Patented continuous multistage process for citric acid production.
Halama 1974 (34)	Chemicals such as pentachloralphenolate, semicarbozone, tetracycline, etc., used for control of bacterial growth.
Ohtsuka et al. 1975 (35)	Addition of malic hydrazide increased yield from 30 to 70%.
Wold & Suzuki 1973 (36)	The rate of acid production increased when AMP was added to the culture medium.
Wold & Suzuki 1976 (37)	Zn^{2+} regulated growth and production.
Brezhnoi et al. 1976 (38)	Continuous production of citric acid from molasses.
Choudhary 1978 (39)	A 3-day-old slant culture was as good as 7- to 8-day-old spore.
Krishnan & Natarajan 1987 (40)	Liquid–liquid extraction of citric acid.
Krishnan & Annadurai 1992 (41)	Precipitation of citric acid.
Krishnan et al. 1996 (42)	Reduction in cycle time for citric acid production.

[a] vvm, volume air per volume medium per minute.

A. niger was allowed to grow in the entire volume of the culture solution in large deep tanks. This process was introduced in Mexico in 1959 and in Israel in 1961.

The earliest attempt to produce citric acid by the submerged growth technique was that by Amelung, who aerated the culture by bubbling air through the solution. Numerous studies, subsequent to the work of Amelung, resulted in the development of two approaches to the control of citric acid production in submerged cultures. In the first approach, Perquin produced citric acid by inducing the establishment of a deficiency of one of the major nutrients, such as phosphate. In another approach, Johnson et al. studied the effect of manganese and iron and related the concentration of these to the amount of inhibition of citric acid accumulation.

Further development of methods for the control of iron and manganese levels in the culture solution led to the popularity of the submerged fermentation process, now used in the United States, Mexico, and Israel. Excess iron may be precipitated by ferrocyanide. The removal of iron and the presence of specific enzyme inhibitors that control the destruction of citric acid form the basis for the successful submerged fermentation process. Other mineral nutrients required for the submerged fermentation are the same as those for surface fermentation.

The most interesting change in citric acid manufacture was done by Miles Laboratories. They changed the raw materials from molasses to glucose, doubling the production. The company patented a process in which starchy material is first converted to sugar by enzymatic action and then used as the raw material for citric acid production. They also claim that with the addition of various phenols and other compounds to check the growth of *A. niger*, it is possible to use less highly purified starting material. This is similar to the use of the lower alcohols, which has been known for many years and is being investigated now as an additive to blackstrap molasses.

STORAGE OF CITRIC ACID

The popular forms of citric acid are the anhydrous forms, the monohydrate form and sodium salts of the acid. Both citric acid and citric monohydrate are available in variety of sieve sizes. They are conventionally available as granular, fine granular, and powder forms.

Crystalline anhydrous citric acid can be stored in dry form without difficulty, although highly humid conditions and elevated temperatures should be avoided to prevent caking. The product should be stored in tight containers to prevent exposure to moist air. Several granulations are commercially available with larger particle sizes having less tendency toward caking. Materials packed with desiccants are also available. Solutions of citric acid are corrosive to normal concrete, aluminium, carbon steel, copper alloys and should not be used with nylon, polycarbonates, polyamides, polyimides, or acrylics. Recommended materials of construction for pipes, tanks, and pumps handling citric acid solutions are 316 stainless steel, fiberglass-reinforced polyester, polyethylene, polypropylene, and polyvinyl chloride. Sodium and potassium citrates are not as corrosive as citric acid but they should be handled in the same type of equipment as citric acid.

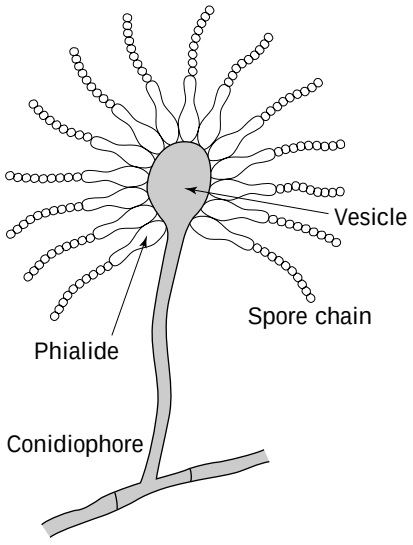


Figure 1. *Aspergillus niger*.

PRODUCT APPLICATIONS

The food and pharmaceutical industries use citric acid extensively because of its high solubility, pleasant sour taste, very low toxicity, and ready assimilability (Table 3). Citric acid also finds application in some cosmetic preparations, metal cleaning, electropickling, copper plating, secondary oil recovery, and other industrial uses (Table 4).

Foods and Beverages

Citric acid is widely used in foods, beverages, and jams and jellies owing to its ability

Table 3. Global Production Pattern of Citric Acid

Place	Production (%)
Western Europe	41
North America	28
South and Central America	11
Far East, Australia, and New Zealand	11
Other	9

Table 4. World Consumption Pattern of Citric Acid

Uses	Consumption (%)
Food and beverages	60
Pharmaceuticals	12
Detergents	12
Metal cleaning	6
Textile dyes	5
Cosmetics	3
Others	2

1. To maintain desired pH levels
2. To impart refreshing and tingling taste
3. To act as a flavor-enhancing agent
4. To act as a color stabilizer

In candy manufacture, citric acid helps to enhance the flavor and taste of berries and other ingredients. As a very dependable agent for maintenance of pH, citric acid finds wide application in the production of jams, jellies, preserves, soft drinks, syrup, and soft drink tablets. For pH control and color stabilization, citric acid is used in fruit and vegetable juices. The pH prevents juice spoilage, and natural taste and flavor are greatly enhanced. Lowering of pH inactivates certain oxidative enzymes, thus improving the keeping qualities of frozen fruits, peaches, apricots, plums, pears, and cherries against flavor or color spoilage. Citric acid also combines with trace metals, helping avoid the undesirable oxidation changes.

Pharmaceuticals

Citric acid's wide application in the manufacture of drugs and pharmaceuticals is for enhancing the taste and flavor. It is used as an anticoagulant for blood. Powders and tablets owe their effervescent property to citric acid. The free acid and its sodium and potassium salts are used as mild acidulants in astringent preparations. Several salts of citric acid are used in the manufacture of antianemic tonics, vitamins, liver tonics, antipyretic agents, cough syrups, antidysentery, and antidiarrheal agents. It is also used in inducing fertility.

Cosmetics

Hair rinses and hair setting agents derive benefit from citric acid. It is also used in astringent lotions, bleaching lotions, and others.

Metallurgy

Citric acid finds extensive use in metallurgical application and as a sequestering agent for such metals as iron, copper, zinc, nickel, cobalt, chromium, and manganese. The acid and ammonium salts are added for scale removal in boilers; cleaning of reactors could be done very well using the acid. Citric acid is widely used in electropickling of copper and its alloys and in copper coating. Any iron plugging in the oil-flow line could be removed by the acid. Tanning liquors, bottle washing compounds, and diazo printing paper all contain citric acid.

Industrial Uses

It is used in the manufacture of several esters, linoleum, inks, silvering compounds, and fabric dyes (Table 5). Citric acid enjoys a superior status above phosphates in the manufacturing of detergents.

Other Applications

Some of the algicides, pesticides, fish preservation, chemicals, and insecticides are based on citric acid. Preservation of latex from certain natural plants involves the use of citric acid.

Table 5. Some Leading Manufacturers and Suppliers of Citric Acid

Manufacturers
Miles Laboratories Inc., Elkhart, Indiana
Joh. A. Benckshiser Gmbh, Ludwingshafen/Rhein, Germany
Rhone-Poulenc S.A., France
John & E. Sturge Ltd., Birmingham, England
Cargill Inc., Minneapolis, Minnesota
San Fu Chemical Co., Ltd., China
Boehringer Ingelheim KG, Rhein, Germany
Pfizer Inc., New Jersey
Fermenta Products, Quimicos, SA, Brazil
Showa Chemical Co., Pvt Ltd., Osaka, Japan
Suppliers
Lurgi AG. Frankfurt am Main, Germany
Snamprogetty SPA, San Donato, Milan, Italy
Mannesman Anlagenbau, Dusseldorf, Germany
Boehringer Ingelheim KG, Burgerstrasse, Germany
Vogelbusch GmbH, Austria

THE TECHNOLOGY OF CITRIC ACID PRODUCTION

Citric acid production centers around one strain, *Aspergillus niger*, which is still extensively used. In the late 1970s, processes involving *Candida* yeast became commercial. *Candida guillermondi* has unique benefits over *A. niger* in that trace metal removal is not essential. It can also tolerate higher pHs (3.5 to 5) and sugar concentrations for acid production. The rate of acid production is also faster. As in the case of *A. niger*, nitrogen limitation triggers acid production. Certain bacteria also produce citric acid on a variety of substrates.

Citric acid can be manufactured commercially by three methods:

1. Surface culture technique
2. Submerged culture technique
3. Solid substrate technique

Commercially used substrates are citrus fruits (not adapted presently because the fruits are costly and seasonal), sucrose, glucose, molasses cane juice, and certain petroleum fractions.

Shallow pan reactors, stirred tank fermenters, and airlift fermenters are commonly used depending on the process requirements. Fermenters work in a semibatch mode or a continuous mode. It will be of interest to state here that yields of product vary with the pretreatment to which the media are subjected. For example, yields for pretreatment with suitable exchange resins is 98% in the case of sucrose and 75% for molasses; ferrocyanide-treated molasses yields 68% as contrasted to the untreated molasses, for which the yield is 62%.

Citrus Fruits and Pineapples

Lemon peels and the white layer are removed and separated for the recovery of oil and pectin, respectively. The remaining portion is pulped and filtered. The filtrate is likely to contain pectin and albumin, which are removed by spontaneous fermentation. In the case of pineapples, the waste and low quality fruits are crushed and filtered. The lemon juice and the pineapple juice, recovered separately, could contain 4% and 0.75% citric acid, respectively. Calcium carbonate or hydroxide is added to the individual juice to the required amount, and the temperature is maintained between 55 and 95 °C. Any oxalate present is filtered preferentially before citric acid. The calcium citrate cake obtained is washed and reacted with a small amount of sulfuric acid. The resultant slurry is filtered to recover citric acid solution from the precipitated calcium sulfate. The clear filtrate is subjected to decolorization before recovering the citric acid crystals by evaporation. This process is on the decline, as stated earlier.

Cane Molasses and Beet Molasses

Blackstrap molasses, about 40 °Bé with a sugar content of 52 to 57% sugar, or beet molasses, 41 °Bé containing 48 to 52% sugar, could be used for citric acid production (Table 6). However, cane molasses should be purified before subjecting it to fermentation. Alternate methods for molasses purification are discussed below:

1. *Sulfuric acid treatment*. The pH of the molasses (10% total reducing sugar) is adjusted to 3.0 by adding 0.1 N sulphuric acid. This is allowed to stand for 1.5 h and then centrifuged at 3,000 rpm for 15 min. The supernatant is collected and used.
2. *Potassium ferrocyanide treatment*. The molasses (10% TRS) is heated to 85 °C for 30 minutes and centrifuged at 3,000 rpm for 15 min; 100 mL of the supernatant is collected, and 0.5 mL of K_4 $Fe(CN)_6$ (10% solution) is added. The pH is adjusted to 6.5.
3. *Tricalcium phosphate treatment*. The pH of the molasses (10% TRS) was adjusted to 7.0 by the addition of 0.1 N NaOH and treated with 2% (w/v) tricalcium phosphate followed by heating at 105 °C for 5 minutes. The mixture was cooled and centrifuged at 3,000 rpm for 15 minutes. The supernatant was used.
4. *Tricalcium phosphate with hydrochloric acid treatment*. The TCP-treated liquor was adjusted to pH 2.0 by the addition of 0.1 N Hcl followed by vigorous shaking. The mixture was allowed to stand for 6 h. The supernatant was used after centrifugation at 3,000 rpm for 20 min.
5. *Bentonite treatment*. The pH of molasses (10% TRS) was adjusted to 7.0. Bentonite 2% (w/v) was added and kept in a boiling water bath for 30 min. The solution was then centrifuged at 3,000 rpm for 15 min. The supernatant was collected and used.

Table 6. Analyses of Cane and Sugar Beet Molasses

Component	Beet (%)	Cane (%)
Water	20.0	16.5
Sugar	62.0	53.0
Nonsugars	10.0	19.0
Ash	8.0	11.5

Molasses Medium

Beet molasses is the most widely used raw material in the United States and Europe. Latin American and Caribbean plants use sugar as raw material (3 tons of sucrose per ton of citric acid). Smaller plants in the Caribbean use citric wastes from citrus fruits. The production rates are Argentina, 2,000 tons per annum, Mexico, 1,000 tons per annum; and Uruguay, 500 tons per annum.

The beet molasses or pretreated cane molasses is taken in a mixing vessel where dilute sulphuric acid is added to register a pH of 5.5 of 6.5. Phosphorous, potassium, and nitrogen are added as nutrients to the required amount for growth and citric acid production. This mixture is sterilized with live steam. After this, a requisite amount of sterile water is added to give a sugar percentage of 15 to 20%. The feed is admitted into shallow aluminium pans arranged in convenient stacks in a sterile room. Each tray is shallow (about 75-mm deep) but big enough to hold solutions up to 500 L. The temperature (28 to 32 °C) and humidity (40 to 60%) are controlled for maximum yield. This sugar medium is inoculated with the spores of a selected strain of *A. niger*. In about 8 to 10 days, the fermentation process is over (pH being about 2). The acid might be contaminated with oxalic acid and gluconic acid. The tray contents are sent out for recovery of the acid, and the trays and the chamber are sterilized with steam. Dilute formic acid or sulfur dioxide could also be used for positive sterilization.

The recovery consists of adding hydrated lime (1 part of lime to every 2 parts of liquor) to the fermenter broth kept at 95 °C. The precipitated calcium citrate could be reacted with enough dilute sulfuric acid when citric acid remains in solution, leaving behind calcium sulfate. Based on the sugar content of the raw materials, the yield could be 35 to 65% by weight of sugar. In this process, pH maintenance is difficult.

Submerged Fermentation

The majority of industries have adopted this process. Conventionally, the growth phase and the production phase are two distinct regions. Two different media are used, one for growth and the other for production (Table 7). After 3 to 4 days of lush growth, the mycelia are separated from the growth medium and added to the production fermenter under stirred conditions. In the production fermenter, the pH and temperature are maintained around 3 and 30 °C, respectively. Air is sparged without interruption into the solution at about 1 to 1.2 volume air per volume medium per minute (vvm). In about 4 to 5 days, the production becomes nearly complete with about 65 to 70% of sugar being converted to the final product.

The submerged process is admirably suited for a flexible operation. It is possible to introduce certain process accel-

Table 7. Medium for Citric Acid

Component	Sporulation (g/L)	Production (g/L)
Sucrose	140	140
Bactoagar	20	0.0
Ammonium nitrate	2.5	2.5
Potassium hydrogen phosphate	1.0	2.5
Magnesium sulphate heptahydrate	0.25	0.25
Copper ion	0.0048	0.00006
Zinc ion	0.0038	0.00025
Ferrous ion	0.0022	0.0013
Manganous ion	0.001	0.001

erators or nutrients at desired and convenient time intervals; activated carbon or ascorbic acid could be used as the accelerator. The flexibility of such a fermenter is so good that sugar concentration and pH could be maintained at desired values. The acid recovery is through the calcium citrate process. For every ton of citric acid produced, one needs about 3,650 kg of molasses, 5 to 14 kgs of nutrient, 650 kg of sulfuric acid, and 450 kg of lime. Although one might use several organisms (*A. niger* is still the most sought after), the yield depends mainly on the nutrient used, presence of trace metals, and above all, the pH and air supply.

In the new process by Miles Laboratories mentioned previously, sucrose has been replaced with glucose for acid production. Starch is converted to sugar through enzymatic reaction, and the resulting sugar is the raw material for acid production. Specific chemicals control the growth of the mold at the expense of citrate production. An interesting point is that their starting material is not of high purity. They use very large fermentation vats into which sucrose solution is pumped, and fungal spores are added to help the growth and production cycles. After the cycle is over, the vats are charged with fresh medium.

Solid-Substrate Fermentation

In the solid-substrate fermentation process, the mold *A. niger* ferments molasses or sucrose adsorbed on beet or cane pulp. Carriers such as beet or cane pulp offer large areas for microbial contact with the medium and ensure good aeration. Bagasse is preferred over beet pulp because of its larger surface area and lower cost. In addition, bagasse can be used several times over. Bagasse is broken down into fine pieces, 1 to 2 mm long, sterilized, and soaked in molasses or sucrose solution before being inoculated with the selected strain. The citric acid yields are about 45% of the sugar content of molasses and 55% of sucrose. This process is labor intensive and needs more floor space, plus there is the likelihood of contamination. It is difficult to maintain pH. The process requires good aeration with controlled humidities.

Relative Merits between Submerged and Surface Fermentation

The starting sugar cane concentration and the yield percentage are more or less the same for submerged and surface fermentation. However, the surface process is highly labor intensive and not easily amenable for pH or tem-

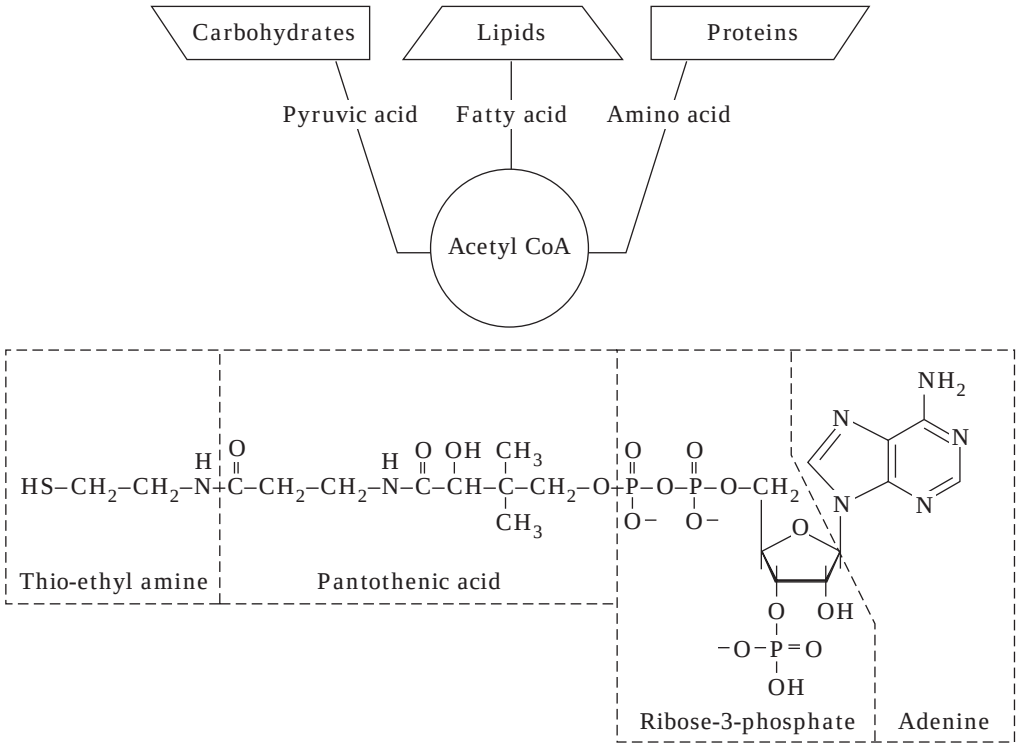


Figure 2. Enzyme CoA.

perature control. Flexibility of operation is higher in the submerged process because addition or removal of chemicals is easier. The risk of contamination is less in the submerged process. The surface process needs a larger working area, and the chance of pollution (caused by spores) is also high. Oxalic acid formation is higher in surface fermentation.

One advantage of the surface process is that it is less energy intensive. Although automation of the submerged process may be easier, in the event of contamination, the losses suffered will be enormous.

BIOCHEMISTRY OF CITRIC ACID PRODUCTION

Successful production of citric acid depends on several factors. External factors can be controlled directly, and biochemical factors, the various biological reactions that go

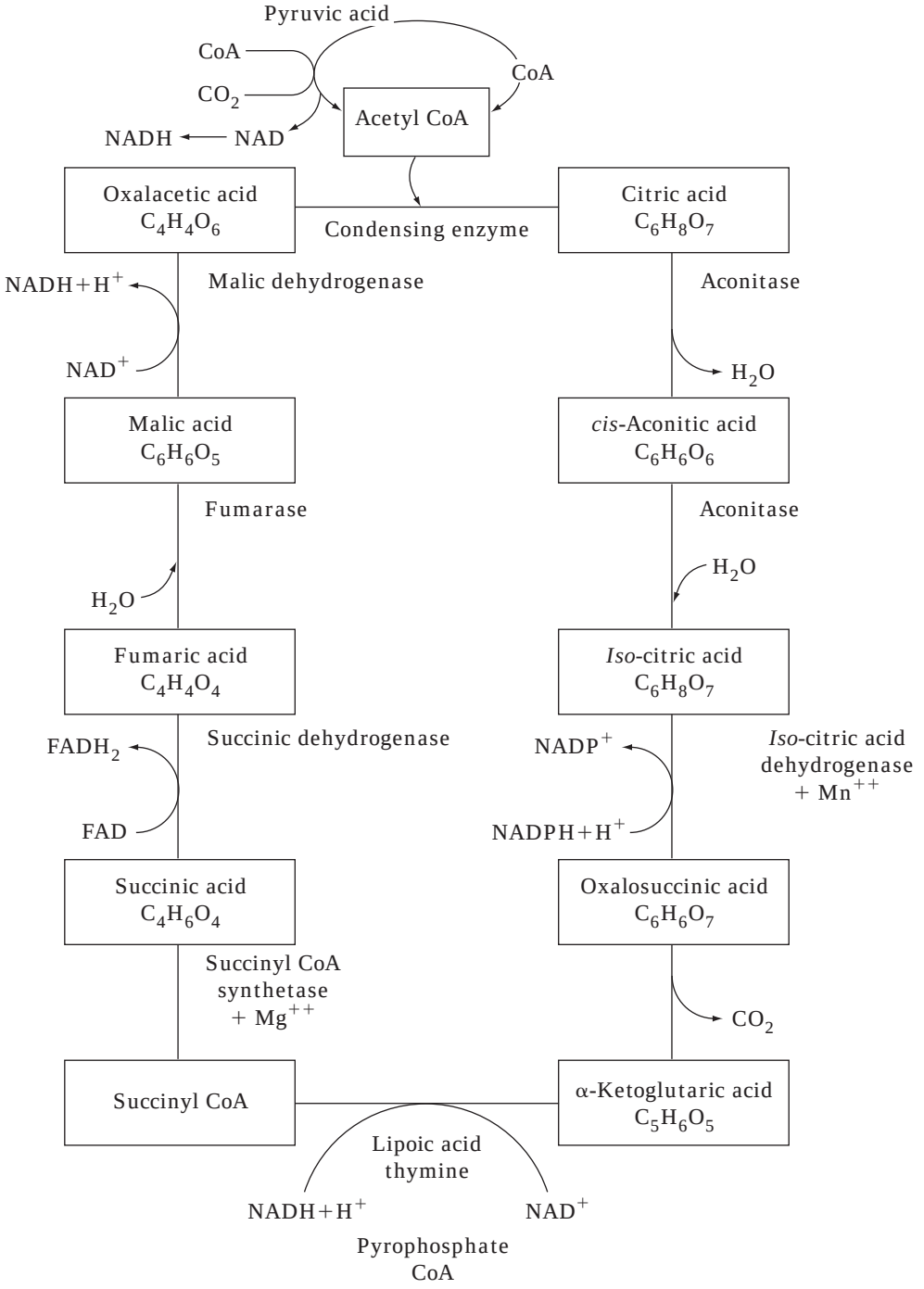


Figure 3. Tricarboxylic acid-cycle.

on inside the cell, can be controlled indirectly by controlling the external factors.

External Factors

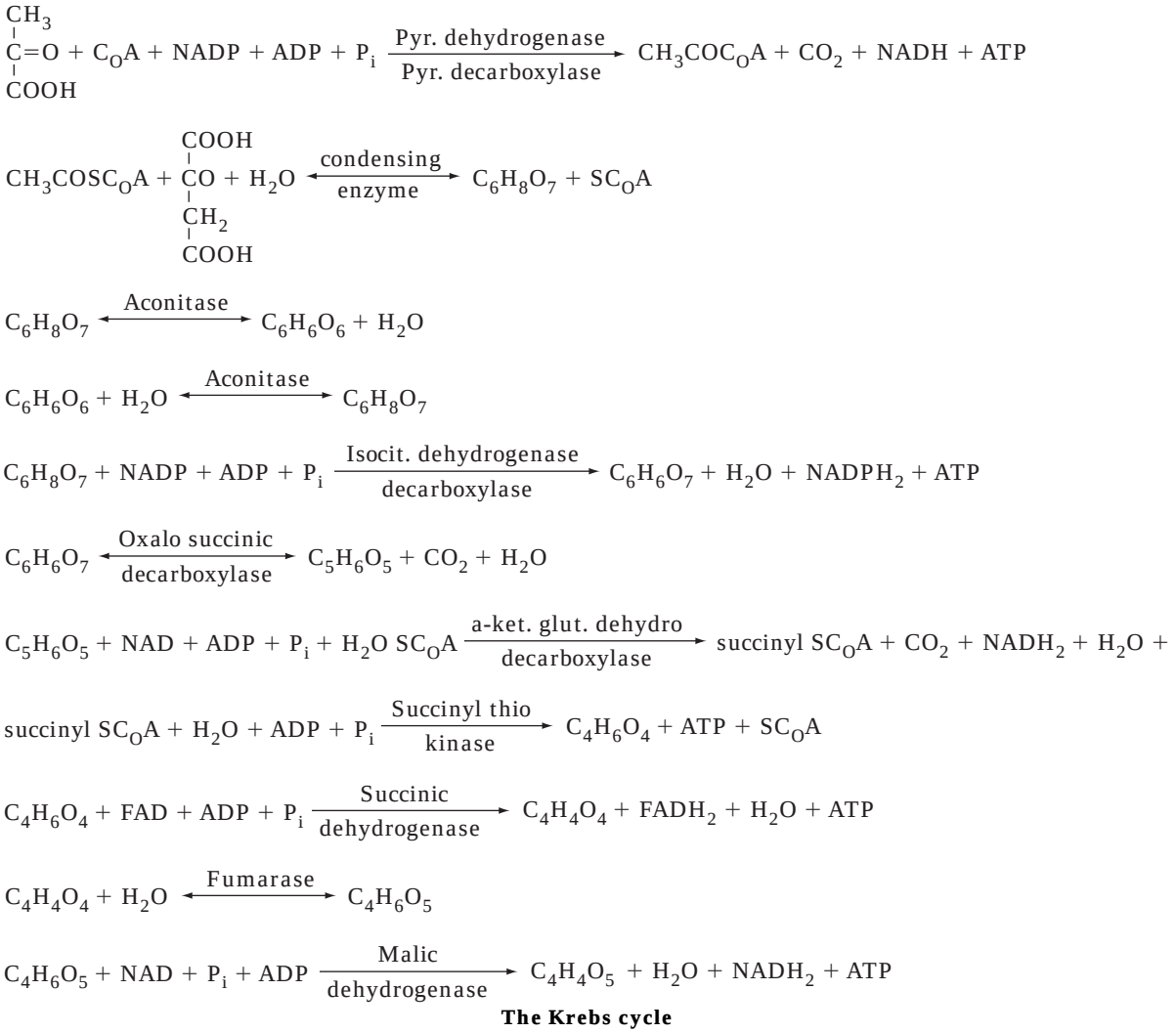
- The type of sugar used and sugar levels
- The pH in the milieu
- The temperature of production
- The presence or absence of trace metals such as Fe, Cu, Zn, Mn
- The presence or absence of phosphates and nitrates in the medium
- The oxygen concentration in the environment

Biochemical Factors

- Breakdown of sugars to pyruvic acid and acetyl CoA (Fig. 2)
- Formation of oxalacetic acid from pyruvic acid and carbon dioxide
- Promotion of the enzymatic reaction necessary for citric acid accumulation
- Suppression of the enzymatic reaction that represses citric acid production
- Oxygen concentration in the medium and cell (citric acid production needs a copious supply of oxygen)
- Favoring citric acid accumulation at the expense of cell growth

Based on the external and biochemical factors, one could draw the following conclusions:

1. Glycolytic and pyruvate enzymes should be activated for the citrate synthesis by maintenance of high sugar levels.
2. Repression of enzymes in the tricarboxylic acid (TCA) cycle would inhibit citrate production.
3. Manganese deficiency should be ensured for NH_4^{2+} generation, which would counter the inhibition of phosphofructokinase by citrate.
4. A copious oxygen supply will overcome possible cell starvation, even under emergent conditions.
5. Ideal pH maintenance encourages citrate production rather than oxalic or gluconic acid formation.



The Krebs cycle

In most cells, catabolism of simple sugar is the major source of energy. It starts with the glycolytic pathway. Glucose is broken down into two molecules of three-carbon compounds called pyruvic acid. The cell gains two ATP molecules for each molecule of glucose entering the glycolytic pathway. The ATP molecules are storehouses of energy, releasing energy necessary for cell work. The catabolic reaction also gives rise to NADH, which is useful for several anabolic reactions.

Depending on the type of cell and availability of oxygen, the pyruvic acid produced earlier can undergo two different types of reactions. When anaerobic conditions prevail, pyruvic acid is converted to lactic acid or ethyl alcohol. Under aerobic conditions, the pyruvic acid is converted to carbon dioxide and acetyl CoA (Fig. 2). This acetyl CoA is further degraded in the (Fig. 3) TCA cycle or election transport system to produce energy for cell work. Several other substrates exist that can give rise to acetyl CoA. For example, proteins and lipids can be degraded to provide acetyl CoA. Thus, cells can degrade several chemical compounds to acetyl CoA and subject them to the TCA cycle and electron transport chain for obtaining energy.

The TCA cycle together with the electron transport system constitutes the cells' primary metabolic furnace. These two combine to burn acetyl CoA in oxygen to liberate carbon dioxide, water, and useful energy. This energy from the TCA cycle and electron transport chain is captured and taken up in ATP for storage of energy.

During oxidation of one molecule of acetyl CoA (by 0the TCA cycle occurring in the mitochondrial matrix), one molecule of flavoprotein (FP or FAB) and three molecules of nicotinamide adenine dinucleotide (NAD) are reduced. These reduced coenzymes are oxidized by molecular oxygen by way of a system of enzymes and coenzymes called the respiratory chain or electron transport system, occurring in the inner mitochondrial membrane. During this oxidation process, enormous amounts of energy are released, some of which is utilized by the inner membrane subunits.

The TCA cycle occurs in the following ways:

- The pyruvic acid is decarboxylated; the acetyl group combines with CoA to give acetyl CoA. The H and NAD combine to give NADH. Energy transfer occurs, yielding ATP.
- Active acetyl combines with oxalacetic acid to form citric acid. The released SCoA can now combine with more CH_3CO.
- Citric acid looses water to give rise to aconitic acid.
- Isocitric acid results as *cis*-aconitic acid gets rehydrated.
- Isocitric acid loses hydrogen to yield oxalosuccinic acid. The H combines with NADP to form NADPH. Energy transport takes place, yielding ATP. Loss of CO_2 also occurs.
- Decarboxylation of oxalosuccinic acid takes place to yield α-ketoglutaric acid. Hydrogen reacts with NAD to form $NADH_2$. Energy is released to convert ADP to ATP.
- SCoA combines with the succinyl group of α-ketoglutaric acid and gets converted to succinic acid, yielding SCoA and ATP.
- Dehydrogenation of succinic acid leads to fumaric acid. Hydrogen joins FAD to form $FADH_2$ and energy. This leads to formation of ATP.
- Malic acid is formed as fumaric acid gets hydrated.
- Malic acid is dehydrogenated to form oxalacetic acid. Oxalacetic acid combines with active acetyl to form citric acid and the cycle continues. The released hydrogen forms $NADH_2$ from NAD and energy released gives rise to ATP.

Several yeasts and bacteria can produce citric acid from specific substrates:

Citric Acid Production by Yeasts

Candida lipolytica
Candida tropicalis
Candida zylenoides
Candida fibrae
Candida intermedia
Candida parpsilosis
Candida petrophylum
Candida subtropicalis
Candida oleophila
Candida hitachinica
Candida citra
Candida guillermondi
Candida sucrosa

A variety of substrates can be utilized by the *Candida* species. They could utilize glucose, acetic acid, calcium acetate, hydrocarbons, molasses, alcohols, fatty acids, and natural oils (such as coconut oil). For example, *Candida* can utilize glucose or molasses as indicated in the list. *C. lipolytica* can use *n*-paraffins, n-alkanes, and alkenes.

From Glucose	From Molasses
Glucose, 10–18%	Sugar cane/beet molasses, 5–250 g/L
NH_4Cl	$(NH_4)_2 SO_4$, NH_4Cl, NH_4NO_3
KH_2PO_4	0.3–1.5 VVM, 6 days
$MgSO_4 \cdot 7H_2O$	5.5–6.50 pH for growth
22–30 °C, 3–6 days	2.8–4 pH for production
	3,154 kg sugar → 1,119 kg monocitrate

Citric Acid Production by Bacteria

Certain bacteria are preferred for producing citric acid because of their low doubling time. Examples include:

Bacillus licheniformis *Bacillus subtilis* *Brevibacterium flavum*	They utilize glucose or isocitric acid or hydrocarbons
Arthrobacter paraffinens can utilize dodecane (C_{12}–C_{14}) at 28 °C over a period of 72 h; yield 28 mg/mL	*Bacillus licheniformis* can utilize glucose medium: 30–37 °C 36–120 h; pH 7; nitrogen salts act as nutrients; yield 42 g/L

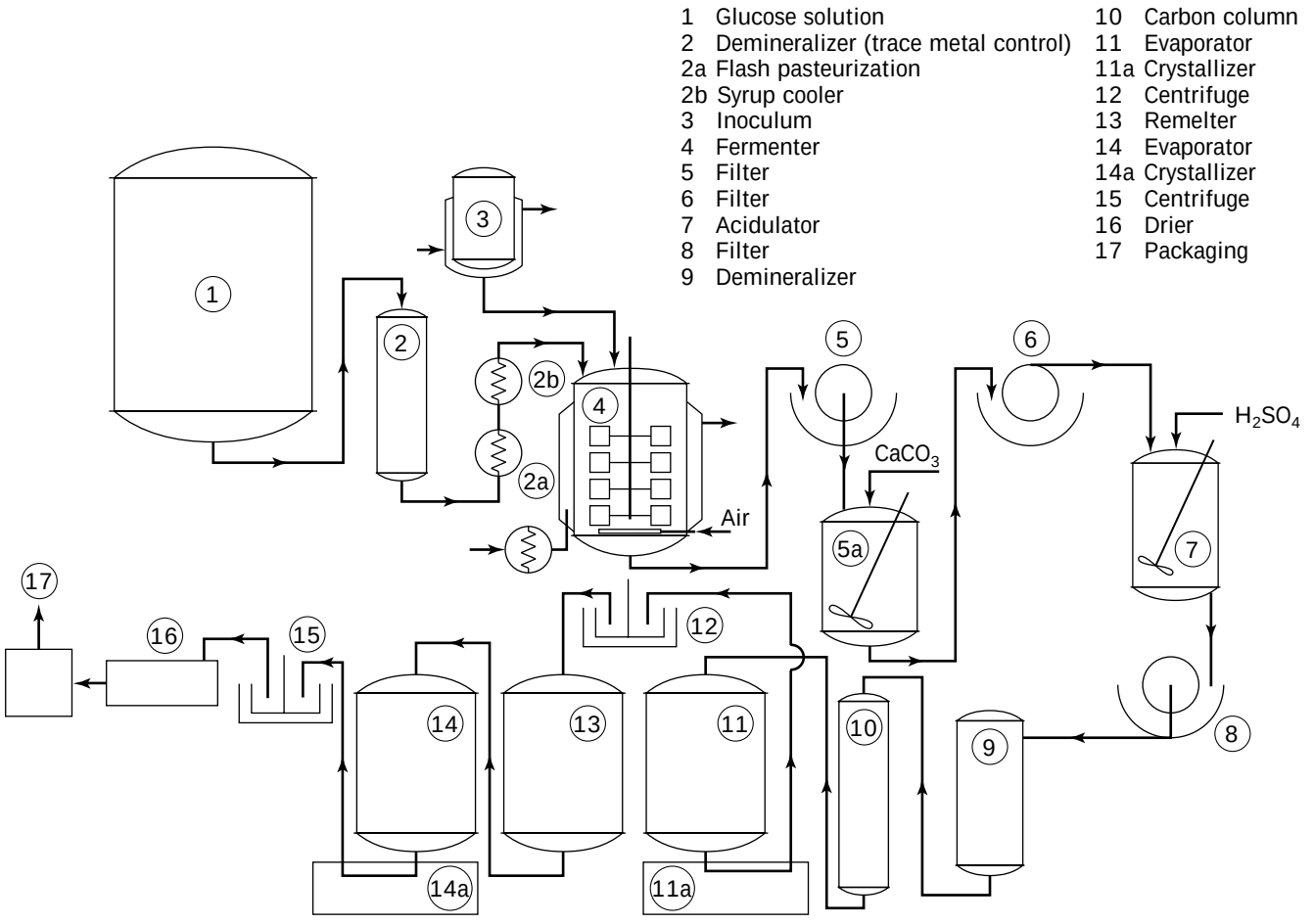


Figure 4. Citric acid production and separation flowsheet.

Recovery of Citric Acid

Two important routes are followed. One is the precipitation as calcium citrate (Fig. 4). The second is the liquid–liquid extraction technique as discussed by Krishnan and Natarajan (40). Solvents such as esters, ketones, amines, and alcohols have been successfully used for the recovery, as outlined by Rieger and Kioustelidis (43).

CONCLUSIONS

Citric acid enjoys a unique place in several important everyday applications. The capabilities of *A. niger* have been almost totally exploited by using several strains of its culture and their mutants. The application of *A. niger* for citric acid production has been stretched nearly to the limit. It is the wishful thinking of the author that some more work could be undertaken in producing such stains that could utilize unclarified raw materials in shorter periods. No effort should be spared in developing newer and cheaper chemicals (promoters) that could accelerate the production rate.

Care has been taken to cover the entire world literature on citric acid production. However, some references might have been missed more because of specific constraints rather than because of oversight or willful omissions. I should like to end by saying that some more work on the application of bacterial culture would be rewarding. It may not be right to say that the last word has been spoken on citric acid production.

Dissolve away difficulties,
Filter out gloom
And from the mother liquor of life
Gather crystals of (citric acid) joy.

Adapted from Krishnan.

BIBLIOGRAPHY

1. S.C. Prescott and C.G. Dunn, *Industrial Microbiology*, McGraw-Hill, New York, 1959.
2. J.N. Currie, *J. Biol. Chem.* **31**, 15–37 (1917).
3. W.P. Doelger and S.C. Prescott, *Ind. Eng. Chem.* **26**, 1142 (1934).
4. F.J. Cahn, *Ind. Eng. Chem.* **27**, 201–203 (1935).
5. Fr. Pat. 833,631 (Oct. 26, 1938), G. Mazzadroli.
6. U.S. Pat. 2,353,771 (July 18, 1944) J. Szucs.
7. A.J. Moyer, *Appl. Microbiol.* 1–7 (1953).

8. S. Usami and N. Takatomi, *Kogyo Kagaku Zasshi* **61**, 1494–1497 (1958).
9. S. Usami, H. Suzuki, and N. Takatomi, *Kogyo Kagaku Zasshi* **63**, 1766–1768 (1960).
10. J. Kovats, *Acta Microbiol. Pol.* **9**, 275–287 (1960).
11. U.S. Pat. 297,084 (Jan. 31, 1961), L.B. Schweiger.
12. Y. Naguchi and Y. Bando, *Hakko Kogaku Zasshi* **38**, 485–488 (1960).
13. A. Sanchez-Marroquin et al., *Rev. Soc. Quim. Mex.* **7**, 191–198 (1963).
14. D.S. Clark, *Biotechnol. Bioeng.* **4**, 17–21 (1962).
15. N.F. Millis, B.H. Trumphy, and B.M. Palmer, *J. Gen. Microbiol.* **30**, 365–379 (1963).
16. Fr. Pat. 1,355,922 (March 20, 1964), Czech. Acad. Sci.
17. Czech. Pat. 142,338 (Aug. 15, 1971), J. Leopold.
18. E.E. Bruchmann, *Naturwissenschaften* **53**, 226–227 (1966).
19. A.S. Sussman and H.O. Halvorson, *Spores: Their Dormancy and Germination*, Academic Press, New York, 1966.
20. D.G. Wendel, *Zucker* **20**, 238–246 (1967).
21. T. Tabuchi, M. Tanaka, and M. Aee, *Nippon Nogei Kagaku Kaishi* **443**, 154–158 (1969).
22. M.A.A. Khan, M.M. Hussain, M.A. Khalique, and M.A. Rahman, Pak. *J. Sci. Ind. Res.* **13**, 439–444 (1970).
23. V.F. Fedoseev et al., *Khlebopek. Konditer. Promst.* **14**, 33–35 (1970).
24. Br. Pat. 1,187,610 (April 8, 1970), Kyowa Fermentation Industry.
25. A. Sanchez-Marroquin et al., *Appl. Microbiol.* **20**, 888 (1970).
26. Ger. Pat. 2,003,221 (Dec. 3, 1970), H. Fukuda, T. Suzuki, Y. Sumino, and S. Akiyama.
27. H.S. Dhankar, M.Sc. Thesis, Haryana Agric. Univ., Hissar, India, 1972.
28. Jpn. Pat. 7,234,955 (Sept. 2, 1972), H. Kumamoto and T. Okamura.
29. Fr. Pat. 2,113,668 (July, 1972), J.L. Sardinas.
30. K. Chaudhary et al., *J. Res. Haryana Agric. Univ.* **1**, 48 (1972).
31. Jpn. Pat. 7,443,157 (Nov. 19, 1973), I. Ohmori and Y. Ikeno.
32. H.S. Dhankar, S. Ethiraj, and S.R. Vyas, *Ind. J. Technol.* **12**, 316–317 (1974).
33. U.S.S.R. Pat. 432,186 (June 15, 1974), G.I. Zhuravskii et al.
34. D. Halama, *Biotechnol. Bioeng. Symp.* **4**, 891–898, 1974.
35. Jpn. Pat. 7,530,158 (Sept. 29, 1975), M. Ohtsuka, Y. Ohzaki, and H. Aramiya.
36. W.S.M. Wold and I. Suzuki, *Biochem. Biophys. Res. Commun.* **50**, 237–244 (1973).
37. W.S.M. Wold and I. Suzuki, *Can. J. Microbiol.* **22**, 1083–1092 (1976).
38. U.S.S.R. Pat. 510,509 (April 15, 1976), Y.D. Brehznoi, L.A. Novikova, Y.L. Ignatov, and Y.E. Kachanov.
39. K. Chaudhary et al., *J. Ferment. Technol. Jpn.* **56**, 554 (1978).
40. M.R.V. Krishnan and M. Natarajan, M. Tech. Thesis Anna Univ. Madras, India, 1987.
41. M.R.V. Krishnan and G. Annadurai, M. Tech. Thesis Anna Univ., Madras, India, 1992.
42. M.R.V. Krishnan, et al. *Bioprocess Eng.* **15**, 323, 1996.
43. Ger. Pat. 2,355,039 (May 7, 1975), M. Rieger and J. Kioustelidis.

ADDITIONAL READING

A.Q. Chaudhary and S.J. Pirt, *J. Gen. Microbiol.* **43**, 71–81 (1966).
U.S. Pat. 3,118,821 (Jan. 21, 1964), D.S. Clark.
Ger. Pat. 2,108,094 (Sept. 2, 1971), H. Fukuda, T. Suzuki, S. Akiyama, and Y. Sumino.
Can. Pat. 973,824 (Sept. 2, 1975), H. Fukuda, T. Suzuki, S. Akiyama, and Y. Sumino.

CLEANING, CLEANING VALIDATION

GAIL SOFER
BioReliance
Rockville, Maryland

KEY WORDS

Assays
Cleaning
Cleaning methods
Equipment
Multiproduct
Nucleic acids
Proteins
Unit operations
Validation
Viruses

OUTLINE

INTRODUCTION

The importance of adequate cleaning in a biological process cannot be overemphasized. The required stringency of cleaning and cleaning validation is best evaluated by performing a risk analysis. The risks associated with the use of improperly cleaned equipment and materials (such as filters and chromatography media) vary with the product source, intended use of end product, the position of the equipment and materials in the processing train, the phase of development of the process, and the type of unit operation. A further consideration is whether the equipment or facility is used for more than one product. Robust cleaning methods, suitable cleaning agents, and cleaning validation ensure that a product with the desired properties is manufactured without contamination from processing additives, well-defined impurities, unexpected contaminants, and previously run product.

PRODUCT SOURCES

For certain complex product sources such as plasma, the potential safety risk from inadequate cleaning is clearly very high. The impurities in the source material are not fully characterized, making evaluation of cleaning efficiency more difficult. For other products, such as secreted products from a mammalian cell culture such as chinese hamster ovary (CHO) cells, the impurities are clearly defined, cleaning efficacy is more readily assessed, and the risks are fewer than with products derived from multiple sources such as blood donors. Robust cleaning procedures are also required for synthetic processes such as those used to produce oligonucleotides used for antisense therapy and synthetic peptides, but the cleaning methods and assessment of their efficacy may be even simpler.

INTENDED USE OF PRODUCT

If the end product is intended for use as a therapeutic agent delivered in a large dose for an extended time in a healthy individual, the criticality of cleaning is obvious. For a product intended for in vitro diagnostic use, the potential contamination that can arise from inadequate cleaning can result in an incorrect diagnosis and even result in a dangerous treatment for a patient. For industrial enzymes, contamination from previously run product may or may not have a detrimental effect on product performance, but even in this case the limits of cleanliness should be defined.

POSITION IN PROCESSING TRAIN

A typical processing scheme will consist of several unit operations (Table 1). The closer one gets to final product, the more critical the cleaning becomes. For example, it has been suggested by some regulators during oral presentations that firms using contract manufacturers for formulation and filling consider supplying their own dedicated equipment to ensure no risk of contamination from other products caused by insufficient cleaning. But even upstream in fermentation or synthetic processes, the lack of adequate cleaning can lead to contamination of the next batch with unknown impurities that adulterate the product in such a way that its safety, potency, and efficacy are compromised.

Table 1. Unit Operations in a Typical Biotechnology Process

Cell culture or fermentation
Isolation
Clarification
Purification
Formulation or final filling

PHASE OF PRODUCT DEVELOPMENT

From research to full-scale manufacturing, proper cleaning is essential for reproducible processes. Not surprisingly, there is a great rush to get a product into the clinic to demonstrate efficacy. In this rush, cleaning is sometimes overlooked, resulting in laboratory data that are unreliable. For example, when a researcher uses a chromatography column for developing a purification process and fails to adequately clean precipitated impurities, the purification process in the next run is likely to yield a product that does not represent what is intended. Good manufacturing practices (GMP) are required for all clinical trials (1). This requirement is also applied in many nations other than the United States. Although cleaning validation might not be complete at this stage, routine cleaning is essential, and it is advantageous to have a cleaning validation plan.

Cleaning and cleaning validation are essential in the manufacture of marketed products and are part of good manufacturing practices. For therapeutics, the legal requirements are given in the U.S. Code of Federal Regulations in the form of general guidance (21 CFR 211.67 and 21 CFR 600.11).

MULTIPRODUCT VERSUS DEDICATED EQUIPMENT AND FACILITIES

In the past, most facilities used for manufacturing therapeutics were dedicated to one product or a family of products from the same source. For complex biologicals, such as plasma products and traditional vaccines, this is still the case. For these biologics, starting materials, and even the final products, may not be sufficiently well defined to evaluate the effectiveness of the cleaning regimes in removing all traces of previously run product and processing materials. Advances in analytical technologies, however, are allowing firms to better assess cleaning effectiveness, especially for well-characterized products. Products that are considered well characterized include therapeutic DNA plasmids, therapeutic synthetic peptides of 40 or fewer amino acids, monoclonal antibodies for in vivo use, and therapeutic recombinant DNA-derived products (2).

Today, there is a trend to reduce processing costs by employing contract manufacturers to produce clinical trial

materials, thereby avoiding the construction of costly facilities for products that may not ultimately obtain regulatory approval. There are also many firms that campaign products (produce one at a time) to fully utilize space, equipment, and personnel. Regulators carefully scrutinize cleaning and cleaning validation in such facilities to ensure that there is no cross-contamination of products. Cleaning validation using sensitive, specific, and nonspecific analytical methods is essential for such operations (see "Methods for Evaluating Cleanliness").

UNIT OPERATIONS

General Considerations

Each unit operation has somewhat different cleaning requirements, which as previously stated are dependent on the nature of the product source and its intended use, its position in the process train, and whether it is used for more than one product. Cleaning is not an issue to be addressed after a process is finalized; it should be designed into the process. In fact, the selection of the most suitable equipment and materials may be dependent on their ability to be cleaned. Records kept during research and development may be valuable in the design of a cleaning process. For example, a researcher may find that a standard cleaning procedure recommended by an equipment vendor is not sufficiently harsh for a particular product feedstream.

Although each feedstream and each unit operation have unique features that must be considered before the most efficient cleaning protocol can be designed, there are many commonalities that can used to minimize development time. References such as regulatory guidelines and PDA's book on cleaning should be consulted (3). Keeping up with the latest technical and regulatory publications and comments at meetings and workshops is essential.

In the design of the cleaning regime for each unit operation, critical factors to consider include the type of cleaning method, compatibility, and detection methods. Cleaning may be automated or manual. Automated cleaning can provide more consistent results. Manual cleaning, on the other hand, is clearly much more difficult to validate, and operators must be certified to perform the task. Cleaning-in-place (CIP) is preferred, but for some components, cleaning-out-of-place (COP) is necessary. Cleaning may be accomplished by chemical or physical methods, or, most likely, by a combination of both.

Chemical Methods

When chemical methods are used, the contact time, temperature, and cleaning agent concentration must be specified. The order of adding cleaning agents and rinsing agents may be critical. Additionally, if the chemical agent is flowing through the equipment, then flow must be specified. It is essential to ensure that the cleaning agent is compatible not only with the specified equipment, but also with any ancillary components that might be exposed to the agent. The nature of the soil must also be considered. Whereas hot alkali and detergents might work for one unit operation, for another they may actually exacerbate the cleaning problem by causing deterioration of the equipment or aggregation of impurities. Organic solvents used for removal of lipids may extract potentially harmful chemicals from some equipment components in a given unit operation. If a packed chromatography column containing a soluble hydrophobic protein is cleaned by treatment with a salt solution, the protein is likely to precipitate and actually leave behind a larger residue. For filters and chromatography media, compatibility with the cleaning agent is of particular concern. For example, when incompatible cleaning conditions are used, an immunoadsorbent might leach significant amounts of antibody, leading to reduced capacity and hence inconsistent results in the next batch. Table 2 shows some common cleaning agents and their mechanisms.

Physical Methods

Physical methods include high velocity flow, jet sprays, and agitation. If physical methods are used, it is essential to ensure compatibility with the equipment. Physical methods commonly used for cleaning tanks and fermentation vessels (e.g., spray balls) are clearly not suitable for any packed chromatography columns or thin piping.

Combining Methods

A combination of physical and chemical methods is usually used. Mechanical forces remove gross soils and deliver cleaning solutions to large areas, but chemicals (such as alkali, acid, and detergents) are generally required as well. According to Rohsner and Serve, cleaning processes require energy as well as chemical and physical reactions (4). Cleaning mechanisms are further addressed by LeBlanc et al. (5).

Some development work is necessary to put into place the most efficient and cost-effective cleaning methods for each unit operation. It is important to recognize that the nature of the soil may change somewhat upon scale-up of a process, necessitating a modification of the cleaning procedure. This is most common in the earlier processing steps such as fermentation. Ranges of acceptable conditions should be established and validated to accommodate the worst-case soil.

Fermentation and Cell Culture. At this early step in the process, the nature of the soil to be cleaned is not well defined, but it may contain substances such as lipids; soluble and precipitated proteins; endotoxins (if Gram-negative bacterial fermentation is used); small, soluble

Table 2. Common Cleaning Mechanisms and Cleaning Agents

Mechanism	Agent
Dissolve	Water, alkali, acid
Saponify	Alkali
Degradation of proteins	Alkali, acid
Wetting	Surfactants
Emulsification, suspension	Phosphates, surfactants
Sequester	Chelating agents

highly charged molecules; large DNA; particulate cell debris; and even whole cells. The amount of each substance to be removed depends on the fermentation or cell culture process. Filamentous fungi generally require harsh cleaning, vessels used for yeasts and bacteria are more easily cleaned, and animal cell cultures in which the products are secreted require even less stringent conditions. Assessing the cleanliness at this early stage requires the use of nonspecific assays such as total organic carbon, and in some cases specific assays are also used (see "Methods for Evaluating Cleanliness").

If the product is secreted and very few cells die during the process, the substances to be removed from the bioreactor may be limited, and relatively mild cleaning procedures might be very effective. However, if the cells die and lyse at unknown rates in a process that is not well controlled, there will be more debris and released cellular components to remove. The heat used to kill cells may also cause aggregation and precipitation, making the cleaning more difficult. The type of cell culture media (i.e., serum, serum-free, defined) must also be considered.

In addition to designing the cleaning procedure to accommodate the type of cells or microorganisms, it is necessary to evaluate the bioreactor design and ensure that the CIP mechanisms are suitable. CIP system design and operation for bioreactors have been addressed by Chisti and Moo-Young (6).

Cell Removal and Clarification. Cell removal and clarification most commonly use filtration and centrifugation operations. Techniques such as expanded bed absorption, two-phase separations, and precipitation are also used. Each one of these methods may require its own unique cleaning procedures. Cleaning must ensure removal of cells, cell membranes, cell walls, nucleic acids, proteins, lipids, etc.

Filtration. In filtration operations, one must consider the membranes as well as the hardware. Chemical cleaning is most common, and the cleaning agents must be compatible with both the membrane and the wetted components of the hardware. Filter housing designs that should be considered for ease of cleaning are discussed in Ref. 3. In addition to the piping and membranes, wetted components may include spacers, gaskets, valves, and seals. Adhesives may also be present. The cleaning agent should not damage any of these over time. Some membranes are incompatible with cleaning agents such as sodium hydroxide; some may be cleaned with mild detergents and then sanitized with bleach. Furthermore, some filter materials have pressure and temperature requirements that must be considered when designing the cleaning protocol. And as with other unit operations, the number of uses of wetted components may need to be defined to ensure that with extended exposure to cleaning agents extractables are not released and that the membrane performance is not changed.

The stringency of cleaning required is dependent on the level of gel polarization and membrane fouling. Brose and Waibel have addressed the adsorption of proteins to microfiltration capsules (7).

The cleanliness of tangential flow membranes is often determined by measuring the flux of clean water through the system at a standard transmembrane pressure. Guidance and cautions for performing this measurement have been described by Michaels (8). Other methods of assessing cleanliness include total organic carbon, physical inspection, and comparing the specifications of a water-for-injection (WFI) rinse.

Centrifugation. Because centrifuges represent a significant capital cost, they are frequently used for more than one product. Cleaning validation is essential to prevent cross-contamination. Although continuous desludging disk centrifuges can be designed so that they can be cleaned in place, those that do not desludge continuously cannot. The bowl in tubular bowl centrifuges must usually be cleaned out of place after disassembly (3). Aerosols generated during centrifugation may require further evaluation of the equipment area to be cleaned. For low-volume products, cleaning can be minimized by using disposable self-contained tubes. This approach may be preferable, for example, when centrifugation is used to prepare multiple products over a relatively short period of time, such as for monoclonal antibodies used for in vitro diagnostics.

Phase Partitioning and Precipitation. With the exception of aerosol generation, phase partitioning and precipitation present the same general cleaning issues as found in tanks and centrifuges. Equipment design plays a major role in determining how well these pieces of equipment can be cleaned.

Expanded Bed Adsorption and Big Beads. Expanded bed adsorption and the use of large (on the order of 300 μm) beads are two techniques in which product isolation is achieved by binding to chromatographic media. Unlike standard chromatography (see next section), particulates and even whole cells are applied to the chromatography media. Cleaning is essential to ensure that the equipment does not retain soils from previous runs and that media can be reused. Chromatography media, whether used for isolation or purification, have a high degree of surface area that allows for adsorption and even entrapment of impurities. Even though chromatography media are dedicated to one product, cleaning must address the specific feedstream—both impurities and product—and its interaction with the media. Furthermore, the equipment must be designed to ensure that cleaning procedures will remove all traces of previously run product and that no residual live organisms remain.

In a study to evaluate the effectiveness of a cleaning and sanitizing method, a challenge test was performed by adding an *Escherichia coli* homogenate containing approximately 10^9 colony forming units (cfu) per milliliter to two STREAMLINE expanded bed columns. One column contained 250 mL of STREAMLINE SP and the other 250 mL of STREAMLINE DEAE. Sodium chloride cleaning solutions of increasing concentrations from 0.1 to 0.5 M were added, and the reduction of living microorganisms measured. As seen in Figure 1, the relative reduction in living cells is approximately 10^5. The increase in the number of cells in the eluate from the DEAE illustrates the necessity to understand the nature of the impurity to be removed and its interaction with the component being cleaned. In this case, the increase occurs when the increased sodium

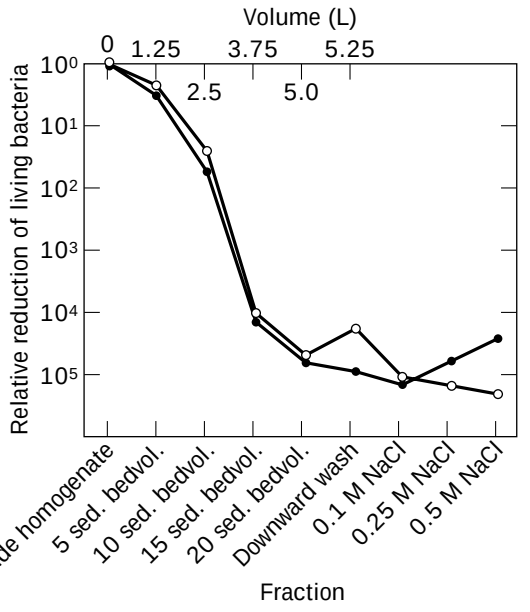


Figure 1. Relative reduction of living *E. coli* cells for STREAMLINE DEAE (●) and SP (○) in relation to the volume of the washing solution. *Source*: With kind permission from Kluwer; Ref. 9.

chloride concentration releases negatively charged *E. coli* cells from the positively charged DEAE (9).

Chromatography. A robust chromatographic process provides the purity required for proteins, peptides, oligonucleotides, plasmids, viral vectors, etc. But it is essential to select equipment and media that can be cleaned in place.

Most of today's chromatographic media can tolerate rather harsh conditions, such as exposure to sodium hydroxide. In affinity chromatography, the ligand usually determines the cleaning solutions that can be used. Protein A, for example, has been shown to withstand even sodium hydroxide (10). On the other hand, with immobilized monoclonal antibodies, it is often necessary to protect the column by using only clarified, partially purified feed. The cost of preparing a monoclonal antibody for use as an immunoadsorbent used in the purification of a therapeutic product is extraordinarily high, because the immobilized antibody must have the same purity as the therapeutic itself (11). One example of the highly successful use of such a column is described by Feldman et al. for the purification of factor VIII from human plasma (12).

As with other unit operations, specific impurities must be addressed in designing the cleaning regimes to ensure that there is no residue carried over from batch to batch of product. Clearly, efficient cleaning is even more important the closer one gets to final product purity. All cleaning procedures, contact time, and temperature must be specified. To maintain column lifetime, the least harsh method that provides a safety window should be implemented.

From a typical bacterial fermentation, one might have to remove soluble and precipitated proteins, hydrophobic proteins and lipids, nucleic acids, viruses, and endotoxins.

Proteins and lipids. Soluble proteins are the easiest to remove. For ion exchange, gel filtration, and most affinity chromatography media, a high salt wash (e.g., 1 to 2 M NaCl) is most common. For hydrophobic interaction chromatography (HIC), low-ionic-strength buffers or water are usually sufficient. In some cases, a detergent must be used, but this approach is more costly and requires validation of detergent removal.

Precipitated proteins frequently require up to 1 M sodium hydroxide or 1 M acetic acid. This is often followed by a water rinse and then a salt wash (except for HIC). Like precipitated proteins, hydrophobic proteins usually require sodium hydroxide. Other approaches include the use of ethanol, detergents, and a combination of cleaning agents. For example, ethanol (20%) in 1 M sodium hydroxide has been used at 40 °C for 30 to 60 minutes to remove host cell proteins from an ion exchanger.

For lipids, very little information is available. Commonly used solubilizing agents may destroy some chromatography media. Detergents and organic solvents (i.e., ethanol and acetonitrile) have been used with some success, but it is necessary to be aware that explosion-proof areas may be required for organic solvents. For those already working in such an area, this does not present a problem, but for firms used to working only in aqueous conditions, this may increase costs significantly. Furthermore, solvent disposal or recycling must be implemented. Removal of lipids before chromatography is strongly recommended to alleviate these concerns and added expenses.

Nucleic Acids. Nucleic acids bind tightly to anion exchangers. Some earlier work to evaluate their removal was performed by loading radiolabeled calf thymus DNA onto a DEAE Sepharose® Fast Flow column. Combinations of sodium chloride and sodium hydroxide were insufficient to totally remove all of the DNA. More recently, the ability of different cleaning agents to remove nucleic acids from a quaternary amine anion exchanger was tested using three different samples (13). The data from two of the samples are shown in Table 3 and 4. It is interesting to observe that for the monoclonal antibody (Table 4), DNase was required to thoroughly remove all traces of nucleic acids.

Viruses. Regulatory agencies have expressed concern about the retention of viruses on columns. As stated at the International Conference on Harmonization Guideline on Viral Safety, "Assurance should be provided that any virus potentially retained by the production system would be adequately destroyed or removed prior to reuse of the system. For example, such evidence may be provided by demon-

Table 3. Mass Balance of Calf Thymus DNA Clearance from Q-Sepharose® Fast Flow

Fraction tested	μg DNA	% of total DNA
Starting material	699.75	100
Buffer A wash	0.79	0.11
100% B eluate	253.80	36.27
Acid wash	0.64	0.09
1 M NaOH/1 M NaCl	449.60	64.25
Water wash	3.92	0.56
Total DNA eluted	708.75	101.2

Table 4. Mass Balance of Total DNA Clearance from a Monoclonal Antibody on Q-Sepharose Fast Flow

Fraction tested	μg DNA	% of total DNA
Total DNa in MAb	481.50	100
Flow through	0.31	0.06
Wash	0.12	0.02
Peak 1	10.00	2.08
Peak 2	270.68	56.21
100% B wash	6.08	1.26
Acid wash	4.00	0.83
1 M NaOH/1 M NaCl	64.00	13.29
2 M NaOH	3.21	0.67
3 M NaCl	0.65	0.13
Water wash	0.29	0.06
Subtotal	359.34	74.61
Dnase 1 treatment	125.00	25.96
Total DNA eluted	485.9	100.6

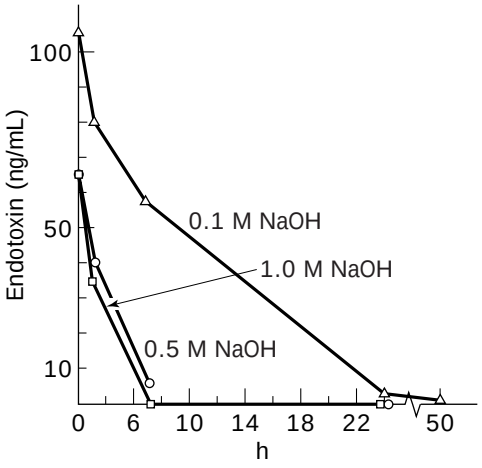


Figure 2. Inactivation of *E. coli* endotoxin by sodium hydroxide.

strating that the cleaning and regeneration procedures do inactivate or remove virus" (14). Table 5 shows work from Q-One Biotech summarizing the inactivation of eight different viruses by 0.1 M and 0.5 M NaOH. The kinetics of inactivation, which must be taken into account for specifying contact time, are also shown.

Endotoxins. Like nucleic acids, endotoxins are highly negatively charged and may bind tenaciously to anion exchangers. Sodium hydroxide has been shown to be very effective in inactivating endotoxins. Figure 2 shows the inactivation of *E. coli* endotoxin in solution at room temperature. Although a contact time of up to 5 hours with 1 M NaOH is indicated here to inactivate endotoxin, the added action of flow may enhance removal. In a study on a chromatography system challenged with a solution containing 1,200 endotoxin units (EU) per milliliter, it was found that circulating 1 M NaOH for 60 minutes was sufficient to inactivate endotoxin in the both the system and the packed column of a quaternary amine anion exchanger that was placed in line. In addition to the *Limulus* amebocyte lysate (LAL) test for endotoxin, a total organic carbon assay was used to analyze the rinse fluids for any carbon-containing residues remaining from the endotoxin. No residue was detected (15).

Combining Cleaning and Sanitization Methods. The ability to remove microorganisms from chromatography media is an issue that should be addressed in development. Depending on the microorganisms, they may produce endotoxins, enterotoxins, spores, and proteases—all of which could adulterate a product. Sanitization of chromatography columns is readily accomplished with sodium hydroxide (16). Combining cleaning and sanitization is the most cost-effective approach. In many cases, addition of sodium chloride to a sodium hydroxide solution provides an efficient, simple method that is easy to validate for removal of bioburden (tested by routine monitoring) and impurities from the feedstream. The importance of compatibility of the wetted materials with such combinations should be considered at the design phase.

Formulation and Final Filling. Cleaning issues in formulation and filling operations are addressed in depth elsewhere. Some relevant texts include those by Olson and Groves (17) and Cole (18). A recent publication that ad-

Table 5. Virus Titles and Inactivation values (log_{10})

		HIV	BVD	CPV	BHV	POL	SV-40	MLV	ADV
0.1M NaOH	Spike	2.0×10^8	9.5×10^5	2.0×10^9	6.9×10^9	7.1×10^8	1.7×10^8	2.6×10^5	2.2×10^8
	t = 0 min	[b]*5.9 × 10²*	2.7×10^7	1.9×10^3	1.2×10^2	3.5×10^4	1.5×10^5	*3.7 × 10¹*	1.7×10^2
	t = 10 min	*5.7 × 10²*	2.7×10^5	2.4×10^3	1.5×10^1	2.7×10^3	3.6×10^5	*3.8 × 10¹*	6.0×10^1
	t = 20 min	*5.8 × 10²*	1.5×10^4	9.6×10^2	4.5×10^1	2.0×10^4	4.7×10^4	*4.0 × 10¹*	6.3×10^1
	t = 60 min	*5.8 × 10²*	2.7×10^4	5.0×10^3	4.5×10^1	2.1×10^3	2.0×10^4	*4.3 × 10¹*	*2.9 × 10¹*
Inactivation (log_{10})		>3.5	2.5	5.6	8.2	5.5	3.9	>3.8	>6.9
0.5M NaOH	Spike	2.0×10^6	9.5×10^6	2.0×10^9	6.9×10^9	7.1×10^8	1.7×10^8	2.6×10^5	2.2×10^8
	t = 0 min	*5.7 × 10²*	1.9×10^4	9.4×10^2	*5.9 × 10¹*	1.1×10^5	1.5×10^5	*6.3 × 10¹*	9.4×10^1
	t = 10 min	*5.6 × 10²*	1.3×10^2	1.2×10^3	*5.9 × 10¹*	1.1×10^5	1.7×10^3	*4.7 × 10¹*	7.5×10^1
	t = 20 min	*5.6 × 10²*	*1.7 × 10²*	1.5×10^3	*5.9 × 10¹*	2.0×10^4	8.4×10^3	*4.7 × 10¹*	*2.0 × 10¹*
	t = 60 min	*6.7 × 10²*	*1.7 × 10²*	1.5×10^3	*5.9 × 10¹*	6.2×10^3	1.0×10^2	*5.5 × 10¹*	*2.2 × 10¹*
Inactivation (log_{10})		>3.5	>4.7	6.1	>8.1	5.1	6.2	>3.7	>7.0

Notes: Virus titers expressed in $TCID_{50}$ units for all viruses except BHV and MLV (expressed in pfu). Italicized titers: No virus was detected in these samples and values listed are theoretical minimum detectable titers.

HIV = human immunodeficiency virus; BVDA = bovine viral diarrhea virus; CPV = canine parvovirus; BHV = bovine herpes virus; POL = poliovirus; SV-40 = simian virus-40; MLV = murine leukemia virus; ADV = adenovirus.

dresses reducing pyrogens in cleanroom wiping materials provides valuable information on cleaning (19).

METHODS FOR EVALUATING CLEANLINESS

Several methods for evaluating cleanliness were discussed previously in conjunction with the different unit operations. There are three basic methods to evaluate cleanliness: visual inspection, assaying rinse water, and assaying surfaces by swabbing. The nature of the unit operation and equipment design frequently determine which method will be most appropriate. A combination of methods is usually necessary for validation. Visual inspection may be quite useful as a qualitative assessment of cleanliness for equipment such as tanks. As noted by several regulators, "you can always see the ring around the bathtub." However, more sensitive, quantitative methods are also required. For multiproduct facilities, it is absolutely necessary to use product-specific assays. Product-specific assays are also used for dedicated facilities.

Rinse water can be collected after cleaning and examined for residues. In many cases, the quality of the rinse water is compared to WFI that has not been run through the component being tested. Other tests commonly run on rinse water (3) include LAL for endotoxin, total organic carbon (TOC) (20,21), protein assays, sodium dodecylsulfate–polyacrylamide gel electrophoresis (SDS-PAGE), spectrophotometric analysis, and product-specific assays such as ELISAs.

Swabbing is performed by using non-carbon-containing materials that do not interfere with subsequent testing but from which swabbed soils can be extracted. The swabs themselves should not generate particles, background carbon, or nonvolatile residues. Swabbing of surfaces allows evaluation of the cleanliness of areas in which non-water-soluble residues might remain. However, there may be components or areas of systems that are hard to reach by swabbing. In this case, a soil can be applied to pieces of the equipment, called coupons, that are cut out or provided by a vendor. The soil is allowed to dry, and then the coupon is swabbed after cleaning. Coupons can be useful in designing the cleaning protocol. The coupon can be placed inside equipment and cleaned during a normal cleaning cycle. Sometimes, the coupon and the soil are heated to simulate a difficult cleaning problem. Cleaning methods can be tested on the baked-on soil. After the swabbing is performed, the swab is extracted, and the extract is analyzed by assays such as TOC, HPLC, and ELISA.

A survey of five U.S. biotechnology companies revealed which assays are commonly performed and how they are used at different stages of manufacturing (22). The report presents the purpose of each test, detection limits, acceptance criteria, and methodology in a concise tabular form.

SETTING LIMITS

How clean is clean? As analytical methods become more sensitive, the likelihood of finding something is increased. Attempts to define acceptance limits of cleanliness have been made by Fourman and Mullen (23) and Zeller (24) and discussed in Reference 3. Fourman and Mullen suggested that "No more than 10 ppm of any product should appear in another product" and "No more than 0.001 dose of any product will appear in the maximum daily dose of another product." But clearly, a risk analysis should be performed to assess the potential hazard for each product.

Whereas for some therapeutics these low levels suggested by Fourman and Mullen may be accepted by regulatory authorities, for many multiproduct facilities, "not detectable" may be required. On the other hand, for diagnostic products, the real issue should be whether the product is consistent and whether the carryover affects its accuracy.

The acceptance criteria for upstream unit operations such as fermentation may not be as stringent as those for final purification and processing steps. Typical acceptance criteria for CIP of fermenters has been discussed by Naglak et al. (25).

In addition to setting specifications for residual impurities derived from the product feedstream and final product, it is necessary to consider the removal of cleaning agents. Analytical strategies for organic cleaning agent residues have been addressed by Gavlick et al. (26). The questions: "What is the level of detergent residue that would be acceptable to the Food and Drug Administration (FDA)? What is the basis for arriving at this level, if any?" were answered by the FDA's Office of Compliance as follows: "FDA has repeatedly stated that it is the firm's responsibility to establish acceptance limits and be prepared to provide the basis for those limits to FDA. Thus, there is no fixed standard for levels of detergent residue. Any residues must not adversely alter drug product safety, efficacy, quality, or stability" (27).

CLEANING VALIDATION

Although it is not a regulatory requirement, a master plan for cleaning validation can be quite valuable. The master plan will provide a time frame for performing the work, prioritizing tasks, designating responsibilities, describing protocols, and listing all facility, equipment, and process validation requirements. Although the time frame will most likely change, the plan will ultimately save time and may prevent regulatory delays by ensuring that no critical items are overlooked. A final validation report should be signed by management.

Agalloco described a series of questions that provide insight into cleaning validation (28): What items are being removed? How are samples taken? What analytical tests are utilized? When shall cleaning validation be performed? Which physical parameters shall be evaluated? How much residual will be permitted? These questions should be addressed during the preparation of the master plan.

Two cleaning validation guides written by the U.S. FDA are very useful (29,30). Although the intent of these guides is to provide FDA investigators with information that results in consistency in inspections, there are many relevant points that should be considered by any firm that has to perform cleaning validation. For example, in the mid-

Atlantic version, it is stated that operators performing cleaning operations should be aware of problems and have special training, and the length of time between the end of processing and each cleaning step should be controlled. The documents further address issues such as equipment design, documentation, analytical methods, sampling, and establishment of limits. Although no one approach is required, the guidance provided in these documents should be followed or a reason why it was not adhered to provided.

One of the more difficult techniques to validate is swabbing. Some of the variables include operator swabbing technique, extraction efficiency from the swab, and recovery of residue from different wetted materials. In spite of its difficulties, this technique provides valuable information that cannot be obtained from visual inspection and rinse water analysis. Total protein analysis and total carbon analysis of swab samples for the cleaning validation of bioprocess equipment have been addressed by Strege et al. (31,32). Use of this technique for validating detergent residues has also been presented (33).

Most cleaning validation is performed at pilot or full scale, but for certain unit operations such as chromatography, the parameters that are to be validated can often be established with scaled-down models using production feedstream. Seely et al. scaled down a chromatography system and used cleaning conditions that were harsher than those used in normal production to determine if residues remained on the packed column (34). At the same time, production columns were also evaluated for cleaning efficacy by collecting relevant product fractions from blank runs and assaying them for UV280 absorbing material and endotoxin. In addition, reversed-phase HPLC, ELISA, and silver-stained SDS-PAGE assays were performed. No endotoxin or UV280 absorbing material was found. A trace amount of the product was found, but the level (about 0.3% of the amount normally collected) was thought not to represent a significant cycle-to-cycle carryover risk to the final product. Furthermore, the eluate after storage was evaluated by TOC for evidence of further cleaning effects of the storage solution.

Once cleaning is validated, it is much more convenient to monitor key process conditions such as cleaning agent concentration, pH, conductivity, temperature, contact time, and flow if applicable. Revalidation of cleaning may have to be performed when there is a change in a process or equipment, after maintenance or shutdown periods, or after area contamination. Each firm should establish the frequency of cleaning revalidation.

CASE STUDIES

Some case studies further illustrate the critical issues in cleaning and cleaning validation.

Validation of CIP of a Bioreactor

In a 1994 paper, Geigert et al. presented a case study on the CIP of a bioreactor (35). They addressed five key questions: What components need to be cleaned from the bioreactor? What is a scientifically sound and appropriate sampling plan? How will the test samples be handled prior to analysis? Are the test methods chosen scientifically sound and appropriate for the intended purpose? What is clean?

The components to be cleaned included active protein, host cell components, media ingredients, and cleaning agents. The sampling plan included swabbing of a few select surface sites, performing a worst-case cleaning challenge using artificial soil or reducing cleaning parameters, and evaluating liquid from the rinse. Both final rinse and earlier rinse samples were taken to show that the cleaning process was effective. Samples were evaluated for their stability during storage and handling. As an example, it was noted that hydrophobic proteins may adhere to certain plastic storage containers.

Several assays, both specific and nonspecific, were evaluated. Three assays were selected: SDS-PAGE, a bioassay, and pH. TOC was not yet set up at the facility; SDS-PAGE and the bioassay had already been validated by quality control and only needed to be further validated for the limit of detection for the cleaning studies. The absence of cleaning agent was determined by pH, and the absence of final product by the bioassay. Four protein products were analyzed by SDS-PAGE with silver staining. The limit of detection varied from 1 ppm to 10 ppm. The bioassay was performed for the three products that might be campaigned in a single bioreactor. Two products were from yeast; one was from *E. coli*. The two yeast products could be detected at levels as low as 4 ppb, but the *E. coli* product, which was very hydrophobic, aggregated. This study illustrates that the test procedures to evaluate cleanliness are critical in supporting the validation work. By using validated test methods, the firm was able to provide a high degree of assurance that a 1,600-L bioreactor could be cleaned sufficiently between product campaigns, thereby avoiding the additional costs of purchasing costly dedicated equipment.

Changeover of Equipment from Development to a Dedicated Unit

In this example, a firm had used a pilot-scale chromatography unit for development work in the purification of monoclonal antibodies (IgA and IgM) (36). To avoid the expense of purchasing a new unit, the firm decided to use this system for the manufacture of clinical trial material. This required extensive cleaning validation. Because the unit contained tubings and threaded fittings, it was decided that replacing low-cost wetted components was less expensive than validating their cleaning. After the new components were installed, the cleaning of the system was validated by demonstrating absence of carryover on previously exposed sites and by using new surface sites to establish a baseline for future cleaning validation. Samples were taken by swabbing and collecting rinse water, and samples were analyzed by both TOC and ELISAs. The IgM ELISA has a limit of detection of 25 ng/mL, and the IgA ELISA has a limit of 4 ng/mL. The results showed that the 16 used surfaces sampled were below the detection levels of the ELISAs.

CIP Validation in a Multiproduct Facility

A multiproduct biopharmaceutical contract manufactur-

ing facility must address cleaning and cleaning validation to satisfy its own requirements as well as the requirements of their customers and regulators. Sherwood has presented a description of CIP experiences in such a facility (37). The facility manufacturers monoclonal antibodies and recombinant proteins by mammalian cell culture. The author points out that "Each company must determine their own cleaning policies and acceptance criteria based upon knowledge of the equipment and processes employed."

Visual inspection, final rinse water, and swabbing of surface samples were used to evaluate CIP of fermenters and purification holding vessels. Because visual inspection is subjective, two observers are always used. For the rinse water samples, sample containers are single use, and sample storage temperatures and duration were validated. The following tests were performed on rinse water: pH and conductivity, nitrate and protein concentration, bioburden, TOC, endotoxin, and an enzyme immunoassay (EIA). The EIA was only performed for the purification holding vessels. In general, nonspecific assays are used for screening both product and manufacturing materials in fermentation, whereas in the purification area more specific assays are used. For the swab samples, TOC and protein concentration were performed for the fermentation, and in the purification area, the EIA was added.

Acceptance criteria are more difficult to establish in such a facility because the final product dose is not known. Acceptance limits have to be based on analytical method performance criteria, pharmacopeial requirements, and historical data.

The usefulness of performing cleaning validation was illustrated by the performance qualification of the CIP of the fermenter vessel. In one run, a sight glass swab showed a high TOC result. After investigation, a manual cleaning of the sight glass was implemented.

For the initial performance qualification, three consecutive CIP runs were monitored and found to meet acceptance criteria. On an annual basis, one CIP run is fully monitored. The author noted that ongoing monitoring in such a contract manufacturing facility may also require final rinse and swab sampling.

SUMMARY

Cleaning and cleaning validation require sound scientific judgment. Cleaning is not just an afterthought, but should be designed into a process. Today there is a trend to use contract manufacturing, and most of these facilities will manufacture more than one product. Cleaning validation is particularly critical for such facilities. Regulatory agencies are spending more time investigating cleaning and cleaning validation. Ultimately, adequate cleaning leads to enhanced product quality, fewer batch failures, and hence better profitability.

BIBLIOGRAPHY

1. *Guidelines on the Preparation of Investigational New Drug Products*, U.S. Food and Drug Administration, Rockville, Md., March 1991.
2. Elimination of Establishment License Application for Specified Biotechnology and Specified Synthetic Biological Products. *Fed. Regist.* **61**, 24227–24233 (1996).
3. PDA Biotechnology Cleaning Validation Subcommittee, *Cleaning and Cleaning Validation: A Biotechnology Perspective*, PDA, Bethesda, Md., 1996.
4. D. Rohsner and W. Serve, *Pharm. Eng.* **15**, 20–28 (1995).
5. D.A. LeBlanc, D.D. Danforth, and J.M. Smith, *Pharm. Technol.* **17**, 84–92 (1993).
6. Y. Chisti and M. Moo-Young, *J. Ind. Microbiol.* **13**, 201–207 (1994).
7. D.J. Brose and P. Waibel, *Pharm. Technol.* **20**, 48–52 (1996).
8. S.L. Michaels, *BioPharm* **7**, 38–45 (1994).
9. R. Hjorth, S. Kåmpe, and M. Carlsson, *Bioseparation* **5**, 217–233 (1995).
10. G. Hale, A. Drumm, P. Harrison, and J. Phillips, *J. Immunol. Methods* **171**, 15–21 (1994).
11. Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products for Human Use, U.S. Food and Drug Administration, Rockville, Md., 1994.
12. F. Feldman, S. Chandra, M. Hrinda, A. Schreiber, in *Quality Assurance in Transfusion Medicine, Vol II. Methodological Advances and Clinical Aspects*, CRC Press, Boca Raton, Fla., 1993, pp. 259–284.
13. Y. Dasarathy, *BioPharm* **9**, 41–44 (1996).
14. International Conference on Harmonization; Guidance on Viral Safety Evaluation of Biotechnology Products Derived from Cell Lines of Human or Animal Origin, *Fed. Regist.* **63**, 51074–51084 (1998).
15. Pharmacia Biotech, *Depyrogenation of a BioProcess™ Modular System and a Packed BPG Column with Sodium Hydroxide*, Technical Note 216, Research Report, 1996.
16. G. Sofer and L. Hagel, *Handbook of Process Chromatography: A Guide to Optimization, Scale up, and Validation*, Academic Press, London, 1997.
17. W.P. Olson and M.J. Groves, *Aseptic Pharmaceutical Processing*, Interpharm Press, Prairie View, Ill., 1987.
18. G.C. Cole, *Pharmaceutical Production Facilities Design and Applications*, Ellis Horwood, New York, 1990.
19. D.W. Cooper, *Pharm. Eng.* **16**, 51–61 (1996).
20. R. Baffi, G. Dolch, R. Garnick, Y.F. Huang, B. Mar, D. Matsuhiro, B. Niepelt, C. Parra, and M. Stephan, *J. Parenteral Sci. Technol.* **45**, 13–19 (1991).
21. K.M. Jenkins, A.J. Vanderwielen, J.A. Armstrong, L.M. Leonard, G.P. Murphy, and N.A. Piros, *PDA J. Pharm. Sci. Technol.* **50**, 6–15 (1996).
22. P.R. McArthur and M. Vasilevsky, *Pharm. Eng.* **15**, 24–31 (1995).
23. G.L. Fourman and M.V. Mullen, *Pharm. Technol.* **17**, 54–60 (1993).
24. A.O. Zeller, *Pharm. Technol.* **17**, 70–80 (1993).
25. T.J. Naglak, M.G. Keith, and D.R. Omstead, *BioPharm* **7**, 28–36 (1994).
26. W.K. Gavlick, L.A. Ohlemeier, and H.J. Kaiser, *Pharm. Technol.* **19**, 136–144 (1995).
27. Division of Manufacturing and Product Quality, HFD-320, Office of Compliance, Center for Drug Evaluation and Research, *Human Drug CGMP Notes* **3**, 1995.
28. J. Agalloco, *J. Parenteral Sci. Technol.* **46**, 163–168 (1992).
29. Mid-Atlantic Region Inspection Guide, Cleaning Validation, U.S. Food and Drug Administration, Rockville, Md., July 28, 1992.

30. Division of Field Investigations, Guide to Inspections of Validation of Cleaning Processes, U.S. Food and Drug Administration, Rockville, Md., July 1993.
31. M.A. Strege, J.J. Dougherty, R. Green, and A.L. Lagu, *BioPharm*. **7**, 40–42 (1994).
32. M.A. Strege, T.L. Stinger, B.T. Farrell, and A.L. Lagu, *BioPharm* **9**, 42–45 (1996).
33. J.M. Smith, *Pharm. Technol.* **16**, 60–66 (1992).
34. R.J. Seely, H.D. Wight, H.H. Fry, S.R. Rudge, and G.F. Slaff, *BioPharm* **7**, 41–48 (1994).
35. J. Giegert, R. Klinke, K. Carter, and A. Vahratian, *PDA J. Pharm. Sci. Technol.* **48**, 236–240 (1994).
36. P. Gariepy, R. Moore, and G. Sofer, *Biopharm* **11**, 32–36 (1998).
37. D. Sherwood, *PDA Fourth Int. Congress*, Vienna, Austria, February 19–23, 1996.

See also ASTM STANDARDS FOR BIOTECHNOLOGY; CHROMATOGRAPHY, COMPUTER-AIDED DESIGN; GOOD MANUFACTURING PRACTICE (GMP) AND GOOD INDUSTRIAL LARGE SCALE PRACTICE (GLSP); PILOT PLANTS, DESIGN AND OPERATION; PROCESS VALIDATION.

CLOSTRIDIA, SOLVENT FORMATION

PALMER ROGERS
University of Minnesota
Minneapolis, Minnesota

KEY WORDS

Biochemistry
Clostridium
Clostridium genetics
Physiology
Solvent fermentations
Solvent production
Strain development

OUTLINE

INTRODUCTION

The general biology of the genus *Clostridium* has been treated recently in detail by Hippe et al. (1). More than 100 species of bacteria fall into this group, probably because only four simple criteria need be met for membership: *(1)* ability to form heat-resistant endospores, *(2)* obligate requirement for anaerobic energy metabolism; *(3)* inability to reduce sulfate as a final electron acceptor, and *(4)* a gram-positive type of cell wall. The extreme heterogeneity of this huge "genus" has become more and more evident over the past 20 years. First, the GC content of the DNA of more than 100 species of *Clostridium* ranges between 21 and 54 mol %. Second, the 16S rRNA gene sequences have been determined for most of these species together with their close relatives. A recent analysis of these data indicates that the genus *Clostridium* is in fact mixed in with some of the closely related bacteria to form several very divergent "clusters", providing a strong basis for a taxonomic revision (2).

The largest rRNA homology group, which is newly designated cluster I (2), contains all the butanol-producing species of *Clostridium* (Table 1) (2–7). Fifty-five strains classified as *Clostridium acetobutylicum* in national collections were recently examined by means of DNA–DNA hybridization (3) and a combination of DNA fingerprinting and 16S RNA gene sequence analysis (4). The results show that these strains fall into four taxonomic groups. These taxa, or species, are listed in Table 1, together with three other butanol-producing species. Also listed in Table 1 are five ethanol-producing clostridia, which are considered important because of their potential to produce ethanol from fermentation of specific substrates such as crystalline cellulose or at a high temperature. Reclassification suggests that these "clostridia" fall into three or four distinct phylogenetic clusters (2). *C. thermocellum* is the most thoroughly studied ethanol-producing cellulolytic clostridial species. It is grouped with seven other cellulolytic species in cluster III (2) or group E (5). *C. thermohydrosulfuricum* strains have recently been divided into two groups and reclassified as *Thermoanaerobacter thermohydrosulfuricus* or *T. ethanolicus* (6) and placed in cluster V (2) or group A (5). *C. thermosaccharolyticum* most likely belongs in the new genus *Thermoanaerobacterium* (6) in cluster VII (2) or group C (5). The phylogenetic position of *C. saccharolyticum* has not been determined (7). Table 1 also includes strain numbers and culture collection numbers that appear most frequently in the literature dealing with solvent production by these clostridia.

The major role of the solvent- and acid-forming saccharolytic clostridia in nature is the degradation of organic materials, namely the biopolymers starch, cellulose, hemicelluloses, and proteins. Utilization of these polymers requires that these bacteria produce and secrete enzymes such as amylases, cellulases, proteinases, and xylanases. Some of these enzymes are of potential importance to commercial processes (8). The anaerobic fermentation of sug-

Table 1. Clostridial Species Important for Solventogenic Fermentations

Former name(s)	Proposed (new) name	Ref.
	Ethanol-producing species	
C. thermocellum		
LQRI, ATCC 35609, DSM 2360	No change (cluster III)	2,5
C. thermohydrosulfuricum		
E100-69^T, ATCC 35045, DSM 567	*Thermoanaerobacter thermohydrosulfuricus* (cluster V)	2,5,6
C. thermohydrosulfuricum, Thermoanaerobacter ethanolicus		
JW200^T, 39E, ATCC 33223	*Thermoanaerobacter ethanolicus* (cluster V)	2,5,6
C. thermosaccharolyticum	Misclassified, belongs in genus *Thermoanaerobacterium*	
NCA 3814, ATCC 7956	(cluster VII)	2,5,6
C. saccharolyticum		
NRCC 2533, ATCC 35040	No change	7
	Butanol-producing species[a]	
C. acetobutylicum		
ATCC 824, 4259 Weizmann strain; DSM 792, and 8 strains	No change, "taxon I"	3,4
C. acetobutylicum		
NRRL B643, NCP 262, and four NCP strains	*C* (new species) "taxon II"	3,4
"*C. saccharoperbutylacetonium*"		
NI-4, NI-504, ATCC 27021	No change, "taxon III"	3,4
C. beijerinckii, C. acetobutylicum		
ATCC 25732^T, NRRL B 592, NCIMB 9362^T, 8052^T, and 19 strains	*C. beijerinckii*, "taxon IV"	4
C. puniceum		
NCIMB 11596	No change; closely related to taxa II and III	4
C. aurantibutyricum		
ATCC 17777	No change	4
C. tetanomorphum		
MG-1, ATCC 49273, DSM 4474	No change	4

Note: Culture collections: ATCC, American Type Culture Collection, Rockville, Md.; DSM, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Göttingen, Germany; NCIMB, National Collection of Industrial and Marine Bacteria, Aberdeen, Scotland; NRRL, Northern Utilization and Development Division, Peoria, Ill.; NI, Nagao Institute, Tokyo, Japan; NRCC, National Research Council, Ottawa; NCA, National Canner's Association, Washington, D.C.; NCP, Department of Microbiology, University of Capetown, Rondebosch, South Africa.

[a]All butanol-producing species fall into cluster I(2).

ars from digestion of polymers or other sources is called a saccharolytic fermentation. The major products that accumulate are organic acids, neutral solvents, CO_2, and H_2. The solvents and acids are listed in Table 2 for typical fermentations by 11 solvent-producing *Clostridium* species (9–19). Production of CO_2 and H_2 gases is always high in these fermentations (see Table 2 note *a*); however, the exact amount produced is not listed for each species. In a natural setting these clostridia occur in conjunction with a consortium of other anaerobic bacteria, which utilize these fermentation products as rapidly as they are formed for their own energy and growth needs. In this way most of the organic starting materials are mineralized. Only a few isolated strains of *Clostridium* produce significant amounts of solvents and hence are potential candidates for industrial fermentations.

The reason for the high level of interest in basic research and development of fermentation technology for solvent production by the clostridia is straightforward. First, although there is no immediate threat that the world's petroleum reserves will be used up, a severe dwindling of these nonrenewable resources, which serve as the major source of fuels and chemicals, will most likely occur in the mid-twenty-first century (20). Second, it is now technologically possible to produce essentially all commodity chemicals from renewable biomass feedstocks such as starch and cellulose. Third, from about 1915 until the mid-1940s, fermentations employing *Clostridium acetobutylicum* made a significant impact on the commercial solvents industry. Typically, more than 1000 kg of butanol, 500 kg of acetone, and 175 kg of ethanol plus other utilizable byproducts were produced in 90,000-L fermentors in North America and elsewhere (21).

From about 1925 to 1980 research on the physiology and biochemistry of the clostridial fermentations was centered on the unraveling of the metabolic pathways essential to our current understanding of the biology of solvent formation (8). Over the past 10 years, basic research on clostridial solventogenesis has made great strides in the following areas: development of genetic methods applied to the solventogenic *Clostridium* species, regulation of pathways of solvent production, degradation and use of biopolymers, and relation of solvent production to developmental processes such as sporulation. This summary treatment of the biology of the solventogenic clostridia will emphasize our current understanding of the clostridial fermentation strategies, physiologic control and genetic mechanisms regulating solvent production, and future directions for development of new solvent-producing strains. The advantages and problems inherent in the biology of these bacteria relevant to their use in commercial solvent production also are presented.

Table 2. Ethanol and Butanol Fermentations by *Clostridium* Species

Organism	Substrates utilized	
	Complex polymers	Sugars
	Ethanol fermentations	
C. thermocellum	Cellulose, xylan	Cellobiose, few hexoses and pentoses
C. thermohydrosulfuricum (T. ethanolicus)	Starch, pectin, salicin, xylan	Many dissacharides, hexoses, pentoses, etc.
C. thermosaccharolyticum	Starch, xylan	Dissacharides, hexoses, and pentoses
C. saccharolyticum		Dissacharides, hexoses, and pentoses
	Butanol fermentations	
C. acetobutylicum taxon I (corn-starch-utilizing strains)	Starch, xylan	Cellobiose, lactose, some hexoses and pentoses
C. acetobutylicum taxon II		
C. saccharoperbutylacetonicum taxon III (molasses-utilizing strains)	Starch	Sucrose, cellobiose, some hexoses and pentoses
C. beijerinckii taxon IV	Starch	Sucrose, cellobiose, some hexoses and pentoses
C. aurantibutyricum	Starch	Sucrose, hexoses, and pentoses
C. puniceum	Starch, pectin	Hexoses, pentoses
C. tetanomorphum		Cellobiose, some hexoses and pentoses

[a]The gases CO_2 and H_2 are also produced in these fermentations. Production per 100 mmol of hexose CO_2, 150–220 mmol; H_2, 50–100 mmol.
[b]N, not reported.
[c]Amount of butyrate is variable depending on conditions.
[d]Acetone only is produced.
[e]Either isopropanol or acetone is produced, depending on the strain.
[f]Both isopropanol and acetone are produced.

PHYSIOLOGY AND BIOCHEMISTRY OF THE SOLVENT FERMENTATIONS

Clostridium's General Fermentation Strategy

Thauer et al. (22) outlined the general energy metabolism of the chemotrophic anaerobic bacteria, which include the clostridia. The clostridia depend almost exclusively on the fructose biphosphate pathway (Embden–Meyerhof pathway) for conversion of one hexose to two pyruvates with the net production of two ATPs and two NADHs (Fig. 1, pathway 1). In the fermentation of pentoses, the intermediates from 3 mol of pentose-5-phosphate, are one glyceraldehyde 3-phosphate and two fructose 6-phosphates. A combination of the enzymes transaldolase and transketolase is used (Fig. 1, pathway 6). The sugar-phosphates enter the fructose biphosphate pathway and have the capability of producing five ATPs and five NADHs per 3 mol of pentose fermented. The phosphorolytic 3-2 cleavage of xylulose-phosphate yielding acetyl-phosphate is not found in the clostridia or in other obligate anaerobes. Also, the clostridia do not have the enzyme pathway for oxidation of glucose 6-phosphate to NADPH, pentose 5-phosphate, and CO_2. However, high activities of glucose 6-phosphate dehydrogenase and of 6-phosphogluconate dehydrogenase were found to be induced in sporulating cells of *C. thermosaccharolyticum* (12). Possibly this pathway (Fig. 1, pathway 7) is induced during sporulation to provide extra NADPH for biosynthesis and has been overlooked in other clostridia. Yet, NADPH–ferredoxin oxidoreductase (Fig. 1, reaction 4), which is ubiquitous in all clostridia investigated (23), probably provides the NADPH for biosynthesis during both vegetative growth and sporulation. Most pyruvate produced from sugars during the clostridial fermentation is cleaved by pyruvate–ferredoxin oxidoreductase in a coenzyme A (CoA) dependent reaction yielding CO_2, acetyl CoA, and reduced ferredoxin (Fd_{red}), as shown in Figure 1, reaction 2. Acetyl CoA is central to all clostridial fermentations. The thioester bond of acetyl CoA is a very high energy bond; thus it is an important source of ATP in most clostridial fermentations, because for every mole of butyrate or acetate produced, a mole of ATP is formed (Fig. 2, reactions 2 and 6). Acetyl CoA is also the most important precursor of all alcohols and organic acids synthesized by these organisms: it thus serves as a major intermediate in these fermentations (Fig. 1). The NADH, which is formed during oxidation of glucose, is thus reoxidized and recycled so that the fermentation can continue.

The second key element in the clostridial fermentation strategy is the system that manages electron flux diagrammed in Figure 1, reactions 3–5. This system consists of ferredoxin and/or other iron–sulfur proteins that can accept or donate electrons at a very low potential ($E^{0\prime} = -0.41$ V) close to the hydrogen electrode; a ferredoxin–NADH oxidoreductase for passing electrons between NADH and Fd; a ferredoxin–NADPH oxidoreductase for controlled production of NADPH needed for biosynthesis of bacterial mass; and one or more hydrogenases permitting the use of protons as final electron acceptors, producing hydrogen gas, typical of most clostridial fermentations.

The electron distribution system functions in coordination with the branched fermentation pathways typical of clostridia, which eventually yield from two to five major acids and solvents (Table 2). The amounts of reduced versus neutral and oxidized products always are balanced with the amounts of H_2 and ATP produced, and there is a

Major fermentation products (mmol/100 mmol hexose)[a]						
Ethanol	Butanol	Acetone isopropanol	Acetate	Butyrate	Lactate	Ref.
		Ethanol fermentations				
85			31		26	9
194			11		2	10
109			75	2[c]	38	11,12
180			19		4	13
		Butanol fermentations				
6	55	28[d]	11	9		14
5	46	24[d]	3	8		15
N[b]	43	26[e]	3	4		16
N	42	25[f]	N	N		17
2	68	15[d]	16	5.5		18
43	47	0	23	5		19

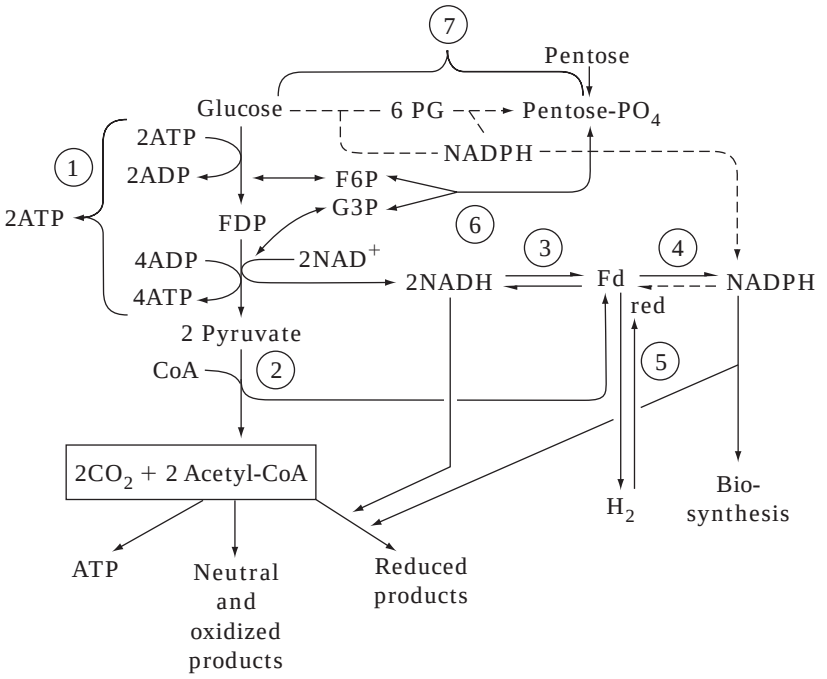


Figure 1. The clostridial fermentation strategy. Numbers indicate either pathways or enzymes as follows: 1, fructose biphosphate (Embden–Meyerhof) pathway; 2, pyruvate–ferredoxin oxidoreductase; 3, NADH–ferredoxin oxidoreductase and ferredoxin–NAD oxidoreductase; 4, NADPH–ferredoxin oxidoreductase; 5, hydrogenase; 6, transketolase + transaldolase; and 7, oxidative pentose phosphate pathway. *Source*: Reproduced with permission from Ref. 24, p. 6.

great deal of natural variation dependent on species and fermentation conditions (24). Figure 2 diagrams an example of this principle, in which *C. acetobutylicum* produces five organic fermentation products from acetyl CoA by utilizing three branch points in the fermentation pathways. Certainly, this variety of pathways available to the clostridia provides physiologic versatility and thus adaptability to different growth conditions. Presently, research is focused on elucidating the physiologic signals and details of molecular sensor and responder systems that control the direction of flow toward specific products in these fermentations (see "Regulation of Solvent Production"). The usefulness of these organisms will depend on directing these fermentations toward one or perhaps two major products. Another consequence of the biology of *Clostridium* (and perhaps all solvent-forming microbes) is that the yield of

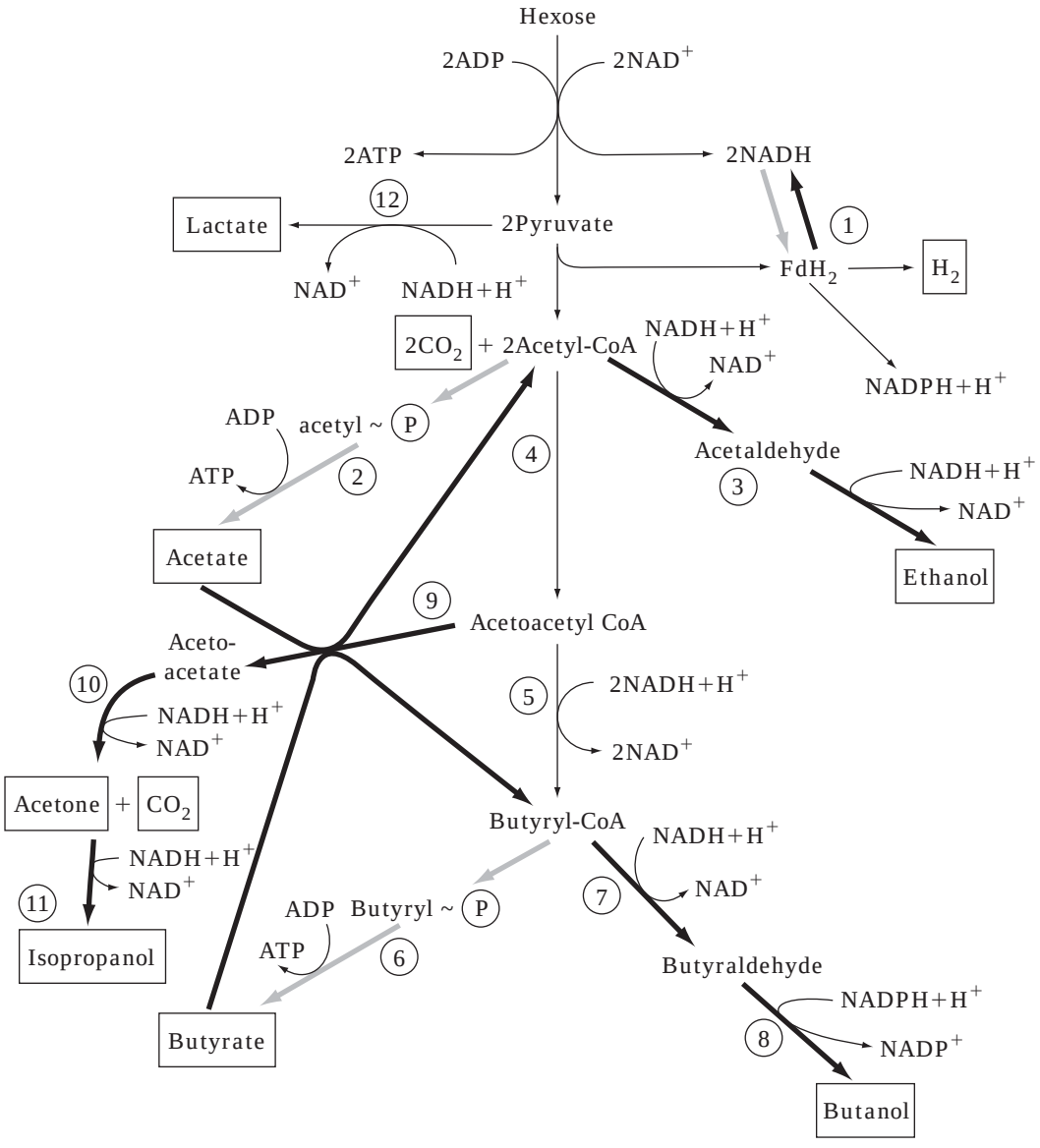


Figure 2. The butanol fermentations. Numbers indicate enzymes as follows: 1, three enzymes that exchange electrons with reduced ferredoxin (FdH_2), NAD^+: and $NADP^+$: ferredoxin oxidoreductases, and hydrogenase; 2, phosphotransacetylase and acetate kinase; 3, acetaldehyde dehydrogenase and ethanol dehydrogenase; 4, acetyl CoA acetyltransferase (or thiolase); 5, three enzymes producing butyryl CoA; 6, phosphotransbutyrylase and butyrate kinase; 7, butyraldehyde dehydrogenase; 8, butanol dehydrogenase; 9, acetoacetyl CoA: acyl CoA transferase; 10, acetoacetate decarboxylase; 11, isopropanol dehydrogenase; 12, lactate dehydrogenase. Thick gray arrows indicate reactions that predominate during acidogenesis; thick black arrows indicate reactions that predominate during solventogenesis. *Source*: Reproduced with permission from Ref. 25, p. 26.

organic solvents and acids formed is limited to only two-thirds of the substrate sugars fermented. As diagrammed in Figures 2 and 3, all fermented hexoses are funneled into CO_2, which is expelled, and acetyl CoA, which is the precursor of solvents and organic acids. Thus, only two-thirds of the substrate carbon is recovered as useful organic products. Exceptions to this one-third loss of productivity are the fermentations of sugars to produce two lactic acids by *C. thermolacticum* and the fermentation of sugars to produce three acetic acids by *C. thermoaceticum* (25).

Ethanol Fermentations

Tables 1 and 2 list three thermophilic clostridia, *C. thermohydrosulfuricum*, *C. thermocellum*, and *C. thermosaccharolyticum*, that produce significant amounts of ethanol together with acetate and some lactate during growth at 55–70 °C (26). *C. saccharolyticum*, first described in 1982, is a mesophile (37 °C optimum for growth) that grows on 20 different sugars and sugar alcohols including cellobiose, but it does not utilize starch or cellulose (7). In contrast,

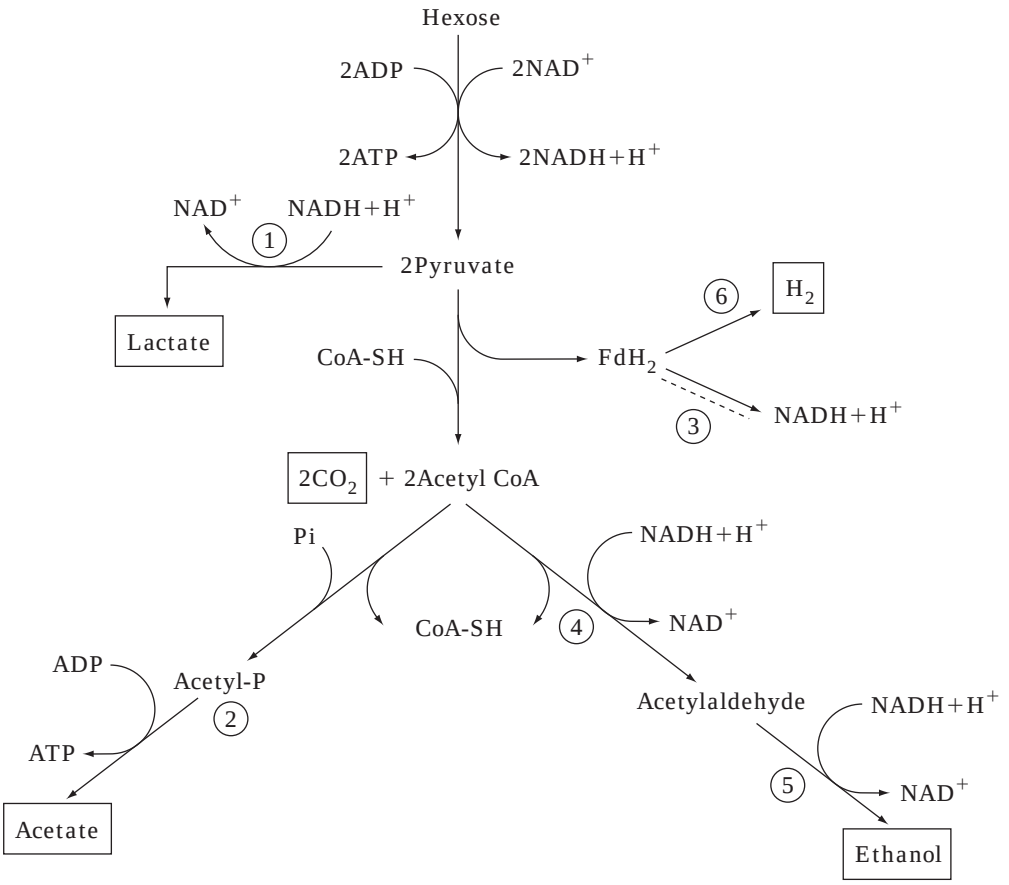


Figure 3. Clostridial ethanol fermentations. Numbers indicate enzymes as follows: 1, lactate dehydrogenase; 2, phosphotransacetylase + acetate kinase; 3, NADH: ferredoxin oxidoreductase and ferredoxin: NAD oxidoreductase; 4, acetaldehyde dehydrogenase; 5, ethanol dehydrogenase, and 6, hydrogenase. *Source*: Reproduced with permission from Ref. 25, p. 35.

C. thermocellum, which will utilize crystalline cellulose and hemicelluloses as feedstocks, is an excellent candidate for future bioconversion of cellulosic biomass to solvents. All these thermophilic clostridia utilize starch but are more limited in their ability to ferment sugars than is *C. saccharolyticum*. During the ethanolic fermentation of hexoses, the bulk of the pyruvate is converted to acetyl CoA, reduced ferredoxin (FdH_2) and CO_2 (Fig. 3). Depending on fermentation conditions, some of the pyruvate is diverted to lactate by lactate dehydrogenase. In one branch of the major fermentation pathway, acetyl CoA is reduced to acetaldehyde and then to ethanol by means of two NAD^+-linked dehydrogenases (Fig. 3, reactions 4 and 5). Acetate is formed via a second branch yielding a stoichiometric amount of ATP as in other clostridial fermentations (Fig. 3, reaction 2). Using this three-branched pathway, *C. thermocellum* produces about 0.6–1.0 mol ethanol/mol hexose, whereas the *C. thermohydrosulfuricum* and *C. saccharolyticum* fermentations yield 1.8–1.9 mol ethanol/mol hexose, which makes the latter two species attractive ethanol production organisms (Table 2).

The biochemical basis for the different product ratios produced by different species using the same branched pathway (Fig. 3) remains unknown, although the problem has been studied indirectly. Batch cultures of *C. thermosaccharolyticum* subjected to slow feeding of glucose show restricted growth, producing excess ethanol (1.0 mol/mol glucose) and little acetate or butyrate. On the other hand, rapidly growing cultures of this bacterium produce mostly acetate, lactate, and butyrate (12). Batch fermentation of cellulose by *C. thermocellum* yielded ethanol as the major product for the first two days, while lactate and acetate accumulated from days 2–5. Maintaining the pH at 7.2 rather than 6.5 favored a high ethanol/lactate ratio (9). Thus, in batch fermentation, the amount of ethanol observed depends on the pH and the time of sampling. Referring to Figure 3, it is clear that when these clostridia produce more than 1 mol of ethanol/mol of hexose fermented, a significant transfer of electrons must occur from FdH_2 to NAD^+ catalyzed by Fd:NAD^+ oxidoreductase. The role of this electron-balancing function was made clear when strain LQRI of *C. thermocellum*, which produced exactly 1:1 ethanol–acetate from cellobiose, was found to be devoid of Fd:NAD^+ oxidoreductase (27). Both strain AS39 of *C. thermocellum* and *Thermoanaerobacter brockii* contained this enzyme and could produce excess ethanol and low H_2 in experimental fermentations. Because *T. brockii* is in the same genus as *C. thermohydrosulfuricum* (5,6), it

is likely that the high ethanol production by this clostridium is a result of the same electron-balancing pathways (see Table 2).

Because of the effective degradation of various crystalline cellulose sources by *C. thermocellum*, the cellulose enzyme complex has been extensively studied, as discussed later in "Biology of Substrate Utilization." To take advantage of this capability, the approach has been to develop mutants or devise both batch and continuous culture fermentation conditions that will increase the relative yield of ethanol to other products and increase tolerance of the organism to produce ethanol at 15 g/L rather than the normal 5 g/L (9). A second approach, which uses the principle of symbiotic coculture, is also available. Coculture fermentations have been developed at the experimental level to exploit the advantageous properties of two thermophilic clostridia to produce ethanol from biomass (26,28). A metabolic explanation for a stable communal coculture of *C. thermocellum* and *C. thermohydrosulfuricum* fermenting cellulose to ethanol has been presented. *C. thermohydrosulfuricum* depends on the products cellobiose and glucose formed from degradation of cellulose by *C. thermocellum*. Glucose utilization by *C. thermocellum* is inhibited by cellobiose, and both glucose and cellobiose are more rapidly utilized by *C. thermohydrosulfuricum*, which then allows some growth of *C. thermocellum*; however, the major fermentation, which is carried out by *C. thermohydrosulfuricum*, yields a high ratio of ethanol to acetate plus some lactate. Recent example cocultures of strains of these two bacteria fermenting 10 g of cellulose per liter yielded an end-product ethanol to acetate ratio of 10.5 to 4.3 in 19 hours (28). Since the thermophilic clostridia also secrete potent amylases and pullulunases, starch biomass is clearly an important feedstock for ethanol production catalyzed by these organisms, as discussed later (see "Development of New Solvent-Producing Strains").

Recently, 10 or more mesophilic cellulose-fermenting clostridia have been isolated, which degrade cellulose and ferment cellobiose to a mixture of acetate, ethanol, and other products (29). *C. saccharolyticum*, which was originally isolated from a cellulose-enriched sewage sludge, has been grown on cellulose in stable coculture with "*Bacteroides cellosolvens*" (30). Since *C. saccharolyticum* ferments cellobiose primarily to ethanol at 25 °C (Table 2), cellulose fermentation employing this strain in coculture with a mesophilic cellulolytic clostridium may form the future basis for direct bioconversion of cellulose to ethanol.

Butanol Fermentations

Very few species of the butyric acid clostridia are capable of producing significant amounts of butanol. The seven organisms listed in Table 1 have quite distinctive properties. *C. acetobutylicum*, taxon I, is the amylolytic type strain, ATCC 824^T or DMS 792; and phylogenetically it is only distantly related to the saccharolytic strains belonging to taxonomic groups II, III, and IV (4). *C. acetobutylicum*, taxon II (strains NCP 262 or NRRLB 643), contain a very small genome of 2.9 Mbp (versus 4.0 Mbp for strain ATCC 824 or 6.5 Mbp for a strain of *C. beijerinckii*, strain NCIMB 8052T) (31). Taxonomic group III, *C. saccharoperbutylacetonicum*, strain N1-4, is closely related to the other saccharolytic strains (4). In addition to *n*-butanol, all four of these strains produce butyrate, acetate, ethanol, and acetone (Table 2). Alternatively, some *C. beijerinckii* strains typically produce isopropanol instead of acetone. Even though these four organisms carry out basically the same fermentation, they are quite distinct biologically as to rapid growth and fermentation with a high butanol yield on different substrates in industrial-scale batches. The taxonomic group I strains utilize maize or potato starch effectively, whereas organisms of taxa II and III were selected to carry out a saccharolytic fermentation with a molasses substrate. *C. puniceum* carries out the same amylolytic fermentation, producing mostly butanol under the proper conditions (18). *C. puniceum* has the added ability to produce pectinolytic enzymes, which would increase access to plant starch important for some starting substrates. Taxonomic group IV, *C. beijerinckii* (NRRL B592) and *C. aurantibutyricum*, which ferment starch and a variety of sugars to butanol, ethanol, and isopropanol, have been the subjects of research but not commercial fermentations. *C. tetanomorphum* differs from the other organisms in that, during a saccharolytic fermentation, it produces equal amounts of butanol and ethanol but not acetone or isopropanol. Both the more recently described strains, *C. tetanomorphum* and *C. puniceum*, differ from the other five organisms in that they produce significant butanol and ethanol throughout exponential growth rather than only during the later stages of batch fermentations (18,19).

The earlier studies up to 1988 on the biochemistry, genetics, and development of butanol fermentation by the clostridia have been reviewed (24,32,33). Recent advances in the molecular biology, regulation of the fermentation, and other new developments are reviewed by Rogers and Gottschalk (25) and in a review of a meeting on solventogenic clostridia (34). The major enzymatic pathways for the products of butanol fermentation are now known, and the physiologic controls of the switch from an acidogenic to a solventogenic fermentation observed in some of the organisms have been identified. Figure 2 depicts the five major pathways of this fermentation. A characteristic shift in the fermentation pattern was known for some time by microbiologists running industrial fermentations employing strains of "*C. acetobutylicum*" (21,24). In a typical experimental batch culture fermenting glucose, there is an early accumulation of acids (acidogenic phase) followed by butanol formation (solventogenic phase) and a reutilization of butyrate and acetate (Fig. 4). *C. acetobutylicum* growing exponentially on sugars or starch at a pH 5.6 or greater produces butyrate, acetate, CO_2, and H_2 as major fermentation products (Fig. 2, thick gray arrows). During this acidogenic fermentation phase, the bacteria are running a basic butyrate–acetate fermentation. To maintain electron balance, for every mole of acetate formed, an extra mole of H_2 is formed, similar to the case of ethanolic fermentation. Therefore, about 2.4 mol of H_2 is produced from a fermentation of 1 mol of hexose that produces 0.4 mol of acetate and 0.8 mol of butyrate. The extra 0.4 mol of H_2 comes from transfer of electrons from NADH through NADH:Fd oxidoreductase and hydrogenase to H_2 (Fig. 2, reaction 1).

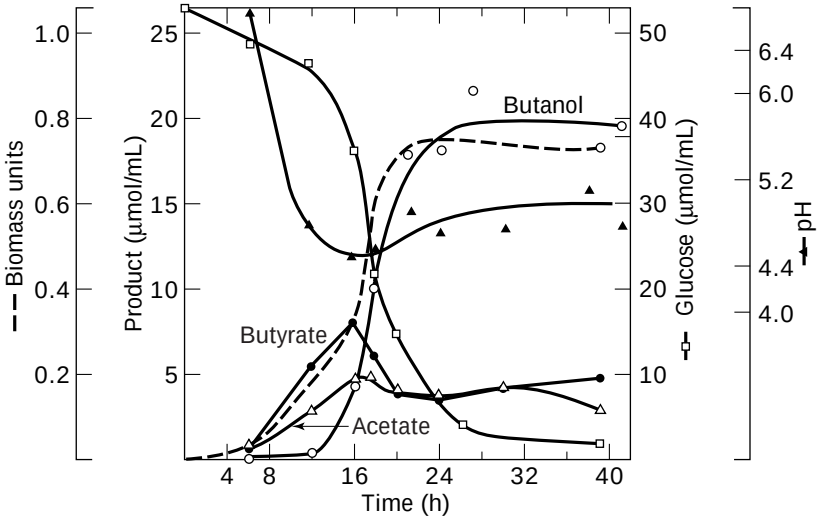


Figure 4. The switch from acidogenic to solventogenic fermentation by *C. acetobutylicum*, strain NRRL B643, in batch culture. *C. acetobutylicum* was grown in a yeast extract, amino acid medium at 37 °C with glucose as an energy source. Samples withdrawn at times shown were analyzed for biomass (dashed line), glucose (open squares), and pH (solid triangles); products in the broth were determined using gas chromatography: acetate (open triangles), butyrate (solid circles), and butanol (open circles). Acetone and ethanol determinations are omitted. *Source*: Reproduced with permission from Ref. 24, p. 6.

With accumulation of acids in a batch culture, the pH drops to pH 4.0–4.5 (depending on the strain), the growth becomes linear, then stops, and the classical switch in the fermentation occurs. The triggering of solvent formation requires the induction of new enzyme pathways in the cells (Fig. 2, thick black arrows) catalyzing formation of butanol, acetone, and ethanol. The connection among the processes of organic acid accumulation, drop in pH, slowing of growth, and induction of solvent-forming enzymes is addressed in the next section. During this phase, there is a net uptake of butyrate and acetate from the fermentation beers back into the cells (Fig. 4) and their recycling and conversion to butanol or ethanol (Fig. 2, arrow, reaction 9). Finally, acetone or isopropanol is produced as a result of decarboxylation of acetoacetate (Fig. 2, reactions 10 and 11).

Depending on the clostridial strain, one or two CoA-dependent aldehyde dehydrogenases and two or three alcohol dehydrogenases are needed for butanol, ethanol, and isopropanol formation, and some of these enzymes are induced 40 to 200-fold at the shift (Fig. 2, reactions 3, 7, 8, and 11). Three additional changes occur as a result of the shift from acidogenesis to solventogenesis:

- There is a lower ATP yield per mol of glucose fermented because net accumulation of acetate and butyrate allows additional production from acetyl phosphate and butyryl phosphate during exponential growth (Fig. 2, reactions 2 and 6).
- Operation of the alcohol pathways requires 2 mol additional NADH per mol of alcohol produced.
- The amount of H_2 made is reduced to about 1.4 mol of H_2 per mol of glucose fermented, since electrons must be shunted from FdH_2 to NADH (Fig. 2, reaction 1).

These changes in electron flux may be involved in regulatory signals for promoting the induction of the shift. This is discussed later ("Regulation of Solvent Production").

Acetone and eventually isopropanol production requires two enzymes: acetoacetyl–CoA transferase (ACT) and acetoacetate decarboxylase (Fig. 2, reactions 9 and 10). In *C. acetobutylicum* and *C. beijerinckii* strains, these enzymes are induced 50- to 100-fold during the shift to solventogenesis. The production of acetone by many of the solvent-forming clostridia is biologically interesting because the pathway neither produces extra ATP nor acts as an electron sink. Even a reduction of acetone to isopropanol by *C. beijerinckii* strains is oddly inefficient, requiring only one NADH, while reduction of acetoacetyl CoA to butyryl CoA recycles two NADHs (Fig. 2, cf. pathways 9–11 and 5). The best guess is that the physiologic role of this pathway is recycling of accumulated organic acids for detoxification and pH control. Thus, after induction of these enzymes, acetate and butyrate are taken up by the bacteria, the ACT enzyme transfers CoA from acetoacetyl–CoA to these acids to form their acyl–CoA intermediates (Fig. 2, reaction 9), and then they are funneled through the alcohol pathways to butanol and ethanol. The decarboxylase pulls the reaction, forming CO_2 and acetone (Fig. 2, reaction 10), or is reduced to isopropanol (Fig. 2, reaction 11) and excreted. During the solvent production phase, reduction in concentration of acids in the medium occurs early, while acetone continues to accumulate long after, when there is no net increase in acids. It appears that in these clostridia, acid production and acid recycling occur throughout solvent formation. High levels of the butyrate pathway enzymes are present during this period, and added [^{14}C]butyrate is rapidly recycled to [^{14}C]butanol continuously. In contrast, *C. tetanomorphum*, which produces both butanol and butyrate simultaneously throughout the fermentation of sugars, contains no ACT enzyme, produces no acetone, and does not recycle [^{14}C]butyrate to [^{14}C]butanol at any time (25).

Other Solvent Products

The butanol-fermenting clostridia can be forced to produce several alternative solvents by growth under special conditions with specific carbon sources. For example, growth of *C. beijerinckii* in a chemostat culture at pH 6.5–6 with glycerol as a carbon source yielded 61% conversion to 1,3-propanediol (35). Glycerol fermentation by 17 of 21 solvent-producing clostridia produced 1,3-propanediol. *C.*

acetobutylicum and *C. beijerinckii* grown in defined media with rhamnose, with a gas phase (CO_2) plus 0.2% $NaHCO_3$, yielded 1,2-propanediol, propionic acid, and propanol as major fermentation products. Three strains of *C. thermosaccharolyticum* ferment glucose yielding up to 7.9 g/L of *R*(–)-1,2-propanediol with a best yield ratio of 0.27 g per gram of glucose. *C. sphenoides* normally produces lactate, acetate, and ethanol from fermentation of glucose. During phosphate limitation, this organism yields up to 34 mM D(–)-1,2-propanediol per 100 mM glucose fermented. New fermentation pathways are probably induced under these conditions. However, the biochemistry has not been carefully studied (see Ref. 25 for specific references).

Another strategy for unusual solvent production uses the principle of enzyme recruitment. It has been demonstrated that *C. acetobutylicum*, grown until it has shifted into the solventogenic phase, can convert several organic acids to their analogous alcohols when added to the growth medium through the recycle mechanism outlined earlier. For example, propionate, valerate, and 4-hydroxybutyrate are taken up by this bacterium and converted to *n*-propanol, *n*-pentanol, and 1,4-butanediol, respectively (25).

REGULATION OF SOLVENT PRODUCTION

Physiology of Control of Solvent Formation

To obtain a high yield of a specific product in a fermentation for solvents, it is often necessary simply to adjust the growth medium, carbon source, and other conditions that will ensure maximum amounts of the desired solvent and minimize production of other products. However, research on the cellular regulatory mechanisms that govern the fermentation not only will allow more effective control of culture conditions but will open the door to development of new highly productive strains. Because many clostridial fermentations typically depend on branched fermentation pathways, analysis of the control elements becomes particularly important. Because most of the recent research is centered on the *C. acetobutylicum* butanol fermentation, this work is emphasized here, although a few observations from the clostridial ethanol fermentations are mentioned. The pathways, products, and characteristics of the acidogenic and the solventogenic fermentations of *C. acetobutylicum* were presented in the preceding section. The culture conditions that promote the shift away from acid production and favor optimal butanol and acetone production are listed in Table 3. These conditions are about the same for batch or continuous cultures of *C. acetobutylicum* (24,32,33) and of *C. puniceum* (18). The conditions that favor butanol and ethanol production without acetone (the alcoholic fermentation) by continuous cultures of *C. acetobutylicum* (36) are also listed in Table 3. The internal status of the bacterium has been measured for each of the three different fermentations (36–38). During the acidogenic fermentation in batch or continuous cultures, the following conditions are required: adequate glucose, a pH_o greater than pH 6.0; and a low level of organic acids. The intracellular pH (about pH_i 6.8) is maintained close to neutral either by excretion of H_2 through the action of hydrogenase or by hydrolysis of ATP and pumping H^+ out of the cell. At a pH_i 6.8 the cell maintains a low positive ΔpH and a very low level of undissociated butyric acid. ATP is produced and utilized at a high rate in this fermentation. The low NADH/NAD ratio is probably a result of the high NADH:Fd oxidoreductase activity, whereby excess electrons are dumped onto FdH_2 and exit through hydrogenase as molecular hydrogen. H_2 is produced in excess in the acidogenic fermentation (Table 3 and Fig. 2, reaction 1, gray arrow).

The shift to a solventogenic fermentation requires an adequate carbon source (glucose), an acid pH_o (pH 4.8–4.4), and high concentrations of butyrate and acetate in batch cultures. Continuous chemostat cultures limited for PO_4 or nitrogen source can be forced to carry out a solventogenic fermentation either by adding a high concentration of organic acids or by setting the pH_o to pH 4.8–4.3, depending on the strain of clostridium (36,37). The intracellular pH (pH_i 5.9) is more acid and results in a greater than 10-fold higher concentration of undissociated butyric and acetic acids in the cell. Also a very high positive ΔpH of 1.1–1.5 is maintained by these cells. Curiously, this high ΔpH persists even when solvent-producing cells are treated with *N*,*N*′-dicyclohexylcarbodiimide (an ATPase inhibitor) or after CO gassing (a hydrogenase inhibitor) (36). A high concentration of ATP is maintained, and turnover of ATP is slowed in the cells, because growth is slower during solvent production. The NADH/NAD ratio is maintained similar to acidogenic cells, even though production of butanol results in new sites for oxidation of NADH. This is due to a net transfer of electrons from FdH_2 as mentioned earlier. This physiologic state is somehow essential for the induction of the enzymes required for solvent fermentation (Fig. 2, reactions 7–10).

The alcoholic fermentation, in which high amounts of butanol and ethanol are produced without acetone, has been studied only recently (36). The alcoholic fermentation in continuous cultures requires high glycerol and limited glucose, at a neutral pH. Table 3 lists three inhibitors of hydrogenase activity that, when added to continuous cultures with adequate glucose, also force cells to produce butanol, butyrate, and ethanol, but no acetone (37). These conditions cause an overload of NADH in the cell, resulting in a high NADH/NAD ratio and various levels of ATP (Table 3). The high level of NADH leading to butanol and ethanol formation is accompanied by increased activities of NADH-dependent alcohol and aldehyde dehydrogenases, which differ from the NADPH-dependent dehydrogenases that are expressed at high levels during solventogenesis in the bacterium. For example, the NADH-dependent butyraldehyde dehydrogenase activity is 100-fold higher in alcohologenic cells, while the NADPH-dependent enzyme activity is induced in solventogenic cells. This suggests that different sets of genes are expressed in these three fermentation conditions (36). The inversion of the transmembrane pH gradient ($-\Delta$pH 0.3) observed during the alcoholic fermentation at neutral pH (Table 3) is apparently due to the low rate of H_2 production. But this inversion does not seem to create an energy crisis for the bacteria, since a similar proton motive force (pmf) was measured for both alcohologenic and acidogenic cells (36). An increase in the electrical potential ($\Delta\Psi$) compensates for the low chem-

Table 3. Conditions for Three Fermentations by *Clostridium acetobutylicum*

	Acidogenic	Solventogenic	Alcoholic
Intracellular status			
$[H^+]$	pH_i 6.8 ΔpH 0.3–0.6	pH_i 5.9 ΔpH 1.1–1.5	pH_i 6.2 ΔpH − 0.3
Undissociated butyrate, acetate	Low concentration	High concentration	Low concentration
ATP (μmol/g cell mass)[a]	Low (0.62), High turnover	High (1.6), Slow turnover	Variable (0.3–1.3), slow turnover
NADH/NAD (ratio)[a]	Low (0.14)	Low (0.1–0.2)	High (0.5–0.8)
Enzymes	Acid pathways (2 and 6 in Fig. 2) NADH:Fd oxidoreductase very high (gray arrow at 1, Fig. 2)	Solvent pathways, induced: butanol (7 and 8 in Fig. 2), acetone (9 and 10 in Fig. 2)	Alcohol pathways, induced: NADH-dependent alcohol and aldehyde dehydrogenases, Fd:NAD $(P)^+$ reductases
Major products	Butyrate, acetate, excess H_2	Butanol, acetone, low H_2	Butanol, ethanol, very low H_2
Fermentation conditions	Adequate glucose: pH_o 6.5–6.2; low $[organic\ acid]_o$	Adequate glucose: (1) pH_o 4.8–4.4, high $[organic\ acids]_o$; (2) Limited PO_4 either neutral pH at high $[organic\ acids]_o$ or pH 4.3 at low $[organic\ acids]_o$	(1) Glycerol and low glucose, pH_o 6.5; (2) Adequate glucose, pH_o 6.5–5.6 + neutral red (1 μM) or + CO gassing or + methyl viologen (4 μM) or limited Fe

[a]Data from Ref. 21.

ical potential (ΔpH). This may be related to the distribution of K^+ across the membrane.

Thus, in *C. acetobutylicum* each of the three fermentation conditions produces a distinct intracellular state, which in turn causes the bacterium to develop the enzyme network needed for each type of fermentation. The current view of physiologic control of *C. acetobutylicum* metabolism is a dual signal model in which a high internal ATP and/or butyric acid level generates one signal and a high NADH level provides the other (37). But how these internal physiologic signals are sensed by molecular sensors in the cell and transferred to molecular responders that specifically regulate gene expression by *C. acetobutylicum* is still unknown.

Genetics of Solvent Formation

Mutations That Affect Solvent Production and Its Regulation. Isolation and characterization of mutants of the solventogenic clostridia has provided valuable information on gene function and regulation of the solvent-producing pathways and of the linkage of this process with developmental changes in the bacterial cell (39). During industrial production of acetone and butanol by *C. acetobutylicum*, the early brewmasters recognized that high solvent-producing strains produced a high percentage of spores. Thus, the high solvent-producing strains were selected by subjecting cultures to cycles of sporulation, heat activation, and outgrowth (32). Mutants that fail to sporulate often also fail to produce solvents. These mutant cells tend to accumulate when *C. acetobutylicum* and related clostridia are repeatedly subcultured in batch cultures or grown in continuous culture, rather than starting cultures from fresh heat-activated spores. When the culture loses the capacity to produce solvents and form spores, this change is called "strain degeneration." Some strains, such as *C. beijerinckii* (NCIMB 8052), tend to degenerate more rapidly than others. In contrast, *C. acetobutylicum* (ATCC 824) will often give rise to asporogenous cells that continue to form acids or solvents (40). Recent studies on the phenomenon of clostridial strain degeneration have been reviewed (41). No satisfactory molecular explanation for the permanent degeneration of clostridial cells was available until just recently. In *C. acetobutylicum*, strains ATCC 824 and ATCC 4259, a large 210-kb plasmid (pWEIZ or pSOL-1) has been demonstrated that carries four solvent-forming genes of the *sol* operon, *ctfA*, *ctfB*, *adc*, and *aad* (see Fig. 5) (42). Degenerate strains that produce no butanol or acetone have lost the pSOL-1 plasmid. Since degenerate strains lacking pSOL-1 are also unable to sporulate or carry out any of the coordinate developmental functions, it is likely that major regulatory genes controlling development are located on this plasmid as well. A similar link between the induction of endospore formation and a fermentation shift occurs in *C. thermosaccharolyticum*. Exponentially growing cells produce acetate, butyrate, and lactate. A change in culture conditions produces elongate forms that ferment sugars to primarily ethanol, CO_2, and H_2 before the cells proceed to sporulate. Asporogenous mutants were selected that remained in the intermediate stage and continued to produce ethanol (43).

Mutants of solvent-forming clostridia that have defects in the pathways leading to acid production have been isolated to bias the carbon flow toward solvent production. Following mutagenesis, pathway mutants are often selected through the use of suicide substrates, which are toxic analogs of normal end products formed during the fermentation. For example, allyl alcohol acts as a suicide substrate when oxidized to the toxic aldehyde, acrolein, by the activity of alcohol dehydrogenases. Members of one set of allyl alcohol resistant mutants of *C. acetobutylicum* are

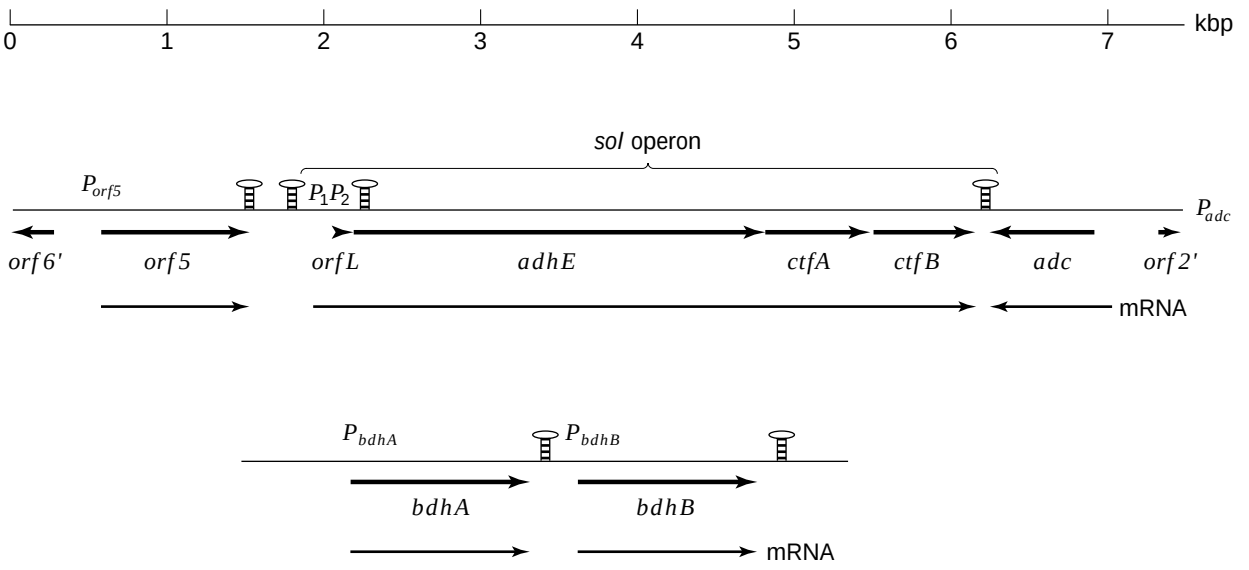


Figure 5. Organization of chromosomal gene regions encoding solventogenic enzymes in *C. acetobutylicum*, strains DSM 792 and ATTC 824. Arrows and arrowheads represent lengths, locations, and orientation of genes; lines with arrows indicate primary mRNA transcripts. The *sol* operon is marked off by a dashed line. A prime at the end of an open reading frame (ORF) indicates that the ORF is truncated. Promoter positions are indicated by P_{adc}, P_1, P_2, P_{bdhA}, P_{bdhB}, and P_{orf5}; possible stem–loop structures are indicated by hairpin symbols. *Source*: Reproduced with permission from Ref. 51, p. 253.

defective in butanol and ethanol production because of altered alcohol dehydrogenase enzymes. However, these mutants contain normal butyraldehyde dehydrogenase activity and produce significant amounts of butyraldehyde (15) (see Fig. 2, reactions 7 and 8). Fluoroacetate-resistant mutants of *C. thermosaccharolyticum* were found to be deficient in acetate kinase or phosphotransacetylase enzyme activities and thus produce less acetate and more ethanol than the parental strains (see Fig. 3, reaction 2) (44). An asporogenous ethanol-tolerant mutant of *C. thermocellum* (LD 1) was selected by means of fluoropyruvate. When fermenting cellobiose, this mutant strain produced a high ratio of ethanol to acetic acid, in contrast to the parental strains, which produce mostly lactic acid and acetic acid (9). Mutants of *C. acetobutylicum* (B643) and *C. beijerinckii* (NCIMB 8052) that produce excess butanol and acetone and more importantly also recycle virtually all acetate and butyrate have been isolated (15,16). These super recycling strains may be used in future production methods covered in a later section "Development of New Solvent-Producing Strains").

Conjugative Transposons and Transfer of Plasmid Vectors. Conjugative transposons are a class of mobile genetic elements usually found as residents in the chromosomes or plasmids of gram-positive bacteria (see reviews, Refs. 45,46). Originally discovered in *Enterococcus faecalis*, these elements have a broad host range and are important tools for genetic analysis. These mobile elements can integrate almost at random into the chromosome of the recipient bacterium, therewith creating an insertion that disrupts the sequence of the DNA information in a specific gene. Thus, the transposon mutagenesis differs from mutagenesis due to physical or chemical means in that only one specific alteration in the host cell DNA occurs, rather than a possibility of two or more random changes in the DNA that obscure the alteration in the phenotype and gene under study. Genetic techniques ensure that only one copy of the transposon has been integrated into the cell for each mutational event. The gram-positive bacterium *Enterococcus faecalis* is the usual donor in matings with *Clostridium*, yielding a transfer frequency of 10^{-3} to 10^{-4} per recipient. The transferred transposon, Tn916 (16.4 kbp), Tn925 (14.5 kbp), or Tn1545 (23.5 kbp), carries an antibiotic resistance marker such as tetracycline resistance (*tetM*) that allows selection of the transconjugants. Because the genes required for the conjugation and DNA transfer are located within the transposon, a transposon carried either on a plasmid or on the chromosome of the donor will be transferred. A number of specific Tn916-induced mutants have been identified in *C. acetobutylicum*. Both structural gene mutants in the fermentation pathway and regulatory mutants preventing solventogenesis, sporulation, and granulose formation have been selected (45,46).

Development of methods for transfer of plasmid vectors into *Clostridium* is vital to the application of genetics to further our understanding of the molecular basis of regulation of solvent formation. A summary of these transformation methods and the clostridial cloning vectors used with *C. perfringens* and *C. acetobutylicum* is available (48). There are two effective methods of transformation of *C. acetobutylicum*. Transformation of protoplasts requires conversion of cells to wall-less protoplasts followed by uptake of plasmid or bacteriophage DNA induced by polyethylene glycol (PEG), regeneration back to rod-shaped cells,

and selection of transformed cells. Using this procedure with strain N1-4 of *C. saccharoperbutylacetonicum*, 10^4–10^5 transformants per microgram of plasmid or phage DNA are obtained. Transformation of whole cells requires suspension of washed cells together with plasmid DNA, electroporation in a cuvette at high voltage (5.0 kV/cm) using a BioRad Gene Pulser, and, after a recovery period of 2–4 hours, selection for transformants. When *C. beijerinckii* (NCIMB 8052) is used, about 10^2–10^4 transformants are obtained per microgram of DNA. Basically, the short pulse of electric current punches temporary holes in the cell membrane, permitting the plasmid DNA to enter the cell. To be useful and identifiable, the plasmid DNA must be able to survive and replicate along with cell division of the host bacterium. In addition, there must be selectable markers such as the *ermC* gene, for antibiotic (erythromycin) resistance, so that colonies of transformed cells can be identified.

Vectors for *C. acetobutylicum* are chimeric cloning vectors that contain an origin of replication that operates in *Clostridium* and a second origin of replication that permits the plasmid to replicate in *E. coli*. Thus, specific pieces of *Clostridium* DNA can be ligated (cloned) into specific sites of this vector, replicated and multiplied in *E. coli*, extracted, and then transferred to mutant *Clostridium* cells to study their effects on the physiology of the host cell. This technique has permitted identification of the roles of suspected regulatory genes and overexpression of specific fermentation pathways in *Clostridium acetobutylicum* (49).

Unfortunately, the successful transformation process is strain specific, as is plasmid survival. Originally, transformation procedures were successful only when *Clostridium* strains in taxa III and IV (see Table 1) were used. Later, electroporative transformation of *C. acetobutylicum* (ATCC 824, taxon I) with plasmid pIM13 from *Bacillus subtilis* was successful. But when a shuttle plasmid was prepared by ligating pIM13 together with *E. coli* plasmids, no transformants were seen. It was discovered that strain ATCC 824 contains a restriction endonuclease (*Cac* 824 I) that cut these shuttle vectors in more than 40 sites, thus destroying the plasmid in the cell. The strain 8052 (taxon IV) does not carry this restriction endonuclease, which explains why the same shuttle plasmids survive in these bacteria (49). The clever solution to survival of these shuttle vectors in the *C. acetobutylicum* (ATCC 824) host requires replication of the shuttle plasmid in *E. coli* in the presence of a second plasmid that expresses a specific transmethylase. This Φ3TI transmethylase adds a methyl group to a cytosine base in the sensitive sites of the shuttle vector DNA. This protects the plasmid DNA, and when extracted and electrotransferred into *C. acetobutylicum* (ATCC 824), the methylated plasmid survives at a high frequency (10^5 transformants/μg DNA) and replicates (49). To date, the important production strains of *C. acetobutylicum*, NRRL B643 or NCP 262 (taxon II), have not been successfully transformed by means of the methods and vectors available. Perhaps there are additional restriction endonucleases that act as barriers to vector survival in these strains.

Cloning, Genetic Organization, and Expression of Solvent and Acid Pathway Genes. Application of molecular genetic techniques together with basic enzymology to the steps in the *C. acetobutylicum* fermentation pathway has revealed a great deal about how this network is organized and regulated. Most of the genes coding for the enzymes involved in acid and solvent fermentation have been cloned in *E. coli*, since this organism expresses many foreign proteins from plasmid-borne genes on specially constructed expression vectors (50). Sequencing of the DNA regions coding for the solvent genes together with the adjacent upstream promoter regions has revealed the basic organization of these genes (Fig. 5, Ref. 51). In strains ATCC 824 and DMS 792, acetoacetate decarboxylase is encoded by a gene (*adc*) in a single operon. Adjacent is the convergently transcribed *sol* operon, which encodes an aldehyde/alcohol dehydrogenase (*adhE*), and the two subunits of acyl CoA transferase (*ctfA* and *ctfB*). The promoters P, P_2, and P*adc*, all have high homology to typical gram-negative and gram-positive bacterial promoter sequences and control expression of these genes. A typical terminator sequence (Fig. 5, hairpin symbols) between the two regions stops transcription from both directions. This entire region of the DNA that includes the *sol* operon and other linked genes such as *adc* is carried on the 210 kbp plasmid (pSOL-1) in *C. acetobutylicum* (ATCC 824), along with other yet undetermined regulatory genes. The remaining genes coding for solvent- and acid-forming enzymes appear to be carried on the chromosome in this strain.

At present, the distribution of the solvent genes in the other six groups of butanol-producing clostridia is not known. The DNA sequences reveal that these solvent genes as well as the genes encoding for butyrate production are transcribed by the RNA polymerase with sigma H (σ^H) that carries out expression of many housekeeping genes in gram-positive bacteria. Alternative sigma factors do not appear to play a role in regulation of solventogenesis in *C. acetobutylicum* (51).

The expression of a synthetic acetone operon carried on a shuttle plasmid (pFNK6) and transferred back into *C. acetobutylicum* (ATCC 824) reveals some additional properties of regulation of solvent production (49). The synthetic acetone operon was constructed from the *adc*, *ctfa*, and *ctfb* genes arranged to transcribe in the same direction from the P*adc* promoter (see Fig. 5). In cell extracts from strain ATCC 824 (pFNK6), the activities of the two acetone enzymes were more than fourfold higher than in cell extracts of strain ATCC 824. The course of the batch fermentation by ATCC 824 (pFNK6) differed from ATCC 824, with almost double the yield of both butanol and acetone and the complete disappearance of butyrate and acetate from the broth by 20 h (Fig. 6). The increase in butanol is consistent with the view that diversion of acetoacetyl CoA to acetone formation and away from butyryl CoA might cause a temporary high rise in NADH/NAD ratio and trigger early induction of the enzymes for butanol synthesis (see Fig. 2, reactions 5, 9, and 10, and 7 and 8). Two "degenerate" mutants of ATCC 824, namely M5 and DG1, which had lost the pSOL plasmid and produce none of the enzymes encoded by the *sol* operon, were transformed with plasmid pCAAD containing the *adcE* gene or with pFNK6 containing the synthetic acetone operon D (42). The degenerate mutants with pCAAD produced significant butanol and butyric acid but no acetone. The same degenerate mutants with pFNK6 produced butyric acid and acetone but

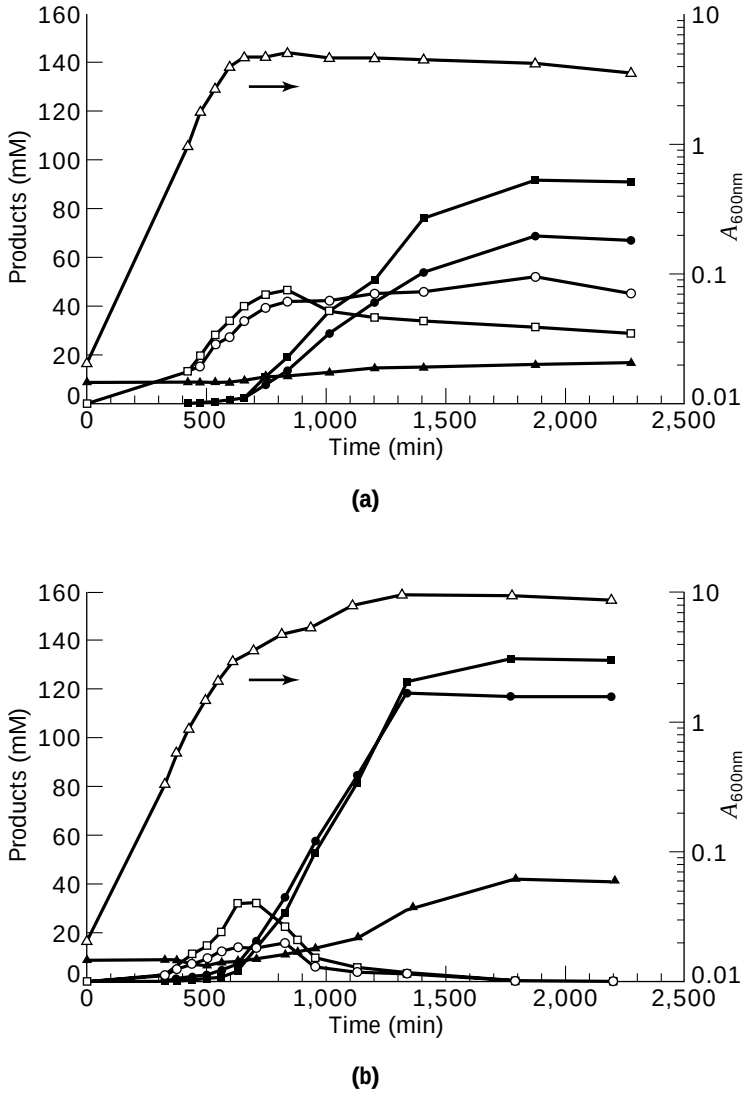


Figure 6. Optical density and product profiles in pH 4.5 batch fermentations for *C. acetobutylicum* strains: (**a**) ATCC 824 and (**b**) ATCC 824 (pFNK6). Open triangles, optical density at 600 nm; solid squares, butanol; solid circles, acetone; solid triangles, ethanol; open squares, butyrate; open circles, acetate. *Source*: Reproduced with permission from Ref. 49, p. 62.

no butanol. The butyraldehyde dehydrogenase (BYDH) activity was produced only in late growth phases in M5 (pCAAD), as is typical for strain 824 (50). M5 cells alone produced no butanol or BYDH activity. Thus the regulatory signals for the *sol* operon must operate "in trans" and be able to control multiple copies of the plasmid-borne genes for solvent production.

Chromosome Mapping and Gene Sequencing. When gene transfer methods and a number of cloned genes became available, construction of a genetic map of the *C. beijerinckii* chromosome was begun to expand genetic analysis of a solventogenic clostridia. About 32 genes have been positioned relative to one another on the 6.5-Mbp circular chromosome map (31). The process of mapping depended on preparation of the genomic DNA by lysis of cells within agar blocks, treatment with exonucleases that cut this DNA in just a few places, and then pulsed-field gel electrophoresis. The large chromosome pieces displayed on the gel are than probed with cloned genes to position each gene. As indicated earlier ("Butanol Fermentations"), this method has revealed wide divergence in strains of *C. acetobutylicum* (4). Other differences in arrangement of solvent genes and regulatory genes may emerge later.

The complete genomic DNA sequences of a number of Eubacteria and of Archaea, such as *Methanococcus jannaschii* are now available (52). The complete DNA sequences of the genome and the pSOL-1 plasmid of *C. acetobutylicum* (ATCC 824) are currently being determined (53). These sequences should provide valuable information on the genetic basis for expressing the proteins and regulatory networks essential for cellular vegetative and developmental changes carried out by this organism. Future successful strain development employing *Clostridium* will depend on complete analysis of the genetic potential of these organisms.

Insertional Mutagenesis with Integrational Plasmids. Recently, it has been possible to develop gene transfer involving homologous recombination in the solvent-producing strains *C. beijerinckii* and *C. acetobutylicum* (ATCC 824). Conjugation procedures with *C. beijerinckii* and electro-

transformation with *C. acetobutylicum* were used to obtain transconjugants or transformed cells at high efficiency, leading to the detection of the integration of plasmids that lack a replication function active in *Clostridium* (Rep^-). Transformants or transconjugants arise by a Campbell-type recombination event between the chromosomal segment in the plasmid and the same region in the bacterial chromosome (31). This insertional event results in an altered or inactivated gene. Nonreplicative (R^-) plasmid constructs containing parts of the acetate kinase gene (*ack*) or butyrate kinase (*buk*) gene fragments were integrated into normal regions of the chromosome of *C. acetobutylicum*. The altered strains produced more butanol during fermentation of sugars (54).

Is Solvent Production a Cellular Stress Response?

Endospore Formation and Solventogenesis. The ultimate response of a number of bacteria to physical or chemical stress or starvation is formation of stress-resistant, metabolically inactive resting cells. Species of the gram-positive *Bacillus* and *Clostridium* differentiate into highly resistant endospores. The biology of endospore formation, including the molecular events that take place during this developmental process, has been studied primarily in *B. subtilis* by many microbiologists over the past 30 years (55). The physiologic changes inside the cell that signal the shift of *C. acetobutylicum* to solventogenesis are most likely either a shift in the redox state of elements in the electron transfer system, such as NADH or reduced Fd, or an increase in the concentration of butyric acid or H^+, or maybe both, as suggested earlier (see "Physiology of Control of Solvent Formation"). In this bacterium, the shift to solventogenesis is associated with other rather massive changes, which include induction of endospore formation and must involve the induction and activation of complex regulatory networks, proteins, and pathways (25). What are the molecules that receive the physiologic signals, and how are all these events triggered in this bacterium? As mutant studies cited earlier ("Genetics of Solvent Formation") indicate, initiation of solvent production and sporulation are tightly coupled, although clearly solventogenesis is not a prerequisite to sporulation in *C. acetobutylicum*, *C. beijerinckii*, and *C. thermosaccharolyticum*. Hence there must be both separate and common control mechanisms. Two lines of recent research support the view that these bacteria are regulated by two classes of the molecular signaling systems that normally operate in bacteria for stress control and adaptation to environmental change.

First, in many bacteria, coordinate transcription of sets of genes is regulated by alternative sigma factors that confer new promoter specificities on RNA polymerase. Examples are heat shock, or stationary phase in *E. coli*, sporulation in *Bacillus* and streptomycetes, and flagella synthesis in many bacteria. Development of endospores in *Bacillus* requires a temporal sequencing of gene expression regulated by four sigma factors (σ^F, σ^E, σ^G, and σ^K) and a protease (*spo* II GA) that activates the σ^E protein. All the genes encoding these sigma factors and the protease have been identified in *C. acetobutylicum* by cloning and sequencing (56). Thus, even though the physiologic signals for sporulation in *Bacillus* (nutrient starvation) are different from those for *Clostridium* (blocking of growth in the presence of excess nutrients), they utilize the same general regulatory features for development. As pointed out above, σ^H the sigma factor that directs RNA polymerase to general cellular promoters in *C. acetobutylicum*, similar to σ^A in other bacteria, apparently also directs the expression of solvent and acid enzyme genes in this bacterium. Therefore, it is clear that the genes concerned with solventogenesis (*sol* genes) are not part of the sporulation cascade. Indeed, a different mechanism coordinates sporulation and solventogenesis, along with other key physiological and morphological changes that occur during the metabolic "switch."

The second major signal system required for endospore formation is the phosphorelay pathway essential for initiation of sporulation (57). In *B. subtilis* a number of response regulator proteins respond to environmental, cell cycle, and metabolic signals, resulting in the phosphorylation of the SpoOA transcription factor. The SpoOA~P controls the transcription of numerous genes by activating or inhibiting the levels of other transcription regulators, which in turn regulate transcription of proteins such as sigma factors required to enter the sporulation cycle. The SpoOA protein acts as a transcriptional regulator both as a repressor of transcription and as an activator by binding to DNA sequences called the "OA boxes" (5-TGNCGAA-3) (57). Only the phosphorylated form, SpoOA~P, is active. PCR primers of portions of SpoAO averaged from sequences of four bacilli, were used to isolate SpoOA segments from the DNA of 11 other species of *Clostridium* and *Bacillus* including *C. acetobutylicum*, *C. beijerinckii*, and *C. pasteuranium*. The complete sequence of the *spoOA* gene from *C. beijerinckii* has been determined and, as in the case of two *Bacillus* strains, the promoter region of the gene has a σ^H promoter with an OA box (58). This is consistent with the autoregulation of *spoOA* by SpoOA~P, first shown in *B. subtilis*. Further, possible "OA box motifs" were found in all the promoter regions for the solvent genes *adc*, *ctfAB*, *adhE*, *bdhA*, and *bdhB* (see Fig. 5) and for the genes *ptb-butk* of *C. acetobutylicum* (31). Expression of these genes is either activated or repressed during the switch to solventogenesis.

Thus, in *C. acetobutylicum* (ATCC 824), direct control by SpoOA is likely. This view was tested in *C. beijerinckii* (NCIMB 8052) by means of the insertional mutagenesis technique described earlier ("Genetics of Solvent Formation"). An R^- (unable to replicate) integrational plasmid containing a DNA segment in the *spoOA* coding region was prepared and transferred into the cell. Recombination occurred by the Campbell-type mechanism, yielding an insert of the plasmid with its antibiotic resistance marker (Em^R) within the chromosomal *spoOA* gene inactivating its expression. The $SpoOA^-$ strain was asporogenic, unable to produce granulose, and unable to produce solvents (31). This indicates that all three of these "switch" responses are most certainly regulated by the phosphorelay system.

Heat Shock Response System in Clostridium. Exposure to a heat shock of all organisms, from bacteria to humans, results in the rapid burst of synthesis of from 15 to 40

proteins. This transient upshift in synthesis of this set of heat shock proteins is also triggered by exposure to toxic chemicals, virus infection, ethanol, and other agents. The major proteins induced include two important categories. First to be induced are the chaperone proteins such as GroES, GroEL, DnaK, DnaJ, and GrpE, which are required for folding newly synthesized proteins and preventing aggregation of denatured proteins destined to be renatured or digested by proteases. Second, a series of ATP-dependent proteases (e.g., Lon, ClpP, ClpX) is induced, since such enzymes must play a role in protein processing or in digesting damaged proteins. There are also a number of sigma factors and other proteins induced. These heat shock proteins function as well in cells growing in normal conditions by "chaperoning" the proper folding of many proteins and digesting abnormal or denatured proteins. Thus the general function of this stress-induced family of proteins is to ensure a high yield of correctly folded proteins even in the presence of protein-denaturing physical and chemical shocks (for general review, see Ref. 59).

Since solvent production in *C. acetobutylicum* and other clostridia is often enhanced by stress conditions such as accumulation of toxic acids that inhibit growth as outlined earlier, it is important to ask how regulation of this heat shock system is related to induction of solventogenesis. The two major chaperon families (the GroE and the Dna K-J groups), as well as 10 or so other heat shock proteins, are transiently induced when *C. acetobutylicum* is upshifted from 30 °C to 42 °C for 5 min. (60). When the same organism shifts to solvent production by a drop in pH at a high butyric acid concentration, the same transient induction of heat shock proteins occurs. Thus, this stress response is coordinated with the induction of the expression of genes encoding solvent-forming enzymes. In the case of *E. coli*, the expression of the heat shock genes depends on a special sigma factor, σ^{32}, which forces cellular RNA polymerase to recognize specific promoter sequences of the heat shock gene family. However, the regulatory mechanism for heat shock gene expression in *C. acetobutylicum* appears to be different. No evidence for an alternative sigma factor has been found, and the promoter sequences of the *groE* and *dnaK* operons conform to the common promoter region for gram-positive bacteria. However, in both *C. acetobutylicum* and *B. subtilis* there is a conserved inverted repeat (IR) of 9 bp separated by a 9-bp spacer within the promoter region of the *dnaK* and *groE* operons. This IR has been identified in 27 bacterial species (both gram-negative and gram-positive), always occurring only before the *dnaK* and *groE* operons. There is evidence in *B. subtilis* that this IR is a site that interacts with a regulatory protein (ORF 39) that may act alternatively as an activator and a repressor (61). A similar protein (ORF A) identified in *C. acetobutylicum* is a candidate for the same regulatory function.

How the mechanisms regulating heat shock and general stress response in the clostridia may be connected with the *SpoOA* phosphorelay system involved in the initiation of endospore formation and solventogenesis is still not known. In any case, the stress response in gram-positive bacteria is designed to combat a variety of stress situations. In summary, the overall stress response of the solvent-forming *C. acetobutylicum* and perhaps other solvent-producing clostridia appears to involve three response systems. First, heat shock genes are induced, expressing a 10- to 20-fold increase in chaperone proteins protecting vital cellular enzymes and other proteins against denaturing effects of internal butyric acid. Second, solvent genes are induced, expressing a 40- to 100-fold increase in butanol and acetone pathway enzymes and resulting in recycling of butyric and acetic acids, thus detoxifying the environment. Finally, the first two stress responses permit the cell time to induce the complex machinery required for differentiation to endosporulation, which is the ultimate stress response available to all the clostridia.

DEVELOPMENT OF NEW SOLVENT-PRODUCING STRAINS

The Product Tolerance Problem

When a *Clostridium*-based fermentation is contemplated for a commercial solvent production system, the factors such as rate and efficiency of substrate conversion to solvents must be considered. But the major problem to be faced for all clostridial fermentations is extraction of low concentrations of solvents from water. It has been calculated that when distillation is used in the recovery of ethanol at less than 6–8%, or *n*-butanol at less than 2–3% from fermentation beers, the energy required exceeds the net energy recovery in the final products obtained in the commercial process. Yeasts and strains of *Zymomonas mobilis* convert sugars to an ethanol concentration high enough (10–20%) to permit separation of the solvent from the fermentation beer on an energy budget sufficiently low to allow profitable industrial production. In contrast, cocultures of *C. thermocellum* and *C. thermohydrosulfuricum* fermenting liquefied wood to ethanol produce a low final yield of 2% ethanol, a concentration above which growth of both these bacteria is inhibited. Likewise, the butanol–acetone fermentation by *C. acetobutylicum* stops when only 2% (w/v) solvents is reached, primarily owing to the toxicity of butanol.

There has been only limited research on the biological basis of low tolerance of clostridia to solvents involving strains of three organisms: *C. thermocellum*, *C. thermohydrosulfuricum*, and *C. acetobutylicum* (24). Growth of *C. thermocellum* at low concentration of ethanol results in an increase in the ratio of unsaturated/saturated fatty acids in the cytoplasmic membrane consistent with work with other bacteria. A mutant strain, C9, adapted to growth at higher levels of ethanol (2%) exhibited even greater increase in the amount of unsaturated fatty acids. The view is that ethanol enters the cell, causing a decrease in membrane fluidity, and the bacteria adapt by altering the fatty acid content to increase fluidity. The low ethanol tolerance of *C. thermocellum* (5 g/L) has been ascribed to specific inhibition of a number of enzymes required for glycolysis. Glucose fermentation by *C. thermohydrosulfuricum* (strain 39E) was inhibited by 0.5–2% ethanol. A mutant strain (39EA) grew in 8% ethanol at 45 °C and 3.3% ethanol at 68 °C. Unfortunately, the mutant strain produced a 1:1 ratio of ethanol to lactate, while the parent strain (39E) pro-

duces an ethanol-to-lactate ratio of 8:1. The ethanol-tolerant mutant also, incomprehensibly, lost the ability to utilize starch for growth (see review, Ref. 24).

The studies with *C. acetobutylicum* on product tolerance reveal that the effect of butanol appears to be opposite of ethanol in that butanol increases the cell membrane fluidity (24). Growth of *C. acetobutylicum* in 1% butanol increases the ratio of saturated to unsaturated fatty acids in the membrane. Thus, this adaptation is similar to fatty acid changes in mesophilic bacteria in response to maximum growth temperatures. Growth of this organism is inhibited at 11, 51, and 43 g/L of acetone, ethanol, and butanol, respectively, while as little as 6–8 g/L of butyrate and acetate inhibit growth. Thus, the view that the metabolic shift of these clostridia from an acidogenic to a solventogenic metabolism with recycling of organic acids is a detoxification stratagem is consistent with these observations. However, the fermentation is finally restricted by a low concentration of the major solvent, butanol.

Preliminary experimental approaches to improving butanol tolerance have been reported. Physiologic replacement of normal membrane fatty acids with long-chain fatty acids such as oleic acid (C_{18}) and elaidic acid (C_{20}) in cell membranes of *C. acetobutylicum* has been partially successful. Adding oleic acid or elaidic acid to growing cells in a biotin-free medium results in bacteria twice as tolerant to butanol inhibition. Butanol-resistant mutants selected from *C. acetobutylicum* (ATCC 824) will ferment corn starch to yield 7.9 g per liter of butanol (24). These butanol-tolerant mutants yielded 14 g per liter of butanol during fermentation and tolerated 16 g per liter of butanol. Autolysin-resistant mutants of strain NCP 262 are somewhat more butanol tolerant but produce only 20 g per liter of total solvents during fermentation. Clearly, the mechanism of solvent tolerance has not been elucidated. Thus, the low solvent yields characteristic of all *Clostridium* fermentations remain a major stumbling block for commercialization.

Biochemical engineers have developed new designs for continuous fermentation combined with product recovery technologies to cope with solvent intolerance inherent in clostridial fermentations. Application of these techniques for the acetone–butanol fermentation was reviewed by Maddox et al. (62). A combination of the new separation technologies and specialized mutant strains may be the best approach for future solvent production systems. For example, a pervaporation technology was combined with a fed-batch fermentation employing the superrecycling mutant B18 of *C. acetobutylicum* (NRRL 643) (15). Continuous removal of acetone and butanol through silicone tubing maintained the level of butanol in the broth below 4.5 g/L while hourly glucose consumption was held at about 2 g/L for 80 h. There was little accumulation of organic acids, and a high solvent yield (63). A long-term (2,000 h) continuous cultivation of *C. beijerinckii* (NRRLB 592) in a two-stage chemostat with on-line solvent removal by membrane evaporation produced solvents at a rate of 9.27 g/L/h without degeneration (64). The substitution of the high solvent, superrecycling *C. beijerinckii* strain BA101 should improve yield and productivity (16). The combined application of either mutant or genetically altered clostridia with integrated fermentation and product recovery technologies may successfully sidestep the low product tolerance characteristic of the clostridia.

Biology of Substrate Utilization

Solventogenic clostridia have the potential for converting carbohydrate substrates from excess agricultural crops or forest and agricultural wastes into solvents and fuels. The carbohydrate feedstocks for these fermentations are primarily starch, cellulose, xylans (or hemicelluloses), and lactose. Three recent reviews present the details of the enzymology, and genetics of utilization of these carbohydrates by the clostridia (62,65,66). Most of the solvent-producing saccharolytic clostridia utilize starch as a carbon source, reflecting their synthesis of extracellular and cell-bound amylolytic enzymes. Each species carries an array of three or four of the following enzymes: α-amylase, which hydrolyzes α-1,4-glucosidic linkages of amylose at random; β-amylase, which hydrolyzes α-1,4-glucosidic linkages from the ends of amalose, releasing disaccharides; glucoamylase, a hydrolyzer like β-amylase, but one releasing glucose; and pullulanase, which hydrolyzes α-1,6-glucosidic linkages or branch points of amylopectin. Some thermophilic clostridia, such as *C. thermohydrosulfuricum*, produce a novel double-headed amylopullulanase that has both α-amylase and pullulanase activity. Strain development for utilizing specific starch-containing feedstocks will require selection of strains that contain the best array of these enzymes for the effective utilization of the selected carbon source. In addition, *C. puniceum* produces pectinolytic enzymes that seem to aid in exposing the starch feedstock to the amylolytic enzymes (8). Because most of the genes involved in expressing the enzymes for starch utilization have been cloned, it is now possible to genetically engineer solventogenic clostridial strains for each starch source.

The thermophilic *C. thermocellum* has been the central object of research on bacterial utilization of cellulose as a carbon source for solvent production (66). There are also seven or so mesophilic cellulolytic clostridia that are closely related but with distinctly different arrangements of cellulases (29). Although as noted earlier ("Ethanol Fermentation"), *C. thermocellum* has major drawbacks as a candidate for commercial production of ethanol, extensive study of the genetics and biochemistry of the extracellular cellulose complex indicates that it may play a role in new strain development. During batch growth of *C. thermocellum* in media with insoluble cellulose, most of the cellulose activity is located in large-cell-associated multiprotein aggregates (cellulosomes) that also bind to cellulose fibers. Clusters of cellulosomes are observed, and the basic subunit in *C. thermocellum* strain YS as a 2.1-MDa complex including 14 different polypeptides. In other strains, cellulosomes up to 3.5 MDa consisting of 50 polypeptides have been described. Multiple forms of β-1,4-endogluconases and exogluconases plus a 210-kDa scaffolding protein are organized in the cellulosome structure.

Other separate components in *C. thermocellum* include two xylanases and two β-glucosidases. Many of the genes for these proteins have been cloned, partially sequenced,

and also expressed in *E. coli*. Thus, in future strain development, many or all of the component parts of the cellulosome from *C. thermocellum* could be genetically transferred and expressed in other solventogenic clostridia such as *C. acetobutylicum*. This later organism already produces xylanases and can grow slowly on xylans as a source of fermentable sugars. Six different strains of *C. acetobutylicum* contain both extracellular and cell bound β-1,4-exogluconases; but cannot utilize cellulose as a carbon source (65). The genetic technology is now available to develop a cellulolytic *C. acetobutylicum* that will convert cellulose to acetone and butanol.

Metabolic Pathway Inactivation and Amplification

As noted earlier ("Genetics of Solvent Formation"), the genetic methods for DNA transfer and insertional mutagenesis now available for some of the clostridial strains that produce solvents together with recent advances in cloning and sequencing of many of the important genes encoding solventogenesis and pathways of substrate utilization promise a new era of strain development. The best experimental examples for changing the ratios of fermentation products come from studies with *C. acetobutylicum*. Normally, the bacteria carrying out this fermentation produce five different organic products (as well as H_2 and CO_2) in different amounts, depending on growth conditions (see earlier: "Butanol Fermentations" and "Physiology of Control of Solvent Formation"). Ideally, during solvent production, a mutant that produced only acetone, ethanol, and butanol, with no residual organic acids, would be desirable. By transferring a plasmid carrying the three genes required for acetone production to a normal *C. acetobutylicum* (ATCC 824), one can produce an excess of the two acetone enzymes in the cell (49). This amplifies the capability to recycle the organic acids and produce a high yield of both butanol and acetone (see Fig. 6). The same result is obtained with a mutant strain, B18, of strain NRRL 643, selected for low acid production following mutagenesis (15).

The second approach to directing the flow of carbon into specific products in the *C. acetobutylicum* fermentation is "gene knockout" by insertional mutagenesis. Recent experiments indicate that inactivation of acetyl kinase or of butyrate kinase prevents acetate or butyrate production, respectively, and permits an earlier and increasing flux toward butanol. Since autolysin, an enzyme produced by *C. acetobutylicum*, causes a late growth phase self-lysis, inactivation of this enzyme might also prolong solvent production. Autolysin-deficient mutants of *C. acetobutylicum* strain P262 show increased cell stability and a higher butanol tolerance.

Another problem facing the use of *C. acetobutylicum* for a continuous fermentation is the natural coupling of solventogenesis with development of endospores. The normal biology of the organism has coordinated these and other processes to adapt to organic acid stress and other stress situations. In *C. acetobutylicum* (ATCC 824), the loss of the large 210-kb pSOL-1 plasmid causes loss of the *sol* operon as well as certain unknown genes required for sporulation and granulose formation (42). Thus, in strain ATTC 824, loss of the pSOL-1 plasmid may serve as an advantage for development of a continuously fermenting, solvent-producing strain. By transferring the genes of the *sol* operon and the *adc* gene on a plasmid back into a pSol-1-negative strain, it is possible to induce solvent formation uncoupled from spore formation (42). In a number of strains of *Clostridium* as well as the butanol-producing strains, the initiation of sporulation and solvent formation is controlled by a complex phosphorelay mechanism as in *B. subtilis*, as noted earlier ("Is Solvent Production a Cellular Response?"). Further study of this regulatory network in the clostridia may reveal ways to upregulate solvent formation and also block initiation of endospore formation.

Since microbiologists estimate that only about 0.1% or less of all the bacteria in nature have ever been successfully isolated, there must be a large untapped natural pool of uncharacterized clostridia with valuable capabilities worthy of the attention of industrial microbiologists. For instance, in 1979 only five clostridial thermophilic bacteria were known. By 1995 more than 16 clostridial thermophiles had been purified (26). Recent reclassification demonstrates that *Thermoanerobium* sp. and *Thermoanaerobacter* sp. are close relatives to the clostridia, which expands the number of candidates for new strain development to more than 25. Thus, just considering growth temperatures in the range of 55–70 °C as a desirable property, many natural species are only recently isolated from nature. Perhaps a program of genetic engineering should be preceded by the application of new enrichment techniques to a search for natural species with new pathways, increased solvent production, or increased tolerance to chemical or physical stress.

ACKNOWLEDGMENT

The work reported from our laboratory was supported by grant DE-FG02-86ER-13512 from the U.S. Department of Energy.

BIBLIOGRAPHY

1. H. Hippe, J.R. Andreesen, and G. Gottschalk, in A. Balows, H.G. Trüper, M. Dworkin, W. Harder, and K.-H. Schleifer eds., *The Prokaryotes*, 2nd ed., Springer-Verlag, New York, 1991, pp. 1800–1866.
2. M.D. Collins, P.A. Lawson, A. Willems, J.J. Cordoba, J. Fernandez-Garayzabal, P. Garcia, J. Cai, H. Hippe, and J.A.E. Farrow, *Int. J. Syst. Bacteriol.* **44**, 812–826 (1994).
3. J.L. Johnson and J.-S. Chen, *FEMS Microbiol. Rev.* **17**, 223–240 (1995).
4. S. Keis, C.F. Bennett, V.K. Ward, and D.T. Jones, *Int. J. Syst. Bacteriol.* **45**, 693–705 (1995).
5. F.A. Rainey, N.L. Ward, H.W. Morgan, R. Toalster, and E. Stackebrandt, *J. Bacteriol.* **175**, 4772–4779 (1993).
6. Y.-E. Lee, M.K. Jain, C. Lee, S.E. Lowe, and J.G. Zeikus, *Int. J. Syst. Bacteriol.* **43**, 41–51 (1993).
7. W.D. Murray, A.W. Kahn, and L. Van den Berg, *Int. J. Syst. Bacteriol.* **32**, 132–135 (1982).
8. J.G. Morris, in D.R. Woods ed., *The Clostridia and Biotechnology*, Butterworth-Heinemann, Boston, Mass., 1993, pp. 1–24.

9. P. Tailliez, H. Girard, J. Millet, and P. Beguin, *Appl. Environ. Microbio.* **55**, 207–211 (1989).
10. J.G. Zeikus, *Annu. Rev. Microbiol.* **34**, 423–464 (1980).
11. C.K. Lee and J. Ordal, *J. Bacteriol.* **94**, 530–536 (1967).
12. E.J. Hsu and Z.J. Ordal, *J. Bacteriol.* **102**, 369–376 (1970).
13. W.D. Murray and A.W. Kahn, *Can. J. Microbiol.* **29**, 342–347 (1983).
14. H. Bahl and G. Gottschalk, *Biotechnol. Bioeng. Symp.* **14**, 215–223 (1984).
15. P. Rogers and N. Palosaari, *Appl. Environ. Microbiol.* **53**, 2761–2766 (1987).
16. J. Formanek, R. Machi, and H.P. Blaschek, *Appl. Environ. Microbiol.* **63**, 2306–2310 (1997).
17. H.A. George, J.L. Johnson, W.E.C. Moore, L.V. Holdeman, and J.S. Chen, *Appl. Environ. Microbiol.* **45**, 1160–1163 (1983).
18. R.A. Holt, A.J. Carins, and J.G. Morris, *Appl. Microbiol. Biotechnol.* **27**, 319–324 (1988).
19. M. Gottwald, H. Hippe, and G. Gottschalk, *Appl. Environ. Microbiol.* **48**, 573–576 (1984).
20. U.S. Congress, Office of Technology Assessment, *Commercial Biotechnology: An International Analysis*, OTA-BA-218, Government Printing Office, Washington, D.C., 1984, pp. 237–250.
21. S.C. Prescott and C.G. Dunn, *Industrial Microbiology*, 3rd ed., McGraw-Hill, New York, 1959, pp. 240–284.
22. R.K. Thauer, K. Jungermann, and K. Decker, *Bacteriol. Rev.* **41**, 100–180 (1977).
23. H. Petitdemange, C. Cherrier, G. Raval, and R. Gay, *Biochim. Biophys. Acta* **421**, 334–347 (1976).
24. P. Rogers, *Adv. Appl. Microbiol.* **31**, 1–60 (1986).
25. P. Rogers and G. Gottschalk, in D.R. Woods ed. *The Clostridia and Biotechnology*, Butterworth-Heinemann, Boston, Mass., 1993, pp. 25–50.
26. F. Canganella and J. Wiegel, in D.R. Woods ed., *The Clostridia and Biotechnology*, Butterworth-Heinemann, Boston, Mass., 1993, pp. 393–430.
27. R. Lamed and J.G. Zeikus, *J. Bacteriol.* **144**, 569–578 (1980).
28. Y. Mori, *Appl. Environ. Microbiol.* **56**, 37–42 (1990).
29. L.B. Leshine, *Annu. Rev. Microbiol.* **49**, 399–426 (1995).
30. W.D. Murray, *Appl. Environ. Microbiol.* **51**, 710–714 (1986).
31. S.R. Wilkinson, D.I. Young, G. Morris, and M. Young, *FEMS Microbiol. Rev.* **17**, 275–285 (1995).
32. D.T. Jones and D.R. Woods, *Microbiol. Rev.* **50**, 484–524 (1986).
33. H. Bahl and G. Gottschalk, in H.J. Rehm and G. Reed eds., *Biotechnology*, vol. 6B, VCH, Weinheim, Germany, 1988, pp. 1–30.
34. P. Dürre, N.P. Minton, E.T. Papoutsakis, and D.R. Woods eds., Papers presented at a meeting on solventogenic *Clostridia*. *FEMS Microbiol. Rev.* **17**, 221–364 (1995).
35. C.W. Forsberg, *Appl. Environ. Microbiol.* **53**, 639–643 (1987).
36. L. Girbal, C. Croux, I. Vasconcelos, and P. Soucaille, *FEMS Microbiol. Rev.* **17**, 287–297 (1995).
37. H. Grupe and G. Gottschalk, *Appl. Environ. Microbiol.* **58**, 3896–3902 (1992).
38. C.L. Meyer and E.T. Papoutsakis, *Appl. Microbiol. Biotechnol.* **30**, 450–459 (1989).
39. D.T. Jones, in D.R. Woods ed., *The Clostridia and Biotechnology*, Butterworth-Heinemann, Boston, Mass., 1993, pp. 78–98.
40. R.C. Woolley and J.G. Morris, *J. Appl. Bacteriol.* **69**, 718–728 (1990).
41. E.R. Kashket and Z.-Y. Cao, *FEMS Microbiol. Rev.* **17**, 307–315 (1995).
42. E. Cornillot, R.V. Nair, E.T. Papoutsakis, and P. Soucaille, *J. Bacteriol.* **179**, 5442–5447 (1997).
43. S.L. Landuyt, E.J. Hsu, and M. Lu, *Ann. N.Y. Acad. Sci.* **413**, 473–478 (1983).
44. D.M. Rothstein, *J. Bacteriol.* **165**, 319–320 (1986).
45. P. Dürre, in D.R. Woods ed., *The Clostridia and Biotechnology*, Butterworth-Heinemann, Boston, Mass., 1993, pp. 228–246.
46. D.M. Mattsson and P. Rogers, *J. Ind. Microbiol.* **13**, 258–268 (1994).
47. M. Young, in D.R. Woods ed., *The Clostridia and Biotechnology*, Butterworth-Heinemann, Boston, Mass., 1993, pp. 99–118.
48. N.P. Minton, J.K. Brehm, T.-J. Swinfield, S.W. Whelan, M.L. Mauchline, N. Bodsworth, and J.D. Oultrum, in D.R. Woods ed., *The Clostridia and Biotechnology*, Butterworth-Heinemann, Boston, Mass., 1993, pp. 119–150.
49. L.D. Mermelstein, N.E. Welker, D.J. Petersen, G.N. Bennett, and E.T. Papoutsakis, *Ann. N.Y. Acad. Sci.* **721**, 54–68 (1994).
50. E.T. Papoutsakis and G.N. Bennett, in D.R. Woods ed., *The Clostridia and Biotechnology*, Butterworth-Heinemann, Boston, Mass., 1993, pp. 157–200.
51. P. Dürre, R.-J. Fischer, A. Kuhn, K. Lorenz, W. Schreiber, B. Stürzenhofecker, S. Ullmann, K. Winzer, and U. Sauer, *FEMS Microbiol. Rev.* **17**, 251–262 (1995).
52. C.J. Bult, O. White, G.J. Olsen, and 37 other authors, *Science*, **273**, 1058–1073 (1996).
53. D. Holzman, *Am. Soc. Microbiol. News* **62**, 8–9 (1996).
54. E.M. Green, Z.L. Boynton, L.M. Harris, F.B. Rudolf, E.T. Papoutsakis, and G.N. Bennett, *Microbiology* **142**, 2079–2086 (1996).
55. P.J. Piggot, C.P. Moran, Jr., and P. Youngman, *Regulation of Bacterial Differentiation*, American Society for Microbiology, Washington, D.C. 1994.
56. U. Sauer, J.D. Santangelo, A. Trever, M. Bucholz, and P. Dürre, *FEMS Microbiol. Rev.* **17**, 331–340 (1995).
57. J.A. Hoch, *Annu. Rev. Microbiol.* **47**, 441–465 (1993).
58. D.P. Brown, L. Ganova-Raevn, B.D. Green, S.R. Wilkinson, M. Young, and P. Youngman, *Mol. Microbiol.* **14**, 411–426 (1994).
59. C.A. Gross, in F.C. Neidhardt ed., *Escherichia coli and Salmonella Cellular and Molecular Biology*, vol. 1, 2nd ed. ASM Press, Washington, D.C., 1996, pp. 1382–1399.
60. H. Bahl, H. Müller, S. Behrens, H. Joseph, and F. Narberhaus, *FEMS Microbiol. Rev.* **17**, 341–348 (1995).
61. M. Hecker, W. Schumann, and U. Volker, *Mol. Microbiol.* **19**, 417–428 (1996).
62. I.S. Maddox, N. Qureshi, and N.A. Gutierrez, in D.R. Woods ed., *The Clostridia and Biotechnology*, Butterworth-Heinemann, Boston, Mass., 1993, pp. 343–370.
63. Q. Geng and C.-H. Park, *Biotechnol. Bioeng.* **43**, 978–985 (1994).
64. J.R. Gapes, D. Nimcevic, and A. Friedl, *Appl. Environ. Microbiol.* **62**, 3210–3219 (1996).
65. K. Bronnenmeier, and W.L. Staudenbauer, in D.R. Woods ed., *The Clostridia and Biotechnology*, Butterworth-Heinemann, Boston, Mass., 1993, pp. 261–310.
66. G.P. Hazlewood and H.J. Gilbert, in D.R. Woods ed., *The Clostridia and Biotechnology*, Butterworth-Heinemann, Boston, Mass., 1993, pp. 311–342.

See also ANAEROBES; CULTURE COLLECTIONS.

COAGULATION FACTORS, THERAPEUTIC

DAVID L. ARONSON
Consultant
Bethesda, Maryland

KEY WORDS

Antithrombin
CPMP
Factor VIII
Food and Drug Administration
Gene therapy
Hemophilia
Recombinant coagulants
Transgenics
von Willebrand factor

OUTLINE

INTRODUCTION

For the past 40 years, therapeutic preparations of plasma coagulation factors have been derived from plasma. The plasma fractionation industry developed during and after the second World War and was often a partnership between a governmental or not-for-profit organization that obtained the blood from volunteer donors and the nascent commercial plasma industry. Since the beginning of blood banking and the production of plasma products, the regulatory oversight has been done by the national organizations responsible for biologics control. The authority and the procedures of these organizations are different than those for pharmaceutical regulation.

Regulation of the plasma products in the United States was based on the Public Health Service Act of 1902. This act was promulgated to regulate the production of vaccines and therapeutic serums. The latter phrase was construed to include blood and plasma products when this industry developed later on. Compliance was monitored by the Public Health Service Laboratories, later to become the National Institutes of Health. In 1972, the regulation of U.S. biologics was moved to the Food and Drug Administration (FDA).

One of the major efforts in any of the biological control agencies was the development of standards and assays for impure and poorly characterized natural products. This major component of biologics regulation is seen in the name of the former U.S. control organization, Division of Biologics Standards at NIH, and in the U.K., the National Institute of Biological Standards and Control (NIBSC).

Although regulation has properly belonged to national regulatory bodies, there are other inputs to the regulatory process from supranational organizations, such as the European Community and the International Conference on Harmonisation (ICH). In the United States, the FDA requires that an advisory committee votes for approval of licensure. These committees include representatives of advocacy groups that have voting membership in the FDA blood product advisory committee (BPAC). The recommendations of professional organizations are also quoted in regulatory documents or regulatory decisions. The Scientific and Standardization Committee of the International Society of Thrombosis and Haemostasis has provided recommendations for the hemophilias and thrombotic disease. In addition, civil liability law has an impact on all therapeutic producers, particularly in the United States.

Regulatory documents of the FDA are published in the Code of Federal Regulations (CFR), and updated guidance documents can be obtained from the FDA World Wide Web site (*http://www.fda.gov*). European documents are found in the European Pharmacopeia, and guidance documents are obtained from the Committee on Proprietary Medicinal Products (CPMP).

Blood Coagulation

Blood coagulation results from a series of proteolytic reactions that first activate a procoagulant protein, and then the resulting coagulant is inactivated either proteolytically or by interaction with an inhibitor (1) (Fig. 1). Although presented as the coagulation cascade (Fig. 2), implying amplification, the complexity results in many points where there is sensitive biochemical control. Individual components are identified by roman numerals, and the activated form is indicated by a lowercase "a."

Factor IX, the defective protein in hemophilia B, is the proenzyme of a serine protease that is cleaved to activated factor IX (IXa). Factor VIII, the defective or deficient pro-

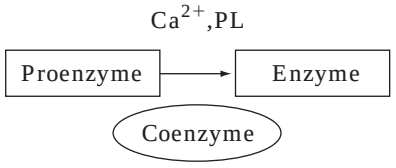


Figure 1. Prototype coagulation step.

tein in hemophilia A, is a protein cofactor that circulates bound to a large polymeric protein, von Willebrand's factor, which is defective in von Willebrand's disease. In the bound state the factor VIII remains in the plasma; in the presence of small amounts of thrombin, it is released from the von Willebrand factor and activated. In the presence of factors IXa and VIIIa there is activation of the downstream enzymes in the coagulation cascade (Fig. 2). Absence or defects of either factor IX or factor VIII leads to inability to form fibrin clots. Absence of the carrier protein, von Willebrand's factor, leads to low circulating levels of factor VIII as well as decreased recruitment of platelets to damaged endothelium and leads to a hemorrhagic state.

The factor VIII gene product is ~250,000 kDa derived from a gene of 186 kb in 26 exons. The gene product is glycosylated intracellularly, and the tyrosine at position 1680 must be sulfated. During passage through the Golgi apparatus there is excision of much of the central portion of the protein (B-domain), resulting in a heavy and light chain held together by Ca^{2+}. This is the form found in the circulation.

Genetic Hemorrhagic Disorders

Hemophilia A and hemophilia B are the result of mutation of genes located on the X chromosome. As such they are sex-linked, and sons of carrier mothers have a 50% likelihood of carrying the defective gene and presenting with the clinical phenotype of abnormal hemorrhage. This leads to crippling and death in those severely afflicted.

Treatment of these diseases began with plasma and evolved to more purified and concentrated plasma products (2). Adequate treatment for those with severe hemophilia involves lifetime intravenous administration given approximately weekly. Although the treatment is effective, there are two major safety issues with plasma products: blood-borne viral infections and the development of antibodies that inhibit the therapeutic effect. Hemophilia A and B were one of the early targets of the biotechnology industry.

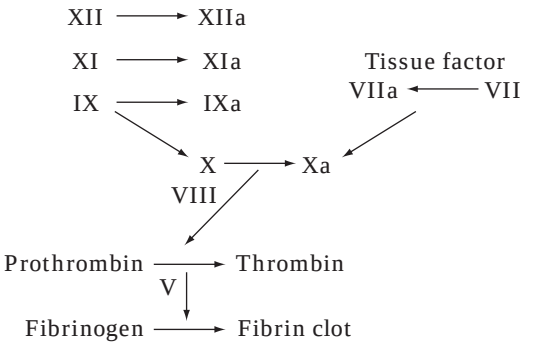


Figure 2. Coagulation sequence.

The patient population is small, with 16,000 patients with hemophilia A (deficiency of factor VIII) and hemophilia B (deficiency of coagulation factor IX) in the United States. Of these, about 8,000 have severe disease and need frequent treatment. Treatment with plasma products temporarily improved the well-being of this patient population until the HIV epidemic; 6,600 U.S. patients were infected (3). Patients with severe hemophilia A need frequent and lifelong treatment; this group uses ~50,000 to 100,000 factor VIII units annually.

Viral infection from plasma products has been the major problem in the treatment of the congenital bleeding disorders. Infection with most blood-borne viruses has been minimized during the past 10 years by viral inactivation. The most common procedure is the destruction of the lipid-coated viruses (HIV, hepatitis B and C) by extraction with a solvent-detergent mixture. Another procedure is the use of heating either in solution or in the dry state. Many regulatory authorities are demanding two viral inactivation and removal steps for recombinant and plasma-derived factor VIII and factor IX.

The second problem for the treatment of hemophilia A is that 15% of patients will develop inhibiting antibodies to factor VIII and become refractory to standard treatment. Although most patients who develop antibodies may be genetically predisposed, in other cases this is because of neoantigenicity induced in the factor VIII during production and viral inactivation (4).

A third hemorrhagic disease, von Willebrand's disease, is the result of heterogeneous mutations of the gene for a polymeric carrier protein of factor VIII, von Willebrand's factor.

REGULATORY ISSUES FOR RECOMBINANT COAGULATION FACTORS

Statutory drug law is based on the principles of safety and efficacy. Regulatory oversight of pharmaceutical production is done in large part through good manufacturing practices (GMPs) and good laboratory practices (GLPs). This is a body of regulation and policy that can change rapidly. In addition, in the regulation of biologics there is a need for product consistency whether the product is derived from animals or from cells. The particular origin of biologics leads to special needs in the control of the starting materials, including cells, genes, etc. This is best exemplified in the regulations for the establishment of intensively characterized master cell banks for cell-derived products (5). This cell bank will be expanded in number under standard conditions to produce a consistent product over decades.

The advent of the biotechnology products for coagulation deficiencies (i.e., cell-derived rather than plasma-derived coagulation factors) raised few unique regulatory issues. Traditional biologics regulation had already dealt with cell-derived products, including the potentially dangerous live virus vaccines. A regulatory framework for the characterization and control of cell lines had already been

constructed, and in 1984, a workshop sponsored by the Bureau of Biologics dealt with many of the residual issues. This was followed by a WHO report in 1987 (6,7). This was followed by the development of *Points-to Consider* for products derived from mammalian cells (8) and the EEC equivalent (9). The *Points-to-Consider* are not legal regulatory documents, but they inform those who are dealing with the agency of the current views of the regulatory authority. There is an effort to have as much commonality as possible between the international regulatory agencies, and extensive efforts to harmonize the guidelines are ongoing by the ICH (10). Of particular concern originally was the quality and quantity of residual DNA in the final product (11); a fear that has diminished with time.

The FDA does not produce specific documents on coagulation or fibrinolytic products. In contrast, the European Community has specific and detailed regulatory documents for the coagulation products (12). The basis for regulation of recombinant coagulation factors is similar to plasma-derived coagulation factors: the assessment of safety, efficacy and consistency. More particularly, there is intensive and extensive evaluation of the cell production systems, the control and standardization of the final product, and clinical effectiveness and safety. The main point of the original regulation for cell-derived coagulation factors was that the products should be identical with plasma-derived products. More recently, products with similar but not identical structures have been received by regulatory authorities with little additional concern.

The GLPs and GMPs are in principle identical around the world; however, there can be significant differences in detail. There has been rapid change in pharmaceutical science during the past decade, including filling procedures, air and water systems, and analytical methods. Approved production facilities of today may not meet GMPs in 5 years. In addition, regulatory policy can change without any formal notification or publication. For these reasons, it is important that the manufacturer maintain constant communication with the regulatory agency from inception to final approval.

Many facilities for recombinant products will be designed and maintained as single product manufacturing areas. This will include raw materials, cell lines, expansion of cell lines, fermenters, purification equipment, and the filling line. It will only be for the postfilling operations that there will be contact with other products. Some useful products may have such small demand that a few lots will meet needs for several years. Such a case may be recombinant activated factor VII (VIIa) for use in hemophilics with inhibitors not responsive to other modes of therapy. In this case, the plant must have a validated procedure for producing a different product.

Cellular Issues in Recombinant Coagulation Products

All cells used for the production of recombinant coagulation factors must either meet the *Points-to-Consider* documents by the FDA or have information indicating that there is no compromise to the safety and efficacy of the product if some of the points are modified. The relevant *Points-to-Consider* are those involved with production of the master cell banks (5) and the *Points-to-Consider* for products derived from mammalian cells (8). The European equivalents are CPMP *Note for Guidance: Production and Quality Control for Medicinal Products Derived by Recombinant DNA Technology* (9) and the ICH documents on the analysis of the expression construct (10).

The cells must not only synthesize the desired protein, but efforts must be made to simplify the purification process and minimize the presence of adventitious infectious agents by simplification of the growth medium. In particular, the use of fetal calf serum is avoided if at all possible to minimize a significant source of microbial and antigenic contamination. It is likely that regulatory agencies in the future will require more effort to develop culture systems free of serum or other animal protein.

During the production time, there must be no evidence of either microbial contamination or of change in the cell or the cell product. Both the U.S. and European documents require characterization of the vector to demonstrate the stability of the inserted gene over the time of a production run or longer. The development of PCR technology has allowed the characterization of cDNA and RNA. Frequent sampling and testing of the conditioned media during production and the final purified product are a requirement.

The cells must be capable of appropriate post-translational modifications. In the case of factor IX, there must be carboxylation of the 10 to 11 glutamic acids so that the molecule has full function. This has been accomplished by inclusion of a gene for a protease to remove the propeptide and allow carboxylation of the glutamic acids at the NH_2 terminal of the factor IX molecule. Other significant post-translation modifications include the formation of a β-OH aspartate residue, sulfation, phosphorylation, and glycosylation. In addition, there must not be the production of any protease that will degrade or activate factor IX.

A similar constraint is true for cells producing factor VIII. This is a heavily glycosylated molecule, and each production system could yield different products, some of which might be immunogenic. Although this is a concern, no such immunogenic derivatives have been found as yet in factor VIII produced by recombinant technology.

Von Willebrand's factor must be produced in cells allowing polymerization of the monomer and secretion of these virus-sized protein molecules into the media. The cell culture system must also be able to sulfate a specific tyrosine residue to allow binding to factor VIII and thrombin activation of the factor VIII.

Purification Issues in Recombinant Coagulation Products

Purification of coagulation products revolves around three issues: the removal of foreign allergens, the removal of DNA, and the removal of adventitious agents originating from the host cell and the media. The issue of allergens is of importance because the recipients need lifelong therapy. In contrast to the older plasma products, the purification process using mouse immunoglobulin affinity columns may be a new hazard, and at this time, monoclonal antibody purification is the universal purification method for factor VIII. The monoclonal antibodies do not have to meet the full requirements for a therapeutic monoclonal antibody but must have been examined for murine viruses.

There are no absolute requirements or regulations for protein purity within the U.S. or European regulations.

However, the manufacturer is obligated to have satisfactory test systems to detect nontherapeutic proteins derived either from the cell source or the purification procedure (e.g., leached monoclonal antibodies). The ultimate test will be in recipients of the product (see "Clinical Trials").

The factor IX product being developed not only is grown in media without protein but the purification steps do not involve monoclonal antibodies. This strategy simplifies the purification process and is also deemed desirable in terms of the introduction of viruses either in tissue culture or in the final product (13). Although this is apparently the first coagulation product from cells grown in the absence of protein in the media and purified without the use of monoclonal antibodies, it may set a new standard.

The same procedures for the removal and inactivation of viruses that are required for cell-derived products are required for the plasma products. Although the use of cell-derived products in contrast to plasma-derived products will only rarely contain viruses transmissible to humans, the newer knowledge of virology and the public perception demands extreme measures. The manufacturer must show that the purification methodology has the capability to remove or inactivate viruses of different types. However, the European guidelines are much more specific than the comparable U.S. documents (14).

One of the first issues for cell derived products was the establishment of limits for DNA. This was set at 10 pg/dose for all products in the original FDA documents (10), but the current WHO level of 100 pg/dose (7) seems to be the actual figure used by the FDA. Products used for the treatment of hemophilias are infused 1,000 times in a lifetime; however, there is no change in the dosage for products used a few times as opposed to 1,000 times in a lifetime. The requirements for DNA removal revolve around the theoretical ability of cell-derived DNA to be incorporated into somatic cells in the recipient (11).

Final Product

The structure of the gene product must be extensively studied. This will include extensive peptide mapping, molecular weight analysis by electrophoretic assessment, and perhaps mass spectrometry. In some cases such as factor IX, the complete amino acid structure is feasible, whereas for factor VIII the total amino acid sequence is more onerous than useful in the light of the established gene structure. Disulfide bonding should be characterized and shown to be identical to plasma-derived coagulants. As noted earlier, there are specific post-translational modifications for given products important for their function, such as glutamyl carboxylation of factor IX, that should be characterized.

The final product will include degradation products of the gene product. These should be identified, analyzed for their sequence, and discriminated from other protein contaminants.

A specific issue for the coagulation factors is that they are in the appropriate state (i.e., nonactivated for factor VIII or IX). For von Willebrand's factor, the polymeric forms must be functional. As noted earlier, the circulating form of the coagulant proteins should be the proenzyme; factor IXa can induce thrombus formation. In the case of the cofactor, factor VIII, the uncleaved, nonactivated form is necessary; the presence of the activated form yields inaccurate potency measurements and is rapidly cleared. Thus, the purification system should remove the activated forms and be done in a manner that does not produce these forms.

Preclinical Testing

In general, the preclinical testing of recombinant products is similar to plasma-derived products and includes routine pyrogen, safety, and potency testing. Specific and accurate potency measurements for labeling and dosing must be available.

Potency Measurement

The assumed efficacy of the final product is assessed by the potency assay, the results of which yield the labeled value for treatment. The standard assay is a bioassay with intrinsic variability (2,15). The unit definition for coagulation products is the functional activity found in 1 mL of normal plasma containing a 10th volume of anticoagulant. This operational definition is translated into WHO standards for coagulants by the preparation of plasma or plasma fractions that originally were compared to normal plasmas. All WHO coagulation standards have been prepared by the NIBSC in the United Kingdom. These are calibrated in multicenter studies organized by the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis. New standards are compared to previous WHO standards and also to plasma standards. As both factor VIII and factor IX therapeutic products become more purified, there are inconsistencies with older standards (15). A further complication is the development of what are apparently more accurate quantitative potency tests based on better biochemical knowledge but that can show very substantial differences with the standard reference tests (16–18). The European Pharmacopia designates a reference methodology for the measurement of coagulation factors. In general, this seems to be the assay used by the control laboratory rather than that with which there is the most experience. At this writing, the reference methodology for factor VIII in the European Pharmacopiea is the chromogenic assay. In contrast, the FDA has no designated reference method for coagulation factors but does distribute national references. There is a consistent difference in the pharmacokinetics when assessed by different methods (17,18).

The development of variant proteins such as the B-domain deleted factor VIII aggravates the difference seen between the different methods of measuring potency. The specific activity of the B-domainless variant is reported as being substantially higher than plasma-derived factor VIII (17). The in vitro potency assays are approximately equivalent on a molar basis. Whether they are the same therapeutic quantity must be answered.

A recombinant von Willebrand factor is in development (19). Von Willebrand's factor has several different functional attributes of importance to its efficacy. These include the support of platelet aggregation, binding to collagen, and the binding to factor VIII. At this time there is no

consensus as to the appropriate potency test for a therapeutic product to treat von Willebrand's disease. The WHO standard for von Willebrand's factor is plasma and not a purified preparation.

Animal Models

Animal models add useful information before clinical testing of the products. A dog model of hemophilia A has been used extensively in the preclinical testing of factor VIII. One of the most immediate questions was the comparison of the pharmacokinetics of recombinant factor VIII with the plasma product. If the proposed product was more rapidly removed from the circulation than the plasma product, this would raise serious doubts about the efficacy of the product. Experiments in the dog hemophilia model showed normal pharmacokinetics and rapid binding of factor VIII to canine von Willebrand's factor (20). Various approaches have been and are being assessed to determine abnormal immunogenicity with animal models: induction of immune tolerance in newborn animals in addition to laboratory comparison of the epitopes of the plasma and recombinant product (21).

Thrombogenicity of the factor IX preparations has been assessed by both laboratory testing and assessment of the ability to form clots in animals (22). Although this concern arises from thrombosis associated with the use of older, less-purified products, there is still concern that even the purified products and recombinant factor IX can produce thrombosis. Animal pharmacokinetics comparing plasma-derived and recombinant factor IX are a necessary prelude to clinical studies. Dog models of hemophilia B are available, but the animal pharmacokinetic studies are generally performed in smaller laboratory animals. Because there is endogenous factor IX coagulant activity in the samples, the plasmas are assayed for the quantity of human factor IX antigen present over time (23).

Animal models for coagulation factors have been developed over the years for specific known adverse effects. In addition, animal testing is of value for identifying unsuspected effects.

Clinical Trial Information

Clinical trials are the final step in the assessment of the efficacy and safety of a new therapeutic product. The FDA regulations outline the information needed to begin human studies and the type of studies needed to establish safety and efficacy (24). European regulations are similar. Efficacy of treatment of the hemophilias has never been established in a placebo-controlled study because observational studies coupled with historical data clearly established the effectiveness originally of both factor VIII and factor IX replacement therapy for preventing or treating hemorrhagic episodes. This type of data was acceptable to the regulatory authorities around the world in the mid-60s and was consistent with the biologics definition of efficacy (25) based on the pharmacologic and pharmacokinetic studies. Current policy would, however, require comparison with an approved efficacious product for the factor VIII and factor IX products. The FDA does not prescribe an exact protocol but describes the general types of clinical trials that could give efficacy data (26). In contrast, the EEC guidelines describe in detail the clinical trials to be undertaken (12). Many of these details are derived from articles promulgated by a nonregulatory professional organization, the International Society on Thrombosis and Haemostasis (ISTH) (27,28).

In addition to establishing efficacy, clinical trials are designed to detect adverse events. The most common concern is for known or unknown infectious agents. Recommendations of the ISTH serve as guidelines for some of the European community policies.

The original protocols for recombinant factor VIII started with pharmacokinetic trials in older patients with frequent need for treatment. The strategy was to keep these patients solely on recombinant factor VIII for several years. This would give ample opportunity to test any abnormal immunogenicity of the product. In addition, there would be interim pharmacokinetic studies as the most accurate way of assessing low-level inhibitors. A trial was also done in previously untreated patients (PUPs) independent of any request from a regulatory agency (29). This study yielded valuable results on the treatment of hemophilia in very young patients without the confounding of immunologic or viral infection by other products. This set a precedent for the evaluation of all future recombinant products and has been incorporated into the EEC requirements for all recombinant factor VIII and factor IX. This will probably also be required by the FDA in the future. Although the immunogenicity of recombinant factor VIII and plasma factor VIII is a major concern, it is not clear what information comes from the routine PUP studies, because any excess immunogenicity should be observed in older patients. There are no specific European clinical guidelines for recombinant coagulation factors. There are the general clinical guidelines (12) for the coagulation products, which specifically exclude recombinant products (6,8). Based on the historical knowledge and precedent, there is no requirement for a blinded efficacy trial. Efficacy assessment is based on the comparison of pharmacokinetics with an approved product. For factor VIII, the requirements are to study 12 pretreated (PTP) subjects over 1.5 days with eight sampling times. A minimum of three lots are to studied. These subjects should be maintained solely on the experimental product for a further 3 to 6 months and then have a repeat pharmacokinetic study. Before approval, a study of pharmacokinetics in 20 PUPs must be started using the same protocol as in the other patients. There is a requirement for at least 10 surgical procedures in a minimum of 5 patients and clinical assessment of 30 immunocompetent patients after licensure.

In addition, information should be submitted on 30 patients who have been treated on at least 10 days with at least 6 months of follow-up. The factor VIII inhibitor titer must be repeated every 3 months. After approval, there is a periodic update on the laboratory parameters of at least 50 patients treated for at least 2 years.

With plasma-derived factor IX, the major adverse event other than viral disease transmission was thrombosis associated mostly with the earlier less-purified plasma fractions containing large amounts of the other vitamin K coagulants. The clinical trials for a recombinant factor IX

must test for evidence of in vivo coagulation with currently available tests for thrombin–antithrombin complex and prothrombin fragment 1.2. In addition, there will probably be required testing for the enzymatically active form of factor IX (IXa). Although the FDA does not specifically deal with this issue, this still must be addressed for U.S. licensure. The European guidance documents call for tests for in vivo activation of coagulation.

PUP studies will be required by the FDA and by the European Community for factor VIII and factor IX products. The European community requirements for plasma-derived factor IX require pharmacokinetic studies in 10 subjects with 10 sampling times. Treatment with the same product should be maintained for 6 months, and repeat pharmacokinetic studies should be performed in 3 to 6 months. Current published clinical experience with recombinant factor IX are limited to abstracts indicating pharmacokinetics and clinical effectiveness similar to plasma-derived factor IX (30). These studies are not inconsistent with the European guidelines and were accepted by the BPAC (December 19, 1996).

Von Willebrand's factor is a new molecule to deal with, and there are no national regulatory comments addressing the therapeutic use of this particular protein. Although the pharmacokinetic analysis of the clearance of the coagulation activity of factors VIII and IX is accepted throughout the world, there is no accepted or validated functional assay for von Willebrand's factor. Should the functional assay be based on the ristocetin cofactor activity or the ability to bind to platelets or the content of multimers or the content of intact monomers or the ability to bind factor VIII?

One can anticipate that the regulatory agencies will use the SSC/ISTH guidelines for appropriate clinical studies (31). The current FDA policy is to request comparative clinical data with another product, but there is no product licensed in the United States for the treatment of von Willebrand's disease.

Although von Willebrand's disease is a common genetic deficiency, the clinical phenotype is extremely variable, and for the most part, the hemorrhages are less severe and less frequent than in hemophilia A or B. This leads to difficulty in recruiting patients into a structured protocol. Heretofore, the clinical trials of von Willebrand's factor have been uncontrolled, and current treatment strategies are based on overall experience. The surrogate markers that have been used traditionally are the level of factor VIII, the shortening of the bleeding time, and the functional level of the von Willebrand factor as assessed by the ability to support platelet aggregation in the presence of ristocetin. Inhibitor development is an issue in von Willebrand's disease, and this will have to be assessed. With recombinant von Willebrand's factor, the factor VIII level will be endogenously regulated and will increase to peak levels 10 to 20 hours after infusion. This is in contrast to the current practice of infusing patients with plasma products containing von Willebrand's factor and factor VIII with instantaneous increase in the factor VIII level.

Two recombinant factor VIII products have been licensed in the United States and in Europe. Recombinant factor IX is in clinical trials and the license was submitted to the FDA in 1996. Recombinant factor VIIa, a product proposed for use in factor VIII and IX patients with inhibitors, is being reviewed by the FDA (1997) and has been licensed in Europe and Canada for the treatment of patients who acquire inhibitors to factor VIII. A recombinant von Willebrand factor is undergoing preclinical development.

Tissue Plasminogen Activator (Alteplase)—Regulatory Issues

The in vivo formation of blood clots is usually a process limited in time and space. The formation of thrombi, pathologic blood clots, are the result of subacute or chronic disease processes resulting in blockage of blood vessels. The development of fibrinolytic agents, enzyme that digest fibrin clots, was the first major advance in the treatment of thrombotic diseases in 50 years. The biologics used for this purpose are activators of the endogenous proenzyme, plasminogen. Tissue plasminogen activator (tPA) (alteplase) was the first fibrinolytic agent produced by recombinant technology.

The regulatory issues for the licensure of tPA are documented in contemporary news reports (31) and in a recent book by Beebe and Murano (32).

Production of tPA

One technical issue was emphasized: the effect on product of altering culture conditions. Early production of tPA in roller bottles was used for the initial clinical investigation, including the initial pharmacokinetic studies. When there was change to fluid-phase culture, there was a significant decrease in the pharmacokinetics, leading to the need for an increase in the clinical dosage. It was originally proposed that this was due to the chain structure of the tPA but that has been ruled out (33). Despite investigation of the carbohydrates and other possible structural changes, the post-translational change resulting in the pharmacokinetic change is still a mystery.

CLINICAL TRIALS

Gene Therapy in Hemorrhagic Diseases

Gene therapy is the procedure in which a new gene is transplanted into the body of a host to improve its health. Recombinant proteins are produced by placing the desired gene in a cell system and producing the desired gene product in vitro. Ex vivo gene therapy is accomplished by transplanting cells containing the desired gene into the host; these cells will then propagate and produce the desired biologic effect. In vivo gene therapy injects a gene construct that will enter the cell of the host and produce the desired protein.

Hemophilias A and B are prime targets for gene therapy. The ideal of a one-time or infrequent treatment as opposed to weekly intravenous injections is attractive, but despite substantial advances there are many hurdles to overcome (34,35). Gene therapy procedures are changing rapidly, and this is noted in the regulatory documents with the suggestion that there be constant contact between the regulators and developers of any product proposed for gene therapy (36,37). Many of the molecular biology issues are

covered by the regulatory documents on recombinant products (5,7–10).

Vectors

Vectors and gene product are characterized in the same manner as for cells producing rDNA proteins (8–10). For smaller viral vectors this would include sequencing the entire viral sequence; for larger viral vectors this may not be feasible. The gene product should be completely characterized, however, because the vectors used for in vivo gene transfer will not be expressed in the same cell, and many of the post-translational changes may differ. As with other cell-derived products, the development and characterization of master cell banks will be required (5,7). In addition, there must be establishment of a vector bank, that is, the final product for most of the proposed vectors is the result of the interaction of a cell and the vector. The vector bank should be tested for the presence of adventitious agents. Stability of the host cell–vector system must be established.

The most promising vectors under study for treatment of the hemophilias include replication of defective retroviral and adenoviral vectors, adeno-associated virus, and nonviral vectors. Retrovirus vectors have promise for in vitro exchange with a variety of cells, including fibroblasts, endothelium, and myoblasts. The nonviral vectors would seem to raise fewer issues albeit much lower in vivo expression. The strategy issues to be resolved include choice of vector and in vitro or in vivo gene transfer.

Production Issues

There are different regulatory issues and pathways for the licensure of vectors used in vitro and in vivo. Products for in vivo transduction will be produced purified and packaged in the final form for administration to the patients in a facility meeting all the standards for biologics production (38). For vectors where the transduction will be done in vitro, the producing institution will be responsible for the production of the vector and will have to meet all GMPs; however, the actual product administered will be done after expansion of the patient's transduced cells and will be a service for an individual patient similar to treatment with blood.

Current standards for the production of pharmaceutical products encourage the development of single product production facilities. However, the target group for gene therapy in the hemophilias is small, and only a few production runs may be needed if the hope of long-lasting or permanent effectiveness is achieved. This would lead to a multiuse production area with stringent validation of the facility after changeover from one product to another to prevent cross-contamination.

There are no specific guidelines as yet for the degree of purification for a gene therapy product; rather, it is left to the producer to show adequate purification and justify the adequateness of the procedure. The consistency of the purification method must be demonstrated. Published purification methods for viral vectors involves density gradient ultracentrifugation; it is unlikely that this will be implemented as a production method. Published purification procedures give little information on the degree of purity achieved. Because the desired outcome with gene therapy would be one-time or infrequent exposure to the agent, the presence of small amounts of cell-derived or process-derived protein is less of an issue than for recombinant proteins administered frequently.

The fidelity of reproduction of the vector is an issue because an error in replication could yield a toxic gene product. Where possible the correct nucleotide sequence should be verified.

Lot release testing suggested by the FDA includes (35):

1. Test for RNA or DNA content
2. Test for homogeneity of size and structure of RNA/DNA
3. Test for contamination with RNA or host DNA
4. Test for noninfectious viruses
5. Tests for toxic materials
6. Identity tests by restriction enzyme mapping
7. Sterility including testing for adventitious agents and replication competent viruses
8. Potency

The only test that would be unique for the coagulation products would be the potency test. It will be the responsibility of the manufacturer to design and validate a potency test during the development phase. It is likely that the standard coagulation tests of conditioned media produced by cells treated with a given amount of vector would be adequate for the validation of a lot. Whole animal methods could be developed but could be anticipated to be less quantitative and too time consuming to be useful for lot release.

Possible Adverse Effects

Viral mutagenesis and tumorigenicity are of concern even in replication of incompetent vectors because the genes could be inserted into chromosomes and disrupt growth regulation genes. Identification of the insertion mode and testing for replication competent viruses is recommended for the master cell bank, the working cell bank, and lot release testing on the adenoviral vector products.

The FDA and EEC guidelines (36,37) specifically are concerned with the ability of replication-incompetent viral genomes to form recombinants with cellular viral components. Strategies suggested include separation of the structural and enzymatic genes, removal of complementary viral sequences from the packaging cell lines, and prevention of infection of the packaging cell line with wild-type viruses that could lead to the formation of replication-competent recombinant viruses.

Specific issues with viral vectors are the pathology induced by the viral components. As many pathologic viral sequences as possible should be removed. The induction of virus antibodies preventing retreatment is an issue that can be modulated by minimization of the viral structure and perhaps regulation of the administered dose.

Animal Studies

Preclinical animal testing will be of importance not just for determining efficacy but for supporting safety and study-

ing tissue distribution. All viral vectors will have to be examined by appropriate routes of administration and dose effects in the most sensitive animal models. Species will depend on sensitivity to particular viral toxicity. There is no requirement or indication that regulatory authorities will require primate studies.

Studies are mandated of the vector distribution in animals. A specific issue will be whether there is uptake into germ-line cells. Although this can be studied in animals by study of ovarian and testicular cells, it is not clear whether this will be a reliable indicator of human cell or organ location. Although there is public and political concern about germ-line modification, the FDA does not have any specific prohibition of a product with germ-line modification. The organ of synthesis for both factor VIII and IX is the liver, and this is the site of uptake of the majority of adenoviral vectors administered intravenously.

The cell or organ location is of importance to the final gene product; post-translation modification will be affected by the cell type transduced. There is at this time, however, no set experimental protocol for such studies, and the manufacturer will be expected to use all appropriate tools and develop specific tools as necessary for a given product.

Retroviral transduction and implantation of the ex vivo transduced cells should give an identifiable localized expression of the gene product. There may, however, be specific cell–cell interactions.

Presence of functional clotting should ensure adequate in vivo function. Human factor IX and factor VIII have been maintained in the therapeutic range for up to a half a year in various animal models with various vectors (35). The most promising are the ex vivo retrovirus-transformed fibroblasts or muscle cells and the adenovirus vectors with in vivo transfer.

As noted above, measurable shifts in clearance occur with recombinant products. Although the steady-state levels are measurable in both animals and humans, it will be difficult to measure clearance in patients, although this might be a regulatory issue.

A particular problem with factor VIII may be the need to have the gene product secrete directly into the plasma compartment. In the absence of the stabilizing effect of von Willebrand's factor, it is possible that factor VIII will be proteolyzed to an inactive form before reaching the plasma compartment where it functions.

Persistence of the transgene depends on absence of antibody to the vector-infected cell or the gene product. The antibody response to the vector and the toxicity of the vector can be limited by introducing lower quantities of the viral DNA. More recent experience indicates that immunity to the adenovirus may be circumvented.

Adverse effects of the vector and the implanted gene may be detected before patient exposure only in animal models, and at this time there have been no unexpected events. As expected, there are reports of liver damage with the use of adenoviral vectors, which can be minimized by dose adjustment (39). Adverse effects of the vector and the implanted gene may not be detected in an animal model.

Clinical Trial Issues

The initial decision to attempt gene therapy in hemophilia will be difficult and controversial and will be reviewed by regulatory groups, patient advocates, lawyers, and politicians. In the United States, any clinical use directly or indirectly involved with government funding formerly was approved by the NIH's Recombinant DNA Advisory Committee (RAC) and subsequently by the Office of Recombinant DNA Activities after approval of the the originating institution's internal review board. The results of these reviews were placed in the public domain, and it was understood that all adverse effects would be open information. In addition, the RAC and the FDA often had parallel review processes for other projects. In 1996, the role of the RAC was made more limited, and the FDA assumed all regulatory review and approval responsibility.

Certain principles in regard to informed consent have already been established by the RAC and the FDA. These include many of the usual points such as the trial goal and known risks. In addition, there must be clear discussion of the methodology and the duration of the therapy and of follow-up. One specific point mentioned is the provision for autopsy. The patient must have a commitment and understanding of the trial and the importance of each patient's value to the study as well as the advantages he or she may achieve.

The hemophilia patient who would benefit most would be the patient without access to adequate means of alternative treatment. Because 80% of the hemophilic population in the world does not have access to treatment, one can argue that this is a potential target group. Outside the United States, there has been an effort to treat human factor IX deficiency by gene therapy in China (40).

A second group to be considered for early treatment with gene therapy might be patients with inhibitors (antibodies) to factor VIII. These patients are already undergoing frequent exposure to induce tolerance. Treatment with gene therapy would simply be a modification of tolerance therapy, which could have great benefit for patients with inhibitors. Assessment of therapy would be more difficult if there is no measurable plasma factor VIII. Alternatives could be evidence of a decrease in the inhibitor titer, similar to that seen in classic tolerance regimens, or evidence that there is long-term production of gene product or mRNA. A third group would be those without adequate venous access for administration of current products.

The group that might benefit the most would be the very young or newborn hemophiliacs who, in principle, would not have any of the damage caused by lack of factor VIII or factor IX. It can be anticipated that gene therapy treatment of this group will not be approved in the absence of substantial long-term safety data from older patients.

There will be a regulatory requirement for long-term follow-up: how long and for what are not discussed in the *Points-to-Consider*. The obvious points are evidence of functional expression, unusual immunologic outcomes, viral shedding if a viral vector is used, and cell trafficking. Semen samples will be collected to assess germ-line effects. The patients currently undergoing gene therapy have been informed of the importance of autopsy information.

It is easy to state that the clinical trial must be designed to assess clinical safety and efficacy; it is far more difficult at this time to anticipate with precision the scientific or regulatory requirements, such as the number of patients

to be included, the duration before licensure by a regulatory agency, and the statistical design and the number of patients in a monogenic disease where the assessment of bleeding reduction can be done with retrospective data.

The first major effectiveness issue will be the expression level. Clinical efficacy initially will be a plasma level deemed sufficient to minimize hemorrhage, 5 to 10% of the normal plasma level. The need for other replacement therapy will be the ultimate test of efficacy. The initial trials will begin with a dosage based on the animal data or some fraction thereof, and there may well need to be adjustment. Expression should be seen in days to weeks; duration of expression must be monitored frequently. There are no historical data to support toxicity of supranormal levels of factor VIII in the patient with hemophilia. However, nonhemophiliac subjects with supranormal levels of factor VIII have increased incidence of venous thrombosis (40,41).

The immunologic issues, inhibitor formation, and immune complex disease will have to be studied. The worst possible outcome would be the induction of chronic immune complex disease, which has never been reported in the 30-year history of the treatment of hemophilia despite studies indicating increased levels of immune complexes. The other side of the issue is that for years large daily doses of factor VIII have been used to develop immune tolerance in patients with factor VIII inhibitors, and no chronic immune complex disease has been reported. If recipients of gene therapy do develop immune complex disease, it would be observed within days to months and perhaps should be part of the normal treatment protocol as well as play a role in the investigative use of gene therapy in hemophilia.

Although hemophilia patients with inhibitors are routinely treated with factor VIII, these are patients with no intrinsic synthesis of the factor VIII. A concern raised by the FDA is that the cells producing the gene product could be destroyed by an immune cytolysis. This potential outcome would argue against the use of known inhibitor patients for the initial trials. Patients without hemophilia A who develop antibodies to factor VIII have not been reported to have any organ damage despite the presence of cells producing normal amounts of factor VIII.

The development of inhibitors usually occurs with the first 20 treatments, but the current European guidelines request prolonged inhibitor follow-up, and it is likely that long-term follow-up will be requested in gene therapy. Follow-up for the development of antibodies to the vector may also be suggested. If therapy is by a modified viral vector, there may be limited antibody formation. Methods should be developed to detect antibodies to this vector in the presence of wild-type virus infection.

Surveillance of the impact of the vector will be mandatory. The older hemophilia patient population is infected with a variety of chronic viral diseases, and many already have liver disease. Adenoviral vectors targeted to the liver will at least temporarily induce some inflammatory reaction. Other unknown adverse effects resulting from the vector may occur with this novel and dramatic therapeutic approach.

Gene therapy science is moving rapidly, and as old problems are solved, new areas of concern arise. The developers of any such product must be aware of all the issues. Preclinical and clinical studies must be rigorous; at this time, the regulatory guidelines are appropriately rigorous. The best updated source of information on the regulatory issues may be found at the FDA's Center for Biologics Evaluation and Review (CBER) internet site (*www.CBER.FDA.gov*).

TRANSGENIC ANIMAL PRODUCTS

The term *transgenic animal* is used for animals with either somatic cell DNA alterations or germ-line alterations. The potential value of transgenic animals for the production of therapeutic proteins would involve the production of animal strains with germ-line DNA alterations, a regulatory base exists both in the United States and Europe (42,43). The use of transgenic animals to produce proteins secreted into milk is an attractive alternative to the large investment in a sophisticated plant to produce therapeutic proteins. At this time, there is little evidence of commercial interest in this approach. Many coagulation proteins have been produced in the milk of transgenic animals (44).

The regulation in the United States of transgenic products would involve not only the FDA drug or biologics regulatory groups but in addition the veterinary medicine and food regulators of the FDA in addition to the Department of Agriculture. The FDA regulatory issues are spelled out in Reference 42.

Many of the issues of blood-derived products would be involved, including screening for viruses, assessing the health of the donor, etc. The possibility of adapting unknown animal viruses to a human host will not be a trivial issue and will have to be addressed by the producer.

The actual gene and the regulatory element characterization are covered by the *Points-to-Consider* for cell-derived products (8). The regulatory elements are of particular interest, and extensive characterization of the enhancers, promoters, suppressors, and presence of dominant control regions should be reported.

There must be a full record of all the procedures in obtaining the ova, methods for in vitro fertilization, and introduction of the altered DNA. After homologous recombination, it is important to show loss of function of the wild-type gene.

Both the history of the donors of the gametes and the recipient (foster parent) animals need to be documented in detail. Intensive veterinary evaluation of health with particular emphasis on diseases related to the species and the breed will be required.

The founder animal must be studied for the presence of the introduced gene and for the expression of the gene product in addition to the absence of functional product from the inactivated wild-type gene. The characterization of the gene expression should be evaluated as a function of age, season, etc. The acceptable range for the gene product should be established to allow for all the systematic variations. Anatomical localization and function of the transgene should be evaluated, and that the transgene is at the desired location must be verified. There is particular concern about the post-translational modifications in a species and at a tissue site different from the normal human prod-

uct, which could lead to functional and immulogic problems.

The maintenance of a given transgenic line is presumed desirable, and methods similar to the master working cell bank for cell-derived products is suggested.

In addition to study of the transgene in the founder, there must be extensive study of the transgene stability in the succeeding generations until stability of the gene and the gene expression are demonstrated. In the founder, there may well be multiple gene copies inserted which over succeeding generations will be rearranged or deleted. It is desirable to demonstrate at the founder stage that there is integration at a single chromosomal site. If this is not possible, studies of the chromosomal integration must be done in the later generations. In addition, the RNA transcript should be examined for quality and quantity and identification of the tissues or cells producing the transcript.

The animals must be kept isolated from sources of known and unknown infections during their productive life. Animals would only be acceptable in the absence of disease and with a consistent level of production of the gene product.

ACKNOWLEDGMENTS

The author gratefully acknowledges the helpful guidance and education on gene therapy and international regulatory insights and issues by Terry Snape, Sheila Connelly, Milt Mozen, and Arthur Thompson.

BIBLIOGRAPHY

1. Y. Nemerson, in W.J. Williams, E. Beutler, A.J. Erslev, M.A. Lichtman eds., *Hematology*, McGraw-Hill, New York, 1990, pp. 1295–1304.
2. D.L. Aronson, *Transfusion* **30**, 748–758 (1990).
3. B.L. Evatt, *Thromb. Haemostasis* **74**, 36–39 (1995).
4. K. Peerlick, J. Arnout, J.G. Gilles, J.M. Saint-Remy, J. Vermylen, *Thromb. Haemostasis* **69**, 115–118 (1993).
5. *Points-to-Consider in the Characterization of Cell Lines Used to Produce Biologics*. Center of Biologics Evaluation and Review, U.S. Food and Drug Administration, Bethesda, Md., 1993.
6. H. Hopps, J.C. Petricciani, *Abnormal Cells, New Products and Risks*. Tissue Culture Association, Gaithersburg, Md., 1985.
7. *Acceptability of Cell Substrates for Production of Biologicals*. World Health Organization Tech. Report Ser. **747**, 1 (1987).
8. *Points-to-Consider for Products Derived from Mammalian Cells*. Center of Biologics Evaluation and Review, U.S. Food and Drug Administration, Bethesda, Md., 1987.
9. Committee for Proprietary Medicinal Products, *Note for Guidance: Production and Quality Control of Medicinal Products Derived by Recombinant DNA Technology*, 1995 (*http://www.eurdra.org/emea.html*).
10. *Quality of Biotechnological Products: Analysis of the Expression Construct in Cells Used for Production of R-DNA Derived Protein Products*, International Conference on Harmonization (*http://ifpma.org/ich/*).
11. J. Petricciani, F. Horaud, *Biologicals* **23**, 233–238 (1995).
12. European Agency for the Evaluation of Medicinal Products, Human Medicines Evaluation Unit, *Note for Guidance to Assess Efficacy and Safety of Human Plasma Derived Factor VIII: C and Factor IX: C Products in Clinical Trials in Haemophiliacs before and after Authorisation*, 1996 (*http://www.eurdra.org/emea.html*).
13. S. Harrison, B. Clancy, S. Brodeur, P. Oakes, D. Miller, D. Drapeau, M. Hamilton, T. Charlebois, M. Leonard, M. McCarthy, R. Zollner, S.R. Adamson, *Thromb. Haemostasis* **73**, 1235 (1995).
14. Committee for Proprietary Medicinal Products, *Note for Guidance on Viral Validation Studies: The Design, Contribution, and Interpretation of Studies Validating the Inactivation and Removal of Viruses*, 1996, (*http://www.eurdra.org/emea.html*).
15. A.B. Heath and T.W. Barrowcliffe, *Thromb. Haemostasis* **68**, 155–159 (1992).
16. G. Kemball-Cook, J.E. Tubbs, N.J. Dawson, T.W. Barrowcliffe, *Br. J. Haematol.* **84**, 273–278 (1993).
17. K. Fijnvandaat, E. Berntorp, J.W. ten Cate, H. Johnson, M. Peters, G. Savidge, L. Tengborn, J. Spira, C. Stahl, *Thromb. Haemostasis* **77**, 298–302 (1997).
18. C.A. Lee, T. Barrowcliffe, G. Bray, E. Gomperts, A. Hubbard, G. Kemball-Cook, P. Lilley, D. Owens, I. Von tilberg, J. Pasi, *Thromb. Haemostasis* **76**, 950–956 (1996).
19. H.P. Schwarz, P.L. Turecek, L. Pichler, A. Mitterer, W. Mundt, E. Dorner, J. Roussi, L. Drouet, *Thromb. Haemostasis* **78**, 571–576 (1997).
20. A.R. Giles, S. Tinlin, H. Hoogedorn, M. Fournel, P. Ng, N. Pancham, *Blood* **72**, 335– (1988).
21. I.R. MacGregor, L.E. McLauglin, M.C. MacGregor, C.V. Prowse, D.S. Pepper, *Vox Sanguinis* **69**, 319–327 (1995).
22. I.R. MacGregor, J.M. Ferguson, L.E. McLaughlin, T. Burnouf, C.V. Prowse, *Thromb. Haemostasis* **66**, 609–613 (1991).
23. J.C. Keith, T.J. Ferranti, B. Misra, T. Frederick, B. Rup, K. McCarthy, R. Faulkner, L. Bush, R.G. Schaub, *Thromb. Haemostasis* **73**, 101–105 (1994).
24. Investigational New Drug Applications, *Code of Federal Regulations*, Title 21, Part 312.
25. *Code of Federal Regulations*, Title 21, Part 601.25.
26. *Code of Federal Regulations*, Title 21, Part 314.126.
27. M. Morfini, M. Lee, A. Messori, *Thromb. Haemostasis* **66**, 384–386 (1991).
28. P.M. Mannucci, M. Colombo, *Thromb. Haemostasis* **61**, 532–534 (1989).
29. J.M. Lusher, S. Arkin, C.F. Abilgaard, R.S. Schwartz, *N. Engl. J. Med* **328**, 453–459 (1993).
30. G. White, J. Lusher, A. Shapiro, K. Tubridy, S. Courter, *Blood* **88**, 1296 (1996).
31. B. Brody, *Ethical Issues in Drug Testing, Approval, and Pricing*, Oxford Univ. Press, Oxford, U.K., 1995.
32. D. Beebe, G. Murano, in A.S. Lubiniecki, S.A. Vargo eds., *Tissue Plasminogen Activator: Regulatory Issues in Regulatory Practice for Biopharmaceutical Production*, Wiley-Liss, New York, 1994, pp. 281–294.
33. J.J. Emeis, C.M. van den Hoogen, D. Jense, *Thromb. Haemostasis* **54**, 661–664 (1985).
34. Thompson A., *Thromb. Haemostasis* **74**, 45–51 (1995).
35. S. Connelly, M. Kaleko, *Thromb. Haemostasis* **78**, 31–41 (1997).
36. *Points to Consider in Human Somatic Cell and Gene Therapy*, Center of Biologics Evaluation and Review, U.S. Food and Drug Administration, Bethesda, Md., 1996.

37. Committee for Proprietary Medicinal Products, *Note for Guidance: Gene Therapy Products—Quality Aspects in the Production of Vectors and Genetically Modified Somatic Cells* (III/5863/93 Final) (*http://www.eurdra.org/emea.html*).
38. Current Good Manufacturing Practices for Finished Pharmaceuticals, *Code of Federal Regulations*, Title 21, Part 211.
39. S. Connelly, J.M. Gardner, R.M. Lyons, A. McClelland, M. Kaleko, *Blood* **87**, 4671–4677 (1996).
40. D.-R. Lu, J.-M. Zhou, B. Zheng, X.F. Qui, J.L. Xue, J.M. Wang, P.L. Meng, F.L. Han, B.H. Ming, X.P. Wang, *Sci. China [B Series]* **36**, 1342–1351 (1993).
41. T. Koster, A.D. Blann, E. Briet, J.P. Vandenbroucke, and F.R. Rosendaal, *Lancet* **345**, 152–155 (1995).
42. *Points-to-Consider in the Manufacture and Testing of Therapeutic Products for Human Use Derived from Transgenic Animals*, Center of Biologics Evaluation and Review, U.S. Food and Drug Administration, Bethesda, Md.
43. Committee for Proprietary Medicinal Products, *Note for Guidance: Use of Transgenic Animals in the Manufacture of Biological Medicinal Products for Human Use*, 3612/93 (*http://www.eurdra.org/emea.html*).
44. W.N. Drohan, *Thromb. Haemostasis* **78**, 543–547 (1997).

COFACTOR REGENERATION, NICOTINAMIDE COENZYMES

JOERG PETERS
Bayer AG
Wuppertal, Germany

KEY WORDS

Coenzyme recycling
Enzyme-coupled approach
Immobilization
Inhibition
Reaction kinetics
Reaction mechanisms
Reactor systems
Regeneration methods
Substrate-coupled approach
Total turnover number

OUTLINE

INTRODUCTION

Today, there are about 3500 known enzymes (1), of which about 70% are cofactor dependent and about 15% are commercially available. Cofactors such as adenosine triphosphate (ATP), nicotinamide adenine dinucleotide (NADPH), acyl coenzyme A, *S*-adenosylmethionine (SAM), sugar nucleotides, and 3′-phosphoadenosine-5′-phosphosulfate (PAPS) are involved in stoichiometric amounts in cofactor-dependent enzymatic reactions (2). The very high cost of these cofactors prohibits their use as stoichiometric reagents. For economic reasons, regeneration of the cofactors from their reaction products is an essential requirement in making use of cofactor-dependent enzymes in preparative synthesis of organic compounds.

Using coenzyme recycling, the expensive cofactor is needed only in catalytic amount, so there is a drastic reduction in cost of the respective reactions. Additionally, coenzyme regeneration can accomplish three major objectives. First, it can influence the position of equilibrium of a reaction. A thermodynamically unfavorable reaction can be driven by coupling with a favorable cofactor-regeneration reaction. Second, regeneration affords only catalytic amounts of coenzyme, preventing the accumulation of cofactor by-product, which may inhibit the reaction of interest. Third, coenzyme regeneration can simplify the downstream processing of the reaction mixture.

The following chapter deals with methods for the regeneration of nicotinamide coenzymes, because this type of cofactor is the most significant. Hitherto, more than 650 oxidoreductases are known (3), of which about 90 are commercially available. For the majority of these redox enzymes, about 80% and 10% require nicotinamide adenine dinucleotide (NADH) or its respective phosphate (NADPH) respectively.

REGENERATION OF NICOTINAMIDE COENZYMES

Introduction

Nicotinamide adenine dinucleotide (NAD) and the analogous 2′-phosphate (NADP) are involved in redox reactions catalyzed by alcohol dehydrogenases. Two electrons and a proton (hydride) are transferred from the reduced coenzyme NADPH to the carbonyl compound in a stoichiometric reaction. Using coenzyme recycling, the expensive cofactor is needed only in catalytic amounts, leading to a drastic reduction in the cost of dehydrogenase-catalyzed reactions. This chapter discusses methods for regeneration of nicotinamide coenzymes.

Cost of Coenzymes. Nicotinamide adenine dinucleotides (NADH, NADPH) are expensive compounds. In Table 1 the actual prices per mole are listed and compared with the costs in the year 1976. If these coenzymes were used in stoichiometric amounts, enzymatic reductions with dehydrogenases would be prohibitively expensive. Therefore, the application of dehydrogenases on the industrial scale requires effective methods for coenzyme regeneration. The economic barrier to large-scale reactions posed by cofactor costs has been recognized for many years. A range of publications address this topic (2,5,6).

The turnover number (TN) is defined as the number of moles of product formed per mole of cofactor or enzyme per unit time. This number is indicative of the productivity of

Table 1. Prices of Nicotinamide Coenzymes

Coenzymes	Price in dollars per mole (depending on purity, 1976)[a]	Price in dollars per mole (mean catalog prices, 1998)
NAD^+	1.722 to 49.550	710
NADH	6.030 to 52.766	3
$NADP^+$	20.667 to 168.604	25
NADPH	177.032 to 535.197	215

[a]from Ref. 4.

the process. The total turnover number (TTN) is defined as the total number of moles of product formed per mole of coenzyme or enzyme during the course of the entire reaction. Hence, the TTN reflects the loss of coenzyme and enzyme throughout the process, and thus it is a useful estimate of the operational cost of the coenzyme or the enzyme. In order to overcome economic limitations, total turnover numbers of greater than 100 are desirable. For a batch synthesis, the cost of the coenzyme or enzyme is calculated on the basis of its initial cost. In contrast, for a continuous process, the cost may be calculated on the basis of the activity loss over the entire process time. When estimating the cost of a given reaction per unit time, the rate of the reaction must also be considered. The regeneration cost is defined as the cost of the components (enzymes, reagents, and coenzymes) required to regenerate one mole of coenzyme per unit time (7).

The highest TTN can be achieved as follows:

- The substrate concentration is as high as possible, determined by the substrate solubility, product inhibition, and enzyme stability.
- The concentration of coenzyme and enzyme to achieve acceptable reaction rates is as low as possible.
- The coenzymes and enzymes are highly stable.

Thus, both the initial costs of the cofactor and the enzyme, and the efficiency of their regeneration and utilization, determine their contribution to the cost of the product.

Origin of Coenzymes. Nicotinamide adenine dinucleotide (NAD) is presently isolated from yeast (8), and its 2′-phosphate NADP is normally synthesized by enzymatic phosphorylation of NAD using NAD-kinase (EC 2.7.1.23) and ATP (9,10). NADPH can be prepared from NADP by three different methods: chemical (11), enzymatical (12,13), or microbial reduction (14). Additionally, a combined chemical and enzymatic route for the synthesis of NADP has been developed that may be useful for the preparation of NADP analogs (15–17).

Requirements of an Economical Coenzyme Regeneration System. Chenault and Whitesides (4) compiled the following criteria for an ideal coenzyme regeneration system:

- Compatibility with the synthetic reaction
- Commercial availability of inexpensive and stable enzymes with high specific activity
- Simple and cheap auxiliary substrates; the corresponding products should not interfere with the isolation of the product of interest. They should not influence the stability of enzymes and coenzymes.
- High turnover numbers (TN) and total turnover numbers (TTN).
- The equilibrium constants of the coupled coenzyme regeneration system should be as high as possible.
- The analysis of the yield should not be effected by the substrates or products of the coenzyme regeneration.

As a rule of thumb, 10^3–10^4 cycles are sufficient for laboratory-scale redox reactions, whereas at least 10^5–10^6 cycles are desirable for industrial-scale synthesis. The exact turnover number required depends on the initial cost of the coenzyme and the value of the product of interest. A high turnover number requires a high selectivity for the formation of enzymatically active coenzyme. 1,4-dihydro-NAD^+, or $NADP^+$, is the only enzymatically active form; 1,2- and 1,6-dihydro species are not active. For instance, if 50% of the original coenzyme activity is to remain after 1,000 cycles, the regeneration reaction must be 99.3% regioselective. In the case of industrial applications of dehydrogenases, 10^6 turnovers are desired, resulting in a requirement for the regioselectivity of 99.99993%. Chemical, electrochemical, photochemical, or enzymatic methods (see later) show different selectivities in the formation of 1,4-dihydro $NADP^+$. Electro- and photochemical methods especially suffer from poor regioselectivities. Thus, the requirement for very high regioselectivity necessitates enzymatic catalysis, particularly in the recycling of NADH and NADPH (2,6,18).

Stability of Nicotinamide Coenzymes. *Nicotinamide coenzymes contain three labile bonds that may be the target of general or specific acid-or base-catalyzed hydrolysis*:

- the N-glycosidic bond between nicotinamide and ribose
- The phosphodiester bond between the two riboses
- The bond between ribose and phosphate (NADP)

Oppenheimer (19) and Chenault and Whitesides (6) compiled data on the influence of different buffers on the stability, and the possible products from the hydrolysis of NADPH. Under basic conditions, the reduced nicotinamide coenzymes are stable; under acidic conditions, they are unstable. In contrast, the oxidized coenzymes are stable under acidic conditions, and labile under basic conditions (19,20). As a compromise, NAD and NADH are used at pH 7–7.5, whereas NADP and NADPH are used at pH 8–8.5. The mechanisms of the decomposition of NADPH were studied in great detail (20–22). The N-glycosidic bond between nicotinamide and ribose is the most labile bond of the molecule. Base-catalyzed hydrolysis of this bond results in the removal of the nicotinamide ring and, thus, the coenzymatic activity is destroyed. On the other hand, the general acid-catalyzed protonation of the nicotinamide ring at position C_5 is followed by a rapid rearrangement to a cyclic ether product. Additionally, the 2′-phosphate group

of NADPH catalyzes this decomposition intramolecularly. The half lifes for NADH and NADPH in 0.1 M phosphate (pH 7, 25 °C) are 27 h and 13 h, respectively (20). Organic buffers such as imidazole, Tris, Hepes, and triethanolamine, partially stabilize reduced coenzymes, resulting in a half life that is three times as long. The major pathway for the decomposition of $NADP^+$ is nucleophilic addition at C_4 of the nicotinamide ring, yielding a 1,4-dihydropyridine structure (23–25). In Tris-buffered solution at pH 7–8, $NADP^+$ is stable over at least 14 days (Peters, unpublished data). There is a strong dependence of the stability of the oxidized coenzymes on the concentration of the buffer (Peters, unpublished data). At low molarity (50 mM), the half-life in Tris buffer is lower compared with the half-life in buffer of high molarity (500 mM). Poly(ethylene)-glycol-coupled $NADP^+$ is very stable at pH 8 in Tris buffer of any molarity. Half-lifes of more than 240 days have been reported (Peters, unpublished data).

Principles for the Regeneration of Nicotinamide Coenzymes

Nonenzymatic Regeneration. Nonenzymatic methods for coenzyme regeneration can be divided in chemical, electrochemical, and photochemical methods. Chemical reduction of $NADP^+$ using a reducing agent like sodium dithionite ($Na_2S_2O_4$) suffers from very low total turnover number of less than 100 (26). Moreover, this agent can deactivate enzymes, presumably by modification of thiol groups in the protein (27). The major disadvantages of most electrochemical and photochemical regeneration methods (28–31) are low regioselectivity, leading to coenzyme inactivation (as already discussed) occurrence of side reactions, and low total turnover numbers of less than 1,000 (2,6). However, Simon and coworkers (32) stated that electro- and photochemical methods should not be underestimated. Schummer and coworkers (33) reported the synthesis of polyfunctional (*R*)-2-hydroxycarbonic acids on the preparative scale with resting cells of *Proteus vulgaris* by using electrochemical regeneration. Several groups (34–37) reported on the electrochemical regeneration of NADPH by using a bipyridine rhodium-(I) complex as an electron-transfer agent and formate as an electron donor. The bipyridine-rhodium complex can be coupled to poly(ethylene) glycol and thereby retained in an enzyme-membrane reactor. Dicosimo and coworkers (38) described the electrochemical regeneration of NADPH using the dye methyl viologen and flavoenzymes. However, enzyme denaturation on the surface of the electrode or enzyme deactivation by redox agents may occur.

In summary, enzymatic methods meet most of the requirements for an ideal coenzyme regeneration system as described earlier. The difficulties encountered with enzymatic coenzyme regeneration are mainly competitive or mixed inhibition by the substrate or product (39,40). In the following sections, the two general concepts for coenzyme regeneration using enzymes are presented (18). Focus will then turn to the most attractive methods for the recycling of reduced and oxidized nicotinamide coenzymes.

Substrate-Coupled Regeneration. In the substrate-coupled approach, both the main reaction and the regeneration of the coenzyme are catalyzed by a single enzyme (Fig. 1). The auxiliary substrate (S'_{red}) of the coenzyme regeneration reaction serves as the hydride donor. In order to shift the equilibrium of the reaction to the desired position, the hydride donor, usually a low-cost alcohol, is applied in large excess. However, yield of the system is limited by equilibrium of the coupled system. For the reduction of acetophenone using 2-propanol as auxiliary substrate, the equilibrium was reached at 60% yield (41).

In the case of the reduction of 2-acetyl-naphthaline using 2-propanol as hydride donor, the equilibrium could be shifted by the use of cyclodextrins, and thereby the total yield was increased to 93% of 1-(2-naphthyl)-ethanol (42). In the continuous synthesis, yields of more than 80% were obtained, and total turnover numbers for the coenzyme were in the range of 10^4. In this application the native coenzyme was not recovered from the product solution but was directly used in substoichiometric amounts. The total turnover number for the enzyme was greater than 10^6, and the productivity was 2.5 times as high as the comparable system with enzyme-coupled cofactor regeneration (formate dehydrogenase system). This demonstrates the potential of substrate-coupled systems for coenzyme regeneration.

Due to the lack of a well-established method for NADPH regeneration, the substrate-coupling method is mainly used for this purpose. However, there are some drawbacks of this method:

- The auxiliary substrate has to be present in a large excess, which may lead to complications in product recovery, or which may cause enzyme deactivation or inhibition.
- Enzyme deactivation may also occur as the result of the accumulation of highly reactive carbonyl species like acetaldehyde or cyclohexenone.

A special enzyme reactor employing a membrane that is only permeable to gases (comparable to Gore-Tex) may avoid most of these drawbacks. This type of reactor was successfully applied to improve cofactor recycling via the yeast alcohol dehydrogenase–ethanol system in the stereoselective reduction catalyzed by lactate dehydrogenase (43). The reaction scheme is shown in Figure 2.

Enzyme-Coupled Regeneration. In the case of the enzyme-coupled approach to NAD regeneration (Fig. 3), the reduction of the main substrate and the coenzyme re-

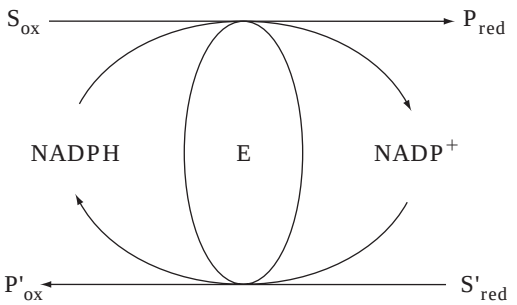


Figure 1.

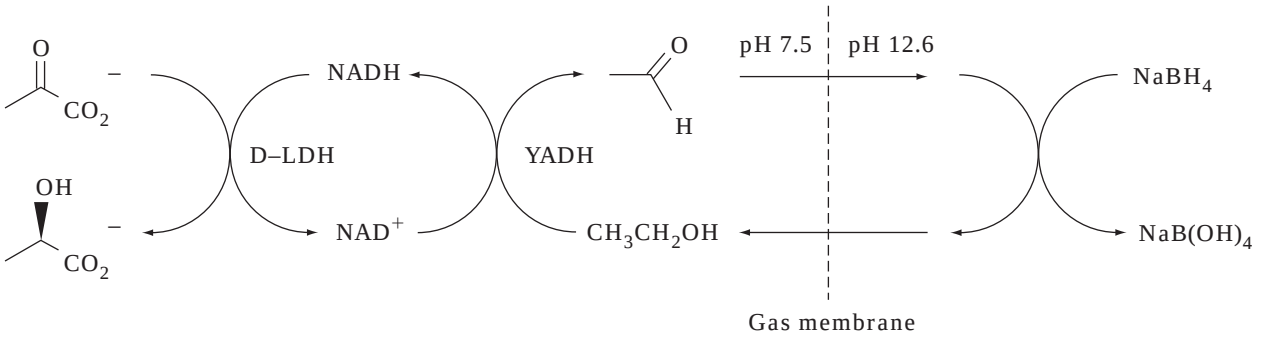


Figure 2.

generation reaction are catalyzed by two different enzymes (44). The enzymes employed should have sufficiently different substrate specificities to avoid interferences. The enzyme-coupled cofactor recycling has the following major advantages:

- The position of the equilibrium of a redox reaction can be influenced. Thermodynamically unfavorable reactions can be driven by coupling with a favorable cofactor regeneration reaction.
- High turnover numbers can be achieved: >600,000 (45).
- The activity ratio has to be optimized for each pair of enzymes to ensure neither lack nor excess of reduction equivalents (46).

However, there are also some disadvantages to this type of cofactor-recycling method:

- Cross inhibition between the main substrate and the coenzyme regeneration system may occur (46).
- A second enzyme must be used, generating additional costs.
- The characteristics of the second enzyme, such as pH and temperature optimum, must be compatible with the main reaction.

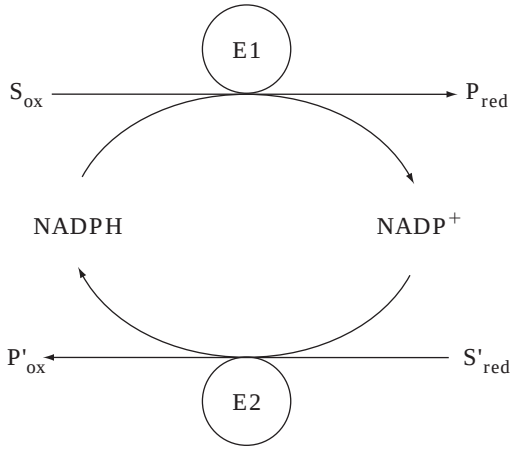


Figure 3.

Survey of Coenzyme Regeneration Methods and Their Assessment

Different methods for the regeneration of NADPH have been discussed in great detail by several authors (2,6,7,18,20,32,47–52). In this chapter, only the most useful methods are discussed and assessed with respect to their applicability in batch and continuous synthesis. Furthermore, recently developed new methods for nicotinamide coenzyme recycling are highlighted.

Regeneration of NADPH. In Table 2, the major disadvantages of a selection of available cofactor-recycling systems are listed, and each method is assessed with respect to its applicability in batch or continuous processes.

For the regeneration of native and poly(ethylene) glycol-bound NADH (PEG-NADH), the formate dehydrogenase (FDH) from *Candida boidinii* (53) has been widely and successfully applied (45,54–57). This enzyme catalyzes the oxidation of inexpensive formate (HCOOH) to carbon dioxide (CO_2). The reverse reaction is hardly detectable (53). Therefore, the FDH reaction can shift the equilibrium of unfavorable reactions to the product side. Both the auxiliary substrate and the coproduct are innocuous to enzymes and are easily removed from the reaction. Therefore, product isolation is easy. FDH has a broad pH optimum for activity, so it can be easily implemented in coupled enzymatic synthesis. The major disadvantages are the low specific activity of 4 U/mg and product inhibition by NADH. However, these drawbacks can be circumvented or minimized: the former by using immobilized or membrane-retained systems (54,58); the latter by adjusting the optimal ratio of enzymatic activities for the main reaction and the coenzyme-recycling reaction (46). Overall, the formate–FDH system is the most economical method for regeneration of NADH for batch and continuous processes. The total turnover numbers (TTN) range typically from 10^3 to 10^5 (45,56,58). Unfortunately, the *C. boidinii* formate dehydrogenase does not accept $NADP^+$.

Although several NADP-dependent formate dehydrogenases have been discovered (59–63), only the recently described mutant formate dehydrogenase from *Pseudomonas* sp. 101 (64,65), using $NADP^+$ as coenzyme, was applied in batch and continuous synthesis. Acetophenone was reduced by the NADPH-dependent alcohol dehydrogenase from *Lactobacillus* sp. (EC 1.1.1.99) with this enzyme for the regeneration of NADPH, at a space–time yield

Table 2. Methods for In Situ Regeneration of NADPH

Methods	Major disadvantages	Assessment
Glucose/glucose dehydrogenase	Not applicable in the enzyme-membrane reactor K_m value for PEG-NADP$^+$ 1,000 times that of native NADP$^+$ Complication of product isolation caused by gluconate High amounts of K^+ or Na^+ necessary to stabilize enzyme	Useful only for lab-scale reactions in batch reactors with native NADH or NADPH, but not for PEG-NAD(P)H
Glucose-6-P/glucose-6P dehydrogenase	High costs for glucose-6-phosphate (~5.500 $/mol) Coenzyme deactivation by 6-phosphogluconate No efficient usage of the coenzyme Complication of product recovery caused by 6-phosphogluconate	Useful only for lab-scale reactions, regeneration of NADH and NADPH
Glucose-6-sulfate/glucose-6-P dehydrogenase	Complication of product recovery caused by 6-sulfogluconate Substrate is accepted only by glucose-6-phosphate dehydrogenase from yeast	Circumvents some limitations of the former method; useful for preparative applications in continuous reactors; regeneration of NADH and NADPH
Ethanol/alcohol and aldehyde dehydrogenase	Relatively complicate and unstable multienzyme system with low TTN Oxygen sensitivity of alcohol and aldehyde dehydrogenases Low reactivity of the aldehyde dehydrogenase with NADP$^+$ Only activated carbonyl substrates are reduced at good yields, which may cause enzyme deactivation	Useful only for lab-scale reactions; regeneration of NADH and NADPH
Formate/formate: NAD$^+$ dehydrogenase	Low specific activity (3 U/mg) Deactivated by autooxidation Product inhibition caused by PEG-NADH	Simple and stable system for regeneration of PEG-NADH in batch or continuous mode of operation up to the process scale; very good compatibility with other enzymes
Formate/formate: NADP$^+$ dehydrogenase	Enzyme not commercially available yet	Simple and stable system for regeneration of NADPH in batch or continuous mode of operation; very good compatibility with other enzymes
Sec. alcohol/*Th. brockii* alcohol dehydrogenase	Deactivation by autooxidation possible Relatively expensive enzyme (58 cents/unit)	Simple and stable system for regeneration of PEG-NADPH in batch or continuous mode of operation; very good compatibility with other enzymes

of 8 g/(L day) and a conversion rate of 90–95% (66,67). The new FDH is quite stable, showing no loss in activity over one year at 4 °C (67). Unfortunately, this enzyme is not yet commercially available. A further disadvantage is the relatively high K_m value of the mutant FDH for NADP (320 μM). Additional work is required to engineer FDH mutants with improved kinetic properties. In conclusion, the new formate–NADP dehydrogenase (EC 1.2.1.2) has a large potential for use as a regeneration system for NADPH.

The secondary alcohol dehydrogenases from *Thermoanaerobium brockii* (TBADH) accept a broad range of secondary alcohols and NADP$^+$ as coenzyme (68). The enzyme is commercially available and contains a very low NADP-2′-phosphatase activity (Peters, unpublished data) (NADP(H) is hydrolyzed at the 2′-ribose phosphate bond at a rate of 0.1% per day in the continuous synthesis of NADPH (69).) This enzyme has been applied in the reduction of numerous carbonyl compounds using the substrate-coupled approach for coenzyme recycling (70–75). Since this enzyme tolerates up to 30% isopropanol, the substrate-coupled regeneration of NADPH is highly favorable. The TBADH-enzyme family is also an excellent system for the regeneration of native and poly(ethylene) glycol-bound NADPH (Peters, unpublished data; 69). In a model system, glutamate dehydrogenase (GluDH) was employed for the synthesis of glutamate from α-ketoglutarate. A complete set of kinetic data was recorded and the effectivity of the regeneration of PEG-NADPH by the TBADH was studied in a batch reactor. The optimal activity ratio of GluDH/(GluDH + TBADH) was 0.6. At this ratio, 95% of the total coenzyme is reduced. The conversion of α-ketoglutarate is affected negatively only at a TBADH portion of less than 20%. For comparison, in the leucine dehydrogenase–FDH system studied by Wichmann and coworkers (55), only 10% PEG-NADH is present at the optimal activity ratio of 1:1. The reason for this different behavior is the product inhibition of FDH by PEG-NADH. In contrast, TBADH shows only slight product inhibitions by 2-butanone and PEG-NADPH. The loss of coenzyme in the enzyme-membrane reactor is a function of the absolute coenzyme concentration (76) and, therefore should be as low as possible. The K_m value of the TBADH for native and

PEG-bound NADPH is 10 μM. Hence, only 0.2–0.3 mM NADPH is usually sufficient to ensure a maximum usage of the catalytic activities in the reactor.

Overall, the TBADH is recommended for the recycling of NADPH for the following reasons. First, the enzyme is commercially available and is stable in a polypropylene enzyme-membrane reactor. Second, the specific activity is almost 10 times as high as that of formate dehydrogenase. Favorably, the K_m values for native and PEG-bound NADPH are low (10 μM), and the product inhibitions by 2-butanone and PEG-NADPH are also low. Third, the TBADH is compatible with high amounts of organic solvent (e.g., 30% isopropanol) and is stable at elevated temperatures.

Regeneration of NADP$^+$. Oxidized nicotinamide coenzymes are used for the synthesis of carbonyl compounds from the corresponding racemic hydroxy compounds. For their regeneration, both enzymatic and nonenzymatic methods have been reported (for review see Ref. 7). Enzymatic methods seem to be preferred because of their compatibility with biological systems and their simplicity. The regeneration of oxidized nicotinamide coenzymes is somewhat problematic due to thermodynamics that are often unfavorable. Additionally, product inhibition by the reduced coenzyme may cause problems (see later).

In Table 3, the major disadvantages of various available cofactor-recycling systems are listed, and each method is assessed with respect to its applicability in batch or continuous processes. The alcohol dehydrogenase from *Pseudomonas* sp. which is not yet commercially available, has potential for the recycling of NAD$^+$ (77,78). A range of simple ketones can be used as hydrogen acceptors, which are reduced at high velocity (up to 270 U/mg). However, the kinetic limitations of this enzyme must be explored in more detail.

For recycling of NADP$^+$, the secondary alcohol dehydrogenase from *Thermoanaerobium brockii* (TBADH) is problematic due to the low product-inhibition constant K_{ip} for NADP$^+$ of 50 μM (Peters, unpublished data; 69). The K_m value of NADPH is 10 μM. Therefore, the oxidation of alcohols producing the reduced coenzyme is the favorable reaction.

Acetaldehyde and the commercially available alcohol dehydrogenase from baker's yeast (YADH) have also been used to regenerate NAD$^+$ from NADH (79). The total turnover numbers were 10^3 to 10^4. However, deactivation of the enzyme by the highly reactive acetaldehyde and the self-condensation of acetaldehyde outweigh the advantages of the inexpensive enzyme and the volatility of the reagents involved.

The most widely applied method for regeneration of oxidized nicotinamide coenzymes involves glutamate dehydrogenase (GluDH), which catalyzes the thermodynamically favorable reductive amination of inexpensive α-ketoglutarate to L-glutamate (40,80). Both the substrate and the product are innocuous to enzymes. The commercially available and inexpensive GluDH accepts NADH, NADPH, and their respective poly(ethylene) glycol-bound coenzymes (Peters, unpublished data) and has a reasonably high specific activity of 40 U/mg protein. The products of the reductive amination of α-ketoglutarate, PEG-NADP$^+$ and L-glutamate, do not strongly inhibit the reaction (Peters, unpublished data). The ratio of the product inhibition constant to the Michaelis–Menten constant (K_{ip}/K_m) will determine the efficacy of the reaction (40). At a K_{ip}/K_m ratio greater than 1, the reaction may proceed to acceptable conversions. If this ratio is less than 1, the reaction can never proceed efficiently. The K_{ip}/K_m ratio for the substrate and product of the reductive amination of α-ketoglutarate to L-glutamate is 10.5. For the PEG-NADPH and PEG-NADP$^+$, the K_{ip}/K_m ratio is about 30 (Peters, unpublished data). Hence, product inhibition by L-glutamate and PEG-NADP$^+$ does not play a significant role in the efficacy of the regeneration reaction. The disadvantage of this regeneration method is that L-glutamate may complicate product isolation from the reaction mixture. This must be addressed case by case.

Pyruvate and commercially available and inexpensive lactate dehydrogenase (LDH) have been used for recycling of NAD$^+$ (6,81). The LDH is stable and has a high specific activity of about 1,000 U/mg protein. However, the redox potential is less favorable and, in contrast to GluDH, LDH does not accept NADPH. A further disadvantage is that pyruvate tends to polymerize in solution and reacts with NAD$^+$ in a process catalyzed by LDH. Despite these draw-

Table 3. Methods for In Situ Regeneration of NADP$^+$

Methods	Major disadvantages	Assessment
Alcohol/*Pseudomonas* sp. alcohol dehydrogenase	Enzyme not commercially available	
Sec. alcohol/*Thermoanaerobium brockii* dehydrogenase	Strong product inhibition by PEG-NADP$^+$; K_{ip} (NADP), 50 μM; K_m (NADPH), 10 μM	Not useful due to the severe product inhibition by PEG-NADP$^+$
Acetaldehyde/yeast alcohol dehydrogenase	Self-condensation of acetaldehyde Enzyme deactivation by highly reactive acetaldehyde	Not useful for the regeneration of NAD$^+$ due to side reactions and deactivation by acetaldehyde
α-Ketoglutarate glutamate dehydrogenase	Glutamate may complicate product isolation	Simple and stable system for regeneration of PEG-NADPH, good compatibility with other enzymes
Pyruvate/lactate dehydrogenase	Pyruvate tends to polymerize in solution and reacts with NAD$^+$ in a process catalyzed by lactate dehydrogenase	Simple and stable system for regeneration of NAD$^+$; side reactions may happen

backs, the pyruvate–LDH system has been applied in enzymatic oxidation reactions of 10–100 mmol of material.

Immobilized Coenzymes and Reactor Concepts

Enzyme-catalyzed reactions are often accompanied by the following disadvantages:

- Enzymes may not be sufficiently stable under the reaction conditions employed. Some lose their catalytic activity due to autooxidation, others due to self-digestion and/or denaturation by the solvent, and others due to mechanical shear forces.
- Isolated enzymes are water soluble and must be used repeatedly due to the economics of the processes for which they are used (82). They must be separated from the reaction mixture without destroying their catalytic activity.
- The productivity, expressed as the space–time yield, is often low due to the limited tolerance of enzymes to high concentrations of substrate(s) and product(s).

Some of these problems may be overcome by "immobilization" of the enzyme. However, a special problem is encountered with NADPH-dependent dehydrogenases because the coenzyme has to associate and dissociate freely to and from the active site(s) of the enzyme(s). Therefore, coimmobilization of the dehydrogenase and the cofactor is necessary to make the system practical (83). Either dehydrogenase, coenzyme, or both may decompose, leading to the replacement of the immobilized system.

If the K_m value for the coenzyme is in the micromolar range and the substrate-coupled approach for coenzyme regeneration is employed (as discussed earlier), a column reactor can be used with a small amount of free coenzyme dissolved in the mobile phase (72). This situation is particularly efficient if the reaction mechanism is an ordered bi-bi- or Theorell–Chance mechanism where the coenzyme is bound first to the active site, followed by the substrate. These mechanisms were found for a broad range of alcohol dehydrogenases.

The immobilization techniques most widely used for coenzyme-dependent dehydrogenases are shown in Figure 4. Immobilization techniques for coenzymes may be divided into coupling and entrapment methods.

Coupling Methods. For dehydrogenases, coupling methods can be further divided into methods for coupling the coenzyme on cross-linked enzymes or methods for coupling the enzyme and coenzyme on a carrier.

Coupling of the Coenzyme to Cross-Linked Enzymes. Enzymes can be attached to each other by covalent bonds, termed *cross-linking*. Bifunctional reagents such as glutardialdehyde, dimethyl adipimidate and dimethyl suberimidate can be used. The coenzyme can be bound to cross-linked enzymes using a spacer long enough so that the coenzyme has access to the active sites of both enzymes, thereby transferring the hydride. These requirements are very difficult to meet in practice. Lortie and coworkers (84) reported on the coimmobilization of horse liver alcohol dehydrogenase (HLADH), albumin, and NAD^+ using glutaraldehyde as the cross-linking agent. Long-chain aldehydes were prepared using this immobilized HLADH/NAD^+ system in a packed-bed reactor (glass column). The conversions were performed in organic solvent (hexane) saturated with buffer. Because proteins and coenzymes are insoluble in the water immiscible hexane, they remained in the aqueous phase. The long-chain alcohols (substrate) diffused into the aqueous phase, underwent enzymatic conversion, and the product (aldehyde) finally diffused back into the organic phase. This system was limited by the mass transfer between the organic phase and the water-containing porous particles. Lortie and coworkers (84) generalized this result to all types of fixed-bed enzyme reactors in biphasic systems. Importantly, the regeneration of the coenzyme was not a rate-limiting step.

Coupling of the Coenzyme on a Carrier. Another option for immobilizing coenzymes is to bind both the dehydrogenase(s) and the coenzyme to an insoluble carrier such as cellulose, dextrans, starch, chitin, agarose, or synthetic copolymers (85,86). The activation of the hydroxyl groups of polysaccharides is usually done by reaction with cyanogen bromide, leading to the formation of reactive imidocarbonates. Synthetic polymers can be activated using partial hydrolysis of the acetate groups followed by activation with epichlorhydrin. A range of preactivated polysaccharides and synthetic polymers (Biosynth, Eupergit, etc.) may be obtained from commercial sources. The enzyme is coupled via its amino groups. Also, in this case, a spacer long enough to allow the coenzyme to have access to the active sites of the enzyme(s) must be used. Horse liver alcohol dehydrogenase (HLADH) and an NADH analogue, N^6-[(6-aminohexyl)carbamoyl-methyl]-NADH, have been successfully coimmobilized to Sepharose 4B (83). It was shown that the coenzyme could be regenerated at a cycling rate of 3,400 cycles/h using the substrate-coupled approach. Both thermal and storage stability of the HLADH were increased substantially. It was not necessary to add soluble coenzyme to the substrate solution to maintain activity.

Entrapment Methods. Polymer-linked coenzyme derivatives are often introduced in enzyme reactors with coenzyme-dependent processes, because they can be retained upon ultrafiltration or dialysis and are recycled (Fig. 5) (35,39,45,47,56,57,76,87–93). Small substrate and product molecules can freely pass through the membrane (e.g., polyamide, polyethersulfone), but large enzymes and the polymer-linked coenzyme cannot. Synthetic membranes of defined pore size covering the range between 500 and 300,000 Da are commercially available at a reasonable cost. Additionally, a variety of synthetic membrane shapes, such as membrane discs or hollow fibers, are on the market. A simple form of a membrane reactor that does not require any special equipment is also shown in Figure 4. In this simple technique, termed *membrane-enclosed enzymatic catalysis* (MEEC), the enzymes and polymer-bound coenzyme are entrapped in a dialysis bag, which is mounted on a gently rotating magnetic stirring bar (81).

One modification of the enzyme membrane-reactor concept makes use of charged ultrafiltration membranes in order to retain the native coenzyme in the reactor (94–97). The application of charged ultrafiltration membranes in membrane reactors is limited to the production of non-charged products. In contrast, nanofiltration membranes

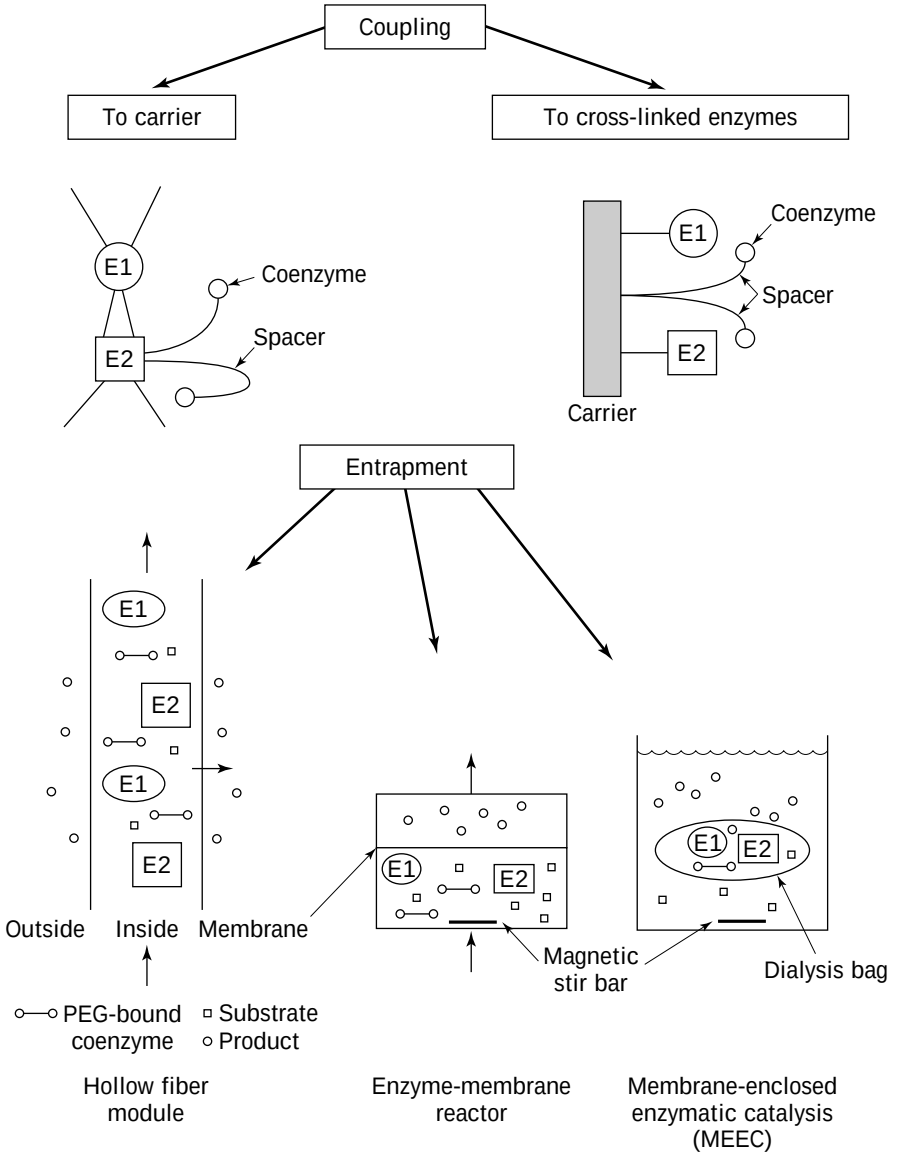


Figure 4.

can be used to retain the coenzyme in the reactor, whereas the reaction product(s) may freely pass the membrane (98). The separation of coenzyme and reaction product is due to the differences in the molecular weight, without a negative net charge of the membrane. For a recent overview of the applications of dehydrogenases using the enzyme membrane reactor technology, the reader is referred to Ref. 99.

Many reports on the preparation of polymer-linked coenzyme derivatives are available (100–109). The dextran-bound coenzymes have been demonstrated to the serious problem of protein leakage, due to the instability of the isourea linkage between coenzyme derivatives and dextran (110). However, more stable dextran derivatives can be synthesized by using bromohydroxypropyl derivatives of dextrans (111). These derivatives are used mostly in enzyme electrodes or as stationary phases in affinity chromatography (112).

The polyethyleneimine and polylysine derivatives of coenzymes (109) have the disadvantage of having electrostatic interactions between the multicharged polymers and enzymes, thus interfering with their effectiveness as active coenzymes. The relative reaction rates compared to native NAD^+ were 2–25% (polyethyleneimine) and 2–60% (polylysine), depending on the enzyme tested. The synthesis of polyacryl copolymers has the advantage that parameters like solubility, charge and the functional group can be varied just by exchange of the monomer (106). However, polyacryl-bound NAD^+ showed only 30% of the V_{max} value of the native coenzyme (113).

Another approach was established by Bückmann and coworkers using poly(ethylene) glycol derivatives with molecular weights of 3,000 to 20,000 kDa and modified coenzymes. The yield of the coupling of *N*(1)-(2-aminoethyl)-NAD^+ with the carboxy-PEG derivative was 90% (103). The alkylation of NADPH with 1,4-bis-(2,3-epoxypropoxy)-butane was investigated by Stein (114). This method allows the direct alkylation of NADPH in a one-step reaction, whereas all other described alkylation methods are mul-

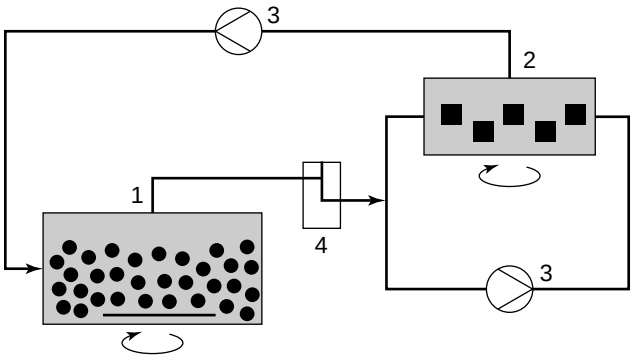


Figure 5.

tistep reactions starting from the oxidized coenzymes. Moreover, the coenzymatic activities of the NADPH derivatives were almost identical to the native coenzymes. The NADPH derivatives carry a hydrophilic spacer and a terminal epoxy group, which can be coupled easily to PEG-NH_2, resulting in a stable, covalent bond between PEG and the coenzyme derivative. For most of the enzymes tested, the coenzymatic activity of PEG-NADH is comparable to the native coenzyme (88,103,115). However, there are also exceptions (105). The PEG–coenzyme derivatives have been successfully applied in the abovementioned enzyme membrane reactor (35,39,45,56, 57,76,87–93). PEG-NADH, as well as other PEG-derivatives, are commercially available from spin-offs of German research centers in Braunschweig and Jülich (ASA Spezialenzyme GmbH; Jülich Enzyme Products).

Modern reactor concepts for the application of dehydrogenases and native coenzymes in organic solvents have been worked out by several groups (89,116–118). These concepts circumvent the immobilization of the coenzyme by using hydrophobic microfiltration membranes and two-phase systems, an emulsion membrane reactor or a differential circulation reactor with continuous extraction, respectively.

Conclusions

Enzymatic reductions catalyzed by NADP-dependent dehydrogenases require stoichiometric amounts of nicotinamide coenzymes, which are very expensive compounds. The application of dehydrogenases on a preparative scale therefore requires effective methods for the regeneration of these coenzymes.

Today, coenzyme regeneration of both NADH and NADPH are very straightforward and the cofactor is no longer the dominant cost in preparative reductions. Nicotinamide coenzymes are labile compounds, but if they are used under optimal conditions, the half-life can be extended to more than 14 days. Coupling of NADPH to poly(ethylene) glycols stabilizes the coenzymes, by which means the half-life can be extended to more than 240 days. Enzymatic methods for coenzyme regeneration are preferable to nonenzymatic methods because they meet most of the requirements for an ideal recycling system. Two different principles can be distinguished: the substrate-coupled and the enzyme-coupled approaches. The enzyme-coupled method is advantageous because even unfavorable reactions can be driven to the product side by coupling with a favorable cofactor-regeneration reaction.

The most economical method for PEG-NADH recycling is the formate–formate dehydrogenase system, which has been applied up to the commercial-scale process.

For the regeneration of (PEG-NADPH, the alcohol–alcohol dehydrogenase from the *Thermoanaerobium brockii* system or the formate–formate:NADP dehydrogenase system are recommended.

The α-ketoglutarate–glutamate dehydrogenase system catalyzes the thermodynamically favorable reductive amination of α-ketoglutarate. Both NAD^+ and $NADP^+$, as well as PEG-$NAD(P)^+$, can be regenerated effectively.

The use of the enzyme membrane reactor technology in conjunction with poly(ethylene) glycol-linked coenzymes made possible the scale-up of enzymatic processes to preparative scale (91,119,120). Among others, this technology was successfully applied for the synthesis of different L-amino acids and a range of chiral hydroxy compounds of high value.

BIBLIOGRAPHY

1. H. Bielka, B.L. Horecker, W.B. Jakoby, P. Karlson, B. Keil, C. Lièbecq, B. Lindberg, and E.C. Webb eds., *Enzyme Nomenclature*, Academic Press Inc., New York, 1992.
2. H.K. Chenault, E.S. Simon, and G.M. Whitesides, *Biotechnol. Genet Eng Rev.* **6**, 221–267 (1988).
3. H.G. Davies, R.H. Green, D.R. Kelly, and S.M. Roberts, *Biotransformations in Preparative Organic Chemistry: the Use of Isolated Enzymes and Whole Cell Systems in Synthesis*, Academic, London, 1989.
4. J.B. Jones and J.F. Beck, in J.B. Jones, C.J. Sih, and D. Perlman eds., *Applications of Biochemical Systems in Organic Chemistry*, Wiley, New York, 1976, pp. 107–402.
5. I. Willner and D. Mandler, *Enzyme Microb. Technol.* **11**, 467 (1989).
6. H.K. Chenault and G.M. Whitesides, *Appl. Biochem. Biotechnol.* **14**, 147–197 (1987).
7. C.H. Wong and G.M. Whitesides, in J.E. Baldwin, and P.D. Magnus eds., *Tetrahedron Organic Chemistry Series: Enzymes in Synthetic Organic Chemistry*, Pergamon, Oxford, 1994, pp. 1–370.
8. T. Sakai, T. Uchida, and I. Chibata, *Agric. Biol. Chem.* **37**, 1049–1956 (1973).
9. K. Murata, J. Kato, and I. Chibata, *Biotechnol. Bioeng.* **21**, 887 (1979).
10. T. Hayashi, Y. Tanaka, and K. Kawashima, *Biotechnol. Bioeng.* **21**, 1019 (1979).
11. A.L. Lehninger, *Methods Enzymol.* **3**, 885–887 (1957).
12. G.W. Rafter and S.P. Colowick, *Methods Enzymol.* **3**, 887–889 (1957).
13. S. Suye and S. Yokoyama, *Enzyme Microb. Technol.* **7**, 418 (1985).
14. S.Y. Eguchi, N. Nishio, and S. Nagai, *Agric. Biol. Chem.* **47**, 2941–2943 (1983).
15. A. Traub, E. Kaufmann, and Y. Teitz, *Anal. Biochem.* **28**, 469–476 (1969).
16. A.J. Kornberg, *J. Biol. Chem.* **182**, 779–793 (1950).

17. D.R. Walt, M.A. Findeis, V.M. Rios-Mercadillo, J. Auge, and G.M. Whitesides, *J. Am. Chem. Soc.* **106**, 234–239 (1984).
18. W. Hummel and M.R. Kula, *Eur. J. Biochem.* **184**, 1–13 (1989).
19. N.J. Oppenheimer, in J. Everse, B. Anderson, and K.S. You eds., *The Pyridine Nucleotide Coenzymes*, Academic, New York, 1982, pp. 51–89.
20. C.H. Wong and G.M. Whitesides, *J. Am. Chem. Soc.* **103**, 4890–4899 (1981).
21. S.P. Colowick, N.O. Kaplan, and M.M. Ciotti, *J Biol Chem.* **191**, 447–459 (1951).
22. N.J. Oppenheimer and N.O. Kaplan, *Biochemistry* **13**, 4675–4685 (1974).
23. S.L. Johnson and K.W. Smith, *Biochemistry* **15**, 553–559 (1976).
24. J.F. Biellman, C. Lapinte, E. Haid, and G. Wiemann, *Biochemistry* **18**, 1212–1217 (1979).
25. J. Everse, E.C. Zoll, L. Kahan, and N.O. Kaplan, *Bioorg. Chem.* **1**, 207–233 (1971).
26. J.B. Jones, D.W. Sneddon, W. Higgins, and A.J. Lewis, *J. Chem. Soc. Chem. Commun.* 856–857 (1972).
27. R. Ramio and E.M. Lilius, *Enzymologia* **40**, 360–368 (1971).
28. D. Mandler and I. Willner, *J. Chem. Soc. Perkin Trans.* **II**, 805–811 (1986).
29. J.B. Jones and K.E. Taylor, *Can. J. Chem.* **54**, 2969–2973 (1976).
30. M.D. Legoy, V. Laretta-Garde, J.M. LeMoullec, F. Ergan, and D. Thomas, *Biochimie (Paris)* **62**, 341–345 (1980).
31. M. Julliard, J. Le Petit, and P. Ritz, *Biotechnol. Bioeng.* **28**, 1774 (1986).
32. H. Simon, H. Günther, J. Bader, and S. Neumann, *Ciba Found. Symp.* **111**, 97–111 (1985).
33. A. Schummer, H. Yu, and H. Simon, *Tetrahedron* **47**, 9019–9034, (1991).
34. R. Ruppert, S. Herrmann, and E. Steckhan, *Tetrahedron Lett.* **28**, 6583 (1987).
35. E. Steckhan, S. Herrmann, R. Ruppert, J. Thömes, and C. Wandrey, *Angew. Chem., Int. Ed. Engl.* **102**, 445–447 (1990).
36. D. Westerhausen, S. Herrmann, W. Hummel, and E. Steckhan, *Angew. Chem.* **104**, 1496–1498 (1992).
37. R. Wienkamp and E. Steckhan, *Angew Chem. Int. Ed. Engl.* **21**, 782–783 (1982).
38. R. Dicosimo, C.H. Wong, L. Daniels, and G.M. Whitesides, *J. Org. Chem.* **46**, 4622–4623 (1981).
39. M.R. Kula, R. Wichmann, U. Oden, and C. Wandrey, *Biochimie (Paris)* **62**, 523–536 (1980).
40. L.G. Lee and G.M. Whitesides, *J. Org. Chem.* **51**, 25–36 (1986).
41. J. Peters, T. Minuth, and M.R. Kula, *Tetrahedron Asymm.* **4**, 1683–1692 (1993).
42. T. Zelinski, Ph.D. Thesis, University of Düsseldorf, Germany, 1995.
43. P. van Eikeren, D.J. Brose, D.C. Muchmore, and J.B. West, *Ann. N.Y. Acad. Sci.* **613**, 796–801 (1990).
44. H.R. Levy, F.A. Loewus, and B. Vennesland, *J. Am. Chem. Soc.* **79**, 2949–2953 (1957).
45. E. Schmidt, W. Hummel, H. Schütte, and M.R. Kula, in H. Chmiel, W.P. Hammes, and J.E. Bailey eds., *Biochemical Engineering—A Challenge for Interdisciplinary Cooperation*, Gustav Fischer, Stuttgart, 1987, pp. 487–490.
46. J. Peters, Ph.D. Thesis, University of Düsseldorf, Germany, 1993.
47. S.S. Wang and C.K. King, in T.K. Ghose, A. Fiechter, and N. Blakebrough eds., *Advances in Biochemical Engineering*, Springer, Berlin, 1979, pp. 119–146.
48. J. Bader, H. Günther, S. Nagata, H.J. Schuetz, M.L. Link, and H. Simon, *J. Biotechnol.* **1**, 95–109 (1984).
49. C.H. Wong, D.G. Drueckhammer, and H.M. Sweers, *J. Am. Chem. Soc.* **107**, 4028–4031 (1985).
50. L.G. Lee and G.M. Whitesides, *J. Am. Chem. Soc.* **107**, 6999 (1985).
51. L. Poppe and L. Novak, *Selective Biocatalysis. A Synthetic Approach*, VCH Verlagsgesellschaft mbH, Weinheim, 1992.
52. K. Faber, *Biotransformations in Organic Chemistry*, Springer, 1995.
53. H. Schütte, J. Flossdorf, H. Sahm, and M.R. Kula, *Eur. J. Biochem.* **62**, 151–160 (1976).
54. Z. Shaked and G.M. Whitesides, *J. Am. Chem. Soc.* **102**, 7105–7107 (1980).
55. R. Wichmann, C. Wandrey, A.F. Bückmann, and M.R. Kula, *J. Biotechnol.* **23**, 2789–2802 (1981).
56. E. Schmidt, D. Vasic-Racki, and C. Wandrey, *Appl. Microbiol. Biotechnol.* **26**, 42–48 (1987).
57. D. Vasic-Racki, M. Jonas, C. Wandrey, W. Hummel, and M.R. Kula, *Appl. Microbiol. Biotechnol.* **31**, 215–222 (1989).
58. W. Hummel, H. Schütte, E. Schmidt, C. Wandrey, and M.R. Kula, *Appl. Microbiol. Biotechnol.* **26**, 409–416 (1987).
59. I. Yamamoto, T. Saiki, S.M. Liu, and L.G. Ljundahl, *J. Biol. Chem.* **258**, 1826–1832 (1983).
60. J.R. Andreesen, *J. Bacteriol.* **120**, 6–14 (1974).
61. L.R. Krumholz, R.L. Crawford, M.E. Hemling, and M.P. Bryant, *J. Bacteriol.* **169**, 1886–1890 (1987).
62. L.G. Ljundahl and J.R. Andreesen, *FEBS Letts.* **54**, 279–282 (1975).
63. S.W. Tannenbaum, *Biochim. Biophys. Acta* **21**, 335–342 (1956).
64. V. Tishkov, A. Galskin, L.B. Kulakova, and A. Egorov, *Enzyme Engineering XII*, Deauville, France, 1993.
65. V.I. Tishkov, A.G. Galkin, G.N. Marchenko, Y.D. Tsygankov, and H.M. Ergorov, *Biotechnol. Appl. Biochem.* **18**, 201–207 (1993).
66. K. Seelbach, Diplomarbeit, University of Bonn, Germany, 1994.
67. K. Seelbach, B. Riebel, W. Hummel, M.R. Kula, V.I. Tishkov, A.M. Egorov, C. Wandrey, and U. Kragl, *Tetrahedron Lett.* **37**, 1377–1380 (1996).
68. R.J. Lamed and J.G. Zeikus, *Biochem. J.* **195**, 183–190 (1981).
69. J. Peters and M.R. Kula, *Biotechnol. Appl. Biochem.* **13**, 363–370 (1991).
70. E. Keinan, K.K. Seth, R.J. Lamed, R. Ghirlando, and S.P. Singh, *Biocatalysis* **3**, 57–71 (1990).
71. E. Keinan, S.C. Sinha, and A. Sinha-Bagchi, *J. Org. Chem.* **57**, 3631–3636 (1992).
72. E. Keinan, K.K. Seth, and R.J. Lamed, *J. Am. Chem. Soc.* **108**, 3474–3480 (1986).
73. E. Keinan, E.K. Hafeli, K.K. Seth, and R.J. Lamed, *J. Am. Chem. Soc.* **108**, 162–169 (1986).
74. E. Keinan, K.K. Seth, and R.J. Lamed, *Ann. N.Y. Acad. Sci.* **501**, 130–150 (1987).
75. R.J. Lamed, E. Keinan, and J.G. Zeikus, *Enzyme Microb. Technol.* **3**, 144–148 (1981).

76. M.R. Kula, and C. Wandrey, in S.P. Colowick and N.O. Kaplan eds., *Methods in Enzymology Part C*, Academic, New York, 1987, pp. 9–21.
77. C.W. Bradshaw, H. Fu, G.J. Shen, and C.H. Wong, *J. Org. Chem.* **57**, 1526–1532 (1992).
78. G.J. Shen, Y.F. Wang, and C.H. Wong, *J. Chem. Soc. Chem. Commun.* **9**, 677–679 (1990).
79. G.L. Lemiere, J.A. Lepoivre, and F.C. Alderweireldt, *Tetrahedron Lett.* **26**, 4527–4528 (1985).
80. G. Carrea, R. Bovara, P. Cremonesi, and R. Lodi, *Biotechnol. Bioeng.* **26**, 560–563 (1984).
81. M.D. Bednarski, H.K. Chenault, E.S. Simon, and G.M. Whitesides, *J. Am. Chem. Soc.* **109**, 1283–1285 (1987).
82. C.J. Suckling and K.E. Suckling, *Chem. Soc. Rev.* **3**, 387–406 (1974).
83. S. Gestrelius, M.O. Mansson, and K. Mosbach, *Eur. J. Biochem.* **57**, 529–535 (1975).
84. R. Lortie, I. Villaume, M.D. Legoy, and D. Thomas, *Biotechnol. Bioeng.* **33**, 229–232 (1989).
85. P.W. Carr, and L.D. Bowers, *Immobilized Enzymes in Analytical and Clinical Chemistry*, Wiley, New York, 1980.
86. C.Y. Lee, and A.F. Chen, in J. Everse, B. Anderson, and K.S. You eds., *The pyridine Nucleotide Coenzymes*, Academic, New York, 1982, pp. 189–224.
87. K. Hosono, S. Kajiwara, Y. Yamazaki, and H. Maeda, *J. Biotechnol.* **14**, 149–156 (1990).
88. N. Katayama, I. Urabe, and H. Okada, *Eur. J. Biochem.* **132**, 403–409 (1983).
89. S. Kise and M. Hayashida, *J. Biotechnol.* **14**, 221 (1990).
90. German Patent 3937892 C2 (November 15, 1989), U. Kragl, J. Peters, C. Wandrey, and M.R. Kula (to the Regents of the Forschungszentrum Jülich GmbH); Swiss Patent CH679676A5 (November 5, 1990), U. Kragl, J. Peters, C. Wandrey, and M.R. Kula (to the Regents of the Forschungszentrum Jülich GmbH).
91. W. Leuchtenberger, M. Karrenbauer, and U. Plöcker, *Ann. N.Y. Acad. Sci.* **434**, 78–86 (1984).
92. C. Wandrey and R. Wichmann, in A.I. Laskin ed., *Enzymes and Immobilized Cells in Biotechnology*, Benjamin/Cummings, Menlo Park, Calif., 1985, pp. 177–208.
93. C. Wandrey, R. Wichmann, M.R. Kula, and A.F. Bückmann, *Die Umschau*. **3**, 88–91 (1984).
94. M. Ikemi and Y. Ishimatsu, *Biotechnol. Bioeng.* **36**, 155–165 (1990).
95. M. Ikemi, N. Koizumi, and Y. Ishimatsu, *Biotechnol. Bioeng.* **36**, 149–154 (1990).
96. B. Nidetzky, D. Haltrich, and K.D. Kulbe, *Chemtech*. 31–36 (1996).
97. T.R. Röthig, K.D. Kulbe, F. Bückmann, and G. Carrea, *Biotechnol Lett.* **12**, 353–356 (1990).
98. K. Seelbach and U. Kragl, *Enzyme Microb. Technol.* **20**, 389–392 (1997).
99. J. Peters, in H.J. Rehm, and G. Reed eds., *Biotechnology*, Wiley-VCH, Weinheim, 1998, pp. 391–473.
100. A.F. Bückmann and G. Carrea, *Adv. Biochem. Eng. Biotechnol.* **39**, 98–152 (1989).
101. A.F. Bückmann, *Biocatalysis* **1**, 173–186 (1987).
102. A.F. Bückmann, *Heterocycles* **27**, 1623–1628 (1988).
103. A.F. Bückmann, M.R. Kula, R. Wichmann, and C. Wandrey, *J. Appl. Biochem.* **3**, 301–315 (1981).
104. A.F. Bückmann, M. Morr, and G. Johansson, *Makromol. Chem.* **182**, 1379–1384 (1981).
105. A.F. Bückmann, M. Morr, and M.R. Kula, *Biotechnol. Appl. Biochem.* **9**, 258–268 (1987).
106. C.W. Fuller and H.J. Bright, *J. Biol. Chem.* **252**, 6631–6639 (1977).
107. P.O. Larsson and K. Mosbach, *FEBS Letts.* **46**, 119–122 (1974).
108. K. Okuda, I. Urabe, and H. Okada, *Eur. J. Biochem.* **151**, 33–38 (1985).
109. P. Zapelli, A. Rossodivita, and L. Re, *Eur. J. Biochem.* **54**, 475–482 (1975).
110. R. Axen, J. Porath, and S. Ernback, *Nature (London)* **214**, 1302–1304 (1967).
111. C.Y. Lee, A. Leigh-Brown, C.H. Langley, and D. Charles, *J. Solid-Phase Biochem.* **2**, 213–224 (1978).
112. C.R. Lowe and K. Mosbach, *Eur. J. Biochem.* **49**, 511–520 (1974).
113. C.W. Fuller, J.R. Rubin, and H.J. Bright, *Eur. J. Biochem.* **103**, 421–430 (1980).
114. A. Stein, Ph.D. Thesis, University of Münster, Germany, 1990.
115. N. Katayama, K. Hayakawa, I. Urabe, and H. Okada, *Enzyme Microb. Technol.* **6**, 538–542 (1984).
116. W. Kruse, W. Hummel, and U. Kragl, *Recl. Trav. Chim Pays-Bas.* **115**, 239–243 (1996).
117. A. Liese, M. Karutz, J. Kamphuis, C. Wandrey, and U. Kragl, *Biotechnol. Bioeng.* **51**, 544–550 (1996).
118. A. Liese, T. Zelinski, M.R. Kula, H. Kierkels, M. Karutz, B. Schulze, U. Kragl, and C. Wandrey, *J. Mol. Cat. B.* **4**, 91–99 (1998).
119. A. Bommarius, M. Schwarm, K. Stingl, M. Kottenhahn, K. Huthmacher, and K.H. Drauz, *Tetrahedron Asymm.* **6**, 2851–2888 (1995).
120. U. Kragl, D. Vasic-Racki, and C. Wandrey, *Bioprocess Eng.* **14**, 291–297 (1996).

CONDUCTIVITY

ALESSANDRO D'APRANO
University of Rome, La Sapienza
Rome, Italy

KEY WORDS

Carrier mechanism
Channel mechanism
Electrolytes
Electrolytic conductors
Electronic conductors
Ionic transport
Membrane
Molar conductivity
Resistance
Solid-state polymer electrolytes

OUTLINE

INTRODUCTION

Transport properties of electrolyte solutions (i.e., *conductance, diffusion, viscosity, thermal diffusion*, etc.) play a relevant role in basic and applied researches because they control most processes occurring in many fields of pure and applied science, such as chemistry, biology, medicine, colloids science, engineering, and materials science. In this article, electrolytic conductance will be emphasized, considering in particular the physical phenomena on which this quantity is based, the experimental apparatus used for its measurement, and some of the theories proposed to explain the variation of conductance with the electrolyte concentration. It will also be shown that conductance is a macroscopic property yielding both kinetic information in the form of ionic mobilities and thermodynamic information in the form of thermodynamic equilibrium constants for the formation of ion pairs. Conductance behavior also yields information on electrolytic systems at the microscopic level when the experimental data are analyzed with the theoretical equations based on models in which ion–solvent and ion–ion interactions are included.

CONDUCTION OF ELECTRICITY THROUGH MATTER

The transport of electricity through matter is effected by the movement of charge carriers. Depending on the nature of the charge carriers involved in the process, two basic classes of conductors can be distinguished. In the first class (*electronic conductors*), electricity is carried by electrons; in the second class (*electrolytic conductors*), electric current is carried by ions. Electronic conductors include metals and semiconductors.

The properties of the electronic conductors follow solid-state band theory. The rigorous treatment of electronic conductors is beyond the scope of this article, but to provide some background information for the reader, the mechanism of electrical conduction in metals can be summarized as follows. The energy levels of isolated atoms have discrete values. The filling of these energy levels by electrons is governed by the laws of quantum mechanics. As atoms approach one another to form a crystal lattice, the original energy levels combine to form energy bands. In metals, the conductivity bands correspond to the highest, partially occupied, energy levels of the metal atom in the ground state and contain a sufficient number of loosely bound electrons in random motion. Under the influence of an external electric field, the original random motions of electrons become oriented, and an electron flux (*electrical current*) runs along the conductor.

Quite different is the mechanism of conduction in electrolytic conductors. In electrolyte solutions, ions and solvent molecules are in rapid thermal motion, and any net ionic migration is essentially a small perturbation along the direction of the electric field superimposed on the former. Because ionic velocities are only of the order of 10^{-6} s^{-1} for an electrical field of unity (1 V/cm^{-1}), the solvent is usually approximated as a continuum in which the ions are assumed to move at a constant velocity, which is the time average equal to the sum of their components of motion in the direction of the applied field divided by the time of observation. Electrolytic conductance is, in last analysis, a measure of the steady-state flux achieved between the effects of an external electrical force and those of the internal hydrodynamic resistance to ionic motion. As such, it is an important research tool with which molecular interactions may be conveniently investigated.

In many technological processes (i.e., *electrodialysis, ion-selective electrode production, environmental and waste applications*, etc.) as well as in some biological processes (i.e., *permeation of ions through cellular membrane*), an artificial or natural membrane is interposed between two electrolyte solutions. From a physical point of view, the *membrane* is a potential barrier for the ionic flux; thus, the mechanism of conduction in these systems differs from that occurring in the simple electrolytic systems. Membranes can be divided in two main types: membranes that have a preferential permeability for a particular ion and membranes that are totally unpermeable to any kind of ions. The membranes of the first type are used in the technological applications where ion selectivity is the main factor of the process. The ionic conduction in these electrolyte membrane systems is only allowed for a particular ionic species. Quite different is the conduction mechanism for the systems containing a membrane of the second type. In these systems, the flux of ions (*hydrophilic*) is hindered by the hydrophobic character of the membrane (e.g., *lipidic bilayers*). To overcome this handicap, two alternative mechanisms are possible (*carrier and channel mechanisms*). Ionic transport mediated by carrier is obtained by using macrocyclic compounds (1,2) (e.g., natural antibiotics such as valinomycin and enniatin B or synthetic macrocyclic compounds such as crown ethers, cryptands, etc.). These carrier compounds form intrusion complexes with alkali metal ions. In these complexes, the central hydrophilic ion is hosted into a cavity containing oxygen atoms, and the exterior of the complex is hydrophobic. Such a feature, decreasing the energy barrier imposed by the membrane to the ionic flux, allows for the transport of the hosted ionic species across the hydrophobic membrane. It has been observed that the conductance of lipidic bilayers (10^{-10} S/cm) linearly increases with the carrier concentration. Such a result suggests that current-carrying species are 1:1 complexes between alkali metal cation and macrocyclic molecules. A transport mechanism consistent with the experimental results is depicted in Figure 1. According to this mechanism, the ionic transport across the membrane occurs in four distinct steps: (*1*) Association of ion M^+ and carrier S in the interface electrolyte solution–membrane (rate constant K_R); (*2*) Translocation of the complex M^+S to the opposite side of the membrane (rate con-

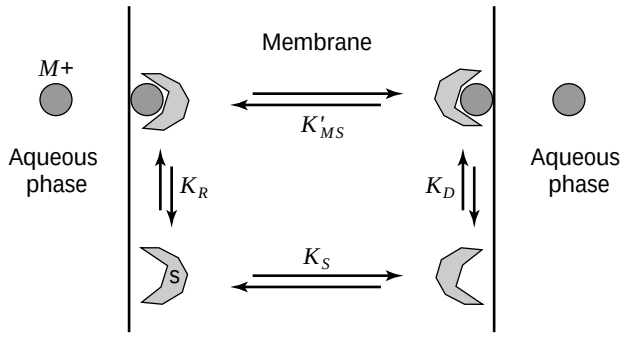


Figure 1. Carrier mechanism.

stant K'_{MS}); (*3*) Release of the ion to the solution (rate constant K_D); and (*4*) Back transport of the unloaded carrier (rate constant K_S). The kinetics of this carrier mechanism for a given system can be analyzed by studying the electric conductance behavior (3).

Different from the ionic transport mediated by carrier compounds is the channel mechanism. A channel consists of one or several binding sites arranged in a transmembrane sequence, and it is accessible from both sides at the same time. A number of simple *molecules* (e.g., the linear pentadecapeptide gramicidin A [4]) introduced into a biological or artificial membrane form channels permeable to cations. A detailed model of the gramicidin A channel was proposed by Urry (5). Information on the mechanism of ion transport through channels has been obtained in experiments (3) in which small amounts of gramicidin A were added to a planar bilayer membrane. Under this condition, the membrane current under a constant applied voltage shows discrete fluctuations with more or less the same amplitude. The experiment has been explained assuming that the single current fluctuation results from the formation and disappearance of a single channel permeable to ions.

In addition to electronic and electrolytic conductors, there is a special class of *solid-state polymer electrolytes* (6) that have recently attracted considerable worldwide interest for their prospective applications in batteries, sensors, and other practical devices. Polymer electrolytes are simply formed by dissolving an inorganic salt in suitable polymeric hosts. These host polymers contain electronegative heteroatoms (*oxygen*, *nitrogen*, *sulfur*, or *phosphorus*) that interact with cations by donor–acceptor bonds to form polymer–cation complexes. Poly(ethylene oxide) (PEO) and poly(propylene oxide) (PPO) are the most-used polymers that form solid-state polymer electrolytes.

PEO: -CH-CH-O-

PPO: -CH-CH-O-
|
CH

The mechanism of ionic transport in such systems was initially attributed to ion hopping between fixed coordination sites of polymer. Recently, the percolation theory (7) formally describes these ion conductors as a random mixture of conductive islands interconnected by a nonconductive matrix in terms of ion size, charge, and interactions with other ions and the host matrix.

As stated before for the electronic conductors, a detailed description of the conductivity properties of polymer electrolytes and the transmembrane ionic transport is beyond the scope of this article, which focuses mainly on electrolyte conductance, namely, *conductance in liquid electrolyte solutions*.

ELECTROLYTIC CONDUCTANCE

In an electrolytic conductor, the relationship between the steady-state flux (J) of ions and the driving forces of the electric field (X) can be represented by the power series

$$J = A + BX + CX^2 + DX^3 + \ldots \quad (1)$$

where A, B, C, and so forth are constants, and the units of J and X are $cm^3/s/K$ and volt/cm, respectively. For small applied fields, the terms containing the powers of X greater than unity can be neglected, and equation 1 therefore becomes

$$J = A + BX \quad (2)$$

The constant A in equation 2 must be zero because the ionic flux vanishes when $X = 0$. Hence, for small electric fields, the ionic flux is linearly related to the driving force

$$J = BX \quad (3)$$

where B is a constant specific of the system under investigation. The ionic flux J, representing the number of moles of ions crossing a unit area per second, can be easily converted into the charge flux or the current density i (number of electrical charge crossing unit area per second) by the relationship

$$i = JzF = zFBX \quad (4)$$

In the preceding equation, z is the charge number, i has units of ampere (A) per cm^{-2}, and F is the faraday equal to $F = eN$ (e is the electron charge equal to $e = 1.602177 \times 10^{-19}$ coulomb, and N is the Avogadro's number equal to $N = 6.02214 \times 10^{23}$ mol^{-1}). For a given electrolytic system, zFB is constant. Setting this constant equal to χ, equation 4 becomes

$$i = \chi X \quad (5)$$

that can be rearranged in the form

$$\chi = i/X \quad (6)$$

where χ is the *electrolytic conductance* as of a solution of known concentration (also referred to as the *specific conductance* in older literature) and has units of siemen (S) cm^{-1}. Table 1 compares representative values of the specific conductance of electrolytic solutions with those of metallic conductors and nonelectrolytic systems.

Table 1. Representative Values of Electrolytic Conductance at 25 °C

Substance	Type of conductor	Electrolytic conductance (S/cm)
Acetonitrile	Nonelectrolyte	1×10^{-8}
Water	Nonelectrolyte	2×10^{-7}
0.1 M KCl	Electrolytic	1×10^{-2}
4 M H_2SO_4	Electrolytic	8×10^{-1}
Iron	Metallic	1×10^{5}
Copper	Metallic	6×10^{5}

By definition, the electric field X is the ratio of the potential difference ΔV between two regions of the electrolyte and the distance l between them.

$$X = \Delta V/l \tag{7}$$

On the other hand, the density current (i) is the ratio of the total current (I) and the area (A) of the electrodes

$$i = I/A \tag{8}$$

Thus, combining equations 7 and 8 with equation 6 one has

$$\Delta V = I(l/\chi A) \tag{9}$$

According to Ohm's law, the quantity $(l/\chi A)$ must be a resistance, that is,

$$R = (1/\chi)(l/A) \tag{10}$$

The modern units for resistance is reciprocal siemen (S^{-1}), and the older unit of ohm (Ω) is no longer used. Equation 10 indicates that electrolytic conductors, like metallic conductors are ohmic conductors (i.e., they obey the Ohm's laws) and can therefore be represented in an electrical circuit as a resistor.

Measurements of Electrolytic Conductance

In conductance experiments, the electrolyte is placed in a conductivity cell of controlled geometry, and the resistances (R) of solutions of known concentration are measured. The cells are generally constructed in Pyrex glass with sealed platinum electrodes. Other noble or passive metals, such as gold, silver, titanium, stainless steel, and tungsten, as well as graphite can be used as electrodes for technical applications of conductivity (i.e., routine or in situ industrial measurements). Cells designs are available for a variety of research and industrial applications. The different types of conductance cells have in common an electrode compartment, shown schematically in Figure 2. The geometry factor, l/A (cm^{-1}), known as the cell constant, κ, is generally determined by calibrating the cell with a solution of known conductivity. Aqueous potassium chloride solutions are usually used for this purpose (8) because the specific conductance of this electrolyte at different concentrations and temperatures has been determined with a very high degree of accuracy. Cells are designed

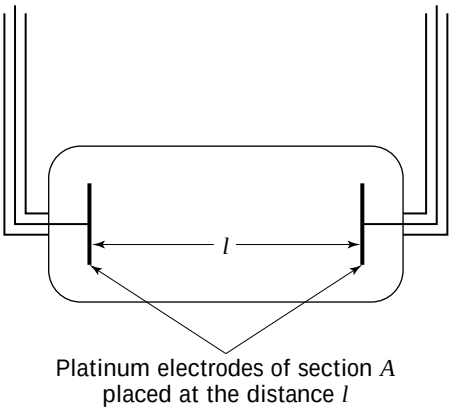


Figure 2. Schematic representation of conductometric cell.

with various cell constants so that the resistance of different electrolytic systems ranges between 1,000 and 30,000 S^{-1} in order to avoid errors arising from high density current and insulation linkage. Table 2 summarizes recommended cell constants for various conductance ranges.

The electrical resistance of electrolyte solutions cannot be measured with a traditional bridge circuit (e.g., a Wheatstone bridge) operating with direct current (dc) because this would give rise to polarization at the electrodes and in some cases to the electrolysis of the solution. Polarization produces undesirable changes in the electrolyte concentration during measurements that must be avoided because the resistance of the electrolytic system depends on ionic concentration. The simplest way to achieve this requirement is to use alternating current (ac) in the frequency range between 500 and 20,000 Hz. This rapid reversible change in applied potential effectively eliminates polarization effects.

The theory and design of ac bridges suitable for electrolyte conductance measurements were discussed by Jones and Josephs (9) and by Shedlovsky (10). A crude representation of the different ac bridges commercially available can be made in reference to the Kohlraush's bridge circuit, schematically depicted in Figure 3. The conductivity cell forms one arm of the bridge and a standard variable resistance box (R) forms another arm of the bridge. The other two arms of the bridge consist of two closely equal resistances, R_a and R_b. With ac bridge impedances are determined. The impedance Z is defined by the relation

$$Z = (R + X_R)^{1/2} \tag{11}$$

where R is the resistance and X_R the reactance given by

Table 2. Recommended Cell Constants for Various Conductance Ranges

Conductance Range (S cm)	Cell constant (cm^{-1})
$(0.5–20) \times 10^{-6}$	0.01
$(0.1–20) \times 10^{-5}$	0.1
$(0.2–20) \times 10^{-4}$	1.0
$(0.1–20) \times 10^{-3}$	10.0
$(0.1–20) \times 10^{-2}$	50.0

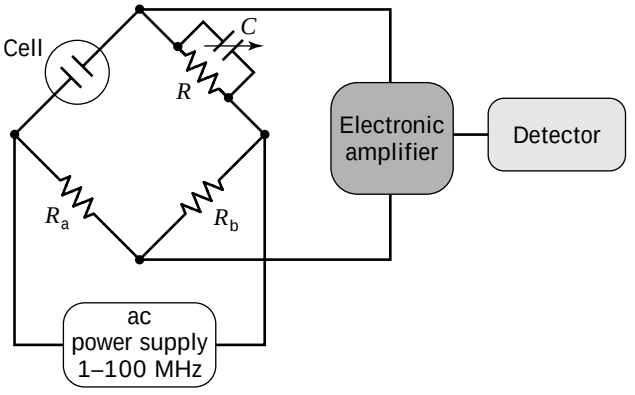


Figure 3. Schematic representation of a conductance apparatus based on Kohlraush's bridge circuit.

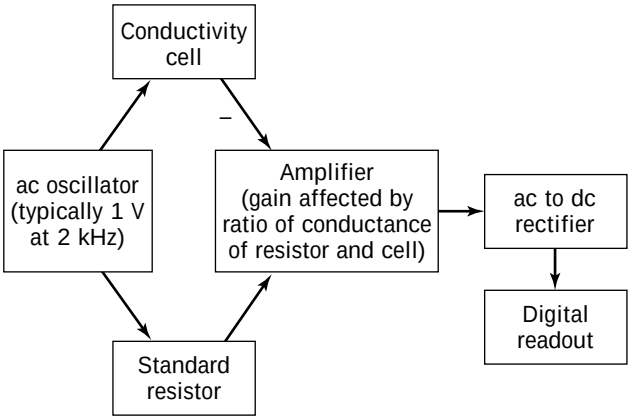


Figure 4. Block diagram of a typical digital conductivity meter.

$$X_R = X_L - X_C \tag{12}$$

where $X_L = 2\pi fL$ is the inductive reactance and $X_C = 1/(2\pi fC)$ is the capacitive reactance. In the preceding equations, R, X_L, and X_C are in reciprocal siemen (S^{-1}), f is the frequency in hertz, L is the inductance in henry, and C is the capacitance in faraday. Because the resistances used in the bridges for conductometric measurements are generally noninductance, capacities and resistances are the variables determined by balancing the bridge. Thus, only a standard variable condenser box (C) needs to be placed in parallel to the resistance box (R). For the measurement, both C and R are adjusted until zero signal from the electronic amplifier is detected. For this null condition, the following relations for the reactance capacities $X_C = 1/(2\pi fC)$ as well as for the resistances (R) can be written

$$\begin{aligned} [1/(2\pi fC)]_{cell}/[1/(2\pi fC)]_a &= [1/(2\pi fC)]/[1/(2\pi fC)]_b \\ &= [1/C)]_{cell}/[1/C)]_a \\ &= [1/C)]/[1/C]_b \\ R_{cell}/R_a &= R/R_b \end{aligned} \tag{13}$$

Because $[1/C)]_a = [1/C]_b$ and $R_a = R_b$, one obtains

$$\begin{aligned} C_{cell} &= C \\ R_{cell} &= R \end{aligned} \tag{14}$$

The ac bridge circuit is, in other words, a simple RC circuit allowing us to measure the individual impedance contributions from the resistance and the capacitance reactance. The block diagram for typical digital instruments used for routine resistance measurements is shown in Figure 4.

The conductance of electrolytic solutions changes by as much as 2% per degree Kelvin (Fig. 5). Thus, it is very important to have accurate temperature control within ± 0.001 K for high-precision measurements. Although the thermostatic bath may have any convenient and appropriate design, mineral oil rather than water is recommended for use in the bath. This is because the high dielectric constant and conductivity of water compared with mineral oil may result in leakage of electrical energy from the cell sufficient to introduce errors of appreciable magnitude. For routine or industrial applications such as in situ conductivity measurements, temperature control need not be very precise, and temperature regulation within ± 0.1 K is sufficient in most cases. Cells are available with built-in temperature sensors and associated instrumentation that automatically convert conductance values to a reference temperature, generally 25 °C. This feature on commercial bridges is referred to as automatic temperature compensation of conductivity.

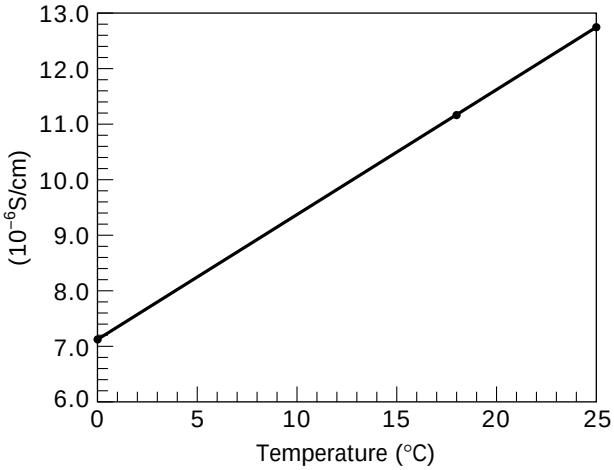


Figure 5. Dependence of the specific conductance of aqueous KCl solution (0.1 mol/L) with temperature.

For high-precision and high-accuracy conductivity measurements of dilute electrolyte solutions, the conductance of the solvent, χ_o, must be subtracted from the *apparent* (i.e., *experimental*) conductance, χ_a, as measured in the cell to obtain the real conductance, χ, of ionic species in solution.

$$\chi = \chi_a - \chi_o \tag{15}$$

The electrolytic conductance of commonly studied solvents is summarized in Table 3. Because conductivity measurements can be extremely accurate, by comparing measured

Table 3. Specific Conductance of Common Solvents at 25 °C

Solvent	Specific conductance (χ_0[S/cm])
Water	$1–2 \times 10^{-7}$
Methanol	$1–2 \times 10^{-8}$
Glycerol	$1–3 \times 10^{-8}$
Ethylene glycol	$6–8 \times 10^{-8}$
Acetonitrile	$3–4 \times 10^{-8}$
Propylene carbonate	$1–2 \times 10^{-6}$
Nitromethane	$1–3 \times 10^{-8}$
N-*N*-Dimethylformamide	$2–4 \times 10^{-8}$

conductances to those determined for highly purified solvents (see Table 3) the purities of solvents can be easily determined.

Molar, Equivalent, and Ionic Conductivity

The current density (i) across unit area per second can be expressed in terms of ionic velocity (v_i) by

$$i = F\Sigma_i c_i z_i v_i \tag{16}$$

where c_i, z_i, and v_i are the number of moles per cubic centimeter, the charge number, and the velocity in cm/s of the nth ionic species. Dividing equation 16 for the strength of electrical field (X) gives

$$i/X = (F/X)\Sigma_i c_i z_i v_i \tag{17}$$

Substituting equation 6 in equation 17 and setting the mobility of the nth ions $U_i = (v_i/X)$ one obtains

$$\chi = F\Sigma_i c_i z_i U_i \tag{18}$$

Restricting the treatment to 1:1 symmetrical electrolytes for which $z^+ = z^- = z$ and $c^+ = c^- = c$, equation 18 becomes

$$\chi = Fzc(U^+ + U^-) \tag{19}$$

Equation 19 shows that, contrary to electronic conductors, the specific conductance of electrolyte solutions depends on the ionic concentration. Thus, in order to normalize the electrolytic conductance for the concentration, a new quantity, the molar conductivity (Λ_m) is defined as

$$\Lambda_m = 10^3\chi/c_m \tag{20}$$

where c_m is the concentration (mol dm^{-3} units) of the electrolyte. The units of Λ_m are S cm^2 mol^{-1}. Molar conductivities taking into account only the molar concentration of the electrolyte do not always normalize the conductance, as in the cases where multiply charged ions are involved. To account for unsymmetrical electrolytes, such as $MgCl_2$, normalization is achieved by considering equivalent concentrations, $c = c_m/z$. In terms of equivalent concentrations, the normalized quantity is the *equivalent conductance* Λ (S cm^2 eq^{-1}), defined similarly to equation 20 by

$$\Lambda = 10^3\chi/c \tag{21}$$

Figure 6 shows that the equivalent conductance for a typical electrolyte (KCl in water) decreases as the ionic concentration increases. Thus, for meaningful comparisons of conductivity differences for different electrolytes, it is necessary to define a common reference state for equivalent conductances. The reference state is that of infinite dilution where the equivalent conductance at zero electrolyte concentration is referred to as the *limiting molar conductivity* and given the symbol Λ_o. As will be shown in the next section, the dependence of conductance on ionic concentration results from interionic interactions. Because at infinite dilution such interactions are negligible, only ionic velocity is affected by the external electrical field. Under these circumstances, Λ_o represents the conductivity of the electrolyte solution in its ideal state.

Substituting equation 21 in equation 19 gives

$$\Lambda = F(U^+ + U^-) = \lambda^+ + \lambda^- \tag{22}$$

where λ^+ and λ^- are the ionic equivalent conductance of the cation and anion at the given concentration. At infinite dilution, equation 22 for the limiting equivalent conductance becomes

$$\Lambda_o = \lambda_o^+ + \lambda_o^- \tag{23}$$

An important implication of this equation is that at infinite dilution, molar conductivities are *additive*. This property, discovered by Kohlrausch, is known as the *Kohlrausch law of independent ionic conductivity*. Thus, the limiting equivalent conductance for any electrolytic system can be obtained from the sum of the contributions of the individual ions. Limiting equivalent conductivity for a large variety of ionic species in different solvents as a function of temperature has been measured by many researchers. Table

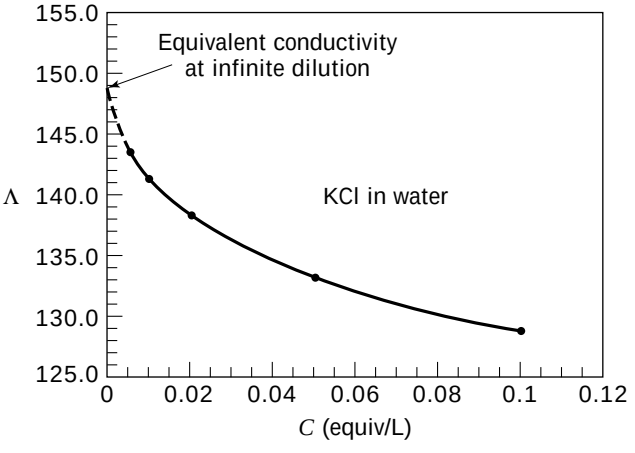


Figure 6. Dependence of Λ(S cm^2 equiv^{-1}) with electrolyte concentration C (equiv/L).

4 reports individual ionic λ_o values for selected ions in water at 25 °C.

THEORIES OF ELECTROLYTIC CONDUCTANCE

The search for a mathematical expression able to theoretically predict the concentration dependence of the molar conductivity has been the object of numerous scientists since the time of Arrhenius (1883). An ideal theory of conductance would predict Λ from fused salt systems (*solute mole fraction equal to unity*) to infinite dilution (*solvent mole fraction equal to unity*). In spite of this intense effort, the different conductance equations so far proposed take into account only the influence of the ion–ion and ion–solvent interactions on the ionic mobility and the concentration of free ionic species in solution. Thus, they are limited to concentrations up to where the effect of short range three-ion interactions (i.e., *triple ions formation*) becomes significant.

Debye–Onsager–Fuoss Theory

Because of the electrical charge on the ions, the ionic distribution in an electrolyte solution is modulated by the ion–ion interaction field superimposed to the ionic thermal motion. In order to quantitatively express the contribution of electrostatic interactions on ionic distribution, it is necessary to choose a model to represent the system. Although the model must include all the relevant properties of solute and solvent, the problem is to select a sufficiently simple model that is amenable to available mathematical treatments. All models for conductance behavior, though different in some details, are essentially based on the *primitive model* proposed by Debye and Hückel (11) in which

1. The electrolyte is completely dissociated at all the concentrations.
2. The solution is considered to be a collection of point charges (ions) in a homogeneous medium of viscosity (η) and relative permittivity (ϵ) (hydrodynamic continuum).
3. The Coulombic field surrounding the ions has spherical symmetry.
4. Only ion–ion long-range electrostatic interactions are considered.
5. The electrostatic potential energy of each ion is smaller than the thermal energy, that is, $|z_i e\psi| \ll kT$.

Table 4. Equivalent Ionic Conductivity of Selected Ions at Infinite Dilution in Water at 25 °C

Cations	λ_0^+(S cm^2 equiv^{-1})	Anions	λ_0^-(S cm^2 equiv^{-1})
H^+	349.8	OH^-	198.30
Li^+	38.7	Cl^-	76.4
Na^+	50.1	I^-	76.8
K^+	73.5	Br^-	78.1
Rb^+	77.8	NO_3^-	71.4
Cs^+	77.3	ClO_4^-	67.3
Ag^+	61.9	ClO_3^-	64.0
NH_4^+	73.5	HCO_3^-	44.5
$(n\text{-}Bu)_4N^+$	19.5	$CH_3CO_2^-$	40.9
Ca^{2+}	59.5	$HC_2O_4^-$	40.2
Ba^{2+}	63.6	$C_6H_5CO_2^-$	32.4
Mg^{2+}	53.1	$Picrate^-$	30.4
Zn^{2+}	52.8	CO_3^{2-}	69.3
Cu^{2+}	53.6	$C_2O_4^{2-}$	74.2
Pb^{2+}	69.5	SO_4^{2-}	80.0
Fe^{2+}	54.0	$Fe(CN)_6^{3-}$	101.0
Fe^{3+}	68.0	$Co(CN)_6^{3-}$	98.9
$Co(NH_3)_6^{3+}$	102.3	$Fe(CN)_6^{4-}$	110.5

In the Debye–Hückel theory, the central ion is considered, from a time-averaged point of view, at rest. The ion is surrounded by a symmetrical ionic cloud localized at the distance (*Debye distance*) given by

$$x^{-1} = [(\varepsilon T/\Sigma z_i c_i)(1{,}000\ \kappa/8\pi e^2 N]^{1/2} \qquad (24)$$

where z_i and c_i are the charge number and the concentration of the nth ion, respectively, ϵ is the static dielectric constant of the solvent, T is the temperature (Kelvin), κ is the Boltzman constant ($1.38\ 10^{-6}$ erg/degree), e is the electron charge ($4.80\ 10^{-10}$ absolute electrostatic unity [e.s.u.]), and N is the Avogadro's number. To the extent that one seeks to interpret the equilibrium properties of electrolytic solutions (i.e., *activity coefficients*), this picture of a static central ion surrounded by a ionic cloud with spherical symmetry is quite reasonable. Ionic interactions yielding the formation of spherically symmetrical ion clouds in the absence of external fields change from spherical to axial symmetry (*egg shape*) when a uniform external field (i.e., electric field as in the conductance or chemical potential field as in the diffusion) is applied to the electrolyte solution (Fig. 7).

Onsager (12) accounts for this asymmetrical distribution by considering the mutual interactions between an ion and its asymmetrical cloud. In the mathematical treatment of the migration of ions, he assumed that the drift velocity of an ion, besides the electric force arising from the externally applied field, is also determined by two other extra forces (i.e., *relaxation* and *electrophoretic forces*). The basic conductivity equation obtained by Onsager is

$$\Lambda = \Lambda_{0^\circ} - S\sqrt{c} \qquad (25)$$

where Λ_0 is the limiting conductance due to the unper-

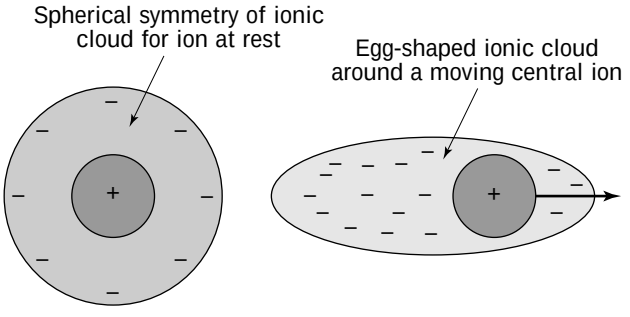


Figure 7. Ionic cloud for central ion at rest and in movement.

turbed ion movement in the external field at infinite dilution, and S is a constant that depends on absolute temperature, valence type of solute, dielectric constant, and viscosity of solvent. S includes the relaxation and the electrophoretic effect calculated via two particle correlation functions, assuming ions to be point charges and truncating series terms at the term $c^{1/2}$. It is important to realize that the constant S (and the constant E to be described in the next paragraph) does not account for finite size of ions and their solvation shells. For solvents of high dielectric constants ($\varepsilon > 45$), equation 25 can be considered an acceptable extrapolation formula, and we refer to it as the *limiting law*. However, for solvents of lower dielectric constant, deviations from equation 25 become large even in the case of completely dissociated electrolytes. The reasons for this drastic limitation arises from the crude assumption of the Debye–Hückel physical model, which represents the ions as point charges.

A further development of the conductance theory was made by Fuoss and Onsager (13), who treated the ions as finite-size particles. The resulting conductance equation derived by Fuoss and Onsager is

$$\Lambda = \Lambda^{o} - S\sqrt{c} + Ec \log c + J_a c \qquad (26)$$

where Λ_0 and S have the same meaning as in equation 25, E is a new constant dependent on the same variables as S, and J_a is a parameter that according to the new model depends on the cation–anion distance a, which can be used as a variable in fitting data to equation 26 or simply estimated as the sum of crystal radii of cation and anion; a is therefore referred to as the *distance of closest approach*. The logarithmic term in equation 26 derives from the integration of the higher terms of the continuity equation. Although Equation 26 is based on a more realistic model (*rigid charged spheres in a continuum*) it does have limitations because it often fails to accurately describe conductance behavior in many solvents, particularly solvents with low dielectric constants. However, for solutions containing salts with large ions, such as tetraalchylammoniun salts in high dielectric constant solvents, equation 26 does in fact reproduce the experimental data with acceptable accuracy. The problems encountered with equation 26 arise when electrolytes with small ions in solvents of moderate dielectric constants are involved, and where ion pair formation (Fig. 8) becomes significant. Ion pairs behaving as a dipole in the external electric field do not contribute to the conductance current.

In 1926, J. Bjerrum (14) developed a theory of ionic association. This theory was based on the observation that for certain systems the ratio of measured activity coefficients to the stoichiometric concentration was less than the activity coefficients calculated by the Debye–Hückel theory. To account for such discrepancies, Bjerrum postulated that some ions associated to ion pairs that as dipoles are no longer included in the ionic atmosphere. In order to obtain the degree of association of ions into neutral ion pairs, Bjerrum calculated the probability function [$W(r)$] of finding an anion in the spherical shell surrounding the reference cation. By applying a mathematical treatment analogous to that used by Debye– Hückel in the derivation of the activity coefficients, he found an analytical form for $W(r)$ that after a sharp minimum, diverged to infinity (Fig. 9). In order to mathematically treat this function, Bjerrum arbitrarily cut the integral off at a certain critical distance between the cation and the anion at which the electrostatic attractive force is balanced by the mean force corresponding to the thermal motion. Such a distance, in which the ion is at minimum energy, depends on the valence type (z), electron charge (e), dielectric constant of solvent (ε), Kelvin temperature (T), and Boltzman constant (k). For uniunivalent electrolytes, this critical distance (*Bjerrum distance*) is represented by the equation

$$R = e^2/2\varepsilon kT \qquad (27)$$

Integration of the probability function from the distance of closest approach (a) of the ions to R allows one to calculate the fraction of ions present in solution in the associated state. Thus, the degree of ion-pair association and the ion-pair association constant can be calculated.

Taking ion pair formation into account, Fuoss (15) extended the previous conductivity equation 26 to

$$\Lambda = \Lambda^{o} - S\sqrt{c\gamma} + Ec\gamma \log c\gamma + jc\gamma - K_a c\gamma f_{\pm}^{2}\Lambda \qquad (28)$$

where $f_{\pm}^{2}$ is the mean ionic activity coefficient calculated by the Debye–Hückel theory, and γ is the fraction of the electrolyte that exists as unassociated ions. In terms of γ, the thermodynamic equilibrium constant for the formation of ion pairs is determined by the mass equation

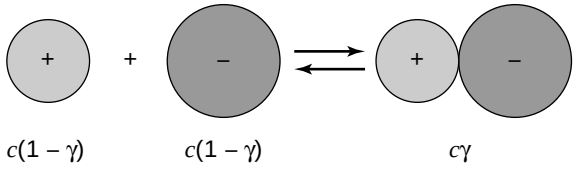


Figure 8. Schematic representation for ion-pair formation.

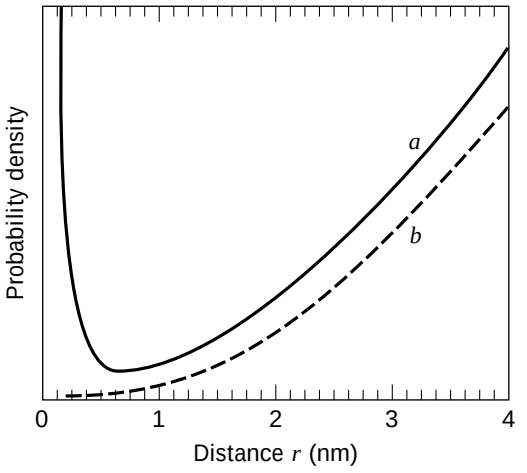


Figure 9. Bjerrum's curves of probability. (**a**) Two ions of opposite charges; (**b**) uncharged particles.

$$K_a = (1 - \gamma)/c\gamma^2 f^2 \quad (29)$$

Subsequently, Quint and Viallard (16) and Lee and Wheaton (17) extended the above equation to mixed electrolytes and for unsymmetrical electrolytes (e.g., $MgCl_2$). After the publication of equation 28, the theoretical treatment of the conductance of electrolyte solutions seemed to be completely solved in term of the electrostatic theory. However, numerous accurate experiments studies during the past 25 years have shown that equation 28 still has severe limitations. In particular, the systematic investigation on the conductivities of electrolytes in different solvents carried out by many researchers confirmed the limitation of equation 28 (or similar conductivity equations derived by Pitts [18], Kraeft [19], and Fuoss et al. [20]). Part of these limitations were eliminated by refining the mathematical treatment of the electrophoretic and relaxation terms [21]. Nevertheless, the most important limitation of the theory was the small concentration range over which equation 28 can be applied.

Several mathematical improvements were made to push the concentration limit of the conductivity equation to higher values. Fuoss and Hsia (22), by a new integration of the equations of continuity and motion with retention of terms of order $c^{3/2}$, presented in 1967 the empirical $\Lambda(c)$ function

$$\Lambda + \Lambda^o - S\sqrt{c} + Ec\gamma \log c + Ac + Bc^{3/2} \quad (30)$$

This equation is valid for 1:1 electrolytes up to the concentration $C_{max} = 3\ 10^{-7}\ \varepsilon^3$ mol dm^3, above which ion pair formation becomes statistically undefinable due to the presence of triple ions (23). Thereafter, Fernandez-Prini (24) expanded the function of the Fuoss–Hsia equation using, as suggested by Justice (25), the Bjerrum radius (R) as the lower limit in the integration instead of the contact distances a. The modified conductivity equation derived by Justice and Fernandez-Prini is

$$\Lambda = \Lambda_o - S\sqrt{c\gamma} + Ec\gamma \log c\gamma + J_1(R)c\gamma + J_2(R)c^{3/2}\gamma^{3/2} - K_a c\gamma f^2_{\pm}\Lambda \quad (31)$$

It turned out that equation 32 was also not free from theoretical and experimental problems. The most serious was the prediction that a given electrolyte in different solvents of the same dielectric constants should have the same association to ion pairs.

Many violations of the failure of this *isodielectric rule* (26) show that the primitive model can be considered an oversimplified physical model adequate to describe the conductometric behavior of electrolytic systems, in which long-range ion–ion interactions play the main role, but it completely ignores additional effects caused by anion–cation, anion–solvent, and cation–solvent short-range interactions. Recently, Fuoss (27) proposed a new model for ions in solution, based on the concept of the Gurney cosphere (28) surrounding each ion. This *chemical model* takes into consideration short-range specific ion–solvent interactions (*solvation*) related to the polarizability, dipole moment, and shape and size of ions and solvent molecules, both of which are explicitly ignored in the primitive model. This new model involves the dual roles played by the ions involved in ion-pair formation. As long as the distance between the two ions is greater than the diameter of the Gurney cosphere, an ion involved in pairing is an anonymous member of an enormous population; once inside the cosphere, the ions behave like particles and move from site to site by interchanges with solvent molecules. Because of this duality, the long-range interactions (*relaxation field* and *electrophoresis*) can be determined by using continuum electrostatic and hydrodynamic theories, whereas short-range interactions can be determined by the thermodynamic changes of the free energy of the pairing process involving both solvent separate ion pairs and contact ion pairs. For symmetrical electrolytes, the conductance equation based on the chemical model can be written in the following symbolic form

$$\Lambda = \gamma\{[\Lambda_o(1 + (\Delta E/E)] + \Delta\Lambda^v + \Delta\Lambda^H + \Delta\Lambda^{el}]\} \quad (32)$$

In this equation, where γ is the fraction of solute that contributes to the conductance current, $(\Delta E/E)$ is the part of the relaxation field generated by purely electrostatic interactions between ions, $\Delta\Lambda^v$ is the part of the relaxation field caused by the change in the electrophoretic current caused by perturbation potential, $\Delta\Lambda^H$ is the Sanding–Feistel (29) hydrodynamic–hydrodynamic interactions term, $\Delta\Lambda^{el}$ is the electrophoretic countercurrent, and Λ_o is the limiting conductance. Starting from this chemical model proposed by Fuoss, Lee and Wheaton, in a series of papers (30), derived a conductivity equation for the general case where any number of ionic species of any valence type may be presented in solutions. Thus, not only 1:1 electrolytes but also asymmetrical and mixed electrolytes can be analyzed.

BIBLIOGRAPHY

1. B. Sesta and A. D'Aprano, *J. Solution Chem.* **18**, 1163–1171 (1989).
2. A.Y. Ovvhinnikov, V.T. Ivanov, and A.M. Skrob, *Membrane-Active Complexones*, Elsevier, Amsterdam, 1974.
3. R. Benz and P. Lauger, *J. Membr. Biol.* **27**, 171–191 (1976), P. Lauger, *Science* **178**, 24–30 (1972).
4. S.B. Hladky and D.A. Haydon, *Biochim. Biophys. Acta* **274**, 294–312 (1972).
5. C.W. Urry, *Proc. Natl. Acad. Sci. USA* **68**, 672–676 (1971).
6. J.R. MacCallum and C.A. Vincent eds., *Polymer Electrolyte Reviews*, Elsevier, London, 1987.
7. M.A. Ratner and D.F. Shriver, *Chem. Rev.* **88**, 109–112 (1988).
8. J.E. Lind Jr., J.J. Zwolenik, and R.M. Fuoss, *J. Am. Chem. Soc.* **81**, 1557–1559 (1959).
9. G. Jones and R.C. Josephs, *J. Am. Chem. Soc.* **50**, 1049–1092 (1928).
10. T. Shedlovsky, *J. Am. Chem. Soc.* **52**, 1806–1811 (1930).
11. P. Debye and E. Hückel, *Z. Phys.* **24**, 305–325 (1923).
12. L. Onsager, *Z. Phys.* **28**, 277–284 (1927).
13. R.M. Fuoss and L. Onsager, *Proc. Natl. Acad. Sci. USA* **41**, 1010–1025 (1955).
14. J. Bjerrum, *Kgl. Danske Videnskab. Selskab.* **7**, 9–16 (1926).
15. R.M. Fuoss, *J. Am. Chem Soc.* **79**, 3301–3303 (1957); R.M. Fuoss and F. Accascina, *Electrolytic Conductance*, Interscience Publishers, New York, 1959.

16. J. Quint and A. Viallard, *J. Chim. Phys.* **72**, 335–342 (1975).
17. W.H. Lee and R.J. Wheaton, *J. Chim. Phys.* **74**, 689–692 (1977).
18. E. Pitts, *Proc. R. Soc.* **217**, 43–48 (1953).
19. W. Kraeft, *Z. Phys. Chem.* **237**, 289–295 (1968).
20. R.M. Fuoss, L. Onsager, and J. Shinner, *J. Phys. Chem.* **69**, 2581–2594 (1965).
21. R.M. Fuoss, *J. Phys. Chem.* **63**, 633–636 (1959); R.M. Fuoss and L. Onsager, *J. Phys. Chem.* **67**, 628–632 (1963); R.M. Fuoss and L. Onsager, *J. Phys. Chem.* **68**, 1–8 (1964); M.S. Chen, Ph.D. Thesis, Yale Univ., New Haven, Conn., 1969.
22. R.M. Fuoss and K.L. Hsia, *Proc. Natl. Acad. Sci. U.S.A.* **57**, 1550–1601 (1975).
23. R.M. Fuoss, *J. Am Chem. Soc.* **56**, 1857–1859 (1934).
24. R. Fernandez-Prini and J.C. Justice, *Pure Appl. Chem.* **56**, 541–563 (1984).
25. J. C. Justice, *J. Chim. Phys. Phys.-Chim. Biol.* **65**, 353–362 (1968).
26. A. D'Aprano and R. M. Fuoss, *J. Solution Chem.* **3**, 45–55 (1974); A. D'Aprano, I. D. Donato, and E. Caponetti, *J. Solution Chem.* **3**, 371–376 (1974); A. D'Aprano, M. Goffredi, and R. Triolo, *J. Chem. Soc. Faraday Trans.* 1 **71**, 1188–1191 (1975).
27. R.M. Fuoss, *Proc. Natl. Acad. Sci. USA* **77**, 34–46 (1980); R.M. Fuoss, *J. Solution Chem.* **15**, 231–236 (1986); A. D'Aprano, F. Accascina, and R.M. Fuoss, *J. Solution Chem.* **19**, 65–78 (1990).
28. G. Gurney, *Ionic Process in Solution*, Dover, New York, 1953.
29. R. Sanding and R. Feistel, *J. Solution Chem.* **8**, 141–145 (1979).
30. W.H. Lee and R.J. Wheaton, *J. Chem. Soc., Faraday Trans.* 2 **74**, 743–766 (1978); W.H. Lee and R.J. Wheaton, *J. Chem. Soc., Faraday Trans.* 2 **74**, 1456–1482 (1978); W.H. Lee and R.J. Wheaton, *J. Chem. Soc., Faraday Trans.* 2 **75**, 1128–1145 (1979).

See also PROCESS MONITORING.

CORROSION, MICROBIAL

PATRIZIA PEREGO
BRUNO FABIANO
University of Genoa
Genoa, Italy

KEY WORDS

Biofouling
Cathodic reaction
Caulobacter
Chemio-lithotrophic microorganisms
Depassivation of metal surfaces
Desulfotomaculum
Desulphovibrio
Gallionella
Sphaerotilus natans
Thiobacillus

OUTLINE

INTRODUCTION

In its extensive meaning, biological corrosion, or biodeterioraton, refers to any corrosion caused by biological activity. Indeed, the original meaning of the term *corrosion*, as derived from *rodere*, the Latin root for "act of gnawing," includes biological activity. Microbiological corrosion refers, by definition, to the degradation of metallic structures due to the activity of a wide variety of microorganisms, usually embedded in a gel matrix, the biofilm, attached to the metal surface.

Before further discussion on this subject, a short review of the milestones in its historical development is presented. Evidence of anaerobic corrosion was observed as early as 1840, when Mallet noted sulfide corrosion products typical of anaerobic process on an ancient anchor on the bottom of the Seine River (1).

The phenomenon of microbially influenced corrosion (MIC) was reported by the end of the nineteenth century by Garret (2), who postulated that corrosion of lead in water was caused by metabolites produced by bacterial degradation of organic matter. He suggested that biooxidation of nitrite to nitrate enhanced the corrosive action of water.

Sulfate-reducing bacteria were discovered in 1895 by Beijerinck. In 1910 Gaines (3) indicated that iron bacteria and sulfur bacteria were responsible for the corrosion of buried ferrous metals.

The formation of deposits in waterpipes by iron bacteria (in particular aerobic species *Sphaerotilus natans*, *Caulobacter*, and *Gallionella*) was investigated by Harder (4). In 1921 Grant et al. (5) produced evidence that the rate of pitting of brass was highly increased by ammonia, which could be produced by bacterial activity.

In 1934, Von Wolzogen Kuhr and Van Der Vlugt (6) proposed a theory on the direct role of sulfate-reducing bac-

teria in the corrosion mechanism by their utilization of cathodic hydrogen and stimulation of the anodic dissolution of iron.

In 1953 Uhlig reported on the role of microorganisms, such as slime-forming bacteria, fungi, algae diatoms, and protozoa, attributing their corrosive action to the formation and conservation of differential aeration and concentration cells (7).

Several reviews dealing with MIC have been published since 1971 (8–11), with a particular effort in research during the so-called energy crisis. In fact, marine fouling on offshore platforms and biofouling in industrial cooling systems, as well as the problem of oil and gas reservoir contamination by sulfide, enhanced the need for a deeper insight into the role of, and the mechanism associated with, bacterial corrosion of metal. The adoption of accelerated tests, designed ad hoc according to newly developed measurement techniques, can address the right selection of materials, surface treatments, and welding filler materials, as well as the optimization of chemical and mechanical protective or preventive measures.

MICROORGANISMS IMPLICATED IN BIOLOGICAL CORROSION

To indicate proper defensive means against corrosion of metal materials, an accurate study and a deeper understanding of the mechanisms of action of the microorganisms involved are undoubtedly of the utmost importance. An essential role in the corrosion process is often played by the cell structures, for instance, the capsular material, essentially of a polysaccaridic nature, which has a protective function and allows surface adhesion. It is in this way that microorganisms create stratified deposits under which the cell's metabolic activity takes place. The capsular material contains substances in such concentrations as to actually exert a corrosive action on metals and, further, to ensure suitable conditions for development of some anaerobic microorganisms.

Microorganisms from a wide range of genera and species are involved in corrosion problems. They may be classified in three general groups: bacteria, fungi, and algae.

Bacteria

The bacteria associated with metal corrosion are unicellular, with a rigid cell wall, and divide by binary fission. They can be either autotrophic or heterotrophic, and either aerobic or anaerobic. There are three basic shapes of unicellular bacteria: spherical, rodlike, and curved or spirilloid. They vary considerably in size. Some iron bacteria are filamentous, obtaining the energy for carbon dioxide fixation by oxidation of ferrous ions to ferric ions with the consequent precipitation of ferric hydroxide. They are responsible for the formation of tubercles in potable-water-transmission pipelines. The taxonomy and physiology of these bacteria have been reviewed by Pringsheim (12) and Stokes (13).

The most important bacteria associated with the corrosion process are those whose metabolism is closely related to the well-known sulfur cycle reproduced in Figure 1. Aerobic bacteria of the *Thiobacillus* genus, known as sulfur-oxidizing bacteria, undertake the upper part of the cycle, (i.e., the oxidation of sulfur to sulfuric acid), and the central part of the cycle (i.e., reduction of sulfate to sulfide) is carried out by anaerobic bacteria (*Desulphovibrio* and *Desulfotomaculum* genus), indicated as sulfate-reducing bacteria (SRB). A list of the SRB and their main characteristics can be found in Table 1 (modified from Ref. 14). A list of other bacteria relevant to biological corrosion is found in Table 2.

Fungi

The fungi associated with corroding metal are generally filamentous or yeastlike. Unlike some bacteria, fungi require oxygen for growth, and obtain energy for growth through aerobic oxidation of an organic substrate, which releases organic acid metabolites as the end product, or through anaerobic fermentation. The fungi can utilize hydrocarbons as a source of carbon; they may be active in largely hydrocarbon environments and have proven to be the agents responsible for the corrosion of aluminium aircraft fuel tanks in the presence of condensed water.

Algae

Algae are a heterogeneous group of chlorophyll-containing plants found in seawaters and freshwaters. They are autotrophic organisms, and algae are commonly found in cooling water systems. Their role is similar to that of the

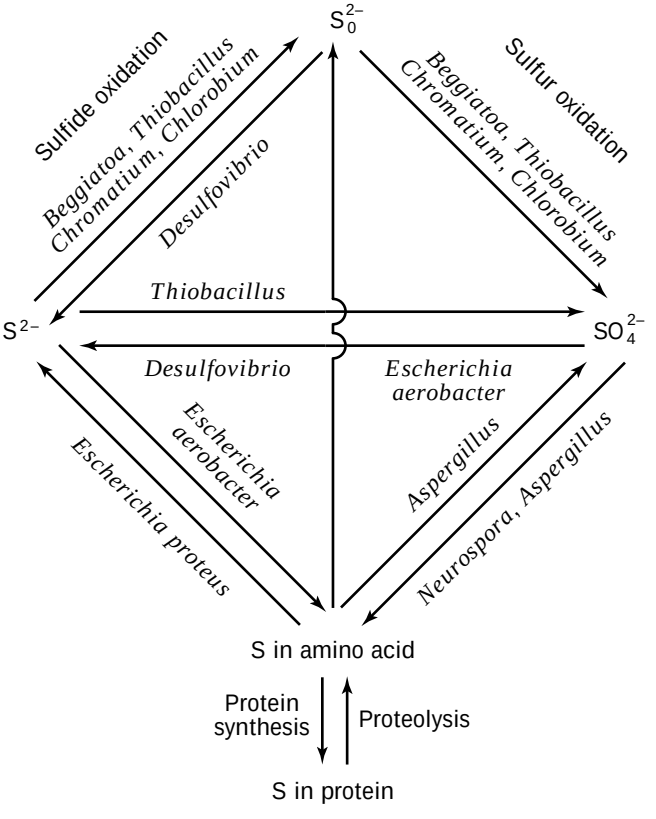


Figure 1. The sulfur cycle and microorganisms involved.

Table 1. List of Sulfate-Reducing Bacteria

Genus or species	pH	Temperature (°C)	Characteristics	Metal influenced	Mechanism
Desulfovibrio	4–8	10–40	Anaerobic	Iron and steel, stainless steel, aluminum, zinc, copper alloys	H_2 utilization to reduce SO_4^{2-} to S^{2-} and H_2S; promotion of sulfide film formation; no spore formation
Dv. desulfphuricans					Growth on pyruvate or choline in sulfate-free media
Dv. vulgaris					
Dv. salexigens					Obligate saltwater species (Cl^-)
Dv. gigas					
Dv. africanus					
Desulfotomaculum	6–8	10–40	Anaerobic	Iron and steel, stainless steel	Reduction of SO_4^{2-} to S^{2-} and H_2S; spore formation
Dt. nigrificans		55 optimum			Hydrogenase activity variable; not coupled to sulfate reduction; growth on pyruvate in sulfate-free media
Dt. orientis		10–40			Hydrogenase apparently absent
Dt. rumnis		10–40			Growth on pyruvate in sulfate-free media
Desulfomonas		10–40	Anaerobic	Iron and steel	Reduction of SO_4^{2-} to S^{2-} and H_2S

Note: From Ref. 15.

fungi: By adhering to the metal surface they encourage the formation of differential aeration and concentrate cells, thereby accelerating already-existing corrosion processes. The corrosive effect of algae attached to metal surfaces was investigated by Terry and Edyvean (15), who demonstrated that microfouling organisms readily attach to unprotected steels, cathodically protected steels, and painted steels within 100 days of exposure, causing alternating pH values night and day and enhancing pitting corrosion. Evidence of siliceous skeletons of diatoms causing defects in electroplated coatings of metals was shown by Ebneth and Ritterskamp (16).

MICROBIOLOGICAL CORROSION FEATURES

Metal corrosion in an aqueous environment is recognised as an electrochemical phenomenon where the anode reaction takes place when part of the substratum (metal) is oxidized and transferred into solution. The balancing cathode reaction consists of the contemporaneous reduction of components in the corrosive environment to fully preserve electroneutrality. Metal ions in solution might subsequently precipitate as free insoluble products, or may firmly adhere to the metal surface, where they can have a protective effect.

In favorable circumstances and under natural conditions, these films may have a protective function; a typical example of this function was seen in archaeological findings consisting of iron handiwork, which is well preserved even though found in corrosive ground. These well-preserved items were covered with a slightly compact, tightly adherent, thin, black film of ferric and ferrous phosphate.

When the metal is submersed in water, wet soil, or any other aqueous system, the initial reaction is metal dissolution to cation, which involves freeing of electrons:

$$Fe \rightarrow Fe^{++} + 2e^-$$

For the corrosion to continue, these electrons need to be consumed by reduction reactions that take place on the same metal on the so-called cathodic sites. Generally, the oxygen reduction to hydroxyl ion is the most common cathode reaction in a neutral environment:

$$O_2 + 2H_2O + 4e^- \rightarrow 4OH^-$$

In a neutral environment, the overall reaction is formation, and possibly subsequent precipitation, of insoluble products by the reaction of the ferrous ions with the hydroxyl ion (and, often, oxygen) in solution. As a rule, reactions occur evenly over the metal surface. Anode sites may be determined by either heterogeneous material (e.g., stress areas, inclusions) or preexisting, unevenly distributed oxide films. The overall corrosion rate may depend on quite a number of factors: Anode and cathode reactions, however, must remain balanced in the process to fully maintain electroneutrality. Lacking oxygen, the cathode reaction that may replace the preceding reaction in aqueous ambient in sustaining the corrosion process is reduction of the ion H^+, as follows:

$$2H^+ + 2e^- \rightarrow H_2$$

The reaction, however, is quickly checked in neutral solutions so that the corrosion process is rapidly stopped. At lower pH values, hydrogen evolution usually prevails, in both the presence and absence of oxygen. Besides these basic reactions, other factors affect the whole process such as, in particular, the metal surface microgeometric characteristics and the ambient oxygen concentration variation that creates differential aeration cells. Cathode sites are, in this case, the highest oxygenation areas. Similar concentrations of cells may occur on metal in the case of con-

Table 2. List of Other Bacteria Relevant to Microbial Corrosion

Genus or species	pH	Temperature (°C)	Characteristics	Metal influenced	Mechanism
Thiobacillus thiooxidans	0.5–8	20–40	Aerobic	Iron and steel, stainless steel	Sulfur and sulfide oxidation to H_2SO_4; protective coating damage
Thiobacillus ferrooxidans	1–7			Iron and steel	Fe^{2+} oxidation to Fe^{3+}
Gallionella	7–10			Iron and steel, stainless steel	Fe^{2+} oxidation to Fe^{3+}; manganous oxidation to manganic; tubercule formation promotion
Sphaerotilus	7–10	20–40	Aerobic	Iron and steel, stainless steel	Fe^{2+} oxidation to Fe^{2+} oxidation to Fe^{3+}; manganous oxidation to manganic; tubercule formation promotion
S. natans				Aluminium alloys	
Pseudomonas	4–9	20–40	Aerobic	Iron and steel, stainless steel	Some strains reduce Fe^{3+} to Fe^{2+}
P. aeruginosa	4–8			Aluminium alloys	

tact with soils or solutions where the concentration of aggressive ions is not homogeneous.

Many articles (10,11,17,18) on this subject have originally assumed that in order to participate in the corrosion process, microorganisms ought to be in a vegetative phase and produce oxydizing agents through their metabolism, or else directly intervene in one or both of the metal's electrochemical reactions. It has always been known that all bacterial colonies (not only the ones consisting of vegetative cells) may be responsible for differential aeration cell formation on metal surfaces. Microhabitats generated by these colonies often create anaerobic conditions that stimulate, in turn, the activity of SRB.

There are several ways of classifying MIC, such as on the basis of the metabolism of implicated microorganisms, as follows (10,11,19):

- Absorption of nutrients by microbial growth in adhesion to the metal surface
- Evolution of corrosive metabolites or end products of fermentative growth
- Sulfuric acid production
- Interference in the cathodic process under anaerobic conditions by obligate anaerobes

Another method of classification is on the basis of the mechanism involved (20):

- Corrosion by SRB
- Corrosion by acid producers
- Corrosion due to heavy microbial growth and deposits
- Corrosion associated with the rupture of protective coatings

Even though such classifications are useful, the mechanism and means of recognition of MIC are discussed, in view of industrial application, with reference to aerobic or anaerobic environmental conditions. Some practical examples are reported as well to provide the reader with the visual appearance of different MIC phenomena and practical ways of identification.

Bacterial Corrosion under Aerobic Conditions

Many aerobic microorganisms are capable of affecting the corrosion process. Some microorganisms are responsible for surface cell-deposit formation (bioscaling phenomena). In a marine environment, the biofouling phenomenon is mostly due to cirripeds, mussels, and so on. The corrosion process, in such cases, consists of formation of a differential-aeration cell due to oxygen assumption by the microbial colonies, tending to exhaust oxygen concentration under such mass. Poorly aerated surfaces behave as anodes, while better aerated surfaces, farther from the deposit, provide the cathode balancing of the reaction. The simple mechanism of aerobic corrosion is schematized in Figure 2 (after Ref. 10).

The formation of an occluded area on a metal surface is a mechanism potentially involved in MIC. It is based on the observation that when microorganisms form colonies on the surface of a metal, they do not constitute uniform layers but local colonization sites related to such metallurgical features as roughness, inclusions, surface charge, and superficial defects. The development of an occluded cell formed by sticky polymers is reproduced in Figure 3; biological and nonbiological (e.g., metals and chloride) species tend to aggregate to the colony site.

Microbiological corrosion on a reactor made of steel, with no addition of particular alloy elements and belonging to the FE33 commercial category, can be so marked under aerobic conditions as to be detectable by simple optical microscopy. A case study, taken from a reactor employed for yeast biomass industrial production and showing uniform generalized corrosion, is shown in Figure 4. The corroded surface, examined under a scanning electron microscope and reproduced in Figure 5, shows how the difference in molar volumes existing between oxide flakes and metal causes the process of spalling. Spalling of the corrosion products and superficial pitting due to the action of localized cellular deposits are shown in Figure 6.

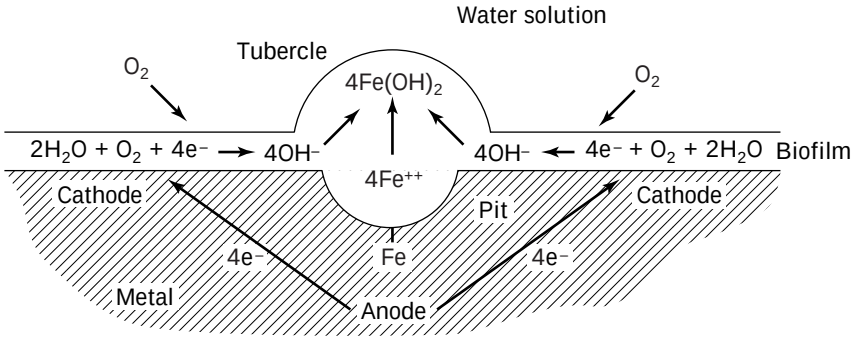


Figure 2. Aerobic corrosion mechanism by differential aeration cell (after Ref. 10)

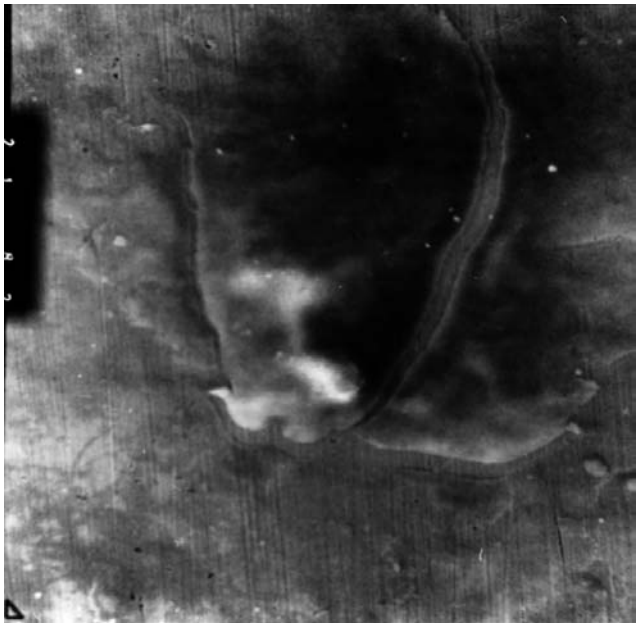


Figure 3. AISI 304 (500×): development of occluded cell under aerobic conditions.

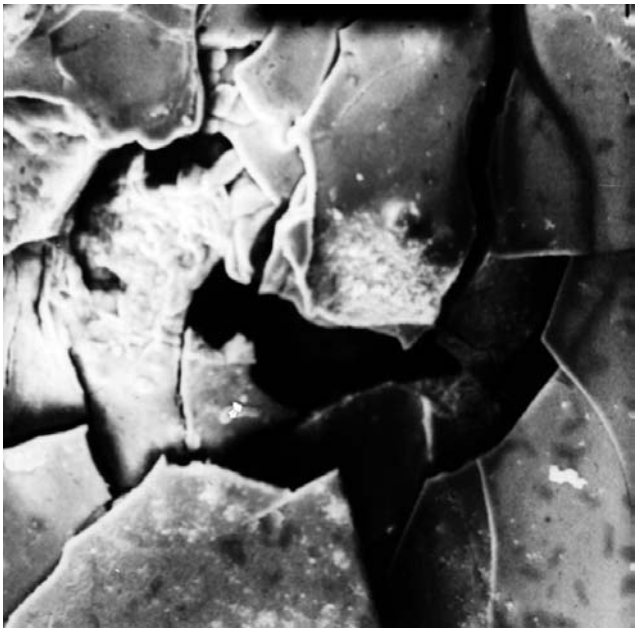


Figure 5. FE33 (500×): morphology of aerobic corrosion products showing spalling phenomenon.

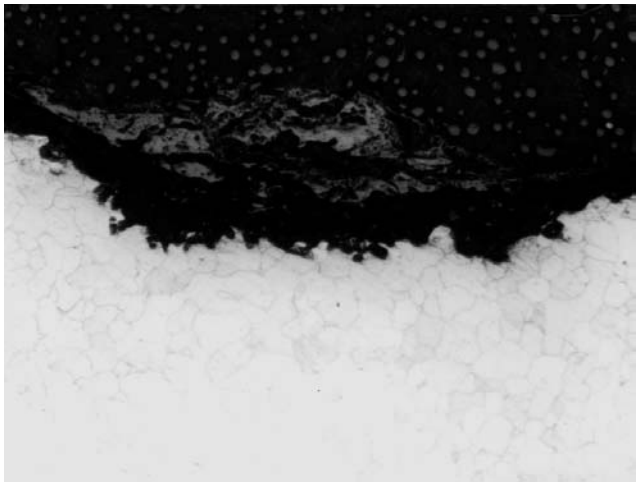


Figure 4. FE33 (120×): uniform generalized aerobic corrosion phenomena.

One particular type of aerobic corrosion is due to microorganisms belonging to the *Thiobacillus* genus. These are chemio-lithotrophic microorganisms that draw the energy needed for their vital processes from oxidation of inorganic substances. The nutrition functioning of these microorganisms, in particular *Thiobacillus ferrooxidans*, was investigated; various models have been proposed, among which one of the most feasible is that iron as a complex is absorbed through the microorganism's cell wall and donates the electron to the respiratory tract. The energy liberated during the subsequent metabolic phases is employed to form ATP, and the ferric iron and water formed are expelled from the cell, forming colloidal ferric hydroxide (21). So little energy is made available through the ferrous iron oxidation reaction (about 6.6 kcal/g-atom) that the organism is compelled to oxidize vast quantities of iron to obtain what it needs to grow.

Another type of microorganism, *Thiobacillus thiooxidans*, is capable of producing sulfuric acid. The relevant

Figure 6. FE33 (300×): spalling of aerobic corrosion products and superficial pitting due to localized cellular deposit.

mechanism has been studied in detail by Booth (22) and Purkiss (23): The generally accepted path involves the following reactions that may depend on the substrate and the environment:

$$2H_2S + 2O_2 \rightarrow H_2S_2O_3 + H_2O$$

$$5Na_2S_2O_3 + 4O_2 + H_2O \rightarrow 5Na_2SO_4 + H_2SO_4 + 4S$$

$$4S + 6O_2 + 4H_2O \rightarrow 4H_2SO_4$$

Bacterial Corrosion under Anaerobic Conditions

According to the previously mentioned considerations concerning electrochemical reactions, a corrosion process involving ferrous metals would seem unlikely in a near pH-neutral, oxygen-deficient environment, because neither of the two common cathode reactions can take place under these conditions at a sustained rate. This, however, is not always true, and when a corrosion phenomenon takes place in such conditions it is much more serious than under aerobic conditions.

The main factor responsible for this problem is the activity of SRB as well as their sulfured metabolic products. Buried metals corroded by this sort of process are one example of the severe economic losses and damage caused by this type of corrosion.

The process of anaerobic corrosion is characterized by a coating on corroding steel consisting of strongly reduced black sulfide–containing corrosion products. Graphitization takes place in cast iron, and little superficial evidence of corrosion can be noticed; the iron migrates from the material, leaving a soft residue, mainly carbon-made. As with bacterial involvement in the corrosion process, the explanation proposed by Wolzogen Kuhr and Van der Vlugt (6) is based on the assumption that cathode hydrogen may be utilized for sulfate reduction; the equations describing the process are as follows:

$$4Fe \rightarrow 4Fe^{2+} + 8e^- \quad \text{anode reaction}$$

$$8H_2O \rightarrow 8H^+ + 8OH^- \quad \text{water dissociation}$$

$$8H^+ + 8e^- \rightarrow 8H \quad \text{cathode reaction (*)}$$

$$SO_4^{2-} + 8H \rightarrow S^{2-} + 4H_2O \quad \text{cathode depolarization (*)}$$

$$Fe^{2+} + S^{2-} \rightarrow FeS \quad \text{corrosion product}$$

$$3Fe^{2+} + 6\,OH^- \rightarrow 3Fe(OH)_2 \quad \text{corrosion product}$$

The overall reaction can be written as

$$4Fe + SO_4^{2-} + 4H_2O \rightarrow 3Fe(OH)_2 + FeS + 2OH^-$$

Through the reactions marked by an asterisk, a biological action develops, bringing about hydrogen activation that, in turn, causes sulfate reduction. Several research projects carried out in recent years seem to support this assumption (24,25); lately, however, some studies point to an alternative process involving sulfide and the iron sulfide film (26,27).

Laboratory studies employing cultures of SRB are unlikely to produce heavily-pitted corrosion, and normal iron oxide and hydroxide corrosion products are nearly always absent. A similar phenomenon also occurs in the presence of elementary sulfur, which often forms, under natural conditions, in the periphery of pitted areas; it very seldom does so under laboratory conditions.

A combination of these and other observations related by King and Miller (27) has led to a more critical approach to the classic theory. In view of the particularly complex problems involved, we find it advisable to analyze the classic theory and cathode depolarization alternative mechanism separately.

Classic Theory. The Classic theory is based on the fact that some SRB strains possess the hydrogenase enzyme, which is capable of catalyzing both the reversible reactions involving hydrogen; that is,

$$H \leftrightarrow 2H \leftrightarrow 2H^+ + 2e^-$$

as well as sulfide reduction using molecular hydrogen, according to following:

$$SO_4^{2-} + 4H_2 \rightarrow S^{2-} + 4H_2O$$

Hydrogenase-positive strains of alophile, thermophile, and mesophile bacteria were shown by Booth and Tiller (28) to be capable of cathode depolarization, whereas hydrogenase-negative bacteria proved completely inactive in this process. Thus, strains such as *Desulphovibrio desulphuricans* and *Desulphovibrio vulgaris*, both hydrogenase positive, use the cathode hydrogen for substrate reduction, whereas hydrogenase-negative strains, such as *Desulfotomaculum orientis*, were found to play no role at all in this mechanism.

Among other hypotheses, a relation has been supposed to exist between the microorganism hydrogenase coefficient and the corrosion rate. In this manner, the steel corrosion rate would be related to hydrogenase activity, even if only in the limited field of experimental conditions. However, subsequent experiments that employed active cell cultures of several different strains of SRB indicated that the corrosion rate did not depend on the hydrogenase activity. And, further, hydrogenase-negative bacteria proved to be at least as aggressive as the *Desulphovibrio* species. This clearly indicates that other factors must be involved.

Cathode Depolarization Alternative Mechanism. Studies on cathode depolarization always refer to the acknowledged property of sulfides, both in solution and in solid state, (for instance, elemental sulfur) to act as corrosive agents. Further, it might be noted that the mild steel corrosion rate increases by up to 230% when in contact with iron sulfide. Serious corrosions in acidic environments were likewise found to be imputable to the presence of sulfured compounds and to elemental sulfur, with the corrosion rate varying from 5.3 mm per annum for iron sulfide to 1,100 mm per annum for sulfur.

Laboratory studies on the influence of ferrous iron on anaerobic mild steel corrosion by active SRB cultures hinted that a protective film was formed on the metal surface at low iron-ion concentration. However, the presence in solution of sufficiently high iron concentrations caused all of the bacterially produced sulfide to precipitate as iron sulfide, so that no protective layer was formed in the medium on the metal and the corrosion rate increased considerably (29). Electrochemical studies showed that substantial and energetic cathode depolarizations took place, even if the ferrous iron concentration was subsequently reduced.

In a previous study (30) associated with bacterial growth in a medium containing 0.5% ferrous iron, the corrosion rate of mild steel was shown to be an approximately linear function of time and independent of the hydrogenase activity of bacteria. The results of further tests clearly showed that the highest contribution to the corrosion process is provided by the ferrous sulfide mechanism. A comparison between mild steel corrosion by means of chemically prepared and biogenic ferrous sulfide indicated that at high ferrous iron concentrations most of the corrosion in bacterial cultures may be ascribed to biogenic ferrous sulfide. King et al. (31) showed that corrosiveness of chemically prepared ferrous sulfide is a stoichiometric function of sulfur, and that stimulation of corrosion due to iron sulfide decreases with time.

High corrosion rates in free bacterial systems may only be obtained by means of a continuous iron sulfide supply. The efficiency decrease of depolarization due to iron sulfide as a function of time without the presence of bacteria is ascribed to the hydrogen atomic bond of the iron sulfide structure. However, the iron sulfide activity may be renewed in the presence of SRB because of their capability of using the hydrogen contained in the film. In this way an explanation can be provided for this alternative mechanism, which is considered responsible for the high corrosion rate observed in the ground.

Under anaerobic conditions, too, the mechanism of formation of occluded areas on the metal surface is often associated with microbiologically induced corrosion. In connection with the development of occluded cells, microorganisms create conditions under the colony very different from those on the surrounding metal, leading to the formation of crevices (as shown in Fig. 7) as well as of ion concentration cells. This hypothesis is supported, in the presented case study, by energy dispersion spectrometer EDAX™ investigation revealing the presence of a series of deposits containing Fe, Cr, Ni, Mo, Ca, and S.

In parallel with the previously mentioned phenomenon, it may be noted that the microbiological community usually remains fixed to the colonization site, thus fixing the anodic site. The resulting pitting corrosion is characterized by the formation of small anodes and corresponding large cathodes, which allow the formation of pits in the metal matrix. Figure 8 shows incipient pitting triggered by microbiological activity on AISI 316. Incipient pitting engendered by microbiological activity on AISI 347 is shown in Figure 9; a detail of pits in the growing phase on the same metal is shown in Figure 10.

An example of intergranular corrosion, usually taking place on the border of uniform corroded areas, is reproduced in Figure 11. EDAX™ investigation (Fig. 12) on the selected corroded point revealed the presence of the following elements (% w/w): S: 1.58, Ca: 0.17, Cr: 34.7, Fe: 48.4, and Ni: 15.15. In the surrounding area the following elements were present (% w/w): Nb: 2.3, Cr: 14.7, Fe: 70.7, and Ni: 12.3.

DETECTION AND MONITORING OF MIC

In the recent past, the most important developments related to MIC research were provided by a multidisciplinary

Figure 7. AISI 304 (1,000×): crevice corrosion due to the formation of occluded area.

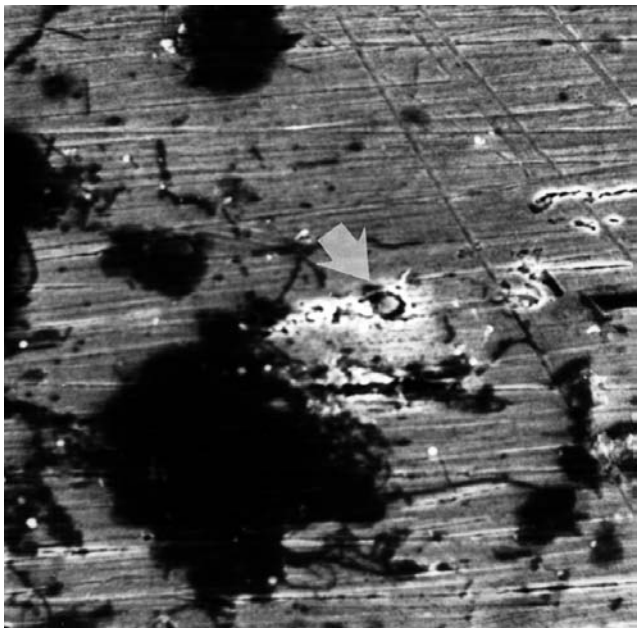


Figure 8. AISI 316 (1,000×): anaerobic incipient localized pitting corrosion.

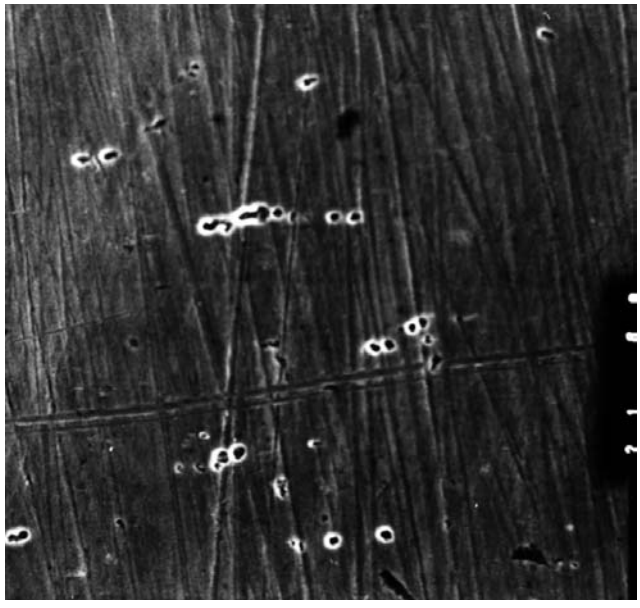


Figure 9. AISI 347 (1,000×): anaerobic incipient pitting corrosion.

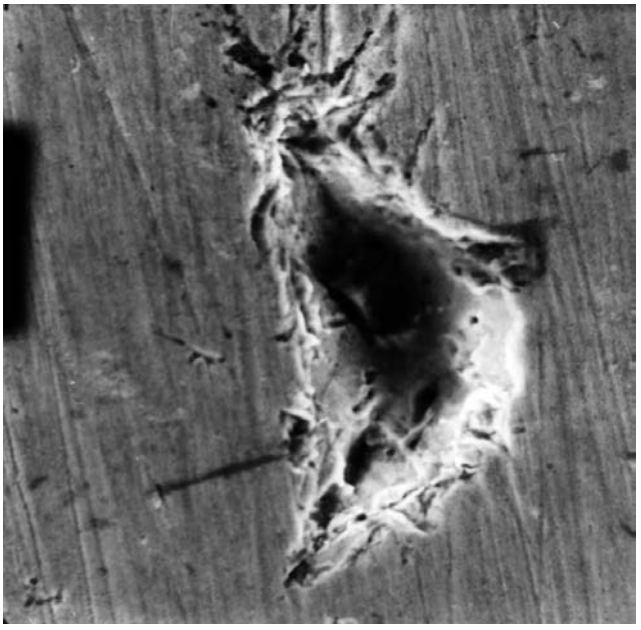


Figure 10. AISI 347 (1,500×): particular of anaerobic pitting in the growing phase.

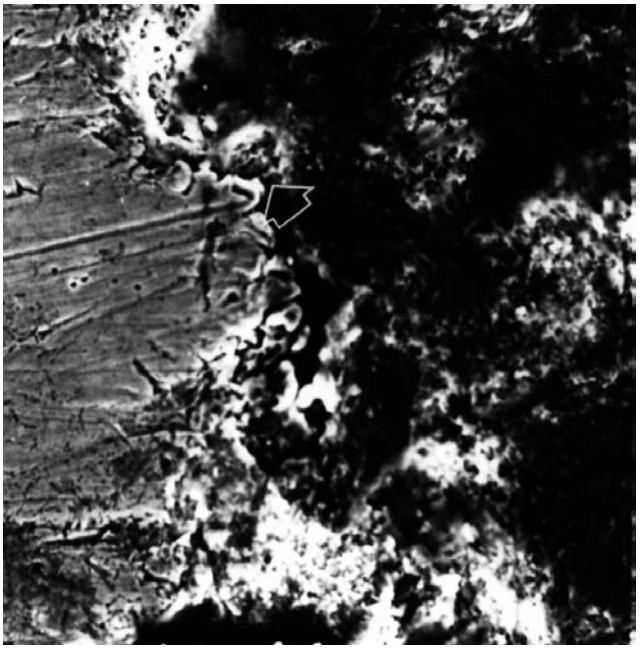


Figure 11. AISI 347 (1,000×): intergranular anaerobic corrosion on the border of uniform corroded area.

approach combining nondestructive analytical and electrochemical investigation methods. Indeed, appropriate techniques for detection and monitoring of the phenomenon are essential to dealing successfully with the nature of mutual interactions and to setting up effective control methods. Evidence of MIC lies in the conspicuous presence of microorganisms in those areas where the integrity of the material has deteriorated markedly. However, the electrochemistry of MIC has not been clearly defined because of the limits of the techniques employed in detection and monitoring (32). Some of the most important developments in MIC research over the past years were the result of a combination of nondestructive analytical and electrochemical methods aimed at characterizing microbial biofilms and corrosion reactions (33). In the following, the main traditional and innovative techniques are briefly presented and discussed.

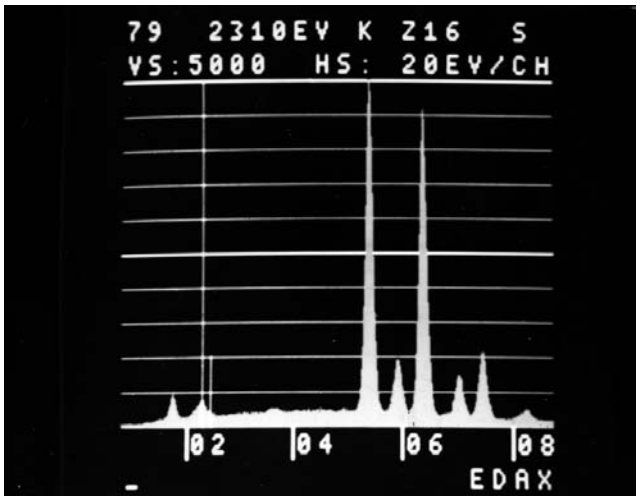


Figure 12. EDAX™ investigation in correspondence to the corroded point shown in Figure 11.

Microscopic Analysis

The identification of bacteria in pitted areas by means of scanning electron microscopy (SEM) was performed by several researchers (34–37), even though microbially induced pitting on stainless steel, the degree of interaction on this material being limited to a few surface monolayers, is not easily revealed. Moreover, it requires large amounts of corrosion products.

Energy-dispersive X-ray spectroscopy is widely applied in analyzing pit composition and is commonly employed for failure analysis of corroded metals. Recently developed (38) atomic force microscopy was applied to investigate sample topography in corroded pipework, allowing the analysis of profiles, depth, and size of pits, even in the presence of a nonconductive sample. In situ biofilm formation occurring under MIC conditions may be monitored by confocal laser microscopy (39); it allows the attainment of three-dimensional biofilm structure and has the potential of becoming an important method in the study of MIC kinetics.

Spectroscopic Analysis

Fourier transform infrared spectroscopy has been applied to monitor copper corrosion under biofilm (40,41), adopting Ge (an IR transparent internal reflection element) as substratum to be coated by a thin Cu film. The technique is severely limited by IR absorption by water, as well as by its inability to distinguish dead cells from live ones. X-ray photoelectron spectroscopy (XPS) and Auger electron spectroscopy (AES) may provide meaningful information about corroded surface composition and chemical states, as binding energies of the photoelectrons and Auger electrons are atom specific and changes in the binding energies of electrons are related to electron distribution around the atom. The limitations of these techniques, recently applied to study MIC (42–44), are connected with the kinetic energies of the Auger electrons and photoelectrons, which permit characterization of only few monolayers of the top surface, as well as with the need for cooling equipment to prevent bioproduct evaporation and decomposition under operative conditions (ultrahigh vacuum $<10^{-9}$ torr). XPS can provide composition and valence states of surface elements, as well as explore the mechanicistic nature of MIC over a relatively large area; Auger electron spectroscopy may be used to analyze smaller areas, obtaining chemical information of pits (31).

Extended X-ray absorption fine structure (EXAFS) may be employed to characterize enzymatic corrosion products created by bacterial consortium on metal surfaces (45,46), exploiting the oscillation of X-ray absorption spectrum in the range of 1 keV above the K-absorption edge (47). Limitations are related to in situ monitoring, as the X-ray energy may be too high.

Electrochemical Techniques

Conventional electrochemical techniques for detection and monitoring of MIC have been used for a long time and include corrosion and pitting potential measurements, linear polarization resistance (48), and electrochemical direct current (DC) methods (49,50); these techniques provide evidence of MIC and partial mechanistic information on loss of corrosion resistance and on electrochemical interactions (51), even if DC techniques require high applied currents and therefore can alter biofilms. Large polarization techniques have limited applications in MIC studies because of the destructive nature of applied currents to microbial biofilms. Consequently, electrochemistry of the biofilm–metal interface can be better investigated by means of techniques requiring small or zero applied potential.

Time-dependent changes in open circuit potential (OCP) were used to prove 316L stainless steel oxidation under a photosynthetic biofilm producing oxygen under illumination (52). Combined OCP and redox potential measurements of the bacterial medium permit accurate estimation of the corrosion tendency of a metal in microbially active environments (53), as well as of changes occurring in the passive films. Time-dependent changes of galvanic current between identical electrodes in dual cells demonstrated the effects of biofilms on induction of corrosion (54,55).

The concentric electrodes system, consisting of a small area anode surrounded by a large-area annular cathode, allowing maintenance of a uniform current profile with reduced overpotential, was successfully applied to monitor pitting commencement and development in SRB-containing anaerobic artificial seawater (56). Because the passive film has a blocking effect on anodic dissolution, this test apparatus, if employed with stainless steel, can result in overpotential to the anode so much higher than the redox potential as to impair bacterial growth.

The scanning electrode technique (SVET) has provided a nondestructive, on-line method to define the magnitude and sign of current densities in solution over freely corroding metals; in particular, it was applied to locate the presence and position of anodic electrochemical activity on the surface of carbon-steel samples by analysis of current maps and OCP (57), demonstrating the possibility of ob-

taining qualitative and quantitative information regarding localized corrosion influenced by bacteria.

Electrochemical impedance spectroscopy (EIS) was used to investigate the impedance of metal specimens exposed to microbial biofilms, in conjunction with biofilm analyses. It consists of the measurement of amplitude and phase shift of sinusoidal currents resulting from the application of small-amplitude sinusoidal signals to metal samples over a large frequency range (10 kHz to 5 mHz) (58). EIS can provide polarization resistance values (inversely proportional to corrosion rate), double-layer capacitance, and solution resistance so as to evaluate localized corrosion and three-dimensional surface deposits (51).

Future studies of MIC electrochemistry will focus on the combined use of different techniques, includings, confocal laser microscopy to characterize biofilm structure, EIS to investigate initiation of pits too small to be observed microscopically, and XPS and AES to analyze the surface chemistry and structure resulting from the corrosive phenomenon.

PROTECTIVE AND PREVENTIVE MEASURES

Replacement costs, maintenance shutdown, and loss of profits caused by microbiologically influenced corrosion of metal structures are problems in many industries. MIC, once recognized, is a question of optimizing prevention measures and minimizing maintenance and countermeasure efforts. Studies of MIC have addressed both understanding the complexities of microbial interactions with metal substrata and setting up protective and preventive measures. In particular, the involvement of microbial biofilms was amply investigated in the corrosion of several metals, including aluminium (58), copper (59), mild steel, and stainless steel (60). No universal approach to prevent MIC, has been developed so far, however, and the problem is tackled more from the corrosion than from the microbiological viewpoint.

Cathodic Protection

Cathodic protection principles and theory have been discussed extensively in the literature (61,62). Two techniques are commonly adopted: sacrificial anodes usually made of magnesium, zinc, or aluminium alloys, which corrode preferentially to provide structure protection, or impressed current from a rectifier via an inert anode. The cathodic protection method is used extensively to protect structures buried in soil or in marine environments.

In the absence of SRB, protection of steel is achieved at a potential depression of -0.85 V versus $Cu/CuSO_4$. In the presence of a soil environment presenting risk of microbial corrosion, a further -0.1-V potential depression is required to ensure the protection of the structure (63,64). The achievement of this value allows the suppression of SRB growth as well, even if, in the case of bare steel in direct contact with soil, the formation of iron sulfide can lead to extraordinary current requirements. On the other hand, far lower potential values must be avoided, because hydrogen cracking and alkaline blistering of protective coatings can occur.

In the presence of cell suspensions of sulfate reducers, a linear relationship between the hydrogenase-enzyme microbial activity and the additional current requirement to maintain different cathodic potential was verified (63). A focal conclusion is that the corrosion rate is enhanced in the case of an incompletely protected sample because of the accelerating effect of bacteria (65).

Protective Coatings

Protective coatings provide a barrier, impermeable to soluble aggressive species, between the metal and the surrounding environment. The range of available protective coatings is wide, and includes products based on coal tar, asphalt bitumens, epoxy resins and coal tar epoxies, thermoplastic polymers, synthetic polymeric coatings, and cementitious materials. Paints are also used as coating materials. Many of these coatings, being of organic nature, may biodeteriorate. Moreover, they usually cannot prevent gas diffusion and, for instance, allow hydrogen sulfide to pass the coating layer and reach the metal.

In marine environments, metallic coatings, especially aluminium and zinc, are often employed. Protection of steel by these metals is partly due to both their action as physical barriers and their anodic effect on the basic material.

In terms of industrial applications, when contamination connected with corrosion products must absolutely be avoided (e.g., in the food and pharmaceutical industries), the inner sides of plant sections or apparatuses should not be lined with coatings or resins (e.g., epoxy resins). In fact, the resulting unevenness (pores, microcrevices, blisters) might cause an extremely nonuniform, seriously corrosive process under the coating, due to an unfavorable relation between anodic and cathodic zones (66). Figure 13, referring to a FE33 steel sample from a reactor employed in yeast-biomass production, clearly demonstrates how crevice or surface unevenness of the protective paint leads to corrosion under the coating—a process detectable only under an optical microscope.

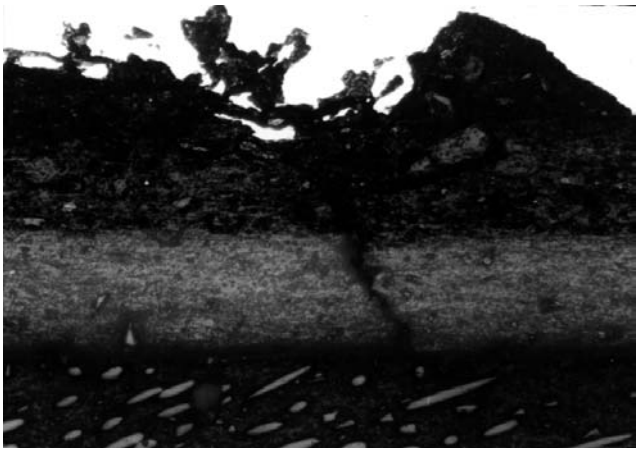


Figure 13. FE33 painted (300×): start of aerobic MIC below the paint due to protective coating unevenness.

Inhibitors and Biocide

In closed systems it is possible to use corrosion inhibitors and biocides to stop or limit the biological activity. The use of such chemicals has limitations because biocides are potential hazards to health and environment. Organic biocides can take part in complex and uncontrolled reactions, and completely unexpected compounds can be formed. Oxidizing biocides, for example, hydrogen peroxide, which cause death of cells by attacking cell enzymes, can be used in cooling systems and drinking water mains. Oxygen-based biocides (e.g., ozone) are used to prevent bacterial hydrogen-sulfide production. Nonoxidizing biocides, which cause cell-wall rupture, blockage of enzyme systems, or denaturation of proteins, include heavy-metal organic compounds, organic copper salts, organic sulfur compounds, amines, chlorinated phenols, and quaternary ammonium salts. The treatment is ineffective if the bacteria have previously formed nodules as their habitat, so in this case a preventive mechanical removal of nodules is necessary.

Careful and well-documentated observations by plant operating engineers have recognized the efficacy of regular scraping in pipeline corrosion control; experimental runs showed that well-established consortia of SRB are deeply disturbed by the mechanical treatment, with reestablishment only after 1–2 weeks (67). On the basis of these remarks, microbial corrosion in pipelines can be effectively prevented or controlled by combining mechanical scraping to remove protective nodules with proper dose treatment using a penetrating biocide (68).

Remedial action in the case of SRB-induced corrosion is usually approached by a combined action consisting of the application of a biocide and a corrosion inhibitor. Compounds such as molybdate show a double effect, inhibiting both SRB and the corrosion (69). It was observed (70) that 18 mM 2-*n*-Heptyl-4-hydroxyquinoline *N*-oxide inhibited respiration in *Hemophilus parainfluenza* by blocking electron transport between cytochromes b and c_1. Inhibition of the periplasmic hydrogenase enzyme in *Desulfovibrio vulgaris* was verified by both carbon monoxide and copper II chloride (71). In case of *Desulfovibrio desulphuricans*, characterized by a cytoplasmatic hydrogenase, only carbon monoxide was found effective in inhibiting the hydrogenase, and copper II chloride had no effect (72). Sodium molybdate evidenced inhibiting action on stainless steel corrosion by making passive the anodic reaction of an active corrosion cell and inhibition of the growth of a culture of *D. vulgaris* and unidentified *Vibrio* sp. (73). The main conclusion is that viable bacteria are necessary only in the early stages of pit formation, and not for the maintenance of galvanic current; in any case, the effective treatment of actively pitting systems requires the adoption of both a corrosion inhibitor and a biocide.

Alternative Materials

Solution of the problem of MIC based on safety considerations (74) may consist of the application of the substitution principle, in which a safer material is used in place of a more hazardous one, thus decreasing the need for added protective equipment and also decreasing plant cost and complexity. The use of common steel should be avoided in the construction of industrial plants for bioleaching of mine drainage. Biologic desulfurization of pit coal and the use of stainless steel, such as AISI 304, is compulsory to avoid safety problems connected with possible metallic structure failures (75). A number of resistant alloys providing adequate corrosion strength are available (e.g., all experimental runs on higher nickel–chromium alloys and titanium highlight their resistance to MIC), although their cost makes their use uneconomical in most situations. Steel and iron will be replaced extensively by nonmetallic materials, such as composite fiber reinforced resin systems, due to the availability of reliable and long-lasting materials. Other plastic materials, such as PVC and polyethylene, must be used with care, because they may contain plasticizers susceptible to biodeterioration, resulting in progressive embrittlement of the polymer. Cement products, including spun concrete and reinforced or prestressed concrete, are widely used for buried piping; however, their selection must be site specific, because they may have problems in some heavily contaminated soils and waters (63,76).

CONCLUSIONS

Microbiological corrosion is an interesting but costly and frustrating problem more widespread than one would expect, and it is often associated with the more common problem of biofouling. Microorganisms play a very complex role, and several mechanisms were suggested to highlight their involvement in the corrosion process. In many instances their role consists of the formation of differential aeration and concentration cells, in the production of aggressive metabolites, or in the acceleration of an existing corrosion process. Microbial involvement in the process can be indicated by localized corrosion associated with tubercles, discolorations, slime, anaerobic metabolism odors, sludges, and underdeposit corrosion phenomena.

The majority of these problems can be ascribed to SRB. The knowledge of aspects of SRB physiology and of their ability to grow on hydrogen, as well as the role of enzyme hydrogenase, has shown the most probable mechanism of corrosion due to these organisms.

From a practical industrial viewpoint, avoidance or minimization of microbial corrosion problems can be faced first by optimal selection of materials and, subsequently, by interposition of barrier systems or modification to the surrounding environment. The combined use of chemical inhibitors and biocides represents the most common solution to dealing with internal problems in process systems.

MIC problems will be positively contained and safely controlled in industrial situations only if designers and plant operators properly consider projects, design features, operating and maintenance procedures, and system monitoring.

BIBLIOGRAPHY

1. R. Mallet, *Br. Assoc. Adv. Sci. Rpt.* **9**, 221 (1940).
2. J.H. Garrett, *The Action of Water on Lead*, H.K. Lewwis, London, 1891.

3. R.H. Gaines, *J. Eng. Ind. Chem.* **2**, 128 (1910).
4. E.C. Harder, *Iron Depositing Bacteria and Their Geologic Relations*, Governmental Printing Office, Washington, D.C., 1919.
5. R. Grant, E. Bate, and W.H. Myers, Inst. Eng. Australia Sydney Div., Paper no. 8, 1921.
6. C.A.H. von Wolzogen Kuhr and L.S. van der Vlugt, *Water* (den Haag) **18**, 147 (1934).
7. H.H. Uhlig, *Corrosion Handbook*, 4th ed., Wiley, New York, 1953.
8. J.D.A. Miller ed., *Microbial aspects of Metallurgy*, MTP, Aylesbury, 1971.
9. G.H. Booth, *Microbial Corrosion*. M and B Monograph CE/1, Mills & Boon, London, 1971.
10. W.P. Iverson, in M.G. Fontana and R.W. Staehle, eds., *Advances in Corrosion Science and Technology*, vol. 2, Plenum Press, London, 1972, pp. 70–94.
11. J.D.A. Miller and R.A. King, in D.W. Lovelock and R.J. Gilbert, eds., *Microbial Aspects of the Deterioration of Material*, Academic Press, London, 1975, p. 83.
12. E.C. Pringsheim, *Biol. Rev. Cambridge Phil. Soc.* **24** 200 (1949).
13. J.S. Stokes, *J. Bact.* **67**, 278 (1954).
14. J.D.A. Miller and A.K. Tiller, in J.D.A. Miller, ed., *Microbial Aspects of Metallurgy*, MTP, Aylesbury, 1971, p. 61.
15. L.A. Terry and G.J. Edyvean, *Bot. Mar.* **24**, 177–183 (1981).
16. H. Ebneth, D. Rempel, and W. Hoyer, *Galvanotechnik* **63**, 27 (1972).
17. W.P. Iverson, *Science* **151**, 986 (1966).
18. G.H. Booth, A.W. Cooper, and P.M. Cooper, *Chem. Ind.* **86**, 2084 (1967).
19. D.J. Crombie, G.J. Moody, and J.D.R. Thomas, *Chem Ind.* **21**, 500 (1980).
20. J. Chantereau, *Corrosion Bacterienne. Bactéries de la Corrosion*, 2nd edition, Technique et Documentation ed., Paris, 1980.
21. G. Rossi, *Resoconti dell' Associazione Mineraria Sarda* **79**, 1–21 (1974).
22. G.H. Booth, *Microbial Corrosion*, MandB Monograph CE/1, Mills & Boon, London, 1971.
23. B.E. Purkiss, in J.D.A. Miller, ed., *Aspects of Metallurgy*, MTP, Aylesbury, 1971, p. 107.
24. J. Horvath, *Acta Chem. Acad. Sci.* **25**, 65 (1960).
25. G.H. Booth and F. Wormwell, in *1st Proc. Int. Congr. Metallic Corrosion*, Butterworth, London, 1962, p. 341.
26. G.H. Booth, L. Elford, and D.S. Wakerely, *Br. Corr. J.* **3**, 242 (1968).
27. R.A. King and J.D.A. Miller, *Nature* **233**, 491 (1971).
28. G.H. Booth and A.K. Tiller, *Trans Faraday Soc.* **58**, 2510 (1962).
29. R.A. King, J.D.A. Miller, and D.S. Wakerley, *Br. Corr. J.* **8**, 89 (1973).
30. G.H. Booth, P.M. Cooper, and D.S. Warkerley, *Br. Corr. J.* **1**, 345 (1966).
31. R.A. King, J.D.A. Miller, and J.S. Smith, *Br. Corr. J.* **8**, 137 (1973).
32. G. Chen, R.J. Palmer, and D.C. White, *Biodegradation* **8**, 189–200 (1997).
33. M.J. Franklin and D.C. White, *Current Opinion in Biotechnology* **2**, 450–456 (1991).
34. D.E. Nivens, P.D. Nichols, J.M. Henson, G.G. Geesey, and D.C. White, *Corrosion* **42**, 204–210 (1986).
35. T.R. Jack, G. Van Boven, and M. Wilmott, *Material Performance* **35**, 39–45 (1996).
36. W. Lee and W.G. Charaklis, *Corrosion* **49**, 186–99 (1992).
37. B.J. Little, P. Wagner, S.M. Gerchakov, M. Walch, and R. Mitchell, *Corrosion*, **42**, 533 (1986).
38. A. Steele, D.T. Goddard, and I.B. Beech, in *International Biodeterioration and Biodegradation*, Elsevier, London, 1994, pp. 35–46.
39. J.R. Lawrence, G.M. Wolfaardt, and D.R. Korber, *Appl. Environ. Microbiol.* **60**, 1166–1173 (1994).
40. P.J. Bremer and G.G. Geesey, *Appl. Environ. Microbiol.* **57**, 1956–1962 (1991).
41. G.G. Geesey and P.J. Bremer, *Mar. Technol. Soc. J.* **24**, 36–43 (1990).
42. P.J. Bremer, G.G. Geesey, B. Drake, J.G. Jolley, and M.R. Hankins, *Surface Interface Anal.* **17**, 767–772 (1991).
43. G. Chen, R.A. Sadowski, and C.R. Clayton, in *CORROSION/95*, paper 217, NACE National Association of Corrosion Engineers, Houston, Tex., 1995.
44. G. Chen, S. Kagwade, G.E. French, C.R. Clayton, T.E. Ford, and R. Mitchell, *Corrosion* **52**, 891–899 (1996).
45. A.J. Davenport, J.A. Bardwell, and C.M. Vitus, *J. Electrochem. Soc.* **142**, 721–724 (1995).
46. A.J. Davenport and M. Sansone, *J. Electrochem. Soc.* **142**, 725–730 (1995).
47. L.C. Feldman and J.W. Mayer, in *Fundamentals of Surface and Thin Film Analysis*, North-Holland, New York, 1986, pp. 201–204.
48. J.S. Luo, X. Campaignolle, D.C. White, and I. Vance, in *CORROSION/93*, paper 301, NACE National Association of Corrosion Engineers, Houston, Tex., 1993.
49. S.C. Dexter, D.J. Duquette, O.W. Siebert, and H.A. Videla, *Corrosion* **47**, 308, (1991).
50. B.J. Little, P. Wagner, and F. Mansfeld, *Electrochimica Acta* **37**, 2185–2194 (1991).
51. F. Mansfeld and B.J. Little, *Corrosion Sci.* **32**, 247–272 (1991).
52. N.J.E. Dowling, J. Guezennec, J. Bullen, B.J. Little, and D.C. White, *Biofouling* **5**, 315–322 (1992).
53. X. Zhang, R.A. Buchanan, E.E. Stansbury, and N.J.E. Dowling, in *CORROSION/89*, paper 512, NACE National Association of Corrosion Engineers, Houston, Tex., 1989.
54. N.J.E. Dowling, J. Guezennec, M.L. Lemoine, A. Tunlid, and D.C. White, *Corrosion* **44**, 869–874 (1988).
55. S.M. Gerchakov, B.J. Little, and P. Wagner, *Corrosion* **42**, 689 (1986).
56. P. Angell, J.-S. Luo, and D.C. White, *Corrosion Sci.* **37**, 1085–1096 (1995).
57. M.J. Franklin, D.C. White, and H.S. Isaacs, *Corrosion Sci.* **32**, 945–952 (1991).
58. E.S. Ayllon and B.M. Rosales, *Corrosion* **44**, 638–644 (1988).
59. P. Angell, Ph.D., Thesis, University of Surrey, Surrey, OK, 1992.
60. B. Little, P. Wagner, and F. Mansfeld, *Int. Mat. Rev.* **36**, 253–272 (1991).
61. J.H. Morgan, *Cathodic Protection*, Leonard Hill, London, 1959.
62. L.L. Shrier ed, *Corrosion*, vol. 1, 2nd ed., Newnes Butterworth, London, 1976.
63. G.H. Booth and A.K. Tiller, *Corrosion Sci.* **8**, 583 (1968).
64. J. Horvath and M. Novak, *Corrosion Sci.* **8**, 538 (1968).

65. J. Guezennec, N.J.E. Dowling, J. Bullen, and D.C. White, *Biofouling* **8**, 133–136 (1994).

66. P. Perego, B. Fabiano, A. Ferraiolo, and R. Pastorino, *Chem. Biochem. Eng. Q.* **11**, 101–106 (1997).

67. J.W. Costerton and G.G. Geesey, in *Biologically Induced Corrosion NACE 8*, NACE National Association of Corrosion Engineers, Houston, Tex., 1986, pp. 223–232.

68. G.N. Kirby, *Chem. Eng. Prog.* **6**, 29–38 (1995).

69. J.S. Luo, X. Campaignolle, J. Bullen, M.W. Mittleman, D.C. White, and J.F. Zibrida, in *CORROSION/92*, paper 186, NACE National Association of Corrosion Engineers, Houston, Tex., 1992.

70. D.C. White and L. Smith, *J. Biol. Chem.* **239**, 3956–3963 (1964).

71. R.M. Fitz and H. Cypionka, *Arch. Microbiol.* **155**, 444–448 (1991).

72. R.M. Fitz and H. Cypionka, *Arch. Microbiol.* **152**, 369–376 (1989).

73. P. Angell and D.C. White, *J. Ind. Microbiol.* **15**, 329–332 (1995).

74. T. Kletz, *Plant Design for Safety. A User-Friendly Approach*, Hemisphere Publishing, New York, 1991, pp. 53–66.

75. P. Perego, B. Fabiano, R. Pastorino, and G. Randi, *Bioprocess Eng.* **17**, 103–109 (1997).

76. P. Perego, Ph.D. Thesis, University of Milan, Milan, Italy, 1986.

CORYNEBACTERIA, BREVIBACTERIA

P. A. LESSARD
S. GUILLOUET
L. B. WILLIS
A. J. SINSKEY
Massachusetts Institute of Technology
Cambridge, Massachusetts

KEY WORDS

Corynebacterium
Brevibacterium
Fermentation
Glutamic acid
Lysine
Threonine
Isoleucine
Biochemical engineering
Metabolic engineering

OUTLINE

INTRODUCTION

Corynebacteria and brevibacteria are Gram-positive, aerobic actinomycetes. The primary industrial interest in corynebacteria and brevibacteria relates to the production of amino acids by fermentation. In 1956, a researcher at Kyowa Hakko Kogyo Co. in Japan isolated soil bacteria capable of producing large amounts of glutamic acid (1). Until that time, glutamic acid was produced commercially from *konbu* a type of seaweed used in popular Japanese cuisine. The discovery of these bacteria proved to be a major innovation in the industry at the time, and today these bacteria account for worldwide production of more than 800,000 metric tons of glutamic acid (primarily for use as a flavor enhancer) and 330,000 metric tons of lysine (for feed additives) annually.

SPECIES RELATIONSHIPS AMONG GLUTAMIC ACID–PRODUCING CORYNEBACTERIA

When first identified, the species we now call *Corynebacterium glutamicum* was classified as *Micrococcus glutamicus*. Several "species" of glutamic acid–producing bacteria were similarly identified by a number of groups, and classified into genera such as *Brevibacterium*, *Corynebacterium*, *Arthrobacter*, and *Microbacterium*. During the subsequent four decades, considerable effort has been spent on reconciling the phylogeny of these bacteria. Nonetheless, the scientific literature regarding these bacteria is still laden with inconsistencies in the nomenclature, and an opportunity exists to establish conventions for species designation among these bacteria.

Systematics among the glutamic acid–producing bacteria have evolved from the classical morphological and physiological analyses to include molecular genetic characterizations. It can be argued that morphological analyses caused much of the confusion encountered in the literature because corynebacteria and brevibacteria belong to a group of bacteria, frequently called the "coryneform" bacteria, whose common characteristic appears to be pleomorphism. Abe and coworkers (2) identified several characteristics that are useful for distinguishing the corynebacteria and brevibacteria from sister clades. In addition to their prolific glutamate production under aerobic conditions, they require biotin, and individual bacteria are

nonmotile, nonsporulating, and occur as short clubs or rods in V-shaped aggregates of two or more. Chemotaxonomic information has permitted a more methodical classification of these bacteria (3). Based on the presence of relatively short-chain mycolic acids (containing 22–36 carbon atoms) in the cell walls, and the types of lipids that comprise their cell membranes, these bacteria cluster with members of the genera *Rhodococcus, Mycobacterium*, and *Nocardia* into an informal collection known as the CMN group, which includes many medically significant pathogens.

In recent years, molecular systematics have been used to more precisely define the glutamic acid–producing bacteria. Chromosomal DNA of the glutamic acid–producing bacteria comprises approximately 52–56 mol % G + C. Comparison of 16S and 23S ribosomal RNA structures has revealed that many of the glutamic acid–producing species that have traditionally been classified into the different genera *Corynebacterium* and *Brevibacterium* (e.g., *B. divaricatum, B. flavum, B. lactofermentum, C. glutamicum*, and *C. lilium*) may all actually be members of the same species, *Corynebacterium glutamicum* (4). In the same studies, other species within the genus *Brevibacterium* (e.g., *B. linens*) were found to be much more distantly related, containing a higher G + C content in their genomic DNA, showing greater divergence among their ribosomal RNA structures, and lacking the mycolic acids that are characteristic of the CMN group. For this reason, in this article we shall refer to all of the glutamic acid–producing bacteria simply as corynebacteria, with emphasis on the well-characterized *C. glutamicum*.

COMMERCIAL INTERESTS IN CORYNEBACTERIA

The main industrial interest in corynebacteria concerns the production of amino acids. The twenty amino acids can be commercially manufacture and are used in a variety of products including food additives, animal feed supplements, cosmetics, and medicines, as well as some industrial applications. Currently, the highest revenues are from the animal feed industries and human food uses. The overall global market for amino acid animal feed additives is estimated to be worth more than $2 billion, and totals about 700,000 metric tons of material. Lysine and methionine account for an overwhelming majority of the market, which also includes lower-volume products such as threonine (Table 1). The market for monosodium glutamate alone, used in the human food industry as a flavor enhancer, is also estimated to be $2 billion. The worldwide market growth rate is around 8–12% annually for amino acids (5), but individual amino acids could experience different spurts of growth. For example, the price of cysteine doubled during 1996 due to increased demand that sparked a worldwide shortage (6). In 1990, the world production of threonine was 1,000 metric tons per year, but already new projects will give BioKyowa, for example, a total capacity of 10,000 metric tons annually for tryptophan and threonine in 1998 (7).

In 1990, aspartic acid was produced at around 10,000 metric tons per year, mainly for the production of aspartame. Most of the world's aspartic acid still goes into aspartame, but Donlar Corp. and Rohm & Haas believe the emerging application for polyaspartate, a newer specialty polymer with a wide range of applications, will drive new production of aspartic acid (8).

Table 1. Industrial Production of Amino Acids

Amino acid	Worldwide production (ton/year)	Production process
L-Glutamic acid	>800,000	Production by corynebacteria
L-Lysine	330,000	Production by corynebacteria
DL-Methionine	350,000	Chemical synthesis
L-Aspartic acid	10,000[a]	Bioconversion by *E. coli*
L-Phenylalanine	5,000[a]	Production by corynebacteria
L-Threonine	1,000[a]	Production by corynebacteria and *E. coli*

[a]Production data from 1990; all other data from 1997.

Amino acids are produced industrially by four different processes: production in reactors by bacteria (glutamate, lysine, phenylalanine, threonine, tryptophan, valine), bioconversion (aspartic acid), chemical synthesis (methionine), and extraction from natural materials (cysteine). Corynebacteria in particular are involved in the production of the two most important amino acids, glutamic acid and lysine (Table 1).

Corynebacteria are able to accumulate 8–10 times as much intracellular amino acid, such as glutamic acid, as accumulated by Gram-negative bacteria such as *Escherichia coli* (9). In addition, the amino acid biosynthetic pathways in corynebacteria are usually less strongly regulated than those in *E. coli*, for instance (Fig. 1). *E. coli* has three isoenzymes for aspartokinase and as many as five for acetohydroxy acid synthase (AHAS). The abundance of isoenzymes complicates the prospects of using genetic engineering to improve amino acid production in *E. coli*. These observations explain why corynebacteria are preferred for the production of the major amino acids.

Other industrial applications for corynebacteria concern the production of vaccines from *Corynebacterium tetani* and *C. diphtheriae* (10), the production of L-malic acid from fumaric acid by the fumarase of immobilized *Brevibacterium ammoniagenes* (11) and the production of L-ascorbic acid by 2,5-diketo-D-gluconic acid reductase of *Corynebacterium* sp. (11). Corynebacteria may also be used to produce recombinant proteins such as gamma interferon, proteases, and cellulase (12).

Corynebacteria, and specifically *Brevibacterium linens*, are involved in the cheese ripening process. Although the lactic acid bacteria play a key role in cheesemaking by catabolizing nitrogen compounds, the roles of other bacteria, such as brevibacteria, are also important. Possessing powerful and varied enzymatic systems, they contribute to the aging of ripened cheeses by producing unique flavors (13). Flavored compounds produced by these bacteria, including peptides, amino acids, and sulfured compounds, have found applications in the food industry as additives in cheese-flavored products (sauces, crackers, etc.) (14).

Not surprisingly, the *Corynebacterium* strains that are used in industrial fermentations are not the original iso-

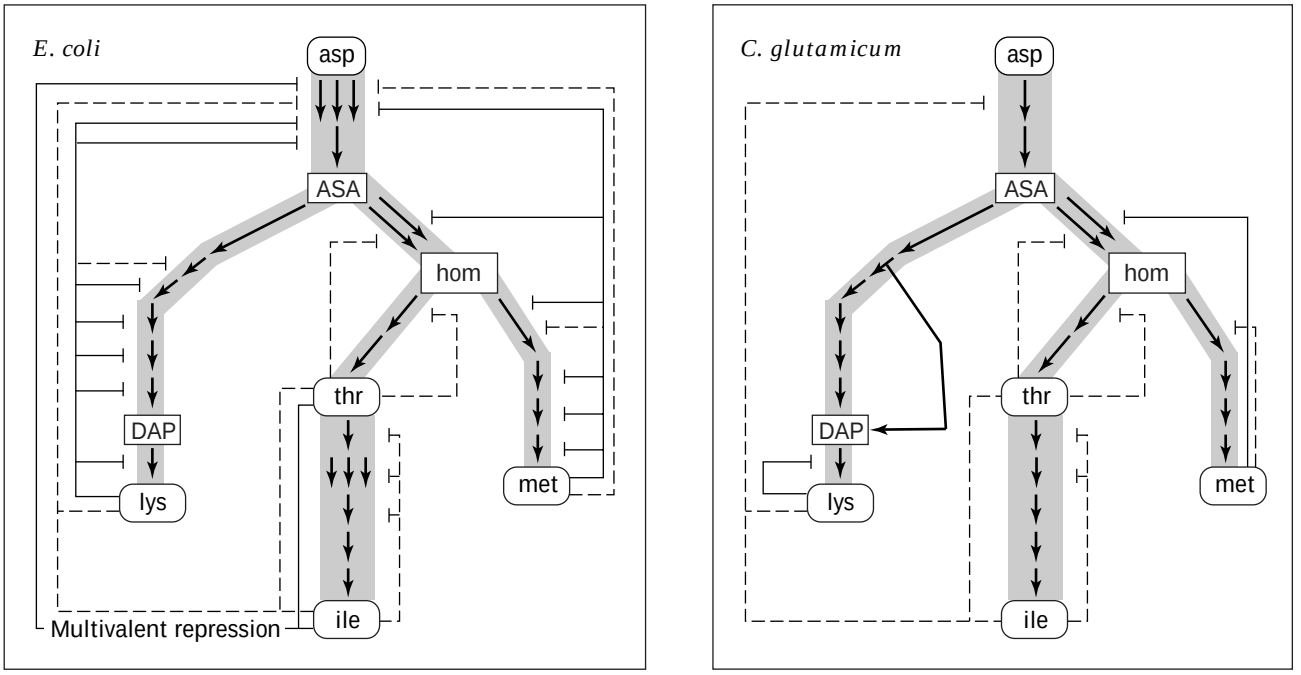


Figure 1. Summary of feedback inhibition and genetic repression operating in pathways for synthesis of aspartate derived amino acids in *E. coli* and *Corynebacterium*. Dashed lines denote feedback inhibition, solid lines denote genetic repression. Arrows indicate individual, enzymatically catalyzed steps. Steps that are catalyzed by multiple isozymes in *E. coli* are indicated by more than one arrow. Note that control of these pathways in *E. coli* is much more complex than that known in *Corynebacterium*. *E. coli* lacks the parallel pathways leading to lysine biosynthesis (described later in the text). More detailed depictions of the enzymes and intermediates involved in these pathways are depicted later in this chapter. ASA, aspartate semialdehyde; hom, homoserine; DAP, *meso*-diaminopimelate.

lates that were identified for their prolific amino acid production. A great deal of research has focused on improving the properties of the production strains. Glutamate excretion from the cells appears to be the major limitation to yield in glutamate production. Because excretion can be enhanced by manipulating the culture conditions, biochemical engineering and reactor design have been the primary routes toward improving glutamate yield. In lysine fermentations, however, the primary bottleneck appears to be in the rate of synthesis of the amino acid. For this reason, improvements in lysine production have come mainly from metabolic engineering. Following, we discuss both aspects, describing first biochemical engineering of the glutamate production process, and second, metabolic engineering of the pathways leading to lysine and other aspartate-derived amino acids.

BIOCHEMICAL ENGINEERING FOR IMPROVED AMINO ACID PRODUCTION

General Characteristics of Industrial Amino Acid Fermentations

Corynebacteria are mainly involved in industrial production of amino acids by fermentation processes. The volume of industrial fermentors for amino acid production may reach up to 100 m^3 in order to support the high demand of the market. Some production facilities produce up to 100,000 metric tons per year of amino acids. These large cultures are usually realized in fed-batch mode, in order to reach the highest productivities. Although continuous cultures also lead to high productivities, they are rarely used in industry due to the higher risks of bacterial contamination. The industrial raw materials are complex media such as cane or beet molasses, corn steep liquor, or starch hydrolysates. The carbon sources are thus mainly sucrose (from beet molasses), or sucrose, glucose, and fructose (from cane molasses). The nitrogen sources are ammonium salts provided by the medium in the first stage of the cultures, and later during the fermentation as ammonium hydroxide, which is used to regulate pH in the reactor.

Glutamic Acid Fermentation

Glutamic acid fermentation provides a good example of the development of a process in order to answer industrial problems. Glutamic acid is produced industrially by fed-batch fermentations of 25–40 h. The final concentration of glutamic acid ranges between 90 and 120 g/L with a biomass production of 15–25 g/L. Depending on the raw materials used, the conversion yields vary between 50 and 60%. Glutamic acid production is an aerobic process. Therefore it is very important to control the dissolved ox-

ygen at its optimal level. Oxygen limitation leads to a drastic decrease in glutamic acid titer, and to significant production of succinic and lactic acids. Moreover, an ammonium limitation with sufficient oxygen supply results in a deviation of the fermentation process to a production of 2-oxoglutarate. In industrial-scale processes, automatic addition of gaseous ammonia avoids ammonium limitation and enables the pH regulation of the broth without the harmful effects of liquid ammonia and the undesirable dilution of the fermentation liquid.

Corynebacteria usually excrete little or no glutamic acid into the medium. Therefore, different processes were developed to increase the excretion of this amino acid. The first industrial process was based on biotin limitation. Takinami et al. (15) reported that in media with high biotin concentrations, *B. lactofermentum* did not produce glutamic acid. But when the biotin concentration was low (<3 μg/L), glutamic acid was excreted in the external medium (Fig. 2). Glutamic acid fermentation by biotin limitation occurs in two phases (Fig. 3). Cultures begin with a relatively high concentration of biotin to permit cell growth. When the biotin becomes limiting, bacterial growth slows down and glutamic acid is excreted. Using biotin limitation to enhance glutamate, industrial production has given glutamate titers up to 90 g/L.

The process of biotin limitation introduces different problems on an industrial scale: it does not allow the use of complex media that often contain high biotin concentrations, and it necessitates a precise and tedious monitoring of the biotin concentration. These problems were partly solved by the use of antibiotics or surfactants. The addition of sublethal concentrations of penicillin to a culture of *C. glutamicum* led to excretion of glutamic acid (16). This process was used in industry but was quickly replaced by a process involving surfactants. Addition of either penicillin or surfactants leads to an excretion of glutamic acid. Current modern production processes are fed-batch fermentations consisting of an initial period of rapid growth of the bacteria followed by a phase of extensive glutamate excretion, which is induced by the addition of surfactants such as amine surfactant (17) or Tween (18) (Fig. 4). Glutamate production via surfactant addition may reach titers of 120 g/L in industrial media.

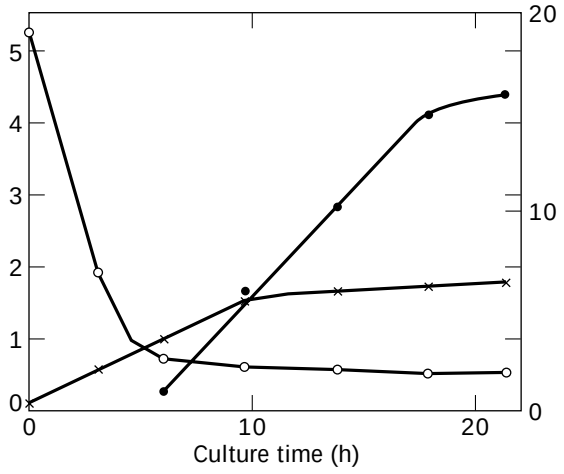


Figure 3. Production of glutamic acid by biotin limitation process. Closed circles, glutamic acid accumulated (g/L); x, dry cell concentration (g/L). *Source*: Adapted from Ref. 15.

The development of the fermentations that rely on penicillin or surfactants led to benefits for the amino acid industries. These processes have the main advantage of allowing the use of cheap, biotin-rich materials, such as molasses. Nevertheless, the monitoring of these fermentations remains delicate due to the sensitivity of the bacteria to surfactants. The ratio of surfactant to biomass concentration, and the physiological state of the cells when the surfactants are added, are critical determinants for production of high levels of glutamic acid (19).

The high levels of glutamic acid reached in industrial processes are usually explained by the fact that the excretion of glutamate leads to a reduction in the internal pool

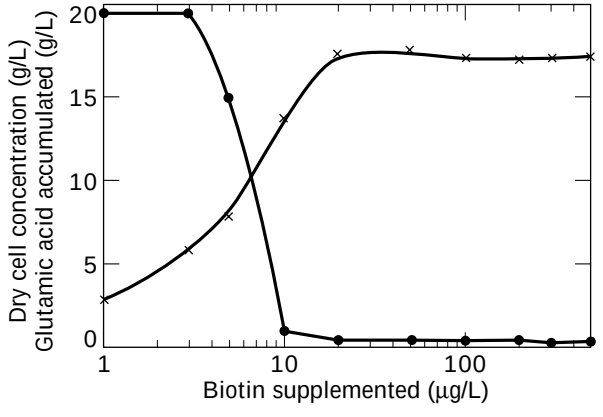


Figure 2. Relationship between the level of external biotin and the accumulation of glutamic acid in *Brevibacterium lactofermentum*. Closed circles, glutamic acid; open circles, dry cell concentration. *Source*: Adapted from Ref. 15.

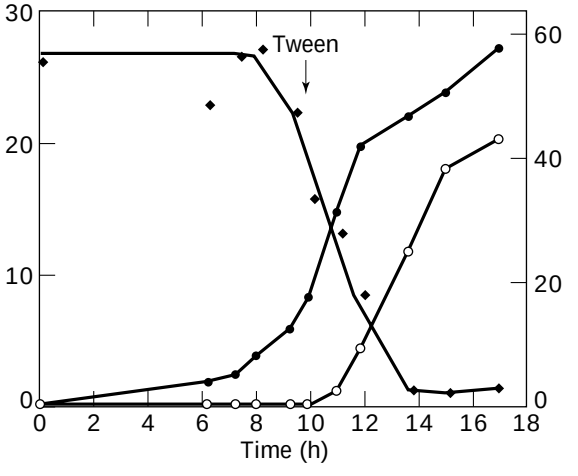


Figure 4. Production of glutamate by fed-batch culture of *Corynebacterium glutamicum* with addition of Tween. Closed diamond, glucose (g/L); open circles, glutamate (g/L). *Source*: Adapted from Ref. 18.

of this amino acid, which in turn relieves feedback control mechanisms. Regarding the regulatory aspects, 2-oxoglutarate dehydrogenase, which converts 2-oxoglutarate to succinyl-CoA at the branch point where glutamate formation diverges from the TCA cycle, is a key enzyme for glutamic acid biosynthesis (Fig. 5). The properties of the enzyme must be favorable for deviation of the carbon flux towards the glutamic acid pathway in order to prevent further oxidation through the TCA cycle. It has been shown that in overproducing strains, the V_{max} of the glutamate dehydrogenase is 150 times higher than that of 2-oxoglutarate dehydrogenase. Also, the intracellular 2-oxoglutarate concentration is high enough to direct the flux to glutamic acid biosynthesis, even with a 70-fold lower K_m value of 2-oxoglutarate dehydrogenase for 2-oxoglutarate (20). Recently, Kawahara et al. (21) showed that the activity of 2-oxoglutarate dehydrogenase was reduced in cells cultured under glutamate-production conditions (biotin limitation and addition of penicillin or surfactant) as compared to those in glutamate nonproductive conditions. In contrast, glutamate dehydrogenase activity, which draws carbon away from the TCA cycle into glutamate, was unchanged under these conditions. Thus, a high flow of carbon to the TCA cycle could be easily converted to glutamate after exhaustion of biotin, or addition of antibiotic or surfactants.

The second aspect of regulatory mechanisms of glutamic acid biosynthesis concerns the carbon dioxide fixation to form oxaloacetate. The efficiency of glutamic acid biosynthesis depends on the supply of oxaloacetate to the first part of the TCA cycle (Fig. 5). Anaplerotic pathways, such as phosphoenolpyruvate carboxylase and pyruvate carboxylase, may then play an important role in the production of glutamic acid. In the future, increases in carbon dioxide fixation might improve the production yield of the strains.

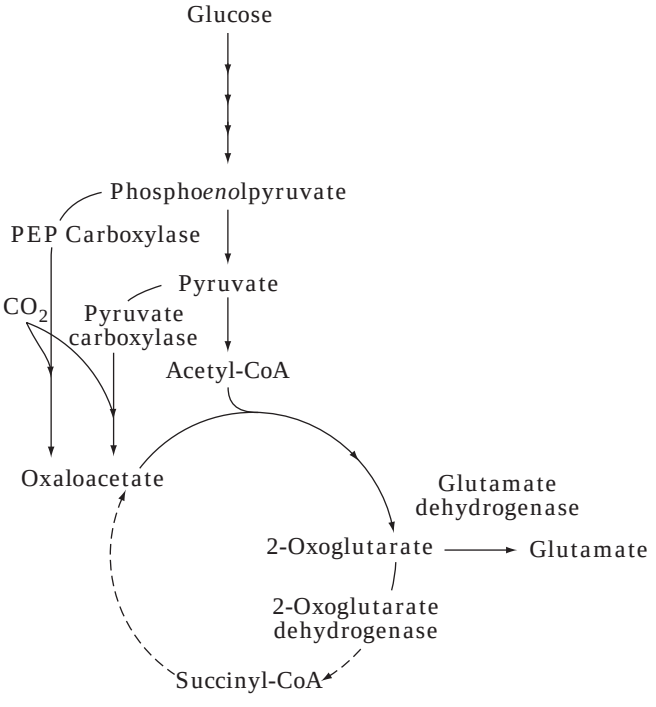


Figure 5. Simplified scheme of glutamic acid biosynthesis pathway through anaplerotic pathways as oxaloacetate generating system.

The mechanism of glutamate excretion into the medium has long been a subject of controversy in the literature. Three different mechanisms have been proposed to describe glutamate efflux. The "leak" model proposes that glutamate is excreted because the cell membrane is physically altered after different treatments, such as biotin limitation, addition of sublethal antibiotic concentration, or addition of surfactants. The second hypothesis is based on a general model for many efflux processes in bacteria, namely, that inversion of the influx system of glutamate is induced by changes in chemical potential or regulation, or by uncoupling due to changes in the membrane structure. Finally, the last hypothesis proposes the existence of a specific glutamate carrier analogous to the recently characterized lysine transport system. These different hypotheses are discussed in a recent review (22).

Momose and Takagi (23) isolated mutant strains of *C. lactofermentum* sensitive to temperature and able to excrete glutamic acid after a thermal upshock from 30 °C to 37 °C in biotin-rich media. One of their mutants produced 20 g/L glutamic acid from molasses. In the future, with such strains it could be possible to produce glutamate from biotin-rich materials without the use of expensive additives, such as antibiotics or surfactants, just by simple control of the temperature.

METABOLIC ENGINEERING OF CORYNEBACTERIA

Through classical mutagenesis and selection, it has been possible to identify *Corynebacterium* mutants that overproduce lysine. By 1958, Kinoshita and colleagues (24) had learned that homoserine auxotrophs secreted higher amounts of lysine than did wild-type strains. Mutants with altered feedback inhibition in the lysine biosynthetic pathway have also been found to overproduce the amino acid. Thus, resistance to the lysine analog 2-aminoethyl-L-cysteine (AEC) has been a hallmark of the lysine overproduction phenotype. Other mutants that overproduce lysine have been found to contain lesions in pathways, such as those for leucine and alanine biosynthesis, which compete for the pyruvate that is required for lysine biosynthesis. Lysine production by *Corynebacterium* strains that are resistant to one or more lysine analogs and that carry auxotrophies for one or more additional amino acids has been described previously (25). The introduction of molecular genetics coupled with a greater understanding of how *Corynebacterium* regulates lysine synthesis has made it possible to enhance lysine productivity even further through metabolic engineering.

Metabolic engineering is an emerging field of research that combines the analysis of complex biochemical networks with biochemistry, enzymology, and molecular biology. Through metabolic engineering it is possible, at least in principle, to identify processes within the cell that limit overall lysine productivity, for example, and introduce spe-

cific changes into the cell to alleviate these limitations. Metabolic engineering requires both a detailed understanding of biosynthetic pathways and their regulation, and the ability to alter specific steps in these pathways. The most direct method for making these alterations is genetic manipulation. In this way, *E. coli* strains, with their well-characterized physiology and wide array of molecular genetic resources, have been engineered to produce large amounts of threonine, valine, or tryptophan. Metabolic engineering in *Corynebacterium*, however, cannot draw on the same diversity of resources, and although the physiology of these bacteria is becoming ever more understood, methods for introducing changes into these strains at the genetic level are not as well developed as they are for *E. coli*.

Genetic Tools for Manipulating *Corynebacterium*

To make the genetic changes necessary for metabolic engineering in *Corynebacterium*, researchers need to be able to identify and clone the genes that are involved in the target pathway. They also need methods for altering these genes to affect the regulation or level of expression of the enzymes they encode, and for subsequently reintroducing the altered genes into *Corynebacterium* to monitor their effects on amino acid biosynthesis. Therefore, metabolic engineers must have at their disposal an array of plasmids that can replicate in both *Corynebacterium* and other, more easily manipulated hosts, such as *E. coli*. Also required are a collection of selectable markers encoding, for example, antibiotic resistance, well-characterized transcriptional promoters that permit regulation of the altered genes, and efficient transformation or conjugation systems that allow the plasmids to be inserted into the target *Corynebacterium* strain.

Plasmids. Several different plasmids have been isolated and developed for the introduction and expression of genes in *Corynebacterium* (26). The majority of these were originally identified as small (3–5 kbp), cryptic plasmids from *C. glutamicum*, *C. callunae*, and *C. lactofermentum*. They fall into four compatibility groups, exemplified by the plasmids pCC1, pBL1, pHM1519, and pGA1. Shuttle vectors, plasmids that are capable of replicating in both *Corynebacterium* and *E. coli*, have been developed from these cryptic plasmids by incorporating elements from known *E. coli* plasmids (particularly the ColE1 origin of replication from pBR322 or pUC18), as well as antibiotic resistance markers. A fifth class of plasmids that is very useful for manipulating *Corynebacterium* is based on pNG2, a plasmid originally isolated from *Corynebacterium diphtheriae* (27). This plasmid and its derivatives replicate efficiently in many species of corynebacteria, as well as in *E. coli*. Since the sole origin of replication in pNG2 (an element of only 1.8 kbp) functions in both the Gram-positive and Gram-negative host, there is no need to add an additional ColE1-type element to it. As a result, pNG2 derivatives (e.g., pEP2) are much smaller than other *Corynebacterium* shuttle vectors and are therefore more easily manipulated.

Selectable Markers. Several genes conferring antibiotic resistance have proven useful for plasmid selection and other recombinant DNA work in corynebacteria. These include the kanamycin resistance determinant from Tn*903*, a hygromycin resistance marker isolated from *Streptomyces hygroscopicus*, a tetracycline resistance gene from *Streptococcus faecalis*, a bleomycin resistance gene from Tn*5*, and a chloramphenicol resistance marker from *Streptomyces acrimycini*. Work in our laboratory has shown that this latter gene does not confer chloramphenicol resistance to *Corynebacterium* unless its native promoter is substituted by a promoter that is known to work in *Corynebacterium*. The β-lactamase gene that is employed in many *E. coli* plasmids such as pBR322 does not confer ampicillin resistance in *Corynebacterium*.

Transformation Systems. Several methods have been devised for introducing foreign DNA into *Corynebacterium*. The earliest method to be employed routinely was based on protocols that had been successful for other Gram-positive species involving incubation of spheroplasts in the presence of DNA and polyethylene glycol (28). While useful, these methods were generally inefficient, often yielding fewer than 10^5 transformants per milligram of DNA. Electroporation of *Corynebacterium* spheroplasts has proven to be a much more efficient and reliable means of transformation. Spheroplasts are generated by growing the cells in rich media containing glycine and/or low concentrations of other inhibitors of cell wall biosynthesis, such as isonicotinic acid hydrazide (isoniazid), ampicillin, penicillin G, or Tween-80. The spheroplasts are then washed in low-salt buffers containing glycerol, concentrated, and mixed with DNA before being subjected to electroporation. Efficiencies as high as 10^7 transformants per microgram of plasmid DNA have been reported with this protocol.

A third method for DNA transfer into corynebacteria involves transconjugation. This method takes advantage of the promiscuity of *E. coli* strains carrying derivatives of the plasmid RP4. In *E. coli*, RP4 encodes many functions that mediate the conjugal transfer of plasmids from the host strain to other recipient strains of *E. coli*, or even to other species. These "*tra* functions" mediate pilus formation and plasmid transfer. RP4 also carries an origin of transfer, *oriT*, a *cis*-acting element that is recognized by the transfer apparatus that allows the plasmid to be conducted through the pilus and into the recipient strain. From this system Simon et al. (29) have developed a useful transconjugation tool that allows the transfer of plasmids from *E. coli* to *Corynebacterium*. They relocated the *tra* functions from RP4 into the *E. coli* chromosome in a strain called S17-1. Plasmids carrying the RP4 *oriT* can be mobilized from S17-1 into other recipients very efficiently. Although this method has proven useful for introducing replicating plasmids into *Corynebacterium*, it has proven even more useful for generating gene disruptions. This is accomplished by introducing a selectable marker into a clone of the *Corynebacterium* gene that is targeted for disruption. This construct is then ligated into an *E. coli* plasmid that carries the RP4 *oriT* but lacks an origin to support replication in *Corynebacterium*. S17-1 carrying this plasmid is then incubated with the recipient strain and the mixture is later transferred to a selective medium. Because the

plasmid that was introduced is unable to replicate in corynebacteria, transconjugants that express the selectable marker are most likely to have undergone a cross-over recombination within the genomic DNA.

Restriction-Deficient Strains. Regardless of the transformation system used, there is clear precedent in the literature that corynebacteria are able to recognize *E. coli*–derived DNA as foreign and will most often degrade it. This ability has been attributed to the *Corynebacterium* restriction and modification system. To overcome this system, some transformation and transconjugation protocols call for briefly heating the recipient strain prior to transformation. The heat treatment presumably inactivates the enzymes responsible for the restriction system, allowing the introduced DNA to become established before the enzymes are turned over. Another strategy for improving the efficiency of DNA transfer has been to isolate *Corynebacterium* mutants that are deficient in the restriction system. These strains will incorporate plasmids that had been propagated in *E. coli* with almost the same efficiency as plasmids that had been propagated in *Corynebacterium*. In an alternate strategy used to circumvent the restriction system in *Corynebacterium*, Leblon and coworkers (30) developed an "integron" system for gene disruption. Integrons are DNA molecules that have the same restriction/modification properties as the target host's DNA, carry DNA that is homologous to a portion of the host genome (i.e., a region of the genome that is to be disrupted), and are unable to replicate in the host cell. A cloned gene from *Corynebacterium* is first interrupted with a selectable marker in a plasmid that is propagated in one *Corynebacterium* strain. This construct is then excised from the corynebacterial plasmid and self-ligated to form a nonreplicating circular molecule. This "integron" is then electroporated into the restrictive host. Modification of the DNA allows the integron to elude the host restriction system, and recombination into the host genome permits expression of the selectable marker.

Promoters. Reliable transcriptional promoters are required for efficient expression of foreign genes in *Corynebacterium*. For certain experiments, there is also a need for regulated promoters whose activity can be induced under specific culture conditions. Promoters such as the *fda*, *thrC*, and *hom* promoters derived from *Corynebacterium* genes have proven useful for heterologous gene expression. Inducible promoters from *E. coli*, such as P_{lac}, P_{tac}, and P_{trc}, which are induced by isopropylthiogalactopyranoside (IPTG) when the *lac* repressor (*lacI*) is present; P_{trp}, which responds to the inducer indole acrylic acid when the *trp* repressor (*trpR*) is present; and lambda P_L, which is repressed in the presence of the temperature-sensitive lambda repressor (cI857), have all been used to modulate gene expression in *Corynebacterium*. There is also intense interest in identifying developmentally regulated promoters from *Corynebacterium*, such as promoters that are activated during late exponential or stationary phases of culture, although this work has proceeded slowly.

Gene Identification. With all other genetic tools in place, there still remains the challenge of identifying relevant genes from *Corynebacterium*. In *E. coli*, some of the resources that have been used to isolate genes are transducing phage, transposable elements, genetic maps of the *E. coli* chromosome from transduction and transconjugation experiments, and more recently, complete physical and sequence maps of the chromosome. To date, the most successful method for identifying and recovering genes from *Corynebacterium* has been to use *Corynebacterium* genomic DNA to complement known auxotrophs of *E. coli*. In this exercise, libraries of plasmids carrying fragments of the *Corynebacterium* genome are introduced into *E. coli* strains that are deficient in a particular enzyme or function. Transformants that no longer display the auxotrophy (e.g., homoserine deficiency) are likely to carry the complementing gene from *Corynebacterium*. This strategy has led to isolation of numerous *Corynebacterium* genes, including several from the pathways responsible for synthesis of aspartate-derived and aromatic amino acids, intermediary metabolism, and other cellular processes. One limitation to this strategy is that not all genes from *Corynebacterium* will be expressed in the *E. coli* host. Thus, although a gene may be represented in the plasmid library, it may be unable to complement the *E. coli* mutation and therefore would not be recovered during selection. Overcoming this limitation, a smaller number of genes have been identified with a similar strategy in which a plasmid library from wild-type *Corynebacterium* was used to directly complement mutations in other *Corynebacterium* strains. Although this strategy avoids the concern of insufficient gene expression in the auxotrophic host, its utility is limited by poor plasmid transformation efficiency in the auxotrophs. Still other genes have been identified by hybridization with nucleic acid probes based upon homologous genes from other species, and direct amplification of genes using the polymerase chain reaction and degenerate oligonucleotide primers.

Transposable Elements. Transposable elements are extremely powerful tools in gene identification because they couple mutagenesis with gene recovery. Unlike classical mutagenesis techniques, which generate point mutations or small deletions within a gene, when transposable elements insert within a gene they form large disruptions, thereby "tagging" the altered gene for easier identification. A number of transposable elements have been found to transpose in *Corynebacterium*. Transposons found in the plasmids pTP10 of *C. xerosis* and pNG2 of *C. diphtheriae* have been shown to transpose in *C. glutamicum* and confer resistance to erythromycin. A group from the Mitsubishi Chemical Company in Japan developed a series of artificial transposons from an insertion sequence, IS*31831*, that they discovered in *C. glutamicum* (31). After inserting a selectable marker between the inverted repeats of IS*31831*, these researchers were able to introduce the resulting transposon into *C. glutamicum* strains on an *E. coli* plasmid (unable to replicate in *Corynebacterium*) via electroporation. They found that the selectable marker had inserted into the genome of the target cell at a frequency of approximately 4×10^4 mutants/μg DNA. The use of such transposons to generate *Corynebacterium* auxotrophs has led to the isolation of several genes responsible for amino

acid biosynthesis, as well as other functions in corynebacteria.

Transducing Phage. Transducing phage have been used in other systems for mapping genetic loci and for isolating genes. In 1976, researchers at Ajinomoto Co. in Japan surveyed 150 strains of characterized and uncharacterized glutamic acid–producing coryneform bacteria to identify phage that might be useful for transduction (32). Of 24 different phage isolates recovered from this screen, only three were able to transduce a *trp* marker from a trp^+ donor to a trp^- recipient with any appreciable frequency, although even this efficiency was only 10^{-7} or less. These researchers were able to improve transduction efficiency slightly by including 4 mM cyclic adenosine monophosphate (cAMP) or 1.2 M magnesium chloride. Several different researchers have attempted to develop reliable transduction methods by isolating corynephages from sources such as contaminated industrial fermentations, soil, and animal waste. Although many phage have been isolated and characterized, few have been associated with transduction, and an opportunity still exists to develop a reliable, high-efficiency transduction system for general use with the glutamic acid–producing bacteria.

Metabolic Engineering for Aspartate-Derived Amino Acids

Classical mutagenesis and selection has given rise to amino acid–overproducing strains that carry mutations that alleviate feedback inhibition in critical regulatory enzymes, eliminate reactions that compete for substrates, alter control of gene expression, or somehow modify other aspects of the cell's physiology. Metabolic engineering of the pathways for the aspartate family of amino acids has followed very similar themes. Aspartate is derived from oxaloacetate, a tricarboxylic acid cycle (TCA) intermediate. All of a cell's lysine, methionine, threonine, and isoleucine are produced from aspartate (Fig. 1). *Corynebacterium* modulates the partitioning of precursors into pathways specific for the individual amino acids in this family at several levels, including feedback sensitivity of several key enzymes, differences in the specific activities of competing enzymes at nodes where the pathways for two amino acids diverge, and genetic repression. Although it is complex, regulation of the aspartate pathways in *Corynebacterium* does not appear to be as complex as that in *E. coli* (Fig. 1).

Lysine. Although there appears to be some level of transcriptional regulation of the operon that encodes diaminopimelate decarboxylase, the enzyme that carries out the final step in lysine biosynthesis (Fig. 6), the main control in lysine biosynthesis in *Corynebacterium* is exercised by aspartokinase (AK), which catalyzes the first common step toward the synthesis of all of the aspartate-derived amino acids. Many AEC-resistant strains of *Corynebacterium* have been shown to produce an aspartokinase that is insensitive to the feedback inhibition by lysine and threonine that is displayed by the wild-type enzyme. Sequence analysis of cloned aspartokinase genes from AEC-resistant *Corynebacterium* strains has revealed the molecular basis

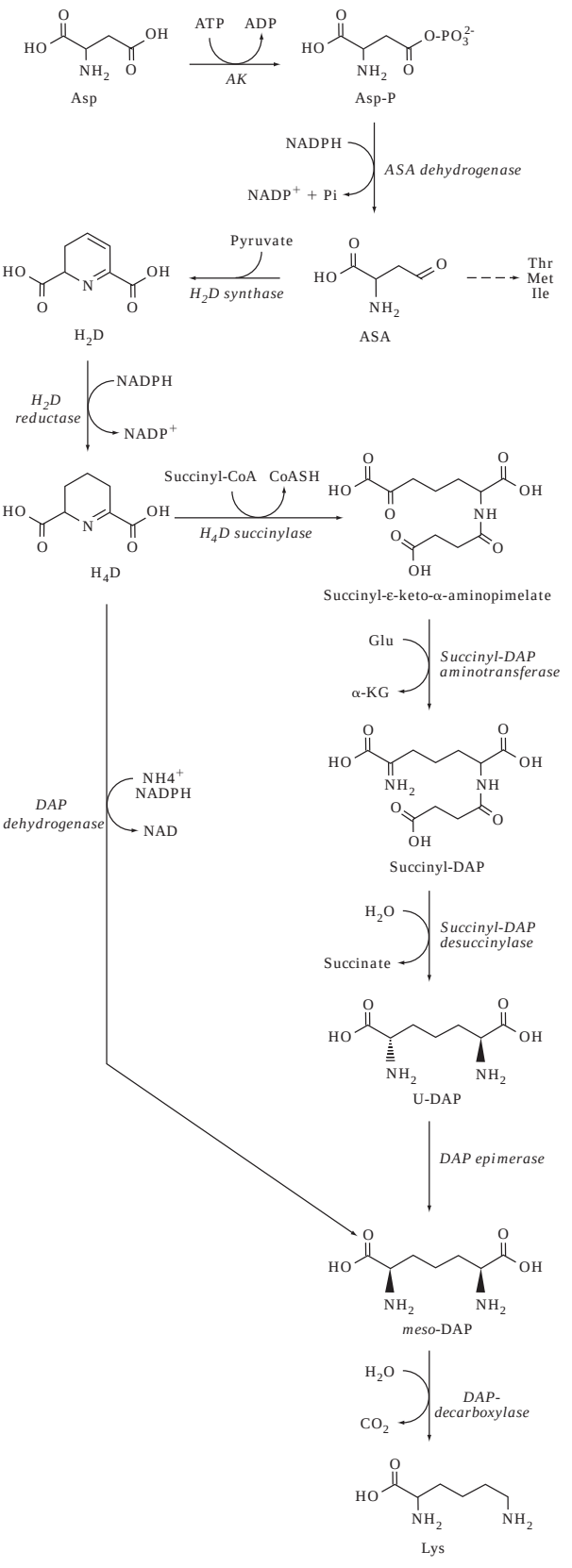


Figure 6. Lysine synthesis in *Corynebacterium glutamicum*. AK, aspartokinase; ASA, aspartate semialdehyde; H_2D, dihydrodipicolinate; H_4D, tetrahydrodipicolinate; DAP, diaminopimelate; α-KG, α-ketoglutarate.

of lysine overproduction in these strains. In *Corynebacterium*, aspartokinase is synthesized as two polypeptides, AKα and AKβ, which are encoded by overlapping reading frames and assemble as an $\alpha_2\beta_2$ heterotetramer. Feedback resistance is due to mutations that lead to amino acid substitutions within the AKβ subunit: specifically, substitution of tyrosine in place of serine at position 301, aspartate in place of glycine at position 345, or glycine in place of arginine at position 349 (22). From this it is postulated that the β-subunit of aspartokinase mediates regulation by lysine and threonine, although it is not clear whether these amino acids are directly involved in binding of either lysine, threonine, or both. Regardless, these mutations alone allow *Corynebacterium* to accumulate and secrete high levels of lysine.

However, it is clear that feedback inhibition of aspartokinase is not the only limitation to lysine overproduction. In AEC-resistant lysine overproducers, amplifying expression of the deregulated aspartokinase gene through recombinant DNA techniques further increases the synthesis of lysine, indicating that the specific activity of this enzyme is also limiting. A further limitation in lysine production, especially at the industrial scale, comes at the point of lysine secretion (33). Wild-type *Corynebacterium* cannot secrete significant amounts of lysine. Many lysine-overproducing strains have been found to possess enhanced lysine export systems in addition to deregulated aspartokinase activity. The lysine export mechanism apparently involves symport of two hydroxyl ions and can proceed against a lysine gradient and against the membrane potential (33). At this time, the molecular basis for increased lysine export during overproduction is not yet fully understood.

Unlike *E. coli*, *C. glutamicum* possesses two biochemical pathways for synthesizing lysine or, more precisely, for synthesizing *meso*-diaminopimelate (*meso*-DAP), the immediate precursor to lysine. The first pathway (the succinylase pathway) is initiated by the enzyme tetrahydrodipicolinate succinylase and involves a total of four enzymatic steps to produce *meso*-DAP. The second pathway (the dehydrogenase pathway) converts tetrahydrodipicolinate to *meso*-DAP in a single step through the action of DAP dehydrogenase. Because these two pathways incorporate nitrogen via different substrates (the succinylase pathway uses glutamate while the dehydrogenase pathway uses ammonium), the operation of these two mechanisms allows *C. glutamicum* to grow in the presence of different nitrogen sources. *meso*-DAP is an important cell wall component in corynebacteria, providing the crosslinking of the peptidoglycan material in this structure. Strains of *C. glutamicum* that are deficient in the succinylase pathway were found to have altered morphology and cell wall rigidity (34). Deficiency in the dehydrogenase pathway reduces productivity during lysine fermentations without significantly altering growth rates (35). The dehydrogenase enzyme has a lower affinity for the common substrate than does the succinylase, and it is constitutively expressed. The succinylase pathway, on the other hand, is energetically more expensive. ^{13}C-nuclear magnetic resonance studies on carbon flux (36) have confirmed that during normal growth the majority of the *meso*-DAP is produced via the succinylase pathway, and that during conditions of high lysine production the dehydrogenase pathway accounts for the bulk of this flux.

Threonine. Aspartate semialdehyde is the last common intermediate in the pathways for the synthesis of lysine and the other aspartate-derived amino acids (Fig. 7). Given

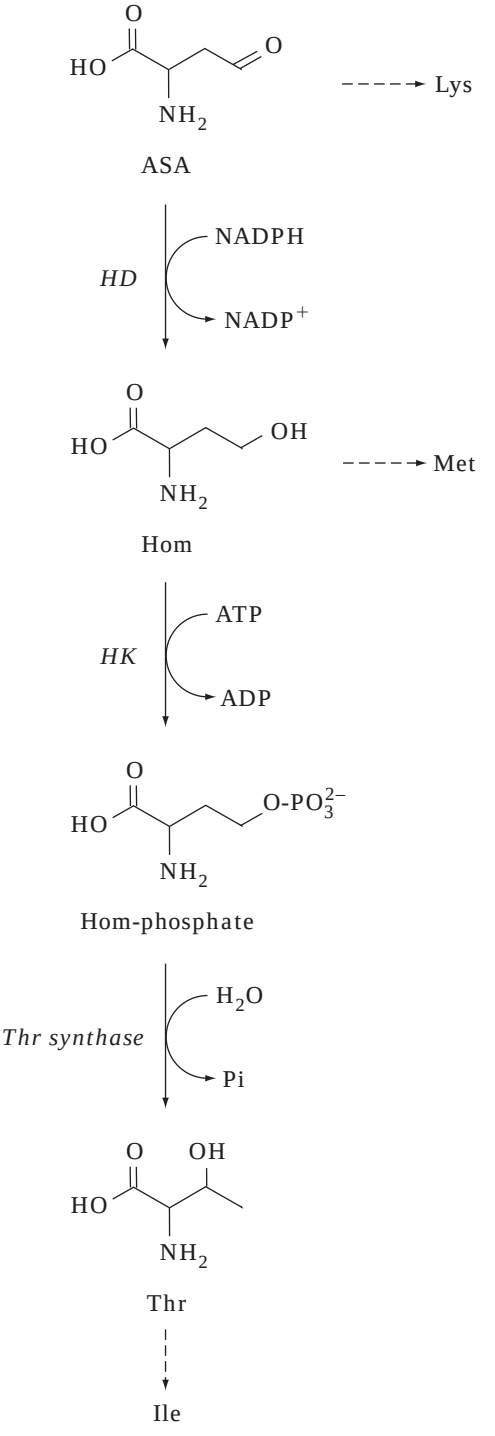


Figure 7. Threonine synthesis in *Corynebacterium glutamicum*. ASA, aspartate semialdehyde; Hom, homoserine; HD, homoserine dehydrogenase; HK, homoserine kinase.

the fact that lysine, methionine, threonine, and isoleucine are all derived from aspartate, it would seem to follow that deregulation of aspartokinase should boost synthesis of these other products along with lysine. The observation that the first committed step toward threonine, methionine, and isoleucine is catalyzed by an enzyme (homoserine dehydrogenase) with a stronger affinity for aspartate semialdehyde than the enzyme (dihydropicolinate synthase) that leads to lysine would seem to support this notion. However, when feedback inhibition of aspartokinase is relieved, the predominant product to accumulate is lysine because, unlike the pathway for lysine biosynthesis, the threonine and methionine biosynthetic pathway contains additional regulatory mechanisms. Homoserine dehydrogenase itself is inhibited by accumulation of threonine, and expression of the homoserine dehydrogenase gene is repressed by methionine. The result is that, even though homoserine dehydrogenase can outcompete dihydrodipicolinate synthase for aspartate semialdehyde, accumulation of the products of this branch in the pathway is quickly squelched. Lysine synthesis, which has almost no subsequent direct regulation, then resumes unhindered.

Because of its strong feedback regulation, homoserine dehydrogenase is an obvious target for metabolic engineering to improve threonine biosynthesis in *Corynebacterium*. Mutations near the carboxy terminus of this enzyme have been found to confer insensitivity to feedback regulation. When the deregulated homoserine dehydrogenase gene, hom^{dr}, and copies of *thrB*, which encodes homoserine kinase (the next enzyme in the pathway toward threonine), were integrated into the genomes of AEC-resistant strains of *Corynebacterium*, threonine accumulated at the expense of lysine (22,23). Inefficient export of the excess threonine in these cells triggered feedback inhibition of homoserine kinase, causing homoserine to accumulate as well. Concomitant overexpression of *thrB* from an inducible promoter was found to partially compensate for the deficiencies in threonine export by allowing more homoserine to be converted to threonine, but this effort led to the discovery that excess threonine was rapidly lost as the degradation products isoleucine and glycine. Disruption of the gene encoding threonine dehydratase, the first enzyme in the pathway from threonine to isoleucine, permitted more threonine to accumulate. However, threonine was still lost as glycine through two additional pathways, the first involving threonine dehydrogenase, and the second involving serine hydroxymethyl transferase, which is also active in serine catabolism (22). The difficulties encountered in development of *Corynebacterium* strains that overproduce threonine illustrates the complex and elegant regulation of this pathway.

Isoleucine. Overproduction of isoleucine has been accomplished by introducing excess threonine dehydratase (encoded by *ilvA*) into threonine production strains (37) (Fig. 8). Thus, combining a deregulated aspartokinase, deregulated homoserine dehydrogenase, overexpressed homoserine kinase, and overexpressed threonine dehydratase will result in abnormally high levels of isoleucine accumulation. Threonine dehydratase is normally feedback-inhibited by isoleucine, so it is not surprising that

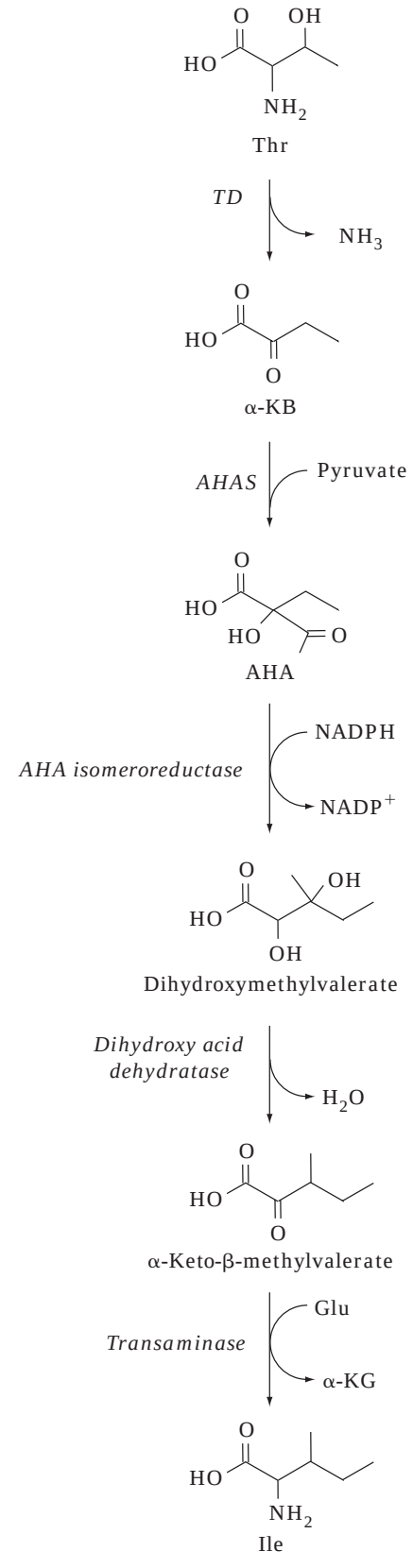


Figure 8. Isoleucine synthesis in *Corynebacterium glutamicum*. TD, threonine dehydratase; α-KB, α-ketobutyrate; AHA, acetohydroxy acid; AHAS, acetohydroxy acid synthase; α-KG, α-ketoglutarate.

mutant derivatives of threonine dehydratase with reduced sensitivity to isoleucine provided an additional dividend in this isoleucine production system (38). Despite these gains, it appears that amino acid export is once again a serious limitation to the effectiveness of amino acid production (33). An increase in threonine dehydratase activity causes α-ketobutyrate to accumulate. This induces the activity of the next enzyme in the isoleucine pathway, acetohydroxy acid synthase (AHAS), which might be expected to promote the conversion of threonine to isoleucine. Indeed, the addition of α-ketobutyrate alone to the medium enhanced flow through the latter parts of this pathway. However, when threonine dehydratase is overexpressed in the threonine production system, α-ketobutyrate is excreted into the medium. AHAS is also involved in valine and leucine synthesis and is itself feedback inhibited by the branched-chain amino acids. It appears that under isoleucine fermentation conditions a concomitant accumulation of valine may inhibit AHAS, thereby overcoming the stimulatory effects of α-ketobutyrate accumulation (37). Observations of intracellular concentrations of isoleucine suggest that isoleucine export is also a major rate-limiting step in isoleucine overproduction (39).

PERSPECTIVES

Considerable progress has been made in the production of amino acids by corynebacteria since the isolation of the first strain of *C. glutamicum*. In only four decades, the field has advanced from classical genetic manipulation of strains with uncharacterized lesions, through elegant biochemical analyses of enzymes from wild-type and overproducing strains, and on to the current molecular approaches. Aided by information gathered from sophisticated metabolic flux analyses and a greater understanding of biochemical pathways and cellular processes, the field has begun to apply a more rational strategy based on metabolic engineering of improved strains. These new strategies, facilitated by the development of improved genetic tools for the manipulation of corynebacteria, should permit the unraveling of complicated regulatory pathways. Using these methods, researchers have overcome many obstacles to amino acid overproduction and gained valuable insights that have benefited the broader scientific community.

Challenges to further enhancements in amino acid production remain in the areas of amino acid transport and global regulatory mechanisms (40), particularly those relating to central carbon metabolism. There is also opportunity for further improvement of large-scale fermentation techniques. The availability of the complete genomic sequence of closely related Gram-positive bacterial strains, including *Mycobacterium tuberculosis* and *M. leprae*, will undoubtedly aid in the isolation of homologous genes from corynebacteria. As the analysis of bacterial genomes and proteomes continues to mature, functional genomics and high-throughput chip-based technologies for the visualization of gene expression and cellular protein content hold out the promise for synthesis of an accurate and detailed picture of cellular metabolism. The study of the glutamic acid–producing corynebacteria will likely lead to continuous gains in productivity while offering new insights into microbial physiology.

BIBLIOGRAPHY

1. S. Udaka, *J. Bacteriol.* **79**, 754–755 (1960).
2. S. Abe, K.-I. Takayama, and S. Kinoshita, *J. Gen. Appl. Microbiol.* **13**, 279–301 (1967).
3. W. Liebl, in A. Balows, H.G. Trueper, M. Dworkin, W. Harder, and K.H. Schleifer eds., *The Prokaryotes*, Springer-Verlag, Berlin, 1991, pp. 1157–1171.
4. W. Liebl, M. Ehrmann, W. Ludwig, and K.H. Schleifer, *Int. J. Syst. Bacteriol.* **41**, 255–260 (1991).
5. S. Muirheid, *Feedstuffs* **69**, 1 (1997).
6. *Medical & Healthcare Marketplace Guide*: 1997–1998, 13th ed., Dorland's Biomedical, Philadelphia, Pa., 1997, pp. III–237.
7. M. Lerner, *Chemical Market Reporter* **253**, January 26, 1998, p. 4.
8. M. Lerner, *Chemical Marketing Reporter* **250**, September 16, 1996, p. 18.
9. J.C. Measures, *Nature* **257**, 398–400 (1975).
10. A.J. Makoff, I.G. Charles, and N.F. Fairweather, in Y. Murooka and T. Imanaka eds., *Recombinant Microbes for Industrial and Agricultural Applications*, Marcel Dekker, New York, 1994, pp. 205–232.
11. M. Ikemi, in Y. Murooka and T. Imanaka eds., *Recombinant Microbes for Industrial and Agricultural Applications*, Marcel Dekker, New York, 1994, pp. 797–814.
12. H. Billman-Jacobe, L. Wang, A. Kortt, D. Stewart, and A. Radford, *Appl. Environ. Microbiol.* **61**, 1610–1613 (1995).
13. M.E. Johnson and J.L. Steele, in M.P. Doyle, L.R. Beuchat and T.J. Montville, eds., *Food Microbiology: Fundamentals and Frontiers*, ASM Press, Washington D.C. 1997, pp. 581–594.
14. J.Y. Leveau, *Industries Alimentaires Agricoles* March 102 (1992).
15. K. Takinami, Y. Yamada, and H. Okada, *Agric. Biol. Chem.* **30**, 674–682 (1966).
16. T.D. Nunheimer, J. Birnbaum, E.D. Ihnen, and A.L. Demain, *Appl. Microbiol.* **20**, 215–217 (1970).
17. M. Marquet, J.L. Uribelarrea, A. Huchenq, G. Laneelle, and G. Goma, *Appl. Microbiol. Biotechnol.* **25**, 220–223 (1986).
18. D. Leyval, F. Debay, J.M. Engasser, and J.L. Goergen, *J. Microbiol. Methods* **29**, 121–127 (1997).
19. F. Debay, D. Leyval, J.M. Engasser, and J.L. Goergen, *Récents Progrès en Génie des Procédés* **57**, 43–48 (1997).
20. I. Shiio and K. Ujigawa, *Agric. Biol. Chem.* **42**, 1897–1904 (1978).
21. Y. Kawahara, K. Takahashi-Tuke, E. Shimizu, T. Nakamatsu and S. Nakamori, *Biosci. Biotech. Biochem.* **61**, 1109–1112 (1997).
22. M.S.M. Jetten and A.J. Sinskey, *Crit. Rev. Biotechnol.* **15**, 73–103 (1995).
23. H. Momose and T. Takagi, *Agric. Biol. Chem.* **42**, 1911–1917 (1978).
24. S. Konishita, K. Nakayama, and S. Kitada, *J. Gen. Appl. Microbiol.* **4**, 128 (1958).
25. K. Nakayama, in ed Blanch, S. Drew and Wang, eds., *Comprehensive Biotechnology*, vol. 3, 1985, pp. 607–620.

26. H. Sonnen, G. Thierbach, S. Kautz, J. Kalinowski, J. Schneider, A. Pühler, and H.J. Kutzner, *Gene* **107**, 69–74 (1991).
27. T.M. Serwold-Davis, N. Groman, and M. Rabin, *Proc. Natl. Acad. Sci. USA* **84**, 4964–4968 (1987).
28. M. Yoshihama, K. Higashiro, E.A. Rao, M. Akedo, W.G. Shanabruch, M.T. Follettie, G.C. Walker, and A.J. Sinskey, *J. Bacteriol.* **162**, 591–597 (1985).
29. R. Simon, U. Preifer, and A. Pühler, *Bio/Technology* **1**, 784–791 (1983).
30. O. Reyes, A. Guyonvarch, C. Bonamy, V. Salti, F. David, and G. Leblon, *Gene* **107**, 61–68 (1991).
31. A.A. Vertes, Y. Asai, M. Inui, M. Kobayashi, Y. Kurusu, and H. Yukawa, *Mol. Gen. Genet.* **245**, 397–405 (1994).
32. H. Momose, S. Miyashiro, and M. Oba, *J. Gen. Appl. Microbiol.* **22**, 119–129 (1976).
33. H. Sahm, L. Eggeling, B. Eikmanns, and R. Kramer, *FEMS Microbiol. Rev.* **16**, 243–252 (1995).
34. A. Wehrmann, B. Phillipp, H. Sahm, and L. Eggeling, *J. Bacteriol.* **180**, 3159–3165 (1998).
35. B. Schrumpf, L. Eggeling, and H. Sahm, *Appl. Microbiol. Biotechnol.* **37**, 566–571 (1992).
36. C.A. Shaw-Reid, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, Mass., 1997.
37. G.E. Colón, T.T. Nguyen, M.S.M. Jetten, A.J. Sinskey, and G. Stephanopoulos, *Appl. Microbiol. Biotechnol.* **43**, 482–488 (1995).
38. S. Morbach, H. Sahm, and L. Eggeling, *Appl. Environ. Microbiol.* **61**, 4315–4320 (1995).
39. S. Morbach, H. Sahm, and L. Eggeling, *Appl. Environ. Microbiol.* **62**, 4345–4351 (1996).
40. R. Krämer, *J. Biotechnol.* **45**, 1–21 (1996).

CRYSTALLIZATION, BULK, MACROMOLECULES

ANA PAULINA BARBA DE LA ROSA
ALEXANDER MCPHERSON
University of California, Irvine
Irvine, California

KEY WORDS

Bulk purification
Crystals
Drug delivery
Drug design
Protein engineering
Proteins
X-ray diffraction

OUTLINE

INTRODUCTION

There are three principle areas of modern biotechnology where crystallization plays important roles. These are (*1*) The crystallization of proteins as well as viruses, nucleic acids, and macromolecular complexes for X-ray diffraction analysis; (*2*) the crystallization of specific proteins as a part of their formulation in a therapeutic compound, agricultural chemical, or drug; and (*3*) the crystallization of enzymes, and in some cases other proteins, as a means of purification from bulk processes. Although the objectives in each of these situations may be quite different, with the exception of scale, the underlying principles and approaches are fundamentally the same.

In cases of crystallization for X-ray analysis, one is usually dealing with quite pure, often exceptionally pure molecules, and the goal is to grow only a few large, perfect crystals like those shown in Figure 1. The X-ray analysis is a singular event or study, confined to the research laboratory, and the product is basic scientific knowledge. The crystals themselves have no material or medicinal value, but serve as intermediaries in the crystallography process. They provide X-ray diffraction patterns that in turn serve as raw data allowing the visualization of the macromolecules composing the crystals. Thus, they serve as an essential component in the elucidation of molecular structure.

Macromolecular structure is of formidable value in biotechnology because it is the essential knowledge required to apply the technique of rational drug design in the creation and discovery of new drugs and pharmaceutical products. It is the basis of powerful approaches now being applied in small, emerging biotechnology companies as well as major pharmaceutical companies to identify lead compounds to treat a host of human ailments, veterinary problems, and crop diseases in agriculture. The underlying hypothesis is that if the structure of the active site of a salient enzyme in a metabolic or regulatory pathway is known, then chemical compounds, such as drugs, can be rationally designed to inhibit or otherwise affect the behavior of that enzyme.

A second approach, of equal importance to biotechnology, that also requires X-ray quality crystals is the genetic engineering of proteins. Although recombinant DNA techniques provide the essential synthetic role that permits modification of the proteins, crystallography provides the analytical function. It serves as the structural guide for the introduction of intelligent and purposeful changes, in the place of random and chance amino acid substitutions. Crystallography allows direct visualization of the structural changes introduced and offers new directions for chemical or physical enhancements.

Bulk crystallization, along with the detailed consideration of crystallization processes and products, has long been a staple in the pharmaceutical and chemical industry (1), where it has been applied to conventional small mol-

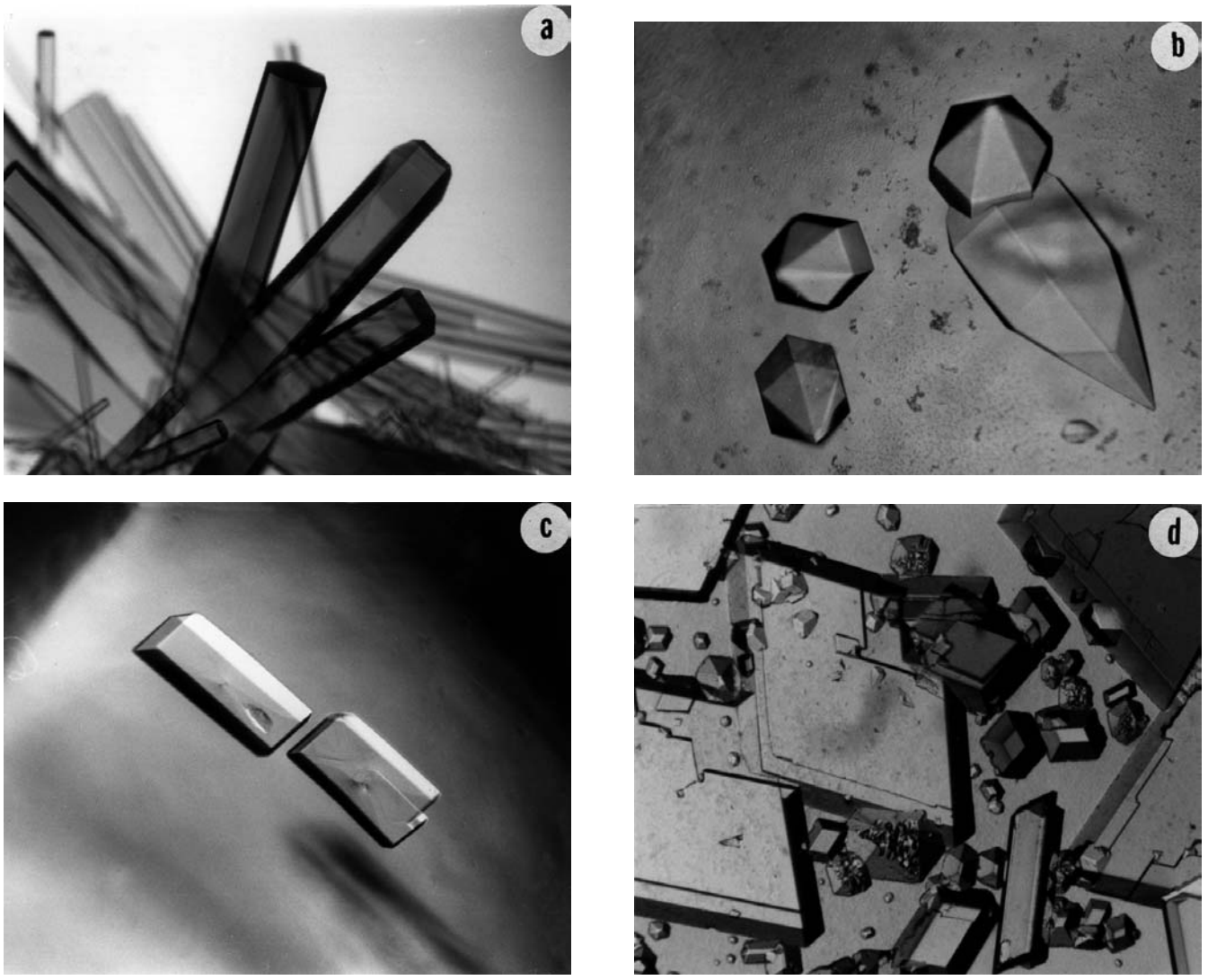


Figure 1. The crystals shown here were grown for X-ray diffraction studies that require a relatively small number of crystals, but of large size and high quality. (**a**) Beef liver catalase, a detoxifying enzyme; (**b**) α_1-acid glycoprotein, a human serum protein; (**c**) Fab fragments of an antiviral mouse antibody; and (**d**) canavalin, the major storage protein from legumes.

ecules and even more complex biological molecules such as glycosides, peptides, and antibiotics. A few examples from the literature where the crystallization procedures are described in some detail are for L-alanine (2), fructosylxyloside from *Scopulariopsis brevicaulis* (3), cephalosporin (4), and a flavanone, hesperidin from orange peels (5). Here, it has long been recognized that manipulation of crystalline formulations could dramatically affect solubility, shelf life, impurity content, thermal stability, release, efficacy, or a host of other commercially significant properties. Much research has, therefore, been devoted to the crystallization of commercially important conventional, small molecules during the past hundred years.

Only more recently has it become evident that crystalline formulations hold substantial promise for the administration and delivery of biological macromolecules as well (6,7). The first major example was that of insulin, a small protein of striking physiological importance to nearly all animals and crucial to the treatment of diabetes. In this instance, it was found that injection of crystalline protein resulted in a more prolonged, slower release of the protein drug and a significant improvement in its efficacy. This timed-delivery feature, first realized for insulin, now appears possible for other protein drugs.

In the coming years, a majority of biotechnology products will likely be proteins, particularly genetically modified proteins designed for introduction into living organisms, including humans. Their delivery and release will be a major problem, and their formulation as crystals will serve a major role in solving those problems. Here, unlike the growth of protein crystals for X-ray diffraction analysis, the crystals are in fact the product, and their produc-

tion must be reproducible. The physical properties of the crystals, such as size, morphology, solubility, and impurity levels will be rigorously specified. They will be produced, under strict regulatory control, as is any drug. Thus, the objectives are quite different than for crystals produced for diffraction analysis.

In terms of pharmaceutically important proteins, natural insulin from animal tissues was certainly the first to be produced in bulk quantities by crystallization (8), and recombinant insulin followed (9). Others, however, are now known, including a large-scale facility for the production of crystalline Fab fragments from monoclonal antibodies (10).

The third area for crystal growth, the downstream processing of proteins produced in large fermentations for the purpose of purification, has associated with it yet another set of objectives and problems. Crystallization is one of the oldest chemical purification technologies and one that has been seriously underutilized to this point in biotechnology. Here, the goal is to promote the crystallization of macromolecules from extremely large volumes, commonly reaching several thousands of liters, of rather impure solutions or culture broths. The ideal is to achieve a very high degree of purification in a single, economical step, and at the same time generate a concentrated, stable, and active product that appeals in its form to a broad customer base. Although certain crystal sizes, morphologies, size distribution, and other features of the crystals may be preferred, these are not the most important parameters. Often crystals such as those shown in Figure 2 are quite acceptable. The major considerations are yield, reproducibility, degree of purification, and the ability to crystallize the material in spite of very high and variable contaminant levels.

Major difficulties accompany the crystallization of proteins in downstream processing and purification that are not seen in X-ray diffraction or formulation crystal growth. For example, processes such as vapor diffusion (discussed later), which is applicable to microliter volumes, are completely inapplicable to thousand liter volumes. Exotic precipitants may not be economical, sensitive procedures such as column chromatography may be out of the question, and speed is often of the essence. Waste product disposal can be a significant consideration. The process must be carefully conceived and developed to be as economical and efficient as possible and still yield the desired crystalline product at the required purity.

There are many examples in the literature of crystallization used for the purification of proteins from bulk fermentations, and these provide useful guides to the design of new procedures for other macromolecules. In one of the earliest reports, Fukumoto et al. (11) reported the purification of lipase from a strain of *Aspergillus niger* by fractionation with ammonium sulfate and its ultimate crystallization from acetone. Tsujisaka et al. (12) in 1973 purified the same enzyme from *Geotrichum candidum* using salt fractionation, DEAE Sephadex, and gel chromotography. The protein was crystallized from its aqueous solution. Buckel et al. (13) in 1981 and Mack et al. (14) describe a scheme based on salt fractionations and chromatography for the production of glutaconate CoA-transferase from *Acidaminococcus fermentans*.

In the 1990s, the appearance of reports of bulk crystallizations accelerated, including those on procedures for protein iminopeptidase from *Bacillus coagulans* (15), chymosin mutants from *Trichoderma reesei* (16), and pyroglutamyl peptidase from *Bacillus amyloliquefaciens* (17). Although only abstracts are currently available, some additional enzymes have also been crystallized for commercial production, these include cellulase (18,19), subtilisin protease and glucose isomerase (18), and alcohol oxidase (20).

THE UNIQUENESS OF MACROMOLECULAR CRYSTALS

Macromolecules, being unique in their properties, both in terms of size and complexity, give rise to crystals that are unique as well (21,22). We cannot, therefore, expect that

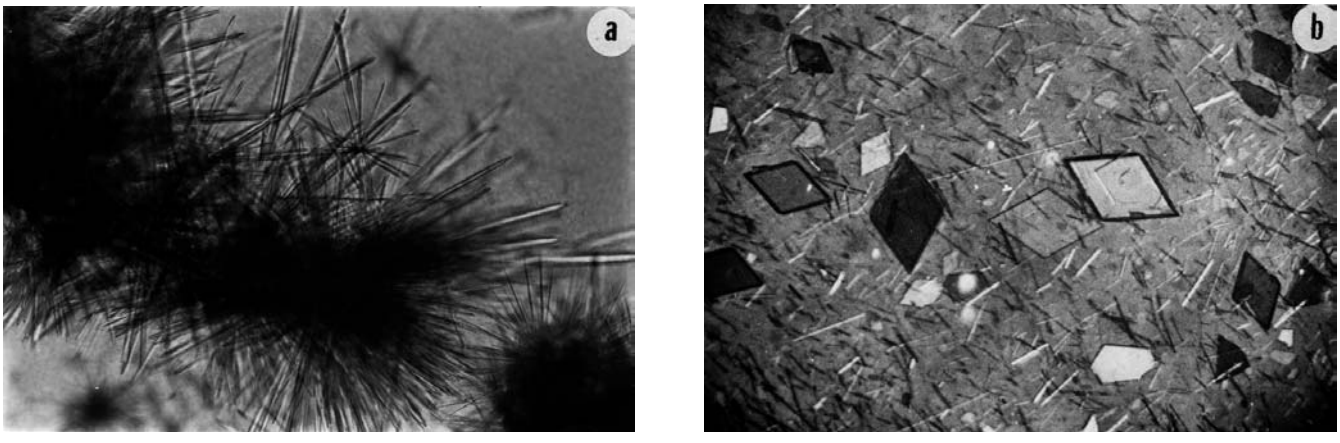


Figure 2. Needle crystals of α-amylase (**a**) and platelike microcrystals and needles of pineapple bromelin (**b**) are entirely suitable for downstream processing of enzymes produced in bulk for industrial purposes but are not of sufficient size nor of useful morphology for analytical studies.

macromolecular crystals will necessarily develop according to precisely the same mechanisms and principles as do conventional crystals (23). Indeed, recent investigations using advanced physical techniques such as inelastic light scattering, interferometry, and atomic force microscopy (22) have yielded evidence of new mechanisms and properties not before encountered with conventional crystals.

Macromolecular crystals are relatively small in comparison with conventional crystals, rarely exceeding 1 mm on an edge and generally smaller. Because only one stereoisomer of a biological macromolecule naturally exists, they do not form crystals possessing inversion symmetry and, therefore, generally exhibit simple shapes that lack the polyhedral character of small molecule crystals. They are extremely fragile, often crushing at the touch; degrade outside a narrow temperature, ionic strength, or pH range; generally exhibit weak optical properties (24); and diffract X-rays to resolutions far short of the theoretical limit (25). The reason for most of these character deficiencies is that macromolecular crystals incorporate large amounts of solvent in their lattices, ranging from about 30% at the lower limit to 90% or more in the most extreme cases (26,27). Proteins also have, as individual molecules, an array of water molecules surrounding them that are relatively tightly bound both in solution and in the crystal (28).

There are two other crucial differences between macromolecular and conventional crystal growth that have important practical consequences. The first is that macromolecular crystals are usually nucleated at extremely high levels of supersaturation, often several hundred to a thousand percent. Small molecule crystals, on the other hand, usually nucleate at only a few percent supersaturation. Virtually every quantitative aspect of crystal growth is a direct function of supersaturation (29). Although high supersaturation may be essential to promote nucleation, it is far from ideal for growth, and the many problems observed for macromolecular crystals attest to this. Furthermore, supersaturated macromolecular solutions, in addition to crystal nuclei, produce alternate solid states that we refer to collectively as amorphous precipitates. Unlike conventional systems, competition exists at both the nucleation and growth stages between crystals and precipitate. This is particularly acute because competition is promoted by high levels of supersaturation. Because amorphous precipitates are kinetically favored, though of less-favored energy states, they tend to dominate the equilibration process and often inhibit or preclude crystal formation.

Given the complexities that beset macromolecules, can we reasonably expect their crystallization to resemble that of conventional molecules? Evidence to this point suggests that the answer is in principle yes, but in practice no. It appears that the fundamental mechanisms and pathways of macromolecular crystal growth are the same as for conventional crystals (30,31) but that the magnitudes of the underlying kinetic and thermodynamic parameters that govern the process differ dramatically.

METHODS FOR GROWTH OF PROTEIN CRYSTALS

The growth of macromolecular crystals has been studied for more than 150 years, initially as a purification tool, and more recently for other purposes. It has been the topic of several texts and reviews and, with the increasing importance of protein crystallography, has become a major factor in the biotechnology industry. Because this body of literature exists, no attempt will be made here to exhaustively review all methods and aspects of macromolecular crystal growth. An effort will be made, however, to describe the underlying principles, present the most commonly employed laboratory approaches, summarized in Table 1, and suggest ways in which the field can be addressed by novice crystal growers.

The strategy used to bring about crystallization (22,32–37) is to direct a macromolecular solution slowly toward a state of supersaturation by modifying the properties of the solvent or the solute. This is generally achieved by altering the concentration of a precipitating agent, such as salt, or by altering some physical property, such as pH. This is illustrated by a version of the classical phase diagram for supersaturation, modified somewhat for the case of proteins in Figure 3.

In concentrated solutions, macromolecules, unlike most conventional molecules, may aggregate as an amorphous precipitate, a result to be avoided. It is usually indicative that supersaturation has proceeded too far or been reached too swiftly. One would like to slowly and marginally exceed the equilibrium point of saturation and thereby allow the macromolecules reasonable opportunity to assemble into a crystalline lattice.

The traditional procedure, illustrated by Figure 4, for inducing proteins to separate from solution as a solid phase is to gradually increase the level of saturation of a salt. Classically, the salt was ammonium sulfate, but others are also in common use (see, for example, Table 2 and the databases described in Refs. 26 and 27). Although the protein may separate as an amorphous precipitate, with appropriate care, manipulation of salt concentration is an effective means for growing protein crystals. This approach has, in fact, probably yielded more varieties of protein crystals than any other (22,26,27).

The accepted explanation for this salting-out phenomenon is that the salt ions and macromolecules compete for chemical interactions, principally hydrogen bonds, with the solvent molecules, that is, water. This occurs because both salt ions and protein molecules depend on hydration layers to maintain solubility. When competition between

Table 1. Methods for Attaining a Solubility Minimum

Bulk crystallization
Batch method in vials
Evaporation
Bulk dialysis
Concentration dialysis
Microdialysis
Liquid bridge
Free-interface diffusion
Vapor diffusion on plates or slides
Vapor diffusion in hanging drops
Sequential extraction
pH-Induced crystallization
Temperature-induced crystallization
Crystallization by effector addition

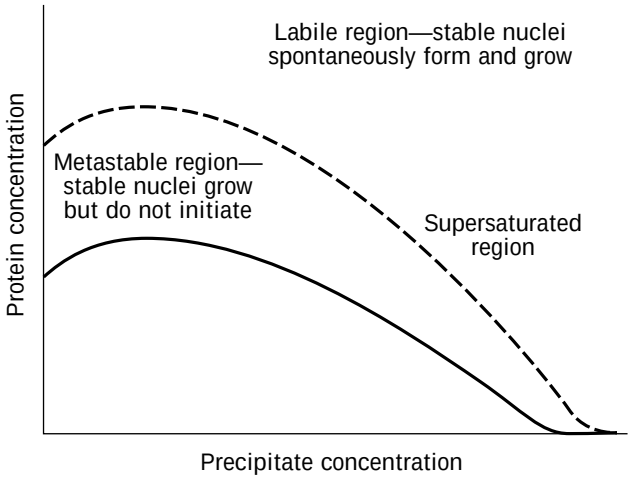


Figure 3. A phase diagram for a hypothetical protein showing its solubility as a function of precipitant concentration. The solid line represents the maximum solubility, or saturation curve, for the protein. Note that the protein is less soluble at very low and very high concentrations of the precipitant, corresponding to the salting-in and salting-out regions, respectively. The supersaturated region lies above the maximum solubility curve, and is, in turn, demarcated by a boundary discriminating the metastable region of supersaturation from the labile region. In the labile region, crystal nuclei can both spontaneously form and grow, whereas in the metastable region they can only grow.

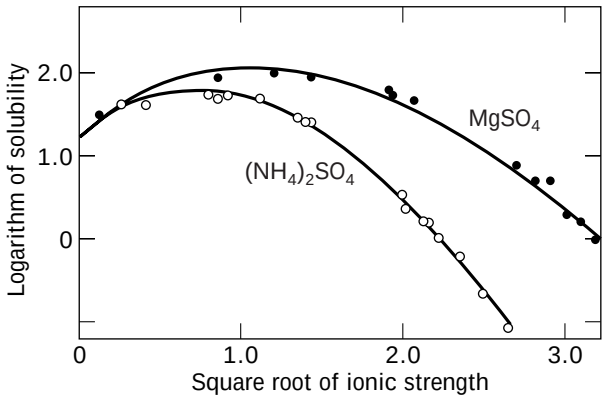


Figure 4. The solubility of the protein enolase at constant pH and temperature as a function of ionic strength produced by two separate salts. The two ends of the curves correspond to the salting-in and the salting-out regions of the solubility diagram.

ions and proteins becomes sufficiently intense, the protein molecules begin to self-associate in order to satisfy, by intermolecular interactions, their electrostatic requirements. Thus, dehydration, or the elimination and perturbation of solvent layers around protein molecules, reduces solubility. For a specific protein, the solubility minima are usually quite dependent on the pH, temperature, chemical composition of the precipitant, and properties of both the protein and the solvent (for more extensive discussions, see Refs. 38–42).

In the low ionic strength range, as seen in Figure 4, a phenomenon known as salting-in occurs in which the solubility of the protein increases as the ionic strength rises from zero. Protein solubility at low ionic strength is diminished by the removal of ions essential for satisfying the electrostatic requirements of the molecules. As ions are removed, and in this region of low ionic strength cations are most important (38–41), the protein molecules again balance their electrostatic requirements by forming interactions among themselves.

The salting-in effect can be applied in the opposite sense, that is, reducing ionic strength can serve as a crystallization tool. In practice, a protein is optimally solubilized at moderate ionic strength and then extensively dialyzed against distilled water. Many proteins, such as catalase, concanavalin B, immunoglobulins, and plant seed proteins, have been crystallized by this means (26,27,32,33).

Proteins may be forced from solution, at constant pH and temperature, by increases in ionic strength, or they can be crystallized or precipitated at constant ionic strength by changes in pH or temperature (38–41). This is illustrated by Figures 5 and 6. Reduction in solubility occurs because the charge distribution on a macromolecule, its surface features, or its conformation may change as a function of these variables (39). By virtue of its ability to inhabit a range of states, a protein may exhibit a number

Table 2. Common Precipitants Used in Macromolecular Crystallization

Salts	Organic solvents	Polymers
Ammonium or sodium sulfate	Ethanol	Poly(ethylene glycol) 1,000, 3,350, 6,000, 20,000
Lithium sulfate	Isopropanol	Jeffamine T
Lithium chloride	1,3-Propanediol	Polyamine
Sodium or ammonium citrate	2-Methyl-2,4-pentanediol	
Sodium or potassium phosphate	Dioxane	
Sodium or potassium or ammonium chloride	Acetone	
Sodium or ammonium acetate	Butanol	
Magnesium or calcium sulfate	Acetonitrile	
Cetyltriethyl ammonium salts	Dimethyl sulfoxide	
Calcium chloride	2,5-Hexandediol	
Ammonium nitrate	Methanol	
Sodium formate	1,3-Butyrolactone	
	Ethylene glycol 400	

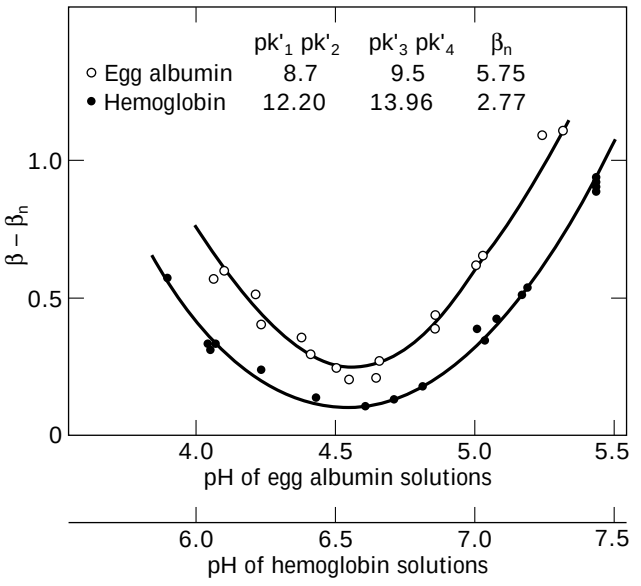


Figure 5. The solubility behavior of two proteins, hemoglobin and egg ovalbumin, at constant ionic strength as a function of pH. The solubility minima are often good targets for the crystallization of proteins by variation of pH.

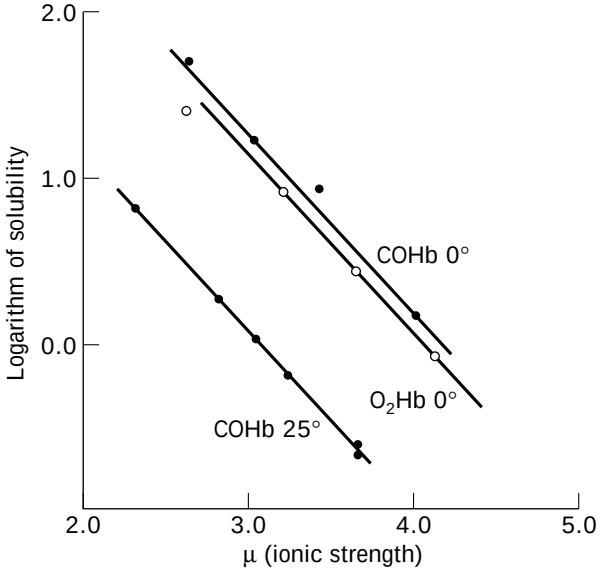


Figure 6. The solubility of hemoglobin in concentrated phosphate buffer at various temperatures.

of environmentally dependent solubility minima, and each of these minima may afford the opportunity for crystal growth.

Similar effects may be achieved as well by the slow addition to the mother liquor of some relatively mild alcohols or other organic solvents such as ethanol, acetone, or methylpentanediol. The only essential requirement for the precipitant is that at the specific temperature and pH of the experiment, the additive not adversely affect the structure and integrity of the protein. This is a stringent requirement and deserves some consideration. An organic solvent will compete, to some extent like salt, for water molecules. It also reduces the dielectric shielding capacity of intervening solvent. Reduction of bulk dielectric increases the effective strength of the electrostatic forces responsible for one protein molecule associating with another.

Polymers such as polyethylene glycol also serve to dehydrate macromolecules in aqueous solution, as do salts, and they furthermore alter the dielectric properties in a manner similar to organic solvents. They produce, however, an additional important effect. Polyethylene glycol perturbs the natural structure of the solvent and creates a more complex network having both water and itself as structural elements. A consequence of this restructuring of solvent is that macromolecules, particularly proteins, tend to be excluded, and phase separation is promoted (43,44). This is illustrated for several proteins in Figure 7.

Combinations of approaches using both precipitating agents, such as those seen in Table 2, and variation of physical parameters are frequently effective in promoting the crystallization of macromolecules. Increasing the concentration of a precipitating agent to a point just below supersaturation and then adjusting the pH or temperature to further reduce the solubility of the protein is a common approach. Modification of pH can be brought about quite well using the vapor diffusion technique (see later section) with volatile acids and bases such as acetic acid and ammonium hydroxide.

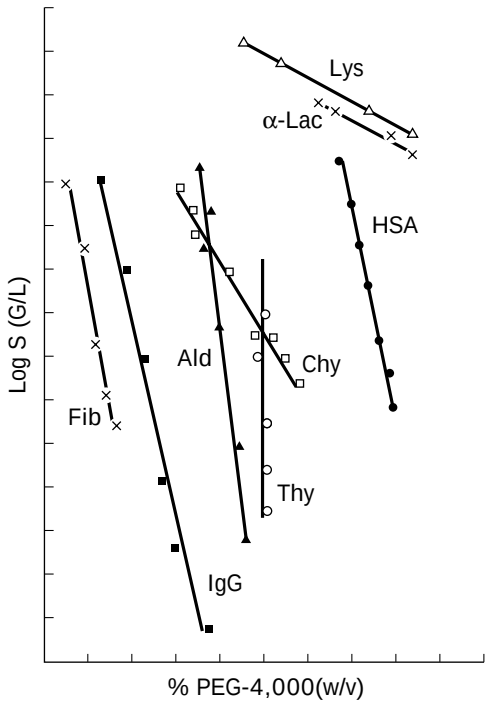


Figure 7. Solubility of various proteins in PEG-4,000. Measurements were made in 0.05 M potassium phosphate, pH 7.0, containing 0.1 M KCl. Human fibrinogen (×), human γ-globulin (■), lysozyme (△), α-lactalbumin (+), chymotrypsin (□), aldolase (▲), thyroglobulin (○), human serum albumin (•).

Because of the complexity of interactions between solute and solvent and the shifting character of the protein (45), the methods of crystallization must usually be applied over a broad set of conditions with the objective of discovering the particular minimum (or minima) that yield crystals. In practice, one determines the precipitation points of the protein at sequential pH values with a given precipitant, repeats the procedure at different temperatures, and then examines the effects of different precipitating agents and their concentrations.

METHODS FOR PROMOTING SUPERSATURATION

The oldest technique used for the crystallization of proteins and nucleic acids (see Table 1), with a history of more than 150 years (46), is what is commonly called batch crystallization (32,33,38,47–49). It is the simplest, most thoroughly tested, and when optimized, the most reliable of all crystallization methods. Lysozyme, for example, is most generally crystallized by the batch method. It relies on the mixing of a protein solution at undersaturated concentration with precipitant components. The latter alter the electrolyte properties of the solution to produce a mother liquor supersaturated with respect to protein. From such a solution the solid state may be expected to eventually form, and if the conditions and physical parameters have been chosen correctly, this solid state will be crystalline rather than amorphous. Alternatively, the physical state of an undersaturated protein solution may be altered by, for example, lowering or raising the temperature. The supersaturated solution produced as a result may yield crystals.

The batch method is attractive because of its inherent simplicity and reproducibility. It requires nothing more than the combination of two or more solutions into one and a period of time until nucleation commences and growth follows. The concept is appealing because it can be carried out with an absolute minimum of human intervention. Because the apparatus is nothing more than a container for a small quantity of fluid, which can be varied as circumstances require from a few microliters to hundreds of milliliters, it is fully amenable to interface with both simple and complex observational systems designed to monitor and record the process of protein crystal growth. Thus, the technique is very useful, for example, for inelastic light scattering studies of nucleation (50–57), or time-lapse photography investigations of protein crystal growth (58,59). It is particularly attractive as well for macromolecular crystallization experiments in space.

Batch crystallization requires no complex device to produce equilibration, either through the liquid or the vapor phase. It relies strictly on the energetics of the system—the activation barriers to nucleation—to regulate the process. Although it has not, in recent years, been extensively used to screen for protein crystallization conditions, this reflects more a prejudice of investigators and the popularity of other methods rather than any deficiency in the batch method itself. Because it was commonly used in the past with fairly large volumes of protein solution, investigators have mistakenly come to believe that it is inappropriate for small samples, which is untrue. It can even be carried out using microliter amounts in glass capillaries or under oil. It can be a useful screening technique as well as a good method for the preparation of large quantities of protein crystals once optimal values for parameters are established.

There are a number of ingenuous devices, procedures, and methods for dynamically altering the degree of saturation of a protein solution. Generally, these rely on the slow increase in concentration of some precipitant such as salt or polyethylene glycol. These same approaches can of course also be used for salting-in, modification of pH, and the introduction of materials that alter protein solubility. These techniques have been extensively reviewed elsewhere (32–36,60–63), and only a selection will be dealt with here.

Dialysis is familiar to nearly all biochemists as a means of modifying the components of a protein solution (32,33). As illustrated in Figure 8, the solution containing the protein is enclosed in a membrane casing or in a container having a semipermeable membrane partition. The membrane selectively allows the passage of small molecules and ions, but pore size prevents passage of the much larger protein molecules. The vessel or dialysis tube containing the protein is submerged in a larger volume of liquid that has the desired properties of pH, ionic strength, ligands, and so forth. With successive changes of the exterior solution and concomitant diffusion of small molecules and ions across the semipermeable membrane, the protein solution gradually acquires defined properties.

The dialysis approach, in some manifestation, has been used to crystallize many proteins (27,33). It was usually applied on a large scale, and only when substantial

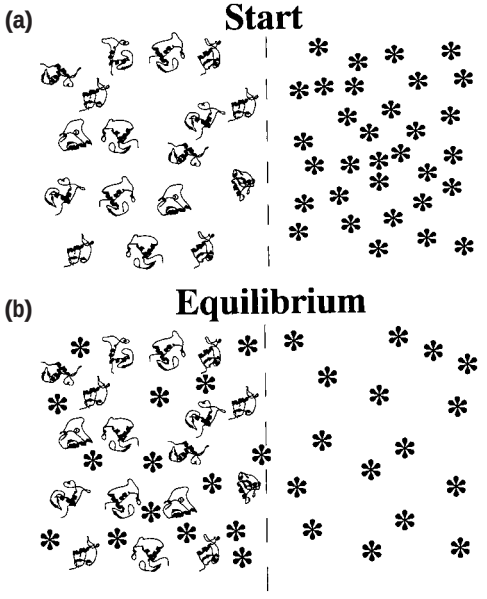


Figure 8. A schematic description of the dialysis process. (top) Protein molecules are separated from conventional small molecules or ions by a semipermeable membrane that prevents the macromolecules from passing. (bottom) The small molecules or ions have passed through the membrane and at equilibrium are equally distributed on its two sides.

amounts of protein were available. It has the advantage that liquid–liquid diffusion through a semipermeable membrane exposes the protein to a continuum of potential crystal-producing conditions without physically perturbing the mother liquor. Diffusion through the membrane is slow and controlled. Because the rate of exchange of substituents in the mother liquor is a function of the concentration gradient across the membrane, the nearer the system to equilibrium, the more slowly it changes.

This method has been adapted to much smaller amounts of protein by crystallographers who use almost exclusively microtechniques. These typically involve no more than 5 to 50 μl of protein solution in each trial. First described by Zeppenzauer et al. and Zeppenzauer (64,65) and subsequently modified and refined by numerous others, the method confines a protein solution within a glass capillary or the microcavity of a small plexiglass button. The cavity of the button or the ends of the microcapillary tube are closed off by a dialysis membrane. The whole arrangement, charged with protein solution, is then submerged in a much larger volume of an exterior liquid that is held within a closed vessel such as a test tube or vial. The microdyalysis method is shown schematically in Figure 9. If the exterior solution is at an ionic strength or pH that induces the mother liquor to become supersaturated, crystals may grow. If not, the exterior solution may be exchanged for another and the experiment repeated.

The dialysis buttons, seen in Figure 10, are particularly ingenious. Not only are they compact and easy to examine with a dissecting microscope, but they have a shallow groove about their waist. A section of moistened dialysis membrane is placed over the mother-liquor-filled cavity, and it is held securely in place by simply slipping a rubber O-ring over the top of the button and seating it in the groove. These buttons are in wide use and have proven quite successful.

The technique of liquid–liquid diffusion (Figs. 11 and 12), also known as free-interface diffusion, for the crystallization of macromolecules (25–27,62,63,66) is in principle and practice simple and straightforward. It relies on the direct interdiffusion of a precipitant solution into a protein solution across their common liquid–liquid interface. When the two solutions are juxtaposed, precipitant gradients form in the region of the interface and produce local supersaturation. This promotes the formation of crystal nuclei and subsequent growth, often to sizes useful for analysis.

The fundamental idea of liquid–liquid diffusion of a solvent or solute into a macromolecular solution is similar to dialysis. The method is particularly applicable in microgravity because the necessity of a membrane is eliminated and the two liquids can be maintained relatively stable with respect to one another for long periods of time. Because the diffusion rate of macromolecules is so much less than that of the small molecules of most precipitants, there is relatively little penetration of protein into the precipitant during the course of most experiments.

The technique of liquid–liquid diffusion has been applied in laboratories on earth to grow macromolecular crystals, and with some success. A search of the National Institute of Standards and Technology (NIST) database (26,27) shows 35 entries using this technique with another 193 ascribed to various kinds of dialysis. Direct liquid diffusion has not, however, been as extensively used in conventional laboratories as have some other techniques, such as batch or vapor diffusion, because of some inherent disadvantages. Significantly, most of these methodological deficiencies are relieved or eliminated in the absence of gravity.

On earth, density-driven convective currents produce immediate and continuous turbulence at the liquid–liquid interface. Transient mixing, if layering of solutions is used, or the necessity of a membrane, as in dialysis, lessens the methods' effectiveness. Sedimentation produces a particularly deleterious effect with the free-interface method because nuclei and small growing crystals, under the influence of gravity, fall from the liquid–liquid interface. Thus, as a crystal forms, gravity increasingly pulls it downward from the region of the solution where growth would most favorably continue. Gravity essentially removes growing crystals from an optimal growth environment and thereby imposes a self-limiting character on the process, in turn restricting crystal size.

Currently, the most widely used method for bringing about supersaturation in microdrops of mother liquid is vapor diffusion (26,27) (Fig. 13). This approach (32,33,35, 60,62,63,67) has provided a diversity of applications and may be divided into those procedures that use a "sitting

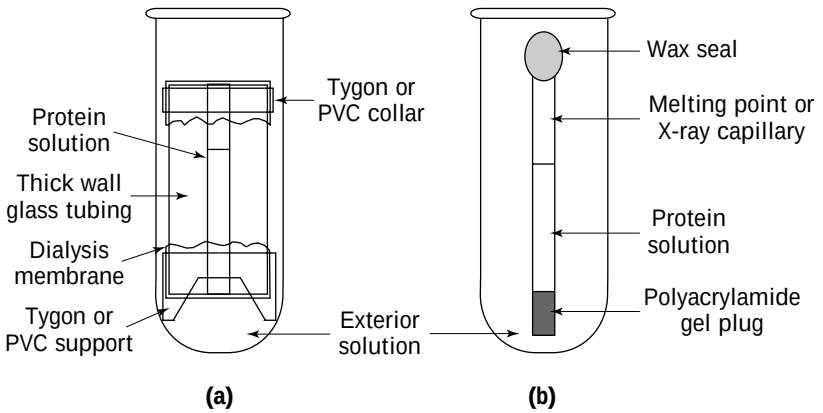


Figure 9. A microdialysis cell (**a**), for crystallization of small protein samples, is closed at one end by a section of dialysis membrane held in place by a Tygon ring that also serves as a support. (**b**) A similar arrangement using a thin glass capillary, but the membrane has been replaced by a polyacrylamide plug that is permeable to small molecules and ions.

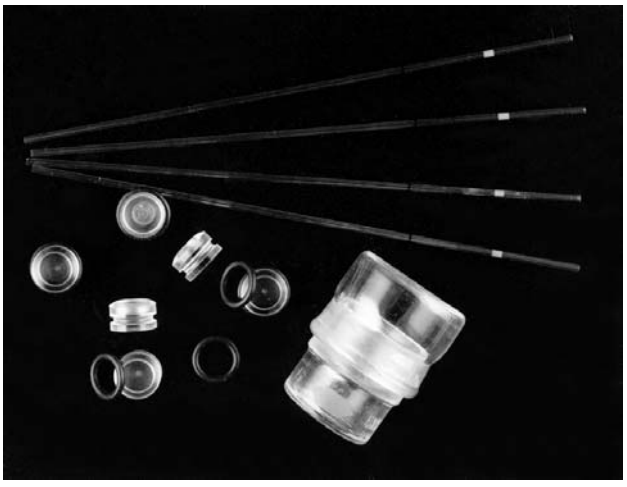


Figure 10. The thin glass capillaries shown here may be used for either batch crystallization or crystallization by dialysis with volumes in the microliter range. The dialysis buttons seen both face on and on edge can be closed by dialysis membrane held in place by the rubber O-rings. Once filled with protein solution they can be submerged in a larger volume of solution contained, for example, in the small weighing bottle.

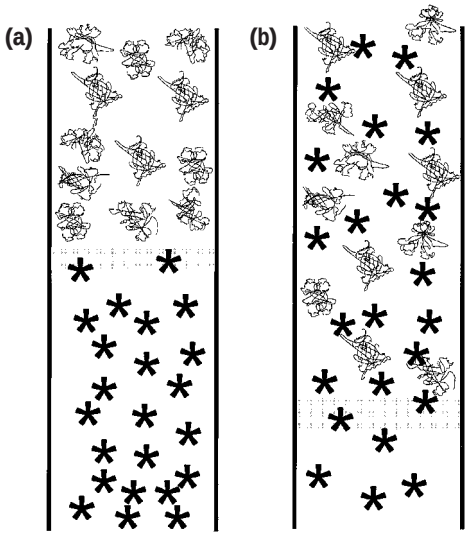


Figure 11. In free-interface diffusion the protein solution is (left) initially layered atop a precipitant solution, such as a highly concentrated salt. Diffusion of the small molecules into the protein solution (right) induces crystallization to occur.

drop" and those that use a "hanging drop." In either form, the method relies on the transport of water or some volatile agent between a microdrop of mother liquor, generally 2 to 25 μL volume, and a much larger reservoir solution of 0.75 to 25 mL in volume. Through the vapor phase, the droplet and reservoir come to equilibrium. Because the reservoir is of much larger volume, the final conditions at equilibrium are essentially those of the initial reservoir state. A variety of devices currently in use for protein crystal growth by vapor diffusion are shown in Figure 14.

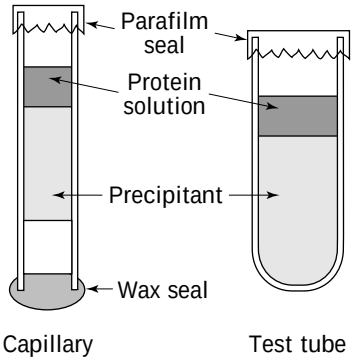


Figure 12. Diagram illustrating the free-interface diffusion technique for the crystallization of microliter amounts of protein solution on the left and milliliter amounts on the right. The method relies on the direct diffusion of precipitant and small molecules across the liquid–liquid interface.

In the vapor phase, water can be removed slowly from droplets of mother liquor, pH may be changed, or volatile components such as ethanol may be gradually introduced. As with those methods already described, the procedure may be carried out to further advantage at a number of different temperatures.

According to one popular procedure (32,33,67) (Fig. 15), droplets of 10 to 20 μL are placed in the nine wells of siliconized glass depression plates. The samples are then sealed in transparent containers, such as Pyrex dishes or plastic boxes that also hold reservoirs of 20 to 50 mL of the precipitating solution. The plates bearing the protein or nucleic acid samples are supported above the reservoirs by the inverted halves of disposable Petri dishes. Through the vapor phase, the concentration of salt or organic solvent in the reservoir equilibrates with that in the sample. In the case of salt precipitation, the droplet of mother liquor must initially contain a level of precipitant lower than the reservoir, and equilibration proceeds by diffusion of water from the droplet into the reservoir. This holds true as well for nonvolatile organic solvents, such as methylpentanediol and for polyethylene glycol. In the case of volatile precipitants, none need be added initially to the microdroplet, because distillation and equilibration proceed in the opposite direction.

The method described here has the advantage that it requires only microliter amounts of material and is well suited for screening large numbers of conditions. A disadvantage is that all the samples in a single container must be equilibrated against the identical reservoir solution. It does, however, permit some flexibility in varying conditions once the samples have been dispensed, by modification of the concentration or pH of precipitants in the reservoir. When clear plastic boxes are used, large numbers of samples can be quickly inspected under a dissecting microscope and conveniently stored.

The disadvantage of identical reservoir conditions throughout a single container was overcome by the introduction of plastic plates specifically designed for protein crystallization. One of these, sponsored by the American

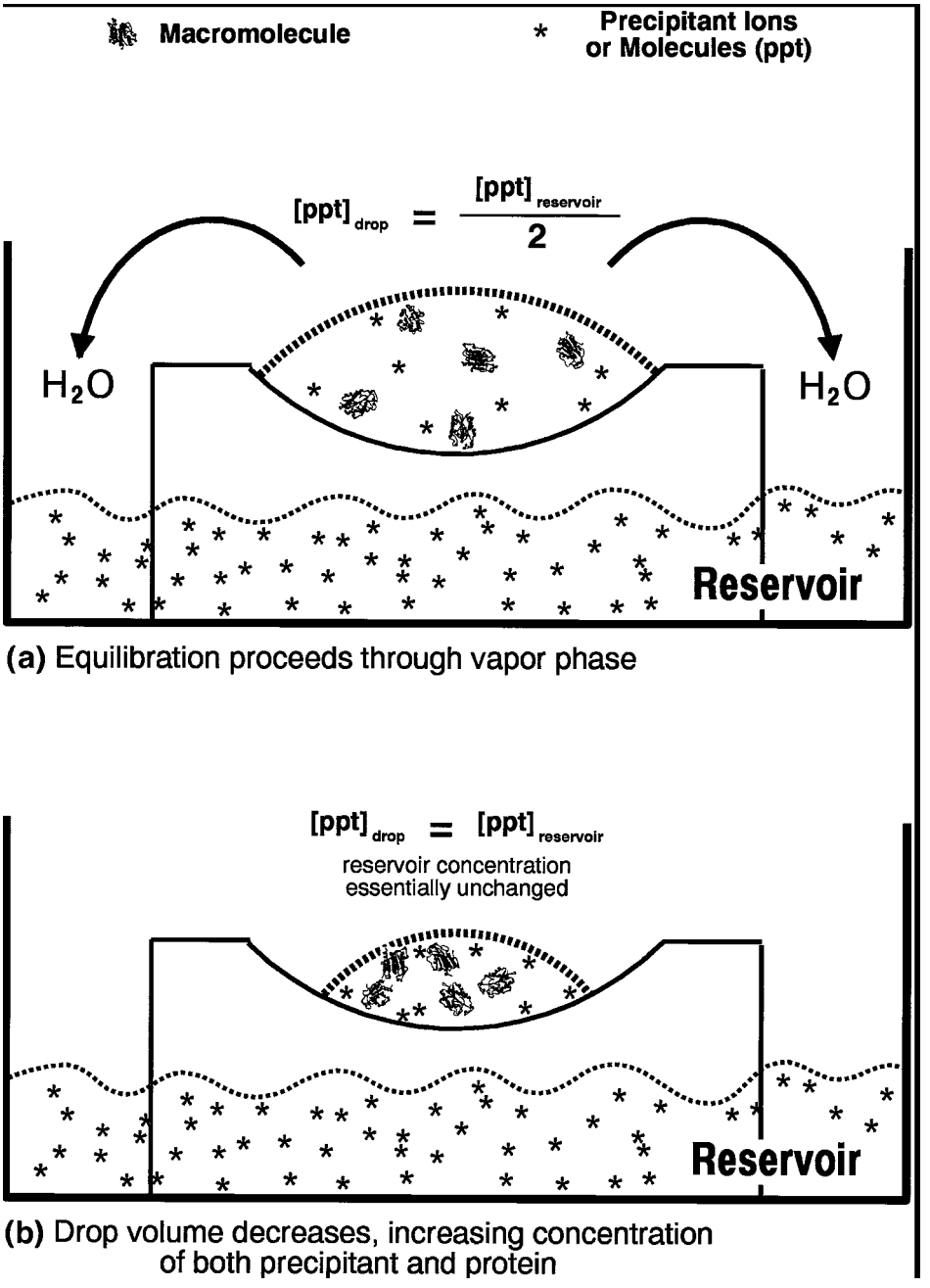


Figure 13. In vapor diffusion, a small drop of mother liquor containing protein and a low concentration of precipitant is, in (**a**), initially exposed to a much larger reservoir solution containing a high concentration of precipitant. At equilibrium in (**b**), through loss of water and reduction in volume, the mother liquor achieves a precipitant concentration approximately equal to that in the reservoir, thereby promoting crystallization of the protein to occur.

Crystallographic Association, is a plastic plate having accommodation for 15 protein samples. Each chamber has a separate reservoir compartment and the mother liquor microdroplet may be either suspended from the underside of a glass cover slip, as in the hanging drop method described later, or sandwiched between two glass cover slips. Sealing of the chambers from air requires silicon grease or oil between cover slips and the plastic rims of the chambers. With these plates, the optical properties are very good but equilibration tends to be slow, and the setup is quite tedious.

A second crystallization plate (68), also seen in Figure 14, is supplied by Charles Supper. With these plates, the drop sits atop a clear support post that protrudes upward from a circular moat containing the reservoir solution. The chambers can be rapidly and conveniently sealed from air by transparent plastic tape pressed over the upper surface of the plate after the reservoirs have been filled and the drops of mother liquor dispensed. Equilibration, as with the other plate, is through the vapor phase. Although the optical properties are less favorable with these devices, they are convenient and compact and can be used for rapid screens of crystallization conditions.

The hanging drop procedure (Fig. 16) also uses vapor phase equilibration, but with this approach a microdroplet of mother liquor (as small as 2 μL) is suspended from the underside of a microscope cover slip, which is then placed over a small well containing a milliliter or less of the pre-

Figure 14. Four examples of trays, plates, or combinations used for crystallization by vapor diffusion. All are commercially available and can be used for either the hanging-drop or the sitting-drop approach.

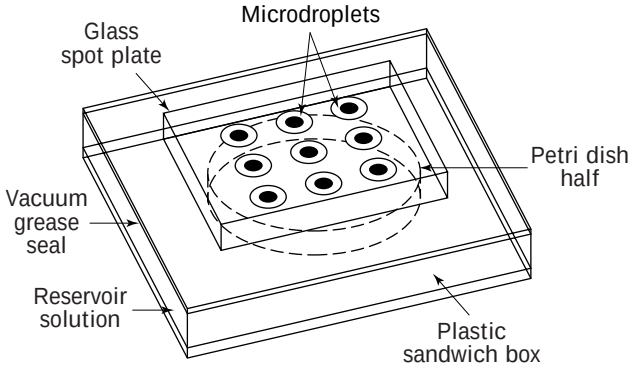


Figure 15. Schematic diagram of a nine-well depression spot plate enclosed in a plastic sandwich box used for the crystallization of macromolecules by the vapor diffusion method. Microdrops of 10 to 25 μL can be placed in the depressions with 25-mL reservoirs of precipitating solution in the bottom of the box. Equilibration proceeds through the vapor phase.

cipitating solution. The wells are most conveniently supplied by disposable plastic tissue culture plates that have 24 wells with rims that permit sealing by application of silicone vacuum grease around the circumference. The hanging drop technique can be used both for the optimization of conditions and for the growth of large single crystals.

Although the principle of equilibration with both the sitting drop and the hanging drop is essentially the same, they frequently do not yield the same results even though the reservoir and protein solutions may be identical. Presumably because of geometrical differences in the apparatus used to achieve equilibration, the path to supersaturation is different, even though the end point may be the same. In some cases there are striking differences in the degree of reproducibility, final crystal size, morphology, or time required for growth. These observations emphasize the point that in crystal growth the route leading to supersaturation and the kinetics of the process may be as

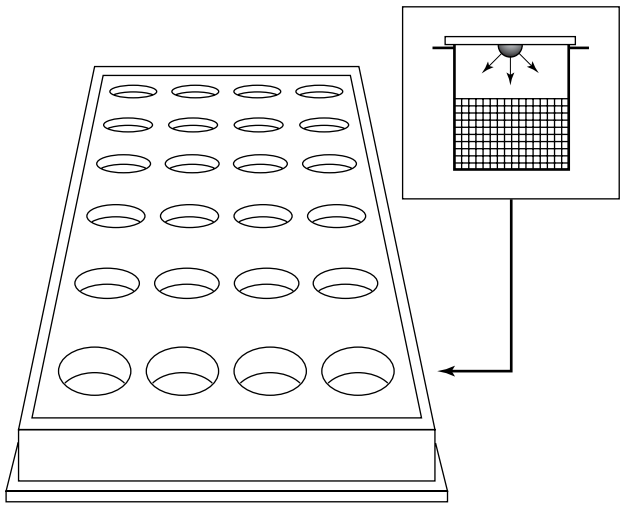


Figure 16. The standard configuration for hanging drop protein crystallization experiments is shown here. A tissue culture plate provides 24 wells for reservoirs of about 0.5 to 0.75 mL. Each well can be covered by a glass slip with a drop of protein solution hanging from its underside. The drop of 2 to 20 μL equilibrates with the reservoir solution over time through the vapor phase, causing precipitant concentrations to increase in the drop and induce crystallization.

important as the end point achieved. There are nearly limitless variations on these themes for liquid–liquid or vapor diffusion, ranging from the very simple to the unnecessarily complex. You need only use a bit of imagination to devise a novel approach for yourself.

With some exceptions, the methodologies used for protein crystallization, most of those described above, have been devised for the growth of crystals from very small volumes of mother liquor. Indeed, reducing the amount of material needed has been a principal objective in X-ray crystallography laboratories for some time. Crystals can now be obtained from only 2 to 4 μL of solution, and entire screening matrixes (see later discussion) can be explored using only 2 to 5 mg of protein or nucleic acid.

Two major applications of protein crystallization, for formulations, and for downstream processing of macromolecules do not, however, benefit from reduction in scale. On the contrary, they profit from very large volumes and high yields of the crystalline product. As a consequence, the microtechniques described earlier may represent model approaches, but they are, in themselves, quite inappropriate to bulk processes.

Large-scale, industrial crystallizations are, at least currently, limited to methods based essentially on batch crystallization in combination with manipulation of pH and temperature. In addition, many chemical compounds that might be used as precipitants or additives in laboratory settings cannot be used in many cases in bulk processes because of expense, hazard, or the creation of unacceptable by-products (7). Although this might appear severely limiting, remarkable success has been achieved using only these simple tools. The large number of biological macromolecules in crystalline form and produced in kilogram (or

larger) amounts available on the commercial market attests to this.

The restriction on techniques is compensated to a great extent in industrial processes by other factors. For example, the objective protein is usually in relatively high concentration and comprises a significant fraction of the total protein present. It is usually produced by genetically engineered or overproducing strains of simple microorganisms; hence, it does not suffer from many of the sources of microheterogeneity, seen in Table 3, that afflict many natural products. In addition, crystals produced in large-scale processes are usually very small crystals of marginal perfection that would be unsuitable for analytical work. Such crystals may, however, be entirely appropriate for the intended purpose of purification and product formulation.

THE CRYSTALLIZATION OF PROTEINS BY VARIATION OF PH OR TEMPERATURE

One of the most useful techniques for producing a supersaturated solution is adjustment of pH to values where the protein is substantially less soluble. This may be done in the presence of a variety of precipitants so that a spectrum of possibilities are created whereby crystals might form. The gradual alteration of pH is particularly subtle because it may be accomplished by a variety of gentle approaches that do not otherwise perturb the system. Although microdialysis is probably equally suitable, more success with pH has been achieved by the vapor diffusion method, using "sitting" microdroplets on spot plates or in one of the plastic plates available for protein crystallization. The ambient salt, effector, or buffer conditions are established before dispensing the microdroplets in the depressions on the plate. The pH is then slowly raised or lowered by adding a small amount of volatile acid or base to the reservoir. Diffusion of the acid or base then occurs from reservoir to sample, just as for a volatile precipitant.

If the pH is to be raised, for example, a small drop of concentrated ammonium hydroxide can be added to the reservoir; a drop of acetic acid may be used to lower it. The pH can also be gradually lowered over a period of days by simply placing a tiny chip of dry ice in the reservoir. The liberated CO_2 diffuses and dissolves in the mother liquor to form weak carbonic acid. When a specific pH end point is required, the mother liquor may be buffered at that point and then moved significantly away by addition of acid or base. The microdroplets of mother liquor may then be returned to the buffer point by addition of an appropriate volatile acid or base to the reservoir.

As shown in Figure 4, individual proteins display distinct solubilities at varying levels of salt concentration and can be selectively separated from solution as amorphous precipitate or crystals. At very low ionic strengths where salting-in occurs, the solubility of the protein increases with the ionic strength because of a decrease in its activity coefficient. The salting-out effect, which Hofmeister (39,41) suggested to be caused by a reduction of the chemical activity of water by salt, occurs as the ionic strength increases. Because ionic strength is a function of the second power of the valence, multivalent ions are most efficient. In the salting-out region of a protein's solubility profile, the logarithm of protein solubility as a function of the ionic strength is linear and can be expressed by

$$\log S = \beta - K_s(I/2)$$

where I is the ionic strength in moles per kilogram of water, S is the solubility of the protein in grams per kilogram of water, and K_s is a constant. K_s serves as a measure of the slope of the solubility extrapolated to zero ionic strength and is the logarithm of the solubility at that point (39).

In principle, the precipitation or crystallization of a protein is effected by manipulating the factors that appear in the above equation. One can crystallize a macromolecule by increasing the ionic strength at constant pH and temperature, a K_s fractionation procedure, or it can be brought about at constant ionic strength by manipulating the pH and temperature, which in most cases strongly influence solubility.

With all techniques, the objective is to produce a supersaturated protein solution by very gradual and nonperturbing adjustment of conditions. An example is shown in Figure 17. In practice, a solubility minimum is determined for the protein at some specific pH, temperature, and ionic strength. The protein solution is set initially at the requisite salt concentration for this solubility minimum, but either the pH or temperature is set initially to maintain protein solubility. The temperature is then slowly relaxed to minimize protein solubility, or the pH is gradually equilibrated via the liquid or vapor phase to an appropriate value.

In the case of temperature shifts, the protein is generally raised to a higher temperature, as, for example, with squash globulin, insulin, or excelsin (41,42,50,52,69,70), a

Table 3. Sources of Microheterogeneity

Presence, absence, or variation in a bound prosthetic group, substrate, coenzyme, inhibitor, or metal ion
Variation in the length or composition of the carbohydrate moiety of a glycoprotein
Proteolytic modification of the protein during the course of isolation or crystallization
Oxidation of sulfhydryl groups during isolation
Reaction with heavy metal ions during isolation or storage
Presence, absence, or variation in post-translational side-chain modifications such as methylation, amidiation, and phosphorylation
Microheterogeneity in the amino or carboxy terminus or modification of termini
Variation in the aggregation or oligomer state of the protein caused by association or dissociation
Conformational instability caused by the dynamic nature of the molecule
Microheterogeneity caused by the dynamic nature of the molecule
Partial denaturation of sample
Genetically different animals, plants, or microorganisms that make up the source of protein preparations
Bound lipid, nucleic acid, or carbohydrate material or substances such as detergents used in the isolation

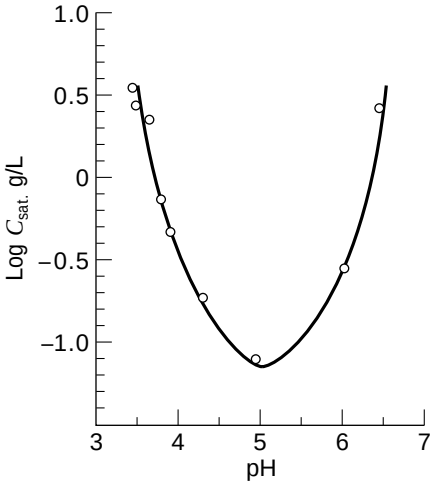


Figure 17. The solubility of insulin falls dramatically in the immediate neighborhood of pH 5, providing a ready means for creation of a supersaturated solution. The protein is simply dissolved to high concentration at pH 7 or higher, and the pH subsequently lowered through addition of H_3O^+ by vapor diffusion or dialysis.

process that increases solubility. The solution is then cooled slowly to where the protein is no longer soluble. This can be achieved by placing the warm solution in an insulated or thermally controlled container or by placing the sample in a Dewar flask. The approach to supersaturation thus occurs at an almost imperceptible rate. Exactly the same technique can be applied if the protein is more soluble at cold than warm temperature. Under these circumstances, the protein solution is cooled under established ionic strength conditions and then allowed to warm slowly. This approach was used, for example, to grow crystals of DNase by Kunitz (71) and is the basis of Jakoby's sequential extraction procedure using decreasing concentrations of ammonium sulfate (72).

The use of pH gradients in space and time has been successful in many investigations (26,27); they can be used in association with direct liquid dialysis or with vapor diffusion methods as well. They can, furthermore, be used in conjunction with a host of different precipitating agents. Temperature can be used in an analogous manner (73).

Proteins and nucleic acids are structurally flexible and are believed to alternate between several conformational states. In particular, they may assume substantially different conformations when they have bound coenzyme, substrate, or some other ligand. As a consequence, a protein with bound effector may exhibit different solubility properties than the nonliganded form. In addition, if multiple conformational states are available, the presence of effector may be used to select for only one of these, thereby imposing a degree of conformity of structure that would otherwise be absent.

The addition of ligands can sometimes be used to induce supersaturation and crystallization in those cases where its binding to the protein produces solubility differences under a given set of ambient conditions. The effector may be slowly and gently combined with the protein, for example, by dialysis, so that a resulting complex becomes supersaturated. The addition of ligands, substrates, and other small molecules has seen widespread use in protein crystallography, because it provides useful alternatives if the apoprotein itself cannot be crystallized.

No matter what the ultimate objective of macromolecular crystallization is, it is first necessary to (*1*) identify a set of physical and chemical conditions that yield some kind of crystals, (*2*) optimize those conditions, and (*3*) reproducibly apply the conditions according to reliable procedures that will consistently produce the requisite crystals. Thus, the initial step consists of screening large numbers of possible precipitants, buffers, effectors, pHs, temperatures, etc. The second is a careful screening using a finely sampled matrix of conditions focused on those that yielded crystals, and the third is on devising reproducible procedures for application. Clearly, the first stage is the greatest hurdle that must be overcome, but it is not necessarily the step demanding the most effort. Table 4 is included here, not to alarm or intimidate the novice, but to alert those interested in macromolecular crystal growth of the complexities and the many factors that may have an impact on a crystal growth investigation. Fortunately, only a few of these factors may be simultaneously operative for a particular problem; the trick is to discover which they are for your specific case.

How does one begin the process, that is, how do you discover those initial conditions that yield crystals and provide starting points? Over the years, a vast number of reports and observations entered the literature, without much form or even clear purpose (46). Attempts were made to sort and classify (48,74); to shape ad hoc experiments into more rational approaches (26,32,34,55,75); to gather,

Table 4. Factors That Do or Could Affect Protein Crystal Growth

pH and buffer
Ionic strength
Temperature and temperature fluctuations
Concentration and nature of precipitant
Concentration of macromolecule
Purity of macromolecules (see Table 3)
Additives, effectors, and ligands
Organism source of macromolecule
Substrates, coenzymes, inhibitors
Reducing or oxidizing environment
Metal and other specific ions
Rate of equilibration and rate of growth
Surfactants or detergents
Gravity, convection, and sedimentation
Vibrations and sound
Volume of crystallization sample
Presence of amorphous or particulate material
Surfaces of crystallization vessels
Proteolysis
Contamination by microbes
Pressure
Electric and magnetic fields
Handling by investigator and cleanliness
Viscosity of mother liquor
Heterogeneous or epitaxial nucleating agents

collate, and systematize the experiences of thousands of investigators into rationally constructed databases (27,37); and finally to extract from those bases arrays of conditions most likely to yield protein and nucleic acid crystals.

These sets of conditions form the foundation for the crystal screens (57,76–82), now commercially available (Hampton Research), that have virtually revolutionized the initial search for crystallization conditions. Although some important macromolecules continue to slip through these still imperfect nets, the number of successes is impressive. Improved screening techniques alone are to a great extent responsible for the explosion of new crystals, grown not only by crystallographers, but by the myriad of biochemists and molecular biologists whose need for structure is acute. The search for novel, and perhaps even better, precipitants and conditions for crystal growth continues, and we can reasonably expect even more useful developments in this area.

With arrays of high-probability conditions that need be tested against no more than a few milligrams of protein or nucleic acid, attention has focused on the macromolecules themselves. Three principles have clearly come forward: purity, stability, and solubility. Failure in crystallization can probably be attributed, in general, to one or more of these.

Homogeneity has always been recognized as important, though not always essential. The trick is to recognize when it is the crucial factor and direct one's efforts there. Important advances have been made with regard to purification. Not only are there now more powerful techniques and instruments available (36,83), but crystallography has benefited to an extraordinary degree from the refinements of recombinant DNA technology and the development of high-yield expression systems (84–86). With the introduction of endogenous purification tools such as the histidine tag or glutathione sulfur transferase (GST) chimers (87), even greater availability of ultrapure crystallizable protein seems on the horizon. Recombinant systems provide us not only with pure, native proteins, but a way to introduce mutations and thereby create almost limitless alternatives should the native molecule prove unyielding. They offer systematic approaches as well to the intelligent engineering of solubility properties and stability.

Chemical and conformational stability of the target macromolecule, too, has long been recognized and given serious attention. Ligands are used to lock discrete conformations, disulfides introduced by mutation or chemical modifications are made to reduce mobility (88), proteins and ribosomes from thermophilic organisms are used for their natural stability, and proteolytic or mutational truncations are made to remove naturally flexible domains or elements (36). We are only beginning to understand how to manipulate conformation to enhance crystallization properties (89), and this area also promises much for the future, particularly for challenging topics such as antibodies, multiple domain enzymes, and multimolecular complexes.

Perhaps the principle least recognized but now receiving the most attention is solubility, meaning here not only dissolution of macromolecules in a liquid, but in a disaggregated, monodisperse molecular form. Indeed, we have come to recognize aggregation and polydispersity as a principal reason why many macromolecules cannot be crystallized. This includes not only membrane and lipophilic proteins (90) but macromolecules that have a strong propensity to aggregate, such as antibodies (91), hormone polypeptides, and proteins whose purpose is ordered aggregation, such as those from viral capsids.

Extending observations that nonionic detergents promote the growth of hydrophilic (92) as well as membrane proteins (90,93), light-scattering studies have confirmed the salient importance of monodispersity in macromolecular crystallization (50,53,78,94,95). The current consensus among those involved in crystallization research is that if by some means monodispersity can be ensured, then there is better than an 80% chance that crystals will be obtained. To promote monodispersity, the common approaches are to incorporate detergents in crystallization trials or to crystallize at low protein concentrations, at higher temperatures, or in the presence of chaotropic agents. This is an area that continues to be under active investigation.

Additional factors have emerged from recent crystal growth research and from the empirical results generated by the current revolution in crystallography. Among them are the importance of surfaces and their promotion of heterogeneous nucleation (96–100), the value of unique environments such as microgravity (101–103), gels (104–106), thin capillaries (107), the broad range of polymeric precipitants useful in growing crystals (108,109), and new approaches to solving the problem of persistent microcrystals (110).

BIBLIOGRAPHY

1. A.S. Myerson, *Handbook of Industrial Crystallization*, Butterworth-Heinemann Series in Chemical Engineering, Stoneham, Mass., 1993.
2. S. Takamatsu and D.D.Y. Ryu, *Biotechnol. Bioeng.* **32**, 184–192 (1988).
3. H. Takeda, T. Hidano, and S. Kinoshita, *J. Ferment. Bioeng.* **80**, 492–498 (1995).
4. M.J. Akers, and T. Miyakawa, *Pharm. Res.* **11**, S133, PT6012 (1994).
5. N. Inaba, S. Ayano, Y. Ozaki, M. Mitake, H. Maeda, and Y. Ifuku, *J. Jpn. Soc. Food Sci. Technol.* **40**, 833–840 (1993).
6. B. Cudney and S. Patel, *Am. Biotechnol. Lab.* **12**, 42 (1994).
7. M.D. Weiss, *Gen. Eng. News* **8**, 15 (1995).
8. M.M. Long, L.S. DeLucas, C. Smith, M. Carson, K. Moore, M.D. Harrington, D.J. Pillion, S.P. Bishop, W.M. Rosenblum, R.J. Naumann, A. Chait, J. Prahl, and C.E. Bugg, *Microgravity Sci. Technol.* **7**, 196–202 (1994).
9. M. Butler, *Biotechnology Series Animal Cell Technology: Principles and Products*. Taylor and Francis Inc., Philadelphia, Pa., Open University Press, 1991.
10. W. Schiweck and A. Skerra, *Proteins: Struct. Funct. Genet.* **23**, 561–565 (1995).
11. J. Fukumoto, M. Iwai, and Y. Tsujisaka, *J. Gen. Appl. Microbiol.* **9**, 353–361 (1963).
12. Y. Tsujisaka, M. Iwai, and Y. Tominaga, *Agr. Biol. Chem.* **37**, 1457–1464 (1973).

13. W. Buckel, U. Dorn, and R. Semmler, *Eur. J. Biochem.* **118**, 315–321 (1981).
14. M. Mack, K. Bendrat, O. Zelder, E. Eckel, D. Linder, and W. Buckel, *Eur. J. Biochem.* **226**, 41–51 (1994).
15. A. Kitazono, T. Yoshimoto, and D. Tsuru, *J. Bacteriol.* **174**, 7919–7925 (1992).
16. J.E. Pitts, J.M. Uusitalo, D. Mantafounis, P.G. Nugent, D.D. Quinn, P. Orprayoon, and M.E. Penttila, *J. Biotechnol.* **28**, 69–83 (1993).
17. T. Yoshimoto, T. Shimoda, A. Kitazono, T. Kabashima, K. Ito, and D. Tsuru, *J. Biochem.* **113**, 67–73 (1993).
18. H. Gros, *Abstracts of Papers, American Chemical Society* **21**, Biot 114 (1996).
19. M.H. Heng, N.T. Becker, E.L. Braunstein, R.C. Fewkes, and H. Gros, *Abstracts of Papers, American Chemical Society* **21**, Biot 177 (1996).
20. R.G. Harrison and L.P. Nelles, *16-Fermentations* **113**, 585 (1990).
21. A. McPherson, *Sci. Am.* **260**, 62–69 (1989).
22. A. McPherson, A.J. Malkin, and Y.G. Kuznetsov, *Structure* **3**, 759–768 (1995).
23. A.J. Malkin, Y.G. Kuznetsov, T.A. Land, J.J. DeYoreo, and A. McPherson, *Nat. Struct. Biol.* **2**, 956–959 (1995).
24. E.A. Wood, *Crystals and Light: An Introduction to Optica Crystallography*, 2nd ed., Dover, New York, 1977.
25. J.M. Cowley, *Diffraction Physics*, 2nd ed., North Holland, Amsterdam, 1984.
26. G.L. Gilliland, *J. Cryst. Growth* **90**, 51–59 (1988).
27. G.L. Gilliland, M. Tung, D.M. Blakeselee, and J.E. Ladner, *Acta. Crystallogr.* **D50**, 408–413 (1994).
28. M. Frey, *Acta Crystallogr.* **D50**, 663–666 (1994).
29. A.A. Chernov, *Modern Crystallography III: Crystal Growth*, Springer-Verlag, New York, 1984.
30. H.E. Buckley, *Crystal Growth*, Wiley, London, 1951.
31. W.K. Burton, N. Cabrera, and F.C. Frank, *Philos. Trans. R. Soc. London, Ser. A* **243**, 299 (1951).
32. A. McPherson, in D. Glick ed., *Methods of Biochemical Analysis*, vol. 23, Academic Press, New York, 1976, pp. 249–345.
33. A. McPherson, *The Preparation and Analysis of Protein Crystals*, Wiley, New York, 1982.
34. P.C. Weber, *Adv. Protein Chem.* **41**, 1–36 (1991).
35. B. Lorber and R. Giege, *Crystallization of Nucleic Acids and Proteins: A Practical Approach*, Oxford Univ. Press, Oxford, U.K., 1992.
36. R. Giege, B. Lorber, and A. Theobald-Dietrich, *Acta Crystallogr.* **D50**, 339–350 (1994).
37. A.M. Hassell, T.J. Harrocks, E.H. Dashman, and A. Mistry, *Acta. Crystallogr.* **D50**, 459–465 (1994).
38. J.H. Northrop, M. Kunitz, and R.M. Herriott, *Crystalline Enzymes*, Columbia Univ. Press, New York, 1948.
39. E.J. Cohn and J.D. Ferry, in E.J. Cohn and J.T. Edsall eds., *Proteins, Amino Acids, and Peptides*, Reinhold, New York, 1950.
40. R. Czok and T. Bucher, *Adv. Protein Chem.* **15**, 315–415 (1960).
41. F. Hofmeister and Z. Hoppe-Seylers, *Physiol. Chem.* **14**, 165 (1890).
42. S.N. Timasheff and T. Arakawa, *J. Cryst. Growth* **90**, 39–46 (1988).
43. D.H. Atha and K.C. Ingham, *J. Biol. Chem.* **256**, 12108–12117 (1981).
44. J.C. Lee and L.L.Y. Lee, *J. Biol. Chem.* **256**, 625–631 (1981).
45. T. Arakawa and S.N. Timasheff, *Methods Enzymol.* **114**, 49–77 (1985).
46. A. McPherson, *J. Cryst. Growth* **110**, 1–10 (1991).
47. K. Bailey, *Nature*, **145**, 934 (1940).
48. K. Bailey, *Trans. Faraday Soc.* **38**, 186 (1942).
49. J.B. Sumner and G.F. Somers, *The Enzymes*, Academic Press, New York, 1943.
50. A.J. Malkin, J. Cheung, and A. McPherson, *J. Cryst. Growth*, **126**, 544–554 (1993).
51. Z. Kam, A.B. Shore, and G. Feher, *J. Mol. Biol.* **23**, 539–555 (1978).
52. G. Feher and Z. Kam, *Methods Enzymol.* **114**, 77–112 (1985).
53. E.T. Baldwin, K.V. Crumly, and C.W. Carter, *Biophys. J.* **49**, 47–48 (1986).
54. S.D. Durbin and G. Feher, *J. Cryst. Growth* **76**, 383–392 (1986).
55. C.W. Carter, E.T. Baldwin, and L. Frick, *J. Cryst. Growth* **90**, 60–73 (1988).
56. L. Betts, L. Frick, R. Wolfenden, and C.W. Carter, *J. Biol. Chem.* **264**, 6737–6740 (1989).
57. W. Kadima, A. McPherson, M.F. Dunn, and F.A. Jurnak, *Biophysical J.* **57**, 125–132 (1990).
58. S. Koszelak and A. McPherson, *J. Cryst. Growth* **90**, 340–343 (1989).
59. S. Koszelak, D. Martin, J. Ng, and A. McPherson, *J. Cryst. Growth*, **10**, 177–181 (1991).
60. T.L. Blundell and L.N. Johnson, *Protein Crystallography*, Academic Press, New York, 1976.
61. A. McPherson, *Methods Enzymol.* **114**, 112–120, (1985).
62. A. McPherson, in H. Michel ed., *Crystallization of Membrane Proteins*. CRC Press, Boca Raton, Fla., 1991, pp. 1–52.
63. A. McPherson, *Eur. J. Biochem.* **189**, 249–345 (1990).
64. M. Zeppenzauer, H. Eklund, and E. Zeppenzauer, *Arch. Biochem. Biophys.* **126**, 564 (1968).
65. M. Zeppenzauer, *Methods Enzymol.* **22**, 253 (1971).
66. B.A. Weber and P.E. Goodkin, *Arch. Biochem. Biophys.* **141**, 489–498 (1970).
67. A. Hampel, M. Labanauskas, P.G. Conners, L. Kirkegard, U.L. RajBhandary, P.B. Sigler, and R.M. Bock, *Science* **162**, 1384 (1968).
68. D. Morris, C.Y. Kim, and A. McPherson, *Biotechniques* **7**, 522–527 (1989).
69. G.E. Rindone, *Materials Processing in the Reduced Gravity Environment of Space*, North-Holland, New York, 1982.
70. A.J. Malkin and A. McPherson, *J. Cryst. Growth* **133**, 29–37 (1993).
71. M. Kunitz, *J. Gen. Physiol.* **35**, 423 (1952).
72. W.B. Jakoby, *Anal. Biochem.* **26**, 295 (1968).
73. F. Rosenberger and E.J. Meehan, *J. Cryst. Growth* **90**, 74–78 (1988).
74. E.T. Reichert and A.P. Brown, *The Differentiation and Specificity of Corresponding Proteins and Other Vital Substances in Relation to Biological Classification and Organic Evolution: The Crystallography of Hemoglobins*, Carnegie Ins., Washington, D.C., 1909.
75. C.W. Carter, Jr. and C.W. Cater, *J. Biol. Chem.* **254**, 12219–12223 (1979).
76. A. McPherson, *J. Cryst. Growth* **122**, 161–167 (1992).
77. R. Cudney, S. Patel, K. Weisgraber, Y. Newhouse, and A. McPherson, *Acta Crystallogr.* **D50**, 414–423 (1994).

78. A. D'Arcy, *Acta Crystallogr.* **D50**, 469–471 (1994).
79. P.D. Shaw Stewart and M. Khimasia, *Acta Crystallogr.* **D50**, 441–442 (1994).
80. E.A. Stura, A.C. Satterthwait, J.C. Calvo, D.C. Kaslow, and I.A. Wilson, *Acta Crystallogr.* **D50**, 448–455 (1994).
81. L.J. Harris, E. Skaletsky, and A. McPherson, *Proteins: Struct., Funct., Genet.* **23**, 285–289 (1995).
82. J. Jancarik and S.H. Kim, *J. Appl. Crystallogr.* **24**, 409–411 (1991).
83. R.F. Boyer, *Modern Experimental Biochemistry*, Benjamin/Cummings, New York, 1993.
84. B. Nilsson, G. Forsberg, T. Moks, M. Hartmanis, and M. Uhlén. *Curr. Opin. Struct. Biol.* **2**, 569–575 (1992).
85. H. Ago, N. Habuka, J. Kataoka, M. Furuno, H. Tsuge, M. Noma, M. Miyano, B.C. Wang, and N.H. Xuong, *Acta Crystallogr.* **D50**, 355–360 (1994).
86. A.R. Rees, M.J.E. Sternberg, and R. Wetzel, *Protein Engineering: A Practical Approach*, IRL Press, Oxford Univ. Press, New York, 1992.
87. G.G. Privé, G.E. Verner, C. Weitzman, K.H. Zen, D. Eisenberg, and H.R. Kaback, *Acta Crystallogr.* **D50**, 375–379 (1994).
88. I. Rayment, W.R. Ryppniewski, K. Schmidtbase, R. Smith, *Science* **261**, 50–58 (1993).
89. F. Dyda, A.B. Hickman, T.M. Jenkins, A. Engelman, R. Craigie, and D.R. Davies, *Science* **266**, 1981–1986 (1981).
90. H. Michel, in H. Michel ed., *Crystallization of Membrane Proteins*, CRC Press, Orlando, Fla., 1990, pp. 73–88.
91. L.J. Harris, S.B. Larson, K.W. Hassel, J. Day, A. Greenwood, and A. McPherson, *Nature* **360**, 369–372 (1992).
92. A. McPherson, S. Koszelak, H. Axelrod, J. Day, R. Williams, M. McGrath, L. Robinson, and D. Cascio, *J. Biol. Chem.* **261**, 1969–1975 (1986).
93. R.M. Garavito, in H. Michel ed., *Crystallization of Membrane Proteins*, CRC Press, Orlando, Fla., 1990, pp. 89–106.
94. V. Mikol, E. Hirsh, and R. Giege, *J. Mol. Biol.* **213**, 187–195 (1990).
95. S. Veesler, S. Marcq, S. Lafont, J.P. Astier, and R. Boistelle, *Acta Crystallogr.* **D50**, 355–360 (1994).
96. W.J. Ray, Jr. C.E. Bracker, *J. Cryst. Growth* **76**, 562–576 (1986).
97. A. McPherson and P. Shlichta, *Science* **239**, 385–387 (1988).
98. C.J. Leung, B.T. Nail, and G.D. Brayer, *J. Mol. Biol.* **206**, 783 (1989).
99. S.P. Wood, R.W. Janes, E. Sweeney, and R.A. Palmer, *J. Cryst. Growth* **122**, 204–207 (1992).
100. W. Littke and C. John, *Science* **225**, 203–204 (1984).
101. L.J. DeLucas, C.D. Smith, H.W. Smith, V.K. Senagdi, S.E. Senadhi, S.E. Ealick, C.E. Bugg, D.C. Carter, R.S. Snyder, P.C. Weber, F.R. Salemme, D.H. Ohlendorf, H.M. Einspahr, L. Clancy, M.A. Navia, B.M. Mckeever, T.L. Nagabhushan, G. Nelson, Y.S. Babu, A. McPherson, S. Koszelak, D. Stammers, K. Powell, and G. Darby, *Science* **246**, 651–654 (1989).
102. J. Day and A. McPherson, *Protein Sci.* **1**, 254–1268 (1992).
103. S. Koszelak, J. Day, C. Leja, R. Cudney, and A. McPherson, *Biophysical J.* **69**, 13–19 (1995).
104. S. Patel, R. Cudney, and A. McPherson, *Acta. Crystallogr.* **D50**, 479–483 (1994).
105. M.C. Robert, Y. Bernard, and F. Lefaucheux, *Acta Crystallogr.* **D50**, 496–503 (1994).
106. K.J. Thiessen, *Acta Crystallogr.* **D50**, 491–495 (1994).
107. J.M. Garcia-Ruiz and A. Moreno, *Acta Crystallogr.* **D50**, 484–490 (1994).
108. A.M. Brzozowski and S.P. Tolley, *Acta Crystallogr.* **D50**, 466–468 (1994).
109. S. Patel, B. Cudney,and A. McPherson, *Biochem. Biophys. Res. Comm.* **207**, 819–828 (1995).
110. A. McPherson, *J. Appl. Crystallogr.* **28**, 362–365 (1995).

CRYSTALLIZATION, PROTEINS, KINETICS

MICHAEL R. JOHNS
University of Queensland
Brisbane, Australia

KEY WORDS

Crystal growth kinetics
Nucleation
Osmotic second virial coefficient
Precipitation
Protein crystallization
Protein solubility
Purification

OUTLINE

INTRODUCTION

Crystallization is a mass transfer separation process commonly used in many chemical industries. It permits highly selective purification of a material from an impure mixture, and the crystalline state of the final product is frequently advantageous. Whereas bulk crystallization has been applied for the purification of many small organic biological products, including antibiotics and amino and or-

ganic acids, and has played an important role in chiral resolutions of organic materials, until recently it has been applied only rarely to the bulk purification of proteins. It is well known that proteins, despite their large size, can be crystallized. Nevertheless, a survey of the vast protein crystallization literature, almost entirely devoted to producing crystals for X-ray studies, reveals that protein crystallization has been perceived generally to be tedious, difficult, and disrupted by the presence of impurities. These perceptions have perhaps arisen because of the unique requirements of protein X-ray crystallography, which remains the primary method for the determination of protein structure. The requirements for X-ray crystallography and bulk protein crystallization are contrasted in Table 1. The significant differences highlight the fact that many of the well-known issues that bedevil the production of protein crystals for X-ray crystallography are not germane to the bulk crystallization of proteins for manufacturing uses. In fact, the bulk crystallization of many proteins is found to be reasonably straightforward, similar in nature to the crystallization of inorganic molecules (1), and capable of achieving high degrees of purification at good yield and large throughput. This article will focus on existing knowledge relevant to the bulk crystallization of proteins from solution and largely bypass the more specialized field of protein crystallization for X-ray crystallography, for which excellent reviews already exist (2,3). Useful general reference books on bulk crystallization of materials include those by Randolph and Larson (4) and Mullin (5).

THE PROTEIN CRYSTAL

In addition to the beauty of the crystalline form (Fig. 1), protein crystals share many features in common with smaller inorganic and organic molecule crystals. They consist of highly ordered molecular arrays with distinct crystal faces. Proteins usually exhibit polymorphism, leading to several crystallographic forms, depending on various factors including the nature of the precipitant or the temperature (Table 2). Obtaining the most appropriate crystal form is often important for process manufacture, because "chunkier" crystal forms are easier to wash, store, and transport without excessive crystal damage and fragmentation. Needle forms, which are difficult to process, occur commonly in industrial practice. Transition boundaries between different polymorphic forms must be established experimentally and may be affected by many factors, including temperature, pH, and precipitant type and concentration. A common example is the tetragonal-orthorhombic transition for lysozyme, which occurs at about 25 °C (6). Further, the crystals may undergo significant habit change in the presence of impurities or under varying conditions (7) and exhibit dendritic morphologies during very high crystal growth rates.

A significant feature of all protein crystals is their large degree of incorporated solvent, which is typically 50% v/v (8), but varies generally between 30 and 80%. This has the practical result of making protein crystals unstable when exposed to air, resulting in their rapid collapse to an amorphous state. Consequently, protein crystals are usually maintained as a liquid crystal slurry.

SOLUBILITY

Crystallization processes depend on the production of a supersaturated solution that leads to the formation of the crystalline solid phase, consisting primarily of solute. Consequently, an accurate protein solubility data set determined for the conditions chosen for the crystallization process is essential. Two processes are fundamental to crystallization: nucleation and subsequent crystal growth by addition of additional solute to the nucleated crystal. The extent of supersaturation is the primary determinant of both the yield of crystalline solute from the solution and the rate of crystal growth.

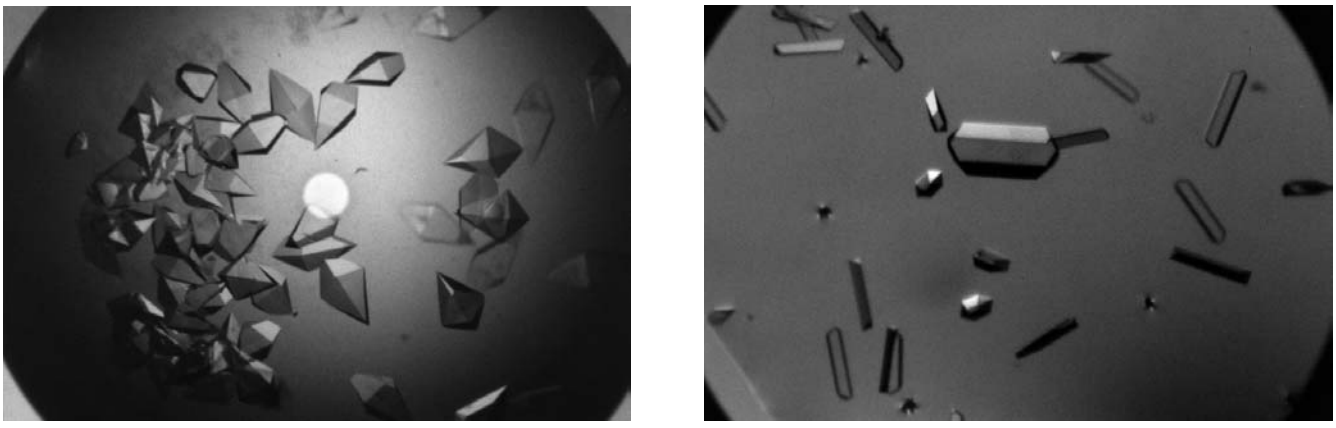
Nucleation

The influence of solubility on the crystallization process is illustrated by Figure 2. Under any set of conditions, there is a unique protein solubility given by the solubility curve. This curve represents the thermodynamic boundary at which there is stable coexistence of solid and dissolved protein forms. Below the solubility line, solution conditions thermodynamically favor the maintenance of a single, homogeneous liquid phase, and protein crystals added to this solution will dissolve. Solution conditions that lie above the solubility curve represent supersaturated conditions under which the formation of a protein-rich solid phase and a saturated liquid phase are favored. The solubility data represent the lower limit of the crystallization process in that once the protein concentration in solution declines to this concentration, the system is at equilibrium and no further crystallization is possible.

Metastable Region. The supersaturation region can be further divided into a metastable region and a nucleation

Table 1. The Conflicting Crystallization Needs of X-ray Crystallographers and Bulk Protein Manufacturers

	X-ray crystallography	Bulk manufacture
Aim of crystallization	Protein structure determination	Protein purification
Protein used	Usually novel	Usually well characterized
Protein quantity available	Usually limited (mg)	Large (grams or more)
Number of crystals desired	One	Unlimited
Crystal size	Generally $>300\ \mu$m	$25\ \mu$m or more
Lattice order	Very highly ordered	Ordered
Solution purity	Preferably free of impurities and protein heterogeneity	Usually contains many impurities



(a) **(b)**

Figure 1. (**a**) Crystals of the commercial sweetener protein, thaumatin, produced from 0.5 M sodium/potassium tartrate. *Source*: Courtesy of W.W. Wilson and A. Arabshahi, Mississippi State Univ. (**b**) Crystals of subtilisin. *Source*: Courtesy of Richard Bott, Genencor International, Inc.

Table 2. Polymorphic Forms of Proteins

Protein	Crystal form	Condition	Reference
Lysozyme	Tetragonal, orthorhombic	Temperature	6
Lysozyme	Monoclinic; triclinic; tetragonal	Nature of salt	9
Cytochrome c_2	Hexagonal; triclinic; tetragonal	Ionic strength, pH	10
Ovalbumin	Monoclinic; triclinic	pH	11

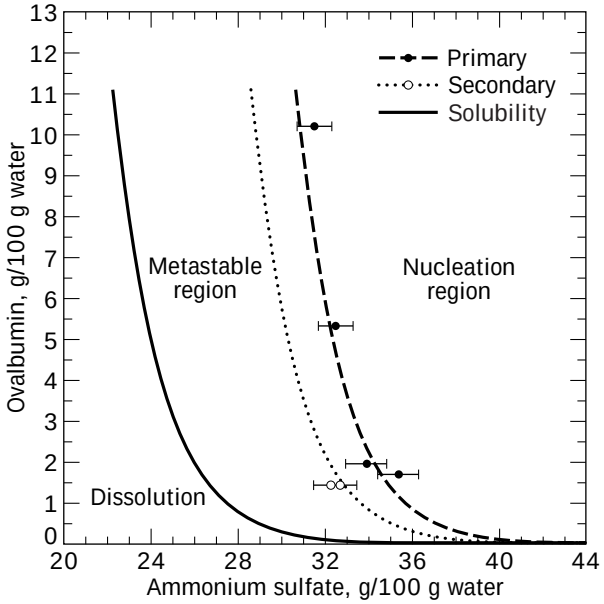


Figure 2. Solubility and nucleation data for ovalbumin at pH 5.2 and 30 °C (11).

region. The metastable region represents a region of low supersaturation in which nucleation does not occur spontaneously within times useful for crystallization processes, despite the existence of supersaturation. However, if crystals are already present, crystal growth will occur. This is the preferred region for the operation of industrial crystallization processes, since it provides the greatest degree of control over the crystal number in the crystallizer and, hence, crystal size distribution and quality. The width of the metastable region is of great significance for crystallization processes and appears to be reasonable generous for proteins (Table 3), although the available data are extremely limited.

When crystallization processes are operated in the metastable region, the crystallization process must be initiated by providing crystal nuclei to the crystallizer (seeding), so that crystallization can begin. The number, mean size, and size distribution of these seeds are variables that permit the quality of the final crystal product and the process crystallization time to be manipulated (Fig. 3).

Nucleation Region. In some crystallization processes, seeding is not preferred, usually because of requirements

Table 3. Typical Size and Metastable Regions for Various Molecular Species

Molecular category	Size of metastable region (c/c*)
Ionic, inorganic salts	≅1.1 or less
Small organics (e.g., sugars)	≅5–30
Proteins	≅1–20
Large proteins (MW > 200 kDa)	≅1.1 or less

Note: c, concentration; c*, concentration at solubility limit under same conditions.

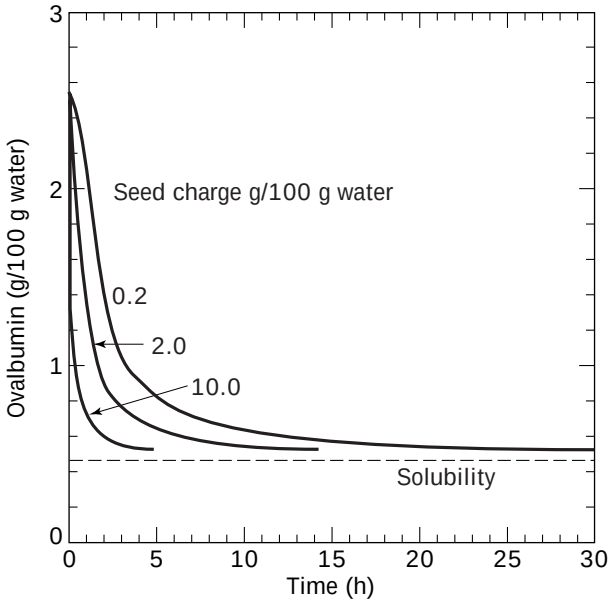


Figure 3. The effect of initial crystal seed charge on the rate of crystallization (given as decrease in ovalbumin concentration in solution) during ovalbumin crystallization from aqueous ammonium sulfate (27% w/v) at pH 4.9 and 30 °C (11).

for aseptic operation (e.g., pharmaceutical preparations) or when a very small (<30 μm diameter) final mean crystal size is desired. In these cases, batch crystallization processes must be initiated in the nucleation region. In this region of higher supersaturation, spontaneous production of large numbers of tiny crystal nuclei will occur after an initial induction time. The control of such processes can be difficult, because rates of crystal nucleation and growth are usually very high at the large supersaturation typical of this region. The nucleation region can be further subdivided into a primary nucleation region, in which nucleation will occur spontaneously from the supersaturated solution (heterogeneous nucleation if active foreign nuclei, such as dust, are present, or homogeneous nucleation if they are not) or secondary region (at lower supersaturation) if crystals are already present. In industrial crystallization, secondary nucleation predominates. However, the threshold between these regions is fuzzy.

Precipitation. At very high supersaturation, conditions will favor the production of a disordered solid phase (precipitation), which usually forms rapidly on nucleation. Such operation generates amorphous protein precipitates, which lack the purity and structural properties of protein crystals. Precipitation processes have been widely used for protein recovery and are covered elsewhere in this encyclopedia. Recent dynamic light scattering (DLS) studies of supersaturated protein solutions have suggested that there is a significant difference in the manner in which homogeneous nucleation occurs for crystallization and precipitation processes (12).

Factors Affecting Protein Solubility

Proteins vary widely in their solubility in aqueous solutions. Most commercially manufactured commodity proteins exhibit high solubilities (40 to 100 mg/mL) in aqueous buffers, but more esoteric proteins can have unusual behavior. For example, some proteins (prolamins) are more soluble in aqueous ethanol than in water (13). Accurate measurement of protein solubility is required for proper design of crystallization processes. This is frequently a tedious process, made difficult by the challenge of obtaining sufficient quantities of pure protein material and the common contamination of the material with heterogeneous forms of the protein, whose effect on the solubility is unclear. For real crystallization processes, the effect of impurities present, often in significant concentration, on protein solubility must not be overlooked. Issues and protocols for protein solubility determination have been discussed by Ries-Kautt and Ducruix (9). Recently, a useful apparatus for the rapid measurement of protein solubility (Fig. 4) using only small quantities of protein has been developed in which solutions near the solubility limit are equilibrated with small volume crystal beds of the protein (14). Protein solubility is affected by many factors, and the major factors are discussed below.

Solution pH. The pH of solution usually has a marked influence on protein solubility. The solubility often exhibits a minimum at the isoelectric point (IEP), with a sharp increase in solubility within small pH changes either side of the IEP (Fig. 5). In concentrated protein solutions, typical of those necessary for protein crystallization, this generalization is fallible. For example, glucose isomerase has an IEP of 4.0, but has minimum solubility in ammonium sulfate at about pH 7 (15).

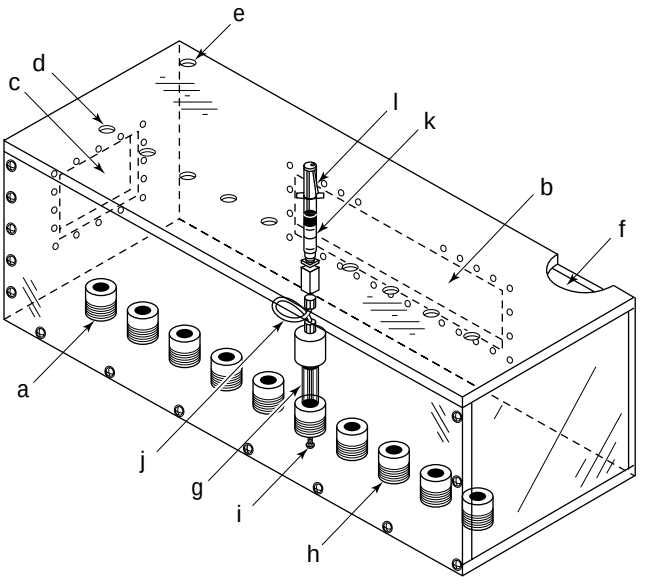


Figure 4. A temperature-controlled apparatus for the rapid determination of protein solubility using small amounts of material. a, column mounts; b, cooling slot; c, circulation pump slot; d, holes for reservoir access; e, temperature probe access; f, heater port; g, crystal column holder; h, base column mount; i, eluent tube; j, tube from column to reservoir; k, reservoir syringe; l, syringe pump. *Source*: Reprinted from Ref. 14 with kind permission of Elsevier Science-NL, Amsterdam.

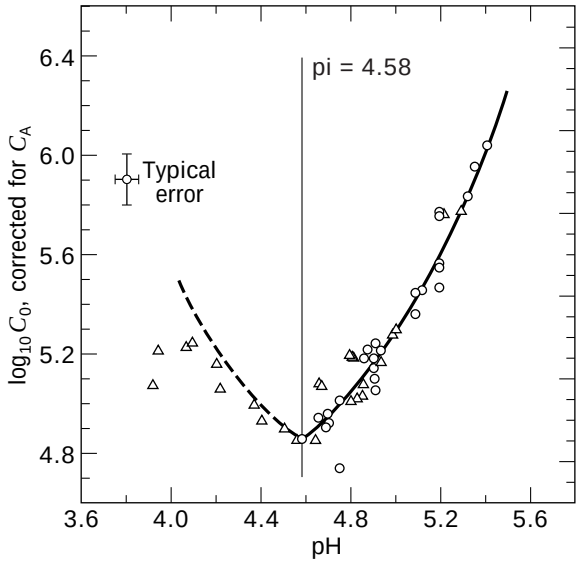


Figure 5. The effect of pH on ovalbumin solubility (C_o) near the isoelectric (pI) point at 18 and 30 °C. *Source*: Reprinted with permission from Ref. 51. Copyright © 1996 American Chemical Society.

Temperature. Proteins exhibit a wide range of solubility responses to temperature. The solubility of some proteins is largely unaffected by temperature; others, such as glucose isomerase (15) and lysozyme (16), show marked increase with temperature; and others exhibit a retrograde response. The effect of temperature and salt concentration on lysozyme solubility is presented in Figure 6. The two common crystal forms of lysozyme have been noted to have a markedly different solubility response to temperature (6).

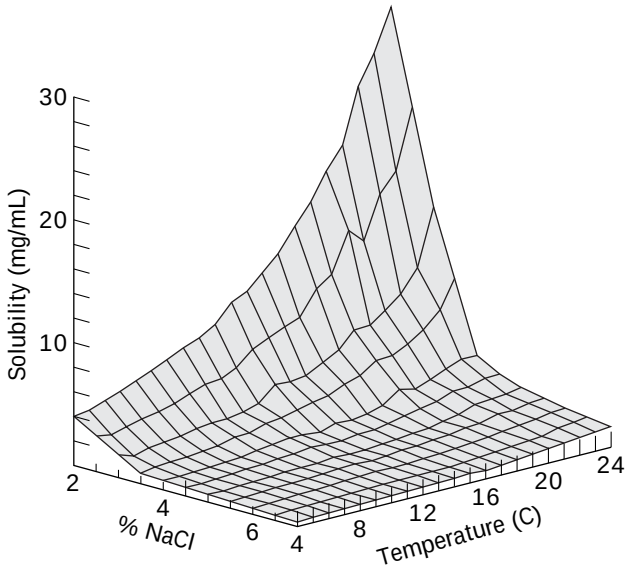


Figure 6. The solubility of tetragonal lysozyme as a function of salt concentration and temperature at pH 4.6 in 0.1 M sodium acetate. *Source*: Reprinted from Ref. 16 with kind permission of Elsevier Science-NL, Amsterdam.

Ionic Strength. Increasing ionic strength decreases globular protein solubility, although diverse effects occur, including regions of low ionic strength ($I < 0.15$ M) at which salting-in phenomena are commonly observed. In some cases, a solubility minimum with concentration has been observed for some protein–salt combinations, particularly for divalent salts (15,17,18). Protein solubility during precipitation processes has often been analyzed by the Cohn plot in which protein solubility is graphed as a logarithmic function of ionic strength. It has been frequently noted, however, that protein solubility in crystal suspensions is lower than for amorphous precipitate suspensions and that the Cohn relationship exhibits poor fit for the former (6,19).

Salt Type. The varying ability of different salts to alter protein solubility is well known and is described by the Hofmeister series (lyotropic series), which ranks ions according to their salting-in or salting-out effect. However, care is needed in applying the series, as shown for lysozyme (20).

Nonelectrolytes. A wide variety of nonelectrolytes can influence protein solubility, usually in a complex manner. Organic solvents (ethanol, acetone), polyethylene glycol (PEG), sugars, and even amino acids have been used to salt-out proteins. Where waste disposal of inorganic salt-rich mother liquors is difficult or expensive, nonelectrolyte nonsolvents may be a good alternative. The complex behavior of these systems is illustrated well by ethanol, which salts-out at low temperatures and low concentrations, but salts-in at all concentrations as temperature rises. These effects do not depend on the dielectric constant (polarity) of the solvent, but on the unusual properties of ethanol and water mixtures (13).

PREDICTING CONDITIONS FOR CRYSTALLIZATION

Obtaining the appropriate conditions for protein crystallization is more challenging than for many small organic and inorganic molecules, because relatively little is known about protein solubilities, often only small quantities of the protein are available early in process development, protein crystallization even under optimal conditions can be slow (days, months), and the use of a precipitant, rather than temperature, to create supersaturation adds an extra degree of freedom in locating conditions for good crystallization. Because X-ray crystallography has until recently been the sole method for protein structure determination, immense effort has been put into the development of protocols for quickly finding suitable conditions for protein crystallization at the microscale. Typical techniques used are well known and include dialysis, vapor diffusion, and batch crystallization methods. Various statistical search strategies, such as sparse matrix screens using commercially available solution kits (21), and robotic technology for repetitive microscale tests (22) are used to reduce the

amount of material and time required and optimize the likelihood of finding a suitable set of conditions.

This body of work exists as an accessible database (23) and serves as an excellent resource for quickly narrowing down suitable systems for the bulk crystallization of proteins. Unfortunately, the transient solution environment typical of microscale systems means that additional tests are required to define crystallization conditions exactly. Process engineering issues can then be applied as an additional set of criteria. Precipitants used for the crystallization will need to be available in bulk at low cost and, for some sites, waste disposal concerns may impose constraints on the use of concentrated salt solutions. In this vein, the use of biodegradable alternative precipitants has been investigated.

Both DLS (12,17) and fluorescence spectroscopy techniques (24,25) have been applied recently to attempts to differentiate conditions leading to precipitation, or crystallization. Although changes in solute structure are seen, the predictive ability of these techniques remains elusive. An alternative approach for predicting optimal solution conditions for protein crystallization is to measure the osmotic second virial coefficient (B_{22}) of the protein in the crystallization mixture, but at dilute (<8 g/L) protein concentrations (26). It was discovered that for the proteins tested (Table 4) there exists a crystallization slot characterized by B_{22} values in the range of -1 to -8×10^{-4} mol/mL/g^2. At more negative values, the solution was found to promote the formation of amorphous protein aggregates, whereas at positive values, the protein preferred to remain in solution. Although independent experimental validation of this approach has not yet been published, subsequent work by Wilson to extend the method to more proteins and by others investigating the application of lattice model (8) or colloidal theory (27) to describe protein interaction in concentrated solutions has supported the validity of this approach. Measurements of B_{22} may provide a faster approach to optimal conditions than do microscale tests, which require several days for the result to be known.

Table 4. Values of B_{22} Corresponding to Crystallization Conditions

Protein	Precipitant and pH conditions[a]	$B_{22} \times 10^{-4}$ mol/mL/g^2
Lysozyme	2% NaCl, pH 4.6	−3.0
Canavalin	0.7% NaCl, pH 7	−0.8
Concanavalin A	1 M $(NH_4)_2SO_4$, pH 7	−2.5
Concanavalin A	0.1 M NaCl, pH 6	−1.9
BSA	52% satd $(NH_4)_2SO_4$, pH 6.2	−2.0
Ovostatin	7.5% PEG 8000, pH 7.5, 20°C	−7.1
Ribonuclease A	50% n-propanol, pH 5, 24°C	−4.1
α-Chymotrypsin	10% PEG 3350, pH 4.6	−8.4
STMV	12.5% SAS, pH 6.5	−1.8
Ovalbumin	43% SAS, pH 5.4	−6.1

Note: BSA = bovine serum albumin; SAS = saturated ammonium sulfate.
[a] −298 K unless stated otherwise; original buffer conditions given in reference. *Source*: Ref. 26.

KINETIC STUDIES

The knowledge of rates of protein crystal growth is of primary importance in industrial crystallization processes, but except for lysozyme, few systematic studies have been published. Most kinetic studies have been performed using static crystal measurements, the most common method of which involves nucleating several crystals in a growth cell located on a microscope and following the increase in crystal dimension with time as supersaturated solution is pumped gently around the crystals (28,29). These are useful for fundamental studies of growth mechanisms, but have a number of disadvantages for process design:

- The growth rate is typically expressed as a face (e.g., two-dimensional increase) growth rate whose applicability to three-dimensional crystal growth in suspended crystal slurries in industrial crystallizers, in which the hydrodynamic conditions may be vastly different, has not been validated.
- Because crystals grow at different rates, a large number of crystals must be followed to obtain a reasonable approximation of mean population growth rates.

Nevertheless, the method has benefits in that data can be obtained on individual crystals. Few kinetic studies of protein crystallization have been performed in bulk crystallizers in which the growth of a large crystal population can be measured under conditions more typical of industrial conditions (11,30).

Factors Affecting Rates of Protein Crystal Growth

Supersaturation. Chief among the factors determining the rate of crystal growth is the degree of supersaturation, which defines the driving force for crystallization. Supersaturation (s) can be defined a number of ways and is typically defined in the manner most convenient for the user. The relative supersation (σ) = s/c^* is commonly used. In industrial practice, it is convenient to express the relationship between the growth rate (G, μm/h) and supersaturation (s) as a power law:

$$G = ks^n$$

where k is the nth-order growth rate constant and n is the kinetics order. From the studies performed on protein growth kinetics, a range of dependencies of growth rate on supersaturation have been reported, but most have revealed a second-order dependence (Table 5). This may be interpreted as surface integration kinetics control of the rate of crystal growth. In isolated cases, a first-order dependence was measured, for example, for concanavalin A (Table 5), which is more indicative of a diffusion-controlled growth mechanism. The growth rates of protein crystals and their dependence on supersaturation are contrasted in Figure 7. Protein crystal growth rate appears to be independent of crystal size (31,32).

In the case of lysozyme, which has been extensively studied, the usual growth rates experienced in static cell studies suggest that about two to three layers of molecule are added to the crystal face every second, amounting to a

Table 5. The Relationship between Protein Growth Rate and Supersaturation

Protein	Kinetics order, n
Concanavalin A[a], porcine pancreatic α-amylase[b]	1
Insulin[c], ovalbumin[d], canavalin[e], lysozyme[f]	2

[a]Ref. 7.
[b]Ref. 33.
[c]Ref. 30.
[d]Ref. 31.
[e]Recalculated from Ref. 28.
[f]Ref. 29.
Source: Ref. 31.

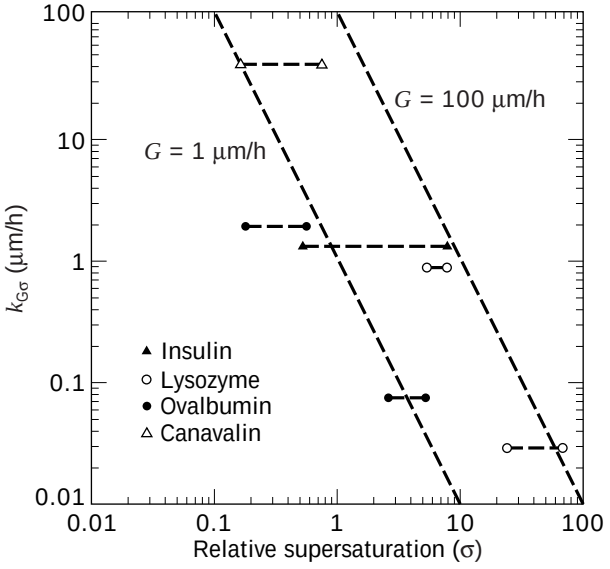


Figure 7. Values of the second-order growth rate constant, $k_{G\sigma}$, against relative supersaturation for various proteins. Most average crystal growth rates, G (expressed on a volume equivalent diameter basis), fall between 1 and 100 μm/h over the range of supersaturation commonly used.

face growth rate of 0.1 to 1 nm/s (0.4–4 μm/h) (34). Recent work has proposed that protein crystal growth occurs by the addition of oligomers of lysozyme into the crystal, rather than by monomer incorporation (35). There is a large body of supporting evidence for this view, both with lysozyme and for other proteins (1,7), although this finding remains controversial (36).

A comparison of protein growth rates with those of other commonly crystallized materials reveals that proteins are slower to crystallize relative to most smaller molecules for a given relative supersaturation (Fig. 8). Nevertheless, known protein crystal growth kinetics are comparable to those of aluminum trihydroxide, which is crystallized industrially on a very large scale. Furthermore, the metastable zone available for protein crystallization (e.g., the values of supersaturation in which it is convenient to operate a crystallizer) is many fold more than that possible in many small molecule crystallizations (Table 3).

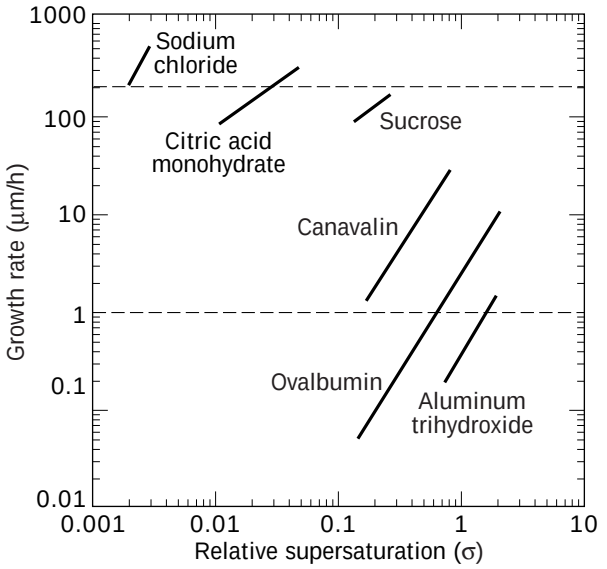


Figure 8. Comparison of average crystal growth rates for various values of supersaturation for industrially crystallized small molecules and the proteins ovalbumin and canavalin.

Most commercial protein crystallization processes are batch operations. During isothermal batch crystallizations, the supersaturation will fall rapidly from an initially high value to much lower values, especially because supersaturation is typically achieved by the addition of precipitant. In these cases, it may be uneconomical to wait for the system to reach equilibrium before terminating the crystallization. Although low supersaturation is undesirable because it will result in long crystallization times, excessive supersaturation is also undesirable and generally results in any of a number of inconvenient features, including greater inclusion of impurities in the crystal lattice, unusual and dendritic crystal forms that are hard to process (Fig. 9), and amorphous precipitation. Where the solubility of the protein being crystallized exhibits sensitivity to pH, temperature, or other factors, the supersatu-

Figure 9. Rosette of α-amylase crystals. *Source*: Courtesy of Richard Bott, Genencor International, Inc.

ration may be maintained at a moderate value by changing these factors as crystallization occurs, provided that there are no deleterious results such as decreased rate, change in crystal form, denaturation, etc. This can only be determined by thorough solubility studies.

Temperature and pH. The effect of temperature and pH on the kinetics of protein crystallization have been little studied. There are frequent comments in the literature that different pH and temperature conditions, or even changes in the primary structure of the protein, altered the rate of crystallization or made crystallization easier. In most cases, there are insufficient data to confirm that the experimental variable (pH, temp, etc.) had a direct influence on the rate of crystallization, rather than an indirect influence through some other effect, such as an altered protein solubility, which in turn would result in increased supersaturation for the same initial protein concentration and hence a faster crystallization rate.

The pH has been observed to have a marked effect on the crystallization rate, independent of its effect on solubility, in the case of ovalbumin (11). An increase in pH at values up to 1 pH unit higher than the isoelectric point (pH 4.6) resulted in an increase in the rate of crystallization of up to 10-fold. Whether this trend continued at larger deviations from the IEP was not studied. Reports on the effect of temperature are rare, but the probable dependence of the rate of protein crystallization on kinetic mechanisms implies that the rate should increase with temperature at a given supersaturation. Although the results above are too limited to generalize to all commercially produced proteins, it is interesting to note that conditions that have been observed to enhance crystallization rate (increased temperature, pH values away from the IEP) also correspond to conditions of greater protein solubility. Consequently, a compromise is necessary between high crystallization rates and the need for high recovery of protein as crystal.

Impurities. The presence of some impurities in a crystallizing solution is known to result in depressed rates of crystal growth for many molecules. In most cases, such poisoning is difficult to prove unequivocally. The impurities likely to be present in protein solutions undergoing crystallization can be arbitrarily divided into three categories:

- Nonprotein impurities
- Protein contaminants not derived from the product protein
- Heterogeneous forms of the product protein

The term *pure* typically means crystals devoid of the latter two categories, because large quantities of the precipitant inevitably will be present in the solvent channels within the protein crystal. In fact, it is exactly this property that makes protein crystals amenable to X-ray crystallography.

Nonprotein Impurities. The rate of protein crystallization appears to be remarkably unaffected by the presence of small compounds, such as salts, precipitants, sugars, and pigments, compared to small molecule crystallization processes (1), although they may have a pronounced effect on solubility of the protein (15). This resilience may be because of the comparatively small number of contact points between adjacent protein molecules in the crystal, which makes disruption of the ordered lattice by small molecules more unlikely compared to the tightly packed lattices of salts and other small molecules. This is convenient for protein bioprocessing. However, polysaccharides derived from microbial cell wall or components of the media are commonly present in fermentation broths and have been implicated as causing difficulties in bulk crystallization processes, particularly affecting the crystal morphology and properties of the final product (37). This problem was resolved by the use of carbohydrate-hydrolyzing enzymes.

Contaminating Proteins. The presence of contaminating proteins has often been considered a primary reason for not adopting crystallization as a protein purification process. In fact, in the few cases studied to date, the presence of contaminating proteins had little or no deleterious effect on the rate of crystallization. The rate of crystallization of ovalbumin was unaffected by the presence of up to 0.22-mol fraction of total protein as lysozyme and conalbumin (11); similarly, the rate of lysozyme crystallization was unaffected by the presence of up to 5% w/w ovalbumin (36). However, the literature on this issue is scarce.

Heterogeneous Forms of the Product Protein. An important issue is the effect of microheterogeneous forms of the product protein on rates of crystallization. There is mounting evidence that even small changes in the primary sequence of a polypeptide may lead to large alterations in its crystallizability. Unfortunately, the term *crystallizability* is vague and may refer to an improved propensity for nucleation, altered solubility under the conditions used for crystallization, or an actual enhancement in the rate of growth of the crystal face for a given supersaturation. Discerning between these effects requires extensive characterization of protein behavior, which typically is not performed. Nevertheless, numerous reports of protein variants that have either improved crystallizability or that inhibit crystallization of the preferred protein exist. Thus, variants of recombinant human insulin were generated by site-directed mutagenesis that exhibited improved crystallizability under physiological conditions (38), whereas for ovalbumin, the subtilisin-treated variant, plakalbumin (which possesses 6 fewer amino acid residues out of the original 385), is reported to exhibit greatly improved crystallizability (39). This opens the possibility of engineering protein variants with improved crystallization characteristics to improve properties germane to their final use or to reduce processing cost. Some forms of heterogeneity, however, may also be responsible for poisoning rates of protein crystallization. This has been reported for lysozyme in mixtures of that protein from different sources (40). This adverse effect of variant forms is more likely, because they would share a high degree of homology with the native protein and might be expected to integrate into the lattice in such a way as to poison or disrupt further growth. The

ability of crystallization process to select for crystals of only one microheterogeneous form of a protein is unclear.

Other Factors. The effect of a number of other operating variables on the rate of protein crystallization have not been investigated. It would be expected that the precipitant used, or its concentration, would alter the rate of crystallization. However, data are scarce. Certainly, the effect of salt concentration (ammonium sulfate) on the rate of ovalbumin crystallization was negligible, independent of its effect on protein solubility, which is large (11). Agitation, or a convective environment, has been argued to have a deleterious effect on crystal growth in small-scale studies performed on microscope cells (29), although other studies have contradicted these findings (41,42). Nevertheless, some agitation is essential in industrial crystallizers to prevent the crystals settling in the vessel. A more serious effect of agitation is crystal breakage. Protein crystals appear to be relatively fragile, probably because of the small density of intermolecular contact points between neighboring molecules in the crystal lattice. Consequently, rough handling of the crystals should be minimized. High pressure (up to 3,000 bar) was reported to increase the crystallization rate of glucose isomerase relative to that of solutions at atmospheric pressure (43); however, it is more likely that the enhanced rates observed were the result of higher rates of nucleation at high pressure, which would generate larger crystal surface areas for growth (44). This latter effect would be advantageous for the production of large numbers of small (<10 μm) protein crystals in the absence of seeding, because induction times for nucleation were also reduced by high pressure.

Growth Dispersion

Growth rate dispersion (GRD) refers to the existence of a population of crystal growth rates under identical crystallization conditions. The primary cause of GRD remains unclear, but it is now accepted that crystals, under conditions of constant supersaturation, will grow at different, but constant, growth rates. The most significant effect of GRD is to broaden the crystal size distribution with time, which is generally undesirable for most industrial applications. The existence of GRD in protein crystallization is largely unstudied, but appears to be small relative to crystallizations that exhibit the effect strongly, such as sugars.

PURIFICATION BY CRYSTALLIZATION

Crystallization is capable of highly selective purification of most compounds from impure mixtures. Protein crystals sought for X-ray crystallography need to be highly pure, because the crystal is required for structure determination. This requirement has commonly led to the comment that protein solutions must be highly pure to attain crystallization. However, for bulk crystallization purposes, this requirement is not relevant. Because of the highly open structure of protein crystals, small impurities, such as salts, sugars, etc., are likely to be present in the final crystals to a similar degree as in solution, although protein crystals appear to exclude the pigments and colorants found routinely in fermentation broths (37). Successful protein crystallization has been routinely obtained in the presence of other contaminating proteins with degrees of purification similar to those achieved by ion exchange. Most of the carryover of contaminating protein is in the liquid adhering to the exterior surface of the crystals. Studies showed that ovalbumin crystals (MW 45,000 Da) produced in bulk crystallization were substantially free of both smaller and larger protein contaminants, which comprised 22 mol % of the total protein initially present in solution (11), after a single wash of the crystals (Fig. 10). Similar results have been observed for lysozyme (36).

PROCESS APPLICATIONS OF PROTEIN CRYSTALLIZATION

Despite the ubiquity of crystallization for the structural determination of proteins, actual process applications of crystallization as a technology for the purification and recovery of proteins remained relatively rare until the last few years. An increasing number of proteins are now produced in multikilogram or multiton quantities by crystallization from clarified fermentation broth. These include insulin, glucose isomerase, asparaginase, subtilisin, lipases, thermolysin, and penicillin acylase (45–47). The first major use of crystallization was in the purification of insulin, which was crystallized as a zinc complex from salt solution. This technology was developed in the 1950s and simulated the natural process found for storage of insulin as zinc hexameric crystalline arrays in the pancreas (30). More recently, the ability of insulin to be crystallized has been used in slow release forms to achieve a more sustained titer of the hormone in the blood (38), due presumably to the regulation of insulin adsorption by slow dissolution kinetics at physiological pH.

The salient features of protein crystallization processes reported in the patent literature are listed in Table 6, al-

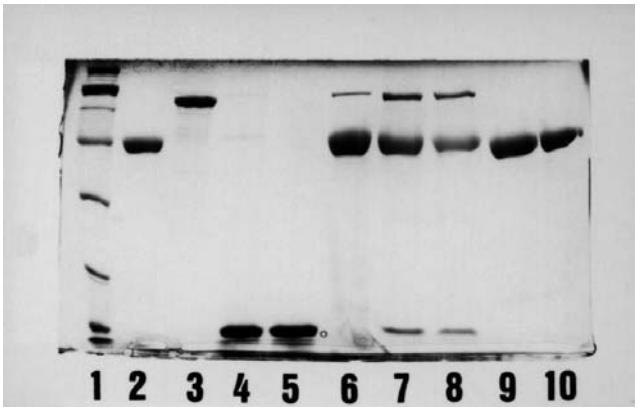


Figure 10. Crystallization of ovalbumin from a solution mixture of ovalbumin (0.784 mol fraction of protein), lysozyme (0.182), and conalbumin (0.034) results in high purity crystals by overloaded SDS-PAGE analysis with Coomassie blue R-250 staining. Lanes: 1, molecular weight; 2, ovalbumin; 3, conalbumin; 4 and 5, lysozyme; 6, initial solution; 7 and 8, final mother liquor; 9 and 10, washed ovalbumin crystals separated from the mother liquor (11).

Table 6. Selected Protein Crystallization Processes from Patent Literature

Protein	Conditions	Yield (%)	Time (h)	Company	Reference
Subtilisin	pH 5.2–5.8; 0.5M NaCl	84	24–48	Genencor	18
Glucose isomerase	1.5% $MgSO_4$	97	3	Stabra AG	47
Peroxidase	pH 4.0; 1.5% $CaCl_2$; 20% PEG 4,000	90	4	Novo-Nordisk	48
Alkaline protease	pH 7.5; 0.5M NaCl	76	4	Solvay	37
Amylase	pH 7.6; 8% NaCl	91.5	24	Solvay	49

though the process may not be used to produce the protein in commercial practice. In the majority of cases, halide salts are preferred for producing the supersaturation needed to obtain crystallization, although a recent report suggests that the use of PEG may greatly enhance the process at lower salt contents (48).

Crystallization is usually performed using clarified fermentation broth that has been concentrated normally by ultrafiltration so that supersaturation is attained on addition of precipitant. The crystallization pH is generally adjusted to values at or near the isoelectric point of the protein to take advantage of the lower solubility at these pH values. Alkaline proteases, such as those from *Bacillus* species (e.g., subtilisin) and lysozyme, are exceptions. These enzymes are typically crystallized at lower pH values to minimize autolysis during the process (46). Process temperatures vary widely. Temperature change plays a minor role compared to precipitant addition for obtaining supersaturation. This is usually the result of the constraints imposed on process temperature by protein stability requirements or the relative insensitivity of protein solubility to temperature, particularly at high salt concentrations. Typically, the crystals produced in commercial crystallization processes are small (30 to 100 μm in diameter) to aid their more rapid dissolution, in either product applications or subsequent purification steps. By the end of the crystallization process, the crystal content of the liquor can be up to 20 to 30% v/v (18).

The crystalline product can be separated from the mother liquor by settling of the denser crystal phase and decantation of the mother liquor from the crystals, or by usual solid–liquid separation technologies such as microfiltration, pressure filtration, or centrifugation. Washing to remove the mother liquor from the crystals improves crystal purity and the overall purification ability of the process. Generally, the crystals must be maintained as a concentrated slurry in a suitable salt solution to avoid redissolution or drying out and reversion to an amorphous state.

A novel approach to protein crystallization, which overcomes the problem of maintaining the crystal form, is the subsequent cross-linking of the crystals by, for example, glutaraldehyde, which renders the crystals insoluble in aqueous solutions. This process has been used to develop mechanically robust and stable enzyme crystals called CLECs (cross-linked enzyme crystals) suitable for repeated use and easy separation from the reaction liquor in organic syntheses (45,50). A photograph of CLECs is given in Figure 11. A further advantage of CLECs is their enhanced stability relative to the noncrystalline form and even to the un-cross-linked crystal form of the enzyme under adverse conditions, such as low water environments, elevated temperatures or the presence of proteases, which makes them of value in the synthesis of many chirally pure drugs and chemical intermediates. A recent example (50) comprising a lipase revealed the half-life of the soluble enzyme to be 5 h compared to more than 13 days for the cross-linked crystal form.

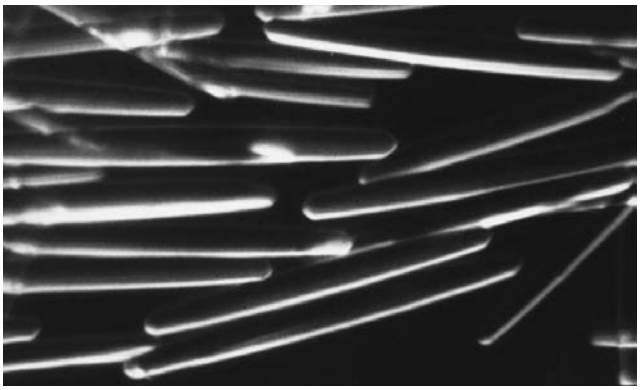


Figure 11. Cross-linked enzyme crystals (CLECs) of thermolysin (100 μm length). *Source*: Courtesy of Alex Margolin, Altus Biologics Inc. Reprinted from Ref. 45 with kind permission of Elsevier Science-NL, Amsterdam.

CONCLUSION

In addition to its critical importance for protein structure determination by X-ray crystallography, protein crystallization is now emerging as a useful bulk technology for protein purification. It is particularly well suited to commodity proteins where bioprocessing costs must be contained in the face of demands for improved product quality. Protein crystallization offers potentially high degrees of purification at any scale in reasonable process times and cost. Nevertheless, the gaps in knowledge concerning the application of protein crystallization at process scale remain large, despite its successful use in the production of commercially significant proteins. Protein crystals also offer exciting prospects as novel forms of catalysts in controlled-release formulations and as biomaterials. The sensitivity of protein crystallizability to even single amino acid changes in the polypeptide chain further suggests that fruitful results can be obtained using tools of modern molecular biology and protein engineering, already well established in the production of commodity proteins. Perhaps what is needed most, however, is the realization that crystallization, long overlooked in production processes for

proteins and far more elegant than its wilder sibling—precipitation—deserves to be considered alongside the many other operations commonly selected for protein purification.

ACKNOWLEDGMENTS

I would like to thank Professor Ted White and Dr. Russell Judge for their participation in much of the work described in this article and for their reading of the drafts. I also give my heartfelt thanks to Todd Becker, John Kan, Alfred Gaertner, Rick Bott, and all the team at the California Technology Center of Genencor International, Inc., Palo Alto, Calif., for their hospitality during my stay with them and for their insight on protein crystallization and proteins. Finally, I acknowledge the many useful discussions with researchers in the field, especially Marc Pusey (NASA, Huntsville, Ala.), Bill Wilson (Mississippi State Univ.), Harvey Blanch (Univ of California, Berkeley), Kris Berglund (Michigan State Univ.), Chuck Glatz (Iowa State Univ.), and many others. Thank you.

BIBLIOGRAPHY

1. F. Rosenberger, *J. Cryst. Growth* **76**, 618–636 (1986).
2. R. Giegé, B. Lorber, and A. Théobald-Dietrich, *Acta Crystallogr.* **D50**, 339–350 (1994).
3. A. McPherson, *Eur. J. Biochem.* **189**, 1–23 (1990).
4. A.D. Randolph and M.A. Larson, *Theory of Particulate Processes. Analysis and Techniques of Continuous Crystallization* 2nd ed, Academic Press, San Diego, 1988.
5. J.W. Mullin, *Crystallization* 3rd ed., Butterworth-Heinemann, Oxford, U.K., 1993.
6. F. Ewing, E. Forsythe, and M.L. Pusey, *Acta Crystallogr.* **D50**, 424–428 (1994).
7. S. Moré and W. Saenger, *J. Cryst. Growth* **153**, 35–41 (1995).
8. C. Haas and J. Drenth, *J. Cryst. Growth*, **154**, 126–135 (1995).
9. M. Ries-Kautt and A. Ducruix, in A. Ducruix and R. Giegé eds., *Crystallization of Nucleic Acids and Proteins*, Oxford Univ. Press, Oxford, U.K., 1992, pp. 195–218.
10. H.L. Axelrod, G. Feher, J.P. Allen, A.J. Chirino, M.W. Day, B.T. Hsu, and D.C. Rees, *Acta Crystallogr.* **D50**, 596–602 (1994).
11. R. Judge, M. Johns, and E. White, *Biotechnol. Bioeng.* **48**, 316–323 (1995).
12. A.J. Malkin and A. McPherson, *Acta Crystallogr.* **D50**, 385–395 (1994).
13. F. Franks, in F. Franks ed., *Protein Biotechnology*, Humana Press, Totowa, N.J. 1993, pp. 133–189.
14. M. Pusey and S. Munson, *J. Cryst. Growth* **113**, 385–389 (1991).
15. K. Visuri, in *10th Int. Conf. of Enzyme Engineering*, Kashikojima, Japan, 1989.
16. E. Cacioppo and M.L. Pusey, *J. Cryst. Growth* **114**, 286–292 (1991).
17. Y. Georgalis and W. Saenger, *Adv. Colloid Interface Sci.* **46**, 165–183 (1993).
18. U.S. Pat. 5,041,377 (August 20, 1991), T. Becker and V.B. Lawlis (to Genencor International Inc.).
19. C.W. Bunn, P.C. Moews, and M.E. Baumber, *Proc. R. Soc. London, Ser. B* **178**, 245–258 (1971).
20. M. Ries-Kautt and A. Ducruix, *J. Biol. Chem.* **264**, 745 (1989).
21. J. Jancarik and S.H. Kim, *J. Appl. Crystallogr.* **24**, 409–411 (1991).
22. H.A. Kelders, K.H. Kalk, P. Gros, and W.G.J. Hol, *Protein Eng.* **1**, 301–303 (1987).
23. G.L. Gilliland, *J. Cryst. Growth* **90**, 51–59 (1988).
24. M. Jullien and M.-P. Crosio, *J. Cryst. Growth* **110**, 182–187 (1991).
25. B. Pan and K. Berglund, *J. Cryst. Growth* **130**, 587 (1993).
26. A. George and W.W. Wilson, *Acta Crystallogr.* **D50**, 361–365 (1994).
27. D. Rosenbaum, P.C. Zamora, and C.F. Zukoski, *Phys. Rev. Lett.* **76**, 150–153 (1996).
28. R. DeMattei and R. Feigelson, *J. Cryst. Growth* **97**, 333–336 (1989).
29. E. Forsythe, F. Ewing, and M. Pusey, *Acta Crystallogr.* **D50**, 614–619 (1994).
30. J. Schlichtkrull, *Acta Chem. Scand.* **11**, 439–460 (1957).
31. R.A. Judge, Ph.D. Thesis, University of Queensland, Brisbane, Australia, 1995.
32. M. Pusey and R. Naumann, *J. Cryst. Growth* **76**, 593–599 (1986).
33. R. Boistelle, J.P. Astier, G. Marchis-Mouren, V. Desseaux, and R. Haser, *J. Cryst. Growth* **123**, 109–120 (1992).
34. N. Niimura, Y. Minezaki, M. Ataka, and T. Katsura, *J. Cryst. Growth* **154**, 136–144 (1995).
35. A. Nadarajah, E. Forsythe, and M. Pusey, *J. Cryst. Growth* **151**, 163–172 (1995).
36. M. Skouri, B. Lorber, R. Giegé, J.-P. Munch, and J. Candau, *J. Cryst. Growth* **152**, 209–220 (1995).
37. U.S. Pat. 5,439,817 (August 8, 1995), J.K. Shetty, C.P. Patel, and M.A. Nicholson (to Solvay Enzymes, Inc.).
38. J. Markussen, P. Hougaard, U. Ribel, A.R. Sorensen, and E. Sorensen, *Protein Eng.* **1**, 205–213 (1987).
39. M. Miller, J. Weinstein, and A. Wlodawer, *J. Biol. Chem.* **258**, 5864–5866 (1983).
40. C. Abergel, M.P. Nesa, and J.C. Fontecilla-Camps, *J. Cryst. Growth* **110**, 11–18 (1991).
41. M.L. Grant and D.A. Saville, *J. Cryst. Growth* **153**, 42–54 (1995).
42. H. Lin, F. Rosenberger, J. Alexander, and A. Nadarajah, *J. Cryst. Growth* **151**, 153–162 (1995).
43. K. Visuri, E. Kaipainen, J. Kivimaki, H. Niemi, M. Leisola, and S. Polasaari, *Biotechnology* **8**, 547–549 (1990).
44. M. Saikumar, C. Glatz, and M. Larson, *J. Cryst. Growth* **151**, 173–179 (1995).
45. A.L. Margolin, *Trends Biotechnol.* **14**, 223–230 (1996).
46. U.S. Pat. 5,256,557 (October 26, 1993), J.K. Shetty, C.P. Patel, and M.A. Nicholson (to Solvay Enzymes, Inc.).
47. U.S. Pat. 5,120,650 (June 9, 1992), K. Visuri (to Stabra AG).
48. World. Pat. WO 95/01989 (January 19, 1995), B.M. Nilsson, M.A. Laustsen, and A. Rancke-Madsen (to Novo-Nordisk A/S).
49. U.S. Pat. 5,281,526 (January 25, 1994), I.C. Good, C.Y. Kim, J.K. Shetty, and K.M. Sproat (to Solvay Enzymes, Inc.).
50. J. Lalonde, C. Govardhan, N. Khalaf, A. Martinez, K. Visuri, and A. Margolin, *J. Am. Chem. Soc.* **117**, 6845–6852 (1995).
51. R. Judge, M. Johns, and E. White, *J. Chem. Eng. Data* **41**, 422–424 (1996).

See also PROTEIN PURIFICATION, AQUEOUS LIQUID EXTRACTION.

CULTURE COLLECTIONS

PHILIP PACKER
ALAN DOYLE
Centre for Applied Microbiology and Research
Salisbury, Wilts
U.K.

KEY WORDS

Culture collection
Culture supply
Gene banks
Identification
Quality Control
Resource centers

OUTLINE

INTRODUCTION

The Increased Importance of Microorganisms and Cell Lines

Microorganisms have a significant effect on human life, whether as a result of the diseases or benefits they cause humans or to their food source (plants and animals) and through their effect on the environment (food spoilage, water pollution). In addition, microorganisms and cell lines are used increasingly within biotechnology and are therefore an integral part of the infrastructure of microbiology and fundamental in underpinning scientific research. They are used in crop improvement, food production, pharmaceutical production, and treatment of waste. Industry has the capacity to screen isolates for particular biochemical properties and metabolite production, to evaluate and develop commercial kits and rapid detection kits for microbially produced compounds, and to evaluate organisms' capacity for use in bioremediation and as biocontrol agents.

With the decline in the use of animals in scientific experimentation over recent years, both for economic and ethical reasons, many modern standards and specifications require the use of microorganisms and cell lines in the testing of products and quality control procedures. Culture collections, therefore, supply microorganisms and cell lines for the British, European, and U.S. testing standards, disinfection testing, the assessment of mutagenicity against a range of microorganisms, toxicity testing, NAMAS accreditation, quality accreditation, water and environmental reference strains, and biological and biomedical research.

The supplier is responsible for adequate quality control procedures to ensure the authenticity of these cultures. Failure to do so may result in potential loss incurred as a result of contamination, work on the wrong organism or biochemical product, or invalid manufacturing quality control procedures.

Preservation of Existing Biodiversity

Over recent years, there has been increased concern regarding the conservation of biodiversity and the wealth of genetic information still to be harnessed. This matter has been found to be so important that it has been taken up governmentally at an international level. The Convention on Biological Diversity was signed by the governments of 155 countries and the European Community at the Earth Summit in Rio de Janeiro in 1992. The objectives of the convention were the conservation of biological diversity, the sustainable use of its components, and the fair and equitable sharing of benefits arising from the use of biological resources. The consequences of the introduction of this complex convention has been to generate interest in the scientific aspects of diversity and to begin to establish a number of major programs and projects aimed at increasing knowledge and improving conservation procedures, particularly in situ methodologies.

At a legal level, the common inheritance principle regarding ownership of biological resources has changed to sovereign rights, and access to genetic resources will be controlled by the contracting parties and subject to national legislation. Scientifically, countries are required to establish research and training programs, develop guidelines and procedures for conservation (ex situ and in situ), introduce procedures for assessing and monitoring environmental impact, establish emergency response procedures, and establish partnerships that support capacity building and technology transfer in developing countries.

These regulations are far from being neatly juxtaposed to the Budapest Treaty legislating the patenting of microorganisms used in the industrial microbiology industry because the patent system limits the access to biological re-

sources. The role of the Budapest Treaty is discussed later.

Conservation of microorganisms is obviously a very important component of the biological diversity initiative because they are the most biochemically adaptable organisms on the planet on which they are ubiquitous, occupying every ecological niche. Thus, culture collections have a primary role in conserving microorganisms and cells.

THE DEVELOPMENT OF CULTURE COLLECTIONS AS A RESOURCE

Since Krael established the first culture collection in Prague in 1890 (1), many others have been set up, not only to conserve living examples of microbes (e.g., bacteria, fungi, protozoa, algae, viruses), but also cell lines (human, animal, and plant cells) and DNA probes. Although the preservation of genetic diversity will fundamentally benefit humankind anyway, it is not known which materials may be useful or, indeed, for what purpose they may be required in the future. It is also impossible to predict changes in special habitats or in the general environment, nor what effect such changes may have on the bacterial genome or chromosomal elements. Therefore, some cultures that are of no importance today may be of significance in the future, and it is therefore of fundamental importance to preserve today's biodiversity.

There are three main types of culture collections, the classification of which largely depends on the size of the collection, its documentation, and the service provided. The first type is one where conservation of a small collection is a secondary activity of the laboratory. As such culture collections evolve and increase in size, a second type that is larger and more specialized will arise. These collections often result from specific research programs that are vulnerable when projects come to an end or individual senior scientists retire. Such collections need correct documentation to be a valuable resource and a greater effort to maintain the resource successfully.

The third type of culture collection is the nationally and internationally recognized service-supply culture collections where conservation is the primary function. In order that such service-culture collections can become a recognized international resource, three principal elements must be satisfied. First, the culture collection has to be preserved in a viable state. Second, the history, original source, and properties of the culture must be documented. Third, the cultures must be available to the wider scientific community.

With the increasing demand on such culture collections for authenticated, reliable biological material, the World Federation for Culture Collections (WFCC) has published recommendations for good practice in culture collections so that new and existing collections have guidance and approved standards of operation (2).

Service-supply culture collections exist primarily to maintain and supply on request cataloged, authenticated cultures and to give expert advice on cultivation and preservation to scientists, technologists, industrialists, and teachers. The primary task of the staff therefore is curatorial, thus distinguishing it from secondary collections. Through continuity of work, these culture collections become expert in documentation and preservation techniques and in the knowledge of the behavioral properties of the holdings.

For a service collection to be a truly valuable resource center, it must keep abreast of the developing science it serves. There must be an active accessioning policy, ongoing reassessment of preservation procedures, awareness of taxonomic changes, identification of new requirements, and an original research program. Recently, as a consequence of this, other services have become equally important, including cryopreservation, bulk supply of cultures for screening for products and quality control, patent and safe deposit facilities, identification of cultures, and the delivery of training courses.

OPERATION OF CULTURE COLLECTIONS

It is important that all culture collections operate good practice in all aspects of their function, from microbiology to dispatch of samples. The various activities of culture collections and their successful management are discussed in this section.

Organization, Funding, and Objectives

It is of fundamental importance that the establishment that houses a culture collection realizes its importance, accepts the responsibility inherent in maintaining a public service to appropriate standards, and is committed to its long-term maintenance. In the case of existing collections, this aspect should be continually clarified with the director of the parent institute, its scientific council, senior university officials, governing board, or other appropriate authorities.

Intrinsic to such a commitment is appropriate funding. The operation of such culture collections is a long-term commitment and requires core-funding from the overseeing parent organization. Support solely in the form of short-term contracts is totally inappropriate, because necessary levels of funding need to be guaranteed for the future activities of the culture collection, including the services being planned and the standards users would expect.

All culture collections should have a mission statement and long- and short-term objectives relating to the scope of the holdings and to the range of services that it will offer. If these statements are not agreed upon by the parent institute and other funding bodies, then the culture collection will not be able to operate effectively.

A considerable degree of income generation is possible from culture collection activities, varying from as little as 25% to as much as 85%. This can depend on a mixture of sources of income, culture sales, and contract services.

Holdings

The diversity of microorganisms and cells and their number should be continually addressed by a culture collection because appropriate funding will be needed to maintain the collection. Microorganisms that require particular con-

tainment facilities are those that are potentially pathogenic to humans, animals, or plants or those that are toxic.

When a new culture collection is being established, it often relates to the interests of the parent organization. However, it is advisable to aim to complement existing collections rather than duplicating existing ones. It is necessary, therefore, to have a clearly defined accession policy on which new strains are to be added to the collection. Lack of such a policy will result in many unsolicited strains entering a collection uncritically without due regard to collection objectives, storage capacity, personnel, and financial resources. Conversely, the range of a collection should not be restricted as to limit the effectiveness of the service provided to the scientific community.

New accessions may arise as a result of any of the following: (*1*) From the collection's own research; (*2*) From samples sent to the collection for identification; (*3*) From active solicitation by the collection; (*4*) They may arise unsolicited; (*5*) As a formal patent deposit under the Budapest Treaty.

Staff

The effective curation, management, and staffing of a culture collection is a demanding task. Routine accessions, preservation, maintenance, viability checking, culture supply, and additional services are all extremely time consuming. The staff have to have considerable knowledge not only about the organisms themselves, but also their growth and preservation requirements, properties, and potential applications. Where a culture collection offers other services, such as identification and taxonomic expertise, additional specialized staff are required. Experience in culture collection management in addition to taxonomy is a fundamental prerequisite to expertise, and it is therefore important to attract and retain staff of a high caliber.

Preservation and Maintenance

There are a number of techniques available for the preservation of microorganisms, each of which has particular advantages and disadvantages. In addition, different organisms require special preservation techniques in order to ensure optimal storage and maintain purity. The choice of technique should be determined by relating the features of each method to the needs of the user. A number of considerations must be taken into account when deciding which method is most appropriate.

Purity of samples is of paramount importance, and the preservation method should minimize the chance of contamination. The process of preservation and storage should be fully validated. Because preservation could result in loss of viability and hence cell death, a process causing minimum damage should be chosen. Viability should be ascertained immediately after preservation and reassessed during storage.

Preservation may introduce the possibility of a change in the characteristics of a culture. This can happen either as a result of a significant number of cells dying during the preservation process, thus selecting for a resistant population of surviving cells, or by genetic drift as a result of mutation or, for example, loss of plasmids.

The expense of preserving and storing cultures may also influence the choice of method, and this may be further influenced by the number of cultures and factors, including staffing, equipment, materials, and storage space. The operation time for the preservation and maintenance of a small collection may be too labor intensive for a large collection. The high capital cost for methods such as freeze-drying might be considered unsuitable for a small collection, although once preserved the samples require little maintenance. Storage in liquid nitrogen vapor may be cheaper, but for large collections, a sizeable storage space is required. Considerable attention must therefore be taken when the choice of preservation is being made, particularly with regard to the long-term objectives and size of a collection.

The consequence of loss of a culture is another factor to be considered when choosing a preservation method. Valuable samples should be preserved by methods that minimize loss and should ideally be preserved by more than one technique.

Finally, the choice of preservation technique will depend on the frequency of distribution of a sample. Samples that are regularly distributed should be preserved by a method amenable to bulk storage and that does not put them at risk to loss of viability or contamination as a result of frequent manipulation. Cultures that are to be supplied through the mail should be in a form suitable for packaging and must survive transportation time and temperature in accordance with IATA (1988) Dangerous Goods Regulations for Transport of Infectious Material (3).

There are four main storage methods briefly outlined in the next section. This subject has previously been reviewed in depth, for further information see the work by Kirsop and Doyle (4).

Subculture. Preservation of samples by continued subculture is open to the risk of complete loss on the one hand and contamination on the other. Samples are incubated on a suitable medium, at an appropriate temperature, and for a suitable time. The procedure is repeated at intervals that ensure that a fresh culture is prepared before the old one dies. Samples are therefore continually in danger through loss of viability, particularly because the survival time of a particular microorganism varies considerably. *Staphylococcus* spp. and coliforms will survive for several years, but *Neisseria* spp. require subculture within a few weeks.

The majority of media used for the storage of microorganisms are unenriched and have limited nutrients. Excess carbohydrates should not be included in the medium because the acid produced might kill the culture.

Storage periods can be extended by reducing the metabolic rate of the microorganism. This can be achieved by reducing the storage temperature to 5 °C in many instances. However, exceptions occur; *Neisseria* spp. survive better at 37 °C. Restricting the availability of air with parafilm can also limit metabolism.

Thus, repeated subculturing, together with mislabeling or transposition of cultures, can result in significant contamination of deposits. Genetic drift and instability of characters also increases with each subculture. A large inoculum reduces the risk of selection but increases the risk

of contamination. Therefore, this method is not to be recommended.

Freeze-Drying. Freeze-drying has found popularity with many service culture collections primarily because it is suitable for batch production and distribution. Maintenance of viability requires very little attention, storage being relatively simple and effective and allowing preservation for 50 years for some microorganisms. The capital cost of freeze-drying equipment is initially high, and the preparation of samples is quite labor intensive, but the fact that large batches of samples can be prepared, significantly reducing labor hours per ampule, quickly recoups the initial financial outlay.

The process of freeze-drying is the removal of water from frozen samples by sublimation. Cultures are suspended in a suitable medium (generally mannitol), frozen, and exposed to a vacuum. The water vapor is trapped in phosphorous pentoxide or a refrigerated condenser. The microorganisms are stored in individual ampoules or vials in an inert gas or under vacuum. There are two types of freeze-dryers in general use; however, the centrifugal rather than shelf dryer is more popular and has more advantages.

Although shelf dryers automatically stopper the ampules, the bungs can be naturally permeable to air or water vapor or theromas simply leak. The glass ampoules produced by centrifugal freezing are completely constricted, preventing leakage, and they can be plugged with cotton-wool, eliminating cross-contamination and acting as a filter to prevent scatter of microorganisms into the environment when the ampoules are opened.

Even though freeze-drying has been widely used to preserve a variety of microorganisms, the process always needs refining to optimize preservation in a particular microorganism. Factors that can be adjusted and controlled are composition of medium, growth temperature, growth phase of the culture, rate of freeze-drying, final temperature, duration of drying, and the final moisture content of the culture.

In general, freeze-drying does not cause instability of characteristics and genetic drift.

Drying. Desiccation has been used across a range of microorganisms, although the process is rather a specialized application and there is less information available on the method. Long-term viability appears to be good, when samples are stored between 5 and 10 °C, and contamination is not a major problem, although strain stability is less well documented. The simplicity of the process makes it suitable for storing large numbers of cultures.

L-Drying. The term *L-drying* originates from the term *drying from the liquid state* (5). Suspensions of microorganisms are dried in ampoules under pressure, but the vacuum is adjusted to allow rapid drying without freezing. The method is widely applied and allows a significant reduction in drying time and thus labor costs when a number of microorganisms are preserved. Samples should be stored at 5 °C.

Freezing. Preservation by freezing has been carried out for a variety of different microorganisms over temperatures ranging from −20 °C (refrigerator) to −140 °C (vapor phase nitrogen) to −196 °C (liquid nitrogen). Freezing and thawing can cause injury to cells by the formation of crystals that may damage the cells' integrity or by the concentration of electrolytes through the removal of water as ice. Eutectic mixtures form at −30 °C and above, exposing cells to high salt concentrations. Temperatures of −70 °C and below have been found to be increasingly successful, and storage in the vapor phase of liquid nitrogen is now ubiquitous as a preservation method.

Storage in the liquid phase of nitrogen results in considerable safety problems resulting from nitrogen invading vials that subsequently become explosive (Fig. 1). The freezing and thawing process itself may cause a significant loss in microorganisms. However, once stored, virtually no further loss occurs. The addition of cryoprotectants, the adjustment of growth conditions, the rate of cooling and warming, and freezing methodology can reduce losses (4).

The temperature of samples should be lowered at a controlled rate of 1 to 3 °C/min to −30 °C followed by a more rapid rate of 15 to 30 °C/min to −100 °C or lower. Ideally, this should be carried out using a commercial programmable freezer, which usually operate on a differential thermocouple principle. If such equipment is not available, many microorganisms, including bacteria and fungi, can be successfully frozen by placing the ampules in a dry chest, in a mechanical refrigerator set at below −65 °C, or

Figure 1. Storage of cells in the gaseous phase of liquid nitrogen.

in the top of a liquid nitrogen storage tank for 30 minutes, and then placing the samples at the required temperature.

Preprepared plastic ampoules containing beads and cryopreservant are now commercially available to enhance preservation (Prolab Diagnostics, Microbank). A vial of 25 beads is inoculated and agitated and the microorganisms attach to the beads. Excess cryoprotectant is then removed and the ampoule frozen. Beads can then be removed when a culture is required. The vials are placed in a cryoblock before they are removed from the storage vessel. This ensures that the remaining beads do not thaw out while one bead is removed. Thus, loss of viability does not result from repeated thawing.

The major disadvantage of freezing is refrigerator failure or interruption to the supply of liquid nitrogen. Both will result in the loss of cultures; however, this can largely be overcome by 24-h alarm systems warning of refrigerator failure or loss of liquid nitrogen supply. The initial capital cost is high, but the process is not labor intensive.

Preservation of Cell Lines

It is impractical for most laboratories to maintain cell lines in culture indefinitely; moreover, cell cultures will undergo genetic drift with continuous passage and risk losing their differentiated characteristics. Therefore, it is necessary to store adequate cell stocks for future use. It is recommended practice that a limit is set for the number of passages any cell line should undergo before replacing from cryopreserved stocks. This is of less importance for undifferentiated cell lines having infinite life span, although it still remains a consideration.

Nearly all cell lines can be cryopreserved successfully in liquid nitrogen at -196 °C. A few simple criteria must be adhered to in order to provide sufficient cells for all future needs.

The essentials for efficient cryopreservation are slow freezing (i.e., at a rate between 1 and 3 °C per minute) and fast thawing, which is achieved by placing ampoules in a water bath at the temperature required for growth. The addition of a cryoprotectant, such as glycerol or dimethyl sulphoxide (DMSO), enhances survival. Nevertheless, DMSO does have certain toxic properties. Therefore, care must be taken in handling the cryoprotectant, and some consideration may be necessary as to whether cells should be washed immediately on thawing.

Sample Safety and Security. For security, and in order to minimize the possibility of strains being lost, each culture should be maintained by at least two different procedures. Cryopreservation and lyophilization are the best methods for minimizing the risk of genetic drift, therefore, at least one of these methods should be used. Where only freezing is available, duplicates should be stored with different electrical supply. Duplicates of important or irreplaceable strains should be securely held in different buildings or ideally on a different site. The documentation of samples is equally important because loss owing to natural disaster or fire would make the corresponding cultures meaningless and worthless.

Cell Banking Procedures. When a culture is received in the laboratory, it is important that a strict accessioning scheme be followed (Fig. 2). The method described prevents substantial genetic variation from the original deposit. Starter cultures or ampoules received should be propagated according to the conditions provided by the originator, and a token freeze should be produced (usually up to 5 ampoules). Cultures derived from this token freeze should then be subjected to detailed quality control and characterization. If these results are satisfactory, then a master (or seed) stock should be made. Major authentication of the culture should then be undertaken on the master cells because it is these ampoules that will be used to provide the working or distribution stock.

One ampoule from the master stock is used to produce a working or distribution stock. The number of ampoules in each bank is dependent on how often it will be used. Because the production of these banks is both expensive and time consuming, the numbers required should not be underestimated. For some industrial processes, master banks may contain up to 200 ampoules, and working banks may contain 1,000 ampoules.

Once the working bank is depleted, a fresh one is made from an ampoule from the master bank; similarly, a new master bank is made from the frozen freeze stock. If no more frozen freeze stock remains, then one of the current master bank ampoules should be used.

The level of characterization and quality control that should be carried out at each stage will depend on what it is to be used for and what the necessary regulatory bodies require (Table 1). Quality control implications are obvi-

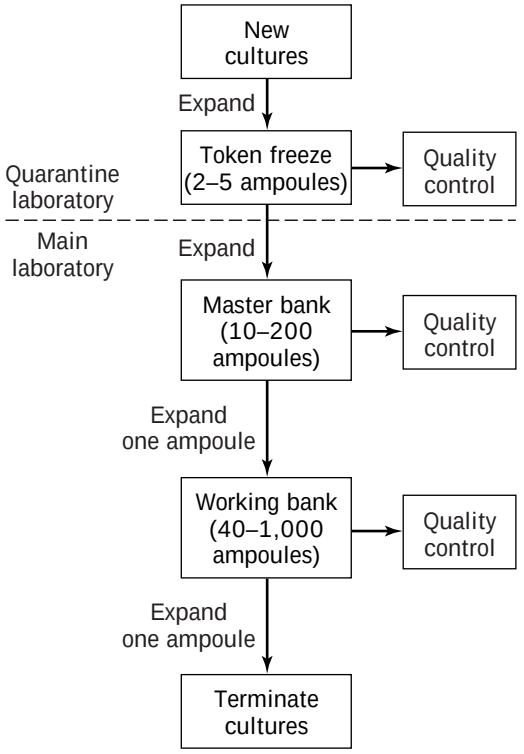


Figure 2. Cell banking procedures.

Table 1. Table of Quality Control Requirements for Cultures within Collections

	Culture type			
Cell lines	Bacteria	Fungi	Protozoa	Algae
Viability	Viability	Viability	Viability	Viability
Purity	Purity	Purity	Purity	Purity
Test for bacteria, yeast, and other fungi				
Test for mycoplasma				
Test for viruses				
Authenticity	Authenticity	Authenticity	Authenticity	Authenticity
Isoenzyme analysis	Gram Stain			
Cytogentic testing	Description of morphology			
DNA fingerprinting				

ously more intense for cultures used in the manufacturing industry.

CULTURE COLLECTION SERVICES

The aim of culture collections is to supply authenticated cultures to bona fide scientists, on request, promptly and without restriction on their ultimate application. The staff of culture collections have expertise and knowledge of the appropriate scientific literature and as such can advise on strain selection. In addition, they provide a range of scientific services, largely in response to the needs of the user. These include safe and patent deposit facilities and identification and preservation services. The expertise gained by culture collections and their staff enables them to run training courses and seminars for other scientists.

Culture Supply and Pricing Policy

Culture collections should be able to readily distribute cultures listed in their catalog on request. Arrangements for culture supply vary according to the financial basis and policies of the legal owner of the collection. Service-supply collections usually levy fees for supplying cultures. This charge should strike a balance between definable measurable costs and those that yield an income that is significant in relation to costs, but that do not result in a declining effective use of the collection.

Consideration should be taken as to the fact that many collections are partly government funded, unit costs would be far too high for what is a reusable commodity (i.e., they can be subcultured), and fees should not discourage the use of the collection. Therefore, culture collections operate on a nonprofit basis. It is often common practice to charge commercial companies higher fees than nonprofit organizations. Further charges may be levied for postage and packaging and any special documentation.

Great care should be given to distribution of pathogenic or toxic cultures. Originators of a request should be able to guarantee that their establishment has stipulated equipment and staff to handle pathogenic or toxic cultures. Dispatch of cultures should be in accordance with IATA (1988) Dangerous Goods Regulations for Transport of Infectious Material (3).

Restrictions to which countries cultures can be distributed also exist. Because the threat exists that microorganisms could be developed for use in biological warfare, cultures can only be issued to countries in the Australia Group of nations. The nations belonging to this group and restricted groups are detailed in *The End-Use Control* (6).

Catalogs

Catalogs are fundamental to the operation of culture collections. They describe the strains available and their salient properties, including literature citations, propagation properties, etc. Stocks are usually arranged alphabetically, in the case of bacteria by genus and species name. Increasingly catalogs are becoming available in computerized form, either on-line or on disk, because, they can can easily be updated whereas paper catalogs cannot. Such catalogs adopt a field structure and definitions that enable the data to be integrated into the international and major regional schemes that are in operation, including the Internet, on which many resource centers are accessable through the World Wide Web, Microbial Strain Data Network (MSDN), World Data Centre on Micro-organisms (WDC), Microbial Information Network Europe (MINE), and more recently CABRI (Common Access to Biotechnological Resources and Information).

CABRI is researching and developing the formation of an integrated resource center service linking databases of different organism type, genetic materials, and other biologicals in Europe so that users worldwide can access these relevant catalogs during one searching session through a common entry point and request or order products to be delivered to their place of work.

Safe Deposits

Safe deposits are held for a variety of reasons, including as a back-up to existing collections and for the retaining of ownership and confidentiality of strains. These collections are not listed in catalogs and their taxonomic status may be unclear. In some cases, the service collection may hold stocks in its collection and market them on behalf of the owner.

Patent Deposits

The rapid advance in genetic manipulation techniques and the allied boom in biotechnology over recent years has led to increased awareness of the enormous industrial poten-

tial of microorganisms. Competition is fierce in the race to exploit this potential, and large sums of money are being risked in commercial development of new microbial processes. In such circumstances, patent protection naturally becomes a major concern of any organization involved in modern industrial microbiology.

An important principle of the patent system is that an invention should be disclosed to the public. Although most kinds of invention can be disclosed simply by means of written description, those involving the use of new microorganisms cannot. No matter how detailed the written description may be, a new microbiological process cannot be tested without a culture of the organism itself. Therefore, in addition to a written disclosure, the patent offices of most industrial countries require the strain used in a microbiological invention to be deposited in a recognized public culture collection, where it will eventually be available to the public.

Patent deposits are now regulated internationally by the Budapest Treaty on the International Recognition of Deposit of Microorganisms for the purpose of patent procedures and patent depository facilities in Europe, under which certain culture collections of a generally acknowledged high scientific standard are officially recognized as international depository authorities (IDAs). The main feature of the treaty is that all countries party to it, and hence their patent offices, must recognize a deposit that is made in any one of these IDAs as being valid for the purposes of their patent procedure. This means that one deposit is all that is needed, regardless of the number of countries in which patent applications citing the strain are filed.

The Budapest Treaty contains detailed provisions on the procedure to be followed by the applicant in making a deposit and by the IDA in accepting, storing, and subsequently making the material available to the public. Broadly speaking, these provisions are designed to ensure that the applicant makes a viable deposit and provides enough information on it to enable the IDA to handle and maintain it safely and correctly. The IDA, for its part, must process the deposit promptly, issue the depositor with a receipt and certificate as proof of deposit and viability, keep the deposit for at least 30 years, and provide cultures to those parties entitled to receive them but most importantly, *to no one else.* Entitled parties in this respect are those who have the depositor's written authorization or who have an official certificate of entitlement from the appropriate patent office. The IDA is also bound by strict rules of secrecy and may not give information about a deposit to anyone not entitled to receive the patent-deposited cultures.

Microorganisms must become available to the public at some stage of the patenting procedure. To date, there are three kinds of systems in operation. In the Netherlands and Japan, patent applications are published twice, first 18 months after the filing date of priority date and before the application has been examined, and second when the application has been accepted. Once the patent office has decided to grant a patent, after the second publication date, the microorganism must be made available. The major advantage of this system to inventors is that their microorganisms do not have to be released until they has an enforceable right. In Japan, the patent owner is given a further measure of protection in that the recipients of cultures must not pass them on to a third party and must use them only for experimental purposes.

In the United States, patent applications are not published until the patent is granted, and any microorganism deposited for patent purposes need not be available until then. Thus, under the U.S. system, inventors again have an enforceable right at the time they are required to make the organism available. Additionally, under the U.S. mechanism an inventor is never put in the position of having to allow access to the organism when no legal protection is available because if a microorganism is not granted a patent, then it need never become available.

The EPC (European Patent Convention) also operates a dual publication system, but stipulates that a microorganism for which a patent is being sought must be made available at the date of first publication, before any enforceable right exists. This mechanism reflects the principle that the microorganism is an integral part of the patent application and should therefore be available at the same time as the written description. Availability at first publication can be restricted to an independant expert acting on behalf of a third party. This expert has to be selected from a list held by the EPC and is not permitted to pass the strain on to anyone else. On second publication, the strain becomes generally available, whereupon the inventor has an enforceable right.

There are 23 IDAs. A full description of these is described by Halluin (7).

Culture Authentication and Identification

Culture collections have a critical responsibility when it comes to the authentication of microorganisms and cells from their initial receipt through preservation procedures and to dispatch to external scientists. Quality control of cultures should be as extensive as possible and have three primary targets; authenticity, purity, and viability. Research using the wrong organism wastes time, is expensive, and leads to publication of invalid results.

The identification of a new accession should be ideally confirmed by a competent specialist (e.g., the depositor) or confirmed by the culture collection itself to ensure that it agrees with published descriptions. Culture collections with specialized expertise in identification can also offer an identification consultancy to the scientific community. The microorganisms identified would not necessarily enter the culture collection.

Documentation and Quality

Increasingly, many culture collections operate using a quality management system such as BS EN ISO 9001. Essentially, these regulations ensure that all procedures have proper protocols that are fully documented, thus allowing all activities to be fully traced if the need arises. All equipment has to have records of calibration and servicing, and all laboratory records, worksheets, and equipment log

books are regularly audited. All staff are trained in these procedures.

In order for a culture collection to run completely successfully, it has to keep good records about the microorganisms and cells it holds, including the following categories of information, place, substrate or host, date of isolation, name or person isolating the strain, depositor, name of person identifying the strain, preservation procedure used, optimal growth media and temperature, any data on biochemical or other characteristics, and any regulatory conditions applying (e.g., contamination level).

Duplicate copies of computer files or paper records of all information should be kept to prevent unexpected loss of information.

Research

It is important that culture collections have an ongoing research program. This ensures that staff are kept abreast of current developments and are aware of the needs of the wider scientific community. Lack of adequate research programs discourages high-caliber staff from joining a culture collection and may encourage the loss of existing quality staff. As well as the obvious research and development into preservation and taxonomy, culture collections have the necessary skills to undertake a diverse spectrum of research, whether by themselves or acting as consultants. As discussed earlier, culture collections have the capacity to screen isolates for particular biochemical properties and metabolite production, evaluate and develop commercial kits and rapid detection kits for microbially produced compounds, evaluate organisms' capacity for use in bioremediation and as biocontrol agents, and select test organisms used in the evaluation of materials, media, culture vessels, and diagnostic reagents.

Table 2. The European Culture Collection Organization and the World Federation for Culture Collections

President and Secretary of ECCO and WFCC

Dr. Alan Doyle
CAMR
Porton Down
Salisbury SP4 0JG
United Kingdom
Tel: +44-1980-612512
Fax: +44-1980-611315
E-mail: *alan.doyle@camr.org.uk*

Secretary of ECCO

Dr. Maija-Liisa Suihko
VTT
Biotechnology and Food Research
PO Box 1501
FIN-02044 VTT Finland
Tel: +358-0-4565133
Fax: +358-0-4552028
E-mail: *maija-liisa.suihko@vtt.fi*

WFCC's Web Page: http://www.wdcm.riken.go.jp/wfcc.html

Safety

It is the responsibility of all culture collections to handle cultures and carry out procedures as safely as possible and within the relevant guidelines. This is particularly important with regard to microorganisms that are pathogenic to man or are genetically manipulated. Within the United Kingdom, all laboratory procedures should be performed in line with COSHH (Classification of Substances Hazardous to Health), and all matters of safety are enforced by the HSE (Health and Safety Executive), which also offers an advisory service.

All culture collections are responsible to their own staff, the delivery services, and receiving laboratories and must therefore address all matters of safety properly. All service collections have regulations for both the deposit and dispatch of cultures that have to be fully abided by before samples can be transported following the IATA (1988) Dangerous Goods Regulations for Transport of Infectious Material (3).

Within the United Kingdom, microorganisms are categorized in accordance with the Advisory Committee on Dangerous Pathogens (1995) Categorization of Pathogens According to Hazard and Categories of Containment. In the United States, a similar categorization has been devised (U.S. Department of Health, Education and Welfare, 1974; American Type Culture Collection, 1986). Both sets of guidelines have been accepted by the WHO, and other countries tend to adopt these systems. Information is given on the degree of containment and protective clothing, which should be applied during handling of such organisms in the laboratory.

The control of genetically manipulated organisms is somewhat more complicated in the United States than it is in the United Kingdom. U.S. researchers must adhere to the policy that research in this field must conform to the requirements of the coordinated framework, such as the National Institutes of Health Recombinant DNA Guidelines. In the United Kingdom, genetic manipulations are the responsibility of the Advisory Committee in Genetic Manipulations. Other countries likewise tend to adopt these systems.

All culture collections must additionally follow the regulations of their own countries and the institutes within which they work.

Training and Publications

Culture collection staff require appropriate training for all techniques and procedures they carry out. If an organization has a quality management system, such as ISO 9000, then this will be performed as a mandatory requirement. Once the staff have become skilled, they are then in an ideal position to train others in techniques relating to culture preservation, growth, and identification given that they have adequate teaching facilities and supervision.

Regular publication relating to culture collection activities and the research activities carried out within an organization is necessary to enhance its activities.

Table 3. Major Culture Collection Repositories

Country and organization	Address	Collection
	Belgium	
BCCM/IHEM	Mycology Laboratory Rue Juliette Wystman 14 B-1050 Brussel Belgium Fax: +32 2 6425519 E-mail: Dialcom Telecom Gold: 75: DB I0248	Fungi
BCCM/LMBP	Plasmid Collection University of Gent K.L. Ledeganckstraat 35 B-9000 Gent Belgium Fax: +329 2645348 E-mail: *martine@lmbl.rug.ac.be*	Plasmids
BCCM/LMG	Bacterial Collection University of Gent K.L. Ledeganckstraat 35 B-9000 Gent Belgium Fax: +32 9 2645346 E-mail: *danielle.janssens@rug.ac.be*	Bacteria
BCCM/MCLK	Faculte des Sciences Agronomiques Flace Croix du Sud 3 B-1348 Louvain-La-Neuve Belgium Fax: +3210451501 E-mail: *mucl@mbla.ucl.ac.be*	Fungi
	Brazil	
CCT	Tropical Culture Collection Rua Latino Coelho 1301 Parque Taquaral CEP 13087-010 Campinas - SP Brazil Fax: +55 (019) 242 7827 E-mail: *cct.@bdt.org.br* *World Wide Web: http://www.bdt.org.br/cct/*	Bacteria Fungi Yeasts
	France	
CIP	Institut Pasteur BP52 25 rue du Docteur Roux F-75724 Paris Cedex 15 France Fax: +33 1 40613007	Bacteria
LCP	Museum National D'Histoire Natwelle Laboratoire de Cryptogenie 12 rue Buffon F-75005 Paris France Fax: +33 1 40793594 E-mail: *roqueber@frmnhn.fr*	Fungi
	Germany	
DSMZ	Mascheroder Weg 1b D-38124 Braunschweig Germany Fax: +49 531 2616418 E-mail: dsmz@gbf.de World Wide Web: http://www.gb.de.bl/DSMZ	Bacteria Plant cells Animal and human cells Plant viruses Plasmids

Table 3. Major Culture Collection Repositories (*continued*)

Country and organization	Address	Collection
	Italy	
ICLC	CBA L.go R Benzi 10 16132 Genova Italy Fax: +39 10 5737295 E-mail: *iclc@istge.ist.unige.it* World Wide Web: *http://arginine.umdrj.edu/home.html*	Human and animal cells
	Japan	
IFO	17-85 Juso Honmachi 2-Chome Yodogawa-Ku Osaka 532 Japan Fax: +81 6 3006555 E-mail: *cellbank@ub.infaveb.ne.ip*	Human and animal cells
Riken Gere Bank	3-1-1 Koyadai Tsukuba Science City 305 Iboraki Japan Fax: +81 2975 4 2616	
	The Netherlands	
CBS	PO Box 273 Oosterstraat 1 NL 3740AG Baarn Netherlands Fax: +31 355 416142 E-mail: *Stalpers@cbs.knaw.nl* World Wide Web: *http://www.cbs.knaw.nl*	Fungi Yeasts
	Spain	
CECT	Universidad de Valencia Edificio de Investigation Campus de Burjasot Valencia Spain Fax: +3496 3983187 E-mail: *cect@uv.es*	Bacteria Fungi
CCUG	University of Goteborg Dept of Clinical Bacteriology Guldhedsg 10 S413 46 Goteborg Sweden Fax: +46 31 825484 E-mail: *ccugef@ccug.gu.se*	Bacteria
	United Kingdom	
CCAP	Institute of Freshwater Ecology The Windermere Laboratory Far Sawrey Ambleside Cumbria LA22 OLP UK Fax: +44 15394 46914 E-mail: *ccap@ife.ac.uk*	Protozoa Freshwater algae

Table 3. Major Culture Collection Repositories (*continued*)

Country and organization	Address	Collection
CCAP	Dunstaffnage Marine Laboratory PO Box 3 OBAN Argyll PA34 4AD UK Fax: +44 1631 562244 E-mail: *ccap@dml.ac.uk*	Marine algae
NCIMB	National Collection of Industrial and Marine Bacteria 23 St. Machar Drive Aberdeen AB2 1RY UK Fax: +44 1224 487658	Bacteria
NCTC	National Collection of Type Cultures Institute of Food Research Colney Lane Norwich NR4 7UA UK Fax: +44 1603 458414 E-mail: *ian.roberts@bbsrc.ac.uk*	Bacteria
	United States	
ATCC	American Type Culture Collection PO Box 1549 Manassas Virginia 20108-1549 USA Fax: +1 703 365 2750 World Wide Web: *http://www.atcc.org/*	Bacteria Viruses Animal and human cells Plasmids Plant Cells
GIMR	Coriell Institute for Medical Research 401 Hadden Avenue Camden NJ 08103 USA Fax: +1 609 757 9737 World Wide Web: *http://locus.umdnj.edu/home*	Human cells/DNA

National and International Collaboration

Microbiologists having interests in culture collections have banded and assembled in formal or informal national federations or associations of the culture collections within them, such as UKFCC (U.K. Federation of Culture Collections). These federations tend to be further affiliated with international bodies, such as ECCO (European Culture Collections Organisation), which comprises 54 collections from 21 European countries. This organization itself is affiliated to WFCC (World Federation for Culture Collections). These federations provide excellent opportunities for discussion and exchange of information in all aspects of culture collection curation.

European Culture Collection Organization. ECCO was established in 1981. The aim of the organization is to promote collaboration and trade ideas and information about all aspects of culture collection activity. Membership is open to representatives of any resource centers that provide a professional service on demand and without restriction, that accept cultures for deposit, that provide catalogs, and that are housed in countries with microbiological societies affiliated to the Federation of the European Microbiological Society (FEMS). After 14 years of activity, ECCO comprises 54 members from 21 European countries. The total holdings of the collections are about 300,000 deposits representing a large pool of biodiversity from different sources (from humans, animals, plants, foodstuffs, environmental) as well as various sources of biotechnological interest, including GMO strains and organisms of interest for genetic engineering.

World Federation for Culture Collections. Most of the culture collections organized in national or regional federations are members of the WFCC. This international organization has been set up as a multidisciplinary commission of the International Union of Biological Science and is a federation within the International Union of Microbiological Societies (IUBS). Various committees exist within the WFCC for carrying out specific tasks in the different aspects of culture collection work. These include the publication of books, provision of training courses, collaboration with postal and patent organizations, setting up databases, providing help for endangered culture collections,

and designating quality standards and guidelines for collections. One of the major activities of WFCC is to organize its quadrennial meetings (International Congress of Culture Collections).

The World Data Centre (WDC), now the World Centre on Microorganisms (WDCM), covers a central role in WFCC. Initiated in 1972 by Prof. V.B.D. Skerman and sponsored by UNEP and UNESCO, the database pioneered the organization of microbial culture collections worldwide. Relocated to the RIKEN Institute in Japan in 1986, the WDC, under the directorship of Dr. H. Sugawara, developed into a valuable information resource at the microbial level. Several databases are available as hard copies and are also accessible through the Internet. The core databases, CCINFO and STRAIN, are regularly updated and published as the WDC Directory. CCINFO covers information on nearly 500 microbial resource centers in the world, their organization, the kinds of cultures held, the expertise available, and the services offered. Strain lists all microbial species held by registered collections. The WDCM activity of continual data collection and dissemination plays an important role in the improvement of data quality of collections and for making the resource centers known to the scientific community.

Any individual with an interest in culture collection activities is eligible for membership. Culture collections as institutions may become affiliate members if they meet the standards established by the federation. Individuals or organizations may be accepted as sustaining members on the basis of extraordinary support of the objectives and activities of the Federation. Further information can be obtained from the WFCC (Table 2).

Access to Resources

Hard copy catalogs get out-of-date rapidly; the best and most up-to-date information can be obtained from most resource centers on the World Wide Web. Table 3 gives a comprehensive list of contact information for the major repositories.

BIBLIOGRAPHY

1. M. Kocur, *Proc. of the Kral Symposium*, Institute for Fermentation, Osaka, Japan, 1990, pp. 4–12
2. WFCC, *Guidelines for the Establishment and Operation of Collections of Microorganisms*. Simworth Press, Richmond, Surrey, U.K., 1990.
3. G. Stacey, A. Doyle, and P. Hambleton eds., *Safety in Tissue Culture*, Kluwer, London, 1998; IATA, 33 route de l'Aeroport, PO Box 672, CH-1216 Geneva 15, Switzerland.
4. B.E. Kirsop and A. Doyle, *Maintenance of Micro-Organisms and Cultured Cells. A Manual of Laboratory Methods*, 2nd ed., Academic Press, London, 1991.
5. D.I. Annear, *Aust. J. Exp. Biol.* **36**, 211–222 (1958).
6. DTI, *The End-Use Control*, Department of Trade and Industry, U.K. 1995, pp. 1–16.
7. A.P. Halluin, in J.C. Hunter-Cevera, and A. Belt eds., *Maintaining Cultures for Biotechnology and Industry*, Academic Press, London, 1996, pp. 1–14.

See also ASPERGILLUS; CELL CYCLE; CELL CYCLE, EUKARYOTES; FREEZE DRYING, PHARMACEUTICALS; SECONDARY METABOLITE PRODUCTION, ACTINOMYCETES, OTHER THAN STREPTOMYCES.

CULTURE MEDIA, ANIMAL CELLS, LARGE SCALE PRODUCTION

JENNIE P. MATHER
Rauen Biotechnologies, Inc.
Mountain View, California

ALISON MOORE
Amgen, Inc.
Newbury Park, California

KEY WORDS

Cell line
Fermenter
Large scale
Nutrient optimization
Product titers
Serum-free
Suspension

OUTLINE

INTRODUCTION

The process of choosing and optimizing a cell culture medium for a large-scale production process shares many biological issues and experimental approaches with the process of optimizing medium for any in vitro model system at the laboratory scale. However, there are some significant differences. The cell culture aspect of any large-scale process should fit in seamlessly with the upstream process of the transfection and selection of the protein of interest in the desired parental cell line, the equipment and scale available and desirable to run the process, and the downstream purification process. Issues surrounding the speed

of development and the cost, safety, and reproducibility of the process are of major importance in developing a commercial process.

The importance of seeing the entire process—from selection of the host cell line through transfection and selection of the production line, to the growth of the seed train cells and the actual production culture, harvest, and purification of the protein—as an integrated whole cannot be overemphasized. Failure to do so will result in an unworkable process, at worst, and missed opportunities for maximizing the reliability and minimizing the cost, at best. If the end goal is kept in mind from the first, many opportunities for trade-offs in various aspects of the process will arise that allow decisions to be made in a timely manner. For example, if two candidate production lines exist, one with high specific productivity but slow doubling time and low saturation density and the other with lower specific productivity but better growth characteristics, the choice between the lines should be made taking into consideration the titer from each line and the total tank and equipment usage (seed train, production, and purification) required to obtain a given final yield of protein. It might be surprising to some how taking this broader view will change decisions made early in the process that are much harder to change later. In another case, one medium formulation was found to increase the stability of a cell line even though the titer was reduced compared to the standard medium formulation. The advantages of the increased stability were seen to far outweigh the minor loss of product in the short term.

This chapter will concentrate on specific aspects of choosing and optimizing the cell culture medium for large-scale processes. However, as will be emphasized again in the section at the end, this cannot be achieved optimally without taking into consideration the product (as well as the cell line) to be secreted and the purification process and equipment to be used. Some of the basic parameters that influence protein yield from a mammalian cell culture process are outlined in Table 1 and are discussed in detail subsequently.

OPTIMIZING CELL GROWTH AND SURVIVAL

The aim of optimizing the growth and survival of cells in production culture is to obtain the maximum average viable cell number per volume of medium over the duration of the culture (viable cell days [VCD]). Characteristics of the cell that will influence the total VCD are listed in Table 2. When selecting the production line to be used for large-scale culture, these parameters and specific productivity should be taken into consideration. A transfected cell line with high specific productivity is useless as a production line if it cannot be grown in the conditions necessary for large-scale culture. For example, a tetraploid line may have a high specific productivity but be too fragile to do well in a fermenter. Or, a low-cell inoculum may be acceptable with a rapidly growing cell line, but not with a slow growing line producing a labile product. Continuous culture systems strive to produce a steady-state situation where the cell number is constant and the growth rate and death rate are balanced. Ideally, one would wish to achieve something like a contact-inhibited G_0-arrested state in which the cells are neither growing or dying. This is based on the as-yet-unproved assumption that suspension cells producing recombinant proteins will have a specific productivity in G_0 equal to or higher than that in log phase growth. In our experience, however, it is difficult to achieve a true G_0 growth arrest state in suspension-cultured cells (1).

Although a great deal is known about the optimal growth conditions for specific types of cells, including Chinese hamster ovary (CHO), baby hamster kidney (BHK), and transformed human embryonic kidney (293) cells, it is still impossible to exactly predict the optimal medium formulation for every cell type in every culture configuration. In addition, the process of obtaining a stably transfected cell line expressing a recombinant protein can change the growth phenotype of the cells. Thus, the empirical approach is still the only method of ensuring that the medium being used is optimal for a particular production line. The optimal medium components or concentrations may vary from one CHO production line to the next. This is not surprising given that these lines are transfected with foreign DNA, selected using rather stringent conditions, and frequently amplified in the presence of a mutagen such as methotrexate. Additionally, the lines are engineered to secrete biologically active molecules that may act on the cell

Table 1. Parameters Influencing Final Product Yield in a Production Cell Line

Medium or culture configuration
Optimize nutrients
Optimize growth supplements
Minimize inhibitors
Inducers
Protein production
Specific productivity
Protein quality
Compatibility with purification process

Table 2. Characteristics to Be Considered When Choosing a Production Cell Line

Cell growth
Inoculum density
Saturation density
Doubling time
Death rate
Fragility
Suspension growth
Growth in serum-free medium
Other cell characteristics
Free of adventitious agents
Normal glycosylation
Has appropriate processing of proteins
Protease secretion low

or the medium components, altering the environment of the culture. Generally speaking, however, two media optimal for the growth of two CHO-derived cell lines will be similar to each other and differ from the medium optimized for the growth of a dissimilar cell line such as a 293-derived line. For this reason, it is best to concentrate on one or a few cell lines for large-scale production efforts. Currently, most recombinant proteins produced commercially in mammalian cells use CHO as the host cell line.

Measuring Viable Cell Number

Accurate measurement of cell viability or characterization of cell death is key to the optimization of any production cell line. Viability can be readily measured by a variety of well-known techniques; however, it has more recently been established that many cell lines grown under batch culture conditions undergo cell death via apoptosis (1,2). Apoptosis is an active programmed cell death that exhibits specific morphological traits, such as nuclear condensation and fragmentation, maintenance of organelle and membrane integrity, and gross cell shrinkage (3). This contrasts with necrotic cell death, which is passive in nature and recognized by the degradation of cellular membranes and organelles and cell swelling (4). These different forms of cell death are distinct in morphology, mechanisms, and the way in which they can be identified. An understanding of such is required for accurate reciprocal viability quantification, because the observation of simple loss of membrane integrity (e.g., using trypan blue exclusion) may be insufficient to account for all cell death caused by apoptosis. This may also be important in the accurate calculation of specific productivity. Finally, if one desires to improve the yield in an industrial process by decreasing cell death in the culture, it is crucial that one understand the mechanism of cell death in the cultures and the extent to which death decreases total VCD.

Viability can be assessed most simply by the exclusion of charged dyes such as trypan blue or propidium iodide (PI) that move freely into dead cells and exhibit minimal uptake into live cells. The extent of cell death can then be quantified by microscopy, or in the case of PI, by flow cytometry (1,5). Loss of plasma membrane integrity also results in the release of intracellular contents into the media that may then be quantified, such as the enzyme lactate dehydrogenase (LDH) (6). LDH content can be assayed in both the cellular and supernatant fractions, allowing relative quantification of cell death.

Apoptotic morphology specifically can be readily identified by staining fixed cells with a fluorescent dye such as acridine orange, which fluoresces green when bound to double-stranded nucleic acids (7). The classic apoptotic morphology of condensed or fragmented nuclear material can then be observed. An assessment of the percent apoptotic CHO cells during the course of a production run in a 2-L fermenter is shown in Figure 1. From day 7 there is an abrupt increase in the percentage of cells that show the characteristics of apoptotic cell death even though the total cell number in the culture does not change appreciably over this time period. However, the percentage of apoptotic cells increased from day 7 to a value of 50% apoptotic cells on day 11. Clearly, this large proportion of apoptotic cells will adversely affect the titers late in the run.

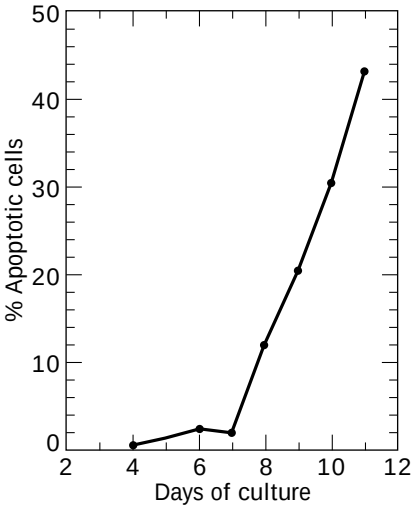


Figure 1. Apoptotic cell death in fermenters. Apoptosis was identified by light microscopy after acridine orange staining. Values are expressed as percent of total cell number that are apoptotic on days 4 to 11 of a 2-L minifermenter run using CHO cells.

An additional characteristic of apoptotic cell death that can be used for quantification is that of DNA strand breaks, which are introduced at the internucleosomal linker sections of DNA in apoptotic cells (8). The ultimate loss of short DNA fragments, coupled with reduced accessibility to intercalating fluorochromes such as PI because of the condensation of chromatin, often results in the congregation of apoptotic cells within a fixed cell population to a hypodiploid peak, when analyzed by flow cytometry (5). The percentage of cells localized to this peak can be quantified if not obscured by excessive amounts of cellular debris.

The free 3′-hydroxyl groups at the ends of the apoptotic DNA strand breaks may also be labeled with a fluorescent nucleotide using terminal deoxynucleotidyl transferase (9) or annexin (10). These techniques allow the fluorescelnation of apoptotic cells with only a low level (or no) labeling of nonapoptotic cells. Cells labeled in such a way may be counterstained with PI, allowing concurrent cell cycle analysis, so that exit from the cell cycle can be tracked by flow cytometry (1). These methods can be used to trace the effects of manipulating the culture parameters on the amount of cell death in the cultures (11). Such experiments are an invaluable adjunct to direct empirical determination of conditions that optimize titer because they can help in deciding which experimental manipulations, among the many possibilities, might yield benefits.

Considerations of Scale

Virtually all industrial productions of recombinant proteins or antibodies in mammalian cell culture are done on a large scale (100 to 12,000 L). The few exceptions may be production of antibodies or proteins for research use only. Although even these small- and intermediate-scale pro-

cesses will benefit from using an optimized medium, any medium optimization will obviously be aimed at use in large-scale processes. We have found that suspension-adapted cells and suspension processes scale well from plates to fermenters. Spinners are actually less predictive, in some cases, than static plate cultures using suspension cells. (Suspension-adapted cells grown in tissue culture plates in medium containing additives such as F-68 frequently will not attach but remain in suspension.) The experiments described in this article can be performed, preferably using the suspension-adapted cells, in plates, including 96-well plates for high throughput screening. However, larger (e.g., 60 mm) plates allow multiple sampling from the same plates over time. The type of careful empirical derivation of optimal process parameters described herein really cannot be performed entirely in fermenters because too many conditions are required. However, these experiments can help narrow the number of parameters to be tested on a larger scale. Figure 2 presents data on yield of identical cells in identical medium cultured in small- and large-scale vessels. The cells in 60-mm plates, roller bottles, and spinners all performed similarly to those in the 1,000-L fermenter. The static shake flask cells that performed very poorly may have been suffering from oxygen deprivation. A few experiments, such as adjustments in medium volume, might allow closer congruence between the small-scale model selected and the performance in large-scale fermenters. All major changes or improvements to a process based on results from small-scale work should be confirmed using fermenters. The use of 2-L minifermenters allows replication of conditions and good statistical analyses of results. If these titers are then confirmed in larger-scale fermenters, one can be confident that a rational approach using small-scale cultures, which allows complete analysis of the interacting variables in the medium, has led to a robust and optimal process.

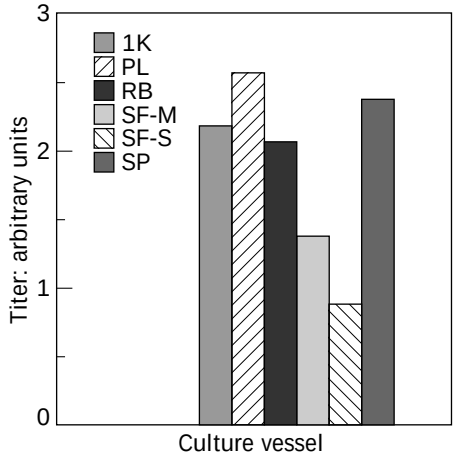


Figure 2. Cells were inoculated into a 1,000-L production fermenter. An aliquot of these cells was removed and dispensed into the various culture vessels on day 0. All vessels were sampled and the titers assayed throughout the culture period. The data shown compare the titers of product produced by cells in the different vessels on day 5 of culture. 1K, 1,000-L fermenter; SP, 1-L spinner; RB, roller bottle; PL, 60-mm tissue culture plate in CO_2 incubator; SF-S, static shake flask in incubator; SF-M, shake flask mixed using a stir bar in incubator.

Although attached cell culture systems can be scaled up via roller bottle microcarrier culture, hollow fiber systems, and so forth, these devices have generally been used for medium-scale rather than very-large-scale production. Each introduces a different set of variables that should be understood to design a small-scale equivalent system for optimizing medium, as described later. Again, any improvements should be compared directly to the standard conditions before adoption.

Some parameters change, of necessity, when going from static cultures to fermenters. The method of pH control differs because base and acid addition are used to regulate pH in fermenters in addition to the bicarbonate buffer. This results in a more stable pH over the course of a run in fermenters, compared to plates, but at the cost of a greater increase in osmolarity in the fermenters because of the added salts. Shear and foaming are also factors in fermenters, but not plates. Components of the medium that are added to fermenters to prevent shear and foaming should also be added to plate experiments when optimizing medium to allow for better concurrence between large- and small-scale results. Finally, the control of P_{O_2} is quite different in plates (diffusion) and fermenters (air or oxygen sparging). Adjusting the total surface area to volume in plates and spinners can have a significant effect on protein production and cell number and help in creating closer agreement between small- and large-scale data.

Some parameters such as pH, temperature, and P_{O_2} can be adjusted or controlled more easily in a fermenter. These parameters are best optimized directly in experiments performed in fermenters after the nutrient and growth supplements have been optimized at small scale and tested at large scale. It is important to optimize these parameters both for improved cell growth and, on occasion, to improve product quality. With a few experiments to directly compare small- and large-scale culture, and an understanding of the principle underlying differences in plate and fermenter cultures, the former can be an invaluable tool in the process of obtaining optimized media for large-scale industrial processes.

Nutrient Component of the Medium

For convenience, the culture medium can be divided into two components: the nutrient portion and the added growth factors and growth-promoting substances. The nutrient components are generally part of the commercially available powdered nutrient mixtures, such as MEM (minimal essential medium), DMEM (Dulbecco's Modified Eagle's Medium), Ham's F12 nutrient mixture, RPMI 1640, or F12/DME. These media were all empirically derived for very specific cell types and growth conditions (12), none of which are very similar to a suspension culture of genetically modified cells in a fermenter. However, they still make a good starting point for optimizing the culture medium for the specific production line of choice. An initial screen should compare several commercially available or proprietary media to determine which is optimal for the

production line. If the medium is being optimized before the cells are transfected, cell number can be used as the end point. If the line is already producing the protein of interest, protein titers should be measured. If the conditions for obtaining maximal titers and maximal VCD are the same, this is a good starting point for optimizing the nutrient portion of the medium. If different media give optimal VCD and titer, then more studies should be undertaken to understand the biology behind this phenomenon. This will prove useful in designing the optimal production system.

Once a medium is found that works well, there is a temptation to move on the assumption that this adequate. However, large-scale commercial culture is one situation in which it is almost always worth the time and effort required to obtain a truly optimal medium. Optimizing the nutrient portion of the medium is important not only for the increased yields it will afford directly but because optimizing nutrients will lead to a decreased requirement for parts of the growth supplements such as serum or growth factors (13). Additionally, optimizing VCD will lead to the ability to prolong culture time while minimizing damage to the product caused by enzymes released from dying cells. Even a 10% increase in VCD can be significant, and improvements of 2-fold to 10-fold in titer are frequently seen with optimized media. Because it is very difficult to change a process after clinical trials have begun, these early improvements will result in accrued benefits extending over many years that are often worth millions of dollars during the life of a production process.

The available nutrient mixtures generally contain amino acids, a carbohydrate source (usually glucose), inorganic salts, vitamins, some type of buffering system (usually CO_2, bicarbonate, and phosphate), fatty acids or more complex lipids, and other components that may vary from medium to medium (e.g., nucleosides, nucleotides, coenzymes, detergents). These components interact with each other so that the optimal concentration will depend to some extent on concentrations of other components in the mixture. Many producers of cell culture media and components have developed their own proprietary cell culture media for specific cell types used in production of recombinant products (e.g., CHO-S-SFM, #12050 or Hybridoma-SFM, #12045 Gibco/BRL, and products from other companies such as JRH, Hana Biologics, and Sigma). These media may offer a shortcut when proteins must be produced in short order; however, incomplete disclosure of the medium components can be a cause for concern, and costs are frequently higher than those of media formulated in-house.

Optimization of the medium starts with choosing the best existing commercial or proprietary nutrient mixture. Then, media components are made up individually and in groups (12,13), and individual media are prepared missing only one component. That component is then titrated using the desired end point of viable cell number or product titer, if a good high throughput assay is available. Figure 3 shows the response of CHO cells to three media components: inositol, KCl, and isoleucine. It is obvious that the omission of KCl or isoleucine results in cell death over the production period and consequently very low protein yields. However, the omission of inositol had, if anything, a positive effect. Conversely, increasing the inositol concentration 20-fold had little detrimental effect, whereas increasing isoleucine was beneficial and high levels of KCl were toxic. Using these data, one might then adjust the inositol concentration downward by 10-fold, the KCl downward by 2-fold, and the isoleucine upward. By using the data from all the titrations, one can then determine the optimal concentrations for each component and make up the best medium. This titration can then be performed again, using the best medium as a starting point. Each successive round of optimization would be expected to yield less benefit as the optimum is approached. A comparison of titers from a production cell line in the starting (control) and round-one optimized medium is shown in Figure 4. This one round of optimization more than doubled the titers during the 8-day production period. This benefit was achieved in large part by increasing the VCD in the culture. Note that significant changes in the culture process, such as changing cell lines, changing inoculum density, changing the supplements, or going from attached to suspension culture, might well require reoptimization of the medium. With experience, it is possible to identify a few medium components that are most important and check only these after the initial optimization. Conversely, once an optimal medium is designed for a given cell line (e.g., CHO), the cells can be grown in that medium during transfection and amplification in order to select for cells that grow optimally in the medium designed.

Growth Factor Additions

The nutrient mixtures discussed in the previous section are generally supplemented with either undefined mixtures, such as serum, that promote the growth of cells or defined mixtures of growth factors, transport proteins, and hormones. The choice of whether one will attempt to reduce or eliminate serum and which supplements to use is based on the characteristics of the product and the cell line as well as cost and safety considerations. Most established cell lines have been carried in serum for years and are therefore well adapted to growth in serum. It may be difficult to achieve equivalent growth rates or titers in a serum-free medium. Table 3 illustrates the effect of medium supplementation on tissue plasminogen activator (tPA) production by Bowes melanoma cells. Even though higher specific productivities can be reached in defined medium, when compared to serum-containing medium, the total yield is lower because of the much-reduced cell growth. Note also that some of the additions primarily affect cell number, whereas others (e.g., nerve growth factor) increase specific productivity with little effect on cell yield. The cells in Table 3 are not recombinant; however, similar effects can be seen using recombinant proteins and 293, CHO, or other host cell lines. In this case, a subset of factors will increase cell number (via decreased cell death or increased cell division), and another set of factors might increase titer via increases in specific productivity of the cells or increase protein stability. The former factors might be useful for all CHO-derived production lines, whereas the latter may be product specific.

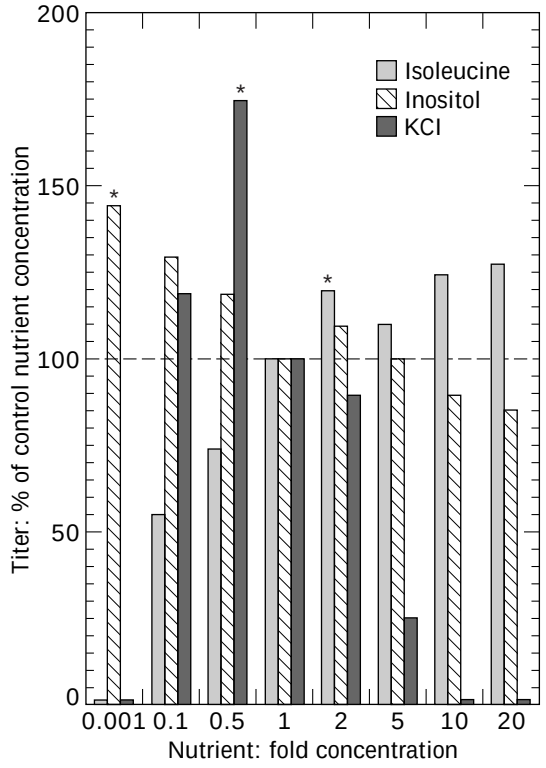


Figure 3. Titration of individual medium nutrients. CHO cells transfected to produce a recombinant protein were grown in production medium from which each of the indicated components had been omitted. These components were then titrated at the indicated concentrations. Cells were plated at standard inoculum and media collected to assay for product on day 6. All cultures were in 60-mm plates at 37 °C in triplicate. Values are expressed as percent of control where the control is the normal concentration (1×, *dotted line*) of each component in the medium. The ordinate values are the fold increase or decrease of the single component compared to control. The asterisks indicate the optimal concentration for each component that will then be chosen for use in the next round of medium optimization.

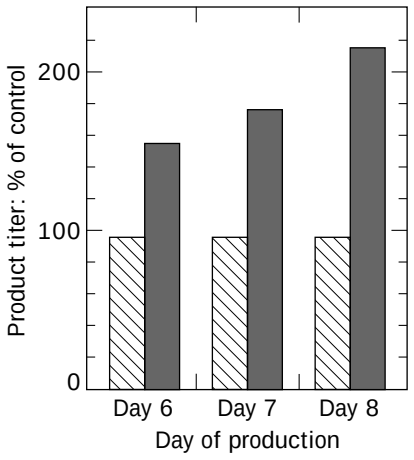


Figure 4. Comparison of titers obtained using the same cells before (control, *hatched bars*) and after (*filled bars*) optimization of the nutrient portion of the medium. These differences result entirely from changes in concentrations of the nutrients. Results are expressed as percentage of control where the control is the starting F12/DME medium.

Table 3. Cell Number, Titer, and Specific Productivity of Cells Producing tPA

		Plasminogen activator	
Supplement	Cell no. × 10^5	ng/10^5	ng/mL
Serum free	0.69	<30	Lts
+I	0.86	<30	Lts
+I+TF	1.08	<30	Lts
+I+TF+NGF	1.06	54	57
+I+TF+NGF+P	1.09	55	59
7.5% FBS	6.00	30	178

Note: Bowes melanoma cells were grown in F12/DME supplemented as indicated. Medium was collected after 5 days in culture. Lts, less than standard; I, insulin; TF, transferrin; NGF, nerve growth factor; P, progesterone; FBS, fetal bovine serum.

However, there are several reasons to eliminate serum, where possible, even if this results in a decrease in titer. Serum is expensive, contains high levels of many unknown proteins (about 40 mg/mL) that must then be purified away from the recombinant product, and contains enzymes and binding proteins that may affect product quality. For example, the production of tPA in the presence of serum will yield predominantly two-chain cleaved tPA rather than single-chain material. This is because the recombinant plasminogen activator produces plasmin from the plasminogen in the serum, which in turn cleaves the tPA. Thus, the tPA produced from the cells in Table 3 in the serum and serum-free conditions would have different characteristics. If single-chain material is desired, it is best to try and improve the cell yield by adding other supplements or adjusting the nutrient mixture.

One straightforward method for reducing contaminating proteins from serum is to simply reduce the serum concentration. This makes the medium less expensive in many cases, simplifies purification by eliminating serum proteins, and often improves product quality or cell growth. The general practice is to wean the cells by reducing the serum concentration over several passages (e.g., from 10% to 5%, 2%, 1%, 0.5%, 0.1%, etc.) allowing two or more passages at each concentration if a reduction in growth rate is observed. This method often works, but it is exceptionally time consuming. Eliminating serum altogether over one to two passages is often a reasonable solution for cells that do not have very fastidious growth requirements, although the resultant titers might not be optimal. A better solution is to derive a set of supplements that substitute for serum without any adaptation being required. In our experience, one can usually obtain titers in serum-free medium similar to, or even better than, those obtained in serum, although some thought must be taken as to the parental cell lines used for expression (see "The Cell Line").

With the advent of bovine spongiform encephalitis (BSE), it is likely that more processes will be driven to

become serum free for safety considerations. Because testing for BSE on each lot of serum is time consuming and expensive, elimination of serum from the process is an easy way to avoid this problem. However, the entire process must be carried out serum free if one is to avoid the chance of infection. This is a more stringent growth condition than growing cells for only one passage serum free and requires more complete supplementation.

The question of which supplements to add is again one that can only be answered in the context of the desired end point. It should be emphasized in this context that optimizing the nutrient portion of the medium can reduce growth factor (and serum [13]) requirements. Optimizing the nutrient portion of the medium is generally less expensive than adding supplements such as serum or growth factors. For production, minimizing expense and optimizing product quality may be more important than maximizing titer. It is the final yield of purified product as a function of fermentation plus purification cost that is the determining factor in deciding what the best medium will be. In an optimized serum-free culture process, the recombinant product may comprise 10 to 80% of the total protein in the harvest medium. This makes purification easier and less costly. In contrast, if one is growing up cells for the seed train, minimizing the doubling time to have cells ready for inoculation may be of more importance than the cost of the medium because cells are grown in smaller volumes. The quality of the protein produced is also not an issue in the early passages after thaw. However, it is generally best to not make too many major changes in media while bringing cells from the thawed vial to the production tank in order to avoid lag time as the cells adjust to new conditions.

Much has been written concerning optimizing the supplements for serum-free media. Our preference is to use small-scale culture experiments to obtain defined supplements that will provide cell numbers and product titers equivalent to or higher than those obtained in serum-supplemented medium. The elimination of serum necessitates the addition of hormones, growth factors, trace elements, and lipids (12,14–16). Because of the low concentrations of these components, the absence of binding and carrier proteins provided by serum, and the instability of the components in culture medium, these must be added to the media after filtration and just before use. The most commonly required additives are insulin (1 to 10 μg/mL), transferrin (1 to 100 μg/mL), and selenium (10 to 30 nM); however, some cell lines have an added requirement for lipids. One can then determine, based on experimental data, whether the increased product yield obtained by adding the most expensive components of the medium (for example, an epidermal growth factor requirement) are worth the expense. Using such data, one can then determine a minimum required medium that contains the most cost-effective supplements. A more complete description of deriving the appropriate hormone supplement in order to eliminate serum is given by Barnes and Sato (15) and Hewlett (17). For cell lines that are particularly fragile, Pluronic F-68 (18) or Pluronic F-127 can be added to the medium to prevent shear-associated cell lysis. Finally, if the product is sensitive to proteolysis, one might wish to add protease inhibitors, as discussed more fully later.

The empirical derivation of medium formulations optimizing both the nutrient portion of the medium and the supplements is time and labor consuming. It is seldom undertaken in a thorough fashion. One often-expressed wish is the desire to be able to predict what the optimal medium will be for a given cell line. Unfortunately, at this time, it is not possible to do this with any degree of assurance. We can make good guesses, largely based on experience gained by performing previous experiments on similar cells, but this will probably result in a good enough, not an optimal, medium. The only cases in which a full optimization of the medium is undoubtedly justified are (*1*) those research laboratories interested in understanding the nutritional and hormonal requirements of cells in vitro and (*2*) in the development of a commercial process where small gains in titer, viability, and stability would, over many years of production, be expected to repay the initial effort expended in the optimization. Most large-scale culture processes fall into this category.

OPTIMIZING SPECIFIC PRODUCTIVITY

Specific productivity is the amount of recombinant protein produced per cell per specified time period. Given that the cells are growing in an optimal medium, this will be a function of the expression system (promoter, enhancer, etc.) chosen and how well that particular system works in the host cell chosen, the integration site and orientation of the recombinant plasmid, the copy number of the gene to be expressed, and any specific or nonspecific inducers used. A discussion of the many expression systems available (19) is beyond the scope of this article. However, the method used to select and clone transfected cells and amplify gene copy number frequently impacts the performance of the cells in the production culture medium and will be discussed briefly.

The process of obtaining a stable cell line expressing a foreign protein involves a method of permeabilizing the cells to allow the plasmid containing the foreign DNA to enter the cell, a method of selecting for the minority of cells that have integrated the DNA into the nucleus, cloning to select a clonal cell line producing high levels of protein, and frequently several rounds of amplification of the DNA, each followed by cloning. This process allows ample opportunity for both genetic changes in the host line and the selection of a subpopulation from the initial population. Usually, the primary, if not the sole, concern is to maintain the gene containing the recombinant protein in the cells. However, the phenotype of a good production cell line also includes the ability to grow rapidly and to high saturation density while maintaining a high viability. If one has made the effort to optimize the medium for the growth of the parental cell line, it is desirable to maintain the ability to grow well in this medium, including serum-free and suspension growth. To maintain the desired phenotype, selective pressure should be maintained for these growth characteristics during the process of transfection, amplification, and cloning. If the cells are grown in the process medium whenever possible, they will be much easier to handle when they reach the production setting. To the ex-

tent that this is not possible, the production cell line selected might benefit from a period of adaptation to the process medium and growth conditions before being banked.

OPTIMIZING PRODUCT STABILITY

The production medium is designed to provide the maximum number of cells secreting protein over the maximum amount of time. The longer the productive run time, the longer the product can accumulate in the medium and the higher the resultant titer. However, this medium, along with the metabolites and proteins being secreted into it by the cells, is also the storage medium for the proteins being produced and secreted. An environment that is 37 °C, neutral pH, fully oxygenated and stirred and where proteases and other enzymes are separated from the proteins by, at best, a layer or two of cell membrane would hardly be considered optimal for protein storage by any biochemist. This is, however, what recombinant proteins produced in the first days of the production run have to withstand for the duration of the run. Therefore, some thought should be given to protein stability in the production medium. This will be in large part dictated by the characteristics of the protein being produced. Figure 5 illustrates the accumulation of titer during a production run with two different proteins. One is stable and accumulates linearly during the run; the other is easily proteolyzed and begins to degrade faster than it is produced when the cells in the culture begin to die. Continuation of the culture at this point leads to loss of product. The most important change in this process would be to optimize the medium components to prolong cell viability. One might then adjust the inoculum density so that the number of viable cell days is optimized. Too high an initial inoculum can result in early cell lysis and degraded product, whereas too low an inoculum results in too long a production period to reach the desired titer, resulting in overly long product residence time in the medium. One might also adjust the temperature to minimize cell death (11), shorten the run (including using processes that shorten the product residence time, such as perfusion), or add protease inhibitors to the medium. Protease inhibitors can be nonspecific (e.g., serum albumin) or specific (e.g., aproteinin or small molecular weight inhibitors). The small molecular weight inhibitors are frequently toxic to cells in culture, whereas inhibitors such as aproteinin are well tolerated but very expensive for production use.

THE MARRIAGE OF THE MEDIUM AND THE PROCESS

As discussed in the introduction, medium formulation and optimization should not proceed in a vacuum but should be an integral part of the overall process design. From the initial choice of cell line through to the final purification, there should be consideration of the entire final process that will result.

The Cell Line

The production cell line should be chosen not only for its protein secretion rate measured under laboratory conditions but also for its ability to perform well in the special conditions that occur in large-scale culture (Table 2). The cells may be adapted to these conditions after transfection, as discussed earlier, or the parental cell line may be engineered to optimize its performance as a production cell line before transfection. Cells can thus be adapted to grow in suspension or serum-free medium and then be banked for use as parental cell lines. If they are maintained in these conditions during transfection and selection, then the recombinant protein producing the cell lines will have these important characteristics. Even if they cannot be handled in suspension or serum-free conditions at every step of the

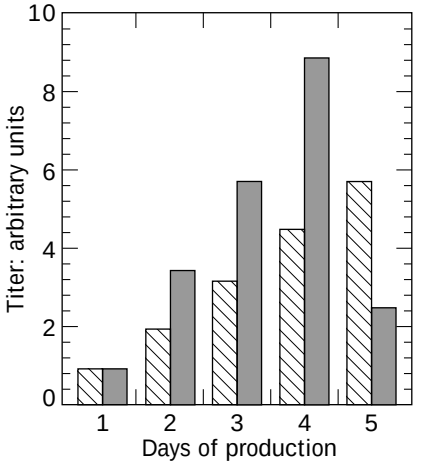


Figure 5. Titer as a function of run length. The figure shows the cumulative titers during two runs of recombinant cells, each expressing a different protein (*open bars* and *hatched bars*). The assays used only measure intact protein. Data are expressed as fold increase over control, where the control is the titer on day 1 of the culture.

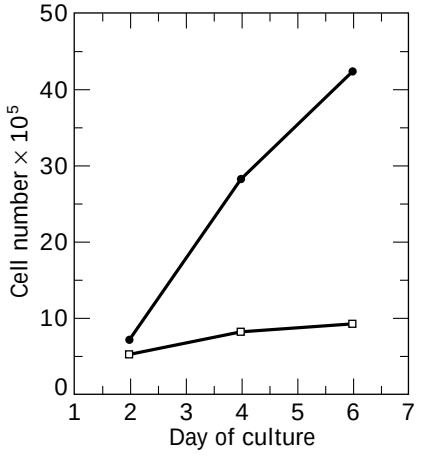


Figure 6. Insulin-independent growth of CHO cells transfected with the proinsulin gene. These cells can grow as well in the absence of insulin as their parent line can grow in medium supplemented with 10 μg/mL insulin. CHO-parent line (*squares*) and dp12-proinsulin transfected clone (*circles*) were grown in the absence of insulin.

transfection process, they will be more likely to retain the desired characteristics.

Additionally, one can genetically engineer the parental cell line to have traits that are optimal for a production cell line. One example of this is to transfect the production cell line with a required growth factor, allowing autocrine growth at large scale without exogenous addition of the factor. Figure 6 demonstrates that CHO cells that have been transfected with the proinsulin gene have become insulin independent for growth. This will, of course, alter the composition of the medium used to grow these cells.

Equipment

It is obvious that the optimal medium for a process will depend, in part, on the equipment used and the configuration of the process. Clearly, suspension cultures and attached cells might require different media. The optimal nutrient mixtures for batch cultures and for perfusion cultures would be quite different. The inoculum density and amount of growth that will be required of the cells at each stage of the scale-up process will depend on the tank sizes used to inoculate each stage. The equipment to be used for the final process should be considered from the first stages of medium optimization so that one is sure that the medium designed does, in fact, support the cell growth and viability required in the process.

Downstream Purification

Because the final goal of the process is to obtain a highly purified active protein, the downstream purification methods to be used should be considered from the beginning. Some protein characterization should be performed throughout the medium development process, in addition to a simple measurement of titer. Medium changes can alter protein characteristics as well as titers. It is therefore important to make certain that the desired product characteristics are maintained when medium changes are made to improve titer.

Finally, everything that is put into the medium by the formulation or by the cells in the culture must be removed from the final protein product. Wherever possible, it is always best to use supplements that will not raise safety considerations. Additionally, it does not make sense to add supplements to increase titer if the resulting product will be lost because of the extra purification steps that must be added. There is not as much room to alter pH, salt concentration, and so forth in a cell culture medium as there is in protein purification processes, but some understanding of the constraints and problems of both ends of the process will help make for a smooth transition and a reproducible process.

In summary, the time and effort required to optimize all components of the production medium will pay off many times over in a mammalian cell culture process. Optimal medium not only optimizes protein titers but can lead to improvements in protein quality and stability, more stable cell lines, more reproducibility from run to run, and improved use of facilities. To obtain maximal benefit from medium design it must be done in concert with the transfection and selection of the production cell line, the configuration of the manufacturing plant, and the design of the downstream purification process. All these factors working together have made large-scale mammalian cell culture an economically important part of the modern biotech industry.

ACKNOWLEDGMENTS

We would like to acknowledge the invaluable interactions with other members of Genentech's process sciences division, especially the cell culture R&D department, the cell genetics department, and assay services group. All these people contributed to the experiments and experiences described in this article.

BIBLIOGRAPHY

1. A. Moore, C.J. Donahue, J. Hooley, D.L. Stocks, K.D. Bauer, and J.P. Mather, *Cytotechnology* **17**, 1–11 (1995).
2. F. Franek, T. Vomastek, and J. Doinikova, *Cytotechnology* **9**, 117–123 (1992).
3. M.J. Arends and A.H. Wyllie, *Int. Rev. Exp. Pathol.* **32**, 223–254 (1991).
4. J.V. Moore, in C.S. Potten ed., *Perspectives in Mammalian Cell Death*. Oxford Univ. Press, Oxford, U.K., 1987, pp. 295–325.
5. W.G. Telford, L.E. King, and P.J. Fraker, *Cytometry* **13**, 137–143 (1992).
6. J.L. Goergen, A. Marc, and J.M. Engasser, *Cytotechnology* **11**, 189–195 (1993).
7. C. Dive and A.H. Wyllie, in J.A. Hickman and T. Tritton eds., *Frontiers in Pharmacology: Cancer Chemotherapy*, Oxford Univ. Press, Oxford, U.K., 1993, pp. 21–56.
8. R. Duke, R. Chervenak, and J. Cohen, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 6361–6365 (1983).
9. W. Gorczyca, S. Bruno, R.J. Darzynkiewicz, J. Gong, and Z. Darzynkiewicz, *Int. J. Oncol.* **1**, 639–648 (1992).
10. R. Pitti, S. Marsters, S. Ruppert, C. Donahue, A. Moore, and A. Ashkenazi, *J. Biol. Chem.* **271**, 12687–12690 (1996).
11. A. Moore, J. Mercer, G. Dutina, C. Donahue, K. Bauer, J. Mather, T. Etcheverry, and T. Ryll, *Cytotechnology* **27**, 47–54 (1997).
12. R.G. Ham and W.L. McKeehan, *Methods Enzymol.* **58**, 44–93 (1979).
13. W.L. McKeehan and R.G. Ham, *In Vitro* **13**, 399–416 (1979).
14. J. Bottenstein, I. Hayashi, S.H. Hutchings, H. Masul, J.P. Mather, D.B. McClure, S. Okasa, A. Rizzino, G.H. Sato, G. Serroro, R. Wolfe, and R. Wu, *Methods Enzymol.* **58**, 94–109 (1979).
15. D. Barnes and G.H. Sato, *Anal. Biochem.* **102**, 255–270 (1980).
16. J.P. Mather ed., *Mammalian Cell Culture: The Use of Serum Free and Hormone Supplemented Media*, Plenum Press, New York, 1984.
17. G. Hewlett, *Cytotechnology* **5**, 3–14 (1991).
18. J.E. Swim and R.F. Parker, *Proc. Sci. Exp. Biol. Med.* **103**, 252–254 (1960).
19. D.V. Goeddel, *Methods Enzymol.* **185**, 567–577 (1991).

See also ADSORPTION, PROTEIN, BATCH; ANTIBODY PURIFICATION; MEDIUM FORMULATION AND DESIGN E. COLI AND BACILLUS SPP.; FREEZE DRYING, PHARMACEUTICALS; ROLLER BOTTLE CULTURE, MIXING.

CULTURE PRESERVATION, BACTERIA, FUNGI, YEAST, AND CELL LINES

ROBERT L. GHERNA
American Type Culture Collection
Rockville, Maryland

KEY WORDS

Commercial freeze-drying
Component freeze-drying
Control Freeze
Dimethyl sulfoxide (DMSO)
Freeze-drying
Glycerol
Lyophilization
Manifold freeze-drying
Sublimation
Ultrafreezing
Manifold freeze-drying

OUTLINE

INTRODUCTION

The ubiquitous presence of microorganisms, as well as their metabolic diversity, has made them important factors in industrial applications and product development. Bacteria, fungi, and yeasts have been used for the production of food, medicinals, solvents, enzymes, and numerous other products. They have also been used as standards for bioassay and bioremediation applications. In addition, advances in bioengineering have enabled the incorporation of genes into microorganisms for the production of mammalian proteins and peptides.

During the past 20 years, cell line technology grew dramatically to enable the isolation and maintenance of a wide variety of tissue lines from any species of interest. The initial use of cell lines for vaccine production was expanded by developments in cell hybridization, transduction and transfection, and genetic manipulation to include the production of monoclonal antibodies, cytokines, and other pharmaceutical products. Lines that have been genetically engineered have been successfully expanded and grown in fermenters for large-scale production.

The development of microbial and cell cultures represents an enormous investment in time and money and is an important asset to industry. These cultures must be protected from accidental loss by means of standardized preservation techniques. The primary aim of culture preservation is to maintain the culture such that it is as close as possible to the original strain or line exhibiting the desired phenotypic or genetic traits. Often, industrial cultures are not analogous to taxonomic strains in that they have been modified to have special properties, either rapid growth, more active metabolic rates, or other features.

Many methods have been used to preserve bacteria and fungi, but not all species respond in a similar way to a given method. Unfortunately there is no universal method that can be applied to all microorganisms. It should be emphasized, however, that the success or failure of any preservation technique also depends on the use of the proper growth medium and cultivation procedures and on the age of the culture at the time of preservation. This is particularly true when working with microorganisms that contain plasmids or recombinant DNA, or that exhibit growth phases such as morphogenesis or spore formation.

The following provides practical information for the maintenance and preservation of bacteria, fungi, yeast, and cell lines. It should be noted, however, that not all microorganisms will be preserved by these methods, and some experimentation with cryoprotective agents and other factors may be necessary for success. Additional information on bacterial, fungal, and cell-line preservation may be found elsewhere (1–3).

CULTURE AND PRESERVATION OF BACTERIA

Short-Term Preservation

Subculture. Most bacteria can be preserved by periodic serial transfer to fresh medium. The period between transfers varies with the organism, medium employed, and the external conditions. In general, minimal media are preferred for subculturing because they lower the metabolic rate of the organism and thus prolong the interval between transfers. In rich media, growth is often much faster, with the accumulation of metabolic products. These compounds may alter environmental conditions such as pH and gas phase and shorten the interval between transfers. Some bacteria such as pathogens need complex media (4) for growth, or their retention of specific physiological properties require the addition of complex compounds to the medium. The subculturing interval for these organisms must be determined by experience.

When preserving bacteria by serial transfer, it is important to use containers that can be securely sealed. Glass test tubes or 6-mL bottles with rubber-lined screw caps are widely used. The bacteria can be grown with slightly loose caps until adequate growth is obtained and then tighten to prevent dehydration. Solid media are preferred to broth because contaminants can be readily seen. Duplicate tubes should be maintained as a precaution against loss, at least until the subculturing interval is determined. The cultures should be examined for purity after each transfer, and abbreviated characterization tests must be run periodically

to ensure retention of desired traits. Storage in a refrigerator is the preferred method for subcultures. Many bacteria can be kept for 3 to 5 months between transfers if proper precautions are taken to avoid dehydration of the medium.

It should be noted, however, that the maintenance of cultures by serial transfer is precarious, time-consuming, and expensive. The frequent transfer of cultures can result in contamination, mislabeling, and the selection of strain variants.

Other methods involve storing cultures at refrigeration temperatures, drying in sterile soil, glass beads, filter paper, and other substrates. Detailed descriptions of these methods can be found in Refs. 5–7.

Short-term maintenance and preservation procedures are not recommended for industrial strains because a number of problems can occur using these methods. The main hazards are contamination, population changes through selection resulting in the loss of key genetic and physiological properties, and the expense of maintaining cultures under these conditions.

Ordinary Freezing. The use of the freezer compartment of a refrigerator or an ordinary freezer with a temperature range of 0 to −20 °C to preserve bacteria is not recommended because the freezing process (see "Long-Term Methods", Deep Freeze) damages the cells. In addition, most refrigerators that have a freezer compartment in the range of 0 to −20 °C are frost free. The frost-free state is achieved by an alternating warming and cooling cycle, which can result in further damage to the cells and a loss of viability.

Long-Term Methods

Freeze-Drying, or Lyophilization. Freeze-drying, or lyophilization, is one of the most economical and widely used methods for long-term preservation of bacterial cultures and other microorganisms. Many metabolically diverse bacterial species have been successfully preserved by this technique and have remained viable and unaltered for more than 60 years.

Freeze-drying is a complex process that involves the removal of water and other solvents from a frozen product by sublimation (8,9). Sublimation occurs when a frozen liquid goes directly to a gaseous phase without passing through the liquid phase. There are three stages in the freeze-drying process: *(1)* prefreezing of the suspension to ensure a solid frozen starting material, *(2)* primary drying during which most of the water is removed through sublimation, and *(3)* secondary drying to remove the bound water.

The rate of cooling during the prefreezing step and the final temperature of the frozen material can affect the freeze-drying process. The removal of water during the primary drying stage requires an environment that allows the water molecules to migrate from the frozen material. This is achieved by using a vacuum pump. A large-capacity pump must be used to prevent the water vapor load from compromising the efficiency of the system. The procedure also requires a moisture trap or condenser to collect and remove the water molecules before they enter the vacuum system. The condenser temperature must be lower than the frozen material for sublimation to occur. The sublimation rate is dependent on the differential vapor pressure between the frozen material and the condenser. Because the vapor pressure varies with the temperature, heat can be carefully applied to the frozen cell suspension to increase the vapor pressure and promote a faster primary drying rate. Precautions must be taken during the application of heat to prevent the frozen cell suspension from thawing. Most commercial instruments have sophisticated controllers that allow the input of heat to the frozen material to achieve maximum drying rates without damaging the product.

When the primary drying is complete, some bound water remains. This water cannot be eliminated by sublimation and requires the application of external heat to the material under low pressure and low condenser temperature. The moisture remaining in freezed-dried cells is termed residual moisture. A residual moisture content of 1–3% seems to be required for the long-term shelf life of bacteria. For an in-depth review of the complex freeze-drying process, the reader is referred to several excellent reviews (10–12).

Centrifugal freeze-drying and prefreezing are two of the most commonly used methods of lyophilization of bacteria. In centrifugal freeze-drying, the suspensions are initially frozen by evaporative freezing under a vacuum while centrifuging to prevent frothing due to removal of dissolved gases. After primary freeze-drying, the vials are constricted using a narrow flame and then placed on a manifold for secondary freeze-drying. A detailed discussion of centrifugal freeze-drying can be found in articles by Rudge (13) and Lapage (14). The prefreezing method employs a controlled rate of cooling to freeze the cell suspension and then a primary drying phase, as discussed earlier.

The simplest system consists of a vacuum pump capable of an ultimate pressure of less than 10 μmHg (a pump rated at 35–50 L/min is usually adequate), a small stainless steel condenser for cooling with dry ice, a thermocouple vacuum gauge to monitor the vacuum system, and heavy-walled vacuum/pressure tubing to connect the components. The connection of these modules into a single system is often termed a component freeze-dryer. The main components are the vacuum pump, thermocouple vacuum gauge, and a condenser that can be attached to several manifolds, to which glass vials can be attached by means of vacuum-tubing nipples, or the manifolds can be connected to a stainless steel pan containing the vials. The manifold system is relatively simple, and it has been used successfully by many laboratories to preserve cultures. A small amount of cell suspension (0.2 mL) containing a cryoprotectant is dispensed into sterile ampules, which are plugged with sterile cotton. A 1-inch piece of nonpowdered amber latex IV tubing is attached to the rim of each ampule (by using a tube stretcher) and then to the manifold. The ampules are then immersed in an ethylene glycol (50%) dry-ice bath at −40 °C to freeze the cell suspension. The freeze-drying cycle is initiated by starting the vacuum pump and allowing the temperature of the bath to rise of its own accord to ambient temperature. In general, runs

can be started early in the afternoon and allowed to proceed overnight. At the end of the drying cycle, the ampules are sealed using a double-flame air/gas torch and stored at 2–8 °C. For specific details on the equipment and methods see articles by Gherna (15) and Simione and Brown (16). Details of other procedures and equipment are given by Haynes et al. (17).

The American Type Culture Collection (ATCC) has successfully lyophilized many of the bacteria described in the literature. It currently uses three methods:

1. *Component freeze-dryer*. Samples are freeze-dried in cotton-plugged inner vials, which are then sealed in outer vials under vacuum.
2. *Commercial freeze-dryer*. Cultures are freezed-dried in the double-vial system in a commercial freeze-dryer.
3. *Preceptrol cultures*. Cultures are lyophilized in a commercial freeze-dryer using glass serum vials that are sealed with a rubber stopper and metal cap.

It should be noted that after the drying cycle is complete in the commercial freeze-dryer and Preceptrol methods, the vacuum in the drying chamber is replaced with cooled nitrogen gas (2–8 °C) that has been passed through a 0.2-μm Pall filter. The condenser temperature must not be allowed to rise above −50 °C. The nitrogen gas may have to be passed through a copper coil immersed in liquid nitrogen.

Culture Preparation. Successful freeze-drying depends on using healthy cells grown under optimum conditions on the medium of choice for each strain, which ensures the retention of the desired features of the bacterium. Bacteria can be grown on agar or in shaken or static broth cultures. The cells are harvested at maximum stability and viability, in the late logarithmic or early stationary phase, and suspended in a medium. The cell suspension should contain at least 10^7–10^{10} cells/mL in order to achieve optimal results. Anaerobic cultures must be grown, harvested, and dispensed under anaerobic conditions (18).

The viability of freeze-dried cultures is greatly influenced by the suspending medium. The choice of suspending medium for freeze-drying depends on the type of bacteria and on the method used.

The ATCC uses two suspending media for the commercial freeze-drying method:

1. Reagent 18
   0.75 g trypticase soy broth
   10.0 g sucrose
   5.0 g bovine serum albumin fraction V
   100.0 mL distilled water
   Filter sterilized through a 0.2-μm filter
2. Reagent 20
   10.0 g bovine serum albumin fraction V
   20.0 g sucrose
   100 mL distilled water
   Filter sterilized through a 0.20-μm filter

Reagent 18 is more commonly used for freeze-drying bacteria and fungi. Reagent 20 is employed for bacteria that are adversely affected by trypticase soy broth and is added to an equal volume of cell suspension in the broth.

In the component freeze-drying method (double-vial), the additive is a 20% (wt/vol) sterile solution of skim milk, prepared by autoclaving a 20% skim milk solution at 116 °C for 20 min in 10-mL tubes. For cultures grown on agar surfaces, the cells are washed off with the 20% skim milk solution. Broth cultures are centrifuged, and the cell pellet is resuspended with the sterile skim milk to give a suspension of 10^8 cells/mL. Skim milk should not be used for bacteria that are inhibited by milk, instead use a 24% (wt/vol) sterile sucrose solution diluted equally with growth medium to yield a 12% (wt/vol) sucrose solution (final concentration) for the single-vial manifold or commercial freeze-dryer method. As soon as the cell suspension is prepared, it should be dispensed into the ampules. The interval between dispensing and freeze-drying should be kept to a minimum to avoid possible alteration of the culture. Other investigators have used 10% (wt/vol) dextran, horse serum, inositol, raffinose, trehalose, and other cryoprotective chemicals in the suspending medium (19–21).

Most bacteria can be lyophilized using these methods; for those cultures that require special conditions, the reader is referred to Ref. 16. The same methods have been employed successfully to freeze-dry most bacteriophages; exceptions are stored in liquid nitrogen. In general, bacteriophages can be grown in a soft agar layer or in broth. The phages are harvested by scraping off the soft agar with a sterile glass rod or rubber policeman, macerating and dispensing into a sterile centrifuge tube, and centrifuging at low speed to sediment the agar and most of the unlysed bacteria. Broth-gown cultures are centrifuged at low speed to remove the unlysed bacteria. The supernatant is then filtered through a 0.45-μm and then through a 0.2-μm Millipore filter. The filtrate should be titrated to determine the concentration of phages; a titer of 10^8 pfu/mL is desirable. The phage can be freeze-dried using 20% skim milk and the component freeze-dryer method.

Freeze-dried bacteria and bacteriophages can be stored at 2–8 °C. A longer shelf life has been obtained when cultures are stored at −30 or −70 °C in a mechanical freezer. Freezed-dried cultures should not be stored in a freezer compartment of a frost-free refrigerator because of the alternate heating and freezing cycles, nor should they be stored at room temperature.

Although lyophilization has facilitated the long-term preservation of bacteria, viability checks must be done before and after lyophilization to determine the effectiveness of the process. Quality control procedures must be in place to check for purity, retention of essential characteristics and periodic viability to determine the shelf life of the culture.

Deep Freeze. The long-term preservation of bacterial species not amenable to freeze-drying has been achieved through storage in the frozen state at a temperature of −70 °C or, more advisable, at the temperature of liquid nitrogen (−196 °C) or liquid vapor phase (−150 °C).

The freezing process results in several events that can be destructive to the cells. As the bacterial suspension begins to cool, the liquid water turns to ice at 0 °C (the exact temperature depending on the nature of the bacterial suspension). Ice formation occurs first in the external aqueous environment, resulting in an increased concentration of solutes outside the cells. The differential osmotic pressure causes water to migrate from the cells. The rate and extent of the water loss are dependent on the rate of cooling and the permeability of the cells. Removal of too much water leads to a concentration of solutes within the cells that may be harmful. If too much water is left in the cells, internal ice crystals may form, which would result in intracellular damage. It is important to maintain the critical balance between the two events by controlling the rate of cooling while freezing the cells, and warming as rapidly as possible when the cells are thawed. Chemical compounds such as glycerol and dimethyl sulfoxide (DMSO) can be added to the cells to minimize the damage caused by freezing. These compounds, labeled cryoprotective agents (see "Cryoprotectants"), exhibit certain properties that are essential for good cryoprotection (22,23). They must be nontoxic to the cells, show good permeability, and bind either the electrolytes that accumulate or the water molecule to delay freezing.

The physiological condition of the cells plays an important role in survival of freezing. In general, actively growing cultures harvested at the mid to late logarithmic phase of growth will survive the freezing process better than those harvested at an earlier or later phase. The bacteria should be grown under optimum conditions in the appropriate medium that best retains the prominent characteristics of the strain. Bacteria can be grown on broth that is static or shaken, or on agar. Shaken cultures generally reach optimal cell density faster than agar-grown cultures. Some strict anaerobes require prereduced media and anaerobic procedures for harvesting and dispensing.

Cells grown in broth cultures are harvested by aseptic centrifugation, and the resultant pellet is resuspended in sterile broth containing 30–50 % (vol/vol) sterile glycerol. Bacteria grown on agar are harvested by aseptically washing the growth with sterile broth containing the cryoprotectant. Dispense 0.4 ml of the cell suspension, containing at least 10^8 cell/mL into each sterile, prelabeled vial. Plastic presterilized screw-capped vials (such as Nunc) are recommended for bacteria. Place the filled vials into a mechanical freezer set at −70 °C. The cells are recovered by rapidly thawing the frozen cell suspensions in a 37 °C water bath.

A wide variety of bacterial have been successfully preserved at −70 °C using glass beads and 15 % glycerol as the cryoprotectant (24). The glass beads are prepared as described in Ref. 15, and 30 beads are placed in 2-mL glass screw-cap vials and sterilized for 15 min at 121 °C. The vials must be prelabeled with ink that will not come off at ultralow temperatures. Cell suspensions are prepared containing approximately 10^8 organisms/mL in sterile 15% glycerol, and 0.5 mL is dispensed into each vial. The vials are shaken gently to ensure that all the beads are wetted with the bacterial suspension, and then the excess liquid is aseptically removed with a sterile pasteur pipette. The vials are placed into a mechanical freezer at −70 °C. The cells are recovered by aseptically removing a bead from the vial with a sterile spatula or forceps and placing the bead in sterile broth. It should be noted, however, that the contents of the vial should not be allowed to thaw; this can be prevented by placing the vial in crushed dry ice and removing the bead after ensuring the beads will remain frozen. This technique has been used to preserve bacteria obtained from extreme environments such as hypersaline pools.

Precautions must be taken against electrical shutdowns and compressor malfunctions. Adequate back-up systems (such as alarms, back-up freezers, or an electrical generator) will help to prevent the loss of a valuable collection. However, additional alarm systems are recommended, such as the sound/off power-temperature monitor (Arthur H. Thomas Co.), which has a temperature range of −75 to + 200 °C. The unit is battery powered and can be mounted on a wall.

Ultrafreezing, or Cryopreservation. Although the long-term maintenance of bacteria in liquid nitrogen has been considered expensive, the successful preservation of physiologically diverse bacteria by this method, along with minimal handling and lower labor cost, makes this method feasible to use. The ATCC has successfully preserved many bacteria and bacteriophages, including fastidious ones, in liquid nitrogen for over 36 years without the loss of phenotypic properties.

A large variety of liquid-nitrogen refrigerators are now commercially available, with a wide assortment of features and storage capacities. The sizes range from 10 to 1,000 L, allowing storage of 300–40,000 ampules.

The physiological condition of the cultures plays an important role in the survival of bacteria under liquid-nitrogen freezing. Bacteria are grown in the appropriate medium and harvested at mid-late logarithmic phase of growth (25). Cultures grown in broth are harvested by aseptic centrifugation, and the resultant pellet is resuspended with sterile fresh medium containing either glycerol or DMSO. For agar-grown cells, the growth is washed off from the agar surface with sterile broth containing the appropriate cryoprotectant. Cryoprotective agents must be added to the suspending medium to protect cells from freeze damage. Cryoprotectants fall into two main classes: compounds such as glycerol and DMSO, which are permeable and appear to provide both intracellular and extracellular protection against freezing; and agents such as dextran, glucose, lactose, mannitol, polyglycol, polyvinylpyrrolidone, and sucrose, which seem to exert their protective effect external to the cell membrane. The first class of cryoprotectants has proven to be the most effective, and in general, glycerol and DMSO seem to be equally effective in protecting a wide range of bacteria. It is advisable to conduct a tolerance test on new species to ascertain whether the cryoprotectant is toxic or beneficial.

DMSO and glycerol are routinely used at concentrations of 5 % (vol/vol) and 10% (vol/vol), respectively, in the proper suspending medium. Glycerol is normally prepared as a double-strength (20 %, vol/vol) solution and then mixed with an equal amount of the cell suspension. Glycerol is

sterilized by autoclaving at 121 °C for 15 min and stored in 6-mL volumes at 2–8 °C. DMSO is filter sterilized with a 0.22-μm-pore-size Teflon (polytetrafluoroethylene) membrane (Gelman or Millipore) that has been prewashed with methanol and DMSO, collected in 10 to 15-mL quantities in sterile test tubes, and stored in the frozen state at 5 °C (DMSO freezes at 18 °C) and protected from light. An opened bottle of DMSO should not be used for more than 1 month because of the accumulation of oxidative breakdown products.

Cell suspensions containing at least 10^8 cells/mL are dispensed into prelabeled sterile vials. There are a variety of vials available commercially, ranging from prescored glass to sterile plastic. With glass vials, care must be taken when sealing and storing because improperly sealed glass vials can explode when retrieved from the liquid nitrogen as a result of the rapid expansion of the liquid nitrogen that enters the vials through microscopic holes. Plastic screw-cap vials that are not properly sealed can fill with liquid nitrogen if stored in liquid phase; retrieval of the vials to a warmer temperature can result in liquid nitrogen expansion, causing the contents of the plastic vials to spray into the laboratory environment. Plastic screw-cap vials should always be stored in the vapor phase. Protective gloves and face shields should be worn when handling frozen liquid-nitrogen vials.

The rate of cooling is important in liquid-nitrogen storage. In general, the best results for preservation of bacteria have been achieved with slow cooling. This step can be controlled using programmable controllers such as the Linde BF 3-2 (Union Carbide Corp.), which has an adjustable cooling rate. The cells are frozen at a controlled rate of 1–3 °C/min to −40 °C and then at a more rapid rate of 10 °C/min to −90 °C. After this temperature is reached, the vials are transferred to a liquid-nitrogen tank and stored in the liquid phase at −196 °C or above in the vapor phase at −150 °C.

If a programmable freezer is not available, slow cooling can be obtained by placing the filled glass or plastic vials in a stainless steel pan at the bottom of a mechanical freezer at −60 °C for 1 h and then plunging the vials into a liquid-nitrogen bath for 5 min. The rate of cooling to −60 °C is approximately 1.5 °C/min. The frozen vials can then be stored in the liquid-nitrogen tank in the liquid or vapor phase.

It has been reported that the best recovery of cultures from liquid-nitrogen storage is obtained by rapid thawing (26,27). This has also been the experience with cultures frozen at the ATCC. The events during thawing are complex, and slow warming rates can lead to ice crystal formation; a rapid warming rate of approximately 100 °C/min will minimize the damage due to ice formation. A rapid rate of warming can be achieved by quickly immersing the frozen vials in a 37 °C water bath with moderate agitation until all the ice melts. This usually takes about 50 s for glass vials and 90 s for polypropylene ones.

Cryopreservation of strict anaerobic bacteria must be performed under anaerobic conditions. Technical details of the procedures involved have been described by Hill et al. (28). The reader should consult various sections of *Bergey's Manual of Systematic Bacteriology*, volumes 1 through 4 (29–32), for information on the maintenance and preservation of representative bacterial genera.

The successful cryopreservation of plasmid-containing bacteria depends on the stability of their plasmids (33). These organisms are less genetically stable and more likely to lose their foreign genetic material (34). In general, bacteria containing unstable plasmids should be grown on antibiotic-containing liquid medium at 30 °C, frozen with 10 % (vol/vol) glycerol, and stored in the vapor over liquid nitrogen.

CULTURE AND PRESERVATION OF FUNGI AND YEAST

Short-Term Preservation

Subculture. Many filamentous fungi can be successfully maintained by serial transfer on suitable media. It is important, however, to use optimal media and growth conditions to ensure healthy cells that retain their morphological and physiological characteristics. A list of suitable media and incubation temperatures for a variety of fungi is provided by Onions and Pitt (35). In general, a medium that promotes good sporulation is considered to be the most desirable. When possible, it is best to cultivate the fungi on agar slants in test tubes or culture bottles. Transfers should be made only from the growing edges of the culture, taking precautions to ensure that contamination is prevented. Some fungi degenerate when maintained on the same medium for extended periods, thus media should be alternated from time to time. A minimal medium such as potato/carrot agar can often be used to induce sporulation in a culture that has deteriorated (36). Also, the continual transfer of spores from some aging fungal cultures such as *Mucorales* can result in a deterioration of the culture; the transfer of mycelium and spores seems to minimize this problem.

The interval between transfers varies from fungus to fungus; some require transfer every 2–4 weeks, the majority every 2–4 months, while others may survive for 12 months without transfer. Cultures grown on agar slants can be stored at room temperature. Storage in the refrigerator or cold room at 5–8 °C can extend the transfer period to 4–6 months. However, there are some fungi, such as the thermophiles, that are sensitive to storage at these temperatures.

The maintenance of fungi by serial transfer has many disadvantages. The most notable are contamination, selection of variants, mislabeling of cultures, infestation with mites, and labor intensity.

Yeasts have been successfully maintained by serial transfer on either solid or liquid media. In general, many yeasts will survive for longer periods when grown on solid media rather than in broth, especially nonfermentative yeast. Fresh medium is inoculated with actively growing cells and incubated at the appropriate temperature. After sufficient growth occurs, the cells are stored at refrigerated temperature (+4 °C). The majority of yeast species will remain viable for a least 6 months. The interval between transfers must be determined by experimentation. Serial transfer is not recommended for extended maintenance because genetic drift has been observed. In addition, asco-

sporogenous strains may sporulate on agar slant media stored for prolonged periods, resulting in strain variability.

Many fungi and yeast can be preserved for years by drying on a suitable menstruum such as soil, silica gel, and filter paper (37), and drying under vacuum from the liquid state (L-drying) (38).

Silica Gel. Sterile anhydrous silica gel (6–22 mesh, nonindicator) is prepared by dispensing the gel into growth bottles so that they are one-quarter filled, and sterilizing in an oven at 180 °C for 2–3 h. The bottles containing the sterile gel are cooled by placing in a pan containing ethylene gycol and dry ice to the depth of the gel layer or by placing the bottles in water in a mechanical freezer at a temperature of −17 to −24 °C. Fungal spores are suspended in a sterile 5% skim milk solution that has been cooled to 4 °C. The spore suspension is added to the cooled silica gel until three-quarters of the gel has been moistened. It is important to avoid saturating the gel. The growth bottles are left in the ice bath for approximately 20–30 min and then agitated to ensure thorough dispersion of the suspension. The bottles are held at 25 °C for approximately 1–2 weeks until the crystals easily separate when shaken. Inoculated bottles are sealed with screw caps and stored over indicator silica gel in an airtight container at 4 °C. To recover the fungal spores, a few crystals are sprinkled on an appropriate medium and incubated. This method has been used successfully for a variety of fungi (39). This method has produced variable results with yeast; some are not viable after 3 months of storage, whereas other survive after 2–5 years (40).

Soil. A variety of fungi have been maintained in soil (41). The method involves the inoculation of approximately 1 mL of a spore suspension into soil that has been sterilized by autoclaving twice, 24 h apart, at 121 °C for 15 min. The soil is incubated at room temperature for 5–10 days and then stored in a refrigerator at 4–8 °C. Cultures are recovered by inoculating a suitable medium with a few grains of soil.

Paper. Yeasts have been preserved by drying on paper discs or squares and storage over the desicant silica gel (42). Whatman No. 4 filter paper is cut into small squares or discs about 10 mm across and placed on a small aluminum foil, which is folded into a packet and autoclaved at 121 °C for 15 minutes. The sterile discs or squares are inoculated by immersing into drops of a heavy yeast suspension prepared in 5% skim milk. The packets are dried in a desiccator for 2–3 weeks at 4 °C and then placed in an airtight container and stored at 4 °C. Cultures are recovered by aseptically removing a piece of filter paper and inoculating the appropriate solid or liquid medium. Yeast maintained in this manner have been viable after 2–3 years.

L-Drying. L-drying refers to liquid-state drying in such a way as to prevent freezing. Drying occurs under vacuum at temperatures generally no colder than 5–10 °C. The method has been improved and simplified by Malik (43) for the preservation of sensitive microorganisms that are damaged by freezing or freeze-drying. The technique has been used to successfully preserve some yeasts and filamentous fungal spores.

Long-Term Methods

Freeze-Drying, or Lyophilization. The equipment, cryoprotectants, and theoretical aspects of freeze-drying have already been discussed. Fungi are grown under optimum conditions using media that will produce maximum sporulation. Mature spores are harvested by flooding the agar cultures with approximately 2 ml of a 20% sterile skim milk solution. The skim milk must be cold when used, and thus it is stored at 2–8 °C until required. The spores are gently scraped from the surface of the culture to yield a suspension containing at least 10^6 spores/mL. The spore suspension is pipetted back into the tube containing the remaining milk and mixed thoroughly. If more than one tube is used, repeat the procedure and pool the suspension in one tube to yield a concentration of at least 10^6 spores/mL. The spores of cultures grown in broth medium are centrifuged and resuspended in sterile 20% skim solution. Approximately 0.2 mL of the spore suspension is dispensed into sterile vials for freeze-drying. The interval between harvesting the spores and dispensing must be minimized because many spores begin to germinate when suspended in liquid. Spores should not be in the skim milk for more than 2 h before being processed. Filled vials should be refrigerated while awaiting further processing.

The fungi are freeze-dried using one of the ATCC methods already described. Lyophilized spores can be stored at 2–8 °C. Extended shelf life can be achieved by storing the vials at −70 °C in a mechanical freezer.

Freeze-dried fungal cultures are rehydrated with 0.5–0.9 mL of sterile water, then transferred to tubes containing 6 mL sterile water and allowed to soak for 50–60 min before transferring to solid medium.

At the ATCC, yeasts are processed in the same manner as bacteria for the component freeze-drying method. Cultures are grown on the appropriate solid medium and suspended in sterile 20% skim milk or reagent 18 to a concentration of at least 10^6 cells/mL. The suspension is mixed thoroughly, and 0.2 mL is dispensed into each vial. The cultures are freezed-dried using the component freeze-drying method described earlier.

Ultrafreezing, or Cryopreservation. The preservation of fungi at ultralow temperatures of liquid nitrogen (−196 °C for liquid phase and −150 °C for vapor phase) is presently considered the best method of storage. The method can be applied to both sporulating and nonsporulating strains, as well as those fungi that have been genetically altered or harbor plasmids. The freezing rates and cryoprotective agents described for bacteria are also applicable to the freezing of fungi. Spores or mycelial fragments are harvested by flooding slants or plates with 10% glycerol or 5% DMSO and gently scraping the surface of the cultures. The fungal suspension is dispensed in 0.5-mL amounts into sterile plastic freezing vials and frozen at a controlled rate. Fungi that produce sticky mycelia or whose mycelia grow embedded in the agar can be prepared for freezing by cutting agar plugs containing new growth (hyphyl tips) with

a sterile cork borer (5 mm). Three or four plugs are place into each sterile plastic vial with 0.4 mL of 10% glycerol and frozen at a slow controlled rate.

Cultures grown in broth are fragmented in a sterile Waring blender in a biological hood and suspended in equal parts of 20% glycerol and growth medium or equal parts of 10% DMSO and growth medium to give a final concentration of 10% glycerol (vol/vol) or 5% (vol/vol) DMSO, respectively. Some strains must be concentrated by centrifugation in order to obtain sufficient material for freezing. Pathogens must not be macerated in a mechanical blender because of the hazard of aerosol dispersion.

Some cultures, such as *Agaricus* strains, can consist of seeds, grains, or pollen and be frozen without the presence of a cryoprotectant. Slime molds can be preserved by freeze-drying or freezing. Spores, microcysts, and spherules are preserved by a controlled freeze in 10% glycerol and stored in the liquid-nitrogen vapor phase. Thawing of the frozen culture should be done rapidly in a 37 °C water bath with moderate agitation. Once the ice melts, the culture can be transferred to a suitable medium.

Yeast cells have been successfully preserved by controlled freezing to liquid-nitrogen temperature (−196 °C liquid, −150 °C vapor phase). The yeast cultures are handled in the same manner as bacteria and grown as described in the freeze-drying section for yeast. Cell suspensions are prepared in 10% glycerol to a concentration of at least 10^6 cells/mL. The yeast are frozen under controlled conditions as described in the bacterial section.

Several reports have described the loss of plasmid expression or the induction of respiratory-deficient mutants in yeast as a result of freeze-drying or improper freezing (44). Nierman and Feldblyum (33) reported that when plasmid-containing *Saccharomyces cerevisiae* were preserved by freezing in liquid nitrogen, no loss of viability was observed, and the fraction of plasmid-containing cells was the same before and after freezing. The preservation of plasmid-containing yeast is superior in liquid nitrogen with a controlled freezing rate than by freeze-drying.

CULTURE AND PRESERVATION OF CELL CULTURES

Long-Term Preservation

Ultrafreezing, or Cryopreservation. Storage in liquid nitrogen is currently the best method of preserving cultured cells. Cryopreservation methods for a variety of cell lines are well developed and used extensively for culture storage. For a detailed description of standard cryopreservation protocols see reports by Freshney (45) and Hay (46).

The general events that occur during freezing and the effect of the cooling rate, cryoprotective agents, and a controlled rate of cooling have been discussed in the bacterial section.

Cell suspensions are prepared in the same manner used for routine subcultivation; trypsin is added if necessary to produce a uniform suspension. The cell suspension is centrifuged at approximately 100 *g* for 10 min, and the resultant pellet is resuspended in an appropriate amount of fresh medium containing a cryoprotectant. Normally, 5–10% DMSO or glycerol is employed. The freeze medium should be prepared just prior to use by mixing fresh growth medium and the cryoprotective agent and kept at room temperature. The remaining cryoprotectant is discarded because oxidative contaminants can accumulate. A cell suspension containing approximately 10^6–10^7 cells/mL is satisfactory. The cell suspension is dispensed in 1-mL amounts into each glass vial or plastic ampule and frozen at a controlled rate of 1 °C/min to about −50 °C. The vials are immediately transferred into liquid nitrogen for storage in the liquid or vapor phase. Rapid thawing is essential for recovery of frozen cells, thus the frozen vials are immersed directly into a 37 °C water bath and agitated until the suspension is completely thawed. The thawed cell suspension is transferred to 10 mL of sterile growth medium and centrifuged at 100 *g* for 10 min. The supernatant containing the cryoprotectant is discarded, and the cell pellet is resuspended in fresh medium and propagated using standard procedures.

Quality Control. It is important that all preservation methods include standard operating procedures for the determination of culture purity, viability, retention of morphological and physiological characteristics, production of desirable products, and other key attributes. Purity checks must be conducted before and after processing microorganisms for freeze-drying and microorganisms and cell lines for freezing. Generally, the cell suspension is diluted and inoculated into suitable medium and incubated at optimum temperature. After growth appears, the cultures are examined for purity. It is also important to conduct pre- and postpreservation viability checks for freeze-drying and freezing to determine the effectiveness of the process. In addition, periodic viability assays should be performed to ascertain the shelf life of the culture.

Cell lines should also be examined for *Mycoplasma* contamination, because it has been estimated that about 10% of cell lines are contaminated (47). It is also important to conduct tests to verify cell-line identity; several rapid techniques have been developed that use DNA probes to identify the cell lines (48).

Documentation. Documentation plays an important role in the maintenance and preservation of cultures. Important data include the identity of the culture, isolation source and methods, the individual who isolated it, geographical location, morphological and physiological characteristics, maintenance data, and other salient features. If the culture was obtained from another investigator, the documentation should include the name of the investigator, culture history if available, and date of acquisition. The greater the amount of data recorded and regularly updated, the greater the value of the culture.

BIBLIOGRAPHY

1. R.J. Heckly, *Adv. Appl. Microbiol.* **24**, 1–53 (1978).
2. D. Smith and J. Kolkowski, in J.C. Hunter-Cevera and A. Belt eds., *Maintaining Cultures for Biotechnology and Industry*, Academic, New York, 1996, pp. 101–132.
3. R.J. Hay, in I. Freshney ed., *Animal Cell culture: A Practical Approach*, IRL Press, Oxford, England, 1986, pp. 71–112.

4. F. Kauffman, *The Bacteriology of Enterobacteriaceae*, Williams & Wilkins, Baltimore, 1966, pp. 1–400.
5. L.R. Hill and B.E. Kirsop eds., *Living Resources for Biotechnology*, Cambridge University Press, Cambridge, 1991.
6. B.E. Kirsop and J.J.S. Snell eds., *Maintenance of Microorganisms: A Manual of Laboratory Methods*, Academic, London, 1984.
7. R.E. Miller and L.A. Simmons, *J. Bacteriol.* **84**, 1111–1114 (1962).
8. L.R. Rey, in V.J. Cabasso and R.H. Regamey eds., *Development in Biological Standardization. International Symposium of Freeze Drying of Biological Products*, S. Karger, Basel, 1977, 19–27.
9. T.W.G. Rowe and J.W. Snowman, *Edwards Freeze-drying Handbook*, Edwards High Vacuum, Inc., Grand Island, New York, 1976.
10. R.J. Heckly, *Adv. Appl. Microbiol.* **3**, 1–76 (1961).
11. R.J. Heckly, *Dev. Ind. Microbiol.* **26**, 379–395 (1985).
12. H.T. Merryman, in H.T. Merryman ed., *Cryobiology*, Academic, New York, 1966, pp. 609–663.
13. R.H. Rudge, in B.E. Kirsop and A. Doyle eds., *Maintenance of Microorganisms and Cultured Cells*, 2nd ed., Academic, London, 1991, pp. 31–44.
14. S.P. Lapage, J.E. Shelton, T.G. Mitchell, and A.R. Mackenzie, in J.R. Norris and D.W. Ribbons eds., *Methods in Microbiol*, vol. 3A, Academic, London, 1970, pp. 135–228.
15. R.L. Gherna, in P. Gerhardt, R.G.E. Murray, W.A. Wood, and N.R. Krieg eds., *Methods for General and Molecular Bateriology*, American Society for Microbiology, Washington, D.C., 1994, pp. 278–292.
16. F.P. Simione and E.M. Brown eds., *ATCC Preservation Methods: Freezing and Freeze-drying*, American Type Culture Collection, Rockville, Md., 1991.
17. W.C. Haynes, L.J. Wickerham, and C.W. Hesseltine, *Appl. Microbiol.* **3**, 361–368 (1955).
18. H. Hippe, in B.E. Kirsop and A. Doyle eds., *Maintenance of Microorganisms and Cultured Cells*, 2nd ed., Academic, London, 1991, pp. 101–113.
19. V.J. Cabasso and R.H. Regamy eds., *International Symposium on Freeze-Drying of Biological Products*, Dev. Biol. Stand. vol. 36, 1977.
20. H.T. Merryman, in H.T. Merryman ed., *Cryobiology*, Academic, New York, 1966, pp. 1–114.
21. K.F. Redway and S.P. Lapage, *Cryobiology* **11**, 73–75 (1974).
22. G.M. Fahy, *Cryobiology* **23**, 1–13 (1986).
23. H.T. Merryman, *Cryobiology* **8**, 173–183 (1971).
24. R.K.A. Feltham, A.K. Power, P.A. Pell, and P.H.A. Sneath, *J. Appl. Bacteriol.* **44**, 313–316 (1978).
25. M.L. Speck and R.A. Cowan, in H. Iizuka and T. Hasegawa eds., *Proceedings of the First International Conference on Culture Collections*, Univ. of Tokyo Press, Tokyo, 1970, pp. 241–250.
26. A.P. Mackenzie, in V.J. Cabasso and R.H. Regamey eds., *International Symposium on Freeze-Drying of Biological Products*, S. Kargel AG, Basel, 1977, pp. 263–277.
27. P. Mazur, in H.T. Merryman ed., *Cryobiology*, Academic, New York, 1966, pp. 213–315.
28. L.R. Hill, M. Kocur, and K.A. Malik, in L.R. Hill and B.E. Kirsop eds., *Living Resources for Biotechnology*, Cambridge University Press, Cambridge, 1991, pp. 62–80.
29. N.R. Krieg and J.G. Holt eds., *Bergey's Manual of Systematic Bacteriology*, vol. 1, Williams & Wilkins, Baltimore, 1984.
30. P.H.A. Sneath, N.S. Mair, M.E. Sharpe, and J.G. Holt eds., *Bergey's Manual of Systematic Bacteriology*, vol. 2, Williams & Wilkins, Baltimore, 1986.
31. J.T. Staley, M.P. Bryant, N. Pfennig, and J.G. Holt eds., *Bergey's Manual of Systematic Bacteriology*, vol. 3, Williams & Wilkins, Baltimore, 1989.
32. S.T. Williams, M.E. Sharpe, and J.G. Holt eds., *Bergey's Manual of Systematic Bacteriology*, vol. 4, Williams & Wilkins, Baltimore, 1989.
33. W.C. Nierman and T. Feldblyum, *Dev. Ind. Microbiol.* **26**, 423–434 (1985).
34. T. Imanaka and S. Aiba, *Ann. N.Y. Acad. Sci.* **36**, 1–14 (1981).
35. A.H.S. Onions and J.I. Pitt, in D.L. Hawksworth and B.E. Kirsop eds., *Living Resources for Biotechnology*, Cambridge University Press, Cambridge, 1988, pp. 188–193.
36. D. Smith and J. Kolkowski, in J.C. Hunter-Cevera and A. Belt eds., *Maintaining Cultures for Biotechnology and Industry*, Academic, San Diego, 1996, pp. 101–132.
37. D. Smith, in D.L. Hawsworth and B.E. Kirsop eds., *Living Resources for Biotechnology*, Cambridge University Press, Cambridge, 1988, pp. 75–99.
38. I.C. Tommerup and D.K. Kidby, *Appl. Environ. Microbiol.* **37**, 831–835 (1979).
39. D. Smith, in B.E. Kirsop and J.J.S. Snell eds., *Maintenance of Microorganisms: A Manual of Laboratory Methods*, Academic, London, 1984, pp. 83–107.
40. B.E. Kirsop, in B.E. Kirsop and J.J.S. Snell eds., *Maintenance of Microorganisms: A Manual of Laboratory Methods*, Academic, London, 1984, pp. 109–130.
41. R.G. Atkinson, *Can. J. Bot.* **32**, 542–547 (1953).
42. R.G. Atkinson, *Can. J. Bot.* **32**, 673–678 (1954).
43. K.A. Malik, *J. Microbiol. Methods* **12**, 125–132 (1990).
44. B.M. Pearson, P.J.H. Jackman, K.A. Painting, and G.J. Morris, *Cryo-Letters* **11**, 205–210 (1990).
45. R.I. Freshney, *Culture of Animal Cells: A Manual of Basic Technique*, 3rd ed., Wily-Liss, New York, 1994.
46. R.J. Hay, in R.I. Freshney ed., *Animal Cell Culture: A practical Approach*, IRL Press, Oxford, England, 1986, pp. 71–112.
47. R.J. Hay, in J.C. Hunter-Cevera and A. Belt eds., *Maintaining Cultures for Biotechnology and Industry*, Academic, San Diego, 1996, pp. 161–178.
48. D.A. Gilbert, Y.A. Reid, M.H. Gail, D. Pee, C. White, R.J. Hay, S.J. O'Brien, *Am J. Hum. Genet.* **47**, 499–514.

DENATURATION, AGGREGATION.

See DENATURATION, PROTEINS, SOLVENT MEDIATED; PROTEIN AGGREGATION, DENATURATION; THERMAL UNFOLDING, PROTEINS.

DENATURATION, PROTEINS, SOLVENT MEDIATED

VADIM V. MOZHAEV
University of Iowa
Iowa City, Iowa

KEY WORDS

Catalytic activity
Conformation
Denaturation
Enzymes
Inactivation
Organic solvents
Proteins
Stability
Stabilization

OUTLINE

INTRODUCTION

The use of enzymes in organic solvents for biotransformations is now commonplace. The nonaqueous milieu permits the advantages of organic solvents as the reaction media, and enzymes provide the opportunity for high selectivity and catalytic rates under mild reaction conditions (1). Indeed, enzymes are fantastic catalysts in terms of effectiveness—enzyme-catalyzed reactions provide up to 10^{15}-fold rate accelerations as compared with uncatalyzed processes—and reaction specificity (e.g., enantio-, regio-, and chemioselectivity). Unfortunately, exploiting such advantages is often limited by the low stability of the biocatalysts. Even in aqueous environments, enzymes undergo denaturation that is, they loose their unique three-dimensional structure and catalytic activity. The risk of denaturation is even greater when proteins are placed into conditions much less "natural" than the conventional aqueous environment. The goal of this review is to consider the general mechanisms behind the often low stabilities of proteins in the presence of organic solvents, and to evaluate more specific mechanisms responsible for denaturation of proteins in systems with organic solvents. Special attention is given to enzymes as the most important class of proteins in biotransformations. Successful approaches elaborated to increase protein stability are also discussed.

GENERAL DEFINITIONS

Protein Denaturation

Mechanisms of protein denaturation in aqueous media have been studied for decades and are mostly understood now. Denaturation of proteins is caused by various factors such as high and low temperatures, high pressures, ultrasound, high-intensity irradiation (including microwaves), addition of organic solvents, certain salts, and specific compounds such as urea. Despite the high structural diversity of proteins, their denaturation is described by very similar features. Denaturation, or more specifically inactivation of enzymes (i.e., loss of their catalytic function) can be either reversible or irreversible. For reversible denaturation, the enzyme completely recovers the native structure and catalytic activity after the denaturing action is eliminated, for example, by reestablishing room temperature after heating or by removing chemical denaturants during dialysis. Reversible denaturation is often a two-state transition between the folded native state, N, and the unfolded denatured state, U (2). Because structural differences between N and U are significant, concentration of both forms can be followed in the transition zone quite easily by circular dichroism or fluorescence spectroscopy, or other physical methods. This ability to follow concentrations of both protein forms during denaturation enables researchers to obtain the thermodynamic characteristics of protein stability and quantitatively compare the stability under a variety of external conditions including organic solvents.

Often after incubation under denaturing conditions (especially at temperatures higher than 50–60 °C), the enzyme does not recover its native structure and catalytic activity; this is irreversible denaturation (3,4). The reason may be purely kinetic, for example, heat-denatured enzyme remains in a catalytically inactive conformation after

cooling because at room temperature the protein molecule is not flexible enough to surpass the high kinetic barrier that exists between the denatured and the native states. On the other hand, in the unfolded state, the protein becomes more susceptible to aggregation and chemical changes in the structure that also lead to irreversible inactivation. Table 1 lists inactivation mechanisms.

Studying Reversible and Irreversible Denaturation. When protein is reversibly denatured, the two protein forms, N and U, are in equilibrium, which is governed by the equilibrium constant K:

$$\mathrm{N} \overset{K}{\leftrightarrow} \mathrm{U},\ K = [\mathrm{U}]/[\mathrm{N}]$$

where [N] and [U] represent concentrations of the native and unfolded proteins, respectively, and are determined experimentally under denaturing conditions by measuring catalytic activity or physical characteristics of the protein. Other thermodynamic parameters of denaturation are determined from the value of K by studying the temperature dependence of denaturation:

$$\Delta G_{\mathrm{D}} = -RT\ (\ln K); \qquad \Delta H_{\mathrm{D}}/RT^2 = d(\ln K)/dT$$
$$\Delta S_{\mathrm{D}} = (\Delta G_{\mathrm{D}} - \Delta H_{\mathrm{D}})/T; \qquad \Delta c_{\mathrm{p}} = d(\Delta H_{\mathrm{D}})/dT$$

where ΔG_{D} is the free energy of denaturation, R is the gas constant, and T is the temperature in degrees kelvin. ΔH_{D} is the enthalpy of denaturation, ΔS_{D} is the change in entropy of denaturation, which is a characteristic of the degree of protein unfolding, and Δc_{p} is the change in heat capacity on denaturation and is a measure of the hydration effects in a protein. This formalism also applies to denaturation of proteins in organic solvents, at least in cases when the denaturation is not complicated by irreversible processes such as aggregation.

Irreversible inactivation is studied by following the rate, v_{in}, of disappearance of the native form of the protein:

$$\mathrm{N} \overset{k_{\mathrm{in}}}{\rightarrow} \mathrm{I};\ v_{\mathrm{in}} = -d[\mathrm{N}]/dt = k_{\mathrm{in}}[\mathrm{N}]$$

where I is the inactivated enzyme and k_{in} is the rate constant of inactivation. Aliquots of solution, or suspensions with enzyme in organic solvents are sampled at time intervals, and the residual enzymatic activity is assayed in aqueous solution (or in organic solvent).

Protein Stability

Closely related to low stability is the problem of stabilization of enzymes. In the last two decades, numerous approaches have been elaborated to increase stability of proteins in aqueous media; see Refs. 3–6 for reviews. These approaches employ immobilization and cross-linking of proteins, the methods of site-directed and random mutagenesis, and covalent modification of proteins; also, stabilization is achieved by addition of low molecular weight compounds. Some of these approaches have been successfully used in organic solvents and are discussed in this review.

Despite the variety of organic solvent systems containing enzymes, two of them are most important in the practical sense and for understanding protein denaturation by organic solvents. First, enzymes are often dissolved in binary mixtures of water and polar organic solvents with concentrations of the solvent as high as several dozens volume percent. Second, lyophilized or immobilized enzymes are studied when suspended in neat organic solvent or in the solvent containing as little as a few volume percent of water. These two types of systems are the focus of this review.

Table 1. Inactivation Mechanisms and Experimental Methods for Their Study

Inactivation mechanism	Methods for identifying inactivation mechanism
1. Aggregation (sometimes followed by formation of intermolecular S-S bonds)	Formation of precipitates visible by eye; smaller aggregates are detected by electrophoresis or sedimentation
2. Changes in primary structure	
a. Hydrolysis of peptide bonds	Detection of low molecular weight protein fragments by SDS electrophoresis
b. Oxidation of Cys, Trp, Met	Deceleration of protein inactivation by inert gases
c. Reduction of S-S bonds by OH^-, followed by β-elimination and reactions of formed SH groups with Cys or other amino acids	Detection of emerging SH groups by Ellman's reagent; disappearance of Cys, Lys, and Arg, and appearance of lysinoalanine, dehydroalanine, and ornithinoalanine, detected by amino acid analysis
d. Modification of SH groups by metal ions to form mercaptides	Decrease in free SH groups, detected by Ellman's reagent
e. Deamidation of Asn and Gln	Formation of proteins with smaller isoelectric points; detection of ammonia with glutamate dehydrogenase
f. Racemization of amino acids	Only above 100 °C
3. Elmination of coenzyme from the active site	Disappearance of spectral bands characteristic of coenzyme attached to proteins
4. Dissociation of oligomeric proteins into monomers	Detection of dissociation products by chromatography or electrophoresis
5. Adsorption on the vessel surface	Sometimes desorption by added denaturants
6. "Incorrect refolding," with production of mismatched forms	Reactivation of mismatched protein forms by unfolding and subsequent refolding

Source: Ref. 3.

ENZYMES IN MONOPHASIC WATER–COSOLVENT MIXTURES

Reversible Denaturation

Exploiting enzymes in peptide synthesis and other biotransformations demands the use of mixtures of water with polar organic solvents (i.e., alcohols and formamides, among others) as the reaction media. However, almost all organic solvents in binary mixtures with water have unfavorable denaturing effects on enzymes (7). Even polyhydric alcohols (glycerol and ethylene glycol, among the others), which at low concentrations are known to stabilize proteins against heat treatment, induce denaturation at concentrations higher than 50 vol % (8). Catalytic activity of enzymes dissolved in mixtures of water with a polar organic solvent are described by a characteristic profile, shown in Figure 1. Addition of cosolvent to aqueous solution of enzyme up to some critical concentration does not change or increase enzymatic activity. This activation could often be explained by more favorable solvation of the polar transition state of enzymatic reaction by water–cosolvent mixture than by water. However, beginning with this critical concentration of a cosolvent, enzymatic activity drops (often reversibly) in a threshold manner, reaching a very low or even zero value. This threshold decrease in activity is accompanied by an abrupt change in spectroscopic or other physical properties of the dissolved enzyme, and this fact has been explained by a conformational transition in the protein molecule. The fact that the plummet in enzymatic activity occurs at the same concentration as the spectral changes do, implies that it is the denaturation that causes the threshold inactivation of the dissolved enzymes.

Mechanistic description of this inactivation is based on the important prerequisite that the protein molecule in aqueous solution is surrounded by a shell formed by water molecules that are bound by hydrogen bonds to the protein surface. According to a widely accepted notion (9), this hydration shell, or at least a portion of it, represents an integral part of the protein and is essential for protein structure and function. Consequently, displacement of bound water molecules by an organic solvent results in a dramatic change of the whole structure and leads to protein denaturation. In line with these ideas, reversible inactivation of enzymes in water–cosolvent mixtures consists of three major steps, described in Figure 2: *(1)* dehydration of the protein molecule, that is, removal of a certain critical amount of water molecules from the protein hydration shell; *(2)* binding of the organic cosolvent by the partially dehydrated protein molecule; and *(3)* conformational transition in the protein molecule, resulting in the formation of a denatured form, D, with concomitant loss of catalytic activity. Thermodynamic consideration of this model the details of which are beyond the scope of this article, leads to the following expression for the free energy of denaturation (ΔG_D) of proteins by organic solvents (8):

$$\Delta G_D = B_0 + B_1 n + B_2 E_T(30) + B_3 n \log P$$

where B_0, B_1, B_2, and B_3 are numerical coefficients, which remain unchanged for a given protein at constant temperature, regardless of the nature of the solvent; n is a numerical parameter that depends on the size and shape of the molecules of organic solvents and gives an idea of how many molecules of water can be displaced from the protein surface by one molecule of the organic solvent; $E_T(30)$ is the Dimroth–Reichardt parameter, which is directly related to the free energy of the solvation process of the organic cosolvent; P, the partition coefficient of the solvent in a water–octanol biphasic system, is a measure of organic solvent hydrophobicity. The values of $E_T(30)$ and P for different solvents are well documented, and the parameter n can be calculated from the information on the organic solvents contained in comprehensive handbooks on organic and physical chemistry.

Equation 1 predicts that cosolvents having high values of $E_T(30)$ and $\log P$ possess greater denaturing ability. In

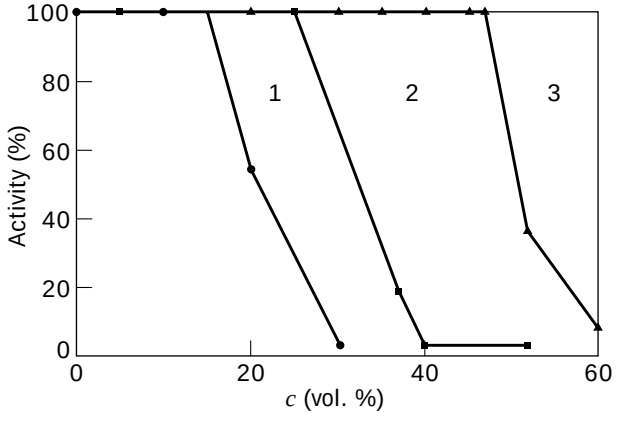


Figure 1. Dependence of relative catalytic activity of dissolved enzymes on the concentration of organic solvents in binary mixtures with water. (1) laccase in water–acetone mixtures; (2) α-chymotrypsin in water–ethanol mixtures; (3) trypsin in water–acetonitrile mixtures.

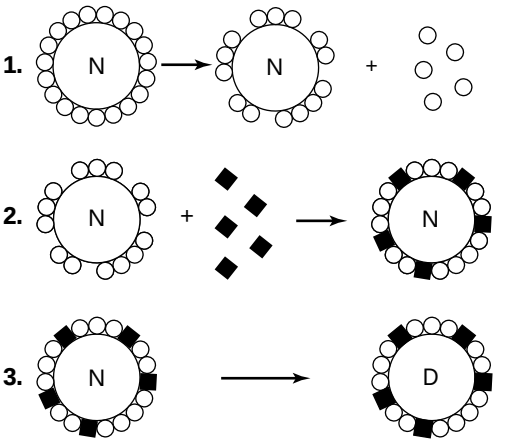


Figure 2. Schematic representation of major molecular steps involved in reversible protein denaturation by organic solvents. ○, water molecule; ■, organic cosolvent molecule; *N* states for the native protein, *D* for the denatured protein.

fact, the solvents such as 1,4-dioxane, tetrahydrofuran, and higher alcohols (e.g., isomers of butanol), which are both hydrophobic and have a high solvation capacity, appear to be strong denaturants and cause enzyme inactivation at concentrations as low as 10–30% (v/v). Inversely, Equation 1 predicts that hydrophilic solvents, such as glycerol, ethylene glycol, and formamide, have a low denaturing capacity. Indeed, it has been experimentally confirmed that many enzymes are not inactivated by these solvents at concentrations as high as 50–60 vol % (8). Thus, solutions with high concentrations of these solvents are convenient as a media for enzyme-catalyzed reactions carried out in binary water–cosolvent mixtures.

Stabilization of Enzymes against Denaturation by Organic Cosolvents

Molecular events behind enzyme inactivation in water–cosolvent mixtures also prompt ideas of how to stabilize enzymes against this denaturation. Because protein unfolding is an indispensable step in the inactivation by solvents (Fig. 2), one could try the approaches recommended as best for deceleration of protein unfolding in aqueous solutions. By using the method of covalent immobilization, proteolytic enzymes were stabilized against denaturation by alcohols and dimethylformamide (10,11); for other examples, see the literature cited in Ref. 11. The stabilization effect expressed as an increase in the threshold concentration at which enzymatic activity dropped abruptly, rose with the number of covalent bonds between the enzyme and support. Stabilization against denaturation by polar organic solvents has been also obtained when noncovalent complexes of enzymes with polyelectrolytes were formed due to multiple electrostatic interactions between oppositely charged groups in the interacting macromolecules (12). The delay in activity decrease in this case was also accompanied by a shift in the concentration of organic solvents, at which point significant structural changes in the protein molecule (unfolding) occurred. Probably the same reason (i.e., increase in rigidity of the protein molecule in water–cosolvent mixtures) underlies stabilization of enzymes by chemical cross-linking with bifunctional reagents, as was the case with cross-linked crystals of thermolysin in dimethylformamide (13). Equilibrium between the native and denatured states of a protein (step 3 in Fig. 2) can be efficiently shifted toward state N by addition of substrates and specific ligands that are preferentially bound with the native enzyme. Due to this reason, enzymes are often more stable in organic solvents in the presence of saturating concentrations of their substrates.

Another productive idea for stabilization of enzymes in water–cosolvent mixtures is to strengthen the binding of the hydration shell to the protein surface. As seen from Figure 2, protein denaturation includes the step of dehydration, namely, dissociation of essential water molecules from the hydration shell. It has been assumed that if water molecules are more tightly bound with the protein, denaturation could be decelerated. Indeed, introduction of polar and charged groups into the protein by covalent modification led to a significant increase in the concentration of organic solvents that caused denaturation of α-chymotrypsin. Stabilization was observed when enzymes were modified with low molecular weight compounds (pyromellitic anhydride, glyceraldehyde) (14) and poly(ethylene glycol) (PEG) (15). This stabilization was explained by the ability of the newly attached hydrophilic (carboxylic and hydroxylic) groups to hold water, thus preventing protein dehydration by organic cosolvent molecules. Tighter binding of the hydration shell could also be achieved due to replacements of the proteins' amino acid residues by site-directed or random mutagenesis. Recent studies from Arnold's and Wong's groups showed that such replacements increased the stability of proteases in the presence of a strong denaturant, dimethylformamide (16,17). The stabilization was explained by numerous reasons, such as enhancement of internal hydrogen bonds and ionic forces, optimization of internal hydrophobic interactions, reduction of solvation energy, and introduction of conformational restrictions to the protein molecule. Also it is possible that chemical modification or replacement of amino acid residues on the protein surface both change the protein's affinity to organic solvent molecules and, according to the schemes in Figure 2, decrease the protein's sensitivity to denaturation. A stabilizing approach recently developed by Arnold (18) and Wong (19) suggests increasing the rigidity of a protein by introducing amino acid residues (e.g., histidines) that are able to specifically bind with metal ions. In fact, in organic solvents, catalytic activity and stability of such metal-chelated proteases was significantly higher than in their nonmodified analogs.

Irreversible Inactivation: Precipitation

It has been commonly noticed that the initially reversible denaturation of proteins by organic solvents is further accompanied by irreversible inactivation during long (a dozen minutes to hours) incubation. This irreversible inactivation becomes especially important when the concentration of cosolvent is higher than the denaturation threshold. Proteins may even precipitate if their concentration in water–cosolvent solution is of the order of micromoles per liter or lower; normally these concentrations are used to detect catalytic activities. The general features of this inactivation were described by Singer (7), who found that proteins lose only a part of their native structure in organic solvents—the denatured protein retains a large portion of secondary structure, although the tertiary structure is destroyed. This conformation, analogous to the recently discovered molten globule state, turns out to be very "sticky," and proteins have a strong tendency to form inactive aggregates due to formation of interprotein hydrogen bonds and hydrophobic interactions. However, even in the insoluble form, enzymes may show a small, though detectable, activity in the presence of high concentrations (60–80 %, v/v) of polar solvents. For example, such an activity was detected when solid α-chymotrypsin was added to a water–cosolvent mixture (20). On the other hand, the enzyme does not show any activity when the precipitate is formed after an aqueous solution of the enzyme is added to the organic solvent. This difference does not seem too surprising in view of the recent fact (21) that the enzyme directly suspended in water–cosolvent mixture contains a higher

percentage of the native secondary structure than the enzyme first dissolved in water and then mixed with the solvent to form a precipitate. Most likely, in the former case, only the enzyme molecules forming the surface layer of particles are inactivated due to direct contact with the organic solvent, whereas the enzyme molecules inside the particles are protected from denaturation by interprotein contacts. Oppositely, in the latter system, enzyme molecules are first denatured and then aggregate, giving rise to inactive precipitates.

This inactivation caused by aggregation has been successfully combated by protein immobilization or by modification with polymers such as PEG; both are highly efficient in preventing interprotein contacts. Aggregate formation and protein precipitation can also be delayed by chemical modification of proteins with low molecular weight compounds (for example, charged reagents) or by replacement of amino acid residues on the protein surface through mutagenesis. Sometimes these modified enzymes remain catalytically active without noticeable precipitation in the significantly wider range of cosolvent concentrations (by 30–40 vol %) than their nonmodified counterparts (V.V. Mozhaev and E.V. Kudryashova, 1996).

SUSPENSIONS OF ENZYMES IN ANHYDROUS MEDIA

Neat organic solvents or solvents containing only a few volume percent of water are very convenient media for many enzymatic reactions (Table 2). Most proteins are insoluble or sparingly soluble in neat organic solvents and most often are used as suspensions of freeze-dried protein powders. These systems have found numerous applications in chemical processing, food-related conversions, and analyses (1).

Table 2. Potential Advantages of Enzymatic Processes in Nonaqueous Media

1. Increased solubility of apolar substrates.
2. Enhanced thermostability.
3. Easier product recovery from low-boiling sovents.
4. Immobilization is often unnecessary because the enzyme is insoluble.
5. Easier recovery of the enzyme by simple filtration or centrifugation.
6. If immobilization is required for better performance, simple adsorption onto nonporous surfaces may be sufficient. Enzymes do not desorb from these surfaces in nonaqueous media.
7. Alteration in substrate specificity.
8. Suppression of water-dependent side reactions, e.g., nonenzymatic blank reaction in enzymatic hydrocyanations of aldehydes, which leads to racemic product.
9. Shifting thermodynamic equilibria to favor synthesis over hydrolysis, e.g., esterification and peptide formation.
10. The possibility of carrying out reactions with other neucleophiles, which would not be feasible in water, e.g., transesterification and ester aminolysis.
11. Easier integration of enzymatic with conventional chemical steps in a multistep synthesis (chemoenzymatic processes).

Source: Ref. 22.

The main problem that still remains with nonaqueous enzymology is a rather low catalytic activity of enzymes in anhydrous media in comparison with activity in water. Many factors such as diffusional (mass transfer) limitations, sterical shielding of the enzyme's active site from the access by a substrate, and unfavorable energetics of substrate desolvation may contribute to this decrease in activity (23). Here we focus on aspects of the problem that are related to the stability of enzymes.

Protein Denaturation During Lyophilization

It has been established that many proteins are denatured during lyophilization (24–26). This denaturation can be either reversible (lyophilized proteins regain their native three-dimensional structure and function when placed in aqueous solution) or irreversible. Especially sensitive to irreversible inactivation during lyophilization are oligomeric proteins, which often dissociate into individual subunits (25). Carpenter and coauthors have found that enzymes can lose their activity both at freezing and at drying, which constitute the lyophilization protocol, and that denaturation at these steps follows essentially different mechanisms.

Freezing. This process includes a number of "stresses" such as protein exposure to low temperatures and ice, contact with the test tube surface, concentration of the dissolved compounds, and eutectic crystallization of solution components, and each of these stresses may decrease enzyme activity. For example, eutectic crystallization from a buffer-containing solution may result in dramatic alterations in pH, for example, up to 3 pH units (27). Such changes largely influence the catalytic activity of enzymes, which is very sensitive to the ionization state of the enzyme's functional groups participating in catalysis. As shown for subtilisin Carlsberg in anhydrous media, the enzyme activity increased more than 300 times upon changing the pH of the aqueous solution (from which the enzyme was lyophilized), from 5 to 11 (28). However, addition of compounds such as polyvinylpyrrolidone and bovine serum albumin (BSA) during freezing is known to prevent such large shifts in pH (25) and may be helpful to increase the activity of enzymes in organic solvents.

Inactivation of enzymes in the freezing process is dominated mainly by unfolding of the protein molecule. Protection of labile proteins during freezing may be achieved by adding a variety of solutes, such as polyols, sugars, amino acids (normally at concentrations of a few hundred millimolar), synthetic polymer PEG, and a protein BSA, effective at concentrations of 1% (w/w) or smaller (25). This cryoprotection is explained by the mechanism of preferential hydration of proteins (6). Sugars, polyols, and some other compounds are preferentially excluded from contact with the surface of a protein because their interaction with a protein's functional groups is unfavorable in comparison with interactions of the same compounds with water molecules. Consequently, in the presence of these compounds, the protein is preferentially hydrated, that is, the hydration shell becomes more tightly bound to the protein molecule. This leads to an energetically unfavorable situation

as chemical potentials of both the protein and the additive increase. As a result, the compact native structure is stabilized by the additives because denaturation (unfolding) would lead to a greater surface area of contact between the protein and the solvent and would increase this thermodynamically unfavorable tendency.

Drying. The frozen protein sample is already significantly dehydrated due to formation of ice from the bulk water. However, an essential amount of water comprising the protein's hydration shell (several hundreds of water molecules per protein of molecular weight of 20–25 kDa) remains unfrozen and is removed only by sublimation under high vacuum (i.e., by lyophilization). How much water remains bound to a protein after lyophilization is a question that is not resolved. Most likely, only a few water molecules are bound by several hydrogen bonds to polar groups in the protein's interior, whereas almost all water on the protein's surface is stripped. This latter fact leads to the appearance of a great number of polar groups with the noncompensated ability to form hydrogen bonds, which is satisfied by the formation (mainly intermolecularly) of β-sheet structures because the β-structure allows for the highest degree of hydrogen bonding in proteins (25,29). Due to low molecular flexibility in anhydrous media, these structural changes in freeze-dried enzymes are conserved, and the catalytically active conformation is not completely recovered after the enzyme is placed in organic solvent. Actually, active site titration of chymotrypsin in organic solvents has shown that about 40% of the active enzyme was destroyed at the drying step (26).

Effect of Excipients. Stabilization during the lyophilization procedure demands the protection of proteins against inactivation at both the freezing and drying steps. As discussed in the previous section, many additives are effective as protectors during freezing. Fewer compounds are known to protect enzymes against dehydration stress during drying; sugars are one example. Due to several hydroxyl groups, sugars substitute for molecules of water to form multiple hydrogen bonds with polar groups on the protein surface, thus preventing alterations in the protein's secondary structure (29). Especially successful are combinations of several additives. For example, PEG was shown to protect lactate dehydrogenase and phosphofructokinase from inactivation during freezing, but not to stabilize these proteins during drying (24). This is not surprising because PEG crystallizes during freezing and is then not able to form the hydrogen bonds with proteins that are necessary for protein protection at the drying step. However, in the presence of 1% PEG and small amounts of glucose or trehalose, which alone provide only minimal protection, the activity of both lyophilized enzymes recovers almost completely after rehydration in aqueous solution. In this case, PEG stabilizes the proteins by the mechanism of preferential hydration during freezing, while sugars act as water substitutes during drying because of their hydrogen bonding with the proteins. However, protection during freeze-drying is not the sole mechanism responsible for higher catalytic activity in organic solvents of enzymes lyophilized with excipients. For example, several proteases were shown to be activated a thousandfold in anhydrous media when they were lyophilized with an excess of KCl or other salts (30). Although certain salts are known to be good cryoprotectants, the activation of enzymes in that case was shown to be due to some other effect than protection during freeze-drying. The most plausible explanation, suggested by Khmelnitsky et al. (30), is that the activation was due to stabilization of the reaction transition state by a polar matrix microenvironment.

Inactivation at the steps of freezing and drying can, however, be avoided if a method other than lyophilization is used for preparing enzyme catalysts. Vacuum evaporation of aqueous solutions of enzymes or precipitation by acetone did not lead to better catalytic activities in anhydrous media. Much higher activities have been obtained with crystals of enzymes (31). In a crystalline microenvironment, additional polar and hydrophobic contacts between neighboring protein molecules significantly increase stability by preventing unfolding, aggregation, or dissociation in the case of oligomeric proteins. "Mild" cross-linking with bifunctional chemical reagents makes the crystals more solid (cross-linked enzyme crystals) and also increases protein stability. One more way to avoid lyophilization is to obtain noncovalent complexes of enzymes with surfactants, which exist due to ion-pairing between the molecules of surfactant and enzyme (32). Such complexes are quite soluble in nonpolar solvents (up to concentrations of several milligrams per milliliter) and possess very high catalytic activities. For example, the value of k_{cat}/K_m for transesterification of *N*-acetyl-L-phenylalanine ethyl ester catalyzed by the complex of subtilisin with Aerosol OT dissolved in octane was equal to 70% of the value for the hydrolysis reaction in water.

Effect of Organic Solvents on Structure and Activity of Enzymes

In a dry state almost all enzymes are inactive; they need a small quantity of water to display catalytic activity. This water provides enzymes with conformational mobility, which is necessary for catalytic function. Because of this, lyophilized (dry) enzymes are often preequilibrated with small amounts of water before starting reactions in organic solvents. However, contact with anhydrous solvents may lead to significant dehydration of enzymes ("stripping" essential water of the protein) and inactivate them (33). This tendency is especially strong with polar solvents. Some of them were shown to dehydrate enzymes on a very rapid time scale—5 min or faster (34). The driving force for dehydration is the higher affinity of water to molecules of polar solvent than to the protein's functional groups. The interaction energy for a hydrogen bond between water and dimethyl sulfoxide was calculated as equal to 8.8 kcal/mol; the solvent interaction with an amide H has the same energy, and both are much stronger than hydrogen bonds between water and a carbonyl O or an amide H and between a backbone carbonyl O and amide H (35). These data explain why dimethyl sulfoxide, as well as other polar solvents, has the tendency to strip water molecules from the protein surface and form new hydrogen bonds with protein functional groups. On the other hand, contact with non-

polar solvents leaves the major portion of the protein molecule hydrated and causes only small rearrangements of side chains on the protein surface (36).

Effects of polar solvents on the structure of suspended enzymes depend mainly on the nature of the solvent (37) and on enzyme stability (38). For example, acetonitrile, butanol, ethyl acetate, and tetrahydrofurane, each containing 1% (v/v) of water, produced few changes in the structure of bovine pancreatic trypsin inhibitor in addition to those already brought about by lyophilization (37). On the other hand, this protein was strongly denatured by methanol, dimethyl formamide, and dimethyl sulfoxide—solvents known for their ability to dissolve proteins. Contact with polar solvents is accompanied by significant changes in the protein's hydration. For example, 4% (v/v) isopropanol added to crystalline chymotrypsin in hexane nearly doubled the number of water molecules observable in the protein's crystal structure (36). These additional water molecules did not contact directly with the protein's surface groups but were integrated within the second hydration sphere of the protein. Structural changes in proteins caused by organic solvents have been successfully inhibited by the simple addition of low molecular weight compounds during lyophilization. For example, lyoprotectors such as ammonium sulfate or sucrose did not only stabilize enzymes during lyophilization but were quite efficient in protecting enzymes from inactivation by organic solvents (33), probably by maintaining water in the enzyme's active site.

Polar solvents may significantly increase molecular mobility of proteins in nonaqueous media by eliminating flexibility constraints imposed by protein–protein contacts. It has been shown that the rates of interconversion between several conformations of a peptide, gramicidin A, in an anhydrous milieu were increased in protic solvents due to a faster exchange between hydrogen bonds (39). Molecular dynamic simulations also showed that subtilisin is much more flexible in polar versus nonpolar solvents (35). Probably, such an increase in molecular flexibility by added polar cosolvents (dimethyl sulfoxide and formamide) is responsible for the more than hundredfold acceleration of the reactions catalyzed by subtilisin in neat nonpolar solvents (40), although this explanation has been disputed (36).

Recent studies with Fourrier-transform infrared spectroscopy have pointed out that enzymes conserve a much more nativelike secondary structure in neat solvents than they do in water–cosolvent solutions with a solvent concentration of 60 vol % (21). The explanation is that protein denaturation (unfolding) is significantly inhibited by the more severe conformational restrictions imparted on the enzyme molecules in anhydrous versus water–cosolvent media. This may explain why enzymes are often more catalytically active in neat organic solvents than in water–cosolvent mixtures, that is, why the addition of water is detrimental rather than beneficial for stability of the native enzyme's conformation (21).

High Thermal Stability of Enzymes in Anhydrous Media

The fact of the enormously high stability of lipase suspended in organic solvents was mentioned in one of the first works in nonaqueous enzymology: this enzyme retained high activity after several hours of incubation in anhydrous media at 100 °C (41). Further studies confirmed that high thermal stability is a hallmark of proteins in nonpolar solvents, as opposed to their low stability in aqueous media; for such comparison, see the recent review by Zaks (42). For example, the melting temperature for proteins suspended in nonpolar solvents is 110–145 °C, which is 50–70 °C higher than in aqueous solutions (43). Thermal stability of ribonuclease A was nearly the same when measured in either a free-of-solvent state or hydrophobic solvents (44). The explanation was that the conformational flexibility of proteins is significantly retarded in systems with a small water content (45), as the latter acts as a lubricant in all macromolecular motions, owing to its high propensity to form multiple hydrogen bonds with a protein. Water removal also increases the strength of intramolecular hydrogen bonds and salt bridges, due to the higher energy of electrostatic interactions in the media with low dielectrics. All this creates a high kinetic barrier between the native and unfolded states and stabilizes native conformations of proteins (42). Also molecular mobility is suppressed by protein–protein interactions, which exist in multitude in lyophilized and crystalline protein samples, where the packing density of macromolecules is very high. Elimination of water decelerates almost all detrimental chemical reactions (Table 1), which are also responsible for irreversible inactivation of proteins in nonaqueous media. For example, a very slow inactivation of proteins in nonpolar solvents at temperatures higher than 100 °C is caused by the formation of aggregates cross-linked by intermolecular reactions of transamidation and disulfide exchange (43).

On the other hand, proteins are significantly less thermostable in polar versus nonpolar solvents (46). The reason may be that in hydrophobic solvents a protein has more intraprotein hydrogen bonds, which make the protein molecule more rigid and more resistant to high-temperature unfolding. Thermal stability of proteins in organic solvents often declines on addition of even small amounts of water, which are quite sufficient to revive molecular dynamics in proteins. An almost linear decrease in the melting temperature has been observed as a function of the amount water in anhydrous media (44). This stability decrease was explained by stabilization of the unfolded protein state, rather than destabilization of the folded native state. High-temperature inactivation in anhydrous media is accompanied by migration into the solvent of water bound to protein; it was suggested that this dehydration could be the rate-limiting step of the whole inactivation process (47). The tendency to dehydrate proteins essentially depends on the hydrophobicity of organic solvents. Molecular dynamic simulation has shown that more water molecules leave the surface of a protein placed in dimethyl sulfoxide than in the acetonitrile simulation, whereas no water at all leaves the protein in carbon tetrachloride (35). Finally, protein thermal stability in anhydrous solvents significantly depends on conditions of lyophilization, such as pH of the buffer or presence of excipients (48). For example, the complex of chymotrypsin with sorbitol denatures in organic solvent at much higher

temperatures than the enzyme in the absence of sorbitol (49).

CONCLUSION

The problem of low activity and stability of enzymes in organic solvents led biochemists and biochemical engineers to the necessity of protein stabilization, and some recent advances were discussed in this review. A variety of methods have been applied, including familiar immobilization and chemical modification, and recently designed genetic procedures, which have become routine in the last decade. Essentially, not only protein stability is increased by methods of enzyme engineering, but also a variety of other important enzyme properties are changed, such as pH profiles of catalytic activity and stability, substrate specificity, including chemio-, regio- and enantioselectivity, and others. It is a task of the next decade to learn how to tailor in the most rational way the whole combination of enzyme properties that are necessary for successful utilization of enzymes in biotransformations.

BIBLIOGRAPHY

1. A.M.P. Koskinen and A.M. Klibanov, eds., *Enzymatic Reactions in Organic Media*, Blackie Academic & Professional, London, 1986.
2. P.L. Privalov and S.J. Gill, *Adv. Prot. Chem.* **39**, 191–234 (1988).
3. V.V. Mozhaev, *Trends Biotechnol.* **11**, 88–95 (1993).
4. A.M. Klibanov, *Adv. Appl. Microbiol.* **29**, 1–28 (1983).
5. B.W. Matthews, *Annu. Rev. Biochem.* **62**, 139–160 (1993).
6. S.N. Timasheff, *Annu. Rev. Biophys. Biomol. Struct.* **22**, 67–97 (1993).
7. S.J. Singer, *Adv. Prot. Chem.* **17**, 1–69 (1962).
8. Yu.L. Khmelnitsky, V.V. Mozhaev, A.B. Belova, M.V. Sergeeva, and K. Martinek, *Eur. J. Biochem.* **198**, 31–41 (1991).
9. J.A. Rupley, E. Gratton, and G. Careri, *Trends Biochem. Sci.* **8**, 18–22 (1983).
10. D.S. Clark and J.E. Bailey, *Biochim. Biophys. Acta* **788**, 181–188 (1984).
11. V.V. Mozhaev, M.V. Sergeeva, A.B. Belova, and Yu. L. Khmelnitsky, *Biotechnol. Bioeng.* **35**, 653–659 (1990).
12. A.K. Gladilin, E.V. Kudryashova, A.V. Vakurov, V.A. Izumrudov, V.V. Mozhaev, and A.V. Levashov, *Biotechnol. Lett.* **17**, 1329–1334 (1995).
13. N.L. St Clair and M. Navia, *J. Am. Chem. Soc.* **114**, 7314–7316 (1992).
14. Yu.L. Khmelnitsky, V.V. Mozhaev, A.B. Belova, and A.V. Levashov, *FEBS Lett.* **284**, 267–269 (1991).
15. V.V. Mozhaev, E.V. Kudryashova, N.V. Efremova, and I.N. Topchieva, *Biotechnol. Tech.* **10**, 849–854 (1996).
16. P. Martinez and F.H. Arnold, *J. Am. Chem. Soc.* **113**, 6336–6337 (1991).
17. C.-H. Wong, *Trends Biotechnol.* **10**, 378–381 (1992).
18. F.H. Arnold and J.-H. Zhang, *Trends Biotechnol.* **12**, 189–192 (1994).
19. R.D. Kidd, H.P. Yennawar, P. Sears, C.-H. Wong, and G.K. Farber, *J. Am. Chem. Soc.* **118**, 1645–1650 (1996).
20. Y. Tomiuchi, K. Ohshima, and H. Kise, *Bull. Chem. Soc. Jpn.* **65**, 2599–2603 (1992).
21. K. Griebenow and A.M. Klibanov, *J. Am. Chem. Soc.* **118**, 11695–11698 (1996).
22. R.A. Sheldon, in *Enzymatic Reactions in Organic Media*, A.M.P. Koskinen and A.M. Klibanov eds., Blackie Academic & Professional, London, 1996, pp. 266–307.
23. J.L. Schmitke, C.R. Wescott, and A.M. Klibanov, *J. Am. Chem. Soc.* **118**, 3360–3365 (1996).
24. J.F. Carpenter, S.J. Prestrelski, and T. Arakawa, *Arch. Biochem. Biophys.* **303**, 456–464 (1993).
25. J.F. Carpenter, S.J. Prestrelski, T.J. Anchordoguy, and T. Arakawa, in *Formulation and Delivery of Proteins and Peptides*, American Chemical Society Symposium Series 567, J.L. Cleland and R. Langer, eds., American Chemical Society, Washington, D.C., 1994, pp. 134–147.
26. U.R. Desai, J.J. Osterhout, and A.M. Klibanov, *J. Am. Chem. Soc.* **116**, 9420–9422 (1994).
27. T.J. Anchordoguy and J.F. Carpenter, *Arch. Biochem. Biophys.* **332**, 231–238 (1996).
28. K. Xu and A.M. Klibanov, *J. Am. Chem. Soc.* **118**, 9815–9819 (1996).
29. K. Griebenow and A.M. Klibanov, *Proc. Nat. Acad. Sci. U.S.A.* **92**, 10969–10976 (1995).
30. Yu.L. Khmelnitsky, S.H. Welch, D.S. Clark, and J.S. Dordick, *J. Am. Chem. Soc.* **116**, 2647–2648 (1994).
31. A.L. Margolin, *Trends Biotechnol.* **14**, 223–230 (1996).
32. V.M. Paradkar and J.S. Dordick, *J. Am. Chem. Soc.* **116**, 5009–5010 (1994).
33. P.A. Burke, R.G. Griffin, and A.M. Klibanov, *J. Biol. Chem.* **267**, 20057–20064 (1992).
34. L.A.S. Gorman and J.S. Dordick, *Biotechnol. Bioeng.* **39**, 392–397 (1992).
35. Y.-J. Zheng and R.L. Ornstein, *J. Am. Chem. Soc.* **118**, 4175–4180 (1996).
36. G.K. Farber, *Ann. N. Y. Acad. Sci.* **799**, 85–89 (1996).
37. U.R. Desai and A.M. Klibanov, *J. Am. Chem. Soc.* **117**, 3940–3945 (1995).
38. A. Dong, J.D. Meyer, B.S. Kendrick, M.C. Manning, and J.F. Carpenter, *Arch. Biochem. Biophys.* **334**, 406–411 (1996).
39. F. Xu, A. Wang, J.B. Vaughn, and T.A. Cross, *J. Am. Chem. Soc.* **118**, 9176–9177 (1996).
40. O. Almarsson and A.M. Klibanov, *Biotechnol. Bioeng.* **49**, 87–92 (1996).
41. A. Zaks and A.M. Klibanov, *Science* **224**, 1249–1251 (1984).
42. A. Zaks, in *Enzymatic Reactions in Organic Media*, A.M.P. Koskinen and A.M. Klibanov eds., Blackie Academic & Professional, London, 1996, pp. 70–93.
43. D.B. Volkin, A. Staubli, R. Langer, and A.M. Klibanov, *Biotechnol. Bioeng.* **37**, 843–853 (1991).
44. N.A. Turner, D.B. Duchateau, and E.N. Vulfson, *Biotechnol. Lett.* **17**, 371–376 (1995).
45. R. Affleck, Z.-F. Xu, V. Suzawa, K. Focht, D.S. Clark, and J.S. Dordick, *Proc. Nat. Acad. Sci. U.S.A.* **89**, 1100–1104 (1992).
46. P.P. Wangikar, P.C. Michels, D.S. Clark, and J.S. Dordick, *J. Am. Chem. Soc.* **119**, 70–76 (1997).
47. G. Greco, Jr., D. Pirozzi, G. Toscano, and M. Maremonti, *Ann. N. Y. Acad. Sci.* **799**, 108–114 (1996).
48. B. Schulze and A.M. Klibanov, *Biotechnol. Bioeng.* **38**, 1001–1006 (1991).

49. A.O. Triantafyllou, E. Wehtje, P. Adlercreutz, and B. Mattiasson, *Biotechnol. Bioeng.* **45**, 406–414 (1995).

See also ENZYMES, DETERGENT; ENZYMES, DIRECTED EVOLUTION; ENZYMES, EXTREMELY THERMOSTABLE; ENZYMES, PULP AND PAPER; PEPTIDE; PROTEIN AGGREGATION, DENATURATION.

DENATURATION, THERMAL.

See THERMAL UNFOLDING, PROTEINS.

DEXTRAN, MICROBIAL PRODUCTION METHODS

R.S. KARTHIKEYAN
T. SWAMINATHAN
A. BARADARAJAN
Indian Institute of Technology–Madras
Chennai, India

KEY WORDS

Biosynthesis
Dextransucrase
Exopolysaccharide
Immobilization
Leuconostoc mesenteroides
Rheology

OUTLINE

INTRODUCTION

Over the past 25 years a new class of microbial products, the microbial exopolysaccharides, has grown in industrial importance. These products can be used as alternatives to other synthetic and natural water-soluble polymers that are used as emulsifiers, stabilizers, binders, coagulants, lubricants, film formers, and thickening, gelling, and suspending agents. Despite their costly collection and extraction, plant and algal polysaccharides such as gum guar, locust bean gum, starch, alginate, carageenan, and agar are widely used in the food and pharmaceutical industries. These polymers, although valuable, have certain drawbacks such as rheological properties that may not match exactly with a specific requirement, production that is vulnerable to crop failure, climatic conditions and marine pollution that cause the quality, supply, and cost to fluctuate (1,2).

The production of microbial exopolysaccharides provides a valid alternative to plant and algal polysaccharides. The exopolysaccharides, produced by a wide variety of microorganisms, show considerable diversity in their composition, structure, and unique physical properties. Their industrial production has been stimulated by the opportunity to provide a guaranteed supply, constant quality, superior properties, and stable price. Among all of the exopolysaccharides, dextran exemplifies attainment of these objectives.

Dextran produced from a bacterial strain of *Leuconostoc mesenteroides* NRRL-B512F has a unique value in medical applications as a blood plasma substitute and a blood flow improver, properties that have provided the impetus for initiating much work on its production and further applications (3). Dextran is a microbial exopolysaccharide made up of α-D-glucose units linked mainly by α, 1 → 6 bonds. Traditionally, dextran is produced by fermenting dextran-producing bacteria in sucrose-rich media (4). The prime bacterial product of fermentation in the presence of substrate sucrose is not dextran itself but the enzyme dextransucrase. The enzyme is secreted by growing cells into the medium, where it converts the sucrose to dextran according to the following equation:

$$n\ \text{Sucrose} \xrightarrow{\text{Dextransucrase}} \underset{\text{Dextran}}{(\text{Glucose})_n} + n\ \text{Fructose} \quad (1)$$

Therefore, dextran can be produced from sucrose in cell-

free environments by using the enzyme dextransucrase (5). The main dextransucrase-producing bacteria belongs to the genera *Leuconostoc* and *Streptococcus*. The enzyme dextransucrase uses sucrose as a high-energy glucopyranosyl donor for dextran synthesis, and the properties of the enzyme and the dextran produced are specific to the microorganisms used (6).

In addition to sucrose, α-D-glucopyransosyl, α-L-sorbofuranoside (7), lactulosucrose (8), α-D-glucopyranosyl fluoride (9), and *P*-nirophenyl α-D-glucopyranoside (10) also serve as glucopyranosyl donors, with sucrose having the lower Michaelis constant. These substrates have glucosidic bond energies of 5.0–5.5 kcal/mol. Hehre and Hamilton (11) and Barker et al. (12) reported that *Acetobactor capsulatum* and *A. viscosum* can transform amylodextrins to dextran. A chemical synthesis of dextran has been reported by Ruckel and Schuerch (13).

Dextran produced from microbial fermentation is called native dextran and typically has a molecular weight of 5×10^7 to 5×10^9. Native dextran is usually separated from the producing cells and precipitated by alcohol. This native dextran may then be subjected to controlled acid hydrolysis to obtain the desired molecular weight for specific applications (14,15).

HISTORICAL BACKGROUND

Dextran first came under close examination more than 100 years ago. Dextrans develop naturally in sucrose-containing solutions that have become inoculated with dextran-producing bacteria from air, plants, or soil. The resulting transformation of sucrose solution to syrupy, viscous, or ropey fluids or even gelled masses doubtlessly has plagued humankind since the inception of accumulating and storing sucrose-containing food and beverages and has caused much trouble in sugar beet industries.

In 1861 Louis Pasteur (16) identified the viscous polymer as a polysaccharide. In 1874 Scheubler (17) determined its empirical formula to be $C_6H_{10}O_5$ and named it dextran because it was closely related to starch and dextrin. Van Tieghem (18) named the causative bacteria *Leuconostoc mesenterides*. Cell-free enzymatic synthesis of dextran was first demonstrated by Hehre and Sugg (5,19). The enzyme involved in this synthesis was given the name dextransucrase by Hestrin and Shapiro-Avineri (20), and it was soon found to be a member of the general class of transglycosylases (21–23). In 1937 Flower and coworkers made the first systematic investigation of the chemical structure of dextran (24). A large body of work carried out between 1950 and 1970 on dextran production and its structure, properties, and uses is reviewed by Neely (25), Jeanes (4,26), Murphy and Whistlkr (27), and and Sidebotham (28).

CLASSIFICATION AND STRUCTURE OF DEXTRANS

The term *dextran* collectively describes a large class of bacterial extracellular hydrocolloid homopolysaccharides consisting of α-D-glucopyranosyl units polymerized predominantly in an α, $1 \rightarrow 6$ linkage. The diversity within the class of dextran molecules derives from the existence of non–α, $1 \rightarrow 6$ linkages that lead to branching from the main polymer chain. Branching mainly occurs by α, $1 \rightarrow 3$, α, $1 \rightarrow 4$, and α, $1 \rightarrow 2$ linkages. The type, amount, length, and arrangements of branches define the differences in dextran structure and influence the water solubility and other rheological properties of dextran (29).

Hucker and Pederson (30) classified the dextran-producing microorganisms into family, tribe, genus, and species. In 1954 Jeanes et al. (6) reported the classification and characterization of dextrans by examining 96 strains of bacteria. First they classified dextrans into three groups: water-soluble dextrans, water-insoluble dextrans, and heterogeneous dextrans. Water-soluble dextrans are further divided into less-soluble dextrans (can be precipitated by 34–37% of alcohol) and more-soluble dextrans (can be precipitated by 40–44% of alcohol). Different kinds of dextrans produced by the same strain of bacteria are called heterogeneous dextrans, and these mixtures of heterogeneous dextrans can be separated by differential alcohol precipitation (31). Further, Jeanes et al. classified the dextrans into class A, class B, and class C dextrans on the basis of the percentage of their α, $1 \rightarrow 3$ linkages. They characterized the dextrans by optical relation, viscosity, periodate oxidation, and physical appearance after alcohol precipitation. Dextran structures have two common features:

1. Main chain and branching chain (side chain) are made by α-D-glucose units linked together by α, $1 \rightarrow 6$ bonds.
2. Branching chains (or branching units) are connected to the main chain by α, $1 \rightarrow 3$, α, $1 \rightarrow 4$, or α, $1 \rightarrow 2$ bonds, called branching linkages.

Examples of α, $1 \rightarrow 3$ Branch–Linked Dextrans

Commercially produced *L. mesenteroides* B-512F dextran has 95% of the total glucouspyranosyl linkages as α, $1 \rightarrow 6$ linkages and 5% as α, $1 \rightarrow 3$ linkages (32,33) (structure A in Fig. 1). The α, $1 \rightarrow 3$ branching linkages occur in two ways:

1. A single α-D-glucose unit attached to the main chain by an α, $1 \rightarrow 3$ bond
2. A branching chain attached to the main chain by an α, $1 \rightarrow 3$ bond (see Fig. 1)

This dextran appears as a translucent gel. Senti et al. reported that 77% of the α, $(1 \rightarrow 3)$ linkages are single glucose residues and the remaining branch chains have 50 or more glucose residues (34). Walker and Pulkownik, studying the hydrolysis of B512F dextran by dextranglucosidase, also concluded that it has 85% single glucose branches and 33 or more glucose residues in the remaining branch chains to give the yield of glucose that they observed (35).

Streptococcus mutans 6715 produces water-soluble and water-insoluble dextrans. The water-soluble fraction has 35% α, $1 \rightarrow 3$ linkages and 65% α, $1 \rightarrow 6$ linkages. Branching single glucose units are uniformly connected to the main chain and to extended side chains (structure B in Fig. 1). The resulting structure was named an alternating bifur-

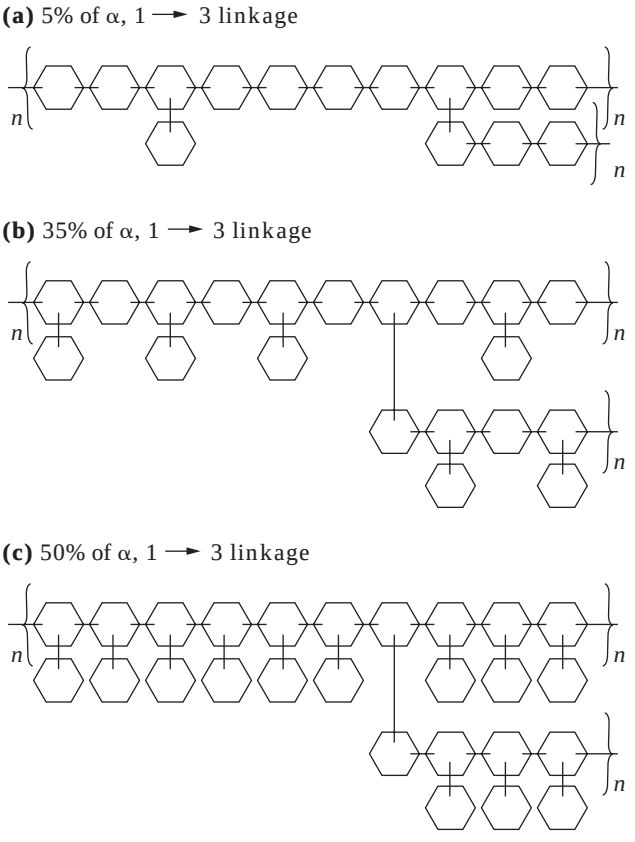


Figure 1. Structures of dextrans: Hexagonal units represent α-D-glucose units, horizontal lines joining glucose units are α, 1 → 6 bonds, vertical lines joining glucose units are α, 1 → 3 bonds: (**a**) clinical dextran, (**b**) alternating bifurcated comb structure, and (**c**) bifurcated comb structure. (Not drawn to statistically accurate scale.)

cated comb polymer by Robyt et al. (36,37), is resistant to hydrolysis by endodextranase, and appears as a heavy, opaque material (38).

L. mesenteroides B-742 strain produced more-soluble and less-soluble dextrans. The more-soluble fraction of dextrans has 50% α, 1 → 6 linkages and 50% α, 1 → 3 branching linkages. Some branching chains are attached to the main chain by α, 1 → 3 branching linkages, but single glucose units are attached to all other glucose units in the main chain by α, 1 → 3 bonds. This structure results in the most highly branched dextrans. Robyt et al. (36,37) named this structure a (structure C in Fig. 1), in which each single α, 1 → 3 linked glucose branch is like a tooth of a comb and the α, 1 → 6 linked chains are the backbone of the comb, bifurcated comb polymer (39,40). This structure is slightly resistant to hydrolysis by endodextranase. This dextran appears as a finely divided material. The α, 1 → 4 and α, 1 → 2 branched dextrans are less common (39–44).

CHARACTERISTICS OF DEXTRANSUCRASE

The unit of dextransucrase (EC 2.4.1.5, α, 1 → 6-D-glucan: D-fructose 2-glucosyltransferase) activity, DSU, is defined as the amount of enzyme that will convert 1 mg of sucrose to dextran in 1 h at 30 °C in acetate buffer solution at pH 5.2. Another unit, U, is defined as the amount of enzyme that liberates 1 μmol of reducing sugar from sucrose in 1 min at pH 5.2 and 30 °C. The relation between these two units is 1 U/mL = 20.27 DSU/mL (14).

The extensive purification of dextransucrase was first reported by Ebert and Schenk, and the molecular weight of the enzyme was determined to be 280,000 by ultracentrifugation methods (45). In 1980 Kobayashi and Matsuda separated two components of dextransucrase having different affinities for dextran gel by a column of Sephadex G-100. The major component of the purified enzyme had a molecular weight of 64,000–65,000, an isoelectric point (pI) of 4.1, and was 17% carbohydrates (46). Most of the carbohydrates in the purified enzyme were found to be endogenous dextran, covalently bound to the enzyme molecule and released from it by mild treatment with endodextranase. They concluded that dextransucrase exists in multiple forms and that endogenous dextran molecules might be responsible for the multiple forms (47). Miller et al. (48) and Fu and Robyt (49) found that the molecular weight of dextransucrase produced from a mutant strain of *Leuconostoc mesenteroides* B-512 FM initially secreted into the culture medium was 177,000; later, however, the enzyme exhibited a drop in molecular weight to 158,000 without any loss in activity.

Mode of Dextransucrase Action

The irreversible reaction of dextransucrase with sucrose is simply formulated in equation 1. Dextransucrase is one of a very few synthesizing enzymes that do not require the presence of cofactors or high-energy phosphorylated intermediates. The energy required for the condensation of two glucosyl units is provided by the hydrolysis of sucrose, distinguishing it from other mechanisms of polysaccharide synthesis (36,37).

Mechanism of Main Chain Synthesis. When dextransucrase was first described in 1941 by Hehre (5), Cori and Cori were studying the action of muscle phosphorylase (50) and Hanes was studying potato phosphorylase (51). These investigators observed that phosphorylase could elongate glycogen and starch chains by the transfer of glucose from α-D-glucopyranosyl phosphate (Glc-1-P) to the nonreducing end (C_6 moiety of glucose) of the polysaccharide chain (52) according to the following equation:

$$\begin{bmatrix}\text{Polysaccharide}\\ \text{chain (primer)}\end{bmatrix} + \begin{bmatrix}\alpha\text{-D-glucopyranosyl}\\ \text{phosphate (Glc-1-P)}\end{bmatrix} \xrightarrow{\text{Phosphorylase}} \begin{bmatrix}\text{One glucose unit added}\\ \text{to the polysacchardie chain}\end{bmatrix} \quad (2)$$

It was believed that the enzyme phosphorylase that catalyzed the preceding reaction requires a preformed polysaccharide chain, called a primer, and a high-energy monosaccharide donor, Glc-1-P. Later it was proven that the enzyme phosphorylase was a degradative enzyme that catalyzed the reaction of inorganic phosphate with the non-

reducing end of the polysaccharide chain to give Glc-1-P according to the following equation:

$$\begin{bmatrix}\text{Polysaccharide}\\ \text{chain}\end{bmatrix} + \begin{bmatrix}\text{Inorganic}\\ \text{phosphate}\end{bmatrix} \overset{\text{Phosphorylase}}{\rightarrow} [\text{Glc-1-P}] + \begin{bmatrix}\text{One glucose unit reduced in the}\\ \text{polysaccharide chain (primer)}\end{bmatrix} \quad (3)$$

The reverse of the preceding reaction, which requires a primer (a polysaccharide chain), produces a polysaccharide chain by adding one glucose monomer to the primer at the nonreducing end (C_6 moiety). Thus in the 1940s and 1950s the primer mechanism was assumed for dextran synthesis (53,54).

Evidence for the primer-dependent mechanism for dextran synthesis was strengthened by Germaine and coworkers in the 1970s. They found that addition of dextran (as a primer) to dextransucrase increased the rate of dextran synthesis and that the rate reached a maximum at an average size of the added dextran of 30 glucose residues (55–57). Kobayashi and Matsuda also reported that purified dextransucrase was stimulated by dextran (as a primer), although dextransucrase could synthesize dextran without the addition of primer (46,47). Based on these results, Kobayashi et al. (58) proposed a primer-dependent mechanism for dextran synthesis. They proposed that the active site has two types of substrate binding sites, one for sucrose and the other for primer dextran. The elongation of the primer was hypothesized to take place when an enzyme nucleophile (presumably a carboxylate anion) attacks C_1 of the glucose moiety to give the glucosyl–enzyme covalent intermediate. This glucose is then transferred to the 6-hydroxyl group of the nonreducing end of the primer chain attached to the primer binding site of the enzyme, forming the next α, $1 \rightarrow 6$ linkage. For the synthesis to continue, the primer, with newly added glucose residue, must dissociate from the primer binding site and a new primer molecule must bind in the primer binding site with its nonreducing glucose residue in position to accept the next glucose unit. The synthesis in this scheme is thus a discontinuous process.

However, Robyt and Corrigan found that a dextran (as a primer) modified by a blocking group (triisoproprylbenzenesulfonyl, tripsyl) on the C_6 position of the nonreducing end of the glucose residue was as effective in increasing the rate of dextran synthesis by dextransucrase as the unmodified dextran primer. The modified dextran could not act as a primer because its sixth position was blocked, proving that the added modified dextran stimulated synthesis by some other mechanism (59). Robyt et al. have shown that the activation of dextransucrase by addition of dextran is by an allosteric mechanism rather than by an extending primer mechanism (37) as distinguished from the chain elongation of glycogen and starch by phosphorylase.

In 1962 Ebert and Patat proposed another mechanism for dextran synthesis, insertion type polymerization, in which the addition of glucose units takes place at the reducing end (C_1 of glucose moiety) of the growing dextran polymer (60). In their mechanism they assumed that the growing chain and the enzyme remain in contact in the form of a complex throughout the chain elongation cycle. This enzyme–polymer complex (EP_n) reacts in the first step with substrate to form a substrate complex $(ESP_n)_2$, which is transformed in a second step in such a way that the polymer chain in the reaction complex is extended by one unit. This can only happen if the new monomer unit is inserted into the covalent bond between the enzyme and the polymer chain in the EP_n complex.

$$EP_n + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ESP_n \quad (4)$$

$$ESP_n \xrightarrow{k_2} EP_{n+1} + F \quad (5)$$

(F: fructose)

The bond between the enzyme and polymer chain does not dissociate during the reaction but must be flexible enough so that the new unit can be inserted into it. From their kinetic studies, Ebert and Schenk (45) assumed that the substrate undergoes a second reaction with the EP_n complex in which it reacts as an acceptor.

$$ESP_n + S^* \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} S^*ESP_n \rightarrow ES + S^*P_n \quad (6)$$

In the preceding reaction, the acceptor (sucrose molecule acting as an acceptor identified by the asterisk) does not compete with the substrate for the enzyme but reacts with the intermediate ESP_n complex, terminating polymer extension. This would explain why the molecular weights of the products decrease with increasing substrate concentration. From the preceding mechanism Ebert and Patat proposed the following reaction rate expressions:

$$v = \frac{V_m[S]}{K_m + [S]} \cdot \frac{K_a}{K_a + [S]} \quad (7)$$

where V_m is the maximum reaction rate (k_2E_0), K_m is the Michaelis constant of the chain extension reaction, and K_a is the corresponding constant of acceptor reaction. Ebert and Schenk conclude that there is strong evidence for their proposed mechanism because the results of their kinetic studies on dextran formation match the proposed rate in equation 7 quite well. They further conclude that primers are not necessary to initiate polymer growth, and sucrose alone can initiate polymerization (45).

The early theories of the mechanisms of dextran synthesis were described by Neely (61). Ebert and Schenk propose the sequential addition of monomer units to the reducing end by insertion between a carrier (the enzyme) and growing polysaccharide chain (45). Sidebotham (28) continued the discussion in 1974, and a series of papers by Robyt and coworkers appears to have conclusively resolved the mechanism (59,62,63).

In 1974 Robyt and coworkers proposed a two-site insertion mechanism for dextran synthesis in which the glucose unit is added to the growing polymer chain at the reducing end (C_1 moiety of glucose). They used pulse and chase techniques with ^{14}C-labeled sucrose and dextransucrase immobilized in Bio-Gel-P2 (64). In their two-site insertion

mechanism (see Fig. 2), there are two sucrose binding sites (A and B) that attack two sucrose molecules to give two covalently attached glucosyl–enzyme intermediates (C). The C_6 hydroxyl of one of these glucosyl–enzyme intermediates makes a nucleophilic attack on C_1 of the other glucosyl–enzyme intermediate with the formation of an α, $1 \rightarrow 6$ glucosidic linkage and results in an isomaltosyl–enzyme intermediate (D). The vacant sucrose binding site (A) attacks another sucrose molecule to form a new glucosyl–enzyme intermediate to give a second α, $1 \rightarrow 6$ linkage and the formation of the isomaltotriosyl–enzyme intermediate (E). This process continues in a similar fashion by inserting a glucose residue between the enzyme and growing chain, which results in a dextranyl–enzyme complex (F). In this mechanism the growing dextran chain is transferred from one site to another during the synthesis. Robyt and Eklund (65,66) considered the stereochemistry of the reaction and concluded that the linkages of glucosyl and dextranyl units to the enzyme must be β to retain the configuration of the glucose residue in growing from sucrose to dextran.

Mechanism of Branch Linkage Formation. In 1959 Bovey investigated the formation of branching in dextrans using data on molecular weights obtained from light-scattering measurements. He proposed that a separate branching enzyme (acting in a way similar to Q enzyme on amylose) is responsible for the formation of branches (67). The branching enzyme, however, has never been found. In 1967 Ebert and Brosche proposed a reaction mechanism for the formation of branches in which a dextran chain itself acts as an acceptor, attacking an enzyme–dextran complex (ESP_n) (see equations 4–6). The acceptor dextran then becomes the main chain, and the dextran chain from the enzyme becomes the side-branched chain. Using a ^{3}H-labeled acceptor dextran of low molecular weight and assuming an average molecular weight of 4×10^5 for synthesized dextran, they concluded from the specific activity of synthesized product that there was only one labeled acceptor dextran molecule in each synthesized dextran (68). Although this seemed to be proof for the proposed mechanism, some doubt was cast on the mechanism because of assumptions and the circular arguments that were made.

In 1976 Robyt and Taniguchi (63) proposed a mechanism based on their study using ^{14}C-labeled sucrose and dextransucrase immobilized in Bio-Gel-P2. In their mechanism (see Fig. 3), a C_3 hydroxyl group of an interior glucose unit of an acceptor dextran (D) makes a nucleophilic attack at C_1 of either the glucosyl–enzyme complex (C) or at C_1 of the dextranyl–enzyme complex (E) to form an α, $1 \rightarrow 3$ branch linkage by displacing glucose or dextran from the enzyme. Thus, branching can take place without a separate enzyme by the action of an acceptor dextran on a glucosyl–enzyme complex (C) to give single glucose unit branch linkage (F) and on a dextranyl–enzyme complex (E) to give a long side chain branch linkage (G) (36,66).

Acceptor Reactions and Acceptor Products

In addition to catalyzing the synthesis of dextran from sucrose, dextransucrase also catalyzes the transfer of the glucose moiety from sucrose to other carbohydrates that are present or added to the reaction mixture (69). These other carbohydrates are acceptors (see Table 1). Actually, there are two classes of acceptors—those that give a homologous series of oligosaccharides, each differing one from the other by one glucose residue, and those that form only a single acceptor product that contains one glucose residue more than the acceptor (36,66).

Koepsell et al. (53) and Tsuchiya et al. (54) first observed that the presence of low molecular weight carbohydrates (acceptors) shifted the course of dextransucrase reaction from synthesizing a high molecular weight dextran to synthesizing of a low molecular weight dextran. These acceptors not only terminate the dextran synthesis but also produce a novel series of oligosaccharides. Su and Robyt systematically studied maltose acceptor reactions with dextransucrase (70). In this maltose acceptor reaction the D-glucose moiety of sucrose is diverted from synthesis of dextran but transferred to the nonreducing end of maltose (a disaccharide) to form panose (a trisaccharide). This first acceptor product, panose, can itself act as an acceptor to give a tetrasaccharide. This tetrasaccharide can again act as an acceptor to give a pentasaccharide and so forth. This series of homologous acceptor products is produced by acceptor reactions.

Mechanism of Acceptor Reaction. As described earlier, dextransucrase catalyzes the synthesis of dextran by initially forming two covalent glucosyl–enzyme complexes (Fig. 2) in which the C_6 hydroxyl moiety of one glucose attacks the C_1 of the other to form an α, $1 \rightarrow 6$ glucosidic bond to produce an isomaltosyl–enzyme complex. This is the beginning of the dextran main chain growth; the dextran chain is covalently attached to the enzyme active site and elongated by the insertion mechanism (36,66). In the presence of an acceptor, such as maltose, the transfer of glucose into a growing dextran chain is interrupted as the maltose binds to an acceptor binding site.

Robyt and Walseth proposed a mechanism for the acceptor reaction (see Fig. 4). In this mechanism, the acceptor bound at an acceptor binding site, and if the maltose molecule is an acceptor, its C_6 hydroxyl group at the nonreducing end attacks C_1 of the glucosyl–enzyme (C) or dextranyl–enzyme (E) complex to give an oligosaccharide (F) or a dextran acceptor product (G), respectively (62). Su and Robyt, using maltose as an acceptor in an equilibrium dialysis experiment, showed that there was one acceptor binding site and two sucrose binding sites (A and B) (70).

Effect of Acceptor on Dextran Production. The acceptors affect dextran production in two ways:

1. They terminate the growth of the dextran chain by reacting with the dextranyl–enzyme complex, which results in the low molecular weight dextran (product G in Fig. 4).
2. They interfere with the transfer of the glucose moiety of sucrose to the growing dextran chain by reacting with the glucosyl–enzyme complex, which results in production of a homologous series of acceptor

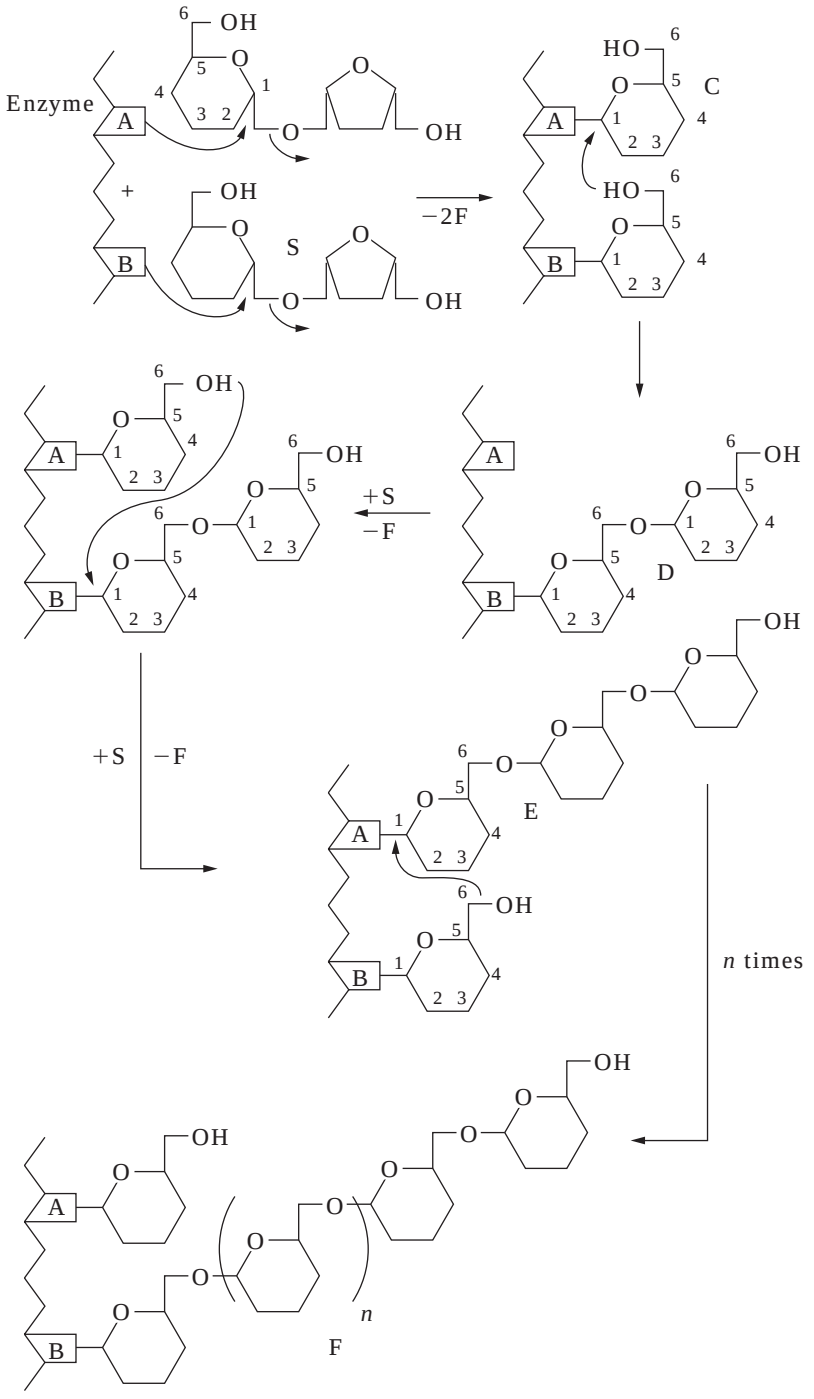


Figure 2. Two-site insertion mechanism: A and B, enzyme active sites; S, sucrose; F, fructose; C, glucosyl–enzyme complex; D, isomaltosyl–enzyme complex; E, isomaltotriosyl–enzyme complex; F, dextranyl–enzyme complex (—OH groups are omitted for clarity.)

products and reduction in the amount of dextran produced (product F in Fig. 4).

Paul et al. demonstrated the effect of the sucrose-to-acceptor ratio in the synthesis of controlled molecular weight dextrans (71). Robyt and Eklund showed that the amount of dextran synthesized decreased as the molecular ratio of maltose to sucrose increased (69). Not all of the acceptors reacted with equal efficiency. In a series of reactions with different acceptors at a 1:1 ratio of acceptor concentration to sucrose concentration, they found that the most effective acceptor for decreasing the amount of dextran synthesized was maltose and that the amount of glucose residues diverted from dextran synthesis into acceptor products depended on the ratio of maltose to sucrose used. The efficiency of other acceptors is given in Table 1, where maltose efficiency is defind as 100% (66). Su and Robyt found that an increase in the ratio of maltose to sucrose increased the amount (concentration) of acceptor products, with a concomitant decrease in dextran synthesis. Further, as the ratio was increased, the number of acceptor products (e.g., tri-, tetra-, and pentasaccharides) de-

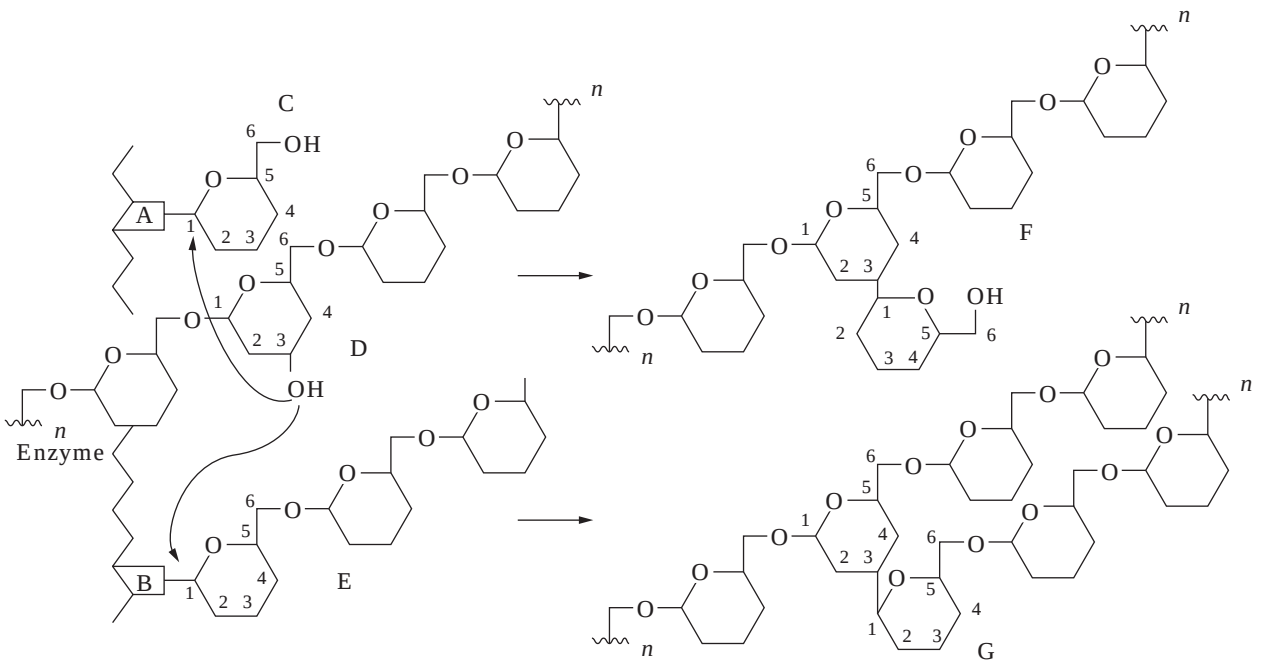


Figure 3. Mechanism of α, 1 → 3 branch linkage synthesis: D, acceptor dextran main chain; C, glucosyl–enzyme complex; E, dextranyl–enzyme complex; F, resulting α, 1 → 3 branch–linked product when C and D are reacted together; G, resulting α, 1 → 3 branch–linked product when E and D are reacted together. (—OH groups are omitted for clarity.)

Table 1. Acceptors and Their Efficiency on Dextransucrase

Acceptors[a]	Efficiency (%)
Maltose	100
Isomaltose	89
Nigerose	58
Methyl α-D-glucopyranoside	52
D-Glucose	17
Lactose	11
Cellobiose	9
D-Fructose	6.4
Maltooctaose	6.2

[a]The products from these acceptors can also act as acceptors to produce a homologous series of acceptor products.

creased. When the concentration of maltose and sucrose increased and the ratio was maintained at 1:1, there also a decrease in the amount of dextran and an increase in the amount (concentration) of acceptor products observed. In addition, the amount of dextran decreased and the amount and number of acceptor products increased when the amount of enzyme increased. From these studies, they reported that the first acceptor product (panose) can be exclusively obtained without formation of dextran by using a specific ratio and concentration of maltose and sucrose and a specific amount of enzyme (70).

Part et al. reported on the effect of borate ions on the acceptor reaction products of dextransucrase. They explained that the fructose released during dextran synthesis reacts as an acceptor with a glucosyl–enzyme complex to form fructose acceptor products (e.g., leucrose and isomaltulose), which in turn reduce the dextran yield. This fructose can be complexed with borate ions, and its acceptor effects can be minimized to increase dextran yield. They concluded that with 0.115 M sodium tetraborate in the reaction mixture, leucrose and isomaltulose formation can be minimized up to 75% with consequent increase in dextran yield (72). Therefore, the acceptor reaction between fructose and the dextranyl–enzyme complex, which leads to diminished low molecular weight dextran, can be overcome by adding borate ions to the reaction medium.

Lee et al. (73) used the idea of producing low molecular weight dextran by adding an acceptor to the reaction medium to directly produce clinical dextran (MW 70,000). They studied the effect of dextran T-10, dextran T-40, and maltose as acceptors in the batch and continuous reactors. They reported that increasing the concentration of dextran T-10 reduces the production of high molecular weight dextran and increases the production of clinical dextran, whereas increasing the dextran T-40 concentration decreases the production of clinical dextran. Maltose reduces high molecular weight dextran production but increases the production of acceptor products, thereby reducing clinical dextran production (73).

Activation, Inhibition, and Stability of Dextransucrase

Germaine et al. observed that addition of dextran to the dextransucrase (produced from the *Streptococcus mutans* reaction mixture) increased the rate of dextran synthesis. They found that the rate depended on the size of the dextran chain and reached a maximum when the average size

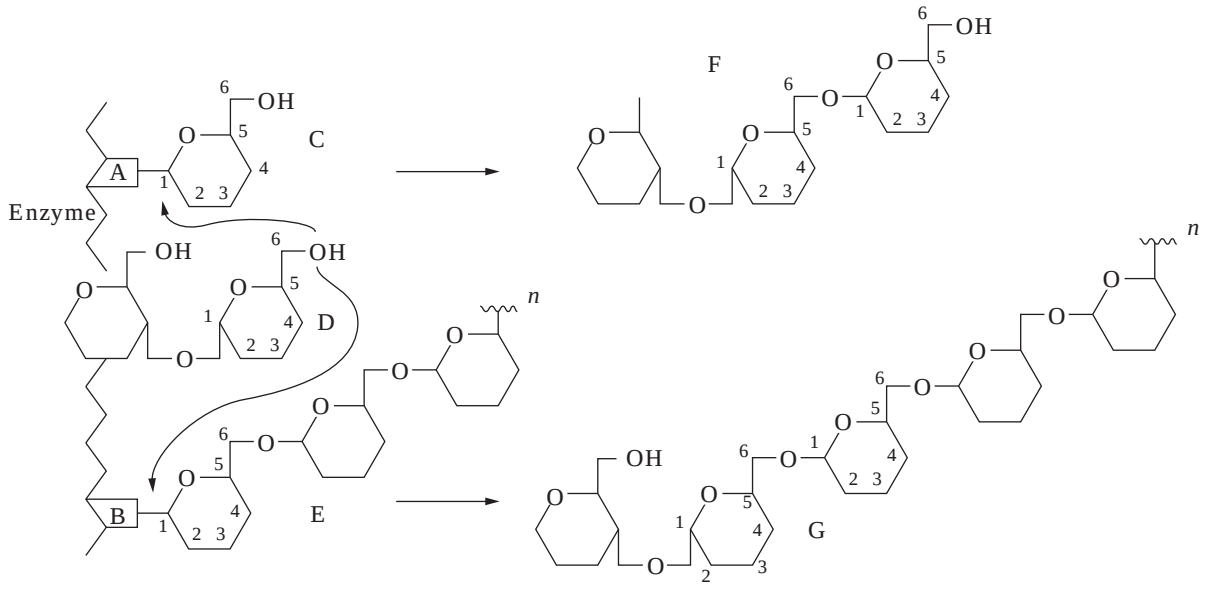


Figure 4. Mechanism of the acceptor reactions: D, maltose molecule bound at acceptor binding site; C, glucosyl–enzyme complex; E, dextranyl–enzyme complex; F, resulting acceptor product when C and D are reacted together; G, resulting acceptor product when E and D are reacted together. (—OH groups are omitted for clarity.)

of the added dextran was 30 glucose residues (55–57). Kobayashi and Matsuda also observed that purified dextransucrase was stimulated by addition of dextran, although dextransucrase could synthesize dextran without addition of dextran to the reaction mixture. The rate of dextran synthesis in dextran-free reaction mixtures was accompanied by a lag period that could be eliminated by addition of exogenous dextran (58). Robyt et al. reported that the addition of dextran to dextransucrase reaction mixtures affects the initial velocity of dextran synthesis in a sigmoidal manner characteristic of an allosteric binding of dextran to the enzyme. Further, they proposed that the allosteric activation of dextransucrase occurs when the dextran binds to the dextran binding site, where it acts as a matrix to join or induce the subunits to associate into an active form (37). Kobayashi and Matsuda reported that purified B-512 F dextransucrase contains 17% carbohydrates, mostly as endogenous dextran covalently bound to the enzyme, which can be eliminated by the mild action of endodextranase. They suggested that exogenously added dextran could stabilize the enzyme not only during storage but also during long-term incubation (46,47).

Miller and Robyt (74), from their rate kinetics of dextran synthesis, showed that concentrations of Ca^{2+} ion below 1 mM activate the enzyme by increasing V_m and decreasing the K_m for the sucrose. Above 1 mM Ca^{2+}, the ions act as weak competitive inhibitors. Although Ca^{2+} is an activator at low concentrations, it is not required for dextran synthesis. Miller and Robyt proposed a model for Ca^{2+} ion activation and inhibition behavior in which the enzyme has two Ca^{2+} binding sites. One Ca^{2+} site is an activating site and the other is an inhibitory site at which Ca^{2+} ions prevent the binding of sucrose to the sucrose binding site (48,74). Kobayashi and Matsuda reported that the purified enzyme can be activated by Ca^{2+} ions and strongly inhibited by EDTA. The inhibitory effect of EDTA can be reduced by adding Ca^{2+} ions (75,76). However, they further observed that the purified enzyme was strongly inhibited by Cu^{2+}, Fe^{2+}, and Mn^{2+} ions in this order (46). Contrary to the crude enzyme, which contains endogenous dextran, the purified enzyme shows optimum activity at pH 6.0 (46). Miller and Robyt found that the dextransucrase could be stabilized against the activity losses caused by heating or by dilution through the addition of low concentrations nonionic polymers (e.g., dextran, PEG 20000, and methyl cellulose) or of nonionic detergents at or slightly below their critical micelle concentrations (77). The effect of pH and temperature on dextransucrase is discussed in the later section "Effect of pH and Temperature".

Effect of Enzyme Concentration

In 1962 the effect of enzyme concentration on dextran synthesis was reported by Braswell et al. (78). They observed that at higher concentrations of enzyme the molecular weight of dextran synthesized was decreased and the intrinsic viscosity and branching ratio (α, 1 → 6 linkages to α, 1 → 3 linkages) was increased. Tsuchiya et al. made a related observation that increasing the enzyme concentration decreased the amount of high molecular weight dextran synthesis (54).

At the time, neither Braswell and coworkers nor Tsuchiya and coworkers put forward an explanation. However, in 1993 Su and Robyt explained in their report that the results are due to the increased acceptor reactions at higher enzyme concentrations (70) (see the earlier section "Acceptor Reactions and Acceptor Product"). Another important point is that improving the yield of enzyme from a

given fermentation does not necessarily produce an improved yield of dextran because the conditions under which the enzyme is allowed to act on sucrose solution determine the dextran yield (14).

PRODUCTION OF DEXTRANSUCRASE AND DEXTRAN

The dextran fermentation process involves three phases: *(1)* the cell growth phase, *(2)* the dextransucrase production phase, and *(3)* the enzyme action phase. The enzyme action phase that produces dextran is the third phase (i.e., it occurs after cell and enzyme production). Thus, the rate of dextran production under given conditions is a function of the enzyme production rate, which in turn is a function of the cell production rate. Moreover, dextransucrase stability and its activity under given conditions influence dextran synthesis. To optimize the dextran fermentation process, one must consider the following rates of reactions: *(1)* cell growth rate, *(2)* enzyme production rate, *(3)* enzyme deactivation rate (stability), and *(4)* enzyme action rate.

The interdependency of these reaction rates under specific fermentation conditions is a key factor for in situ dextran production. Because each rate of reaction has its own optimum level, a set of conditions given for a fermentation may not be optimal for all four rates of reaction. For example, a temperature of 30 °C is optimal for cell production and enzyme action but not for enzyme production and its stability. The three phases of dextran fermentation can be separated and the optimum level of each phase can be used one by one for maximum dextran yield. Therefore, each phase and its reaction rates must be understood separately and in great detail to employ process control techniques, to find optimal conditions, and to design a high-performance reactor.

Factors That Influence Dextran Fermentation

Bacterial Strains. The genera of bacteria recognized to produce the enzyme dextransucrase, which is capable of synthesizing dextran from sucrose, are primarily *Leuconostoc* and *Streptococcus*. These two genera are gram-positive, facultative anaerobic cocci that are closely related. The notable difference between these two genera is that *Leuconostoc* species require sucrose in the culture medium to induce the formation of dextransucrase, whereas the *Streptococcus* species do not require sucrose in the culture medium to induce enzyme formation (79,80). Thus, the *Leuconostoc* species are normally inducible for the formation of dextransucrase and the *Streptococcus* species are normally constitutive for their formation. However, Kim and Robyt have obtained mutant strains of *Leuconostoc mesenteroides* B-512 FM, B-742, B-1299, and B-1355 that are constitutive for dextransucrase formation (81).

Effect of Sucrose Concentration. Sucrose is an essential substrate for dextransucrase and dextran synthesis. Although *L. mesenteroides* can grow on various carbohydrates such as glucose, fructose, and maltose, it prefers sucrose over other carbohydrates as a carbon and energy source (82,83). The carbohydrates are fermented through the hexose monophosphate pathway, giving lactic acid, carbon dioxide, and ethanol under anaerobic conditions. An aerobic oxidative metabolism that yields carbon dioxide, lactic acid, and acetic acids also exists (84).

Tsuchiya et al. found in batch culture that increasing sucrose concentration from 5 to 50 g/L resulted in an increase in enzyme productivity from 6 to 120 DSU/mL. However, they recommended 20 g/L sucrose for optimum dextransucrase production because higher sucrose concentrations increase the viscosity of the broth due to dextran synthesis, which makes cell removal difficult. The enzyme activity that they achieved at 20 g/L sucrose is 86 DSU/mL (85). It is essential to maintain the lower level of sucrose concentration (5–10 g/L) throughout fermentation to produce a higher activity of dextransucrase (86). This is possible in semicontinuous or continuous operational mode. Lopaz and Monsan achieved the enzyme activity of 180 DSU/mL by addition of sucrose at the rate of 20 g/L/h over the cell growth phase (87,88). Lawford reported that the enzyme can be synthesized only after a critical concentration of sucrose, the inducer, has been reached or surpassed (82). The highest activities of dextransucrase reported in the literature are by Schneider et al. (86) and Barker and Ajongwen (89). The former obtained 400 DSU/mL in laboratory scale and 289 DSU/mL at pilot plant scale. Barker and Ajongwen obtained activity of 450 DSU/mL under nonaerated fed-batch fermentation and 350 DSU/mL under aerated fed-batch fermentation.

At higher sucrose concentrations, the enzyme is inhibited and the conversion of sucrose to dextran decreases (see Fig. 5). Moreover, the molecular weight of synthesized dextran decreases at higher sucrose concentrations, where sucrose itself acts as an acceptor (see the earlier section "Acceptor Reactions and Acceptor Products") to terminate polymer growth. Braswell et al. observed that increasing the sucrose concentration from 50 to 300 g/L increases the intrinsic viscosity and the branching ratio of dextran produced while decreasing the molecular weight of dextran and the rate of dextran formation (78,91).

Alsop reported that higher dextran yields have been obtained at lower initial sucrose concentrations, reducing costs as a function of sucrose alone. However, the net amount (and so the concentration) of dextran produced

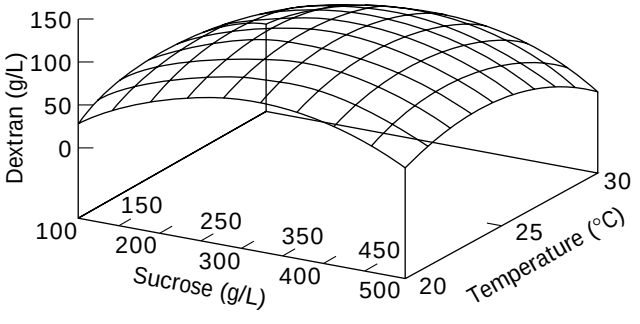


Figure 5. A calculated response surface plot of dextran yield (g/L) as a function of sucrose and temperature with a constant level of yeast extract (10 g/L), (30 g/L) K_2HPO_4, and initial pH 8.3. *Source*: From Ref. 90.

was lower, reducing plant productivity and substantially increasing ethanol and labor costs in downstream processing. Alsop calculated the optimum value on the basis of ethanol precipitation costs (which depend on the molecular weight of the dextran synthesized; high molecular weight dextrans can be precipitated with lower concentrations of alcohol) and reported that the optimum value is 170 g/L of initial sucrose concentration. However, the same author reported another optimal value, 125 g/L of sucrose, based on the operating costs of the plant (14). Martinez-Espindola and Lopez-Munguina reported that in cell-free enzymatic synthesis of dextran, an initial sucrose concentration of 150 g/L is optimum for 99% conversion of sucrose to dextran. The optimum value was based on reaction kinetics of dextransucrase and a specific design for a batch reactor (92).

In our work, we found the optimum value (based on yield) of sucrose to be 300 g/L for batch dextran production in whole cell culture based on response surface methodology (90). We observed a yield of 135 g/L dextran at this optimum level. Further, we observed substrate inhibition kinetics for cell growth at more that 100 g/L sucrose and that at more than 300 g/L sucrose the dextran yield decreases considerably (93,94) (see Fig. 5).

Effect of Nitrogen Source. The nitrogen sources used to grow *L. mesenteroides* and produce enzyme are corn steep liquor, yeast extract, soya meal preparations, and peptone. These complex nutrients also provide essential growth factors and trace elements. Ammonium ions used as the nitrogen source appear to have an adverse effect on cell growth (85). Tsuchiya et al. reported that an increase in the concentration of corn steep liquor as a nitrogen source from 2.5 to 20 g/L increases the enzyme yield but only when higher amounts of phosphate are present in the medium (85).

Barker and Ajongwen, who obtained the highest enzyme activity, used 40 g/L yeast extract as a nitrogen source in fed-batch culture. However, they observed variable performance with different yeast extract types, because different yeast extracts contain varying levels of vital nutrient components (89). In our work, we optimized yeast extract concentration to 10 g/L for conversion of 300 g/L sucrose to dextran (90) and further observed that below 10 g/L yeast extract the growth of cells and enzyme production were affected considerably.

Effect of Phosphate Concentration. Phosphates such as K_2HPO_4 and KH_2PO_4 act as buffers in the fermentation medium in addition to being nutrients. Tsuchiya et al. reported that 35 g/L KH_2PO_4 is optimum for enzyme production (85). We observed that 30 g/L K_2HPO_4 is optimum for batch dextran production to convert 300 g/L sucrose to dextran (90). Lawford et al. (82) reported that larger amounts of phosphate can remove the Ca^{2+} ions in the medium by forming a complex with Ca^{2+} ions. Calcium is an essential element for dextransucrase stability (see the earlier section "Activation, Inhibition, and Stability of Dextransucrase"). They observed loss of activity of dextransucrase during batch culture at 10 g/L potassium phosphate, whereas at 1 g/L phosphate the produced activity remained stable for a long time.

London and Webb optimized the concentration of Na_2HPO_4 to 20 g/L (95,96), and Barker and Ajongwen used K_2HPO_4 at 20 g/L in their fed-batch fermentation for higher enzyme production (89). In batch fermentation without controlled pH the pH of the medium starts to fall as the cell starts to grow. However, the rate of fall can be controlled by using higher phosphate concentrations. Cell growth, enzyme production, and dextran production are faster at pH 7.0, 6.7, and 5.2, respectively (14). The interaction between initial pH and phosphate can be seen in Figure 6 (90). However, in pH-controlled fermentations, the higher amounts of phosphates are not required and the adverse effect of Ca^{2+} ion removal can be avoided. pH control is recommended over a higher phosphate concentration in the culture.

Effect of Growth Factors and Trace Elements. Carlson and Carlson reported that the growth factors nicotinic acid, thiamine, pantothenic acid, and biotin and the amino acids valine and glutamic acid are required for *L. mesenteroides* growth (97–99). However, the complex nitrogen sources (e.g., yeast extract, corn steep liquor, and peptone) are expected to supply those vital nutrients and necessary trace elements. In 1952 a salt solution called R salts that contained 4.0% $MgSO_4 \cdot 7H_2O$, 0.2% NaCl, 0.2% $FeSO_4 \cdot 7H_2O$, and 0.2% $MnSO_4 \cdot H_2O$, was first introduced by Koepsell and Tsuchiya (100). London reported that the trace inorganic salts generally used were not found to be necessary (95,96), suggesting that nutrient requirements are culture dependent. The effect of Ca^{2+} ions on enzyme stability was discussed earlier.

Effect of pH and Temperature. The optimum levels of pH and temperature for cell, enzyme, and dextran production are given in Table 2, and the enzyme dextransucrase is stable at pH 5.2 and low temperature. The effects of pH and temperature on enzyme production and its action and stability are interrelated. Although pH 6.7 is optimum for enzyme production in this example, this value is applicable only at the optimum temperature of 23 °C. At higher temperatures, pH 6.7 inactivates the enzyme rapidly (85).

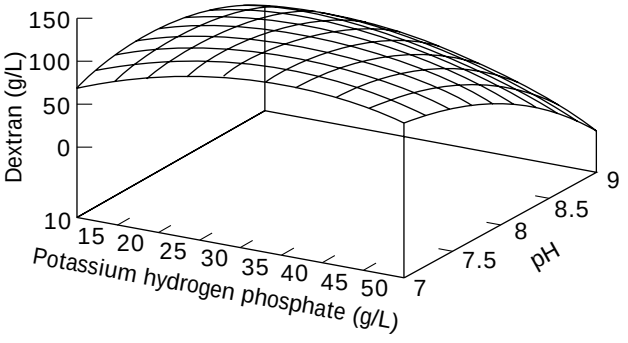


Figure 6. A calculated response surface plot of dextran yield (g/L) as function of initial pH and K_2HPO_4 with a constant level of sucrose (300 g/L) 10 g/L yeast extract, and 23 °C. *Source*: From Ref. 90.

Table 2. Optimum pH and Temperature for Cell Growth, Enzyme Production, and Dextran Production by *Leuconostoc mesenteroides*

Parameter	Cell growth	Enzyme production	Enzyme action
pH	7.0	6.7	5.2
Temperature (°C)	30	23	30

Similarly, the optimal temperature of 30 °C for cell growth and enzyme activity is too high for enzyme production and stability. The enzyme activity in culture filtrate is retained at pH 5.2 for at least 24 h at 25 °C, but 35% of the activity is lost at pH 6.7 within 1 h at 25 °C. Moreover, at higher temperatures the deleterious effect of pH is accentuated markedly. Kobali and Reilly (101) reported that the half-lives of enzyme at pH 4.9 are 13.75, 7.02, and 3.28 min at 30, 32, and 35 °C, respectively. London and Webb observed that the maximum enzyme production rate was achievable at pH 6.1, although the maximum activity was observed at pH 6.7. They further stated that at pH 6.0 and below, the enzyme was relatively stable, retaining 100% of activity for 5 h at pH 5.4 and 26 °C. Above pH 6.0 the enzyme was far less stable, retaining only 10% of its original activity at pH 6.6 and 26 °C (95,96).

The effect of temperature as the velocity of the reaction corresponds to an Arrhenius activation energy of 10.6 kcal/mol (78). However, Kobali and Reilly reported a lesser value of 8.6 kcal/mol of activation energy (101). Dextran synthesis temperature also plays an important role in the properties of the dextran produced. Braswell et al. reported that decreasing the reaction temperature from 30 to 0 °C lowers the molecular weight of dextran and increases the intrinsic viscosity. One would normally expect lesser intrinsic viscosity for lower molecular weight polymers, but a more linear polymer is produced at lower temperatures, increasing the intrinsic viscosity observed (78). Sabatie and Choplin (102) also reported that dextran produced at 5 °C was more linear than dextran produced at 30 °C. They observed 5% α, 1 → 3 branch linkages at 5 °C of synthesis temperature compared with 4.8% at 30 °C (78).

Effect of Aeration and Agitation. The maximum productivity and product concentration achievable in industrial microbial processes for production of biomass as product are typically limited by the rate of oxygen transfer. *L. mesenteroides* is a facultative anaerobe. However, some researchers used the aerobic fermentation for dextransucrase production. Schneider and coworkers observed a higher enzyme activity of 400 DSU/mL under highly aerated conditions in which the oxygen level was maintained at 40–80% of the initial saturated value throughout the course of fermentation (86). However, Barker and Ajongwen observed the higher activity of 450 DSU/mL under nonaerated fed-batch fermentation but achieved only 350 DSU/mL of activity under aerated fermentation (89). Foster also found that aeration gives a lower yield of enzyme (103,104).

Koepsell and Tsuchiya (104) and Tsuchiya et al. (86) noticed that enzyme yields were higher in shaken than in still flasks, and Johnson and McCleskey (105) also observed similar results. This may be due to the mixing requirements for limiting nutrients. London and Webb compared the stirred and unstirred cultures for enzyme production. They observed a lower cell growth rate and higher enzyme production in unstirred culture than in stirred culture. This is due to the limiting substrate concentration around the cells being lower under poor mixing conditions (i.e., the transfer from the liquid to the cells is insufficient to meet their demand, which results in poor cell growth rate). But as the cells produce the enzyme and acid, the enzyme and acid are expected to collect in pockets around the cells in nonagitated conditions. This lowers the pH around the enzyme to its stability optimum (95,96).

As the dextran concentration increases in the fermenter during its synthesis, the viscosity of the broth also increases, which requires proper mixing conditions for better conversion of substrate to polysaccharide and scale-up of polysaccharide fermentation (106). Margaritis and Bokkel studied the rheological behavior of dextrans (low and high molecular weight) in different concentrations (107). Their data fit well the following power law equation, which represents the pseudoplastic behavior of fluid:

$$\log\tau = \log K + n \log\gamma \tag{8}$$

where τ is the shear stress applied on the fluid (N/m^2), K is the consistency index, n is the flow behavior index, and γ is the shear rate (s^{-1}).

From Margaritis and Bokkel's observations, the flow index n of the high molecular weight dextran solution was much less than the flow index for the low molecular weight solution. The consistency index K of high molecular weight dextran solutions was found to be higher than the K values of low molecular weight solutions for the same concentrations (1). These results show that high molecular weight dextran shows a higher degree of pseudoplastic behavior than low molecular weight dextran.

Batch Culture

Traditionally, dextran is produced by anaerobic, unstirred batch culture using dextransucrase-producing bacteria in sucrose-rich media (14). Sugar solutions are fortified with a nitrogen source such as yeast extract or corn steep liquor, a phosphate source such as K_2HPO_4 or KH_2PO_4, and trace metal salts. Little further process control is employed. The pH, which was neutral at the start of the fermentation process, falls as cells grow. Dextransucrase is not produced until near the beginning of the stationary phase. At this point the enzyme is secreted rapidly into the culture medium and reaches a maximum after a few hours (76).

Dextransucrase kinetics have generally been described by the Michaelis–Menten model and applied for sucrose levels below 100–150 g/L, where the saturation of the enzyme is achieved. Martinez-Espindola and Lopez-Munguina proposed a simple substrate inhibition model for enzyme action at higher sucrose levels as follows:

$$V_i = [V_{max} + S]/[K_m + S + (S^2/K_i)] \tag{9}$$

For enzymatic synthesis of dextran in a batch reactor,

Martinez-Espindola and Lopez-Munguina derived the following design equation:

$$V_{max}t = S_0X - K_m \ln(1 - X) + (K_iS_0^2)(2X - X^2)/2 \quad (10)$$

where S_0 is the initial sucrose concentration, X is the conversion, and K_i is the inhibitor constant. They predicted that 150 g/L of sucrose concentration is optimum for 99% conversion of sucrose to dextran in terms of productivity in a batch reactor (92).

London and Webb (96) optimized both the medium and inoculum for batch production of dextransucrase in an unaerated and unstirred culture. The optimum conditions are given in Table 3. They achieved a final enzyme activity of 142 DSU/mL. Although it is lower than the fed-batch achievement, it was the best reported value for batch operation. However, London and Webb's scaled-up process to 1,000 L yielded only 45 DSU/mL of activity (see Table 3).

We reported the optimum conditions for batch dextran production by using response surface methodology (RSM) (109) and multistage Monte Carlo optimization (MSMCO) (110) optimization techniques with a predicted yield of 154 g/L dextran. The optimum values are given in Table 3; under these conditions we observed 135 g/L dextran yield (90). Typical batch dextran fermentation profiles can be seen in Figure 7. The optimum initial pH of 8.3, which was set before inoculation and sterilization, dropped to 7.5 at the start of fermentation after adding 10% v/v acidic inoculum (see Figs. 5 and 6). Batch operational mode is a suitable method to produce dextran through either whole cell fermentation or cell-free enzymatic synthesis due to the possibility of a high final dextran concentration (hence the productivity), which minimizes the alcohol precipitation costs in recovery of the product. However, if our aim is to produce maximum enzyme activity, fed-batch operation that keeps lower levels of sucrose concentrations in the fermenter is the best operational mode.

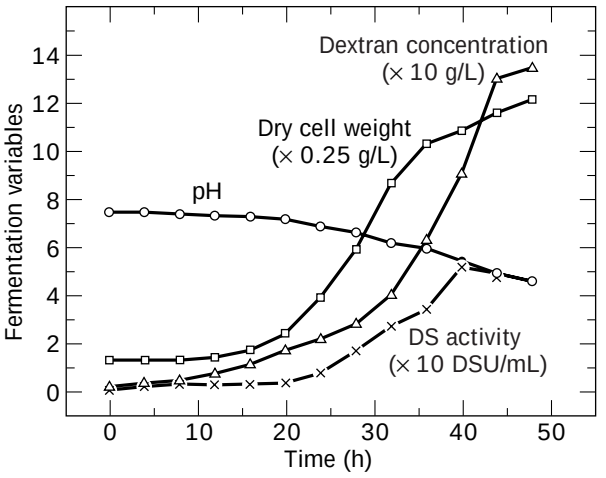


Figure 7. Typical batch dextran fermentation profiles. Fermentation conditions are given in Table 3. *Source*: From Ref. 90.

Semicontinuous Culture

Because dextran fermentation shows substrate inhibition kinetics, the substrate level can be maintained at the desired level by using fed-batch operational mode. This is shown by the various authors who produced higher enzyme activity than in batch cultures (87,89,111). Lopez and Monsan (87,88) and Monsan et al. (111,112) achieved 180 DSU/mL enzyme activity by fed-batch fermentation. They fed sucrose solution at a rate of 20 g/L/h and achieved maximum cell density and enzyme activity. Barker and Ajongwen (89) achieved the highest enzyme activity (450 DSU/mL) under low agitated, nonaerated fed-batch fermentation (see Table 3). El-Sayed and coworkers observed higher enzyme activity in fed-batch operation than batch mode in both immobilized-cell and free-cell cultures. They observed that when the sucrose feed rate is increased from 5 to 10 g/L/h, enzyme activity increases from 73 to 118 DSU/mL in free-cell culture and 110 to 223 DSU/mL

Table 3. Optimal Conditions for Production of Dextransucrase Used by Various Authors

Variables	El-Sayed et al. (108)	London and Webb (96)	Barker and Ajongwen (89)	Karthikeyan et al. (90)
Sucrose	20 g/L	40 g/L	10 g/L	300 g/L
Yeast extract	20 g/L	20 g/L	40 g/L	10 g/L
Tween 80	2 g/L	—	—	—
KH_2PO_4	2 g/L	—	—	—
K_2HPO_4	—	—	20 g/L	30 g/L
Na_2HPO_4	—	16 g/L	—	—
$CaCO_3$	10 g/L	—	—	—
R salts[a]	5 mL	—	0.5% v/v	—
pH	6.7[b]	7.1[c]	6.7[b]	8.3[c,d]
Temperature	26 °C	26 °C	23 °C	23 °C
Inoculum	—	1% v/v	—	10% v/v
Mode	Fed-batch	Batch	Fed-batch	Batch
Sucrose feed	10 g/L/h	—	Connected with alkali addition	—
Product	Dextransucrase	Dextransucrase	Dextransucrase	Dextran

[a]R salts contain 40 g $MgSO_4 \cdot 7H_2O$, 2 g NaCl, 2 g $FeSO_4 \cdot 7H_2O$, and 2 g $MnSO_4 \cdot H_2O$, dissolved in 1 L.
[b]pH-controlled fermentation.
[c]Indicates initial pH, which is uncontrolled during the fermentation.
[d]pH will become 7.5 after adding 10% v/v of acidic inoculum.

in immobilized-cell culture (113–115). Although fed-batch culture (used to maintain a low substrate level) is suitable for enzyme production, it may not be suitable for dextran production because of lower plant productivity and increased alcohol precipitation costs in downstream processing.

Continuous Culture

Ringpfeil and Selenia (116) theoretically calculated from batch culture experiments the maximum yield of dextransucrase that would be obtained from sucrose-limited continuous culture at low dilution rate. The maximum enzyme activity achieved at a dilution rate of 0.53 h^{-1} was 16 DSU/mL. However, Lawford et al. (82) reported that the enzyme can be synthesize only after a critical concentration of sucrose has been reached because they could not synthesize dextransucrase at dilution rates less than 0.2 h^{-1}. They calculated the theoretical dilution rate for maximum cell growth to be 0.67 h^{-1}. However, they achieved maximum cell growth rate at 0.35 h^{-1}, concluding that either some essential metabolite other than the energy source limited the growth or some components of the growth medium inhibited the growth. Lee et al. (73) produced low molecular weight clinical dextran in a two-stage continuous reactor using acceptors such as maltose, dextran T-40, and dextran T-10. Cells cultivated in the first reactor were transferred to the second reactor, where sucrose and acceptor were added for clinical dextran production.

Immobilized Cells and Enzymes

Use of immobilized dextransucrase, if it were of sufficient stability, would allow the production of dextran without the requirement for repetitive fermentation. Earlier reports of dextransucrase immobilization were issued from several groups. The enzyme has been immobilized ionically to DEAE sephadex A-50 (117), covalently to Bio-Gel-P2 (63,64), and covalently to polyacrylamide gel, cellulose acetate membranes, and polysulfone hollow fibers (118). In 1980 Kobali and Reilly (101) examined 13 carriers (e.g., Enacryls, DEAE cellulose, DEAE sephadex, SP sephadex, and porous silica) for dextransucrase immobilization. Among the 13 carriers surveyed, only alkaline amine porous silica retained more than 10% of activity on immobilization (10 DSU/g carrier). The optimum pH (5.2) and temperature (30 °C) of enzyme activity are not changed on immobilization, but the activation energy below 30 °C was 1.7 kcal/mol for immobilized enzyme, which is lower than for soluble enzyme (8.6 kcal/mol).

Monsan and Lopez immobilized dextransucrase on an aminoporous silica support activated with 2% glutaraldehyde by covalent coupling (111). They examined the specific surface area of the carrier silica for immobilization efficiency and observed that *(1)* lowering the specific surface area increased the coupling efficiency for dextransucrase and that *(2)* adding maltose to the dextransucrase preparation during enzyme coupling greatly increased the coupling efficiency. The production of high molecular weight dextran with immobilized dextransucrase results in extremely viscous solutions, which makes recovery of the support difficult at the end of the synthesis. They also observed from their kinetic study diffusional limitation in the immobilized catalyst because of high molecular weight dextran synthesis. Chang et al. (119) successfully immobilized the enzyme on porous phenoxy acetyl cellulose beads by hydrophobic adsorption with high efficiency. The enzyme was immobilized at more than 100 times the activity attained on porous silica beads by Kobali and Reilly (101) and at nearly double the activity found by Monsan and Lopez (111).

El-Sayed and coworkers published a series of papers (113–115) in which they studied the effect of different carriers (i.e., calcium alginate, three types of celites, celites coated with calcium alginate, stainless steel, and stainless steel coated with calcium alginate) for immobilization, mode of operation (batch and fed-batch), and the performance of different reactors (bubble column and airlift bioreactors). They further compared all the results with whole cell fermentation under the same conditions. The medium they used for dextransucrase production is given in Table 3.

El-Sayed and coworkers used the immobilized cells in calcium alginate beads for sequential production of dextransucrase followed by dextran production in three cycles. Each cycle consisted of fed-batch enzyme production for 24 h followed by batch dextran production for 24 h. The maximum dextran yields reached 40%, 44%, and 46% based on the added sucrose and were achieved 15, 7, and 3.5 h from the beginning of the first, second, and third dextran production periods, respectively. Increasing the sucrose feed rate from 5 to 10 g/L/h led to an increase in the total enzyme activity of about 88%. El-Sayed et al. compared the results with free-cell fermentation where only one cycle was possible. They observed that immobilized cells produced more than three times as much dextran as free cells during one cycle (113). When comparing different carriers they observed that the immobilized cell cultures were significantly affected by pore and particle size. Enzyme production with cells immobilized in celite R 648 resulted in the highest total enzyme activity followed by celite R 633, alginate-coated celite R 630, celite R 630, and the calcium alginate beads. A layer of calcium alginate on porous celite R 630 particles improved their mechanical stability, increased the amount of soluble dextransucrase activity, and decreased the cell leakage from the highly porous support (114). Immobilizing *L. mesenteroides* on a porous stainless steel support also increased dextransucrase and dextran production. Coating calcium alginate layers on stainless steel supports enhanced the cell loading capacity of the support and increased dextransucrase and dextran production (115).

Other Operational Strategies

Dextran fermentations consist of three phases—cell, enzyme, and dextran production—and each phase shows different optima for pH, temperature, and sucrose concentration. A major advantage in dextran fermentation is that these individual phases can be performed separately and sequentially in a reactor system that allows flexibility in operational strategy (120,121). Their separation could be achieved by either of the following simple techniques:

1. Single reactor system with multistage operation: a stepwise operation of the individual phases with their individual optimum condition.
2. Multistage reactor systems in series: simultaneous operation of individual phases (i.e., each individual phase performed with unique optimum levels in separate reactors in series).

In Figure 8, reactors a, b, and c represent single reactor–multistage operational systems with batch, fed-batch, and continuous mode of operation, respectively. In these systems, for example, if we divide the total fermentation time into three stages so as to perform cell, enzyme, and dextran production in a stepwise fashion, we can control the pH at 7 and temperature at 30 °C for the first stage to maximize cell growth. In the second stage, after reaching a high biomass concentration, we set the pH and temperature to 6.7 and 23 °C, respectively, for enzyme production. Similarly, in the third stage we can set the pH at 5.2 and temperature at 30 °C for dextran production. Further, we can control sucrose concentration in fed-batch and continuous cultures for these three stages. El-Sayed and coworkers attempted the sequential production of dextransucrase followed by dextran production using immobilized cells in an airlift bioreactor (113–115). If we divide the fermentation time for individual phases in the stepwise process the questions to be answered are as follows:

1. What is the required amount of biomass and enzyme activity to be produced to get a maximum conversion of given sucrose to dextran in the shortest time?
2. How many stages are required? Are three stages necessary, because dextran fermentation consists of three phases?
3. What is the duration required for each stage?

The term *required amount* in question 1 is important because the aim of fermentation is to produce maximum dextran, not cells or enzyme. If we produce more cells and enzymes than the required amount, it will result in low yield from the fermentation medium. Therefore, in the first and second stages we need not produce maximum cells and enzyme but a required amount to optimally convert sucrose to dextran in the overall process. Optimization can be solved by employing mathematical techniques such as dynamic programming and the continuous maximum principle (122–125). These techniques require the mathematical models for cell growth, enzyme formation and deactivation, and dextran synthesis. The model parameters should explain the effect of pH and temperature. In the stepwise process, we need not necessarily divide the total time into only the three stages suggested earlier. In fact, more than three stages will prove advantageous. The number of stages required for maximum production can also be obtained through the previously mentioned mathematical techniques.

Multistage reactors in series allow simultaneous operations. In all simultaneous reactor systems in series shown in Figure 9, the symbol F_a represents additional feed of sucrose to the respective reactors. This feed can be utilized to control sucrose concentration in the respective reactor. The optimal operational conditions for simultaneous operation can also be obtained by the previously mentioned mathematical techniques. In reactor system f in Figure 9, three reactors labeled A, B, and C are connected in series. Reactor A is kept at optimum conditions for cell growth and, similarly, B and C are kept at optimum conditions for

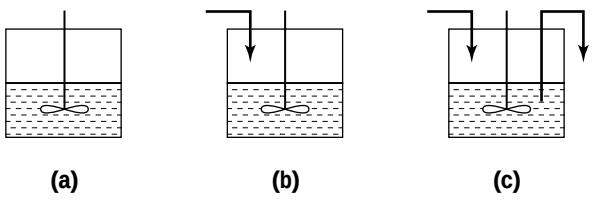


Figure 8. Reactor designs for stepwise process: (**a**) batch reactor, (**b**) fed-batch reactor, (**c**) continuous reactor.

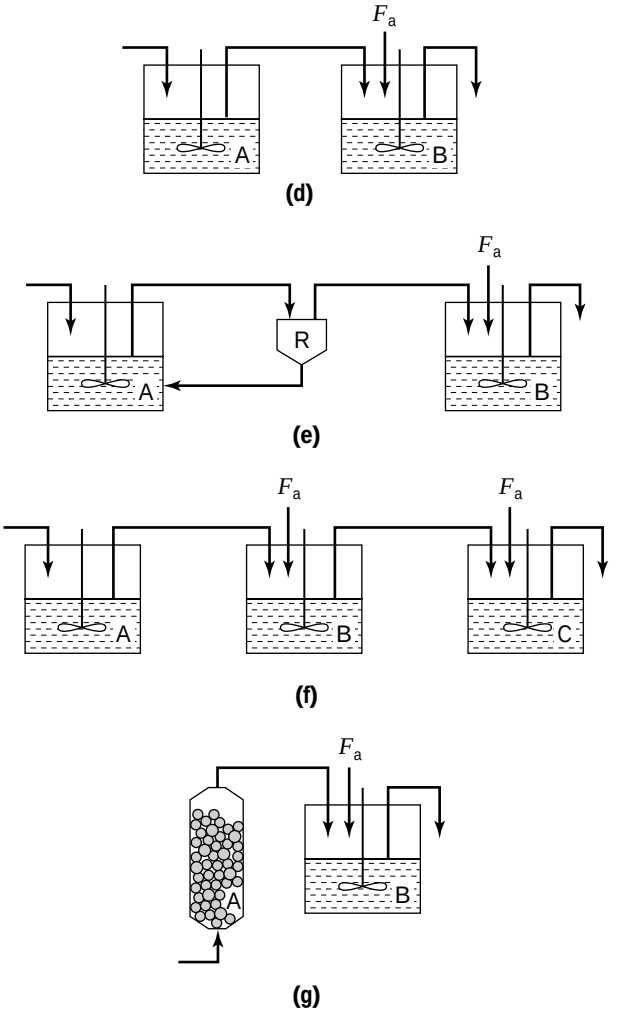


Figure 9. Reactor designs for simultaneous process (see text for description of reactors d–g).

enzyme and dextran production, respectively. Another reactor system shown in Figure 9d has only two reactors, labeled A and B. Here, reactor A supports both cell growth and enzyme production and reactor B performs the dextran synthesis. Separate sucrose feed (F_a) can also be employed in reactor B. Lee et al. (73) reported the performance of a two-stage continuous reactor for dextran synthesis. In the d reactor system in Figure 9, reactor A can also be utilized with immobilized cell systems to produce enzyme, which can be supplied to reactor B for dextran production. The reactor system shown in Figure 9e has a cell recycle unit R that separates the cells in the medium before it goes to reactor B and maximizes the cell concentration in reactor A. Another interesting system shown in Figure 9g consists of a packed-bed reactor (A) connected to a stirred tank reactor (B). Here immobilized cells are packed and are controlled at the optimum conditions for enzyme production. Using the enzyme from reactor A, reactor B supports dextran synthesis under optimum conditions. The stirred tank reactors shown in Figure 9 can be replaced by other reactors such as airlift, bubble column, and so forth.

A major advantage in reactor systems e and g in Figure 9 is that dilution of the final viscous broth for cell removal by centrifugation, is not needed, thereby saving on alcohol precipitation costs. The additional investment in terms of equipment and control systems and the operational difficulties are major disadvantages in systems with reactors in series. The overall economics of such a process should be worked out after taking into consideration the product value for such a process.

USES AND APPLICATIONS OF DEXTRANS

Depending on the microbial species employed and the fermentation conditions, a wide variety of dextrans can be produced. For example, dextrans commercially produced from *L. mesenteroides* NRRL-B512F consist of 95% α, $1 \rightarrow 6$ linkages and 5% α, $1 \rightarrow 3$ linkages. Many types of applications of this particular dextran and its derivatives have been developed (4,14,27,126–128). The important applications of dextran and its derivatives are described in the following sections.

Clinical Dextran

Dextran fractions of average molecular weight 70,000 can be used as blood plasma substitutes (or blood volume expanders). This single most important application initiated extensive research in dextran production. Clinical dextran (6% w/v) in isotonic saline provides viscous, colloidal, and osmotic properties equivalent to those of plasma proteins. Clinical dextran solutions can be administered intravenously for treatment of shock and burns, especially when blood is not available, to restore blood pressure and circulatory volume. The advantages of clinical dextran solution are its stability to sterilization by heat and in storage, independence of the blood types of recipients, low incidence of adverse reactions, and freedom from transmissible diseases (128,129). Clinical dextran (10–12% w/v) in aqueous or dextrose solutions administered intravenously provides high colloidal and osmotic activity and is used for treatment of hypoproteinemia and nephrotic and cerebral edema (128).

Rheomacrodex or Low-Viscosity Dextran

Dextran fractions of average molecular weight 40,000 can be used as blood flow improvers. Dextran (10% w/v) in sodium chloride solution (0.9% w/v) shows hematorheological properties that counteract capillary stagnation and improve the suspension stability of blood. When intravenously administered, it increases the fluidity of erythrocytes (130).

Cross-Linked Dextran Gel

Dextran is used in the manufacture of molecular sieves. It can be cross-linked with reagents such as epichlorohydrin in the presence of sodium hydroxide. The degree of cross-linking of dextran determines the porosity of the gel structure, water regain value, and molecular sieveing properties. These cross-linked gels, having three-dimensional structure, revolutionized the purification and separation, of proteins, nucleic acids, and polysaccharides based on differences in molecular size. These gels have well-defined exclusion limits in the molecular weight range of 5,000–200,000 (131). Commercial cross-linked dextran was introduced in the 1960s by Pharmacia Fine Chemical Limited under the tradename Sephadex. Hydroxyl propyl derivatives of Sephadex have been made to separate materials according to molecular size in organic solvents and have had use in separation of lipids, fatty acids, steroids, and so forth (132).

Diethylaminoethyl and carboxymethyl derivatives of cross-linked dextran gels provide adsorbants with different porosities and fractionation ranges used in the fractionation of proteins, nucleic acids, and polysaccharides on the basis of different adsorption characteristics. Cross-linked dextrans and their derivatives have found wide application in the separation and purification of biological molecules based on charge, molecular size, and so forth on both laboratory and preparative scales. Proteins are not usually denatured by hydrophilic polymer networks, and nonspecific adsorption is usually very low.

Dextran Sulfate

Dextran sulfate can be prepared by esterification of polysaccharide of molecular weight 5×10^5 with chlorosulfonic acid. This product, a polyelectrolyte containing 17% sulfur (an average 2.3 sulfate groups on each glucosyl residue), has various applications. It shows reversible interaction with plasma proteins, and under some conditions selective precipitation, and thus permits isolation of substances such as β-lipoproteins and human antihemophilic globulin (133).

Dextran of sufficiently low molecular weight (5,000–10,000), substituted with sulfate, has anticoagulant activity (128). It mimics the action of heparin but with 15% of the activity of heparin. Dextran with a molecular weight of 7,000–8,000, substituted with sulfate, is used for selective stabilization of γ-globulin and fibrinogen to heat (134).

The wide spectrum of biological properties of dextran sulfate has produced studies involving enzyme inhibition, activation, and release; adjuvant effects on immune responses; cellular interactions; and decreased virus infectivity and multiplication. It has shown some efficacy as a drug in the treatment of AIDS and tumors.

Dextran and Polyethylene Glycol in Aqueous Two-Phase Systems

The need for an extracting phase that does not harm labile enzymes, proteins, cellular organelles, and macromolecules has motivated consideration of liquid–liquid two-phase extraction systems in which each phase is primarily aqueous. The production of separate aqueous phases is achieved by dissolution of two incompatible polymers such as dextran and polyethylene glycol (135). Dextran may be replaced by its derivatives, including sodium dextran sulfate, sodium carboxymethyl dextran, or diethylaminoethyl dextran, in this two-phase system. These systems offer a new approach for extraction that is different from conventional techniques (136–138). Pullulanase and phosphorylase are extracted from *Klebsiella pneumoniae* using the dextran–polyethylene glycol (PEG) system (139).

Activated Dextrans

Several attributes (119) can be achieved by coupling ligands to dextran, including the following:

1. Stability and lifetime of ligands can be increased.
2. Because of the general compatibility of the dextran fraction in vivo, drugs that require a longer circulatory lifetime can be covalently coupled with dextran. Insulin has been coupled with dextran, and the hypoglycemic action of the complex was reported to be superior to native insulin. The complex also had a longer lifetime in the bloodstream than free insulin (140,141).
3. The antigenicity of protein ligands can be decreased.
4. Enzymes can be immobilized for enzyme reactors to produce various products (142).
5. Covalent linking of ligands for affinity chromotography can be used in purification processes (143).

Because dextran, in a well-defined molecular size range, is an established blood plasma substitute, the coupling with hemaglobin may provide a blood substitute. The preparation of hemaglobin covalently coupled with dextran was studied (144). These preparations were cleared slowly by the kidney, and it was found that the oxygen molecule bound more tightly than to free hemaglobin. Wong and associates have published a number of papers in which the preparation and experimental use of these dextran–hemaglobin products are described (145–149).

Industrial Uses

Dilute aqueous solutions of dextrans with molecular weights 18,000, 74,000, 228,000, and 7,400,000 are used as superior media for suspension polymerization of unsaturated monomers (150). Dextran (5–50% w/w; MW 10,000) incorporated into the total solids in gelatin-silver halide emulsions has significant use in the photography industry. It improves film surface coverage, maximum density contrast, and speed of gelatin-silver halide emulsions for radiologic and lithographic films (151,152).

Native Dextrans

Linear-type, water-soluble native dextrans are effective for a number of other potential uses. Dextran is superior to many polymers tested in viscous water flooding for secondary recovery of petroleum. The dextran used for this purpose has the requisite rheological properties—thermal, chemical, and solution stability and innertness to ionic effects; it does not plug earth formations and is not adsorbed on rock surfaces (153–155). Dextran is superior to carboxymethylcellulose and other common colloidal agents when used in high-calcium drilling muds even at elevated temperatures. Muds containing dextran or aldehyde-modified dextran show superior stability and performance at high pH and in saturated brines (156–159). Dextran shows relatively high resistance to deterioration in soil aggregates (160). It also makes effective protective coatings for seeds (161) and deflocculants for production of fine papers (162). Certain film-forming dextrans have value for protecting selected areas from deposition of metals in gas plating processes while remaining water soluble for ease of subsequent removal (163).

Other Uses

An iron–dextran complex has achieved significant commercial success in the treatment of iron-deficiency anemia (164). This complex is prepared by mixing a hot alkaline solution of dextran with $FeCl_3$ solution followed by neutralization. This complex containing 5% iron and 20% dextran is used for intramuscular or intravenous injections for treating iron deficiency. It is currently being used to treat anemia in newborn piglets. Dextran is used as a stabilizer in freeze-drying sensitive materials such as vaccines and bacteria (165). As recently reported, it is one of the best preservative and stabilizing agents for transplant organs (119). Compatibility of linear-type dextrans with living tissues and their ability to be metabolized permits their use in binding multiple collagen fibers into strands for surgical suture (166). Dextran's use in foods is based on its ability to stabilize syrups, confections, and the like against crystallization; to stabilize and provide viscosity to liquid sugars (167); to act as humectants and adhesives; and to show high tolerance for a low pH. Acting alone, dextran is superior for stabilizing texture in ice creams and sherbets (15).

BIBLIOGRAPHY

1. R. Margaritis and G.W. Pace, in M. Moo-Young ed., *Comprehensive Biotechnology*, vol. 3, Pergamon Press, London, 1985, pp. 1005–1044.
2. I.W. Sutherland, in H.J. Rehm and G. Reed eds., *Biotechnology*, vol. 3, Verlag Chemie, Weinheim, Germany, 1983, pp. 531–574.

3. I.W. Sutherland, *Biotechnology of Microbial Polysaccharides*, Cambridge Univ. Press, Cambridge, U.K., 1990.
4. A. Jeanes, *Encyclopedia of Polymer Science and Technology*, vol. 4, Wiley, New York, 1966, pp. 805–824.
5. E.J. Hehre, *Science* **93**, 237–238 (1941).
6. A. Jeanes, et al., *J. Am. Chem. Soc.* **76**, 5041–5052 (1954).
7. J.C. Mazza, et al., *Carbohydr. Res.* **40**, 402–406 (1975).
8. E.J. Hehre and H. Suzuki, *Arch. Biochem. Biophys.* **113**, 675–683 (1966).
9. D.S. Genghof and E.J. Hehre, *Proc. Sco. Expt. Biol. Med.* **140**, 1298–1301 (1972).
10. T.P. Binder and J.F. Robyt, *Carbohydr. Res.* **124**, 287–299 (1983).
11. E.J. Hehre and D. Hamilton, *Proc. Soc. Exp. Biol. Med.* **71**, 336 (1949).
12. S.A. Barker, et al., *J. Chem. Soc.* 4414 (1958).
13. E.R. Ruckel and C. Schuerch, *Biopolymers* **5**, 515 (1967).
14. R.M. Alsop, *Prog. Industrial Microbiol.* **18**, 1–44 (1983).
15. P.J. Barker, Jr., in R.L. Whister and J.N. Be Miller eds., *Industrial Gums: Polysaccharides and Their Derivatives*, Academic Press, New York, 1959, pp. 531–563.
16. L. Pasteur, *Bull. Soc. Chim. Fr.* 30–31 (1861).
17. C. Scheubler, *Z. Ver. Dtsch. Zuker-Ind.* **24**, 309–317 (1874).
18. P. Van Tieghem, *Ann. Sci. Nat. Bot. Biol. Veg.* **7**, 180–203 (1878).
19. E.J. Hehre and J.Y. Sugg, *J. Exp. Med.* **75**, 339 (1942).
20. S. Hestrin and S. Shapiro-Avineri, *Biochem. J.* **38**, 2 (1944).
21. P.E. Barker and E.J. Bourne, *Q. Rev.* **7**, 56 (1953).
22. E.J. Hehre, *Adv. Enzymol.* **11**, 297 (1951).
23. J. Edelman, *Adv. Enzymol.* **17**, 189 (1956).
24. F.L. Flower, et al., *Can. J. Res.* **B15**, 486 (1937).
25. W.B. Neely, *Adv. Carbohydr. Chem.* **15**, 341–364 (1960).
26. A. Jeanes, *A.C.S. Symp. Ser.* **45**, 284–298 (1977).
27. P.T. Murphy and R.L. Whistlkr, *Dextran in Industrial Gums: Polysaccharides and Their Derivatives*, 2nd ed., Academic Press, New York, 1973, pp. 513–542.
28. R.L. Sidebotham, *Adv. Carbohydr. Chem. Biochem.* **30**, 371–444 (1974).
29. M. Wales, et al., *J. Polym. Sci.* **10**, 229 (1953).
30. G.J. Hucker and C.S. Pederson, *N.Y. State Agr. Expt. Sta. Tech. Bull.* **167**, 3 (1930).
31. C.A. Wilham, et al., *Arch. Biochem. Biophys.* **59**, 61 (1955).
32. J.W. Van Cleve, et al., *J. Am. Chem. Soc.* **78**, 4435–4440 (1956).
33. A. Jeanes and F.R. Seymour, *Carbohydr. Res.* **74**, 31–40 (1979).
34. F.R. Senti, et al., *J. Polym. Sci.* **17**, 527–546 (1955).
35. G.J. Walker and A. Pulkownik, *Carbohydr. Res.* **29**, 1–4 (1973).
36. J.F. Robyt, in S.B. Peterson, B. Svensson, and S. Pedersen eds., *Carbohydrate Bioengineering*, Elsevier, Amsterdam, 1995.
37. J.F. Robyt, et al., *Carbohydr. Res.* **266**, 293–299 (1995).
38. A. Shimamura, et al., *Biochem. Biophys. Acta* **702**, 72–80 (1982).
39. F.R. Seymour, et al., *Carbohydr. Res.* **68**, 1130–121 (1979).
40. F.R. Seymour, et al., *Carbohydr. Res.* **68**, 123–140 (1979).
41. F.R. Seymour, et al., *Carbohydr. Res.* **74**, 41–47 (1979).
42. F.R. Seymour, et al., *Carbohydr. Res.* **75**, 275–294 (1979).
43. F.R. Seymono and R.D. Knapp, *Carbohydr. Res.* **81**, 105–129 (1980).
44. M.D. Hare, et al., *Carbohydr. Res.* **66**, 245–255 (1978).
45. K.H. Ebert and G. Schenk, *Adv. Enzymol.* **30**, 179–221 (1968).
46. M. Kobayashi and K. Matsuda, *Biochim. Biophys. Acta* **614**, 46–62 (1980).
47. M. Kobayashi and K. Matsuda, *J. Biochem.* **100**, 615–621 (1986).
48. A.W. Miller, et al., *Carbohydr. Res.* **147**, 119–133 (1986).
49. D. Fu and J.F. Robyt, *Prep. Biochem.* **20**, 93–106 (1990).
50. G.T. Cori and C.F. Cori, *J. Biol. Chem.* **131**, 397–398 (1939).
51. C.S. Hanes, *Proc. R. Soc. London, Ser. B* **129**, 174–208 (1940).
52. M.A. Swanson and C.F. Cori, *J. Biol. Chem.* **172**, 815–829 (1948).
53. H.J. Koepsell, et al., *J. Biol. Chem.* **200**, 793–801 (1953).
54. H.M. Tsuchiya, et al., *J. Am. Chem. Soc.* **77**, 2412–2419 (1955).
55. G.R. Germaine, et al., *J. Bacteriol.* **120**, 287–294 (1974).
56. G.R. Germaine, et al., *Infect. Immun.* **16**, 637–248 (1977).
57. A.M. Chludzinski, et al., *J. Dent. Res.* **55**, C75–C86 (1976).
58. M. Kobayashi, et al., *Agric. Biol. Chem.* **50**, 2585–2590 (1986).
59. J.F. Robyt and A.J. Corrigan, *Arch. Biochem. Biophys.* **174**, 129–135 (1976).
60. K.H. Ebert and F. Patat, *Z. Naturforsch.* **17b**, 738–745 (1962).
61. W.B. Neely, *Adv. Carbohydr. Chem.* **15**, 341–364 (1960).
62. J.F. Robyt and T.F. Walseth, *Carbohydr. Res.* **61**, 433–445 (1978).
63. J.F. Robyt and H. Taniguchi, *Arch. Biochem. Biophys.* **174**, 129–135 (1976).
64. J.F. Robyt, B.K. Kimble, and J.F. Walseth, *Arch. Biochem. Biophys.* **165**, 634–640 (1974).
65. J.F. Robyt and S.H. Eklund, *Bioorg. Chem.* **11**, 115 (1982).
66. J.F. Robyt, *Adv. Carbohydr. Chem. Biochem.* **51**, 133–168 (1995).
67. F.A. Bovey, *J. Polym. Sci.* **35**, 423 (1967).
68. K.H. Ebert and M. Brosche, *Biopolymers* **5**, 423 (1967).
69. J.F. Robyt and S.H. Eklund, *Carbohydr. Res.* **121**, 279–286 (1983).
70. D. Su and J.F. Robyt, *Carbohydr. Res.* **248**, 339–348 (1993).
71. P. Paul, et al., *Carbohydr. Res.* **149**, 433–441 (1986).
72. D. Part, et al., *Biotech. Lett.* **9**, 1–6 (1987).
73. S. Lee, et al., *Biotech. Lett.* **14**, 379–384 (1992).
74. A.W. Miller and J.F. Robyt, *Biochim. Biophys. Acta* **880**, 32–39 (1976).
75. M. Kobayashi and K. Matsuda, *J. Biochem.* **79**, 1301–1308 (1976).
76. J.F. Robyt and T.F. Walseth, *Carbohydr. Res.* **68**, 95–111 (1979).
77. A.W. Miller and J.F. Robyt, *Biochim. Biophys. Acta* **785**, 89–46 (1984).
78. E. Braswell, et al., *J. Polym. Sci.* **61**, 143–151 (1962).
79. A. Jeanes, *Methods Carbohydr. Chem.* **5**, 118–127 (1965).
80. J.E. Ciardi, et al., *Infect. Immun.* **18**, 237–241 (1977).
81. D. Kim and J.F. Robyt, *Enzyme Microb. Technol.* **16**, 659–664 (1994).
82. G.R. Lawford, *Biotech. Bioeng.* **21**, 1121–1131 (1979).

83. D.H. Bergey, *Manual of Determinative Bacteriology*, 8th ed., 1974.
84. M.K. Johnson and C.S. McCleskey, *J. Bacteriol.* **74**, 22–25 (1957).
85. H.M. Tsuchiya, et al., *J. Bacteriol.* **64**, 521–526 (1952).
86. U.K. Pat. 2079292 B (May 16, 1984), M. Schneider, et al.
87. A. Lopez and P. Monsan, *Biochemie* **62**, 325 (1980).
88. A. Lopez and P. Monsan, *Adv. Biotechnol.* **1**, 679 (1981).
89. P.E. Barker and N.J. Ajongwen, *Biotech. Bioeng.* **37**, 703–707 (1991).
90. R.S. Karthikeyan, et al., *Bioprocess Eng.* **15**, 247–251 (1995).
91. E. Braswell, et al., *J. Polym. Sci.* **41**, 467 (1959).
92. J.P. Martinez-Espindola and C.A. Lopez-Munguina, *Biotech. Lett.* **7**, 483–486 (1985).
93. R.S. Karthikeyan, et al., *Chemical Engineering Congr.*, Bombay, India, December 1993.
94. R.S. Karthikeyan, Ph.D. Thesis, I.I.T-Madras, India, 1999.
95. R.S. London, Ph.D. Thesis, UMIST, Manchester, U.K., 1982.
96. R.S. London and C. Webb, *Proc. Biochem.* 19–23 (February 1990).
97. V.W. Carlson and W.W. Carlson, *J. Bacteriol.* **58**, 135–140 (1949).
98. V.W. Carlson and W.W. Carlson, *J. Bacteriol.* **58**, 143–149 (1949).
99. V.W. Carlson and W.W. Carlson, *J. Bacteriol.* **62**, 583–589 (1951).
100. H.J. Koepsell and H.M. Tsuchiya, *J. Bacteriol.* **63**, 293–295 (1952).
101. H. Kobali and P.J. Reilly, *Biotech. Bioeng.* **22**, 1055–1069 (1980).
102. J. Sabatie and L. Choplin, *Biotech. Lett.* **8**, 425–430 (1986).
103. F.H. Foster, *Process Biotech.* **3**, 15–19 (1968).
104. F.H. Foster, *Process Biochem.* **3**, 55–62 (1968).
105. M.K. Johnson and C.S. McCleskey, *J. Bacteriol.* **74**, 22–25 (1957).
106. A. Margaritis and J.E. Zajic, *Biotech. Bioeng.* **20**, 939–1001 (1978).
107. A. Margaritis and D. Bokkel, *187th Annu. Meeting American Chemical Society*, Philadelphia, Pa., August 26–31, 1984.
108. A.M.M. El-Sayed, et al., *Biotech. Bioeng.* **36**, 338–345 (1990).
109. S. Akhnazarova and V. Kafarov, *Experiment Optimization in Chemistry and Chemical Engineering*, Moscow, Mir Publishers, 1982.
110. W.C. Conley, *Computer Optimization Techniques*, rev. ed., Princeton, N.J., 1984.
111. P. Monsan and A. Lopez, *Biotech. Bioeng.* **23**, 2027–2037 (1981).
112. P. Monsan, et al., *Ann. N.Y. Acad. Sci.* **434**, 276 (1983).
113. A.M.M. El-Sayed, et al., *Biotech. Bioeng.* **36**, 346–353 (1990).
114. A.M.M. El-Sayed, et al., *Biotech. Bioeng.* **36**, 83–91 (1990).
115. A.M.M. El-Sayed, et al., *Biotech. Bioeng.* **40**, 617–624 (1992).
116. M. Ringpfeil and M. Selenia, in I. Meleik, K. Beran, and J. Hospedka eds., *Continuous Culture of Microorganisms*, Pub. House of Zech. Acad. Sci., Prague, 1964, pp. 145–150.
117. S. Ognio, *Agric. Biol. Chem.* **34**, 1268 (1970).
118. C.R. Edwards and S.W. Drew, *Abstr. Ann. Meet. Am. Soc. Microb.* **77**, 251 (1977).
119. H.N. Chang, et al., *Biotech. Bioeng.* **23**, 2647–2653 (1981).
120. D.J. Widde, *Optimum Seeking Methods*, Prentice-Hall, Englewood Cliffs, N.J., 1964.
121. H. Ohno, E. Nakanishi, and T. Takamatsu, *Biotech. Bioeng.* **20**, 625 (1978).
122. G.S. Beveridge and R.S. Schechter, *Optimization Theory and Practice*, McGraw-Hill, New York, 1969.
123. F.S. Hiller and G.L. Liebermann, *Introduction to Operations Research*, McGraw-Hill, New York, 1989.
124. R.E. Bellman and S.E. Dreyfus, *Applied Dynamic Programming*, Princeton Univ. Press, Princeton, N.J., 1962.
125. L.T. Fan, *The Continuous Maximum Principle*, Wiley, New York, 1966.
126. J.F. Robyt, *Proc. Am. Chem. Soc.* **72**, 139–140 (1955).
127. G.H. Bixler, et al., *Ind. Eng. Chem.* **45**, 682 (1953).
128. A.J.T. Gronwall, *Dextran and Its Uses in Colloidal Infusion Solutions*, Academic Press, New York, 1957, p. 156.
129. A.J.T. Gronwall and B.G.A. Inglelman, *Acta Phys. Scand.* **7**, 97 (1945).
130. L.E. Gelin, et al., *Acta Chir. Scand.* **122**, 294–302 (1961).
131. P. Flodin, Meijels Bokindustri, Halmstad, Uppsala, 1962, p. 85.
132. Sephadex, Pharmacia Fine Chemicals Inc., New York, 1964.
133. J.L. Oneley, et al., *J. Am. Chem. Soc.* **76**, 4666–4671 (1957).
134. K.W. Walton and R.B. Pennell, *Arch. Biochem. Biophys.* **56**, 553–554 (1955).
135. P.A. Albertsson, in D. Glick ed., *Methods of Biochemical Analysis*, vol. 10, Wiley, New York, 1962, pp. 229–262.
136. M.R. Kula, *Appl. Biochem. Bioeng.* **2**, 71–94 (1979).
137. M.R. Kula, et al., *Enzyme Eng.* **4**, 47 (1978).
138. U. Matsumoto, et al., *J. Chromotogr.* **285**, 69 (1984).
139. H. Walter, et al., eds., *Partitioning in Aqueous Two-Phase Systems: Theory, Methods, Uses, and Applications in Biotechnology*, Academic Press, New York, 1985.
140. K. Armstrong, et al., *Biochem. Biophys. Acta* **498**, 102 (1977).
141. Y. Sakamoto, et al., *Biochem. Biophys. Acta* **498**, 102 (1977).
142. J. Woodward, *Immobilized Cells and Enzymes: A Practical Approach*, IRL Press, Oxford, U.K., 1996.
143. C.M. Yang and G.T. Tsao, *Adv. Biochem. Eng.* **15**, 19–42 (1982).
144. J. Blumenstein, et al., *Prog. Clin. Biol. Res.* **19**, 205 (1978).
145. J.E. Chang and J.T. Wong, *Can. J. Biochem.* **55**, 388–403 (1977).
146. P.G. Cunnington, et al., *Biochem. J.* **193**, 361–366 (1981).
147. S. Tam, et al., *Can. J. Biochem.* **56**, 981–984 (1978).
148. S. Tam, et al., *Can. J. Biochem.* **58**, 732–736 (1980).
149. S. Tam, et al., *Proc. Natl. Acad. Sci. U.S.A.* **73**, 2128–2131 (1976).
150. U.S. Pat. 2,587,367 (1958), J.J. Kearney.
151. U.S. Pat. 3,063,838 (1962), A.B. Jennings.
152. U.S. Pat. 3,203,804 (1965), A.B. Cohen and C.D. Schecklett.
153. U.S. Pat. 3,042,611 (1962), J.T. Patton.
154. U.S. Pat. 3,053,765 (1962), W.J. Sparks.
155. U.S. Pat. 3,084,122 (1963), J.C. Cybert and J.T. Patlon.
156. P.H. Monaghan and J.L. Gidley, *Oil Gas J.* **57**, 100–103 (1959).
157. U.S. Pat. 3,065,170 (1962), P.H. Monaghan and G.K. Dumbauld.
158. E.P. Mueller and Z. Angew, *Geol.* **9**, 213–217 (1963).
157. U.S. Pat. 2,602,002 (1952), W.L. Owen.
160. L.J. Novak, et al., *Agric. Food Chem.* **3**, 1028–1033 (1955).
161. U.S. Pat. 2,764,843 (1956), P.G. Peake.
162. A. De Waele, *Chem. Ind. (London)* **32**, 253 (1945).

163. U.S. Pat. 2,970,064 (1961), W.M. Bolton.
164. U.S. Pat. 2,820,740 (1958), E. London and G.D. Twigg.
165. U.S. Pat. 2,908,614 (1959), P.W. Muggleton and J. Ungar.
166. U.S. Pat. 2,748,774 (1956), L.J. Novak.
167. U.S. Pat. 2,409,816 (1946), D.V. Wardsworth and H.F. Hughes.

See also XANTHAN GUM.

DICARBOXYLIC ACID. See CITRIC ACID, PROCESSES; ENZYMES, FRUIT JUICE PROCESSING; OPINE DEHYDROGENASE, SECONDARY AMINE DICARBOXYLIC ACIDS.

DIHYDROXYPHENYLALANINE (4), PRODUCED BY MICROORGANISMS

HIDEHIKO KUMAGAI
Kyoto University
Kyoto, Japan

KEY WORDS

Ammonia
Erwinia herbicola
L-DOPA
L-Tyrosine
Pyrocatechol
Pyruvate
Tyrosine phenol-lyase (TPL)

OUTLINE

INTRODUCTION

Tyrosine phenol-lyase (TPL) (deaminating) (EC 4.1.99.2) catalyzes 3,4-dihydroxyphenyl-L-alanine (L-DOPA) synthesis from pyrocatechol, pyruvate, and ammonia in a one-step reaction. In the investigation to produce L-DOPA by this enzyme reaction, a bacterial strain (*Erwinia herbicola*) has been chosen as the best producer of TPL out of more than 1,000 others tested. Culture conditions for the preparation of cells containing high TPL activity and reaction condition for the synthesis of L-DOPA were optimized with this bacterium to support L-DOPA, appearing as the crystals in the reaction mixture, at final amounts reaching 110 g/L.

L-DOPA is the precursor of the neurotransmitter dopamine and is an effective agent in the treatment of Parkinson's disease. In this disease, the amount of dopamine in the brain of the patient is insufficient. L-DOPA had been produced mainly by a synthetic method that contains many chemical reaction steps, including a step for optical resolution. TPL is a pyridoxal phosphate dependent multifunctional enzyme obtained from bacteria that catalyzes the synthesis of L-DOPA from pyruvate, ammonia, and pyrocatechol in a one-step reaction. The L-DOPA from *E. herbicola*, is produced in cells containing high amount of TPL. This biocatalytic method is one of the most economical processes to date for the production of L-DOPA.

SELECTION OF MICROORGANISMS HAVING HIGH TPL ACTIVITY

The activity of TPL in microorganisms was investigated with 646 strains of bacteria, 140 strains of yeast, 138 strains of fungi, and 117 strains of actinomycetes (1). TPL activity was determined by measuring the amount of L-tyrosine or L-DOPA synthesized through the β-replacement reaction between L-serine and phenol or pyrocatechol, respectively. Cells of microorganisms were directly added to the reaction mixture as the enzyme for practical applications. TPL activity occurred widely in a variety of bacteria, most of which belong to the Enterobacteriaceae, especially to the genera *Escherichia, Aerobacter, Proteus*, and *Erwinia*. No activity was observed in yeasts, fungi, and actinomycetes. A large number of bacterial strains isolated from rice and corn plants showed high TPL activity. Cells of *E. herbicola* ATCC 21434, which showed the highest activity, were selected as a promising source of enzyme for the preparation of L-DOPA.

Aeration, pH, and temperature affect the formation of TPL. Cell growth requires oxygen transfer, but the condition for TPL synthesis may have a different optimum. For example, the optimum volume of the medium for the formation of TPL was 40 to 60 mL in a 500-mL shaking flask, but cell growth was rather repressed in this volume. The optimum pH for enzyme formation in this example was in the range of 7.0 to 7.5, and the optimum temperature for enzyme formation was in the range of 27 to 32 °C.

TPL activity occurs during cultivation of *E. herbicola*, and continues beyond the active growth phase. Cells were grown at 31.5 °C for 32 h with reciprocal shaking. The medium contained 0.2% tyrosine, 0.2% KH_2PO_4, 0.1% $MgSO_4 \cdot 7H_2O$, 1.0% yeast extract, 0.5% polypeptone, and 0.5% meat extract (pH 7.5) and supported maximum formation of the enzyme during the early stationary phase (cultivated for 26 to 30 h); thereafter, the enzyme gradually disappeared with the consumption of L-tyrosine added to the medium. Additions of yeast extract, meat extract, polypeptone, and the hydrolyzate of soybean protein to the basal medium enhance cell growth as well as the formation of

TPL. DL-Methionine, DL-alanine, and glycine stimulate enzyme formation. However, L-serine, L-ornithine, and L-cystine inhibit enzyme formation. Pyridoxine is added as a required precursor of coenzyme for enzyme formation.

Catabolite repression of formation of TPL by *E. herbicola* has been observed on adding glucose, pyruvate, and α-ketoglutarate to the medium at high concentrations. Glycerol is a suitable carbon source for cell growth as well as for the accumulation of the enzyme in growing cells. A marked increase in enzyme formation has been seen when glycerol is added together with succinate, fumarate or malate. Among the metallic ions tested, Fe^{2+} ion was the only one effective for the formation of TPL.

TPL is an inducible enzyme, and the addition of L-tyrosine to the medium is essential for formation of the enzyme. However, when large amounts of L-tyrosine were added, inhibition of enzyme formation and repression of cell growth were observed. The phenomena seem to be caused by the phenol that is liberated from the L-tyrosine.

Various analogs of L-tyrosine that were not decomposed during cultivation were investigated in the search for an inducer of TPL, but no effective, nondecomposable inducer was found among the L-tyrosine analogs tested. L-Phenylalanine, D-phenylalanine, and phenylpyruvate did not induce this enzyme by themselves, but they show a synergistic effect on the induction of the enzyme by L-tyrosine. L- and D-phenylalanine are competitive inhibitors of TPL.

Cells of *E. herbicola* with high TPL activity are typically prepared by growing them at 28 °C for 28 h in a medium containing 0.2% KH_2PO_4, 0.1% $MgSO_4$ $7H_2O$, 2 ppm Fe^{2+} ($FeSO_4 \cdot 7H_2O$), 0.01% pyridoxine · HCl, 0.6% glycerol, 0.5% succinic acid, 0.1% DL-methionine, 0.2% DL-alanine, 0.05% glycine, 0.1% L-phenylalanine, and 12 mL of hydrolyzed soybean protein in 100 mL of tap water, with the pH controlled at 7.5 throughout cultivation. Under these conditions, TPL is efficiently accumulated in the cells of *E. herbicola* and makes up about 10% of the total soluble cellular protein.

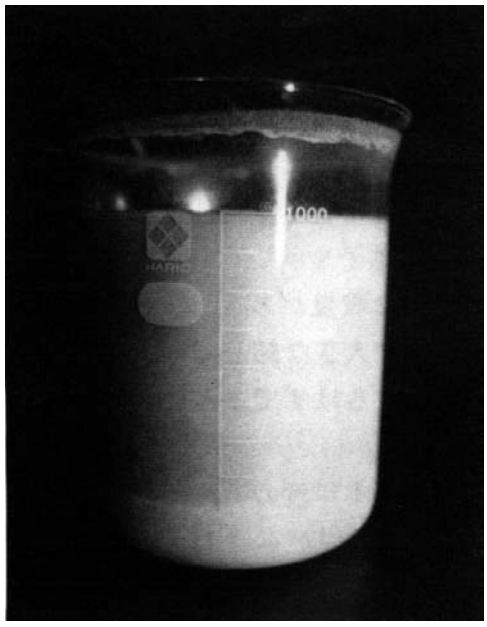


Figure 1. The reaction mixture for the L-DOPA synthesis by TPL. Crystals of L-DOPA are precipitated on the bottom of flask.

Figure 2. Microphotograph of crystals of L-DOPA.

REACTION CONDITIONS FOR THE SYNTHESIS OF L-DOPA

Synthesis of L-DOPA from Pyruvate, Ammonia, and Pyrocatechol

The optimum pH for this reaction is around 8.0, and the optimum temperature is 16 °C. Because pyruvate is

Table 1. Comparison of Chemical and Enzymatic Methods of L-DOPA Production

	Chemical Method	Enzymatic Method
Starting materials	Vanillin, hydantoin, acetic anhydride, hydrogen	Pyrocatechol, pyruvate, ammonia
Total number of unit-reaction	8	1
By-products	Ammonia, CO_2, acetate	H_2O
Necessity of optical resolution	Yes	No
Equipment	Specific plant	Commonly usable fermenter
Period (day)	15	3
Contaminating amino acids	L-Tyrosine, 3-methoxytyrosine, trihydroxy-phenylalanine	L-Tyrosine

unstable in the reaction mixture at a high temperature, low temperature is favorable for the synthesis of L-DOPA. The reaction is carried out at 16 °C for 48 h in a reaction mixture containing various amounts of sodium pyruvate, 5 g of ammonium acetate, 0.6 g of pyrocatechol, 0.2 g of sodium sulfite, 0.1 g of EDTA, and cells harvested from 100 mL of broth, in a total volume of 100 mL. The pH was adjusted to 8.0 by the addition of ammonia. At 2-h intervals, sodium pyruvate and pyrocatechol were added to the reaction mixture to maintain the initial concentrations. The maximum synthesis of L-DOPA was obtained when the concentration of sodium pyruvate was kept at 0.5%. L-DOPA appears as the crystals in the reaction mixture, as shown in Figures 1 and 2.

When a higher concentration of pyruvate was added to the reaction mixture during the synthesis of L-DOPA, two by-products, 1-methyl-1,3-dicarboxy-6,7-dihydroxy-1,2,3,4-tetrahydroisoquinoline and 1-methyl-3-carboxy-6,7-dihydroxy-3,4-dihydroisoquinoline, were formed. The by-products have been shown to be formed through a spontaneous condensation reaction.

TPL catalyzes a variety of α,β-elimination, β-replacement, and racemization reactions and the synthesis of L-tyrosine and its related amino acids. The mechanism for these reactions has been studied in some detail, by adopting the general mechanism proposed for pyridoxal-dependent reactions (1).

An enzymatic method for preparing L-DOPA was developed using bacterial cells with high TPL activity. This method is simple and is one of the most economical processes to date (Table 1) for preparing L-DOPA. L-DOPA is currently produced by this method with 60-m^3 scale fermenters. The yield of L-DOPA compared to pyrocatechol is 98% and to pyruvate, 90%, with the final accumulation of L-DOPA reaching 110 g/L (2). Certain L-tyrosine-related amino acids can be synthesized from the starting materials of pyruvic acid and appropriate derivatives of phenol, such as pyrocatechol (1).

BIBLIOGRAPHY

1. H. Yamada and H. Kumagai, in D. Perlman, ed. *Advances in Applied Microbiology*, vol. 19, Academic Press, New York, 1975, pp. 249–288.
2. H. Kumagai, H. Enei, H. Nakazawa, T. Tsuchida, and T. Namekawa, *Biosci. Industry* **54**, 9–15 (1996).

DILTIAZEM SYNTHESIS

HIROAKI MATSUMAE
HIROYUKI AKATSUKA
TAKEJI SHIBATANI
Tanabe Seiyaku Co., Ltd.
Osaka, Japan

KEY WORDS

Asymmetric hydrolysis
Biocatalyst
Diltiazem
Lipase
Membrane bioreactor
Methyl(2*R*,3*S*)-3-(4-methoxyphenyl)glycidate
Recombinant

OUTLINE

INTRODUCTION

Diltiazem, (2S,3*S*)-3-acetoxy-5-[2-(dimethylamino)ethyl]-2,3-dihydro-2-(4-methoxyphenyl)-1,5-benzothiazepin-4(5*H*)-one hydrochloride, is a calcium channel blocker and has been clinically used since 1974 as an antianginal and antihypertensive agent in more than 100 countries (1). Because diltiazem has two stereo centers at positions 2 and 3 in the structure shown in Figure 1, four isomers are possible. Among them, only the (2*S*,3*S*)-isomer exhibits the desired coronary vasodilating activity. The conventional production has been a nine-step process starting with *p*-anisaldehyde and methyl chloroacetate (1). The fourth step is a chemical resolution of the advanced intermediate 6. Resolution at an earlier step should save raw materials and reduce the amount of waste. (±)-MPGM is the first intermediate containing stereo centers. Two of the four isomers are excluded through the chemical synthesis of (±)-MPGM, which yields only *trans* type MPGM. Furthermore, no racemization occurs on the subsequent steps from (−)-

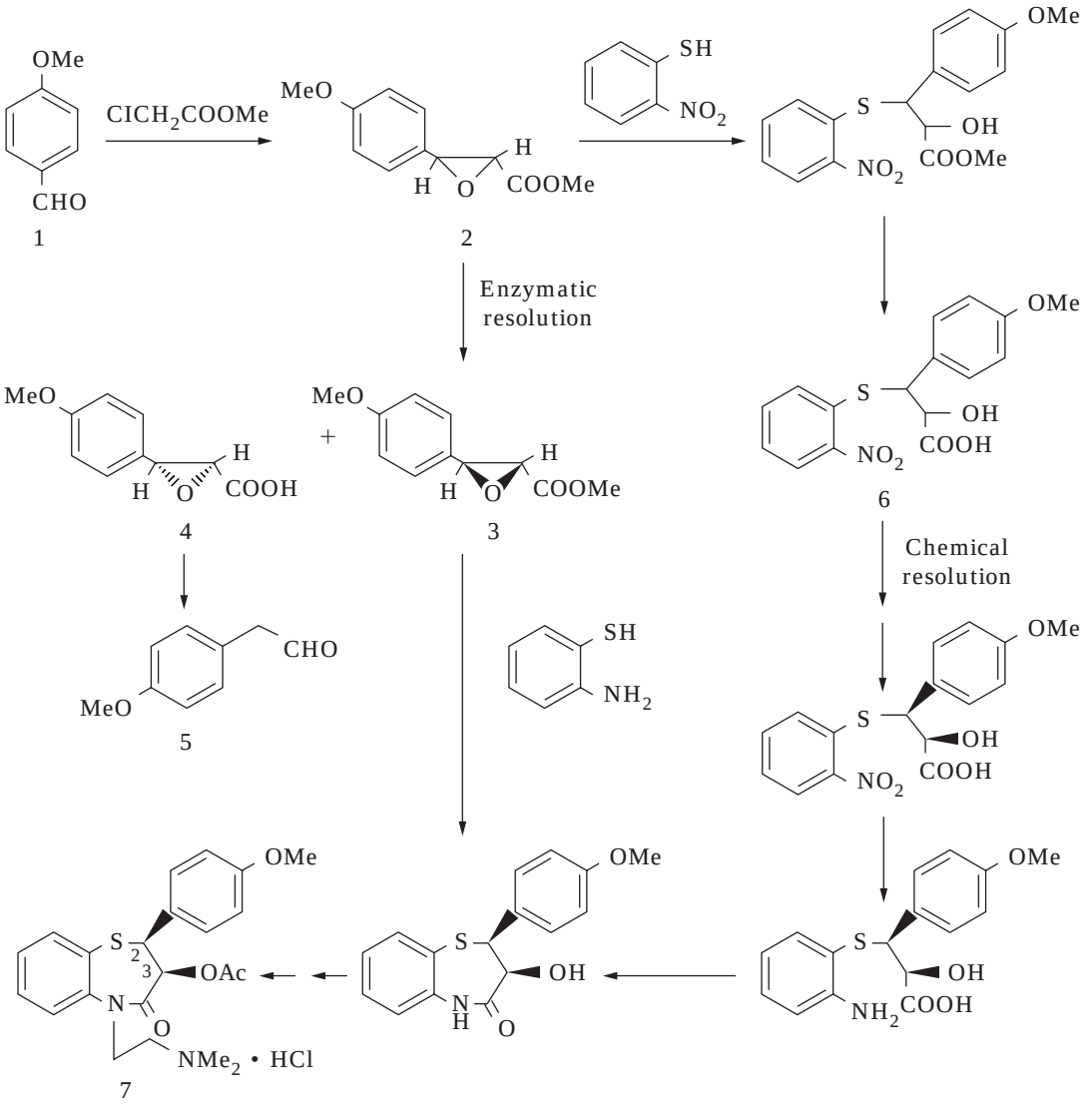


Figure 1. The synthetic process for diltiazem hydrochloride. Numerals: 1, *p*-anisaldehyde; 2, (±)-(2*RS*,3*SR*)-3-(4-methoxyphenyl)glycidic acid methyl ester [(±)-MPGM]; 3, (−)-(2*R*,3*S*)-3-(4-methoxyphenyl)glycidic acid methyl ester [(−)-MPGM]; 4, (+)-(2*S*,3*R*)-3-(4-methoxyphenyl)glycidic acid; 5, 4-methoxyphenylacetaldehyde; 6, *threo*-2-hydroxy-3-(4-methoxyphenyl)-3-(2-nitrophenylthio)propanoic acid; 7, (+)-(2*S*,3*S*)-3-acetoxy-5-[2-(dimethylamino)ethyl]-2,3-dihydro-2-(4-methoxyphenyl)-1,5-benzothiazepin-4(5*H*)-one hydrochloride (diltiazem).

MPGM to diltiazem. Therefore, the asymmetric syntheses of (−)-MPGM (2–5) and resolutions of (±)-MPGM (6,7) have been investigated for the synthesis of diltiazem. We focus on the resolution of (±)-MPGM as the more effective method of (−)-MPGM production (8,9).

Previous workers resolved (±)-MPGM by first hydrolyzing to the carboxylic acid, resolving the carboxylic acid with an optically active amine, and then esterifying (10,11). This method involves many steps and yields oily MPGM with low purity. To solve these problems, we resolved (±)-MPGM using an enzyme. Enzymes are now widely recognized as practical catalysts for asymmetric synthesis (12). Previous examples of enzyme-catalyzed enantioselective reactions of glycidate compounds include enantioselective hydrolysis of one ester group in *meso-cis*-oxirane dicarboxylic acid diester (13) and of one enantiomer of racemic glycidyl esters (14). Although these glycidate compounds are stable in aqueous solution, our substrate, (±)-MPGM, is not. The 4-methoxyphenyl group increases the reactivity of the oxirane so that (±)-MPGM decomposes easily in aqueous solution under neutral to acidic conditions. (±)-MPGM is more stable in an organic–aqueous biphasic system. In addition, the two-phase mixtures also simplify separation of reactants and products. Enzyme recovery is also very important for cost reduction in industrial enzyme reactions, but effective recovery is difficult from two-phase mixtures due to the formation of stable emulsions. We used the Sepracor mem-

brane reactor based on the principle that permits immobilization of the enzyme and contact between organic and aqueous phases for both efficient hydrolysis and simple separation of products. Furthermore, we overproduced the enzyme using mutation, optimization of culture medium, and recombinant DNA techniques.

Our objectives were production of (−)-MPGM without chemical degradation of (±)-MPGM, construction of a lipase-hyperproducing strain, an effective isolation of (−)-MPGM from an organic–aqueous biphasic system, and a repeated use of enzyme. The screening of hydrolytic enzymes, overproduction of the enzyme, optimization of the enzyme reaction, and application of membrane reactor are reported herein.

SCREENING FOR ENANTIOSELECTIVE HYDROLASES

Screening Method

The instability of (±)-MPGM in an aqueous solution is a serious problem in the enzymatic resolution of (±)-MPGM. However, as shown in Table 1, (±)-MPGM is both more stable and more soluble in a mixture of toluene–water. Chemical hydrolysis of the oxirane ring in (±)-MPGM increases at a pH lower than 8, yielding 2,3-dihydroxy-3-(4-methoxyphenyl)propanoic acid methyl ester. We used the toluene–water biphasic system at pH 8 to screen for hydrolase with high activity and enantioselectivity toward (±)-MPGM (15). Because the reaction system contains an organic solvent, we focused on lipases that are more stable in organic solvents and can act on insoluble substrates such as oils by adsorbing to the oil droplet (16). (+)-MPGM hydrolyzing activity is assayed using the emulsion method (17) shown in Figure 2. One unit of esterase activity is defined as the amount of enzyme that hydrolyzes 1 μmol (+)-MPGM/min. One unit of lipase activity is defined as the amount of enzyme that hydrolyzes 1 μmol of ester group from 2,3-dimercaptopropan-1-ol tributyrate, as measured with the Lipase Kit S (18). The enzyme activity in "Overproduction of Enzyme" is shown by the lipase activity. *E*-values are calculated from the enantiomeric excess of the recovered (−)-MPGM and extent of conversion (19) and are not corrected for chemical hydrolysis of the methyl ester.

Table 1. The Solubility and Stability of (±)-MPGM in Organic Solvents

Organic solvent (organic solvent:water = 1:1)	Solubility (mg/mL)	Stability (%)
Water only (suspension)	<0.1	10
N,*N*-Dimethylformamide	>20	29
Ethanol	>20	21
Acetone	>20	70
Tetrahydrofuran	>20	85
Methyl isobutyl ketone	219	90
Ethyl acetate	560	90
n-Butyl acetate	267	90
tert-Butyl methyl ether	79	95
Toluene	251	95
Carbon tetrachloride	205	95
p-Xylene	171	95

Note: Four milligrams of (±)-MPGM is dissolved in 2 mL of various organic solvent–water (1:1) biphasic systems, and the biphase is incubated at 30 °C for 24 h. The stability of (±)-MPGM is expressed as a percentage of remaining (±)-MPGM. The solubility of (±)-MPGM is expressed as the concentration of dissolved (±)-MPGM in organic solvent at 30 °C.

Asymmetric Hydrolysis of (±)-MPGM by Various Microorganisms

Among the enzymes from 700 type strains and 30 commercially available enzymes, 453 enzymes hydrolyzed the ester of (±)-MPGM. Hydrolytic enzymes for (±)-MPGM were widely distributed in bacteria, yeasts, and molds. Many favored the (+) enantiomer of MPGM. The *E*-values of 12 enzymes were more than 10, and *Serratia marcescens* lipase has the highest enantioselectivity (E = 135) and hydrolyzing activity, as shown in Table 2. The ratio of the esterase activity to the hydrolyzing activity on olive oil is the highest for *S. marcescens* lipase. This indicates that *S. marcescens* lipase is the most effective enzyme for hydrolysis of (+)-MPGM. In addition, since the *E*-value includes chemical degradation of (±)-MPGM, it shows a lower value than the true value of the enzyme. *S. marcescens* lipase does not enzymatically hydrolyze (−)-MPGM; the *E*-value of this lipase itself is considered to be infinity (17).

Purification and Characterization of *S. marcescens* Lipase

Three main protein bands exist in the supernatant after centrifugation of the culture broth. The molecular masses of these bands are 62,000, 51,000, and 40,000 Da on SDS-PAGE. The hydrolytic activity toward (+)-MPGM exists in the 62,000-Da band. The esterase can be purified from culture supernatant (about 21-fold) at a yield of 20%, and its (±)-MPGM hydrolyzing activity is 570 units/mg protein. Esterase corresponds to 5% of the extracellular proteins in cells grown with dextrin as the carbon source and Meast S (beer yeast extract) as the nitrogen source. The activity of the purified enzyme is 3,100 units/mg protein for olive oil, and 520,000 units/mg protein for 2,3-dimercaptopropan-1-ol tributyrate. Because this esterase selectively hydrolyzes the 1,3-positional ester of triglyceride, the esterase is a lipase. The protein exhibiting a molecular mass of 51,000 on SDS-PAGE is a protease with hydrolytic activity toward casein. The protease does not have (±)-MPGM hydrolyzing activity.

(±)-MPGM as well as other esters having different alcohol moieties may be used as a key intermediate in the synthesis of diltiazem. The enantioselectivity of the lipase did not change upon altering the chain length of the alcohol moiety, except for the acid ethyl ester. (±)-MPGM crystallizes more easily from toluene and is less expensive than other esters, so it was selected for preparing enantiomerically pure 3-(4-methoxyphenyl)glycidic acid ester (17).

OVERPRODUCTION OF ENZYME

Cloning of *S. marcescens lipA* Gene

S. marcescens produces the lipase at a level of 3×10^3 units/mL when grown in LB medium. The amount of lipase increased 50-fold by optimizing the culture condition (15),

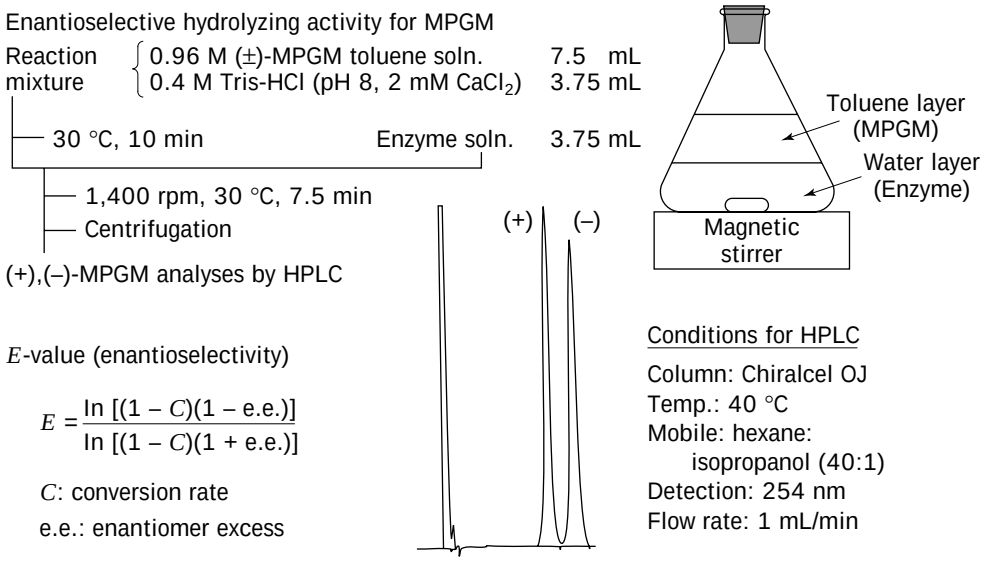


Figure 2. Analytical methods of esterase activity and enantioselectivity.

Table 2. Asymmetric Hydrolysis of (±)-MPGM by Enzymes from Various Microorganisms

Enzyme source[a]	Enzyme used[b] (mg)	Esterase activity (units/mg protein)	Conversion (%)	e.e. (%)	E^c
Serratia marcescens Sr41 8000[d]	1.4 mL	36	48.2	89.0	135
Pseudomonas putida ATCC 17426[d]	3.75 mL	18	44.2	75.1	85
Corynebacterium primorioxydans ATCC 31015[d]	3.75 mL	16	44.6	75.6	73
Pseudomonas mutabilis ATCC 31014[d]	3.75 mL	13	44.7	74.7	57
Corynebacterium alkanolyticum ATCC 21511[d]	3.75 mL	18	45.8	77.4	53
Candida cylindracea (Lipase OF)	4.0	34	50.0	82.2	26
Mucor javanicus (Lipase M)	150	0.82	48.0	75.0	22
Bacillus licheniformis (Alcalase)	1.9 mL	0.91	41.2	57.0	17
Rhizopus japonicus (Lipase saiken 100)	100	2.3	42.8	58.4	15
Bacillus subtilis (Protease 27)	300	–	45.1	62.6	14
Mucor miehei (Palatase M)	0.25 mL	27	45.3	60.3	12
Pseudomonas sp. (Lipase P)	300	0.20	49.8	68.6	11

[a]Trade name of the commercially available enzyme is shown in parentheses.
[b]The extent of hydrolysis (%) of (±)-MPGM and the optical purity (%) of the recovered (−)-MPGM are examined using this amount of enzymes.
[c]*E*-value $\{E = \ln[(1 - C)(1 - \text{e.e.})]/\ln[(1 - C)(1 + \text{e.e.})]\}$ is calculated on the basis of the enantiomeric excess of the recovered (−)-MPGM (e.e.) and extent of conversion (*C*).
[d]Cells are grown in the shaking flasks (500 mL) with 2% Meast S, 1% dextrin, 0.2% $(NH_4)_2SO_4$, 0.1% kH_2PO_4, 0.05% $MgSO_4 \cdot 7H_2O$, 0.01% $CaCl_2 \cdot 2H_2O$, 0.001% $FeSO_4 \cdot 7H_2O$, 0.5% Tween 80, and 0.1% Colorin 102 (pH7) at 30 °C for 18 h (140 rpm). The esterase activity in the culture supernatant is measured.

and 2.5-fold as a result of mutation using *N*-methyl-*N*-nitro-*N*-nitrosoguanidin. Extracellular lipase activity reaches 400×10^3 units/mL of culture medium when the lipase-hyperproducing mutant is cultured in a fed-batch culture on proline. To further increase production of the lipase, the *lipA* gene, encoding the lipase, was cloned from *S. marcescens*. The nucleotide sequence of the *lipA* gene showed a major open reading frame encoding a 64.9-kDa protein of 613 amino acid residues; the deduced amino acid sequence contained a lipase consensus sequence, GXSXG. The *S. marcescens* lipase has 66 and 56% homologies with the lipases of *Pseudomonas fluorescens* B52 (20) and *P. fluorescens* SIK W1 (21), respectively, but does not show any overall homology with other lipases. The *Escherichia coli* cells carrying the *lipA* gene do not secrete the lipase into the medium. The *S. marcescens* lipase lacks a conventional N-terminal signal sequence and is not subjected to any processing at either the N-terminal or C-terminal regions. A specific short region, GGXGXDXXX, which is often found in secretory proteins having no N-terminal signal peptide, is observed in the amino acid sequence. The region is characterized as a glycine- and aspartic acid–rich region. Thus, these results reveal that *S. marcescens* lipase is secreted by the N-terminal-signal-peptide-independent system (22).

Secretion Mechanism of *S. marcescens* Lipase

The 20-kb *Sac*I DNA fragment allowing the lipase secretion was cloned from *S. marcescens* by lipase selection phe-

notype based on the clear halo formation by recombinant *E. coli* cells, as shown in Figure 3. The subcloned 6.5-kb *Eco*RV fragment contains three open reading frames, *lipB*, *lipC*, and *lipD*, which are composed of 588, 433, and 437 amino acid residues and constitute an operon (Fig. 3). The *E. coli* cells carrying the *lipA*, *lipB*, *lipC*, and *lipD* genes secreted the *S. marcescens* lipase, but the cells carrying only the *lipA* gene did not. Deletion mutation analysis shows that all three genes are essential for the extracellular secretion of the *S. marcescens* lipase. Thus, the secretion of the lipase requires three proteins: LipB, LipC, and LipD (23). Recent papers reported that the genes encoding the secretion device for the extracellular proteins lacking an N-terminal signal sequence are found in several Gram-negative bacteria and has been categorized as being in the ABC (ATP binding cassette) transporter family: the *hlyB*, *hlyD*, and *tolC* genes (24–27) for *E. coli* hemolysin; the $prtD_{EC}$, and $prtE_{EC}$, and $prtF_{EC}$ genes (28) for *Erwinia chrysanthemi* metalloprotease; the *aprD*, *aprE*, and *aprF* genes (29) for *Pseudomonas aeruginosa* alkaline protease; the *cyaB*, *cyaD*, and *cyaE* genes (30) for *Bordetella pertussis* cyclolysin; the *lktB* and *lktD* genes (31) for *P. haemolytica* leukotoxin; and the *hasD*, *hasE*, and *hasF* genes (32,33) for *S. marcescens* heme-binding protein and metalloprotease. The deduced amino acid sequences of the *lipB* and *lipC* genes showed high homology to those of the *E. chrysanthemi* $prtD_{EC}$ and $prtE_{EC}$ genes (56 and 46% identity, respectively), and also the *S. marcescens hasD* and *hasE* genes (54 and 45% identity, respectively). The LipB protein belongs to a member of the ABC protein family, and its putative ATP-binding site was found in the central region of the C-terminal half. No N-terminal signal sequence was identified in the LipB and LipC proteins. The LipB and LipC proteins were revealed to be inner membrane components involved in transport across the membrane. The outer membrane component corresponding to the *E. chrysanthemi* $prtF_{EC}$ gene product was encoded by the *lipD* gene in *S. marcescens* (42% identity). The products of the *lipB*, *lipC*, and *lipD* genes also showed 26, 20, and 22% identity with the *hlyB*, *hlyD*, and *tolC* gene products of *E. coli*, respectively. These results indicates that LipB-LipC-LipD is included in the ABC transporter family and is the secretion device of the *S. marcescens* lipase, as shown in Figure 4.

Overproduction of Lipase by a Recombinant Strain of *S. marcescens*

The *S. marcescens lipA* gene was cloned into the high-copy-number vector pUC19, resulting in pLIPE121 (22). This plasmid was introduced into the cells of *S. marcescens* 8000, a wild-type strain of Sr41. The extracellular lipase activity of the recombinant strain (34) reached 220×10^3 units/mL in the lipase production medium and is about nine fold greater than that of the strain carrying pUC19, as shown in Table 3. However, we expected an even greater increase, based on the high copy number of the plasmid. The *S. marcescens lipB*, *lipC*, and *lipD* genes were cloned into the low-copy-number vector pMW219 (35), resulting in pMWBCD10 (23). The *S. marcescens lipA* gene was cloned into the high-copy-number vector pBR322, resulting in pBRE121 (34). The lipase productivity of *S. marcescens* cells carrying pMWBCD10 and pLIPE121 (or pBRE121) is higher than that carrying the *lipA* plasmid only. The extracellular lipase activity of the hyperproducing strain carrying pMWBCD10 and pBRE121 reaches 420

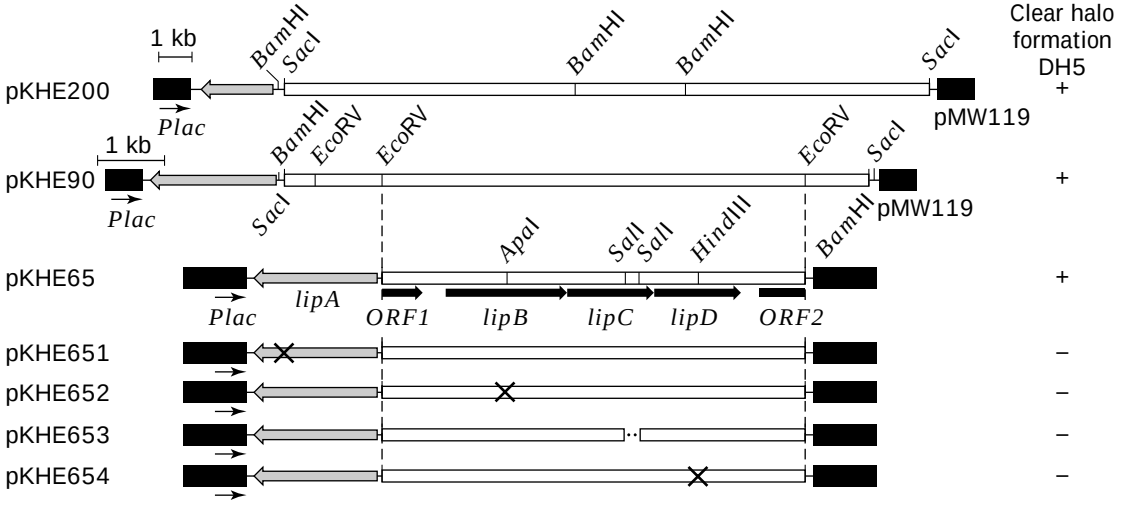


Figure 3. Restriction endonuclease maps of the inserted DNA of pKHE200, deletion derivatives, and the ability of the recombinant plasmids to give a lipase phenotype to the *E. coli* $tolC^+$ strain DH5 (22). The lipase phenotype was detected by formation of a clear halo on a tributyrin agar plate. The thick arrows indicate the open reading frame determined by sequencing. The shaded arrows represent the lipase gene. Open boxes and black bars represent the chromosomal DNA inserts and vector pMW119, respectively. The *lac* promoter carried on the plasmid pMW119 is shown by thin arrows. Only the restriction sites used for subsequent experiments are shown. The crosses symbolize the location of the mutations introduced into the inserted DNA.

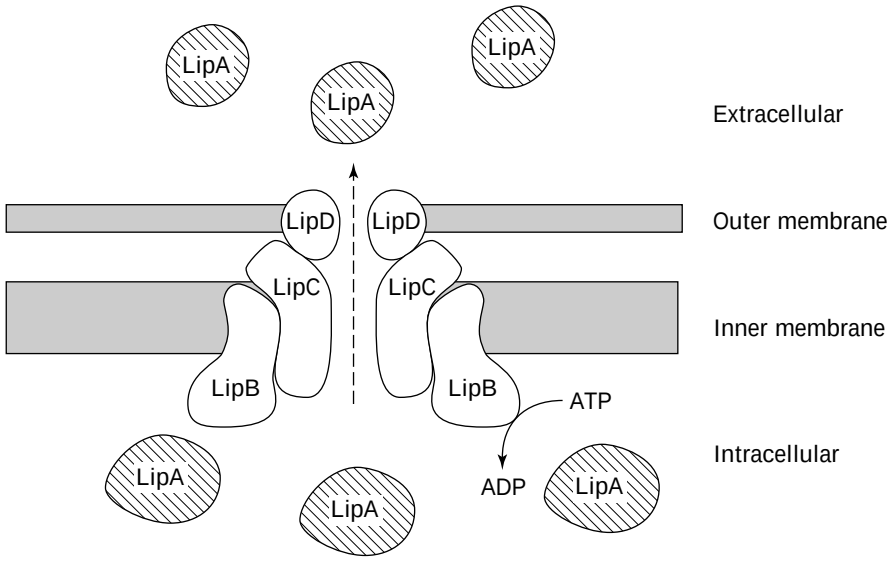


Figure 4. The secretion mechanism for the *S. marcescens* lipase and metalloprotease. LipA indicates the *S. marcescens* lipase. PrtA indicates the *S. marcescens* metalloprotease. LipB, LipC, and LipD are the gene products of the *lipB*, *lipC*, and *lipD* genes, which encode the ABC transporter for *S. marcescens* lipase and metalloprotease.

Table 3. Effect of *lipA* and *lipBCD* Expression on Lipase Production

Plasmid		Growth[a]	Lipase activity[b]
lipA	*lipBCD*	(OD_{660})	($\times 10^{-3}$ units/mL)
–	–	12.2	31.8
pBRE121	–	11.9	258
pBRE121	pMWBCD10	11.2	421
pLIPE121	–	12.2	220
pLIPE121	pMWBCD10	10.9	820

[a]The recombinant strains of *S. marcescens* Sr41 8000 carrying various plasmids are cultivated at 30 °C for 25.5 h in lipase production medium. Antibiotics are added at specified concentrations: ampicillin, 500 μg/mL; kanamycin, 100 μg/mL.
[b]The lipase activity in the culture supernatant is measured.

Table 4. Esterase Activity of *S. marcescens* Lipase in Various Organic Solvent–Water (1:1) Biphasic Systems

Organic solvent (organic solvent:water = 1:1)	Dielectric constant of solvent (ϵ)	Esterase activity (units/mg protein)
Carbon tetrachloride	2.2	85
Toluene	2.2	36
p-Xylene	2.2	31
n-Butyl acetate	5.0	26
Ethyl acetate	6.1	19
Methyl isobutyl ketone	13.1	11
Trichloroethane	7.5	3
Ethanol	24.6	0

$\times$ 10^3 units/mL of the culture medium and is 13-fold greater than that of the wild-type strains. The extracellular lipase activity of the hyperproducing strain carrying pMWBCD10 and pLIPE121 reaches 820 $\times$ 10^3 units/mL of the culture medium and is 25-fold greater than that of the wild-type strains. We are now investigating large-scale lipase production using this recombinant strain.

OPTIMIZATION OF ENZYMATIC REACTION

Conditions for Enzymatic Reaction and Asymmetric Hydrolyzing Activity

Screening for hydrolases, as already described, used the toluene–water biphasic system, which suppresses chemical degradation of (±)-MPGM and readily dissolves (±)-MPGM. In this section, (±)-MPGM hydrolyzing activities of the *S. marcescens* lipase are compared in various biphasic systems (Table 4). The hydrolyzing activities differ with the kind of organic solvent. The lower the dielectric constant of the solvent, the higher the enzyme activity obtained. Among the biphasic systems tested, the carbon tetrachloride–water biphasic system gives the highest hydrolyzing activity. However, carbon tetrachloride is not suitable for industrial application because of its toxicity. Consequently, the toluene–water biphasic system, which gives the second highest hydrolytic rate, was selected as the reaction solvent.

The lipase shows maximum activity at about pH 8. When the lipase was kept at 30 °C for 24 h at different pHs, we found it to be stable between pH 6 and 8. The apparent enantioselectivity of hydrolysis of (±)-MPGM by this lipase in a toluene–water biphasic system decreases at pH values below 8 because the chemical degradation of the oxirane ring in (±)-MPGM increases below pH 8. An Arrhenius plot indicates that the apparent activation energy of the enzymatic hydrolysis is 26 kJ/mol. The thermal inactivation of the enzyme occurs above 50 °C. On the other hand, the apparent activation energy of chemical degradation of the oxirane ring in (−)-MPGM or (+)-MPGM is 66 kJ/mol. These findings show that reaction at a lower temperature minimizes the relative amount of chemical degradation of (±)-MPGM. The best conditions were a toluene–water emulsion at pH8 and 30 °C.

As the toluene to water ratio in the reaction mixture decreases, the rate of enzymatic hydrolysis also decreases. At a toluene to water ratio lower than 1:1, the chemical

degradation of (±)-MPGM increases. Therefore, we use the condition of a toluene to water ratio of 1:1 as a typical reaction for the asymmetric hydrolysis of (±)-MPGM (15).

Effect of Substrates and Solvent on the Stability of the Enzyme

Attempted reuse of the *S. marcescens* lipase showed much slower reaction rates, indicating deactivation of the lipase. When fresh lipase is incubated with the toluene–water biphase without substrate before the enzymatic reaction is started, the progress of reaction is nearly equal that of untreated enzyme, as shown in Figure 5. Therefore, the esterase activity of the reaction decreased only slightly. However, when the first incubations are carried out in the toluene–water biphase containing (±)-MPGM or (−)-MPGM, the progress of reaction shown in Figure 5 is slowly toward that of untreated enzyme. The esterase activity of the second incubation decreases from 36 units/mg protein to 6 and 12 units/mg protein, respectively. Inactivation of the lipase is greatly influenced by MPGM. When the first incubation is carried out in the toluene–water biphase containing (±)-MPGM, the degree of inactivation of the lipase is higher than that of the lipase treated with the toluene–water biphase containing (−)-MPGM. This is because *p*-methoxyphenylacetaldehyde (aldehyde 5) derived from chemical degradation of the hydrolyzed product of (+)-MPGM probably inactivates the lipase. Treatment of the lipase with the toluene–water biphase containing 20 mM and 100 mM aldehyde 5 lowers the esterase activity of the second incubation to 28 and 18 units/mg protein, respectively. Although the lipase is highly stable in a toluene–water biphase, it is unstable under the biphasic system containing (±)-MPGM and aldehyde 5. Hence, the repeated use of the lipase is difficult for this batch process (15).

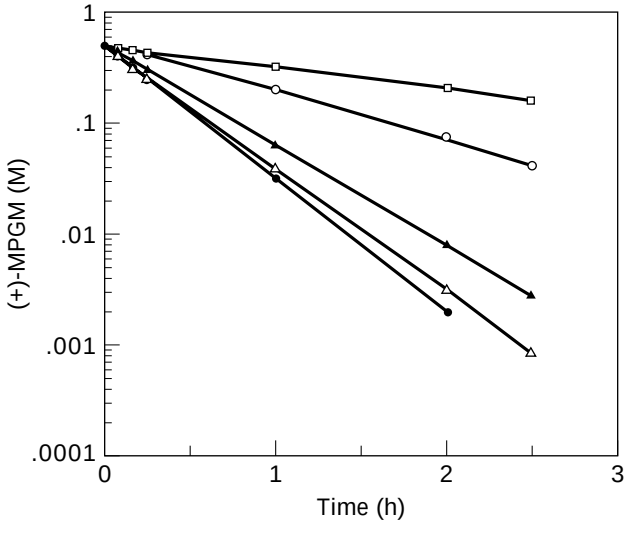


Figure 5. Repeated use of *S. marcescens* lipase for the asymmetric hydrolysis of (±)-MPGM. The toluene solution (50 mL) of 0.96 M (±)-MPGM, 0.48 M (−)-MPGM, or 20 mM *p*-methoxyphenylacetaldehyde is added to 50 mL of 0.2 M Tris-HCl buffer (pH 8, 1 mM $CaCl_2$) containing the lipase. The first incubation is carried out at 30 °C for 3 h with vigorous stirring (900 rpm). After the incubation, the mixture is centrifuged, and the water layer containing the lipase is recovered. The second incubation is carried out with the above water layer (50 mL) containing recovered lipase and a fresh toluene layer (50 mL) containing 0.96 M (±)-MPGM. The concentration of residual (+)-MPGM in the second incubation is measured by the HPLC method described in the analytical methods. Symbols: ●, without first incubation; △, toluene-water (1:1); ▲, aldehyde 5 and toluene–water (1:1); ○, (−)-MPGM and toluene–water (1:1); □, (±)-MPGM and toluene–water (1:1).

Destruction of the Emulsion and Isolation of (−)-MPGM

In the organic–aqueous biphasic systems, the reaction rate is influenced by the outer diffusion and contact area on the biphase, and the highest rate is obtained by the formation of fine emulsion. However, the fine emulsion was unfavorable for industrial purposes because the destruction of the emulsion after the enzymatic reaction and subsequent phase separation are difficult, and the desired product cannot be recovered in high yields. The destruction of the emulsion after the enzymatic reaction and subsequent phase separation is a serious problem for industrial applications of the emulsion bioreactor. Several means involving an electrical method such as sonication, a physical method such as filtration, centrifugation, or heating, and a chemical method such as addition of surfactant or solvent have already been proposed to overcome this problem. Among them, the chemical method using surfactants would be the most favorable for industrial purposes because of the ease of operation. When a 0.01% (w/v) sodium dodecyl sulfate is added to the reaction mixture, the destruction of emulsion and phase separation are accomplished. In the hydrolysis of (±)-MPGM with the lipase, aldehyde 5 is accumulated in the toluene phase, with the desired (−)-MPGM. Aldehyde 5 suppresses crystallization of (−)-MPGM. To isolate (−)-MPGM efficiently, it is necessary to remove the aldehyde 5 from the toluene solution. It is well known that aldehyde compounds generally form a water-soluble adduct through the reaction with bisulfite anion (36). The aldehyde 5 can be removed by transfer to the aqueous phase by the formation of an adduct ($RCHOH \cdot SO_3Na$) with the addition of sodium bisulfite (pH 6). After the phase separation and removal of the aldehyde, the toluene phase is neutralized with sodium hydrogen carbonate, washed with water, dried over anhydrous sodium sulfate, and the (−)-MPGM is crystallized. The yield of (−)-MPGM isolated from (±)-MPGM is 40–43%, and the optical purity of (−)-MPGM is nearly 100% e.e.

USE OF A MEMBRANE REACTOR

Production of (−)-MPGM Using a Membrane Reactor

The enzymatic hydrolysis of (±)-MPGM proceeds efficiently in a conventional emulsion reactor with a toluene–aqueous biphasic system (15). However, destruction of the emulsion and separation of the biphase are required for effective isolation of the product. Enzyme stability and recovery are also required for its repeated use. In particular, the recovery of enzyme is very important for cost reduction

in industrial processes involving enzyme reactions. However, efficient recovery is difficult in the case of a stable emulsion. We proposed an immobilized enzyme system to solve these disadvantages of the emulsion reactor.

With respect to the immobilization of lipase, various methods have already been proposed (37–40). When these lipase-immobilized reactors are utilized, lipase can be efficiently immobilized. However, it is difficult to achieve reaction and product separation simultaneously. Furthermore, because the hydrolysis of oily fat by lipase is an interfacial reaction involving the adsorption of enzyme to the surface of oil droplets (16), the activities of most immobilized lipases are low compared with other immobilized enzymes (39). A membrane reactor permits both physical adsorption of enzyme and contact between the organic and aqueous phases; this configuration enables efficient hydrolysis and simplifies product separation. Since a pioneering work on the membrane reactor with liquid–liquid contact mode by Hoq et al. (41), several membrane bioreactors with industrial promise have been developed (42,43). We used the ultrafiltration membrane reactor technology commercialized by Sepracor, Inc., in Massachusetts, on the basis of Matson's investigation (44).

Structure of the Membrane and Reaction Conditions

The membrane reactor used in this study is a commercially available Sepracor membrane reactor model MBR 500-tm (Sepracor Inc., U.S.A.), Figure 6. The membrane module is a hydrophilic hollow-fiber ultrafiltration membrane with a molecular weight cutoff of 50,000. The membrane module contains approximately 8,500 fibers with an internal diameter of 0.21 mm and a wall thickness of 0.05 mm. The effective membrane area is 0.75 m^2. The saturated adsorption amount of proteins is 3.0 g/m^2. The membrane, composed of polyacrylonitrile, is an asymmetric membrane having a tight skin layer on the lumen side and a loose spongy layer on the shell side. Immobilization of enzyme is carried out by ultrafiltration of the lipase solution from the shell side to the lumen side. The lipase is physically immobilized on the spongy layer of the shell side. As shown in Figure 6, the reaction using a membrane reactor is achieved by circulating both toluene solution containing (±)-MPGM in the shell loop and an equimolar solution (pH 8.5) of sodium sulfite in the lumen loop. Flow rates in both loops are maintained at 750 mL/min and 500 mL/min, respectively; these are sufficient to keep mass transfer resistances at a minimum. The pressure in the shell loop is maintained at overpressure at 0.4–0.5 kg/cm^2 over the pressure in the lumen loop, to avoid the filtration of aqueous-phase solution into the shell loop through the hydrophilic membrane. The enzyme reaction is carried out at 20 °C and pH 8.5 by controlling the pH of the aqueous phase. As discussed earlier, 30 °C was selected for the optimum temperature of reaction. As for the lipase-catalyzed asymmetric hydrolysis on a membrane reactor, a long period of reaction was required, and 20 °C was selected in order to suppress chemical degradation of (±)-MPGM. If sodium bisulfite was not added to the aqueous phase in the lumen loop, the aldehyde 5 accumulated in the toluene phase, as shown in Figure 7, and deactivated the lipase. The stability of the lipase is not influenced when the sodium hydrogen sulfite concentration is below 0.1 M at pH 8.5 and 22 °C. From these results, 0.0935 M sodium hydrogen sulfite solution (pH 8.5) was applied to the aqueous phase to prevent the lowering of enzyme activity. Another by-product, methanol, is transferred from the toluene phase to the aqueous phase. Because the concentration of methanol in the aqueous phase was below 0.2% (w/v), the hydrolysis of (±)-MPGM was not influenced (45).

Rate Analysis and Optimization

Time course for hydrolysis of (±)-MPGM in a membrane reactor with *S. marcescens* lipase is shown in Figure 8. The residual concentration of (+)-MPGM decreases with re-

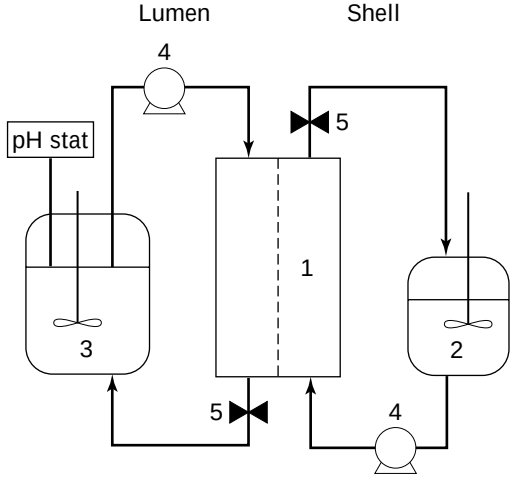


Figure 6. Flow diagram of a membrane reactor. Numerals: 1, hydrophilic membrane; 2, 1.15 L of toluene phase containing 1.15 mol of (±)-MPGM; 3, 12.3 L of aqueous phase containing 1.15 mol of sodium hydrogen sulfite; 4, recycle pump; 5, throttle valve.

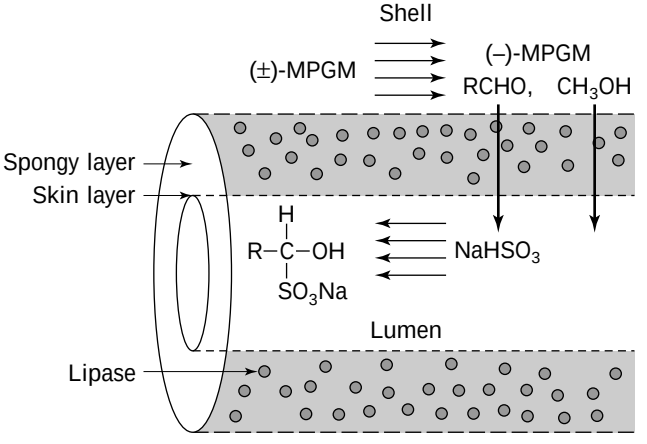


Figure 7. Behavior of products in membrane reactor. Pressure in the shell side is kept at excess compared with pressure in the lumen side. RCHO, *p*-methoxyphenylacetaldehyde; RCHOH · SO_3Na, an adduct formed between sodium hydrogen sulfite and *p*-methoxyphenylacetaldehyde.

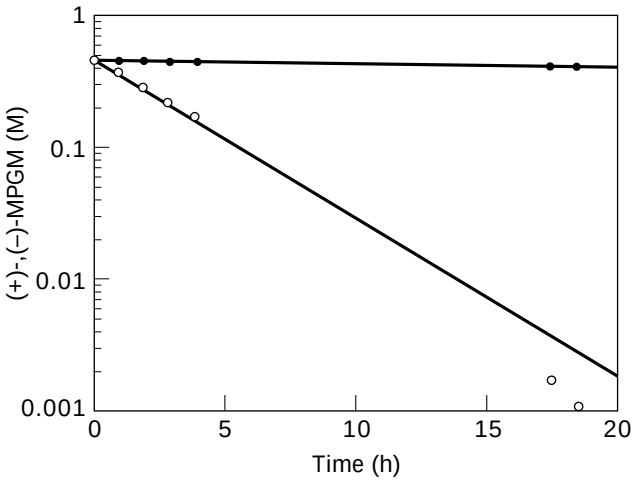


Figure 8. Time course for hydrolysis of (±)-MPGM in a membrane bioreactor with *S. marcescens* lipase. 0.6×10^5 units of lipase is loaded onto the membrane. The reaction is carried out at 22 °C using 1.15 L of 1 M (±)-MPGM toluene solution and 12.3 L of 0.094 M sodium hydrogen sulfite solution.

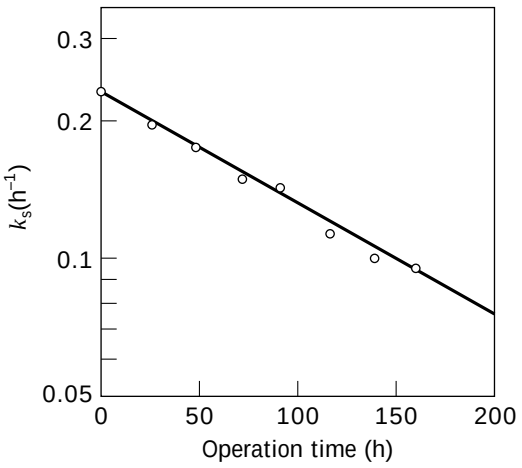


Figure 9. Operational stability of lipase immobilized on a membrane. k_s is calculated from the results of repeated batch reaction; loaded enzyme, 0.8×10^5 units/m^2; (±)-MPGM, 1 M; reaction time in individual batch, 23 h; temp., 22 °C.

action time and reaches less than 1% of initial concentration of (+)-MPGM during the first 15 h. The degradation of (−)-MPGM in merely trace amounts is observed. Thus this hydrolyzing rate of the lipase obeys the following first-order reaction kinetics:

$$r_S = -dC_S/dt = k_S C_S$$
$$r_R = -dC_R/dt = k_R C_R$$

where r_S and r_R are reaction rates for (+)- and (−)-MPGMs (M/h); C_S and C_R are (+)- and (−)-MPGM concentrations (M); k_S and k_R are reaction rate constants (h^{-1}); and t is reaction time (h). k_S and k_R are calculated to be 0.23 h^{-1} and 0.007 h^{-1}, respectively. The ratio of k_S to k_R is about 33, indicating that the lipase from *S. marcescens* has a high enantioselectivity on hydrolysis of (+)-MPGM.

Optimization of the asymmetric hydrolysis reaction with the following parameters was investigated using the reaction rate constant k (h^{-1}) as the indication of reaction. This hydrolyzing reaction is influenced by pH, temperature, phase ratio, and immobilized enzyme amount. The optimized conditions are pH 8.5; temperature 20 °C; phase ratio (=aqueous:toluene phase) 2:1; immobilized enzyme amount 0.8–1.6 $\times 10^5$ unit/m^2 (based on esterase activity). In the industrial application of the membrane bioreactor, one of the most important factors is the stabilization of enzyme. The half-life of enzyme activity was examined in a long-term continuous experiment, as shown in Figure 9. The reaction rate constants for the asymmetric hydrolysis of (±)-MPGM exponentially decrease with time, and the half-life of enzyme activity is 127 h. This half-life is much lower than for enzymes employed for the industrial production of various amino acids (46–48). However, one is about 30 times higher than that in the emulsion reactor. In this membrane bioreactor, one is able to remove inactivated enzymes by a simple back flash. When the enzyme activity in the membrane reactor lowers, one can immobilize new enzyme without regeneration of the membrane. As a result, it has become possible to achieve repeated runs of the enzyme reaction for periods up to a month. When the regeneration of the membrane after the every eight successive batch runs is repeated 10 times over a period of 4 months, the water flux rate of the old membrane is the same as that of new membrane. Thus the production method of (−)-MPGM using the membrane reactor can be industrialized (45).

Scale-Up

Advantages of the production process of (−)-MPGM using the membrane bioreactor are as follows: *(1)* the reaction and products separation are achieved simultaneously, and crystalline (−)-MPGM with a high yield of 40–43% and optical purity of 100% e.e. is obtained by just concentrating the toluene solution; *(2)* the working time taken to isolate crystalline (−)-MPGM is reduced in comparison with that of the emulsion bioreactor, *(3)* stabilization and recovery of the enzyme are achieved in laboratory-scale experiments.

Table 5. Comparison of Reactor Performances for Laboratory- and Pilot-Scale Membrane Reactors

Reactor membrane area	Laboratory scale 0.75 m^2	Pilot scale 7.8 m^2	Pilot scale 57 m^2
L/D^a (−)	3.9	5.5	4.4
Loaded enzyme^b (units/m^2)	0.8×10^5	1.1×10^5	0.9×10^5
k_S value (h^{-1})	0.23	0.23	0.22
Half-life of enzyme activity (h)	127	137	144

[a]L, effective length of the membrane module; D, diameter of the membrane module.
[b]The amount of loaded enzyme is shown as the esterase activity.

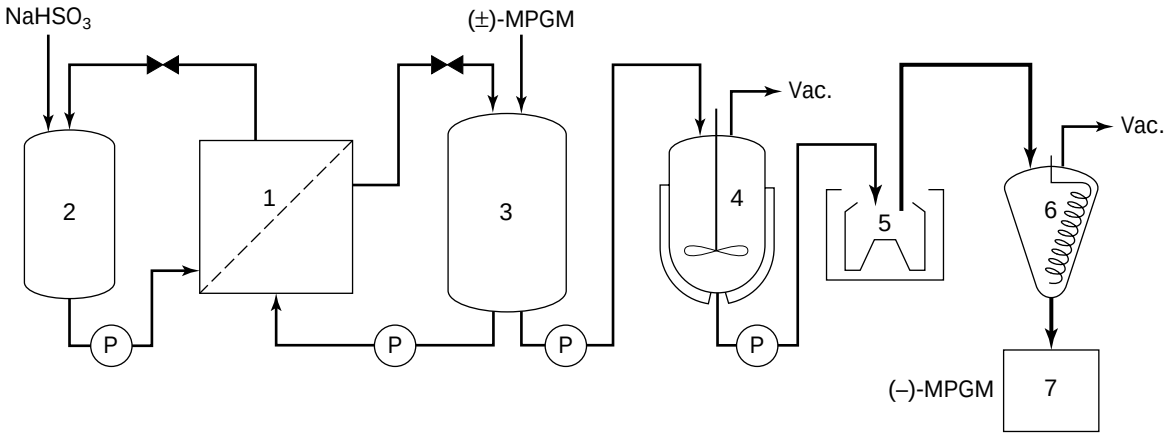


Figure 10. Schematic diagram of industrial production system of (−)-MPGM. Numerals: 1, membrane bioreactor skid; 2, aqueous solution tank; 3, organic solution tank; 4, crysytallizer; 5, centrifuge; 6, dryer; 7, product container.

On the basis of these results, the scale-up experiments were started using membrane modules having an effective membrane area of 7.8 m^2 and 57 m^2. Both membranes were magnified from that of the laboratory scale on the basis of geometrical analogy. The basic strategy of scale-up of the membrane reactor is as follows: *(1)* the flow rate of supplied fluids to the membrane area is kept constant, *(2)* when a lot of modules are used in series, a pressure gradient with respect to flow direction among the modules is not allowed to form. We have tried the reaction in pilot-scale experiments under a liquid line rate comparable to that in the laboratory-scale experiments. The functions of pilot-scale membrane reactors are identical with those of laboratory-scale membrane reactors, as shown in Table 5. From the result, it is believed that scale-up should proceed very smoothly without any trouble. On the basis of the obtained data, we made a design for an apparatus producing (−)-MPGM (Fig. 10). It is estimated that the membrane bioreactor using membrane modules having the effective membrane area of 57 m^2 produces about 50 kg (−)-MPGM m^{-2} $year^{-1}$. The pilot-plant experiments were carried out in collaboration with Sepracor, Inc., and this technique was commercialized in 1993.

BIBLIOGRAPHY

1. K. Abe, H. Inoue, and T. Nagao, *Yakugaku Zasshi* **108**, 716–732 (1988).
2. K.G. Watson, Y.M. Fung, M. Gredley, G.J. Bird, W.R. Jackson, H. Gountzos, and B.R. Matthews, *J. Chem. Soc. Chem. Commun.* 1018–1019 (1990).
3. O. Miyata, T. Shinada, I. Ninomiya, and T. Naito, *Tetrahedron Lett.* **32**, 3519–3522 (1991).
4. A. Schwartz, P.B. Madan, E. Mohacsi, J.P. O'Brien, L.J. Todaro, and D.L. Coffen, *J. Org. Chem.* **57**, 851–856 (1992).
5. K. Matsuki, M. Sobukawa, A. Kawai, H. Inoue, and M. Takeda, *Chem. Pharm. Bull.* **41**, 643–648 (1993).
6. M. Yamamoto, M. Hayashi, M. Masaki, and H. Nohira, *Tetrahedron Asymmetry* **2**, 403–406 (1991).
7. L.T. Kanerva and O. Sundholm, *J. Chem. Soc. Perkin Trans. 1*, 1385–1389 (1993).
8. Eur. Pat. 0362556 A1 (Apr 11, 1990), T. Shibatani, K. Nakamichi, and H. Matsumae (to the Regents of Tanabe Seiyaku).
9. Eur. Pat. 0343714 A1 (May 17, 1989), L.A. Hulshof and J.H. Roskam (to the Regents of Stamicarbon B. Y.).
10. Japanese Pat. Publication 60-13775 (January 24, 1985), H. Sawai, J. Yoshinaga, Y. Ikuta, and K. Nakamura (to the Regents of Sawai Seiyaku).
11. Japanese Pat. Publication 60-13776 (January 24, 1985), H. Sawai, J. Yoshinaga, and T. Masagaki (to the Regents of Sawai Seiyaku).
12. J.B. Jones, *Tetrahedron* **42**, 3351–3403 (1986).
13. G. Sabbioni and J.B. Jones, *J. Org. Chem.* **52**, 4565–4570 (1987).
14. U.S. Pat. 4,732,853 (March 22, 1988), G.M. Whitesides and W. Ladner (to the Regents of the Harvard University).
15. H. Matsumae, M. Furui, and T. Shibatani, *J. Ferment. Bioeng.* **75**, 93–98 (1993).
16. Y.J. Wang, J.Y. Sheu, F.F. Wang, and J.-F. Shaw, *Biotechnol. Bioeng.* **31**, 628–633 (1988).
17. H. Matsumae and T. Shibatani, *J. Ferment. Bioeng.* **77**, 152–158 (1994).
18. S. Kurooka, S. Okamoto, and M. Hashimoto, *J. Biochem.* **81**, 361–369 (1977).
19. C.-S. Chen, Y. Fujimoto, G. Girdaukas, and C.J. Sih, *J. Am. Chem. Soc.* **104**, 7294–7299 (1982).
20. K. Tan and K.J. Miller, *Appl. Environ. Microbiol.* **58**, 1402–1407 (1994).
21. G.H. Chung, Y.D. Lee, G.H. Jeohn, O.J. Yoo, and J.S. Rhee, *Agric. Biol. Chem.* **55**, 2359–2365 (1991).
22. H. Akatsuka, E. Kawai, K. Omori, S. Komatsubara, T. Shibatani, and T. Tosa, *J. Bacteriol.* **176**, 1949–1956 (1994).
23. H. Akatsuka, E. Kawai, K. Omori, and T. Shibatani, *J. Bacteriol.* **177**, 6381–6389 (1995).
24. M.A. Blight and I.B. Holland, *Mol. Microbiol.* **4**, 873–880 (1990).
25. N. Mackman, J.-M. Nicaud, L. Gray, and I.B. Holland, *Curr. Top. Microbiol. Immunol.* **125**, 159–181 (1986).

26. W. Wagner, M. Vogel, and W. Goebel, *J. Bacteriol.* **154**, 200–210 (1983).

27. C. Wandersman and P. Delepelaire, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4776–4780 (1983).

28. S. Letoffe, P. Delepelaire, and C. Wandersman, *EMBO J.* **9**, 1375–1382 (1990).

29. J. Guzzo, F. Duong, C. Wandersman, M. Murgier, and A. Lazdunski, *Mol. Microbiol.* **5**, 447–453 (1991).

30. L. Gilson, K. Mahanty, and R. Kolter, *EMBO J.* **9**, 3875–3884 (1990).

31. C.A. Strathdee and R.Y.C. Lo, *J. Bacteriol.* **171**, 916–928 (1989).

32. S. Letoffe, J.M. Ghigo, and C. Wandersman, *J. Bacteriol.* **176**, 5372–5377 (1994).

33. R. Binet and C. Wandersman, *Mol. Microbiol.* **22**, 265–273 (1996).

34. H. Akatsuka, E. Kawai, K. Omori, S. Komatsubara, and T. Shibatani, *J. Ferment. Bioeng.* **81**, 115–120 (1996).

35. K. Yamaguchi and Y. Masamune, *Mol. Gen. Genet.* **200**, 362–367 (1985).

36. T. Furutani, M. Furui, T. Mori, and T. Shibatani, *Appl. Biochem. Biotechnol.* **59**, 319–328 (1996).

37. R.B. Lieberman and D.F. Ollis, *Biotechnol. Bioeng.* **17**, 1401–1419 (1975).

38. Y. Kimura, A. Tanaka, K. Sonomoto, T. Nihira, and S. Fukui, *Eur. J. Appl. Microbiol. Biotechnol.* **17**, 107–112 (1983).

39. J.-F. Shaw, R.-C. Chang, F.F. Wang, and Y.J. Wang, *Biotechnol. Bioeng.* **35**, 132–137 (1990).

40. C.O. Ibrahim, H. Saeki, N. Nishio, and S. Nagai, *Agric. Biol. Chem.* **52**, 99–105 (1988).

41. M.M. Hoq, T. Yamane, and S. Shimizu, *J. Am. Oil Chem. Soc.* **61**, 776–781 (1984).

42. F. Taylor, C.C. Panzer, J.C. Craig, Jr., and D.J. O'Brien, *Biotechnol. Bioeng.* **28**, 1318–1322 (1986).

43. W. Ponk, P.J.A.M. Kerkhof, C. van Helden, and K. van't Riet, *Biotechnol. Bioeng.* **32**, 512–518 (1988).

44. S.L. Matson and J.A. Quinn, *Ann. N.Y. Acad. Sci.* **469**, 152–165 (1986).

45. H. Matsumae, M. Furui, T. Shibatani, and T. Tosa, *J. Ferment. Bioeng.* **78**, 59–63 (1994).

46. T. Sato, T. Mori, T. Tosa, I. Chibata, M. Furui, K. Yamashita, and A. Sumi, *Biotechnol. Bioeng.* **17**, 1797–1804 (1975).

47. M. Furui and K. Yamashita, *J. Ferment. Technol.* **61**, 587–591 (1983).

48. M. Senuma, O. Otsuki, N. Sakata, M. Furui, and T. Tosa, *J. Ferment. Bioeng.* **67**, 233–237 (1989).

DILTIAZEM SYNTHESIS, MICROBIAL ASYMMETRIC REDUCTION

TAKUO NISHIDA
TAKEJI SHIBATANI
Tanabe Seiyaku Co., Ltd.
Osaka, Japan

KEY WORDS

Asymmetric reduction
Baker's yeast
Biocatalysis
(2*S*,3*S*)-2,3-Dihydro-3-hydroxy-2-(4-methoxyphenyl)-1,5-benzothiazepin-4(5*H*)-one
Diltiazem
Methyl (2*R*,3*S*)-3-(4-methoxyphenyl)glycidate
Methyl (2*RS*,3*S*)-2-chloro-3-hydroxy-3-(4-methoxyphenyl)propionate
Microbial conversion
Microbial reduction

OUTLINE

INTRODUCTION

Diltiazem hydrochloride, (2*S*,3*S*)-3-acetoxy-5-[2-(dimethylamino)ethyl]-2,3-dihydro-2-(4-methoxyphenyl)-1,5-benzothiazepin-4(5*H*)-one hydrochloride (**7**; reaction products are described in Figs. 1 and 2), is a typical calcium channel blocker and is used clinically as an effective antianginal and antihypertensive agent in more than 100 countries (1). Many methods have been investigated for the synthesis of this optically active medicine (2–11). On a commercial basis, the standard process of nine chemical synthetic steps had been used for a long time. Later, lipase-catalyzed asymmetric resolution of a key intermediate, methyl *trans*-3-(4-methoxyphenyl)glycidate [(±)-MPGM] was developed, and the synthetic process was reduced to five steps (Fig. 1) (12,13). As a result, diltiazem has been produced on a commercial basis by this method since 1993. However, one drawback of the method is that the yield of the desired optically active intermediate, methyl (2*R*,3*S*)-3-(4-methoxyphenyl)glycidate [(−)-MPGM], does not exceed essentially 50% of the starting material, which means that half of the substrate goes to waste.

The microbial reduction of carbonyl compounds is a useful method in which 100% theoretical yield can be expected (14), and the procedure for the reaction is very simple. Its use has been limited to preparation of chiral alcohols on a large scale because of the following disadvantages: *(1)* a relatively low concentration of substrate in the reaction medium, *(2)* a tedious procedure for isolation and purifi-

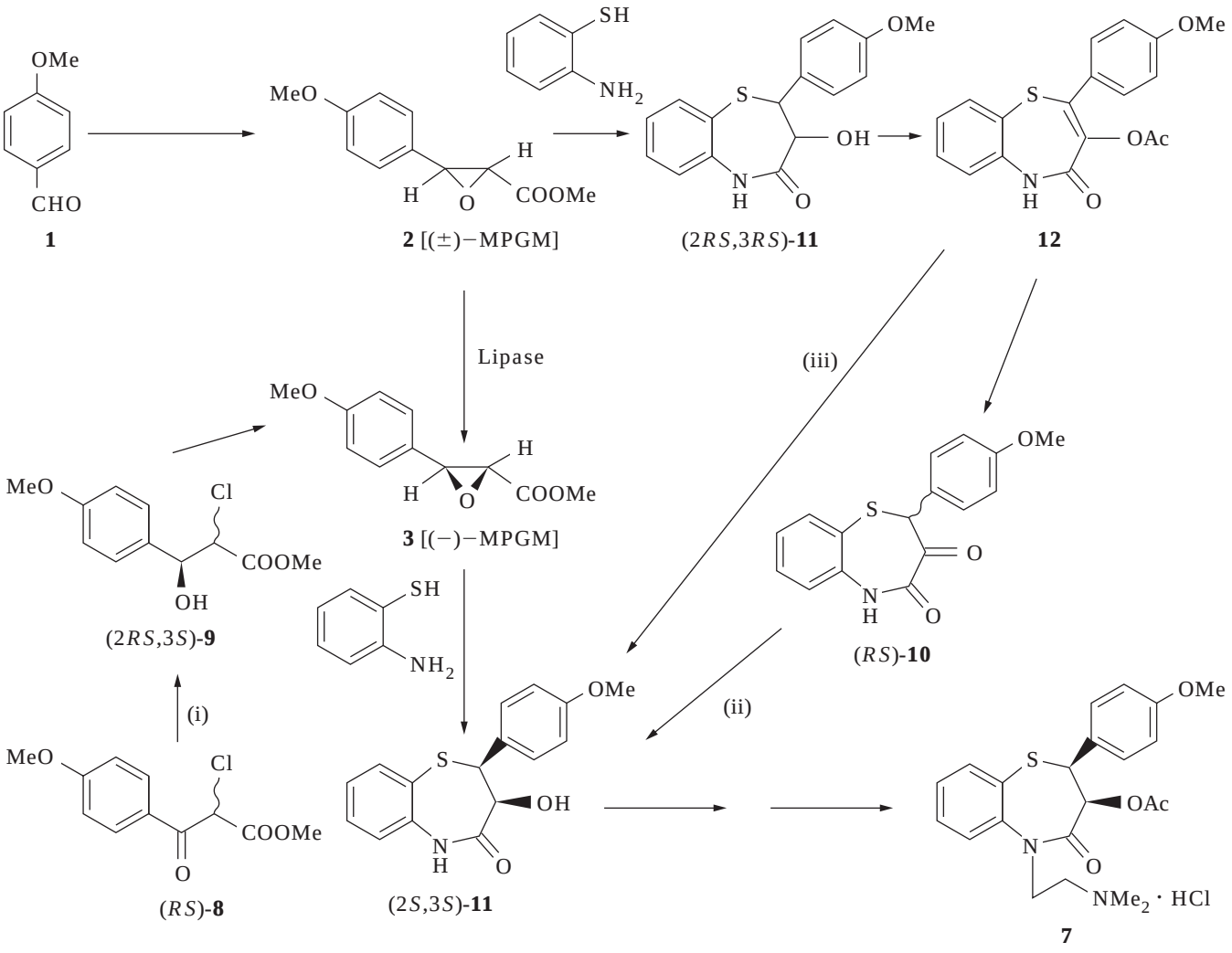


Figure 1. Synthetic process for diltiazem hydrochloride.

cation of the product, and sometimes *(3)* a relatively low optical purity of the product (15).

In order to eliminate these disadvantages, microorganisms having a high enzymatic activity and a high enantioselectivity were sought to produce the optically active diltiazem intermediates for three production methods: the reduction of methyl (*RS*)-2-chloro-3-(4-methoxyphenyl)-3-oxopropionate [(*RS*)-**8**] to methyl (2*RS*,3*S*)-2-chloro-3-hydroxy-3-(4-methoxyphenyl) propionate [(2*RS*,3*S*)-**9**] [Fig. 1(i)], the reduction of (*RS*)-2-(4-methoxyphenyl)-1,5-benzothiazepin-3,4(2*H*,5*H*)-dione [(*RS*)-**10**] to (2*S*,3*S*)-2,3-dihydro-3-hydroxy-2-(4-methoxyphenyl)-1,5-benzothiazepin-4(5*H*)-one [(2*S*,3*S*)-**11**] [Fig. 1(ii)]; and the microbial conversion of 3-acetoxy-2-(4-methoxyphenyl)-1,5-benzothiazepin-4(5*H*)-one (**12**) to (2*S*,3*S*)-**11** [Fig. 1(iii)].

Among these three routes, the second is the most promising; (2*S*,3*S*)-**11** was obtained from (*RS*)-**10** in a high yield with high enantioselectivity from baker's yeast, and was able to be easily separated from the reaction mixture. This route was also introduced to the production of (2*S*,3*S*)-**11** in a high concentration for industrial application.

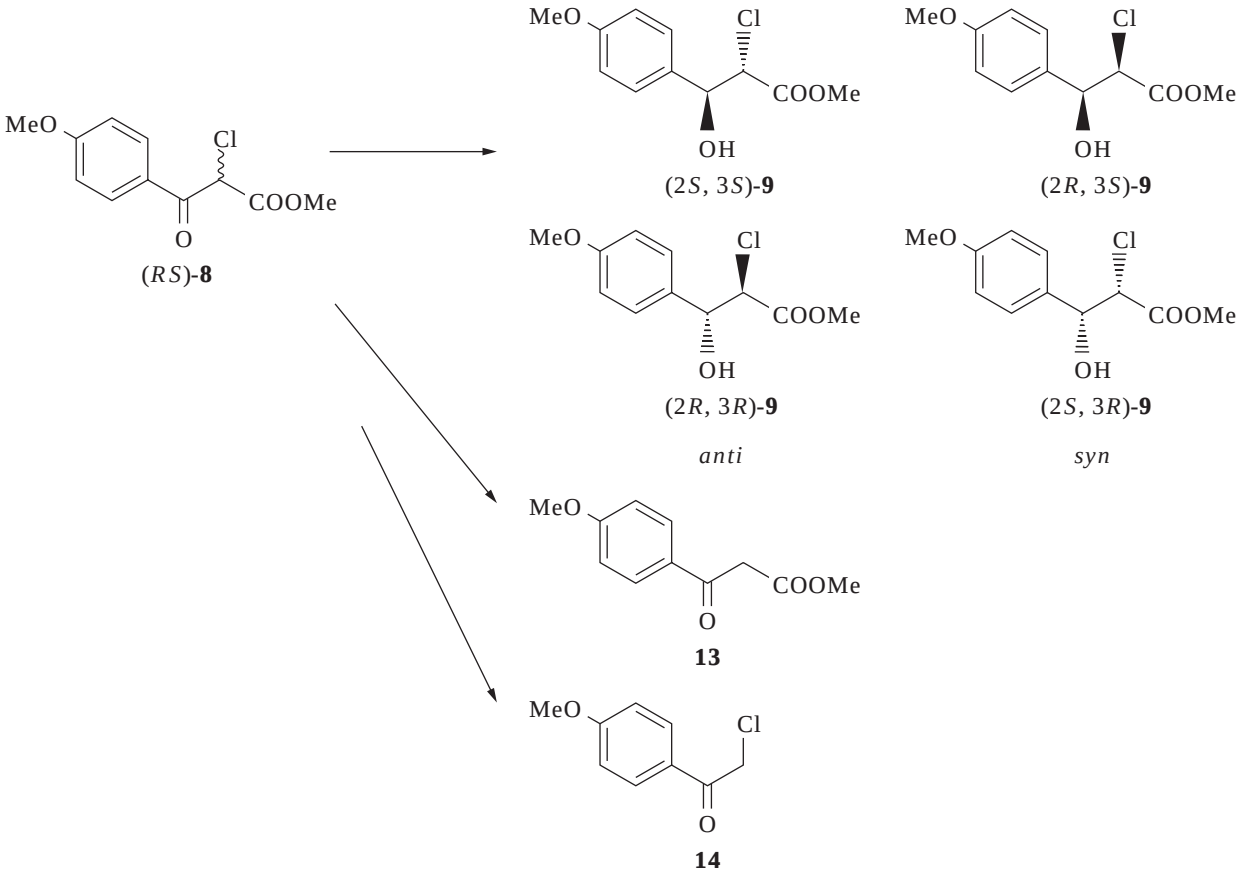
SCREENING OF MICROORGANISMS FOR THE PRODUCTION OF OPTICALLY ACTIVE DILTIAZEM INTERMEDIATE

To successfully apply microbial reduction to industrial scale production it is important to have microorganisms with a high enzymatic activity and a high enantioselectivity. Three production routes for optically active diltiazem intermediates were examined (Fig. 1).

Microbial Reduction of Methyl (*RS*)-2-Chloro-3-(4-methoxyphenyl)-3-oxopropionate [(*RS*)-8]

Methyl (2*R*,3*S*)-3-(4-methoxyphenyl)glycidate [(−)-MPGM] is an important optically active intermediate for the industrial synthesis of diltiazem (12,13). (−)-MPGM can be also obtained selectively from methyl (2*RS*,3*S*)-2-chloro-3-hydroxy-3-(4-methoxyphenyl)propionate [(2*RS*, 3*S*)-**9**] by treating it with NaOMe in methanol; there is epimerization in the second position of **9** (9,16) (Fig. 1).

Table 1. Asymmetric Reduction of (*RS*)-8 by Various Microorganisms



Microorganism	Product **9** Yield[a] (%)	*anti*/*syn*	Absolute configuration	% e.e.[a]
Absidia corymbifera IFO 4009	18	1/>99	*syn*-2*S*,3*R*	>99
Cryptococcus laurentii OUT 6027	43	80/20	*anti*-2*R*,3*R*	>99
			syn-2*S*,3*R*	99
Mucor ambiguus IFO 6742	58	1/>99	*syn*-2*S*,3*R*	>99
Mucor hiemalis OUT 1045	20	12/88	*anti*-2*R*,3*R*	68
			syn-2*S*,3*R*	>99
Mucor janssenii OUT 1050	23	5/95	*anti*-2*R*,3*R*	>99
			syn-2*S*,3*R*	>99
Mucor javanicus IFO 4570	40	7/93	*anti*-2*R*,3*R*	>99
			syn-2*S*,3*R*	>99
Trichoderma viride OUT 4642	7[b]	1/2	*anti*-2*S*,3*S*	87
			syn-2*R*,3*S*	96

The microorganisms were inoculated into test tubes (15 mm diameter × 12 cm) containing 3 mL of medium A[c] or B[d] and cultured at 30 °C for 24 h (yeast and bacteria) or for 72 h (fungi and actinomycetes) with continuous shaking (300 strokes/min). The cells were collected by centrifugation and suspended in 3 mL McIlvaine buffer (pH 5) containing 5% glucose. Then 6 μL of *N*,*N*-dimethylformamide (DMF) solution containing 6 mg of (*RS*)-**8** was added to the test tube. After the reaction was carried out at 30 °C for 24 h with continuous shaking (300 strokes/min), the reaction products were extracted from the reaction mixture with 2 mL of ethyl acetate. The yield of **9** was calculated by using HPLC equipped with Zorbax CN™ column (Du Pont Company). The e.e. of **9**, which was isolated by silica gel TLC (solvent: *n*-hexane: chloroform: ethyl acetate, 4:1:1) was also calculated by HPLC equipped with Chiralcel OJ™ column (Daicel Chemical Industries, Ltd.).

Source: Ref. 17.

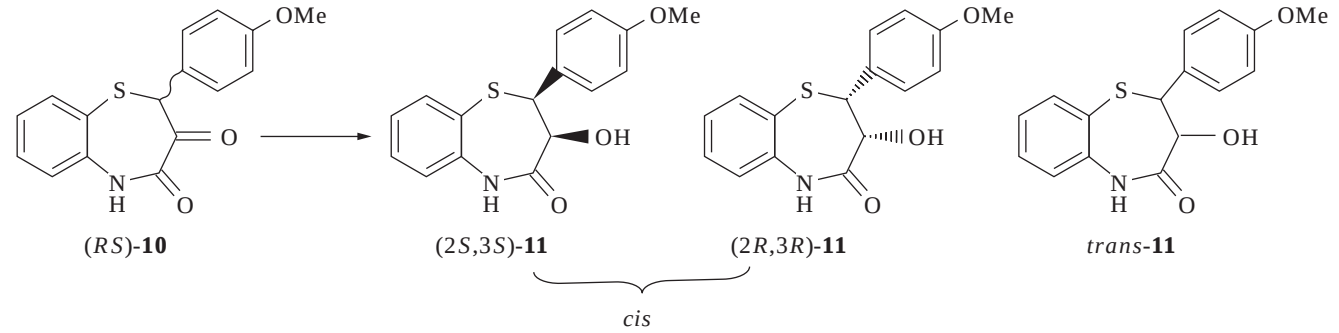
[a]Yield and e.e. of **9** were determined by HPLC.

[b]Isolated yield.

[c]Medium A comprised 5% glucose, 0.1% KH_2PO_4, 0.05% $MgSO_4 \cdot 7H_2O$, 0.05% $CaCl_2 \cdot 2H_2O$, 0.1% yeast extract, 0.1% $(NH_4)_2SO_4$, 0.05% urea, 0.0002% $FeSO_4 \cdot 7H_2O$, 0.0002% $MnCl_2 \cdot 4H_2O$, and 0.0002% $ZnSO_4 \cdot 7H_2O$ (yeast and fungi, pH 6.5; actinomycetes, pH 7.3).

[d]Medium B comprised 1% glucose, 1% Polypepton™ (Nikon Seiyaku Co. Ltd.), 0.1% KH_2PO_4, 0.05% $MgSO_4 \cdot 7H_2O$, 0.05% $CaCl_2 \cdot 2H_2O$, 0.1% yeast extract, 0.1% $(NH_4)_2SO_4$, 0.05% urea, 0.0002% $FeSO_4 \cdot 7H_2O$, 0.0002% $MnCl_2 \cdot 4H_2O$, and 0.0002% $ZnSO_4 \cdot 7H_2O$ (bacteria, pH 7).

Table 2. Asymmetric Reduction of (*RS*)-10 by Various Microorganisms



Microorganism	Product **11** Yield (%)	*cis*/*trans*	Absolute configuration	% e.e.
Brevibacterium ketoglutamicum ATCC 15588	33	98/2	2*S*,3*S*	96
Candida melinii IFO 0747	50	97/3	2*S*,3*S*	99
Candida saitoana IFO 0380	49	99/1	2*S*,3*S*	>99
Debaryomyces polymorphus IFO 1166	51	99/1	2*S*,3*S*	>99
Pichia capsulata IFO 0721	56	99/1	2*S*,3*S*	>99
Pichia pinus IFO 1793	48	>99/1	2*S*,3*S*	>99
Pseudomonas putida ATCC 17484	27	98/2	2*S*,3*S*	97
Rhizopus arrhizus IFO 5780	22	>99/1	2*R*,3*R*	99
Rhizopus japonicus IFO 4758	11	>99/1	2*R*,3*R*	>99
Rhodosporidium toruloides IFO 1638	36	99/1	2*S*,3*S*	99
Saccharomyces cerevisiae (baker's yeast)	67	95/5	2*S*,3*S*	>99
Saccharomyces cerevisiae IFO 1346	54	98/2	2*S*,3*S*	>99
Streptomyces lavendulae IFO 3145	29	98/2	2*S*,3*S*	>99

The microorganisms were inoculated into test tubes (15 mm diameter × 12 cm) containing 3 mL of medium A or B (same as those described in Table 1) and cultured at 30 °C for 24 h with continuous shaking (300 strokes/min). The cells were collected by centrifugation, and the reaction mixture contained 0.6 mmol McIlvaine buffer (pH 5.0), 0.83 mmol glucose, and cells (150 mg baker's yeast) in a total volume of 3.0 mL. Six microliters of DMF solution containing 0.01 mmol (*RS*)-**10** was added. The reaction was carried out at 30 °C for 24 h with continuous shaking (300 strokes/min); the reaction products were extracted from the reaction mixture with 4 mL of ethyl acetate. Yield and e.e. of **11** were analyzed by HPLCs equipped with ChemcoPak™ column (Develosil ODS-7) (Chemco Scientific Co., Ltd.) and Chiralcel OD™ column (Daicel Chemical Industries, Ltd.), respectively.

Source: Ref. 19.

Microbial reduction of (*RS*)-**8** has been studied for the production of an optically active **9** (17). The screening of microorganisms was carried out in a buffer adjusted to pH 5, which contained 5% glucose, substrate **8** and product **9** are unstable above pH 6. Table 1 summarizes the microorganisms that have the ability to reduce (*RS*)-**8**.

The yield and e.e. of **9** were calculated with HPLC, where the yield stands for the total yield of four optically active isomers of **9**, and the e.e. of **9** was determined for each diastereomer of *anti*-**9** and *syn*-**9**. The terms *syn* and *anti* were defined according to Masamune et al. (18). The microorganisms that produced alcohol **9** with the 3*R* configuration, which can be used for the synthesis of the enantiomer of diltiazem with high enantioselectivity, are in the genera *Absidia*, *Cryptococcus*, and *Mucor*. *Mucor ambiguus* IFO 6742 (IFO: Institute for Fermentation, Osaka, Japan) gives only (2*S*,3*R*)-**9** in 58% yield (>99% e.e.) among the four isomers, through keto–enol tautomerization of substrate (*RS*)-**8**. With reaction conditions of pH 4 and 30 °C, the yield of (2*S*,3*R*)-**9** increases up to 68% (isolated yield), and the e.e. of (2*S*,3*R*)-**9** is over 99%.

Furthermore, screening of microorganisms for the ability to conduct asymmetric reduction of (*RS*)-**8**, revealed that *Trichoderma viride* OUT 4642 (OUT: Department of Biotechnology, Faculty of Engineering, Osaka University, Japan) gives **9** with the 3*S* configuration, which is a key diltiazem intermediate with high enantioselectivity. The production ratio of *syn*-**9** to *anti*-**9** was 2:1, and the e.e. of the (2*R*,3*S*)-**9** and (2*S*,3*S*)-**9** is 96% and 87%, respectively. However, the total yield of (2*R*,3*S*)-**9** and (2*S*,3*S*)-**9** is low (7%) due to the formation of the by-products methyl 3-(4-methoxyphenyl)-3-oxopropionate (**13**) and 2-chloro-1-(4-methoxyphenyl)ethanone (**14**) (Table 1).

When whole cells are used as a catalyst, there are often unwanted side reactions. In this case, biotransformation by the extracted free enzyme is thought to be more effective than whole-cell biotransformation. However, if microbial reduction is performed with whole cell, an industrial device for cofactor recycling must be introduced. Desiring microbial reduction with high enzymatic activity and high enantioselectivity, two whole-cell reactions were investigated.

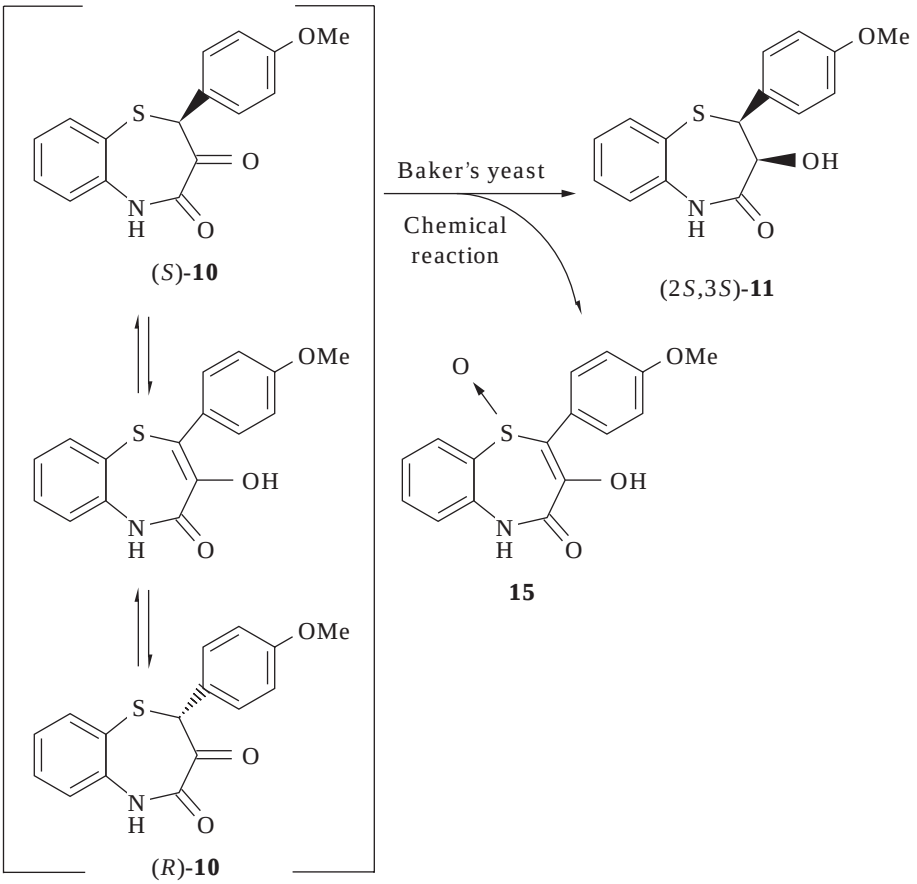


Figure 2. Reduction of (*RS*)-**10** by baker's yeast. *Source*: Ref. 19.

Microbial Reduction of (*RS*)-2-(4-Methoxyphenyl)-1,5-benzothiazepin-3,4(2*H*,5*H*)-dione [(*RS*)-10]

The microbial asymmetric reduction of (*RS*)-2-(4-methoxyphenyl)-1,5-benzothiazepin-3,4(2*H*,5*H*)-dione [(*RS*)-**10**] is a useful method for producing another optically active key intermediate, (2*S*,3*S*)-2,3-dihydro-3-hydroxy-2-(4-methoxyphenyl)-1,5-benzothiazepin-4(5*H*)-one [(2*S*,3*S*)-**11**] (19). Substrate (*RS*)-**10** is easily obtained from (±)-MPGM through three steps of chemical ring closure with aminothiophenol, chemical oxidation, and chemical hydrolysis (Fig. 1) (20).

Table 2 shows microorganisms that have the ability for the asymmetric reduction of (*RS*)-**10** (19). The e.e. of *cis*-**11** was determined by using HPLC.

Microorganisms that produce the key intermediate (2*S*,3*S*)-**11** are widely distributed in the genera *Brevibacterium*, *Candida*, *Debaryomyces*, *Pichia*, *Pseudomonas*, *Rhodosporidium*, *Saccharomyces*, and *Streptomyces*. On the other hand, only *Rhizopus* reduces (*RS*)-**10** to (2*R*,3*R*)-**11**. Among these, baker's yeast (*Saccharomyces cerevisiae*) has the highest ability for the formation of (2*S*,3*S*)-**11**, which is useful for the synthesis of diltiazem. Baker's yeast itself has been widely used as a reagent in organic synthesis (21,22) and the large-scale preparation of, for example, ethyl (*S*)-3-hydroxybutanoate using baker's yeast has already been reported (15).

In the microbial asymmetric reduction of (*RS*)-**10**, baker's yeast strictly recognizes (*S*)-**10**, and (2*S*,3*S*)-**11** forms selectively through keto–enol tautomerization of substrate (Fig. 2). A very small amount of 3-hydroxy-2-(4-methoxyphenyl)-1,5-benzothiazepin-4(5*H*)-one *S*-oxide (**15**) is formed as a by-product of the chemical reaction. In the baker's yeast reduction system, (2*S*,3*S*)-**11** is accumulated with 94% conversion under the following the low substrate concentration conditions: substrate concentration 1 g/L at pH 7 and 42 °C for 24 h. Sometimes, it is difficult to separate a product from the enzymatic reaction mixture. In this case, separation is carried out with the ease as follows: At the end of the reaction, acetone was added as a volume of three times the reaction mixture. The cells in acetone were dried to powder form and easily removed by filtration through Celite™. Then, acetone in the filtrate was evaporated under reduced pressure. The deposited precipitate in the aqueous layer was recovered by filtration and washed with water in the amount of four times the volume of the reaction mixture. The precipitate was redissolved in acetone, and then the solution was dehydrated by the addition of anhydrous magnesium sulfate. After filtration, crude crystals were obtained by evaporation of the organic layer to dryness under reduced pressure. The residue was recrystallized from methanol (10 mL methanol/1 g residue). The isolated yield was 85% based on (*RS*)-**10**, and the e.e. was over 99%.

Microbial Conversion of 3-Acetoxy-2-(4-methoxyphenyl)-1,5-benzothiazepin-4(5*H*)-one (12)

An optically active key intermediate, (2*S*,3*S*)-**11**, can be also obtained by the microbial conversion of **12** (Fig. 1) (23).

For this type biotransformation, the microbial conversion of enol ester using *Pichia farinosa* has been reported (24). The microorganisms that have the ability to catalyze the conversion of **12** into (2*S*,3*S*)-**11**, with excellent enantioselectivity, are summarized in Table 3.

The conversion yield of **12** to **11**, and the e.e. of **11** are calculated on the basis of HPLC analysis, and the isolated yield is based on **12**. No accumulation of **15** (Fig. 2) was observed in this reaction. By employing this microbial conversion, diltiazem can be synthesized in only six steps with a high yield (Fig. 1).

Figure 3 shows the time courses of the microbial conversion of (*RS*)-**10** and **12** using *Bacillus sphaericus* FERM P-6746 (FERM: Fermentation Research Institute, Agency of Industrial Science and Technology, Japan). In both cases, the products are optically active (2*S*,3*S*)-**11** (>99% e.e.) and isolated easily by the same procedure already described. With (*RS*)-**10** as the substrate, the reduction proceeds directly from the beginning of the reaction, whereas with **12**, an obvious lag phase is observed at the beginning of the reaction. Later, however, the yield of (2*S*,3*S*)-**11** from **12** increases to the same level as that from (*RS*)-**10**. Therefore, in the reaction pathway of microbial conversion of **12** to (2*S*,3*S*)-**11**, it is thought that (*RS*)-**10** is an intermediate, that hydrolysis and reduction proceed in turn, and that the hydrolysis of **12** is a rate-determining step. Furthermore, it is also confirmed that the hydrolytic enzyme is induced by **12** and that asymmetry is introduced during the (*RS*)-**10** reduction step. Because all the other microorganisms shown in Table 3 also reduce (*RS*)-**10** to (2*S*,3*S*)-**11** (79–88% conversion, >99% e.e.), the microbial conversion of **12** into (2*S*,3*S*)-**11** is thought to proceed through a similar pathway in these microorganisms.

Another approach to the synthesis of (2*S*,3*S*)-**11** from **12** in the presence of the hydrolytic enzyme from *B. sphaericus* and baker's yeast gives (2*S*,3*S*)-**11**, with a conversion of 65% (99% e.e.).

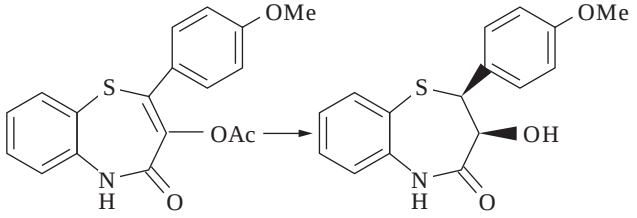
INVESTIGATION FOR INDUSTRIAL APPLICATION

In general, industrial application of enzymatic reduction is difficult due to the low concentration of the substrate in the reaction medium. Among the three reactions catalyzed by microorganisms mentioned earlier, baker's yeast–mediated reduction of (*RS*)-**10** is the most promising, for industrial application if the substrate concentration is increased, because of the high optical purity of the product, easy product isolation, and a simple reaction system (not a plural enzyme system) (25).

Because the substrate (*RS*)-**10**, obtained as crystals by recrystallization from organic solvent, is almost insoluble in an aqueous medium, the reaction was carried out by adding an *N*,*N*-dimethylformamide (DMF) solution of (*RS*)-**10**, and by using ethanol as an energy source under aerobic conditions. Although it is widely accepted that sugar is essential as an energy source, the use of ethanol as the energy source has also been considered because it is clean and efficient compared with sugar (15,26–30).

In the reduction of (*RS*)-**10**, the reduction activity of baker's yeast decreased when DMF was added at a concentration of above 10% (v/v) in order to increase the substrate concentration. By-product, *S*-oxide of the substrate (**15**), also appeared at these high DMF concentrations. Therefore, the large-scale production of (2*S*,3*S*)-**11** by a fed-batch operation on a laboratory scale was examined. A suspension of 11.2 g of baker's yeast in 50 mL of 200 mM ethanol in a 300-mL Erlenmeyer flask was shaken at 30

Table 3. Microorganisms with the Ability to Produce (2*S*,3*S*)-11 from 12



12 (2*S*,3*S*)-**11**

	Product (2*S*,3*S*)-**11**		
Microorganisms	Conversion yield (%)	% e.e.	Isolated yield (%)
Alcaligenes faecalis FERM P-6745	76	99	72
Arthrobacter paraffineus ATCC 21219	76	>99	66
Bacillus sphaericus FERM P-6746	77	>99	74
Rhodococcus sp. ATCC 15592	67	>99	65

The microorganism were cultivated in 500-mL shaking flasks with 100 mL of medium, which comprised 1% glucose, 0.7% Polypepton™ (Nikon Seikayu Co. Ltd.), 0.5% K_2HPO_4, and 0.5% yeast extract (pH 7.2), at 30 °C for 24 h with continuous shaking (140 strokes/min). To each shaking flask was added 1 mL of DMF solution containing 100 mg **12** and 10 mL 1 M potassium phosphate buffer (pH 6.5). The mixtures were shaken at 30 °C (140 strokes/min). After 40 h, the conversion yield and e.e. of (2*S*,3*S*)-**11** were checked by HPLC equipped with Chiralcel OF™ column (Daicel Chemical Industries, Ltd.). The mixtures were shaken for an additional 8 h. According to the method already described, (2*S*,3*S*)-**11** was isolated.

Source: Ref. 23.

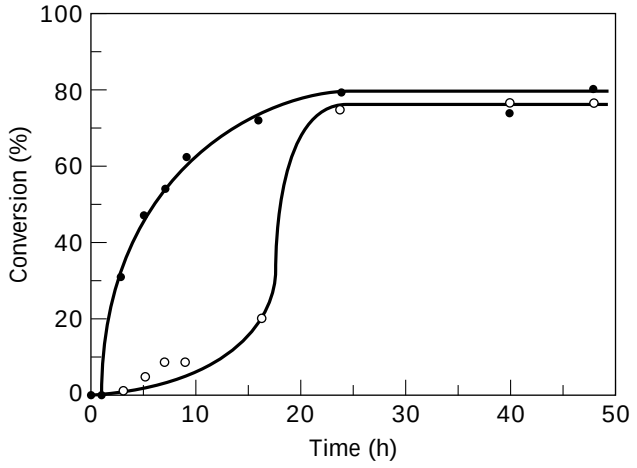


Figure 3. Time courses of (2*S*,3*S*)-**11** production from **12** (○) and (*RS*)-**10** (●) by *B. sphaericus*. *Source*: Ref. 23.

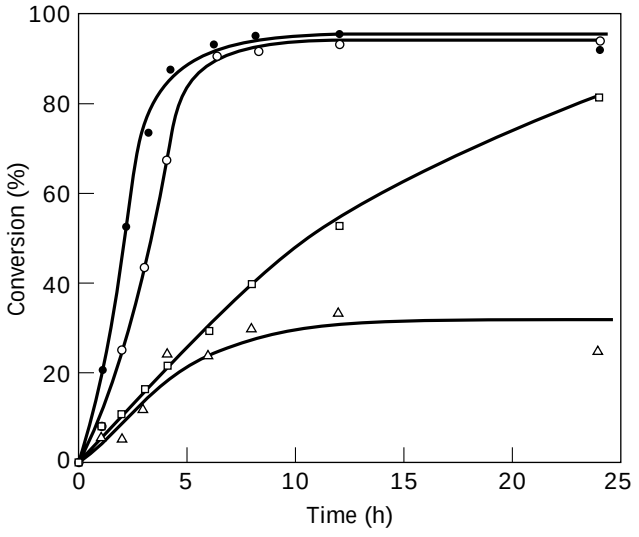


Figure 4. Conversion of (*RS*)-**10** to (2*S*,3*S*)-**11** by baker's yeast. *Source*: Ref. 25. A suspension of baker's yeast (11.2 g) and substrate (*RS*)-**10** in 200 mM ethanol (50 mL) was aerobically shaken at 30 °C at 80 rpm. Substrates using the reaction were as follows: ●, DMF solution (600 μL) containing 100 mg of (*RS*)-**10**; ○, 100 mg of the dry aggregate; □, 100 mg of the fine powder (average grain size of 2.3 μm); △, 100 mg of the crystalline form (average grain size of 21.3 μm). An aggregate of (*RS*)-**10** was prepared by the addition of 6 mL of 0.56 M (*RS*)-**10** in DMF [1.0 g of (*RS*)-**10**] to 100 mL of water for 5 min followed by stirring for 1 h at room temperature. The aggregate was collected by centrifugation at 3,000 rpm for 3 min and washed with water.

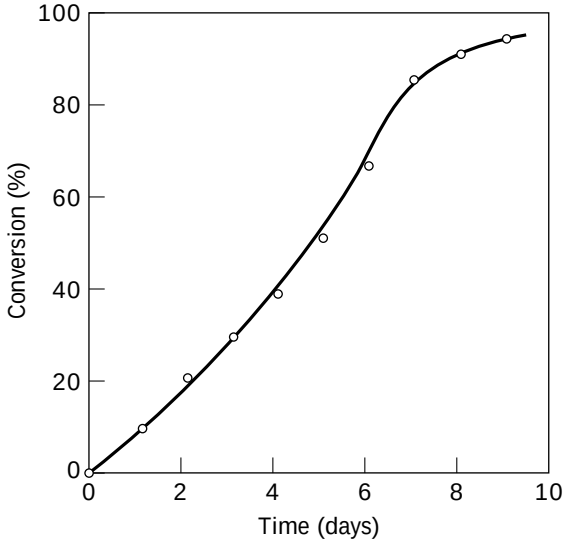


Figure 5. Changes in conversion ratio of (*RS*)-**10** to (2*S*,3*S*)-**11** in a high-concentration reaction. *Source*: Ref. 25.

°C, and 6 mL of DMF solution containing 1.0 g of (*RS*)-**10** (0.56 M) was added in 0.15 mL aliquots every 1 h. Since ethanol was consumed within 12 h, 0.575 mL of ethanol was added three times. After 42 h, the reaction was stopped and the product was isolated and purified. The yield of (2*S*,3*S*)-**11** was 0.83 g (83%), and e.e. was over 99%. By fed-batch operation, substrate concentration increased to 20.0 g/L.

In order to increase the substrate concentration further, as a yeast-mediated reduction would be inhibited by a high concentration of DMF, the reactions were also carried out in an aqueous solution without DMF. As shown in Figure 4, however, in the case of crystalline (*RS*)-**10** (an average grain size of 21.3 μm) in an aqueous solution directly, reduction did not proceed successively. Increasing the surface area of the crystalline (*RS*)-**10** by grinding (an average grain size of 2.3 μm) was not very effective in increasing the reduction rate. On the other hand, when a DMF solution of (*RS*)-**10** was added to water with stirring, a dispersed colloid of (*RS*)-**10**, an emulsion was formed, and this emulsion gradually changed to a suspension, producing small precipitates. These precipitates of (*RS*)-**10**, called aggregates, were easily collected by centrifugation (an average grain size of 13.0 μm) and were reduced with a rate comparable to that of a DMF solution of (*RS*)-**10** under low concentration of substrate, as shown in Figure 4.

The aggregate was soluble in an aqueous medium at about 20 mg/L, and DMF did not need to be added to the reaction medium. By using the aggregate as the substrate, substrate concentration was increased up to an industrial production level (100 g/L, yield 80%) by means of the following procedure: A suspension of 11.2 g of baker's yeast and the aggregate, prepared from 5.0 g of the crystalline (*RS*)-**10**, in 50 mL of 200 mM ethanol solution was shaken at 30 °C for 9 days. Ethanol (0.46 g) was added every 12 h. The reduction of (*RS*)-**10** proceeded successively as shown in Figure 5. After processing as described earlier, over 99% e.e. of (2*S*,3*S*)-**11** was obtained with 80% yield.

The substrate concentration was improved by applying this procedure. The disadvantages of *(1)* a relatively low concentration of substrate, *(2)* a tedious procedure for isolation and purification of the product, and *(3)* a relatively low optical purity of the product in the microbial reduction were resolved. This method of using baker's yeast, which is commercially available and cheap, is expected to be developed for industrial application.

BIBLIOGRAPHY

1. K. Abe, H. Inoue, and T. Nagao, *Yakugaku Zasshi* **108**, 716–732 (1988).
2. H. Inoue, S. Takeo, M. Kawazu, and H. Kugita, *Yakugaku Zasshi* **93**, 729–732 (1973).
3. M. Senuma, M. Shibazaki, S. Nishimoto, K. Shibata, K. Okamura, and T. Date, *Chem. Pharm. Bull.* **37**, 3204–3208 (1989).
4. E.N. Jacobsen, L. Deng, Y. Furukawa, and L.E. Martínez, *Tetrahedron* **50**, 4323–4334 (1994).
5. K.G. Watson, Y.M. Fung, M. Gredley, G.J. Bird, W.R. Jackson, H. Gountzos, and B.R. Matthews, *J. Chem. Soc. Chem. Commun.* 1018–1019 (1990).
6. B.R. Matthews, H. Gountzos, W.R. Jackson, and K.G. Watson, *Tetrahedron Lett.* **30**, 5157–5158 (1989).
7. A. Schwartz, P.B. Madan, E. Mohacsi, J.P. O'Brien, L.J. Todaro, and D.L. Coffen, *J. Org. Chem.* **57**, 851–856 (1992).
8. O. Miyata, T. Shinada, I. Ninomiya, and T. Naito, *Tetrahedron Lett.* **32**, 3519–3522 (1991).

9. K. Matsuki, M. Sobukawa, A. Kawai, H. Inoue, and M. Takeda, *Chem. Pharm. Bull.* **41**, 643–648 (1993).
10. H. Akita, Y. Enoki, H. Yamada, and T. Oishi, *Chem. Pharm. Bull.* **37**, 2876–2878 (1989).
11. H. Akita, I. Umezawa, H. Matsukura, and T. Oishi, *Chem. Pharm. Bull.* **39**, 1632–1633 (1991).
12. H. Matsumae, M. Furui, and T. Shibatani, *J. Ferment. Bioeng.* **75**, 93–98 (1993).
13. H. Matsumae, M. Furui, T. Shibatani, and T. Tosa, *J. Ferment. Bioeng.* **78**, 59–63 (1994).
14. C.J. Shi and C.-S. Chen, *Angew. Chem. Int. Ed. Engl.* **23**, 570–578 (1984).
15. T. Kometani, H. Yoshii, E. Kitatsuji, H. Nishimura, and R. Matsuno, *J. Ferment. Bioeng.* **76**, 33–37 (1993).
16. T. Mukaiyama and M. Murakami, *Chem. Lett.* 1129–1132 (1981).
17. T. Nishida, H. Matsumae, I. Machida, and T. Shibatani, *Biocatal. and Biotransf.* **12**, 205–214 (1995).
18. S. Masamune, S.A. Ali, D.L. Snitman, and D.S. Garvey, *Angew. Chem. Int. Ed. Engl.* **19**, 557–558 (1980).
19. H. Matsumae, H. Douno, S. Yamada, T. Nishida, Y. Ozaki, T. Shibatani, and T. Tosa, *J. Ferment. Bioeng.* **79**, 28–32 (1995).
20. S. Yamada, Y. Mori, K. Morimatsu, Y. Ishizu, Y. Ozaki, R. Yoshioka, T. Nakatani, and H. Seko, *J. Org. Chem.* **61**, 8586–8590 (1996).
21. S. Servi, *Synthesis* 1–25 (1990).
22. R. Csuk and B.I. Glänzer, *Chem. Rev.* **91**, 49–97 (1991).
23. T. Nishida, I. Machida, H. Matsumae, S. Yamada, Y. Ozaki, and T. Shibatani, *J. Ferment. Bioeng.* **81**, 534–537 (1996).
24. T. Sugai, D. Sakuma, N. Kobayashi, and H. Ohta, *Tetrahedron* **47**, 7237–7244 (1991).
25. T. Kometani, Y. Sakai, H. Matsumae, T. Shibatani, and R. Matsuno, *J. Ferment. Bioeng.* **84**, 195–199 (1997).
26. T. Kometani, E. Kitatsuji, and R. Matsuno, *Chem. Lett.* 1465–1466 (1989).
27. T. Kometani, E. Kitatsuji, and R. Matsuno, *J. Ferment. Bioeng.* **71**, 197–199 (1991).
28. T. Kometani, H. Yoshii, Y. Takeuchi, and R. Matsuno, *J. Ferment. Bioeng.* **76**, 414–415 (1993).
29. T. Kometani, Y. Morita, H. Furui, H. Yoshii, and R. Matsuno, *J. Ferment. Bioeng.* **77**, 13–16 (1994).
30. T. Kometani, Y. Morita, Y. Kiyama, H. Yoshii, and R. Matsuno, *J. Ferment. Bioeng.* **80**, 208–210 (1995).

DIMENSIONAL ANALYSIS, SCALE-UP

MARKO ZLOKARNIK
Graz, Austria

KEY WORDS

Complete and partial similarity
Dimensional analysis
Dimensionless representation
Dimensions and physical quantities
Intermediate quantities
Non-Newtonian liquids
Relevance list
Scale-up
Theory of models
Viscolastic liquids

OUTLINE

INTRODUCTION

A chemical engineer is generally concerned with the industrial implementation of processes in which chemical or

microbiological conversion of material takes place in conjunction with the transfer of mass, heat, and momentum. These processes are scale dependent; that is, they behave differently on a small scale (in laboratories or pilot plants) and a large scale (in production). They include heterogeneous chemical reactions and most unit operations. Understandably, chemical engineers have always wanted to find ways of simulating these processes in models to gain insights that will assist them in designing new industrial plants. Occasionally, they are faced with the same problem for another reason: an industrial facility already exists but will not function properly, if at all, and suitable measures have to be carried out to discover the cause of the difficulties and provide a solution.

Irrespective of whether the model involved represents a "scale-up" or a "scale-down," certain important questions always apply:

1. How small can the model be? Is one model sufficient, or should tests be carried out in models of different sizes?
2. When must or when can physical properties differ? When must the measurements be carried out on the model with the original system of materials?
3. Which rules govern the adaptation of the process parameters in the model measurements to those of the full-scale plant?
4. Is it possible to achieve complete similarity between the processes in the model and those in its full-scale counterpart? If not, how should one proceed?

These questions touch on the fundamentals of the theory of models, which are based on dimensional analysis. Although they have been used in the field of fluid dynamics and heat transfer for more than a century (cars, aircraft, vessels, and heat exchangers were scaled up according to these principles), these methods have gained only modest acceptance in chemical engineering. University graduates are usually not skilled enough to deal with such problems at all. On the other hand, there is no motivation for this type of research at universities, which generally are not confronted with scale-up tasks and are not equipped with the necessary apparatus at the bench scale. This gives the totally wrong impression that the methods are, at most, of marginal importance in practical chemical engineering. It is most unfortunate that these methods are not taught and dealt with in greater depth.

DIMENSIONAL ANALYSIS

The Fundamental Principle

Dimensional analysis is based on the recognition that a mathematical formulation of a physicotechnological problem can be of general validity only when the process equation is dimensionally homogeneous.

What Is a Dimension?

A dimension is a purely qualitative description of a perception of a physical entity or a natural appearance. A length can be experienced as a height, a depth, a breadth. A mass presents itself as a light or heavy body, time as a short moment or a long period. The dimension of a length is L, the dimension of a mass is M, and so on.

What Is a Physical Quantity?

Unlike a dimension, a physical quantity represents a quantitative description of a physical quality (e.g., a mass of 5 kg). It consists of a measuring unit and a numerical value. The measuring unit of length can be a meter, a foot, a cubit, a yardstick, a nautical mile, a light-year, and so on. The measuring unit of temperature can be °C (centigrade), °R, °F, or K (kelvin). (It is therefore necessary to establish the measuring units in an appropriate measuring system.)

Basic and Derived Quantities and Dimensional Constants

We distinguish between basic and secondary quantities, the latter being often referred to as derived quantities. The basic quantities are based on standards and are quantified by comparison with them. The secondary units are derived from the primary ones according to physical laws (e.g., velocity = length/time). The borderline separating basic and derived quantities is largely arbitrary: 40 years ago a measuring system was used in which force was a primary dimension instead of mass!

All secondary units must be coherent with the basic units (Table 1); for example, the measuring unit of velocity must not be miles per hour or kilometers per hour, but meters per second. A secondary unit that has been established by a physical law may sometimes contradict another one. For example, according to Newton's second law of motion, the force F is expressed as a product of mass m and acceleration a: $F = ma$, having the measuring unit of kg m/s^2 ≡ N. According to Newton's law of gravitation, the force is defined by $F \propto m_1\, m_2/r^2$, thus leading to a completely different measuring unit (kg^2/m^2). To remedy this, the gravitational constant G—a dimensional constant—had to be introduced to ensure the dimensional homogeneity of this equation: $F = G\, m_1\, m_2/r^2$. Another example entails the universal gas constant R, the introduction of which ensures that in the perfect gas equation of state $pV = nRT$ the secondary unit for work $W = pV[M\, L^2/T^2]$ is not offended.

Another class of derived quantities is represented by the coefficients in diverse physical equations (e.g., transfer equations). They are established by the respective equations and determined via measurement of their constituents, (heat transfer coefficients etc.; see examples 7–10).

Table 1. Base Quantities: Their Dimensions and SI Units

Basis quantity	Basic dimension	Basic unit	
Length	L	m	meter
Mass	M	kg	kilogram
Time	T	s	second
Thermodynamic temperature	Θ	K	kelvin
Amount of substance	N	mol	mole
Electric current	I	A	ampere
Luminous intensity	I_v	cd	candela

Dimensional Systems

A dimensional system consists of all the primary and secondary dimensions and corresponding measuring units. The currently used International System of Dimensions (*Système International d'Unités*; SI) is based on seven basic dimensions. They are presented in Table 1, together with their corresponding basic units. For some of them a few explanatory remarks may be necessary.

Temperature expresses the thermal level of a system, not its energetic contents. (A fivefold mass of matter has a fivefold thermal energy at the same temperature!) The thermal energy of a system can indeed be converted into mechanical energy (base unit joule). Moles are amount of matter and must not be confused with the quantity of mass. The molecules react as individual entities regardless of their mass: one mole of hydrogen (2 g/mol) reacts with one mole of chlorine (71 g/mol) to produce 2 moles of hydrochloric acid, HCl.

Table 2 shows the most important secondary dimensions. Table 3 refers to some very frequently used secondary units that are named after famous researchers.

Dimensional Homogeneity of a Physical Content

The aim of dimensional analysis is to check whether the physical content of the case under examination can be formulated in a dimensionally homogeneous manner. The procedure necessary to accomplish this consists of two parts:

1. All physical parameters necessary to describe the problem are listed. This "relevance list" of the problem consists of the quantity in question and all the parameters that influence it. In each case only *one* target quantity must be considered; it is the only dependent variable. On the other hand, all the influencing parameters must be primarily independent of each other.
2. The dimensional homogeneity of the physical content is checked by transferring it in a dimensionless form. (*Note*: A physical content that can be transformed in dimensionless expressions is dimensionally homogeneous.)

The information given thus far can be made clear by a famous example (1).

Table 2. Often-Used Physical Quantities and Their Dimensions According to Currently Used SI Formulas in Mechanical and Thermal Problems

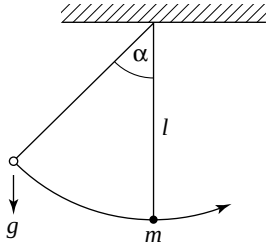
Physical quantity	Dimension
Angular velocity	$1/T$
Shear rate, frequency	
Mass transfer coefficient, $K_L a$	
Velocity	L/T
Acceleration	L/T^2
Kinematic viscosity	L^2/T
Diffusion coefficient	
Thermal diffusivity	
Density	M/L^3
Surface tension	M/T^2
Dynamic viscosity	$M/(LT)$
Momentum	$M\,L/T$
Force	$M\,L/T^2$
Pressure, stress	$M/(L\,T^2)$
Angular momentum	$M\,L^2/T$
Energy, work, torque	$M\,L^2/T^2$
Power	$M\,L^2/T^3$
Heat capacity	$L^2/T^2\Theta$
Thermal conductivity	$M\,L/T^3\Theta$
Heat transfer coefficient	$M/T^3\Theta$

Table 3. Important Secondary Measuring Units in Mechanics, Named After Famous Researchers

Secondary quantity	Dimension	Measuring unit	Eponymous researcher
Force	$M\,L/T^2$	kg m/s^2 ≡ N	Newton
Pressure	$M\,L/T^2$	kg (ms^2) ≡ Pa	Pascal
Energy	$M\,L^2/T^2$	kg m^2/s^2 ≡ J	Joule
Power	$M\,L^2/T^3$	kg m^2/s^3 ≡ W	Watt

EXAMPLE 1

What determines the period of oscillation of a pendulum?



Physical quantity	Symbol	Dimension
Period of oscillation	t	T
Length of pendulum	l	L
Mass of pendulum	m	M
Gravitational acceleration	g	L/T^2
Amplitude (angle)	α	—

We first sketch a pendulum and write down all the quantities that could be involved in this question. It may be assumed that the period of oscillation of a pendulum depends on the length and mass of the pendulum, the gravitational acceleration, and the amplitude of swing.

Our aim is to express t as a function of l, m, g, and α:

$$t = f(l, m, g, \alpha)$$

Can this dependency be dimensionally homogeneous? No! The first thing that now becomes clear is that the basic unit for mass occurs only in the mass, m, itself. Changing this basic unit (e.g., from kilograms to pounds) would change the numerical value of the function. This is unacceptable. Either our list should have included a further variable containing M, or mass is not a relevant variable. If we assume the latter, the relationship is reduced to:

$$t = f(l, g, \alpha)$$

Both l and g incorporate the basic unit of length. When combined

as a ratio (l/g), they become dimensionless with regard to L and thus independent of changes in the basic unit of length:

$$t = f(l/g, \alpha)$$

Because the angle α has no dimension, we are left with the dimension T on the left-hand side of the equation, and T^2 on the right. To remedy this, we will have to write $\sqrt{l/g}$. This expression will keep its dimension T only if it remains unchanged, so we have to take it out of the function f, and we obtain the following:

$$t = \sqrt{l/g}\, f(\alpha) \qquad \text{or} \qquad t\sqrt{g/l} = f(\alpha)$$

The dependency between four dimensional quantities, containing two basic dimensions (L and T) in their dimensions, is reduced to $4 - 2 = 2$ parametric relationship between dimensionless expressions ("numbers").

This equation is the only statement that dimensional analysis can offer in this case. It is not capable of producing information on the form of f. The integration of Newton's equation of motion for small amplitudes leads to $f = 2\pi$ and is independent of α. The relationship can now be expressed as

$$t\sqrt{g/l} = 2\pi$$

The elegant solution of this first example should not tempt the reader to believe that dimensional analysis can be used to solve every problem at hand. To treat this example by dimensional analysis, the law of free fall had to be known. This knowledge was gained empirically by G. Galilei in 1604 through his experiments with the inclined plane. Bridgman's (1, p. 12) comment on this situation is particularly appropriate:

> "The problem cannot be solved by the philosopher in his armchair, but the knowledge involved was gathered only by someone at some time soiling his hands with direct contact."

This transparent and easy example clearly shows how dimensional analysis deals with specific problems and what conclusions it allows. It should now be easier to understand the sarcastic comment with which Lord Rayleigh began his short essay on "The Principle of Similitude" (2):

> I have often been impressed by the scanty attention paid even by original workers in physics to the great principle of similitude. It happens not infrequently that results in the form of "laws" are put forward as novelties on the basis of elaborate experiments, which might have been predicted a priori after a few minutes' consideration.

From example 1 we also learn that a transformation of a physical dependency from a dimensional into a dimensionless form is automatically accompanied by an essential compression of the statement: the set of the dimensionless numbers is smaller than the set of the quantities contained in them, but it describes the problem equally comprehensively. In the example 1, the dependency between four dimensional parameters is reduced to a dependency between only two dimensionless numbers. This is the proof of the so-called pi theorem (pi after Π, the sign used for products), which states the following: Every physical relationship between n physical quantities can be reduced to a relationship between $m = n - r$ mutually independent dimensionless groups, whereby r stands for the rank of the dimensional matrix, made up of the physical quantities in question and generally equal to the number of the basic quantities contained in them. (*Note*: The pi theorem is often associated with E. Buckingham, who introduced this term in 1914, but the proof of it had been accomplished in 1911 in course of a mathematical analysis of partial differential equations by A. Federmann already [3, p. 7].)

THE DETERMINATION OF A PI SET BY MATRIX CALCULATION

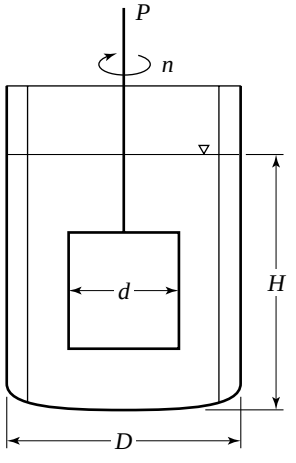
The Establishment of a Relevance List for a Problem

As a rule, more than two dimensionless numbers are necessary to describe a physicotechnological problem, and therefore they cannot be derived by the method just described. In this case the easy and transparent matrix calculation introduced by J. Pawlowski (4) is increasingly used. Example 2, which demonstrates this calculation, treats a problem that is important in industrial chemistry and biotechnology because the mixing vessel belongs to the most important production equipment.

EXAMPLE 2

Which is the pi set for the mixing power of a stirrer?

In this case, too, we first sketch the apparatus. Then we facilitate our work by proceeding systematically toward our target quantity, the mixing power P, which depends on the following parameters:



Geometrical	Stirrer diameter	d
Physical properties	Fluid density	ρ
	Kinematic viscosity	v
Process-related	Stirrer speed	n

We obtain the following relevance list:

$$\{P;\ d;\ \rho,\ v;\ n\} \tag{1}$$

We interrupt the procedure by asking an important question concerning the determination of the characteristic geometric parameter.

Determination of the Characteristic Geometric Parameter

Obviously, we could name all the geometric parameters indicated in the sketch. They were all the geometric parameters of the stirrer and of the vessel, especially its diameter D and the liquid height H. In case of complex geometry, such a procedure would deflect from the problem. It is therefore advisable to introduce only one (or a few) characteristic geometric parameters, knowing that all the others can be transformed into dimensionless geometric numbers by division with this one of a few of them.

In the case of example 2, we introduced the stirrer diameter as the characteristic geometric parameter. This is reasonable. One could only think about how the mixing power would react to an increase of the vessel diameter D: it is obvious that above a certain size, D would have no influence, but a small change of the stirrer diameter d would always have an impact.

We now set forth a procedure for solving example 2.

Constructing and Solving the Dimensional Matrix

In transforming the relevance list, equation 1, of the five physical quantities into a dimensional matrix, we can minimize the calculations required by recalling the following points.

1. The dimensional matrix consists of a square core matrix and a residual matrix.
2. The rows of the matrix are formed of basic dimensions, contained in the dimensions of the quantities, and they will determine the rank r of the matrix. The columns of the matrix are presented by the physical quantities or parameters.
3. Quantities of the square core matrix may appear later on in all the dimensionless numbers as "fillers," whereas each element of the residual matrix will appear only in one dimensionless number. For this reason the residual matrix should be loaded with essential variables such as the target quantity and the most important physical properties and process-related parameters.
4. By the (extremely easy) matrix calculation (linear transformations), the core matrix is transformed into a matrix of unity: the main diagonal consists only of ones, and the remaining elements are all zero. One should therefore arrange the quantities in the core matrix in a way that will facilitate this procedure.
5. After the matrix of unity has been generated, the dimensionless numbers are created as follows: each element of the residual matrix forms the numerator of a fraction, while the denominator consists of the fillers from the matrix of unity with the exponents indicated in the residual matrix.

Let us return to example 2. The dimensional matrix reads as follows:

	Core matrix			Residual matrix	
	ρ	d	n	P	v
Mass M	1	0	0	1	0
Length L	-3	1	0	2	2
Time T	0	0	-1	-3	-1

Only one linear transformation is necessary to transform -3 in the cell formed by the L row and the ρ-column into zero. The subsequent multiplication of the T row by -1 transfers -1 to 1:

	Unity matrix			Residual matrix	
	ρ	d	n	P	v
M	**1**	0	0	1	0
$3M + L$	0	**1**	0	5	2
$-T$	0	0	**1**	3	1

The residual matrix contains only two parameters, therefore only two Π numbers result:

$$\Pi_1 \equiv \frac{P}{\rho^1 n^3 d^5} = \frac{P}{\rho n^3 d^5} \equiv Ne \text{ (Newton number)}$$

$$\Pi_2 \equiv \frac{v}{\rho^0 n^1 d^2} = \frac{v}{nd^2} \equiv Re^{-1} \text{ (Reynolds number)}$$

The interdependence of five dimensional quantities of the relevance list, equation 1, reduces to a set of only $5 - 3 = 2$ dimensionless numbers:

$$\{Ne, Re\} \quad \text{or} \quad f(Ne, Re) = 0 \tag{2}$$

thus again confirming the pi theorem.

Determination of the Process Characteristics

The functional dependency, equation 2, is the maximum that dimensional analysis can offer here. It cannot provide any information about the form of the function f. This can be accomplished solely by experiments. But such experiments can now most certainly be conducted far more easily because it is the Reynolds number that must be varied, and this can be accomplished by altering the characteristic velocity v (here $v \propto nd$), or a characteristic length l (here d), or the kinematic viscosity ϑ. By choosing appropriate model fluids, the viscosity can be very easily altered by several orders of magnitude. This way measurements were conducted to determine the power characteristic $Ne(Re)$ of a blade stirrer of fixed geometry in a vessel. From Figure 1 it follows:

1. In the range $Re < 20$, the proportionality $Ne \propto Re^{-1}$ is found, thus resulting in the expression $NeRe \equiv P/(\eta n^2 d^3) = \text{const}$. Density is irrelevant here—we are dealing with the laminar flow region.
2. In the range $Re > 50$ (vessel with baffles) or $Re > 5 \times 10^4$ (unbaffled vessel), a constant Newton number $Ne \equiv P/(\rho\, n^3\, d^5)$ is found. In this case, viscosity is irrelevant—we are dealing with a turbulent flow region.
3. Understandably, the baffles do not influence the power characteristic in the laminar flow region. However, their influence is extremely strong at $Re > 5 \times 10^4$. Here, the installation of baffles under oth-

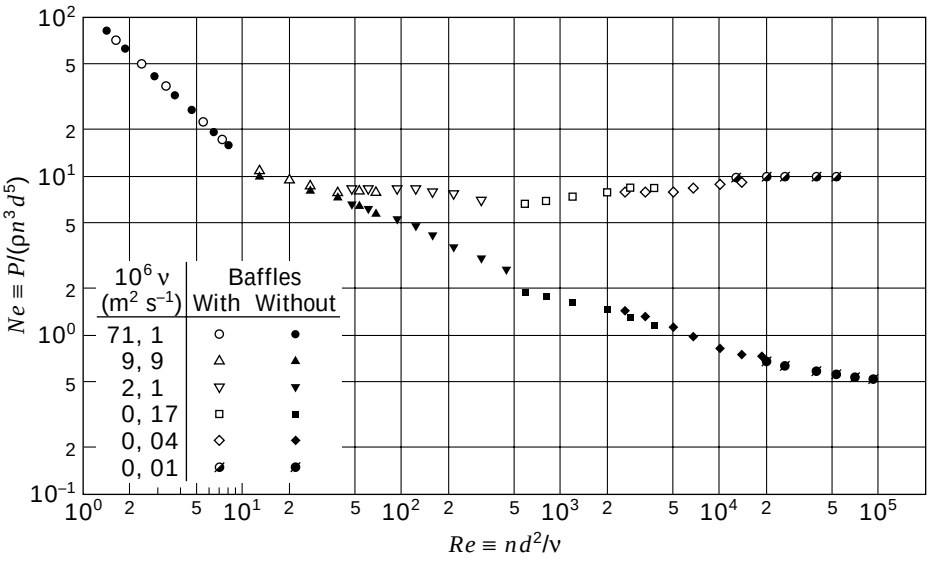


Figure 1. Power characteristics of a leaf stirrer of a given geometry and given installation conditions. *Source*: From Ref. 3.

erwise unchanged operating conditions increases the power consumption of the stirrer by a factor of 20.

4. When liquid is stirred in an unbaffled vessel, it begins to circulate, and a vortex is formed. Gravitational acceleration g, and hence the Froude number $Fr \equiv n^2\, d/g$, plays no role under these circumstances as determined experimentally (Fig. 1): For confirmation, one need only look at the points on the lower $Ne(Re)$ curve where the *same* Re value was set for fluids with *different* viscosities. This was possible only by a proportional alteration of the rotational speed of the stirrer. Where Re = idem (= identical value), Fr was clearly not idem, but this has no influence on Ne: g is therefore irrelevant.

FUNDAMENTALS OF THE THEORY OF MODELS AND SCALE-UP

Theory of Models

The results in Figure 1 also could have been acquired by changing the stirrer diameter: the means by which a relevant number is changed does not matter. This is clear evidence that the representation of a physicotechnological problem in a dimensionless form is independent of scale ("scale-invariant"), and this presents the basis for a reliable scale-up: two processes may be considered completely similar if they take place in similar geometrical space and if all the dimensionless numbers necessary to describe them have the same numerical value (Π_i = identical or idem).

Clearly, the scale-up of a desired process condition from a model to industrial scale can be accomplished reliably only if the problem was formulated and dealt with according to the dimensional analysis.

Model Experiments and Scale-Up

In example 2 we evaluated process characteristics that presented a comprehensive description of the process. This often expensive and time-consuming method is certainly not necessary if one has only to scale up a given process condition from the model to the industrial plant (or vice versa). With example 2, and assuming that the $Ne(Re)$ characteristics like those in Figure 1 are not known, the task is to predict the power consumption of a leaf stirrer of $d = 2$ m, installed in a baffled vesser of $D = 4$ m ($D/d = 2$) and rotating with $n = 10\ \text{min}^{-1}$. The kinematic viscosity of the liquid is $\nu = 1 \times 10^{-4}\ \text{m}^2/\text{s}$.

One need know only—and this is essential—that the hydrodynamics in this case are governed solely by the Reynolds number and that the process is described by an unknown dependency $Ne(Re)$. Then one can calculate Re of the industrial plant:

$$Re \equiv nd^2/\nu = 6.7 \times 10^3$$

Let us assume that we have a geometrically similar laboratory device of $D = 0.2$ m and $d = 0.1$ m and that the rotational speed of the stirrer can be arbitrarily chosen. Which must be the rotational speed to obtain Re = idem using water $\nu = 1 \times 10^{-6}\ \text{m}^2/\text{s}$ as model liquid? The answer is

$$nd^2/\nu = 6.7 \times 10^3 \rightarrow n = 40\ \text{min}^{-1}$$

Under these conditions the stirrer power is measured and the power number $Ne \equiv P/(\rho n^3 d^5)$ calculated. We find $Ne = 8.0$. Because Re = idem results in Ne = idem, the power consumption P_T of the industrial stirrer can be obtained:

$$Ne = \text{idem} \rightarrow Ne_T = Ne_M \rightarrow \left(\frac{P}{\rho n^3 d^5}\right)_T = \left(\frac{P}{\rho n^3 d^5}\right)_M$$

From $Ne = 8.0$ found in the laboratory measurement, it follows for the leaf stirrer having $d = 2$ m and rotational speed $n = 10\ \text{min}^{-1}$:

$$P = Ne\ \rho\ n^3 d^5 = 8.0 \times 1 \times 10^3 \times (10/60)^3 \times 2^5 = 1{,}185\ \text{W} \cong 1.2\ \text{kW}$$

We realize that in scale-up, knowledge of the functional dependency $f(\Pi_i) = 0$ is not necessary. All we need is to know which pi space describes the process.

FURTHER PROCEDURES TO ESTABLISH A RELEVANCE LIST

Consideration of the Acceleration Due to Gravity *g*

If a natural or universal physical constant has an impact on the process, it must to be incorporated into the relevant list, regardless of whether it will be altered. In this context the greatest mistakes are made with regard to the gravitational constant *g*. Lord Rayleigh (2) complained bitterly, saying:

> I refer to the manner in which gravity is treated. When the question under consideration depends essentially upon gravity, the symbol of gravity (g) makes no appearance, but when gravity does not enter the question at all, g obtrudes itself conspicuously.

This is all the more surprising in as much as the relevance of this quantity is easy enough to recognize if one asks the following question: would the process function differently if it took place on the moon instead of on Earth? If the answer to this question is yes, *g* is a relevant variable.

The gravitational acceleration *g* can be effective solely in connection with the density as gravity $g\rho$. When inertial forces play a role, the density ρ must be listed additionally. Thus it follows:

1. In cases involving the ballistic movement of bodies, the formation of vortices in stirring, the bow wave of a ship, the movement of a pendulum, and other processes affected by the Earth's gravity, the relevance list comprises $g\rho$ *and* ρ.
2. Creeping flow in a gravitational field is governed by the gravity $g\rho$ alone.
3. In heterogeneous physical systems with density differences (sedimentation or buoyancy), the difference in gravity $g\Delta\rho$ *and* ρ plays a decisive role.

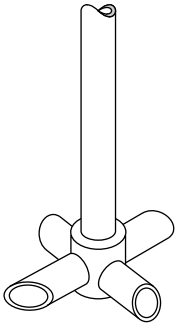
Example 3 treats a problem in which the gravitational constant is of prime importance owing to extreme difference in densities in the gas–liquid system. This example offers opportunities to improve our skills in solving dimensional matrices and to gain deeper insights into the possibilities made available by dimensional analysis.

EXAMPLE 3

Process characteristics of a self-aspirating hollow stirrer and the determination of its optimum running conditions

As a result of their form, hollow stirrers utilize the suction generated behind their edges (Bernoulli effect) to suck in gas from the head space above the liquid. As "rotating ejectors," they are stirrers and gas pumps in one and are therefore particularly suitable for laboratory use (especially in high-pressure autoclaves) because they achieve intensive gas liquid-contact via internal gas recycling without a separate gas pump (3,5). A particularly effective stirrer of this type—the pipe stirrer—is depicted here.

In operating a hollow stirrer, the target quantities gas throughput *q* and stirrer power *P* adapt themselves simultaneously. Both target quantities depend on the following parameters:



Geometry	Stirrer diameter	d
Physical properties	Density	ρ
	Dynamic viscosity	μ
Process parameters	Rotational speed	n
	Gravitational constant	g

Thus we obtain two separate relevance lists:

$$\{q;\ d,\ \rho,\ \mu;\ n,\ g\} \tag{3}$$

$$\{P;\ d,\ \rho,\ \mu;\ n,\ g\} \tag{4}$$

The problem is of a mechanical nature, and dimensions of the quantities describing it will contain only three basic quantities (*M, L, T*). It is to be expected that each process characteristics will therefore involve $6 - 3 = 3$ numbers.

The first relevance list, equation 3, leads to the following dimensional matrix:

	ρ	d	n	q	μ	g
Mass M	1	0	0	0	1	0
Length L	−3	1	0	3	−1	1
Time T	0	0	−1	−1	−1	−2
M	1	0	0	0	1	0
$3M + L$	0	1	0	3	2	1
$-T$	0	0	1	1	1	2

From this the following three numbers result:

$$\Pi_1 \equiv \frac{q}{\rho^0 d^3 n^1} = \frac{q}{nd^3} \equiv Q$$

$$\Pi_2 \equiv \frac{\mu}{\rho^1 d^2 n^1}\frac{\mu}{\rho d^2 n} \equiv \text{Re}^{-1}$$

$$\Pi_3 \equiv \frac{\text{g}}{\rho^0 d^1 n^2} = \frac{\text{g}}{dn^2} \equiv \text{Fr}^{-1}$$

where *Q* is the gas throughput number, *Re* the Reynolds number, and *Fr* the Froude number.

The gas throughput characteristic of a hollow stirrer thus reads as follows:

$$\frac{q}{nd^3} = f_1\left(\frac{nd^2\rho}{\mu}, \frac{n^2 d}{g}\right) \rightarrow Q = f_1(\text{Re, Fr}) \tag{5}$$

The power characteristics—obtained from the relevance list equa-

tion 4 by a similar dimensional matrix—containing P instead of q—results in the following:

$$\frac{P}{\rho n^3 d^5} = f_2\left(\frac{nd^2\rho}{\mu}, \frac{n^2 d}{g}\right) \rightarrow Ne = f_2(Re, Fr) \qquad (6)$$

Model experiments with another type of hollow stirrer (three-edged stirrer, see sketch in Fig. 2) were performed in a water–air system under defined experimental conditions, and the scale factor d_T/d_M changed in the range of 1:2:3:4:5. The results (Figs. 2 and 3) demonstrate that the Reynolds number is irrelevant. In the turbulent range of motion ($Re > 10^4$) the process is governed exclusively by the Froude number.

These process characteristics present the reliable basis for a scale-up of this hollow stirrer under given geometric conditions. But they also allow a further optimization of this process, as will be demonstrated by the following treatment. The question must be answered regarding process conditions under which this stirrer will achieve a given gas throughput with minimum power, P/q = min. The answer can easily be given by the combination of the characteristics presented in equations 5 and 6 in such a way as to produce a dimensionless expression for P/q. This is the combination

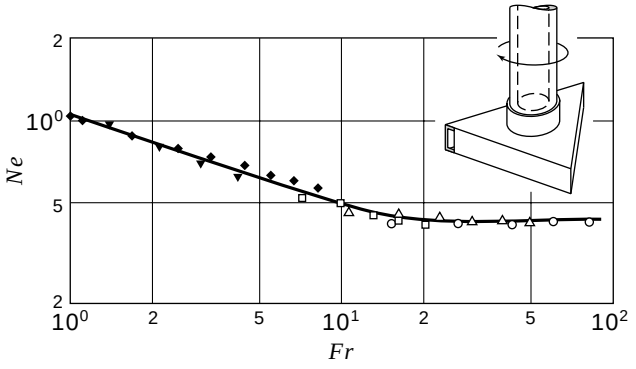


Figure 2. Power characteristics Ne (Fr) of the three-edged stirrer depicted in the inset. *Source*: From Ref. 3.

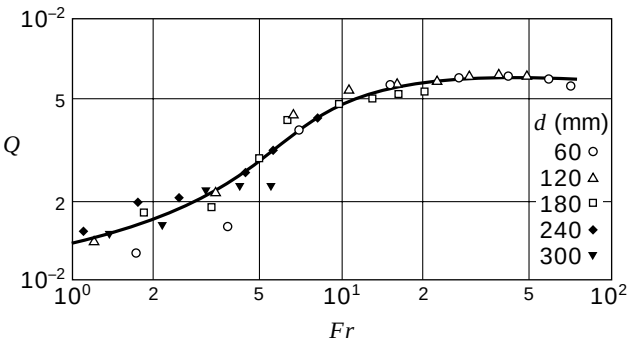


Figure 3. Gas throughput characteristics Q(Fr) of a three-edged stirrer (see the sketch in Fig. 2). *Source*: From Ref. 3.

$$\frac{NeFr}{Q} \equiv \frac{P}{qd\rho g} = f_3(\mathrm{Fr}) \qquad (7)$$

which is, as are their constituents, also a function of the Froude number. The dependency, equation 7, is plotted in Figure 4. Two sections can be distinguished:

$$Fr \leq 10: \; NeFr/Q = \text{const} \rightarrow P/q \propto d$$

$$Fr \geq 10: \; NeFr/Q \sim Fr \rightarrow P/q \propto d^2$$

In the first section P/q increases with direct proportionality to the stirrer diameter. This means that in an industrial plant having a scale-up factor, say, d_T/d_M of 10 (T, industrial scale; M, model scale), the same gas throughput requires a power consumption ten-fold higher than the model scale. At $Fr \geq 10$ the hollow stirrer would be still less efficient.

This procedure clearly shows that hollow stirrers, despite their excellent performance on the laboratory scale, do not perform economically on an industrial scale. In this case the gas throughput and the stirrer power have to be separated by choosing an appropriate stirrer (a Rushton turbine is a good choice) and providing the gas from below the stirrer by a blower.

Introduction of Intermediate Quantities

Many engineering problems involve several parameters, a condition that impedes the elaboration of the pi space. Fortunately, in some cases a closer look at a problem (or previous experience) facilitates reduction of the number of physical quantities in the relevance list. This is the case when some relevant variables affect the process by way of a so-called intermediate quantity. Assuming that this intermediate variable can be measured experimentally, it should be included in the problem relevance list if this would facilitate the removal of more than one variable from the list.

The fluid velocity v in pipes—or the superficial gas velocity v_G in mixing vessels or in bubble columns—present

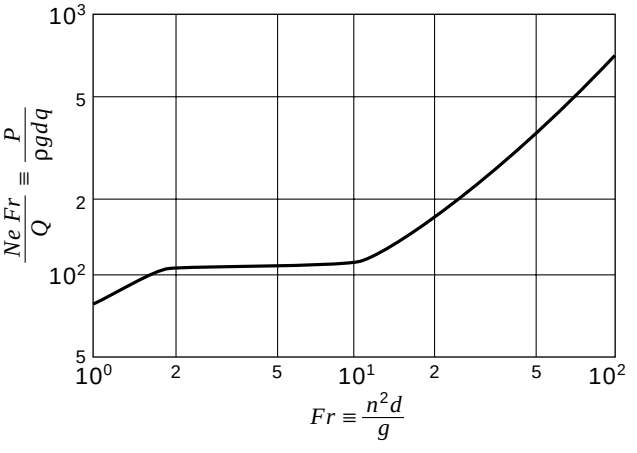


Figure 4. Optimization of a hollow stirrer according to P/q = min. The result reads: Small is beautiful.

a well-known intermediate quantity. The introduction of any such velocity into the relevance list removes two other quantities (throughput q and the diameter D) because of $v \propto q/D^2$ and $v_G \propto q_G/D^2$, respectively. The potential impact of the introduction of intermediate quantities on the relevance list is demonstrated in the following elegant examples.

EXAMPLE 4

Mixing characteristics for liquid mixtures with differences in density and viscosity

Mixing time θ necessary to achieve a molecular homogeneity of a liquid mixture (normally measured by decolorization methods) depends in material systems without differences in density and viscosity on only four parameters: stirrer diameter d, density ρ, dynamic viscosity η, rotational speed n:

$$\{\theta;\ d;\ \rho,\ v;\ n\} \tag{8}$$

From this, the mixing time characteristics are found to be

$$n\theta = f(Re) \qquad Re \equiv nd^2/v \tag{9}$$

In material systems with differences in density and viscosity, the relevance list, equation 8, enlarges by the physical properties of the second mixing component, by the volume ratio of both phases $\varphi = V_2/V_1$ and, because of the density differences, inevitably by the gravity difference $g\Delta\rho$ to nine parameters:

$$\{\theta;\ d;\ \rho_1,\ v_1,\ \rho_2,\ v_2,\ \varphi;\ g\Delta\rho,\ n\} \tag{10}$$

This results in mixing time characteristics incorporating six numbers:

$$n\theta = f(Re,\ Ar,\ \rho_2/\rho_1,\ v_2v_1,\ \varphi) \tag{11}$$

where $Re \equiv nd^2/v_1$ is the Reynolds number and $Ar \equiv g\Delta\rho d^3/(\rho_1 v_1^2)$ is the Archimedes number.

Meticulous observation of this mixing process (the slow disappearance of the Schlieren patterns as result of the disappearance of density differences), reveals that macromixing is quickly accomplished compared to micromixing. This time-consuming process already takes place in a material system that can be fully described by the physical properties of the mixture:

$$v^* = f(v_1,\ v_2,\ \varphi) \qquad \text{and} \qquad \rho^* = f(\rho_1,\ \rho_2,\ \varphi)$$

By introducing these intermediate quantities v^* and ρ^*, we can reduce the relevance list, equation 10, by three parameters to

$$\{\theta;\ d;\ \rho^*,\ v^*;\ g\Delta\rho,\ n\} \tag{12}$$

and get a mixing characteristic of only three numbers:

$$n\theta = f(\mathrm{Re},\ \mathrm{Ar}) \tag{13}$$

(In this case, Re and Ar have to be formed by ρ^* and v^*)

The process characteristics of a cross-beam stirrer was established in this pi space by evaluation of corresponding measurements in two differently sized mixing vessels ($D = 0.3$ and 0.6 m) using different liquid mixtures ($\Delta\rho/\rho^* = 0.01 - 0.29$ and $v_2/v_1 = 1 - 5{,}300$). It reads (6) as follows:

$$\sqrt{n\theta} = 51.6\ Re^{-1}(Ar^{1/3} + 3) \qquad Re = 10^1\text{–}10^5;\ Ar = 10^2\text{–}10^{11}$$

Example 4, which clearly shows the great advantages achieved by the introduction of intermediate quantities, is amplified by example 5.

EXAMPLE 5

Description of the dissolved air flotation process

In the flotation process, fine gas bubbles adhere to the hydrophobic surface of solid particles and float them to the surface of the liquid, from which they are removed as foam. Dissolved air flotation has been introduced to the treatment of wastewater for the removal and subsequent concentration of biological sludge. It utilizes air that has been dissolved under a higher pressure and is released by decompression to atmospheric pressure, thus producing the finest possible gas bubbles. The appliances in use are either troughs exhibiting longitudinal flow, or flotation cells. The latter are advantageously divided by a concentric wall into two spaces (3, p. 87; 7). In the inner chamber, gas bubbles are released and brought into intimate contact with solids in suspension. Here the principal part of the flotation process occurs. The annular space surrounding the inner chamber exhibits vertical flow from top to bottom. It serves for calming down the flow, thus facilitating the final clarification (removal of the residual particles).

Flotation, a depletion process of first-order kinetics, is described by the flotation rate constant k_F [T^{-1}] as the target quantity. A continuous process without back-mixing (plug flow) of the liquid is described in Ref. 7:

$$-\ln \frac{\varphi_t - \varphi_\infty}{\varphi_0 - \varphi_\infty} = k_F\tau$$

where φ is the solids content and τ is the mean residence time of the liquid throughput in the inner chamber or in the annulus, respectively.

In case of full back-mixing in the flotation unit, solids discharge A (which corresponds to the conversion factor X in first-order chemical reactions) is utilized (3, p. 87):

$$A = 1 - \frac{\varphi_{out}}{\varphi_{in}} = 1 - \frac{1}{1 + k_F\tau}$$

In constructing the relevance list of this unit operation, it is easy to list the process parameters. They are liquid feed q_{in}, which leaves the flotation cell and splits into floated output q_{fl} and "clear" output: $q_{out} = q_{in} - q_{fl}$; releasable gas content of the liquid feed q_G/q_{in}; gravitational acceleration g: $\{q_{in},\ q_{out},\ q_G/q_{in};\ g\}$.

The problem lies is the listing of all physical properties of the system, since these properties fluctuate strongly during the process (very much in contrast to the mineral flotation) and are only partially known. There is a degree of hydrophobicity (wettability) of the particle surface, described by the contact angle Θ. Another important parameter is the electrical charge of the particles, which greatly depends on the pH. Furthermore, finer particles float to the surface only after they have been agglomerated to flocs with the aid of a polyelectrolyte. Therefore the surfactant (flocculant) concentration c_f and its chemical origin (type) must be incorporated in the relevance list, too. Further material parameters will be the gravity difference $g\Delta\rho$, the solids content φ in the liquid, and the average gas bubble size (i.e., all liquid properties S_i de-

scribing bubble coalescence, which is still unknown by number and kind): {Θ, pH, c_f, type, $g\Delta\rho$, φ, S_i}

The tentative relevance list will therefore read as follows:

$$\{A;\ D;\ \Theta,\ \text{pH},\ c_f,\ \text{type},\ \varphi,\ S_i;\ q_{in},\ q_{out},\ q_G/q_{in},\ g\Delta\rho\} \qquad (14)$$

This relevance list can be streamlined significantly by introducing two intermediate quantities: the superficial velocity $v \propto q_{in}/D^2$ in the flotation cell and the rising velocity w of the flocs.

The rising velocity w of the flocs as an intermediate variable should allow *all* relevant physical properties to be taken into account; however, separate measurements of this intermediate quantity in suitable equipment are necessary.

The complete relevance list is then

$$\{A;\ w,\ v,\ q_{in},\ q_{out}\} \qquad (15)$$

and from it we have the following pi set of three parameters:

$$\{A;\ w/v,\ q_{out}/q_{in}\} \qquad (16)$$

In case of $q_{out}/_{in}$ = const (this parameter governs the flotate concentration), the solids discharge A will depend only on w/v:

$$A = f(w/v) \qquad (17)$$

To prove this, measurements were performed in a bench-scale flotation cell (D = 1.6 m); the sample consisted of biologically clarified wastewater from a large chemical factory (160 production plants). It turned out that at q_{out}/q_{in} = 0.9, maintaining $w/v \approx 2$ was sufficient to ascertain a complete removal of the biosludge; see Example A 6 in Reference 3.

In the preceding example one need only to compare the tentative relevance list (equation 14) with the final five-parameter relevance list (equation 15) to fully realize the advantages of introducing intermediate quantities in complex interrelations.

Material Systems of Unknown Physical Properties

With foams, sludges, and slimes often encountered in biotechnology, we are confronted with the problem of not being able to list the physical properties because they are still unknown and therefore cannot be quantified. This situation often leads to the opinion that dimensional analysis would fail in such cases.

It is obvious that this conclusion is not valid: dimensional analysis is a *method* based on logical and mathematical fundamentals (1,3,4). If relevant parameters cannot be listed because they are unknown, one cannot blame the method. The only solution is to perform the model measurements with the same material system and to change the model scales.

EXAMPLE 6

Scale-up of a mechanical foam breaker

To obtain reliable information on dimensioning and then scale-up of a given type of mechanical foam breaker (foam centrifuge, see Fig. 5), we wish to ascertain the mode of performing and to evaluate the model measurements of the device. Preliminary experiments have shown that for each foam emergence—proportional to the gas throughput q_G—for each foam breaker of diameter d, a minimum rotational speed n_{min} exists that is necessary to control it. The dynamic properties of the foam (density and viscosity, elasticity of the foam lamella, etc.) cannot be named or measured. We will have to content ourselves by listing them wholesale as material properties S_i. In our model experiments we will of course be able to replace S_i by the known type of surfactant (foamer) and its concentration c_f (ppm).

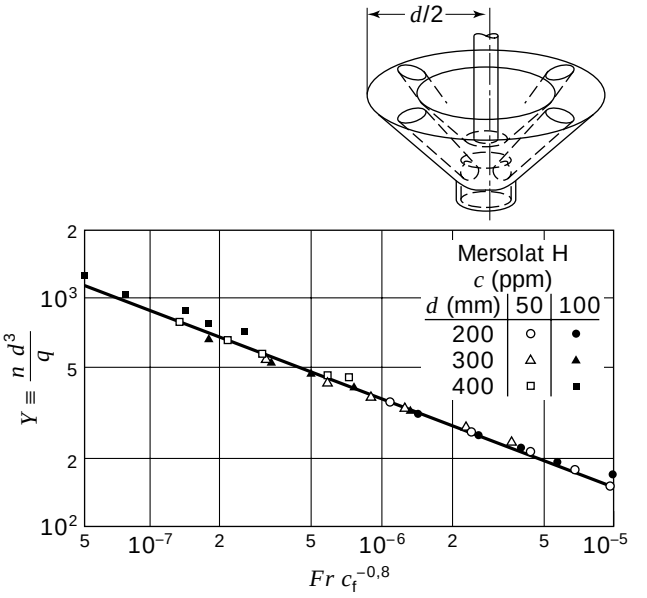


Figure 5. Process characteristics of the foam centrifuge (inset) for a particular foamer (Mersolat H, Bayer AG, Germany). *Source*: From Ref. 8.

In discerning the process parameters, we realize that the gravitational acceleration g has no impact on the foam breaking *within* the foam centrifuge: the centrifugal acceleration n^2d exceeds the gravitational one (g) by far. However, the water content of the foam entering the centrifuge depends very much on the gravitational acceleration: on the moon, the water drainage would be far less effective. In contrast to the dimensional analysis presented elsewhere (8), we are well advised to add g to the relevance list:

$$\{n_{min};\ d;\ \text{type of foamer},\ c_f;\ q_G,\ g\} \qquad (18)$$

For the sake of simplicity, in the following, n_{min} is replaced by n, and q_G by q. For each type of foamer we obtain the following pi space:

$$\left\{\frac{nd^3}{q},\ \frac{n^2d}{g},\ c_f\right\} \quad \text{or, abbreviated} \quad \{Y,\ Fr,\ c_f\} \qquad (19)$$

To prove this pi space, measurements must be made in model equipment of different sizes, to produce a reliable process characteristics. Figure 5 gives results for a commercial foamer (Mersolat H, Bayer AG, Germany); the pi space of equation 19 is seen to be fully satisfied.

The straight line in Figure 5 corresponds to the analytic expression

$$Y = Fr^{-0.4}c_f^{0.32} \qquad (20)$$

which can be reduced to

$$nd = \text{const } q^{0.2} f(c_f) \tag{21}$$

Here, the foam breaker will be scaled up according to its tip speed $u = \pi nd$ in model experiments, which will also moderately depend on the foam yield (q).

In all the other foamers examined (9), the correspondence $Y \propto Fr^{-0.45}$ was found. If—by chance—the correlation

$$Y \propto \text{Fr}^{-0.5} f(c_f)$$

proved to be true, then this one can be deduced to

$$n^2 d/g = \text{const}(c_f)$$

In this case the centrifugal acceleration ($n^2 d$) would present the scale-up criterion and would depend only on the foamer concentration, not on foam yield (q).

Short Summary of the Essentials of the Dimensional Analysis and Scale-Up

The advantages made possible by the correct and timely use of dimensional analysis can be stated briefly as follows.

1. *Reduction of the number of parameters required to define the problem*. The Π theorem states that a physical problem can always be described in dimensionless terms. Thus the number of dimensionless groups that fully describe it is much smaller than the number of dimensional physical quantities and is generally equal to the number of physical quantities minus the number of basic units contained within them.
2. *Reliable scale-up of the desired operating conditions from the model to the full-scale plant*. According to the theory of models, two processes may be considered to be similar if they take place under geometrically similar conditions and all dimensionless numbers that describe the processes have the same numerical value.
3. *A deeper insight into the physical nature of the process*. When experimental data are presented in a dimensionless form, distinct physical states can be isolated from one another (e.g., turbulent or laminar flow regions), and the effect of individual physical variables can be identified.
4. *Flexibility in the choice of parameters and their reliable extrapolation within the range covered by the dimensionless numbers*. These advantages become clear if one considers the well-known Reynolds number, $Re = vl/\nu$, which can be varied by altering the characteristic velocity v, or a characteristic length l, or the kinematic viscosity ν. By choosing appropriate model fluids, the viscosity can be very easily altered by several orders of magnitude. Once the effect of the Reynolds number is known, extrapolation of both v and l is allowed within the examined range of Re.

Area of Applicability of Dimensional Analysis

The application of dimensional analysis is indeed heavily dependent on the available knowledge. The following five steps (see Fig. 6) can be outlined.

1. The physics of the basic phenomenon is unknown: dimensional analysis cannot be applied.
2. Enough is known about the physics of the basic phenomenon to compile a first, tentative relevance list: the resultant pi set is unreliable.
3. All the relevant physical variables describing the problem are known: the application of dimensional analysis is unproblematic.
4. The problem can be expressed in terms of a mathematical equation: a closer insight into the pi relationship is feasible and may facilitate a reduction of the set of dimensionless numbers.
5. A mathematical solution of the problem exists: the application of dimensional analysis is superfluous.

It must, of course, be said that approaching a problem from the point of view of dimensional analysis also remains useful even if all the variables relevant to the problem are not yet known: the timely application of dimensional analysis often lead to the discovery of forgotten variables or the exclusion of artifacts.

Experimental Methods for Scale-Up

A number of questions posed at the beginning of this article are often asked in connection with model experiments.

How small can a model be? The size of a model depends on the scale factor L_T/L_M and on the experimental precision of measurement. Where $L_T/L_M = 10$, a $\pm$ 10 % margin of error may already be excessive. A larger scale

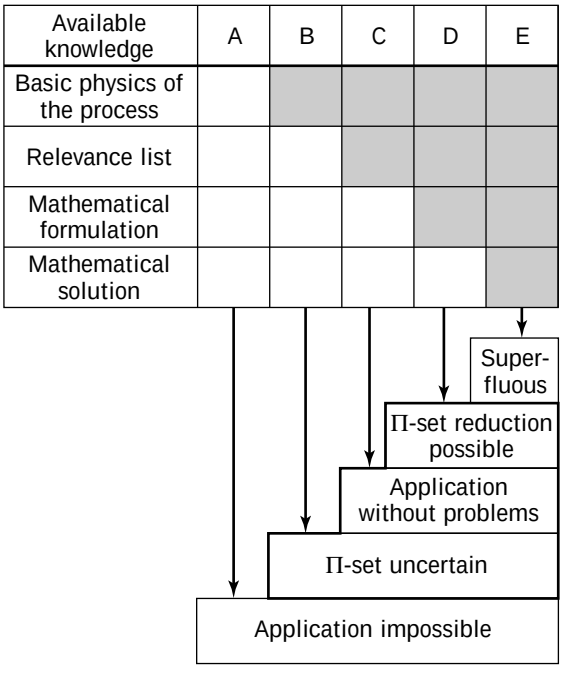


Figure 6. Graphical representation of the four levels of knowledge and their impact on the treatment of the problem by dimensional analysis. *Source*: After J. Pawlowski, Ref. 3.

for the model will therefore have to be chosen to reduce the error.

Is one model scale sufficient, or should tests be carried out in models of different sizes? One model scale is sufficient if the relevant numerical values of the dimensionless numbers necessary to describe the problem (the "process point" in the pi space describing the operational condition of the technical plant) can be adjusted by choosing the appropriate process parameters or physical properties of the model material system. If this is not possible, the process characteristics must be determined in a series of models of different sizes, or the process point must be extrapolated from experiments in technical plants of different sizes.

When must model experiments be carried out exclusively with the original material system? Where the material model system is unavailable (e.g., in the case of non-Newtonian fluids) or where the relevant physical properties are unknown (e.g., foams, sludges, slimes), model experiments must be carried out with the original material system. In such cases, measurements must be performed in models of various sizes (cf. example 6).

Partial Similarity

In the scale-up from a model to industrial scale, not only must the geometric similarity be ensured, but also all dimensionless numbers describing the problem must retain the same numerical values (Π_i = idem). This means, for example, that in the scale-up of boats or ships the dimensionless numbers governing the hydrodynamics

$$Fr \equiv \frac{v^2}{lg} \qquad \text{and} \qquad Re \equiv \frac{vl}{v}$$

must retain their numerical values: $Fr_T = Fr_M$ and $Re_T = Re_M$. It can easily be shown that this requirement cannot be fulfilled here.

Since the gravitational acceleration g cannot be varied on Earth, the Froude number, Fr, of the model can be adjusted to that of the full-scale vessel only by its velocity v_M. Subsequently, Re = idem can be achieved only by adjusting the viscosity of the model fluid. If the model size is only 10% of the full size (scale factor $L_T/L_M = 10$), Fr = idem is achieved in the model at $v_M = 0.32\ v_T$. To fulfill Re = idem for the kinematic viscosity of the model fluid ν_M, it follows

$$\frac{\nu_M}{\nu_T} = \frac{v_M}{v_T}\frac{L_M}{L_T} = 0.32 \times 0.1 = 0.032$$

No liquid exists whose viscosity would be only 3% of that of water!

Some requirements concerning physical properties of model materials cannot be implemented. In such cases only a partial similarity can be realized. For this, essentially only two procedures are available (for details, see Refs. 3 and 10). One consists of a well-planned experimental strategy in which the process is divided into parts, which are then investigated separately under conditions of complete similarity. This approach was first applied by William Froude (1810–1879) in his efforts to scale up the drag resistance of a ship's hull.

The second approach consists in deliberately abandoning certain similarity criteria and checking the effect on the entire process. This technique was used by Gerhard Damköhler (1908–1944) in his attempts to treat a chemical reaction in a catalytic fixed-bed reactor by means of dimensional analysis. Here the problem of a simultaneous mass and heat transfer arises—they are two processes that obey completely different fundamental principles.

It is seldom realized that many rules of thumb utilized for scale-up of different types of equipment are represented by quantities fulfill only a partial similarity. For example, only the volume-related mixing power P/V (widely used for scaling up mixing vessels) and the superficial velocity v (normally used for scale-up of bubble columns) should be mentioned here.

The volume-related mixing power P/V presents an adequate scale-up criterion only in L/L dispersion processes and can be deduced from the pertinent process characteristics $\delta/d \propto We^{-0.6}$ (δ, droplet diameter; We, Weber number). In the most common mixing operation, the homogenization of miscible liquids, where a macro- and back-mixing are required, this criterion fails completely (10).

Similarly, the superficial velocity v_G of the gas throughput as an intensity quantity is a reliable scale-up criterion only in mass transfer in gas–liquid systems in bubble columns. In mixing operations in bubble columns, requiring the whole liquid content be back-mixed (e.g., in homogenization), this criterion completely loses its validity (10).

Thus we must conclude that a particular scale-up criterion that is valid in a given type of apparatus for a particular process is not necessarily applicable to other processes occurring in the same device.

TREATMENT OF VARIABLE PHYSICAL PROPERTIES BY DIMENSIONAL ANALYSIS

Biotechnology deals with nonhomogeneous material systems exhibiting mostly non-Newtonian viscosity behaviors that often change their properties under shear stress. Thus a more comprehensive treatment of variable physical properties by dimensional analysis is essential.

It is generally assumed that the physical properties of the material system remain unaltered in the course of the process. Process equations (e.g., the heat characteristics of a mixing vessel as in Example 9 below, or a smooth straight pipe)

$$Nu = f(Re, Pr) \tag{22}$$

are valid for any material system with Newtonian viscosity and for any constant process temperature (i.e., for any one-time *constant* physical properties).

However, constancy of physical properties cannot be assumed in every physical process. A temperature field may well generate a viscosity field or even a density field in the material system treated. In non-Newtonian (structurally viscous or viscoelastic) liquids, a shear rate can also produce a viscosity field.

Although most physical properties (e.g., viscosity, density, heat conductivity and capacity, surface tension) must be regarded as variable, the value of viscosity in particular can be varied by many orders of magnitude under certain process conditions. In the following, we apply dimensional analysis solely to describe the variability of this one physical property. However, the approach can be adapted for any other physical property.

Dimensionless Representation of the Material Function

Similar behavior of a certain physical property common to different material systems can only be visualized by dimensionless representation of the material function of that property. It is furthermore desirable to formulate this function as uniformly as possible. This can be achieved by the "standard representation" of the material function (4,11) in which a standardized transformation of the material function $\eta(\vartheta)$ is defined such that the expression produced

$$\mu/\mu_0 = \varphi\{-\gamma_0(\vartheta - \vartheta_0)\} \tag{23}$$

meets the requirement

$$\varphi(0) = \varphi'(0) = 1$$

where $\gamma_0 \equiv [(1/\mu)(\delta\mu/\delta\vartheta)]_{\vartheta_0}$ is the temperature coefficient of the viscosity, $\mu_0 \equiv \mu(\vartheta_0)$, and ϑ_0 is *any* reference temperature. Figure 7 depicts the dramatic effect of such standard transformations.

Engineers prefer the representation

$$\mu/\mu_0 = \exp\{-\gamma_0(\vartheta - \vartheta_0)\} \tag{24}$$

which is a special case of equation 23, although this is not the best possible approximation (see dashed line in Fig. 7b). A slightly better approximation of the material function $\mu(\vartheta)$ is provided by the well-known Arrhenius relationship:

$$\frac{\mu}{\mu_0} = \exp\left\{-\frac{E_0}{R}\left(\frac{\vartheta_0}{\vartheta} - 1\right)\right\} \tag{25}$$

To describe the process by dimensional analysis, it is advisable to formulate the reference temperature ϑ_0 in a process-related manner, using a characteristic, possibly mean process temperature, as a reference. A class of functions exists, however, that are independent (invariant) of the reference point ϑ_0. The solid line in Figure 7b shows a function of this type. (*Note*: Function equations 23 and 24 are special cases of invariant functions.)

Pi Set for Variable Physical Properties

The type of dimensionless representation of the material function affects the (extended) pi set within which the process relationship is formulated (for more information see Refs. 3 and 11). When the standard representation is used, the relevance list must include the reference viscosity η_0 instead of η and incorporate two additional parameters γ_0, ϑ_0. This leads to *two* additional dimensionless numbers in the process characteristic. With regard to the heat transfer characteristic, equation 22 of a mixing vessel, it now follows that

$$Nu = f(Re_0, Pr_0, \gamma_0\Delta\vartheta, \Delta\vartheta/\vartheta_0) \tag{26}$$

where the subscript to *Re* and *Pr* denotes that these two dimensionless numbers are to be formed with μ_0 (which is the numerical value of μ at ϑ_0).

If we consider that the standard transformation of the material function can be expressed invariantly with regard to the reference temperature ϑ_0 (Fig. 7b), the relevance list is extended by only *one* additional parameter, γ_0. This, in turn, leads to only *one* additional dimensionless number. For our problem, it now follows that

$$Nu = f(Re_0, Pr_0, \gamma_0\Delta\vartheta) \tag{27}$$

Non-Newtonian Liquids

The main characteristic of Newtonian liquids is that simple shear flow (e.g., Couette flow) generates shear stress τ, which is proportional to the shear rate $\dot{\gamma} \equiv dv/dy$, (s^{-1}). The proportionality constant, the dynamic viscosity μ, is the only material constant in the Newton's law of motion:

$$\tau = \mu\dot{\gamma} \tag{28}$$

where μ depends only on pressure and temperature.

In the case of non-Newtonian liquids, μ depends on $\dot{\gamma}$ as well. These liquids can be classified in various categories of materials depending on their flow behavior: $\eta(\dot{\gamma})$, flow curve; $\mu(\tau)$, viscosity curve.

J. Pawlowski (12) points out that the rheological properties of many non-Newtonian liquids can be described by material parameters whose dimensional matrix has a rank of 2. These physical properties can be usefully linked to produce two dimensional constants: a characteristic viscosity constant H and a characteristic time constant θ, and possibly a set of dimensionless material numbers Π_{rheol}. The relevance list is thus formed by

$$\{H, \theta, \Pi_{\text{rheol}}\} \tag{29}$$

In a change from Newtonian to non-Newtonian liquids, this interrelation has the following consequences for the complete pi set:

1. Every dimensionless number incorporating η must now be formulated with the dimensional equivalent H (e.g., μ_0, η_∞).
2. A single process-related number containing θ is added (e.g., $\dot{\gamma}$).
3. The pure material numbers are increased by Π_{rheol}.

We will demonstrate these rules on the heat transfer characteristic of a smooth, straight pipe. In the following table, line *A* stands for temperature-independent viscosity, and line *B* for a temperature-dependent viscosity.

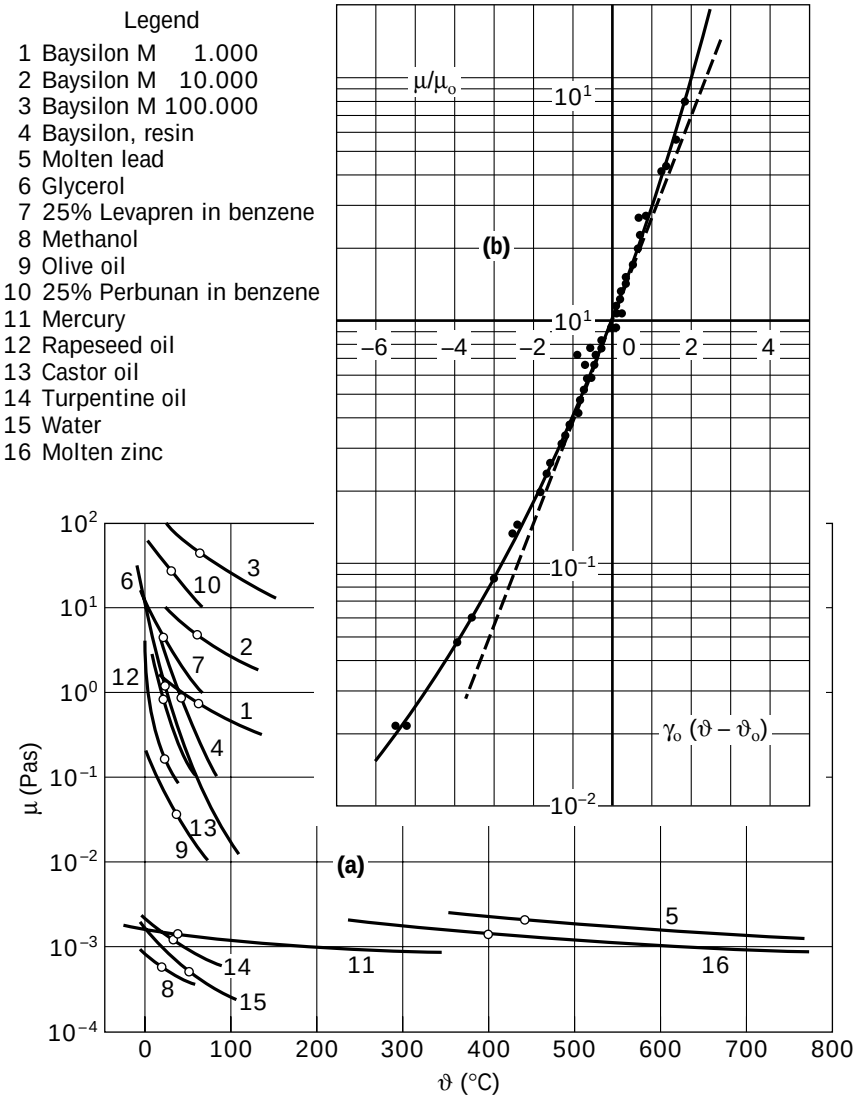


Figure 7. Temperature dependency of the viscosity $\mu(\vartheta)$ and (**b**) its standard representation, which leads to the reference-invariant approximation of the material function. Data points on the curves in (**a**) mark the reference temperatures ϑ_o chosen. From (4, 11). *Source*: From Ref. 3, and Ref. 11.

	Newtonian liquid	**Non-Newtonian liquid**
A	Nu, Re, Pr	$Nu, Re_H, Pr_H, \nu\theta/L$
B	$Nu, Re, Pr, \gamma_0 \Delta\vartheta$	$Nu, Re_H, Pr_H, \nu\theta/L, \gamma_H \Delta\vartheta, \gamma_\theta/\gamma_H$

In *B*, $\gamma_H \equiv \delta \ln H/\delta\vartheta$ and $\gamma_\theta \equiv \delta \ln \theta/\delta\vartheta$ must be added; $\Delta\vartheta$ denotes the temperature difference between the bulk of the liquid and the wall. The conventionally used expression μ/μ_w (where subscript w indicates wall) and $\gamma\Delta\vartheta$ are related by

$$\mu/\mu_w = \exp[-\gamma(\vartheta - \vartheta_w)]$$

Since little is known about the rheological properties of material systems, model experiments must be carried out with the same substance to be used in the full-scale plant. Since

$$\Pi_{\text{material}}(\text{here } Pr_H) = \text{idem} \quad \text{and} \quad \Pi_{\text{rheol}} = \text{idem}$$

the process will then take place within a Π range with only one more dimensionless number ($\nu\theta/L$) than in the case of Newtonian liquids.

However, when scaling up from the model to the full-scale plant, it is impossible to attain complete similarity using the same material system. In the example just described, it is obvious that in keeping ρ, H, θ = idem, we cannot ensure that $Re_H \equiv \rho vL/H$ and $\nu\theta/L$ also remain identical. It is therefore advisable to retain one and the same substance but to alter the scale of the model in the experiments. Pure fluid mechanical processes involving creeping (ρ being irrelevant), steady-state, and isothermal flow are exceptions. In these cases mechanical similarity can be obtained in spite of the constancy of physical properties.

Non-Newtonian liquids with flow curves obeying the so-called power law

$$\tau = K\dot{\gamma}^m$$

are known as Ostwald–de Waele fluids. The dimension of K depends on the numerical value of the exponent m and

is therefore not a consistent physical quantity. To produce a material constant X having the dimension of viscosity, K must be combined with a characteristic velocity, v, and a characteristic length, l:

$$X \equiv K(v/l)^{m-1}$$

With this sort of liquid, the extension of the Π set has the following consequences:

- X replaces μ in every dimensionless number in which it appears.
- Pure material numbers are extended by m (m belongs to Π_{rheol}).

In the case of an Ostwald–de Waele fluid, the heat transfer characteristic of a smooth straight pipe is as follows:

$$Nu = f(Re_X, Pr_X, m)$$

where $Re_X = \rho v^{2-m} l^m / X$.

Viscoelastic Liquids

Almost every biological solution of low viscosity (but also viscous biopolymers such as xanthan and dilute solutions of long chain polymers—carboxymethyl cellulose CMC, polyacrylamide, polyacrylonitrile, etc.) displays not only viscous but also viscoelastic flow behavior. These liquids are capable of storing a part of the deformation energy elastically and reversibly. They evade mechanical stress by contracting like rubber bands. This behavior causes secondary flow, which often runs contrary to the flow produced by mass forces (e.g., the liquid "climbs" the shaft of a stirrer—the so-called Weissenberg effect).

Elastic behavior of liquids is characterized mainly by the ratio of first differences in normal stress, N_1, to the shear stress, τ. This ratio, the Weissenberg number $Wi = N_1/\tau$, is usually represented as a function of the rate of shear $\dot{\gamma}$.

Henzler (13) takes an approach similar to that of the Ostwald–de Waele law

$$Wi = A\dot{\gamma}^a \qquad \text{or preferably} \qquad Wi = A\dot{\gamma}^a + B\dot{\gamma}^b$$

To transform these material functions into dimensionless forms, a reference Weissenberg number Wi_o is chosen. This leads to the generation of the characteristic rates of shear $\dot{\gamma}_1$ and $\dot{\gamma}_2$. It thus follows that

$$\frac{Wi}{Wi_o} = \left(\frac{\dot{\gamma}}{\dot{\gamma}_1}\right)^a + \left(\frac{\dot{\gamma}}{\dot{\gamma}_2}\right)^b$$

In this case the following constituent parts of the dimensionless material function must be incorporated in the relevance list:

$$Wi_o, \dot{\gamma}_1, \dot{\gamma}_2, a, b$$

In the case of a viscoelastic liquid, the material functions of both the viscous and the elastic behaviors must be considered. Assuming that the viscous behavior of the liquid concerned is described as $\mu/\mu_o = f(\dot{\gamma}/\dot{\gamma}_1)$ and the Wi/Wi_o function is valid for its elastic behavior (Fig. 8), the Π set must be modified as follows:

- μ_o replaces μ in every dimensionless number in which it appears.
- The pure material-related numbers are extended by the following six purely rheological numbers: $\Pi_{\text{rheol}} = Wi_o, \dot{\gamma}/\dot{\gamma}_1, \dot{\gamma}/\dot{\gamma}_2, m, a, b$.

Examples of the Use of Dimensional Analysis in Mass and Heat Transfer in Gas–Liquid Systems

Examples 7–10 are presented for two reasons. The first lies in the fact that the compilation of the relevance list (i.e., listing all the relevant parameters of the problem) represents the only real difficulty in using this method, which can be overcome by practice. The other concerns their contents (mass and heat transfer in gas–liquid systems), which are of eminent importance to biotechnology and at the same time show the large potential of information contained in their dimensionless presentation. Clearly, information of this kind cannot be obtained by computational methods.

EXAMPLE 7

Mass transfer in a gas–liquid system in a mixing vessel

According to the so-called two-film theory (Lewis and Whiteman, 1923–1924), the mass transfer in a gas–liquid system (absorption, desorption: "sorption") is described by the overall transfer equation

$$G = k_L A \Delta c \tag{30}$$

G is mass transfer rate through the interface (M/T), k_L is liquid-side mass transfer coefficient (L/T), A is interfacial area (sum of the surfaces of all gas bubbles) (L^2), and Δc is characteristic concentration difference (M/L^3) of the dissolved gas between the interface and the liquid bulk. In the case of surface aeration the difference Δc is given by $c_s - c$. It is assumed that the interface is always saturated with the gas to be dissolved (c_s) according to the Henry's law.

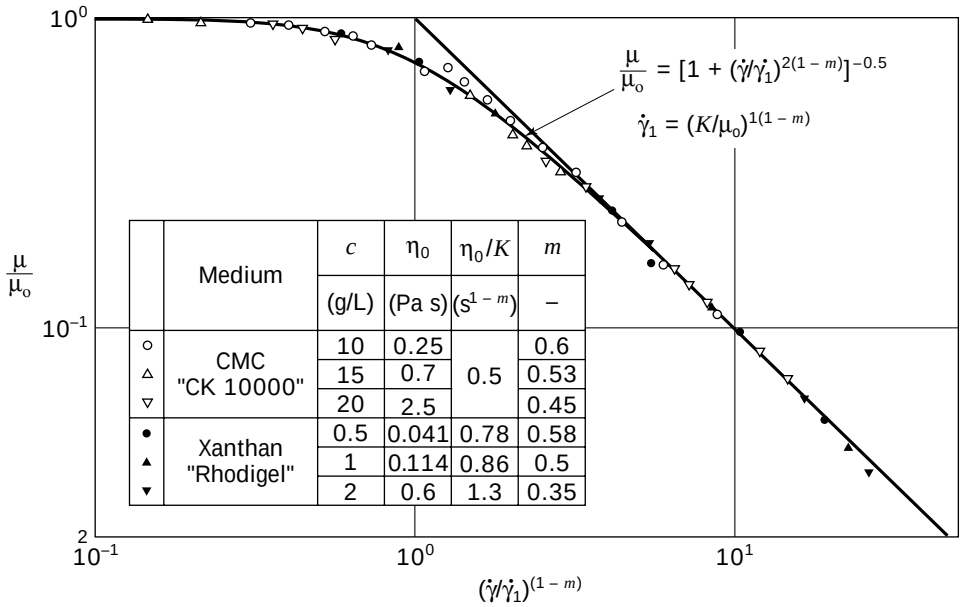
Equation 30 and the foregoing definition of Δc fit well the surface aeration situation. In the commonly used volume aeration systems, it is established practice to assume a uniform distribution of gas bubbles in the liquid and to relate the mass transfer to the volume of the liquid:

$$G/V = k_L(A/V)\Delta c_m = k_L a \Delta c_m \tag{31}$$

The liquid-side mass transfer coefficient k_L as well as the volume-related interface $a \equiv A/V$ can be measured separately only with difficulty. Furthermore, both depend on same physical properties and process parameters. Therefore they are combined to form an overall mass transfer coefficient:

$$k_L a \equiv \frac{G}{V \Delta c_m} \tag{32}$$

The gas composition of bubbles rising in the column of liquid



	Medium	c (g/L)	η_0 (Pa s)	η_0/K (s^{1-m})	m –
○	CMC "CK 10000"	10	0.25	0.5	0.6
△		15	0.7		0.53
▽		20	2.5		0.45
●	Xanthan "Rhodigel"	0.5	0.041	0.78	0.58
▲		1	0.114	0.86	0.5
▼		2	0.6	1.3	0.35

Figure 8. Dimensionless material function of some viscoelastic chemical and biological polymers. *Source*: Ref. 13.

changes as a result of the sorption processes (e.g., depletion of O_2 and absorption of CO_2). Beside this, the system pressure changes along the liquid height. These differences in partial pressures are taken into account by the logarithmic mean concentration difference Δc_m:

$$\Delta c_m = \frac{c_1 - c_2}{\ln[(c_1 - c)/(c_2 - c)]} \tag{33}$$

where c_1 and c_2 are the saturation concentrations under the (p, ϑ, x) conditions at the gas inlet and outlet, respectively, c is the concentration of the dissolved gas in the bulk of the liquid, x is a mole fraction of the dissolved gas.

The use of the volume-related and thus intensively formulated quantity k_La as the target quantity of the mass transfer implies the following consequences:

- k_La should not depend on geometric parameters, since a quasi-uniform material system is assumed.
- k_La must be independent of the physical properties of the gas phase, given that $k_G \gg k_L$.
- Because the target quantity k_La is an intensity variable, the process parameters must also be formulated intensively.

The relevance list must thus be formed with the following parameters:

1. *Target quantity*: k_La
2. *Physical properties*: density ρ, kinematic viscosity ν, diffusivity $\mathcal{D}$, and the coalescence parameters S_i of the liquid phase. (The latter are still unknown by the nature and number for dynamic processes such as aeration in mixing vessels and bubble columns.)
3. *Process parameters*: volume-related mixing power P/V, gas throughput q_G related to the volume of liquid $q_G/V_L \propto q_G/(D^2 H)$ *or* to the cross-sectional area of the vessel (superficial velocity) $v_G \propto q_G/D^2$, and—owing to extreme differences in density $\Delta\rho$—acceleration due to gravity g or $g\Delta\rho$.

We have to come to terms with two problems.

The first one is of a physical nature (S_i unknown). The nature of the liquid treated must be well known (details such as ρ, ν, $\mathcal{D}$ are not sufficient), and model measurements must be performed with the same liquid system used on the large scale.

The second problem concerns the pertinent formulation of the gas throughput. An accurate answer requires measurements with the same material system in differently scaled equipment. The decision in favor of $v_G \propto q_G/D^2$ instead of q_G/V_L was brought about by Judat (14) by comparing extensive research results of the last decade (Fig. 9). Anticipating this work, we compose the relevance list for volume aeration as follows:

$$\{k_La;\ \rho,\ \nu,\ \mathcal{D},\ S_i;\ P/V,\ v_G,\ g\} \tag{34}$$

The resulting pi set reads

$$\{(k_La)^*,\ (P/V)^*,\ v_G^*,\ Sc,\ S_i^*\} \tag{35}$$

where

$$(k_La)^* \equiv k_La(\nu/g^2)^{1/3}$$

$$(P/V)^* \equiv \frac{P/V}{\rho(\nu g^4)^{1/3}}$$

$$v^* \equiv \frac{v_G}{(\nu g)^{1/3}}$$

$$Sc \equiv \nu/\mathcal{D}$$

Quantities S_i are unknown by number and nature. Nevertheless, with the help of known physical properties (ρ, ν, $\mathcal{D}$), we can easily transform them into dimensionless numbers S_i^*.

The extent of influence of the coalescence phenomena

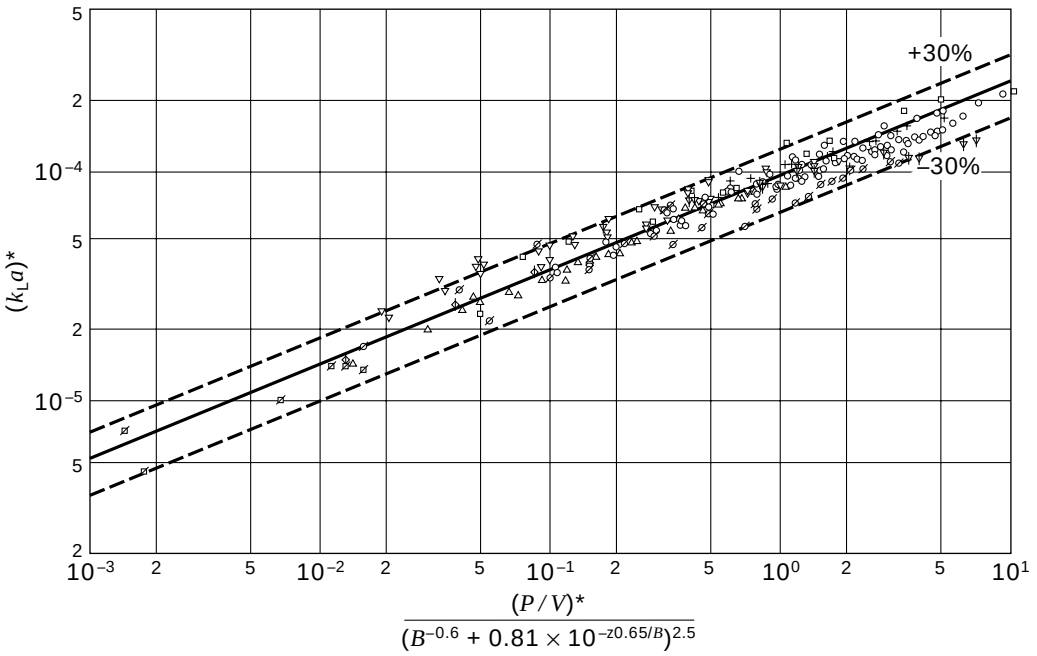


Figure 9. Sorption characteristics for a mixing vessel with turbine stirrer (Rushton turbine) for a fully coalescent water–air system. Comparative evaluation of several investigations covering water volumes of $V = 0.0025 - 906\ m^3$, $B \equiv (\pi/4)\ v^*$. *Source*: From Ref. 14.

on the mass transfer in the gas–liquid system is clearly visible upon comparison of the mass transfer characteristics for the water–air system (unrestricted coalescence; Fig. 9) and for a system consisting of a 1 N solution of sodium sulfite in air (70 g salt/L; fully suppressed coalescence, Fig. 10) (15). Here self-aspirating stirrers (cf. sketches in example 3 and Fig. 2) have been used in which the mixing power and the gas throughput are coupled by the rotational speed of the stirrer. Therefore v_G is not an independent parameter here, and the pertinent pi space is given by

$$(k_La)^* \equiv k_La(\nu/g^2)^{1/3}$$

$$(P/V)^* \equiv \frac{P/V}{\rho(\nu g^4)^{1/3}}$$

$$Sc \equiv \nu/\mathcal{D}$$

For scale-up of surface aerators see Refs. 3 and 16.

EXAMPLE 8

Mass transfer in a gas–liquid system in bubble columns

Bubble columns are preferable in microbiological conversions under aerobic conditions. Perforated coils or sintered plates serving as gas distributors and the superficial velocity v_G present the only process parameters (besides the acceleration due to gravity, g).

When the O_2 uptake in this gas–liquid contacting method is not sufficient, stirrers must provide additional mixing. The turbine stirrers are mounted on the same shaft. The bubble column is thus transformed into a mixing vessel with $H/D \gg 1$ (mostly $H/D = 3$–5), and P/V is an additional process parameter.

In the particularly large bubble columns (tower-shaped bioreactors) increasingly used today for aerobic wastewater treatment (17), the intensification of mass transfer by stirrers would be costly: the stirrer must ensure that the entire contents of the vessel is in motion to generate in the liquid the shear rate needed for gas dispersion. In this case, injectors are the aeration devices of choice: these two-component nozzles utilize the kinetic energy of the liquid propulsion jet to disperse the gas phase into very fine gas bubbles and to distribute them into the liquid. The power of the propulsion jet P_L needed for dispersion of the gas throughput q_G in fine bubbles, P_L/q and the superficial velocity v_G are the process parameters.

Interestingly, in laboratory measurements with bubble columns, the proportionality $k_La \propto v_G$ has been found. This connection is immediately understandable if one considers the constituents of which these quantities are built:

$$\frac{k_La}{v_G} = \frac{G}{V\Delta c_m}\frac{S}{q_G} = \frac{G}{q_GH\Delta c_m} = \text{const}$$

where S is the surface area of the tank, H is the liquid height, and $V = SH$.

It turns out that the absorption velocity G is directly proportional to the gas throughput q_G, to the liquid height H (which determines the residence time of the gas bubbles in the liquid), and to the logarithmic mean concentration difference Δc_m (cf. equation 33).

The pi space in the three mentioned technological implementations is

1. Bubble column with a gas distributor (9): {Y; Sc, S_i^*; geometry of the distributor, aeration density}
2. Bubble column with additional stirrers (9): {Y; $(P/V)^*$ Sc, S_i^*; H/D; type of stirrer}
3. Bubble column with injectors (17, 18): {Y; $(P_L/q)^*$; Sc, S_i^*; H/D; type of injector}

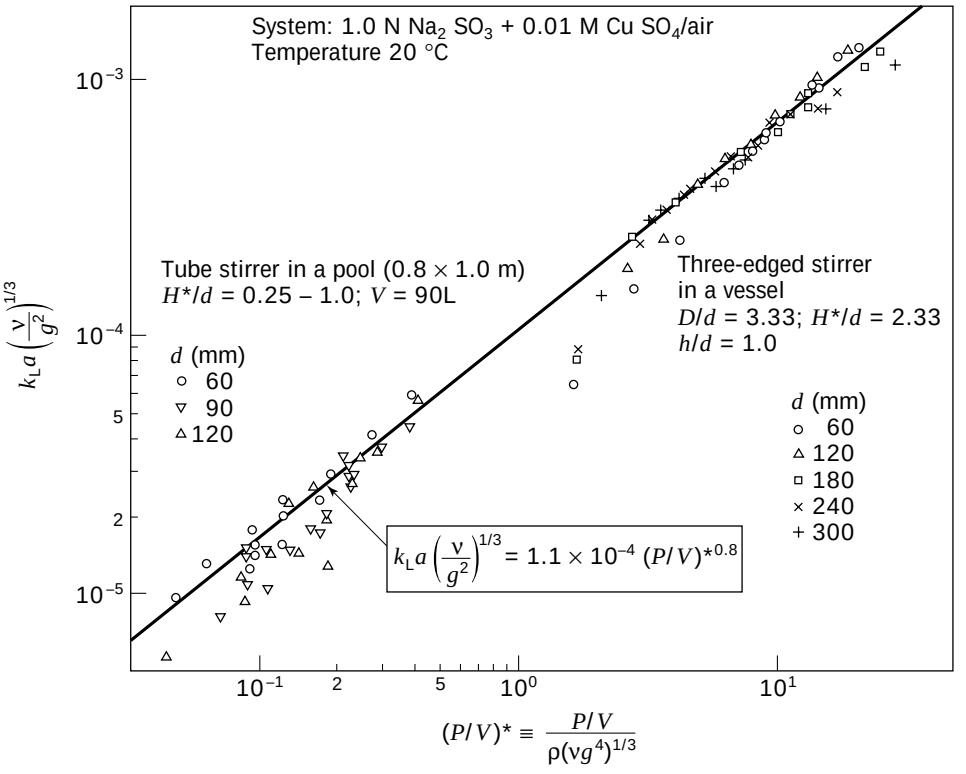


Figure 10. Sorption characteristics for a mixing vessel with two different hollow stirrers (scale factor 1:5) for a fully noncoalescing system. 1 N sodium sulfite solution in air. *Source*: From Ref. 15.

Test liquid	$10^{-6}\nu$(m²/s) 75 °C	25 °C	Rotational speed n (min⁻¹) 5	6	9	12	17	24	32	48	60	87
Water	0.4	0.9									▪	◓
Glycerine, 47%	1.1	4				◐				▽		
Glycerine, 80%	5	35		△		◑		▪		▽		
Glycerine, pure	33	750					◆					
Glucose syrup	74	2180	▽		△			□	◇			

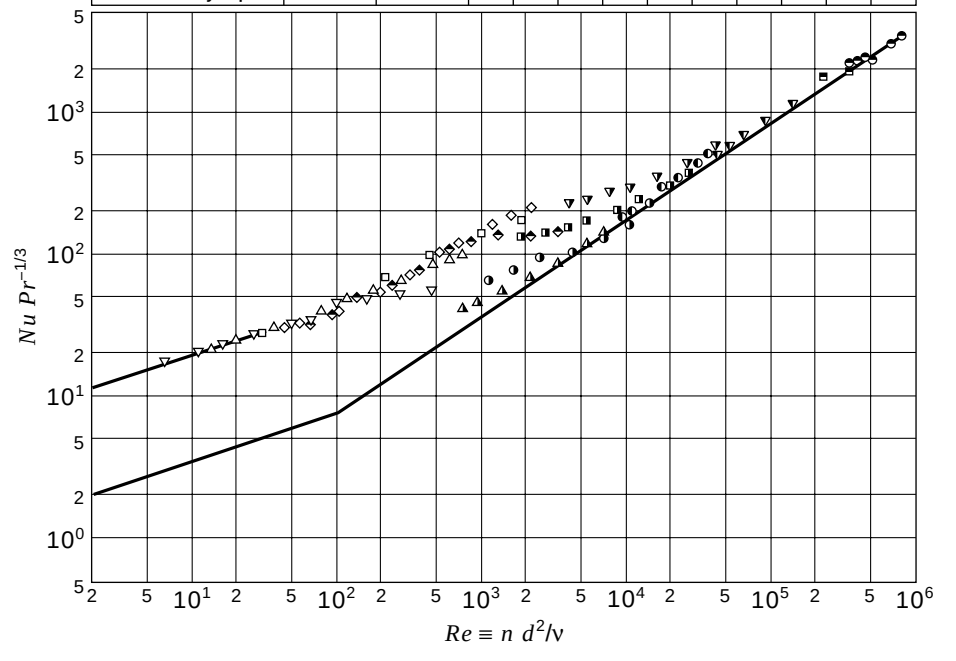


Figure 11. Enhancement of heat transfer at cooling in the laminar flow range ($Re < 10^3$) by wiping the heat transfer surface. Measurements: anchor stirrer with four arms equipped with wipers. Solid curve: anchor stirrer with two arms, D/d = 1.02 (clearance ≈ 2%). *Source*: From Ref. 20.

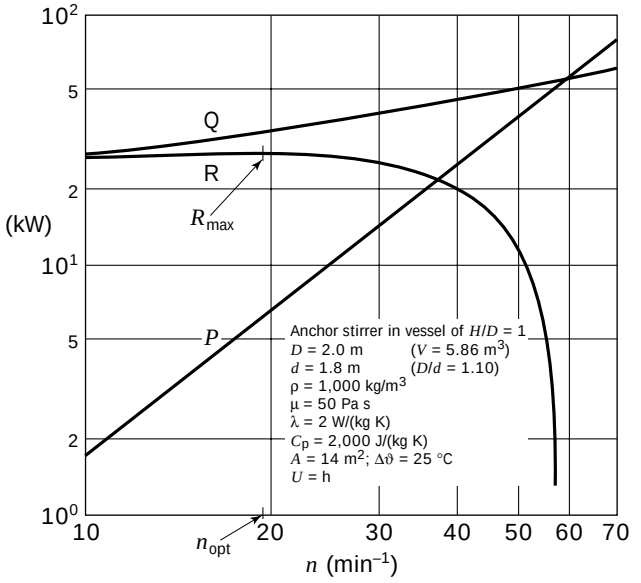


Figure 12. Determination of optimum stirring conditions (n_{opt}) for maximum removal of the heat of reaction R_{max}. *Source*: Refs. 3 and 21.

where

$$Y \equiv \frac{k_L a}{v}\left(\frac{v^2}{g}\right)^{1/3}$$

$$(P/V)^* \equiv \frac{P/V}{\rho(vg^4)^{1/3}}$$

$$(P_L/q_G)^* \equiv \frac{P_L/q_G}{\rho(vg)^{2/3}}$$

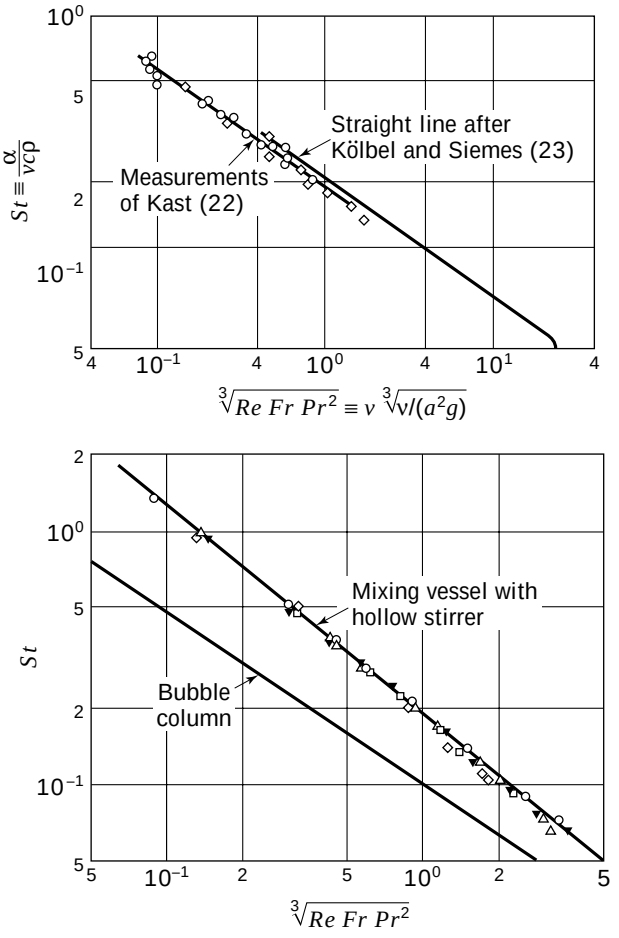


Figure 13. *Top*: Heat transfer characteristic of a bubble column. *Source*: from Ref. 22. Bottom: Comparison of bubble column heat transfer characteristic to that of a mixing vessel equipped with a hollow stirrer *Source*: Ref. 3.

EXAMPLE 9

Heat transfer in mixing vessels

Contrary to mass transfer in fluid–fluid systems, the exchange area A in heat transfer is given and constant, which presents an advantage in dealing with mixing vessels. Here one considers three subsequent steps of heat transfer: namely, the convective heat transport to the outer side of the wall, the heat conduction through the solid wall, and the heat transport from the inner side of the wall to the fluid in the vessel that is to be cooled or heated.

The overall heat transfer equation

$$Q = UA\Delta\vartheta \quad \text{with} \quad 1/U = \frac{1}{h_o} + \frac{\delta}{\lambda} + \frac{1}{h_i} \tag{36}$$

contains thus the overall heat transfer coefficient U as a proportionality constant, which incorporates convective heat transfer coefficients h on both sides of the exchange area and the thermal conductivity λ of the wall of the thickness δ. Each of the heat transfer steps obeys different process conditions and must be described and dealt with separately. For heat transfer at the inner side of the wall, the heat transfer equation reads

$$Q_i = h_i A\Delta\vartheta_i \tag{37}$$

The heat transfer coefficient h_i as the proportionality constant in this heat transfer step is defined by this equation and can be determined only by measurement of its constituents:

$$h_i \equiv \frac{Q_i}{(A\Delta\vartheta_i)} \tag{38}$$

The heat transfer coefficient h_i in a mixing vessel of diameter D, equipped with a stirrer of a given geometry and of diameter d, will certainly depend on the physical properties of the liquid ρ, μ, μ_w, λ, C_p, (λ, heat conductivity; C_p, heat capacity, μ and μ_w, viscosity in the bulk of liquid and at the wall) and on the rotational speed n of the stirrer. (It must not depend on the temperature difference $\Delta\vartheta$, because it is defined by this value: equation 38.) It is assumed that the liquid in the vessel displays Newtonian viscosity behavior.

Then the relevance list is as follows:

$$\{h_i;\ D,\ d;\ \rho,\ \mu,\ \mu_w,\ \lambda,\ C_p;\ n\} \tag{39}$$

In connection with four basic dimensions (M, L, T, Θ) contained in the dimensions of the quantities of this relevance list, we will have to deal with $8 - 4 = 4$ dimensionless numbers:

Table 4. Important Named Dimensionless Numbers

Name	Symbol	Group	Remarks
		Mechanical unit operations	
Archimedes	Ar	$g\Delta\rho l^3/\nu^2\rho$	$\equiv Ga\ (\Delta\rho/\rho)$
Euler	Eu	$\Delta p/(\rho v^2)$	
Froude	Fr	$v^2/(lg)$	
	Fr^*	$v^2\rho/(lg\ \Delta\rho)$	$\equiv Fr\ (\rho/\Delta\rho)$
Galilei	Ga	gl^3/ν^2	$\equiv Re^2/Fr$
Knudsen	Kn	λ_m/l	
Mach	Ma	v/v_s	
Newton	Ne	$F/(\rho v^2 l^2)$	Force
		$P/(\rho v^3 l^2)$	Power
Ohnesorg	Oh	$\eta/(\rho\sigma l)^{1/2}$	$\equiv We^{1/2}/Re$
Reynolds	Re	vl/ν	$\nu \equiv \mu/\rho$
Weber	We	$\rho v^2 l/\sigma$	
		Thermal unit operations (heat transfer)	
Fourier	Fo	at/l^2	
Grashof	Gr	$\beta\Delta\vartheta g l^3/\nu^2$	$\equiv \beta\Delta\vartheta Ga$
Nusselt	Nu	hl/λ	
Péclet	Pe	vl/a	$\equiv RePr$
Prandtl	Pr	ν/a	$a \equiv \lambda/(\rho C_p)$
Rayleigh	Ra	$\beta\Delta\vartheta g l^3/(a\nu)$	$\equiv GrPr$
Stanton	St	$h/(v\rho C_p)$	$\equiv Nu/(RePr)$
		Thermal unit operations (mass transfer)	
Bodenstein	Bo	$vl/\mathcal{D}_{ax}$	
Lewis	Le	$a/\mathcal{D}$	$\equiv Sc/Pr$
Sherwood	Sh	$kl/\mathcal{D}$	
Schmidt	Sc	$\nu/\mathcal{D}$	
Stanton	St	k/v	$\equiv Sh/(Re\ Sc)$
		Chemical reaction engineering	
Arrhenius	Arr	$E/(R\vartheta)$	
Damköhler	Da	$\frac{c\Delta H_r}{C_p\rho\vartheta_0}$	Genuine (cf. Ref. 3, p. 148)
	Da_I	$k_1\tau$	
	Da_{II}	$k_1 l^2/\mathcal{D}$	$\equiv Da_I Bo$
			$\equiv Da_I ReSc$
	Da_{III}	$k_1\tau\left(\frac{c\Delta H_r}{C_p\rho\vartheta_0}\right)$	$\equiv Da_I\left(\frac{c\Delta H_r}{c_p\rho\vartheta_0}\right)$
	Da_{IV}	$\frac{k_1 c\Delta H_r l^2}{\lambda\vartheta_0}$	$\equiv Da_I RePr\left(\frac{c\Delta H_r}{c_p\rho\vartheta_0}\right)$
Hatta	Hat_1	$(k_1 D)^{1/2}/k_L$	First-order reaction
	Hat_2	$(k_2 c_2 \mathcal{D})^{1/2}/k_L$	Second-order reaction

$$\{Nu,\ Re,\ Pr,\ D/d,\ \mu_w/\mu\} \tag{40}$$

where

$\mathrm{Nu} \equiv h_i(D/\lambda)$ Nusselt number

$\mathrm{Re} \equiv nd_2(\rho/\mu)$ Reynolds number

$\mathrm{Pr} \equiv C_p(\mu/\lambda)$ Prandtl number

The heat transfer characteristics $Nu = f(Re, Pr, D/d, \mu_w/\mu)$ are known today for almost all the stirrer types in use. They can be summarized by following process equations:

$$Re = 10^2 - 10^6 \qquad Nu = \text{const}\ Re^{2/3} Pr^{1/3} (\mu_w/\mu)^{-a}$$

$$Re < 10^2 \qquad Nu = \text{const}(RePr)^{1/3} (\mu_w/\mu)^{-a}$$

The constant in these equations depends on the geometry of the vessel and on the type of stirrer. The exponent a of the viscosity term is mostly given as = 0, 14, but in fact it is a function of the μ_w/μ term itself (9, 19):

$$\mu_w/\mu = 10^1 \quad a = 0.18$$

$$\mu_w/\mu = 10^3 \quad a = 0.12$$

Two additional aspects of great interest in biotechnology take into account the requirement, in most fermentations, that the up-

per temperature may not exceed normal physiological temperatures (30–37 °C). This results in a very low $\Delta\vartheta$ when a cryostat should be avoided on an industrial scale.

1. In the turbulent region, h_i can be essentially increased by implementing cooling coils in the vessel and bubble column, respectively. In the laminar range the increase of heat flow is achieved by fastening wipers on arms of an anchor stirrer (20). By this latter relatively simple implementation, the Nu number can be increased by a factor of 10, a demonstrated in Figure 11.
2. Increasing the rotational speed of a stirrer results in an steeper increase of the mixing power P (corresponding to heat flux into the vessel) than that of h_i (corresponding to the heat removal Q from the vessel). Therefore in each case a critical rotational speed exists at which the total heat removal equals the heat flux introduced by the stirrer minus that transferred via the vessel wall. The removable process heat is given by $R = Q - P$ (Fig. 12). The dimensional analytical treatment of this problem is given elsewhere (21).

EXAMPLE 10

Heat transfer in a bubble column

In bubble columns the pertinent process variable is the superficial velocity v_G. This intermediate quantity replaces the gas throughput and the column diameter. Therefore the relevance list reads here:

$$\{h_i;\ \rho,\ \mu,\ \lambda,\ C_p;\ v_G,\ g\} \tag{40}$$

providing $7 - 4 = 3$ dimensionless numbers (22):

$$\{St,\ Pr,\ ReFr\} \tag{41}$$

where

$St\ [\equiv h_i/(v_G C_p \rho)]$	Stanton number
$Pr \equiv C_p \mu/\lambda$	Prandtl number
$ReFr \equiv v_G^3/(v\ g)$	Product of the Reynolds and Froude numbers

Figure 13 (22,23) depicts the dimensionless presentation of the heat transfer in bubble columns as well as in mixing vessels in which the liquid is aerated by hollow stirrers. It is shown that additional mixing roughly doubles the heat transfer. (*Note*: Such comparisons are possible solely by the use of dimensional analysis.)

Table 4 presents a list of often-used, named dimensionless numbers.

NOMENCLATURE

a	Thermal diffusivity ($\equiv \lambda/\rho C_p$)
A	Area
c, c_1, c_2	Concentrations
$\Delta c, \Delta c_m, c_f$	Concentration difference, logarithmic mean difference, and concentration of foamer and flocculant, respectively
C_p	Heat capacity at constant pressure
d	Stirrer diameter
D	Vessel diameter
$\mathscr{D}$	Diffusivity
$\mathscr{D}_{ax}$	Effective axial dispersion coefficient
E	Activation energy
F	Force
g	Gravitational acceleration
G	Mass flow, gravitational constant
h	Convective heat transfer coefficient
ΔH_R	Heat of reaction
k_L	Mass transfer coefficient (subscript L, liquid side)
k_1, k_2	Reaction rate constant (reaction order)
k_F	Flotation rate constant
l	Characteristic length
L	Dimension of length
m	Mass
M	Dimension of mass
n	Rotational speed; number of moles
$p, \Delta p$	Pressure, pressure difference
P	Power
q	Throughput
Q_i	Heat flow on the inner side of the wall
R	Universal gas constant
S	Surface area
t	Time
T	Dimension of time
U	Overall heat transver coefficient, equation 35
v	Velocity, superficial velocity
v_s	Velocity of sound
V	Liquid volume
w	Rising velocity (of gas bubbles or flocs)

Greek characters

β	Temperature coefficient of density
γ	Temperature coefficient of dynamic viscosity
$\dot{\gamma}$	Shear rate
$\vartheta, \Delta\vartheta$	Temperature, temperature difference
Θ	Dimension of temperature
λ	Thermal conductivity
λ_m	Molecular free path length
μ	Dynamic viscosity
ν	Kinematic viscosity
Π	Dimensionless product
$\rho, \Delta\rho$	Density, density difference
σ	(Interfacial) surface tension
τ	Residence time; shear stress

Subscripts

G	Gas
L	Liquid

M	Model, laboratory scale
T	Technological, industrial scale

BIBLIOGRAPHY

1. P.W. Bridgman, *Dimensional Analysis*, Yale University Press, New Haven, Conn., 1922, 1931, 1951. Reprint by AMS Press, New York, 1978.
2. Lord Rayleigh, *Nature* **95**, 2368 66–68 (1915).
3. M. Zlokarnik, *Dimensional Analysis and Scale-Up in Chemical Engineering*, Springer-Verlag, Berlin, 1991.
4. J. Pawlowski, *Die Ähnlichkeitstheorie in der physikalisch technischen Forschung* (Theory of Similarity in Physical and Technological Research; in German), Springer-Verlag, Berlin, 1971.
5. M. Zlokarnik and H. Judat, in *Ullmann's Encyclopedia of Industrial Chemistry*, vol. B2, VCH, Weinheim, Germany, 1988, pp. 1–33.
6. M. Zlokarnik, *Chem. Ing. Tech.* **42**, 1009–1011 (1970).
7. M. Zlokarnik, *Water Res.* **32**, 1095–1102 (1998).
8. M. Zlokarnik, *Ger. Chem. Eng.* **9**, 314–320 (1986).
9. M. Zlokarnik, *Rührtechnik—Theorie und Praxis* (Stirring Technology—Theory and Practice; in German), Springer-Verlag, Berlin, 1999.
10. M. Zlokarnik, *Int. Chem. Eng.* **27**, 1–9 (1987).
11. J. Pawlowski, *Veränderliche Stoffgrössen in der Ähnlichkeitstheorie* (Variable Physical Properties in the Theory of Similarity; in German), Salle Sauerländer, Aarau and Frankfurt am Main, 1991.
12. J. Pawlowski, *AIChE J.* **15**, 303–305 (1969).
13. H.-J. Henzler, *Chem. Ing. Tech.* **60**, 1–8 (1988).
14. H. Judat, *Ger. Chem. Eng.* **5**, 357–363 (1982).
15. M. Zlokarnik, *Adv. Biochem. Eng.* **8**, 133–151 (1978).
16. M. Zlokarnik, *Adv. Biochem. Eng.* **11**, 158–180 (1979).
17. M. Zlokarnik, in H.-J. Rehm and G. Reed eds., *Biotechnology*, vol. 2, VCH, Weinheim, Germany, 1985, pp. 538–569.
18. M. Zlokarnik, *Chem. Eng. Sci.* **34**, 1265/1271 (1979).
19. M. Hruby, *Int. Chem. Eng.* **7**, 86–90 (1967).
20. H. Judat and U. Johenneken, *GVC Ann. Convention*, Stuttgart, Oct. 3–5, 1990.
21. J. Pawlowski and M. Zlokarnik, *Chem. Ing. Tech.* **44**, 982–986 (1972).
22. W. Kast, *Chem. Ing. Tech.* **35**, 785–788 (1963).
23. H. Kölbel and W. Siemes, *Chem. Ing. Tech.* **30**, 400–404, 729–734 (1958).

See also BIOREACTORS, AIR-LIFT REACTORS; BIOREACTORS, FLUIDIZED-BED; BIOREACTORS, CONTINUOUS STIRRED-TANK REACTORS; MASS TRANSFER; SCALE-UP, STIRRED TANK REACTORS.

DISSOLVED OXYGEN. See FERMENTATION MONITORING, DESIGN AND OPTIMIZATION; FERMENTER DESIGN; HYPOXIA, EFFECTS ON ANIMAL CELLS; MASS TRANSFER.

ECONOMICS

HAROLD REISMAN
Biotechnology Results
Weston, Connecticut

KEY WORDS

Aseptic processing costs
Capital investments
Compensation issues
Contract manufacturing
Cost-of-goods
Economics
Equipment cost
Regulatory issues and validation
Scale-up
Start-up
Technology transfer

OUTLINE

INTRODUCTION

The biotechnology industry requires large investment of funds. The market capitalization of U.S. biotech companies was estimated at $93 billion in 1997 (1). Even if the profit motive were to vanish, allocation and expenditure of resources would assume great importance. Assembly and distribution of economic resources in biotechnology are required in at least six general areas. First, there is the need for discovery. Next comes the need for development, scale-up, and pilot production. Following (or concurrently) come testing, evaluation, bioactivity determination, and clinical efficacy. Manufacturing facilities must be funded. Because regulatory and validation issues are becoming more important and more complex, more resources are needed. Finally, compensation issues—which were once included in the relatively minor category of marketing and sales—are now encompassed within new study and negotiation matters that have spawned a subindustry of its own. As third-party payers become larger and as government involvement and regulations grow, economic issues and cost–benefit analyses will grow accordingly.

The current health care system is market driven rather than value driven. It is apparent that economic matters, while always present in biotechnology, continue to grow in importance and are intruding more and more into what was a less complex world of discovery, followed by production and sale of bioactive molecules. Coming decades will emphasize gene-based methods of drug discovery, targeted diagnostic applications, and speedier drug development, and more powerful and more precise treatments can be expected. It seems clear that creation of this new pipeline of pharmaceuticals and devices will be costly for the producer and for the user-payer. The profusion of small biotech companies (their creation—and in some cases demise—can be expected to continue) only adds to the complexity of funding and resource allocation for each of the following phases:

Discovery	Manufacturing
Development and scale-up	Validation and regulatory issues
Bioactivity and clinical issues	Payment and compensation

DISCOVERY

In 1993 there were believed to be some 1,100 specialty biotechnology firms worldwide (with an additional 625–725 firms with an interest in biotech) (2). According to a more recent estimate, there are now 1,274 biotech companies in the United States alone (1). Sales of biotech pharmaceuticals reached $5.9 billion in 1992. Since 1980, biotech companies have raised $20 billion in public markets. Another estimate is more sobering, if not frightening: for each U.S. public biotech company to bring one product to market, a further $40 billion would be needed. These dollar figures point to either excellent discovery or superb programs that will result in discovery. In recent years, the knowledge has circulated that even with superb drug discovery, drug development and drug creation will take much longer than first thought. Venture dollars are apparently shifting from therapeutics to devices and diagnostics (3).

The *Forbes* article (1) gives estimates of total investment (initial public offerings, secondary offerings, private capital, partnering, and venture capital) in 1994 ($3.3 billion) and in 1995 ($6.3 billion). The drug discovery process (including approval) continues to take a longer time. With a lengthening time scale, expenditures rise. The complexity of the clinical testing process coupled with the ever-problematic review procedure means that costs often grow exponentially rather than linearly. A recent time and cost estimate points to an average time for drug discovery and approval of 15 years and a total cost that may reach $400 million (4). More recently, J.L. LaMattina, vice president for discovery research at Pfizer, broached the half-billion-

dollar mark by stating, "The investment, from start to finish, is on the order of $500 million (per new product)" (5). It has been estimated that for the year 1996, expenditures for research and development reached $7.9 billion for the U.S. biotech industry as a whole (1). The entire drug development and approval process is shown in Figure 1 for the general case of a novel drug.

In a decade, the cost estimate for discovery, development, and approval processes has risen by a factor of 3–4. There is no immediate expectation that this ever-increasing cost will be capped, much less reversed. It is clear, then, that expected return (dependent on sales price, cost of production, and market volume) entails greater hurdles and more complexities. The chances for failure exist at any step along the way, and financial loss can also be expected to be greater if any of the hurdles are not overcome.

Discovery does not spawn only technical articles; full-length books have been published on the process of tying science to dollars. One such view is given of the birth of Vertex (6). Even allowing for some melodrama in the narration, this book is a useful guide to the various threads that connect research ideas, academia, and Wall Street. Interesting questions of program selection, methods of concurrent negotiation, and relations with a major academic institution are detailed. As a follow-on, it is reported that Vertex has three candidates in clinical trials and has added a number of research agreements (7).

It has always been clear that genetic information is valuable. Quantifying the value has been difficult. A relatively recent sale of the genetic sequence of *Helicobacter pylori* (by Genome Therapeutics to Astra of Sweden) was made for $22 million (8). Since the fully loaded cost of a research scientist may fall in the range of $150,000 to $250,000 annually, it is incumbent on corporate managers, as well as research managers, to not only make the correct program selections but also to follow up with active direction and necessary alterations to the programs. The *burn rate*—or the dollar outflow in initial stages of a company's history during the period of no sales revenue—is followed by insiders and outsiders alike as a measure of both how much science is being performed and for how long this rate can continue. Twenty-five people in a start-up, while considered quite small, might have a burn rate of $5 million annually.

There are good sources of ideas on how to make appropriate decisions (9) and also on how to manage research and development (R&D) more effectively (10). These references describe results of extensive surveys of industrial (including biotech) organizations and give useful directions for improvement and appropriate management of R&D and technology transfer.

Current methods for discovery of leads are discussed by Norrington (11). He estimates that discovery and preclinical work could cost $10 million. Many candidates will fail at this point and others will continue in development, cost more money, and fail at a later date. The estimates given for cost of phase I trials is $2.3–7.5 million, for phase II, $7.5–15 million, and for phase III, $125–250 million. High-throughout screening (HTS) is discussed; devices for screening 1,500 samples/h with one or more readouts (spectrometry, radiometry, fluorometry/luminometry) are described. Instruments are now being developed to handle 864 well plates; concurrent testing using multiple screens is strongly suggested. Bevan et al. (12) also stress HTS screening (23 references) but provide, as well, comparisons with rational design and combinatorial synthetic methods.

A very detailed article on resource allocation at SmithKline Beecham should be read (13). Although there

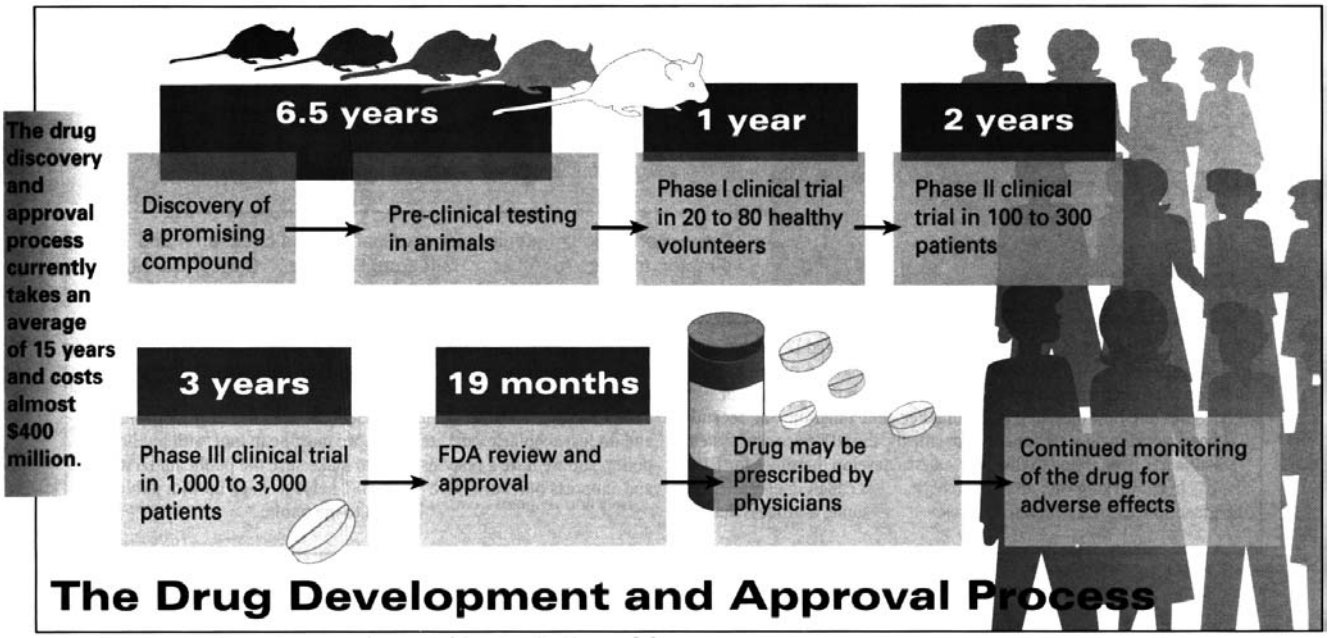


Figure 1. The drug development and approval process. *Source*: Pharmaceutical Research and Manufacturers of America, Washington, D.C.

are alternate scenarios for resource allocation, the authors carefully describe the thought processes in designing a better method for decision making. The basic steps are generation of alternatives, valuation of alternatives, and, finally, creation of a portfolio and allocation of available resources. The very important soft issues (e.g., information quality, credibility, trust) are faced and, so far as possible, defined and controlled. The effort to reach the improved allocation process took about 2 years and was implemented in late 1995; its success can only be evaluated by future results.

Although every potential area of biodiscovery cannot be covered here, two examples point to the need for continual updating of existing technology. The antibiotic market worldwide is estimated at $2.3 billion (14). In the decades following World War II, there were predictions that microbial infection could be halted, or suppressed to the extent that further work on finding novel antibiotics would be unnecessary. The facts are otherwise; drug-resistant strains have evolved in diseases such as tuberculosis, malaria, cholera, and pneumonia. The 160 antibiotics on the market may have been overprescribed or improperly used, or both. The result is a growing research effort in areas of new drug discovery, genome sequencing, and development of new vaccines. Large companies have formed partnerships (Pfizer with Microcide and Schering-Plough with Genome Therapeutics) to speed discovery.

Current antimicrobial therapies and newer agreements are described in a short review (15). The regrowth of chiral chemistry in biotechnology has been related to greater regulatory stress, tighter environmental controls, and cost control (16). Both resolution and asymmetric synthesis are being pursued. Atenolol (for hypertension) and albuterol (for asthma) are enantiomers, each of which exceeds the billion-dollar mark in annual sales. Patent protection is another reason for moving toward the single isomer (17). Drug information derived from use of the racemate may be used, in many cases, to speed clearance of the active (or more active) isomer. Schuster and Menke report that of 1,327 therapeutic compounds, 528 were chiral; of the 528, only 61 were used therapeutically as the single isomer (17).

It should be recognized that the recent emergence of "fully integrated discovery organizations" is also altering the discovery landscape (18). A company (it can be very large or very small) acts as a coordinator of discovery and/or development that takes place outside its boundaries. External companies (there can be one or more working on the same program) supply the research expertise that was once the prerogative of the pharmaceutical company itself. Turner et al. stress biocatalytic processes and report on annual production of various bioactive compounds (18). A cautionary note should be included. Discovery is not easy, technology transfer is no less complex, and scale-up has its own set of major hazards. The culture chasm (6,19) between scientists and business managers in biotech firms was vast a decade or so ago and, in many ways, bridging the gap has not been wholly successful. The scientific "set" and the business "set" tended to run their own shows, had different time scales, had few "common points of reference," and experienced trouble communicating and cooperating. Unless these real and potential difficulties are recognized, confronted, and overcome, no amount of money coupled with "good science" will result in a useful and usable device, drug, or therapeutic agent.

DEVELOPMENT AND SCALE-UP

In 1980, 2% of all prescriptions were generic; by 1993, half of all prescriptions involved generic drugs (20). In the 1993–1999 interval, patents will expire covering branded drugs with annual sales totaling $20 billion. Pisano and Wheelright give the current estimate of $359 million (as of 1992) to develop a new drug (20). This is about threefold higher than the 1987 estimate ($120 million). In the 1990s, manufacturing cost has risen to 20% of sales. All these drivers point to the need for more speed and more cleverness in process R&D. A general review is useful to cover rules of successful development (21).

Mukesh gives 15 steps for improving the success rate of process development (22). Although the overview is general, two points require special emphasis in biotech development. The first concerns the need to collect information on raw materials' sources and purities (and, one might add, impurities as well). The second emphasizes the desirability of having scientists and engineers who develop the process involved (or at least available) at various stages of advancement (or scale-up) until the process is fully commercialized. These two points are often neglected or glossed over in biotech scale-up.

One specific and useful example of a complex and successful scale-up of a bioprocess is available. Prior et al. discuss process development in the manufacture of inactivated HIV-1 (23). A process flow diagram (PFD) is given; all operations occurred within a BSL-3 facility. Cost effectiveness in scale-up is emphasized. A scalable chromatography process is described.

Another review (with 81 references) covers newer concepts in bioprocess R&D (24). Certain operations are covered in depth (mixing, foam control, use of membranes), and other subjects are reviewed and appropriate references given. Subjects include on-line analysis, estimation and control of the physiological state, automation, and error reduction. No developmental program can be expedited and relied on to generate meaningful data without a modern pilot plant. A dozen authors contribute to a review that provides 20 detailed tips on how to keep a pilot plant on the fast track (25).

Although more mundane, an article by Merck workers (26) should prove useful to anyone involved in a cell culture or fermentation pilot plant. The Merck facility was first commissioned in the early 1980s, and significant improvements have been made in hardware, maintenance, equipment replacement, instrumentation, and cleaning. Examples are given of selected upgrades. The facility has fermentors ranging from 0.28 to 19 kL in scale. The message highlighting the need for continuous improvement is clearly conveyed. In biotech operations, what was voluntary but sensible has now come under regulatory review. That is, appropriate documentation (even at the pilot scale) is more essential than ever.

Abate et al., who give careful detail on how and why effective documentation of pilot plant operations is to be

achieved and offer pointers that are directly applicable to production-scale operations as well (27). The process operating log should not only include methods involved in change control, it should give the project background and list equipment to be used, as well as auxiliary systems and services. The log also should describe the process itself, identifying materials, safety issues, and process preparations. An equipment log, data acquisition and detailed process operations, and start-up and shutdown checklists should be included. Examples of forms are given (27).

In any aseptic culturing, agitation issues are invariably involved. Change is commonplace, and in recent years there has been a shift to the use of hydrofoils for improved mixing with reduced (or constant) power consumption. McFarlane and Nienow present an excellent review on this subject and follow with air–water experimental studies (28). Dream has reviewed the various bioreactors available for bioprocessing (29). Not only is there a discussion of various fermentor designs, there is further analysis of problems related to the use of serum, controls to use (including ORP), sterilization, effluent handling, surface polishing details, and relief devices. Economic factors, as well as safety issues, are discussed under the overall heading of continuous mammalian cell culture (30).

The use of batch or continuous processing is based on capital investment (usually lower for continuous processing), quantity of material to be processed, genetic stability, reliability of asepsis, and degree of flexibility needed in the production plant. Design and scale-up of monoclonal antibody production is also reviewed (31,32). Reference 32 has an exhaustive bibliography with 156 references. Reactor designs are reviewed and scale-up considerations discussed in light of the following inputs: government regulation, market requirements, supplier capabilities, capital budget limitations, cell line idiosyncrasies, existing plant layout, and alternate process systems. Actual cost data (using dimensionless factors) in an industrial setting are given by Lonza workers (33). Both fermentation and isolation are considered, and optimization covers the total process rather than one part or another. A fine chemical, L-carnitine, is the subject, and costs of raw material, personnel, overhead, and depreciation are included. A fed-batch operation gives the lowest relative cost. Since this nutrient is actually being synthesized at production scale, the article by the Lonza workers is not only fairly unique but interesting as well.

A very fine article by Middelberg (34), incorporating economic analysis in production of a recombinant protein, once again emphasizes that maximum yield in one step (biosynthesis) will not always give the optimal global solution; in the case study, maximum yield will lead to a suboptimal solution. Protein refolding introduces significant complications. The analysis incorporates direct fixed capital, consumables, waste treatment, and other costs (e.g., labor, utilities, and R&D). Many plots that show interaction of parameter estimates are given. Another paper includes the very important concept of process controllability as a major input in conceptual design (35). Conventional steady-state economic design does not include operability and control capability of each design, that is, the usual tendency is to merely select a minimum cost alternative, assuming that the plans will give the same degree of process controllability. This is a hazardous assumption. The methodology presented incorporates dynamic controllability into the overall assessment and so is a unique contribution. There are two additional examples of preliminary process design including some economic analysis. The enzyme polygalacturonase is manufactured with a 90% recovery from whole broth (36). A continuous lactulose process (including separation and purification) is described (37).

In design and scale-up of bioprocesses, one relatively new development should be considered whenever any aseptic operation is planned. In the past, filling or bottling was the only thing that came to mind when the word *asepsis* was noted. Now, however, novel cell therapies are in use and medical products are incorporating living cells. Products, for one reason or another, cannot undergo terminal sterilization. Clean rooms—or better, controlled environments—are now often considered processing space. It has been estimated that worldwide sales of clean rooms and related equipment reached $2.7 billion in 1994 (38). By the turn of the century, this figure is expected to exceed $4 billion.

The newer concept of a minienvironment or microenvironment (under the general heading of barrier technology) can result in lower investment, lower operating costs, higher product yield, and increased operator comfort. Costs for clean-room space escalate dramatically as particle size control becomes more stringent. Table 1 gives estimated cost data. Currently, class 1 and class 10 areas are found in the electronics industry; in general, class 100 is found to be acceptable for aseptic operations. Given these costs, it is obvious that overdesign would be prohibitively expensive, and efforts are under way to limit the working volumes for class 100 (or better) spaces. Regulatory requirements worldwide are being harmonized, but one should strive to satisfy current regulations of the various governing bodies; it is possible to meet the U.S. and EC guidelines without incurring onerous costs (39).

Issues in design of production facilities for cell and gene therapy are reviewed in detail (40). A range of $250–350/ft^2 is given as total installed cost for a class 10,000 space, but this includes provision and commissioning of walls, ceilings, HEPA filters, low wall returns, lights, fire protection, roll cove flooring, HVAC, controls, ductwork, and minor process support services. The requirements (such as 100% once-through air) for cell and gene therapy (same

Table 1. Cleanliness Is Costly

Class	Cost per square foot
100,000	$100
10,000	$200
1,000	$300
100	$400
10	500^+$
1	$1,000–4,000$^+$

Source: Excerpted by special permission from *Chemical Engineering*, May 1995. Copyright © 1995, by McGraw-Hill, Inc., New York.

space classification) would increase costs to $400–450/ft^2. There are a number of useful references concerning design and cost of barrier isolation systems (41–43) and the use of isolator systems for sterility testing (44). A good overview by Akers (45) discusses numerous applications and gives comparisons of start-up costs and operating expenses for conventional clear rooms and isolator designs. Table 2 gives details of this comparison and shows the significant operating savings attainable with use of isolator technology.

The final major product of a development program is successful technology transfer. (This is equally important during the transfer of information from the research to the development organization.) A number of reports will be included, and there should be continuous interfacing between the groups involved in the transfer of the technology. An appropriate schedule should be established. Each organization has its own methodology; however, Table 3 represents an attempt to list all the information and factors that should be documented and communicated before and during the technology transfer. Any reader who has been involved in such a transfer will no doubt be able to insert additional factors.

It is surprising that insufficient emphasis is often placed on the process of technology transfer. Large, even enormous, sums are expended in research and discovery, even more might be expended in building a new facility, and an overarching goal is established to complete the technology transfer in the shortest possible time with minimum planning input. One might even view some cases of technology transfer as a sort of necessary evil to be completed as quickly as possibly so that the product can be made. It is as though time spent on technology transfer were somehow wasted or unnecessary. The results of such an approach are often painful to see, and the costs incurred (although speeding the process) sometimes dwarf other, more carefully planned costs.

Technology transfer is one of the most important steps in bringing a product to market and at minimum, the planning effort expended here should equal that devoted to all aspects of research, development, scale-up, and construction of a production plant. It is suggested that a budget (time and manpower) be created as part of the technology transfer process and the entire process receive the same high level managerial attention as other parts of the project cycle. An interesting case study is available for a real-life manufacturing development project (46). The product is a recombinant β-interferon. Subjects discussed include team creation, project scope, and timing and feedback. The plan covered some 2 years. Not surprisingly, not everything went according to plan. The difficulties noted are not unique to Biogen, and the lessons learned are probably applicable to many other biotech and regulated companies. There is also a useful appendix that summarizes the U.S. regulatory process for pharmaceuticals.

BIOACTIVITY AND CLINICAL ISSUES

This section could as easily be located after “Manufacturing” (later) but in the normal course of events, bioactivity

Table 2. Operating Expense Comparison: Isolator System versus Conventional Clean Room

	Isolator system	Clean room
Direct operating expenses (annual):	$10,300	$23,040
	Replacement suits, sleeves, gloves, gaskets, etc.	Full gown set ($12)
		2 operators/day, 240 day/yr
		4 gown sets/operator/day
Gown-changing time:	$530	$5,330
	1 min/operator/change	10 min/operator/change
	4 min/day/operator	40 min/day/operator
	53 h/yr @ $10/h wage	533 h/year @ $10/h wage
Maintenance/decontamination/labor:	$4,420	$14,050
	3 h/week = 156 h/year	16 h/week = 832 h/yr
	$10/h wage	$10/h wage
	1 VHP cartridge/week	52 disinfections/yr
	$55/cartridge	4 gown sets/week
Utilities cost (kW h/yr):	$405 (3,372/yr)	$2,775 (23,126/yr)
@ $0.12/kilowatt hour	7 blowers @ 55 W	8 blowers @ 330 W
	24 h/day	24 h/day
	365 days/yr	365 days/yr
DOP testing:	$60	$480
	Every 6 months	Every 6 months
	1 operator/3 h, $10/h wage	2 operators/12 h each, $10/h wage
	Expense summary	
Start-up expenses	$265,100	$284,140
Annual direct expenses	15,715	45,675
Annual indirect expenses	None	300,000
	$280,815	$629,815

Source: This table originally appeared in J.E. Akers, *BioPharm* **7**, 43–47 (1994), with the permission of Advanstar Communications, Cleveland, Ohio. This is the corrected version (October 1994 issue, vol. 7, no. 8, p. 10).

Table 3. Documentation and Factors Involved in Bioprocess Technology Transfer

1. *Raw materials and medium composition*	
Specifications and QC tests	Shelf life
Order of addition	Storage conditions
2. *Cell strains*	
Validation and QC tests	Growth curves
In-process tests	Preparation of banks
3. *Process flow diagram (PFD)*	
Cycle times	Operating conditions
Labor use	Record (data) keeping
Sterilization sequences	Shipping studies
Compounding/packaging	
4. *Documentation*	
Written procedures	Pictures (or videos)
Process response data	In-process tests (and frequency)
Emergency situations	Start-up and shutdown
5. *Training requirements and schedule*	
6. *Acceptable performance*	
Critical variables	Inventory min/max
Acceptance criteria	Operating parameters/ranges
Finished goods variation/testing	Troubleshooting
7. *Process validation*	
Document review	Time per step
Data comparison (various scales)	Batching
QC tests	Stability and shipping studies
Master batch record	Permissible recycle
8. *Instructions for use*	
User manual	Package insert
Labels	Storage by customer
Accessories	Warranty
9. *Safety and sterility*	
Hazards	Environmental testing
Asepsis (precautions)	Other precautions
Special equipment	Cleaning (and test) procedures
Record of asepsis achieved	Off-spec material handling

determination and preclinical activities occur prior to start-up of a manufacturing plant. Initial quantities (which could amount to grams but might be hundreds of kilograms) are usually produced at large-laboratory scale or in a pilot plant. If preliminary data are promising, production plant design could be accelerated and orders might be placed for long-delivery items, but that would also mean creation of greater financial exposure. Early clinical trials (even phase I studies) might involve a few to a few dozen subjects; material to be evaluated normally is pilot plant product. Of course, many steps occur concurrently; we consider this subject at this point for convenience, with the recognition that clinical trials and clinical issues continue through full-scale plant construction and operation and for months or years afterward.

One reference already noted (43) gives the needs (including minimal capital outlay) for creation of validatable clinical production facilities. In a 1994 review, a time and cost flowchart is presented (47). Preliminary development is said to take 1–3 years, and cost is in the range of $0.5–5 million. The success rate is stated to be 5–20%. This presumably means that from 1 in 5 to 1 in 20 candidates is deemed worthy of proceeding to the next hurdle. The next step is called investigational development, and the period of this second stage (of three) is anywhere from 3 to 6 years; cost is $5–100 million. The success rate here is given as total of approximately 20%, or 1 in 5. The final stage is review by the U.S. Food and Drug Administration (FDA); cost range is $3–100 million, and time frame is 1–3 years. Success is 1–5%. It would seem that this, too, is an overall rate (of 1 in 100 to 1 in 20). There might be drugs or devices that would achieve clearance in 5 years (from discovery) and incur a total cost of $8.5 million, but these are either very simple devices or molecules, or they represent a vanishingly small segment of materials or substances that weather the difficulties described. Fitzpatrick and Mackler posit that before 1990, biotech companies could measure or expect success based on a technology portfolio, an intellectual property position, and access to capital (43). In the 1990s, regulatory management has apparently become the primary element of success. There are three reasons for the failure of the FDA to approve a given therapeutic after a clinical trial:

1. Failure in clinical design
2. Improper (or failure in) protocol compliance
3. Deficient GCP documentation

Since moving along the success curve (discovery to investigation to FDA review) means an exponentially rising dollar outlay, it makes sense to

- Plan from end to beginning (i.e., understand and define the end product, and all its requirements, and plan backward from there).
- Integrate all elements of the process.
- Conduct more phase I and II trials or include more variables; speed in starting phase III is not an unalloyed blessing.
- Be responsible and flexible (i.e., have alternatives for every potential holdup or problem).
- Seek the right people within and without the organization.
- Validate and reality-check *often* (i.e., perform many in-house reviews, studies, and audits). Find problems and defects before others do.

The clinical supply process requires a dedicated team and demands dedicated protocols. Ehrich et al. cover manufacturing, packaging, and labeling (48). The preapproval inspection of the FDA now includes an evaluation of the development history of the product. This invariably includes an evaluation of the clinical supply operation. Ehrich et al. stress the need for a manufacturing, packaging, and labeling protocol (called an MPL) and provide a sample table of contents for a typical MPL protocol. Required details are listed for the test drug, any placebo, and any positive control drug in the evaluation. It should be clear that if all this effort is properly applied, the next-generation clinical formulations will be more rapidly created (in a documented manner), and technology transfer to full-scale manufacturing and concurrent phase III trials will be speeded and, in fact, simplified.

Another report describes a semiautomatic system for R&D and clinical use for liquid-filled hard gelatin encapsulation (49). The authors and client believe that the device is well suited for small-scale batches and phase I and II studies. This reference is useful not only for the specific technology but also for the necessary procedures, testing, and validation, which can be extrapolated to any such device used for small test quantities.

It is also not uncommon to use a contract research organization (CRO) to assist in performing a clinical trial. The CRO's contribution may include formulation, dose development, protocol development, shipping, report gathering, and even analyzing results. Of course, the CRO may perform one or more of the functions involved in the overall scheme. For smaller companies, it probably is more efficient in use of time and money to hire a CRO than to do the work in house. Many of the issues relating to CROs are covered by Vogel et al. (50). The contract between the sponsor and the CRO is very complex, and its preparation is very time-consuming. Elements of the contract are given in a very detailed fashion, and relevant regulatory documents and references are shown. In 1995, *The Biomedical Engineering Handbook* was published (51). This work contains a wealth of information (relating to medical products, in the main), but two sections are relevant in the context of bioactivity (or utility as far as a medical product is concerned). A section by Oberg covers assessments for effective product development and another segment by Holohan covers health technology assessment (51). Payment and compensation issues are covered in the following, but a reasonable model of value of a proposed product should be in hand even before early clinical testing begins. Although the model certainly will be modified, it is well understood that product approval without subsequent sales is not a successful outcome.

MANUFACTURING

Starting some 15 years ago, there was a marked increase in both quantity and quality of articles and other publications involving economic issues in biotechnology. The number of these publications remains quite low compared to the burgeoning growth of all biotechnology literature. Still, the improvement noted is a positive development and speaks for maturity of the field. It is likely that before 1980, there was a tendency to protect real or perceived economic secrets. With growing interest in the field (including academia and the government) and with the understanding that many of the so-called proprietary matters were becoming common knowledge, there was an opening that continues to evolve. One well-known handbook (52) has a chapter entitled "Principles for Costing and Economic Evaluation for Bioprocesses." The first edition of this handbook (1983) had no such chapter, and very little economic information was included in the chapter entitled "Processes." The latest edition (1991) includes definitions of economic terms, various economic evaluators, indices for estimation, equipment, and operating costs (mostly 1980s), and useful information on downstream processing. Economic aspects of some traditional processes (ethanol, citric acid, single-cell protein, monosodium glutamate) are included.

A recent book (1995) on principles of fermentation technology, which could easily be used as a text, has a good chapter on fermentation economics (53). Containment issues are covered, and the point is once again made that alternative processes in extraction and purification should be evaluated concurrently in the development phase before commitment is made to a single design. The importance (and cost impact) of effluent treatment and the need for control of waste streams are emphasized. Economic matters in biotechnology were either poorly covered or not mentioned in many of the earlier introductory reference books. This was due more to the lack of information in the literature than to any disinterest on the part of the authors.

Cost estimation for biotech purification is covered by Dwyer (54), in a chapter in a book on protein biochemistry. Economic issues and the processes covered relate to high-value peptides and proteins. There is a review of available equipment and various potential processes; scale-up and related costs of chromatographic purification processes are given. Overall process design is discussed, and both equipment and operating costs are given. In an entire book devoted to economic analysis of fermentation processes, this author takes as his theme the project cycle and how economic matters impinge at every step from discovery (research planning) through plant commissioning and commercial product shipment (55). The specific example

followed is a citric acid process and plant. Profitability matters are covered. One or more of these references (52–55) should be consulted so that terms are understood and available sources for further estimation are known. The many examples provided will be helpful for economic evaluation and estimation of any novel biotechnology process.

Related publications should be consulted for an understanding of where manufacturing fits and for an appreciation of facility planning including estimation of costs. Durkin (56) reiterates the point that cost of goods (COG) as a percentage of pharmaceutical sales has doubled in the last 20 years. Both cost pressure and selling price pressure demand that manufacturing strategy be developed and incorporated into the business plan. In the past, manufacturing was more or less remote from mainstream business activity; organizational inefficiency is now being remedied. The results should be speedier transition from R&D and more facile product launch. Durkin presents an interesting play on words: a biotech organization can either be "leading edge" or "bleeding edge."

Part of the determinant is the success with which manufacturing strategy is incorporated as a key element in the overall plan. Biotech facility planning has a complex set of unique requirements that must be incorporated into the more mundane and better known set of needs. Hubbard (57) gives key steps in preparing a design brief, defining project goals, and stating needs prior to preparation of a design. One step involves preparation of individual room data sheets with appropriate details; another covers the adjacency matrix and a spatial relationship diagram. A tremendous effort is needed prior to final design and commitment of major capital. A few of the issues to be mutually determined (between the biotech firm and the engineering/architectural firm) are plant room size, room location, horizontal services, service chases, vertical services distribution, and maintenance access. Many firms continue to face the need for conversion or adaptation of existing facilities.

A real-life example involving a European firm, Boehringer–Ingelheim, is available (58). A dedicated facility (single-product use) was built in the period 1985–1987 and totaled 110,000 ft^2. It was decided to change to a multiuse facility, where three different products could be made concurrently. Recognized limitations were noted in inoculum development, protein purification, and formulations areas; plans were developed to remedy the situations and also convert the HVAC system to give once-through air. Some 50,000 ft^2 of space was added. The project took 3 years to complete; successful start-up occurred in 1995. Planning and execution included a strategic capacity capability (equal to 30% of overall throughput). This is a good point to remember regardless of whether the plant being built from the ground up or converted. Added capacity for developmental runs or unforeseen new products should always be considered; the added investment may prove itself many times over.

Another case study in facility design concerns a Fujisawa product, FK 506 (tacrolimus) (59). Investment in the plant, located in Ireland, was £17 million. Conceptual and basic engineering occurred in late 1990, and the grassroots facility involved 6,000 m^2 of space. The pure bulk material was made in Japan and shipped to the site. Isolators and full air suits were installed for certain steps. Detailed design started in January 1991, and 15 months later the plant was handed over to the owner. The plant underwent a successful FDA inspection in 1993.

At some point in the developmental cycle, usually after positive bioactivity and/or clinical data are available, it becomes necessary to determine capital needs for a production facility. Cost estimating factors have been presented specifically for biopharmaceutical process equipment (60). Exponential scaling factors are given for 58 different types and sizes of equipment. Use of indices is explained. A cost scaling factor of 0.63 was the average value for all equipment covered, but the standard deviation was 0.21. Although the historic 0.6–0.7 scale factor can be used when no other information is available, overuse or inappropriate use might result in introduction of large errors in final dollar cost estimates.

Other articles, while not specific for biotech processes, are useful for newer insights in estimating costs. Use of a single scale factor can lead to problems unless there is very careful definition of auxiliary items and off-sites (61). Biotech plants often involve extensive investment in items usually grouped as outside battery limits. It is best, therefore, to calculate capital for auxiliaries in terms of product capacity in order to determine economic significance.

Life-cycle costing estimates the total cost of ownership over the entire life of the system (62). The system with the lower total cost (not lowest initial cost, necessarily) over the plant life is the optimal solution. Life-cycle costing includes process control alternatives, failure event and associated costs, and, perhaps most important, costs of lost production. Normally this technique requires use of a computer spreadsheet. Costs are usually divided into six categories: initial acquisition, initial installation, operating, maintenance, lost production, and decommissioning.

A general review on biotech manufacturing economics is available (63). Building and retaining inefficient or expensive-to-operate manufacturing facilities (often due to the need for speed to market) will cause serious problems in the long run. Multiple-track (or at least two-track) development strategy is stressed to counter such an occurrence. McKown et al. provide no hard numbers, but the information, especially as it relates to purification, should be studied for insights into the early design phase (64).

Cost modeling as a management tool is explained by Busch (64). This development tool starts with a defined process and a PFD. The assumptions used in creating the cost algorithm (materials, tooling, labor, space, energy, maintenance, capital equipment, time value of money) must be defined and understood. Comparative cost estimates and sensitivity analyses can be performed; the economic bottlenecks can be identified. The model and underlying assumptions can be followed and modified as needed in the course of development. This modeling method can and should be continued into full manufacturing, to facilitate the improvement of subsequent models as more realistic assumptions (based on experience) become available.

For those wishing to prepare an early cost estimate (without initiating a full-fledged engineering project), it is often difficult to find costs and installation factors. This is

especially true for small projects (investment of less than $5 million) or for projects involving retrofitting or internal alterations. One engineering firm that is involved in biotechnology projects has supplied certain cost data, including installation, that will be of use for initial estimation purposes. The information will also be of use in negotiating details with in-house or external architectural and engineering firms. The inevitable differences between estimators notwithstanding, the sponsor should be aware of why costs are far from either published information or any other up-to-date cost estimate. Table 4 gives costs for the northeastern United States as of the middle of 1996. Escalation factors are available (noted earlier) for changes due to inflation.

New tools continue to be refined and expanded. Steinberger discusses reengineering the capital investment process (65). The key point is to produce a detailed evaluation at a very early phase (2–5% of completion) of design engineering. Computer programs are now available to unify the stages of design; process simulation software can be used in conjunction with detailed capital evaluation software. Rapid changes can be made, and many alternatives can be evaluated in a short time. This is important because about 80% of a project's capital investment is determined when conceptual design is completed. Screening multiple potential designs, rather than one or two, is clearly advantageous in selecting an optimum configuration. This capability is of great utility in discarding certain potential pathways (reasons might be high capital, excessive waste treatment cost, or scheduling problems). Furthermore, selected optima can drive further development or pilot-plant experimentation to determine real-time data and agreement to the model. Any promising model can be verified by design rather than by search. Concurrent engineering and integrated process and product development are more than concepts and become readily feasible, since the software models can be shared among different groups. A few firms have software usable in biotech design; some have the capacity to connect design information to capital estimation software. Some firms can do both under one roof. Companies that have one or more capabilities in these areas include the following:

Intelligen, Inc.
Scotch Plains, N.J.
908-654-0088

BioPro designer (plus superset and other subsets for special cases—also see Ref. 66)

- Process simulation with economic evaluation
- Models more than 70 unit operations with database of process equipment
- Waste stream characterization
- Process scheduling
- Rapid economic evaluation of alternate manufacturing processes
- Sensitivity analysis and scenario evaluation for process optimization
- Example given: β-galactosidase production by *Escherichia coli*
- Profitability analysis
- demo discs available

Icarus Corp.
Rockville, MD
301-881-9350

Process evaluator

- Evaluates process economics and design alternatives
- Capital and operating costs, schedules, investment analyses
- Rapid estimates of feasibility and profitability of alternatives
- CPM-based planning schedules (general overview, engineering, construction)
- Works from process simulator software (e.g., Aspen Plus, Chemcad III, Design II)
- Oil, chemical, petrochemical, specialty chemical, pulp and paper, and mining industries
- Expert system knowledge bases (does not use indices)

Aspen Technology
Cambridge, MA
617-577-0100

AspenPlus

- Complete library of physical properties and unit operation models
- Demo disk available
- Creation of process flowsheet and simulation model—can create PFD-style drawings
- Sensitivity analyses, optimization, and data fit (owner's or databases)
- Costing and economic analyses
- Runs on networks or stand-alone hardware (workstations and PCs)
- BioProcess Simulator (BPS) module simulates performance of bioprocesses, such as enzymes, vitamins, and antibiotics (equipment includes batch and/or continuous fermentors, sterilizer, membrane and/or depth filtration, cell disruption, disc centrifuge, ion exchange, chromatography)

Batch Process Technology
West Lafayette, IN
317-463-6473

- Simulation system to manage multiproduct batch and semicontinuous processes in biochemical, food, pharmaceutical, and specialty chemical industries
- Optimize number and size of equipment, retrofit, debottleneck, and schedule processes
- Plans for increased throughput by evaluating process capacities
- Inputs: equipment network, recipe network, and sequencing information

Table 4. Construction and Equipment Costs: Ranges Based on Mid-1996 Pricing for the Northeastern United States

Item number	Description	Size or unit of measure	Unit cost range	Comments
		Selected construction tasks		
1.	Slab demolition and removal	6 in. thick	$8–12/ft^2	
2.	Underslab piping			
	Cast iron	4 in. diameter	$26–32/linear ft	
	Polypro (single wall)	4 in. diameter	$38–55/linear ft	
	Polypro (double wall)	6 in./4 in. diameter	$75–90/linear ft	
	Stainless steel	4 in. diameter	$60–65/linear ft	Type 304L
3.	Concrete slab			
	>1,000 ft^2	6 in. thick	$3–5/ft^2	Dowel into adjacent slab
4.	Concrete slab			
	Trench infills	6 in. thick	$6–8/ft^2	Dowel into adjacent slab
5.	Masonry—interior	8 in. thick	$5–7/ft^2	
6.	Masonry—exterior	12 in. thick	$7–9/ft^2	
7.	Carpentry			
	Hollow metal door, frame and hardware	Each	$600–800	3 ft × 7 ft nominal
8.	Exterior windows			
	Strip	Square foot	$20–28	
9.	Roof, 4 ft high			
	Membrane	45 mil.	$3.50–5.50/ft^2	Approx. 1,000 ft^2
	Built up	3 ply	$3.50–5.50/ft^2	Approx. 1,000 ft^2
10.	Insulation			
	Building—walls (6 in. thick)	Square foot	$0.50–0.75	
	Piping insulation (6 in. diameter)	Linear foot	$8–12	
11.	Drywall partition			
	10 ft high	Linear foot	$32–36	
12.	Acoustic ceiling	Square foot	$1.75–2.00	2 × 4 lay-in
13.	Painting			
	Standard latex	3 coats	$0.35–0.55/ft^2	
	Exterior	3 coats	$0.55–0.75/ft^2	
	Epoxy	2 coats	$0.75–0.95/ft^2	
14.	Flooring			
	Vinyl composition tile	Square foot	$1.25–1.50/ft^2	
	Wood	Square foot	$7.50–9.50/ft^2	
	Seamless epoxy	Square foot	$7–10/ft^2	Includes integral coves
	Carpet	Square yard	$20–28	
	Ceramic tile	Square foot	$6–8	
	Seamless vinyl	Square foot	$5–7	Includes integral coves
15.	Laboratory casework	Linear foot	$250–350/linear foot	Includes “Chemsurf” countertop and approx. 40% base cabinet
16.	Process piping			
	Cast iron	2 1/2 in.	$28–36/linear ft	
	Carbon steel	6 in.	$32–42/linear ft	
	Stainless steel	1 in.	$90–225/linear ft	316L, includes valves and fittings

Selected systems (includes installation)

1.	Clean rooms (packaged rooms, ≤1,500 ft^2)			
	Class 100,000	Square foot	$150–175	Sidewall return
	Class 10,000	Square foot	$165–200	Sidewall return
	Class 1,000	Square foot	$185–225	Sidewall return
	Class 100	Square foot	$215–250	Sidewall return
2.	Alarm system	Per point	$750–1,000	
3.	Building demolition (interior)	Cubic foot	$0.25–0.75	
	Building demolition (including structure)	Cubic foot	$0.50–1.00	
4.	Fire protection	Square foot	$1.25–2.50	Water system
5.	Electrical system	Square foot	$10–15	Base building/no process

Selected equipment and utilities (includes placement and installation)

1.	Fume hoods 4 ft	Each	$6,000–7,000	Stainless steel ducted w/SS work station
2.	Culture vessels			Agitated, 30 psig, full vacuum
	100 L	Each	37,000–40,000	
	1,000 L	Each	$55,000–60,000	
	10,000 L	Each	$530,000–600,000	
3.	Air compressor	Lump sum	$78,000–85,000	600 scfm
4.	Chiller	Ton	$210–260	
5.	Freezer room	Square foot	$160–200	Excluding furnishings
6.	Emergency generator			
	100 kW	Lump sum	$30,000–32,000	Diesel, outside
	300 kW	Lump sum	$41,000–44,000	Diesel, outside
	500 kW	Lump sum	$70,000–75,000	Diesel, outside
7.	Tankage			
	Carbon steel	Gallon	$5.00–7.00	<1,000 gal
	Stainless steel	Gallon	$8.50–10.00	<1,000 gal
8.	Environmental rooms (including cold rooms)	Square foot	$120–160	Temperature controlled, excluding furnishings
9.	Steam boilers			150 psig design
	100 hp	Lump sum	$38,000–46,000	
	250 hp	Lump sum	$70,000–80,000	
	300 hp	Lump sum	$80,000–90,000	
10.	Heat exchangers			
	Carbon steel	gal/min	$100–150	
	Stainless	gal/min	$300–350	
11.	Vacuum pumps	Scfm	$900–1,100	Liquid ring
12.	RO/DI water,			
	Package system	gal/min	$1,200–1,500	Generation equipment only
13.	WFI system, package system	L/h	$110–120	Generation equipment only

Source: Data supplied by Marshall Contractors, Inc., Rumford, R.I.

- Outputs: mass balance, equipment and resource utilization, cycle time, and waiting time
- Examples: expansion of an agricultural products plant, capacity expansion of a pharmaceutical process, and optimize operating conditions for an esterification

Richardson Engineering Services
Mesa, AZ
602-497-2062

Rapid Access Cost Estimating (RACE) Spreadsheet System

- Material and manpower estimation for general and process plant construction industries
- Extensive databases are updated annually (>100,000 line items)
- Estimate new construction costs and maintenance project costs
- Create unit cost assemblies rapidly
- Multiple database access (5) while in use
- Can operate with more than one currency at a time, also English and metric interchange
- User-defined estimate spreadsheet
- Import and export: ASCII and dBASE II

Because biotech process plants are expensive and most (or all) biotech companies understand the need for a multicandidate pipeline, it would be useful to incorporate flexibility in plant design. Young presents an approach for selection of the right capacity for multiproduct plants (67). Given the unknowns of product mix, sales forecasts, and manufacturing inefficiencies, it is still better to apply some logic to the process. Young gives an example of three novel products (although greater numbers can be introduced), with the expectation that one of them will move into manufacturing; however, there is a finite chance that all three will succeed. Risks are quantified and a decision tree for probable events is created. Both dedicated and campaigned (mixed operation) programs are established. The solution gives a best estimate for long-lead items, construction cost, time of construction, and start-up. Before moving to specific examples of economic analysis, an interesting phenomenon should be noted (Fig. 2). For many diverse input variables that have a cost component, there is an economic minimum somewhere in the study region. Each set of variables requires its own specific calculations (based on theoretical formulations or empirical data); the point here is that left to chance, one will not be operating at the minimum cost point in all likelihood. One example will be given, and Table 5 lists other possible variables to study.

Suppose one is designing a plant. Let the quality of the facility be the abscissa (moving from poor at the left to perfection at the right). The left ordinate would be recurring cost. This value would be highest at the poorest design and lowest closest to the perfect design. Then label the right ordinate "initial cost": The lowest value would correspond to the poorest design, and cost would increase as one move from good to perfection in quality of design.

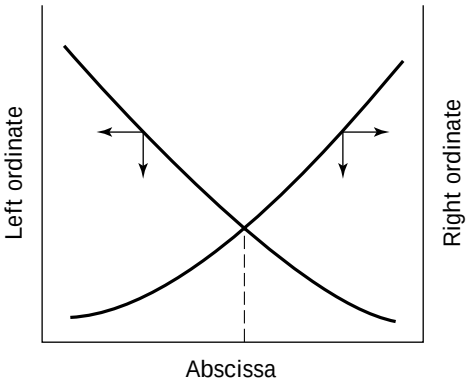


Figure 2. Determination of economic minimum.

Somewhere in the region (near or at the intersection of the curves), there is an overall cost minimum. There may be reasons for not selecting this point, but costs still must be reviewed as a function of one or more input variables. Other typical inputs are given in Table 5.

Specific cases involving economic analysis and costing can be found in the literature. The contributions of Datar are noteworthy (68,69). The earlier paper (1986) presents economics of enzyme recovery from an *E. coli*–based fermentation and extends the treatment to recombinant products. There is a breakdown of production costs for each step, capital and operating costs are developed, approaches are given for reducing manufacturing costs, and there is a clear delineation of assumptions and required off-sites. The second paper (1993) focuses on tissue plasminogen activator (tPA). Here, different processes are compared and detailed economic analyses are performed for alternative processes. Both papers are highly recommended.

A useful review exists on processing to prepare recombinant human insulin (70). The entire process is detailed, but emphasis is placed on large-scale purification. The high cost of purification is noted, but no dollar value or economic analysis is given. Another paper gives process analysis and economic evaluation for production of biosynthetic human insulin (66). Here, *E. coli* is used for expression of a precursor of proinsulin (intracellular) as a Trp-E fusion protein. The process results in 1,500 kg/yr of purified biosynthetic human insulin. Equipment is thoroughly detailed and specifications are shown. The fixed capital estimate (including contingency) comes to $142 million, using 1994 costs. Annual operating cost (including depreciation) is $56 million. Selling price is set at $110/g; a cash flow analysis is given and payback time is less than 2 years. BioPro Designer (a product of Intelligen, see preceding listing) was employed.

Economic analyses and publications continue to be generated on ethanol production from various substrates (71,72). The interesting historical phenomenon for ethanol biosynthesis (for fuel or an intermediate, not as food) is that just a little more R&D will be needed to make the specific (or any) process economically viable. This has been true for a number of years. The ethanol references noted are useful for methodology employed and probably will not result in construction of a facility in the short term.

Table 5. Determination of Diverse Cost Minima

Left ordinate	Abscissa	Right ordinate
Cost of mechanical failure	Frequency of preventive maintenance	Cost of preventive maintenance
Cost of imperfect product	Number of tests or inspections	Cost of inspection (testing)
Cost of production	Number or extent of capital items	Cost of capital
Cost of accidents	Number of safety inspections	Cost of safety programs
Cost of contamination	Number of cleaning cycles	Cost of cleaning
Cost of estimated investment	Number of planning iterations (cases)	Cost of planning or engineering
Cost of aeration	Cost of power	Cost of agitation
Cost of unplanned shutdown	Number of devices	Cost of backup or redundant equipment
Cost of energy	Extent (size) of heat exchangers or insulation	Cost of heat exchanges or insulation
Cost of product failure	Number of tests	Cost of testing (QC)
Project cost	Number of iterations	Cost of planning/engineering
Recurring cost	Quality level (poor to perfect)	Initial capital cost

Any review of economics should include some of the recent contributions to selection of the best process using payout parameters. Ward (73) stresses discounted rate of return while including consideration of investment flows over several years. Woinsky (74) presents a few different measures and seems to prefer payout time. Derivations are given and various plots are presented showing payout time as a function of pretax internal rate of return, plant economic life, and the engineering–procurement–construction period. Other methodologies are given in the reviews and books noted earlier in this section.

Another early stage technological development that may have the greatest economic impact on the biotechnology industry includes the areas of design, cost estimation, scheduling, and process control. Rapid change in computer control and especially in knowledge-based systems (KBS) offers the promise of significant cost savings by means of adaptive or predictive control of the process. These new control methods are especially useful when reaction–process mechanisms are understood poorly or not at all. The KBS spectrum includes expert systems, fuzzy logic, and neural networks.

An introductory book (75) is useful for its glossary of terms and for giving early applications in industry. Each of the separate concepts has strengths and weaknesses, and these are discussed. Software programs already exist for each of the classifications, and newer concepts are being presented with increasing frequency. For reasons of space, only a few references are given here; these works can be consulted for more publications and background. The first thing needed in consideration of computer systems is a careful definition of what the user wants. Furthermore, for biotech systems, there is the requirement for validation. The validation program should be considered in parallel with computer (and software) selection. There is, not surprisingly, a semantics and communication problem between hardware and software companies and users.

Chapman (76) gives very useful guidelines in discussing user requirements and gives a specific example for a multiproduct bulk pharmaceutical chemical plant. Documents needed for system validation are included. A paper on the use of neural networks in fermentation monitoring and control discusses optimization of inoculum development in a real case (77). Neural networks can model highly nonlinear, multivariable processes. Artificial intelligence (covering the three general classes of KBS noted above) is explained by other authors (78). Stress is placed on process control and support systems for operators. Quality improvement is reviewed. Examples are given in diagnosis and repair and alarm analysis (predictive control). One or more subgroups of KBS seem to be well suited to bioprocessing, since the programs can be used in production, staff and batch scheduling, engineering design, and project management. A Chiron computer control system (for a 1,500-L culture vessel producing a recombinant product) is discussed and explained (79). Design, building, validation, and start-up are reviewed. A supervisory control and data acquisition system above a control PC (with appropriate backup) was installed in 1992 for about $400,000.

Guidelines for computer control systems have been published (see Table 6). A large system might be found in a production facility or a pilot plant with multiple culture vessels or many purification trains and packaging capabilities. Medium scale would probably cover most pilot plants. If one is fortunate enough to have justification for a distributed-control system, provision should be made for incorporation of both the pilot plant and production equipment in a single system. Data collection and comparison are simplified, and scale-up is performed much more efficiently.

Another relevant point in this part of the review concerns contract manufacture. Just as it is sometimes efficient and prudent to use a contract research organization in clinical trials, it may be more effective to use a contract manufacturer (CMO) when product is needed. A useful starting point in this determination is an article on points to consider and methods to employ in evaluating a CMO (80). There are 24 references as well. The information relates to initial interviews, meetings, and inspections, and many issues to be incorporated into a contractual agreement are covered. In another article, describing an actual contact manufacturer in Puerto Rico, the author relates attributes to be sought in a contact partner (81). Seaver gives considerable insight into working with contract manufacturers (82). His suggestion is that a company's first protein-based product be made by a CMO. The estimate given for producing material for phase I and II trials in-

Table 6. Guidelines for Computer Control Systems

Process size	Input/output points	Number of units		Control option
		Continuous	Batch	
Large	>500	>3	>10	Distributed-control system
Medium	100–500	3	2–10	Programmable logic controller
Small	≤100	1	1–2	Personal computer, front-end system

Source: Excerpted by special permission from *Chemical Engineering*, November 1993. Copyright © 1993, by McGraw-Hill, Inc., New York.

house is 12–24 months in time and $0.5–1.5 million in cost (this latter range may be low).

Use of a CMO is not simple and requires numerous joint meetings and much effort on everyone's part to develop an agreement. More joint effort is expended during the production phase. A good suggestion is to plan on making 5–10 times more product than is needed for all the initial trials. Also, perform a test (or scale-up) run prior to the production run and save a considerable amount of product at multiple locations to serve as a long-term reference standard.

Other points related to manufacturing should be emphasized. Complete characterization of a master cell bank or a master working cell bank takes 4–6 months and costs "at least $50 thousand" (82). With newer pathogens appearing and the desirability of showing viral absence or inactivation by ever more complicated means, complete characterization might cost 1.5–2 times the figure given. It should be remembered that pooling of sublots (some of which have failed testing) to achieve an average "pass" is frowned on by regulatory bodies and is undesirable. In Table 7 the author attempts to list all issues and concerns that should be reviewed, considered, and resolved (in writing) in developing a contract with a CMO. It goes without saying that legal assistance is required in drafting or reviewing documents either during the negotiation of all conditions and terms or afterward.

VALIDATION AND REGULATORY ISSUES

Validation and regulatory issues are understood by almost all persons intimately involved in biotechnology; however, the time and cost impacts of these essential functions are woefully underestimated. The rapid growth of firms (or divisions within engineering/contractor firms) in the area of validation is not an anomaly; it is the result of ever-more intensive and in-depth requirements by regulatory agencies. Even though many large companies have in-house validation groups, variations in need (created by a new facility, a major retrofit, or a new process introduction) create a demand for outside services and consultants. Smaller companies rarely have the money to maintain a fully functional validation group that could satisfy even a majority of the company's needs. If one considers a validation technician or writer at a cost of $40–60/h, an experienced validation supervisor or technical expert at $70–120/h, a senior expert at $150–300/h, and a project that could consume several thousand man-hours for merely writing the master plan and protocols, it is easy to see that merely having an acceptable set of validation plans could cost many hundreds of thousands of dollars.

With more complicated control schemes (including programmable logic controllers or PLCs) requiring special (and often proprietary) software, there is a greater need for specialists to write the protocols and oversee their execution. Validation is the process by which documented evidence is accumulated to provide a high degree of assurance that a specific process (or part of a process) will consistently produce a product (or an intermediate) meeting its predetermined specifications and quality characteristics. Therefore, it is no wonder that this important feature comes under such careful and extensive regulatory review.

Validation cannot begin when a plant is up and running. Commissioning of a plant should include extensive written documentation, not the least part of which must include installation qualification (IQ) and operational qualification (OQ). Finally, as part of the commissioning, or hopefully soon thereafter, comes performance qualification (PQ). PQ provides the rigorous and documented testing that demonstrates the effectiveness, reproducibility, and allowable performance ranges of the process. Goswami and Lorenz discuss facility commissioning (83). A main point is that major savings are realizable if commissioning protocols are properly integrated into the overall validation plan. During design and construction, terse narrative system descriptions and single line drawings or indicators should be prepared to show cGMP control and compliance. These will prove invaluable during the preapproval inspection by the FDA.

A series of articles has been collected in a single bound issue; all articles refer to validation issues (84). This compendium will prove very useful to anyone new to the subject and also to anyone heavily involved in an ongoing validation project. Although no formula exists for cost determination, these articles can serve as background and as the framework for preparing an estimate. Furthermore, a number of firms offer to draft time and cost estimates for preparing protocols, for assisting in execution, or for completion of the entire validation process. The latter obviously would be the most costly, but there is another disadvantage. Regulatory bodies prefer some (but appreciable) direct involvement of the owner in the validation process, as opposed to providing only money and ultimate oversight.

Another area of ongoing cost (also often seriously underestimated) is that for environmental control. Even if the production process were not aseptic, many in-process steps are now performed in controlled environments. In many if not most cases, biologics, vaccines, and parenterals

Table 7. Contract Manufacturing Issues and Concerns: Audit and Contract Preparation

1. *Specifications*		
Raw materials	Cleaning	Formulation/packaging
Specialized technology	Documentation	Validation
Cell strains	Product specifications	Certificate of analysis
Testing	Safety requirements	Personnel protection
Hazard analysis	Sample retention	Labels
2. *Quantities and timing*		
Forecasting		Current planning program
Current capacity utilization		Geographic location (movement/shipping)
3. *Personnel*		
Qualifications		Training programs
Degree of automation		Training documentation
4. *Inspection* (by manufacturing, QC, QA groups of sponsor)		
Contractor's validation master plan		FDA inspection reports and status
Materials of construction		cGMP inspection (by sponsor)
5. *Regulatory review*		
Outside audits		Inside audits
Accreditation		Consultants used (plus qualifications)
6. *Experience*		
Need for specialized technology		Segregation of processes/products
History (experience in similar area)		Management capabilities
7. *Data analysis*		
Computer systems and software		Backup systems
LIMS and PLCs		Extent of (plans for) process development
8. *Sponsor supervision or participation*		
Emergencies		Timing of coverage (and access)
Planned/unplanned checks		Reports (How often? How many?)
9. *Scale-up schedule*		
Technology transfer		Additional capital needs (also ownership)
Ownership of process improvements		Program (plus explanation of deviations)
10. *Patent rights*		
11. *Payment schedule*		
Additional efforts		Criteria for payment and terms
Transfer to third party (if needed)		Change orders/changeover procedures
Termination clauses		Approvals needed

call for sterile packaging or bottling. The owner is responsible not only for design, control, and maintenance of environmental quality; the processor is responsible for evaluation of environmental quality. Alert and action levels for each controlled environment must be established; monitoring devices must be checked and qualified. It is not enough to monitor and control particulates, temperature, and humidity; additional monitoring must be in place for airborne bioburden, surface bioburden, supply and exhaust air volumes, pressure differentials between environments, and (for laminar flow only) airflow patterns and airflow uniformity. These special facilities are costly to install, costly to maintain, and costly to monitor. Cleaning methods must be verified and recertification periods must be established and followed. Generally, systems and components of the environmental control system (in aseptic processing) must be validated and requalified at these times:

Every time there is a significant system or control component repair or change

Before commissioning

Every time there is a process change or a new process is introduced

Every time out-of-control limits are noted in the operational parameters

Tedesco and Titus discuss other hidden costs that should be considered in planning, building, and operating biopharmaceutical facilities (85). They include engineering and environmental assessment fees and the cost of any required cleanup, permits, and change orders. The same article also presents generalized estimates that are useful for very early estimating:

Turnkey cost of GMP biopharmaceutical facility	\$800–1,200/ft^2
Turnkey cost for pilot facilities	\$400–600/ft^2

Another area of insufficient planning input is the warehouse. Appropriate inventory space (under appropriate conditions) is needed, along with quarantine and release areas. One should add space for reject materials and the various kinds of waste streams generated (chemical, biological, radioactive). Ideally, there should be no chance for mix-up or cross-contamination; proper plant layout and material flows should be carefully planned and provided.

As noted earlier, the physical plant itself, even if superbly designed and built, makes up only one part of the essential preapproval inspection by the FDA. Table 8 lists

Table 8. Preapproval Inspection Checklist

Organizational chart(s)	Integrity policy
cGMP audits	Integrity audits
Operating procedures (SOPs)	Validation documents (plus master plan)
Material specifications	Product/process flowcharts
Facility diagrams (esp. HVAC) and site plan	Computer hardware and software validation (PLCs)
Cleaning and disinfection	Product stability
Utilities	Release criteria
Employees (résumés, training)	QC review
Change control	Drug development information
Equipment list	In-process controls
Failure investigations	Labels and labeling
R&D laboratories	Internal audit summaries
Batch record(s) review	Development reports
Clinical lots	Scale-up (pilot vs. full-scale)
Documentation	Observed practices versus SOPs
Trend(s) analysis	Vendor qualification and analysis
Storage	Sampling plan(s)
Maintenance logs	Technology transfer
Review: All submissions relating to this inspection	
For an existing firm: FDA-483, seizures, warning letters and responses, recalls, complaints, NDA deficiencies	

a host of factors to be planned, organized, and developed prior to the inspection. A large amount of documentation (and other paperwork, in general) is required. Remember that there is a cost associated with each factor shown. The earlier the tasks and costs are recognized, planned, and budgeted, the greater the potential for a successful inspection accomplished in minimum time, at minimal cost, and with minimal aggravation.

REGULATORY ISSUES

Persons involved in biotechnology (in whatever discipline) know more or less about regulatory issues, but few come into direct contact with regulatory personnel. Therefore, those with a more academic or somewhat displaced knowledge of regulatory matters cannot appreciate the enormous amounts of time and effort that must be expended in the following:

Planning a biotech facility with compliance always an issue

Building such a facility (including records, pictures, and videotapes to show compliance)

Performing animal experiments within or outside this facility

Negotiating the many hurdles in planning a drug test

Achieving facility clearance by all appropriate regulatory bodies

Receiving product clearance (ability to sell)

Postapproval monitoring for minor and major side effects

The amount of money (and energy) needed to achieve regulatory compliance is invariably underestimated. No attempt will be made to cover all programs and costs related to regulatory clearances, but some measure of the complexity is suggested. Although everyone in a commercial biotech setting knows (or has heard about) the EPA, the FDA, and the USDA, other federal agencies have either an interest or a jurisdiction in biotechnology matters. From the simplest issues of syringe storage and the control of undenatured ethanol to animal housing and animal studies to transport of biomaterials, there are federal agencies involved (Bureau of Alcohol, Tobacco, and Firearms [BATF], National Institutes of Health, Centers for Disease Control, Department of Transportation, Occupational Safety and Health Administration, and Department of Labor). One should not forget that each state has its own regulatory bodies.

Consider the reviews needed for plant production: local, state, and federal agencies are involved. Not all the following issues apply in every case, but many will need review and approval: coastal zones, wetlands, tidelands, waterways, and historic preservation. Further, approvals will be needed for or from planning and zoning boards, boards of health, regional planning boards, discharge permitting, hazardous waste, infectious or biowaste, use and disposal of radioactive materials, transport of one or more of the foregoing, and laboratory licensing. Essentially, there is broad regulatory oversight in these areas:

- Licensing of *all* biotechnology products (biologics, drugs, and devices)
- Use of recombinant DNA and cell hybridization techniques
- Any alterations of the human gene (and related human testing and therapy)
- Planned releases of *any* genetically engineered plant or animal
- Worker protection

The cost of regulation, in its broadest sense, depends on the location of the laboratory and production plant, the process, the product and its intended use, and postmarket surveillance. Therefore, it is not possible to arrive at a total cost by formulaic means. At the lowest end (production of devices that have been in use for many decades), a minor change might involve regulatory costs of tens of thousands of dollars. For a new, simple, and noninvasive device, costs might be estimated at from a few thousand to many hundreds of thousands of dollars. Some relatively simple devices cannot be marketed because the cost of required double-blind studies to prove efficacy (there is no safety issue) is well beyond the capacity of the developer or exceeds any foreseeable financial return. For novel biologics or pharmaceuticals, regulatory costs could easily start in the tens of millions of dollars and could readily rise to the hundreds of millions of dollars, depending on test group size and overall clinical test requirements. Postmarket surveillance must be added to the total; note that this phase of the overall study is becoming ever more complex, lengthy, and costly.

Although change in regulations is common, some relatively recent reviews are useful for the state of the art at least to that point in time. An encyclopedia published in

1996 contains an enormous amount of information on the entire spectrum of pharmaceutical technology (86); however, two sections in Volume 13 are directly applicable to validation and regulatory issues. Even though the emphasis is on drugs, the tendency in U.S. regulatory affairs has been to blur the distinctions between devices and drugs. Therefore, while not every regulation covers both areas, the articles mentioned can serve as guides for the manufacture of medical products as well as for pharmaceutical development and production. The materials on registration of drugs and quality assurance as applied to drug products are very good starting points in their respective areas.

PAYMENT AND COMPENSATION

Although every phase in the foregoing discussion has elements of a more or less quantitative nature, and reasonable cost and time estimates can be made for each stage in biotech development and processing, the area of return on investment or payment for a product or service presents problems in economic analysis. It would appear that the bench scientist or the pilot-plant engineer (just to name two of many individuals involved) should have little to no interest in rate of return or, indeed, in how this rate is to be established. If that were ever true, it is no longer. From selection of a physiologic target to process design to scale-up and to all elements of operations, there is a continuing need for feedback on what range of compensation is attainable, assuming project success. Day-to-day decisions cannot be based on compensation issues; however, periodic review is essential to compare cost of good estimates with payment potential.

Given the global costs of health care and the ever-increasing involvement of third-party payers as well as government bodies, the situation is becoming critical. To assess the dollar magnitude, total 1997 sales for only five major drug firms (Merck, Johnson and Johnson, Novartis, Bristol-Myers Squibb, and Rhône-Poulenc) approaches $100 billion. Each of these firms achieved a double-digit percentage increase in sales compared to the prior year (5). Thayer also details revenues for 20 of the new major biotech firms for 1997; the total is $6.6 billion (5). With the advent of high-volume screening, genetic analysis, and the exponential growth of life-enhancement drugs, one can expect continued (and rapid) growth in sales of bioproducts and medical devices.

The complex issues of payment for biotech products or services have already elicited volumes giving fact and opinion. There is no attempt here to cover the topic completely, but some measure of the problem is needed by all workers in the field, not only those who attempt to set pricing. Fortunately, some references give both a reasoned introduction to the interplay of factors and groups (very often in conflict) and more detailed exposition of medical cost problems and potential solutions. At the least, ideas for alleviation are given.

A short article by K. Keller discusses cost of care with special attention to ethical choices sometimes overlooked (87). Keller recognizes that while government is a major factor in forcing changes in health care, government's views and approaches are hardly the same as those of either health care givers and suppliers or those who need the products and processes. Government's attention is primarily regulatory: The focus is on the process (rather than the outcome), and in general government bodies are concerned with short time scales relative to biotechnology development. Keller indicates the strong need to move beyond the "technology push" mentality; cost has now become as important as performance. Hence the creation of novel measurements (e.g., quality-adjusted life years). The point is made that all persons involved in the development of the technology must have some grasp of the complex interactions coming to the fore; these knowledgeable individuals should attempt to work constructively within the evolving framework to assist members of society in dealing with the interplay of factors and associated dislocations.

Another question has arisen that was either an insignificant problem or a nonexistent one a decade ago: Who pays for clinical trials, and just how are they to be funded? (88). Managed care organizations (MCOs) are claiming that clinical trial expenses are not part of their responsibilities. Therefore, coverage is often denied, hospital stays are shortened, and additional diagnostic tests (very often essential in the course of a trial) are denied. MCOs are taking the position that therapy in phase I, II, and III trials is unproven and that others (preferably the sponsor) should assume all associated (and additional) costs incurred during these phases. Some MCOs relent on so-called phase IV studies, which are outcome studies involving cost-effectiveness and quality of life. Not unexpectedly, elected officials (responding to biotech product suppliers, patients, or both) are becoming heavily involved with legislative ideas, proposals, and even threats. The outcome is far from clear; the problem is relatively new, and whatever the solutions, biotechnology companies will be affected.

Two separate volumes should be consulted for an overview of the cost situation in health care as well as various analyses and proposals to avoid further difficulties. Each volume is a compilation of presentations; the same sponsor arranged the meetings—one in April 1994 and the other in May 1995. The contents of the first set of proceedings (89) can be gauged by a selection of titles (the overall meeting topic is "The Role of Technology in the Cost of Healthcare"):

"Technology: Quality at What Cost?"

"How Do We Evaluate the Cost of Health Care Technology?"

"Crisis of the Uncontrolled and Misunderstood Numbers of Health Care Costs"

"Outcomes Research: Economics and Quality of Care"

The many articles in the collection cover the evolving situation, ideas on resource allocation, and the ever-increasing cost of health care. Of special interest is the review of outcomes analysis; no longer are morbidity and mortality the only issues to study and review. One must add patterns of recurrence, rehospitalization, relapse rate,

procedural results, and patient satisfaction. From a number of presenters comes the unsurprising finding that the perspectives of those making cost–benefit analyses have a large impact on the findings and results. There are other articles on trial design, ethics, and current and proposed means of analysis. The second volume continues the discussion with the addition to the title of the phrase "Providing the Solutions" (90). Needless to say, solutions were offered but the problems persist. Of special note is a review by A.M. Frendrick entitled "Outcomes Research in Evaluating the Benefit of Health Care" (pp. 68–73). There is a large section with a number of papers on the impact of regulation on health care costs. Near the end of the proceedings, there were a few papers on a subject that is rather emotional. One had the provocative title "Is There a Need for Health Care Rationing?" The conclusions were rather mild, but at least the organizers had been willing to broach the subject. The authors conclude that rationing has always existed and will continue to exist. It is our job to establish appropriate limits, inasmuch as there is a definite need for setting such bounds.

A review article on disease management has been written by Merck–Medco personnel (91), who define *disease management* as "a systematic population-based approach to identifying those at risk, intervening using information from the growing field of evidence-based medicine, and measuring patient outcomes once an intervention is in effect." Economic outcomes weigh heavily in the perspective of almost every group involved in disease management. (It is odd that the only outcome listed for the perspective of the physician is clinical outcomes. This is at best idealistic and at worst naive.) The perspectives of each party to the process are discussed and analyzed. Not surprisingly, not all outlooks are congruent. Components (guidelines, patient identification, intervention methods, and outcomes assessment) are reviewed and various possible alternatives given. Disease management has as its primary goal population-wide improvements in health; an unstated subordinate goal is improved productivity—and possibly cost control—in delivery of health care.

These various issues are raised here because solutions—whether they come from HMOs, MCOs, the government, hospital conglomerates, or a combination of these bodies—will have an enormous impact on the other topics discussed earlier. Any significant change in compensation (which relates to rate of return) must have an impact on discovery, development, and plant investment. This, in turn, means that either the quality or the quantity (or likely both) of biotechnology research and product and service delivery must change. At this writing, it appears that limits may be placed on compensation. At the least, the rate of growth of health care costs will be reduced. It does not take a leap of intuition to project the impact on the biotechnology product cycle. The concluding message is that biotechnology industry participants must become involved in resource allocation in every stage from discovery to payment. The involvement and input should be reasoned; if background and understanding of the full spectrum of economic needs and impacts are clean and logical, the resulting positions will be heard.

BIBLIOGRAPHY

1. Biotechnology Industry Organization, *Forbes*, May 4, 1998, pp. 97–120.
2. J. Hodgson and K. Barlow, *CHEMTECH* **23**, 45–47 (1993).
3. Anon., *Genet. Eng. News* **16**, 3 (1996).
4. Biotechnology Medicines in Development, 1996 Survey, Pharmaceutical Research and Manufacturers of America, Washington, DC.
5. A.M. Thayer, *Chem. Eng. News* **76**, 25–37 (1998).
6. B. Werth, *The Billion-Dollar Molecule*, Simon & Schuster, New York, 1994.
7. V. Glaser, *Genet. Eng. News* **12**, 33 (1996).
8. The biotech century, *Business Week*, March 10, 1997, p. 84.
9. D. Matheson, J.E. Matheson, and M.M. Menke, *Res. Technol. Manage.* **37**, 21–24 (1994).
10. R. Szakonyi, *Res. Technol. Manage*, **33**, 31–36 (1990); **33**, 41–46 (1990).
11. I. Norrington, *Pharm. Manuf. Int.* 37–40 (1996).
12. P. Bevan, H. Ryder, and I. Shaw, *Trends Biotechnol.* **13**, 115–121 (1995).
13. P. Sharpe and T. Keelin, *Harvard Business Rev.* **76**, 45–57 (1998).
14. K.S. Brown, *Scientist*, **10**, 1, 8–9 (1996).
15. N. Pfeiffer, *Genet. Eng. News* **16**, 1, 18, 19 (1996).
16. G. Samdani and G. Ondrey, *Chem. Eng.* **100**, 35–39 (1993).
17. O. Schuster and G. Menke, *Pharm. Manuf. Int.* 85–86 (1996).
18. C. Turner, M. Turner, and K. O'Toole, *Chem. Eng.* May 16, 1996, pp. 25–27.
19. F.A. Dubinskas, *Technol. Rev.* **74**, 24–30, (1985).
20. G.P. Pisano and S.C. Wheelright, *Genet. Eng. News* **16**, 4, 32, 33 (1996).
21. R. Frumerman, *Res. Technol. Manage.* **33**, 10–11 (1990).
22. D. Mukesh, *Chem. Eng. Prog.* **92**, 110–112 (1996).
23. C. Prior, P. Bay, B. Ebert, R. Gore, J. Holt, T. Irish, F. Jensen, C. Leone, J. Mitschelen, M. Stiglitz, C. Tarr, R. Trauger, D. Weber, and M. Hrinda, *Pharm. Technol.* **19**, 30–52 (1995).
24. B. Sonnleitner, *Adv. Biochem. Engineer. Biotechnol.* **54**, 155–188 (1996).
25. J. Jones, W. Asher, J. Bomben, D. Bomberger, C. Marynowski, R. Murray, R. Phillips, D. Roberts, K. Semrau, R. Swidler, C. Witham, and J. Zuegel, *Chem. Eng.* **100**, 98–106 (1993).
26. B.H. Junker, J. Lynch, J. Leporati, J. Schmitt, J. Gieger, T. Garah, M. Stober, and P. Salmon, *Bioprocess Eng.* **13**, 279–287 (1995).
27. J. Abate, A. Chen, W.J. Dale, and D. McNeil, *Chem. Eng. Prog.* **91**, 82–87 (1995).
28. C.M. McFarlane and A.W. Nienow, *Biotechnol. Prog.* **11**, 601–607 (1995).
29. R.F. Dream, *Pharm. Eng.* **13**, 56–66 (1993).
30. R.G. Werner, F. Walz, W. Noe, and A. Konrad, *J. Bacteriol.* **22**, 51–68 (1992).
31. R.L. Bruno and P.E. Scott, *Pharm. Eng.* **13**, 42–46 (1993).
32. T.A. Bibila and D.K. Robinson, *Biotechnol. Prog.* **11**, 1–13 (1995).
33. F.W.J.M.M. Hoeks, J. Muhle, L. Bohlen, and I. Psenicka, *Chem. Eng. J.* **61**, 53–61 (1996).
34. A.P.J. Middelberg, *Chem. Eng. J.* **61**, 41–52 (1996).
35. T.R. Elliott and W.L. Luyben, *Ind. Eng. Chem. Res.* **34**, 3907–3915 (1995).

36. S. Harsa, C.A. Zaror, and D.L. Pyle, *Proc. Biochem.* **28**, 187–195 (1993).
37. M.F. Kozempel, M.J. Kurantz, and J.C. Craig, Jr., *Biotechnol. Prog.* **11**, 592–595 (1995).
38. S. Moore, G. Parkinson, and G. Ondrey, *Chem. Eng.* **102**, 33–37 (1995).
39. M.A. del Valle, *Pharm. Eng.* **15**, 14–22 (1995).
40. S.E. Mackler, *Genet. Eng. News* **16**, 3 (1996).
41. P.J. Haas, *Pharm. Technol.* **19**, 26–38 (1995).
42. W.R. Brader, R.T. Hsu, and P.E. Lorenz, *Pharm. Eng.* **15**, 30–42 (1995).
43. S.W. Fitzpatrick and S.E. Mackler, *Pharm. Technol.* **19**, 118–126 (1995).
44. J.E. Akers, J.P. Agalloco, and C.M. Kennedy, *J. Pharm. Sci. Technol.* **49**, 140–144 (1995).
45. J. Akers, *BioPharm* **7**, 43–47 (1994).
46. Harvard Business School Case Study 9-696-083, Biogen, Inc.: rBeta Interferon Manufacturing Process Development, January 29, 1996.
47. G.E. Gamerman and B.F. Mackler, *CHEMTECH* **24**, 37–41 (1994).
48. J.A. Ehrich, D.F. Bernstein, and K.H. Sills, *Pharm. Technol.* **19**, 98–112 (1995).
49. G.J. Wiley, I. Ullah, and S.N. Agharkar, *Pharm. Technol.* **19**, 72–76 (1995).
50. J.R. Vogel, R.A. Schober, and R.B. Olson, *Pharm. Technol.* **19**, 40–52 (1995).
51. J.D. Bronzino ed., *The Biomedical Engineering Handbook*, CRC Press, Boca Raton, Fla., 1995; P.A. Oberg, in J.D. Bronzino ed., *The Biomedical Engineering Handbook*, CRC Press, Boca Raton, Fla., 1995, pp. 2549–2556; T.V. Holohan, in J.D. Bronzino ed., *The Biomedical Engineering Handbook*, CRC Press, Boca Raton, Fla., 1995, pp. 2794–2801.
52. B. Atkinson and F. Mavituna, *Biochemical Engineering and Biotechnology Handbook*, 2nd ed., Stockton Press, New York, 1991. See especially Chapter 19, pp. 1059–1109.
53. P.F. Stanbury, A. Whitaker, and S.J. Hall eds., *Principles of Fermentation Technology*, Elsevier Science, Oxford, U.K., 1995, pp. 331–349.
54. J.L. Dwyer, in F. Franks ed. *Protein Biotechnology*, Humana Press, Totowa, N.J., 1993, pp. 533–572.
55. H.B. Reisman, *Economic Analysis of Fermentation Processes*, CRC Press, Boca Raton, Fla., 1988.
56. J. Durkin, *Pharm. Manuf. Int.* 103–108 (1995).
57. S. Hubbard, *Pharm. Manuf. Int.* 109–112 (1995).
58. H. Allgaier, *Pharm. Biotech. Int.* 47–48 (1996).
59. R. Maycock, *Pharm. Manuf. Int.* 113–118 (1995).
60. D.S. Remer and J.H. Idrovo, *BioPharm* **3**, 36–42 (1990).
61. D.M. Haseltine, *Chem. Eng. Prog.* **92**, 26–32 (1996).
62. W.M. Goble and B.O. Paul, *Chem. Proc.* **58**, 35–39 (1995).
63. R. McKown, R. Teutonico, and S. Fox, *Genet. Eng. News* **16**, 6, 7, 28 (1996).
64. J. Busch, *Res. Technol. Manage.* **37**, 50–56 (1994).
65. R.L. Steinberger, *Chem Eng.* **101**, 114–117 (1994).
66. D. Petrides, E. Sapidou, and J. Calandranis, *Biotech. Bioeng.* **48**, 529–541.
67. R.A. Young, *Chem. Eng. Prog.* **91**, 64–59 (1995).
68. R. Datar, *Proc. Biochem.* **21**, 19–26 (1986).
69. R.V. Datar and T. Cartwright, *Bio/Technology* **11**, 349–357 (1993).
70. M.R. Ladisch and K.L. Kohlmann, *Biotechnol. Prog.* **8**, 469–478 (1992).
71. N.D. Hinman, D.J. Schell, C.J. Riley, P.W. Bergeron, and P.J. Walter, *Appl. Biochem. Biotechnol*, **34**, 639–649 (1992).
72. M. von Sivers, G. Zacchi, L. Olson, and B. Hahn-Hagerdal, *Biotechnol. Prog.* **10**, 555–560 (1994).
73. T.J. Ward, *Chem. Eng.* **101**, 102–107 (1994).
74. S.G. Woinsky, *Chem. Eng. Prog.* **92**, 33–37 (1996).
75. W.H. Ver Duin, *Better Products Faster: A Practical Guide to Knowledge-Based Systems for Manufacturers*, Irwin Professional Publishing, Burr Ridge, Ill., 1995.
76. K.G. Chapman, *Pharm. Eng.* **16**, 8–18 (1996).
77. M.S. Slusher, R.L. Bennett, and D.L. Deitz, *Pharm. Eng.* **14**, 8–12 (1994).
78. E.R. Crowe and C.A. Vassiliadis, *Chem. Eng. Prog.* **91**, 22–31 (1995).
79. D. Mosher and P. McArthur, *Pharm. Eng.* **14**, 86–92 (1994).
80. J.L. Tedesco, T. Olson, S. Tse, and J.M. Patrick, *Pharm. Technol.* **18**, 174–187 (1994).
81. M. Borrero, *Pharm. Technol.* **19**, 113–115 (1995).
82. S.S. Seaver, *BioPharm* **8**, 20–25 (1995).
83. D. Goswami and M.R. Lorenz, *Pharm. Eng.* **16**, 14–18 (1996).
84. G. Sofer, J. McEntire, and J. Akers, *BioPharm* **8**, 60–63 (1995).
85. J.L. Tedesco and M.J. Titus, *Pharm. Eng.* **15**, 22–28 (1995).
86. J. Swarbrick and J.C. Boylan eds., *Encyclopedia of Pharmaceutical Technology*, Marcel Dekker, New York, 1996; W.J.C. Currie and R.W.J. Currie, in J. Swarbrick and J.C. Boylan eds., *Encyclopedia of Pharmaceutical Technology*, Marcel Dekker, New York, 1996, pp. 353–370; R. Kaplan, in J. Swarbrick and J.C. Boylan eds., *Encyclopedia of Pharmaceutical Technology*, Marcel Dekker, New York, 1996, pp. 207–231.
87. K.H. Keller, *CHEMTECH* **25**, 7–11 (1995).
88. M.E. Watanabe, *Scientist* June 24, 1996, pp. 1, 4–5.
89. W.S. Grundfest, ed., *Health Care Technology Policy I: The Role of Technology in the Cost of Healthcare*, International Society for Optical Engineering, Bellingham, Wash., 1994.
90. W.S. Grundfest, ed., *Health Care Technology Policy II: The Role of Technology in the Cost of Healthcare: Providing the Solutions*, International Society for Optical Engineering, Bellingham, Wash., 1995.
91. S. Epstein and M.G. McGlynn, *Dis. Manage. Health Outcomes* **1**, 3–10 (1997).

ADDITIONAL READING

A. Kornberg, *The Golden Helix: Inside Biotech Ventures*, University Science Books, Sausalito, Calif., 1995. Relates the story of DNAX and its relations with Schering-Plough as seen by a key scientific adviser. Also comments on other companies and biotech future.

V. Moses and R.E. Cape eds., *Biotechnology: The Science and the Business*, Harwood Academic Publishers, Chur, Switzerland, 1991. Covers business matters, underlying technology, and some economics; various industries (including biomass, waste treatment, foods, and pesticides).

V. Moses and S. Moses, *Exploiting Biotechnology*, Harwood Academic Publishers, Chur, Switzerland, 1995. Discusses the technology base, organizational framework, and biotech businesses; contains an extensive glossary.

A. Prokop, in S. Neidleman and A.I. Laskin eds., *Advances in Applied Microbiology*, vol. 40, Academic Press, San Diego, Calif., 1995, pp. 95–154, 155–236. An excellent and up-to-date summary of the project cycle from discovery through development to marketing. Articles combine the technical and business aspects of biotechnology, touching on many key areas and needs.

ELECTRON TRANSPORT

DAVID WHITE
Indiana University
Bloomington, Indiana

KEY WORDS

ATP synthesis
Bacteria
Chemiosmosis
Coupling sites
Electrode potentials
Electron carriers
Mitochondria
Protonmotive force
Respiration
Uncouplers

OUTLINE

INTRODUCTION

There is a current of electrons in the cell membranes of bacteria and in mitochondrial and chloroplast membranes. The electrons enter the membrane from an electron donor usually metabolically generated or supplied in the growth medium, travel through the membrane via a series of electron carriers, and leave the membrane upon being transferred to a terminal electron acceptor such as oxygen. (Photosynthetic bacteria carry out a light-driven cyclic electron flow where the electrons do not leave the membrane.) According to the chemiosmotic theory, the current of electrons, called electron transport, produces a vectoral flow of positive charges, that is, protons (H^+) to the outside of cells and organelles (1) (Fig. 1). In bacteria, the protons are translocated to the outside of the cell membrane, in mitochondria the protons are translocated from the mitochondrial matrix across the inner membrane to the cytosolic side, and in chloroplasts the protons are translocated across the membrane from the stroma to the inner thylakoid space. In all three cases, a membrane potential, outside positive, develops as does a ΔpH, outside acidic (more concentrated in protons). The return of the protons through appropriate proteins toward the negative inner side and higher pH (lower proton concentration) drives the synthesis of ATP as well as other membrane-associated processes such as solute transport. The resultant ATP is the major source of biochemical energy for the synthesis of cell material. Thus, electron transport fuels cell metabolism. This chapter describes the general routes that the electrons take in various nonphotosynthetic membranes on their way to a terminal electron acceptor such as oxygen, as well as the differences among organisms, and how the the proton current is produced. Archaeal* electron transport chains are less well studied, but they appear to be similar to bacterial electron transport chains (2).

Some Terminology and Definitions

Redox, Reactions, Electron Transport, and Respiration. The basic reaction in electron transport is the transfer of electrons from one molecule to another. The transfer of electrons from one molecule to another is called an *oxidation-reduction reaction* (or redox reaction), where the molecule that gives up the electron becomes oxidized, and the molecule that accepts the electron becomes reduced. Electron transport is actually a series of such redox reactions in membranes where the electron is passed from one molecule to another. The molecules that transfer the electrons are called *electron carriers*, and these are described later. Electron transport takes place in cell membranes of prokaryotes (organisms such as bacteria that have no organelles such as nuclei or mitochondria) and in mitochondrial and chloroplast membranes of eukaryotes (organisms with organelles). Electron transport in chloroplast membranes is called photosynthetic electron transport, whereas electron transport in other membranes is called

*There are two phylogenetic lines of prokaryotes (microorganisms without organelles such as nuclei). These are the archaea and the bacteria.

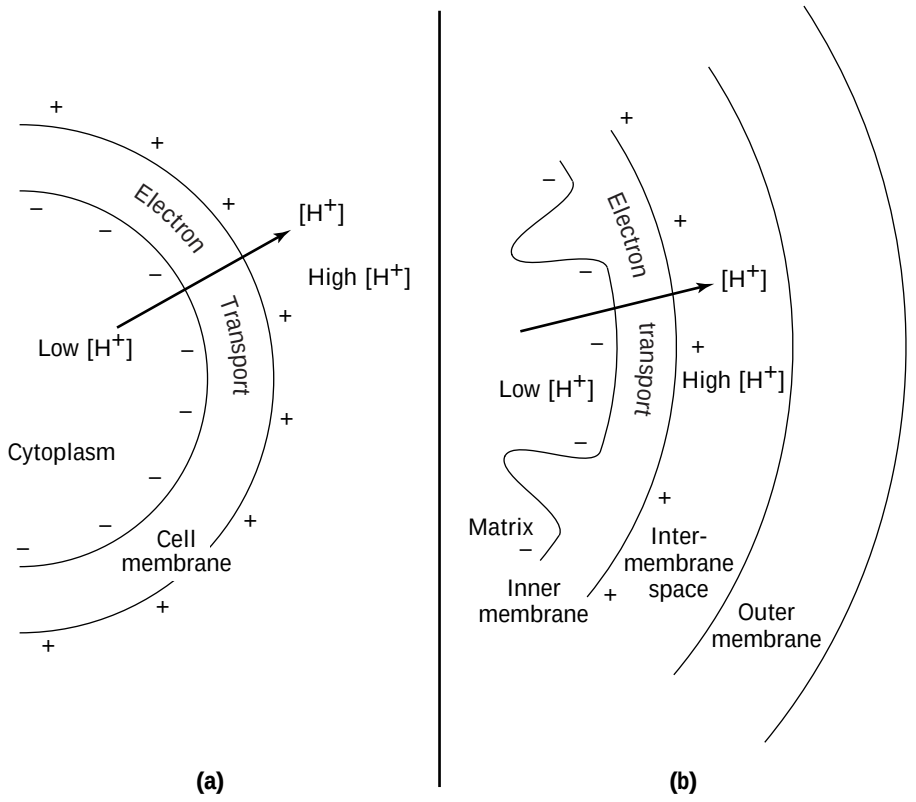


Figure 1. Electron transport in bacterial cell membranes (**a**) and mitochondrial membranes (**b**) result in proton translocation accompanied by a membrane potential (outside positive) and a pH gradient (outside acidic). The molecular mechanisms and the intermediates are not shown.

respiration. If the terminal electron acceptor in respiration is oxygen, then electron transport is called *aerobic respiration*. If it is not oxygen, then it is called *anaerobic respiration*. Mitochondria carry out only aerobic respiration, but many prokaryotes are capable of anaerobic respiration and use organic compounds such as fumaric acid or inorganic compounds such as nitrate or sulfate as terminal electron acceptors.

Electron Energy and the Redox Potential Difference. During electron transport, electrons spontaneously flow from donor to acceptor down a potential gradient called the ΔE, and it is the energy released during this process that is used to generate the vectoral flow of protons from inside to outside that is ultimately used to make ATP. By convention in biology, electron *donors* in a redox reaction that releases energy are said to have a *a more negative* redox value (E) than the molecules that accept the electron. That is to say, electrons always flow spontaneously toward the molecule with the higher redox value. The actual redox potential of a molecule depends on the ratio of oxidized to reduced forms.* The difference in electrical potential between the acceptor of the electron and the donor of the electron ($E_{\text{acceptor}} - E_{\text{donor}}$) is called the *redox potential* and is written as ΔE. The ΔE is always positive when work can be done by the moving electrons, that is, when the electrons are moving down an energy gradient. The ΔE is generally spoken of as an *electrical force* (also called the *electromotive force*) that maintains the current of electrons, and it is expressed in units of volts. The *work* that can be done by n moles of moving electrons depends on the total charge carried by the electrons and is given by $nF\Delta E$ joules where F is the charge carried by 1 mol of electrons or monovalent ion. (F is the *faraday* constant, which is equal to 96,485 coulombs per mole.) The unit of work is called the joule. In summary, then, electrons flowing over a potential gradient of ΔE volts can do a maximum of $nF\Delta E$ joules of work or exert a force of $n\Delta E$ volts. As described next, the work that is done is the creation of a protonmotive force that is, a Δp, by translocating protons to the outside to a higher electrochemical potential.

The Protonmotive Force. The protonmotive force, Δp, refers to the electrochemical potential difference of the proton across bacterial, mitochondrial, and chloroplast membranes. Its units are volts. In order to understand what this means, it is necessary to realize that a membrane potential, $\Delta\Psi$, exists across these membranes such that they are positively charged on the outer surface and negatively charged on the inner surface (Fig. 1.) (How the membrane potential is created will be explained later.) The size of the $\Delta\Psi$ across bacterial cell membranes varies from about -60 mV to -200 mV, depending on the growth conditions and

*E is sometimes called the electrode potential or the reduction potential. The value of E at any particular temperature depends on the pH and relative concentrations of oxidized and reduced forms of the compound. E_o is the redox potential in the standard state, where the oxidized and reduced forms of the compound both have an activity of 1, as does the hydrogen ion concentration. E'_o refers to the standard state at pH 7 and is usually called the midpoint potential ($E_{m,7}$). The actual redox potential of a compound depends on the ratio of oxidized to reduced forms in the cell and the pH and is called E_h when the activities are not 1, that is, $E_h = E_o + RT \ln[ox]/[red]$.

the bacterium. (Note that by convention the $\Delta\Psi$ is negative when the inside surface is negative with respect to the outside surface.) Because protons themselves are positively charged, it requires energy to move them to the positive side. That is to say, they are being pushed away by the $\Delta\Psi$ from the positive side toward the negative side. Another factor that will influence the direction that protons will spontaneously move is the concentration difference (ΔpH). That is to say, energy is required to move protons from the less concentrated to the more concentrated side of the membrane, and protons give up energy when they move from the more concentrated to the less concentrated side. The sum of the energies involved in the movement of protons along an electrical and concentration gradient is called the *electrochemical energy*, and it takes $+yF\Delta p$ joules to move y moles of protons against their electrochemical energy gradient from inside to outside. Accordingly, $-yF\Delta p$ joules is released and work can be done when y moles of protons return to the inside of the cell or organelle. This can be expressed as volts (a force) by dividing by the faraday ($y\Delta p$ volts). As discussed later, the energy to move protons against their electrochemical gradient is provided by respiration. In summary, then, the protonmotive force consists of contributions from the membrane potential and the ΔpH and is given in equation 1. Note that the Δp is analogous to the redox potential, ΔE, which describes the electrical force for moving electrons, except the Δp applies to moving protons and is due to both electrical and chemical energy.

$$\Delta p = \Delta\Psi - 60\ \Delta\text{pH mV at 30 °C} \qquad (1)$$

For most bacteria, the ΔpH is positive and thus contributes to the Δp. An exception are bacteria that live in environments more alkaline than the cytoplasm. Under these circumstances, the ΔpH is negative and lowers the energy available from the Δp.

The Relationship between ΔE and Δp

As mentioned, in prokaryotic cell membranes and in mitochondria (and chloroplasts), the work done by the current of electrons is to translocate protons (H^+) across the membrane to a higher electrochemical potential to create the Δp (Fig. 1). Proton translocation is done at certain redox reactions called coupling sites in the respiratory chain. The means by which this is done will be described later ("Q Loops, Q Cycles, and Proton Pumps"). The numerical relationship between the ΔE of the redox reaction at the coupling site and the Δp is given in equation 2.

$$-n\Delta E = y\Delta p \qquad (2)$$

In equation 2, n is the number of electrons transferred from the electron donor to the acceptor during the redox reaction, ΔE is the redox potential difference between the electron donor and acceptor, y is the number of protons translocated across the membrane when n electrons are transferred, and Δp is the protonmotive force. The ΔE is always positive when electrons can do work, and the Δp is always negative when protons can do work, hence the negative sign in equation 2.

Work Done by the Δp

When protons reenter the cell (or cross mitochondrial or chloroplast membranes) down their electrochemical gradient, work can be done driven by the Δp. The type of work done at the expense of the Δp depends on the specific route taken by the protons as they return across the membrane toward their lower electrochemical potential. ATP synthesis occurs when the protons cross the membrane via the membrane-bound ATP synthase, which is a complex of proteins that synthesizes ATP (see Fig. 10). Solute transport occurs when protons traverse the membrane via solute transporters, which are proteins that transport protons down their electrochemical gradient at the same time as solutes are transported against their concentration gradient, and flagellar rotation in prokaryotes occurs when the protons return to the inside of the cell via the flagellar motors, which consist of a complex system of proteins. (Flagellar motors in eukaryotes are driven by ATP rather than the Δp.) Thus, prokaryotic cell membranes as well as mitochondrial and chloroplast membranes can be viewed as containing electric machines that are powered by currents of protons, which themselves are maintained by electron transport.*

THE ELECTRON CARRIERS

Before describing the electron transport chains and the generation of the Δp, it is necessary to introduce the electron carriers. During electron transport, the electrons flow through protein and lipid electron carriers in the membranes. Some of these carry electrons only, whereas others carry both electrons and hydrogen atoms in combination. They are

1. Flavoproteins (hydrogen and electron carrier)
2. Quinones (hydrogen and electron carrier)
3. Iron–sulfur proteins (electron carrier)
4. Cytochromes (electron carrier)

The quinones are lipids (molecules that have a relatively poor solubility in water but a good solubility in lipid environments such as the interior of membranes), whereas the other electron carriers are proteins, which exist in multiprotein enzyme complexes called *oxidoreductases*.† The electrons and hydrogen are not carried in the protein per se, but in a nonprotein molecule tightly bound to the protein. The nonprotein portion that carries the electron and hydrogen is called a *prosthetic group*. The prosthetic group

*The sodium ion current also drives solute transport and in some cases flagellar rotation. Usually the Δp provides the energy to create the sodium gradient, but some marine bacteria couple electron transport to the creation of a sodium ion gradient rather than a proton gradient.

†For example, the enzyme complex that oxidizes NADH and reduces quinone is an NADH-quinone oxidoreductase. It consists of a flavoprotein and iron–sulfur proteins. (Reduced quinone is called quinol.) An enzyme complex that oxidizes quinol and reduces cytochrome c is called a quinol-cytochrome c oxidoreductase. It consists of cytochromes and an iron–sulfur protein.

in flavoproteins (fp) is a flavin that can be either flavin adenine dinucleotide (FAD) or flavin mononucleotide (FMN). The prosthetic group in iron–sulfur proteins is a cluster of iron–sulfide, which is abbreviated FeS. The prosthetic group in cytochromes is heme.

Each of the electron carriers has a different redox potential, and the electrons are transferred sequentially to a carrier of a higher redox potential. The standard redox potentials at pH 7 (E'_o) of the electron carriers and some electron donors and acceptors are listed in Table 1. Notice that redox couples are generally written in the form oxidized/reduced for biologists and biochemists. Several of the oxidation–reduction reactions in the electron transport chain can be reversed using excess energy provided by the Δp. During reversal of electron transport, the protons *enter* the cells driven by the Δp and the electrons travel to the more *negative* potential creating a ΔE. (See equation 2.)* For example, in Figure 7, the Δp can drive electrons from succinate through ubiquinone (UQ) to NAD^+. (NAD^+ is the oxidized form of NADH and will accept two electrons and a proton to become NADH.) This allows succinate to be used as an electron donor to reduce NAD^+. The NADH that is formed is used for biosynthesis. The reversal of electron transport is discussed later ("Coupling Sites").

*This means that the ratio of oxidation to reduction for several of the electron carriers (fp, cytochromes, quinones, FeS proteins) must be close to 1. Thus, for these reactions the E_h (actual potential at pH 7) of the redox couples are close to their midpoint potentials, E'_m, which is the potential at pH 7 when the couple is 50% reduced (i.e., [ox] = [red]).

Table 1. Standard Electrode Potentials (pH 7)

Couple	E'_o (mV)
Fd_{ox}/Fd_{red} (spinach)	−432
CO_2/formate	−432
H^+/H_2	−410
Fd_{ox}/Fd_{red} (*Clostridium*)	−410
NAD^+/NADH	−320
FeS(ox/red) in mitoch.	−305
Lipoic/dihydrolipoic	−290
S^o/H_2S	−270
$FAD/FADH_2$	−220
Acetaldehyde/ethanol	−197
$FMN/FMNH_2$	−190
Pyruvate/lactate	−185
OAA/malate	−170
Menaquinone(ox/red)	−74
$Cytb_{558}$(ox/red)	−75 to −43
Fum/succ	+33
Ubiquinone(ox/red)	+100
$Cytb_{556}$(ox/red)	+46 to +129
$Cytb_{562}$(ox/red)	+125 to +260
Cytd(ox/red)	+260 to +280
Cytc(ox/red)	+250
FeS(ox/red) in mitoch.	+280
Cyta(ox/red)	+290
$Cytc_{555}$(ox/red)	+355
$Cyta_3$(ox/red) in mitoch.	+385
NO_3^-/NO_2^-	+421
Fe^{3+}/Fe^{2+}	+771
O_2 (1 atm)/H_2O	+815

Note: Fd = ferredoxin; OAA = oxaloacetate; Fum = fumarate; succ = succinate.
Sources: Refs. 3 and 4.

Flavoproteins

Flavoproteins (fp) are electron carriers that have as their prosthetic group an organic molecule called flavin. There are two flavins, FMN and FAD (Fig. 2). The flavins are made from the vitamin riboflavin. The addition of phosphate to riboflavin yields FMN, and the addition of ADP to FMN yields FAD. When flavins are reduced they carry 2[H] (two electrons and two hydrogens), one on each of two ring nitrogens (Fig. 2). There are many different fp, and they differ with respect to their redox potentials and where they are situated in the electron transport chain. The differences in the redox potentials result from differences in the protein component of the enzyme, not in the flavin itself.

Quinones

Quinones are fat soluble, that is, they are soluble in the interior (lipid phase) of the membrane, which is nonaqueous. They are small and mobile and shuttle electrons between the complexes of protein electron carriers. Their structure and oxidation–reduction reactions are shown in Figure 3. Bacteria make two types of quinones that function during respiration: ubiquinone (UQ), a quinone also found in mitochondria; and menaquinone (MQ or sometimes MK). Menaquinones have a much lower electrode potential than ubiquinones and are used predominantly during anaerobic respiration where the electron acceptor has a low potential (e.g., during fumarate respiration). A third type of quinone, plastoquinone, occurs in chloroplasts

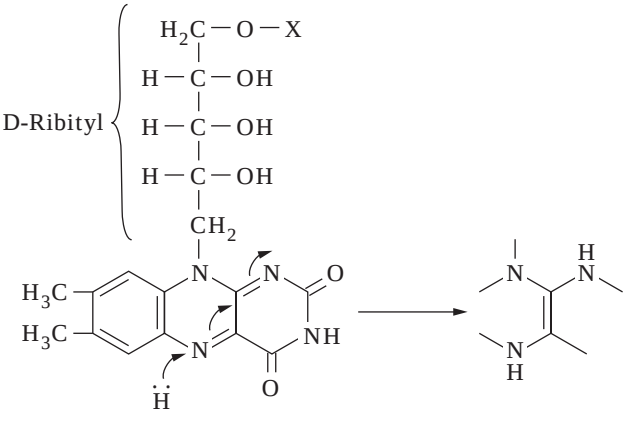


Figure 2. Structures of riboflavin, FMN, and FAD. The electronic shift upon reduction is shown. Riboflavin, X = H; FMN, X = PO_3H_2; FAD, X = ADP. For the sake of convenience, the reduction reaction is drawn as proceeding via a hydride ion even though this need not be the actual mechanism in all flavin reductions. *Source*: Ref. 5.

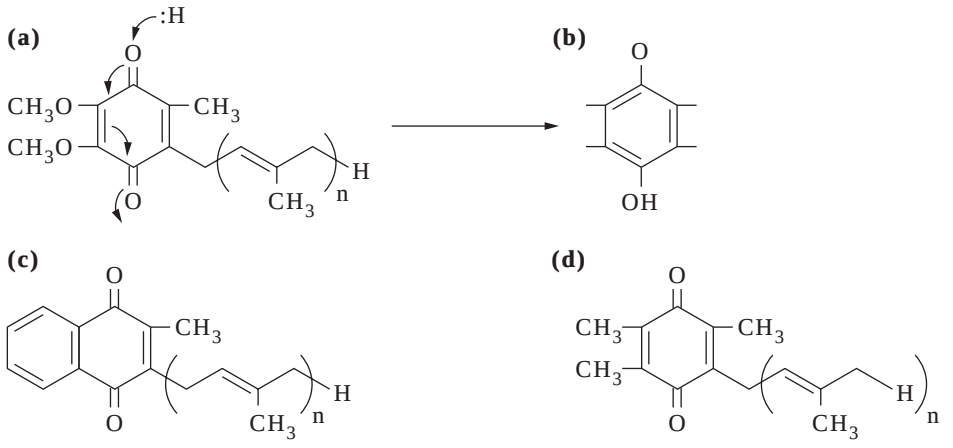


Figure 3. The structure of quinones. (**a**) Oxidized ubiquinone; (**b**) a portion of reduced ubiquinone; (**c**) oxidized menaquinone; (**d**) oxidized plastoquinone. The value of *n* can be 4–10 and is 8 for both quinones in *E. coli*. In *E. coli* ubiquinone plays a major role in aerobic and nitrate respiration, whereas menaquinone is dominant during fumarate respiration. One reason for this is that ubiquinone has a potential (E'_o) of +100 mV compared to + 30 mV for fumarate. It is therefore at too high of a potential to deliver electrons to fumarate. Menaquinone has a low potential, −74 mV, and is thus able to deliver electrons to fumarate. Plastoquinone is used in chloroplast and cyanobacterial photosynthetic electron transport. *Source*: Ref. 5.

and cyanobacteria, and functions in photosynthetic electron transport.

Iron–Sulfur Proteins

Iron–sulfur proteins contain nonheme iron and usually acid-labile sulfur (Fig. 4). The term *acid-labile* sulfur means that H_2S is released from the protein when the pH is lowered to approximately 1. This is because there is sulfide attached to iron by bonds that are ruptured in acid. Generally, the proteins contain clusters in which iron and acid-labile sulfur are present in a ratio of 1:1. However, there may be more than 1 FeS cluster per protein. For example, in mitochondria the enzyme complex that oxidizes NADH has at least four FeS clusters (see Fig. 9). The FeS clusters have different E_h values, and the electron travels from one FeS cluster to the next toward the higher E_h. (The E_h is the redox potential at physiological concentrations of oxidized and reduced forms.) It appears that the electron may not be localized on any particular iron atom, and the entire FeS cluster should be thought of as carrying one electron, regardless of the number of Fe atoms. These proteins also contain cysteine sulfur, which is not acid labile and which bonds the iron to the protein. There are several different types of FeS proteins, and these catalyze numerous oxidation–reduction reactions in the cytoplasm as well as in the membranes. The FeS proteins cover a very wide range of potentials, from approximately −400 mV to +350 mV. They therefore can carry out oxidation–reduction reactions at both the low potential end and the high potential end of the electron transport chain, and indeed they are found in several locations. An example of an FeS cluster is shown in Figure 4. Note that each Fe is bound to two acid-labile S and two cysteine S. This would be called an Fe_2S_2 cluster.

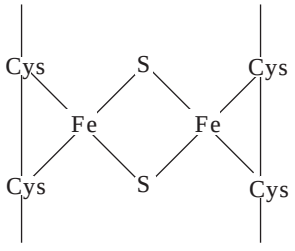


Figure 4. FeS cluster. This is an Fe_2S_2 cluster. More than one cluster may be present per protein. The sulfur atoms held only by the iron are acid labile. The iron is bonded to the protein via sulfur in cysteine residues. *Source*: Ref. 5.

Cytochromes

Cytochromes are electron carriers that have heme as the prosthetic group. Heme consists of four pyrrole rings attached to each other by methene bridges (Fig. 5). Because hemes have four pyrroles, they are called *tetrapyrroles*. Each of the pyrrole rings is substituted by a side chain. Substituted tetrapyrroles are called *porphyrins*. Therefore, hemes are also called porphyrins. (An unsubstituted tetrapyrrole is called a porphin.) Hemes are placed in different classes, described next, on the basis of the side chains attached to the pyrrole rings. In the center of each heme there is an iron atom that is bound to the nitrogen of the pyrrole rings. The iron is the electron carrier and is oxidized to ferric or reduced to ferrous ion during electron transport.* Cytochromes are therefore one-electron carri-

*The electron carried by the iron is partly delocalized over the π-electrons of the heme.

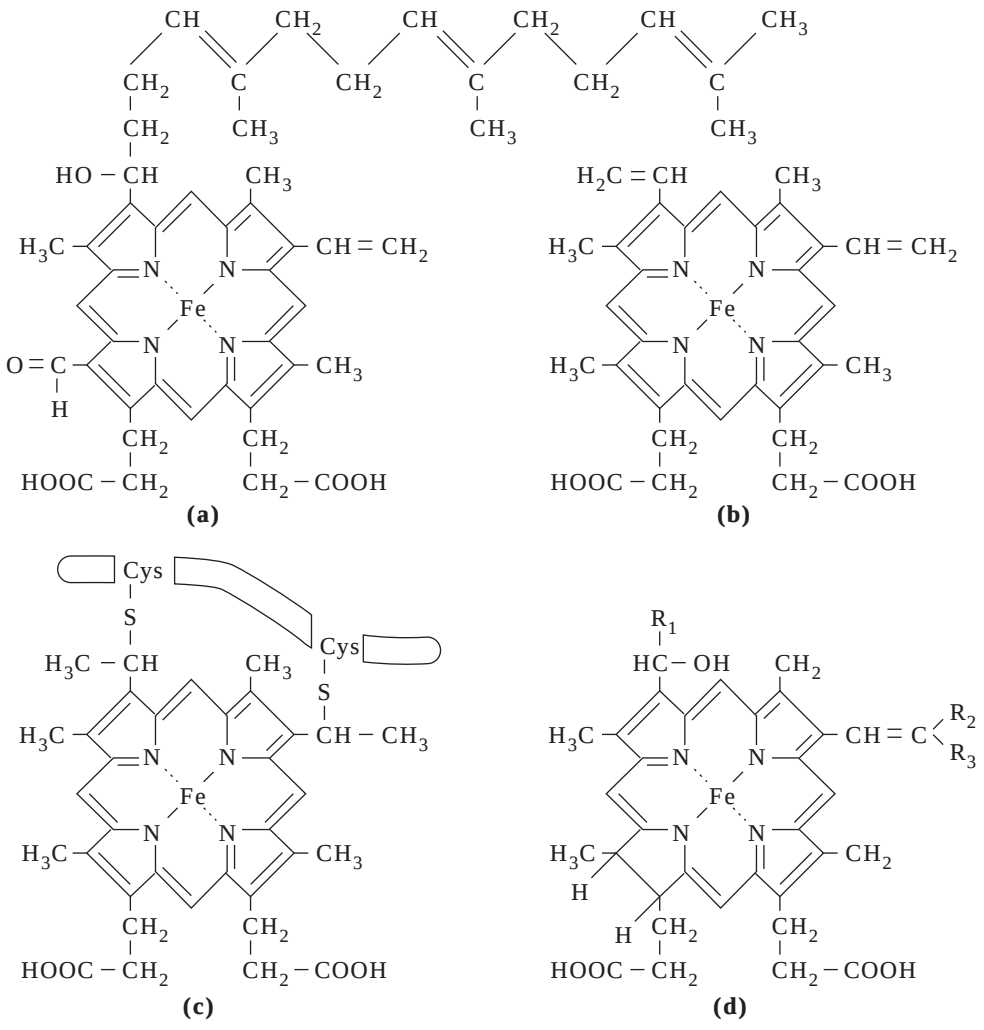


Figure 5. The prosthetic groups of the different classes of cytochromes. The hemes vary according to their side groups. Heme c is covalently bound to the protein via a sulfur bridge to a cysteine residue on the protein. The structures of R_1, R_2, and R_3 are not known. *Source*: Ref. 6.

ers. The E_h of the different cytochromes differ depending on the protein and the molecular interactions with surrounding molecules.

Classes of Cytochromes. There are four classes of cytochromes (a, b, c, and d) that have hemes a, b, c, and d. (Cytochrome o in bacteria is a b-type cytochrome.) As mentioned previously, the hemes can be distinguished according to the side groups that they possess. Hemes can also be distinguished spectrophotometrically. When cytochromes are in the reduced state they have characteristic light absorption bands in the visible range because of absorption by the heme. These are the α, the β, and the γ bands. The α bands absorb light between 500 and 600 nm, the β absorb at a lower wavelength, and the γ bands are in the blue region of the spectrum. The spectrum for a cytochrome c is shown in Fig. 6. Cytochromes are distinguished, in part, by the position of the maximum in the α band. For example, cyt. b_{556} and cyt. b_{558} differ because the former has a peak at 556 nm and the latter a peak at 558 nm.

ORGANIZATION OF THE ELECTRON CARRIERS IN MITOCHONDRIA

The electron carriers are organized as an electron transport chain in the inner mitochondrial membrane. The scheme is shown in Figure 7. The protein electron carriers (with the exception of cytochrome c) are organized in the membrane as individual complexes (complexes I–IV) that can be isolated from each other by appropriate separation techniques after mild detergent extraction. Four complexes can be recognized in mitochondria. They are complex I (NADH–ubiquinone oxidoreductase), complex II (succinate dehydrogenase), complex III (ubiquinol–cytochrome c oxidoreductase, also called the bc_1 complex), and complex IV (cytochrome c oxidase, which is cytochrome

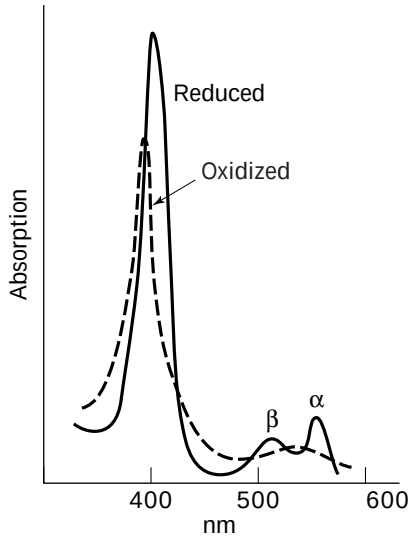


Figure 6. Absorption spectra of oxidized and reduced cytochrome c. The α band in the reduced form is used to identify cytochromes. *Source*: Ref. 5.

aa_3). Cytochrome c exists as a peripheral membrane protein that is easily extracted with water from the inner membrane. Complexes I, III, and IV are coupling sites, a Δp is generated by these complexes (see "Coupling Sites"). The electron donors are either NADH or succinic acid. NADH is generated during metabolism by various oxidation reactions in the cytoplasm. It must be reoxidized to NAD^+ in order that the cytoplasmic oxidations continue, and this is done through electron transport in respiring organisms. Succinic acid is generated from various organic compounds such as carbohydrates, amino acids, and fatty acids. It is actually an intermediate stage in the oxidation of these compounds to CO_2 (in the citric acid cycle). Succinic acid can also serve as an exogenous source of carbon and energy for growth for many microorganisms. The electrons and hydrogen are removed from the NADH or succinic acid by enzyme complexes called *dehydrogenases* (complexes I and II). The dehydrogenases are named after the substrate that donates the electrons. For example, NADH donates electrons and hydrogen to NADH dehydrogenase (complex I), and succinic acid donates electrons and hydrogen to succinate dehydrogenase (complex II). The dehydrogenases then transfer the electrons to UQ, which then transfers the electrons to cytochromes (complex III). There is a pattern here that is worth emphasizing because it is seen in bacteria as well. Dehydrogenases remove electrons and pass these to quinones which in turn pass the electrons to cytochrome oxidase complexes that transfer the electrons to oxygen. The value of the quinone is that it is a small diffusible molecule that connects the dehydrogenase complexes with the cytochrome complexes.

Each complex can have several proteins. The most intricate is complex I from mammalian mitochondria, which has about 40 polypeptide subunits, at least four FeS centers, one FMN, and one or two bound UQs. Analogous complexes have been isolated from bacteria, but in some cases (e.g., NADH–ubiquinone oxidoreductase and the bc_1 complex), they have fewer protein components (7–10).

Inhibitors of Electron Transport

There are several well-known inhibitors of the electron transport enzymes. These include rotenone and amytal, which inhibit complex I, antimycin A, stigmatellin, and myxothiazol, which inhibit complex III, and cyanide, azide, and carbon monoxide, all of which inhibit complex IV.

ORGANIZATION OF THE ELECTRON CARRIERS IN BACTERIA

Bacterial electron transport chains vary among the different bacteria and also according to the growth conditions.

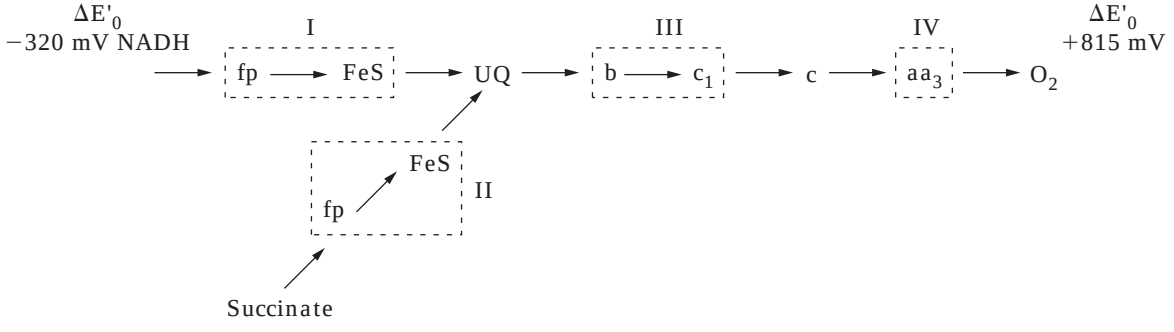


Figure 7. Electron transport scheme in mitochondria. Electrons travel in the electron transport chain from a low to a high electrode potential. Complexes I to IV are bracketed by dotted lines. Complex I is NADH dehydrogenase, also called NADH-ubiquinone oxidoreductase. Complex II is succinate dehydrogenase, also called succinate–ubiquinone oxidoreductase. Complex III is the bc_1, complex, also called ubiquinol–cytochrome c oxidoreductase. Complex IV is the cytochrome aa_3 oxidase, also called cytochrome c oxidase. There are several FeS clusters in complexes I and II and an FeS protein in complex III. Complex II also has a cytochrome b. fp, flavoprotein; FeS, iron–sulfur protein; UQ, ubiquinone; b, cytochrome b; c_1, cytochrome c_1; c, cytochrome c; aa_3, cytochrome aa_3. *Source*: Ref. 5.

However, there are certain features that bacterial electron transport chains have in common with each other and with mitochondria. As with the mitochondrial electron transport chain, the bacterial chains are organized into dehydrogenase and oxidase complexes connected by quinones (Fig. 8). The quinones accept electrons from dehydrogenases and transfer these to oxidase complexes via cytochromes, for example, via cytochromes b and c to cytochrome o or cytochrome aa_3 oxidase. As with mitochondria, the oxidase complexes reduce oxygen.

As opposed to mitochondria, bacteria are capable of using electron acceptors other than oxygen, denoted as Y in Figure 8b, during *anaerobic* respiration. A commonly used electron acceptor other than oxygen is nitrate, which can be reduced to nitrite, ammonia, or nitrogen gas. In fact, the depletion of nitrate from anaerobic soils is due to a large extent to bacteria carrying out anaerobic respiration and reducing the nitrate to nitrogen gas that escapes into the atmosphere. Sulfate is also used as an electron acceptor by certain anaerobic bacteria that are responsible for the production of hydrogen sulfide in various ecological niches, including sulfur springs. The enzyme complexes that reduce electron acceptors other than oxygen are called *reductases*, rather than oxidases.

The dehydrogenase complexes are more varied than found in mitochondria. Some of the dehydrogenase complexes are NADH and succinate dehydrogenase complexes analogous to complex I and II in mitochondria. However, in addition to these dehydrogenases, there are several others that reflect the diversity of substrates oxidized by the bacteria as sources of carbon, energy and electrons. For example, there are formic acid dehydrogenase, lactic acid dehydrogenase, methanol dehydrogenase, methylamine dehydrogenase, and so on. Many bacteria also have hydrogen gas (H_2) dehydrogenases (called hydrogenases) and use hydrogen gas as a source of energy and electrons.

Some bacteria grow on inorganic compounds as a source of electrons and energy and therefore have enzymes to oxidize the inorganic compounds and transfer the electrons to the electron transport scheme. Inorganic compounds that can be used for this purpose by certain bacteria include ammonia, nitrite, ferrous ion, hydrogen sulfide, sulfur, and (as mentioned) hydrogen gas. These bacteria have an enormous ecological impact and are crucial in the recycling of inorganic materials in the biosphere.

Depending on the source of electrons and electron acceptors, bacteria can synthesize and substitute one dehydrogenase complex for another, or reductase complexes for oxidase complexes. For example, when growing anaerobically, *Escherichia coli* makes the reductase complexes instead of the oxidase complexes, represses the synthesis of some dehydrogenases, and stimulates the synthesis of others. For this reason, the electron carrier complexes in bacteria are sometimes referred to as modules because they can be synthesized and plugged into the respiratory chain when needed.

Another important difference between mitochondrial and bacterial electron transport chains is that the routes to oxygen in the bacteria are branched, the branch point being at the quinone or one of the cytochromes. Under aerobic conditions there are often two or three branches leading to different oxidases. For example, a two-branched

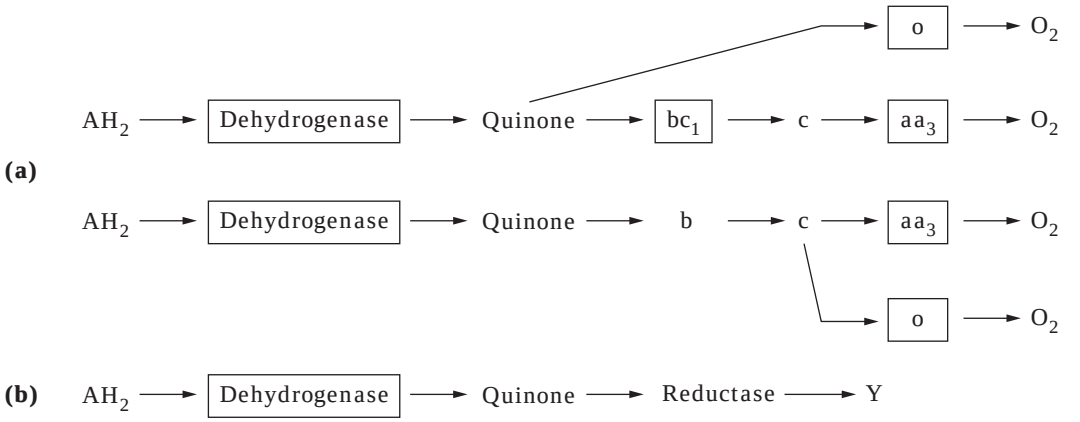


Figure 8. Generalized electron transport pathways found in bacteria. The details will vary depending on the bacterium and the growth conditions. (**a**) Aerobic respiration. A dehydrogenase complex removes electrons from an electron donor and transfers these to a quinone. The electrons are transferred to an oxidase complex via a branched pathway. Depending on the bacterium, the pathway may branch at the quinone or at cytochrome. Many bacteria have bc_1, cytochrome c, and cytochrome aa_3 in one of the branches and in this way resemble mitochondria. Other bacteria do not have a bc_1 complex, and may or may not have cytochrome aa_3. (**b**) anaerobic respiration. Under anaerobic conditions the electrons are transferred to reductase complexes that are synthesized anaerobically. Several reductases exist, each one specific for the electron acceptor. Y represents either an inorganic electron acceptor other than oxygen, such as nitrate, or an organic electron acceptor, such as fumarate. More than one reductase can simultaneously exist in a bacterium. *Source*: Ref. 5.

electron transport chain might contain a branch leading to cytochrome o oxidase and a branch leading to cytochrome aa_3 oxidase (cytochrome c oxidase). Other bacteria (e.g., *E. coli*) have cytochrome o and d oxidase branches but do not have the cytochrome aa_3 branch. The ability to synthesize branched electron transport pathways to oxygen confers flexibility on the bacteria, because the branches may differ not only in the Δp that can be generated (because they may differ in the number of coupling sites) but their terminal oxidases may also differ with respect to affinities for oxygen. Switching to an oxidase with a higher affinity for oxygen allows the cells to continue to respire even when oxygen tensions fall to very low values. This is important to ensure the reoxidation of electron carriers such as NADH in order that the various oxidative steps in metabolism continue at an acceptable rate.

COUPLING SITES

Sites in the electron transport pathway where redox reactions are coupled to proton translocation creating a Δp are called *coupling sites* (11). In mitochondria there are three coupling sites, and they are called sites 1, 2, and 3. These are shown in Figure 7. Site 1 is the NADH dehydrogenase complex (complex I), site 2 is the bc_1 complex (complex III), and site 3 is the cytochrome aa_3 complex (complex IV). The succinate dehydrogenase complex (complex II) is not a coupling site. The ratios of protons translocated per two electrons vary, depending on the complex. The bc_1 complex translocates four protons per two electrons, and depending on the reported value, complexes I and IV translocate from two to four protons per two electrons. (The consensus value for mitochondrial complex I is $4H^+/2e^-$.) Parts of the electron transport chain are reversible, a necessary situation if the source of electrons for biosynthesis is a molecule with a redox potential higher than that of the cell material being synthesized. For example, bacteria growing on ammonia or nitrite as a source of electrons must reverse electron transport in order to synthesize NADH, which is used as a source of electrons for biosynthesis. Electron transport must also be reversed when succinate is the electron donor for biosynthesis. During reversed-electron flow, protons enter through coupling sites 1 and 2, driven by the Δp, and the electrons are driven toward the lower redox potential. This creates a positive ΔE at the expense of the Δp (see equation 1.) Coupling site 3 is not physiologically reversible. Thus, water cannot serve as a source of electrons for NAD^+ reduction using reversed electron flow. However, during oxygenic photosynthesis, light energy can drive electrons from water to $NADP^+$, producing NADPH and O_2. The mechanism of photoreduction of $NADP^+$ is different from reversed electron flow.

P/O And P/2e Ratios

The synthesis of ATP during respiration is called *oxidative phosphorylation* or *respiratory phosphorylation*. The number of ATPs made per 2 e^- transfer to oxygen is called the *P/O ratio*. It is equal to the number of ATP molecules formed per atom of oxygen taken up when the Δp is used solely for ATP synthesis. When an electron acceptor other than oxygen is used, then P/2e^- is substituted for P/O. In mitochondria, the oxidation of NADH results in a P/O ratio of about 3 reflecting the three coupling sites that exist between NADH and O_2. These are complexes I, III, and IV. The use of succinate as an electron donor results in a P/O ratio of approximately 2 because electrons from succinate enter at the UQ level, bypassing coupling site 1 (complex I) that occurs between NADH and UQ (i.e., the NADH-ubiquinone oxidoreductase reaction [Fig. 7]). The other two coupling sites occur between UQ and oxygen. When electrons enter the respiratory chain after the bc_1 complex, the P/O ratio is reduced to 1, indicating that the bc_1 complex (complex III) is the second coupling site and that site 3 is cytochrome aa_3 oxidase (complex IV). Each of these sites is characterized by a drop in midpoint potential of about 200 mV, which is sufficient for generating the Δp required for ATP synthesis (Fig. 9). The size of the Δp that can be generated with respect to the ΔE is given in equation 1.

Number of ATPs Made According to Chemiosmotic Theory

Although it is usually stated that the number of ATPs made per coupling site is 1, the number of ATPs made at a coupling site need not be a whole number. This is because the amount of ATP that can be made per coupling site is equal to the ratio of protons translocated at the coupling site to protons that reenter via the ATP synthase, and this ratio need not be 1 (Fig. 10). For example, if 2 mol of protons is translocated at a coupling site and 3 mol of protons enters through the ATP synthase, then 2/3 mol of ATP can be made when 2 mol of electrons passes through the coupling site. The ratio of protons translocated per two electrons traveling through all the coupling sites to O_2 is called the H^+/O ratio. It can be measured by administering a pulse of a known amount of oxygen to an anaerobic suspension of mitochondria or bacteria and measuring the initial efflux of protons with a pH electrode as the small amount of oxygen is used up. A permeable anion such as SCN^- must be present so that proton extrusion is electroneutral, thereby allowing protons to accumulate in the bulk phase. The reported values for H^+/O for NADH oxidation vary. However, there is a consensus that the true value is probably around 10. The ratio of protons entering via the ATP synthase to ATP made is called the H^+/ATP ratio. Values of H^+/ATP from 2 to 4 have been reported, and a consensus value of 3 can be used for calculations. For intact mitochondria, an additional H^+ is required to bring P_i electroneutrally from the cytosol into the mitochondrial matrix in symport with H^+, so H^+/ATP would be 4. A value of 10 for H^+/O predicts a maximum P/O ratio of 2.5 (10/4) for mitochondria. This means that the often-stated value of 3 for a P/O ratio for NADH oxidation by mitochondria may be a little too high. The number of protons translocated per 2 *e*-traveling between succinate and oxygen (couplings sites 2 and 3) in mitochondria is 6. Therefore, the maximal P/O ratio for this segment of the electron transport chain in mitochondria may be 6/4 or 1.5, rather than 2.

Uncouplers

Uncouplers of oxidative phosphorylation are toxic compounds that short-circuit the proton current so that res-

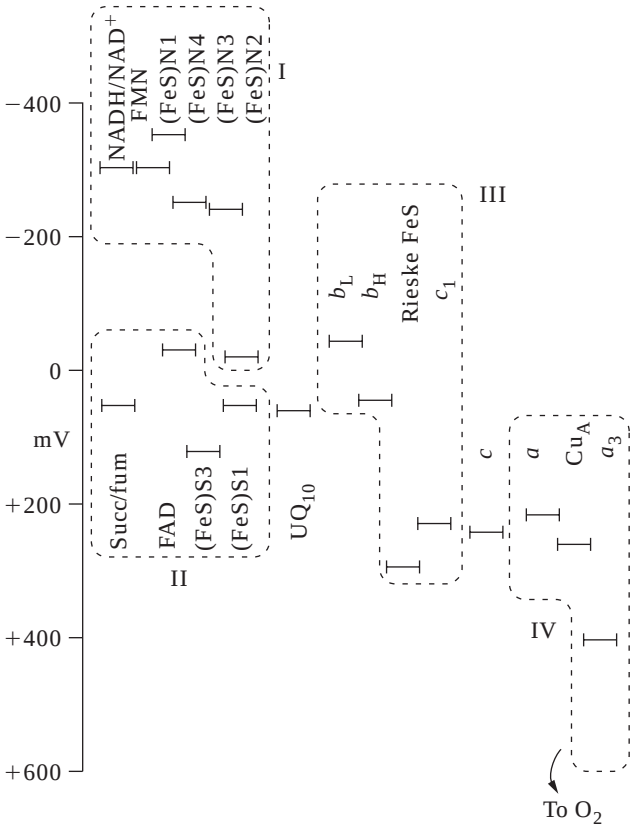


Figure 9. Average midpoint potentials, E'_m, for components of the mitochondrial respiratory chain. The complexes are in dotted boxes. The actual potentials (E_h) for most of the components is not very different from the midpoint potentials. An exception is cytochrome aa_3 whose E_h is much more positive than its midpoint potential. There are three sites where there are changes in potential of 200 mV or more. These drive proton translocation. One site is within complex I, a second within complex III, and the third between complex IV and oxygen. *Source*: Ref. 12.

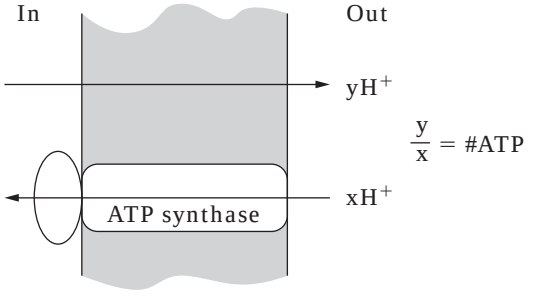


Figure 10. The ratio of protons extruded to protons translocated through the ATPase determines the amount of ATP made. The depiction of the ATP synthase does not attempt to describe the known complexity of its structure. *Source*: Ref. 5.

piration continues (usually at an increased rate) but ATP synthesis stops. The reason that ATP synthesis stops is that the protons enter via the uncoupler rather than through the ATP synthase. Two examples of uncouplers are dinitrophenol (DNP) and carbonyl cyanide-*p*-trifluoromethylhydrazone (FCCP). They are organic compounds that equilibrate protons across the membrane. As a consequence, the Δp is collapsed.

Q Loops, Q Cycles, and Proton Pumps

The previous discussion points out that proton translocation takes place at coupling sites when electrons travel downhill over a potential gradient of at least 200 mV (Fig. 9.) However, the mechanism by which the redox reaction is actually coupled to proton translocation was not explained. There are two ways in which this is thought to occur: (*1*) a Q loop or Q cycle, and (*2*) a proton pump. A Δp is established in a similar way in both the Q loop and Q cycle, as described next. However, some important differences do exist between these two mechanisms, and these are discussed separately.

In the Q loop or Q cycle, reduced quinone carries hydrogen atoms and electrons (the equivalent of H_2) across the membrane and becomes oxidized, releasing only protons on the external face of the membrane because the electrons return electrogenically via electron transport carriers to the inner membrane surface. When this occurs a $\Delta\Psi$ is established, outside positive because a positive charge (the proton) has been left on the outer surface while a negative charge (the electron) moves across the membrane toward the inside surface. (Electrogenic means that a membrane potential is established.) On the cytoplasmic side, the same number of protons are taken up as were released on the outside as the terminal oxidant (e.g., oxygen) is reduced by the electrons. The result is the net translocation of protons from the inside to the outside, although protons per se do not actually traverse the membrane, as well as the generation of a $\Delta\Psi$. A ΔpH can be established if the protons accumulate in the bulk phase outside of the cell. This can occur only if an equal number of anions (negatively charged ions) also accumulate to balance the positive charge of the protons, or if the protons exchange with cations in the medium. For most bacteria, the major portion of the Δp is due to the $\Delta\Psi$, although a small ΔpH is also created. An exception are bacteria living in environments that are much more acidic than their cytoplasm. For these bacteria, the Δp is due mostly to the ΔpH.

The Q Loop. The essential feature of the Q loop model is that the electron carriers alternate between those that carry both hydrogen and electrons (flavoproteins and quinones) and those that carry only electrons (FeS proteins and cytochromes). This is illustrated in Figure 11. The electron carriers and their sequence are flavoprotein (H carrier), FeS protein (e^- carrier), quinone (H carrier), and cytochromes (e^- carriers). The flavoprotein and FeS protein comprise the NADH dehydrogenase, which can be a coupling site, although the mechanism of proton translocation by the NADH dehydrogenase is not understood. Perhaps it is a proton pump. When electrons are transferred

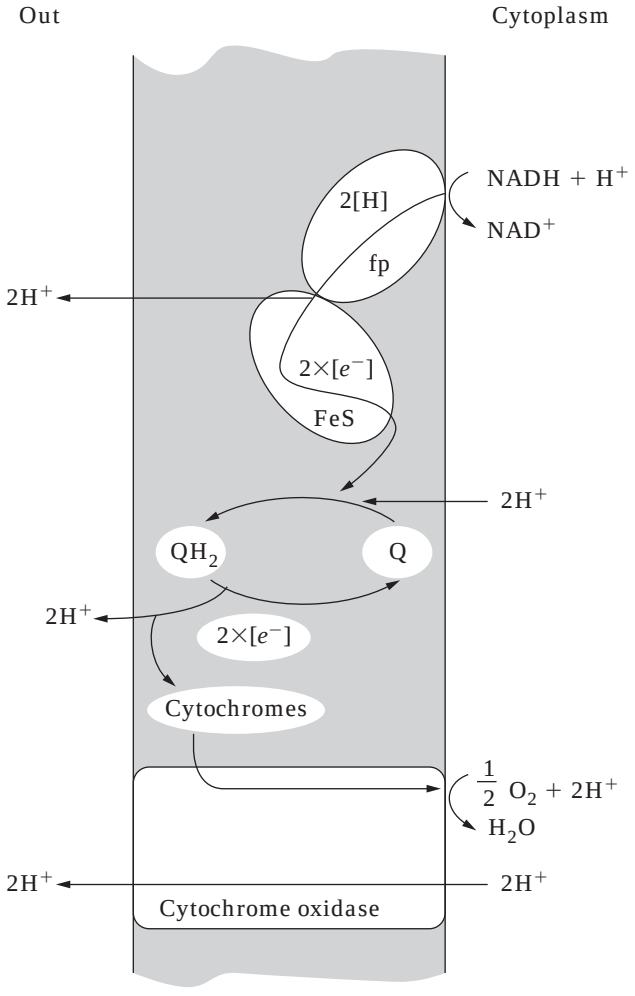


Figure 11. Proton translocation showing a Q loop and a proton pump. The diagram does not attempt to show the actual locations of the electron carriers within the membrane. It is proposed that the electron carriers exist in an alternating sequence of hydrogen (H) and electron (e-) carriers. This would be (H)flavoprotein (fp), (e-) iron–sulfur protein (FeS), (H)quinone (Q), and (e-)cytochromes. Oxidation of the fp deposits two protons on the outer membrane surface. The electrons return to the inner membrane surface where a quinone is reduced, taking up two protons from the cytoplasm. The reduced quinone diffuses to the outer surface of the membrane where it is oxidized, depositing two more protons on the surface. The electrons return to the cytoplasmic surface via cytochromes where they reduce oxygen in a reaction that consumes protons. Some cytochrome oxidases function as proton pumps. During anaerobic respiration the cytochrome oxidase is replaced by a reductase, and the electrons reduce some other electron acceptor, such as nitrate or fumarate. It should be noted that electrons can also enter at the level of quinone, such as from succinate dehydrogenase. *Source*: Ref. 5.

from the FeS protein to quinone (Q) on the inner side of the membrane, two protons are acquired from the cytoplasm. According to the model, the reduced quinone (QH_2), called quinol, then diffuses to the outer membrane surface and becomes oxidized, releasing the two protons. The electrons then return electrogenically via cytochromes to the inner membrane surface where they reduce oxygen. This would create a $\Delta\Psi$ because the protons are left on the outer surface of the membrane as the electrons move electrogenically to the inner surface. Thus, quinol oxidation is a second coupling site. As mentioned previously, the energy to create the membrane potential is derived from the ΔE between the oxidant and the reductant. One way to view this is that the energy from the ΔE pushes the electron to the negative membrane potential on the inside surface. Note that the role of the quinone is to ferry the hydrogens across the membrane, presumably by diffusing from a reduction site on the cytoplasmic side of the membrane to an oxidation site on the outer side.

During reduction of the terminal electron acceptor, cytoplasmic protons are consumed. For example, during reduction of $1/2O_2$ to H_2O, two protons are taken up from the cytoplasm. Thus, the result is a net translocation of two protons to the outside per QH_2 oxidation, even though there was no transmembrane movement of protons, per se. For the purpose of calculating the expected Δp generated at a coupling site, it makes no difference whether one postulates the transmembrane movement of protons in one direction or electrons in the opposite direction, because they both carry the same charge, and equation 1 can be used.

The Q Cycle. Although the linear Q loop as described above for the oxidation of quinol may accurately describe quinol oxidation in *E. coli* and some other bacteria, it is inconsistent with experimental observations of electron transport in mitochondria, chloroplasts, and many bacteria. For example, the Q loop predicts that the ratio of H^+ released per QH_2 oxidized is 2, whereas the measured ratio in mitochondria and many bacteria is actually 4. In order to account for the extra two protons, Mitchell suggested a new pathway for the oxidation of quinol called the *Q cycle*, which is more complicated than the Q loop (13) (Fig. 12.) The Q cycle operates in an enzyme complex called the bc_1 complex (complex III). The bc_1 complex from bacteria contains three polypeptides. These are cytochrome b with two b-type hemes, an FeS protein containing a single 2Fe-2S cluster (the Rieske protein), and cytochrome c_1 with one heme. The complex spans the membrane and has a site for binding reduced ubiquinone, UQH_2, on the outer surface of the membrane called site P (for positive); and a second site on the inner surface for binding UQ, site N (for negative) (Fig. 12). At site P, UQH_2 is oxidized to the semiquinone anion, UQ^- by the removal of one electron (Fig. 12**a**), and the two protons are released to the membrane surface. The electron travels to the FeS protein, and from there to cytochrome c_1 on its way to oxygen. The UQ^- is then oxidized to UQ by the removal of the second electron, which is transferred to b_{556}, also called b_L because of its relatively low *E*. The electron is then transferred across the membrane to b_{560}, also called b_H because of its relatively high *E*. Heme b_H transfers the electron to UQ bound at site N, reducing it to the semiquinone anion, UQ: *Electron flow from the Q_P site to the Q_N site is transmembrane and creates a membrane potential.* A second UQH_2 is oxidized at the P site and the UQ_N^- is reduced to UQH_2, picking up two protons from the cytoplasm (Fig. 12**b**). The UQH_2 enters the

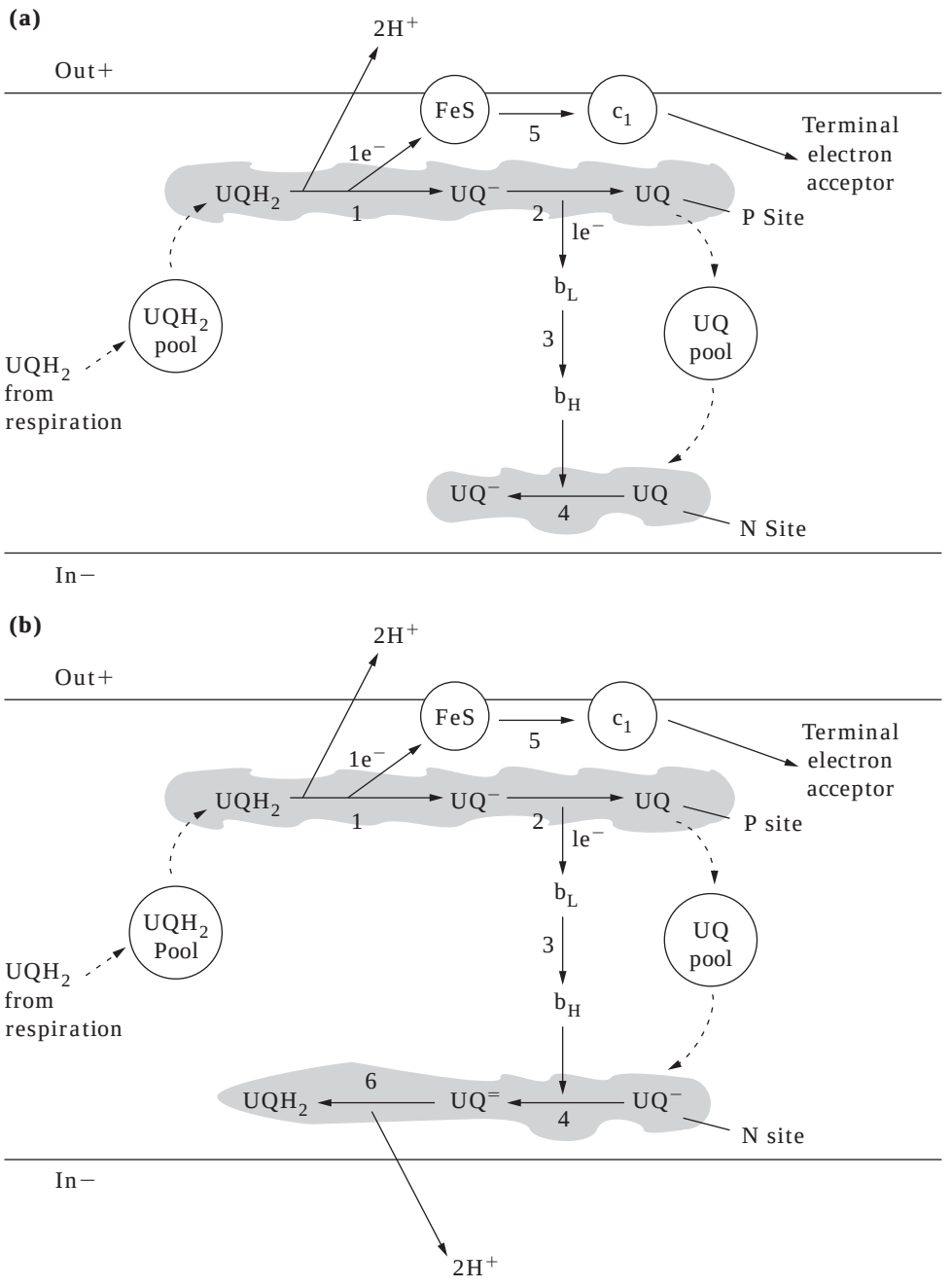


Figure 12. The bc_1 complex. The bc_1 complex isolated from bacteria contains three polypeptides, which are a cytochrome b containing two heme b groups, b_H and b_L, an iron-sulfur protein, FeS (the Rieske protein), and cytochrome c_1. It is widely distributed among the bacteria, including the photosynthetic bacteria. (Mitochondrial bc_1 complexes are similar but contain an additional 6 to 8 polypeptides without prosthetic groups.) The FeS protein and cytochrome c_1 are thought to be located on the outside (positive or p) surface of the membrane and the cytochrome b is believed to span the membrane acting as an electron conductor. On the outer surface there is a binding site in the bc_1 complex for ubiquinol, UQH_2 (the P site). On the inner surface (negative or n) there is a binding site in the bc_1 complex for ubiquinone, UQ (the N site). (**a**) Reduced ubiquinone binds to the P site and one electron is removed forming the semiquinone anion, UQ^- (reaction 1). At this time two protons are released on the outer membrane surface. The electron that is removed is transferred to the FeS protein and from there to cytochrome c1 (step 5). (Cytochrome c_1 transfers the electron to the terminal electron acceptor via a series of electron carriers in other reactions.) In reaction 2 the second electron is removed from the semiquinone anion producing the fully oxidized quinone, UQ. The second electron is transferred transmembrane via cytochrome b to ubiquinone at site N (steps 3 and 4). Because the electron travels transmembrane a membrane potential is created, outside positive. (**B**). A second reduced ubiquinone is oxidized at the P site, releasing two more protons, and the sequence of electron transfers is repeated. The UQ^- at site N becomes reduced to UQH_2, having acquired two protons from the cytoplasm. Note that for every two UQH_2 molecules that are oxidized, four protons are released, and one UQH_2 is regenerated. Therefore, four protons are released per one UQH_2 oxidized. Another way of saying this is that the ratio, H^+/e^-, is 2. *Source*: Ref. 5.

quinone pool. Thus, for every two UQH_2 that are oxidized releasing four protons, one UQH_2 is regenerated. The net result is the oxidation of one UQH_2 to UQ with the release of four protons, two from UQH_2 and two from the cytoplasm. The situation can be summarized as the following:

$$\text{P site: } 2UQH_2 + 2\text{cyt.c}_{ox} \rightarrow 2UQ + 2\text{cyt.c}_{red} + 4H^+_{out} + 2e^-$$
$$\text{N site: } UQ + 2e^- + 2H^+_{in} \rightarrow UQH_2$$
$$UQH_2 + 2\text{cyt.c}_{ox} + 2H^+_{in} \rightarrow UQ + 2\text{cyt.c}_{red} + 4H^+_{out}$$

In the pre–steady state, UQH_2 can be oxidized at the N site (by reversal) as well as the P site. The P site is inhibited by myxathiozol and by stigmatellin, whereas the N site is inhibited by antimycin. Therefore, a combination of inhibitors is required to completely inhibit the oxidation of UQH_2 in the pre–steady state.

Advantage of the Q Cycle Over the Q loop. Because the Q cycle translocates two protons per electron that flows to the terminal electron acceptor, it generates a larger proton current than the Q loop that translocates only one proton per electron. This can result in more ATP synthesis. Consider the situation where an ATP synthase requires the influx of three protons to make one ATP. If the transfer of an electron through the electron transport pathway resulted in the translocation of one proton, then one-third of an ATP could be made per electron. On the other hand, if electron transport resulted in the translocation of two protons per electron, then two-thirds of an ATP could be made per electron. In other words, the size of the proton current generated by respiration determines the upper value of the amount of ATP that can be made.

Pumps. Proton pumps are electron carrier proteins that couple electron transfer to the electrogenic translocation of protons (rather than electrons) through the membrane. It is not understood exactly how these pumps work at the molecular level, but perhaps the proton is passed from one amino acid side group that can be reversably protonated to another. A model for proton pumping is the light-driven proton pump bacteriorhodopsin (14). Proton extrusion accompanies the cytochrome aa_3 oxidase reaction when cytochrome c is oxidized by oxygen in mitochondria and some bacteria (Fig. 11). This can be observed by feeding electrons into the respiratory chain at the level of cytochrome c, from where they travel directly to the cytochrome aa_3 oxidase. The experimental procedure is to incubate the cells in lightly buffered anaerobic media with a reductant for cytochrome c and a permeant anion (e.g., SCN^-) or an exchangeable cation such as valinomycin plus K^+, so that protons can exit electroneutrally and accumulate in the medium. Changes in pH are measured with a pH electrode. Upon the addition of a pulse of oxygen, given as an air-saturated salt solution, a sharp, transient acidification of the medium occurs, and the H^+/O can be calculated. Because the only enzymatic reaction is the cytochrome aa_3 oxidase reaction, it can be concluded that the cytochrome aa_3 oxidase pumps protons out of the cell during the redox reaction. When similar experiments were done with *Paracoccus denitrificans* it was found that the *P. denitrificans* cytochrome aa_3 oxidase pumps protons with a stoichiometry of $1H^+/1e$ (15). Cytochrome oxidase pumping activity can also be demonstrated in proteoliposomes* made with purified cytochrome aa_3. Because the proton pumps move a positive charge across the membrane leaving behind a negative charge, a membrane potential, outside positive, develops. The membrane potential should be the same as when an electron moves inward since the proton and the electron carry the same charge (i.e., 1.6×10^{-19} C).

The mitochondrial NADH dehydrogenase complex (NADH ubiquinol oxidoreductase) translocates four protons per NADH oxidized (16). Two types of NADH ubiquinol oxidoreductases exist in bacteria (9–11,17). One of these, called NDH-1, is similar to the mitochondrial complex I in that it is a multisubunit enzyme complex (approximately 14 polypeptide subunits) consisting of FMN and FeS clusters and translocates protons across the membrane during NADH oxidation ($4H^+/2e^-$). The mechanism of proton translocation is not known. (The marine bacterium, *Vibrio alginolyticus*, has a sodium-translocating NDH [Na-NADH]). A second NADH/ubiquinol oxidoreductase, called NDH-2, is also present in bacteria. NDH-2 differs from NDH-1 in consisting of a single polypeptide and FAD and not being an energy-coupling site. In *E. coli*, NDH-1 and NDH-2 are simultaneously present (18). How *E. coli* regulates the partitioning of electrons between the two NAD dehydrogenases is not known, but clearly has important energetic consequences.

PATTERNS OF ELECTRON FLOW IN INDIVIDUAL BACTERIAL SPECIES

Although the major principles of electron transport as outlined above apply to nonphotosynthetic bacteria in general, several different patterns of electron flow exist in particular bacteria, often within the same bacterium grown under different conditions (19). The patterns of electron flow reflect the different sources of electrons and electron acceptors that are used by the bacteria. For example, bacteria may synthesize two or three different oxidases in the presence of air and several reductases anaerobically. In addition, certain dehydrogenases are made only anaerobically because they are part of an anaerobic respiratory chain. Furthermore, whereas electron donors such as NADH and $FADH_2$ generated in the cytoplasm are oxidized inside the cell, there are many instances of oxidations of other electron donors in the periplasm in Gram-negative bacteria. (Gram-negative bacteria are called thus because they fail to stain with the Gram's stain. They all possess a cell wall with common features and a periplasm. The periplasm is a compartment that exists external to the cytoplasm between the cell membrane and the cell wall. It is filled with a proteinacious liquid called periplasm, and the cell wall's rigid layer is called murein.) Examples of substances found in the growth medium or the environment that are oxidized in the periplasm include hydrogen gas, methane, methanol, methylamine, formate, perhaps ferrous ion, reduced inorganic sulfur, and elemental sulfur. In most of these instances periplasmic cytochromes c accept the electrons from the

*Proteoliposomes are lipoprotein vesicles made in the laboratory with purified proteins and phospholipids.

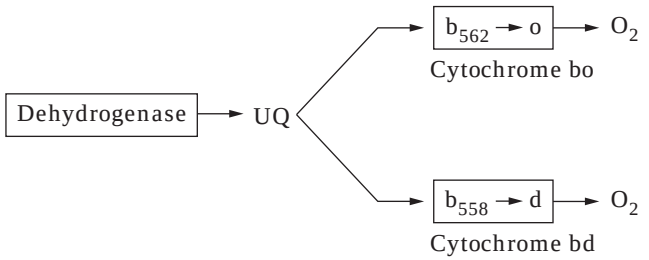


Figure 13. Aerobic respiratory chain in *E. coli.* A variety of dehydrogenase complexes are possible, such as NADH dehydrogenase, succinate dehydrogenase, lactate dehydrogenase, and glycerol-3-phosphate dehydrogenase. The chain branches at the level of ubiquinone (UQ) to two alternate quinol oxidases, cytochrome bo and bd. Cytochrome bd complex has a higher affinity for oxygen and is synthesized under low oxygen tensions where it becomes the major route to oxygen. Coupling sites, where protons are translocated to the outer surface, occur during NADH oxidation, catalyzed by NADH ubiquinone oxidoreductase (also called NDH-1), and quinol oxidation, catalyzed by cytochrome bo complex or cytochrome bd complex. Additionally, cytochrome bo is a proton pump. *E. coli* has a second NADH dehydrogenase, called NADH-2, that does not translocate protons. Therefore, the number of protons translocated per NADH oxidized can vary from two (NADH-2 and bd complex) to eight (NADH-1 and bo complex). *Source*: Ref. 5.

electron donor and transfer them to electron carriers in the cell membrane. The discussion that follows is not complete, but is meant to convey the diversity of electron transport systems found in bacteria.

Escherichia coli

E. coli is a Gram-negative heterotrophic facultative anaerobe. Heterotrophic means that it grows on organic carbon as a source of carbon. (Autotrophic means that it grows on carbon dioxide as its major source of carbon.) Facultative means that it can be grown aerobically using oxygen as an electron acceptor, anaerobically (e.g., using nitrate or fumarate as the electron acceptor), or anaerobically via fermentation. Fermentation means that it does not respire but instead reoxidizes the NADH produced during metabolism in the cytoplasm using organic compounds provided by metabolism as electron acceptors, and synthesizes ATP in the cytoplasm via a mechanism called substrate-level phosphorylation, which is different from oxidative phosphorylation. The bacteria adapt to their surroundings, and the electron transport system that the cells assemble reflects the electron acceptor that is available. When none is available, they ferment.

Aerobic Respiratory Chains. When grown aerobically, *E. coli* makes two different cytochrome oxidase complexes, cytochrome o complex and cytochrome d complex, resulting in a branched respiratory chain to oxygen (Fig. 13) (20). In these pathways, electrons flow from ubiquinol to the ter-

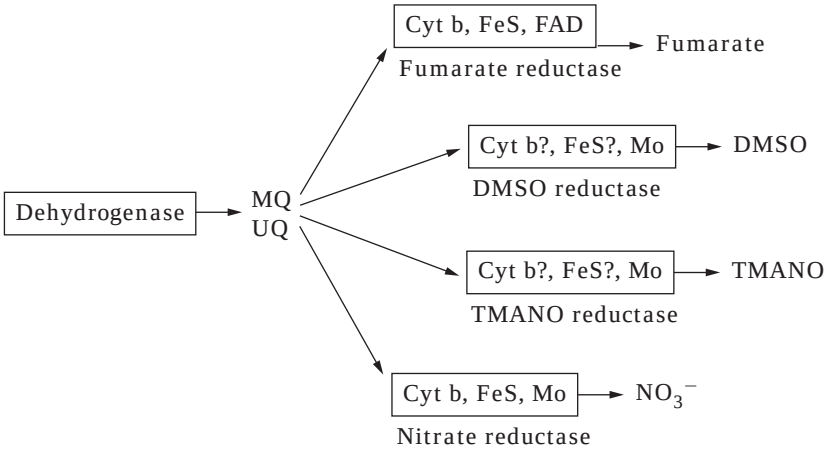


Figure 14. Anaerobic respiratory chains in *E. coli*. When oxygen is absent, *E. coli* synthesizes any one of several membrane-bound reductase complexes depending on the presence of the electron acceptors. Nitrate induces the synthesis of nitrate reductase and represses the synthesis of the other reductases. Menaquinone must be used to reduce some of the reductases, such as fumarate reductase, because it has a sufficiently low midpoint potential (−74 mV). Ubiquinol can reduce nitrate reductase. Each reductase may be a complex of several proteins and prosthetic groups through which the electrons travel to the terminal electron acceptor. The transfer of electrons from the dehydrogenases to the reductases results in the establishment of a proton potential. If the dehydrogenase has site one activity, there can theoretically be two coupling sites, one at the dehydrogenase step and one linked to quinol oxidation at the reductase step. Cyt b, cytochrome b; FeS, nonheme iron–sulfur protein; FAD, flavoprotein with flavin adenine dinucleotide as the prosthetic group; Mo, molybdenum; TMANO, trimethylamine N-oxide; DMSO, dimethylsulfoxide; MQ, menaquinone; UQ, ubiquinone. *Source*: Ref. 5.

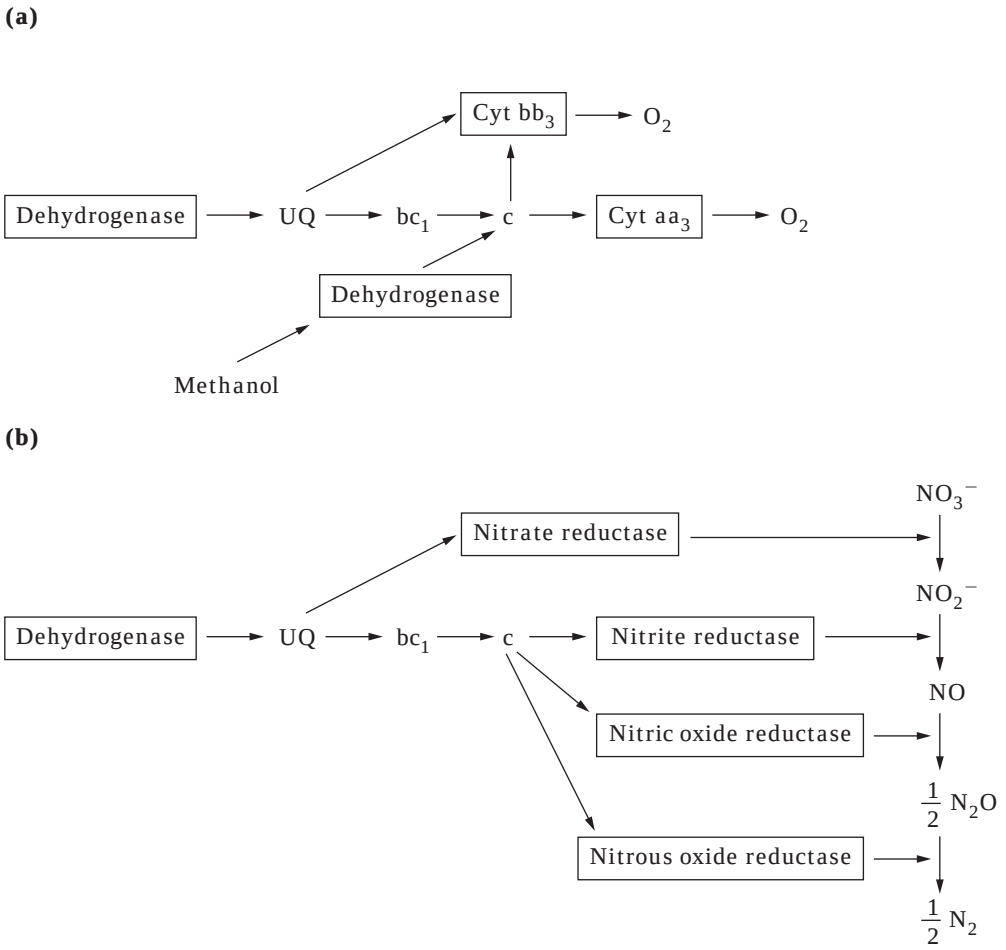


Figure 15. A model for electron transport pathways in *Paracoccus denitrificans*. (**a**) Aerobic. The pathway has two branch points. One branch is at the level of ubiquinone leading to one of two ubiquinol oxidases, that is, the cyt bc_1 complex or cyt bb_3. Both of these quinol oxidases are coupling sites. The bc_1 complex extrudes two protons per electron via the Q cycle. Cyt bb_3 is a proton pump and extrudes one proton per electron vectorially, whereas the second proton is extruded via a Q loop. A second branch point occurs at the level of the bc_1 complex. Electrons can flow either to cyt aa_3, which is a proton pump, or to cyt bb_3, which has been reported to pump protons under certain experimental conditions. NDH-1 is also a coupling site. (**b**) Anaerobic. When the bacteria are grown anaerobically using nitrate as the electron acceptor, the cytochrome aa_3 levels are very low, and the electrons travel from ubiquinone to nitrate reductase and also through the bc_1 complex to nitrite reductase, nitric oxide reductase, and nitrous oxide reductase. *Source*: Ref. 5.

minal oxidase complex, and therefore the oxidases are also called ubiquinol oxidases. The cytochrome o complex from *E. coli* is a proton pump, with a stoichiometry of $1H^+/e^-$. Cytochrome o complex is the predominant cytochrome oxidase when the oxygen levels are high. When the oxygen tensions are lowered, *E. coli* makes cytochrome oxidase d. The cytochrome d oxidase complex has a higher affinity for oxygen than does the cytochrome o complex, suggesting a rationale for why it is the dominant oxidase during growth at low oxygen tensions. Cytochrome d is not a proton pump.

The cytochrome o pathway is more energetically efficient (21). For every two electrons traveling from ubiquinol to oxygen via cytochrome o, four protons are translocated (i.e., two in the Q loop and two by the cytochrome o pump). Assuming a H^+/ATP of 3 (i.e., the ATP synthase translocates inwardly three protons per ATP made), then four-thirds or 1.33 ATPs can be made. Because cytochrome d is not a proton pump, only two-thirds or 0.67 ATPs can be made per two electrons during quinol oxidation via this route.

Physiological Significance of Alternate Electron Routes That Differ in the Number of Coupling Sites. Because the NDH-1 dehydrogenase may translocate as many as four protons per two e^-, whereas the NDH-2 dehydrogenase translocates 0, the number of protons translocated per NADH oxidized can theoretically vary from two (NDH-2 and cytochrome d complex) to eight (NDH-1 and cytochrome o complex). This means that the ATP yields per

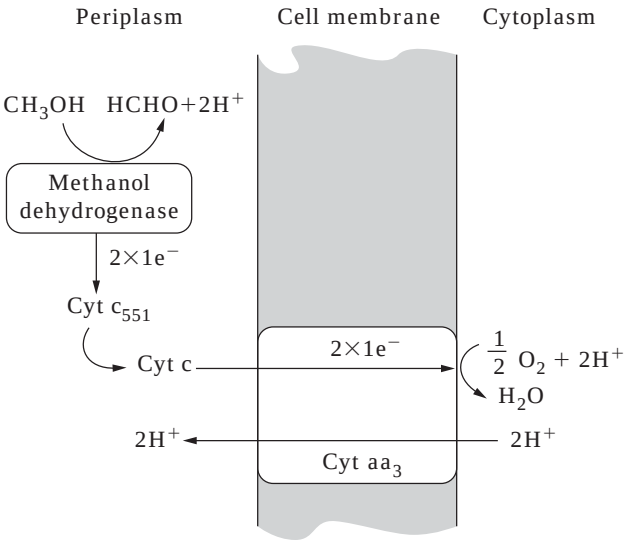


Figure 16. Oxidation of methanol by *P. denitrificans*. Methanol is oxidized to formaldehyde by a periplasmic methanol dehydrogenase. The electrons are transferred to periplasmic cytochromes c and to a membrane bound cytochrome aa_3, which is also a proton pump. A Δp is created due to the electrogenic influx of electrons and the electrogenic efflux of protons, accompanied by the release of protons in the periplasm (methanol oxidation) and uptake in the cytoplasm (oxygen reduction). *Source*: Ref. 5.

NADH oxidized can vary fourfold, from two-thirds or 0.67 to eight-thirds or 2.7. It also means that *E. coli* has great latitude in adjusting the Δp generated during respiration. Because a large Δp can drive electron transport in reverse and thus slow down oxidation of NADH, it may be an advantage to be able to direct electrons along alternate routes that bypass coupling sites and translocate fewer protons. This could ensure adequate rates of reoxidation of NADH so that the oxidative steps in metabolism can continue at a reasonable rate.

Anaerobic Respiratory Chains. In the absence of oxygen, *E. coli* can use either nitrate or fumarate as an electron acceptor. The nitrate is reduced to nitrite or further reduced to ammonia, and the fumarate is reduced to succinate, all of which are excreted into the medium* (22). The anaerobic respiratory chains consists of a dehydrogenase, a reductase, and a diffusible quinone to mediate the transfer of electrons between the dehydrogenase and the reductase (Fig. 14). There can be two coupling sites, one at the dehydrogenase step (site 1) and one at the quinone (i.e., a quinone loop). Anaerobic respiration occurs only in the absence of oxygen because oxygen represses the synthesis of the nitrate and fumarate reductases. When presented with nitrate anaerobically, only the nitrate reductase is made because nitrate represses the synthesis of the fumarate reductase. Thus, there exists a hierarchy of electron acceptors for *E. coli* (i.e., oxygen > nitrate > fumarate). This may reflect the energy yields of the electron transport pathways, because oxygen has the highest electrode potential and fumarate the lowest (Table 1).

*Fumarate is reduced to succinate using the membrane-bound enzyme fumarate reductase. Nitrate is reduced to nitrite using the membrane-bound nitrate reductase, which resembles the dissimilatory nitrate reductase in *P. denitrificans*. Both fumarate and nitrate reductases are coupled to the generation of a Δp, probably via a quinone loop. Nitrite is reduced to ammonia using an NADH-linked cytoplasmic enzyme (NADH-nitrite oxidoreductase). The ammonia is not assimilated, nor is energy conserved in the reaction. Nitrite is toxic to bacteria and probably the reduction to ammonia functions to reduce toxic levels.

Paracoccus denitrificans

P. denitrificans is a nonfermenting Gram-negative facultative anaerobe that can obtain energy from either aerobic respiration or nitrate respiration (13,23). The bacterium can grow heterotrophically on a wide variety of organic carbon sources or autotrophically on H_2 and CO_2 under anaerobic conditions using nitrate as the electron acceptor. It can be isolated from soil by anaerobic enrichment with media containing H_2 as the source of energy and electrons, Na_2CO_3 as the source of carbon, and nitrate as the electron acceptor. Electron transport in *P. denitrificans* receives a great deal of research attention because certain features closely resemble electron transport in mitochondria.

Aerobic Pathway. *P. denitrificans* differs from *E. coli* in that it has a bc_1 complex and a cytochrome aa_3 oxidase (cytochrome c oxidase), and in this way resembles mitochondria (Fig. 15). The composition of the electron transport chain varies with the growth conditions and other oxidases can be present.* The aerobic pathway is thought by some to branch at UQ or perhaps at cytochrome c.† It is not known to what extent the cytochrome bb_3 branch functions during aerobic growth. It has been suggested that the cytochrome bb_3 pathway serves as a Δp release valve. The idea is that under conditions where ATP consumption diminishes, electron flow through the two coupling sites, the bc_1 complex and cytochrome aa_3 complex, generates a Δp that can build to a point of retarding electron flow through the bc_1 complex. (Recall that electron flow can be driven in the reverse direction by the Δp.) Under these conditions, the electrons would be shunted to cytochrome o. In this way, the rates of oxidation of NADH and ubiquinol could be maintained.

P. denitrificans has three coupling sites. Site 1 is the NADH dehydrogenase (NADH-ubiquinone oxidoreductase) and this translocates two to three protons per two electrons. Site 2 is the bc_1 complex that translocates four protons per two electrons due to the Q cycle. Site 3 is the cytochrome aa_3 oxidase that pumps two protons per two electrons across the membrane. As described later, *P. denitrificans* also oxidizes methanol and in this case the electrons enter at cytochrome c, thus bypassing the bc_1 site.

*For example, when the bacteria are grown under low oxygen tensions, cytochrome aa_3 is decreased, and cytochrome o, cytochrome cd (nitrite reductase with oxidase activity), and cytochrome a_1 are increased.

†There are several cytochromes c (at least eight), some of which are in the membrane, and some in the periplasm.

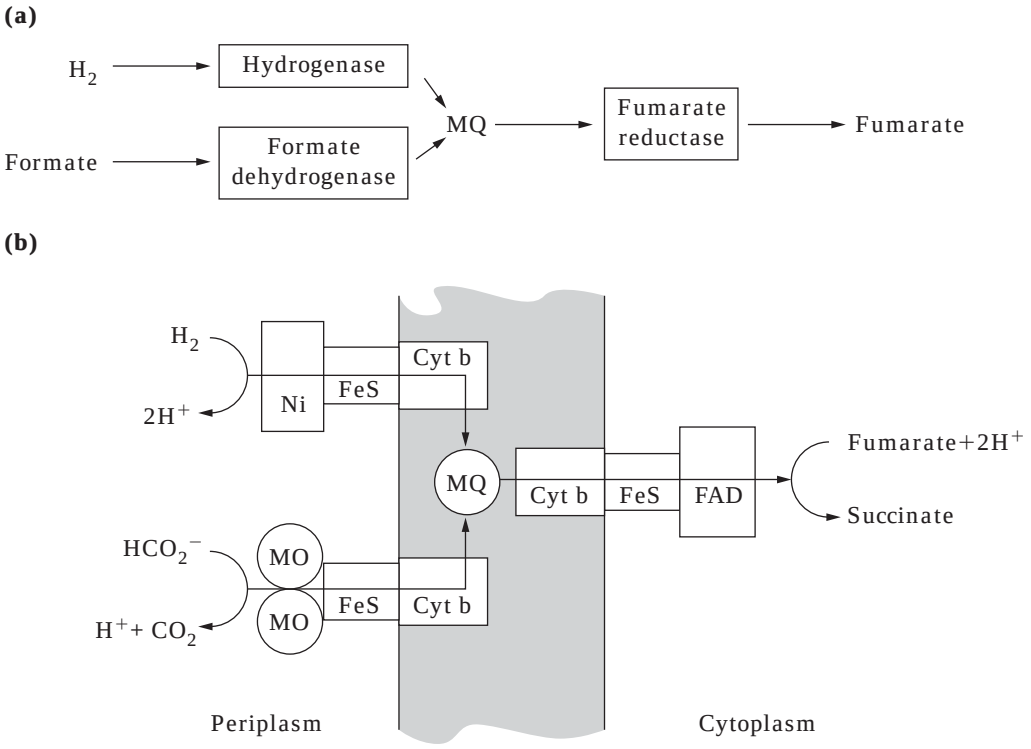


Figure 17. A model for the electron transport system of *W. succinogenes*. (**a**) Electrons flow from H_2 and formate through menaquinone (MQ) to fumarate reductase. (**b**) Illustration showing that the catalytic portions of hydrogenase and formate dehydrogenase are periplasmic, whereas fumarate reductase reduces fumarate on the cytoplasmic side. Electrons flow electrogenically to fumarate. A Δp is created because of electrogenic influx of electrons together with the release of protons in the periplasm, and their consumption in the cytoplasm. *Source*: Ref. 24.

Anaerobic Pathway. *P. denitrificans* can also grow anaerobically using nitrate as an electron acceptor, reducing it to nitrogen gas in a process called *denitrification* (Fig. 15) (25,26). During anaerobic growth on nitrate the electron transport chain is very different than during aerobic growth. The cells contain cytochrome o, nitrate reductase, nitrite reductase, nitric oxide reductase, and nitrous oxide reductase.* The levels of cytochrome aa_3 are very low. The nitrate (NO_3^-) is reduced to nitrite (NO_2^-) in a two-electron transfer via a membrane-bound nitrate reductase. The NO_2^- is reduced to nitric oxide (NO) in a one-electron transfer via a periplasmic* nitrite reductase. The NO is reduced to ½ mol of nitrous acid (½ N_2O) in a one-electron step via a membrane-bound nitric oxide reductase. And the 1/2 N_2O is reduced to 1/2 mol of nitrogen gas (1/2 N_2) in a one-electron step by a periplasmic nitrous oxide reductase. Thus, a total of five electrons flow in and out of the cell membrane through membraneous electron and periplasmic electron carriers from ubiquinol to the various reductases in order to reduce 1 mol of NO_3^- to 1/2 N_2. As shown in Figure 15, the electron transport pathway includes several branches to the individual reductases. The first branch site is at UQH_2, where electrons can flow either to nitrate reductase or to the bc_1 complex. Then there are three branches to the three other reductases after the bc_1 complex at the level of cytochrome c. In agreement with the model, electron flow to nitrate reductase is not sensitive to inhibitors of the bc_1 complex, whereas electron flow to the other reductases is sensitive to the inhibitors.

In *P. denitrificans*, *Pseudomonas aeruginosa*, and many other facultative anaerobes, both the synthesis and activ-

*Nitrate reductase is a membrane-bound molybdenum protein with three subunits, α, β, and γ. The γ subunit is a cytochrome b. A proposed route for electrons is from ubiquinol to the FeS centres of the β subunit, and from there to the molybdenum center at the active site in the α subunit. The α subunit then reduces the nitrate. The molybdenum is part of a cofactor (prosthetic group) called the molybdenum cofactor (Moco), which consists of molybdenum bound to a pterin. The Moco prosthetic group is in all molybdenum-containing proteins, except nitrogenase. (The molybdenum cofactor from nitrogenase contains nonheme iron and is called FeMo cofactor or FeMoco.) Nitrite reductase is a periplasmic enzyme that reduces nitrite to nitric oxide. Nitrite reductase from *P. denitrificans* has two identical subunits. These have hemes c and d_1. Nitric oxide reductase is a membrane-bound enzyme that reduces nitric oxide to nitrous oxide. It has heme b and heme c. Nitrous oxide reductase is a periplasmic enzyme that reduces nitrous oxide to nitrogen gas. It has two identical subunits containing Cu.

*The periplasm is an area in Gram-negative bacteria that exists between the cell membrane and the cell wall.

ity of the denitrifying enzymes are prevented by oxygen. However, in certain other facultative anaerobes, including *Comamonas* sp. and certain species of *Pseudomonas*, *Thiosphaera pantotropha*, and *Alcaligenes faecalis*, denitrifying enzymes are made and are active in the presence of oxygen (27). In these systems, both oxygen and nitrate are used simultaneously as electron acceptors, although aeration can significantly decrease the rate of nitrate reduction. The advantage of corespiration using both oxygen and nitrate is not obvious.

Periplasmic Oxidation of Methanol. Many Gram-negative bacteria oxidize substances in the periplasm and transfer the electrons to membrane-bound electron carriers, often via periplasmic cytochromes c. An example is *P. denitrificans*, which can grow aerobically on methanol (CH_3OH) by oxidizing it to formaldehyde (HCHO) and $2H^+$ with a periplasmic dehydrogenase, methanol dehydrogenase (Fig. 16).

$$CH_3OH \rightarrow HCHO + 2H^+$$

The electrons are transferred from the dehydrogenase to c-type cytochromes in the periplasm. The cytochromes c transfer the electrons to cytochrome aa_3 oxidase in the membrane, which reduces oxygen to water on the cytoplasmic surface. A Δp is established due to the inward flow of electrons and the outward pumping of protons by the cytochrome aa_3 oxidase, as well as the release of protons in the periplasm during methanol oxidation and consumption in the cytoplasm during oxygen reduction. Because the oxidation of methanol bypasses the bc_1 coupling site, the ATP yields are lower. In addition to methanol, the bacteria can grow on methylamine ($CH_3NH_3^+$), which is oxidized by methylamine dehydrogenase to formaldehyde, NH_4^+, and $2H^+$. The formaldehyde is eventually oxidized to CO_2, which is assimilated into cell material via the Calvin cycle. (The Calvin cycle is the pathway that green plants, cyanobacteria, and many bacteria use to grow on CO_2.)

$$CH_3NH_3^+ + H_2O \rightarrow HCHO + NH_4^+ + 2H^+$$

Methylamine dehydrogenase is also located in the periplasm and donates electrons via cytochromes c to cytochrome aa_3, bypassing the bc_1 complex.

Wolinella succinogenes

Fumarate respiration occurs in a wide range of bacteria growing anaerobically (24). This is probably because fumarate itself is formed from carbohydrates and protein during growth. *W. succinogenes* is a Gram-negative anaerobe isolated from the rumen. It can grow at the expense of H_2 or formate, both produced in the rumen by other bacteria. The electron transport pathway is shown in Figure 17. The active sites for both the hydrogenase and formate dehydrogenase are periplasmic, whereas the active site for the fumarate reductase is cytoplasmic. A Δp is produced when electrons travel electrogenically across the membrane from the electron donor on the outside to fumarate on the inside.

BIBLIOGRAPHY

1. P. Mitchell, *Nature (London)* **191**, 144–148 (1961).
2. M. Lubben, *Biochim. Biophys. Acta* **1229**, 1–22 (1995).
3. R.K. Thauer, K. Jungerman, and K. Decker, *Bacteriol. Rev.* **41**, 100–180 (1977).
4. D.E. Metzler, *Biochemistry: The Chemical Reactions of Living Cells*, Academic Press, New York, 1977.
5. D. White, *The Physiology and Biochemistry of Prokaryotes*, Oxford University Press, New York, 1995.
6. G. Gottschalk, *Bacterial Metabolism*, Springer-Verlag, New York, 1986.
7. T. Yagi, X. Xu, and A. Matsuno-Yagi, *Biochim. Biophys. Acta* **1101**, 181–183 (1992).
8. M. Finel, *J. Bioenerg. Biomemb.* **25**, 357–366 (1993).
9. T. Ohnishi, *J. Bioenerg. Biomemb.* **25**, 325–329 (1993).
10. T. Yagi, T. Yano, and A. Matsuno-Yagi, *J. Bioenerg. Biomemb.* **25**, 339–345 (1993).
11. M. Finel, *J. Bioenerg. Biomemb.* **25**, 357–366 (1993).
12. D.G. Nicholls and S.J. Ferguson, *Bioenergetics 2*, Academic Press, London, 1992.
13. B.L. Trumpower, *Microbiol. Rev.* **54**, 101–129 (1990).
14. M.P. Krebs and H.G. Khorana, *J. Bacteriol.* **175**, 1555–1560 (1993).
15. H.W. Van Verseveld, K. Krab, and A.H. Stouthamer, *Biochem. Biophys. Acta* **635**, 525–534 (1981).
16. P.C. Hinkle, M.A. Kumar, A. Resetar, and D.L. Harrs, *Biochemistry* **30**, 3576–3582 (1991).
17. T. Yagi, X. Xu, and A. Matsuno-Yagi, *Biochim. Biophys. Acta* **1101**, 181–183 (1992).
18. K. Matsushita, T. Ohnishi, and H.R. Kaback, *Biochemistry* **26**, 7732–7737 (1987).
19. V.D. Sled, T. Freidrich, H. Leif, H. Weiss, S.W. Meinhardt, Y. Fukumori, M.W. Calhoun, R.B. Gennis, and T. Ohnishi, *J. Bioenerg. Biomemb.* **25**, 347–356 (1993).
20. R.K. Poole and W.J. Ingledew, in F.C. Neidhardt, J.L. Ingraham, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger eds., *Cellular and Molecular Biology*, vol. 1, ASM Press, Washington, D.C., 1987.
21. M.W. Calhoun, K.L. Oden, R.B. Gennis, M.J. Teixeira de Mattos, and O.M. Neijssel, *J. Bacteriol.* **175**, 3020–3025 (1993).
22. B.C. Berks, S.J. Ferguson, J.W.B. Moir, and D.J. Richardson, *Biochem. Biophys. Acta* **1232**, 97–173 (1995).
23. H.W. Van Verseveld and A.H. Stouthamer, in A. Balows, H.G. Truper, M. Dworkin, W. Harder, and K.-H. Schleifer eds., *The Prokaryotes*, vol. III, 2nd ed., Springer-Verlag, Berlin, 1992.
24. A. Kroger, V. Geisler, E. Lemma, F. Theis, and R. Lenger, *Microbiology* **158**, 311–314 (1992).
25. S.J. Ferguson, *TIBS* **12**, 354–357 (1987).
26. B.C. Berks, S.J. Ferguson, J.W.B. Moir, and D.J. Richardson. *Biochem. Biophys. Acta* **1232**, 97–173 (1995).
27. D. Patureau, N. Bernet, and R. Moletta. *Curr. Microbiol.* **32**, 25–32 (1996).

See also ANAEROBES; BIOENERGETICS OF MICROBIAL GROWTH; ENERGY METABOLISM, MICROBIAL AND ANIMAL CELLS.

ELECTROPHORESIS OF PROTEINS AND NUCLEIC ACIDS

SEAN R. GALLAGHER
Motorola Phoenix Corporate Research Laboratories
Tempe, Arizona

KEY WORDS

Electrophoresis
Genomics
Isoelectric focusing
Microelectrophoresis
Polyacrylamide gel electrophoresis (PAGE)
Proteomics
Pulsed-field electrophoresis
SDS PAGE
Southern blotting
Zone electrophoresis

OUTLINE

INTRODUCTION

Electrophoresis is the movement in a liquid of a charged particle under the influence of an electric field. Because of its unique ability to separate charged biological macromolecules such as proteins and nucleic acids, electrophoresis is a mainstay in all fields of academic and industrial biological sciences. Over the last 50 years, a large body of work has been generated detailing the development of various electrophoretic analysis techniques. Among the applications currently in use, one-dimensional SDS polyacrylamide gel electrophoresis (SDS PAGE) is by far the most common for protein separations. Using the negatively charged detergent sodium dodecyl sulfate (SDS) to denature and give a constant negative charged-to-mass ratio to the proteins, SDS PAGE separates according to size and is frequently used to check purity or complexity of a sample. Isoelectric focusing (IEF), which separates either native or denatured proteins by their isoelectric point, provides a different criteria for assessing purity or complexity of a sample and is the critical first step or dimension in two-dimensional SDS PAGE (2D SDS PAGE). The second dimension is SDS PAGE, giving an IEF- and size-based separation for the protein sample. 2D SDS PAGE is a mainstay in the field of proteomics, which studies gene expression at the protein level.

RNA and DNA separations use both agarose for larger molecules and acrylamide for smaller-sized nucleic acid fragments. In addition to the purification, isolation, and sizing of nucleic acids, applications relying on electrophoresis include DNA sequencing and gene analysis. DNA sequencing, which uses thin rectangular acrylamide separation gels to separate short ($<$500 bases) DNA oligonucleotides, is performed with both manual and automated sequencing systems and is central to the numerous genome initiatives aimed at determining the DNA sequence of an organism's entire genome.

BACKGROUND

Electrophoresis has been studied for nearly two centuries (1). Starting in the early 1800s, researchers noted that colloidal clay particles are negatively charged and will move toward the positively charged electrode (anode). This observation was expanded by many to include positively charged particles moving toward negatively charged electrodes. Furthermore, the number of charges on the particle influenced the speed of movement. The foundation of modern electrophoresis started with the Ph.D. thesis of Arne Tiselius in 1930 (University of Uppsala, Uppsala, Sweden), in which he demonstrated moving zones of serum proteins in free solution. For the first time, protein components in serum could be separated out, giving the current designation of α-, β-, and γ-globulin (2,3). His work on electrophoresis and adsorption led to the Nobel Prize in Chemistry in 1948.

By the 1950s, several variations of electrophoresis were in use. The movement of bacteria, protozoa, blood cells, and even protein-coated quartz particles could be observed directly and studied in detail with a microscope. Although useful, the study of electrophoresis using the microscope was rapidly replaced by the technique developed by Tiselius, referred to as the moving boundary procedure. The moving boundary technique, instead of observing a single particle with a microscope, monitored the bulk movement of a protein-containing solution during electrophoresis. By placing the protein solution in a U-shaped tube and then overlaying with the appropriate buffer solution, the original protein solution will begin to migrate as a boundary into the buffer solution when subjected to an electric field. This boundary represents the leading or trailing edge of

the charged proteins as they move to the positive or the negative electrode.

The moving boundary procedure greatly expanded our knowledge of protein behavior in solution and permitted charged-based separation of complex mixtures and recovery of purified proteins. It is interesting to note that when the first separation of human serum was achieved with the moving boundary technique, the terminology for α-, β-, and γ-globulin was given by Tiselius to indicate the fastest, next fastest, and slowest globulin seen. By altering the pH of the buffered solutions, proteins would migrate either toward the positive or the negative electrode. By studying the movement as influenced by the pH of the buffer, the isoelectric points (or the pH where the net charge on the protein is 0) of many proteins were determined. The moving boundary procedure also brought other needed technology to bear on the problem of separations of proteins, including image analysis. Using schlieren scanning, the profile of the separated proteins could be visualized on a screen and then recorded.

The work by Tiselius was a triumph over the many physical factors that interfere with electrophoretic separations. Heat dissipation in particular continues to be a major problem in modern electrophoresis. Convective mixing is a severe problem in free solution and is caused by differential heating of the buffer or sample. When warmed, the solution expands and becomes more buoyant, leading to mixing and a loss of resolution. To minimize heat build up in the Tiselius apparatus, rectangular tubes were used to maximize the heat exchange from the U-shaped tube during electrophoresis, and the tube was further bathed in a water bath at 4 °C. At this temperature, water is most dense and is less prone to the effects of heat fluctuation.

However, moving molecules in free solution have a number of drawbacks, including mixing of zones that limit resolution. The addition of a solid support, such as agarose or acrylamide, stabilizes the separation against convective mixing, allowing higher-resolution separations. Additionally, the porosity and chemistry of the gel separation medium influences molecule migration and plays a significant role in defining separations. Electrophoresis on solid supports was originally referred to as zone electrophoresis, simply meaning that in a solid medium, proteins separate into individual component bands, or zones, rather than moving as a boundary in free solution. The zone technique in its original form used filter paper or silica gel as a solid support and was the earliest example of the modern techniques of electrophoresis that are currently being used today in laboratories for separating proteins and nucleic acids. Currently polyacrylamide or agarose is the solid support of choice for electrophoresis.

ELECTROPHORESIS THEORY

The force driving the charged molecule through the separation medium is the product of the charge Q on the protein or nucleic macro-ion and the potential gradient E across the electrophoresis chamber:

$$QE$$

Resistance from the surrounding medium counteracts the movement, with Newton's law indicating that the movement will experience an equal but opposite force. As a starting point to understand electrophoretic behavior of a molecule, Stokes's law, which approximates particle movement in a free solution, defines the resistance to electrophoresis as

$$f = 6\pi r v \eta$$

where f is the resistance of the medium to electrophoresis of the molecule, r represents radius of the particle (assumed to be a sphere), v is velocity of the particle, and η is viscosity of the liquid.

These two basic equations indicate that molecule charge, size, and shape, solution viscosity, and applied voltage gradient all play a role in electrophoresis. Although these equations are idealized, a number of general conclusions can be drawn. For a given size, more charge on the molecule will increase the separation speed, and for a given charge a larger molecule will migrate more slowly. Increasing the applied voltage gradient will speed separation as well. Several other factors also affect separation. Proteins and nucleic acid fragments are typically not spheres and can have complex structures. For example, a protein with a compact native structure will become extended upon full denaturation and reduction of internal disulfide bonds. The extended protein interacts more with the surrounding gel and buffer, experiencing more frictional resistance, and moves more slowly. Furthermore, separations usually take place in porous gels that sieve and interact with the molecules during electrophoresis. From a practical point of view, heat generation is among the most difficult problems to contend with. Increased temperatures decrease viscosity and the electrical resistance of the buffer and if not controlled will lead to localized heating, causing poor results and limited reproducibility. Typically buffer concentration is determined as a compromise between providing enough buffering capability to keep the pH of the solution stable and a low enough ionic strength to minimize heat generation at a voltage high enough to result in a rapid separation.

ELECTRICAL PARAMETERS, SAFETY, AND ELECTROPHORESIS

Ohm's law is important for understanding the practical day-to-day use of electrophoresis and can be stated in a variety of forms:

$$V = IR$$

$$I = V/R$$

$$P = VI$$

$$P = I^2R$$

where V is voltage in volts, I is current in amps, R is in ohms, and P is power in watts. A full appreciation of how the basic electrical parameters of electrophoresis interact is required to ensure safe operation and day-to-day reproducibility. A gel can be viewed as a resistor, and the power

supply as the voltage and current source. Most power supplies deliver constant voltage or constant current, and more specialized power supplies give constant power. The multiple outlets on the power supply are wired in parallel, so current is additive. That is, doubling the number of separation tanks will require doubling the current output of the power supply to give the same starting voltage.

Typical separation conditions for proteins and DNA in acrylamide and agarose require <500 V and <200 mA. Exceptions are found in isoelectric focusing where voltages up to 5,000 V are used, and DNA sequencing where up to 2,000 *V* and 80 W are used. Keeping careful records of the starting voltage, current, and power will quickly identify problems. For example, excess heating will cause the current to go up dramatically, while atypical electrical readings at the beginning indicate an incorrect buffer concentration or pH.

When assembling and using electrophoretic instrumentation, care must be exercised to avoid a potentially lethal shock. To start, the power supply should be turned off. Note that in all manipulations connecting and disconnecting the power supply, use one hand only. Using both hands will provide a path through the chest should a shock occur. Typically the red lead from the separation chamber is connected to the red input on the power supply, and the black lead is connected to the black input. Note that proteins can be positively charged under native conditions and that the polarity of the jacks may need to be reversed in order to have electrophoresis proceed in the correct direction. This is simply because manufactures have based their standards on separation of negatively charged proteins found in SDS PAGE (see later) and nucleic acids, which are intrinsically negative due to their phosphate backbone. Once connected, turn the power supply on and program or dial in the appropriate current or voltage. When the run is complete, turn the power supply down to zero before turning off, to ensure no excess charge is left in the capacitors in the power supply. Once the display reads zero, turn off and disconnect.

PORE SIZE AND SEPARATION MATERIAL

The choice of acrylamide or agarose gel depends on the application and the size the molecule and is governed in large part by the porosity of the two matrices. Acrylamide is composed of long chains of acrylamide monomers crosslinked with the bifunctional reagent *N*,*N*′-methylene bisacrylamide to form a mesh of pores (4–6) and the sieving network that separates larger from smaller molecules and provides size discrimination. Polyacrylamide is formed by polymerizing the acrylamide monomers and the crosslinking agent *N*,*N*′-methylene bisacrylamide with the addition of the initiator ammonium persulfate and the accelerator *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (TEMED). Polyacrylamide is typically used at concentrations ranging from 5 to 20 % total monomer (Table 1). Acrylamide concentrations below 3–4 % T at 2–3 % bisacrylamide produce gels that are very difficult to handle and still have pores too small for many nucleic acid applications.

Table 1. Separation Size Range for Proteins in Acrylamide

% Acrylamide	Size separation range of protein standard (Da)
5	25,000–200,000
10	14,000–200,000
15	14,000–66,000
20	14,000–45,000
5–20% Gradient	14,000–300,000

Source: Refs. 7 and 8.

In contrast, agarose is a galactan hydrocolloid from agar derived from marine algae (5). To prepare a separation gel, powdered agarose is dissolved by heat in the electrophoresis buffer and then allowed to solidify into a gel by cooling. Typical concentrations of agarose (0.75–1.5%) have very large pores that can fractionate much larger molecules than those in acrylamide. Agarose is typically used for nucleic acid separations and to a lesser extent isoelectric focusing (see later).

STANDARDIZATION AND CALIBRATION

To simplify gel-to-gel comparisons, common definitions are used to describe acrylamide gel concentration and molecule mobility (4–7):

R_r = Distance migrated by band/distance migrated by reference marker

%T = Total monomer (acrylamide + bisacrylamide) in g/100 mL final gel solution

%C = Bisacrylamide g/ (acrylamide + bisacrylamide) g × 100

Relative mobility, R_f, is a useful way to compare gels because it places the migration relative to an internal reference marker or tracking dye, usually bromophenol blue. Tracking dyes are also a useful visual indicator of gel performance, and any distortions in the separation can be seen before the separation is complete.

To determine molecule size, molecular weight or size standards are required for both protein and nucleic acid separations (Table 2 and Fig. 1). These are added to a lane on the gel along with the sample for comparison after separation and visualization, and are available from a variety of commercial companies.

VISUALIZATION

Separated proteins are typically visualized through equilibration with a protein-binding blue dye, Coomassie blue R 250, or by silver staining (7,10). Alternative techniques include prelabeling the proteins with a radioactive amino acid, and then analysis by exposure to X-ray film in a procedure called autoradiography. Postseparation analysis and labeling of the proteins with a fluorescent tag and then finally analyzing the proteins through fluorescent detec-

Table 2. Protein Molecular Weight Standards

Protein	Molecular weight (Da)
α-Lactalbumin	14,200
β-Lactoglobulin	18,400
Carbonic anhydrase	29,000
Ovalbumin	45,000
Bovine serum albumin	66,000
Phosphorylase b (rabbit muscle)	97,400
β-Galactosidase	116,000
RNA polymerase, *E. coli*	155,000
	165,000
Myosin, heavy chain (rabbit muscle)	205,000

Source: Ref. 7.

tion is also used. Nucleic acids are in general visualized with fluorescence after soaking the gel with the intercalating dye ethidium bromide (11), but they can also be stained with silver (12). Other procedures with nucleic acids make use of radioactive labeling and autoradiography.

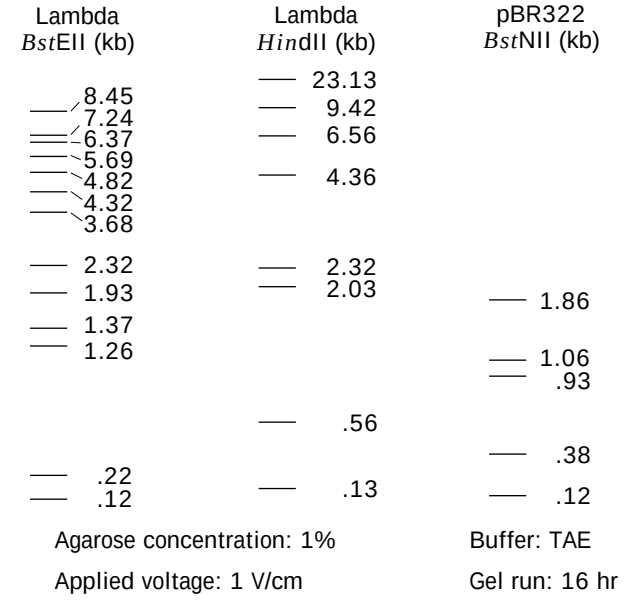


Figure 1. Migration pattern and fragment sizes of common DNA molecular weight markers. *Source*: Ref. 9.

ELECTROPHORESIS APPLICATIONS

Over the last 50 years, a large body of work has been generated detailing the development of various electrophoretic analysis techniques. Among the applications currently in use, one-dimensional SDS polyacrylamide gel electrophoresis (SDS PAGE) is by far the most common for protein separations. Using the negatively charged detergent sodium dodecyl sulfate (SDS) to denature and give a constant negative charge-to-mass ratio to the proteins, SDS PAGE separates according to size and is frequently used to check purity or complexity of a sample (Table 3). Isoelectric focusing (IEF), which separates either native or denatured proteins by their isoelectric point, provides a different criteria for assessing purity or complexity of a sample and is the critical first step, or dimension, in two-dimensional SDS PAGE (2D SDS PAGE). The second dimension is SDS PAGE, giving an IEF- and size-based separation for the protein sample. 2D SDS PAGE is a mainstay in the field of proteomics, which studies gene expression at the protein level.

RNA and DNA separations use both agarose for larger molecules and acrylamide for smaller-sized nucleic acid fragments. In addition to the purification, isolation, and sizing of nucleic acids, applications relying on electrophoresis include DNA sequencing and gene analysis. DNA sequencing, which uses thin rectangular acrylamide separation gels to separate short (<500 bases) DNA oligonucleotides, is performed with both manual and automated sequencing systems and is central to the numerous genome initiatives aimed at determining the DNA sequence of an organism's entire genome.

NUCLEIC ACID ELECTROPHORESIS

RNA and DNA electrophoretic separations are fundamental to modern molecular biology and biotechnology (Table 4). Depending on the application, the electrophoresis is performed in either a vertical or horizontal format. Polyacrylamide is used for smaller fragments of DNA in vertical (13) or horizontal (12) slab gels, whereas agarose is used for separating larger fragments of DNA in horizontal submarine separation chambers (9) (Tables 5 and 6). Due to the phosphate backbone, nucleic acids are negatively charged, with a constant charge-to-mass ratio. Although nucleic acid mobility is largely size dependent up to 30 kb,

Table 3. Key Applications of Protein Electrophoresis

Technique	Separation principle	Application
Native PAGE	Native charge, size, shape	Purification, size of native protein and protein complex
SDS PAGE	Size dependent; SDS imparts a negative charge to proteins, giving a constant charge-to-mass ratio	Size estimation, purity, purification, subunit composition, protein expression
IEF	Intrinsic charge with both native and denatured proteins	Purification, purity check, isoenzyme analysis
2D SDS PAGE	Isoelectric point in the first dimension, and size in the second	Protein expression, purification, posttranslational analysis, proteomics

Table 4. Key Applications of Nucleic Acid Electrophoresis

Technique	Separation principle	Application
Large fragment DNA and RNA separation	Size; due to the negatively charged phosphate backbone, nucleic acids have a constant charge-to-mass ratio	Purification, determination of purity and size, genetic analysis, preparation of DNA probes
Small fragment DNA and RNA separation	Size; see above	Purification, determination of purity and size, genetic analysis, preparation of DNA probes
DNA sequencing	Size of single-stranded oligonucleotide; denaturing gels and high temperature keep DNA single stranded	DNA sequence determination, mutation, and polymorphism analysis

Table 5. Separation Size Range for DNA in Agarose

Agarose (%)	Separation range of linear DNA (kbp)
0.5	30–1
1.0	10–0.5
1.5	3–0.2

Source: Ref. 9.

Table 6. Separation Size Range for DNA in Acrylamide

Acrylamide (%)	Separation range of linear DNA (bp)
3.5	100–1,000
5.0	100–500
8.0	60–400
12	50–200
20	5–100

Source: Ref. 13.

larger fragments of DNA require the use of pulsed-field gel electrophoresis for separation (see later). Furthermore, conformation of both DNA and RNA will influence mobility. The circular DNA plasmid pBR322, for example, migrates differently depending on its form (closed circular supercoiled Type I, nicked or open circular Type II, or straight linear Type III) even though all have the same molecular weight (9) (Fig. 2). The supercoiled form has a smaller hydrodynamic radius than the linear and will move more quickly through the gel.

Pulsed-Field Electrophoresis

DNA molecules larger than 20–30 kb migrate as a single species, or band, during separation and cannot be resolved with traditional agarose gel electrophoresis. Separation of very large DNA is, however, possible through the use of pulsed-field electrophoresis (14–18). DNA molecules are forced by the alternating electric fields to change their direction through the gel, and depending on the size of the DNA molecule, the percentage of agarose, and the frequency of the alternating electric field, separation of megabase-sized DNA is possible. Pulsed-field electrophoresis is dramatically affected by temperature and field strength, particularly for large molecules. A low field strength of <2 V/cm for DNA >5 Mb is required. Cool temperatures tend to enhance resolution but also prolong the length of the separation, and a compromise of room temperature is typically used. DNA larger than 50 kb is very difficult to handle in liquid without breaking during isolation and loading and is prepared in agarose plugs. The agarose plug is loaded into the sample well prior to separation.

There are two basic methods in common use for achieving pulse-field electrophoresis. One, field inversion gel electrophoresis (FIGE) works by simply inverting the polarity on the standard electrode found in most gel electrophoresis units. Specialized electronics are either incorporated into the power supply or placed between the power supply and the gel box to achieve the switching. Thus the orientation of the DNA during separation changes by 180 degrees (14). Separation with pulsed-field electrophoresis using an angle shallower than 180 degrees is typified by the technique called (contour clamp homogenous electric field) (CHEF), allowing a separation up to 6 Mb (16,19).

Southern Blotting

Once separated, DNA is frequently transferred out of the gel and onto a flat nitrocellulose or nylon membrane for further analysis. The technique is known as Southern blotting, after E. M. Southern (20,21), or capillary blotting. Briefly, DNA is denatured and nicked to enable it to migrate from the gel. Typically capillary blotting involves loading the gel on top of paper wicks that are suspended in a tray filled with transfer buffer. On top of the gel is the nitrocellulose or nylon membrane, a buffer-saturated sheet of blotter paper, dried blotter paper, and a weight, which draws the liquid from the wicks through the gel and onto the blotter paper. This pulls the DNA out of the gel and onto the membrane where it attaches. The DNA can then be analyzed with a radiolabeled or fluorescently tagged probe containing a complementary sequence to determine the presence or absence of the sequence in the sample.

Sequencing

Determination of the DNA sequence is a prerequisite to a number of key areas of research in the biological sciences (22–25). This includes development of genetic detection technologies, identification of disease-related genes, identification of genes that indicate a individual's predisposition toward either sensitivity or insensitivity to drugs, and the precise determination of similarities among genes from different organisms.

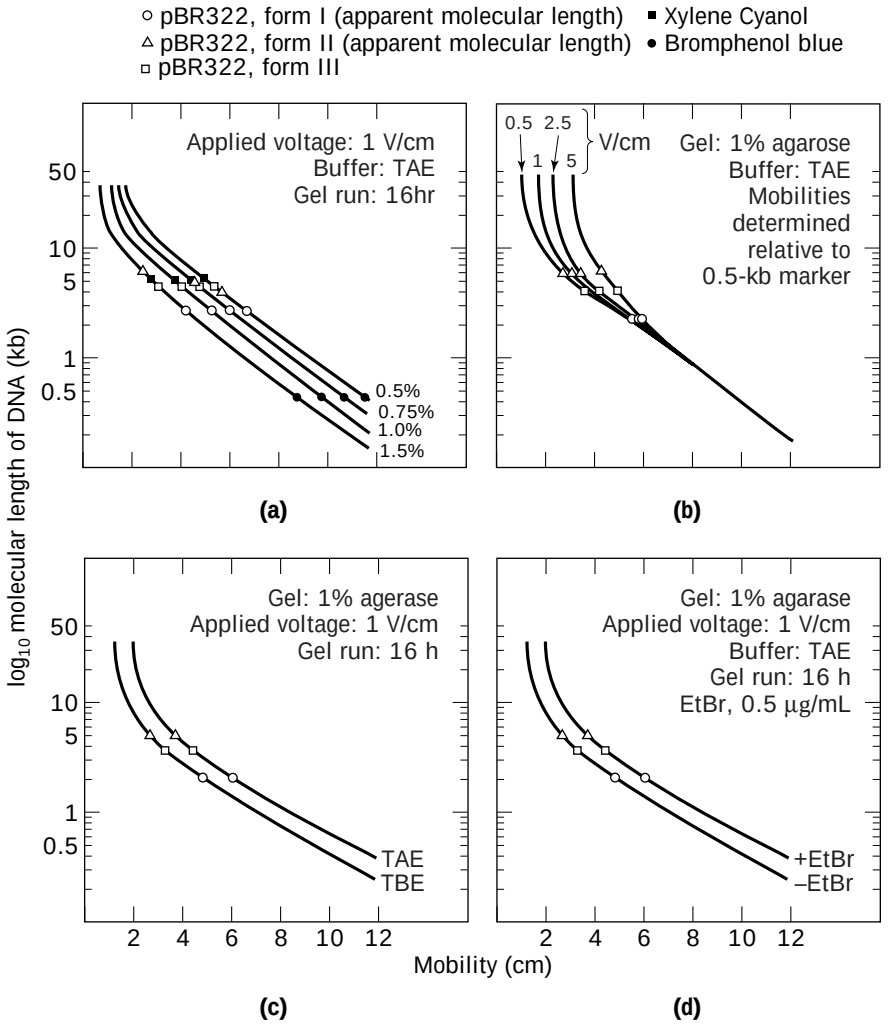


Figure 2. Effect of (**a**) agarose concentration, (**b**) applied voltage, (**c**) electrophoresis buffer, and (**d**) ethidium bromide on migration of DNA through agarose gels. *Source*: Used with Permission from Wiley.

DNA sequencing relies on electrophoretic separation of DNA fragments, oligonucleotides, that range from 10 to 500 bp in length. The separation, which is performed in thin (0.1- to 1.5-mm-thick) acrylamide gels at high temperature under denaturing conditions, yields 1-bp resolution. For most applications, dideoxyribonucleoside triphosphates (ddNTPs) chain termination sequencing is currently used (22,24). Briefly, the DNA sequence to be analyzed is annealed with a primer, which serves as the initiation point for the added DNA polymerase. With the appropriate fluorescently tagged or ^{35}S-labeled deoxyribonucleoside triphosphates (dNTPs) in the reaction, the polymerase proceeds to copy the template DNA. In each of the four separate dideoxy chain termination sequencing reactions (GATC), modified GATC nucleotides or chain terminators are added along with the unmodified nucleotides. As the chain terminators are incorporated into the growing chain, the reaction stops at A, T, C, or G. Using four separate enzymatic reactions, a set of DNA fragments with nucleotides that end with either an A, T, C, or G are placed in lanes labeled A, T, C, or G and separated by electrophoresis. The gel is then dried and placed against X-ray film to detect the radioactive nucleotides. The sequence is simply read off the resulting autoradiogram (Fig. 3). Alternatively the DNA fragments can be labeled with a fluorescently tagged nucleotide and then automatically read in an automated DNA sequencer that uses a laser to excite the fluorophore and a detector to read the fluorescence as the DNA fragment moves past a detector. For DNA larger than 300–500 bp, the DNA is generally broken into smaller fragments that can be individually sequenced.

PROTEIN ELECTROPHORESIS

Proteins are polyampholytes composed of amino acids containing a variety of ionizable acidic and basic groups that give the protein either a positive or a negative charge, depending on the pH. The acidic amino acids (aspartate and glutamate) and basic amino acids (lysine, histidine, and arginine) contribute along with other factors, such as the native structure, to the overall charge on the protein at a particular pH. Thus in acidic solutions amino groups be-

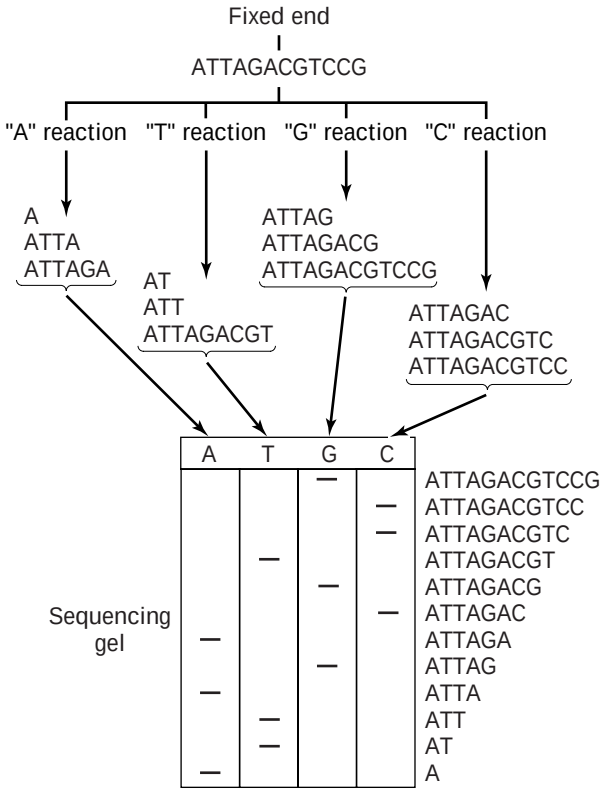


Figure 3. General strategy for DNA sequencing. To sequence a fragment of DNA, a set of radiolabeled single-stranded oligonucleotides is generated in four separate reactions. In each of the four reactions, the oligonucleotides have one fixed end and one end that terminates sequentially at each A, T, G, or C, respectively. The products of each reaction are fractionated by electrophoresis on adjacent lanes of a high-resolution polyacrylamide gel. After autoradiography, the DNA sequence can be "read" directly from the gel. *Source*: Ref. 22.

come ionized but carboxyl groups remain neutral, resulting in overall positive charge. In basic solutions, the carboxyl groups become ionized and the amino groups uncharged, creating an overall negative charge. The protein thus has an isoelectric point (pI), which is the pH at which the net charge on the protein is zero, reflecting the balance between the ionized positive and negative groups on the protein. The pI also depends on the conformation of the protein and any charged groups such as phosphorus that were added as a posttranslational modification. In the presence of denaturants or reductants that break disulfide linkages, proteins unfold and expose buried charged groups, causing an altered pI. When a protein is in its fully reduced and denatured state, both size and charge separation are reasonably accurate. SDS binds to a protein at a constant 1.4 g/g protein, completely masking the native charge, giving the protein (when fully denatured and reduced) a constant charged-to-mass ratio and making size-dependent separation possible (26).

Native Gel Electrophoresis

Under native or nondenaturing conditions, the native size, number of subunits, and isoforms can be studied in detail. Factors influencing native gel electrophoresis include native charge and shape and size of the protein (4,7,27,28). Ferguson plots (27), first applied to starch gels and later to polyacrylamide, gels provide a simple way to evaluate size and other characteristics of a protein under native conditions. First the protein sample is run under varied gel concentrations (typically 5, 10, 15, and 20%) and the slope of the log relative mobility (log R_f) versus acrylamide concentration (%T), called K_r, or the retardation coefficient, is plotted against the size of a series of standard proteins. By comparison to the standards, the approximate size of the unknown is determined.

Isoelectric focusing

Isoelectric focusing (IEF) separates proteins based on the intrinsic charge of the protein (6) and is performed by adding the protein sample to a gel containing a pH gradient. Chemically synthesized additives to the electrophoresis gel (ampholytes) generate a stable pH gradient from the basic (cathode or negative electrode) to the acidic (anode or positive electrode) end of the gel with an applied voltage. Ampholytes not only carry charge, but also are good buffers, containing a mix of acidic (carboxyl) and basic (amino) groups. Under the influence of an electric field, ampholytes migrate either as anions or cations toward the positive or negative electrode, respectively. Ampholytes gain or lose protons, and thus charge, until the isoelectric point is reached, where the molecule has a net charge of zero. Commercially available ampholytes have many species with slightly different isoelectric points that produce relatively smooth transitions in the pH gradient over a variety of ranges. The sample proteins move either toward the acid or the basic end, depending on the native charge of the protein. When the protein gets to the pH that represents its isoelectric point, or the point where the positive and negative charges on the protein balance to give a net charge of zero, the protein stops migrating and is considered focused. Initially, IEF used mobile ampholytes, but it is now possible to create a pH gradient grafted directly onto the acrylamide matrix by pouring a gradient of ionizable groups covalently attached to the acrylamide using Immobiline chemistry (12,29,30).

SDS PAGE

The most commonly used electrophoresis technique for protein separation is based on discontinuous SDS PAGE (7,31). Discontinuous or multiphasic systems separate proteins through a gel with two sections, the stacking and resolving regions. The stacking gel contains a pH 6.8 Tris-HCl buffer as well as a nonrestrictive large-pore acrylamide gel through which the sample moves. The chloride ions (called the leading ions) have an electrophoretic mobility greater than the sample proteins. Glycine ions (called the trailing ions) in the Tris-glycine electrode, or reservoir, buffer are less mobile than the sample proteins. The high mobility chloride ions leave a zone of lower conductivity between themselves and the migrating protein, causing a zone with a high voltage gradient. The proteins are then forced to "stack" in the zone between the leading and trailing ions. The proteins "destack" upon entering the resolv-

ing, or separating, gel. In the Tris-HCl pH 8.8 separating-gel buffer, glycine has a faster mobility and migrates past the proteins. The proteins then separate according to either molecular size in a denaturing SDS gel, or molecular shape, size, and charge in a nondenaturing gel.

To fully disassociate the subunits and denature the protein, the sample is mixed with SDS and a reductant, such as dithiothreitol, and heated to 100°C. SDS binds at 1.4 g/g protein, giving a constant charge-to-mass-ratio, leading to identical charge densities for the denatured proteins. Thus, the SDS-protein complexes separate according to size, not charge.

Optimal separation depends on the polyacrylamide concentration, with most proteins separated on gels containing from 5 to 15% acrylamide and 0.2 to 0.5% bisacrylamide (Table 1). The relationship between the relative mobility and log molecular weight is not linear over the entire separation range (Fig. 4). However, by using the linear portion of the curve, the molecular weight of an unknown protein is determined by comparison with known protein standards (Table 2). Gradient gels expand the linear range of separation over a much wider size range than single-concentration gels and are prepared by casting a gradient from high to low concentration of acrylamide. Proteins move into a constantly increasing concentration of acrylamide and smaller pore size, with the effect of sharpening the band and giving unparalleled resolution (Fig. 5) (5,7).

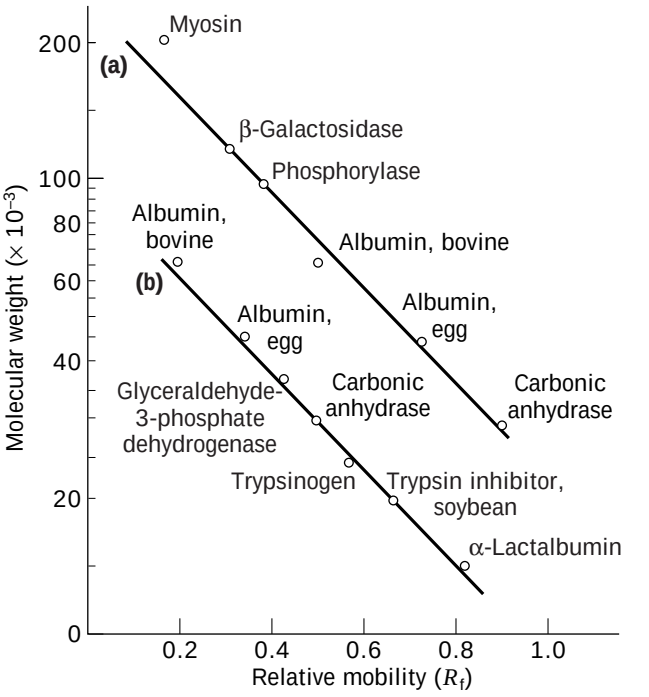


Figure 4. Typical calibration curves obtained with standard proteins separated by nongradient denaturing (SDS) discontinuous gel electrophoresis based on the method of Laemmli (31). (**a**) Gel with 7% polyacrylamide. (**b**) Gel with 11% polyacrylamide. *Source*: Redrawn with permission from Sigma.

2D SDS PAGE

Two-dimensional SDS polyacrylamide gel electrophoresis (2D SDS PAGE) is a fundamental technology for many areas of basic and applied biological research including molecular biology, protein chemistry, clinical chemistry, and toxicology (32–38). In 2D SDS PAGE, proteins are first separated by isoelectric point via IEF and then by size via SDS PAGE (Fig. 6). 2D SDS PAGE simultaneously identifies the properties such as size, isoelectric point, and post-translational modifications of thousands proteins. The separation yields a high-purity product that is amenable to mass spectroscopy analysis, providing further details about the protein structure.

With the publication of work by O'Farrell (37), the basic form of current two-dimensional electrophoresis took shape. There are many applications of 2D SDS PAGE analysis in biology (32), including understanding complex interactions between expressed proteins in a cell and the environment; identifying proteins that are immunoactive; toxicology testing (where exposure to a particular drug will increase or decrease the level of one or many proteins), which is a particularly critical area where gene analysis alone does not suffice; and quality control separation of proteins in various products for both agriculture and human consumption.

2D SDS PAGE is performed by first solubilizing and denaturing the protein using nonionic detergents, chaotropic agents, and a reductant. Typically these are NP40, urea, and DTT. This basic sample buffer, however, does not work for all samples, and many variations on solubilization have been developed to handle bacterial proteins, serum samples, membrane proteins, and plant samples. The proteins

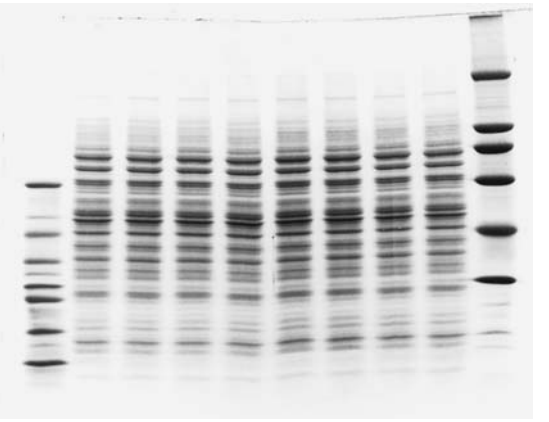


Figure 5. Separation of soluble proteins by 5.1 to 20.5%T polyacrylamide gradient SDS PAGE. Approximately 30 μL of 1 × SDS sample buffer containing 30 μg of Alaskan pea (*Pisum sativum*) soluble proteins was loaded in wells of a 14 × 14-cm, 0.75-mm-thick gel. Standard proteins were included in the outside lanes. The gel was run at 4 mA for ~15 h. *Source*: Used with permission from Wiley.

are separated in the first dimension by focusing either in thin strips horizontally, or in vertical glass tubes. After the first-dimension separation, the proteins are then separated with discontinuous SDS PAGE in vertical or horizontal (12,29) systems. In the first dimension, the acrylamide matrix is porous to keep sieving effects to a minimum,

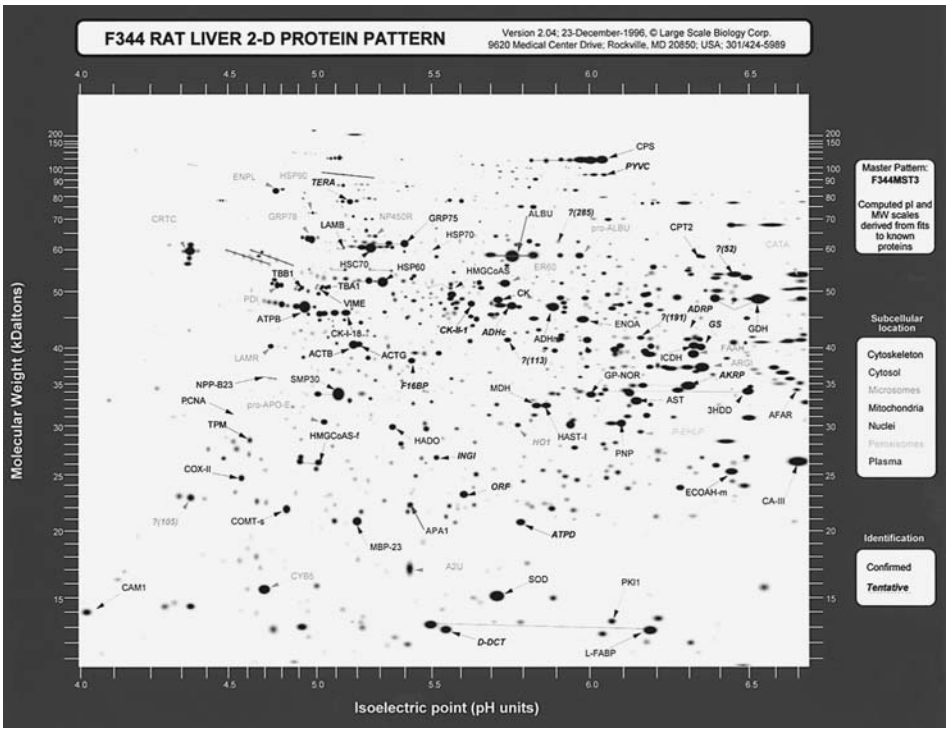


Figure 6. 2D protein pattern of rat liver. The proteins are arrayed from acidic to basic isoelectric points (right to left) and from large to small size (top to bottom). *Source*: Used with permission from Large Scale Biology, Rockville, Maryland.

whereas the second dimension is performed under conditions that emphasize the size separation by using single concentration or pore-gradient gel electrophoresis. 2D SDS PAGE is the only technique currently in use that is capable of parallel separation of literally thousands of proteins in a single separation run. The resulting gel represents an array of proteins being currently expressed in one particular cell type at that particular moment.

Although still popular, mobile carrier ampholytes present a number of challenges for researchers. Although the buffering capacity of ampholytes exceeds that of proteins, large quantities of proteins tend to interfere with the generation of the pH gradient. In addition, synthesis of the ampholytes is complex and not as reproducible as one would like from batch to batch. Furthermore, carrier ampholytes have a tendency to drift, shifting the pH gradient during separation and contributing to uncertainty about the protein position in the final separation. A new technology, immobilized pH gradient (IPG) gels (12,29,30, 36,39), helped considerably to enhance the technology of 2D SDS PAGE and make it more reproducible. With IPG gels, the pH gradient is covalently immobilized in the gel matrix during the preparation of the gel itself. When separation of the proteins occurs under electrophoresis, the pH gradient is stable. Another tremendous advantage is that, in contrast to carrier ampholytes, large quantities of proteins can be loaded without significant disruption of the separation pattern, permitting detailed analysis of the spots subsequent to separation by techniques such as mass spectroscopy. The extreme resolution capable with 2D SDS PAGE permits the identification of proteins not only by their position on the map, but by analysis of co- and posttranslational modifications (40), including phosphorylation, glycosylation, proteolysis, and prenylation.

NEW APPLICATIONS

Genomics and Proteomics

First used in 1994, the term *proteome* is defined as the study of the expressed proteins in a cell and is derived from *prote*ins expressed by the gen*ome* (32). Currently, the only way to separate and visualize all proteins in a cell is with 2D SDS PAGE. Proteomics is critical to modern cell and molecular biology because, in a single separation, number, abundance, and posttranslational modifications are apparent. Furthermore, determination of a partial sequence with mass spectrometry is now a standard technique that gives an accurate identification of the protein, or if no record of the protein is apparent, a DNA probe for identifying and isolating the gene can be developed.

The connection between genomics, drug discovery, and proteomics derives from the observation that mRNA expression and protein abundance representing mRNA do not show a good correspondence. Furthermore, because most drug targets are proteins, the analysis of drug effects must start with the expressed protein complement. However, to fully understand the protein target and associated gene family, the gene must also be identified.

Microelectrophoresis and Microfluidics

Current research in electrophoresis is directed toward automating manually intensive steps such as sample preparation, and shrinking the size of electrophoretic analysis devices to increase throughput, lower use of reagents, and generally automate electrophoretic analysis (41).

With the development of a large body of genetic analysis information and techniques, a number of diseases have

been distinguished by DNA analysis. Developing assays for the diagnosis of genetic disease is a primary driver in instrumentation development. Current electrophoretic instrumentation found in research laboratories is largely manual, requiring operators to run gels either a few at a time in parallel or in a serial operation. The exception is automated DNA sequencing apparatus. However, this is also large and requires a high level of operator expertise. Recently, there has been a strong drive to not only simplify electrophoretic analysis, but also lower the cost and increase the number of samples that can be run in a given day. In order to accomplish this, researchers have turned toward microminiaturization using semiconductor microfabrication technology. Although not commercially available, microminiaturized devices for DNA and protein separation are available at the research level and have a number of attractive features. Sometimes referred to as lab-on-chip devices (LOC), these microminiaturized laboratories and electrophoresis chambers have the capabilities of not only preparing the sample but mixing the sample, amplifying the genetic material through PCR if necessary, and then providing electrophoretic separation and subsequent analysis. In fact, separation of DNA for sequencing purposes on chips has been demonstrated. As advances in microminiaturization continue, the use of microfluidics and the integration of the movement of small amounts of liquid with electrophoretic separation and detection will become increasingly important for diagnostics.

BIBLIOGRAPHY

1. G.W. Gray, *Sci. Am.* **185**, 45–53 (1951).
2. A. Tiselius, *Biochem. J.* **31**, 1464–1477 (1937).
3. A. Tiselius, *Trans. Faraday Soc.* **33**, 524 (1937).
4. A.T. Andrews, *Electrophoresis: Theory, Techniques and Biochemical and Clinical Applications*, 2nd ed., Oxford University Press, New York, 1986.
5. B.D. Hames, in B.D. Hames, D. Rickwood eds., *Gel Electrophoresis of Proteins*, Oxford University Press, Oxford, 1990, pp. 1–147.
6. P.G. Righetti, *Isoelectric Focusing: Theory, Methodology and Applications*, Elsevier Biomedical, Amsterdam, 1983.
7. S.R. Gallagher, in F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl eds., *Current Protocols in Molecular Biology*, Wiley, New York, 1997, Unit 10.2.
8. G. Smejkal, S. Gallagher, *Biotechniques* **16**, 196–198, 200–202 (1994).
9. D. Voytas, in F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl eds., *Current Protocols in Molecular Biology*, Wiley, New York, 1997, Unit 2.5.
10. J.H. Morrissey, *Anal. Biochem.* **117**, 307–310 (1981).
11. F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl eds., *Current Protocols in Molecular Biology*, Wiley, New York, 1997.
12. R. Westermeier, *Electrophoresis in Practice*, VCH, Weinheim, 1997.
13. J. Chory, A.S. Baldwin, Jr., in F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl eds., *Current Protocols in Molecular Biology*, Wiley, New York, 1997, Unit 2.7.
14. G.F. Carle, M. Frank, M.V. Olson, *Science* **232**, 65–68 (1986).
15. B.W. Birren, E. Lai, S.M. Clark, L. Hood, M.I. Simon, *Nucleic Acids Res.* **16**, 7563–7582 (1988).
16. G. Chu, *Electrophoresis* **10**, 290–295 (1989).
17. S.M. Clark, E. Lai, B.W. Birren, L. Hood, *Science* **241**, 1203–1205 (1988).
18. D.C. Schwartz, C.R. Cantor, *Cell* **37**, 67–75 (1984).
19. G. Chu, D. Vollrath, R.W. Davis, *Science* **234**, 1582–1585 (1986).
20. E.M. Southern, *J. Mol. Biol.* **98**, 503–517 (1975).
21. T. Brown, in F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl eds., *Current Protocols in Molecular Biology*, Wiley, New York, 1997, Unit 2.9.
22. B.E. Slatko, R.L. Eckert, L.M. Albright, F.M. Ausubel, in F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl eds., *Current Protocols in Molecular Biology*, Wiley, New York, 1997, Unit 7.1.
23. A.M. Maxam, W. Gilbert, *Methods Enzymol.* **65**, 499–560 (1980).
24. F. Sanger, S. Nicklen, A.R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
25. F. Sanger, A.R. Coulson, B.G. Barrell, A.J. Smith, B.A. Roe, *J. Mol. Biol.* **143**, 161–178 (1980).
26. W.W. Fish, J.A. Reynolds, C. Tanford, *J. Biol. Chem.* **245**, 5166–5168 (1970).
27. K.A. Ferguson, *Metabolism* **13**, 985–1002 (1964).
28. J.L. Hedrick, A.J. Smith, *Arch. Biochem. Biophys.* **126**, 155–164 (1968).
29. A. Gorg, *Biochem. Soc. Trans.* **21**, 130–132 (1993).
30. B. Bjellqvist, K. Ek, P.G. Righetti, E. Gianazza, A. Gorg, R. Westermeier, and W. Postel, *J. Biochem. Biophys. Methods* **6**, 317–339 (1982).
31. U.K. Laemmli, *Nature* **227**, 680–685 (1970).
32. M.R. Wilkins, K.L. Williams, R.D. Appel, D.F. Hochstrasser eds., *Proteome Research: New Frontiers in Functional Genomics*, Springer-Verlag, Heidelberg, 1997.
33. N.L. Anderson, N.G. Anderson, *Anal. Biochem.* **85**, 341–354 (1978).
34. N.G. Anderson, N.L. Anderson, *Anal. Biochem.* **85**, 331–340 (1978).
35. N.G. Anderson, N.L. Anderson, *Electrophoresis* **17**, 443–453 (1996).
36. A. Gorg, G. Boguth, C. Obermaier, A. Posch, W. Weiss, *Electrophoresis* **16**, 1079–1086 (1995).
37. P.H. O'Farrell, *J. Biol. Chem.* **250**, 4007–4021 (1975).
38. P.Z. O'Farrell, H.M. Goodman, P.H. O'Farrell, *Cell* **12**, 1133–1141 (1977).
39. A. Gorg, G. Boguth, C. Obermaier, W. Weiss, *Electrophoresis* **19**, 1516–1519 (1998).
40. A.A. Gooley, N.H. Packer, in M.R. Wilkins, K.L. Williams, R.D. Appel, D.F. Hochstrasser eds., *Proteome Research: New Frontiers in Functional Genomics*, Springer-Verlag, Heidelberg, 1997, pp. 65–91.
41. C.H. Mastrangero, M.A. Burns, D.T. Burke, *Proc. IEEE* **86**, 1769–1787 (1998).

ELECTROPHORESIS, PROTEINS, BATCH AND CONTINUOUS

CORNELIUS F. IVORY
Washington State University
Pullman, Washington

KEY WORDS

Analytical
Batch
Capillary
Continuous
Dispersion
Electrophoresis
PAGE
Preparative

OUTLINE

INTRODUCTION

This article is intended as a guide for scientists and engineers who are considering using electrophoresis to process proteins or other biologicals for the biotechnology industry at bench or larger scales. An overview of several types of equipment, a description of what each piece of equipment is designed for, and the problems the end user should anticipate before operating that equipment are discussed.

The discussion of each class of apparatus focuses on design principles rather than experimental protocols, emphasizing the engineering aspects of each chamber rather than the details of operation. This type of treatment provides insight into the general operating characteristics of each device and is intended to aid the reader in *(1)* choosing an appropriate instrument for a given separation and/or *(2)* setting operating conditions (e.g., column length, sample volume, or power levels) that avoid excessive dispersion (e.g., band smearing).

For the purposes of illustration, it is assumed that the buffer has the kinematic and electrical properties of pure water at 10 °C, as given in Table 1, with the exception of its electrical conductivity, σ_e, which is determined from electrolyte composition. Most of the values used are sufficiently accurate for purposes of illustration. However, the mobilities and diffusion coefficients found in practice may be quite different from those listed, since they depend primarily on the solutes processed.

The free-solution values of the mobility and diffusion coefficient correspond roughly to those of serum albumin and tend to be high compared with other proteins. The values of the mobility and diffusion coefficient given for gels are reduced from the free-solution values by a factor of 5 in order to exaggerate the retardation effects found in gels. In practice, the retardation factor depends on the composition and degree of cross-linking in the gel, and can be substantially larger or smaller than 5. The ratio of the electrophoretic mobility to the diffusion coefficient in the gel and in free solution, a parameter that frequently appears in theoretical models of electrophoresis, has been kept constant in these calculations.

Band Broadening (Dispersive) Effects

Diffusion. Before describing the various apparatuses, it is worthwhile to briefly review those phenomena that can lead to deterioration of a separation, that is, band broadening or dispersion (which is a general term for these types of effects). The most ubiquitous dispersive phenomenon is diffusion; one should always take precautions to minimize the effects of diffusion, especially when other band-broadening effects are small.

An important characteristic of diffusion-based dispersion is that it increases band width in proportion to the square root of the residence time according to the formula

$$\chi^2_{\text{diffusion}} = 2D_m t \text{ or } \sigma^2_{\text{diffusion}} = \frac{2D_m Z}{\langle u \rangle^3} \qquad (1)$$

where the spatial variance χ^2 represents band spreading in space, as would be observed by scanning bands on a gel that had been run for time t. The temporal variance σ^2

Table 1. Physical Constants Used in Illustrations

Thermal conductivity	k_{gel} (W/cm °C)	0.00578
Thermal conductivity	k_{glass} (W/cm °C)	0.011
Thermal conductivity coefficient	λ[equation 4] (°C^{-1})	0.02874
Thermal viscosity coefficient	α[equation 30] (°C^{-1})	−0.02874
Thermal expansion coefficient	β[equation 40] (°C^{-1})	0.000088
Electrophoretic mobility in gel	μ (cm^2/V s) (gel)	0.00002
Electrophoretic mobility in free solution	μ (cm^2/V s) (buffer)	0.0001
Diffusion coefficient in gel	D_{gel} (cm^2/s) (gel)	1×10^{-7}
Diffusion coefficient in free solution	D_m (cm^2/s) (buffer)	5×10^{-7}

represents band spreading in time, as would be observed as bands migrate at a mean electrophoretic velocity $\langle u \rangle$ past a fixed detector located at mean position $z = Z$ relative to a starting point $z = 0$. The width of a solute band in space can be approximated as 4χ and its duration in time as 4σ. When several different mechanisms contribute to dispersion, their net effect is usually taken into account by summing the appropriate variances, in either space or time. This assumes that the individual effects are independent and is often a reasonable approximation.

Thermal Dispersion. The next major source of dispersion comes from the effect of spatial temperature variations on the electrophoretic mobility. All forms of electrophoresis require that a current be passed through a resistive medium, so electrical energy is necessarily converted to thermal energy in the medium. To efficiently remove the joule heat generated by the current, one spatial dimension of the chamber is generally kept thin, and heat is removed through the bounding surfaces along this axis.

The mobilities of charged solutes and aqueous electrolytes usually vary inversely with the buffer viscosity, which is itself quite similar to water in its temperature dependence. The viscosity of water decreases by about 25% with a 10 °C rise in temperature, so the warmer parts of a solute band migrate faster than the cooler parts and often take on parabolic or crescent shapes. This deformation results in a broadening of the fractionated sample, which usually overwhelms diffusive spreading, especially at the field strengths used in high-performance equipment (e.g., >200 V/cm).

Electroosmosis. Electroosmosis can be a severe problem in free-flow systems (i.e., those with no anticonvective medium). It arises when mobile counterions in the double layer adjacent to a charged wall are moved parallel to the wall surface by an electric field (1). These ions drag solvent with them, setting up a bulk fluid motion that can smear solute bands. Although it is difficult to eliminate electroosmosis entirely, for example, by using surface coatings or additives, steps may be taken to control it or even to take advantage of it (2). In homogeneous gels, electroosmosis plays a lesser role in band distortion but may act to displace purified bands, thereby giving inaccurate estimates of solute mobilities. In capillary zone electrophoresis at alkaline pH, electroosmosis typically provides the convective flow used to transport solutes through the capillary and does not contribute significantly to dispersion.

Other problems associated with Kohlrausch (3) effects and conductivity gradients (4,5) can have a major impact on band broadening. The former gives rise to stable band broadening or narrowing due to differences in mobilities between the target proteins and the electrolyte ions, but without inducing bulk flows. The latter may generate stable or unstable flows that can deform solute bands in a number of ways. The instability associated with conductivity gradients is particularly deleterious. As a rule of thumb, to avoid these problems with native zone electrophoresis, the electrolyte composition and conductivity should closely match those of the sample and should provide adequate buffering; the sample proteins should not be too heavily concentrated and the electric power applied should not exceed the effective cooling capacity of the column.

Natural Convection. Although problems with thermally induced natural convection (6) have long hampered scale-up of the thin-film continuous-flow electrophoresis device, these problems are largely mitigated in recycling electrophoresis equipment (7,8) and the membrane-based chambers developed by Faupel et al. (9) and Horvath et al. (10). This is not meant to imply that natural convection is a trivial problem but rather that dispersion due to natural convection can be ameliorated by innovative and careful design.

Conductive-Dielectric Flows. Although flows induced by gradients in the conductivity and dielectric constant are still largely a curiosity in electrophoretic separations, they will become major barriers to device development as electric fields are increased to boost chamber performance. Figure 1a shows the cusplike distortion of conductive dye streaks containing 100 mM NaCl in a clear solution containing 10 mM NaCl with the flow stopped and the streamlines exposed to a nominal electric field of about 26 V/cm. When the salt is replaced with different concentrations of ethanol, the induced flows have essentially the same morphology, suggesting a common mechanism for both flows. Figure 1b shows both cusps and putative roll cells in a similar experiment, but with vertical flow from top to bottom. One important feature of this instability is that conductive dispersion can be explosive, smearing finely resolved protein bands over a distance of several centimeters along the electric field lines in a few seconds.

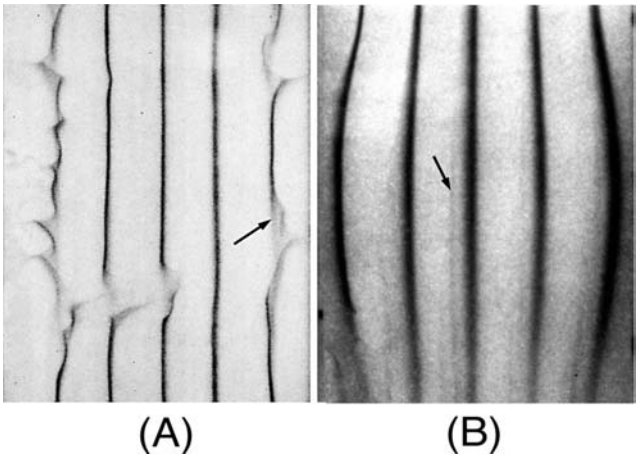


Figure 1. (a) Conductive–dielectric instability in a stop-flow experiment in a Plexiglas slit, 2 mm thick by 10 cm wide. Five streaks of water containing 10 mM NaCl and traces of bromophenol blue dye were injected into a moving stream of distilled water flowing between a pair of electrodes. The flow was stopped and a nominal field of 26 V/cm applied between the electrodes. The cusplike streaks are a hallmark of this instability, but roll cells *(arrow)* are also seen. **(b)** In water flowing at 1 cm/s the instability appears at a higher voltage and produces primarily roll cells *(arrow)*. Some cusps are visible in the streakline on the left.

A stable form of these flows has been described and analyzed by Roberts and coworkers (4,11) and the unstable form recently analyzed by Baygents and Baldessari (12). These papers support the empirical rule of thumb that bands in zone electrophoresis or isoelectric focusing should have conductivities as close to those of the surrounding buffer as possible to avoid conductivity gradients. These flows can be stabilized by applied shear stresses, but at fields above 400 V/cm even shear rates in excess of about 100 s^{-1} are only marginally effective at damping these flows.

Performance Metrics

Number of Theoretical Plates. In order to facilitate comparison of chamber designs and separation techniques, it is useful to have a yardstick or metric that can be universally applied not only to electrophoresis but also to other techniques. The simplest and most commonly used performance metric is the plate number N:

$$N_{\text{spatial}} = \frac{Z^2}{\chi^2} \text{ or } N_{\text{temporal}} = \frac{t^2}{\sigma^2} \tag{2}$$

for spatial or temporal analysis, respectively, where the various mean displacements and variances are taken from a single peak.

Strictly speaking, the plate efficiency is only applicable to rate-based separations, for example, zone electrophoresis or isocratic chromatography, as opposed to (pseudo)equilibrium-based separations, for example, isoelectric focusing or gradient-elution chromatography. Typical values of the plate number in zone electrophoresis range from below 1,000 for the high-capacity BioStream® apparatus (13,14) to above 10^7 for PAGE in gel-filled capillaries (15,16).

A closely related concept to plate number is plate height (17), HETP $= Z/N_{\text{spatial}}$ or $= (\langle u \rangle t)/N_{\text{temporal}}$. Although neither plate height nor plate number can be considered a truly universal metric for comparing different separation columns, they are both useful ways of checking performance within an individual column, for example, after a packing has been changed in an electrochromatography column.

Resolution A more universal metric for separation columns that can be applied to both rate-based and equilibrium-based techniques is resolution. An approximate form for the resolution that is sufficiently accurate for most applications is (17)

$$R_{\text{spatial}} = \frac{|Z_1 - Z_2\rangle}{\frac{1}{2}(W_1 + W_2)} \approx \frac{|Z_1 - Z_2|}{2(\chi_1 + \chi_2)} \quad \text{or}$$

$$R_{\text{temporal}} = \frac{|t_1 - t_2|}{\frac{1}{2}(W_1 + W_2)} \approx \frac{|t_1 - t_2|}{2(\sigma_1 + \sigma_2)} \tag{3}$$

for spatial or temporal analysis, respectively, where the relationship between the peak width at its base and the variance is $W = 4\chi$ (spatial) or 4σ (temporal). Two key components with an R equal to or slightly greater than 1 can be considered to be baseline separated, so R can be used either to compare data from different separation devices or to modify the column so that complete separation can be achieved.

BATCH ANALYTICAL TOOLS

Capillary Electrophoresis

Capillary electrophoresis (CE) offers important advantages over conventional gel electrophoresis (GE) in the characterization of proteins and nucleic acids. CE is much faster than GE, and can be fully automated and instrumented to include sample injection, detection, and collection. In addition, CE can separate solutes at a resolution significantly greater than that of HPLC. Very small sample loads (e.g., <10 ngs/run) make CE suitable primarily for analytical work. Automated equipment for CE is commercially available from several vendors, including units that allow collection of fractionated samples (18,19) or direct coupling to a secondary separation device (e.g., HPLC or mass spectrometer).

Basic Operation. In principle CE is quite simple. It can be carried out at very high power densities because the large surface area to volume ratio of the capillary lumen allows it to shed joule heat efficiently. However, in practice it has not yet reached the level of maturity of automated chromatography (i.e., where stand-alone columns can be operated by individual staff on an as-needed basis). CE is best operated by a trained technician who retains responsibility for developing protocols. The basic apparatus consists of a fused-silica capillary with a lumen diameter typically ranging from 25 to 150 μm, an outer diameter of several hundred microns (e.g., 400 μm) and a thin coat of polyimide that makes the otherwise brittle capillary quite flexible. A small window is burned in the polyimide near one end of the capillary, and this window is then placed in the beam of a spectrophotometer. To avoid damaging the sensitive electronics in the detector, this end of the capillary is dipped into a vial filled with running buffer and the electrode placed in this vial is set to electric ground (20).

Before loading a sample into a bare capillary the lumen is usually precleaned with 0.1–0.5 M sodium hydroxide and then rinsed sequentially with pure water and running buffer. The capillary is loaded by dipping the high-voltage end into a vial containing the sample and using a pressure head or electric field to load 10–100 nL of sample, which is roughly 0.01–0.1 μg at a concentration of 1 mg/mL. The high-voltage end of the capillary is then moved to a vial containing running buffer, the hot electrode is dipped into this vial, and the voltage is increased until the electric field falls to a range of 200–400 V/cm. These large fields drive solutes through the capillary at electrophoretic velocities on the order of 1–4 cm/min, yielding run times of a few minutes with plate efficiencies approaching 10^6 plates per meter of capillary in free solution.

Fundamental Principles. Two important points that must be considered when discussing CE performance are

the temperature profile in the capillary and dispersion due to temperature variations. The temperature profile can be calculated using the steady-state form of the thermal energy equation in cylindrical coordinates

$$\frac{k_{\text{buffer}}}{r}\frac{d}{dr}\,r\,\frac{dT}{dr} + \sigma_{\text{buffer}}T_{\text{ref}}E^2[1 + \lambda(T - T_{\text{ref}})] = 0$$

$$\text{where } \lambda \equiv \left[\frac{d\log_e(\sigma_{\text{buffer}})}{dT}\right]_{T=T_{\text{ref}}} \tag{4}$$

where the first term accounts for radial conduction and the second for generation of joule heat. The second term in brackets accounts for the autothermal effect (21,22), that is, the tendency of warm buffer to conduct more current than cold buffer. The thermal energy equation, equation 4, must be solved together with the boundary conditions

$$\text{at } r = 0 \quad T \text{ is finite}$$

$$\text{at } r = a \quad -k_{\text{buffer}}\frac{dT}{dr} = \frac{1}{2}h_{\text{CE}}(T - T_{\text{ref}}) = \frac{1}{2a}\frac{T - T_{\text{ref}}}{\left[\frac{1}{k_{\text{g}}}\log(a_{\text{g}}/a) + \frac{1}{k_{\text{p}}}\log(a_{\text{p}}/a_{\text{g}}) + \frac{1}{a_{\text{p}}h_{\text{fluid}}}\right]} \tag{5}$$

The boundary condition at $r = 0$ states that the temperature must remain finite at the centerline, and the wall condition at $r = a$ indicates that heat transfer from the capillary buffer to the environment outside the capillary is controlled by the serial resistances of the glass capillary, the polyimide coat, and the boundary layer in the external cooling fluid surrounding the capillary. These three resistances correspond respectively to the three terms in the denominator of the boundary condition at $r = a$. The overall heat transfer coefficient between the lumen OD and the coolant is h_{CE}.

Equation 4 with the boundary conditions in equation 5 has the solution

$$\Theta_{\text{buffer}}\eta = \frac{1}{\kappa^2}\left(\frac{BiJ_0(\kappa\eta)}{BiJ_0(\kappa) - 2\kappa J_1(\kappa)} - 1\right) \tag{6}$$

with dimensionless parameters defined as

$$\Theta_{\text{buffer}} = \frac{k_{\text{buffer}}(T - T_{\text{ref}})}{\sigma_{\text{buffer}}E^2a^2} \quad \eta = \frac{r}{a}$$

$$Bi = \frac{2h_{\text{CE}}a}{k_{\text{buffer}}} \quad \kappa^2 = \lambda\frac{\sigma_{\text{buffer}}E^2a^2}{k_{\text{buffer}}}$$

where Bi is the Biot number. J_0 and J_1 are Bessel functions of the first kind of order 0 and 1, respectively. The temperature profile in the capillary lumen is roughly parabolic with its maximum at the centerline. As seen in Figure 2, the temperature in the lumen can vary markedly with applied voltage, as well as with the overall heat transfer coefficient h_{CE}. However, the variation across the radius of the lumen is usually less than 1 °C, and even this small variation can have a marked impact on dispersion in CE. This occurs because the electrophoretic mobility varies inversely with viscosity. Viscosity, in turn, decreases with increasing temperature, so solutes migrate faster in the hot center of the lumen than near the lumen wall, distorting the solute band into a parabolic shape. When this is taken into account, the effective dispersion coefficient in the capillary becomes (23,24)

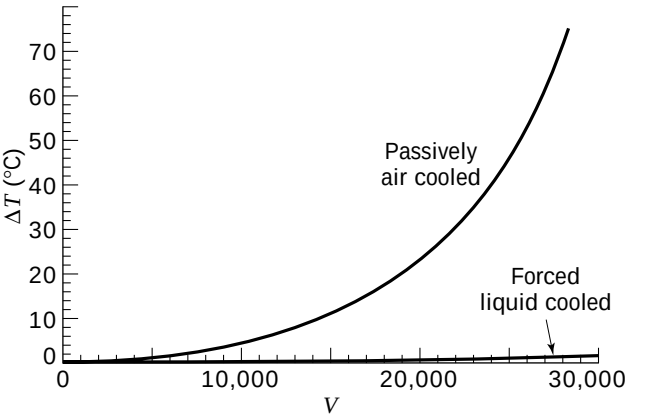


Figure 2. Temperature rise in the CZE lumen as a function of applied voltage V. Both the centerline and wall temperatures are shown but, since they differ by less than 1 °C, only a single line is apparent in this graph. The buffer is assumed to have a conductivity, $\sigma_{\text{e}} = 0.01$ mho/cm.

$$D_{\text{CE}} = D_{\text{m}} + \frac{a^2(\mu(T_{\text{ref}})E)^2}{3072D_{\text{m}}}\frac{\kappa^4 Bi}{(BiJ_0(\kappa) - 2\kappa J_1(\kappa))} \tag{7}$$

The first term on the right accounts for pure axial diffusion, and if only this term is included it predicts a temporal variance equal to that given in equation 1. Under normal operating conditions the second term is roughly the same order of magnitude as the first, indicating that thermal band broadening is generally significant in CE. More detailed information about the range of the Biot number, Bi, and the autothermal coefficient, κ, is available in Ref. 24.

Performance Benchmarks. Modifying the expression for the number of theoretical plates to include sample volume, wall electroosmosis, and the extended dispersion coefficient, yields the formula

$$\frac{1}{N_{\text{CE}}} = \frac{2D_{\text{CE}}}{|(\mu + \mu_{\text{eo}})\Delta\Phi|} + \frac{1}{12}\frac{(q/\pi a^2)^2}{Z^2} \tag{8}$$

where μ_{eo} is the electroosmotic mobility and q is the volume of sample loaded into the capillary. The first term on the right gives the contribution from distortion due to thermal effects and the second term gives the contribution from the variance of the initial sample injection volume q. When hydrostatic loading is used, the injection volume can be calculated as

$$q = \frac{\rho g H}{8\nu Z}\pi a^4 t_{\text{load}} \tag{9}$$

where ρ and ν are the density and viscosity of the buffer,

respectively, and H is the difference in height between the liquid levels in the vials into which the two ends of the capillary are immersed, assuming both vials are at atmospheric pressure. Alternatively, if a pressure differential is used to load the capillary, then ρgH can be replaced with ΔP.

As illustrated in Figure 3, for capillaries whose lumen ID is less than 100 μm and for a loading of about 1% of the column volume, the number of theoretical plates levels out near 1×10^5. This result is approximately true regardless of how the capillaries are cooled, for example, by air in natural convection or by forced liquid flow. In small capillaries, the sample injection volume sets an upper limit on capillary performance because the sample volume can be reduced only to the point where the signal is still detectable. After that, to improve performance, the sample must be preconcentrated or stacked or the sensitivity of the detection system somehow improved. Clearly CE will benefit from the development of protocols for stacking or focusing injected solutes in free solution before starting a run (25,26).

The main advantages of forced cooling are that it *(1)* dramatically reduces the temperature rise in the capillary lumen by reducing the heat transfer resistance in the external fluid and *(2)* reduces the thermal fluctuations associated with natural convection. Using conventional sample-loading protocols, there is little benefit in plate efficiency gained by forced liquid cooling. However, as stacking procedures become more common, liquid-cooled CE systems should eventually outperform both passively and actively air-cooled units.

Polyacrylamide Gel Electrophoresis

For several decades now, polymer gel electrophoresis has been a standard method used for laboratory analysis of proteins, nucleic acids, and other biological materials (27,28). The primary advantages of GE over other applicable techniques are *(1)* moderately high resolution, that is, native PAGE routinely provides more than 10,000 theoretical plates and *(2)* the protocols are well established and widely accepted in the field (29). All engineers and scientists who enter the biotechnology industry will, at some point in their careers, find themselves running or reading results from electrophoresis gels.

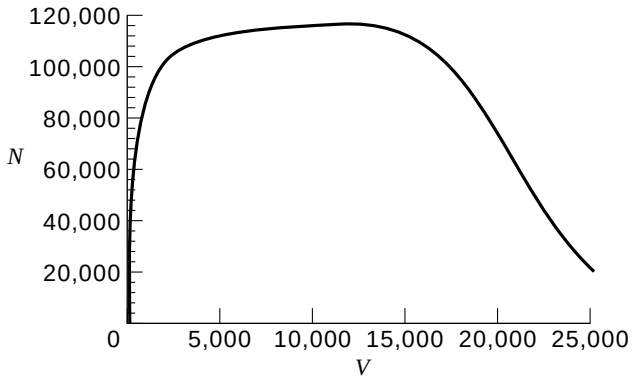


Figure 3. Predicted plate number as a function of applied voltage for a 75-μm ID $\times$ 350-μm OD $\times$ 100-cm capillary in air using an injection volume, $q = 50$ nL. In this example the excessive volume of the hydrostatic load limits the number of theoretical plates that can be obtained for this system. Operating at voltages near 15 kV will give the best combination of plate efficiency and speed.

The primary disadvantages of conventional GE are that it is *(1)* labor intensive; *(2)* slow, often requiring 24 h between initial buffer preparation and final destain; *(3)* difficult to recover purified solutes for subsequent steps; and *(4)* difficult to scale beyond a few milligrams of sample. My experience with PAGE is that it takes several weeks of practice to be able to produce publishable gels, and even then users may experience problems with gel-to-gel reproducibility, especially between departments within a company.

Given these limitations, it is not surprising that an effort would be made to develop instrumentation that retains the advantages of GE while minimizing its disadvantages. One can reduce the size and thickness of the gel, actively remove joule heat from one or both sides of the gel, and automate the stain and destain cycle. This greatly reduces labor requirements, turnaround time, and many of the problems associated with gel reproducibility. Also, a variety of prefabricated electrophoresis gels are now commercially available from a number of vendors for about $10 each.

Basic Operation. Detailed protocols for preparing, running, and analyzing gels are available in any number of references (e.g., Ref. 29) and are not discussed further here. In all likelihood, the end user will run some combination of SDS or native PAGE with the sample in reduced or nonreduced form, in Tris or tricine buffer, with or without a gradient in cross-linking agent, using one of several stacking procedures and Coomassie blue or silver stain to visualize the fractionated bands. Within a biotech company, it is also likely that the gel protocols, including stain and destain procedures, will be specified by a quality assurance or analytical chemistry department or their equivalent. These protocols will almost certainly include a ladder of molecular weight markers in one lane and additional lanes containing controls appropriate to the separation.

IEF-PAGE gels or capillary IEF should be run in conjunction with the molecular weight gels to obtain pIs for the various protein isoforms and glycoforms that inevitably accompany production of recombinant therapeutics. Not only does this give a clearer picture of both contamination and heterogeneity of the purified protein, but it may also help in preparing a rational design for downstream processing.

Fundamental Principles. Although some progress has been made in understanding how proteins migrate through tight gels (30–32) from a practical standpoint it is not necessary to understand the underlying theory to produce publishable gels. Once an appropriate gel and running buffer have been selected for the molecular weight range of interest, it is best to run the gel as quickly as possible to minimize diffusive band spreading. As a rule of thumb, an electrical power density of 0.1 W/cm^3 per gel should be reasonably low for use with air-cooled, glass-walled, vertical slab PAGE. (*Note*: To calculate watts, mul-

tiply amps times voltage. If the gel is set up correctly, nearly all of this power will be dissipated in the working portion of the gel slab.) However, if the apparatus comes with a run protocol, these directions should be followed closely, at least during the initial stacking phase of the run. Once the proteins have entered the separating gel, the user has a bit more flexibility with the applied voltage, depending on how well the gel box is cooled, but the voltage should probably not exceed 0.2 W/cm^3 per gel in an air-cooled box.

Performance Benchmarks. Once the stacked proteins have entered the separating gel, the conventional wisdom is that band spreading in near-isothermal gels is controlled by diffusion. Because noneluting gels are analyzed in space, the number of theoretical plates is given by equation 2 as

$$N_{\text{SDS-PAGE}} = \frac{Z^2}{\chi^2} = \frac{Z^2}{2D_{\text{eff}}t} = \frac{(\mu_{\text{eff}}Et)Z}{2D_{\text{eff}}t} = \frac{(\mu_{\text{eff}}\Delta\Phi)Z}{2D_{\text{eff}}L_{\text{gel}}} = \frac{\mu_{\text{eff}}\Delta\Phi}{2D_{\text{eff}}}\frac{Z}{L_{\text{gel}}} \approx 10{,}000 \quad (10)$$

where the approximation is for a 5-cm displacement. The penultimate term, which is a classical result for electrophoresis, shows how applied voltage, gel length, and so on, affect performance.

This formula also implies that high performance can be maintained as the gel is scaled down to about the size of a glass slide (33–36) if the voltage drop is fixed during miniaturization. The higher-power densities generated can be accommodated by replacing the passive air-cooled, glass-encased gel by an actively cooled (e.g., Peltier system), ceramic-encased gel and by making the gel as thin as is reasonable for handling and detection. Actively cooling both sides of the gel would allow the user to apply about twice the voltage and would therefore be expected to double performance, that is, double the run speed and increase resolution by 41%.

BATCH PREPARATIVE TOOLS

Preparative Gel Electrophoresis

Continuous elution preparative gel electrophoresis (pGE) is a relatively new entry in the field of preparative bioprocessing and is intended to address the limitations of slab GE. The essence of this technique is that a pulse of multicomponent sample can be loaded onto a column and fractionated using well-known protocols on a standard electrophoresis gel and then continuously recovered in free solution as it migrates off the end of the gel. This technique may be viewed as a melding of conventional gel electrophoresis with electroelution. In combining these two steps, it is possible to relieve several of the disadvantages of GE mentioned earlier, especially by automating and accelerating the otherwise labor-intensive sample-recovery step.

Basic Operation. A generic pGE apparatus consists of three basic parts: (*1*) a tube- or annular-shaped electrophoresis gel, (*2*) an elution chamber at the base of the gel, and (*3*) a product offtake line leading to a detector and/or fraction collector. In practice, a sample is loaded into the liquid layer immediately adjacent to the top of a standard gel and then electrophoresed, at low power to avoid mixing due to natural convection, into the top of the gel. Once inside the gel, electrophoresis of the solutes is carried out at power densities at or below 1 W/cm^3. The solutes migrate through the gel at different velocities, and as each solute band reaches the end of the gel it elutes into free solution in the elution chamber and is carried to a fraction collector by the elution buffer. Once the effluent has been aliquoted by the fraction collector, the recovered solutes, which are suspended in the elution buffer, are ready for biochemical analysis or for further processing (37).

Although only a single batch sample can be processed in each run, the gel can be reused several times before it must be replaced. A typical column load ranges from less than 100 μg to more than 10 mg per run depending on the cross-sectional area of the gel; typical run times are somewhat less than 30 min per centimeter of column.

The Electrophoresis Gel. In the electrophoresis gel, the central issue is resolution. This is the heart of the apparatus, and once the purified solutes have eluted from the gel, whatever resolution was achieved there can only be degraded in subsequent processing steps. In this part of the system the major contributors to peak variance are sample volume, molecular diffusion, and thermally induced dispersion. Sample bandwidth is directly related to the volume of the sample band as it begins electromigration through the gel. However, if a stacking gel is used, sample volume is usually not an important contributor to peak variance.

Assuming isothermal conditions and ignoring the contribution to the variance due to sample injection volume, molecular diffusion dominates band spreading by contributing a temporal variance

$$\sigma^2_{\text{diffusion}} = \frac{2D_{\text{gel}}Z}{\langle u_{\text{gel}}\rangle^3} \quad (11)$$

as measured by a detector located at a fixed position Z. Here D_{gel} is the diffusion coefficient of the solute in the gel and $\langle u_{\text{gel}}\rangle = \langle \mu_{\text{gel}}\rangle E$ is the average electrophoretic velocity of the solute in the gel. Noting that the average holdup time for a gel of length Z is $Z/\langle u_{\text{gel}}\rangle$, the number of theoretical plates is

$$N_{\text{diffusion}} = \frac{t^2}{\sigma^2_{\text{diffusion}}} = \frac{u_{\text{gel}}Z}{2D_{\text{gel}}} \quad (12)$$

For native PAGE, the number of theoretical plates routinely falls in the range $10^3 < N < 10^4$.

The Elution Chamber. As solutes leave the bottom of the gel tube, they are rapidly swept away from the gel, out of the electric field, and into an elution stream. In a properly designed elution chamber, dispersion is not an important issue, but dilution is. If the elution stream is too slow, some or all of each solute will be entrained by the electric field and permanently removed from the elution stream. If the elution stream flow rate is too high, the solute will be di-

luted, making it difficult to detect or use in subsequent processing steps.

Many different strategies could be used to reduce dilution in the elution chamber, and this part of the system offers the greatest potential for innovation. Here only two generic approaches to elution, cross-flow and counterflow, are considered to illustrate how an elution chamber is designed.

In cross-flow elution (Fig. 4) a hydrodynamic flow is directed perpendicular to the electric field at the base of the gel tube and flows out of the elution chamber through an opening on the right. The minimum elution flow rate is determined by finding the conditions under which a molecule exiting the gel tube and entering the elution stream at the farthest point from the elution channel just avoids being captured by the electric field. In Figure 4, the lower trajectory is captured on the electrode membrane, and the upper trajectory is cleared into the elution stream.

In counterflow elution, the hydrodynamic flow is directed opposite to the electrophoretic migration in the elution chamber. The minimum elution flow is found when a molecule's electrophoretic velocity in the elution buffer exactly balances the counterflow. As in Figure 4a, the lower trajectory is captured on the electrode membrane and the upper trajectory is cleared into the elution stream.

It turns out that in both of these cases the minimum elution flow rate that discourages capture of the solute by the electrode is

$$Q_{\min} = \langle u_{\text{buffer}} \rangle A_c \tag{13}$$

where A_c is the cross-sectional area of the elution chamber through which the current flows and $\langle u_{\text{buffer}} \rangle$ is the electrophoretic velocity in free solution. This is the free solution electrophoresis velocity in the elution buffer and may be substantially different from the velocity in the gel. In practice, the elution flow rate is often several times higher than this minimum value, leading to unnecessary dilution of the sample.

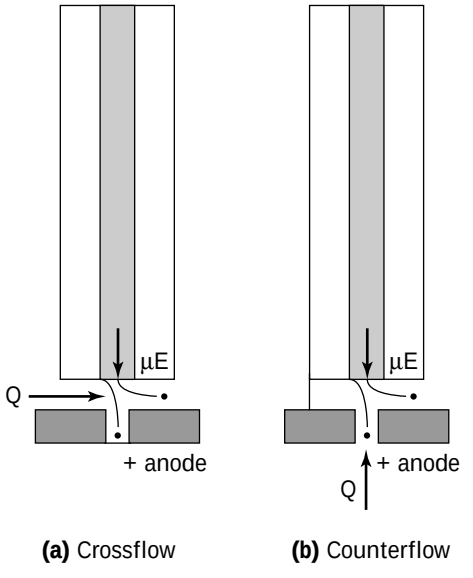


Figure 4. In cross-flow elution **(a)** the sample is swept away from the electrode port by a perpendicular flow of elution buffer. In counterflow elution **(b)** the elution buffer flow rate is directly opposed to the electrophoresis of the solute toward the lower electrode. In both cases if the elution flow rate is too low, the solute may be "captured" by the electrode. This is represented by the particle that follows the lower trajectory in this figure.

The Product Offtake Line. After being flushed from the elution chamber, the solute is carried by the elution buffer through a length of tubing to a detector and/or fraction collector. In this stage the main consideration is loss of resolution, since the purified solute bands are now subject to Taylor (diffusive–convective) dispersion (38) as they travel through the tubing. This occurs because the flow in the system is laminar, parabolic with a maximum velocity at the tube centerline. Solutes near the wall move slowly compared with those near the centerline so the solute band is elongated by the flow as it travels through the tube.

Aris–Taylor theory (39), which is strictly valid only for long tubes, predicts a dispersion coefficient that goes as

$$D_{\text{Taylor}} = D_m + \frac{\langle v \rangle^2 d_t^2}{192 D_m} = D_m + \frac{Q^2}{12\pi^2 d_t^2 D_m} \tag{14}$$

where $\langle v \rangle$ is the average hydrodynamic velocity of the buffer in the tube. Because of Taylor dispersion the offtake line can generate significant band spreading if it is improperly designed. The variance predicted by Taylor's theory for dispersion in a tube of length L is

$$\sigma^2_{\text{Taylor}} = \frac{D_m L}{32} \frac{\pi^3 d_t^6}{Q^3} + C \frac{\pi d_t^4 L}{384 Q D_m} \tag{15}$$

which indicates that Taylor dispersion may be reduced by using short, thin tubes. For tubes of length $L < Q/\pi D_m$, Aris–Taylor theory overpredicts dispersion and must be adjusted before it can be applied to the lengths of elution channel typically used with these instruments. To put this in perspective, for Aris–Taylor theory to apply in commercial pGE equipment, one would need $L > 1$ m when $Q = 10\,\mu$L/min and $L > 100$ m for $Q = 1$ mL/min. The constant C, which corrects for short tubes, can be estimated from Figure 5, where $\zeta = D_m L/\langle v \rangle a_t^2$.

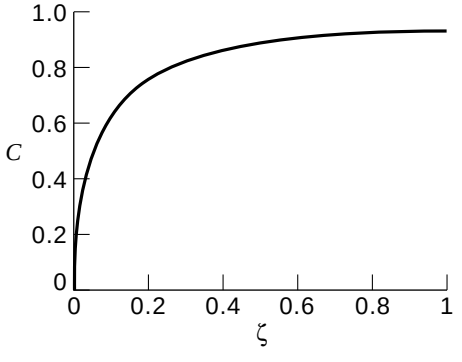


Figure 5. Ratio of the dispersion coefficient in tubes of arbitrary length to the dispersion coefficient in long tubes as a function of the dimensionless distance ζ.

For example, a 1-m-long tube with a flow rate of 1 mL/min has an adjusted variance of 8,214 s^2 at a dimensionless distance of $\zeta = 0.009425$ ($C = 0.185803$). A pulse of solute placed at the entrance of the tube will take about 6.1 min to flush through the tube entirely. The implication here is that solute bands that come off the gel less than 6.1 min apart will tend to overlap as they come out of the elution stream tube. If this becomes a problem, it is best addressed by decreasing the length of the tube.

The effect of Taylor dispersion in short tubes is summarized in Figure 6 for the two cases, $Q \approx 1$ mL/min, which corresponds roughly to the annular device, and $Q \approx 10$ μL/min, which corresponds roughly to the tubular device. The temporal bandwidth, that is, the time it takes for a pulse of solute to completely elute from the tube, is approximately equal to 4σ, so this has been plotted against the tube length for each of these flow rates. Note that the temporal dispersion is independent of tube diameter (a counterintuitive result) for a fixed volumetric flow and that dispersion can be decreased by reducing Q and L.

Although existing continuous elution systems do not meet the long-tube condition required for Taylor's analysis to apply rigorously, the variance can still be significantly decreased by reducing the inner diameter and length of the elution stream tube.

Fundamental Principles. As with CE, temperature profiles in pGE can be calculated using equation 4, except that the electrical parameters used are those that pertain to the gel. The steady-state form of the thermal energy equation in cylindrical coordinates is

$$\frac{k_{gel}}{r}\frac{d}{dr}r\frac{dT}{dr} + \sigma_{gel}(T_{ref})E^2[1 + \lambda_{gel}(T - T_{ref})] = 0 \quad (16)$$

For a tubular gel, this equation is put in dimensionless form using the transformations

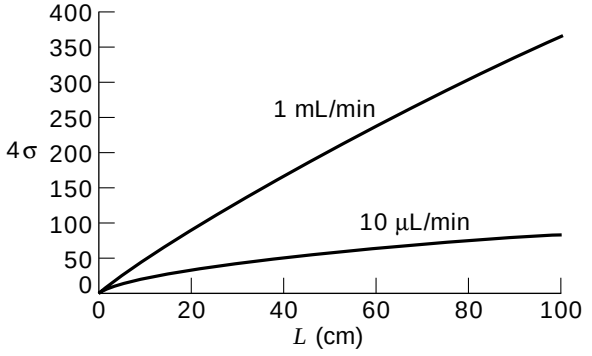


Figure 6. Contribution of the variance computed to the baseline width, in seconds, for flow in the elution stream tube. Although this is generally not a significant contribution to the overall dispersion, it can smear bands that elute within a few minutes of one another. The dispersion is alleviated by reducing the elution flowrate Q, and keeping the elution tube as short as possible.

$$\Theta_{gel}(\eta) = \frac{k_{gel}(T - T_{ref})}{\sigma_{gel}E^2a_{gel}^2} \quad Bi = \frac{2h_{pGE}a^{gel}}{k_{gel}} \approx \frac{2k_{glass}}{k_{gel}\log\left[\frac{a_{gel}}{a_{glass}}\right]}$$

$$\kappa^2 = \frac{\sigma_{gel}E^2a_{gel}^2}{k_{gel}}\left[\frac{d\log(\sigma)}{dT}\right]_{T_{ref}}$$

where a_{gel} is the radius of the gel and $\eta = r/a_{gel}$. The heat transfer coefficient h_{pGE} for the micropreparative apparatus includes resistances due to the glass wall of the tube and the external cooling fluid. If cooling is by forced air and the glass wall is thick, the heat transfer resistance is typically controlled by the glass. This is reflected in the approximate term on the far right in the relation for Bi.

These transformations yield a second-order ordinary differential equation that, when solved with the boundary conditions from equation 5, has a solution of the form

$$\Theta_{gel}(\eta) = \frac{1}{\kappa^2}\left(\frac{BiJ_0(\kappa\eta)}{BiJ_0(\kappa) - 2\kappa J_1(\kappa)} - 1\right) \quad (17)$$

which is identical to equation 6 except that the definitions of the dimensionless parameters have been changed.

Table 2 contains the physical data used to predict temperature profiles in pGE at a base power dissipation rate of 1 W/cm^3. As is clear in Figure 7, at this power setting the centerline temperature rise is only about 1 °C greater than the coolant temperature. The 1 °C variation in temperature implies that, for practical purposes, thermal excursions do not have a significant impact on dispersion or resolution when compared with contributions from other sources. Consequently, it is acceptable to ignore temperature effects when optimizing pGE performance at power densities below 1 W/cm^3.

Annular Gels. In cases where larger amounts of material must be processed in a single run, the basic tube gel used can be significantly scaled by switching to an annular gel. The gel in the annulus should be kept thin to avoid excessive temperature excursions, but the annulus perimeter can be enlarged to increase capacity. For example, an annulus with a 2-cm inner diameter and a 3-cm outer diameter would have a typical column load of roughly 20 mg/run. An annulus with a 4-cm outer diameter and a 2-cm inner diameter could process about 50 mg/run.

For a thin annulus with a large diameter and an annular gap equal to the diameter of a tubular column, the relative increase in scale of the annulus over the tube is approximately $4a_o/a_{gel}$, where a_o is the outer diameter of the annulus. In the more general case where the annulus

Table 2. Physical Constants Used in the Tubular pGE Illustrations

	Gel tube
$\sigma(T_{ref})E^2$ (W/cm^3)	1.0
Z (cm)	5.0
$2a_{gel}$ (cm)	0.25
d_i (cm)	0.8
Run time (h)	3–5

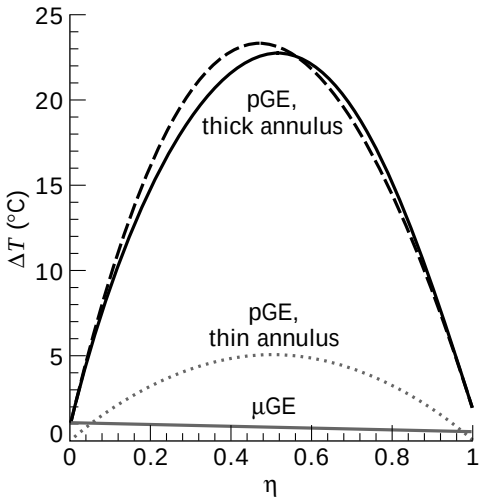


Figure 7. Temperature rise (°C) in continuous elution gel electrophoresis at a base power density, $P_w = 1$ W/cm^3 in a tube gel (———), a thin annulus (– – – –), and a thick annulus (· · · ·). Note the slight asymmetry in the annular temperature profiles. The line next to each curve is the corresponding parabolic Galerkin approximation used in the dispersion calculations.

is not thin, the relative increase in scale is given by the ratio S of the cross-sectional areas of the two columns

$$S = \frac{a_o^2 - a_i^2}{a_{gel}^2} \tag{18}$$

The capacity can be increased further by increasing the thickness of the annulus and by lengthening the column, but the former admits some degradation of performance due to thermal dispersion and the latter requires longer run times.

For an annulus, equation 16 is put in dimensionless form by using different transformations than in the previous cases. These are

$$\Theta(\eta) = \frac{k_{gel}(T - T_{ref})}{\sigma(T_{ref})E^2(a_o - a_i)^2} \quad \eta = \frac{r - a_i}{a_o - a_i}$$
$$\kappa^2 = \frac{\sigma(T_{ref})E^2(a_o - a_i)^2}{k_{gel}}\left[\frac{d\,\log(\sigma)}{dT}\right]_{T=T_{ref}} \quad \Lambda = \frac{a_i}{a_o - a_i}$$

where a_i is the radius of the inner surface of the gel and a_o is the radius of the outer surface. Note that *(1)* the dimensionless radial coordinate η is defined so that the domain extends over $0 < \eta < 1$ and *(2)* a new geometric parameter, Λ, appears in the governing equations.

These transformations yield a second-order ordinary differential equation

$$\frac{d^2\Theta}{d\eta^2} + \frac{1}{\eta + \Lambda}\frac{d\Theta}{d\eta} + \kappa^2\Theta = -1 \tag{19}$$

that has solutions of the form

$$\Theta(\eta) = AJ_0[\kappa(\eta + \Lambda)] + BY_0[\kappa(\eta + \Lambda)] - \frac{1}{\kappa^2} \tag{20}$$

where A and B are constants that must be fit using appropriate boundary conditions. J_o is a Bessel function of the first kind of order 0, and Y_o is a Bessel function of the second kind of order 0.

Appropriate boundary conditions in this case are

$$\text{at } \eta = 0 \quad +\frac{\partial\Theta(\eta)}{\partial\eta} = Bi_i\Theta(\eta) \quad \text{where } Bi_i = \frac{k_i}{k_{gel}}\frac{2}{\log(2a_i/d_i)}$$
$$\text{at } \eta = 1 \quad -\frac{\partial\Theta(\eta)}{\partial\eta} = Bi_o\Theta(\eta) \quad \text{where } Bi_o = \frac{k_o}{k_{gel}}\frac{2}{\log(d_o/2a_o)} \tag{21}$$

where d_i is the inside diameter of the cooling finger and d_o is the outer diameter of the glass tube. For the annular device, the constants of integration, A and B, are obtained by solution of the matrix equation

$$\begin{bmatrix} -\kappa J_1(\kappa(1+\Lambda)) + Bi_o J_0(\kappa(1+\Lambda)) & -\kappa Y_1(\kappa(1+\Lambda)) + Bi_o Y_0(\kappa(1+\Lambda)) \\ +\kappa J_1(\kappa\Lambda) + Bi_i J_0(\kappa\Lambda) & +\kappa Y_1(\kappa\Lambda) + Bi_i Y_0(\kappa\Lambda) \end{bmatrix}\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \frac{Bi_o}{\kappa^2} \\ \frac{Bi_i}{\kappa^2} \end{bmatrix}$$

and autothermal runaway is predicted at values of κ for which the determinate of the 2×2 matrix on the left vanishes.

Table 3 contains the physical data used to predict temperature profiles in the preparative chamber at a base power dissipation rate of 1 W/cm^3. For purposes of illustration, two different cases are shown. The first is an instrument with a thin annular gap and the second has a thick annular gap. The predicted temperature rise in each instrument under these conditions is shown in Figure 7. Since the temperature rise in each of these devices is roughly proportional to $\sigma(T_{ref})E^2(a_o - a_i)^2/k_{gel}$, these results can be readily extrapolated to other operating conditions and gel geometries. To maintain performance levels while scaling up, the product of the electric field times the gap thickness, $E(a_o - a_i)$, should be held constant in order to limit temperature variations across the transverse axis, thereby discouraging excessive thermal dispersion.

Table 3. Physical Constants Used in the Preparative GE Illustrations

	Thin annulus	Thick annulus
L (cm)	10.0	10.0
$2\,a_i$ (cm)	1.905	1.905
$2\,a_o$ (cm)	2.8	3.7
d_o (cm)	1.565	1.565
d_i (cm)	3.14	4.04
k_{gel} (W/cm °C)	0.00578	0.00578
k_{glass} (W/cm °C)	0.011	0.011
$k_{ceramic}$ (W/cm °C)	0.0578	0.0578
Run time (h)	5–15	5–15

This implies that, while the scale may be doubled by doubling the annular gap thickness, the field must be decreased by half so the run takes twice as long. Consequently, overall productivity in preparative GE remains roughly independent of gap thickness, with the major savings coming from the reductions in the time and effort associated with reloading and purging the gel. As stated earlier, the gel apparatus is most effectively scaled by increasing the perimeter of the annulus while keeping the annular gap thickness constant.

Solute Concentration Profiles. In the absence of thermal dispersion, the concentration profiles of solutes exiting the gel can be calculated by solving the diffusion–convection equation

$$\frac{\partial C}{\partial t} + u_{\text{gel}} \frac{\partial C}{\partial z} = D_{\text{gel}} \frac{\partial^2 C}{\partial z^2} \tag{22}$$

subject to the boundary and initial conditions

$$\begin{array}{lll} \text{at } t = 0 & \text{all } z & C = 0 \\ \text{at } z = 0 & \text{all } t & C = H(0, \tau)C_o \\ \text{as } z \to \infty & \text{all } t & C \text{ is bounded} \end{array} \tag{23}$$

where the distribution function $H(0, t) = 1$ if $0 < t < \tau$ and $H(0, t) = 0$ if $t > \tau$. This problem is most conveniently solved using Laplace transforms to obtain

$$\begin{aligned} C = \frac{C_o}{2}\Bigg[& e^{\text{Pe}_{\text{gel}} Z/a_{\text{gel}}\zeta}\Bigg(\text{Erfc}\left[\sqrt{\frac{Pe_{\text{gel}} Z}{4a_{\text{gel}}}}\left(\frac{\zeta}{\sqrt{\theta}} + \sqrt{\theta}\right)\right] \\ & - \text{Erfc}\left[\sqrt{\frac{Pe_{\text{gel}} Z}{4a_{\text{gel}}}}\left(\frac{\zeta}{\sqrt{\theta - \Theta}} + \sqrt{\theta - \Theta}\right)\right]\Bigg) \\ & + \text{Erfc}\left[\sqrt{\frac{Pe_{\text{gel}} Z}{4a_{\text{gel}}}}\left(\frac{\zeta}{\sqrt{\theta}} - \sqrt{\theta}\right)\right] \\ & - \text{Erfc}\left[\sqrt{\frac{Pe_{\text{gel}} Z}{4a_{\text{gel}}}}\left(\frac{\zeta}{\sqrt{\theta - \Theta}} - \sqrt{\theta - \Theta}\right)\right]\Bigg] \end{aligned} \tag{24}$$

where the dimensionless parameters in equation 24 are defined as

$$\theta = \frac{\langle u_{\text{gel}}\rangle t}{Z} \quad \zeta = \frac{z}{Z} \quad \Theta = \frac{q}{\pi a_{\text{gel}}^2 Z}\frac{\langle u_{\text{gel}}\rangle}{u_{\text{buffer}}} \quad Pe_{\text{gel}} = \frac{\langle u_{\text{gel}}\rangle a_{\text{gel}}}{D_{\text{gel}}}$$

Here θ is the dimensionless time, ζ is the dimensionless distance, Pe_{gel} is the gel Peclet number, and Θ is a dimensionless loading factor. Note that the leading fraction in the definition of Θ is the ratio of the load volume to the gel volume and is normally kept below 5%. The trailing fraction is the ratio of the electrophoretic velocities in the gel and in the buffer. If a stacking gel is used, the estimated volume of the sample as it leaves the stacking gel should be used in calculating Θ.

Figure 8 illustrates the effect of decreasing the injection volume from 50 μl to 5 μL. At the lower injection volumes (leading peak) the peak concentration drops below the value in the feed, its breadth decreases, and the center of the peak approaches a symmetric distribution about the mean elution time, $\theta = 1$, that is, $t = Z/\langle u_{\text{gel}}\rangle$. At higher loadings the peak width expands to accommodate the added mass of solute, thus making it more difficult to resolve closely spaced peaks. To minimize the influence of the injection volume, q should be adjusted so that Θ ;< 1/300. This is most conveniently done by using a stacking procedure.

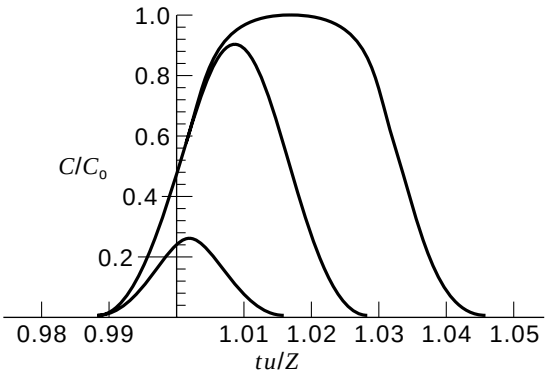


Figure 8. The effect of injection volume q, on effluent concentration profiles from the micropreparative gel electrophoresis chamber. The ordinate is dimensionless concentration C/C_o and the abscissa is dimensionless time θ. This simulation was carried out with $Pe_{\text{gel}} = 2{,}000$, $Z = 5$ cm, $a_{\text{gel}} = 0.125$ cm, and $Q = 1/30$, 1/60, and 1/300 ($q \approx 50$, 25, and 5 μL, respectively) and clearly shows the reduction in bandwidth and effluent concentration that accompanies decreasing q. The smallest injection volume is shown in the leading curve.

Thermal Dispersion. To estimate the effect of radial temperature variations on axial dispersion, it is necessary to solve the extended form of the convective–diffusion equation

$$\frac{\partial C}{\partial t} + u(r)\frac{\partial C}{\partial z} = \frac{D_{\text{gel}}}{r}\frac{\partial}{\partial r} r \frac{\partial C}{\partial r} + D_{\text{gel}}\frac{\partial^2 C}{\partial z^2} \tag{25}$$

where $u(r)$ is the axial electrophoretic velocity distribution in the gel. We can estimate the contribution of axial velocity distribution to the variance by dropping the diffusion terms, assuming that the injection volume is infinitesimally small and then using moment analysis to compute the variance. On eliminating the diffusive terms, the mass transport equation reduces to

$$\frac{\partial C}{\partial t} + u(r)\frac{\partial C}{\partial z} = 0 \tag{26}$$

A very thin volume of concentrated solute injected onto the column can be represented by the Dirac distribution (impulse function) $C_o\delta(z)$ as an initial condition and the resulting system readily solved (e.g., using Laplace transforms) to obtain

$$C(t, r, z) = \begin{bmatrix} C_o\delta\left(t - \dfrac{z}{u(r)}\right) & \text{when } \dfrac{z}{u_{\max}} < t < \dfrac{z}{u_{\min}} \\ 0 & \text{when } \dfrac{z}{u_{\max}} > t > \dfrac{z}{u_{\min}} \end{bmatrix} \tag{27}$$

These solutions are then used to calculate the radially av-

eraged flux J and the temporal moments M_n at the outlet of the gel. The formulas for these are

Radially averaged flux

$$J(t, z) = \int_{a_t}^{a_o} u(r)C(t, r, z)rdr \Big/ \int_{a_t}^{a_o} rdr \quad (28)$$

$$\text{Temporal moments} \quad M_n(z) = \int_0^{\infty} J(t, z)t^n dt \quad (29)$$

The electrophoretic velocity distribution $u(r)$ varies with the buffer viscosity, which is a function of the temperature. For relatively small temperature variations (i.e., less than 10 °C), the velocity may be calculated by using a Taylor expansion in the temperature rise to obtain the formula

$$\begin{aligned} u(r) &= \frac{v(T_{ref})}{v(T)}\mu(T_{ref})E \\ &= \mu(T_{ref})E\left[1 - (T - T_{ref})\left(\frac{d\log(v)}{dT}\right)_{T_{ref}}\right] \\ &= \mu(T_{ref})E[1 - \alpha(T - T_{ref})] \end{aligned} \quad (30)$$

where μ is the electrophoretic mobility and v is the viscosity of the buffer.

Strictly speaking, the temperature profiles are given by equation 20 for annular gels. However, to generate a relatively simple analytical form for the fluxes, Galerkin's method has been used to find an approximate solution of equation 19 in the form of a parabolic polynomial in η

$$\Theta(\eta) \approx A_0[Bi_b(\eta - \eta^2) + 1] \quad (31)$$

where

$$\begin{aligned} A_0 = &-7[10(\Lambda^2 + \Lambda)(Bi_b + 6) + 3Bi_b + 20] \\ &\div \{14(\Lambda^2 + \Lambda)[Bi_b^2(\kappa^2 - 10) + 10Bi_b(\kappa^2 - 6) + 30\kappa^2] \\ &+ Bi_b^2(4\kappa^2 - 49) + 14Bi_b(3\kappa^2 - 25) + 140\kappa^2\} \end{aligned}$$

To simplify the analysis somewhat, only the symmetric form of the solution for the temperature profile is considered here. Although this introduces a small error in the variance computed for thick-gap devices, as seen in Figure 7, it greatly simplifies the analysis, allows closed-form solution of the resulting equations, and gives the correct qualitative tendency as the power density is increased.

For an electrophoretic velocity profile that is symmetric about the centerline of the annulus and that has the parabolic form $u(\eta) = A + B\eta + C\eta^2$, the radially averaged concentration and flux measured at position Z are

$$C(t, Z) = \begin{cases} \dfrac{\left(B\sqrt{t} + \sqrt{B^2t - 4C(At - Z)}\right)Z}{Ct^2\sqrt{B^2t - 4C(At - Z)}(a_o^2 - a_i^2)} C_o \\ 0 \end{cases}$$

$$\begin{aligned} &\text{when } t_{min} = \frac{Z}{u_{max}} < t < \frac{Z}{u_{min}} = t_{max} \\ &\text{when } t_{min} = \frac{Z}{u_{max}} > t > \frac{Z}{u_{min}} = t_{max} \end{aligned} \quad (32)$$

$$J(t, Z) = \begin{cases} \dfrac{\left(B\sqrt{t} + \sqrt{B^2t - 4C(At - Z)}\right)Z^2}{Ct^3\sqrt{B^2t - 4C(At - Z)}(a_o^2 - a_i^2)} C_o \\ 0 \end{cases}$$

$$\begin{aligned} &\text{when } t_{min} = \frac{Z}{u_{max}} < t < \frac{Z}{u_{min}} = t_{max} \\ &\text{when } t_{min} = \frac{Z}{u_{max}} > t > \frac{Z}{u_{min}} = t_{max} \end{aligned} \quad (33)$$

where the parameters, A, B, and C, calculated for the annular columns, are given in Table 4.

Equation 28 for the radially averaged flux is used to compute the first three temporal moments, and these are then used to compute the mean elution time and the temporal variance. The results of these calculations are summarized in Figure 9, which shows the reduction in plate efficiency due to thermal dispersion in pGE as a function of the power density.

The most interesting aspect of the thermal dispersion is its dependence on column length. As shown in Table 5, the variance associated with thermal dispersion varies with the square of the column length rather than the first power, as is common with convective–diffusive dispersive

Table 4. Parameters Used in the Preparative Electrophoresis Moment Calculations

Annular gel ($Bi_a = Bi_b$)
$A = +\mu(T_{ref})E\left[1 - \alpha(T_{ref})A_0\dfrac{\sigma(T_{ref})E^2b^2}{k}\right]$
$B = -C = -\mu(T_{ref})E\left[\alpha(T_{ref})A_0\dfrac{\sigma(T_{ref})E^2b^2}{k}Bi_b\right]$
$t_{min} = Z/(A + B/2 + C/4)$
$t_{max} = Z/(A + B + C)$

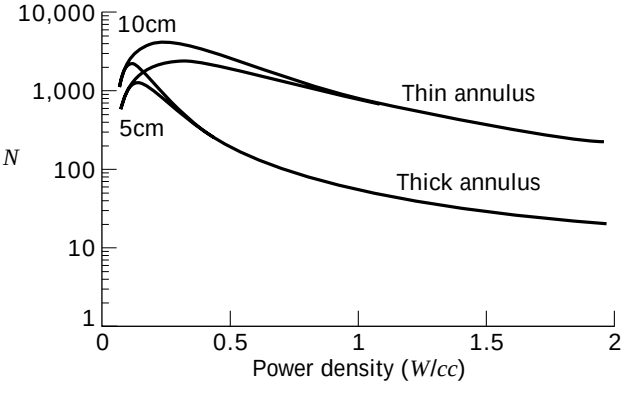


Figure 9. Plate efficiency as a function of power density in the preparative GE. The curves that meet on the upper left are calculated for $Z = 10$ cm; those on the lower left are for $Z = 5$ cm. The curves that meet on the lower right are calculated for the thick annulus; those that meet on the upper right are for the thin annulus. The effect of sample injection volume is not included in these calculations, but would tend to substantially reduce the plate efficiency of the shorter column.

Table 5. Mean Residence Time and Variance Due to Thermal Dispersion in pGE

$$t_{\text{mean}}(Z) = Z\frac{M_1(Z)}{M_o} = \frac{72Z}{72A + 11B}$$

$$\sigma^2(Z) = \frac{M_2(Z)}{M_0} - \frac{M_1(Z)^2}{M_0^2} = 192Z^2\left(\frac{\log\left[\sqrt{\frac{4A + B}{4A}}\right]}{B(72A + 11B)} + \frac{\log\left[\sqrt{\frac{4A + B}{4A}} + \sqrt{\frac{B}{4A}}\right]}{\sqrt{B(4A + B)}(72A + 11B)} - \frac{27}{(72A + 11B)^2}\right)$$

processes. As a result, thermal dispersion dominates other sources of dispersion in long columns and at high power densities, as shown in Figure 9. As a rule of thumb, the optimum power density in existing commercial equipment for pGE should occur at a characteristic temperature rise, $\sigma E^2(a_o - a_i)^2/k_{gel} \approx 8$ °C.

Although the formulas given are approximate to the extent that molecular diffusion is ignored and only symmetric parabolic temperature distributions are admitted, they provide an excellent qualitative description of the available pGE columns and a rational quantitative basis for designing new units or upgrading existing columns. Figure 9 shows that the thick annular chamber must operate in a very narrow power dissipation range to obtain reasonable plate efficiencies. If higher resolution is needed, the user has a choice of reloading the thin-annulus device several times or building a new instrument with a larger perimeter.

Solute Dilution in the Elution Chamber. As solute leaves the gel phase, it enters the elution chamber and is then flushed into the elution stream. There is some loss of resolution at this point, but the main consideration here is dilution of the fractionated samples as they come off the gel. The dilution factor F_{dilution} is calculated by taking the ratio of the solute velocity J/C as it comes off the gel to the elution stream velocity Q/A_c. This is

$$F_{\text{dilution}} = \frac{J(Z, t)}{C(Z, t)}\frac{A_c}{Q} = \frac{Z}{t_{\text{elution}}}\frac{A_c}{Q} \gtrsim \frac{\langle u_{\text{gel}}\rangle}{\langle u_{\text{buffer}}\rangle} \qquad (34)$$

from which it is seen that the dilution factor is a function of elution time.

The approximate inequality on the far right has been included to give perspective to the dilution factor. This approximation is based on the assumption that one would like to operate under conditions in which no solute is captured by the electrodes in the elution chamber. Roughly speaking, the largest attainable dilution factor (i.e., that giving least dilution) is equal to the ratio of the electrophoretic mobility in the gel near the outlet to the electrophoretic mobility in the elution buffer. The utility of this expression is that it may be rearranged to compute the lowest admissible value of the elution stream flow rate Q

$$Q \gtrsim Q_{\text{min}} = \frac{ZA_c}{t_{\text{elution}}}\frac{\langle u_{\text{buffer}}\rangle}{\langle u_{\text{gel}}\rangle} \approx A_c\langle u_{\text{buffer}}\rangle \qquad (35)$$

Then, if the electrophoretic velocity of a solute in free solution is known, the elution buffer flow rate can be set very near this minimum value. If a collection of solutes is being processed at a constant elution flow rate, then Q should be calculated using the mobility of the fastest solute (i.e., the first one off the gel) to avoid the loss of fast solutes to the electrode.

Note that the dilution factor depends not only on the tightness (i.e., the reduction of the solute's diffusion coefficient) of the gel being run but also on the individual electrophoretic mobilities of the solutes being processed. This implies that, to keep the dilution factor constant during a run, the elution buffer flow rate would need to be decreased in inverse proportion to time. This latter type of correction could be important in applications where peak width or height must be accurately measured, for example, so that a direct comparison of bands may be made with those on a standard gel. It may also be important in minimizing dilution of the slower solutes.

Remarks. The continuous elution pGE annular apparatus is capable of processing several hundred times more solute in a single pass than the tubular apparatus described earlier. To accomplish this, some resolution is sacrificed for capacity, and because the optimal run conditions in the annulus are found at a fraction of the power density used in the tube the runs generally take longer to complete. It is up to the user to decide for a given separation whether this trade-off between time and resolution is worthwhile.

PREPARATIVE CONTINUOUS-FLOW ELECTROPHORESIS

Thin-Film Apparatus

When still greater amounts of material must be processed, the apparatus that has the next largest capacity is a thin-film continuous electrophoresis device, such as the VaP 22, which was until recently available through Bender-Hobein, FRG. This type of apparatus, which dates back to the early 1960s, consists of a thin film of buffer that flows down between two closely spaced (e.g., 0.5-mm transverse) vertical glass plates about 20 cm across (lateral) and 50 cm long (axial), as illustrated in Figure 10. The CFE is usually run in free-flow mode, that is, without anticonvectant media.

Multicomponent sample is injected into the carrier buffer at the top of the column through a thin capillary. The solutes are convected down with the buffer, flow between electrodes that are located at the lateral edges of the plates, and are deflected by the electric field according to their electrophoretic mobilities. As the fractionated solutes

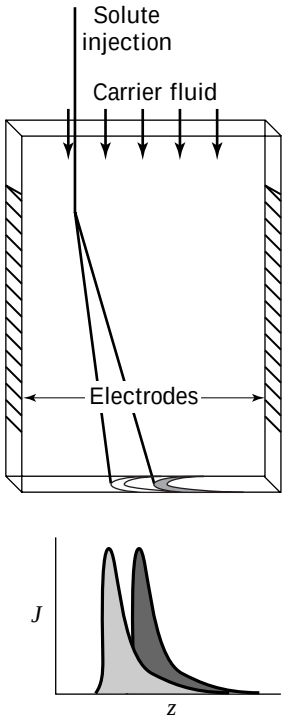


Figure 10. Schematic of the CFE showing electrode placement, multicomponent sample injection, carrier fluid flow, and product distribution. Typical elution bands are shown beneath the chamber corresponding to the two crescent-shaped bands on the lower surface of the chamber. The skewing is a deformation of the sample bands referred to as the crescent phenomena (2,40), which is caused by the interaction of hydrodynamic and electrokinetic effects.

exit the electric field, they are swept into discrete effluent ports by the carrier buffer and collected in separate vials (41).

A major advantage of the CFE over equipment described earlier is that it operates in true continuous mode, whereas the others operate in semibatch mode (i.e., samples are loaded batchwise, but product fractions are eluted continuously). Processing rates can approach 50 mg/h, and theoretical plate efficiencies can in theory exceed $N = 10{,}000$, although in practice they are generally somewhat lower.

Fundamental Principles. The first step in analyzing the performance of the CFE is to calculate the temperature profile along its thin (transverse) axis. The governing differential equation is

$$k_{\text{buffer}} \frac{d}{dy}\left(\frac{dT}{dy}\right) + \sigma_{\text{buffer}}(T_{\text{ref}})E^2[1 + \lambda(T - T_{\text{ref}})] = 0 \quad (36)$$

with the boundary conditions

$$\text{at } y = \pm b$$
$$-k_{\text{buffer}} \frac{dT}{dy} = \pm h_{\text{CFE}}(T - T_{\text{ref}}) \approx \pm \frac{k_{\text{glass}}}{d_{\text{glass}}}(T - T_{\text{ref}}) \quad (37)$$

Here it is assumed that both transverse surfaces are cooled, that the major resistance to heat transfer is the thickness of the transverse surfaces, and that the heat transfer coefficients are equal on both sides of the unit. Note that asymmetric conditions could significantly impair CFE performance, so every effort should be made to match the electrokinetic and thermal properties of these surfaces.

Equation 36 with symmetric boundary conditions as in equation 37 has the solution

$$\Theta_{\text{CFE}}(\eta) = \frac{1}{\kappa^2}\left[\frac{Bi\cos(\kappa\eta)}{Bi\cos(\kappa) - \kappa\sin(\kappa)} - 1\right] \quad (38)$$

where the dimensionless parameters used are

$$\Theta_{\text{CFE}}(\eta) = \frac{k_{\text{buffer}}(T - T_{\text{ref}})}{\sigma_{\text{buffer}}E^2b^2} \quad \eta = \frac{y}{b} \quad Bi = \frac{h_{\text{CFE}}b}{k_{\text{buffer}}}$$
$$\kappa^2 = \lambda\frac{\sigma_{\text{buffer}}E^2b^2}{k_{\text{buffer}}}$$

and autothermal runaway occurs when $Bi\cos(\kappa) - \kappa\sin(\kappa) = 0$. The parabolic approximation to equation 38 obtained, using Galerkin's method is

$$\Theta_{\text{CFE}}(\eta) \approx \frac{5(Bi + 2)(Bi + 3)}{2[5Bi(Bi + 3) - \kappa^2(2Bi^2 + 10Bi + 15)]} \cdot \left(1 - \frac{Bi\eta^2}{Bi + 2}\right) \quad (39)$$

and the predictions of both of these formulas are shown in Figure 11 for typical operating conditions. Because of the symmetry inherent in this problem, the exact and approximate solutions differ by only a few percent over most of their useful range. The agreement between equations 38 and 39 breaks down as κ approaches the critical value for autothermal runaway.

Axial Velocity Profile. Because the CFE does not normally use an anticonvective packing, the velocity profiles

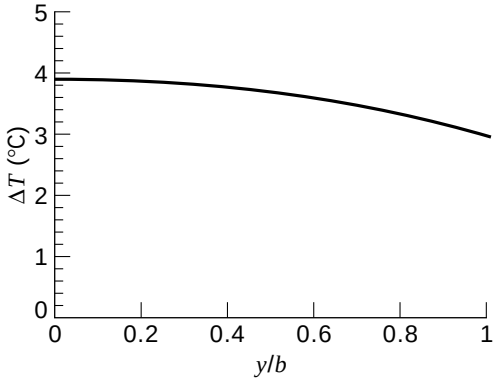


Figure 11. Predicted CFE temperature rise, in degrees centrigrade from the center line at $y = 0$ to the wall at $y = b$. Since the bounding transverse walls dominate heat transfer resistance, there is a 3 °C rise above the coolant temperature at the transverse wall. The temperature difference across the slit is less than 1 °C.

in the axial and lateral directions have a significant impact on performance. Along the x axis, the velocity profile between the front and back plates is, to a first approximation, laminar and parabolic. However, because of the small temperature difference between the walls and the centerline, there is a slight distortion of the flow due to stable natural convection (42,43), which can either improve or degrade performance. Assuming that the base flow is parabolic, the perturbation to the flow is found by solving the Boussinesq (44,45) approximation to the equations of motion

$$\nu_{\text{buffer}} \frac{d}{dy}\left(\frac{dv}{dy}\right) + \rho_{\text{buffer}} g\beta(T - T_{\text{ref}}) = \frac{\Delta P'}{L} \qquad (40)$$

where

$$\beta = \left[-\frac{d\,\log(\rho_{\text{buffer}})}{dT}\right)_{T=T_{\text{ref}}}$$

with no-slip conditions on the bounding surfaces

$$\text{at } y = \pm b \quad v(y) = 0 \qquad (41)$$

and subject to the integral constraint that there is no net increase in the average velocity associated with natural convection, that is,

$$\int_{-b}^{b} v(y)dy = 0 \qquad (42)$$

Note that P' is a perturbation pressure that determined by the integral constraint of equation 42. When the exact form for the temperature profile (see equation 46) is used, this system of equations yields the solution

$$v(y) = \frac{3}{2}\langle v\rangle\left((1 - \eta^2) + \frac{Gr}{Re} \cdot \frac{Bi\{3(\eta^2 - 1)[\sin(\kappa) - \kappa\cos(\kappa)] + 2\kappa[\cos(\kappa\eta) - \cos(\kappa)]\}}{3\kappa^5[Bi\cos(\kappa) - \kappa\sin(\kappa)]}\right) \qquad (43)$$

while the polynomial approximation yields

$$v(y) = \frac{3}{2}\langle v\rangle\left\{(1 - \eta^2) + \frac{Gr}{Re} \cdot \frac{Bi(Bi + 3)[5(1 + \eta^2) - 6](1 - \eta^2)}{36[\kappa^2(2Bi^2 + 10Bi + 15) - 5Bi(Bi + 3)]}\right\} \qquad (44)$$

where

$$Re = \frac{\rho_{\text{buffer}}\langle v\rangle b}{\nu_{\text{buffer}}} \quad \text{and} \quad Gr = \frac{\rho_{\text{buffer}}^2 g\beta}{\nu_{\text{buffer}}^2} \frac{\sigma_{\text{buffer}} E^2 b^5}{k_{\text{buffer}}}$$

Under normal CFE operating conditions the ratio of the Grashof number to the Reynolds number in equations 43 and 44 is somewhat less than 1, and, as illustrated in Figure 12, the functional portion of the perturbation velocity never exceeds ± 0.01, so the contribution of natural convection in commercial CFEs, $b < 1$ mm, can be neglected when $Gr/Re < 1$. Note, however, that $Gr/Re \propto b^4$, so this term rapidly becomes important as the slit thickness is increased.

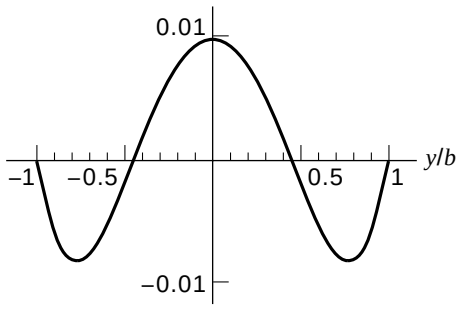


Figure 12. Predicted perturbation velocity profile due to stable natural convection in a slit. Both the exact solution, equation 43, and the polynomial approximation, equation 44, are shown here. The warmer fluid in the center of the slit rises while the cooler fluid near the wall falls.

Lateral Velocity Profile. Because the transverse surfaces of the CFE carry a fixed charge, the electric field induces wall electroosmosis, which drives a bulk recirculating flow along the lateral axis. The shape and magnitude of this flow can be calculated from the equation of motion

$$\nu_{\text{buffer}} \frac{d}{dy}\left(\frac{dw}{dy}\right) = \frac{\Delta P''}{L} \qquad (45)$$

together with the integral constraint

$$\int_{-b}^{b} w(y)dy = 0 \qquad (46)$$

Assuming that the induced electroosmosis is identical on both transverse surfaces, perhaps due to application of a surface coating, a suitable set of first-kind boundary conditions is

$$\text{at} \quad y = \pm b \quad w(y) = w_{\text{eo}} \qquad (47)$$

Solving the differential equation 45 subject to boundary conditions as in equation 47 and constraint from equation 46 yields the polynomial

$$w(\eta) = -\frac{3}{2} w_{\text{eo}}(1 - \eta^2) + (w_{\text{eo}} + \langle u\rangle) \qquad (48)$$

where the electrophoretic velocity of the solute $\langle u\rangle$ has been tacked on to account for displacement of the solute relative to the electroosmotic flow. Note that when $w_{\text{eo}} = -u$, the term on the far right vanishes and the lateral flow becomes parabolic.

Solute Dispersion in the CFE. In practice, a multicomponent feed is injected into the CFE carrier flow through a capillary that extends over one-third to one-half of the transverse thickness of the chamber. In most instances it

is possible to neglect the effects of molecular diffusion on dispersion because convective distortion of the sample band dominates band spreading. The criterion for neglecting molecular diffusion is that the time required for a molecule to diffuse from the centerline to the transverse will be much greater than the solutes' average holdup time in the chamber. This ratio is characterized by the dimensionless Graetz number

$$Gz = \frac{t_{\text{diffusion}}}{t_{\text{convection}}} = \frac{b^2/D_{\text{m}}}{L/\langle v\rangle} = \frac{\langle v\rangle b}{D_{\text{m}}} \cdot \frac{b}{L} \geq 10$$

which for proteins typically has values well in excess of 10.

Because this apparatus operates continuously and solutes are separated in space along the lateral axis rather than in time, it is appropriate to use spatial moments rather than temporal moments to characterize performance. If diffusive dispersion is ignored, the spatial moments can be found using the formula

$$M_n(L) = \left(\frac{L}{b}\right)^n C_{\text{o}} \int_{-\eta^*}^{+\eta^*} \frac{w(\eta)^n}{v(\eta)^{n-1}}\, d\eta \tag{49}$$

where η^* is the fraction of the transverse thickness over which sample injection extends. In this calculation the feed band is assumed to be infinitesimally thin along the lateral axis. Neglecting the effects of natural convection in equation 44, we can readily generate analytical expressions for the mean lateral displacement and spatial variance.

Table 6 gives analytical formulas for the mean lateral displacement, the spatial variance, the theoretical plate number, and the resolution in the CFE, taking into account only pure convective dispersion. Several points are brought to light in the moment analysis that are crucial to CFE operation. First, as is clear from the premultiplicative terms in the variance and the plate number, performance is optimized when $\langle u\rangle = -w_{\text{eo}}$, a condition that can be approximated by coating the transverse surfaces with a material that has a zeta potential similar to that of the key solutes being processed (2).

Second, the plate number N_{CFE} is a function neither of column length nor of the carrier fluid velocity, and is therefore not a function of residence time. This occurs because, like thermal dispersion, convective band spreading due to electroosmosis is faster than a diffusion-like process; that is, CFE variance increases with the first power of the residence time rather than the square root.

Third, as $\eta^* \to 1$, the variance increases rapidly, and, from a purely theoretical standpoint, if diffusive effects are ignored it becomes infinite in this limit. Performance degrades rapidly as η^* approaches 1, so that there is a stark trade-off between capacity and resolution. Although plate efficiencies in excess of $N = 20{,}000$ can be obtained in continuous operation, this is only possible under rather stringent conditions. However, for $\eta^* < 3/4$ and $Gz > 100$, the analysis given earlier should be sufficiently accurate to allow the user to optimize a given separation.

The contribution of the breadth of the sample injection stream to the variance is

$$\chi^2_{\text{injection}} = \frac{W^2}{12} \tag{50}$$

where W is the lateral breadth of the injected stream. This expression is readily incorporated into the calculation for the plate efficiency N by adding this contribution to the spatial variance in Table 6.

Figure 13 is a three-dimensional plot of N versus $\langle u\rangle$ and η^* for $w_{\text{eo}} = -0.01$, which shows how CFE performance varies with operating conditions. Here we see that N is greatest along the line $w_{\text{eo}} = -\langle u\rangle$ and for small values of η^*, but for reasonable injection capacities ($\eta^* > 1/3$) rarely exceeds 20,000 plates. In practice one should choose conditions that give at least baseline resolution, $R_{\text{CFE}} > 1$, of key components and differences in the mean lateral displacements z, which are at least as broad as the effluent ports.

Solute Dilution in CFE. For the CFE, the dilution factor that corresponds to that discussed with the GE systems is determined from the size of the sample injection stream as $F_{\text{dilution}} = \eta^{*-1}$, and typically ranges from 2 to 4 depending on the injection flow rate. However dilution by band smearing along the lateral axis in the CFE is often an important consideration and is relatively easy to estimate by using the formula

$$F_{\text{gross}} = \frac{1}{\eta^*} \frac{\sqrt{\chi^2_{\text{injection}} + \chi^2}}{\chi_{\text{injection}}} \tag{51}$$

where $\chi_{\text{injection}}$ is given by equation 50 and χ comes from Table 6. The numerator in equation 51 is proportional to the breadth of the product in the effluent stream, and the

Table 6. Statistical Expressions for the CFE Subject to Pure Convective Dispersion

Mean displacement	$z_{\text{mean}}(x) = x\,\dfrac{w_{\text{eo}}(\eta^{*3} - \eta^*) + 2\eta^*\langle u\rangle}{(3\eta^* - \eta^{*3})\langle v\rangle}$
Spatial variance	$\chi^2(x) = x^2\,\dfrac{(w_{\text{eo}} + \langle u\rangle)^2}{\langle v\rangle^2}\left(\dfrac{2}{3\eta^*(3 - \eta^{*2})}\log\left[\dfrac{1 + \eta^*}{1 - \eta^*}\right] - \dfrac{4}{(3 - \eta^{*2})^2}\right)$
Plate number	$N_{\text{CFE}} = \dfrac{z^2_{\text{mean}}}{\chi^2} = \dfrac{3\eta^*(w_{\text{eo}}(\eta^{*2} - 1) + 2\langle u\rangle)^2}{2(w_{\text{eo}} + \langle u\rangle)^2\left[(3 - \eta^{*2})\log\left(\dfrac{1 + \eta^*}{1 - \eta^*}\right) - 6\eta^*\right]}$
Resolution	$R_{\text{CFE}} = \dfrac{\Delta z}{2\chi} \approx \sqrt{N_{\text{CFE}}}\left\lvert\dfrac{\Delta\langle u\rangle}{w_{\text{eo}}(\eta^{*2} - 1) + 2\langle u\rangle}\right\rvert$

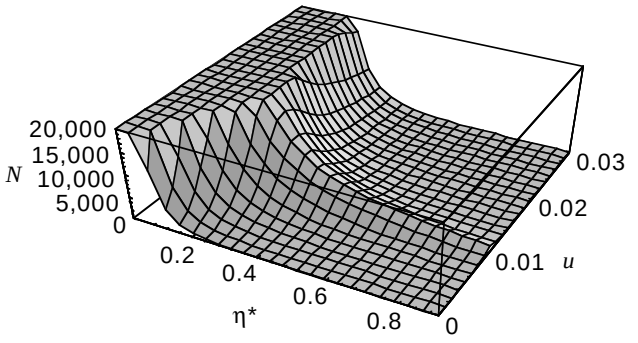


Figure 13. Isothermal plate efficiency N as a function of injection cross-section η^* and electrophoretic velocity u. N was computed for $w_{eo} = -0.01$ cm/s, $v = 1$ cm/s, $L = 100$ cm, and $2b = 0.1$ cm. In this figure it is assumed that a square injection capillary is used. This shows an optimal plate efficiency when $u = -w_{eo}$ but illustrates how small the operating regime is where plate efficiencies are greater than about 20,000 in the CFE. Note that this figure has been capped at $N = 20{,}000$.

denominator is proportional to the breadth of the sample injection stream. The premultiplicative factor $1/\eta^*$ is simply the dilution factor given earlier. As shown in Figure 14, minimum dilution occurs when $\langle u \rangle = -w_{eo}$, with the broadest operating region centered near $\eta^* = 1/3$. In practice, plate efficiency and dilution must be weighed against capacity.

Continuous-Flow Electrophoresis with Recycle (RCFE)

Effluent recycle was introduced for isoelectric focusing by Bier and Egen in 1979 (46). At that time it was believed that recycle could not be adapted to zone electrophoresis because of fundamental differences in the transport processes governing these two techniques. Specifically, isoelectric focusing is an equilibration process in which solutes placed in a pH gradient migrate to the point in the gradient where they have zero net charge. Separation depends primarily on the steepness of the pH gradient and the difference in isoelectric points (pIs) of the solutes. Zone electrophoresis, on the other hand, is a rate process in which solutes migrate at fixed velocities, so separation depends primarily on the length of the run, the electric field strength, and the difference in electrophoretic mobilities.

Recycle was first applied to zone electrophoresis in 1983 (47,48). The major obstacle that had to be overcome to accomplish this was to find a way to get two solutes with like charges but different electrophoretic mobilities to move in opposite directions. Our solution to this problem works in the following way: Two proteins with similar electrophoretic mobilities can be made to migrate in opposite directions by opposing the electrophoretic motion with a counterflow velocity set close to the average of the electrophoretic velocities of two key components.

Solutes migrate at their effective velocities (i.e., electrophoretic plus counterflow) until they reach the "regenerators" (see Figure 15) on either side of the recycle section. In the regenerators the counterflow of solvent is altered by means of the recycle lines, to keep solutes from migrating past the outlet port. At the regenerator on the left, the cross-flow rate is decreased; it is increased at the regenerator on the right and adjusted so that solutes cannot escape the offtake ports. With proper design, complete separation can be achieved with virtually no loss of biological activity.

The RCFE was set up so that on each cycle through the unit the solutes would separate slightly. However, after recycling through the chamber 100 or more times, the solutes are completely separated and can be continuously drawn off at either end of the slit with very low losses. The ultimate resolution of solutes in the RCFE depends largely on the number of recycle stages designed into the chamber. Because the design engineer has complete control of this variable, any degree of product purity can, in principle, be attained.

The primary advantage of this configuration is that it can process feeds in excess of 1 g/h and operates in true

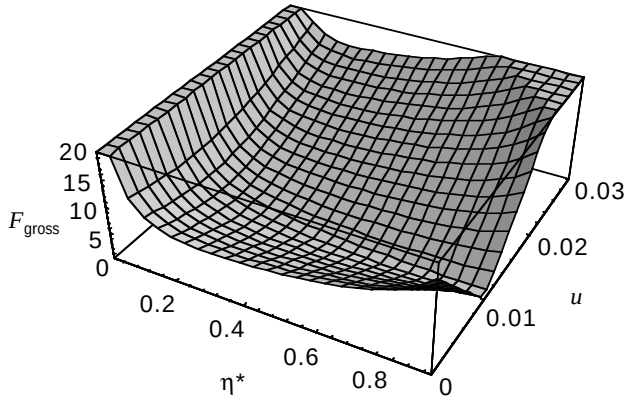


Figure 14. Gross dilution factor vs. η^* and u. This function gives the degree of dilution a solute will undergo during processing in the CFE. As seen in the center of the surface plot, under near-optimal operating conditions it is possible to obtain gross dilution factors below 4.

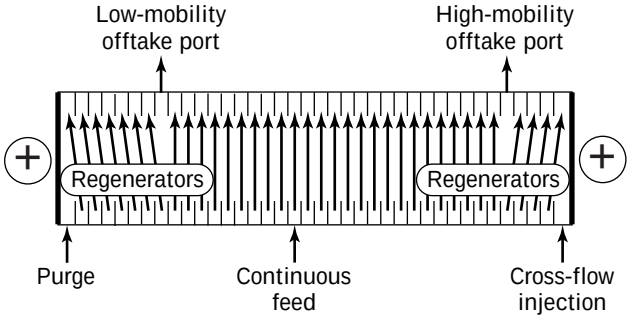


Figure 15. The RCFE with regenerators. The regenerators keep solute from migrating past the outlet ports, and this helps keep solute losses to a minimum. Note that opening up offtake ports automatically creates the regenerators and that the magnitude of the shift and the strength of the countercurrent flow is altered at each regenerator. The arrows within the chamber indicate the connection between the inlet and outlet ports. Note the change in plumbing in the right and lefthand regenerators.

continuous mode. Its major drawback is that it will split a multicomponent feed into just two products—one that contains components that are faster than the cross-flow and the other one having components that are slower than the cross-flow. Consequently, complex feeds will generally require a minimum of two passes to isolate the desired product: The first pass could be used to remove faster contaminants and the second pass to remove slower contaminants.

Separation is carried out in native buffer with background conductivities near 0.01 mho/cm (50 mM buffer), so low protein solubility is not generally a problem. Power requirements are kept low by using small electric fields (less than 40 V/cm), so the temperature rise in the chamber on a single pass is typically less than 5 °C. A multichannel heat exchanger removes joule heat from the recycle streams and sets the base temperature of the separation. In addition, the RCFE is hydrodynamically stable because the chamber is operated adiabatically with upflow.

The RCFE is not commercially available. However, commercial recycle IEF chambers, such as Rainin's MinipHor, can be readily modified to operate in RCFE mode.

Remarks. When the RCFE was first conceived, it was believed that zone electrophoresis, which can provide fine resolution of proteins in analytical equipment, would find its niche in the final polishing stages of industrial purification processes. In practice, this apparatus appears to be more useful in the early stages of downstream processing, where complex mixtures must be handled and emphasis is placed on removing solutes that are difficult to handle by other means, for example due to fouling or precipitation, or removing bulk proteins (such as serum albumin from blood plasma) inexpensively and without organic solvents or large amounts of precipitating salts.

DISCUSSION

Since 1980, the field of applied electrophoresis has undergone revolutionary changes, spurred in large part by the development and automation of capillary electrophoresis. Instrumentation is available for application of zone electrophoresis at scales up to pilot, so it is now possible to decide, on a rational basis, whether to incorporate electrophoresis as part of a purification train.

One could imagine, for instance, applying CE (along with analytical HPLC) early in process development to determine which contaminants could be removed from a product stream by electrophoresis and the optimum pH and buffer conditions at which this might be carried out. The results of this inexpensive and rapid preliminary screening step would allow the user to decide whether zone electrophoresis should be considered at all. As a rule of thumb, if the mobility difference between key components, typically the desired product and a major contaminant, measured by CE is greater than 25%, then zone electrophoresis may be considered further.

Following CE, one could scale up to native pGE in order to recover enough purified protein to check for product heterogeneity, activity, and toxicity, as well as to determine the amino acid sequence. When carried out in PAGE, the resolution achieved is boosted by the filtering power of the gel; however, separations conducted in native agarose gels can give insight to the presence (or absence) of problems with zone electrophoresis at larger scales. As before, one should look for a 25% difference in mobilities of key components before considering continuous zone electrophoresis as a possible process step. Mobility differences greater than 25%, as measured in native agarose gels, should yield a straightforward and clean separation by zone electrophoresis at pilot scale. This preparative step is somewhat more expensive and time-consuming than CE but is a better predictor of success or failure at larger scales. If an automated GE unit is not available, the user can always resort to vertical or horizontal slab gels.

Assuming the preparative step is successful, further scale-up requires a return to free-solution zone electrophoresis in either the CFE or the RCFE. (*Note*: In this step one can also use an RF3 or MiniPhor that has been outfitted for isotachophoresis. However, both the equipment itself and the protocols used must be modified for zone electrophoresis. The advantage of using this equipment is that it is commercially available and relatively inexpensive.) This phase of the study is again more expensive and time-consuming than previous ones, and should be carried out at processing rates between 10 and 100 mg/h. In free solution, the user is forced to confront potential problems that might arise as a result of heat generation or electroosmotic dispersion, hopefully for the last time during scale-up. As the conductivity of the buffer is reduced, the electric power and chamber cooling requirements decrease, and throughput generally increases because electrophoretic mobilities increase. However, as the conductivity is reduced, wall electroosmosis eventually increases to the point where the separation is ruined; an optimum conductivity should thus be sought.

In our experience, it is best to begin this search with a buffer that has a conductivity near 500 μmho/cm, which corresponds roughly to 10 mM Tris-acetate at pH 8.0, and raise the electrolyte concentration if it is necessary to reduce electroosmotic dispersion or to discourage protein aggregation. If the separation is successful at 500 μmho/cm, then the conductivity should be halved and the experiment repeated until a suitable electrolyte concentration is found. Keep in mind that whether Tris or more exotic electrolytes or additives are used, these materials are likely to constitute the major electrophoretic processing costs in zone recycle electrophoresis rather than, say, electric power or cooling.

Closing Remarks. Over the past two decades there have been a number of important developments in the area of analytical electrophoresis, the most important and most active of these centering on the commercialization of capillary electrophoresis. In addition, gel electrophoresis has metamorphosed from a labor-intensive laboratory technique to one where the instrumentation is automated and the gels are mass-produced, changes that significantly reduce operator training and improve reproducibility.

During this same period, preparative and larger-scale zone electrophoresis has remained largely stagnant, with few innovations surviving to commercialization. There is no agreement over the precise reasons for the lack of interest in electrokinetic processing, although it is clear that the biotech industry is convinced that the tools it already has, specifically the chromatographies, are superior to those currently available via electrophoresis. It will be nearly impossible to move this technology forward without the cooperation and support of the industry it will serve.

What will it take to change this mindset? In my opinion it will take at least one case study where electrophoresis shows itself to be clearly superior to existing separations technologies and offers a straightforward route to scale-up.

A number of different electrophoretic devices have not been discussed in this article. For example, neither Hoefer's Isoprime® (49–52) nor BioRad's Rotofor® (53,54) are discussed because they are designed strictly for isoelectric focusing, a topic treated elsewhere in this book. Margolis' Gradiflo® (10,55,56) is not treated both because it is not yet commercially available in this country and because I am not sufficiently familiar with that device to give it due justice.

CACE (57) and various other forms of electrochromatography (58–61) have been omitted here because they could not be adequately treated in the space allotted to this article. Other recent innovations by Shiue and Pearlstein (62), Lochmüller and Ronsick (63), and Shea et al. (64) were omitted from this article because they are still in the early stages of development.

NOMENCLATURE

Notation

a	CE capillary lumen radius
a_g	Glass outer radius
a_{gel}	Gel radius
a_i	Annulus inner radius
a_o	Annulus outer radius
a_p	Polyimide outer radius
A_c	Gel cross-sectional area
b	CFE slit half-thickness
Bi	Biot number
C	Solute concentration
C_o	Solute concentration in sample
$C(Z, t)$	Solute concentration at Z
d_{glass}	CFE glass wall thickness
d_i	Inner diameter of annulus cooling finger
d_o	Outer diameter of annulus glass tube
d_t	Diameter of elution tube
D_m	Molecular diffusivity in free solution
D_{gel}	Molecular diffusivity in gel
D_{CE}	CE dispersion coefficient
E	Electric field strength
$F_{dilution}$	Dilution factor
Gr	Grashof number
Gz	Graetz number
g	Gravitational acceleration
h	Slit thickness
h_{CE}	CE overall heat transfer coefficient
h_{CFE}	CFE overall heat transfer coefficient
h_{fluid}	Cooling fluid heat transfer coefficient
h_{pGE}	pGE heat transfer coefficient
$HETP$	Height equivalent to a theoretical plate
$hetp$	Reduced $HETP$
$J(Z, t)$	Solute flux at Z
J_0	Bessel function of first kind, order 0
J_1	Bessel function of first kind, order 1
k	Thermal conductivity
L	Elution tube length
N	Theoretical plate number
M_n	nth moment, averaged across thin axis
P	Pressure
Pe	Peclet number
q	Sample injection volume
Q	Elution flow rate
r	Radial coordinate
R	Resolution
Re	Reynolds number
S	pGE scale factor
T	Temperature
t	Time
t_{load}	CE sample injection time
u	Electrophoretic velocity
u_{max}	Maximum electrophoretic velocity in r
u_{min}	Minimum electrophoretic velocity in r
$u(r)$	Electrophoretic velocity profile in r
v	Hydrodynamic velocity
$v(r)$	Hydrodynamic velocity profile in r
w	CFE lateral velocity
w_{eo}	CFE electroosmotic wall velocity
W	Sample load width in CFE
x	Electrophoretic velocity
y	Transverse axis
Y_0	Bessel function of second kind, order 0
Y_1	Bessel function of second kind, order 1
z	Lateral or electrophoresis axis
Z	Gel column length

Greek Symbols

α	Thermal viscosity coefficient
β	Thermal volumetric expansion coefficient
δ	Dirac impulse distribution
χ^2	Spatial variance
κ	Autothermal parameter
λ	Eigenvalue or thermal conductivity coefficient
Λ	pGE geometric factor
θ	Dimensionless time

Θ	Dimensionless temperature
η	Dimensionless transverse or radial coordinate
μ	Electrophoretic mobility
μ_{eo}	Electroosmotic mobility
ν	Molecular viscosity
Φ	Electric potential
σ^2	Temporal variance
σ	Electrical conductivity
ρ	Fluid density
τ	Dimensionless time
ζ	Dimensionless distance along z

BIBLIOGRAPHY

1. R.J. Hunter, *Zeta Potential in Colloid Science*, Academic Press, New York, 1981.
2. A. Strickler and T. Sacks, *Ann. N.Y. Acad. Sci.* **209**, 497–514 (1973).
3. F.M. Everaerts, F.E.P. Mikkers, and Th.P.M. Verheggen, *Sep. Purif. Methods* **6**, 287–351 (1977).
4. P.H. Rhodes, R.S. Snyder, and G.O. Roberts, *J. Colloid Interface Sci.* **129**, 78–90 (1989).
5. J.F. Hoburg and J.R. Melcher, *Phys. Fluids* **20**, 903–911 (1977).
6. C.F. Ivory, W.A. Gobie, J.B. Beckwith, R. Hergenrother, and M. Malec, *Science* **238**, 58–61 (1987).
7. M. Bier, G.E. Twitty, and J.E. Sloan, *J. Chromatogr.* **470**, 369–376 (1989).
8. C.F. Ivory, *Electrophoresis* **11**, 919–926 (1990).
9. M. Faupel, B. Barazaghi, C. Gelfi, and P.G. Righetti, *J. Biochem. Biophys. Methods* **15**, 147–162 (1987).
10. Z.S. Horvath, G. Corthals, C.W. Wrigley, and J. Margolis, *Electrophoresis* **15**, 968–971 (1994).
11. D.A. Saville, *Phys. Rev. Lett.* **71**, 2907–2910 (1993).
12. J.C. Baygents and F. Baldessari, *Phys. Fluids* **10**, 301–311 (1998).
13. P.T. Noble, *Biotech. Progr.* **1**, 237–241 (1985).
14. J.B. Beckwith and C.F. Ivory, *Chem. Eng. Commun.* **54**, 301–331 (1987).
15. A.S. Cohen, D.R. Najarian, A. Paulus, A. Guttman, J.A. Smith, and B.L. Karger, *Proc. Nat. Acad. Sci. U.S.A.* **85**, 9660–9663 (1989).
16. A. Guttman, A.S. Cohen, D.N. Heiger, and B.L. Karger, *Anal. Chem.* **62**, 137–141 (1990).
17. J.C. Giddings, *Unified Separation Science*, Wiley, New York, 1991.
18. W.G. Kuhr and C.A. Monnig, *Anal. Chem.* **64**, 389–407R (1992).
19. Y. Xu, *Anal. Chem.* **65**(12), 426–433R (1993).
20. J.W. Jorgenson, in J.W. Jorgenson and M. Phillips eds., *New Directions in Electrophoretic Methods*, ACS Symposium Series 335, American Chemical Society, Washington, D.C., 1987, pp. 182–198.
21. E.D. Lynch and D.A. Saville, *Chem. Eng. Commun.* **9**, 201–211 (1981).
22. J.O.N. Hinckley, *J. Chromatogr.* **109**, 209–217 (1975).
23. A.E. Jones and E. Grushka, *J. Chromatogr.* **466**, 219–225 (1989).
24. W.A. Gobie and C.F. Ivory, *J. Chromatogr.* **516**, 191–210 (1990).
25. M. Albin, P.D. Grossman, and S. E. Moring, *Anal. Chem.* **65**, 489–497A (1993).
26. C. Schwer and F. Lottspeich, *J. Chromatogr.* **623**, 345–355 (1992).
27. M.P. Deutscher ed., *Guide to Protein Purification*, Academic Press, New York, 1990.
28. A.T. Andrews, *Electrophoresis: Theory, Techniques, and Biochemical and Clinical Applications*, Oxford, U.K., Clarendon Press, 1986.
29. B.D. Hames and D. Rickwood eds., *Gel Electrophoresis of Proteins: A Practical Approach*, Oxford University Press, New York, 1990.
30. J.R. Boyack and J.C. Giddings, *Arch. Biochem. Biophys.* **100**, 16 (1963).
31. A.G. Ogston, B.N. Preston, J.D. Wells, and J.McK. Snowden, *Proc. R. Soc. Lond.* **333**, 297–316 (1973).
32. P.G. De Gennes, *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca, New York, 1979.
33. D.J. Harrison, A. Manz, Z. Fan, H., Ludi, and H.M. Widmer, *Am. Chem. Soc.* **64**, 1926–1932 (1992).
34. D.J. Harrison, K. Fluri, K. Seiler, Z. Fan, C.S. Effenhauser, and A. Manz, *Science* **261**, 895–897 (1993).
35. S.C. Jacobson, R. Hergenroder, L.B. Koutny, and J.M. Ramsey, *Anal. Chem.* **66**, 1114–1118 (1994).
36. S.C. Jacobson, R. Hergenroder, L.B. Koutny, and J.M. Ramsey, *Anal. Chem.* **66**, 2369–2373 (1994).
37. B.B. Rosenblum, *J. Liquid Chromatogr.* **14**, 1017–1024 (1991).
38. G.I. Taylor, *Proc. R. Soc. Lond.* **A219**, 186–203 (1953).
39. R. Aris, *Proc. R. Soc. A* **235**, 67–77 (1956).
40. C.F. Ivory, *J. Chromatography* **195**, 165–179 (1980).
41. K. Hannig, in J.R. Norris and D.W. Ribbons eds., *Methods in Microbiology*, vol. 5B Academic Press, New York, 1971, pp. 513–548.
42. S. Ostrach, *J. Chromatogr.* **140**, 187–195 (1977).
43. D.A. Saville and S. Ostrach, TR NASA-1 Final Report on Contract NAS-8-31349 Code 361, Dept. Chemical Engineering, Princeton University, Princeton, New Jersey, 1978.
44. S. Chandrasekhar, *Hydrodynamic and Hydromagnetic Stability*, Dover Publications, New York, 1961.
45. J.M. Mihaljan, *Astrophys. J.* **136**, 1126–1133 (1962).
46. M. Bier and N. Egen, in H. Haglund, J.G. Westerfield, and J.T. Ball eds., *Electrofocus / 78*, Elsevier/North Holland, New York, 1979, pp. 35–48.
47. C.F. Ivory, W.A. Gobie, and R.S. Turk, in H. Hirai ed., *Electrophoresis '83*, Walter de Gruyter, Berlin, 1984, pp.
48. W.A. Gobie, J.B. Beckwith, and C.F. Ivory, *Biotech. Progr.* **1**, 60–68 (1985).
49. P.G. Righetti, B. Barazaghi, M. Luzenna, G. Manfredi, and M. Faupel, *J. Biochem. Biophys. Methods* **15**, 199–206 (1987).
50. P.G. Righetti, E. Wenisch, and M. Faupel, *J. Chromatogr.* **475**, 293–309 (1989).
51. P.G. Righetti, M. Faupel, and E. Wenish, in A. Chrambach, M.J. Dunn, and B.J. Radola eds., *Advances in Electrophoresis, vol. 5*, VCH, New York, 1991, pp. 159–200.
52. P.G. Righetti and A. Bossi, *Anal. Biochem.* **247**, 1–10 (1997).
53. N.B. Egen, W. Thormann, G.E. Twitty, and M. Bier, in H. Hirai ed., *Electrophoresis '83*, Walter de Gruyter, Berlin, 1984, pp. 547–550.
54. U.S. Pat. 4,588,492 (May 13, 1986), M. Bier (to University Patents, Inc., Westport, Conn.).

55. J. Margolis, G. Corthals, and Z.S. Horvath, *Electrophoresis* **16**, 98–100 (1995).
56. G. Corthals, J. Margolis, K.L. Williams, and A.A. Gooley, *Electrophoresis* **17**, 9–11 (1996).
57. P.H. O'Farrell, *Science* **227**, 1586–1589 (1985).
58. S.R. Rudge and M.R. Ladisch, in J.A. Asenjo and J. Hong eds., *Separation, Recovery, and Purification in Biotechnology*, ACS Symposium Series 314, American Chemical Society, Washington, D.C., 1986, pp. 122–152.
59. S.R. Rudge, S.K. Basak, and M.R. Ladisch, *AIChE J.* **39**, 797–808 (1993).
60. S.K. Basak and M.R. Ladisch, *AICHE J.* **41**, 2499–2507 (1995).
61. S.K. Basak, A. Velayudhan, K. Kohlmann, and M. Ladisch, *Chromatography* **707** 69–76 (1995).
62. M.-P. Shiue and A.J. Pearlstein, *J. Chromatogr.* **707**, 87–103 (1995).
63. C.H. Lochmüller and C. Ronsick, *Anal. Chim. Acta* **249**, 297–302 (1991).
64. L.D. Shea, D.L. Feke, and U. Landau, *Biotechnol Progr.* **10**, 246–252 (1994).

ENERGY METABOLISM, MICROBIAL AND ANIMAL CELLS

MICHAEL BUTLER
RICHARD SPARLING
University of Manitoba
Winnipeg, Canada

XINFA XIAO
Apotex Fermentation, Inc.
Winnipeg, Canada

KEY WORDS

Energy charge
Glycolysis
Glutaminolysis
Heterolactic fermentation
Metabolic engineering
Methanogenesis
Oxidative phosphorylation
Pentose phosphate pathway
Secondary metabolites
TCA cycle

OUTLINE

INTRODUCTION

Every cell has a requirement for energy metabolism to maintain growth as well as to synthesize macromolecules or metabolic by-products. In the development of an industrial fermentation process it is important to be able to channel the metabolic capabilities of a selected cell to the production of a desired product. In this quest it is necessary to have a full understanding of the energy metabolism of the cell in order to be able to control the fermentation parameters to meet the cellular requirements and maximize the yield of the end product.

In this article we consider the energy metabolism of three quite distinct microbial cells—bacterial, fungal, and animal—all of which are used in large-scale commercial fermentation processes. The inclusion of animal cells as a category of microbial cells may be surprising to some, but in terms of their behavior as independent cells in homogeneous culture operations they do not differ from bacterial or fungal cells. Of course it has to be recognized that all cells used in industrial processes are likely to be selected and genetically altered in a way that marks their genotype as quite different from that found in wild strains or of cells contained within a multicellular organism.

In this article we point out the similarities and differences in energy metabolism of these cells and the aspects of the metabolism that are important for control in production processes of selected commercial products.

Common Pathways of Energy Metabolism for Bacterial, Fungal, and Animal Cells

Catabolic pathways provide energy for biosynthesis and in most instances also provide reducing equivalents for carbon reduction reactions. There are three major groups of compounds that link anabolic and catabolic pathways: adenosine phosphates (AMP, ADP, ATP), which link energy yielding and energy requiring reactions; nicotinamide adenine dinucleotide (NAD, NADH); and nicotinamide adenine dinucleotide phosphate (NADP, NADPH). The latter

two link oxidative reactions of catabolic pathways to reductive reactions. The adenylate energy charge is an indicator of the pool of high-energy phosphate bonds carried on adenosyl residues:

$$\frac{[ATP] + (1/2)[ADP]}{[ATP] + [ADP] + [AMP]}$$

This index could theoretically vary between 0 (100% AMP) and 1 (100% ATP). Measurements of this charge have been shown to correlate well with the cells' physiological state (1). The energy charge of actively growing cells ranges from 0.80 to 0.95 but drops significantly when energy substrates are depleted.

From a biotechnological point of view, when the product of interest is derived from biosynthetic pathways, the adenylate energy charge is of specific importance because it participates in the regulation of metabolic pathways. There is a relationship between the energy charge and reaction rates of regulated biosynthetic (energy-requiring) and catabolic (energy-regenerating) enzymes and pathways (2). The activity of biosynthetic enzymes is low and the activity of catabolic enzymes high, at low energy charges (<0.8). The rate of protein synthesis declines in response to decreases in energy charge (3).

Because of the involvement of a variety of oxidative and reductive pathways using nicotinamides as the electron carrier, NADH/NAD and NADPH/NADP ratios, also known as the catabolic and anabolic reduction charges, respectively, have also been measured for their potential as a means of measuring the physiological state of cells (1). In *Escherichia coli*, the catabolic energy charge does not vary significantly in response to the various growth conditions tested; however, the anabolic reduction charge decreases under carbon and energy source starvation and decreases transiently under conditions of O_2 starvation (4). Changes in anabolic reduction charge may have an impact on rates of biosynthesis as well.

The Embden–Meyerhof–Parnas (EM) pathway, the pentose-phosphate (PP) pathway, and the Entner–Doudoroff (ED) pathway are three routes for the utilization of hexoses such as glucose. The EM and PP pathways are widespread in fungi and animal cells. The ED pathway is common in bacteria and in some fungi such as *Tilletia caries* and *Caldariomyces fumago* (5,6).

The EM pathway takes place in the cytoplasm and is the most common means of energy production (Fig. 1). Two molecules of ATP are generated in this pathway, whereas only one is generated the ED pathway (6). Three enzymes are involved in the initial glucose phosphorylation step of the EM pathway: hexokinase P-I, hexokinase P-II, and glucokinase. Hexokinase acts on fructose as well as glucose and mannose. These enzymes are generally present in fungal, bacterial, and animal cells. However, doubt has been expressed as to the existence of the EM pathway in some fungi such as *Candida* 107 and *Rhodotorula glutinis* because of the low activity of phosphofructokinase (7).

The ED pathway was first discovered in *Pseudomonas* and serves as a mechanism of glucose breakdown in the absence of phosphofructokinase (Fig. 2). In this pathway, glucose-6-phosphate is oxidized to 6-phosphogluconate

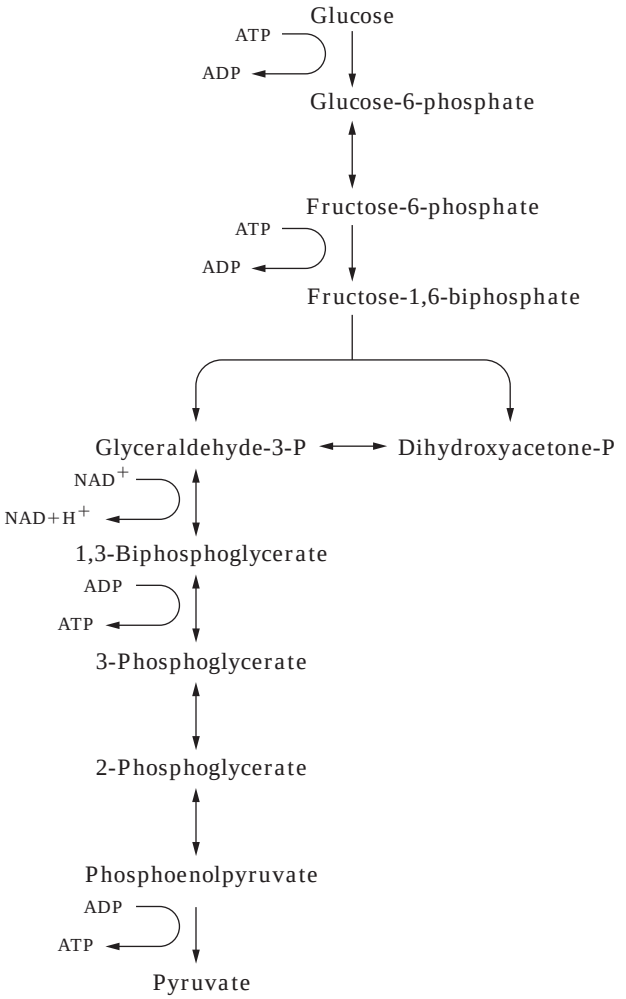


Figure 1. Embden–Meyerhof–Parnas pathway. This is the glycolysis pathway in which 1 mol glucose is degraded to 2 mol pyruvate. The net energy gain is of 2 mol ATP. The pyruvate is either reduced to lactate under anaerobic conditions or converted into acetyl-CoA for entry into the TCA cycle.

with the enzyme glucose-6-phosphate dehydrogenase. The presence of two other key enzymes (6-phosphogluconate dehydratase and 2-oxo-3-deoxy-6-phosphogluconate aldolase) allow further degradation to the three-carbon metabolites, pyruvate and glyceraldehyde-3-phosphate.

The PP pathway can be considered a side branch of the EM pathway that exists in fungal, bacterial, and animal cells (Fig. 3). Its principle role is the production of reducing power for biosynthetic reactions including nucleic acid and cell wall biosynthesis. In reductive biosynthesis, such as in fungal sporulation, the PP pathway is very important. Mutations of the PP pathway enzymes cause drastic morphological effects (6). Ribose-5-phosphate and erythrose-4-phosphate are produced from the PP pathway and are precursors for the synthesis of nucleotides and aromatic compounds, respectively (5).

The phosphoketolase (PK) pathway is an alternative pathway for sugar breakdown in fungal and bacterial cells (Fig. 4) (6). The key enzyme of the pathway (phosphoke-

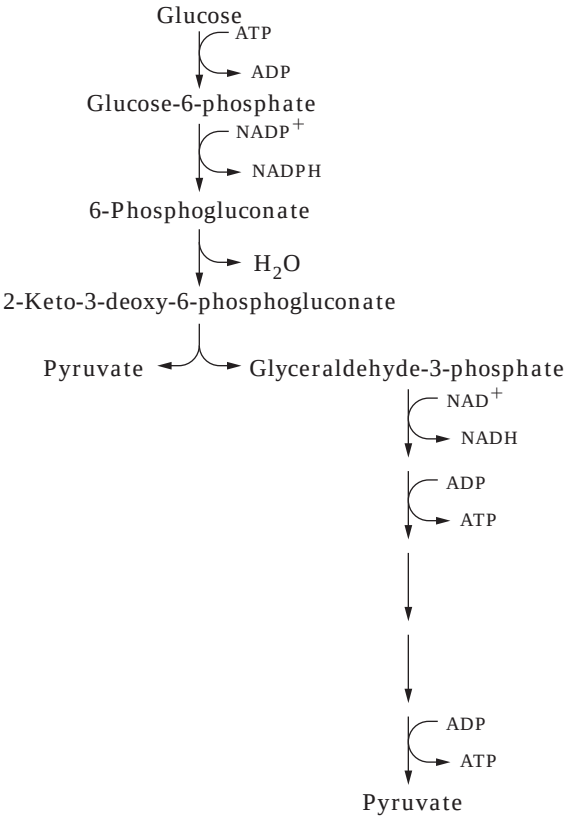


Figure 2. Entner–Duordoff pathway. Glucose is degraded in the absence of phosphofructokinase to pyruvate and glyceraldehyde-3-phosphate. The latter can be degraded further to pyruvate by the enzymes of glycolysis. This pathway is less efficient that glycolysis, yielding only 1 mol ATP/mol glucose.

tolase) converts a pentose such as xylulose-5-phosphate into acetyl-phosphate and glyceraldehyde-3-phosphate. The pathway is characteristically observed in heterolactic bacteria such as *Leuconostoc*, in which glucose may be converted into lactate, ethanol, and carbon dioxide.

Respiration in eukaryotes involves the complete oxidation of carbon-containing molecules to CO_2 and H_2O (Fig. 5). This includes three interrelated processes: the tricarboxylic acid (TCA) cycle, electron transport, and oxidative phosphorylation. These aerobic processes occur in the mitochondrion. The TCA cycle is the most efficient metabolic pathway for ATP generation (9,10) and is much more efficient than anaerobic metabolism. Using the EM pathway together with the TCA cycle under aerobic conditions, 30 to 38 molecules of ATP can be generated from one molecule of glucose, depending upon the degree of coupling of phosphorylation to the electron transport chain. By comparison, the EM pathway generates only 2 molecules of ATP from one molecule of glucose. Inorganic phosphate (P_i) supply is essential for ATP production. According to the energy requirements of cells, a mitochondrial inorganic phosphate carrier (P_ic) delivers P_i to the mitochondrial ATP synthase complex very rapidly, with a rate much higher than the rates of most mitochondrial reactions (10).

The TCA cycle is initiated when the pyruvate resulting from glycolysis is transported into the mitochondria and is converted to acetyl-CoA. Acetyl-CoA may also be formed from acetate by acetyl-CoA synthetase, and from the β-oxidation of fatty acids. When the acetyl group is transferred to oxaloacetate, the initial compound of the TCA cycle (citrate) results and is metabolized by a series of enzymatic reactions leading to the regeneration of oxaloacetate and production of CO_2, with the release of metabolic energy (5,8).

The TCA cycle is a central hub of both catabolic and anabolic metabolism. In addition to its role in the oxidation of substrates to generate energy, the TCA cycle provides important intermediates for biosynthesis. Glutamic acid, aspartic acid, and glycine are all directly derived from the TCA cycle and glyoxylate cycles, and they are important precursors for biosynthesis of other amino acids and nucleotides. The intermediates of the TCA cycle also contribute to the biosynthesis of secondary metabolites. The result of these biosynthetic processes is the depletion of intermediates from the cycle, and the TCA cycle would soon come to a halt without replenishment of intermediary metabolites. This may be accomplished through the fixation of CO_2 into oxaloacetate with phosphoenolpyruvate from glycolysis (8).

The glyoxylate cycle forms an apparent shortcut across the TCA cycle in fungi and some bacteria, linking isocitrate, succinate, and malate (Fig. 6). The net effect of this

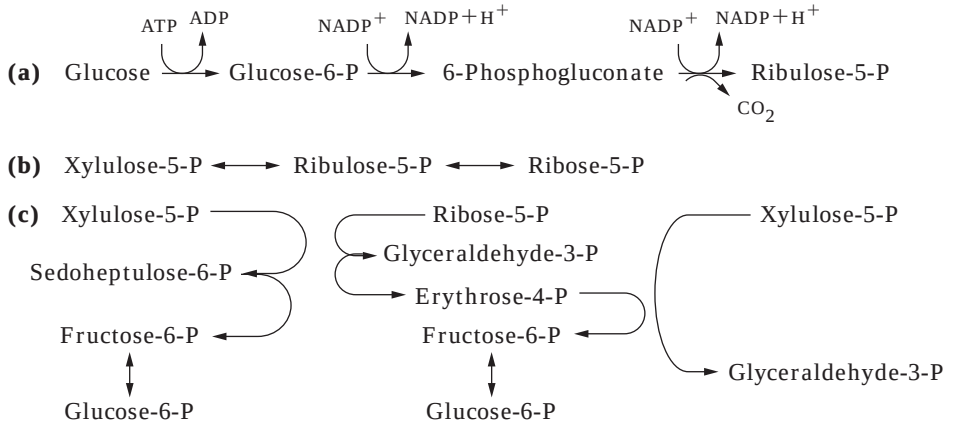


Figure 3. Pentose-phosphate pathway. This pathway is divided into three parts: (**a**) An oxidative sequence of reactions in which glucose is converted to the five-carbon, ribulose-5-phosphate. (**b**) Epimerization and isomerization reactions in which the important nucleic acid precursor, ribose-5-phosphate is formed. (**c**) A nonoxidative sequence of transaldolase and transketolase reactions.

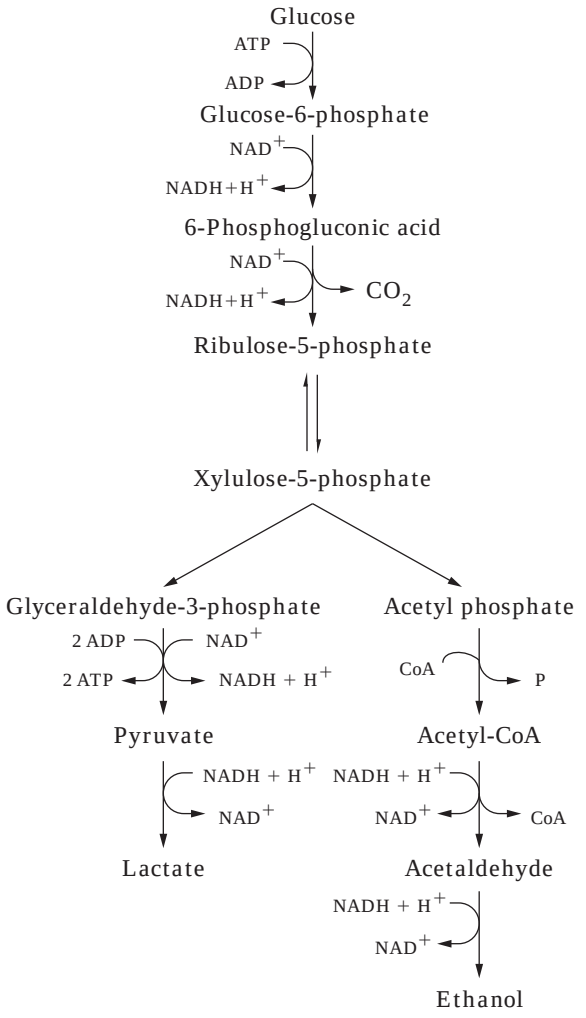


Figure 4. Phosphoketolase pathway. This pathway allows heterolactic fermentation in certain bacteria and fungi. The key enzyme is phosphoketolase, which degrades the five-carbon xylulose into a two-carbon, acetyl and three-carbon, glyceraldehyde. Lactate and ethanol are carbon products of the pathway. The net energy gain is 1 mol ATP/mol glucose utilized.

pathway is to produce one mole of succinate from two moles of acetate. The two enzymes required for this are isocitrate lyase and malate synthase. The pathway is called anaplerotic (from the Greek for "filling up") and serves to replenish the intermediates of the TCA cycle. In turn this may allow the TCA cycle intermediates to be used as precursors for biosynthetic reactions.

Mitochondrial reactions are essentially similar in plant, animal, and fungal cells. Electron transport and oxidative phosphorylation represent the final stages in the flow of electrons from organic substrates to oxygen (Fig. 7). Hydrogen atom pairs resulting from enzymatic dehydrogenation of certain intermediates of carbohydrate, fat, and protein catabolism, together with those from the TCA cycle, enter the electron transport chain (ETC), located in the inner mitochondrial membrane. The electrons of the hydrogen atom pairs are transferred to carrier components of the chain while the accompanying hydrogen ions enter the aqueous environment. The electrons are transferred by a series of oxidation–reduction reactions along the chain of redox components as far as cytochrome oxidase (cytochrome aa_3), which catalyzes electron transfer to oxygen. Each oxygen atom accepts two electrons and takes up two hydrogen ions from the environment to form water. At three sites in the ETC, oxidative phosphorylation may occur. That means ADP is phosphorylated, resulting in the formation of up to three ATPs from each electron pair transported to oxygen. For each NADH that transfers its electrons to the ETC, a maximum of three molecules of ATP result; for each $FADH_2$, a maximum of two ATP molecules result (5,8).

Besides the ubiquitous cytochrome-dependent and ATP-generating path for the oxidation of NADH + H^+, there are alternative respiration pathways. An alternative pathway, which branches from ubiquinone and transfers electrons via an unknown oxidase toward oxygen, has been described in all major classes of fungi (11). This alternative pathway is cyanide- and azide-insensitive but can be inhibited by salycilhydroxamate (SHAM). It transports electrons to oxygen without proton transport, and therefore without phosphorylation of ADP to form ATP; only heat is generated. Another alternative pathway is cyanide-insensitive and azide-sensitive and is also believed to be mitochondrial and to lack proton transport capability. This pathway accepts electrons from external NADH but not from succinate, implying lack of access to the ubiquinol pool, but whether it has access to internal mitochondrial NADH is not known (6).

In bacteria and fungi, respiration may function with an electron acceptor other than oxygen. For example, nitrate may be used as an electron acceptor and reduced to nitrite by nitrate reductase. This process, which is often called dissimilatory nitrate reduction, is coupled to the generation of ATP. Denitrification involves the complete reduction of nitrate to nitrogen and prevents the accumulation of potentially toxic levels of nitrite. Denitrification is carried out by *Pseudomonas* or *Bacillus* and also in the fungi, *Fusarium oxysporum* and *Cylindrocarpon tonkenense*. The respiratory substrates such as malate, pyruvate, succinate, and formate are effective donors of electrons. Rotenone, antimycin A, and thenoyltrifluoroacetone inhibit the reductase activity (12).

The range of energy substrates and catabolic pathways in bacteria is extremely wide. These pathways may lead to the generation of ATP or to primary ion gradients that are easily interconverted through the ubiquitous presence of an ion-coupled ATPase or synthase (13). The large variety of metabolic pathways available in prokaryotic organisms has led to their use in the production of a range of metabolic end products and enzymes.

NUTRIENTS REQUIRED FOR GROWTH

Bacteria and Fungi

Although some bacteria are autotrophic and may use carbon dioxide as a carbon source, many are heterotrophic and require some organic carbon for growth in culture. It has

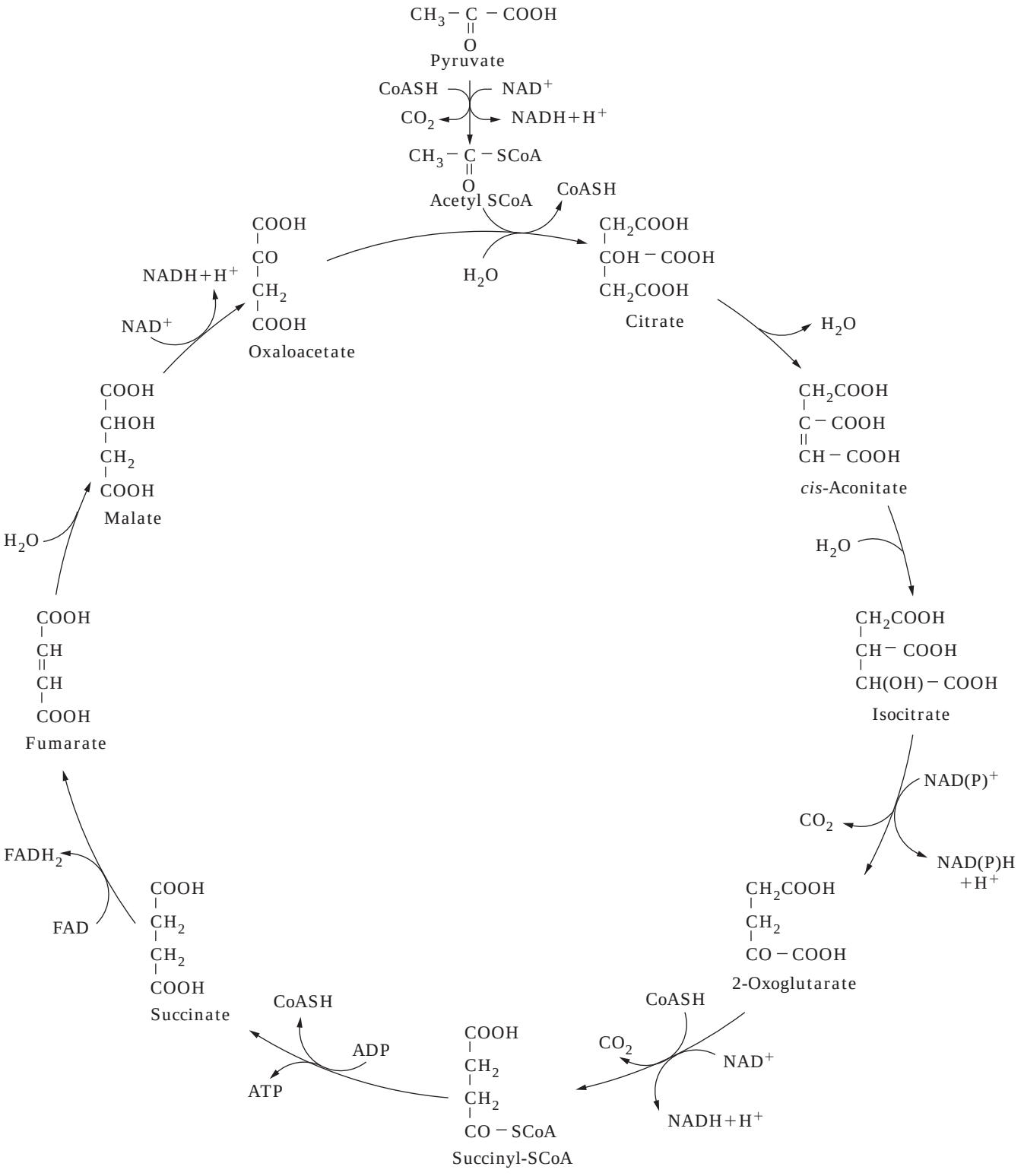


Figure 5. Tricarboxylic acid cycle. The TCA pathway operates in the mitochondria of eukaryotes. It allows the complete oxidation of pyruvate, which enters the cycle from glycolysis. The reduced coenzymes are reoxidized via the electron transport chain. The complete aerobic metabolism of glucose via glycolysis and the TCA cycle yields between 30 and 38 mol ATP/mol glucose, depending on the extent of the coupling of phosphorylation in the electron transport chain.

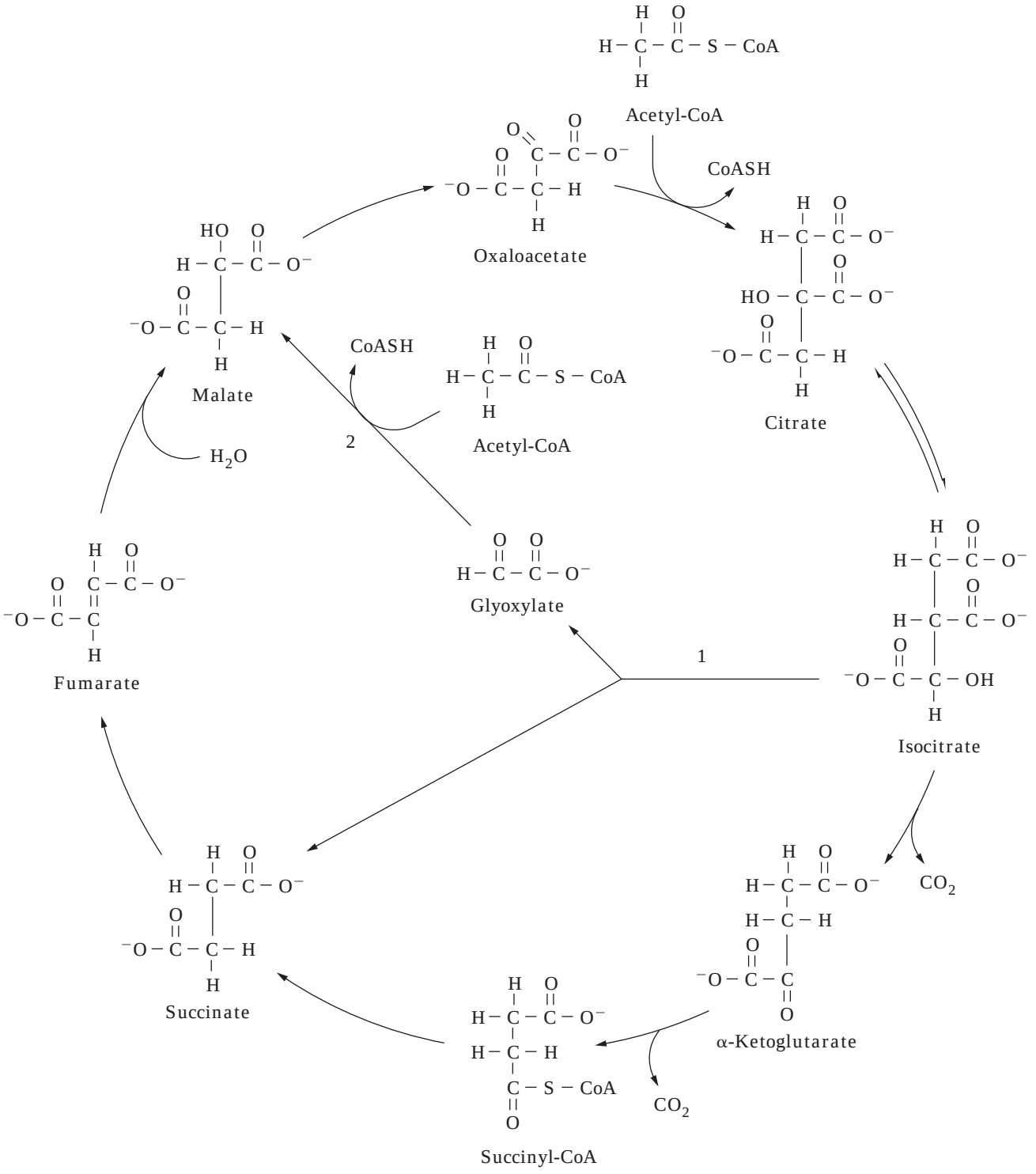


Figure 6. Glyoxalate pathway. The presence of the two enzymes, isocitrate lyase (1) and malate synthase (2) in some bacteria and fungi enables the replenishment of TCA cycle intermediates, which may be used as precursors for biosynthetic pathways. This is termed an anaplerotic pathway.

been common in laboratory cultures to use non–chemically defined sources of organic carbon such as peptones, meat extract, or yeast extract. However, to maintain consistency in large-scale cultures it is important to use a chemically defined nutrient broth consisting of an energy source in an isotonic salt solution.

Glucose and glucose-yielding carbohydrates are the most commonly used carbon energy sources for bacterial or fungal cell growth. With the breakdown of glucose through the pathways of glycolysis, fermentation, or respiration, energy and a series of small compounds are generated to support growth and biosynthesis. ATP is the energy carrier mediating between the energy donor and energy-requiring reactions. Hexoses can be converted to glucose, and pentose can be metabolized through alternative glycolytic pathways. Fatty acids, amino acids, organic acids, and alcohols enter the major catabolic pathways at various points (6,8).

Because of the extensive functional biosynthetic pathways in bacteria and fungi, the energy source and culture medium are relatively simple. Some examples of single energy sources for fungal producer cells are shown in Table 1.

Unusual Energy Sources

Methanol or alkanes can be used as a sole source of carbon and energy by some species of bacteria and fungi. Methanol is oxidized to formaldehyde, which is further metabolized by two pathways. The first path oxidizes formaldehyde to formic acid and then to CO_2, with the generation of ATP and NADH. The second path combines formaldehyde with xylulose-5-phosphate to form dihydroxyacetone (DHA) and glyceraldehyde-3-phosphate. This path requires ATP to phosphorylate the DHA and provides intermediates for the pentose cycle or the EM pathway. The enzymes DHA synthase and DHA kinase are unique to this pathway and are formed only with methanol utilization. The *n*-alkanes are oxidized to corresponding alkanols by a monooxygenase enzyme complex containing cytochrome P450 and NADPH cytochrome P450 reductase. Peroxisomal enzymes carry out the remaining degradation to intermediates that are transferrable to the mitochondria. NAD-long-chain alcohol and aldehyde dehydrogenase further oxidize the alkanols to fatty acids (9). *Candida tenuis* was reported to be able to metabolize the whole range of fuel hydrocarbons. In a mixture of them, the relatively short-chain-length *n*-alkanes were used first, and then progressively longer *n*-alkanes. When all *n*-alkanes were used up, methylalkanes were exploited, but only relatively slowly (19).

Some strains of *Candida tropicalis*, *Rhodotorula* sp., *Pullularia pullulans*, and *Oidium* can use phenol at low concentrations as the single source of carbon and energy. *Aspergillus fumigatus* can grow on *p*-cresol or 4-ethylphenol as its sole carbon and energy source. *p*-Cresol is metabolized by two different routes, both of which converge on protocatechuate as the ring-fission substrate. One of these proceeds by initial hydroxylation of the methyl group of *p*-cresol and its further oxidation to 4-hydroxybenzoic acid, followed by ring hydroxylation; the other involves ring hydroxylation of *p*-cresol to form 4-methylcatechol as the first step, followed by oxidation of the methyl group. The protocatechuate is cleaved by *ortho*-fission, leading to formation of 3-oxoadipate. In contrast, the ring-fission substrate in the pathway for 4-ethylphenol degradation is hydroxyquinol (1,2,4-trihydroxybenzene), which is cleaved by *ortho*-fission, leading to maleylacetate. The pathway involves hydroxylation of the methylene group of 4-ethylphenol, followed by oxidation to give 4-hydroxyacetophenone. This undergoes a Baeyer–Villiger type of oxygenation to form the ester, 4-hydroxyphenyl acetate. Hydrolysis of the ester gives quinol (hydroquinone), which is converted to the ring-fission substrate by hydroxylation (20).

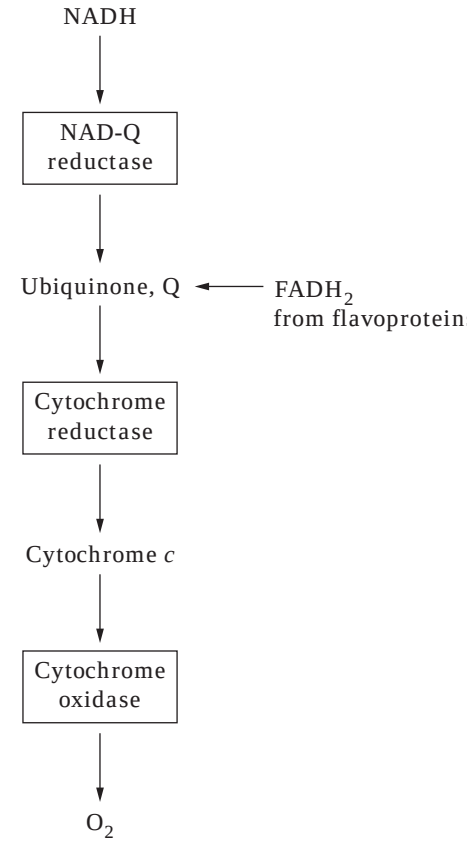


Figure 7. Electron transport chain. In this pathway electrons are transferred from NADH to oxygen through a chain of three large protein complexes (NAD-Q reductase, cytochrome reductase, and cytochrome oxidase). Ubiquinone and cytochrome *c* are relatively small molecules that shuttle between these complexes.

Toluene can be used as a sole source of carbon and energy by *Cladosporium sphaerospermum*, which is the first eukaryotic organism found to do so (21). This fungus was isolated from a biofilter used to remove toluene from waste gases. The oxygen consumption rates, as well as the measured enzyme activities, of *C. sphaerospermum* indicate that toluene is degraded by an initial attack on the methyl group.

The Composition of Animal Cell Culture Media

The substrate requirements for growth and production from animal cells is far more complex than for bacteria or

Table 1. Energy Sources for Producer Fungal Cells

Product	Producer	Source of energy and carbon	g/L	Reference
Citric acid	*Aspergillus niger*	Sugar	125–200	14
Itaconic acid	*Aspergillus terreus*	Glucose	50–150	14
Gluconic acid	*Aspergillus niger*	Dextrose	150–200	14
Penicillin G/V	*Penicillium chrysogenum*	Glucose	0–10 or 5–15	
		Lactose	20–50 or 15–60	15
Lovastatin	*Aspergillus terreus*	Dextrose	45	16
Compactin	*Penicillium citrinum*	Glucose	50	
		Glucose/maltose	10/40	17
Cyclosporin A	*Tolypocladium inflatum*	Maltose, glucose, etc.	30–50	18

fungi. Prior to the 1950s animal cells were grown in culture in undefined biological fluids, with varying degrees of success. Subsequently, chemically defined culture medium was based on the analysis of plasma and other growth-supporting biological fluids (22). The reduction of the defined components to minimum concentrations then led to "minimal essential media" such as that developed by Eagle in 1955 (23). Similar media formulations such as Dulbecco's modification of Eagle's medium (DMEM), Ham's F-12, and RPMI were later developed to promote the growth of specific cell lines. Table 2 shows the composition of DMEM as an example of a typical culture medium suitable for the growth of a wide range of animal cell lines.

The nature and concentration of the carbon and nitrogen substrates in the culture medium affect both the energy status and metabolism of the cells (Fig. 8). Medium generally contains both a carbohydrate source and a range of amino acids to satisfy cellular growth requirements. Most standard media formulations would be expected to support cell growth to a maximum density of 1–2 $\times$ 10^6 cells/ml in batch culture. The limitation to growth is normally a reflection of the deterioration of the culture medium due to nutrient depletion or the accumulation of toxic metabolites. The specific factor affecting the final cell yield may vary with cell line, media, and growth conditions. However, the identification of such factors and an understanding of the interaction between the cells and the media components is essential for designing a fully optimized and controllable culture system.

Carbohydrate Source. Glucose is the most commonly used carbohydrate source for animal cells and is normally included in culture at a concentration between 5 and 25 mM. At low glucose concentration ($<$25 μM) its primary function is to provide ribose for nucleic acid synthesis (24). It has also been shown that glucose may not be necessary if a source of pentose sugar is provided in the culture. Uridine and inosine have been found suitable, the carbohydrate requirement being derived from the ribose moiety of these nucleosides (25).

The specific glucose consumption rate of cells increases with the glucose concentration of the medium up to a maximum rate that occurs at around 5 mM glucose (24). The specific growth rate and antibody production rate of hybridomas are independent of glucose concentration over a wide range. The specific glucose consumption rate may vary with growth phase, culture pH, or cell type. During

Table 2. Dulbecco's Modification of Eagle's Medium (DMEM): An Example of a Basal Medium Commonly Used for the Growth of Animal Cells in Culture

Inorganic salts (mM)	
NaCl	110.30
KCl	5.40
$CaCl_2 \cdot 2H_2O$	1.80
$MgSO_4 \cdot 7H_2O$	0.80
$NaH_2PO_4 \cdot 2H_2O$	0.91
$NaHCO_3$	44.00
L-Amino acids (mM)	
Arginine·HCl	0.40
Cystine·diNa	0.20
Glutamine	4.00
Glycine	0.40
Histidine·HCl·H_2O	0.20
Isoleucine	0.80
Leucine	0.80
Lysine·HCl	0.80
Methionine	0.20
Phenylalanine	0.40
Threonine	0.80
Tryptophan	0.08
Tyrosine·diNa	0.40
Valine	0.80
Trace elements (μM)	
$Fe(NO_3)_3 \cdot 9H_2O$	0.25
Vitamins/cofactors (μM)	
Choline·Cl	28.60
Folic acid	9.10
Inositol	38.90
Nicotinamide	32.80
Pantothenate·Ca	8.40
Pyridoxal·HCl	19.60
Riboflavin	1.10
Thiamine·HCl	11.90
Other components	
Phenol red (uM)	26.55
Glucose (mM)	25.00
Pyruvate·Na (mM)	1.00
CO_2 (gas phase)	10%

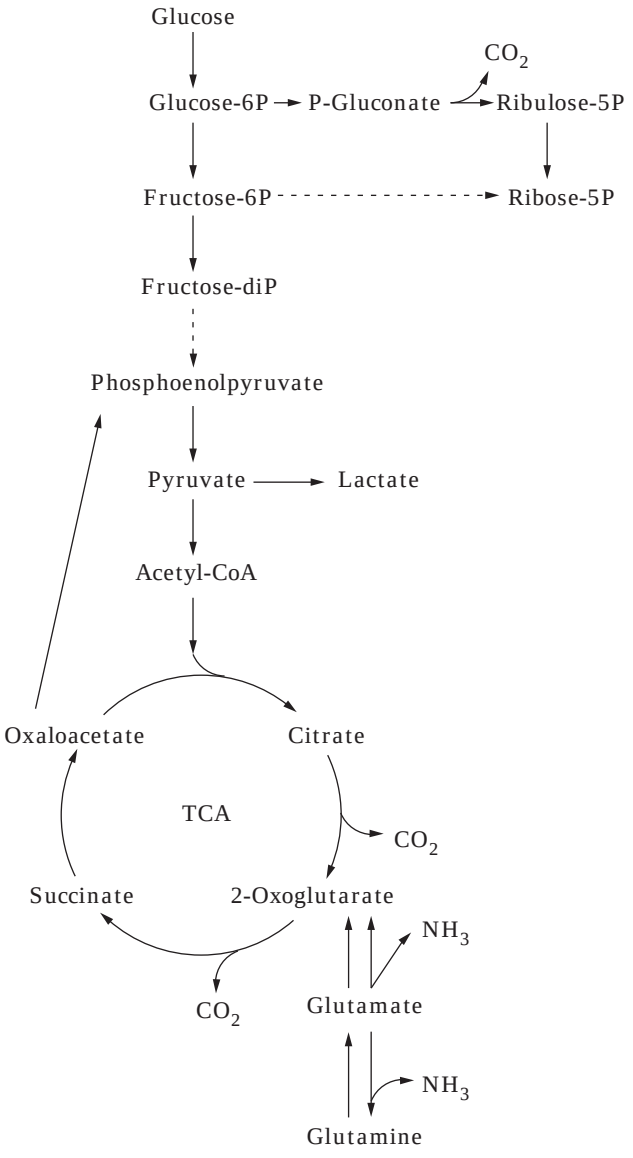


Figure 8. Interrelated pathways of energy metabolism. The two main substrates for animal cell growth in culture are glucose and glutamine. Glucose may be metabolized via glycolysis, the pentose-phosphate pathway, or the TCA cycle. Glutamine enters the TCA cycle via 2-oxoglutarate (α-ketoglutarate), which may lead to complete oxidation. Alternatively, glutaminolysis leads to incomplete oxidation of glutamine and the formation of a range of carbon by-products including lactate, alanine, pyruvate, or aspartate.

the culture of many animal cells, glucose is not completely depleted from the medium, suggesting that the glucose concentration is not the limiting factor.

Glucose utilization in culture leads to a concomitant increase in lactate concentration in the medium, which may lead to a decrease in pH if adequate buffering is not provided. The molar ratio of lactate produced to glucose consumed is often >1.8 for growing cells and is an indication of the anaerobic metabolism of glucose (26). If glucose is substituted by either fructose, galactose, or maltose, cell growth will continue, but the glycolytic rate and hence the production of lactate will decrease (27). In this situation cell growth will depend more heavily on other available substrates such as glutamine.

Glutamine. Next to carbohydrates, amino acids are the most abundant source of reduced carbon available for energy metabolism. Glutamine is normally included at a concentration higher than the other amino acids (0.7–6 mM). It is recognized as an important precursor for the synthesis of purines, pyrimidines, amino sugars, and asparagine, as well as a substrate for oxidative metabolism (28).

Glucose and glutamine utilization has been shown to be coordinately controlled in a number of cell lines (24,29). Glutamine was shown to provide 40–50% of the intracellular energy in a variety of cell lines including HeLa, human fibroblasts, and hybridoma cells. When fructose or galactose is the carbohydrate source, the proportion of cellular energy provided by glutamine is even higher.

One of the major end products of glutamine utilization is ammonia. This can increase in concentration in the medium as a result of the spontaneous chemical degradation of glutamine, as well as from its metabolic deamination. The molar ratio of ammonia produced to glutamine consumed is typically >0.75 (26). The accumulation of ammonia can be inhibitory to cell growth at >3 mM, but the sensitivity of cell lines to ammonia is variable.

There are various strategies that may be used to overcome the inhibitory effect of ammonia, including removal from the culture medium by ion exchange resins or by gas-permeable membranes (30). Alternatively, the substitution of glutamine by equivalent molar concentrations of glutamate (26) or glutamine-based dipeptides (31) can lead to reduced ammonia accumulation and, under appropriate culture conditions, to higher cell yields.

Amino Acids Other Than Glutamine. A range of amino acids is normally provided in culture medium, each at a concentration of 0.1–1.0 mM. The amino acids act primarily as precursors for protein synthesis, but some may be deaminated and degraded via the tricarboxylic acid cycle, particularly if the primary energy substrates (glucose and glutamine) have been utilized.

The branched-chain amino acids are consumed particularly rapidly by a number of cell lines including MDCK cells, human fibroblasts, mouse myeloma cells, and BHK cells. Hybridomas have been shown to metabolize serine and branched-chain amino acids when the glutamine level of the culture is low. During cell growth some amino acids may be released into the culture medium, particularly alanine and glycine. These may be products of a detoxification mechanism that prevents an excessive accumulation of ammonia in the cell.

Lipids. The inclusion of a lipid fraction or defined fatty acids in serum-free culture medium has been shown to improve cell proliferation (32). The most likely role of such fatty acids, which are normally supplemented at micromolar concentrations, is a source of precursors for the structural components of cell membranes. However, it has been shown that the availability of unsaturated fatty acids

such as oleic and linoleic acids can significantly alter the metabolic profile of hybridoma cells by causing changes in the specific uptake rate of glutamine (33).

PATTERNS OF ENERGY METABOLISM

Patterns of Fungal Energy Metabolism

According to their different patterns of energy procurement, fungi can be classified into several metabolic groups. Many yeasts and most filamentous fungi belong to the group of obligate respirers. Chitridiomycetes rumen fungi belong to the group of obligate anaerobes. Between these two extreme groups are aerobic respirers, aerobic fermenters, facultative aerobic fermenters, and aeroneutrial fermenters. Few yeasts and filamentous fungi belong to the group of aerobic respirers. There is no anaerobic growth in this group, and anaerobic fermentation is possible only in pregrown cells. Aerobic fermenters and facultative aerobic fermenters have capabilities for both aerobic and anaerobic fermentation, but their respiratory capabilities are limited. Fungi in the group of aerobic fermenters, such as *Schizosaccharomyces pombe*, do not have the capability for anaerobic growth. Fungi in the group of facultative aerobic fermenters, such as *Saccharomyces*, have the capability for anaerobic growth. Some Chytridiomycetes and Oomycetes (e.g., *Blastocladia* and *Aqualinderella*) belong to the group of aeroneutrial fermenters, which lack respiration and have the capability of both aerobic and anaerobic fermentation or growth (6).

Several fungi can perform anaerobic fermentation in the presence of O_2. There are three principal types of fungal fermentation: alcoholic, lactic acid, and mixed acid fermentation. Lactic acid fermenters are found primarily among the Chytridiomycetes, Oomycetes, and Zygomycetes. Some of these fungi are obligate aerobes (e.g., *Allomyces*, *Sapromyces*, and *Apodachlya*), whereas others grow equally well anaerobically (e.g., *Blastocladia* and *Mindiniella*). *Rhizopus*, a member of the Zygomycetes, carries out lactic acid fermentation together with a simultaneous alcoholic fermentation. Mixed acid fermentation is found in a small group of obligately anaerobic Chytridiomycetes (6).

For most industrial fungi, oxygen is essential for both growth and production of metabolites. In an industrial aerobic fermentation process, an adequate oxygen supply is usually obtained by sparging sterile air into the fermentation broth. Because of the low solubility of oxygen it is important to maximise the mass transfer between the gas and liquid phase. The mass transfer is measured by the volumetric mass transfer coefficient, k_La, where k_L is the mass transfer coefficient in the liquid film (cm/h), and a is the specific interfacial area for air–liquid mass transfer (cm^2/cm^3). The mass transfer coefficient is related to the volumetric power input and superficial gas velocity. In large-scale aerobic fermentation, oxygen enrichment of the inlet air stream may be advantageous to meet the oxygen requirement. The profile of the oxygen uptake rate throughout the fermentation process is an important parameter for scale-up (34). The oxygen uptake rate (OUR) may be calculated by

$$\mathrm{OUR} = k_La(C^* - C)$$

where C^* is the equilibrium oxygen concentration, and C is the dissolved oxygen in the culture broth.

Pattern of Energy Metabolism in *E. coli:* A Model Bacterium for Commercial Production

E. coli has been one of the most intensely studied species of bacteria over the last several decades and has been used extensively as the bacterium of choice for the expression of a wide variety of foreign proteins including many of industrial importance. Numerous human proteins have been expressed in *E. coli* by means of recombinant DNA technology, and some are currently in commercial production, including the examples shown in Table 3 (35). This bacterium's importance in the field of biotechnology requires that its energy conservation pathways for ATP generation be described in more detail. *E. coli* is a heterotrophic facultative anaerobe with diverse strategies for generating ATP from pathways leading to substrate level phosphorylation and proton-motive-force-coupled ATP synthesis. The anaerobic fermentative pathways of *E. coli* yield 3 ATP/mol glucose using a combination of the EM pathway of glycolysis to pyruvate followed by various pathways of pyruvate reduction leading to the formation of acetate, ethanol, formate, and lactate as the major end products (35). *E. coli* also possesses different electron transport chains coupled to a variety of electron acceptors, including O_2, NO_3^-, fumarate, and dimethylsulfoxide, for the generation of a primary proton motive force. While this versatility is of great ecological importance for the survival of this organism in the environment, most of it is of little use from a biotechnological perspective in the expression of foreign proteins, because the anaerobic pathways provide less ATP per mole of substrate used, and therefore, less cell material or product per mole of substrate.

Most of the applications of *E. coli* involve aerobic growth of the cells with O_2 as terminal electron acceptor of the respiratory chain. When using glucose as the primary energy source, oxidation of the glucose molecule to $6CO_2$ proceeds via glycolysis and the Krebs cycle, yielding 10 NADH and 2 $FADH_2$ (36). From the perspective of substrate-level phosphorylation, 4 ATP is generated within glycolysis and the Krebs cycle. Further ATP is generated in association

Table 3. Examples of Currently Licensed Therapeutic Protein Products from Animal Cell and Bacterial Cultures

Animal cell cultures	*E. coli* cultures
Interferon alpha	Insulin
OKT3 monoclonal antibody	Interferon alpha, beta, and gamma
Tissue plasminogen activator	
Erythropoietin	Human and bovine growth hormone
Human growth hormone	
Hepatitis B surface antigen	Rennin
Granulocyte colony stimulating factor	Interleukin 2
Clotting factor VIII	
DNase I	

with the electron transport chain and the generation of a primary proton motive force. The electron transport chain of *E. coli* differs from that of mitochondria in eukaryotes in that cytochrome *c* is not present (35). The chain is divided into two steps. The first step is the transfer of electrons from NADH to ubiquinone (a membrane-soluble cofactor carrying two electrons). This electron transfer is performed via the enzyme NDH1 (NADH-ubiquinone oxidoreductase I), a complex multisubunit transmembrane enzyme capable of proton translocation. The product of this reaction is ubiquinol. There are other electron transfer reactions that yield ubiquinol but are not associated with proton translocation, for example, the transfer of electrons from succinate to ubiquinone via the FAD-containing succinate dehydrogenase. In contrast to the pathway in mitochondria, where the electrons from ubiquinol are donated to cytochrome *c*, the two electrons of ubiquinol are donated directly to O_2, the reduction of which takes place on a quinol oxidase. *E. coli* can synthesize two different quinol oxidases, depending on the concentration of oxygen in the growth medium. At high concentrations of O_2, a cytochrome *bo*$_3$–containing quinol oxidase is produced, whereas a cytochrome *bd*–containing quinol oxidase, which has a higher affinity for oxygen, is produced at low O_2 concentrations. This difference is significant in that although both enzymes are involved in proton translocation, the first oxidase translocates 4 $H^+/2e^-$, whereas the second translocates only 2. Under optimal conditions (when cytochrome *bo*$_3$ quinol oxidase is used) a maximum 8 $H^+/2e^-$ has been estimated, as compared with 6 when the oxygen tension is low (35). By comparison the electron transport chain in mitochondria releases between 10 and 12 $H^+/2e^-$. These extruded protons can be used for several functions including active transport, motility, and the synthesis of ATP via the F_1F_0 ATPase. The F_1F_0 transmembrane ATPase of *E. coli* uses 3 H^+ per ATP synthesized.

The number of protons translocated per $2e^-$, and therefore also the number of ATPs that can be synthesized per $2e^-$, can be said to be dependent on both the substrate utilized and the O_2 concentration. Furthermore, the selection of bioenergetic substrate does not only have an impact on the amount of ATP generated per mole of substrate but also on the amount of ATP needed for cell biosynthesis, as demonstrated when experimentally measuring maximal cell yields per millimole of ATP (Y_{ATP}^{max}) when cells are grown on different carbon and energy sources (37). The Y_{ATP}^{max} for cells grown on glucose is twice that of cells grown on acetate. This corresponds well with theoretical calculations of the requirement for ATP for the synthesis of macromolecules (proteins, lipids, nucleotides) on the same substrates. For example the ATP requirement for protein synthesis doubles when cells use acetate as their carbon and energy source rather than glucose (42.7 vs. 20.5 mmol ATP/g cells). Thus, the choice of carbon and energy substrate not only has an impact on the amount of ATP synthesized but also on the amount of ATP needed for biosynthesis.

Patterns of Energy Metabolism in Animal Cell Culture

The metabolic pathways for the utilization of glucose in animal cells include glycolysis, the PP pathway, and the TCA cycle. The complete oxidation of glucose via glycolysis and the TCA cycle generates the greatest energy yield. However, the rate of aerobic metabolism of glucose is extremely low during the growth of many cell lines (38). By metabolic flux analysis, the proportion of glucose metabolized via the TCA cycle in cultured cells has been determined to be less than 5%, whereas over 90% is utilized via glycolysis to lactate, pyruvate, or alanine (39). The essential synthetic role of glucose is to feed the PP pathway, which leads to the formation of ribose, which is essential to maintain high rates of DNA synthesis during cell growth.

The addition of glucose to cells in culture may cause a decrease in oxygen consumption, a phenomenon commonly referred to as the Crabtree effect (40) and widely observed for many cell lines (41). This indicates a limited metabolic capacity for conversion to TCA cycle intermediates as the flux through glycolysis increases; this may be a result of a high cytoplasmic NADH/NAD ratio. The aerobic metabolism of cells in culture has also been shown to decrease at elevated dissolved oxygen concentrations (42). This may be related to the inhibition of pyruvate dehydrogenase, which is an enzyme complex that is required for linking glycolysis with the TCA cycle.

Glutamine provides a significant proportion of cellular energy by means of a metabolic network of up to eight individual metabolic pathways collectively described as glutaminolysis (43). This network is initiated by the deamination of glutamine by phosphate-activated glutaminase, which may regulate the activity of this extensive metabolic pathway. Metabolic analysis of HeLa and CHO cells by ^{15}N-NMR has shown that all the ammonia produced from glutamine originates from the amide nitrogen rather than the α-amino group (44). Because the activity of glutamate dehydrogenase is low in most cells, subsequent conversion of glutamate into α-ketoglutarate takes place by aminotransferases, notably aspartate transaminase (45). There follows a pathway of complete or partial oxidation leading to the end products of glutaminolysis, which include carbon dioxide, ammonia, alanine, aspartate, and lactate. The complete breakdown of the carbon skeleton of glutamine via the TCA cycle is the most efficient energy-releasing pathway (27 mol ATP/mol glutamine). However, it is probable that a large proportion of glutamine is incompletely oxidized to pyruvate, lactate, or alanine, resulting in a lower energy yield (9 mol ATP/mol glutamine) (43). A study of the metabolism of a murine hybridoma shows that alanine is the major end product of glutaminolysis (39). From this work it can be estimated that glutamine accounts for over 40% of the ATP generated in energy metabolism.

THE IMPORTANCE OF ENERGY METABOLISM TO SPECIFIC CELL PRODUCTS

Catabolic End Products of Fungi

Since L. Pasteur published his work on alcoholic fermentation in 1860, ethanol production from pyruvate has been best known from the studies on *Saccharomyces* (46). Many obligate aerobic fungi, such as common molds of the genera *Aspergillus, Fusarium*, and *Mucor*, are also well known for their ability to make ethanol (6).

There are excellent reviews about fermentation production of organic acids, such as citric acid, itaconic acid, gluconic acid, and fumaric acid. In economic terms, citric acid is the most important organic acid produced by fungi. Commercial citric acid production from *Aspergillus niger* was first demonstrated as feasible in 1917 (47). Strains of *Candida lipolytica* with reduced aconitase activity (which converts citrate to aconitate) are also used for citric acid production (48). It is believed that citric acid production depends on a disturbance of the metabolism of the TCA cycle. An important factor in the high citric acid productivities obtained is the existence of a powerful supplementary route to oxaloacetic acid by the carboxylation of phosphoenolpyruvate (49).

Secondary Metabolites of Fungi

The metabolic precursors of most secondary metabolites are products or intermediates of energy metabolism pathways (6). The shikimic acid pathway depends on the biosynthesis of aromatic amino acids, the polyketide biosynthetic pathway on acetyl CoA, and the mevalonic acid pathway on acetyl-CoA, polysaccharides, and peptidopolysaccharides.

Three amino acids (L-α-aminoadipic acid, L-cysteine, and L-valine) are precursors for biosynthesis of the β-lactam antibiotics penicillin G or V and cephalosporin C. The fungi *Penecillium chrysogenum* and *Acremonium chrysogenum* (also known as *Cephalosporium acremonium*) produce these two antibiotics, respectively, and use a multifunctional enzyme, δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase (ACV synthetase), to convert the three amino acids to a tripeptide that is a common intermediate of β-lactam antibiotics. Cyclosporin A, an important immunosuppressant, is produced by the fungus *Tolypocladium inflatum* (also called *Beauveria nivea*) from 11 amino acids.

In recent years, the cholesterol-lowering agent lovastatin (also called mevinolin, monacolin K, or MB-530B) has been widely used for treatment of hypercholesterolemia. The fungi *Aspergillus terreus* and *Monascus ruber* produce lovastatin through a polyketide pathway from acetyl-CoA. Nine intact acetate units are incorporated into the lovastatin molecule, and the methyl group of lovastatin is derived from methionine. Acetyl-CoA is a precursor for the biosynthesis of lovastatin and fatty acids, although these two pathways occur at different times during the course of the fungal culture. Fatty acid biosynthesis is performed actively with growth at an early period of fermentation and reaches maximum on day 1 of growth, while lovastatin biosynthesis increases at a later time, with a maximum between days 3 and 5. Fatty acid biosynthesis is minimal on day 3 of growth and is associated with a decrease in the activity of fatty acid synthetase. This allows an increase in lovastatin biosynthesis, although this may be limited by a concomitant decrease in activity of the citrate cleavage enzyme that generates acetyl-CoA from citrate (50).

ATP Requirements for Fungal Secondary Metabolites. All secondary processes need energy (ATP) support. For instance, in the biosynthesis of penicillin G by *Penicillium chrosygenum*, the three precursor amino acids have to be activated by ATP at the cost of two high-energy phosphate bonds for each amino acid. Two high-energy phosphate bonds are also needed to activate phenylacetic acid (PAA) by CoA-SH. Not including energy consumption due to intracellular compartmentation and transport of penicillin from the cell to the fermentation broth, 3 ATP and 1 ADP are consumed to synthesize one molecule of penicillin G from the three amino acids (51). Cyclosporin A biosynthesis by the fungus *Tolypocladium inflatum* has a high rate of ATP consumption. The ATP concentration can be increased in a cell-free reaction mixture in order to obtain a higher cyclosporin A yield (52). A single multifunctional cyclosporin synthetase catalyses at least 40 reaction steps to complete the synthesis, and 11 amino acids need to be activated by using ATP before building up the peptide chain.

There is evidence that too much ATP production or accumulation in cells causes inhibition of the biosynthesis of secondary metabolites. Regulation of secondary metabolism by readily utilized carbon and energy sources, by high concentration of phosphate, or by the high phosphate concentration combined with an increase of consumption of the carbon source has been well reviewed, and it has been speculated that ATP could be the intracellular effector controlling the synthesis of secondary metabolism (53). A high fusidic acid–producing mutant of *Fusidium cocineum* was found to contain less intracellular ATP during production and less polyphosphate throughout the fermentation than the low producer. Similar phenomena were observed in prokaryotic antibiotic producers. Within 5 min after adding phosphate to candicidin-producing cells, the concentration of ATP doubles or triples before inhibition of antibiotic synthesis. Improved tetracyclin-producing strains have lower intracellular ATP levels than their poor-producing ancestral strains. The difference of ATP levels between the high- and low-producing strains does not mean a difference of the adenylate energy charge. In fact, the energy charge is similar between the strains with higher or lower productivity (54). In research on relationships among cellular energy status and the induction and initiation of aflatoxin synthesis in *Aspergillus parasiticus*, there was no evidence for the regulation of aflatoxin synthesis by the overall energy status of the fungal cell. However, electron microscopy indicates that aflatoxin synthesis occurs in association with a glucose-mediated inactivation of mitochondria. This suggests a possible regulation of aflatoxin synthesis by the energy status of specific subcellular compartments (55).

Fungicides and Drug Resistance Relating to Energy Metabolism

Inhibition of energy generation is one of the important mechanisms of fungicides (56). Membrane-active substances such as Ag^+, Hg^+, Cu^{2+}, Cu^+, uranyl ions, bis(tributyltin) oxide, and other organic tin compounds or polyenic antibiotics (nystatin, candicidin, amphotericin, fungimycin, etc.) indirectly inhibit fungal glycolysis by causing sublethal damage to the cytoplasmic membrane. Fluoride ions are potent inhibitors of glycolysis because of their ability to form stable complexes with magnesium and

phosphate in certain enzymes involved in glycolysis. Fungicides showdomycin, gramicidin, 2,4-dinitrophenol, $CuSO_4$, and oleic acid inhibit respiration by uncoupling of oxidative phosphorylation. The fungicides carboxin, dexon, tridemorph, and antimycin A interfere with the function of the electron transport system in mitochondria. These inhibitory effects by fungicides can be attributed to a disturbance of membrane permeability and chemical reaction with sulfhydryl, carboxyl, phosphate, or amino groups of sensitive enzymes in fermentation or respiration metabolism. The fungicides terrazole and chloroneb disturb the respiratory process by causing lysis of mitochondria (56).

A mechanism of fungal resistance to the ergosterol-biosynthesis-inhibiting fungicides is energy-dependent efflux activity. For instance, the relatively high activity of energy-dependent efflux of fenarimol in resistant mutants of *Aspergillus nidulans* might be connected with constitutive membrane transport activities, which may be "fueled" with energy at the cost of other cell processes (e.g., germination, growth, and sporulation) (57). According to this principle, one might be able to manipulate fenarimol uptake by inhibiting efflux activity with other chemicals, possibly enhancing fungal toxicity of fenarimol against pathogenic fungi in agriculture. Recently, the energy-dependent drug efflux was identified as the mechanism of fluconazole resistance in *Candida glabrata* from a case of infection in which pre- and posttreatment isolates were available for comparison (58).

Catabolic End Products of Bacteria

Products of Sugar Fermentation. Typically, the EM pathway of glycolysis is used as the primary means of energy supply, yielding pyruvate, NADH, and 2 ATP from glucose (59). For the most part, species-specific variations in end products will depend on the selection of the means of recycling the electron carrying cofactor NAD and the selected substrate as an electron sink. The reductive steps are not associated with the supply of energy.

Lactic Acid. A familiar fermentation product is lactic acid (60), which is derived from the reduction of pyruvate with NADH to lactate by lactate dehydrogenase. A further means of energy conservation observed in several homolactic fermenting organisms, but which is also found in other organisms, is the generation of a proton motive force in association with the excretion of lactic acid out of the cells via a secondary transport mechanism (61). Another similar example is found in organisms performing malolactic fermentation, for example, *Lactobacillus lactis*. In this organism a combination of electrogenic malate uptake and malate/lactate antiport leads to the net efflux of 1 H^+ per malate molecule metabolized.

Ethanol. The formation of ethanol as the sole organic fermentation product is a two-step process. The first step is the decarboxylation of pyruvate to acetaldehyde plus CO_2 by pyruvate decarboxylase. NAD is recycled in the next step, the NADH-dependent reduction of acetaldehyde by alcohol dehydrogenase. Only organisms possessing pyruvate decarboxylase are capable of producing ethanol as a major fermentation product. In bacteria this enzyme is rare, and the sole group of organisms of industrial interest performing this fermentation comprises members of the genus *Zymomonas* (62), which use the E–D pathway of glucose oxidation to pyruvate, and therefore make only 1 ATP per glucose. This is in contrast to the production of ethanol in yeasts, where the E–M pathway is used for the synthesis of 2 ATP per glucose. Although *Zymomonas mobilis* is highly efficient in ethanol production, its range of fermentable substrates is very limited. In order to widen the range of usable fermentative substrates for ethanol synthesis, the genes for pyruvate decarboxylase and alcohol dehydrogenase have been transfected into *E. coli* (63).

Alcohols, Ketones, and Organic Acids. Traditionally, the biological generation of industrially useful short-chain fatty acids and solvents (64) has involved the anaerobic degradation of sugars by organisms dependent only on substrate-level phosphorylation to form ATP. From pyruvate, a further ATP can be generated via the synthesis of acetyl-CoA and acetyl phosphate, yielding acetate plus CO_2 as the end products. This reaction, however, does not recycle the NAD reduced during glycolysis. The electrons are used to reduce protons via a hydrogenase, yielding gaseous hydrogen (H_2) as an end product. An added benefit of this reaction is that the consumption of H^+ from inside the cell increases the internal pH and thus participates in the maintenance of the transmembrane ΔpH. Hydrogen may be generated from the fermentation of organic wastes (65). The buildup of hydrogen in the growth environment, however, causes a decrease in the thermodynamic efficiency of acetate production such that other electron sinks must be sought, but these various alternate end products are not necessarily associated with ATP synthesis. In the case of organisms capable of butyric acid synthesis, two molecules of acetyl-CoA can condense, forming acetoacetyl-CoA, which can then be reduced to butyryl-CoA, and finally butyrate, with the production of 1 ATP per 2 acetyl groups. Other products derived from the reduction of acetyl-CoA (ethanol, acetone, propanol, and butanol) are not associated with ATP synthesis (Fig. 9).

This shift in fermentation products, depending on hydrogen concentration, has been described in *Ruminococcus albus*, where cellulose fermentation normally yields acetate, ethanol, and hydrogen in a ratio of 1:1:2, but acetate and ethanol in a ratio of 7:1 when hydrogen is continuously removed (66). Conversely, a metabolic shift from acid production to enhanced solvent production has been observed in *Clostridium acetobutylicum* when hydrogen production is blocked by carbon monoxide, a hydrogenase inhibitor. It has also been noted that the addition of glycerol to glucose-fermenting cultures of *Clostridium acetobutylicum* also inhibits hydrogen and acid generation but enhances butanol production (67).

High concentrations of these various products (undissociated volatile fatty acids, ketones, and alcohols), because of their amphipathic character, dissolve into membrane lipids, affecting membrane fluidity and increasing membrane permeability. This results in a decrease of the transmembrane pH gradient and a reduction of the intracellular level of ATP (68) and thus limits the concentration of these products the cells can tolerate.

In certain species capable of solvent formation (alcohols, ketones), for example, *Clostridium acetobutylicum*, the

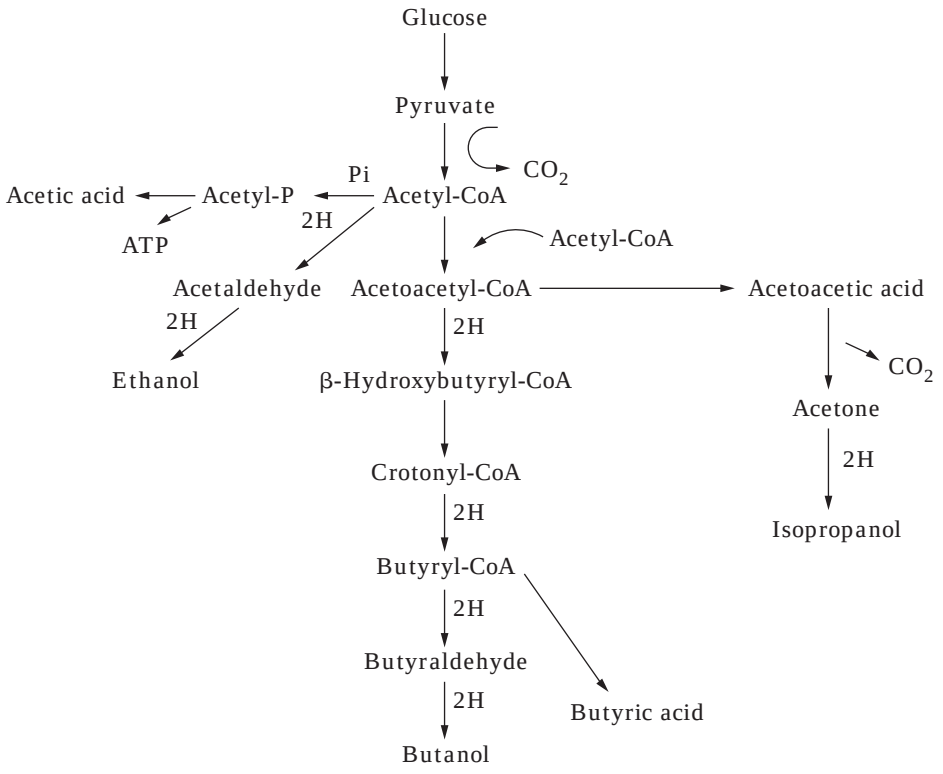


Figure 9. Butanol/acetone fermentation. This anaerobic metabolism is typical of the genus *Clostridium*. Various fermentation products are formed by reduction using NADH derived from glycolysis. The proportion of each product formed is dependent on the fermentation conditions.

drop in internal pH caused by the initial acidogenic phase can act as part of the triggering mechanism for solvent production (69). This reduction in proton motive force causes a drop in internal ATP concentration, which has been shown to act as one of the regulatory factors in turning on solvent production (70). Nonetheless, excessive butanol formation eventually leads to inability of maintaining a transmembrane pH gradient, lowers intracellular ATP levels, and prevents glucose uptake. Higher concentrations of alcohols in general have a negative impact on membrane integrity and membrane functions (71).

Acetic Acid. Acetic acid bacteria are obligate aerobes and members of the genera *Acetobacter* and *Gluconobacter*. They are capable of the conversion of sugars or other carbon substrates by fermentation to acetate with a maximum conversion of 66%, with the remaining carbon being evolved as CO_2 and the electrons serving to reduce protons to H_2 (72). These organisms have been used primarily for the synthesis of acetic acid from ethanol, whereby the oxidation of ethanol to acetate provides the electrons for O_2 reduction. In *Gluconobacter* the cells lack succinate dehydrogenase and therefore have an incomplete TCA cycle. Although *Acetobacter* does have a full TCA cycle and therefore can oxidize acetate to CO_2, its function is repressed in the presence of ethanol.

The oxidation of ethanol to acetate is a four-electron oxidation and involves two membrane-bound enzymes: alcohol dehydrogenase and acetaldehyde dehydrogenase. In *Gluconobacter suboxydans* and *Acetobacter aceti* both enzymes are localized on the outer surface of the cytoplasmic membrane and contain cytochrome *c* as well as a novel electron-transferring prosthetic group, pyrroloquinoline quinone (PQQ), with ubiquinone as the ultimate electron acceptor for these reactions (73). In view of the location of these enzymes on the outer surface of the membrane, two protons are left behind during each oxidation reaction participating in the formation of a proton gradient.

Electrons from reduced ubiquinone are then transferred to terminal ubiquinol oxidases. Depending on the growth conditions, the terminal oxidase of *Acetobacter aceti* contains a cytochrome a_1–type chromophore, which has been shown in proteoliposomes to translocate protons or a cytochrome *o* (74). In *Gluconobacter suboxydans* a cytochrome *o*–type ubiquinol oxidase can translocate protons in proteoliposomes in the presence of O_2 and ubiquinol (75).

In the overall synthesis of ascorbic acid, one step, the conversion of D-sorbitol to D-sorbose, is generally performed by *Gluconobacter suboxydans*, which possesses a D-sorbitol dehydrogenase that is bound to the outer surface of the cytoplasmic membrane. This is a cytochrome *c*–and flavin-containing protein that uses ubiquinone as an electron acceptor, which is reoxidized by terminal oxidases.

Homoacetogens. Homoacetogens are a set of anaerobes capable of stoichiometric conversion of glucose to three acetates (76). These organisms are also capable of acetate synthesis and growth as autotrophs on H_2/CO_2. The E–M pathway is used to yield the first two acetates from glucose, and the third acetate is derived from the reduction by hydrogen of the evolved CO_2 to methyl or formyl derivatives. In the reduction of CO_2 to the methyl level, ATP is hydrolyzed for the generation of formyl tetrahydrofolate, which is subsequently reduced to the methyl level. This methyl group is then transferred to a correnoid protein. The CO_2 in the carboxyl position is reduced to carbon monoxide by a carbon monoxide dehydrogenase (CODH). This same en-

zyme then catalyzes the formation of acetyl-CoA from this CO, the methyl group from the correnoid protein and HS-CoA. The release of acetate from acetyl-CoA via acetyl phosphate generates 1 ATP. Because of the consumption of ATP in the transfer of formate onto tetrahydrofolate, this pathway does not yield any net ATP via substrate-level phosphorylation.

Net energy is conserved in the form of a primary transmembrane ion gradient. In *Clostridium thermoaceticum*, acetogenesis from H_2 plus CO_2 generates a proton motive force. The $\Delta\mu_{H^+}$ is generated via a hydrogenase located on the outside surface of the membrane and a rudimentary electron transport chain consisting of two *b*-type cytochromes and menaquinone. The hydrogenase releases protons from the oxidation of H_2 on the outside of the membrane, and the electrons are carried through the membrane to the inner side where they are used for the generation of the methyl group of the acetate. The proton motive force thus generated can be used to drive ATP synthesis via an F_1F_0 ATPase.

The production of acetic acid may have a toxic effect. Although there is little diffusion through the cell membrane of the ion form (i.e., $R\text{-}COO^-$), weak acids can act as protonophores when they are present in their nondissociated form. In *C. thermoaceticum*, for example, it has been shown that, at the lower limits of the physiological pH (~pH 5), the concentration of undissociated acetic acid is sufficiently elevated to act as an effective protonophore, bringing protons into the cell and thus eliminating the proton motive force and decreasing the ATP pool size. This in turn stops growth and fermentation because ATP is necessary in the activation steps of glycolysis (77).

In *Acetobacterium woodii*, acetogenesis from H_2 plus CO_2 generates a primary transmembrane Na^+ gradient, which is used directly for ATP generation via a sodium-dependent ATPase (78). Na^+ translocation is probably associated with the transfer of the methyl group from tetrahydrofolate to the correnoid protein. The $\Delta\mu_{Na^+}$ is also necessary for flagellar motility in this organism. The capacity of acetogenic bacteria to use CO as well as CO_2 and H_2 can be used in the synthesis of organic carbon from syngas (synthesis gas) generated from the pyrolysis of coal.

Methane. Methane is one of the major end products of anaerobic biodegradation of organic material, including municipal and industrial organic wastes. Organisms that generate methane, the methanogens, are important in these environments because they allow mineralization of organic materials and generate an inflammable gas (79). Although they act on a limited range of substrates (H_2/CO_2, formate, acetate, methanol, methylamines), these represent the major products of organic polymer biodegradation. The consumption of hydrogen by the methanogens maintains a low environmental concentration, thereby reducing the amounts of alcohols and ketones produced by fermentative bacteria.

Methane is generated by H_2-dependent CO_2 reduction or conversion from methanol (80). The pathway of CO_2 reduction involves four two-electron reduction steps similar to those of the acetogenic pathway (Fig. 10). The major difference is the use of novel C_1-carrying cofactors. The reduction of CO_2 to the formyl level is catalyzed by formylmethanofuran dehydrogenase, which releases formyl methanofuran; the C_1 is then transferred to tetrahydromethanopterin, an analog of tetrahydrofolate, which becomes methylated. The methyl group is transferred to thioethanesulfonate methyl-CoM, which is then reduced to release methane.

The first step of methanogenesis from H_2 plus CO_2 involves a membrane-associated enzyme and may require energy in the form of an ion motive force to drive the reaction. The sites of energy conservation associated with the methanogenic pathway, as well as the protonophore sensitivity and the requirement for sodium, have been elucidated for *Methanosarcina*. There are two sites of ion translocation: the heterodisulfide reductase, which is associated with proton translocation, and the methyl tetrahydromethanopterin, HS-CoM methyltransferase, which is associated with sodium translocation. Thus both a proton and a sodium motive force are generated. Two different ion motive force–dependent ATP synthases have also been uncovered: a $\Delta\mu_{H^+}$-dependent A_0A_1 ATP synthase and a $\Delta\mu_{Na^+}$-dependent F_0F_1 ATP synthase (81). Proton translocation is now proposed to be associated with electron transport via a membrane-bound carrier called methanophenazine (82).

Poly-β-Hydroxyalkanoates

Poly-*β*-hydroxyalkanoates (PHA) serve as metabolic carbon reserves in various microorganisms and have found commercial value as biodegradable substitutes for petroleum-based plastics (83). The polymers are produced under conditions of metabolic stress in the presence of excess carbon or other energy source when, for example, nitrogen, phosphate, or oxygen limitation prevents balanced growth and acetyl-CoA accumulates in the cell. The condensation of acetyl groups leads to the production of *β*-hydroxybutyrate, which is esterified into long chains. Under such conditions, PHA can accumulate in the cells, forming up to 90% of the cell mass. Depending on the organism used, there can be simultaneous growth and PHA formation, or a serial process whereby the organism is allowed to buildup biomass, followed by a second stage of PHA accumulation. Although nitrogen limitation is commonly used as trigger for PHA production, the continued addition of a complex organic nitrogen source (e.g., tryptone or casein) during PHA synthesis from glucose was shown to enhance PHA synthesis without significant effect on residual cell mass yield. It has been suggested that this enhanced yield is due to the prevention of use of acetyl-CoA for amino acid and protein synthesis, thus permitting a greater proportion of the acetyl-CoA produced from glucose to form PHA rather than other cell products. Varying substrates or growth conditions can also lead to an increased yield of product and variation in the type of alkanoates formed. This permits the formation of polyesters with different physical properties. For example, *β*-OH-valerate is formed from the condensation of propionyl-CoA and acetyl-CoA when propionate is added to the growth medium.

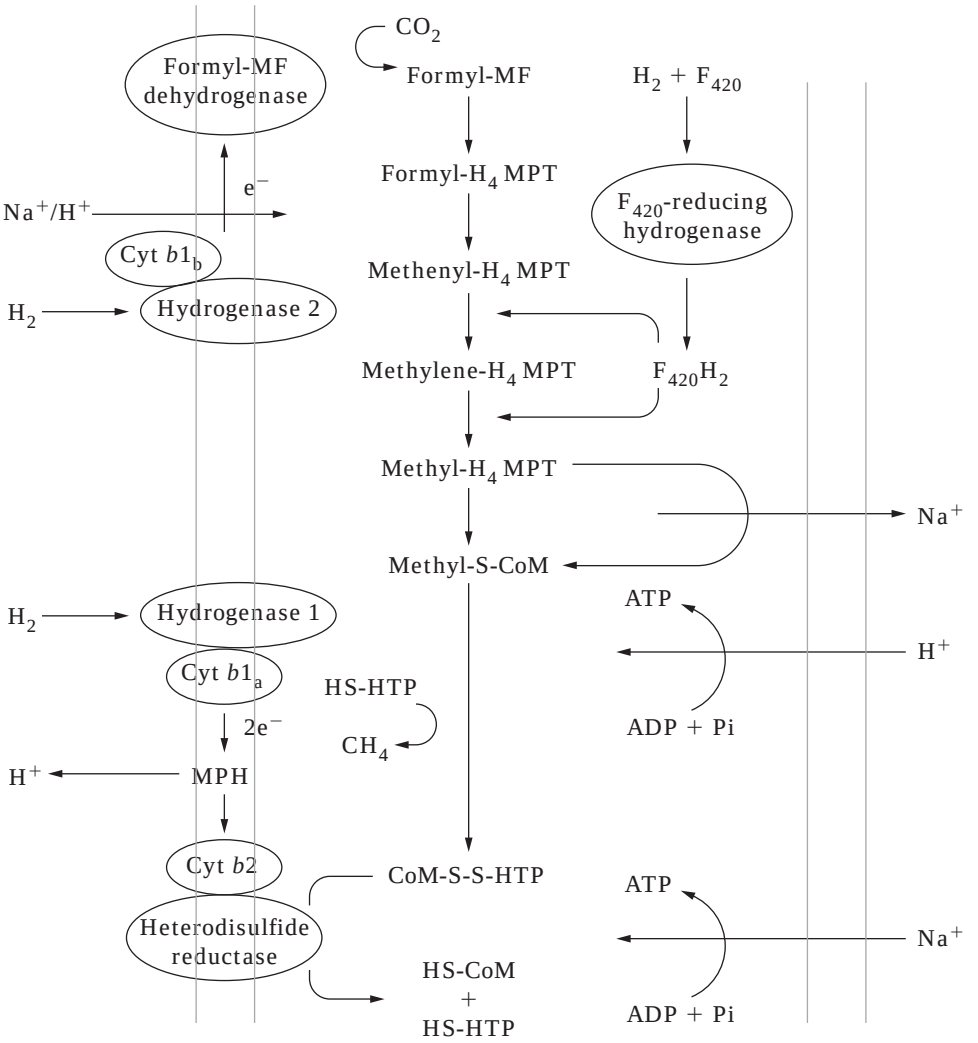


Figure 10. Scheme of methanogenesis. Methane is formed from hydrogen and carbon dioxide in a metabolic pathway present in *Methanosarcina*. The dashed lines represent the cytoplasmic membrane. A similar scheme is expected for other methanogens. H_4MPT, tetrahydromethanopterin; HS-CoM, coenzyme M (2-mercaptoethanesulfonate); HS-HTP, 7-mercaptohectanoyl threonine phosphate; MPH, methanophenazine; MF, methanofuran.

Enzyme Production in Bacteria

Aside from the proteins and enzymes expressed and produced in *E. coli* by recombinant DNA technology, there are various enzymes produced by a wide variety of bacteria that are used commercially (α-amylases, xylanases, lipases, DNA polymerases, etc.). Many of these are degradative enzymes used to break down growth substrates that cannot enter directly into the cells. Such enzymes are produced in large quantities and exported out of the cell. The need for highly stable and robust proteins capable of functioning at high temperatures and/or pH has led to the isolation and characterization of novel extremophilic organisms capable of degrading organic polymers (84,85).

A representative example of such organisms is *Pyrococcus furiosus*, a hyperthermophilic archaean that grows optimally at 100 °C and pH 6.4 and produces a thermostable α-amylase of potential use in the starch hydrolysis industry, as well as a high-fidelity DNA polymerase (86). The need for optimization of the production of these enzymes has required the characterization of the associated fermentation pathway (87). A modified EM pathway of glycolysis that employs ADP-dependent glucokinase and phosphofructokinase is used rather than the conventional ATP-utilizing enzymes. A further modification is that the oxidation of glyceraldehyde-3-phosphate to 3-phosphoglycerate occurs directly, using a glyceraldehyde-3-phosphate: ferredoxin oxidoreductase and does not involve the synthesis of 1,3-bisphosphoglycerate and the subsequent synthesis of ATP. ATP is generated from phosphoenolpyruvate via pyruvate kinase, the 2 ATP formed per glucose in this reaction balance the two high-energy phosphate bonds

used to activate glucose. The only net ATP produced is from the oxidation of pyruvate to CO_2 and acetate via acetyl-CoA and acetyl phosphate. The flux of carbon through acetate, however, is dependent on the presence or absence of electron sinks. In the presence of sulfur (S^0), the electrons released during glycolysis are passed on to sulfur via a sulfhydrogenase, generating H_2S (88). In the absence of sulfur, the same enzyme is capable of reducing protons to hydrogen. However, as the hydrogen concentration in the medium increases, alanine accumulates in the growth medium and is associated with a decreased growth rate (89).

A major difficulty in batch cultures of this organism in the absence of S^0, is the accumulation of H_2, the production of alanine, and the accumulation of acetate, which can be toxic at higher concentrations. Fermenter design strategies have been developed to enhance growth yields and increase cell densities. The use of a dialysis membrane reactor permits the cells to reach 10-fold higher cell densities and a lower alanine level (89).

Production from Animal Cells

Commercial products derived from animal cells in culture include viral vaccines, monoclonal antibodies (Mabs), and recombinant proteins. Table 3 shows examples of the protein products of animal cell cultures licensed for therapeutic use. In all cases, productivity is dependent upon protein biosynthesis, which in turn must be supported by an adequate provision of intracellular energy. From data based on the macromolecular composition of a cell, it can be estimated that 60% of cellular energy is utilized for protein synthesis, ATP being required for the condensation of amino acids (90). Although some of these proteins can be synthesized in bacteria, one of the major advantages of using animal cells is that the resulting proteins are glycosylated (i.e., have attached carbohydrate structures). This is often important for the therapeutic use of the proteins.

During culture of a cell line that has been genetically engineered to secrete a specific protein, the energy demands of cell growth must be balanced with the demand for synthesis of the selected protein. There are many reports that show an inverse relationship between growth and antibody production (91). Typically, the specific rate of Mab production (q_{Mab}) occurs at 0.05 pg/cell min during the growth of a hybridoma, and this amounts to about 20% of the total cellular protein synthesis. If the growth rate is decreased, for example, in continuous culture, then the q_{Mab} may double in value as more of the cellular energy is diverted from growth to synthesis of the single secreted protein. However, there is a limit to the reduction in the growth rate of lymphocyte hybridomas beyond which apoptosis (programmed cell death) causes the rapid loss of the cell population (92).

In recombinant protein synthesis there may be an energetic burden placed on the metabolism of a cell following gene amplification. A balance must be established in the cell between the requirements of growth and protein productivity. Gu et al. (93) studied a recombinant CHO cell line that showed high expression of β-galactosidase by coamplification of the *lacZ* and *dhfr* genes. The genes were amplified 100 times by selective pressure from increasing concentrations of methotrexate, which is an inhibitor of the essential enzyme dihydrofolate reductase, the product of the *dhfr* gene. However, as a result of gene amplification, the specific cell growth rate was reduced by 62% because of the metabolic burden placed upon the cell line.

ENHANCEMENT OF PRODUCTIVITY: METABOLIC MANAGEMENT OR ENGINEERING

In the development of an industrial bioprocess, there is a strong incentive to maximize the synthetic capacity of the producer cell line by a form of metabolic management. This may be achieved by an understanding and control of the limiting factors for growth or cell yield. For example, the availability of oxygen may be critical in the profile of the energy pathways of a microorganism and is an important parameter for the production of metabolic end products. In the case of animal cell cultures, controlling the availability of glutamine is essential in attempting to limit the accumulation of ammonia, which is a major parameter in realizing a respectable cell yield. This process of metabolic management may be put to greatest effect by the development of mathematical programming models based upon metabolic flux data obtained by radioactive tracers or ^{13}C-NMR (94).

Further enhancement of productivity of a bioprocess may be gained by strain improvement, which is often achieved by genetic mutation. For example, high-lysine-producer strains of *Corynebacterium glutamicum* can be obtained by selective mutation resulting from feedback resistance of the aspartate kinase enzyme (95). Further improvement may result from mutations in the enzymes of energy metabolism. Often the metabolic basis of a synthetic enhancement is unknown but may be analyzed by changes in the pattern of metabolic flux.

A more selective form of productivity enhancement may be achieved by metabolic engineering as a means to maximize the flow of carbon and energy to the specific pathway of interest (96,97). This involves transfection of the producer cell line with a gene to manipulate the enzymatic, transport, or regulatory function of the cell. A particularly successful application of this technique is in the ability to grow a Chinese hamster ovary (CHO) cell line in a protein-free medium following transfection with genes capable of expressing essential factors such as insulin-like growth factor, transferrin, or cyclin E (98,99). Another application involves the expression of the gene for hemoglobin isolated from *Vitreoscilla* in a variety of aerobic industrial microorganisms (99). Because of the central importance of oxygen to energy metabolism, this technique has been successful in causing a significant productivity enhancement in a long list of cells including bacteria, fungi, and animal cells. With an increased understanding of the energy metabolism of industrially important producer cell lines, there comes a greater likelihood of success in designing the required phenotypic characteristics by metabolic engineering.

BIBLIOGRAPHY

1. D.M. Karl, *Microbiol. Rev.* **44**, 739–796 (1980).
2. D.E. Atkinson, *Biochemistry* **7**, 4030–4034 (1968).

3. J.S. Swedes, R.J. Sedo, and D.E. Atkinson, *J. Biol. Chem.* **250**, 6930–6938 (1975).
4. K.B. Anderson and K. von Meyenburg, *J. Biol. Chem.* **252**, 4151–4156 (1977).
5. G.M. Gadd, in D.R. Berry ed., *Physiology of Industrial Fungi*, Blackwell Scientific Publications, Oxford, 1988, pp. 21–57.
6. D.H. Griffin, *Fungal Physiology*, 2nd ed., Wiley-Liss, New York, 1994.
7. P.A. Botham and C. Ratledge, *J. Gen. Microbiol.* **114**, 361–375 (1979).
8. E.A. Dawes, *Microbial Energetics*, Blackie & Son, Glasgow, 1986, pp. 21–51.
9. T.G. Cartledge, in D.R. Berry, I. Russell, and G.G. Stewart eds., *Yeast Biotechnology*, Allen and Unwin, London, 1987, pp. 311–342.
10. G.C. Ferreira and P.L. Pedersen, *J. Bioenerg. Biomembranes* **25**, 483–491 (1993)
11. N.A.R. Gow and G.M. Gadd, *The Growing Fungus*, Chapman & Hall, London, 1995, p. 225.
12. M. Kobayashi, Y. Matsuo, A. Takimoto, S. Suzuki, F. Maruo, and H. Shoun, *J. Biol. Chem.* **271**, 16262–16267 (1996).
13. V.P. Skulachev, *Eur. J. Biochem.* **208**, 203–209 (1992).
14. M.J. Zidwick, in D.B. Finkelstein and C. Ball eds., *Biotechnology of Filamentous Fungi*, Butterworth-Heinemann, Boston, 1991, pp. 303–334.
15. G.J.M. Hersbach, C.P. Van der Beek, and P.W.M. Van Dijck, in E.J. Vandamme ed., *Biotechnology of Industrial Antibiotics*, Marcel Dekker, New York, 1984, pp. 45–140.
16. B. Buckland, K. Gbewonyo, T. Hallada, L. Kaplan, and P. Masurekar, in A.L. Demain, G.A. Somkuti, J.C. Hunter-Cevera, and H.W. Rossmoore eds., *Novel Microbial Products for Medicine and Agriculture*, Elsevier on behalf of Society for Industrial Microbiology, 1989, pp. 161–169.
17. M. Hosobuchi, T. Shioiri, J. Ohyama, M. Arai, S. Iwado, and H. Yoshikawa, *Biosci. Biotech. Biochem.* **57**, 1414–1419 (1993).
18. S.N. Agathos, C. Madhosingh, J.W. Marshall, and J. Lee, in *Annals New York Academy of Sciences*, 1987, pp. 657–662.
19. N.D. Lindley, J.F. Pedley, S.P. Kay, and M.T. Heydeman, *Int. Biodeterior.* **22**, 281–287 (1986.).
20. K.H. Jones, P.W. Trudgill, and D.J. Hopper, in A. Bousher, M. Chandra, and R. Edyvean eds., *Biodeterior. Biodegrad. 9. Int. Biodeterior. Biodegrad. Symp., [Proc.], 9th* (meeting date 1993), Institution of Chemical Engineers, Rugby, U.K., 1995, pp. 470–475.
21. F.J. Weber, K.C. Hage, and J.A. de Bont, *Appl. Environ. Microbiol.* **61**, 3562–3566 (1995).
22. J.F. Morgan, H.J. Morton, and R.C. Parker, *Proc. Soc. Exp. Biol. Med.* **73**, 1–8 (1950).
23. H. Eagle, *Science* **122**, 501 (1995).
24. H.R. Zielke, P.T. Ozand, J.T. Tildon, D.A. Sevdalian, and M. Cornblath, *J. Cell Physiol.* **95**, 41–48 (1978).
25. B.M. Wice and D. Kennell, *J. Biol. Chem.* **258**, 13134–13140 (1983).
26. T.E. Hassell and M. Butler, *J. Cell Sci.* **96**, 501–508 (1990).
27. T. Imamura, C.L. Crespi, W.G. Thilly, and H. Brunengraber, *Anal. Biochem.* **124**, 353–358 (1982).
28. W.L. McKeehan, in Morgan, M.J. ed. *Carbohydrate Metabolism in Cultured Cells*, Plenum, New York, 1986, pp. 111–150.
29. K.W. Lanks and P-W. Li, *J. Cell Physiol.* **135**, 151–155 (1988).
30. M. Schneider, I.W. Marison, U. von Stockar, *J. Biotechnol.* **46**, 161–185 (1996).
31. M. Butler and A. Christie, *Cytotechnology* **15**, 87–94 (1994).
32. M.C. Glassy, J.P. Tharakan, and P.C. Chau, *Biotech. Bioeng.* **32**, 1015–1028 (1988).
33. M. Butler, N. Huzel, and N. Barnabé, *Biochem J.* **322**, 615–623 (1997).
34. R.L. Greasham, in D.B. Finkelstein and C. Ball eds., *Biotechnology of Filamentous Fungi*, Butterworth-Heinemann, Boston, 1992, pp. 65–87.
35. F.C. Neidhardt ed., Escherichia coli *and* Salmonella *Cellular and Molecular Biology*, ASM Press, Washington D.C., 1996.
36. G. Gottschalk, *Bacterial Metabolism*, Springer-Verlag, New York, 1986.
37. A.H. Stouthamer, *Int. Rev. Biochem.* **21**, 1–47 (1979).
38. M.W. Glacken, R.J. Fleischaker, and A.J. Sinskey, *Batik. Bioeng.* **28**, 1376–1389 (1986).
39. D. Petch and M. Butler, *J. Cell Physiol.* **161**, 71–76 1994.
40. H.D. Crabtree, *Biochem. J.* **23**, 536–545 (1929).
41. M. Guppy, E. Greiner, and K. Brand, *Eur. J. Biochem.* **212**, 95–99 (1993).
42. D.C.H. Jan, D. Petch, N. Huzel, and M. Butler, *Biotech. Bioeng.* **54**, 153–164 (1997).
43. L. Haggstrom, in R.E. Spier, J.B. Griffiths, and B. Meignier eds., *Production of Biologicals from Animal Cells in Culture* (ESACT 10th Mtg.) Butterworth-Heinmann Oxford, 1991, pp. 79–81.
44. J.C. Street, A.M. Delort, P.S.H. Braddock, and K.M. Brindle, *Biochem. J.* **291**, 485–492 (1993).
45. H.A. Jenkins, M. Butler, and A.J. Dickson, *J. Biotechnol.* **23**, 167–182 (1992).
46. H.W. Doelle, *Microbial Metabolism*, Dowden, Hutchinson & Ross, Stroudsburg, Pa., 1974, pp. 8–79.
47. M.J. Zidwick, in D.B. Finkelstein and C. Ball eds., *Biotechnology of Filamentous Fungi*, Butterworth-Heinemann, Boston, 1991, pp. 303–334.
48. J.W. Bennett, in A.L. Demain and N.A. Solomon eds., *Biology of Industrial Microorganisms*, Butterworths, Boston, 1985, pp. 359–406.
49. H.L. Kornberg and A.H. Romano, *Process Biochem.* **3**, 40–42 (1968).
50. M.D. Greenspan and J.B. Yudkovitz, *J. Bacteriol.* **162**, 704–707 (1985).
51. G.J.M. Hersbach, C.P. Van der Beek, and P.W.M. Van Dijck, in E.J. Vandamme ed., *Biotechnology of Industrial Antibiotics*, Marcel Dekker, New York, 1984, pp. 45–140.
52. A. Lawen, R. Traber, D. Geyl, R. Zocher, and H. Kleinkauf, *J. Antibiot.* **42**, 1283–1289 (1989).
53. A.L. Demain, Y. Aharonowitz, and J.F. Martin, in L.C. Vining ed., *Biochemistry and Genetic Regulation of Commercially Important Antibiotics*, Addison-Wesley, Advanced Book Program/World Science Division, Reading, Mass., 1983, pp. 49–72.
54. S.W. Drew and D.A. Wallis, in J.W. Bennett and A. Ciegler eds., *Secondary Metabolism and Differentiation in Fungi*, Marcel Dekker, New York, 1983, pp. 35–54.
55. R.J. Buchanan, S.B. Jones, W.V. Gerasimowicz, L.L. Zaika, H.G. Stahl, and L.A. Ocker, *Appl. Environ. Microbiol.* **53**, 1224–1231.
56. H. Lyr, in M.R. Siegel and H.D. Sisler eds., *Antifungal Compounds*, Marcel Dekker, New York, 1977, pp. 301–332.

57. M.A. de Waard and A. Fuchs, in J. Dekker and S.G. Georgopoulos eds., *Fungicide Resistance in Crop Protection*, Centre for Agricultural Publishing and Documentation, Wageningen, The Netherlands, 1982, pp. 87–100.
58. T. Parkinson, D.J. Falconer, and C.A. Hitchcock, *Antimicrob. Agents Chemother.* **39**, 1696–1699 (1995).
59. G. Gottschalk, *Bacterial Metabolism*, 2nd ed., Springer-Verlag, New York, 1986.
60. J.H. Litchfield, *Adv. Appl. Microbiol.* **42**, 4595 (1996)
61. R. Otto, A.S.M. Sonnenberg, H. Veldkamp, and W.N. Konings, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5502–5506 (1980).
62. J. Swings and J. de Ley, *Bacteriol. Rev.* **41**, 1–46 (1977).
63. L.O. Ingram, F. Alterthum, K. Ohta, and D.S. Beall, *Develop. Ind. Microbiol.* **31**, 21–30 (1990).
64. J.G. Zeikus, *Ann. Rev. Microbiol.* **34**, 423–464 (1980).
65. S. Roychowdhury, D. Cox, and M. Levandowsky, *Int. J. Hydrogen Energy* **13**, 407–410 (1988).
66. S.G. Pavlostathis, T.L. Miller, and M. Wolin, *Appl. Microbiol. Biotechnol.* **33**, 109–116 (1990).
67. I. Vasconcelos, L. Girbal, and P. Soucaille, *J. Bacteriol.* **176**, 1443–1450 (1994).
68. P. Dürre, H. Bahl, and G. Gottschalk, in L.E. Erickson and D.Y. Fung eds., *Handbook of Anaerobic Fermentations*, Marcel Dekker, New York, 1988, pp. 187–206.
69. B. McNeil and B. Kristiansen, *Adv. Appl. Microbiol.* **31**, 61–92 (1986).
70. H. Grupe and G. Gottschalk, *Appl. Environ. Microbiol.* **58**, 3896–3902 (1992).
71. L.O. Ingram and T. Buttke, *Adv. Microbiol. Phys.* **25**, 253–300 (1984).
72. T. Asai, *Acetic Acid Bacteria Classification and Biochemical Activities*, University Park Press, Baltimore, 1968.
73. K. Matsushita, H. Ebisuya, M. Ameyama, and O. Adachi, *J. Bacteriol.* **174**, 122–129 (1992).
74. K. Matsushita, E. Shinagawa, O. Adachi, and M. Ameyama, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9863–9867 (1990)
75. K. Matsushita, E. Shinagawa, O. Adachi, and M. Ameyama, *Biochim. Biophys. Acta* **894**, 304–312 (1987).
76. H.L. Drake ed., *Acetogenesis*, Chapman & Hall, New York, 1994.
77. G. Wang and D.I.C. Wang, *Appl. Environ. Microbiol.* **47**, 294–298 (1984).
78. V. Muller and S. Bowien, *Arch. Microbiol.* **164**, 363–369 (1995).
79. L. Daniels, *Trends Biotechnol.* **2**, 91–98 (1984).
80. J.G. Ferry ed., *Methanogenesis*, Chapman & Hall, New York, 1993.
81. U. Deppenmeier, V. Müller, and G. Gottschalk, *Arch. Microbiol.* **165**, 149–163 (1996).
82. H.-J. Abken, M. Tietze, J. Brodensen, S. Baumer, U. Beifuss, and U. Deppenmeier, *J. Bacteriol.* **180**, 2027–2032 (1998).
83. C. Sasikala and C.V. Ramana, *Adv. Appl. Microbiol.* **42**, 97–218 (1996).
84. R.A. Herbert, *Trends Biotechnol.* **10**, 395–402 (1992).
85. S.E. Lowe, M.K. Jain, and J.G. Zeikus, *Microbiol. Rev.* **57**, 451–509 (1993).
86. K.S. Lundberg, D.D. Shoemaker, M.W.W. Adams, J.M. Short, J.A. Sorge, and E.J. Mathur, *Gene* **108**, 1–6 (1991).
87. S.W.M. Kengen, A.J.M. Stams, and W.M. de Vos, *FEMS Microbiol. Rev.* **18**, 119–137 (1996).
88. R.N. Schicho, K. Ma, M.W.W. Adams, and R.M. Kelly, *J. Bacteriol.* **175**, 1823–1830 (1993).
89. M. Krahe, G. Antranikian, and H. Märkl, *FEMS Microbiol. Rev.* **18**, 271–285 (1996).
90. A.H. Stouthamer, in J.R. Quayle ed., *International Review in Biochemistry, Microbial Biochemistry*, University Park Press, Baltimore, 1979, pp. 1–47.
91. M. Leno, O-W. Merten, and J. Hache, *Biotechnol. Bioeng.* **39**, 596–606 (1992).
92. R.P. Singh, M. Al-Rubeai, C.D. Gregory, and A.N. Emery, *Biotech. Bioeng.* **44**, 720–726 (1994).
93. M.B. Gu, P. Todd, and D.S. Kompala, *Cytotechnology* **18**, 159–166 (1996).
94. S.M. See, J.P. Dean and G. Dervakos, *Appl. Biochem. Biotechnol.* **60**, 251–301 (1996).
95. J.J. Vallino and G. Stephanopolous, *Biotech. Bioeng.* **41**, 633–646 (1993).
96. J.E. Bailey, *Science* **252**, 1668–1675 (1991).
97. G. Stephanopoulos and J.J. Vallino, *Science* **252**, 1675–1681 (1991).
98. S.C.O. Pak, A.M.N. Hunt, M.W. Bridges, M.J. Sleigh, and P.P. Gray *Cytotechnology* **22**, 139–146 (1996).
99. J.E. Bailey, A. Sburlati, V. Hatzimanikatis, K. Lee, W.A. Renner, and P.S. Tsai *Biotech. Bioeng.* **52**, 109–121 (1996).

See also AMMONIA TOXICITY, ANIMAL CELLS; BIOENERGETICS OF MICROBIAL GROWTH; CELL CYCLE; CELL CYCLE, EUKARYOTES; ELECTRON TRANSPORT; PYRUVATE, PRODUCTION USING DEFECTIVE ATPASE ACTIVITY.

ENZYMES, BAKING, BREAD MAKING

JOAN QI SI
Novo Nordisk Ferment Ltd.
Dittingen, Switzerland

KEY WORDS

Antistaling
Crumb elasticity
Dough
Enzyme
Firmness
Gluten strengthening
Lipase
Maltogenic α-amylase
Rheology
Xylanase

OUTLINE

INTRODUCTION

Bread is one of the most common and low-cost foods and is associated with traditions in many different countries. Bread is also closely related to *biotechnology*, which is synonymous with *high technology*. The term that relates these two subjects is *enzyme*.

Consumers have certain criteria for bread, including appearance, freshness, taste, flavor, variety, and a consistent quality. It is a great challenge for the baking industry to fulfill these criteria for several reasons. Firstly, the main ingredient of the bread—flour—varies due to varied wheat qualities, weather, and milling technology, although millers attempt to blend the wheat sorts to produce a good and consistent baking flour quality. It often proves difficult to satisfy both high-quality and low-cost standards at the same time. Secondly, because bread traditions differ, the baking industry needs various qualities of ingredients and procedures. For instance, an English toast bread with fine crumb structure and very soft texture is not popular with the French, who want baguettes with crispy crusts, large holes, and good chewiness in the crumb. Thirdly, healthier products are becoming more popular due to more intensive cultural exchanges and changes in consumer preferences. Some new variety breads can be made by simple adjustment of the formulation or baking procedure. However, others require the bakers to develop new techniques. Therefore, both millers and bakers need agents or process aids such as chemical oxidants, emulsifiers, and enzymes to standardize the quality of the products and diversify the product range.

For decades enzymes such as malt and microbial α-amylases have been used for bread making. Due to the changes in the baking industry and the demand for more varied and natural products, enzymes have gained more and more importance in bread formulations. Through new and rapid developments in biotechnology, a number of new enzymes have recently been made available to the baking industry. One example is pure xylanase, with single activity instead of traditional hemicellulase preparations, which improves the dough machinability. A lipase has a gluten-strengthening effect that results in more stable dough and improved bread quality, and a maltogenic α-amylase has a unique antistaling effect.

This article reviews the effects of these enzymes. It also presents synergistic effects when these enzymes are combined with either each other or traditional enzymes such as fungal α-amylase.

FUNGAL α-AMYLASES FOR BREAD MAKING

Wheat and thus wheat flour contain endogenous and indigenous enzymes, mainly amylases. However, the level of amylase activity varies from one type of wheat to another. The amount of α-amylases in most sound, ungerminated wheat or rye flours is negligible (1). Therefore, most bread flours must be supplemented with α-amylases, added in the form of malt flour or fungal enzymes.

Many methods are available for the determination of amylase activities, as reviewed by Kruger et al. (2). Other methods such as falling number (FN) and Brabender amylograph are used by the baking industry and millers to determine the amylase content correlating to bread making. Although FN is excellent for measuring the activity of cereal amylases including added malt flour, it is not suitable for measuring the activity of fungal α-amylases. Because fungal α-amylases are generally less thermostable, they are inactivated at temperatures near 65 °C. Therefore, fungal α-amylases cannot be detected by the standard FN method, which is conducted at 100 °C. Perten et al. (3) reported a modified FN method for measuring the fungal α-amylase activity in flour. However, in practice (B. Allvin, personal communication, January 1996), the method is less suitable for routine analysis because it is necessary to define the measuring conditions for each type of flour (because the indigenous and endogenous amylases vary in different flours) to determine the modified FN with added fungal α-amylases.

Fungal α-amylases act on the damaged starch, which varies depending on wheat sort and milling condition. Generally, flour made from hard winter wheat contains more damaged starch than soft wheat. The α-amylases widely used in the baking industry can hydrolyze amylose and amylopectin to release soluble intermediate-size dextrins of DP2–DP12 (4). α-Amylases provide fermentable sugar, which results in an increased volume, better crust color, and improved flavor. Due to hydrolysis of the damaged starch, a suitable dosage of α-amylases results in a desirable dough softening. However, extensive degradation of the damaged starch due to an overdose of α-amylases leads to sticky dough. Fungal glucoamylases are a possibility but less common as baking enzymes. Literature on glucoamylases for baking is scarce (5). Figure 1 illustrates the effect of a fungal amylase on bread quality in terms of bread volume, crumb structure, and dough characteristics.

The volume and crumb structure improve with increasing dosage of fungal α-amylases. Although a high dosage can provide a larger volume increase, the dough would be too sticky to work with. The optimum dosage is thus defined as the dosage with maximum reachable volume without a sticky dough. For the examples in Figure 1, the optimum dosage for both flours is 15 FAU/kg flour.

MALTOGENIC α-AMYLASE FOR ANTISTALING

Bread staling is responsible for significant financial losses. An estimated 3–5% of all baked goods produced in the United States are discarded every year, representing a value of $1 billion (6). A maltogenic α-amylase has been

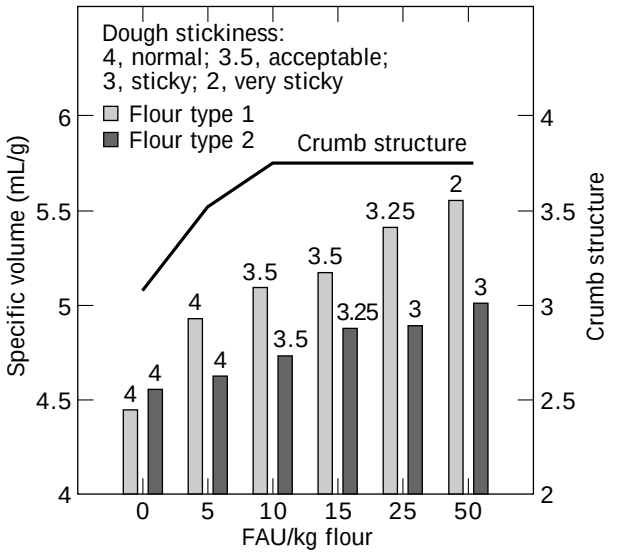


Figure 1. Dosage response of fungal amylase in different flours. FAU, units of fungal α-amylase. *Source*: Novo Nordisk Ferment.

found to have unique antistaling effects (7,8). It can degrade both amylose and amylopectin at the gelatinization temperature and produces mainly α-maltoses and a small amount of dextrins of DP1–DP12.

A study of the mechanism of antistaling effects of various amylases compared with monoglycerides (4) showed that this maltogenic α-amylase not only has a substantially improved antistaling effect compared with fungal α-amylases but also improves the elasticity of the bread crumb, which is an important factor that influences the palatability of bread. Both crumb softness and crumb elasticity are important characteristics for the description of crumb freshness perceived by consumers. Softness indicates the force needed to compress the crumbs, and elasticity indicates the resiliency or resistance given by the crumb while being pressed. The two texture characteristics may not necessarily correspond with each other; that is, the softest bread may not necessarily have the most elastic crumb texture or vice versa. Bread staling results in reduction of crumb elasticity and increase of crumb firmness.

Commercial amylases used in the baking industry are generally α-amylases that specifically hydrolyze the α-1,4 glycosidic linkages of amylose and amylopectin molecules in starch. Most known products are cereal enzymes derived from barley malt and the fungal α-amylase from *Aspergillus oryzae*. These amylases are effective in partially hydrolyzing damaged starch and are often added to flour as flour correction to develop desirable properties, such as oven spring and brown color to the crust. These amylases have limited antistaling effect due to their limited thermostability and are for the most part inactivated before the onset of starch gelatinization during baking. It was reported by Si and Simonsen (4) that fungal α-amylase, although it reduces the crumb firmness as shown in Figure 2, has no effect on reducing starch retrogradation.

Pullulanase, which specifically hydrolyzes the α-1,6 glycosidic linkages of amylopectin, does not have a positive effect and may even have a negative effect on antistaling.

The maltogenic α-amylase is a unique α-amylase that produces a significantly softer bread and maintains a high level of crumb elasticity during storage, as shown by Figure 2. Addition of maltogenic α-amylase in sponge and dough breads can prolong shelf life for at least 4 days longer compared with 0.5% powdered distilled monoglycerides, as shown by Figure 3. The crumb softness of the loaf with maltogenic α-amylase at day 7 was as soft as the loaf with 0.5% distilled monoglycerides at day 3, whereas the elasticity of the loaf with maltogenic α-amylase at day 7 is even higher than the loaf with distilled monoglycerides at day 3.

Maltogenic α-amylase has a thermostability between that of fungal α-amylase and thermostable bacterial α-amylase. Therefore, it is able to hydrolyze the glycosidic linkages during the gelatinization of starch during the baking process, but it does not excessively degrade the starch because it is inactivated during the later stage of baking.

Another major advantage of maltogenic α-amylase is its tolerance of overdosing during the bread-making process in the bakery. Too much fungal α-amylase can cause sticky dough or overbrowning of the crust. The bacterial α-amylase can easily be overdosed. Its pure endo action excessively degrades the starch during baking, causing collapse of the bread immediately after baking and sticky crumb during retail storage. Maltogenic α-amylase does not affect the dough rheological property because of its low activity at a temperature less than 35 °C. It is highly active only at the temperature during starch gelatinization and does not excessively degrade the starch but mainly produces small, soluble dextrins. The overdosing risk is thus much lower than with the other two types of amylases.

The bacterial α-amylases from *Bacillus subtilis* are able to inhibit staling by hydrolyzing glycosidic linkages within the amorphous areas of gelatinized starch. However, because of the high degree of thermostability, the enzymes can persist throughout baking and produce an excessive level of soluble dextrins. As a result the final product is often unacceptable, with a gummy or even sticky crumb texture that causes problems during slicing and retail storage. Figure 3 shows that the bacterial α-amylase gave the softest crumb but also a very gummy crumb with extremely low elasticity. The excessive degradation of starch during baking and storage creates a sticky crumb.

Soy β-amylase has been reported to have an antistaling effect. Our results, based on both sponge and dough and straight dough procedures, show that 3,000 BANU/kg flour, which is a quite high dosage, has a similar antistaling effect as 0.5% powdered distilled monoglycerides. The acid fungal α-amylase at a dosage of 100–200 AFAU/kg flour has an effect similar to 0.5% distilled monoglycerides (Fig. 3).

To understand the mode of action of this maltogenic α-amylase on bread staling, the effects of enzymes on the starch retrogradation were measured by differential scanning colorimetry (DSC), as shown in Figure 4. The endotherm area indicates the retrogradation of wheat starch. The endotherm area per gram starch gel containing monoglycerides was lower than the control, indicating its effect on retarding starch retrogradation. The endotherm area per gram of the starch treated with maltogenic α-amylase

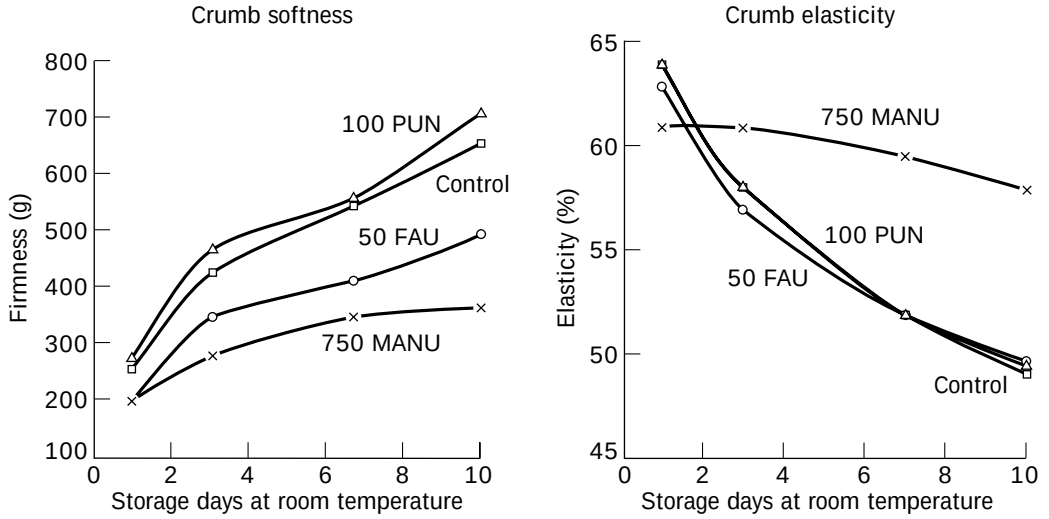


Figure 2. Effect of maltogenic α-amylase compared with fungal α-amylase and pullulanase on crumb softness and elasticity in toast bread. Straight dough procedure, controlled volume. FAU, units of fungal α-amylase/kg flour; MANU, units of maltogenic α-amylase/kg flour; PUN, units of pullulanase/kg flour.

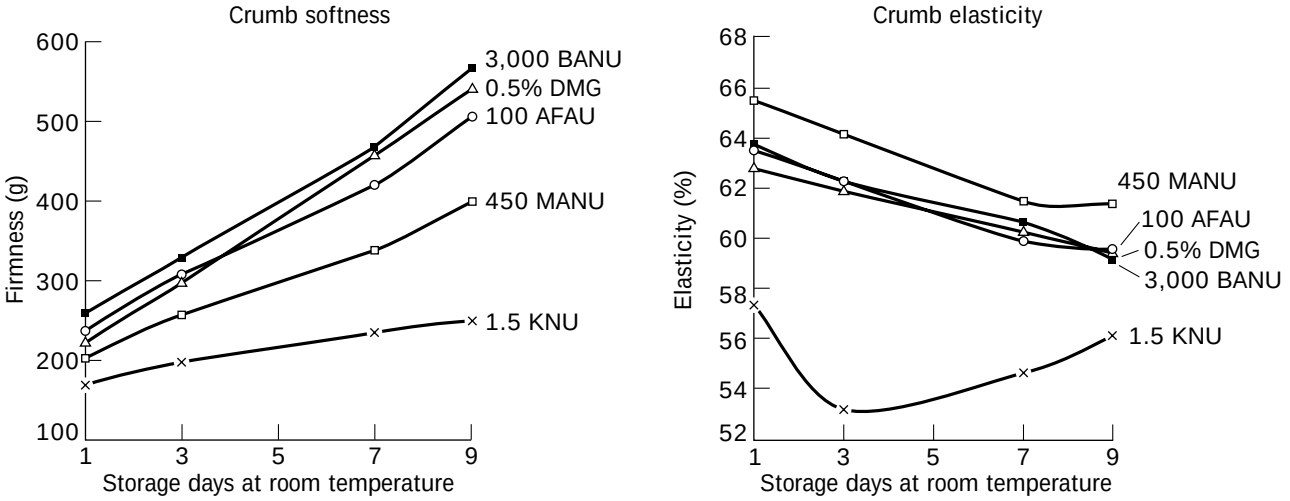


Figure 3. Effect of different amylases on crumb softness and elasticity. Sponge and dough procedure using American flour. Volume difference: maximum +/−3%, DMG, fine powdered distilled monoglycerides; MANU, units of maltogenic α-amylase/kg flour; KNU, units of bacterial α-amylase/kg flour; BANU, units of soy β-amylase/kg flour; AFAU, units of acid fungal α-amylase/kg flour. *Source*: Ref. 9.

was significantly lower than with the other samples; the difference increased after 4 days of storage and was even more pronounced after 9 days. This indicated a significantly slower rate of starch retrogradation. The fungal α-amylase had no effect on retarding starch retrogradation. The results indicate that an enzyme can reduce the crumb firmness by means of volume or crumb structure improvement, but it has no effect on retarding the retrogradation of starch during storage. Pullulanase had no effect on retarding starch retrogradation. This means that reducing the number of branches of amylopectin does not reduce the rate of bread staling and may even increase the staling rate, as reported earlier (10).

XYLANASES AND PENTOSANASES AS DOUGH CONDITIONERS

Xylanase and pentosanase, also being called hemicellulases, have long been used as dough-conditioning enzymes, especially in European-type bread. Pentosanases improve dough machinability and oven spring, resulting in a volume increase.

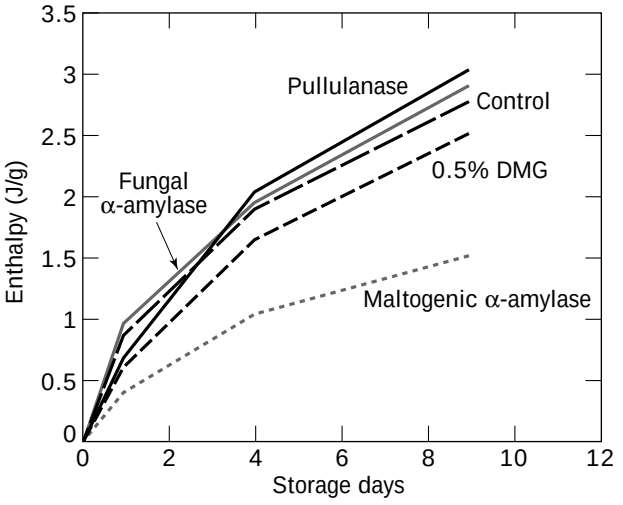


Figure 4. Effect of amylases on starch retrogradation compared with distilled monoglycerides measured by differential scanning colorimetry. *Source*: From Ref. 4.

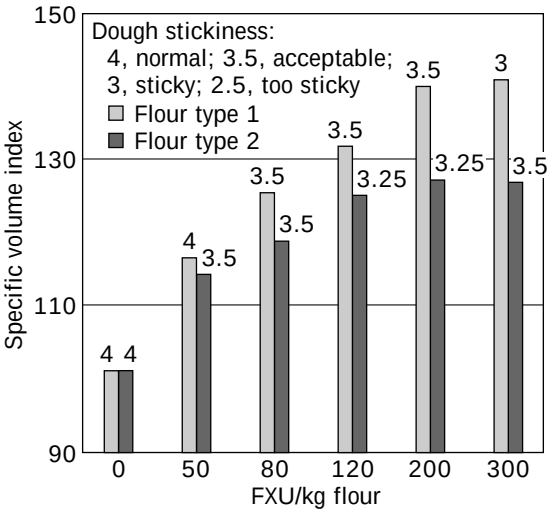


Figure 5. Dosage response of xylanase in different flours. *Source*: Novo Nordisk.

Through recombinant gene technology, some xylanases made from genetically modified organisms (GMO) are coming on the market (11). The benefit of using a xylanase instead of a traditional pentosanase preparation is that there are far fewer side activities in the xylanase product. Consequently, a lower dosage is needed to achieve the same effect with less risk of possible interference from side activities. The enzyme quality is also more consistent.

Pentosanases and xylanases can be overdosed due to overdegradation of wheat pentosans, thereby destroying the water-binding capability of the wheat pentosans. The result of overdosing is dough stickiness. A suitable dosage of these enzymes results in a desirable softening of dough, thereby improving the machinability. An optimum dosage is therefore defined as the dosage that gives the maximum improvement of the bread properties without causing dough stickiness. The optimum dosage of pentosanases or xylanases varies according to the flour, as illustrated in Figure 5. The optimum dosage for type 1 flour is 120–200 and for type 2 flour is 80 fungal xylanase units (FXU).

The true mechanism of xylanase in bread making has not been clearly elucidated, although a number of publications have attempted different approaches. The composition of pentosans varies depending on the flour type (12). The interaction between pentosans and gluten plays an important role that has not yet been elucidated. Most commercially available enzyme preparations used for studies consist of several enzyme activities including amylase, protease, and several hemicellulases. Works by Hamer (5,13) indicate that the use of pentosanases increased gluten coagulation in a diluted dough system. Si and Goddik (14) reported that a good baking xylanase increases the gluten strength measured by dynamic rheological methods on a Bohlin rheometer VOR. By measuring the gluten extracted from a dough system, the good baking xylanase increased storage modulus G′ and G″. At the same time the gluten became more elastic because the phase angle δ decreased, whereas the two other xylanases, which did not show good baking performance, had no significant effect on gluten strengthening.

Figures 6 and 7 show the effect of a commercially available baking xylanase on the rheological properties of gluten extracted from a wheat flour dough measured on Bohlin rheometer VOR (11). At the dosages that give good baking performance, 50–100 FXU/kg flour, the gluten was strengthened and exhibited a more elastic property. These improved rheological properties of the gluten explain the positive effect of the xylanase on bread making (i.e. increased oven spring, resulting in larger volume and improved dough stability). At an overdosed level of 400 FXU/kg flour, which provided a very sticky dough, the G′ is at the same level as the control and the phase angle shows a more viscous property of the gluten.

Different xylanase or pentosanase preparations have different effects on arabinoxylans in terms of their scission points and reaction products; therefore, they have different effects on bread making (15). Although many analytical methods are available for assaying pentosanases or xylanases using a wide range of substrates, it is still not possible to have an analytical method of which the enzyme units can correlate to the baking performance when comparing different types of xylanases or pentosanases. It has been reported (15) that no correlation to the baking performance was found when different enzymes from different sources were submitted to the same standardized assay and baking tests using purified substrate. One possible explanation is that once the wheat pentosans, both soluble and insoluble, were extracted from the flour, the mechanism of the interactions between wheat pentosans and gluten could no longer be assayed. Xylanase influences the gluten, as discussed earlier.

A possible mechanism of pentosanases mentioned by Hamer (5) is that the enzyme can offset the negative effects of insoluble pentosans present in the flour, because insoluble pentosans are regarded as having a negative effect on loaf volume and crumb structure. A study conducted by

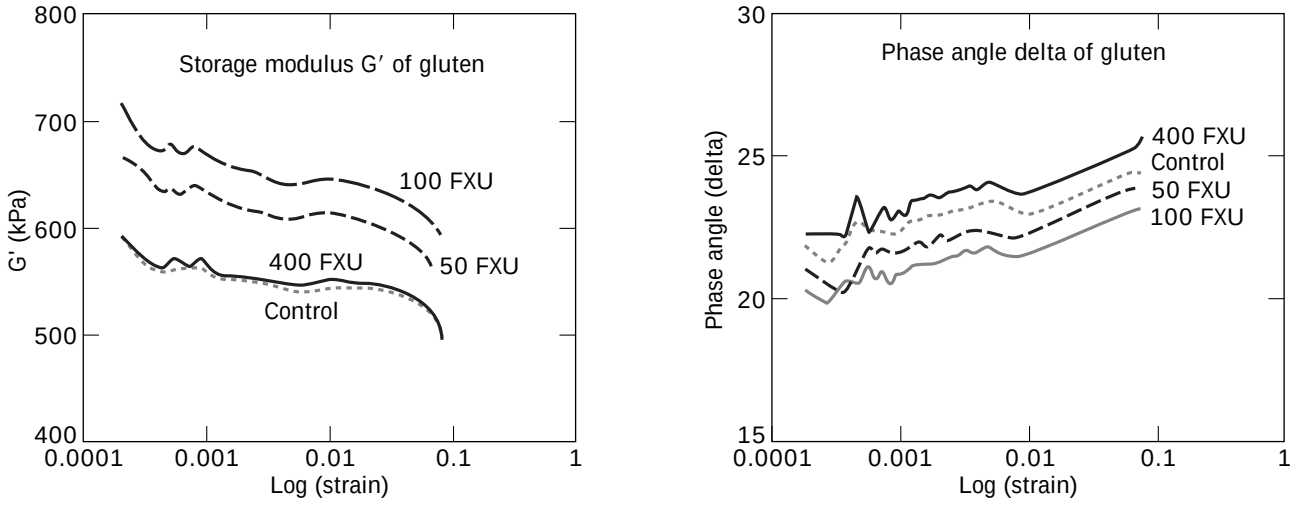


Figure 6. The effect of xylanase on the rheological properties of gluten.

Jakobsen and Si (15) using four different pure xylanases concluded that the best enzyme in terms of baking performance is the xylanase that has a certain level of activities toward both soluble and insoluble wheat arabinoxylans in a dough system. A xylanase having the major activity toward the insoluble arabinoxylans gave a dough too sticky to be accepted.

LIPASE FOR DOUGH CONDITIONING

The use of lipase for bread making was almost unknown just a few years ago, although it was reported as early as 1968 that the use of certain lipase preparations in a bread dough significantly retards the tendency of bread to become stale (16). Until recently, some 1,3-specific lipases were found to have a good dough-conditioning effect, in terms of improved dough rheological properties, increased dough stability upon overfermentation, increased oven spring resulting in a larger volume, and improved crumb structure in dough systems without added shortening (17,18).

Like most other enzymes, the effect of lipase also depends on the type of flour and the baking formulation. All three types of flour shown in Figure 7 are European baking flours with similar characteristics in terms of lipids, gluten content, and water absorption. For Meneba flour the optimum dosage was 2,000 LU/kg flour, and for Manitoba and Baguepi flours it was 1,000 LU/kg flour. An addition of 500–2,000 LU/kg flour results in a volume increase of 20–30%. At too high a dosage of lipase the dough becomes dry and stiff with a reduced volume (11).

Figure 8 shows the effect of the lipase in formulations with added fat or oil. Whereas lipase significantly improves bread quality in fat-free formulations or those with added oil, it has no beneficial effect when the formulation contains added hydrogenated shortening in terms of volume increase.

Because of its effect on volume increase and especially on the improved, more uniform crumb structure, the lipase

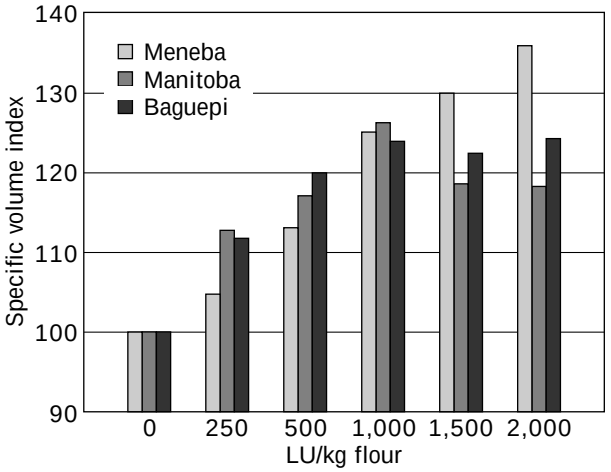


Figure 7. Dosage response of lipase in different flours.

improves crumb softness during storage. As discussed earlier, larger volume and more uniform crumb structure contribute to a softer bread. As illustrated in Figure 9, lipase reduces the crumb firmness compared with the control. It can provide a crumb softness as good as 3–6% shortening. However, it has no significant effect on improving crumb elasticity.

The mode of action of the lipase is still unknown. The postulation about producing monoglycerides in situ cannot explain the effects of the lipase demonstrated earlier. Monoglycerides are known to have an antistaling effect but have limited dough-conditioning effects compared with lipase. Furthermore, the contents of total lipids in most wheat flours is in the range of 1–1.5%. Regarding the limited hydrolysis degree (20) and limited quantity of triglycerides of saturated fatty acids in dough systems, an insufficient amount of monoglycerides of saturated fatty acids is produced to exhibit the effects described earlier. Besides, the monoglycerides produced by this 1,3-specific lipase

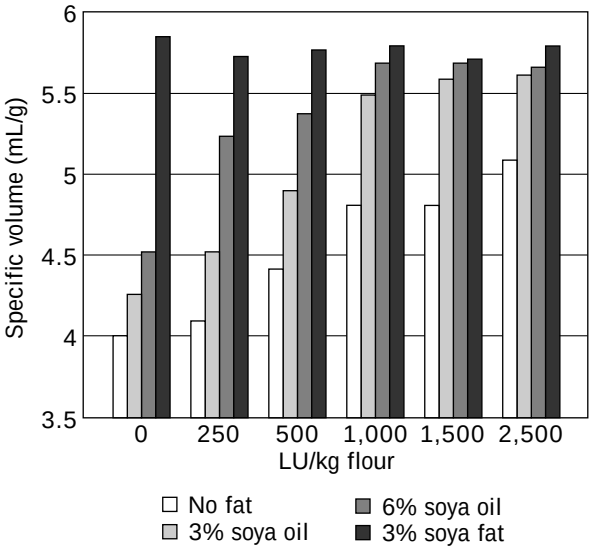


Figure 8. Effect of lipase in formulation with added lipids.

would be the second-position monoglycerides. It was reported (21) that first- and third-position monoglycerides are more able to form complexes with starch and thus having a retarding effect on starch staling.

Addition of lipase to a flour dough does not change the rheological properties of the dough measured by both Farinograph and Extensograph (9) with optimum water absorption. This is a good characteristic for bread making, because most bakers do not like major changes in the dough system. However, when we measured the rheological property of the gluten treated by the lipase using the dynamic rheological method, we found that the reason for the good dough-conditioning effect of this lipase is the increase of gluten strength, as shown in Figure 10. The gluten taken from lipase-treated wheat flour dough is significantly stronger, with an increased G′ and more elasticity, a lower δ than the control. An overdosing of lipase results in too strong of a gluten complex with a very high G′, which gives a dough that is too stiff and a smaller volume increase, as mentioned earlier. The gluten treated with lipase has a high G filt and increased relaxation time, indicating an increase of cross-links in the gluten networks measured by the dynamic stress relaxation method (20). It is not yet possible to elucidate the nature of these extra cross-links.

The knowledge of the chemistry of wheat lipids in bread making is still fragmentary. It was assumed that glycolipids, through hydrogen bonds and hydrophobic interactions, form a linkage between gliadin and glutenin (22,23). Bekes et al. (24) reported that there is a strong positive correlation between the lipid-mediated aggregates and loaf volume. However, overaggregates of the gluten may be responsible for the deterioration of bread-making quality as reported by Weegels and Hamer (25). A possible explanation of why this lipase has the positive effective is that it degrades some of the lipids, namely triglycerides, thereby reducing the overaggregation and forming a more stable gluten network.

A recent study conducted at Lund University (26) indicated that lipase increased the thermostability of the reversed hexagonal phase of the liquid–crystalline phase during heating up to 100 °C. This observation is suggested (27) as a possibility for one of the mechanisms of lipase in bread making.

Because lipases are gaining more and more attention in the baking industry, studies in the near future should give us a better understanding of the mode of action of lipase for bread making.

OXIDASES FOR BREAD MAKING: REPLACING CHEMICAL OXIDANTS

Oxidants, such as ascorbic acid, bromate, and azodicarbonamide (ADA), are widely used for bread making and have

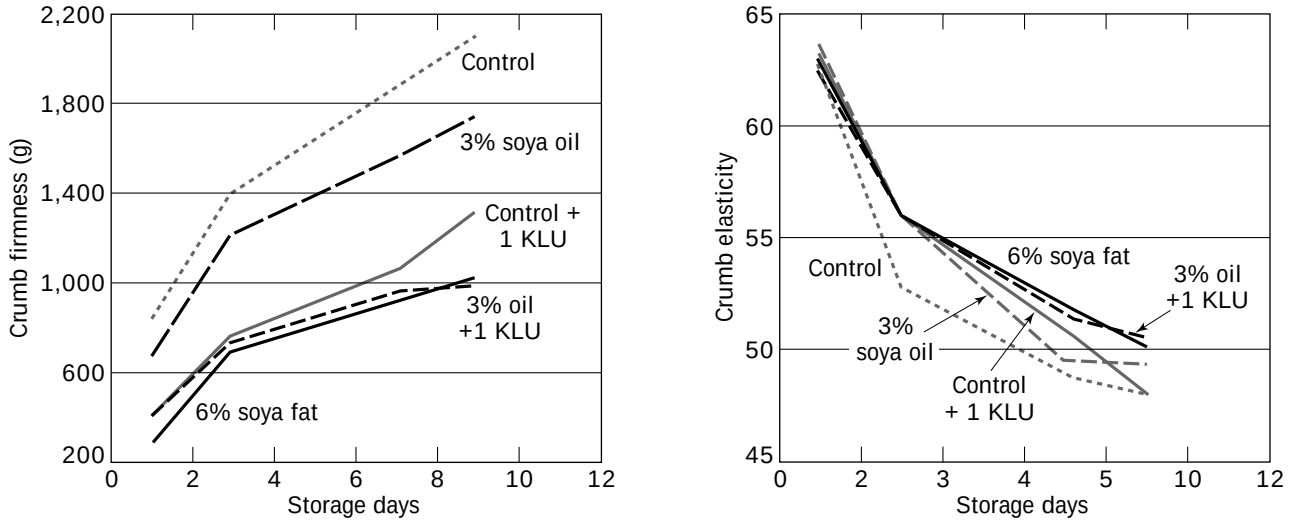


Figure 9. The effect of lipase on crumb softness compared with shortening. 1 KLU = 1,000 LU/kg flour of lipase. *Source*: From Ref. 19.

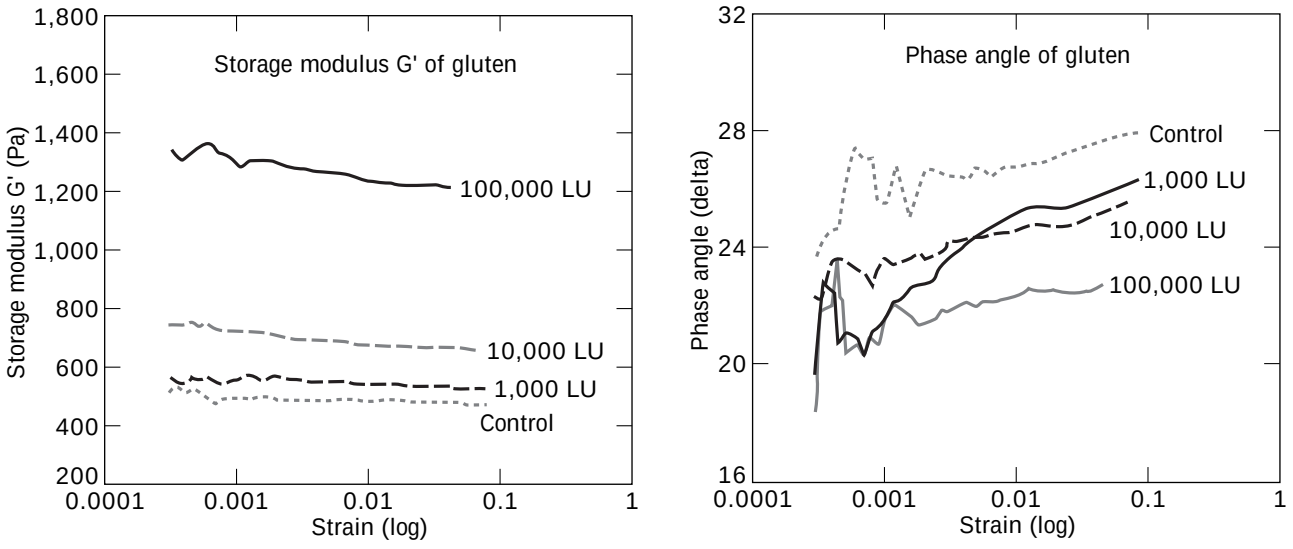


Figure 10. The effect of lipase on the rheological properties of gluten.

been well studied and reported on. However, the true mechanism of the oxidants in bread making has not been established, although a number of hypotheses were offered by numerous studies. It is generally known that bromate is a slow-acting oxidant and becomes active at high temperatures. In relation to baking, it has the maximum effect in the later stages of proofing and in the early stages of baking, whereas ascorbic acid and other oxidants are fast-acting oxidants and having the maximum effect during mixing and proofing.

Increasing demands by consumers for more natural products with fewer chemicals and especially the possible risks with bromate in food have created the need for bromate replacements. Therefore, oxidases are gaining more and more attention within the baking industry. Although glucose oxidase for bread making has been known since 1957 (28) and lipoxygenase, lysyl oxidase, sulfhydryl oxidase (29), peroxidase (30), laccase (31), and transglutaminase have been reported to have good oxidizing effects, the knowledge about oxidases for bread making is scarce and only glucose oxidase is currently available commercially.

Glucose oxidase has good oxidizing effects that result in a stronger dough. It can be used to replace oxidants such as bromate and ascorbic acid in some baking formulations and procedures. In other formulations it is an excellent dough strengthener along with ascorbic acid. Examples of applications of glucose oxidase for bread making are given elsewhere in this volume.

SYNERGISTIC EFFECTS OF ENZYMES

Enzymes for Dough Conditioning

Using enzyme combinations for bread making is not new. It is well known that hemicellulase or xylanase used in combination with fungal α-amylase has synergistic effects.

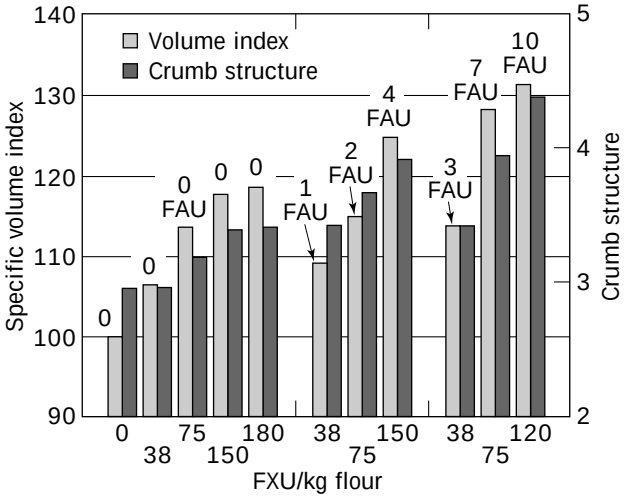


Figure 11. Synergistic effect of amylase and xylanase. FAU, fungal α-amylase units; FXU, fungal xylanase units. *Source*: Novo Nordisk, Pentopan Mono.

As shown in Figure 11, a high dosage of a pure xylanase may give some volume increase, but, the dough with this dosage of xylanase would be too sticky to be handled in practice. When the xylanase is combined with even a very small amount of fungal α-amylase, a lower dosage of xylanase with the α-amylase provides a larger volume increase and better overall score without the dough stickiness problem. The combination of xylanase and fungal amylase still has its limitations because of the dough stickiness easily caused by these enzymes.

Because of its effect on gluten strengthening, lipase improves dough stability against overfermentation. Figure 12 shows the combination of the enzymes for a French ba-

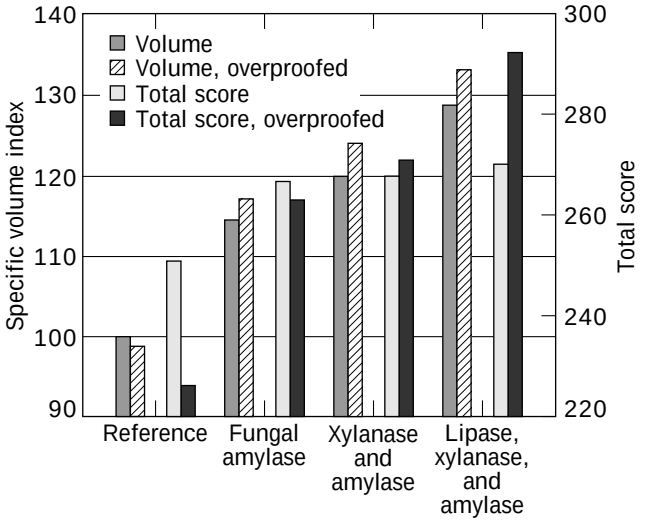


Figure 12. The synergistic effects of enzymes for the French baguette. *Source*: Novo Nordisk, Biotimes.

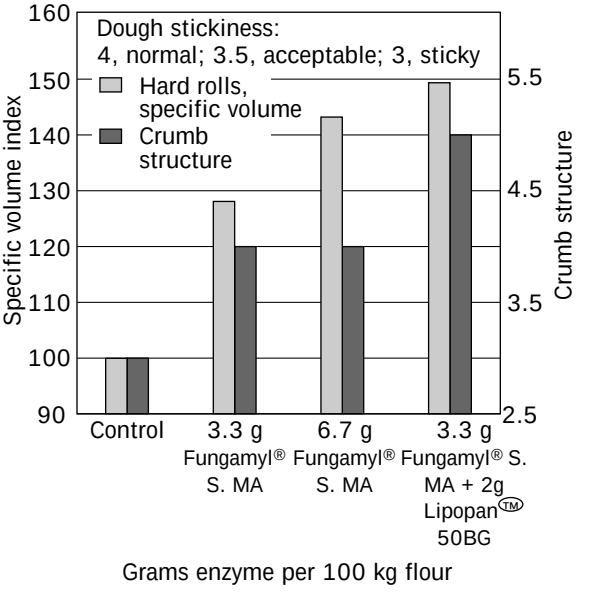


Figure 13. The combination of amylase or xylanase with lipase for hard rolls and pan bread. *Source*: Novo Nordisk Ferment.

guette (trials conducted at ENSMIC Baking Laboratory, Paris). Without any enzyme, the baguette had poor stability on overfermentation. With the addition of α-amylase alone or in combination with xylanase, the volume and overall scores of the bread improved. When α-amylase, xylanase, and lipase were used together, the overall scores were further improved. The dough exhibited especially good stability on overfermentation, which resulted in significantly better baguettes.

Because lipase does not make the dough sticky and significantly improves dough stability and crumb structure, the synergistic effects between xylanase or amylase and lipase provide many possibilities for improved bread quality. Figure 13 shows the results of hard rolls from Meneba flour. A dosage of 3.3 g Fungamyl® Super MA consisting of fungal α-amylase and xylanase gave satisfactory dough consistency but a limited volume increase. A doubling of this dosage gave a larger volume increase but a dough too sticky to be handled. When combining the low dosage of Fungamyl® Super MA with the lipase, the volume increase was significantly higher without any dough stickiness. Figure 14 shows that the combination of lipase with α-amylase or xylanase can provide a fine, silky, and uniform crumb structure to loaves based on a straight dough process.

Enzymes for Extending Shelf Life and Improving Crumb Softness and Elasticity

As mentioned earlier, the maltogenic α-amylase is a true antistaling enzyme that affects neither bread volume nor

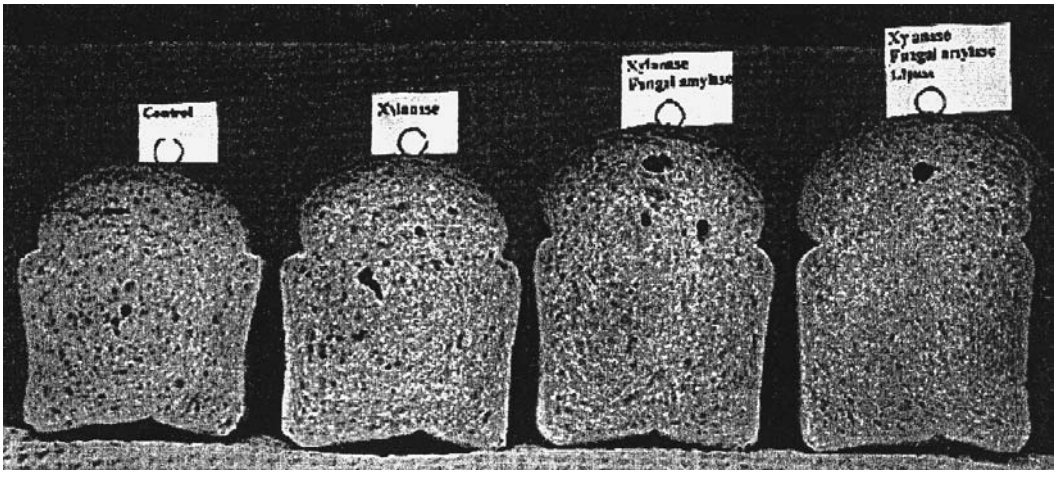


Figure 14. Synergistic effect of lipase in combination with α-amylase and xylanase.

crumb structure. Therefore, it is more feasible to use this enzyme in combination with enzymes such as fungal α-amylase, xylanase, and lipase to ensure the other bread quality parameters such as volume, dough stability, and crumb structure.

Most enzymes or other bread-improving ingredients including emulsifiers can improve the crumb softness due to the effects on bread volume and/or crumb structure, but they have a limited effect on crumb elasticity. The exception is the maltogenic α-amylase. Although the fungal α-amylase and xylanase or that in combination with thermostable bacteria α-amylase can reduce the crumb firmness during storage, as shown in Figure 15, the bread with the maltogenic α-amylase was significantly softer and had more elastic crumb than the other samples.

The preceding example also illustrates that the crumb firmness and elasticity may not necessarily correspond with each other. Many ingredients can reduce crumb firmness without affecting crumb elasticity, such as the bacterial α-amylase shown earlier. Our studies have also found that emulsifiers such as sodium stearoyl lactylate (SSL) may also have a negative effect on crumb elasticity.

Figure 16 shows different enzyme combinations compared with distilled monoglycerides in a sponge and dough system using a strong Canadian flour. The maltogenic α-amylase has the synergistic effect together with lipase, xylanase, and fungal α-amylase: 30 ppm of the maltogenic α-amylase together with the other enzymes gave a softer crumb during storage than using 45 or 75 ppm of the maltogenic α-amylase alone. The combination of the four enzymes (fungal α-amylase, xylanase, lipase, and maltogenic α-amylase) gave the softest crumb throughout the entire storage period of 9 days. The softness of the bread with these four enzymes at day 9 was at the same level as the loaf with 0.5% distilled monoglycerides at day 3. However, because only the maltogenic α-amylase has an effect on improving crumb elasticity, the more maltogenic α-amylase added, the more elastic is the bread crumb.

Enzymes for Dough Strengthening

Glucose oxidase has good oxidizing effects that result in stronger dough. It can be used to replace oxidants such as bromate and ADA in some baking formulations and procedures. In other formulations such as with ascorbic acid, it is an excellent dough strengthener (Fig. 17).

Use of glucose oxidase creates a dry and strong dough. A high dosage of fungal α-amylase gives the dough extensibility. The combination of these two enzymes can therefore achieve a synergistic effect. Figure 18 shows the combination of glucose oxidase and fungal α-amylase compared with ascorbic acid. When the two enzymes are used together with a smaller amount of ascorbic acid, the dough is not only very stable but also absorbs 1–2% more water, resulting in a greater volume increase and a crispier crust.

Glucose oxidase combined with fungal α-amylase can replace the bromate in some bread formulations. Figure 19 shows a type of South American bread, Marraqueta. This procedure requires a very stable dough because after two series of 50 min of proofing the dough is turned around before baking. A dough without strong dough stability will collapse when turned. The basic formulation contains both ascorbic acid and K-bromate. By adding glucose oxidase and fungal α-amylase instead of bromate, the final bread has a much improved appearance and a volume increase of approximately 40%.

In the previous sections each of the most commonly used enzymes for bread making has been reviewed, because of the rapid development of biochemistry and enzyme technology and the increasing interest in enzyme applications for the baking industry, more and improved enzyme products will be introduced in the near future. There will also be more reports and studies that will bring

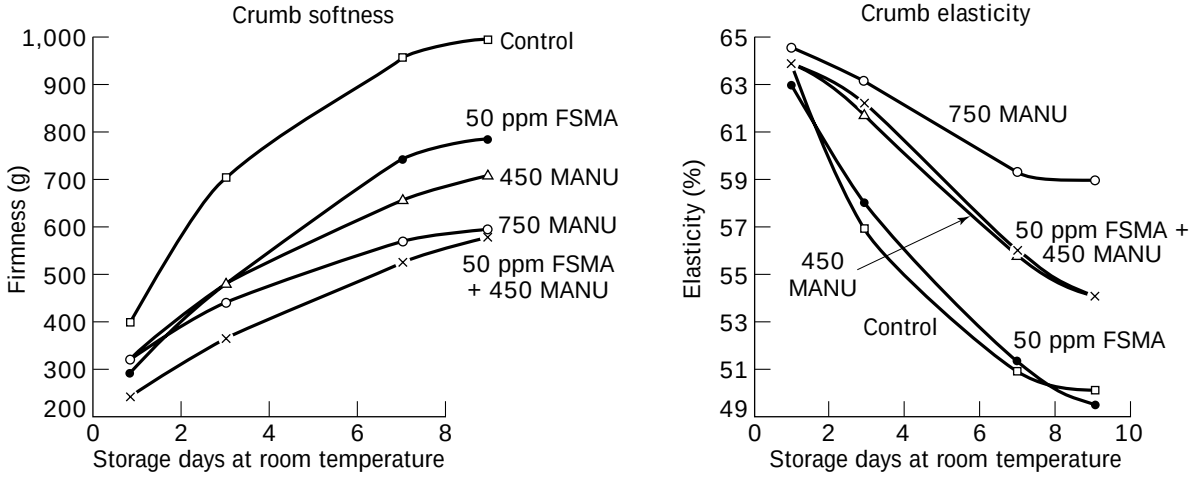


Figure 15. Antistaling effect of maltogenic α-amylase compared with thermostable bacteria amylase. 50 ppm FSMA (Novo Nordisk Fungamyl® Super MA) corresponds to 12.5 FAU and 112.5 FXU/kg flour; MANU, units of maltogenic α-amylase/kg flour.

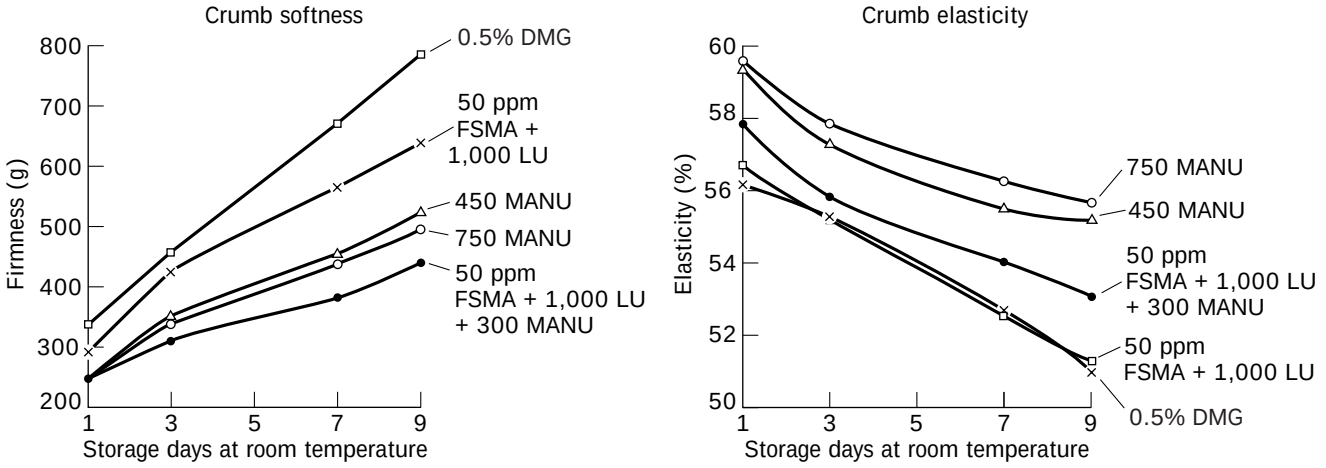


Figure 16. The synergistic effect of maltogenic α-amylase in combination with fungal α-amylase, xylanase, and lipase compared with distilled monoglycerides in American types of sponge and dough bread. FSMA (Novo Nordisk Fungamyl® Super MA) corresponds to 12.5 FAU and 112.5 FXU/kg flour; LU, units of lipase (Lipopan™)/kg flour; MANU, units of maltogenic α-amylase/kg flour; DMG, fine powder distilled monoglycerides. *Source*: Novo Nordisk Ferment.

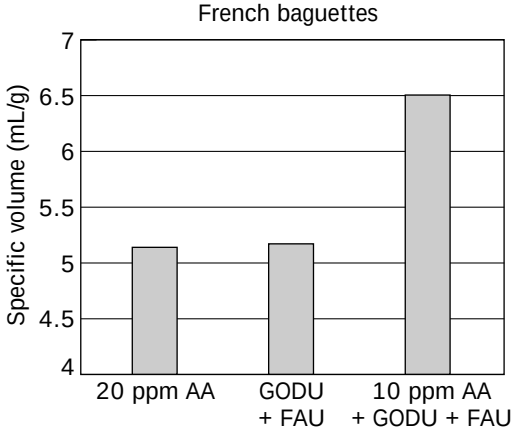


Figure 17. The combination of glucose oxidase and fungal amylase compared with ascorbic acid. The dosage of glucose oxidase is 300 GODU/kg flour plus 50 FAU. *Source*: Novo Nordisk Ferment.

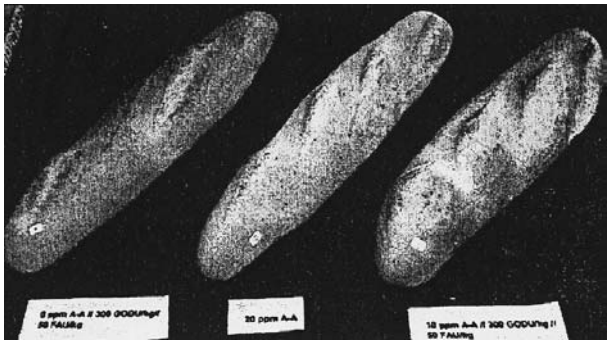


Figure 18. Glucose oxidase and fungal α-amylase for French baguettes.

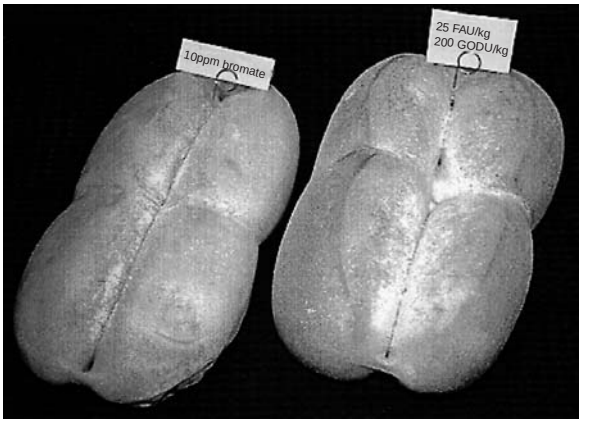


Figure 19. Glucose oxidase and fungal amylase as bromate replacer in Marraqueta. *Source*: Novo Nordisk.

us a better understanding of the mechanism of baking enzymes. In turn this understanding will lead to further improved products. An exciting future for baking enzymes lies ahead.

BIBLIOGRAPHY

1. H.W. van Dam and J.D.R. Hille, *Cereal Foods World* **37**, 245 (1992).
2. J.E. Kruger, D. Lineback, and C.E. Stauffer, in *Enzymes and Their Role in Cereal Technology*, American Association of Cereal Chemists, 1987, pp. 117–130.
3. H. Perten, *Cereal Chem.* **61**, (1984).
4. J.Q. Si and R. Simonsen, *Proc. Int. Symp. AACC/ICC/CCOA*, Beijing, November 16–19, 1994.
5. R.J. Hamer, in G. Tucker and L.F.J. Woods eds., *Enzymes in Food Process*, 1991.

6. R.E. Hebeda, et al. *Cereal Foods World* **35**, 453 (1990).
7. H. Outtrup and B.E. Norman, *Starch* **36**, 405 (1984).
8. European Pat. EP494233 (September 26, 1990), T. Olsen.
9. J.Q. Si, *Annual Meeting of the American Association of Cereal Chemists*, Baltimore, Md., September 15–19, 1996.
10. M.L. Martin and R.C. Hoseney, *Cereal Chem.* **68**, 503 (1991).
11. J.Q. Si, *Proc. 6th Mtg. Industrial Applications Enzymes*, Barcelona, November 29–30, 1995.
12. G. Cleemput, S.P. Roels, M. Van Oort, P.J. Grobet, and J.A. Delcour, *Cereal Chem.* **70**, 324 (1993).
13. R.J. Hamer, P.L. Weegels, J.P. Marseille, and M. Kelfkens, in Y. Pomeranz ed., *Wheat Is Unique*, American Association of Cereal Chemists, 1989, p. 467.
14. J.Q. Si and Goddik, *Annual Meeting of the American Association of Cereal Chemists*, Miami, Fla., October 1993.
15. T.S. Jakobsen and J.Q. Si, *Proc. Congr. Wheat Structure Biochemistry and Functions*, Reading, U.K.
16. U.S. Pat. 3,512,992 (April 2, 1968), A. Cooke and H. Johannson.
17. PCT Pat. WO 94/04035 (August 21, 1992), T. Olsen and J.Q. Si.
18. J.Q. Si, *Proc. Int. Symp. AACC/ICC/CCOA*, Beijing, November 16–19, 1994.
19. J.Q. Si, *Food Tech. Eur.* **3**, 60–64 (1996).
20. J.Q. Si and T.T. Hansen, *Proc. Int. Symp. AACC/ICC/CCOA*, Beijing, November 16–19, 1994.
21. N. Krog, *AOCS*, Hawaii.
22. R.C. Hoseney, et al. *Cereal Chem.* **47**, 135 (1970).
23. W. Bushuk, F. Bekes, G.J. MacMaster, and U. Zawistonwska, in A. Graveland and J.H. Moonen eds., *Gluten Protein: The Proceedings of the International Workshop on Gluten Protein*, 2nd ed., Wagenigen, 1984, p. 101.
24. F. Bekes, et al., *J. Cereal Sci.* **16**, 129 (1992).
25. P.L. Weegels and R.J. Hamer, *Cereal Food World* **37**, 379 (1992).
26. J. Borné, M.S. Thesis, Lund University, Sweden, 1996.
27. L. Dybdal, J.Q. Si, J. Born, and A. Eliasson, *Proc. First Eur. Symp. Enzymes Grain Processing*, Noordwijkerhout, The Netherlands, December 2–4, 1996.
28. U.S. Pat. 2,783,150 (February 26, 1957), H.G. Luther.
29. G. Matheis and J.R. Whitaker, *J. Food Biochem.* **11**, 309 (1987).
30. PCT Pat. WO 94/28729 (June 11, 1993), J.Q. Si.
31. PCT Pat. WO 94/28728 (June 13, 1994), J.Q. Si.

ADDITIONAL READING

V. Bolnedi, and C. Hoseney, Ph.D. Dissertation, Kansas State Univ., Manhattan, Kans., 1997.
J. Maat, et al. *Xylanases and Their Application in Bakery*, 1991.
Novo Nordisk, *Biotimes* (September 1996).

See also ENZYMES, PROTEIN HYDROLYSIS; FOOD PROCESS ENGINEERING; YEAST, BAKER'S.

ENZYME KINETICS. See KINETICS, ENZYMES.

ENZYMES, DETERGENT

M. S. SHOWELL
Procter & Gamble
Cincinnati, Ohio

KEY WORDS

Amylase
Cellulase
Detergents
Granulates
Lipase
Peroxidase
Protease
Stability

OUTLINE

INTRODUCTION

The use of enzymes as purposefully added ingredients to laundry products dates back to 1913 when Otto Röhm patented the use of crude pancreatic extracts in laundry presoak compositions to improve the removal of biological stains (1). It wasn't until the early 1960s, however, with the isolation and commercial production of alkaline active proteases, that enzymes became widely used in laundry products. By 1969, at least 50% of European detergents contained a protease (2). Within 6 years of the introduction of alkaline proteases in detergents, 80% of all European laundry detergents contained enzymes (3). Today, the penetration of enzymes into laundry detergents across the globe is quite high. Most premium and lower-tier laundry products contain at least one enzyme, usually a protease, and many are formulated with up to four different enzyme activities—protease, amylase, lipase, and cellulase. Enzymes in laundry and cleaning products account for about 40% of the more than 1 billion dollar industrial enzyme

market, with detergent proteases accounting for about 50% of the volume (4). In the United States, the detergent sector is the largest market for enzymes, accounting for about 20% of total enzyme demand and more than 25% of industrial enzyme sales (5). The move to compact detergents, beginning in 1987, helped raise the demand for detergent enzymes, and global enzyme sales doubled between 1989 and 1994.

Four factors are responsible for the increased reliance of detergent manufacturers on enzyme technology. First, they provide consumer-recognizable cleaning benefits. Historically, proteases and amylase have been used to improve the removal of proteinaceous and starch-based soils, respectively. More recently, detergent-compatible lipases have been introduced to help clean greasy and oily stains arising from foods and body soils. Second, some enzymes enable the detergent formulator to add completely new performance benefits. Cellulases are the best and most recent example where the development of cost-effective, alkaline-active cellulases has enabled the delivery of consumer-valued benefits such as color maintenance and fabric restoration. Third, enzymes are highly weight efficient. The ability to act catalytically imparts a space efficiency to enzymes that is highly desirable to the formulator of compact detergents. The fourth factor contributing to the increased use of enzymes in detergents is that in recent years the performance/cost ratio for enzymes has improved significantly as a result of the availability of more efficient enzymes and the industry trend toward reduced pricing.

Current market trends and consumer needs are influencing the development of enzymes for detergent applications, with the emphasis on enzymes that have improved performance/cost ratios, increased cold water activity, and improved compatibility to other detergent ingredients. In addition, enzyme suppliers and detergent manufacturers are actively pursuing the development of new enzyme activities that address the consumer-expressed need for improved cleaning, fabric care, and antimicrobial benefits. For applications beyond laundry and cleaning products, enzymes with low allergenicity are being developed to minimize worker and consumer safety concerns.

THE DETERGENT ENZYME MARKET

As noted in the "Introduction", enzymes have long been of interest to the detergent industry for their ability to aid in the removal of soils and stains as well as to deliver unique benefits that cannot otherwise be obtained with conventional detergent technologies. As a result, the detergent enzyme market has grown nearly 10-fold during the past 20 years. In 1997, sales of detergent enzymes totaled about $550 million. U.S. demand accounted for about $160 million and is expected to grow to $260 million by 2001 (6). Through the 1980s and early 1990s the market was split between five major suppliers: Novo-Nordisk A/S of Denmark with the dominant market share (>55%); Gist-Brocades N.V. in The Netherlands; Genencor International in the United States; Solvay in Belgium; and; Showa-Denko in Japan. These suppliers marketed a full range enzymes for liquid and powder detergents. Beginning in 1995, however, there was considerable rationalization in the detergent enzyme market owing to the relatively high cost of manufacturing coupled with increased pressure from detergent manufacturers to drive down raw material costs. Genencor International purchased the detergent enzyme businesses of Gist-Brocades and Solvay, and Novo-Nordisk acquired Showa-Denko's detergent enzyme business. Today, Novo and Genencor are the main suppliers of detergent enzymes, supplying more than 95% of the global market. Table 1 provides examples of the types of enzymes available from these two suppliers.

CURRENT DETERGENT ENZYMES

The majority of enzymes used in detergents today are hydrolases; the catalytic event involves addition of water across a chemical bond. In addition, most detergent enzymes exhibit *endo* specificity, that is, they preferentially attack internal bonds as opposed to terminal linkages.

Before the conversion of the detergent market to compact product forms, the use of enzymes in detergents was limited primarily to proteases. The market conversion to compacts, which require high levels of cleaning performance from low product dosages, fueled the interest in formulating new enzyme activities that can provide a more robust cleaning profile. This, coupled with recent advances in genetic and protein engineering, has led to new classes of enzymes being available for detergents, including amylases, lipases, cellulases, and peroxidases (Table 1). Each of these general enzyme classes is discussed in detail in the following sections.

Proteases

Most powder and liquid laundry detergents on the market today, both low density and compact, contain a protease (7). Protease enzymes enhance the cleaning of protein-based soils, such as grass and blood, by catalyzing the breakdown of the constituent proteins in these soils through hydrolysis of the amide bonds between individual amino acids. Proteases serve a multifunctional role in the overall cleaning process. They boost the efficacy of surfactant and dispersant systems by degrading proteinaceous soils into smaller, more dispersible fragments, thereby directly facilitating the soil removal process. In addition, they decrease the redeposition of proteins, especially hydrophobic proteins such as those found in blood, resulting in improved overall whiteness (8).

Figure 1 shows the dose-response curves for removal of grass stain with typical detergent proteases in standard North American liquid and powder detergent formulations. The plateau observed at higher enzyme concentrations is typical of most detergency evaluations of enzymes. Typically, detergent formulators dose enzymes at the breaking point of such curves in order to obtain optimum cleaning performance at lowest possible cost.

Figure 2 further illustrates the cleaning benefits of protease. Figure 2**a** shows an electron micrograph of blood-stained fibers from a polyester swatch washed under North American conditions in a liquid detergent without protease. The hemoglobin in the blood stain has been imaged with a colloidal-gold-labeled antibody marker to help vi-

Table 1. Detergent Enzyme Types and Suppliers: 1997

	Protease			Amylase			Cellulase		
Supplier	Alkaline	High alkaline	Bleach stable	Conventional	Bleach stable	Lipase	Multi-component	Mono-component	Peroxidase
Genencor	Optimase	Opticlean Purafect Properase	Purafect OxP	Maxamyl Rapidase	Purafect OxAm Purafect HpAm	Lumafast Lipomax		Puradax	
Novo	Alcalase	Savinase Esperase Kannase	Everlase	Termamyl BAN	Duramyl	Lipolase Lipolase Ultra LipoPrime	Celluzyme	Carezyme Endolase	Guardzyme

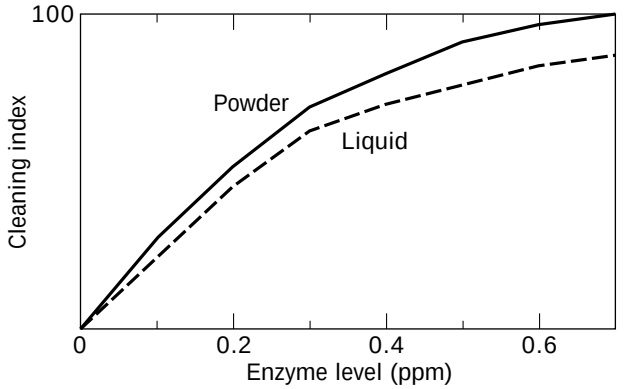
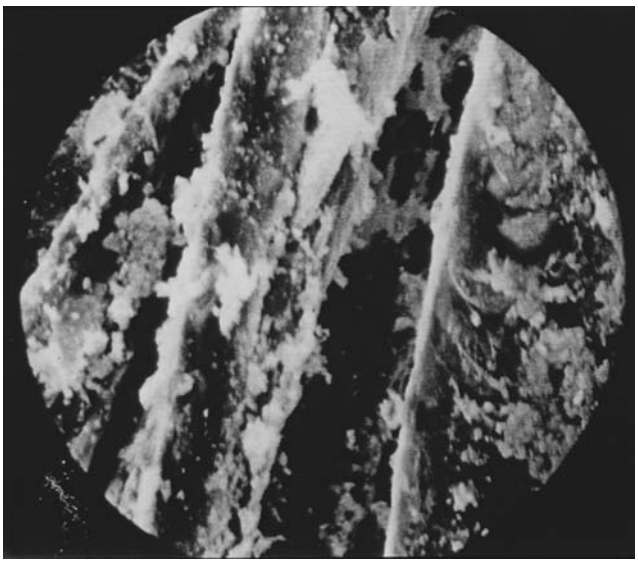


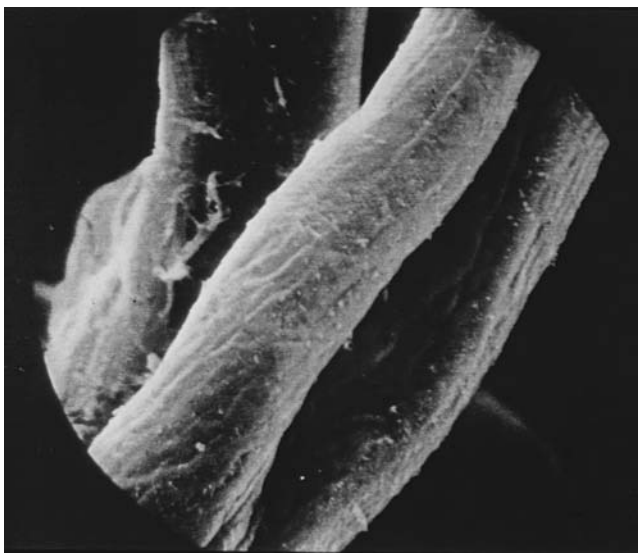
Figure 1. Protease dose response on grass. Detergent tests run at 20 °C with 1 mM mixed Ca/Mg hardness and manufacturer's recommended detergent dosage. Removal of grass on cotton based on visual grades from expert graders.

sualize remaining protein. There is a considerable amount of residual blood protein left on the fibers after washing. Figure 2**b** shows similarly stained and treated fibers that have been washed under the same conditions, in the same liquid detergent, but with 0.5 ppm protease added. There is a marked reduction in the amount of residual blood protein consistent with the visual observation of enhanced stain removal.

As shown in Table 1, a number of different commercial detergent proteases are currently being marketed. However, they are all closely related, nonspecific endopeptidases, with molecular weights around 27 kDa and highly homologous primary amino acid sequences (Fig. 3). Current commercial detergent proteases are derived from various species of *Bacillus subtilis* and are characterized by a common catalytic triad consisting of serine, aspartic acid, and histidine. The serine hydroxyl group functions as a nucleophile, attacking the peptide amide bond to form a tetrahedral intermediate that undergoes hydrogen trans-



(a)



(b)

Figure 2. (**a**) Transmission electron micrograph of blood-stained fibers after washing with a nil-protease liquid detergent. The residual hemoglobin has been marked using colloidal-gold-labeled antibodies. (**b**) Transmission electron micrograph of blood-stained fibers after washing with a protease-containing liquid detergent. The residual hemoglobin has been marked using colloidal-gold-labeled antibodies.

Figure 3. Sequence alignments of common detergent proteases: Based on BPN′ numbering. 1 = protease from *Bacillus lentus* (Savinase); 2 = protease from *Bacillus amyloliquefaciens* (BPN′); 3 = protease from *Bacillus licheniformis* (Alcalase).

fer facilitated by the aspartyl and histidine functional groups acting as general base catalysts. The result is addition of water across the amide bond, generating two peptide fragments (9). These so-called *serine* proteases (EC 3.4.21.62) are particularly well suited for laundry applications owing to their alkaline pH optima (Table 2), stability to detergents, and broad substrate specificity.

Table 2. pH Profiles of Common Detergent Enzymes

Enzyme	Source (species)	pH Optimum	Range (>80% maximal activity)
Alcalase®	*Bacillus licheniformis*	8.5	8.0–10.0
BPN′	*Bacillus amyloliquefaciens*	8.5	8.0–9.5
Savinase®	*Bacillus lentus*	10.0	8.0–11.0
Esperase®	*Bacillus alcalophilus*	10.0	9.0–12.0

Source: Internal data from Proctor & Gamble; Novo-Nordisk A/S, Denmark.

Transmission electron microscopy-immunocytochemical (TEM-ICC) analysis of residual soils remaining on fabrics after washing provides more detailed clues on the mechanisms of detergent protease action. Table 3 shows the frequency and mean particle sizes (MPS) of two types of protein—chlorophyll-binding protein (CBP) and β-lipoprotein (BLP)—and four types of carbohydrate—glucose/mannose, glucosamine, fucose, and galactose/galactosamine—remaining on residual grass stains on cotton and polycotton fabrics after washing in a liquid detergent with and without protease. The general trend is for the protease to reduce the frequency but not the size of the stain components. For example, with addition of protease, the frequency of occurrence of CBP on cotton drops from 2 of every 5 fibers to 1 in 10, and on polycotton, it drops from 4 markers in every 5 fibers to 1 in 6, with little or no change in particle size. Similar effects are observed for BLP and the carbohydrates.

Although the subtilisin proteases are stable to most of the surfactants used in typical laundry detergents, they

Table 3. TEM-ICC Data for Grass Removal

Component	Fabric	Without protease		With protease	
		Frequency[a]	MPS (μm)	Frequency[a]	MPS (μm)
Chlorophyll binding protein	Cotton	2/5	0.5	1/10	0.3
	Polycotton	4/5	0.3	1/6	0.3
β-Lipoprotein	Cotton	1/2	0.2	1/5	0.3
	Polycotton	3/4	0.3	1/3	0.3
Glucose/mannose	Cotton	1/7	0.2	1/10	0.2
	Polycotton	1/4	0.2	1/15	0.2
Glucosamine	Cotton	1/7	0.5	1/10	0.2
	Polycotton	1/5	0.4	1/10	0.3
Fucose	Cotton	1/25	0.2	1/30	0.3
	Polycotton	1/20	0.3	0/10	—
Galactose/galactosamine	Cotton	1/10	0.5	1/8	0.3
	Polycotton	1/10	0.3	0/10	—

[a]Frequency is determined from the number of fibers displaying the market out of all those examined (typically 100 fibers per sample).

are susceptible to oxygen bleaches and calcium sequesterants. Oxidative attack by peroxide or peracids on the methionine residue adjacent to the catalytic serine results in a nearly 90% loss of enzymatic activity (10). Replacement of the methionine with an oxidatively stable amino acid such as alanine significantly improves the stability of the enzyme toward oxygen bleach (11). In addition, the subtilisin proteases strongly bind at least one calcium ion, which contributes to maintaining the overall three-dimensional configuration of the enzyme. The calcium sequestering agents used in many laundry products to control water hardness can remove this calcium, resulting in decreased thermal and autolytic stability. Introduction of negatively charged residues near the calcium binding site can increase the binding affinity of the enzyme for the calcium, resulting in improved stability toward calcium sequesterants (12).

Because proteases degrade proteins, there is some risk associated with washing garments constructed from natural protein fibers such as wool and silk in protease-containing detergents. Proteolytic attack against such fibers leads to increased wear and reduced tensile strength. The degree of damage is dictated by the type and level of protease as well as the wash process. Accordingly, many detergents advise against use on wool and silk garments.

In powder detergents, proteases are added in an encapsulated or granulated form (see "Powders"), which protects them from other detergent ingredients and eliminates the potential for them to degrade each other (autolysis) or other enzymes (proteolysis) in the formulation. However, in aqueous-based liquid detergents, autolysis and proteolysis are key problems requiring the formulation of reversible protease inhibitors that can shut down the activity of the enzyme in product, but are diluted away from the enzyme when the detergent is added to the wash. Boric acid and polyols, such as propylene glycol, are commonly used to inhibit protease activity in liquid detergents (3).

Detergent proteases can be generally classified into two main groups, alkaline and high alkaline. Alkaline proteases, such as those derived from *Bacillus licheniformis*, show optimum activity around pH 8, making them ideal for liquid laundry products, which typically have wash pH's between 7.0 and 8.5. Examples include Alcalase® from Novo-Nordisk and Optimase® from Genencor International.

Several high alkaline proteases are available and are typically characterized by a pH-activity optimum around 10.0. These enzymes find widespread use in powder laundry detergents and automatic dishwashing formulations. Examples include Savinase® from Novo-Nordisk and Purafect® from Genencor.

A number of analytical methods are available for assaying proteolytic activity in product and in the wash. Most methods use colormetric detection of the hydrolysis of a specific reagent. Typical reagents include dimethylcasein, azocasein, azocoll, or hemoglobin. By far the most common assay method is based on hydrolysis of small synthetic peptides functionalized with the *p*-nitroanilide chromophore. For most detergent proteases, the synthetic tetrapeptide Suc-AAPF-pNA, or succinylalanylalanylprolyl(phenylalanine)-*p*-nitroanilide (PNA), is used. Hydrolysis of the anilide bond liberates the *p*-nitroaniline group, resulting in the appearance of a yellow color that can be followed spectrophotometrically. This assay is particularly useful for establishing pH and temperature profiles of proteases and following their stability through the wash or during in-product storage (13). However, in general, PNA activity does not correlate with wash performance. Activity against larger substrates, such as casein, which more closely resemble the soil proteins found in common laundry stains, shows a better correlation with protease wash performance.

Amylases

Amylases facilitate the removal of starch-based soils, particularly food soils, by catalyzing the hydrolysis of the glycosidic linkages in starch polymers. This general class of enzymes encompasses several enzymatic activities responsible for starch hydrolysis, including α-amylase, β-amylase, pullulanase, isoamylase, and glucoamylase. The α-amylases (1,4-α-D-glucan-glucanohydrolase, EC 3.2.1.1) are most commonly used in detergents. They catalyze enhohydrolysis of 1,4-α-D-glucosidic bonds in high molecular weight starches of the type commonly found in food soils. As such, amylases find utility in both laundry and dish

detergent formulations. Figure 4 shows the dose-dependent response of starch soil removal using a commercial α-amylase in a typical automatic dishwashing detergent formulation.

Current commercial detergent α-amylases are derived from various species of *Bacillus subtilis*. The α-amylase from *B. amyloliquiefacians*, is particularly well suited for detergent applications because of its thermal stability, broad pH activity, and resistance to proteolytic attack. Examples include Maxamyl® from Genencor and Termamyl® from Novo (14).

The detergent α-amylases in use today have highly homologous primary amino acid sequences and a tertiary structure comprised of three domains, A, B, and C (Fig. 5). The active site is typically located in a cleft between domains A and B and is usually comprised of acidic amino acids such as aspartic and glutamic acids. Like the subtilisin proteases, α-amylases contain an essential calcium ion that is important in maintaining the tertiary structure. Accordingly, the α-amylases show some degree of sensitivity to calcium sequesterants, with solution stability severely compromised in well built (low calcium) environments. In addition, most wild-type α-amylases are sensitive to oxygen bleach. Recently, protein engineering has been used to improve bleach stability of amylases (15). Both Novo and Genencor offer bleach-stable versions of their α-amylase (Table 1).

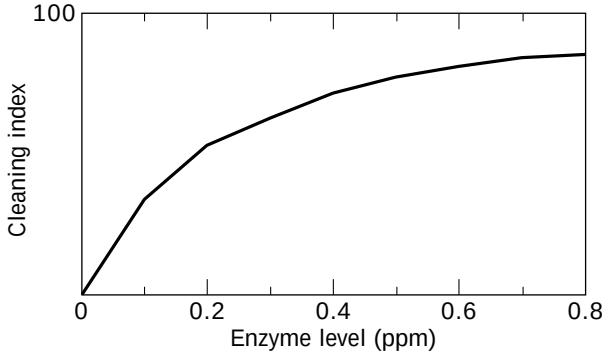


Figure 4. Dose response of α-amylase on starch soil in automatic dishwashing application. Cleaning tests run at 55 °C with 1 mM mixed Ca/Mg hardness and manufacturer's recommended detergent dosage. Starch soil removal measured gravimetrically.

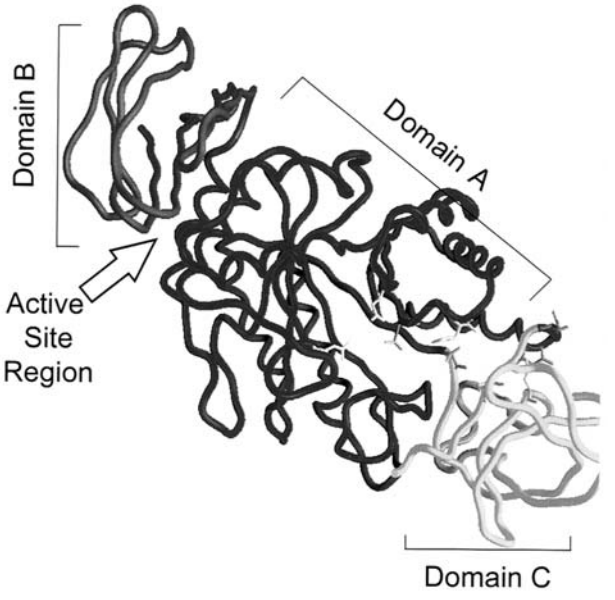


Figure 5. Three-dimensional structure of α-amylase.

The use of amylase and protease combinations for improved stain removal and general cleaning was fairly well established in the mid-1980s and practiced in a number of European and U.S. brands. The use of these two enzymes in compact detergents is now fairly standard, especially in Europe. Figure 6 shows the benefit of combining an α-amylase with a serine protease for removal of food stains from cotton in a typical North American powder laundry detergent.

As with proteases, there are a number of methods available for analyzing amylase activity (16). Probably the simplest, and most common, involves monitoring the color change during hydrolysis of iodine-stained starch. Iodine staining of a standard starch substrate results in a blue color that changes to reddish brown upon hydrolysis with α-amylase. Additional methods include hydrolysis of synthetic substrates like *p*-nitrophenylmaltoheptoaside (pNP-G7). Amylolytic attack on pNP-G7 produces pNP-G3 and pNP-G4. Reacting the pNP-G3 with an excess of α-glucosidase generates glucose and the characteristic yellow color of *p*-nitrophenyl. Alternative methods involve detection of the reducing sugars formed during starch hydrolysis using 3,5-dinitrosalicylate (17), alkaline ferricyanide (18), or alkaline copper (19).

Lipases

Lipases are fairly recent entries into the detergent enzyme market. Their use in detergents is illustrated in several patents from the 1970s (20). However, the lack of a viable commercial source of alkaline-active lipases with stability

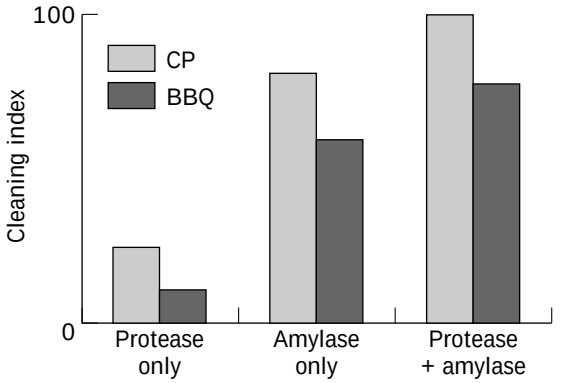


Figure 6. Cleaning benefit for protease + amylase on chocolate pudding and barbecue stains on cotton. Detergent tests run at 20 °C with 1 mM mixed Ca/Mg hardness and manufacturer's recommended detergent dosage. Enzyme dosages were 0.1 ppm each. Removal of chocolate pudding (CP) and barbecue (BBQ) stains from cotton measured by image analysis.

toward anionic surfactants essentially prohibited practical application for almost 10 years until the late 1980s when Novo Industries announced the commercial sale of the first detergent-compatable lipase, Lipolase® (21), obtained by cloning the lipase gene from *Humicola lanuginosa* into a high-expression fungal host. Lipolase® first appeared in compact powders in Japan and has since found fairly broad application in the global detergent market.

Lipases (triacylglycerol acylhydrolases, EC 3.1.1.3) degrade triglycerides by catalyzing the hydrolysis of the ester linkage between the glycerol moeity and the constituent fatty acids, generating glycerol, free fatty acids, mono- and diglycerides. Triglycerides are common components of some of the most-difficult-to-remove soils and stains, including greasy foods, cosmetics, and body soils. Owing to their generally hydrophobic nature, triglycerides are difficult to remove from fabric and may be present even after several wash cycles. As the residual triglycerides age they can oxidize, resulting in the generation of a rancid malodor as well as an overall yellowing or discoloration of the fabric. Lipolytic attack increases the water solubility of the original triglyceride and decreases the melting point of the fat, allowing for increased soil removal during the washing process.

Lipases are most active at the water–substrate interface. This phenomenon is referred to as interfacial activation (22). Lipase activity does not follow classical enzyme kinetics, which generally are only valid for reactions that occur in a homogeneous phase. Rather, lipase activity appears to follow a two-step process (23). In the first step, lipase physically binds to the surface. This is believed to involve activation of the enzyme via a structural change that exposes the hydrophobic binding domain to the substrate. The second step involves formation of the enzyme–substrate complex, leading to hydrolysis of the ester bond. Because physical adsorption to the interface is key to lipase activity, anything that interferes with this process, such as surfactants, can dramatically affect substrate hydrolysis. Exclusion from the soil–water interface is believed to be one of the key mechanisms involved in surfactant inhibition of lipase activity. Given the tendency of surfactants to accumulate at an oil and water interface, lipase can be denied access to the substrate by electrostatic repulsion and steric hindrance of the aggregated surfactant (24). Accordingly, when formulating lipase into detergents it is important to select a surfactant system that minimizes this effect. Generally, nonionic surfactants are less inhibiting than anionics.

Most commercial detergent lipases require more than one wash cycle to show a benefit. Figure 7 shows this multicycle performance effect for Lipolase Ultra®. It is generally thought this traces to deposition of the lipase onto the hydrophobic stain during the wash such that, during the drying cycle, when the water content is decreased, the enzyme is activated and can hydrolyze triglycerides in the stain (25), facilitating removal in the second wash cycle. For Lipolase®, maximum enzymatic activity against a standard lipid soil such as olive oil is obtained when the moisture content of the fabric is 20 to 30% by weight (26), suggesting that significant decomposition of the oily soil occurs during the drying process; hence the need for a second wash cycle to remove the hydrolytic fragments.

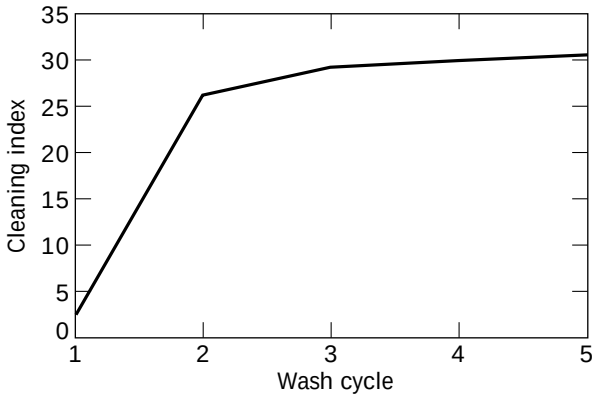


Figure 7. Multicycle stain removal of hamburger grease from cotton fabric with Lipolase® under typical North American wash conditions. Cleaning tests run at 30 °C with 1 mM mixed Ca/Mg hardness and manufacturer's recommended dosage of powder detergent. Lipolase Ultra ® dosed at 3000 LU/I. Removal of hamburger grease from cotton was measured by image analysis.

The most commonly used method to measure lipase activity is a titrimetric assay in which the fatty acid liberated during hydrolysis of an emulsified substrate, such as Triolein, is titrated with a standardized base solution. The milliequivalents of base required to maintain the solution pH at a given value is directly related to the amount of fatty acid generated. Alternate methods rely on colormetric detection of the hydrolytic products generated from synthetic substrates such as *p*-nitrophenylbutyrate.

Cellulases

Cellulases are unique among the hydrolase enzymes used in detergency. Rather than attacking substrates found in soils and stains, cellulases hydrolyze the cellulose in cotton and cotton blends to deliver cleaning benefits such as whitening and dingy cleanup as well as fabric care benefits such as color maintenance and fabric restoration. Like lipases, cellulases have only recently been introduced broadly into detergents. The first use of cellulases in detergents was described by Murata in 1981 (27). Commercial application began in 1987 with the introduction of the alkaline cellulase from *Bacillus* sp. KSM-635 in Kao's compact detergent Attack® (28). Since 1991, a number of European and North American detergents have incorporated cellulases. In 1993, Procter & Gamble introduced Novo-Nordisk's Carezyme®, an enzyme reported to help keep cotton and cotton blends looking newer longer by removing the pills and fuzz that build up on cotton garments over time, making them look worn and faded (29).

Cellulases deliver a number of benefits including cleaning (via removal of particulate and oily soils), fabric softeness, color clarification, and depilling. The cleaning benefits are believed to derive from the removal of surface fibrils that serve as anchor points for soil. Softness benefits are believed to result from the ability of cellulases to loosen surface fibrils, preventing them from intertwining and forming a rigid network with negative hand feel. Color

clarification and depilling derive from the same mechanism. Repeated washing and wearing of cotton garments generates fuzz and pill buildup as surface fibrils loosen and become entwined. The fuzz and pills scatter light, making garments look worn and faded. Cellulases weaken the attachment points of the fuzz and pills, which subsequently break off in later wash cycles. The result is brighter, more colorful garments. Figure 8 shows this benefit on cotton socks after 30 wash cycles in a North American powder detergent formulation with and without Carezyme®.

Cellulases (1,4-β-D-glucan-4-glucanohydrolases, EC 3.2.1.4) catalyze the endohydrolysis of 1,4-β-D-glucosidic linkages in cellulose. Cellulose is a linear, unbranched polymer of anhydroglucose units linked through β-1,4 bonds. Each polymer contains between 10,000 and 15,000 glucose units. Up to 60 polymers are arranged in parallel bundles, called elementary fibrils, and stabilized through extensive hydrogen bond and van der Waals interactions, which promotes crystallinity of the bundle. The elementary fibrils can further aggregate to form microfibrils that, in turn, aggregate into macrofibrils of up to 500 nm in diameter. Cotton fibers are formed by the accumulation and aggregation of macrofibrils (30). Disruption of the hydrogen bonding network within these structures, either through chemical or mechanical means, leads to the formation of amorphous regions. In new cotton garments, the ratio of crystalline to amorphous regions is about 70:30. In regenerated cellulose fabrics such as rayon, the structure is much more disordered, leading to a shift in the crystalline amorphous ratio to about 30:70. The amorphous regions are the site of cellulytic attack owing to the greater accessibility and flexibility of the substrate. As a result, aged garments that have experienced a number of wash and wear cycles and, hence, have more amorphous regions, tend to show greater sensitivity to cellulase than new garments.

Figure 8. Effect of Carezyme® after 30 wash cycles. (**a**) 30-cycle wash in detergent only; (**b**) 30-cycle wash in detergent plus cellulase.

Cellulose hydrolysis is an acid-catalyzed process. The O-glycosidic bond is cleaved by donation of a proton from the carboxyl group of a glutamic acid residue in the active site of the enzyme, with the resulting cabonium-ion intermediate stabilized by an aspartic acid or other nucleophilic residue (31). The formation of a carbonium-ion intermediate opens the possibility of stereochemistry, with the configuration of the anomeric carbon atom of the product either being retained or inverted. The configuration of the active site of the particular cellulase enzyme determines the sterochemistry of the hydrolysis products, leading to a classification of cellulases as being either retaining or inverting types of enzymes.

Current detergent cellulases are derived from fungal sources, such as *Humicola insolens*, or bacterial sources, such as the *Bacillus* species. The fungal enzymes typically represent a mixture of *exo*- and *endo*-glucanase activities with pH optima in the acidic to neutral range. This mix of activities facilitates the degradation of crystalline cellulose. Celluzyme® from Novo-Nordisk (Table 1) is derived from *H. insolens* strain DSM 1800 and consists of a mixture of *endo*-glucanases, cellobiohydrolases, and β-glucosidases. Carezyme®, on the other hand, is a monocomponent alkaline-active *endo*-glucanase with low affinity to crystalline cellulose.

The detergent cellulases from alkalophilic *Bacillus* species are generally monocomponent *endo*-glucanases with good stability against proteases. Relative to the fungal cellulases, bacterial cellulases tend to show somewhat reduced fabric care benefits and increased detergency effects on pigmented dirt and sebum (30).

Generally, cellulases are multidomain structures comprised of a core catalytic domain, containing the active site, and a cellulose binding domain, or CBD, which is responsible for anchoring the enzyme to the cellulose substrate. Typically, the CBD is attached to the catalytic domain through a short peptide linker. In the absence of the CBD, the core will show catalytic activity against soluble substrates but drastically reduced activity against insoluble cellulose. Because the benefits derived from cellulase clearly rely on the enzyme binding to fabric surfaces, the CBD is a critical structural component. Like lipases, adsorption of cellulase onto fabric is the first step of the hydrolytic process, corresponding to a phase transfer of enzyme from solution to the fabric surface (32). Those enzymes that adsorb tightly are generally the most efficient with respect to cellulose hydrolysis (33). Thus, the CBD is an important structural feature controlling the efficacy of a particular cellulase with respect to detergency benefits. Table 4 shows activity on soluble and insoluble substrates and the depilling benefit for cellulases with (EG V, Carezyme), and without (EG 1, Endolase) a CBD as well as the catalytic core of an EG V enzyme. The importance of the CBD both to hydrolysis of insoluble substrate and in delivering a depilling benefit is apparent in the higher activity of the CBD-containing Carezyme on crystalline cellulose (Avicel) and the more than fivefold increase in depilling benefit versus the Carezyme core.

A number of analytical methods are available for assaying cellulase activity (26,34). The two most common are the viscometric method and measurement of reducing sug-

Table 4. Activity and Performance of Various Cellulases: Role of the Cellulase Binding Domain

Enzyme	CMCU/mg (soluble substrate)	Red-*Avicel*/mg (insoluble substrate)	Relative depilling benefit[a]
EG I (Endolase)	80	0.2	4
EG V (Carezyme)	100	22	100
EG V core	180	8	19

[a]Depilling measured by visual inspection after 25 wash cycles.

ars. Because of their tendency to cleave internal glycosidic linkages, detergent cellulases cause a rapid decrease in chain length of soluble cellulose polymers, resulting in a drop of solution viscosity. This change in viscosity can be measured via standard viscometeric methods and, when compared to a standard sample, can be used as a relative measure of cellulytic activity. Substituted, water-soluble cellulose substrates are commonly used in this assay, such as carboxymethylcellulose (CMC) or hydroxyethylcellulose.

Measurement of the reducing sugars formed upon hydrolysis of a cellulose substrate is often used in conjunction with the viscometric method to assess cellulase activity. Reducing sugars can be detected colorimetrically by ferricyanide.

Peroxidase

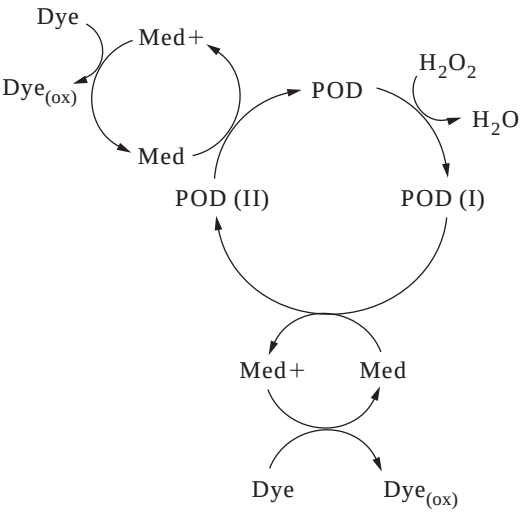
This represents the newest class of enzyme developed for detergent. There is currently only one commercially available peroxidase for detergents, Guardzyme® from Novo-Nordisk A/S (Table 1). Guardzyme® was isolated from the inkcap mushroom, *Coprinus cinereus*, and has been cloned into a production host for large-scale production (35). It is a heme-containing protein that, in the presence of hydrogen peroxide, can mediate the oxidation of fugitive dyes in solution, thereby providing a dye-transfer inhibition benefit.

Peroxidases (EC 1.11.1.7) are a subclass of the general class of enzymes known as oxidoreductases. Oxidoreductases catalyze oxidation–reduction reactions on a broad range of substrates, including alcohols, carboxyl groups, alkenes, and amines. Peroxidases specifically catalyze the reaction:

$$\text{Donor} + H_2O_2 \rightarrow \text{Oxidized donor} + 2H_2O$$

where the donor molecule can be anything with an oxidizable group. Specifically, if the donor molecule is a fugitive dye that has bled off a colored garment, peroxidase can be used to bleach the dye in solution, effectively eliminating the transfer of the colored dye species to other fabrics.

For Guardzyme®, the dye-transfer inhibition system consists of three components: the enzyme, a mediator, and hydrogen peroxide. The mediator is generally a substituted phenolic compound that shuttles electrons from the substrate to peroxidase. Through selection of the appropriate mediator, a catalytic oxidation cycle can be established as follows:



where POD (I) and POD (II) are oxidized forms of the peroxidase. Figure 9 shows the dye-transfer inhibition benefit for a Guardzyme®-mediator system, where the mediator is phenothiazine-10-propionic acid. The peroxidase-based dye-transfer inhibition system essentially eliminates transfer of dye from a blue bleeder (Direct Blue 1 on cotton) to a white cotton tracer.

FORMULATION OF DETERGENT ENZYMES

The key challenge to the use of enzymes in detergents is stability. This is particularly true in liquid detergents, where the enzyme is in intimate contact with the formulation ingredients. For powder detergents, enzymes are typically supplied as a coated granulate, thereby separating the enzyme from other detergent ingredients and generally delivering a stable composition. Liquid detergents, however, must be formulated to minimize damage to the enzyme. Denaturation by surfactants and builders, unfavorable interactions with perfume ingredients, and deg-

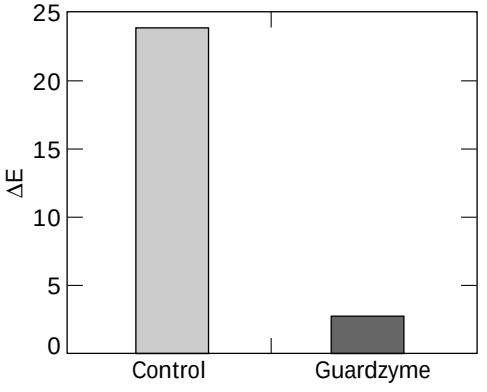


Figure 9. Dye transfer inhibition benefit for Guardzyme®/mediator under North American wash conditions. Dye transfer was measured by recording the difference in reflected light (reported as ΔE from a Hunter colorimeter) from a white cotton tracer swatch before and after washing with a bleeder fabric (Direct Blue 1 on cotton) in a North American powder detergent at 35 °C with 1-mM water hardness.

radation by protease are the principal processes responsible for loss of enzyme activity upon storage.

Liquids

Most modern liquid detergents contain fairly high concentrations of anionic surfactants for cleaning. These surfactants are particularly good at denaturing proteins, including enzymes. In addition, most heavy-duty liquids contain multiple enzyme systems, including proteases, amylases, lipases, and cellulases. Proteolytic degradation of these enzymes by the added protease poses a significant long-term stability issue.

In order to achieve acceptable shelf stability and ensure delivery of a high level of cleaning performance over time, detergent manufacturers and enzyme suppliers have taken three basic approaches to improve the stability of enzymes in liquid detergents:

1. *Formulation of stabilizing additives*. A number of ingredients can be added to the liquid detergent to improve enzyme stability. Calcium is known to stabilize both proteases and amylases. The added calcium serves to tie up detergent actives that could, otherwise remove the essential calcium from the enzyme, resulting in a loss of the catalytic three-dimensional structure. Figure 10 shows the effect added calcium has on the stability of a protease in a heavy-duty liquid detergent. The half-life of the enzyme under stressed storage conditions increases as calcium concentration is increased.

   In addition, short-chain carboxylates have been shown to promote protease stability (36) through a noncompetitive-inhibition mechanism. Accordingly, calcium and sodium formates are commonly added stabilizing agents in enzyme-containing liquid detergents. Reversible protease inhibitors can also be added to shut down the activity of the protease in product. The combination of a polyol with vicinal hydroxyl groups, such as propylene glycol, and a boron compound, such as boric acid, results in formation of a borate-diol complex that effectively inhibits proteolytic activity. Upon dilution of the product into the wash, the inhibitor is diluted away from the protease yielding a fully active enzyme. Substituted boronic acids (37) and peptide derivatives (38) have also been shown to be effective reversible inhibitors.
2. *Selection of more stable wild-type enzymes*. Enzymes with enhanced stability to detergent ingredients can be obtained by classical screening methods. For example, the Kao Corporation in Japan has reported isolation of protease-resistant cellulases from cultures of *Streptomyces* strains (39). Kawai et al. (40) identified a number of new cellulases from *Bacillus* species with good stability to surfactants, proteases, and chelants. Similarly, new α-amylases have been found that do not appear to be substantially affected by anionic surfactants (41). Natural isolate screening for new wild-type lipases with improved detergent compatibility has met with limited success to date, although lipases from *Fusarium culmorium* have been isolated and reportedly show improved activity versus Lipolase® in the presence of alcohol ethoxylate surfactants (42).
3. *Protein engineering*. Another way to improve stability of enzymes for detergents is to engineer in specific stabilizing mutations. For example, replacing a neutral or positively charged amino acid in the calcium-binding domain of an enzyme with a negatively charged amino acid would be expected to increase electrostatic interactions, thereby minimizing stability loss by removal of the essential calcium ion. Aehle et al. (43) used this approach to modify the calcium binding site of an α-amylase. The resulting variant showed a 2.5-fold increase in stability versus the native enzyme in low calcium environments. Similarly, amino acid substitutions in the vicinity of a hydrophobic domain in the alkaline *Bacillus lentus* protease are reported to improve storage stability of the resulting variants in both liquid detergents and soap bars (44). Schülein et al. (45) reported a number of modifications to Carezyme® that improved its alkaline activity and detergent compatibility.

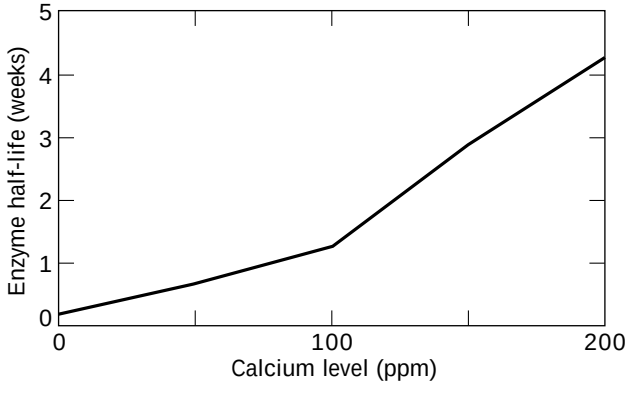


Figure 10. Half-life of protease in a heavy-duty liquid detergent as a function of calcium level. Liquid detergent stored at 35 °C. Enzyme activity measured by PNA assay.

In addition to stability, aesthetics are another major consideration for the formulation of enzymes into liquid products (46). Odor, color, and clarity are key aesthetic signals for consumers. Enzyme raw materials, being products of microbial fermentation, often contain a number of smelly by-products, including volatile amines and low molecular weight carboxylic acids, that must be removed to avoid consumer malodor complaints. Likewise, fermentation-derived by-products may cause turbidity, flocculation, or color changes in liquid detergent formulas. Therefore, the formulation of enzymes into liquid detergents requires extensive evaluations of odor, color, and clarity in order to select enzyme formulations that minimize aesthetic negatives.

Powders

Many of the same considerations for the formulation of enzymes in liquids apply to powders. Odor, color, and stability are important considerations although to a somewhat lesser extent owing to the physical separation of the en-

zyme from the bulk detergent via granulation. In addition, particle size, shape, flowability, and bulk density become important when considering mixing and segregation of enzyme particles with powdered detergents. Probably the most critical factor in formulating solid enzymes is dust generation during detergent manufacture. This was clearly evident in the experiences of the detergent industry in the late 1960s and early 1970s, with the raw powdered enzyme formulations originally supplied for detergent use. These formulations led to high airborne dust levels in the manufacturing plants and an increase in worker sensitization. As a result, enzyme manufactures replaced these early powder and agglomerated enzymes with various granulate formulations (13,46), the most important of which are summarized below:

1. *Enzyme prills*. In this process, dry enzyme is dispersed into a molten wax, nonionic surfactant, or polymer matrix, then sprayed into a cooling tower to form solid, spherical particles. The molten phase is typically comprised of a water-soluble or water-dispersable material with a melting point above 50 °C. The key advantages of prilling are high throughput and the ability to recycle particles that fall outside the desired size range. The key disadvantage is the relatively poor physical strength of the particle, leading to break up and high dust generation in subsequent processing. Overcoating with an inert, mechanically tough material such as polyethylene glycol can be used to improve physical strength and reduce dustiness. However, this typically is not sufficient to ensure no break up of the particles during detergent processing and, as a result, enzyme prills are not in widespread use today.
2. *Extruded granulates*. By combining an enzyme powder or liquid concentrate with binders, such as clay, sugar, or starch, an extrudable dough can be produced that can then be pressed through a perforated metal plate. Upon drying, the extrudates are reasonably durable. In order to produce spherical particles of the appropriate size the wet extrudate is fed into a Marumerizer, where a rotating textured plate breaks up the noodles into smaller particles and rounds the corners to yield a more spherical shape. After sieving, the particles are coated with pigments, such as TiO_2, and a protective outer layer to achieve the desired appearance and further improve granulate integrity. Key disadvantages of this method are the high capital investment in a multistep process and the sensitivity of the process to variation in feedstock moisture and composition.
3. *High shear granulation*. This method is used by Novo-Nordisk to produce their T-granulate products (47). In this process, enzyme is mixed with controlled amounts of water, binders, such as polyethylene glycol, ethoxylated fatty alcohols, fatty acids, bentonite, kaoline, or other clays, and other granulation ingredients to form a low-moisture agglomerate. The agglomerate is fed into a high-shear mixer that breaks it up into smaller particles. The particles are dried in a fluidized bed and coated with a final protective layer of polymers and pigments. By incorporating cellulose fibers into the original agglomerate, particles with high mechanical stability and physical strength can be produced (48). Enzyme T-granulates are widely used in dry laundry products and are among the lowest-dusting enzyme raw materials available. The process is reasonably flexible, especially with respect to enzyme feedstock concentration, allowing granulation of a given enzyme, such as Savinase®, across a wide range of enzyme payload.
4. *Fluid bed coating*. In this process, liquid enzyme feedstock is sprayed onto a solid, inert support, or core, such as NaCl or a sucrose-starch nonpareil. Once the enzyme layer has dried, additional coatings can be applied, including stabilizers, chelants, antioxidants, and pigments. The outer coating typically consists of a film-forming polymer, such as polyvinyl alcohol, and a pigment, such as TiO_2. Genencor International uses this granulation process to deliver their Enzoguard line of dry enzyme formulations. The key advantage of fluid bed coating is that a single piece of equipment, the fluid bed coater, can be used to produce the final product.

Figure 11 shows scanning electron micrographs of typical granulates produced by the high shear granulation process (**a**) and fluid bed coating (**b**). Both processes produce spherical granulates with fairly uniform particle size.

Industrial Hygiene

Controlling protein dust from enzyme raw materials is a key consideration in the manufacture of enzyme-containing laundry products because enzymes are proteins, like animal dander or pollen, with the potential to act as respiratory allergens. Enzyme formulations, such as the granulate compositions described above, that minimize protein dust or reduce the tendency of enzyme protein to form aerosols are critical to protein dust control. Just as important is the need for appropriate industrial hygiene practices and in-plant exposure control technology to further minimize the potential for exposure to enzyme protein. This involves compliance to operational guidelines that require no visible dust in the manufacturing environment, no recurring enzyme spills or spills of enzyme-containing detergent, no gross skin contact with raw material enzyme, and cleanup processes that minimize dust and aerosol generation. Design of dust handling equipment should follow current best practices for enclosure and local exhaust ventilation. Spill-containment systems should be in place for those areas where either spills of raw material enzyme or enzyme-containing product are likely. Additionally, an appropriate preventive maintenance program should be implemented to ensure the equipment performs as intended. Strict adherence to guidelines such as these has resulted in a drop in airborne enzyme dust levels in North American and Western European detergent manufacturing plants by more than three orders of magnitude versus 1970 values (48).

Figure 12 shows a simplified schematic for safe handling of bulk solid enzymes in a detergent manufacturing

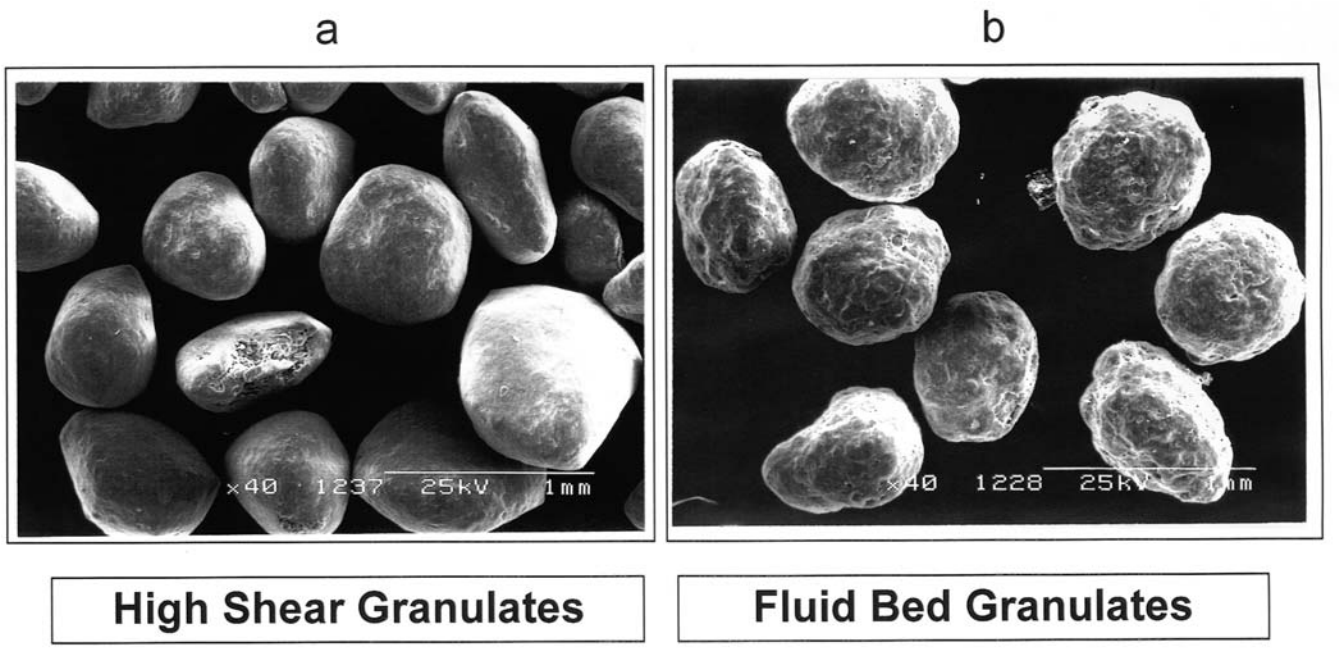


Figure 11. Electron micrographs (40×) of enzyme granulates. (**a**) High shear granulates; (**b**) Fluid bed granulates.

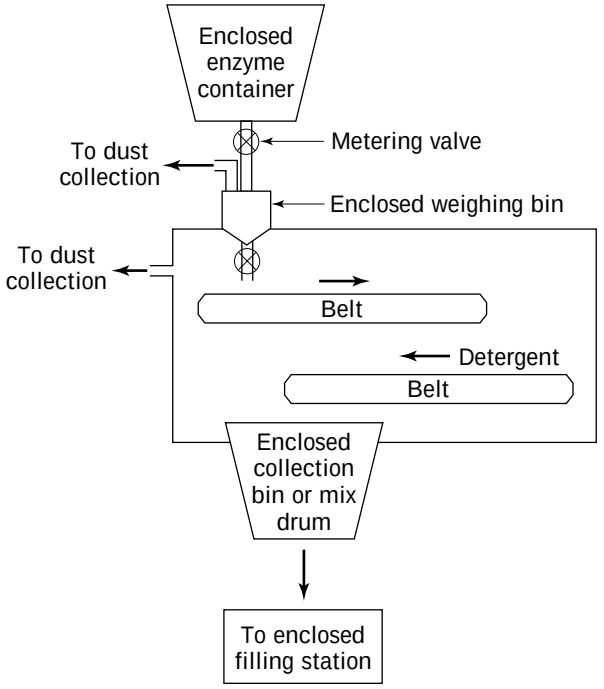


Figure 12. Simplified schematic of dry enzyme handling station.

environment. Enzymes are received from the supplier in an enclosed container that can be coupled to a metering system for accurate dosing. From the weigh station, enzyme is mixed with a stream of detergent base granules in a dust-proof enclosure and sent to either a mix drum or holding bin in preparation for pack-out to cartons. The entire process, including packaging of the final product into boxes or cartons, is housed in dust-proof enclosures to minimize airborne protein dust. Systems such as these ensure low risk for enzyme handling.

MANUFACTURE OF DETERGENT ENZYMES

Nearly all detergent enzymes in use today are produced through large-scale fermentation of microorganisms. As discussed in "Current Detergent Enzymes", detergent enzymes are derived from bacterial and fungal sources. Large-scale fermentation is typically carried out with optimized strains of appropriate bacterial (*Bacillus* or *Pseudomonas* strains) or fungal species (*Aspergillus*) that have been genetically engineered to express high levels of the target enzyme, excrete that enzyme into the bulk fermentation medium for ease of recovery, and minimize production of other, unwanted proteins. After fermentation, the enzyme is recovered from the broth by removal of the whole cells and other particulates through centrifugation or filtration, concentrated via ultrafiltration and diafiltration techniques, and, if necessary, further purified by selective precipitation of nonactive proteins (13).

Because multiton quantities of low-cost enzymes are needed to support the requirements of the global detergent business, enzyme manufacturers closely adhere to the following principles (49):

1. The production organism must be capable of secreting the enzyme extracellularly into the bulk fermentation medium. This greatly facilitates recovery because the cost, both in time and money, to recover intracellular enzymes is to great to support the needs of the detergent industry. Enzyme secretion

can be maximized by appropriate optimization of fermentation conditions.

2. The production organisms should be able to be tuned to give the highest possible yields. Strain optimization can be accomplished either through classical mutagenesis and screening methods or via genetic engineering, where multiple copies of the gene coding for the enzyme of interest are inserted into the host DNA so that each microbe in the strain can produce several copies of the enzyme of interest.
3. The number of steps in the downstream recovery process should be minimized to avoid yield losses and keep overall processing costs down.
4. The production organisms should produce the target enzyme in relatively high purity with no contaminating side activities or proteins. This ensures that only basic recovery methods, such as filtration and centrifugation, are needed to obtain the final product and eliminates the need for high-cost process steps like solvent precipitation. By deleting the genes that code for unwanted enzymes and proteins, strains can be prepared that secrete only the enzyme of interest into the bulk media.

Herrmann et al. (49) provide a thorough review of the principles and practices of detergent enzyme manufacture.

TRENDS IN DETERGENT ENZYMES

A number of market trends and consumer needs are currently driving research on new enzyme technologies for detergent application. Two critical changes occurring in the marketplace today that are expected to have a major impact on future enzyme technologies are the continued move to cooler wash temperatures and fundamental changes in washing machine design leading to reduced water consumption. In the United States, there has been a slow but steady decline in wash temperatures since the early 1980s. This trend has been matched in Europe where more loads are being washed at 40 °C. Asia and Latin America have historically been cold-water-wash geographies. In order to deliver consumer valued benefits under these conditions, enzymes with improved activity at low temperatures will be required, particularly for those laundry stains and soils that rely on thermal energy for removal, such as mechanical or food grease and oil.

In response to anticipated legislation on energy consumption, U.S. washing machines will become more energy efficient, work at lower temperatures, and require less water per load. The net effect will be a much higher ratio of fabric and soil to wash liquor. These new machine designs will likely require enzymes with increased stability to other detergent ingredients in low water environments. In addition, changes in enzyme raw material formulation may be necessary for improved solubility.

There is a trend in the detergent industry toward development of technologies that can deliver fabric care benefits. This ranges from enzymes that can maintain the color of a garment over successive wash cycles to systems capable of reducing the wear associated with the laundry process. As a result, there is increased emphasis on identification and development of enzymes that can extend the useful wear life of fabrics. Cellulases are probably the best example of this for cotton garments. Recently, proteases have been considered as a way to deliver comparable benefits to washable wools and silks.

Finally, there is growing concern among consumers about communicable disease and the ability of laundry and cleaning products to eliminate this risk. The trend toward colder wash temperatures is expected to decrease the efficacy of traditional detergent ingredients, especially oxygen bleaches, toward microbial kill. Recent patent literature has cited the use of *endo*-glycosidases alone for removal of bacteria from fabric and, in combination with traditional antimicrobial actives, for improved microbial kill during the wash process (50). Antimicrobial detergent compositions in which the active is a peroxidase or laccase have also been disclosed (51).

Current trends in the development of new detergent enzymes are focused on improved activity in cold water, increased stability and compatibility with other detergent actives, and the identification of new enzymes that can deliver new benefits such as fabric care and santization.

ACKNOWLEDGMENTS

The author wishes to thank the following for their help in preparing this chapter: Chanchal Ghosh, Genevieve Kipte, Yong Zhu, Eric Wells, Paul Correa, Michael Fedak, Stanton Boyer, and Christian Thoen.

BIBLIOGRAPHY

1. German Pat. DE 28923 (1913), Otto Röhm.
2. J.H. Houston, in J.H. van Ee, O. Misset, E.J. Baas eds., *Enzymes in Detergency*, vol. 69, Surfactant Science Series, Dekker, New York, 1997, pp. 11–21.
3. T. Krawczyk, *INFORM*, **8**, 6–14, (1997).
4. P. Van Arnum, *Chem. Market. Reporter*, **251** (suppl.), 5–6 (Jan. 27, 1997).
5. Business Communications Company, Inc., Study No. C-147NA: *Enzyme for Industrial Applications: Products, Uses, Technologies and Economic Values*, P. Rotheim (1998).
6. E.M. Kirschner, *Chem. Eng. News*, p. 46 (Jan. 26, 1998).
7. *Chem. Market. Reporter*, SR18 (Jan. 16, 1995).
8. M.G. Venegas, in M.S. Showell ed., *Powdered Detergents*, vol. 71, Surfactant Science Series, Dekker, New York 1997, pp. 295.
9. E.J. Bass, M.M.P. Bollier, P.F. Plank, and D.S. Winetzky, in J.H. van Ee, O. Misset, E.J. Baas eds., *Enzymes in Detergency*, vol. 69, Surfactant Science Series, Dekker, New York, 1997, pp. 23–59.
10. C.E. Stauffer and D. Etson, *J. Biol. Chem.* **244**, 5333–5338 (1969).
11. J.A. Wells, D.B. Powers, R.R. Bott, B.A. Katz, M.H. Ultsch, A.A. Kossiakoff, S.D. Power, R.M. Adams, H.H. Heyneker, B.C. Cunningham, J.V. Miller, T.P. Graycar, and D.A. Estell, in *Protein Engineering*, Alan R. Liss, 1987, pp. 279–287.
12. M. Egmond, in J.H. van Ee, O. Misset, E.J. Baas, eds., *Enzymes in Detergency*, vol. 69, Surfactant Science Series, Dekker, New York, 1997, pp. 61–74.

13. E. Gormsen, E. Marcussen, and T. Damhus, in *Enzymes in Detergency*, J.H. van Ee, O. Misset, E.J. Baas eds., vol. 69, Surfactant Science Series, Dekker, New York, 1997, pp. 137–163.
14. W. Aehle, in J.H. van Ee, O. Misset, E.J. Baas eds., *Enzymes in Detergency*, vol. 69, Surfactant Science Series, Dekker, New York, 1997, pp. 213–229.
15. PCT Pat. Application WO 95/10603 (1995) H. Bisgaard-Frantzen, T. Borchert, A. Svendsen, M.H. Thellersen, and P. Van Der Zee (to Novo-Nordisk A/S).
16. T. Takagi, H. Toda, and T. Isemura, in P.D. Boyer ed., *The Enzymes*, vol. 5, Academic Press, New York, 1971, pp. 265–266.
17. E.H. Fischer and E. Stein, *Biochem. Prep.* **8**, 27 (1961).
18. J.F. Robyt and S. Bemis, *Anal. Biochem.* **19**, 56 (1967).
19. S. Dygert, L.H. Li, D. Florida, and J.A. Thoma, *Anal. Biochem.* **13**, 367 (1965).
20. East German Pat. 14,296 (1958), F. Leidholdt. French Pat. 2,011,578 (1970), C.M. Brooke and R.T. Littleton. British Pat. 1,442,418 (1976), R.E. Montgomery.
21. *Chem. Eng. News*, p. 6 (Feb. 8, 1988).
22. O. Misset, in J.H. van Ee, O. Misset, E.J. Baas, eds., *Enzymes in Detergency*, vol. 69, Surfactant Science Series, Dekker, New York, 1997, pp. 107–131.
23. R. Verger and G.H. de Haas, *Annu. Rev. Biophys. Bioeng.* **5**, 77 (1976).
24. A. Svendsen et al., in *Protein Engineering of Microbial Lipases*, Academic Press, New York, 1997.
25. P.N. Christensen and E. Gormsen, *Chim. Oggi.* **11**, 15 (1993).
26. N. Eriksen, in T. Godfrey and S. West eds., *Industrial Enzymology*, 2nd ed. Macmillan London, 1996.
27. Japanese Pat. 5,631,674 (1981), M. Murata.
28. E. Hoshino and S. Ito, in J.H. van Ee, O. Misset, E.J. Baas eds., *Enzymes in Detergency*, vol. 69, Surfactant Science Series, Dekker, New York, 1997, pp. 149–174.
29. S.M. Colwell, *Soap Cosmetics Chem. Specialties*, **70**, 28 (1994).
30. K.-H. Maurer, in J.H. van Ee, O. Misset, E.J. Baas eds., *Enzymes in Detergency*, vol. 69, Surfactant Science Series, Dekker, New York, 1997, pp. 175–202.
31. J. Knowles, P. Lehtovaara, and T. Teeri, *Trends Biotechnol.* **5**, 255 (1987).
32. B. Henrissat, *Cellulose* **1**, 169 (1994).
33. A.A. Klyosov, *Biochemistry* **29**, 10577 (1990).
34. K.R. Sharrock, *J. Biochem. Biophys. Methods* **17**, 81 (1988).
35. E. Gormsen, *JAOCS/AOCS Conference*, Kuala Lumpur, Malaysia, September 1997.
36. M.C. Crossin, *JAOCS*, **66**, 1010–1014 (1989).
37. U.S. Pat. 5,442,100 (1995), D.W. Bjorkquist and R.K. Panandiker (to Procter & Gamble Co.) U.S. Pat. 5,472,628 (1995, R.K. Panandiker et al. (to Procter & Gamble Co.) U.S. Pat. 5,488,157 (1996), D.W. Bjorkquist and R.K. Panandiker, (to Procter & Gamble Co.).
38. WO 98/13461 A1 (1998) (to Procter & Gamble Co.).
39. JP0629575 A (1994) (to Kao Corp).
40. EP 270974 and EP 269977 (1988), S. Kawai et al. (to Kao Corp.).
41. JP09206073 (1998) (to Kao Corp.).
42. WO 96/19570 A1 (1994) (to Novo-Nordisk A/S).
43. W. Aehle, *American Oil Chemists Society 86th Annual Meeting*, San Antonio, 1995.
44. WO 96/34935 (1996) (to Unilever).
45. WO 94/07998 (1994), M. Schülein, H. Fredholm, C. Hjort, G. Rasmussen, E. Nielsen, and P. Rosholm (to Novo-Nordisk A/S).
46. T. Becker, G. Park, and A. Gaertner, in J.H. van Ee, O. Misset, E.J. Baas eds., *Enzymes in Detergency*, vol. 69, Surfactant Science Series, Dekker, New York, 1997.
47. U.S. Pat. 4,106,991 (1978), E.K. Markusseen and A.W. Schmidt, (to Novo Industries A/S).
48. G. Peters and D.P. Mackenzie, in J.H. van Ee, O. Misset, E.J. Baas, eds., *Enzymes in Detergency*, vol. 69, Surfactant Science Series, Dekker, New York, 1997.
49. H.A. Herrmann, I. Good, and A. Läufer, in J.H. van Ee, O. Misset, E.J. Baas eds., *Enzymes in Detergency*, vol. 69, Surfactant Science Series, Dekker, New York, 1997.
50. U.S. Pat. 5,238,843 (1993), R.S. Carpenter, I.J. Goldstein, and P.J. Lad (to Genencor Int. and Procter & Gamble Co.); U.S. Pat. 5,258,304 (1993), R.S. Carpenter, P.J. Lad, and A.M. Wolff (to Genencor Int. and Procter & Gamble Co.).
51. WO 97/43383 A-1 and WO 97/42825—A1 (1997) (to The Procter & Gamble Co.).

See also PROTEIN EXPRESSION, SOLUBLE; PROTEIN PURIFICATION, AQUEOUS LIQUID EXTRACTION; PROTEIN SECRETION, *SACCHAROMYCES CEREVISIAE*; PROTEIN ULTRAFILTRATION.

ENZYMES, DIRECTED EVOLUTION

FRANCES H. ARNOLD
PATRICK L. WINTRODE
California Institute of Technology
Pasadena, California

KEY WORDS

Biocatalysis
Catalyst improvement
DNA shuffling
High-throughput screening
Protein engineering
Random mutagenesis
Recombination
Selection

OUTLINE

INTRODUCTION

Staunch Darwinists attribute all the complexity of living things to an algorithm of mutation and natural selection. The exquisite products of this evolution algorithm are apparent at all levels, from the amazing diversity of life all the way down to individual protein molecules. Scientists and engineers who wish to redesign these same molecules are now implementing their own versions of the algorithm. Directed evolution allows us to explore enzyme functions never required in the natural environment and for which the molecular basis is poorly understood. This bottom-up design approach contrasts with the more conventional, top-down one in which proteins are tamed "rationally" using computers and site-directed mutagenesis. We will describe how molecular evolution can be directed in the test tube in order to produce useful biocatalysts. It is not possible to provide a complete account of all the methods proposed for in vitro evolution; this article therefore introduces methods and strategies used successfully in our laboratory for directing the evolution of enzymes. Some alternative methods for biocatalyst evolution are also be discussed. The reader should also be aware that there is a rather substantial and largely separate literature on combinatorial approaches to engineering binding molecules. Recent advances in the ability to create genetic diversity and to screen or select for improved functions in large libraries of enzyme variants are being combined in a robust approach to solving difficult molecular design problems. With directed evolution we now have the ability to tailor individual proteins as well as whole biosynthetic and biodegradation pathways for biotechnology applications.

WHY DIRECTED EVOLUTION?

The Limitations of Nature's Biocatalysts Present Insurmountable Challenges to Rational Design

When natural enzymes are recruited for industrial applications—from serving as catalysts in chemicals synthesis to additives for laundry detergents—we discover that they are often not well suited to these tasks. Due to poor substrate solubility, breakdown of unstable products, or competing chemical reactions, the conditions for an enzyme reaction may be unsuitable for large-scale applications. Reflecting their participation in complex biochemical networks inside living cells, enzymes are often inhibited by their own substrates or products, either of which may severely limit the productivity of a biocatalytic process. Evolution is usually the culprit: enzymes are optimized and often highly specialized for specific biological functions within the context of a living organism. Biotechnology, in contrast, needs enzymes that are stable and active over long periods of time (a feature that might clash with the need for rapid protein turnover inside a cell), are active in nonaqueous solvents (a feature probably not required in most biological milieu), and can accept different substrates (substrates not present in nature).

It is possible to produce new enzymes in recombinant organisms, altering the amino acid sequence and therefore the properties through appropriate modifications at the DNA level. We are hobbled, however, by near complete ignorance of how the amino acid sequence affects every aspect of enzyme performance, from its ability to be expressed in a heterologous host to its catalytic activity in nonnatural environments. Numerous protein engineering experiments have demonstrated that changes in protein properties are brought about by the cumulative effects of many small adjustments, many of which are distributed or propagated over significant distances. Furthermore, proteins are usually teetering on the brink of instability, with folded structures that are more stable than unfolded—and therefore inactive—ones by the equivalent of a few hydrogen bonds out of the hundreds that form. Superimposing on the need to retain this relatively fragile folded state, the additional requirements of having to fold in the first place, and the need of maintaining or even reengineering a catalytic site that is affected at some level by virtually any modification yields a design problem of such complexity that any rational design effort will require enormous inputs of structural, mechanistic, and dynamic information. Information that is available for but a tiny fraction of interesting catalysts. Even if one trait is successfully designed (e.g., enhanced stability), it is virtually impossible to predict the cost to another (e.g., catalytic activity or expression level). The relatively few examples where rational design has yielded *useful* enzymes do not negate the view that rational enzyme design is often a fruitless exercise.

Extending Natural Diversity by Laboratory Evolution

All these hurdles to the rational design of enzymes are bypassed by evolution. The power of evolution as an algorithm for molecular design is perhaps best appreciated by studying its products. By constructing the evolutionary histories of today's proteins we have learned that they are highly adaptable molecules, at least on evolutionary time scales. Well illustrated by the panoply of α/β-barrel enzymes (1), enzymes catalyzing very different reactions have evolved divergently from a common ancestral protein of the same general structure, acquiring diverse capabilities by processes of random mutation, recombination, and natural selection. We also know that enzymes sharing a common function (for example, all catalyzing a particular

step in a metabolic pathway) in addition to three-dimensional structure can exhibit widely different properties (stability, solubility, tolerance to pH, etc.), depending on where they are found.

Enzymes evolve, and adapt, at the molecular level. The structures of protein modules are conserved (although the modules themselves are often shuffled to create new, multifunctional proteins). Function, however, can vary. Specific features such as substrate specificity or thermostability vary significantly. Amino acid sequences can vary to such an extent that evolutionary relationships may no longer be apparent from sequences alone (1).

Evolution is a powerful algorithm with proven ability to alter enzyme function and especially to "tune" enzyme properties. It is also an algorithm that can be implemented in the laboratory for redesign. The challenge is to collapse the time scale to months, or even weeks.

CHOOSING AN EVOLUTIONARY STRATEGY

Evolutionary mechanisms at work in nature assure adaptability to ever-changing environments. Evolution does not work toward any particular direction, nor is there a goal; the underlying processes occur spontaneously during reproduction and survival of the whole organism. In contrast, a directed molecular evolution experiment has a defined goal, and the key processes (mutation, recombination, and screening or selection) are controlled by the experimenter. Although there may be multiple ways to reach a defined goal (i.e., a desired enzyme function), the approach that minimizes the effort is preferred.

The major steps in a typical directed enzyme evolution experiment are outlined in Figure 1. The genetic diversity for evolution is created by mutagenesis and/or recombination of one or more parent sequences. These altered genes are inserted into a plasmid for expression in a suitable host organism (bacteria or yeast). Clones expressing improved enzymes are identified in a high-throughput screen or by selection, and the gene(s) encoding those improved enzymes are isolated and recycled to the next round

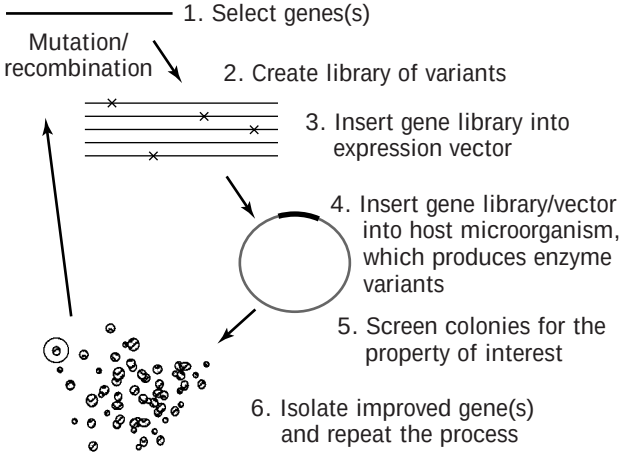


Figure 1. Key steps of a typical directed enzyme evolution experiment.

of directed evolution. Approaches for carrying out these key steps are discussed in some detail in the "Methods of Directed Evolution" section. Here we focus on more fundamental considerations that help to define workable strategies for directed evolution.

To appreciate the challenge of designing and carrying out a successful directed protein evolution experiment, it is important to underscore the powerful combinatorial features of this system. A typical enzyme is a linear chain of N amino acids (N is usually several hundred), and there are 20 possible amino acids at each position in the chain. Thus the "sequence space" of possible proteins is huge beyond the imagination (N^{20}). Even in 4 billion years, nature has had a chance to explore but a tiny fraction of these possibilities. A laboratory exploration of this vast space of sequences and their corresponding functions must obviously be severely limited and carefully guided (2). Because much of sequence space will be devoid of the desired function, and probably even of folded proteins, it is best to direct the evolution of an existing enzyme rather than look for function in random peptide libraries.

Evolution is often referred to as a hill-climbing exercise in the fitness landscape of sequence space (3,4). The fitnesses (performance, for laboratory evolution) of the proteins in sequence space make up this landscape, whose most basic features are still quite unknown. The landscape for laboratory evolution will be different for each property or collection of properties undergoing evolution. An uphill climb in a protein landscape is more likely to be successful if it can take place in small steps (one or two amino acid substitutions). The high dimensionality of the surface (there are $19N$ one-mutant neighbors of any given sequence) offers many opportunities to find improved mutants. Although we may never reach the "global optimum," the improvements achieved by taking even a simple random up-hill walk via single amino acid mutations often yield useful results. A widely effective evolutionary strategy, illustrated in Figure 2, is one in which the steps are small (preferably one or two amino acid substitutions in each generation), and multiple such mutations are accumulated either sequentially (5) or by recombination (6,7) to acquire the desired function. Such an approach is compatible with a low level of random mutagenesis over the entire gene. An alternative approach is to direct a much higher level of random mutation to a relatively small region of the gene (8). Both approaches have their advantages. Mutagenesis over the entire gene allows discovery of unanticipated solutions (a common experience). More intense, directed mutation, however, may yield novel combinations of amino acid substitutions, combinations that would be inaccessible by single-step walks because the intermediates are unfavorable. Details of the evolutionary exploration are ultimately dictated by a combination of *(1)* the power of the search tool, *(2)* the frequency of beneficial mutations (usually small!), and *(3)* the choice of starting point(s).

A different approach to creating diversity for directed evolution is the in vitro shuffling of homologous genes, or "family shuffling" (9), illustrated in Figure 3. Here, recombination of two or more parent genes yields a chimeric gene library for evolution of the desired features. Because the recombined sequences are related through divergent evo-

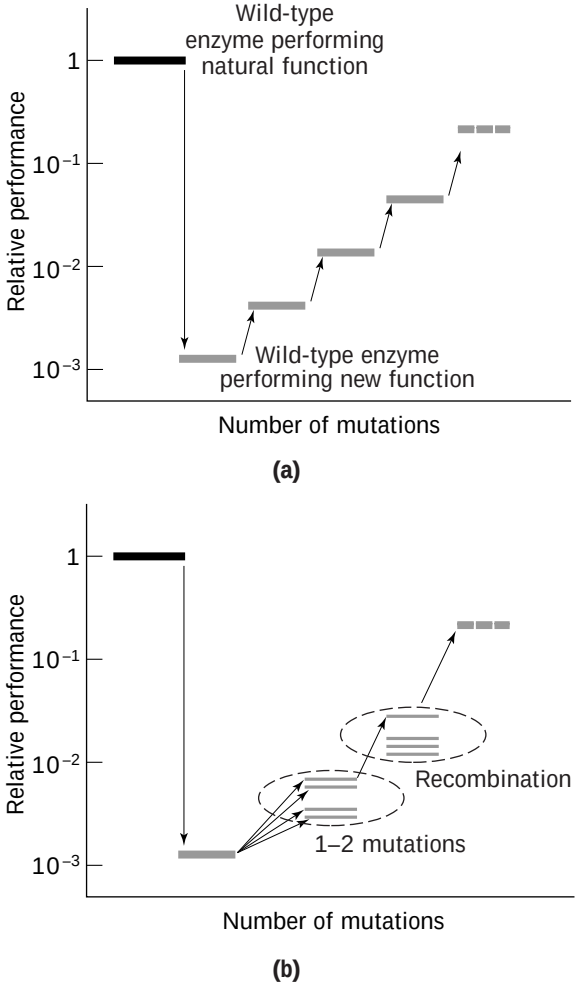


Figure 2. An effective strategy for directed enzyme evolution is one in which small changes associated with one or two amino acid substitutions are accumulated sequentially (**a**) or by recombination of improved genes (**b**).

lution from a common ancestor of similar structure and function (and therefore the sequence differences are to some extent neutral with respect to structure and function), it appears that very large jumps in sequence space can yield functional proteins (9,10). In vivo recombination can also yield interesting new chimeric enzymes (11–14).

Good Problems for Directed Enzyme Evolution

There are four requirements for successful directed evolution. *(1)* The desired function must be physically possible. *(2)* The function must also be biologically, or evolutionarily, feasible. In practice, this means that there exists a mutational pathway to get from here to there through ever-improving variants (see earlier discussion). Although we cannot know a priori that the path exists, a good experiment will maximize the likelihood. *(3)* One must be able to make libraries of mutants complex enough to contain rare beneficial mutations. This usually means functional expression in a suitable microorganism such as *Escherichia coli* or *Saccharomyces cerevisiae*. *(4)* One must have a rapid screen or selection that reflects the desired function. Just how rapid the screen must be depends on how rare mutations leading to the desired property are and how many must be accumulated to achieve the desired result.

Whether directed evolution will solve a particular problem depends to some extent on how hard natural evolution has already worked at it. If a particular trait is already under selective pressure (e.g., catalytic activity), it is unlikely that further improvements can be obtained in the laboratory by small mutational steps. However, if biological function has imposed additional constraints, for example the trait is coupled to another trait that is also under selective pressure (e.g., high thermostability), then this balance can be altered during laboratory evolution (15). While selected traits are often difficult to improve, they should be relatively easy to remove (e.g., product inhibition). Many traits are not under selective pressure; they may be changing as a result of random genetic drift,

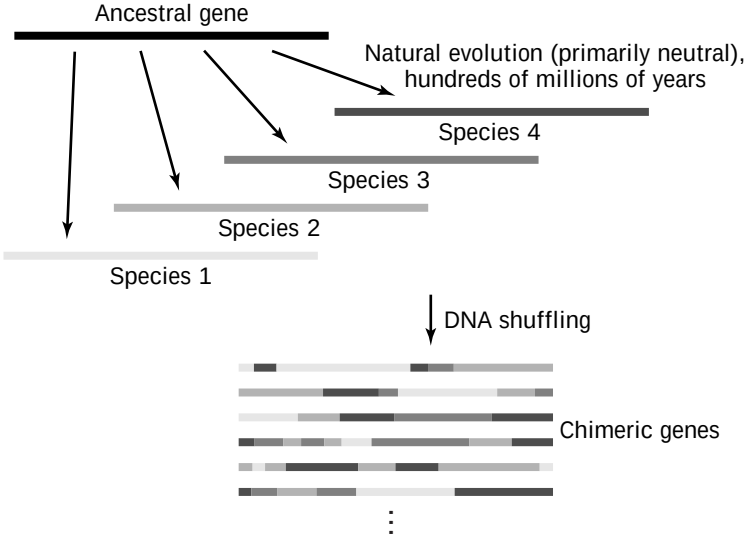


Figure 3. Homologous proteins are descended by divergent evolution from a common ancestral protein and share its overall three-dimensional structure. Recombining homologous genes creates chimeras, some fraction of which should also fold into that structure. Such family-shuffled libraries could be rich in novel function.

they may be vestigial—reflecting the enzyme's history—or they may be coupled to selected traits (16). In general, nonselected traits are easier to improve, but may be more difficult to remove. It is especially easy to improve traits never required for biological function, such as stability or activity in a nonnatural environment or activity toward a new substrate; there is often much room for improvement, and small changes in sequence and function can be accumulated. As expected for any hill-climbing exercise, the number of pathways leading uphill diminishes as a peak is approached. Thus the ease with which improved mutants are identified (the frequency of improved clones) should eventually decrease as the sequences move closer to an optimum.

The preceding discussion focuses only on altering existing enzyme traits. Evolving a completely new function is a risky venture because we rarely know how far in sequence space we have to go in order to create the new function, or how frequent the solutions will be. If there is good reason to believe that a new function, for example, activity toward a substrate not accepted at all by the wild-type enzyme, can be obtained (at a level measurable during a high-throughput screen) by making one or two amino acid substitutions, then evolution makes sense. However, if the new function requires the simultaneous placement of multiple new amino acid residues, it is unlikely to appear in a random library of mutants. Such a problem is probably a better candidate for a combination of rational design and combinatorial tuning (17).

Table 1 summarizes the key features of selected directed enzyme evolution experiments. In nearly all cases, the desired trait(s) was at a measurable, albeit low, level in the starting enzyme(s). The problems can be roughly divided into a few major categories: improving function in nonnatural or extreme environments (where activities or stabilities are low), improving activity toward a new substrate, tuning specificity (enantioselectivity), and increasing functional expression in a heterologous host.

METHODS OF DIRECTED EVOLUTION

Methods for Generating Genetic Diversity

The first step in a directed evolution experiment is the creation of a mutant library containing an appropriate degree of molecular diversity. For this, the mutation rate must be tuned to the power of the sorting method. Useful, reasonably sized libraries can be created by introducing multiple mutations in a particular region or across a limited number of positions (e.g., by combinatorial mutagenesis using oligonucleotide cassettes [42]). However, such an approach excludes the many useful solutions found in unexpected places. We therefore usually try to evolve the entire gene, rather than target particular positions. Whole-gene evolution requires a correspondingly lower mutation rate (see later).

Diversity is created by two operations: point mutagenesis and recombination (Fig. 4). Most random point mutagenesis methods create mutations in single bases. Due to the degeneracy of the genetic code, this provides access to only about 6 amino acid substitutions instead of 19, significantly reducing the potential diversity. In addition, many mutagenesis methods are not really random, further limiting the number of amino acid substitutions actually accessible in a given experiment. For example, a commonly used method, error-prone PCR, shows a strong bias for transitions over transversions (43).

Recombination methods combine sequences from multiple parent genes. In vitro recombination methods include DNA shuffling (6,7,44), random-priming recombination (45), and the staggered extension process (StEP) (46). In yeast, in vivo recombination is particularly simple (11,14). All these methods have a (controllable) level of associated point mutagenesis. Directed evolution can begin from multiple, closely related starting points rather than a single sequence. Recombination of existing functional sequences (i.e., homologous enzymes) will create another level of diversity that point mutagenesis cannot generate (9).

Point Mutagenesis. The most important factors in random mutagenesis are the mutation frequency and mutation bias. Mutation frequency is the average number of mutations per gene and is usually reported as a percentage. The optimal target mutation frequency can be calculated from the length of the DNA coding sequence and an estimate of the desired number of mutations per sequence. In deciding on the optimal number of mutations, one must balance the cost of combinatorially searching large libraries against the likelihood of finding mutants with improved properties. The number of possible variants increases drastically with the number of mutations introduced. The number of possible variants that can be produced by simultaneously substituting M amino acids in a protein of length N is $19^M[N!/(N - M)!M!]$ (47). For example there are only 3,800 possible single mutants, but 7,183,900 double mutants, and 9,008,610,600 triple mutants of a protein of only 200 amino acids. The single-mutant library of an enzyme is generally within our ability to screen exhaustively. Double-mutant libraries are already on the outer limit of current screening technology, and triple-mutant libraries are well beyond our capability for the foreseeable future. To these numerical considerations we add the biochemical fact that beneficial mutations are rare, and combinations of beneficial mutations are extremely rare. In light of these considerations it appears that, at present, single-mutant libraries usually represent the best compromise between search effort and diversity: searching single-mutant libraries maximizes the chance of finding beneficial mutations in a reasonable amount of time. If high-throughput screening technology is available, or if significant fractions of possible mutations are neutral, then double-mutant libraries may be the optimal choice. An average of one amino acid substitution per sequence corresponds to approximately three base substitutions per gene. Thus for a 1-kb sequence, the mutation frequency should be ~0.3–0.5%.

Although numerous methods for making random DNA mutations exist, error-prone PCR is often preferable because the procedure is simple, rapid, robust, and most importantly, the mutation frequency can be precisely controlled. Error-prone PCR does have some intrinsic bias as to the location of mutations (mutations at AT base pairs

Table 1. Selected Examples of Directed Enzyme Evolution

Target enzyme	Target function	Change effected	Approach	Organism	Reference
Kanamycin nucleotidyltransferase	Thermostability	>200-fold increase in half-life at 60–65 °C	Mutator strain + selection	*B. stearothermophilus*	18
Subtilisin E	Activity in organic solvents	~170-fold increase in −60% dimethylformamide	Error-prone PCR + screening	*B. subtilis*	5,19
β-Lactamase	Activity toward new substrate	32,000-fold greater resistance to cefotaxime	DNA shuffling + selection	*E. coli*	6
Subtilisin BPN′	Stability in the absence of Ca^{2+}	1,000-fold increase in half-life	Loop removed +casette mutagenesis + screening	*B. subtilis*	20
para-Nitrobenzyl esterase	Activity toward pNB esters; activity in organic solvent	60 to 150-fold increase	Error-prone PCR and DNA shuffling + screening	*E. coli*	21,22,23
Thymidine kinase	Substrate specificity (gene therapy)	43-fold increase in sensitivity to gancyclovir in hamster cells	Cassette mutagenesis + selection and screening	*E. coli*	8
β-Galactosidase	Activity toward new substrate; substrate specificity	66-fold increased activity; 1,000-fold increase in substrate specificity	DNA shuffling + screening	*E. coli*	24
Subtilisin E	Expression level; activity in organic solvents	500-fold increase in total activity	Error-prone PCR + screening	*B. subtilis*	25
*O*6-Alkylguanine-DNA alkyltransferase	Protection against alkylating agents (gene therapy)	10-fold increased protection against toxic methylating agent	Cassette mutagenesis + selection	*E. coli*	26
Arsenate detoxification pathway	Arsenic resistance	12-fold increased rate of arsenate reduction	DNA shuffling + screening	*E. coli*	27
Aminoacyl-tRNA synthetase	Aminoacylation of a modified tRNA	55-fold increase in activity	DNA shuffling + selection	*E. coli*	28
Aspartate aminotransferase	Activity toward *β*-branched amino and 2-oxo-acids	10^5 increase	DNA shuffling + selection	*E. coli*	29
Lipase	Wash performance	Improved performance in one-cycle wash	Mutagenesis and in vivo recombination + screening	*S. cerevisiae*	30
Lipase	Enantioselectivity in hydrolysis of *p*-nitrophenyl 2-methyldecanoate	Increase in enantiomeric excess from 2% to 81%	Error-prone PCR + screening	*Pseudomonas aeruginosa* PAOI	31
Lipases	Activity toward long chain *p*-nitrophenyl esters	3-fold increase	In vivo recombination of homologous genes + screening	*E. coli*	13

Table 1. Selected Examples of Directed Enzyme Evolution (continued)

Target enzyme	Target function	Change effected	Approach	Organism	Reference
pNB esterase	Thermostability	14 °C increase in T_m + increased activity at all temperatures	Error-prone PCR, DNA shuffling + screening	*E. coli*	15
Esterase	Enantioselectivity of hydrolysis of a sterically hindered 3-hydroxy ester	Increase in enantiomeric excess from 0% to 25%	Mutator strain + selection	*E. coli*	32
Subtilisin E	Thermostability	17 °C increase in T_m + increased activity at all temperatures	Error-prone PCR, DNA shuffling + screening	*B. subtilis*	33
Subtilisin E	Thermostability	50-fold increase in half-life at 65 °C	DNA shuffling + screening	*B. subtilis*	34
B. lentus subtilisin	Expression level (total activity of secreted enzyme)	50% increase	Error-prone PCR + enrichment in hollow fibers	*B. subtilis*	35
Subtilisin BPN′	Activity at 10 °C	2-fold increase	Chemical mutagenesis + screening	*B. subtilis E. coli*	36
3-Isopropylmalate dehydrogenase	Thermostability	3.4-fold increase in activity at 70 °C	Spontaneous mutations + selection	*T. thermophilus*	37
Cephalosporinases	Activity toward moxalactam	270 to 540-fold increased resistance	DNA shuffling of homologous genes + selection	*E. coli*	9
Chorismate mutase	Conversion to monomeric enzyme (solubility)	Functional monomeric enzyme	Oligonucleotide-directed codon mutagenesis + selection	*E. coli*	17
Biphenyl dioxygenases	Degradation of polychlorinated biphenyls (PCBs)	Gained activity toward substrates, poorly degraded by native enzymes, improved activity toward various substrates	DNA shuffling of homologous genes + screening	*E. coli*	38
FLP recombinase	In vivo recombination efficiency at elevated temperatures in *E. coli* and mammalian cells; in vitro thermostability	Improved recombination efficiency in *E. coli* and mammalian cells	Error-prone PCR and DNA shuffling + screening	*E. coli*	39
Echinocandin B deacylase	Activity	3-fold increase	Error-prone PCR, random-priming recombination and screening	*Streptomyces lividans*	40
Horseradish peroxidase	Functional expression in *E. coli*	11-fold increase in expressed activity	Error-prone PCR and screening	*E. coli*	41

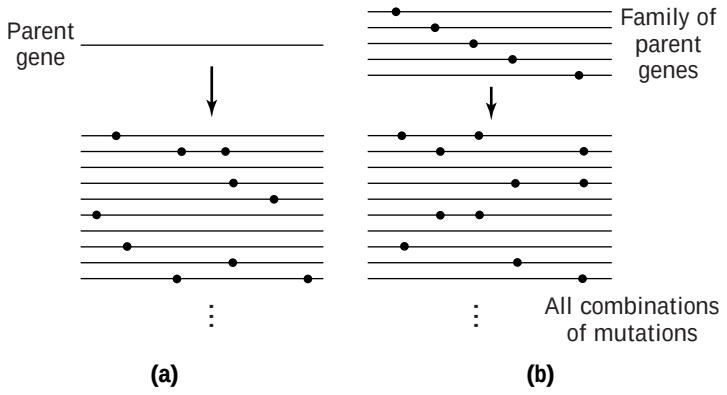


Figure 4. Two general approaches for generating molecular diversity: random point mutagenesis (**a**) and in vitro recombination (**b**).

occur much more frequently than mutations at GC base pairs) as well as their type (A is more frequently substituted with G). PCR modifications reduce bias, but do not eliminate it (48).

The error-prone PCR method routinely used was originally outlined by Leung et al. (49) and further examined by Cadwell and Joyce (43) and Shafikhani et al. (50). Mutation frequencies ranging from 0.11 to 2% have been obtained under different reaction conditions. In particular, the mutation frequency can be controlled simply by adjusting the concentration of manganese ions in the reaction mixture (48).

Because polymerase fidelity also depends on the nature of the target sequences, different mutation rates will be observed on different sequences, even when the exact same reaction conditions are used. A straightforward way to assess the overall mutation frequency and the nature of mutations is to sequence a few random clones from the amplified population. However, sequencing is time-consuming and expensive. A simple and efficient alternative is to estimate the mutation frequency from the fraction of active clones from the amplified population (51). Activity profiles of clones sampled from libraries of subtilisin E variants produced by error-prone PCR with different $MnCl_2$ concentrations are plotted in Figure 5. The clones are sorted and plotted in descending order to create local profiles, or landscapes, of the enzyme's fitness. The higher the Mn^{2+} concentration, the higher the mutation rate, which manifests itself in a higher proportion of inactivated clones. Even though different enzymes show different activity profiles for different mutation rates, the fraction of active clones is a convenient index of mutation frequency and can be used as a diagnostic check for the successful creation of the desired randomly mutated library.

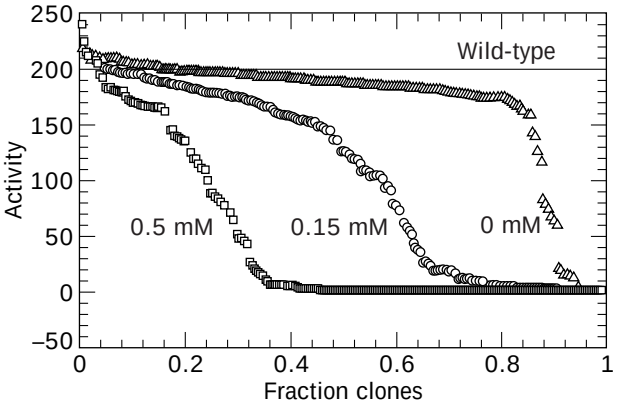


Figure 5. Activity profiles of subtilisin E mutant libraries generated by error-prone PCR with different concentrations of $MnCl_2$. Mutants are sorted according to activity in descending order. Horizontal line indicates the activity of the wild-type enzyme. *Source*: Ref. 51.

Recombination. When an enzyme is evolved by sequential generations of random mutagenesis and screening, only the best variant identified in each generation is used to parent the next generation. Other improved variants are set aside and must be rediscovered in subsequent generations in order to become incorporated. This is wasteful because screening is time-consuming. If the mutation rate is high, this approach can also accumulate deleterious mutations, possibly limiting the fitness that can be reached. "DNA shuffling" (6,7,44) and related in vitro recombination methods (45,46,48) can overcome these limitations of the sequential approach. Recombining parental genes to produce libraries of different mutation combinations, and screening for improved variants quickly accumulates the beneficial mutations, while removing deleterious ones.

The goal is to create gene libraries containing novel combinations of mutations present in the parent genes (Fig. 4). This library can be screened in order to find the combinations of mutations giving rise to the best enzymes. Screening eliminates unfavorable combinations of mutations. For example, if two point mutations that are beneficial by themselves become harmful when combined, the variant carrying this combination will be removed during screening. We now discuss some general considerations for recombination experiments before describing different recombination methods.

Statistics of Random Recombination. When recombining genes from multiple improved variants it is important to consider the recombined library size. The various in vitro recombination methods can recombine any number of parent genes. However, the resulting libraries may be so impossibly large that the probability of finding large improvements in function is effectively zero.

For the recombination of N sequences and M total mutations, the probability of generating progeny sequences containing μ mutations is equal to the number of ways a μ mutation sequence can be generated (C_μ^M) multiplied by the probability of generating any single μ-mutation sequence (22)

$$P_\mu = C_\mu^M \left(\frac{1}{N}\right)^\mu \left(\frac{N-1}{N}\right)^{M-\mu} = \frac{M!}{(M-\mu)!\mu!}\left(\frac{1}{N}\right)^\mu \left(\frac{N-1}{N}\right)^{M-\mu}$$

For recombination of N single mutants the probability of generating single-mutant parents or wild-type grandparents is approximately 75%, and only 25% of the library consists of new sequences. The probability of generating individual sequences decreases precipitously with increasing numbers of parents. The least-frequent sequences (often the most desirable) are those containing the majority of mutations from the parent population. The rarest sequence will be the one containing all mutations ($\mu = M$). The probability P_M of generating this sequence is $1/N^m$.

Some degree of oversampling is required in practice to maximize the chance of discovering a given variant. The sampling S required in order to achieve a given level of confidence of having sampled the rarest variant in a library is given by (22)

$$(1 - P_M)^S < 1 - (\text{confidence limit})$$

Generally, for 95% certainty that a specific clone has been sampled, the oversampling is between 2.6 and 3.0.

Recombination Methods. Stemmer described the first method for in vitro recombination of DNA sequences, called DNA shuffling (6,44). This method, illustrated in Figure 6, involves enzymatic digestion of the parental DNA into short fragments, followed by reassembly of the fragments into full-length genes. Because the fragments are free to associate with complementary fragments from other, similar genes, mutations from one parent can be combined with mutations from other parent(s) to generate novel combinations. The fidelity of this process is not perfect, and new point mutations not present in any of the parents will occur at a finite rate. The original method has been modified to simplify it and to yield better control over the associated mutagenic rate (48,52). The mutagenic frequency can be controlled over a wide range, from about 0.05 to 0.7%, by the inclusion of Mn^{2+} or Mg^{2+}, by the choice of DNA polymerase, and/or by using restriction enzyme digestion to prepare the starting DNA (48). The finite error frequency associated with DNA shuffling has been used by Stemmer and coworkers to supply point mutations for recombination and evolution. When starting with a single sequence, we prefer to use error-prone PCR under controlled conditions to generate libraries of variants. Improved sequences identified during screening can then be recombined.

An alternative method for recombination was developed by Shao et al. (45). In this method the template DNA sequences are primed with random-sequence primers, which are then extended by DNA polymerase to generate a pool of short DNA fragments. After removal of the template, these fragments can be reassembled, as with the Stemmer shuffling method, to give a library of full-length sequences containing information from multiple parents. Yet another method for in vitro recombination, the staggered extension process (StEP), does not require fragmentation and reassembly of the parent sequences (34). This method is illustrated in Figure 7 for recombination of two parent genes. In StEP recombination the template genes are primed and very briefly extended before denaturation and reannealing. The growing fragments can reanneal to different templates and therefore pick up sequence information from different parents as they grow into full-length sequences over hundreds of very brief cycles. StEP (in principle) involves only a single relatively short protocol as opposed to the multiple steps of fragmentation, isolation, and annealing required in DNA shuffling.

"Family Shuffling": Recombination of Homologous Genes. The approaches already described (point mutation and recombination of improved variants) allow a local exploration of sequence space. Only sequences that are quite close to the parent sequence are sampled. One would like to sample distant regions of sequence space to search for new functions, but this is not readily accomplished using these methods. The simultaneous introduction of more than a few random mutations into a wild-type sequence will almost always result in variants that are inactive or unfolded. The probability that an improved enzyme will be found is usually much smaller than the screening capacity. An alternative approach to making large jumps in sequence space is by "family shuffling" (9) (see Fig. 3). In this approach, homologous genes from different organisms are recombined to create a library of chimeric molecules. Family shuffling exploits the fact that the sequence differences between two homologous proteins are *not* random. Consider the two proteases subtilisin E and subtilisin thermitase. These two enzymes differ at 157 of their amino acid positions, and yet they fold to the same overall three-dimensional structure and catalyze the same proteolytic reaction. The amino acid substitutions in these and other naturally occurring subtilisins have been preselected to be largely neutral with respect to folding and gross function. Mutations resulting in a misfolded or inactive enzyme have been eliminated by natural selection. Accordingly, recombination of this pool of largely neutral mutations will be far more likely to result in folded, active proteins than recombination of purely random mutations. The use of family shuffling to access remote unexplored regions of sequence space may yield proteins with a variety of desirable, but not yet found, functions.

Searching for Improved Enzymes

In general, the most time-consuming and expensive part of directed enzyme evolution is setting up and implementing the search for improved variants. Developing an optimal search strategy is crucial for success. A fundamental rule of directed evolution, "you get what you screen for," specifies that the screen (or selection) should reflect the desired result as closely as possible.

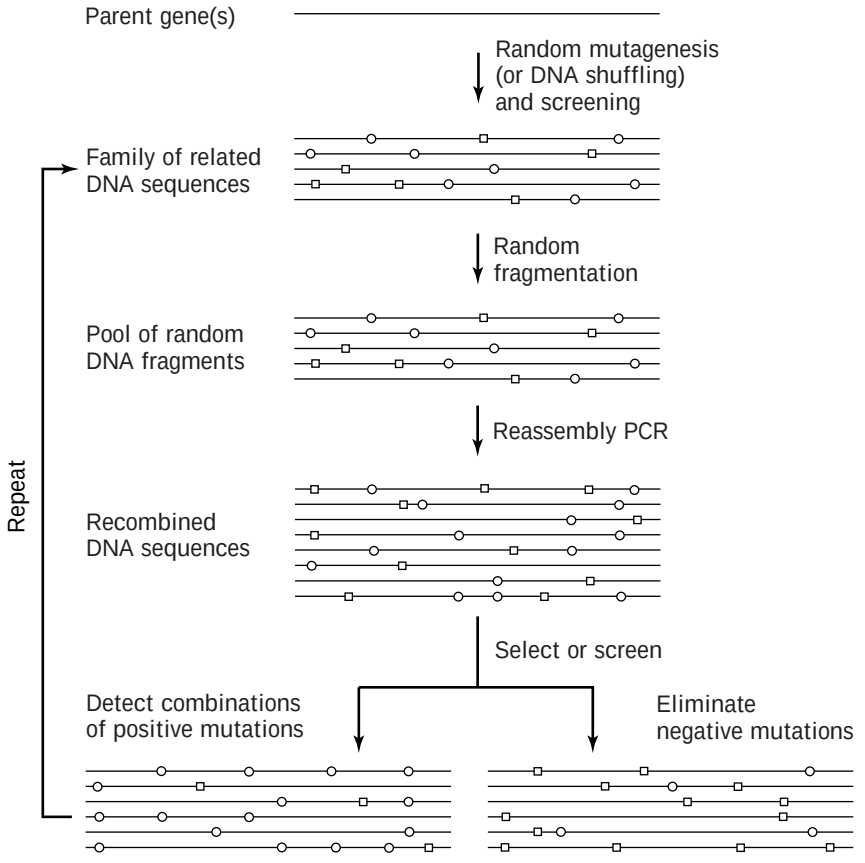


Figure 6. Recombination of parent genes by DNA shuffling. *Source*: Ref. 44.

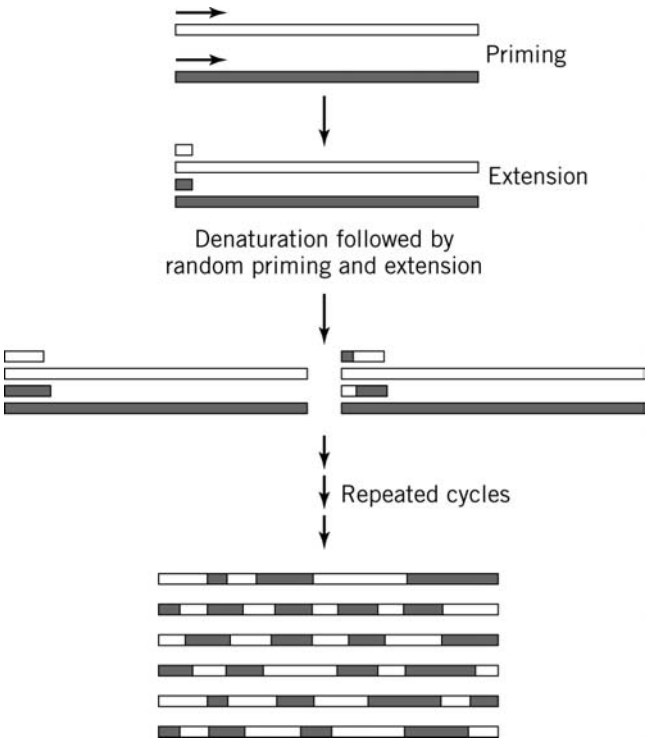


Figure 7. Recombination of two parent genes by StEP. *Source*: Ref. 34.

Mutant libraries are searched by screening or selection. In a genetic selection, the ability of an enzyme to perform the desired function is directly coupled to the survival and reproduction of the host organism. A straightforward example is selection for drug resistance: clones expressing an enzyme capable of degrading an antibiotic are selected for their ability to grow on antibiotic-containing plates (53,54). Another example is the selection of functional mutants of a human methyltransferase. Active mutants were selected based on their ability to protect alkyltransferase-deficient *E. coli* cells from a methylating agent (26). Such approaches are attractive in principle because they allow one to search larger libraries than does screening (10^6–10^7 for selection, as opposed to 10^4–10^5 for screening). Kast and Hilvert (55) recently reviewed genetic selections in directed evolution.

However, selections suffer serious drawbacks. The types of new properties that can be searched for are limited. For example, enzymes tolerate a number of environments that cannot sustain life (e.g., organic solvents). It becomes extremely difficult to tie the desired function to the survival or growth of an organism when the organism cannot survive. The main drawback, however, is the complexity of the organisms themselves and the surprising solutions that spontaneously arise during adaptation. Given our limited understanding of even simple, well-studied organisms such as *E. coli*, it is difficult to ensure that improving a particular enzyme is the *only* way for an organism to adapt to the selective pressure applied. One may go

through several rounds of selection only to find that the organism has found some alternate way to solve the problem.

Various in vitro selections have been proposed based on column binding or "panning" (56). In these methods, the proteins are displayed on the surface of a phage or a cell. Very large libraries ($>10^9$) can be searched based on the ability of the displayed protein to bind to a substrate immobilized on a column. The greatest limitation to this approach is that it is not generally applicable to screening for properties other than binding.

For many problems, and especially those of practical interest, libraries of variants must be screened rather than selected. A typical screening strategy involves the construction of an arrayed enzyme library and application of a rapid assay of the desired property. The screen can be more or less sensitive, depending on the willingness of the researcher to accept false positives (and to apply additional tests) (57).

Prerequisites for an effective screening strategy include the following:

1. The screen should be sufficiently rapid that large numbers of clones can be searched in a reasonable amount of time. It should be feasible to screen at least a few hundred, and preferably a few thousand, in a day. For obvious reasons, the screen should not require purification of the enzyme. A second-level assay can eliminate false positives.
2. Only rarely do very large improvements result from a single amino acid substitution. Directed evolution most often succeeds by accumulating a number of modest improvements over several generations. Therefore a screen must be accurate enough to allow modest improvements to be identified against the background variation.
3. The phenotype (the enzyme property that is being screened) must be physically coupled to the genotype (the DNA or RNA sequence coding for the enzyme). This is most often accomplished trivially by the fact that proteins are expressed in host organisms: the same cell expressing the enzyme contains the gene coding for its sequence. Coupling can also be accomplished by display on the surface of a virus containing the genetic information in the form of DNA or RNA (58). Recently developed in vitro translation systems such as ribosome display (59) or RNA–peptide fusions (60) link the enzyme directly to the genetic information in the form of RNA without using cells. Tawfik and Griffiths (61) have encapsulated the in vitro translation apparatus and genes within reverse micelles.

Simple visual screens are widely used when the function of interest can generate a visible signal. Various oxidases, for example, produce hydrogen peroxide as a byproduct of their oxidation reaction. As illustrated in Figure 8, coupling this to a second enzyme that is easily assayed colorimetrically (horseradish peroxidase) provides a visual screen for the oxidase activity. Clones secreting active proteases produce a zone of clearing or "halo," the size of which is proportional to the hydrolytic activity when the organism is grown on agar plates containing casein or skim milk proteins. This simple method has been used extensively in screening protease libraries (25,62).

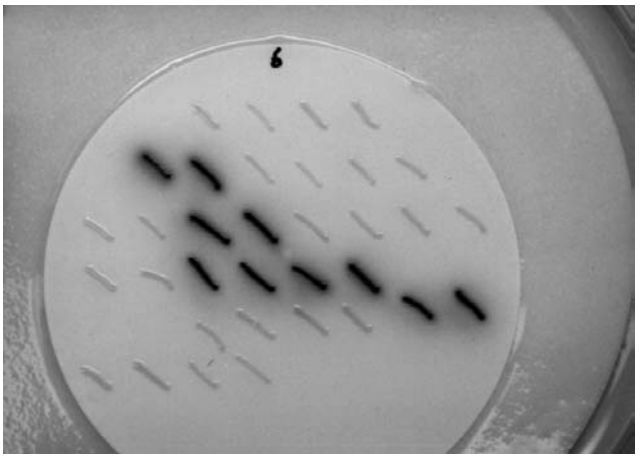


Figure 8. Example data from a visual screen for galactose oxidase activity toward D-galactose. Colonies are grown on agar plates and transferred to a membrane. Active colonies generate a purple color when incubated with 2,2′-azinobis(3-ethylbenzothiazline-6-sulfonic acid), horseradish peroxidase, and D-galactose. *Source*: M. Yagasaki, unpublished results, 1998.

Although visual screening based on color or halo formation is rapid and efficient, it is also nonquantitative and often relatively insensitive. Digital imaging spectroscopy has been developed to increase the sensitivity and throughput of filter- and agar-plate-based screens (63). However, the 96-well microtiter plate remains the standard format for automated, high-throughput screening. Screening automation and quantification of the results are highly desirable; 96- and 384-well plates appear to be the format most compatible with currently available robotic arms, liquid handling systems, and plate readers. Data collection and analysis are greatly facilitated by computerized data acquisition (computerized data are easily sorted into local enzyme fitness landscapes such as those shown in Fig. 5). Computerized data acquisition will certainly facilitate exploitation of the enormous amounts of information available in protein libraries. Many conventional enzyme assays can be readily converted into automated formats.

If activity against the desired substrate does not result in an easily detectable signal (such as a change in absorbance or fluorescence), it is sometimes possible to screen against a surrogate substrate that does. Naturally, this approach requires caution. It must be verified that activity against the replacement substrate reflects activity against the real substrate under the conditions of interest. A successful application of this approach is described in the example provided later of the evolution of a *p*-nitrobenzyl esterase (21).

Enzyme properties other than activity are often targeted for improvement. A convenient approach to screening for thermostable variants of irreversibly inactivated

enzymes is shown in Figure 9. Duplicate 96-well plates are made from a single master plate. One is incubated at high temperature for a fixed period. Clones in both are then assayed for activity, and the ratio of activity after heating to activity without heating provides a measure of enzyme stability at the temperature of incubation.

An attractive approach to screening large libraries is to couple functional complementation with screening. By requiring that at least some of the biological function of the protein is retained, functional complementation can greatly reduce the subsequent screening requirements. The initial selection asks an essentially binary question (is there any activity? yes or no), and only in the later screening stage are differences in the level of activity detected. Loeb and coworkers have applied this strategy to several enzymes of interest in gene therapy and other applications (8,26). However, the major limitation is finding or constructing an appropriate complementation system. Furthermore, retention of biological function may preclude acquisition of other functions.

Extracting Useful Information from the Results of Evolution

Two major types of information can be extracted from the results of directed evolution experiments. The first and most obvious is the information contained in the sequences of the evolved proteins. The second consists of the overall activity profiles, or local fitness landscapes, that can be generated by analysis of the entire screened library.

Clearly, we would like to identify the specific amino acid substitutions that are responsible for conferring new properties. In contrast to comparisons of evolutionarily related proteins found in nature, this identification can be accomplished relatively easily in directed evolution experiments. In nature, related enzymes that are separated by millions of years of evolution will generally have accumulated a large number of neutral mutations in addition to those that are responsible for changes in specific properties. In contrast, in laboratory evolution we strive to allow only functional mutations to survive and carry on to the next generation. A striking example of the difference between natural and laboratory evolution can be seen in a recent study in which subtilisin E was converted into a functional equivalent of the thermophilic enzyme thermitase (33). Subtilisin E differs from thermitase at 157 positions, but only eight substitutions were required to turn it into a thermitase-like enzyme, with an 18 °C increase in its temperature optimum and a 250-fold increase in its half life at 65 °C.

However, it is not always possible to identify functional mutations simply by examining the sequences of the evolved enzymes. In the experimental procedures for introducing random point mutations, only the *average* error rate can be controlled. Thus, some variants will accumulate multiple mutations in a single generation. If a given variant has multiple mutations, only one of which is functional, the remaining neutral mutations can "hitchhike" their way into subsequent generations and complicate analysis of the final evolved sequences. A method for rapidly identifying functional mutations by back-crossing with wild type has been described (64).

It is hoped that detailed structural analysis of evolved enzymes will provide insight into the underlying physical–chemical mechanisms responsible for protein adaptation for new functions or environments. However, the results of single point mutations can be very subtle, and the analysis is not straightforward even when the three-dimensional structure is known. For example, both subtilisin E and pNB esterase have been successfully evolved for activity in aqueous organic solvents, but functional mutations in

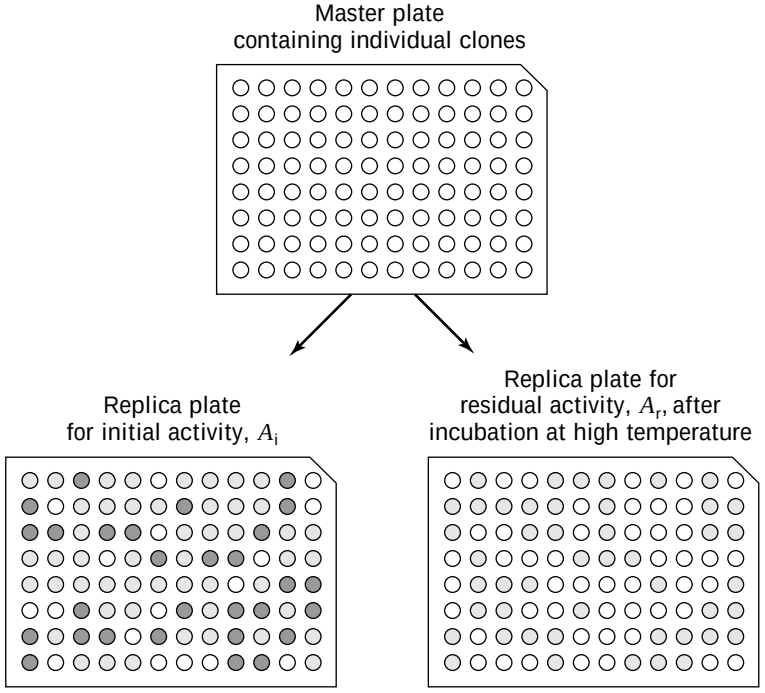


Figure 9. A rapid screen for enzyme thermostability based on catalytic activity before and after incubation at high temperature. *Source*: Ref. 34.

the evolved enzymes follow no discernible pattern: large residues are replaced by small, small by large, charged by uncharged, uncharged by charged (5,21,25). Further, mutations are found both close to and far from the enzyme active sites. Numerous studies now attest to the subtlety of the effects of amino acid substitutions on enzyme structure and properties (8,24,38,39).

The distribution of fitnesses (defined in terms of activity, stability, etc.) in a mutant library can provide useful information for designing evolutionary strategies. One way to access this information is by presenting the data in a local fitness landscape such as those shown in Figure 5. As mentioned previously, the mutation rate can be estimated from the fraction of inactive mutants. More generally, such a plot represents (at least for single- or double-mutant libraries) the distribution of fitness in the region of sequence space immediately surrounding the wild-type sequence. Most mutants will be equally or less fit than the wild type; a small number may be better. The fraction of improved mutants can be used to estimate the evolvability of the property of interest. When two or more properties are being evolved simultaneously, plots of one versus the other can be informative. Typically, single mutations that simultaneously improve two independent properties are extremely rare. However, individuals may be recombined to yield an enzyme with multiple improved properties (Fig. 10) (65).

EXAMPLES OF DIRECTED EVOLUTION APPLICATIONS

Evolution of Lipases with High Activity in Detergent Solutions

New lipase variants that give greatly improved residual lipid removal after one wash cycle have been developed by directed evolution (30). Lipases are used as detergent enzymes to remove lipid or fatty stains from clothes and other textiles. A drawback of the known detergent lipases is that they exert the best lipid-removing effect after more than one wash cycle, presumably because they are more active during a certain period of the drying process than during the wash itself. Therefore at least two wash cycles, separated by a sufficiently long drying period, are required to remove fatty stains.

An intense protein engineering effort involving both site-directed and random mutagenesis yielded a large number of variants of *Thermomyces lanuginosa* lipase that were strongly improved in multicycle wash performance, but not significantly improved in one-cycle wash tests. Random mutant libraries were screened using filter assays containing detergent and low calcium concentrations, to mimic wash conditions. Consecutive rounds of mutagenesis and screening with increasing amounts of detergent were performed.

New variants with a strong one-cycle wash effect were obtained by recombining 20 variants with the best performance in multicycle wash tests and screening. The recombination was performed using a simple in vivo method in *S. cerevisiae* in which PCR fragments from the 20 variants mixed with the opened vector are used to transform competent *S. cerevisiae* cells. Upon transformation, the vector and fragments recombine and shuffle the variants (14). Screening yielded seven variants with a strong one-cycle wash effect; the best of these is now commercially available.

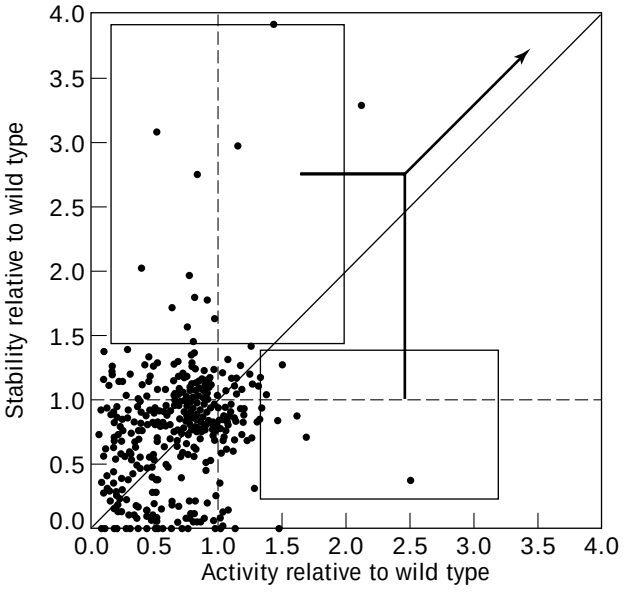


Figure 10. Activity and stability to hydrogen peroxide of a randomly mutagenized enzyme library. The individual populations shown in boxes can be evolved in parallel and recombined to obtain enzymes that are both more active and more stable. *Source*: Reproduced with permission from Ref. 65.

Evolution of an Esterase for Deprotection of *p*-Nitrobenzyl Esters and Enhanced Thermostability

By screening microbial cultures, scientists at Eli Lilly identified an enzyme in *Bacillus subtilis* that could remove the *p*-nitrobenzyl ester protecting group used during the large-scale synthesis of certain β-lactam antibiotics. However, the enzyme's relatively poor activity, particularly in the solvents required to solubilize the ester substrate, made it a poor competitor to the chemical catalyst for deprotection of a loracarbef synthetic intermediate. The natural function of *B. subtilis* pNB esterase is unknown.

Moore and Arnold were able to evolve highly active pNB esterases that also function well in mixed aqueous–organic solvents (21,22). A surrogate, chromogenic substrate (the antibiotic *p*-nitrophenyl ester) was used in a rapid screen to identify potential positives. Variants identified during the rapid screen could be verified during a second-level screen on the *p*-nitrobenzyl ester using HPLC. Four generations of PCR mutagenesis and two rounds of recombination by DNA shuffling yielded a clone with more than 100 times the total activity of wild type in 15–20% dimethylformamide (DMF) toward loracarbef-*p*-nitrobenzyl ester. The enzyme's catalytic efficiency increased more than 50-fold (Fig. 11). The total activity toward the screening substrate (loracarbef-*p*-nitrophenyl ester) increased more than 150-fold. Although the contributions of individual effective amino acid substitutions to enhanced activity were small (usually approximately two-fold increases in activity), the accumulation of multiple mutations over a num-

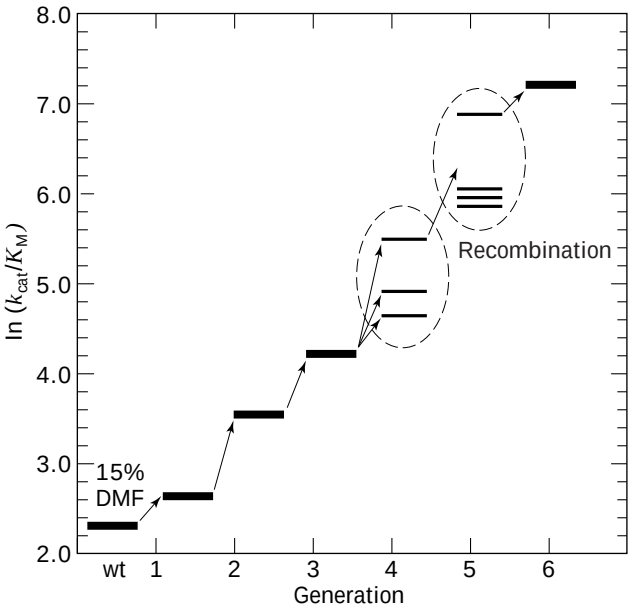


Figure 11. Evolutionary progression of catalytic efficiency of pNB esterase (toward LCN-pNB) in 15% DMF through four generations of random mutagenesis, followed by two rounds of recombination of circled populations. *Source*: Ref. 66.

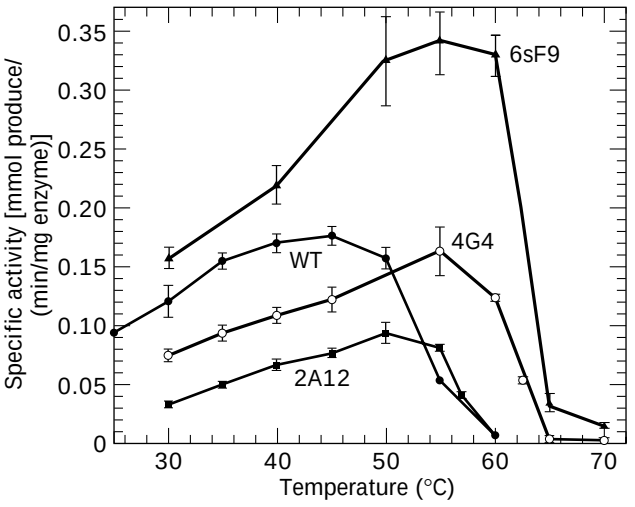


Figure 12. Activities of wild-type and evolved pNB esterases as a function of temperature. *Source*: Ref. 15 and A. Gershenson, unpublished results.

ber of generations significantly improved the biocatalyst for this nonnatural reaction. Moore and Arnold concluded that none of the mutations accumulated in the evolved enzyme are in direct contact with the substrate. Some are as far away as 20 Å. Limiting the search for possible solutions to residues that line the substrate binding site would have overlooked important beneficial mutations (and may not have succeeded).

Directed evolution has also significantly increased the thermostability of the *B. subtilis* pNB esterase (15). A simple screen was developed based on retention of catalytic activity after incubation at high temperature. Positives were then subjected to differential scanning calorimetry (DSC) to verify an increase in melting temperature (T_m) relative to the parent enzyme. Accumulating mutations over eight generations of random mutagenesis and DNA shuffling yielded an increase in T_m of 17 °C (15, A. Gershenson, unpublished results). This large increase in thermostability is equivalent to the difference between proteins from mesophiles and many thermophilic organisms. Furthermore, because only those variants that retained their catalytic activity as well were chosen, the increased thermostability was accompanied by a very significant increase in enzyme activity at elevated temperatures (Fig. 12). Reflecting the fact that the thermostability screen involved activity toward *p*-nitrophenyl acetate, the resulting enzymes are highly active toward this substrate and less active toward the antibiotic substrate and *p*-nitrobenzyl esters. The most thermostable pNB esterase variant has 13 amino acid substitutions (out of 490).

Evolution of Herpes Simplex Virus Thymidine Kinase for Gene Therapy

The Loeb group has pioneered the use of directed evolution for developing improved enzymes for cancer gene therapy. Here, greater activity toward nonnatural substrates is desired, with the goal of selectively protecting bone marrow cells (26) or sensitizing tumor cells (8) to toxic agents. This group consistently directs intense mutagenesis of a limited region of the gene and genetic selection to identify the sequences that are still biologically functional. Screening this much smaller subset of sequences identifies those with the desired properties (altered substrate specificity, ability to protect or sensitize cells).

In one study (8), directed evolution was used to create herpes simplex virus thymidine kinase (HSV TK) variants that enhance cell killing by the guanine analogs gancyclovir and acyclovir, which terminate DNA synthesis and prevent viral and host cell replication. Mutant HSV TKs that selectively phosphorylate gancyclovir and/or acyclovir were obtained after complete randomization of a limited set of amino acid positions (six) adjacent to the putative nucleoside binding site and previously identified to tolerate some substitution. Of the approximately 10^6 *E. coli* transformants, 426 expressed active TK, as determined by genetic complementation. These were then assayed for their ability to enhance the sensitivity of the *E. coli* to the guanine analogs. Hamster cells transfected with one of the improved variants were more than 43 times more sensitive to gancyclovir and 20 times more sensitive to acyclovir than cells expressing the wild-type HSV TK. This mutant contained four amino acid substitutions, all within the TK active site pocket.

Evolution of Biphenyl Dioxygenases for Bioremediation of PCBs

Biphenyl dioxygenases (BPDO) are responsible for the initial oxidation of polychlorinated biphenyls (PCBs) during their biodegradation by various organisms. In an effort to create BPDOs with enhanced ability to oxidize a wider range of PCB congeners, Furukawa and coworkers used DNA shuffling to create a chimeric library from two genes

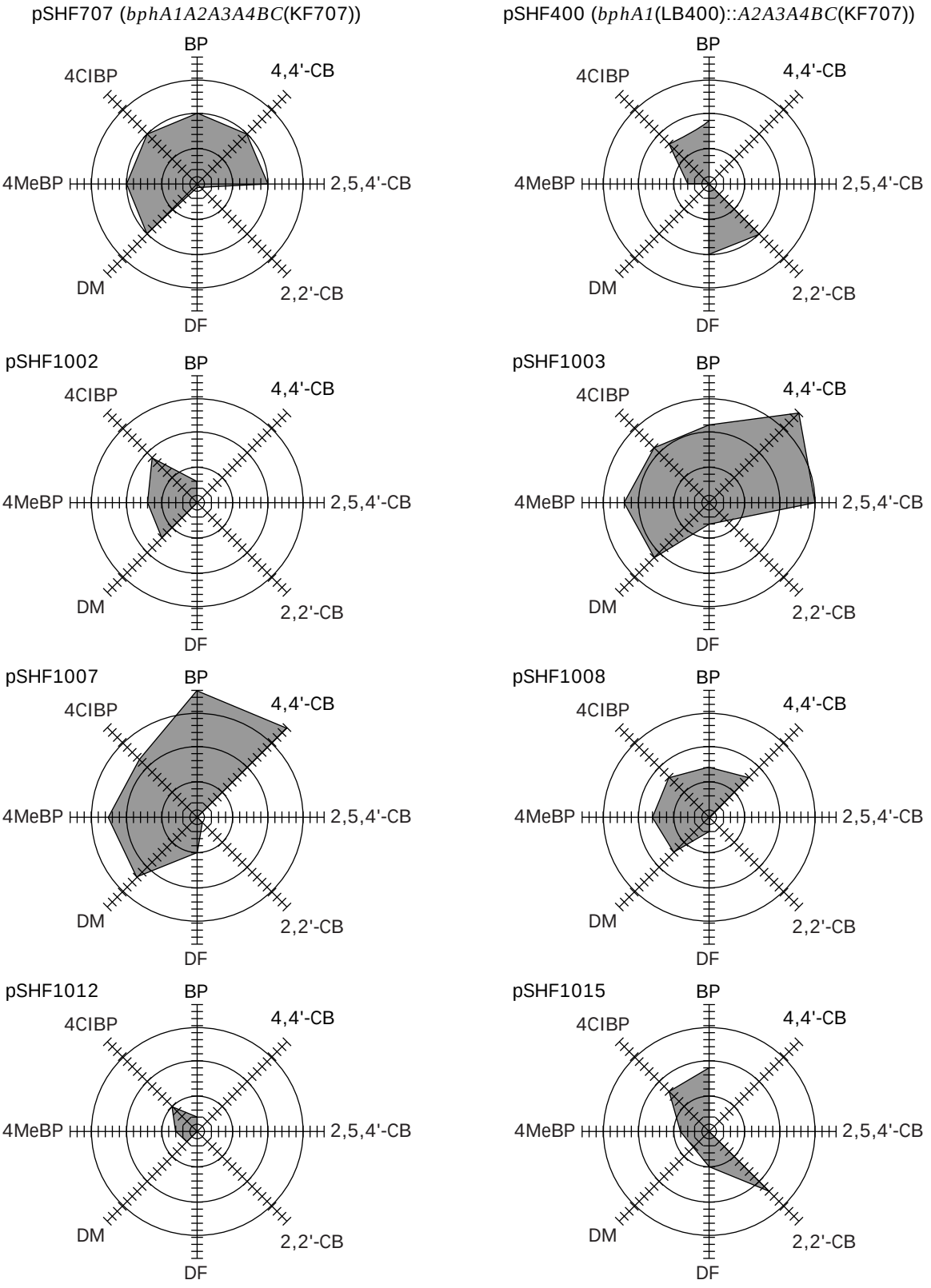


Figure 13. Formation of *meta*-cleavage yellow products from a variety of biphenyl compounds by *E. coli* expressing chimeric biphenyl dioxygenases. BP, biphenyl; 4CIBP, 4-chlorobiphenyl; 2.2′-CB, 2,2′-dichlorobiphenyl; 4,4′-CB, 4,4′-dichlorobiphenyl; 2,5,4′-CB, 2,5,4′-trichlorobiphenyl; MeBP, 4-methylbiphenyl; DM, diphenylmethane; DF, dibenzofuran. The relative values of yellow products formed from each compound are plotted for *E. coli* strains carrying the evolved enzymes. *Source*: Reprinted with permission from Ref. 38.

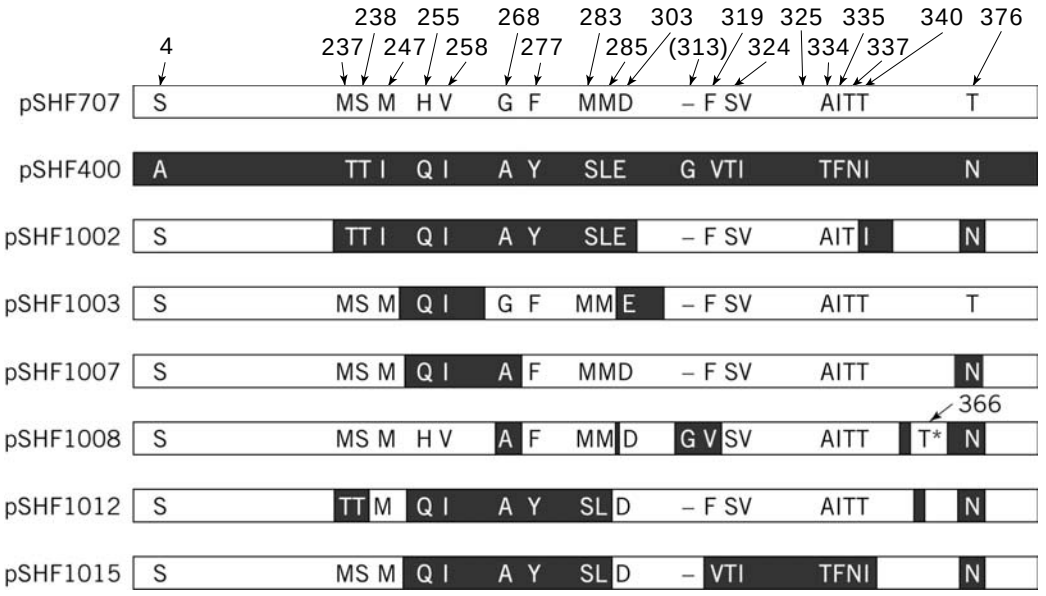


Figure 14. Sequence analysis of the shuffled *bphA1* genes. Twenty different amino acids between KF707-BphA1 and LB400-BphA1 are shown. *, Indicates the mutational change (I366T) that occurred during PCR; –, indicates an amino acid lacking in KF707 relative to the LB400 sequence. *Source*: Reprinted with permission from Ref. 38.

for the terminal dioxygenases of BPDO (*bphA1*) from *Pseudomonas pseudoalcaligenes* KF707 and *Burkholderia cepacia* LB400 (38). A high-fidelity DNA shuffling protocol (48) was found to increase the frequency of clones expressing active PBDO. The two BphA1 protein sequences are extremely similar, differing only at 20 amino acids. The genes are 96.4% identical.

E. coli containing the chimeric genes were screened for their ability to produce yellow *meta*-cleavage products from a variety of biphenyl compounds in order to identify functional chimeras with enhanced activities and altered activity profiles. Enzymes with greater activity toward various biphenyl compounds than either parent were identified (Fig. 13), as were enzymes with relaxed specificity and enzymes active toward smaller aromatic hydrocarbons such as benzene and toluene. These latter compounds are poor substrates for the two parent enzymes.

Sequence analysis of six shuffled functional *bphA1* genes demonstrated how the parent sequences recombined to yield these new enzymes (Fig. 14). A very low level of point mutagenesis was observed. The specific roles of the amino acids in determining substrate specificity are difficult to pinpoint in the absence of a three-dimensional structure of the enzyme. However, the results are generally in agreement with previous studies of BPDO chimeras and site-directed mutagenesis.

ACKNOWLEDGMENTS

F.H.A. wishes to thank the many talented undergraduate, graduate, and postdoctoral students who have contributed to directed evolution research at Caltech. The financial support of the Office of Naval Research, the Department of Energy, and the Army Research Office is gratefully acknowledged.

BIBLIOGRAPHY

1. L. Holm and C. Sander, *Proteins: Struct. Funct. Genet.* **28**, 72–82 (1997).
2. F.H. Arnold, *Acct. Chem. Res.* **31**, 125–131 (1998).
3. M. Eigen, *Chemica Scripta* **26B**, 13–26 (1986).
4. S.A. Kauffman, *The Origins of Order*, Oxford University Press, Oxford, 1993.
5. K. Chen and F.H. Arnold, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5618–5622 (1993).
6. W.P.C. Stemmer, *Nature* **370**, 389–391 (1994).
7. U.S. Pat. 5,605,793 (February 25, 1997) W.P.C. Stemmer (to Affymax Technologies N.V.).
8. M.E. Black, T.G. Newcomb, H.M.P. Wilson, and L.A. Loeb, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 3525–3529 (1996).
9. A. Crameri, S.A. Raillard, E. Bermudez, and W.P.C. Stemmer, *Nature* **391**, 288–291 (1998).
10. F.H. Arnold, *Nat. Biotechnol.* **16**, 617–618 (1998).
11. D. Pompon and A. Nicolas, *Gene* **83**, 15–24 (1989).
12. U.S. Pat. 5,093,257 (March 3, 1992) G.L. Gray (to Genencor International, Inc.).
13. M.D. van Kampen, N. Dekker, M. Egmond, and H. Verheij, *Biochemistry* **37**, 3459–3466 (1998).
14. PCT Pat. WO 97/07205 (August 12, 1996), J.S. Okkels (to Novo Nordisk A/S).
15. L. Giver, A. Gershenson, P.O. Freskgard, and F.H. Arnold, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12809–12813 (1998).
16. S.A. Benner, *Chem Rev.* **89**, 789–806 (1989).
17. G. MacBeath, P. Kast, and D. Hilvert, *Science* **279**, 1958–1961 (1998).
18. H. Liao, T. McKenzie, and R. Hageman, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 576–580 (1986).

19. U.S. Pat. 5,316,935 (May 31, 1994), F.H. Arnold and K. Chen (to California Institute of Technology).
20. S.L. Strausberg, P.A. Alexander, D.T. Gallagher, G.L. Gilliland, B.L. Barnett, and P.N. Bryan, *Biotechnology* **13**, 669–673 (1995).
21. J.C. Moore and F.H. Arnold, *Nat. Biotechnol.* **14**, 458–467 (1996).
22. J.C. Moore, H.M. Jin, O. Kuchner, and F.H. Arnold, *J. Mol. Biol.* **272**, 336–347 (1997).
23. U.S. Pat. 5,741,691 (April 21, 1998), F.H. Arnold and J.C. Moore (to California Institute of Technology).
24. J-H. Zhang, G. Dawes, and W.P.C. Stemmer, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4504–4509 (1997).
25. L. You and F.H. Arnold, *Protein Eng.* **9**, 77–83 (1996).
26. F.C. Christians and L.A. Loeb, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6124–6128 (1996).
27. A. Crameri, G. Dawes, E. Rodrigues, Jr., S. Silver, and W.P.C. Stemmer, *Nat. Biotechnol.* **15**, 436–438 (1997).
28. D.R. Liu, T.J. Magliery, M. Pasternk, and P.G. Schultz, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10092–10097 (1997).
29. T. Yano, S. Oue, and H. Kagamiyama, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5511–5515 (1997).
30. PCT Pat. WO 97/07202 (February 27, 1997), J.S. Okkels, M. Thellersen, D.A. Petersen, A. Svendsen, S.A. Patkar, and K. Borch (to Novo Nordisk A/S).
31. M.T. Reetz, A. Zonta, K. Schimossek, K. Liebeton, and K.E. Jaeger, *Angew. Chem. Int. Ed. Engl.* **36**, 2830–2832 (1997).
32. U.T. Bornscheuer, J. Altenbuchner, H.H. Meyer, *Biotechnol. Bioeng.* **58**, 554–555 (1998).
33. H. Zhao and F.H. Arnold, *Protein Eng.* **12**, 47–53 (1999).
34. H. Zhao, L. Giver, Z. Shao, J.A. Affholter, and F.H. Arnold, *Nat. Biotechnol.* **16**, 258–235 (1998).
35. D. Naki, C. Paech, G. Ganshaw, and V. Schellenberger, *Appl. Microbiol. Biotechnol.* **49**, 290–294 (1998).
36. S. Taguchi, A. Ozaki, and H. Momose, *Appl. Environ. Microbiol.* **64**, 492–495 (1998).
37. S. Akanuma, A. Yamagushi, N. Tanaka, and T. Ooshima, *Protein Sci.* **7**, 698–705 (1998).
38. T. Kumamaru, H. Suenaga, M. Mitsuoka, T. Watanabe, and K. Furukawa, *Nat. Biotechnol.* **16**, 663–666 (1998).
39. F. Buchholz, P.O. Angrand, and A.F. Steward, *Nat. Biotechnol.* **16**, 657–662 (1998).
40. Z. Shao, M. Callahan, and F.H. Arnold, *Appl. Environ. Microbiol.* (in press).
41. Z. Lin, T. Thorsen, and F.H. Arnold, *Biotechnol. Prog.* (in press).
42. J. Reidhaar-Olson and R. Sauer, *Science* **241**, 53–57 (1988).
43. R. Cadwell, and G. Joyce, *PCR Methods Appl.* **3**, S136–S140 (1994).
44. W.P.C. Stemmer, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10747–10751 (1994).
45. X. Shao, H. Zhao, L. Giver, and F.H. Arnold, *Nucleic Acids Res.* **26**, 681–683 (1998).
46. H. Zhao, L. Giver, X. Shao, J. Affholter, and F.H. Arnold, *Nat. Biotechnol.* **16**, 258–261 (1998).
47. F.H. Arnold, *Chem. Eng. Sci.* **51**, 5091–5102 (1996).
48. H. Zhao and F.H. Arnold, *Nucleic Acids Res.* **25**, 1307–1308 (1997).
49. D. Leung, E. Chen, and D. Goeddel, *Technique* **1**, 11–15 (1989).
50. S. Shafikhani, R. Siegel, E. Ferrari, and V. Schellenberger, *Biotechniques* **23**, 304–310 (1997).
51. H. Zhao, J. Moore, and F.H. Arnold, in *ASM Manual of Industrial Microbiology and Biotechnology*, ASM Press, Washington D.C., 1999, pp. 597–604.
52. I.A. Lorimer and I. Pastsn, *Nucleic Acids Res.* **23**, 3067–3068 (1995).
53. J. Petrosino and T. Palzkill, *J. Bacteriol.* **178**, 1821–1828 (1996).
54. W. Huang, J. Petrosino, M. Hirsch, P. Shenkin, and Y. Palzkill, *J. Mol. Biol.* **258**, 688–703 (1996).
55. P. Kast and D. Hilvert, *Curr. Opin. Struct. Biol.* **7**, 470–479 (1997).
56. C. Wang, Q. Yang, and C. Craik, *Methods Enzymol.* **267**, 52–68 (1996).
57. H. Zhao and F.H. Arnold, *Curr. Opin. Struct. Biol.* **7**, 480–485 (1996).
58. K. O'Neil and R. Hoess, *Curr. Opin. Struct. Biol.* **5**, 443–449 (1995).
59. J. Hanes and A. Pluckthun, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4937–4949 (1997).
60. R. Roberts and J. Szostak, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12297–12302 (1997).
61. D. Tawfik and A. Griffiths, *Nat. Biotechnol.* **16**, 652–656 (1998).
62. S. Sidhu and T. Borgford, *J. Mol. Biol.* **257**, 233–245 (1996).
63. D. Youvan, E. Goldman, S. Elgrave, and M. Yang, *Methods Enzymol.* **246**, 732–749 (1995).
64. H. Zhao, and F.H. Arnold, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7997–8000 (1997).
65. O. Kuchner and F.H. Arnold, *TIBTECH* **15**, 523–530 (1997).
66. F.H. Arnold and J. Moore, *Adv. Biochem. Eng. Biotechnol.* **58**, 2–14 (1997).

See also ENZYMES, BAKING, BREAD MAKING; ENZYMES, DETERGENT; GOOD MANUFACTURING PRACTICE (GMP) AND GOOD INDUSTRIAL; PROTEIN EXPRESSION, SOLUBLE; PROTEIN SECRETION, *SACCHAROMYCES CEREVISIAE*.

ENZYMES, EXTREMELY THERMOSTABLE

PAULA M. HICKS
North Carolina State University
Raleigh, North Carolina

MICHAEL W. W. ADAMS
University of Georgia
Athens, Georgia

ROBERT M. KELLY
North Carolina State University
Raleigh, North Carolina

KEY WORDS

Archaea
Biomass
Enzymes
Hyperthermophilic
Proteases
Protein stability
Thermostable

OUTLINE

INTRODUCTION

The influence of temperature on the structure and function of proteins has been extensively studied. However, it is still not possible to determine in any definitive sense which factors underlie the vast differences in thermostability and thermoactivity between seemingly similar proteins (1,2). The discovery of hyperthermophilic microorganisms (those with growth temperatures of 90 °C and above) and the subsequent purification and characterization of their extremely thermostable and thermoactive enzymes have underscored the complex nature of protein stability. It is clear now that differences of 100 °C or more may separate the temperature optima of enzymes carrying out identical biocatalytic functions. Information available on amino acid sequences of enzymes with widely varying temperature optima, and even three-dimensional structural data, may not provide much insight into how seemingly similar enzymes can respond to thermal energy so differently (3). Thus, enhancing the thermostability of a given enzyme to a significant extent remains an elusive objective left to fortuitous improvements through random mutagenesis, or to extensive screening efforts from natural high temperature biotopes (4).

In cellular environments, protein stability is a property that has important physiological implications. Protein architecture must be consistent with in vivo requirements. For example, the free energy of stabilization of an enzyme (ΔG), a balance between large but opposing entropic ($T\Delta S$) and enthalpic forces (ΔH), is equivalent to only a few weak bonds (5), presumably so that it can be regulated in response to changing cellular conditions. Once an enzyme has performed a needed cellular function that is no longer necessary or becomes detrimental, it may denature, be deattenuated through a regulatory response, or be proteolytically inactivated.

Although enzymes are desirable in industry because of their specificity and availability, in order to be useful for commercial purposes, some degree of stability is required to make economic sense. In fact, proteins may need to not only withstand thermal forces, but also be resistant to nonaqueous solvents, high ionic strengths, detergents, and extremes of pH as well. In addition to stability, it is important that the enzyme be functional in the specific extreme environment in which it is to be used. Thus, a highly thermostable and thermoactive enzyme may not be useful in applications at very low temperatures, even though it may be very stable under these conditions. Clearly, the maintenance of an enzyme's structural integrity and its catalytic efficiency need both be considered for possible application of these biocatalysts in an industrial setting.

Even though site-directed and random mutagenesis have yet to become a routine approach for conferring thermostability on an enzyme, nature has fortunately provided a source of thermostable biocatalysts in extremely thermophilic microorganisms (optimal growth temperatures above 75 °C). An ever-increasing list of such enzymes shows that most significant activities from mesophiles (optimal growth around 37 °C) have high-temperature, and presumably thermostable, counterparts. The biochemical and biophysical characteristics of many of these enzymes are currently being examined for both scientific and technological purposes (4,6). In this article, the methodology being used to discover extremely thermostable enzymes will be discussed in addition to information on their function and use. It is expected that the extraordinary stability of these enzymes will make them candidates for replacing many currently used biocatalysts in addition to opening up many new applications.

EXTREMELY THERMOSTABLE ENZYME DISCOVERY

The methodology for identifying enzymes with technological importance has gone through considerable changes in recent years with the ever-expanding use of molecular biological tools. Concomitantly, the search for enzymes from microorganisms in extreme environments has also expanded and has made use of new discovery techniques (7). Gone are the days when industrial enzymes were exclusively mined from microorganisms isolated from soil samples and grown in pure culture. Now, only the gene encoding an enzyme of interest is, in principle, necessary if it can be overexpressed in an appropriate host. Another development that is sure to impact enzyme discovery is the increasing availability of genome sequences for a variety of mesophilic and thermophilic microorganisms (8). Various approaches for the identification of extremely thermostable enzymes are discussed in the following sections and summarized in Figure 1.

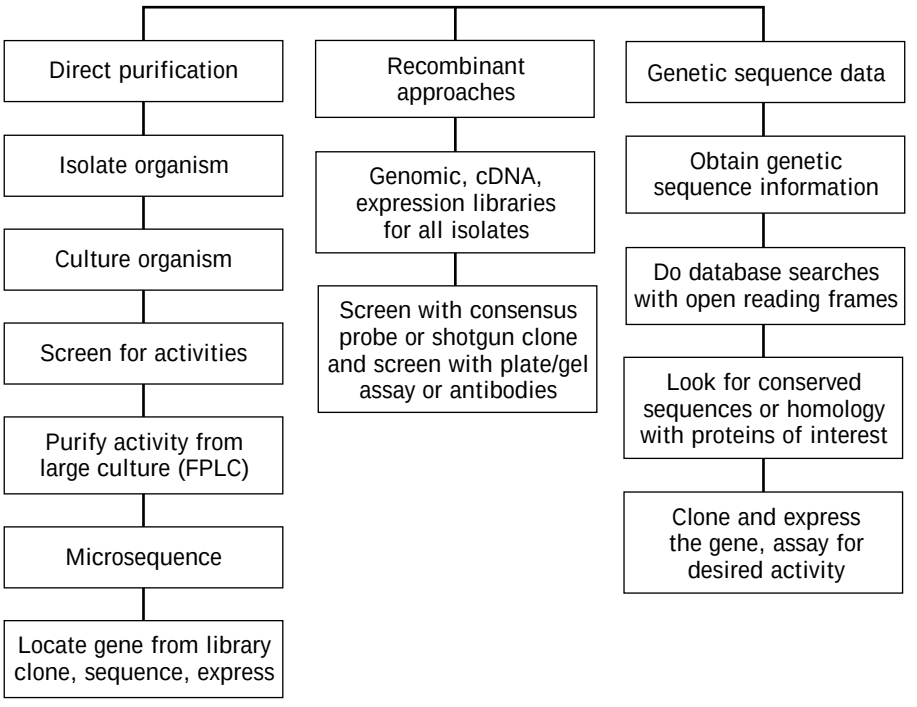


Figure 1. Screening for extremely thermostabile enzymes.

Direct Purification From Extreme Thermophile Biomass

Expeditions to geothermal habitats, as accessible as hot springs in Yellowstone National Park (9) and as inaccessible as deep sea geothermal vents (10), have led to the isolation of an array of high-temperature microorganisms covering a diversity of physiological types (11). Typically, isolates are obtained through serial dilution techniques because high temperatures preclude cultivation on solid media in many cases. Most microorganisms classified as extreme thermophiles ($T_{opt} \geq 75$ °C) are anaerobic, which is not surprising considering the low solubility of O_2 at elevated temperatures. Notable exceptions are members of the order Sulfolobales, many of which are thermoacidophiles (11). A subset of the extreme thermophiles, the hyperthermophiles (capable of growth at 90 °C or higher) are almost exclusively anaerobic, and many reduce sulfur to sulfide during growth (11). Additional details on these microorganisms can be found elsewhere in this volume (THERMOPHILIC MICROORGANISMS).

Cultivation of extreme thermophiles on a scale sufficient to provide biomass for enzyme purification efforts presents many challenges (12,13). In short, lack of knowledge of microbial physiology, low biomass yields, generation of copious amounts of hydrogen sulfide, explosive substrate and product gases, salt-laden media, anaerobic conditions, and cell sensitivity to shear are all potential obstacles. A few heterotrophic hyperthermophiles, such as *Pyrococcus furiosus* (14), *Thermococcus litoralis* (15), and *Thermotoga maritima* (16), can be cultivated without sulfur and have, thus, been the primary focus of direct purification efforts. These organisms can be grown to biomass yields of 1 kg (wet) or higher in 600-L fermentation systems (17). From such efforts, more than 20 enzymes have been purified from the biomass generated from a single fermentation (13).

Once sufficient biomass has been generated, enzyme purification can proceed along conventional lines using multistep protocols involving various types of liquid chromatography. Cells can be broken in a variety of ways including freeze-thaw cycles, French pressing, or sonification. Because many extreme thermophiles are members of the domain archaea, which have unusual cell membrane features, including a lack of peptidoglycan, lysozyme treatments are typically not effective. Most types of prep-scale chromatography have been used in the purification of these enzymes. These include ion exchange, hydrophobic interaction, gel filtration, and affinity methods. The extreme thermostability of these enzymes does provide a convenience: Purification can be carried out at ambient temperatures rather than in refrigerated environments. On the other hand, thermostability does present some difficulties in assessing the homogeneity of products of the purification process. Conventional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) can be misleading because extremely thermostable enzymes may not be sufficiently denatured in this process. The result is that both the native and partially denatured forms of multimeric proteins may appear on SDS-PAGE, making assessment of M_r (relative molecular mass) and homogeneity difficult. Unfortunately, extended periods of heating to ensure complete denaturation may also lead to covalent modifications of labile amino acid residues.

Because it is economically impractical to produce large amounts of extremely thermostable enzymes in their native hosts, recombinant approaches must be used. The purified enzyme may be N-terminally or C-terminally (or both) sequenced, and conventional methods for identifying the corresponding gene in genomic preparations by nucleic acid probe hybridization or polymerase chain reaction (PCR) techniques can be used. Overexpression of genes en-

coding extremely thermophilic enzymes is still in its earliest stages, and most efforts to date have used a variety of vectors in *Escherichia coli* (4). Although many efforts have yielded a recombinant enzyme seemingly identical to the native at low expression levels (18,19), it is not known whether existing expression systems will be suitable to produce the large amounts of enzyme needed for certain industrial uses. This is a critical issue for the development of applications making use of these enzymes.

Expression Cloning

Although not specific to extremely thermophilic enzymes, molecular biology has provided an array of techniques that circumvent the tedious process of obtaining biomass from large fermentations done in pure culture. If an assay is available for a particular enzyme activity, expression libraries containing DNA from an organism of interest may be screened to isolate the gene encoding this activity. This assumes that the gene manifests itself as an expression product. This same approach has been used to mine enzymes directly from environmental samples for which no effort is made to sort out individual organisms in pure culture (7). For extreme thermophiles, the intrinsic thermostability of the enzyme made as the expression product facilitates partial purification. A heating step will serve to denature the less-stable enzymes produced by a mesophilic host so that the thermostable enzyme is enriched in the resulting mixture. Unfortunately, expression cloning will not always work because some genes may not be expressed as an active enzyme or a suitable screening assay may not be easily used. However, this method can be a particularly attractive alternative to large-scale production of extreme thermophile biomass and can also be used as the basis for high throughput screening and directed evolution (7).

Sequence Analysis of Genomic DNA

Recent efforts to sequence entire genomes for microorganisms are sure to have an impact on the enzyme discovery process (8). Because many of the initial microbial sequencing projects involve extreme thermophiles, this bank of information can be used to identity putative proteins of industrial importance. The *Methanococcus jannaschii* genome has been sequenced (20), and many other extreme thermophiles are currently being sequenced, such as *Archaeoglobus fulgidus*, *Aquifex aeolicus*, *Methanobacterium thermoautotrophicum*, *Pyrobaculum aerophilum*, *Pyrococcus furiosus*, *Pyrococcus horikoshi*, *Sulfolobus solfataricus*, and *Thermotoga maritima* (21). Computer databases can be searched using amino acid or nucleotide sequences corresponding to known motifs, and resulting homologous sequences can be the basis for cloning the gene of interest. At this time, however, 50% or more of the genome sequence obtained from specific microorganisms contain open reading frames (ORFs) that are not attributable to a protein whose function is known. Furthermore, there are relatively little sequence data available for many hydrolytic enzyme types of potential commercial significance so that identification of homologous versions of certain enzymes can be an imprecise exercise. Nevertheless, as entire genome sequence data expand, this approach should not be ruled out because it may be difficult to purify a particular enzyme directly from the native source.

ENZYMES FROM EXTREME THERMOPHILES

General Characteristics

Enzymes of all known classes have been purified from extremely thermophilic microorganisms and characterized biochemically. Many of these proteins are intrinsically thermostable (Table 1). In many cases, the genes encoding these enzymes have been sequenced and in some cases expressed in active form in a foreign host, usually *E. coli* (Table 2). With the availability of genome sequence data (20), we also now know that these microorganisms contain thermostable counterparts to many proteins previously characterized from mesophilic sources, many of which are linked to central metabolic pathways (162). Also, because most existing enzyme applications make use of hydrolytic enzymes, these enzyme types have been sought in extreme thermophiles (163–165).

Enzymes from Extreme Thermophiles Involved in Intermediate Metabolism

Because of the evolutionary placement of extreme thermophiles, based on 16S rRNA phylogeny (8,166), interest in the origins of metabolic pathways has led to the isolation and characterization of enzymes directly involved in intermediate metabolism and bioenergetics (162). These include a number of oxidoreductases (167), hydrogenases (71,72,74), dehydrogenases (60–65), and redox proteins (167). Many such enzymes from the most studied hyperthermophile *P. furiosus* (14) appear to be intrinsically thermostable and function in vitro at the organism's optimal growth temperature of 98 to 100 °C (4). The study of these enzymes in vitro presents certain challenges because assays to detect activity may involve thermally labile substrates, anaerobic conditions, and coenzymes. For example, even though *P. furiosus* contains several dehydrogenases, these catalyze reactions involving NADH as a cofactor, which itself is not very thermostable. This organism is also known to contain a modified Emden-Myerhoff-Parnos pathway for glycolysis, which implies that glucose is processed at high temperatures through several intermediate steps (168). The mechanisms by which simple sugars, which may be sensitive to thermal modifications, can be processed at high temperatures are not known.

Proteases from Extreme Thermophiles

Proteolysis (hydrolysis of peptide bonds) is essential to cellular function and is implicated in nutrition, heat shock response, immune response, and certain housekeeping roles (169). Historically, proteases have found important industrial uses in detergents, leather bating, food processing, and much more, collectively comprising the largest volume use of commercial enzymes (170). But in the past 10 years, these enzymes have also been shown to be active in organic solvents for peptide synthesis and are often selective enough to produce pharmaceuticals (171,172). Recently, protease inhibitors have received much attention in

medical applications for combatting the HIV virus and some cancerous tumors (173,174).

There are many potential benefits to be derived from the use of proteases from extreme thermophiles. If mesophilic proteins are to be hydrolyzed, it is very likely that at elevated temperatures their denatured forms will facilitate access of thermostable proteases to specific peptide bonds. Also, at higher temperatures, many substrates are more soluble, solution viscosity is reduced, and diffusion rates are higher, which can lead to higher reaction rates (175). Because these enzymes appear to also be more stable to denaturing forces other than heat, proteases from extreme thermophiles should be amenable to use in organic solvents for synthetic reactions.

Many proteases have been purified and characterized from extreme thermophiles (163). From an evolutionary viewpoint, these proteases are of interest because they presumably predate those found in less thermophilic cells and organisms in terms of structure and function. Some of these proteases appear to be novel with no previously known mesophilic counterparts, whereas others represent more stable forms of known proteases.

Proteolytic activities produced by the hyperthermophilic archaeon *P. furiosus* (14) have been the most extensively studied among the extreme thermophiles. This organism appears to produce a number of proteases (176), which is consistent with its excellent growth on peptide-based media (177). The properties of several proteases from this organism are discussed below.

***P. furiosus* Protease I (PfpI).** *P. furiosus* protease I (PfpI) is a major intracellular proteases in *P. furiosus*. Previously, PfpI was named S66 because, after heating at 95 °C in 1% SDS for 24 hours, it migrates to a molecular mass corresponding to 66 kDa on a denaturing gel (25). It was subsequently found to occur in at least two active forms (hexamer and trimer composed of 18.8-kDa subunits) in vitro (26) and appears to have at least one active larger form (Hicks and Kelly, unpublished data). The calculated pI (from amino acid sequence data) of the 18.8-kDa subunit is 6.1, whereas the experimentally determined pIs of the hexamer and trimer are 6.1 and 3.7 to 3.9, respectively. If all asparagine and glutamine residues are changed to their corresponding acidic residues, the pI of the monomer is calculated to be 4.8. Thus, at high temperatures, it is possible that the protein is deamidated, especially in its trimeric form, explaining in part the difference from the measured and predicted pI values (26). Upon long exposures to heat (24 to 48 hours), the predominant band on an overlay gel containing gelatin is in fact the smaller form (26). PfpI is specific toward basic and bulky, hydrophobic P1 amino acid residues in peptide substrates (26) and most active toward *N*-succinyl-alanine-alanine-phenylalanine-7-amido-4-methylcoumarin (AAF-MCA) (26). Large proteins, such as azocasein and gelatin, are also degraded by PfpI, although specific cleavage sites in these proteins have not yet been identified. The gene for PfpI (*pfpI*) has been expressed in *E. coli* using the T7 promoter and infection with λ phage CE6 (107). Vector constructs are often unstable, and toxicity has been a problem in producing recombinant PfpI. Expression levels were highest (1 mg per 100 mL culture) when a histidine tag was used to make a fusion protein, which could then purified using an affinity column (107). The amino acid sequence of PfpI was found to be highly homologous to putative proteins of unknown function contained in representatives of all three domains of life (e.g., *E. coli* (bacteria), *M. jannaschii* (archaea), *Arabidopsis thaliana* (eukarya)) (178). However, the role of this protease in *P. furiosus* or in other cells and organisms is not clear at this time.

Pyrolysin. Pyrolysin is a cell-envelope-associated protease that can be purified from *P. furiosus* membrane fractions (27,179). Two versions of this enzyme, one based on a 150-kDa subunit and one that was apparently cleaved by carboxy-terminal autoproteolysis, have been reported. The *pls* gene encoding the subunit has been cloned and sequenced (27). The gene corresponds to a 4194-bp open reading frame, containing a prepro sequence, whereas the mature protein is encoded by a 1249-bp region. The N-terminus shows high homology to subtilisin-like serine proteases, whereas multiple glycosylation sites are proposed at the carboxy terminus. Pyrolysin shows highest homology at the amino acid level with eucaryl tripeptidyl peptidases (TPP), although it has different cleavage specificities; TPPs are typically exopeptidases (27). Pyrolysin appears to be a version of the surface protein assembly (designated tetrabrachion) identified in another hyperthermophile, *Staphylothermus marinus*, which contains a protease with trypsin- and chymotrypsin-like activities (28).

Archaeal Proteasome. Several thermophilic archaea contain a version of the eukaryotic proteasome (29,33,180). *P. furiosus* produces a version of the archaeal proteasome that appears to be composed of α (25 kDa) and β (22 kDa) subunits with an overall M_r of 640 kDa ($\alpha_{14}\beta_{14}$) (33). The *P. furiosus* proteasome has a temperature optimum of 95 °C, and preliminary biochemical analysis showed that methionine was preferred in the P1 position for hydrolysis of short peptides, although aromatic residues were also cleaved (33). Because the purification fold of the proteasome from *P. furiosus* cell extracts was approximately 16, it appears that it is a major cytosolic protease. Studies to assess its physiological role are under way. Unlike PfpI, no evidence for smaller active forms of the proteasome in *P. furiosus* was detected.

Other Pyrococcal Proteases. There have been reports of other proteases produced by pyrococcal species. A prolyl endopeptidase was identified in *P. furiosus*, the gene for which was cloned and expressed in *E. coli* (108). The recombinant enzyme was found to be optimally active between 85 and 90 °C. The gene encoding the enzyme corresponds to a polypeptide of 70 kDa, which was consistent with SDS-PAGE analysis of the recombinant version. A thiol protease from *Pyrococcus* strain KOD1 was isolated and characterized (181). This protease, one of at least three extracellular proteases produced by this strain, had a molecular mass of 44 kDa and was optimally active at pH 7.0 and 110 °C. Its threefold activation in the presence of 8 mM cysteine led to its classification as a thiol protease.

Table 1. Types of Enzymes Purified from Extreme Thermophiles

Enzyme	M_r[a]	T_{opt}[b]	$t_{50\%}$[c]	Function	Source[d]	Reference
			Hydrolases			
Aminopeptidase	320 (α_4)	75°	NA	Peptide hyrolysis	Ss	22
Carboxypeptidase	170 (α_4)	85°	<15 min 95°	Peptide hydrolysis	Ss	23
Archaelysin	52 (α)	98°	1.5 h/95°	Peptide hydrolysis	DM	24
Protease PfpI-H	66 (α_3)	105°	33 h/98°	Peptide hydrolysis	Pf	25
Protease PfpI-Cl	113 (α_6)	85°	<0.5 h/98°	Peptide hydrolysis	Pf	26
Protease pyrolysin	150 (α)	115°	9 h/95°	Peptide hydrolysis	Pf	27
Tetrabrachion	130 (α)	>90°	NA	Structural/proteolytic	Sm	28
Thiol protease	44 (α)	110°	1 h/100°	Peptide hydrolysis	P KOD1	29
Proteinase	68 (α)	85°	22 h/90°	Peptide hydrolysis	Ts	30
Thermopsin	46 (α)	75°	90 min/90°	Peptide hydrolysis	Sa	31
Proteinase I	108 (α_2)	90°	5.7 h/92°	Peptide hydrolysis	Ss	32
Proteasome	640 ($\alpha_{14}\beta_{14}$)	95°	>24 h/95°	Protein turnover	Pf	33
α-Glucosidase	125 (α)	110°	48 h/98°	Maltose hydrolysis	Pf	34
α-Glucosidase	80 (α)	105°	3.0 h/95°	Maltose hydrolysis	Ss	35
Amylase	132 (α_2)	100°	2 h/120°	Starch hydrolysis	Pf	36,37
Amylase	70 (α)	100°	6 h/100°	Starch hydrolysis	Pw	38
Amylopullulanase	140 (α)	118°	20 h/98°	Starch degradation	ES	39
β-Glucosidase	230 (α_4)	105°	13 h/110°	Cellobiose hydrolysis	Pf	40
Sucrose α-glucohydrolase	114 (α)	105°	48 h/95°	Unknown	Pf	41
β-Mannosidase	240 (α_4)	105°	1.1 h/98°	Mannobiose hydrolysis	Pf	42
β-Mannosidase	200 (α_2)	87°	17 h/85°	Mannobiose hydrolysis	Tn	43
β-Mannanase	55 (α)	92°	34 h/85°	Mannan hydrolysis	Tn	43
α-Galactosidase	66 (α)	105°	9 h/85°	Galactomannan hydrolysis	Tn	43
Endo-β-1,4-glucanase	27 (α)	95°	>6 h/80°	Glucan hydrolysis	Tm	44
Exo-β-1,4-glucanase	29 (α)	95°	>6 h/80°	Glucan hydrolysis	Tm	44
Xylanase A	120 (α)	92°	45 min/90°	Xylan/xylose hydrolysis	Tm	45
Xylanase B	40 (α)	105°	125 min/95°	Xylan/xylose hydrolysis	Tm	45
Exo-1,4-β-cellobiohydrolase	36 (α)	105°	70 min/108°	Cellulose hydrolysis	TF	46
β-Xylosidase	174 (α_2)	80°	2 h/80°	Xylobiose hydrolysis	TF	47
α-Arabinofuranosidase	92 (α)	80°	4 h/98°	Arabinoxylan hydrolysis	TF	48
Fumarate hydratase	170 (α_4)	85°	15 min/90°	Sugar metabolism	Ss	49
N^5N^{10}-methenyltetrahydro-methanopterin cyclohydrolase	41.5 (α)	95°	>1 h/90°	Methanogenesis	Mk	50

Table 1. Types of Enzymes Purified from Extreme Thermophiles (continued)

Enzyme	M_r[a]	T_{opt}[b]	$t_{50\%}$[c]	Function	Source[d]	Reference
			Oxidoreductases			
Aldehyde OR	85 (α)	>95°	6 h/80°	Glycolytic enzyme	Pf	51
Pyruvate OR	120 ($\alpha\beta\gamma\delta$)	>95°	0.3 h/90°	Pyruvate oxidation	Pf	52
Formaldehyde OR	280 (α_4)	>95°	2 h/80°	Unknown	Tl	53
Indolepyruvate OR	180 ($\alpha_2\beta_2$)	>95°	6 h/80°	Peptide fermentation	Pf	54
Gly. 3-phosphate OR	63 (α)	>95°	NA	Glycolytic enzyme	Pf	55
Isovalerate OR	220 $(\alpha\beta\gamma\delta)_2$	>95°	0.6 h/90°	Peptide fermentation	Pf	56
2-Ketoglutarate	220 $(\alpha\beta\gamma\delta)_2$	>95°	0.3 h/90°	Peptide fermentation	Pf	57
Pyruvate OR	200 $(\alpha\beta\gamma\delta)_2$	>95°	15 h/80°	Pyruvate oxidation	Tm	58
Quinone OR	280 ($\alpha\beta\gamma\delta\epsilon\zeta\eta$)	65°	5 min/75°	Electron donor	Af	59
Glutamate dehydrogenase	270 (α_6)	95°	10 h/100°	Glutamate oxidation	ES	60
GAPDH	150 (α_4)	NA	0.7 h/100°	Glycolysis (?)	Pw	61
GAPDH (NAD)	196 (α_4)	NA	0.3 h/100°	Unknown	Tt	62
GAPDH (NADP)	156 (α_4)	NA	0.5 h/100°	Unknown	Tt	62
Sulfide dehydrogenase	92 ($\alpha\beta$)	>95°	12 h/95°	Sulfur reduction	Pf	63
Alcohol dehydrogenase	200 (α_4)	80°	2.0 h/85°	Unknown	Tl	64
Isocitrate dehydrogenase	90 (α_2)	NA	NA	Sugar metabolism	Ss	65
Quinol oxidase	149 ($\alpha\beta\gamma\delta\epsilon$)	NA	NA	Proton gradient	Aa	66
Methyl coenzyme M reductase	270 $(\alpha\beta\gamma)_2$	NA	NA	Methanogenesis	Mk	67
N^5N^{10}-methylenetretrahydromethanopterin reductase	300 (α_2)	90°	>1 h/100°	Methanogenesis	Mk	68
N^5N^{10}-methylenetetrahydromethanopterin dehydrogenase	170 (α_4)	>90°	Very labile	Methanogenesis	Mk	69
Adenylylsulphate reductase	160 (α_2)	85°	NA	Electron donor	Af	70
			Hydrogenases / redox proteins			
Hydrogenase	118 ($\alpha\beta$)	>90°	1 h/98°	H_2 production	Pf	71,72,73
Hydrogenase	280 (α_4)	>95°	1 h/90°	H_2 production	Tm	74
Ferredoxin	7 (α)	>95°	>24 h/95°	Electron transfer	Pf, Tl	75
Rubredoxin	5 (α)	>95°	>24 h/95°	Unknown	Pf	76
Thioredoxin (glutaredoxin)	24.8 (α)	NA	>3 h/90°	Reduces disulfides	Ss, Pf	77,78
			Isomerases / invertases			
Xylose isomerase	204 (α_4)	97°	2.3 h/95°	Sugar metabolism	Tm	79
Triose phosphate isomerase	100 (α_4)	NA	4.7 h/104°	Unknown	Pw	80
Prolyl *cis*/*trans* isomerase	18 (α)	NA	NA	Unknown	Ss	81

Table 1. Types of Enzymes Purified from Extreme Thermophiles (*continued*)

Enzyme	M_r[a]	T_{opt}[b]	$t_{50\%}$[c]	Function	Source[d]	Reference
			Nucleic-acid modifying enzymes			
DNA polymerase	93 (α)	>75°	20 h/95°	Replication	Pf	82,83
Endonuclease SuaI	NA	65°	30 min 80°	Restriction	Sa	84
Reverse gyrase	200 ($\alpha\beta$)	NA	NA	Supercoiling	Mk	85
DNA topoisomerase II	230 ($\alpha_2\beta_2$)	80°	35 m/80°	DNA relaxation	Ssh	86
DNA topoisomerase III	108 (α)	NA	NA	DNA relaxation	Damy	87
DNA toposisomerase V	110 (α)	75°	NA	DNA relaxation	Mk	88,89
RNAse p2	13 (α_2)	NA	<15 min/90°	RNA hydrolysis	Ss	90
			Synthetases			
ACS I	140 ($\alpha_2\beta_2$)	>95°	18 h/80°	Acetyl-CoA formation	Pf	91
ACS II	140 ($\alpha_2\beta_2$)	>95°	8 h/80°	Aryl-CoA formation	Pf	91
Carbamoyl-P synthetase	46.6 (α)	>90°	3 h/95°	Pyrimidine pathway	Paby	92
PEP-synthase	2000 (α_7)	NA	NA	Metabolism	Sm	93
			Transferases			
Aromatic aminotransferase	90 (α_2)	>95°	NA	Peptide fermentation	Pf	94
Aspartate transcarbamoylase	NA	70°	>6 h/90°	Nucleotide synthesis	Paby	95
Propylamine transferase	110 (α_3)	90°	>1 h/100°	Amino acid metabolism	Ss	96
Formyltransferase	35 (α)	90°	>1 h/90°	Methanogenesis	Mk	97
			Other			
Glucokinase	93 (α_2)	105	3.7 h/100°	Sugar metabolism	Pf	98
Enolase	180 (α_4)	>95°	0.7 h/100°	Glycolysis	Pf	99
Chaperonin	920 (α_{16})	80°	NA	Protein refolding	Ss	100
PEP carboxylase	240 (α_4)	88°	>2 h/80°	Metabolism	Ms	101
5′Methylthioadenosine phosphorylase	160 (α_6)	120°	15 min/130°	Nucleoside catabolysis	Ss	102

Note: NA = not available (not determined or not reported); GADPH = glyceraldehyde-3-phosphate dehydrogenase; ACS = acetyl CoA synthetase. PEP = phophoenolypyruvate.

[a]Molecular weight in kDa (the number of subunits is given in parentheses).

[b]Optimum temperature (°C) for catalytic activity in vitro.

[c]Time required to lose 50% of catalytic activity after incubation at the indicated temperature.

[d]The sources of the enzymes are Aa = *Acidianus ambivalens*; Af = *Archaeoglobus fulgidus*; Damy = *Desulfurococcus amylolyticus*; Dm = *Desulfurococcus mucosus*; ES = Es-4; Mk = *Methanopyrus kandleri*, Ms = *Methanothermus sociabilis*; Paby = *Pyrococcus abyssi*, Pf = *Pyrococcus furiosus*; P KOD1 = *Pyrococcus* KOD1; Pw = *P. woesei*; Sa = *Sulfolobus acidocaldarius*; Sm = *Staphylothermus marinus*; Ss = *Sulfolobus solfataricus*; Ssh = *Sulfolobus shibatae*; Tf = *Thermotoga* FjSS3-B.1; Tl = *Thermococcus litoralis*; Tm = *Thermotoga maritima*, Tn = *Thermotoga neapolitana*; Ts = *Thermococcus stetteri*; Tt = *Thermoproteus tenax*.

Table 2. Proteins from Extreme Thermophiles Whose Genes Have Been Cloned and Expressed

Organism/Enzyme (host)	Reference
Pyrococcus furiosus	
Ferredoxin	103
Glutaredoxin	78
DNA polymerase	104
α-Amylase	105
Glutamate dehydrogenase	106
Protease I	107
Prolyl endopeptidase	108
Citrate synthase	109
β-Glucosidase	42,110
β-Mannosidase	42
Pyrococcus woeseii	
Glyceraldehyde-3-phosphate dehydrogenase	61
3-Phosphoglycerate kinase	111
Pullulanase	112
GTP-binding protein (expression library)	113
Transcription factor	114
Pyrococcus strain ES-4	
Glutamate dehydrogenase	60
Pyrococcus sp. KOD1	
Aspartyl-tRNA synthetase	115
TATA-binding protein	116
α-Amylase	117
RecA/RAD51 protein	118
Molecular chaperonin beta subunit	119
Pyrococcus sp. strain ST700	
Adenylosuccinate	120
Pyrodictium occultum	
DNA polymerase 1	121
DNA polymerase 2	121
Thermococcus litoralis	
DNA polymerase	122
Thermococcus sp. 9 degree N-7	
DNA polymerase	123
Thermotoga neapolitana$^+$	
Phosphoglycerate kinase	124
Triosephosphate isomerase	124
Enolase	124
Xylose isomerase	125
Endoxylanase A	126
Thermotoga maritima$^+$	
Dihydrofolate reductase	127
L-Lactate dehydrogenase	128
D-Glyceraldehyde-3-phosphate dehydrogenase	129
Chemotaxis protein Y	130
Chemotaxis protein W	130
Ferredoxin	131
RuvB	132
trp Gene products (tryptophan pathway)	133
4-α-Glucanotransferase	134
α-Amylase	135
3-Phosphoglycerate kinase	136
Triosephosphate isomerase	136
β-glucosidase	137
β-galactosidase	137
β-glucanases (2)	138
Histidine protein kinase	139
Response regulator	139
Elongation factor Tu	140
Thermotoga FjSS3-B.1	
Endoxylanase A (expression library)	141
Desulfurococcus SY	
Aspartate racemase	142
Sulfolobus acidocaldarius	
Sac 7d gene product	143
Genanylgeranyl-diphosphate synthase	144
Adenylate kinase	145
Sulfolobus sp. strain 7	
Ferredoxin	146
Sulfolobus solfataricus	
Anthranilate synthase	147
β-Glycosidase (Ec), (Sc), (mammalian)	148–150
Phosphoglycerate kinase	151
Glyceraldehyde-3 phosphate dehydrogenase	151
Elongation factor I	152
Glycosyltransferase	153
α-Amylase	153
Aspartate aminotransferase	154
3-Hydroxy-3-methylglutaryl CoA reductase	155
Sulfolobus shibatae	
ATP(CTP):tRNA nucleotidyl transferase	156
Methanothermus fervidus	
2-Phosphoglycerate kinase	157
Methanobacterium thermoautotrophicum	
Methylene-5,6,7,8-tetrahydromethanopterin dehydrogenase	158
HMt 1 and 2 (histone proteins)	159
Methanopyrus kandleri	
Formylmethanofuran:tetrahydromethanopterin Formyl transferase (Ftr)	160
Caldococcus noboribetus	
Isocitrate dehydrogenase	161

Note: All genes expressed in *E. coli* unless otherwise indicated.
$^+$ = denote hyperthermophilic bacterium; Ec = *E. coli*; Sc = *Saccharomyces cerevisiae.*

Glycosyl Hydrolases from Extreme Thermophiles

The hydrolysis of natural polymers, such as cellulose, hemicellulose, and starch, by glycosyl hydrolases is the focus of a significant portion of all current commercial uses of enzymes as biocatalysts. It is desirable to hydrolyze natural polymers with these enzymes at elevated temperatures because of increased substrate solubility and lowered solution viscosity, which improves enzyme access to susceptible bonds (165). It is somewhat surprising that the source of these thermostable enzymes, extreme thermophiles, are thought to be primitive life forms and presumably predate the emergence of higher life forms, such as plants, which produce most naturally occurring polysaccharides. Nonetheless, these organisms have proved to be a fertile source of an array of glycosyl hydrolases, many of which have potential as commercial enzymes (163,165). Classification schemes for glycosyl hydrolases have advanced considerably in recent years. Because these enzymes are often capable of hydrolyzing a range of glycosidic bonds, classification in terms of substrate specificity can be confusing. Amino acid sequence information has been used to fortify traditional classifications schemes so that glycosyl hydrolases have now been arranged in over 60 families (182). As three-dimensional structures and detailed kinetic mechanisms are resolved for enzymes within these families, a comprehensive basis for distinguishing among various glycosyl hydrolases can be developed.

Hydrolysis of α-Specific Glycosidic Bonds. Starch, glycogen, pullulan, and maltose have proved to be good growth substrates for a variety of heterotrophic extreme thermophiles. These microorganisms contain a variety of *exo*-acting and *endo*-acting α-specific glycosyl hydrolases that are apparently used to hydrolyze complex carbohydrates to glucose. *P. furiosus*, for example, produces an extracellular amylopullulanase that can hydrolyze either α-1,4 or α-1,6 linkages (183), both intracellular (36) and extracellular (165) α-amylases, and an α-glucosidase (34). Versions of these enzymes have been found in many other heterotrophic extreme thermophiles (163,165). In fact, glycosyl hydrolases may be more widely distributed among the extreme thermophiles than previously thought. The genome sequence of *M. jannaschii* indicates that this methanogen also produces an intracellular α-amylase and glucoamylase, although this organism has not been reported to grow on starch-based carbon sources (20).

Hydrolysis of β-Specific Glycosidic Bonds. Because of the natural abundance of cellulose and hemicellulose, enzyme systems that are capable of hydrolyzing these polysaccharides to simple sugars have long been sought to facilitate recycling of renewable resources (184). The structure and molecular composition of β-linked natural biopolymers are highly variable, often necessitating the use mixtures of enzymes to get substantial hydrolysis. On the other hand, unsubstituted cellulose and hemicellulose (e.g., xylan and mannan) are sparingly soluble in water, and enzymes capable of hydrolyzing these compounds often contain binding domains used to pry open the structure for enzymatic access (184).

Exo-acting, β-specific glycosyl hydrolases have been found among extremely thermophilic bacteria and archaea (163,165). *P. furiosus* produces an intracellular β-glucosidase (40,185) and an intracellular β-mannosidase (42); although the former appears to be induced by cellobiose in the growth media and has broad substrate specificity, the function of the latter is not clear. β-glucosidases, β-galactosidases, β-xylosidases, and cellobiohydrolases have also been identified in other extreme thermophiles (163,165), which suggests that biopolymeric substrates associated with these specificities are present in geothermal environments.

The wide occurrence of *exo*-acting, β-specific glycosyl hydrolases among the extreme thermophilic archaea suggests that *endo*-acting enzymes capable of hydrolyzing β-linked polysaccharides are also present. However, no such *endo*-acting, β-specific glycosyl hydrolases have been identified to this point in the extremely thermophilic archaea. Although some evidence for *endo*-acting hemicellulase activity toward xylan-based polysaccharides among the extremely thermophilic archaea has been mentioned (165), a xylanase corresponding to this activity has not been isolated. The hyperthermophilic bacterial genus *Thermotoga*, however, has been a rich source of *endo*-acting, β-specific enzymes. Glucanases (44,186), xylanases (45,141,186–189), and mannanases (43,190) have all been purified and characterized from various *Thermotoga* species.

MECHANISMS OF THERMOSTABILITY

Although there has been much progress in isolating and characterizing the biochemical features of extremely thermophilic enzymes (4), little is yet known about the intrinsic basis for their astounding thermostability. It was only a few years ago that the amino acid sequences of these enzymes were eagerly awaited, because it was anticipated that they could be used as templates for the design of stable forms of industrially important, but thermally labile, enzymes. However, as more and more genes encoding hyperthermophilic enzymes have become available and their amino acid translations scrutinized, it is evident that the present sophistication of protein chemistry provides little encouragement that strategies for protein stabilization in a generic sense are at hand (7,191). In fact, three-dimensional structures are now available for four hyperthermophilic proteins (192–195) but these have not yielded specific mechanisms for stabilization (6). At present, it appears that the most fruitful approach to obtaining extremely stable proteins is to isolate them from extreme thermophiles or to express the genes encoding these proteins in mesophilic hosts. But even in these cases, it is important to examine both intrinsic and extrinsic factors to ensure that the recombinant protein so produced resembles its natural counterpart.

Much has been done to examine intrinsic factors that lead to the thermostabilization of hyperthermophilic enzymes. However, much less has been done to examine extrinsic factors, such as small molecules and other proteins present in the microenvironment in which the protein folds. By mimicking the intracellular environment encoun-

tered by these enzymes, their in vitro stability and function might be optimized. In addition, such conditions may be useful in improving the stability of less thermophilic proteins, particularly those that are used industrially where their stability is a limiting factor.

Intrinsic Factors Influencing Protein Stability

A number of thermostable enzymes from extremely thermophilic microorganisms have been discovered during the past decade (Table 1). In most cases, it is not unusual to find the $t_{1/2}$ of these enzymes to be on the order of a day or more at temperatures near or even above 100 °C. Just as is the case with mesophilic organisms, enzymes from a particular hyperthermophile vary in their catalytic efficiency (k_{cat}/K_m), T_{opt}, and thermostability because of factors presumably related to their physiological roles. However, the intrinsic bases for protein thermostabilization at high temperatures are no better understood than the bases for stabilization at moderate conditions. No one feature seems to underlie the fact that proteins exist in stable conformations as a result of a delicate balance of large, opposing enthalpic and entropic forces. Additional stabilization arises from the efficient packing of amino acid residues, optimization of charged interactions within the structure, and minimization of exposure of hydrophobic elements to the solvent (5). This all implies a degree of cooperativity within the protein structure (196), which is difficult to attribute in a quantitative way to its component parts.

Extrinsic Factors Influencing Protein Stability

Extrinsic factors clearly play an important role in protein stabilization. Environmental conditions such as temperature and pressure have a universal effect on the collection of proteins within a given cell, but other extrinsic factors that influence protein thermostability are more specific. Some proteins benefit from the presence of critical concentrations of certain cations, whereas others function best and are the most stable in concert with coenzymes and cofactors, such as ATP. It is difficult to be definitive about how the in vitro stability of an intracellular protein relates to its in vivo stability, because the latter can be influenced by so many intracellular factors. This assessment may be especially difficult for extremely thermophilic proteins because in vitro, they are subject to an array of potentially damaging chemical reactions, involving covalent modifications to thermally labile amino acid residues, such as deamidation and oxidation. Protection against such destabilizing forces is presumably provided in the intracellular environment and may or may not be important in vitro.

There are ways that protein stability can be influenced in vitro. Of course, some conditions that optimize enzyme activity can be readily determined, such as pH, ionic strength, temperature and enzyme concentration. Other factors that may be important in vivo are less easily reproduced. These include the presence of compatible solutes (197,198) or molecular chaperones (199). Each of these will be discussed with regard to extremely thermophilic proteins.

Compatible Solutes. Although enzymes purified from extreme thermophiles are intrinsically more thermostable than those purified from their less thermophilic counterparts, the intracellular environments of extreme thermophiles thus far examined contain high levels of certain unusual solutes. These compounds are thought to play a role in the thermoprotection of at least some of their constituent biomolecules in vivo (200). High concentrations of a novel phosphodiester, di-*myo*-inositol-1,1′-phosphate (DIP), have been found in a number of hyperthermophiles, and the potassium salt of this compound has been shown to offer a significant degree of thermoprotection to at least one hyperthermophilic enzyme in vitro (201). The mechanisms for protein stabilization and the pathways used to synthesize DIP have not been determined, nor have the factors that influence its intracellular production. These studies are necessary to understand the intrinsic basis for the function of organisms at extremely high temperatures as well as to determine factors responsible for the stabilization of proteins in a general sense.

Scholz et al. (202) reported that the S°-dependent hyperthermophile *Pyrococcus woesii* ($T_{opt} \sim 100$ °C) contains high concentrations (~0.6 M) of DIP, and Ciulla et al. (203) subsequently found that the hyperthermophilic methanogen *Methanococcus igneus* ($T_{opt} \sim 80$ °C) also produced DIP when grown at temperatures above 80 °C or when exposed to high external salt concentrations. They also showed that DIP is not used as a metabolic intermediate. Martins and Santos (204) determined the intracellular concentration of DIP to be ~60 mM when *Pyrococcus furiosus* (T_{opt} 100 °C) is grown at 95 °C, but this increases to above 1.0 M at supraoptimal temperatures. In addition, they found that at high salinity, the organism accumulated the unusual sugar derivative β-mannosyl glycerate (~30 mM). In contrast to DIP, the latter compound had previously been found in some moderately thermophilic bacteria (205), suggesting that it may play a role in both osmolysis and thermal protection in these organisms. Moreover, the extraordinary increase in DIP concentrations in *P. furiosus* at the higher growth temperatures strongly suggests that it plays some role in thermoprotection (204). Accordingly, Hensel and Jakob (201) showed that a crude preparation of DIP (~115 mM) dramatically increased the thermal stability of glyceraldehyde-3-phosphate dehydrogenase (61) from *P. woesii* (the time for a 50% activity loss at 105 °C increased by 10-fold). They also showed that compounds related in structure to DIP, such as citrate and phosphate, stabilized two hyperthermophilic enzymes (201).

Molecular Chaperones in Extreme Thermophiles and Protein Aggregation. Although many proteins have been shown to fold spontaneously into their native "correct" structures in vitro, it is also clear that molecular chaperones often assist in the folding process in vitro and, presumably, in vivo. In one sense, the temperatures at which extreme thermophiles grow would seem to create a perpetual situation of thermal stress, yet these organisms respond to heat shock in many of the same ways as less thermophilic microorganisms (206). Members of the extremely thermophilic archaea have been shown to produce heat shock proteins (HSPs) at supraoptimal temperatures (207–210). Kagawa et al. (211) showed that *Sulfolobus shibatae* produces an HSP60-like protein, similar to the one

identified in bacteria (199), that is a double-ring complex formed from two subunits. Similar structures have been isolated from another hyperthermophilic archaeon, *Pyrodictium occultum* (212,213) and from the moderately thermophilic archaeon *Thermoplasma acidophilum* (214). All these proteins appear to have ATPase activity and have been shown to be involved in the folding process in vitro. The genes for the archaeal HSP60s that have been sequenced show that they are closely related to a family of eukaryotic proteins termed TCP1s (215).

Although the response to thermal stress in extreme thermophiles thus far appears to be similar to the stress response of more conventional organisms at lower temperatures, the degree to which molecular chaperones are involved in protein folding in vivo is not clear. One general characteristic that has been proposed that differentiates extremely thermophilic proteins from others is increased hydrophobicity (4). If so, this may mean that the tendency for extremely thermophilic proteins to aggregate during the folding process is enhanced. Thus, molecular chaperones that can bind to the nascent protein as it folds may take on added importance relative to mesophilic proteins. Aggregation will be especially problematic when high protein concentrations are present and no means to prevent collisions between interactive protein surfaces exist (5). Hence, approaches that would prevent aggregation of extremely thermophilic proteins through the addition of extrinsic factors are highly desirable, especially in systems where high levels of overexpression are needed.

Recombinant Extremely Thermophilic Enzymes

There have been several literature reports of the successful cloning and expression of the genes encoding extremely thermophilic proteins in foreign hosts (4), a summary of which is provided in Table 2. In addition, expression cloning of genomic DNA from hyperthermophiles, coupled with robotic screening for enzyme activities of interest, has led to the discovery of a number of thermostable enzymes (7). This approach, however, says little about what genes were not successfully cloned and expressed. Moreover, of the proteins listed in Table 2, there have been only a few reports on the comparative properties of the native and recombinant forms. In fact, only in a limited number of cases has the native protein been characterized and the recombinant protein purified. For example, only DNA polymerase, glyceraldehyde-3-phosphate dehydrogenase, glutaredoxin, and β-glucosidase from *P. furiosus* have been shown, thus far, to have comparable thermal stability and molecular and catalytic properties in their native and recombinant forms. In contrast, native glutamate dehydrogenase (GDH) from *P. furiosus* has a half-life at 100 °C of about 10 h (216), whereas the recombinant form has a half-life of minutes at 100 °C, suggesting that the protein is incompletely folded or not assembled (106). The latter was supported by the fact that the native enzyme is a homohexamer (α_6), whereas the recombinant form was comprised of inactive monomers. Interestingly, some conversion of the latter into the active hexameric form was accomplished by heating (85 °C for 15 min). The opposite situation appears to exist with the α-amylase of *P. furiosus* (105), as the recombinant form of this enzyme, although not purified, was of higher molecular weight than the native form, suggesting some type of aggregation had occurred. Last, the recombinant form of *P. furiosus* citrate synthase was reported to have a half-life of only 1 min at 100 °C (109). Unfortunately, the native enzyme has not been purified but one could anticipate that it would be much more stable than that produced by *E. coli*. Clearly, some recombinant extremely thermophilic enzymes are as stable as their native counterparts, but it is also apparent that this is not a general phenomenon.

USES OF EXTREMELY THERMOSTABLE ENZYMES

General Considerations

The increasing variety of enzymes purified from extreme thermophiles (Table 1) supports the prospect of their future use in industrial applications. In some situations, existing enzyme-based approaches could potentially be improved by the use of substantially more thermostable versions of the same enzyme type. This presumes that the increase in enzyme stability achieved with an extremely thermophilic enzyme is not offset by a loss of enzyme activity at the desired operating temperature; this may or may not be the case. The fact that extremely thermophilic enzymes tend to be less active at ambient temperatures may be desirable in some instances, especially in cases where it is necessary to have strict control over the timing or extent of the reaction. In other instances, strategic opportunities that make use of improved biocatalyst stability and also advantageously use higher operating temperatures are desirable. Probably the most intriguing aspect is the use of extremely thermophilic enzymes to catalyze reactions that otherwise were not amenable to biocatalysis. Examples of these cases are presented below.

Polymerase Chain Reaction

The revolution in molecular biology has to a significant extent relied on the ability to amplify small amounts of DNA to levels that facilitate cloning, sequencing, and expression of particular genes of interest. This can be accomplished through the use of thermostable DNA polymerases that are produced by thermophilic microorganisms (122). To be effective in catalyzing the DNA polymerization reaction, the polymerase must be able to withstand repeated thermal swings from low to high temperatures so that the DNA template can denature to be copied and then reanneal before the next cycle. Although the DNA polymerase from a moderately thermophilic bacterium, *Thermus aquaticus*, was initially used in PCR applications, a host of polymerases from bacterial and archaeal sources are now available (122). Recent advances include the ability to replicate long sequences (20 to 40 kb) by adding to the reaction mixture an additional DNA polymerase with proofreading (3′ to 5′ *exo*-nuclease activity) capability to repair mistakes that might otherwise terminate the replication process (217). PCR is an example of the case in which protein thermostability is used strategically because thermally labile polymerases would not survive the repeated exposures to high

temperatures, thus requiring addition of enzyme at each cycle.

Replacements of Existing Industrial Enzymes

Glucose Isomerase. The biocatalyzed conversion of glucose to fructose for the production of high-fructose corn syrup (HFCS) by xylose (glucose) isomerase represents the largest existing industrial application of an immobilized enzyme. Because of enzyme thermostability limitations, this process is currently operated at approximately 60 °C, at which temperature the equilibrium conditions limit the HFCS concentration of fructose to a maximum of 40 to 42% (218). Because 55% fructose concentration is desired for sweetening properties, a chromatographic separation step is subsequently used to enrich the mixture to this level by the removal of unreacted glucose. However, because the equilibrium yield of fructose increases with temperature (55% fructose estimated at 95 °C) (218), thermostable versions of glucose isomerase (GI) that could be used at higher temperatures are desirable to minimize the need for the glucose removal step.

Brown et al. (79) reported the presence of xylose isomerase in members of the hyperthermophilic eubacterial genus *Thermotoga*. For example, *T. neapolitana* produces a version of GI with an optimal temperature of at least 95 °C, which was also found to be highly active over a broad pH range (125). Compared to many other GIs that have been used commercially, the *T. neapolitana* GI possesses a 50 amino acid insert at the N-terminus, the function of which is unknown. The catalytic efficiency of this enzyme was superior to those determined for other GIs from less thermophilic sources at their respective optimal temperature ranges (125).

The GI from *T. neapolitana* again illustrates the potential strategic benefits that arise from the availability of thermostable versions of enzyme used in current industrial applications. By operating at elevated temperatures, the separation step needed for fructose enrichment can be minimized or eliminated because of the improved reaction yield. However, many other issues must be resolved before this enzyme can be considered as a replacement for the one used in existing processes. For example, consideration must be given to the levels of overexpression that can be attained in a foreign host, the impact of and approach used for immobilization, problems with potential side reactions of sugars at high temperatures, and regulatory approvals that might be required in the production of a substance for human consumption.

Hydrolysis of Natural Polymers. Interest in renewable resources has driven efforts to identify enzyme systems capable of hydrolyzing natural biopolymers, such as cellulose and hemicelluse, to more readily usable saccharides. This may be done in a nonspecific way so that the biopolymer is extensively hydrolyzed to simple sugars or in a specific way so that the properties of the biopolymer are systematically modified (219). As mentioned earlier, solutions containing natural polymers are less viscous at high temperatures compared to ambient conditions so that enzyme access to specific sites for hydrolysis is facilitated. Therefore, natural polymers may be hydrolyzed more readily by extremely thermophilic enzymes. Additionally, because these enzymes are typically stable at temperatures much lower than their optimal temperatures for hydrolysis, these hydrolases may retain their function at suboptimal temperatures for extended periods of time.

In the stimulation of oil and gas wells by hydraulic fracturing, aqueous solutions of natural polymers such as guar gum (a molecule consisting of a mannan backbone substituted with galactose through α-1,6 linkages) are used to transport particles to the site of the fracture (220). The objective is to prop open fissures in the bedrock perpendicular to wellbore and, hence, promote the flow of oil or gas to the point where it is readily recovered. The viscous properties of the guar gum solution used advantageously to transport the particles become problematic once the particles are in place. Thus, various methods, including the use of hemicellulases, are used to hydrolyze the guar to reduce viscosity. The high temperatures encountered in deep wells reduce the effectiveness of conventional enzyme systems but create an opportunity for extremely thermophilic enzymes that can withstand temperatures above 100 °C (43). An additional advantage of using thermostable enzymes is their lower activity at ambient conditions that otherwise can lead to premature hydrolysis of the guar (190). Such an enzyme system, consisting of β-mannanase, α-galactosidase, and β-mannosidase from the hyperthermophile *T. neapolitana*, has been shown to work to hydrolyze guar at high temperatures but much less so at moderate temperatures (190).

New Opportunities

To date, extremely thermophilic enzymes have been considered mostly as replacement enzymes to improve existing bioprocesses. However, their high degree of stability to thermal and other denaturing forces makes them candidates for catalyzing reactions not previously amenable to biocatalysis (6). This includes reactions conducted in nonaqueous environments such as organic solvents and supercritical fluids. The fact that their operational ranges are consistent with conditions used in chemical processing environments presents some intriguing possibilities for biocatalysis. It remains to be seen if these enzymes will be used for these purposes. Certainly, any insights gained into the intrinsic or extrinsic basis for enzyme stability that can be translated into bioprocessing improvements for enzymes from any source would be a welcome development.

SUMMARY

Extremely thermophilic microorganisms have proved to be a rich source of thermostable enzymes. The diversity of such enzymes and the associated biochemical properties are only beginning to be understood. The discovery process has made good use of the tools of molecular biology, and these will also be critical for producing sufficient enzymes for various applications. The promise that these enzymes hold for the expansion of uses of biocatalysis is considerable but will require input from those outside of this field to find the most suitable applications. Nonetheless, biocatalysis has added a new dimension with the increased

availability of extremely thermostable enzymes, which bodes well for the development of more efficient and effective bioprocesses.

BIBLIOGRAPHY

1. R. Jaenicke, *Eur. J. Biochem.* **202**, 715–728 (1991).
2. R. Jaenicke, H. Schurig, N. Beaucamp, and R. Ostendorp, in M.W.W. Adams ed., *Advances in Protein Chemistry* 48, Academic Press, San Diego, 1996, pp. 181–269.
3. R. Jaenicke, *FASEB J.* **10**, 84–92 (January 1996).
4. M.W.W. Adams, F.B. Perler, and R.M. Kelly, *Bio/Technol.* **13**, 662–668 (1995).
5. R. Jaenicke, *FEMS Microbiol. Rev.* **18**, 215–224 (1996).
6. M.W.W. Adams and R.M. Kelly, *Chem. Eng. News* **73**, 32–42 (1995).
7. D.E. Robertson, E.J. Mathur, R.V. Swanson, B.L. Marrs, and J.M. Short, *SIM News* **46**, (1996).
8. E.J. Strauss and S. Falkow, *Science* **276**, 707–712 (1997).
9. T.D. Brock, *Thermophilic Microorganisms and Life at High Temperatures*, Springer-Verlag, New York, 1978.
10. J.A. Baross and J.W. Deming, in D.M. Karl ed., *Deep-Sea Hydrothermal Vents*, CRC Press, Boca Raton, 1995, pp. 169–217.
11. K.O. Stetter, *FEMS Microbiol. Rev.* **18**, 149–158 (1996).
12. R.M. Kelly, S.H. Brown, I.I. Blumentals, and M.W.W. Adams, in M.W.W. Adams and R.M. Kelly ed., *Biocatalysis at Extreme Temperatures*, American Chemical Society, Washington, D.C., 1992, pp. 23–41.
13. K.D. Rinker, C.J. Han, M.W.W. Adams, and R.M. Kelly, in *ASM Manual of Industrial Microbiology and Biotechnology*, in press.
14. G. Fiala and K.O. Stetter, *Arch. Microbiol.* **145**, 56–61 (1986).
15. A. Neuner, H.W. Jannasch, S. Belkin, and K.O. Stetter, *Arch. Microbiol.* **153**, 205–207, (1990).
16. R. Huber, T.A. Langworthy, H. Konig, M. Thomm, C.R. Woese, U.B. Sleytr, and K.O. Stetter, *Arch. Microbiol.* **144**, 324–333 (1986).
17. M.W.W. Adams, in F.T. Robb, A.R. Place, K.R. Sowers, H.J. Schreier, S. DasSarma, and E.M. Fleischmann ed., *Archaea: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York, 1995, pp. 47–49.
18. C.-C. Chen, M.W.W. Adams, and J. Westpheling, *4th Annual Hyperthermophile Symposium*, Del Mar, Calif., May 14–16, 1997.
19. D. Hess and R. Hensel, *Gene* **172**, 121–124 (1996).
20. C.J. Bult, O. White, G.J. Olsen, L. Zhou, R.D. Fleischmann, G.G. Sutton, J.A. Blake, L.M. Fitzgerald, R.A. Clayton, J.D. Gocayne, A.R. Kerlavage, B.A. Dougherty, J.F. Tomb, M.D. Adams, C.I. Reich, R. Overbeek, E.F. Kirkness, K.G. Weinstock, J.M. Merrick, A. Glodek, J.L. Scott, N.S.M. Geoghagen, and J.C. Venter, *Science* **273**, 1058–1073 (1996).
21. The Institute for Genomic Research (TIGR), 1996, 1997, (*http://www.tigr.org/*).
22. M. Hanner, B. Redl, and G. Stoffler, *Biochim. Biophys. Acta 1033*, 148–153 (1990).
23. S. Colombo, S. D'Auria, P. Fusi, L. Zecca, C.A. Raia, and P. Tortora, *Eur. J. Biochem.* **206**, 349–357 (1992).
24. D.A. Cowan, K.A. Smolenski, R.M. Daniel, and H.W. Morgan, *J. Biochem.* **247**, 121–133 (1987).
25. I.I. Blumentals, A.S. Robinson, and R.M. Kelly, *Appl. Environ. Microbiol.* **56**, 1992–1998 (1990).
26. S.B. Halio, M.W. Bauer, S. Mukund, M.W.W. Adams, and R.M. Kelly, *Appl. Environ. Microbiol.* **63**, 289–295 (1997).
27. W.G.B. Voorhorst, R.I.L. Eggen, A.C.M. Geerling, C. Platteeuw, R.J. Siezen, and W.M. de Vos, *J. Biol. Chem.* **271**, 20426–20431 (1996).
28. J. Peters, M. Nitsch, B. Kuhlmorgen, R. Golbik, A. Lupas, J. Kellerman, H. Engelhardt, J.P. Pfander, S. Muller, K. Goldie, A. Engel, K.O. Stetter, and W. Baumeister, *J. Mol. Biol.* **245**, 385–401 (1995).
29. J.A. Maupin-Furlow and J.G. Ferry, *J. Biol. Chem.* **270**, 28617–28622 (1995).
30. M. Klingeberg, B. Galunsky, C. Sjoholm, V. Kasche, and G. Antranikian, *Appl. Environ. Microbiol.* **61**, 3098–3104 (1995).
31. M. Fusek, X. Lin, and J. Tang, *J. Biol. Chem.* **265**, 1496–1501 (1990).
32. N. Burlini, P. Magnani, A. Villa, F. Macchi, P. Tortora, and A. Guerritore, *Biochim. Biophys. Acta* **1122**, 283–292 (1992).
33. M.W. Bauer, S.H. Bauer, and R.M. Kelly, *Appl. Environ. Microbiol.* **63**, 1160–1164 (1997).
34. H.R. Costantino, S.H. Brown, and R.M. Kelly, *J. Bacteriol.* **172**, 3654–3660 (1990).
35. M. Rolfsmeier and P. Blum, *J. Bacteriol.* **177**, 482–485 (1995).
36. K.A. Laderman, B.R. Davis, H.C. Krutzsch, M.S. Lewis, Y.V. Griko, P.L. Privalov, and C.B. Anfinsen, *J. Biol. Chem.* **268**, 24394–24401 (1993).
37. R. Koch, P. Zablowski, A. Spreinat, and G. Antranikian, *FEMS Mirobiol. Lett.* **71**, 21–26 (1990).
38. R. Koch, A. Spreinat, K. Lemke, and G. Antranikian, *Arch. Microbiol.* **155**, 572–578, (1991).
39. J.W. Schuliger, S.H. Brown, J.A. Baross, and R.M. Kelly, *Mol. Mar. Biol. Biotechnol.* **2**, 76–87 (1993).
40. S.W.M. Kengen, E.J. Luesink, A.J.M. Stams, and A.J.B. Zehnder, *Eur. J. Biochem.* **213**, 305–312 (1993).
41. H.R. Badr, K.A. Sims, and M.W.W. Adams, *Syst. Appl. Microbiol.* **17**, 1–6 (1994).
42. M.W. Bauer, E.J. Bylina, R.V. Swanson, and R.M. Kelly, *J. Biol. Chem.* **271**, 23749–23755 (1996).
43. G.D. Duffaud, C.M. McCutchen, P. Leduc, K.N. Parker, and R.M. Kelly, *Appl. Environ. Microbiol.* **63**, 169–177 (1997).
44. K. Bronnenmeier, A. Kern, W. Liebl, and W.L. Staudenbauer, *Appl. Environ. Microbiol.* **61**, 1399–1407 (1995).
45. C. Winterhalter and W. Liebl, *Appl. Environ. Microbiol.* **61**, 1810–1815 (1995).
46. L.D. Ruttersmith and R.M. Daniel, *J. Biochem.* **277**, 887–890 (1991).
47. L.D. Ruttersmith and R.M. Daniel, *Biochim. Biophys. Acta* **1156**, 167–172 (1993).
48. L.D. Ruttersmith, R.M. Daniel, and H.D. Simpson, *Ann. N. Y. Acad. Sci.* **672**, 137–141 (1992).
49. S. Puchegger, B. Redl, and G. Stoffler, *J. Gen. Microbiol.* **136**, 1537–1541 (1990).
50. J. Breitung, R.A. Schmitz, K.O. Stetter, and R.K. Thauer, *Arch. Microbiol.* **156**, 517–524 (1991).
51. S. Mukund and M.W.W. Adams, *J. Biol. Chem.* **266**, 14208–14216 (1991).
52. E.T. Smith, J.M. Blamey, and M.W.W. Adams, *Biochemistry* **33**, 1008–1016 (1994).

53. S. Mukund and M.W. Adams, *J. Biol. Chem.* **268**, 13592–13600 (1993).
54. X. Mai and M.W.W. Adams, *J. Biol. Chem.* **269**, 16726–16732 (1994).
55. S. Mukund and M.W.W. Adams, *J. Biol. Chem.* **270**, 8389–8405 (1995).
56. J. Heider, X. Mai, and M.W. Adams, *J. Bacteriol.* **178**, 780–787 (1996).
57. X. Mai and M.W. Adams, *J. Bacteriol.* **178**, 5890–5896 (1996).
58. J.M. Blamey and M.W. Adams, *Biochemistry*, **33**, 1000–1007 (1994).
59. J. Kunow, D. Linder, K.O. Stetter, and R.K. Thauer, *Eur. J. Biochem.* **223**, 503–511 (1994).
60. J. DiRuggiero, F.T. Robb, R. Jagus, H.K. Klump, K.M. Borges, X. Mai, M. Kessel, and M.W.W. Adams, *J. Biol. Chem.* **268**, 17767–17774 (1993).
61. P. Zwickl, S. Fabry, C. Bogedain, A. Haas, and R. Hensel, *J. Bacteriol.* **172**, 4329–4338 (1990).
62. R. Hensel, S. Laumann, J. Lang, H. Heumann, and F. Lottspeich, *Eur. J. Biochem.* **170**, 325–333 (1987).
63. K. Ma and M.W.W. Adams, *J. Bacteriol.* **176**, 6509–6517 (1994).
64. K. Ma, F.T. Robb, and M.W. Adams, *Appl. Environ. Microbiol.* **60**, 562–568 (1994).
65. M.L. Camacho, R.A. Brown, M.-J. Bonete, M.J. Danson, and D.W. Hough, *FEMS Microbiol. Lett.* **134**, 85–90 (1995).
66. W.G. Purschke, C.L. Schmidt, A. Peterson, and G. Schafer, *J. Bacteriol.* **179**, 1344–1353 (1997).
67. S. Rospert, J. Breitung, K. Ma, B. Schwoerer, C. Zirngibl, R. K. Thauer, D. Linder, R. Huber, and K.O. Stetter, *Arch. Microbiol.* **156**, 49–55 (1991).
68. K. Ma, D. Linder, K.O. Stetter, and R.K. Thauer, *Arch. Microbiol.* **155**, 593–600 (1991).
69. K. Ma, C. Zirngibl, D. Linder, K.O. Stetter, and R.K. Thauer, *Arch. Microbiol.* **156**, 43–48 (1991).
70. N. Speich and H.G. Truper, *J. Gen. Microbiol.* **134**, 1419–1425 (1988).
71. F.O. Bryant and M.W.W. Adams, *J. Biol. Chem.* **264**, 5070–5079 (1989).
72. K. Ma, R.N. Schicho, R.M. Kelly, and M.W.W. Adams, *Proc. Natl. Acad. Sci. USA* **90**, 5341–5344 (1993).
73. A.F. Arendsen, P.T.M. Veenhuizern, and W.R. Hagen, *FEBS Lett.* **368**, 117–121 (1995).
74. A. Juszczak, S. Aono, and M.W. Adams, *J. Biol. Chem.* **266**, 13834–13841 (1991).
75. S. Aono, F.O. Bryant, and M.W.W. Adams, *J. Bacteriol.* **171**, 3433–3439 (1989).
76. P.R. Blake, J.B. Park, F.O. Bryant, S. Aono, J.K. Magnuson, E. Eccleston, J.B. Howard, M.F. Summers, and M.W.W. Adams, *Biochemistry* **30**, 10885–10895 (1991).
77. A. Guagliardi, V. Nobile, S. Bartolucci, and M. Rossi, *Int. J. Biochem.* **26**, 375–380 (1994).
78. A. Guagliardi, D. de Pascale, R. Cannio, V. Nobile, S. Bartolucci, and M. Rossi, *J. Biol. Chem.* **270**, 5748–5755 (1995).
79. S.H. Brown, C. Sjøholm, and R.M. Kelly, *Biotechnol. Bioeng.* **41**, 878–886 (1993).
80. M. Kohlhoff, A. Dahm, and R. Hensel, *FEBS Lett.* **383**, 245–250 (1996).
81. A. Guagliardi, L. Cerchia, S. Bartolucci, M. Rossi, P. Gallo, and C. Pedone, *Prot. Peptide Lett.* **1**, 5–8 (1994).
82. T. Uemori, Y. Ishino, H. Toh, K. Asada, and I. Kato, *Nucleic Acids Res.* **21, 25–265** (1993).
83. M. Imamura, T. Uemori, I. Kato, and Y. Ishino, *Biol. Pharm. Bull.* **18**, 1647–1652 (1995).
84. D.A. Prangishvili, R.P. Vashakidze, M.G. Chelidze, and I.Y. Gabriadze, *FEBS Lett.* **192**, 57–60 (1985).
85. S.A. Kozyavkin, R. Krah, M. Gellert, K.O. Stetter, J.A. Lake, and A.I. Slesarev, *J. Biol. Chem.* **269**, 11081–11089 (1994).
86. A. Bergerat, D. Gadelle, and P. Forterre, *J. Biol. Chem.* **269**, 27663–27669 (1994).
87. A.I. Slesarev, D.A. Zaitzev, V.M. Kopylov, K.O. Stetter, and S.A. Kozyavkin, *J. Biol. Chem.* **266**, 12321–12328 (1991).
88. A.I. Slesarev, K.O. Stetter, J.A. Lake, M. Gellert, R. Krah, and S.A. Kozyavkin, *Nature* **364**, 735–736 (1993).
89. A.I. Slesarev, J.A. Lake, K.O. Stetter, M. Gellert, and S.A. Kozyavkin, *J. Biol. Chem.* **269**, 3295–3303 (1994).
90. P. Fusi, G. Tedeschi, A. Aliverti, S. Ronchi, P. Tortora, and A. Guerritore, *Eur. J. Biochem.* **211**, 305–310 (1993).
91. X. Mai and M.W.W. Adams, *J. Bacteriol.* **178**, 5897–5903 (1996).
92. C. Purcarea, V. Simon, D. Prieur, and G. Herve, *Eur. J. Biochem.* **236**, 189–199 (1996).
93. C. Cicicopol, J. Peters, J. Kellerman, and W. Baumeister, *FEBS Lett.* **356**, 345–350 (1994).
94. G. Andreotti, M.V. Cubellis, G. Nitti, G. Sannia, X. Mai, M.W. Adams, and G. Marino, *Biochim. Biophys. Acta* **1247**, 90–96 (1995).
95. C. Purcarea, G. Erauso, D. Prieur, and G. Herve, *Microbiology* **140**, 1967–1975 (1994).
96. G. Cacciapuoti, M. Porcelli, M. Carteni-Farina, A. Gambacorta, and V. Zappia, *Eur. J. Biochem.* **161**, 263–271 (1986).
97. J. Breitung, G. Borner, S. Scholz, D. Linder, K.O. Stetter, and R.K. Thauer, *Eur. J. Biochem.* **210**, 971–981 (1992).
98. S.W.M. Kengen, J.E. Tuininga, F.A.M. de Bok, A.J.M. Stams, and W.M. de Vos, *J. Biol. Chem.* **270**, 30453–30457 (1995).
99. M.J. Peak, J.G. Peak, F.J. Stevens, J. Blamey, X. Mai, Z.H. Zhou, and M.W. Adams, *Arch. Biochem. Biophys.* **313**, 280–286 (1994).
100. A. Guagliardi, L. Cerchia, S. Bartolucci, and M. Rossi, *Protein Sci.* **3**, 1436–1443 (1994).
101. Y. Sako, K. Takai, A. Uchida, Y. Ishida, *FEBS Lett.* **392**, 148–152 (1996).
102. G. Cacciapuoti, M. Porcelli, C. Bertoldo, M. De Rosa, and V. Zappia, *J. Biol. Chem.* **269**, 24762–24769 (1994).
103. A. Heltzel, E. T. Smith, Z.H. Zhou, J.M. Blamey, and M.W.W. Adams, *J. Bacteriol.* **176**, 4790–4793 (1994).
104. B.S. Mroczkowski, A. Huvar, W. Lernhardt, K. Misono, K. Nielson, and B. Scott, *J. Biol. Chem.* **269**, 13522–13528 (1994).
105. K.A. Laderman, K. Asada, T. Uemori, H. Mukai, Y. Taguchi, I. Kato, and C.B. Anfinsen, *J. Biol. Chem.* **268**, 24402–24407 (1993).
106. J. DiRuggiero and F.T. Robb, *Appl. Environ. Microbiol.* **61**, 159–164 (1995).
107. S.B. Halio, I.I. Blumentals, S.A. Short, B.M. Merrill, and R.M. Kelly, *J. Bacteriol.* **178**, 2605–2612 (1996).
108. V.J. Harwood, J.D. Denson, K.A. Robinson-Bidle, and H.J. Schreier, *J. Bacteriol.* **179**, 3613–3618 (1997).
109. J.M. Muir, R.J. Russell, D.W. Hough, and M.J. Danson, *Protein Eng.* **8**, 583–592 (1995).
110. W.G. Voorhorst, R.I. Eggen, E.J. Luesink, and W.M. de Vos, *J. Bacteriol.* **177**, 7105–7111 (1995).

111. D. Hess, K. Kruger, A. Knappik, P. Palm, and R. Hensel, *Eur. J. Biochem.* **233**, 227–237 (1995).

112. A. Rudiger, P.L. Jorgensen, and G. Antranikian, *Appl. Environ. Microbiol.* **61**, 567–575 (1995).

113. P. Baumann and S. Jackson, *Proc. Natl. Acad. Sci. USA* **933**, 6726–6730 (1996).

114. T. Rowlands, P. Baumann, and S.P. Jackson, *Science* **264**, 1326–1328 (1994).

115. T. Imanaka, S. Lee, M. Takagi, and S. Fujiwara, *Gene* **164**, 153–156 (1995).

116. N. Rashid, M. Morikawa, and T. Imanaka, *Gene* **166**, 139–143 (1995).

117. Y. Tachibana, M.M. Leclere, S. Fujiwara, M. Takagi, and T. Imanaka, *J. Ferment. Bioeng.* **82**, 224–232 (1996).

118. N. Rashid, M. Morikawa, and T. Imanaka, *Mol. Gen. Genet.* **253**, 397–400 (1996).

119. Z. Yan, S. Fujuwara, K. Kohda, M. Takagi, and T. Imanaka, *Appl. Environ. Microbiol.* **63**, 785–789 (1997).

120. A. Bouyoub, G. Barbier, P. Forterre, and B. Labedan, *J. Mol. Biol.* **261**, 144–154 (1996).

121. T. Uemori, Y. Ishino, H. Doi, and I. Kato, *J. Bacteriol.* **177**, 2164–2177 (1995).

122. F.B. Perler, D.G. Comb, W.E. Jack, L.S. Moran, B. Qiang, R.B. Kucera, J. Benner, B.E. Slatko, D.O. Nwankwo, S.K. Hempstead, C.K.S. Carlow, and H. Jannasch, *Proc. Natl. Acad. Sci. USA* **89**, 5577–5581 (1992).

123. M.W. Southworth, H. Kong, R.B. Kucera, J. Ware, H.W. Jannasch, and F. Perler, *Proc. Natl. Acad. Sci. USA* **93**, 5281–5285 (1996).

124. J.S. Yu and K.M. Noll, *FEMS Microbiol. Lett.* **131**, 307–312 (1995).

125. C. Vieille, J.M. Hess, R.M. Kelly, and J.G. Zeikus, *Appl. Environ. Microbiol.* **61**, 1867–1875 (1995).

126. V. Zverlov, K. Piotukh, O. Dakhova, G. Velikodvorskaya, and R. Borris, *Microbiol. Biotechnol.* **45**, 245–247 (1996).

127. M. Van de Casteele, C. Legrain, V. Wilquet, and N. Glansdorff, *Gene* **158**, 101–105 (1995).

128. R. Ostendorp, W. Liebl, H. Schurig, and R. Jaenicke, *Eur. J. Biochem.* **216**, 709–715 (1993).

129. A. Tomschy, R. Glockshuber, and R. Jaenicke, *Eur. J. Biochem.* **214**, 43–50 (1993).

130. R.V. Swanson, M.G. Sanna, and M.I. Simon, *J. Bacteriol.* **178**, 484–489 (1996).

131. B. Darimont and R. Sterner, *EMBO J.* **13**, 1772–1781 (1994).

132. J. Tong and J.G. Wetmur, *J. Bacteriol.* **178**, 2695–2700 (1996).

133. R. Sterner, A. Dahm, B. Darimont, A. Ivens, W. Liebl, and K. Kirshner, *EMBO J.* **14**, 4395–4402 (1995).

134. W. Liebl, R. Feil, J. Gabelsberger, J. Kellermann, and K.H. Schleifer, *Eur. J. Biochem.* **207**, 81–88 (1992).

135. W. Liebl, I. Stemplinger, and P. Ruile, *J. Bacteriol.* **179**, 941–948 (1997).

136. N. Beaucamp, R. Ostendorp, H. Schurig, and R. Jaenicke, *Prot. Peptide Lett.* **2**, 281–286 (1995).

137. J. Gabelsberger, W. Liebl, and K.H. Schleifer, *FEMS Microbiol. Lett.* **109**, 131–138 (1993).

138. W. Liebl, P. Ruile, K. Bronnenmeier, K. Riedel, F. Lottspeich, and I. Greif, *Microbiology* **142**, 2533–2542 (1996).

139. P.-J. Lee and A.M. Stock, *J. Bacteriol.* **178**, 5579–5585 (1996).

140. A.M. Sanangelantoni, P. Cammarano, and O. Tiboni, *Microbiology* **142**, 2525–2532 (1996).

141. D.J. Saul, L.C. Williams, R.A. Reeves, M.D. Gibbs, and P.L. Bergquist, *Appl. Environ. Microbiol.* **61**, 4110–4113 (1995).

142. M. Yohda, I. Endo, Y. Abe, T. Ohta, T. Iida, T. Maruyama, and Y. Kagawa, *J. Biol. Chem.* **271**, 22017–22021 (1996).

143. J.G. McAfee, S.P. Edmondson, P.K. Datta, J.W. Shriver, and R. Gupta, *Biochem.* **34**, 10063–10077 (1995).

144. S. Ohnuma, M. Suzuki, and T. Nishino, *J. Biol. Chem.* **269**, 14792–14797 (1994).

145. T. Kath, R. Schmid, and G. Schafer, *Arch. Biochem. Biophys.* **307**, 405–410 (1993).

146. T. Wakagi, T. Fujii, and T. Oshima, *Biochem. Biophys. Res. Comm.* **225**, 489–493 (1996).

147. M.L. Tutino, A. Tosco, G. Marino, and G. Sannia, *Biochem. Biophys. Res. Comm.* **230**, 306–310 (1997).

148. M. Moracci, A. La Volpe, J.F. Pulitzer, M. Rossi, and M. Ciaramella, *J. Bacteriol.* **174**, 873–882 (1992).

149. M. Moracci, M. Ciaramella, R. Nucci, L.H. Pearl, I. Sanderson, A. Tincone, and M. Rossi, *Biocatalysis* **11**, 89–103 (1994).

150. R. Cannio, D. De Pacale, M. Rossi, and S. Bartolucci, *Biotechnol. Appl. Biochem.* **19**, 233–244 (1994).

151. C.E. Jones, T.M. Fleming, D.A. Cowan, J.A. Littlechild, and P.W. Piper, *Eur. J. Biochem.* **233**, 800–808 (1995).

152. G. Ianniciello, M. Masullo, M. Gallo, P. Arcari, and V. Bocchini, *Biotechnol. Appl. Biochem.* **23**, 41–45 (1996).

153. K. Kobayashi, M. Kato, Y. Miura, M. Kettoku, T. Komeda, and A. Iwamatsu, *Biosci. Biotech. Biochem.* **60**, 1882–1885 (1996).

154. M.I. Arnone, L. Birolo, M.V. Cubellis, G. Nitti, G. Marino, and G. Sannia, *Biochim. Biophys. Acta* **1160**, 206–212 (1992).

155. D.A. Bochar, J.R. Brown, W.F. Doolittle, H.-P. Klenk, W. Lam, M.E. Schenk, C.V. Stauffacher, and V.W. Rodwell, *J. Bacteriol.* **179**, 3632–3638 (1997).

156. D. Yue, N. Maizels, and A.M. Weiner, *RNA* **2**, 895–908 (1996).

157. A. Lehmacher and R. Hensel, *Mol. Gen. Genet.* **242**, 163–168 (1994).

158. B. Mukhopadhyay, E. Purwantini, T.D. Pihl, J.N. Reeve, and L. Daniels, *J. Biol. Chem.* **270**, 2827–2832 (1995).

159. R. Tabassum, K.M. Sandman, and J.N. Reeve, *J. Bacteriol.* **174**, 7890–7895 (1992).

160. S. Shima, D.S. Weiss, and R.K. Thauer, *Eur. J. Biochem.* **230**, 906–913 (1995).

161. M. Aoshima, A. Yamagishi, and T. Oshima, *Arch. Biochem. Biophys.* **336**, 77–85 (1996).

162. R.M. Kelly and M.W.W. Adams, *Antonie van Leeuwenhoek* **66**, 247–270 (1994).

163. M.W. Bauer, S.B. Halio, and R.M. Kelly, in M.W.W. Adams ed., *Advances in Protein Chemistry* 48, Academic Press, San Diego, 1996, pp. 271–310.

164. C. Leuschner and G. Antranikian, *W. J. Microbiol. Biotechnol.* **11**, 95–114 (1995).

165. A. Sunna, M. Moracci, M. Rossi, and G. Antranikian, *Extremophiles* **1**, 2–13 (1997).

166. C.R. Woese, O. Kandler, and M.L. Wheelis, *Proc. Natl. Acad. Sci. USA* **87**, 4576–4579 (1990).

167. M.W.W. Adams and A. Kletzin, in M.W.W. Adams ed., *Advances in Protein Chemistry* 48, Academic Press, San Diego, 1996, pp. 101–173.

168. S.W.M. Kengen, A.J.M. Stams, and W.M. de Vos, *FEMS Microbiol. Rev.* **18**, 119–137 (1996).

169. C.G. Miller, in F. Neidhardt ed., E. coli and Salmonella *typhimurium. Cellular and Molecular Biology*, vol. 1, American Society for Microbiology, Washington D.C., 1987, pp. 680–691.

170. *Enzymes at Work*, product literature from Novo Nordisk A/S, Bagsvaerd, Denmark, 1995.

171. O. Almarsson and A.M. Klibanov, *Biotechnol. Bioeng.* **49**, 87–92 (1996).

172. H. Kise, A. Hayakawa, and H. Noritomi, *J. Biotechnol.* **14**, 239–254 (1990).

173. H.L. Sham, C. Zhao, K.D. Stewart, D.A. Betebenner, S. Lin, C.H. Park, X.P. Kong, W. Rosenbrook, T. Herrin, D. Madigan, S. Vasavanonda, N. Lyons, A. Molla, A. Saldivar, K.C. Marsh, E. McDonald, N.E. Wideburg, J.F. Denissen, T. Robins, D.J. Kempf, J.J. Plattner, and D.W. Norbeck, *J. Med. Chem.* **39**, 392–397 (1996).

174. B.F. Sloane, *Nat. Biotechnol.* **14**, 826–827 (1996).

175. K. Peek, S.A. Wilson, M. Prescott, and R.M. Daniel, *Ann. N. Y. Acad. Sci.* **672**, 471–477 (1992).

176. H. Connaris, D.A. Cowan, and R.J. Sharp, *J. Gen. Microbiol.* **137**, 1193–1199 (1991).

177. L.J. Snowden, I.I. Blumentals, and R.M. Kelly, *Appl. Environ. Microbiol.* **58**, 1134–1140 (1992).

178. P.M. Hicks and R.M. Kelly, in F. Woessner, N. Rawlings, and A. Barret ed., *Handbook of Proteolytic Enzymes*, Academic Press, London, 1999, pp. 1579–1581.

179. R. Eggen, A. Geerling, J. Watts, and W.M. de Vos, *FEMS Microbiol. Lett.* **71**, 17–20 (1990).

180. J. Lowe, D. Stock, B. Jap, P. Zwickl, W. Baumeister, and R. Huber, *Science* **268**, 533–539 (1995).

181. M. Morikawa, Y. Izawa, N. Rashid, T. Hoaki, and T. Imanaka, *Appl. Environ. Microbiol.* **60**, 4559–4566 (1994).

182. B. Henrissat, *Biochem. J.* **280**, 309–316 (1991).

183. S.H. Brown and R.M. Kelly, *Appl. Environ. Microbiol.* **59**, 2614–2621 (1993).

184. G.P. Hazlewood and H.J. Gilbert, in M.P. Coughlan and G.P. Hazlewood ed., *Hemicellulose and Hemicellulases*, Portland Press, Chapel Hill, N.C., 1993, pp. 103–126.

185. S.W.M. Kengen and A.J.M. Stams, *Biocatalysis* **11**, 79–88 (1994).

186. J.D. Bok, S.K. Goers, and D.E. Eveleigh, *ACS Symp. Ser.* **566**, 54–65 (1994).

187. H.D. Simpson, U.R. Haufler, and R.M. Daniel, *J. Biochem.* **277**, 413–417 (1991).

188. C. Winterhalter, P. Heinrich, A. Candussio, G. Wich, and W. Liebl, *Mol. Microbiol.* **15**, 431–444 (1995).

189. A. Sunna, J. Puls, and G. Antranikian, *Biotechnol. Appl. Biochem.* **24**, 177 (1996).

190. C.M. McCutchen, G.D. Duffaud, P. Leduc, A.R.H. Petersen, A. Tayal, S.A. Khan, and R.M. Kelly, *Biotechnol. Bioeng.* **52**, 332–339 (1996).

191. M.W.W. Adams, *Annu. Rev. Microbiol.* **47**, 627–658 (1993).

192. M.W. Day, B.T. Hsu, L. Joshua-Tor, J.-B. Park, Z.H. Zhou, M.W.W. Adams, and D.C. Rees, *Prot. Sci.* **1**, 1494–1507 (1992).

193. M.K. Chan, S. Mukund, A. Kletzin, M.W.W. Adams, and D.C. Rees, *Science* **267**, 1463–1469 (1995).

194. K.S.P. Yip, T.J. Stillman, K.L. Britton, P.J. Artymiuk, P.J. Baker, S.E. Sedelnikova, P.C. Engel, A. Pasquo, R. Chiaraluce, V. Consalvi, R. Scandurra, and D.W. Rice, *Structure* **3**, 1147–1158 (1995).

195. I. Korndorfer, B. Steipe, R. Huber, A. Tomschy, and R. Jaenicke, *J. Mol. Biol.* **246**, 511–521 (1995).

196. K.A. Dill, *Biochemistry* **29**, 7133–7155 (1990).

197. J.F. Carpenter, J.S. Clegg, J.H. Crowe, and C.N. Somero, *Cryobiology* **30**, 2001–2041 (1993).

198. R.J. Ellis and S.M. van der Vies, *Annu. Rev. Biochem.* **60**, 321–347 (1991).

199. D.A. Parsell and S. Lindquist, in R.I. Morimoto, A. Tissieres, and C. Georgopoulos eds., *The Biology of Heat Shock Proteins and Molecular Chaperones*. Cold Spring Harbor Laboratory Press, Plainview, N.Y., 1994, pp. 457–494.

200. G. Zellner and H. Kneifel, *Arch. Microbiol.* **159**, 472–476 (1993).

201. R. Hensel and I. Jakob, *System. Appl. Microbiol.* **16**, 742–745 (1994).

202. S. Scholz, J. Sonnenbichler, W. Schafer, and R. Hensel, *FEBS Lett.* **306**, 239–242 (1992).

203. R.A. Ciulla, S. Burggraf, K.O. Stetter, and M.F. Roberts, *Appl. Environ. Microbiol.* **60**, 3660–3664 (1994).

204. L.O. Martins and H. Santos, *Appl. Environ. Microbiol.* **61**, 3299–3303 (1995).

205. O.C. Nunes, C.M. Manaia, M.S. DaCosta, and H. Santos, *Appl. Environ. Microbiol.* **61**, 2351–2357 (1995).

206. J.D. Trent, *FEMS microbiol. Rev.* **18**, 249–258 (1996).

207. J.D. Trent, J. Osipiuk, and T. Pinkau, *J. Bacteriol.* **172**, 1478–1484 (1990).

208. J.F. Holden and J.A. Baross, *J. Bacteriol.* **175**, 2389–2843 (1993).

209. J.D. Trent, M. Gabrielsen, B. Jensen, J. Neuhard, and J. Olsen, *J. Bacteriol.* **176**, 6148–6152 (1994).

210. C.J. Han, S.H. Park, and R.M. Kelly, *Appl. Environ. Microbiol.* **63**, 2391–2396 (1997).

211. H.K. Kagawa, J. Osipiuk, N. Maltsev, R. Overbeek, E. Quaite-Randall, A. Joachimisk, and J.D. Trent, *J. Mol. Biol.* **253**, 712–725 (1995).

212. B.M. Phipps, A. Hoffmann, K.O. Stetter, and W. Baumeister, *EMBO J.* **10**, 1711–1722 (1991).

213. B.M. Phipps, D. Typke, R. Hegerl, S. Volker, A. Hoffmann, K.O. Stetter, and W. Baumeister, *Nature* **361**, 475–477 (1993).

214. T. Waldmann, E. Nimmesgrern, M. Nitsch, J. Peters, G. Pfeifer, S. Müller, J. Kellermann, A. Engel, F.-U. Hartl, and W. Baumeister, *Eur. J. Biochem.* **227**, 848–856 (1995).

215. J.D. Trent, E. Nimmesgern, J.S. Wall, F.-U. Hartl, and A. Horwich, *Nature* **354**, 490–493 (1991).

216. F.T. Robb, J.B. Park, and M.W.W. Adams, *Biochim. Biophys. Acta* **1120**, 267–272 (1992).

217. W.M. Barnes, *Proc. Natl. Acad. Sci. USA* **91**, 2216–2220 (1994).

218. Y.B. Tewari and R.N. Goldberg, *Appl. Biochem. Biotechnol.* **11**, 17–24 (1985).

219. E.J. VanDamme and W. Soetaert, *FEMS Microbiol. Rev.* **16**, 163–186 (1995).

220. A. Tayal, R.M. Kelly, and S.A. Khan, *J. Soc. Petr. Eng.*, SPE 38432 (1997).

See also ENZYMES, DIRECTED EVOLUTION; PROTEIN SECRETION, *SACCHAROMYCES CEREVISIAE*; THERMOPHILIC MICROORGANISMS.

ENZYMES, FOR FLAVOR PRODUCTION

PETER S. J. CHEETHAM
Zylepsis, Ltd.
Ashford, United Kingdom

KEY WORDS

Bioactive molecules
Biocatalysis
Downstream processing
Fermentation
Isolated enzymes
Natural flavoring substances
Nature-identical flavoring substances
Receptor

OUTLINE

INTRODUCTION

Flavor and aroma chemicals are exceptionally bioactive molecules that exert their very characteristic taste and smell effects even when present at very low concentrations. This is because of their highly selective interactions with receptors in the mouth and nose. It is this high activity, in combination with the usually very pleasant, evocative, and identity-conferring properties that make flavor and aroma chemicals so valuable. In particular their tastes and aromas are very characteristic of particular foods, beverages, flowers, and so forth. This is made possible because of the many combinations and permutations of different flavor and aroma chemical molecules, with different characteristic mixtures found in each particular natural source. Hence complex mixtures of flavor chemicals are generally much more valuable because of their more rounded tastes. These organoleptic characteristics are generally very precisely determined by the exact molecular structures of the chemicals involved. As a result, even slight modifications, such as another hydroxyl group, a *trans* rather than a *cis* double bond, or an *R* rather than an *S* isomer, can completely change or abolish a molecule's taste and smell properties, both in terms of its character and identity, and also the threshold concentration at which it can be detected.

Many flavor and aroma chemicals are produced naturally in foods by the action of endogenous enzymes and/or by naturally occurring microorganisms such as those contributing to the tastes and smell of bread, cheese, wine, beer, and many oriental products such as soy. The need to make processed foods with taste and smell characteristics redolent of traditional products creates a derived demand for natural flavors, which in turn, creates a need for processes to make them. This is particularly the case for many highly processed foods that lose many of their volatile or heat-unstable flavor chemicals during processing. It is very appropriate to use modern enzyme and microbial technologies to manufacture flavor and fragrance materials, particularly as biocatalysis techniques provide the opportunity to carry out new reactions, and especially more selective reactions not possible with conventional organic chemical synthesis.

Use of biocatalysis methods is especially advantageous for the production of flavors because, when carried out rigorously, such techniques allow them to be described as "natural," and so command the premium prices conferred on them as a result of the insistent consumer demand for natural or "green" food ingredients. This advantage is, however, less strong for aroma chemicals because the consumer demand for natural fragrance materials is, so far, less well established. Biocatalysis processes are also very useful when alternative methods such as chemical synthesis or agricultural production are either not possible or not cost-effective. Perhaps the biggest challenge is in developing processes that can reliably and reproducibly deliver products at cost-effective prices.

The normal range of isolated enzyme, fermentation, and bioconversion processes are used. However, biocatalytic processes for flavors and fragrances are different from those in other areas of biotechnology in that they generally do not utilize genetic engineering techniques. This is because the products of genetic engineering don't fully meet the stringent consumer criteria for natural products. However, in all other respects this article on the use of bioprocessing to make flavors has been written in the context of,

and with reference to, other successful bioprocesses used in the drug, agrochemical, food, and chemicals industries.

Ultimately the goal is products that are of good quality, are cost-effective and safe to use, and that meet customer demands. So far this combination of market pull, the consumer demand for natural flavors, and the technical push of enzyme and microbial technology has been successful in bringing a wide range of natural flavours to market, produced by enzyme, bioconversion, and fermentation techniques—whichever is most effective for each particular flavor material. Notable exceptions include a microbially produced natural vanilla flavor, and indeed any product using plant cell culture as a manufacturing technology. For reviews see Refs. 1–3.

The world market for flavor sales has been estimated at $4.5 billion in 1994 (4,5); this figure excludes the very considerable in-house production and use of flavors by some major food and beverage manufacturers. Europe still constitutes 40% of these sales (ca. $1.8 billion per year), followed by the United States, with sales in Asia having expanded most rapidly and with eastern European markets offering potential for the future. The major uses of flavors are in soft drinks (29%), followed by savory flavors such as meat and snack products (26%), and then applications in many other categories of products such as alcoholic beverages.

Major commercial trends in the flavor industry include globalization, consolidation, compliance, coping with the pressure of flavor costs, and customer service. Globalization trends are especially aimed at gaining a commercial presence in emerging markets such as China. Consolidation is proceeding apace as companies join together and are acquired so as to create the critical mass in research and development, purchasing, sales, and so forth in order to compete worldwide and against larger competitors. Globalization and consolidation trends combine as the smaller local companies with particular skills and niche markets are purchased by multinational flavor companies. This often results in a rationalization of manufacturing and fewer but much larger flavor-producing plants.

Compliance with regulatory, legislative, and customer requirements is increasing rapidly, and the resources that must be devoted to compliance issues are increasing much more rapidly than the growth in sales revenues. This is despite growth in flavor sales of 6–7% per year over the last decade, which is significantly higher than for the economy in general.

Due to the pressure on large food and beverage manufactures to reduce the contribution of flavor ingredients to the overall cost of their final products, and increasing competition between different flavor manufacturers, there has been an interesting trend to provide a more valuable service, for instance, in the form of more highly formulated flavors that are easier for the end-manufacturer to use. Other trends are to focus on the absolutely key "character-conferring" flavor chemicals, especially when they are sufficiently powerful to allow use at only low concentrations in the final product; to search for generic manufacturing processes whereby a small range of different flavors, such as C6 alcohols and aldehydes, can all be made using the same manufacturing plant; and to anticipate emerging consumer demands, such as for ethnic food influences or particular types of healthy ingredients.

Despite recent consolidation, the flavor ingredient supply market is still very fragmented, with the top 10 flavor companies having together still only 54% of the 1994 market of $4.5 billion in sales, and with the 5 leading companies, IFF, Quest, Givaudan, Haarmann & Reimer, and Firmenich, having 14, 10, 9, and 7% market shares, respectively (Table 1). This means that even these market leader companies have turnovers that are much smaller than their biggest customers such as Nestlé, Unilever, and Sara Lee, who are among the largest companies in the world. Another indication of economic importance is that European Community estimates show that eventually the impact of biotechnology in food and related businesses could be greater than in pharmaceuticals.

Any successful product requires a strong market demand and also a cost-effective manufacturing process to meet that demand. This article concentrates on the latter, but it is important to touch on some aspects of market demand first. Any bioprocess to make a natural flavor or aroma chemical exists only because of the market need for that flavor or aroma chemical. In turn the need for the flavor or aroma chemical arises as a result of a derived demand from the consumer's choice of products that contain the flavor or aroma chemical in a fully formulated and often branded product. The current world consumption of flavor and fragrance products (6) is summarized in Figure 1, and Hausler and Munch (7) report more than 65% of all

Table 1. Estimated Flavor and Fragrance Product Sales of Leading Companies

Company	Millions of dollars[a]	% of total industry sales
International Flavours and Fragrances (U.S.)	1,315	13.6
Givaudan-Roure (Switzerland)	980	10.1
Quest International (U.K./Netherlands)	930	9.6
Bayer (Haarmann & Reimer, & Florasynth, U.S. and Germany)	850	8.8
Firmenich (Switzerland)	708	7.3
Takasago International (Japan)	500	5.2
Bush Boken Allen (U.S.)	375	3.9
T. Hasegawa Co. (Japan)	356	3.7
Dragoco Gerberding and Co. (Germany)	260	2.7
Tastemaker (U.S.) (now acquired by Givaudan-Roure)	210	2.2
Universal Flavours (U.S.)	180	1.9
V. Mane Fils (France)	160	1.7
Ogawa & Co. (Japan)	158	1.6
Robertet (France)	130	1.3
Subtotal	7,112	73.4
Other companies	2,575	26.6
Total	9,687	100

Source: Ref. 6.

[a]Flavor and fragrance sales only; does not include sales of nonflavor and fragrance products such as food additives, speciality chemicals, juices, ice cream, and beverage bases, into which many of these companies have expanded as a logical diversification of their original product ranges.

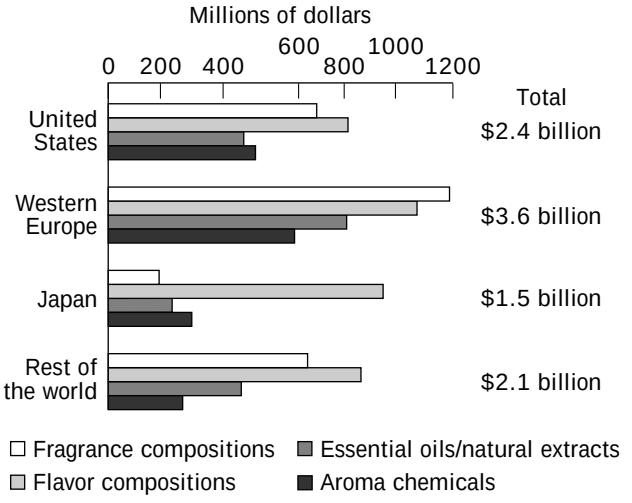


Figure 1. World consumption. Estimated consumption of flavor and aroma products, 1994.

flavoring ingredients used commercially in the United States are labeled as natural and have a food market potential exceeding $9 billion. This estimates the worldwide market for flavors alone at approaching $4 billion per year, and the total for all flavor, aroma, and essential oil ingredients at almost $10 billion per year. Another illustration is the high percentage of naturally flavored foods sold in Europe and the United States that contain natural flavor chemicals (8) (Table 2).

The economic value of biocatalytic processes is not just in terms of the materials produced, but also in the very much higher economic value of the flavored end products, which include branded consumer goods that depend on availability of flavors of the right quality and cost.

HISTORY

Flavors have a long history. The first products involving the bioproduction of flavors included beer, wine, and bread; yogurt made using lactobacilli; cheeses matured using a variety of microorganism including fungi. Also in Asia soy, miso, tempeh, and other products made using *Aspergillus oryzae* and other strains. Very often the success of these processes depends on the use of fast-growing anaerobes that accumulate products that discourage the growth of contaminants, such as ethanol or lactic acid, and also supplementation with additives such as hops and spices. This first phase of products was refined over many millennia.

Many of the traditional flavors and flavored products still set the pattern for modern products because of deeply ingrained consumer preferences. For instance, beer and cheese were first made at least 6,000 years ago, that is some 250 generations ago. So there has been plenty of time for humans to develop inate flavor preferences, as well as preferences for other qualities such as color, clarity, foaming behavior, and mouthfeel properties. Some of these additional properties are the direct result of microbial biochemistry, for instance, glycerol is important in conferring mouthfeel, and it is produced from dihyroxyacetone by yeast.

It was not until the nineteenth century and the rise of a technological culture that second-phase products arrived in the form of much more chemically defined flavor products such as monosodium glutamate, citric acid, lactic acid, acetic acid in the form of improved vinegar producing processes, and also the products of a range of microbial enzymes. These processes required strain selection, development of special growth media, and the maintenance of hygienic or sterile conditions.

In 1837 Liebig and Wöhler identified the first flavor compound, benzaldehyde. The isolation and synthesis of vanillin by Tiemann, Haarmann, and Reimer in 1874–1876 began the modern flavor industry, and microbially produced flavors were first described by Omelianski in 1923. Also, once the technology for producing enzymes and cells had become widespread, the use of these enzymes and cells as biocatalysts in biotransformation processes became possible.

Some other indications of the high sensory effects of flavors and the impact that biomanufacturing has already had are the estimates that only about 400 aroma chemicals are manufactured in quantities of greater than 1 tonne per year, and only 68 materials are reported to be used at a rate greater than 3 tonne per year, but that some 50–100 flavor materials are reported to be made by bioprocesses (9). Several key considerations need to be recognized because they strongly influence the type of processes used (see following).

Table 2. Proportions of Natural Aromas (%) in Flavored Foods

Product group	Europe	U.S.
Nonalcoholic beverages	90	83
Dairy products	50	75
Savory products	80	80

Source: Ref. 8.

Note: It has been estimated that, in 1990, 30 metric tons of aroma compounds was produced by biotechnological processes, sufficient to flavor several million tons of food products.

SOME CHARACTERISTIC FEATURES OF FOOD BIOTECHNOLOGY

Raw materials often contain endogenous enzyme activities and are often physically and chemically complex and diverse. They include oils and fats, proteins, and carbohydrates. Many are biopolymers that are usually produced from substainable agricultural sources, rather than petrochemicals, and contain many chiral molecules.

Products have the following characteristics:

- Have a wide range of value, from £ several 000/kg for some flavors to less than £100/tonne for animal feed (compare the prices of citric acid and vitamin B_{12}).
- Product value depends on consumer satisfaction (flavor, mouthfeel, etc.), which can vary with age, culture, and so forth.

- Genetic engineering, especially genetically modified organisms, have poor consumer acceptability.
- Flavor/aroma chemicals with comparatively similar chemical structures usually have quite different characteristics and intensities. For example, both diacetyl (2,3-butadione), which gives a buttery taste to dairy products, and hop-α-iso acids, which gives beer its bitter taste, are both vicinal diketones.
- It is not possible to record taste and smell sensations in the same way as sight and hearing sensations, and so the flavor and aroma quality of a material cannot be accurately assessed on-line, or even easily off-line. In a process, only the proportions of the various chemicals that contribute to the overall taste and smell can be analyzed. True flavor quality can be properly assessed only by an expert flavorist or trained taste panel, or on a gross basis by the reaction of consumers to a particular flavored product. However, some progress toward a taste–aroma sensor is being made, with olfactory neurones being shown to respond to physiological levels of volatiles in culture media (10) and a bovine binding protein being coupled to a resin (11). A functional odor receptor has been expressed in the noses of rats and shown to exhibit specificity toward just four C7–C10 aliphatic aldehydes and with the greatest response being to octanal (12). This will hopefully lead to a proper working knowledge of how the 500–1,000 odor receptors can recognize the thousands of distinct aromas, probably by acting in combinational mechanisms, with any one receptor responding to several different aroma chemicals and any one such aroma chemical capable of triggering several receptors.
- Flavor and aroma chemicals are often very potent materials, well suited to small-scale processing, and are usually diluted considerably when added to the final product. Their flavor value is defined as the concentration of flavor used, divided by its flavor threshold (that is, the concentration at which it can be just perceived). Because of the potent activities of flavors, care must be taken in multipurpose processing operations that produce a range of flavor products to ensure that there is no cross-contamination of one particular flavor by residues of another flavor produced in the preceding run.
- The actual perceived flavor of a molecule depends to some degree on the physical form in which it is eaten, thus the same flavor molecule formulated in products based on fat or oil, carbohydrate, water (as in carbonated beverage), and ethanol (as in an alcoholic beverage) can have quite different tastes, for instance due to solubility effects and the influence of mouthfeel and other factors on taste perception.
- Often one single flavor chemical will substantially define the characteristic of a particular flavor or aroma chemical, such as benzaldehyde for almond taste and vanillin for vanilla. These character-impact flavor chemicals can often command particularly high prices because of their high flavor value.
- Even quite closely related molecules can have very different flavor sensations, for instance geraniol (flower aroma) can be microbially converted into 6-methyl heptene-2-one, which is a component of tomato flavor.
- The natural or nature-identical status of flavors is important and deserves a more detailed consideration because of the impact of such definitions on the design and operation of flavor bioprocesses.
- Flavor and aroma chemicals must meet both industry and government standards for safety such as those provided by the International Organization of the Flavor Industry and the U.S. Food and Drug Administration, respectively. Increasingly, kosher and halal standards are also being used as criteria of good quality. For a more detailed discussion of the issues involved see work by Manley (9).
- Very often different isomeric forms of the same molecule have quite different flavor and aroma sensations. For instance (*S*)-(+)- and (*R*)-(−)-linalool are described as sweet and lavender, respectively, *R* and (−)-*trans*-α-ionone are violet-fruity and woody, (*S*)-(+)- and (*R*)-(−)-carvone are caraway and spearmint, (+)- and (−)-*cis*-rose oxide are sweet and fruity, (+)- and (−)-nootkatone are grapefruit and woody, (+)- and (−)-limonene are orange and turpentine, and (+)- and (−)-*p*-menthene-8-thiol are fruity and grapefruit tasting, respectively. In addition marked differences in flavor threshold often occur. Thus the (+)- and (−)-nootkatones have odor thresholds of 0.6–1.0 ppm and 400–800 ppm, respectively, and the (*R*) and (*S*)-isomers of δ-decalactone have odor thresholds of 1.5 and 5.6 ppb, respectively (13).

Natural and nature-identical flavors are defined as follows. These definitions obviously very much influence the types of processing operations that can be used.

Natural flavoring substances. These are materials obtained by appropriate physical processes (including distillation and solvent extraction) or enzymatic or microbiological processes from material of vegetable or animal origin, either in the raw state or after processing for human consumption by traditional food-preparation processes (including drying, torrefaction, and fermentation).

Nature-identical flavoring substances. These are materials obtained by chemical synthesis or isolated by chemical processes and are chemically identical to a substance naturally present in material of vegetable or animal origin.

Market Need

In order to succeed, a product made by any process needs to meet market demand factors such as the following:

- The declining availability and quality of many traditional materials, especially when they cannot meet the demand of modern high-volume-, high-quality-specification mass-market products, for instance, some plant spices and colors

- Changes in the technologies used for cooking and food processing (e.g., microwaved foods, extruded products, and liposome delivery of cosmetic ingredients)
- Increasing size of low-cost mass markets, especially in branded and own-brand consumer products, detergents, air fresheners, washing powders, and so on
- Changes in consumer demands, for example, healthy eating trends that include preferences for low saturated fat, salt, sugar, azo dyes, high fiber, and vitamins
- Increased legislative concerns that amplify the drive toward safety and healthiness

These factors create opportunities for technical innovations such as the following:

- Making available alternative raw materials
- New reaction and processing techniques, as well as processes adapted to use new raw materials
- Improved analytical techniques including those useful for process monitoring and control, and improved formulation for better performance (e.g., encapsulation and knowledge of synergy effects)

If technical innovations are successful, the flavor will then be used in formulated products recommended by flavorists on the basis of availability and cost of raw materials, the organoleptic value of the product, and the product's physical properties that affect ease of formulation and so on.

A more detailed business analysis will include a consideration of the following factors:

- Applications
- Customer needs and product segmentation
- Volume of demand and likely longevity of demand, price, and especially likely profit margins
- Specifications for different grades of product
- Competitors, both current and prospective
- Raw materials prices and availabilities
- Processing costs
- Patent position, including freedom to patent and noninfringements position
- Likely time and investment required to reach first sales
- Return on investment, especially allowing for risk factors, regulatory and safety considerations, and strategic benefits such as other uses for the technology
- Manufacturing capacity required, including downstream processing and any special requirements such as continuous operation

The major flavor and fragrance companies, most of which have significant in-house bioresearch and bioprocess development capabilities, are as given in Table 1.

The dynamic nature of the industry is illustrated by recent changes. The Tastemaker company has recently been acquired by Hoffman LaRoche (Givaudan-Roure), Quest International has now been sold by Unilever to ICI, Bayer has integrated Haarmann and Reimer and Florasynth, and Danisco has purchased both Borthwicks and Becks, is an interesting example of a major food ingredients company expanding its flavor businesses, rather than the more usual diversification of flavor companies into food ingredients that will move them into the top dozen flavor companies (Table 1).

Bioprocesses to Make Flavor and Aroma Chemicals

The necessary bioprocesses to make flavor and aroma chemicals fall into four broad categories: *(1)* traditional processes, *(2)* processes for low-intensity flavors, *(3)* processes for high-intensity flavors, and *(4)* production of taste enhancers.

Traditional Processes. Traditional processes are used for products such as beer, wine, cheese, yogurt, and soy sauce. These generally involve microbial action on chemically complex plant- or sometimes animal-derived raw materials, with the formation of relatively low concentrations of flavor chemicals that may very well have good preservative properties as well. These can be described mostly as flavored foods produced by naturally occurring microorganisms and/or enzymes derived from them. For instance, in the beer-brewing process, flavor is derived both from the ingredients used, such as hops and malt, and from the processing: the action of the yeasts and also the heating that forms the butter-tasting α-iso acids from precursors present in the hops (Fig. 2). The traditional brewing process is typical of more recently developed bioprocesses that consist of three main stages; work-up of raw materials, the microbial reaction step, and product processing, which puts it into the form required by the customer. This last step may involve formulation with other flavor materials and/or bulking agents.

In many cases traditional processes have been industrialized and adapted. For instance, intense cheese flavors can be manufactured by the hydrolysis of casein with proprietary proteases and peptidases combined with the hydrolysis of butter fat by lipases and esterases. *Aspergillus niger* and *Rhizomuor miehei* are the preferred sources for lipases for cheese making.

Other bioprocesses for the dairy flavor chemicals diacetyl, acetaldehyde, and various lactones, as well as for accelerated milk souring, have been developed. From Asian countries there are also many excellent solid-state fermentation products (Table 3).

Low-Intensity Flavors. Manufacturing has already been established for a range of low-intensity flavoring materials such as high-fructose corn syrups (10 billion tons/year produced worldwide), citric and acetic acids (acidulants), protein hydrolysates, and cheeses and cheese flavors (14). The bioprocesses that have been developed for their large-scale manufacture are well described elsewhere: citric acid (15), milk souring (16), protein hydrolysates (17). More recently processes involving the sequential use of enzymes to modified cheese flavors have been developed (Fig. 3).

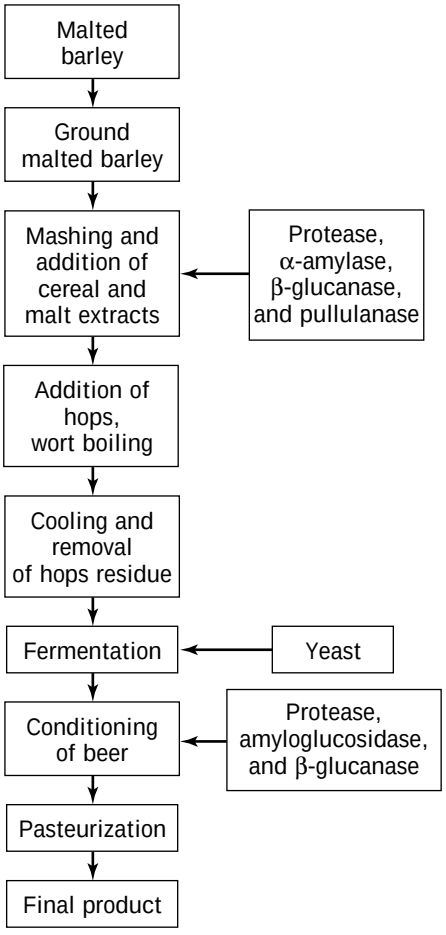


Figure 2. Beer manufacture. A flow diagram of the main operations. Note that pasteurization is omitted in the brewing of live beers. *Source*: Ref. 2.

High-Intensity Flavor Chemicals. More recent developments include the production of high-intensity flavors, such as methylketones (blue cheese flavor) and γ-decalactone (peach and other fruit flavors), in the form of pure natural chemicals by fermentation and bioconversion processes.

Taste Enhancers. Taste enhancers include materials such as monosodium glutamate (MSG) and mononucleotide flavor enhancers such as guanosine monophosphate (GMP) and inosine monophosphate (IMP), which are also produced as pure chemicals.

Important Considerations in Developing a Successful Process

It is important to appreciate that scaling up a process to meet cost and product specification targets can be at least as challenging as the original discovery research necessary to invent the process.

Much lower flavor concentrations need to be achieved for a flavor chemical when it is actually created in the food, in the way that flavor chemicals are produced in traditional products such as beer, cheese, and bread, rather than as a pure chemical, which is then added to the food or beverage at the required level. This does, of course, presume that any microorganisms involved are completely safe and acceptable, and also that flavor production will occur under the very different conditions existing in the food. Where processed foods are concerned, production in as pure and concentrated a form as possible is the most economic. The flavor can then be diluted as required for each end-use product and also combined with other flavors and flavor chemicals to give improved flavor quality. Some indication of the relationship between the concentration at which a flavor is produced and the concentration at which it is used in the final food or beverage product is given by the production-to-use ratio (Table 4), which can be illustrated by the following. If a material is produced at a concentration of 10 g L^{-1} and then used at 0.5 g L^{-1} (500 mg L^{-1}) in the product, or is present in a natural food or beverage at a concentration of 0.5 g L^{-1}, then its production-to-use ratio will be 20:1.

Key variables include the type of biocatalyst to be used, its performance, and in particular the medium used to carry out the reaction. The key features of the main methods are summarized as follows:

- *Fermentation production*. Fermentation is very suitable for multistep reactions. When using unstable enzymes and when cofactor regeneration is needed, it requires expensive fermenters and associated equipment. Operation is relatively expensive, because of the requirement for sterile conditions. The product is usually dilute and heavily contaminated with cells and by-products, and product quality can be variable due to batch-to-batch variations.
- *Use of isolated enzymes*. This method is especially suitable when extracellular stable enzymes are available, when single-step reactions are required that do not require cofactors (up to 10^3- to 10^4-fold regeneration of cofactors is necessary to enable them to be used cost-effectively, making synthetic reactions,

Table 3. Examples of Flavored Foods Obtained by Solid-State Fermentation

Product	Primary genus	Substrate	Further processing required?	Countries of origin
Soy sauce	*Aspergillus*	Soybean, wheat	Yes	China, Japan, southeast Asia
Sake	*Aspergillus*	Rice	Yes	Japan
Miso	*Aspergilus*	Rice, soybean	Yes	Japan, China, Taiwan, Indonesia
Tempeh	*Rhizopus*	Soybean	No	Indonesia, Surinam, New Guinea
Sufu	*Actinomucor*	Tofu	Yes	China
Ontjom	*Neurospora*	Peanut, coconut cakes	No	Indonesia

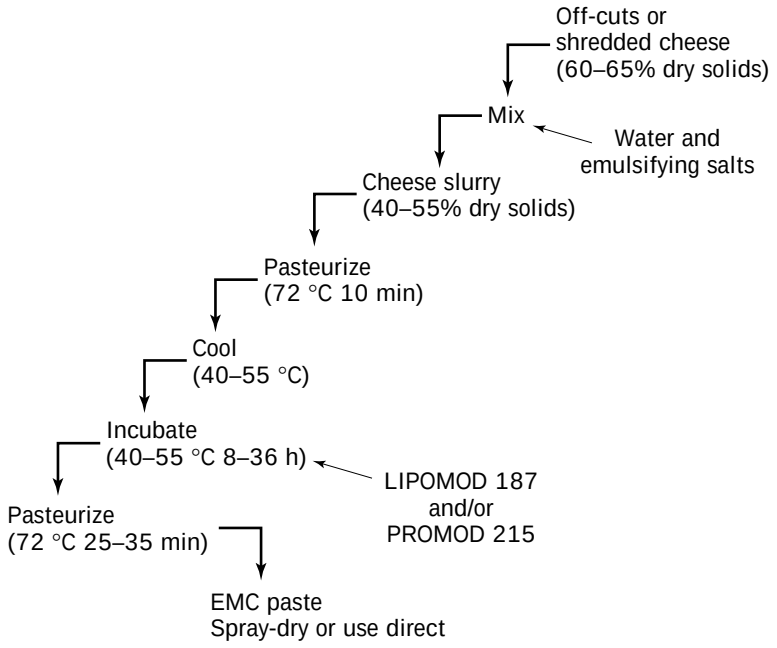


Figure 3. Outline of scheme for enzyme-modified cheese production. *Source*: Ref. 14.

Table 4. Production-to-Use Ratio

Traditional food products, e.g., beer, wine, bread, cheese	1.0 or greater if used as an ingredient in a more complex food
Industrialized forms of traditional products, e.g., enzyme-modified cheeses (EMCs) have flavor intensities about 20-fold greater than traditional cheeses	ca. 20
Traditional condiments, e.g., vinegar	ca. < 1.0
High-impact flavor chemicals, e.g., if γ-decalactone is made at 5 gL^{-1} and used at 100 ppm	50
trans-2-Hexenol	3,000 as compared with the concentration present in free-living plants

among others, very difficult to carry out using isolated enzyme preparations), and when hygienic but not sterile processing conditions can be used (although, of course, fermentation is necessary to produce the original cells and/or enzymes).

- Screening. In both cases, screening is a vitally important technique for discovering new and proprietary biocatalysts with the required activities and other properties required to make the needed molecules.

Key Performance Parameters

Cost-effective productivities can be met by biocatalysts that have the required activities in terms of their stability, specificity, and activity (grams of product per gram of cells per hour, or gram of product per liter reactor per hour), and with the right combination of the following characteristics (18,19):

- Cost
- Regulatory approval (safe and acceptable sources)
- Easy fit into process, including downstream processing and process tolerance (process integration)
- Availability of raw materials, including access to new materials
- Easy scale-up
- Broad specificity
- Activity on concentrated reactants
- Equilibrium point (high degree of conversion)
- Good stability
- High selectivity (few side products)
- Sterile conditions *not* required (or fermentation not susceptible to contamination)
- Ease of strain improvement by mutation or genetic engineering
- Rapid growth rates and nonfastidious nutritional requirements
- Robust and easily reproducible performance

The complexities involved in selecting biocatalysts with the optimal performance is illustrated by the cascade of enzymes used to create cheese flavors (20) (Fig. 3), and by a comparison of the rates of reaction of some common lipases when used either to hydrolyze triglycerides, or in the reverse reaction, to synthesis a flavor ester (Table 5).

Similarly there is a very large variation in the ability of different lipases to make a range of esters, as shown by the variabilities in degrees of conversion shown in Table 6 (21). These differences in reactivity are less surprising. It

Table 5. A Comparison of Lipolytic and Esterolytic Activities of Some Microbial Lipases

Microbial strain	Lipase activity (U g^{-1})	Esterase activity (U g^{-1})	Ratio of esterase: lipase activities
Chromobacter viscosium	144,000	2,650,000	18.4
Mucor miehei	420	22,000	5.24
Mucor miehi[a]	8,100	150	54
Pseudomonas fluorescens	8,500	26,300	3.1
Pseudomonas fluorescens[a]	15,000	150	0.01
Penicillium roquerfortii	290	625	2.2
Rhizopus japonicus	8,750	13,300	1.5
Rhizopus japonicus	36,000	39,400	1.09
Rhizopus japonicus[a]	47,000	0	∞
Candida cylindracae	85,700	53,600	0.6
Candida cylindracae[a]	95,000	370	257
Aspergillus niger	1,580	940	0.6
Geotrichum candidum	4,650	170	0.04
Candida lipolytica	10,500	150	0.01

[a]Data used is from Ref. 13 together with independant data for four of the enzymes in which esterase activity was assayed using geranyl proprionate in hexane and hydrolysis using olive oil emulsion as the substrate.

Table 6. Ester Formation Using Lipases

Lipase	Acetate	Propionate	Butyrate	Acetate	Propionate	Butyrate
Alcaligenes (Meito-Sangyo)	0.5	3	18	<0.5	4	24
Aspergillus (Rohm)	23[a]	33	40[a]	6[a]	15	6[a]
Fungal (Rohm)	0.5	40	65[a]	<0.5	48	61
A. niger (Palatase)	0.5[a]	22	98[a]	<0.5[a]	5	40[a]
C. rugosa (cylindracea) (MY)	3.5	95	98	<0.5	95	90
G. candidum (SNEA)	<0.5	2	18[a]	<0.5	2	5
M. miehei (Gist-Brocades)	80[a]	92	97	14	96	96
M. miehi (Lipozyme)	15	42	98	<0.5	30	93
P. cyclopium A (SNEA)	0.5	4	48	<0.5	2.5	38
P. cyclopium B (SNEA)	18	3.5	45	14	3.5	38
Porcine pancreas (Rohm)	0.5	55	95	<0.5	94	85
Porcine pancreas (Kochlight)	0.5	0.5	35	0.5	<0.5	12
R. arrhizus (SNEA)	5	70	97	3.5	58	97

Note: Data was obtained using 0.2 g lipase with analysis at 24 h, unless otherwise indicated.
[a]0.5 g of lipase was used instead of 0.2 g.

is noted that the catalytic structure of various lipases are different. Thus, whereas the catalytic triad in the active site of *Homo sapiens* and *Mucor miehei* lipases is Asp, His, Ser, in *Humicola lanuginosa* it is Asp, His, Try, and in *Geotrichum candidum* it is Glu, His, Ser.

Fermentation medium requirements include the following important considerations when developing a production medium that will be as inexpensive and reliable as possible (19):

- Selective advantages for the production strain, allowing fast growth, high yields, and no remaining by-products (a big challenge, not least because most wild-type strains are surface growers, and considerable effort may be necessary to get them to adapt to efficient production in submerged fermentations)
- Defined and balanced medium (no excess of any medium component)
- No by-products, color, odor, or substrates, remaining at the end of biotransformation
- Substrate(s) stable, soluble, and of constant quality
- Reduced risk of contamination due to medium composition
- Product not degraded by use as C- or N-source for cell growth in any circumstances
- Oxygen requirements (this can become rate-limiting for rapidly growing cultures; provision of air for growth is expensive because of filtration costs to ensure sterility and mixing power costs necessary to ensure good gas transfer throughout the fermenter)
- Little or no foaming problem
- Short lag period

Overall the bioprocess will be quantified and costed on the basis of the following seven factors:

- Specific growth rate
- Biomass concentration (g dry wt/L)
- Oxygen transfer rates and power consumption rates (usually turbine impellors are used so as to minimize power input costs and maximize mixing)

- Grams of product produced per liter of reactor per hour
- Concentration of product achieved (g/L)

Yields are very dependant on the type of metabolite being formed by the cells. Estimates of rates vary from 100 to 500 nmol/g cells^{-1} h^{-1} for amino acid formation, to 1 to 5 nmol/g cells^{-1} for vitamin and coenzyme synthesis (Table 7). Degree of conversion of precursor(s) into product (%), and purity of product (especially with respect to other chemically similar materials that could make its purification difficult) are also important. This can be also described as the selectivity of the reaction: grams of product required per gram of all products formed.

Problems in bioprocessing are well illustrated by the terpenes. Whereas terpenes are very common flavor and aroma materials with a very wide range of tastes, smells, and applications, production to date is entirely by extraction from plant raw materials. It appears that no true bioproduction process for a terpene has yet been established. This is probably because monoterpenes have low water solubilities, combined with high cytotoxicities, together with high volatilities and often poor chemical stabilities. By contrast, sesquiterpenes can be biotransformed much more successfully. For instance, the processes for producing Ambrox and sclareolide are described elsewhere in this article. Various bioengineering aspects of flavor process are summarized in Table 8.

Of course, each product and process can be expected to have some individual features. These include the optimization of fermentation media to maximize productivity and minimize cost, and to allow for particular features, for instance, often the product concentration achieved can vary with the specific growth rate of the fermentation. Some indication of the challenges encountered in developing a bioprocess are illustrated by the examples given in this article.

Technicocommercial Factors

In addition to the scientific factors just discussed, the following technicocommercial factors also have a big influence in many instances.

Flavors and fragrances are manufactured only when required by end-user companies, who are usually the manufacturers of consumer end products, that is, they are subject to a derived demand. This is accomplished following the issuing of a competitive brief to a number of companies. Thus, the requirement for the individual flavor chemicals and materials required to create the flavor is subject to a derived demand, and as the actual flavor is usually tailored specifically to the needs of one end-user company, the amounts of any one flavor chemical or material required can vary considerably from customer to customer.

Compounding and formulation of flavor chemicals is very often the last step in the manufacturing process and is a comparatively simple operation that adds significant value because of the improved complexity and character that can be created. This step depends on the creative input of skilled and experienced flavorists who create the formulations that dictate the types and amounts of flavor chemicals that are used. High flavor/aroma impact and character- or identity-conferring chemicals are very important because of their high value, even when they are expensive to make.

Control of organoleptic purity involves far more than just maintenance of chemical purity, as a trace impurity with a high flavor impact present at instrumentally undetectable levels can still have a very detrimental effect on flavor and aroma quality. This factor is very important when considering the batch-to-batch variations in quality because very few flavor chemicals are required in large enough volumes to justify production by continuous processes. Control of organoleptic quality also extends to the storage of ingredient chemicals and formulated flavors because cross-contamination can occur from ingredient to ingredient (vanillin is especially prone to this) and also from materials leached out of containers and packaging.

Economies of scale of manufacture are difficult to achieve for all but the largest-volume flavor chemicals. Instead, the small scale of manufacture creates the opportunity for generic manufacture using the same multipurpose equipment to make a range of ingredient chemicals on a campaign basis in direct response to market demands. This requires careful scheduling so as to properly match demand with supply, and scrupulous cleaning of equipment between runs.

Patent protection is important, especially for new flavor materials, but the enforcement of process patents is sometimes difficult, which is a particular problem for key producer microbial strains. Because of the individual nature of the flavor chemicals, including extraction from natural raw materials and complex formulation, special manufacturing know-how is often involved that cannot be protected by patents.

Table 7. A Comparison of the Relative Rates of Metabolic Processes Producing Different Classes of Metabolites as Bioprocess Products

Substrate family	Reaction type	Turnover rate (mmol/h/g cells)
Carbon source	Uptake, oxidation	10
Monomers, polymers	Synthesis, polymerization	1
Amino acids	Uptake, synthesis, incorporation in protein	0.1–0.5
Bases, nucleotides	Uptake, synthesis	
	Incorporation in RNA	0.15
	Incorporation in DNA	0.01
Vitamins, coenzymes	Uptake, synthesis	0.001–0.005

Table 8. Bioengineering Aspects of Bioflavor Production

γ-Decalactone	The precursor is an oil, and so water-miscibility issues had to be solved. Supply of the precursor in the form of an ester reduces foaming problems in the fermentor.
Benzaldehyde	Toxic HCN is produced when using cyanogenic nitrile glycoside precursors appropriate for microbial benzaldehyde production. Product inhibition reduces the rates of formation and final product concentrations, in situ product recovery by perevaporation has been successfully trialed.
Vanillin	A complex pathway is involved. When eugenol is used as the precursor then precursor toxicity to the cells must be overcome. Easy recovery of the product as a Schiff's base is incompatible with natural processing.
Soy sauce	Several strains of microorganism have been used, separately immobilized and operated in series.
Milk souring	In order to achieve balanced productivity, two fermentors have to be used in series. Because this is a rare example of a continuous fermentation process and because a food product is being made, special care has to be taken to ensure aseptic operation.
Hexenals and other flavor aldehydes	Use of an alcohol oxidase in two-phase water–hexanol reaction has been demonstrated to be high yielding.
Ethylisovalerate	Recovery of product from the exit gases from the reactor by adsorption to charcoal has been reported as successful.
Monosodium glutamate	Rod-shaped β-crystals of monosodium glutamate are easier to filter and wash, but the α-form crystalizes preferentially and so is redissolved and then recrystallized in the β-form.
Methylbutyric acids	Fed-batch reactions are used because of the toxicity of the methylbutanol precursors.
Methylketones (blue cheese flavor)	Extractive fermentation can be employed in which product is removed during the reaction to encourage continued production of the methylketones.
Offensive odors in exit gases	Mixed microbial culture biofilters used to degrade off-odors in exit gas streams.
Enzyme-modified cheese flavors	A number of different types of enzyme have to be added in the right sequence in order to obtain the desired flavor.
δ-Decalactone	Flavor quality was significantly reduced by trace impurities carried over from the fermentation media. Therefore before the bioconversion step, the 11-hydroxypalmitic acid precursor had to be purified.

The very extensive process validation requirements for biopharmaceutical processes are not generally required for flavor chemicals. The standards that are required are consistent with best practice in the food and fine chemicals industries, such as food hygiene standards and the sterility necessary to perform uncontaminated fermentions.

Manufacture of high-volume flavor and aroma chemicals is increasingly becoming attractive to chemical companies, rather than just flavour and fragrance companies, and internationalization is occurring fast, as indicated by the entry of Chinese manufacturers.

Because of the volatile nature of flavor and aroma chemicals, distillation is a dominant downstream processing operation in their manufacture. Usually, sophisticated equipment is used, controlled by computers, and upgraded to minimize oxidation, overheating, and so forth that can so easily produce off-flavors. Solvent extraction is common, and new methods are gaining credibility, for example, the use of supercritical carbon dioxide as an extraction technique; it has a polarity similar to that of hexane, and because of its volatility, leaves no residue in the product. Solvent extraction is limited to a short list of solvents that are classed as being consistent with natural processing (Table 9). Formulation of flavors often begins with spray-drying, often onto a carrier such as maltodextrins, which also serve to stabilize the flavor.

Raw materials are usually purchased from a great variety of sources, rather than produced in-house. Sometimes the availability of a raw material is a key driving force in a company's decision to produce a flavor. One example is that a good supply of by-product yeast is very useful for making savory flavors. Variabilities in the supply, price, and quality of natural raw materials, such as vanilla beans, can frequently pose big problems. In other cases it is the availability of the natural precursor molecule for a bioprocess that is the problem that can prevent the commercialization of that process. Examples include the 11-hydroxypalmitic acid required to make δ-decalactone, the methionine required for methional and related flavor chemicals, and also precursors for raspberry ketone (Fig. 5).

Effluent disposal from manufacturing plants can create some novel problems because of the objectionable smells that can be created for neighbors. One widely adopted process to combat this problem is to route all exit gases through a bed of supported microorganisms that metabo-

Table 9. Comparison of Various Solvents for the Extraction of 2-Phenylethanol Produced by Bioconversion

Solvent	K_d
Ethylacetate	0.95
Butylacetate	34
Butanol-1	11
Butanol-2	10
Diethylether	12
Hexane	0.5
Methyl propanol-1	22

Source: Ref. 22.
Note: K_d is the ratio of the dissolved substance in the extracting phase to that in the aqueous phase. Data assumes 20 °C with an aqueous solution containing 1 g/L 2-phenylethyl alcohol.

lize much of the objectionable materials, thereby valorizing the exit gases to acceptable emission levels. One such design of biofilter is the Bioton, supplied by Monsanto Enviro-Chem Systems.

In any process, its performance, and therefore the cost and quantity of the product, is dependent on a complex interrelationship between a large number of process variables, as illustrated for the use of glucose isomerase (GI) to make high-fructose corn syrup bulk sweeteners (23). This is a splendid example of the high degree of process integration that must be achieved in order for a mature bioprocess to continue to be successful (Fig. 4). The process involves a semicontinuous process with three enzyme reaction steps together with raw material processing, downstreaming, and sophisticated chromatographic separation of glucose and fructose sugars. Since GI is now a mature product, new versions of this enzyme demonstrate the big improvements necessary for second-generation processes to succeed. GI from *Streptomyces murinus* has about a 60-fold greater productivity in terms of tons of product produced/kg enzyme, which results in a 30-fold lower cost of product.

Protection of Technological Advantages

Another important factor is the protection of the technology by patenting (24). The key objectives are to avoid infringement of existing patents and to gain patent protection oneself. In this respect biochemical and microbiological inventions create some particular problems:

- Functionally identical biological materials (e.g., enzymes) can have significantly different structures and amino acid sequences.

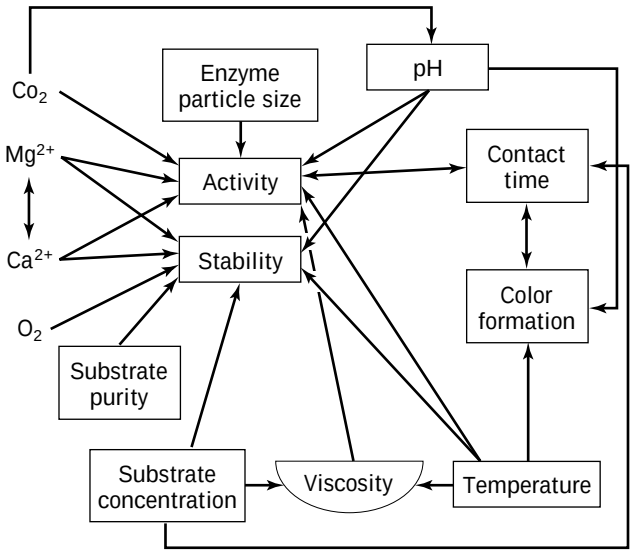


Figure 4. The interactions between the various process variables that can influence the productivity of glucose isomerase when used for the industrial production of high-fructose syrups from hydrolyzed starch.

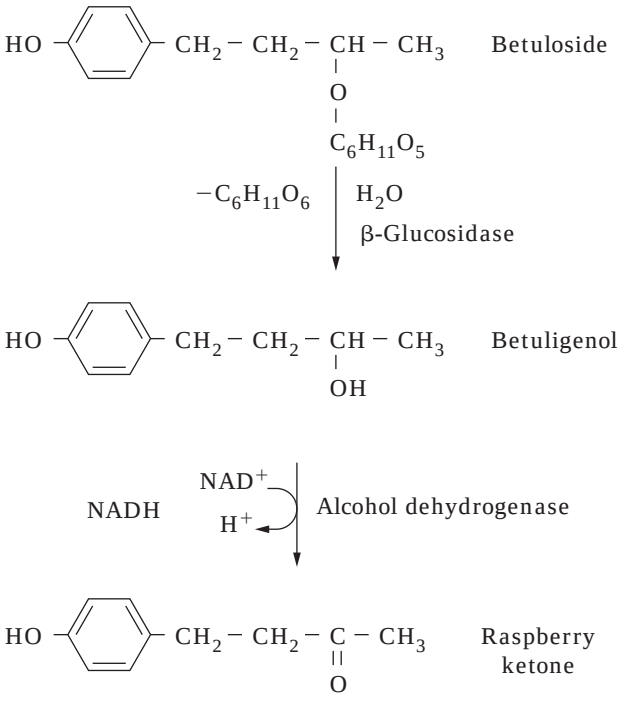


Figure 5. Scheme for making raspberry ketone flavor chemical from the plant-derived precursor betuloside.

- Functionally identical biological materials can be obtained from very different biological sources (e.g., enzymes with the same substrate specificity from different microbial sources).
- The requirement to deposit samples of microorganisms is a substantial overdisclosure as compared to patenting of other inventions. This often leads to some processes being kept as in-house secrets.
- Biological materials, especially living organisms, are too complex to be described in sufficient detail and with enough precision to easily satisfy the requirement of patent law. For instance, existing microbial classifications are not always directly applicable to newly isolated microorganisms.
- Varying criteria and definitions are used to distinguish biocatalysts, for example, substrate and inhibitor specificity and immunochemical cross-reactivity in the case of enzymes, and nutrient requirements and fatty acid composition and so on for microorganisms.

Categorization of Bioflavor Processes

Depending on the complexity of the biochemical reactions required to produce the desired products, bioflavor processes can be categorized along the lines of the following examples:

Single-Enzyme Reactions. Examples of single-enzyme reactions include glutaminase, 3′–5′ nucleotide diesterase (e.g., from *Penicillium citrinium*), adenylic deaminase (e.g., from *Brevibacterium ammoniagenes*), glycosidases

for the release of terpene flavors from grape and other flavor sources, lipases and esterases for flavor ester synthesis and for resolution of racemic mixtures (such as obtaining *l*-menthol from *dl*-menthol).

Mixed-Enzyme Reactions. Examples of mixed-enzyme reactions include use of mixed proteases for producing protein hydrolysate flavors; β-glucosidase and nitrile lyase to make benzaldehyde from amygdalin; lipoxygenase, hydroperoxide lyase, and oxidoreductase to make, for instance, *cis*-3-hexenol from linoleic acid via *cis*-3-hexenal), and proteases, peptidases, lipases, and esterases to make enzyme-modified cheese flavors (EMCs) (Fig. 3). Recently, mixed-enzyme preparations have become widely used in the extraction of fruit flavors.

Use of Mixed Enzymes to Enhance Fruit Flavors. Recently it has been found that very many fruit flavor chemicals such as the 13-norisoprenoids and terpenes (e.g., linalool and geraniol) occur in plants as glycosides (24). Geraniol was the first glycosidically bound aroma chemical to be identified as such in rose extracts, with a structure of geranyl-xylose-glucose, and the C13-norisoprenoids (e.g., ionones, raspberry ketones, and others such as *trans*-3-hexenol are present in the fruit of origin, predominantly in glycosidially bound forms; see Ref. 25). Usually the proportions of these glycosidically bound flavor chemicals to free flavor chemicals are in the ratio of 2:1 to 5:1. This is possibly because the glycosides are more water soluble, more resistant to acid hydrolysis, and effectively detoxify these terpenes, phenols, and similar compounds.

The glycosides have similar structures with the aglycone bound first to glucose and then to a another sugar such as α-L-arabinose, α-L-rhamnose, β-D-xylose, β-D-apiose, or another β-D-glucose unit, depending on the plant source. Hydrolysis of such glycosides liberates the aglycose, thereby enhancing the flavor content of a conventional fruit product and even allowing the generation of flavor from waste materials such as peel and stems. However, the enzymes required to achieve these advantages must have some special characteristic such as good activity at low pH (about pH 3.5), resistance to product inhibition by glucose, and even resistance to moderate concentrations of ethanol in the case of fermented products. Most importantly, flavor chemicals are released, and flavor is enhanced only when two enzyme activities are present: a β-glucosidase and also another glycosidase such as α-L-arabinosidase or α-L-rhamnosidase, depending on the plant material being treated. Because of these requirements, glycosidically bound flavors are released only rarely in existing processes, such as fermentation, because yeast β-glucosidase is insufficiently active. However, one process that absolutely depends on endogenous glycosidase action is in flowers, where it appears that aroma materials are stored as glycosidases and then, upon the flower opening, glycosidases are produced that hydrolyze the glycosides, releasing the aroma chemicals that create the fragrance of the flower.

Use of such special glycosidase combinations have now been widely adapted, for instance, in grapefruit juice extraction, and the cloning of the appropriate enzymes into suitable host strains is underway. An example is the *A. niger* enzymes used by Gist-brocades (26). A very good example of the improved yields of flavor that can be achieved by the use of glycosidase is the improved yields of vanilla flavor that can be obtained by treating vanilla beans with emulsin, which is a mixed enzyme preparation especially rich in β-glucosidase that is obtained as a crude extract of almond meal.

Use of Single Microbial Strains. Single microbial strains are used to produce the following: methylketones with spores of *A. niger*, γ-decalactone with *S. cerevisiae* strains, δ-decalactone with *Cladosporiun suaveolens*, diacetyl with *Streptococcus cremoris* and *Streptococcus diacetylactis*, soured milk for margarine production with *S. cremoris*, methylbutyric acids with *Acetobacter acetii*, and pyrazines with *Corynebacterium glutamicum*.

Combined use of Several Microorganisms. A combination of several microorganisms is used for such processes as the Unilever two-stage continuous process for manufacturing soured milk, which makes use of *Lactobacillus* and *Streptococcus* strains (27) (Fig. 6), and the accelerated process for soy sauce manufacture using *Pediococcus halophilus*, *Saccharomyces rouxii*, and *Torulopsis / Candida versatilis*. Also several microbial strains are used to achieve a synergistic effect in cheese fermentations. For instance, propionic acid is an important flavor component of Emmental and some other Swiss cheeses. It can be produced by first forming lactic acid with strains such as *Lactobacillus lactis* or *Streptococcus thermophillus* and then metabolizing the lactic acid into propionic acid by *Propionibacter* strains. Using this two-stage method, concentrations of propionic acid of 40 g L^{-1} can be made. Propionic acid is also useful to form flavor esters such as ethyl, benzyl, citronelyl, and geranyl propionates, as is butyric acid, also produced by fermentation in concentrations of 20–30 g L^{-1}, which is used to make ethyl, isobutyl, and amyl butyrate flavor esters.

Combined use of Enzymes and Microorganisms

Savory Flavor. Production of savory flavors using proteases, lipases, ribonucleases, and polysaccharidases to break down plant raw materials, and then microorganisms to convert the products of hydrolysis into flavor chemicals are used in processes such as the Biosol process (28). Improved flavors using this approach could be possible by genetically engineering yeast strains to contain higher concentrations of the savory flavor enhancers GMP and IMP, and also of the flavorsome octapeptide "beefy meaty peptide."

Fruit and Dairy Flavor. The traditional use of the acetic acid bacterium *A. acetii* to make vinegar by oxidizing the ethanol in wine into acetic acid has been redeployed by modern flavor biotechnologists. This time *A. acetii* is used for the bioconversion of isobutyric alcohols present in fusil oil into isobutyric acids. Thus both the 2- and 3-methylbutanols are oxidized into 2- and 3-methylbutyric acids. This oxidation takes place with very little racemization.

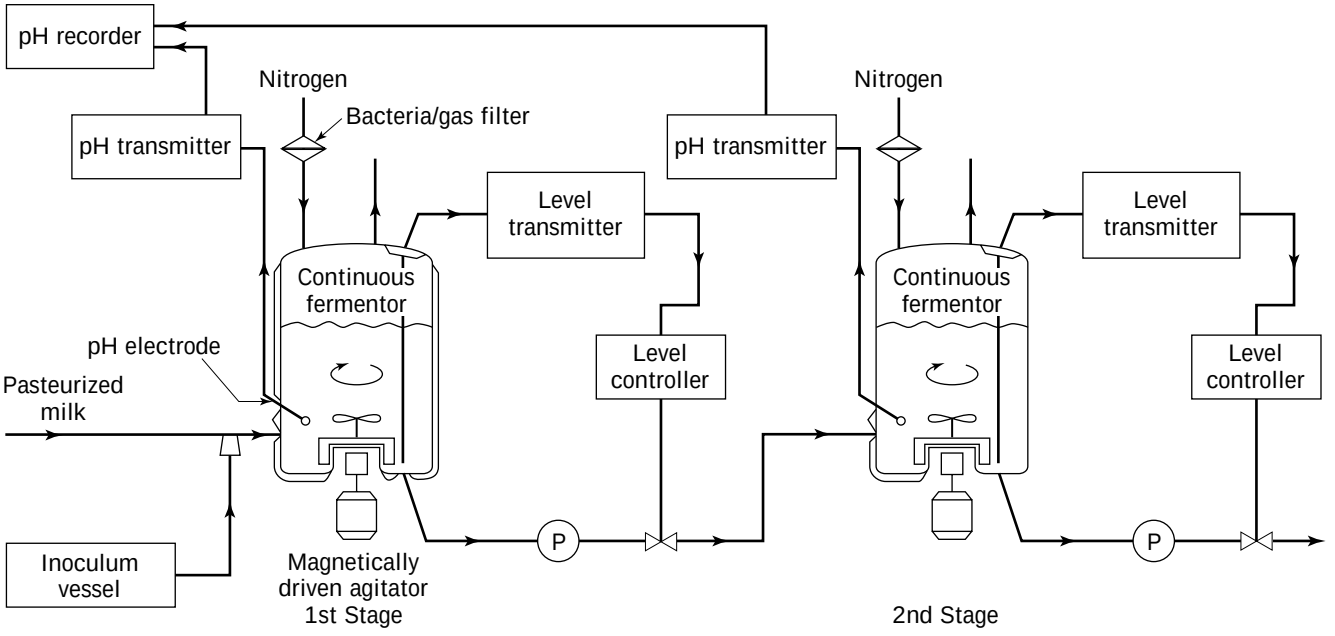


Figure 6. Two-stage continuous microbial souring of milk. *Source*: Ref. 16.

The overall reaction requires two sequential reactions: a dehydrogenation into the aldehyde intermediate than can tautomerise easily, and then oxidation to the acid. These isobutyric acids are good flavor chemicals in their own right, but in addition their value and range of uses can be increased and expanded by esterification either by using lipases or esterases "in reverse," or by azeotropic distillation (29) (Fig. 7).

Plant-Derived Biocatalysts. Plant-derived biocatalysts include lipoxygenase for the production of *β*-ionone, *cis*-3-hexenol, and so forth. Soy lioxygenase produces 13-hydroperoxides that break down to C6 aldehydes, whereas tomato lipoxygenase has a different substrate specificity and produces 9-hydroperoxides from linolenic acid, rather than the 13-hydroperoxides, with the 9-hydroperoxides breaking down to form C9 aldehydes such as *cis*-3-nonenal. By contrast, Fusarium lipoxygenase forms 9- and 13-hydroperoxides. *Morchella* sp. are used for 1-octen-3-ol production and hydroperoxide lyase obtained from plant sources such as green pepper. So, in general there would appear to be a historical trend from traditional fermented products to the use of isolated enzymes, submerged fermentation, and now bioconversion methods. Obviously the complexity of the biocatalytic system used has a direct effect on the ease of operation and costs of the process, and so a more technically sophisticated process can be justified only when the resulting product is sufficiently valuable, both in terms of profit margins created and markets that can be gained, to justify its use.

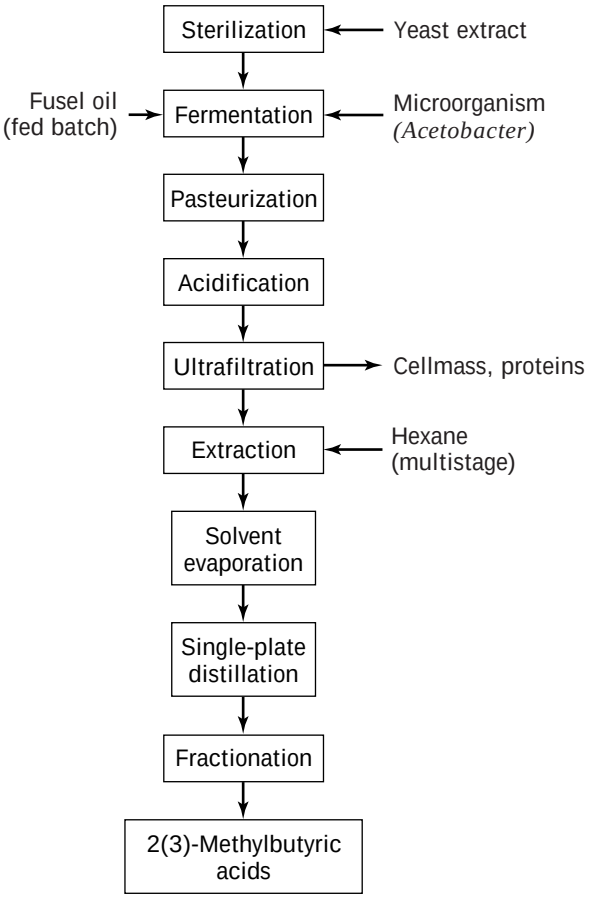


Figure 7. Large-scale process for the manufacture of 2- and 3-methylbutyric acids for flavor chemicals and flavor precursors, using fusel oil as the source of the 2- and 3-methylbutanol precursors. *Source*: Ref. 29.

Examples of Actual Bioflavor Processes

Isolated enzymes are used to make a variety of flavor products, as described in the following sections.

Protein Hydrolysates. Different mixtures of microorganisms and plant and enzyme proteases are used to produce protein hydrolysates, depending on the source of the protein, the process used, and the product needed, especially because different end-use applications require different degrees of hydrolysis of the proteins. A particular challenge is to minimize the bitter taste of the hydrolysate, which is caused by peptides that have N-terminal hydrophobic amino acids. This can now be achieved using specially developed peptidases, especially aminopeptidases, which are usually obtained from lactic acid bacteria such as *Leuconostoc lactis* and from *Aspergillus oryzae*. These were discovered as enzymes from some dairy starter cultures that were found to have debittering activity (30) and are being cloned to make them available more cheaply and in greater quantities. Some carboxypeptidases are also used. Note that the process lends itself to the creating of a range of savory products, depending on whether the material is dried, concentrated, or supplied as a dilute liquor.

Monosodium Glutamate. Over 500,000 tonnes of monosodium glutamate the taste enhancer is made worldwide each year, mostly by fermentation using specially selected strains of *Corynebacterium glutamicum*. These strains have a low α-ketoglutarate dehydrogenase activity that prevents the formation of succinate, and so they instead form glutamate by reductive amidation of the α-ketoglutarate. In addition, successful and high-yielding fermentations depend on the use of media that have a very low biotin content (less than 5 μg L^{-1}). This makes the cell walls of the microorganism leaky so that the glutamate passes out of the cell and accumulates in the medium. After fermentation the cells are removed, and the glutamate is precipitated (by the addition of HCl) as α-form crystals. These are heated in water to convert them to β-form rod-shaped crystals that are easier to filter and wash. This is followed by neutralization, decolorization, and recrystallization (17).

Flavor Nucleotides. Hydrolysis of yeast RNA into flavor-potentiating nucleotides such as guanosine-5′-monophosphate is carried out by processes similar to the following (17). Yeast is grown such that it has a high RNA content. It is then harvested in late log phase with a hot alkaline solution (Fig. 8). Since inosine monophosphate (IMP) is a potent flavor enhancer, the activity of the hydrolysate is frequently enhanced by treatment with an adenylic deaminase, which converts 5′-AMP into 5′-IMP, as shown in Figure 9.

Use of Enzymes in the Extraction of Plant Materials. Plant materials are a traditional source of flavoring materials. Recently it has been found that many plant flavor chemicals, especially terpenols such as linalol, nerol, and geraniol, occur glycosidically bound to the plant tissue. As a result, specialized glycosidase enzyme preparations have been developed to augment traditional processing by enabling increased rates of extraction and higher yields of products. This approach has been increasingly employed in wine making; enzyme preparations such as those from *A. niger* are added to the must.

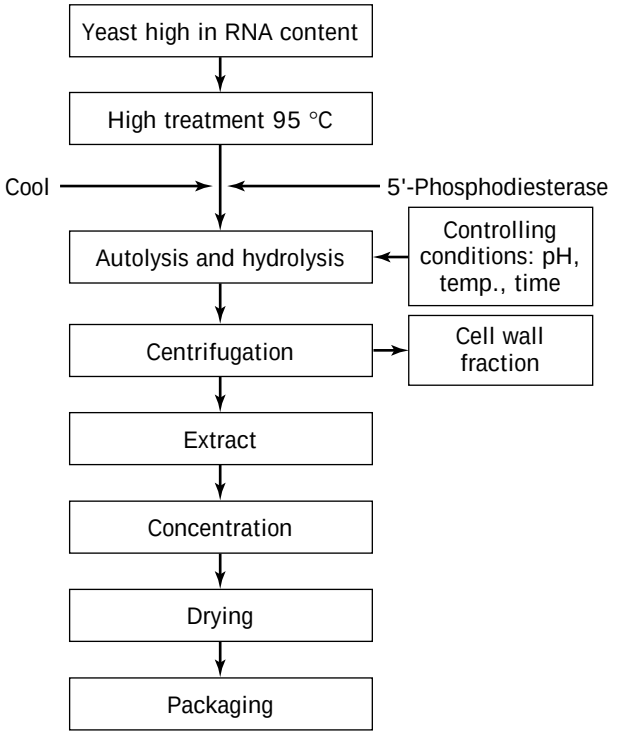


Figure 8. Flow diagram of the manufacturing process for AMP flavor-enhancer production.

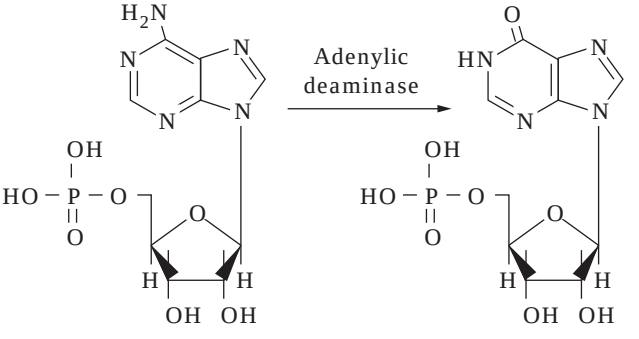


Figure 9. Conversion of 5′-AMP to 5′-IMP by the action of adenylic deaminase.

Curiously, endogenous plant glycosidases and the glycosidases of wine and beer yeasts are barely active against these glycosidically bound flavors, which explains the low yields obtained during traditional extraction and fermentation. This is because the glycosidases present in plants and yeast tend to be active at a moderate pH (around pH 5), are inhibited by glucose and ethanol, and have only a narrow aglycone substrate specificity. Such specialized enzymes have now been made available by enzyme supplier companies such as Gist-brocades.

One good example of extraction is the preparation of L-rhamnose, which is the key precursor molecule for the very important flavor Furaneol® (2,5,-dimethyl-4-hydroxy-(2*H*)-furan-3-one), which is prepared by heating rhamnose

with an amine source such as proline. L-Rhamnose can be isolated by the enzyme hydrolysis of glycosides such as naringin, neohesperidin, and rutin present in citrus peel and other plant sources. These molecules are flavanone glycosides with rhamnose, in the terminal position. Enzymes are preferred that exclusively or, more usually, preferentially hydrolyse rhamnose, while minimizing the release of other sugars. This is to avoid an off-taste that develops during the reaction of glucose with the proline. Therefore any glucose released is removed by fermentation to ethanol and CO_2 by yeast, or into 5-ketogluconic acid, with its removal by precipitation as its calcium salt (31). Taking this approach a stage further, by coupling hydrolysis of a precursor glycoside with its bioconversion into a useful flavor chemical, raspberry ketone (4-(4′-hydroxyphenyl)-butan-2-one) has been synthesized (32) (Fig. 5).

Enzyme treatment can also be used to reduce an undesirable flavor of a plant juice, rather than to enhance a pleasant taste. For instance, the bitterness of citrus juices is due to the flavanones limonin, naringin, and neohesperidin, which are often present in fruit juice as glycosides. Bitterness is a particular problem for grapefruit, in which naringin predominates and has a taste threshold of just 20 ppm. Commercially available naringinase, which is a mixture of glucosidase and rhamnosidase, acts sequentially by converting the naringin to prunin and to then the much less bitter naringenin, has been proved in trials to provide an effective solution to this problem.

Benzaldehyde. Another source of an enzymatically treated plant-derived flavor is natural benzaldehyde (33). A number of companies have operated a bioprocess to make natural benzaldehyde, based on the combined use of β-glucosidase and mandelonitrile lyase, which are both conveniently present in almond meal (emulsin). This is used to treat the mandelonitrile (a nitrile glycoside) present in certain plant materials. A preferred industrial source is the cherry stones remaining after processing to remove the cherry flesh. However, the liberation of poisonous hydrogen cyanide by the lyase creates safety problems and has led to a search for alternative processes. L-Phenylalanine is a suitable precursor, especially since it is now available cheaply from a fermentation process developed to make phenylalanine as a component of the high-intensity sweetner Aspartame. The basidiomycete *Ischnoderma benzoinum*, combined with effective in situ product recovery, produces 1 g benzaldehyde/L, together with some 3-phenylpropanol.

Methylketone Production. Methylketone provides the characteristic flavor of blue cheeses such as Roquefort, Camembert, and Stilton (34) and consists of mixtures of C5 to C11 2-alkanones, especially 2-pentanone, 2-heptanone, 2-nonanone, and 2-undecanone. This is because only medium-chain-length fatty acids are converted into methyl ketones. This may be because only such medium-chain fatty acids, and not longer-chain fatty acids, can enter the cells as free acids and seem to cause significant morphological changes to the cells that take them up. Such flavor chemicals are commercially important in the manufacture of consumer products, such as salad dressings, soups, and pizza toppings, in which a cheese flavor can predominate, and also as a component of other flavors, especially savory flavors.

Methylketones are manufactured using the spores of microorganisms, such as *Penicillium roquefortii*, that still retain the required enzyme activities despite being dormant. This makes the production-scale use of spores very easy, as once grown they are easy to store until required and then require no growth media. They are eminently suitable for use in a logistically convenient and cost-effective biotransformation process. This process is much simpler than a conventional fermentation because of, for instance, very moderate oxygen transfer requirements. Using this method, lipolyzed fats or oils can be transformed into cheese flavors with 10 times the methylketone content of normal blue cheeses (Fig. 10).

The details of the biochemical pathway whereby microorganisms convert fatty acids into methylketones have been elucidated and used as the basis for large-scale fermentation processes, although the detailed physiologically control processes are still not properly understood. Commercially successful microbial processes for making methylketones have been used for some time by companies such as Unilever and Bush Boake Allen (35). The way microorganisms such as *P. roquefortii* produce methyl ketones from medium-chain fatty acids is of considerable scientific interest because it resembles conventional β-oxidation until the final step, at which a decarboxylase produces the methyl ketone, replacing the thioesterase found in β-oxidation. Methylketone formation is favored under con-

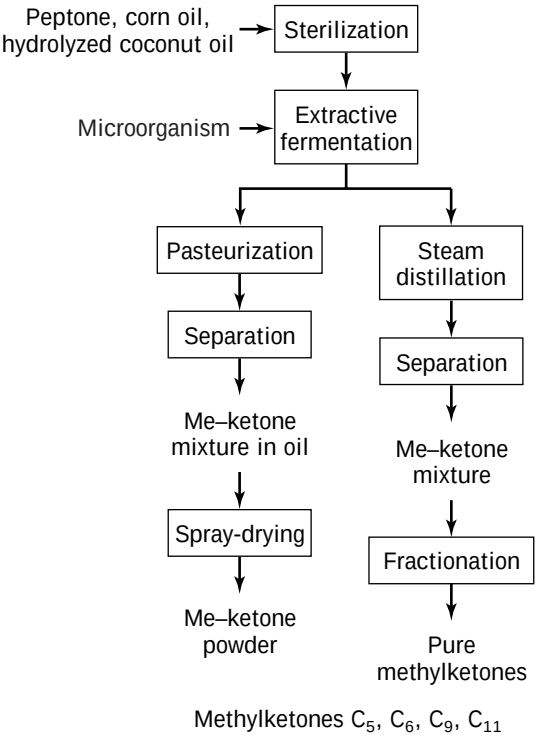


Figure 10. The unit operations involved in manufacturing processes for the methyl ketone blue cheese flavoring materials.

ditions in which cell growth is constrained. Spores are more active in producing methyl ketones than mycelium; use of immobilized cells have been described, and solid-state fermentations are often used. For instance, Bush Boake Allen's process involves mixing cellulose powder with coconut oil, which contains a high concentration of octanoic acid, and inoculating with *A. niger* spores, which converts the octanoic acid into 2-heptanone, which has strong blue cheese aroma and flavor (20). The yield is reported to be about 40%, and 2-undecanone to be 60% of the total products. By comparison, using spores of *P. roqueforti* in solid-state fermentation, Larroche and Gross (36) achieved product concentrations of 75 g 2-heptanone L^{-1}, with a yield of almost 80%. This appears to be the only manufacturing process for flavor chemicals that employs solid-state fermentation, despite the advantageous features of this technique. This may be because the "culture" of applied microbiology at the manufacturing level is dominated by the stirred-tank approach, developed with such success for antibiotic production, except for traditional processes such as cheese ripening and silage making.

γ- and δ-Decalactones. The γ- and δ-decalactones are commercially very important flavor chemicals for which successful industrial bioconversion processes have been established. This is despite the fact that de novo production by fermentation of glucose and other carbon sources is very low yielding. An outline of the process (29) for making γ-decalactone, useful for peach and other flavors, and a typical fermentation profile (37) are shown in Figures 11 and 12.

The yields of lactones obtained by this process are greatly enhanced by acidification during product recovery. Only about 25% of the 4-hydroxydecanoic acid is lactonized during the fermentation at pH 6.0–7.0, but full yields are reported when extraction by steam distillation and solvent extraction are performed at pH 2 (38).

By-product formation can also be an important consideration affecting the yield of a product and the ease of its isolation. For example, 3-hydroxy-γ-decalactone continues to be formed even after production of the desired product, γ-decalactone, has ceased. It is associated with the fermentation becoming more aerobic (Fig. 11), although some 3,4-dihydroxydecanoic acid can be cyclized into 3-hydroxy-γ-decalactone, which has no positive flavor value, or it can be dehydrated into a mixture of 2-decen-4-olide and some 3-decen-4-olide, which together make up some 5% of the microbial metabolites. Upon reduction by yeast, these metabolites give δ-decalactone and the *R*-γ-decalactone, respectively. In addition, minor amounts of other lactones are also formed from hydroxylated fatty acid intermediates in the β-oxidation pathway; 3-dodecen-6-olide is formed from the C12 fatty acid intermediate.

Yields of the desired γ-decalactone product can be maximized by terminating the fermentation at the right point, or the 3,4-dihydroxy-γ-decalactone can be converted into 3,4-unsaturated-γ-decalactones by distillation and then converted into some additional γ-decalactone by a stereoselective *S. cerevisiae* reduction.

In all of the processes developed, yeasts are used as the biocatalyst. This is in part because yeast allows fatty acids to diffuse freely across their cell membranes, making them readily available for reaction. By contrast most bacteria find fatty acids to be toxic, particularly in the undissociated forms that prevail at low pH, and cellular uptake is deliberately limited and requires specific uptake proteins.

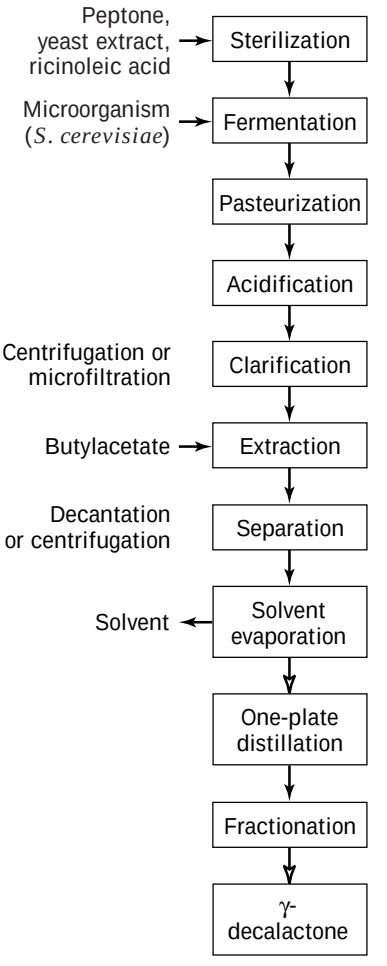


Figure 11. Outline of the process unit operations required for large-scale production of the flavor chemical γ-decalactone.

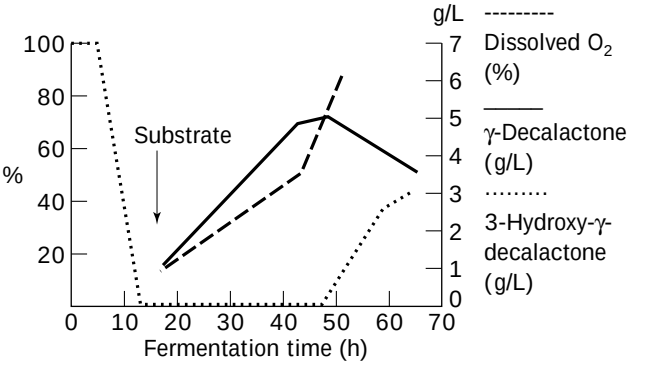


Figure 12. Fermentation profile for γ-decalactone formation from caster oil substrate.

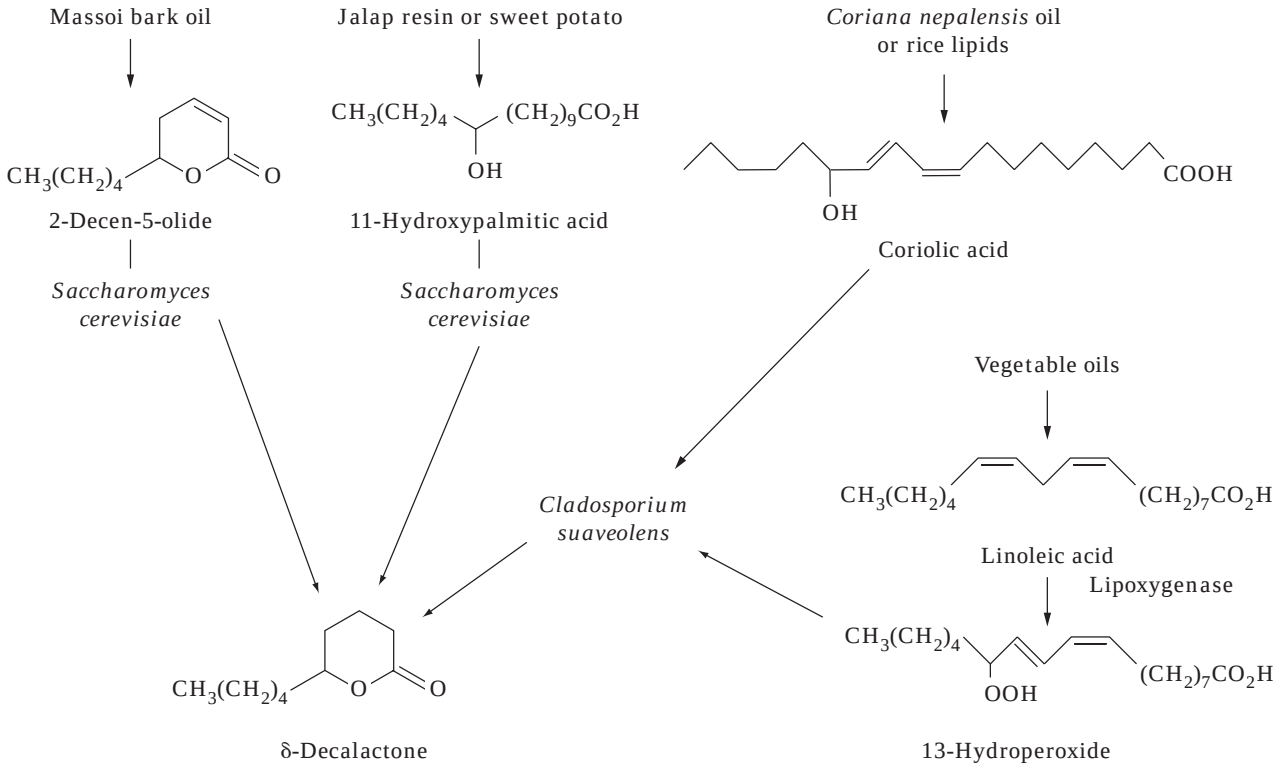


Figure 13. Various precursor molecules and biotransformation routes used for making δ-decalactone flavor.

It should also be remembered that the process from the microorganisms' viewpoint consists of diffusion of the precursor molecule across the cell membrane and into the cytosol, thence into the peroxisome, where reaction takes place, with subsequent movement of the lactone out of the peroxisome and then out of the cell. Hence, the concentration of fatty acid precursor supplied is important: too little and it is all consumed to meet microbial growth and maintenance requirements, too much and the excessive fatty acid concentration is toxic, which also reduces yields of the desired product. Considerable research has been devoted to finding the best raw material for γ-decalactone production. Castor oil is a cheap and readily available source of *R*-ricinoleic acid, but plant sources of oleic acid have also been converted into ricinoleic acid by microbial hydration. In addition, alkyl ricinoleate has been shown to have the advantage of reducing the amount of foaming that takes place in the fermenter. Even more ingenuity has been displayed in developing sources of the precursors and biosynthetic pathways to δ-decalactone, as demonstrated in Figure 13. Bioconversations have also been carried out using water-immiscible extracting solvents separated from the aqueous phase by a semipermeable membrane to allow selective passage of the δ-decalactone into extraction solvent (39).

These routes to δ-decalactone are also of interest because *S. cerevisiae* is used to carry out two different reactions: reduction of the 2-decen-5-olide and lactonization of the 11-hydroxypalmitic acid. Jalap resin also contains some 3,11-dihydroxymyristic acid, and this is converted by *S. cerevisiae* during the process into its corresponding lactone δ-decalactone.

Diacetyl. Shukunobe and Takato (40) described the production of diacetyl, useful as a component of dairy flavors, from citric acid using *Streptococcus cremoris* and *S. diacetylactis,* achieving yields of 14 g L^{-1} of product. A feature of this process was the requirement of oxidizing agents for the effective recovery of the diacetyl by distillation.

1-Octen-3-ol. Schindler and Seipenbusch (41) reported an industrial process for making 1-octen-3-ol, which is a character impact flavor chemical of mushrooms, by fermenting mycelia of morchella and then adding linoleic acid. Under high shear conditions, about 2.5 g product/kg dry w mycelia was formed, presumably by lipoxygenase action on the linoleic acid.

Methylbutyric Acids. 2-Methylbutyric acid and 3-methylbutyric acid are useful flavor components, especially when esterified. They are manufactured from their corresponding 2- and 3-methylbutanols, which are extracted from fusel oils by oxidation using *A. acetii* (vinegar bacteria) by the process shown in Figure 7 (29).

C6 "Green" Flavor Aldehydes and Alcohols As an example of a more recently developed bioprocess, Firmenich claims to have developed a flexible and viable industrial-scale biotechnological process for the production of natural

C6 aldehydes and alcohols. These include *trans*-2-hexenol, which has a leafy green and fruity taste, and *cis*-3-hexen-1-al, which has a grasslike green and fresh taste (42). The current world market for these materials is estimated to be high, at up to $50M per year. Another stimulus to develop a bioprocess is that although demand for C6 "green" flavors is growing, the traditional sources such as mint oil continue to be in short supply compared with demand. However, there are no good microbial sources of the key hydroperoxide lyase enzymes required; plant sources have hitherto been too low in activity, and cloning of these plant enzyme genes to provide more active enzyme sources is difficult because of the reluctance of consumers to accept gene modification as natural.

The Firmenich process simulates the metabolic path whereby these flavor molecules are made in plants: from polyunsaturated fatty acids by the sequential action of lipoxygenase, hydroperoxide lyase, and oxidoreductase enzymes. Firmenich uses soy bean lipoxygenase to make the hydroperoxide fatty acid derivatives, with the reaction carried out in aqueous medium at pH 9.5, and supplying pure oxygen to ensure that the reaction is completed within 1 h. The next step, cleavage of the hydroperoxides, was carried out at pH 8.5 using a plant extract specially selected to have a high lyase activity. For instance, for producing *cis*-3- and *trans*-2-hexenals, green pepper and cucumber extracts had the highest activities.

If desired these flavor aldehydes could be converted to their corresponding alcohols, such as *trans*-2-hexenol, using the alcohol dehydrogenase activity of *S. cerevisiae*. Using this process, typical results include the production of *cis*-3-hexenol at 4.2 g L^{-1}, with a yield of 42% based on the fatty acid supplied, and *trans*-2-hexenal at 1.8 g L^{-1} at a yield of 24%. These values compare with typical concentrations of these flavor chemicals in fruits and vegetables of only about 2.1 mg kg^{-1} and 6.2 mg kg^{-1}, respectively; in this respect, the bioprocesses are almost 2,000- and 200-fold more effective, respectively (Table 4). However, this method is not so easily applicable to *cis*-3-hexenol because, although *trans*-2-hexenal and *cis*-3-hexenal are initially produced as a mixture, the *cis*-3-hexenal is then subject to rapid nonenzymic conversion into *trans*-2-hexenal, and then the former is easily reduced by the yeast first into *trans*-2-hexenol and then into *n*-hexanol. In addition, *trans*-2-hexenal can react with acetaldehyde to form 4-octen-2,3-diol (37). Therefore, in order to make *cis*-3-hexenol, the 13-hydroperoxide, plant lyase, and yeast must be incubated simultaneously so that *cis*-3-hexenal is immediately reduced into *cis*-3-hexenol without any isomerization.

A similar approach is used in the manufacture of 1-octen-3-ol and other related C8 alcohols and aldehydes that are very characteristic of mushroom flavors from *Agaricus bisporus* and related species. This fungus has a lipoxygenase with a special substrate specificity that converts linoleic acid into the 10-hydroperoxide, which is then cleaned to give the 1-octen-3-ol. This flavor is manufactured using unwanted mushroom stems as the source of the biocatalyst.

β-Ionone. β-Ionone production is a good example of an emerging bioprocess that also uses lipoxygenase, but requires the development of novel reaction conditions. Both α- and β-ionones are widely used in the perfume and flavor industries, and are probably produced in nature by oxidation of carotenes and fatty acids. In this process soya lipoxygenase is used to convert carotenes, from carrots or algae, in a very concentrated reaction together with fatty acids. The reaction is carried out in a kneading trough under pressure and with air renewal. Products include α- and β-ionones and aldehydes such as hexanal and 2,4-decadienal, produced presumably from the carotene and fatty acids, respectively (43).

Vanillin. An example of a new and emerging approach, but using a different technical approach is the Harmaan and Reimer process for vanillin, the key flavor constituent of vanilla. Screening allowed the isolation of a novel *Pseudomonas* strain capable of converting eugenol, which is readily extracted from clary sage, into a number of products, several of which could be isolated in good yields, the choice of product being dictated by the fermentation conditions employed (44) (Fig. 14). From a biochemical perspective, there is an interesting parallel between phenylpropenoid degradation, such as of ferulic acid and fatty acid β-oxidation. A recent advance has been achieved by Gasson et al. (45). They have characterized the pathway of ferulic acid degradation in a strain of a soil bacterium, *Pseudomonas fluorescens*, which they isolated by its ability to grow on ferulic acid as sole carbon source. A key feature of the pathway is the hydration and retroaldol cleavage of feruloyl-CoA to produce vanillin and acetyl-CoA, catalyzed

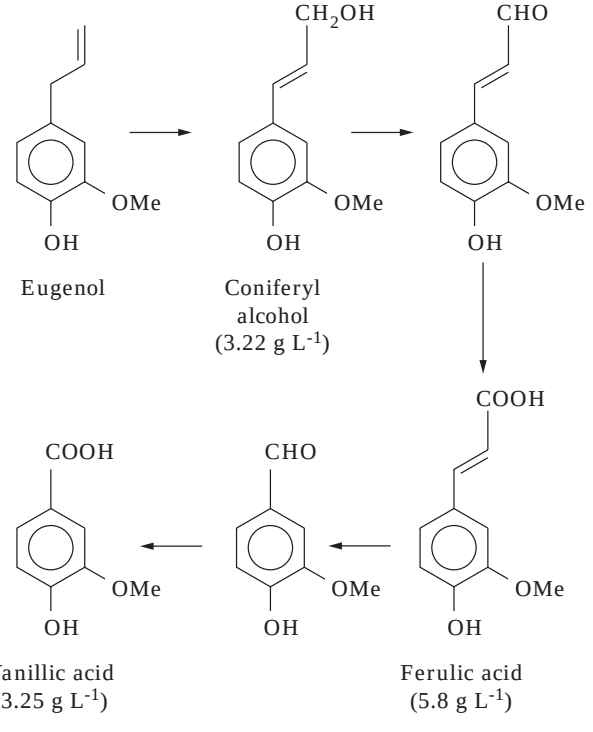


Figure 14. Eugenol metabolism using newly isolated *Pseudomonas* sp. and using different fermentation media to produce the various metabolites as isolated products. *Source*: Ref. 44.

by a newly discovered enzyme, (+)hydroxycinnamoyl-CoA hydratase/lyase. Comparison of the amino acid sequence of this enzyme with other known enzymes reveals that it is a member of the enoyl-CoA hydratase/isomerase superfamily (45).

Pyrazines. Pyrazines contribute substantially to the characteristic tastes of roasted and toasted foods and also contribute to many cheese, vegetable, nut, and fermented food flavors. *Bacillus subtilis* and *Brevibacterium linens* have been used to produce 2,5-dimethylpyrazine and tetramethylpyrazine at 4 g L^{-1} and 1 g L^{-1}, respectively, by fermentation of soy beans or flour enriched with the precursors L-threonine, acetoin, and ammonia and then supplied in a fed-batch mode, followed by extraction and distillation. The highest reported yield of a pyrazine as produced by a bioprocess is 12.5 g L^{-1} of 2-isobutyl-3-methoxypyrazine, produced by a mutant strain of *Pseudomonas perolens*. This pyrazine has a characteristic bell pepper flavor. The report provides a good example of strain-to-strain variations in flavor production by microorganisms, as the parent strain of *P. perolens* produced only 46 mg L^{-1} of this pyrazine when grown on the same medium (46).

Irones. Irones are degradation products of carotenoids that are valuable because of their very high flavor impact and low flavor thresholds. Traditionally they have been isolated from stored iris rhizomes, but now accelerated microbial processes have been developed, for instance, by Givaudan-Roure using microbial strains isolated from rhizomes such as *Serratia liquifaciens*. Using this method, 1 g mixed irones/kg rhizome could be obtained after 8 days, compared with 0.4 g/kg after 3 years for the traditional process (a 340-fold increase in productivity), and with the proportions of *trans*-α-, *cis*-α-, *cis*-γ-, and β-irones being very similar in both products (47).

Terpenes. Despite terpenes comprising a class of molecules that is very important in flavor and aroma products, very little real progress toward developing robust industrial processes appears to have been made. This is in part because of the water insolubility and cytotoxic characteristics of terpenes due to their effect on cell membranes. Toxicity seems to be related to their log *P* (partition coefficient) values, and terpenoids are often more toxic than hydrocarbon terpenes. However, a couple of promising attempts to develop productive process have been made. (*R*)-(+)-limonene used at 20 g/L has been converted on a 70-L scale into 900 g of pure 1*S*,2*S*4(*R*)-*p*-menth-8-ene-1,2-diol, with very few side products, using *Corynespora cassiicola* (Fig. 15). So as to avoid toxicity problems, the limonene was fed continuously; glucose, which is required to maintain enzyme activities, was also fed continuously in response to the CO_2 level in the exit gases (an indicator of cellular respiration rates). (Fig. 16). Using a second microorganism, *Penicillium digitatum*, the (*R*)-(+)-limonene component of racemic limonene was transformed by the same workers into α-terpineol in a 46% yield (48) (Fig. 15). Another promising terpene biotransformation is of the sesquiterpene patchoulol into 10-hydroxypatchoulol, a key intermediate to nor-patchoulenol, which is the major odiferous component of patchouli oil; 2–4 g L^{-1} of the 10-hydroxypatchoulol was obtained using strains of *Pithomyces* obtained by a selective screening exercise (49). Another promising terpene biotransformation is the use of immobilized 3-β-hydroxysteroid dehydrogenase from *Cel-*

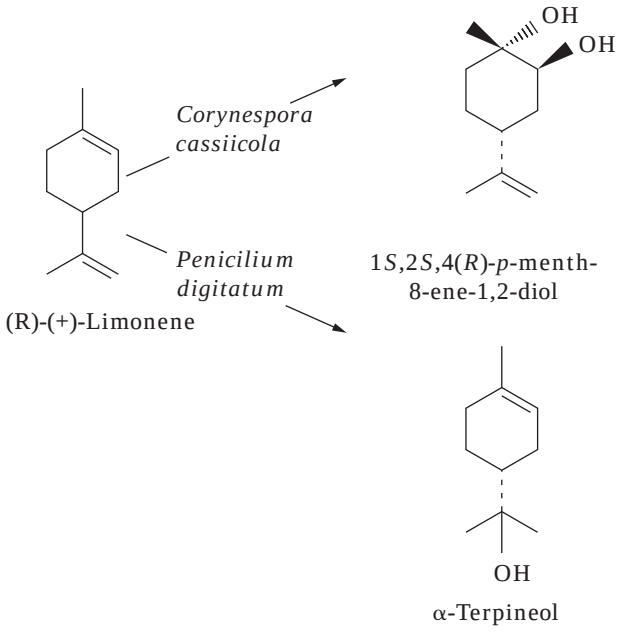


Figure 15. Microbial biotransformation of the abundant monoterpene (*R*)-+-limonene.

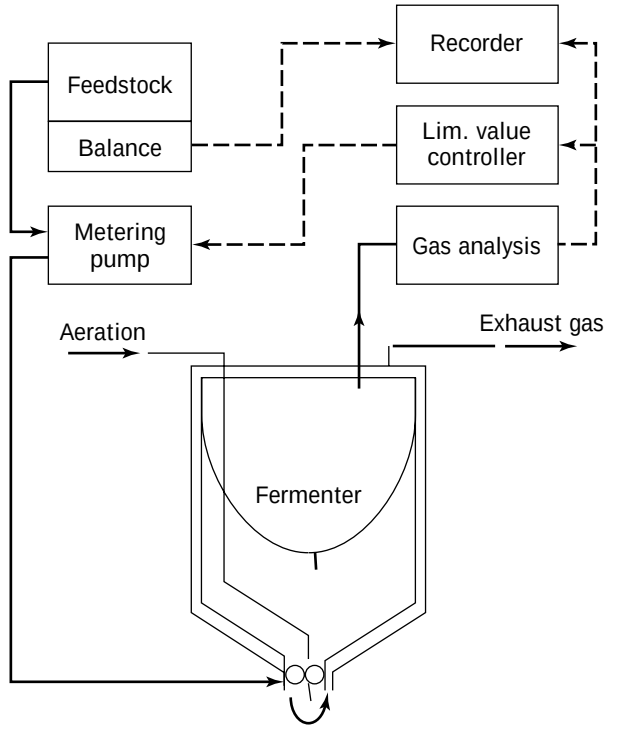


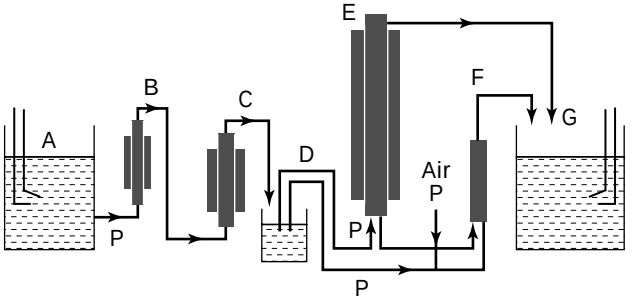
Figure 16. Pilot-scale reactor developed for the biotransformation of terpenes.

lulomonas turubata to convert L-menthone into L-menthol (50). In this reaction, NADH was recycled by oxidation of L-methylisobutyl carbinol by the same enzyme. Although this dehydrogenase was active only on L-menthone at 0.3% of its activity toward its natural steroid substrate, it was stable, with a half-life for the immobilized enzyme of 25 days, and produced 25 g L^{-1} of L-menthol product. Reactor productivity was improved by continuous in situ product removal by adding solvent to extract the L-menthol, with selective passage of solvent out of the reactor through a hydrophobic membrane. This process is potentially very useful for industrial use because immature peppermint plants produce high yields of an essential oil that consists mostly of L-menthone, only 40% of which is converted into the L-menthol and is normally harvested, which opens up the possibility of a higher-yielding process involving harvesting the immature peppermint oil and then producing L-menthol by the sort of dehydrogenase reaction described earlier.

Accelerated Soy Sauce Production. Traditional soy sauce takes 6 months to make, which adds to the costs of the process. For a description of the traditional process, see Springham (51). Osaki et al. (52) succeeded in making soy sauce of acceptable flavor with a system of sequential immobilized cell columns. The system shown in Figure 17 has been operated on a 280-L scale. Best results were obtained using *P. halophilus*, *Zygosaccharomyces rouxii*, and *Candida versatilis* cells immobilized on ceramic supports and run in series, which required just a 6-day reaction period.

Dairy Flavors. The variety of different yogurt, cheese, and other flavors is well known, each being due to a different mixture of flavor chemicals, most of which are produced by microbial action (53). These include proprionic and butyric acids, peptides and amino acids, δ- and γ-lactones, methylketones, pyrazines, diacetyl and acetoin, acetaldehyde, phenylethanol, methanediol, dimethylsulfide and even phenol. Another important factor is control of the degree of bitterness caused by peptides containing hydrophobic amino acids by use of the most appropriate mixture of proteases and peptidases.

The original objective of bioprocesses was to reduce costs by accelerating the natural fermentation and thereby reducing storage times and costs. The next objective has now become a quest for improved flavor quality, and even the generation of novel flavors. Since the actual flavor development process remains the same, the goal is to develop superior starter cultures. Microbial improvement of cheese making is a challenging task for several interrelated reasons. First, mixed cultures are common, so there is often sequential growth of different strains, with growth of the second strain taking place after the pH has been changed and using metabolites produced by the first strain. Then there is the solid nature of the cheese as the fermentation medium, with both surface growth of strains and growth within the cheese taking place. Consequently both microbial growth and flavor production can be very slow compared with submerged fermentation. The type of microbial growth that takes place also seems to influence the flavor of the cheese produced. For surface-ripened cheeses, such as Camembert and Brie, flavor development is dominated by proteolysis and the resulting peptides formed, whereas for cheeses ripened by internal mold growth, such as Roquefort and Danish Blue, the flavor is dominated by methyl ketones and fatty acids, respectively. A well-established innovation is the so-called enzyme modified cheese flavors available from a wide range of suppliers. These are made by treating cheeses with selected lipases and proteases and result in well-balanced cheese flavors with intensities up to 20 times those of the conventional mature cheeses.



A: Feed tank
B: Glutaminase reactor
C: *P. halophilus* reactor
D: Buffer tank
E: *Z. rouxii* reactor
F: *C. versatillis* reactor
G: Product tank
P: Feed pump
Air P: Air pump

Figure 17. Pilot-scale process for the accelerated production of soy sauce using three different immobilized cell reactors in series. Preceded by immobilized glutaminase reactor.

Savory Flavor Processes. *Gluconobacter suboxydans* is used to manufacture 5-ketogluconic acid from glucose as the key precursor for the important meat flavor 4-hydroxy-5-methyl-2,3-(2*H*)-furanone. The precursor is easily recovered from the bioconversion as its calcium salt and is then converted into the furanone by heating. A related sulfur-containing furanone is formed by reacting in the presence of cysteine (54).

The process described in Figure 18 has been developed by Quest International. It can be seen as a more advanced form of yeast flavor built up from traditional yeast extracts (Fig. 7). By varying the types of enzyme and microorganisms used, the flavor profile of the product can be modified to best match the needs of particular applications (28). Additional flavor character and range can also be created by a subsequent reaction flavor step.

THE APPLICATION OF GENETIC ENGINEERING TECHNIQUES TO THE PRODUCTION OF BIOFLAVORS

The public acceptance of genetically engineered products in the food industry is not very well developed. Also there is considerable opposition to being able to assign the label "natural" to flavors produced by genetic engineering. Hence, there are still not very many good examples of successful processes using this technology. However, there are

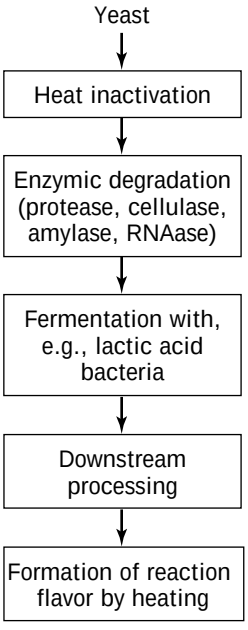


Figure 18. The major operations involved in the manufacture of savory flavors by combination of enzyme treatment of yeast and subsequent reaction flavor formation.

many instances of well-researched processes that are waiting in the wings to be used once consumer perceptions have changed. What is quite well established is the use of enzymes produced on an industrial scale using engineered producer strains in which the gene for the required enzyme has been cloned into a host strain with good fermentation characteristics. These include lipases, and chymosin for milk clotting. A comprehensive review by Muheim et al. (55) also describes the use of genetic engineering technology to make debittering proteases, to produce a flavorsome beefy meaty octapeptide, and to make *cis*-3-hexenol and methional.

One very effective use of genetic engineering is to remove the off-flavor caused in beers by the accumulation of diacetyl and thereby reduce the time required for the slow metabolism of diacetyl by the maturing yeast, which converts it into acetoin. The diacetyl is formed from α-acetolactate, and it can either be prevented from forming, or be rapidly removed before it can be converted into diacetyl. In order to prevent its formation, acetolactate decarboxylase (ALDC) can be cloned into the yeast. Yeasts transformed with the *Acetobacter aceti* ssp. xylinum ALDC gene have been used for beer production at the pilot scale and produce much less diacetyl during fermentation than their untransformed parents, resulting in a significant saving in maturation time and therefore production costs (56).

Diacetyl formation can be reduced by modifying genes in the amino acid biosynthetic pathway that lead to α-acetolactate, especially the IL V2 gene that codes for acetolactate synthetase, followed by mating of these strains to generate hybrids that produce both little diacetyl and good quality beer (57).

Another success is the development of *B. subtilis* strains that produce the flavor enhancer IMP in enhanced yields because of an absence of IMP dehydrogenase that is responsible for its metabolism into guanosine synthesis. Alternatively guanosine production can be enhanced by increasing the number of gene copies of this enzyme (58).

DOWNSTREAM PROCESSING

Biochemical engineering has the key role of transforming promising laboratory discoveries into cost-effective and reproducible large-scale manufacturing processes in a timely fashion and with minimal expenditure. This usually involves real-time analysis of pH and dissolved oxygen and carbon dioxide levels and sophisticated data logging and analysis as well as a careful consideration of the chemical and physical properties of the materials involved, for instance, the separation characteristics of biological materials are very often pH dependent. Also of critical importance is maintaining good sterile or hygenic conditions (e.g., by using cleaning-in-place methods), product specifications and cost targets, by-products and their disposal costs, the intended scale of production and any economies of scale possible, effects on the sizing of process equipment, operating regime (e.g., batch vs. continuous reactors, distillation vs. solvent extraction of product), and safety and good management practice constraints including process validation requirements. Of particular consideration for flavor and aroma products is that they often consist of a mixture of different chemicals, many of which are necessary to give the product its real value. As well as downstream processing, upstream processing is also important. This is because so many flavor and aroma products are manufactured from natural raw materials extracted from plants, so upstream processing operations for such intrinsically variable raw materials also pose special challenges, especially for successful process integration.

The main problems that are usually associated with bioprocessing are the low concentrations of product(s) that are formed, and consequently the need to remove large amounts of water, and the instability of these biochemical products with respect to heat and other extreme conditions that may occur during their purification, concentration, and recovery. Consequently the costs of downstream processing can be greater than those of the fermentation or bioconversion unit operation and can be as great as 50% of the overall processing costs.

Because the value of a flavor or aroma chemical is so dependent on its taste and smell qualities, the effects on these of trace impurities carried over from fermentation media and other contaminants can be very significant. For example, in a process to produce δ-decalactone flavor, considerable efforts were made to purify the 11-hydroxypalmitic acid precursor to an acceptable purity so as not to compromise the flavor quality of the product (29). This involved extraction of the botanical source (sweet potato residue or jalap roots) followed by hydrolysis and purification.

Obviously it is the performance of the overall process that determines the success of the process and the product it produces. Therefore the integration of all the individual unit operations, whether bioproduction steps or product isolation operations, is vitally important. Also important

is the reproducibility of the process; an important factor in many processes for natural flavors is the variability in the quality and prices of raw materials.

As compared with chemical processing, precursor and product inhibition effects often dominate bioreactions, necessitating specialized techniques including fed-batch reactions, continuous supply of precursors, use of slow-release forms of products, two-phase reactions (in which the second phase or resin acts as a product reservoir for product sequestered out of the aqueous phase), loop reactors (in which product is recovered by an in-line resin column or membrane extraction unit), and the recovery of product by its adsorption from the exit gases of a reactor. The potential for such sophisticated bioengineering is illustrated by the good productivities already achieved for some nonflavor products such as L-tertiary leucine and *N*-acetylneuraminic acid, which have productivities of 638 and 470 g product (L day)$^{-1}$, respectively, using membrane reactors in which the product is continuously removed and the cofactor is regenerated in situ, allowing the enzymes to work at as high a rate as possible (59).

Other important considerations are the requirements of good manufacturing practice, including process validation and safety audit procedures and for meeting "natural flavor" and kosher criteria. For instance, a European Community directive on natural flavors exists, and the list of extraction solvents that may be used during the processing of raw materials for food stuffs, of food components, or of food ingredients is given in Table 9 (60).

On occasions the maximum permitted amounts of residue solvents are defined, and multiple extractions are required to obtain good yields. It may also be necessary to use more sophisticated extraction equipment, such as rotating disc extractors. In general the solvents used should be nonvolatile (for safety and to prevent contamination of the aroma of the product), not prone to emulsion formation, and immiscible with water to facilitate solvent recovery and reuse and minimize environmental problems, and cells should be dried prior to solvent extraction so as to prevent emulsion formation with residual water.

In any consideration of the use of solvents, the importance of good quality water is vital. An early illustration of this in the history of biotechnology is that the best beers were brewed using the special, high-quality water obtained at Burton-on-Trent, U.K. The subsequent introduction of ion exchange for the purification of water for brewing allowed the production of superior quality beers to become widespread.

A good account of the biochemical engineering aspects of process development is given by Chisti and Moo-Young (61). These include the following unit operations.

Chromatography

Chromatography is very often expensive and difficult to fit into processes, and so it is used only for high-value products. Reverse-phase chromatography is very useful because the desired material can be absorbed straight from the aqueous phase in which it has been produced. Chromatography is usually used after clarification to remove any impurities that would foul the column, such as proteins released by any cell lysis, and then product is recovered by washing with ethanol, in which form the flavor can often be used directly. An example is the malt flavor produced by fermentation using *Streptococcus lactis* var. maltigenes, which (following ultrafiltration to remove cell mass and protein remaining from the fermentation media) can either be subject to reverse osmosis and then spray-dried with a carrier to make a powder form of the flavor, or steam distilled and the reverse-phase chromatography carried out with ethanol as the eluent to make a liquid flavor (29).

Cell Disruption

Cell disruption is necessary for the liberation of intracellular products. This imposes two constraints on the process. First, the amount of product that can be made is limited to that which can be accumulated within the cell, whether the intracellular product is an enzyme or a small molecule. Second, cell disruption liberates most of the contents of the cell together with the desired product, and so considerable purification is often required. It is also interesting to note that cell disruption is a unit operation that has had to be developed de novo in response to the need from bioprocesses. Therefore, unlike many other operations such as filtration, centrifugation, and crystallization, cell disruption has no equivalent in conventional chemical engineering.

Centrifugation

Centrifugation is required variously to recover cells, for the clarification of liquids prior to extraction, chromatography etc., for breaking emulsions (such as are often formed by cell disruption), or for washing cells prior to another operation such as disruption or immobilization.

OTHER CONSIDERATIONS

Many other biological factors are also important in the production and processing of flavor chemicals. Trace contaminant strains can introduce traces of undesirable flavor component that make the product less valuable or even unusable. The BOD of fermentor wastes needs to be minimized. Exit gases from fermentation need to be sterilized often by microfiltration. Noxious smells should be removed prior to emission into the local environment, although flavor factories are often sited with respect to prevailing winds so that emissions are less likely to reach major sites of population (62). In biotransformation processes the precursors will ideally be appreciably water soluble and with minimal cytotoxicity. As in any bioprocess, the media will be formulated so as to achieve a number of aims, including minimization of the lag phase, the inoculum size required, and the duration of the fermentation. Other objectives are to minimize power input costs and to maximize mixing so as to maximize the utilization of expensive filtered air.

Perhaps the newest innovation in flavor recovery technology are the use of in situ solvent extraction (Table 10), and of perevaporation, which has been demonstrated to be effective for the in situ recovery of benzaldehyde from fer-

Table 10. List of Solvents Accepted for Use in the Manufacturing of Natural Flavors

Name
Propane
Butane
Butyl acetate
Ethyl acetate
Ethanol
Carbon dioxide
Acetone
Nitrous oxide

mentation broth using *Bjerkandera adusta* (63). Another continuous in situ product recovery method is the continuous addition of an extracting solvent to the reaction vessel, with its selective removal together with the product using a hydrophobic membrane (50). In addition pervaporation has been used to recover 6-pentyl-α-pyrone from *Trichoderma viride* cultures, obtaining an 85% yield of product and a 20-fold enrichment after 2 weeks of fermentation (64).

COMPARISON OF DIFFERENT PROCESSES

In comparing various processes, the productivity of the biocatalyst is very important. As a general guide, Table 10 provides an indication of the typical rates of enzyme-catalyzed bioorganic reactions of various substrate classes in bacteria.

Production of flavors by plant cells is a very desirable method that still requires major advances to make it both practicable and cost-effective. These advances may be developed for pharmaceutical products and then transferred to flavor processes. For instance, Phyton has made big advances in developing plant cell fermentations to make paclitaxel, the active ingredient in Bristol-Myers Squibb's taxol anticancer drug. Therefore, it is interesting to note that, by comparison with the microbial systems already described, it is reported that the rates of capsaicin biosynthesis in ripening pepper plants is ca. 0.5 mg g dry wt^{-1} day^{-1}, and in pepper plant cell culture it is ca. 0.1 mg g dry wt^{-1} day^{-1}, which is equivalent to a turnover rate of 0.00006 mmol g^{-1} h^{-1}. To make progress it is usually necessary to elucidate the biochemical pathway the plant uses to make the flavor. For instance, the biosynthesis of Furaneol by strawberries (2,5-dimethyl-4-hydroxy-(2*H*)-furan-3-one) has been recently described (65), as has the biosynthesis of raspberry ketone by coupling coumaroyl-CoA and malonyl-CoA, followed by reduction, by raspberry fruit (66).

If sufficient research and development resources are available, biocatalysts with very high productivities can be developed, and these have created vast industries. For example, consider the commodity products made by biocatalysts, as shown in Table 11. Because flavor and aroma chemicals are much higher-value products, the bioprocesses can afford to be less productive but still need to achieve good productivities because the volume of demand is not as great as for other products. For instance ca. 10^7 tons/year of high fructose corn syrup are made, but processes for flavors and aroma chemicals do not benefit from

Table 11. Examples of Productivities That Have Been Achieved in Bioprocesses for Flavor and Aroma Chemicals

Products	Concentration of product in reactor (g L^{-1})	Scale of production (t year^{-1})
Diol precursor of Ambrox	150	
Citric acid	120–150	500,000
Acetic acid	130	
C13 dicarboxylic acid precursor of musk	140	
ω-Hydroxy hexadecanoic acid (hexadecanolide precursor)	>100	
Glutamic acid, especially as monosodium glutamate	80	300,000
Methylketone (2-heptanone)	75	
Coriolic acid (precursor of δ lactones)	25	
Esters	30	
Acetaldehyde	28	
Butyric acid	20–30	
L-Phenylalanine[a]	25	
Fructose in high-fructose corn syrup[a]	17	7,000,000
Diacetyl	14	
Pyrazines	12.5	
Penicillin[a]	10–15	5,000
β-Cyclohomogeraniol	9	
Rosmarinic acid[a]	6.7	
Proprionic acid	6.3	
Gamma decalactone	5	
Trans-2-hexenol	4.2	
Tetramethylpyrazine	3	
Lactic acid	2.8	
Ionones	1	
6-Pentyl-2-pyrone	0.03	

[a]Nonflavor/aroma benchmark products.

such big economies of scale. The concentration at which a product can be produced in an enzyme or microbial reaction is a good guide to how cost-effectively it can be manufactured. Table 11 is illustrated by examples of high productivities that biocatalysts must reach to achieve a successful and widespread industrial use (Table 12). High reaction productivity must be complemented by effective downstream processing and by good process intensity in each unit operation of the manufacturing process. This is because product concentration reflects the size of equipment, especially of reactors, and the hydraulic throughput rates that are required to make a given amount of product, and thus influences both capital costs and costs of mixing, drying, pumping, and so forth. Also the product concentration indicates the volumes of fluid that will have to be processed to obtain the required amount of product and the cost of concentrating and isolating the product. The relationship between production concentration and sale price (Table 11) is well established for a wide range of bioproducts ranging from very high value biopharmaceuticals to low-cost food ingredients such as citric acid. Thus, if a flavor material can be produced only at quite low concentrations, it is very likely that the only market opportunity for it will be in prestige and/or high-value niche markets, if they exist. A major constraint on the product concentrations that can be achieved is product inhibition. Therefore, a determination of inhibition constants is as important as determining the rate of reaction parameters, and work on developing effective in situ product recovery methods that will remove product as it is made are very important. For flavor molecules, such in situ product recovery techniques can include the recovery of volatile flavors from the exit gases of bioreactors. Product inhibition is, of course, exacerbated if it is difficult for the product to diffuse out of the cell in which it has been produced.

FLAVOR APPLICATIONS CONSTRAINTS

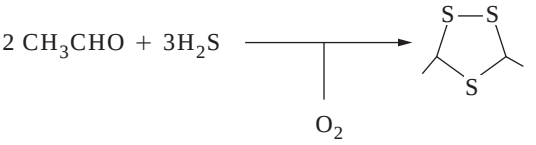
Once a flavor has been produced it must always be remembered that its value will vary with its applications. A graphic example is provided by vanilla flavor when used in low-fat ice creams. Reformulation of the flavor is required because the less oil soluble flavor chemicals are less well perceived in the low-fat products and so need to be present in relatively higher concentration than does the more oil soluble flavor component in traditional high-fat ice creams (67).

Table 12. Comparison of the Productivities of Three of the Most Economically Important Biotransformation Manufacturing Processes for Chemical, Pharmaceutical, and Food Applications

Nitrile hydratase (Nitto Chemicals)	50 g Acrylamide/L reactor/h
Penicillin G amidase (SmithKline Beecham and others)	1–2 tonnes 6-APA/kg of immobilized enzyme
Glucose isomerase (Novo Nordisk and others)	20 tonnes HFCS/kg of immobilized enzyme

Big advantages can be obtained if positive applications results can be obtained, for example, in creating reductions in fat and salt contents of products, improved flavor stability and controlled release, emulsion stabilisation, as well in as the growing area of health benefits (nutraceuticals).

Changes can also occur during food processing and cooking, especially when extreme conditions are used. For instance, so-called reaction flavors can be formed from flavor components during heating, particularly via the Mailard reaction that occurs between reducing sugars and the amino groups of amino acids. Condensate flavors can also arise upon cooling, especially when local high concentrations of reactants occur. An example is the reaction of acetaldehyde and hydrogen sulfide to form 3,5-dimethyl-1,2,4-trithiolene.



Product Approval and Registration

The necessary tests and procedures for product approval and registration can be both lengthy and expensive, and different countries often have different requirements. The most difficult task is to gain approval for a new chemical; approval for new processes to make existing materials with a proven record of safety is easier to gain. The ideal is “generally regarded as safe” approval, which indicates that long-term and significant consumption in the diet results in no evidence of any adverse effects.

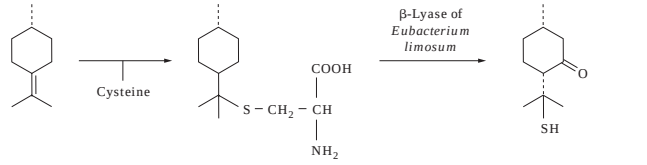
Consumer Approval

However novel the underlying science and however efficient the performance of the manufacturing process, the product will be successful only if it achieves both regulatory approval and consumer acceptability, and indeed, generates a marked consumer preference because of the perceived value of the products to individual consumers and interest groups. This is a difficult area not least because of the marked difference between lay and expert perceptions of risk, with the consumer tending to put more weight on the risk of new uncontrolled factors and the effects directly on themselves and their families, rather than relying on the statistical risk factors based on broad demographics. In particular individual consumers fear loss of control, particularly to organizations that they find difficult to trust because of their vested interests that may act counter to their own interests, such as companies' profit motive. So in this respect there is a very big difference between the use of bioprocesses for flavor and for pharmaceuticals. Whereas, both types of product have very high biological activities and are discovered and then developed along similar lines, the consumer needs and decisions for their uses are quite different. There is a strong “cure” imperative for drugs, which are usually used only an occasionally, whereas the use of food flavors is much more an everyday pleasure-based activity, leading to an ethical concern expressed as a preference for natural-flavored products. For

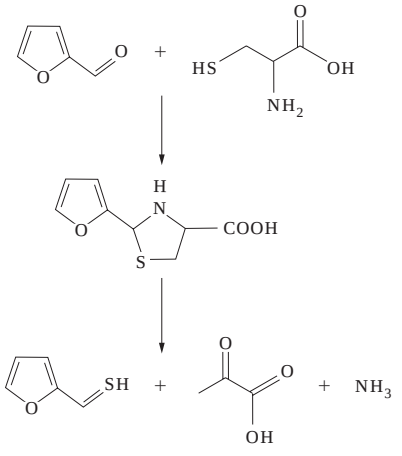
instance, Abraham et al. (66) reports that about 70% of the food flavors used in Germany in 1990 were natural.

FUTURE OPPORTUNITIES

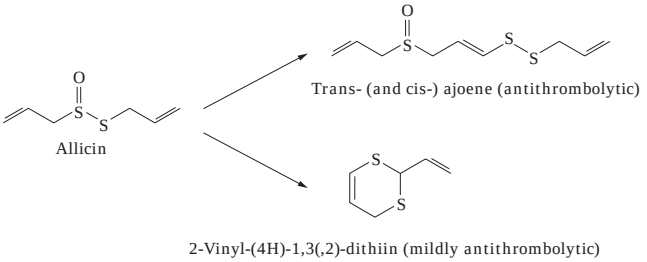
One of the recurrent themes in flavor and fragrance biotechnology is the need to oxygenate molecules, such as terpenes, in order to convert them into molecules with higher-intensity flavors. This can be extended to the introduction of sulfur, as not only is sulfur close to oxygen in the periodic table, but many important flavor molecules contain sulfur. For instance, an important component of blackcurrant flavour has been biosynthesized as follows (68):



Similarly there is the biosynthesis of furfurythiol using the β-lyase activity of an *Enterobacter* strain (69):



Healthy, indeed so-called nutraceutical food ingredients, are gaining increasing consumer recognition and commercial momentum. Natural flavors with some health benefits may be possible. Allicin is an important precursor of flavor compounds in garlic and onion; allicin has some antibacterial properties but also forms molecules with marked antithrombolytic properties.



CONCLUSIONS

Despite biopharmaceuticals being widely regarded as the cutting edge of modern biotechnology, a surprisingly high proportion of current flavors and food ingredients are manufactured using bioprocesses. These include not only refined and developed versions of traditional processes, such as for cheese, yogurt, and protein hydrolysates, but also many new and sophisticated bioprocesses for key chemically defined materials such as the flavor chemicals methyl ketones, and γ-decalactone, and flavor enhancers such as GMP and IMP. This article has attempted to explain the technical basis for these products as used in industry and indicate exciting future developments in this field that are based on the pull of market demand forces but still constrained by cost, patents, labeling, and regulatory considerations. However, a comprehensive account is not yet possible because full details are not available concerning the bioprocesses now known to be operating. In many cases it is even uncertain at what scale processes are operating.

In conclusion, some important commercial factors that must be considered including the following:

- Significant market demand (especially for label-friendly natural ingredients and for products with superior performance such as fat reduction, health benefits, or improved convenience)
- Substantial profit margins
- Economies of scale of production
- Availability of low-cost raw materials
- Patentable (proprietary) technologies
- Weak competition
- A range of products with different market uses from a single process
- Ease of regulatory approval

As regards cost, Welsh (Ref. 3) has suggested a relationship between selling price and market size, which has been reinterpreted in Figure 19. This shows that although the highest-priced products, up to $1,250 per kg, appear superficially most attractive, the largest markets are for the

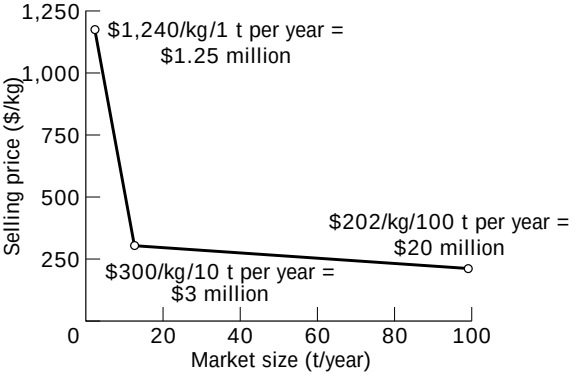


Figure 19. A comparison between the selling price and market size for flavor chemicals, with the market values of the products rising from $1.25 million, to $3 million, to $20 million as the tons sold per year increases and the sale price decreases.

lower-priced products, with an approximate market size of $20 million, as compared with only $3 million and $1.25 million for the higher-priced products. These values of natural flavors do not compare unfavourably with the price of gold, currently about $12,000 per kilogram. Much exciting research is taking place, and the pace of change in the food and beverage industries is intense, which is creating many new opportunities for bioflavors, especially considering the sustained demand from consumers for products containing natural ingredients. However, these opportunities must be balanced against constraints, including the availability of natural raw materials in sufficient quantity, quality and low price, the difficulty in developing bioprocesses that give high conversions of concentrated substrate streams into products with no by-products, and the difficulties in easily concentrating and purifying the required bioproducts. Perhaps in this last respect more efforts should be made to obtain the benefits of biomolecular selectivity for product isolation and recovery.

The objective of bioprocesses in the flavor and fragrance industry is to produce the highest quality product at the minimum processing costs. I hope this chapter illustrates the complex interrelationships between what is needed (consumer demands), how to make it (the development of bioprocesses), and also the important conditioning factors, such as what it is possible to patent, safe operation, and the need to meet other demands, such as for naturalness. Perhaps this chapter can in some small way help others in this exciting area of technology to succeed in developing more new and improved bioprocesses for flavors.

BIBLIOGRAPHY

1. T. Godfrey and S. West eds., *Industrial Enzymology*, 2nd ed., Macmillan, New York, 1996.
2. T. Scheper managing ed. and R.G. Berger volume ed., *Advances in Biochemical Engineering*, vol. 55, Springer, Berlin, 1997.
3. A. Gabelman ed., *Bioprocess Production of Flavour, Fragrance, and Color Ingredients*, Wiley Interscience, New York, 1994.
4. H. Hartmann, Perfumer and Flavourist **20**, 35–42 (1995).
5. H. Hartmann, Perfumer and Flavourist **22**, 21–23 (1996).
6. L.P. Somogyi, *Chem. Ind.* 170–173 (March 1996).
7. A. Hausler and T. Munch, *ASM News* **63**, 551–559 (1997).
8. J. Hariel, data of 1988–1990 quoted in R.G. Berger, *Aroma Biotechnology*, Springer, Berlin, 1995, pp. 170.
9. C.H. Manley, in A. Gabelman ed., *Bioprocess Production of Flavour, Fragrance, and Color Ingredients*, Wiley Interscience, New York, 1994, pp. 19–40.
10. P.P.C. Graziadei, *Cytotechnology* **11**, 3 (1993).
11. L. Bussolati, R. Ramoni, S. Grolli, G. Donotrio, and E.J. Bignetti, *Biotechnology* **30**, 225 (1993).
12. H. Zhao and S. Firestein, *Science* **279**, 237 (1998).
13. R.G. Berger, *Aroma Biotechnology*, Springer-Verlag, Berlin, 1995.
14. S. West and D. Pawlett, in T. Godfrey and S. West eds., *Industrial Enzymology*, 2nd ed., Macmillan, New York, 1996, pp. 209–225.
15. P.E. Milson, in R.D. King and P.S.D. Cheetham, eds., *Food Biotechnology*, vol. 1, Elsevier Applied Science, London, 1987, pp. 273–308.
16. H.L.M. Lelieveld, *Process Biochem.* **19**, 112–113 (1982).
17. T.W. Nagodawithana, in A. Gabelman ed., *Bioprocess Production of Flavour, Fragrance, and Color Ingredients*, Wiley Interscience, New York, 1994, pp. 135–168.
18. P.S.J. Cheetham, in A. Wiseman ed., *Handbook of Enzyme Biotechnology*, 3rd ed., Ellis Horwood, Chichester, 1995, pp. 83–234.
19. H.-P. Meyer, *Biotech. Forum Eur.* **8**, 602–606 (1991).
20. S. West, in T. Godfrey and S. West eds., *Industrial Enzymology*, 2nd ed., Macmillan, New York, 1996, pp. 155–175.
21. D.W. Armstrong and L.A. Brown, in A. Gabelman ed., *Bioprocess Production of Flavour, Fragrance, and Color Ingredients*, Wiley Interscience, New York, 1994, pp. 41–94.
22. C.E. Fare, P.J. Blanc, A. Marty, G. Goma, I. Souchen, and A. Voilley, *Perfumer and Flavourist* **21**, 27–39 (1996).
23. P. Fulbrook, in T. Godfrey and J. Reichelt eds., *Industrial Enzymology*, 1st ed., Nature Press, London, 1983, pp. 8–110.
24. P.S.J. Cheetham, *Biocatalysis* **6**, 1–30 (1992).
25. P. Winterhalter and P. Schreirer, *Flavour and Fragrance* **J9**, 281–287 (1994).
26. European Pat. EP 416713 A1 (1991), Gist-brocades.
27. H.M. Lelieveld, *Proc. Biochem.* **19**, 112 (1984).
28. European Pat. 850245 (1986), J.F.M. de Rooij and M.J.J. Hakkaart.
29. H.J. Takken, N.M. Groesbeck, and R. Roos, in R.L.S. Patterson, B.V. Charlwood, G. Macleod, and A.A. Williams eds., *Biotrans of Flavours*, Royal Society of Chemists, Cambridge, 1992, pp. 186–200.
30. N.J. Izawa, *Agric. Food* **45**, 543 (1997).
31. U.S. Pat. 5,077,206 (1991), P.S.J. Cheetham and M.A. Quail.
32. P. Hugueny, B. Dumont, F. Ropert, and J.M. Belin, in P. Etievant and P. Schreier eds., *Bioflavour '95*, INRA, Paris, 1995, p. 269.
33. U. Krings, M. Hinz, and R.G. Berger, *J. Biotechnol.* **51**, 123–129 (1996).
34. M. Starkle, *Biochemistry* **2**, 151, 371 (1924).
35. U.S. Pat. 5,185,252 (1993), A. Humphrey et al.
36. C. Larroche and J.B. Gross, in R.G. Berger, ed., *Advances in Biochemical Engineering*, 1997, pp. 174–220.
37. I. Gatfield, *Perfumer and Flavourist* **20**, 5–14 (1995).
38. I.L. Gatfield, M. Guntert, H. Sommer, and P. Werkhoff, *Chem. Mikrobiol. Technol. Lebensm.* **15**, 165–170 (1993).
39. French Pat. 2 705 971 (1993), H.E. Spinnier, L. Dufosse, I. Souchen, A. Latrasse, C. Piffaut-Juffard, C. Voilley, and A. Philippe.
40. Japanese Pat. 01 168 256 (1989), Y. Shukunobe and S. Takato.
41. F. Schindler and R. Seipenbusch, *Food Biotech.* **4**, 77 (1990).
42. B.L. Muller, C. Dean, and I.M. Whitehead, in P. Etievant and P. Schreier eds., *Bioflavour '95*, INRA, Paris, 1995, pp. 339–344.
43. F. Ropert, B. Dumont, and J.M. Belin, in P. Etievant and P. Schreier eds., *Bioflavour '95*, INRA, Paris, 1995, pp. 275–278.
44. J. Rabbenhorst, *Appl. Microbiol. Biotechnol.* **46**, 470–474 (1996).
45. M.J. Gasson, Y. Kitamura, W.R. McLauchan, A. Narbad, J. Parr, E.L. Parsons, J. Payne, M.J.C. Rhodes, and N.J. Walton, *J. Biol. Chem.* **273**, 4163–4170 (1998).
46. R.C. McIver and G.A. Reineccius, *ACS Symp. Ser.* **317**, 266–274 (1986).
47. U.S. Pat. 4,963,480 (1990), B. Belcour, D. Coutois, and C. Ehret.

48. W.R. Abraham, H.M.R. Hoffmann, K. Kieslich, G. Reng, and B. Stumpf, in *Enzs in Orig. Syn. CIBA Foundation Symposium* III, 1985, pp. 146–160.

49. Y. Suhara, S. Itoh, M. Ogawa, K. Yokose, T. Sawada, T. Sano, R. Ninomiya, and B. Maruyama, *Appl. Environ. Microbiol.* **42**, 187–191 (1981).

50. S. Kise and M.J. Hayashida, *Biotechnology* **14**, 221 (1990).

51. D.G. Springham, in V. Moses and R.E. Cape eds., *Biotechnology: The Science and the Business*, Harwood Academic, 1991, pp. 249–296.

52. K. Osaki, T. Ishiyama, and H. Motai, *Appl. Microbiol. Biotechnol.* **31**, 346 (1989).

53. D.C. Eaton, in A. Gabelman ed., *Bioprocess Production of Flavour, Fragrance, and Color Ingredients*, Wiley Interscience, New York, 1994, pp. 169–203.

54. European Pat. 58870 (1982) (to Unilever).

55. A. Muheim, A. Hausler, B. Scilling, and K. Lerch, *Perfumer and Flavourist* **23**, 21–27 (1998).

56. S. Tada, T. Takeuchi, H. Sone, S. Yamono, M.A. Schofield, J.R.M. Hammond, and T. Inoune, *Proc. Eur. Brewery Conv. Cong.*, Brussels, 1995, p. 369.

57. M.C. Kielland-Brandt, C. Gjermansen, T. Nilsson-Tilgren, and S. Holmberg, *Proc. Eur Brewing Conv. Cong.*, Zurich, 1987, p. 37.

58. K. Miyagawa, et al. *Biotechnology* **7**, 821–825 (1989).

59. U. Kragl, M. Kittelmann, O. Ghisalba, and C. Wandrey, *Ann. N.Y. Acad. Sci.* **750**, 300–305 (1995).

60. EEC Council Directive 13.6.88 (88/334/EEC).

61. Y. Christi and M. Moo-Young, in V. Moses and R.E. Cape eds., *Biotechnology: The Science and the Business*, Harwood Academic Publishers, 1991, pp. 167–209.

62. T.O. Williams and F.C. Miller, *Biocycle* **33**, 72 (1992).

63. T. Lamer, H.E. Spinnler, I. Souchon, A. Voilley, *Process Biochem.* **31**, 533–542 (1996).

64. W. Bonejsza-Wysocki and G. Hrazdina, *Phytochemistry* **35**, 623 (1994).

65. I. Zabetakis, J.W. Gramshaw, and D.S. Robinson, in H. Maarse and D.G. vd Heij eds., *Trends Flav. Res. Proc. 7th Weurman Symp.*, Elsevier, Amsterdam, pp. 91–93.

66. B. Abraham, U. Krings, and R.G. Berger, *GIT Fachz Lab.* **38**, 370–375 (1994).

67. E. Graf and K.B. de Roos, in R.I. McGorrin and I.V. Leland, eds., *Flavour–Food Interactions*, American Chemical Society, Washington, D.C., 1996.

68. U.S. Pat. 5,182,194 (1988), A. Kerkenaar, D.J.M. Schmedding, and J. Berg (to Int. Flavors and Fragrances).

69. P. Van der Schaft, I. van Gee, G. de Jong, and N. ter Berg, in H. Maarse and D.G. van de Heil eds., *Trends in Flavour Research*, Elsevier, Amsterdam, 1994, p. 437.

ENZYMES, FRUIT JUICE PROCESSING

CATHERINE GRASSIN
PIERRE FAUGUENBERGUE
Gist-brocades
Seclin Cedex, France

KEY WORDS

Cell wall
Enzyme
Fruit
Juice
Pectin
Process

OUTLINE

INTRODUCTION

The fruit juice industry has to deal with many intricate problems in dealing with material whose composition varies widely depending on the variety of the fruit, the year, and the fruit ripeness. The main objective of the fruit juice industry is to process as much fruit as possible at the lowest cost, while maintaining, or if possible improving, the organoleptic quality and the stability of the finished product. Because it is very viscous and forms a gel, fruit pectin is responsible for numerous problems encountered during processing. The first application of pectinases for apple juice clarification started in the 1930s (1) (Table 1).

The induction of the juice clarification after the pectin breakdown by pectinases and the decrease of the juice viscosity have resulted in a shorter process, improving greatly the quality of industrial apple juice. Enzymes have been applied to fruits other than apples, such as pear, black currant, strawberry, raspberry, cherry, cranberry, and grape. In the production of clear and stable red berry juice or concentrate, the hot maceration with enzymes ensures a higher juice yield and a better color extraction. The use of pectinases and amylases, for apple starch degradation, during the hot clarification stage has prevented postbottling haze formation, and for the first time, the concentration of the juice by a factor of six has been possible. This has resulted in a smaller storage volume and cheaper transportation together with a better concentrate stability by preventing spoilage. Apple pulp enzyming with pectinases and hemicellulases appeared later (2): by lowering the viscosity of the pulp, the press capacity and the yield were significantly improved at the same time. Juice depectinization with pectinases (and arabanases) has improved the product quality even more by preventing araban haze formation after juice concentration.

Table 1. History of Enzymes in Fruit Processing

	Product innovation	Enzyme innnovation
1930	Clear apple juice	Introduction of pectinases
1953	Berry processing/hot maceration	Heat-stable pectinases
1960	Apple juice concentrate	Amylases to prevent starch haze
1976	Hot clarification process	Pectin and starch degradation at 50 °C
1980	Apple pulp enzyming for yield increase	Pectinases + hemicellulases
1983	Stable apple juice concentrate	Arabanases to prevent araban haze
1985	Advanced maceration with decanter	Pectinases + hemicellulases + cellulases
1994	Apple liquefaction	Pectinases + rhamnogalacturonase + hemicellulases + cellulases
1995	Enzymes from genetically modified organisms	Single enzyme activity

The addition of cellulases to the fruit pulp has resulted in a liquefaction in which solids are easily separated from the juice with decanters instead of presses. The maximal degradation of apple tissue with a broad range of enzymes has been obtained with a total liquefaction process and decanters. Today, enzyme suppliers can offer to fruit juice producers tailor-made enzyme preparations, optimally blended on the base of the fruit structure for every fruit process—citrus, tropical fruit, or even vegetable. Numerous new enzymes have recently been discovered based on fruit cell wall biochemical structure. The improvement in the knowledge of fruit composition, cell wall chemistry, and enzymology in parallel with better control of large-scale enzyme production has made possible the supply of almost pure enzymes from genetically modified organisms (GMOs) blended in optimal composition based on the fruit composition. Fruit juice technologies now require more accurate transformations than the simple mechanical pressing or physical extractions. Application of exogenous enzymes leads to the degradation of fruits or the selective extraction of some of their components, allowing the creation of new types of finished products and fruit derivatives. They improve quality and stability of finished products together with a shorter process duration and a larger plant capacity. Associated with new equipments and processing technologies, industrial enzymes allow processors to add value to raw material for food and feed and to reduce waste quantity.

In this article, fruit juice processes are described nonexhaustively; in fact, processes are very variable and numerous. We will describe the most common of them, but variations exist in fruit processing companies and factories. Processing of apple, pear, red berry, stone fruit, tropical fruit, and citrus fruit is detailed. The aspect of juice authentification is discussed.

BIOCHEMISTRY OF FRUIT CELL WALL

The pulp of fruits and vegetables is composed of cells. They are surrounded by a cell wall. This is what separates and maintains the integrity of the cells in the pulp and gives them their rigidity. This cell wall stands up to internal pressure and external shocks. Polysaccharides constitute 90 to 100% of the structural polymers of the walls of growing plant cells called *primary cell walls*. *Secondary cell walls* develop from primary cell walls during cell growth. Cell wall composition varies among fruit species and evolves depending on agronomic and climatic conditions, fruit ripeness, and the method and duration of fruit storage (cell growth and cell senescence). Plant cell wall composition has been widely studied, and numerous models of three-dimensional structure have been proposed (3,4). Three major independent domains are distinguished: the xyloglucan network, the pectin matrix, and the structural proteins. The cellulose–xyloglucan network is embedded in the pectin matrix. Pectin is the major structural polysac-

charide component of the fruit lamella and cell wall. Three pectic polysaccharides are present in all primary cell walls that have been described: homogalacturonan, rhamnogalacturonan I (RG-I), and rhamnogalacturonan II (RG-II) (5). Recent models divide pectin in smooth regions as unbranched homogalacturonan and hairy regions as highly branched RG-I (6,7) (Fig. 1). Smooth region or homogalacturonan is a homopolymer of (1→4)-α-D-galactosyluronic acid residues, renowned for its ability to form gels (Fig. 1**a**). Carboxyl groups of the galactosyluronic acid residues of primary cell wall homogalacturonan can be methyl esterified at C6 position and acetyl esterified at C2 or C3 position (i.e., sugar beet). Helical chains of homogalacturonan that are less than 50% methyl esterified can form a gel-like structure and condense by cross-linking with calcium ions, which are present in primary cell walls, to form junction zones.

Methyl and acetyl esterification degrees and molecular weight of pectin are specific of fruit species (Table 2). Both decrease during fruit ripeness. With time, the protopectin in primary cell wall is slowly transformed into soluble pectin by the action of endogenous pectinases present in the fruit. The decrease of molecular weight is due to pectin depolymerases, and the decrease of esterification degree is due to pectinmethylesterase and pectinacetylesterase. However, these activities are very low. Homogalacturonan contains about 200 galacturonic acid unit chains 100 nm long, with some rhamnose making a kink in the backbone. A xylogalacturonan subunit (XGa1A) has been identified as part of galacturonan backbone substituted with xylose at the C3 position of the galacturonic acid residue. It can be methylesterified. In apples, XGa1A content is high. Hairy regions are composed of RGI. RG-I has a backbone

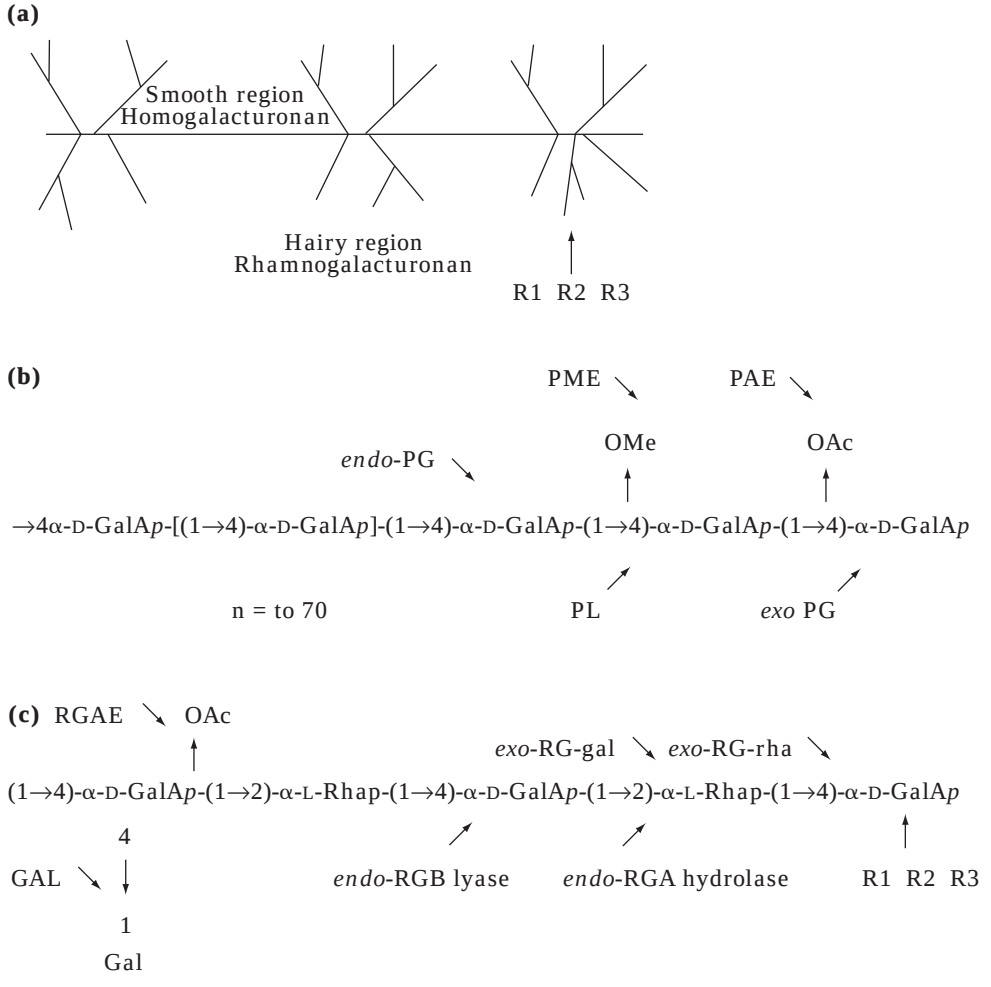


Figure 1. (**a**) Pectin structure and related degrading enzymes. R1 = arabinan; R2 = arabinogalactan type I; R3 = arabinogalactan type II. (**b**) Smooth region, homogalacturonan; PL = pectinlyase; PME = pectinmethylesterase; PAE = pectinacetylesterase; PG = polygalacturonase. (**c**) Hairy region, rhamnogalacturonan type I. RGAE = rhamnogalacturonase acetylesterase; RG-gal = rhamnogalacturonase galacturonohydrolase; GAL = galactosidase; RG-rha = rhamnogalacturonase rhamnohydrolase; RGB = rhamnogalacturonase lyase; RGA = rhamnogalacturonase hydrolase; R1 = arabinan; R2 = arabinogalactan type 1; R3 = arabinogalactan type II. (**d**) Hairy region, arabinan (R1). ARA = arabanase; ARFA = arabinofuranosidase A; ARFB = arabinofuranosidase B. (**e**) Hairy region arabinogalactan type I (R2). AGal = $\beta(1 \rightarrow 4)$ arabinogalactanase. (**f**) Hairy region, arabinogalactan type II (R3). AGal = $\beta(1 \rightarrow 3)$, $\beta(1 \rightarrow 6)$ arabinogalactanase.

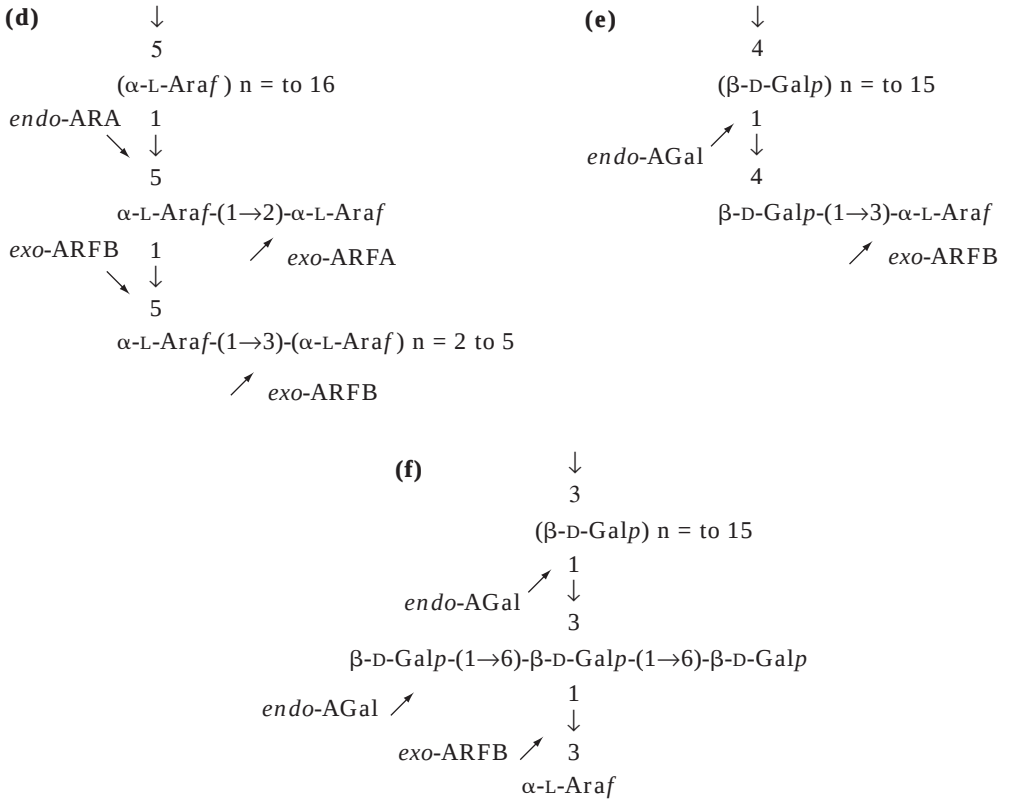


Figure 1. (*continued*).

of as many as 100 repeats of the disaccharide [→2)-α-L-rhamnosyl-(1→4)-α-D-galactosyluronic acid-(1] (5) (Fig. 1**b**). Arabinosyl- and galactosyl-rich side chains are attached to the *O*-4 of rhamnosyl residues. The side chain can vary in size from a single glycosyl residue to 50 or more glycosyl residues (5). In general, about half of the rhamnosyl units of RGI have side chains, but this can vary with cell type and physiological state. Arabinans are mostly five-linked arabinofuranosyl units forming helical chains, but arabinosyl units can be connected to each other at every free position *O*-2, *O*-3, and *O*-5, forming a diverse group of branched arabinans (8) (Fig. 1**c**). RG-I of primary cell walls is branched with (1→5)-α-L-arabinan (1→3) or (1→2)-α-L-arabinosyl residues (Fig. 1**c**), (1→4)-β-D-galactan, arabinogalactan (AG) type I with a (1→4)-β-D-galactan backbone (Fig. 1**d**), and AG type II with a (1→3),(1→6)-β-D-galactan backbone; both AGs have (1→3)α-L-arabinosyl residues (9) (Fig. 1**e**).

Arabinan and galactan of primary cell walls are present as covalently attached side chains of RGI (5) (Fig. 1). But AG would be unattached to cellulose or xyloglucan, and arabinogalactan-proteins (AGPs) are associated with the plasma membrane (10). RG-II is a low molecular weight (about 4.8 kDa), complex polysaccharide composed of 12 different sugars (Fig. 2). These different sugars are apiose, 2-*O*-methyl-1-fucose, 2-*O*-methyl-D-xylose, aceric acid (3-*C*-carboxy-5-deoxy-1-xylose), Kdo (3-deoxy-D-manno-octulosonic acid), Dha (3-deoxy-D-lyxo-heptulosaric acid) in more than 20 different linkages. RG-II has a homogalacturonan backbone composed of about nine (1→4)-α-D-galactosyluronic acid residues (11). Four different complex side chains are attached to *O*-2 or *O*-3 of four of the backbone residues. Two main hemicelluloses of all primary cell walls are xyloglucan and arabinoxylan. Minor components such as glucomannan or galactoglucomannan have also been described (5). Hemicelluloses bind tightly via hydrogen bonds to the surface of cellulose-linking or cross-linking microfibrils, creating a cellulose-hemicellulose network. Interconnections with the pectic polysaccharides are of first importance concerning the integrity of the pectin network. Cellulose accounts for about 20 to 30% of the dry matter of most primary cell walls and is particularly abundant in secondary cell walls. This is a (1→4)-β-D-glucan polysaccharide component (Fig. 3). Cellulose microfibrils give the shape and tensile strength to primary cell wall and move relative to another for cell growth but form intractable aggregates in secondary cell walls (5). Microfibrils are coated with hemicelluloses that bind tightly to their surface (12). In the primary cell wall, glucan chains are randomly arranged; the cellulose is said to be amorphous (Fig. 3). This makes the primary cell wall probably more susceptible to enzymatic degradation.

The cellulose of secondary cell walls has a rather high degree of crystallinity, with the chains running in the same direction in a parallel arrangement. In 1993–1994, Vincken et al. demonstrated that xyloglucan was a key structure of apple cell wall for the degradation of cell wall embedded cellulose (13). The cellulose–xyloglucan network

Table 2. Fruit Composition

Fruit	Pectic substances (% wet weight)	Methylation (degree %)	Endogenous enzymes	Main neutral sugars	Pectin[a]	Hemicellulose[a]	Cellulose[a]	Lignin[a]	Proteins[a]	Total[a]	Titrable acidity (% as tartaric)
Apple	0.5–1.6	65–92	PE, PG, PPO	Glu, Ara, Gal	272	169	349	2	76	868	0.5–1.4
Banana	0.7–1.2		PE, PG, PPO								0.3–0.4
Black currant	1.1	50–80	PE								3.0–4.0
Cherry	0.2–0.3	40	PE, PG, CEL	Glu, Ara, Gal	396	49	130	169	244	988	0.4–0.6
Grape	0.1–0.4	50–65	PE, PG, PPO	Glu, Gal, Ara							0.4–1.3
Guava	0.8–1.0	54–62	PE, PG, CEL, PPO, AMY	Glu, Xyl							0.9–1.3
Lemon pulp	2.5–4.0		PE								4.0–4.5
Mango	0.3–0.4	78–85	PE, PG, CEL	Glu, Xyl, Ara	408	91	236	27	127	889	0.2–1.2
Orange peel	3.5–5.5	65	PE								0.8–1.1
Papaya	0.7–1.0		PE, PG, CEL, PPO	Glu, Gal, Man	364	165	124	4	127	784	0.5–0.8
Peach	0.1–0.9	60–80	PG, CEL, PPO								0.8–1.1
Pear	0.7–0.9	50–70	PE, PG	Glu, Xyl, Ara	281	148	267	69	82	847	0.2–0.4
Pineapple	0.04–0.1	22–40	PG	Glu, Xyl, Ara	163	267	210	85	94	819	0.8–1.3
Plum	0.7–0.9	75	PE, PPO								1.4–1.7
Raspberry	0.4–0.5	20	PE, PG, CEL, PPO	Glu, Xyl, Ara	168	89	177	73	277	784	1.4–1.6
Strawberry	0.5–0.7	20–60	PE, PG, CEL, PPO	Glu, Ara, Gal	411	66	232	11	253	975	0.6–1.5

Note: PE = pectinesterase; PG = polygalacturonase; CEL = cellulase; PPO = polyphenoloxidase; AMY = amylase; Glu = glucose (from cellulose); Ara = Arabinose; Xyl = Xylose; Gal = galactose; Man = mannose.

[a]Components in g/kg of ethanol insoluble residue from fresh matter.

determines the strength of the cell wall and is embedded in an independent pectin matrix, hemicelluloses, and proteins. The backbone is (1→4)β-D-glucopyranose bound to α-D-xylopyranose residues attached to the 6-position of glucose. The apple xyloglucan fraction makes up about 24% of the total amount of sugar. The cellulose–xyloglucan complex accounts for approximatively 57% of the apple cell wall matrix. In fruit liquefaction process, efficient enzymic degradation of this complex is crucial. This involves at least endoglucanase and cellobiohydrolase (13). About 5% of the mass of some primary cell walls is made up of a structural glycoprotein hydroxyprolin-rich *extensin* binding some polysaccharides together. Nature and localization of polysaccharides and macromolecules within the fruit cell wall are summarized in the Table 3. Starch is present in unripe apple in amyloplasts. This is the largest biological molecule, with 10^5 to 10^9 Da molecular weight. Starch is composed of two components: amylose and amylopectin (Fig. 4). Amylose is the minor component of the two and has a linear structure. This is a 1→4-α-D-glucose polymer. Various degrees of polymerization have been ascribed to this fraction, with chain lengths in the range of 100 to 1,000 glucose units. Amylopectin contains both α-1→6 plus α-1→4 glucose linkages. Amylopectin exhibits branching at the 1→6 position, and the degree of polymerization is far greater in amylopectin than in amylose. The ratio of amylose to amylopectin can vary in natural starches in the general range of 1:3 to 1:4.

CELL WALL DEGRADING ENZYMES

Pectinases

Recent progress in enzymology has been so fast that certain activities are not yet described in enzyme classification (Table 4). Many microorganisms produce enzymes that degrade the fruit cell wall. Commercial pectinases for the fruit juice industry come from selected strains of *Aspergillus* sp. Enzymes are purified and concentrated. Pectinases are defined and classified on the basis of their action toward the pectin. Two main groups are distinguished: pectinesterases and pectin depolymerases (14). The pectin lyase (PL) is a pectin depolymerase of endo type that has a great affinity for long methylated chains and acts by β-elimination with the formation of C4-C5 unsaturated oligouronides (15) (Fig. 1**a**). *Aspergillus* produces several endo-PL isoenzymes with an average pH optimum of 5.0 to 5.5. Acting on the same substrate as the PL, the pectinmethylesterase (PME) removes methoxyl groups from methylated pectin, decreasing at the same time the affinity of PL for this substrate. This results in the formation of methanol and lower methylated pectin, but PME has very low effect on pectin viscosity (Fig. 1**a**). PME from *Aspergillus* has a pH optimum around 4.0 to 4.5 and has a strong affinity for high methoxylated pectin such as apple pectin. Fruit PME from orange, for example, acts according to a blockwise mechanism, but *Aspergillus* PME acts according to a multichain mechanism (16). Demethylation with PME and deacetylation with pectinacetylesterase (PAE) reveals free carboxylic acid groups and the pectin becomes negatively charged. Polygalacturonase (PG) exists in two forms:

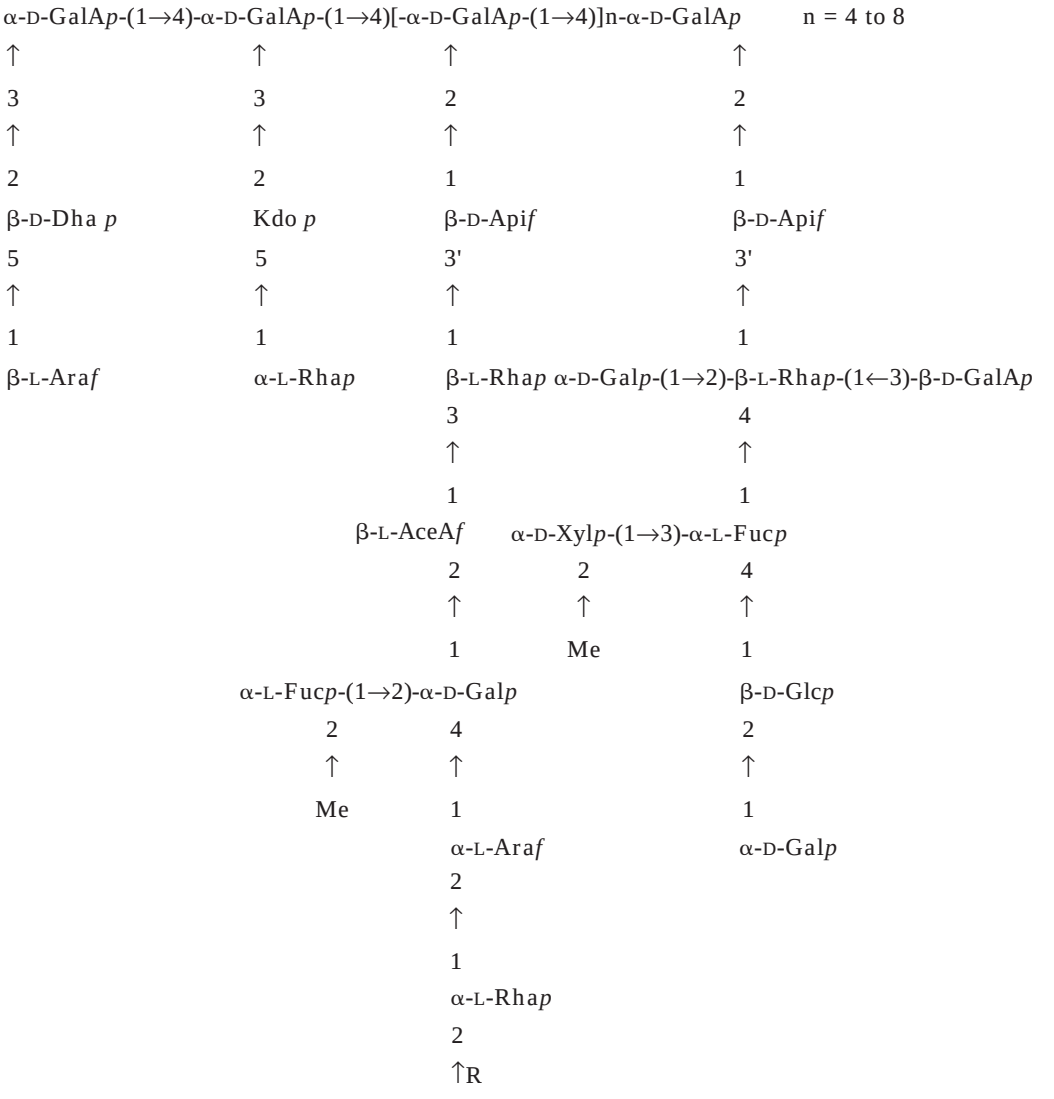


Figure 2. Rhamnogalacturonan type II.

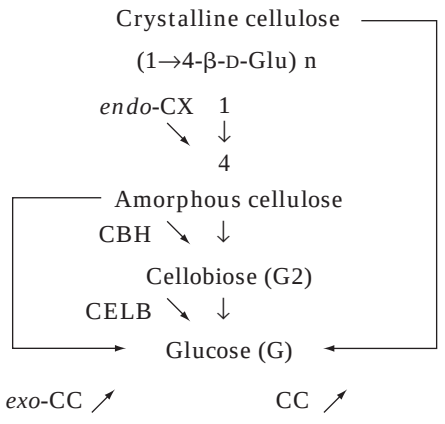


Figure 3. Cellulose structure and related degrading enzymes. CX = β(1→4) *endo*-glucanase; cc = glucohydrolase; CBH = cellobiohydrolase; CELB = cellobiase.

endo-PG and *exo*-PG. Both enzymes are pectin depolymerases and act only on low-esterified pectin (below 50 to 60%) (Fig. 1**a**).

endo-PG acts randomly on the polygalacturonic backbone and *exo*-PG acts on the relevant bond at the nonreducing end of the chain. *exo*-PG releases small fragments from the chain and does not significantly reduce the viscosity. Five *endo*-PG isoenzymes and one *exo*-PG isoenzymes have been described from *A. niger* (17). *endo*-PGs from *Aspergillus* have a pH optimum around 4.0 to 4.5, and *exo*-PG has a pH optimum of 5.0. *endo*-PG rapidly induces pectin depolymerization and reduces its viscosity. Different enzymes active on RG-I were identified and purified from *Aspergillus aculeatus* using apple-pectin-modified hairy regions as substrate (18) (Fig. 1**b**). RGase A was identified as a hydrolase, splitting the α-D-GalA*p*-(1→2)-α-L-Rha*p* linkage of RG-I, while RGase B appeared to be a lyase splitting the α-L-Rha*p*-(1→4)-α-D-GalA*p* linkage by β-elimination. Two novel enzymes were also identified: a rhamnogalacturonan rhamnohydrolase and a rhamnoga-

Table 3. Nature and Localization of Fruit Cell Wall Polysaccharides

Polysaccharides	Class	Middle lamella	Primary cell wall	Secondary cell wall
Cellulose	β1,4-Glucan		+	+
Pectins	Galacturonan	+	+	
	Rhamnogalacturonan	+	+	
	Arabinans	+	+	
	Galactans	+	+	
	Arabinogalactans	+	+	
Hemicelluloses	Xyloglucans		+	
	Arabinoxylans		+	+
	Xylans			+
	Glucomannans			+
	Galactoglucomannans			+
Glycoproteins		+	+	
Lignin		+	+	+

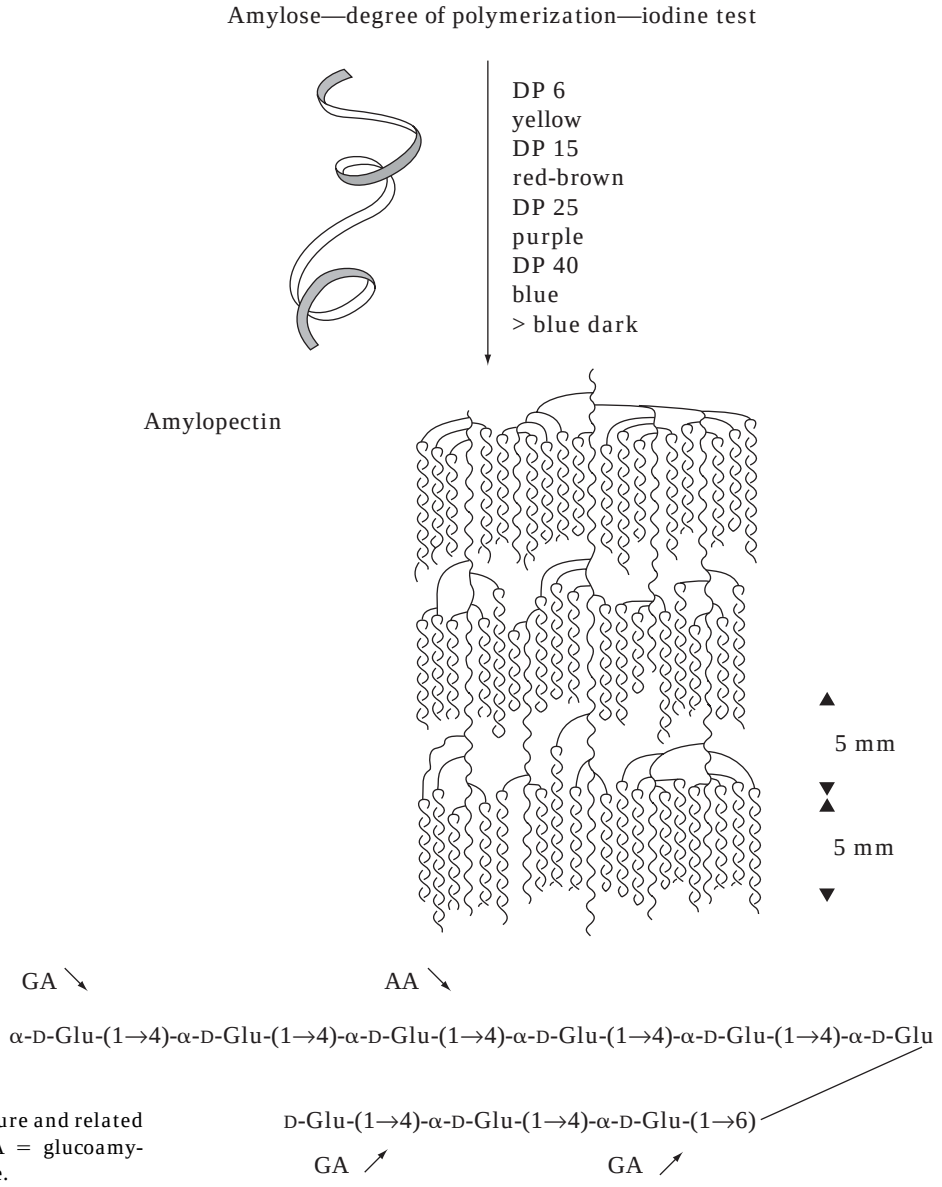


Figure 4. Starch structure and related degrading amylases. GA = glucoamylase; AA = acid amylase.

Table 4. Enzyme Classification

Enzyme	Symbol	EC No.	Substrate	Action mode
Pectin lyase	PL	4.2.2.10	Highly methylated homogalacturonan	*endo*-α1 → 4 β-elmination
Pectinmethylesterase	PME	3.1.1.11	Highly methylated homogalacturonan	Deesterification multichain
Pectinacetylesterase	PAE	3.1.1.6	Acetylated homogalacturonan	Deacetylation
Polygalacturonases	PG	3.2.1.15	Deesterified homogalacturonan	*endo*-α1 → 4 depolymerization
	PG	3.2.1.67	Polygalacturonic acid	*exo*-α-1 → 4
Rhamnogalacturonases	RHG A		Rhamnogalacturonan	Gal α-1 → 2 Rha hydrolysis
	RHG B		Rhamnogalacturonan	Rha α-1 → 4 Gal β-elmination
			Rhamnogalacturonan	*exo*-Rhan-Rha
			Rhamnogalacturonan	*exo*-Rhan-GalA
	RGAE		Rhamnogalacturonan	Deacetylation
Arabinosidase	ARA	3.2.1.99	Pectin araban	*endo*-α 1 → 5
	ARF A	3.2.1.55	Pectin side chains	*exo*-α 1 → 2, 1 → 3
	ARF B	3.2.1.55	Pectin side chains	*exo*-α 1 →, 1 → 5
Galactanase	GAL	3.2.1.89	Pectin side chains	*endo*-β1 → 4
Galactosidase	Gal	3.2.1.23	Pectin side chains	*exp*
Arabinogalactanase	AGal	3.2.1.90	RG side chains	*endo*-β1 → 3
	AGal	3.2.1.90	RG side chains	*endo*-β1 → 6
Xylanase	XYL	3.2.1.8	Xyloglucan	*endo*-β1 → 4
Xylosidase	Xyl	3.2.1.37	Xyloglucan	*exp*
Amylase	AMY	3.2.1.1	Amylose	*endo*-α 1 → 4
Amyloglucosidase	AG	3.2.1.3	Amylose amylopectin	Nonreducing end *exo*-α 1 → 4, 1 → 6
Mannanase	MAN	3.2.1.78	Mannan	*endo* β1 → 4
Galactomannanase	GMan		Galactomannan	*endo*
Cellulase	CC	3.2.1.4	Amorphous cellulose	*endo*-β1 → 4
	CX	3.2.1.74	Crystalline cellulose	*exo*-β1 → 4
Cellobiohydrolase	CBH	3.2.1.91	Crystalline cellulose	Dimer
Cellobiase	CELB	3.2.1.21	Cellobiose	*exo*

lacturonan galacturonohydrolase (Fig. 1**b**). As an accessory enzyme for the RGases, rhamnogalacturonan acetyl esterase (RGAE) was discovered in the same preparation. Although a hypothetical structure of RG-II has been described, RG-II degrading enzymes are still to be discovered. Arabanases are pectinases because they remove arabinose covalently bound to the homogalacturonan backbone. Three enzymes have been described: one *endo*-arabinanase (α-1→5) (*endo*-ARA) and two arabinofuranosidases, among which one *exo*-arabinofuranosidase A (α-1→2; α-1→3) (ARFA) and one *exo*-arabinofuranosidase B (α-1→3; α-1→5) (ARFB) (Fig. 1**c**). All three are produced by *Aspergillus* sp. (19). Activity level and ratio must vary in enzyme preparations, depending on the fruit species processed (i.e., high activities are required for apple and pear processing). Highly branched arabinans are not a substrate for *endo*-ARA. However, when *exo*-ARFA activity is too high, the linearization of arabinan chains leads to a reassociation of these chains at low water activity, forming a haze in apple and pear concentrates (20). *A. niger* ARAs have pH optimum at 4.0.

Hemicellulases

Hemicellulases are enzymes hydrolyzing arabinogalactans, galactans, xyloglucans, and xylans. Arabinanases are active on arabinans, and also on arabinogalactans or arabinoxylans. *Aspergillus* produces enzymes able to hydrolyze arabino (1→4)-β-D-galactans type I (Fig. 1**d**) and arabino (1→3)-(1→6)-β-D-galactans type II (Fig. 1**e**) (9,21,22). The *exo*-(1→3)-β-D-galactanase is able to release galactose and (1→6)-β-D-galactobiose. *Aspergillus* enzyme is able to bypass a branch point in a β-(1→3) backbone. The action of this enzyme is enhanced by the presence of ARAB. Xylanases hydrolyze the (1→4)-β-D-xylans in synergy with ARAs, glucuronidase, and acetylesterase (23). They are important enzymes for pear cell wall breakdown. Other hemicellulases such as mannanases hydrolyze the β-(1→4) mannans, galactomannans, and glucomannans present in coconut or pineapple.

Cellulases

Commercial cellulases come from selected strains of *Trichoderma* sp. Beldman has widely described cellulases from *Trichoderma* (24). Cellulase is a multienzyme system composed of several enzymes with numerous isoenzymes that act in synergy (Fig. 3). Nevertheless, products of reaction can retroinhibit the enzymes (i.e., glucose and cellobiose toward *exo*-glucanases and cellobiohydrolases).

endo-Glucanase isoenzymes (CX) hydrolyze the (1→4)-β-D-glucan chain of cellulose at random (Fig. 3). Vincken et al. in 1994 demonstrated that removal of the xyloglucan coating from cellulose of isolated apple cell walls by a xyloglucanase (XGase) purified from *Trichoderma viride* or *endo*-IV glucanase (CX) enhanced the enzymic degradation of cell wall embedded cellulose (5). XGase can work in synergy with pectic enzymes such as PL and is a key activity for the enzymatic liquefaction of apple (25). *exo*-Cellulase (CC) hydrolyzes the β-D-(1→4) glucan chain of cellulose at its reducing extremity with an *exo* behavior. After CX action, cellobiohydrolases (CBHs) hydrolyze bonds at the

nonreducing end of the crystalline cellulosic chain, producing cellobiose. Cellobiase (CELB) or β-glucosidase splits the cellobiose to glucose. Many models of cellulose hydrolysis have been proposed (26). Native crystalline cellulose is opened by *endo*-glucanase, which acts on regions of low crystallinity to form amorphous cellulose. *endo*-CX plus CBH can degrade amorphous cellulose to cellobiose easily. This cellulolytic complex is of first importance in every process of fruit liquefaction.

Amylases

In the fruit juice industry, fungal acid amylases are used to process fruits that contain starch. This is the case of unripe apples just after the crop period. *A. niger* produces acid α-amylase and amyloglucosidase (27). These exogenous enzymes are applied after starch gelatinization (occurring above 75 °C (167 °F) for preventing postbottling haze formation. Acid *endo*-amylase (AA) is active on (1→4)-α-D-glucose chain or amylose (Fig. 4). AA from *A. niger* has an optimal pH of 3.0 to 4.0, and the optimal temperature is 70 to 75 °C (158 to 167 °F).

After action of AA that gives dextrins, glucoamylase or AG (which is an *exo*-hydrolase) acts on gelatinized starch by releasing glucose from the nonreducing end of the α-(1→4) bonds of amylose chains but can also hydrolyze α-(1,6) bonds of amylopectin chains (Table 4). AG from *A. niger* has an optimal pH of 4.0 to 4.5, and the optimal temperature is 55 to 60 °C (131 to 140 °F).

INDUSTRIAL ENZYMES

Standardization of Enzyme Production

Commercial enzymes are produced and sold by several companies, such as Gist-Brocades, Novo Nordisk, Solvay, Röhm, Quest, and Danisco. The enzymes most frequently offered to fruit juice producers are pectinases coming from *Aspergillus* species, the most common from *A. niger*. Next to pectinases, numerous side-activities-type hemicellulases or glycosidases are secreted, all working in synergy with the pectinases on the cell wall substrate. The spectrum of enzymes and their activities are natural to the strain (constitutive) and to the culture conditions (inducible). This is why the enzyme spectrum differs from one commercial product to another. Pectinases are exocellular enzymes and are the main activities produced on carbohydrate substrates. These enzymes are concentrated and sold mainly as liquids, but also as powders or granulates. In the past few years, some juices producers have chosen the new technology of fruit liquefaction. In addition to pectinases, hemicellulases and cellulases are needed for complete cell wall degradation. These enzymes are produced from *Aspergillus* or *Trichoderma* spp. It is possible to blend cell wall degrading enzymes coming from different cultures optimally, depending on the application and the fruit composition (Table 5).

Various microorganisms are used by enzyme producers as enzyme source. All the microorganisms used to produce food enzymes bear the GRAS status (generally recognized as safe). Enzymes are produced under extremely well defined and reproducible conditions (Gist-Brocades has ISO 9002 certification and applies HACCP). Quality assurance and standardization of industrial enzymes are strictly defined concerning the enzyme activity, its concentration, the presence of chemical and microbiological contaminents (heavy metals, toxins, yeasts, bacteria, molds) depending on legislature requirements. Enzymes are proteins considered as processing aids; they are inactivated in fruit processes and are not present or active in the final product. Good manufacturing practices (GMP) in the microbial food enzyme industry ensures that enzyme preparations are produced, packed, and handled in a hygienic way. The key issues in GMP are the microbiological control of the microorganism selected for enzyme production, the control and monitoring systems ensuring pure culture and optimum enzyme productivity conditions during fermentation, and the control of the hygienic conditions throughout recovery and finishing of the enzyme preparations. GMP will ensure that contamination of the product during recovery is minimized. The final product has to meet the criteria formulated by the WHO/FAO Joint Experts Committee on Food Additives (JECFA) and the Food Chemicals Codex (FCC). Gist-Brocades and other enzyme suppliers are members of AMFEP (Association of Manufacturers of Fermentation Enzyme Products). AMFEP makes sure that regulatory aspects of enzyme applications in the food industry are respected by its members.

Genetically Modified Organisms

Enzymes produced by recombinant DNA (rDNA) technology will increasingly find application in fruit juice production. Application of so-called rDNA enzymes has been slow for a number of reasons, mainly because of the technical complexity of the processes involved, but it can be expected to increase because knowledge about enzyme mixtures is increasing rapidly. The application of rDNA technology to produce enzymes for fruit juices involves the following steps: (*1*) The genes coding for these enzymes have to be cloned. This is not a small task, because many enzymes are involved and for many enzymatic activities more than one gene can be found in the traditional strains that are being used. Today, more than 25 genes encoding so-called plant cell wall degrading enzymes that are important for this process are known. (*2*) The genes have to be overexpressed in a suitable safe host strain. This strain is preferably belonging to the same genus as the production strain, for instance *A. niger*. The technology for introducing these genes in production strains resulting in overexpression of the relevant enzymes are well established now. It results in a genetically modified organism (GMO) that has a small change in its genetic makeup, thereby overproducing one of the enzymes of interest, for instance, a pectinlyase. If no DNA of other species is introduced in this process, it is called homologous recombination, which can be compared to the classical genetic techniques. If foreign DNA is introduced in the receptor strain as well, the process is called heterologous recombination. (*3*) The enzyme produced by the GMO has to be tested in application studies. This is, of course, a very critical process because a balanced mixture of enzymes is required to obtain a juice of the right quality and yield.

Table 5. Fruit, Enzymes, Process

Fruit/process	Juice extraction	Juice depectinization concentration	Fruit liquefaction
Apple	PL or PME + PGs + RGs + ARFs + ARA	PL + PME + PGs + RGs + ARFs + ARA + AGals + AMY + AG	PL + PME + PGs + RGs + ARFs + ARA + CX + C1 + CBH + CELB
Pear	PL or PME + PGs + RGs + ARFs + ARA	PL + PME + PGs + RGs + ARFs + ARA + AGals + AMY + AG	PL + PME + PGS + RGS + ARFs + ARA + XYL + CX + C1 + CBH + CELB
Red Berry	PME + PGs + ARA + ARFs + AGals + CX	PME + PGs + ARAs + AGals + XYL	
Grape	PG + ARFs + ARA + AGals	PL + PME + PGs + RGs + ARFs + ARA + AGals	
Pineapple		PG + GMan + AGals + XYL	
Stone fruit			PL + PME + PGS + RGS + ARFS + ARA + XYL + CX + C1 + CBH + CELB
Citrus		PL	

Note: PL = pectin lyase; PME = pectinmethylesterase; PG = polygalacturonase; RG = rhamnogalacturonase; ARF = arabinofuranosidase; ARA = arabanase; CX = endoglucanase; C1 = exoglucanase; CBH = cellobiohydrolase; CELB = cellobiase; XYL = xylanase; GMan = galactomannanase; AGal = arabinogalactanase; AMY = amylase; AG = amyloglucosidase.

The genes that have been cloned from *Aspergillus* species that are currently relevant in this respect include enzymes that degrade pectin (*endo*- and *exo*-polygalacturonase, pectinlyase, pectinmethylesterase, rhamnogalacturonase). Other examples are arabinan-degrading enzymes (*endo* and *exo*), cellulases (*endo*-glucanases and β-glucosidases) and galactanases. Starch-degrading enzymes and proteases are relevant as well. Despite the impressive list of known enzymes, a number of others, for instance those degrading the so-called hairy regions of pectin, still have to be identified and their genes cloned. But, in theory it will be possible in the future to make optimal mixtures of enzymes produced by GMOs for every application (Table 5). Fruit juice enzymes are usually produced from *Aspergillus* and *Trichoderma* strains. The GMO technology guarantees that enzyme proteins produced from GMOs are identical to those produced from non-GMOs and allows the production of single pure enzyme activity. GMOs are absent from final products. Various GMO-derived food enzymes have already been approved and used for more than 15 years in Europe and the United States in the starch, dairy, and beverages industries. In summary, it can be predicted that enzymes produced by GMOs will start a new era in fruit juice manufacture, provided that it will be possible to further unravel the complex mixtures that are used today. Another important prerequisite is that the products will present an advantage for both the fruit juice manufacturer and the consumer.

ENZYMES AND APPLE JUICE PROCESSES

Apple is by far the most important raw material worldwide for the production of clear juice and clear concentrate. Annual production is around 54 million tons of apples, from which about 20% are processed mainly as clear concentrate, but a small part is processed as cloudy juice and apple sauce. The major producers of apple juice concentrate are located in the United States, Poland, Argentina, Italy, Chile, China, and Germany. In the traditional process, enzymes are used at two different stages. Pectinase preparations containing numerous side activities are added to the fruit pulp for a maceration during which soluble pectin and cell walls are partially hydrolyzed, resulting in a dramatic decrease of the viscosity, increasing the yield because the free-run juice volume is larger and the pressing easier. Mash enzyming increases also the productivity of the plant. Pectinases with amylases are added to the juice after pressing and juice pasteurization. These enzymes decrease the juice viscosity, increase its filterability, making the clarification and the concentration easier. Both enzyme treatments allow production of clear and stable concentrate. At the same time, enzyme treatments prevent postbottling haze formation (starch haze or araban haze) in concentrates. A few producers in Germany and the United States make natural or cloudy apple juice. Until now, they have not added enzymes because the existing commercial pectinases induce a fast clarification of the juice. Gist-Brocades sells a purified pectinmethylesterase without pectin depolymerase activity. It can be used in the cloudy juice process and also in the French cider process with flotation. Finally, the new technology of total liquefaction, achieved by using pectinases, hemicellulases, and cellulases in blend, is performed with only one enzyme addition. Solids are removed from the liquefied pulp with decanter or press. The yield is maximal.

Clear Apple Juice Concentrate

Process Description. The traditional process (Fig. 5) (28) consists of the following stages: apples are selected, washed, and crushed. Macerating enzymes are added continuously during crushing after their dilution in water through a dosing pump. The enzymed pulp is then pumped into the maceration tank where the incubation takes place during 30 to 60 minutes at ambient temperature. Pectinases such as Pectinex UltraSP (Novo Nordisk) or Rapidase® Press (Gist-Brocades) are added at 40 to 75 g/

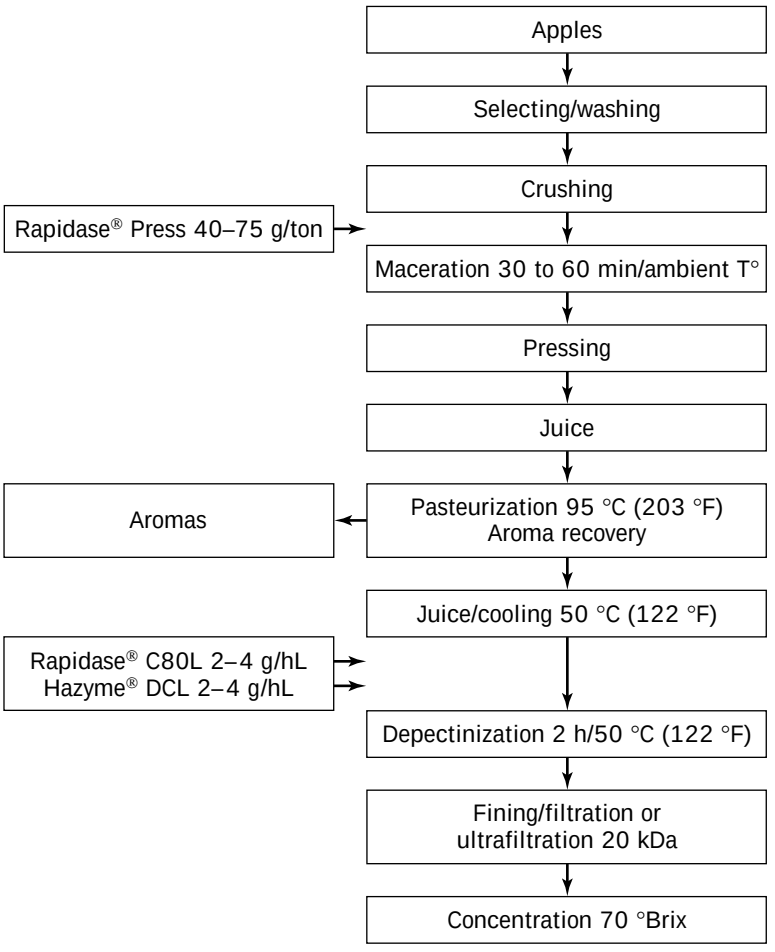


Figure 5. Clear apple juice concentrate process.

ton of apples depending on fruit ripeness. In the apple pulp, macerating pectinases from *Aspergillus* are stable at about 75% of their optimal activity at pH 3.5 to 4.0. After this enzymatic maceration, the mash is pumped to the pressing equipment. After pressing, the juice is flash pasteurized for aroma recovery and cooled down. Pectinases are added to the juice at 45 to 50 °C (113 to 122 °F) at 2 to 4 g/hL. Pectinases Pectinex 3XL (Novo Nordisk) or Rapidase® C80L (Gist-Brocades) and amylases AMG (Novo Nordisk) or Hazyme®, DCL (Gist-Brocades) are added at the same dosage at the beginning of the season when apples contain starch. After 2 h, the juice no longer contains any pectin or starch. In recent years, the development of tangential ultrafiltration with membranes of 20 to 30 kDa cutoff has enabled the fast clarification of the juice by removal of insoluble particles and macromolecules. The brilliant and clear ultrafiltered juice is then concentrated at 72 ° Brix and stored in the dark at low temperature.

Apple Pulp Maceration

Fruit juice producers look for an optimization of the yield and of the plant capacity to make more profit from raw material. Press efficacy is very much higher after pulp enzyming. The application of exogenous pectinases from *Aspergillus* sp. on apple mash is a necessity because endogenous enzymes present in apple such as pectinmethylesterase and polygalacturonase have too low activity to cause immediate visible effect (29,30). Because apple pectin is approximately 90% methylated, commercial enzyme preparation must contain a high level of pectinylase or pectinmethylesterase in association with polygalacturonase and arabanases (Table 5). In 1990, Schols et al. described a new pectinase rhamnogalacturonase produced by *Aspergillus* sp. and its important role on the maceration of apple tissue (7). In 1996, new enzymes are active toward pectic structures, among which different rhamnogalacturonases have been described (18). Rapidase® Press (Gist-Brocades), Pectinex UltraSP (Novo Nordisk) or Rohapect MA plus (Röhm), sold as, mash enzyme preparations, contain activities that are necessary to obtain high yields throughout the entire season. Commercial products contain more or less these enzyme activities but with a different ratio. About 70% of the producers use horizontal presses; 15%, belt presses; and others, decanter or screw presses. After pulp enzyming, extraction of the juice can increase the yield up to 90% with hydraulic press (with pomace leaching), compared to a maximum 80% yield without enzyme treatment. Apple varieties processed just at the crop period (fresh apples) are easily pressed with a relatively high yield. The current trend is to process apple juice from table varieties (golden delicious, granny smith, jonagold, red delicious) with quality defects. These apples are rich in pectin and poor in tannins in contrast to old apple varieties that

have almost disappeared (cider-type apples). These apple varieties are often stored at low temperature for several months in controlled atmosphere (stored apples), and overripe apples are processed in juice throughout the year. During storage, bonds between different polysaccharide fractions such as pectin, hemicelluloses, and cellulose are modified. The insoluble protopectin is slowly transformed into soluble pectin through the endogenous pectinases of the apple, and starch is degraded by endogenous amylases (Fig. 6). In the apple, the elasticity and the cohesion of the cells decrease with time. The water in the fruit moves to the crystalline cellulose in the periphery, linked to lignin in a more organized and thick structure (31). Moreover, as apples become overripe, especially after being stored for a long time at a low temperature and in a controlled atmosphere, they give a decreased juice yield (Fig. 6). The mash pressability is bad, and the pomace sticks to the press cloth. Apples are impossible to press without enzymatic maceration because of the high soluble pectin content. This is why the dosage of enzymes, added to decrease the viscosity of the pulp, must be higher at the end of the season.

Apple Juice Depectinization

Pectin is the main cause of turbidity. One liter of juice with a dry matter of 13% can contain 2 to 5 g of pectin after pressing, depending on fruit ripeness. After aroma recovery and pasteurization, pectinases and amylases (only at the beginning of the processing season) are added simultaneously to the juice during 1 to 2 h at 45 to 60 °C (113 to 140 °F). *Aspergillus* pectinases are stable at 50 to 55 °C (122 to 131 °F) in apple juice at pH 3.0 to 4.5, and some of them are stable until 60 °C (140 °F). Pectinase dosage is determined at the laboratory by performing the alcohol test. PL, or pectinmethylesterase plus polygalacturonases, and arabanases are the most important activities.

Arabinose represents 55% of neutral sugars in apple pectin hairy regions. Arabanases prevent the haze formation after the juice concentration. The juice clarification is actually a secondary effect of the enzymatic depectinization (32). It is caused by enzymatic and physicochemical mechanisms. The first stage consists of the destabilization of the cloud first because of PL, without visible effect but that results in a strong decrease of the viscosity of the juice (Fig. 7). The PL cuts the pectin at random, and cutting 1 to 2% linkages is enough to reduce apple juice viscosity by 50%. PL affinity for apple pectin is high because of the high methylation degree. PGs become active only after the action of PME because they cannot depolymerize the pectin when the methylation degree is higher than 50 to 60% and have a greater affinity for the active site after its partial demethylation (steric hindrance). At this stage, the PL is no longer active and the pectin hydrolysis results from the system PME and PGs. The second stage is flocculation and sedimentation. The sedimentation necessary to lead to the clarification of an apple juice occurs only after these enzymatic degradations of pectin and starch. The cloud is composed of proteins with positive charges when the pH of the juice is 3.5 to 4.0 because their isoelectric point is in-between 4.0 and 5.0. These proteins are bound to hemicelluloses surrounded by the pectin as a protective, negatively charged colloid layer. Pectinases partially hydrolyse the pectin gel, resulting in the aggregation of oppositely charged particles, such as tannins, and in the flocculation of the cloud and then clarification of the juice. The optimal pH for this mechanism is 3.6. The clarification occurs after the solubilization of the insoluble protopectin by pulp enzymes, the decrease of the juice's viscosity (in fact, in a first stage, the viscosity increases first because of the progressive solubilization on the part of protopectin and then decreases after pectin hydrolysis), and the partial hydrolysis of the colloid protector.

The sedimentation occurs quickly during the neutralization electrostatic phase inducing the clarification. In the past, this step occurred after a very long time and the juice was sensitive to contamination. The use of pectinases saves time. The alcohol test shows if residual pectin is still present in the juice. We recommend using acidified ethanol (2 vol of ethanol containing 0.5% HCl + 1 vol of apple juice) to precipitate the residual pectin with a molecular weight equal or lower than 3,000 Da. It is important to use acidified ethanol; if not, the precipitation of organic acids or calcium pectate occurs because of the modification of the pH after mixing of alcohol with the juice, giving a false-positive answer. At the beginning of the processing season, unripe apples contain starch, 5 to 7 g/L of juice. At this time, the iodine test will give a blue dark color, and the starch will precipitate with iodine. Starch is in a granular

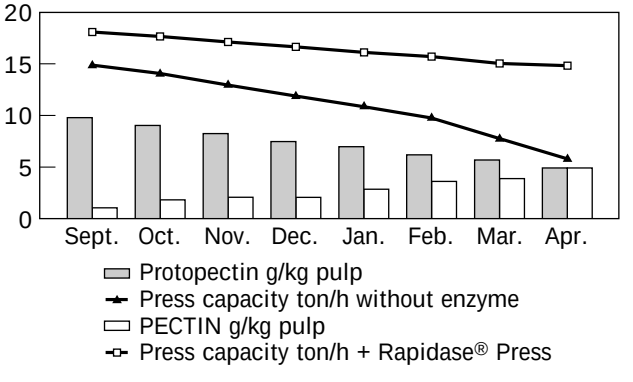


Figure 6. Apple ripeness, protopectin, pectin, and pressability (in northern hemisphere).

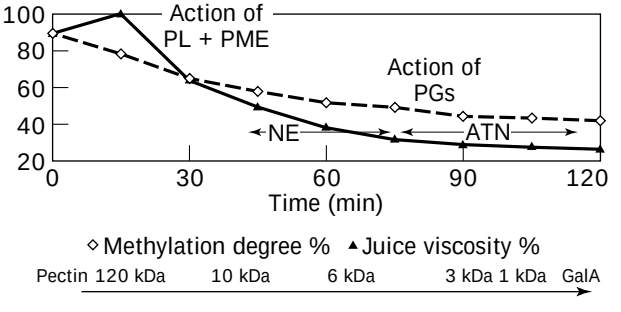


Figure 7. Apple juice depectinization. NE = neutralization of electrostatic charges; ATN = alcohol test negative.

form, ranging from 2 to 13 μm and is composed of 30% amylose and 70% amylopectin chains. Amylopectin can fix 20% its weight in iodine (Fig. 4). Apples contain endogenous amylases; however, the process is too short and activities are too low to degrade the starch. When the juice is heated up to 75 to 80 °C (67 to 176 °F), starch moves from an insoluble form to a soluble form (gelatinization with water). If no amylase is added, these molecules can retrograde later and make big aggregates that are difficult to hydrolyze and clott the filters. The addition of fungal amylase results in the hydrolysis of soluble starch until glucose monomer. Retrogradation and postbottling haze formation are prevented. Clarification and filterability are improved. Preparations such as Hazyme® DCL (Gist-Brocades) or AMG (Novo Nordisk) contain α-amylase and amyloglucosidase activities from *A. niger*. The dosage is determined by the iodine test.

Amylases are added during the depectinization stage at 2 to 4 g/hL simultaneously with pectinases, for 2 h at 50 °C (122 °F). At the end of the hydrolysis, the iodine test and alcohol tests are negative. The third stage is the clarification with fining agents (gelatin and bentonite), filtration, or ultrafiltration (UF). UF membranes made of inorganic substances have an average cutoff of 20,000 Da. The juice goes through the membrane as filtrate. Insoluble particles and soluble colloidal material of molecular weight higher than the pore size (nondegraded part of RG-I, RG-II, proteins such as enzymes, polyphenols) are removed as retentate. Recently, enzymes active toward the rhamnogalacturonan part of pectin have gained attention, because it was found that in juice processing, rhamnogalacturonan parts can foul UF membranes. For this reason, the UF stage was sometimes a bottleneck in the process, especially when the extraction of cell wall components was too strong. We know that a more complete pectin degradation is required to increase the UF flux rate. This is why enzyme preparations proposed for apple juice depectinization must contain a broad range of pectinases, including rhamnogalacturonases, hemicellulases, and amylases (Table 5). With the all these enzymes, the juice clarification is optimal, and its concentration up to 70 °Brix is easy, without risk of gel or haze formation.

Apple Total Liquefaction

The enzymatic liquefaction of apple consists of the hydrolysis of cell wall polymers by exogenous enzymes into soluble components, mainly acidic and neutral sugars, thus increasing the extraction level and the yield. In relation to the cell wall composition and in comparison with maceration, more activities are needed for cell wall liquefaction (33).

The apple cell wall can be solubilized more than 80% by the combined action of pectinases, cellulases, and hemicellulases (Table 5) (34). Recently, new enzymes such as rhamnogalacturonases and xyloglucanases have been found to be very important for apple liquefaction yield and UF flux rate. They also strongly decrease the viscosity of the mash, and solids are discarded from the juice by centrifugation more easily. The new trend is a process including a hot break stage.

Process Description. Apples are selected, washed, and crushed, and then the pulp is heated up to 85 °C (185 °F) for 4 to 5 min (pectinases work more easily after the physical degradation of pulp cells) (Fig. 8). During cooling, the pulp goes to a finisher for seed peel and core removal. The liquefaction is then carried out for 2 to 3 h at 50 °C (122 °F) with a continuous slow stirring at 40 to 60 rpm. During the tank filling, commercial pectinases such as AFPL (Novo Nordisk), Grindamyl (Danisco) or Rapidase® LIQ EXTRA (Gist-Brocades) are added, from 100 to 250 g/ton of pulp. In this process, no cellulase is needed, and there is no cellulose degradation to cellobiose and glucose in the juice. The yield is still very high. Enzymes are stable in the process conditions, and pectolytic and hemicellulolytic activities are high enough to depectinize the juice during the liquefaction time. After about 3 h, the liquefied pulp, which contains 25 to 30% solids (in volume), is easily centrifuged with a decanter, giving a first juice. The pomace is mixed with hot water and pumped to a second decanter or press, giving a second juice. Both juices are mixed and contain less than 1.5% solids. After pasteurization, active carbon and gelatin are often added to decolorize the juice and to improve UF flux rate. Fining agents are removed by vacuum filtration. Then the juice is ultrafiltered and concentrated up to 70 °Brix. The total liquefaction process results in a high yield increase, with an average of 95 to 96% along the whole processing season and according to the process described in Figure 8. The yield is calculated as follows:

$$[(\text{juice n°1 weight} \times \text{°brix}) + (\text{juice n°2 weight} \times \text{°Brix})] \div [(\text{apples weight} \times \text{°Brix of apple})]$$

The titrable acidity of the final juice is high because of the galacturonic acid content. Liquefied apple juice can contain methanol, up to 300 to 500 mg/L of juice. However, its major part is removed during the concentration stage. The yield after apple total liquefaction process is compared with the traditional process in Table 6. The quantity of residual waste is very low. Production cost is decreased (higher yield, cheaper equipment, less labor in a continuous process). This process could be applied to extract the juice from fruits such as mango or guava for which no equipment has been developed.

Apple Pomace Liquefaction

Process Description. Gist-Brocades is one of the first enzyme suppliers that has proposed enzymes dedicated to apple pomace liquefaction (35) (Fig. 9). Depending on the equipment, the first extraction yield can be adjusted precisely by using macerating pectinases, and we recommend 65 to 70% yield. After addition of water (Fig. 10), the pomace is liquefied with enzymes. Rapidase® Pomaliq Extra is used at 150 to 250 g/ton calculated according to the weight of apples at the start of the process. The pomace is stirred at 40 to 60 rpm for about 3 h at 50 °C (122 °F). The liquefied pomace is pumped to the press or the decanter for obtaining a second juice. During the pomace liquefaction, the Brix degree of the juice increases from 6.0 to about 7.5, mainly because of the degradation of the pectin.

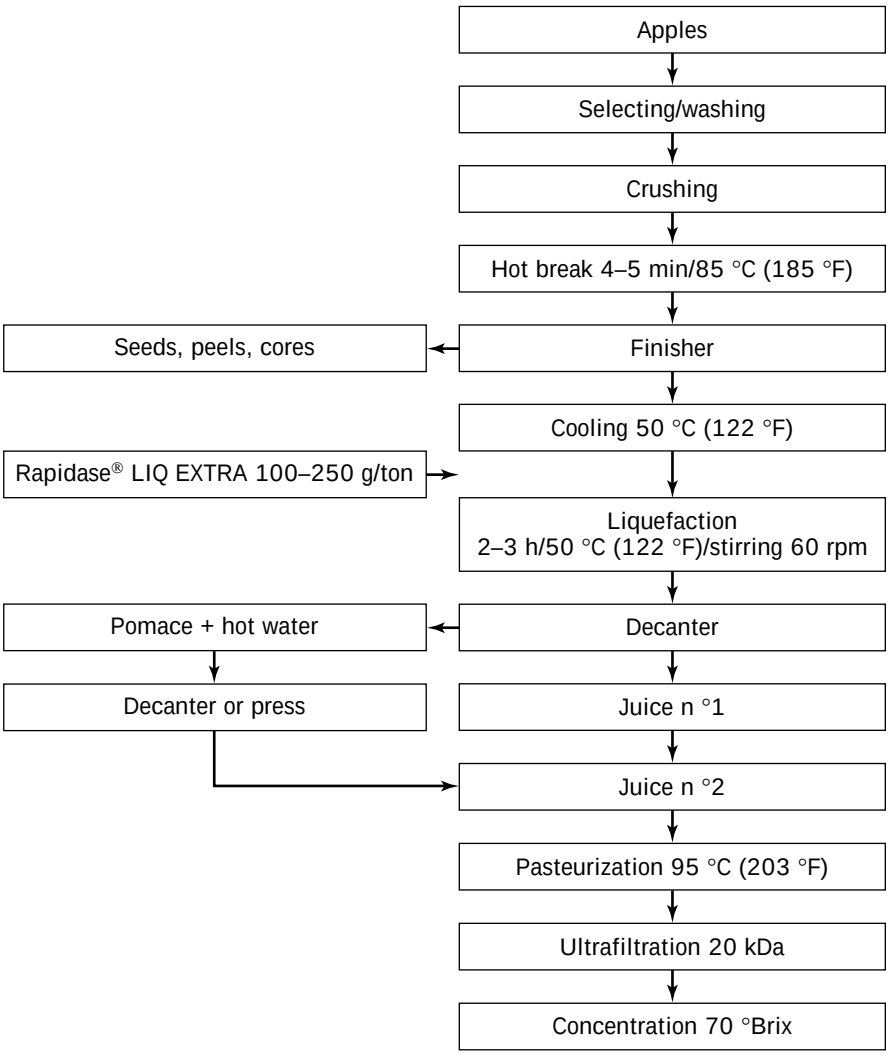


Figure 8. Apple total liquefaction process.

This juice can be blended to the first pressing juice that has been depectinized in the case of clear juice production. The second juice, which is very acidic, can be used separately for other applications (to increase the acidity, cider, vinegar). Further stages are pasteurization, UF, and concentration. In this process, the quantity of waste pomace is dramatically reduced (around 60 kg pomace per ton of apples using horizontal press). The final yield is around 97 to 98% (Table 6). This technology is very flexible because the producer can produce cloudy juice, clear juice, or clear concentrate, gaining greater value from the raw material. Pomace liquefaction can be performed with the existing pressing equipment (only a heating system and a tank with a stirrer are needed). This results in a dramatic increase of the output of the plant with an important additional weight of concentrate.

French Apple Cider

French cider is produced from cider varieties of apples, bitter, bittersweet, sweet, or sour, with high tannin content (Bedan, Bisquet, Domaine, Douce Coetligné). In the traditional French process, the fermentation of the apple juice is very slow (a few weeks to 3 months). According to the French government, residual sugar content depends on the type of cider (brut, $<$28 g/L; demi-sec, 28 to 42 g/L; doux, $>$35 g/L with less than 3% ethanol). Because of the variability of raw material and the lack of hygiene, quality was difficult to manage. The use of continuous pressing and the diffusion system, the improvement of filtration equipment, and the control of bacteria and yeast strains resulted in improving the quality of French cider (36). In 1909, Warcollier described cider production, and the technology is almost the same today. The major part of French cider is produced by the technology of juice "defecation."

Traditional French Apple Cider Process. The existing process can be divided in three stages (Fig. 11). The first stage consists of the preparation of the apple must. Apples are crushed and pressed, and then a second extraction is performed by diffusion (countercurrent of cold water in the pomace) or second pressing. The second stage is the traditional clarification by "defecation." It consists of the pectin demethylation by the apple endogenous PME, giving pectic acid, which binds to apple calcium ions and becomes

Table 6. Yield and Technology

For 1 ton apples at 12.5 °C Brix	Pressing depectinization	Maceration depectinization	Total liquefactor	Pomace liquefaction
Enzyme	No Rapidase® C8OL 0	Rapidase® Press Rapidase® C80L 50 g/ton	Rapidase® LIQ EXTRA	Rapidase® POMALIQ EXTRA
Dose	3 g/hL	3 g/hL	150 g/ton apples	150 g/ton apples
Juice (kg)	750	850	860	750 + 455
Brix °	12.5	12.5	13.5	12.5 + 7.0
Concentrate (kg) 70 ° brix	134.0	151.8	165.9	179.5
Increase	–	+17.8 kg	+3.1 kg	+45.5 kg

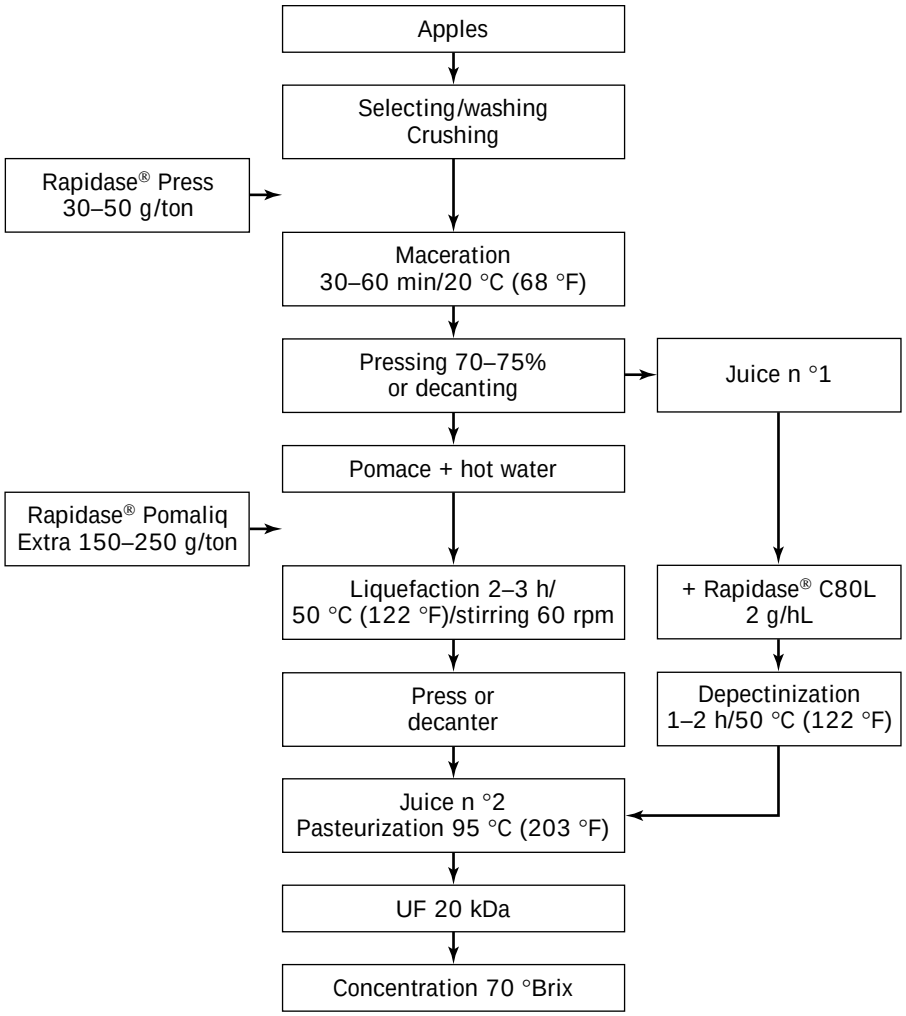


Figure 9. Apple pomace liquefaction process.

insoluble in the form of calcium pectate. The so-called coagulum, which entraps pectate, proteins, and tannins, rises to the surface because of CO_2 bubbles formed by the alcoholic fermentation that has started by this time. This coagulum is then discarded. This type of clarification lowers the nitrogen content, resulting in a slower fermentation that gives a more aromatic cider. This process, which was unreliable because of the variability of endogenous PME activity present in the juice, is now better controlled by the use of commercial fungal PME. Gist-Brocades and INRA Le Rheu have developed a new process of defecation using a purified pectin methylesterase Rapidase® CPE from *A. niger* that enables the cider producer to control and to accelerate the mechanism of coagulum formation. Rapidase® CPE is added at 7 to 10 mL/hL, plus anhydrous calcium chloride, 44 g/hL. A gel of insoluble calcium pectate is formed after about 36 h at 15 °C (59 °F) or a few days at low temperature that rises to the top of the tank on the surface of the must. This coagulum is then discarded. Before the fermentation, sulphur dioxide (50 mg/

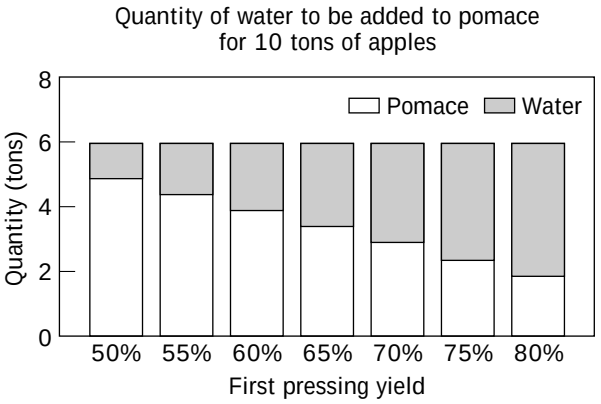


Figure 10. Water addition for apple pomace liquefaction.

L) can be added to the must to inhibit undesired microorganisms. The alcoholic fermentation takes place over 1 to 2 weeks at cellar temperature or up to 3 months in cold storage conditions.

The third stage consists of post-treatments after alcoholic and possibly malolactic fermentation. The cider is centrifuged to discard yeasts at a gravity selected according to the type of cider. It is filtered, and carbon dioxide (standard is above 5 g/L) is added just before bottling. Most of the ciders are pasteurized after bottling to ensure microbiological stability, but traditional ciders are not. In conclusion, the use of a purified PME like Rapidase® CPE (Gist-Brocades) results in producing French cider of high and constant quality whatever the raw material and process conditions. Recently, a new enzyme has been introduced in the cider industry that can be used as a macerating enzyme, improving the juice yield and at the same time making the clarification with flotation possible. This is a highly purified PME without side activity. Because of its high purity, this enzyme does not hydrolyze apple pectin, but the pectin can be extracted from the dry pomace after cider processing by pectin manufacturers.

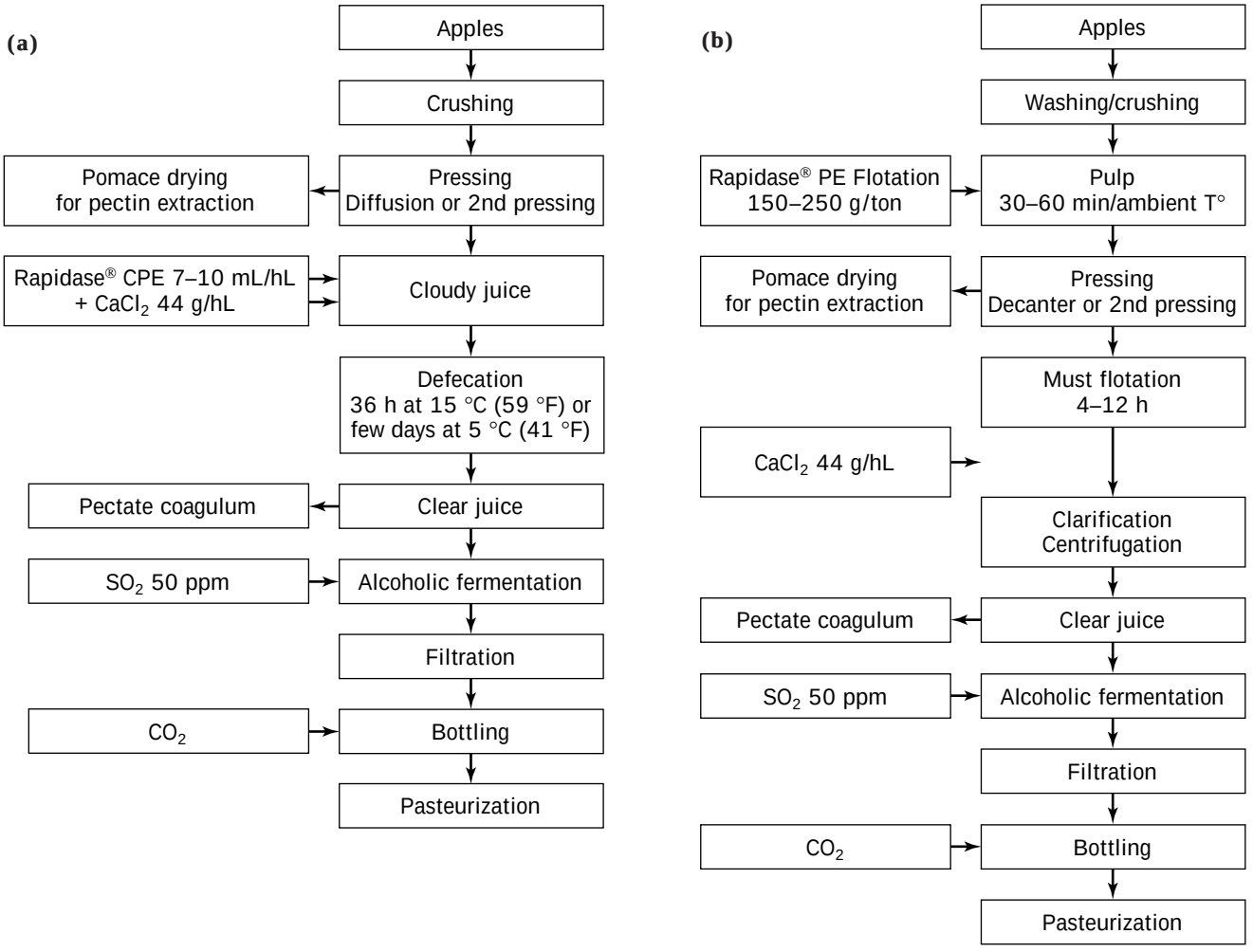


Figure 11. (**a**) Traditional French apple cider process. (**b**) French apple cider process with pulp enzyming.

French Apple Cider Process with Pulp Enzyming. The process using pulp enzymes is shown in Figure 11*b*. Apples are washed and crushed. The pulp is macerated with 150 to 250 g/ton Rapidase® PE FLOTATION (Gist-Brocades) at ambient temperature for 30 to 60 minutes. The juice extraction is performed with press or decanter, depending on available machinery and expected yield. After pressing and leaching, the pomace is dried and sold to pectin manufacturers. The juice is demethylated in 4 to 12 hours after extraction, depending on the juice temperature and the enzyme dosage. The demethylated juice is then clarified by flotation. Anhydrous $CaCl_2$ is added at 44 g/hL. The clarified juice is centrifuged and can be fermented at low temperature or stored. Fermentation is stopped at the required sugar gravity by centrifugation and earth filtration or microfiltration.

Cloudy Apple Juice

Stable cloudy juice is difficult to process. Because of the natural clarifying action of actual commercial pectinases, cloudy juice production with enzymes has not been possible until now. Prevention of juice oxidation is also important to avoid problems after bottling. To conciliate quality preservation together with a good extraction yield, the right enzyme and process must be chosen. The following process has been designed to improve the production of highly cloudy apple juice (about 1,500 NTU) with improved yield and reduced cloud loss risk.

Process Description. The process for making cloudy apple juice is shown in Figure 12*a*. Apples are washed sorted out and crushed. Pulp is added with 400 ppm of ascorbic acid to prevent oxidation. Maceration is carried out with 25 g/ton Rapidase® CLOUDY (Gist-Brocades) for 15 to 30 min at ambient temperature. Stirring must be avoided to prevent oxidation. Juice extraction can be performed with decanters, belt presses, or horizontal presses, depending on available equipment and required yield. Juice must be pasteurized as soon as possible after extraction (20 min is a maximum delay between pressing and pasteurization) to prevent any clarification. The juice is centrifuged for cloudiness standardization. It is then stored in aseptic tanks or immediately bottled. Some processors want to valorize apple pomace for pectin production. The process described next has been designed to improve juice yield by enzymatic treatment of the pulp without degradation of the protopectin of the pomace (37). Enzymes used in this process for pulp maceration do not contain pectin depolymerases but only a purified PME. In the mash, the pectin demethylation is followed by the formation of insoluble calcium pectate, which remains within the pomace after pressing.

The stronger firmness of the pulp containing pectate results in an increase of free-run juice and an easier release of the juice when high pressure is applied, such as with a hydraulic press. The partially demethylated pectin binds to other cloud components such as proteins. It remains in suspension in the juice after pressing, providing a better cloud stability. The extracted juice can be processed either to slightly cloudy juice (200 NTU) or to clear juice or concentrate.

Cloudy Apple Juice Process with Pomace Recovery. Apples are washed, sorted, and crushed (Fig. 12**b**). Pulp is macerated with 100 to 300 g/ton during crushing to apple mash, simultaneously with ascorbic acid at 200 to 400 ppm Rapidase® PE MASH for 30 to 60 min at ambient temperature. If the temperature is too low, the pulp can be heated up to 20 °C (68 °F). Stirring should be avoided to prevent oxidation. Juice extraction can be performed with decanters or presses. The juice is pasteurized and can be later processed to cloudy single-strength juice or clear juice. It must be centrifuged if processed as cloudy juice. It is then bottled or stored in aseptic tanks. It is depectinized if processed as clear juice with 2 to 4 g/hL Rapidase® PRO (Gist-Brocades) for 1 to 2 h at 50 °C (122 °F). The depectinized juice is clarified with traditional fining and filtration or UF and then concentrated.

ENZYMES AND PEAR JUICE PROCESSES

Commercial enzymes sold for maceration, depectinization, and liquefaction of apples work well to degrade pear pulp. However, juice producers sometimes have problems in processing such fruit. The pear is very fragile, and it ripens fast, resulting in a decrease of pressability and yield. Changes during ripening and postharvest storage are similar to those described for apples. The soluble pectin content increases, and there is a pectin de-esterification (endogenous PME) and partial pectinolysis to lower molecular weight uronides (endogenous PG) (Table 2). Pectate content increases. Phenolic compounds are oxidized during ripening and form insoluble compounds (lignin). In pear, the presence of scleroids or stone cells in the pulp results in a lower textural quality. These stone cells are lignocellulosic and contain approximatively 18% lignin and 82% carbohydrates, which makes filtration or UF difficult (38). Araban, hemicellulose, and cellulose contents are higher than in apples. Linear arabans can also reassociate after concentration, producing araban haze. This is why it is important to use pectinases with strong arabanase and arabinosidase activities. Enzymes such as Rapidase® Pear (Gist-Brocades), Pectinex AR (Novo Nordisk) or Pearex (Solvay) prevent haze formation and improve the stability of the concentrate.

Clear Pear Juice Concentrate

Pear Pulp Maceration and Juice Depectinization. The process is similar to that for apples (Fig. 5). Pears are crushed and enzymes, for example, Rapidase® Press (Gist-Brocades), are added to the pulp at about 75 g/ton for 30 to 60 min at ambient temperature.

Pectinases PL, PME, and PG and arabanases contained in Rapidase® Press improve the extraction and the juice yield. After pressing, the juice is depectinized with pectinases containing high *endo*- and *exo*-arabanase activities (e.g., Rapidase® Pear from Gist-Brocades) for about 1 h at 45 to 50 °C (113 to 122 °F). The juice is then fined and filtered or ultrafiltered and concentrated up to 70 °Brix. Pears can also be processed to purees or nectars.

Pear Total Liquefaction. Next to pectinases, hemicellulases are of first importance for pear liquefaction (Table 5).

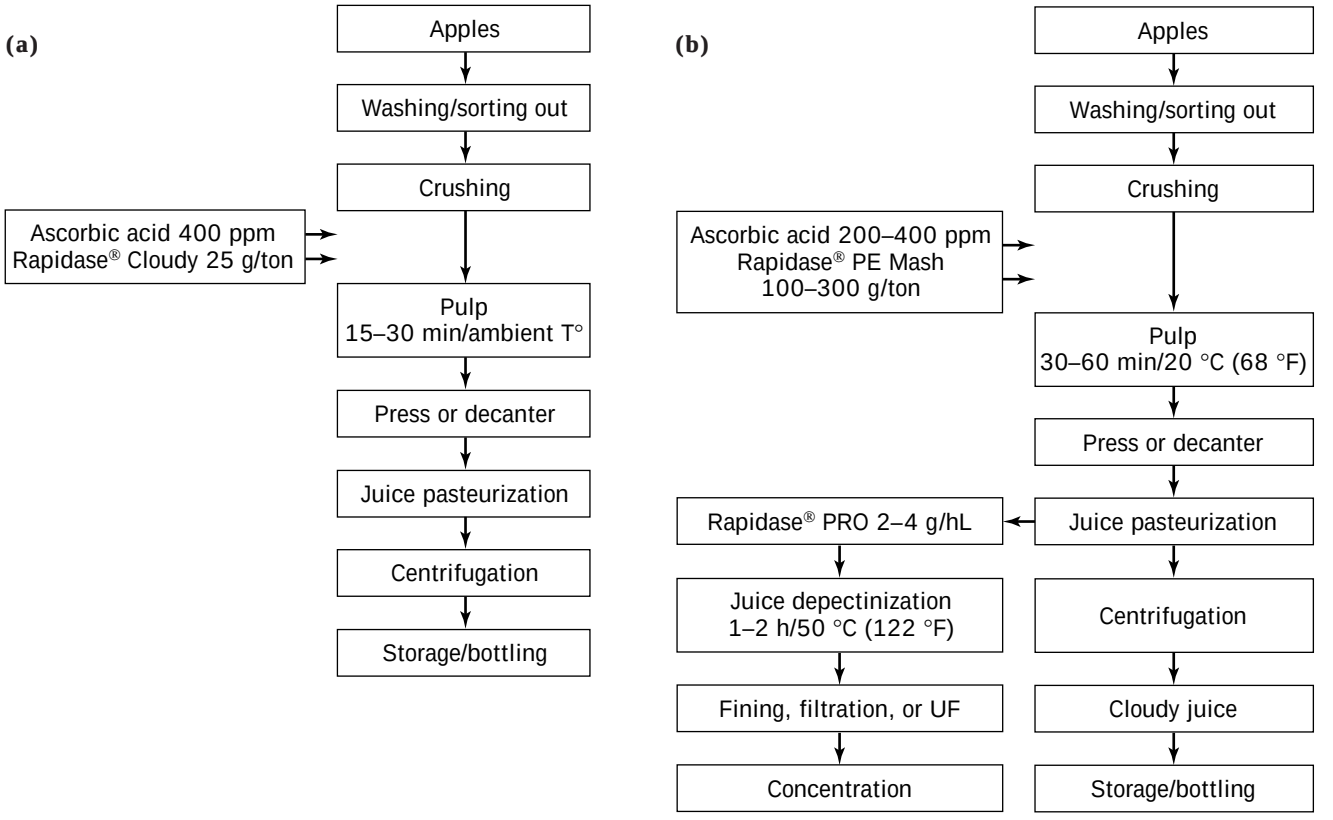


Figure 12. (**a**) Cloudy apple juice process. (**b**) Cloudy apple juice process with pomace recovery.

The pear total liquefaction process is similar to the apple total liquefaction process, including a hot break before enzyme addition (Fig. 8). Pears are crushed and heated up to 85 °C (185 °F) for 4 to 5 min to destroy the endogenous PPO and to open the cell wall. Cooled down at 50 °C (122 °F), the pulp goes through a finisher that discards seeds, stems, and stone cells. The liquefaction is carried out during 2.5 h at 52 °C (125.6 °F), with enzymes such as Rapidase® Pear Liq (Gist-Brocades) or Pearex (Solvay) added from 200 to 250 g/ton with continuous stirring at 60 rpm. The juice is centrifuged with a decanter and pasteurized. Then, aroma recovery is performed. The juice is either fined to remove proteins and tannins and filtered or ultrafiltered and then concentrated up to 70 °Brix. Pear total liquefaction is carried out in only one step; no depectinization is required.

ENZYMES AND RED BERRY JUICE PROCESSES

The production of clear juices and concentrates from strawberries or blackberries requires enzymatic depectinization. The clarification, filtration, and concentration remains difficult because these juices contain high amounts of pectic substances (39). It is always difficult to obtain a complete depectinization of the juices, and it is normal to find a residual low molecular weight pectin level of 7 g/L, compared with about 0.5 g/L for apple juice. The assumption is that hairy regions of the strawberry pectin remain as colloidally dissolved residual pectins in the juice, whereas rhamnogalaturonans are the main group of residual pectins in raspberry juice. Neutral sugars are also important in the depectinization. With typical molecular weights from 10,000 to 1 million Da, polysaccharides, arabinose, and galactose are the main sugar constituents, probably in arabinogalactans. Moreover, residual pectins and hemicelluloses tend to bind to phenolics and proteins during processing and storage. The result is the formation of irreversibly linked complexes that enzymes cannot break down anymore. An additional problem is related to the frequent contamination of red berries, mainly strawberries and raspberries, by *Botrytis cinerea*. This parasitic fungus, which in many cases cannot be seen from inspection of the fruit, secretes a β1,3–1,6-linked glucan into the berries with molecular weight around 10^6 Da. This gum reduces the filterability and the clarity of the juice. It is possible to hydrolyze this glucan with β-glucanases such as Glucanex (Novo Nordisk). Berry juices must have a good flavor and are attractive in color. For this reason, they have a long history of use in the preparation of fruit derivatives and other fruit beverages. With the exception of blackberry and cranberry juices, none is commonly used as a table beverage. Common berries have low sugar content (from 4 to 8%).

The average acidity is about 1% calculated as citric acid. Cranberry is an exception, with about 1.9% or more. It is for this reason that most berry juices are diluted with a half volume of light sugar syrup. Pectinases are used to

improve the pressing stage and to extract the red color, while retaining the organoleptic properties of the fruit. The color is located in the skin of the berries. It can be released by heat, alcohol, or enzymes. However, it is not advantageous to extract the color if it is then destroyed by oxidation. Oxidation can be chemical, caused by atmospheric oxygen, or enzymatic, from the endogenous polyphenoloxidase (PPO) and peroxidase in the fruit. These enzymes can be inhibited with sulfur dioxide addition or with heating. It is therefore recommended that the pulp be heated to over 85 °C (185 °F) to inhibit these endogenous enzymes. This action not only improves the release of color but also the protopectin solubilization, and this is why pectinases are required. Black currants are very acid (pH of 2.6 to 2.8) and contain high levels of phenolics and anthocyanins. Both acids and polyphenolics inhibit fungal pectinases. This results in high doses of enzymes that are added and emphasizes the importance of enzyme stability in the juice.

Red Berries Single Stage Process. The single-stage process (Fig. 13**a**) is recommended for berries such as blackberry, bilberry, elderberry, gooseberry, raspberry, and black currant. This single-stage process allows the performance of maceration and depectinization at the same time. Yield is maximized, and good quality is preserved. This process is recommended when decanters are used for juice extraction. Fruits are inspected and sorted out, then thawed if necessary and crushed. The pulp is heated up to 90 °C (194 °F) for a thermal break of 2 min minimum to increase color extraction and to destroy endogenous polyphenoloxidase (PPO). Then the pulp is cooled down to 50 °C (122 °F) for enzymatic maceration. Enzymatic maceration is carried out with 200 to 400 g/ton Rapidase® BE Super (Gist-Brocades) or Pectinex BE (Novo Nordsik) at 50 °C (122 °F) for 1 to 2 h. If the pulp pH is below 3.2, we recommend the use of Klerzyme® 120 (Gist-Brocades) at same dosage instead of Rapidase® BE Super. Juice separation can be performed either with press or with decanter. The extracted juice should be pectin free (acidified alcohol test or acetone test should be negative). No further enzymatic treatment is required. Juice is pasteurized after extraction for protein denaturation. A minimal temperature of 90 °C (194 °F) is necessary. Cooling down below 50 °C (122 °F) leads to protein insolubilization. Pasteurized juice can be clarified either with UF or with traditional fining with gelatin, bentonite, and silicasol. In the case of UF being used, membrane cutoff should not be below 30,000 Da to prevent color loss. The juice is concentrated.

Red Berry Two-Stage Process. Applied to the same berries as described for the single-stage process, the two stages process consists of an enzymatic maceration of fruit pulp followed by a second enzymatic treatment of the extracted juice for its depectinization (Fig. 13**b**). The low processing temperature after the hot break prevents aroma losses, and the best quality juices and concentrates can be produced. Fruits are inspected and sorted, thawed if necessary, and crushed. Fruit pulp is heated up to 90 °C (194 °F) for a thermal break of 2 min minimum to increase color extraction and to destroy endogenous PPO. Then the pulp is cooled down to 20 to 25 °C (68 to 75 °F) for enzymatic maceration, which is carried out with Pectinex BE or Rapidase® BE Super at 100 to 200 g/ton at 20 °C (68 °F) for 30 to 60 mins. If the pulp pH is below 3.2, we recommend the use of Klerzyme® 120 at same dosage instead of Rapidase® BE Super.

Juice separation can be performed either with the horizontal or belt press. After extraction, juice is pasteurized at minimum 90 °C (194 °F) for protein denaturation. Cooling down to 20 °C (68 °F) leads to protein insolubilization. Depectinization is carried out with Rapidase® BE Super or Pectinex BE at 4 to 8 g/hL at 20 °C (68 °F) for 1 to 2 h. If the pulp pH is below 3.2, we recommend the use of Klerzyme® 120 at same dosage instead of Rapidase® BE Super. The acidified ethanol or acetone test should be negative at the end of the treatment. The juice can be clarified either with UF or with traditional fining with gelatin, bentonite, and silicasol. In the case of UF being used, membrane cutoff should not be below 30,000 Da to prevent color loss. The juice is concentrated.

Strawberry

In contrast to black currant or red currant, it is impossible to heat the strawberry pulp because it would create an unpressable puree. Furthermore, many aromas would be lost during heating, oxidation, destruction, or evaporation. Because it is not recommended to heat such pulp, endogenous PPO cannot be destroyed, which can result in condensation reactions between oxidized products and anthocyanins and proteins and pectin, leading to great loss of color. Additionally, anthocyanin pigments are temperature and oxidation sensitive. Thus, it is essential to work at low temperatures to process delicate fruit such as strawberry. Special care is required to prevent oxidation and to keep aromas. Quality in term of the sanitary state of processed fruit is of first importance. Contaminated strawberries could contain very viscous β-glucan from *B. cinerea*. Prevention of oxidation with quick processing is also of key importance.

Process Description. Strawberries are inspected and sorted, thawed, and crushed with a hammer mill fitted with a 0.635-cm screen (Fig. 14). Pectinex BE (Novo Nordisk) or Rapidase® BE Super (Gist-Brocades) can be added at 100 to 150 g/ton to the pulp at 20 °C (68 °F) for 30 to 60 min. For the processing of frozen fruits, enzymes are added after heating when the temperature has decreased to about 20 °C. The juice is extracted with a horizontal press, belt press, or decanter. Extraction yield is around 70 to 80%. The juice is flash pasteurized at minimum 90 °C (194 °F) for protein denaturation and a simultaneous aroma stripping. The juice is then cooled down at 30 °C (86 °F) to be depectinized with the same enzymes at 2 to 5 g/hL over 1 to 2 h. In countries where glucanase treatment is authorized, Cytolase CL (Gist-Brocades) or Glucanex (Novo Nordisk) glucanase-containing products can be added for glucan hydrolysis at 2 to 5 g/hL together with pectinases. To control depectinization and glucan hydrolysis, two specific tests are performed. The specific alcohol test is per-

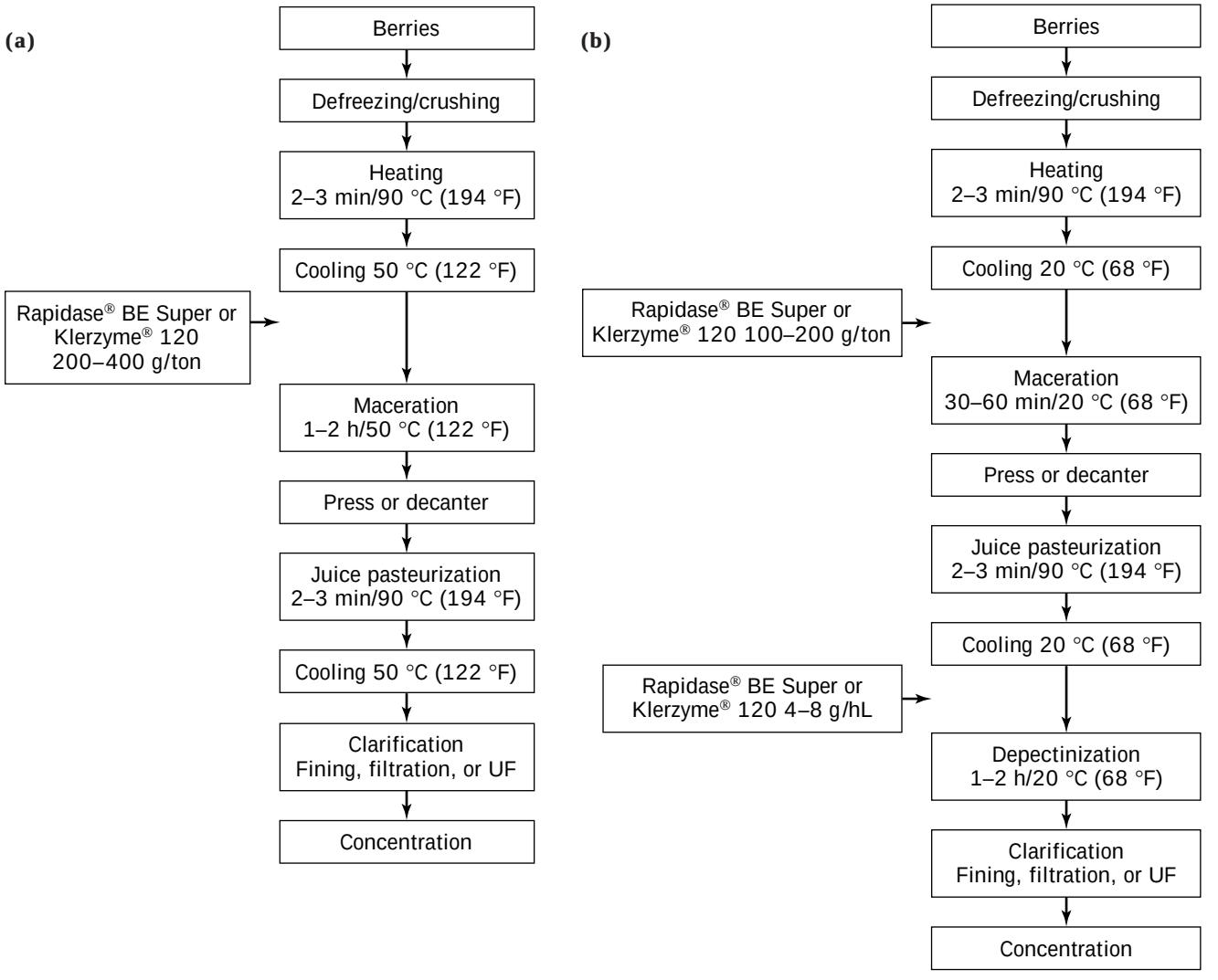


Figure 13. (**a**) Red berry juice, single-stage process. (**b**) Red berry juice, Two-stage process.

formed to check glucan degradation. One vol ethanol plus 1 vol juice are gently mixed. White filaments appear when glucanes are present. The specific acetone test is performed to check pectin degradation. The alcohol test used to check the depectinization of apple juice is not suitable for red berries because anthocyanins and other compounds can also precipitate together with pectin. We recommend mixing 1.5 vol juice with 1 vol acetone. The test is negative when the sediment is less than 1:10 of the control or when the amount of precipitated colloids is less than 1 g/L (when strawberry concentrates are blended for the liquor industry, no precipitation or haze formation should occur with this test). Strawberry juice is clarified either with UF or with traditional fining treatment with gelatin, bentonite, and silicasol. In UF, the membrane cutoff should not be below 30,000 Da to prevent any color loss. The clear juice is then concentrated. *Note*: It is important to have stainless steel equipment to avoid oxidation catalyzed by metallic ions such as iron, copper, or aluminium. A final UF step before concentration is recommended.

Cherry

The global production of cherries, supplied by more than 55 countries, varies from year to year between 2 and 3 million tons, more than 60% of which is sour cherries. Fruit processing involves an average of 40% of sour cherries and 15% of sweet cherries. The most important world producers of cherries (in thousands of tons) are Germany (280), the United States (170), Italy (140), and the former Soviet Union (130). Production in China is growing fast. Cherry pectin has smooth and hairy regions, with arabinose and galactose as main neutral sugars.

Process Description. After washing and sorting, cherries are destoned and crushed (Fig. 15). The pulp is heated up to 85 °C (185 °F) for 3 min to destroy endogenous PPO and to increase color extraction, and then cooled down to 50 °C (122 °F) for enzymatic maceration. Maceration is carried out with Klerzyme® 120 at 300 to 400 g/ton for 2 h at 50 °C (122 °F). The juice is extracted with a horizontal

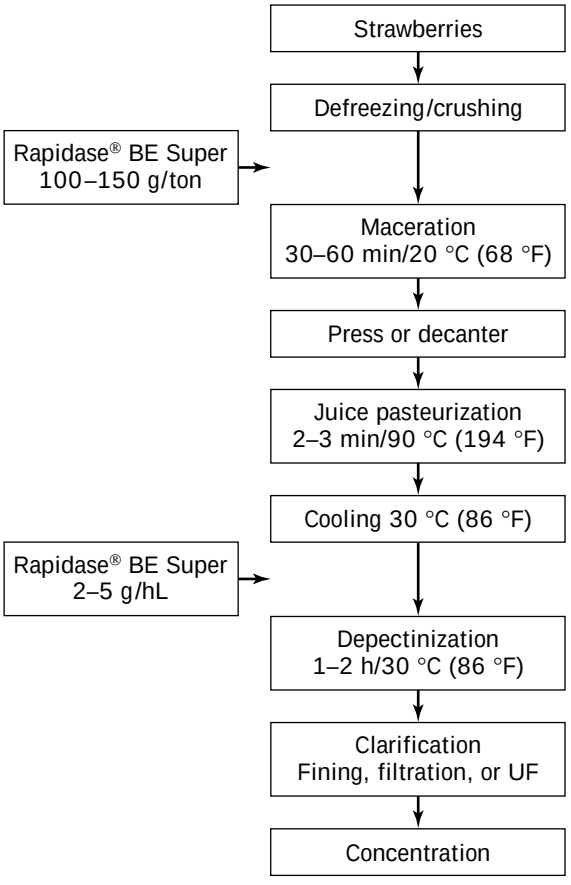


Figure 14. Strawberry juice process.

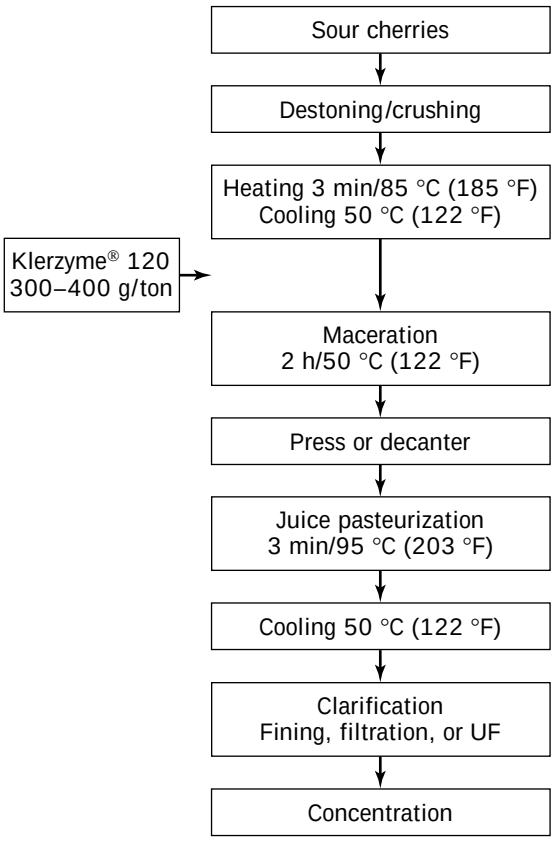


Figure 15. Cherry juice process.

press or decanter. Extracted juice should be pectin free (alcohol test negative). It must be pasteurized to insolubilize proteins at minimum 95 °C (203 °F). The juice is then cooled down to 50 °C (122 °F) for clarification either by traditional fining and earth filtering or by UF, with the membrane cutoff above 30,000 Da. The juice is then concentrated.

Cranberry

Cranberry juice is very acidic, with an acid content exceeding 2% w/w, and a pH of 2.6 or lower. The juice is very astringent because of a high natural level of tannins and leukoanthocyannins. Because the single-strength juice, with a Brix-to-acid ratio of 3.75, is unpalatable (the ratio for grapefruit is 10), formulated cranberry juice blends have been developed by companies such as Ocean Spray. Many other cranberry-based juice drink blends have been marketed. They have a fresh taste and attractive pink color. Cranberry is reported to be the third favorite flavor in the United States after orange and apple (40). Cranberry pectin exhibits different properties than citrus or apple pectin. Cranberry sauce is manufactured with a solid level of only 40%, whereas other fruits need 65% solids. In addition to citric and malic acid, the fruit contains also quinic and benzoic acids. With a low carbohydrate content and significant quantities of vitamin C, potassium, and phenolic glycosides, this juice is considered healthy.

Process Description. Fresh or frozen fruits are milled and heated up to 85 °C (185 °F) for 5 min, and then pulp is cooled down to 50 °C (122 °F) for enzymatic maceration (Fig. 16). Biopectinase BP300 (Quest) or Rapidase® Cranberry (Gist-Brocades) pectinase is added to the pulp at 200 to 400 g/ton for 1 to 2 h. The juice is extacted with a decanter or horizontal press, pasteurized, ultrafiltered, and concentrated up to 50 °Brix.

Concord Grape

Grape juice is not consumed in large amounts because it is too sweet (about 200 g/L sugars) or too acidic (up to 10 g/L tartaric acid). It is mixed with other fruit juices such as apple. Pederson in 1980 reviewed the technology of grape juice processing (41). Red grapes, especially Concord variety and its hybrids, are the principal grapes for the U.S. industry. These grapes are processed to juice. The Concord variety is rich in pectin and anthocyanins and has a unique flavor due to methyl anthranilate. The acidity is 13 g/L as tartaric acid. This juice is often recommended in special diets. Because of the high pectin content (5 to 10 g/L), berries are difficult to crush and to press.

Process Description. Grapes are destemmed, crushed, and heated up to 85 °C (185 °F) for 3 min to release red color pigments from the skins and to destroy PPO (Fig. 17). After cooling down to 60 °C (140 °F), enzymes such as Cytolase® PCL5 (Gist-Brocades) or Ultrazym (Novo Nor-

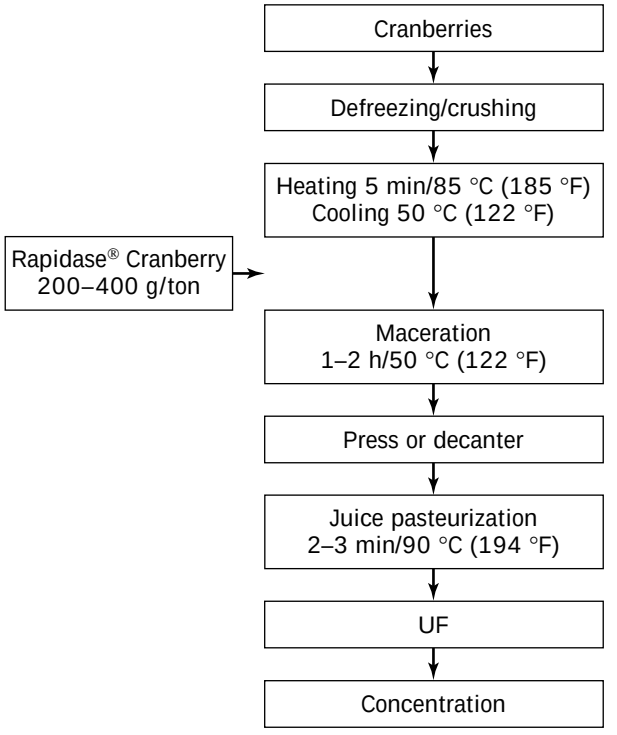


Figure 16. Cranberry juice process.

disk) are added at 50 g/ton for 1 h. Precise monitoring of temperature is important; it should not be below 55 °C (131 °F). After 1 h, juice extraction can be carried out with decanters, a screw press, or a horizontal press, after addition of cellulose paper for absorbing particles. The juice is pasteurized at 88 °C (190 °F) for 30 sec and then cooled down to −1 °C (30 °F) to prevent start of alcoholic fermentation. The juice is clarified at low temperature for 7 to 14 days with Klerzyme® 200 (Gist-Brocades) at 200 to 400 ppm. Enzymatic clarification of grape juice as well as other juices is based on the partial hydrolysis of soluble pectin that acts as a protective colloid. Pectinases, mainly PG and hemicellulases-like arabinogalactanases, help insoluble particles to flocculate.

At the same time, cold temperature induces the insolubilization and sedimentation of tartaric acid salt, reducing the total acidity to acceptable levels. The juice is then filtered with diatomaceous earth on vacuum filter and then on pressure filter and concentrated or pasteurized and bottled. Concord grapes can also be processed with enzymatic liquefaction. A thermal break is necessary. Pulp is heated up to 85 °C (185 °F) for 3 min and then cooled down to 60 °C (140 °F) for enzymatic maceration, which is carried out with 300 g/ton Klerzyme® LIQ. Temperature monitoring is important and should not be below 55 °C (131 °F) nor above 60 °C (140 °F). Juice extraction is carried out with decanters. It is pasteurized at 88 °C (190 °F) for 30 sec. Clarification takes place at cold temperature, −1 °C (30 °F) to 2 °C (36 °F) for 3 to 7 days using 10 to 20 g/hL Klerzyme® 200. Tartaric stabilization occurs during this period. Clarified juice is usually vacuum filtered and then pressure filtered and concentrated.

White Grape

Different red and white grape varieties such as Niagara, Thomson seedless, Sultana, and Muscat are processed to juice and concentrate. Because grape sugar content is very high and because the berry has endogenous PPO activity, it is necessary to process this type of fruit very quickly to prevent start of alcoholic fermentation and juice oxidation.

Process Description. White grapes are destemmed and crushed (Fig. 18**a**). Special care should be taken to avoid rotten fruits, frequent contamination by *B. cinerea* results in the presence of glucan and laccase in the fruit. Glucans cause pressing problems, and laccase is a fungal polyphenoloxidase causing juice browning. Grape pulp is macerated at ambient temperature for 1 to 2 h with 40 to 80 g/ton Klerzyme® 120. If the temperature is too warm, cooling down to 15 to 20 °C (59 to 68 °F) is required to prevent fermentation. Pressing can be carried out with a horizontal press, pneumatic press, or screw press. The juice is quickly pasteurized. After cooling at 50 °C (122 °F), depectinization is carried out with 2 g/hL Klerzyme® 120 or Ultrazym (Novo Nordisk) for 1 to 2 h. Enzyme is added at the beginning of tank filling. No stirring is required. Depectinized juice can be filtered or ultrafiltered. It is then concentrated.

Red Grape

A main concern for red grape juice producers is color extraction. Process conditions, the sanitary state of grapes, and enzymes are key factors for good color stabilization in the final juice. Anthocyanidins, responsible for the red color, are glycosylated with glucose. These pigments turn from red to brown depending on light, pH, temperature, enzyme oxidases (naturally present in the grape), and β-glucosidases (present in grape and a side activity of pectinase products) that destabilize anthocyanins after hydrolysis of their glycosidic part.

Process Description. Red grapes are sorted out before crushing (Fig. 18**b**). Berries are heated up to 90 to 95 °C (194 to 203 °F) for 3 min to increase color extraction and to destroy endogenous PPO. Pulp is cooled down to 50 °C (122 °F) for the enzymatic maceration. Rohapect VR (Röhm) or Klerzyme® Color is added at 40 to 100 g/ton for 2 to 5 h. Pulp should be carefully mixed with enzymes during tank filling and just before pressing. No stirring is required during maceration. Color extraction is monitored, and the juice should be pectin free (alcohol test). Juice extraction can be carried out with horizontal, pneumatic, screw presses, or decanters. Juice should be pasteurized quickly after extraction for a better color stability. Juice is vacuum filtered, and then pressure filtered or ultrafiltered (membrane cutoff higher than 50,000 Da) and concentrated.

ENZYMES, STONE FRUIT AND TROPICAL FRUIT PROCESSES

The production of fruit puree, nectar, pulpy juices, and fruit blends are of great commercial importance. Many

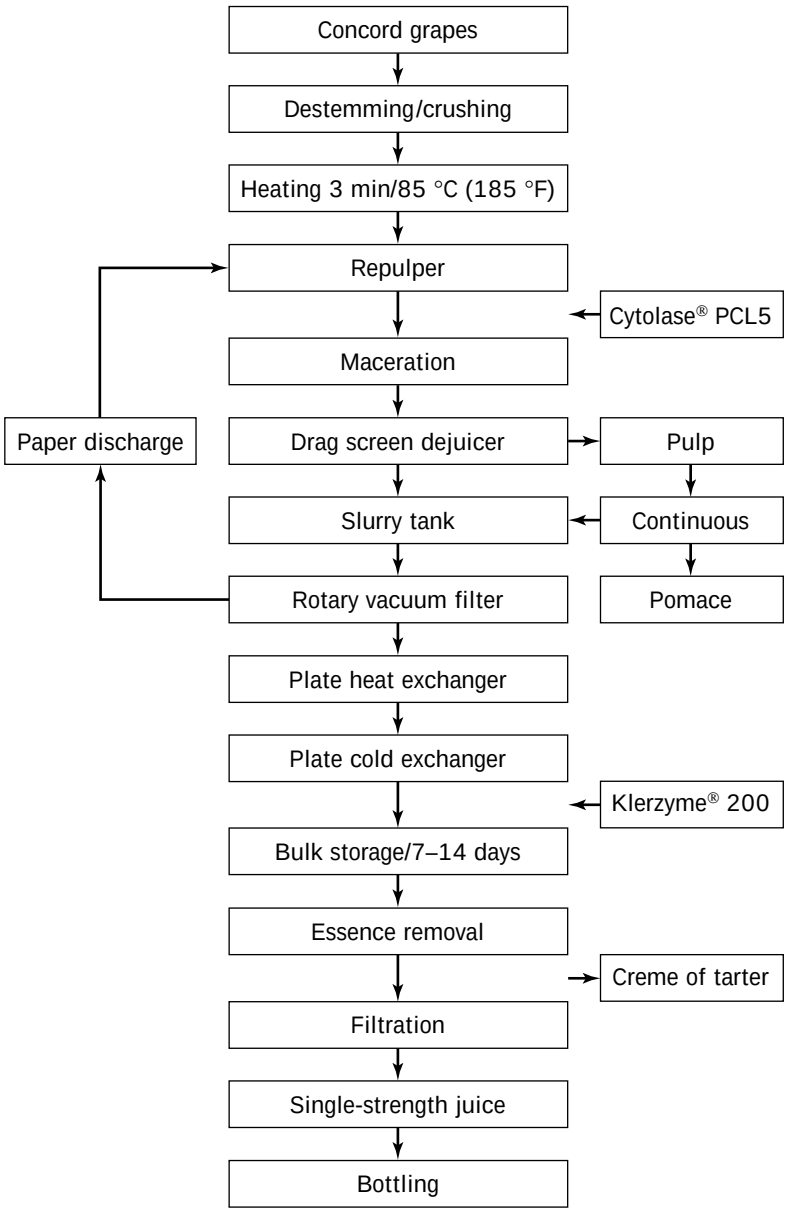


Figure 17. Concord juice process.

fruits are not pressable, or they are too acidic or too strongly flavored to be pleasant beverages without dilution or blending or both. These juices are delicious after diluting with thin syrup or blending with other juices; good examples of this are apricot, peach, pear, and tropical fruit such as mango, guava, papaya, and banana. Other fruits may not have enough flavor, and they are greatly improved if the entire fruit is processed into a smooth pulpy beverage. Fruit nectars are pulpy fruit juices blended with sugar syrup water and citric acid to produce a ready-to-drink beverage. Their composition is strictly defined in the Code of Federal Regulations (1984). They come from diluted purees or pulpy juice process. Cloudy beverages have a high content of fruit pulp (mango and guava have more than 35% pulp). The most important quality factor of these nectars is that the cloud be stable. After mixing with water, sugar, and acid, the cloud particles tend to settle down and sometimes form a gel with a clear supernatant layer. This separation is a serious defect that reduces the attractiveness of the product. Fruit polysaccharide structures change during ripening because of endogenous enzymes (Table 2). The fruit texture is modified and has an effect on the texture, viscosity, and stability of further-processed nectars. Prevention of cloud loss or instability in nectars has focused on the use of homogenization as a mechanical means to reduce the particle size. The use of exogenous enzymes can improve the cloud stability in nectars or make puree concentration easier. The general process is used to produce puree, nectar, cloudy, or clear juice from pear, peach, apricot, mango, or guava.

Process Description. Fruits are washed, sorted, halved, and destoned (Fig. 19). Destoned fruits are heated with

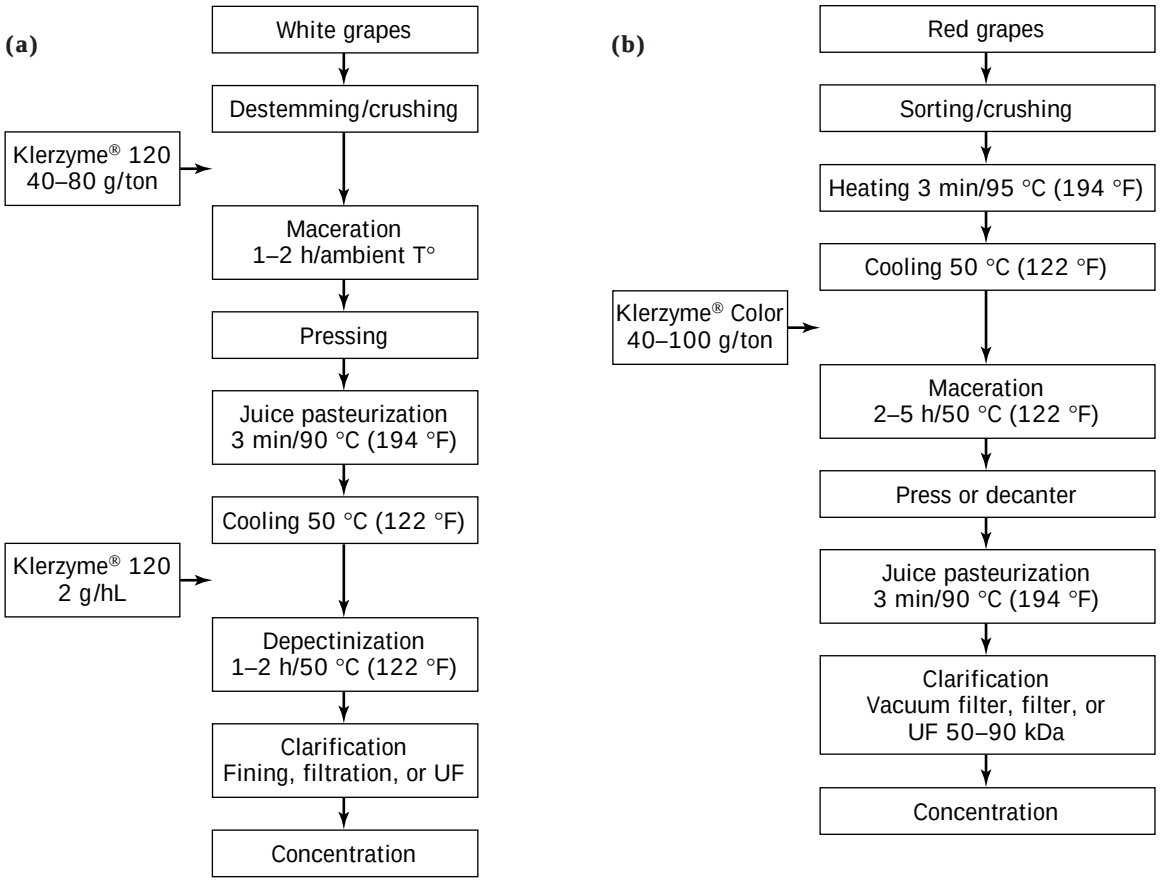


Figure 18. (**a**) White grape juice process. (**b**) Red grape juice process.

steam in a continuous blanching step at about 85 °C (185 °F) for 5 min to inactivate endogenous PPO, to avoid puree browning, and to kill microorganisms. Enzyme preparations should have high PG or PL activities combined with cellulases and hemicellulases. For puree processing, maceration is carried out with Pectinex UltraSP (Novo Nordisk), Rohapect TF (Röhm), or Rapidase® TF (Gist-Brocades) at 50 to 100 g/ton for 1 to 2 h at 50 °C (122 °F). These enzymes have been found to decrease the viscosity of puree or cloudy juice without destabilizing the cloud and to improve the ability to concentrate the product. Puree extraction is performed with a finisher (with a 0.6 to 0.8-mm screen). The resulting puree is concentrated or deaerated and passed through a steam-heated tubular heat exchanger where it is quickly heated to pasteurization temperature, 88 to 95 °C (190.4 to 203 °F), before canning. For the production of nectar, the puree is sweetened with sugar syrup and acidified with citric acid to maintain a constant ratio of sugars to acid. The nectar is then canned or bottled, sealed and processed at 100 °C (212 °F) for 2 to 3 min. The Tetra-Pak system is now widely used for aseptic packing of nectars. Cloud instability in nectars depends mainly on the presence of large pulp particles associating together, and it is also related to the endogenous PME. This enzyme leads to the formation of calcium pectates that form sediment and result in clarification. It is necessary to inactivate this activity as soon as possible in the process, because even a low residual activity after processing would destabilize the nectar during storage after bottling. It has been found that PME from guava or mango (at pH greater than 4) had a high heat stability, and it was necessary to heat the nectar for at least 1 min at 97 °C for its inactivation.

For clear juice processing, 150 to 200 g/ton Rapidase® TF for 2 to 2.5 h at 50 °C (122 °F) is added to the puree. The juice is extracted with a horizontal press or decanter and then pasteurized at 95 °C (203 °F) to denaturate proteins. Clarification is carried out with diatomaceous earth filter or UF. The clear juice can be concentrated.

Peach and Apricot

The role of various polysaccharide degrading enzymes for cloud stabilization in peach and apricot nectar has been demonstrated. Fruit primary cell walls contain around 90% polysaccharide and 10% protein. Apricot nectar exhibits a phenomenon of cloud loss. This separation can be avoided by grinding the pulp very finely (42). The addition of pectolytic enzymes separately or in combination with homogenization results in a cloud-stable apricot nectar, but homogenization alone fails to achieve cloud stability. The cellulose microfibrils liberated by mechanical action become more hydrated and interact with constituents in the serum, increasing the yield and the cloud stability. En-

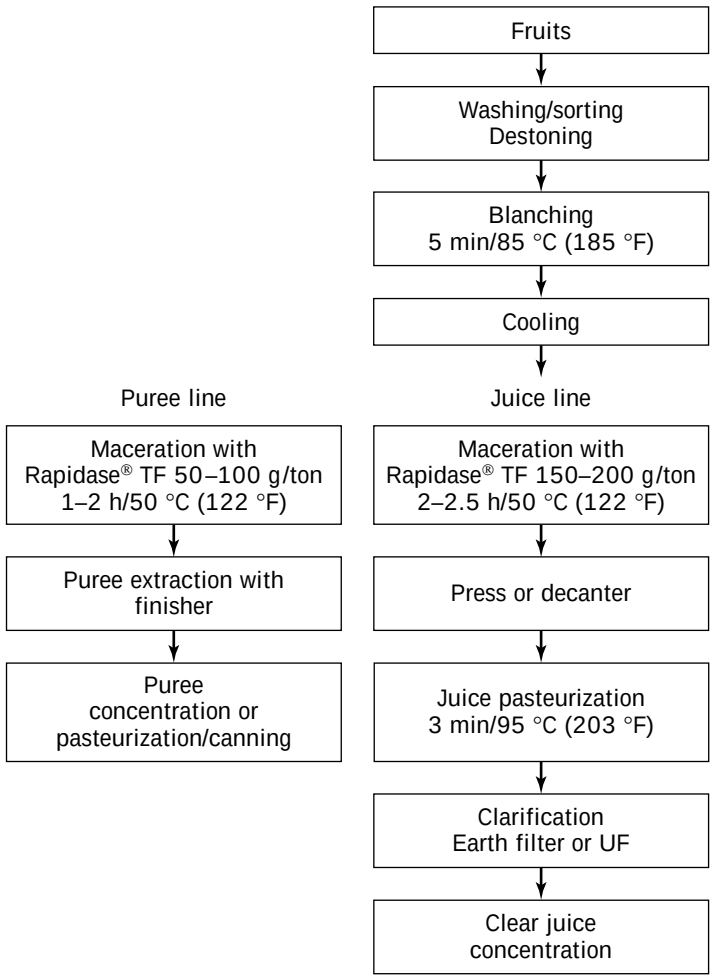


Figure 19. Stone fruit and tropical fruit juice process.

zymes such as Rapidase®TF are able to degrade the arabinan-rich pectin, open cells in the pulp, and assist in the production of clear peach and apricot juice concentrate.

Mango

Mango is the most important tropical fruit. India is the largest producer, but it is also produced in Mexico, China, Brazil, and the Philippines. The annual world production is about 20 million metric tons. The skin color of ripe fruits is yellow-red, under which the flesh is white, pink, green, or orange-yellow and juicy. The taste varies from sweet to sour. It is a source of vitamin C (100 mg per 100 g) and nicotinic acid. Mango is processed as a puree or nectar, juice, and various juice drink blends. A typical mango nectar is one with 20 to 30% puree, 12 to 18 ° Brix, pH around 3.5, and titratable acidity 0.2 to 0.3%. Exogenous enzymes are used to increase the concentration of mango puree or to produce clear juice and concentrate. Enzymatic treatment of mango destoned flesh provides a yield of around 80% at 15 °Brix. Pectinases and hemicellulases are important activities to degrade mango pulp (Tables 2 and 5).

Guava

Guava is a popular fruit throughout the tropics, eaten fresh or made into juice, nectar, puree, and jam or jelly after removal of the seeds. The major producers are South Africa, India, and Hawai. Guava puree is reprocessed into nectars, juice drink blends, syrups, jams, or jellies. Fruit pulp is increasingly treated with exogenous enzymes to improve juice extraction. Guava pericarp has two main cellular types: parenchyma and stone cells, which are also encountered in pear, plum, and olive mesocarps (43). Stone cells are strongly lignified woody material. Cell walls from stone cells represent the major proportion of total cell wall material found to be insoluble in alcohol (74% of insoluble substances). The mesocarp portion contains 90% of total cell wall material of guava pulp, 77% of which are stone cells. Half of stone cells are made up of neutral polysaccharides, 94% of which are cellulosic glucose and xylose (Table 2). Uronic acids are also present and *O*-acetyl groups are important and could be related to xylose-containing polymers. The dominant structural feature in parenchyma cell walls is cellulose, with side chains of glucans and xyloglucans, highly branched arabinans, and type I arabinogalactans; the rhamnose comes from rhamnogalacturonans.

Guava stone cells contain large amounts of β-(1→4) xylan as the main constituent. Compared to other fruit juices, an additional centrifugation is required to remove stone cells.

Yellow Banana

Banana production is a key activity in a lot of equatorial countries. A considerable amount of bananas are rejected because they do not meet quality standards; this waste can be substantial (160,000 tons of rejected bananas in Costa Rica in 1978). The quantity of processed bananas is very small because of the competition with fresh fruits throughout the year. Many banana products have been described, such as powder, flour, chips, puree, canned slices, jams, beverages, and alcohol. The banana is widely appreciated for its flavor and aroma (isoamylacetate) either as banana juice or a mixture with other juices. However, banana juices are too pulpy and too viscous to yield beverage juices easily, even if they contain endogenous pectinases such as PME and PG (44) (Table 2). Bananas contain 3 to 4% fiber, 1% pectin, and 2% hemicellulose (45). The composition of banana polysaccharides changes considerably during ripening, with a large decrease in hemicellulose, cellulose, pectin, and starch. Bananas can be processed at the green stage, mainly for alcohol production, or at the ripe stage, for the production of clear juice, requiring different enzyme treatments.

Process Description. Yellow bananas are peeled and ground (Fig. 20). Banana PPO, which is very active, must be inhibited by heating or sulfite treatment (100 mg/L of potassium metabisulfite) to prevent browning. Addition of ascorbic acid is also possible. After heating at 85 °C (185 °F) for 3 min and then cooling at 45 °C (113 °F), Rapidase® Press (Gist-Brocades) can be added at 75 to 100 g/ton over 1 to 2 h. After decanting, the clear juice at about 85% yield is pasteurized, filtered, and concentrated.

Clear Pineapple Juice

Annual pineapple production worldwide is 5 million tons in Thailand, the Philippines, Indonesia, and Kenya. Smooth Cayenne is the main variety. Thailand was the world's leading producer of commercial canned pineapple (420,500 metric tons) and pineapple juice concentrate (87,000 metric tons) in 1991. There are two types of pineapple juices: beverage juice (cloudy) from eradicated fruit and mild juice, the clear pineapple juice from wastes. The processing of pineapple mainly involves the fruit being sliced or cubed and then canned. A clear concentrate can be processed. The latter is used to cover the slices and cubes in the cans.

Process Description. Pineapple juice results from the pressing of the following by-products of the canning factory: the surplus of juice drained from the eradicator fruit, juice drained during crushed pineapple preparation, fruits that are too small for slicing, and fruit shell peel cores, trimmings, or end cuts (Fig. 21). All are blended, crushed, and pressed with a screw press to obtain a 8 to 10 °Brix

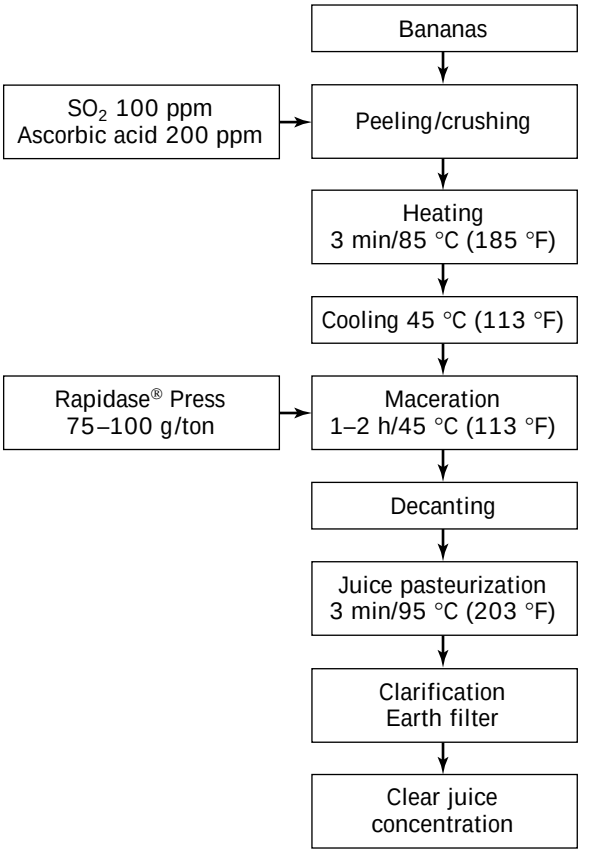


Figure 20. Banana juice process.

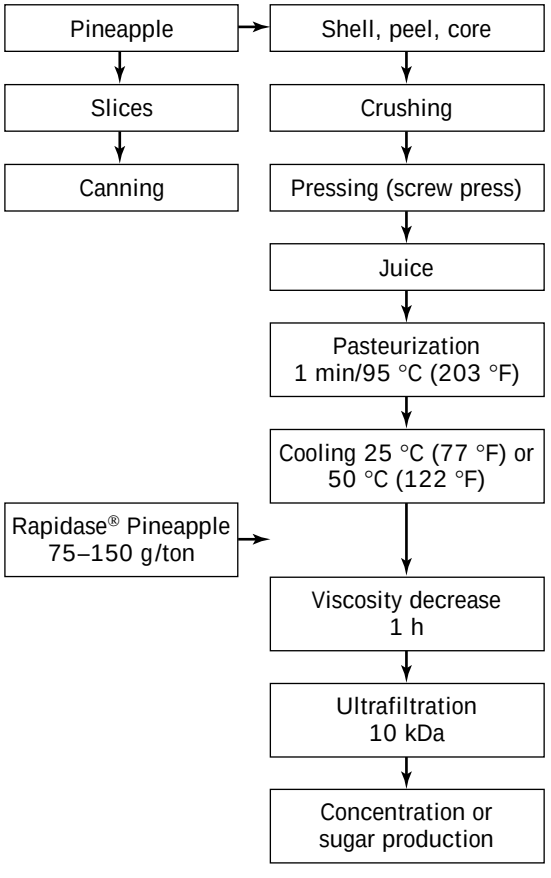


Figure 21. Clear pineapple juice processing.

juice with a of pH 3.5 to 4.0. The juice must be centrifuged to reduce its fiber content; the percentage of insoluble solids (spin test at 3,000 rpm for 5 min) should be below 1.0 to 1.5%. The juice is pasteurized to insolubilize bromelain protein at 95 °C (203 °F) for 1 min.

The juice is cooled down and depectinized at 25 °C (77 °F) for 1 h. Enzymes such as Rapidase® Pineapple (Gist-Brocades) or Rohapect B1L (Röhm) are added for the depectinization from 7.5 to 15 g/hL. Rapidase® Pineapple contains pectinases and specific hemicellulases that hydrolyze pineapple gum. This pineapple juice contains a small amount of pectin but has a high hemicellulose content (galactomannans, arabinogalactans, and galactoglucomannans), whereas the insoluble parts are rich in arabinoxylans (Table 5). The natural gum in pineapple juice was found to decrease the UF capabilities and increase pulp suspension and foaming properties of the juice (46). This gum is a neutral polysaccharide containing 70% sugars, which are predominantly galactomannans (mannose-to-galactose ratio of 2.25:1). Because of this gum, the UF flux rate quickly drops and becomes a bottleneck in the process when using UF membranes with a cutoff of 10,000 Da. Enzymes in Rapidase® Pineapple improve the UF flux rate because the xylanase and other hemicellulase activities reduce the polysaccharides' molecular weight. The juice viscosity is decreased, the UF flux rate is improved, and the juice is more easily concentrated at 62 °Brix.

Kiwi

The processing of kiwi puree requires no enzyme application. The processing of kiwi juice and concentrate is a new demand, especially in Korea, New Zealand, and Chile. This fruit is very rich in ascorbic acid (vitamin C) and contains more than 0.5% proteins. The fruit pulp is very viscous, and pectinases are required for processing because of a high pectin content of about 260 mg%.

Process Description. Kiwis are washed and sorted out then crushed with a hammer mill (Fig. 22). SO_2 is added to prevent pulp browning at 30 to 50 ppm. Pectinase products such as Rohapect D5L (Röhm) or Rapidase® Press (Gist-Brocades) at 100 to 150 g/ton are added to the pulp to improve the juice extraction. Maceration is carried out for 1 to 2 h at ambient temperature. After pressing or decanting, the juice is pasteurized at 95 to 98 °C (203 to 210 °F) for 1 min for protein denaturation. Depectinization is carried out with Cytolase® PCL5 at 2 to 4 g/hL for 1 to 2 h at 50 °C (122 °F). Centrifugation may be useful to remove heat-stable proteins. Clarification is obtained after fining with bentonite (100 to 200 g/hL) or with bentonite plus silicasol (75 g each) followed by earth filtration or UF. The clear juice is concentrated up to 65 °Brix.

Cloudy Passion Fruit Juice

Most of the passion fruit juice is sold as cloudy concentrate juice. Cloud stability juice yield, and concentrability are the main concerns of juice producers. Enzymes, pectinases, hemicellulases, and cellulases can increase the extraction juice yield and improve the concentration after decreasing the viscosity of the juice.

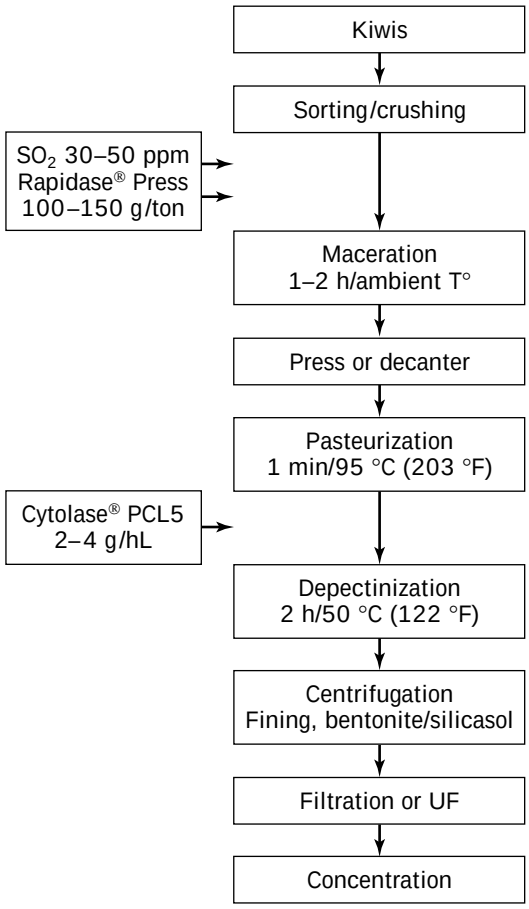


Figure 22. Kiwi juice process.

Process Description. Fruits are inspected, washed, brushed, and rinsed before crushing (Fig. 23). Pulp extraction is carried out on specific extractors with a desteming system. The pulp is heated up to 95 °C (203 °F) for 2 min to inactivate endogenous PPO and is cooled down at 50 °C (122 °C). Enzymatic maceration is carried out with 100 to 200 g/ton Rapidase® Tropical Cloud (Gist-brocades) for 1 to 2 h. Juice is extracted with a pulp finisher, centrifuged, and pasteurized at 85 °C minimum (185 °F) for 30 to 45 sec and concentrated up to 55 °Brix.

Orange, Grapefruit, and Tangerine

The commercial production of citrus for processing into fruit sections and juice began in the 1920s in Florida. The development of frozen concentrated orange juice (FCOJ) started in the 1940s and now exceeds all other citrus products combined, and it is the largest single frozen food item. Although citrus juice is the most important product, the by-products, of which pectin and essential oils are the most significant, provide added value for producers. Although some countries forbid the use of enzymes in the production of premium citrus juices from orange, tangerine, and grapefruit, enzymes can be used to increase the pulp wash yield, to make the concentration easier, to improve essential oil recovery, and to debitter or clarify the juice (lemon

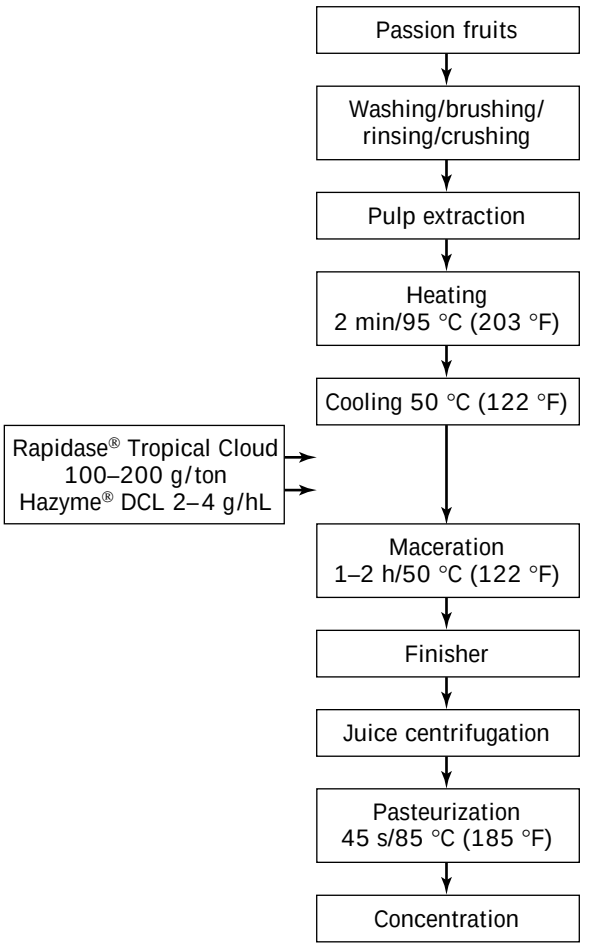


Figure 23. Passion fruit juice process.

juice). After incoming fruits are sampled and tested (the ripeness standard is the ratio of Brix to acid, defined in the 1985 Florida Citrus Laws), oranges enter the extractor. A schematic diagram of extraction and finishing in citrus juice plant is shown in Figure 24. Methods of extraction and finishing operations affect the yield, quality, and characteristics of the citrus juice produced. To maintain high quality, the yields obtainable by processors in Florida are highly regulated by citrus laws (e.g., size and type of extractors). Officials decide what processors are allowed to recover as pure juice. There are two main types of extractors and finishers, manufactured by FMC and Brown Companies in the United States. The expressed juice from the extractors contains pulp and cell membrane residue, and processing in a finisher (rotating screw principle) is required to reduce the pulp content. Raw juice enters at one end, pulp is squeezed out at the other end, and the juice passes radially through sieving cylindrical-shaped screens. Standard finishing reduces the pulp content of juice to about 12%. The excess pulp and solid by-products of the extracted juice from the primary finishers are processed through a pulp wash system (47).

Pulp Wash

The pulp wash system consists of a series of five to seven stages of secondary finishers. Water is added in a multistage countercurrent washing system. This water, along with a high finisher screen pressure, removes the remaining soluble solids containing a high level of pectin from the fruit pulp. About 5 to 7% of the additional solids can be recovered. After removal of the soluble solids, the product is concentrated. Rapidase® Citrus 2000 is added before or after pulp wash extraction at 2.0 g per 100 kg pulp (or per hL) to reduce the viscosity (mainly from the pectin) for achieving the desired 65 °Brix concentrate (Fig. 24). The result is a better extraction of sugars and soluble solids, resulting in a higher yield and therefore a lower viscosity. The pulp wash product can be easily concentrated to 65 °Brix without risk of gelling and keeping a good cloud stability. The pulp wash product is added back to sodas as the cloud source, stable after centrifugation and pasteurization. After this treatment, the pectin degradation is limited to lowering the viscosity without attacking the insoluble pectin that maintains the stability of the cloud. It is very important to use pectinases with the lowest PME activity as possible, to avoid clarification of the pulp wash. The ideal enzyme is a pure PL. Next to PL, Rapidase® Citrus 2000 contains glycosidases that decrease the pulp wash bitterness and improve its taste, mouthfeel, and flavor level.

Oil Recovery

The essential oil content is 2 to 5 kg/ton of processed raw material, but the oil is mainly located in the fruit peel. Essential oils of citrus are located in special cells of the flavedo. The flavedo contains hydrocarbons (terpenes and sesquiterpenes), oxygenated compounds (aldehydes, esters, alcohols, ketones, phenols), and nonvolatile residues (waxes, coumarins, flavonoids, fatty acids). Although the Brown system requires a separate machine to extract oil before juice extraction, the FMC extractor is capable of extracting juice and peel oil in the same operation.

Process Description. After extraction, the juice is sent to the finisher for juice production, and the peel particles containing the oil are washed in a separate stream. After juice extraction, the peel particles and the oil–water emulsion are separated through a finisher. The proportion of oil in the emulsion is between 1 and 3% (9 lb/ton fruit). Enzymes can speed the process time and increase the yield of oil from the emulsion (cream) and the quality of the final product. Enzymes needed for a better separation of oil and water from the emulsion are pectinases and acid proteases. By hydrolyzing the complex pectin proteins, oil is released more easily from the aqueous phase. Enzymes preparations such as Citrozym CEO (Novo), Rohapect VR (Röhm), or Rapidase® Citrus Oil (Gist-Brocades) are added at 10 to 20 g/hL of emulsion before the desludger centrifugation for 30 to 60 min at room temperature. Oil yield can be increased by 10 to 15%.

The finished emulsion passes through a sand cyclone and is fed to a desludger centrifuge to produce an oil-rich emulsion that is sent to a polisher that recovers the clear oil. After the desludger stage, the oil is concentrated up to 60 to 80%. In the second stage, the polisher removes the remaining water and the very fine solid particles. A cold storage allows dewaxing of the oil for better purification.

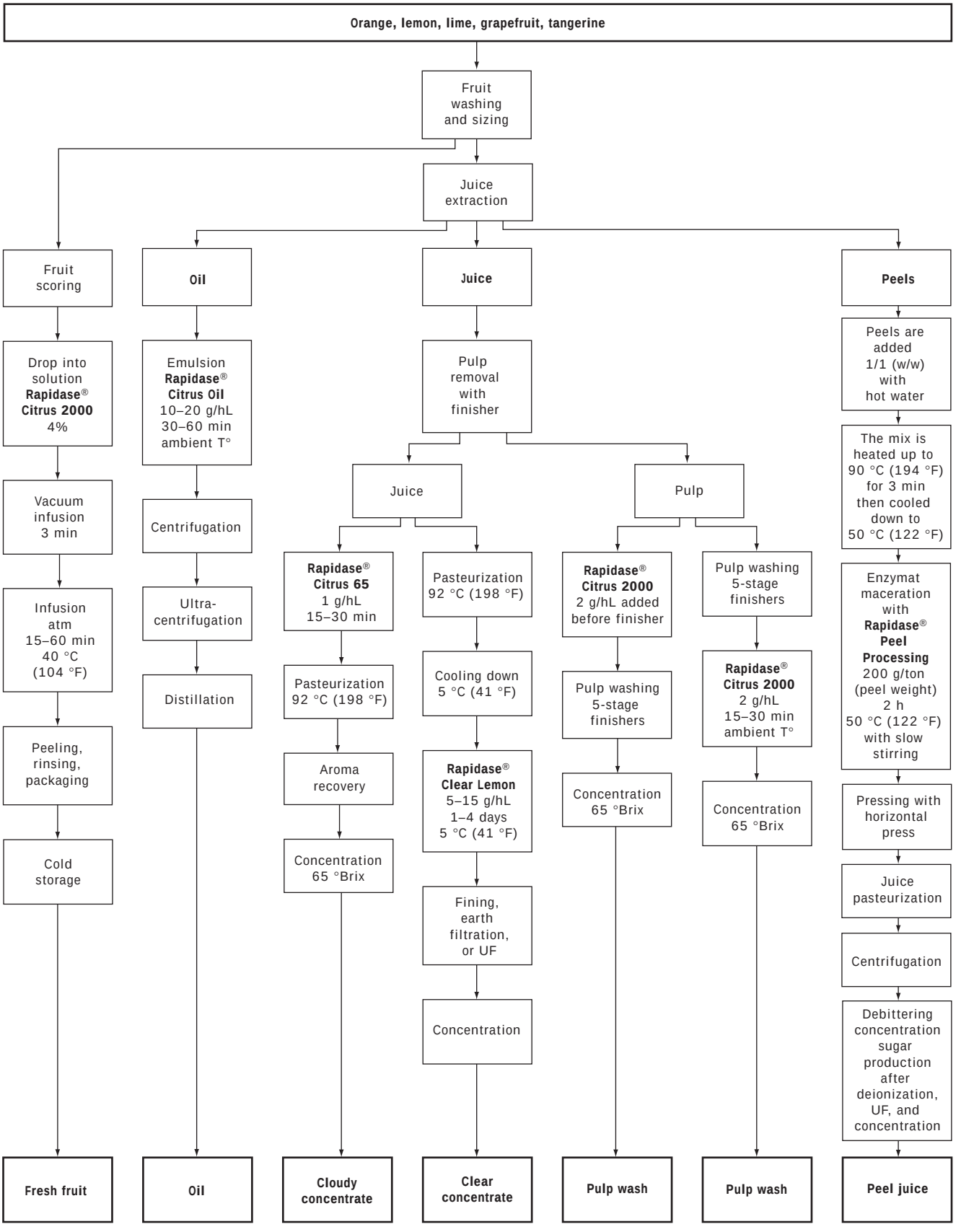


Figure 24. Citrus processing.

Highly Concentrated Citrus Base

Extracted finished juice is sensitive to microbiological contaminations, enzymatic activities, and chemical reactions, all leading to a quick quality degradation. Problems with unpasteurized juices are cloud loss and juice separation caused by citrus PME, gelation of the concentrate, fermentation, and spoilage, even if the pH is low (3.0 to 4.0). These problems are prevented by fast pasteurization, although it causes undesirable browning reactions and flavor and taste changes in the juices before aseptically packaging. Inactivation of citrus PME by heating has been determined to occur at 90 °C (194°F) for 1 min (48). The juice is then concentrated with multieffect evaporators at 65 °Brix. Volatiles recovered during juice concentration are referred to as essences. Concentrate is chilled immediately after evaporation to maintain high flavor quality. Citrus juices can be concentrated after decrease of the juice viscosity by special pectinases in countries where it is authorized, such as Italy. The soluble pectin, which is responsible for many problems in processing, must be slightly degraded to avoid the risk of gelatinization in the concentrator.

Process Description. Added after the first finisher, enzymes such as Rohapect SK1 (Röhm) or Rapidase® Citrus 65 (Gist-Brocades) at 1 g/hL for 15 to 30 min make the concentration of the premium juice easier without clarification and enable the production of concentrated citrus base above 65 °Brix (Fig. 24).

Prepeeling

Citrus fruits are sometimes treated with enzymes that digest white albedo, which binds the flavedo peel to fruit segments. After treatment, hand peeling is easier and faster. Fruit peel is first scored, and fruits are then dropped into an enzyme solution with 4% Peelzym (Novo Nordisk) or Rapidase® Citrus 2000 (Gist-Brocades). Fruits are vacuum infused for 2 to 3 min in this enzyme solution (Fig. 24). After vacuum break, fruits are maintained at 40 °C (104 °F) for 15 to 60 min for albedo digestion inside the fruit (49). Fruits will be later be very easily hand peeled. Fruits are then rinsed and cooled in a water bath at 5 °C (41 °F) and packed.

Peel Pressing

Orange juice yield is low, around 40%. This is why peel valorization is of high interest. After juice extraction, peels are added with same weight of hot water (Fig. 24). The blend is heated up to 90 °C (194 °F) for 3 min and then cooled down to 50 °C (122 °F). Enzymes such as Pectinex UltraSP (Novo Nordisk) or Rapidase® Peel Pressing (Gist-Brocades) are added at 200 g/ton for 2 h at 50 °C (112 °F) with slow stirring. Pressing is carried out with a horizontal press or decanter. The extracted juice is pasteurized, centrifuged, possibly debittered, and concentrated.

Lemon

Lemon juice is extracted with FMC or Brown extractors. Before concentration, seeds, bits of peel, and pulp must be removed (Fig. 25). Lemon pulp, which is added back to lemonade, is recovered by hydrocyclone or cylindrical perforated screen. Primary oil emulsion discharged from extractors still contains peel fragments that must be discarded by finishers before essential oil is recovered from the emulsion by centrifugation. Cold-pressed lemon oil has a large range of applications, and the market value of lemon oil is very high. There is a strong demand for lemon oil as flavoring in foods and soft drinks and in cosmetics, pharmaceuticals, and household products. Citral is the predominant aldehyde in lemon oil and has long been considered the chief flavor constituent of the product. Lemon juice contains a complex colloidal system of particles composing the cloud. These particles, remaining in suspension after the largest ones have been removed, play a major role in the organoleptic qualities of the juice. Lemon cloud is composed of approximatively equal amounts of proteins and pectin with flavonoids and phospholipids (50). Lemons can be processed as cloudy juice. In this case, it is necessary to destroy the endogenous PME, which could lead to a cloud destabilization. The temperature of lemon PME inactivation is 74 °C (165.2 °F) for 14 s. Heating the pulpy lemon juice for too long may result in browning. The juice is then concentrated by evaporation or reverse osmosis at 55 to 65 °Brix. Demand for clarified lemon juice concentrate has increased in recent years. Clarification of single-strength lemon juice may be carried out directly by UF or addition of polygalacturonic acid or commercial pectinases active at low pH. Lemon juice is not easily clarified and concentrated.

In the past, juice clarification took as long as 4 to 16 weeks and occurred only after the addition of huge amounts of sulfur dioxide (about 1,000 to 2,000 mg/L) to

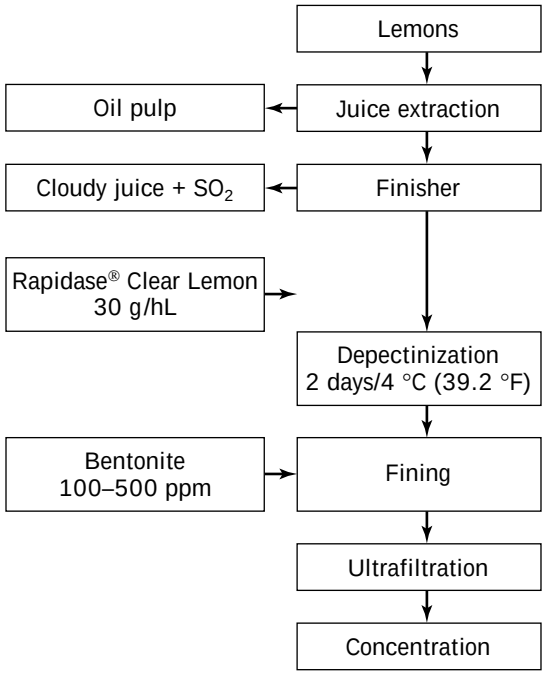


Figure 25. Lemon juice process.

aid the action of lemon PME. Depectinization requires pectinases to be stable at low pH (2.2 to 2.8) with high citric acid concentration. Moreover, pectinases alone did not allow juice clarification because the cloud is caused by pectin, hemicelluloses, and cellulose, with a high protein content.

Process Description. Enzymes proposed for lemon juice production are Rapidase® Clear Lemon (Gist-Brocades), Rohapect C (Röhm), or Citrozym (Novo). After lemon juice extraction, the juice goes through a finisher and is then depectinized (Fig. 25). Rapidase® Clear Lemon is added at 30 g/hL to the juice at low temperature, 4 to 8 °C (39.2 to 46.4 °F), to avoid juice oxidation. Lemon albedo contains about 33% of its dry weight as pectin. The complete degradation of pectin that is required is finished after about 2 days at low temperature or a few hours at 40 °C (104 °F). However, even if the juice is depectinized, it is not clarified. At this stage, it is still possible to produce a cloudy concentrate or a clear juice and concentrate (Fig. 24). A fining step is necessary for the clarification, with high dosis of bentonite (about 100 g/hL or 500 g/hL without enzyme treatment) or silicasol. Then, flocculation and sedimentation of the cloud can occur. After a few hours, the juice is centrifuged, ultrafiltered, and concentrated easily up to 52 °Brix.

FRUIT JUICE AUTHENTIFICATION

Today, quality, authenticity, and detection of adulteration of fruit juices are absolute requirements from juice purchasers and consumers. The evolution of technology and processes using enzymes in parallel with more sensitive analytical techniques, especially in the field of liquid and gas chromatography, and new analytical methods such as SNIF-NMR reveal a broad spectrum of components in marketed fruit juices. It is sometimes difficult to identify these different compounds and precisely determine their origin. Most of them result from the physical and enzymatic degradation of natural substrates released during the fruit processing, but deviations from the standard ranges may indicate adulteration, and expert interpretation is required. In Europe, a code of practice has been developed to maintain quality and authenticity standards. Members of Association of the Industry of Juices and Nectars (AIJN) have established reference guidelines to judge the quality and authenticity of fruit juices (51). The values should reflect the commercial standards. The most widely used of these are the German Fruit Juice Industry guide values known as RSK values, introduced in the 1970s. Fruit juice authenticity is assessed by taking into account numerous chemical analyses. Apple has been extensively studied, but products coming from the degradation of cell wall polymers by enzymes are not yet fully described, which creates problems in juice authenticity: what is a "standard" apple juice and what is the reference process? At the same time, the development of enzymatic liquefaction technology with pectinases, hemicellulases, and cellulases and the development of new analytical techniques with a very high sensitivity show new and atypical analytical profiles of juices, revealing more numerous chromatography peaks caused by sugar oligomers that are not easily identified and from which the origin is not obvious.

For example, starch, pectin, and cellulose are polysaccharides naturally present in most of fruits. Granular starch is mainly present at the crop period in immature apples. It is hydrolyzed during fruit ripening by endogenous amylases. Apple starch is composed of 70% amylopectin and 30% amylose. Starch degradation results in the formation of simple sugars such as glucose, maltose, and maltitol. Because endogenous enzyme activity is very low and the process duration is short, it is necessary to add high exogenous fungal amylases that release glucose, maltose, and other isomers such as isomaltose and pannose. We assume that side enzyme activities, which are present in commercial amylase preparations such as maltase and α-glucosidase, can synthesize these compounds. Pectin is the main substrate being degraded by exogenous enzymes during processing. Pectinases from *A. niger* release all sizes of oligomers and monomers of galacturonic acid, arabinose, and galactose and also parts of nondegraded RG-I and RG-II. Analyses of juice composition with new techniques can now reveal all these different compounds. With the purpose of optimizing fruit sugar extraction, new technologies such as fruit liquefaction (total or pomace) have appeared. The related new processes increase the juice yield highly. The yield is high because of the almost total hydrolysis of protopectin in oligomers and monomers of galacturonic acid and neutral sugars, but also because of partial hydrolysis of cellulose into oligomers and monomers of glucose, including cellobiose. Cellulases, by decreasing the mash viscosity, improve pectinases efficiency. The sugar profile depends also on physical conditions of the process submitted to the fruit. One sugar molecule can easily epimerize at a slightly alkaline pH of 8.0–8.5. This can be the case during fruit pulp or pomace heating with steam condensate water being at alkaline pH. The contact between glucose in the fruit pulp and alkaline water can form fructose, thus changing the ratio of glucose to fructose that is a criterion of fruit juice authenticity. A high temperature could also change the sugar profile in the final juice.

In conclusion, numerous laboratories of authentification have recently appeared that can detect the adulteration of juices with cheap sugars coming from cereals. However, it is urgent now that analysts at these laboratories enlarge the description of standard juice composition, including the following parameters: the fruit variety, the geographical area where raw material originates, climatic conditions of the year, fruit ripeness (endogenous enzymes), and all process conditions. Communication among analysts, enzyme producers, fruit juice producer organisations, and legislators is necessary, and studies are ongoing. The role of exogenous enzymes that are necessary to produce clear and stable concentrate is crucial in determining the final composition of the juice. All these parameters should be integrated in the description of new standards in order to avoid confusion in the interpretation of results.

BIBLIOGRAPHY

1. A. Mehlitz, *Biochem. Z.* **221**, 217–231 (1930).
2. J.G. Faigh, *Food Technol.* **49**(9), 79–83 (1995).
3. M.C. MacCann, and K. Roberts, *Agro-Food Industry Hi-Tech*. 43–46 (March–April 1994).
4. N.C. Carpita and D.M. Gibeaut, *Plant J.* **3**, 1–30 (1993).
5. P. Albersheim, A.G. Darvill, M.A. O'Neill, H.A. Schols, and A.G.J. Voragen, in J. Visser and A.G.J. Voragen ed., *Pectin and Pectinases*, Elsevier Science, Amsterdam, 1996, pp. 47–55.
6. J. De Vries, A. Voragen, F. Rombouts, and W. Pilnik, in M. Fishman and J. Jen ed., *Chemistry and Function of Pectins*, ACS Symposium Series 310, Am. Chem. Soc., Washington, D.C., 1986, pp. 38–48.
7. H. Schols, M.A. Posthumus, and A. Voragen, *Carbohydr. Res.* **206**, 117–129 (1990).
8. P. Dey and K. Brinson, *Adv. Carbohydr. Chem. Biochem.* **42**, 265–282 (1984).
9. F. Will, Ph.D. Thesis, University of Giessen, Germany, 1990.
10. P. Pellerin, S. Vidal, P. Williams, and J.M. Brillouet, *Carbohydrate Res.* **277**, 135–143 (1995).
11. T. Doco, P. Williams, S. Vidal, and P. Pellerin, *Carbohydr. Res.* **297**, 181–186 (1997).
12. F. Barnoud, in B. Monties ed., *Les Polymères Végétaux*, Gauthier-Villars, Paris, 1980, pp. 66–86.
13. J.P. Vincken, G. Beldman, and A. Voragen, *Plant Physiol.* **104**, 99–107 (1994).
14. L. Rexova-Benkova and O. Markovic, *Adv. Carbohydr. Chem. Biochem.* **33**, 323–385 (1976).
15. A. Voragen, Ph.D. Thesis, Wageningen University, The Netherlands, 1972.
16. C. Versteeg, Ph.D. Thesis, Wageningen University, The Netherlands, 1979.
17. J.A.E. Benen, H.C.M. Kester, Lucie Parenicova, and J. Visser, in J. Visser and A.G.J. Voragen ed., *Pectin and Pectinases*, Elsevier Science, Amsterdam, 1996, pp. 221–230.
18. G. Beldman, G. Mutter, M.J.F. Searle-van-leeuwen, L.A.M. Van den Broek, H.A. Schols, and A.G.J. Voragen, in J. Visser and A.G.J. Voragen ed., *Pectin and Pectinases*, Elsevier Science, Amsterdam, 1996, pp. 231–245.
19. F.M. Rombouts, A. Voragen, M. Searle-van-leeuwen, C. Gereads, H. Schols, and W. Pilnik, *Carbohydr. Polym.* **9**, 5–47 (1988).
20. P. Ducroo, *Flüssiges Obst.* **5**, 265–269 (1987).
21. H. Schols, P. Veld, W. Van Deelen, and A. Voragen, *Z. Lebensm.-Unters.-Forsch.* **192**, 142–148 (1991).
22. L. Saulnier, J.M. Brillouet, M. Moutounet, C. Herve du Penhoat, and V. Michon, *Carbohydr. Res.* **224**, 19–235 (1992).
23. F. Kormelink, M. Searle-van-Leeuwen, T. Wood, and A. Voragen, *J. Biotechnol.* **27**, 249–265 (1993).
24. G. Beldman, M. Searle-van-Leeuwen, F.M. Rombouts, and A. Voragen, *Eur. J. Biochem.* **146**, 301–308 (1985).
25. A. Kliosov and M. Rabinovitch, in L. Wingard, J. Berezin, and A. Kliosov ed. *Enzyme Engineering: Future Directions*, Plenum Press, New York, 1980, pp. 82–165.
26. C. Renard, M. Searle-van-Leeuwen, A. Voragen, J.F. Thibault, and W. Pilnik, *Carbohydr. Polym.* **14**, 295–314 (1991).
27. W. Fogarty and C. Kelly, in A. Rose ed., *Economic Microbiology*, vol. 5, Academic Press, New York, 1980, pp. 115–170.
28. C. Grassin and P. Fauquembergue, in Godfrey T. and S. West ed., *Industrial Enzymology*, 2nd ed., MacMillan, London, 1995, pp. 227–264.
29. M. Yamasaki, T. Yasui, and K. Arima, *Agric. Biol. Chem.* **28**, 779–787 (1964).
30. Q. Wu, M. Szabacs-Dobozi, M. Hemmat, and G. Hrazdina, *Plant Physiol.* **102**, 219–225 (1993).
31. I. Bartley and M. Knee, *Food Chem.* **9**, 47–58 (1982).
32. C. Grassin, *Flüssiges Obst.* **57**, 519–522 (1992).
33. C. Grassin and P. Fauquemberque, *Fruit Process.* **7**, 242–245 (1993).
34. A.G.J. Voragen, R. Heuting, and W. Pilnik, *J. Appl. Biochem.* 452–468 (1980).
35. C. Grassin and P. Fauquembergue, *Fruit Process.* **12**, 2–7 (1996).
36. J.F. Drilleau, *IAA* 47–58 (Sept. 1985).
37. European Pat. Specification EP 0547,648,B1 (November 6, 1992), C. Grassin and P. Fauquembergue (to Gist-Brocades).
38. A. Ranadive and N. Haard, *J. Food Sci.* **38**, 331–333 (1973).
39. F. Will, G. Hasselbeck, and H. Dietrich, *Flüssiges Obst.* **59**, 352–353 (1992).
40. A. Holmes and M. Starr, in S. Nagy, C.S. Chen, and P.E. Shaw eds., *Fruit Juice Processing Technology*, Agscience Inc., Auburndale, Fla., 1993, pp. 515–531.
41. C. Pederson, in P. Nelson and D. Tresslier ed., *Fruit and Vegetable Juice Processing Technology*, AVI Publishing, Westport, Conn., 1980, pp. 268–309.
42. H. Siliha, Ph.D. Thesis, Wageningen University, The Netherlands, 1985.
43. O. Marcellin, P. Williams, and J.M. Brillouet, *Carbohydr. Res.* **240**, 233–243 (1993).
44. O. Markovic, K. Heinrichova, and B. Lenkey, *Collect. Czech. Chem. Commun.* **40**, 769 (1976).
45. E. Koffi, C. Sims, and R. Bates, *J. Food Quality* **14**, 209–218 (1991).
46. K. Chen Chin and H. Yamamoto, *J. Food Sci.* **43**, 1261–1263 (1978).
47. S. Chen, P. Shaw, and M. Parish, in Nagy, Chen, and Shaw ed., *Fruit Juice Processing Technology*, Agscience, Auburndale, Fla., 1993, pp. 110–165.
48. B. Eagerman and A. Rouse, *J. Food Sci.* **41**, 1396–1397 (1976).
49. U.S. Pat. 5,200,217 (April 6, 1993), R. Elliot and J. Tinibel (to Sunkist Growers).
50. B. Carter, in Nagy, Chen, and Shaw ed., *Fruit Juice Processing Technology*, Agscience, Auburndale, Fla., 1993, pp. 215–270.
51. G.G. Martin, E. Jamin, J. Gonzales, G. Remaud, V. Hanote, P. Ströber, and N. Naulet, *Fruit Process.* **9**, 344–349 1997.

See also ENZYMES, PROTEIN HYDROLYSIS; FOOD PROCESS ENGINEERING.

ENZYMES, IMMOBILIZATION METHODS

TADASHI SATO
TETSUYA TOSA
Tanabe Seiyaku Co., Ltd.
Osaka, Japan

KEY WORDS

Carrier binding
Cross-linking
Entrapping enzyme
Immobilization method
Immobilized enzyme

OUTLINE

In 1916, Nelson and Griffin (1) found that an enzyme in water-insoluble form showed catalytic activity. They reported that invertase extracted from yeast was adsorbed on charcoal, and the adsorbed enzyme showed the same activity as native enzyme. In 1948, Sumner (2) found that urease from jack bean became water insoluble on standing in 30% alcohol and sodium chloride for 1 to 2 days at room temperature and that the water-insoluble urease was active. Thus, it has been known for a long time that enzymes in water-insoluble form show catalytic activity. However, the first attempt to immobilize an enzyme to improve its properties for a particular application was not made until 1953, when Grubhofer and Schleith (3) immobilized enzymes such as carboxypeptidase, diastase, pepsin, and ribonuclease by using diazotized polyaminopolystyrene resin. On the other hand, before this (1949), immobilization of physiologically active protein was carried out by Micheel and Ewers (4). Further, in 1951, Campbell et al. (5) prepared immobilized antigen by binding albumin to diazotized *p*-aminobenzylcellulose. Subsequently, a number of reports on the preparation and application of immobilized antigens and antibodies appeared. Such studies may also be applicable to immobilized enzymes.

In the 1960s, many papers appeared; in particular, Professor Katzir-Katchalski and his co-workers at the Weizmann Institute of Science, Israel, carried out extensive studies on new immobilization techniques, and on the enzymatic, physical, and chemical properties of immobilized enzymes.

In 1969, Dr. Chibata et al. at Tanabe Seiyaku Co., Ltd., in Japan succeeded in the industrialization of a continuous process for the optical resolution of DL-amino acids using immobilized aminoacylase (6). This was the first industrial application of immobilized enzymes in the world.

In 1971, the first Enzyme Engineering Conference was held at Henniker, New Hampshire, and the predominant theme of this conference was immobilized enzymes. A definition and classification of immobilized enzymes were proposed at the conference. This conference is biannual, and the main topics have continued to be immobilized enzymes.

Some books on immobilized enzymes have been published (7,8) as well.

Immobilized enzymes are defined as "enzymes physically confined or localized in a certain defined region of space with retention of their catalytic activities, and which can be used repeatedly and continuously." Accordingly, enzymes modified to a water-insoluble form by suitable techniques satisfy this definition of immobilized enzymes. In addition, when an enzyme reaction using a substrate of high molecular weight is carried out in a reactor equipped with a semipermeable ultrafiltration membrane without leakage of the enzyme from the reactor, this can also be considered as a kind of immobilized enzyme system.

The term *immobilized enzyme* was recommended at the First Enzyme Engineering Conference in 1971. Before that time, various terms such as water-insoluble enzyme, trapped enzyme, fixed enzyme, and matrix-supported enzyme had been used.

As functional groups involved in enzyme immobilization, free, amino and carboxyl groups, the sulfhydryl group of cystein, the imidazole group of histidine, phenolic groups, and hydroxyl groups of serine and threonine may be considered. For immobilization of an enzyme, it is necessary that functional groups in the active center should not be involved in the reaction leading to immobilization of the enzyme. Further, because the tertiary structure of enzyme protein is maintained by relatively weak binding forces, such as hydrogen, hydrophobic, and ionic bonds, it is necessary to carry out the immobilization reaction under mild conditions.

Methods for enzyme immobilization can be classified into three basic categories as follows.

1. Carrier-binding method: the binding of enzymes to water-insoluble carriers such as polysaccharide derivatives, synthetic polymers, and porous glass, etc.
2. Cross-linking method: intermolecular cross-linking of enzymes by means of bifunctional or multifunctional reagents such as glutaraldehyde, bisdiazobenzidine, and hexamethylene diisocyanate, etc.
3. Entrapping method: incorporating enzymes into the lattice of a semipermeable gel on enclosing the enzymes in a semipermeable polymer membrane, such as collagen, gelatin cellulose triacetate polyacrylamide, and κ-carrageenan, etc.

These methods are shown schematically in Figure 1.

The carrier-binding method is the oldest immobilization method for enzymes, and many papers have been published on this method. When enzymes are immobilized in this way, care is required regarding the selection of carriers as well as in binding techniques. The carrier-binding method can be further divided into three categories according to the binding mode of the enzyme, that is, physical adsorption, ionic binding, and covalent binding.

The cross-linking method is based on the formation of chemical bonds, but water-insoluble carriers are not used in this method. The immobilization of enzymes is performed by the formation of intermolecular cross-linkages between the enzyme molecules by means of bifunctional or multifunctional reagents.

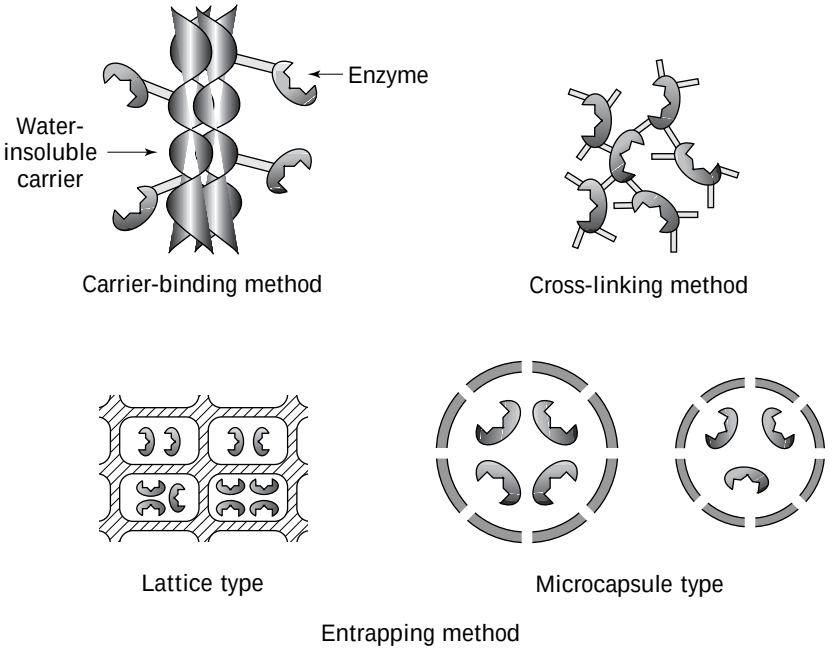


Figure 1. Schematic diagrams of immobilized enzymes.

The entrapping method is based on confining enzymes in the lattice of a polymer matrix or enclosing enzymes in semipermeable membranes, and it can be classified into the lattice and microcapsule types. This method differs from the covalent binding and cross-linking method in that the enzyme itself does not bind to the gel matrix or membrane. Thus, this method may have wide applicability. However, if a chemical polymerization reaction is used for entrapping, relatively severe conditions are required and lose of enzyme activity occurs in some cases. Therefore, it is necessary to select the most suitable conditions for the immobilization of various enzymes.

Preparation methods and characteristics of immobilized enzymes are summarized broadly in Table 1, though there are many exceptions.

Immobilization of enzymes by covalent binding or cross-linking is carried out under relatively severe conditions in comparison with those of physical adsorption or ionic binding. Thus, in the former cases, conformational change of the enzyme structure and partial destruction of the active center may occur. Accordingly, unless immobilization of an enzyme by covalent binding is carried out under well-controlled conditions, immobilized enzyme having high activity cannot be obtained. However, the binding forces between the enzyme and carrier are strong, and the enzyme cannot easily be lost from carriers even in the presence of substrates or salts at high concentrations. However, when the activity of enzymes immobilized by covalent binding decreases on long-term operation, regeneration is impossible.

On the other hand, immobilization of enzymes by the ionic-binding method can be achieved simply under mild conditions. Accordingly, in this case, preparations having relatively high activity are obtained. However, the binding forces between enzyme and carrier are weak in comparison with those in the covalent-binding methods. Therefore, leakage of the enzyme from the carrier may occur after changes in the ionic strength, pH of the substrate, or product solution. This type of immobilized enzyme can be regenerated when the enzyme activity decreases after prolonged operation. Thus, the ionic-binding method is advantageous in comparison with the covalent-binding method, particularly when expensive carriers or enzymes are used.

Unlike the ionic-binding, covalent-binding, and cross-linking methods in the entrapping method, no binding

Table 1. Preparation and Characteristics of Immobilized Enzyme

	Carrier-binding method				
Characteristic	Physical adsorption	Ionic binding	Covalent binding	Cross-linking method	Entrapping method
Preparation	Easy	Easy	Difficult	Difficult	Difficult
Enzyme activity	Low	High	High	Moderate	High
Substrate specificity	Unchangeable	Unchangeable	Changeable	Changeable	Unchangeable
Binding force	Weak	Moderate	Strong	Strong	Strong
Regeneration	Possible	Possible	Impossible	Impossible	Impossible
General applicability	Low	Moderate	Moderate	Low	High
Cost of immobilization	Low	Low	High	Moderate	Low

between enzyme and carrier should occur in theory. Therefore, in many cases, preparations having high activity are obtained. However, in this method, regeneration of activity losses is impossible, as in the case of the covalent-binding method. A major disadvantage of the entrapping method is that its application is limited to small molecular substrate and product molecules; the entrapped enzyme shows little or no activity toward macromolecular substrates.

Although a number of enzyme immobilization methods have been studied, no ideal general methods applicable for the immobilization of many enzymes have yet been developed. Each method has specific disadvantages. Therefore, in practice, it is necessary to find a suitable method and conditions for the immobilization of a particular enzyme in the light of the intended application.

Because of the special characteristics of immobilized enzymes, a considerable amount of work has been carried out on application of immobilized enzymes as solid catalysts having high substrate specificity. Immobilized enzymes are used in various fields, such as chemical processes, analysis, medical treatment, food processing, affinity chromatography, and so forth. For the efficient utilization of such immobilized systems, chemical engineering studies are necessary.

BIBLIOGRAPHY

1. J.M. Nelson and E.G. Griffin, *J. Am. Chem. Soc.* **38**, 1109–1115 (1916).
2. J.B. Sumner, *Science* **108**, 410–413 (1948).
3. N. Grubhofer and L. Schleith, *Naturwissenschaften* **40**, 508 (1953).
4. F. Micheel and J. Ewers, *Makromol. Chem.* **3**, 200–209 (1949).
5. D.H. Campbell, E. Luescher, and L.S. Lerman, *Proc. Nat. Acad. Sci.* **37**, 575–581 (1951).
6. I. Chibata, T. Tosa, T. Sato, T. Mori, and Y. Matuo, *Proc. of the 4th Int. Fermentation Symp.: Fermentation Technology Today*, 1972, p. 383–389.
7. L.B. Wingard Jr. ed., *Enzyme Engineering*, Wiley, New York, 1972.
8. O.R. Zaborsky, *Immobilized Enzymes*, Chemical Rubber Co. Press, Ohio, 1973.

See also BIOFILMS, MICROBIAL; BIOREACTORS, AIR-LIFT REACTORS; BIOREACTORS, FLUIDIZED-BED; BIOREMEDIATION.

ENZYMES, IMMOBILIZED, REACTORS

UDO KRAGL
LASSE GREINER
CHRISTIAN WANDREY
Forschungszentrum Jülich GmbH
Jülich, Germany

KEY WORDS

Amino acylase
Enzyme, immobilized
Glucose-fructose isomerase
Glycerol kinase
Lipase
Mass transport limitation
Membrane
Nitrile hydratase
PenG acylase
Reactor

OUTLINE

INTRODUCTION

The use of enzymes in food as well as in pharmaceutical and chemical industries has increased steadily since the early 1970s. Applications range from synthesis in the milligram scale for natural or nonnatural oligosaccharides up to industrial synthesis for the production of bulk chemicals such as acrylamide. Some examples are listed in Table 1. This has been well documented by a large number of textbooks, reviews, and original papers (1,5–20). This article highlights some of the most important aspects for the application of immobilized enzymes and presents selected examples.

The most common definition for immobilized enzymes is that proposed by Katchalski-Katzir in the 1960s (16): "Enzymes physically confined or localized in a certain defined region of space with retention of their catalytic activities, which can be used repeatedly and continuously." According to this definition, three types of immobilized enzymes can be distinguished:

Table 1. Examples of the Industrial Use of Immobilized Enzymes

Enzyme (ref.)	Substrate	Product	Scale (ton a^{-1})
Penicillin amidase (1,2,3)	Penicillin G	6-Amino penicillanic acid	6,000
Cephalosporin amidase (1)	Glutaryl-7-ACA	7-ACA	?
Amino acylase (1)	Acyl-D-L-amino acid	L-amino acid	approx. 300
Aspartase (1)	Fumaric acid	L-aspartic acid	1,200
Fumarase (1)	Fumaric acid	L-malic acid	360
Lactase (1) (galactosidase)	Lactose GalGlc	Low lactose milk Gal + Glc	?
Aspartase β-decarboxylase (1)	Aspartic acid	L-alanine	120
Nitrile hydratase (H. Yamada, personal communication)	Acrylonitrile	Acrylamide	30,000
Glucose isomerase (1)	Glucose	High fructose corn syrup	8,000,000
Lipase (4) (expected)	Rac-1-phenylethyl-amine	(*S*)-1-phenylethyl-amine	200

1. Heterogenization of the soluble enzyme by coupling to an insoluble support by adsorption or covalent binding, by cross-linking of the enzyme or entrapment in a lattice or in microcapsules such as alginate beads
2. Retention of the enzyme by means of ultrafiltration membranes
3. Use of whole cells for biotransformations using their enzyme apparatus

Because of the scope of this article, only the first type is considered. Membrane reactors have been reviewed quite recently (21–28), as have applications of whole cell processes (1,5,13,14,17,29–33). Retention of the cells within the reactor may be achieved by membrane separation or by the same immobilization methods that are used for isolated enzymes (34). In principle, the cell itself can be regarded as a form of native immobilization of enzymes. Biosensors are a very special form of carrier-fixed biocatalysts, for which the reader is referred to the literature (35–37).

The major goal behind immobilization (in its most common definition) is the recovery of the biocatalysts, separation from products and reactants, and subsequent reuse in batch or continuous processes. The latter is especially important for a reduction of the catalyst costs. Enzymes immobilized on a support often show enhanced stability when compared with the soluble form (38–41). When considering the use of soluble or carrier-fixed enzymes, the following topics have to be addressed:

1. Additional costs for support and chemicals performing the immobilization have to be balanced against the increase of stability.
2. Loss of activity during the immobilization step.
3. When the catalyst is immobilized only by adsorption or entrapment without covalent attachment, its leakage from the carrier support has to be examined and compared with the overall deactivation rate (42).
4. Mass transfer limitations for enzymes on a support may cause problems when adjusting of the pH is necessary during the reaction.
5. With soluble enzymes, higher volumetric activities at high catalyst concentrations are possible, enabling conversion of poor substrates at reasonable rates.
6. Whereas membrane reactors can be easily sterilized before use, this is not possible for reactors with carrier-fixed enzymes. To prevent microbial contamination, these processes are quite often operated at higher temperatures. These are discussed in more detail later.
7. When enzymes are to be used together with organic solvents to increase reactant or product solubility or to alter their kinetics, it may become necessary to immobilize them on a support (8,18,43–48). The support will at the same time act as a water pool to maintain the enzymatic activity (49,50). In such systems, water-insoluble organic solvents have less effect on the enzyme stability than water-soluble solvents.

Table 2 summarizes some arguments in favor or against immobilization of biocatalysts, namely enzymes. The final decision for a certain reactor design should be based on an optimization process covering all relevant factors contributing to the overall costs, including investment, catalyst consumption, or productivity.

Reactors for Immobilized Enzymes

The methods for the heterogenisation (or localization) of enzymes by coupling them to insoluble supports or by entrapment have been discussed in the literature

Table 2. Comparison of Processes Using Soluble or Carrier-Fixed Enzymes

Enzyme fixed on a carrier or membrane	Soluble enzyme retained by a membrane
Continuous or batch process	Continuous or batch process
Improved stability	Lower enzyme stability
Loss of activity during immobilization	No loss of activity during immobilization
Immobilization costs	No costs for immobilization
Mass transfer limitations (heterogeneous reaction)	No mass transfer limitation (homogeneous reaction)
Established technology[a]	New technology[b]
	Sterile process possible
Higher investment costs	Easy supply of new enzyme

[a]Immobilization protocol may be different for each enzyme.
[b]Can be used for different enzymes without modification.

(5,15,34,51,52). The types of reactors used for immobilized enzymes are summarized in Figure 1. The principles developed for general heterogeneous catalysis in synthetic chemistry are valid, resulting in well-known reactor configurations. Differences between enzyme catalysis and other systems result from the nature of the biocatalyst and reaction medium. For example, soft particles containing the biocatalyst, such as alginate beads, may limit the pressure drop in fixed-bed reactors.

The decision as to specific reactor design will be based on a careful analysis of the kinetic properties of the reaction system (53–55). For example, if the enzyme shows a strong substrate-surplus inhibition, a continuously operated reactor with complete backmixing working at high conversion is advantageaous. A reaction with strong product inhibition may utilize a batch reactor or a plug flow reactor to achieve higher volume and catalyst specific productivities. An extractive bioreactor may be used if substrates and products show different solubilities (Fig. 1**i** and **j**) (see "Selected Examples" section). By using this reactor configuration, the destabilizing effect of organic solvents may also be overcome, because the enzyme is separated from the organic phase, which is used to extract the insoluble product (56,57). The aqueous phase containing the enzyme will be saturated until the maximum solubility of with the substrate is reached.

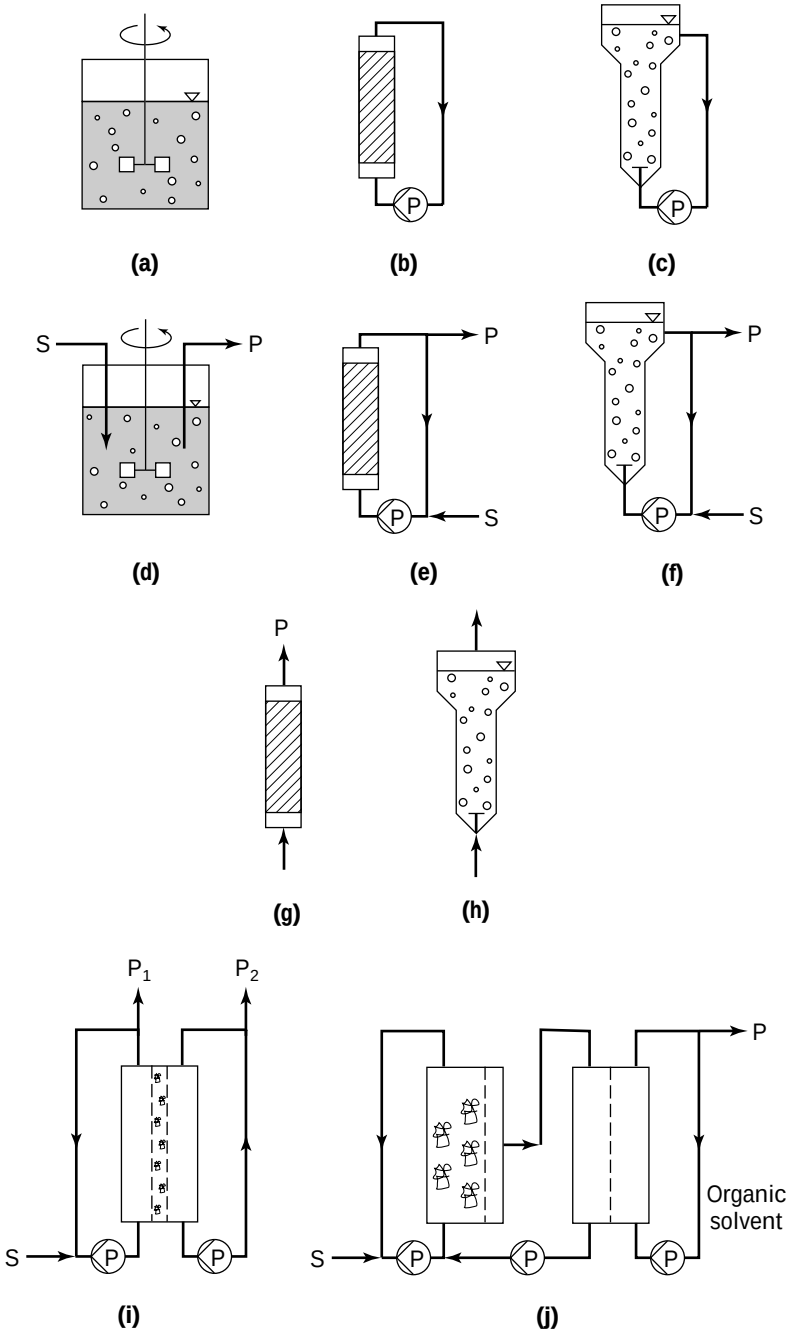


Figure 1. Reactors for immobilized enzymes. (**a–c**) Batch reactors with complete backmixing; (**a**) Stirred-tank reactor; (**b**) Fixed-bed reactor; (**c**) Fluidized-bed reactor. (**d–f**) Continuously operated reactors with complete backmixing. (**g–h**) Continuously operated reactors with plug-flow behavior. (**i**) Reactor with the enzyme immobilized in or on a membrane that may at the same time separate two phases such as water and organic solvent. (**j**) reactor with physically separated enzyme and organic solvent in order to prevent denaturation of the protein (the membrane reactor shown may be replaced by one of the types for immobilized enzymes).

Reactions using biocatalysts are normally performed in aqueous solution at temperatures between 10 and 80 °C and at ambient pressure. Due to the inhibition of some enzymes by heavy metals, the materials of construction must not release these elements. Reactors are operated under conditions that prevent microbial contamination. The reactor itself as well as the substrate may be sterilized prior to reaction by using chemical agents (ethanol, formaldehyde, ethylenoxide, Velcorin™) or steam. Ultraviolet rays or γ-rays may be used to sterilize the immobilized enzyme on its support (58). Alternatively, the immobilization may be performed under sterile conditions. Antibacterial agents may be added to the reaction mixture to prevent microbial growth while the reactor is running. In some cases, the reactants may act as sterilants or inhibitors of microbial growth, such as ketones or alcohols. At higher concentrations (more than 500 mmol/L), solutions may become autosterile because of osmotic pressure effects. Industrial processes are often performed at elevated temperatures, above 55 °C, reducing the danger of microbial contamination.

For a constant product quality and reproducibility of downstream processing, the reactor should be operated at constant conversion. To overcome the deactivation per unit of time that shows all biocatalysts as a result of denaturation processes, either the residence time has to be increased or fresh enzyme has to be supplied. The latter is especially easy for soluble enzymes. For carrier-fixed enzymes, a combination of both methods is used, as discussed later.

CHARACTERIZATION OF IMMOBILIZED ENZYMES

An exhaustive treatment of all aspects of immobilized enzymes, including a complete mathematical description, is beyond the scope of this article. This information may be found elsewhere (22,25,53,59–70).

The well-known Michaelis-Menten equation (equation 1) describes the enzymatic activity as function of the substrate concentration for an irreversible one-substrate reaction without inhibition. This idealized situation is rarely found in nature. Equation 2 gives the rate for a reversible reaction A⇔B. Using this form, product inhibition is considered as a competition of A and B for the active site as expressed by the two K_M values. This form might be extended to cover substrate surplus inhibition as well as multisubstrate reactions by adding additional terms in the numerator (53–55). The form of equation 2 resembles the kinetics of heterogeneously catalyzed reactions (Hougen-Watson kinetics) (71).

$$v = c_E \times \frac{A_{max} \times c_A}{K_M + c_A} \quad (1)$$

$$v = c_E \times \frac{\frac{A_{max,A} \times c_A}{K_{M,A}} - \frac{A_{max,B} \times c_B}{K_{M,B}}}{1 + \frac{c_A}{K_{M,A}} + \frac{c_B}{K_{M,B}}} \quad (2)$$

Upon immobilization of the enzyme on a support, the intrinsic properties expressed as kinetic constants A_{max} and K_M may be altered. This may be because of to conformational changes of the protein by the attachment resulting in a variation of the active site. More often, however, there might be a direct influence of the support or the microenvironment (59). For example, the polymer matrix may cause a pH difference between the microenvironment of the immobilized enzyme and the bulk fluid. Thus, the activity as a function of pH is often different for the soluble and the immobilized enzyme. In a similar way, the apparent K_M value of an immobilized enzyme may be altered compared to that of the soluble form. Additional attracting or repulsing interactions of the support matrix with the substrates or products may lead to different values when compared with the native enzyme. A similar problem arises when lipases are used in a heterogeneous reaction system. Because of their natural properties, these enzymes enrich at hydrophopic/hydrophilic interfaces as found in oil-in-water emulsions. If the inhibiting product enriches there, too, the enzyme will experience a very high local product concentration that will drastically alter its activity (72).

In heterogeneous catalysis, mass transport affects control kinetics at the solid–liquid interphase (Fig. 2). The external mass transfer or film diffusion is determined by the velocity difference between the particle and the solution. The influence of this type of mass transfer resistance can be minimized by rapid flow through a fixed bed or fluidized bed reactor or vigorous mixing in a stirred tank reactor. The limiting property will be the mechanical stability of the support used.

The example given in Scheme 1 illustrates the influence of film diffusion (73,74). The enzyme fructose-1,6-bisphosphate aldolase from rabbit muscle catalyzes the reaction between dihydroxyacetone phosphate (**1**) and various aldehydes (**2**) (10). The enzyme has been immobilized on an insoluble support, thereby enhancing its stability and enabling the reaction in a continuously operated slurry reactor (Fig. 1**d**). If the mass transport is only diffusion controlled (relative velocity between particle and medium = 0), the Sherwood number describing the mass transport with spheres is defined as

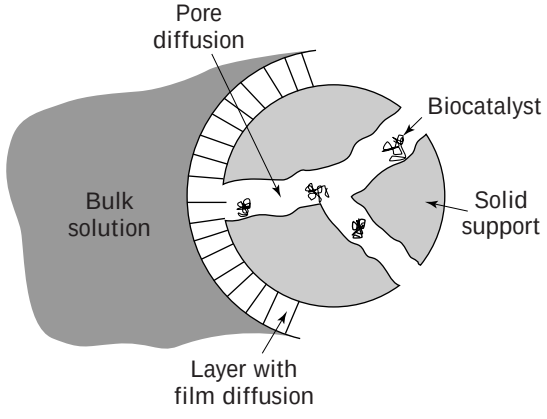
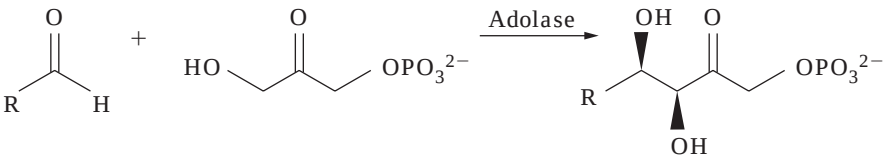


Figure 2. Scheme of a particle with immobilized biocatalyst.



Scheme 1. **1** **2** **3**

$$Sh = \frac{k \times d}{D} = 2 \tag{3}$$

The molar flux of mass transport is given by

$$v_{mt} = k \times (c_0 - c) \tag{4}$$

Assuming Michaelis-Menten-type kinetics, the enzymatic reaction will be rate limiting as long as

$$k > \frac{A_{max,A}}{K_M} \tag{5}$$

For rabbit muscle aldolase bound to Eupergit™ C1Z (nonporous carrier, diameter 1 μm), typical values for $A_{max,A}$ are 50 μmol m^{-2} s^{-1} corresponding to 405.9 nkat g^{-1} (24.3 U g^{-1}). K_M values range between 0.1 and 10 mmol L^{-1}. From equation 3, typical k values of 2×10^{-3} m s^{-1} result. Thus, the mass transport is much faster than the enzymatic reaction, since

$$5 \cdot 10^{-6}\ ms^{-1} < \frac{A_{max,A}}{K_M} < 500 \cdot 10^{-6}\ ms^{-1}$$

If, however, particles with a diameter of 100 μm were used for immobilization, film diffusion would be rate limiting (Fig. 3). The mass-transport resistance might be reduced to a certain extent by establishing a high velocity between particle and solution. The influence of internal mass transfer or pore diffusion becomes more dominant with increasing particle size. This may reduce the effectiveness by substrate limitation or product inhibition in the inner part of the particle, and when the reaction causes a strong pH-shift, local high values may deactivate the enzyme. This may be avoided by higher buffer concentrations. An example is the production of 6-aminopenicillanic acid (see "Selected Examples" section) (75, 76).

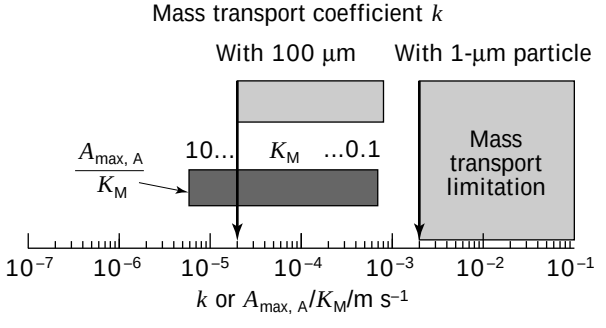


Figure 3. Influence of particle diameter on mass transport in comparison with enzymatic reaction.

In a very simple way the effect of internal mass transfer may be considered as an efficiency coefficient η, which is a function of size, shape, and porosity of the support matrix and the diffusion of the compounds involved:

$$\eta = \frac{\text{activity of immobilized catalyst}}{\text{activity of catalyst without mass transfer}} \tag{6}$$

$$v_{app} = \eta \times v \tag{7}$$

Pore diffusional resistance is indicated when apparent activity rises after grinding the catalyst particles (homogeneous catalyst distribution). If no increase of activity is observed, film diffusion is already limiting. These effects, together with the kinetics of the immobilized catalyst, can be easily investigated in a differential recirculation reactor (62,66). The basic setup is the same as in Figure 1**b**. Only a small amount of catalyst is used to minimize conversion per pass through the bed, as measured in the bypass. By adjusting the flow rate, film diffusion and pore diffusion may be easily distinguished.

For a more elaborate description of pore diffusion, the Thiele modulus, interpreted as the reaction rate relative to the diffusion rate within the particle (59,68–70), has to be used.

The number of catalytically active reaction sites per unit mass of catalyst can be easily determined for immobilized enzymes. For this active-site test, a suitable substrate, such as an ester, is converted by the enzyme and the product is determined analytically (77,78). This test is independent from mass transport phenomena because only the additional amount of liberated product after addition of the catalyst is used for calculation. According to the Michaelis-Menten kinetics of an enzymatic reaction with varying reaction order, the efficiency coefficient η is a function of conversion. This has been demonstrated for immobilized α-chymotrypsin in the cleavage of amino acid esters (Fig. 4) (79). The use of immobilized enzymes in processes with precipitated substrates or products has been addressed recently (80).

SELECTED EXAMPLES OF REACTIONS USING IMMOBILIZED ENZYMES

One of the first enzymes shown to be active after immobilization by adsorption on charcoal was invertase (12,81). This enzyme is still used for the large-scale production of invert sugar by hydrolysis of sucrose. One kilogram of im-

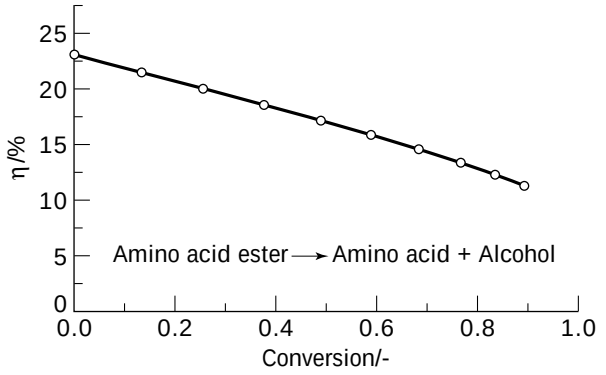


Figure 4. Efficiency coefficient η as function of conversion; initial ester concentration 200 mmol L^{-1}; Thiele modulus 10.1.

mobilized enzyme supports production of approximately 10,000 kg of invert sugar (12).

Production of L-Amino Acids Using Amino Acylase

Enantiomerically pure L-amino acids are required for use in parenteral nutrition, as food additives or as chiral synthons. In addition to isolation from natural sources and direct fermentation, racemic mixtures of L- and D-amino acid can be resolved by enzymatic reaction with one of the two enantiomers. Scheme 2 illustrates this for a racemic amino acid (4) using amino acylase I from *Aspergillus* (EC 3.5.1.14). (*Note*: the EC number, in parentheses, is the internationally recognized Enzyme Commission number (82). The six main classes are oxidoreductases, transferases, hydrolases, lyases, isomerases, and lipases). The unreacted D-enantiomer (**6**) can be racemized easily and recycled to increase the overall yield. The first industrial process of this type was established by Tanabe Seiyaku Co. in 1969 using enzyme immobilized on an insoluble support (1,83).

One of the best methods to obtain a stable catalyst in their example has been to immobilize by ionic binding to DEAE-Sephadex™. To compensate for the loss of activity during operation (20% after 1 month), fresh enzyme has been passed through the column periodically to establish the former activity again. The main products of this process are L-methionine, L-phenylalanine, and L-valine (9). Typically, a 0.2 mol L^{-1} solution of D, L-acetyl amino acids is feed into a 1-m^3 fixed-bed reactor with various flow rates. For L-methionine, the space-time yield reported is about 715 g L^{-1} d^{-1}.

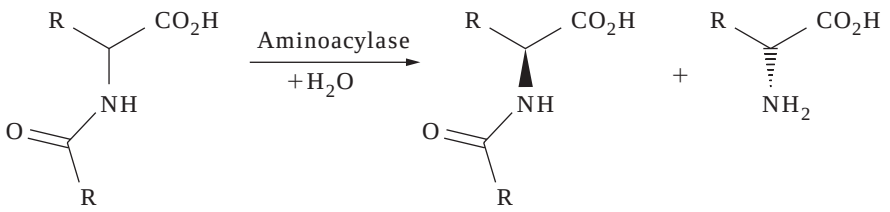
Production of 6-Amino Penicillanic Acid and 7-Amino Cephalosporanic Acid Using Amidases

6-Amino penicillanic acid (6-APA) (**8**) and 7-amino cephalosporanic acid (7-ACA) (**11**) are important intermediates for semisynthetic β-lactam antibiotics. Both compounds can be obtained readily by amidase-catalyzed cleavage of the corresponding amides (**7** and **10**) that are accessible by fermentation (Scheme 3).

Whereas penicillin amidase (EC 3.5.1.11) has been used for production of 6-APA since the early 1970s, cephalosporin C amidase (EC 3.5.1.3) has been used just recently in a chemoenzymatic synthesis of 7-APA (1,84). Amidases have been prepared from several different sources, such as *Pseudomonas*, *Bacillus*, and (recombinant) *Escherichia coli* (1,85,86). In general, the enzyme is immobilized on and in various supports such as polyacrylonitrile fibers, polyacrylamid beads, or Sephadex™. For 6-APA batch production, typically a 10% solution of penicillin G is circulated through the bioreactor or processed in a stirred-tank reactor. One cycle lasts about 3 h. The life time for the enzyme is more than 360 cycles. The space-time yield for this process is about 800 g L^{-1} d^{-1} (1). One of the main problems for both processes is the liberation of a strong acid. To avoid large pH shift, an effective method to control and maintain the desired pH is needed. High recirculation rates are used for the enzyme immobilized on hollow fibers. For the enzyme immobilized in beads, both a small bead diameter (typically around 100 to 600 μm) and good control of film and pore diffusion by high fluid velocities are needed.

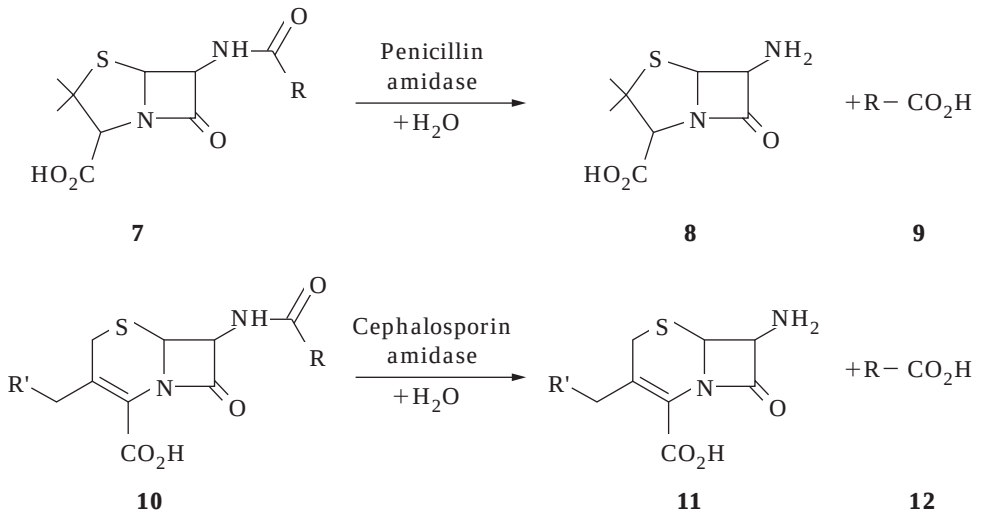
Production of High-Fructose Corn Syrup Using Glucose Isomerase

The production of high-fructose corn syrup, a mixture of glucose (**13**) and fructose (**14**), is the largest enzymatic process in commercial practice so far utilizing immobilized biocatalysts (Scheme 4). The process using glucose isomerase (EC 5.3.1.5) was first commercialized in 1970 (1,12). The enzyme is available from different bacterial sources. By site-directed mutagenesis, more thermostable mutants have been obtained. Most of the enzyme preparations commercially available are cross-linked with glutaraldehyde (1). The enzyme is used in fixed-bed reactors, and the feed flow is adjusted to overcome enzyme deactivation and en-



4 **5** **6**

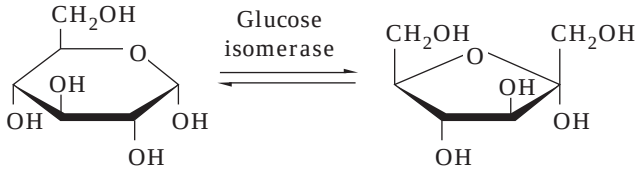
Scheme 2.



7 **8** **9**

10 **11** **12**

Scheme 3.



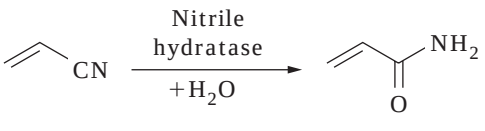
13 **14**

Scheme 4.

sure a constant product quality (conversion). Different columns of different enzyme age are operated in parallel, resulting in an average conversion of 45% in a process of uniform throughput. The equilibrium conversion is about 50%. If the residual activity has dropped to about 10% of the initial activity, the enzyme is replaced. For the most-stable commercially available enzyme preparations, the half-life is more than 100 days. Typical temperatures range from 50 to 60 °C with 50% substrate solution.

Production of Acrylamide Using Nitrile Hydratase

Biotransformations have been also established for the production of bulk chemicals. The enzymatic production of acrylamide (**16**) from acrylonitrile (**15**) using nitrile hydratase (EC 4.2.1.84) from *Rhodococcus rhodochrous* has replaced the copper-catalyzed process in Japan (Scheme 5) (1,87). For this process, whole cells were immobilized in a cationic acrylamide-based polymer gel. Typical reaction conditions range from 0 to 15 °C and pH 6.5 to 8.0. Acrylonitrile is added gradually in order to achieve a final concentration of acrylamide of 400 g L^{-1}. The conversion of acrylonitrile is more than 99%, with a selectivity for acrylamide of more than 99%. Because of its broad substrate specificity, the nitrile hydratase may be also used for production of nicotinamide, isonicotinamide, or pyrazinamide (87).



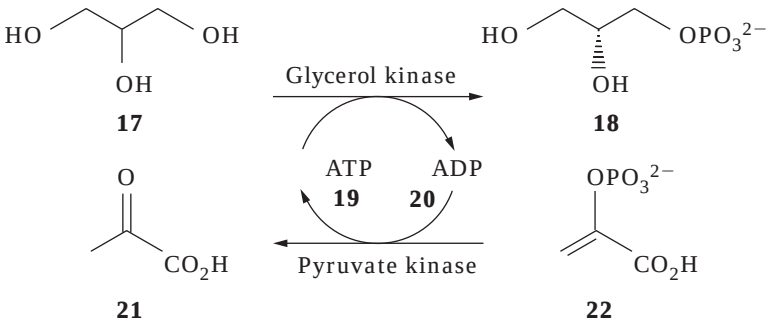
15 **16**

Scheme 5.

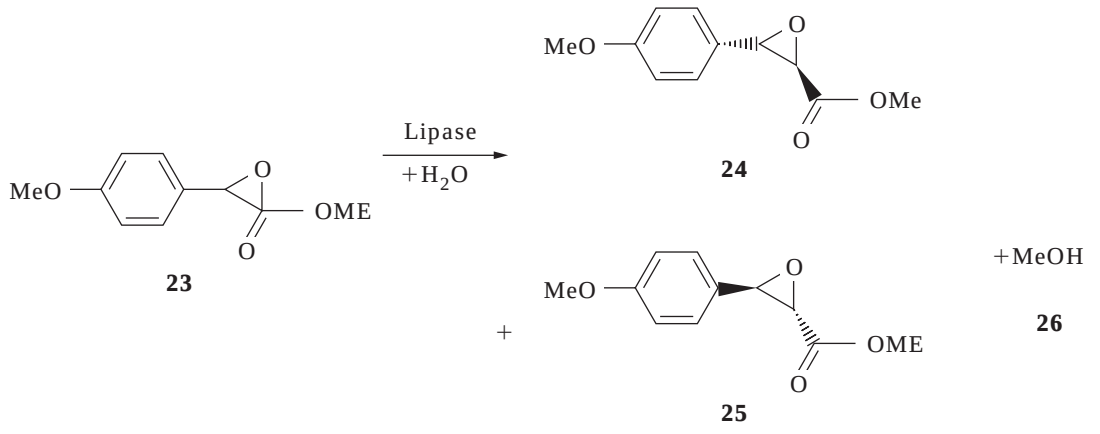
Production of Glycerol-3-Phosphate Using Glycerol Kinase

Glycerol-3-phosphate (G3P) (**18**) is a useful starting material for the synthesis of phospholipids. These compounds are of importance for numerous medicinal, diagnostic, and research applications. Their potential applications include microemulsion and liposomal drug delivery systems (80). The enzymic synthesis of G3P using glycerol kinase (EC 2.7.1.30) has been commercialized by Genzyme (88–90). Adenosine triphosphate (ATP) is used as the phosphorylating agent (Scheme 6), and regeneration of ATP is achieved by pyruvate kinase (EC 2.7.1.40) using phosphoenol pyruvate (PEP) (**22**).

The enzymes used are immobilized on a solid support. Especially for the glycerol kinase the stability was increased dramatically when compared with the native form. Its half-life in aqueous solution is about 15 min, because of dissociation of the subunits. The reaction is typically performed in a batch reactor containing the immobilized enzyme in a fixed bed. Multikilogram-scale production typically uses 1 mol L^{-1} glycerol and 1 mol L^{-1} phosphoenol pyruvate. For this process, space-time yield is about 81 g L^{-1} d^{-1}. Catalytic amounts of ATP (1 to 10 mol%) are used, resulting in total turnover numbers up to 100. These and even higher values are attainable without requiring coimmobilization of the cofactors. NADH values of about 1,000 may be reached easily (27,56). These processes are economically feasible because of the low cost of the cofactors. (*Note*: Cofactors in bulk quantities are available from Oriental Yeast, Japan.)



Scheme 6.



Scheme 7.

Table 3. Further Examples for Application of Immobilized Enzymes

Enzyme	Substrate	Product	Ref.
Glucose kinase	Glc, ATP	Glucose-6-phosphate	97
Oxynitrilase	Aldehydes, HCN	Cyanohydrines	44
Transaminase	Ketone	Amine	98,99
Pyranose oxidase	Sugar	Keto-sugars	100
Hydantoinase	Hydantoin	Amino acid	101
Oxidoreductase	Benzoylformate	Mandelate	102
Sialidase	Neu5Ac, Gal	Neu5Ac-Gal	103
Glucosidase	Glc, octanol	Octylglycoside	104
Chymotrypsin	Amino acids	Oligopeptides	105
Glycosyltransferase	Carbohydrate	Oligosaccharide	10,106

Production of (2R,3S)-Methyl-*p*-Methoxyphenylglycidate Using a Lipase

Optically pure (2R,3S)-methyl-*p*-methoxyphenylglycidate (**24**) is used as chiral building block for the synthesis of diltiazem. Sepracor has developed a process for enantioselective hydrolysis of the racemic ester using a lipase, which has reached a 50 t a^{-1} scale (91,92) (Scheme 7). The reaction is performed in a multiphase membrane reactor with the enzyme reversibly entraped within the porous matrix of an asymmetric hollow fiber ultrafiltration membrane (92–94). This reactor is shown in Figure 1**i**. An aqueous buffer is circulated on one side of the hydrophilic membrane, providing an aqueous environment for the enzyme. A 17% solution of the racemic ester (**23**) dissolved in toluene is circulated on the other side of the membrane. Despite its lower enantioselectivity, *Candida cylindracea* lipase (EC 3.1.1.3) is being used; lipases from other sources were not retained sufficiently within the membrane. The reactor is used in a batch mode. The space-time yield, based on the membrane area, is about 125 g m^{-2} d^{-1}.

This type of reactor has also been used for the hydrolysis of butter oil (95) or the production of peroxycarboxylic acids (96).

Miscellaneous Processes Using Immobilized Enzymes

Table 3 lists some reactions that have been performed using immobilized enzymes to give an impression of the range of enzymatic biotransformations. For a comprehen-

sive overview the reader is refered to reviews and textbooks (1,5–20). Some of the examples are already used on industrial level. Application of soluble enzymes retained by ultrafiltration membranes have been successful as well (21–28).

NOMENCLATURE

Symbols and Abbreviations

Note:	$1\ kat = 1\ mol\ s^{-1}$, $1\ U = 1\ \mu mol\ min^{-1}$
	$1\ kat = 60 * 106\ U$
A_{max}	Maximum mass specific activity, $U\ g^{-1}$, $kat\ g^{-1}$, or $mol\ s^{-1}\ g^{-1}$
$A_{max,A}$	Maximum area specific activity, $mol\ s^{-1}\ m^{-2}$
ATP	Adenosine triphosphate
c	Concentration, $mol\ L^{-1}$, $g\ L^{-1}$
D	Diffusion coefficient, $m^2\ s^{-1}$
d	Particle size, m
Gal	Galactose
Glc	Glucose
k	Stoff transport coefficient, $m\ s^{-1}$
K_m	Michaelis-Menten constant, $mol\ L^{-1}$
η	Efficiency coefficient, −
Neu5Ac	*N*-acetylneuraminic acid
Sh	Sherwood number, −
v	Reaction rate, $kat\ L^{-1}$
v_{mt}	Mass transport, $mol\ m^{-2}\ s^{-1}$

Indices

0	Time zero or in bulk phase
A	Substance A
app	Apparent
B	Substance B
E	Enzyme

BIBLIOGRAPHY

This is a modified version of the article first published as U. Kragl and C. Wandrey, "Biocatalytic Reactions with Immobilized Enzymes" in G. Ertl, H. Knözinger, J. Weitkamp eds., *Handbook of Heterogeneous Catalysis*, Wiley-VCH Verlag GmbH, D-69451 Weinheim, Germany, 1997, pp. 2436–2447.

1. A. Tanaka, T. Tosa, and T. Kobayashi eds., *Industrial Application of Immobilized Biocatalysts*, Decker, New York, 1993.
2. J.G. Shewale, B.S. Deshpande, V.K. Sudhakran, and S.S. Ambedkar, *Process. Biochem. Int.* **25**, 97–103 (1990).
3. V. Kasche, W. Hass, C. Niebuhr-Redder, and E. Piotratscke, *Bioengineering* **2**, 69–77 (1992).
4. F. Balkenhol, K. Ditrich, B. Hauer, and W. Ladner, *J. Prakt. Chemie* **339**, 381–384 (1997).
5. I. Chibata ed., *Immobilized Enzymes, Research and Development*, Kodansha, Tokyo, 1978.
6. H.G. Davies and R.H. Green, *Biotransformations in Preparative Organic Chemistry: The Use of Isolated Enzymes and Whole Cell Systems in Synthesis*, Academic Press, London, 1989.
7. M.W. Blanch and D.S. Clark eds., *Applied Biocatalysis*, vol. 1, Decker, New York, 1991.
8. K. Faber, *Biotransformations in Organic Chemistry*, Springer-Verlag, Berlin, 1995.
9. S.M. Roberts, K. Wiggins, and G. Casy, *Whole Cells and Isolated Enzymes in Organic Chemistry*, Wiley, Chichester, U.K., 1992.
10. C.H. Wong and G.M. Whitesides, *Enzymes in Synthetic Organic Chemistry*, Pergamon Press, Oxford, U.K., 1994.
11. W. Gerhartz, Y.S. Yamamoto, L. Kandy, J.I. Rounsaville, and G. Schulz eds., *Ullman's Encyclopedia Industrial Chemistry*, 5th ed., vol. 9, Wiley-VCH, Weinheim, 1987, pp. 341–530.
12. W. Gerhartz, *Enzymes in Industry*, VCH, Weinheim, 1990.
13. A.N. Collins, G.N. Sheldrake, and J. Crosby eds., *Chirality in Industry*, Wiley, Chichester, U.K., 1992.
14. D. Rozzell and F. Wagner, *Biocatalytic Production of Amino Acids & Derivatives*, Hanser, Munich, 1992.
15. K. Mosbach ed., *Methods Enzymol.* **135B**, **136C**, **137D**, 1987.
16. E. Katchalski-Katzir, *Trends Biotechnol.* **11**, 471–478 (1993).
17. R.A. Sheldon, *Chirotechnology: Industrial Synthesis of Optically Active Compounds*, Decker, New York, 1993.
18. K. Drauz and H. Waldmann eds., *Enzyme Catalysis in Organic Synthesis*, VCH, Weinheim, Germany, 1995.
19. B.P. Sharma, L.F. Bailey, and R.A. Messing, *Angew. Chem. Int. Ed. Engl.* **21**, 837 (1982).
20. A. Rosevar, *J. Chem. Tech. Biotechnol.* **34B**, 127–150 (1984).
21. M.-R. Kula and C. Wandrey, *Methods Enzymol.* **136**, 9–21 (1987).
22. U. Kragl, D. Vasic-Racki, and C. Wandrey, *Chem. Ing. Tech.* **64**, 499–509 (1992).
23. U. Kragl, D. Vasic-Racki, and C. Wandrey, *Ind. J. Chem.* **32B**, 103–117 (1993).
24. A.S. Bommarius, K. Drauz,U. Groeger, and C. Wandrey, in A.N. Collins, G.N. Sheldrake, J. Crosby eds., *Chirality in Industry*, Wiley, Chichester, U.K., 1992, pp. 371–397.
25. A.S. Bommarius, in H.-J. Rehm, G. Reed, A. Pöhler, P. Stadtler, G. Stephanopoulus eds., *Biotechnology* vol. 3, VCH, Weinheim, 1993, pp. 427–466.
26. D.M.F. Prazeres and J.M.S. Cabral, *Enzyme Microb. Technol.* **16**, 738–750 (1994).
27. U. Kragl, in T. Godfrey, M. Daniels, S. West eds.,*Industrial Enzymology*, Macmillan, London, 1996, pp. 271–283.
28. U. Kragl, A. Gödde, C. Wandrey, W. Kinzy, J.J. Cappon, and J. Lugtenburg, *Tetrahedron Asymmetry* **4**, 1193–1202 (1993).
29. P. Brodelius and E.J. Vandamme, in H.J. Rehm, G. Reed, J.F. Kennedy eds., *Biotechnology*, vol. 7a, VCH, Weinheim, 1987, pp. 405–464.
30. A. Grobiollot, D.K. Boadi, D. Poncelet, and R. Neufeld, *Crit. Rev. Biotechnol.* **14**, 75–107 (1994).
31. S. Norton and J.C. Vuillemard, *Crit. Rev. Biotechnol.* **14**, 193–224 (1994).
32. J.A.M. de Bont, J. Visser, B. Mattiasson, and J. Tramper eds., *Physiology of Immobilized Cells*, Elsevier, Amsterdam, 1990.
33. J.F. Kennedy and J.M.S. Cabral, in H.J. Rehm, G. Reed, J.F. Kennedy eds., *Biotechnology*, vol. 7a, VCH, Weinheim, 1987, pp. 347–404.

34. W. Keim and B. Dreißen-Hölscher, in G. Ertl, H. Knözinger, J. Weitkamp eds., *Handbook of Heterogeneous Catalysis*, Wiley-VCH, Weinheim, 1997, pp. 231–240.
35. L.J. Blum and P.R. Coulet eds., *Biosensor Principles and Applications*, Decker, New York, 1991.
36. R.D. Schmid and F. Scheller eds., *Biosensors: Fundamentals, Technologies and Applications*, VCH, Weinheim, 1992.
37. J.E. Zull, J. Reed-Mundell, Y.W. Lee, D. Vezenow, N.P. Ziats, J.M. Anderson, and C.N. Sukenik, *J. Industr. Microbiol.* **13**, 137–143 (1994).
38. P. Wang, T.G. Hill, C.A. Wartchow, M.E. Huston, L.M. Oehler, M.B. Smith, M.D. Bednarski, and M.R. Callstrom, *J. Am. Chem. Soc.* **114**, 378–380 (1992).
39. D. Leckband and R. Langer, *Biotechnol. Bioeng.* **37**, 227–237 (1991).
40. V.V. Mozhaev, N.S. Melik-Nubarov, M.V. Sergeeva, V. Siksnis, and K. Martinek, *Biocatalysis* **3**, 179–187 (1990).
41. R.D. Schmidt, in T.K. Gose, A. Fiechter, N. Blakebrough eds., *Advances in Biochemical Engineering*, vol. 12, Springer-Verlag, Berlin, 1979, pp. 41–118.
42. R. Ulbrich, R. Golbik, and A. Schellenberger, *Biotechnol. Bioeng.* **37**, 280–287 (1991).
43. C. Laane, J. Tramper, and M.D. Lilly eds., *Biocatalysis in Organic Media*, Elsevier, Amsterdam, 1987.
44. F. Effenberger, *Angew. Chem. Int. Ed. Engl.* **33**, 1555–1564 (1994).
45. L.E.S. Brink, J. Tramper, K.C.A.M. Luyben, and K. Van't Riet, *Enzyme Microb. Technol.* **10**, 736–743 (1988).
46. M.N. Gupta, *J. Biochem.* **203**, 25–32 (1992).
47. A.M. Klibanov, *Acc. Chem. Res.* **23**, 114–120 (1990).
48. W.F.M. Stöcklein and F.W. Scheller, *Chem.-Ing. Techn.* **67**, 69–77 (1995).
49. E. Wehtje, P. Adlerkreutz, and B. Mattiasson, *Biotechnol. Bioeng.* **41**, 171–178 (1993).
50. P. Adlerkreutz, *Eur. J. Biochem.* **199**, 609–614 (1991).
51. I. Chibata, T. Tosa, and T. Sato, in H.-J. Rehm, G. Reed, and J.F. Kennedy eds., *Biotechnology*, vol. 7a, VCH, Weinheim, 1987, pp. 653–684.
52. W. Gerhartz, Y.S. Yamamoto, L. Kandy, J.F. Rounsaville, and G. Schulz eds., *Ullman's Encyclopedia Industrial Chemistry*, 5th ed., vol. A9, Wiley-VCH, Weinheim, 1987, pp. 382–388.
53. M. Biselli, U. Kragl, and C. Wandrey, in K. Drauz and H. Waldmann eds., *Enzyme Catalysis in Organic Synthesis*, VCH, Weinheim, 1995.
54. I.H. Segel, *Enzyme Kinetics*, Wiley, New York, 1975.
55. A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, Butterworths, London, 1981.
56. W. Kruse, W. Hummel, and U. Kragl, *Recl. Trac. Chim. Pays-Bas* **115**, 239–243 (1996).
57. A. Liese, M. Karutz, J. Kamphuis, C. Wandrey, and U. Kragl, *Biotechnol. Bioeng.* **51**, 544–550 (1996).
58. P. Surmann and E. Nürnberg eds., *Hagers Handbuch der Pharmazeutischen Praxis*, vol. 2, Springer-Verlag, Berlin, 1991, pp. 426–428.
59. D.S. Clark, *Trends Biotechnol.* **12**, 439–443 (1994).
60. J.E. Prenosil, I.J. Dunn, and E. Heinzle, in H.J. Rehm, G. Reed, J.F. Kennedy eds., *Biotechnology*, Vol. 7a, VCH, Weinheim, 1987, pp. 489–545.
61. B. Elvers, S. Hawkins, M. Ravenscraft, and G. Schulz eds., *Ullmann's Encyclopedia Industrial Chemistry*, 5th ed., vol. 14, Wiley-VCH, Weinheim, 1989, pp. 1–48.
62. C. Wandrey and E. Flaschel, in T.K. Ghose, A. Fiechter, N. Blakebrough, and A. Fiechter eds., *Advances in Biochemical Engineering*, Vol. 12, Springer-Verlag, Berlin, 1979, pp. 147–218.
63. O. Levenspiel, *Chemical Reaction Engineering*, Wiley, New York, 1972.
64. J.E. Bailey and D.F. Ollis, *Biochemical Reaction Fundamentals*, McGraw-Hill, New York, 1986.
65. H. Chmiel ed., *Bioprozeßtechnik Band 1 und 2*, Gustav Fischer, Stuttgart, 1991.
66. K. Buchholz, *Characterisation of Immobilized Biocatalysis*, vol. 84, VCH, Weinheim, 1979.
67. F. Shadman-Yazdi and E.E. Petersen, *Chem. Eng. Sci.* **27**, 227–237 (1972).
68. B.J. Rovito and J.R. Kittrell, *Biotechnol. Bioeng.* **15**, 143–161 (1973).
69. R. Lortie and D. Thomas, *Biotechnol. Bioeng.* **28**, 1256–1260 (1986).
70. Y.H. Park, M.H. Han, and H.-K. Rhee, *J. Chem. Tech. Biotechnol.* **34B**, 57–69 (1984).
71. M. Baerns, H. Hoffmann, and A. Renken, *Chemische Reaktionstechnik*, Georg Thieme, Stuttgart, 1992.
72. J.G.T. Kierkels, L.F.W. Vleugels, E.T.F. Gelad, D.P. Vermeulen, J. Kamphuis, and C. Wandrey, *Enzyme Microb. Technol.* **16**, 513–521 (1994).
73. B. Bossow-Berke, Ph.D. Thesis, Bonn University, Bonn, 1989.
74. B. Bossow-Berke, W. Berke, and C. Wandrey, *Bioengineering* **8**, 12–19 (1992).
75. J. Danzig, W. Tischer, and C. Wandrey, *Ind. J. Chem* **32B**, 40–43 (1993).
76. J. Danzig, W. Tischer, and C. Wandrey, *Chem. Eng. Technol.* **18**, 256–259 (1995).
77. M.L. Bender, M.L. Begecanton, R.L. Blakely, L.J. Brubacher, J. Feder, C.R. Gunter, F.J. Kezdy, J.V. Killhefter, T.H. Marshall, C.G. Miller, R.W. Roeske, and J.K. Stoops, *J. Am. Chem. Soc.* **88**, 24 (1966).
78. J.R. Ford, R.P. Chambers, and W. Cohen, *Biochim. Biophys. Acta* **309**, 175–180 (1973).
79. W. Halwachs, C. Wandrey, and K. Schügerl, *Biotechnol. Bioeng.* **20**, 541–554 (1978).
80. V. Kasche and B. Galunsky, *Biotechnol. Bioeng.* **45**, 261–267 (1995).
81. J.M. Nelson and E.G. Griffin, *J. Am. Chem. Soc.* **83**, 1109–1115 (1916).
82. Int. Union of Biochemistry and Molecular Biology, *Enzyme Nomenclature*, Academic Press, San Diego, 1992.
83. T. Tosa, T. Mori, N. Fuse, and I. Chibata, *Agric. Biol. Chem.* **33**, 1047–1052 (1969).
84. W. Tischer, U. Gieseke, G. Lang, A. Röder, and F. Wedekind, *Ann. N.Y. Acad. Sci.* **672**, 502–509 (1992).
85. J.G. Shewale and H. Sivaram, *Process Biochem.* **24**, 146–154, (1989).
86. *United Nations Industrial Development Organization, Enzyme Engineering at the Industrial Level—Present Status and Future Prospects*, Wien, 1989, pp. 39–47.
87. T. Nagasawa and H. Yamada, *Trends Biotechnol.* **7**, 153–158 (1989).
88. A.E. Walts, D.G. Shena, E.M. Fox, J.T. Davis, and M.R. Mische, in A.N. Collins, G.N. Sheldrake, J. Crosby eds., *Chirality in Industry*, Wiley, New York, 1992, pp. 223–235.
89. D.C. Crans and G.M. Whitesides, *J. Am. Chem. Soc.* **107**, 7019–7027 (1985).

90. A.E. Walts, D.G. Shena, E.M. Fox, J.T. Davies, and M.R. Mischke, *Chim. Oggi.* **9**, 29–33 (1991).
91. J.L. Lopez, *First Int. Symposium on Technologies for the Production of Enantiomerically Pure Compounds*, Amelia, Florida, 1993.
92. M. Furui, T. Furutani, T. Shibatani, Y. Nakamoto, and T. Mori, *J. Ferment. Bioeng.* **81**, 21–25 (1996).
93. WO 88/07582 (Oct. 6, 1988), S.L. Matson (to the regents of Sepracor Inc.).
94. J.L. Lopez, S.A. Wald, S.L. Matson, and J.A. Quinn, *Ann. N.Y. Acad. Sci.* **613**, 155–166 (1990).
95. F.X. Malcata, C.G. Hill Jr., and C.H. Amundsen, *Biotechnol. Bioeng.* **39**, 984–1001 (1992).
96. F.P. Cuperus, S.T. Bouver, G.F.H. Kramer, and J.T.P. Derksen, *Biocatalysis* **9**, 89–96 (1994).
97. H. Nakajima, H. Kondo, R. Tsurutani, M. Dombou, I. Tomioka, and K. Tomita, *ACS Symp. Ser.* **466**, 111–120 (1991).
98. G. Matcham, A. Bowen, S. Lee, P. Pienkos, and A. Zeitlin, *Proc. Chiral 94 USA Symposium*, Spring Innovations, Reston, Va., 1994, pp. 309–315.
99. D.I. Stirling, in A.N. Collins, G.N. Sheldrake, J. Crosby eds., *Chirality in Industry*, Wiley, New York, 1992, pp. 209–222.
100. A. Huwig, H.J. Daneel, and F. Griffhorn, *J. Biotechnol.* **32**, 309–315 (1994).
101. E.U. Pat. 92,118,795 (1992), F. Wagner, M. Pietsch, and C. Syldatk.
102. K. Hosono, S. Kajiwara, Y. Yamazaki, and H. Maeda, *J. Biotechnol.* **14**, 149–156 (1990).
103. J. Thiem and B. Sauerbrei, *Angew. Chem. Int. Ed. Engl.* **30**, 1503 (1991).
104. G. Ljunger, P. Adlerkreutz, and B. Mattiason, *Enzyme Microb. Technol.* 16, 751–755 (1994).
105. A.O. Richards, I.S. Gill, and E.N. Vulfson, *Enzyme Microb. Technol.* **15**, 928–935 (1993).
106. J. Hodgson, *Biotechnology* **13**, 38–41 (1995).

ENZYMES, PROTEIN HYDROLYSIS

HANS SEJR OLSEN
Novo Nordisk
Bagsvaerd, Denmark

KEY WORDS

Dairy products
Detergents
Enzymes
Food processing
Hydrolysis
Industrial cleaning
Leather
Osmometry
pH stat
Whipping agents

OUTLINE

INTRODUCTION

Industrial enzymology is an important branch of bioprocess technology. Enzymatic processes permit natural raw materials to be upgraded and finished. Using enzymes as a functional ingredient in detergents, laundry and dishes, for example, are cleaned in an efficient, environmentally

mild, and energy-saving way. The most used industrial enzymes are proteases for hydrolysis reactions within the detergent, dairy, protein, pet food, feed, and leather tanning industries.

In detergents proteases are the most widely used enzymes. In laundry detergents, protein stains such as grass, blood, egg, and human sweat are removed through proteolysis. In automatic dishwashing detergents, proteases secure the removal of proteinaceous food films, which are a particular problem with glassware and cutlery.

Enzymes offer alternative ways of making products previously made using conventional chemical processes. Proteases, in particular, offer attractive ways of performing hydrolysis reactions needed for most tasks within the mentioned industries. Historically, proteases were used in ancient Greece for the production of cheese, as referenced in Greek epic poems dating from about 800 B.C.

The development of the submerged fermentation technique has resulted in tremendous progress in the field of industrial enzymology. Due to their efficiency, specific action, ability to work under mild conditions, high purification, and standardization, proteases are ideal catalysts for the food industry. Simple equipment can be used, and because moderate temperature and pH values are required for the reactions, only a few by-products influencing taste and color are formed. Furthermore, reactions using proteases are easily controlled and can be stopped when the desired degree of conversion is reached.

World Consumption of Industrial Proteases

Worldwide consumption of proteases amounted to approximately $600 million in 1996. The total growth (including nonfood applications) in volume of the enzyme business from 1974 to 1986 was 10–15% per year, although no major new applications emerged in this period. From 1980 to 1990 the growth was estimated as 5–10% per year, and from 1992 to 1996 as 8–10%.

The principal producers of industrial proteases are Novo Nordisk A/S (headquartered in Denmark) and Genencor International, Inc. (headquartered in the United States). These companies provide more than 90% of the total volume of enzyme products to the detergent enzyme market.

Regulatory Aspects and Quality Assurance of Detergent Enzymes

In most countries, regulatory status classification, and labeling of enzymes are determined in accordance with existing schemes for chemicals. Many enzyme types are listed on chemical inventories, for example, European Inventory of Existing Commercial Chemical Substances (EINECS) in the European Union and Toxic Substances Control Act (TSCA) in the United States. In some cases, enzymes are considered natural substances exempt from listing. In other cases enzymes are regulated by specific biotechnology products legislation.

The Association of Manufacturers of Fermentation Enzyme Products (AMFEP) has defined good manufacturing practice (GMP) for microbial food enzymes. This practice is generally followed for detergent enzymes as well. The most important element is to ensure a pure culture of the production organism.

When an enzyme is used for a nonfood and nonfeed industrial technical application, its regulatory status is determined by its properties as a naturally occurring substance. These properties determine the classification and consequent labeling in accordance with existing schemes for chemicals.

Regulatory Aspects and Quality Assurance of Food Enzymes

The application of enzymes in food processing is governed by food laws. Within the European Union large parts of food laws have been harmonized by directives and regulations. For general purposes, the Joint FAO/WHO Expert Committee on Food Additives (JECFA) and the Food Chemicals Codex (FCC) have made guidelines available for the application of enzymes as food additives. AMFEP (in Europe) and the Enzyme Technical Association (ETA) (in the United States) work nationally, as well as internationally, for harmonization of regulations with regard to enzymes.

AMFEP members ensure that the enzymes used in food processing are obtained from nonpathogenic and nontoxicogenic microorganisms, that is, microorganisms that have clean safety records, with no reported cases of pathogenicity or toxicosis attributed to the species in question. When the production strain contains recombinant DNA, the characteristics and safety record of each donor organism contributing genetic information to the production strain are assessed.

The majority of food enzymes are used as processing aids and have no function in the final food. In this case they need not be declared on the label, because they will not be present in the final food in any significant amount. Some enzymes are used as both processing aids and food additives. When used as additives, they must be declared on the food label.

GMP is used in the microbial food enzyme industry. The key issues in GMP are the microbial control of the microorganism selected for enzyme production, the control and monitoring systems ensuring pure culture and optimum enzyme productivity conditions during fermentation, and control of the hygienic conditions throughout recovery and finishing of the enzyme preparation.

Commercial enzyme products are usually formulated in aqueous solutions or processed to nondusty dry products (granulates or microgranulates). Both types of preparations must be formulated to be suitable for the final application.

Requirements mainly regarding the storage stability (e.g., enzyme activity stability, microbial stability and physical stability) and formulations as significant quality parameters of industrial enzyme products must be taken into consideration by both the producer and the user.

Safety

Like many other proteins foreign to the human body, enzymes are potential inhalation allergens. Inhalation of even small concentrations of a foreign protein in the form of dust or aerosol can stimulate the body's immune system to produce antibodies. In some individuals, increased concentrations of antibody–enzyme protein complexes can

trigger increased concentrations of histamine. The latter compounds can cause hay fever–like symptoms, such as watery eyes, runny nose, and sore throat. When exposure ceases, these effects cease also.

Enzymes must be inhaled to present a risk of causing sensitization, which may lead to an allergic reaction. Working environments in which enzymes are used are therefore subjected to extensive monitoring to confirm that threshold limit values (TLV) for atmospheric enzymes are not exceeded. In many countries, the TLV for enzymes is based on the proteolytic enzyme subtilisin, and is stated as 0.00006 mg/m^3 of pure crystalline subtilisin in air (1).

Successful experience in controlling enzyme exposure in detergent manufacturing and in protecting workers' health has been reported by Procter & Gamble (2). As early as 1971, a National Research Council (NRC) report concluded that consumers of enzyme-containing laundry products did not develop respiratory allergies (3). Further studies over the years have confirmed that enzyme-containing laundry and dishwashing detergents are safe for consumer use.

Use of proteases as aids in food processing virtually never represent safety problems for consumers. Because liquid enzyme preparations are normally used in processing plants in very small concentrations, dust- or aerosol-containing enzymes are usually not released from process equipment based on closed systems.

INDUSTRIAL PROTEASES

It is difficult to tell which applications spurred the development of new proteases in the early days. Besides for production of foods such as cheese, beer, and shoyu (soy sauce), proteases have been developed for the tanning industry, and in the past the detergent industry has been the driving force in development. Scientific interest in proteolytic action on different food proteins has resulted in the development of new proteases for both extraction processes and production of functional ingredients for the food industry. By far the largest tonnage production of protease is based on microbial sources. Recombinant DNA methods and protein engineering are today's means to develop important proteases for both food and nonfood uses.

Proteases are classified according to their source of origin (animal, plant, microbial), their catalytic action (endopeptidase or exopeptidase), and the nature of the catalytic site (active site). They are characterized by common names and trade names, typical pH range, and preferential specificity. Based on a comparison of active sites, catalytic residues, and three-dimensional structures, four major protease families are recognized: serine, thiol, aspartic, and metalloprotease. The serine protease family contains two subgroups: the chymotrypsin-like and the subtilisin-like proteases—the most important industrial products. Many commercial proteases are mixtures of the different types. This is especially the case for pancreatin, papain (crude), and some proteases from *Bacillus amyloliquefaciens*, *Aspergillus oryzae*, *Streptomyces*, and *Penicillium duponti*, mainly used within the food area.

Some commercial proteases have the characteristics shown in Table 1. The enzymes Alcalase, Esperase, Savinase, and Trypsin are serine proteases, while the enzyme Neutrase is a metalloprotease having Zn^{++} in its active site. Furthermore, this enzyme is stabilized by Ca^{++}. Durazym and Everlase are protein-engineered variants of Savinase. The milk-clotting enzyme product Rennilase or Fromase is an aspartic protease. Examples of thiol (or cysteine) proteases are plant proteases from papaya latex (papain), from pineapple stem (bromelin), and from fig latex (ficin).

Table 2 shows the status of today's merchant market for detergents: Practical selections of proteases are made on the basis of performance and stability tests carried out under certain standard conditions chosen in conjunction with detergent formulators.

Protein Hydrolysis in Enzyme Applications

Proteases catalyze the hydrolytic degradation of the peptide chain, as shown for reactants and products in Figure 1. In aqueous solutions and suspensions of protein the equilibrium lies so far to the right that degradation rather than synthesis of larger molecules is thermodynamically favored.

When a protease acts on a protein substrate (see Fig. 1) a catalytic reaction actually consists of three consecutive reactions. This reaction mechanism is simplified by the scheme shown in Figure 2.

1. Formation of the Michaelis complex between the original peptide chain (the substrate) and the enzyme (referred to as ES)
2. Cleavage of the peptide bond to titrate one of the two resulting peptides
3. A nucleophilic attack on the remains of the complex to split off the other peptide and to reconstitute the free enzyme

The rate-determining step is the acylation step characterized by the reaction velocity constant k_{+2}. The ratio k_{-1}/k_{+1} is equal to the Michaelis constant k_w for serine proteases. In other words classical Michaelis–Menten kinetics are valid for protein hydrolysis.

The principal difference of the hydrolysis conditions of food and feed proteins and of protein stains in detergent applications is shown in Table 3. In detergent applications the substrate and enzyme concentrations are very small. Strangely enough, the $E/S\%$ ratio is of the same order of magnitude.

Controlling the Hydrolysis Reaction

The technical discipline of enzymatic hydrolysis of food proteins has been developed over the past 20 years to become a significant area of modern food processing. The number of literature references has grown enormously. Descriptions of the principles, methods of controlling the hydrolysis reaction, ways of carrying out industrial processing, and characterization of functional properties of proteins treated with proteases have been intensively researched and developed at Novo Nordisk A/S. A great deal of this work was introduced by Adler-Nissen (4). Examples and highlights are described in the following.

Table 1. Some Commercial Proteolytic Enzymes

Product names	Microorganism or other origin	State of product	Practical application range pH	Practical application range temperature (°C)
Alcalase®	*Bacillus spp.*	Liquid or granulate	6–10	10–80
Esperase®	*Bacillus spp.*	Liquid or granulate	7–12	10–80
Everlase™	*Bacillus GMO* (engineered)	Liquid or granulate	8–11	15–30
Savinase®	*Bacillus GMO*	Liquid or granulate	8–11	15–75
Durazym®	*Bacillus* spp. *GMO* (engineered)	Liquid or granulate	8–11	15–70
Neutrase®	*Bacillus* spp.	Liquid or granulate	6–8	10–65
Protamex™	*Bacillus* spp.	Microgranulate	6–8	10–65
Flavourzyme™	*Aspergillus* spp.	Liquid or granulate	4–8	10–55
Rennilase® or Fromase®	*Rhizomucor meihei*	Liquid or granulate	3–6	10–50
Trypsin	Pancreas	Granulate	7–9	10–55

Table 2. Merchant Market Detergent Proteases

Producer	Alkaline	High alkaline	High bleach stability	Cold water
Genencor International	Maxatase® Optimase®	Purafect® Maxacal® Opticlean®	Maxapem® Purafect® OxP Opticlean® plus	Properase®
Novo Nordisk A/S	Alcalase	Esperase Savinase	Durazym Everlase	New developments

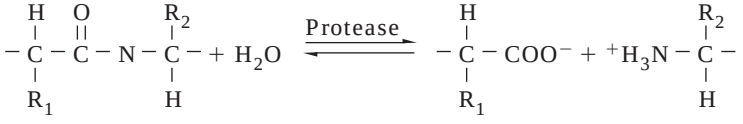


Figure 1. The hydrolysis reaction: Enzymatic cleavage of a peptide bond.

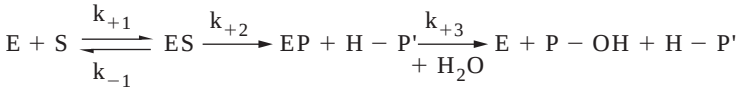


Figure 2. A catalytic mechanism of proteolysis: E, enzyme, S, substrate, P and P′, resulting peptide, *k*, reaction velocity constant.

The most important parameters of hydrolysis reactions are *S* (percent protein in the reaction mixture), *E*/*S* (enzyme–substrate ratio in activity units per kilogram protein), pH, and temperature. Together with the specificity and properties of the enzyme itself, these parameters are responsible for the course of reaction on a given protein raw material. The quantitative criterion of a proteolytic reaction is the degree of hydrolysis, which is calculated from determination of the number of peptide bonds cleaved and the total number of peptide bonds in the intact protein.

As a quantitative measure for the hydrolytic reaction, hydrolysis equivalents are applied, termed *h*. The unity of *h* is equivalent peptide bonds cleaved per kilogram protein. The term *h* is used for kinetic investigations of protein hydrolysis, but in other usages the derivative quantity *DH* (degree of hydrolysis) is preferred. *DH* is calculated from the following equation:

$$DH = \frac{h}{h_{tot}} \times 100\% \qquad (1)$$

h_{tot} is estimated on the basis of the amino acid composition of the protein. For most food proteins, the average molecular weight of amino acids is about 125 g/mol, making h_{tot} about 8 g equiv/kg protein (calculated as $6.25 \times N$). Table 4 gives more exact figures for Kjeldahl conversion factors f_N and for h_{tot} for some common proteins.

The degree of hydrolysis of enzyme-treated proteins determines the properties of relevance to food applications. Therefore, it is of the utmost importance that the degree of hydrolysis can be measured while the reaction is going on. Only in this way is it possible to stop the reactions at a well-defined stage, when the desired property of the product has been obtained. Within detergent applications (laundry or dishwashing), no estimates of this parameter have been reported, presumably because the substrate concentration is considerably low and therefore difficult to measure.

When sufficient substrate is available, relatively simple analytical tools that can be applied directly during the reactions are often used. Some of them are the pH-stat technique, pH drops, osmometry, viscosimetry, and chemical determination of free amino groups, using, for example, a measurement of liberated α-amino groups by reaction with

Table 3. Properties and Parameters Used during Protein Hydrolyses

Property or parameter	In food or feed hydrolysis processes	In detergent applications
Type of proteases in use	endoprotease and/or exoprotease activities	Endoprotease
Substrate concentration	Usually high (>6% protein)	Usually low (<0.1% protein)
Substrate state	Native or denature	Denatured
Enzyme concentration during application.	e.g., 0.025 g enzyme protein/L	e.g., 0.0005 g enzyme protein/L
E/*S*% ratio	0.03% (enzyme protein to substrate)	0.05% (enzyme protein to substrate)
Temperature	45–80 °C	15–95 °C
pH	4–8	10–12
Reaction time	1–4 hr	15–45 min

Table 4. Kjeidahl Conversion Factors and Content of Peptide Bonds for Various Food

Protein	Kjeldahl conversion factor, f_N	h_{tot} equiv/kg ($N \times f_N$)
Casein	6.38	8.2
Whey protein isolate	6.38	8.8
Meat	6.25	7.6
Fish muscle	6.25	8.6
Egg white	6.25	approx. 8
Soy meal, concentrate and isolate	6.25	7.8
Cotton seed	6.25	7.8
Red blood cells	6.25	7.8
Wheat protein	5.7	8.3
Gelatin	5.55	11.1

trinitrobenzenesulfonic acid (TNBS), whereby a colored derivative of the peptide is formed. The TNBS method has been adapted for use for hydrolysates of food proteins, which usually are partly soluble (i.e., the sample should be dispersed in sodium dodecyl sulphate (SDS)) (6). The *O*-phthaldialdehyde and the ninhydrin reaction with liberated α-amino groups may also be used, but with some limitations (7). The pH-stat technique and the osmometry technique are quicker and less labor-demanding than the "chemical" methods. When simultaneous solubilization of protein occurs as an effect of the protein hydrolysis measurements of soluble nitrogen, for example, by Kjeldahl analysis or a simple detection of soluble dry matter by measuring of the increase of refraction index (°BRIX) can be used to describe the hydrolysis reaction.

Calculation of Degree of Hydrolysis of Proteins Using the pH-Stat Technique

Originally, the pH-stat method was developed at the Carlsberg Laboratory (8). It is based on the principle that pH is kept constant during hydrolysis by means of automatic titration with a base, when the hydrolyses are carried out under neutral to alkaline conditions. When the hydrolysis is carried out under acidic conditions, (e.g., by use of pepsin at pH 3) the titration must be made with acid. The high buffer capacity of proteins means that at extremely pH values (i.e., pH > 11 or pH < 3) the pH stat is inoperable (8).

The free carboxyl and free amino groups found after hydrolyses will be more or less ionized depending on the pH of the hydrolysis reaction. Working in the pH range of 6–9.5, the amino groups will be partially protonated. This means that the hydrolysis of proteins in this pH region is accompanied by a release of H^+, as already noted by Sörensen (9). Consequently, at pH values above 7.5–7.8 (the p*K* values at 25 °C), the amino group will be less than half protonated, but the carboxyl groups will be fully dissociated. This leads to a net release of 0.5–1 mol H^+ for each mole of peptide bonds cleaved. In the pH range of 4–6.5, amino groups and carboxyl groups will be approximately equally dissociated. Therefore, in this pH range the pH state is not accurate. *DH* is calculated on the basis of the titration equations as follows:

$$DH = \frac{h}{h_{tot}} \times 100\% \quad (1)$$

$$DH = B \times N_b \times \frac{1}{\alpha} \times \frac{1}{MP} \times \frac{1}{h_{tot}} \times 100\% \quad (2)$$

where *B* is base consumption in mL (or L), N_b is normality of base, α is average degree of dissociation of the α-NH_2 groups (see following), *MP* is mass of protein ($N \times f_N$) in g (or kg), *h* is hydrolysis equivalents in mequiv/g protein (or equiv/kg protein), and h_{tot} is total number of peptide bonds in the protein substrate (mequiv/g protein or equiv/kg) (see Table 4). The degree of dissociation is

$$\alpha = \frac{10^{pH-pK}}{1 + 10^{pH-pK}} \quad (3)$$

p*K* is the average p*K* value of the α-amino groups liberated during the hydrolysis, and it can be determined from a direct assay of these amino groups by use of the TNBS methods. p*K* also varies significantly with temperature because the ionization enthalpy of the amino group is considerable (4). Inserting the ionization enthalpy (+45 kJ/mol) in the Gibbs–Helmholtz equation, it was found that p*K* changed about 0.23 pH units for a change of 10 °C in the hydrolysis temperature (10). Table 5 gives results of calculations of the degrees of dissociation α from equation 3 for various pH values as a function of the temperature using the previously described method.

Calculation of Degree of Hydrolysis of Proteins Using the Osmometry Technique

The osmometry technique can be used to calculate the degree of hydrolysis of proteins when no soluble components are added during the reaction (i.e., in ranges of the pH

Table 5. Degree of Dissociation (α Values)

	T(°C)					
	40	50	60	70	75	80
pK	7.3	7.1	6.9	6.7	6.6	6.5
pH						
6.5	—	0.20	0.29	0.39	0.44	0.50
7.0	0.33	0.44	0.55	0.67	0.71	0.76
7.5	0.61	0.71	0.80	0.86	0.89	0.91
8.0	0.83	0.89	0.93	0.95	0.96	0.97
8.5	0.94	0.96	0.97	0.98	0.99	0.99

Source: From Ref. 5

scale where the pH-stat principle is not functioning) (11). A freezing-point osmometer can record the osmolality of a sample in a few minutes on a sample taken directly from the hydrolysis mixture. By drawing samples as a function of time during a hydrolysis reaction, a good reaction curve can be constructed. With the osmometer ΔT is measured and converted to osmole though the simple proportionality $\Delta T = K_f \times$ osmolality.

From the increase in osmolality ΔC, DH is calculated by use of the following equation:

$$DH = \frac{\Delta C}{S\% \times f_{osm}} \times \frac{1}{\omega} \times \frac{1}{h_{tot}} \times 100\% \qquad (4)$$

where ΔC is the increase in osmolality measured in milliosmol/kg H_2O, S% is the protein substrate concentration ($N \times 6.25$), w/w%, ω is the osmotic coefficient ($\omega = 0.96$ for most actual concentrations of protein [4]), and f_{osm} is the factor to convert percent to grams per kilogram H_2O, which can be calculated using the following equation:

$$f_{osm} = \frac{1000}{100 - D\%} \qquad (5)$$

where D% is percent dry matter present in the reaction mixture, and h_{tot} is total number of peptide bonds in the protein substrate (mequiv/g protein or equiv/kg).

Some Kinetic Aspects of Protein Hydrolysis

The pH-stat technique has been utilized at Novo Nordisk to study the proteolytic degradation of hemoglobin, for example, and to understand the reaction mechanism.

In discussions of initial proteolysis, Linderstrøm-Lang (12) referred to the reaction of a protease on native hemoglobin molecules as one-by-one, indicating that a particular protease molecule degraded one substrate molecule at a time. No appreciable amounts of intermediary products will be present. The reaction mixture will consist of native proteins and end products only.

In another reaction mechanism discussed, the native protein molecules were rapidly converted into intermediary forms, which degraded more slowly to end products—the zipper reaction.

As seen on the hydrolysis curves made by using Alcalase in Figure 3 (see Table 1) at 50 and at 55 °C in the pH stat,

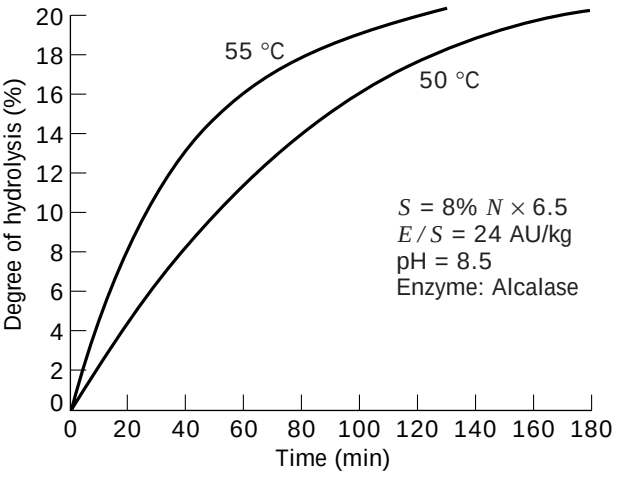


Figure 3. Hydrolysis curves for a hemolyzed red blood cell fraction.

the slight bending at low DH values of the blood-cell-fraction hydrolysis curves indicates a thorough initial degradation to small peptides (a zero-order reaction). This degradation is close to the ideal one-by-one reaction. The hydrolysis proceeds rapidly with little decrease in velocity until the substrate concentration is so low that first-order kinetics take over.

A gel chromatogram of the soluble part of the hemoglobin hydrolysate at DH 10% and DH 20% (Fig. 4) fully confirms the degradation pattern of the hemoglobin substrate. At DH 10%, the hydrolysate consists of high molecular weight material and small peptides in accordance with the one-by-one mechanism. At DH 20% the high molecular weight material has completely disappeared, as predicted by Linderstrøm-Lang as early as 1952 (12).

This is an important factor behind the application of ultrafiltration in the industrial process for separation of the hydrolysate from precipitated hemin, called the sludge (13). If the hydrolysis was not of the one-by-one type, large

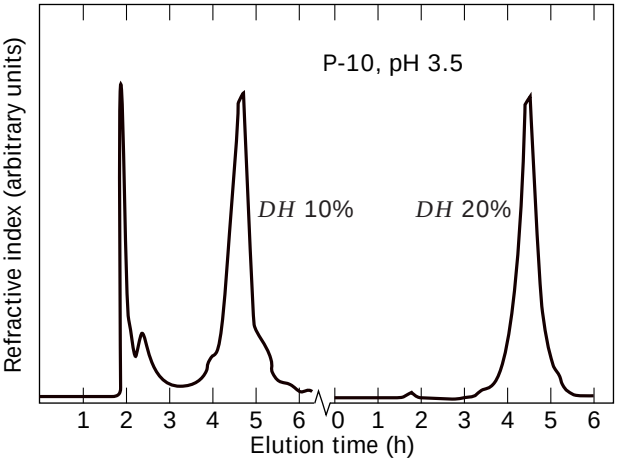


Figure 4. Gelchromatography of hemoglobin hydrolysate. *Source*: From Ref. 4.

peptides would be present also, and consequently the yields and fluxes would have been much lower. One should be careful not to draw too simple conclusions about kinetic models on enzymatic hydrolysis on even a fairly simple and well-defined substrate as hemoglobin. The HPLC spectrum of blood cell hydrolysates (BCH) is shown in Figure 5. At least 50 different peptide compounds are present under the single peak in the gel chromatogram of *DH* 20%.

The Effect of Proteolysis

When the protein structure is modified by proteolysis, the various effects are considered either positive or negative, depending on whether the effect is wanted or unwanted. The developments within enzymatic hydrolysis have often been within both these criteria. The uses and importance of wanted effects including increased solubility, reduced surface tension, increased emulsification, increased foaming properties, and detergency of textile fibers and hard surfaces have been studied. Unwanted effects, such as increased bitterness and off-flavor developments, have been key problems to solve within the food area.

During an enzymatic hydrolysis of proteins, the size character of the soluble peptides produced is changed. Key indices such as %*DH*, protein solubility index (PSI), and protein solubility in 0.8 M trichloroacetic acid (TCA index) increase, whereas peptide chain length (PCL) is reduced, as shown in Figure 6.

Protease Inhibitors

Some proteins are natural inhibitors of proteases. Trypsin and chymotrypsin inhibitors of legume and cereal grains have thus been reviewed and studied (14). Protease inhibitors for bacterial proteases, such as serine proteases (e.g., subtilisin) have been purified and characterized from legume seeds (15) and potatoes (16).

The probably best-known and most-studied protease inhibitors are the trypsin inhibitors of soybeans. The nutritional significance of the antitryptic activity of soybeans, soybean foods, and feed materials has been known for a long time. The way to inactivate these inhibitors has been an inherent part of most manufacturing processes for processes based on soybean material (17).

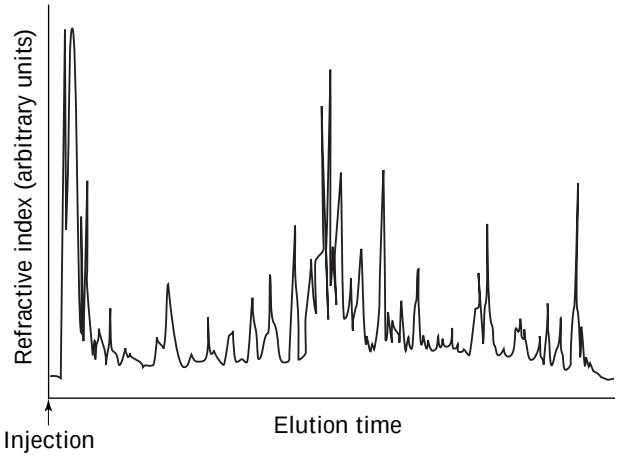


Figure 5. HPLC chromatogram of blood cell hydrolysates, *DH* 20%. *Source*: From Ref. 4.

Addition of protease inhibitors is known to protect proteases in detergents during storage. In powder detergents, where proteases are granulated and encapsulated, all enzymes are stable in the absence of water. In liquid detergents (more than 40% of the U.S. market [18]), the aqueous environment is challenging enzyme stability, and inhibitors are necessary to prevent autodigestion of the proteases themselves or to prevent other enzymes from being degraded. Polyols such as glycerin and boric acid have been used in combination to inhibit protease activity and reduce autodigestion. Borate works as a reversible protease inhibitor by interacting with groups in the active site of the enzymes. Once in the wash water, these inhibitors are diluted and enzyme activity increases (18). Protein hydrolysate products and amino acids are also known to be able to stabilize proteases by binding to the active site of the enzyme (19).

In liquid formulations used in the food industry, the way to stabilize proteases against loss of catalytic activity by denaturation and against microbial growth is by combining low water activity and a reduced pH.

PROTEASES FOR DETERGENTS

Over the past 30 years, the importance of enzymes in detergents has changed from being minor additives to being key ingredients. The main areas where the use of proteases has grown are household laundry, automatic dishwashing, and industrial and institutional cleaning (I&I). I&I covers laundry, dishwashing, hard-surface cleaning, cleaning-in-place, and membrane cleaning.

All commercial proteases from the early 1990s are serine endoproteases produced by *Bacillus* strains (Tables 1 and 2). Some have a maximum activity in the high pH range (pH 8–12) and others have a maximum activity in the pH range 7–10. Protease performance is influenced by factors such as detergent pH, ionic strength, wash temperature, wash time, detergent composition, and mechanical handling. Both performance and stability are influenced by some detergent surfactants and some bleach systems.

To illustrate the reaction conditions under which enzymes must operate, Table 6 shows the components of household detergents in broad terms. Heavy-duty powdered detergent formulations may vary across geographic regions, with a tendency for higher content of builders in North America than in Europe and Japan (20). The ratio of anionic to nonionic surfactant usage also varies from region to region. Furthermore, the content of surfactants is higher in Japan than elsewhere (20). The remarkable difference between detergent formulations for laundry and dishwashing is a much lower content of surfactants in dishwashing powder.

The variable compositions in different regions should be seen in terms of the laundry practices common in North America, Europe, and Japan. Table 7 shows some general laundry washing conditions and washing machine types in

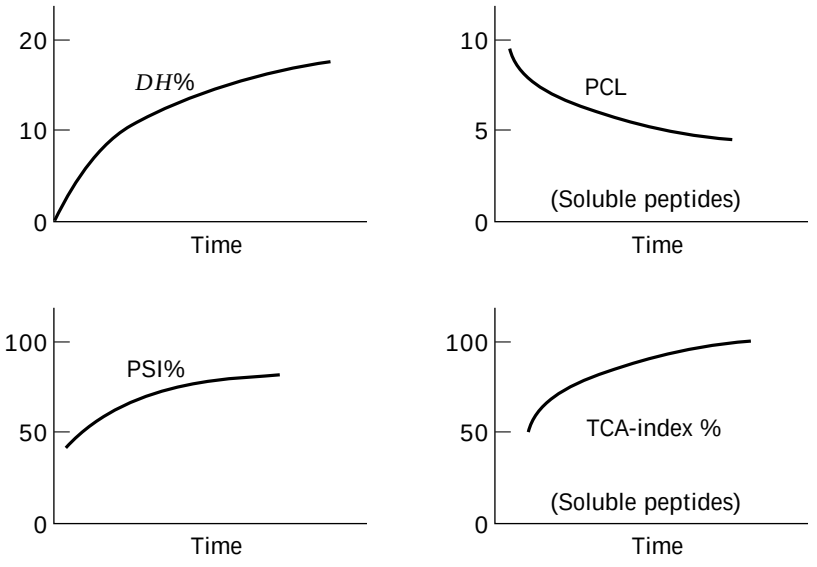


Figure 6. Qualitative illustration showing the time dependence of key indices during enzymatic hydrolysis of protein. *Source*: From Ref 4.

the three principal geographical regions. The ratio of wash liquor to fabric load, temperatures, the usual dosage of detergents, and water hardness around the world are significantly different, as are detergent compositions. Most of these parameters may influence the detergency of proteases in a washing machine or automatic dishwasher. Particularly in laundry, the water hardness (i.e., the calcium concentration) is of concern. Calcium-ions are known to stabilize some enzymes and destabilize others. Usually, sufficient amounts of enzymes are present in the powders to secure the desired detergency power.

Washing Performance Evaluations Laundry

Proteases are usually evaluated by wash trials of test pieces soiled with milk, blood, or grass. Commercial presoiled test pieces are available from a number of research and test institutes. In laboratories, washing performance studies are carried out at relevant temperatures using Terg-O-tometers, which simulate the U.S. top-loading type of washing machine, and Launder-O-meters, which simulate European drum-type machines. Launder-O-meters to some degree simulate washing mechanics and temperature profiles, including initial temperatures, heating rate, and maximum temperatures. Washing results depend on a number of factors, including detergents (Table 6), water, type of soils, textiles, washing machine (temperature, time, mechanics, and number of rinses), and the conditions mentioned in Table 7. Dose-response trials may be made in ordinary washing machines with ballast laundry and artificially soiled fabrics. A sufficient amount of standard swatches may be fixed onto the ballast cloth. Evaluation of the effects on the test pieces may be made visually or by measuring the reflectance of light remitted at 460 nm when illuminating the test pieces. The intensity of reflected light % R (% remission) can thus be measured. The value $\Delta R = R_{\text{washed}} - R_{\text{unwashed}}$ is a measure of the total detergency. $\Delta R_{\text{Enz}} = R_{\text{washed}} - R_{\text{washed without enzyme}}$ reflects the enzyme's effect.

Figure 7 is an example of a dose-response curve of Savinase (Table 1) in a European machine at 25 °C (21). Using a particular standard test textile swatch, EMPA 117 (cotton and polyester), containing blood, milk, and carbon black, Figure 7 shows that the optimal effect of the enzyme was almost reached at a dosage of about 0.2% of the Savinase 6.0 T formulation. In practice, a dosage of about 1% is used to compensate for the detergency factors mentioned earlier.

Figure 8 is an example of a dose curve made in a Launder-O-meter. At $\Delta R < 22$ the fabric was clean. This means that the maximum measured R value was almost identical to that of an unsoiled test piece. Figure 8 shows that the performance is better at 60 °C than at 40 °C until a dosage of about 3% Savinase 6.0 T (granulate) in the tested detergent was used.

Stored detergents in open packages may become humid. At storage conditions of about 55% relative humidity, the bleaches (percarbonate or perborate) become activated. After some weeks of storage the residual enzyme activity may become reduced (22). However, stability of proteases in activated bleach systems has been improved significantly in the past few years by the use of genetic and protein engineering. Using these techniques, oxidation-sensitive regions of amino acids in protease molecules have been replaced by more stable amino acids (22).

Washing Performance Evaluations: Automatic Dishwashing

A standard laboratory procedure for evaluating the cleaning efficacy of enzymes, mainly proteases and amylases, has been published (23). To test the efficiency of proteases, stainless steel plates are soiled with a baked egg–milk film. After washing light reflectance values, R, are measured directly on the plates. Calculation of percentage of removed, protein film, RPF%, is shown by the formula

$$RPF\% = \frac{(R_{\text{after wash}} - R_{\text{before wash}})}{(R_{\text{clean plate}} - R_{\text{before wash}})} \times 100\% \qquad (6)$$

The soil-removal values are affected by the temperature at which the protein film has been baked. The effect of the baking temperature is shown in Figure 9 using a citrate-

Table 6. General Composition of Detergents for Laundry and Dishwashing

Raw materials	General functions	Type of compounds	Contents (w/w%) Laundry	Dishwashing
Builders	Washing alkalis	Phosphates (sodium triphosphate)	30–60	10–40
	Water softening	Phosphonates		
	Buffer action	Na citrates		
	Stabilizers	Zeolites		
	Corrosion protection	Soda		
		Silicates		
		Complex formers (e.g., nitrilotriacetate [NTA])		
Surfactants	Emulsify particles	Surfactants	10–30	1–4
	Interfacial activity	Anionics		
	Reduce surface tension	Nonionics		
		Cationics		
Soap	Prevents overfoaming, in the machine	Soap	1–5	0
Bleach as peroxygen or activator	Oxidize polyphenolics from fruits, coffee, tea, wine, and vegetables	Na perborates	0–25	4–15
		Na percarbonates		
		Activators (e.g., tetraacetylethylenediamine [TAED])		
Enzymes	Removal of starch-based stains	Amylases	0.4–1	1–3
	Removal of protein complexes	Proteases	0.4–2	0.5–2
	Removal of fat stains	Lipases	0.2–1	0.5
	Textile color brightening, softening, soil removal, whiteness maintenance	Cellulases	1–3	0
Salt	Process aid to promote flow and loading properties (powders)	Na sulfate	Balance to 100	Balance to 100
Perfume	Adds pleasant smell to fabrics	Various esters	0.1–0.5	0–0.1
Other organics	Graying inhibitors	Carboxymethylcellulose (CMC), polycarboxylates	0.5–5	0–1
	Optical brighteners			
	Solvents	Alcohols		

Note: Data from general brochures, package information, encyclopedias, and other general literature, including Ref. 20.

Table 7. Laundry Washing Conditions and Procedures by Region

Conditions	United States and Canada	Japan	Europe
Machine type	Top loading Vertical axis Agitator	Top loading Vertical axis Impeller/pulsator	Front loading Horizontal axis Drum type
Fabric load (kg)	2–3	1–1.5	3–5
Wash liquor (L)	35–80	30–45	15–25
Wash temperature (°C)	20–50	10–30	30–95
Washing time (min)	10–15	15–35	40–60
Water hardness (ppm $CaCO_3$) (simplified)	100	50	250
Detergent dosage levels (g/L)			
Powder	1–5	0.5–1.5	5–10
Liquid	2–4	0.75–1.0	7.5–10

Note: Information from various general brochures, package information, encyclopedias, and other general literature.

based model detergent. In Figure 10, a dosage response (activity basis) for Savinase is shown for removal of protein film baked at 120 °C for 60 min.

Protease Effects for Cleaning-in-Place and Membrane Cleaning

Proteinaceous soils often contain complexes formed between proteins, fats, and carbohydrates. Thus it may be advantageous to combine enzymes to obtain a synergism during the hydrolysis of the protein. This has been demonstrated by the addition of lipase to protease in a model hydrolysis experiment.

A homogeneous freeze-dried model substrate called burned whole milk was prepared from whole milk stirred in a pressure cooker at 120 °C for 30 min. This substrate was tested for protein degradation using combinations of Esperase (Table 1) and Lipolase™ (a lipase produced by Novo Nordisk A/S). Surprisingly, it was found that when Esperase was combined with Lipolase, a synergistic effect on the release of free amino groups was found (TNBS method). Esperase used without Lipolase resulted in a

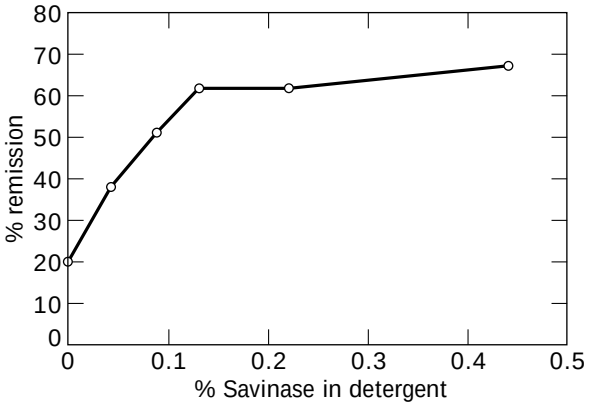


Figure 7. Dose-response curve of Savinase® in a European machine at 25 °C. Test system: test fabric, EMPA117; 4 g/L European powder detergent; 18 German degree of hardness (°dH) water hardness (equivalent to 320.4 ppm $CaCO_3$). *Source*: From Ref. 21.

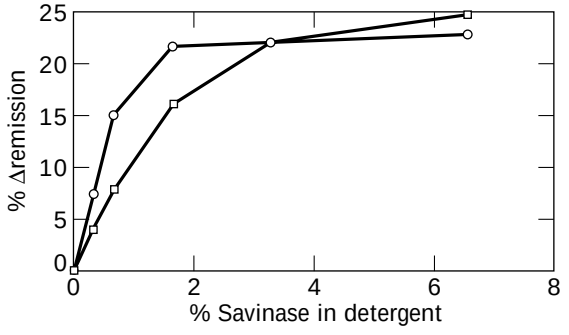


Figure 8. Launder-O-meter data of Savinase 6.0 T. Test system: EMPA 117; 4 g/L European powder detergent: —□—, 40 °C for 47 min; —○—, 60 °C for 40 min.

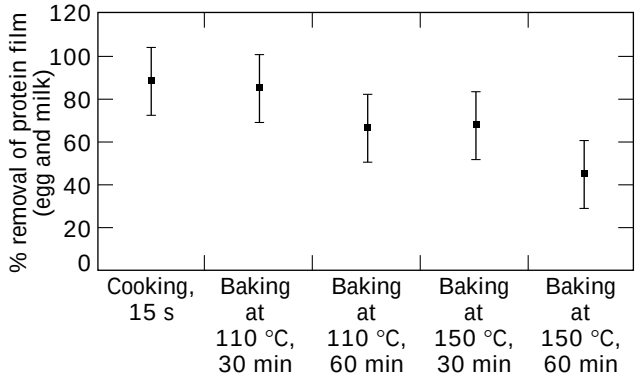


Figure 9. Effect of baking temperature on removal of protein film in automatic dishwashing. Enzyme: 0.5% Savinase 16L added to a citrate-based model automatic dishwashing detergent (ADD); 95% confidence intervals are shown based on five plates/machine; Each plate was measured at six points.

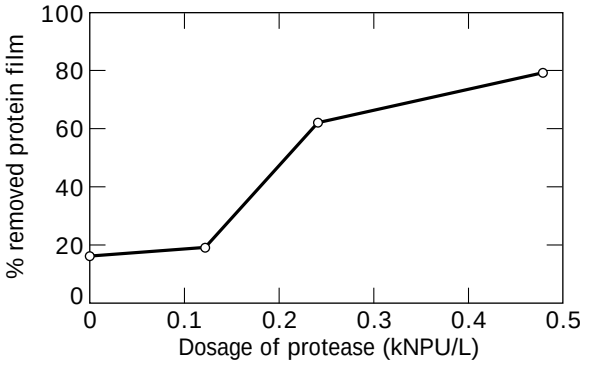


Figure 10. Dosage response (activity basis) of Savinase® at 55 °C for automatic dishwashing. Soiling system: Film of egg and milk on stainless steel plates baked at 120 °C for 60 min; detergent: European ADD–powder; enzyme: Savinase 6.0 T. kNPU is the proteolytic activity in kilo Novo protease units measured as the hydrolysis rate of dimethyl casein and compared with a standard reference protease preparation.

lower solubilization of the burned whole milk (Fig. 11). Using fresh milk, high levels of Lipolase did not greatly improve the degradation of the protein (Fig. 11).

A practical utilization of this hydrolysis synergism was demonstrated for removing burned-on soil on a heat ex-

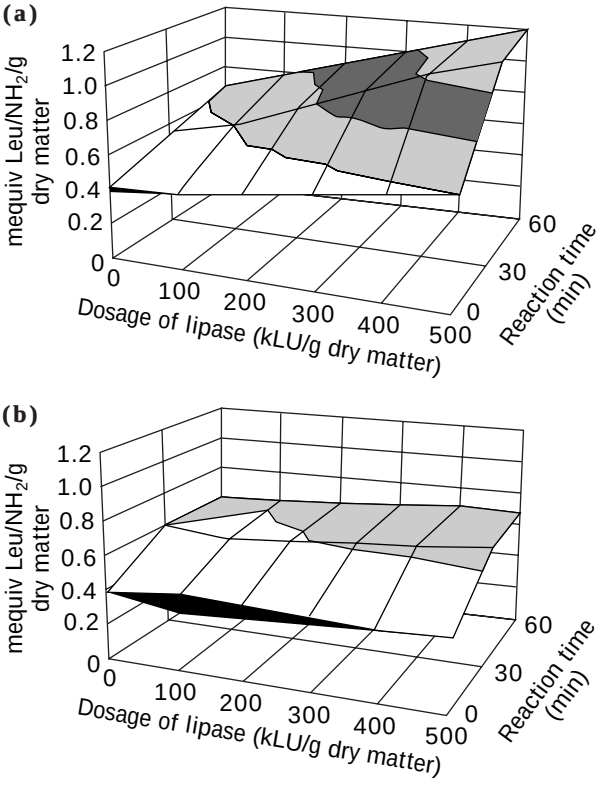


Figure 11. (**a**) Protein hydrolysis of burned whole milk assisted by use of lipase. (**b**) Protein hydrolyses of fresh whole milk. Enzymes used: Esperase and Lipolase; substrate concentration 0.5% dry matter; T, 50 °C; pH, 8.0; protease activity: kNPU/g; lipase activity: kLU/g.

changer used for high pasteurization of whole milk. A 2- to 3-mm layer of burned milk was removed using a protease and lipase at pH 8 (phosphate buffer), 50 °C for 2 h (24). Scale was removed using 0.5% nitric acid after the enzyme treatment.

During the enzyme cleaning, surfactants (soaps) were produced in situ from the fat, and emulsifiers and foaming agents were formed from the proteins. This was demonstrated by measuring the surface tension of a reaction mixture as a function of the consumed amount of base in a pH-stat trial using lipase and protease (Fig. 12). A significant reduction of the surface tension was shown.

A complete cleaning-in-place (CIP) program using enzymatic cleaning of heat exchangers is estimated to be economical and useful when there are problems with conventional chemical cleaning processes.

Membrane cleaning was tested in the laboratory and pilot plant using a pure enzyme–containing medium in dairy applications. Spiral-wound modules are very sensitive to harsh chemicals and in general not easy to clean by use of traditional cleaning products. Protease and lipase used at 40–60 °C, pH 7–10, for 10–120 min under recirculation at high speed seem to solve the problems. In dairy plants enzymatic CIP are usually carried out according to the following scheme: rinsing → enzymes → rinsing → acid → rinsing.

Understanding Enzymatic Cleaning Mechanisms

To understand the mechanism of performance of proteases during washing of cloth it is important to know the molecular binding properties of the surfaces. Issues such as adsorption and desorption of stains to textile, composition of dirt substrate, specificity of the enzyme, net charge, and isoelectric properties of the enzyme in relation to the pH value of the detergent solution must be considered (22,25). Research within this field may involve activity tests on soluble and insoluble substrates and wash trials using test textile swatches with common stains, such as milk, blood, grass, and egg. Detergent proteases hydrolyze the insoluble proteins in stains into fragments that can be easily removed or dissolved when assisted by the detergent. Larger peptide-protein fragments from hydrolysis with specific proteases could be removed more efficiently from the surface than small hydrophobic peptides (22). The simple analytical tools mentioned earlier may be used for evaluation of the protein hydrolysis reaction during a washing process (26). The one-by-one reaction, described earlier, may, together with the absorption theory, be part of the explanation of why the surfaces may be cleaned in a short time (15–45 min).

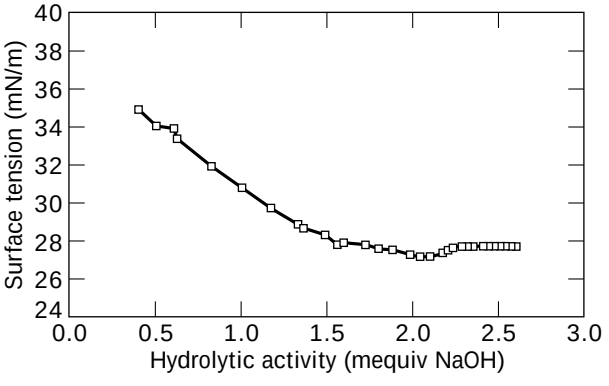


Figure 12. Surface tension versus hydrolytic activity using protease and lipase on whole milk (0.4% dry matter) at 50 °C and pH 8.0.

Protein Hydrolysates for Detergents

Enzymatically produced protein hydrolysates made by processes similar to those described in the following can, in combination with anionic surfactants, modify viscosity, foaming ability, foam stability, and cleaning efficiency for fat on textile fabric (27). Such protein hydrolysates have been acylated (condensated) with fatty acids of various lengths (28,29) and have been found to be powerful amphiphiles in the sense that they can lower surface and interfacial tension, enhance wetting and foaming ability, enhance skin compatibility, and reduce viscosity (28). Such products are produced from wheat gluten and are used in shampoos and cosmetics (30). Enhanced emulsifying properties were demonstrated in laundry and cosmetic products, such as bath soap, conditioning shampoo, and cream hand cleanser, when enzymatic soy protein hydrolysates were used to partly replace detergent components (31,32).

N-acyl protein hydrolysates made from enzymatic hydrolysates of keratin by acylation with fatty acid chlorides may be used as detergents during scouring of raw wool or other wool-cleaning operations or as fiber-protective agents during wool dyeing and bleaching (33).

HYDROLYSIS OF PROTEINS FOR FOOD PROCESSING

Hydrolysis of proteins with enzymes is an attractive means of obtaining better functional and nutritional properties of food proteins from vegetable origin or from by-products (e.g., from slaughterhouses). Some important properties of proteins and their applications in foods are shown in Table 8. Industrial baking of biscuits, conversion of milk to cheese, and fermentation of soy sauce are examples of applications of proteases to produce food.

Proteases in Baking

The proteins of wheat flour form a network with disulfide bridges. The structure of gluten becomes stronger as more

Table 8. Functional Properties of Proteins in Foods and Their Applications

Property	Application
Emulsification	Meats, coffee whiteners, salad dressing
Hydration	Dough, meats
Viscosity	Beverages, doughs
Gelation	Sausages, gel desserts
Foaming	Toppings, meringues, angel food cakes
Cohesion binding	Textured products, dough
Textural properties	Textured foods
Solubility	Beverages

sulfur bridges are present. Oxidizing agents, such as bromate and dehydroascorbic acid, strengthen the dough in this manner. By hydrolysis of the gluten fraction of wheat, neutral bacterial endoproteases may weaken a flour that is too strong or may provide the plastic properties required in a dough to be used for biscuits. Figure 13 shows a simplified model of the gluten network with disulfide bridges and the possibilities for chemical and enzymatic modification, which have been further discussed elsewhere (34). In biscuit manufacturing, either sodium meta-bisulfite or a neutral bacterial protease, such as Neutrase® (Table 1) is used to weaken the gluten, so that the biscuits, cookies, or crackers keep their shape and size.

Dairy Products

Milk is clotted to cheese by the application of proteases. Proteases of various kinds are used for acceleration of cheese ripening, for modification of the functional properties, and for modification of milk proteins to reduce allergic properties of cow milk products.

Production of Cheese. For industrial cheese production, the milk must be standardized with respect to protein and fat content. This is normally achieved by blending milk batches, skimming (by centrifugation) to remove fat, or adding fat. The standardized milk is normally pasteurized at 72 °C for 15 s. Subsequently, the milk is cooled to 30 °C before being transferred to the cheese tank. Starter culture is added to the pasteurized milk to initiate fermentation. Rennet (the milk-clotting enzyme) and calcium chloride are added to promote the milk protein clotting reaction.

The milk-clotting effect of rennet is due to a limited hydrolysis of the κ-casein surrounding the protein micelles. Thereby, the micelles become electrostatically discharged and able to aggregate with the help of calcium and phosphate ions and can form a network entrapping the fat micelles. A gel structure is thus formed.

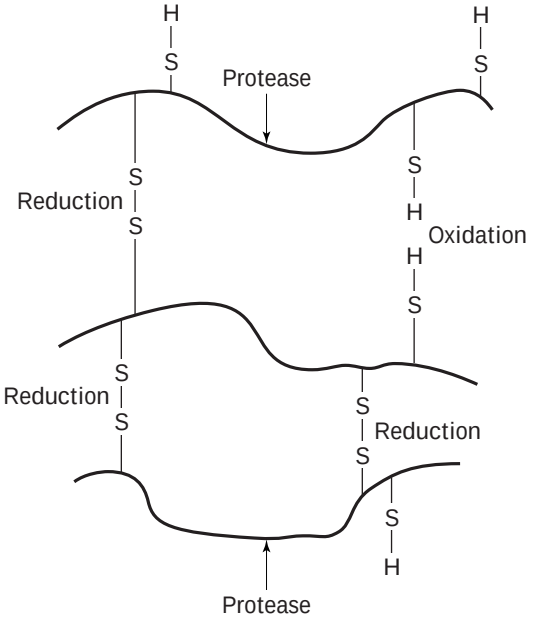


Figure 13. Model of gluten network and chemical and enzymatic modifications.

After about 25 min at 30 °C, the gelled milk (the coagulum) is cut or stirred to promote syneresis, or exudation of the whey. Syneresis is further promoted by heating the curd–whey mixture. Specialized equipment for drainage of the whey is used. After drainage, the cheese is pressed, and lactic acid fermentation continues during pressing and subsequent storage. The cheeses are then steeped in brine. The fermentation of lactose continues until all lactose has been fermented, so that after removal of the brine the cheese is free of lactose. The cheese is then ripened during an appropriate period of storage.

Chymous Enzymes. Rennet, also called rennin, is a mixture of chymosin and pepsin. Rennet is extracted from the gastric mucosa of the abomasum of young mammals (e.g., calves and lambs). The pepsin-to-chymosin ratios differ in different rennet preparations because chymosin is present only in the stomachs of unweaned mammals and is later replaced by pepsin. Thus, the content of pure chymosin depends on the age and species of the animal. Chymosin rapidly splits κ-casein into a glycomacropeptide and *para*-$\kappa\kappa$-casein by selectively cleaving the 105–106 bond between phenylalanine and methionine. Pepsin and other proteolytic enzymes are less specific and can give rise to a number of degradation products, which tend to taste bitter (35).

Rennet Substitutes. Microbial rennets from a number of producers are marketed under various trade names (e.g., Fromase®, Hannilase®, Marzyme®, and Rennilase®). The enzyme costs for treating the same volume of milk are considerably lower than using standard rennet (36).

Microbial rennets are produced by submerged fermentation of selected strains of fungi, often *Rhizomucor miehei*. The properties of microbial rennets have proven similar to those of chymosin, and only slight modifications of the cheese-making technique are made in practice (37).

The gene for calf chymosin has been cloned into selected bacteria, yeasts, and molds. Chymosins made with recombinant DNA techniques on *Kluyveromyces lactis* (Gist Brocades), *Escherichia coli* (Pfizer), and *Aspergillus nidulans* (Hansen) are now commercially available (38).

Extraction or Modification of Protein Substrates

Protein structure can be modified to improve solubility, emulsification, and foaming properties. Chemical modification is not desirable for food applications because of harsh reaction conditions, nonspecific chemical reagents, and difficulties of removing residual reagents from the final product. On the other hand, enzymes provide several advantages, including fast reaction rates, mild conditions, and, most importantly, high specificity. There is also a need for gentle methods that modify the food itself, in order to limit the use of additives. Over the years many different protein raw materials have been used, with different objectives. Examples of extraction processes with enhanced yields include production of soya milk, recovery of scrap

meat, cleaning of bone from slaughterhouses, recovery of gelatin, and production of meat (flavor) and yeast extracts. Furthermore, processes such as evaporation of fish or meat stickwater, rendering of fat, deskinning of fish roe, and tenderizing of meat are assisted by proteases. Functional food ingredients from protein sources are made by use of proteases, including soluble protein hydrolysates used for nutrition, and foaming and emulsifying ingredients. Examples of such products are isoelectric soluble soy protein (ISSPH), egg white substitute of soy protein, emulsifier of soy protein, soluble wheat gluten, foaming wheat gluten, blood cell hydrolysate, whey protein hydrolysates, casein hydrolysates, soluble meat proteins, and gelatin hydrolysates.

Functional Protein Hydrolysates

A limited hydrolysis of soy protein isolates was made at a very early stage with *Aspergillus oryzae* protease at varying ratios of enzyme to substrate (39). The *DH* was not directly measured, but from the data given by Puski in 1975 (39) the maximum *DH* value obtained was about 5%. It was shown that whipping and emulsifying capacities were improved.

The first application of the pH-stat technique for a controlled production of enzymatically hydrolyzed soy proteins was by use of Alcalase at pH 8 and Neutrase at pH 7 (Table 1) (40). The isoelectric solubility, the foaming capacity, and the emulsifying capacity were all increased. If the reactions were continued to *DH* values higher than 5%, the foaming and emulsifying capacities again decreased. Also, a significant difference of the functional capacities was observed, depending on which of the two different proteases was applied. Alcalase had a much better performance, in general. The reason for this effect could be due to the fact that foams and oil emulsions are stabilized better by peptides having a higher content of nonterminal hydrophobic amino acids than by peptides having terminal hydrophobic amino acids. Figure 14 shows the whipping expansion versus *DH* from the previously mentioned early demonstration.

In further research on functional ingredients based on enzymatic modification, improved products were developed with other raw soy materials or when physical treatments, such as ultrafiltration, were used during purification of the active components responsible for the specific effect. When a soy protein concentrate was hydrolyzed to *DH* 3–5%, with Alcalase, a functional protein hydrolysate with high foam expansion and remarkably good foam stability was obtained (41). On the other hand, the emulsifying capacity was considerably lower than if soy isolate was modified the same way.

Using a native soy protein isolate produced by ultrafiltration of an aqueous extract of defatted soybean meal (42,43), the shape of the hydrolysis curve was remarkably different compared to that obtained for acid-precipitated protein, as seen in Figure 15. The hydrolysis curves are shown for acid-precipitated soy protein isolate (denatured and nearly insoluble in water) and ultrafiltered soy protein isolate (native and soluble in water). From Figure 15 it appears that the ultrafiltered protein isolate is hydrolyzed

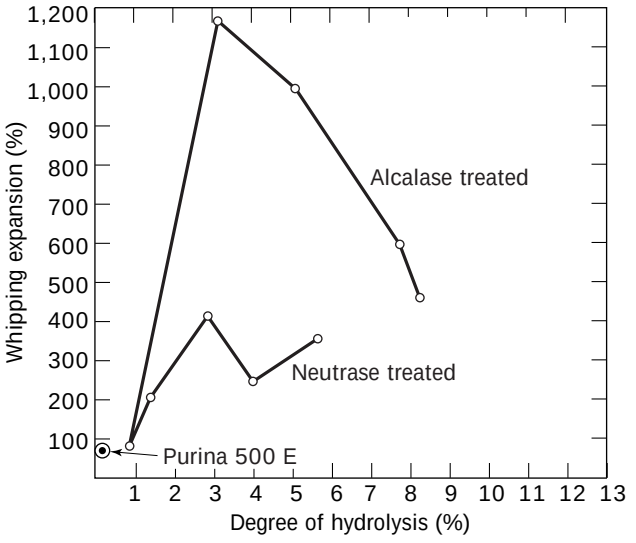


Figure 14. Whipping expansion as a function of the degree of hydrolysis (DH) for soy protein hydrolysates. *Source*: From Ref. 40.

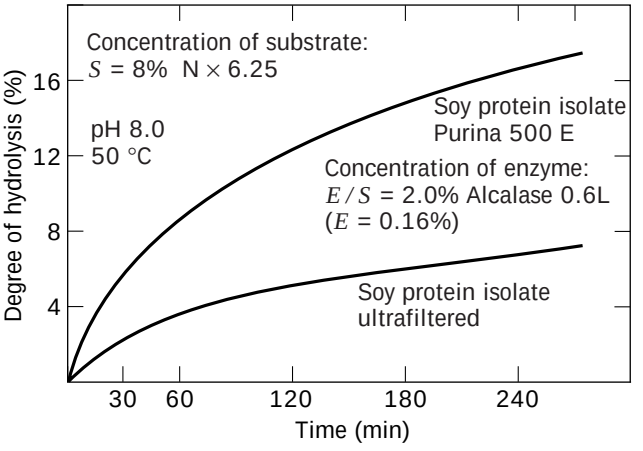


Figure 15. Hydrolysis curves for soy protein isolates. *Source*: From Ref. 43.

more slowly than the acid-precipitated protein. This is due to the compact molecular structure of the ultrafiltered protein, which is still in its native stage. The degree of denaturation of a protein substrate has a profound influence on the kinetics of proteolysis, as has been known for a long time (44). Using this soy protein isolate, functional protein hydrolysate with better foam expansion and better foam stability was obtained compared to those from the acid-precipitated protein (Fig. 14). If the small peptides in this hydrolysate were removed by ultrafiltration, the foam stability was improved even further. The use of ultrafiltration before or after the enzymatic hydrolysis was investigated in pilot-plant scale in a number of process combinations (43,45). The high molecular weight hydrolysates had excellent whipping properties; furthermore, they were also

able to heat-coagulate, so that they could substitute for egg white 100% on a protein basis in a meringue batter.

Enzymatic Hydrolysis of Soy Protein Concentrate in Alcohol. Soya protein concentrate is produced by leach of defatted soybean grits using a water and alcohol mixture of 20–40% H_2O in ethanol. By this process the protein denatures so that the protein solubility is reduced and does not function adequately. A way to increase the solubility might be to enzymatically hydrolyze the separated protein before the alcohol is stripped off. To test if a protease such as Alcalase® could operate in relatively high ethanolic environments, a series of hydrolysis curves were made under various conditions. In Figure 16, pH-stat hydrolysis curves are shown for a soy protein concentrate at various concentrations of ethanol. Figure 17 shows the effects on the protein solubility index (PSI%) of percent ethanol, temperature, and percent DH during hydrolysis. For Figure 17, freeze-dried samples were produced to the percent DH as indicated. The alcohol reduced the enzyme's effect considerably, but even when the percent DH was low, an improvement could be obtained on the protein solubility index. The practical utilization of these data could be to add water to an alcohol-precipitated protein sediment from a soy concentrate factory and to perform a protein hydrolysis before drying. A product having improved functional properties may thereby be produced.

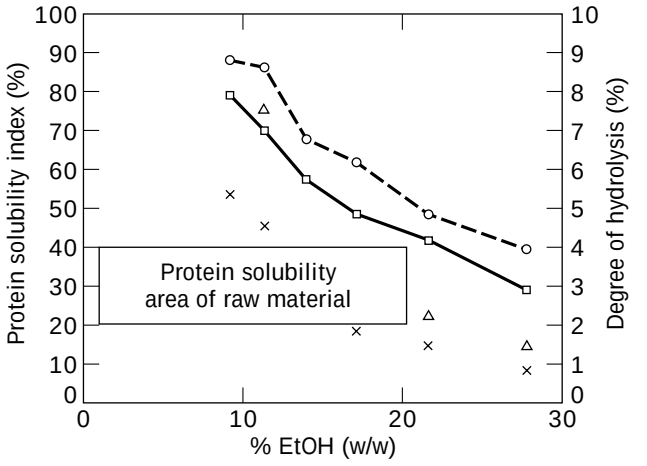


Figure 17. Enzymatic hydrolysis of soya protein concentrate in alcoholic solution. Protein solubility index and degree of hydrolysis at two different hydrolysis temperatures. —○—, 57 °C; —□—, 62 °C; △, % DH at 57 °C; x, % DH at 62 °C.

Practical Aspects of Enzymatic Processing

Enzymatic processing methods have become an important area for new processes used by the food industry to produce a large and diversified range of products based on protein raw materials. The principal advantages offered by enzymatic processes are the enzymes' specific mode of attack and the mild temperature and pH conditions required. Gentle treatment of the raw material is achieved, and only a few by-products that may affect the taste and smell of the final products are produced.

Considering a typical enzyme application in details, several effects influence the enzyme reaction and the final product. The scheme shown in Figure 18 shows a selection of parameters that can be varied when optimizing and developing new enzyme applications.

Inactivation of the enzymes is of particular concern. In most enzymatically treated food products, active enzyme is undesirable because of changes that could appear in the final product. Therefore an efficient method for irreversible inactivation of the enzyme must be developed. Usually an efficient heat treatment is used, but neutral and alkaline

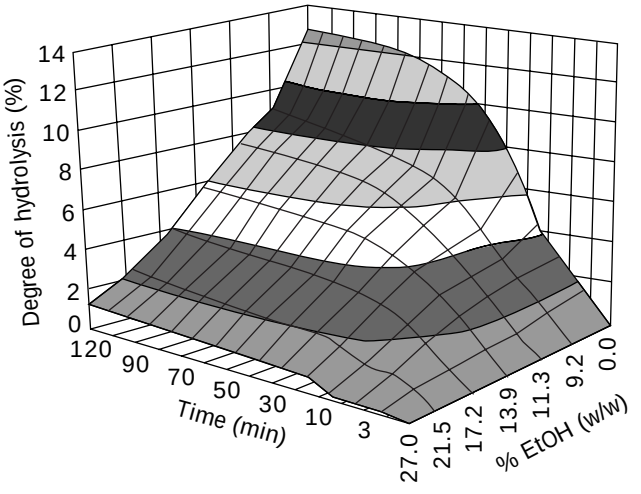


Figure 16. Percentage DH measured in pH stat by use of Alcalase® on soy protein concentrate versus alcohol concentration. Enzyme, Alcalase 2.4 L; *E*/*S*, 2.0%; pH 8.0 (pH stat); temperature, 55–60 °C; substrate, soya concentrate; danpro H; S = 8.0% protein; EtOH concentrations, w/w: 0, 9.2, 11.3, 13.9, 17.2, 21.5, 27.7.

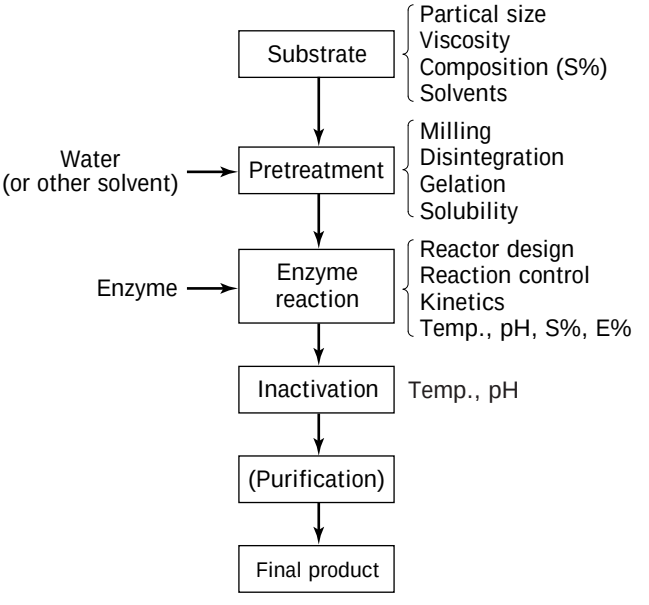


Figure 18. Practical aspects of enzyme applications.

proteases are also effectively inactivated when pH is adjusted to 4 or below.

Engineering aspects with regard to the up-scaling of inactivation must be taken into account for enzyme treatments carried out in production scale. Examples of inactivation procedures are discussed in the following. Purification of the products made enzymatically in which the enzymatic method does not produce the product in situ may in most cases be wanted. The selection of unit operations and the understanding of the performance of downstream processes are therefore also important parts of the application technology.

EXAMPLES OF APPLIED PROCESSING USING PROTEASES

Hydrolysis Processes and Simultaneous Inactivation of the Protease

If the reaction is carried out at a temperature above the enzyme's denaturation temperature, the protease is inactivated simultaneously during the hydrolysis (46). Examples of processes using this principle are shown for modification of wheat gluten and, as an alternative to rendering, for upgrading fresh bones from slaughterhouses to valuable products. The meat protein hydrolysates from the latter process are used as functional ingredients in meat products or as a clear, soluble meat extract for soups and seasonings. The cleaned bones make an excellent raw material for production of gelatin (46).

Modification of Wheat Gluten. Figure 19 shows examples of hydrolysis curves of wheat gluten when using a pH stat at 72.5 °C. After a certain time, which seems to be dependent on the enzyme dosage used, the curves end at a plateau, showing that no further hydrolysis is occurring. Two product types from wheat gluten are considered.

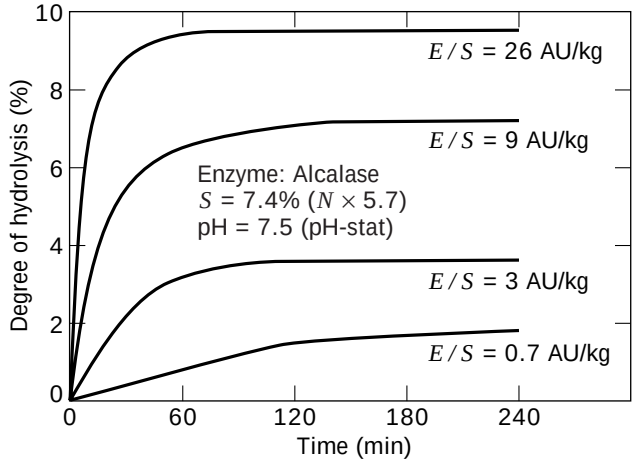


Figure 19. Enzymatic hydrolysis of wheat gluten at 72.5 °C. One Anson unit (AU) is the amount of enzyme that under standard conditions digests hemoglobin at an initial rate, liberating per minute an amount of TCA soluble products that give the same color with phenol reagent as 1 mequiv of tyrosine. *Source*: From Ref. 47.

The first is soluble wheat gluten. High protein solubility may be obtained at higher degrees of hydrolysis, as shown in Figure 20. A *DH* value of about 10% results in solubilization of more than 90% of the gluten. A 100% soluble, bland-tasting wheat gluten hydrolysate with high yield may be recovered by centrifugation and concentration.

A completely soluble wheat gluten hydrolysate is desirable for application as protein enrichment of food (e.g., soft drinks). Also, such protein hydrolysates may be excellent raw material for production of enzymatically produced hydrolyzed vegetable protein (HVP) for flavoring using, for example, exopeptidases. Nonfood applications were mentioned earlier.

The second type of wheat gluten is whipping wheat gluten. In Figure 20 whipping expansions are shown for wheat gluten hydrolysates made at different *DH* percentages. The optimal whipping properties have been found at a degree of hydrolysis of 2–3%. The active whipping protein is recovered by centrifugation and drying. A whipping protein product is required for baked goods and for different types of candy.

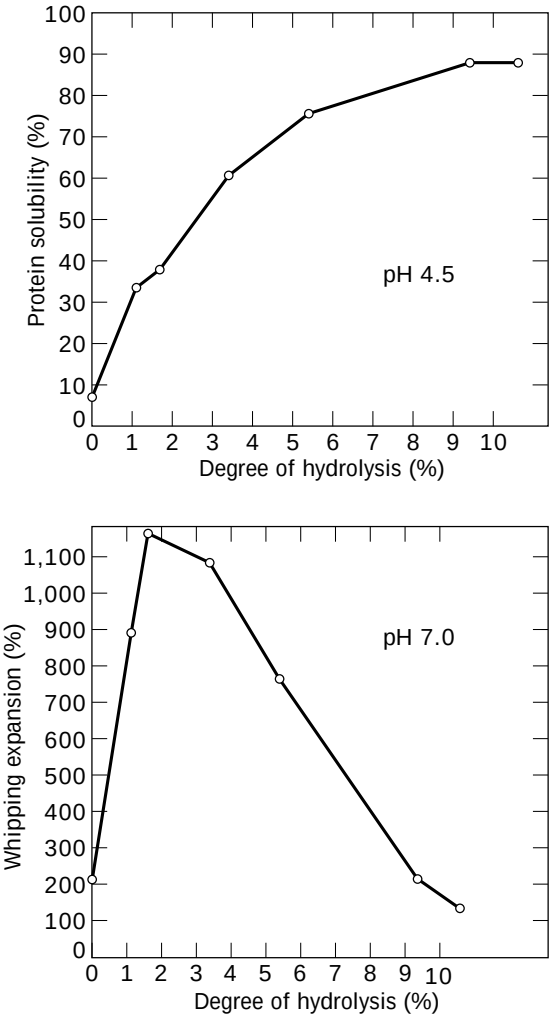


Figure 20. Protein solubility (top curve) and whipping expansion (bottom curve) for high-temperature enzyme-hydrolyzed wheat gluten.

Bone Cleaning. A two-step enzyme process in which functional meat extracts are produced by the first enzymatic extraction and cleaned bones by the second enzymatic extraction can be performed (46).

The fresh bone material (beef or pork) is crushed and mixed with hot water. The enzymes Neutrase, Alcalase, Esperase, and Flavourzyme, used alone or in combination, are added, and the slurry is agitated and extracted batchwise or continuously at 60–75 °C at neutral pH (6.5–7). Poultry material may also be used, but the bone material is not in general used for gelatin production.

Bone cleaning trials have been carried out in Novo Nordisk's pilot plant for process development. Using a bone crusher, 250 kg of beef bones were milled to a particle size of approximately 20 mm. This material was poured into 250 kg of 80 °C water and the temperature was adjusted to 65–70 °C. Hereafter the enzymatic hydrolysis reaction was carried out under powerful stirring. Samples, drawn as a function of time, were centrifuged for 3 min before determination of °BRIX (dry matter), osmolality, and pH. The total reaction mixture was sieved and the liquid phase was pasteurized, defatted by centrifugation, concentrated, and dried. The bone material was washed and dried. Hydrolysis data recorded from such a trial using Esperase are shown in Figure 21

Within the 20 min of reaction the °BRIX value and the osmolality were still increasing. This indicates that the enzyme was not completely inactivated at this temperature (− 70 °C). The value mOSM/°BRIX is a measurement of the molecular size of the soluble phase. In the beginning, when the content of soluble protein is low, mOSM/°BRIX is high because of a high content of low molecular weight compounds. Later, this value stabilizes because no further degradation occurs when the protein is solubilized from the surface of the bone material (Fig. 21). The one-by-one reaction system seems to be confirmed.

Inactivation of the enzyme Esperase may simultaneously be achieved at a higher temperature (approximately 75 °C). Centrifuging at 90 °C during defatting of the hydrolysate inactivates the enzyme and pasteurizes the product.

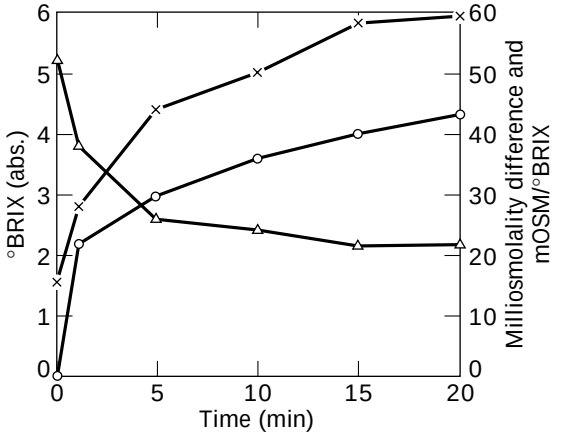


Figure 21. Data from a pilot-plant bone-cleaning trial. Enzyme, Esperase 7.5 FG; dosage, 2.0 kg/t bone —x—, °BRIX; —○—, mOSM (diff.); —△—, mOSM/°Brix; T, 70 °C; pH, 6.7–7.1.

Hydrolysis of Meat Protein in Relation to Bone Cleaning. Based on the osmolality difference, ΔC in equation 4, *DH* can be calculated and compared with those values from hydrolyses of lean beef. Hydrolyses using Esperase were carried out at 55, 65, and 75 °C on minced beef (raw or cooked). Both pH-stat trials and hydrolyses followed by osmometer were done at temperatures in the high region. The pH was adjusted to 7.5 so there was no appreciable hydrolysis of triglycerides. Figure 22a shows the hydrolyses curves and corresponding protein solubility for the osmometer trials. The percent *DH* was higher for raw meat, and the highest *DH* value for raw meat was obtained at 55 °C. For both the raw and cooked meat the highest solubilization was found at 65 °C (Fig. 22b). This was also seen in pH-stat trials. In Table 9 the hydrolysis data are compared for the bone cleaning trial (Fig. 21) and for hydrolysis of raw meat. The enzyme dosage based on protein was doubled for the bone trial as compared to the meat trial, yet a high protein-solubilizing effect of Esperase was seen in both cases. But, surprisingly, a lower percent *DH* secured a high solubility for the raw meat than for the meat on bones.

Use of Membranes in Protein Hydrolysis Processes

Membrane processes are used within many areas of industrial enzymatic hydrolysis of proteins. Table 10 shows membrane processes can be applied for pretreatment of proteins, for both the reaction step and the purification step.

Functional Protein Hydrolysates. Examples of the application of ultrafiltration processes in various process combinations for production of highly functional protein hydrolysates of soy protein were mentioned earlier. Principally, the technique of removing small peptides from total hydrolysates may be used on many substrates other than soy. Thereby, excellent foaming and coagulating protein products may be isolated (purified). The first process step after inactivation is normally a centrifugation carried out to remove insoluble material, which may reduce the whipping expansion and the foam stability. To purify the supernatant as effectively as possible, ultrafiltration can be carried out in three steps. In the first step, the concentrate reaches a protein content of about 10%. Then, a diafiltration process is performed for washing out small, low-function peptides. A final ultrafiltration can bring the protein content up to approximately 20%. Finally, the concentrate can be spray-dried. The efficiency of these processes and the overall yields are dependent on the actual ultrafiltration equipment used, the membrane cutoff value, and the character of the hydrolysate (enzyme used, *DH*%, pH, protein type) (43,45).

Low Molecular Weight Hydrolysates. Low molecular weight hydrolysates are produced from raw materials that are hydrolyzed to a high degree of hydrolysis. Examples of such products are the mild-tasting decolorized BCH developed from the hygienically collected red cell fraction of slaughterhouse blood (13) and the isoelectric soluble soy protein hydrolysate (ISSPH) (48). Flowcharts for these

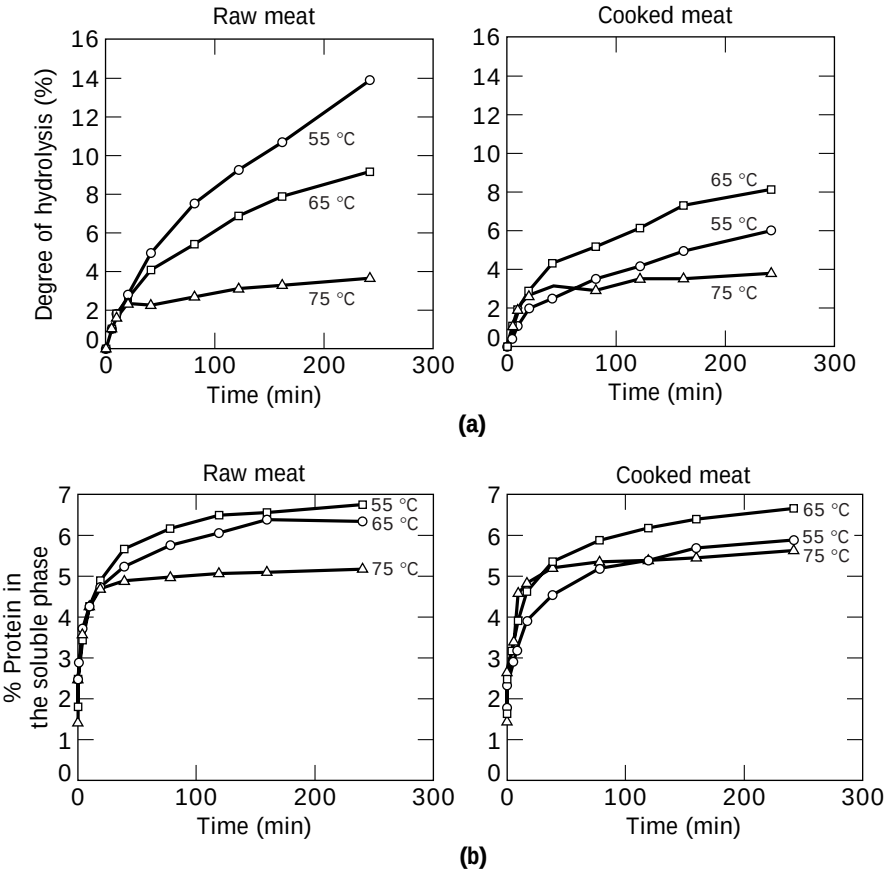


Figure 22. (**a**) Hydrolyses curves (osmometer) at 8% beef, protein. enzyme, Esperase 7.5 FG, dosage 0.5% of protein, —○—, 55 °C; —□—, 65 °C; —△—, 75 °C. (**b**) Solubilization of protein by hydrolyses of beef; meat, enzyme, Esperase 7.5 FG; dosage, 2.0 kg/t bone.

Table 9. Bone or Meat Treated with Esperase at 65–70 °C

	Bone			Raw meat		
Time (min)	% *DH*	pH	% Soluble protein	% *DH*	pH	% Soluble protein
0	0	7.14	1.2	0	7.5	1.8
1	5.7	6.82	2	0.1	7.33	2.5
5	7.7	6.81	3.2	1	7	3.4
10	9.3	6.67	3.7	2.1	6.85	4.1
15	10.3	6.67	4.2	2.5	6.77	4.5
20	11.1	6.68	4.3	2.8	6.71	4.8

hydrolysates are shown in Figure 23 (BCH) and Figure 24 (ISSPH). These two examples demonstrate how the membrane processes can be used for different purposes.

In the BCH process, the pH is 4.2 and insoluble hem product may be removed by centrifugation or ultrafiltration. Comparison of the centrifugation process and the ultrafiltration process shows that the ultrafiltration process was most economical when new plants were installed (13,49). Ultrafiltration equipment from various suppliers has been tried with technical success.

Application Trials Using Blood Cell Hydrolysate. The earliest application trials were for pumping of meat cuts. Us-

Table 10. Application of Membrane Processes during Enzymatic Modification of Proteins

Types of enzymatic modified proteins	Pretreatment	Enzymatic reaction step	Posttreatment
Highly functional proteins	Production of native protein isolate by ultrafiltration		Molecular separation by ultrafiltration
Low molecular weight protein hyrolysates		Membrane reactors by ultrafiltration	Concentration or desalination by hyperfiltration or nanofiltration

Source: From Ref. 43.

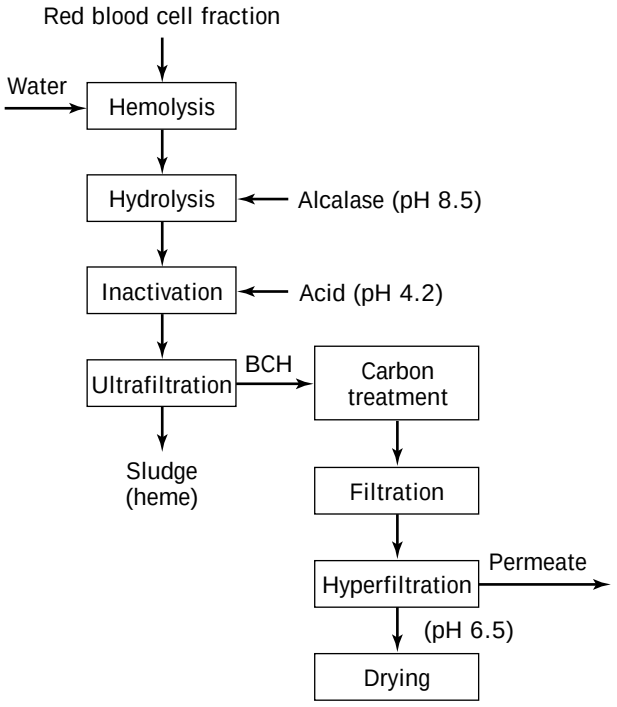


Figure 23. Flowchart for production of blood cell hydrolysate.

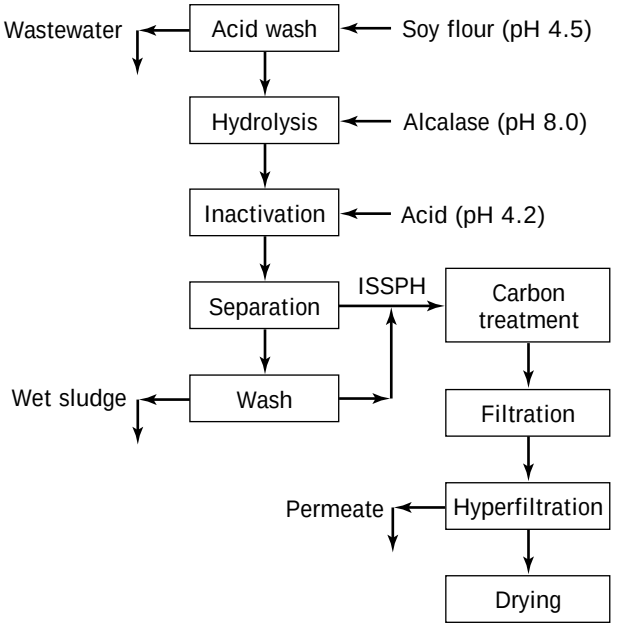


Figure 24. Flowchart for production of isoelectric soluble soy protein hydrolysate (ISSPH).

ing up to 3% BCH on a final meat product basis, nutritional protein equivalence was obtained, yields were increased, costs were reduced, toughness of meat was reduced, and the fat to protein ratio was reduced.

In a study using BCH in low concentrations in meat, it was demonstrated that BCH has a polyphosphate-like character in water-added meat systems. BCH has also been compared to sodium triphosphate in luncheon meat heat-treated at 75 °C and 108 °C, respectively (50).

During production of the soy protein hydrolysate (Fig. 24), the clear filtrate after carbon treatment is concentrated by hyperfiltration as an alternative to evaporation. A falling film evaporator can concentrate the hydrolysate to above 50% dry matter. Hyperfiltration can only be used efficiently up to about 25% dry matter but allows the simultaneous removal of salt if nanofiltration membranes are used for the concentration. Thus, about 75% of the salt in ISSPH was removed by direct concentration and 93% if a diafiltration was included in a pilot-plant trial on a selected cellulose acetate membrane (43).

Membrane Reactor. The membrane reactor is an ultrafiltration system in which a high concentration of hydrolytic enzyme is confined. High molecular weight substrate is fed continuously to the reactor, and the low molecular weight products are removed simultaneously as they permeate through the membrane. Ideally, a steady state is reached in which the degradation of substrate is carried out indefinitely with high efficiency and negligible loss of enzyme. A considerable number of publications on the use of membrane reactors for enzymatic hydrolysis of proteins appeared 10–15 years ago. However practical use of the system has not been widespread. One of the reasons may be that the economical advantage of the membrane reactor depends critically on the need for purging of the reactor. A small part of the substrate is always nondegradable. Therefore inert material will rapidly build up in the reactor, causing mechanical problems. A considerable purge is thus necessary if the concentration of this material is to be kept at a reasonably low level. This has a drastic, negative influence on the yield of product. Also, the enzyme loss during purging is considerable. Kinetic investigations of protein hydrolysis in the membrane reactor have been presented and a mathematical model established (43). Based on these calculations, it was found that the concomitant loss of enzyme in certain cases may lead to an overall enzyme consumption close to that of the batch hydrolysis process.

The Bitterness Problem

Unpleasant bitterness has been observed for some protein hydrolysates. This problem can be managed by proper selection of the hydrolysis parameters and the enzymes used. Bitterness is a complex problem that can be influenced by the following variables (4,51):

1. Hydrophobicity of the substrate, hydrolysates become bitter if amino acid side chains containing hydrophobic groups become exposed due to hydrolysis of the protein (52)
2. High *DH* and, therefore, the concentration of soluble hydrophobic amino acids and their peptide chain lengths
3. The specificity of the protease, to some extent, determines the bitterness

4. Separation steps applied in downstream processing, for example, centrifuging at isoelectric pH and carbon treatment
5. Masking effects from other components in the protein hydrolysate; organic acids such as malic acid and citric acid exert a significant masking effect in soy protein hydrolysates (53) that can be utilized in the final formulation of the food, where the hydrolysate is incorporated

In addition, recent recognition of the following effects points out other solutions to solve bitterness problems in production of protein hydrolysates with high *DH* values. A debittering of protein hydrolysates can be performed by a selective extraction or removal of hydrophobic peptides to a 2-butanol phase (4,51). This knowledge can be used to explain why it was surprisingly possible to produce a nonbitter *DH* 20% soy protein hydrolysate from full-fat soybean meal using Alcalase (54). This *DH* 20% hydrolysate was fully acceptable from an organoleptic point of view. The oil served as the hydrophobic extraction medium and eliminated the bitter peptides. The oil was simultaneously recovered by centrifugation.

Reaction parameters such as pH and temperature, by which the specificity of the proteases are changed, also play an important role in managing bitterness. Furthermore, combinations of enzymes can change the overall mode of degradation of a protein so that bitterness is reduced. The application of exopeptidases is a generally recognized way to remove bitterness of high-*DH* hydrolysates.

Other Off-Flavor Problems in Relation to Bitterness

Animal raw materials considered useful for protein hydrolysis usually do not contain many in situ bitter compounds of significance. Bitterness can appear from rancid traces of fat or from extended hydrolysis of proteins.

Oilseeds contain a variety of phenolic compounds, in particular, bitter-tasting phenolic acids (55). Some volatile compounds are hydrophobic and bind to the protein phase, especially to denatured protein. Off-flavors may be released by enzymatic hydrolysis of such proteins, some of which, but usually not all of which can be removed. A solution to such problems has been described for industrial production of ISSPH (48). The white flakes were washed at pH 4.5 with water in a four-step extraction process to produce a soy concentrate with a low level of beany off-flavor. This soy concentrate was immediately hydrolyzed with Alcalase to a specific degree of hydrolysis. The hydrolysis reaction was terminated by lowering the pH, and the supernatant was recovered by centrifugation at the isoelectric point. The hydrolysate was carbon treated, concentrated (Fig. 24), and used for protein-fortified fruit juices and for extension of cured, whole meat products. This was possible because the ISSPH had a bland taste and no bitterness.

Enzymatic Tenderization of Meat

The use of the vegetable proteinases papain, ficin, and bromelain for enzymatic tenderization of meat has been known for many years. These enzymes have a relatively strong potency on muscle tissue components such as collagen and elastin. Microbial proteinases have also shown the ability to tenderize meat, but not to the same extent as the vegetable proteinases.

The most successful meat, tenderizing enzymes are those having the ability to hydrolyze the connective tissue proteins as well as the proteins of the muscle fibers. The enzymes are applied to the meat by sprinkling the enzyme powder on a thin slice of meat, dipping the meat in an enzyme solution, or injecting an enzyme solution into the meat. A method has been developed in which the enzyme is introduced directly into the circulatory system of the animal shortly before slaughter (56) or after stunning the animal to cause brain death (57).

Whatever method is used, the difficulty in using external enzymes to obtain an even distribution of the enzyme throughout the tissues seems to be in the control of the hydrolysis reaction.

Production of Flavor Enhancers

Proteins in their natural state do not contribute chemically to formation of flavor in foods. The products of the hydrolysis of proteins, peptides, and amino acids are much more reactive, as are other breakdown products in general.

Hydrolyzed protein, mainly produced by the use of hydrochloric acid, is one of the largest contributors to flavoring produced from protein. Using various sources, many savory products are on the market. The hydrolysis with hydrochloric acid usually takes place in the presence of triglyceride fat material. It is well known that chloro- and dichloropropanols are formed during acid hydrolysis under harsh conditions, although they are found in small quantities in the final food products. Concern for the safety of the products of hydrochloric acid hydrolysis of proteins has led to investigations into the safety of these chlorocompounds (58).

Glutamic acid, as monosodium glutamate (MSG), is by far the most consumed single flavor enhancer of proteinogenic origin. About 270,000 tons are produced per year (59). Glutamates are known as the fifth basic taste (in addition to sweet, sour, salty, and bitter). This taste is called *umami* by Japanese and savoriness by Americans. MSG is used at concentrations of 0.2–0.8% in a variety of foods, such as soups, broths, sauces, gravies, flavoring and spice blends, canned and frozen meats, poultry, vegetables, and finished dishes. To be able to reduce the intake of sodium and to produce in situ flavor enhancers in protein hydrolysates, peptidoglutaminases are of interest for industrial purposes (60–62).

Through their reactions with taste sites, peptides may give flavors that are bitter, sweet, salty, or *umami*. Also, sour taste and astringency have been attributed to peptides isolated from protein hydrolysates or by synthesis. Many studies have been made of the effects of flavor in dairy products, meat and fish products, and yeast extracts. The direct production of flavor products can be made with proteases used alone, or in combination with fermentation processes. Chemical hydrolysis of protein and carbohydrates combined with heat treatments form Maillard reactions that allow many flavors to be produced. Soy and

wheat hydrolysates are used for flavoring and flavor enhancement of soups and seasonings; meat hydrolysates for addition of flavor to meat products, soups, bouillon, and sauces; and milk and cheese hydrolysates for production of cheese flavors (62).

Flavors from Hydrolyzed Vegetable Proteins

Enzymatic hydrolysis usually leads to products that are light in color and have a much less pronounced meaty or savory flavor than, for example, yeast extracts (34). Meat flavor and similar condiments are often produced from many sources, and their production often involves a natural microbial fermentation. Among these can be included the Asian products based on soy protein. Wheat and maize gluten are also hydrolyzed to flavor intense peptides.

The content of free glutamic acid is important for the *umami* flavor. A way to exhibit flavor enhancement characteristics by using mild acid hydrolysis of the protein, followed by enzymatic hydrolysis, has been described (63). The preferred protein for producing savory flavors is wheat gluten. The mildly acidic hydrolysis is done to obtain a deamidation in the first step. This step was carried out for 1 h at about 95 °C at a hydrogen ion concentration of about 1.0 M. The enzymatic hydrolysis was preferably carried out by use of an *Aspergillus oryzae* protease (Prozyme 6 from Amano) having both endoproteases and exoproteases. The overall degree of hydrolysis was 50–70%. The conditions during the acid treatment were mild enough to avoid the formation of substantial amounts of monochloropropanols.

Flavourzyme™ used in combination with Alcalase may obtain a degree of hydrolysis of 60–70% for hydrolysis of soy protein isolate (62). Except for the deamidation step, the same degree of hydrolysis could thus be obtained without the acid treatment.

An enzymatically hydrolyzed soy protein, so-called eHVP, has been developed by Miwon C., Ltd. (64). This product is suitable for special uses, such as infant diets, sports nutrition, and medical diets. eHVP gave no limitations of dosage in formulation as a flavor enhancer.

Flavors from Hydrolyzed Animal Proteins

Products that have a high flavor of meat extract are used in soups, sauces, and prepared meals. Proteinaceous material is recovered with proteases from coarse and fine scrap-bone residues from mechanical fleshing of beef, pig, turkey, or chicken bones. The flavor intensity depends on the content of free amino acids and peptides and their reaction products. Reactions that develop flavor are, for example, Maillard reactions between reducing sugars and amino acids, thermal degradation due to the Maillard second stage, and deamination, decarboxylation, or degradation of cysteine and glutathione. The latter reaction can give rise to a large number of volatile compounds of importance to aroma and taste (58).

For the production of protein hydrolysates from meat, the first step is an efficient solubilization of the product. For this purpose the bacterial serine proteases and metalloproteases are very efficient. Unless certain treatments are undertaken, meat hydrolysates are usually found to be bitter when the degree of hydrolysis is above 10%, which is needed for sufficient solubilization (65). A method comprising a series of steps using raw meat hydrolyzed with a specific combination of neutral and alkaline proteases has been developed (66). This meat hydrolysate exhibits excellent organoleptic properties and can be used as a meat-flavored additive in soup concentrates. The degree of hydrolysis above 20% did not show bitterness when this specific combination of enzymes was used. This effect may be due to the preferential proteolytic specificity being favorable when metalloprotease and serine protease are used simultaneously.

A method for further enhancement of the flavor of a natural beef-juice concentrate is hydrolysis using endoproteases (metalloprotease, serine protease, or combinations) followed by treatment with an exoprotease such as Flavourzyme™ or by incubation with a culture of a food-grade microorganism that is capable of producing exoproteases (e.g., aminopeptidase) (67).

PROTEASES FOR LEATHER MANUFACTURING

The processing of skin and hides for leather has been based on enzymes since 1908, when Otto Röhm patented the first standardized bate based on pancreatic enzymes (68). Leather chemistry research helped to improve understanding of the bating process and at the same time spurred on developments to improve leather processing (69).

The stages in the processing of hides to leather include curing, soaking, liming or dehairing, deliming, bating, pickling, and tanning. The benefits of using enzymes in the different stages of leather manufacturing are mainly reduction of process time, increased opening of the fibrous structure, improved clarity of the surface, increased softness, increased uptake of water, improvement of area yield, and reduction of the need for chemicals.

The animal hide is composed of three main layers. The grain is the outer layer in which the hairs are embedded. This layer also contains the small glands that produce oil and sweat. The corium represents the principal part of the hide. To convert raw hide into leather a great deal of interfibrillar material must be removed from this layer. Collagen fibers, elastin, and proteoglycans, such as dermatan sulfate, can be regarded as a kind of cement, that binds collagen fibrils. The connective tissue of the underside of the hide contains the blood vessels, fat tissue, and some meat proteins.

Raw or cured hides must be properly soaked to obtain a satisfactory hydration and removal of unwanted materials, including meat and fat. Bacterial or pancreatic alkaline proteases are normally preferred for hydrolysis of interfibrillar proteins.

Enzymes for Dehairing, Liming, and Bating

During liming, the highly charged dermatan sulfate may be removed from nearly each collagen fibril by use of alkaline-stable proteolytic enzyme preparations such as NUE 0.6 MPX from Novo Nordisk A/S (69,70). An accelerated dehairing process and an up to 40% reduction of the requirement for sulfide have been measured (69). Further-

more, increased strength of grain leather and increased area yield are useful results.

Before the bating process, the hides are delimed with ammonium sulfate or ammonium chloride. When proteases are applied a thickness similar to that of hairless pelt is reestablished. The early preparations proposed by Röhm were pancreatic trypsins. Some of the microbial proteases shown in Table 1 are used today as bating enzymes to make the hides soft and supple before tanning.

OTHER BIOPROCESSES BASED ON ENZYMES AND PROTEIN HYDROLYSIS

A number of applications of proteases have been mentioned. The emphasis of this article has been on my research over the years and on the most significant industrial applications, as seen from an enzyme producer's point of view. Some applications of proteases have not been mentioned in detail. This is the case, for example, with regard to the use of enzymes in fish processing, where proteases are used for production of fish sauce, for salted matjes herring, for deskinning of fish, for production of protein hydrolysate as a cryoprotective agent, and for roe caviar purification (71).

Using a glutamic acid–specific protease from *Bacillus licheniformis*, the viscosity of a solution of whey protein concentrate could increase dramatically during hydrolysis due to production of plastein aggregates (7). An increased water-binding capacity and increased viscosity during hydrolysis of mechanically deboned meat was also reported (7). Extending hydrolysis to smaller peptides reduced viscosity again. A practical application of these findings remains to be further developed.

Use of proteases as processing aids for reduction of viscosity of stickwater before and during evaporation in a fishmeal plant or a meat-rendering plant could reduce the fouling on heat surfaces. The capacity of the evaporators was then increased due to the reduction of the overall heat transfer coefficients and because a greater amount of dry solids could be obtained. A direct economic advantage was shown (72,73).

Evaporation of raw fruit water from potato production represented an insurmountable problem, because gelling due to the protein components occurs at about 15% dry matter. It was obviously necessary to add proteolytic enzymes, as mentioned earlier, for concentration of stickwater. However, it was surprisingly found that no effects of the enzymes could be observed until heat treatment, preferably jet cooking to at least 125 °C using a holding time of about 3 min, had inactivated a powerful protease inhibitor in the raw fruit water (74).

CONCLUSION

Proteases are by far the most researched enzymes for industrial bioprocessing. The major uses of these enzymes are divided among many applications, all based on hydrolysis as the catalytic concept. Similarities in the catalytic mechanism within the widespread uses of proteases for detergents, for modification of food proteins, for better utilization of waste fractions of the food and feed industry, and for production of basic food have been shown. No doubt more development within proteolysis, as broadly defined, will be seen in the future. To meet challenges within the detergent industry and the food–feed industry, discovery of new protease molecules must be made through protein engineering and recombinant DNA techniques.

BIBLIOGRAPHY

1. American Conference of Governmental Industrial Hygienists, *Documentation of the Threshold Limit Values*, 5th ed., ACGIH, Cincinnati, Ohio, 1986, pp. 540–541.
2. G. Peters and D.P. Mackenzie, in J.H. van Ee, O. Misset, and E.J. Baas eds., *Enzymes in Detergency*, Marcel Dekker, New York, 1997, pp. 327–340.
3. National Research Council (NRC), *Enzyme-Containing Laundering Compounds and Consumer Health*, Report no. PB 204 118 of the ad hoc Committee on Enzyme Detergents, Division of Medical Sciences, National Academy of Science—National Research Council, Washington, D.C., 1971, pp. 1–31.
4. J. Adler-Nissen, *Enzymic Hydrolysis of Food Proteins*, Applied Science Publishers, London, 1986.
5. Novo Nordisk, *Enzymatic Modification of Proteins Using Novo Proteinases*, Bagsvaerd, Denmark, 1989.
6. J. Adler-Nissen, *J. Agric. Food Chem.* **27**, 1256–1262 (1979).
7. P.M. Nielsen, in S. Damodaran and A. Paraf eds., *Food Proteins and Their Applications*, Marcel Dekker, New York, 1997, pp. 443–472.
8. C.F. Jacobsen, J. Léonis, K. Linderstrøm-Lang, and M. Ottesen, *Methods Biochem. Anal.* **4**, 171–210 (1957).
9. S.P.L. Sörensen, *Biochem. Z.* **7**, 45–101 (1908).
10. J. Adler-Nissen, *J. Chem. Technol. Biotechnol.* **32**, 138–156 (1982).
11. J. Adler-Nissen, *J. Chem. Technol. Biotechnol.* **34B**, 215–222 (1984).
12. K. Linderstrøm-Lang, *Proteins and Enzymes, III: The Initial Stages in the Breakdown of Proteins by Enzymes*, vol. 6, Stanford Univ. Press, Stanford, Calif., 1952, pp. 53–72.
13. H.S. Olsen. *ZFL: International Zeitschrift für Lebensmittel-Technologie und Verfahrenstechnik*, **34**, 406–412 (1983).
14. Y. Birk, *Methods Enzymol.* **45**, 695–739 (1976).
15. J.K. Chavan and J. Hejgaard, *J. Sci. Food Agric.* **32**, 857–862 (1981).
16. A. Shinzaki, Y. Aizono, H. Yamagata, T. Iwasaki, and M. Sawano, *Agric. Biol. Chem.* **55**, 1521–1530 (1991).
17. I.E. Liener, in J. Adler-Nissen, B.O. Eggum, L. Munck, and H. Sejr Olsen eds., *Proceedings of the 11th FEBS Meeting, vol. 44. Biochemical Aspects of New Protein Food*, Pergamon Press, Oxford, U.K. 1978, pp. 129–138.
18. T. Krawczyk, *INFORM* **8**, 7–14 (1997).
19. U.S. Pat. 4,497,897 (February 5, 1985), J.H. Eilertsen, A.D. Fog, and K. Gibson (to Novo Nordisk).
20. M.S. Showell, in M.S. Showell and E.J. Baas eds., *Powdered Detergents*, Marcel Dekker, New York, 1998, pp. 1–19.
21. R. Fleuren, P. Bauditz, S. Schløsler and H. Bodenhoff, *88th AOCS Annual Meeting*, Seattle, Wash., May 14–18, 1997.
22. N. Eriksen, in T. Godfrey and S. West eds., *Industrial Enzymology, 2nd ed.*, Stockton Press, New York, 1996, pp. 189–200.
23. G. Jensen, *SÖFW J.* **123**, 723–731 (1997).
24. Pat. WO 97/02753 (January 30, 1997), H.S. Olsen (to Novo Nordisk A/S).

25. D. Aaslyng, E. Gormsen, and H. Malmos. *J. Chem. Tech. Biotechnol.* **50**, 321–330 (1991).
26. H.S. Olsen and P. Falholt, *J. Surfact. Deterg.* **1**, 555–567 (1998).
27. H.-P. Welzel, R. Neumann, and K. Haage. *Tenside Surf. Det.* **31**, 286–293 (1994).
28. K. Haage, B. Weiland, C. Wedler, and H.-P. Welzel. *Tenside Surf. Det.* **32**, 45–54 (1995).
29. Pat. DE 44 10 000 C1 (March 2, 1995), E.V. Kries, E. Eberhard, and A. Sander (to Henkel KGaA).
30. Anonymous, *SÖFW J.*, **121**, 383 (1995).
31. W. Wu and N.S. Hettiarachchy, *J. Surfact. Deterg.* **1**, 241–246 (1998).
32. W. Wu, N.S. Hettiarachchy, and K.C. Rhee. *J. Surfact. Deterg.* **1**, No. 3. 387–392 (1998).
33. K. Schäfer and H. Höcker, *Proceedings of the 38th International Detergency Conference*, wfk-Forschungsinstitut für Reinigungstechnologie, e.V., D-47798, Krefeld, Germany, 1998, pp. 116–121.
34. H. Sejr Olsen, in G. Reed and T.W. Nagodawithana eds., *Biotechnology*, vol. 9: *Enzymes, Biomass, Food, and Feed*, VCH, 2nd ed., Weinheim, Germany, 1995, pp. 666–736.
35. P.F. Fox, *Neth. Milk Dairy J.* **35**, 233–253 (1981).
36. K. Burgess and M. Shaw, in T. Godfrey and J. Reichelt eds., *Industrial Enzymology*, Macmillan, Surrey, England, 1983, pp. 260–283.
37. I.A. Fedrick and S.C. Fuller, *Aust. J. Dairy Tech.* **43**, 12–15 (1988).
38. P.F. Fox, *J. Food Biochem.* **17**, 173–199 (1993).
39. G. Puski, *Cereal Chem.* **52**, 655–666 (1975).
40. J. Adler-Nissen and H.S. Olsen, *Am. Chem. Soc. Symp. Ser.* **92**, 125–146 (1979).
41. J. Adler-Nissen, S. Eriksen, and H.S. Olsen, *Qual. Plant. Plant Foods Hum. Nutr.* **32**, 411–423 (1983).
42. H.S. Olsen, *International Zeitschrift für Lebensmittel—Technologie und Verfahrenstechnik* **11**, 57–64 (1978).
43. H.S. Olsen and J. Adler-Nissen, *ACS Symp. Ser.* **154**, 133–169 (1981).
44. L.K. Christensen, *Comptes Rendus des Travaux du Laboratoire Carlsberg—Série Chimique*, **28**, 39–169 (1952).
45. U.S. Pat. 4,431,629 (February 14, 1984), H.S. Olsen (to Novo Industri A/S).
46. H.S. Olsen, *Food Technology International Europe 1988*, Sterling Publications Limited, London, 1988 pp. 245–250.
47. L. Anson, *J. Gen. Physiol.* **22**, 79–89 (1939).
48. H.S. Olsen and J. Adler-Nissen, *Process Biochem.* **14**, 6–11 (1979).
49. S. Kristensen, A/S De Danske Sukkerfabrikker File 1972-GB-1182, 1985.
50. H.S. Olsen, *Meat Symposium*, Zalaergerszeg, Hungary, April 3, 1991 (Novo Nordisk file no. A-06143, 1991.)
51. J. Adler-Nissen, in G. Charalambous ed., *Frontiers of Flavor*, Elsevier, Amsterdam, 1988, pp. 63–77.
52. K.H. Ney, *Z. Lebensm Untersuch. Forsch.* **147**, 64–71 (1971).
53. J. Adler-Nissen and H.S. Olsen, in G. Charalambous and G. Inglett eds., *Chemistry of Foods and Beverages:—Recent Developments*, Academic Press, New York, 1982, pp. 149–169.
54. U.S. Pat. 4,324,805 (April 13, 1982), H.S. Olsen (to Novo Industri A/S).
55. F. Sosulski, *J. Am. Oil Chem. Soc.* **56**, 711–715 (1979).
56. H.F. Bernholdt, in G. Reed ed., *Enzymes in Food Processing*, 2nd ed., Academic Press, New York, 1975, pp. 473–492.
57. European Pat. EP 0471 470 A2 (February 19, 1992), S.J. Warren (to British Beef Company Ltd.).
58. G.S.D. Weir, in B.J.F. Hudson ed., *Biochemistry of Food Proteins*, Elsevier, London, 1992, pp. 363–408.
59. BIOTOL, *Biotechnological Innovations in Food Processing*, Oxford, U.K., Butterworth-Heinemann, 1991, pp. 211–252.
60. J.S. Hamada, in B.J.F. Hudson ed., *Biochemistry of Food Proteins*, Elsevier, New York, 1992, pp. 249–270.
61. J.S. Hamada, *Crit. Rev. Food Sci. Nut.* **34**, 283–292 (1994).
62. P.M. Nielsen, *Food Ingredients Europe '94*, Earls Court, London, October 4–6, 1994.
63. European Pat. EP 0495390 AI (January 9, 1992), D.J. Hamm (to CPC International).
64. H.J. Chae, M. In, and M.H. Kim, *Agric. Chem. Biotechnol.* **40**, 404–408 (1997).
65. G.M. O'Meara and P.A. Monro, *Meat Sci.* **11**, 227–238 (1984).
66. Pat. WO94/01003 (January 20, 1994), H.H. Pedersen,, H.S. Olsen, and P.M. Nielsen (to Novo Nordisk A/S).
67. European Pat. EP 0 505 733 A1 (February 21, 1992), S.S-Y. Kwon, M.A. Marsico, and D.V. Vadehra (to Societe des Produits Nestle S.A.).
68. U.S. Pat. 886,411 (May 5, 1908), O. Röhm (to Otto Röhm of Esslingen, Germany).
69. K.T.W. Alexander, *JALCA*, **83**, 287–316 (1988).
70. Application of NUE 0.6 MPX for liming and unhairing of hides, Application sheet, Novo Nordisk A/S, 1991.
71. G. Stefansson, in T. Nagodawithana and G. Reed eds., *Enzymes in Food Processing*, 3rd ed., Academic Press, New York, 1993, pp. 459–470.
72. F. Jacobsen, *Process Biochem.* **20**, 103–108 (1985).
73. F.M. Christensen, *Biotechnol. Appl. Biochem.* **11**, 249 (1989).
74. U.S. Pat. 5,573,795 (November 12, 1996), H.S. Olsen (to Novo Nordisk A/S).

See also ENZYMES, DETERGENTS; ENZYMES, FRUIT JUICE PROCESSING; ENZYMES, PULP AND PAPER PROCESSING; FOOD PROCESS ENGINEERING.

ENZYMES, PULP AND PAPER PROCESSING

J.S. TOLAN
Iogen Corporation
Ottawa, Canada

KEY WORDS

Bleaching
Deinking
Dewatering
Enzymes
Ligninase
Lipase
Papermaking
Pulp
Refining
Xylanase

OUTLINE

INTRODUCTION

The use of enzymes in the pulp and paper industry started in the late 1980s and has grown rapidly since then. This growth has been fueled by several factors, including (*1*) an improved understanding of the interactions between enzymes and the constituents of pulp and paper processing, (*2*) an increased need for the industry to adopt environmentally benign technology, and (*3*) development of cost-effective production technology for the relevant enzymes. The pulp and paper industry processes wood and other cellulose-containing fiber, such as straw. Typically, pulp is made by processing the fiber feedstock, and paper is made by combining one or more types of pulp with various non-pulp additives designed to achieve the desired properties, such as smoothness, opacity, or color.

Enzymes are used to aid pulp and paper manufacturing by modifying the following types of substrates:

1. The cellulose fiber
2. Noncellulosic constituents of the fiber, such as hemicellulose and lipids
3. Paper additives, such as starch
4. Contaminants, such as slime

This article provides general background on pulp and paper processing and describes the use of enzymes in each of the processing areas.

PULPING PROCESSES

Generalized flow sheets for three major pulp production processes are shown in Figure 1. Although these processes have common elements, the differences are sufficiently large that enzymes are used for different purposes.

Mechanical Pulps

Newsprint and other bulk papers are made primarily by mechanical means, that is, the raw fiber is processed primarily by mechanical action. First, the fiber feedstock is cleaned and chipped to uniform size. The raw fiber is then torn apart by mechanical means in the pulping. Closely associated with the pulping is refining, which improves the strength of the pulp. Enzymes are used to improve the dewatering and refining of mechanical pulps. After refining, the pulp is mildly bleached, usually with peroxide. Enzymes are not used to aid in the bleaching of mechanical pulps, although enzymes are used in chemical pulp bleaching. After bleaching, the pulp is ready for papermaking. Mechanical pulp is produced at lower cost than chemical pulp, but the strength and surface properties are not suitable for more demanding applications, such as writing paper or wrapping paper.

Chemical Pulps

These higher performance papers are made from chemical pulps, most notably by cooking in alkali (kraft pulp) or in acid (sulfite pulp). For chemical pulps, the fiber source and initial handling are similar as for mechanical pulps. However, the chemical pulping processes remove impurities, principally lignin, from the fiber, which improves the strength and makes the surface more uniform. After pulping, bleaching is carried out at harsher conditions and with more aggressive oxidizing chemicals than used for bleaching mechanical pulps; the resulting product is very white. Enzymes are widely used to decrease the usage of these bleaching chemicals, particularly chlorine gas and other chlorine-based oxidizing chemicals. After bleaching, the pulp is refined (the reverse of the order for mechanical pulp). Data suggest that enzymes can enhance the refining of chemical pulps, but this is not as widely reported as with mechanical pulps.

Secondary Fiber

A third source of pulp that is becoming more prominent involves the use of recycled paper as the feedstock. This is known as secondary fiber processing. After mechanical shredding and repulping, the ink is removed from the pulp, a process known as deinking. There are laboratory data and limited mill results that show benefits of enzyme use in deinking. The deinked stock is then lightly refined and blended with pure mechanical pulps for paper production. Enzymes are also reported to be beneficial in the refining of secondary fiber.

Papermaking

Regardless of the pulping process, the finished pulp is combined with inorganic fillers, paper additives to improve properties, and process additives to improve the performance of the paper machine. The pulp and additives are formed into sheets and dried on a paper machine. Enzymes are used in papermaking to modify the starches used in paper sizing, to remove pitch, and to prevent slime accumulation. Although an important benefit of enzyme use in the industry is a decreased amount and toxicity of effluent

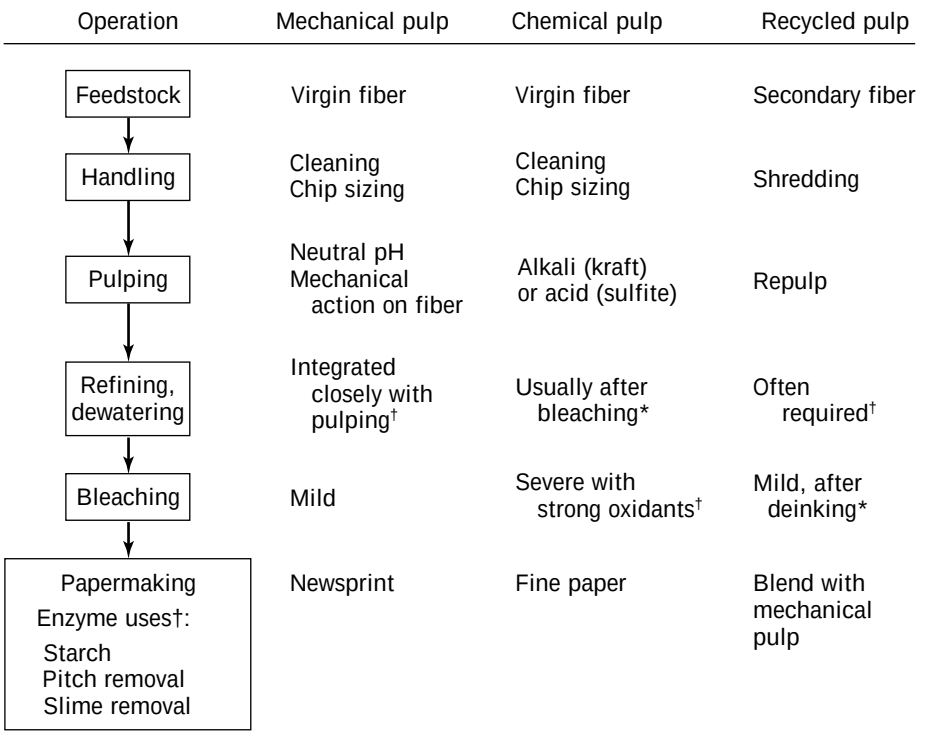


Figure 1. Pulping processes. *Widely reported enzyme effect. †Commercial enzyme process.

discharges, there is little use of enzymes for direct treatment of waste streams. The environmental benefits of enzyme use in the industry are mentioned in the following relevant process sections, and each processing area and the associated use of enzymes is described in more detail.

ENZYMES IN PULP PRODUCTION

Fiber Handling

Historically, wood has been by far the most prominent cellulose-containing feedstock for the industry. Numerous species of hardwood and softwood are used around the world. Nonwoody fibers such as bamboo, kenaf, and straw are also widely used, particularly in China and tropical areas. Regardless of the fiber source, a certain amount of physical handling is carried out to obtain cleanliness and particle size suitable for pulping. This processing takes the form of debarking, washing, and screening, to remove the foreign matter, and mechanical chipping for most pulping processes. Nonwoody fibers require removal of seeds and other nonfiber constituents. In virtually all cases, the intact fiber structure is not very susceptible to enzyme attack. Consequently, there have been few laboratory scale, let alone commercial, enzyme processes involving treating intact fiber.

There has been much research focused on the use of whole cells to digest wood chips. This is targeted particularly to remove lignin to improve the subsequent pulping and bleaching (1) and to remove pitch by digesting lipids, of which at least one commercial product is available (2). Whole cell applications on intact fiber are plagued by the long (several day) reaction times required, the extreme variability in day-to-day chip storage times, and the vulnerability of the microbes to subfreezing temperatures. Whole cell applications will not be discussed further in this article.

Pulping

Pulping is the process by which the macroscopic structure of raw wood fiber is broken apart, rendering the fiber much more pliable and thereby suitable for papermaking. Two general classes of pulping operations are chemical pulping and mechanical pulping. These operations will be discussed in terms of wood processing, although the same operations apply to the processing of nonwoody fibers.

Chemical Pulping. In chemical pulping, the fiber is heated in the presence of chemicals that dissolve a portion of the fiber, particularly the lignin. Today, most of the chemical pulp is from the alkali or kraft process, in which wood chips are cooked at 160 to 190 °C (320–374 °F) for about 3 h in a concentrated liquor of sodium hydroxide and sodium sulfide. After the alkali cook, the pulp is a slurry with the liquor; the cooking liquor is reused through an extensive recovery process. The kraft cooking treatment greatly increases the cellulose content of the fibers (Table 1).

The other significant chemical pulping process is acid (sulfite) pulping. Sulfite pulping is particularly prevalent in central Europe. In sulfite pulping, the wood chips are cooked in sulfurous acid analogous to the alkali liquor of kraft pulp. The cellulose content is higher than that of kraft pulp, but the resulting pulp is not as strong. Sulfite pulp is used primarily as a cellulose chemical feedstock. The conditions of chemical pulping are too harsh to be tol-

Table 1. Typical Chemical Pulp Composition

Constituent	Feedstock	Pulp
Cellulose	60	85
Hemicellulose	15	10
Lignin	20	4
Miscellaneous: resin, ash, etc.	5	1

erated by commercial enzymes, so there has not been significant attention paid to enzymes in this area. The most significant use of enzymes in chemical pulp has been in conjunction with bleaching rather than pulping.

Mechanical Pulping. In mechanical pulping, the fiber is refined by rotating plates; today, this is usually carried out at temperatures of about 140 °C (284 °F) and designated as thermomechanical pulping. The power input onto the fiber and the spacing between the plates that the pulp is forced through determine the severity of the pulping. Refining is a series of compressions and decompressions that gently separate and fibrillate individual fibers. This enhances pulp strength relative to the older stone groundwood process, which involved coarse grinding on a large stone. As with chemical pulp, the conditions of mechanical pulping are too harsh for enzyme treatment. However, enzyme treatment of mechanical pulp has important implications in the pulping and refining, as described in the next section.

Pulp Dewatering and Refining

One process parameter that is of primary interest in mechanical pulping and in secondary fiber processing is the rate of drainage of water from the pulp. This is important because the subsequent rate of water removal on the paper machine is often rate limiting in mill production throughput. The presence of fine particles in the pulp markedly decreases the rate of drainage. Typically, pulps consist of about 12% to 15% fines that pass through a 200-mesh screen; particle size is usually based on a Bauer-McNett classifier. Secondary fiber tends to have particularly high levels of fines, and the level of fines increases with repeated recycling of stock. Refining, which is required in many cases to increase the strength of the pulp, also creates fines that decrease the rate of drainage.

Mixtures of cellulase enzymes and hemicellulase enzymes are used to increase the drainage rate of mechanical pulp and secondary fiber. The speed of the paper machine can then be increased. Alternatively, additional refining or pulping can be carried out to obtain pulp of increased strength with the original rate of drainage. This is shown in Table 2.

The fines consist largely of cellulose and hemicellulose and, relative to the remainder of the pulp, have a high surface area for enzyme attack. There is some evidence for three distinct actions by the enzymes: (*1*) a portion of the fines are solubilized, (*2*) there is a flocculation of the particles, and (*3*) there is a cleaning of the small fibrils on the fiber surface. All these would tend to increase the drainage of the pulp.

Table 2. Pulp Dewatering by Enzyme (3)

Parameter	Control[a]	Enzyme-Treated[b]	Refine + Enzyme[b]
Freeness (mL CSF)	192	274	156
Burst (kPa m^2/g)	1.83	1.84	2.07
Corrugation (newtons force)	129	124	141

[a]Corrugated container board refined on Sprout Waldron disc refiner.
[b]Enzyme treatment 3% consistency, 50 °C (122 °F,) pH 4.8, 30 min with 0.1% enzyme.

A commercial product of Nalco (Naperville, Ill.) that is used to aid pulp dewatering is a mixture of cellulase and hemicellulases from the fungus *Trichoderma*. The enzyme is supplied in liquid solution and is pumped into the storage chest before the pulp machine headbox or, in the case of secondary fiber, just after repulping (Figure 2). The pulp is diluted in water (2 to 4% solids consistency) in a stock chest, and the enzyme is simply added to this slurry; the flow and mixing are sufficient to disperse the enzyme into the pulp. The process pH of 5 to 7 is in the optimal range for the enzyme; the process temperature of 35 to 50 °C (95 to 122 °F) is adequate. The retention time is at least 30 mins, and the enzyme dosage is 1.5 L (~1/2 U.S. gallon) per ton of pulp. The drying or refining of the sheet destroys the enzyme, preventing further action on the pulp. Careful control of the enzyme dosage is important, excessive enzyme treatment can weaken the pulp because of cellulose hydrolysis. In addition, excessive fines creation or destruction can affect the opacity of the sheet.

Bleaching

The process by which pulp is brightened or made completely white with oxidizing chemicals is bleaching. En-

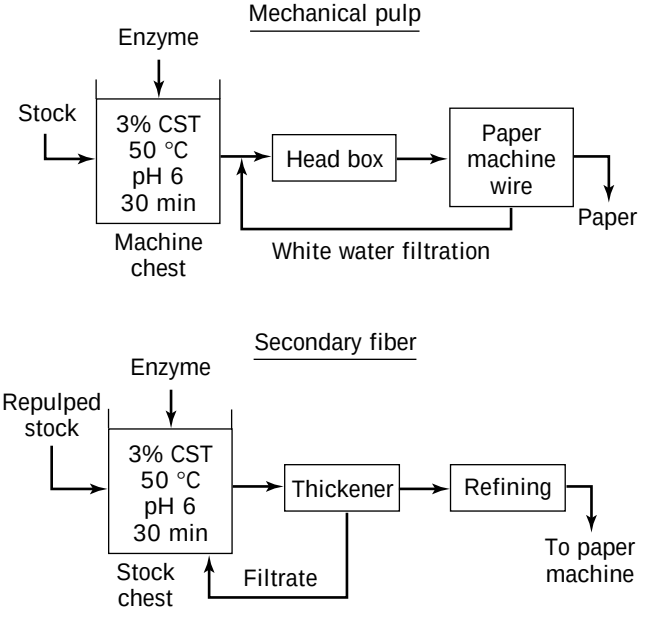


Figure 2. Enzymes for pulp dewatering.

zymes are used to aid in the bleaching of kraft pulp. In most cases, mechanical pulp is only mildly bleached with hydrogen peroxide and sodium hydrosulfite. There are few effluent problems with these bleaching chemicals. The major usage of bleaching chemicals and associated effluent problems are with kraft pulp.

In bleaching kraft pulp to produce bright white finished pulp, the residual lignin is virtually removed. Bleaching is carried out as a continuous process. The pulp is reacted with oxidizing compounds that preferentially attack the aromatic rings on the lignin while damaging the cellulose as little as possible. Bleaching is carried out in 3 to 6 discrete stages, which taken together are known as a bleaching sequence. A classical three-stage sequence uses chlorine gas in the first stage, followed by an alkali extraction, then finally reaction with chlorine dioxide, a chemical that is particularly suited to brightening kraft pulp.

In recent years, the bleached pulp industry has been motivated to decrease the use of chlorine-based bleaching compounds for environmental, regulatory, and market reasons. There are several options for mills to do this, but most require the expenditure of large sums of capital. Enzymes can be used to eliminate some of the chlorine-based compounds with little capital investment and are thus attractive to the industry.

Xylanase. Xylanase enzymes, which act on the xylan in the pulp, have been used commercially by the industry and are the example of the growing use of enzymes to aid in the bleaching of kraft pulp. At this writing (late 1997), the percentage of the bleached kraft pulp that is treated with xylanase enzymes, by country, is: Canada, 12%; Finland, 6%; Sweden, 3%; United States, 2%. These enzymes are marketed to mills in North America by Iogen Corporation of Ottawa, Canada; ICI of Toronto, Canada; and Clariant of Charlotte, N.C.

The beneficial effect of xylanase enzymes in bleaching was reported by Viikari et al. (5). The commercial enzymes used for enhancing the bleaching of kraft pulp all contain β-1,4 endo-xylanase activity that performs most of the action on the pulp. These enzymes may also contain other hemicellulase enzymes. Xylanase enzymes act on the xylan in the pulp; this is a portion of the hemicellulose with a β-1,4-linked xylose backbone, with arabinosyl and other side chains. The enzymes do not bleach the pulp, but rather change the pulp structure such that the oxidizing bleaching chemicals (whether chlorine-based or otherwise) work more effectively on the pulp. Xylanase enzymes can decrease the use of oxidizing chemicals by 15 to 20%; more typically, a decrease of 10 to 15% is observed. The enzyme treatment can also increase the brightness of the bleached pulp and decrease the quantity of off-white fiber particles, known variously as shives, sclereids, and vessels.

Figure 3 shows the process streams at the point of xylanase enzyme treatment in a kraft mill. The unbleached pulp, known as brownstock, is pumped from the final washer into a storage tank before it enters the bleach plant. The pulp is a mat about 2 inches thick, of 8 to 14% solids consistency. Although the pulp is well washed, it has sufficient residual alkali soda that its pH is 9 to 12. This is higher than the effective pH range of the commercial enzymes, so the pulp must be acidified, usually with sulfuric acid, to the appropriate pH. The temperature of the pulp is typically 50 to 70 °C (122 to 158 °F). The capacity in the storage tank usually is sufficient for at least 1 to 2 hours of stock retention, which is sufficient time for the enzyme to act on the pulp. The pulp is pumped from the storage tank into the bleach plant, where contact with the bleaching chemicals destroys the enzyme. The enzyme-treated pulp then proceeds as normal through the bleach plant, except with a decreased requirement for bleaching chemicals. The decrease in demand for bleaching chemicals requires changes to be made in the bleach plant control strategy. The details of the how this is done are covered elsewhere (6).

The enzymes are supplied in liquid solution, often in 1,000-L (275 gallon) totes. At the typical enzyme usage of 0.2 to 0.6 L per ton of dry pulp, a mill that processes 1,000 tons of pulp per day requires 2 to 5 totes per week of enzyme. The temperature on the operating floor can reach 35 °C (95 °F), so the enzyme must withstand this temperature during storage or be maintained in a refrigerated tank. It is imperative that the enzyme, which is added to pulp in such small quantities, be evenly dispersed into the pulp mat to obtain uniform treatment and bleaching response. The enzyme solution is usually added to the pulp with some dilution water to increase the bulk liquid dispersion. The combined stream is added before a stock pump to take advantage of the mixing provided by these pumps. The equipment that is used to deliver this stream to the pulp varies from mill to mill but has included simple hoses, shower bars, and other devices.

The pH of the pulp must be decreased for effective enzyme treatment. This is accomplished by adding acid to the pulp, usually well in advance of the point of enzyme addition. The most common acid used is sulfuric, but sulfurous acid, carbon dioxide, and waste acid from various mill process streams have been used. When sulfuric acid is used, it is stored as a 93 to 98% concentrate. The typical acid requirement is 3 to 5 kg/ton (6.6 to 11 lb/ton) pulp, which is an acid flow of 1 L of concentrate per minute on a 1,000 ton per day bleaching line. As with the enzyme, the acid must be dispersed uniformly throughout the pulp. The acid stream is pumped through a hose or shower bar and onto the pulp. The choice of location and equipment to handle the acid and add it to the pulp takes into account the mill's equipment configuration and desires for safe acid handling and avoidance of equipment corrosion.

Several properties of xylanase enzymes are important in bleaching applications. These include:

1. *The bleaching benefit.* The precise mechanism of the enzyme's action on pulp is not known. Hence, no single enzyme activity has been identified that characterizes the enhanced bleaching of pulp. The bleaching benefit for a given enzyme is measured by treating pulp and bleaching. That having been said, xylanase activity is present in all the commercial enzymes and is important in enhancing pulp bleaching. However, it is by no means the only active enzyme present; enzymes of the same xylanase activity can have very different effects in bleaching (7).

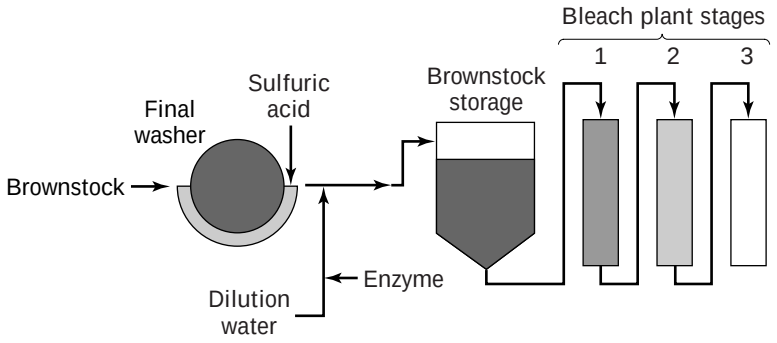


Figure 3. Enzyme treatment in pulp bleaching.

2. *The conditions of use.* All the commercial xylanases are characterized by a range of pH and temperature that must be maintained for full effectiveness. The pH range is typically 1 unit broad, and the temperature range spans about 10 °C (18 °F). Figure 4 shows the implications of various pH and temperature ranges on mill operation; different enzymes are preferred for different mills. Enzymes that can operate at pH 10 to 12 and at temperatures up to 70 °C (158 °F) are the subject of much research.
3. *The absence of deleterious effects on the pulp.* This includes degradation of the pulp by cellulase; the enzymes must have little or no cellulase activity.

Enzyme manufacturers have used several different technologies to produce xylanase enzymes that comply with these requirements. These technologies are as follows:

1. *Produce xylanase using a natural microbe that does not make cellulase.* This seemingly straightforward approach has not generally been successful, because most such microbes make only a limited amount of protein. Examples of such microbes are species of *Bacillus* and *Streptomyces* bacteria and *Aspergillus* fungi.
2. *Clone and express xylanase in a microbe that does not make cellulase.* This approach has been somewhat more successful than the first, with host organisms including *Aspergillus oryzae, Bacillus pumilis*, and *Streptomyces lividans*.
3. *Produce xylanase along with cellulase and remove the cellulase using downstream processing.* This approach has worked well using *Trichoderma* fungi, which are well known as high protein-producing microbes. Methods of cellulase removal include size-differentiated filtration (8), polyethylene glycol precipitation (9), precipitation with metal ions (10), ion exchange chromatography (11), and salting out (12).
4. *Delete cellulase genes from* Trichoderma. This approach has also worked well for commercial products (13).

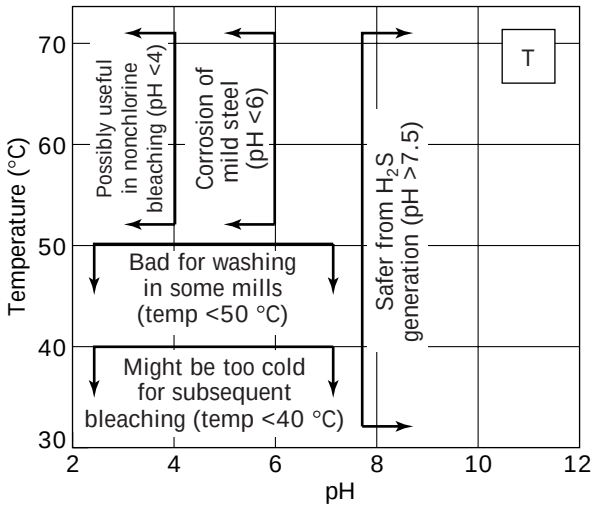


Figure 4. pH and temperature for enzymes in pulp bleaching: Effects on mill operations. The box labeled T indicates typical brownstock conditions.

In addition to xylanase, many other hemicellulase enzymes have been evaluated for their ability to enhance the bleaching of pulp. Of these, only mannanase appears to deliver benefits comparable to xylanase (14). Mannanase enzymes are not used commercially for bleaching at present.

Ligninase. Research over 3 decades focused on the use of ligninase enzymes to oxidize and remove the lignin from the pulp. The motivation for this research was straightforward: lignin removal is the objective of bleaching. However, for many years ligninase technology was not suitable for mill applications because reaction times of several days were required. A breakthrough discovery (15) decreased the time required to bleach pulp with ligninase enzymes to 1 hour, which has motivated much recent effort in the area.

Four proteins have been identified that participate in enzymatic delignification: lignin peroxidase, cellobiose dehydrogenase, manganese peroxidase, and laccase. Lignin peroxidase is a manganese-containing enzyme that requires hydrogen peroxide to oxidize the aromatic rings and side chains of lignin. Cellobiose dehydrogenase is a multifunctional enzyme that has recently been shown to be able to participate in enzymatic delignification (16).

From the point of view of commercial-scale use, laccase is the furthest along in development of the ligninase en-

zymes. Laccase is a copper-containing enzyme that catalyzes the oxidation of phenolic groups on lignin. The most prominent microbe for laccase production is the white-rot fungus *Trametes versicolor*, and production of the enzyme is induced by phenolic compounds. The copper that forms the redox-active center of laccase is part of the functional enzyme as it is produced and does not need to be added to the system for lignin degradation to occur.

Call (15) determined that inclusion of a mediator compound dramatically increases the rate of lignin oxidation. The mediator reported by Call is 1-hydroxybenzotriazole; many other effective mediators have since been reported (17–19). Laccase and mediator are proposed to be used in a mill as follows. Pulp at an intermediate brownstock washer is adjusted to pH 4.5 and 45 °C (113 °F), then exposed to oxygen at moderate pressure. The enzyme and mediator are then added and act on the pulp for 1 hour. After treatment, the pulp is washed to remove the soluble lignin. The laccase-mediator process is reported to remove 50% of the lignin in a single-stage reaction and a further 25% in a second reaction. This magnitude of effect is much greater than that possible with xylanase and would be an extremely important step forward for mills to eliminate chlorine-based oxidizing chemicals entirely. The laccase-mediator bleaching has been demonstrated in pilot plant trials, but it has not been used commercially because of high costs of production of the enzyme and the mediator.

The fourth ligninase enzyme, manganese peroxidase, offers the potential advantage over laccase in that it does not require a mediator. Manganese peroxidase is an iron-containing enzyme that uses hydrogen peroxide to oxidize chelated manganese, which then acts on lignin. The enzyme is most effective at 45 °C (113 °F), pH 4.5. The rate and extent of lignin removal of manganese peroxidase is not as great as with a laccase-mediator system, but it is still very impressive (i.e., 40% delignification in 4 h). The areas for development of this technology are the need for careful control of the levels of peroxide, manganese, chelating agent, and enzyme and the costs.

Deinking

The removal of ink is an important part of secondary fiber processing. Large-scale deinking is new and evolving rapidly. Deinking is carried out by repulping at 3 to 4% solids consistency, then diluting to 1% consistency and aerating the slurry while adding flocculating surfactants and ink solvents. The ink particles float to the surface and are collected and removed.

Several reports indicate that cellulase enzymes can increase the efficiency of deinking processes (20,21). It appears that cellulase can release ink particles, both bound to fiber fines and separately from fiber, and thereby enhance the removal of ink by flotation. This in turn decreases the ink count, allows a higher brightness to be achieved, or alternatively decreases the usage of the deinking chemicals. There is some question, however, as to whether the enzyme treatment decreases the strength of the pulp (22). Some reports state that the enzymatic deinking process is best carried out under alkaline conditions (23). There has been one published report of a mill scale trial using this enzymatic deinking; the results confirmed the laboratory data (24). EDT of Norcross, Ga., has marketed cellulase and xylanase enzymes for mill-scale deinking applications.

ENZYMES IN PAPERMAKING

Most of the content of paper is finished pulp (which can be blends of both mechanical and chemical pulps) and inorganic fillers such as clay. The remainder of the paper, perhaps up to 10% of the total weight, consists of a variety of additives designed to achieve certain properties in the paper and other additives to improve the operation of the paper machine in the mill. Among the additives to influence paper properties is modified starch for coated paper. Amylase enzymes are used in the starch modification. Enzymes are also used to remove pitch from the paper and to eliminate slime buildup on the paper machine.

Starch Modification

The oldest use of enzymes in the pulp and paper industry is in the modification of starches for coated paper. This practice dates back to the early 1970s and is now carried out more frequently within the starch processing plants than in the paper mills. Starch imparts many beneficial properties to paper, including strength, stiffness, and erasability. The starches used in papermaking are often modifications of the natural starch polymers found in the locale of the mill, such as corn starch in the United States, potato starch in Britain, and rice starch in Japan. Of particular interest in enzyme use is starch for sizing. Sizing refers to an increased stiffness of the paper that is useful as a coating or for further paper processing.

The starch is added to the paper partway through the drying section of a paper machine. The starch is added at a size press, which is two rollers through which the paper picks up starch solution while passing. The presses run best at 45 to 60 °C (113 to 140 °F), so the viscosity of the starch solution must be in an acceptable range at this temperature. The standard starch products for the paper industry are suitable for some grades of paper, but in other cases the viscosity is too high for good spreading through the press. In other cases, a mill desires the flexibility to vary the starch viscosity for different paper grades. More recent work has suggested that amylase enzymes can form unique starch products for the paper industry (25).

The generic process for decreasing starch viscosity is to decrease its molecular weight by hydrolyzing the polymer to shorter chains. The first method to do this involved a dilute acid cooking of the starch; this is still practiced in some mills today. The second method, hydrolysis with α-amylase enzymes, was originally carried out in the late 1960s. These enzymes catalyze the addition of a water molecule across the linkages between two glucose units in the starch polymer. The desired reactions for paper starch is to shorten the starch polymer. Complete conversion to glucose, which would be achieved using glucoamylase and pullulanase enzymes, is undesirable.

The two types of reactor systems used for enzymatic starch hydrolysis in paper mills are batch and continuous.

In batch systems (Fig. 5), roughly 1,000 kg (2,200 lb) of starch is diluted in 4,000 L (1,000 gallons) of water in an agitated tank. The temperature is maintained at, depending on the enzyme used, 25 to 90 °C (77 to 194 °F). The pH is controlled to 4.5 to 7. The treatment time is typically 30 minutes, after which the system is heated to a suitable temperature to destroy the enzyme's activity, or, for the thermostable enzymes, chemical denaturants are added. Batch systems offer the advantages of simplicity of control and the ability to run at solids levels of up to 40 %.

In continuous systems (Fig. 6), the starch and water are mixed in a slurry tank and pumped through a steam-heated tube. Enzyme is injected at the mouth of the tube. The running temperature is 125 °C (257 °F) for 2 min at pH 6. At this point, the target viscosity is reached, and the enzyme, which is added to a concentration of 0.5 g/L, is fully denatured. This system offers advantages of easy automation and high volumetric productivity. The continuous systems are generally for the larger starch users.

Pitch Removal

Pitch is the sticky substance that is prevalent in softwoods, especially pine. Pitch is a collection of compounds known as lipids, which are long-chain fatty acids connected by glycerol linkages. These compounds are analogous to the fatty tissue of mammals and have both hydrophobic and hydrophilic sites on the molecule. This character contributes to the processing difficulties caused by pitch; the compounds are slightly soluble in some points of the process, and they can form a difficult precipitate in other areas. The most troublesome pitch formation is from Japanese Red Pine. Pitch comprises typically less than 1% of the weight of wood, but can reach 3% of the total on this species.

The removal of pitch by chemical pulping and bleaching is not particularly efficient, but it is usually sufficient to prevent severe deposition. In mechanical pulping, the pitch problems are more severe, particularly with pitch exposed on the paper machine. At this point, pitch interacts with other additives, decreasing their effectiveness as well as decreasing the rate of drainage of the sheet. Chemical dispersants have been used for pitch control, and the pitch content is decreased by longer storage periods in the woodyard, which is limited by fungal growth and chip deterioration.

A third method of decreasing pitch deposition is by using lipase enzymes to digest the pitch. Lipases hydrolyze the lipids to their constituent fatty acids and substituted polyol (Fig. 7). The polyol (glycerol for triglycerides) is water soluble, and the fatty acids can be bound with talc or other charged compounds. Lipase for pitch removal has been particularly successful in Japan, where 120,000 to 140,000 tons of groundwood pulp are treated annually with lipase, representing 20% of the total groundwood pulp produced in Japan (26).

Lipase has been reported to be added to the groundwood pulping line and to the storage chest before the paper machine (26). When added to the pulping line (just before the post refiner), a mixture of two lipases added at levels of about 100 ppm and about 600 ppm reduced the triglyceride content of the pulp by 70% over a 1-month mill trial, with an enzyme reaction time of 3 to 3.5 h. This corresponded to a decease in pitch deposits on the paper and equipment and defects in the resulting paper sheets. However, one disadvantage of this location for lipase treatment is the frequent use of alum, particularly in American mills, which can inactivate the enzyme.

In another mill trial, lipase was added at 1 to 5 ppm in the storage chest before the paper machine, for a retention time of at least 1 h. The process temperature and pH of 40 °C (104 °C) and 6.0, respectively, were compatible with the enzyme treatment. The enzyme was added to the dilute slurry in the chest without additional mixing, although in at least one mill lack of good enzyme dispersion caused substandard results (27). Results of a mill trial using lipase for pitch removal are shown in Table 3. Lipase does not affect the strength or brightness of the pulp. The yeasts *Candida cylindracae* and *Candida rugosa* and *Aspergillus* fungi have been reported to make lipase in high productivity and are used commercially to make lipase.

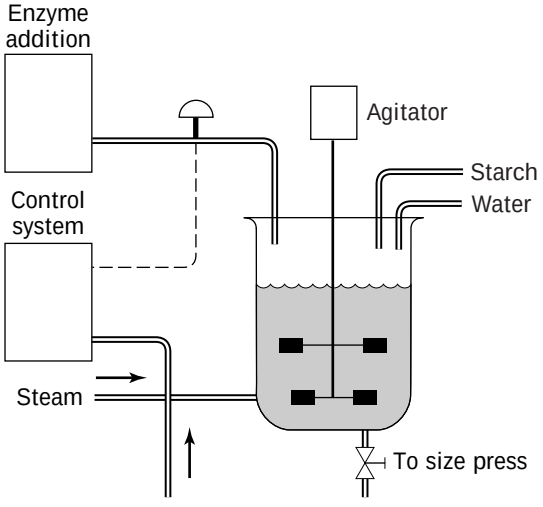


Figure 5. Batch starch conversion system.

Slime Removal

Like pitch, slime deposition causes significant operating problems around a paper machine. Slime is the generic name for deposits of a microbial origin in a paper mill. It is impractical to run a paper mill as a sterile system. As a result, a vast array of microbes contaminate the mills, and many of the resulting slime compounds have not been characterized. When confronted with a slime, often the strategy is to try every available biocide until one is found that targets the microbe and destroys the source of the slime.

In some cases, however, specific slime compounds have been characterized, and efficient strategies for their removal can be used (28). One such case is with levan, which is a β-2,6-linked polymer of fructose that forms a slime film. This compound is secreted by several species of *Bacillus* and *Pseudomonas* bacteria that can grow in the recirculated water around the paper machine, especially for fine paper, where the level of inhibiting compounds is low. The

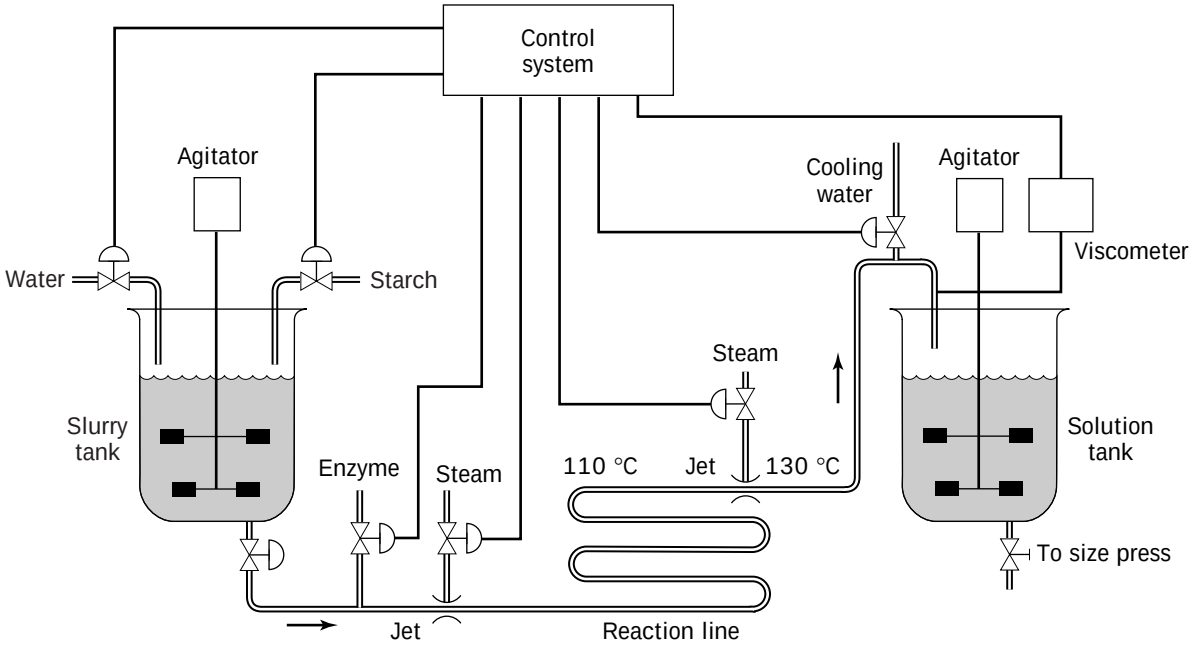


Figure 6. Continuous starch conversion system.

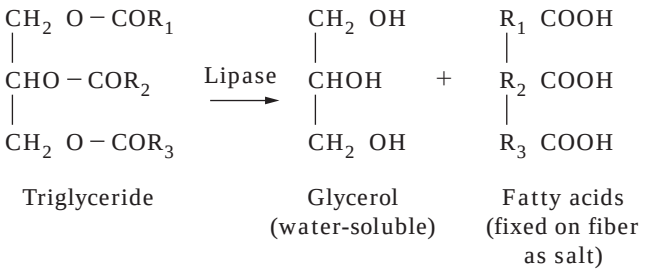


Figure 7. Hydrolysis of pitch by lipase.

Table 3. Lipase for Pitch Removal

	Newsprint		Telephone book	
Parameter	Control[a]	Lipase[b]	Control	Lipase
Pitch, g/day	242.6	32.6	110.2	46.1
Paper web, holes/day	11.0	1.4	4.5	3.4

Source: Data from Irie et al. (27).
[a]Red pine groundwood pulp, 200 ton/day, paper machine speed 830 m/min.
[b]Enzyme added to mixing chest, retention time 60–120 min, lipase 3 ppm.

enzyme levan hydrolase can hydrolyze this polymer to low molecular weight polymers that are water soluble, thereby cleaning the slime out of the system. Another advantage of the enzymatic treatment is that it destroys the ability of the microbe to bind to the surface, thereby increasing the effectiveness of chemical biocides.

Commercial levan hydrolase is supplied as the product EDC-1 by Henkel Corporation (Morristown, N.J.). The enzyme does not hydrolyze cellulose, so is not harmful to the paper. The enzyme is usually added at the headbox of the paper machine, although in some cases it has been added at the dryer discharge. The enzyme is effective at pH 4 to 8 and runs best at 50 °C. The enzyme can be used alone or in combination with chemical biocides.

Other enzymes reported to be effective for slime removal are proteases, cellulase, amylase, and pentosanase (28). These enzymes are used more typically to treat the process waters of the mill to prevent slime buildup in cooling towers, heat exchanger surfaces, and other surfaces exposed to water.

BIBLIOGRAPHY

1. T.K. Kirk, H.M. Chang, *4th Int. Conf. on Biotechnology in the Pulp and Paper Industry Proc.*, Raleigh, N.C., May 16–19, 1989.
2. R.L. Farrell, R.A. Blanchette, T.S. Brush, B. Gysin, Y. Hadar, J.J. Perollaz, P.A. Wendler, W. Zimmerman, in *Proc. of the Fifth Int. Conf. of Biotechnology in the Pulp and Paper Industry*, Kyoto, Japan, May 27–30, 1992, pp. 27–32.
3. J.C. Pommier, J.L. Fuentes, G. Goma, *TAPPI* 187–191 (June 1989).
4. L.S. Jackson, J.A. Heitmann, T.W. Joyce, TAPPI **76**, 147–154 (1993).
5. L. Viikari, M. Ranua, A. Kantelinen, J. Sundquist, M. Linko, *Proc. Int. Conf. Biotechnol. Pulp and Paper Ind.*, Stockholm, 1986, pp. 67–69.
6. J.S. Tolan, *Proc. Non-Chlorine Bleaching Conf.*, Hilton Head, S.C. March 12–17, 1993.
7. J.S. Tolan, *TAPPI Pulping Conference Proc.*, Boston, November 6–10, 1992.
8. L.U.L. Tan, K.K.Y. Wong, J.N. Saddler, *Enz. Microb. Technol.* **7**, 425–430 (1985).
9. U.S. Pat. 5,320,960 (June 14, 1994), B. Bower (to Genencor International).

10. J.M. Chauvet, J. Comtat, P. Noe, *Paris Symposium Wood and Pulping Chemistry*, 1987, pp. 325–327.
11. A. Lappalainen, *VTT Publication* **50**, 22 (1988).
12. K. Poutanen, *VTT Publication* **47**, 13–20 (1988).
13. U.S. Pat. 5,298,405 (March 29, 1994), H. Nevalainen, J. Knowles, P. Suominen, M. Pentilla, A. Mantyla (to Alko Ltd.).
14. A. Suurnakki, M. Tenkanen, J. Buchert, L. Viikari, L., *Adv. Biochem. Eng. Biotechn.* **57**, 261–287 (1997).
15. H.P. Call, Proc. of the 1994 TAPPI Pulping Conference, San Diego.
16. P. Ander, G. Daniel, H. Pettersson, U. Westermark, in T. Jeffries and L. Viikari eds., *Enzymes for Pulp and Paper Processing*, ACS Symposium Series 655, 1997, pp. 297–307.
17. Eur. Pat. Appl. 717,143 (June 20, 1996), H.P. Call (to Lignozym GmbH).
18. PCT Pat. Appl. WO 95/01426 (January 12, 1995), P. Schneider, A. Pedersen (to Novo Nordisk).
19. M.G. Paice, R. Bourbonnais, I. Reid, F.S. Archibald, 9th *Int. Symposium on Wood and Pulping Chemistry Proc.* Montreal, 1997, pp. 1–4.
20. J.A. Heitmann, T.W. Joyce, D.Y. Prasad, in *Proc. of the Fifth Int. Conf. of Biotechnology in the Pulp and Paper Industry*, Kyoto, Japan, May 27–30, 1992, pp. 175–180.
21. K. Rutledge-Cropsey, T. Jeffries, J.H. Klungness, M. Sykes, *TAPPI 1994 Recycling Symposium*, Boston, May 1994, pp. 103–105.
22. G. Stork, H. Preira, T.M. Wood, E.M. Dusterhoft, A. Toft, J. Puls, *TAPPI 1994 Recycling Symposium*, Boston, pp. 103–105.
23. U.S. Pat. 5,364,501 (November 15, 1994), J.A.G. Baret, M. Leclerc, J.P. Lamort (to Novo Nordisk).
24. T.J. Kim, S.S.K. Ow, T.J. Eom, in *Pulping Conference 1991 TAPPI Proc.*, October 1991, pp. 1023–1027.
25. S.E. Nachtergale, F.J. van der Meeren, *Stark* **39**, 135–141 (1987).
26. K. Hata, M. Matsukura, H. Taneda, Y. Fujita, in T. Jeffries and L. Viikari eds., *Enzymes for Pulp and Paper Processing*, ACS Symposium Series 655, 1997, pp. 280–296.
27. Y. Irie, M. Matsukura, M. Usui, K. Hata, *Papermakers Conference / TAPPI Proc.* 1990, p. 10.
28. S.C. Johnsrud, *Adv. Biochem. Eng. Biotechn.* **57**, 311–328 (1997).

See also BIOREMEDIATION; ENZYMES, EXTREMELY THERMOSTABLE; ENZYMES, PROTEINS HYDROLYSIS; HEMICELLULASES; HEMICELLULOSE CONVERSION.

ENZYMES, STARCH CONVERSION

I.S. BENTLEY
ABM Brewing and Enzymes Group/Rhone-Poulenc
Stockport, U.K.

KEY WORDS

Corn
Enzymes
Fructose
Glucose
Maize
Maltodextrins
Maltose
Starch
Syrups
Wheat

OUTLINE

INTRODUCTION

Starch is the main storage carbohydrate of plants and a primary source of energy in the human diet. It can be found in cereals such as wheat or barley, in the seeds of the maize plant, and in the tuberous roots of potato and cassava. Industrially, starch and starch-derived sweeteners are mainly produced from maize, wheat, potato, and tapioca. The typical starch contents of these raw materials are given in Table 1.

The use of carbohydrate hydrolyzing enzymes to convert starch to nutritive sweeteners is not a new technology. For centuries malt has been used to produce a sweet syrup from starch, although it was not known at the time that the enzymes in malt were responsible for the reaction. It was the development of amylases from fungal and bacterial sources at the turn of the twentieth century that paved the way for the use of enzymes in starch hydrolysis. Even so, it was not until the 1960s that enzymes began to replace mineral acids in the production of dextrose from starch.

The term *glucose syrup* is often used to describe the refined, concentrated, aqueous solution of saccharides obtained by the controlled partial hydrolysis of edible starch.

Table 1. Starch Content of Basic Raw Materials

Raw material	Starch content (%)
Maize	72
Wheat	64
Potato	15
Tapioca	20–30

This conversion may be catalyzed by food-grade acids and enzymes. Depending on the degree of hydrolysis, the resultant syrup contains varying amounts of D-glucose (1). Syrup produced from maize starch is often called corn syrup rather than glucose syrup.

The term *glucose syrup* is ambiguous. In the glucose syrup industry, glucose is used to describe any starch-derived syrup; however, from a chemical viewpoint, it is the monomeric saccharide α-D-glucose. To distinguish the monomer from the syrup, the term *dextrose* is used to describe the monomer. Because starch is composed of D-glucose units, dextrose is the ultimate product of nondestructive starch hydrolysis (2).

The extent to which the starch molecules are broken down is described by the term *dextrose equivalent* (DE). Starch has a DE of 0, whereas the ultimate product, dextrose, has a DE of 100. In general, sweetness will increase with increasing DE value, although the nutritive value will remain the same. It is the DE value that is important in describing a glucose syrup, since the characteristics of the syrup will vary with the DE.

Products with a DE of less than 20 are referred to as maltodextrins, those with a DE of between 20 and 80 as glucose syrups, and those with a DE greater than 80 as hydrolysates. At a DE lower than 20, the functionality of the maltodextrin is often related to the botanical source of the starch. This is related to the ratio of amylose and amylopectin within the particular starch structure.

The Structure of Starch

In its native state, starch is present in plant cells as microscopic granules arranged in layers. These granules vary in size and shape—from the relatively large granules found in potato (50–150 μm) to the smaller granules of maize and wheat (2–40 μm). Starch is made up of two types of high molecular weight polymers of glucose, amylose and amylopectin. Amylose and amylopectin are found concomitantly within the starch granule, and the ratio varies depending on the source of the starch. Table 2 shows typical levels of amylose and amylopectin in some common starches.

Starches from tubers, such as potato and tapioca, tend to have lower levels of amylose compared to starches from grain. The ratio of amylose to amylopectin also varies with the variety of the crop. For example, some maize varieties yield up to 99% amylopectin (waxy maize), whereas other starches have a high amylose content (high amylose starches). These are specialist starches, and unlikely to be used in the production of glucose syrup.

The basic building block of both amylose and amylopectin is α-D-glucose monomer linked through an α-1,4 glucosidic bond. In amylose, the glucose monomers are linked solely through the α-1,4 glucosidic bonds to yield linear chains of several hundred glucose units. In solution, the amylose chains may be coiled in the form of a helix. In amylopectin, the basic structure is still linear chains of glucose monomers linked through α-1,4 glucosidic bonds, but these are linked to other chains through α-1,6 glucosidic bonds. These branch points occur every 10–12 glucose units. This gives amylopectin a branched structure that may contain many thousands of glucose units per molecule.

Table 2. Level of Amylose and Amylopectin in Common Starches

	Amylose (%)	Amylopectin (%)
Maize	28	72
Wheat	28	72
Potato	20	80
Tapioca	17	83

In a starch granule, amylose and amylopectin are linked together through hydrogen bonding to form structures that contain both crystalline and amorphous regions. It is the intra- and intermolecular hydrogen bonding that gives starch its insolubility in water. On heating an aqueous suspension of starch, the hydrogen bonds are disrupted and the starch disperses. On cooling, the hydrogen bonds reform and the starch precipitates. The rate at which these hydrogen bonds form is related to the concentration of starch in the dispersion and the temperature. This precipitation is called retrogradation and can cause low yields and processing difficulties if allowed to occur in the production of glucose syrups.

USE OF ENZYMES IN THE CONVERSION OF STARCH

The manufacture of glucose syrup and nutritive sweeteners from starch-containing raw materials can be divided into a number of individual processes:

1. Starch separation
2. Gelatinization and liquefaction
3. Dextrinization
4. Saccharification
5. Isomerization

Within each of these defined stages, the use of microbial enzymes facilitates the process to give a product that in turn becomes the raw material for the subsequent stage. The first of the stages that makes use of enzymes is the separation of the starch from the starch-containing raw material.

Starch Separation

The separation of starch is primarily based on physical processes. With tubers such as potatoes and tapioca, the starch is separated using a process of chopping, disintegration (rasping), and sieving. It is unusual to use enzymes to facilitate the separation of starch from tuberous raw materials.

The processes of starch separation for maize and wheat are much more complex than those for potato and tapioca. In recent years enzymes have come to be used extensively to improve the starch separation from wheat. Enzymes can be used to improve the yield of starch from maize (3). In both processes the enzymes employed act on hemicellulose compounds such as arabinoxylans and β-glucans. The particular enzyme depends on the type of hemicellulose pres-

ent in the raw material. Typical values for arabinoxylans and β-glucans in maize and wheat are given in Table 3.

In the case of maize, whole maize kernels are steeped in an aqueous environment under controlled conditions of temperature, time, and concentrations of sulfur dioxide and lactic acid. The main purpose of the steeping process is to soften the maize kernel for milling while inhibiting microbial growth to enable optimum starch recovery (4). It has been suggested that the addition of enzymes to the steep tank improves the separation of the starch from the fiber by degrading the arabinoxylans that bind the starch to the fiber.

Studies have shown that the use of an enzyme preparation in the steeping of maize can result in increased starch yield, increased grind capacity, and decreased steep times (5). The effectiveness of enzyme preparations when added to the steeping of maize is a subject of some debate due to the increase in starch yield versus the amount of enzyme added. It has also been suggested that enzymes are not absorbed along with the water into the maize kernel because the enzyme is a large molecule that could not penetrate the kernel through the tip cap.

This problem could be overcome by adding the enzyme during the first grind (6), after the initial cracking of the maize, allowing the enzyme access to the arabinoxylan gums that bind the starch. Many maize starch processors do not have the facilities to have a holding time between the first and second grind to allow enzyme action. It is up to the individual maize starch processor, therefore, to determine the feasibility of using enzymes.

In contrast to the separation of corn starch, wheat starch is produced by at least five different commercial processes. In all of these, wheat is ground and combined with water to form either a dough or a batter. The fractionation is achieved by physical means, and vital wheat gluten and starch are products of high value.

Wheat flour contains a high level of arabinoxylans (Table 3), a large proportion of which are soluble in water. These high molecular weight gums can bind significant amounts of water and hence give rise to high viscosities in doughs or batters. Addition of an enzyme (xylanase) to the dough or batter can reduce the viscosity by degrading the arabinoxylan and, hence, improve the separation process. The enzymes used not only improve the separation of gluten and starch but also enhance gluten coagulation.

In all enzymes used to aid starch separation, the presence of proteinase or amylase side activities is detrimental to the process, since they will break down protein and starch, cause material loss, and cause deterioration in the quality of the end products, gluten and starch.

Once separated, the starch may be dried as a product in itself or further processed to nutritive sweeteners. The conversion of starch to nutritive sweeteners is the main use of enzymes in the starch industry.

Table 3. Percentage of β-Glucan and Arabinoxylan in Maize and Wheat

	β-glucan (%)	Arabinoxylan (%)
Maize	0.3	3.5
Wheat	0.4	8.5

Gelatinization

At ambient temperatures starch granules are insoluble in water, however, if an aqueous starch suspension is heated to above 60 °C the granules begin to swell and disrupt. Eventually the starch molecules are dispersed into solution. This process is known as gelatinization.

The temperature required for complete dispersion or gelatinization depends on the source of the starch. For most starches, a temperature of 105–110 °C is sufficient. At 25–40% dry solids, the viscosity of gelatinized starch is extremely high, and so a thinning agent is required. Viscosity reduction may be achieved by a combination of the shear forces exerted on the slurry as it is continuously pumped through a jet cooker (mechanical thinning) and the action of acid or enzyme. The gelatinization temperatures of starches from the most commonly used raw materials are given in Table 4.

During the process of gelatinization and thinning of a starch slurry, proteins, lipids, and hemicellulases that complex with the amylose and amylopectin are released and coagulate to some extent. The lipids that are complexed mainly with the amylose are dissociated at higher temperatures. Subsequent cooling may again lead to the precipitation of the amylose–lipid complexes. These precipitates cause hazes and flocculation later in the process and are difficult to remove. They are often the cause of a positive iodine reaction to starch in saccharification reactors (7).

To prevent retrogradation and reprecipitation of starch–amylose complexes, it is necessary to continue the thinning process until partial hydrolysis occurs. This process is called liquefaction.

Liquefaction

Liquefaction is the first primary action in hydrolysis or thinning of the starch slurry whereby viscosity is reduced and first-stage degradation of starch molecules takes place (Fig. 1).

Acid Liquefaction Process. Traditionally, the process of thinning, liquefaction, and dextrinization of starch was

Table 4. Gelatinization Temperature for Some Common Starches

Starch type	Gelatinization temperature (°C)
Maize	72–76
Wheat	60–64
Potato	65–70
Tapioca	70–75

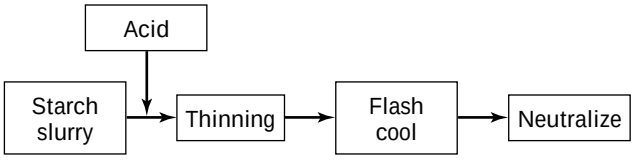


Figure 1. Scheme for the acid liquefaction of starch; thinning conditions: pH 1.5–2.0, 140–150 °C, 5–8 min.

carried out using acid. In acid liquefaction, a starch slurry of between 30 and 40% dry solids is adjusted to a pH of 1.5–2 with addition of an acid, usually hydrochloric acid.

The acidified slurry is then heated to 140–150 °C by either direct steam injection in a jet cooker or indirect heating in a reactor for approximately 5–8 min. The hydrolysate is then flash-cooled to atmospheric pressure and neutralized (Fig. 1).

Acid liquefaction results in complete gelatinization, thinning, and a liquefact that can be easily filtered. However, the process also results in a product with a high color and ash content, and significant purification costs. The action of the acid on the starch molecules is nonspecific, and by-products: such as furfuraldehyde derivatives are formed, resulting in a lower yield of desired products.

The use of a heat-stable endo-α-amylase allows for less rigorous conditions to be employed in the liquefaction process. By-product formation becomes less of a problem and refining costs are reduced.

Two-Stage Enzyme Liquefaction Process. The first enzyme to be employed in starch liquefaction was the endo-α-amylase produced by *Bacillus subtilis* var. *amyloliquefaciens*. The enzyme is utilized in a two-stage enzyme liquefaction process. A starch slurry at 30–40% dry solids is thinned at a temperature of 85 °C for 1 h at pH 6.0–6.5. This is followed by a short heat treatment at 140 °C, flash-cooled to 85 °C (at which point a second enzyme addition is made), and the reaction continued until liquefaction is complete. For optimum enzyme stability a calcium concentration of 150 ppm is required; this is usually added as calcium hydroxide or calcium chloride into the starch slurry (Fig. 2).

If a thermostable α-amylase is employed, such as that from *Bacillus licheniformis* or *Bacillus stearothermophilus*, the heat treatment at 140 °C may be eliminated. The use of thermostable α-amylases has largely replaced acid liquefaction and processing using the less heat-stable α-amylase from *B. subtilis*.

Single-Stage Enzyme Liquefaction Process. A starch slurry at 30–40% dry solids is prepared, and the pH is adjusted to 5.8–6.2 by the addition of sodium of calcium hydroxide. Calcium salts are added, if necessary, to give a level of 75–100 ppm of free calcium ions. The slurry is pumped continuously through a jet cooker operating at 103–107 °C, enzyme having already been added to the slurry stream. The higher shear produced by the action of pumping through the jet cooker gives a degree of mechanical thinning, and this, in addition to the action of the α-amylase, means that peak viscosities are avoided. The slurry is then transferred to a holding coil, which maintains the temperature for 5–10 min before the liquefied starch is flash-cooled to 95 °C, where dextrinization continues.

In the single-stage liquefaction process, the control of the jet cooker is critical (Fig. 3). Control of the temperature of the slurry through the jet cooker and into the holding coil must be accurate. If the temperature is allowed to drop much below 105 °C, incomplete gelatinization will occur and filtration problems may be encountered later in the process. However, at temperatures much above 105 °C, the enzyme will be inactivated and the viscosity of the slurry is not effectively reduced. Operating pH should be between 5.8 and 6.2, although lower pH values (down to 5.5) have been reported with the use of *B. stearothermophilus* heat-stable α-amylase. At lower pH values the enzymes are less stable and, in combination with high temperatures, lead to rapid inactivation; on the other hand, higher pH levels may cause color formation and production of by-products, resulting in lower end product yields.

Other factors that affect the performance of the liquefaction process are dry substance and calcium ion concentration. Higher levels of dry substance have a stabilizing effect on the enzyme but may also lead to incomplete gelatinization and subsequent filtration problems. A dry substance level of 30–35% gives adequate enzyme stability, and at the same time ensures complete gelatinization. Calcium is required to give the heat-stable α-amylases stability at higher temperatures; for *B. licheniformis* enzyme the requirement is 75–100 ppm free calcium ions, and for *B. stearothermophilus* it is 100–150 ppm. The level of calcium ions is usually adjusted by addition of calcium hydroxide or calcium chloride to adjust the pH of the slurry.

The final parameter that requires consideration is enzyme dose rate: Too little enzyme leads to incomplete liquefaction and starch positive reaction after dextrinization; too much enzyme results in a large enzyme carryover into dextrinization and a dextrose equivalent higher than desired. The DE is used to monitor the degree of starch breakdown. It is calculated by expressing the reducing sugars, such as dextrose, as a percentage of the dry substance. Typical enzyme dose rates for the single-stage liquefaction process range from 0.6 to 1.0 L of a heat-stable α-amylase per ton of dry substance.

Enzymes produced from *B. licheniformis* give different starch breakdown products to heat-stable α-amylases produced using *B. stearothermophilus*. The difference shows in the measurement of DE for the same degree of starch degradation. The DE given by *B. stearothermophilus* enzyme can be as much as two percentage points higher than *B. licheniformis* enzyme.

Because of their different breakdown products, mixtures of heat-stable α-amylases from *B. licheniformis* and *B. stearothermophilus* give a more efficient breakdown of starch during liquefaction. A further potential use of the

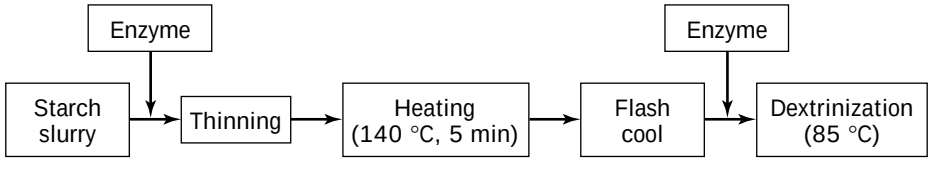


Figure 2. Liquefaction using a non-heat-stable α-amylase (*Bacillus subtilis*); thinning conditions: pH 6.0–6.5, 80–90 °C, 5–8 min.

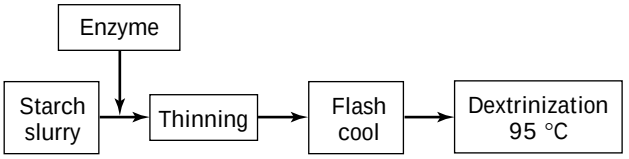


Figure 3. Single-stage liquefaction process; thinning conditions: pH 5.8–6.2, 105–107 °C, 5–11 min.

B. stearothermophilus heat-stable α-amylase is using lower pH liquefaction; the enzyme is more stable than *B. licheniformis* enzymes at pH values down to 5.5 (8). The advantage of lower-pH liquefaction is reduced by-products and color formation.

Acid or Enzyme Liquefaction. A process that takes advantage of acid thinning (which can improve filtration downstream due to better separation of fat and protein) and enzyme liquefaction is the acid or enzyme liquefaction process.

A starch slurry containing between 30 and 40% dry substance is adjusted to give a pH of 2–5. The slurry is jet cooked at 140 °C for about 5 min, where mechanical thinning due to both shear and acid thinning takes place. The pH is very important, if it is too low, color and by-product formation occurs, but if it is too high, insufficient thinning results.

After this initial cooling, the thinned starch is cooled to 95–100 °C, the pH is adjusted to 5.8–6.2, and enzyme is added. The major disadvantages of this process are increased energy consumption and potential by-product formation.

Dextrinization

After liquefaction of the starch slurry, the action of the heat-stable α-amylase on the starch is allowed to continue in a process called dextrinization. This process involves further starch degradation until a syrup with the required dextrose equivalent is obtained. Dextrinization may involve the further addition of α-amylase, as in the case of maltodextrin production, or there may be sufficient enzyme remaining after the cooling process.

The degree to which dextrinization is allowed to continue is determined by the type of product manufactured. In the production of high-dextrose syrups (i.e., syrups containing 94–97% dextrose), dextrinization is allowed to continue to a DE of 15–18. However, in the production of very-high-maltose syrups (68–75% maltose) the DE must be controlled to a maximum of 12. Addition of a further quantity of bacterial α-amylase allows maltodextrin with intermediate DEs up to 30 to be prepared.

Once the desired DE has been reached during dextrinization, the α-amylase may require inactivation. This can be achieved by a second jet-cooking step at temperatures of 120 °C and above with a low pH, preferably between 3.8 and 4.5. The resulting hydrolysate may then be filtered, clarified with carbon, concentrated, and dried or spray-dried to give a maltodextrin.

The process of liquefaction and dextrinization produces a hydrolyzed starch suspension of low viscosity that is stable to retrogradation and suitable for further enzyme processing to high-dextrose, glucose, or high-maltose syrups.

Saccharification

The process in which the dextrinized starch is converted to nutritive sweeteners is called saccharification (Fig. 4). Further hydrolysis of the oligosaccharides produced by the process of liquefaction and dextrinization may be carried out by a number of different exoamylases and debranching enzymes, depending on the required product (Table 5).

Dextrose (Glucose) Production. During the production of glucose, the enzyme amyloglucosidase is used in saccharification (Fig. 5). Amyloglucosidase is an exo-α-amylase that catalyzes the hydrolysis of α-glucosidic bonds by successively recovering glucose from the nonreducing end of oligosaccharides in a stepwise manner. The enzyme is able to hydrolyze both α-1,4 and α-1,6 glucosidic bonds; however, the rate of hydrolysis of α-1,6 bonds is much slower than the rate of hydrolysis of α-1,4 bonds. It has been reported that amyloglucosidase hydrolyzes α-1,4 bonds between 20 and 30 times faster than α-1,6 bonds.

Amyloglucosidase preparations are usually available at activities of either 200 or 300 U mL^{-1}. After liquefaction and dextrinization the solution is cooled to 60 °C and the pHs adjusted to 4.2–4.6 with acid. After dextrinization, the dry substance should be between 25 and 37%. Amyloglu-

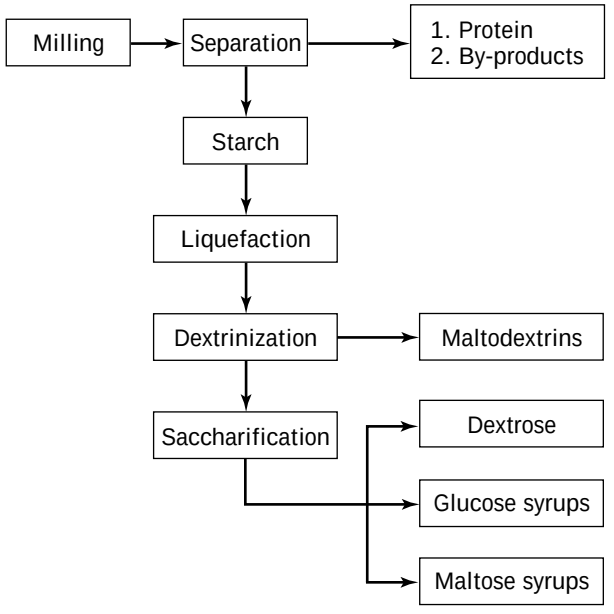


Figure 4. Production of nutritive sweeteners from starch.

Table 5. Listing of the Main Saccharification Enzyme Required for Various Desired Products

Product required	Main saccharification enzyme
Glucose	Amyloglucosidase
Maltose (50%)	Fungal α-amylase
Maltose (≥50%)	Cereal β-amylase

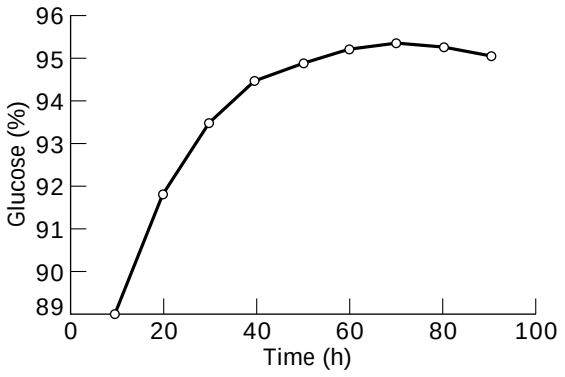


Figure 5. Typical saccharification profile with amyloglucosidase.

cosidase (300 U mL^{-1}) is added, typically at dose rates of between 0.5 and 1.1 L ton^{-1} dry substance, depending on the required saccharification time. The saccharification reaction takes place in stirred reactors and is allowed to continue until peak dextrose is reached, typically 40–72 h.

Given that the enzyme is able to hydrolyze both types of bonds present in starch, it would be reasonable to expect that amyloglucosidase is able to produce quantitative amounts of glucose from liquefied starch. At very low solid levels, glucose levels approaching 100% are possible, but at the higher solid levels used in industrial saccharification this is not the case.

Amyloglucosidase can polymerize glucose by the reverse reaction of hydrolysis. This reversion is a condensation reaction and is accompanied by the elimination of water. The products of this reversion are maltose and isomaltose. The reversion reaction takes place at a much slower rate than the hydrolysis reaction and is dependent on the dry substance level and the residual enzyme activity at the end of saccharification. At lower dry substance the rate of reversion is slower and higher levels of dextrose can be attained (Fig. 6).

Peak dextrose levels may also be reduced by the presence of a secondary enzyme activity within amyloglucosidase preparations, namely transglucosidase. This enzyme is a glucose transferase and catalyzes the formation of α-1,6 branched oligosaccharides such as isomaltose and panose from glucose and maltose. The buildup of isomaltose and panose during saccharification leads to an overall reduction in final dextrose levels. Transglucosidase activities have recently been essentially removed from amyloglucosidase preparations either after the fermentation process or by selection of an *Aspergillus niger* strain that does not produce the activity.

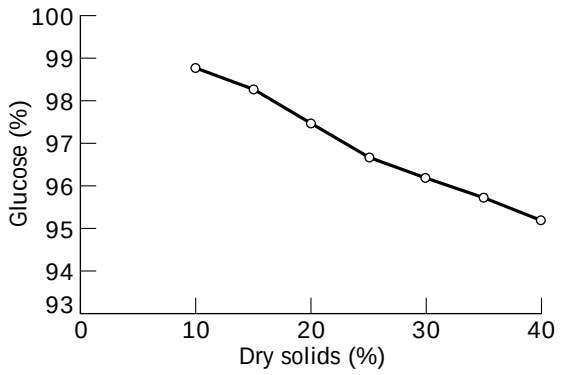


Figure 6. The effect of solids on maximum dextrose yield.

Amyloglucosidase itself is very inefficient at hydrolyzing the products of the reverse reaction, such as isomaltose, isomaltotriose, and panose. The formation of these products is the major factor in governing the dextrose level after saccharification. It is important, therefore, to limit the rate of the reversion reaction as much as possible. One of the ways to achieve this is to lower the level of amyloglucosidase in saccharification.

The relative inefficiency of amyloglucosidase in hydrolyzing α-1,6 glucosidic linkages in the branched oligosaccharides is a major obstacle to dextrose production. An excess of enzyme is therefore required to achieve the desired saccharification rates. This excess results in a higher rate of reversion and ultimately lower dextrose levels.

It is possible to reduce the level of amyloglucosidase and maintain the reaction rate by the addition of an enzyme that can specifically hydrolyze the α-1,6 links of the branched oligosaccharides. This enzyme is called pullulanase, and blends of pullulanase and amyloglucosidase are now available. The addition of a specific α-1,6 debranching enzyme augments the α-1,6 activity of amyloglucosidase and, therefore, the addition rate of the glucose-producing enzyme may be lowered. The reduction in amyloglucosidase in the saccharification tank decreases the rate of the reversion reaction. The overall effect is an increase in the peak dextrose level. The dose rate and conditions of use of the blended products should be the same as the dose rate for amyloglucosidase alone.

The choice of an amyloglucosidase or a blend of pullulanase and amyloglucosidase depends on the individual saccharification requirements. A fast saccharification can be achieved by a high addition of amyloglucosidase, but this in turn will mean quicker reversion rates; therefore, if dextrose levels are important, the enzyme blend is the preferred route. Often the residual oligosaccharides referred to as DP_4+ are considered undesirable, and again the enzyme blend has been shown to lower the levels of this oligosaccharide fraction. The use of an amyloglucosidase–pullulanase blend gives greater flexibility and control over the saccharification process.

Further alternatives are available to achieve specific saccharification requirements. For example, all amyloglucosidase preparations contain acidic α-amylase activity. The presence of this activity enhances the saccharification rate by degrading the higher molecular weight oligosaccharides into lower molecular weight fractions, which also results in a reduction of DP_4+ oligosaccharides in the final product.

Higher dextrose levels may also be obtained by addition of an α-1,4-glucosyl transferase to amyloglucosidase (9). Unlike transglucosidase, which catalyzes the formation of α-1,6 glucosyl linkages, α-1,4 glucosyl transferase produces α-1,4 glucosyl linkages. Such an enzyme is produced by the fermentation of *Bacillus megaterium*. Substrates for

the enzyme include isomaltose and isomaltotriose; these are combined with glucose to form molecules that can then be hydrolyzed by amyloglucosidase.

A combination of amyloglucosidase and α-1,4 glucosyl transferase results in higher dextrose levels by hydrolysis of the products of the reversion reaction. The enzyme may be added at the start of saccharification along with amyloglucosidase (10), however, the pH optimum of Megadex is above that of amyloglucosidase and the enzyme may be added to the process after a pH adjustment.

Maltose Syrups. Syrups containing high levels of maltose have different properties than those containing high dextrose levels. Maltose has a different sweetness characteristic, and its syrups have a resistance to crystallization. Unlike the production of high-dextrose syrups, it is difficult to achieve maltose levels in excess of 80–85% at high dry substance levels.

There are two main maltose-producing enzymes. The first is a maltogenic α-amylase produced by *Aspergillus oryzae* and is often referred to simply as fungal α-amylase. The second is a β-amylase, which is found in cereals such as barley, wheat, rye, oats, and sorghum, and also in soya beans and sweet potato. Recently a microbial β-amylase has been produced from a *Bacillus* species.

The production of high-maltose-containing syrups using *A. oryzae* (fungal) maltogenic α-amylase follows a similar process as the production of high-dextrose syrups. After dextrinization using a thermostable α-amylase the solution is cooled to 55–60 °C and the pH adjusted to 4.7–5.0. Fungal α-amylase is added to the stirred reactor at a level of between 0.3 and 0.7 L ton^{-1} dry solids. The reaction is allowed to continue for 24 to 40 h.

After saccharification the crude syrup may be filtered and ion exchanged to remove ash. The syrup can then be decolorized by activated charcoal and finally evaporated to the desired solids level. Syrups produced by fungal α-amylase can contain up to 60% maltose based on dry substance. Dextrose levels may be controlled by limiting dextrinization with the thermostable α-amylase. One disadvantage of maltose syrup production using a fungal α-amylase is the high level of maltotriose obtained.

Maltose syrups produced by cereal β-amylases have lower levels of maltotriose than those made with fungal α-amylases. After dextrinization, the starch solution is cooled to 54–58 °C and the pH lowered to 5.2–5.6. Cereal β-amylase is added between 0.25 and 0.75 L ton^{-1} dry solids and the reaction is allowed to continue for 36–48 h, depending on enzyme dose rate. The level of maltose produced by cereal β-amylases is limited to between 50 and 55% on dry substance.

β-Amylases produce maltose sequentially from the nonreducing ends of the oligosaccharides formed by the action of the thermostable α-amylase. Once the α-1,6 branch points are reached, the enzyme is unable to continue. The major products of the reaction are limit dextrins and maltose itself. Table 6 shows a typical carbohydrate spectrum of maltose syrups produced using fungal α-amylase and cereal β-amylase.

Syrups containing between 50 and 60% of maltose on a dry solids basis are referred to as high-maltose syrups.

Table 6. Carbohydrate Spectrum of Maltose Syrups Produced Using Fungal α-Amylase and Cereal β-Amylase

	Fungal α-amylase	Cereal β-amylase
Glucose	3	1
Maltose	50	60
Maltotriose	26	8
Higher dextrins	21	31

Higher levels of maltose may be achieved by addition of a debranching enzyme (pullulanase). The pullulanase hydrolyzes the α-1,6 glucosidic linkages in branched dextrins and amylopectin and allows the β-amylase to continue producing maltose. With careful control of the process, up to 75% maltose may be achieved. Such syrups are called very-high-maltose syrups (VHMS).

Careful control of dextrinization is imperative to produce maltose in excess of 70%. After liquefaction, dextrinization is controlled to give a DE of between 7 and 10, the solution is cooled to 54–58 °C, and the pH adjusted to 5.2–5.6. Cereal β-amylase is added at 0.5 and 0.75 L ton^{-1} dry solids and the pullulanase is added in a similar amount. The reaction is allowed to continue for 36–58 h and the product refined and evaporated. Maltose-producing enzymes do not catalyze the reverse reaction, so there is no reduction in maltose level if the reaction is allowed to proceed.

Maltotriose is a barrier to higher levels of maltose being achieved. β-Amylases are unable to hydrolyze maltotriose, so it is important to keep the level of this oligosaccharide as low as possible. One of the ways this can be achieved is by controlling the DE to a low level during dextrinization.

An alternative is to use an enzyme that has some action toward maltotriose, unfortunately, these enzymes, in addition to producing maltose, also produce glucose. It is often a requirement of maltose syrups that the glucose level be kept low ($<$3.0% on dry substance).

Filtration Enzymes

Maize starch hydrolyzates generally exhibit good filtration characteristics. High filtration rates make it easy to remove undesirable, insoluble components from the hydrolysates in a relatively short time. However, wheat starch hydrolysates have very poor filtration characteristics due to low filtration rates and obstruction of the filter cake. To overcome this problem so-called filtration enzymes are commonly used in the industry.

The major cause of the poor filtration has been identified as a monoacyl lipid compound that is thought to occur in inclusion complexes with the amylose fraction. In wheat the native starch contains about 0.8% of this internal lipid, the principal component of which is lysophosphatidylcholine (lysolecithin). Amylose molecules are helical, with a hydrophobic interior to the helix, which can accommodate certain lipophilic molecules, such as lysophosphatidylcholine.

On hydrolysis of the amylose fraction of wheat starch, the lysophosphatidylcholine is released into the solution. Lysophospholipids are water soluble and excellent emulsifiers. It has been reported that the lipid forms micelles

above a concentration of 0.025 g kg^{-1}. When the critical micelle concentration is exceeded, the filtration rate of the hydrolysate is decreased (11). The critical micelle concentration is 40 times lower than the amount of lysophosphalipids in wheat starch hydrolysates. Therefore, the concentration of lysophosphalipids needs to be reduced to less than 2.5% of the original amount in wheat starch hydrolysates to achieve acceptable filtration rates.

The reduction in lysophosphalipids can be achieved by the addition to the hydrolysate of the enzyme lysophospholipase, which cleaves the remaining fatty acid to give a free fatty acid and a glycerophosphatide. This cleavage destroys the emulsifying ability of the phospholipid. The molecule no longer contains a hydrophobic part and a hydrophilic part, but instead a water-insoluble molecule (the fatty acid) and a water-soluble molecule (the glycerophosphatide) are formed.

Lysophospholipase preparations are produced by the fermentation of *A. niger*. The enzymes are very low in amyloglucosidase activity and are therefore suitable for use in the production of maltose syrups. Manufacturers of glucose from wheat starch should choose an amyloglucosidase enzyme preparation that has an appreciable amount of lysophospholipase as a side activity. There should be no need to add further quantities of lysophospholipase to aid filtration of glucose syrups produced using amyloglucosidase.

A second cause of poor filtration in wheat starch hydrolysates has been identified as the gummy substances from wheat starch, the arabinoxylans, often referred to as pentosans. Pentosans are important macromolecules of wheat flour because of their ability to bind water. Because of the high water-binding capacity, the soluble pentosans form highly viscous solutions with intrinsic viscosities that are about 15–20 times higher than those of soluble proteins. The addition of an enzyme with the ability to break down the water-soluble pentosans will reduce the viscosity of a wheat starch hydrolysate, improving filtration performance.

A second class of pentosans can be separated by means of centrifugation of a flour–water suspension in the so-called tailings, which form a layer above the primary starch. These are the water insoluble pentosans of wheat flour.

The structure of the water-insoluble pentosans is similar to that of the soluble pentosans, but unlike the soluble pentosans, on hydration the water-insoluble pentosans swell without going into solution. On contact with the filter bed these water-insoluble pentosans impede flow through the filter. A second type of pentosan-degrading enzyme can be added to degrade the insoluble pentosan fraction and improve the flow through the filter. The preferred pentosan-degrading enzyme for wheat starch hydrolysate combines action toward both soluble and insoluble pentosans.

Isomerization

The worldwide demand for fructose syrups has increased in recent years due to the inclusion of enriched fructose corn syrup (EFCS) containing 55% fructose in popular soft drinks. Fructose is a naturally occurring sugar that is approximately 1.5 times as sweet as sucrose on an equal weight basis.

Although fructose can be produced by the process of inversion of sucrose followed by separation of the fructose and glucose by chromatography, it is also produced by the enzymatic hydrolysis of starch. The production of fructose from starch initially involves the production of glucose, as already outlined.

Glucose is converted into fructose by a process called isomerization. The enzyme used to catalyze this conversion is glucose isomerase and is usually used in an immobilized form suitable for a fixed-bed continuous column process. The enzyme-catalyzed reaction is reversible, and under the process conditions the equilibrium constant is approximately 1.

Glucose solution from the saccharification is refined using filtration, carbon treatment, and ion exchange and is evaporated to 40–50% dry matter. Calcium ions inhibit the activity of the enzyme, and their levels should be kept as low as possible. The effect of residual calcium ions can be counteracted by the presence of magnesium. Magnesium sulfate is added to give a magnesium ion concentration above 50 ppm and 20 times greater than the calcium ion concentration.

The pH is adjusted to 8.1–8.3 and the temperature to between 55 and 58 °C. The solution is then pumped through a fixed-bed column containing immobilized glucose isomerase. The flow rate through the column should be adjusted in proportion to the enzyme activity in such a way that the fructose content of the product remains constant. The enzyme activity in the column decreases during use and the flow rate must be increased over the lifetime of the column to maintain the fructose level. The typical carbohydrate profile for the products of isomerization is given in Table 7.

The product from the isomerization columns may be refined and concentrated to approximately 70% solids to give a high-fructose syrup at 42% fructose based on dry solids. The 55% fructose syrup can be achieved by concentrating this 42% high-fructose syrup to 90% fructose by a separation process, such as chromatographic adsorption. This concentrated 90% fructose syrup can then be blended with 42% high-fructose syrup to produce the required 55% enriched fructose syrup. A scheme for the production of 55% fructose is shown in Figure 7.

The stability of the enzyme and its ability to retain its activity under process conditions are essential factors in the productivity of the immobilized enzyme. The operating conditions given in Table 8 illustrate the feedstock parameters that have an influence on enzyme performance and therefore productivity.

Table 8 is not an exhaustive list of parameters that effect enzyme performance, but it gives an indication of the conditions that will give optimal enzyme performance.

Table 7. Carbohydrate Profile for Products of Isomerization

Type of carbohydrate	% w/w on dry solids
Fructose	42
Glucose	50–55
Others	3–8

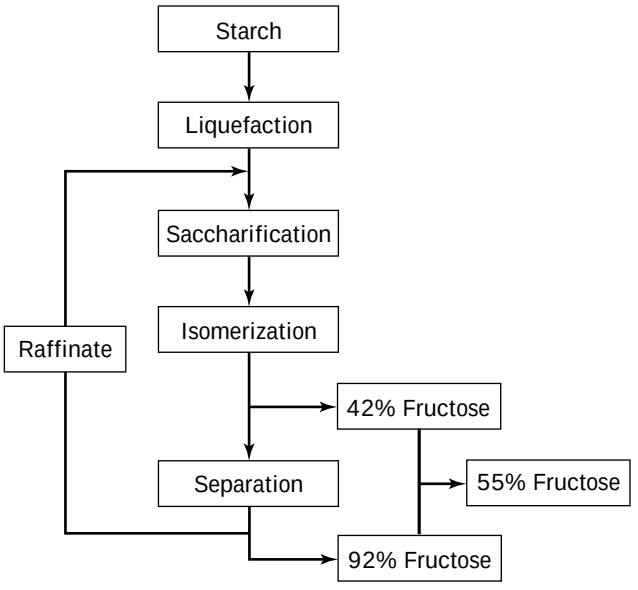


Figure 7. The production of 55% fructose syrup.

Table 8. Conditions Required by Certain Parameters to Promote Optimal Enzyme Performance

Parameter	Range
Dry Substance (%)	40–50
Glucose (% dsb)[a]	≥94
Fructose (% dsb)[a]	≤3
Temperature (°C)	55–58
pH	7.8–8.3
Magnesium (ppm)	≥50
Sulfite (ppm)	≥125

Secondary Product Processing

The processing of the secondary products of the starch industry can be improved with the addition of enzymes. The major use of enzymes in this area is in wheat starch processing.

The yield of quality starch from wheat that is suitable for processing into maltodextrins and sweeteners is lower than the total starch content of the wheat itself. Between 10 and 15% of the total starch of wheat is not separated along with the large particle starch (A starch) during the separation process. This starch, depending on the separation process, is called the B starch or C starch fraction and is usually sold as a low-value animal feed ingredient.

The B or C starch fraction contains high levels of protein and gums (pentosans) that make this fraction unsuitable for processing into sweeteners. These high levels of protein and gums lead to problems with color formation and evaporation of the syrup.

As with other starches, B starch can be liquefied and dextrinized using a heat-stable α-amylase. A saccharification enzyme, such as amyloglucosidase, can be used to produce sugars. The value of the B starch in animal feed is often related to the DE of the syrup. The conditions of liquefaction, dextrinization, and saccharification are no different from those used in the production of nutritive sweeteners.

To aid evaporation, a pentosanase enzyme is also added during saccharification to reduce the viscosity of the syrup. The level of protein can be reduced by introducing a heating step that allows the protein to coagulate. This protein can be dried and sold as a further product in animal feedstuffs.

The enzyme used in the reduction of viscosity can be heat stable, such as that produced by *Penicillium emersonii*, which allows a continued action in the evaporators. Alternatively, there may be no need for a heat-stable enzyme if the saccharification time is long enough to allow the viscosity to be reduced sufficiently.

Using enzymes to reduce the viscosity of B starch syrups has the following benefits.

1. Higher solids content of the final evaporated syrup
2. Improved heat transfer during evaporation and less energy consumption
3. Less fouling of the evaporators, leading to more efficient cleaning
4. Higher evaporation capacity

Hemicellulases (xylanases) may also be used to aid the filtration of maize gluten. The enzyme reduces the viscosity of the gluten suspension and thereby improves the rate of filtration.

CONCLUSION

The development of the enzyme-catalyzed conversion of starch has led to great advancements in the variety and quality of starch-derived sweeteners. The production and use of enzymes is still a relatively new technology, and with the help of modern biotechnology further improvements are not far in the future.

BIBLIOGRAPHY

1. Committee on Food Institute of Medicine ed., *The Food Chemicals Codex*, 4th ed., National Academy Press 1997.
2. E.R. Kooi and F.C. Armbruster vol. 2, in R.L. Whistler ed., *Starch: Chemistry and Technology*, Academic Press, New York, 1984, pp. 553–568.
3. J.D. Steinke and L.A. Johnson, *Cereal Chem.* **68**, 7–12 (1991).
4. D.S. Jackson and D.L. Shandera, Jr., *Adv. Food Nutr. Res.*, **38**, 271–300 (19xx).
5. J. Steinke and L. Johnson, *Cereal Chem.* **68**, 7–17 (1991).
6. A. Caransa, M. Simell, A. Lehmussaari, M. Vaara, and T. Vaara, *Starch* **40**, 409–411 (1988).
7. P.J. Brumm, R.E. Hebeda, and W.M. Teague, *Starch* **41**, 343–348 (1989).
8. P.J. Brumm, R.E. Hebeda, and W.M. Teague, *Food Biotechnol.* **2**, 67–80 (1988).
9. M.-H. David, H. Gunther, and H. Röper, *Starch* **39**, 436–440 (1987).
10. R.E. Hebeda, C.R. Strylund, and W.M. Teague, *Starch* **40**, 33–36 (1988).
11. A.M. Matser and P.A.M. Steeneken, *Detmold Starch Conference*, Germany. 1997.

ERYTHROPOIETIN

SEIJI MASUDA
RYUZO SASAKI
Kyoto University
Kyoto, Japan

KEY WORDS

Carbohydrate structure
Erythropoietin
Monoclonal antibody
Recombinant protein

OUTLINE

INTRODUCTION

Erythropoietin (EPO) stimulates red blood cell production by promoting both the growth of late erythroid precursor cells and their maturation into proerythroblasts, in which globin synthesis starts. It is widely accepted that EPO is a major physiological regulator of the process of erythroid differentiation and that it controls the maintenance of physiological levels of the circulating erythrocyte mass. EPO is a heavily glycosylated protein and produced mainly by the kidney in human adults. Anemia associated with renal failure often results from a decreased level of EPO, and EPO administration results in a dramatic improvement with few side reactions in these anemic patients. Carbohydrates attached to EPO play a critical role in expression of the in vivo activity. In this paper, we mainly focus on the production of recombinant human EPO by mammalian cells and structure and function of carbohydrate chains, comparing with urinary EPO and recombinant product. Basic aspects of EPO, including a novel function of EPO in brain, are also described.

BIOCHEMISTRY AND PHYSIOLOGY OF EPO

It has taken a long time to isolate and characterize EPO. More than 100 years ago, it had been suggested that hypoxia stimulated red blood cell formation. In 1977, human EPO was purified from patients with aplastic anemia (1). The human gene for EPO was cloned in 1985 (2,3), and recombinant EPO (rEPO) is now available. One unit (U) of EPO activity was defined as an erythropoiesis-stimulating activity elicited when 5 μmol of $CoCl_2$ are injected into a starved Sprague Dawler male rat (4). Cobaltous ion promotes erythropoiesis through stimulation of EPO production, because this ion mimics hypoxia in terms of EPO production. The latest standard human rEPO gives 130,000 U/mg protein when it is fully glycosylated. The values of the serum concentration in healthy adult humane range 10 to 30 mU/mL. In 1989, EPO receptor was cloned from erythroleukemia cell line (5), and the signal transduction pathway of EPO has been analyzed in recent decades.

Chemistry

Human EPO gene has been mapped on chromosome 7q11-q22. The mouse EPO gene is localized on chromosome 5. There is a single copy with a length of 4 kb that covers a minimum promoter, five exons, four introns, and hypoxia-inducible enhancer (6,7). The promoter, first intron, and hypoxia-inducible enhancer are highly homologous between human and mouse EPO. The amino acid sequences of EPO in human, monkey, mouse, and rat are presented in Figure 1. Deduction from cDNA indicated 193 amino acid residues in human EPO. The mature EPO isolated from human urine consists of 165 residues and is formed by removal of N-terminal 27 signal peptide and C-terminal Arg residue through the post-translational processing. Two disulfide bond are formed between positions 7 and 161 and 29 and 33. The disulfide linkage between 7 and 161 is essential for EPO bioactivity. Because there is 80% protein homology between mouse and human EPO, human EPO crossbinds with mouse EPO receptor and vice versa. EPO protein has a significant homology to mature thrombopoietin, which is a primary regulator for megakaryopoiesis (8).

A tertiary conformation for EPO has been predicted based on the protein structure modeling. Despite little homology in the amino acid sequence, EPO has a similarity to the three-dimensional structures of growth hormone, prolactin, interleukin-6, and granulocyte-colony stimulating factor (9). By analogy with growth hormone that has been shown to contain four antiparallel α-helices (10), EPO is also thought to contain four antiparallel α-helices and three loops, which is supported by the analysis of site-directed mutation within EPO molecule. Mutations of buried portion of the helix A (first helix from the N-terminal

Figure 1. The amino acid sequences of human, monkey, mouse, and rat EPO. Numbering starts from N-terminus of mature human EPO. N-linked glycosylation sites are shown by *asterisks*, and an O-linked glycosylation site is indicated by *arrows*. The disulfide bonds are drawn as bridges.

of EPO) and D (last helix) show little effect in vitro activity, whereas those of surface portion of the helix A and D reduced the activity (11). The mutational analysis suggests that EPO has two separate receptor binding sites, which is consistent with the ligand-induced receptor homodimerization, an initial step for the signal transmission (12).

EPO Biosynthesis

As far as kidney cells responsible for EPO production are normal, the relationship between concentrations of EPO and hemoglobin in blood is normally inverse. A decrease in the hemoglobin content, for instance from blood loss, is detected by the EPO-producing cells in kidney as a reduction of oxygen supply (Figure 2). Low oxygen stimulates transcription of EPO gene, and the resulting increase in blood EPO promotes the proliferation and differentiation of colony-forming unit erythroids (CFU-Es), erythroid precursor cells on which EPO acts, into matured red blood cells. Recovery of the red blood cell mass and the concomitant oxygen supply suppresses EPO production. The molecular mechanism of hypoxic activation of EPO gene transcription has been extensively studied, although the molecule that senses changes of oxygen concentration is poorly understood (13). An important *cis*-acting sequence required for the hypoxic induction of Epo gene was defined in the 3′-flanking region. This enhancer is a 50-bp element consisting of three important segments. The highly conserved sequence near the 5′ end of the enhancer is a hypoxia-inducible factor-1 (HIF-1) binding site. HIF-1 is a transcription factor consisting of two basic helix-loop-helix-PAS proteins (HIF-1α and HIF-1β). The middle segment containing CA repeats in human gene is not well conserved between human and mouse, but the mutation of this region abolished the inducible behavior of both the human and murine enhancers. Thus far the specific protein that binds to this region has not been demonstrated. The third element is the 3′ segment, which is again highly conserved and amplifies the hypoxic signal. Binding of hepatocyte nuclear factor-4 (HNF-4), which belongs a member of the steroid hormone receptor family, to this element appears to increase the hypoxic inducibility.

EPO Production Sites

The kidney is the primary site of EPO production in adults, whereas the liver is the production site in the fetal stage. Hypoxia activates expression of the EPO gene in kidney

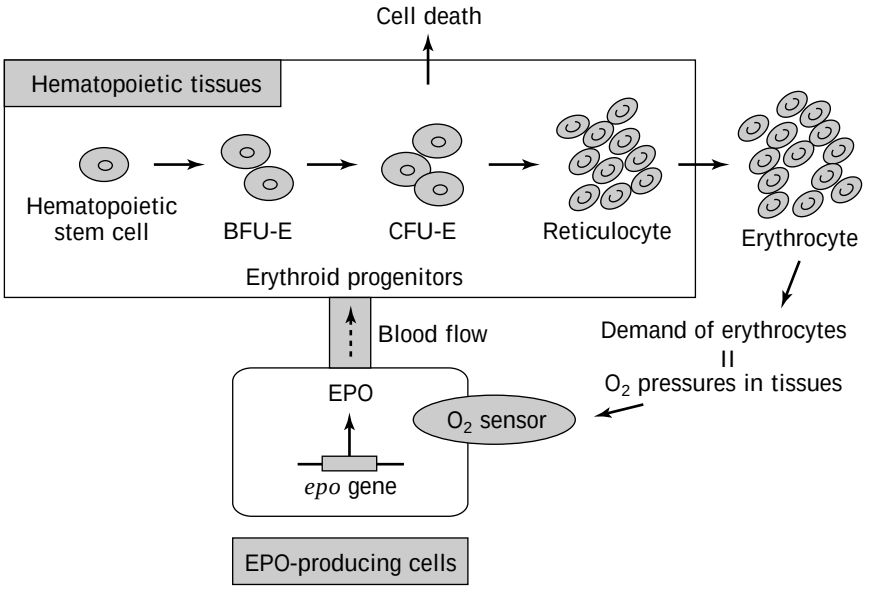


Figure 2. Regulation of erythropoiesis by EPO.

and liver (6,7,12–15). The-EPO producing cells in the kidney have been controversial but recent studies indicate that EPO is produced by peritubular interstitial fibroblasts. The hepatic EPO production is essential for fetal erythropoiesis, because homozygous mice carrying null mutations of EPO gene die around embryonic day 13 due to deficiency in erythropoiesis (16). Although EPO production in liver is hypoxia inducible, the failure of renal EPO production cannot be substituted by the liver, because uremic patients with severe anemia show low levels of serum EPO. EPO has been believed to act exclusively on erythroid precursor cells, but several recent findings indicate that EPO acts on neurons in the central nervous system (CNS) (17–19). Astrocytes produce brain EPO, and low oxygen tension stimulates the production of EPO as it does in liver and kidney. EPO in CNS functions as a neurotrophic factor as described in the next section.

EPO Action

Hematopoiesis is governed by several hematopoietic growth factors. EPO is a glycoprotein that plays a essential role in the regulation of erythropoiesis by stimulating the proliferation and differentiation of erythroid precursor cells (6,7,12–15). All the terminally differentiated blood cells have been derived from pluripotent stem cells with self-renewing activity. Hematopoiesis occurs in bone marrow, and the spleen is also an active hematopoietic site in murine. Pluripotent stem cells are committed to differentiate to multipotent stem cells with differentiation activity to several lineage-specific progenitors that have lost the self-renewality. Each lineage-specific progenitors differentiate to mature blood cells such as erythrocytes, monocytes, granulocytes, and megakaryocytes. The most primitive erythroid-specific progenitor cells derived from multipotent stem cells are burst-forming unit-erythroids (BFU-Es). Although BFU-Es express EPO receptor, their response to EPO is low, and they require additional cytokines such as interleukin 3 or granulocyte-macrophage–colony-stimulating factor for their proliferation and differentiation into CFU-Es that show the highest sensitivity to EPO. Homozygous mice carrying null mutation in EPO receptor gene die around embryonic day 13, because mature erythrocytes are not formed, although CFU-Es abundantly exist (16). Thus, EPO is essential for differentiation and proliferation of CFU-E. Recently, a novel function in the CNS has been proposed. Neuronal cell lines such as PC12 and SN6 express EPO receptor, and binding of EPO to PC12 cells increases the intracellular concentration of monoamines (20). EPO protects primary cultured hippocampal and cerebral cortical neurons from NMDA receptor-mediated glutamate toxicity (19), which is believed to be a major cause of neuron death by ischemia. These findings suggest that EPO plays a neuroprotective role when brain is exposed to hypoxia or ischemia.

PRODUCTION AND PURIFICATION OF RECOMBINANT EPO

In some cases, carbohydrates attached to the proteins play a key role in expression of their biological activities. The carbohydrates may be required for the in vitro activity as well as the in vivo activity. Even when the carbohydrates have nothing to do with the in vitro activity, they may be very important in vivo; carbohydrates protect glycoproteins from heat inactivation and proteolytic degradation, resulting in an increased lifetime in circulating blood. Eukaryotic cells are used for production of recombinant glycoproteins. For clinical applications, not only glycosylation but also the structure of attached carbohydrates are critical; a glycoprotein whose terminal carbohydrate residues are recognized by hepatic receptors would be rapidly eliminated from the circulation. The kind of carbohydrate structure may also affect the antigenicity of the recombinant product. When recombinant glycoproteins are produced in heterologous animal cells, their amino acid sequences govern the positions onto which carbohydrate

chains are attached, but the maturation of the chains may entirely depend on the host cells. It is important, therefore, to compare the properties of recombinant glycoproteins produced in different cell lines with those of naturally occurring ones. The glycoprotein may also be dependent on the culture conditions.

Isolation and Expression of EPO Gene

Peptides derived from urinary EPO (u-EPO) (21) were sequenced to prepare synthetic DNA probes for screening a human fetal liver genomic library for EPO gene clones. Restriction endonuclease mapping and Southern analysis showed the gene fragments contained in three phage clones to be identical. One, designated λEPO41, was used for sequence analysis. The entire amino acid sequence of EPO was determined. The clone λEPO41 contained a whole EPO coding region, although half of 3′-untranslated region was missing. A vector for the expression of EPO in animal cells was constructed by the procedures diagrammed in Figure 3. The EPO expression vector, pZIP-NeoSV(X) 1-EPO, contained the EPO gene preceded by a long terminal repeat (LTR) functioning as transcriptional promoter and followed by DNA sequences derived from the transposon Tn5 (Neo fragment in Fig. 3) that confers G418 resistance upon mammalian cells. The expression vector was introduced into two mammalian cells, BHK cells established from baby hamster kidney, and ψ2 cells derived from NIH/3T3. Both cells were transfected with pZIP-NeoSV(X) 1-EPO, using calcium phosphate method. EPO-producing cells were isolated by selection in G418 media and by two rounds of limiting dilution. The EPO levels in culture supernatants of EPO-producing cells assayed by radioimmunoassay (RIA) were 150 U/mL for BHK cells and 300 U/mL for ψ 2 cells.

Cell Culture and EPO Production

BHK cells were cultured in Eagle basal medium supplemented with 10% calf serum and 10% bacto tryptose phosphate broth. The ψ 2 cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% calf serum. Cell cultures for EPO production were started at a cell density of 2×10^5 cells/mL in plastic culture trays (Cell Factory, 6,000 cm^2, NUNC) containing 2 L of the medium. After 4 days of culture, the cells grew to be nearly confluent. The culture medium was replaced with the low-serum medium (1.5% calf serum) and culture supernatants were harvested every 3 days.

Purification of Human rEPO

Human urinary EPO directed MAb (21) was effective in the purification of EPO from human urine and culture supernatants of BHK cells that produced rEPO (22), but pretreatment of the EPO with SDS was needed for EPO to bind with the immobilized MAb. To develop a more rapid purification procedure, the new MAbs, R2, (see next section) were used for rEPO purification (23).

The culture supernatant (100 L) was filtered with suction and concentrated to 2 L by ultrafiltration on a hollow-fiber device (Amicon DC-10, Danvers, Mass.) with a nominal molecular cutoff of 10,000. The concentrate was put on an immunosorbent column (4.4 × 6.6 cm) containing EPO-directed MAb R2 fixed on Affi-Gel 10 equilibrated with phosphate buffered saline (PBS). The column was extensively washed with PBS, 10 mM NaPi, pH 7.4, containing 0.5 M NaCl and 0.15 M, in this order. EPO was eluted by 0.2 M acetate. pH 2.5, containing 0.15 M NaCl. After neutralization of the eluted fraction, the protein was concentrated by ultrafiltration with a hollow-fiber device (Mini-Module NM-3. Asahikasei. Tokyo) and was precipitated with 90% ethanol. The precipitate was dissolved in a small amount of PBS and put on a Sephadex G-100 column (2.2 × 94 cm) equilibrated with PBS. The EPO in the fractions was concentrated and precipitated with 90% ethanol. The precipitate was dissolved in 10 mM NaPi, pH 6.8, containing 10 μM $CaCl_2$ and put on a hydroxyapatite column (100 × 7.8 mm: Bio-Gel HPHT, Bio-Rad) equilibrated with the same buffer. The adsorbed EPOs were eluted by increasing concentration of NaPi. The rEPO-B appeared in the flowthrough of a column equilibrated with 10 mM NaPi, whereas rEPO-ψ was eluted by 50 mM NaPi.

The large-scale purification of rEPO-B in the supernatant of BHK cells was attempted. Purification with the immunosorbent column was effective (Table 1); most of the protein and little EPO activity in the concentrate of culture supernatants emerged in the flowthrough fractions without being adsorbed, and EPO was eluted sharply at pH 2.5. About 2,800-fold purification was achieved by this single step, and 84% of the activity was recovered. Mab R2 was suitable for purification of EPO with a high yield. Subsequent chromatography with Sephadex G-100 and then hydroxyapatite yielded purified EPO with 3,260-fold purification and 52% recovery of the starting activity. The purified EPO had the specific activity of 137,000 $U/A_{280\,nm}$, which was higher than that of the EPO purified with an earlier preparation of an MAb (21). The final EPO preparation was pure on SDS-polyacrylamide gel electrophoresis and high-pressure liquid chromatography. The immobilized R2 could be used 30 times without any deterioration.

Figure 4 shows the electrophoretic patterns of rEPOs and uEPO on SDS-polyacrylamide gels. Purified uEPO showed two bands (lane 1). The minor one with the lower molecular weight is the asialylated form of uEPO (21). Both rEPOs migrated as a single band, but they differed in electrophoretic mobility (lanes 2 and 3). Recombinant EPO from ψ 2 cells migrated slightly faster than uEPO, and the rEPO-B band was broader than those of rEPO-ψ and uEPO. Thirty amino acids in the N-terminal portion of rEPO-ψ and rEPO-B were consistent with those of uEPO and also with those deduced from the cDNA sequence (2,3). The amino acid sequence in the C-terminal region of both rEPOs and uEPO determined by the carboxypeptidase method was identical to that of uEPO, indicating that the recombinant products had the same primary structure as uEPO. Therefore, differences between purified rEPOs in their electrophoretic mobility on SDS-polyacrylamide gel, therefore, are attributable to the difference in the carbohydrate chains.

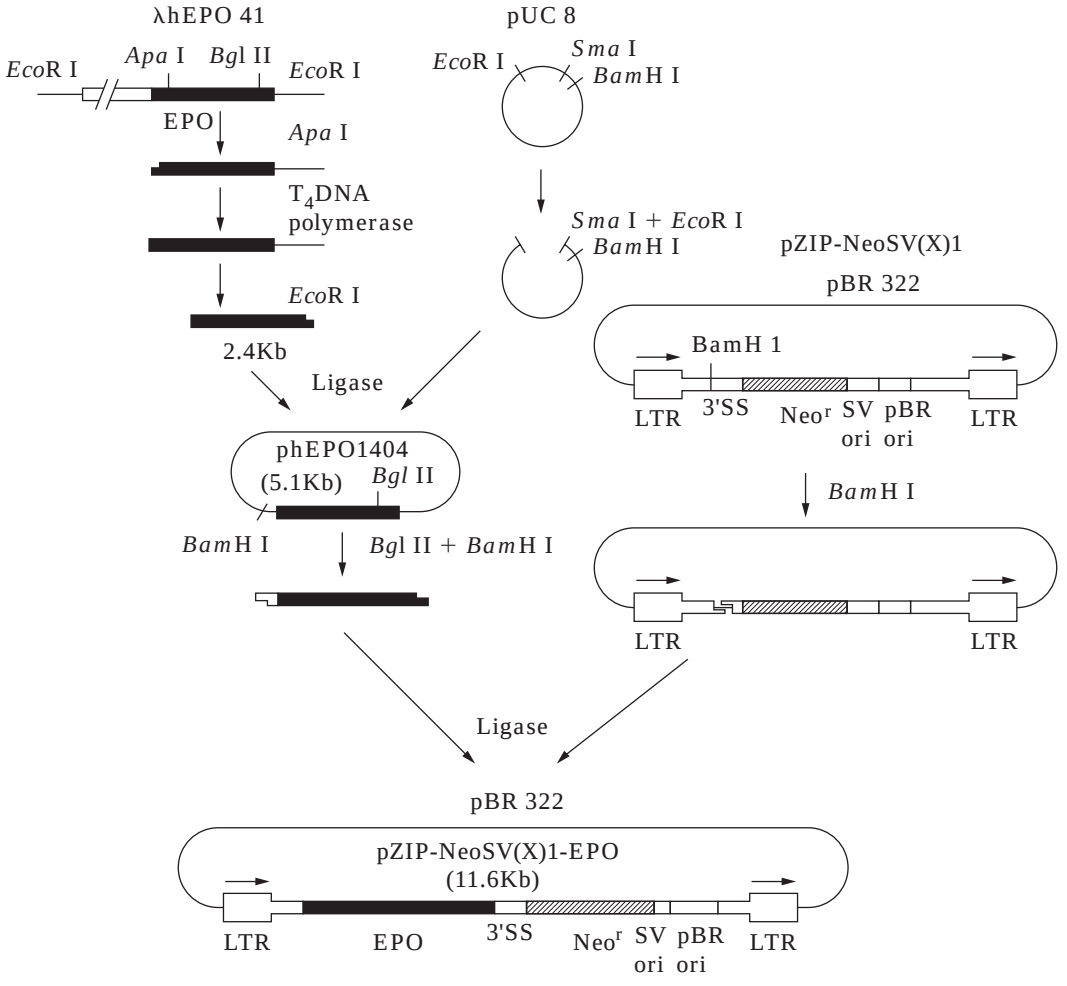


Figure 3. Construction of EPO expression plasmid. The *Apa*l site of the EPO gene in λhEPO41 is a unique site located at 62 bp upstream from the translation initiation codon, and the *Eco*RI site is for inserting human DNA fragments into the bacteriophage cloning vector, Charon 4A. The *Bgl*II site of the EPO gene in phEPO1404 is 185 bp downstream from the translation stop codon. The details in pZIP-NeoSV(X)1 have been described elsewhere (22). LTR = long terminal repeat of Molony murine leukemia virus; 3′SS = 3′ splicing signal; Neo^r = sequence derived from the transposon Tn5; SV ori = sequence of SV40 origin; pBR ori = sequence of pBR322 origin. *Source*: Goto et al. (22).

Table 1. Purification of Recombinant Human EPO from BHK Cells

Procedures	Protein ($A_{280\ nm}$)	Activity ($U \times 10^{-4}$)	Specific activity ($U/A_{280\ nm}$)	Yield (%)	Purification (fold)
Concentrated culture supernatant	473,000	1,990	42	100	1
Immunosorbent column	141	1,680	119,000	84	2,830
Sephadex G-100	97	1,310	135,000	66	3,210
Hydroxyapatite	75	1,030	137,000	52	3,260

Source: Goto et al. (23).

Biological Activity of Isolated EPOs

The Second International Reference Preparation for human EPO obtained from the WHO International Laboratory for Biological Standards, National Institute for Medical Research, London (24), was used as the standard. The in vivo and in vitro activities of three purified EPOs were assayed (Table 2). The stimulatory effect of EPO on the incorporation of ^{3}H-thymidine into DNA of cultured mouse fetal liver cells was used to assay the in vitro activity, and the in vivo activity was assayed with the use of starved rats (4). The stimulatory effect of the EPO preparations were completely repressed by incubation with antiserum against EPO, indicating that the mitogenic effect of a con-

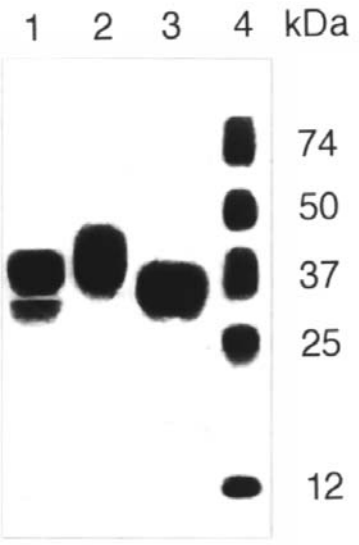


Figure 4. Analysis of human urinary and recombinant EPOs with SDS-polyacrylamide gel electrophoresis. Lanes 1, 2, and 3 are EPOs from urine, the culture supernatant from BHK, and ψ2 cells, respectively. Lane 4 shows the molecular standards. *Source*: Goto et al. (22).

Table 2. Biological Activities of Human EPO

		EPO activity (U/A_{280})	
EPO	Treatment	In vitro	In vivo
u-EPO	Untreated	77,000	72,000
	Sialidase treated	99,000	0
r-EPO-B	Untreated	96,000	72,000
	Sialidase treated	109,000	0
r-EPO-ψ	Untreated	153,000	18,000
	Sialidase treated	173,000	0

Source: Goto et al. (22).

taminant did not contribute to the in vitro activity. The in vivo activity of rEPO-B was equivalent to that of uEPO, and its in vitro activity was slightly higher than that of uEPO. The in vitro activity of rEPO-ψ was 1.6-fold that of uEPO and of rEPO-B, whereas, the in vivo activity was only 25% that of other EPOs. The in vitro activities of the three EPO preparations slightly increased after sialidase treatment, but their in vivo activities were abolished. It seems to be that the lower the amount of sialic acid, the higher the in vitro activity. The high activity of the native rEPO-ψ as compared with other native EPOs may be due to the low content of sialic acid, because the asialo EPOs have an increased affinity to EPO receptor (25,26). However, the low in vivo activity of rEPO-ψ is probably due to the small amount of sialic acid, because the removal of sialic acid from EPO by sialidase treatment abolished in vivo activity. Loss of in vivo activity is caused by the removal of asialoglycoproteins from the circulation by the liver (20). Thus, it is likely that the in vivo lifetime of rEPO-B is longer than that of rEPO-ψ.

ISOLATION OF EPO-DIRECTED MONOCLONAL ANTIBODY AND APPLICATION TO ELISA

We reported an EPO-directed monoclonal antibody (MAb) that binds to the antigen after it is treated with SDS (21). It appears that this antibody binds to the epitope exposed by a conformation change of EPO caused by the SDS. This antibody worked efficiently in the purification of EPO from human urine and human recombinant EPO from the supernatant of cultured cells, but treatment of preparations containing EPO with SDS is time consuming. We have prepared MAbs capable of binding with the native EPO and thereby to achieve the following purposes: (*1*) establishment of a more rapid procedure for the purification of EPO by immobilized MAb; (*2*) development of a sandwich-type ELISA, which is easier to handle and more rapid and sensitive than radioimmunoassay (RIA), by use of MAbs that bind to EPO at different antigenic determinants.

Isolation of EPO-Directed MAb-Producing Hybridomas

Female BALB-c mice 8 weeks old were immunized intraperitoneally with 90 μg of pure rEPO emulsified 1:1 in Freund's complete adjuvant. After three times of immunization, hybridomas were prepared by procedures described previously (21). Hybridoma cells were maintained in RPMI 1640 supplemented with 10% fetal calf serum. Two hybridoma clones (R2 and R6) that were stable in their rapid growth and their production of MAbs to EPO were isolated. Both antibodies were produced in the ascitic fluid of BALB/c mice and purified by the ammonium sulfate precipitation and diethylaminoethyl (DEAE)-cellulose column chromatography. The subclass of both antibodies was IgG1. The apparent dissociation constant, k_d was calculated by the binding to ^{125}I-labeled EPO. Both MAbs bound to EPO with high affinities, with k_d = 1–2 nM, respectively. The fully deglycosylated rEPO bound to the MAbs, indicating that they recognized peptide sequences of the antigen but not the carbohydrates attached to the antigen. Moreover, each antibody recognizes a different epitope.

Enzyme-Linked Immunosorbent Assay

The finding that the antigen site recognized by MAb R2 differed from the sites recognized by R6 prompted us to develop a sandwich-type ELISA for the measurement of EPO levels. EPO was sandwiched between R6 fixed in microtiter wells and the enzyme-linked R2. R2 was linked covalently with alkarine phosphatase with 0.2% glutaraldehyde. The MAb R6 solution (10 μg/mL) was added to the wells of microtiter plates with 96 flat-bottomed wells and incubated for 2 h. After washing, 100 μL/well of the sample solution containing 0.05% Tween 20 was added to the wells, which were incubated for 2 h to form the fixed R6 and EPO complexes. After washing, 100 μL/well of the alkaline phosphatase-linked R2 solution was added and the wells were incubated for 2 h more to form ternary complexes (fixed R6, EPO, and enzyme-linked R2). The wells were washed, and then 100 μL of 2 mg/mL *p*-nitrophenyl phosphate disodium was added. After incubation for 30 min, the reaction was stopped by the addition of 3 M

NaOH. The optical density at 405 nm was measured with a microtiter reader. This procedure is referred to as a two-step reaction. In the two-step reaction, EPO in samples was bound to the fixed R6 and then alkaline phosphatase-linked R2 was added to the wells to form ternary complexes of R6, EPO, and the enzyme-linked R2. In the alternative procedure, the one-step reaction mixtures of EPO samples and enzyme-linked MAb R2 were added to the wells coated with MAb R6.

The relationship of the amounts of rEPO and the enzyme activity was linear regardless of different procedures (Fig. 5); use of the one-step reaction shortened the time needed for completion of the process to 7 h. Once the microtiter wells in which the first antibody, R6, have been fixed, results can be obtained within 3 h. To check the reproducibility of the ELISA, intra-assay and interassay studies were done. In both studies, the coefficients of variation were satisfactory (<20%). The assay measured plasma levels of EPO as low as 2 mU/mL within several hours.

We have developed RIA for EPO by the use of antisera against human u-EPO and measured EPO levels in plasma samples from normal subjects and anemic or polycythemic patients (27). A sample diluted to about 1 U/mL EPO was mixed with 10,000 cpm of ^{125}I-rEPO, rabbit anti-rEPO antiserum. The mixture was incubated at 4 °C overnight and sheep antirabbit γ-globulin (Cappel) was added. The immunocomplexes were precipitated by centrifugation. The radioactivity in the precipitate was counted with a γ-counter. The amounts of EPO in the sample were estimated from a standard curve (2 to 128 mU/assay).

From total 38 samples, EPO levels of normal subjects and patients were measured with ELISA to compare with RIA. The results are shown in Figure 6 and indicate a high degree of correlation between these two methods, with a correlation coefficient of 0.965. The ELISA was comparable with the RIAs in sensitivity. One can assay EPO for a short time without using any radioactive compounds. Moreover, the assay conditions, once established, can be used semipermanently, because MAbs with the constant properties can be supplied continually. The sandwich-type ELISA developed by the use of MAbs thus had advantages over the RIAs in practicality. The findings that the MAbs R6 and R2 bound to EPO with high affinities and recognized different antigenic determinants made it possible to develop a sensitive and easier-to-handle method for EPO assay. The sandwich-type ELISA should be useful to assay EPO levels in human plasma for investigating an etiology of defective erythropoiesis and also to monitor rEPO given to anemic patients.

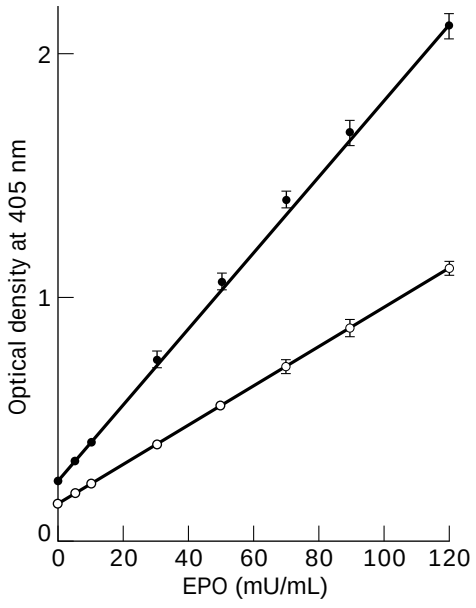


Figure 5. EPO titration by ELISA. ○ indicates one-step reaction; ● indicates two-step reaction. Each point is the mean of triplicate samples. *Source*: Goto et al. (23).

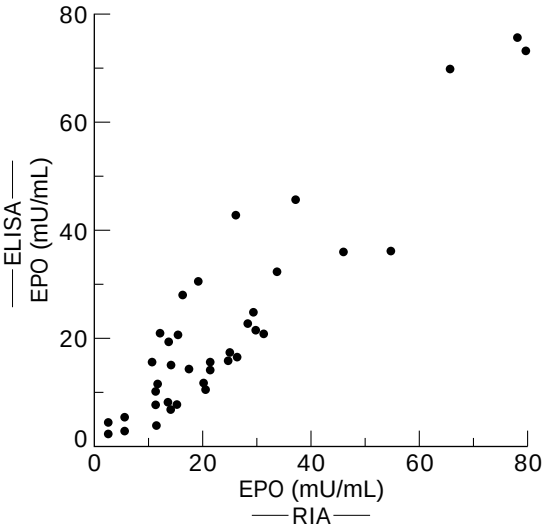


Figure 6. Comparison of RIA and ELISA in human plasma. The ordinate shows EPO in the plasma samples assayed with ELISA by two-step reaction, and the abscissa indicates EPO estimated with RIA by using polyclonal antibody raised in a rabbit. Each point is the mean of triplicate samples. *Source*: Goto et al. (23).

FUNCTION OF CARBOHYDRATES ATTACHED TO EPO IN ERYTHROPOIESIS

The following biological functions can be assigned to carbohydrates attached to proteins such as EPO that act in the endocrine system. First, the carbohydrates of glycoproteins are needed for their biosynthesis and secretion, including folding of the protein during translation and protection of the protein from intracellular degradation. Second, carbohydrates confer stability to the protein so that it is not eliminated from the circulation during its travel from the production site to its target cells. Mammalian cells, particularly liver cells, express proteins (lectins) that bind to carbohydrates in glycoproteins. Glycosylation of proteins thus can be a signal determining their lifetime in the circulation. Third, carbohydrates are involved in the expression of biological activity, including the

steps of the binding of the protein to its target cells and intracellular signal transduction.

To examine these possible functions of carbohydrates, the following strategies can be used. First, there are some mammalian cell lines lacking an enzyme involved in glycosylation. One can use these cells to produce a recombinant glycoprotein and compare with the natural product or the fully glycosylated recombinant product. Second, one may produce the protein in procaryotes incapable of attaching carbohydrates, and if the product has the biological activity, one can get some information of functions of carbohydrates. Third, one can examine the effects of chemical or enzymatic removal of carbohydrate residues from a glycoprotein on its biological properties. Discovery of enzymes that remove carbohydrate residues from a glycoprotein is particularly useful, because the enzymatic treatment is milder by far than the chemical treatment. Fourth, glycosylation sites can be removed by site-directed mutagenesis and the biological properties of the glycosylation-defective mutant glycoproteins are compared with those of wild type.

Table 3. Biological Activities of Recombinant EPO Deglycosylated by Glycosidases

EPO	In vitro activity (%)	Affinity to EPO-R [I_{50} (nM)]	In vivo activity (%)
EPO-	100	1.2	100
EPO-1	252	0.3	0
EPO-2	252	0.3	0
EPO-3	266	0.3	0
EPO-4	81	0.6	0
EPO-5	106	0.3	0

Source: Tsuda et al. (25).

Note: EPO-0-native EPO; EPO-1-sialidase-digested EPO-0; EPO-2-sialidase/*β*-galactosidase-digested; EPO-3-sialidase/*β*-galactosidase/*β*-N-acetylhexosaminidase-digested; EPO-4-*N*-glycanase-digested; EPO-5-fully deglycosylated (sialidase/*endo*-*α*-*N*-acetylgalactosaminidase/*N*-glycanase-digested) (see also Fig. 7). Affinity to EPO receptor (EPO-R) does not represent a dissociation constant but a relative affinity. The lower the value, the higher the affinity.

Enzymatic Removal of Carbohydrates from EPO

Figure 7 shows representative structures of N-linked and O-linked oligosaccharides of recombinant human EPO produced by BHK cells (22,25) and cleavage sites of the enzymes that we used for removal of carbohydrate residues. The native EPO is represented as EPO-0, and deglycosylated forms are as EPO-1∼5 in Table 3. Sialidase removes terminal sialic acids, producing asialo-EPO (EPO-1). *β*-Galactosidase cleaves off the exposed Gal residues of N-linked oligosaccharides in EPO-1 but does not split the b1→3 linkage between Gal and GalNAc in the O-linked oligosaccharide, giving EPO-2. Double digestion of EPO-1 with *β*-galactosidase and *β*-*N*-acetylhexosaminidase removes all the exposed Gal and GlcNAc residues from N-linked oligosaccharides, and the core structure is exposed (EPO-3). N-Glycanase completely eliminates N-linked oligosaccharides from EPO-0, yielding EPO-4. *Endo*-*α*-*N*-acetylgalactosaminidase removes O-linked oligosaccharides from asialo-EPO. Therefore, successive digestion of EPO with sialidase, *endo*-*α*-*N*-acetylgalactosaminidase, and *N*-glycanase yields the fully deglycosylated EPO,

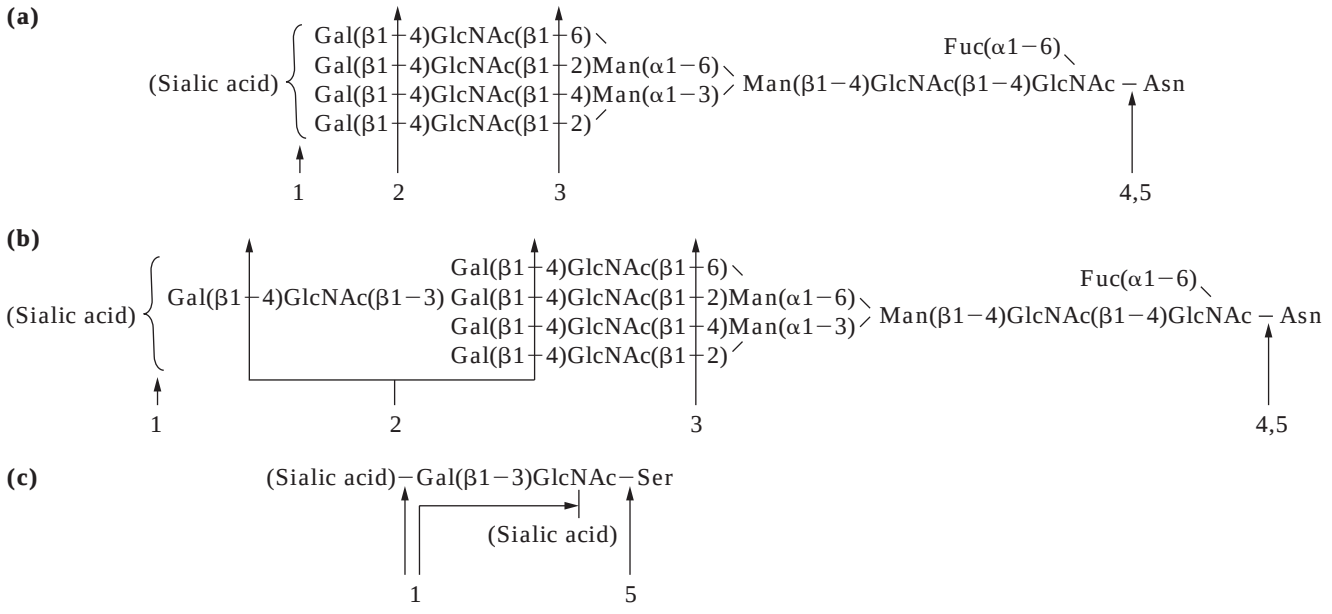


Figure 7. Enzymatic removal of sugars from recombinant human EPO. Cleavage site of sugars in EPO by glycosidases are shown. (**a** and **b**) Structures of representative complex-type N-linked sugars (**c**) O-linked sugar. The lines with arrowheads numbered 1–5 indicate the cleavage sites by glycosidases: 1-sialidase; 2-*β*-galactosidase after sialidase digestion; 3-*β*-galactosidase/*β*-*N*-acetylhexosaminidase after sialidase-digestion; 4-*N*-glycanase; 5-sialidase/*endo*-*α*-*N*-acetylgalactosaminidase/*N*-glycanase. *Source*: Tsuda et al. (25).

EPO-5. Completion of the enzyme reactions were confirmed by SDS-polyacrylamide gel electrophoresis of the products.

Effects of these enzymatic removals on the biological properties are summarized in Table 3. Affinities to EPO receptors do not represent the true dissociation constants but the relative values; the lower value, the higher affinity. Carbohydrates with terminal sialic acids are essential for expression of the in vivo activity. The carbohydrates are not important for expression of the in vitro activity. Deglycosylation increased the in vitro activity by increasing the affinity to EPO receptor. The presence of terminal sialic acids is responsible for an decreased affinity. EPO-4 and 5 have in vitro activity similar to that of the native EPO (EPO-0), although their affinity to EPO receptor is higher than EPO-0. They may be partially inactivated during assay thorough aggregate formation. The removal of terminal sialic acids from EPO has been shown to destroy its in vivo activity (21,22,28–31). Desialylated EPO is cleared from the circulation faster by far than the native EPO (32,33). In contrast, the in vitro activity of EPO increases with desialylation (21,29). All these results are consistent with those in Table 3.

A possible function of carbohydrates in the in vitro activity of EPO was controversial. It has been reported that *N*-glycanase digestion of EPO produced in Chinese hamster ovary (CHO) cells caused almost complete loss of activity (34), which was contradictory to our conclusion that carbohydrates were not important in expression of the in vitro activity. However, all the results published later support our conclusion (26,35–37). We have to be very careful to assay the activity of EPO that has lost N-linked sugars, because the deglycosylated EPO is temperature-sensitive and sticks to experimental vessels (25).

Site-Directed Mutagenesis for Production of Glycosylation-Defective Mutant EPO

We constructed EPO genes mutated by site-directed mutagenesis so that EPOs lacking N-linked carbohydrates in every possible combination of the three N-glycosylation sites were produced (26). Because the codon of Asn (N) at the N-glycosylation sites is replaced with that of Gln (Q) in these mutants, we referred to them as NQ0, NQ1, and so on (Table 4). In this nomenclature, for example, NQ13 means the mutant at positions 24 and 83. NQ0 corresponds to the wild type and NQ123 is the mutant with no N-glycosylation sites. Transient production of mutant EPOs by BHK cells was examined to find roles of sugars in biosynthesis of EPO. The mutant EPOs produced by stably transfected BHK cells were partially purified and their properties, including the biological activity both in vivo and in vitro, were investigated. In some mutants, the in vitro activity was higher than that of wild-type, and the high activity is probably due to an increased affinity to EPO receptor. The in vitro activity of EPO with no glycosylation sites was comparable with that of the wild-type EPO, whereas the in vivo activity severely decreased. The elimination of all three N-glycosylation sites decreased EPO production to 10% of that of the wild-type EPO. These results indicate that N-linked sugars of EPO have two major functions: N-glycosylation is important for (*1*) proper biosynthesis and/or secretion and (*2*) expression of the in vivo activity. N-Linked sugars of EPO affect binding affinity of the ligand to the receptor but do not play a key role in expression of the in vitro activity. It has been reported that replacement of Asn to Gln at position 24 or 38 gives EPOs with the specific activity in vitro of about 20% that of the wild type (38), which is contradictory to our results (26). A possible cause has been discussed (26), and the results in a paper published later (36) are consistent with our data.

Preventing O-linked glycosylation at position 126 by site-directed mutagenesis (38) severely reduced the production into the culture medium, suggesting that O-linked sugar is important for proper biosynthesis and/or secretion. But the result from other laboratories showed the normal production of the mutant (36). The latter was supported by the result obtained using the mutant cells that are deficient in an enzyme involved in the carbohydrate synthesis (39). Involvement of O-linked sugar in expression of the biological activity is unlikely (25,36,38,39). O-linked sugar is present only in primate EPOs but not in rodents (6,7,14).

Others

Interestingly, fractionation of recombinant EPO containing biantennary and tetraantennary N-linked sugars revealed that EPO containing more extensively branched chains has greater in vivo activity (40). This suggests a possible unique role of the N-linked sugar in vivo. Attempt

Table 4. Biological Properties of N-Glycosylation Site-Defective EPO

Mutant EPO	Mutation points (Asn → Gln)	In vitro activity (%)	Affinity to EPO-R [I_{50}(nM)]	In vivo activity (%)	Secreted EPO (%)
NQ0	Wild type	100	0.50	100	100
NQ1	24	180	0.25	129	70
NQ2	38	107	0.28	60	6
NQ3	83	130	0.40	106	78
NQ12	24,38	140	0.15	42	31
NQ13	24,83	170	0.18	63	58
NQ23	38,83	66	0.31	51	41
NQ123	24,38,83	100	0.14	8	10

Source: Yamaguchi et al. (26).
Note: Affinity to EPO receptor (EPO-R) represents a relative value but not a dissociation constant. The lower the value, the higher the affinity.

to produce EPO in *Escherichia coli* has been made (36). They were unable to compare the biological activity of EPO produced in *E. coli* with that of EPO produced in mammalian cells because of the poor stability of *E. coli* EPO, but they confirmed that *E. coli* EPO had the in vitro activity. *Bacillus brevis* cells secrete recombinant human EPO, which has in vitro activity similar to that of the fully glycosylated EPO (42). These results support our conclusion that the N-linked sugars in EPO are not required for expression of the in vitro activity. When human wild-type EPO was expressed in the polarized Madin-Darby canine kidney (MDCK) epithelial cells, EPO was preferentially secreted from the apical domain (43). This polarized secretion was perturbed by the treatment of the cells with tunicamycin, suggesting the involvement of N-linked sugars in the apical sorting mechanism in MDCK cells. N-Linked sugar chain at position 38 is critical for the polarized secretion, but the biological significance of this finding remains unknown.

BIBLIOGRAPHY

1. T. Miyake, C.K.-H. Kung, and E. Goldwasser, *J. Biol. Chem.* **252**, 5558–5564 (1977).
2. K. Jacobs, C. Shoemaker, R. Rudersdorf, S.D. Neill, R.J. Kaufman, A. Mufson, J. Seehra, S.S. Jones, R. Hewick, E.F. Fritsch, M. Kawakita, T. Shimizu, and T. Miyake, *Nature* **313**, 806–810 (1985).
3. F.-K. Lin, S. Suggs, C.-H. Lin, J.K. Browne, R. Smalling, J.C. Egrie, K.K. Chen, G.M. Fox, F. Martin, Z. Stabinsky, S.M. Badrawi, P.-H. Lai, and E. Goldwasser, *Proc. Natl. Acad. Sci. USA* **82**, 7580–7584 (1985).
4. J.F. Garcia, in J.W. Fischer ed., *Kidney Hormones*, vol. 2, Academic Press, London, 1977, p. 7.
5. A.D. D'Andrea, H.F. Lodish, G.G. Wong, *Cell* **57**, 277–285 (1989).
6. W. Jelkmamnn, *Phys. Rev.* **72**, 449–489 (1992).
7. D. Roberts and D.J. Smith, *J. Mol. Endocrinol.* **12**, 131–148 (1993).
8. S. Lok, K. Kaushansky, R.D. Holly, J.L. Kuijper, C.E. Lofton-Day, P.J. Oort, F.J. Grant, M.D. Heipel, S.K. Burkhead, J.M. Kramer, L.A. Bell, C.A. Sprecher, H. Blumberg, R. Johnson, D. Prunkard, A.E.T. Ching, S.L. Mathewes, M.C. Bailey, J.W. Forstrom, M.M. Buddle, S.G. Osborn, S.J. Evans, P.O. Sheppard, S.R. Presnell, P.J. O'Hara, F.S. Hagen, G.J. Roth, and D.C. Foster, *Nature* **369**, 565–568 (1994).
9. J.F. Bazan, *Biochem. Biophys. Res. Commun.* **147**, 392–398 (1989).
10. M.V. Abraham, M. Ultsch, and A.A. Kossiakoff, *Science* **255**, 306–312 (1992).
11. S. Elliott, T. Lorenzini, D. Chang, J. Barzilay, and E. Delorme, *Blood* **89**, 493–502 (1997).
12. J.W. Fisher, *Proc. Soc. Exp. Biol. Med.* **216**, 358–369 (1997).
13. H.F. Bunn and R.O. Poyton, *Physiol. Rev.* **76**, 839–885 (1996).
14. S.B. Krantz, *Blood* **77**, 419–434 (1991).
15. P.J. Ratcliffe, *Kid. Int.* **44**, 887–904 (1993).
16. H. Wu, X. Liu, R. Jaenisch, and H.F. Lodish, *Cell* **83**, 59–67 (1995).
17. S. Masuda, M. Okano, K. Yamagishi, M. Nagao, M. Ueda, and R. Sasaki, *J. Biol. Chem.* **269**, 19488–19493 (1994).
18. M. Digicaylioglu, S. Bichet, H.H. Marti, R.H. Wenger, L.A. Rivas, C. Bauer, and M. Gassmann, *Proc. Natl. Acad. Sci. USA.* **92**, 3717–3720 (1995).
19. E. Morishita, M. Masuda, M. Nagao, Y. Yasuda, and R. Sasaki, *Neuroscience* **76**, 105–116 (1997).
20. S. Masuda, M. Nagao, K. Takahata, Y. Konishi, F. Gallyas Jr., T. Tabira, and R. Sasaki, *J. Biol. Chem.* **268**, 11208–11216 (1993).
21. S. Yanagawa, K. Hirade, H. Ohnota, R. Sasaki, H. Chiba, M. Ueda, and M. Goto, *J. Biol. Chem.* **259**, 2707–2710 (1984).
22. M. Goto, K. Akai, A. Murakami, C. Hashimoto, E. Tsuda, M. Ueda, G. Kawanishi, N. Takahashi, A. Ishimoto, H. Chiba, and R. Sasaki, *Bio/Technology* **6**, 67–71 (1988).
23. M. Goto, A. Murakami, K. Akai, G. Kawanishi, M. Ueda, H. Chiba, and R. Sasaki, *Blood* **74**, 1415–1423 (1989).
24. L.M. Annable, P.M. Cotes, and M.V. Mussett, *Bull. World Health Org.* **47**, 99–112 (1972).
25. E. Tsuda, G. Kawanishi, M. Ueda, S. Masuda, and R. Sasaki, *Eur. J. Biochem.* **188**, 405–411 (1990).
26. K. Yamaguchi, K. Akai, G. Kawanishi, M. Ueda, S. Masuda, and R. Sasaki, *J. Biol. Chem.* **266**, 20434–20439 (1991).
27. H. Mizoguchi, K. Ohta, T. Suzuki, A. Murakami, M. Ueda, R. Sasaki, and H. Chiba, *Acta Hematol. Jpn.* **50**, 15–24 (1987).
28. W.A. Lukowsky and R.H. Painter, *Can. J. Biochem.* **50**, 909–917 (1972).
29. E. Goldwasser, K.H. Kung, and J. Eliasson, *J. Biol. Chem.* **249**, 4202–4206 (1974).
30. M.S. Dordal, F.F. Wang, and E. Goldwasser, *Endocrinology* **116**, 2293–2299 (1985).
31. R. Sasaki, S. Yanagawa, and H. Chiba, *Methods Enzymol.* **147**, 328–340 (1987).
32. M.N. Fukuda, H. Sasaki, L. Lopez, and M. Fukuda, *Blood* **73**, 84–89 (1989).
33. J.L. Spivak and B.B. Hogans, *Blood* **73**, 90–99 (1989).
34. M. Takeuchi, S. Takasaki, M. Shimada, and A. Kobata, *J. Biol. Chem.* **265**, 12127–12130 (1990).
35. A.J. Sytkowski, L. Feldman, and D.J. Zurbuch, *Biochem. Biophys. Res. Commun.* **176**, 698–704 (1991).
36. E. Delmore, T. Lorenzini, J. Giffin, F. Martin, F. Jacobsen, T. Boone, and S. Elliot, *Biochemistry* **31**, 9871–9876 (1992).
37. M. Higuchi, M. Oh-eda, H. Kuboniwa, K. Tomonoh, Y. Shimonaka, and N. Ochi, *J. Biol. Chem.* **267**, 7703–7709 (1992).
38. S. Dube, J.W. Fisher, and J.S. Powell, *J. Biol. Chem.* **263**, 17516–17521 (1988).
39. L.C. Wasley, G. Timony, P. Murtha, J. Stoudemire, A.J. Dorner, J. Caro, M. Krieger, and R.J. Kaufman, *Blood* **77**, 2624–2632 (1991).
40. M. Takeuchi, N. Inoue, T.W. Strickland, M. Kubota, M. Wada, R. Shimizu, S. Hoshi, H. Kozutumi, S. Takasaki, and A. Kobata, *Proc. Natl. Acad. Sci. USA* **86**, 7819–7822 (1989).
42. M. Nagao, K. Inoue, S.K. Moon, S. Masuda, H. Hakagi, S. Udaka, and R. Sasaki, *Biosci. Biotech. Biochem.* **61**, 670–674 (1997).
43. Y. Kitagawa, Y. Sano, M. Ueda, K. Higashio, H. Narita, M. Okano, S. Mastumoto, and R. Sasaki, *Exp. Cell Res.* **213**, 449–457 (1994).

EXPRESSION SYSTEMS. See TRANSIENT EXPRESSION SYSTEMS.

EXPRESSION SYSTEMS, *E. COLI*

SAVVAS C. MAKRIDES
PRAECIS Pharmaceuticals, Inc.
Cambridge, Massachusetts

KEY WORDS

Chaperone
E. coli
Fusion protein
Gene expression
Vector

OUTLINE

INTRODUCTION

Escherichia coli is a widely used organism for the production of recombinant proteins for both basic research and therapeutic applications. In spite of the extensive knowledge of the genetics and molecular biology of *E. coli*, the robust expression of genes in this prokaryote is not necessarily a routine matter. This may be due to the unique structural features of the gene sequence, the stability and translational efficiency of mRNA, the ease of protein folding, degradation of the protein by host cell proteases, major differences in codon usage between the foreign gene and native *E. coli*, and the potential toxicity of the protein to the host. Progress in our understanding of many of these processes is certain to broaden the usefulness of *E. coli* as a tool for gene expression. The major drawbacks of *E. coli* as an expression system are the lack of a secretion mechanism for the efficient release of protein into the culture medium, and the limited ability to facilitate extensive disulfide bond formation. In addition, although bacteria are capable of carrying out a variety of posttranslational modifications, including some types of glycosylation (1), these are not as complex as those performed by eukaryotic cells. On the other hand, many eukaryotic proteins retain their full biological activity in a nonglycosylated form and, therefore, can be produced in *E. coli*.

The objectives here are to focus on expression systems and methodological approaches useful for the high-level production of proteins in *E. coli*. Nutrient composition and fermentation variables are not discussed here, as these have been reviewed in detail (2). Due to space limitations many original references are not cited. The reader is referred to recent reviews for detailed discussion of specific topics (3–7).

EXPRESSION VECTORS

The construction of an expression plasmid requires several elements whose configuration must be carefully considered in order to ensure the highest levels of protein synthesis (7). The promoter is positioned approximately 10–100 bp upstream of the ribosome binding site (RBS) and is under the control of a regulatory gene that may be present on the vector itself or integrated in the host chromosome. Downstream of the promoter is the Shine and Dalgarno (SD) site, which interacts with the 3′ end of 16S rRNA during translation initiation. The distance between the SD region and the start codon ranges from 5 to 13 bases, and its sequence should eliminate the potential for secondary structure formation in the mRNA transcript, which can reduce the efficiency of translation initiation. Both 5′ and 3′ regions of the RBS exhibit a bias toward a high adenine content. The transcription terminator is located downstream of the coding sequence and serves both as a signal to terminate transcription as well as a protective element comprising stem–loop structures thereby protecting the mRNA from exonucleolytic degradation. In addition, vectors contain a gene that confers antibiotic resistance on the host to aid in plasmid selection and propagation. Finally, the copy number of plasmids is determined by the origin of replication. Although the use of runaway replicons has been reported to result in higher yields of plasmid-encoded protein, in general, higher copy number vectors are not advantageous (7).

Transcription

Promoters. There are many promoters available for gene expression in *E. coli*, including those derived from Gram-positive bacteria and bacteriophages (Table 1). A promoter for use in *E. coli* should have certain characteristics to render it suitable for high level protein synthesis. First, it must be strong, resulting in the accumulation of protein comprising 10–30% or more of the total cellular protein. Second, it should exhibit a minimum level of basal transcriptional activity. Large-scale gene expression preferably employs cell growth to high density and minimal promoter activity, followed by induction or derepression of the promoter. The tight regulation of a promoter is essential for the synthesis of proteins that may be detrimental to the host cell. Furthermore, incompletely repressed expression systems can cause plasmid instability, a decrease

Table 1. Promoters Used for the High-Level Expression of Genes in *E. coli*

Promoter (source)	Regulation	Induction
lac (E. coli)	*lacI*, $lacI^Q$	IPTG[a]
	*lacI*ts[b], $lacI^Q$ts[b]	Thermal
	*lacI*ts[c]	Thermal
lpp (E. coli)		IPTG, lactose[d]
trp (E. coli)		Tryptophan depletion, indole acrylic acid
tnaA (E. coli)[e]		Tryptophan, glucose depletion
Mn-SOD *(E. coli)*[f]		Oxygen, paraquat
phoA (E. coli)	*phoB* (positive)	Phosphate depletion
	phoR (negative)	
recA (E. coli)	*lexA*	Nalidixic acid
araBAD (E. coli)	AraC	L-Arabinose
proU (E. coli)		Osmolarity
cst-1 (E. coli)		Glucose depletion
tetA (E. coli)		Tetracycline
cadA (E. coli)	*cadR*	pH
nar (E. coli)	*fnr*	Anaerobic conditions, nitrate ion
nirB (E. coli)[g]		Anaerobic conditions
tac (hybrid) *(E. coli)*	*lacI*, $lacI^Q$	IPTG
	$lacI^h$	Thermal
trc (hybrid) *(E. coli)*	*lacI*, $lacI^Q$	IPTG
	*lacI*ts, $lacI^Q$ts	Thermal
lpp-lac (hybrid) *(E. coli)*	*lacI*	IPTG
$P_{syn(synthetic)}$ *(E. coli)*	*lacI*, $lacI^Q$	IPTG
$P_{LtetO-1}$ *(E. coli)*[i]		Anhydrotetracycline
$P_{lac/ara-1}$ *(E. coli)*[i]		IPTG, arabinose
Starvation promoters *(E. coli)*		
P_L (bacteriophage λ)	*lacI*ts857	Thermal
P_L-9G-50 (mutant) (λ)		Reduced temperature (<20 °C)
cspA (E. coli)		Reduced temperature (<20 °C)
P_R, P_L (tandem) (λ)	*lacI*ts857	Thermal
T7 (T7)	cascaded system[j]	IPTG
T7-*lac* operator (T7)	$lacI^Q$	IPTG
λP_L, P_{T7} (tandem) (λ, T7)	*lacI*ts857, $lacI^Q$	Thermal, IPTG
T3-*lac* operator (T3)	$lacI^Q$	IPTG
T5-*lac* operator (T5)	$lacI^Q$, *lacI*	IPTG
T4 gene 32 (T4)		T4 infection
nprM-lac operator (Bacillus)	$lacI^Q$	IPTG
vhb (Vitreoscilla)	*fnr*	Oxygen, cAMP-CAP[k]
protein A (*Staphylococcus aureus*)		
Pu (Pseudomonas putida)[l]	*xylR*	*m*-Toluic acid
Pm (P. putida)[l]	*xylS*	3-Methylbenzylacohol

Source: Updated from Ref. 7 with permission from the publisher.
Note: Except where indicated, original references may be found in Ref. 7.
[a]Isopropyl-β-D-thiogalactopyranoside.
[b]*lacI* gene with single mutation, Gly187Ser.
[c]*lacI* gene with three mutations, Ala241Thr, Gly265Asp, Ser300Asn.
[d]The constitutive *lpp* promoter (P_{lpp}) was converted into an inducible promoter by insertion of the *lac*UV5 promoter/operator region downstream of the P_{lpp}. Expression occurs only in the presence of a *lac* inducer.
[e]Ref. 8.
[f]Ref. 9.
[g]Ref. 10.
[h]Wild-type *lacI* gene.
[i]Ref. 11.
[j]Expression of T7 RNA polymerase is controlled by the *lacUV5* promoter, which is regulated by the *lac* repressor. Production of T7 RNA polymerase causes transcription of the recombinant gene, which is under the control of the φ10 promoter.
[k]Cyclic AMP-catabolite activator protein.
[l]Ref. 12.

in cell growth rate, and loss of recombinant protein production (13). A third important characteristic of a promoter is its inducibility in a simple and cost-effective manner. The most widely used promoters for large-scale protein production use thermal induction (λP_L) or chemical inducers (tac, trp) (Table 1). IPTG-inducible promoters are widely used in basic research. However, the use of IPTG for the large-scale production of human therapeutic proteins is undesirable due to its toxicity and cost. Moreover, IPTG is detrimental to the physiology of *E. coli* (14). The

availability of a mutant *lacI*ts gene that encodes a thermosensitive *lac* repressor now permits the thermal induction of these promoters (15,16).

Cold-responsive promoters have been shown to facilitate efficient gene expression at reduced temperatures. The rationale behind the use of cold-responsive promoters for gene expression is based on the proposition that the rate of protein folding will be only slightly affected at about 15–20 °C, whereas the rates of transcription and translation, being biochemical reactions, will be substantially decreased. This, in turn, will provide sufficient time for protein folding, yielding active proteins and avoiding the formation of inactive protein aggregates (inclusion bodies) without reducing the final yield of the target protein (17).

Transcription Terminators. Transcription through a promoter may inhibit its function, a phenomenon known as promoter occlusion (18). This interference can be prevented by the proper placement of a transcription terminator downstream of the coding sequence in order to prevent continued transcription through another promoter. Efficient transcription terminators serve several important functions: they minimize background transcription, stabilize plasmids, enhance mRNA stability, and can increase substantially the level of protein production (7).

The advantages of tightly regulated promoters have led to the design of completely repressible expression systems (19) that are particularly useful for the expression of genes whose products are detrimental to host growth. Most of these systems have not yet been used for the high-level production of proteins on a large scale.

Translation

Translational Initiation. The wide range of efficiencies in the translation of different mRNAs is predominantly due to the unique structural features at the 5′ end of each mRNA species. To date no universal sequence for the efficient initiation of translation has been devised. The translational initiation region of most sequenced *E. coli* genes contains the initiation codon AUG (91%). GUG is used by about 8% of the genes, and UUG (1%) is rarely used as a start site (20). The spacing between the SD site and the initiating AUG codon can vary from 5 to 13 nucleotides, and it influences the efficiency of translational initiation. The SD sequence UAAGGAGG is optimal for protein production (21). The secondary structure at the translation initiation region of mRNA plays a crucial role in the efficiency of gene expression. It is believed that the occlusion of the SD region and/or the AUG codon by a stem–loop structure prevents accessibility to the 30S ribosomal subunits and inhibits translation. Several different strategies have been devised to minimize mRNA secondary structure and enhance the expression of genes. These include the enrichment of the RBS with adenine and thymidine residues, the mutation of specific nucleotides upstream or downstream of the SD region, and the use of translational coupling or dicistronic systems (7). The insertion of a single amino acid residue (preferably arginine or lysine) between the N-terminal methionine and the protein of interest was reported to increase expression levels from 0.4% to 22% of the total cell protein (22). This strategy was effective for several types of proteins and different promoters. The recombinant proteins were converted to their native sequences by removal of dipeptides with cathepsin C. The precise mechanism for the observed increase in protein yields is unclear.

Translational Termination. The design of expression vectors frequently includes the insertion of all three stop codons in order to prevent possible ribosome "skipping." In *E. coli* there is a preference for the UAA stop codon (Table 2). A statistical analysis of more than 2,000 *E. coli* genes revealed local nonrandomness both in the stop codon and in the nucleotide immediately following the triplet and concluded that UAAU is the most efficient translational termination sequence in *E. coli* (23).

Codon Usage. Genes in both prokaryotes and eukaryotes show a nonrandom usage of synonymous codons (24). These observations imply that heterologous genes enriched with codons that are rarely used by *E. coli* (Table 2) may not be expressed efficiently in *E. coli*. The minor arginine $tRNA^{Arg(AGG/AGA)}$ has been shown to be a limiting factor in the bacterial expression of several mammalian genes because the codons AGA and AGG are infrequently used in *E. coli*. Many studies have demonstrated that the substitution of high-usage for low-usage codons or the coexpression of the cognate tRNA genes can ameliorate the detrimental effect of the rare codons on gene expression levels (25). The location of low-usage codons in the transcript may adversely impact translation. Thus, the production of β-galactosidase decreased when AGG codons were inserted before the 10th codon from the initiation codon of the *lacZ* gene (26). Similarly, translational inhibition of a test mRNA was much stronger when consecutive low-usage arginine codons were located near the 5′ end of the mRNA (27,28). Moreover, 5′ translational blockage was also observed with a different sequence of consecutive low-usage leucine codons (CUA), making it unlikely that struc-

Table 2. Codon Usage by *E. coli*

UUU	F	21.7	UCU	S	9.1	UAU	Y	16.2	UGU	C	5.1
UUC	F	16.7	UCC	S	8.9	UAC	Y	12.4	UGC	C	6.4
UUA	L	13.4	UCA	S	7.6	UAA	*	2.0	UGA	*	0.9
UUG	L	13.2	UCG	S	8.7	UAG	*	0.3	UGG	W	14.4
CUU	L	11.1	CCU	P	7.0	CAu	H	12.6	CGU	R	21.1
CUC	L	10.7	CCC	P	5.3	CAC	H	9.8	CGC	R	21.5
CUA	L	3.9	CCA	P	8.4	CAA	Q	14.7	CGA	R	3.6
CUG	L	51.9	CCG	P	22.8	CAG	Q	29.0	CGG	R	5.5
AUU	I	29.8	ACU	T	9.4	AAU	N	18.0	AGU	S	8.8
AUC	I	25.1	ACC	T	23.1	AAC	N	21.8	AGC	S	15.7
AUA	I	4.9	ACA	T	7.5	AAA	K	34.4	AGA	R	2.5
AUG	M	27.4	ACG	T	14.0	AAG	K	11.2	AGG	R	1.5
GUU	V	19.0	GCU	A	16.1	GAU	D	32.1	GGU	G	25.3
GUC	V	14.9	GCC	A	25.2	GAC	D	19.5	GGC	G	29.2
GUA	V	11.2	GCA	A	20.4	GAA	E	39.9	GGA	G	8.3
GUG	V	25.8	GCG	A	32.8	GAG	E	18.4	GGG	G	11.0

Source: Taken from GeneBank (*http://www.dna.affrc.go.jp/~nakamura/codon.html*).

Note: Codon usage is expressed as frequency per 1,000 codons, based on a total of 4,143,606 codons.

tural effects on the mRNA were responsible (28). In another case tandem AGG triplets caused a substantial inhibition of gene expression independently of their localization in mRNA, a phenomenon that was attributed to competition of the tandem AGGAGG codons with the natural SD sequence (29). In contrast, other workers reported efficient expression of genes that contained low-usage codons (7).

To date, it has not been possible to formulate precise "rules" to predict whether the content of low-usage codons in a specific gene might adversely affect the efficiency of its expression in *E. coli*. The experimental results may possibly be confounded by several variables such as positional effects, the clustering or interspersion of the rarely used codons, secondary structure of the message, and other effects (25). Nevertheless, it is clear that the codon context of specific genes *can* have adverse effects on both the quantity and quality of the expressed protein. Usually, this problem can be rectified by the alteration of the codons in question, or by the coexpression of the cognate tRNA genes.

mRNA Stability. mRNA degradation is not effected randomly by nonspecific endonucleolytic cleavage, since there is no inverse correlation between mRNA length and half-life (30). Two classes of protective elements are known to stabilize mRNAs in *E. coli*. One class consists of sequences in the 5′ untranslated regions (UTRs) of mRNAs, and the other class includes stem–loop structures from the 3′ UTRs and intercistronic regions. Some of these elements act as stabilizers when fused to heterologous mRNAs, but only under restricted conditions (7). In contrast, the 5′ UTR of the *E. coli ompA* transcript prolongs the half-life of a number of heterologous mRNAs in *E. coli* under normal conditions of rapid cell growth (31). Belasco and colleagues showed that the presence of a stem–loop at or very near the extreme 5′ terminus of the *ompA* 5′ UTR is essential for its stabilizing effect. Furthermore, the half-life of a normally labile mRNA was prolonged in *E. coli* by adding a hairpin structure at its 5′ terminus, and it was proposed that *E. coli* mRNAs beginning with a long single-stranded segment are preferentially targeted by a ribonuclease that interacts with the 5′ terminus before cleaving the mRNA at internal sites (31). It appears, therefore, that the addition of the *ompA* 5′ stabilizer to heterologous genes might enhance gene expression in *E. coli*. However, it is possible that mRNAs that contain internal RNA processing sites may not be protected by the presence of the 5′ stabilizer.

The other class of mRNA protective elements consists of 3′ UTR sequences that can form stem–loop structures, thereby blocking exonucleolytic degradation of the transcript from the 3′ terminus. Such an element was identified within the transcription terminator of the crystal protein gene of *Bacillus thuringiensis*. Its fusion to the 3′ termini of the penicillinase *(penP)* gene of *B. licheniformis* and the human interleukin 2 cDNA increased the half-life of the mRNAs and enhanced the production of the corresponding polypeptides in both *B. subtilis* and *E. coli* (32). However, as in the case of some of the 5′ stabilizers, this 3′ "retroregulator" is unlikely to act as a universal mRNA stabilizer. For example, the replacement of the 3′-terminal hairpins of labile mRNAs with those from stable messages did not enhance the expression of the labile transcripts (32–34). It has been suggested that gene expression might be enhanced by the use of host strains that are deficient in specific ribonucleases, such as RNase II or PNPase. This too is unlikely to be an effective strategy, because the absence of RNase II or PNPase, as well as the overproduction of RNase II, had no effect on the average half-life of *E. coli* bulk mRNA (35). Moreover, strains that were deficient in both RNase II and PNPase were inviable (36).

The deletion of 110 bp in the 3′ untranslated region of a cDNA encoding the human TNF-α was reported to enhance the expression level of this cytokine (37). The authors suggested that a UUUAUUA sequence at the 3′ end of the mRNA may have an inhibitory effect on the expression of TNF-α in *E. coli*. It is well known that 3′ AU-rich sequences in cytokines, growth factors, and oncoproteins have an inhibitory effect on translation and may destabilize the mRNA (reviewed in Ref. 38). However, AU-binding proteins (39) are not known to exist in bacteria (Gary Brewer, personal communication, 1999). Therefore, the observed increase in TNF-α yield (37) is probably due to some other mechanism.

CELLULAR COMPARTMENTS

Cytoplasm

Each of the four "compartments" of *E. coli*, i.e. the cytoplasm, periplasm, membranes/cell surface, and the extracellular medium, offers advantages as well as disadvantages for protein production (Table 3). The formation of inclusion bodies remains a significant barrier to gene expression in the cytosol. Inclusion bodies do offer several advantages (Table 3), particularly the production of proteins that are toxic to *E. coli*. For example, the highly insoluble bovine prochymosin has recently been used as a fusion partner for the production of antimicrobial peptides as inclusion bodies in *E. coli* (40). On the other hand, the advantages of inclusion bodies are counterbalanced by the arduous task of refolding the aggregated protein, the uncertainty of whether the refolded protein retained its biological activity, and the reduction in yield of the refolded and purified protein. To date the precise physicochemical parameters that contribute to the formation of inclusion bodies remain unclear (41). A statistical analysis of the composition of 81 proteins that do and do not form inclusion bodies in *E. coli* concluded that charge average and turn-forming residue fraction are strongly correlated with inclusion body formation (42). These observations led to the development of a model to predict the probability of inclusion body formation solely based on the amino acid composition of a protein (42). The model was used to predict accurately the insolubility of the human T-cell receptor Vβ5.3 in *E. coli* (15), as well as to enhance the solubility of a protein by adding negatively charged residues (43).

Several experimental approaches have been utilized to minimize the formation of inclusion bodies and improve protein folding. These include the growth of bacterial cultures at lower temperatures, selection of different *E. coli* strains, substitution of selected amino acid residues, co-

Table 3. Relative Merits of Different Compartments for Gene Expression in *E. coli*

Advantages	Disadvantages
Cytoplasm	
Inclusion bodies. Facile isolation of protein in high purity and concentration, target protein protected from proteases, desirable for production of proteins that, if active, are lethal to host cell.	*Inclusion bodies*. Protein insolubility, refolding to regain protein activity, refolded protein may not regain its biological activity, reduction in final protein yield, increase in cost of goods.
Higher protein yields	Reducing environment: does not facilitate disulfide bond formation.
Simpler plasmid constructs.	N-terminus authenticity.
	Proteolysis.
	Purification is more complex (more protein types).
Periplasm	
Purification is simpler (fewer protein types).	Signal peptide does not always facilitate transport.
Proteolysis is less extensive.	Inclusion bodies may form.
Improved disulfide bond formation/folding.	
N-terminus authenticity.	
Membranes/cell surface	
Surface display may facilitate drug screening, biocatalysis and other applications.	
Some progress made in overexpression of membrane proteins.	
Extracellular	
Least level of proteolysis.	No secretion, usually.
Purification is simpler (fewest protein types).	Protein diluted, more difficult to purify.
Improved protein folding.	
N-terminus authenticity.	

production of chaperones, use of *E. coli* DsbA or thioredoxin either as fusion partners or coproduced with the protein of interest, growth and induction of the cells under osmotic stress in the presence of sorbitol and glycyl betaine, addition of nonmetabolizable sugars to the growth medium, alteration of the pH of the culture medium, and the use of strains deficient in thioredoxin reductase (7,43).

The reducing potential of the cytoplasmic redox state (44) presents still another problem. Bacterial cytoplasmic proteins contain few cysteine residues and few disulfide bonds (44). Most proteins that contain stable disulfide bonds are exported from the cytoplasm (45). Thus, mammalian proteins whose complex tertiary structure depends in part on disulfide bond formation may not be produced in their correct conformation in the bacterial cytoplasm. Mutant *E. coli* strains were isolated that allow the formation of disulfide bonds in the cytoplasm (46). These mutations inactivate the *trxB* gene that encodes thioredoxin reductase and contributes to the sulfhydryl reducing potential of the cytoplasm. Thioredoxin itself was unnecessary for disulfide bond formation (46). The authors suggested that the cytoplasm may contain another thioredoxin-like protein that can be reduced by thioredoxin reductase; in the absence of thioredoxin reductase, the oxidized form of this protein facilitates formation of disulfide bonds in the cytoplasm. Other workers have utilized *E. coli* strains carrying null mutations in the *trxB* gene, and observed significant amounts of functional disulfide-containing protein in the cytoplasm (47).

Proteins produced in the cytoplasm may not possess an authentic N-terminus. The cytoplasmic expression of a gene without a leader requires the presence of an initiation codon, the most common one encoding methionine. Although this amino acid might have no adverse effect on the protein synthesized, there are specific cases in which the extraneous methionine residue may have profound consequences (7). Bacterial translation is initiated by *N*-formylmethionine, which is deformylated during synthesis (48) but not necessarily removed. The N-terminal methionine might be cleaved off by an endogenous methionine aminopeptidase (MAP) (49), depending on the side-chain length of the second amino acid residue (50,51). Thus, residues with small side chains such as Gly, Ala, Pro, Ser, Thr, Val, Cys, and to a lesser degree, Asn, Asp, Leu, and Ile facilitate the MAP-catalyzed removal of the N-terminal methionine. Several strategies have been successfully used to remove the extra methionine residue from recombinant proteins in vivo, including the coexpression of the *E. coli* methionine aminopeptidase gene and the use of exopeptidases (7).

Protein degradation in the cytosol and in other compartments is another concern, discussed in the section "Protein Stability". Finally, cytosolic gene expression necessitates purification of the target protein from the large pool of intracellular proteins.

Periplasm

The periplasm offers several advantages for protein targeting. In contrast to the cytosolic compartment, the periplasm contains only 4% of the total cell protein (52). The target protein is thus effectively concentrated, and its purification is considerably less onerous. The oxidizing environment of the periplasm facilitates the proper folding of proteins, and the cleaving in vivo of the signal peptide during translocation to the periplasm is more likely to yield the authentic N-terminus of the target protein. Protein degradation in the periplasm may be less extensive than in the cytosol, but it is still a significant issue.

The transport of a protein through the inner membrane to the periplasm normally requires a signal sequence. A wide variety of signal peptides have been employed successfully in *E. coli* for protein translocation to the periplasm. These include prokaryotic signal sequences, such as the *E. coli* PhoA signal, OmpA, OmpT, LamB, OmpF, β-lactamase, enterotoxins ST-II, LT-A, LT-B, protein A from *Staphylococcus aureus*, endoglucanase from *B. subtilis*, PelB from *Erwinia carotovora*, a degenerate PelB signal

sequence, the murine ribonuclease, and the human growth hormone signal (7). However, the presence of a signal peptide does not always ensure efficient protein translocation through the inner membrane. It is clear that besides the signal peptide, other structural features in proteins are involved in membrane transport. Strategies for the improved translocation of proteins to the periplasm include the supply of components involved in protein transport and processing, such as overproduction of signal peptidase I, use of *prlF* mutant strains, coexpression of the *prlA4* and *secE* genes, coexpression of the *prlF* gene, and expression of the *pspA* gene (7).

The intracellular accumulation of unprocessed proteins has often been interpreted as a possible saturation of the *E. coli* translocation system (e.g., Ref. 5). For example, Hsiung et al. (53) observed the intracellular accumulation of a greater amount of human growth hormone precursor after IPTG induction, but no increase in the amount of translocated growth hormone. Similarly, the *Chlamydia trachomatis* major outer membrane protein (MOMP) was successfully targeted to the *E. coli* surface, but upon prolonged induction, the unprocessed higher molecular weight MOMP accumulated (54). Although this may have been caused by saturation of the transport mechanisms, the authors point out that other explanations are also possible, such as differences in the carboxyl terminus of the chlamydial leader peptide, or formation of inclusion bodies, making a large proportion of MOMP unavailable for processing and resulting in the accumulation of unprocessed MOMP (54). Other workers dismiss the possibility of jamming of the protein export machinery, pointing out that under proper conditions correctly processed and fully active proteins can accumulate in the periplasm at a level 20–50% of the total cell protein (6). In agreement with this, insulin-like growth factor I (IGF-I) was expressed in the periplasm at up to 30% of the total cell protein (55). Moreover, there was no unprocessed IGF-I detected in the cytosol (John Joly, personal communication, 1998). However, IGF-I expression was achieved in the presence of the thiol oxidant proteins DsbA or DsbC, which might have increased protein yields by possibly assisting translocation of IGF-I. Interestingly, the overexpression of the disulfide oxidoreductases increased the fraction of aggregated IGF-I for reasons not understood (55).

The basic protein Skp has been identified as a periplasmic molecular chaperone (56). Overexpression of Skp was shown to increase the amount of fusion proteins incorporated into phage (57). The effect was most notable on scFv fragments that are aggregation prone, suggesting that Skp acts by increasing the amount of correctly folded proteins, which can then get incorporated into the phage coat. The chaperone composition of the periplasm is not as well characterized as that of the cytoplasm, and the mechanism of action of Skp is unclear (57). Further studies should demonstrate whether Skp and other periplasmic molecular chaperones have wide applicability in periplasmic protein expression.

Membranes and Cell Surface

The cell surface of *E. coli* may facilitate several applications including vaccine development, biocatalysis, protein–protein interactions, and drug screening (58). Progress in the high-level expression of membrane proteins in *E. coli* has been reviewed in detail (5). In many cases production of membrane proteins is limited by toxicity and bacterial cell death. A procedure has recently been described for the selection of *E. coli* strains that facilitate the overproduction of proteins and their incorporation into the membrane (59).

Extracellular Medium

The targeting of synthesized proteins for secretion to the culture medium presents significant advantages (Table 3). Unfortunately, *E. coli* normally secretes very few proteins, and the manipulation of the various transport pathways to facilitate secretion of foreign proteins remains a formidable task. The methodological approaches to protein secretion in *E. coli* may be conveniently divided into two categories: (*1*) the exploitation of existing pathways for "truly" secreted proteins (60); and (*2*) the use of signal sequences, fusion partners, permeabilizing proteins, nutrients, or other agents that effect protein secretion as a result of "leakage" or selective and limited permeability of the outer membrane. The first approach offers the advantage of specific secretion of the protein of interest, hence minimum contamination by nontarget proteins. For example, the hemolysin gene has been used for construction of secreted hybrid proteins (60). Secretion, however, is not a particularly efficient process. The second approach relies on the induction of limited leakage of the outer membrane to render protein secretion, albeit with modest yields. Examples are the use of the *pelB* leader, *ompA* leader, protein A leader, the coexpression of bacteriocin release protein along with the addition of glycine to the culture medium, and the coexpression of the *kil* gene for membrane permeabilization (7). A mutated bacteriocin release protein was reported to provide more efficient extracellular production of proteins by *E. coli* (61).

PROTEIN STABILITY

Proteolysis is a selective, highly regulated process that plays an important role in cellular physiology. *E. coli* contains at least 40 enzymes that catalyze the hydrolysis of peptide bonds (62). These enzymes participate in a wide range of metabolic processes including the removal of N- or C-terminal amino acid residues, catabolism of dipeptides, oligopeptides, and proteins, the selective removal of abnormal proteins, and the inactivation of regulatory proteins (62,63). Protein damage or alteration may result from a variety of conditions, such as incomplete polypeptides, mutations caused by amino acid substitutions, excessive synthesis of subunits from multimeric complexes, post-translational damage through oxidation or free radical attack, and genetic engineering. Such abnormal proteins are efficiently removed by the bacterial proteolytic machine. *E. coli* proteases are localized in the cytoplasm, the periplasm, and the inner and outer membranes (Table 4).

Some determinants of protein instability have been elucidated. The "N-end rule" relates the metabolic stability of a protein to its amino-terminal residue (64). Thus, in *E.*

Table 4. Localization of Proteases in *E. coli*

Cytoplasm	Inner membrane	Periplasm	Outer membrane
La (Lon)	SppA (protease IV)	Deg P (HtrA, Do)	OmpT (protease VII)
ClpP	DegS (HhoB)	DegQ (HhoA)	OmpP
ClpX	Protease VI	Ptr (Pi, III, pitrilysin)	OmpX
ClpA (Ti)	LepB (signal peptidase I)	Prc (Tsp, Re)	
ClpQ	LspA (signal peptidase II)	Mi	
Clp Y	SohB		
HflB (FtsH, TolZ)	Protease V		
Protease II			
Ci			
Fa			
In			
So			

Source: Adapted from Refs. 62 and 63.
Note: The nomenclature for *E. coli* proteolytic activities is confusing, and some of these proteases are found under multiple names. A few proteins originally identified as proteases turned out to have other primary activities. In addition to the listed proteases, there are at least 12 peptidases in the cytoplasm.

coli, N-terminal Arg, Lys, Leu, Phe, Tyr, and Trp conferred 2-min half-lives on a test protein, whereas all the other amino acids, except proline, conferred greater than 10-h half-lives on the same protein. As already discussed in "Cellular Compartments, Cytoplasm", amino acids with small side chains in the second position of the polypeptide facilitate the MAP-catalyzed removal of the N-terminal methionine. Therefore, these studies suggest that Leu in the second position would probably be exposed by the removal of the methionine residue and destabilize the protein. The second determinant of protein instability is a specific internal lysine residue located near the amino terminus (65). This residue is the acceptor of a multiubiquitin chain that facilitates protein degradation by a ubiquitin-dependent protease in eukaryotes. Another correlation between amino acid content and protein instability is presented in the PEST hypothesis, which proposes that proteins are destabilized by regions enriched in Pro, Glu, Ser, and Thr, flanked by certain amino acid residues (66). It was suggested that PEST-rich proteins may be produced efficiently in *E. coli*, which apparently lacks the PEST proteolytic system.

Strategies for minimizing proteolysis of recombinant proteins in *E. coli* have been reviewed (7,67,68). These include protein targeting to the periplasm or the culture medium, use of protease-deficient host strains, growth of the host cells at low temperature, construction of N- and/or C-terminal fusion proteins, tandem fusion of multiple copies of the target gene, coexpression of molecular chaperones, coexpression of the T4 *pin* gene, replacement of specific amino acid residues in order to eliminate protease cleavage sites, modification of the hydrophobicity of the target protein, and optimization of fermentation conditions. The usefulness of some of these methods may be limited, depending on the intended use of the recombinant protein. Thus, for example, the presence of fusion moieties on the target protein may possibly interfere with functional or structural properties or therapeutic applications of the product. The engineering of enzymatic or chemical cleavage sites for the subsequent removal of the fusion partners involves many considerations: the accessibility of the cleavage sites to enzyme digestion, the purity, specificity and cost of the commercially available enzymes, the authenticity of the N- or C-termini upon enzymatic digestion, the possible modification of the target protein upon chemical treatment, and so forth (e.g., Refs. 69, 70). For the large-scale production of fusion proteins, some of these difficulties are amplified. Similarly, the fusion of multiple copies of the target gene to effect multidomain polypeptides requires the subsequent conversion to monomeric protein units by cyanogen bromide cleavage. In this case, the target protein must not contain internal methionine residues and must be able to withstand harsh reaction conditions. Moreover, a limited extend of amino acid side-chain modification may occur, and the toxicity of CNBr presents a significant issue for large-scale cleavage reactions. Similarly, the "rational" modification of a protein sequence requires structural information that may not be available. Molecular chaperones (see later) have been used successfully to stabilize specific proteins (68), but this approach to date remains a hit or miss affair (71).

The cytoplasm of *E. coli* appears to contain a greater number of proteases than the periplasm. Therefore, proteins located in the periplasm are less likely to be degraded (72). However, proteolytic activity in the periplasm is substantial (62). Secretion into the culture medium would provide a better alternative in terms of protein stability. Unfortunately, the technology for secretion of proteins from *E. coli* into the culture medium is still in its infancy. A major catalyst of protein degradation in bacteria is the induction of heat shock proteins in response to a variety of stress conditions, such as the thermal induction of gene expression or the accumulation of abnormal or heterologous proteins in the cytoplasm (73). Under these conditions the production of the *lon* gene product, protease La, and other proteases, is enhanced. This problem is minimized by the use of host strains deficient in the *rpoH (htpR)* locus (74). The *rpoH* gene encodes the RNA polymerase σ^{32} subunit, which regulates several proteolytic activities in *E. coli* Hosts that carry the *rpoH* mutation have been demonstrated to increase substantially the production of foreign proteins in *E. coli* (7). A large number of protease-deficient hosts exists (e.g., Refs. 74, 75), including those

that are deficient in all known protease loci that affect the stability of secreted proteins (76).

PROTEIN FOLDING

The efficient posttranslational folding of proteins, the assembly of polypeptides into oligomeric structures, and the localization of proteins are mediated by specialized proteins termed molecular chaperones. The demonstration that the production and assembly of prokaryotic ribulose bisphosphate carboxylase in *E. coli* requires both groES and groEL proteins, led to an increasing interest in the use of molecular chaperones for high-level gene expression in *E. coli*. The experimental results from the use of chaperones, however, have been inconsistent, and thus far the effects of chaperone coproduction on gene expression in *E. coli* appear to be protein specific (71, 77). It is also unclear whether the in vivo levels of different chaperone species are limiting under conditions of gene overexpression (78).

Contradictory results may be due to several reasons. Proteins that are drastically destabilized will probably fold incorrectly, even in the presence of chaperones. Thus, the truncation of polypeptides, the production of single domains from multisubunit protein complexes, lack of formation of disulfide bonds (which ordinarily contribute to protein structure), or absence of posttranslational modifications such as glycosylation may make it impossible to attain thermodynamic stability. It was shown that calmodulin coexpression was essential to generate active inducible nitric oxide synthase (iNOS) in *E. coli*, and it was suggested that calmodulin binds to iNOS after translation and aids in folding and stabilizing the enzyme (79). It is now clear that different types of chaperones normally act in concert (80). Therefore, the overproduction of a single chaperone may be ineffective (81). Another variable to consider is growth temperature. For example, GroES/EL coexpression increased the production of β-galactosidase at 30 °C, but not at 37 °C or 42 °C, whereas DnaK and DnaJ were effective at all temperatures tested (82). Finally, the overexpression of chaperones can lead to phenotypic changes, such as cell filamentation, that can be detrimental to cell viability and protein production (81).

The coexpression of the human (83) or rat (84) protein disulfide isomerase (PDI) enhanced the yield of correctly folded protein in the *E. coli* periplasm. Disulfide bond formation in the *E. coli* periplasm is facilitated by a group of proteins that maintain the correct redox potential (85). It is thought that DsbA, a soluble periplasmic protein, directly catalyzes disulfide bond formation in proteins, whereas DsbB, an inner membrane protein, is involved in the reoxidation of DsbA. Eukaryotic PDI was capable of complementing the phenotypes of *dsbA* null mutants (83,84), but its function was virtually abolished in *dsbB* mutants (84). In addition, the ability of PDI to enhance the yield of target proteins was increased in the presence of exogenously added glutathione (84). These observations suggest that PDI depends on the presence of bacterial redox proteins for its reoxidation. The coexpression of rat PDI has also been reported to enhance the correct folding of tissue plasminogen activator (84).

Protein misfolding can be attributed to the intracellular concentration of aggregation-prone intermediates, suggesting that reducing the rate of protein synthesis should disfavor protein misfolding. Indeed, the use of weaker promoters or conditions of partial induction from stronger promoters can result in higher amounts of soluble protein (82). The ability of *E. coli* to secrete a large amount of alkaline phosphatase into the periplasm was due to a lower synthetic rate of the *phoA* gene product (86), and a reduction in promoter strength enhanced the extracellular production of exoglucanase by *E. coli* (87).

FUSION PROTEINS

In recent years there has been a remarkable increase in the variety of fusion proteins for biological research (70,88). Table 5 includes most of the published fusion moieties. Other studies have addressed the design and engineering of excision sites for the chemical or enzymatic cleavage and removal of fusion partners (69,70). This section briefly summarizes the use of selected fusion systems that have a direct impact on high-level production and, in some cases, secretion of target proteins.

Uhlén and colleagues developed a multifunctional fusion partner based on staphylococcal protein A or synthetic derivatives thereof. In addition to its utility as an affinity tag for purification, the protein A moiety acts as a solubilizing partner to improve folding (92), and the presence of the protein A signal peptide causes the secretion of the gene product to the culture medium. An alternative fusion partner is derived from streptococcal protein G (SPG), a bacterial cell wall protein that has separate binding regions for albumin within the amino-terminal domains, and for IgG within the carboxyl-terminal domain. The combination of both protein A and SPG domains formed a tripartite fusion protein, thus providing an additional purification option and further protecting the target protein from proteolytic degradation (93).

Another fusion partner, thioredoxin, has been shown to dramatically increase the solubility of fusion proteins produced in the *E. coli* cytoplasm, and to prevent the formation of inclusion bodies (88). Similarly, the thioredoxin homolog DsbA has been used as a fusion partner to direct the transport of proteins to the periplasm (94). More recently, a mutant form of DsbA ($DsbA^{mut}$) was used as a fusion partner to confer higher solubility to 10 different eukaryotic proteins expressed in the cytoplasm at 37 °C (43). The consistent performance of $DsbA^{mut}$ led the authors to conclude that this protein may be used as a generic fusion partner for enhancing the solubility of a variety of heterologous eukaryotic proteins. An additional advantage of $DsbA^{mut}$ may be its ability to exert its solubilizing effect at 37 °C rather than at the lower growth temperatures often employed for enhancing solubility of proteins: at reduced temperatures the cell growth cycle becomes longer, with a significant impact on fermentation cycle time and a reduced protein yield (43).

The fusion of genes to the ubiquitin sequence significantly increased the yield of proteins (95). The remarkable

Table 5. Fusion Partners

Fusion partner	Ligand/matrix	Purification conditions
Flag®	Anti-Flag monoclonal antibodies M1, M2	Low calcium, EDTA, glycine
His_6	Ni^{2+}-nitrilotriacetic acid	Imidazole
Glutathione-S-transferase	Glutathione-Sepharose	Reduced glutathione
Staphylococcal protein A	IgG-Sepharose	Low pH, IgG-affinity ligand
Streptococcal protein G	Albumin	Low pH, albumin-affinity ligand
Calmodulin	Organic or peptide ligands, DEAE-Sephadex	Low calcium, EGTA
Calmodulin-binding peptides[a]	Calmodulin	Low calcium, EGTA
Thioredoxin	ThioBond™ resin	Ion exchange
Flagellin *(E. coli)*[b]		
β-Galactosidase	TPEG[c] Sepharose	Borate
Ubiquitin		
Chloramphenicol acetyltransferase	Chloramphenicol-Sepharose	Chloramphenicol
S-peptide (RNAse A, residues 1-20)	S-protein (RNAse A, residues 21–124)	Denaturing or nondenaturing conditions
Myosin heavy chain		Differential solubility in low/high salt
DsbA		
Biotin subunit (in vivo biotinylation)		
Avidin	Biotin	Denaturation (urea, heat)
Streptavidin	Biotin	Denaturation (urea, heat)
Strep-tag	Streptavidin	2-Iminobiotin, diaminobiotin
c-*myc*	anti-*myc* antibody	
Dihydrofolate reductase	Methotrexate-agarose	Folate buffer
CKS[d]		
Polyarginine	S-Sepharose	NaCl
Polycysteine	Thiopropyl-Sepharose	Dithiothreitol
Polyphenylalaline	Phenyl-Superose	Ethylene glycol
lac repressor	*lac* operator	Lactose analogue, DNAse, restriction endonuclease
T4 gp55		
Growth hormone, N-terminus		
Maltose-binding protein	Amylose resin	Maltose
Galactose-binding protein	Galactose-Sepharose	Galactose
Cyclomaltodextrin glucanotransferase	*α*-Cyclodextrin-agarose	*α*-Cyclodextrin
Cellulose-binding domain	Cellulose	Water
Hemolysin A *(E. coli)*		
λ cII protein		
TrpE or TrpLE		
Protein kinase site(s)		
$(AlaTrpTrpPro)_n$		Aqueous two-phase extraction
HAI[e]-epitope		
BTag (VP7 protein region of Bluetongue virus)	Anti-BTag antibodies	
Green fluorescent protein		
Prochymosin[f]		
TNFb[g]		

Source: Updated from Ref. 7 with permission from the publisher.
Note: Except where indicated, original references may be found in Ref. 7.
[a]Ref. 89.
[b]Ref. 90.
[c]*p*-Amino-phenyl-*β*-D-thiogalactoside.
[d]CTP:CMP-3-deoxy-D-manno-octulosonate cytidyltransferase.
[e]Influenza virus hemagglutinin; "reverse epitope tagging": tagging of the chromosomal rather than the plasmid-encoded protein, to avoid the need to remove the fusion partner.
[f]Highly insoluble, facilitates formation of inclusion bodies (40).
[g]Mutated tumor necrosis factor gene (91).

increase in protein yield was thought to be due to protection of the target protein from proteolysis, improved folding, and efficient mRNA translation. Ubiquitin or the ubiquitin metabolic pathway is absent in prokaryotic organisms. In order to remove the ubiquitin moiety from fusion proteins, Baker et al. (96) coexpressed the ubiquitin-specific protease Ubp2 in *E. coli*, thus effecting the cotranslational cleavage of ubiquitin from the fusion protein.

SUMMARY

An efficient prokaryotic expression vector contains a strong and tightly regulated promoter, a Shine–Dalgarno

site that is positioned approximately 9 bp 5′ to the translation initiation codon and is complementary to the 3′ end of 16S rRNA, and a transcription terminator positioned 3′ to the protein-coding sequence. In addition, the vector requires an origin of replication, a selection marker, and a gene that facilitates the stringent regulation of promoter activity. The translational initiation region of a gene must be free of secondary structures that may occlude the initiation codon and/or block ribosome binding. UAAU is the most efficient translation termination sequence in *E. coli*.

Each of the four "compartments" for targeted protein production, namely, the cytoplasm, periplasm, membranes/cell surface, and the growth medium, offers advantages and disadvantages for gene expression. The formation of inclusion bodies can be minimized using a variety of techniques, but it remains a significant barrier to high-level protein production in the cytoplasm. The effectiveness of molecular chaperones has been protein specific so far.

Codon usage can have adverse effects on the synthesis and yield of recombinant proteins. Positional effects appear to play an important role in protein synthesis. Thus, the presence of rare codons near the 5′ end of a transcript probably affects translational efficiency. This problem may be rectified by the alteration of the rare codons, and/or the coexpression of the cognate tRNA genes.

To date there is no generally applicable strategy to prevent the degradation of a wide variety of mRNA species in *E. coli*. Although certain 5′ and 3′ stem–loop structures have been shown to block mRNA degradation, these seem to stabilize only specific mRNAs, under restricted conditions. One exception appears to be the 5′ UTR of the *E. coli ompA* transcript, which prolongs the half-life of a number of heterologous mRNAs in *E. coli*.

The wide variety of existing fusion partners have utility in the production, detection, and purification of recombinant proteins. Specific fusion moieties can increase the folding, solubility, resistance to proteolysis, and secretion of recombinant proteins into the growth medium. Protein misfolding, attributed to the intracellular concentration of aggregation-prone intermediates, may be minimized by a combination of experimental approaches: replacement of amino acid residues that cause aggregation, coexpression of molecular chaperones and foldases, reduction of the rate of protein synthesis, the use of solubilizing fusion partners, and the careful optimization of growth conditions.

Proteolysis in *E. coli* may be minimized by the use of protease-deficient host strains, the construction of fusion proteins, the coexpression of molecular chaperones, the elimination of protease cleavage sites through genetic engineering, and the optimization of fermentation conditions. Host strains that are deficient in the *rpoH (htpR)* locus are effective, particularly for thermally induced expression systems.

BIBLIOGRAPHY

1. E.I. Tuomanen, *J. Clin. Invest.* **98**, 2659–2660 (1996).
2. S.Y. Lee, *Trends Biotechnol.* **14**, 98–105 (1996).
3. D.V. Goeddel, *Methods Enzymol.* **185**, 3–7 (1990).
4. A.R. Shatzman, *Curr. Opin. Biotechnol.* **6**, 491–493 (1995).
5. R. Grisshammer and C.G. Tate, *Q. Rev. Biophys.* **28**, 315–422 (1995).
6. G. Georgiou, in J.L. Cleland and C.S. Craik eds., *Protein Engineering: Principles and Practice*, Wiley-Liss, New York, 1996, pp. 101–127.
7. S.C. Makrides, *Microbiol. Rev.* **60**, 512–538 (1996).
8. K.C. Sitney, M.B. Mann, G.W. Stearns, A.D. Menjares, J.L. Stevenson, M.D. Snavely, J.C. Fieschko, C. Curless, and L.B. Tsai, *Ann. NY Acad. Sci.* **782**, 297–310 (1996).
9. B.F. Gao, S.C. Flores, S.K. Bose, and J.M. McCord, *Gene* **176**, 269–272 (1996).
10. S.M.C. Newton, P.E. Klebba, M. Hofnung, and A. Charbit, *Res. Microbiol.* **146**, 193–202 (1995).
11. R. Lutz and H. Bujard, *Nucleic Acids Res.* **25**, 1203–1210 (1997).
12. J.M. Blatny, T. Brautaset, H.C. Winther-Larsen, P. Karunakaran, and S. Valla, *Plasmid* **38**, 35–51 (1997).
13. W.E. Bentley, N. Mirjalili, D.C. Andersen, R.H. Davis, and D.S. Kompala, *Biotechnol. Bioeng.* **35**, 668–681 (1990).
14. M.J. Kosinski, U. Rinas, and J.E. Bailey, *Appl. Microbiol. Biotechnol.* **36**, 782–784 (1992).
15. B. Andrews, H. Adari, G. Hannig, E. Lahue, M. Gosselin, S. Martin, A. Ahmed, P.J. Ford, E.G. Hayman, and S.C. Makrides, *Gene* **182**, 101–109 (1996).
16. N. Hasan and W. Szybalski, *Gene* **163**, 35–40 (1995).
17. Patent WO 96/03521, A.B. Oppenheim, H. Giladi, D. Goldenberg, S. Kobi, and I. Azar.
18. S. Adhya and M. Gottesman, *Cell* **29**, 939–944 (1982).
19. N. Hasan and W. Szybalski, *Gene* **56**, 145–151 (1987).
20. C.O. Gualerzi and C.L. Pon, *Biochemistry* **29**, 5881–5889 (1990).
21. S. Ringquist, S. Shinedling, D. Barrick, L. Green, J. Binkley, G.D. Stormo, and L. Gold, *Mol. Microbiol.* **6**, 1219–1229 (1992).
22. R.M. Belagaje, S.G. Reams, S.C. Ly, and W.F. Prouty, *Protein Sci.* **6**, 1953–1962 (1997).
23. E.S. Poole, C.M. Brown, and W.P. Tate, *EMBO J.* **14**, 151–158 (1995).
24. T. Ikemura, *Mol. Biol. Evol.* **2**, 13–34 (1985).
25. J.F. Kane, *Curr. Opin. Biotechnol.* **6**, 494–500 (1995).
26. G.-F.T. Chen and M. Inouye, *Nucleic Acids Res.* **18**, 1465–1473 (1990).
27. E. Goldman, A.H. Rosenberg, G. Zubay, and F.W. Studier, *J. Mol. Biol.* **245**, 467–473 (1995).
28. W. Gao, S. Tyagi, F.R. Kramer, and E. Goldman, *Mol. Microbiol.* **25**, 707–716 (1997).
29. I. Ivanov, R. Alexandrova, B. Dragulev, A. Saraffova, and M.G. AbouHaidar, *FEBS Lett.* **307**, 173–176 (1992).
30. C.-Y.A. Chen and J.G. Belasco, *J. Bacteriol.* **172**, 4578–4586 (1990).
31. S.A. Emory, P. Bouvet, and J.G. Belasco, *Genes Dev.* **6**, 135–148 (1992).
32. H.C. Wong and S. Chang, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3233–3237 (1986).
33. J.G. Belasco, G. Nilsson, A. von Gabain, and S.N. Cohen, *Cell* **46**, 245–251 (1986).
34. C.-Y.A. Chen, J.T. Beatty, S.N. Cohen, and J.G. Belasco, *Cell* **52**, 609–619 (1988).
35. W.P. Donovan and S.R. Kushner, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 120–124 (1986).

36. W.P. Donovan and S.R. Kushner, *Nucleic Acids Res.* **11**, 265–275 (1983).
37. J. Chang, W. Cai, L. Xu, C. Li, and S. Zhao, *Chin. J. Biotechnol.* **11**, 213–219 (1995).
38. V. Kruys and G. Huez, *Biochimie* **76**, 862–866 (1994).
39. G.M. Wilson and G. Brewer, in K. Moldave ed., *Progress in Nucleic Acid Research and Molecular Biology*, vol. 62, Academic, San Diego, 1999, pp. 257–291.
40. C. Haught, G.D. Davis, R. Subramanian, K.W. Jackson, and R.G. Harrison, *Biotechnol. Bioeng.* **57**, 55–61 (1998).
41. R. Rudolph and H. Lilie, *FASEB J.* **10**, 49–56 (1996).
42. D.L. Wilkinson and R.G. Harrison, *Bio/Technology* **9**, 443–448 (1991).
43. Y. Zhang, D.R. Olsen, K.B. Nguyen, P.S. Olson, E.T. Rhodes, and D. Mascarenhas, *Protein Expr. Purif.* **12**, 159–165 (1998).
44. R.C. Fahey, J.S. Hunt, and G.C. Windham, *J. Mol. Evol.* **10**, 155–160 (1977).
45. J.M. Thornton, *J. Mol. Biol.* **151**, 261–287 (1981).
46. A.I. Derman, W.A. Prinz, D. Belin, and J. Beckwith, *Science* **262**, 1744–1747 (1993).
47. K. Proba, L.M. Ge, and A. Plückthun, *Gene* **159**, 203–207 (1995).
48. J.M. Adams, *J. Mol. Biol.* **33**, 571–589 (1968).
49. A. Ben-Bassat, K. Bauer, S.-Y. Chang, K. Myambo, A. Boosman, and S. Chang, *J. Bacteriol.* **169**, 751–757 (1987).
50. F. Sherman, J.W. Stewart, and S. Tsunasawa, *BioEssays* **3**, 27–31 (1985).
51. P.-H. Hirel, J.-M. Schmitter, P. Dessen, G. Fayat, and S. Blanquet, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8247–8251 (1989).
52. N.G. Nossal and L.A. Heppel, *J. Biol. Chem.* **241**, 3055–3062 (1966).
53. H.M. Hsiung, N.G. Mayne, and G.W. Becker, *Bio/Technology* **4**, 991–995 (1986).
54. J.E. Koehler, S. Birkelund, and R.S. Stephens, *Mol. Microbiol.* **6**, 1087–1094 (1992).
55. J.C. Joly, W.S. Leung, and J.R. Swartz, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2773–2777 (1998).
56. R. Chen and U. Henning, *Mol. Microbiol.* **19**, 1287–1294 (1996).
57. H. Bothmann and A. Plückthun, *Nat. Biotechnol.* **16**, 376–380 (1998).
58. G. Georgiou, H.L. Poetschke, C. Stathopoulos, and J.A. Francisco, *Trends Biotechnol.* **11**, 6–10 (1993).
59. B. Miroux and J.E. Walker, *J. Mol. Biol.* **260**, 289–298 (1996).
60. J.A. Stader and T.J. Silhavy, *Methods Enzymol.* **185**, 166–187 (1990).
61. F.J. van der Wal, G. Koningstein, C.M. ten Hagen, B. Oudega, and J. Luirink, *Appl. Environ. Microbiol.* **64**, 392–398 (1998).
62. C.G. Miller, in F.C. Neidhardt, R. Curtiss III, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger eds., Escherichia Coli and Salmonella. *Cellular and Molecular Biology*, vol. 1, American Society for Microbiology Press, Washington, D.C., 1996, pp. 938–954.
63. S. Gottesman, *Annu. Rev. Genet.* **30**, 465–506 (1996).
64. A. Varshavsky, *Cell* **69**, 725–735 (1992).
65. A. Bachmair and A. Varshavsky, *Cell* **56**, 1019–1032 (1989).
66. S. Rogers, R. Wells, and M. Rechsteiner, *Science* **234**, 364–368 (1986).
67. F. Baneyx and G. Georgiou, in T.J. Ahern, and M.C. Manning eds., *Stability of Protein Pharmaceuticals. Part A: Chemical and Physical Pathways of Protein Degradation*, Plenum, New York, 1992, pp. 69–108.
68. M. Murby, M. Uhlén, and S. Ståhl, *Protein Expr. Purif.* **7**, 129–136 (1996).
69. P. Carter, in M.R. Ladisch, R.C. Willson, C.C. Painton, and S.E. Builder eds., *Protein Purification: From Molecular Mechanisms to Large-Scale Processes*, vol. 427, American Chemical Society, Washington, D.C., 1990, pp. 181–193.
70. P.-Å. Nygren, S. Ståhl, and M. Uhlén, *Trends Biotechnol.* **12**, 184–188 (1994).
71. J.G. Wall and A. Plückthun, *Curr. Opin. Biotechnol.* **6**, 507–516 (1995).
72. K. Talmadge and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1830–1833 (1982).
73. S.A. Goff and A.L. Goldberg, *Cell* **41**, 587–595 (1985).
74. S. Gottesman, *Methods Enzymol.* **185**, 119–129 (1990).
75. F. Baneyx and G. Georgiou, *J. Bacteriol.* **173**, 2696–2703 (1991).
76. H.J. Meerman and G. Georgiou, *Bio/Technology* **12**, 1107–1110 (1994).
77. T. Yasukawa, C. Kaneiishii, T. Maekawa, J. Fujimoto, T. Yamamoto, and S. Ishii, *J. Biol. Chem.* **270**, 25328–25331 (1995).
78. A. Knappik, C. Krebber, and A. Plückthun, *Bio/Technology* **11**, 77–83 (1993).
79. C. Wu, J. Zhang, H. Abu-Soud, D.K. Ghosh, and D.J. Stuehr, *Biochem. Biophys. Res. Commun.* **222**, 439–444 (1996).
80. J. Buchner, *FASEB J.* **10**, 10–19 (1996).
81. P. Blum, J. Ory, J. Bauernfeind, and J. Krska, *J. Bacteriol.* **174**, 7436–7444 (1992).
82. G. Georgiou and P. Valax, *Curr. Opin. Biotechnol.* **7**, 190–197 (1996).
83. D.P. Humphreys, N. Weir, A. Mountain, and P.A. Lund, *J. Biol. Chem.* **270**, 28210–28215 (1995).
84. M. Ostermeier, K. De Sutter, and G. Georgiou, *J. Biol. Chem.* **271**, 10616–10622 (1996).
85. J.C.A. Bardwell, *Mol. Microbiol.* **14**, 199–205 (1994).
86. H. Kadokura, K. Yoda, S. Watanabe, Y. Kikuchi, G. Tamura, and M. Yamasaki, *Appl. Microbiol. Biotechnol.* **41**, 163–169 (1994).
87. T.-L. Lam, R.S.C. Wong, and W.-K.R. Wong, *Enzyme Microb. Technol.* **20**, 482–488 (1997).
88. E.R. LaVallie and J.M. McCoy, *Curr. Opin. Biotechnol.* **6**, 501–506 (1995).
89. C.-F. Zheng, T. Simcox, L. Xu, and P. Vaillancourt, *Gene* **186**, 55–60 (1997).
90. B. Westerlund-Wikström, J. Tanskanen, R. Virkola, J. Hacker, M. Lindberg, M. Skurnik, and T.K. Korhonen, *Protein Eng.* **10**, 1319–1326 (1997).
91. L. Wang, H. Wu, F. Dou, W. Xie, and X. Xu, *Biochem. Mol. Biol. Int.* **41**, 1051–1056 (1997).
92. E. Samuelsson, T. Moks, B. Nilsson, and M. Uhlén, *Biochemistry* **33**, 4207–4211 (1994).
93. M. Murby, L. Cedergren, J. Nilsson, P.-Å. Nygren, B. Hammarberg, B. Nilsson, S.-O. Enfors, and M. Uhlén, *Biotechnol. Appl. Biochem.* **14**, 336–346 (1991).
94. L.A. Collins-Racie, J.M. McColgan, K.L. Grant, E.A. DiBlasio-Smith, J.M. McCoy, and E.R. LaVallie, *Bio/Technology* **13**, 982–987 (1995).
95. A.L. Pilon, P. Yost, T.E. Chase, G.L. Lohnas, and W.E. Bentley, *Biotechnol. Prog.* **12**, 331–337 (1996).

96. R.T. Baker, S.A. Smith, R. Marano, J. McKee, and P.G. Board, *J. Biol. Chem.* **269**, 25381–25386 (1994).

See also CHINESE HAMSTER OVARY CELLS, RECOMBINANT PROTEIN PRODUCTION; EXPRESSION SYSTEMS, MAMMALIAN CELLS; INSECT CELL CULTURE, PROTEIN EXPRESSION; INSECT CELLS AND LARVAE, GENE EXPRESSION SYSTEMS; PICHIA, OPTIMIZATION OF PROTEIN EXPRESSION.

EXPRESSION SYSTEMS, MAMMALIAN CELLS

SUBINAY GANGULY
ALLAN R. SHATZMAN
SmithKline Beecham Pharmaceuticals
King of Prussia, Pennsylvania

KEY WORDS

293 cells
Amplification
CHO cells
COS cells
mAbs
Regulated expression
Selectable markers
Stable expression
Transient expression

OUTLINE

INTRODUCTION

Gene expression systems offer a way to produce a heterologous protein encoded by a specific gene of interest through its introduction into a specific host cell. Heterologous gene expression of recombinant proteins has many applications, for example, in the biochemical and structure and function analysis of complex proteins, promoter analysis of genes of interest, and development of cell/membrane-based assays. One of the most crucial applications of mammalian cell expression systems is the production of therapeutic proteins such as growth factors and recombinant monoclonal antibodies (mAbs). Mammalian gene expression systems are also being used in gene-therapy-related experiments for the treatment of human diseases. Moreover, various expression systems utilizing mammalian cells can be used to determine functions for novel genes that are being identified through genomics. Dozens (if not hundreds) of mammalian cell expression systems have been described in the literature. The efficiency of expression of heterologous genes in mammalian cells is determined by the choice of host cells, method of gene transfer, sites at which the heterologous genes are integrated into the host genome (in the case of stable expression), and *cis*-acting DNA elements (e.g., enhancers) present in the vector carrying the gene of interest. In addition, efficiency of transcription, mRNA transport and stability, efficiency of translation, and secretion of the expressed protein into the media (for secreted proteins) may also play a role in the successful expression of recombinant proteins. Steps such as posttranslational modification, glycosylation, phosphorylation, and proteolytic processing must also be considered because they can affect the quality of the protein made in or secreted by recombinant host cells. This review summarizes various host–vector and expression systems commonly available for transient and stable expression of heterologous proteins in mammalian cells (for review see Refs. 1–5) and aims to illustrate the use of some of the most common ones.

GENERAL CONSIDERATIONS AND STRATEGIES FOR OPTIMIZING EXPRESSION

Host Cell Types

Commonly Used Cell Types. Mammalian cells such as chinese hamster ovary (CHO) cells, myeloma cells, human embryonic kidney (HEK)293 cells, baby hamster kidney (BHK) cells, mouse Ltk^- cells, and NIH3T3 cells have been frequently used for stable expression of heterologous genes. In contrast, cell lines such as Cos and HEK293 (see later) are routinely used for transient expression of recombinant proteins.

CHO-K1 and $DHFR^-$ CHO cells DG44 (6) and DUK-B11 (7) have been extensively used for high-level protein production (see later) because these cells enable the amplification of genes of interest via selection with the amplifiable marker DHFR (8). A wide variety of proteins have been expressed in CHO cells, including enzymes, recom-

binant mAbs, blood coagulation factors, 7TM-receptors, and other cell surface markers, chemokines, adhesion molecules, growth factors, and their receptors (see later for examples).

Myeloma cells such as Sp2/0, YB2/0, NSO, and P3X63.Ag8.653 have also been commonly employed for high-level production of mAbs (9–13). Several examples of this application will be described later in this review.

HEK293 cells, which were derived after transformation of primary human kidney cells by adenovirus (Ad virus) early-region-encoded oncoproteins, have been widely used to stably express many 7TM receptors that can functionally couple to G-proteins. It has been shown that the human cytomegalovirus (hCMV) enhancer/promoter in an expression plasmid is highly active in these cells because Ad virus E1A protein expressed in HEK293 cells transactivates the hCMV enhancer/promoter (14). Examples of 7TM receptor expression in these cells are the *Drosophila* dopamine D1 receptor (15), the human PTH receptor (16), the human muscarinic cholinergic receptor Hm1 (17), the human CGRP type 1 receptor (18), and the rat bradykinin B2 receptor (19). In addition to these receptors, a number of chemokine receptors such as the human CC CKR5 receptor (20), the mouse mCCR2 receptor (21), the mouse eotaxin receptor (22), the human MCP3 receptor (23), and the human MCP-1 receptor (24) have been successfully expressed in HEK293 cells. High-level expression on the cell surface and functional coupling of most the 7TM receptors expressed in HEK293 cells have been achieved. HEK293 cells have also been used for expression of various enzymes, as exemplified by the bovine constitutive nitric oxide synthase (eNOS) (25), and the human prostacyclin synthase (26). HEK293 cells have been used for heterologous expression of a wide variety of proteins, for example, interleukin 1 receptor–IgG fusion protein (immunoadhesion) (27), the human insulin growth factor binding protein 3 (28), platelet integrin $\alpha_{IIb}\beta_3$ (an integral membrane protein that mediates adhesion of cells to the extracellular matrix) (29), and the human tissue factor (a procoagulant factor that serves as a cofactor for the blood coagulation factor VII) (30).

Cos cells have been used to transiently express a large number of heterologous proteins. Cos cells constitutively express simian virus 40 (SV40) T-antigen and therefore support replication of expression plasmids containing the SV40 origin of replication. This results in high copy number ($\sim$40,000 copies cell^{-1}) of the transcribed coding sequences (31). Cos cells are primarily used for three reasons: *(1)* to determine if the expression plasmid produces the protein of interest, *(2)* for rapid expression of recombinant proteins to perform initial functional studies, and *(3)* expression cloning. Hundreds (if not thousands) of proteins have been expressed transiently in Cos cells. For example, Patel and Grundy (32) used a Cos cell expression system to study biogenesis of apolipoprotein B–containing lipoproteins. Functional expression of 7TM-receptors, for example, the guniea pig α2A, α2B, and α2C-adrenoreceptors (33), and the rabbit bradykinin B1 receptor (34), have been successfully achieved in Cos cells. Cos cells have been used commonly to express small amounts of functional enzymes, such as the rat liver microsomal glutathione transferase (35), the human 11β-hydroxylase and aldosterone synthase (36). Functionally active recombinant mAbs have also been expressed transiently in Cos cells (for a review see Ref. 4). It should be remembered however, that the Cos expression system is a transient expression system, and therefore, multiple repeat transfection experiments, which are time-consuming and labor intensive, are needed to produce milligram quantities of proteins for initial studies. More importantly, data obtained from a Cos expression system cannot necessarily be extrapolated to the level of expression expected from stable expression systems.

Modified Hosts. Some of the host cell lines just described have been engineered to produce milligram quantities of recombinant proteins in a rapid manner. In most cases this has been achieved by expressing (constitutively) a transcriptional transactivator that stimulates transcription from the promoter linked to the gene of interest. Following are descriptions of some of these modified hosts.

CHOE1a. To quickly express a heterologous protein, Cockett et al. (37) have established a CHO cell line constitutively expressing the Ad virus E1A, which transactivates the CMV promoter in the expression plasmid. Transfection of a human procollagenase gene into this cell line produced a 13-fold higher level in stable expression compared with that in CHO-K1 cells. This system has been used to express mAbs (13,38) and single-chain Fvs (39).

CHO/Pit-1. CHO cells were engineered to express the rat Pit-1 transactivator (a pituitary specific transcription factor). Stable transfection of a reporter gene (tPA) under the control of the rat prolactin promoter showed an enhancement of expression levels ($\sim$20- to 60-fold) compared with the same promoter in parallel in CHO cells (40).

BHK/VP16. Hippenmeyer and Highkin (41) have established a BHK cell line that stably expresses the herpes simplex (HSV) VP16, a transcriptional transactivator. Using this cell line, tPA and bGH were expressed at high levels (1–20 μg 10^6 cells^{-1} day^{-1}) as compared with that in the parental BHK cell line ($<$0.2 μg 10^6 cells^{-1} day^{-1}). Since then, a wide variety of proteins have been stably expressed in the BHK/VP16 cell line (42).

293/EBNA. HEK293 cell lines have been established after stably transfecting the Epstein–Barr virus (EBV) nuclear antigen 1 (EBNA-1) protein, which is required for the maintenance of a plasmid carrying the EBV origin of replication extrachromosomally (43). Such cell lines have been used for heterologous protein expression. For example, using a secreted form of human placental alkaline phosphatase (SEAP) as a reporter gene, Cachianes et al. (44) obtained 10-fold higher expression in 293 cells expressing EBNA-1 as compared with that in wild-type 293 cells. A number of other proteins (e.g., tPA, vascular endothelial growth factor, hepatocyte growth factor, soluble tumor necrosis factor) have been produced in this cell line at levels ranging from 0.2 to 10 μg ml^{-1} (44).

Mammalian Cells Expressing T7 RNA Polymerase Mouse Ltk$^-$ cells, CHO cells, and mouse myeloma Sp2/0 cells expressing T7 RNA polymerase have been established (45).

In these cells, T7 RNA polymerase has been modified by fusion of a nuclear localization signal to the N-terminus of the polymerase so that it can function in the cell nuclei. Transfection of a heterologous gene under the control of a T7 promoter into these modified host cell lines enables high-level expression of the gene of interest. For example, using this system the human growth hormone has been stably expressed at levels of 20–30 μg 10^6 cells^{-1} day^{-1} (45).

Vector Systems

A typical mammalian expression vector contains a cDNA copy of the gene of interest placed downstream of a promoter sequence to ensure efficient transcription; placed upstream is a polyA addition signal, which is required for efficient transcription termination. Enhancer sequences, which stimulate transcription, are usually placed upstream of the linked promoter, whereas a selectable marker (e.g., amp^r gene) and a plasmid origin of replication are included for propagation of the expression construct in *E. Coli*. In some instances, a drug selectable/amplifiable marker expression cassette (see later) for use in eukaryotic cells is linked to the heterologous gene expression unit. As mentioned, for transient expression in Cos cells, the SV40 origin of replication is included in the expression plasmid. See Figure 1 for examples of mammalian expression vectors.

A wide variety of promoters and enhancers and a combination thereof have been used for heterologous expression in mammalian cells and are included in vectors that are commercially available. Among the more popular promoters are those from *(1)* animal viral genomes such as SV40, human CMV, Ad virus, Rous sarcoma virus (RSV), murine leukemia virus (MuLV), HSV, and so forth; and *(2)* cellular genes such as β-globin, immunoglobulin, heat shock, and metallothionein genes (2,3,46,47). SV40 early and late promoters, hCMV IE promoter, RSV long-terminal repeat (LTR), and SRα promoter (a composite promoter consisting of the SV40 early promoter and a part of the LTR of human T-cell leukemia virus) (48) have all been successfully utilized in CHO-K1 and DHFR$^-$ CHO cells for expression of a variety of proteins. However, the choice of promoter is not usually critical when high-level expression is desired in DHFR$^-$ CHO cells where gene amplification methodology will be employed (see later). Plasmids containing the CMV promoter have been shown to produce high-level expression in HEK293 cells (49). Most of the expression vectors used for mAb expression in myeloma cells contain both H- and L-chain cDNAs under the transcription control of IgG promoter, although the hCMV promoter has been used successfully for mAb expression in NSO myeloma cells (10). A polyA addition signal (AA/TTAAA) and downstream sequences that influence transcription termination can be derived from many sources including the SV40 genome, the bovine growth hormone gene, and Ig genes.

The expression level of the gene of interest can sometimes be increased several fold through the use of an intronic sequence, for example, β-globin (50) or chimeric J/Ck intron (51). It is advisable to include the intron at the 5′ end of the coding sequence because, in some examples (52), aberrant processing of transcript was found when the intron was placed at the 3′ end of the cDNA. However, it should be pointed out that the effect of introns in enhancing expression levels is not a general phenomenon and in most cases does not significantly affect expression levels.

Kozak Sequence

In most vertebrate mRNAs examined to date, a sequence related to GCCACC*AUG*G, called a Kozak sequence (the underlined AUG is the initiator methionine), has been identified at the 5′ end, which can promote efficient translation of a linked open reading frame (ORF) (53). Therefore, many investigators include an optimal Kozak sequence in their expression vector. For example, Ho et al. (54) studied translation initiation and assembly of rat peripherin, an intermediate filament protein in transfected cells. A full-length peripherin cDNA (designated as p61-11), when expressed in an adrenocarcinoma cell line (SW13 c1.2 vim$^-$) without any detectable cytoplasmic intermediate filament, did not yield a functional filamentous network because of low-level expression of p61-11. When the first ATG was placed in the context of a better Kozak consensus sequence, the translational efficiency of p61-11 was increased significantly, and the resulting 57 kDa protein was able to form a functional filamentous network in the vim$^-$ cells.

Translational Enhancers

Regulatory sequences that stimulate translation, hence called translational enhancers, have been identified at the 5′ end of many eukaryotic viral RNAs and some cellular mRNAs (55–57). These sequences bind to translation initiation factors with higher affinity than their cellular counterparts. The Ad virus tripartite leader sequence is such a translational enhancer and has been used in mammalian expression vectors. For example, Sheay et al. (58) compared expression of a reporter CAT gene from expression vectors containing the Ad virus tripartite leader sequence placed downstream of various promoters such as CMV, RSV, and MuLV. In transient transfection assays, expression vectors containing the Ad virus tripartite leader sequence gave 5- to 20-fold higher CAT activity when compared with that from vectors without the leader sequence. A translational enhancer has also been identified in the 5′ UTR of a cellular mRNA (e.g., the the human β-kinesin mRNA encoding the first exon) (59). High levels of reporter gene expression have been achieved from transfection of plasmids carrying the kinesin enhancer sequences when compared with that obtained without the enhancer. For example, in Cos transient transfections, the kinesin enhancer placed downstream of the HSV tK promoter gave a four-fold increase in reporter gene (CAT) activity when compared with that from a plasmid without the enhancer.

A 34-bp-long translational enhancer has been described in the *Xenopus* β-globin mRNA. When placed upstream of a reporter gene, up to 300-fold increases in translational efficiency have been demonstrated in in vitro systems (60).

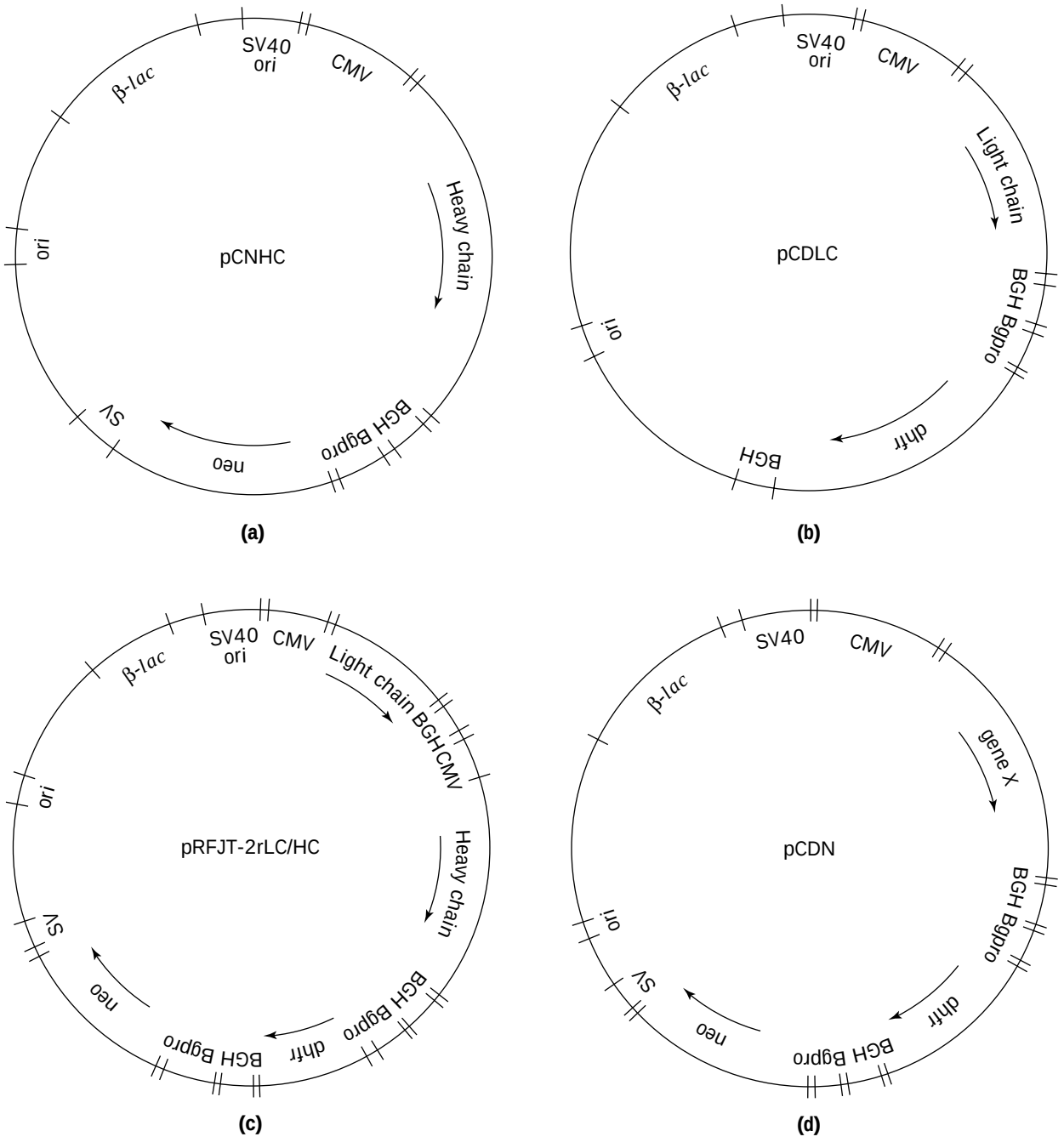


Figure 1. Mammalian expression vectors. pCNHC **(a)** and pCDLC **(b)** are used for mAb expression after subcloning Ig heavy- and light-chain cDNAs separately. **(c)** Single-vector pRFJT-2rLC/HC for mAB expression. **(d)** pCDNX for expression of gene of interest (X) in CHO, Cos, and 293 cells. CMV, cytomegalovirus promoter; BGH, bovine growth hormone polyA addition signal; Bgpro, mouse *β*-globin promoter; dhfr, dihydrofolate reductase cDNA; neo, neomycin phosphotransferase gene; SV, SV40 polyA additional signal; ori, plasmid origin of replication; *β-lac*, *β*-lactamase gene; SV ori, SV40 origin of replication. Orientation of the genes is shown by arrows.

Translational enhancers have also been identified in many plant viral genomes (55), with the tobacco mosaic virus (TMV) enhancer receiving the most attention. However, the use of this and other plant translation enhancers for improving heterologous mammalian expression systems has not yet been explored.

Selectable/Amplifiable Markers

A number of dominant selectable markers such as neomycin resistance (Neo^r), hygromycin resistance (Hyg^r), *E. Coli*. xanthine-guanylyl phosphotransferase (Ecogpt), and glutamine synthetase (GS) have been used for initial se-

lection of clones obtained after transfection of the expression plasmids into appropriate host cells (2,8,61,62). A mouse DHFR cDNA has been used for both initial selection of DHFR resistant CHO cells, as well as gene amplification (8). The selectable/amplifiable marker expression unit (under the control of a weak promoter and polyA addition signal) can either be linked to the heterologous gene expression unit or can be on a separate plasmid that is usually cotransfected into the host cells. With the use of an expression vector containing a linked DHFR gene, the gene copy number in the initially selected $DHFR^+$ clones can be amplified by selecting the clones further in the presence of stepwise increasing concentrations of methotrexate (MTX), a competitive inhibitor of DHFR.

Glutamine synthetase can be used as a dominant selectable/amplifiable marker when cells are grown in the presence of methionine sulfoxamine (MSX), an inhibitor of GS (62) (see later). Other gene amplification markers that have been used to a lesser extent are encoded by adenosine deaminase (ADA) and multiple drug resistance (MDR) genes (8,61).

Transfection Methods

A variety of methods can be used to introduce expression plasmids into mammalian host cells (2,63,64). These include $CaPO_4$, DEAE-dextran, electroporation, and lipofection. Receptor-mediated gene transfer has also been described in the literature (65). Interestingly, a single, standard, optimized method has yet to be described in the literature for establishing mammalian cells expressing heterologous gene products. In a typical stable expression experiment, the plasmid carrying the gene of interest and a drug selectable/amplifiable marker is transfected into host cells using any of the gene transfer methods just mentioned. The transfected cells are selected in the presence of a drug until the drug-resistant colonies are obtained (usually within 2–3 weeks). The resultant colonies are then screened using a variety of methods to determine which are producing the desired heterologous gene product. The detection method is dependent on the nature of the expressed protein, availability of antibodies, and so forth. For secreted proteins, conditioned media can easily be screened by ELISA, western blotting, and so forth. On average, one-third to one-half of the drug-selected colonies produce the protein of interest at a detectable level. The other clones are either expressing the desired gene product at very low levels or not at all, presumably due to inactivation or rearrangements of the heterologous gene.

Growth Media

Growth media and process development play an important role in the successful expression of a heterologous gene product in mammalian cells. Serum-free media has been developed for large-scale growth of CHO and myeloma cells. Serum-free or protein-free media offers the following advantages: *(1)* they are less expensive compared with media containing serum; *(2)* downstream purification of heterologous gene products becomes easier, especially for proteins that are secreted into the media; and *(3)* contaminating agents present in serum are avoided (66–69).

Increases in expression levels of heterologous gene products can normally be achieved through media manipulation and process development. For example, Bebbington et al. (10) have reported a three- to fourfold increase in mAb expression in NSO cells amplified with the GS system when cells were grown in a defined serum-free medium. When an amino acid feed was applied to these cells, an additional increase in expression level up to 560 mg L^{-1} was achieved. Using NSO myeloma cells and the GS system, expression level of a number of mAbs in the range of 0.5–2 g L^{-1} have been reported through medium development, nutrient supplementation, and other process development considerations (70). Similar results have also been obtained with mAbs produced in $DHFR^-$ CHO cells (4,70,71).

Glycosylation patterns of recombinant proteins can vary widely when expressed in different mammalian host cell types. For example, Parekh et al. (72,73) reported different glycosylation patterns for recombinant human tPA expressed in four different cell types: CHO cells, murine epithelial (C127) cells, the Bowes melanoma cells, and human colon fibroblast (CCD-18Co) cells. Cell culture conditions also determine glycosylation pattern of recombinant proteins (for review see Ref. 74). Borys et al. (75) found that culture pH affects expression rate and glycosylation of recombinant mouse placental lactogen proteins produced in CHO cells. The optimum expression rate was found to be in pH range of 7.6 to 8.0. The extent of glycosylation was decreased at low pH (below 6.9) and high pH (above 8.2). N-linked oligosaccarides were found to be different in composition when recombinant human tissue kallikrein was produced by CHO cells grown in serum-free suspension culture or on microcarrier beads (76).

Regulated Expression

Inducible expression systems have been proven to be valuable for elucidating the function of a number of genes in *E. Coli*, yeast, and *Drosophila*. In mammalian cells a tightly regulated expression system may also offer an advantage for expression of those proteins that are potentially cytotoxic. One way to achieve regulated expression of a heterologous gene is through the use of an inducible promoter. A number of inducible promoters have been described in the literature, such as IFN-γ, hsp70, metallothionein (Mtn), and MMTV LTR (for review see Refs. 46 and 77). Following are some popular methods for achieving inducible expression of heterologous genes in mammalian cells.

The Lacswitch System. Bacterial transcription control systems such as the *lac* operator/repressor have been adapted for regulated expression of heterologous genes in mammalian cells (78). The *lac* repressor, made from a *lacI* expression vector, blocks transcription of a reporter gene placed downstream of *lac* operator sequences. Addition of the inducer IPTG results in derepression of reporter gene transcription by up to 10- to 60-fold. Labow et al. (79) modified the *lac* repressor/operator system by converting the *lac* repressor into a transactivator by fusing it to the HSV

VP16 transactivation domain. In the absence of IPTG, this *lac* transactivator stimulates expression from a minimal promoter containing *lac* operator sequences up to 1,000-fold.

The *tet* System. Gossen and Bujard (80) described a regulated expression system based on the control elements of the *E. coli* tetracycline (Tc) resistance operon. In this system (for review see Ref. 81) the *tet* repressor has been fused to the HSV VP16 transactivator domain, and the resultant transactivator, called tTA, stimulates expression from a hCMV minimal promoter linked to *tet* operator sequences. In the presence of Tc (1 μg ml^{-1}), the tTA cannot bind to the operator sequence, and transcription from the minimal promoter shuts off, whereas removal of Tc leads to activation of transcription. In stable HeLa clones expression of a reporter gene has been shown to be modulated up to 10^5-fold. Derivatives of Tc, such as anhydrotetracycline (82) and doxycyclin (83), have also been used to regulate expression using this system.

An expression vector with the *tet*-responsive promoter, tTA expression unit, origin of replication of EBV, EBNA-1 expression unit, and a selectable marker has been described for the use in transient and stable expression of heterologous genes (84). Using this vector, luciferase activity in transiently transfected cells can be regulated by several orders of magnitude by varying both the Tc concentration and incubation time in the presence of Tc.

The *tet*-responsive system has been shown to produce different responses in different cell lines. Howe et al. (85) used the *tet* system to regulate *(1)* expression of dopamine D3 receptor in stable GH3 cells and *(2)* GluR6 (a glutamate receptor subunit) in stable HEK293 cells. In GH3 cells, a high-level induction of dopamine D3 receptor expression was obtained in the absence of Tc, whereas only a modest level induction of GluR6 was seen in HEK293 cells.

Recently, a mutant form of the *tet* repressor has been used to create a transactivator after fusion with the HSV VP16 (83). This transactivator, called rtTA, has reversed DNA binding activity as compared with that of wild-type *tet* repressor or tTA. In the presence doxycycline or anhydrotetracycline, more than 1,000-fold induction has been achieved in stable HeLa clones.

Ecdysone-Inducible System. The *Drosophila* ecdysone (Ec) receptor, a member of the steroid hormone receptor superfamily, can function as a ligand-dependent transcription factor in animal cells (86). Functional EcR is a heterodimer of EcR and the *Drosophila* ultraspiracle gene product (USP). The activity of the EcR is regulated by certain ecdysteroids such as ecdysone (Ec) and not by mammalian hormones such as dexamethasone and estradiol. EcR by itself is incapable of high-affinity DNA binding and transcriptional activation. In the presence of Ec, physical association between EcR and USP is greatly enhanced (87). Structural analysis shows that steroid hormone receptors contain a DNA-binding domain, ligand-binding domain, and transactivation domain. Mammalian retinoic acid receptor (R-X-R) can replace USP to form a functional heterodimer.

Recently, No et al. (88) described an ecdysone-inducible gene expression system for mammalian cells. Mammalian cells cotransfected with EcR and USP gave only a threefold induction upon treatment with a micromolar concentration of muristerone (MurA), a synthetic analog of Ec. Modifications of EcR were made such that the N-terminal transactivation domain in EcR was replaced by the corresponding domain of the glucocorticoid receptor (GR). This hybrid receptor gave a 34-fold induction in a transient reporter gene assay. A more potent heterodimer was made by combining either R-X-R and VpEcR (an N-terminus of EcR was truncated and fused to the VP16 transactivation domain) or R-X-R and VgEcR (which recognizes a hybrid Ec/GRE). Stable lines generated with these regulatory elements and induced with 100 nM MurA gave 20,000-fold induction of reporter gene (β-galactosidase) expression.

All three of the inducible systems are commercially available. For system it is important to determine the level of leaky expression in the absence of inducer (for the *lac* and Ec system) and the presence of Tc (for the *tet* system). The optimal concentration of the inducer and the kinetics of induction need to be determined. More importantly, one needs to determine if the high-fold induction obtained through any inducible system correlates with the desired high-level expression of a recombinant protein.

Dicistronic Vectors

In prokaryotic systems, many mRNAs encode two or more gene products in polycistronic messages. Expression of the distal coding sequences are often regulated by the translation of the upstream coding sequence in these messages. In contrast, eukaryotic mRNAs are not generally polycistronic. However, translation of polycistronic message in eukaryotic systems has now been achieved through the use of an internal ribosome entry site (IRES). IRESs have been identified in the single-stranded RNA genome of picornaviruses such as encephalomyocarditis (EMC) virus, foot and mouth disease virus, and polio virus. Ribosomes bind to a 450- to 600-nt-long secondary structure (IRES) present in a 600- to 1,200-nt-long 5′ UTR followed by an ORF in the viral genome. Kaufman (89) first reported the use of an IRES placed in between the gene of interest (a cDNA-encoding eukaryotic initiation factor 2α [eIF-2α]) and a selectable marker (DHFR). The resultant expression vector is called a dicistronic vector. When transfected into DHFR$^-$ CHO cells, efficient internal translation initiation of DHFR yielded increased transfection efficiency. High-level expression of eIF-2α was observed when the DHFR$^+$ clones were grown in presence of increasing concentration of MTX.

A number of dicistronic vectors have been described in the literature for heterologous gene expression (90,91). In order to increase the expression level of a gene of interest (5′ cistron), vectors have been designed such that the translation initiation of the selectable marker gene (3′ cistron) is reduced. The resultant high-level expression of the heterologous gene in the marker-selected (drug-selected) clones can be due to the integration of the expression plasmid in a transcriptionally active region of the host genome (92). For example, Kobayashi et al. (93) constructed a di-

cistronic vector in which transcription is under the control of the CMV enhancer/chicken β-actin promoter. The vector contains an IRES of EMC virus downstream of a reporter β-galactosidase gene, followed by a neomycin resistance gene. The translation of *neo*^{r} gene was weakened by placing spacers of various lengths in between AUG-11 (located at the 3′ end of IRES) and the AUG of *neo*^{r}. The number of Neo^r 3T3 clones obtained after transfection of the dicistronic vector with the weakened *neo*^{r} was 4- to 7-fold less compared with that from a conventional dicistronic vector, but these clones gave about 10-fold higher levels of *β-gal* activity. A similar strategy of weakening the selectable marker gene expression from a dicistronic vector was described by Rees et al. (94). Expression levels of serotonin receptors in 10 independent stable HEK293 clones varied from 1.5 to 11.9 pmol mg^{-1} cellular protein, which was significantly higher than the reported expression levels (98 fmol mg^{-1} in HeLa to 4 pmol mg^{-1} in Lmtk$^-$ cells).

Lucas et al. (95) reported a dicistronic intron vector (or DI vector) in which the mouse DHFR cDNA is placed into the intron of the pRK vector (96) and is under the transcriptional control of the CMV enhancer/promoter. A cDNA encoding a gene of interest (e.g., tPA or TPO) was placed downstream of the intron. When transfected into DHFR-CHO cells, full-length mRNA produced by DI vectors was translated into DHFR, while spliced mRNA produced the desired recombinant protein. DI-vector-derived CHO clones grown for 5 days gave 11-fold more tPA ($\sim$1 mg L^{-1}) and 13-fold more TPO than were produced by conventional vector-derived cells. The DI vector system has also been used to produce antibodies at levels $\sim$4–5 mg L^{-1} 10 day^{-1}. The weakening of DHFR expression and/or its linkage to product expression gave fewer clones, but they had more productivity.

APPLICATIONS OF MAMMALIAN CELL EXPRESSION TECHNOLOGY

Specific Examples of CHO Cell Expression

Growth Factors. The following examples describe the use of DHFR$^-$ CHO cells for expression of various growth factors. The expression vectors were either cotransfected with a DHFR expression plasmid to select the DHFR$^+$ clones or contained their own DHFR expression cassette. Expression levels of the growth factors were increased via amplification in the presence of increasing concentrations of MTX, as already described.

Madisen et al. (97) achieved high-level expression of TGF-β2 and a precursor TGF-β2 (414) in CHO cells using an expression vector similar to that described by Gentry et al. (98) for TGF-β1. A pSV2 expression vector containing the genes encoding TGF-β2 (414) and DHFR was utilized to obtain DHFR$^+$ clones. These clones were subsequently adapted to 0.5, 2.5, 10.0, and 50 μM MTX. Two clones grown at 10 and 50 μM MTX were analyzed for expression of recombinant mature TGF-β2 (12 kDa under reducing conditions). In addition to the mature protein, proteins of 85, 105, and 130 kDa, consisting of both mature and proregion sequences were also produced, as determined by immunoblotting. The pro-TGF-β2 was found to be both glycosylated and phosphorylated.

Bourdrel et al. (99) expressed recombinant human TGF-β1 after subcloning the cDNA into the pDSVα2 expression vector, a derivative of the pCD vector (100) in which the SV40 polyA addition signal was replaced by the polyA addition signal of the α-subunit of bovine follicular stimulating hormone and a mouse DHFR minigene (downstream of the expression cassette). CHO cells were transfected using the $CaPO_4$ method, and a single DHFR$^+$ clone expressing human TGF-β1 at 7 mg L^{-1} in the conditioned media was identified.

Ferrara et al. (101) used CHO cells to produce the full-length vascular endothelial growth factor (VEGF) after subcloning the cDNA encoding the full-length human VEGF cDNA into a derivative of the vector pCISh (49), which carries a DHFR gene. The transcription of VEGF cDNA and that of the DHFR are both under the control of the SV40 promoter. The initial selected clones, which released the highest mitogenic activity, were amplified in the presence of increasing concentration of MTX, resulting in clones that were resistant to 1 μM MTX and produced 0.5 μg ml^{-1} day^{-1} of VEGF.

The human neurotrophin 3 (NT-3) gene product was purified and characterized after expression in CHO cells (102). A construct that contained a chimera encoding the human nerve growth factor pre-pro region and NT-3 mature region under the control of MuLV-LTR was transfected into CHO cells. DHFR$^+$ clones were serially subcultured in media containing 0.1, 1.0, and 10 μM MTX, yielding clones expressing as much as 1 mg recombinant NT-3 L^{-1}. The purified recombinant NT-3 was biologically active, as shown through the induction of neurite outgrowth in neurons prepared from an 8-day-old chicks.

Burgess et al. (103) used CHO cells to study biosynthetic processing of the rat *neu* differentiation factors (NDFs). pDSRα2-derived vectors (104) containing NDF-β2c and NDF-β4a were transfected using the $CaPO_4$ method, and the selected clones were screened by western blotting. High-expressing clones for NDF-α2c and NDF-β4a were identified. Expression of NDF-β4a was gradually amplified with MTX (70 nM). The precursor of NDF undergoes typical glycosylation, and a portion of the molecule is proteolytically cleaved, leading to secretion of soluble mature NDF. A portion of the newly synthesized precursors escape processing and are retained on the cell surface.

Monoclonal Antibodies. In recent years DHFR$^-$ CHO cells have been successfully used to produce a number of recombinant mAbs at a high level (4,51,71,105–107). For example, Page and Sydenham (105) used DUK-B11 CHO cells to obtain expression of the humanized mAb Campath-1 at levels of 100 μg 10^6 $cells^{-1}$ day^{-1}. Using the same host cells, Fouser et al. (51) reported expression of a chimeric ganglioside GD2 antibody at 80–110 μg 10^6 $cells^{-1}$ day^{-1}. Recently, Reff et al. (106) expressed a chimeric antibody against CD20 in DG44 CHO cells at levels of 30 μg 10^6 $cells^{-1}$ day^{-1}, using only one round of gene amplification with MTX.

A variety of promoters have been used to obtain expression of H- and L-chains in these CHO cells, for example, the β-actin promoter (105), the human CMV MIE promoter (106), the Ad virus major late promoter (MLP) (108), the RSV promoter (4), and a murine leukemia virus LTR (109). A number of vectors for mAb expression are described in the literature in which the two Ig chains are carried by two different plasmids with an independent selectable/amplifiable marker. For example, Fouser et al. (51) linked the Ig H-chain with the ADA gene, and the L-chain with the DHFR gene. Using these vectors, an expression level of 80–110 μg of mAb 10^6 cells^{-1} day^{-1} was achieved. Trill et al. (4) have used vectors containing H-chain linked to a DHFR marker, and an L-chain expression cassette with the Neo^r marker and vice versa to obtain $\sim$180 mg of a humanized mAb L^{-1} 7 day^{-1} in spinner flasks. The methods used for initial selection and subsequent amplification vary from laboratory to laboratory. In general, high-level mAb expression can be obtained using the following steps: initial selection and subsequent amplification of candidate clones, coselection (e.g., in cases where both H-chain and L-chain expression vectors carry DHFR expression unit) and amplification, coamplification using different amplifiable markers, and initial selection and amplification in mass culture, followed by dilution cloning to identify individual high-expressing clones.

Because integration sites may influence the efficiency of H-chain and L-chain expression and overall mAb expression, single vectors have been created in which the two Ig-chain expression units are placed in tandem. These vectors also carry a dominant selectable marker such as Neo^r and the DHFR expression cassette (4,105,107,109). Using the single-vector system and two rounds of amplification, Newman et al. (107) achieved high-level expression (>200 μg ml^{-1}) of a primatized anti-CD4 mAb. Similar high-level expression of two different humanized mAbs has recently been reported by Trill et al. (4).

Enzymes. Recently, Laubach et al. (110) used the GS system to achieve high-level expression of the human inducible nitric oxide synthase (iNOS) in CHO-K1 cells. The iNOS expression plasmid, a derivative of pEE14 vector (Celltech Ltd., U.K.) was transfected into CHO cells using Lipofectamine™ and selected with 25 μM MSX in the presence of ethylisothiourea (EITU), an inhibitor of iNOS (to reduce toxicity due to NO production). Further amplification of initially selected clones in the presence of 400 μM MSX gave stable clones with activity > 3 nmole min^{-1} mg^{-1}. Sodium butyrate treatment of cells increased the level of expression by another 1.5-fold.

Receptors. A coding sequence for the soluble form of CD4 (111) was constructed through the introduction of a translation termination codon at the extracellular boundary of the transmembrane domain, and subcloning the cDNA downstream of the SV40 early promoter and upstream of the bovine growth hormone poly A addition signal. This expression cassette was linked to a mouse DHFR expression cassette, and the resultant plasmid was transfected into DHFR$^-$ CHO DXB-11 cells. The initially selected DHFR$^+$ clones were exposed to increasing concentration of MTX, and the expression level in several clones was >3pg cell^{-1} day^{-1}. This specific productivity translated into an expression level of $\sim$40 mg of sCD4 L^{-1} under high-density suspension culture conditions.

Newman-Tancredi et al. (112) reported high-level expression of recombinant 5-HT1A receptors in CHO cells using a derivative of pSVL (Pharmacia). DHFR-CHO cells were cotransfected with pSV2dhfr (113) using the $CaPO_4$ method. Out of 35 DHFR$^+$ clones tested for expression, 33 had receptor expression at 20–3,000 fmol mg^{-1}, as determined by ^{3}H-8OH-DPAT binding assays. The highest-expressing clone (expressing at 2.8 pmol receptor mg protein^{-1}) was identified, and functional coupling was demonstrated through the action of a series of 5-hydroxytryptamine agonists that inhibited forskoline-stimulated adenyl cyclase (AC) activity. The receptor density and receptor regulation are maintained in suspension grown cells, but the inhibition of AC by the receptor is lost gradually (probably due to a change in G-protein–AC interaction and not due to a change in receptor–G-protein coupling).

The GS system has also been utilized to express a 7TM receptor for human thyrotropin (TSH) in CHO cells (114). Cell lines expressing up to 78,000 receptors, cell^{-1}, which efficiently couple to AC, have been identified.

Adehesion Molecules. The soluble form of human E-selectin (sE-selectin) has been expressed at a high level in DHFR$^-$ CHO cells (115). The sE-selectin cDNA was subcloned into the RLDN expression vector (in which expression of sE-selectin was under the control of RSV promoter) and transfected into CHO cells. The G418^r clones expressing sE-selectin were further selected stepwise in media containing 10 nM and 100 nM MTX. A 100 nM MTX^r clone was identified. In spinner flasks it produced 100 mg L^{-1} of sE-selectin over a 5-day period.

Chemokines. Liao et al. (116) expressed recombinant human Mig, a CXC chemokine, after subcloning the Mig cDNA into the vector pMSXND. Expression of Mig was under the control of the Mtn I promoter. The expression vector contained the SV40 small t-antigen intron and poly A addition signal, a DHFR expression cassette under the control of the SV40 early promoter, and a *neo*^{r} under the control of the tk promoter. The construct was transfected into CHO cells (proline$^-$), and the *neo*^{r} clones were subjected to amplification using 0.2 μM MTX. The MTX^r clones were induced with $CdSO_4$, and the conditioned media screened by SDS-PAGE showed a collection of immunoreactive polypeptides (due to protein processing of the high molecular weight Mig), which were similar to that secreted from IFN-treated peripheral blood monocytes and THP-1 cells.

Other Examples. Cockett et al. (117) developed the GS system for high-level expression of heterologous genes in CHO cells. Transfection of an expression vector containing a tissue inhibitor of metalloproteinases (TIMP) cDNA (under the transcriptional control of the hCMV promoter) and a GS mini gene (under the control of the SV40 late promoter) into CHO-K1 cells (followed by selection with 20 μM

MSX) yielded clones expressing up to 9 μg TIMP 10^6 cells^{-1} day^{-1}. Further selection with higher concentration of MSX (500 μM) produced a clone expressing TIMP at 110 μg 10^6 cells^{-1} day^{-1}.

Yonemura et al. (118) have developed a high-level expression system for blood coagulation factor VIII (consisting of a 90-kDa H-chain and an 80-kDa L-chain). The expression plasmids contained a chicken β-actin/rabbit β-globin hybrid promoter (119) and a DHFR transcription unit under the control of the SV40 early promoter. DHFR$^-$ CHO cells (DG44) were cotransfected with H- and L-chain expression vectors. After selecting for DHFR$^+$ clones (using nucleoside-free media), cells that tested positive in a precoagulation assay were cultured in increasing MTX concentrations (upto 1,000 nM) for gene amplification. A clone resistant to 1,000 nM MTX was established that produced 15 IU 10^6 cells^{-1} day^{-1} of factor VIII activity.

Specific Examples of Myeloma Cell Expression

Monoclonal Antibodies. High-level production of mAbs have been reported by several groups (9–12). For example, Gillies et al. (9) expressed two chimeric mAbs, Ch14.18 and ChB72.3, in Sp2/0 cells. H-chain V-region cDNA and a hybrid human–mouse L-chain V-region cDNA were fused to the their respective genomic constant regions. The H-chain expression unit (under the control of the mouse H-chain enhancer/metallothionein MT I promoter) was linked to a DHFR expression cassette. The L-chain expression unit (under the control of a synthetic murine K-promoter) was subcloned into the H-chain expression plasmid to generate a single vector. The expression vector was transfected into Sp2/0 cells, and the MTX^r clones secreting antibody were adapted to increasing concentration of MTX. Subclones of 5 μM MTX^r clones had expression levels of 120–180 μg ml^{-1}. Shitara et al. (12) constructed a "tandem vector" in which both H- and L-chain cDNA expression units (both under the control of Moloney virus LTR), DHFR, and Neo^r expression units were linked. Expression from this vector was compared in three host cells: mouse Sp2/0, P3X63Ag8U.1, and rat YB2/0 cells. YB2/0 cells gave higher numbers of G418^r clones and higher levels of mAb expression. Subsequent amplifications with 50–200 nM MTX yielded several clones secreting 70–100 μg of mAb 10^6 cells^{-1} day^{-1}.

The GS system has been widely utilized for antibody expression in myeloma cells. Bebbington et al. (10) reported high-level expression of a chimeric antibody in NSO myeloma cells. An expression vector was created after placing in tandem a GS expression cassette, the H-chain transcription unit, and the L-chain transcription unit. The expression of the GS gene was under the control of the SV40 early promoter, while the expression of both H-chain and L-chain was under the transcriptional control of hCMV promoter. Transfection of this expression plasmid into NSO myeloma cells, followed by selection in glutamine free medium, gave a pool of clones expressing up to ~4 pg antibody cell^{-1} day^{-1}. Further selection with 20–100 μM MSX gave rise to the highest-expressing clone, with a specific productivity of about 10 pg cell^{-1} day^{-1}. Southern blot analysis showed that this cell line had approximately four copies of the vector per cell.

CONCLUSIONS

As described in this review, a large number of expression systems are available for heterologous expression of a wide variety of proteins in mammalian cells. The choice of a particular host/vector system depends primarily on the desired expression level, nature of expression (transient vs. stable), intracellular versus extracellular expression, constitutive versus inducibile expression, and so forth. Using the expression systems described, it is now possible to produce therapeutic proteins such as mAbs at very high levels (~1–2 gm L^{-1}). However, the challenge still remains to produce such therapeutic proteins in an even more economical manner. Therefore, alternate expression systems still need to be developed to *(1)* achieve even higher levels of expression and *(3)* reduce cycle time (the time interval between construction of an expression vector to production of the protein of interest in large scale). Further understanding of cellular processes such as transcription, translation, and secretion will be essential to achieve such a goal.

NOMENCLATURE

AC	Adenyl cyclase
ADA	adenosine deaminase
BHK	Baby hamster kidney
$CaPO_4$	Calcium phosphate
CAT	Chloramphenicol acetyl transferase
$CdSO_4$	Cadmium sulfate
CGRP	Calcitonin-gene-related peptide
CHO	Chinese hamster ovary
DHFR	Dihydrofolate reductase
DI	Dicistronic intron
EBV	Epstein–Barr virus
EcR	Ecdysone receptor
EMC	Encephalomyocarditis
GH	Growth hormone
GR	Glucocorticoid receptor
GRE	Glucocorticoid response element
hCMV	Human cytomegalo virus
HEK	Human embryonic kidney
hsp	Heat shock protein
HSV	Herpes simplex virus
IE	Immediate early
IFN	Interferon
IgG	Immunoglobulin
IPTG	Isopropyl β-thiogalactopyranoside
IRES	Internal ribosome entry site
L	Light
LTR	Long terminal repeat
mAB	Monoclonal antibody
MCP	Monocyte chemotactic protein
MLP	Major late promoter
MMTV	Mouse mammary tumor virus

MSX	Methionine sulfoxamine
MTN	Metallothionein
MTX	methotrexate
MuLV	Murine leukemia virus
MurA	Muristerone
NOS	Nitric oxide synthase
8-OH-DPAT-2-(*N*,*N*-dipropylamino)-8-hydroxy-1,2,3,4-tetrahydronapthalate	
ORF	Open reading frame
PTH	Parathyroid hormone
RSV	Rous sarcoma virus
SV	Simian virus
TGF	Transforming growth factor
TIMP	Tissue inhibitor of metalloproteinase
TM	Transmembrane
tPA	Tissue plasminogen activator
TPO	Thrombopoietin
UTR	Untranslated region
VEGF	Vascular endothelial growth factor

BIBLIOGRAPHY

1. A.D. Levinson, *Methods Enzymol.* **185**, 485–487 (1990).
2. M. Kriegler, *Gene Transfer and Expression: A Laboratory Manual*, Stockton Press, New York, 1990.
3. S.E. Kane, in J.K. Setlow ed., *Genetic Engineering*, vol. 13, Plenum, New York, 1991, pp. 167–182.
4. J.J. Trill, A.R. Shatzman, and S. Ganguly, *Curr. Opin. Biotechnol.* **6**, 553–560 (1995).
5. G. Gross and H. Hauser. *J. Biotechnol.* **41**, 91–110 (1995).
6. G. Urlaub, E. Kas, A.M. Carothers, and L.A. Chasin. *Cell* **33**, 405–412 (1983).
7. G. Urlaub and L.A. Chasin, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 4216–4220 (1980).
8. R.J. Kaufman, *Methods Enzymol.* **185**, 537–566 (1990).
9. S.D. Gillies, K-M. Lo, and J. Wesolowski, *J. Immunol. Methods* **125**, 191–202 (1989).
10. C.R. Bebbington, G. Renner, S. Thomson, D. King, D. Abrams, and G.T. Yarronton, *Bio/Technology* **10**, 169–175 (1992).
11. M.A. Walls, K.-c. Hsiao, and L.J. Harris, *Nucleic Acids Res.* **21**, 2921–2929 (1993).
12. K. Shitara, K. Nakamura, Y. Tokutake-Tanaka, M. Fukushima, and N. Hanai, *J. Immunol. Methods* **167**, 271–278 (1994).
13. R.J. Owens and R.J. Young, *J. Immunol. Methods* **168**, 149–165 (1994).
14. C.M. Gorman, D. Gies, G. McCray, and M. Huang, *Virology* **171**, 377–385, (1989).
15. F. Gotzes and A. Baumann, *Biochem. Biophys. Res. Commun.* **222**, 121–126 (1996).
16. M. Pines, S. Fukayama, K. Costas, E. Meurer, P.K. Goldsmith, X. Xu, S. Muallem, V. Behar, M. Chorev, M. Rosenblatt, A.H. Tashijan, Jr., and L.J. Suva, *Bone* **18**, 381–389 (1996).
17. L.M. Tolbert and J. Lameh, *J. Biol. Chem.* **271**, 17335–17342 (1996).
18. N. Aiyar, K. Rand, N.A. Elshourbagy, Z. Zeng, J.E. Adamou, D.J. Bergsma, and Y.A. Li, *J. Biol. Chem.* **271**, 11325–11329 (1996).
19. S. Tippmer, B. Bossenmaier, and H. Haring, *Eur. J. Biochem.* **236**, 953–959 (1996).
20. C. Combadiere, S.K. Ahuja, H.L. Tiffany, and P.M. Murphy, *J. Leukoc. Biol.* **60**, 147–152 (1996).
21. T. Kurihara and R. Bravo, *J. Biol. Chem.* **271**, 11603–11606 (1996).
22. J-L. Gao, A.I. Sen, M. Kitaura, O. Yoshie, M.E. Rothenberg, P.M. Murphy, and A.D. Luster, *Biochem. Biophys. Res. Commun.* **223**, 679–684 (1996).
23. A. Ben-Baruch, L. Xu, P.R. Young, K. Bengali, J.J. Oppenheim, and J.M. Wang, *J. Biol. Chem.* **270**, 22123–22128 (1995).
24. S. Yamagami, Y. Tokuda, K. Ishii, H. Tanaka, and N. Endo. *Biochem. Biophys. Res. Commun.* **202**, 1156–1162 (1994).
25. P. Martasek, Q. Liu, J. Liu, L.J. Roman, S.S. Gross, W.C. Sessa, and B.S.S. Masters, *Biochem. Biophys. Res. Commun.* **219**, 359–365 (1996).
26. T. Hatae, S. Hara, C. Yokoyama, T. Yabuki, H. Inoue, V. Ullrich, and T. Tanabe, *FEBS Lett.* **389**, 268–272 (1996).
27. R.M. Pitti, S.A. Marsters, M. Haak-Frendscho, G.C. Osaka, J. Mordenti, S.M. Chamow, and A. Ashkenazi, *Mol. Immunol.* **31**, 1345–1351 (1994)
28. R.G. Clark, D. Mortensen, D. Reifsynder, M. Mohler, T. Etcheverry, and V. Mukku, *Growth Regul.* **3**, 50–52 (1993)
29. A.J. Pelletier, S.C. Bodary, and A.D. Levinson, *Mol. Biol. Cell.* **3**, 989–998 (1992)
30. L.R. Poborsky, B.M. Fendly, K.L. Fisher, R.M. Lawn, B.J. Marks, G. McCray, K.M. Tate, G.A. Vehar, and C.M. Gorman, *Protein Eng.* **3**, 547–553 (1990).
31. P. Mellon, V. Parker, Y. Gluzman, and T. Maniatis, *Cell* **27**, 279–288 (1981).
32. S.B. Patel and S.M. Grundy, *J. Lipid Res.* **36**, 2090–2103 (1995).
33. S.P.S. Svensson, T.J. Bailey, A.C. Porter, J.G. Richman, and J.W. Regan, *Biochem. Pharmacol.* **51**, 291–300 (1996).
34. T. MacNeil, K.K. Bierilo, J.G. Menke, and J.F. Hess, *Biochim. Biophys. Acta* **1264**, 223–228 (1995).
35. R. Weinander, E. Mosialou, J. DeJong, C-P.D. Tu, J. Dypbukt, T. Bergman, H.J. Barnes, J-O. Hoog, and R. Morgenstern, *Biochem J.* **311**, 861–866 (1995).
36. K. Denner, J. Doehmer, and R. Bernhardt, *Endocrinol. Res.* **21**, 443–448 (1995).
37. M.I. Cockett, C.R. Bebbington, and G.T. Yarronton, *Nucleic Acids Res.* **19**, 319–325 (1991).
38. E. Bender, J.M. Woof, J.D. Atkin, M.D. Barker, C.R. Bebbington, and D.R. Burton, *Hum. Antibod. Hybridomas* **4**, 74–79 (1993).
39. H. Dorai, J.E. McCartney, R.M. Hudziak, M-S. Tai, A.A. Laminet, L.L. Houston, J.S. Huston, and H. Oppermann, *Bio/Technology* **12**, 890–897 (1994).
40. P.J. Hippenmeyer, A.M. Rankin, B.A. Reitz, D.R. McWilliams, B.B. Brightwell, R.A. Wolfe and T.G. Warren, *Mol. Cell. Endocrinol.* **107**, 155–164 (1995).
41. P.J. Hippenmeyer and M. Highkin, *Bio/Technology* **11**, 1037–1041 (1993).
42. P.J. Hippenmeyer and L.E. Pegg, *Curr. Opin. Biotechnol.* **6**, 548–552 (1995).
43. R.F. Margolskee, P. Kavathas, and P. Berg, *Mol. Cell. Biol.* **8**, 2837–2847 (1988).

44. G. Cachianes, C. Ho, R.F. Weber, S.R. Williams, D.V. Goeddel, and D.W. Leung, *Biotechniques* **15**, 255–259 (1993).
45. A. Lieber, V. Sandig, W. Sommer, S. Bahring, and M. Strauss, *Methods Enzymol.* **217**, 47–66 (1993).
46. R.J. Kaufman, *Methods Enzymol.* **185**, 487–511 (1990).
47. M. Kriegler, *Methods Enzymol.* **185**, 512–527 (1990).
48. Y. Takebe, M. Seiki, J-I. Fujisawa, P. Hoy, K. Yokota, K-I. Arai, M. Yoshida, and N. Arai, *Mol. Cell. Biol.* **8**, 466–472 (1988).
49. C.M. Gorman, D.R. Gies, and G. McCray, *DNA Protein Eng. Techn.* **2**, 3–10 (1990).
50. A.R. Buchman and P. Berg, *Mol. Cell. Biol.* **8**, 4395–4405 (1988).
51. L.A. Fouser, S.L. Swanberg, B-Y. Lin, M. Benedict, K. Kelleher, D.A. Cumming, and G.E. Riedel, *Bio / Technology* **10**, 1121–1127 (1992).
52. M.T.F. Huang and C.M. Gorman, *Mol. Cell. Biol.* **10**, 1805–1810 (1990).
53. M. Kozak, *J. Cell Biol.* **115**, 887–903 (1991).
54. C-L. Ho, S.S.M. Chin, K. Carnevale, and R.K.H. Liem, *Eur. J. Cell Biol.* **68**, 103–112 (1995).
55. R. Turner and G.D. Foster, *Mol. Biotecnol.* **3**, 225–236 (1995).
56. W.C. Merrick, *Microbiol. Rev.* **56**, 291–315 (1992).
57. V.M. Pain, *Eur. J. Biochem.* **236**, 747–771 (1996).
58. W. Sheay, S. Nelson, I. Martinez, T.-H.T. Chu, S. Bhatia, and R. Dornburg, *Biotechniques*, **15**, 856–862 (1993).
59. Br. Pat. GB2284210-A (November 25, 1994), British Technology Group, Ltd.
60. D. Falcone and D.W. Andrews, *Mol. Cell. Biol.* **11**, 2656–2664 (1991).
61. R.E. Kellems, *Curr. Opin. Biotechnol.* **2**, 723–729 (1991).
62. C.R. Bebbington and C.C.G. Hentschel, in D. Glover ed., *DNA Cloning*, vol. III, Academic, New York, 1987, pp. 163–188.
63. W.A. Keown, C.R. Campbell, and R.S. Kucherlapati, *Methods Enzymol.* **185**, 527–537 (1990).
64. J. Barsoum, in J.A. Nickoloff ed., *Methods in Molecular Biology*, vol. 48 (*Animal Cell Electroporation and Electrofusion Protocols*), Humana Press, Totowa, New Jersey, 1995, pp. 225–237.
65. E. Wagner, M. Cotten, R. Foisner, and M.L. Birnstiel, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 4255–4259 (1991).
66. M. Ogata, K-I. Wakita, K. Kimura, Y. Marumoto, K. Oh.-i., and S. Shimizu, *Appl. Microbiol. Biotechnol.* **38**, 520–525 (1993).
67. C. Gandor, C. Leist, A. Fiechter, and F.A.M. Asselbergs, *FEBS Lett.* **377**, 290–294 (1995).
68. A.S. Lubiniecki, K. Anumula, J. Callaway, J. L'Italien, M. Oka, B. Okita, G. Wasserman, D. Zabriskie, R. Arathoon, S. Builder, R. Garnick, M. Wiebe, and J. Browne, *Dev. Biol. Stand.* **76**, 105–115 (1992).
69. P.G. Zaworski and G.S. Gill, *Biotechniques*, **15**, 863–866 (1993).
70. T.A. Bibila and D.K. Robinson, *Biotechnol. Prog.* **11**, 1–13 (1995).
71. M. Reff, *Curr. Opin. Biotechnol*, **4**, 573–576 (1993).
72. R.B. Parekh, R.A. Dwek, J.R. Thomas, G. Opdenakker, T.W. Rademacher, A.J. Wittwer, S.C. Howard, R. Nelson, N.R. Siegel, M.G. Jennings, N.K. Harakas, and J. Feder, *Biochemistry* **28**, 7644–7662 (1989).
73. R.B. Parekh, R.A. Dwek, P.M. Rudd, J.R. Thomas, T.W. Rademacher, T. Warren, T.-C. Wun, B. Hebert, B. Reitz, M. Palmier, T. Ramabhadran, and D.C. Tiemeier, *Biochemistry* **28**, 7670–7679 (1989).
74. C.F. Goochee, M.J. Gramer, D.C. Andersen, J.B. Bahr, and J.R. Rasmussen, *Bio / Technology* **9**, 1347–1355 (1991).
75. M. Borys, D.I.H. Linzer, and E.T. Papoutsakis, *Bio / Technology*, **11**, 720–724 (1993).
76. E. Watson, B. Shah, L. Leiderman, Y-R. Hsu, S. Karkare, H.S. Lu, and F-K. Lin, *Biotechnol Prog.* **10**, 39–44 (1994).
77. G.T. Yarronton, *Curr. Opin. Biotechnol.* **3**, 506–511 (1992).
78. M.C.-T. Hu and N. Davidson , *Cell* **48**, 555–566 (1987).
79. M.A. Labow, S.B. Baim, T. Shenk, and A.J. Levine, *Mol. Cell. Biol.* **10**, 3343–3356 (1990).
80. M. Gossen and H. Bujard, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5547–5551 (1992).
81. M. Gossen, A.L. Bonin, S. Freundlieb, and H. Bujard, *Curr. Opin. Biotechnol.* **5**, 516–520 (1994).
82. M. Gossen and H. Bujard, *Nucleic Acids Res.* **21**, 4411–4412 (1993).
83. M. Gossen, S. Freundlieb, G. Bender, G. Muller, W. Hillen, and H. Bujard, *Science*, **268**, 1766–1769 (1995).
84. Z. Lang and J.M. Feingold, *Gene*, **168**, 169–171 (1996).
85. J.R. Howe, B.V. Skryabin, S.M. Belcher, C.A. Zerillo, and C. Schmauss, *J. Biol. Chem.* **270**, 14168–14174 (1995).
86. K.S. Christopherson, M.R. Mark, V. Bajaj, and P.J. Godowski, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6314–6318 (1992).
87. T-P. Yao, B.M. Forman, Z. Jiang, L. Cherbas, J.-D. Chen, M. McKeown, P. Cherbas, and R.M. Evans, *Nature* **366**, 476–479 (1993)
88. D. No, T-P. Yao, and R.M. Evans, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 3346–3351 (1996).
89. R.J. Kaufman, M.V. Davies, L.C. Wasley, and D. Michnick, *Nucleic Acids Res.* **19**, 4485–4490 (1991).
90. W. Dirks, M. Wirth, and H. Hauser, *Gene* **128**, 247–249 (1993).
91. D.L. Cadena and G.N. Gill, *Protein Expr. Purif.* **4**, 177–186 (1993).
92. M.V. Davies and R.J. Kaufman, *J. Virol.* **66**, 1924–1932 (1992).
93. M. Kobayashi, Y. Yamauchi, A. Tanaka, and S. Shimamura, *Biotechniques* **21**, 398–402 (1996).
94. S. Rees, J. Coote, J. Stables, S. Goodson, S. Harris, and M.G. Lee *Biotechniques*, **20**, 102–110 (1996).
95. B.K. Lucas, L.M. Giere, R.A. DeMarco, A. Shen, V. Chisholm, and C.W. Crowley. *Nucleic Acids. Res.* **24**, 1774–1779 (1996).
96. L.J. Suva, G.A. Winslow, R.E.H. Wettenhall, R.G. Hammonds, J.M. Moseley, H. Diefenbach-Jagger, C.P. Rodda, B.E. Kemp, H. Rodriguez, E.Y. Chen, P.J. Hudson, T.J. Martin, and W.I. Wood, *Science* **237**, 893–896 (1987).
97. L. Madisen, M.N. Lioubin, H. Marquardt, and A.F. Purchio. *Growth Factors* **3**, 129–138 (1990).
98. L.E. Gentry, N.R. Webb, G.J. Lim, A.M. Brunner, J.E. Ranchalis, D.R. Twardzik, M.N. Lioubin, H. Marquardt, and A.F. Purchio, *Mol. Cell. Biol.* **7**, 3418–3427 (1987).
99. L. Bourdrel, C-H. Lin, S.L. Lauren, R.H. Elmore, B.J. Sugarman, S. Hu, and K.R. Westcott. *Protein Expr. Purif.* **4**, 130–140 (1993).
100. H. Okayama and P. Berg, *Mol. Cell. Biol.* **3**, 280–289 (1983).
101. N. Ferrara, J. Winer, T. Burton, A. Rowland, M. Siegel, H.S. Phillips, T. Terrell, G.A. Keller, and A.D. Levinson, *J. Clin. Invest.* **91**, 160–170 (1993).

102. M. Iwane, T. Watanabe, A. Shintani, Y. Kaisho, S. Matsumoto, R. Sasada, and K. Igarahasi, *Appl. Microbiol. Biotechnol.* **41**, 225–232 (1994).

103. T.L. Burgess, S.L. Ross, Y.-x. Qian, D. Brankow, and S. Hu. *J. Biol. Chem.* **270**, 19188–19196 (1995).

104. Y.A. DeClerck, T-D. Yean, H.S. Lu, J. Ting, and K.E. Langley, *J. Biol. Chem.* **266**, 3893–3899 (1991).

105. M.J. Page and M.A. Sydenham, *Bio/Technology* **9**, 64–68 (1991).

106. M.E. Reff, K. Carner, K.S. Chambers, P.C. Chinn, J.E. Leonard, R. Raab, R.A. Newman, N. Hanna, and D.R. Anderson, *Blood* **83**, 435–445 (1994).

107. R. Newman, J. Alberts, D. Anderson, K. Carner, C. Heard, F. Norton, R. Raab, M. Reff, S. Shuey, and N. Hanna, *Bio/Technology* **10**, 1455–1460 (1992).

108. C.R. Wood, A.J. Dorner, G.E. Morris, E.M. Alderman, D. Wilson, R.M. O'Hara, and R.J. Kaufman, *J. Inmmunol.* **145**, 3011–3016 (1990).

109. H. Tada, T. Kurokawa, T. Seita, T. Watanabe, and S. Iwasa, *J. Biotechnol.* **33**, 157–174 (1994).

110. V.E. Laubach, E.P. Garvey, and P.A. Sherman, *Biochem. Biophys. Res. Commun.* **218**, 802–807 (1996).

111. K.C. Deen, J.S. McDougal, R. Inacker, G. Folena-Wasserman, J. Arthos, J. Rosenberg, P.J. Maddon, R. Axel, and R.W. Sweet, *Nature* **331**, 82–84 (1988).

112. A. Newman-Tancredi, R. Wootton, and P.G. Strange, *Biochem J.* **285**, 933–938 (1992).

113. S. Subramani, R. Mulligan, and P. Berg, *Mol. Cell. Biol.* **1**, 854–864 (1981).

114. E. Harfst and A.P. Johnstone, *Anal. Biochem.* **207**, 80–84 (1992).

115. P. Hensley, P.J. McDevitt, I. Brooks, J.J. Trill, J.A. Feild, D.E. McNulty, J.R. Connor, D.E. Griswold, V. Kumar, K.D. Kopple, S.A. Carr, B.J. Dalton, and K. Johanson. *J. Biol. Chem.* **269**, 1–10 (1994).

116. F. Liao, R.L. Rabin, J.R. Yannelli, L.G. Koniaris, P. Vanguri, and J.M. Farber, *J. Exp. Med.* **182**, 1301–1314 (1995).

117. M.I. Cockett, C.R. Bebbington, and G.T. Yarranton. *Bio/Technology* **8**, 662–667 (1990).

118. H. Yonemura, K. Sugawara, K. Nakashima, Y. Nakahara, T. Hamamoto, I. Mimaki, K. Yokomizo, Y. Tajima, K.-i. Masuda, and A. Imaizumi, A. Funatsu, and J.-i. Miyazaki, *Protein Eng.* **6**, 669–674 (1993).

119. I.-i. Miyazaki, S. Takaki, K. Araki, F. Tashiro, A. Tominaga, K. Takatsu, and K.-i. Yamamura, *Gene* **79**, 269–277 (1989).

See also CHINESE HAMSTER OVARY CELLS, RECOMBINANT PROTEIN PRODUCTION; EXPRESSION SYSTEMS, *E. COLI*; PICHIA, OPTIMIZATION OF PROTEIN EXPRESSION.

FERMENTATION MONITORING, DESIGN AND OPTIMIZATION

JENS NIELSEN
Technical University of Denmark
Lyngby, Denmark

KEY WORDS

Fermentation
High-performance bioreactors
Metabolic control analysis
Metabolic engineering
Metabolic flux analysis
On-line monitoring
Quantitative physiology

OUTLINE

INTRODUCTION

The term *fermentation* is derived from Latin *fermentum* to ferment, and it has been used to describe the metabolism of sugars by microorganisms since ancient times. Thus, fermentation of fruits is so old that ancient Greeks attributed its discovery to one of their gods, Dionysos. Among the classical fermentation processes are beer brewing, which has documented from as early as 3000 B.C. in Babylonia, soya sauce production in Japan and China, and fermented milk beverages in the Balkans and in Central Asia (1). Before World War II, however, fermentation processes mainly found their application in the food area, and it was with the introduction of the penicillin production that large-scale fermentation was first used in the production of pharmaceuticals. Today fermentation processes are used in the production of many different products and these processes can be divided into nine categories, according to the product (Table 1).

Development of fermentation processes can roughly be divided into four phases (see Fig. 1). First the product is identified. In the case of a pharmaceutical this may be a result of random screening for different therapeutic effects by microbial metabolites (e.g., by high throughput screening of secondary metabolites from *Actinomycetes*) or it may be the result of a targeted identification of a novel product (e.g., a peptide hormone with known function). Outside the pharmaceutical sector the product may also be chosen after a random screening procedure (e.g., screening for a novel enzyme to be used in detergents) or it may be chosen in a more rational fashion. With the rapid progress in the different genomic sequencing programs (the yeast genome was completely sequenced in 1996 and the genome of several bacteria have been completely sequenced), there is much focus on identification of novel products, and today topics such as bioinformatics and gene function are frontline research topics (2).

The next step in the development phase is to choose a strain. In the past this choice was normally obvious after the product had been identified; for example, *Penicillium chrysogenum* was chosen for penicillin production because it was the first organism identified to produce penicillin.

Table 1. Categories and Typical Application of Fermentation Processes

Category	Typical processes
Production of microbial cells	Baker's yeast, single cell protein, lactic acid bacteria
Production of primary metabolites	Ethanol (both in beer and wine production and for technical use), lactic acid, citric acid, acetic acid, amino acids, vitamins
Production of secondary metabolites	Antibiotics (penicillins, cephalosporins, clavulanic acid, tetracycline), skikonin
Production of microbial enzymes	Lipases, proteases, amylases, glucose isomerase
Production of pharmaceutical proteins	tPA, human insulin, erythropoietin, growth hormone, hepatitis B vaccine, monoclonal antibodies
Production of polysaccharides	Xanthan gum, dextran
Production of DNA	Genes for vaccines or gene therapy
Biotransformations	Steroids, L-sorbose (from L-sorbitol)
Production of tissue	Bone marrow, skin

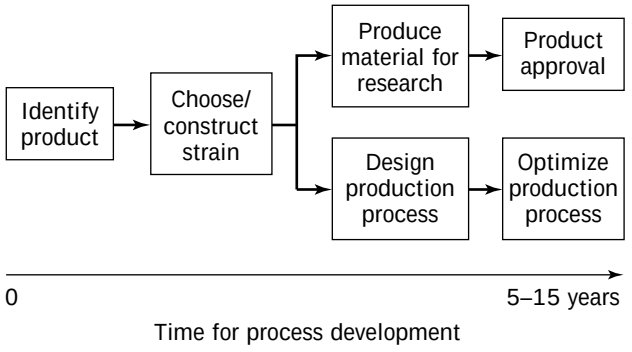


Figure 1. Different phases in the development of a fermentation process.

With the introduction of recombinant DNA (rDNA) technology, it is now possible to choose almost any host for the production. Thus, a strain of *Escherichia coli* has been constructed that can produce ethanol at a high yield (3), and a recombinant strain of *P. chrysogenum* can now be used to produce 7-ADCA (a precursor used for synthesis of cephalosporins) directly by fermentation (4). The choice of strain does, however, often depend much on tradition within the company, and most of the fermentation industries have a set of favorite organisms that are used in the production of many different products. Thus the major players on the industrial enzyme market (Novo Nordisk, Genencor, and Gist-brocades) have chosen a few organisms (typically one to two strains of filamentous fungi and one to two species of bacteria) as production vehicles for a wide range of enzymes. With the production of a heterologous protein, that is, expression of a foreign gene in a given organism, it is also necessary to consider many other aspects (e.g., whether the protein is correctly folded and glycosylated) Table 2 gives an overview of the advantages and disadvantages of different cellular systems for the production of recombinant proteins.

After the strain has been constructed, one of the first aims is to produce sufficient materials for further research, and this is typically done in pilot-plant facilities. For a pharmaceutical compound sufficient material must be produced for clinical trials. For other products it may be necessary to carry out tests of the product and examine any possible toxic effects. In parallel to the continuing research in the application of the product, the production process is also designed. The final steps in development are product approval by the proper authorities and construction of the production facility (which in some cases may be through retrofitting of an existing plant). After the process has been developed, there is continuous optimization of the process, and in some cases better strains are introduced. However, for pharmaceutical products new strains are, only introduced when there is a significant improvement in the process because it is costly to obtain approval from the authorities for redesign of the process.

CRITERIA FOR DESIGN AND OPTIMIZATION

The criteria used for design and optimization of fermentation processes depends on the product. Thus, the criteria used for a high-volume–low-value-added product are normally completely different from the criteria used for a low-volume–high-value-added product. For products belonging to the first category (which includes most whole cell products, most primary metabolites, many secondary metabolites, most industrial enzymes, and most polysaccharides) the three most important design parameters are

- Yield of product on the substrate
- Productivity
- Final titer

A high yield of product on the substrate is for these processes very important, since this gives a good utilization of the raw materials, which often accounts for a significant part of the total costs. Thus, in penicillin production the costs of glucose alone may account for up to 15% of the total production costs (5). Productivity is important because it ensures an efficient utilization of the production capacity (i.e., the bioreactors). In an increasing market it is especially important to increase productivity, because this may prevent new capital investments. Final titer is of importance for the further treatment of the fermentation medium (e.g., purification of the product). Thus, if the product is present in a very low concentration at the end of the fermentation, it may be very expensive to perform purification of the product.

In the production of novel pharmaceuticals, which typically belong to the category of low-volume–high-value-added products, the three previously mentioned design parameters are normally not very important. For these processes quickness to market is generally much more important, and change of the process after implementation is often complicated due to the required U.S. Food and Drug Administration (FDA) approval. In the initial design phase

Table 2. Advantages and Disadvantages of Different Hosts for Production of Recombinant

Host	Advantages	Disadvantages
Bacteria (*E. coli*)	Wide choice of cloning vectors	Posttranslational modifications lacking
	Gene expression easy to control	High endotoxin content
	Large yields possible	Protein aggregation (inclusion bodies)
	Good protein secretion	
Yeast (*S. cerevisiae*)	Generally regarded as safe (GRAS)	Less cloning vectors available
	No pathogens for humans	Glycosylation not identical to mammalian glycosylation
	Large-scale production established	Genetics less understood
	Some posttranslational modifications possible	
Filamentous fungi	Experience with large-scale production	High level of heterologous protein expression has not been achieved
	Source of many industrial enzymes	Genetics not well characterized
	Excellent protein secreters	
Mammalian cells	Same biological activity as natural protein	Cells difficult to grow in bioreactors
	Expression vectors available	Expensive
		Slow growth
		Low productivity
Cultured insect cells	High level of gene expression possible	Not always 100% active proteins
	Posttranlational modification possible	Mechanisms largely unknown

it is, however, still important to keep these three design criteria in mind. The requirement for a high final titer is especially important since the cost for purification (or downstream processing) often accounts for more than 90% of the total production costs.

STRAIN CONSTRUCTION AND STRAIN IMPROVEMENT

Today high-performance production strains to be used in fermentation processes are often constructed using rDNA technology. However, the concept of metabolic pathway manipulation is by no means new. Thus, for decades better strains of *Saccharomyces* to be used for beer fermentation have been obtained through classical breeding and crossing of different strains, and in the production of penicillin productivity has increased by more than 500 times through repeated rounds of mutation and selection of new strains of *P. chrysogenum*. With the introduction of rDNA technology it has become possible to apply a more rational approach to strain improvement–namely by the introduction of targeted genetic changes resulting in strains with a phenotype that gives a better process. This rational approach has been named cellular and metabolic engineering, and different definitions of the approach have been given (6). The term *metabolic engineering* was first introduced by Bailey (7), who defined it as "improvement of cellular activities by manipulation of enzymatic transport and regulatory functions of the cell with the use of recombinant DNA technology." His definition of metabolic engineering includes the following:

1. Inserting new pathways in microorganisms with the aim of producing novel metabolites (e.g., production of polyketides by *Streptomyces* [8,9]), with the aim of degrading toxic compounds (e.g., in bioremediation), or with the aim of constructing a novel biotransformation system
2. Production of heterologous peptides, such as production of human insulin, erythropoitin, and tPA, or industrial enzymes, such as lipases and cellulases
3. Improvement of pathway fluxes leading to higher yields of metabolites (e.g., increasing the flux toward antibiotics or primary metabolites) or yields of biomass (e.g., increasing the cell mass yield in baker's yeast production using industrial media containing mixed sugars [10])

What characterizes metabolic engineering is the rational approach to performing genetic changes, and as with all other fields of engineering it consists of two steps: analysis and synthesis (11). As a consequence of the difficulties in performing detailed analysis of cellular metabolism there has mainly been focus on synthesis in the past, such as expression of new genes in various host cells, amplification of endogenous enzymes, and deletion of genes or modulation of enzyme activities. With modern experimental techniques it has, become possible, however, to perform detailed analysis of cellular function through both in vivo and in vitro measurements, and in the following some of these techniques are described.

MODERN EXPERIMENTAL TECHNIQUES

Whereas the development in rDNA technology has enabled the design of novel high-performance strains that are tailor-made to specific needs, the revolution in computer technology and in analytical chemistry has allowed a much more fundamental characterization of cellular function than previously possible. Thus, flow cytometry allows quantification of population dynamics, while advances in microscopy and image analysis have allowed studies of cell growth and function at the single cell level (12). In studies of fermentation processes it is important to study the function of whole cultures, and here the development of high-performance bioreactors has been an important contribution. When microbial or cell cultures are grown in these bioreactors, it is ensured that all the cells experience the same conditions, and it is therefore possible to perform quantitative physiological studies, such as study of growth and production kinetics, study of gene expression under well-controlled conditions, quantification of metabolic fluxes, and study of the control of flux through the different cellular pathways.

High-Performance Bioreactors

High-performance bioreactors are characterized as being practically ideal bioreactors with a very low mixing time and a very high gas–liquid mass transfer (13,14). In these bioreactors the cells are subjected to an unchanging environment when they are circulated throughout the liquid medium, and the response to imposed variations in the environment is therefore a consequence of the cellular behavior only. High-performance bioreactors are equipped with a large number of in situ sensors, and to control operating variables a flexible direct digital control (DDC) is used, rather than classical single-purpose controllers. This allows precise control of many operating variables, including temperature, pH, dilution rate, and dissolved oxygen concentration, which therefore can be taken to be culture parameters (Table 3) (15). Many commercially available laboratory bioreactors (volumes less than 10 L) normally fulfill the requirements for being high-performance bioreactors. Several bioreactor companies offer systems with flexible DDCs that enable precise control of the culture variables, and when these bioreactors are designed with

Table 3. In Situ Sensors Used for Measuring Culture Parameters

Culture parameter	Sensor	Range	Accuracy
Temperature	Pt-100	0–150 °C	0.1 °C
Pressure	Piezoresistor	0–2 bar	20 mbar
Gas flow	Thermal mass flow meter	0–20 L min^{-1}	20 mL min^{-1}
pH	pH electrode	2–12	0.02
pO_2	Polarografic Clark electrode	0–400 mbar	2 mbar
pCO_2	Membrane-covered pH electrode	0–100 mbar	2 mbar

the proper stirrer and sparger it is possible to ensure a very low mixing time (less than 1 s) and a high gas–liquid mass transfer (k_1a values above 0.1 s^{-1}). In addition to sensors for the culture parameters, modern high-performance bioreactors are normally equipped with in situ sensors and on-line analyzers that allow frequent monitoring of key culture variables. Figure 2 illustrates a typical high-performance bioreactor system. The bioreactor is equipped with on-line flow injection analyzers for monitoring of important culture variables, exhaust gas analysis, and an automatic sampling system. Hereby the most important culture variables can be measured at a high frequency (30–1,000 h^{-1}), whereas other culture parameters can be measured in the automatically withdrawn samples at a lower frequency (1–3 h^{-1}). The bioreactor system can be operated as a constant mass, continuous culture with measurement of the feed flow. It is therefore possible to vary the dilution rate (with a precision of 0.005 h^{-1}), and through proper design of the feed medium it is possible to operate the bioreactor such that there is a single limiting substrate component, such as glucose or ammonia. Finally, the bioreactor is equipped with gas blending, which allows variations in the composition of the inlet gas to the bioreactor, so the influence of, for example, the oxygen concentration on the growth and production kinetics can be studied at the same specific growth rate (16).

With advanced computer control of culture parameters it is possible to perform very reproducible experiment, which is of paramount importance for a fundamental understanding of the underlying cellular reactions. This is illustrated in Figure 3, where the carbon dioxide evolution rate is shown for two different batch cultivations of the filamentous fungi *P. chrysogenum*. With remarkable reproducibility the CER signal can obviously be used to map the different phases of the batch fermentation, which were carried out with sucrose as the carbon source. In the first phase sucrose is hydrolyzed to glucose and fructose and glucose is metabolized. When glucose is exhausted after 45 h there is a distinct shift in the metabolism, and when fructose is exhausted after 60 h there is another shift in the metabolism (i.e., the CER decreases rapidly). Even when both sugars are exhausted there is still some cellular activity due to metabolism of gluconic acid, which was formed in the initial phase during growth on glucose (17).

In situ Sensors. An in situ sensor can be inserted directly into the bioreactor and can be sterilized in place. Table 3 gives the characteristics of in situ sensors used to monitor culture parameters in a typical high-performance bioreactor. Most of these sensors are extremely reliable and are used routinely not only in research laboratories but also in industrial plants.

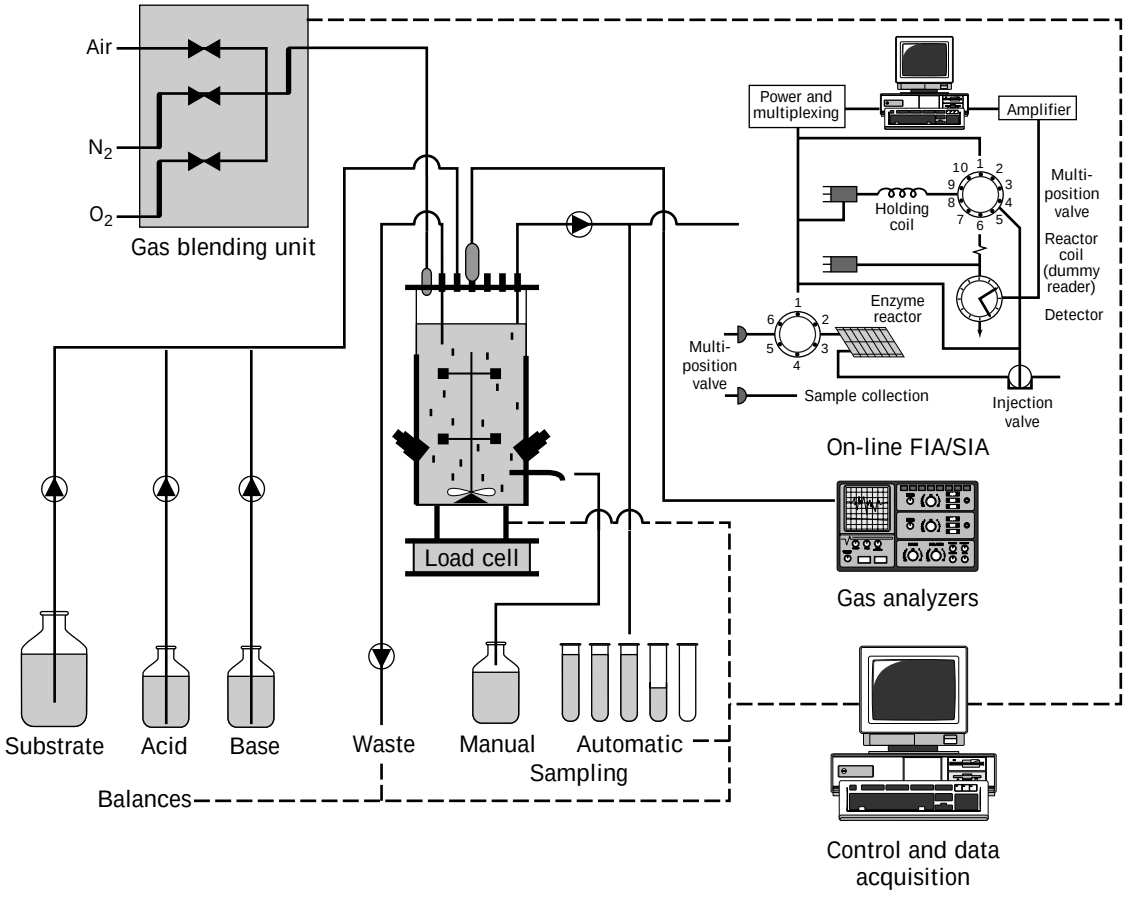


Figure 2. Typical high-performance bioreactor system.

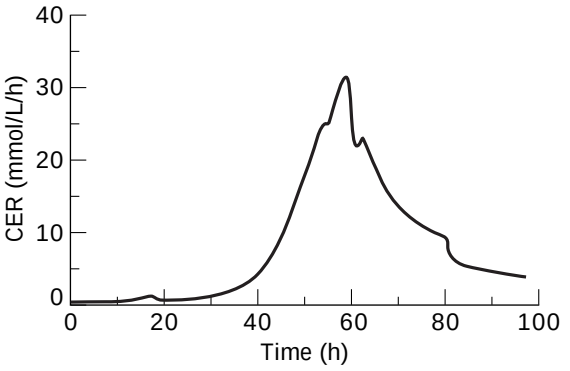


Figure 3. On-line measurement of the carbon dioxide evolution rate (CER) during two consecutive batch cultivations of *P. chrysogenum*. The batch cultures were inoculated with spores at time zero. *Source*: From Ref. 5.

Table 4. Advantages and Disadvantages of Different Measurement Principles for In Situ Biomass Sensors

Principle	Advantages	Disadvantages
Optical density	Wide linear range	Some interference
Culture fluorescence	High sensitivity	Much interference
	Measurement of cellular activity	Signal interpretation difficult
Capacitance	Wide measurement range	Interferance from aeration and agitation
	Measurement of cellular activity	Signal interpretation difficult
Ultrasonic	Wide linear range	Interference from aeration and agitation
	Self-cleaning	Temperature sensitive

For monitoring of culture variables there are, however, only a few commercially available in situ sensors available, and besides a few sensors based on enzymatic analysis of medium components (e.g., for glucose analysis [18]), these sensors measure biomass using different measurement principles (see Table 4) (19). There are several commercially available sensors for in situ monitoring of the optical density, which is linearly correlated with the biomass concentration. These sensors have a wide linear range, but there is interference from gas bubbles and solids in the medium. In some sensors special designs are used to reduce this interference such as by passing the sample into an internal measurement chamber (19). Sensors measuring the culture fluorescence (normally the NAD(P)H-related fluorescence) were first introduced as biomass sensors, but despite high sensitivity and correlation of the culture fluorescence with the total cellular activity, the signal from these sensors is difficult to interpret, and there is generally a poor correlation with the biomass concentration (17,20). Measurement of dielectric properties (or capacitance) is one of the later contributions to measurement of biomass, and it has the advantage of a wide measurement range. However, the signal depends on the cellular activity, which although it is an interesting variable, may not necessarily correlate with the biomass concentration. Even though none of the commercially available in situ biomass sensors stands out as the best biomass sensor, several of the sensors are very useful additions to the sensors normally applied in high-performance bioreactors.

On-Line Analyzers. With the lack of reliable in situ sensors for measurement of important culture variables, including glucose and other medium components, on-line analysis is frequently used. Here a sample is automatically withdrawn and analyzed. For this purpose flow injection analysis (FIA) is well suited due to its high speed, good precision, and good reliability (21). The major drawback of FIA is that it is a single component analysis, and even though a novel technique termed *sequential injection analysis* (SIA) offers the possibility to measure several components using the same hardware (22), other analytical systems such as mass spectrometry (MS), high-performance liquid chromatography (HPLC), and gas chromatography (GC) may sometimes be preferred. These analytical systems offer the ability to measure many components in a single run, but, in general, it is difficult to match the high measurement frequency and reliability of FIA.

A prerequisite for on-line monitoring is automatic sampling, and the sampling system normally serves as the sterile barrier between the bioreactor and the analytical system. Automatic sampling can be performed either by withdrawing a sample directly or by using a membrane module. In direct sample withdrawal, a nonfiltered sample is pumped directly from the bioreactor through a syringe or catheter. Direct sample withdrawal is a very simple solution that allows for on-line measurement of the biomass concentration using FIA (23) and measurement of intracellular components (24). However, a construction with a syringe is not sufficiently robust to be used in an industrial environment. For sampling via a membrane module there are several commercially available designs (25), which can be divided into two groups:

1. *Membrane modules placed in a recycle loop connected to the bioreactor*. The advantage of these modules is that it is possible to change the membrane while the process is running. However, the continuous pumping of the medium through the membrane loop may have a significant effect on the cellular behavior (especially for shear-sensitive organisms).
2. In situ *membrane modules, which are inserted into the bioreactor and sterilized in place*. The advantage of these modules is that the modules do not influence the bioreactor operation and the withdrawn sample is a good representation of the medium in the bioreactor. However, to ensure mechanical breakage of the membrane it is necessary to apply rather thick membranes, which influences the response time (26).

With a reliable membrane module for automatic sampling it is possible to perform routine on-line monitoring of important medium components, and if FIA is used it is especially possible to obtain high frequency measurements, which allows precise quantification of the fluxes in and out

of the cells (e.g., the substrate uptake rates and the product formation rates). These fluxes form the basis for quantitative physiological research, as discussed later, and high-performance bioreactors equipped with on-line analyzers for the most important culture variables are therefore indispensable tools in such research.

Measurement of Intracellular Variables. For quantitative physiological studies, the measurement of intracellular variables is extremely important. These variables should ideally be performed in vivo, while cellular function is maintained, but presently there are no techniques that allow for precise in vivo measurements of intracellular variables; therefore they are taken in vitro. To ensure that the in vitro measurements are good representatives of the in vivo conditions, it is critical to have techniques that allow for rapid sampling and quenching of the cellular activity. For the measurement of intracellular metabolites and cofactors, such as glycolytic intermediates, ATP, and NADH, which have a turnover time of a few seconds, it is extremely important to have an almost instantaneous quenching of the cellular metabolism. Using a specially designed sampling module and rapid cooling of the sample in precooled test tubes containing small glass beads, Theobald et al. (27) were able to measure the intracellular level of many metabolites of the EMP pathway and the cofactors ATP, ADP, and AMP in *S. cerevisiae*. Furthermore, using a freeze–quench method, where the sample is mixed with a reagent and then quenched by pumping it into a reservoir of liquid methanol at -40 °C, de Koning and van Dam studied the uptake kinetics of glucose by *S. cerevisiae* at the scale of 5 ms to 1 s (28). Using these sampling techniques, together with HPLC and enzymatic assays, it is possible to obtain a fundamental understanding of the metabolite levels inside the cells at different growth conditions, which, together with knowledge of the metabolic fluxes (discussed later), will give the ultimate physiological characterization.

Quantification of Morphology. For filamentous microorganisms there is in many cases a relation between morphology and product formation. In the past, detailed studies of the growth mechanisms of these organisms have been hampered by the lack of good experimental techniques. However with the rapid development in image analysis it has become possible to study the growth of these organisms in much detail (29). Yet, because of the many different processes that influence morphology, it is important to combine experimental work and mathematical modeling to extract information about the underlying mechanisms. When the morphology is measured for a given number of hyphal elements in a sample taken from a fermentation with filamentous microorganisms, there will be a certain distribution of properties, governed by the underlying mechanisms for growth and hyphal breakage (30). It is difficult to extract information about the different mechanisms from these measurements directly. However, using small growth chambers, where the developing morphology can be monitored on-line, it is possible to obtain fundamental information about the growth mechanisms (12). When this information is combined with computer simulations of hyphal element populations, it is possible to obtain some insight into the mechanisms behind hyphal breakage during submerged fermentations (31).

Quantitative Physiological Studies

Despite many years of extensive analysis of glycolysis, all the details in regulation of glycolytic flux have not been elucidated. Even in *S. cerevisiae*, where studies of a large number of mutants in the glycolysis have resulted in a meticulous mapping of regulation patterns, all the control structures have probably not been revealed. Thus, it was only a few years ago that fructose-2,6-bisphosphate was discovered not only as an important regulator of phosphofructokinase but of other enzymes as well. Further physiological studies are therefore extremely important to identify yet-unknown regulatory patterns and regulatory compounds. With the recent advances in genome characterization it is now possible to rapidly construct many different mutants by introduction of gene disruptions or overexpression of key enzymes, and this is a good basis for further studies. However, it is important to combine genetic work with quantitative physiological studies in which gene regulation is studied under well-controlled growth conditions, and metabolic fluxes and their control are quantified.

Study of Gene Regulation at Well-Controlled Growth Conditions. To illustrate how gene regulation can be studied in submerged fermentations carried out in high-performance bioreactors, two examples are considered. The first example is regulation of the strong TAKA (or α-amlyase) promoter of *Aspergillus oryzae*, which is often used as expression promoter for production of industrial enzymes by filamentous fungi. In the past there has been much speculation on the regulation of this promoter, but most studies were based on shake-flask experiments, in which it is difficult to distinguish between repression and induction. However, through the use of continuous cultures it can clearly be demonstrated that there is both glucose repression and maltose induction (32,33). Thus, by adding a glucose pulse to a steady-state chemostat operated at a low dilution rate, when the glucose concentration is very low (about 5 mg L^{-1}; i.e., derepressed conditions), it was found that there is an immediate decrease in the α-amylase synthesis (Fig. 4). This clearly indicates that there is glucose repression. Furthermore, by shifting from a glucose-based medium to a maltose-based medium when the cells were grown under derepressed conditions (i.e., at a low glucose concentration) it was possible to show that there is maltose induction (33). Besides allowing qualitative conclusions to be drawn, measurements of the glucose concentration at the milligram per liter level even allowed quantification of the degree of glucose repression (32). These findings are consistent with mapping of the TAKA promoter, which indicates the presence of several *creA* binding sites (creAp is a regulatory protein associated with glucose repression in filamentous fungi) and a binding site for a protein that is induced by starch (or maltose) (34). In a classical approach to analyze the promoter further, one would construct promoter fragments and identify the function of the individual binding sites. In this analysis shake-

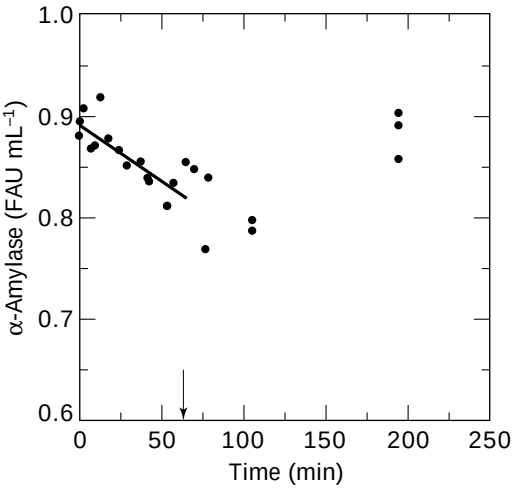


Figure 4. Glucose repression of the TAKA promoter in *A. oryzae*. At time zero a glucose pulse (resulting in a concentration of 600 mg L^{-1}) is added to a steady-state chemostat where the glucose concentration is low (i.e., derepressed conditions). Right after the glucose pulse is added the production of α-amylase decreases instantaneously to the level observed at high glucose concentrations. The α-amylase concentration (given as units per milliliter) therefore decreases due to washout from the bioreactor. When the glucose pulse is exhausted (indicated by the arrow) there is derepression and α-amylase production increases. *Source*: From Ref. 32.

flask cultures may be useful in drawing qualitative conclusions (i.e., whether glucose repression has been eliminated). However, to draw quantitative conclusions these experiments should be followed by another set of continuous fermentation experiments.

The other example is on glucose control in the yeast *S. cerevisiae*. In the production and application of baker's yeast, glucose repression is a serious problem. An often-used carbon source is molasses, which consists of glucose, fructose, sucrose, and raffinose (the trisaccharide galactose–glucose–fructose), and in the uptake and metabolism of the different sugars glucose repression results in lag phases. Furthermore, in the application of baker's yeast in break making, maltose is the typical carbon source and glucose repression of maltose uptake and metabolism results in a reduction in the gassing power. One approach to alleviate glucose repression is to delete the genes coding for key regulatory proteins, for example, *MIG1*, a general repressor that influences the transcription of a number of genes, including the *SUC*, the *GAL*, and the *MAL* genes. Analysis of the effect of deletion of *MIG1* can only be done properly using well-controlled fermentation experiments, and the effect of *MIG1* deletion on sucrose and maltose utilization has been studied using batch cultures (10,35). Two different approaches were applied to silence *MIG1*: *(1)* gene disruption and *(2)* introduction of high copy number plasmids with antisense fragments of *MIG1*. Use of antisense is especially interesting in silencing genes in polyploid industrial yeast strains, but even though the antisense fragments were expressed, there was no change in the degree of glucose repression (10) (Table 5). However,

Table 5. Invertase Activity in Different Haploid Laboratory Strains of *Saccharomyces cerevisiae* during Growth on Glucose[a]

Strain	Genetic construct	Invertase activity [U (g DW)$^{-1}$]
X2180-1A	Wild-type	61
TA8	Antisense, plasmid	55
TA25	Antisense, chromosomal	55
T301	Disruption	263

Note: The antisense fragments were inserted either in a high-copy plasmid or integrated into the chromosome.
Source: From Ref. 10.

with gene disruption of *MIG1* a significantly higher activity of invertase is obtained during glucose consumption (Table 5), resulting in an almost parallel uptake of glucose and sucrose (10), as seen in Figure 5. Disruption of *MIG1* was also found to partly alleviate glucose repression on maltose uptake and metabolism in laboratory strains; there was an almost parallel uptake of glucose and maltose, whereas there was no effect of *MIG1* disruption on maltose uptake in a polyploid industrial strain (35). This was explained to be a consequence of strong catabolite inactivation of the maltose permease in the industrial strain (35). Based on the physiological characterization of the recombinant strains it was concluded that silencing of *MIG1* may alleviate glucose repression, but other effects including catabolite inactivation, also have an important influence. Only through careful physiological characterizations was it possible to identify the different effects.

Quantification of Metabolic Fluxes. In connection with optimization of metabolite production, where the aim is to direct as much carbon as possible from the substrate into the metabolic product to ensure a high yield, it is of prime importance to quantify how the carbon fluxes are distributed through the various cellular pathways. A very powerful technique for calculation of the fluxes through various pathways is the so-called metabolic flux analysis, where the intracellular fluxes are calculated using a stoichiometric model for all the major intracellular reactions and applying mass balances around the intracellular metabolites. A set of measured fluxes as input in the calculations is used, and these fluxes typically are the uptake rates of substrates and secretion rates of metabolites (36–38). If one performs experiments with ^{13}C-enriched carbon sources and measures the fractional enrichment of ^{13}C in intracellular metabolites, it is possible to apply an additional set of constraints in the model, and a better estimation of the intracellular fluxes may be obtained (39). Finally, if one applies information about the isotopomer distributions for the intracellular metabolites much more information is supplied, and it may even be possible to quantify the net fluxes through reactions that are practically reversible (40).

Besides quantification of the various pathway fluxes, and therefore determination of the carbon flows inside the cell, metabolic flux analysis is useful for the following:

- *Identification of possible rigid branch points (or nodes) in the cellular pathways*. Through comparison

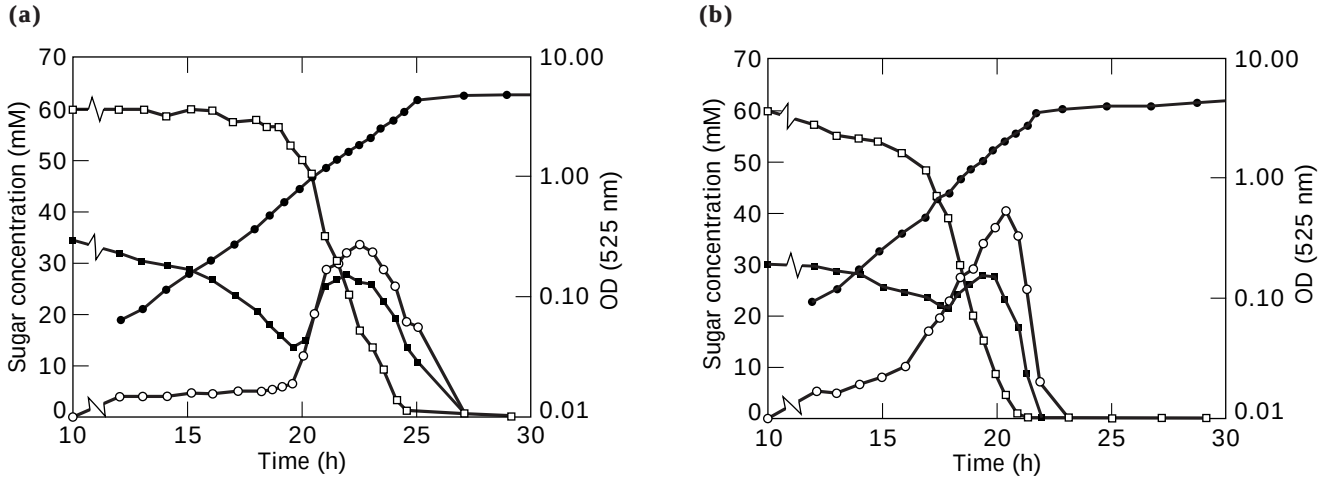


Figure 5. Alleviation of glucose repression by disruption of *MIG1* in *S. cerevisiae*. Results are shown for two batch fermentations carried out on a glucose–sucrose mixture: (**a**) with a wild-type haploid strain and (**b**) with the same strain where *MIG1* has been disrupted. Obviously disruption of *MIG1* results in metabolism of sucrose at a higher glucose concentration as compared with the wild-type strain (i.e., the sucrose concentration begins to decrease right at the start of the fermentation). □, sucrose concentration (mM); ●, biomass concentration (OD); ■, glucose concentration (mM); ○, fructose concentration (mM). *Source*: From Ref. 10.

of the distribution of fluxes at different operating conditions and in different mutants it is possible to identify whether a pathway node is rigid or flexible (41). Thus through metabolic flux analysis of various mutants of *C. glutamicum*, Vallino and Stephanopoulos (42,43) concluded that in lysine production the glucose-6-phosphate node is flexible and the pyruvate node is weakly rigid.

- *Identification of the existence of different pathways*. Formulation of the reaction stoichiometry, which is the basis for metabolic flux analysis, requires detailed information of the biochemistry. However, for many microorganisms certain details in the pathway stoichiometries are unknown, and it may not be known whether a given pathway is active. Furthermore, in many cells there are several isoenzymes whose functions are not known in complete details. By calculating the metabolic fluxes with a different set of cellular pathways it may be possible to identify which pathway is most likely to exist or to obtain indications of the functions of different isoenzymes and/or pathways. Thus, in analysis of anaerobic growth of *S. cerevisiae*, we found that alcohol dehydrogenase III, a mitochondrial enzyme whose physiological function has not been identified, plays an important role in maintaining the redox level inside the mitochondria (38). However, this type of analysis should always be followed by enzyme assays in which the presence of enzyme activity in cell extracts is tested.
- *Calculation of nonmeasured extracellular fluxes*. Normally the number of fluxes that can be measured is larger than what is needed for calculation of the intracellular fluxes. In this case it is possible to calculate some of the extracellular fluxes, such as the rate of production of various by-products, by use of the stoichiometric model and the measured fluxes.
- *Examination of the influence of alternative pathways on the distribution of fluxes*. In connection with optimization of metabolite production it may be possible to identify one or several constraints for increasing the yield of a metabolite on the substrate or the flux leading to the desired metabolite. In these cases it is possible to envision various scenarios, such as examining whether insertion of a new pathway or an isoenzyme (or perhaps deletion of an isoenzyme) can have a positive effect on removing this constraint, leading to an increased flux toward the desired metabolite. In a study of penicillin production we calculated that the yield of penicillin on glucose is likely to be higher if cysteine, one of the precursors for penicillin production, is synthesized by direct sulfhydrylation rather than by the transsulfuration pathway claimed to exist in *P. chrysogenum* (37).
- *Calculation of maximum theoretical yields*. Based on a stoichiometric model it is normally possible to calculate the maximum theoretical yield of a given metabolite if a given set of constraints is specified. This value is of interest because it specifies an upper limit for the yield in the process. This has been illustrated both for lysine production by *C. glutamicum* (36) and for penicillin production by *P. chrysogenum* (37). In these calculations it is necessary to introduce a constraint, for example, that reactions do not run in thermodynamically unfavorable directions.

Control of Metabolic Fluxes. One of the most important aspects of quantitative physiology is control of flux. As dis-

cussed, the concept of metabolic flux analysis is useful for studies of interactions between different pathways and for quantification of flux distributions around branch points, but it does not allow for evaluation of how the fluxes are controlled (i.e., how the rates of synthesis and conversion of metabolites are kept in close balance over a very wide range of external conditions without catastrophic rises or falls in the metabolite concentrations). The discoveries in the 1950s of feedback inhibition, cooperativity, covalent modification of enzymes, and control of enzyme synthesis introduced a number of molecular effects that may play a role in control of fluxes. With these many different mechanisms it is not surprising that disputes often arises over how the flux through a given pathway is controlled. Since the many reports on enzyme regulation typically are verbal and qualitative (e.g., phosphofructokinase is the major flux-controlling enzyme of glycolysis in muscle), it is difficult to discriminate between different findings, and in many cases different findings may seem conflicting. Furthermore, one often finds terms such as *rate-limiting steps* and *bottleneck enzyme* when control of flux is discussed. Thus, the early findings that it is the first enzyme in a pathway or the first enzyme after a branch that is under some type of control (e.g., by feedback inhibition) often results in statements such as "the first step in a pathway is the rate-limiting step."

The concept of metabolic control analysis (MCA) tells us that these kinds of qualitative statements have no meaning, since flux control is distributed over all the steps in a pathway, but that some steps may exert a higher degree of flux control than others. The concept of MCA was developed from the landmark papers of Kacser and Burns (44) and Heinrich and Rapoport (45). Its basis is a set of parameters, called elasticity coefficients and control coefficients, that quantify the control within a reaction network. The elasticity coefficients are given by

$$\epsilon^{i}_{x_j} = \frac{X_j}{v_i}\frac{\partial v_i}{\partial X_j} \quad (1)$$

and they express the sensitivity of the rates of the enzyme-catalyzed reactions (v_i) toward the metabolite concentrations (X_j). The most often used control coefficients are the flux control coefficients (FCCs), which are given by

$$C_i^{J_j} = \frac{v_i}{J_j}\frac{\partial J_i}{\partial v_i} \quad (2)$$

The FCCs express the fractional change in the steady-state flux through the pathway (J_j) that results from an infinitesimal change in the activity of the enzymes (or reaction rates). The FCCs and the elasticity coefficients are related to each other via the summation theorem, which states that the sum of all the FCCs is 1, and the connectivity theorem, which states that the sum of the product of the elasticity coefficients and the FCCs is zero.

Three different groups of experimental methods are available for determination of the FCCs:

- Direct methods, where the control coefficients are determined directly
- Indirect methods, where the elasticity coefficients are determined and the control coefficients are calculated from the theorems of MCA
- Transient metabolite measurements, where the metabolite concentrations are measured during a tran-

Table 6. Overview of Methods for Determination of Flux Control Coefficients

Method	Procedure	Advantages and disadvantages
	Direct	
Genetic manipulations	Alternate the expressed enzyme activity through genetic manipulation, such as by inserting inducible promoters	Robust method that gives direct answers, but is very laborious
Enzyme titration	Vary the enzyme activity through titration with purified enzymes	Simple and straightforward procedure, but it can only be applied for pathway segments that are completely decoupled from the rest of the cell
Inhibitor titration	Vary the enzyme activity through titration with specific inhibitors	Simple and easy to apply, but requires the existence of specific inhibitors
	Indirect	
Double modulation	Measure the metabolite levels at different environmental conditions and determine the elasticity coefficients by calculation of differentials	Elegant approach, but it requires two independent changes in the metabolite levels, which are difficult to obtain due to the high degree of coupling between intracellular reactions
Single modulation	Similar to double modulation, but it is based on knowledge of one of the elasticity coefficients	More robust than double modulation, but it requires knowledge of one elasticity coefficient
Top down approach	Based on grouping of reactions and then using double modulation, for example	Very useful, but do not directly give all the FCCs of the system
Kinetic models	Direct calculation of the elasticity coefficients from a kinetic model	Robust, but relies on the availability of a reliable kinetic model for the individual enzymes in the pathway

sient and this information is used to determine the control coefficients (46)

Table 6 gives an overview of the different direct and indirect methods.

Whereas the elasticity coefficients are properties of the individual enzymes, the FCCs are properties of the system. The FCCs are therefore not fixed but change with the environmental conditions. This has been illustrated in analysis of the penicillin biosynthetic pathway (47,48). Based on a kinetic model for the enzymes in this pathway, the FCCs were calculated at different stages of fed-batch fermentation, and a drastic shift in the flux control was found. During the first part of the cultivation the flux control was mainly at the first step in the pathway, the formation of the tripeptide LLD-ACV by ACV synthetase (ACVS). Later in the cultivation flux control shifted to the second step in the pathway, the conversion of LLD-ACV to isopenicillin N by isopenicillin N synthetase (IPNS). This shift in flux control is due to intracellular accumulation of LLD-ACV, which is an inhibitor of ACVS. Obviously, it makes no sense to talk about a rate-limiting step or a bottleneck enzyme in this process. Besides the shift in flux control, it is interesting to note that most of the flux control is by the first two steps in the pathway (48). Furthermore, through analysis of the kinetic model it was found that the value of the FCCs depends on the dissolved oxygen concentration, which is a substrate in the IPNS catalyzed reaction (48).

BIBLIOGRAPHY

1. B. Sikyta, *Prog. Ind. Microbiol.* **31**, 13–32 (1995).
2. A. Goffeau et al., *Science* **274**, 546–567 (1996).
3. K. Ohta, D.S. Beall, J.P. Mejia, K.T. Shanmugam, and L.O. Ingram, *Appl. Environ. Microbiol.* **57**, 893–900 (1991).
4. L. Crawford, A.M. Stepan, P.C. McAda, J.A. Rambosek, M.J. Conder, V.A. Vinci, and C.D. Reeves, *Biotechnology* **13**, 58–62 (1995).
5. J. Nielsen, *Physiological Engineering Aspects* of Penicillium chrysogenum, World Scientific Publishing, Singapore, 1997.
6. D.C. Cameron and I.-T. Tong, *Appl. Biochem. Biotechnol.* **38**, 105–140 (1993).
7. J.E. Bailey, *Science* **252**, 1668–1674 (1991).
8. R. McDaniel, S. Ebert-Khosla, D.A. Hopwood, and C. Khosla, *Science* **262**, 1546–1550 (1993).
9. C.R. Hutchinson, *Biotechnology* **12**, 375–380 (1994).
10. L. Olsson, M.E. Larsen, B. Rønnov, J.D. Mikkelsen, and J. Nielsen, *Appl. Environ. Microbiol.* **63**, 2366–2371 (1997).
11. G. Stephanopoulos, *Curr. Opinion Biotechnol.* **5**, 196–200 (1994).
12. A.B. Spohr, C.D. Mikkelsen, M. Carlsen, J. Nielsen, and J. Villadsen, *Biotechnol. Bioeng.* **58**, 541–553 (1998).
13. B. Sonnleitner and A. Fiechter, *Anal. Chim. Acta* **213**, 199–205 (1988).
14. J. Nielsen and J. Villadsen, in G. Stephanopoulos ed., *Biotechnology*, vol. 3, 2nd ed., VCH, Weinheim, Germany, 1993, pp. 77–104.
15. B. Sonnleitner and A. Fiechter, *Adv. Biochem. Eng./Biotechnol.* **46**, 141–159 (1992).
16. C.M. Henriksen, J. Nielsen, and J. Villadsen, *Biotechnol. Prog.*, submitted (1997).
17. J. Nielsen, C.L. Johansen, and J. Villadsen, *J. Biotechnol.* **38**, 51–62 (1994).
18. J. Bradley, W. Stöcklein, and R.D. Schmid, *Proc. Control Qual.* **1**, 157–183 (1989).
19. L. Olsson and J. Nielsen, *Trends Biotechnol.*, submitted (1997).
20. G. Lidén, *Pure Appl. Chem.* **65**, 1927–1932 (1993).
21. J. Nielsen, *Proc. Control Qual.* **2**, 371–384 (1992).
22. R.W. Min, J. Nielsen, and J. Villadsen, *Anal. Chim. Acta* **320**, 199–205 (1996).
23. S. Benthin, J. Nielsen, and J. Villadsen, *Anal. Chim. Acta.* **261**, 145–153 (1992).
24. H.-A . Kracke-Helm, L. Brandes, B. Hitzmann, U. Rinas, and K. Schügerl, *J. Biotechnol.* **20**, 95–104 (1991).
25. K. Schügerl, in K. Schügerl ed., *Biotechnology*, vol. 4, 2nd, ed., VCH, Weinheim, Germany, 1991, pp. 149–178.
26. L.H. Christensen, J. Nielsen, and J. Villadsen, *Chem. Eng. Sci.* **46**, 3304–3307 (1991).
27. U. Theobald, W. Mailinger, M. Reuss, and M. Rizzi, *Anal. Biochem.* **214**, 31–37 (1993).
28. W. de Koning and K. van Dam, *Anal. Biochem.* **204**, 118–123 (1992).
29. C.R. Thomas and G.C. Paul, *Curr. Opin. Biotechnol.* **7**, 35–45 (1996).
30. J. Nielsen, *Trends Biotechnol.* **14**, 438–443 (1996).
31. P. Krabben, J. Nielsen, and M. Michelsen, *Chem. Eng. Sci.* **52**, 2641–2652 (1997).
32. M. Carlsen, J. Nielsen, and J. Villadsen, *J. Biotechnol.* **45**, 81–93 (1996).
33. R. Mørkeberg, M. Carlsen, and J. Nielsen, *Microbiol.* **141**, 2449–2454 (1995).
34. N. Tsukagoshi, M. Kato, A. Aoyama, F. Naruse, T. Kawaguchi, and T. Kobayashi, *presented at 3 European Conference on Fungal Genetics*, Münster, March 27–30, 1996.
35. C.J.L. Klein, L. Olsson, B. Rønnov, J.D. Mikkelsen, and J. Nielsen, *Appl. Environ. Microbiol.* **62**, 4441–4449 (1996).
36. J.J. Vallino and G. Stephanopoulos, *Biotechnol. Bioeng.* **41**, 633–646 (1993).
37. H.S. Jørgensen, J. Nielsen, J. Villadsen, and H. Møllgaard, *Biotechnol. Bioeng.* **46**, 117–131 (1995).
38. T.L. Nissen, U. Schulze, J. Nielsen, and J. Villadsen, *Microbiology* **143**, 203–214 (1997).
39. A. Marx, A.A. de Graaf, W. Wiechert, L. Eggeling, and H. Sahm, *Biotechnol. Bioeng.* **49**, 111–129 (1996).
40. K. Schmidt, M. Carlsen, J. Nielsen, and J. Villadsen, *Biotechnol. Bioeng.* **58**, 831–840 (1997).
41. G. Stephanopoulos and J.J. Vallino, *Science* **252**, 1675–1681 (1991).
42. J.J. Vallino and G. Stephanopoulos, *Biotechnol. Prog.* **10**, 320–326 (1994).
43. J.J. Vallino and G. Stephanopoulos, *Biotechnol. Prog.* **10**, 327–334 (1994).
44. H. Kacser and J.A. Burns, *Symp. Soc. Exp. Biol.* **27**, 65–104 (1973).
45. R. Heinrich and T.A. Rapoport, *Eur. J. Biochem.* **42**, 89–95 (1974).
46. J. Delgado and J.C. Liao, *Biochem. J.* **282**, 919–927 (1992).
47. J. Nielsen and H.S. Jørgensen, *Biotechnol. Prog.* **11**, 299–305 (1995).
48. P.N. Pissarra, J. Nielsen, and M.J. Bazin, *Biotechnol. Bioeng.* **51**, 168–176 (1996).

See also AMINO ACIDS, PRODUCTION PROCESSES; AMMONIA TOXICITY, ANIMAL CELLS; BIOREACTORS, AIR-LIFT REACTORS; BIOREACTORS, CONTINUOUS STIRRED-TANK REACTORS; BIOREACTORS, FLUIDIZED-BED; SECONDARY METABOLITE PRODUCTION, ACTINOMYCETES, OTHER THAN STREPTOMYCES.

FERMENTER DESIGN

MARVIN CHARLES
Lehigh University
Bethlehem, Pennsylvania

JACK WILSON
ABEC, Inc.
Allentown, Pennsylvania

KEY WORDS

Bulk mixing
Clean-in-place
Heat transfer
Impellers
Oxygen transfer
Power input
Sampling systems
Seals
Sterilization
Stirred tank

OUTLINE

INTRODUCTION

Scope

The scope of this article is limited to design of agitated vessels used for aerobic, single-cell (bacteria and yeast), and filamentous (bacteria and fungi) fermentations. It presents general design principles, combining basic concepts with practical matters including regulatory compliance, safety, maintenance, cleanability, ease of use, and cost, all of which are strongly interrelated. It focuses primarily on fermenters used to produce human and animal health products or their precursors, but the general approaches discussed are applicable to other types of products (e.g., industrial enzymes, food products); the primary differences are emphasized.

Included in the presentation are fermenter vessels; agitation; aeration; heat transfer; sterilization; cleaning; and piping systems. Control and data acquisition are discussed only to the extent that they influence the items already noted.

It is assumed that the reader is familiar with the basics of fermentation, general fermenter construction, and nomenclature. Those not having such background are advised to consult basic references (1–3).

Design Philosophy

Our basic philosophy comprises the following simple principles:

1. Each design case has unique requirements and calls for individualized application of the basic design concepts and practices discussed herein. The one-size-fits-all approach usually results in sleeves that drape into your food, arm holes that cut off circulation, pant legs that are made to trip over, or some combination of these features and others. The magic number approach—all problems have simple solutions based on codified numbers such as fixed geometric ratios (4)—leads to the same place.
2. Successful design requires a systems approach. A fermenter is a system that is part of a process system, and the process is part of a plant system. All these systems interact with each other, with real people, with control systems, and with other systems related to regulatory compliance, safety, documentation (including protocols, SOPs, etc.), change control, maintenance, and so forth. Failure to take these

interactions into account during fermenter design usually results in considerable pain, not only for those guilty of the omission, but also for innocents who were never asked for planning and/or design input, but who must live with the result. Note that appropriate documentation, protocols, SOPs, change control, and so on should be considered important elements of design and operation of any plant, licensed or not.

3. Compromise is always necessary. Nothing in any project is 100% right, and nothing has to be. He or she who looks for 100% of anything will hold up a project needlessly and will generate a lot of animus. (The words actually used in such cases cannot be used here.)
4. The time to be thinking about all of these points is prior to and during design, not when you're standing on the plant floor trying to validate a continuous mixer that has only one port and no instrumentation.

We discuss next the bases for safety and regulatory considerations, primarily to sensitize readers to these issues early on. They are all too often overlooked during early stages of design, when the focus is on satisfying process requirements (e.g., oxygen transfer). Unfortunately, this frequently leads to important constraints being ignored and hence to designs that require expensive and often cumbersome "fixes."

SAFETY AND REGULATORY COMPLIANCE

All fermenter designs are influenced by safety and regulatory factors, the extent depends on *(1)* the nature of the product and its intended use, *(2)* the process, and *(3)* the nature and location of the facility. For example, even fermentation products not regulated by the U.S. Food and Drug Administration (FDA) must adhere to safety requirements, which in some cases are very strict. Furthermore, no professional design can escape the effects of a long list of code requirements, for example, the ASME pressure vessel code (5). By the same token, not all products that are FDA regulated are equally affected by safety, regulatory, and code requirements. We must consider each case on the basis of its own unique requirements.

Our objective in the rest of this section is to introduce some of the requirements and to provide some rational bases for their implementation, details of which are discussed later. Much of the discussion is directly applicable to fermenters used for production of human or animal health products or for their precursors, but also applies to other cases to varying extents.

Containment: Worker and Community Biosafety

A fermenter must be an integral part of a system designed to insure safety at three levels: product, worker, and community. Many of the methods used to protect the product also serve to protect the plant personnel and the community (e.g., use of closed systems and HEPA filters); however, there are potential points of conflict (6). For example, some containment practices that call for completely welded hard piping to a contained drain line for any condensate that could be exposed to culture fluid could expose the product to drain line contaminants. One resolution of this problem is to use steam locks in all such lines. All such conflicts encountered to date have been resolved, but not always easily or relatively inexpensively.

Insurance of product safety is tied primarily to compliance with FDA (or similar organizations) regulations. This will be considered in the subsection "Product Safety." Here we deal with worker and community safety.

Protection must be provided against potential ill effects of fermentation products (e.g., cytotoxins, allergens), organisms (e.g., pathogens, whether recombinant or not), fermentation byproducts (e.g., pharmacologically active precursors of the active product), or some combination. The practice of providing such protection is called *containment*, which is defined as insuring that deleterious fermentation components can't be transported to any area inside or outside the plant before they have been rendered harmless. To accomplish this, containment must be exercised at several levels (7). We consider in this article only direct or "primary" containment of fermenters; however, one should not lose sight of the fact that the other levels must be considered during fermenter design, if for no other reason than to ensure that the fermenter design and operation will be consistent with the overall containment strategy.

Containment levels have been defined by the National Institutes of Health (NIH) (8), based on the potential danger of organisms. These levels, in increasing order of potential danger are GILSP, BL1-LS, BL2-LS, BL3-LS, and BL4-LS. (BL4-LS must be handled on a case-by-case basis in consultation with the Centers for Disease Control and others. It is, thank goodness, well beyond our scope.) There are other classifications that have been developed in the United Kingdom (9), the European Union (10), and elsewhere, but these do not differ markedly from the U.S. scheme. It is important to know that *(1)* all of these guidelines address what must be accomplished but not how they should be accomplished, and *(2)* there is not complete agreement on what is actually required to implement the guidelines. In addition, the guidelines do not address containment of products. Although many of the considerations for product containment are the same or very similar to those for organism containment, there are some very important differences. Here, we confine our attention to organisms; the interested reader should consult the Refs. 11–14 for more specific information concerning hazardous products. One also should consult U.S. Environmental Protection Agency (EPA) regulations applicable to microorganisms and their products.

Most existing processes use organisms that require either GILSP or BL1-LS containment; nevertheless, many plants are built to satisfy requirements for BL2-LS (actually, somewhere between BL2-LS and BL3-LS). The reasons for this include satisfying FDA requirements (e.g., safety of the product) and the fact that the additional cost is not large and is relatively cheap insurance. There also

is the oft-stated reason that such a route insures greater flexibility for future operation (15). One might add to this the fact that the rules are not well established and that requirements could change substantially between the time of design and the time production begins. It also is worth noting that there is growing interest in using organisms that do, in fact, require BL2-LS and even BL3-LS containment, which will probably focus greater attention on further development of acceptable implementation practices as well as changes in the guidelines. Note that equipment requirements and cost for BL2-LS and BL3-LS are currently about the same (this probably will change); BL3-LS facilities requirements and cost are very much greater than for BL2-LS.

Finally, it is important to note that containment is nothing new. Highly pathogenic organisms have long been used to produce therapeutics and biological warfare components; hence, there is a considerable body of experience dealing with the subject (11,16). There also is considerable guidance to be had from the nuclear industry; nevertheless, the natures, sizes, and large number of new commercial processes, along with new methods, materials, and so on, will force heightened awareness, scrutinity, review, and modernization or change.

Physical Safety

Most physical safety issues are addressed by U.S. Occupational Safety and Health Administration (OSHA) requirements (17) and by various national codes such as the ASME Pressure Vessel Code (5) and the National Electrical Code. There also are many local construction codes that must be satisfied. Among these are various earthquake-resistant construction codes that are also applicable to containment considerations. And then there are the requirements of the final arbitors: the insurance companies.

Product Safety

We focus here on FDA requirements, but the reader is urged to keep in mind that he or she also must consider those of other regulatory agencies and be aware of the unsettling facts that it is not always clear which agency has jurisdiction, and there often are conflicting requirements.

The primary purpose of FDA regulations is to insure product safety, purity, and efficacy. The legal bases for the regulations and their enforcement can be found in the *Code of Federal Regulations* (CFR) (18) and a number of other FDA documents such as *Points to Consider* (19,20), *Inspection Guides* (21,22), and *Guides to Industry* (23,24). Industrial implementation of all the regulations, along with ongoing improvements, is referred to as current good manufacturing practices (cGMP).

There is an endless stream of writings in professional journals, short courses, and such concerning the regulations and cGMP practice. Unfortunately, there is a great deal that is not covered clearly in writing, and there is a constant proliferation of new and/or revised guidelines followed closely by lots of (frequently conflicting) interpretations and opinions. In addition there has been a problem of nonuniform application. For example, regulatory scrutiny has been far stricter for fermenters used to produce active molecules directly than for those used to produce precursors; there has been considerable variation from product to product within a given class; and manufacture of human drugs has been regulated much more rigidly than manufacture of animal drugs. While understandable to some extent, such nonuniformity has caused considerable confusion. These gaps are beginning to be closed, and this may have considerable influence on the design of fermenters—particularly those used to produce precursors and animal drugs.

In simplest terms, cGMP requires that the combination of fermentation process equipment and operating protocols must consistently yield product that meets acceptable specifications and is capable of being converted/purified consistently to final product, meeting approved product specifications. From this general sense of intent it can be argued that failure to have control over the fermentation puts the final product at risk. Few would argue with this general concept; however, there is considerable debate as to interpretation and extent.

Requirements that flow from the need to have control are based on relatively simple ideas:

1. If an organism is subject to environmental, medium, or other conditions outside the range in which it is known to yield product meeting acceptable specifications and capable of being converted/purified consistently to final product meeting approved product specifications, then we have no guarantee that it does not produce other products with which the recovery system can not cope and that can escape detection by the analytical methods in place.
2. If a fermenter becomes contaminated with other microbes, said microbes could produce toxins that could be carried undetected to the final product. It is possible that this could occur without altering the behavior of the process organism. It also is possible that products of the contaminant could cause the process organism to make toxins that could go undetected into the final product. Given these possibilities, plus the fact that it is not possible to prove that the contaminant will always be the same, evidence of contamination is evidence for lack of control.
3. If a fermenter is not cleaned properly, deleterious microbial or nonmicrobial products could remain to contaminate the next batch in such a way that impurities could be carried undetected to the final product. Similar statements can be made about other contaminants introduced via other routes as a result of poor cleaning practices.

Again, it is unlikely that many would argue with these points, but again there is considerable debate as to interpretation and extent.

Requirements 1–3 can be translated to specific requirements for, among other things, controllability and reliability of fermentation conditions, sterilization, and aseptic and cleaning operations. Controllability requires that the fermenter and its subsystems be designed and constructed in ways that make possible control of environ-

mental variables within the operational ranges required. This obviously means that mechanical design must be in harmony with control system design.

Sterilization/aseptic operation translates to *(1)* destroying any microbial contaminants that may be present in any part of the equipment that might contact process fluids and *(2)* insuring that no microbes (other than the production organism) can enter after the equipment has been sterilized. The latter embodies the concept of the "sterile barrier." To these ends, the following apply:

- The fermenter and all its ports and direct attachments must be sterilizable.
- All piping that will contact process fluids (including additives) and/or provide paths into the system (e.g., the air exhaust line) must be sterilizable initially. Some (e.g., sampling lines, addition lines) must also be sterilizable at any time during a fermentation.
- Inlet gases (e.g., air, oxygen) and all additives (e.g., medium components, acid, base, antifoam) must be sterilized before they contact any sterile process piping.
- All penetrations (e.g., drive shaft, probe ports) must be sterilizable.

The preferred sterilization method is automatic sterilization-in-place (SIP) with steam. This is discussed at length later.

Clean operation requires, among other things, the following:

- The system must be designed such that any surface that can contact a process stream can be cleaned consistently to a level that insures that the product will be free of soils resulting from a fermentation.
- Nothing in the system that comes in contact with process fluid can introduce unacceptable and unidentifiable materials.

Note (for example) that 21 CFR 211(Part D) (guidelines for design of equipment for licensed facilities) does not have any specific requirements for fermenter sterilization or cleaning; however, the preceding discussion coupled with the intended use and cGMP concepts makes it difficult to to argue the point. But again, there remain the questions of rational interpretation and extent of application.

Riding along with cGMP is the practice of validation. This is an extensive, expensive, and controversial area for which and about which reams of reams have written. Following is the formal definition of validation:

> Establishing documented evidence that provides a high degree of assurance that a specific process will consistently produce a product meeting its predetermined specifications and quality attributes. (25)

What this means is that not only does the equipment have to be designed, installed, and operated so as to run consistently within well-defined ranges of process variables to consistently yield product that meets acceptable specifications and is capable of being converted/purified consistently to final product, meeting approved product specifications, but that there must be adequate documentation to prove this before product can be sold. In addition, one must demonstrate that this goal can be achieved under so called worst-case conditions; there is often considerable debate as to what this really means (26). It is beyond our scope to consider validation in detail. Suffice it to say that it can translate to checking and testing every component of the equipment and documenting not only its proper function but also its history. The brave of heart are referred to the bulging literature on the subject; Refs. 27–30 provide a start.

All of this means we must take great care to consider *all* the details, such as materials, corrosion issues, machining methods, vendors, welds, and types of steam and water, keeping in mind that all must be considered as part of a system that must interface "seamlessly" with the systems of cGMP and validation. Failure to consider this before and during design is almost always very costly. Design should be done so as to facilitate validation, and subsequently, routine cGMP operation, maintenance, and so forth. "Simple" things such as judicious placement of access valves and piping isolation can go a very long way to ensuring minimum pain and maximum operability. Having said all this, we reiterate that not all fermenters are subject to all these requirements, and even those that are, are subject to them in varying degrees (at least in practice).

DESIGN BASIS AND OTHER GENERAL CONSIDERATIONS

A design basis for fermentation equipment should derive from a facility/process design basis. The latter should include (among other items) the general nature of the facility (e.g., research and development vs. production, single product vs. multiuse); product(s) specifications; regulatory, containment, and other requirements; level of automatic operation; general processing scheme (e.g., batch); nature of individual process steps; productivity, concentrations, and so forth; cleaning requirements; special considerations (e.g., earthquake-proof construction); critical valving and instrumentation; staffing requirements and constraints; architectural and general floor plan constraints; and utilities. Most of these will have some influence on fermenter design—some in more subtle ways than others.

It is difficult to overstate the importance of the initial definition that will derive from the design basis. Obviously, a rational design basis for a fermenter must also be based on fermentation characteristics as well as operating cycle and productivity required (which should derive from the overall process design basis). From these will flow the sizes and number of vessels, and definitions of oxygen transfer, heat transfer, power, and bulk mixing requirements. The design must satisfy these but must also satisfy requirements for regulatory compliance, safety, cleaning, facile operation, and maintenance. All of these factors are highly interactive (e.g., design for oxygen transfer affects design for cleaning); hence, responsible design will almost always require several iterations to ensure the greatest probability of success and to minimize lost time caused by instal-

lation, operation, and various problems. The iterative nature of the design (and, unfortunately, construction) process makes most important the existence of a well-crafted, well-implemented, and well-documented change control process. Finally, as with any engineering project, failure to have a solid basis of design, a complete scope, accurate process flow diagrams and piping and instrumentation diagrams (P&IDs), and accurate process timing will almost certainly result in added cost, lost time, and worse.

PROCESS REQUIREMENTS: BASICS

One of the first steps in fermenter design is translation of process demands to oxygen transfer, heat transfer, bulk mixing, and power requirements—the so-called transport processes (TPs). Sound translation is not straightforward because the TPs interact not only with each other, but also with vessel geometry and other design factors. The nature of each TP and the relationships among them vary with the process requirements. They also vary significantly with the basic nature of the organism and the culture broth. The major factors here are the rheological nature of the broth and the sensitivity of the organism to fluid mechanical forces. In some cases, sensitivity of the organism to temperature and other environmental factors, such as pH, impose tighter constraints (e.g., wall temperature control, uniformity of mixing).

Single-cell organisms (most bacteria and yeasts) tend to tolerate fluid forces very well (there are a few exceptions) and tend to have low-viscosity, Newtonian broths. Mycelial organisms (fungi and some mycelial bacteria such as streptomycetes) tend to be more prone to damage by fluid mechanical forces than are single-cell organisms and tend to have high-viscosity, non-Newtonian broths. Detailed discussion of all these factors is beyond the scope of this article; however, we now discuss a few points concerning rheological behavior, and point to fluid effects wherever appropriate in the remainder of the article.

Broth Rheology and Its Effects on Transport Processes

Newtonian viscosity depends only on composition and temperature; the nature of the fluid motion does not affect it. Non-Newtonian viscosity does depend on the nature of the fluid motion. This dependency is usually expressed in terms of fluid shear rate (a measure of how rapidly fluid velocity changes from one point to another point close by).

There are several types of non-Newtonian behavior (e.g., pseudopalstic, Bingham) that can be described quantitatively by a host of mathematical models. One model used frequently to describe non-Newtonian fermentation broths (31–33) is the so-called power law:

$$\eta = K\gamma^{(n-1)} \tag{1}$$

where η is viscosity, γ is shear rate, and n and K are constants. Such relationships can be useful if one has rheological data for the subject broth and if the design correlations used incorporate rheological properties in a meaningful and accurate way. It is seldom that either of these conditions is satisfied, and almost never that both are satisfied. Also, there is a considerable amount of misleading and/or inaccurate information in the literature concerning broth rheology, its measurement, and its use; the reader is cautioned to tread carefully in this area (4,34). If accurate viscosity information is available, it can be a valuable qualitative and sometimes semiquantitative guide, if interpreted and used properly.

Mass transfer (oxygen), heat transfer, and bulk fluid motion all depend strongly on rheological characteristics. Generally, all are poorer in non-Newtonian than in Newtonian broths; indeed, the rheological characteristics of mycelial broths can and do impose severe constraints. Two of many examples are the following:

1. Heat and oxygen transfer rates tend to be anywhere from 50 to 5% of what they would be for typical (e.g., *Escherichia coli*) bacterial fermentations.
2. Bulk mixing quality (homogeneity) is much poorer than in typical Newtonian broths; therefore, accurate monitoring and control are much more difficult.

Difficulties associated with viscous, non-Newtonian rheology also extend to other areas such as cleaning.

Oxygen Transfer/Aeration

Please note that in this and subsequent sections we include design and operating rules of thumb based on our experience and the experiences of others. They are intended to be helpful guides, not edicts.

Requirements. Oxygen transfer rate (OTR) requirements are usually dictated by conditions during the most active part of the growth phase (other phases require oxygen, but supply rate usually is not as high). The requirement is calculated from

$$\text{OTR} = (Y_{x/o})^{-1}(\mu X)_{\max} \tag{2}$$

where μ is the specific growth rate (h^{-1}), X is the cell mass concentration (g/L dw), and $Y_{x/o}$ is the cell yield coefficient on oxygen (g cells dw/g O_2). Note that the maximum growth rate and maximum cell mass are not always reached simultaneously.

Satisfying the Oxygen Material Balance: Gas Flow and Linear Velocity. The very first thing we must do in satisfying an OTR requirement is to insure that we have a closed oxygen material balance. For most practical operating conditions, the rate of change of oxygen inventory in a fermenter is very small compared to oxygen flows in the inlet and outlet and to the OTR. It also can be demonstrated easily that the total molar flow rate of gas does not change enough to cause any loss of sleep. Given these practical realities, the oxygen material balance is

$$\text{OTR} = (1{,}000 \times 60)\text{F}(y_{\text{in}} - y_{\text{out}})/V_{\text{L}} \tag{3}$$

where, F is the gas flow rate (mol/min), y is the mole fraction of oxygen in the gas, and V_{L} is the liquid volume (L). Note that the usual units for OTR are millimoles per liter

per hour and that the factors of 1,000 and 60 in equation 3 are conversion factors needed to keep unit consistency.

The material balance also can be expressed as

$$F = \text{OTR} \times V_L/(1{,}000 \times 60\ \varepsilon y_{in}) \quad (4)$$

where ε is the oxygen transfer efficiency and y_{in} is the oxygen mole fraction in the inlet gas. Equation 4 is a more useful form because we know that under practical conditions ε will be between 0.15 and 0.35 for typical Newtonian broths, and between 0.05 and 0.15 for typical non-Newtonian broths.

As a general rule, heroic efforts will be required to get OTRs over 300 mmol/(L h) in large fermenters (>5,000 L). Even if the effort is expended and it is successful, it will usually create heat transfer problems (see later) that will require even greater effort to overcome. Our advice is to consider other options very seriously.

The material balance provides information about the gas volumetric flow rate, which often is expressed as the standard flow, v_{std}, in units of standard liters per minute (SLPM):

$$v_{std} = 22.4F \quad (5)$$

This is useful because most flow meters are calibrated in terms of standard flow; however, actual flow is more useful for fermenter design purposes. One problem encountered in calculating the actual flow is that it increases from the bottom to the top of the broth because of pressure change. As a practical matter, however, it is usually adequate to use the average pressure in the tank:

$$v_{act} = v_{std} \times (1/P_{avg})(T_f/298) \quad (6)$$

where v_{act} is the gas flow at average pressure (L/min), P_{avg} is the average pressure (atm, abs), and T_f is the fermentation temperature (K). One might want to revisit this for very tall vessels.

The standard gas flow rate also is expressed in standard gas volumes per liquid volume per minute (VVM).

Among the important information that can be obtained from v_{act} is the gas linear velocity, V_S (cm/min):

$$V_S = 4 \times 1000\ v_{act}/(\pi \times D_t^2) \quad (7)$$

where D_t is the inside diameter of the fermenter. V_S affects mass transfer, deliverable power via the impellers, foaming, gas holdup, and aerosol formation. As a rule of thumb, V_S should be held below 200 cm/min to avoid problems associated with excessive gas holdup, foaming, and aerosol formation. These effects are discussed later.

Mass Transfer. The rate at which oxygen can be transferred is dictated by three factors: the driving force, the resistance to transfer, and the contact area between the gas and liquid phases. This is usually expressed as

$$\text{OTR} = K_g a(\Delta P)_{LM} \quad (8)$$

where K_g is the mass transfer coefficient ([mmole M]/[L h atm]), and a is the interfacial area per unit volume (M^2/M^3). $(\Delta P)_{LM}$ is the log mean pressure driving force, defined as

$$(\Delta P)_{LM} = (P_{O,in} - P_{O,out})/LN[(P_{O,in} - P_O^*)/(P_{O,out} - P_O^*)] \quad (9)$$

where $P_{O,in}$ is the partial pressure of oxygen in the gas inlet (atm, abs), $P_{O,out}$ is the partial pressure of oxygen in the gas outlet (atm, abs), and P_O^* is the partial pressure of oxygen that would be in equilibrium with the dissolved oxygen concentration in the existing liquid (atm, abs). Note that use of the log mean driving force assumes that the liquid is mixed perfectly (homogeneous liquid phase) and that the gas moves in plug flow through the fermenter. These assumptions are acceptable for most low-viscosity Newtonian broths, but they can be quite poor for high-viscosity, non-Newtonian broths. Unfortunately, no useful alternatives have yet been proposed; therefore, extra caution should be exercised in interpreting calculated results for non-Newtonian broths.

Other mass transfer coefficients are used for driving forces other than $(\Delta P)_{LM}$. The most common among these is $k_l a$ (min^{-1}), which is used in conjunction with a dissolved oxygen concentration driving force. The reader is cautioned that $K_g a$ and $k_l a$ are sometimes confused with each other—even in the literature.

A given OTR requirement can be satisfied by various combinations of driving force and mass transfer coefficient. It is important to note, however, the following:

- Factors that influence driving force also can influence $K_g a$.
- Each combination will have different effects on the ultimate design of the vessel, piping system, and support systems, as well as on operational factors related to regulatory compliance cleaning, maintenance, and so on. Such effects are discussed as we proceed.

Driving force is affected by total pressure, gas inlet oxygen mole fraction, total gas flow rate, and the dissolved oxygen concentration. Total pressure and oxygen mole fraction affect the oxygen partial pressure directly. The effect of total gas flow is a bit more subtle. For a given oxygen transfer rate, increasing the total gas (fixed inlet conditions, pressure, etc.) increases the mole fraction of the outlet gas flow, thereby increasing the average mole fraction of oxygen in the gas; hence, the average driving force is increased. This can be seen quantitatively via an oxygen material balance which gives

$$P_{O,out} = P_{O,in} - \text{OTR} \times P/(1{,}000F) \quad (10)$$

where P is the total absolute pressure (atm) and F is the total molar flow (mmol/h) of gas.

The characteristics of K_g and a are complex and not completely understood. Our first clue to this is that $K_g a$ is expressed as a product in terms of empirical correlations (we don't understand enough about them to get reliable quantitative guidance from first principles). Typical correlations have the form

$$K_g a = \delta \times (P_g/V_L)^{\varphi}(V_S)^{\kappa} \quad (11)$$

where P_g is the power delivered to aerated broth (HP), V_L is the unaerated liquid volume (L), and δ, φ, and κ are usually advertised as constants. There also a several variations on this form. A discussion of all the fine points of such correlations is beyond our scope; however, the the reader should take note of the fact that published correlations usually cannot be relied upon to provide accurate predictions. They should be used as qualitative or semiquantitative guides only. Reasons for this caveat include the following:

- They are usually applicable only under conditions at or near the ones used to develop them. Unfortunately, most are developed at scales and under conditions that are not realistic for production.
- They can have pronounced dependence on fermenter geometry.
- They do not scale up well, or at all. The "constants" in equation 11 and its relatives usually aren't constant.

These caveats should be made more emphatically for viscous non-Newtonian broths than for Newtonian broths because non-Newtonian rheological characteristics can have profound effects on $K_g a$ (as well as the other transport properties). Most $K_g a$ correlations do not account for broth rheology. Those that do, usually do so inadequately or in ways valid only for a particular broth. It also is important to note that the broth rheology can vary dramatically during a fermentation and that the rheological characteristics can be affected by the fermentation conditions, ranging from the nature of the inoculum to the history of the agitation speed.

As noted earlier, we can get some qualitative guidance from the form of equation 11. We'll do that, but please keep in mind that the relationships are highly nonlinear, the factors that influence $K_g a$ are dependent on each other in practical operation, and it is not always possible to control variables at the levels you would like. For example, equation 11 predicts that increasing power input per unit liquid volume will increase $K_g a$. This usually turns out to be true in practice if the power can actually be delivered to the fluid and delivered in a manner that will contribute to $K_g a$. Such might not be true in any given case.

The fact that a drive's power rating is 100 hp doesn't mean that 100 hp can be delivered to the broth. Deliverability depends on, among other things, geometry (vessel and agitator), rheology, agitator speed, and gas linear velocity (see later). The effects of all are complex and interactive; the interrelations are more complex for non-Newtonian broths than for Newtonian broths. The manner in which power is delivered depends largely on impeller type and broth rheology. For example, oxygen transfer rates to highly viscous, non-Newtonian mycelial broths have been shown to be larger for A-315 (hydrofoil) impellers than for Rushton (disk turbine) impellers delivering the same power (35). There appears to be no meaningful difference for Newtonian broths.

The bottom line is that one should try to obtain experimentally determined correlations for the subject organism and broth at meaningful scales. Failing that (which usually is the case because of time and resource constraints), one should try to use correlations developed for systems as similar as possible to the one at hand. Considerable art is involved here.

The reader can find more about $K_g a$ in the literature (35–38).

Power

Power delivered to the broth is used for micromixing and gas dispersion, which are related to mass transfer, and macromixing, which provides overall homogeneity (discussed later). Agitated vessel power is delivered via two mechanisms: direct mechanical power from the impellers and gas expansion. The bulk (>90%) of the power comes from the impellers as long as the fluid motion is under their control, a condition that prevails so long as the impellers are not flooded. There are reasonably reliable correlations available to determine flooding conditions for Newtonian broths (39); however, flooding usually is not a problem under typical conditions used in most Newtonian broth fermentation. Flooding is more likely in highly viscous, pseudoplastic broths typical of mycelial and polysaccharide fermentations because the viscosity near the impeller is much lower than in the rest of the broth. As a result (assuming air is introduced under the impeller), air tends to channel toward the impeller, thereby enshrouding it and decreasing the deliverable power (4). This can dramatically decrease overall mass transfer (and heat transfer) rates and quality of bulk mixing quality (see later). Reliable flooding correlations for non-Newtonian broths have not been published, but some companies have accumulated a considerable amount of data for the fermentations they practice.

Calculation of power input relies on empirical correlations. As with $K_g a$, many correlations have been published (see for example Refs. 40–42), but they are not particularly reliable for many of the same reasons given for the unreliability of $K_g a$ correlations; therefore, we give the same advice here as we did for $K_g a$.

Among the more popular approaches used is the one that relies on the aeration number, N_A (dimensionless), defined as

$$N_A = v/ND_i^3 \quad (12)$$

where v is the gas volumetric flow rate (m^3/min), N is the agitation speed (min^{-1}), and D_i is the impeller diameter (m). Published correlations give the ratio of gassed power to ungassed power as empirical functions of N_A:

$$P_g/P_{ug} = \text{Func}(N_A) \quad (13)$$

where P_{ug} is the power delivered to the same broth agitated under the same conditions but unaerated.

Each such correlation pertains to a specific impeller type, a single fermenter geometry, and a fairly narrow range of operating conditions. In some such correlations

(40), a unique function is presented for each impeller type (Fig. 1). In others the relationship is shown not to be unique (43). In most cases, the correlations are for a single aerated impeller. The situation is more confusing for non-Newtonian broths.

Obviously, to use equation 13 one must be able to calculate the ungassed power. This can be done via other empirical correlations that give ungassed power in terms of the dimensionless power number (N_P), defined as

$$N_P = P_{ug} \times g_c/N^3 \times D_i^5 \times \rho \qquad (14)$$

where g_c is Newton's law conversion factor, and ρ is the ungassed broth density (lb/ft^3). There are published correlations that give N_P as a function of Reynolds number for a wide range of single impellers (44). Each is specific to a particular vessel geometry, liquid height, and so forth. Interestingly enough, there are fairly reliable correlations for Newtonian and non-Newtonian fluids so long as the constraints of the correlations are observed.

It is important to note that some impellers that have high power numbers under unaerated conditions can "unload" considerably under aerated conditions. For example, it is fairly typical for a Rushton turbine to deliver, under aerated conditions, only 40% of the power it can deliver when unaerated. On the other hand, the SCABA 6SRGT (45), which is a curved-blade disc turbine, does not tend to unload at all. Some hydrofoils also exhibit very little unloading. This is an important consideration when one is trying to rank impellers in terms of how much power they can deliver for oxygen transfer purposes. (Often the ranking tends to be done qualitatively or "semiquantitatively.")

Multiimpeller Systems. Most fermenters are equipped with more than one impeller. The reasons for this are to improve bulk mixing (see later) and power distribution and to avoid the need for very large impellers and/or very high agitator speeds. Impeller size is limited practically not only by vessel internals but also by the following:

- *Torque transmitted to the drive shaft*. The larger the torque, the stronger and thicker the shaft must be. This also translates to higher torque and more expensive gear boxes.
- *The size of the vessel manway*. This is particularly important if the impellers must be single piece for better cleaning/sterilization characteristics.

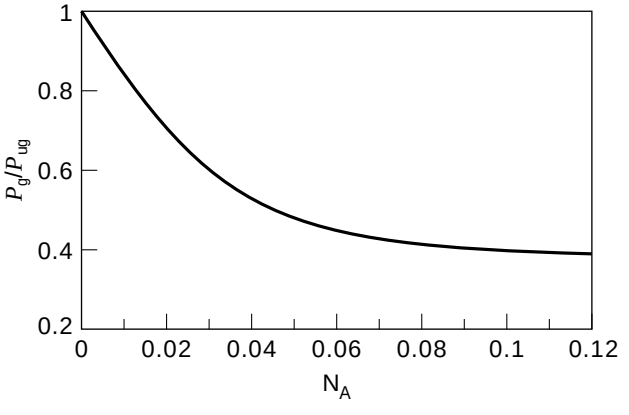


Figure 1. Power ratio as a function of aeration number for a Rushton turbine.

Agitator speed is limited by the natural frequency of the agitation system (4). This is because severe and potentially dangerous vibrations will occur if the rotational speed of the agitator approaches the natural frequency. We recommend that the maximum shaft speed not exceed 70% of the natural frequency.

Having said all this, we are left with the problem of doing the power calculations. This usually involves an iterative calculation in which the parameters are (assuming we have selected impeller types) number of impellers, impeller diameter(s), and shaft speed. (There are other considerations we discuss later.) To do this we use the power correlations already noted. But these usually apply only to single impellers. In most cases, they apply to the lowest impeller, which is the one under which inlet air is introduced. Several approaches have been suggested. The simplest is to treat all the impellers as though each is the only one present. Among other things, this assumes that the impellers do not interact with each other. There is empirical evidence (46pp.264–266) that Rushton turbines will not interact substantively in Newtonian broths if they are spaced at least an impeller diameter apart. Obviously, the assumption of noninteraction is not meaningful for impellers that produce significant axial flow (see later).

Another approach used frequently, but not recommended by the authors, is to use the correlations for the lowest impeller but to use them for the upper impellers:

$$P_g/P_{ug} = (1 + HU)^{-1} \qquad (15)$$

where HU is the gas holdup. You guessed it: there are empirical correlations for HU. One example for Newtonian broths is (47)

$$HU = 1.8(P_g/V_L)^{0.14} V_S^{0.75} \qquad (16)$$

The usual caveats apply.

Typical values for gas holdup range from about 0.1 to 0.3; therefore, equation 15 predicts about 10–20% unloading for the upper impellers. Aeration number correlations predict around 50–60% unloading. Given the uncertainty in all this, and assuming that no other (reliable) information is available, we suggest the more conservative approach of applying the aeration number correlation to each impeller independently.

The added cost for the larger drive will not be a major factor for vessels up to about 5,000 L; this is fairly cheap insurance. One should do pilot agitation studies at a meaningful scale for much larger vessels.

Organism Sensitivity to Fluid Mechanical Forces. Finally, an additional constraint must be imposed if the organism is sensitive to fluid mechanical forces. This is seldom true of unicellular organisms. There are, however, some mycelial organisms that are sensitive. The extent of sensitivity

and the nature of the forces that cause damage should be determined experimentally. One also should keep in mind that the character of fluid forces changes significantly with scale. For example, some turbulent forces that are negligible at small scale can be large and potentially destructive at large scale. Reliable guidance in this area is almost nonexistent. There is, however, a rough rule of thumb that states that damage will occur if impeller tip speed exceeds 1,500 ft/min. Bear in mind that this rule was developed almost 50 years ago and was based on information for organisms used at that time in vessels that had working volumes in the 30,000 to 50,000-gal range. It is clear that many factors (including broth rheology) will affect the potential for damage. Unfortunately, most of the information is seldom available when needed. In such cases, the "any port in a storm" approach to design usually becomes operable. Fortunately, we seldom encounter a circumstance in which the rule of thumb must be violated.

Bulk Mixing

Good bulk mixing is needed to insure homogeneity, which is required for reliable data acquisition and control. There are several important factors that affect mixing quality.

Broth rheology. This is the primary factor affecting our ability to provide good mixing quality, but we seldom have any control over it. We simply must recognize that we will have to go to greater and greater lengths as the broth becomes more viscous and non-Newtonian. Also, one must bear in mind that the nonlinear behavior of most non-Newtonian fluids often results in nonobvious responses to various actions. For example, increasing agitator speed in a highly non-Newtonian, pseudoplastic broth often leads to a decrease in power input and a decrease in bulk mixing quality.

Type(s) of impeller(s). Impeller types fall into three basic categories: radial flow (e.g., Rushton turbine), axial flow (e.g., marine propellor), and mixed axial and radial flow (e.g., Lightnin' A-315). Radial flow impellers tend to deliver the high power required to enhance micromixing and mass transfer but do not promote top-to-bottom mixing; therefore, they are not the best choice for promoting homogeneity. Multiple radial flow impellers often are used in an attempt to compensate for poor bulk mixing. This can work fairly well for low-viscosity Newtonian broths when the ratio of liquid height to tank diameter does not exceed 1.7–2, but it is a very poor solution when the broth is very viscous and highly non-Newtonian. Anyone who has observed classical streptomycete or xanthan gum fermentations can attest to this fact. Pure axial flow impellers do not deliver much power but do tend to promote good top-to-bottom mixing and hence contribute significantly to bulk homogeneity. They do not, however, contribute much to mass transfer. Some mixed flow types appear to provide a good balance between bulk mixing and mass transfer requirements, particularly for non-Newtonian broths. We recommend that they be considered seriously for such cases. We also have had considerable success with combination systems, for example, those with turbines as the lower impellers and a hydrofoil on top. This appears to work well for Newtonian and non-Newtonian broths.

Impeller size(s). In most cases, large-diameter impellers distribute power better and promote bulk mixing better than do small-diameter impellers. As a rough rule of thumb, based on considerable empirical evidence, we recommend D_i/D_t ratios of between 0.4 and 0.5 unless some other constraint (e.g., torque) is violated. Exceptions to this are some cases in which an axial flow impeller is the top impeller. In such cases, the ratio should not exceed 0.35 to avoid the risk of vortexing. This is particularly true for low-viscosity Newtonian broths.

Vessel geometry. We suggest that the ratio of the aerated liquid height to the tank diameter be kept below 2.0 to minimize bulk mixing problems. This must be reconciled with impeller diameter and impeller spacing requirements (see later) and with architectural, shipping, and other constraints.

Impeller spacing. As noted earlier, there is a body of empirical evidence (48) that supports the spacing of Rushton turbines 1–1.5 impeller diameters apart in Newtonian broths. This spacing appears to provide a balance between preventing impellers from interfering with their neighbors and at the same time avoiding dead zones between them. The same sources of empirical data support placing the lowest impeller one impeller diameter from the bottom of the vessel (assuming a standard dished head) and placing the uppermost impeller one impeller diameter below the liquid surface. We have found these rules to work reasonably well for Newtonian broths and a few non-Newtonian broths, and for radial flow impellers other than Rushtons. Please note that these conclusions are based on limited observations. Finally, keep in mind that one cannot avoid interaction among impellers when at least one of them is an axial flow type or has a significant axial flow component, and that spacing effects in highly non-Newtonian broths are not well understood.

Baffling. Baffles are placed in vessels to minimize fluid swirling and vortex formation (see Ref. 48 for flow pattern diagrams). The usual approach is to use four baffles on 90° centers, each baffle having a width equal to 0.1 of the D_t. Baffling tends to increase transmittable power and to improve mixing (except for the dead spots, which tend to form behind the baffles). Elimination of significant vortexing is important for safety as well as for improved bulk mixing. A vortex can reach down to the impeller in an unbaffled vessel. When this happens, a sizable portion of the impeller becomes air enshrouded, thereby decreasing resistance on the impeller. This tends to drive up the impeller speed, at least briefly (or the speed may be increased intentionally in an attempt to increase the power input). Unfortunately, vortices are unstable and can collapse shortly after formation, resulting in very large and potentially catastrophic stresses (impulse) on the impeller, particularly if the impeller speed is increased significantly while it is enshrouded.

Gas flow. Gas flow has a complex effect on bulk mixing. The flow alone does tends to promote bulk mixing (as in bubble tanks), but it also tends to decrease the effect of the impellers, particularly at high values of gas linear velocity. Fortunately, this is not a major problem in most cases, but it can be for very large fermenters and for highly viscous non-Newtonian broths.

Gas Holdup, Foaming, and Aerosol

There are several problems that can interfere seriously with a fermentation. Gas holdup (discussed earlier) decreases the effective volume of a fermenter. Foaming and aerosol formation can constrain operation, cause major cleaning and asepsis problems, and in the extreme cause termination of a fermentation. All three are dependent on broth characteristics, power input, and gas flow rate. There is some information in the literature, but these points have not been given the attention that even begins to reflect their importance. In general, all one can say is that all three become bigger problems for a given fermentation as gas flow and power input increase. Beyond that, each case is a new adventure. More will be said about the practical aspects later.

Heat Transfer

Heat Loads. Heat transfer is required during fermentation to maintain constant temperature conditions, and at other times (i.e., sterilization, induction) to increase or decrease broth temperature. In most cases, cooling is required during most of an active, aerobic fermentation. A good approximation to the total heat load, Q_{tot}, during such fermentations in which the primary carbon source is glucose or a similar carbohydrate is (49)

$$Q_{tot} = Q_{metab} + Q_{mech}$$
$$Q_{tot} = 0.48(\mathrm{OTR})V_L/Y_{x/o} + 2545P_g \qquad (17)$$

where Q_{metab} is the heat generated by metabolism (Btu/h), Q_{mech} is the heat generated by power input (Btu/h), and $Y_{x/o}$ is the yield of cells on oxygen (g dry wt cells per g oxygen consumed). Heat transfer can be more of a limiting factor in aerobic fermentations than is oxygen transfer—particularly in large fermenters. The major heat transfer demand is usually during the maximum growth period (i.e., when μX is greatest). It is always wise, however, to check other loads. For example, rapid cooldown after an induction phase could put enormous demands on the cooling system. Cooling a 10,000-L (wv) fermenter from 38 °C (104 °F) to 10 °C (50 °F) in 30 min would require an average heat transfer rate of approximately 2.4×10^6 Btu/h. The same heat transfer rate would support an oxygen transfer rate of approximately 495 mmol/L(h), which is very high by almost anyone's standards. Similarly, one should check other potentially high loads such as cooldown after sterilization (see later). In any case, careful thought and solid, empirical facts (e.g., product thermal degradation rates) should form the basis of any specification that will require extraordinary transfer rates.

Heat Transfer Rate. The rate at which heat can be transferred is dependent on *(1)* the driving force for heat transfer, *(2)* the area across which transfer must occur and, *(3)* the resistance to heat transfer:

$$Q = (1/\text{resistance}) \times (\text{area}) \times (\text{driving force}) \qquad (18)$$

The driving force for agitated vessels is usually taken to be the log mean temperature difference, ΔT_{LM}, defined as

$$\Delta T_{LM} = (T_{c,o} - T_{c,i})/LN[(T_f - T_{c,i})/(T_f - T_{c,o})] \qquad (19)$$

where $T_{c,o}$ is the coolant outlet temperature (°F), $T_{c,i}$ is the coolant inlet temperature (°F), and T_f is the fermentation temperature (°F). Note that use of equation 19 assumes that the broth is homogeneous and can be characterized by a single temperature.

The transfer area, A_{JA}, is the actual contact area (not necessarily the same as the available area) between the broth and the heat exchange surface, which is usually a jacket and/or an internal coil. The actual area can be found if the height of aerated liquid in the tank is known, along with the geometry of the jacket and any internal exchange area (e.g., a coil). The aerated liquid height measured along the centerline of the vessel, h_{LA} (ft), is

$$h_{LA} = [(1 + \mathrm{HU})V_{LU} - V_D]/(0.785D_t^2) + \mathrm{IDD} \qquad (20)$$

where V_{LU} is the unaerated liquid volume (ft^3), V_D is the lower dish volume (ft^3), IDD is the inside depth of lower dish (ft), and D_t is the tank inside diameter (ft).

The resistance to heat transfer is a little more complicated. It is composed of five resistances in series and is expressed as

$$1/U = 1/h_i + 1/h_o + 1/h_{fi} + 1/h_{fo} + t/k \qquad (21)$$

where U is the overall heat transfer coefficient (Btu/[h °F ft^2]), h_i is the broth film heat transfer coefficient (Btu/[h °F ft^2]), h_o is the jacket fluid heat transfer coefficient (Btu/[h °F ft^2]), h_{fi} is the broth fouling factor (Btu/[h °F ft^2]), h_{fo} is the jacket fluid fouling factor (Btu/[h °F ft^2]), t is the fermenter wall thickness (ft), and k is the fermenter wall thermal conductivity (Btu/[h °F ft]). The resistance of the tank wall can be predicted accurately because both k and wall thickness are known accurately. It should be noted that the k for stainless steel is very low (about 10 Btu/[°F h ft]). This means that the tank wall should be kept as thin as possible to minimize its contribution to heat transfer resistance (see later).

Correlations for h_o can be found in the literature (46pp. 282–283) but are seldom as reliable as those available from vendors for the specific jackets they fabricate.

The problem in predicting U is in predicting reliable values of h_i, h_{fi}, and h_{fo} Published correlations (50) for h_i values for aerated, Newtonian fermentation broths have been published but are not very reliable. Correlations for non-Newtonian broths have also been published (51), but as was the case for mass transfer coefficients, h_i is strongly dependent on broth rheological characteristics. Unfortunately, existing correlations do not do a very good job of accounting adequately for broth rheology, a problem exacerbated by the fact that good rheological information is seldom available. Finally, the fouling factors are essentially impossible to predict. The best one can do is is keep fouling to a minimum by means of good cleaning practices and proper filtration of coolants. Effects of individual resistances on U are illustrated in Figure 2.

The following rough guides (based on many years of experience) are offered because the authors are unable to offer any collection of reliable prediction tools:

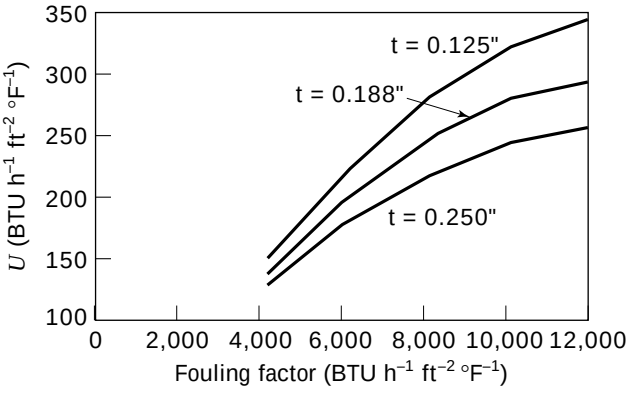


Figure 2. Effects of metal thickness and fouling resistances on the overall heat transfer coefficient.

1. U values for simple, low viscosity Newtonian broths usually range from about 120 to 150 Btu/(h °F ft^2) for new, clean stainless steel tanks.
2. U for viscous, non-Newtonian broths ranges from about 30 to 60 Btu/(h °F ft^2)

It also should be noted that U's for heating usually are a bit higher than U's for cooling.

The heat transfer rate can now be expressed as

$$Q = UA\Delta T_{LM} \tag{22}$$

Similar equations can be developed for heat transfer via an internal coil. The only difference is that the heat transfer coefficients for coils are usually a bit higher than those for jackets. That having been said, we will try everything possible to discourage you from using internal coils. The reasons relate primarily to mechanical and cleaning considerations and are discussed later.

In applying the preceding, we suggest that the reader consider the following suggestions:

- Avoid the use of internal cooling surfaces (e.g., coils). They make cleaning difficult to impossible, are potential sources of contamination (leaking of coolant), put a lot of mechanical strain on the vessel, and can cause deterioration of bulk mixing (particularly for viscous non-Newtonian broths).
- Avoid subfreezing coolant. Sooner or later, subfreezing coolant will cause severe valve freeze-up and worse.
- Try to keep coolant flow rates low enough to avoid pipe diameters greater than 3 in.; cost increases considerably and availability of fittings and so forth decreases as size increases.

Finally, there are some cases in which high wall temperatures have been cited as causing problems with temperature-sensitive organisms. The wall temperature, T_w, depends on the temperatures of the broth and of the jacket fluid, as well as on the relative values of the heat transfer resistances. A simple energy balance across the heat transfer path yields

$$T_w = (T_f + \beta T_j)/(1 + \beta) \tag{23}$$

where

$$\beta = k/h_i/(t + k/h_o) \tag{24}$$

Note that this does not account for locally high wall temperatures that would be caused by steam applied to steam locks during fermentation.

Sterilization

Introduction. Sterilization and aseptic operation taken together (as they must be) have a much greater influence on fermenter design and operation than does any other requirement.

In theory *(1)* sterilization destroys or removes all foreign organisms in all process equipment (including piping, seals, etc.) that might come into contact with the process fluid, and *(2)* aseptic operation insures that *no* contaminating organisms enter the fermenter after sterilization. These ideals are not attainable in practice because (among other reasons) *(1)* there is a finite probability that a very low concentration of contaminating organisms will not be captured in a sample used to test for contamination, and *(2)* there is a finite probability of false positives due to sample contamination and so forth. It also is important to note that there is a continuing debate concerning the definitions of pure culture and sterility (6,51).

The driving forces for sterilization and aseptic operation range from minimizing product losses to insuring strict compliance with regulatory requirements. Among the many specific problems that contamination can cause are the following:

- Production of a toxin that can't be removed by the purification system
- Production of an enzyme that degrades the product
- Decreased product yield due to use of substrate by contaminants
- Production of toxins that inhibit the producer strain
- Production of compounds (e.g., polysaccharides) that interfere with the operation of recovery and purification equipment

Which specific problems will exist and where in the spectrum a particular case will lie depend primarily on the product, its economic value, whether it is regulated, and how it will be used. Given all these variables and the fact that absolute sterility is an unachievable abstraction, the extents to which one should go should be considered on a case-by-case basis. As a practical matter, sterilization has to be interpreted as "effective sterilization," meaning that the design and procedures (including sampling and detection) are suitable for the specific case. For example, sterilization of fermenters used to produce parenterals should be held to a much higher standard than sterilization of

fermenters used to produce amylases for starch hydrolysis. Unfortunately, there remains a lot of room for disagreement among well-intentioned people (see also the section "Product Safety").

Quantification of Sterilization. Mechanical details and other practical considerations of sterilization and aseptic operation are discussed in the subsection "Sterile Piping Systems."

Several methods have been developed to quantify sterilization. The simplest is based on the empirical observation that the death kinetics of many vegetative organisms and spores can be described by a simple first-order expression

$$dN/dt = -kN \tag{25}$$

where N is the number of viable organisms (or spores) at any time, t, and k is the thermal death constant. k is a function of most environmental variables but is most strongly affected by temperature. This dependence is very often expressible as an Arrhenius relationship

$$k = A \exp(-E_A/RT) \tag{26}$$

where A is the preexponential factor (min^{-1}), E_A is the thermal death activation energy (cal/mol), T is the absolute temperature (K), and R is the gas constant. The effect of temperature on k is pronounced for most organisms because their E_A's are so high (e.g., 65,000 cal/mol for *Bacillus stearothermophilus* spores).

Among other methods of quantifying sterilization concepts are the following:

- *Probability-based theories* (52). These are conceptually different from the first-order model, but give essentially the same results for the levels of kill that must be achieved practically.
- *Methods based on the "decimal reduction factor," D.* These are used widely (53). It is important to note, however, that the nature of temperature dependence has no rational, physical basis and differs considerably from an Arrhenius relationship: it is based primarily on an empirical observation for relatively small temperature ranges.

We use the first-order model for the rest of the discussion.

Medium Sterilization. The primary design basis for sterilization systems and procedures is usually the level of sterility that must be achieved. This is most frequently expressed in terms of the probability of sterilization failure (i.e., the probability that a single organism will survive). In most instances probabilities of 10^{-3} to 10^{-4} are quite adequate. This means 1 failure in 1,000 or 10,000 sterilization operations, respectively. It also is important to note that the probability of failure is numerically very close to the value of N in the first-order model.

The reader should also take note of the fact that in some quarters, sterilization criteria are stated in terms of "logs of kill." That is to say, a fixed number of logs (e.g., 12) is taken to be adequate. This is meaningless in that it does not take into account the initial contaminant loading.

Now we consider prediction of requirements for batch sterilizing a fermentation medium as part of the fermenter design process to ensure that we'll have adequate heating and cooling capacity. It should also be done as part of the process design to ensure adequate timing. Batch sterilization involves heating the medium in the fermenter to sterilization temperature (T_{ster}), holding the medium temperature constant at T_{ster} for an adequate time, and then cooling the medium down to fermentation temperature (see the section "Mechanical Design" for additional details). The medium is kept in the fermenter throughout.

It is assumed in what follows that *(1)* heating and cooling are via the jacket only, and *(2)* the energy and condensate contributed by the small amount of steam that flows into the vessel as a result of sterilizing piping, air filters, and so on is negligible for energy balance purposes.

The rate of kill at any instant during the sterilization process is

$$dN/dt = -kN = -A \exp(-E_A/RT)N \tag{27}$$

therefore,

$$N/N_o = A \int \exp(-E_A/RT)dt \tag{28}$$

where N_o is the number of contaminating organisms initially present. The integration of equation 26 must be performed over heating, hold, and cooling phases of sterilization. This is done by numerical integration of equation 28, coupled with the equations that describe broth temperature as a function of time.

For heating

$$T_f = T_{ST} - (T_{ST} - T_{F,O}) \exp(-U * A_{JA} * t/M_F * C_{P,m}) \tag{29}$$

where T_f is the medium temperature (°F), T_{ST} is the steam temperature (°F), $T_{F,O}$ is the initial medium temperature (°F), A_{JA} is the active jacket area (ft^2), t is the time since beginning of heating (h), M_F is the medium mass (lb), and $C_{P,m}$ is the medium heat capacity (Btu/[lb °F]).

For cooling

$$T_f = T_{C,i} + (T_{ster} - T_{C,i}) \exp(t\ (W_J C_{P,c}/(M_F C_{P,m})) \cdot (\exp(UA_{JA}/(W_J C_{P,c})) - 1) \tag{30}$$

where $T_{C,i}$ is the inlet coolant temperature (°F), T_{ster} is the sterilization temperature (°F), t is the time from beginning of cooling (h), W_J is the coolant flow (lb/h), and $C_{P,c}$ is the coolant heat capacity (Btu/[lb °F]). Equations 29 and 30 can be derived via energy balances.

One chooses times and temperatures such that the desired kill, N/N_o, is achieved. Unfortunately, one usually does not know the level of contamination or the nature of the contaminants. Indeed, it is very unlikely that contamination will remain constant from one fermentation to the

next. The usual practice, therefore, is to design for what is thought to be the "worst case." This usually gives results that are quite satisfactory for design specifications. A "typical" worst case is a loading of 5×10^6 spores/mL of the organism *B. stearothermophilus*. Such spores are highly resistant and have the added advantage of being used widely to validate sterilization operations (see later). It is important to note that the purpose of the calculations is to provide a rational basis for obtaining reasonable design information and ensuring a high probability that the desired level of sterilization can be achieved over the whole range of anticipated operating conditions.

Figure 3 illustrates the results of such calculations for the following case:

Medium volume = 15,000 L
Initial spore load = 5×10^6/mL
Initial medium temperature = 70 °F
Steam temperature = 300 °F
Coolant temperature = 35 °F
Coolant flow rate = 100 gpm
Fermentation temperature = 86 °F
Active jacket area = 225 ft^2
U = 150 BTU/(°F h ft^2)
$A = 2.1 \times 10^{36}$/min
E_A = 65,000 cal/mol
N (probability of failure) = 0.0001

As seen in this example, heating and cooling phases usually contribute only a small fraction of the total sterilization kill, but they do affect turnaround time significantly. Long heating and cooling times can also have other effects:

- Cause damage to the fermentation medium to the extent that the fermentation can be compromised. The seriousness of this depends on some combination of economics and regulatory compliance.
- Cause medium changes that have negative effects on recovery and purification without causing fermentation problems. This includes the possibility of introducing foreign materials that may pass undetected into the final product.

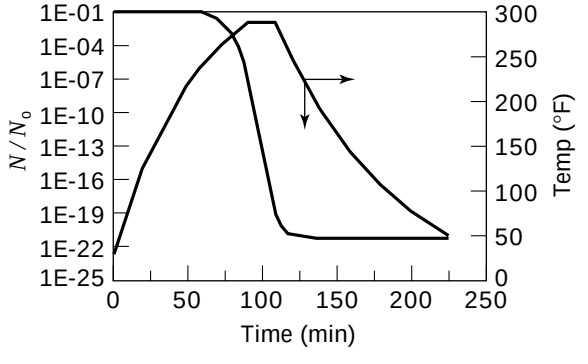


Figure 3. Temperature profile and kill ratio for a batch sterilization.

The example also shows that cooling time is considerably longer than heating time. This results primarily from the lower temperature-driving forces during cooling. The problem is exacerbated considerably when the heat transfer coefficient is very low as is the case for viscous non-Newtonian broths. Another method used for batch sterilization is direct steam injection: live steam is injected directly into the medium as the main source of thermal energy. This decreases heating time as well as the overall steam requirement. It also increases the medium volume by about 20% (as a result of steam condensation), which can cause some problems, including the following:

- *Increased cooling time.*
- *Medium dilution.* This will be a significant problem if the initial medium cannot be made concentrated enough to account for the dilution. Some reasons include low solubility of medium components, increased viscosity, and increased reaction rates among medium components at elevated temperatures.
- *Introduction of impurities.* This depends primarily on the quality of the injected steam. In some cases plant steam is acceptable if boiler cleaning agents do not cause problems. At the other extreme is the requirement to use clean steam (WFI quality). It should be noted with regard to this point that some steam will be injected directly even when jacket heating is used as the main energy source; therefore, one will always be in the position of having to evaluate the effects of contaminants carried by the steam.

For cases in which sterilization times are too long or medium alterations cause too many problems, one might consider continuous sterilization. The interested reader is directed to the literature for additional information (54pp. 159–176).

Piping System Sterilization. Heating and cooling times are not usually significant issues for sterilizing piping; however, there can be some serious heat transfer problems related to piping length, diameter, and orientation. Some of these are discussed in the subsection "Vessel Design." The reader is also referred to literature reports of some practical experiments and theoretical analyses (55,56).

Air Sterilization. Theoretical aspects of air sterilization are not discussed here. Suffice it to say that modern filters, if sized, installed, and maintained properly, will provide very reliable results. We'll discuss some of these practical aspects in the section "Mechanical Design." See Refs. 54pp. 176–183 and 57 for additional information.

Cleaning

Scrupulous cleaning is necessary to decrease nonbiological contamination and prevent cross-contamination of batches. Experience also has shown that reliable sterilization is difficult or impossible to achieve in the absence of rigorous cleaning.

As noted earlier, fermenter cleaning is not mentioned specifically in 21 CFR, but the basis is provided in section

211.67, which requires equipment be cleaned as per a defined plan and by trained people. Furthermore, the FDA has made clear the importance it attaches to cleaning not only in various agency publications (58), but also in more stringent enforcement. This is particularly true in multiproduct facilities (a key issue for most modern biotech plants). This agency initiative and the desire to generally improve operation have caused the industry to develop more reliable cleaning systems and protocols and to move to fully automated cleaning-in-place (CIP) systems, which provide greater assurance of successful validation and long-term compliance.

Developing and designing reliable fermenter cleaning systems is not as straightforward as it might appear, and considerable controversy continues. Among the reasons for this is that while much is known about the basic science of cleaning in general, very little is known about the basic science of cleaning fermenters and little has been published. The approach taken is based primarily on experience derived from the dairy, food, and beverage industries. Such information is useful but is not directly applicable in general to pharmaceutical and biotech processes where soils are different and the cleaning requirements are far more stringent. What meager information there is concerning pharmaceutical and biotech soils has been obtained from experiments done on single soil components and/or studies done under conditions not truly representative of the process conditions (59,60). There have been no significant attempts to develop systematic analyses based on experiments done with complex mixtures typical of real fermentation soils under conditions found in real processes. As a result there are no reliable general methods, tools, or correlations on which to base the development of cleaning agents, protocols and CIP system design; decisions tend to be made based on arbitrary criteria (e.g., coupon bake on studies). In keeping with our general design philosophy, we think it important that such arbitrariness be avoided and that cleaning protocols be developed along with the fermentation. It is clear that this will help to ensure not only proper design but will also minimize cleaning validation studies and any questions concerning the presence of contaminants and/or cleaning residues in the commercial product that could not have been present in clinical trial material.

The selection of fermenter cleaning agents and protocols should be based not only on the specific soil but also on the materials of construction, surface finishes, and so on. In addition, one should consider the issues of *(1)* compatibility of each material with the cleaning agents and protocols, and *(2)* potential materials interaction during cleaning. These are decisions that usually are made during the design phase, but really should be defined much earlier. These issues are discussed in the subsection "Agitation Systems."

Fermenter cleaning validation is beyond the scope of this article; however, the design must be done so as to minimize validation problems and ensure maximum ease in ongoing compliance. The design team must therefore be aware of the procedures that will be used for validation and recognize that this is an issue that has not yet been resolved in the general community.

Putting It Together: Preliminary Design Calculations

We now begin to see some of the interactions that can arise among the transport processes, and between them and the operating variables. This is a necessary step in developing sound designs, and a precursor to preliminary design calculations. As an example, consider a case in which we are trying to find ways to increase the OTR above some base value. One approach would be to increase the gas flow rate (v) keeping other variables constant:

The overall driving force tends to increase (equation 10).
V_s increases (equation 7).
Increased V_s tends to increase $K_g a$ (equation 11).
Increased v also increases N_A (equation 12).
Increased N_A tends to decrease power input (Fig. 1).
Decreased power input tends to decrease $K_g a$ (equation 11).

As another example, consider increasing the vessel pressure while keeping other variables constant:

The overall driving force tends to increase (equation 10).
The actual gas volumetric flow rate (v) tends to decrease (assuming constant molar flow rate) (equation 6).
V_s tends to decrease (equation 7).
The decrease in V_s tends to decrease $K_g a$ (equation 11).
The decrease in v also tends to decrease N_A (equation 12).
This tends to increase power input (Fig. 1).
Increased power input tends to increase $K_g a$ (equation 11).

Such complex interrelationships develop regardless of what steps are taken to increase OTR. The nature of these will change depending on the approach taken and the system being considered. There will also be additional pluses and minuses related to factors such as cost, safety, cleaning, and regulatory compliance. For example, using pure oxygen to increase the driving force carries with it very distinct cost and safety problems. Many of these more practical implications are be discussed in the section "Mechanical Design."

Now let's carry the second approach a little further. We do this because increasing pressure is usually the easiest and seemingly least expensive way to increase OTR. The decrease in V_s caused by the increased pressure has the positive effects of decreasing foaming, aerosol formation, and gas holdup. But it has the negative effect of requiring thicker tank walls, thereby increasing heat transfer resistance in the face of an increased heat load due to increased metabolic activity. This will become a significant problem when the pressure desired is greater than that required for vessel sterilization. The larger the tank diameter, the greater the possible problem. It should also be noted that thicker walls can make it more difficult to get a good surface finish and therefore increases the possibility of cleaning problems. In addition, the higher the pressure and thicker the walls, the shorter the list of qualified vendors.

This type of reasoning process is applicable to any of the many permutations possible. We extend it after we consider some additional basic concepts, mechanical design practices, and a few more important practical constraints.

We now consider two examples of preliminary design calculations, the objectives of which are to, first, define vessel geometry to ensure that it will not violate architectural constraints and be able to satisfy general rules of thumb for good bulk mixing, impeller spacing, and so forth. Some of this will require checking calculations done later (see later) and may require some iteration.

The second objective is to establish ranges of air flow, pressure, oxygen enrichment, and power that will accomplish the following:

- Satisfy the design basis maximum OTR requirement.
- Keep V_S low enough to avoid foaming, holdup, and aerosol problems. We have found a limit of 200 cm/min to be a safe guide; however, there are cases (Newtonian broths) in which considerably higher values up to 300–350 cm/min tolerable and others (very viscous, highly non-Newtonian broths) in which it is not wise to go beyond 75–100 cm/min.
- Not violate compressor and other component constraints.
- Not require heroic efforts (e.g., extremes in pressure, power, etc.).

Please note that we are trying to establish reasonably broad ranges for each of the variables within which we can satisfy the requirements with a reasonable degree of comfort (based on experience). This is important because the correlations usually available have questionable accuracy, and it allows some breathing room for process improvements and so forth.

The third objective is to define impeller type(s), sizes, and speed that will do the following:

- Provide the power required
- Provide good bulk mixing
- Satisfy impeller spacing rules of thumb
- Not violate constraints related to the sensitivity of the organism
- Make possible single-piece impeller construction and so forth (see the section "Mechanical Design" for consideration of mechanical constraints)
- Not violate any other constraints specified

The fourth objective is to define heat exchange area and coolant flow and temperature that achieve the following:

- Satisfy peak heat transfer requirements
- Not violate coolant temperature and flow constraints and such
- Avoid internal heat exchange area
- Not require subfreezing coolant or flow that will require very large pipe diameters

(*Note*: One should check to make sure that other heat loads [e.g., cooling after fermentation] are not greater than the load during maximum growth activity. We assume that to be the case here.)

The fifth objective is to use the results of the calculations not only to define mechanical details (see the section "Mechanical Design"), but also as a basis for recommendations for rational changes in design basis.

Please note that experimentally derived correlations for the subject cases are used in the examples (this is a design gift as rare as hen's teeth). All were obtained at 500-L scale, and we are just within our comfort range concerning their scalability to 15,000 L (design scale). The reader is cautioned, however, that there is no basis to believe that the correlations would be applicable to other cases, nor would we recommend extrapolation to much larger scales. It also is worth noting that there are some good examples in the literature that might be applicable to the problems at hand or to others; however, one must exercise considerable caution if the conditions specified in the reference differ significantly from those for the case being analyzed. Unfortunately, it is often not possible to determine if significant differences exist.

Case 1.

Product	Intracellular protein
Organism	*E. coli*
Final broth volume	15,000 L
Maximum OTR	234 mmol/(L h)
Dissolved oxygen requirement	10% relative to atmospheric conditions
Broth rheological type	Newtonian
Maximum broth viscosity	5 cp
Regulation	CBER
Containment	BL2-LS
Maximum coolant flow	250 gpm
Minimum fluid temperature	35 °F
Internal coils	No
Compressor max press/max flow	50 psig/25,000 SLPM
Minimum heat transfer coefficient	150 Btu/(h °F ft^2)
Ceiling height	336 in.
Floor space	1000 ft^2, unobstructed

The following experimental correlations are applicable:

$$K_g a = 108(P_g/V_L)^{0.6} V_S^{0.2}$$

$$P_g/P_{ug} = 1.0 - 18.56N_A + 215.64N_A^2 - 1082.66N_A^3 + 1859.68N_A^4$$

$$HU = 0.0133(P_g/V_L)^{0.4}(60V_S/^{0.5}) - 0.0333$$

These were developed under the following conditions for single Rushton turbines:

$0.8 < P_g/V_L < 2.9$ hp/100 gal

$100 < V_S < 200$ cm/min

$0 < N_A < 0.1$

$0.33 < D_i/D_t < 0.45$

Vessel geometry calculations are the same for both cases. We start by assuming a vessel diameter and wall thickness. (Note that the thickness will probably be changed a bit later to satisfy code requirements. This will not have a significant effect on height and volume calculations to follow.) In this case we assume a 96-in. outside diameter and a 0.25-in. wall thickness. By straightforward geometry and using tabulated values of volumes and inside depth of dish (IDD) for standard ASME heads, we get the following:

Liquid height	134.2 in.
(Liquid height)/tank diameter	1.41
Total tank height	192.5 in.
Total volume	21.1 L
Percent fill (unaerated)	71.1%

This satisfies rules of thumb concerning liquid height-to-diameter ratio and fill percent. If we add about 72 in. for legs and 30 in. for opening the manway, we get a total height requirement of 294.5 in. We should be able to satisfy the maximum height limit.

The results of oxygen transfer calculations for several operating conditions are given in Table 1. From these results we conclude that the following will satisfy the OTR required and give us a reasonable amount of "breathing room."

Gassed power	100 hp (includes transmission efficiency)
Maximum operating pressure	40 psig
Maximum air flow	10,000 SLPM

We do not recommend oxygen enrichment at this point, but we do suggest that the "holes be punched" to facilitate future installation of an oxygen system. The agitation system calculations are based on the assumption of using two 40-in. diameter Rushton turbines. The results are given in Table 2. A maximum speed of 150 rpm should be adequate. Note also that the impellers can be spaced to satisfy the rules of thumb for bulk mixing rules discussed previously. Although this should give satisfactory results for such a low viscosity broth, we recommend consideration of a slightly larger turbine as the lower impeller and a hydrofoil (e.g., A-315) as the top impeller.

The results of the heat transfer calculations are given in Table 3. A minimum coolant temperature of 35 °F was used in all cases. The results show that the heat loads can be handled, but the conditions are uncomfortably tight. For example, in no case will increasing the coolant flow to the maximum make up for a 10% decrease of the heat transfer coefficient. The potential consequences of inadequate heat transfer in this case should be considered seriously before final design commitments are made. Alternatives (e.g., lower coolant temperature) should be explored.

Case 2.

Product	Antibiotic
Organism	*Streptomyces* sp.
Final broth volume	15,000 L
Maximum OTR	50 mmol/(L h)
Dissolved oxygen requirement	10% relative to atmospheric conditions
Broth rheological type	Non-Newtonian (pseudoplastic)
Maximum broth viscosity	1,000–1,500 cp
Regulation	CDER
Containment	BL1-LS
Max coolant flow	250 gpm
Min fluid temperature	35 °F
Internal coils	No
Compressor max press/max flow	50 psig/25,000 SLPM
Min heat transfer coefficient	40 btu/(h °F ft^2)
Ceiling height	336 in.
Floor space	1000 ft^2, unobstructed

The following empirical equations are applicable:

$$K_g a = 8.0(P_g/V_L)^{0.5} V_S^{0.4}$$

$$P_g/P_{ug} = 1.0 - 26.99N_A + 417.37N_A^2 - 2789.43N_A^3 + 6643.30N_A^4$$

Table 1. Oxygen Transfer Calculation Results (Single Cell)

Pressure (PSIG)	35	20	20	2
O_2 trans eff. (%)	21	30	28	28
O_2 (MOL % in)	21	21	40	40
Gas flow (in, SLPM)	29,762	20,833	11,719	11,719
VVM	1.98	1.39	0.78	0.78
Average V_s (cm/min)	186	183	103	200
hp/100 gal	1.3	2.7	1.0	2.6
Theor. P_g	52.8	105.6	41.1	103.8

Table 2. Citation System Calculation Results (Single Cell)

Shaft speed (rpm)	114	141	88	141
Turb imp dia (in)	40.00	40.00	40.00	40.00
Turb D_i/D_t ratio	0.42	0.42	0.42	0.42
Ungassed hp (turb)	123.1	233.0	56.6	233.0
N_a, aeration numb	0.071	0.057	0.051	0.062
P_g/P_o ratio	0.43	0.46	0.48	0.45
Gassed hp	53	106	42	104

Table 3. Heat Transfer Calculation Results (Single Cell)

Heat load (MMBtu/h)	1.808	1.943	1.779	1.938
Hold-up	0.199	0.271	0.123	0.282
Jacket area (ft^2)	297	317	276	320
UJKT (BTU/ft^2 h °F)	150	150	150	150
Coolant temp in (°F)	35	35	35	35
Coolant flow (GPM)	190	210	235	200

$$HU = 0.012(P_g/V_L)^{0.3}(60V_S/100)^{0.52} - 0.029$$

These were developed under the following conditions for single Rushton turbines:

$1.0 < P_g/V_L < 2.3$ hp/100 gal
$60 < V_S < 130$ cm/min
$0 < N_A < 0.1$
$0.36 < D_i/D_t < 0.42$

The results of oxygen transfer calculations for several operating conditions are given in Table 4. From these results we conclude that the following will satisfy the OTR required.

Gassed power	100 hp (includes transmission efficiency)
Maximum operating pressure	40 psig
Maximum air flow	25,000 SLPM

Oxygen enrichment should be considered in this case despite the fact that the calculations show that it is not necessary. The reason is that the gas flow rate, and hence V_S, is quite high without enrichment. This results primarily from the very low oxygen transfer efficiency typical of very viscous non-Newtonian broths. Although it is true that we haven't broken the rule of thumb for holdup, foaming, and aerosol we should take note of the fact that high air flows in pseudoplastic broths can exacerbate impeller unloading and cause flooding more readily than for non-Newtonian broths. It also is generally true that our correlations become less reliable as conditions become more extreme and further removed from the conditions under which they were developed. Note in particular that our correlations were not developed for the high values of V_S found in the calculations. It also is clear that we should specify higher pressures than shown in the results in Table 4. Additional calculations show that 40–45 psig would bring the V_S into a comfortable range. We suggest, therefore, the following:

Gassed power	100 hp (includes transmission efficiency)
Maximum operating pressure	45 psig
Maximum air flow	25,000 SLPM
Oxygen enrichment	40% oxygen in inlet stream

The agitation system calculations are based on the assumption of using two 40-in.-diameter Rushton turbines. The results are given in Table 5. A maximum speed of 140 rpm appears to be adequate. Note also that the impellers can be spaced to satisfy the bulk mixing rules of thumb discussed previously.

That's very nice, but we have significant reservations about the results. An alternative should be given serious consideration because of the rheological nature of the broth, the profound effect this can have on mixing, and the issues of scalability for such broths. At the very least, the lower turbine should be replaced by one having a D_i/D_t approaching 0.5, and the upper turbine should be replaced by a hydrofoil. The reader should understand that this is based on experience only: we have no solid correlations for this case. The best route to take would be to do a few pilot mixing studies in as large a vessel as possible with the real broth.

The results of the heat transfer calculations are given in Table 6. The minimum coolant temperature of 35 °F was used in all cases. The results show that the heat loads can be handled. It also appears that we are well below the maximum available coolant flow. This may be comforting, but the reader should realize that there is not a lot of breathing room for process improvement. In particular, an increase in OTR up to 60–65 mmol/(L h) would put us over the top even if there were no deterioration in heat transfer coefficient. Again, one should give this serious consideration before committing to a final design.

We trust that the discussion and examples in this section have helped the reader to understand not only the many interactions that must be considered, but also the nature of preliminary design calculations and their limi-

Table 4. Oxygen Transfer Calculation Results (Mycelial)

Pressure (PSIG)	35	25	20	10
Oxygen transfer eff (%)	6	6	6	6
O_2 mol % in	21	21	40	40
Gas flow IN (SLPM)	22,222	22,222	11,667	11,667
VVM	1.48	1.48	0.78	0.78
Average V_s (cm/min)	154	172	102	140
hp/100 gal required	1.7	2.0	1.0	1.6
Theoretical *P*g	60.9	79.9	41.7	61.8

Table 5. Agitation System Calculation Results (Mycelial)

Shaft speed (rpm)	124	137	107	124
Ungassed hp	159	214	102	159
Aeration number	0.049	0.055	0.042	0.049
Gassed hp	61	80	42	62

Table 6. Heat Transfer Calculation Results (Mycelial)

Heat load (MMBtu/h)	0.533	0.560	0.463	0.515
Gas hold-up	0.202	0.230	0.124	0.182
Jacket area (ft^2)	298	306	276	292
UJKT [(Btu/ft^2 h °F)]	40	40	40	40
Coolant temp in (°F)	35	35	35	35
Coolant flow (GPM)	190	210	235	200

tations. The results obtained here must now be translated into a mechanical design that must also take into consideration the items to be discussed in the next section.

MECHANICAL DESIGN

Mechanical Design Basis

Much of the mechanical design follows from the results of calculations discussed in the previous section 4; however, these must be tempered by other considerations that should be included in the process and/or facility design basis. Some of these (e.g., regulatory compliance, containment) have already been discussed and are revisited in greater detail in this section. Other items that should be included and will affect fermenter design include the following:

- *Extent of automation, nature of plant wide control system, etc.* Such items affect not only the instrumentation (beyond our scope) and similar factors, but also valving, piping, vessel ports, and so forth. Among some of the major issues here are identification of critical (as defined by cGMP requirements) valves and other components. It must be kept in mind that failure to identify critical instrumentation and control components frequently leads to very complex valving systems in which the failure of a single switch can bring everything to a grinding halt. Unfortunately, there are too many people who do not think this through (as a system) before final design begins. Much of this results from the human tendency to rationalize and/or push "problems" downstream, the kind of thinking that results in the thought "Why worry? We'll get it whether or not we need it." Such thinking is in no one's best interest (buyer or seller).
- *Staff requirements and the anticipated nature of the operating staff.* This can influence the complexity and physical layout of (among other things) the piping system(s).
- *Plant location*. This will have some affect on overall design and component selection if for no other than service issues.
- *Available utilities*. This can influence such things as the designs of the cooling and aeration systems.
- *Architectural constraints*. These can have profound effects on the ways in which process requirements will be satisfied. Floor space and ceiling height constraints often require that a fermenter be designed in such a way that it violates some or most or even all of the rules of thumb mentioned in the previous section. This is not good but is better than building a vessel that does not fit into the plant.
- *Transportation constraints*. The fermenter has to be moved from the fabricator's shop to the plant—at a cost not greater than the total project budget. This may require some thumb bending or breaking. In extreme cases it may be better to fabricate the fermenter on site; however, this can carry large penalties, particularly in licensed facilities.
- *Maintenance*. Many design decisions can have major effects on the ease of maintenance. It is also often true that some of the design features that ease maintenance are more costly than those that do not. In most cases, the savings in capital expenditure will not come near paying for losses that will result later because of maintenance problems. This is particularly true in licensed facilities where regulators view good, facile maintenance as an integral, indispensible part of cGMP operation.
- *Definition of standards*. Standards for welds, finishes, and so on should be standardized for all parts of the equipment that will contact process fluid. It does not, for example, make sense to call for a high-quality vessel finish (e.g., 320 grit, EP) and at the same time accept unpolished tubing and valves in inoculation and medium addition lines.

Finally, there are a few recurring themes that are encountered in fermenter design:

- *Building-in "versatility."* The types of versatility desired range from wanting a fermenter that can operate well over a very wide range of conditions but with a narrow range of organisms, to wanting a convertible bioreactor that can handle microbes, mam-

malian cells, and maybe (someday) transgenic animals. Most people understand the value of versatility, but many do not understand the attendant problems and costs. As a general rule, converting laboratory glassware directly into large, stainless steel equivalents usually costs more than any rational person should be willing to pay. Also as a general rule, versatility should decrease as a process goes from the lab to the production floor. Versatility in the laboratory is almost a requirement because change is in the nature of laboratory work. Versatility on the plant floor, however, leads to more procedures, more paper work, more testing, and more confusion, particularly when equipment modifications are required to achieve the versatility; change is not in the basic nature of most plant work. Equipment capital savings can justify the added costs of all the preceding for many cases in which the equipment is designed to handle multiple, similar fermentations on a campaign basis and without significant equipment modification. This statement becomes less true as the fermentations become less similar, as more modifications are necessary, and as regulatory scrutiny increases. Bottom line: analyze very carefully any inclination to want versatility built into plant equipment (or even pilot equipment, in some cases).

- *Retrofitting existing equipment.* The usual thinking is that capital and time savings can be had by refurbishing "old faithful." Just how true this is depends not only on the condition of the existing equipment but also on the intended application(s) of the reborn version. Comments similar to the ones made for versatility apply here. The chances of success decrease as one goes from a lab to licensed production facility. There are ample, expensive corpses to prove this point.

Vessel Design

Please note that Figures 4 and 5 are simplified vessel and piping diagrams intended to assist the reader in the discussions to follow.

Materials of Construction. The major choices that must be made are *(1)* the type of metal to be used for the vessel and nozzles, and *(2)* the type of elastomer to be used for static seals (see later for rotating seals). The selections should be based on compatibility with the organism, compatibility with the product, corrosion resistance, cleanability (also related to finish), welding characteristics, and cost and durability. All of these should be determined during process development but seldom are. In almost all cases, the metal selected will be some grade of stainless steel (usually SS304, SS304L, SS316, or SS316L). The choice is usually associated with the nature and value of the product, although some of the other factors noted earlier may be considered. SS304 is usually good enough for lower-value, unlicensed products, whereas SS316L is the material of choice for high-value, licensed products. L-grade is selected when better corrosion resistance and good multipass welding characteristics are required; it adds about 15% to the cost of the vessel.

It is important in making metal selection to keep in mind that clean steam, purified water, and WFI are all highly corrosive. The reader should note that there is considerable controversy concerning metal selection and is directed to Refs. 61–63 for additional information and opinion.

Static seal (O-rings and gaskets) materials are usually limited to silicone, EPDM, Viton, or Teflon. Silicone and EPDM are the materials of choice for headplates and elsewhere, respectively. (The reader should be alert to the fact that there are several grades of both, not all of which are FDA approved.) Teflon has much better temperature resistance than either silicone or EPDM, but it doesn't stretch, and it cold flows. Viton hardens on use.

Finally, one should be aware of that some O-ring/gasket-forming processes can leave very small quantities of metals in the seals. These may not be enough to cause fermentation or product problems per se, but they can be the cause of considerable corrosion.

Finish. In most instances, the need for a high-quality finish is closely related to the need for high-quality cleaning: there is little question that CIP systems function better when surfaces are smoother. Smoothness is expressed on several scales, two of the more common being grit number and surface R_A. Grit number is related directly to the abrasive used to achieve the finish (mechanical); surface R_A is the actual surface roughness as measured by means of a profilometer and expressed in microns:

Grit	R_A (μm)
120	3.2
220	0.8
320	0.4

The two most common ways of achieving smoothness are mechanical polishing and electropolishing. Mechanical polishing is done by means of a series of continually decreasing abrasive particle sizes (grits). Electropolishing is achieved by electrochemical removal of metal from high points on the surface. One is usually best advised to electropolish only after getting a good mechanical polish (240-grit minimum). It has long been held that the very good finishes required for vessels where cleaning is critical could be achieved only by electropolishing a 320-grit mechanical finish (although 240-grit EP is the most common practice). There is some evidence in the literature (62,63) to support this, but it is not overwhelmingly convincing. It also is important to note that recent advances in mechanical finishes have produced surfaces that easily rival the appearance of any electropolished surface (Lee Industries, personal communication, 1996).

It should be (but is not always) evident that cost increases with surface smoothness. It is also important to note that just because a surface is almost perfectly smooth does not guarantee that it will be easy to clean. Cleanability will also depend on the nature of the interaction between foulants and the surface material. The only way to know for sure is to try your broth on various surfaces and

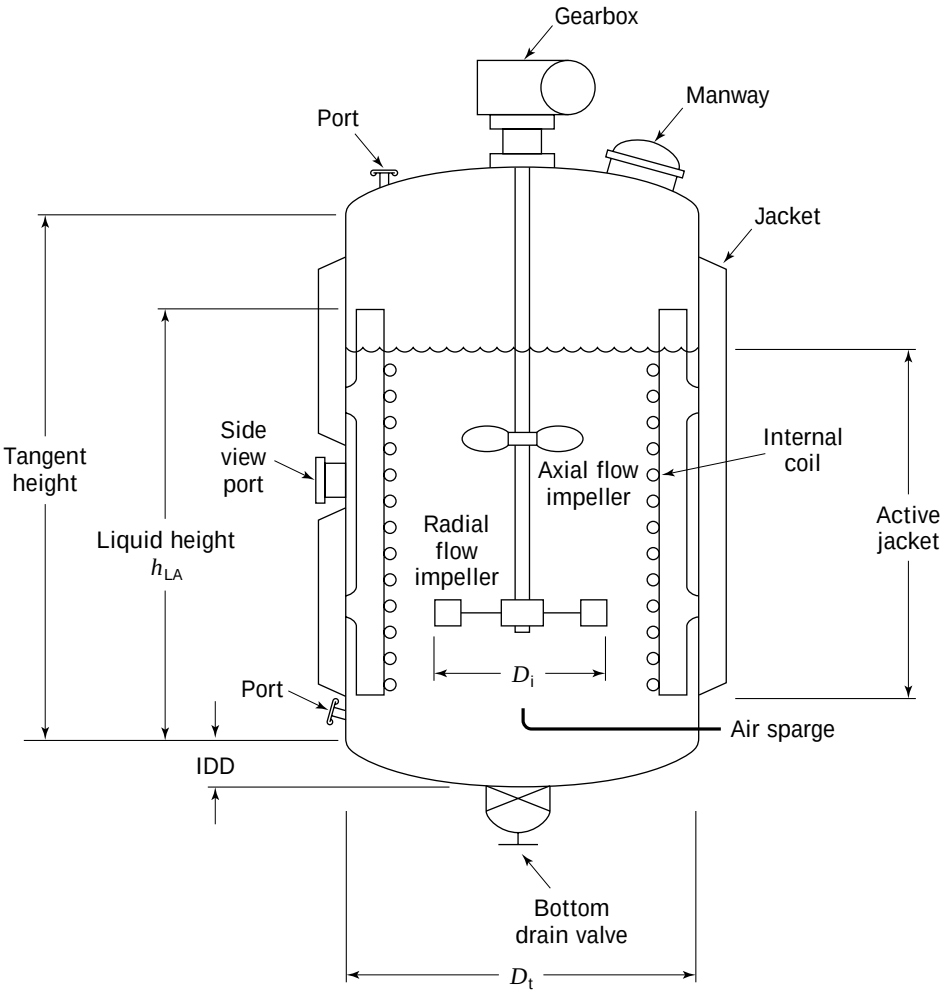


Figure 4. Fermenter vessel schematic and terminology.

then make a specification based on rational evaluation of the experimental information. Surface material and finish, cleaning agents, and cleaning protocols must be considered together—preferably before design.

Passivation must also be considered. This is a treatment that restores oxides that protect stainless steel from corrosion but can be removed during vessel fabrication (primarily welding). The process is rather simple, but the surface chemistry is not and is not completely understood (64,65); hence, there is controversy as to which method is best and whether it is really as necessary for fermenters as it is for high-purity water systems where rouge formation is a serious problem (66). Two methods are in common use: *(1)* treatment with sodium hydroxide, citric acid, and nitric acid; and *(2)* treatment with chelating agents. The second is much milder and does not generate the noxious waste of the first. Both appear to be equally effective; hence, the chelating method is becoming more popular.

Nozzles. Ports for additions and probes must be designed to be sterilizable and cleanable. Among other things, this means they must have reliable seals and be free draining. The major debate here usually focuses on the choice between Ingold ports and those designed for sanitary clamp connection (Fig. 6). There have been sterility problems with Ingolds, most of which have been traced to the ports being out of round as a result of distortions caused by welding. Such distortion causes obvious problems with the O-ring seals. This problem can be corrected quite easily by making the nozzle thick enough to avoid distortion. Our experience has been that such a correction yields nozzles that are just as reliable and more easily cleaned than are sanitary fittings when both are mounted at the usual 15°.

Side View Ports. The first word that comes to mind is "don't." Side view ports let the bugs see out more than they let you see in. The little that's gained by being able to see less than 1% of the total action does not seem to justify the expense, the jacket coving, and the cleaning and sterility problems caused by these little gems. If you must have one, we suggest the circular Metaglas™ type (Fig. 7) attached to the vessel with sanitary clamps; the glass is bonded directly to the metal, which avoids the sealing problems encountered with other types. It does, however, have the disadvantage of having a smaller viewing area than does a standard, circular view port, and it is more expensive. We

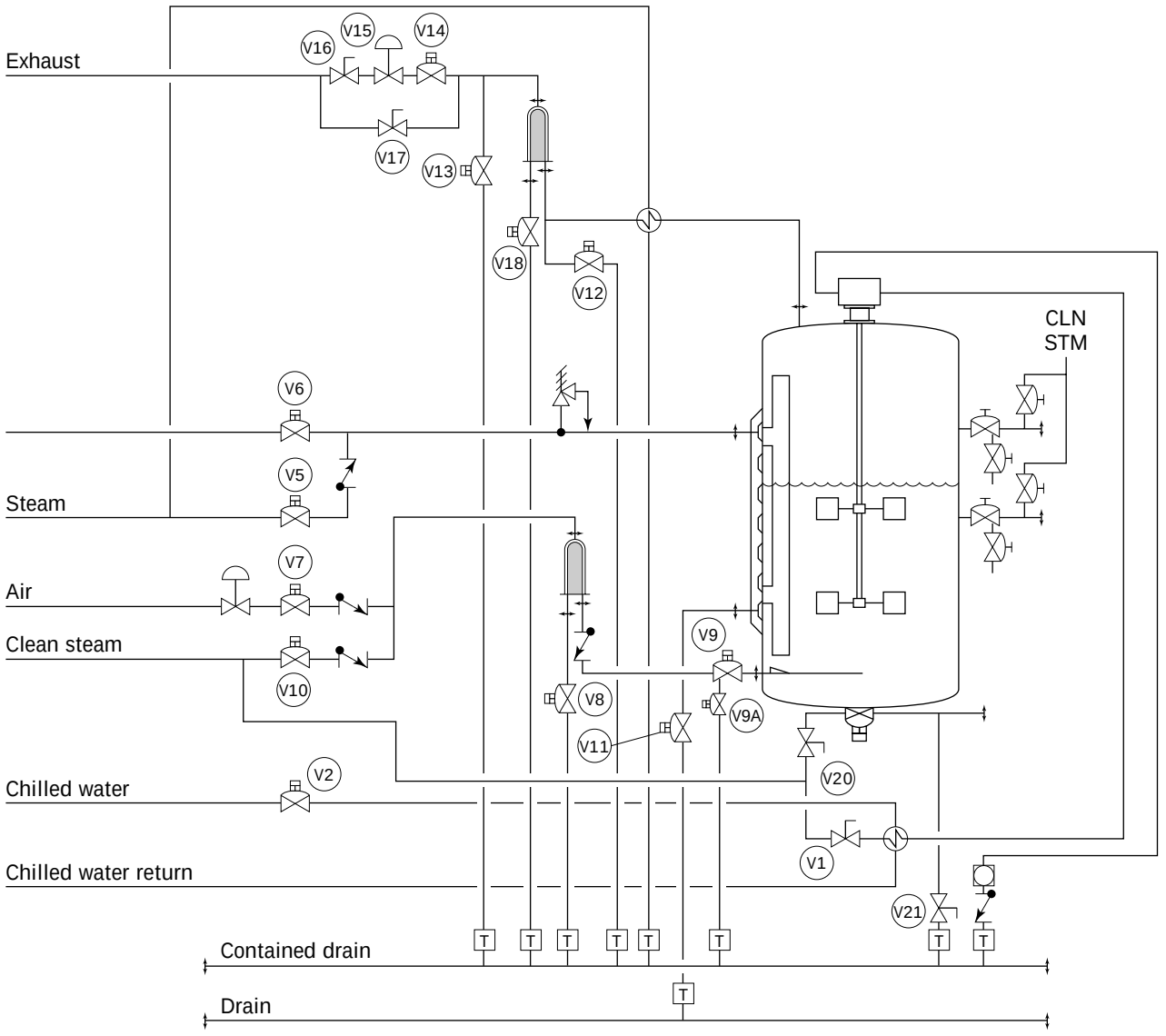


Figure 5. Simplified fermenter piping diagram.

recommend avoiding long rectangular view ports, which are very difficult to keep clean and free draining, and types in which the glass has to be sealed with gaskets and/or O-rings on both sides (Fig. 8). Keep in mind that tolerances for glass are measured in fractions of an inch, not thousandths. That, coupled with the fact that the gaskets will compress means that the probability of repeatedly achieving a good seal and good cleaning is lower than your odds of winning more than once at Atlantic City.

Baffles. Baffles are usually required in high-power systems to prevent swirling and vortexing, thereby increasing the power that can be delivered to the fluid. The usual practice is to use four baffles on 90° centers welded directly to the wall. Each baffle should have a width equal to 10% of the tank diameter and should have long slots cut out of the edge facing the wall so as to prevent solids build up and to make cleaning easier. Removable baffles are used in some cases; however, we discourage this practice, particularly in cases where cleaning is a critical issue, simply because unsealed joints resulting from baffle removal make cleaning more difficult.

Jackets. Several types (67) of jacket are used on fermenters; the choice of type is usually not critical for heat transfer purposes and is best left to the vessel fabricater. In some cases, however, a jacket type (e.g., half pipe) may be chosen to increase the vessel pressure rating. The following have proven to be useful practices:

- The jacket should extend from the probe ring (about 2 in. above the bottom tangent line) to the top tangent line and should be zoned. This allows additional active surface area to come in contact with the broth as the volume increase due to additions and to increas-

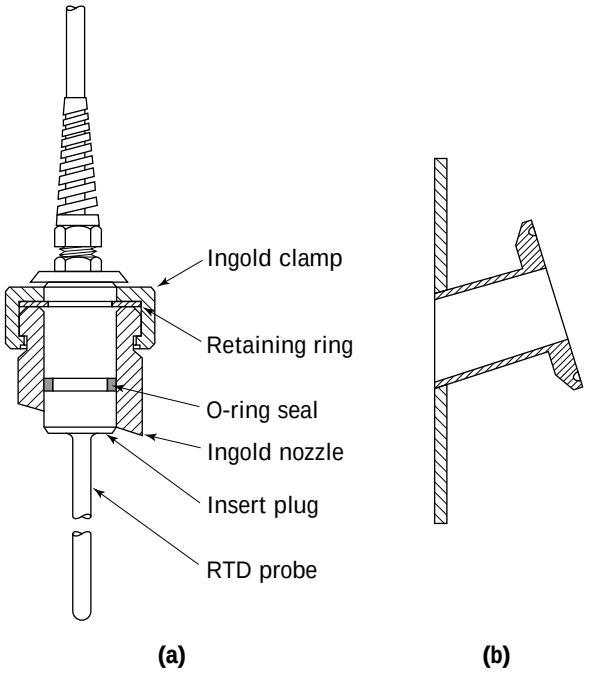


Figure 6. Ingold and sanitary clamp fittings.

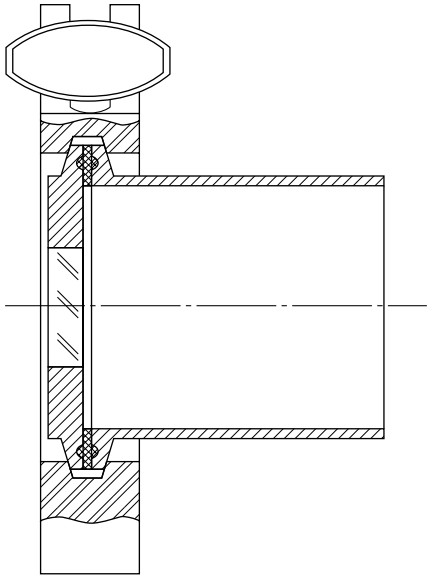


Figure 7. Metaglas™ side viewport.

ing gas holdup, minimizing cooling-loop pressure drop, and minimizing medium bake-on.

- Connections to the jacket should be via sanitary clamps or flanges, not by screwed fittings. This minimizes possible damage to jacket welds when cooling lines are connected or disconnected.
- Coolant should be filtered to avoid solids buildup and/or jacket fouling.
- Bottom jackets should be considered only for vessels larger than 5,000 L. The additional area will be about 10% of the straight side jacket area (allowing for bottom drain, etc. and will add about 5% to the cost of the vessel). In cases where there is an internal coil, bottom jacketing isn't necessary or desirable.
- Jacket coving to accommodate view ports, decreases jacket area and effectiveness and increases cost considerably. It should be avoided.

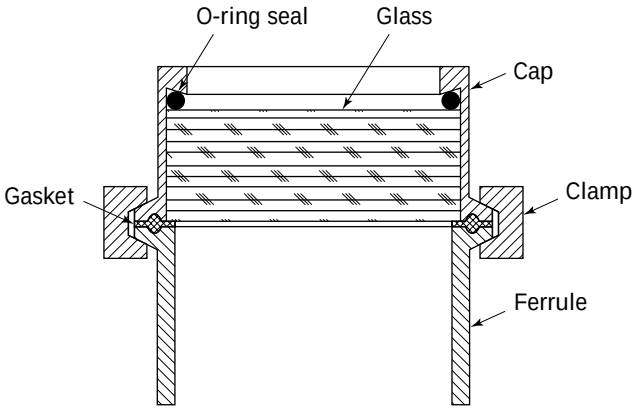


Figure 8. Side viewport with separate glass piece sealed by two gaskets.

Internal Cooling Surfaces. Coils are the most common internal cooling surfaces, although there are other types (68). They can easily double the available heat transfer area and tend to be more effective than jackets; however, they can have big disadvantages:

- They make cleaning very difficult.
- They can cause additional bulk mixing problems, particularly for very viscous non-Newtonian broths.
- They will eventually leak nonsterile coolant into the broth.
- They can add up to 25% to the cost of the vessel.

We recommend that every other alternative (including slight decreases in growth rate and/or cell mass) be considered before yielding to the temptation to use coils. Keep in mind that, once installed, they will be there for the life of the vessel. If you absolutely, positively must use a coil, there are several points to remember:

- Mount it in a way that will insure minimum stress during heat-up and cooldown.
- Weld cladding over the butt welds used to join the coil pipe sections.
- Leak test (69) after construction and build in design features that will simplify leak testing on a regular basis thereafter.
- Space coil turns at least 3 in. apart. Anything closer will insure major cleaning problems.

Spargers. There has been a lot of discussion concerning the pros and cons of ring and single-orifice spargers, and the details of design of each. We have seen both work well

and have not found any evidence for the validity of claims concerning the importance of hole size (for example) for oxygen transfer per se. The primary focus should be on gas distribution (which will depend on other aspects of the agitation system) and on aseptic operation. We usually favor single-orifice spargers.

There also have been advocates of porous (frits) spargers. The rationale presented has focused primarily on the small bubble size such spargers produce. There is some basis for these claims in cases where little mechanical energy is available for bubble breakup (e.g., in mammalian cell systems). There might also be some value in cases for which bubble coalescence is not a problem (far and few between in practical systems). They can be extremely difficult to clean, particularly for mycelial organisms. We usually suggest they not be used except in very special cases, as just noted.

Finally, it is important that the sparger be made and mounted in such a way that it can be removed easily for cleaning.

Piping and valving

Design and construction details of the piping system depend on process, sterility, cleaning, and containment requirements, taken together.

Service Lines for Non-Process-Contacting Fluids. All lines providing fluids that cannot contact process fluid surfaces (e.g., coolant lines, plant steam lines for heating only) fall into this category. Satisfactory service is provided by either copper or stainless steel piping along with a rational combination of welded and compression fittings. Durability, servicability, cost, and corrosion resistance are major considerations. Ball valves are satisfactory on lines not turned on and off frequently (diaphragm valves tend to withstand a greater number of on–off cycles prior to failure).

Sterile Piping Systems. Before we discuss design of sterile piping systems, we reiterate the importance of a systems approach to integrating vessel and piping design. Independent design almost always results in a lot of aggravation, as well as higher cost and lost time.

A sterile piping system can usually be divided into five major subsystems: process, steam/condensate, air, harvest, and seal lubrication. Each has specific requirements dependent on its unique functions as well as the specific demands of the process. None of these systems is completely independent of the others, and all share common components. All also interact with the CIP piping system, which is discussed later.

Details of each of these subsystems follow a simplified description of sterilization procedures (Fig. 5). The reader is advised that the procedure is one of several commonly practiced; there are, for example, different opinions as to when live steam to the air inlet line should be turned on.

Phase 1. After the vessel is filled with set medium, steam flow is started through the jacket. The agitator is run at low speed, and particulate free steam lubricates the shaft seal. The exhaust line is open for venting to the atmosphere. All other valves are closed. During this time not only are the vessel and the set medium being heated, but air is being purged from the medium. If the air is not purged, the vessel pressure during sterilization can become dangerously high, and there will not be a reliable correlation between pressure and temperature. Phase 1 continues until the medium temperature reaches 100 °C.

Phase 2. Steam continues to flow through the jacket. The air exhaust line is closed; the only path out is via steam traps. Steam is now admitted through the air inlet line to sterilize the inlet filter; steam leaving the vessel through the air exhaust line sterilizes the exhaust line piping and the exhaust filter. Steam flow is started through subsurface ports via steam lock assemblies (Fig. 9). (See later for further discussion of steam locks.) Ports above the liquid surface are sterilized by steam escaping through them from the vessel and then to condensate lines. Steam flow also is started through the bottom drain valve. All steam flows continue until the cooldown begins. Phase 2 continues until the medium temperature reaches 121 °C.

Phase 3. Steam flow to the jacket is throttled to hold sterilization temperature constant until cooldown begins.

Phase 4. Cooldown. This is not discussed here.

General Principles. There are several general principles that should be applied to all sterile piping:

- Make all piping system components free draining. This requires special attention to the details of pipe pitches, valve orientations, and so forth
- Design all components and the piping layout so as to eliminate nooks and crannies where contamination (biological and nonbiological) can hide so as to escape cleaning and/or sterilization. Frequently overlooked problems include such things as deadlegs and ridges formed by welding operations.
- Make steam and condensate piping at least 3/8 in. diameter to insure proper steam flow and condensate draining.

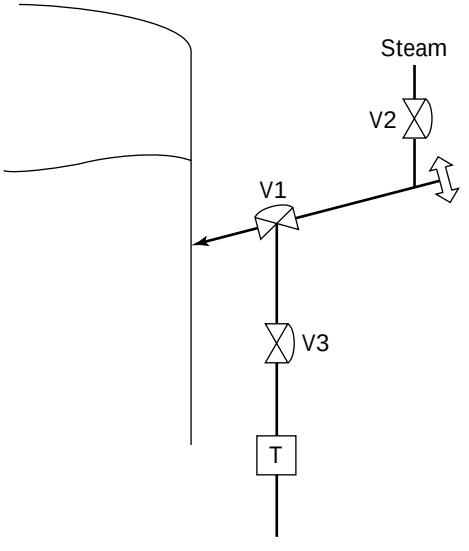


Figure 9. Sterilization of subsurface ports.

- Make all piping lengths as short as possible without interfering with good fabrication practices (GFP), operability, and maintenance.
- Use check valves only when no other solution is possible (e.g., in overpressurized lines).
- Do not use sight glasses in condensate drain lines except in the seal lubricant drain line.
- Pay careful attention to piping orientation to avoid the possibility of air traps and inadequate heating (55,56).
- Make sure that bottom drain valves are flush-mounted diaphragm valves capable of being steamed in the closed position (Fig. 10). They must be free draining.
- Avoid dip tubes unless there is no other way (there almost always is).

Finally, the design should take into consideration the methods that will be used for validation of sterilization operations. These will vary considerably with the product and the organism. At one extreme is virtually no validation; at the other extreme is temperature mapping of the vessel and all the trap lines, as well as the use of spore strips and/or spore ampules in the vessel and in the trap lines. In some cases, tee sections have been included in drain lines for insertion of spore strips.

Welds. Orbital welding of the piping system produces the most reliable results with regard to aseptic operation and cleanability; it is essentially a requirement for many classes of regulated products, particularly those regulated by Center for Biologics Evaluation and Research (CBER). It also is the most expensive to build and to maintain. There are, however, many processes for which an economic combination of welded and compression fittings is acceptable. These can be operated quite satisfactorily if proper attention is paid to cleaning, sterilization, and maintenance protocols. The fact is that such systems have been used for many years to manufacture regulated products safely and efficaciously.

Valves. Diaphragm valves have become essentially required for most classes of regulated fermentation products. There is little argument that their design provides the greatest reliability for clean, aseptic operation. They have the added advantage of integral sterile access ports, which permit very short piping runs, particularly in valve clusters. Historically, diaphragm valves have been more expensive than have ball valves, but this has changed recently, particularly for some materials (e.g., SS316L). It therefore makes sense to use diaphragms for most applications. The primary exceptions are steam lines or other applications in which diaphragm life is a problem; in those cases one should consider ball valves or plug valves, as appropriate.

Another important consideration is the extent to which one uses valve actuators and position indicators. Obviously, this will depend in large measure on the nature of the control systems employed. In any event, one should consider very seriously the problems (e.g., space and orientation constraints, reliability) associated with valve "extras" while control system decisions are being made.

Steamlock Assemblies. The primary purpose of steamlocks (also known as block and bleed) is to allow sterilization of ports and process piping at any time during a fermentation in such a way that connections can be made and broken without risk of breaching the sterile barrier and/or the containment barrier. A basic steamlock assembly is illustrated in Figure 11. In this case, vessel T1 contains presterilized nutrient that must be added at some time during the fermentation. T1 is connected to the fermenter by means of sanitary clamps (other means are possible). Steam flow is then started (open V1, V3, and V5) so as to sterilize the connecting hose and all the valving not previously sterilized. Condensate goes to drain (sanitary, if necessary) via V3 and V5. The steam and condensate valves are closed after sterilization is complete; process valves V2 and V4 can be opened anytime thereafter to make the sterile transfer. Note that the diaphragm valve sterile access ports make it much easier to sterilize the

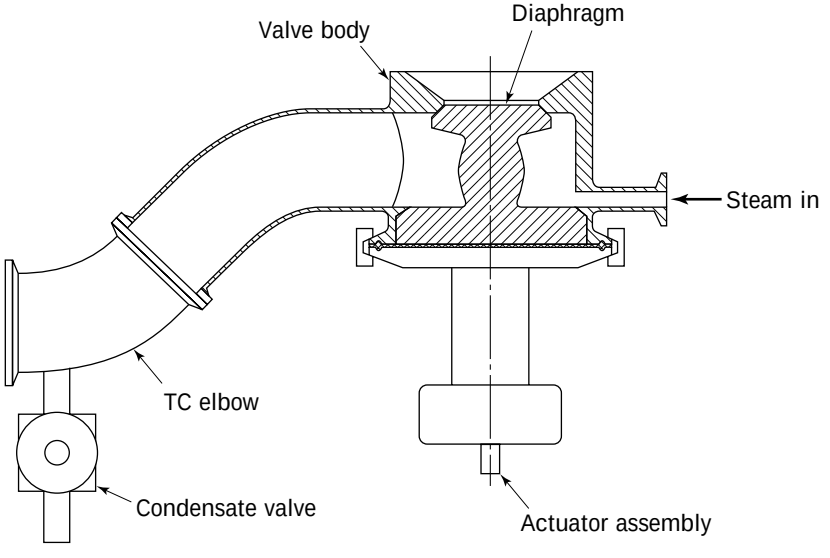


Figure 10. Bottom drain valve.

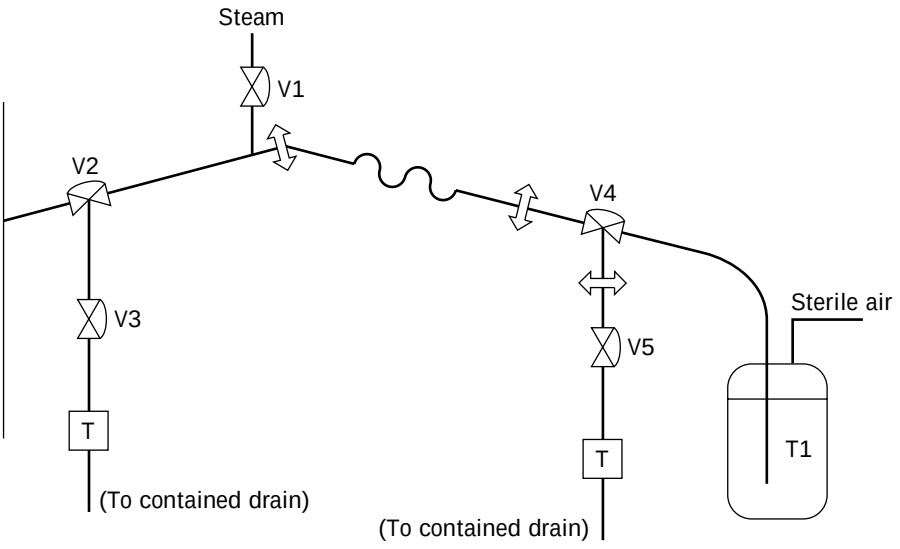


Figure 11. Steam lock assemblies for transfering sterile medium.

steam side of the valves than is possible with other valve types. Variations on this theme are better suited to cases involving long transfer lines or having other special requirements.

Air Inlet and Outlet. Figure 12 is a simplified diagram of a "typical" air inlet/outlet system. The major problems for this system are related mainly to the filters. Chief among these (other than outright failure) is wetting of both filters caused by condensate during sterilization, and wetting and clogging of the exhaust filter by condensate and/or aerosol during fermentation. These problems are not all easy to resolve and should be dealt with most seriously during design. Wetting and/or clogging of the inlet filter by condensate and/or compressor oil during fermentation also can be a problem but is usually the result of not using proper quality air, having no or improper prefilters in the air inlet line, and/or poor maintenance (e.g., of the compressor). All of these are easily remedied.

The extent of condensate problems during sterilization can be remedied best by ensuring that whatever condensate does form can drain freely. This can be complicated by the fact that it is undesirable to have the sterile side of the filter connected directly to the drain line. Another approach is to steam heat the filter housing such that condensation cannot occur. This is effective but has the disadvantages of adding cost and decreasing filter life.

Other factors that can affect condensate problems are the positioning and orientation of the housing. There are differing opinions concerning these; one is best advised to consider the advice the filter and the fermenter vendors for specific cases.

Plugging of the exhaust filter by condensate and aerosols during fermentation require special consideration. Air leaving the fermenter will be essentially saturated with water vapor at fermentation temperature. The exhaust line, filter, and so on usually are colder than the fermenter; therefore, condensation is inevitable. The amount of condensation will depend on temperature differences, air flow rate, and the nature of the surfaces of the components in the exhaust line (the maximum potential is easily calculable). Approaches that have been used to deal with these problems include (70):

- Condensers
- Heat exchangers; used before the exhaust filter to avoid condensation from the fermenter off gases and after the pressure control valve to prevent reflux from the exterior exhaust line
- Steam-heated filter housings
- Heated exhaust lines
- Coalescers
- All of the above

None of these is completely successful, and each has its proponents and detractors. Suffice it to say that each case should be considered on its own merits (e.g., fermentation temperature, duration, and air flow) and that combinations of the approaches should not be ignored.

The condensate problem is exacerbated by the aerosol problem. Aerosols will form in any aerobic fermentation and will carry liquid along with dissolved solids and particulates (including organisms) to the exhaust line. The only questions are how much and what problems they will cause. Among the problems are deposition of dissolved and particulate solids, not only on the filter, but also on heated surfaces or condenser surfaces. Among the problems deposition causes is reduction in heat exchange effectiveness, which reduces the capabilities of the previously mentioned devices to eliminate condensation. In severe cases, exhaust line blockage can lead to other problems related to pressure buildup and/or higher linear velocities in the exhaust line. The problem should be dealt with on a case-by-case basis during process development and fermenter design.

Provisions should be made for in situ integrity testing of sterile filters. This can be done by means of any of several commercial electronic testing devices that are based

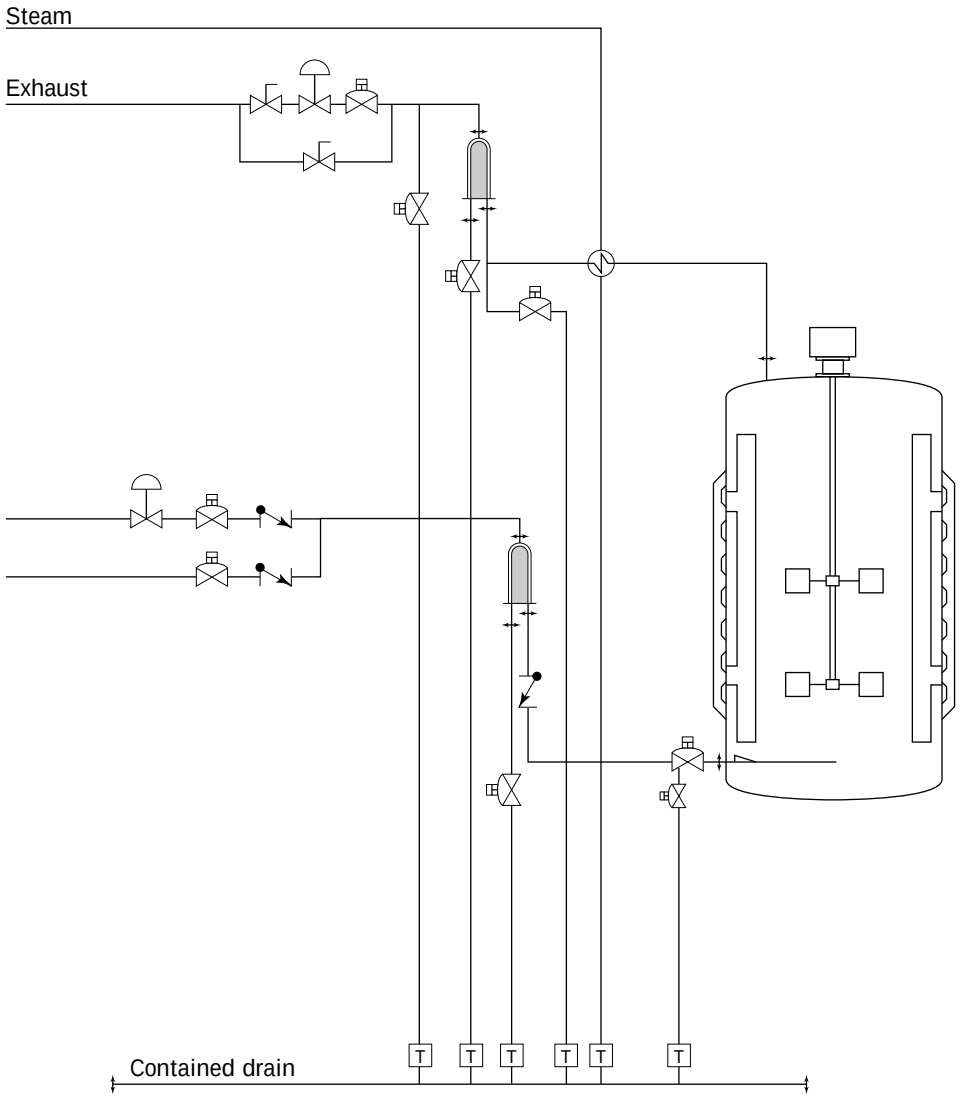


Figure 12. Simplified piping diagram illustrating air inlet and outlet systems.

on liquid intrusion, diffusion, or some other well-established method for testing membrane filter integrity (71–73). One example of the piping necessary to perform the test on an exhaust filter is shown in Figure 13.

Finally, two pressure control valves are required in the case of constant pressure sterilization (continuous flow of steam into the vessel and out to atmosphere). The reason is that the steam flow will be very much lower than the air flow during fermentation; hence, the valve C_v required to maintain steam pressure accurately will be far too small to control air pressure accurately.

Sampling Systems. One example of a simple aseptic sampling system is illustrated in Figure 14. The sample vial, vent filter, and valve V1 are sterilized as a unit in an autoclave. The unit then is connected to the fermenter piping via a sanitary clamp. After the connection is made, steam is introduced via V3 and passes across the nonsterile sides of V2 and V1 and then to trap via V4. Once the valves and line are sterile, the steam and condensate valves are closed. Samples may be taken any time after the sampling system cools. Samples are taken via V1 and V2. Lines and valves are resterilized before the sanitary connection is broken.

The valves selected for this system will depend on the class of service required. We suggest that vessel sampling valve V3 be a flush-mounted diaphragm valve (e.g., Asepco™, NovaSeptic™) designed in such a way as to minimize dead volume and piping lengths; it is an excellent choice for service requiring very high levels of asepsis and cleanability. Less expensive valves can be used for less-demanding service, but we think that the initial extra cost will more than pay for itself. Diaphragm valves are suggested for V1, V3, and V4 in cases of demanding cGMP or containment requirements. This is not required for all classes of service; indeed, it is usually satisfactory to use plug valves for steam and condensate lines even in many demanding circumstances. It should be noted, however, that you may encounter a perception problem if you use them for applications deemed to be sensitive.

There are several other designs that have been used successfully. A few have been designed specifically for systems requiring BL3-LS containment or higher. These are

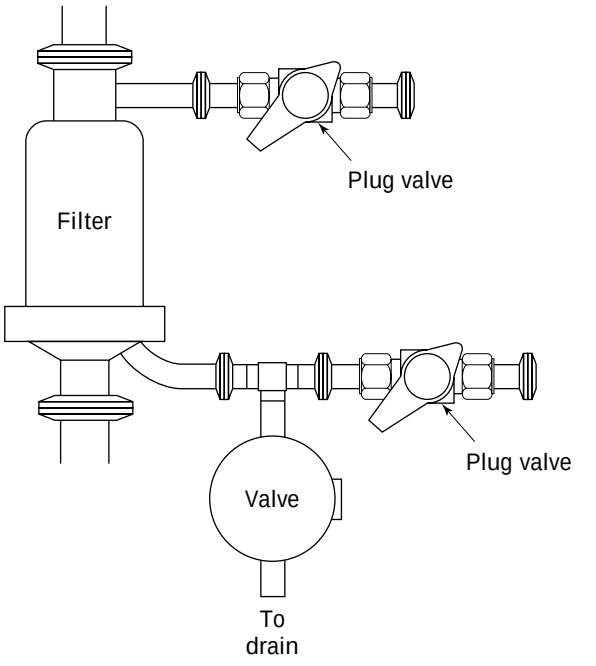


Figure 13. Piping diagram for in situ air outlet integrity test.

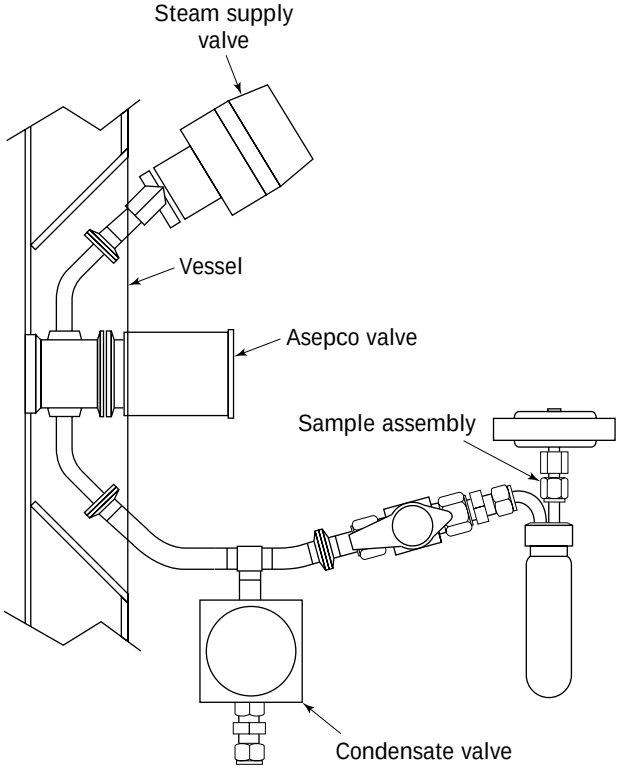


Figure 14. Aseptic sampling system.

discussed later. Finally, please note that addition system design and operation are very similar to those for sample systems and are not discussed here.

Agitation Systems

Agitation can be done by direct mechanical coupling of the shaft to the drive or by magnetic coupling. The latter has been recommended for cases where high levels of containment are required (74). There also is the perception that magnetic drives are more suitable for maintaining more stringent levels of asepsis and cleanliness than is possible with direct drive. Neither claim has any substantive basis; indeed, there is good reason to believe that existing magnetic drives may present greater cleaning difficulties because of the manner in which the driven magnet must be mounted at the bottom of the vessel. In addition, power transfer by magnetic drive is quite low.

Direct drive is the predominant current choice at almost any scale of operation. The major mechanical decisions that must be made in this case (other than power) have been discussed at length elsewhere (4,75). The major arguments for bottom drive are ease of maintenance, shorter shafts, less support structure, and lower overall height. The arguments against bottom (and for top drive) focus primarily on the potential of catastrophic spills resulting from bottom seal failure, seal grinding as a result of broth particulates working into a bottom seal, and greater cleaning difficulties. If the seals are designed and maintained properly, none of these is a problem. We have seldom seen any real difference in aseptic operability between top and bottom drives, and we are reasonably certain that the organisms don't care. The choice probably will continue to be driven primarily by personal preference.

There are very few cases in which double mechanical seals are not (or should not) be used in fermenters. The major debate focuses on seal orientation and the details of individual seal designs. There are basically two orientations used: inline (Fig. 15) and back-to-back (Fig. 16). The details are discussed in Refs. 4 and 75. We recommend inline design because we have found it to operate more cleanly and require a simpler sterilizing/lubricant system (4,75).

Seal lubrication is usually provided by means of sterile steam condensate. It is extremely important that this condensate be free of particulates: their presence guarantees rapid seal failure, contaminated fermentations, and a hyperactive maintenance program. It is also important to note that during sterilization live steam flows through the seal housing. There are some who insist on keeping the steam flowing throughout the fermentation. (Obviously, they have no faith in the seals.) The one thing this will guarantee is much more rapid wearing of the seals (perhaps supporting the lack of faith in the seals).

One must also decide on the means for controlling lubricant flow rate. Most use a valve for this purpose. We suggest an orifice sized to deliver the proper flow. This avoids the cost and maintenance of a valve and insures fiddle-proof operation. It does, however, require the use of particulate-free condensate, but then so does proper operation of the seals. We also recommend including a sight

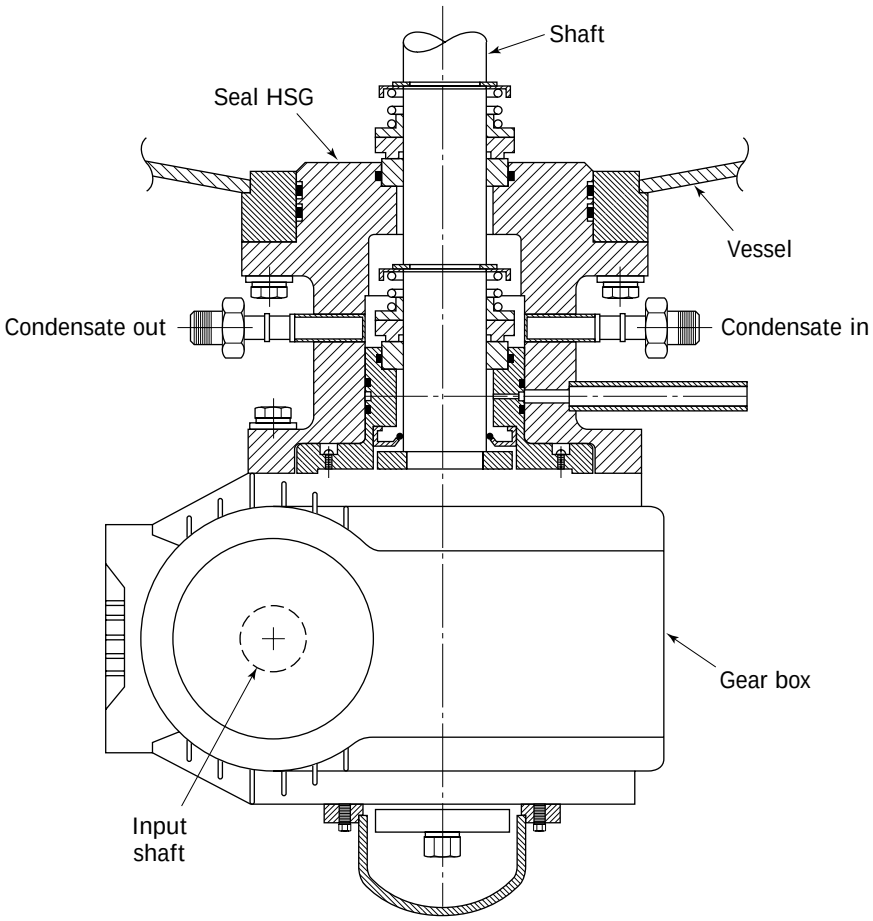


Figure 15. In-line double mechanical fermenter shaft seal.

glass in the lubricant drain line as well as a seal leak detector (Fig. 17). Other more complex detection systems might be considered for specific circumstances (e.g., severe containment requirements).

Finally, preventing shaft vibration is another important factor in agitation system design. Shaft vibration is a safety hazard. It will also cause premature seal failure and other costly mechanical damage. Details have been presented elsewhere (4).

Cleaning Systems

As noted earlier, specific requirements for cleaning depend on several factors including the nature of the fermentation broth. There is a major difference between most microbial broths and most mycelial broths. Single-cell microbial broths, with the exception of those containing a lot of undissolved particulates and high viscosity components (e.g., xanthan broths), tend to be free draining and readily amenable to cleaning based primarily on the physicochemical action of the cleaning agents. Many fungi and mycelial bacteria, on the other hand, tend to cling to fermenter internals and may require mechanical action (e.g., high-velocity jets) in addition to cleaning agents.

The following guidelines are applicable for most systems. These principles must be applied in light of the actual cleaning agents and protocols to be used:

- Eliminate internals and nooks and crannies to the greatest extent possible in the vessel and throughout the piping system. This practice is consistent with design for aseptic operation.
- Drill and position sprayballs to insure complete coverage of all surfaces inside the vessel. This usually requires an empirical approach. Coverage can be tested by means of the riboflavin test. The reader is cautioned, however, that complete coverage is a necessary but not sufficient condition for cleaning. Also, please note the following:
  - Spray balls designed for sanitary operation are self-draining and self-cleaning. They can be sterilized in situ. There is, however, considerable debate concerning whether they should be removed prior to fermentation.
  - High-velocity, rotating devices that may be necessary when large clumps of sticky residue must be removed (e.g., as with a fungus) are not designed to be inherently self-cleaning, self-draining, or sterilizable. They should not be mounted permanently.
- Eliminate deadlegs in piping and deadspaces in valves, fittings, and other system components. This is also consistent with design for aseptic operation.
- Specify the same materials and finishes for process and CIP piping as are specified for the fermenter. Ob-

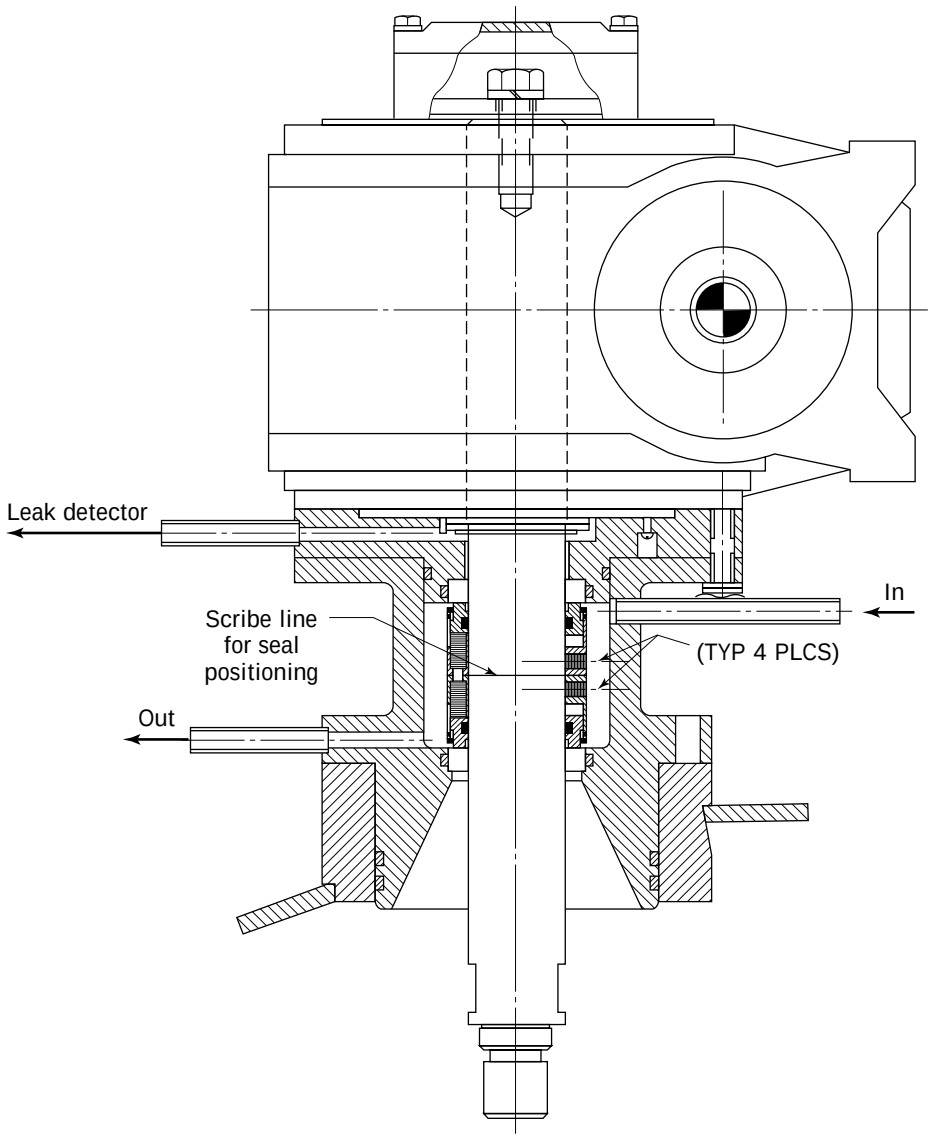


Figure 16. Back-to-back double mechanical fermenter shaft seal.

viously, these must be compatible with the cleansers and conditions used.

- Avoid threaded joints. A completely welded system is best, but compression fittings can be satisfactory when cost is a major consideration, particularly for nonregulated products.
- Make CIP piping as simple as possible. For example, make dual use of process and other piping as much as possible.
- Avoid complex, expensive transfer panels wherever possible. Use swing elbows wherever practical.
- Design and construct the system to facilitate validation and on-going testing for removal of contaminants, including the cleaning agents. This design should provide for swabbing, obtaining rinse samples, or whatever else the cleaning protocol requires.

A portion of one approach to integral CIP piping is illustrated in Figure 18.

There are many cases for which a portable CIP system is preferable to an integral system. For such cases, the fermenter and the portable unit should be designed to allow simple mechanical attachment of the necessary hoses between the two units and with utilities, and a straightforward means for interfacing the instrumentation, logging data, and controlling the systems of the two units.

Finally, we note that classical CIP systems rely on continuous cleanser flow and the maintenance of a shallow puddle in the bottom of the fermenter. To achieve this they rely on special pumping devices such as eductors (74). Aside from the design, control, and other operating problems this causes, it has been observed in some facilities that the stable pool leads to the formation of “cleaning rings” at the bottom of the fermenter. These rings may not be real problems (additional evidence is still required), but they do cause perception problems. We suggest the use of a pulsed-flow system to overcome not only the need for special pumping devices but also the “cleaning ring” problem.

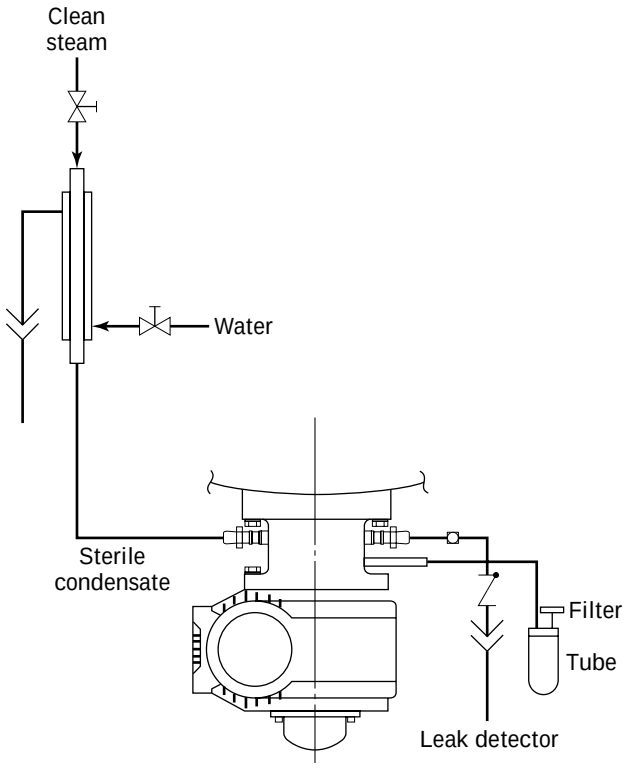


Figure 17. Piping diagram illustrating lubricant flow sight glass and seal leak detector.

Such pulsed systems have been found to accomplish both objectives in practice.

Additional details concerning fluid velocity, vortex elimination, and many other aspects of CIP system design and operation are described elsewhere (59).

Containment

Our scope is limited to containment of fermenters; however, we reiterate here the importance of making certain that the equipment, the facility, and the operating protocols be considered simultaneously to ensure compatibility with regard to containment (just as in the cases of sterility and cleaning). As noted earlier, most new facilities for regulated products are being built to satisfy BL2-LS facility requirements and (in most cases) BL3-LS fermenter containment requirements. With few exceptions (about which there is no agreement) there is no difference with regard to the equipment containment requirements at the two levels—at present.

The guidelines require that BL2-LS/BL3-LS fermenters be designed and constructed to prevent organism release (primary containment) from the vessel and from any of the subsystems noted earlier; hence, they must be built as closed systems. (Note that it also is possible to enclose a fermenter in a biosafety "cabinet," but this is not usually practical or necessary.) Some areas that have been subject to greater scrutiny than has been required for cGMP compliance are as follows. In evaluating the discussion, the reader should bear in mind that there is not much information readily available concerning containment reliability data. Some progress is being made via the Industrial Biosafety Project in the United Kingdom.

- *Static seals*. It has been suggested that double static seals (e.g., in headplates) be used for BL2-LS and that double static seals with a steam trace between the seals be used for BL3-LS (76). There has been little to demonstrate that these are either needed or desirable. Indeed, there have been substantive arguments made against using such devices (76). There is as yet no consensus.
- *Rotating seals*. Arguments have been made in favor of using magnetic drives at anything higher than BL2-LS. This was discussed earlier. There are many who argue that only top drive be considered for BL2-LS and above. Their reasons have mainly to do with avoiding the possibility of large spills in the event of catastrophic seal failure. Such a circumstance is possible but highly unlikely. Leak detectors (which would normally be installed for cGMP requirements) would almost certainly indicate a leak long before there was the remotest possibility of catastrophic failure. One also should consider that a leak in top drive seal might be more difficult to detect and could lead to release of a more insidious nature. Again there is no consensus. Finally, it has been suggested that a low-pressure sensor or flow sensor be used to detect lubricant flow failure. This could help to avoid seal failure (77).
- *Sampling systems*. One system recommended (78) for contained sampling is illustrated in Figure 19. Note that it is similar to the system in Figure 14 recommended for cGMP compliance. It is a bit more complicated and may provide some incremental benefit; however, this might be outweighed by the greater likelihood of operator error. There have also been more complex systems suggested that involve more intricate valving and/or biosafety cabinets. It is apparent that this is an area that needs considerably more development work.
- *Air exhaust system*. It has been suggested that two in-series sterile filters or a sterile filter followed by an incinerator be used at BL2-LS and higher. All other aspects are covered by the earlier discussion of exhaust lines.
- *Pressure relief*. This is an area of some controversy because of the need to satisfy physical and biological safety requirements simultaneously. It has been suggested by many that relief venting be done via a large kill tank protected by a HEPA filter; however, this could be in conflict with physical safety codes that require there be no devices in the relief path (79). We do not have anything approaching universal agreement here; however, it does seem fairly apparent that limiting the supply pressure would help to minimize the risk. There does appear to be agreement on the use of rupture disks rather than pressure relief valves; this is consistent with accepted practice for

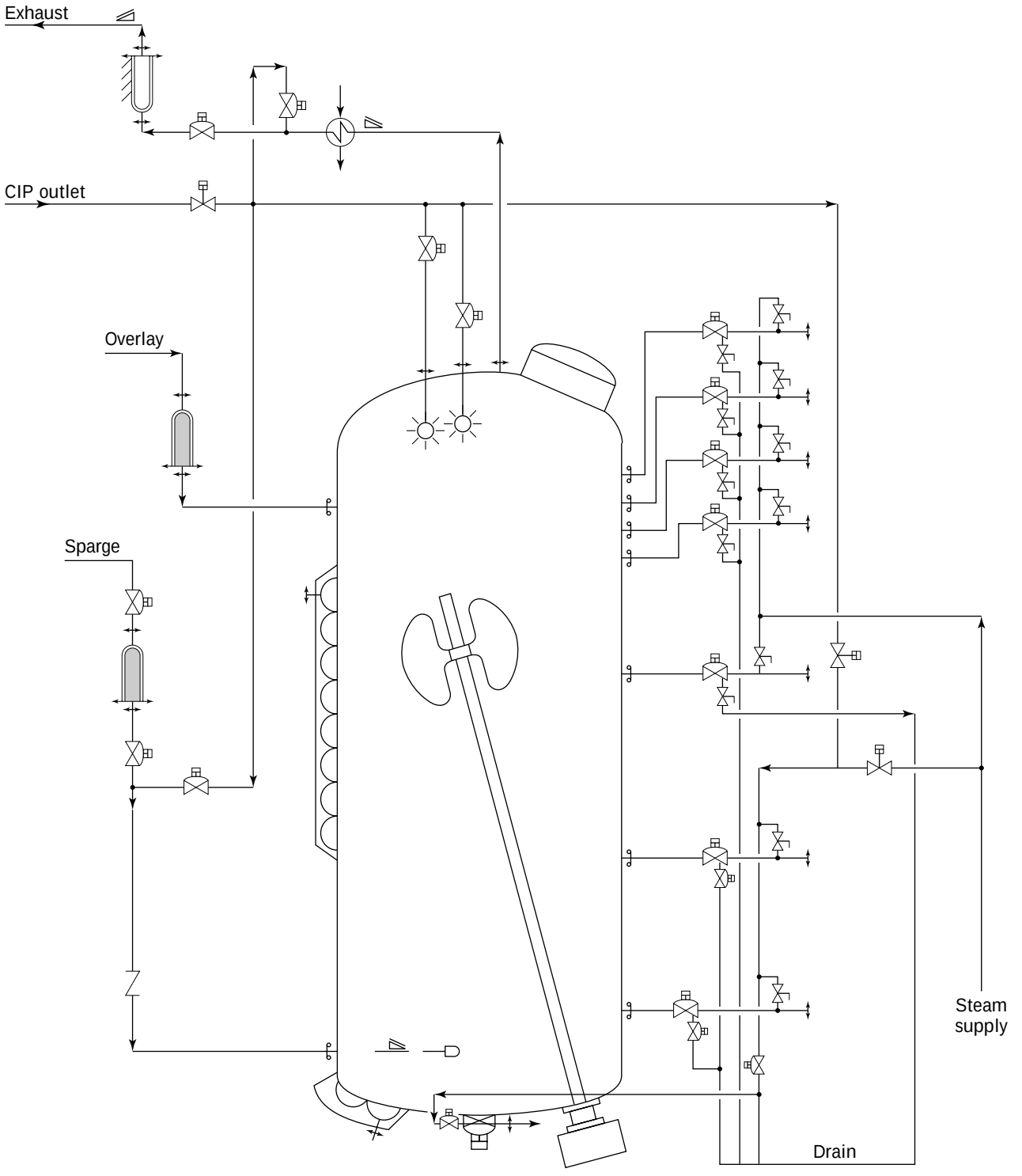


Figure 18. Simplified CIP piping diagram.

cGMP compliance. Rupture disks are cleaner and are not prone to sticking.

Finally, HAZOP evaluations for fermenter containment have been recommended. Such analyses would evaluate potential hazards that might be caused by a wide range of incidents (e.g., fire) not limited to normal operation.

CONCLUSION

Our intent has been to present a balanced view of stirred-tank fermenter design and construction for microbial and and mycelial organisms. There are, however, several important aspects that space and scope considerations prevent us from considering in the depth they warrant. In

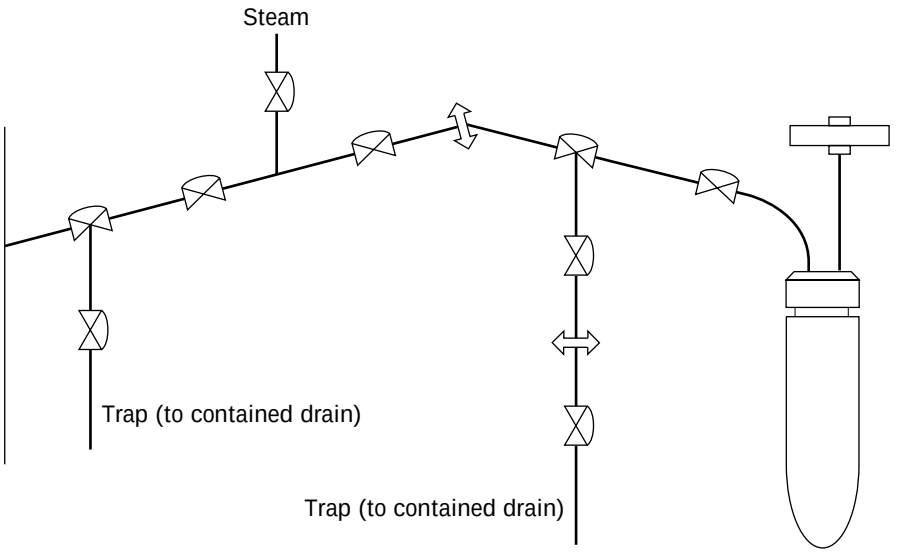


Figure 19. Contained sampling system.
Source: Adapted from Van Houten (78).

particular, we would like to have included more information related to overall process considerations, instrumentation, and control; however, we believe that these are addressed elsewhere in this volume. We trust that we have conveyed the sense that fermenter design is very much an art and that we have convinced you to consider each case separately.

BIBLIOGRAPHY

1. D.I.C. Wang, C.L. Cooney, A.L. Demain, P. Dunill, and A.E. Humphrey, *Fermentation and Enzyme Technology*, Wiley, New York, 1979.
2. M.L. Schuler and F. Kargi, *Bioprocess Engineering: Basic Concepts*, Prentice-Hall, Englewood Cliffs, N.J., 1992.
3. H.W. Blanch and D.S. Clark, *Biochemical Engineering*, Marcel Dekker, New York, 1996.
4. M. Charles and J. Wilson, in B.K. Lydersen, N.A. D'Elia, and K.L. Nelson eds., *Bioprocess Engineering*, Wiley, New York, 1994.
5. S. Yokell, *Chem. Eng.* **93**, 75–83 (1986).
6. S.P. Vranch, in W.C. Hyer ed., *Bioprocessing Safety, Worker and Community Safety and Health Considerations*, STP 1951, ASTM, Philadelphia, 1990, pp. 39–57.
7. C.H. Collins, in C.H. Collins and A.J. Beale eds., *Safety in Industrial Microbiology and Biotechnology*, Butterworth Heineman, Oxford, U.K. 1992, pp. 23–33.
8. Guidelines for Research Involving Recombinant DNA Molecules, *Fed. Regist.* **59**(127), 34496–34547 (July 5 1994).
9. *Management of Health and Safety at Work Regulations*, HMSO Publication Centre, London, 1992.
10. Council Directives on the Protection of Workers from Risks Related to Exposure to Biological Agents at Work (90/679/EEC), *Off. J. Eur. Commun.* **33**, 1–12 (1990).
11. M.C. Flickinger, E.B. Sansone, *Biotech. Bioeng.* **26**, 860–870 (1984).
12. *Design Criteria for Viral Oncology Research Facilities*, DHEW 78-891, Department of Health, Education, and Welfare, Washington, D.C., 1978.
13. M. Dick and E. Hanel, *Design Criteria for Microbiological Facilities, Fort Dietrick, Maryland*, vols. 1 and 2; Dept. of the Army, Technical Engineering Division Project 1B662706A072, 1970.
14. R.S. Runkle and G.B. Phillips, *Microbial Contamination Control Facilities*, Van Nostrand Reinhold, New York, 1969.
15. C.A. Perkowski, in B.K. Lydersen, N.A. D'Elia, and K.L. Nelson eds., *Bioprocess Engineering*, Wiley, New York, 1994, pp. 730–743.
16. E.L. Paul, in W.C. Hyer ed., *Bioprocessing Safety, Worker and Community Safety and Health Considerations*, STP 1951, ASTM, Philadelphia, 1990, pp. 65–73.
17. *Code of Federal Regulations*, Title 29, Washington, D.C. 1996.
18. *Code of Federal Regulations* Title 21, Washington, D.C. 1996.
19. *Points to Consider in the Production and Testing of New Drugs Produced by Recombinant DNA Technology*, Food and Drug Administration Washington, D.C., 1993.
20. *Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products for Human Use*, Food and Drug Administration, Washington, D.C., 1991.
21. *Biotechnology Inspection Guide*, Food and Drug Administration, Division of Field Investigations, Washington, D.C., 1991.
22. *Guide to Inspection of Bulk Pharmaceutical Chemicals*, Food and Drug Administration, Washington, D.C., 1991.
23. *Guidance for Industry for the Submission of Chemistry, Manufacturing, and Controls Information for a Therapeutic Recombinant DNA–Derived Product or a Monoclonal Antibody Product for In Vivo Use*, Food and Drug Administration, Washington, D.C. 1996.
24. *FDA Guidance Document Concerning Use of Pilot Plant Manufacturing Facilities for the Development and Manufacture of Biological Products*, Food and Drug Administration, Washington, D.C., 1995.
25. *Guidelines on the General Principles of Process Validation*, Food and Drug Administration, Washington, D.C., 1987.

26. J.Y. Lee, in I.R. Berry and R.A. Nash eds., *Pharmaceutical Process Validation*, Marcel Dekker, New York, 1993, pp. 573–597.
27. T.J. Naglak, M.G. Keith, and D.R. Omstead, *BioPharm* **7(6)**, 28–36 (1994).
28. R. Baird and P. De Santis, in B.K. Lydersen, N.A. D'Elia, and K.L. Nelson eds., *Bioprocess Engineering*, Wiley, New York, 1994.
29. E.K. White and D.M. Marks, in J.K. Shillenn ed., *Validation Practices for Biotechnology Products*, ASTM STP 1260, ASTM, Philadelphia, 1995.
30. M.G. Beatrice, in G. Stephanopoulos ed., *Biotechnology: Bioprocessing*, vol. 3, VCH Weinheim, 1993.
31. L. Tuanchi and D. Yu, *Biotechnol. Bioeng.* **42**, 777–784 (1993).
32. L. Leong-Poi and D.G. Allen, *Biotechnol. Bioeng.* **40**, 403–412 (1992).
33. Y. Kawase and T. Kumagai, *Bioprocess Eng.* **7**(1), 25–28 (1991).
34. M. Charles, *Adv. Biochem. Eng.* **8**, 1–62 (1978).
35. B.C. Buckland, K. Gwebonyo, D. DiMasi, G. Hunt, G. Westerfield, and A.W. Nienow, *Biotechnol. Bioeng.* **31**, 737–742 (1988).
36. M. Reuss, in G. Stephanopoulos ed., *Biotechnology: Bioprocessing*, vol. 3, VCH, Weinheim, 1993.
37. Y. Kawase, B. Halard, and M. Moo-Young, *Biotechnol. Bioeng.* **39**, 1133–1140 (1992).
38. J.M.T. Vasconcelos and S.S. Alves, *Chem. Eng. J.* **47**, B35–B44 (1991).
39. D.S. Dickey, *72nd AIChE Ann. Meeting*, San Francisco, November 25–29, 1979.
40. A. Bakker, J.M. Smith, and K.J. Myers, *Chem. Eng.* **101**(12), 98–104 (1994).
41. C.M. McFarlane, X.-M. Zhao, and A.W. Nienow, *Biotechnol. Prog.* **11**, 608–618 (1995).
42. G.J. Balmer, I.P.T. Moore, and A.W. Nienow, in C.S. Ho and J.Y. Oldshue eds., *Biotechnology Processes, Scale-Up and Mixing*, AIChE, New York, 1987, pp. 117–127.
43. N.M.G. Oosterhuis and N.W.F. Kossen, *Biotechnol. Lett.* **3**(11), 645–650 (1981).
44. C.J. Geankopolis, *Transport Processes and Unit Operations*, Allyn and Bacon, Boston, Mass., 1983, pp. 154–157, 169–170.
45. A.W. Nienow, *Trends Biotechnol.*, **8**, 224–233 (1990).
46. J.Y. Oldshue, *Fluid Mixing Technology*, McGraw-Hill, New York, 1983.
47. M. Yasukawa, M. Onodero, K. Yamagiwa, and K. Ohkawa, *Biotechnol. Bioeng.* **39**, 629–636 (1991).
48. D.G. Cronin, A.W. Nienow, and G.W. Moody, *Trans. IChemE.* **72** (part C), 35–40 (1994).
49. C.L. Cooney, D.I.C. Wang, and R.I. Mateles, *Biotechnol. Bioeng.* **11**, 269–281 (1968).
50. R. Pollard and H.H. Topiwala, *Biotechnol. Bioeng.* **18**, 1517–1535 (1976).
51. P. Mohan, A.N. Emery, and T. Al-Hassan, *Trans. IChemE.* **70** (part C), 200–204 (1992).
52. S. Aiba and K. Toda, *J. Ferm. Tech. (Japan)* **43**, 527 (1965).
53. U. Pflug and R.G. Holcomb, in S.S. Block ed., *Disinfection, Sterilization and Preservation*, Lea and Feibiger, Philadelphia, 1991.
54. G.K. Raju and C.L. Cooney, in G. Stephanopoulos ed., *Biotechnology: Bioprocessing*, vol. 3, VCH, Weinheim, 1993.
55. J.H. Young, *Biotechnol. Bioeng.* **42**, 125–132 (1993).
56. J.H. Young, W.C. Lasher, and R.P. Gaber, *Bioprocess. Eng.* **12**, 293–304 (1995).
57. T. Oakley, in B.K. Lydersen, N.A. D'Elia, and K.L. Nelson eds., *Bioprocess Engineering*, Wiley, New York, 1994, pp. 500–521.
58. *Guide to Inspection of Validation of Cleaning Processes*, Food and Drug Administration, Division of Field Investigations, Washington, D.C., 1993.
59. R. Brunkow, D. DeLucia, S. Haft, J. Hyde, J. Lidsay, J. McEntire, R. Murphy, J. Myers, K. Nichols, B. Terranova, J. Voss, and E. White, in *Cleaning and Cleaning Validation: A Biotechnology Perspective*, PDA, Bethesda, Md., 1996, pp. 41–62.
60. S. Lombardo, P. Inampudi, A. Scotton, G. Ruezinsky, R. Rupp, and S. Nigam, *Biotechnol. Bioeng.* **48**, 513–519 (1995).
61. C.P. Dillon, D.W. Rahoi, and A.H. Tuthill, *BioProcessing* **8**(5) 32–35, **8**(7), 31–33 (1992).
62. J. Villafranca and E.M. Zambrano, *Pharm. Eng.* **5**(6), 28–30 (1985).
63. J. Butters and A. Reynolds, *Process Eng.* **71**(2), 65–79 (1990).
64. K.B. Balmer and M. Larter, *Pharm. Eng.* **13**(3), 21–28 (1993).
65. D.C. Coleman and R.W. Evans, *Pharm. Eng.* **10**(2), 43–49 (1990).
66. D.C. Coleman and R.W. Evans, *Pharm. Eng.* **11**(4), 75–79 (1991).
67. R.E. Markovitz, *Chem. Eng.* **78**(11), 156–163 (1971).
68. U.S. Pat. 4,670,397 (July 2, 1987), E.H. Wegner and H.R. Hunt (to Phillips Petroleum).
69. C.A. Perkowski, *Biotechnol. Bioeng.* **26**, 857 (1984).
70. M.A. Fogglesong, in W.C. Hyer ed., *Bioprocessing Safety, Worker and Community Safety and Health Considerations*, STP 1951, ASTM, Philadelphia, 1990, pp. 14–19.
71. T.H. Meltzer, M. Jornitz, P.J. Waibel, *Pharm. Technol.* **18**(9), 76–84 (1994).
72. S.F. Emory, *Pharm. Technol.* **13**(9), 68–77 (1989).
73. J.J. Errico, in F.J. Carleton and J.P. Agalloco eds., *Validation of Aseptic Pharmaceutical Process*, Marcel Dekker, New York, 1990, pp. 427–471.
74. P.D. Walker, T.J. Narendranathan, D.C. Brown, F. Woolhouse, and S.P. Vranch, in M.S. Verall and M.J. Hudson, eds., *Separations for Biotechnology*, Ellis Horwood, Chichester, U.K., 1987, pp. 469–479.
75. J.D. Wilson and T.E. Andrews, *Biotechnol. Bioeng.* **25**, 1205–1214 (1983).
76. N.J. Titchener-Hooker, P.A. Sinclair, M. Hoare, S.P. Vranch, A. Cottman, and M.K. Turner, *BioPharm* **6**(8), 32–37 (1993).
77. S.R. Miller and D. Bergmann, *J. Ind. Micro.* **11**, 223–234 (1993).
78. J. Van Houten, in W.C. Hyer ed., *Bioprocessing Safety, Worker and Community Safety and Health Considerations*, STP 1951, ASTM, Philadelphia, 1990, pp. 91–100.
79. G. Leaver, in P. Hambleton, J. Melling, and T.T. Salusbury, eds., *Biosafety in Industrial Biotechnology*, Blackie Academic and Professional, London, 1994, pp. 213–239.

FILTER AIDS

TONY HUNT
Advanced Minerals Corporation
Santa Barbara, California

KEY WORDS

Celpure
Diatomite
Extractables
Microparticulate removal
Solid–liquid separation

OUTLINE

INTRODUCTION

Among the most capable and attractive means of microparticulate (greater than 0.1 μ) separation used in biopharmaceutical processing today is the use of highly porous powdered media in dynamic systems. Powdered media, often called filter aids (Fig. 1), provide versatility, high solids-loading capacity, high product recovery, low cost, and ease of scale-up in any filtration process. Moreover, recent technical advances have stimulated a much greater breadth of applications that use this media filtration in biopharmaceutical processes, with a concomitant exponential increase in product innovation.

PROCESS CLARIFICATION

The first unit operation in downstream processing is clarification. It involves the removal of cells, cell debris, or precipitated components from a fermentation broth or process supernatant. Four types of solid–liquid separations are common to the industry: (*1*) The removal of whole cells in fermentation broths where the product is expressed into the supernatant; (*2*) The removal of cell debris from process broths after cell disruption and extraction to release the product of interest; (*3*) The removal of precipitated contaminants (usually proteins) from a process fluid, for example, in the removal of misfolded forms of recombinant proteins; and (*4*) The selective precipitation of the target protein and the capture of this precipitate from the process supernatant.

Most processes use either centrifugation or filtration for solid–liquid separations. Centrifugation, although widely accepted, has some disadvantages in biopharmaceutical applications. The challenge in many applications is to remove both cellular and subcellular debris without applying excessive shear forces on the solids. Centrifuges depend on differences in density and on the centrifugal force applied for solid–liquid separation. However, centrifugation becomes less practical when the product and waste possess similar densities. In addition, centrifuges can be difficult to maintain, clean, and sterilize. Finally, it is important to avoid the generation of aerosols when dealing with biological fluids, and this is often difficult to avoid when using centrifuges for this unit operation.

Depth filters containing diatomite, perlite, and cellulose are suited to low solids applications, because these products tend to blind quickly when subjected to moderate to high levels of solids. Their capacity can be increased by the addition of filter aid (as body feed) into the unfiltered broth, thus extending the lifetime of the pad.

The use of powdered media offers many advantages in biopharmaceutical applications. These systems are much more flexible and dynamic because powdered media addition can be adjusted based on changes in incoming solids level. In addition, they are extremely efficient at removing suspended solids and are particularly effective at clarifying supernatants heavily contaminated with colloids and small particles. The media can be used in all four applications described above, including the selective precipitation of the target protein and the capture of this precipitate from the process supernatant. Once captured on the filter media, the protein or product of interest can be redissolved by adjusting the pH or salt concentration in the wash buffer and recirculating this solution through the filter media bed. After resolubilization, the product can be eluted from the media with approximately two bed volumes of buffer.

Capital costs for powdered-media-based systems are relatively low when compared with centrifugation or other filtration systems. Because the media are disposable, media cleaning and lifetime studies do not need to be performed, hence the cost of validating the filtration process is substantially reduced. In addition, the ability to postwash the filter cake helps to maximize product recovery, a process that is extremely expensive or impractical with centrifuges. Finally, with improvements in both powder

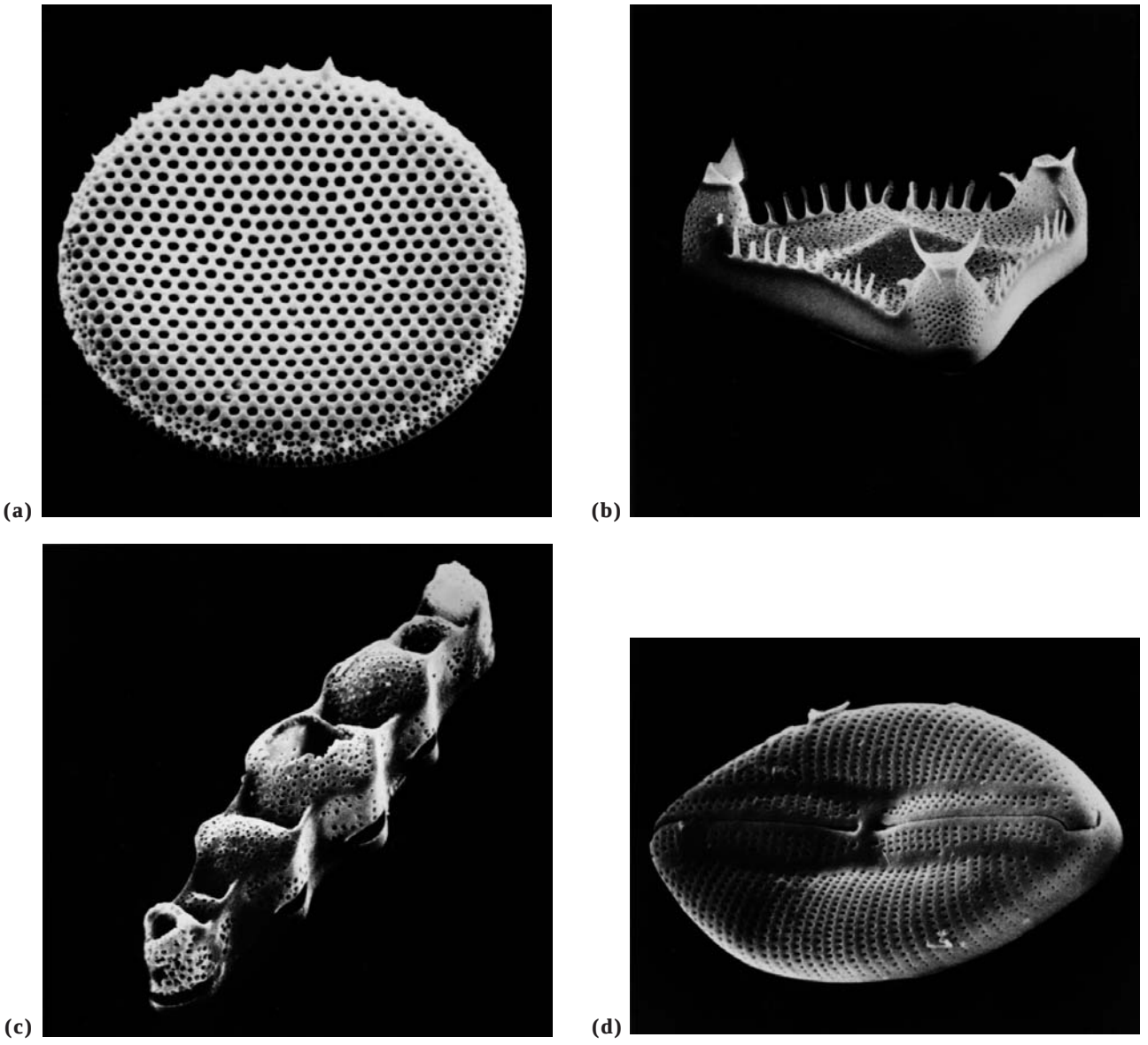


Figure 1. Size and shape of diatoms commonly found in diatomaceous earth deposits.

containment and filter equipment design, complete containment is now easily achieved both in the media preparation area and in GMP suites.

The combination of these improvements in powder containment, filter design, and powdered media technology position filter aids as a very cost-effective and economical approach to clarification.

POROUS MEDIA IN DYNAMIC PROCESS FILTRATIONS

Diatomite, perlite, and cellulose are the most widely used porous media in dynamic process filtrations, with a high percentage of applications using diatomite.

Diatomite

Diatomite is often used in combination with cellulose; however, diatomite is often the key ingredient of static or fixed-bed filters such as filter sheets and filter pads. Diatomite products are especially characterized by an inherently intricate and highly porous structure composed primarily of silica. These products are obtained from diatomaceous earth, a sediment greatly enriched in biogenic silica in the form of siliceous frustules of diatoms, a diverse array of microscopic, single-celled golden-brown algae of the class Bacillariophyceae. Surprisingly, these frustules are sufficiently durable to retain much of their ultrastructure nearly intact through long periods of geologic time when preserved in conditions that maintain chemical equilibrium.

Perlite

Perlite is a naturally occurring volcanic glass that thermally expands upon processing. After milling, porous, com-

plicated structures result (Fig. 2) that are also useful in dynamic process filtrations of biopharmaceuticals. Because its structure is not as intricate as that of diatomite, perlite is better suited to the separation of coarse microparticulates from liquids having high solids loading. Like diatomite, perlite is also useful as a functional filtration component of filter sheets and filter pads.

Cellulose

Cellulose, like perlite, possesses a less intricate structure than diatomite. As a result, the use of cellulose is generally limited to coarse filtrations or in specialized applications where a fibrous precoat on the septum is required.

The rest of this chapter will focus attention on the use of diatomite in solid–liquid separations.

FUNDAMENTAL PRINCIPLES OF DIATOMITE FILTRATION

A discussion on solid–liquid separations can cover a wide range of techniques and applications, but each method is often defined by the relative size of solids being removed or collected from a given process stream. This chapter will

Figure 2. Scanning electron micrograph of perlite.

focus on microparticulate filtration, where the solids are typically greater than 0.1 μ. The key objective is to remove unwanted solids from a process stream, so that the refined solution (filtrate) is adequately clarified and qualifies for further downstream processing.

Solids Composition

In general, solids can either be rigid or compressible in nature. Solids that form rigid cakes have some degree of inherent permeability, dictated by particle size distribution and packing arrangement, but most solids in biopharmaceutical applications are typically compressible. These often gelatinous, highly compressible solids retain a degree of permeability if low flow rates and low differential pressures are used for filtration. Unfortunately, these characteristics are contrary to typical processing requirements where this unit operation needs to be completed quickly and economically.

Dynamic Filtration

The problem in static fixed-bed filtration is that unwanted solids have very limited permeability when they collect or accumulate onto the filter septum (pad, paper, or fabric or metallic or plastic woven wire screen). If these solids build up on a septum, then eventually these solids will lack sufficient permeability for fluid drainage and filtration terminates. This behavior highlights a mistake frequently applied in filtration practice; often, higher differential pressure does not guarantee faster filtration rates precisely because these solids collapse to form an impermeable cake.

Introduction of a filter medium changes the composition of the accumulated cake, and therefore its filtration behavior. Powdered media (filter aids) are essential to this technology because they provide two functions: (*1*) as a precoat applied before the start of a filtration cycle and (*2*) as body feed added to the unfiltered feed throughout the filtration cycle. The precoat layer protects the filter septum, preventing penetration and blinding by the unwanted solids. This layer also facilitates septum cleaning and provides for a high-quality filtrate from the beginning of the filtration cycle.

The addition of powdered media (as body feed) to the process feed increases permeability in the accumulating filter cake, restricts solids movement, provides channels for filtrate recovery, and extends cycle length (Fig. 3). The extended permeability afforded by these media slows the rise in differential pressure for constant flow processes and retards the drop in flow for constant pressure operations. At the conclusion of a filtration cycle, product recovery can be maximized by rinsing or purging the accumulated cake in place. It is extremely important to maximize the body feed in a process. By adding too little, the solids quickly blind the filter cake and the filtration is terminated. Adding too much is not only wasteful but leads to a rapid increase in filter-cake thickness, with a corresponding increase in resistance to flow and a subsequent reduction in cycle length (1).

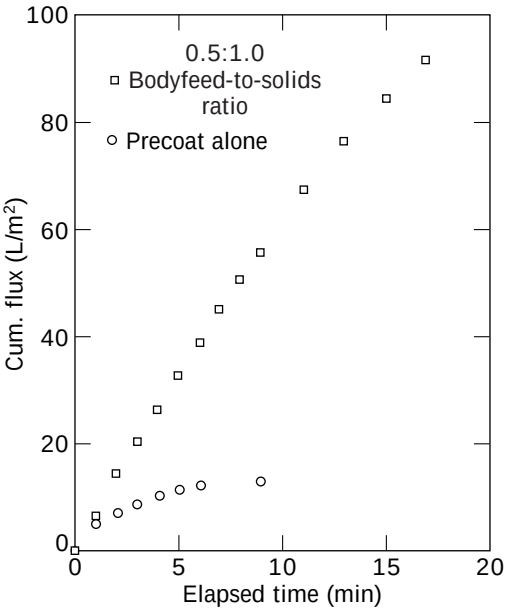


Figure 3. Addition of body feed to a filtration process extends the filtration cycle length.

Filtration Theory Overview

The following section is not intended to be a comprehensive treatment of filtration theory, but rather an overview to describe how the permeability of an accumulated filter cake (solids and filter aid) ultimately affects filtration performance (rate). A thorough discussion of filtration mechanics can be found in numerous reviews (2–4).

Any discussion on porous media filtration can start with a simplified form of the Darcy equation (equation 1), which is used to describe laminar (nonturbulent) fluid flow through a homogeneous, porous medium.

$$k = \frac{Q \times \Delta x \times \mu}{A \times \Delta P} \tag{1}$$

The relative permeability (k) is measured in darcy units, and it is described in terms of the instantaneous volumetric flow rate (Q) and viscosity (μ) of the fluid, the thickness (Δx) and cross-sectional area (A) of the porous medium (accumulating filter cake), and the pressure drop or differential pressure (ΔP) associated with flow. The flow rate (Q) can also be represented by the differential change in volume over time (dV/dt). A value of 1 darcy corresponds to the permeability through a filter media 1-cm thick that allows 1 cm^3 of fluid with a viscosity of 1 cP to pass through an area of 1 cm^2 in 1 s under a pressure differential of 1 atm (i.e., 101.325 kPa). Although fluid viscosity varies with temperature and shear forces, values for it and the other variables can be obtained from tables and experimental observations. Consequently, the permeability (k) is an empirical value (constant) used to characterize the filter cake composition and not a property that can be readily altered by changing the process flow rate or differential pressure. In practice, permeability may appear to be influenced by

flow rate or pressure, but those responses are strictly consequences of composition, and they are often inconsistent with the behavior described by equation 1. The type and relative usage rate of filter aid coupled with the nature and concentration of solids define composition, which in turn controls permeability and filtration performance.

The constancy of k explains why (*1*) the differential pressure increases as filter cake accumulates (Δx) for constant flow (Q_0) processes (equation 2)

$$k = \frac{Q_0 \times \mu}{A}\,(\text{constant}) \times \frac{\Delta x}{\Delta P} \tag{2}$$

and (*2*) flow decreases as cake accumulates for constant pressure (ΔP_0) processes (equation 3)

$$k = \frac{\mu}{A \times \Delta P_0}\,(\text{constant}) \times Q \times \Delta x \tag{3}$$

Permeability measurements are not essential in the application of dynamic filtration technology. In fact, once the unfiltered feed is introduced to the filter, the overall permeability (as measured in darcies) of the filter cake can be reduced by orders of magnitude. Ultimately, the manipulation of permeability (k), through filter aid selection and usage, describes the optimization of this solid–liquid separation process.

As an example, initial (laboratory scale) testing develops enough information to determine the permeability (k_{app}) in a constant pressure application. In fact, a key laboratory test is to determine the elapsed time (t_s) it takes to produce the maximum cake thickness (Δx_f; limited by the physical capacity of the filtration equipment, typically 3 cm) at optimum filter aid usage and terminal pressure (ΔP_f). Hence, the Darcy equation can be put to better use for process scale-up (equation 4):

$$[k_{app}] = \frac{Q_{0\ lab} \times \Delta x_f \times \mu}{A_{lab} \times \Delta P_f} = \frac{Q_{0\ plant} \times \Delta x_f \times \mu}{A_{plant} \times \Delta P_f} \tag{4}$$

Notice that differential pressure (ΔP_f) and filter cake thickness (Δx_f) need not be affected by the scale of operations. Differences in allowable cake thickness (Δx_f) must be accommodated by adjusting filtration area (A_{plant}) requirements. A physical capacity test also determines the total volume of process feed that can be filtered on the given area. In the simplest case of equal cake thickness, scale-up becomes a matter of sizing filter area to accommodate production rate (equation 5).

$$\frac{Q_{0\ lab}}{A_{lab}} = \frac{Q_{0\ plant}}{A_{plant}} \tag{5}$$

GRADE SELECTION AND OPTIMIZATION

An optimized filtration process maximizes throughput and product recovery and minimizes pressure drop while maintaining the desired filtrate clarity, all in a reasonable time. There are five main parameters one needs to consider when selecting an appropriate grade of powder media for a given process.

Product Stability

Careful attention should be paid to soluble metals when selecting a grade. High concentrations of soluble metals in reagents can lead to contamination of biopharmaceutical products, for example, aluminum contamination of albumin (5–7). Also, high concentrations of soluble metals are known to oxidize proteins and enzymes and can activate proteases in fermentation broths (8–10). Therefore, the use of high-purity reagents for filtration will help reduce or eliminate product contamination and degradation issues further downstream.

The industry demand for high purity and high performance reagents has led to the development of a new generation of powdered media. These products, sold under the trade name Celpure™ (Advanced Minerals Corporation, Calif.) offer greater filtration capacity with a corresponding reduction in powdered media consumption and reduction in disposal costs (Fig. 4).

Filtration processes with Celpure grades typically use less media because of a combination of the higher solids loading capacity and improved flow properties of the media. This results in a reduction in overall processing times compared with conventional grades of diatomaceous earth, such as the widely used Celite® grades (Fig. 5). Finally, the products are extremely pure and have very low levels of extractable metals such as aluminum and iron (Table 1) and correspondingly low electrical conductivities.

Filtrate Clarity

Most processes have a clarity specification that needs to be met or exceeded. Achieving stringent clarity specifications can extend the life of downstream filters and protect chromatography columns. The filtrate clarity achieved is dictated by the grade selected and the nature of the turbidity removed. Once the grade is chosen, its level of usage (body feed addition) combined with the available differential pressure to induce flow will control the volume of unfiltered feed that can be processed by a given filtration area.

Product Throughput

Product throughput and filtrate clarity are tightly linked when it comes to grade selection. The goal is to select a grade that achieves the desired clarity and maximizes throughput. By selecting a grade that is too fine, the clarity specification can be exceeded, but the throughput rate may be extremely low with correspondingly high differential pressures.

Operating Pressure

Actual operating pressure can be limited by shear sensitivity of the product or feedstream by equipment constraints or by overall plant design. Many processes have mechanical pressure limits of approximately 30 to 45 psi or 2 to 3 bar.

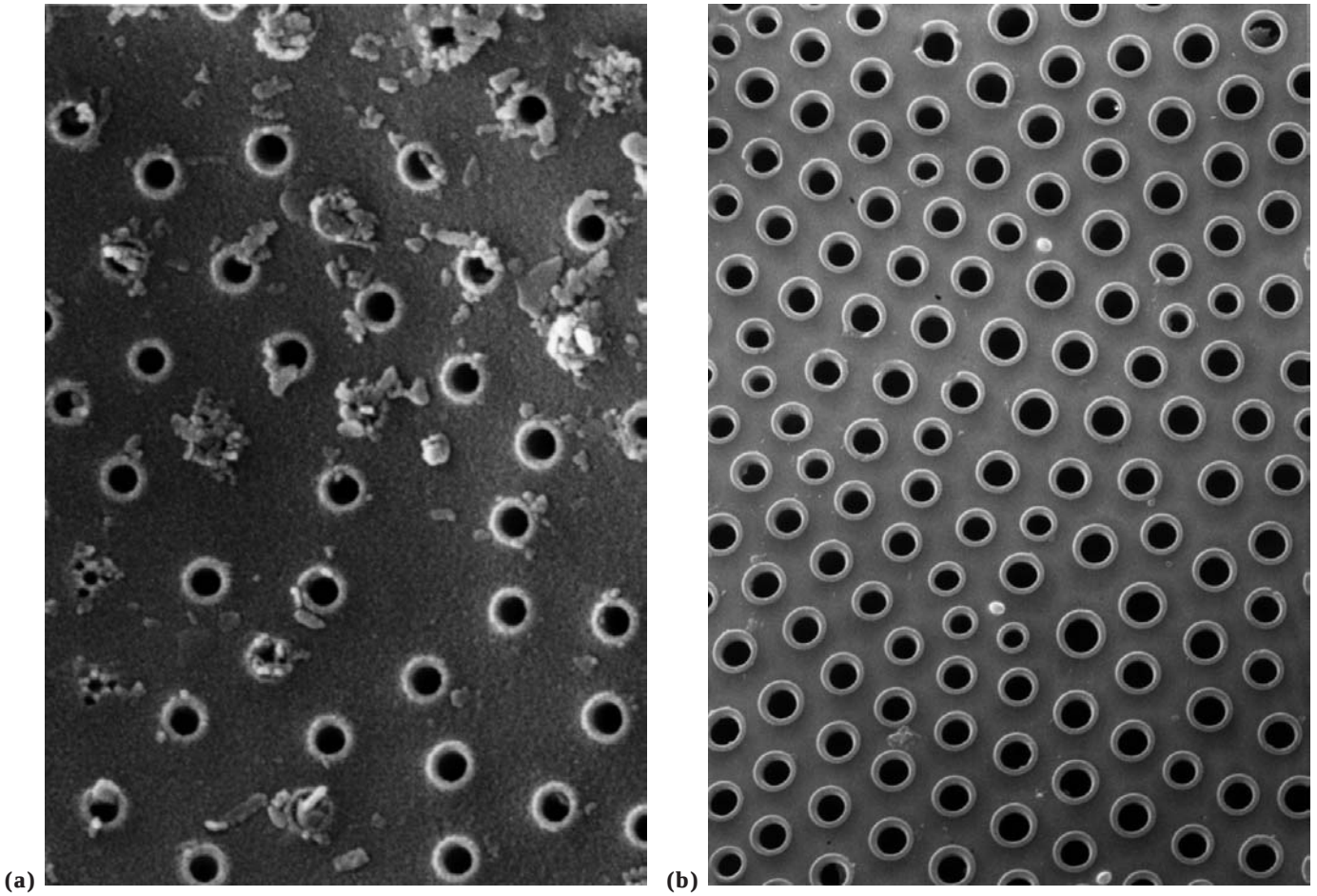


Figure 4. Scanning electron micrograph of the surface of an individual diatom from a conventional Celite grade (**a**) and a Celpure grade (**b**).

Product Recovery

This can be sometimes overlooked in initial filtration studies. At the conclusion of any filtration cycle, the accumulated solids should be washed with a product-compatible buffer to maximize product recovery. Upon completion of the filtration cycle (including rinse), it is important to confirm that the product of interest does not interact with the filtration media. Product recovery issues can often be improved by selecting a more permeable grade of filter aid. As you increase the porosity of a filter aid, the available surface area decreases and any nonspecific interaction decreases. All these factors are tightly linked; therefore, we recommend a systematic approach to grade selection and optimization.

SYSTEMATIC METHODS DEVELOPMENT APPROACH TO GRADE SELECTION

As stated above, the overall goal is to achieve optimum clarity and maximum throughput for a process while minimizing pressure and product losses. The following is a systematic approach to grade selection. Figure 6 shows a typical experimental setup for laboratory-scale constant rate filtration experiments. To perform constant pressure filtrations, a pressurized vessel is required to deliver unfiltered feed to the filtration housing.

Determine the Percent of Solids in Your Feedstream

Any information on distribution and size of feed solids is useful in grade selection. If this information is not available, a recommended procedure is outlined below.

Determine percent solids by weight in your fermentation broth. Filter a small aliquot onto a preweighed filter disk, dry in an oven, and reweigh to determine percent solids in unfiltered feed.

1. Use a Whatman 934-AH filter paper as a septum with the filtration housing. This is an open-pore septum and enables you to precoat without having to custom make a wire screen.
2. Resuspend 1.5 g of media in 100 to 150 mL of buffer (similar in pH to your sample solution; 1.5 g of most grades gives a precoat of approximately 2 mm on 20 cm^2 of surface area). Start with the coarsest grade, Celpure™ 3000 or Celite® 545 AW. By starting with the coarsest grade, you will always achieve maxi-

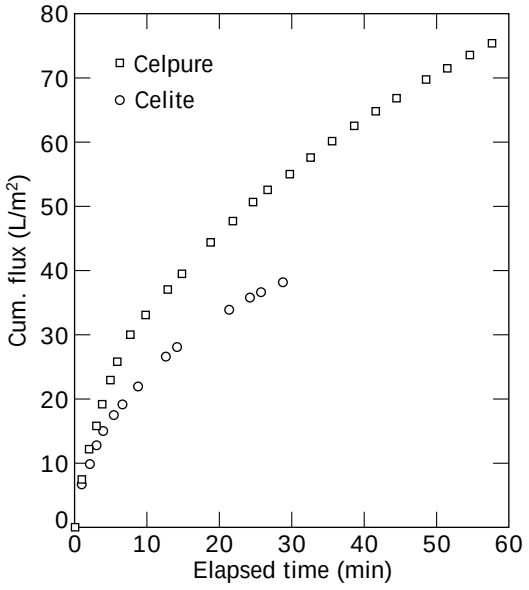


Figure 5. Equal weight comparison of Celpure and Celite grades in the clarification of a fermentation broth with 0.5:1.0 body-feed-to-solids ratio. Runs were terminated at minimum (equal) flow for reasonable production.

Table 1. Typical Celpure™ Properties (Applicable to All Grades)

Property	Value
Soluble aluminum (mg Al/kg)	
Extraction in fermentation broth, pH 4.5.	<3
Soluble iron (mg Fe/kg)	
Extraction in fermentation broth, pH 4.5	<3
Conductivity in H_2O (μS-cm)	<12.5

mum throughput and optimal clarity for a given process.

3. Recirculate the solution until filtrate is clear (typically 2 to 3 mins). The recommended flux rate is 40 L/m^2/min. On a 20-cm^2 surface area, this is equivalent to 80 mL/min.
4. Once the precoat has been formed, reduce the volume of buffer in the chamber until 2 to 3 mm above surface of precoat.
5. Add media to the unfiltered feed solution. As a starting point match the percent of solids in your feed with powdered media addition (2% solids by weight, add 2% by weight of powdered media). Add product/media solution to filter chamber at a flux rate of 1 to 5 L/m^2/min. On a 20-cm^2 surface area, this is equivalent to 2 to 10 mL/min.
6. Measure initial flow rate and monitor both flow rate and pressure throughout the coarse of the filtration. Allow the pressure to rise to 30 psi or 2 bar (or maximum system pressure), and maintain this pressure for the duration of the filtration by reducing flow or until flow rate reaches 20% of initial flow.
7. At the end of the filtration cycle, wash the filter bed with two bed volumes of compatible wash buffer. This will maximize the product recovery for a process. Finally, measure the total volume processed, the time of filtration, filtrate clarity, and the bed height in the filter housing. Save a sample for product recovery analysis.

If clarity is not acceptable, then repeat steps 1 to 7 with a less permeable grade of filter aid. Powdered media are rated based on their Darcy permeability. Celpure 3000 and Celite 545 AW are considered the coarsest grades and Celpure 65 and Celite Filter-Cel® the finest.

For constant pressure filtrations, follow above procedure through step 4. Then switch to a pressurized vessel containing unfiltered broth mixed with appropriate grade of powdered media. Ramp the pressure up to the desired output and monitor decay in flow rate. Repeat experiments with different grades as outlined above.

Body Feed Optimization

Once grade selection is completed, the next step is to optimize the body feed addition. The level of body feed addition controls the permeability and volume of accumulating cake. Higher pressure does not mean greater throughput, especially when body feed usage coupled with the handling characteristics of the solids dictate filter cake

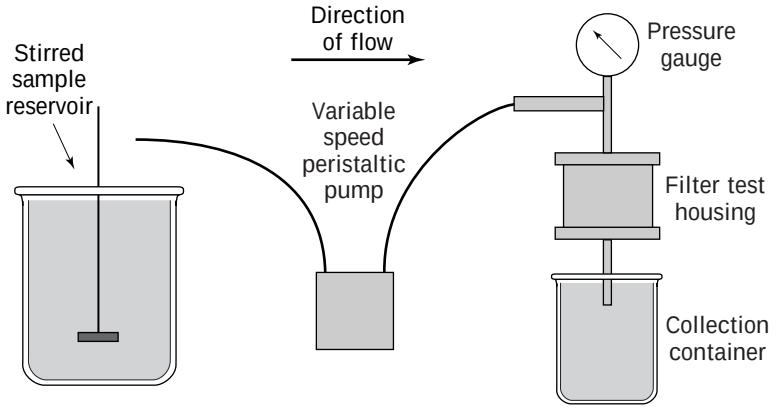


Figure 6. Experimental setup for constant rate experiments.

compressibility. The dilatant nature of diatomite resists cake compression and thereby retains permeability. Feed properties (pH, viscosity, etc.) also affect filtration performance; therefore, these factors must be considered in process development and optimization studies.

Past work with biological solutions at high solids levels (greater than 5% by weight) suggests that constant flux (filtration) rates above 10 L/m^2/min exhibited steep rises in differential pressure because of compression and subsequent loss of cake permeability. Modest flux rates (5.0 L/m^2/min) extended filtration cycles, often producing better economy. Although dilution using a compatible solution can reduce shear and also make filtration easier, the goal to expedite processing implies that it will be used only after others options have failed.

As described earlier, an excellent starting point is to match the percent solids in the process stream on a weight basis with percent media also on a weight basis. If this ratio of 1:1 is not sufficient, i.e., the filtration terminates quickly because of a rapid rise in pressure and subsequent blinding of filter cake with solids, then consider increasing your body feed addition. Repeat the above steps using body feed additions of 1.5:1 or 2:1.

If a body feed addition of 1:1 allows you to process your batch with little or no rise in differential pressure, then you can optimize the body feed by reducing the ratio of powdered media to solids. Reduce the body feed until an acceptable throughput rate is achieved (Fig 7). This approach will allow you to size an appropriate filter and minimize the size of filter needed to process a given batch.

Process Scale-Up and Filter Selection

Candle elements, pressure filters, and filter presses are all used in biopharmaceutical applications. Your choice of filter will depend on a number of factors, including plant design, equipment compatibility with existing systems, clean-in-place and steam-in-place considerations, scale of operation, and overall capital costs. All these factors need to be considered in selecting an appropriate filter. Once the filter design has been selected, the data generated in the methods development stage can be used to predict filter size. A smaller filter can be considered if the cycle time is short, and multiple cycles (including time for cleaning and setup) can be performed, in a typical 8 hour shift, to clarify the batch.

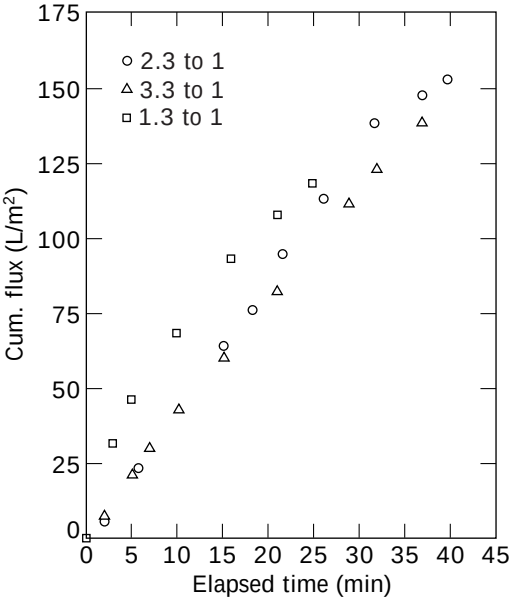


Figure 7. Body feed optimization study for a biological broth with body-feed-to-solids ratios.

SUMMARY

The use of advanced filtration media in dynamic filtration systems is a flexible, efficient, and attractive means for the solid–liquid separation of microparticulates from various process streams. Powdered media provide versatility, high solids-loading capacity, high product recovery, low cost, and ease of scale-up in any filtration process. These systems are cost effective and extremely flexible and efficient at removing suspended solids and clarifying supernatants rich in colloids and small particles. Newer products such as Celpure™ demonstrate improvements in purity and performance and further address the needs of the biopharmaceutical industry.

In addition to the advances that have been made in powdered media, improved systems for media handling and containment are now available from vendors that readily meet the needs of the industry. Finally, a wide variety of economical filter housings of various designs, including candle elements, pressure filters, and filter presses, are manufactured to biopharmaceutical specifications.

BIBLIOGRAPHY

1. R.H. Rees and C.W. Cain, *Chem. Eng.* 72–74 (August 1990).
2. J. Kiefer, *Brauwelt International IV*, 1991, pp. 300–309.
3. J. Bear, *Dynamics of Fluids in Porous Media*, Dover Publications, New York, 1988, pp. 119–195.
4. J. Hermia and S. Brocheton, *Filtration Separation* 721–725 (November 1994).
5. W.P. Olson and R.S. Kent, *Transfusion*, **29**, 86–87 (1989).
6. S.W. King, M.R. Wills, and J. Savory, *Res. Commun. Chem. Pathol. Pharmacol.* **26**, 161–169 (1979).
7. M. Inoue, Y. Gion, H. Itoh, K. Ikariya, K. Takechi, S. Tomioka, K. Furuta, and S. Yabushita, *Vox Sang.* **66**, 249–252 (1994).
8. W.A. Banks, L.M. Maness, M.F. Banks, and A.J. Kastin, *Neurotoxicol. Teratol.* **18**, 671–677 (1996).
9. S.A. Hulea, E. Wasowicz, F.A. Kummerow, *Biochim. Biophys. Acta* **1259**, 29–38 (1995).
10. N. Ahmed, J. Liggins, A.J. Furth, *Biochem Soc. Trans. (England)* **21**, 93S (1993).

See also CELL SEPARATION, SEDIMENTATION; MICROENCAPSULATION.

FILTRATION, AIR

HOLLY HAUGHNEY
Pall Corporation
Long Island, New York

KEY WORDS

Air
Fermentation
Filtration
Membrane
Sterilization
Vent

OUTLINE

INTRODUCTION

In the bioprocessing industry there exist a number of applications for air (or gas) filtration (1–3). The type of filter used for the application will depend on the specific removal requirements. The filtration required can be to the 0.2-μ (sterilizing) level or to a courser filtration level for particulate removal. Some air applications for which sterile filtration can be a requirement are fermenter inlet air, fermenter vent gas, vents on water-for-injection tanks, and vacuum break filters during lyophilization. Courser filtration can 90be used for particulate removal. In many cases, the course filter acts as a prefilter to the sterilizing grade filter. Other applications use coalescers as prefilters for the removal of liquid droplets, such as oil or water.

TYPES OF AIR FILTERS

Packed towers were the first air filters used by the industry for air sterilization. Packed towers consist of beds or pads of a fibrous material, such as paper, cotton wool, glass wool, or mineral slag wool. Membrane filters typically are constructed of a porous, hydrophobic membrane material, such as PVDF (polyvinylidienefluoride) or PTFE (polytetrafluoroethylene).

When filters are used to produce sterile air, the filters can be referred to as sterilizing grade, 0.2-μ filters. Microbial retention tests can be used to verify that the filters produce sterile air. The standard test organism is *Brevundimonas (Pseudomonas) diminuta* (ATCC 19146).

The removal efficiency of filtration media used in packed towers or membrane filters can be dependent on the following mechanisms:

1. Direct interception by the fibers
2. Inertial impaction
3. Brownian motion or diffusional interception
4. Electrostatic attraction between the fibers and particles

Brownian motion (also referred to as diffusional interception) applies to small particles at low face velocities. When air molecules are in a state of random motion, small particles suspended in the particles can be struck by moving air molecules and displaced. The movement of particles resulting from molecular collisions is known as Brownian motion. This phenomenon can increase the probability of capture of the particle by the fibers in a packed tower or by the diffusional interception in a membrane.

The filtration mechanism for *direct interception* is a sieving action that mechanically retains the particles. The filter material acts as a screen that stops particles that are larger than the pores. Direct interception is independent of face velocity and mostly involves particles with large diameters.

Inertial impaction refers to the deviation of a particle from its streamline during flow because of inertia. This deviation can result in the retention of the particle by the fibers in the packed tower or by the membrane in a membrane cartridge filter. As the face velocity increases, the probability of inertial impaction increases.

An *electrostatic attraction* can exist between the fibrous material in the depth filter, or the membrane material, and the particles in the air stream. This attraction can enhance the ability of the filter to remove particles.

A general description of packed towers as well as membrane filters used for air filtration is provided in the next section.

Packed Towers

Description of Packed Towers. Packed towers for air filtration are comprised of beds of fibrous material (1). The diameter of the fibrous material is between 0.5 and 15 μ, and the space between the fibers can be many times this range, depending on how the tower is packed.

Figure 1 illustrates the basic design of a packed tower air filter (5). The filter consists of a steel structure filled with loose fibrous packing. The air inlet is on the bottom and the outlet is on the top of the filter. The packing is supported by a grid or perforated plate. For proper opera-

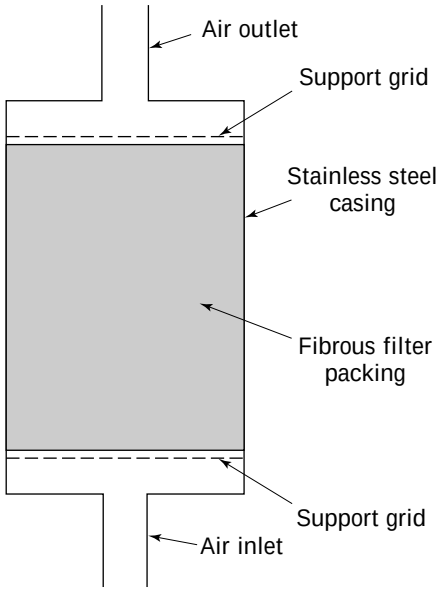


Figure 1. Packed air tower design.

tion, it is necessary to ensure that the appropriate packing density for the application is obtained. If the tower has been packed properly, there will not be movement of the fibers during use. Fiber movement during use can lead to channeling of air, which will lead to an inefficient packed tower, because only a portion of the bed will be acting as a filter. Fiber repositioning can also lead to the dislodging of trapped microorganisms.

Once the filter has been packed with the appropriate fibrous material, a support grid or plate is fitted to the top of the bed to ensure that the bed remains compressed. After the first steam sterilization of a packed tower air filter, the packing will tend to settle further. Additional fibers can be added to the bed after the first steam sterilization to maintain the correct packing density.

Bonded fiber mats have been developed to use in place of loose fibers. Resins are used to bond the fibers, and the mats must be tested to determine if they are compatible with steam sterilization. When mats are used, it is necessary to have a good seal between the mat and the tower wall, so that channeling does not occur. Fiber pads often contain mineral wool with an average fiber diameter of 4 to 5 μ.

Thin sheets of small diameter fibers are another option for packed towers. The sheets are placed on top of each other with a mesh or grid in between each sheet for support. The edges of the sheets are sealed between flanges.

Sterilization of Packed Tower Air Filters. Before using a packed tower for the sterilization of fermenter inlet air, the filter itself must be sterilized. There are two techniques that can be used: steam sterilization and dry heat sterilization (6).

Passing steam at a pressure of 15 psig through a packed tower for 2 h is typically adequate for sterilization. Actual conditions required for a specific installation may vary. The presence of air in the packed tower during the sterilization can prevent complete sterilization of the packing. A drain at the bottom of the packed tower is used to purge the steam and to drain any residual condensate in the filter. It is necessary to remove condensate from the bed because wetted fibers are less efficient for microorganism removal and may also decrease the retention efficiency of the packed tower well below its design value, especially if channeling through the wetted media occurs.

Some fiber material as well as material used to bond fibers can be degraded by steam sterilization. An alternative to steam sterilization is dry heat sterilization, which will avoid the possibility of steam degradation and fiber wetting. This can be accomplished by using a heating device at the inlet of the tower and passing air at a temperature of 160 to 200 °C through the bed for 2 h. During dry heat sterilization, the filter is isolated from the rest of the process. Fiber material can be stable up to 800 °C, thus there is no chance of damage to the bed from this sterilization technique.

Operating Considerations for Air Sterilization by a Packed Tower. The microorganism retention efficiency of packed towers is dependent on the inlet velocity of the air. The relationship between air velocity and filtration efficiency has been determined experimentally in a number of studies (7). In the example shown in Figure 2 (7), the particle retention efficiency may change by a factor of 10 as a result of a relatively small change in the inlet air velocity. The results also show that the most difficult particles to remove are in the size range of small microorganisms. Therefore, it is quite possible to encounter air flow conditions in a packed tower that can reduce the statistical probability for the complete retention of all microorganisms.

The hydrophilic nature of the fibrous material used in packed towers (e.g., glass wool) can contribute to a reduction of the microorganism removal efficiency of the packed tower. Water vapor can enter the system with the air discharge of the compressor. The air and water vapor mixture is initially at an elevated temperature. As the gas stream is cooled, water droplets may condense and wet the hydrophilic glass fibers. The wetted fibers are less efficient for microorganism removal and may decrease the retention

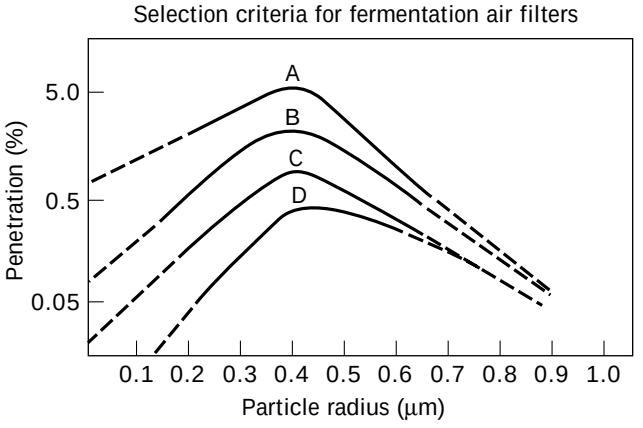


Figure 2. Particle retention efficiency.

efficiency of the packed tower well below its design value, if channeling through the wetted media occurs. Also, organic components present in the compressor exhaust can provide a nutrient source for the retained microorganisms and increase the possibility of bacterial growth and eventual penetration through the depth filter medium. Although a number of approaches have been tried to overcome the wetting problem, such as heat tracing to maintain an elevated temperature, these are expensive to operate and have not consistently resolved the problem. The pressure differential across the tower may also increase significantly with wetting.

Although it is obvious that operating deviations such as these do reduce the reliability of a packed tower for air sterilization, there is no quantitative procedure to determine if the filtration efficiency is adequate to ensure a sterile inlet air condition. The lack of such a quantitative procedure adds an element of uncertainty that affects both the operation and maintenance decisions. Skilled operators are necessary to monitor the packed tower to ensure that it is packed and operating efficiently. The actual source of contamination of a production batch is often difficult to identify, and the air filtration system is always suspect during a contamination outbreak. Without a technique for testing the efficiency of a packed tower, it is difficult to ascertain whether the packed tower is the cause of the contamination outbreak.

Design of Fibrous Air Filters. In order to model a fibrous air filter, several assumptions must be made. It is assumed that (*1*) Once a particle is trapped by a fiber, it then will remain trapped; (*2*) At a particular depth across the filter, the particle concentration does not vary; and (*3*) The removal efficiency at a given depth is equivalent across the filter.

The following equation describes how the concentration of particles varies with the depth position in the filter (6):

$$\frac{dN}{dt} = -Kx \tag{1}$$

where N is the particle concentration, x is the depth, and K is a constant.

Solving the equation between a depth of 0 and x and a particle concentration of N_0 particles entering the filter and particle concentration of N particles leaving the filter yields:

$$\ln\left(\frac{N}{N_0}\right) = -Kx \tag{2}$$

The relationship between depth and the logarithm of the ratio of particles removed to particles incident is known as the log–penetration relationship (6). This relationship has been used in sizing depth filters.

The constant K in equation 2 will vary with the type of packing and is dependent on linear velocity through the packed bed. If the relationship between the constant K and linear velocity through the bed is known, equation 2 can be used to size a packed bed for a given log reduction of particles.

Another consideration for the design of a packed bed is the pressure drop across the bed. Typically, the ΔP is linear with the linear air velocity for a given depth (6). The pressure drop across the packed tower can be dependent on the type of medium, the packing density, the air density, and the linear air velocity through the filter. As an example of an equation for the pressure drop across the packed bed is given in Richards (6):

$$\Delta P = \frac{2\rho v^2 \alpha x C}{\pi D_f} \tag{3}$$

where ΔP is the pressure drop, v is the linear air velocity, α is the ratio of filter density over fiber density, x is the filter bed depth, C is the drag coefficient, D_f is the fiber diameter, and ρ is the air density. The relationship in equation 3 indicates that the pressure drop is proportional to the square of the linear air velocity. At relatively low linear air velocities (2 to 3 ft/s), the relationship is linear.

Membrane Filter Cartridges

Membrane filter cartridges are available as either prefilter (particulate contaminant rating) or sterilizing filter (bacterial contaminant rating) configurations (1–3). Prefilters in air service can be used for particulate removal or aerosol removal. Prefilters are positioned upstream of the final (sterilizing) filters to protect the final filter from premature plugging, thereby prolonging the life of the final filter. The following are brief descriptions of the membrane filters (sterilizing and prefilter types) that can be used for air filtration.

Sterilizing Grade Membrane Filters for Air Service. Membrane filters used for fermenter and bioreactor sterile inlet air and exhaust gas vents, sterile pressure gas, sterile nitrogen blankets, storage tank sterile vents, formulation tank sterile vents, and sterile air for aseptic packaging usually contain a membrane made of hydrophobic materials such as PVDF or PTFE.

Hydrophobic membrane filters are desired in these sterile gas filtration applications because hydrophobic filters do not spontaneously wet with water. When a hydrophilic filter is wetted with water, it will not pass air until the water-wet bubble point of the filter is exceeded, and this water-wet bubble point can be greater than 50 psi. The inherent hydrophobicity of membrane filters used for fermenter air sterilization allows these filters to be able to remove bacteria completely from inlet air, even when exposed to moisture (8). The filter cartridge is expected to remove bacteria and bacteriophage from air streams with 100% efficiency.

The hydrophobic membrane filter material is pleated together with a layer of material (for example, nonwoven polypropylene) on the upstream and downstream side of the membrane. These layers provide mechanical support to the membrane and proper drainage of the fluid. The pleated membrane pack is formed into a cylinder. The longitudinal side seal of the pleated membrane filter pack should be an integral, homogeneous melt seal without any additional or extraneous materials, glues, or resins.

A rigid, perforated inner core is present to provide support against operating pressure. An outer cage placed on the upstream side of the membrane filter pack is provided for additional support and protection during handling. The cage provides retention of structural integrity against accidental reverse pressure (usually up to about 50 psi). Polypropylene and PTFE are examples of materials that can be used for the cage, core, and support material in a membrane filter.

End caps are attached by melt sealing to imbed the medium in the plastic. All hardware should be specified as produced from virgin materials.

Membrane filters are available in a variety of shapes and sizes. The most typical configuration for pharmaceutical air applications are 10-in. elements. These elements can be flame welded together end to end to form multilength configurations (typically up to 40 in.). Multiple elements can be used in a filter housing, as required. A membrane filter is illustrated in Figure 3.

Membrane filter elements may be integrity tested to ensure that the air is sterile. The integrity tests have been correlated with bacteria removal efficiency by direct microbial challenge testing.

Sterilization of Membrane Filters. Hydrophobic membrane process filter cartridges can be subjected to multiple sterilization cycles and must be designed to be repeatedly steam sterilized in either direction of flow or repeatedly autoclaved. The vendor should provide information on filter cartridge sterilization procedures and operating data on process limitations (time and temperature). These filters are typically capable of withstanding multiple in situ steam sterilizations (for example, at least 165 cumulative hours at up to 142 °C). If needed, the membrane filter can be autoclaved and aseptically installed into the process application.

Hydrophobic membrane filters can be in situ steam sterilized in 30 min of exposure time (or longer if the process requires a longer time) at a temperature of 125 °C. Because the filters are hydrophobic, no additional drying time is required. It is necessary for membrane filters as well as for packed towers to drain the entrained condensate on the inlet side of the filter. A membrane filter housing can be isolated under pressure at the end of an aeration period and therefore does not require sterilization for each cycle.

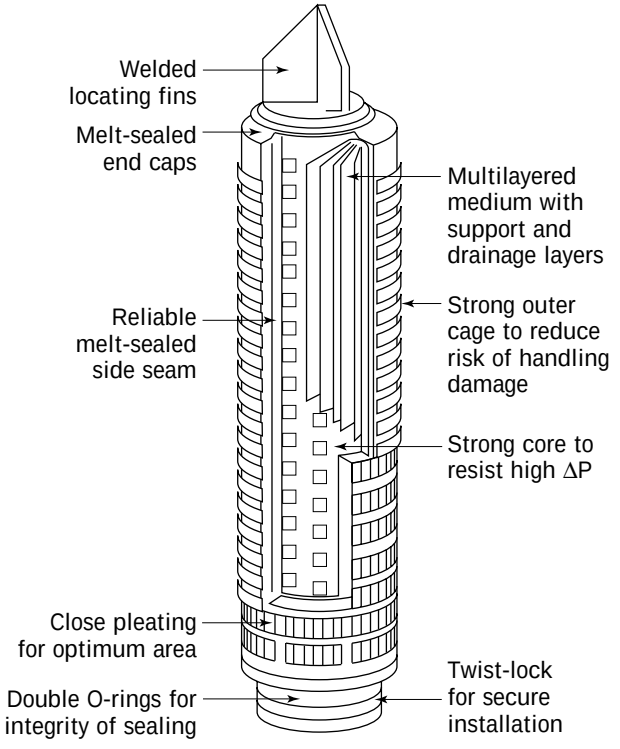


Figure 3. Membrane filter.

Prefilters. For some applications, it is necessary to use a prefilter to remove particulate material or liquid aerosols. It is often desirable to use a prefilter to protect the sterilizing grade filter. Typically, the use of a prefilter will reduce the overall filtration economics.

Particulate Removal. Membrane filters composed of materials such as polypropylene or cellulose can effectively be used to remove particulate material from an air stream. The filter micron ratings range from the order of 1 to the order of 100 μ. The appropriate filter can be selected for the particulate contaminant removal. The following are descriptions of examples of membrane filters that can be used as prefilters in pharmaceutical air filtration, including porous stainless steel filters, cellulose pleated filters, and polypropylene pleated filters.

1. *Porous stainless steel filters cartridges*. Porous stainless steel medium is made by sintering small particles of stainless steel or other high alloy powder together to form a porous metal medium. Porous stainless steel can be formed as a flat sheet or, when used in filter elements, as a seamless cylinder. This special manufacturing process produces a high dirt-capacity medium that is temperature and corrosion resistant. The recommended alloy is type 316LB, which has a higher silicon content than type 316L and provides a stronger, more ductile product with better flow properties. Elements have absolute ratings of 0.4 to 11 μ in gas service applications. Porous stainless steel filter cartridges are chemically or mechanically cleanable, offering economy of reuse. Porous stainless steel filters are used for steam filtration and for air sparging.
2. *Cellulose pleated filter cartridges*. Pleated cellulose filter cartridges (8 μ rating) are applicable as prefilters for inlet air for fermentors and bioreactors. These filter cartridges can be constructed of pure cellulose medium, without resin binders. The cellulose membrane can be pleated into a high area cylinder and has a longitudinal side seal with an appropriate polypropylene (approved for food contact usage by the FDA). Cellulose media cartridges are assembled with hardware components consisting of a perforated inner support core, an outer support cage, and end caps melt sealed to imbed the medium in the plastic. All hardware components should be of pure

polypropylene, without filler or reinforcement, to ensure a minimum of soluble extractables.

3. *Polypropylene pleated filter cartridges*. Polypropylene pleated depth filters (8-μ rating) are applicable as prefilters for prefiltration of exhaust gases. These process filter cartridges are constructed using nonmigrating continuous strands of nonwoven polypropylene filaments. The medium should have a constant pore size section (downstream) for absolute rated filtration and a continuously graded upstream section for effective prefiltration. The thin sheet of polypropylene media is pleated and formed into a cylinder with a longitudinal side seal of melt seal polypropylene. The cylinder is then melt sealed to injection molded polypropylene end caps to ensure no fluid bypass. Polypropylene hardware components consisting of an inner support core and an external protective outer cage are incorporated.

Liquid Aerosol Removal. A coalescer can be used for the removal of liquid aerosols containing water or oil droplets. This is desirable as a prefiltration for a sterilizing air filter, because the liquid aerosol could prevent the flow of air through the sterilizing grade filter.

Coalesers operate efficiently if they are able to separate the liquid and the gas in the liquid aerosol. The three basic steps that are required are (*1*) aerosol capture (*2*) unloading or draining of the liquid, and (*3*) separation of the liquid and gas.

Figure 4 is an illustration of a liquid–gas coalescer. The coalescer has a gravity separator, which allows for the removal of large liquid aerosols (typically >300 μm). The coalescer flow direction is in to out to prevent reentrainment. The liquid is captured through the coalescence of fine aerosols (0.1 to 300 μm) to large droplets (1 to 2 mm). The coalescence of the droplets is illustrated in Figure 5. The large droplets flow downward from a drainage layer. The separation is completed when the liquid is drained from the system. The air leaves the system from the top of the assembly.

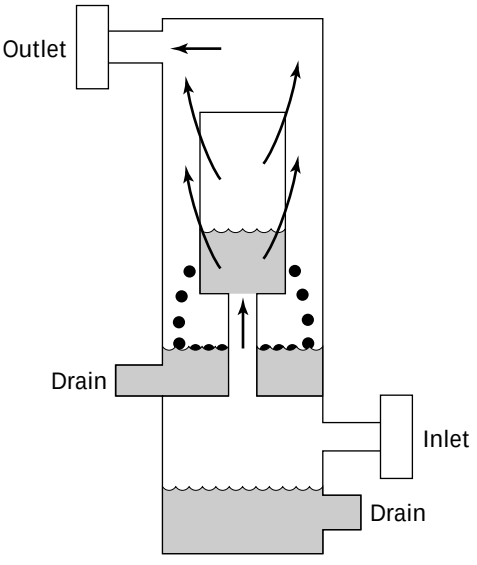


Figure 4. Liquid–gas coalescer.

Filter Housings. Cartridge filters used for air filtration are placed in sanitary, air or gas service filter housings. Cartridge filter housings usually consists of a head and a bowl. The filter is attached to the head, and the bowl is clamped to the head to provide a complete enclosure of the filter.

The size of the housing should be adequate for the flow and differential pressure requirements. Filter housings for bioprocessing applications are typically constructed of stainless steel (e.g., 304, 316, 316L, etc.) or carbon steel, with 316-series stainless steel internal hardware and cartridge seating surfaces. Housings typically have quick-release mechanisms such as V-band clamps or fast-action swing bolts to facilitate filter change-outs.

Design operating pressure of all filter housings should be specified as minimum psig and rated for full vacuum service. Design maximum operating temperature of the housing should also be specified. The housings or pressure vessels that are within the scope of the ASME Boiler and Pressure Vessel Code, Section VIII, Division 1, should be designed and U stamped per the code. TIG weld construction should be used in sanitary-style housings to minimize weld porosity and ensure high-quality, clean joints, with all internal welds ground smooth and flush. All weld procedures and welders should be qualified to ASME/BPVC Section IX.

Housings should be capable of in situ steam sterilization in accordance with the manufacturer's recommended procedures, and housing or system design should provide for condensate drainage. Gasket material and O-ring elastomers must also be capable of withstanding repeated steam sterilization cycles, along with being compatible with process fluids.

Industrial style housings (used for prefilters) provide cartridge mounting on a tie rod and sealing to the tie rod assembly by use of a seal nut at the top of the assembly. Tube sheet adapters should be seal welded to the tube sheet to prevent fluid bypass. Filter cartridges are thereby sealed in the housing independent of any cover assembly, ensuring positive sealing and no fluid bypass. Filter cartridges should be seated on the tube sheet adaptor assemblies above the tube sheet to ensure complete drainage of nonfiltered fluid before cartridge replacement. This prevents potential contamination of downstream surfaces during change-out of filter elements.

QUALIFICATION TESTS FOR MEMBRANE FILTERS USED FOR AIR FILTRATION

Sterilizing Grade Membrane Filters

Organism Retention Tests. Microorganism retention tests can be conducted to verify that membrane filters produce sterile air (1,2). Liquid challenge tests with *Brevundimonas (Pseudomonas) diminuta* (ATCC 19146), measuring 0.3×0.6 to 0.8 μm, is a standard challenge test for the validation of sterilizing grade filters (0.2 μm) in the pharmaceutical industry. Aerosol challenge tests with *P.*

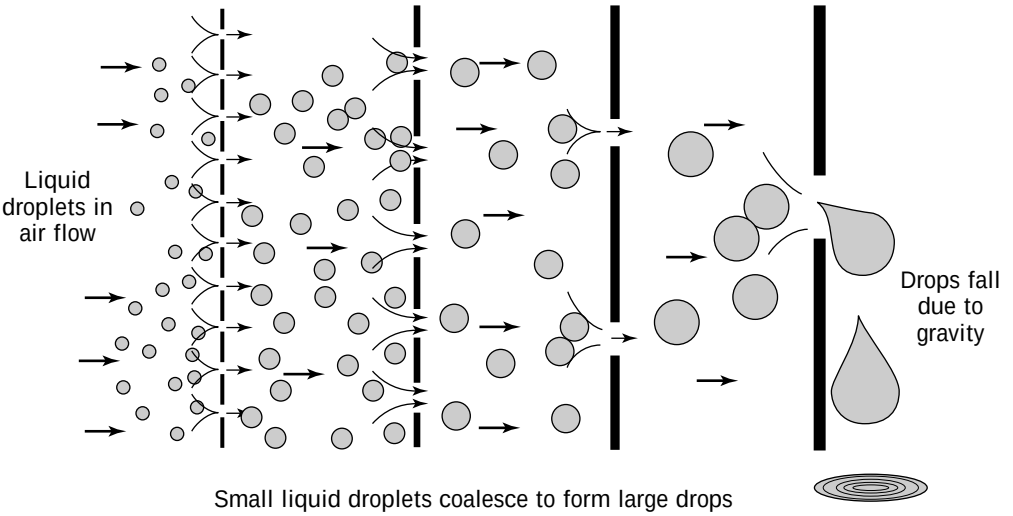


Figure 5. Coalescence of droplets.

diminuta should approximate extreme air flow conditions. Aerosol challenges with T_1 bacteriophage ($0.05 \times 0.1\ \mu$m), for example, can provide a test of a filter's retention efficiency of extremely small organisms, so small that the bacteriophage in a liquid suspension will penetrate a 0.2-μm sterilizing grade filter.

The retention efficiency of a given filter is less when a liquid challenge is used instead of an aerosol challenge. Thus, a liquid challenge test is a more stringent test of a filter's retention capability. A liquid challenge test can also provide retention information for process conditions such as extreme moisture after sterilization or air entrained with water drops. An example of the technique used to perform a liquid bacterial challenge on sterilizing grade membrane cartridges was published by Pall Corporation (9).

The liquid test involves challenging a test filter with a known quantity of *P. diminuta*, no less than 1×10^7 organisms per 2 square centimeter of filter area. The challenge sample is passed through the filter suspended in sterile water at a defined flow rate and time. All of the effluent from the test filter is passed through an analysis membrane. After the challenge is completed, the system is flushed for 5 to 10 minutes with sterile water. The analysis membranes are removed and placed on Mueller Hinton Agar at 32 °C for 48 h. After incubation, the plate is examined for the presence or absence of microbial colonies. Figure 6 illustrates a liquid bacterial challenge test stand.

The aerosol challenge test system can consist of a nebulizer loaded with the challenge microorganism suspension, a separate line for dry air makeup, and split-stream impingers to sample the aerosol challenge with and without the test filter. A schematic of a test stand that can be used for the aerosol challenge procedure described is given in Figure 7. Detailed descriptions of the aerosol test protocols for *P. diminuta* and T_1 bacteriophage Ref. 9.

During the aerosol challenge, an aerosol is generated with a nebulizer. The aerosol is introduced into the test filter at a given flow rate. The filter effluent is collected in dual liquid impingers. Controls are performed simultaneously via a split stream by using a two-channel timer to direct air flow, on an alternating basis, from the test side filter impingers to the unfiltered control side impingers for recovery.

The impingers contain sterile buffer and after the challenge is completed, the buffer can be analyzed for the test organism. If *P. diminuta* is the test organism, then the buffer is analyzed by putting the buffer solutions through an analysis membrane and placing the membrane on Mueller Hinton Agar for 48 h before titering. If T_1 bacteriophage is the test organism, then samples of the buffer are diluted with nutrient broth and mixed with liquid nutrient agar (1.5% agar concentration; 48 °C) and *Escherichia coli* in the log phase of growth. After vortex mixing all three components, the mix is poured over nutrient agar plates and incubated for 3 to 5 h at 37 °C, so that the plaques can be counted.

Integrity Tests for Membrane Filter Cartridges. During the sterilization of fermenter air, it is necessary to achieve the highest possible assurance of filter integrity and removal efficiency. The installation of integrity-testable filters and the performance of routine integrity testing by the user are essential to demonstrate that the system is performing to specification. Tests that qualify the retention characteristics of a membrane filter can be defined as destructive or nondestructive tests.

Destructive tests are performed using an appropriate contaminant to meet a specific claim for retention of the contaminant. The test procedure must be sensitive enough to detect the passage of contaminants of interest. For sterilizing grade 0.2-μm membrane filters, the industry standard test organism (i.e., contaminant) is *P. diminuta* (ATCC 19146). The organism and minimum challenge level (10^7 CFU/cm^2 filter area) are specified in the ASTM standard F383,-83 (10) and referenced in the FDA guideline on sterile drug process produced by aseptic processing (11).

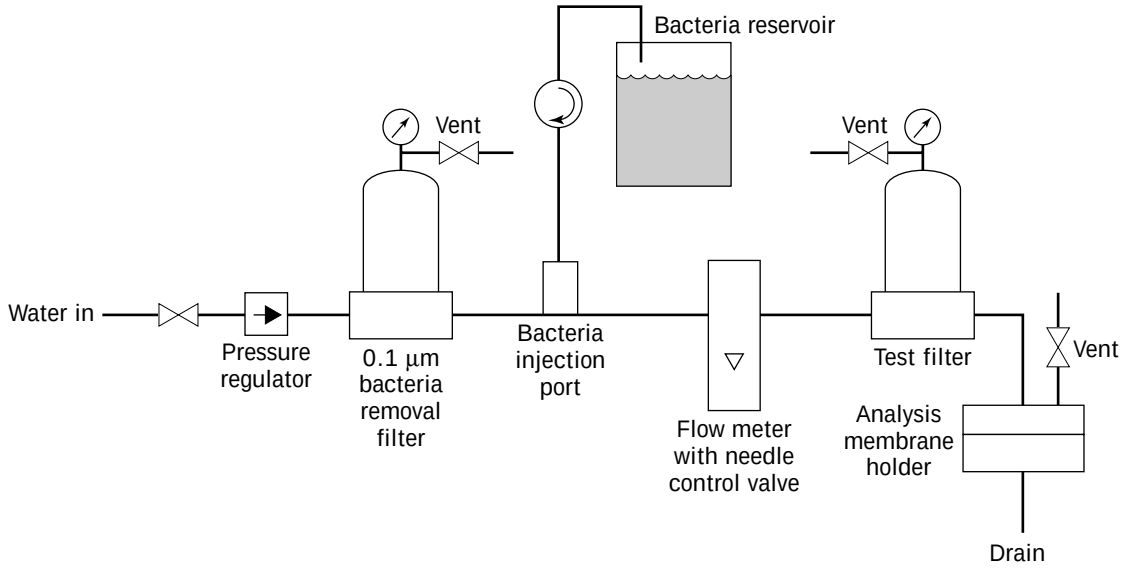


Figure 6. Liquid bacterial challenge test stand.

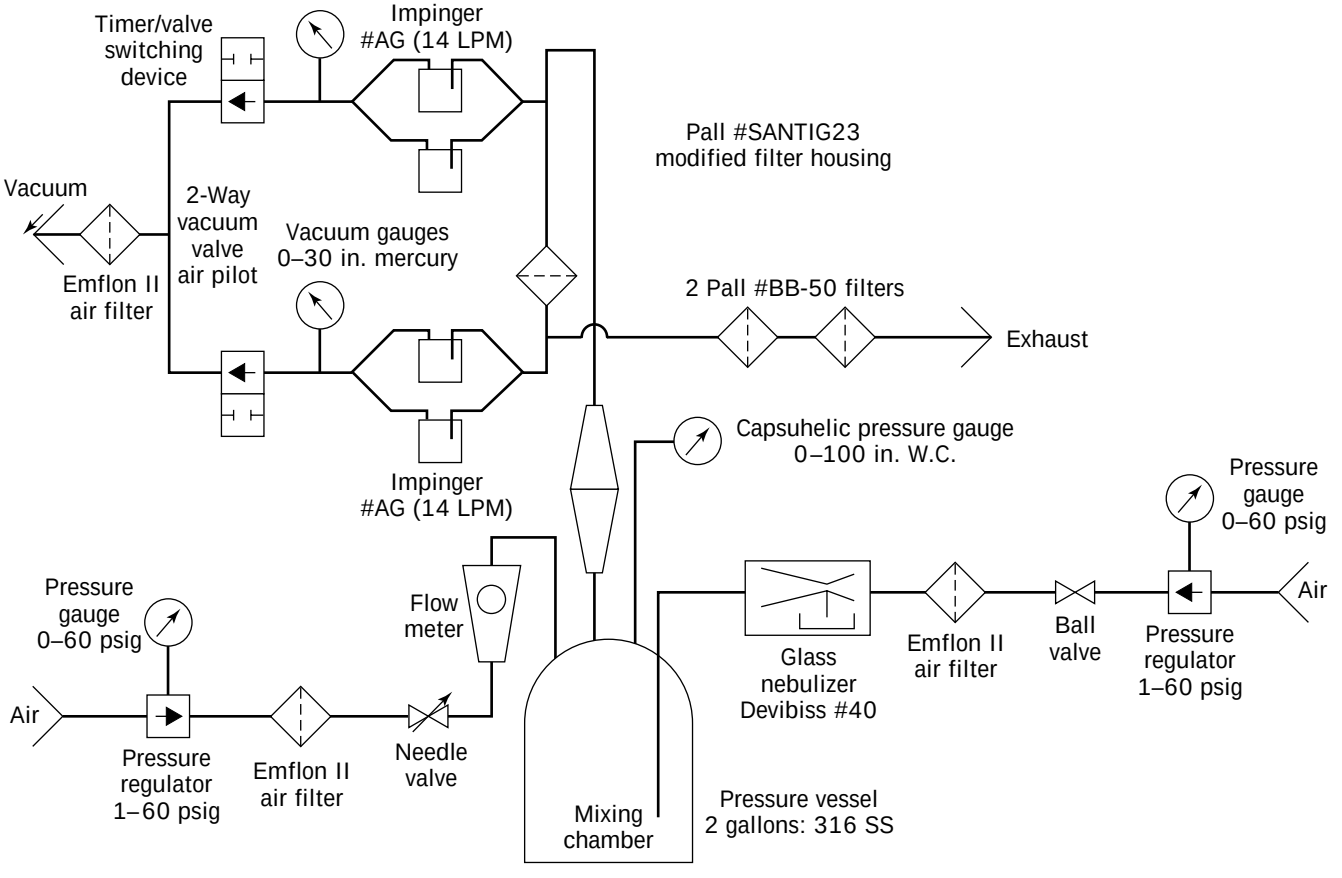


Figure 7. Aerosol challenge test stand.

Because most filter users would not want to perform a destructive test in a process environment and because the filter is not useable after the destructive test, nondestructive tests related to the retention results of the destructive test are used instead. From the relationship developed between a nondestructive and a destructive test, membrane filter performance can be safely and conveniently verified in the production environment. The relationship between a nondestructive integrity test and the assurance of bacterial retention constitutes a filter validation study and is extremely important for microbial retentive filters used in critical fluid processes.

There are four nondestructive integrity tests that can be used for sterilizing grade hydrophobic membrane filters used for air filtration: (*1*) bubble point test, (*2*) forward flow test, (*3*) pressure hold test, and (*4*) water intrusion test. These tests are described in FILTRATION, CARTRIDGE.

For a nondestructive test to be useful, the results of the test must be able to predict the ability of the filter to remove bacterial. The manufacturing quality control of the membrane filters tested to obtain this prediction of bacterial removal for the filters must be used to establish quality control procedures that are consistently maintained in the manufacturing of the filter. The combination of the nondestructive integrity test, the correlated bacterial challenge, and manufacturing quality control allows membrane filters to be used reliably for sterile filtration.

Tests for Prefilters

Many filter manufacturers use a nominal micron rating for removal efficiency. A nominal rating is an arbitrary micron rating assigned by the filter manufacturer. Such ratings are subject to a lack of reproducibility.

An alternative method for rating filters is the Oklahoma State University (OSU) F-2 Test. This rating method (ISO 4572, ANSI B93.31) has received wide acceptance for use on lubricating and hydraulic fluids. Pall Corporation, for example, uses this method for oils extensively and has adapted it for use in water with contaminants ranging from 0.5 to 25 μm.

The test is based on continuous on-line particle counts of different particle sizes, both in the influent and the effluent. The β ratio at a specific particle size is defined as β_x: the number of particles of a given size (X) and larger in the influent, divided by the number of particles of the same size (X) and larger in the effluent, where X is the particle size in microns. The percent removal efficiency can be calculated from the β-value. The percent removal efficiency is $[(\beta_x - 1)/\beta_x)]$ 100 (12).

As an example, Pall prefilters are given a micron rating that corresponds to a 100% removal efficiency or the value in microns at which the OSU F-2 Test gives a β-value of more than 5000.

AIR FILTRATION APPLICATIONS

Fermentation and downstream processing are two major operation categories in bioprocessing. During both fermentation and downstream processing, sterilizing-grade 0.2-μm hydrophobic membrane filters can be required for processing air (or gas) streams.

Sterilizing-grade 0.2-μm hydrophobic membrane filters are used during fermentation for the sterilization of fermentation inlet air and for the filtration of fermentation exhaust gas. During downstream processing, sterilizing grade filters can be found in use as sterile tank vents. In a final purification, membrane filters can be used for vacuum breaking in processes such as lyophilization.

As discussed above, hydrophobic membrane filters are desired in sterile gas filtration applications because hydrophobic filters do not spontaneously wet with water.

Sterile air filtration applications discussed in this section include fermenter inlet air, exhaust gas (vent), high temperature air filtration, and lyophilizer vacuum break filtration. Special attention will be given to the specific requirements for a sterilizing filter when used for these applications. After the discussion of the applications is a general discussion of recommendations for sterilizing-grade filter usage.

Fermentation

During a fermentation process a specific cell (yeast, bacteria, or mammalian) is grown to provide a desired product (1–3). Products can include cells, antibiotics, amino acids, or recombinant proteins. There can be a variety of sizes for the fermenter, ranging from very small (100 L or less) cell culture reactor to very large scale antibiotic production (100,000 L). In these applications, there is often the need to maintain sterility in both liquid and gas (air or nitrogen) feeds to support growth of the desired cells.

Air filtration applications for fermentation are illustrated in Figure 8. Those applications specific to fermentation are described here, and filtration of utilities used in fermentation such as steam, air, and water are discussed in the "Utilities" section.

Prefiltration of Fermentation Air. Compressors are often used to generate air flow for the manufacturing facility. There are two types of compressors, oil free and oil lubricated. In older facilities where oil-lubricated air compressors are commonly used, prefiltration of inlet air is necessary for removal of oil droplets. A coalescing filter can provide greater than 99.9% removal of oil and water droplets in the 0.01- to 0.5-μm range and larger. This also acts as an excellent prefilter for the hydrophobic membrane pleated filters that are commonly used for sterilizing the inlet air to the fermenter. The typical gas flow rate per 10-in. filter module is 200 to 400 standard cubic feet per minute (SCFM).

For oil-free compressors, a prefilter acts to remove dirt in the air system, extending the service life of the final filter. For use with fermentation air, a cellulose pleated filter with an absolute rating of 8.0 μm is normally the filter of choice. Alternately, polypropylene (2.5-μm rated) pleated filters also serve as excellent prefilters for this application. The typical gas flow rate per 10″ filter module is 75 SCFM.

Sterile Air Filtration for Pilot and Production Fermenters and Feed Tanks. One of the largest applications for sterile

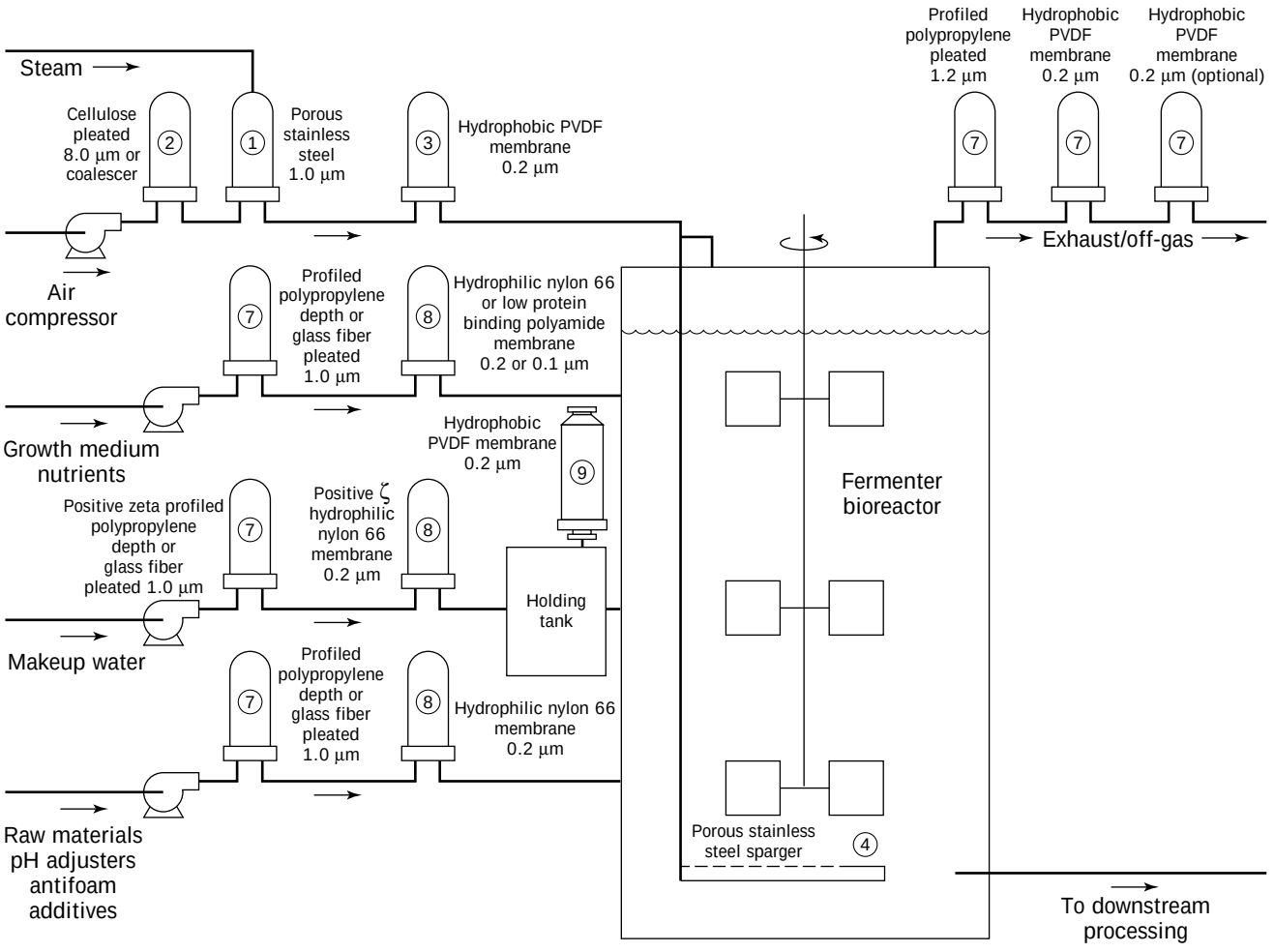


Figure 8. Fermentation air applications.

air filtration is the sterile air used for an aerobic fermenter during a typical production cycle. Typically, 1 vol of air per volume of broth per minute is used. Thus, for a 100,000 L fermenter on line for 48 h a total of 1.01×10^9 cubic feet of air requires sterilization.

The contaminants present in compressed air can include dust, lubricating oil, hydrocarbons, water, rust, and microorganisms including molds, bacteria, and viruses. Microorganisms in air are often associated with carrier particles, such as dust. Water and oil can be present as bulk liquid, vapor, or an aerosol. The air distribution system can give rise to contaminants such as rust and water. The concentration and size distribution of particles in compressed air are variable. The size range is generally between 0.001 and 30 μ, with a concentration between 10^{-2} to 10^{-4} g/m^3 (4,6).

Bacteria and bacteriophage, when present in air feeds, can enter fermentation tanks or bioreactors and contaminate the product. Bacteriophage or other viruses can destroy the producing cells and reduce yields.

The process requirements to supply this sterile air can be quite restrictive. The air sterilization process must (*1*) process a large volume of compressed air, (*2*) provide a high degree of reliability, and (*3*) operate economically. Several methods have been considered for the sterilization of fermenter inlet air. These include filtration, heat, irradiation, washing with sterilizing chemicals, and electrostatic precipitation. Washing and electrostatic precipitation are not effective for the removal of microorganisms. Heat and irradiation are not economical. Filtration is the only technique that meets all the requirements for sterilizing fermenter inlet air.

Packed towers, an early filtration approach, were used widely in the industry. Since the early 1980s, filtration technology has advanced (13), and there has been an ongoing trend since the early 1980s to replace depth filters with hydrophobic membrane cartridge filters (e.g., Hoffman-La Roche [14] and E. R. Squibb and Sons [8]).

The recommended filters for sterilization of air feeds to fermenters and bioreactors are the hydrophobic membrane pleated filters. The hydrophobic (water repelling) nature of these membranes can provide for bacteria and bacteriophage removal with 100% efficiency under moist or dry operating conditions. This is an important benefit over fiberglass towers and cartridges. Filters for sterile air feeds should have a 0.2-μm absolute bacterial rating in liquids

and a 0.01-μm particulate rating in air service. The typical gas flow rate per 10-in. filter module is 75 to 100 SCFM.

For some fermentations, the requirement may be for the filtration of fermenter air at an elevated temperature. If an application involves hot air and a longer service life is desired, then a filter that can withstand the elevated temperature is required. High Temperature Emflon® filters manufactured by Pall Corporation can be used in continuous service at a temperature up to 120 °C. These filters have a 0.2-μm microbial rating in liquid service and a particulate removal rating of 0.01-μm in gas service. The filter membrane is made of inherently hydrophobic PTFE, and the cage, core, and end caps are specifically designed for high temperature applications.

Sparging. Sparging acts to disperse air evenly in the fermenter or bioreactor containing the growth media and product. The product of choice for this application is a porous stainless steel sparging element, which can provide an exceptionally uniform and fine aeration gas dispersion. These elements are fabricated with one face of porous metal and one face of solid metal. If both surfaces of the sparging elements were porous, bubbles from the under surface may coalesce with bubbles from the top surface.

Porous stainless steel sparging elements should be positioned horizontally in the fermentation tank, with the porous stainless steel facing upward. Fine grades of porous stainless steel (e.g., 3.0 μm absolute liquid rated) are ideal even for sheer-sensitive mammalian cell cultures because of their high gas transfer and low shear aeration capability. Elements are typically available in standard and custom designs.

Exhaust Air–Off Gas Filtration. The purpose of a vent filter on a sterile fermentation tank is twofold: to prevent contamination of the tank and to provide containment of the material inside the tank. Prevention of contamination in the tank is desirable for processes that involve long fermentation cycles or require a sensitive fermentation medium (e.g., tissue culture medium). Genetic engineering techniques as well as fermentation of pathogenic organisms (such as organisms used for the manufacture of vaccines) have made it necessary to protect the environment and prevent the escape of microorganisms from the fermentation tank. The exhaust filtration system for a recombinant or mammalian cell fermenter or bioreactor must yield sterile air to the environment and provide a sterile barrier to prevent ingression of contaminants. Additionally, it must be in situ steam sterilizable and typically has a clean differential pressure less than 1 psid.

The removal efficiencies for simple depth filters (as described earlier) are typically poor under wet or variable flow conditions. Therefore, membrane filters are recommended for vent filtration applications.

The fermenter or bioreactor exhaust gas line can be contaminated with microorganisms or cells, growth medium components expelled from the fermenter or bioreactor as droplets or as solid particles, and aerosol condensate droplets formed during cooling of the gas in the exhaust system. These aerosol droplets, when present, can potentially block the final filter and must be removed before reaching the final filter. Mechanical separation devices, such as cyclones, condensers, and demisters, may not achieve effective aerosol removal below 5 μm (15). Removal efficiencies and pressure drops also vary significantly with flow rate in such equipment. Recent studies have shown that aerosols in exhaust lines are predominantly in the very fine 1- to 5-μm range (16).

The contaminants present will depend on the fermentation conditions, the growth medium, and the design of the exhaust gas system. The basic requirements for a vent filter are ability to provide sterility, a low pressure drop, and in situ steam sterilizable.

The recommended exhaust filtration system design entails two stages using a polypropylene pleated depth-filter cartridge as a prefilter to a 0.2-μm absolute rated hydrophobic membrane pleated filter cartridge. The purpose of the prefilter (typically 1.2-μm absolute rated) is to remove aerosolized particles and liquid droplets containing cells or growth media from the fermentation off-gas or exhaust air. This serves to extend the service life of the final sterilizing filter. If the medium contains only fully dissolved components, such as with a sterile filtered cell culture medium, and if the fermentation is run at low temperatures ($<$30 °C) and low aeration rate (1 to 1.5 volume of air per volume of media per minute [VVM]), the prefilter may be optional. The typical gas flow rate per 10-in. filter module is 40 SCFM.

Like the final sterilizing filter, the pleated polypropylene prefilter should be multiple steam sterilizable. As additional benefit of the prefilter is to retard foam-outs from reaching the final sterilizing filter.

Sterilizing-grade 0.2-μm absolute rated hydrophobic membrane pleated final filters, with PVDF or PTFE membranes, can prevent organisms from entering or leaving the controlled reaction zone, even in the presence of water droplets and saturated gas (8,17). Steam sterilizability and integrity test values correlated to microbial retention studies under worst-case liquid challenge conditions provide the highest degree of assurance performance. Redundant systems using a second 0.2-μm rated sterilizing filter in series are recommended for high-risk recombinant organisms.

Condensate control is usually the most critical consideration for this application. In cases in which there is condensate accumulation and if the fermenter is operated with overpressure in the fermenter head, the amount of condensate accumulation can be reduced if a pressure control valve is placed at the fermenter exit, upstream of the exhaust gas filter. An alternative technique for the prevention of condensate accumulation is to use a heating section in the exhaust gas pipe upstream of the filter installation. This can be also be accomplished by specifying steam jacketing on exhaust filter housings. In this case, the exhaust gas temperature at the terminal filter must lie above the temperature of the exhaust gas at the fermenter exit. The heater must be properly sized based on the process parameters.

Downstream Processing

Starting with the cells and conditioned broth medium from the fermenter or bioreactor, the objective of downstream

processing can be to produce a highly purified, biologically active protein product, free of contaminants such as endotoxins, bacteria, particles or other biologically active molecules. This phase of bioprocessing typically comprises a series of unit operations including cell and cell debris separation, fluid clarification and polishing, concentration and purification, and membrane filtration sterilization of the purified product.

Cartridge filters are used in many stages of downstream processing, which involves filtration of both the harvest fluid and product intermediates as well as filtration of air and gases required throughout the process. Air filtration applications include vacuum break filters for lyophilizers, sterile nitrogen blankets, tank vents, and sterile air for container cleaning.

Absolute rated cartridge filters eliminate contaminants and impurities from air, nitrogen, and other gases used in downstream processing to prevent contamination of product and further protect concentration and purification equipment. Vent filtration ensures containment and freedom from product contamination during fluid transfer operations and protects processing equipment during sterilization cycles.

In fermentation, cartridge filters are used to maintain the sterility of the makeup water, feeds, additives, media in holding tanks, and in fermenter or bioreactor exhaust. Cartridge filters are typically used in downstream processing for the filtration of air, gases, and venting applications when it is necessary to vent tanks during fluid transfers; pressurize tanks using inert gases such as nitrogen and argon; protect vacuum lines, sterile vent holding tanks, and lyophilizers; and for gas purging, blanketing, drying, and when sterilizing equipment by in situ steaming or autoclaving.

The recommended filters for nonsterile particulate removal applications are polypropylene pleated filters. Hydrophobic membrane pleated filters such as PVDF or PTFE are recommended for aseptic processing. The absolute removal rating for the latter filters should be 0.2 μm determined under liquid flow conditions. The typical gas flow rate per 10-in. filter module is 75 to 100 SCFM.

In downstream processing applications, Cell and cell debris separation and clarification processes can be broken into a primary separation, secondary separation, and a cell concentrate section. During some primary separations, a cyclone can be used for particulate removal; a sterilizing air filter can be used as a vent on the cyclone. A variety of holding and receiving tanks can be used during the separation and clarification process; these tanks can be fitted with sterile vent filters.

During secondary separation, a nitrogen blanket may be needed; the nitrogen gas can be sterile filtered with a hydrophobic filter. Downstream processing can also involve the concentration and purification of clarified harvest fluid. Applications for air filtration include sterilizing-grade vent filters for solvent or buffer tanks and for holding or buffer tanks needed for ultrafiltration and chromatography.

The final pharmaceutical product will often need to be packaged. During filling processes, a sterile nitrogen blanket and thus a sterilizing grade hydrophobic filter may be needed. The final product is can be placed into a container; sterile air or nitrogen may be needed for container cleaning. A vent filter can be required on holding tanks. The typical gas flow rate per 10-in. filter module for tank vent applications is 75 to 100 SCFM.

Several specific applications for final processing are addressed in the next section.

Vacuum Break Filters. Sterilizing grade filters are used in freeze dryer installations to filter the gases used to maintain the chamber pressure and to break vacuum during operation and in sterilizers for vacuum break purposes.

Blow-Fill-Seal Equipment. Blow-fill-seal equipment can be used for the aseptic filling of pharmaceutical products. The container is formed and sealed aseptically. Air filtration is required to ensure sterility in this unit operation. A typical arrangement is illustrated in Figure 9.

For buffer-tank air hydrophobic membrane filters are used to supply sterile air to a buffer tank on the blow-fill-seal machine. This blanket air is used to drive the sterile solution through a pneumatically controlled dosing system. The air used in the buffer tank is referred to as the gas cushion or buffer tank air.

Hydrophobic membrane filters provide sterile air used to form the hot moldable plastic tube (parison). The air used to form the parison is known as the parison support air. The parison is subsequently blow molded into the shape of a ampule strip or a bottle. Typical requirements for the filters used include that the filters be steam sterilizable, integrity testable, and the proper size to prevent restriction of gas flow. Hydrophobic membranes are used to prevent wetting out and to maintain high flow rates even in moist conditions.

Utilities

There are a number of peripheral unit operations required during a sterile process. The air filtration applications for these applications are described in this section.

Water-for-Injection Tank Vent. Sterilizing grade membrane filters can be used in vent applications in which the fluid in the tank is at an elevated temperature. One such application is the vent used to prevent contamination in a water-for-injection tank. The water is at 80 °C or higher. When a sterilizing-grade air filter is used for this type of vent service, a steam-jacketed housing is typically used. It is only necessary to maintain the temperature of the filter cartridge at a temperature *slightly above* the dew point of the vapor. The steam introduced into the jacket should be at ambient pressure. Continuous operation of the jacket at a significantly higher steam pressure and temperature can reduce the service life of the filter caused by accelerated aging of the hardware by oxidation.

Steam Filtration. Process equipment and final filters are frequently sterilized by direct steam flow in situ, during the normal line sterilizing cycle. This eliminates the need for making aseptic connections and risking recontamination. Filtered steam is required for this sterilizing-in-place

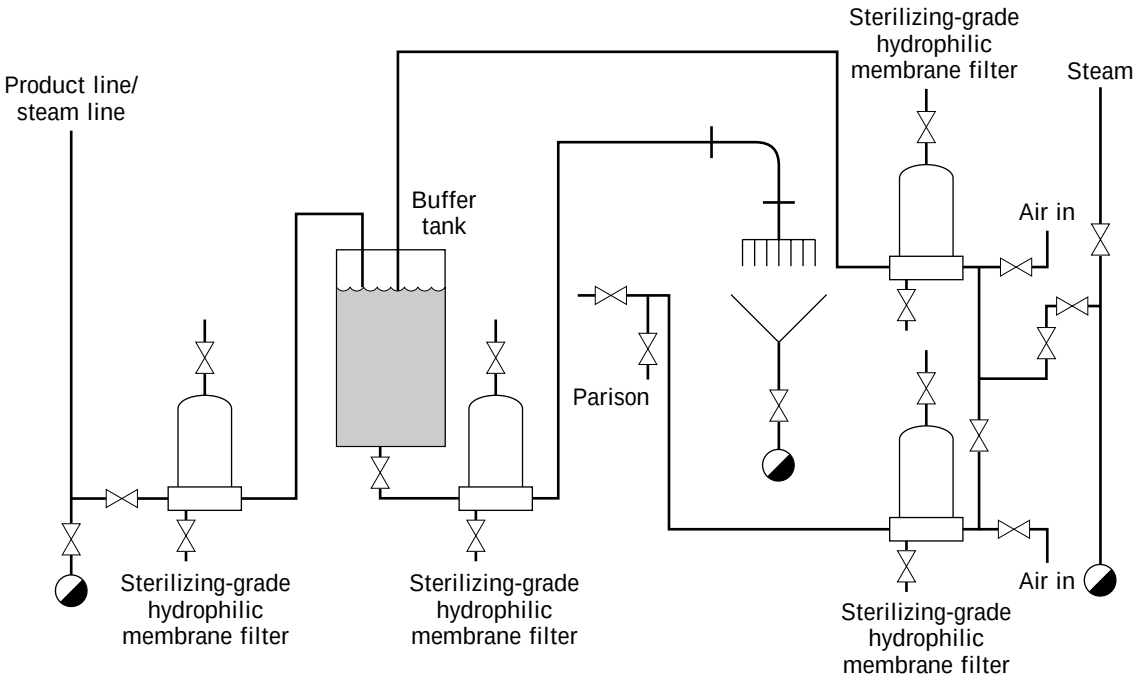


Figure 9. Air filtration applications in a Blow-Fill-Seal operation.

of filters, piping, vessels, and filling equipment. Steam is also required for general equipment cleaning and sterilizing. The steam often contains significant amounts of pipe scale and other corrosion products. This particulate material should be removed in the interest of overall cleanliness and to avoid burdening the prefilters and final filters.

Particulate contamination in process steam is efficiently retained by porous stainless-steel filters with an absolute gas rating of 1.2 μm. Porous stainless-steel filter assemblies are typically sized at steam flow rates of 30 to 40 ACFM per square foot of filter medium.

GENERAL CONSIDERATIONS FOR OPERATION

Integrity Test Considerations and Guidelines

Field experience with the integrity testing of sterilizing grade filters shows that various combinations of integrity test procedures are in use and that different testing schedules, both pre- and post-use, are followed. Perhaps the best approach is to perform the test or combination of tests that provides the highest degree of accuracy commensurate with the economics and the practicalities of the process.

Since the requirements for the use and testing of sterile gas filters can depend on location and application, several different types of filter tests can be used. Filter users typically use the forward flow or pressure decay tests, which require the use of a low surface-tension solvent, such as 60:40 isopropyl alcohol/water or 25:75 *t*-butyl alcohol/water. This is a well-proved, widely accepted approach directly correlatable to bacterial challenge (9).

Air filters are typically integrity tested when they are installed and retested periodically based on service conditions and operating requirements (e.g., once a month). Change-out schedules based on filter life studies have also been used in conjunction with integrity testing. Because filters in these applications are often sterilized in situ and can be damaged during in situ steam sterilization by reverse pressurization if the sterilization procedure is not properly controlled, filter life study data or a periodic integrity test regimen can be used only if the filter sterilization procedure is validated and in control.

In certain cases (such as in facilities in which the use of a nonflammable fluid is required, or when disposal of organic solvents is a concern), some filter users in aseptic processes have recently considered the use of water-based tests for hydrophobic filters.

There are a number of integrity tests possible for a hydrophobic air filter. The selection of the appropriate test and the appropriate test schedule depends on the specific application.

After an integrity test has been completed, it is typically desirable to remove the wetting fluid from the filter. This can be accomplished by blowing clean, dry (-40 °C dew point) air or nitrogen through the filter. It is necessary to qualify this procedure, because every system is different.

Filter Service Life

Filter cartridge change-out is usually based on actual experience, with a safety factor. Filters should be inspected on a monthly basis for oxidation. This should be supplemented by monitoring the pressure drop across the filters during operation to determine if the filters are plugging and routine integrity testing to confirm filter integrity during the service life of the filter. Alternatively, filter life studies, with an appropriate safety factor, could also be used to

set a change-out schedule. Actual conditions for each application should be used during filter life studies.

For elevated temperature applications, conditions leading to oxidation need to be considered. Aging of the membrane filters by oxidation depends on the status of the system. Oxidation does not occur when the cartridge is being steamed, because there should be no air present in a properly operating steam-in-place system. If a cartridge is exposed to air at an elevated temperature, oxidation of the material in the filter, such as polypropylene hardware, will be accelerated. Oxidation will also occur when the filter is in a stagnant situation, that is, it has no air flow going through it. The flow of air through a filter can moderate the temperature environment, whereas under stagnant conditions the temperature of the filter will rise to the temperature of the housing. Stagnant conditions can exist when the tank is not being used or when the tank is empty. To prolong service life, the steam jacket should be turned off when there is no air flow through the filter for extended periods of time, when operating conditions permit.

Steam Sterilization Guidelines for Sterilizing Grade Membrane Filters

Membrane filters can be sterilized by chemical sterilants (such as ethylene oxide, hydrogen peroxide in vapor form, propylene oxide, formaldehyde, and glutaraldehyde), radiant energy sterilization (such as γ-irradiation) or steam sterilization. The most common method of sterilization is steam sterilization.

Steam sterilization of a membrane filter can be accomplished either by an autoclave or by in situ steam sterilization. In situ steam sterilization can be effectively accomplished by a variety of different process arrangements. Steam sterilization is often the most critical portion of the process, and it is important that the procedures followed lead to sterilization of the system and do not impart any damage to the membrane filters. Some general procedural recommendations are included for prevention of either forward or reverse pressurization damage.

Reverse pressurization conditions can be prevented by using a noncondensing gas, such as air or nitrogen, at the end of a steam cycle. If this step is not followed, and the steam valve is shut off without the introduction of air or nitrogen, the filter housing will act like an isolated system. As such, the temperature will be different on the upstream (outside of the housing) and downstream sides of the filter. Due to the temperature difference, steam will condense at different rates on the upstream and downstream sides of the membrane. This can lead to pressure differences of 5 to 15 psi in the reverse direction. Even if this reverse pressure condition exists for a short period of time, the filter can suffer permanent damage.

Condensation during the steam cycle can lead to long cycle times and filter damage by excessive forward pressurization. If during steam sterilization the filter membrane is wet with condensate and in order to overcome the resistance of the wetted membrane the differential pressure in the forward direction is greater than required maximum (typically 5 psi), then the filter can be damaged by excessive forward pressurization. In this case, the filter will appear crushed. Precautions should be taken to minimize condensate accumulation. Drains for condensate removal should be strategically placed and should be cracked during the steam cycle.

CONCLUSION

This chapter has provided information on a variety of air filtration applications: fermenter inlet air, fermenter vent gas, vents on water-for-injection tanks, and vacuum break filters during lyophilization. In addition to a description of the air filtration applications, guidelines were provided on the usage of the filters in these applications.

BIBLIOGRAPHY

This is a modified version of the article first published as H. Haughney, "Air Filtration" in T. Meltzer and M. Jornitz eds. *Filtration in the Biopharmaceutical Industry*, Marcel Dekker, Inc., New York, 1998, Chapter 22. Reprinted by courtesy of Marcel Dekker, Inc.

1. H. Haughney, in T. Meltzer and M. Jornitz eds., *Filtration in the Biopharmaceutical Industry*, Marcel Dekker, Inc., New York, 1998.
2. H. Haughney, in K.E. Avis ed., *Sterile Pharmaceutical Products: Process Engineering Applications*, Interpharm Press, Buffalo Grove, Ill., 1995, pp. 305–355.
3. J.M. Martin, A.M. Trotter, P. Schubert, and H. Katz, in B.K. Lydersen, N.A. D'Elia, K.L. Nelson ed., *Bioprocess Engineering: Systems, Equipment, Facilities*, Wiley Interscience, New York, 1994, pp. 317–370.
4. N.B. Bruckshaw, *Filtr. Sep.* **296** (May/June 1973).
5. P.F. Stanbury and A. Whitaker, *Principles of Fermentation Technology*, Pergamon Press, Oxford, U.K., 1984.
6. J.W. Richards, *Introduction to Industrial Sterilization*, Academic Press, London, 1968.
7. N.A. Fuchs, in C.N. Davies ed., *The Mechanics of Aerosols*, Macmillan, New York, 1964.
8. C.F. Bruno and L.A. Szabo, *Biotechnol. Bioeng.* **25**, 1223 (1983).
9. Pall Corporation, *Validation Guide for Pall Emflon®II Membrane Cartridges for Air and Gas Filtration*, Pub. TR-870b, 1994.
10. American Society for Testing and Materials (ASTM), *Standard Test Method for Determining Bacterial Retention of Membrane Filters Utilized for Liquid Filtration*, ASTM Standard F838-83, 1983.
11. Food and Drug Administration, Center for Drugs and Biologics, Rockville, Md., *Guideline on Sterile Drug Products Produced by Aseptic Processing*, 1987.
12. T. Uberoi, *Chem. Eng.* 75 (March 1992).
13. R.S. Conway, *Biotechnol. Bioeng.* **26**, 844 (1984).
14. C.A. Perkowski, *Biotechnol. Bioeng.* **25**, 1215 (1983).
15. Porter, in C.H. Chilton and R.H. Perry, *Chemical Engineers Handbook*, 5th ed., vol. 26, McGraw-Hill, New York, vol. 26, 1973.
16. Pall Corporation, *Process Parameters on an Economic Use of Membrane Filters in Fermentor Exhaust Gas Filtration*, Scientific and Technical Report (STR) 15, Pall Corporation, East Hills, N.Y., 1992.

17. R.S. Conway, in C. Cooney ed., *Comprehensive Biotechnology*, Academic Press, New York, 1984.

See also STERILIZATION-IN-PLACE.

FILTRATION, CARTRIDGE

HOLLY HAUGHNEY
Pall Corporation
Long Island, New York

KEY WORDS

Cartridge
Filtration
Fluid
Membrane
Sterilization

OUTLINE

INTRODUCTION

Bioprocess technology operations have included membrane cartridge filtration for more than 20 years. Cartridge filters have been used for the removal of submicron particles, such as bacteria, and for the removal of particulate material, usually greater than 1 μ. Many applications involve the removal of bacteria in sterile filtration applications; particulate removal cartridge filters can be used as prefilters to protect the sterilizing grade filters. Some examples of applications for liquid filtration include fermentation media, deionized water, sera, and sanitizing chemicals. The purpose of this article is to provide information on operational considerations for cartridge filters and to explain the usage of cartridge filters in bioprocessing. (1–3)

DESCRIPTION OF MEMBRANE CARTRIDGE FILTRATION

Cartridge filters used in bioprocessing applications are used for the removal of bacteria (sterilizing grade filtration) or for the removal of particulate material (prefiltration). The mechanisms at work during filtration are provided, as well as, a general description of sterilizing grade filters and prefilters. Prefilters are often used upstream of the final (sterilizing) filters to protect the final filter from premature plugging, thereby prolonging the life of the final filter.

Filtration Mechanisms

There are four mechanisms that can govern the particle removal efficiency of membrane filters: (*1*) direct interception, (*2*) inertial impaction, (*3*) Brownian motion or diffusional interception, and (*4*) electrostatic attraction between the membrane and particles (4).

Direct interception involves a sieving action that mechanically retains the particles on the filter surface. The filter acts as a screen that stops particles that are larger than the pores (openings) in the membrane. Direct interception is independent of face velocity and mostly involves particles with relatively large diameters.

Inertial impaction refers to the deviation of a particle from its streamline during flow because of inertia, resulting in the retention of the particle by the membrane. As the face velocity increases, the probability of inertial impaction increases.

Brownian motion or diffusional interception applies to small particles at low face velocities in a gas stream. When air molecules are in a state of random motion, small particles suspended in the air can be struck by moving air molecules and displaced. The movement of particles resulting from molecular collisions is known as Brownian motion. This phenomena can increase the probability of capture of particles by diffusional interception within the membrane.

An electrostatic attraction can exist or be induced between the membrane material and particles in the fluid stream. This attraction can enhance the removal of particles and their retention by the filter membrane.

The pores (open area) in most membrane filter materials are typically not straight through the membrane. In membrane filters, the membrane can be considered to be constructed of multiple screens that provide a tortuous path for the fluid through the membrane.

In bioprocessing applications, there are many processes that require bacterial removal (sterile filtration). The fil-

ters used to provide bacterial removal are generally 0.2-μ sterilizing grade filters. (The tests used to ensure that the filters do remove the bacteria are described below)

Removal of bacteria by a 0.2-μ sterilizing grade filter is not only by a sieving action on the surface, but also by size-exclusion entrapment in the membrane structure. As shown in Figure 1, a membrane filter structure is sponge-like. The ability of such a membrane to remove bacterial contaminants is related to its thickness as well as to the size of its pores. For example, it has been demonstrated that whereas a single 150-μ thick layer of 0.8-μ cellulose ester membrane can provide only a titer reduction of 70 for *Brevundimonas (Pseudomonas) diminuta*, 10 layers can provide a titer reduction of 2.4×10^8 (5).

Because membrane filters are constructed of porous materials, they will have a certain amount of open area. The open area allows for fluid flow through the membrane; this area, expressed as a percent, is known as the porosity of the membrane.

Sterilizing Grade Membrane Filters

Construction. Membrane filters used for sterile liquid filtration are typically constructed of polymeric microporous hydrophilic materials, such as nylon, cellulose acetate, modified polyvinylidenefluoride (PVDF), or other polymers. Most liquids that require sterile filtration are water based, and hydrophilic membranes will spontaneously become wet with the fluid. Absolute rated hydrophilic membrane filter cartridges have been used for sterile filtration of parenterals, diagnostic reagents, purified water and water-for-injection, dry gases, organic solvents, buffers, and biological fluids such as serum, plasma, tissue culture media, nondilute protein solutions, and fermentation harvest fluids (1–3).

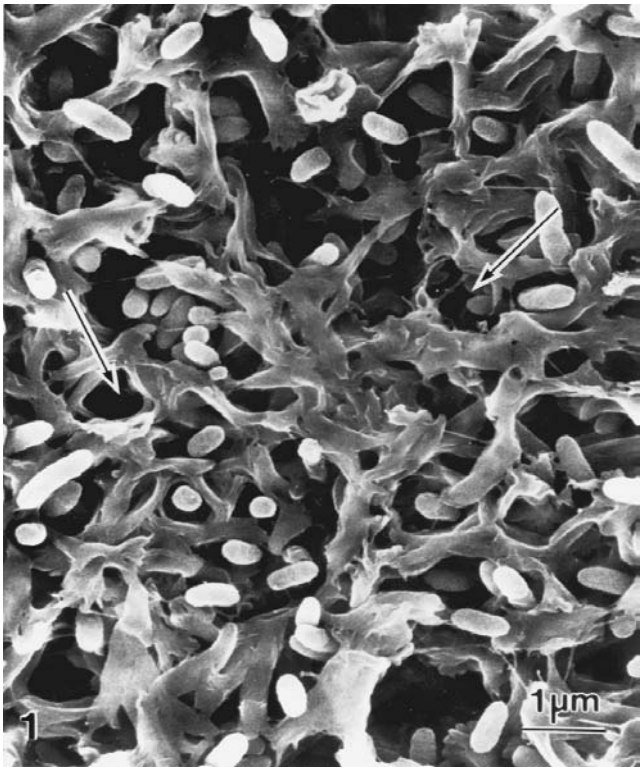


Figure 1. Membrane filter structure. Courtesy of Pall Corporation (20).

Membrane filters used for the sterile filtration of gas streams, such as nitrogen blankets, storage tank sterile vents, formulation tank sterile vents, sterile air for aseptic packaging, sterile filtration of fermenter air, or vent gases, typically contain a membrane made of hydrophobic materials such as PVDF or PTFE (polytetrafluoroethylene). Hydrophobic membrane filters are desired in these sterile gas filtration applications because hydrophobic filters do not spontaneously wet with water. When a hydrophilic filter is wetted with water, it will not pass air until the water-wet bubble point of the filter is exceeded. This water-wet bubble point can be greater than 50 psi (see FILTRATION, AIR).

Membrane materials can be modified to optimize their performance for certain types of applications. Positive-charged hydrophilic membranes contain cationic (positive charged) functional groups that impart a positive ζ-potential when immersed in an aqueous solution and provide enhanced retention of particles smaller than the absolute rating. Most particles have a negative charge, and the positive charge provides a mechanism, in addition to the tortuous path of the membrane, by which the particles are retained. Applications involving charged membrane filters are discussed in the Refs. 6, 7, and 8.

Membrane surfaces can also be modified to expose hydroxyl groups, which tend to prevent the retention of proteins. This is a benefit for dilute protein solutions (typically less than 1 mg/mL), in which the protein is the desired product. Such fluids include serum-free tissue culture media and protein additives, protein-based therapeutics and diagnostics, dilute protein containing diluents and buffers, recombinant proteins, hormones and growth factors, protein chromatography feeds and eluates, vaccines, and other dilute biologicals. Filtration applications involving low protein-binding membrane filters are described in Refs. 9 and 10.

The membranes used for sterile filtration bioprocessing applications are usually pleated and formed into a cylinder. Figure 2 illustrates the pleated cartridge construction.

The filter membrane is typically cast onto a support material (e.g., nonwoven polyester or polypropylene substrate) that provides high tensile strength while retaining flexibility. Layers of the membrane are corrugated, or pleated, along with upstream and downstream layers of courser material (e.g., nonwoven polyester or polypropylene). These layers provide support and drainage for the membrane. The sides of the corrugated membrane pack are sealed, often by a heat seal.

The corrugated filter membrane pack is fitted around the core. The purpose of the rigid, inner core is to provide support for the filter element against pressure in the forward direction. An external cage is provided for additional support and protection during handling. The flow path for the cartridge filter is from the outside (cage) to the inside (core). The material used for the core, cage, and support layers will depend on the intended application for the filter. Polypropylene, polyester, and PTFE are examples of ma-

Figure 2. Pleated cartridge configuration. Courtesy of Pall Corporation (29).

terials that can be used for the cage, core, and support material in a membrane filter. The membrane pack ends are sealed, usually by heat, with end caps. The end caps are constructed of the same types of materials as the core and cage. The final step in the construction of a membrane filter is to attach the proper end closure and O-ring adaptor by an appropriate welding technique.

The end closure is used for placement of the filter in a filter housing. A finned or flat blind end cap on one end is generally attached on one end and an open outlet at the other end. Sealing into the housing is accomplished with a double O-ring seal at the open end to ensure no fluid bypass. Filter elements can have locking tabs at the double O-ring adaptor base for bayonet locking of the element to ensure positive sealing in the housing base.

Most single filter elements are 10 inches in length. For the construction of a multiple length cartridge (20, 30, or 40 inches), the sections are attached by welding 10-in. subassemblies end to end. Figure 3 shows the construction types.

Requirements. For sterile filtration processes, the biological safety of the membrane filter or filter cartridge should be demonstrated by the performance of the USP ⟨88⟩ Class VI (121 °C) *Plastics Test for Biological Reactivity* (11). Another typical qualification used to select filter materials of construction is a listing for food contact in CFR Title 21 Part 177 (12).

For sterile aseptic processes, a typical requirement is sterilization by passage through a 0.2-μm sterilizing grade membrane filter, as defined by ASTM Standard F838-83 (13).

Figure 3. Cartridge filter construction. Courtesy of Pall Corporation (30).

A membrane filter as well as all system components can contribute extractables to the process stream. Many filter manufacturers document the extractables level of a particular filter in an appropriate solvent as a nonvolatile residue (NVR). These extractables are typically composed of the oligimers or additives of the plastic materials present in the filter element. It is important to note that the amount of extractables from a filter element as well as any extractable from the rest of the process system can be reduced by flushing the system before filtration.

Many filter manufacturers address the issue of effluent quality requirements by the performance of appropriate tests on filter samples from manufacturing lots. These include tests for

- *Cleanliness*: Per current USP limits under *Particulate Matter in Injections* ⟨*788*⟩ and conformance with requirements for a non-fiber-releasing filter per Title 21 of the U.S. CFR Title 21, Part 211.72 and Part 210.3(b)(6) (14)
- *Oxidizable substances*: Per current USP requirements under *Purified Water* after flushing (15)
- *pH*: Per current USP requirements under *Purified Water* after flushing (15)
- *Pyrogens*: Per current USP requirements under *Bacterial Endotoxins Test* as determined using the limulus ameobocyte lysate (LAL) reagent with an aliquot from a soak solution (16)

Prefiltration

Cartridge filters used for prefiltration can be either pleated or nonpleated (depth) filters. Standard design depth filter cartridges are open cylinders of thick filter material (double open-ended). Prefiltration filter cartridges are typically available in 2-1/2 or 2-3/4-in. o.d., in multiple lengths of 10 in. up to 40 in. long.

Prefiltration encompasses a wide variety of fluid types and removal requirements. The following is a discussion of some important characteristic of prefilter construction and a sampling of some of the types available.

The particles or fibers in the prefilter membrane should not become dislodged, slough off, shed, or in some other way contaminate the filtrate (media migration). The filter medium should also not unload or release contaminants retained by the filter medium into the effluent as the differential pressure across the membrane increases. Data that indicate filtration efficiency as a function of increasing differential pressure should be available from the vendor to support the claim of nonunloading.

Polypropylene Depth Type Filter Cartridges. Polypropylene depth filters are made using nonmigrating continuous filaments of polypropylene filament medium without resin binders. Absolute rated polypropylene depth filter cartridges are applicable as prefilters for serum, vaccines, diagnostics, tissue culture products, deionized water, container washing, and final product. These filters are used in fermentation for liquid feeds, makeup water, solvents, and antifoam. In downstream processing, they are used for cell and cell debris removal, buffers, cleaning agents, and sanitizing solutions. Positive charged versions of these filters are possible and have the added ability to remove organisms and particles such as bacterial endotoxins (pyrogens) that are much smaller than the absolute pore size rating.

The membrane for these depth filtration cartridges has an inner (downstream) section in which the pore diameter is constant (this section provides absolute rated filtration) and an outer section in which the pore diameter varies continuously from that of the absolute rated section to as much as 90 μm or more. Pore size variation of the filter cartridge is achieved by varying the fiber diameter while maintaining constant void volume throughout the medium. The constant void volume provides increased dirt capacity and a low clean pressure drop of the filter cartridge.

Polypropylene Pleated Filter Cartridges. Polypropylene pleated filters are applicable as prefilters for makeup and rinse water, for serum for culture media, and in downstream processing for cell debris removal and prefiltration of solvents and buffers. These process filter cartridges can be constructed of the same type of membrane material as depth filters.

The thin sheet of polypropylene media is pleated and formed into a cylinder with a longitudinal side seal of melt seal polypropylene. The cylinder is then melt sealed to injection molded polypropylene end caps to ensure no fluid bypass. Polypropylene hardware components consisting of an inner support core and an external protective outer cage are incorporated. The process filter cartridge with polypropylene hardware should be rated to withstand differential pressure of 80 psi up to 50 °C (122 °F) and 60 psi up to 80 °C (176 °F).

Resin-Free Cellulose Pleated Filter Cartridges. Resin-free pleated cellulose filter cartridges are applicable as prefilters for makeup and rinse water, reverse osmosis membranes, deionized water, and inlet air for fermenters and bioreactors. These filter cartridges are constructed of pure cellulose medium without resin binders, which is pleated into a high area cylinder. The longitudinal side seal of the process filter cartridge should be a polypropylene (approved for food contact usage).

Cellulose media cartridges are assembled with hardware components consisting of a perforated inner support core, an outer support cage, and end caps melt sealed to imbed the medium in the plastic. All hardware components should be of pure polypropylene, without filler or reinforcement, to ensure a high quality filtrate with a minimum of soluble extractables. The process filter cartridge with polypropylene hardware should be rated to withstand differential pressure of 80 psi up to 50 °C (122 °F) and 60 psi up to 80 °C (176 °F).

Resin-Impregnated Fiberglass Pleated Filter Cartridges. Resin-impregnated fiberglass filters are applicable as prefilters for deionized water, solvents, serums, diagnostic reagents, and cell debris removal from harvest fluids. The resin impregnation makes available a choice of filters with either a positive or negative ζ-potential when used in aque-

ous service. Filter cartridges with negative ζ-potential are constructed of glass fiber that has a natural negative ζ potential and reinforced with a resin binder that is also naturally negative when immersed in water. Positive ζ potential filter cartridges are constructed using a resin binder that coats the glass fibers and imparts a positive ζ potential to the medium.

The filter cartridges are available with hardware components consisting of end caps and an internal perforated support core made of either polypropylene or stainless steel and in an external polypropylene outer cage or protective net. Stainless-steel end caps are attached to the filter by an inert synthetic resin. Polypropylene end caps are melt sealed to imbed the medium in the plastic. Filter cartridges with polypropylene hardware should be rated to withstand differential pressures of 80 psi up to 50 °C (122 °F) and 60 psi up to 80 °C (176 °F). With stainless-steel hardware, the filter can be rated to 75 psi up to 135 °C (275 °F).

Filter Assemblies Used in Processes

During filtration, the membrane filter cartridges are placed in housings that provide a means of fluid contact with the filter cartridge. The housings are typically composed of 304, 316, and 316L stainless steel; plastic housings are also available. The housing design is typically composed of a head and a bowl. The head is the portion of the housing in which the filter is attached. The bowl is clamped to the head. A housing is illustrated in Figure 4.

Membrane filters can also be obtained in complete assemblies. These disposable assemblies consist of a membrane filter cartridge welded into a plastic filter housing. Disposable assembles or capsule filters are typically used in small-scale applications (e.g., 100-L batches or less, depending on the fluid) and are particularly useful in applications in which operator contact with the fluid is not desirable, because the entire assembly can be removed from the system and discarded without operator contact with the product-wet cartridge. Figure 5 shows disposable assemblies.

Figure 4. Housing configurations. Courtesy of Pall Corporation (31,32).

MEMBRANE FILTER RATINGS AND TESTS

Many filter manufacturers use a nominal micron (μm) rating for particle removal efficiency. This is defined by the American National Standards Institute (ANSI) as an "arbitrary micrometer value indicated by the filter manufacturer. Due to lack of reproducibility this rating is deprecated." Further, nominal rating standards are arbitrary and a comparison of nominally rated filters is imprecise. In addition, nominal ratings can be misleading because the filter can allow passage of particles larger than the rating indicates.

In order to establish a more meaningful filter rating, a test is performed on the filter using an appropriate contaminant to meet a specific claim for retention of the contaminant. The test procedure must be sensitive enough to detect the passage of contaminants of interest. The testing should be performed under carefully controlled conditions using industry-accepted reference standards. These reference standards include silica suspensions, latex beads, or microorganisms.

Particulate Removal Filters

The removal ratings for particulate removal filters (typically used as prefilters in sterile filtration processes) can

Figure 5. Disposable cartridge assembly. Courtesy of Pall Corporation (29).

be established using the Oklahoma State University (OSU) F-2 Test. This rating method (ISO 4572, ANSI B93.31) has received wide acceptance for use on lubricating and hydraulic fluids. The technique has been used for oils extensively and has been adapted for use in water with contaminants ranging from 0.5 to 25 μm.

The test is based on continuous on-line particle counts of different particle sizes, both in the influent and the effluent stream. The β-ratio at a specific particle size is defined as β_x: the number of particles of a given size (X) and larger in the influent, divided by the number of particles of the same size (X) and larger in the effluent, where X is the particle size in micrometers. The percent removal efficiency can be calculated from the β-value. The percent removal efficiency is $[(\beta_x - 1)/\beta_x]$ 100 (17).

Sterilizing-Grade Membrane Filter Ratings and Integrity Tests

Sterilizing-grade hydrophilic membrane filters require testing using a liquid bacterial suspension challenge that is sensitive enough to detect the passage of any microorganisms, for the establishment of the micron rating of the filter. A sterilizing-grade filter is defined by the FDA's *Guideline on Sterile Drug Products by Aseptic Processing* (18) as one that will produce sterile effluent when challenged with the test organism *Brevundimons (Pseudomonas) diminuta* (ATCC 19146) to the level of greater than 10^7 CFU/cm^2 filter area. Such microorganism retention tests are conducted with *P. diminuta* (ATCC 19146), measuring 0.3 $\times$ 0.6 to 0.8 μm, to verify that membrane filters produce sterile fluid filtrates. This is a standard test for the validation of sterilizing-grade filters (0.2-μm pore size rating) in the bioprocessing industry.

In practical terms, there is a limit to how many bacteria a filter element can be challenged with before it becomes plugged. When a filter has been challenged at the level of 5×10^9 CFU/cm^2 it becomes effectively clogged. This challenge level corresponds to about 2×10^{13} CFU for a 5-ft^2 membrane filter (17).

The liquid test involves challenging a test filter with a known quantity of *P. diminuta*, no less than 1×10^7 organisms per cm^2 of filter area. The challenge sample is suspended, for example, in sterile water and then filtered at a defined flow rate and time through a test filter, followed by an analysis membrane. After the challenge is complete, the system is flushed for 5 to 10 min with sterile water. The analysis membranes are removed and placed on Mueller Hinton agar and incubated at 32 °C for 48 h. After incubation, the plate is examined for the presence or absence of microbial colonies. Figure 6 illustrates a liquid bacterial challenge test stand.

For some applications (e.g., mycoplasma control or removal) a membrane with a more stringent rating (0.1 μm) is required. These applications typically involve sera and tissue culture media. One organism that has been used extensively in validating 0.1-μm rated filters is *Acholeplasma laidlawii* (19).

A liquid bacterial challenge is a destructive test; the filter cannot be used after the challenge. Therefore, a correlation is made to a nondestructive test. It is the correlation, not the nondestructive test alone, that provides assurance that the filter performs as required. During the production of a sterile product, the filter should be tested before and after the filtration to ensure that the filter meets the specification, is properly installed and intact, and confirms the filter rating.

The three major tests used to determine the integrity of a membrane filter are the bubble point, forward flow, and pressure-hold integrity tests. These tests are based on the flow of a gas (air or nitrogen gas) through a liquid wetted membrane under applied gas pressure. A fourth test, the water intrusion test, is sometimes used to determine the integrity of hydrophobic membrane filters.

For a nondestructive test to be useful, the results of the test must be able to predict the ability of the filter to remove bacteria. Such a test should also be easy to perform and highly reproducible. The manufacturing quality control of the membrane filters tested to obtain this prediction must be used to establish quality control specifications that are consistently maintained in filter manufacturing. It is the combination of the nondestructive integrity test, the correlated bacteria challenge test, and manufacturing quality control that allows membrane filters to be reliably used for sterile filtration of bioprocessing products.

Forward Flow, Pressure Hold, and Bubble Point Tests. The industry-accepted, nondestructive tests used to verify sterilizing grade filter integrity are the forward flow, pressure hold, and the bubble point tests (20). These tests are performed by applying a preset air (or nitrogen gas) pressure to a filter wet with an appropriate fluid. For hydrophilic filters (filters that wet readily with water) the wetting fluid is typically water or a water-based fluid; for hydrophobic membrane filters (filters that do not spontaneously wet with water and are used in air service) 60/40 isopropyl alcohol/water is a typical wetting fluid.

The wetting fluid fills the voids in the membrane. For this to be occur uniformly, the wetting must be complete. The result of the wetting can be considered as a layer of water supported by the membrane. When gas pressure is applied to one side of the membrane, the test gas will dissolve into the liquid layer to an extent determined by the solubility and pressure of the gas (as described by Henry's law [21]). Downstream of the membrane, the pressure is lower, and the gas in the liquid is driven out of solution. The result is a net flow of gas through the membrane. The forward flow, pressure hold, and the bubble point tests all involve the observation or measurement of a gas flow through a wetted membrane. The measured flow is a com-

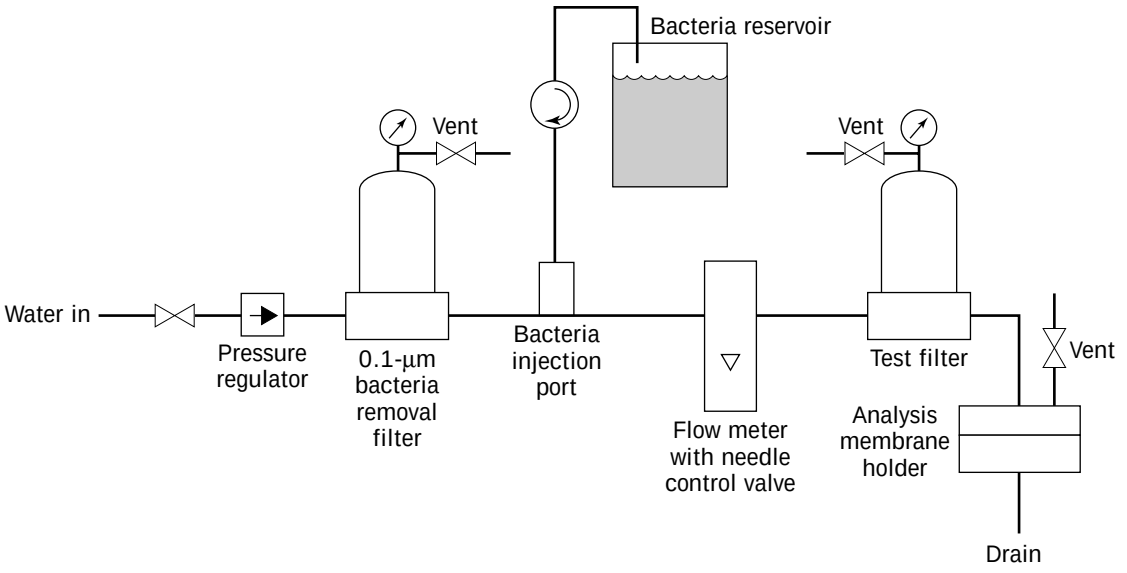


Figure 6. Liquid bacterial challenge. Adapted from "Engineering Considerations in Sterile Filtration Processes" by Holly Haughney in *Sterile Pharmaceutical Products: Process Engineering Applications*, edited by Kenneth E. Avis. Copyright © 1995 by Interpharm Press, Inc. All rights reserved. Adapted by permission.

bination of the diffusion of the gas through the liquid layer and bulk gas flow through any open passageways.

The rate of diffusive flow is a function of the test pressure, the diffusivity of the gas in the liquid, the solubility of the gas in the liquid, the membrane void volume, and the membrane thickness. This diffusive flow is described by Fick's first law (13), which is presented below in an integrated form:

$$N = \frac{DH\Delta P\rho}{L} \tag{1}$$

where N is the permeation rate, D is the diffusivity of the gas in the liquid, H is the solubility coefficient of the gas, L is the thickness of the liquid in the membrane, ΔP is the differential pressure, and ρ is the void volume.

As the applied upstream gas pressure is increased, the diffusive flow increases proportionally. At some point, the pressure becomes great enough to expel the fluid from a passageway and establish a path for bulk flow of gas. As a result, the gas flow through the wetted filter begins to increase in a nonlinear manner due to the initiation of bulk gas flow through these open paths. The pressure at which this occurs can be modeled by a modified capillary rise equation:

$$P = F\,\frac{4\gamma\,\cos\theta}{D} \tag{2}$$

where P is the pressure (dynes/cm^2), γ is the surface tension (dynes/cm), θ is the contact or wetting angle, D is the diameter (cm), and F is the shape factor (to account for the structure of the membrane). As the pressure is further increased, more passages are opened and the proportion of the total gas flow due to the bulk air flow increases. If the pore size distribution is narrow, then the possible passages through the membrane will provide similar resistance to flow.

For a wetted high-area pleated membrane filter, a plot of differential gas flow pressure versus gas flow will show a gradual transition from diffusive to bulk flow. An alternate plot of differential pressure versus gas flow or differential pressure will reveal a rapid transition from diffusive flow to bulk flow, as seen in Figure 7. This graph is referred to as a diffusive flow spectrum or a K_L (knee location) curve. The K_L pressure represents the transition from the diffusive flow region to the bulk gas flow region.

The bubble point test is a test designed to determine the pressure at which a continuous stream of bubbles is initially seen downstream of a wetted filter under gas pressure. There are several ways of performing this test. The first of these is the visual bubble point test. Generally, the

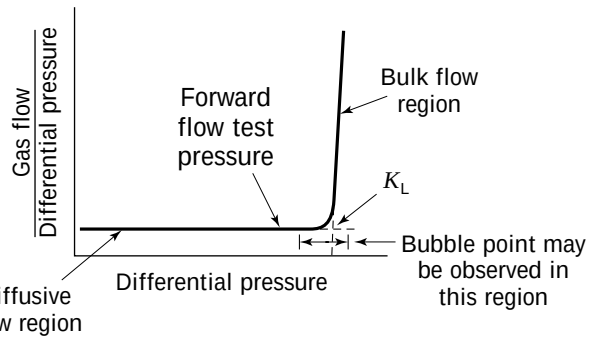


Figure 7. Diffusive flow spectrum. Courtesy of Pall Corporation (20).

visual bubble point test is not used except on small area disks where the whole membrane surface may be examined. To perform a visual bubble point test, gas pressure is applied to one side of a wetted membrane, with a layer of the wetting fluid maintained on the other side. The gas pressure is slowly increased. The rate of increase will affect the measured bubble point value. At some point, the pressure becomes high enough to displace the liquid from one or more passageways. This establishes the beginning of bulk flow of air through the membrane, and a thin stream of bubbles can be observed on the downstream side of the membrane. The pressure at which this stream of bubbles is observed is referred to as the first bubble point.

A manual bubble point is measured with a filter installed in a housing, with tubing running from the housing outlet to a beaker of water as shown in Figure 8. The geometry and position of the filter housing, and the diameter and length of the tubing will affect the measured bubble point value. In this case, the bubble point is considered to be the pressure at which a steady stream of bubbles is seen coming from the tubing; however, the determination of a steady stream is often subjective.

The manual bubble point (Fig. 9) will generally occur at a higher pressure than the first bubble point determined by a visual examination of the filter surface, because greater flow is necessary to obtain a clear stream of bubbles from the volume of the tubing. When determining the manual bubble point on a high area cartridge, the diffusion of air through the wetted filter at pressures below the bubble point can be significant. The complications resulting from the length and diameter of the tubing, the housing, and operator interpretation render bubble points determined in this manner not easily reproducible or comparable to bubble points determined by other techniques.

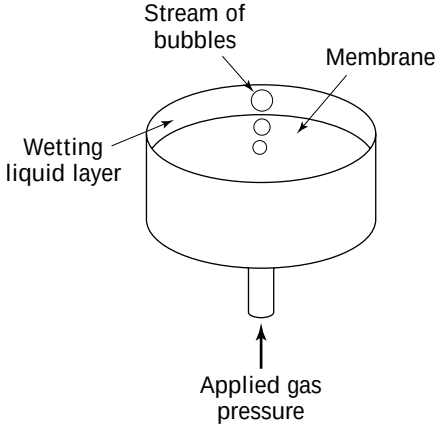


Figure 8. Bubble point test. Courtesy of Pall Corporation (20).

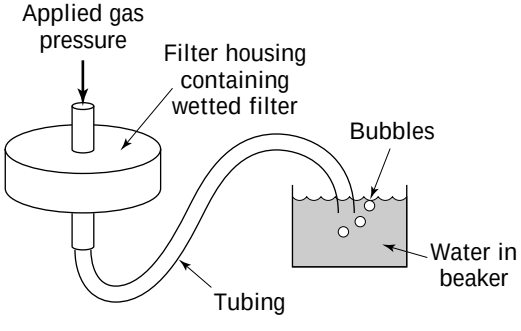


Figure 9. Manual bubble point test. Courtesy of Pall Corporation (20).

The forward flow test and the related pressure hold test were first introduced by Pall Corporation in 1973 (22). During these tests (Fig. 10), a wetted membrane is subjected to a predetermined gas pressure on the upstream side of the membrane. The gas will diffuse through the wetted membrane at a measurable rate. During the forward flow test, the diffusion of the gas and flow through any open pores is measured at the specified test pressure.

The pressure hold test (Fig. 11) is a modification of the forward flow test. The upstream side of filter housing is pressurized at the test pressure and then the filter is isolated from the pressure source. The diffusion of gas through the wetted membrane and flow through any open pore is measured as a decay in pressure over a specified period of time.

The pressure chosen as the test pressure for the forward flow or pressure hold test must be close enough to the K_L to allow differentiation between different grades of membrane and high enough to present a measurable flow. It is also necessary that the test pressure chosen be such so that the background diffusive flow does not interfere with the sensitive detection of defects in the filter. The exact pressure chosen for the forward flow test must be below the K_L. The test pressure is set by filter manufacturers and should be used during tests performed to establish the correlation with bacterial retention.

Integrity tests can be performed with automated equipment. Automated filter integrity test instruments (e.g., TruFlow℠, Integritest℠, Palltronic℠, and Sartocheck℠) have been developed to provide accurate and reproducible filter integrity test values. The instruments traditionally were designed to perform a pressure hold test. The forward flow value can be calculated from the pressure hold value. Newer equipment designs use direct upstream air flow measurements.

The tests are controlled and monitored by a built-in microprocessor, and the equipment often can be used with a programmable logic controller (PLC) for full system automation. An example of an automated integrity test application is provided in Ref. 23.

Automated devices cannot perform a true bubble point test. Instead, a modified bubble point test is performed, which essentially consists of a series of pressure hold tests. In the automated bubble point test, upstream air pressure is increased in steps. Between each pressure increase, the instrument isolates the upstream volume and monitors the pressure decay upstream of the filter. As the pressure decay exceeds a certain value, the software uses a specific algorithm for calculating the bubble point from relative values of this series pressure decays. The instrument uses a series of pressure hold tests as the basis for the calculation of the bubble point rather than measuring the bubble point directly. Figure 12 illustrates how pressure hold

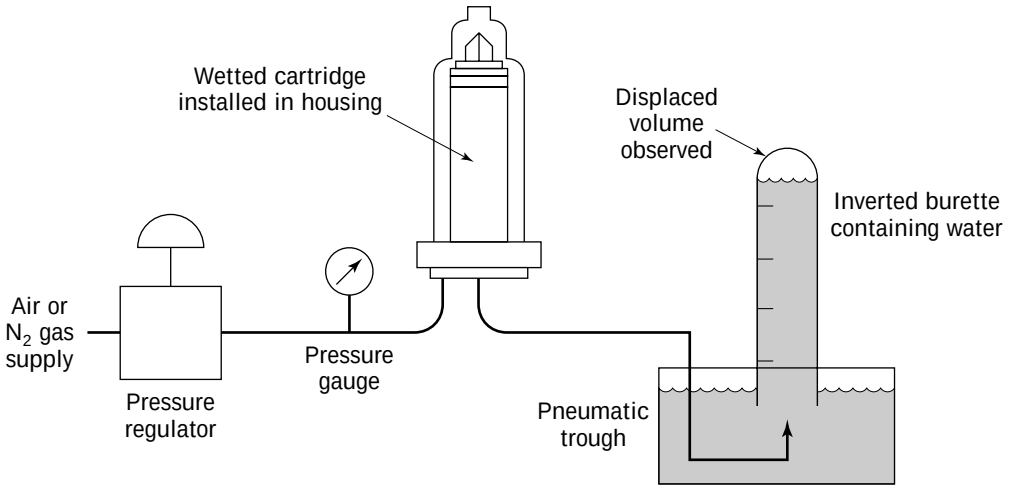


Figure 10. Forward-flow test. Adapted from "Engineering Considerations in Sterile Filtration Processes" by Holly Haughney in *Sterile Pharmaceutical Products: Process Engineering Applications*, edited by Kenneth E. Avis. Copyright © 1995 by Interpharm Press, Inc. All rights reserved. Adapted by permission.

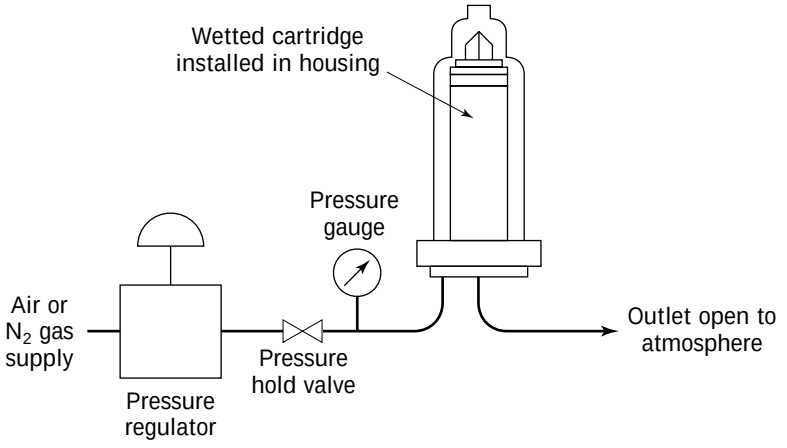


Figure 11. Pressure-hold test. Adapted from "Engineering Considerations in Sterile Filtration Processes" by Holly Haughney in *Sterile Pharmaceutical Products: Process Engineering Applications*, edited by Kenneth E. Avis. Copyright © 1995 by Interpharm Press, Inc. All rights reserved. Adapted by permission.

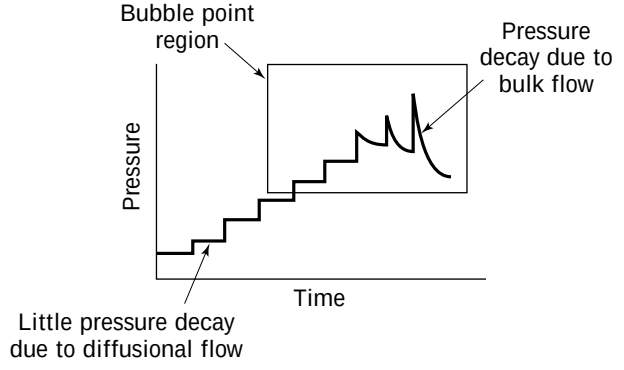


Figure 12. Pressure-hold test for bubble point determination with an integrity test device. Courtesy of Pall Corporation (20).

tests can be used to determine a bubble point. The actual algorithms used in the automated test devices may differ between automated filter integrity test instrument manufacturers.

Water Intrusion Test for Hydrophobic Filters. The water intrusion test can be used to test hydrophobic membrane filters (24–27). To perform a forward flow, pressure hold, or bubble point test on a hydrophobic filter, the membrane must be wet with an alcohol solution (e.g., 60/40 IPA/water). For an in situ integrity test, the wetting agent has to be added to the system to perform the test and then removed to restore air flow through the filter. The potential flammability of the alcohol is a consideration, and for some applications it can be necessary to validate the complete removal of the solvent to avoid contamination of the product. For these reasons, integrity testing of hydrophobic filters in situ has not been widely used, and off-line testing

has been more common. The increasing industry demand and growing regulatory requirement for in situ testing of sterilizing-grade air filters has led to the development of water-based tests.

If the upstream side of a dry hydrophobic filter assembly is filled with water and pressurized, the hydrophobic nature of the porous membrane will prevent the bulk flow of water through the filter until the intrusion pressure is reached. At pressure below the intrusion pressure, a small but measurable flow of water through the membrane occurs.

The presence of larger pores in the filter will be detected by an increased flow due to bulk water flow through these pores (Fig. 13). This principle is the basis of the water intrusion test.

At the start of the test, the upstream side of the filter assembly is filled with water and a predetermined gas pressure is supplied. This causes a fall in the water level within the filter housing because of factors such as pleat compression and the passage of trapped air through the membrane. It is important that the changes in the upstream volume caused by these effects have stabilized so that the very small flows of water through the membrane can be accurately measured. A typical test arrangement is illustrated in Figure 14.

Adequate stabilization time (e.g., 10 mins) is also required to ensure thermal equilibrium. Water flow can be measured by determining the upstream flow of gas required to keep the pressure constant. The use of a direct flow measurement integrity test instrument can be used for water intrusion test measurements. After the test is over, the water can be drained and in many cases the filter is ready for use. Where dry air or gas is required, a drying procedure can be used.

FILTRATION SYSTEM DESIGN AND OPERATIONAL CONSIDERATIONS

Filter Selection

As detailed earlier, there are a number of different types of cartridge filter types for a wide variety of applications (1–3). In order to properly design a filtration system, all the requirements for the system and fluid must be considered. As a starting point, the removal efficiency must be defined. Other considerations can include process temperature, flow rate, pressure drop restrictions, integrity test regimen, and sterilization.

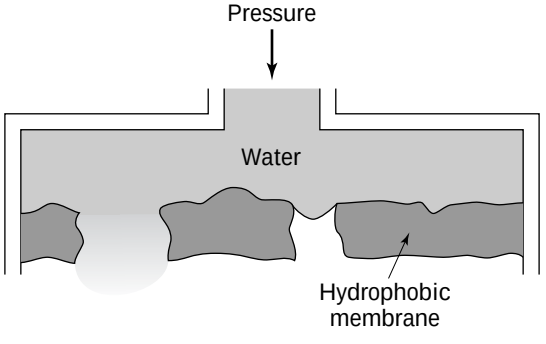


Figure 13. Bulk flow of water through pores. Courtesy of Pall Corporation (27).

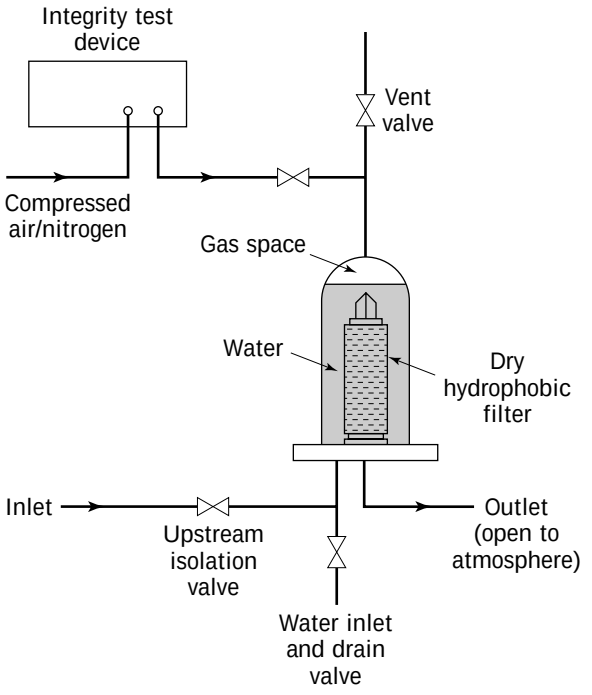


Figure 14. Water intrusion test. Courtesy of Pall Corporation (33).

All cartridge filter materials (membrane, cage, core, support material and housing) should be compatible with the process fluid over the temperature and pressure range indicated for a given process application. In general, most bioprocessing fluids are water based and are compatible with most membrane materials, such as polyamides (nylon), PVDF, polysulfone, and cellulose acetate.

For many filtration processes, it is desirable to have a staged filtration system. A staged system will have one or more prefilters of higher micron rating protecting the final filter. The final filter in the stage is the filter that provides the desired effluent quality. The prefilters are typically courser filters that remove larger contaminants in the fluid and reduce the loading of the final filter. For a sterile filtration process, the final filter is typically a 0.2-μm membrane filter. In a staged system, the life of the final filter is enhanced by the use of the prefilters. This will usually reduce the cost of the system, because the cost of the final filter is generally greater than the cost of the prefilters.

The size of the filtration system will depend upon the volume of the fluid being filtered and the ease with which the fluid is filtered or its filterability. For production applications, the membrane filters are typically provided as 10-in. segments or filter elements. These segments can be welded together to form 20, 30, and 40-in. filter elements. The diameter of the cartridge is approximately 3 in. Filter housings can hold a single cartridge (10- to 40-in. elements) or multiple filter cartridges. For applications that

involve relatively small volumes (for example 10 L), disposable filter assemblies (capsule filters) with smaller membrane areas are available. If prefiltration is required, then the prefilters will also be a part of the filter train. For many applications, there is a one-to-one correspondence of prefilters to final filters. Optimization for a particular process may indicate that fewer, or more, prefilters are needed.

In some production environments, it is desirable to have a dual filtration system (filters arranged in parallel) to allow for a changeover from one system to the other. While one system is being used, the other system can be prepared for use (e.g., it can be cleaned and tested and the filters can be changed out). The decision to use a dual filtration system requires a balancing of the ease of operation with the additional space required for the additional system. For some sterile applications, redundant final 0.2-μm filters (filters in series) are used to provide additional assurance of sterility. If one of the filters fails, because of processing conditions, then the other filter is in place as a backup. If the backup filter is integral, then the fluid has been sterile filtered. The decision point is whether further assurance of maintaining sterility throughout the entire filtration processing time can justify a redundant system.

Flow Rate and Pressure Considerations

All components of a system will cause a resistance to flow, which results in a pressure drop (ΔP). Pressure drops, or pressure losses in a system, can be caused by piping, connections, valves, and filling heads and by the filter and its assembly. The pressure drop through a clean filter assembly is due to the sum of component pressure drops, including the filter housing, the filter hardware, and the filter membrane. The filter membrane will cause a resistance to flow, or a pressure drop, in a filtration system. As the filtration progresses and the filter membrane becomes plugged with the particulate contaminants in the fluid, an increasing pressure drop will occur. The pressure drop will be different, depending on the fluid hydrodynamics (e.g., viscosity) and the filter membrane material characteristics (such as porosity).

An adequate pressure source is needed to overcome the pressure drop across the system, including the filter assembly and to allow for the maintenance of a constant flow rate. If the initial clean pressure drop across a filtration system is close to the maximum available pressure, the flow rate will not be maintained because as the filter plugs the pressure drop will increase and there will not be enough pressure to overcome the increased pressure drop throughout the filtration. This will cause a decrease in flow. If possible, where constant flow is required, pumping capacity should be increased. An alternative remedy would be to increase the membrane filtration area so that the initial pressure drop is reduced.

Ultimately, the filter will become plugged with contaminants. This will depend on the amount of the contaminant in the fluid that is being filtered. For most pharmaceutical sterile filtration applications the load (percent solids) is typically low. Filters upstream of the sterile filter are more likely to become plugged in a properly designed filter train. As plugging occurs, the differential pressure will increase. System design should include a change-out schedule so that the differential pressure constraints for the filtration system are not exceeded. For example, in a batch system, the filter change-out should occur after the batch filtration has been completed. If multiple in-line stages of filtration are needed, then the differential pressure for each stage must be added, and the sum must be less than the maximum differential pressure allowed for the filtration process to allow for increased pressure caused by membrane fouling.

At no time during a process should the pressure drop exceed the recommendations for pressure drop (usually specified for a temperature range of the individual filter unit). At elevated temperatures the recommended pressure drop is typically lower than the recommended pressure drop for ambient conditions. A combination of elevated temperature and a high pressure drop can lead to structural filter damage. Structural damage to the filter is possible under these conditions because at elevated temperatures the polymeric components of the filter can become softer and are more likely to be deformed if the pressure drop is too high. For example, under extreme conditions, the core of the filter as well as the cage can be crushed.

Required flow rate, maximum pressure drop, and available pressure for the filtration system must be considered when the system is designed. Flow rate is typically reported as either gallons per minute (GPM) or liters per minute (LPM). The flow rate achievable through a filtration system is directly related to the applied pressure and inversely related to the resistance to flow. Thus, if the applied pressure is increased, the flow rate will increase. If the resistance to flow is increased, such as by membrane plugging, then the flow rate will decrease. It is important to note that an initial higher pressure (and flow rate) can lead to premature membrane plugging with many products.

The nature of microporous membrane filters is that they will tend to plug rapidly if they are subjected to a relatively high flow rate (directly related to a high applied pressure) during the start-up of a filtration. For fluids with a significant particle loading, under conditions of initial high flow rate, the microporous membrane can become rapidly plugged, or fouled, and the pressure drop will increase and the throughput, or filter life, will be reduced. This is especially true for products that may contain gels, such as biological products. Filter life, or throughput, can be increased if the initial flow rate is reduced. If the initial flow rate is decreased, the pressure drop will increase at a slower rate and the throughput will be increased. Another approach is to increase filter surface area for the given flow rate. Additionally, the pressure drop across a filter assembly can be reduced by increasing the filtration area. Increasing filtration area can often be an economical approach because the increase in throughput is often greater than linear for an increase in filtration area. Figure 15 illustrates the concept of a lower flow rate and differential pressure providing a higher throughput (total volume filtered).

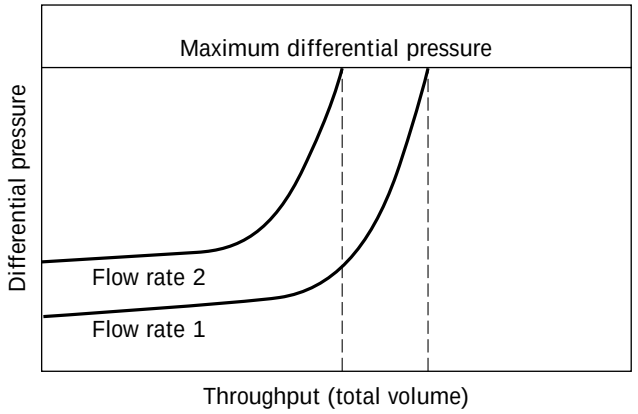


Figure 15. Comparison of a filtration performed with different flow rates. Adapted from "Engineering Considerations in Sterile Filtration Processes" by Holly Haughney in *Sterile Pharmaceutical Products: Process Engineering Applications*, edited by Kenneth E. Avis. Copyright © 1995 by Interpharm Press, Inc. All rights reserved. Adapted by permission.

Fluids with higher viscosities will be more resistant to flow and the pressure drop across the filter will be greater. For fluids with high viscosities, a greater applied pressure will be required to maintain the process flow rate. The smaller the pore size rating of a filter, for a fluid of a given viscosity, the greater the resistance to flow.

Scale-up

Although a particle size distribution evaluation can often provide useful information about the contaminant material in a process stream, the filtration system usually cannot be selected without an actual scaled-down filterability test. Actual testing is required because the particle size distribution fails to provide information on the quantitative particle load, its effect on membrane life, and how the membrane filters will perform with the process fluid under process conditions. In addition, other factors such as gels and high molecular weight biopolymers present in some products will also affect filter life and can only be evaluated by actual test.

The scaled-down, or filterability, test will require a relatively small representative sample of the process fluid (typically 2 to 4 L). Membrane filter disks with a 47-mm diameter are typically used. If a process stream is subject to variable composition, then the worst case fluid should be defined and tested. Only membrane materials that are compatible with the fluid are tested during the filterability test, and test conditions (e.g., temperature and flow rate) should match the process conditions as nearly as feasible. Several filtration schemes, including staged filtration schemes, are usually evaluated to optimize the filtration. The filterability tests can be performed at either a constant pressure or a constant flow rate. The use of constant pressure or constant flow rate must be considered because the system size, system life, process time, and throughput can depend on these process parameters. The flow rate can be scaled linearly based on filtration area. Optimization parameters must be defined and can include effluent quality, time (or throughput) to reach terminal (maximum allowable) differential pressure at a given flow rate, or flow rate for a given applied pressure.

System scaling can be based on the following ratios for the final filter in the system, where it is necessary to solve for full scale system area:

$$\frac{\text{Full scale system area}}{\text{Scaled down system area}} = \frac{\text{Full scale system flow rate}}{\text{Scaled down system flow rate}}$$

$$\frac{\text{Full scale system area}}{\text{Scaled down system area}} = \frac{\text{Full scale system throughput}}{\text{Scaled down system throughput}}$$

Throughput (total fluid filtration volume required), flow rate, and process time requirements must be met by the filtration system.

If the system involves prefiltration stages, the prefilter area will usually be equivalent to the final filter area; however, the size of each stage in a multistage system will depend on the overall requirements for practical and economical filtration of the product. Further optimization is possible by performing a filterability test with the prefilters alone. The issues associated with the sizing of a filter system with pre- and final filters are described next.

Figure 16 illustrates the behavior of a typical optimized (ideal) filtration system. The differential pressure (ΔP) is plotted as a function of volume filtered. The limiting or maximum differential pressure and the total batch volume are indicated on the graph. Ideally, the batch volume (throughput requirement) is reached before the limiting ΔP is reached. This allows a safety factor in the filtration and permits complete processing of the batch without change-out of the filter.

Figure 17 shows the effect of limiting ΔP versus volume for a filtration system in which the prefilter is too coarse and is not able to remove a sufficient number of particles in the fluid (the particles go through the prefilter instead

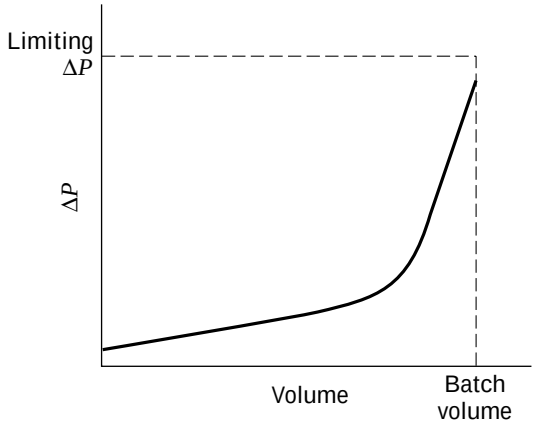


Figure 16. Example of an ideal filtration system. Adapted from "Engineering Considerations in Sterile Filtration Processes" by Holly Haughney in *Sterile Pharmaceutical Products: Process Engineering Applications*, edited by Kenneth E. Avis. Copyright © 1995 by Interpharm Press, Inc. All rights reserved. Adapted by permission.

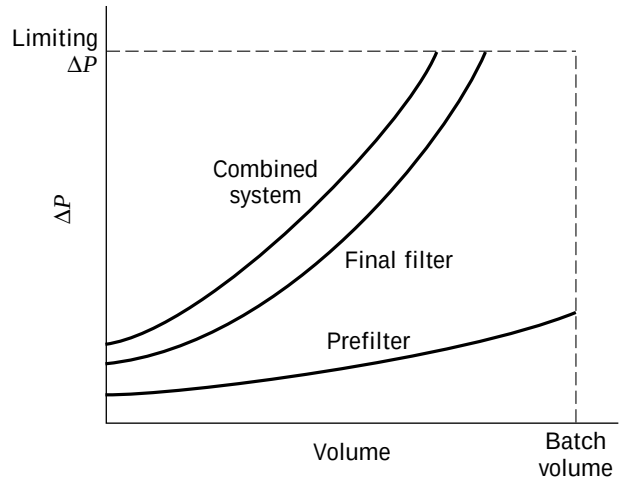


Figure 17. Example of a filtration system with the prefilter too course. Adapted from "Engineering Considerations in Sterile Filtration Processes" by Holly Haughney in *Sterile Pharmaceutical Products: Process Engineering Applications*, edited by Kenneth E. Avis. Copyright © 1995 by Interpharm Press, Inc. All rights reserved. Adapted by permission.

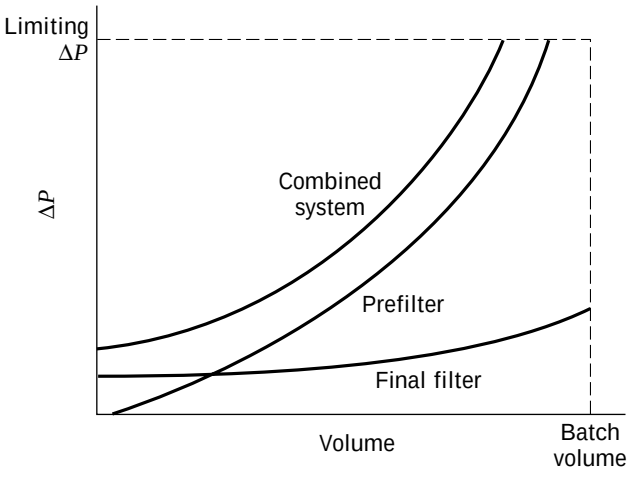


Figure 18. Example of a filtration system with the prefilter too small. Adapted from "Engineering Considerations in Sterile Filtration Processes" by Holly Haughney in *Sterile Pharmaceutical Products: Process Engineering Applications*, edited by Kenneth E. Avis. Copyright © 1995 by Interpharm Press, Inc. All rights reserved. Adapted by permission.

of being retained by the prefilter). The prefilter is unable to adequately protect the final filter. Representative curves are provided for the prefilter, the final filter, and for the combined system. The combined system curve is a sum of the differential pressures for the final filter and the prefilter. The combined system reaches the limiting differential pressure (plugging) before the complete batch is filtered. Thus, a finer prefilter or additional final filter surface area would be required for this batch volume to be filtered at differential pressure less than the limiting differential pressure.

Figure 18 illustrates what occurs in a system in which the prefilter area is too small, but performs its protective function. In this case, the final filter did not reach the limiting ΔP, but the prefilter did. Additional prefilter area would allow the filtration to be optimized. Figure 19 contains representative curves for an optimally staged filtration system. In the optimized system, the combined system is able to completely process the batch at a differential pressure lower that the limiting (plugging) differential pressure. The differential pressures across the final filter are slightly higher that the differential pressure of the prefilter. This is the desired situation for a system that requires prefiltration.

Depending on the knowledge of the filtration process for a particular fluid, it may be desirable to perform a side-stream or a pilot-scale test under process conditions before installing the system at full scale. Further optimization of the system may be possible based on the results of the side-stream test. A side-stream test is recommended when the scaled-down test was limited in its ability to match the process conditions for any stage of the filter system.

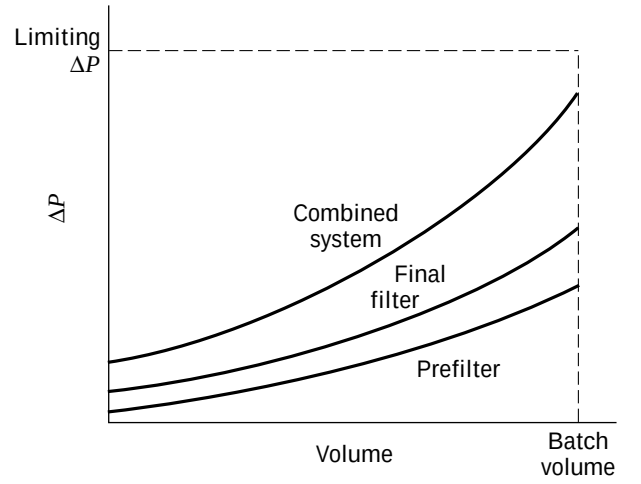


Figure 19. Example of a filtration with optimal prefiltration. Adapted from "Engineering Considerations in Sterile Filtration Processes" by Holly Haughney in *Sterile Pharmaceutical Products: Process Engineering Applications*, edited by Kenneth E. Avis. Copyright © 1995 by Interpharm Press, Inc. All rights reserved. Adapted by permission.

System Design

Once the filters have been specified, the system components must be considered. A filtration system will typically contain the following filters, filter housings, gaskets or O-rings, pressure gauges, thermocouples, pipe or tubing, connections for the piping, and a pump (or gas pressure source). Some pharmaceutical systems may require an inert gas for product stability as well as pressure. The system components must be compatible with the process fluid, the full-scale process fluid temperature, the process pressure, and the sterilization method.

Most sterile filtration applications require a sanitary design. In a sanitary design the components (e.g., sanitary valves) and materials of construction are selected for their

ability to prevent the buildup of contaminants in the system. For a sanitary process, 316L stainless steel is typically used for portions of the system that will have product contact, provided it is compatible with the process fluid. Internal surfaces should be finished to an appropriate microinch R_s specification (e.g., a minimum of 20 to 25 μin.); external surfaces should be mechanically polished to a high quality sanitary finish. Both internal and external surfaces should be electropolished for a smooth adhesion and corrosion resistant surface.

System welds should be constructed in such a way that the weld porosity is minimized and the joints are of high quality and are clean. Any internal welds that have product contact should be ground smooth and flush to reduce the potential for the buildup of material from the process stream. Proper weld procedures need to be followed.

Only sanitary fittings should be used in the system to ensure sterilizibility of the seals and cleanliness. Any ports in the system, such as those required for vents, drains, pressure gauges, and thermocouples, should have a sanitary stem or sanitary valve design. Absolutely no threaded connections should be used in the portions of a sanitary system where there is a potential for contact with the process fluid. Threaded connections can lead to system contamination because of a potential buildup of contaminants.

There are a number of considerations for the filter housing. Housings are essentially pressure vessels, and as such, may be subject to the appropriate ASME boiler and pressure code. A housing should be rated for the appropriate pressure and temperature with a safety margin for the process. A full vacuum rating may also need to be considered, especially for housings used in vent service or lyophilizers. The design of the housing as well as the physical placement of the housing in the production facility, must facilitate filter change-out. The size of the housing must meet the requirements for the flow rate and differential pressure process specifications. Inlet and outlet ports should not lead to excessive pressure losses.

The system and housing design should not contain any portions through which there is no flow (dead legs). All product contact areas need to be accessible for cleaning purposes. Drains should be designed and positioned to minimize holdup, or retention, of the process fluid in the system.

The filter assembly should be optimized for use in a sanitary system, especially where aseptic processing is required. For example, the filter should have an O-ring sealing mechanism in the filter housing instead of a gasket sealing mechanism. For some small batch applications, disposable filter assemblies (in effect filter plus plastic housing) can be used. These assemblies should have proper sanitary fittings.

For most sterile filtration systems a appropriate pump (e.g., a positive displacement pump or a centrifugal pump) is used to provide pumping capacity. The components of the pump head should be of a sanitary design and compatible with the process fluid.

Pressure gauges should be installed upstream and downstream of each filter assembly in the system so that the differential pressure across the filter can be monitored. The ability to monitor ΔP can be used to indicate filter life and that filter change-out is required. For some applications it is desirable to have temperature probes or thermocouples installed in the system. During an in situ steam sterilization of a system, the ability to monitor the temperature is needed to ensure that the temperature conditions for steam sterilization are met. For a process that requires a temperature other than ambient, the temperature of the fluid should be monitored. Temperature measurement should be performed immediately downstream of the filter installation to ensure that the filter installation is at the correct temperature.

Throughout the system, sample ports may be required to allow for sample collection during a process. It is important to ensure that the collection of a sample does not cause contamination in the system. This can be attained by using sample ports that are fitted with septa. For systems that are not hard piped (plumbed with permanent connections), the ability to make sterile (or aseptic) connections must be available.

Sterilization

In sterile filtration processes, the downstream side of the filter must be sterilized and must remain sterile during the process. The sterilization process must be properly validated to ensure that the sterile condition is met for a given system. Filters can be sterilized by a number of techniques, including in situ steam sterilization, autoclave sterilization, irradiation, and chemical treatment (hydrogen peroxide vapors or ethylene oxide). For any method, it is important to ensure that the sterilization is adequate for the system and that the technique does not damage the filter. The Parenteral Drug Association (PDA) published a technical monograph on the validation of steam sterilization cycles (28). The majority of sterile processes employ in situ steam sterilization or autoclave sterilization and will be the subject of this discussion.

Methods that involve steam are validated through the use of thermocouples or biological indicators to ensure that the system has been properly sterilized. The maximum steam temperature for most sterilizing grade filter cartridges is 140 °C; the maximum conditions for temperature and pressure should be obtained from the filter manufacturer. Typically, the steam pressure is held at a minimum of 15 psi at 121 °C for a minimum of 30 min. Although steam temperatures up to 140 °C can be used, the service life of the filter may be shortened, especially where multiple uses require repeat sterilization.

A decision must be made on the appropriate sterilization technique for the process. Larger systems tend to be in situ steam sterilized, while small volume systems tend to use autoclave sterilization. A system that utilizes in situ steam sterilization can be automated, which is becoming a common practice for a new system.

Integrity Testing Design and Operation Considerations

The issues that must be considered for integrity testing of a filtration assembly include the technique (forward flow, pressure hold, bubble point, or water intrusion for hydrophobic filters), the wetting fluid (water or product wet), and the test time (after installation, after sterilization, after

filtration) and the test location (in situ or off line). Once a plan for integrity testing is established, the system must be designed so that the integrity test regimen can be properly carried out.

Field experience with the integrity testing of sterilizing grade filters shows that various combinations of integrity test procedures are in use and that different testing schedules, both pre- and postuse, are followed. In some sterile processes, redundant filters are used for convenience in integrity testing or for increasing assurance of sterility. For others, filter usage is based on filter life studies. A common approach is to perform the test or combination of tests that meet regulatory guidelines and that provide the highest degree of accuracy commensurate with the economics and the practicalities of the process.

The highest degree of confidence in the maintenance of filter integrity can be attained by routinely integrity testing the filter system, either by the forward-flow or pressure-hold integrity test, before sterilization, after sterilization, and after filter use.

Pre- and poststerilization integrity tests are recommended, because both tests can provide information about the process conditions. However, if only one preuse integrity test is to be performed, then an integrity test after sterilization will provide more useful information, because improper sterilization procedures can lead to filter damage. Regardless of any preuse testing performed, a postuse test is always recommended. The assurance of bacterial retention during the filtration of critical fluids, such as parenterals, biological liquids, and media for fermentations, is extremely important and, at the very least, filters are integrity tested after use.

It is often desirable for a filter user to perform a product–wet integrity test after the filter has been used to filter the product. Product–wet testing allows the filter user to perform the integrity test under actual process conditions and does not require a flush step to remove the product from the filter membrane. Flushing out the product and testing with water can lead to a false failure if there is incomplete removal of the product that may have a lower surface tension than water and thus reduce the bubble point to below the test pressure required for a water-wet integrity test.

Appropriate integrity test parameters for specific applications should be obtained from the filter manufacturer in writing because the filter manufacturer can provide the appropriate test parameters for a fluid and can ensure the proper relationship to the claims for the specific filter. The filter user should confirm performance of the tests before incorporating the test parameters into their standard operating procedures. If any anomalies are noted during confirmation, then these must be resolved.

Integrity test parameters can be provided with either air or nitrogen (or other inert gas) pressurization. Carbon dioxide cannot be used as a test gas, because the diffusive flow through the filter will be extremely high and the gas is reactive. Pure oxygen is typically not acceptable as a test gas, because it is a strong oxidizer. Process conditions, including temperature and gas used for pressurization, must be considered for the issuance of integrity test values.

BIBLIOGRAPHY

1. H. Haughney, in T. Meltzer and M. Jornitz ed., *Filtration in the Pharmaceutical Industry*, Dekker, New York, in press.
2. H. Haughney, in K.E. Avis ed., *Sterile Pharmaceutical Products: Process Engineering Applications*, Interpharm Press, Buffalo Grove, Ill., 1995, pp. 305–355.
3. J.M. Martin, A.M. Trotter, P. Schubert, and H. Katz, in B.K. Lydersen, N.A. D'Elia, K.L. Nelson ed., *Bioprocess Engineering: Systems, Equipment, Facilities*, Wiley Interscience, New York, 1994, pp. 317–370.
4. N.B. Bruckshaw, *Filtr. Sep.* **296** (May/June 1973).
5. D.B. Pall and E.A. Kirnbauer, *52nd Colloidal and Surface Science Symposium*, University of Tennessee, Knoxville, Tenn., 1978.
6. H. van Doorne, *J. of Parenter. Sci. Technol.* 192–198 (Sept./Oct. 1993).
7. K. Wickert, *Pharm. Eng.* **11** (1991).
8. S. Brown and A. Fuller, *J. Parenter. Sci. Technol.* 285–288 (Nov./Dec. 1993).
9. J. Martin and R. Manteuffel, *BioPharm* (Nov./Dec. 1988).
10. R. Datar, J. Martin, and R. Manteuffel, *J. Parenter. Sci. Technol.* **46**, 35–42 (1988).
11. United States Pharmacopoeia (USP) XXIII, *Biological Reactivity Tests, In Vivo ⟨88⟩*, The United States Pharmacopoeial Convention, Rockville, Md., 1994, pp. 1699–1703.
12. Code of Federal Regulations, *Indirect Food Additives Subpart B: Substances for Use as Basic Components of Single and Repeat Use Food Contact Surfaces*, Title 21, Part 177, U.S. Government Printing Office, Washington, D.C., 1994.
13. American Society for Testing and Materials (ASTM) *Standard Test Method for Determining Bacterial Retention of Membrane Filters Utilized for Liquid Filtration*, ASTM Standard F838-83, 1983.
14. United States Pharmacopoeia (USP) XXIII, *Particulate Matter in Injections ⟨788⟩*, The United States Pharmacopoeial Convention, Rockville, Md., 1994, 1813–1819.
15. United States Pharmacopoeia (USP) XXIII, *Purified Water*, The United States Pharmacopoeial Convention, Rockville, Md., 1994, p. 1637.
16. United States Pharmacopoeia (USP) XXIII, *Bacterial Endotoxins Test ⟨85⟩*, The United States Pharmacopoeial Convention, Rockville, Md., 1994, p. 1696–1697.
17. T. Uberoi, *Chem. Eng. Prog.* 75–80 (March 1992).
18. Food and Drug Administration, *Guideline on Sterile Drug Products Produced by Aseptic Processing*, Center for Drugs and Biologics, Rockville, Md., June 1987.
19. J. Meeker, E. Hickey, J. Martin, and G. Howard Jr., *BioPharm*, March 1992, Tech Note.
20. J.D. Brantley and J.M. Martin, *Integrity Testing Sterilizing Grade Filters*, Pall Corporation Publication PBB-STR-28, 1997.
21. F.M.H. Villars and G.B. Benedek, *Physics with Illustrative Examples From Medicine and Biology*, 2nd ed., Addison-Wesley, Reading, Mass., 1974.
22. D.B. Pall, *Annual Meeting of Parenteral Drug Association*, New York, November 1973.
23. T. Takegoshi, *Proc. of the Int. Congress: Advanced Technologies for Manufacturing of Aseptic and Terminally Sterilized Pharmaceuticals and Biopharmaceuticals*, Basel, Switzerland, February 17–19, 1993.

24. M.P. Dosmar, K. Wolber, H. Bracht, H. Troger, and P. Waiber, *J. Parenter. Sci. Technol.* (July/August 1992).
25. M.P. Dosmar, K. Wolber, H. Bracht, H. Troger, and P. Waiber, *J. Parenter. Sci. Technol.* (July/August 1992), *Filtr. Sep.* June 1993.
26. R. Jaenchen and J. Malsey, *Analysis of Water Penetration Through Hydrophobic Filters as a Reliable Method for Integrity Testing.*
27. *The Pall Water Intrusion Test for Integrity Testing Sterile Gas Filters* Pall Corporation Publication PBB-STR-29, 1997.
28. Parenteral Drug Association (PDA), Inc. *Validation of Steam Sterilization Cycles*, Technical Monograph No. 1, 1978.
29. Pall corporation, *Filtration Products for the Food and Beverage Industry*, Pub. FBB-100, 1990.
30. Pall Corporation, Pub. E-33, out of print.
31. Pall Corporation, *Selection Guide CLL Style Filter Housings*, Housing Data Sheet H26a, 1987.
32. Pall Corporation, *Selection Guide V-Series Filter Housings*, Housing Data Sheet H28, 1990.
33. Pall Corporation, *Water Intrusion Test for Integrity Testing Hydrophobic Sterile Gas Filters*, Pub. PBB-STR-30, 1996.

FLOW CYTOMETRY

SUSANN MÜLLER
University of Leipzig
Leipzig, Germany
THOMAS BLEY
Dresden University of Technology
Dresden, Germany

KEY WORDS

Cell cycle
Cell sorting
Cytometry
Fluorescence
Immunolabeling
Microbial performances
Population dynamics
Process monitoring
rRNA probes
Viability

OUTLINE

INTRODUCTION

Cytometry is a technique for measuring physical or chemical characteristics of single cells (or other biological particles). Such measurements are usually performed with the aid of a microscope, but there are severe limitations both in the number of cells that can be investigated microscopically and in the number of features that can be observed simultaneously.

In flow cytometry (which avoids these limitations), the cells pass the measuring volume in a fluid stream, preferably in a single file, and in a short time. The first flow cytometers, developed in the 1940s, were based on the Coulter principle, that is, the detection of changes in the electrical conductivity or impedance of a small saline-filled orifice as cells pass through.

The history involved leads most people to think of flow cytometers as instruments that measure optical signals from cells. These optical signals result from light being scattered by the cells (differentially, according to cell type and status) and the fluorescent light emitted by different cell compartments when excited by light of a certain spectral range. The cellular parameters detected can be characterized as intrinsic or extrinsic, depending on whether their measurement requires use of reagents or probes.

A cytometer consists of five main components:

1. *Illumination optics*. Usually consisting of an arc lamp (high-pressure mercury) or a laser (argon ion) combined with crossed cylindrical lenses to focus the light.
2. *Flow cell*. A device for focusing the cells in a linear file by hydrodynamics. The core stream with the cells travels in a laminar flow to the illumination point with the aid of a sheath stream of clean water. The main types are stream-in-air systems and variations thereof passing through a (flat-sided) quartz cuvette.
3. *Collection optics*. Usually consisting of a combination of optical filters for spectral separation, generally set perpendicular to the illuminating beam and in four spectral ranges (green, 510–540 nm; yellow, 560–580 nm; orange, 605–635 nm; and red, 650 nm and above).
4. *Detector electronics*. These consist of photodiodes or photomultiplier tubes (PMTs). The particles of inter-

est in biotechnology, which are frequently very small, (e.g., bacteria), emit only a small number of photons, typically about 1,000 in passage through the detection chamber. Therefore, an appropriate combination of high-power illumination optics and sensitive detectors is necessary for the application of flow cytometry in biotechnology.

5. *Data analysis*. The first flow cytometers displayed and stored the data in multichannel pulse height analyzers. Now the data are stored as standard list mode files in PCs, ready for statistical analysis and sophisticated visualization.

The result of a flow cytometric investigation is a frequency distribution, that is, a histogram. In the channels of this histogram, cells are counted according to the light intensity detected, which corresponds to the relative size or quantity of a specific cellular compartment or component, such as the DNA content. Measuring rates up to 10,000 cells s^{-1} are possible, although about 100 cells s^{-1} is more usual.

Flow cytometry is not recommended for measuring the absolute content or concentration of cell compartments. However, it is the best method for classifying cell types with different physiological properties and for selecting and quantifying subpopulations in a cell ensemble.

If subpopulations of (vital) cells are to be chemically analyzed or cultivated further, the information provided by flow cytometric measurement can be used for selecting appropriate subpopulations by an electrical or mechanical sorting process. Depending on the type of flow cell and the type of illumination involved, three basic constructions are used:

- Stream-in-air
- Contained flow chamber
- Dark-field illumination

The devices range from large constructions for research purposes to small and inexpensive laptop machines on the way to an online approach. The development of new flow cytometric methods is often accomplished using special modular constructions (“cyto mutt” comp. [1]).

A further development is the slit-scanning procedure, which allows the shape of the signal generated by the cell passing through the illumination beam to be used for gathering additional information. In this way, morphological features and the distribution of cellular compartments can be recorded.

Special demands have led to appropriate modifications. The distribution of aquatic microorganisms can be monitored in the size range of 0.1 to 2000 μm by such a device, for instance (2).

Flow cytometers are best used to answer questions concerning the dynamics of subpopulations in a cell population, for example, cells in a mammalian tissue that have a higher DNA content than that of G1-phase cells. Thus, in medicine, flow cytometry was first applied in oncology, where it was quickly (and frequently) used in cancer diagnosis. These results generated a solid basis for mathematical modeling of cell-cycle regulation in eukaryotic organisms.

Using immunofluorescence and multiparameter analysis, flow cytometry became a successful, and standard, method in laboratory hematology and clinical immunology in the 1980s. Several commercial devices are now available, constructed specifically for medical purposes. An excellent and comprehensive documentation of the method of flow cytometry is given by Shapiro (1).

The use of flow cytometry in biotechnology, and especially in bacteriology was for a long time not as successful as it was in medicine. After early, inspiring investigations by Bailey et al. (3), Slater et al. (4), and Hutter and Eipel (5) further development seems to have been delayed. There are probably three major reasons for this:

- The high cost of the method
- The small size of the cells investigated (an *Escherichia coli* cell has only 10^{-3} to 10^{-4} times the size or DNA content, and consequently the potential fluorescence energy gain, of a mammalian cell)
- The belief of most biotechnologists that analysis of physiological state distributions do not have an important role to play in understanding the dynamics of bioprocesses or for running them safely and economically

The first two handicaps have now been overcome because powerful and inexpensive microcomputers have considerably reduced the cost of the technique. Flow cytometry has also benefited from new methods of data analysis and mining, such as neural network analysis, which allow a comprehensive evaluation of the rich and complex information gained by flow cytometry (6). Furthermore, new staining methods combined with improved illumination, detection, and amplification have made flow cytometric analysis highly reliable, even for bacteria and microplankton as small as 1 μm in diameter.

Nevertheless, the key to more widespread use of flow cytometry in biotechnology is more general acknowledgment of the importance of a segregated model concept. Toward this end, Fredrickson et al. (7) proved that the statistics, the distribution of cell populations, and the dynamics of bioprocesses are inseparably connected. Moreover, Munch et al. (8) demonstrated, with a sophisticated approach to flow cytometry, that knowledge concerning the distribution of states in the cell cycle is essential for deeper understanding of the growth dynamics of yeast and the metabolism of a biomass.

BACTERIAL CELLS

Bacteria exist as dynamic and diverse populations in nature. Even the cellular heterogeneity within a pure bacterial culture is far greater than previously assumed. To understand bacterial ecology in nature and in bioprocesses, new techniques are required for answering questions that could not be solved by traditional methods alone. Now more than ever, flow cytometry has been recognized as a valuable asset for microbiologists and is perceived to

be an important and indispensable tool for investigations on the microenvironmental level.

Bacterial flow cytometry enables visualization of cell states and allows the analyst to follow growth, death, replication, cell division, metabolism, and surface phenomena, greatly enhancing the ability to understand and control cell physiology. A great assortment of fluorescent stains is commercially available; their numbers have rapidly increased recently. It is not possible to give an overview of their application in this article, but Lloyd (9) provides a fine introduction to this topic. Summaries are given in Refs. 10 and 11, too.

Generally, however, it should be realized that these techniques need to be handled more carefully for use with bacteria than with eukaryotes to avoid unclear results from artifacts arising either from sampling or from analytical errors.

Growth and the Cell Cycle

All survival strategies should be programmed, above all, to safeguard the genome for the future. Results obtained by flow cytometry allow conclusions to be drawn and tested about how the physiological state of cells is connected with survival strategies under changing environmental conditions. Consequently, the most obvious targets for bacterial staining are the nucleic acids. Steen and Boye (12) and Skarstad et al. (13) presented a model for calculating the relative rates of DNA synthesis per cell from studies based on *E. coli* strains. By means of this model and flow cytometric evaluation of the DNA distribution, the generation time (τ) and the initiation and termination of the replication and postreplication periods (C and D, respectively) can be estimated. Features of replication (e.g., the number of initiation origins) seem to be dependent both on the genetically determined disposition of the bacterial strain analyzed and the growth conditions leading to the initial cell mass. In unperturbed bacterial cultures with high growth rates, several replication forks may be initiated in each cell, if the growth conditions are optimal. Then, after treatment with antibiotics, *E. coli* cells will contain two, four, or eight fully replicated chromosomes. In contrast, bacteria growing slowly, at doubling rates of less than 1 per hour, exhibit an increase in the prereplication period as well as in the C and D periods. This effect can be amplified by further decreasing the growth rate using chemostat cultures under limited growth conditions (12,13).

Besides the considerations just mentioned, it must be stressed that during limited, and even unrestricted, growth, the behavior of individual cells differs. As pointed out by Åkerlund et al. (14), each cell goes through different discontinuous processes in its life (e.g., the cell cycle). They are therefore different both in size and metabolic activity. This is generally true, except in steady-state conditions, to some extent, where only the distribution of the cells and the cell size are time invariant).

Considering these results, it is clear that key events in cell physiology, including initiation of replication and cell division, are tightly bound to the microenvironmental conditions in the surroundings of the bacterial cell. Obtaining information about the replication behavior of biotechnologically useful bacterial strains is, therefore, a crucial prerequisite for controlling industrial bioprocesses.

This knowledge could also be used for investigations into the population dynamics of marine bacteria. Because of fluctuating climatic influences, these organisms must constantly adapt to changing conditions and move between states of growth and starvation. In such cases, major alterations in the level of other cellular components, such as lipids and proteins, may be observed in addition to changes in the nucleic acid contents.

Bioprocessing

Progress in bioprocess engineering depends ultimately on the level of understanding and control of the physiological state of the bacterial population being exploited. Process efficiency is strongly sensitive to changes in the cell state. Characterization of these states via fluorescence monitoring, evaluation of the data obtained, and subsequent regulation of the process regime by controlling the surrounding natural or artificial conditions is one of the most ambitious tasks in the near future for biotechnology. Flow cytometry appears ideally suited for determination of such parameters as cell concentration and size as well as for visualizing intracellular performance in industrial processes promoted by bacteria. Using this technique, the state of the cellular system can be precisely monitored, allowing optimization of process efficiency in production or degradation of the target substances (15).

Although bioprocess control on the cellular level is not being widely practiced as yet, some processes are now being adjusted by flow cytometry. One main limitation seems to be the high content of particulate constituents in most industrial media, because discriminating between microorganisms and noise signals from media is not a trivial undertaking.

Microorganisms respond to an unbalanced supply of nutrients or deviations from optimum physical factors by increasing the synthesis of intermediates. This phenomenon is called overflow metabolism. The metabolites overproduced can be excreted or accumulated intracellularly. Flow cytometric methods have been established for detecting intracellular accumulation of polyhydroxybutyrate (PHB), a typical overflow metabolite in many bacterial strains (16). PHB serves as an energy and carbon reserve and has received attention as a thermoplastic and biologically degradable polymer. Furthermore, researchers are searching for novel bacterial strains that may accumulate novel types of polyhydroxyacid (PHA) possessing better physical and mechanical qualities. Some bacterial yield up to 90% of their dry weight in this material. The inclusion bodies so formed alter both the size of cells and their light-scattering behavior. The effect of different cultivation conditions on PHB formation has been investigated in *Ralstonia eutropha* and recombinant *E. coli* transformed with PHA-synthesis genes. In these cases, cell-sorting technologies can be used to isolate highly efficient strains at a high speed. Flow cytometry can also be used to quantify heterogeneity of PHB production and accumulation, as found in *Methylobacterium rhodesianum* cultures. These investigations are based on the idea that growth and product for-

mation (overflow metabolism) are coupled to specific cell states, which are an expression of maturation of individuals and are connected to stages in the proliferation cycle. It has been found that under growth-limiting conditions, the cells first go though the DNA replication program, thereby safeguarding the genetic information by doubling the chromosome content. Doing this, the organisms maintain the chance of restarting multiplication as a forward strategy of survival if better conditions arise. Cells only lay down PHB as an energy reserve in this kind of situation (17).

Introduction of foreign genes into bacterial cells enables synthesis of desirable products such as proteins, lipids, and a wide variety of other biologically active compounds. Normally, the expression of these genes alters the normal pattern of interaction and synthetic activities in the host cell. It may be observed that highly active cloned gene expression or extremely large plasmid content leads to a markedly slower growth rate in recombinant cells. Generally, however, some of the recombinant cells revert during prolonged cultivation, either (in the case of plasmids), through defective partitioning during cell division, or through changes in plasmid structure. By this process, a subpopulation of nonrecombinant cells develops that has a higher specific growth rate. In chemostat experiments, the recombinant cells are usually replaced by this phenomenon, and the plasmid-free population comes to dominate the dynamics of the reactor (18). It is generally believed that the production of heterologous products is metabolically harmful to the cell. Occasionally, it has been proposed that the decrease in the growth rate seen is more likely to be a result of deleterious and injurious interaction between the expressed product and some component of the cell. Flow cytometry can be used to detect segregational instability of a plasmid-containing expression system as well as to estimate the physiological state of the cells involved.

There have only been a few biotechnological studies investigating the stability of recombinant cells on near-industrial scales. By way of illustration, however, the method has been used to examine recombinant *E. coli* cells forming protein inclusion bodies via wide-angle light-scattering measurements (19).

Fluorescence-activated cell sorting has also been found to be capable of isolating gramicidin S hyperproducing mutant cells of *Bacillus brevis* from wild-type cells (20). Furthermore, flow cytometry can be used to study the progress of gene expression in a wide range of bacterial cell systems. The expression of foreign genes in bacterial cells or the stability of transformed plasmids can generally be determined by the use of bioreporter genes. These systems allow flow cytometric monitoring of the gene expression even when the gene product is difficult to assay. The promoterless reporter gene is placed under the expression control of the promoter of the foreign gene. A typical example is use of the *lacZ* gene from *E. coli*, one of the most widely used bioreporter systems. The *lacZ* gene encodes the enzyme β-galactosidase, which cleaves a fluorogenic product from specially designed substrates. The emission of the resulting fluorescence is proportional to the amount of the enzyme in the cell, corresponding to the expression of the target gene. Consequently, the β-galactosidase-positive cells exhibit a level of fluorescence that is measurable for each cell by flow cytometry (21). However, a primary problem in single-cell β-galactosidase assays is leakage of the intracellular fluorescent marker, although a number of methods have been developed to prevent this movement (22). Other bioreporter systems encode detectable markers such as green fluorescent proteins or surface antigens for monitoring microbial performance.

There have also been attempts to use polymerase chain reaction (PCR) in DNA amplification methods for sensitive detection of specific innate or foreign DNA or RNA sequences inside intact cells (23). Some of the population seem to remain unbroken after such treatment, allowing subsequent flow cytometric assessment. For example, PCR can be used to detect plasmid-encoded gene sequences or poorly expressed mRNA in individual bacterial cells (24).

Vitality

In contrast to biotechnologically processes, bacterial cells in nature live neither in pure cultures nor in surroundings of constant temperature, pH, or nutrient availability. Many of the (microscopically) visible bacterial cells are inactive or dying. Every change in the habitat is potentially a source of stress, which results in a change in growth rates. The sum of the cellular responses to stress is a factor associated with the survival strategy.

Responses to stress take place at two distinct cellular levels: first, activation of existing enzymes and transport processes and, second, gene expression. The biochemical bacterial stress response can be precisely determined using flow cytometry, by analyzing changes in the type of duplication, cellular pH, membrane potential, and the amount and kind of storage products. The physiological links between the environment and the genetic systems with respect to survival of stress situations are often mediated by expression of proteins involved in a range of tolerance mechanisms. It is commonly accepted that specific genetic programs exist for prolonging the survival of nongrowing bacteria exposed to starvation or stress. Bacterial programmed cell death seems to be mostly related to population development, because the lysed cells are usually essential as sources of nutrient compounds, allowing completion of the developmental cycle of the remaining, living population (25).

To date, knowledge of the connections between bacterial processes and bacterial community structure in nature is inadequate and restricted by procedural flaws. Most commonly used methods for the detection and enumeration of bacteria have serious limitations because of their incapacity to allow certain bacteria to grow. The reason may be that the bacteria concerned need symbiotic partners, anaerobic microenvironments, other uncommon conditions, or unknown growth factors.

However, for a variety of reasons, quantifying the bacteria present in the investigated system, independently of growth, is a fundamental requirement for almost all bacteriological studies. Living cells must be distinguished from dormant cells (induced by starvation) as well as from viable (but not culturable) and truly dead bacteria. To char-

acterize the degree of viability of bacteria, flow cytometric detection of basic cell functions such as reproductive activity, metabolic activity, and membrane integrity is an invaluable tool (26,27). Reliable flow cytometric methods for viability assessment that circumvent the need for culture were largely developed by Porter et al. (28). Investigations associated with the state of dormancy in bacteria, in which the cells survived for extended periods of time without growth or multiplication, have been performed by Kell et al. (29). Cationic lipophilic viability dyes should be handled with care in such studies, because they are very often extruded in an energy-dependent manner. There are several indications that a transport system is involved. Excellent correlations can regularly be observed between viability of the cell and the ability to translocate the dye to the cell's surroundings.

A wide variety of bacteria are able to metabolize xenobiotica as sources of carbon and energy for growth and to decontaminate polluted ecosystems in this way. However, the effects these toxic substances have on the artificially introduced bacterial species should be considered in such biotechnologically forced processes. Above certain, critical concentrations the chemicals become toxic, the bacteria are poisoned, the bioremediation process slows down and finally stops. It is well known that most of such substances are membrane active and burden the cell energetically. Consequently, cellular changes in the energetic state and the form of the membrane potential must be analyzed with the solute transport, ATP synthesis, and pH homeostasis. To control and optimize such processes, information is necessary about death characteristics, which depend on the physiological state of the cells (30). Flow cytometric assessment of the membrane potential allows a rapid and differentiated analysis of the ecotoxic potential of pollutants toward bacterial cells and detection of the rapidly changing physiological behavior of the cell in growth-phase-dependent variations of the membrane potential (Fig. 1). Very little information exists regarding the members of bacterial communities with respect to morphology and taxonomy that are responsible for the decontamination process. There is also little information about differences in the physiological state of specific strains under certain conditions. For this reason, methods must be developed that relate specific cellular parameters to taxonomic classes (Fig. 2). The first such investigations of this type were done by Herrmann et al. (31).

In medical science, situations also arise in which living and dead cells must be distinguished. In this way, flow cytometry can be used in the assessment of biocidal drugs, offering rapid assays for prokaryote drug susceptibility. The aim is to use the technique in the clinic for rapid susceptibility testing as an aid for choice of therapy. Common methods for investigating resistance to antibiotics and disinfectants and their effects on bacterial strains are based on estimating the minimum growth inhibitory concentration (MIC) of the compound or the decrease in colony-forming units after exposure. These conventional techniques need at least 24 to 48 h to get information about the action of antibiotics.

Flow cytometry is much faster and, furthermore, gives considerable advantages in drug research for investigating the mode of action of novel antibiotics. The susceptibility of bacterial cells to antimicrobial agents may be so profoundly influenced by age, growth, and metabolism to be heterogeneous within a population. The permeability of the bacterial membrane changes during the life cycle, for instance. Recently, it has become apparent that the permeation of lipophilic substances through the envelope of Gram-negative bacteria has great potential in the development of new chemotherapeutic agents. Generally, using flow cytometry, key drug-induced changes in light-scattering behavior, uptake of vitality stains, DNA ploidy, and protein patterns can be estimated (32,33).

Mixed Populations

Bacterial ecology requires the application of new techniques to help address problems that cannot be solved by traditional methods alone. Efforts to demonstrate the value of flow cytometry in identifying different species within complex populations have relied on immunofluorescence-based methods such as the application of antibodies (34), lectins (35), or rRNA-targeted oligonucleotide probes (36). The methods chosen must be applicable to as wide a range as possible of different organisms. These techniques have also found uses in areas such as quality control of water and foodstuffs as well as in soil and water ecology.

Today, fluorescently labeled oligonucleotide probes are most commonly used for monitoring specific strains. The probes are complementary to group-specific regions of the highly conserved multicopy 16S or 23S ribosomal RNA molecules and are bound using in situ hybridization. In this manner, the bacterial strain of interest can be detected without separating it from the surrounding microflora and contaminating debris. Furthermore, comparative sequencing and the design of nucleic acid hybridization probes have served to establish an alterated phylogenetic framework.

Differentiation between artificial cultures is possible if the members have contrasting proportions or absolute amounts of guanine–cytosine or adenine–thymine. Estimations of guanine–cytosine content vary from nearly 20% up to 80% (37). Using dyes with different binding preferences to DNA and double fluorescence excitation, composition analysis seems to be workable.

Another way of differentiating mixed populations is the genetic insertion of a fluorescent system to split populations with special physiological capacities.

Food

The use of fluorescent labels is accepted as a successful and sensitive approach for detecting bacteria in food. Lactic acid bacteria are fermentative Gram-positive organisms that are widely used as starter cultures for manufacturing a great assortment of foods and drinks, such as cheese, yogurt, and sausages. The quality of a starter culture is highly important. It must be taken into account that industrial fermentations may involve stressful conditions that can affect bacterial gene expression, arrest cell multiplication, interfere with metabolic activity and survival potential, and lead, possibly, to cell death. Studying responses of lactic bacteria to stresses such as freezing and

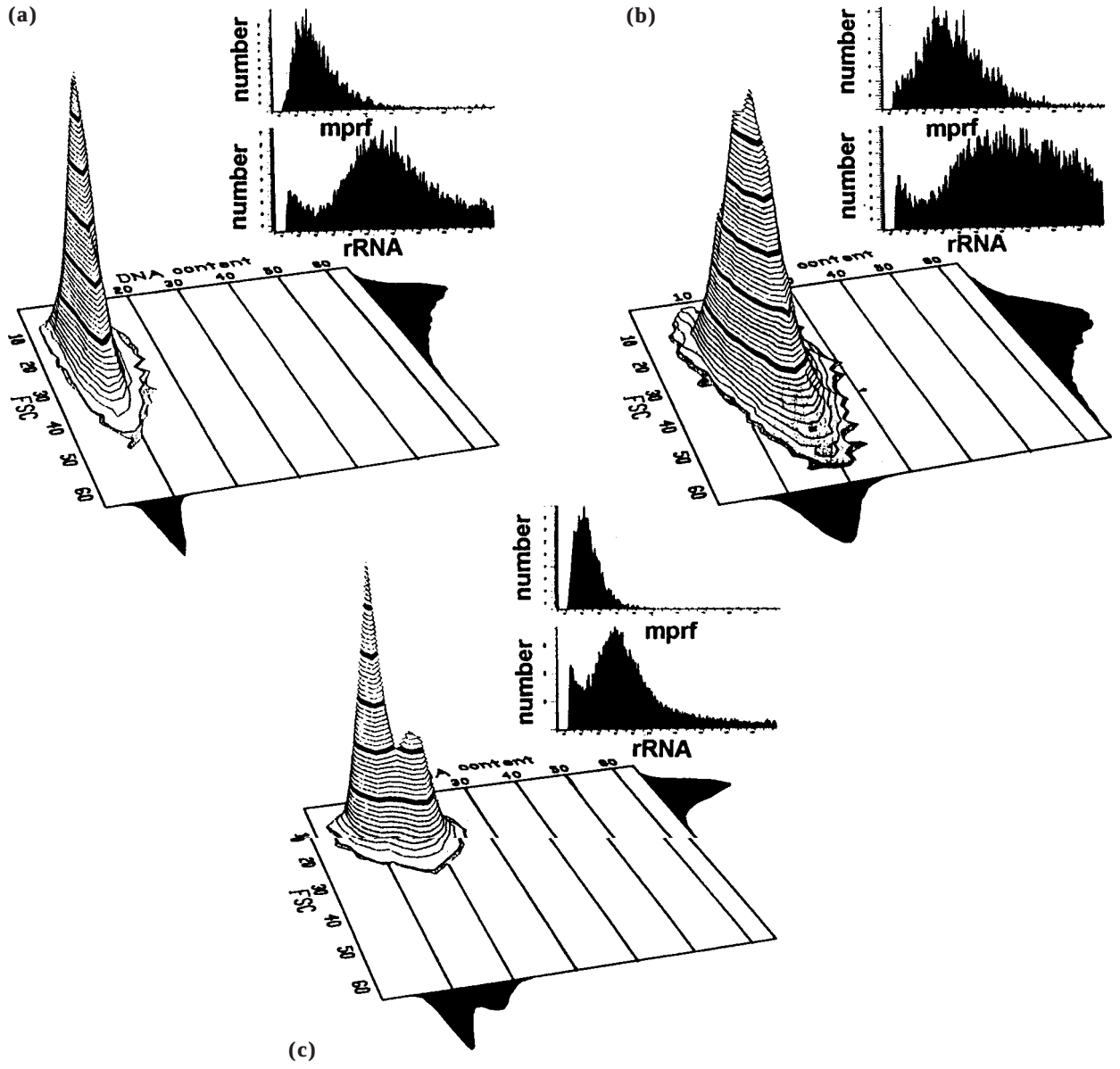


Figure 1. Analysis of the cellular DNA content, membrane-potential-related fluorescence (MPRF), and rRNA content of *Ralstonia eutropha* JMP 134 in relation to the growth rate during batch cultivation on phenol. Samples were harvested at the lag-phase (**a**), early log-phase (**b**), and stationary phase (**c**). The cells were stained with the carbocyanine dye di-OC_6 (3) for MPF, with oligonucleotide probes for analyzing the rRNA, and with 4′,6-diamidinophenylindole (DAPI) for the DNA. The proportions of cells having two chromosomes is estimated to be 3.88% (**a**), 87.80% (**b**), and 38.96% (**c**).

drying, with the aim of monitoring, controlling, and promoting appropriate responses in starter cultures is a typical requirement of flow cytometric investigations in food industries. The method enables determination of cell counts, cell viability, and the effects of stress on the bacterial physiological state. The method can suffer from interference by particulate materials, but multiparameter counting based both on light scattering and fluorescence measurements generally enables separation of bacteria from background material (38).

There are many situations where the total bacterial concentration is the only measurement that is important. Bacteria are a principal cause of food spoilage. Contamination by harmful organisms can cause severe human food-borne diseases, either caused by the bacteria themselves or the toxins released by the bacteria. The time required for conventional tests (plate-count techniques) can lead to substantial delays (of 24 to 72 h), which is a serious disadvantage. Direct epifluorescent filter enumeration is another commonly used technique that allows microscopic

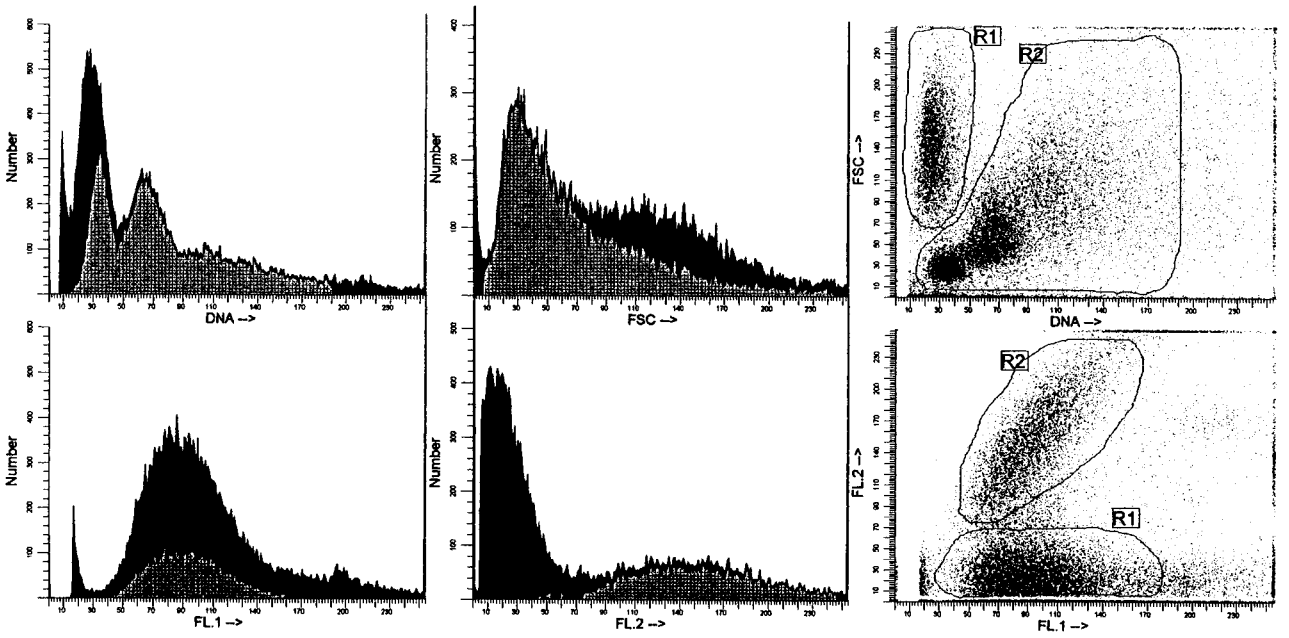


Figure 2. Analysis of the DNA content using DAPI and cell size differentiation via forward light scatter measurements of a mixed population of *Ralstonia eutropha* JMP 134 (R1) and *Acinetobacter calcoaceticus* 69-V (R2) (upper three histograms). Discrimination of the two strains was achieved after in situ hybridization with differently labeled groups and strain-specific oligonucleotide probes using the same sample (lower three histograms). Both strains were harvested during the middle log-phase in mixed batch culture.

counting of bacteria retained on a filter. Over the past 10 to 15 years, this method has been used extensively for studies on the survival of bacteria in food and for estimating biomass in drinks (39).

However, results suggest that flow cytometry also has significant potential for the detection of pathogenic microorganisms in the food industry. Using this method, pathogenic bacterial cells can be detected by applying fluorescently labeled monoclonal antibodies, for rapidly detecting *Salmonella* or *Listeria*, for example. Accurate detection has been demonstrated down to levels of below 10^4 cells/mL or even 1 cell/mL after preenrichment (40,41).

Further Applications

The function of bacterial strains in certain human or animal diseases is increasingly being assessed using flow cytometry. For instance, specific cellular characteristics of bacteria involved in skin diseases are being studied by these techniques (42). Furthermore, flow cytometry offers a rapid method for characterizing anaerobic bacteria present in human fecal suspensions using specific physical and biochemical features (43). Attempts have been made to detect bacterial toxins via fluorescently labeled antibodies bound to polystyrene microbeads in stool specimens (44). Moreover, flow cytometric investigations have been used to identify antisera produced against bacteria capable of specifically binding to surfaces of the target bacteria (45). The resurgence of tuberculosis has caused considerable effort to be focused on developing rapid methods for inhibiting growth of the causal organism by a variety of agents (46). Also, in animal disease investigations, there is growing interest in the capacity of fish-pathogenic bacteria to survive long-term starvation (in a nonculturable mode) in seawater. Knowledge of how long pathogenic bacteria preserve their infective capacity after being released to the environment is clearly of great practical importance (47).

In conjunction with other biochemical and genetical methods, flow cytometry has been used as a rapid method for testing the effects of shear stress, extreme temperatures, action of chemical compounds, sonication, and electroporation on bacterial vitality (48). Bioluminescent *E. coli* harboring lux genes have been used for detecting environmental pollutants, for example. Furthermore, fluorescence-activated cell sorting is often the best available method for obtaining purified samples from natural environments in order to apply molecular techniques or for extracting or enriching samples for detecting low numbers of specific bacteria, such as pathogenic or genetically engineered microorganisms.

The function of bacterial cell surfaces has been investigated in various ways, such as through the interaction of bacterial cells with electrically charged surfaces, macrophages, or other cells through electrostatic interaction. The investigated phenomena may be mediated by bacterial polysaccharides, which are habitually associated with the outer surface of the bacterium. These polysaccharides form a unique class of polyelectrolytes, possessing antigenic properties, for which a lot of fluorescent monoclonal or polyclonal antibodies are available (49).

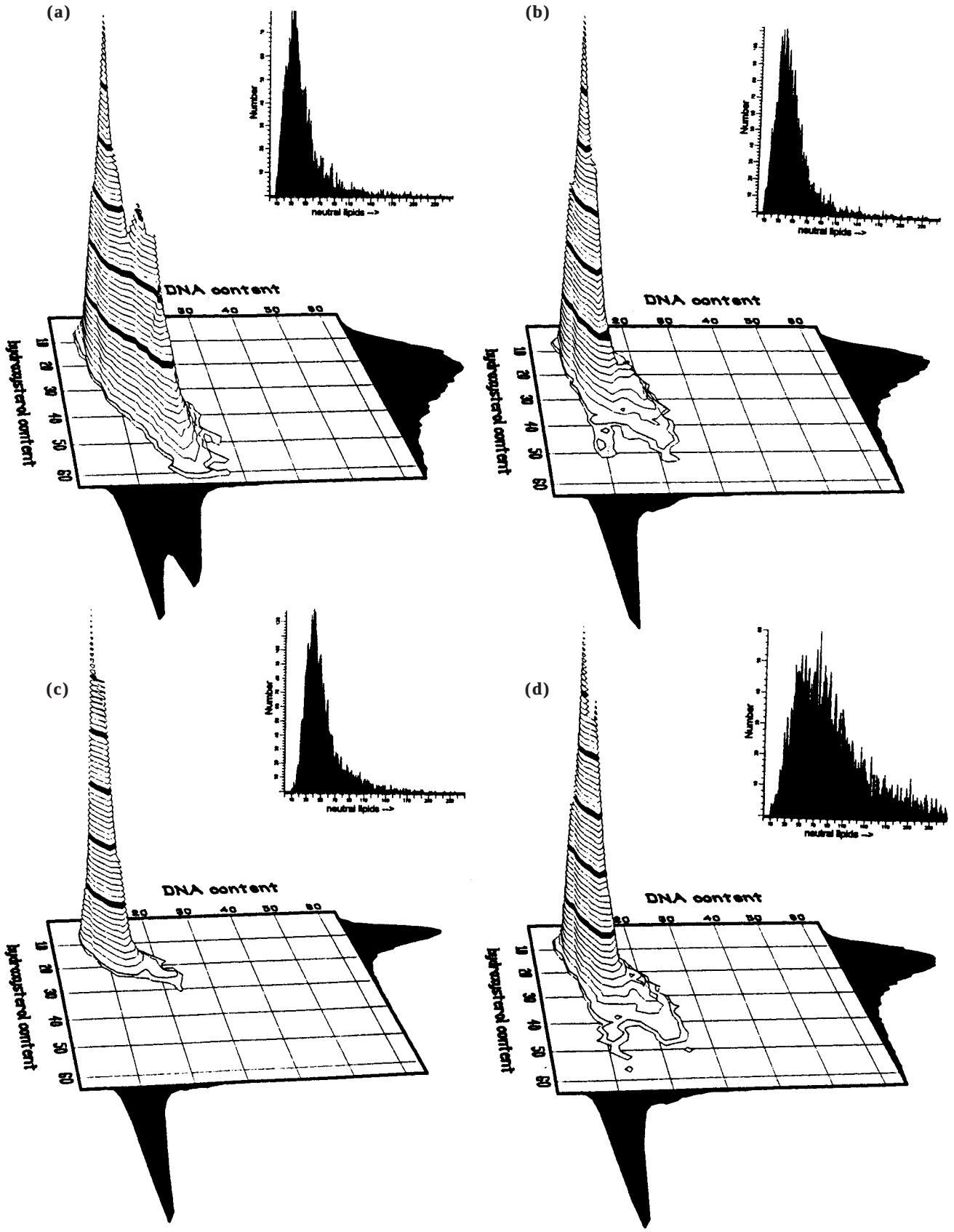


Figure 3. Flow cytometric analysis via double fluorescent measurement of the DNA and 3-*β*-hydroxysterol content of *Saccharomyces carlsbergensis* during fermentation. The first sample (**a**) was harvested 20 after inoculation of the brewing reactor and is characterized by a high rate of proliferation. The second sample (**b**) was harvested the following day, the third (**c**) after 4 days, and the fourth (**d**) after 6 days. The fermentation process was finished on the 7th day by carbon limitation. As shown in the last three samples, most of the cells (especially in the case of **c**) were in the G1 phase of the cell cycle and possessed a limited amount of 3-*β*-hydroxysterols. The insets show the distribution of the neutral lipid content. A large increase in Nile red fluorescence could be observed after 6 days (**d**).

YEASTS

Of all the biotechnologically important microorganisms, the yeast *Saccharomyces cerevisiae* has been the most extensively investigated using flow cytometry to date. As baker's yeast and producer (along with certain related species) of beer and wine, this yeast is of great economic significance. Nevertheless, the chief reasons for the exceptional interest for flow cytometry in this yeast is that it is a widely accepted model organism for eukaryotic cells, because of its ease of cultivation and its well-developed genetics. Investigations of other yeast genera have been very rare and have not been done systematically.

Growth and the Cell Cycle

Basic research into yeast growth and cell cycling by flow cytometry began in the late 1970s with one-dimensional analysis of cell protein content (by FITC-staining) and cell DNA content (using mithramycin staining) by Gilbert et al. (50), Slater (4), and Hutter (51). Both common batch-growth phenomena and the typical diauxic growth pattern of *S. cerevisiae* were investigated. These authors elucidated basic rules concerning the timing of DNA synthesis in cells growing at different rates. The first experimental evidence supporting the hypothesis that the eukaryotic G0/G1 phase is the most variable phase of the cell cycle, and that the duration of the S phase is nearly constant and is not greatly influenced by the cell's environmental conditions, was generated by these flow cytometrically monitored experiments.

In the 1980s, Alberghina et al. (52) and others introduced two-dimensional approaches, involving a light-scattering signal and a fluorescence signal. In particular, methods for the determination of fluorescently stained single-cell protein were added by Scheper et al. (53). Furthermore, sharper detection of the DNA content and, therefore, of the cell cycle state distribution was also made possible by the introduction of more specific staining methods such as propidium iodide (excitation at 488 nm) and DAPI (excitation in the near UV at 351 nm). The theory that DNA replication of *S. cerevisiae* is initiated when a cell has reached a certain critical size was proved experimentally by these investigations.

Asymmetric cell division, subcellular characteristics, and numerous features of the intracellular distribution of cellular components were also shown to be detectable by the application of slit-scanning data acquisition (54). A double flow cytometric tag allows the progress of the cell cycle to be tracked in the different cohorts of mother and newborn daughter cells during balanced exponential growth (55).

Bioprocessing

Substantial progress also came from bioprocess type investigations. Because *S. cerevisiae* populations tend toward autosynchronization under certain process conditions in a continuously stirred tank reactor (CSTR), the changing distribution of states in the cell cycle can be determined by flow cytometry and correlated with the physiological performance of the population. It was demonstrated in this way that the type and rate of substrate utilization and cell compartment synthesis are dependent on the stage of the cell proliferation cycle (53).

In a key paper from 1992, Munch et al. (8) proved on the basis of a comprehensive use of flow cytometric methods that understanding cell cycle behavior is essential for the characterization of growth dynamics in bioprocesses. Using Calcofluor White M2R, the chromosomal and the mitochondrial DNA content in the yeast cells could be separately monitored, and the aging process in the mother cells of the budding yeast could be recorded.

Changing respiration coefficients, combined with changes in substrate fluxes and cell processes, such as the storage and mobilization of carbohydrates and the excretion of ethanol, can now be correlated with different stages in the cell cycle.

A further important cell component, with cell-cycle sparking and bulk membrane functions (3-β-hydroxysterol), became recordable by staining with FITC covalently coupled to the amino group containing polyene-macrolide antibiotic, Nystatin A1 (56,57).

Brewing

Demands of the brewing industry for ever-faster production, without prejudicing the quality of the product, have made flow cytometry more and more significant in this classical branch of biotechnology. The use of cylindrical cone-shaped brewing reactors with volumes of at least 160 m^2 requires increasingly abundant information concerning yeast physiology. Close monitoring of changes in the distribution of DNA, neutral lipid, and 3-β-hydroxysterol contents in *Saccharomyces* cells during propagation, fermentation, and storage enables time-saving process control (58).

DNA distribution monitoring has shown that only cells with a single chromosome content produce ethanol and CO_2 efficiently. On the basis of flow cytometrical investigations, the optimum point for controlling the wort and oxygen supply can be accurately determined. Double-staining techniques prove that cells with a high content of neutral lipids and 3-β-hydroxysterols have a high survival capacity. Only cells with such features are able to generate vigorous subsequent fermentation. However, additional measurements for determining the proportions of living cells or detecting bacterial infections are only sporadically performed in the brewing industry by flow cytometry.

Further Applications

Wittrup and Bailey (59) use flow cytometry to monitor the β-galactosidase activity in *S. cerevisiae* by measuring the fluorescence decay of the fluorogenic substrate resorufin-β-D-galactopyranoside. With this method, the expression of foreign protein, coupled to the β-galactosidase gene, could be measured, and a segregated model mapping β-galactosidase activity as a function of plasmid copy number was presented (60).

Another method for detecting this protein was a quantitative immunofluorescence procedure developed by Eitzman et al. (61). These authors demonstrated that the ac-

tivity of the enzyme is correlated with the stage in the cell cycle that individual cells have reached.

Peroxisomes are inducible organelles that may occupy large proportions of the cell volume when yeasts are growing on methanol-containing media. *Hansenula polymorpha*, a methylotrophic yeast, was investigated via side-scattered light (which was found to depend on cell volume, morphology, and structure) and FITC retention (which was largely dependent on the vacuole.) A wild strain and a strain partially repressed by glucose were compared (62). Furthermore, it has been shown that the viability of the pathogenic *Candida* yeasts can be determined by measuring the membrane potential (63) and that of viable blastospores using tetrabromofluorescein (64).

CELLS OF OTHER ORGANISMS

Aquatic Organisms

In aquatic science, flow cytometry can make an important contribution to ecological and physiological studies of natural populations of microbial plankton. It has been increasingly used to analyze natural communities of aquatic microorganisms because of its sensitivity and quantification capacity. To understand the processes that influence the behavior of any planktonic species, information is required on both the variability between individuals and groups and the major environmental factors.

Flow cytometry was introduced in freshwater and marine sciences at the beginning of the 1980s. However, natural samples of aquatic particles are generally very heterogeneous; there are often immense differences in concentrations and cell sizes (from 0.2 to more than 20,000 μm in diameter) as well as in cell types. Additionally, they may be contaminated by inorganic particles, and artifacts may be generated that can make the data difficult to interpret.

However, the application of flow cytometry to planktology is no longer hampered by shortcomings of commercial instruments. High-resolution and high-speed sorters are available for rapid and precise estimation of cell numbers, differentiation of cell types, and other pertinent problems. Neural network analysis provides an innovative approach for examining taxonomic properties and the state of the marine food web, including its involvement in the surrounding environment.

The prokaryotic fraction of natural marine communities is composed of both heterotrophic and autotrophic organisms. In spite of their very small sizes (0.2 to 2 μm), even the oceanic picoplanktonic cells are now easy to identify. Major groups of photosynthetic prokaryotes can be discriminated by flow cytometry because of their different and, compared to other algae, unusual pigment compositions and their forward light-scatter signal. The heterotrophic organisms are mostly bacteria. Their lack of pigments makes them less easy than autotrophs to analyze by fluorescence methods. Nevertheless, flow cytometry has been used to evaluate kinetic constants of nutrient uptake in relation to cell quantity and, additionally, to characterize individual parameters of the marine bacteria (65). Frequently, total counts of these bacteria include a large fraction of non-nucleoid-containing bacteria (ghosts), which may be cell residues of virus-lysed bacteria or remains of protozoan grazing. Nevertheless, simultaneously measured DNA content allows further discrimination between autotrophic and heterotrophic components of the picoplankton. In some cases, it has been demonstrated that the cell cycle of autotrophic planktonic prokaryotes progresses in phase with the daily light cycle (66), in contrast to cyanobacterial populations, which apparently do not adjust their circadian timekeeping.

Eukaryotic algae exhibit a strong natural autofluorescence emanating from chlorophyll or other pigments such as phycobiliproteins. The simultaneous measurement of DNA and chlorophyll fluorescence, together with cell size, is necessary to distinguish subpopulations of phytoplankton. Flow cytometry allows the quantification of changes in cellular pigmentation (e.g., chlorophyll and phycoerythrin concentrations) of divergent phytoplankton subpopulations. Such alterations arise in response to nonsteady light or nutrient regimes, such as may be encountered along a depth profile.

Information on the physiological status or response induced by different, artificial treatments can be easily assessed by flow cytometry. Biomass production in phytoplankton and its dependence on environmental conditions can be detected via measurement of the activities of various enzymes such as esterases (FDA) or nitrate reductase (by immunolabeling). The recent development of taxon-specific probes should increase the applicability of flow cytometry for rapid identification of cultured phytopico- and nanoplanktonic strains, especially those that lack taxonomically useful morphological features. Moreover, rRNA probes may provide new information on species that are not amenable to culture in the laboratory (67).

Furthermore, contemporary research has shown that protozoa are important in the dynamics of the aquatic system. Grazing by heterotrophic protozoa has now been accepted as the most important factor responsible for limiting the numbers of bacterial cells in aquatic ecosystems and maintaining nutrient availability. There is evidence suggesting that bioremediation processes involving bacteria may be strongly inhibited in the presence of phagotrophic protozoa, for instance. Sometimes, microbial population growth rates are difficult to ascertain because ubiquitious grazers remove cells as quickly as they are produced. These taxonomically heterogeneous organisms (including ciliates, amoebae, and a variety of flagellates) can graze voraciously on phytoplanktonic and bacterial cells that form the foundation of the marine food web. Discrimination has been achieved between prey cells and a dinoflagellate through DNA staining and isolation of the eukaryotic nuclei. Furthermore, differentiation of phagocytosic activity within populations of protozoa is measurable using fluorescently labeled microbeads (68).

Some investigations have been performed on viruses, regarded as ubiquitious, and biologically active members of marine and freshwater microbial communities, in which the viruses may be responsible for the destruction of a considerable proportion of the bacterial and cyanobacterial populations. In some cases, virus and host support each other in a symbiotic manner, in contrast to the observed

annihilation of alternative host populations within the microbial community.

Plants

Application of flow cytometry and cell sorting procedures to study higher plant cells required methods for the production of single-cell suspensions. The need to probe the performance of protoplasts was the main motivation for developing methods for analyzing plant cells. This also required a series of modifications to the flow cytometric instrumentation, because plant cells were sometimes found to be much larger than those of other living organisms.

Flow cytometry has been accepted as a rapid and trustworthy technique for genome analysis; for obtaining information about the size, ploidy, and aneuploidy changes during plant evolution; and for the differentiation and state of replication of plant cells. It provides an accurate method for determining the proportions of cells in the G1, S, and G2/M stages of the cell cycle and for calculating cell cycle times. Consequently, cell-cycle-dependent events such as the expression of distinct proteins can be observed in synchronized cell cultures. For example, the expression of a *Nicotiana tabacum* cyclin, a homologue to the *cdc2* gene, disappears in the G1 and S phases of the cell cycle, but is present in the G2/M phase (69).

Flow karyotyping from mitotic cells and chromosome sorting provide opportunities for gene mapping and the construction of chromosome-enriched DNA libraries. Thus, the analysis and sorting of plant chromosomes is of considerable economic interest. Also, the occurrence and extent of polyploidization, which often accompanies differentiation of plant cells, is important for understanding the regulation of gene expression in differentiated tissues. Furthermore, in situ hybridization methods allow accurate quantification of copy numbers of repetitive DNA sequences on different chromosomes within a species.

Examination of intra- and interspecific variation of DNA content can be important in plant hybridization and genetic manipulation programs. Therefore, the method is widely used in molecular cytogenetic research. For example, it can be used to search for foreign chromosomes added to a species, as in the introduction of resistance genes into wheat from the germplasm of rye. Moreover, one of the major problems of using wheat–rye addition lines is their genetic instability. The absence of the easily lost rye chromosomes must be detected, but counting the chromosomes by methods other than flow cytometry is tiresome and time consuming.

For some properties it is necessary to obtain data concerning distinct compartments within a plant cell, and a fundamental goal is appraisal of the physiological behavior of plant cell protoplasts. For this, there are a number of flow cytometric methods for assessing different intracellular components, such as reserve metabolites as well as enzyme activities. Changes in the cytosolic pH using common pH-specific probes have also been shown using these methods. Indeed, it has been proved that analysis of the quantity and quality of cellular fluorescence can yield important information concerning tissue status and derivation.

Furthermore, plant flow cytometry can be a powerful tool in the examination of symbiotic systems. For instance, *Azolla filiculoides* harbors phototrophic, aerobically grown cyanobacteria in vegetative cells and nitrogen-fixing anaerobic cyanobacteria in heterocysts. Some of these differentiated tissues have been investigated regarding the internal distribution of the blue-green algae and the capacities of the various tissues for nitrogen fixation using DNA estimation (70).

Recently, the green fluorescent protein (GFP) from the jellyfish *Aequora victoria* has been used as a powerful new vital reporter in diverse plant cells. In higher plant protoplast studies, it can be used to monitor gene expression, signal transduction, cotransfection, transformation, protein trafficking and localization, protein–protein interactions, and cell separation and purification. In plant cell systems, the bright green fluorescence of GFP is visible even in the presence of the red chlorophyll autofluorescence. GFP is very stable in plant cells and shows little photobleaching. Viable cells, marked with GFP, can be recovered after fluorescence-activated cell sorting (71).

Animal Cells

Flow cytometry is widely exploited in the study of spermatogenesis and routinely applied to estimate viability. It is also used to determine the chromosome content and sex of the offspring. It is a valuable and adaptable tool for research into the genesis of germ cells in both normal and perturbed situations, and it can be used in the control of reproduction, domestic animal breeding, and wildlife conservation. Furthermore, cytogenetic investigation of eukaryotic organisms is important for understanding evolutionary processes and essential for testing evolutionary models. Improved knowledge of the morphology of germ cells, and the motility and fertility of spermatozoa enhance the efficiency of preinsemination technology. Also, functional heterogeneity can be easily detected, as in the assessment of mitochondrial lesions, which are connected to subfertility. The abundance of live cells, altered sperm morphology, and chromatin structure and the resistance of membrane-bound enzymes and membrane permeability to environmental stress factors can also be calculated from flow cytometric data.

In contrast to other methods, the productivity of the flow cytometric separation of the X or Y chromosome-containing germ cells (using highly efficient staining and sorting methods) yields highly purified samples that can be used for sex control of domestic animals (72). More accurate still is slit-scan technology, which allows the whole X or Y germ cell to be spatially visualized along the direction of flow. This is important, because the intensity of fluorescence detected is affected by the flat ovoid head shape and the compactness of the chromatin. Thus, the orientation of the germ cell during passage through the laser beam, detectable by slit-scan flow cytometry, is critical for a high efficiency sorting procedure.

Another range of applications in flow cytometry lies in improving the control of antibody production. Generally, hybridoma cell lines serve as sufficient carriers. Optimizing the productivity of hybridoma cell lines depends partly

on developing suitable methods for screening and selecting highly productive cultures and on understanding the regulation of antibody production. Surface-antibody immunofluorescence can be used to select high yielding cells by sorting and to gauge the efficacy of the whole population by cytometric cell analysis. Furthermore, information can be obtained about the stability of expression, the yield, and the integrity of the antibodies obtained. Antigen production can be improved using cytostatically growth-arrested cells, because they do not need to consume cellular resources for biomass production (73). Besides the hybridoma lines, other cell lines (often from insects) have been systematically developed for somatic cell genetics, expression of transfected genes, and synthesis of hormone-inducible proteins. Toward this end, flow cytometric-based studies on animal cell systems can be used for developing convenient and widely applicable procedures for gene cloning. There are, for instance, protocols for DNA transfection and subsequent selection of rare transfectants by sorting. For such investigations, the green fluorescent protein (GFP) may be inserted. Several new GFP variants have been generated that are brighter or have altered excitation spectra, facilitating the monitoring of the marked cell lines.

Viruses have a particularly important biotechnological role in the expression of both foreign prokaryotic and eukaryotic genes in host cells. However, only a few virus-cell systems have been analyzed by flow cytometry until now, though the method is competent for broad applications in detecting and quantifying virus-infected cells.

Viruses can be assayed in animal cell cultures by in situ hybridization using fluorescently marked oligonucleotides and quantitatively estimated by flow cytometry. Characterization of the accessible virus host systems offers information about their capability for providing adequate quantities of virus outgrowth. The process of productive and nonproductive virus infection in animal cell lines can also be investigated, for example, by in vivo analysis of promoter activity regarding transcription of the virus gene. The staging of the virus can be estimated using a virus reporter gene, whose product is, for instance, immunofluorescently stainable. Another possibility is to insert an insect luciferase gene, encoding a protein with characteristic light-emission properties that are measurable by flow cytometry. High sensitivity and rapidity of the reaction (in combination with the luminescence, which does not naturally occur in most of the cellular systems used) makes the method valuable for flow cytometric estimation of virus production. It is also possible to get information about the reproductive capabilities of the virus-infected and uninfected host systems. Differences can be analyzed between highly susceptible and less susceptible cells. Furthermore, virus-induced DNA synthesis may be monitored by assisted determination of BrdUrd incorporation, using monoclonal antibodies. An overview is given by McSherry (74).

Another biological field where the prospects offered by flow cytometry seem extremely encouraging, besides clinical diagnostics, is pharmacology. Before the application of a drug is allowed for human use, the mode of action at cellular and subcellular levels must be clear. The experimental conditions must be as close as possible to physiological drug–cell interactions within living target cells. Model systems have been developed to study the operation of curative drug systems. On-line flow cytometry is the most sensitive and fastest method available, and it causes the least possible harm to living cells.

Additionally, the method is appropriate for investigating the mode of action of other biologically active substances. In phytopathological sciences, for instance, insect cell lines are very suitable target systems for studying the cytotoxic potential of microbe-produced cytostatic agents.

BIBLIOGRAPHY

1. H.M. Shapiro, *Practical Flow Cytometry*, Wiley-Liss, New York, 1995.
2. R. Hüller, E. Glossner, S. Schaub, J. Weingärtner, and V. Kachel, *Cytometry* **17**, 109–118 (1994).
3. J.E. Bailey, J. Fazel-Madjlessi, D.N. McQuitty, L.Y. Lee, J.C. Allred, and J.A. Oro, *Science* **198**, 1175–1176 (1977).
4. M.L. Slater, S.O. Sharrow, and J.J. Gart, *Proc. Natl. Acad. Sci. USA* **74**, 3850–3854 (1977).
5. K.J. Hutter and H.E. Eipel, *Eur. J. Appl. Microbiol. Biotechnol.* **6**, 223–231 (1979).
6. L. Boddy, C.W. Morris, M.F. Wilkins, G.A. Tarran, and P.H. Burkill, *Cytometry* **15**, 283–293 (1994).
7. A.G. Fredrickson, D. Ramkrishna, and H.N. Tsuchiya, *Math. Biosci.* **1**, 327–374 (1967).
8. Th. Münch, B. Sonnleitner, and A.J. Fiechter, *Biotechnology* **22**, 329–352 (1992).
9. D. Lloyd, *Flow Cytometry in Microbiology*, Springer, Berlin, 1992.
10. A. Degelau, R. Freitag, F. Linz, C. Middendorf, Th. Scheper, Th. Bley, S. Müller, P. Stoll, and K.F. Reardon, *J. Biotechnol.* **25**, 115–144 (1992).
11. P. Fouchet, C. Jayat, Y. Hechard, M.-H. Ratinaud, and G. Frelat, *Biol. Cell* **78**, 95–109 (1993).
12. H.B. Steen and E. Boye, *Cytometry* **1**, 32–36 (1980).
13. K. Skarstad, H.B. Steen, and E. Boye, *J. Bacteriol.* **163**, 661–668 (1985).
14. T. Åkerlund, K. Nordström, and R. Bernander, *J. Bacteriol.* **177**, 6791–6797 (1995).
15. Th. Scheper, B. Hitzmann, U. Rinas, and K. Schügerl, *J. Biotechnol.* **5**, 139–148 (1987).
16. S. Müller, A. Lösche, Th. Bley, and T. Scheper, *Appl. Microbiol. Biotechnol.* **43**, 93–101 (1995).
17. J.-U. Ackermann, S. Müller, A. Lösche, Th. Bley, and W. Babel, *J. Biotechnol.* **39**, 9–20 (1995).
18. J.-H. Seo and J.E. Bailey, *Biotechnol. Bioeng.* **27**, 156–165 (1985).
19. P. Fouchet, C. Manin, H. Richard, G. Frelat, and J.N. Barbotin, *Appl. Microbiol. Biotechnol.* **41**, 584–590 (1994).
20. T. Azuma, G.I. Harrison, and A.L. Demain, *Appl. Microbiol. Biotechnol.* **38**, 173–178 (1992).
21. J.D. Chung, S. Conner, and G. Stephanopoulos, *Cytometry* **20**, 324–333 (1995).
22. M. Poot and S. Arttamangkul, *Cytometry* **28**, 36–41.
23. J. Porter, R. Pickup, and C. Edwards, *FEMS Microbiol. Lett.* **134**, 51–56 (1995).
24. T. Tolker-Nielsen, K. Holmstrom, and S. Molin, *Appl. Environ. Microbiol.* **63**, 4196–4203 (1997).

25. J. Chaloupka and V. Vinter, *Folia Microbiol.* **41**, 451–464 (1996).
26. G. Nebe-von Caron and R.A. Badley, in *Flow Cytometry Applications in Cell Culture*, Dekker, New York, 1996.
27. D. Lloyd and A.J. Hayes, *FEMS Microbiol. Lett.*, **133**, 1–7 (1995).
28. J. Porter, D. Deere, R. Pickup, and C. Edwards, *Cytometry* **23**, 91–96 (1996).
29. D.B. Kell, H.M. Ryder, A.S. Kaprelyants, and H.V. Westerhoff, *Antonie van Leeuwenhoeck* **60**, 145–158 (1991).
30. S. Müller, N. Loffhagen, Th. Bley, and W. Babel, *Microbiol. Res.* **151**, 127–131 (1996).
31. C. Herrmann, A. Lösche, S. Müller, Th. Bley, and W. Babel, *Acta Biotechnol.* **17**, 19–38 (1997).
32. J.V. Ordóñez and N.M. Wehmann, *Cytometry* **14**, 811–818 (1993).
33. D.J. Masen, R. Lopez-Amoros, R. Allmann, J.M. Stark, and D. Lloyd, *J. Appl. Bacteriol.* **78**, 309–315 (1995).
34. E. Sahar, R. Nir, and R. Lamed, *Cytometry* **15**, 213–221 (1994).
35. J. Yagoda-Shagam, L.L. Barton, W.P. Reed, and R. Chiovetti, *Appl. Environ. Microbiol.* 1831–1837 (July 1988).
36. G. Wallner, R. Amann, and W. Beisker, *Cytometry* **14**, 136–143 (1993).
37. C.A. Sanders, D.M. Yajko, W. Hyun, R.G. Langlois, P.S. Nassos, M.J. Fulwyler, and W.K. Hadley, *J. Gen. Microbiol.* **136**, 359–365 (1990).
38. S.B. Riis, H.M. Pedersen, N.K. Sørensen, and M. Jakobsen, *Food Microbiol.* **12**, 245–250 (1995).
39. C. Laplace-Builhé, K. Hahne, W. Hunger, Y. Tirilly, and J.L. Drocourt, *Biol. Cell.* **78**, 123–128 (1993).
40. R.A. Patchett, J.P. Back, A.C. Pinder, and R.G. Kroll, *Food Microbiol.* **8**, 119–125 (1991).
41. R.G. McClelland and A.C. Pinder, *J. Appl. Bacteriol.* **77**, 440–447 (1994).
42. C.M. Futsaether and A. Johnsson, *J. Microbiol.* **40**, 439–445 (1994).
43. L.A. van der Waaij, G. Mesander, P.C. Limburg, and D. van der Waaij, *Cytometry* **16**, 270–279 (1994).
44. E.D. Renner, *Cytometry* **18**, 103–108 (1994).
45. E.E. Hughes, J.M. Matthews-Greer, and H.E. Gilleland Jr., *Can. J. Microbiol.* **42**, 859–862 (1996).
46. M.A. Norden, T.A. Kurzynski, S.E. Bownds, S.M. Callister, and R.F. Schell, *J. Clin. Microbiol.* **33**, 1231–1237 (1995).
47. B. Husevåg, *FEMS Microbiol. Ecol.* **16**, 25–32 (1995).
48. S. Langsrud and G. Sundheim, *J. Appl. Bacteriol.* **81**, 411–418 (1996).
49. I.S. Roberts, *Annu. Rev. Microbiol.* **50**, 285–315 (1996).
50. M.F. Gilbert, D. McQuitty, and J.E. Bailey, *Appl. Env. Microbiol.* **36**, 615–617 (1978).
51. K.-J. Hutter, *FEMS Microbiol. Lett.* **3**, 35–38 (1978).
52. Alberghina, M. Mariani, E. Martegani, and M. Vanoni, *Biotechnol. Bioeng.* **25**, 1295–1310 (1983).
53. Th. Scheper, H. Hoffmann, K. Schügerl, *Enzyme Microb. Technol.* **9**, 399–405 (1987).
54. D.E. Block, P.D. Eitzmann, J.D. Wangesteen, and F. Srienc, *Biotechnol. Prog.* **6**, 504–512 (1990).
55. D. Porro, B.M. Ranzi, C. Smeraldi, E. Martegani, and L. Alberghina, *Yeast* **11**, 1157–1169 (1995).
56. S. Müller and A. Schmidt, *Acta Biotechnol.* **9**, 71–75 (1989).
57. S. Müller, A. Lösche, and Th. Bley, *Acta Biotechnol.* **12**, 365–375 (1992).
58. S. Müller, K.J. Hutter, Th. Bley, L. Petzold and W. Babel, *Bioproc. Eng.* **17**, 287–293 (1997).
59. K.D. Wittrup and J.E. Bailey, *Cytometry* **9**, 394–404 (1988).
60. K.D. Wittrup, J.E. Bailey, B. Ratzkin, and A. Patel, *Biotechnol. Bioeng.* **35**, 565–577 (1990).
61. P.D. Eitzman, J.L. Hendrick, and F. Srienc, *Cytometry* **10**, 175–483 (1989).
62. C. Smeraldi, E. Berardi, and D. Porro, *Microbiology* **140**, 3161–3166 (1994).
63. J.V. Ordonez and N.M. Wehman, *Cytometry* **22**, 154–157 (1995).
64. P.J. Constantino, D.E. Budd, and N.F. Gare, *Cytometry* **19**, 370–375 (1995).
65. D.K. Button and B.R. Robertson, *Cytometry*, **10**, 558–563 (1989).
66. D. Vaulot, D. Marie, R.J. Olsen, and S.W. Chisholm, *Science* **9**, 1480–1482 (1995).
67. N. Simon, N. LeBot, D. Marie, F. Partensky, and D. Vaulot, *Appl. Environ. Microbiol.* **61**, 2506–2513 (1995).
68. V.S. Avery, J.L. Harwood, and D. Lloyd, *Appl. Environ. Microbiol.* **61**, 1124–1132 (1995).
69. L.X. Qin, C. Perennes, L. Richard, M. Bouvier-Durand, C. Trehin, D. Inze, and C. Bergounioux, *Plant. Mol. Biol.* **32**, 1093–1101 (1996).
70. A. Canini, M.G. Caiola, and L. Ferrucci, *Cytometry* **13**, 299–306 (1992).
71. J. Sheen, S. Hwang, Y. Niwa, H. Kobayashi, and D.W. Galbraith, *Plant J.* **8**, 777–784 (1995).
72. S. Damjanovich, R. Gaspar Jr., L. Tron, and A. Aszalos, *Am. Biotech. Lab.* **3**, 11–21 (1985).
73. M. Fussenegger, X. Mazur, and J.E. Bailey, *Biotechnol. Bioeng.* **55**, 927–939 (1997).
74. J.J. McSherry, *Clin. Microbiol. Rev.* **7**, 576–604 (1994).

FLUORESCENCE TECHNIQUES FOR BIOPROCESS MONITORING

CARSTEN LINDEMANN
STEFAN MAROSE
THOMAS SCHEPER
University of Hannover
Hannover, Germany

HANS OLE NIELSEN
DELTA Light&Optics
Lyngby, Denmark

KENNETH F. REARDON
Colorado State University
Fort Collins, Colorado

KEY WORDS

Bioprocess monitoring
Fluorescence
On-line
Optical sensor

OUTLINE

INTRODUCTION

In recent years, various bioanalytical systems for improved bioprocess monitoring have been developed and investigated. The application of these systems in pilot facilities and production plants can provide the information needed to better understand all types of bioprocesses, including cultivations, biotransformations, and downstream operations. In addition, better documentation of the processes, especially important for the production of pharmaceuticals, can be obtained. However, most of these sensors are invasive, and thus there is the chance of impacting the cultivation process (e.g., contamination) and causing complications in the processes. For this reason, more and more research groups are involved in the development and application of noninvasive bioanalytical tools, which can be interfaced to the bioprocess without causing problems such as contamination or changes in cell metabolism.

Optical sensors offer tremendous advantages for bioprocess monitoring. They can be interfaced noninvasively to the bioprocess via glass fiberoptics. The information can be transported via fiberoptics without interferences caused by electromagnetic fields over long distances, and signals can be obtained instantaneously, because spectro-optical systems have extremely short response times.

Fluorescence sensors have been investigated intensively during the past 15 years for different applications in biotechnology, especially biomass estimation, reactor characterization, metabolic studies, and general bioprocess monitoring. These so-called fluorosensors were developed mainly for monitoring of the reduced form of nicotine adenine dinucleotide (NADH) and its phosphorylated form (NADPH). This coenzyme is present in all living cells and is involved in different metabolic reactions; its level is extremely sensitive to environmental conditions.

PRINCIPLES OF SINGLE-WAVELENGTH FLUOROSENSORS

The NAD(P)H pool gives information about the metabolic state of the cells. In 1957, Duysens and Amesz first showed that relative NAD(P)H measurements could be performed in living cells using spectrofluorometric techniques (1). UV light of 340 to 360 nm was used to excite fluorescence of NAD(P)H in cells in a suspension. The NAD(P)H fluorescence was monitored at 460 nm while metabolic experiments were performed. The first spectrometer device for biotechnological application was designed by Harrison and Chance in 1970 (2). It could be used in glass bioreactors. The excitation light was guided through a quartz window into the bioreactor, and the fluorescence was collected at an angle of 60° through a second quartz window. The penetration depth of the light was only a few centimeters at low cell densities. The sensitivity decreased as the cell density increased. However, it was possible to monitor aerobic and anaerobic transitions in *Klebsiella aerogenes*. Zabriskie and Humphrey used this technique in 1978 for biomass estimation in different cultivation experiments (e.g., *Saccharomyces cerevisiae*, *Streptomyces* spp., and *Thermoactinomyces* spp.) (3).

Based on these experiments, different smaller miniaturized process fluorometers were developed that could be interfaced to a bioreactor via a standard port (Fig. 1). This port is used to guide the excitation light into the bioreactor and to collect the backward (180°) fluorescence. In 1981, Beyeler et al. (4) developed a miniaturized fluorosensor in an open-end detection mode. This design has the advantage of eliminating the inner filter effect by monitoring the surface fluorescence at the observation window. Three commercial versions of the open-end spectrofluorometer were produced by BioChem Technology/USA (FluoroMeasure), Mettler Toledo AG/Switzerland (Fluorosensor), and BioBalance A/S, Denmark (Biobalance).

All these sensors shared two disadvantages. First, they were limited to detection of NAD(P)H fluorescence, it was only possible to excite fluorophores in the range of 340 to 360 nm and to detect the fluorescence intensities at 450 to 460 nm. And they also lacked the ability to distinguish between different fluorophores that might interfere in this region.

In order to obtain more information, a flexible two-channel fluorometer was developed by Scheper and Schügerl (5). In this device, the light of the UV lamp was focused out the optical filters and lenses onto a quartz fiber bundle. This fiberoptic was used to guide the light into the bioreactor. The backward fluorescence light was collected via the same fiber bundle and split into two bundles, passing

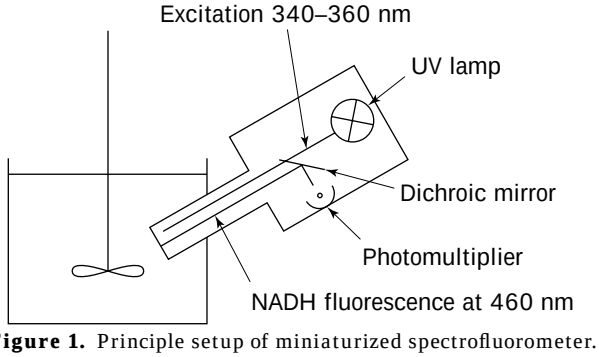


Figure 1. Principle setup of miniaturized spectrofluorometer.

interference filters that selected for different wavelength bands. Thus, it was possible to monitor simultaneously different fluorescence wavelengths.

APPLICATION OF SINGLE-WAVELENGTH FLUOROSENSORS

Single-wavelength fluorosensor devices have been used to monitor different cultivation processes, including suspended as well as immobilized organisms (Table 1). Investigations in which the NAD(P)H fluorescence of these cultivations was monitored had one of two goals: on-line determination of biomass concentration or on-line detection of metabolic changes. A significant advantage of using fluorescence for biomass concentration monitoring is that only living cells are detected. To use culture fluorescence as an indicator for biomass concentration, it is necessary that the culture fluorescence per cell is constant during the cultivation process. In this case, a linear correlation between fluorescence and the biomass concentration could be expected. However, the intracellular NAD(P)H pool changes in a cell during the batch experiment. Thus, it is necessary to linearize the corresponding biomass concentration and fluorescence data to obtain a calibration plot. One problem that must be overcome is that all process conditions affecting the culture fluorescence intensity cause problems with the biomass determination. The two most common interferences are gas bubbles and the presence of compounds with fluorescence spectra that overlap that of NAD(P)H. Interestingly, fouling of observation windows was not observed.

Many investigators have reported efficient biomass concentration determination via culture fluorescence monitoring. Examples include cultivations of *Methylomonas mucosa*, *Zymomonas moblis* in synthetic medium, and *Pseudomonas putida*. In all these applications, suspended cell cultivations were studied, and it was shown that on-line biomass estimation is also possible in technical media under certain process conditions (for detailed information, see Ref. 6).

Table 1. Application of Fluorescence Monitoring for Different Cultivation Processes and Organisms

Suspended
Saccharomyces cerevisiae
Penicillium chrysogenum
Cephalosporium acremonium
Escherichia coli (incl. recombinant)
Zymomonas mobilis
Bacillus licheniformis
Clostridium acetobutolyticum
Spodoptera frugiperda
Baby hamster kidney cells
Wastewater treatment
Immobilized
Cephalosporium acremonium
Saccharomyces cerevisiae
Zymomonas mobilis

Source: See Refs. 6 and 7.

A wide range of metabolic changes have also been studied using NAD(P)H culture fluorescence (Table 2). For example, Ristroph et al. (8) studied the cultivation of *Candida utilis* under different conditions and developed a feeding strategy for fed-batch cultivation of this organism on the basis of NAD(P)H-dependent culture fluorescence. Einsele (9) reported on different applications of culture fluorescence monitoring and determined the mixing-time behavior of different bioreactors and the substrate uptake rate of different microorganisms during cultivation. Reardon et al. (6) studied immobilized *Clostridium acebutylium* in a fixed-bed reactor. The cells were immobilized in calcium alginate beads, and the cultivation medium was circulated through the reactor. The growth and productivity of the cells could be monitored under different cultivation conditions. Scheper et al. reported (10) on the application of this technique for the monitoring of immobilized yeast cells.

NEW DEVELOPMENTS IN FLUOROSENSOR TECHNOLOGY AND APPLICATION

Several approaches have been taken in the past 8 years to circumvent the restrictions of single-wavelength NAD(P)H-dependent culture fluorescence monitoring. In 1991, Li et al. (11) reported on a special fluorosensor with five different wavelengths for the monitoring of *Saccharomyces cerevisiae* and *Candida utilis* cultivations. They showed that other wavelengths could be used for biomass concentration estimation and for metabolic studies. Horvath et al. (12) used a fluorescence spectrometer coupled to a bioreactor via quartz glass fibers for the determination of *S. cerevisiae* biomass concentrations. Based on his results, the tryptophan fluorescence was better correlated to the cell mass concentration than was the NAD(P)H fluo-

Table 2. Different Phenomena Studied via Fluorescence Monitoring

Phenomen studied	Organism
Aerobic–anaerobic transition	*Klebsiella aerogenes* *Saccharomyces cerevisiae* *Candida tropicalis* *Escherichia coli* *C. guilliermondii*
Aeration rate	*Penicillium chrysogenum*
Addition of carbon source to starved cells	*Saccharomyces cerevisiae* *Candida tropicalis* *Escherichia coli*
Diauxic growth	*Saccharomyces cerevisiae*
Dilution rate changes	*Saccharomyces cerevisiae* *Escherichia coli* *Pseudomonas putida*
Glycolytic oscillation	*Saccharomyces cerevisiae*
Culture synchrony	*Saccharomyces cerevisiae*
Metabolic shifts	*Thermoactinomyces* sp. *Clostridium acetobutylicum*
Addition of metabolic uncouplers	*Saccharomyces cerevisiae*

Source: Based on Refs. 3 and 4.

rescence. A more recent development was described by Chenina in 1993 (13) for the cultivation of yeasts (*Cyathus striatus*, *Eurotium cristatum*, and *Crinipellis stipitaria*). A special fluorescence spectrometer was developed that can also be interfaced to a bioreactor via light guide cables. Monochromators were used to monitor the fluorescence of wide excitation and emission ranges.

TWO-DIMENSIONAL FLUORESCENCE

These fluorescence spectrometers can be used to produce a two-dimensional (2-D) fluorescence spectrum, which represents the emission spectra at different excitation wavelengths. Such a spectrum is described by the three parameters of excitation wavelength, emission wavelength, and fluorescence intensity, which can be shown as an excitation-emission-data matrix. In this matrix, each row represents an emission scan and each column an excitation scan. Different compounds can be identified by their position in the matrix. The most illustrative way to show a 2-D fluorescence spectrum is to use an isometric plot, but contour plots are the most useful because no fluorescence peaks are hidden (Fig. 2). Dark shadings in the fluorescence plots indicate high fluorescence intensities.

BIOGENIC FLUOROPHORS

The spectral ranges of important biogenic fluorophores, which occur inside the cell and can be detected simultaneously by 2-D fluorescence spectroscopy, are shown in Figure 3. Almost all biological processes depend on proteins, which fluoresce because of aromatic amino acids such as tryptophan, tyrosine, and phenylalanine. In another region of the 2-D fluorescence spectrum (λ_{Ex} = 310–480 nm, λ_{Em} = 350–550 nm), vitamins (e.g., pyridoxine [vitamin B-6]), riboflavin [vitamin B-2]), and coenzymes (e.g., NADH, NADPH, FMN, FAD) can be found.

INSTRUMENTATION

For on-line monitoring and control of biotechnological processes using 2-D fluorescence, the instrument must be capable of measuring a complete 2-D fluorescence spectrum in a short time (about 1 min), and the time between two measurements should not be longer than 5 to 15 min. This is to ensure that a complete 2-D fluorescence spectrum represents the fluorescence at a certain time and that changes during a cultivation can be monitored in real time.

HITACHI F4500 SPECTROFLUOROMETER

For 2-D fluorescence measurements, fluorescence spectrometers for use in laboratories can be used. One suitable apparatus is the fluorescence spectrophotometer model F4500 (Hitachi, Japan). This spectrometer allows a very fast scan speed (30,000 nm/min) and an entire 2-D spectrum with a wavelength range of 250 to 550 nm for excitation (10-nm step size) and 260 to 600 nm for emission (20-nm step size) can be measured within 1 min. For selection of excitation and emission wavelength, grating monochromators are used. A 150-W Xe lamp is used for excitation, and a photomultiplier at a voltage of 700 V is used for detection of fluorescence.

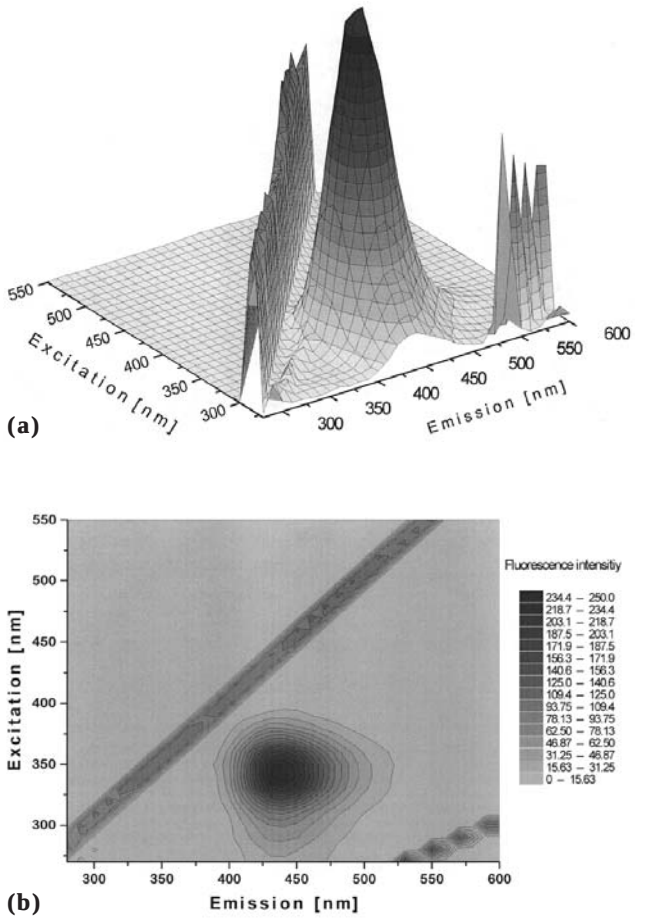


Figure 2. (**a**) Isometric projection of a spectrum of 7-amino-4-methylcoumarin. (**b**) Contour plot of a spectrum of 7-amino-4-methylcoumarin.

The spectrometer is connected by a 1-m bifurcated liquid light conductor (Lumatec, Germany) to a quartz window in a 25-mm port of a bioreactor (Fig. 4). The excitation light is guided via the fiberoptic into the reactor, and the backward fluorescent light is guided via the second fiberoptic to the emission detection unit. With this setup, in situ measurements are possible without interference from light outside the bioreactor, and there is no risk of contamination because the sensor is not in contact with the medium. Furthermore, sterilization is possible without connecting the sensor to the reactor.

BIOVIEW

The BioView sensor (DELTA Light & Optics, Denmark) is a fluorescence sensor optimized for fully automated industrial measurements and applications to monitoring and control of bioprocesses. It is very robust and suitable for harsh environments (e.g., high temperature or moisture)

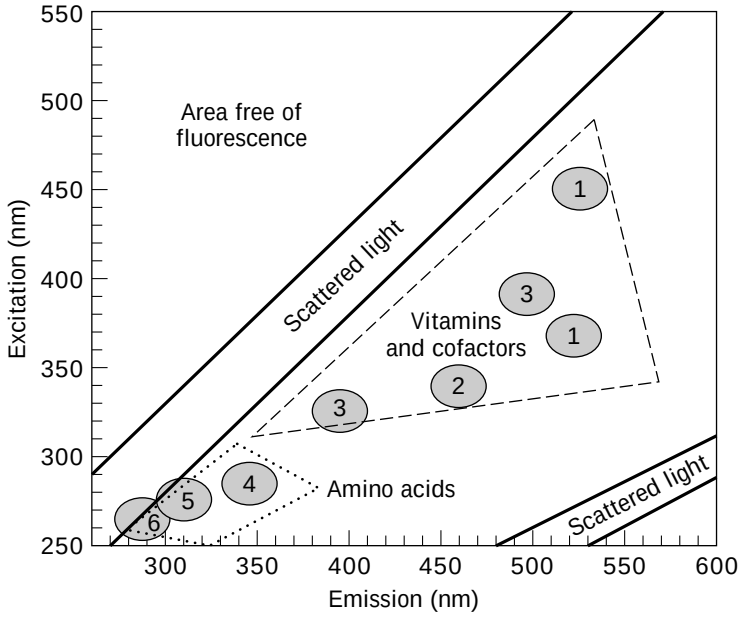


Figure 3. Biogenic fluorophores in a 2-D fluorescence spectrum. 1 = riboflavin, FAD, FMN; 2 = NAD(P)H; 3 = pyridoxine, pyridoxamine, pyridoxal-5′ = phosphate; 4 = tryptophan; 5 = tyrosine; 6 = phenylalanine). *Source*: Adapted from reference 9.

and electromagnetic interference. For data transfer, a single-fiber modem is used, which allows a distance between sensor and computer up to several hundred meters. The optical parts of the BioView are optimized for use in biotechnological processes in that interference filters (DELTA Light & Optics, Denmark), rather than spectrometer gratings, are used by the BioView sensor for selection of excitation and emission wavelengths. Filter wheels (one for excitation and another for emission) with 16 filters each are used (Fig. 5). These filters can be designed and chosen individually. Usually, measurements in steps of 20 nm are performed in a wavelength range of 270 to 550 nm for excitation and 290 to 590 nm for emission. For the measurement of a complete 2-D fluorescence spectrum, the excitation filter wheel is set to the first filter position. With this excitation filter, the fluorescence spectrum is monitored via the emission filters that switch from filter to filter until a cycle is completed. Afterward, the excitation filter switches to the next position and the next fluorescence spectrum is measured. This procedure is continued until the excitation filter wheel has performed a complete cycle. The complete measurement takes about 1 min, depending on the number of different filter positions and measurements for each filter combination, which can be chosen individually. The connection to the bioreactor is the same as that described for the Hitachi instrument, except that a 2-m long liquid light conductor is used.

APPLICATIONS

Two-dimensional fluorescence spectroscopy offers the possibility of simultaneous analysis of several fluorophores. The location of the peak maximum in a 2-D spectrum characterizes each fluorophore, and the relative fluorescence intensity correlates with its concentration. Despite this conceptual simplicity, the interpretation of 2-D fluorescence spectra is not trivial. Interactions of the fluorophores with one another and with other nonfluorescent compounds, bubbles, and other physical influences such as pH, temperature, and viscosity, quenching effect, and changes in morphology of the cells all cause problems in the interpretation of the 2-D spectrum.

To make the system user friendly, automated interpretation by artificial intelligence was developed for 2-D fluorescence spectra. The intelligent BioView sensor is able to predict the course of a cultivation if it is taught with a training cultivation under comparable cultivation conditions. Modern chemometric systems like principle component analysis (PCA) or neural networks are able to consider all influences on the fluorescence signal and lead to a sensor with high capacity, able to measure various cultivation variables.

The changes in fluorescence intensities during cultivation processes are shown in difference spectra, which are obtained by subtraction of a 2-D spectrum taken during the cultivation from the one in the beginning. Through comparison to the spectrum of a mixture of pure biogenic fluorophores, such as riboflavin, NADH, pyridoxine, pyridoxamine, pyridoxal-5′-phosphate, and tryptophan, the peaks could be assigned to the expected fluorophores. All changes have a biological origin and offer information about the cell growth and metabolism. Even small changes in the concentration of a single fluorophore can be seen in a complex mixture of several fluorophores. The regions of changes are regarded as the areas of interest.

Measurements made nearly continuously allow us to construct the course of various culture variables during a bioprocess. One of the most important is the biomass concentration. Off-line analysis of counting cells, measuring optical density or weighting dry mass are time consuming and fail if filamenteous fungi are used. On-line tools such as a turbidimeter often cause problems. The search for a reliable automated technique for on-line monitoring of biomass still continues. In a 2-D fluorescence spectrum, vari-

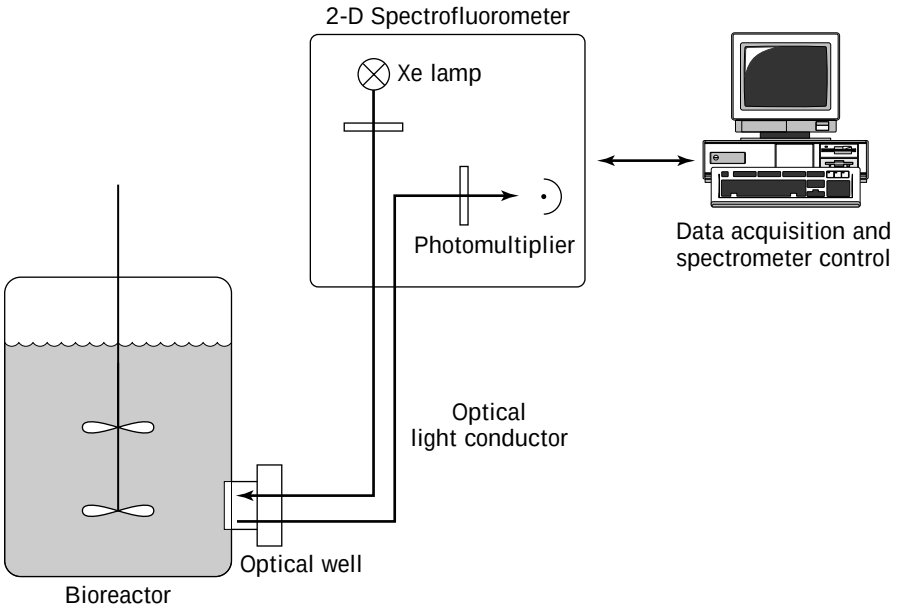


Figure 4. Measuring setup.

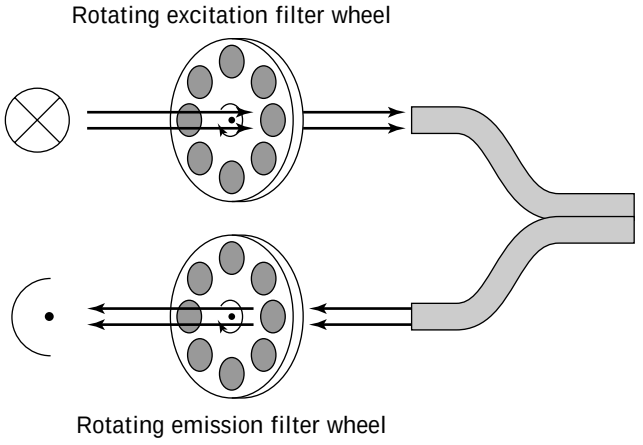


Figure 5. Measuring principle of the BioView sensor.

ous excitation–emission regions can be correlated to biomass concentration. This offers the advantage to change to another region if one cannot be used because of overlapping or other influences. Fluorescence of proteins (because of their high quantum yield, especially in the case of tryptophan), pyridoxine, the area of NAD(P)H, and the area of riboflavine and its derivatives (FAD or FMN) show good correlations with biomass for several organisms (e.g., *Acremonium crysogenum*, *Claviceps purpurea*, *Sphingomonas yanoikuyae*, *Bacillus licheniformis*, different strains of *Escherichia coli*, *Enterobacter aerogenes*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, mammalian cells or *Tetrahymena thermophila* (9,14). For the filamenteous fungi, *Claviceps purpurea* fluorescence intensity in the region of riboflavin and its derivatives (λ_{Ex} = 450 nm, λ_{Em} = 530 nm) allow on-line monitoring of biomass concentration, which is not possible by optical density or cell count because of the mycelial growth. Scattered light is also measured in a 2-D fluorescence spectrum. These data correlate well with data derived from conventional turbidimetric sensors.

Along with the ability to estimate biomass concentration, 2-D fluorescence sensors such as the BioView system offer on-line information about intracellular variables without interference into the process. This is achieved by monitoring changes in the excitation–emission regions associated with the redox-sensitive coenzymes NADH, FAD, and FMN. This approach can be used to study aerobic–anaerobic transitions, the tricarboxylic cycle, effects of uncoupling agents such as 2,4-dinitrophenol (DNP) on oxidative phosphorylation, diauxies, the effects of pulse additions of substrate, and oscillations of a culture.

Fluorescent products can be observed by 2-D fluorescence. The productivity of ergot alkaloids could be estimated by 2-D fluorescence spectroscopy (9). Degradation of fluorescent substrates such as polycyclic aromatic hydrocarbons can be observed. Other extracellular or physical variables such as the microenvironment of the cells are captured by 2-D fluorescence.

Two-dimensional fluorescence spectroscopy is ideal for a wide variety of organisms and installation uses and can therefore be seen as a multianalyzer that is capable of replacing several other sensors in the future.

BIBLIOGRAPHY

1. L.N.M. Duysens and J. Amesz, *Biochim. Biophys. Acta* **24**, 19–26 (1957).
2. D.E.F. Harrison and B. Chance, *Appl. Microbiol.* **19**, 446–450 (1970).
3. D.W. Zabriskie and A.E. Humphrey, *Appl. Eur. Microbiol.* **35** 337–343 (1978).
4. W. Beyeler, K. Gschwend, and A. Fiechter, *Chem. Ing. Technol.* **55**, 869–871 (1983).

5. T. Scheper and K. Schügerl, *J. Biotechnol.* **3**, 221–229 (1986).
6. T. Scheper and K.F. Reardon, in W. Göpel, J. Hesse, and J.N. Memel eds., *Sensors*, vol 2, VCH, Weinheim, Germany, 1991, pp. 1024–1026.
7. K. Reardon and T. Scheper, in H.J. Rehm and G. Reeds eds., *Biotechnology*, 2nd ed., VCH, Weinheim, Germany, 1991, pp. 179–233.
8. D.L. Ristroph, C.M. Watteeuw, W.B. Armiger, and A.E. Humphrey, *J. Ferment. Technol.* **55**, 559–608 (1977).
9. A. Einsele, D.L. Ristroph, and A.E. Humphrey, *Biotechnol. Bioeng.* **20**, 1487–1492 (1978).
10. T. Scheper, K.-D. Anders, M. Busch, W. Müller, and K.F. Reardon, in J.A.M. de Bont, J. Visser, B. Mattiasson, and J. Tramper eds., *Physiology of Immobilized Cells*, Elsevier, Amsterdam, 1990, pp. 625–636.
11. J.-K. Li, E.C. Asali, and A.E. Humphrey, *Biotechnol. Prog.* **7**, 21–27 (1991).
12. J.J. Horvath, S.A. Glazier, and C.J. Spangler, *Biotechnol. Prog.* **9**, 666–670 (1993).
13. R. Chenina, A. Dochnahl, T. Huff, H.-G. Kuball, B. Lange, H. Anke, and T. Anke, *DECHEMA Monographien* **129**, 171–182 (1993).
14. S. Marose, C. Lindemann, and T. Scheper, *Biotechnol. Prog.* **14**, 63–74 (1998).

FOOD PROCESS ENGINEERING

HENRY SCHWARTZBERG
University of Massachusetts
Amherst, Massachusetts

KEY WORDS

Concentration
Drying
Expression
Extraction
Extrusion
Heat transfer
Membrane processes
Preservation
Separations
Thermal processing

OUTLINE

INTRODUCTION

Foods and beverages are processed to increase their safety and shelf life, to favorably modify their sensory and nutritional quality and composition, and to improve ease and reliability of preparation by consumers. Foods must be protected from contamination and stored and transported safely and without adversely affecting quality. Food process engineers employ engineering principles and methods to identify, select, design, develop, and "debug" processes, equipment, systems, and plants used to accomplish these tasks. Food process engineers also supervise the start-up and operation of food processing systems and plants; devise related control, maintenance, and sanitation systems; devise ways to eliminate hazards and minimize waste; and help develop new food products and ways to upgrade existing processes. In carrying out these tasks, engineers must account for the biological activity and chemical instability of foods and for the ability of microorganisms and pests to invade and grow in or on foods.

Many different types of food products are produced. The processes used are extremely diverse (1) and too numerous to cover in detail. Therefore, this article examines only the engineering aspects of a limited number of processes or operations. Many of these are used to preserve foods. Others are used to separate food components, and still others are used to change the shape, size, and texture of foods. Food processing frequently involves heat transfer. This article briefly examines such transfer.

Individual operations may produce effects that fall into several processing categories. Thus drying, which separates water from other food ingredients, is used mainly to preserve foods. Drying often changes the texture of foods and involves transfer of heat from hot air to moist solids. It also involves transfer of moisture (mass) within solids and from solids to air and transfers of momentum, which affects airflow.

HEAT TRANSFER IN FOOD PROCESSING

Food process engineers are frequently asked to model heat transfer, devise ways to control or improve such transfer, and calculate heating and cooling loads and sizes of equipment needed to provide desired amounts of heating or cooling. These tasks are accomplished by means of methods

widely employed by other engineers. However, in using these methods, food process engineers must account for the physical, heat transfer, and sensory properties of foods, and for their temperature sensitivity. Food components rapidly decompose in undesirable ways at high temperatures. The highest product temperature reached during food processing is probably 250 °C, the end-of-roast temperature for very darkly roasted coffee. Maximum tolerable temperatures for many foods are much lower. Higher processing media temperatures may be used (e.g., gas temperatures ranging up to 500 °C are used in roasting coffee). For economic reasons, foods are rarely cooled to temperatures lower than −40 °C.

Foods usually are either solid or complex composites (e.g., moist multicellular tissue, porous materials, stiff doughs) that act like solids. Heat transfer in such foods is modeled by means of partial differential equations based on Fourier's laws of conduction. Modeling also requires specification of boundary conditions based on surface temperatures, surface heat transfer coefficients, or surface rates of heat transfer. Solutions to partial differential equations for conductive heat transfer and many boundary conditions of practical interest can be found in the work of Carslaw and Jaeger (2) and other standard texts. Similar solutions for partial differential equations and boundary conditions describing diffusive mass transfer equations are provided by Crank (3). Many pairs of heat transfer and mass transfer equations and boundary conditions are formally equivalent. Therefore equation solutions for diffusive mass transfer often can be readily converted into solutions that apply for conductive heat transfer and vice versa.

Heat transfer calculations for solid foods may be complicated by a need to account for heats generated by ongoing biological processes, surface evaporation, anisotropic conduction, and internal vaporization and condensation. Often boundary conditions are spatially nonuniform, or they may vary with time during food processing. The thermal conductivities of foods may be difficult to predict. Nevertheless, most problems relating to heat transfer in solid foods can usually be solved fairly easily by means of computer-based numerical methods.

Food engineers also deal with heat transfer that is due to natural convection in liquid-filled containers, in pore spaces in stacks of heat-generating, respiring produce, and in the air that surrounds solid foods discharged from heating devices. In addition to widely encountered problems involving forced convection of heat in heat exchangers, evaporators, and condensers, they deal with convective heat exchange between streams of air or water and solid foods lined up one behind another in rows on shelves. These problems are often solved empirically. Food engineers working on aseptic preservation and ohmic heating processes are now attempting to precisely predict heat transfer to individual solids in groups of solids suspended in heated, flowing fluid.

Food engineers also deal with prediction of temperature changes during microwave and dielectric heating. In doing so they must account for attenuation and focusing of energy as electromagnetic radiation passes through foods, refraction and reflection of such radiation at food surfaces, energy distribution patterns in food-containing microwave cavities, and temperature-induced changes in the energy absorption characteristics of foods.

PRESERVATION OF FOODS

Drying, chilling, freezing, heating-induced inactivation of microorganisms and enzymes, application of bactericides, storage in controlled atmospheres, solute-induced water activity or pH reduction, and combinations of some of these processes are used to preserve foods.

Drying

Drying, the evaporative removal of water from moist solids or solidifiable solutions, may well be the earliest method of preserving foods and still serves that purpose today. It is also used to harden foods that have been shaped in a moist state, to reduce food weight, and volume, and to reduce moisture contents to facilitate processing.

Drying preserves foods because biological activity (e.g., microbial growth) radically slows or stops and chemical reactants become immobile when little water is present. Fresh fruits and vegetables usually contain 5–10 kg water/kg dry solids; stable, dried vegetables, 0.03–0.1 kg water/kg dry solids; and stable, sugar-rich dried fruits, 0.20–0.39 kg water/kg dry solids (4). The water content of freshly harvested grains is up to 0.22–0.43 kg water/kg dry solids; to provide safe storage, this must be reduced to 0.11–0.15 kg water/kg dry solids.

Hot Air Drying. Food drying most frequently is carried out by transferring heat to moist food particles from flowing hot air. This causes food moisture to evaporate and move away with the air. Consequently, the air moisture content (humidity) rises, and the air temperature drops proportionately. Relative humidity (RH) is the partial pressure of water in air divided by the vapor pressure of pure water at the same temperature. Water activity (a_w) is a measure of the chemical activity of water. It is also the equilibrium RH for a food at a given moisture content. Food moisture sorption isotherms, such as those shown for apples and potatoes in Figure 1, are constant-temperature plots of a_w versus a food's dry-basis moisture content, X. Moisture sorption isotherms shift downward moderately as temperature rises. The minimum X obtainable by drying is X_E, the sorption isotherm X at which a_w at the air temperature equals RH.

After start-up transients have decayed, food temperatures during drying depend on a_w at the food's surface. As long as the surface a_w is greater than 0.9, the food temperature will be no more than 1 °C higher than the air wet-bulb temperature; if air conditions are almost constant, the drying rate will be almost constant. Constant-rate drying does not last very long for foods. Surface a_w depends on X_S, X at the surface. As drying removes water from the surface, water diffusing to the surface from the food's interior partially replaces it. Because water diffuses slowly in foods, X_S very rapidly drops enough to drive a_w below 0.9. After this, a_w drops sharply as X_S decreases, the food temperature rises and progressively approaches the air tempera-

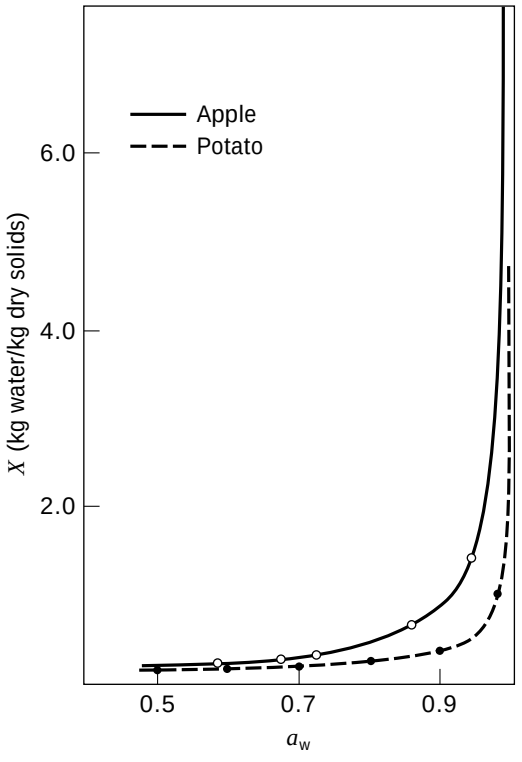


Figure 1. Moisture sorption isotherms for apples and potatoes.

ture, and the drying rate progressively decreases. Drying rates are strongly influenced by heat transfer and airflow rates as long as the air temperature is significantly higher than the food temperature. As food temperature approaches the air temperature, drying is more and more controlled by water diffusion in the food.

Near the end of drying, diffusion completely controls drying rates. The air temperature sets the food temperature. Hence, it affects diffusion and drying rates. Enough airflow must be provided to remove evaporated moisture fast enough to prevent X_E from rising excessively. Other than that, airflow no longer affects drying.

Drying Rates. Water diffusivity in foods often decreases markedly as drying proceeds. Drying also causes shrinkage. Theoretical analysis of water diffusion in drying foods is difficult because of these changes. Though special computational techniques can be used to deal with these complications, in most cases, drying rates are correlated empirically rather than theoretically predicted. After constant-rate drying stops, drying rates at constant air conditions are often proportional to $(X_A - X_E)^n$, where X_A is the average X and n is an empirical constant, which ranges between 0.7 and 2.4.

Drying air temperature and humidity affect food quality in addition to affecting drying rates and thermal efficiency. Therefore, air conditions are adjusted to minimize quality loss and to prevent in-process microbial growth. Dried products may not rehydrate well. This can be partially remedied by vapor-induced puffing, elicited by short-term use of high air temperatures late in the drying process.

Wet particles may stick together during drying. This property can be used to provide desired agglomeration. When sticking occurs and is unwelcome, particles are broken apart by passing them between rubber-faced rolls after drying.

Shallow-Bed Dryers. Foods are air-dried in many different types of equipment. In conveyor dryers (Fig. 2), moist food particles are uniformly spread as a shallow bed on a perforated belt and conveyed through a series of stages where heated air is blown through the bed and belt. Different air temperatures, humidities, and flow rates are often used in the stages. To improve uniformity of drying, upflow is used in some stages and downflow in others.

In conveyor-based drying of vegetables, 60–85% of the food's moisture content is removed in the first 1.5–2 h of drying. An additional 2–8 h usually is needed to complete the process. Air temperatures between 66 and 120 °C are used in early stages of drying. Lower air temperatures (32–52 °C) are used later, at the end of drying. Quite different conditions are used for some foods. Low temperatures are used during first stages of gelatin strip drying to prevent melting. Quick-cooking rice is dried in less than 15 min using temperatures up to 200 °C. Temperature around 175 °C may be used to puff diced potatoes at late stages of drying.

In tunnel dryers, hot air passes over shallow beds of food particles conveyed through a rectangular tunnel on racked trays mounted on carts. Airflows in the direction of cart movement or counter to that direction. Clearances between the carts and the tunnel are small, to minimize bypassing of air. Periodically, a freshly loaded cart is pushed into the tunnel and a cart containing dried material simultaneously leaves. Drying tunnels often are used in series. Cocurrent drying with high inlet-temperature air is followed by countercurrent drying with lower temperature air. Conveyor dryers are used instead of tunnel dryers wherever feasible because tray loading is burdensome and conveyor dryers permit automatic loading and discharging. In addition, drying conditions are more readily controlled in conveyor dryers.

Deep-Bed Drying. Bin dryers in which air passes through a deep bed of particles may be used to complete the drying of partly dried vegetables discharged from shallow-bed dryers. Grain is often directly dried in deep beds in bins, where air is introduced through a false floor or ducts near the floor (5). Ambient air may be used if its relative humidity is less than 0.55, or it may be heated to 11 °C above ambient. In such dryers, moisture desorption waves (Fig. 3) and accompanying temperature-change waves move upward through the grain bed. Most of the time, the exit air temperature equals the initial grain temperature and the exit air RH equals the initial grain a_w. Design engineers select airflow rates that provide drying rapid enough to prevent mold growth in moist region: 0.007 m^3 air/(m^3 grain s) may be used for relatively dry feed; 0.08 m^3 air/(m^3 grain s) may be needed for very moist feed (5).

In other grain dryers, air flows across descending beds of grain that are 0.2–0.4 m thick (5,6). In these cases, air-

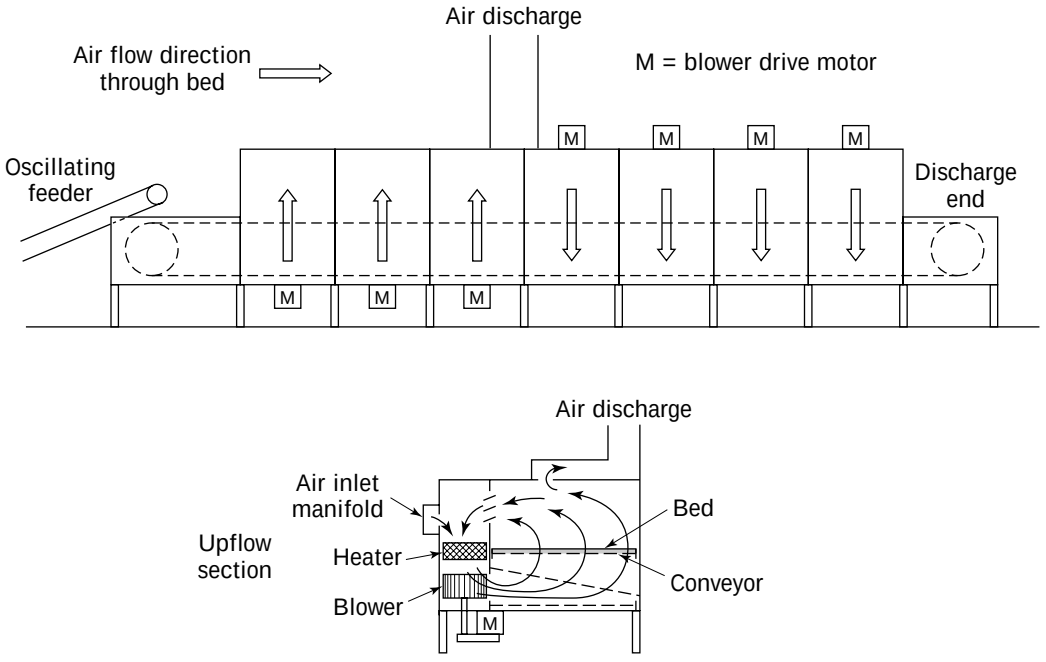


Figure 2. Conveyor dryer.

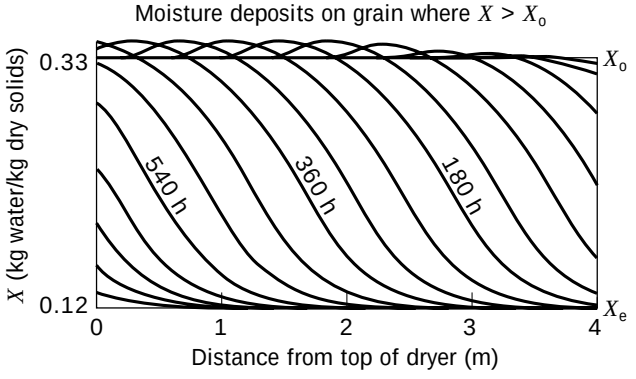


Figure 3. Moisture migration profiles for grain dryer (X vs. position at different times).

flow rates of 0.7–2.7 m^3 air/(m^3 grain s) are used. Inlet air temperatures up to 82 °C are used for animal feeds. Less than 54 °C is used for grain that is to be milled and processed, and less than 45 °C is used for seed grain. Grain is cooled with cool, dry air after such drying. Grain may be partly dried with high temperature air and then transferred to a deep bin where local moisture contents equilibrate without airflow. Final drying and cooling are accomplished by blowing cool, dry air through the grain.

Pasta Dryers. Special dryers are used to dry pasta. Freshly extruded pasta contains roughly 0.45 kg water/kg dry solids. That moisture content is reduced roughly 0.13 kg water/kg dry solids to set pasta's shape and preserve it. Long goods (e.g., spaghetti) are draped over moving rods and carried through the dryer system. Short goods (e.g., macaroni) are conveyed through dryers by simpler means. To prevent adjacent pieces of pasta from sticking together, the pieces are predried using 63 °C, 65% RH air. They leave the predryer and enter a large, final dryer where very humid air is used. Zones in which active drying occurs are followed by zones in which internal equilibration of moisture can take place. To prevent cracking (checking) due to stress buildup, relatively low temperatures (e.g., 50 °C) and very long drying times (up to 18 h) were and often still are used for long goods. Now, based on better modeling of moisture diffusion and stress development, temperatures up to 110 °C and drying times of 5 h are frequently used for long goods. Short goods are less susceptible to cracking and can now be dried in 2 h. Use of higher temperatures also reduces microbial growth.

Spray Dryers. Spray dryers are used to dry liquid foods and pumpable food slurries (7). Such foods are converted into drops with mean diameters ranging between 50 and 250 μm by nozzles or rotary atomizers and sprayed into a large chamber. The drops enter at high velocity but soon decelerate. Hot air is blown in and flows in the same direction as the descending drops. One type of spray dryer is shown in Figure 4. Inlet air temperatures as high as 270 °C may be used; often they range between 160 and 200 °C. Outlet air temperatures usually range between 80 and 110 °C. Drop temperatures quickly rise to the wet-bulb temperature of the inlet air and then progressively approach the local air temperature as drying proceeds.

Because air cools as it moves through the dryer, cocurrent air-drop flow prevents drops from overheating. Dried particles leave the dryer with temperatures 10–20 °C below the exit air temperature. Large particles drop out of the airstream. Smaller particles are collected in cyclones.

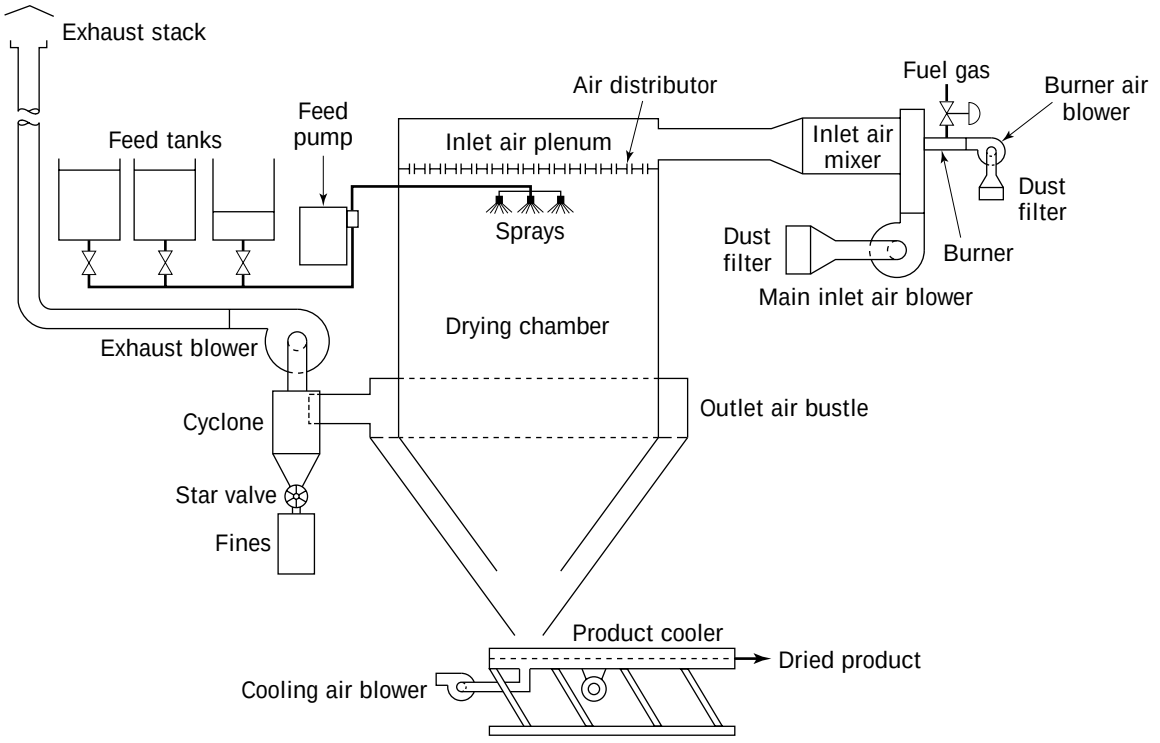


Figure 4. Spray dryer in which large particles drop out of the drying air by gravity and small particles are collected in a cyclone.

Drying times range from 5 s for small drops to 50 s for large drops. Spray-drying involves exchanges of heat, momentum, and mass between the air and the drops, nonlinear diffusion of moisture within drops, and changes in drop size and shape. Complex computer programs have been developed to account for these processes (8). Spray-dried foods are rapidly cooled after drying by sucking cool, dry air into solids discharge lines or by drawing cool, dry air through a screen-supported product bed on a vibrating conveyor.

Freeze-Dryers. Freeze-dryers are used to dry beverage extracts, shrimp, soup ingredients, and military and camper rations. In these dryers, moisture sublimes from frozen food in a vacuum chamber and condenses on very cold refrigerated coils. Heat transfers radiantly from heated panels to the surface of the food and then passes by conduction through a porous dry food layer to a sublimation interface. The interface recedes into the food as drying proceeds (9,10). Vapor created by sublimation escapes from the food through pores produced by prior sublimation.

Temperatures at the product surface and sublimation interface depend on the radiant panel temperature and on the thermal conductivity, vapor permeability, and thickness of the dry layer. Depending on whether heat conduction or vapor permeation controls the drying rate, radiant panel temperatures are programmed either to prevent product surface overheating or to prevent melting at the sublimation interface (11). Small, frozen particles are sometimes used to circumvent drying rate retardation caused by low dried-layer permeability.

Noncondensibles must be removed before and during freeze-drying. If adequate noncondensibles removal is achieved, chamber pressures in the dryer approach ice's vapor pressure at the condenser temperature. Condenser temperatures of -40 °C are frequently used, resulting in absolute pressures around 100 μm Hg. Local chamber pressures in plant-scale freeze-dryers are effected by pressure drops due to high-velocity vapor flows between dryer trays and heating panels (12).

Other Dryers. Foods are also dried in air-lift systems, in fluidized beds, on the surface of rotating, steam-heated drums, by contact with hot air in rotating, hollow cylinders of various types, by passage under or through arrays of high-velocity air jets, and by exposure to the sun in hot dry regions. Sun-dried foods may be treated with SO_2 to prevent browning and fumigated to prevent insect infestation.

a_w Reduction

Microorganism growth can be prevented by sufficiently reducing a_w, the chemical activity of water (13). Thus, *Clostridium botulinum* will not grow at $a_w < 0.94$, most bacteria will not grow at $a_w < 0.905$, most yeasts at $a_w < 0.88$, and most molds at $a_w < 0.8$; even osmiophilic molds and yeasts will not grow at $a_w < 0.6$. Drying preserves foods in part by reducing a_w. Moderate a_w reduction can be used in combination with other steps (e.g., pH reduction, refrigeration or freezing, and use of antimicrobial agents and fungistats to provide storage stability). Water activity values are also reduced by evaporation or by adding solutes

that reduce the mole fraction of water in solution. Different combinations of salting and drying are used to preserve cod. Lightly salted cod is stable when its moisture content is 12.3%; heavily salted cod is stable with a moisture content of 52.4%.

Chilling

Chilled, nonfrozen storage is widely used to extend the storage life of fresh produce, meats, fish, and dairy products (14). Recommended temperatures and humidities and approximate storage lives for the commercial storage of fruits and vegetables are listed in U.S. government publications (15) and by the American Society of Heating, Refrigeration, and Air-Conditioning Engineers (ASHRAE) (16). These sources also provide heat capacities and heats of respiration for many of the commodities listed. Typical rates of heat influx are also given to permit computation of refrigeration loads for storage facilities. Storage temperatures close to 0 °C and humidities ranging between 85 and 95% are frequently recommended; cold-sensitive produce is stored at temperatures between 7 and 13 °C. Storage life at these conditions varies widely, ranging from as short as 2 days to as long as 12 months; storage life is less than 1 month for roughly 60% of the foods surveyed. Design and operating considerations for food storage refrigeration systems are treated by ASHRAE (16). Prechilling before storage is carried out by blowing chilled air on products, by immersing products in chilled water containing chlorine or an approved phenol compound or spraying similar chilled water on them, or by mixing products with ice or slush. Leafy products that conduct heat poorly are often cooled by evaporating a small part of their water content in a vacuum chamber. To minimize weight loss, such products may be prewetted before vacuum cooling.

Controlled Atmospheres. Chilling is combined with use of controlled atmospheres that contain as little as 2–5% O_2 and up to 5% CO_2 to increase the storage life of apples, pears, and cabbage (17). CO_2-rich controlled atmospheres are also used during long-distance shipment of meat. The rooms or shipping containers used must be sealed and airtight. Exceeding tolerable CO_2 levels or falling below tolerable O_2 levels damages apples. These levels significantly differ for different varieties of apples. Respiration can be used to reduce O_2 and raise CO_2 levels. Any excess CO_2 produced is absorbed by hydrated lime, ethanolamine (which can be regenerated by heating), water, or less frequently activated charcoal or molecular sieves. To more rapidly adjust O_2 and CO_2 levels, catalytic or noncatalytic combustion of fuel and flushing with nitrogen are used. Multilocale O_2 and CO_2 sensors and loggers have been used to ensure that improper local environments do not develop in controlled-atmosphere storage rooms. The use of selectively permeable wrapping can generate modified atmospheres in packages and improve shelf life during food distribution. Vacuum packing is also used.

Freezing

Freezing, the cooling-induced conversion of most of a food's water content into ice, is used to preserve virtually all types of moist food (18,19). It is also used as a step in freeze-concentration and freeze-drying and to produce ice cream and sherbet. Freezing preserves foods by reducing rate constants for chemical reactions, radically reducing reactant mobility, and reducing a_w. The water activity of a frozen food depends solely on its temperature and is independent of the nature of the food involved, as long as it is suitably moist; $a_w = 0.908$ at -10 °C, 0.824 at -20 °C, and 0.748 at -30 °C.

Solid foods are frequently frozen by exposing them to suitably cold streams of refrigerated air. This may be done by passing cold air through a shallow, fluidized bed of small particles or by blowing cold air over larger pieces of food carried on perforated conveyors or mounted on trays in racks. Meat carcasses are frozen by passing cold air around them while they hang by the hind legs on hooks. Foods have also been frozen by direct contacting with brine–ice mixtures or by contact with an inert, evaporating refrigerant or cryogenic gas. Packaged foods are often frozen by compressing them between refrigerated plates.

Ice cream and sherbets are partially frozen in scraped-surface freezers. Finish freezing (hardening) is carried out in plate freezers. Ice cream pops are frozen in molds conveyed through baths of refrigerated brine. Descriptions of most types of commercial freezing equipment, their specifications, and operating conditions are provided by Postolski and Gruda (20).

Most water in moist foods will freeze, but part (0.17–0.31 kg water/kg solids, depending on the particular food) is bound to food solutes or solids in nonsolvent form and will not freeze, even at very low temperatures. Food solutes depress food freezing points. Equilibrium initial freezing points for moist foods range between -0.4 and -3.2 °C, and usually lie between -1 and -2 °C. Foods frequently supercool 5 or 6 centrigrade degrees below their equilibrium initial freezing point before ice nucleates and freezing starts. Then the temperature rapidly rises to the initial freezing point. As freezing proceeds, solutes concentrate in residual unfrozen solution, further depressing the freezing point. Thus, food freezing occurs over a range of temperatures rather than at a single temperature.

Fractional conversions of freezable water to ice, fractions of water that remain unfrozen, and thermal properties of partly frozen foods can be predicted with the aid of the freezing point depression equation (21). These predicted values can be used in calculating freezing times for foods and temperature profiles for foods undergoing slow to moderately rapid freezing (i.e., freezing in which local phase equilibrium can be assumed without serious error) (22–24). Freezing involves incorporation of water in ice crystals and diffusion of water to active growth surfaces on ice crystals. These processes occur at rates slow enough to cause significant local deviations from phase equilibrium during very fast freezing. They also largely determine ice crystal size and shape (but not the total amount of ice formed) during slower freezing.

Ice usually grows as dendrites (treelike crystals) when foods freeze. Spacings between ice dendrites tend to be inversely proportional to the square root of the initial freezing rate. Specific surface areas and surface free energies of dendrites are much larger than those of rounded crystals.

Therefore ice dendrites tend to convert into rounded ice. Conversion occurs rapidly in agitated slushes but is quite slow in unagitated frozen food, particularly food that is kept well below its initial freezing point. Specific surface energies are larger for small crystals than for large crystals. Therefore small ice crystals tend to sacrificially melt and cause growth of larger ice crystals. If storage temperatures are not low enough or are allowed to cycle excessively, ice crystal size will increase appreciably after long storage, affecting texture adversely.

When unblanched fruits and vegetables are frozen, ice initially nucleates outside cells. If freezing is slow, much water transfers outward from cell interiors and forms extracellular ice before nucleation occurs inside cells. This affects texture adversely. Good practice calls for rapid freezing. Then, intracellular nucleation occurs quickly, most water freezes inside cells, and adverse textural effects are minimized. Preferential growth of ice in extracellular space usually does not occur during slow freezing of blanched vegetables. Fast freezing of meat and fish helps minimize losses of water-holding power and decreases drip following thawing.

Thermal Processing

Sterilization, pasteurization, and blanching are processes in which foods are heated and held at suitable temperatures long enough to kill the microorganisms and inactivate the enzymes they contain. Commercial sterilization is designed to kill substantially all microorganisms that can grow or spores that can germinate at storage conditions. Because sterilized foods are packed in sealed containers that contain virtually no air, anaerobic organisms are of greatest concern. Destruction of *C. botulinum* spores is often used to evaluate adequacy of sterilization for low-acid foods (i.e., foods with pH $>$ 4.5). Fractional survival levels must be less than 1×10^{-12} for such spores. Adequacy of sterilization for low-acid foods may also be based on destruction of a hardier putrefactive anaerobe called (PA) 3679. Adequacy of sterilization for foods with pH values less than 4.5 is based on destruction of one or another of the facultative anaerobes, *Bacillus coagulans, B. mascerans*, or *B. polymaxa*. Fractional survival levels for these organisms should be less than 1×10^{-5} or 1×10^{-6}. Sterilization of low-acid foods involves maximum temperatures in the 116–149 °C range. Maximum temperatures up to 110 °C are used for more acidic foods.

Sterilization of canned products is carried out batchwise by steam heating in vertical or horizontal retorts. The retorts must be vented for at least 5 min with steam flow to remove air before steam pressurization and starting a heating cycle. Foods in jars are often sterilized in air-pressurized hot water in retorts. Sterilization based on steam heating is also carried out continuously in retorts with sections separated by pressure locks, in rotary sterilizers (where a rotor causes cans to move through the unit along spiral guide rails), and in hydrostatic sterilizers (where entering and leaving cans are conveyed through a long, steam-heated zone confined between vertical water legs that balance the steam pressure). Figure 5 shows an example of a hydrostatic sterilizer.

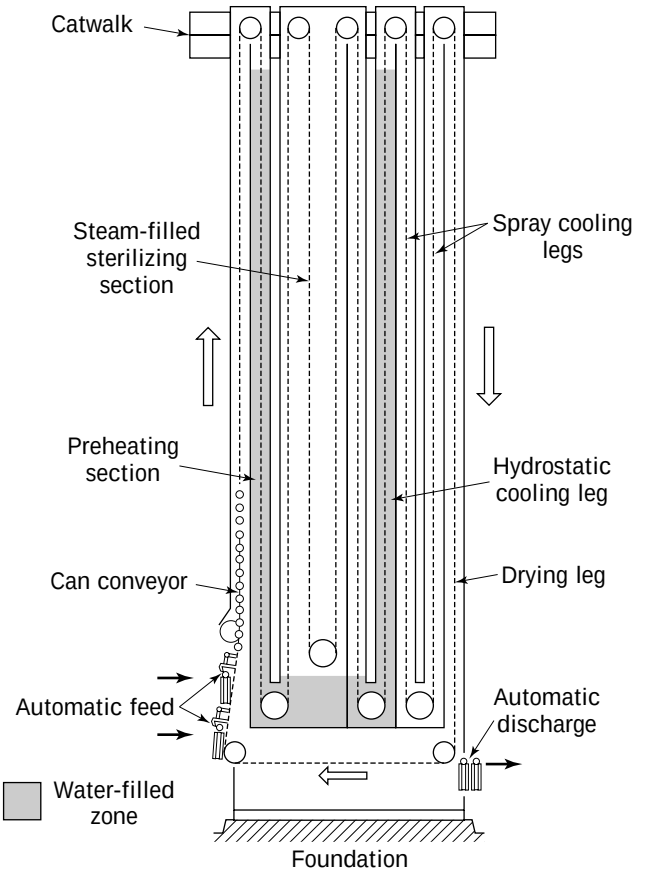


Figure 5. Hydrostatic sterilizer. Reprinted with permission of the Chemetron Corporation.

Headspace is left when food containers are filled. The container walls or tops have elements that flex. These elements and headspace prevent thermal expansion from damaging containers during sterilization. Headspace air and dissolved gases are partly removed from cans, jars, and their contents before sterilization by steam flushing or vacuum evacuation before capping. Thermal exhaustion (i.e., preheating material to expel dissolved gases) is sometimes used instead of or in conjunction with these processes.

Thermally processed products are cooled under air pressure with chlorinated cold water immediately after heating. Air pressure is used to prevent bulging of containers or unseating of caps caused by excess residual pressures inside containers. Cooling is continued until products are just warm enough to fully evaporate water clinging to the container surfaces. Sterilization, container filling, and capping procedures for specific products differ in detail but are described in standard references (25).

Pasteurization. Pasteurization, the thermal treatment at temperatures less than 100 °C, is used to inactivate enzymes, virtually all pathogenic microbes, and many food spoilage organisms. It is used in combination with protective packaging and other treatments (e.g., chilling, pH, and a_w reduction) that greatly retard subsequent microbial

growth. The target organism or enzyme on which adequate inactivation is based depends on the product involved. For milk it is *Coxiella burnetti*, which causes Q fever. For wine or beer it is wild yeasts. For high-acid fruits it is molds and yeasts. For citrus juices it is pectinesterase. Inactivation of pectinesterase also provides more than adequate microbial inactivation. Acceptable holding time-temperature combinations may be specified by government regulations. In the United States, at least 30 min at 62.8 °C or at least 15 s at 71.7 °C is required for pasteurization of milk. Most frequently, the latter conditions are used. Other, equally effective time–temperature combinations can be approved. Ultrahigh temperature pasteurization carried out at 88.3 °C for 1 s or 100 °C for 0.01 s is used sometimes. Liquid mixes used to produce ice cream and frozen desserts are pasteurized for at least 24 s at 79.4 °C.

Milk is pasteurized before it is put in containers. This is done by regeneratively and directly heating milk to a specified temperature in a plate heat exchanger, holding it at that temperature for a required time and then cooling it in the same exchanger. Integrated systems that also provide cream removal by centrifugation and homogenization are often used (see Fig. 6). Other arrangements for pasteurizing milk are described by Kessler (26), who also provides correlations for overall heat transfer coefficients versus fluid velocities or Reynolds numbers for different plate–heat exchanger flow arrangements. Simpler plate–heat exchanger setups are used to pasteurize fruit juices.

Other foods and beverages are pasteurized in containers. Products in glass jars are conveyed on metal-link belts, heated gradually, held at temperature, and cooled gradually in stages to prevent thermal shock to the glass. Long hot water baths or sprays followed by similar cold water baths or sprays are used for jarred pickles, canned fruit, and bottled beer. Heating by passage through steam at atmospheric pressure is sometimes used in pasteurizing products packed in glass.

Blanching. Blanching, heating in water somewhat below it boiling point or in steam at atmospheric pressure, is used to inactivate enzymes that adversely affect food storability and quality. Sterilization and pasteurization processes usually inactivate enzymes as well as microorganisms; blanching is used primarily before drying and freezing for vegetables and some fruits. Blanching also causes expulsion of gases from foods, softens foods, facilitates cutting and peeling, sets some colors, and provides some microbial inactivation. Unfortunately it also leaches nutrients from foods, which in turn increases the biological oxygen demand (BOD) of effluent blanch water. Less leaching occurs when steam blanching is used and still less when very humid hot air is used.

Aseptic Processing. Aseptic processing involves heating liquids, semiliquid products, or liquids containing small particles to sterilization or pasteurization temperatures in heat exchangers or scraped-surface heaters or by injection or infusion of culinary steam and then passing the products through a holding tube. They are then cooled in heat exchangers or, for products heated by steam addition, by vacuum evaporation, and put in sterilized containers. The containers may be cans, jars, or drums that have been heated with superheated culinary steam. They are capped with steam-sterilized covers in chambers filled with super-

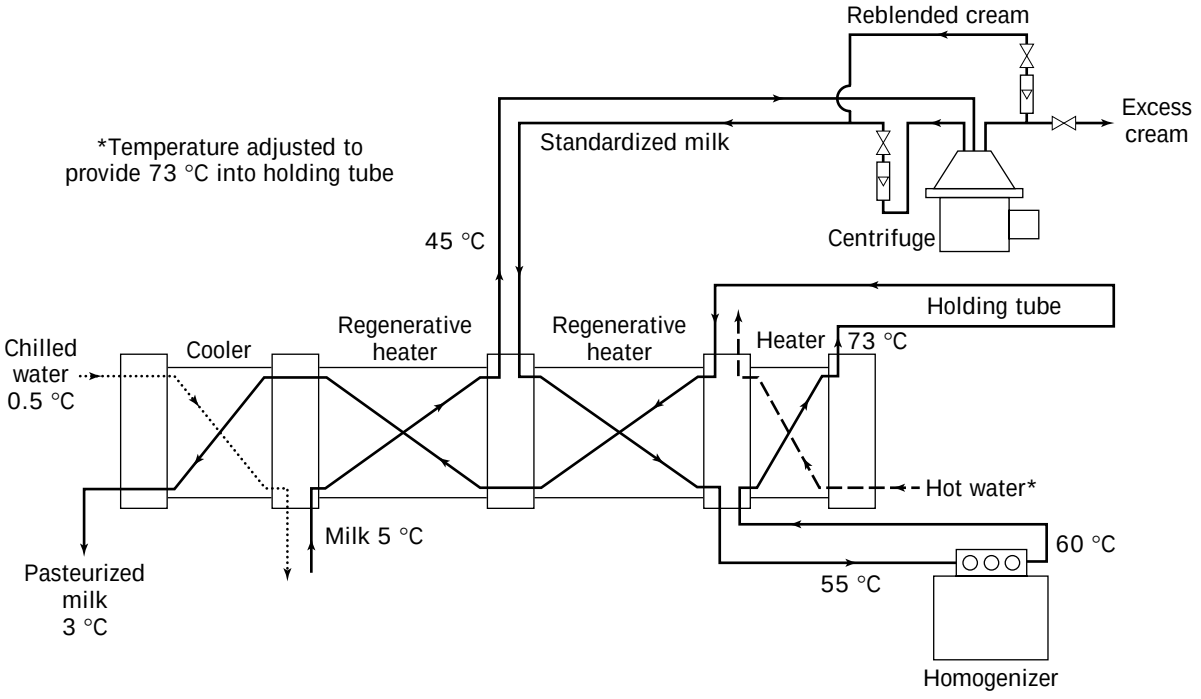


Figure 6. System for standardizing, homogenizing, and pasteurizing milk. Reprinted with permission of Kessler Verlag.

heated culinary steam. Continuously formed laminated containers or plastic cups are also used. Before they are filled, they are bathed in H_2O_2 and dried. These containers are heat-sealed by fusing together narrow bands of a meltable plastic. Heated product temperatures range from slightly less than 100 °C for acidic products to 149 °C for low-acid products.

Inactivation Kinetics. Thermally based microbial destruction and enzyme inactivation are treated as first-order reactions; that is, destruction or inactivation rates are proportional to the residual concentrations of living microbes or active enzymes involved. Inactivation rates are strongly affected by pH and food composition. If $k(T_p)$ is the destruction rate constant for a particular microbe or enzyme in a food expressed as a function of product temperature T_p and T_p is a function of time t, we can write

$$\ln\left(\frac{N}{N_o}\right) = -\int_{t=0}^{t} k(T_p)dt \qquad (1)$$

where N is the residual number of microbes per unit volume or enzyme concentration at t and the spot where T_p and $k(T_p)$ are evaluated and N_o is the initial value of N. Figure 7 depicts how $k(T_p)$, T_p, and N/N_o change during thermally based sterilization. To achieve acceptable destruction or inactivation, the highest N/N_o at any locale in a product for the most critical microbe or enzyme in question must be equal to or less than a required value. For low-acid foods, N/N_o must be less than 10^{-12} for *C. botulinum* spores. For nonmobile canned foods, the greatest N/N_o is assumed to occur at the center of a can's contents. Convection occurs in foods containing free liquid, and the greatest N/N_o is assumed to occur at the slowest heating point. This usually lies along a can's axis and near its bottom, but its locus should be experimentally determined.

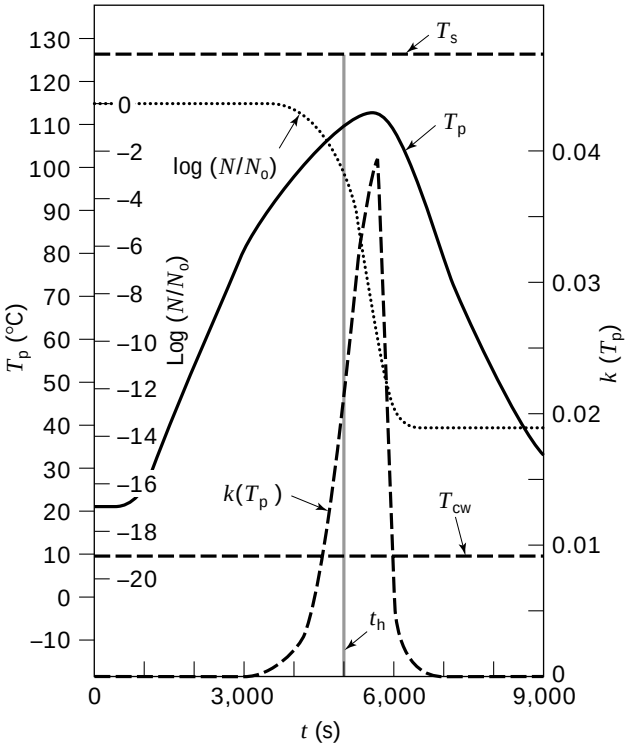


Figure 7. T_p, $k(T_p)$, and N/N_o versus t during sterilization of a canned solid food by thermal processing: T_s and T_{cw} are, respectively, the steam temperature and the cooling water temperature in the retort, T_p is the temperature at the center of the food, and t_h is the time at which heating ends and cooling begins. Reprinted with permission of John Wiley & Sons.

In the food industry, adequacy of thermal processing (27) is often evaluated by determining whether inactivation is equal to or greater than that provided by treatment for a specified time at standard temperature, 121.1 °C (250 °F). Temperature-induced changes in thermal inactivation rates are accounted for in terms of Z, the temperature reduction that will cause $k(T_p)$ to decrease by a factor of 10. A standard Z of 10 °C (18 °F) is often used for *C. botulinum*. Depending on the food and corresponding Z involved, acceptable treatment of low-acid foods should provide *C. botulinum* destruction equal to that provided by 1.2–2.4 min isothermal exposure at 121.1 °C.

In evaluating adequacy of thermal processing, one must account for changes in product temperature that occur during both heating and cooling. Product temperatures predicted from heat transfer analysis are often used in trial-and-error fashion to determine suitable heating and cooling times. These calculations are useful for estimating required processing time, but determination of adequacy of processing is usually based on measured product temperature–time data. This is particularly true for new products, where adequacy of sterilization must be further verified by microbial inactivation tests in inoculated packs.

Formula methods and nomographs based on characteristics of semilog plots of unaccomplished temperature change versus time were once widely used to determine adequacy of sterilization and needed thermal processing time. Now, simple computer programs are used for the same purpose.

Sterilization causes partial destruction of some nutrients and often adversely affects food flavor and color. Destructions of nutrients and some colors and flavors are first-order or pseudo–first-order processes for which kinetic parameters are sometimes available (28). New flavors, colors, and textures that affect product quality also develop by processes with more complex kinetics. Effects of product temperature on nutrient, flavor, and color destruction often are framed in terms of activation energies and the Arrhenius equation instead of Z. In evaluating nutrient destruction and quality reduction, fractional retention of the nutrient or attribute being evaluated is averaged over the whole volume of the container. Sterilization methods and temperature–time conditions that provide adequate pathogen or spoilage agent inactivation but minimize adverse effects should be used. Rates for reactions that destroy nutrients and adversely affect sensory properties increase less sharply than microbial destruction rates as temperature increases. Therefore high-temperature, short-time (HTST) processing often improves nutrient and flavor retention. Heat-resistant enzymes may not be adequately inactivated when microbial destruction is

adequate if very high temperatures are used for very short times.

Outer regions of containers are often overprocessed and suffer excessive quality degradation when the points that heat most slowly are just adequately processed. Overprocessing can be reduced by providing more uniform heating (e.g., by using aseptic processing and microwave and ohmic heating). Other preservation processes that avoid or greatly limit heating (e.g., radiation preservation, microfiltration, and high-pressure treatment) have been extensively investigated or, in some cases, are already in use. Their advantages and limitations were examined in a recent review (29).

Concentration Processes

Evaporators. Water and solvents can be removed from liquid foods by means of evaporators. In multiple-effect evaporators, the most commonly used type, vapor generated by steam or vapor heating in an effect provides heating that causes evaporation in a subsequent effect operating at a lower pressure. In N-effect evaporators, roughly N kg of vapor is generated for each kilogram of steam used. In some cases, a steam ejector is used to compress part of the vapor generated in a second or third effect, and the compressed vapor and motive steam are used to heat the first effect, providing more than N kg of vapor per kilogram of steam used (30). Vapor recompression may also be carried out by large centrifugal compressors. Electrical energy is used to drive the compressor; virtually no steam is used directly. Evaporation, particularly multiple-effect and vapor recompression evaporation, is more thermally efficient than drying. Therefore evaporators are often used to concentrate liquid foods before drying. Evaporation is also used to concentrate juices, syrups, sauces, fermentation residues, stillage bottoms, and absorption refrigeration solutions. It is also used to induce crystallization and produce glassy supersaturated sugar solutions used to make hard candies.

Vacuum evaporation, often combined with single-pass operation in each effect to provide low product holdup, is used to minimize thermal damage to foods (31). In single-pass evaporators, steam condenses on the outer walls of vertical heat transfer tubes 6–12 m long mounted in a steam chest. Water evaporates from liquid flowing in film form down the inner walls of the tubes. Concentrated liquid and vapor leave through the bottom of the tubes. Entrained liquid is separated from the vapor in a cyclone. A single-pass, multieffect system is depicted in Figure 8. High-velocity recirculation in liquid-filled tubes is used to improve heat transfer in evaporators used to concentrate viscous materials such as tomato paste and gelatin and pectin solutions. Liquid superheats as it passes through tubes in such evaporators and flashes as it leaves the tubes. Cyclones are used to separate entrained liquid from the vapor produced by flashing. Scraped-surface, thin-film evaporators are used for very viscous products (e.g., hard candy melts) and for heat-sensitive products.

Natural convection and boiling cause circulation of liquid in evaporators containing short, steam-heated tubes mounted in a wide-diameter assembly (calandria). Vapor separates from the foaming, boiling liquid in a large, vapor-filled chamber above the calandria. Heat transfer coefficients in these evaporators peak at a certain liquid level in the calandria. Multieffect systems containing such evaporators are frequently used to concentrate sugar solutions in sugar refineries. Part of the vapor produced may be withdrawn from effects and used for process heating elsewhere in the refinery.

Most food aromas have high relative volatility with respect to water and almost completely vaporize during initial stages of evaporation. Such aromas can be recovered in concentrated form from first-stage vapor streams and used for add-back purposes.

Membrane Processes. Pressure-driven flow through permselective membranes is used to gently remove water or aqueous solutions of low molecular weight solutes from liquid food products. Reverse osmosis (i.e., selective removal of water) is used to raise the solids content of maple sap from roughly 2% to 10–12%, to recover water and solutes from waste processing streams, and to concentrate dilute aqueous caffeine extracts. Ultrafiltration (selective removal of water and low molecular weight solutes) is used to concentrate proteins in milk, milk whey, and soy whey and to recover protein from potato fruit water. Protein-rich concentrates obtained from heated milk by ultrafiltration are used to make cheese with reduced loss of milk solutes and reduced use of rennet and starter cultures. Diafiltration (ultrafiltration accompanied by water addition) is used to de-ash gelatin. Pervaporation has been used experimentally to produce aroma-rich concentrates by selectively evaporating water through permselective membranes. Mass balances for retained species and equations governing permeation rates are used in designing membrane-based processes. Design calculations must not neglect the effects of concentration polarization on permeation rates and selectivity.

Freeze Concentration. Freezing is sometimes used to gently and selectively remove water from fruit juices, beverage extracts, beer, and wine. Scraped-surface coolers partially freeze the solution being processed, producing small ice crystals. Ostwald ripening (i.e., growth of large crystals induced by sacrificial melting of smaller crystals) is then used to produce large ice crystals, which are separated from residual solution in wash columns (32) or centrifuges. Thermal efficiency and upper concentration limits have been increased, and costs reduced, by means of countercurrent freeze concentration (33), but the use of freeze concentration remains limited because of its high cost.

Solid–Liquid Extraction

Solid–liquid extraction is used to recover sugar from sugar beets and sugarcane, oil from oilseeds, juice solutes from fruits, beverage extracts from coffee and tea, protein-rich materials from soybeans, fermentable solutes from grains, and hydrocolloids, gums, and natural food colors from various sources. It is also used to remove unwanted constituents (e.g., caffeine from coffee, bitter compounds from olives, cyanogenic compounds from cassava, sarcoplasmic

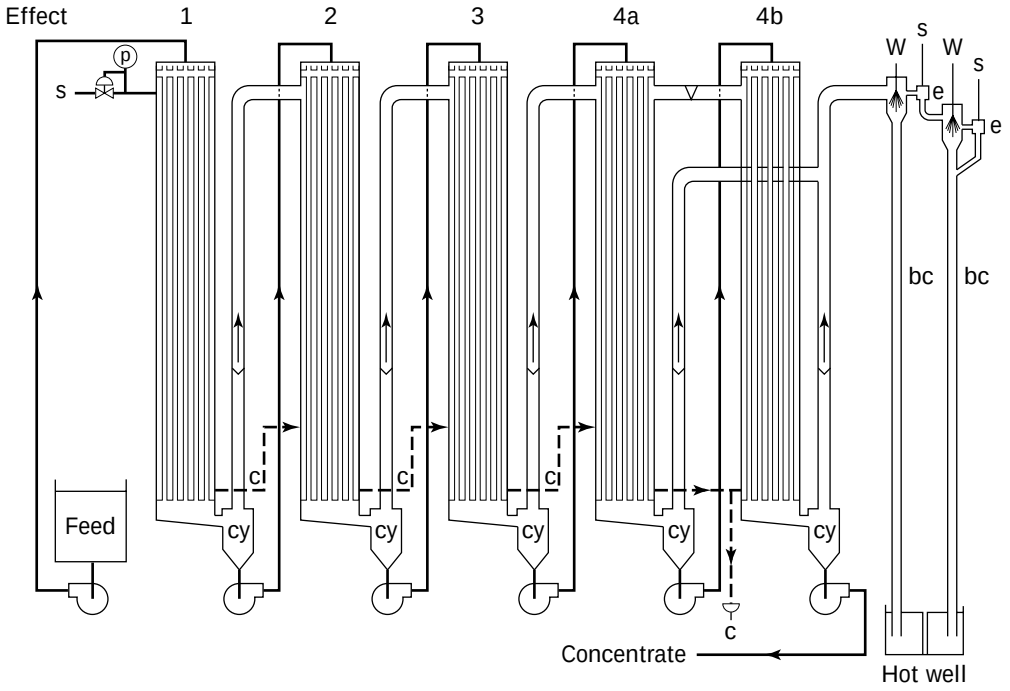


Figure 8. Single-pass, falling-film, multieffect evaporator. The fourth and fifth stages in terms of liquid flow together make up the fourth effect in terms of vapor utilization. bc, barometric condenser; e, ejector; cy, cyclone; s, steam; w, cold water; c, condensate; v, vapor.

protein, and undesired flavors from minced fish, and salt from pickles) (34,35). Extraction also occurs when flavoring agents leach out of charred barrels during the aging of wines and distilled spirits, when solutes leach out of vegetables and fruits during blanching and fluming, and when plasticizers leach out of plastic containers into liquid foods and beverages.

Vegetable or animal matter from which solutes are to be extracted is cut or otherwise divided into small particles. Oilseed grits are flaked or extruded as steam-puffed cylindrical particles. This coalesces isolated oil bodies and creates fissures or pores that allow solvent or extract to reach the oil. During extraction, solutes diffuse out of the solid through internal liquid-filled paths and pass into solvent or extract that contracts the solid. Dry materials (e.g., coffee and tea) must imbibe solvent or extract before diffusion can proceed. In rare cases (e.g., extraction of vanilla from vanilla beans with alcohol), imbibition is very slow and is a rate-controlling step. Plasma membranes are denatured to facilitate water-based extraction of solutes from cellular material. Then, pores in cell walls become the main diffusion bottleneck. In-particle diffusivities for water-soluble extractibles are 0.1–0.5 times as large as in water alone. Required extraction times are inversely proportional to the in-solid solute diffusivity and proportional to the particle size squared.

The extract carries away the solute, following a contacting path that promotes high solute recovery. Extraction efficiency and extract concentration depend, in part, on the stripping factor, the solute distribution coefficient multiplied by the ratio of extract to moist solids. High concentration extracts can be obtained by using a low stripping factor (i.e., low solvent flow rates or small amounts of solvent), but efficient extraction can be obtained only if the stripping factor is greater than 1.0. Stripping factors of 1.2–1.4 are often used in water-based extractions. Higher stripping factors (e.g., 2.4) are used for hexane-based extraction of oil from oilseeds.

Small-sized particles favor fast extraction but cause excessive flow pressure drop and displacement instability. Large particles extract slowly. Particle thicknesses or diameters in the range of 3–5 mm are a good compromise. Where possible, particle shapes that prevent blinding of surfaces by interparticle contact are used.

Commercial extraction is often carried out in very large, continuous or semicontinuous countercurrent extractors where extract percolates downward through a rising bed of solids (see Fig. 9) or in units where a bed of solids moves horizontally while extract repeatedly percolates through it in a flow arrangement that effectively provides countercurrent contact. In other cases, a cyclically loaded and discharged set of interconnected columns is used to provide nearly countercurrent contacting. Most of these processes can be analyzed and designed by means of solutions of partial differential equations describing diffusive mass transfer in solids (3,34) or mass transfer analogs of partial differential equations describing conductive heat transfer in solids (2). Solutes are sometimes produced from insoluble precursors by reactions whose kinetics must be accounted for in designing the extraction process. Extraction processes involving foods often are subject to imperfect contacting (e.g., channeling, unsteady displacement, and axial

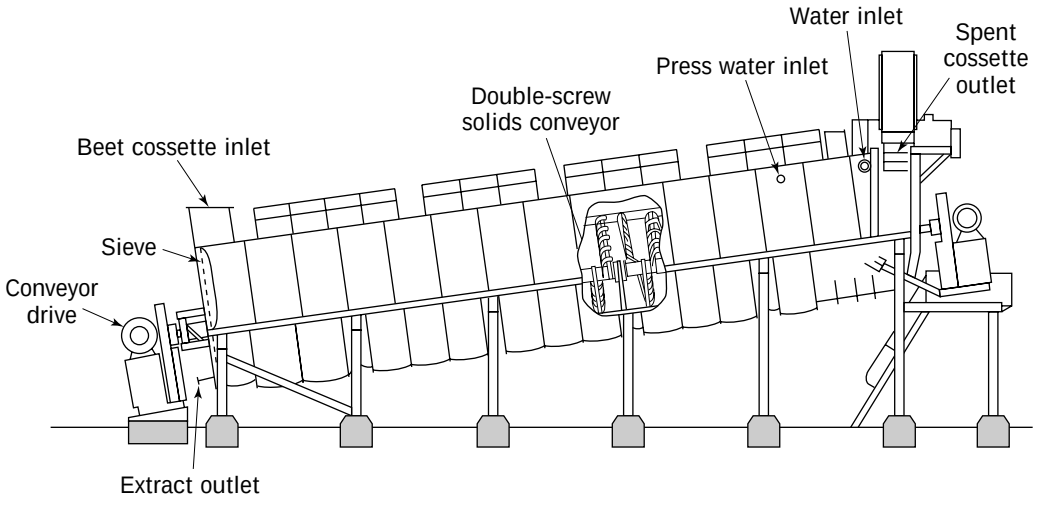


Figure 9. DDS extractor, a type often used to extract sugar from sugar beets. Reprinted with permission of De Danske Sukkerfabrikker.

dispersion). These imperfections must be accounted for in design calculations.

Supercritical carbon dioxide is now used as a solvent in decaffeinating coffee and tea and in recovering solutes from hops. Supercritical carbon dioxide is a particularly attractive solvent because it is nontoxic and can be readily removed from foods. Further, its dissolving capacity and selectivity for solutes such as caffeine can be adjusted by adjusting temperature and pressure. Figure 10 shows a schematic of a large supercritical extraction system used to decaffeinate green coffee beans.

Solid–Liquid Separations

Filtration. Filtration is used to remove solid particles from or to clarify juices, extracts, vegetable and fish oils, fermented beverages, recirculated, cooking oil, flume water, milk, and soy milk. It is also used to separate potato starch from potato fruit water, high-melting fats from vegetable oils in fractionation processes, crystals from mother liquors, and chemically precipitated impurities from sugar juice.

The engineering principles that govern filtrations involving foods are very similar to those that apply to the filtration of chemicals (36), but superior sanitation must be maintained. A fluid that contains particles is passed through a cloth or other porous medium that retains particles. The particles deposit on the medium, forming a porous cake that thickens as deposition continues. As long as total pressure drop remains constant, total flow resistance increases linearly as the cake thickens. Depending on the feed pumping system used, flow resistance increases cause either decreases in flow rates and solids deposition rates, increased fluid pressure drop, or a combined increase in pressure drop and decrease in flow rate. Filtration is stopped when these changes become excessive, and the filter cake is manually or mechanically removed.

Extents of these changes and optimal scheduling of cake removal can be predicted by means of calculations that relate cake depth and flow pressure drop to the initial solids content of the fluid, the amount of filtrate discharged, and the experimentally determined relationships between applied pressure, cake density, specific filtration resistance, and applied pressure. Some food-based filter cakes are highly compressible, and their specific filtration resistance increases markedly as pressure increases. Filter aids (inert, highly porous solids) are sometimes added to filter feeds to reduce cake filtration resistance and compressibility and to improve solids retention. Cake filtration resistance tends to be inversely proportional to the particle size squared. Therefore chemical agents that cause particles to flocculate (stick together in clumps) are often added to filter feeds.

Deep porous media that capture and retain particles within their pores are used for some feeds. Media flow resistance increases as filtration proceeds. Consequently flow pressure drops increase and/or flow rates decrease. The medium is replaced or cleaned by backwashing when these changes become excessive. Unlike cake filtration, where total filtration resistances at constant pressure increase linearly as solids removal progresses, filtration resistance for in-media solids capture is more than linearly proportional to the amount of solid removed. Therefore in-media capture is used only for fluids with light solids loads.

Sterile microfiltration based on use of membrane filters with 0.2-μm pores is used to remove bacteria from beers, thereby eliminating or reducing needs for pasteurization. Ceramic microfilters that can be sterilized and cleaned by back-flushing are used to clarify fruit juices. They are very expensive but will last for 10 years or more if care is taken to avoid heat shock.

Beds of particles used in solid–liquid extraction and adsorption and ion-exchange processes act like deep filter cakes. Pressure drops during flows through such beds often can be predicted with the aid of the Ergun equations or Kozeny–Carman equation. In other cases, flow pressure drops are calculated by means of empirically determined permeabilities or specific filtration resistances.

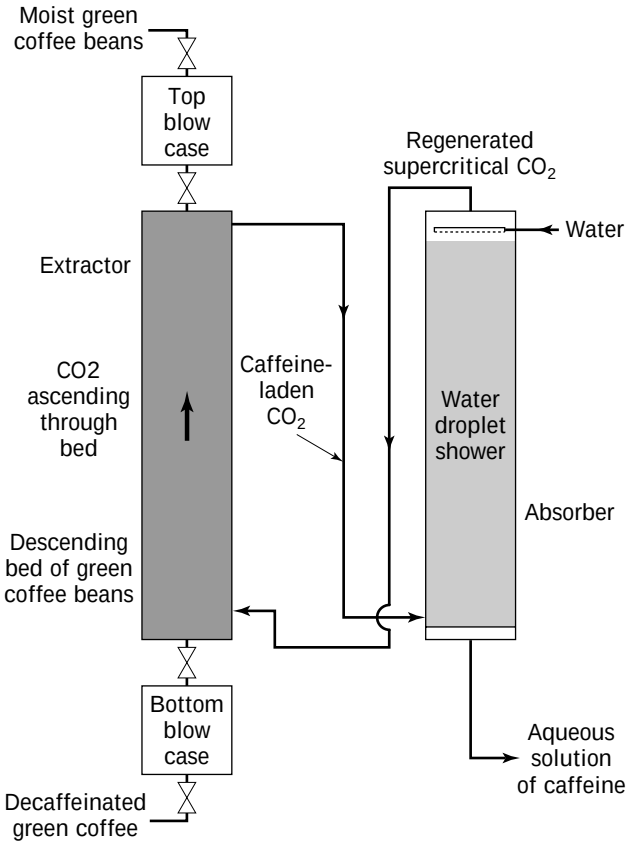


Figure 10. System for extraction of caffeine from green coffee by supercritical carbon dioxide. Reprinted with permission of Chemical Engineering Communications.

Centrifugation. Particles forced to rotate about an axis at high angular velocity in a centrifuge experience a radially acting force proportional to the radial distance from the axis, the angular velocity squared, and the density difference between the particle and surrounding fluid. At rotational speeds used in commercial centrifuges, particles denser than the fluid move outward and particles lighter than the fluid move inward thousands to tens of thousands times faster than they would sink and rise, respectively, under the influence of gravity. Therefore, centrifuges are used to separate and collect dispersed particles and droplets too small to be cleanly and quickly separated by other means. Centrifuges are used, for example, to separate dispersed oil, microorganisms, and fine biological particles and precipitated curds from aqueous solutions (23) and to remove yeasts from fermented beverages and waste treatment sludges from dischargeable water. Density differences driving centrifugal settling are usually much smaller for food processing than for chemical processing; the particles involved are more compressible, and, again higher levels of sanitation are required.

The particles or droplets involved move outward or inward until they strike a collection surface or material previously deposited on that surface. Deposited material then slides along the collection surface or is mechanically conveyed along it to a conduit or port that permits it to leave the centrifuge. Settled solids may leave through peripherally located nozzles as a dense slurry or leave in even denser form through intermittently opened peripherally located slits. Settled solids may also be knifed or augered out of centrifuges. Clarified liquid and separated oil or oil-rich creams leave through ports located closer to the rotational axis of the centrifuge. Centrifuges are sized to provide holdup time sufficient to permit complete settling or a desired level of settling. Large numbers of thin, closely spaced, concentric conical disks are mounted in centrifuges to reduce settling distances and times for extremely fine materials (e.g., butterfat droplets, and microorganisms). This greatly increases usable throughput rates for feeds containing such fine material.

Filtering centrifuges have a finely perforated wall or a perforated wall covered with filter cloth. Centrifugal force causes solids to deposit as a cake on the wall or cloth and then forces liquid through the cake. Wash liquor may also be forced through the cake. Such centrifuges are used to recover crystals (e.g., sugar and citric acid) from concentrated mother liquors, to separate freeze-concentrated vinegar and beverage extracts from ice, and to recover juice from milled fruit.

Centrifugal force generated by rotational flow is produced by tangentially injecting solids-laden liquid or gas into hydroclones or gas cyclones. These units have a cylindrical shell attached to a conical bottom, a central outflow port at their top, and a bottom port through which solids or solids-rich underflows leave. The injected liquid or gas first spirals downward in an outer vortex and then spirals upward in an inner vortex. The vortex flow tends to cause particles to move outward toward the wall. In gas cyclones virtually all the gas leaves through the top port. Gas cyclones are sized to provide enough flow residence time to permit wall impact by all particles greater than a desired size. Gas cyclones are used to separate fine spray-dried material from discharged drying air, pneumatically elutriated materials from carrier gas, chaff from coffee roaster discharge gases, and pneumatically conveyed material from conveying air.

In hydroclones, most entering liquid leaves through the upper port as overflow, but enough leaves through the lower port to provide a flowable slurry, the underflow. The centrifugal force generated by vortex flow is partly counterbalanced by drag force generated by radial flow of liquid from the outer to the inner vortex. Consequently, small particles or low-density particles may migrate inward and be carried away with the overflow. Larger or denser particles leave with the overflow. Large banks of miniature hydroclones used in parallel are employed to separate starches from protein- and solute-rich solutions during the wet milling of corn.

Expression. Pressing of fluid-rich biological matter is used to recover fruit and sugar juice, vegetable and fish oil, pectin released by cooking apple pomace, and protein-rich juice from leaves. It is also used to expel whey and to fuse curds together during cheese manufacture and to dewater spent food processing wastes, by-products, and processing intermediates (37). Expression is carried out in a wide variety of equipment (38). In screw presses, feed is trans-

ported down the length of a perforated barrel by a rotating single screw or by twin screws. The feed compacts, and fluid is expelled through barrel perforations because the screw's root diameter increases, its pitch or rotary speed decreases, or the barrel diameter decreases, or because constricted solids outflow from the barrel causes a progressive buildup of pressure.

In roll mills, a grooved, rotating roll is hydraulically pressed against two similar counterrotating rolls that lie beneath it. As the rolls rotate, they drag feed (usually shredded sugarcane) through nips between the top and bottom rolls. This compacts the feed and expels juice. Some of the juice flows back through the advancing feed; some flows laterally into the grooves and then backward. The process is analyzed by relating local backflow and flow pressure drop to the cumulative extent and rate of compaction and the local specific filtration resistance of the compacted feed.

In belt presses, feed is deposited on a moving cloth belt and pressed between that belt and an opposing, moving cloth belt driven over a series of solid or perforated rolls. In some cases the belts follow a serpentine path. Expelled fluid flows through the cloths and is collected in troughs beneath the rolls. Figure 11 shows a serpentine belt press of a type used to press juice out of milled apples. In presses of other types, feeds confined in a perforated cage or in stacked cloth pouches are pressed by a diaphragm or ram. Fluid is also expelled from cakes compacted by centrifugal force in decanter centrifuges.

Engineering treatment of expression is based on numerical solution of partial differential equations that describe how flow in press cakes is driven by fluid pressure gradients or related solid stress gradients and how rates of local cake volume change as a function of flow rate gradients (38,39). Empirically determined relationships between solid stress, specific filtration resistance, and the solids-based specific volume of the cake are used in solving these equations. Methods have been developed for determining these relationships from constant-pressure or constant-rate pressing tests carried out in suitably instrumented equipment (40).

Some press feeds (e.g., waste treatment sludges and fruit with thin cell walls and weak intercellular bonds) compact excessively as pressure increases. When such feeds are pressed, fluid pressures drop precipitously and solid stress rises sharply over very thin layers near outflow surfaces. Pressing can be improved for such feeds by adding an inert material (e.g., rice hulls) that stiffens the press cake. Other feeds (e.g., spent coffee grounds and oilseeds) have thick cell walls that adhere strongly together. Fluid pressure builds up inside the solid when such feeds are pressed. The extent of in-solids pressure buildup depends on the rate of pressing and the extent of compaction. Pressing of such materials can be improved by treatments that promote cell wall rupture.

PEELING, SKINNING, AND HULLING

Peels, skins, fatty layers, hulls, and shells frequently must be removed from foods. Skins are often manually cut from meat and fish with the assistance of special tools. Wholly mechanical systems are also used. These grab a carcass, carcass part, or fish at a selected spot and drag it past suitably positioned knives and saws. Fat-depth-sensing devices are sometimes used to guide cutting.

Moisture contents of shell-encased and hull-encased foods are often adjusted to increase shell and hull friability and separability. Hulls and shells are then cracked by impaction in rapidly rotating pin mills or by passing the food involved through pairs of rollers or plates that turn at different speeds and produce a pressing, rolling action. Thin shell and hull fragments are then pneumatically separated from more compact usable food particles.

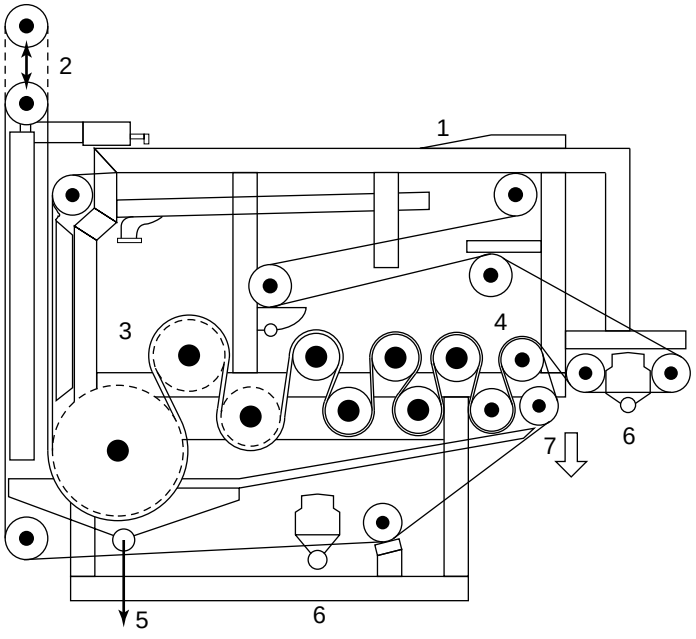


Figure 11. Serpentine belt press. 1, Mash deposition; 2, wedge area; 3, low-pressure press zone; 4, high-pressure press zone; 5, juice runoff; 6, belt cleaning; 7, press-cake discharge. Reprinted with permission of Bellmer GmbH and Co.

Manual peeling of vegetables and fruit is costly; abrasion peeling removes too much peel and is wasteful. Therefore, treatments that break down bonds between cells or decompose cells in or near skins are used. These include short time heating with high pressure steam, short-time immersion in baths of lye, and passing product through a flame. The loosened skin is then brushed off or washed off by water jets, a step that also removes residues of chemical peeling agents. The process can be analyzed in terms of short-term heat or mass transfer in surface layers and the kinetics of protopectin or cell wall breakdown. To sharply limit intercellular bond breakdown, internal temperature gradients near the product surface must be very large and exposure times very short. Even so, undesirable textural and color changes often occur in regions close to the peeled surface.

CLEANING

Washing-based cleaning of food surfaces and cleaning and disinfection of equipment are widely used food processing operations that have been analyzed from an engineering point of view (41–43). Efficient cleaning and rinsing can be achieved only when smooth, highly polished, crevice- and pocket-free equipment is used. Equipment conforming to the sanitary design standards of the Technical Committee of Dairy and Food Industry Supply Association and made of type 304 or type 316 stainless steel is frequently used. Parts may also be made of plastics and synthetic materials approved by the U.S. Food and Drug Administration or Department of Agriculture.

Equipment cleaning frequently involves an initial water rinse, a caustic detergent wash, a second water rinse, an acid wash (if calcium-rich deposits have formed), and a final water rinse. Conditions that promote corrosion and etching (e.g., use of chlorides) must be avoided. Disinfection is provided by cleaning or rinsing at high temperatures, by steam heating of equipment, or by use of chemical agents (e.g., hypochlorites). Cleaning formerly involved manual hosing down and brushing of disassembled pipelines and equipment and one-time use of cleaning agents. Manual cleaning procedures are still frequently used in meat processing plants and in food processing pilot plants. Spray wands operating at pressures up to 3.5 MPa (500 psig) are used to ensure removal of adherent material.

Clean-in-place (CIP) systems that accomplish thorough cleaning without disassembling piping and permit some reuse of cleaning agents have been developed for many plants. Because cleaning agents and rinse water run through pipelines and narrow-bore equipment in CIP systems, they must be laid out to permit hang-up–free drainage. Product contact surfaces in tanks and large-bore equipment are cleaned with the aid of sprays issuing from permanently installed spraying devices. Globelike spray heads or reciprocating sprays are used to provide complete washing and rinsing of internal surfaces. Removal of cleaning agents during rinsing has been successfully modeled so that rinsing can be accomplished without excessive use of water.

Soil deposition and fouling interfere with process heat transfer and greatly reduce permeation rates in membrane-based processes. Therefore fouling and soil deposition mechanisms and kinetics for such processes have been extensively studied. The results of these studies have been used to optimize frequencies of cleaning and to select operating conditions and surface treatments that retard fouling.

HOMOGENIZATION AND EMULSIFICATION

Fat in freshly drawn cow's milk is dispersed as butterfat globules that individually contain a fatty core coated with a lipoprotein layer roughly 10 nm thick. Depending on the type of cow involved and how long the animal has been lactating, mean volume/surface diameters of butterfat globules range between 2.5 and 4.5 μm (43). Unless treated, the globules rise, producing a layer of cream on top of milk. To prevent this, and to minimize butterfat clumping, the globules are broken into droplets whose mean volume/surface diameter ranges between 0.22 and 0.5 μm. This is usually done by passing milk through spring-loaded, narrow-clearance valves in high-pressure piston pumps, called homogenizers (26,43). Cavitation produced by rapid increases in velocity in the pump valve and flow-induced shear break up the globules. The mean size of the resulting droplets is inversely proportional to the homogenization pressure used raised to the 0.6 power. Brownian motion would very rapidly cause collision, clumping, and coalescence of incompletely coated droplets. Therefore the lipoprotein coating must redistribute itself rapidly enough to completely coat freshly created drops before they collide with each other and stick together.

Dispersions of small drops of oil in a continuous aqueous phase are often created by intensive agitation of a mixture of an oil and an aqueous solution. The agitation produces turbulent velocity fluctuations, which in turn cause local shearing intense enough to overcome the interfacial tension forces that tend to hold drops together. Mean drop sizes produced by such agitation are frequently proportional to the interfacial tension raised to the 0.6 power and inversely proportional to agitation power expenditure per unit volume raised to the 0.4 power. Simple dispersions rapidly coalesce and separate when mixing stops. To avoid this, surfactants or agents that coat drops and prevent or greatly retard drop coalescence are added. Dispersions stabilized in this way are called emulsions. Stiff emulsions that act like soft solids (e.g., mayonnaise) contain so much dispersed material that drops butt against each other. A great deal of work has been done to develop gelled low-fat food systems with textural properties resembling those of stiff, oil-rich emulsions.

SOLIDIFICATION

Gelling and Coagulation

Fluid foods often are converted into fluid-rich solids with desirable textures by means of gelling and coagulation. Acid produced by bacterial action and bond site production due to the action of the enzyme rennet serve to coagulate milk casein proteins and produce cheese (44–46). Acidifi-

cation is used to precipitate and gel food proteins with low isoelectric pH values (e.g., casein, milk whey proteins, and soy proteins solubilized by base addition). Calcium salts precipitate soy proteins and cause them to gel during the production of tofu (47). Heat-induced gelation occurs during the cooking of many protein-rich solutions and slurries. Other protein solutions (e.g., gelatin) gel when cooled and melt when heated. High-methoxy pectin causes acidic sugar-rich fruit juices and slurries to gel when jams and jellies are produced. Fully hydrated (gelatinized) starches are used to produce cold-set soft puddings.

Liquid foods gel because reactions or physical chemical forces bond together sites on adjacent, long-chain food molecules or chains of globular proteins (48). Low extents of bonding produce increases in viscosity. Gelling occurs when bond junction densities become large enough to form coherent three-dimensional polymeric networks. Gelling occurs only if the gelling agent concentration C exceeds a critical value C_o. Gelling times are often inversely proportional to $(C - C_o)^n$, where n is a constant. Simple kinetic considerations indicate that n should equal 1, but it is frequently larger than 1. Gelling and coagulation rates may be controlled by the rate of bond site production or by rates of bond forming reactions or events. Cross-linking continues after gels form, and the gels progressively stiffen. This reduces the mobility of network sections, slowing and eventually stopping the bonding process. At gelling agent concentrations substantially greater than C_o, gel strength tends to increase roughly linearly as $(C/C_o)^2$ increases.

Cooling-Induced Solidification

Hard sugar candies are made by evaporating water from a concentrated solution of mixed sugars until the solution is essentially converted into a hot melt containing less than 1 % moisture. Then the melt is cooled. So viscous that crystal growth does not occur, the melt ultimately becomes a vitreous glass. Flavoring and coloring agents are blended into the melt. When the hardening melt becomes stiff enough, it is processed in devices that convert it into pieces of appropriate size and shape. Chewy candies are made by cooling, flavoring, coloring, and shaping evaporatively concentrated mixtures of sugars, corn syrup, fat, and milk solids containing 12–15% moisture. In both cases, solidification is effectively caused by large cooling-induced increases in viscosity and flow yield strength. Cooling is controlled to prevent excess stress from developing and to permit shaping (e.g., molding, rolling, and extrusion) before the completion of hardening.

Chocolate is made by molding and cooling finely milled, thoroughly worked mixtures of molten cocoa butter, cocoa, sugar, and in some cases milk solids. Cocoa butter can exist in four polymorphic forms that solidify or change form at different temperatures. Only one, the β form, is stable at room temperature. Therefore carefully controlled cooling and seeding with fine β-form crystals is used in the solidification of chocolate.

SHAPING AND TEXTURIZING OPERATIONS

Rolling and Flaking

Food processing intermediates used to produce bread, noodles, cookies, and chewing gum are often rolled into sheets before being further shaped by cutting and other forming operations. Cooked cereal grits are rolled to form flakes that are dried and toasted to make ready-to-eat breakfast cereals. Oilseed grits are rolled into flakes to prepare them for extraction. Ropes of partially hardened candy melts are pulled through sequences of pairs of rollers turning at progressively faster speeds to produce narrow-diameter ropes, which are then cut into pieces that are pressed into shape by other devices. Engineering aspects of these operations can be dealt with by methods developed to analyze the calendaring of plastics (48). These methods involve the use of partial differential equations to describe flow fields and pressure gradients that develop in plastic materials passing between rolls. These equations can be solved to determine the pressing force and torque acting on the rolls and the power required to drive them. Dimensional analysis has been applied to circumvent complications due to the complex rheological characteristics of food doughs (49).

Extrusion

Foods are frequently shaped, mixed, cooked, and texturized in screw-driven extruders (50). Flow driven by shear between the rotating screw and the stationary barrel conveys, mixes, and kneads the material being processed and forces it through dies that shape the product. Rotating or oscillating knives cut the extruded material into pieces of suitable size. Action of the screws on the processed material can generate a great deal of frictional heat, particularly when screws with shallow flights are used and the feed's moisture content is low. Supplemental heating is often provided by heated jackets. Cooling jackets are sometimes used. Initially, single-screw extruders were used; now, extruders containing corotating twin screws that contain changeable sequences of conveying, kneading, and pressurizing elements (Fig. 12) are frequently used. Flow, mixing, and power expenditure in extruders of both types have been treated by methods developed for extrusion processing of polymers (49,51). Foods processed in extruders usually contain 15–31% water so that they can be made plastic when worked on by the screw.

Extruders operating close to room temperature and containing screws with deep flights are used to form pasta by forcing semolina-based dough through large arrays of parallel holes. Die hole inserts that provide nonuniform clearance are used in the production of elbow macaroni. High-temperature extrusion, based on use of screws with shallow flights or elements that provide intense kneading, is used to produce confectionery doughs and puffed and textured foods. Temperatures between 130 and 180 °C are frequently used in making these products. Extrusion is used to convert soy protein into products that resemble meat in structure and texture. Such texturization appears to involve the formation of links between ϵ acids and ϵ amines on amino acids in neighboring protein chains. Kinetic models for the reaction indicate that its rate depends on shear rates in the extruder and die as well as the concentrations of the reacting species.

Concentrated purees or slurries containing components capable of entering into setting reactions are pumped or

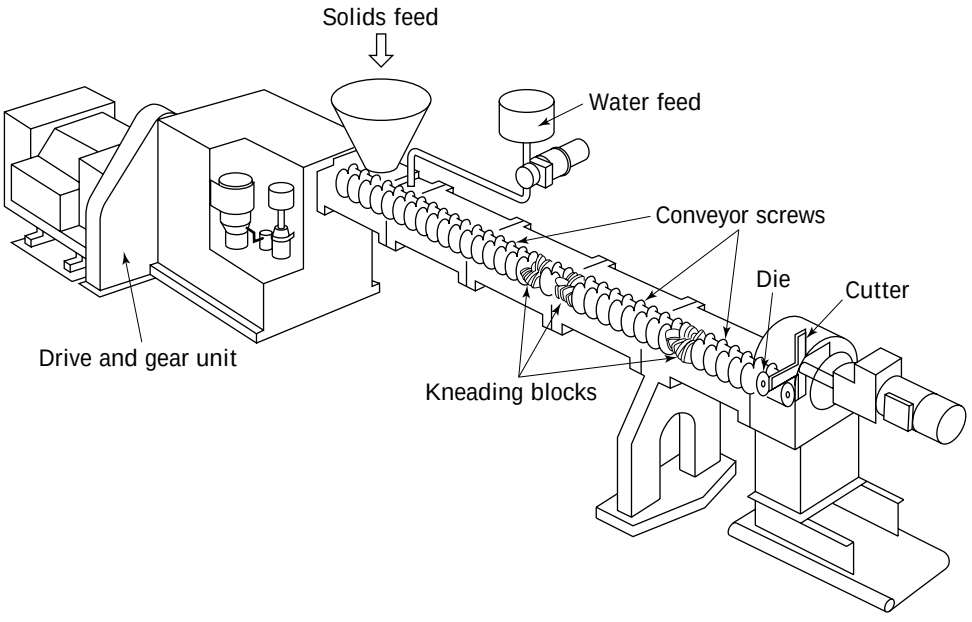


Figure 12. Twin-screw extruder. Reprinted with permission of Krupp Werner and Pfleiderer Corporation.

extruded through dies into baths or vapor-rich atmospheres containing setting agents. Examples include the use of extruded, slurried, partially dissolved collagen or viscose solutions to form sausage casings and use of pumped slurries containing alginates and pureed fruits or vegetables to form artificially shaped fruit and vegetable pieces. Pump- or piston-based extrusion is used to form products that are subsequently heat set (e.g., doughnuts, masa-based puffed products, and certain types of cookie dough). Extrusion is also used to stuff meat emulsions into sausage casings. Coextrusion is used to form fruit-filled bars.

Moist feeds used to produce shaped cereal pieces and pelletized animal feeds and pet food are pushed by rollers through die holes in thick, rotating, perforated rings. Extrudate emerging from the holes is cut into short pieces by knives.

OTHER OPERATIONS

Because of space limitations, many food processing operations have not been covered in this article. These include mixing; grinding, and milling; cutting, slicing, and dicing; screening; pneumatic separation and classification; pumping; mechanical, pneumatic, and hydraulic conveying; slaughtering and carcass disassembly; casting and molding; distillation and deodorization; liquid–liquid extraction; gas absorption; crystallization; adsorption and ion exchange; chemical modification of food ingredients (e.g., hydrogenation of vegetable oils); various types of cooking and baking; treatment of foods with ionizing radiation; pore generation and puffing; coating and enrobing; sorting and inspection; packaging and filling; fermentation, including brewing, wine making, and the production of pickled vegetables and fermented dairy products; aquaculture; shellfish depuration; automated egg production, and automated growing of plants and plant tissue in controlled environments.

BIBLIOGRAPHY

1. N.W. Desrosier, *Elements of Food Technology*, AVI Publishing, Westport, Conn., 1977.
2. H.S. Carslaw and J.C. Jaeger, *Conduction of Heat in Solids*, Oxford Univ. Press, Oxford, U.K., 1959.
3. J. Crank, *Mathematics of Diffusion*, Oxford Univ. Press, London, 1970.
4. B.K. Watt and A.L. Merrill, *Composition of Foods: Agriculture Handbook 8*, U.S. Department of Agriculture, Washington, D.C., 1975.
5. D.B. Brooker, F.W. Bakker-Arkema, and C.W. Hall, *Drying and Storage of Grains and Oil Seeds*, Chapman & Hall, New York, 1992.
6. Midwest Plans Service, *Grain Drying, Handling, and Storage Handbook*, Midwest Plans Service, Iowa State University, Ames, Iowa, 1988, pp. 3.1–3.21.
7. K. Masters, *Spray Drying Handbook*, 5th ed., Wiley, New York, 1991.
8. F.G. Kieviet and P.J.A.M. Kerkhof, in *Proceedings 5th World Congress of Chemical Engineering*, vol. 2, American Institute of Chemical Engineers, New York, 1996, pp. 87–92.
9. J.D. Mellor, *Fundamentals of Freeze Drying*, Academic Press, New York, 1979.

10. C.J. King, *Freeze Drying of Foods*, CRC Press, Cleveland, Ohio, 1971.

11. S.A. Goldblith, L. Rey, and W.W. Rothmayr, *Freeze Drying and Advanced Food Technology*, Academic Press, New York, 1975, pp. 123–480.

12. H.G. Schwartzberg, *69th Annu. American Institute of Chemical Engineers Meeting*, Chicago, Ill., December 1976, pp. 1–54.

13. M. Karel, in M. Karel, O.R. Fennema, and D.B. Lund eds., *Principles of Food Science, II: Physical Principles of Food Preservation*, Marcel Dekker, New York, 1975, pp. 237–263.

14. O.R. Fennema, in M. Karel, O.R. Fennema, and D.B. Lund eds., *Principles of Food Science, II: Physical Principles of Food Preservation*, Marcel Dekker, New York, 1975, pp. 133–171.

15. R.E. Hardenburg, A.E. Watada, and C.Y. Wang, *The Commercial Storage of Fruits, Vegetables, Florist, and Nursery Stocks*, Agriculture Handbook No. 66. U.S. Department of Agriculture, Washington, D.C., 1986, pp. 1–130.

16. *1994 ASHRAE Handbook—Refrigeration* American Society of Heating, Refrigeration, and Air-Conditioning Engineers, Atlanta, Ga., 1994.

17. D.H. Dewey, in S. Thorne, ed., *Developments in Food Preservation*, vol. 2, Applied Science Publishers, London, 1983, pp. 1–24.

18. O.R. Fennema, W.D. Powrie, and E.H. Marth, *Low Temperature Preservation of Foods and Living Matter*, Marcel Dekker, New York, 1973.

19. O.R. Fennema, in M. Karel, O.R. Fennema, and D.B. Lund eds., *Principles of Food Science, II: Physical Principles of Food Preservation*, Marcel Dekker, New York, 1975, pp. 173–215.

20. J. Postolski and Z. Gruda, *Zamraźani Żywnosci*, Wydawnictwa Naukowo-Techiczne, Warsaw, 1985.

21. H.G. Schwartzberg, in *Freezing, Frozen Storage and Freeze Drying of Biological Materials and Foodstuffs*, International Institute of Refrigeration, Paris, 1977, pp. 303–310.

22. H.G. Schwartzberg, *American Institute of Chemical Engineers Summer National Meeting*, Detroit, Mich., 1981, pp. 1–33.

23. A.C. Cleland, *Food Refrigeration Processes, Analysis, Design, and Simulation*, Elsevier Applied Science, New York, 1990.

24. R.P. Singh and J.D. Mannapperuma, in H.G. Schwartzberg and M.A. Rao eds., *Biotechnology and Food Process Engineering*, Marcel Dekker, New York, 1990, pp. 309–358.

25. D.L. Downing, *A Complete Course in Canning and Related Processes*, 13th ed., CTI Publications, Baltimore, Md., 1996.

26. H.G. Kessler, *Food Engineering and Dairy Technology*, Kessler Verlag, Freising, Germany, 1981, pp. 59–81, 119–207.

27. D.B. Lund, in M. Karel, O.R. Fennema, and D.B. Lund eds., *Principles of Food Science, II: Physical Principles of Food Preservation*, Marcel Dekker, New York, 1975, pp. 31–92.

28. R. Villota and J.G. Hawkes, in M. Okos, ed., *Physical and Chemical Properties of Foods*, American Society for Agricultural Engineering, St. Joseph, Mich., 1986, pp. 266–366.

29. H.G. Kessler, in *Proceedings 5th World Congress of Chemical Engineering*, vol. 2, American Institute of Chemical Engineers, New York, 1996, pp. 29–34.

30. H.G. Schwartzberg, *Food Technol.* **31**, 67–76 (1977).

31. H.G. Schwartzberg, in R.P. Singh and A.G. Medina eds., *Food Properties and Computer-Aided Design of Food Processing Systems*, Kluwer Academic Publishers, Boston, Mass., 1989, pp. 443–470.

32. H.A.C. Thijssen, in A. Spicer, ed., *Advances in Preconcentration and Dehydration of Foods*, Wiley, New York, 1973, pp. 115–149.

33. H.G. Schwartzberg, in H.G. Schwartzberg and M.A. Rao eds., *Biotechnology and Food Process Engineering*, Marcel Dekker, New York, 1990, pp. 127–202.

34. H.G. Schwartzberg, in R.W. Roussea, ed., *Handbook of Separation Process Technology*, Wiley, New York, 1987, pp. 540–577.

35. H.G. Schwartzberg, *Chem. Eng. Prog.* **76**, 67–85 (1980).

36. M. Shirato, T. Murase, E. Iritani, F.M. Tiller, and A.F. Alciatore, in M.J. Matteson and C. Orr eds., *Filtration Principles and Practices*, Marcel Dekker, New York, 1987, pp. 299–424.

37. H.G. Schwartzberg, in *McGraw-Hill Encyclopedia of Science and Technology*, vol. 7, McGraw-Hill, New York, 1987, pp. 254–255.

38. H.G. Schwartzberg, *Sep. Purif. Methods* **26**, 1–213 (1997).

39. H.G. Schwartzberg, in M. Peleg and E. Bagley, eds., *Physical Properties of Foods*, AVI Publishing, Westport, Conn., 1983, pp. .

40. B. Hallstrom, D.B. Lund, and C. Trägårdh, eds., *Fundamentals and Applications of Surface Phenomena Associated with Fouling and Cleaning in Food Processing*, Department Food Science, University of Wisconsin, Madison, Wisc., 1981.

41. D.B. Lund, E. Plett, and C. Sandu, eds., *Fouling and Cleaning in Food Processing*, Department of Food Science, University of Wisconsin, Madison, Wisc., 1985.

42. H.G. Kessler and D.B. Lund eds., *Fouling and Cleaning in Food Processing*, Kessler Verlag, Freising, Germany, 1989.

43. H. Mulder and P. Walstra, *The Milk Fat Globule*, Commonwealth Agricultural Bureau, Farnham Royal, England, 1974, pp. 163–194.

44. H.G. Kessler, *Food Engineering and Dairy Technology*, Kessler Verlag, Freising, Germany, 1981, pp. 426–449.

45. F. Kosikowski, *Cheese and Fermented Milk Products*, Brooktendale, New York, 1978.

46. P.F. Fox ed., *Cheese: Chemistry, Physics, and Microbiology*, vol. 1, *General Aspects*, Chapman & Hall, London, 1993.

47. W. Shurtleff and A. Aoyagi, *Tofu and Soymilk Production*, New-Age Foods Study Center, Lafayette, Calif., 1979.

48. S. Middleman, *Fundamentals of Polymer Processing*, McGraw-Hill, New York, 1977, pp. 123–187, 295–353.

49. K. Valentas, L. Levine, and J. Peter Clark, *Food Processing Operations and Scale-Up*, Marcel Dekker, New York, 1991, pp. 359–365.

50. J.M. Harper, in H.G. Schwartzberg and M.A. Rao eds., *Biotechnology and Food Process Engineering*, Marcel Dekker, New York, 1990, pp. 295–308.

51. L.P.B.M. Janssen, *Twin-Screw Extrusion*, Elsevier Science Publishers, New York, 1978.

See also ENZYMES, BAKING, BREAD MAKING; ENZYMES, FRUIT JUICE PROCESSING; FREEZE DRYING, PHARMACEUTICALS; PRODUCTION OF L-AMINO ACIDS BY AMINOACYLASE; SOLID SUBSTRATE FERMENTATIONS, ENZYME PRODUCTION, FOOD ENRICHMENT; YEAST, BAKER'S.

FORMULATION AND DELIVERY, PROTEIN PHARMACEUTICALS

XANTHE M. LAM
JEFFREY L. CLELAND
Genentech, Inc.
South San Francisco, California

KEY WORDS

Degradation
Drug delivery
Formulation
Lyophilization
pH dependence
Recombinant protein
Shelf life
Stability
Stress testing
Temperature dependence

OUTLINE

INTRODUCTION

For the last twenty years, many biotechnology-derived pharmaceuticals have been manufactured for the treatment of diseases such as diabetes, cardiovascular diseases, cancers, dwarfism, anemia, AIDS, cystic fibrosis, chronic granulomatous disease, and kidney diseases. Some of these recombinant protein drugs have already been approved by the U.S. Food and Drug Administration (FDA) for marketing, others are still being tested in clinics or are under review for approval. The success of these protein pharmaceutical drugs greatly depends on the delivery of their biologically active forms to the target site. Therefore, the development of a stable formulation, in which the protein can maintain its native conformation and bioactivity, is one of the important steps in the production of protein pharmaceuticals. In fact, the FDA usually requires the pharmaceutical manufacturers to provide real-time stability data to demonstrate that a pharmaceutical drug contains at least 90% potency throughout the entire shelf life period (usually $\geq$ 2 years) before it can be approved to be a marketed product. In addition, the manufacturers are required to demonstrate that the degradation products that may cause a loss in the potency of the drug do not have any adverse effects on the safety and efficacy of the drug.

In developing a protein drug formulation and designing a delivery system, formulation scientists must consider the physicochemical properties, clinical indication, site of action, pharmacokinetics, and toxicity of the drug. The physicochemical properties of proteins such as the amino acid composition, molecular weight, isoelectric point (pI), glycosylation, and conformation can affect the pharmacokinetics and toxicity as well as the clinical indication. The potential clinical application of a protein drug depends upon the biological function and potency of the product, and the physical and chemical properties of a protein determine its biological function and potency.

The unique physical and chemical properties of each protein also determine its in vitro and in vivo stability. To obtain in vivo information on the pharmacokinetics and toxicity of a drug, the drug must be administered with a stable formulation. In considering the best formulation for the intended applications of a protein drug, the formulation scientist must also consider the route of administration. For initial animal testing, protein drugs are usually administered systemically via an intravenous (i.v.) injection. However, some indications may require a high local drug dose that cannot be achieved by i.v. administration due to toxicity issues. In this case, an alternative route of delivery is necessary. For example, Pulmozyme® (rhDNase) is delivered to patients with cystic fibrosis via the pulmonary route of administration (1). The drug is best administered directly into the lungs of the patients as an aerosol using nebulizers in order to achieve an efficacious dose at the target site (lungs). Thus, the development of both a stable formulation and suitable delivery route or system is critical in determining the success of a final product.

During the past 10 years, several review articles and texts related to protein formulation and stability have been published (2–7). The purpose of this article is to provide a comprehensive overview of formulation and delivery aspects of protein drugs. Protein degradation pathways, analytical characterization of proteins, the process of formulation development, and drug delivery systems are reviewed.

PROTEIN DEGRADATION PATHWAYS

Chemical and physical degradation is the major cause of instability for biopharmaceutical products. There are many degradation routes for proteins. Aggregation, deam-

idation, and oxidation are the three most common ones, others include isomerization, cleavage, thiol disulfide exchange, and β-elimination. To develop a stable protein formulation, protein degradation in the formulation must be inhibited or minimized throughout the shelf life of the product. Fortunately, with the advances in analytical techniques for protein analysis, most of these expected degradation routes can be identified, and the degradation products can also be characterized by formulation scientists. In addition, with the continued success of biotechnology, more recombinant proteins have become available for studying degradation mechanisms, thus increasing the general knowledge in formulating new proteins.

Aggregation

Protein degradation via aggregation pathways is considered a physical reaction that does not involve the breakage or formation of covalent bonds. For a protein to retain its biological function and stability, it must maintain its conformational or tertiary structure. Protein conformation is stabilized mainly by hydrophobic interactions (8–11). Thus, monomeric globular proteins often exist in native or folded conformations such that the hydrophobic groups are not exposed on the surface (12). Loss in the tertiary globular structure or unfolding of proteins due to denaturation results in exposure of hydrophobic residues to the environment. Protein aggregation or denaturation in a formulation (liquid or lyophilized) can be caused by a change in temperature, extreme pH, ionic strength, pressure, or presence of denaturants (11,13). Usually, the native, N, and unfolded, U, states of a protein are in equilibrium during denaturation and folding in a two-state process, as shown in equation 1. The reversibility of the unfolding process depends upon the formulation conditions.

$$\mathrm{N} \rightleftharpoons \mathrm{U} \tag{1}$$

$$\begin{array}{ccccc} \mathrm{N} & \rightleftharpoons & \mathrm{I} & \rightleftharpoons & \mathrm{U} \\ & & \downarrow & & \\ & & \mathrm{A} & & \end{array} \tag{2}$$

Sometimes, a protein denatures via a different pathway that involves the formation of at least one stable, partially unfolded intermediates. As shown in equation 2, the native protein unfolds to form an intermediate, I, that has the internal hydrophobic residues exposed to the environment. This hydrophobic intermediate on the refolding pathway may then form aggregates, A, due to intermolecular hydrophobic interactions. According to this scheme, the protein must be denatured or unfolded before nonnative aggregates can be formed. The formation of a stable intermediate during denaturation and folding has been observed for many proteins (14–18). For example, the dimeric native form of interferon-γ (IFN-γ), unfolds to form a partially denatured monomeric intermediate, below pH 4.5 and in the absence of NaCl (19,20). Upon dialysis, both native IFN-γ and aggregates were obtained. The aggregates lost most of the native tertiary structure as well as bioactivity.

Deamidation

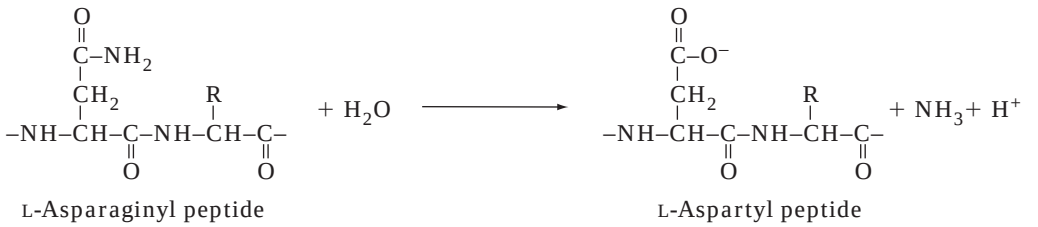
Deamidation of protein is the acid- and base-catalyzed hydrolysis of the side-chain amide on glutamine (Gln) and asparagine (Asn) residues to form a carboxylic acid (Scheme 1). Studies of deamidation in peptides by Robinson and his coworkers suggested that Asn is more susceptible to deamidation than Gln and degrades more readily in the presence of an adjacent glycine (Gly) residue on the C-terminal side of the sequence (21,22). It was also discovered that the deamidation of Asn-Gly was accelerated at pH $\geq$ 7.0. In addition, the studies found that the amino acid sequence plays a major role in determining the rate of deamidation. For example, the presence of polar amino acid residues adjacent to an Asn or Gln enhances the rate of deamidation, whereas adjacent bulky hydrophobic residues decrease the deamidation rate. Serine (Ser) or threonine (Thr) next to an Asn or Gln can act as a general acid by providing a proton in the acid-catalyzed reaction, thereby enhancing the rate of deamidation. The major factors influencing the rate of protein deamidation include pH, temperature, ionic strength, and buffer species. Degradation of proteins via deamidation has been shown to exist in recombinant protein pharmaceuticals such as insulin (23,24) and human growth hormone (25,26) in liquid formulations.

Isomerization

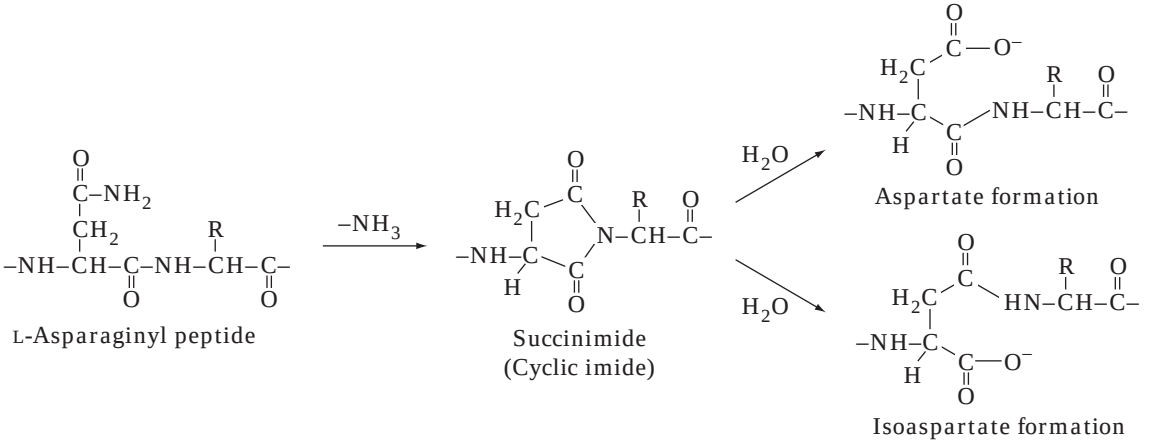
Deamidation of an asparagine (Asn) residue to yield aspartate (Asp) can also lead to the formation of an isoaspartate (iso-Asp) via intermediate succinimide formation (27). Nucleophilic attack of the main-chain amide nitrogen on the carbonyl carbon of the Asn side-chain amide group results in a five-membered succinimide ring. This cyclic imide can be hydrolyzed to form either aspartate or isoaspartate (Scheme 2). An aspartate residue can also undergo deamidation via intermediate succinimide formation to yield isoaspartate. In this case, nucleophilic attack of the main-chain amide nitrogen on the carbonyl carbon of the Asp side-chain carboxylic group results in the formation of a cyclic imide (Scheme 3). The additional CH_2 group in the glutamine (Gln) side chain makes the formation of a cyclic imide intermediate by nucleophilic attack less favorable than for an Asn or Asp residue. Therefore, the deamidation rate of Gln via isomerization is very slow. When a Gly or Ser residue is next to the Asn or Asp residue, the isomerization via cyclic imide formation is accelerated in solution (28). An increase in temperature or pH also increases the rate of isomerization for peptides containing an Asn-Gly or Asp-Gly sequence (29). Cyclic imide formation was demonstrated during the storage of lyophilized recombinant methionyl human growth hormone (met-rhGH) at 45 °C for 4 months (30), and the Asp-130 residue in the protein underwent isomerization (31).

Cleavage

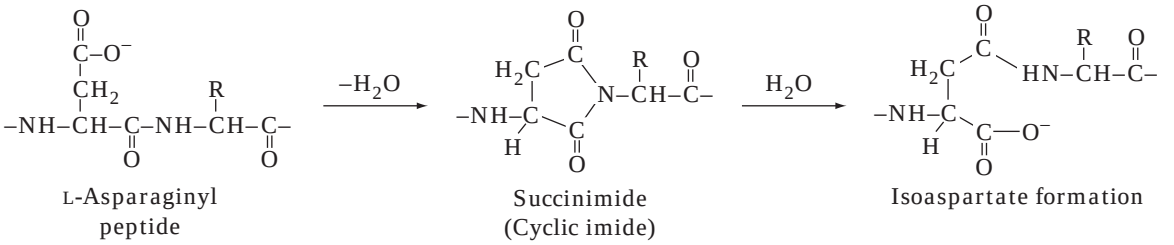
Cleavage of peptides at an asparagine residue can result from the nucleophilic attack of the side-chain amide nitrogen on the main-chain peptide carbonyl to form a C-terminal succinimide (Scheme 4). The reaction is spontaneous and dependent on protein sequence. Asparagine



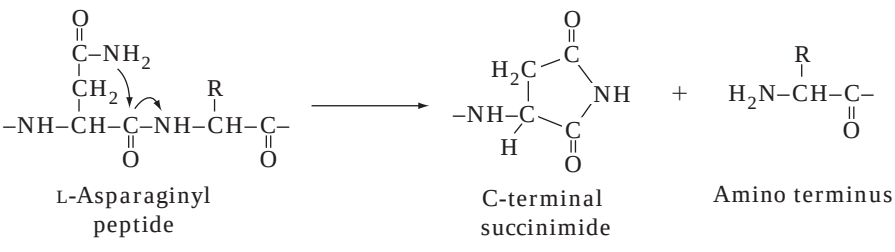
Scheme 1. Deamidation of asparagine residue in a peptide by direct hydrolysis.



Scheme 2. Pathway for asparagine deamidation in peptides or proteins via intermediate succinimide formation.



Scheme 3. Deamidation of aspartate to yield isoaspartate via succinimide formation.



Scheme 4. Spontaneous cleavage of peptide as asparagine residue.

having an adjacent Gly or Ser residue on the C-terminal side is more labile to spontaneous cleavage. Cleavage at an Asn-Ser linkage has been demonstrated in proteins such as bovine and porcine somatotropins (32).

Oxidation

Oxidation is one of the major degradation pathways for protein pharmaceuticals. Amino acids that can undergo

oxidation include methionine, cysteine, histidine, tryptophan, and tyrosine. Oxidation of methionine has been demonstrated in many recombinant proteins such as interleukin 2 (33), relaxin (34), parathyroid hormone (35), and human growth hormone (36). At low pH, the thioether group of methionine is not protonated and is susceptible to oxidation, resulting in methionine sulfoxides. Under extremely oxidative conditions, the sulfoxides can be further oxidized to sulfones. Methionine can react with a variety of oxygen-reactive species such as hydrogen peroxide, alkylhydroperoxides, molecular oxygen, and singlet oxygen generated by heat and light. Methionine oxidation can be catalyzed by the presence of transition metal ions such as Cu^{2+}, Fe^{2+} or Fe^{3+}. Although there are many oxidative pathways for methionine reported in literature (3,7,37,38), oxidation via a free radical, singlet oxygen, or nucleophilic substitution are the three common mechanisms. In contrast to methionine, oxidation of cysteine usually occurs at higher pH where the thiol is deprotonated (39). Cysteine can react with molecular oxygen to form cystine disulfide. The reaction can also be catalyzed by transition metal ions, especially Cu^{2+}. In some proteins cysteine can be oxidized to give sulfenic, sulfinic, or sulfonic acids (40). Histidine, tryptophan, and tyrosine are susceptible to photooxidation via the singlet oxygen pathway. The rate of photooxidation is pH dependent. At neutral pH, histidine and tryptophan photooxidize faster than tyrosine (41). At low pHs, tryptophan and methionine are the most photoreactive amino acids (42).

Oxidation of methionine:

$$\underset{\text{methioine}}{RSCH_3} \rightarrow \underset{\text{methionine sulfoxide}}{RS(O)CH_3} \rightarrow \underset{\text{methionine sulfone}}{RS(OO)CH_3}$$

Oxidation of cysteine by oxygen:

$$RS^- + O_2 \rightarrow RS\cdot + O_2\cdot^-$$

$$RSH + O_2\cdot^- \rightarrow RS\cdot + HOO^-$$

$$RS\cdot + RS\cdot \rightarrow RS\text{-}SR$$

Thiol Disulfide Exchange

Thiol disulfide exchange occurs when a disulfide bond is reduced to two cysteines, and one of them reacts with another cysteine to form a new disulfide (5). The incorrect linkage of two cysteines in a disulfide bond can induce loss of biological activity of the protein. For example, interleukin 2 contains three cysteines at positions 58, 105, and 125. The native protein form has a disulfide linkage between the two cysteines at 58 and 105. In the presence of copper ions, two less active isomers with disulfide linkages between the cysteines at 58 and 125 and the cysteines at 105 and 125 are formed (5). The reaction is also base catalyzed and concentration dependent. The reaction rate can be influenced by pH, temperature, buffer composition, and the presence of other metal ions.

β-Elimination

Inactivation of proteins in an alkaline solution can result from β-elimination of the cystine residue, with the heterolytic cleavage of the disulfide and the formation of dehydroalanine and persulfide (equation 3). The persulfide can further react with hydroxide to form hydrosulfide (HS), as shown in equation 4 (5). The rate of β-elimination increases with increase in temperature and pH. The presence of metal ions also enhances the reaction rate.

$$R'CH_2SSCH_2R'' + OH^- \rightarrow R'CH_2SS^- + CH_2 = CR'' + H_2O \quad (3)$$

$$R'CH_2SS^- + OH^- \rightarrow R'CH_2SO^- + HS^- \quad (4)$$

PROTEIN FORMULATION DEVELOPMENT

After gathering all the background information (physicochemical properties, application, site of action, etc.) on the protein to be formulated, the first step that a formulation scientist usually takes is to characterize the protein in the initial formulation by analytical methods. During this step, analytical methods are also developed or optimized for use in formulation and stability studies. Because protein degradation is often dependent on pH and temperature, the second step of formulation development is to determine the relationship between the major degradation pathways of the protein and these parameters. Formulation scientists usually perform short-term studies of pH and elevated temperature (e.g., 25–40 °C) on initial liquid formulations. After determining the major degradation pathways in these formulations, the formulation scientist will select the most stable formulation that can achieve an optimum balance between the different degradation pathways. Finally, long-term stability testing is performed on the final chosen formulation to obtain real-time stability data. This section reviews the process of developing a stable formulation for protein pharmaceuticals. A simplified process diagram for protein formulation development is also illustrated in Figure 1.

Analytical Methods for Assessment of Protein Formulations

The selection of formulations for protein pharmaceuticals greatly depends on the stability results obtained in the formulation screening. Therefore, it is important to have suitable and sensitive analytical methods to study the degradation of proteins in formulations. Review articles by Jones provide detailed principles and applications of many analytical methods that can be used for protein characterization (43,44). In the following section, the commonly used analytical methods for determining protein formulation stability during formulation development are discussed. A summary of the analytical methods used for identifying protein degradation is shown in Table 1.

High Performance Liquid Chromatography. High performance liquid chromatography (HPLC) methods are widely used in the pharmaceutical industry for the analysis of peptides and proteins. The common HPLC methods for use in formulation development to assess protein stability are size exclusion chromatography (SEC), ion exchange chromatography (IEC), reversed-phase chromatography

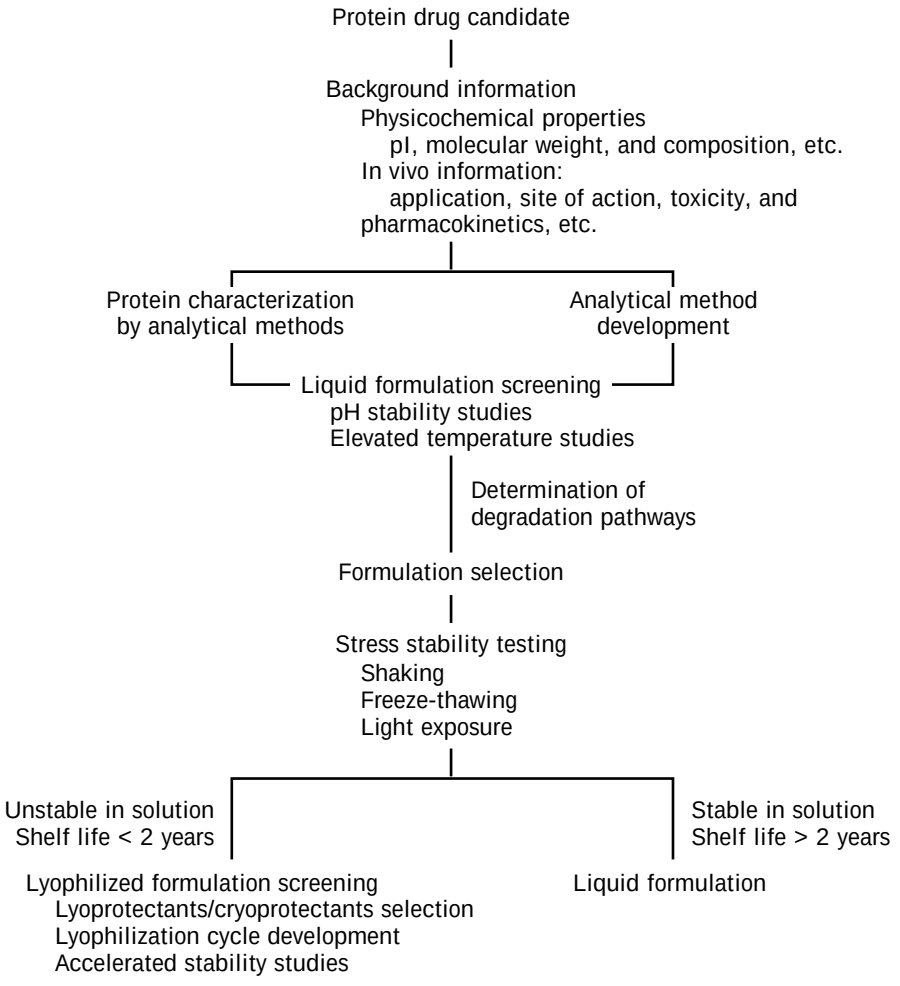


Figure 1. Simplified process diagram for protein formulation development.

(RPC), and hydrophobic interaction chromatography (HIC). Degradation of proteins via aggregation and cleavage results in the formation of soluble and insoluble aggregates (oligomers) and fragments, respectively. These soluble aggregates and fragments have different apparent molecular weights and can be separated from the intact protein on a size exclusion column that contains porous particles with different pore diameters and has been calibrated with molecular weight standards, usually globular proteins of known molecular weight.

IEC is often employed during formulation and stability studies for the separation of degraded proteins with charge heterogeneity resulting from deamidation or isomerization. Charged protein tends to bind to ion-exchanged resin (e.g., sulfopropyl, DEAE) of opposite charge by ionic interaction. Bound proteins can be eluted at their own critical salt concentration, allowing separation of the various charge forms (44). The amount of salt required to elute a charged species depends on its affinity for the resin (e.g., number of binding sites), the net charge, and the ability of the salt to displace the bound protein. By this method, deamidated forms of a protein that exhibit different net charge from each other and the native protein can be separated and quantified.

RPC methods are commonly used for the separation of small molecules and peptides, especially for peptides that result from proteolytic digestion of proteins in peptide mapping (45). A hydrophobic surface on the protein binds to the hydrophobic site of the resin (e.g., alkyl groups such as butyl [C4], hexyl [C6], or octadecyl [C18] or aromatics such as phenyl) in a reversed-phase column by hydrophobic interaction. Such hydrophobic interactions can be weakened by increasing the content of organic modifier in the mobile phase (46). Thus, proteins can be eluted in the order of their hydrophobic interaction strengths by an organic modifier such as acetonitrile. Although the application of RPC for analysis of recombinant proteins is less common than for peptides due to difficulties in resolution of large molecules, it has been used for characterization or separation, as in the case of met-rhGH and rhGH, which differ by the only one amino acid (47). This method is also useful in detecting the formation of more polar degraded proteins resulting from oxidation.

Similar to RPC, HIC is another useful technique for the separation of proteins based on their differences in surface hydrophobicity. Most proteins have some hydrophobic groups exposed to the surface. When the hydrophobic surface of a protein is in contact with the hydrophobic ligands on the beads of a column, water is released from the hydrophobic regions, and the entropy of the system is increased. The release of water increases the entropy of the

Table 1. Analytical Methods Used for Analysis of Protein Degradation During Formulation Development

Analytical method	Protein degradation/alteration
High performance liquid chromatography	
Size exclusion	Aggregation, polypetide cleavage
Ion exchange	Charge alteration (e.g., deamidation)
Reversed phase	Oxidation, disulfide alterations, charge alteration (using peptide map)
Hydrophobic interaction	Neutral alteration (e.g., Met oxidation)
Electrophoresis	
SDS-PAGE	Aggregation, polypeptide cleavage, disulfide alterations
Isoelectric focusing	Charge alteration (e.g., deamidation)
Capillary electrophoresis	
Capillary zone electrophoresis	Charge alteration (e.g., deamidation)
Capillary isoelectric focusing	Charge alteration (e.g., deamidation)
SDS-dynamic sieving capillary electrophoresis	Aggregation, polypeptide cleavage, disulfide alterations
Spectroscopy	
UV absorption	Tertiary structure alteration Change in protein concentration
Circular dichroism	Secondary structure alteration (far-UV CD) Tertiary structure alteration (near-UV CD)
Light scattering	Aggregation

system, causing the free energy to decrease on adsorption (48). Protein adsorption on an HIC column can be adjusted by changing the salt concentration. At high salt concentrations, protein adsorption is mainly due to hydrophobic interactions. By running the column with decreasing salt concentration, protein will be eluted from the column when the salt concentration is not high enough to maintain or stabilize the interaction.

Electrophoresis. Electrophoresis is a common method for the separation of proteins based on their net charge and size. Upon applying an electric field to a protein, it migrates toward the anode or cathode, depending on its net charge. There are two basic forms of electrophoresis: sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and isoelectric focusing (IEF). In SDS-PAGE, proteins are complexed with the denaturing agent SDS to give similar mass-to-charge ratios so that they can have free electrophoretic mobility. When SDS-complexed proteins are electrophoresed in a polyacrylamide gel with the appropriate pore size, their migration rate will be based on the size of the proteins. Thus, a mixture of proteins with different molecular weights can be separated using this method.

Separation of proteins by IEF is based on their isoelectric points (pIs), where the mobility of each protein in an electric field is zero and the net charge of the each protein is also zero. Because the mobility of nondenatured protein is dependent on its net charge, and this charge is determined by the pH of the solution, an electric field in the presence of a pH gradient causes the protein to migrate until the pH is the same as the pI (49). This technique is usually employed by preparing a polyacrylamide gel containing ampholytes of appropriate pI and buffer capacity to produce a pH gradient that is stable in the applied electric field. This method has been used to assess protein deamidation during stability studies of many pharmaceutical products such as recombinant human growth hormone (50).

Capillary Electrophoresis. Due to recent advances in technology, capillary electrophoresis (CE) is being more widely used in the biotechnology industry for protein analysis. The sensitivity and speed of CE make it a useful technique in early protein formulation development when only a small amount of material is available. CE provides a wide range of separation modes that can be performed with a single instrument. The basic separation principles are the same as with conventional electrophoresis techniques.

Capillary zone electrophoresis (CZE) is the most common CE separation method, and it employs a single buffer system in free solution. Separation of proteins using CE is based on their differences in mass-to-charge ratio. Separation takes place in free solution inside a capillary (usually made of fused silica) filled with electrolyte. Proteins with different mass-to-charge ratios are separated into individual zones as electrical force drives them at different rates through the capillary. As the mass-to-charge ratio increases, the electrophoretic mobility decreases, resulting in a longer migration time. Coated capillaries and buffer additives are often used to prevent protein adsorption and electroendosmotic flow. CZE has been used for the detection of the heterogeneity of recombinant proteins (51) as well as the separation of the deamidation products of growth-hormone-releasing factor (52).

Capillary isoelectric focusing (cIEF) is another separation mode in CE for the separation of proteins based on their isoelectric points. There are three basic steps in cIEF: sample injection, focusing, and mobilization. Proteins are premixed with ampholytes and injected into the capillary by pressure. High voltage is applied to the capillary during focusing so that sample ions can migrate to their neutral charge point in the pH gradient formed by the ampholytes along the capillary. After focusing, the focused protein bands are moved to the detection point to generate a signal. The on-line UV detection of cIEF provides a more accurate quantitation than the slab gel technique. A coated capillary is recommended for cIEF because the proteins stay in the capillary for focusing and mobilization. cIEF has been applied to the characterization of charge heterogeneity of recombinant monoclonal antibodies as a result of deamidation (53,54). The high resolving power of the method also makes it useful for the characterization of the

heterogeneity of the glycoforms of glycoproteins such as recombinant tissue plasminogen activator (55).

SDS-dynamic sieving capillary electrophoresis (SDS-DSCE) is a useful separation mode in CE for the molecular weight determination of proteins. Because all proteins have the same mass-to-charge ratio after complexation with SDS, the migration rate of proteins depends only on their molecular size. Although SDS-DSCE is equivalent to the conventional SDS-PAGE, it provides a rapid and automated system for quantitative analysis of proteins based on their molecular weight. Monoclonal antibodies degraded by the formation of aggregates and fragments can easily be separated by this method (54,56,57).

Spectroscopy. Ultraviolet (UV) absorption spectroscopic measurement is the most convenient and accurate method for the determination of protein concentration. Proteins with aromatic amino acids have a maximum in UV absorption between 275 and 282 nm at pH values below 8 (58). Thus, the protein concentration in a sample can be determined by measuring its absorbance at the wavelength maximum near 280 nm, and this absorbance is then compared with the absorbance of a solution with a known protein concentration or, if known, the molar extinction coefficient or absorptivity of the protein may be used. The absorptivity (often incorrectly described as the extinction coefficient) of a protein is defined as the absorbance of a 1 mg/mL protein solution through a 1-cm path at the wavelength maximum (e.g., near 280 nm). The molar extinction coefficient is the absorptivity of a 1.0 M solution of the protein through a 1.0-cm path. It is important to measure the protein concentration accurately for extinction coefficient determination. Quantitative amino acid analysis, nitrogen determination, or dry weight measurement are often used for this purpose.

Circular dichroism (CD) spectroscopy is a useful technique to evaluate the secondary and tertiary structures of protein (59,60). The secondary structure of a protein such as α-helices, β-sheets, and random coils can give rise to CD signals in the far-UV region (170–250 nm). CD spectroscopy is commonly employed for the study of denaturation or unfolding of proteins. The tertiary structure of a protein affects the local environment of the aromatic amino acids such as trytophan, tyrosine, and phenylalanine, as well as the disulfide bonds, and these effects contribute to CD signals in the near-UV region (240–320 nm). Because CD signals can be either positive or negative, it is not easy to obtain detail from CD spectra. However, a change in the far- or near-UV CD spectrum can represent a change in the secondary or tertiary structure of the protein, respectively. Therefore, CD is a useful tool to measure protein denaturation as well as aggregation, both of which result in structural or conformational changes.

Light scattering spectroscopy can be used to detect protein aggregation due to the fact that larger particles scatter more light per weight unit than small ones. Aggregated proteins can be either soluble or insoluble. Insoluble aggregates usually cause opalescence (turbidity) in a protein liquid formulation, yielding light scattering at wavelengths even in the visible region of the spectrum. Turbidity can also be conveniently measured in the near UV (340–360 nm) if no other chromophores are present in the same region. Thus, an increase in absorbance in this region often indicates the formation of aggregates. Correction for light scattering errors is necessary for the measurement of protein concentration made by the absorbance at the peak wavelength near 280 nm (61). Pharmaceutical final product solutions are required by the FDA to be inspected for clarity and degree of opalescence. Therefore, light-scattering determination may be substituted for the visual inspection method. Because native protein in solution also scatters light, light-scattering spectroscopy can be used to estimate the molecular weight of a protein. Dynamic laser light scattering, also known as photon correlation spectroscopy (PCS) or quasi-elastic light scattering (QELS), can be used to calculate the average hydrodynamic diameter and determine the diffusion coefficient of a protein by autocorrelation methods (62,63).

Formulation Screening Methods

The methods commonly used by formulation scientists for protein formulation screening to select the best candidate formulation include the short-term accelerated pH and temperature stability studies. Accelerated conditions are defined as the extremes of the parameter range (e.g., 40 °C and high ($\geq$9.0) or low ($\leq$3.0) pH).

Accelerated pH Stability Studies. Solution pH often plays an important role in the stability of the protein product because most of the major protein degradation pathways are pH dependent. Thus, the stability of a protein in different formulations can be compared by performing stability studies on the protein formulated at a range of pH using a variety of buffer solutions. Examples of buffers that are used to maintain the pH of formulations in commercially available recombinant protein drug products are listed in Table 2. Acetate, succinate, citrate, and phosphate are the most common buffers used in recombinant protein formulations. For rapid formulation screening, a pH stability study is usually performed in a short period of time and at an elevated temperature of 40 °C because the rate

Table 2. Examples of Buffers Commonly Used in Recombinant Protein Formulations for pH Maintenance

Buffers	pH	Commercial products[a]
Sodium acetate	4.0	Neupogen® (Amgen)
Sodium succinate	5.0	Actimmune® (Genentech)
Sodium citrate	6.0	Nutropin® AQ (Genentech)
	6.5	Rituxan® (IDEC, Genentech)
	6.9	Epogen® (Amgen)
L-Histindine HCl	6.0	HERCEPTIN® (Genentech)
Sodium phosphate	7.3	Intron-A® (Schering-Plough)
	7.4	Protropin® (Genentech)
	7.5	Humatrope® (Lilly)
	7.0	Orthoclone OKT®3 (Ortho)
	7.5	Proleukin® (Chiron)
	7.4	Nutropin® (Genentech)
Phosphoric acid	7.3	Activase® (Genentech)

[a]Recombinant biological products approved for marketing in the United States. *Source*: Ref. 64.

of degradation is slow at low temperature. Often, a pH stability profile for each formulation can be developed by plotting the rate constants of all degradation reactions versus pH. The most stable protein formulation selected should be the one with a pH at which the overall degradation reactions are minimal.

Accelerated Temperature Stability Studies. Because temperature also has effects on protein degradation reactions, stability testing at accelerated temperatures is often carried out by formulation scientists for formulation screening, and the degradation rate profiles in various formulations are compared using Arrhenius plots. An Arrhenius plot allows one to predict the degradation rate constants at storage temperatures other than those actually studied, assuming that the protein does not undergo physical changes at a given temperature. Formulation scientists often use the Arrhenius plot to extrapolate results obtained at higher temperatures to 5 °C for product shelf life prediction. However, at least three temperatures are required to obtain an Arrhenius plot, and extrapolation of these data should be done with understanding of the assumptions made regarding the linearity of the response and the accuracy of the rate constants. Small-molecule drugs that often degrade via a single degradation pathway usually follow Arrhenius behavior. However, due to complexity in protein structure, more than one degradation reaction can occur in a protein product. This may result in nonlinear Arrhenius behavior. Sometimes, the degradation pathways for proteins at higher temperatures are not the same as those at lower temperatures. For example, interleukin 1β undergoes deamidation at or below 30 °C and aggregation at or above 39 °C (65). In this case, the formulation selected based on accelerated stability results may not be stable enough to protect the protein from degradation at lower temperatures. However, accelerated stability testing can still be a useful tool for protein formulation screening if the protein primarily degrades via only one degradation pathway within a selected temperature range, or if the degradation pathways are not influenced by one another (e.g., deamidation rate not affected by oxidation).

Stress Testing

After a stable protein formulation is selected from the formulation screening studies, stress tests are usually performed to determine whether additives are required to stabilize the protein against stresses such as agitation, freeze-thawing, and light exposure.

Agitation Studies. Proteins are susceptible to denaturation by mechanical force such shaking or agitation. Shaking of liquid protein products can occur during filling, shipping, and handling. The mechanism of denaturation may involve an increase in area of the air–liquid interface during shaking. Proteins have a tendency to unfold and expose their hydrophobic groups at an air–liquid interface, resulting in denaturation and the formation of aggregates. Proteins also tend to concentrate at the air–liquid interface due to the presence of both polar and nonpolar side chains. Irreversible unfolding or aggregation of the protein occurs when the interface surface area increases (66). Henson explained that the turbidity observed in some protein solutions during vigorous shaking could be due to the formation of aggregates when the air–liquid interface increases (67). Formulation scientists often perform agitation studies on liquid protein formulations to determine if a surfactant is required to protect the protein from denaturation when agitation is unavoidable during manufacturing, shipping, and handling. The addition of surfactants could either reduce the interfacial tension or solubilize the concentrated protein at the interface (2). Surfactants commonly used as stabilizers in the pharmaceutical industry include polysorbate 20, polysorbate 80, pluronic 68, and brij, which are polymers composed of various combinations of ethylene oxide, propyl oxide, sorbitan, and fatty alcohols.

Freeze–Thaw Studies. Under many circumstances, such as processing, storage, shipping, and stability testing, freezing and freeze-thawing stresses may be applied to proteins. For instance, it is very common for a liquid protein bulk to be manufactured and stored as frozen bulk for long-term storage and then thawed before filling into containers for packing as a liquid protein product or for freeze-drying as a lyophilized product. Sometimes, a protein formulation is frozen as a concentrated bulk for storage and then thawed and diluted with placebo before use. The risks of accidental freezing and thawing of proteins in liquid formulations may also exist during shipping and handling. Formulation scientists often perform short-term stability testing such that samples at each timepoint are kept frozen and then thawed at the end of the study for analysis. In lyophilized formulations, proteins are subjected to freezing during the lyophilization process. Therefore, freeze-thawing stability studies (at least three cycles of freezing and thawing) are usually performed on potential formulation candidates during formulation development to ensure the final product is stable upon freeze-thawing.

Protein denaturation during freezing is mainly caused by the phase separation due to the formation of ice. When ice is formed in a liquid protein formulation, the local concentration of protein as well as all excipients increases dramatically. The concentrated excipients in the non-ice phase may become destabilizing agents and cause the protein to denature (68). Some buffer species such as sodium phosphate and sodium succinate exhibit a decrease in pH during freezing. Freezing-induced pH changes may cause instability of recombinant proteins, as observed in IFN-γ formulated in succinate buffer (69).

If a protein formulation is evaluated to be denatured during freeze-thawing studies, formulation scientists may consider adding cryoprotectants such as sugars, polyols, amino acids, and inorganic and organic salts to prevent denaturation. The preferential exclusion mechanism of protection of proteins by these cryoprotectants during freeze-thawing was proposed by Timashelf and his coworkers (70). They suggested that the cryoprotectants, acting as cosolutes in the solution, are preferentially excluded from the protein. Consequently, the protein itself is preferentially hydrated by water. Hydration of a protein plays an important role in maintaining the conformational struc-

ture. Water contributes to the hydrogen bonds and hydrophobic interactions that are responsible for the formation of the secondary structure as well as maintaining the protein in a folded or native form (tertiary structure).

Light Stability Studies. Some amino acid residues in proteins can absorb energy from UV light and become modified to form photooxidized degradation products, especially in the presence of photosensitisers such as molecular oxygen, riboflavin, dyes, and polymers containing ether linkages (e.g., polysorbates). For example, the methionine residue in recombinant human relaxin can undergo photooxidation to form methionine sulfoxide (71), and the tryptophan residue in monoclonal IgG is converted to kynurenine and *N*-formylkynurenine upon exposure to UV light (72). The degree of photolytic degradation in the protein formulation can be influenced by many factors, including the buffer species and its concentration, excipients, and the formulation pH. Light intensity, duration of exposure, and storage temperature can also affect the rate of photolytic degradation.

Light stability testing is often performed during protein formulation development. The aim of a photostability study is to accelerate the formation of photodegradation products, if any, in a particular formulation upon exposure to light. FDA guidelines call for light stability testing of finished drug formulations but do not provide specifics of how to perform the test. A recent review article by Nema et al. provides an extensive overview of the protocols currently being used to perform photostability studies in pharmaceutical industry (73). Fluorescent lamps are usually used by pharmaceutical formulation scientists to simulate sunlight in photostability studies. Finished drug products are stored in a light cabinet with light intensity of 2–180 klx. The total exposure is between 8 and 4,500 klx days. Temperature inside the light cabinet is maintained at 25 ± 2 °C. Control samples, wrapped in aluminum foil, are exposed concurrently with the test samples.

Photolytic degradation of protein can be prevented by adding antioxidants to the formulation. Antioxidants can be classified into four main categories: chelating agents, reducing agents, oxygen scavengers, and chain terminators (74). If the addition of antioxidants in a protein formulation does not inhibit oxidation, the use of light-resistant containers is an alternative way of protecting light-sensitive protein drugs.

Lyophilized Formulation

Protein products that are unstable in solution or whose most-stable liquid formulations have insufficient long-term storage stability for marketing require lyophilization. However, the lyophilized products are usually limited to reconstitution and use in a short period of time. Lyophilized products also require longer time and more labor for formulation development than the ready-to-use liquid products. Lyophilization is also expensive due to the added costs of drying. Therefore, lyophilization is usually considered as an alternative method in protein formulation development when success with a liquid formulation cannot be achieved.

Lyophilization, also known as freeze-drying, is a process usually employed in the protein pharmaceutical industry to convert protein liquid formulations into solids of sufficient stability for storage. The process involves three major steps: freezing, primary drying, and secondary drying. During the initial freezing step, solvent (usually water) from an aqueous solution of protein drug that has been filled into glass vials is frozen and crystallized into solid ice at a low temperature such as −40 °C in a freeze dryer. The solutes including the protein and excipients are usually converted into an amorphous solid phase, which consists of uncrystallized solutes and uncrystallized water below their glass transition temperature, although some excipients such as buffer salts and mannitol may crystallize (75). The second step of the freeze-drying process, in which ice is removed from the frozen product by sublimation and condensation, is called primary drying. After the entire protein solution has been solidified, vacuum is applied to the drying chamber, and the shelf temperature is increased to supply heat to sublime ice from the frozen solid. Water vapor formed as a result of sublimation is condensed on cold surfaces in the condenser chamber. Primary drying is the longest part of the freeze-drying process and may require several days. After the bulk water (ice) has been removed from the partially dried product, secondary drying begins. During the stage of secondary drying, the final step of the freeze-drying process, the shelf temperature is increased to provide an elevated product temperature to remove the unfrozen water remaining in the amorphous solid phase. Secondary drying is usually performed over several hours at a shelf temperature between 25 and 50 °C, and the product temperature is maintained at about 25–35 °C.

The successful development of a lyophilized protein product requires the development of a liquid formulation that is stable to the freeze-drying process. Proteins can undergo denaturation or aggregation during both freezing and drying. The effects of freezing and drying on protein conformation described by Arakawa and Carpenter are illustrated in Figure 2. Formulation scientists often consider adding cryoprotectants and lyoprotectants to formulations to protect the protein from damage by freezing and drying, respectively, in the lyophilization process. Often, lyoprotectants improve the stability of the lyophilized products during subsequent storage. During freezing, all solutes in a formulation are concentrated when ice is formed, and buffer salts may crystallize, resulting in a pH shift (69). These freezing stresses can induce irreversible conformational changes that lead to protein instability. Thus, the use of buffers at low concentration and with minimal pH shift during freezing is recommended for formulations to be prepared by the freeze-drying process (76). The addition of cryoprotectants such as polymers, carbohydrates, and amino acids can provide cosolutes to preserve the native structure of protein in the frozen state. The mechanism for solute-induced protein stabilization during freezing proposed by Arakawa et al. is applicable to these conditions as just described. Amorphous sugars such as trehalose, sucrose, and lactose are the most effective lyoprotectants for stabilizing proteins during drying and storage. Carpenter and his coworkers have found that the formation of an amorphous solid is required for preservation of proteins by sugars, and some sugars such as

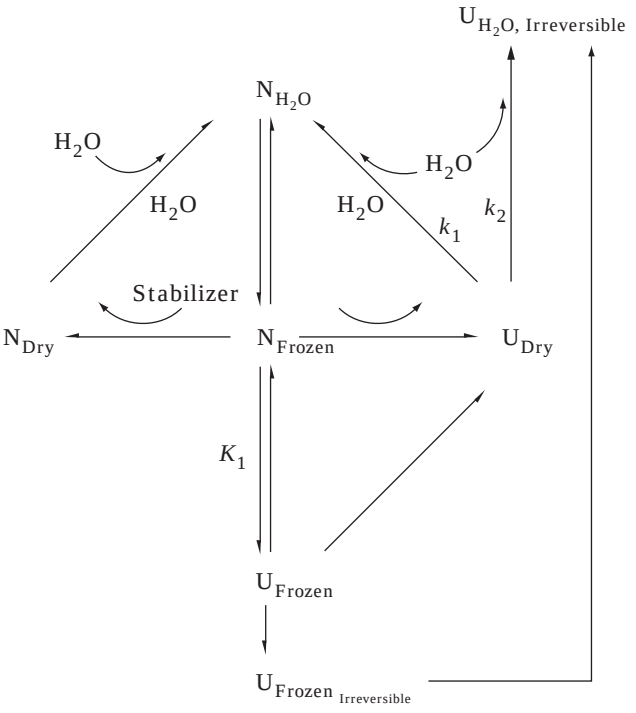


Figure 2. Schematic representation of the conformation of a protein during freezing, drying, and rehydration. N (native), U (unfolded), K_1 (conformational equilibrium formed upon freezing), k_1 (rate constant for refolding), k_2 (rate constant for formation of irreversibly denatured forms). *Source*: Reprinted with permission from Ref. 78.

mannitol fail to protect labile proteins during freeze-drying because the sugar forms a separate crystalline phase during freezing (77). It has been hypothesized that lyoprotectants stabilize proteins by acting as a water replacement through hydrogen bonding to the dried protein, preventing conformational changes during drying (78).

In developing a freeze-drying cycle for lyophilized protein product, temperature control is important in each step of the process. During freezing, the protein solution is frozen so that most of the water is converted to ice and all solutes are converted to either crystalline solids or amphorous solids below their glass transition temperature (79). The temperature at which the concentrated solute crystallizes during freezing is called the eutectic temperature. If the solute crystallizes, the shelf temperature must be reduced to well below the lowest eutectic temperature to ensure complete crystallization. For an amorphous solute, the shelf temperature must be lowered below the glass transition temperature (79). Typically, the final product temperature during freezing is about −40 °C. Usually, rapid ice growth as a result of low freezing temperature produces a high degree of supercooling (79,80). The degree of supercooling determines the number and size of ice crystals and therefore determines the drying characteristics and appearance of the dried cake. A very high degree of supercooling will result in very small pores, which means slower primary drying. Therefore, a very high degree of supercooling is undesirable. The ideal freezing method is to provide a moderate supercooling and rapid ice growth. The time required for primary drying decreases by a factor of 2 for a 5 °C increase in product temperature (79). For minimizing process time, it is important to maintain product temperature as high as possible during primary drying without affecting the product quality. Usually, the product temperature in primary drying is controlled at a safe margin (2–5 °C) below the collapse temperature. After all ice has been removed, the product temperature is usually increased to about 25–35 °C in the secondary drying step to allow the removal of unfrozen water remained in the product. To avoid collapse, the product temperature should not be raised until all the ice has been removed during primary drying.

Long-term Stability Studies.

Long-term premarket stability studies of protein formulations are designed to establish the specific storage conditions and expiration dating for the drug product. It is usually required by the FDA that a pharmaceutical product contain not less than 90% of the label claim of active ingredient at the expiration date. Although real-time stability data are required by the FDA to demonstrate that the pharmaceutical product can meet this requirement prior to approval, an extrapolated expiration date based on the stability results obtained from clinical batches is usually allowed. Thus, stability-indicating assays are important for determining the stability of drug products. Analytical methods such HPLC, electrophoresis, and spectroscopy have been widely used for this purpose.

In designing a long-term stability protocol for drug potency testing, one should consider the storage conditions and storage intervals. For refrigeration (5 °C) and room temperature (15–30 °C, by U.S.P. definition) storage, samples should be stored at 5, 15, and 25–30 °C and monitored at 0, 1, 2, 3, 6, 9, 12, 18, and 24 months, and once yearly after. For lyophilized products, samples can be assayed at longer storage intervals and at higher storage temperatures such as 40 °C. Pharmaceutical drugs are often monitored at high temperatures for the estimation of the degradation constants at low temperatures by Arrhenius plots. For high-temperature storage, samples are usually stored at 37 or 40 °C and assayed at 1, 3, and 6 months. Samples may also be stored at 45 or 50 °C and assayed at 2 weeks and 1 and 2 months. However, these elevated temperatures may approach the denaturation temperature of the protein, resulting in physical denaturation of the protein that would not be encountered in real situations, or extrapolate to lower-temperature behavior. In addition to determining the amount of active ingredient in a drug product, evaluation of the product should also include the following: pH, ionic strength, appearance, color, clarity, particulate matter, sterility, and pyrogenicity.

PROTEIN DRUG DELIVERY

Typically, proteins and many peptides must be administered by injection because of their poor bioavailability (fraction of administered drug in the circulation) by other routes. Initial clinical studies of these molecules is often conducted by dosing the drug i.v., intramuscularly (i.m.),

or subcutaneously (s.c.). These clinical studies assess the safety and tolerability of systemic exposure to the new drug. The results from early clinical trials combined with preclinical studies in animal models may indicate the need for a drug delivery system. Drug delivery systems are often required when the efficacious dose is not well tolerated, indicating a requirement to localize the drug at the site of action or maintain a low systemic steady-state level (e.g., provide same total drug dose at a lower maximum serum concentration). Localization of the drug to the site of action may also be necessary due to biological barriers preventing sufficient delivery of the drug via the systemic circulation (e.g. blood–brain barrier, blood–ocular barrier). In some situations, the disease state may be more readily accessible to an alternative delivery method than traditional parenteral administration. In the case of cystic fibrosis, a disease of the lungs, aerosol delivery of rhDNase has proven to be an effective therapy (81). Another rationale for using a drug delivery system is patient compliance or ease of use. For example, if a patient is required to have multiple injections of a drug per day or per week, it will be difficult to maintain proper compliance unless the drug effects are immediately felt by the patient and/or the disease is debilitating. These issues along with the in vitro (physicochemical properties) and in vivo (pharmacokinetics, pharmacodynamics, and toxicology) characteristics of the drug must be well understood prior to embarking on the development of a drug delivery system.

Selection of a Delivery System

Several potential options are available for drug delivery, as shown in Table 3 (82). These options include both invasive methods involving injections or noninvasive approaches that may or may not provide sufficient bioavailability. In each case, a stable formulation must first be developed, as described in the previous sections. For invasive methods, drug delivery may be applied with either devices or biodegradable depots. Delivery devices such as syringe pens and needle-less injectors (injection by high-pressure gas) are gaining acceptance for many frequently administered drugs (83,84). In addition, implantable pumps are currently used with protein drugs such as insulin (85). Several injectable depot formulations made from the biodegradable polymer, poly(lactic-coglycolic acid) (PLGA), are either approved (86,87) or in clinical trials (88). These depot systems offer the advantage of eliminating multiple injections and localizing drug to the site of action, if necessary. However, a major challenge in the development of these systems is the production of sterile injectable depots, usually microspheres on the order of 20 μm or greater in diameter. Approaches to developing microsphere depot systems for clinical and commercial use have recently been reviewed (89,90).

For noninvasive administration, aerosol delivery of proteins and peptides is the most developed and accepted delivery option. Aerosol delivery offers the unique option of delivering the drug through a single inhalation instead of a single injection. The lung provides a large surface area for absorption of the drug, but the aerodynamic properties of the drug (liquid or solid) play a key role in the deposition and, ultimately, the amount of drug delivered systemically. Several key issues such as drug stability and device selection have been reviewed recently (91). This work demonstrates that if the key parameters are analyzed and the appropriate drug formulation and device are selected, then aerosol delivery of proteins may be an effective method of administration. In contrast, the bioavailability of the protein or peptide when delivered by the other noninvasive routes is usually quite low ($<$1%), indicating that they are not an effective method for achieving significant systemic levels of the drug (for a review see Ref. 92). However, they may be very effective for local delivery of protein drugs if a low dose is efficacious.

Protein Stability in a Delivery System

After selecting a delivery system, the formulation scientist is required to develop a stable formulation that is compatible with the system and its intended use. These studies usually involve the formulation development worked described in the previous sections, as well as protein–delivery system interaction studies. The type of studies undertaken to assure delivery of biologically active and nonimmunogenic protein depend upon the system characteristics.

For injection devices, it is necessary to assess the protein stability in the container used to store the protein (e.g., cartridges) and the effect of the injection itself on the protein stability. These studies often involve storage stability studies of the protein in the device container as already described for protein filled into vials. The protein delivered from the device is also assessed by injecting into an empty sample tube or a buffer solution. Both recombinant human growth hormone (rhGH) (93) and insulin (94) are currently available in injection devices. Typically, high-pressure devices such as the needle-less injectors may cause denaturation of the protein due to the air–water interface and the high shear force applied to the liquid. Surface-induced denaturation may also occur at the device surface or at an air–water interface in the device. This degradation can be reduced and potentially eliminated by the addition of surfactants (e.g., polysorbates). In addition, implantable pump devices require studies to demonstrate maintenance of protein stability during the protein's residence time in the pump at physiological temperatures ($\sim$37 °C). Several studies have been performed with insulin in these pumps (85), indicating the difficulty in designing stable formulations for these pumps. These systems pose a formidable challenge due to the potential for accumulation of degraded protein over time with repeated use of a refillable pump.

Similar issues are also encountered in the development of the injectable depot formulations. In this case, the protein remains in a biodegradable depot at physiological conditions for prolonged periods. The protein must maintain its potency during its in vivo release from the depot. Recent studies with rhGH revealed that while it degrades during incubation at physiological pH, temperatures, and ionic strength, it maintains its biological activity and is not immunogenic (95,96). This result, however, cannot be assumed for each protein. Therefore, stability studies at physiological conditions should be performed prior to the development of a depot formulation to define the feasibility of the system and the acceptable duration of release (i.e.,

Table 3. Routes of Delivery for Proteins and Peptides

Delivery routes	Formulation and device requirements	Commercial products[a]
	Invasive	
Direct Injection	Liquid or reconstituted solid, syringe	
Intravenous		Activase®
Subcutaneous		Nutropin®
Intramuscular		Recombi Vax®
Intracerebral vein		
Depot system	Biodegradable polymes, liposomes, permeable polymers (not degradable), microspheres, implants	
	[LHRH analogs (subucutaneous or intramuscular)]	Lupron Depot®
		Zoladex®
		Decapeptyl®
	Noninvasive	
Pulmonary	Liquid or powder formulations, nebulizers, metered dose inhalers, dry powder inhalers	Pulmozyme®
Oral	Solid, emulsion, microparticulates, absorption enhancers	
Nasal	Liquid, usually requires permeation enhancers	Synarel®
Topical	Emulsion, cream or paste (liposomes)	
Transdermal	Electrophoretic (iontophoresis), electroporation, chemical permeation enhancers, prodrugs, ultrasonic	
Buccal, rectal, vaginal	Gels, suppositories, bioadhesives, particles	

Source: Ref. 8 with permission from the American Chemical Society.
[a]Nutropin® (recombinant human growth hormone), Activase® (recombinant human tissue plasminogen activator), and Pulmozyme® (recombinant human deoxyribonuclease I) are all products of Genentech, Inc. RecombiVax® (recombinant hepatitis B surface antigen) is produced by Merck & Co. Lupron Depot® (leuprolide acetate–PLGA) is a product of Takeda Pharmaceuticals. Zoladex® (goserelin acetate–PLGA) is produced by the Imperial Chemical Industries, Ltd. Decapeptyl® is a manufactured by Debiopharm. Synarel® (nafarelin acetate) is made by Syntex Corporation.

what level of degradation can be tolerated in terms of both potency and immunogenicity). These systems also pose a challenge in protein stability during the manufacture of the depot. For example, PLGA formulations require the dissolution of the polymer in organic solvents or use of high temperatures to melt the polymer, followed by dispersion of the protein in the PLGA and extraction of the organic solvent or reduction of the temperature, respectively. This exposure to harsh conditions often causes denaturation of the protein. Screening studies are performed to select excipients that will stabilize the protein against denaturation under these conditions. An example of this type of study was recently described using the model proteins rhGH and recombinant human IFN-γ (97). These studies revealed that excipients (e.g., sugars), which cause preferential hydration of the protein (cause an increased water layer at the protein surface), also increase protein stability during exposure to organic solvents. If these stability issues are overcome, an injectable depot system such as PLGA microspheres may be developed and tested in clinical trials.

Unlike depot systems, aerosol formulations for proteins do not require long-term stability at physiological conditions or stability in organic solvents. However, the aerosolization of proteins in their liquid state often leads to some protein degradation due to the air–water interface as well as the shear stress generated during the process. Several studies have been performed with traditional nebulizer devices that generate aqueous droplets of the appropriate diameter for efficient delivery to the lung. For rhGH, a surfactant was required to stabilize the protein during the process of aerosol formation (98), whereas rhDNAse was very stable in an unbuffered isotonic solution (1). In many cases, it may not be possible to develop an aqueous protein formulation that both prevents degradation and is not an irritant to the lung. In addition, devices required to generate liquid droplets are usually not easily portable, making them less attractive to patients that are not hospitalized or bed-ridden. However, dry powder protein formulations may be delivered in small hand-held devices. To produce a protein powder with the proper aerodynamic properties, spray-drying of the protein is usually performed (99). This process usually involves high temperatures and aerosolization of the protein, both of which may cause denaturation, as observed for rhGH (100). Once again, a formulation must be developed to prevent this denaturation. In addition, the dry powder protein formulation may be required to have stability at controlled room temperature (15–30 °C) and variable levels of residual humidity. The formulation scientist then carries out formulation screening experiments at different temperatures and residual humidities. If a protein powder is feasible, it offers advantages in ease of use and, perhaps, greater long-term stability.

Manufacturing Issues

The development of drug delivery systems may also require elaborate manufacturing procedures due to the complexity of both the system and the protein. When considering the manufacture of injection devices such as pens and pumps, some of the key issues are the ability to per-

form a sterile filling operation, the stability in the device, and the interaction between the drug, excipients, and the device or manufacturing components. The level of complexity is further increased in the case of pulmonary delivery. A nebulized liquid formulation is usually filled into ampules or other containers that allow easy addition of the liquid to the nebulizer by the patient. The filling of plastic ampules also has unique challenges in the ability to maintain sterility of the final product, the permeability of the material to gases and volatile solvents, and the physical properties of the container. In contrast, if a dry powder is used as the aerosol, the manufacturing process requires large-scale spray-drying equipment and machines for filling the dry powder into blister packs that can be used for metered-dose inhalers (99).

Manufacturing complexity is further increased when considering depot formulations such as PLGA microspheres. These microspheres are typically 20–90 μm in diameter and are therefore not readily sterilized by the traditional filtration methods (e.g., 0.22-μm filters). Terminal sterilization of these microspheres causes degradation of the polymer backbone, altering the release properties and often causing significant degradation of the protein within the microspheres. Microspheres must then be produced aseptically to avoid terminal sterilization while still maintaining the delivery of a sterile product to the patient. An aseptic process may be achieved through the use of barrier technology (isolators) for raw material input into the process and final product removal from the process. The remainder of the process operations is then designed as steam-in-place systems to assure a sterile environment throughout the production process.

If the unique challenges of a drug delivery system are overcome for a given protein, protein delivery may be successfully achieved. To reach commercial use, the drug delivery system must be well matched with the clinical and marketing requirements as well as the physiological requirements of the protein.

BIBLIOGRAPHY

1. D.C. Cipolla, I. Gonda, K.C. Meserve, S. Weck, and S.J. Shire, in J.L. Cleland and R. Langer eds., in *Formulation and Delivery of Proteins and Peptides*, American Chemical Society Symposium Series, ACS Books, New York, 1994, pp. 322–342.
2. Y.-C.J. Wang and M.A. Hanson, *J. Parenter. Sci. Technol.* **42**, S1–S26 (1988).
3. M.C. Manning, K. Patel, and R.T. Borchardt, *Pharm. Res.* **6**, 903–918 (1989).
4. T. Chen, *Drug Dev. Ind. Pharm.* **18**, 1311–1354 (1992).
5. T.J. Ahen and M.C. Manning eds., *Stability of Protein Pharmaceuticals, Part A: Chemical and Physical Pathways of Protein Degradation*, vol. 2, Plenum, New York, 1992.
6. T.J. Ahen and M.C. Manning eds., *Stability of Protein Pharmaceuticals. Part B: In vivo Pathways of Degradation and Strategies for Protein Stabilization*, vol. 3, Plenum, New York, 1992.
7. J.L. Cleland, M.F. Powell, and S.J. Shire, S. Bruck ed., in *Critical Reviews in Therapeutics Drug Carrier Systems*, CRC Press, Boca Raton, Fl., 1993, pp. 307–377.
8. R.T. Sauer and W.A. Lim, *Curr. Opin. Struct. Biol.* **2**, 46–51 (1992).
9. W.A. Lim and R.T. Sauer, *J. Mol. Biol.* **219**, 359–376 (1991).
10. A.E. Eriksson, W.A. Baase, X.-J. Zhang, D.W. Heinz, M. Blaber, E.P. Baldwin, and B.W. Matthews, *Science* **255**, 178–183 (1992).
11. K.A. Dill, *Biochemistry* **29**, 7133–7155 (1990).
12. K.A. Drill, *Biochemistry* **24**, 1501–1509 (1985).
13. C.N. Pace, *Trends Biochem. Sci.* **15**, 14–17 (1990).
14. T. Kiefhaber, R. Rudolp, H.-H. Kohler, and J. Buchner, *Biotechnology* **9**, 825–829 (1991).
15. M.R. DeFelippis, L.A. Alter, A.H. Pekar, H.A. Havel, and D.N. Brems, *Biochemistry* **32**, 1555–1562 (1993).
16. J.L. Cleland and D.I.C. Wang, *Biochemistry* **29**, 11072–11078 (1990).
17. H.A. Havel, E.W. Kauffman, S.M. Plaisted, and D.N. Brems, *Biochemistry* **25**, 6533–6538 (1986).
18. S. Tandon and P.M. Horowitz, *J. Biol. Chem.* **264**, 9856–9866 (1989).
19. T. Arakawa, N.K. Alton, and Y.-R. Hsu, *J. Biol. Chem.* **260**, 14435–14439 (1985).
20. D.A. Yphantis and T. Arakawa, *Biochemistry* **26**, 5422–5427 (1987).
21. A.B. Robinson and C.J. Rudd, in B.L. Horecker and E.R. Stadtman eds., *Current Topics in Cellular Regulations*, vol. 8, Academic, New York, 1974, pp. 247–295.
22. A.B. Robinson, J.W. Schotchler, and J.H. McKerrow, *J. Am. Chem. Soc.* **95**, 8156–8159 (1973).
23. S.A. Berson and R.S. Yalow, *Diabetes* **15**, 875–879 (1966).
24. B.V. Fisher and P.B. Porter, *J. Pharm. Pharmacol.* **33**, 203–206 (1981).
25. U.J. Lewis, E.V. Cheever, and W.C. Hopkins, *Biochim. Biophys. Acta* **214**, 498–508 (1970).
26. G.W. Becker, P.M. Tackitt, W.W. Bromer, D.S. Lefeber, and R.M. Riggin, *Biotechnol. Appl. Biochem.* **10**, 326–337 (1988).
27. E.E. Haley, B.J. Coreoran, F.E. Dorer, and D.L. Buchanan, *Biochemistry* **5**, 3229–3235 (1966).
28. T. Geiger and S. Clarke, *J. Biol. Chem.* **262**, 785–794 (1987).
29. K. Patel, C. Oliyai, R.T. Borchardt, and M.C. Manning, *J. Cell. Biochem.* **13A**, 88 (1989).
30. G. Teshima, J.T. Stults, V. Ling, and E. Canova-Davis, *J. Biol. Chem.* **266**, 13544–13547 (1991).
31. B.A. Johnson, J.M. Shirokawa, W.S. Hancock, M.W. Spellman, L.J. Basa, and D.W. Aswad, *J. Biol. Chem.* **264**, 14262–14271 (1989).
32. B.N. Violand, M.R. Schlittler, P.C. Toren, and N.R. Siegel, *J. Protein Chem.* **9**, 109–117 (1990).
33. K. Sasaoki, T. Hiroshima, S. Kusumoto, and K. Nishi, *Chem. Pharm. Bull.* **37**, 2160–2164 (1989).
34. T.H. Nguyen, J. Burnier, and W. Meng, *Pharm. Res.* **10**, 1563–1571 (1993).
35. Y. Nabuchi, E. Fujiwara, K. Ueno, H. Kuboniwa, Y. Asoh, and H. Ushio, *Pharm. Res.* **12**, 2049–2052 (1995).
36. L.C. Teh, L.J. Murphy, N.L. Huq, A.S. Surus, H.G. Friesen, L. Lazarus, and G.E. Chapman, *J. Biol. Chem.* **262**, 6472–6477 (1987).
37. J.Q. Oeswein and S.J. Shire, in V.H.L. Lee ed., *Peptide and Protein Drug Delivery*, Advances in Parenteral Sciences 4, Marcel Dekker, New York, 1990, pp. 167–202.
38. T.H. Nguyen, in J.L. Cleland and R. Langer eds., *Formulation and Delivery of Proteins and Peptides*, American Chemical Society Symposium Series, ACS Books, New York, 1994, pp. 59–71.
39. L. Philipson, *Biochim. Biophys. Acta* **56**, 375–382 (1962).
40. C. Little and P.J. O'Brien, *Eur. J. Biochem.* **10**, 533–538 (1969).

41. L.A. Holt, B. Milligan, D.E. Rivett, and F.H.C. Stewart, *Biochim. Biophys. Acta* **499**, 131–138 (1977).
42. L.A. Sluyterman, *Biochim. Biophys. Acta* **60**, 557–561 (1962).
43. A.J.S. Jones, in J.L. Cleland and R. Langer eds., *Formulation and Delivery of Proteins and Peptides*, American Chemical Society Symposium Series, ACS Books, New York, 1994, pp. 22–45.
44. A.J.S. Jones, in G. Poste and E. Tomlinson, eds., *Advanced Drug Delivery Reviews*, Elsevier Science Publishers B.V., Amsterdam, 1993, pp. 29–90.
45. C.S. Fullmer and R.H. Wasserman, *J. Biol. Chem*. **254**, 7028–7212 (1979).
46. P.H. Cowan, in R.W.A. Oliver ed., *HPLC of Macromolecules: A Practical Approach*, IRL Press, Oxford, 1989, pp. 127–156.
47. W.J. Kohr, R. Keck, and R.N. Harkins, *Anal. Biochem*. **122**, 348–359 (1982).
48. T.W. Perkins, D.S. MaK, T.W. Root, and E.N. Lightfoot, *J. Chromatogr. A* **766**, 1–14 (1997).
49. P.G. Righetti, *Isoelectric Focusing: Theory, Methodology and Applications*, Elsevier Science Publishers B.V., Amsterdam, 1993.
50. R. Pearlman and T.H. Nguyen, in *Therapeutic Peptides Proteins: Formulation, Delivery and Targeting*, D. Marshak and D. Liu eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989, pp. 23–30.
51. B.L. Karger, A.S. Cohen, and A. Guttman, *J. Chromatogr.* **492**, 585–614 (1989).
52. L. Bongers, T. Lambros, A.M. Felix, and E.P. Heimer, *J. Liquid Chromatogr.* **15**, 1115–1128 (1992).
53. M.A. Costello, C. Woititz, J. DeFeo, D. Stremlo, L.-F.L. Wen, D.J. Palling, K. Iqbal, N.A. Guzman, *J. Liquid Chromatogr.* **15**, 1081–1097 (1992).
54. G. Hunt, K.G. Moorhouse, and A.B. Chen, *J. Chromatogr. A* **744**, 295–301 (1996).
55. A.B. Chen, C.A. Rickel, A. Flanigan, G. Hunt, and F.G. Moorhouse, *J. Chromatogr. A* **744**, 279–284 (1996).
56. L.E. Bennett, W.N. Charman, D.B. Williams, and S.A. Charman, *J. Pharm. Biomed. Anal.* **12**, 1103–1108 (1994).
57. D.J. Kroon, S. Goltra and B. Sharma, *J. Cap. Elec.* **2**, 34–39 (1995).
58. D.B. Wetlaufer, *Adv. Protein Chem*. **17**, 303–390 (1962).
59. R.H. Strick, *CRC Crit. Rev. Biochem*. **2**, 113–175 (1974).
60. T.A. Bewley, *Rec. Prog. Hormone Res.* **35**, 155–213 (1979).
61. A.F. Winder and W.L.G. Gent, *Biopolymers* **10**, 1243–1251 (1971).
62. V.A. Bloomfield, *Ann. Rev. Biophys. Bioeng.* **10**, 421–450 (1981).
63. G.D.J. Phillies, *Anal. Chem*. **62**, 1049A–1057A (1990).
64. *Physicians' Desk Reference (PDR)*, 52nd ed., Medical Economics Company, Montvale, N.J., 1998.
65. L.C. Gu, E.A. Erdos, H.S. Chiang, T. Calderwood, K. Tsai, G.C. Visor, J. Duffy, W.-C. Hsu, and L.C. Foster, *Pharm. Res.* **8**, 485–490 (1991).
66. A.W. Adamson, *Physical Chemistry of Surfaces*, 2nd ed., Interscience Publishers, New York, 1960.
67. A.F. Henson, J.R. Mitchell, and P.R. Musselwhite, *J. Colloid Interface Sci.* **32**, 162–165 (1970).
68. P.H. Von Hippel and T. Shleich, in S.N. Timasheff and G.D. Fasman eds., *Structure and Stability of Macromolecules*, vol. 2, Dekker, New York, 1969, pp. 417–574.
69. X.M. Lam, H.R. Costantino, D.E. Overcashier, T.H. Nguyen, and C.C. Hsu, *Int. J. Pharm*. **142**, 85–95 (1996).
70. S.N. Timasheff and T. Arakawa, in T.E. Creighton ed., *Protein Structure, A Practical Approach*, IRL Press; New York, 1988, pp. 331–345.
71. D.C. Cipolla and S.J. Shire, in J.J. Villafranca ed., *Techniques in Protein Chemistry II*, Academic, New York, 1990, pp. 543–555.
72. H. Mach, C.J. Burke, G. Sanyal, P.-K. Tsai, D.B. Volkin, and C.R. Middaugh, in J.L. Cleland and R. Langer eds., *Formulation and Delivery of Proteins and Peptides*, American Chemical Society Symposium Series, ACS Books, New York, 1994, pp. 72–84.
73. S. Nema, R.J. Washkuhn, and D.R. Beussink, *Pharm. Technol.* **19**, 170–185 (1995).
74. K.A. Connors, G.L. Amidon, and V.J. Stella, in *Chemical Stability of Pharmaceuticals: A Handbook for Pharmacists*, 2nd ed., John Wiley, New York, 1986, pp. 82–104.
75. M.J. Pikal, *BioPharmacy* **3**, 18–27 (1990).
76. M.J. Pikal, *BioPharmacy* **3**, 26–30 (1990).
77. J.F. Carpenter, S.J. Prestrelski, T.J. Anchordoguy, and T. Arakawa, in J.L. Cleland and R. Langer eds., *Formulation and Delivery of Proteins and Peptides*, American Chemical Society Symposium Series, ACS Books, New York, 1994, pp. 132–146.
78. S.J. Prestrelski, T. Arakawa, and J.F. Carpenter, in J.L. Cleland and R. Langer, eds., *Formulation and Delivery of Proteins and Peptides*, American Chemical Society Symposium Series, ACS Books, New York, 1994, pp. 148–169.
79. M.J. Pikal, in J.L. Cleland and R. Langer eds., *Formulation and Delivery of Proteins and Peptides*, American Chemical Society Symposium Series, ACS Books, New York, 1994, pp. 120–132.
80. L.A. Gatlin and S.L. Nail, *Bioprocess Technol.* **18**, 317–367 (1994).
81. R.C. Hubbaard, N.G. McElvaney, P. Birrer, S. Shak, W.W. Robinson, and C. Jolley, *New Engl. J. Med.* **326**, 812–815 (1992).
82. J.L. Cleland and R. Langer, in J.L. Cleland and R. Langer eds., *Formulation and Delivery of Proteins and Peptides*, American Chemical Society Symposium Series, ACS Books, New York, 1994, pp. 1–19.
83. P.D. Gluckman and W.S. Cutfield, *Arch. Dis. Childhood*, **66**, 686–688 (1991).
84. C. Pavia, C. Valls, M. Gallart, R. Velasco, and M.C. Gonzalez, *Acta Ped. Espanola* **53**, 25–28 (1995).
85. J.R. Brennan, S.S.P. Gebhart, and W.G. Blackard, *Diabetes* **34**, 353–359 (1985).
86. R. Sharifi, R.C. Bruskewitz, M.C. Gittleman, S.D. Graham, Jr., P.B. Hudson, and B. Stein, *Clin. Ther.* **18**, 647–657 (1996).
87. O.A. Oyesany, S.K. Teo, E. Quah, N. Abdurazak, F.Y. Lee, and W.C. Cheng, *Hum. Reprod.* **10**, 1042–1044 (1995).
88. J.L. Cleland, O.L. Johnson, S.D. Putney, and A.J.S. Jones, *Adv. Drug Delivery Rev.* **28**, 71–84 (1997).
89. J.L. Cleland, in *Protein Delivery: Physical Systems*, L. Sanders and W. Hendren eds., Plenum, New York, 1997, pp. 1–43.
90. H. Okada and H. Toguchi, *Crit. Rev. Ther. Drug Carrier Systems* **12**, 1–99 (1995).
91. R.W. Niven, *Pharm. Technol.* **17**, 72–74 (1993).
92. L.L. Wearley, *Crit. Rev. Ther. Drug Carrier Systems* **8**, 331–394 (1991).
93. J.T. Jorgensen, *J. Ped. Endocrinol.* **7**, 175–180 (1994).
94. J.O. Hornquist, A. Wikby, U. Stenstrom, P.O. Anderson, *Diabetes*, **28**, 63–72 (1995).
95. J.L. Cleland, A. Mac, B. Boyd, J. Yang, E.T. Duenas, D. Yeung, D. Brooks, C. Hsu, H. Chu, V. Mukku, and A.J.S. Jones, *Pharm. Res.* **14**, 420–425 (1997).

96. J.L. Cleland, E.T. Duenas, and A.L. Daugherty, M. Marian, J. Yang, M. Wilson, A.C. Celniker, A. Shhzamani, V. Quarmby, H. Chu, V. Mukku, A. Mac, M. Roussakis, N. Gillette, B. Boyd, D. Yeung, D. Brooks, Y.-F. Maa, C. Hsu, A.J.S. Jones, *J. Controll. Rel.* **49**, 193–205 (1997).

97. J.L. Cleland and A.J.S. Jones, *Pharm. Res.* **13**, 1462 (1996).

98. J.Q. Oeswein, A.L. Daugherty, E.J. Cornavaca, J.A. Moore, H.M. Gustafson, B.M. Eckhardt, in R.N. Dalby and R. Evans eds., *Proceedings of the Second Respiratory Drug Delivery Symposium*, Continuing Pharmacy Education, University of Kentucky, Lexington, Ky., 1991, pp. 14–49.

99. V. Naini, P.R. Byron, and R.N. Dalby, in R.N. Dalby, P.R. Byron, and S.J. Farr eds., *Respiratory Drug Delivery V*, Interpharm Press, Buffalo Grove, Ill., 1996, pp. 382–384.

100. M. Mummenthaler, C.C. Hsu, R. Pearlman, *Pharm. Res.* **11**, 12–17 (1994).

FREEZE-DRYING, PHARMACEUTICALS

R. BRUTTINI
Criofarma Freeze-drying Equipment
Turin, Italy

KEY WORDS

Amorphous solutes
Bound water
Crystalline solutes
Free water
Freeze-drying
Lyophilization
Primary drying
Secondary drying

OUTLINE

INTRODUCTION

The primary object of freeze-drying is to preserve biological materials and pharmaceuticals without injury by freezing the contained water and then removing the ice by sublimation. Thus one can expect to combine the advantages of both freezing and low-temperature dessication to obtain a more propitious state of preservation. In freeze-drying, the water or other solvent is removed as a vapor by sublimation from the frozen material in a vacuum chamber, leaving a dry porous mass of approximately the same size and shape as the original frozen mass. The resulting product is in a stable form and can be redissolved rapidly in water; the material displays what is called lyophilic behavior, which gave rise to the term *lyophilization* to describe the process of freeze-drying.

Lyophilization is employed soon after production for injectable pharmaceuticals that exhibit poor stability in solution, preventing their deactivation over a period of time and preserving their bioactivity. For example, a compound that is heat sensitive and undergoes rapid decomposition in aqueous solution can be formulated in a stable, rapidly soluble form.

THE FREEZE-DRYING PROCESS

Many pharmaceuticals and biological materials, which may not be heated even to moderate temperatures in ordinary drying, can be freeze dried (1–6). Furthermore, many pharmaceutical products in solution deactivate over a period of time; the bioactivity of such compounds can be preserved by freeze-dried soon after their production, which stabilizes their molecules (2,6). Freeze-drying can also be applied to systems in which interfacial forces during drying need to be eliminated and where homogeneous powder products (derived from multicomponent, multiphase precursors) are required (2,5,7).

As a rule, freeze-drying produces the highest-quality pharmaceutical product obtainable by any drying method. One important factor is the surface structural rigidity of the frozen substance, where sublimation occurs. This rigidity, to a large extent, prevents collapse of the solid matrix remaining after drying. The result is a porous, nonshrunken structure that facilitates rapid and almost complete rehydration when water is added at a later time. Freeze-drying of biological materials and pharmaceuticals also has the advantage of low processing temperatures, the relative absence of liquid water, and the rapid transition of any local region of the material being dried from a fully hydrated to a nearly completely dehydrated state, all of which minimize the degradative reactions that normally occur in ordinary drying processes, such as protein denaturation and enzymatic reactions. In any pharmaceutical or biological material, some nonfrozen water (bound, or sorbed, water) will almost unavoidably be present during freeze-drying, but there is very often a rather sharp transition temperature for the still-wet region during drying (5), below which the product quality improves markedly. This improvement shows that sufficient water is frozen to provide the beneficial characteristics of freeze-drying.

However, freeze-drying is an expensive form of dehydration for pharmaceuticals because of the slow drying rate and the use of vacuum. The cost of processing is offset to some extent by the absence of any need for refrigerated handling and storage.

Systematic freeze-drying is a procedure mainly applied to the following categories of material (1–18):

1. Nonliving matter, such as blood plasma, serum, hormone solutions, foodstuffs, pharmaceuticals (e.g., antibiotics), ceramics, superconducting materials, and materials of historical importance (e.g., archaeological wood)
2. Surgical transplants, which are made nonviable so that the host cells can grow on them and use them as a skeleton (e.g., artery, bone, and skin)
3. Certain living cells that must remain viable for long periods of time, such as bacteria, yeasts, and viruses

Freeze-drying requires very low pressures or high vacuum in order to produce a satisfactory drying rate. If the water were in a pure state, freeze drying at or near 0 °C at an absolute pressure of 4.58 mmHg could be performed. But because the water usually exists in a combined state or a solution, the material must be cooled below 0 °C to keep the water in the solid phase. Most pharmaceutical freeze-drying is done at −20 °C or lower at absolute pressures of about 0.3 mmHg or less.

In short, freeze drying is a multistep operation:

1. The material is frozen hard by low-temperature cooling.
2. It is dried by direct sublimation of the frozen solvent and by desorption of the sorbed or bound solvent (nonfrozen solvent), generally under reduced pressure.
3. It is then stored in the dry state under controlled conditions (free of oxygen and water vapor and usually in airtight, opaque containers, filled with inert dry gas).

If correctly processed, most products can be kept in such a state for an almost unlimited period of time while retaining all their initial physical, chemical, biological, and organoleptic properties, and remaining available at any time for immediate reconstitution. In most cases reconstitution is accomplished by addition of the exact amount of solvent that has been extracted, thus giving to the reconstituted product a structure and appearance as close as possible to the original material. However, in some instances, reconstitution can be monitored in order to yield more concentrated or diluted products by controlling the amount of solvent. Vaccines and pharmaceutical materials are very often reconstituted in physiological solutions that are quite different from the original but are best suited for intramuscular or intravenous injection.

These examples are not exhaustive; detailed presentations on the uses of the freeze-drying process and of freeze-dried products are given in Refs. 1–10, 12, and 17–22.

FREEZE-DRYING STAGES

The freeze-drying separation method (process) involves three stages: *(1)* the freezing stage, *(2)* the primary drying stage, and *(3)* the secondary drying stage.

In the freezing stage, the product or solution to be processed is cooled down to a temperature at which all the material is in a frozen state.

In the primary drying stage, the frozen solvent is removed by sublimation; this requires that the pressure of the system (freeze-dryer) where the product is being dried must be less than or near to the equilibrium vapor pressure of the frozen solvent. If, for instance, frozen pure water (ice) is processed, then sublimation of pure water at or near 0 °C and at an absolute pressure of 4.58 mmHg could occur. But because the water usually exists in a combined state (e.g., gel suspension or biological material) or a solution (e.g., pharmaceutical product), the material must be cooled below 0 °C to keep the water in the frozen state. For this reason, during the primary drying stage the temperature of the frozen layer (Fig. 1) is most often at −20 °C or lower and at absolute pressures of about 0.3 mmHg or less. As the solvent (ice) sublimes, the sublimation interface (plane of sublimation), which started at the outside surface (Fig. 1), recedes, and a porous shell of dried material remains. The heat for the latent heat of sublimation (2,840 kJ/kg ice) can be conducted through the layer of dried material and through the frozen layer, as shown in Fig. 1. The vaporized solvent (water vapor) is transported through the porous layer of dried material. During the primary drying stage, some of the sorbed water (nonfrozen water) in the dried layer may be desorbed. The desorption process in the dried layer could affect the amount of heat that arrives at the sublimation interface, and therefore, it could affect the velocity of the moving sublimation front (interface). The time at which there is no more frozen layer (that is, there is no more sublimation interface) is taken to represent the end of the primary drying stage.

The secondary drying stage involves the removal of solvent (water) that did not freeze (sorbed, or bound, water). The secondary drying stage starts at the end of the primary drying stage, and the desorbed water vapor is transported through the pores of the material being dried.

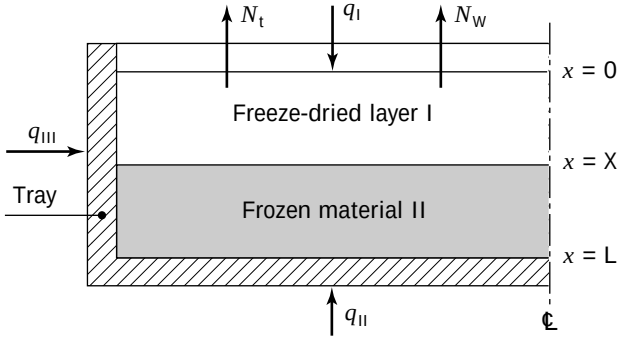


Figure 1. Diagram of a material on a tray during freeze-drying; the variable X denotes the position of the sublimation interface (front) between the freeze-dried layer (layer I) and the frozen material (layer II).

The Freezing Stage

The freezing stage represents the first separation step in the freeze-drying process, and the performance of the overall freeze-drying process depends significantly on this stage. The objective in freezing is to convert most of the water from the system by formation of ice and convert all solutes into solids, either crystalline solids or a glass. The material system to be processed (e.g., gel suspension, pharmaceutical solution, or biological material) is cooled down to a temperature (which depends on the nature of the product) that is always below the solidification temperature of the material system. For instance, if the material to be freeze-dried is a solution whose equilibrium phase diagram presents a eutectic point (e.g., the solution of NaCl and water presents a eutectic point at $-21.6\ ^{\circ}C$), then the value of the final freezing temperature must be below the value of the eutectic temperature; in this case, the material becomes wholly crystalline. In practice, materials display one of two different types of freezing behavior: *(1)* the liquid phase suddenly solidifies (eutectic formation) at a temperature that depends on the nature of the solids in the sample, or *(2)* the liquid phase does not solidify (glass formation), but rather it just becomes more and more viscous until it finally takes the form of a very stiff, highly viscous liquid. In this latter case there is no eutectic temperature, but there is a minimum freezing temperature.

At the end of the freezing step, there already exists a separation between the water to be removed (frozen water in the form of ice crystals) and the solute. In many cases, at the end of the freezing stage about 65–90% of the initial (at the start of the freezing stage) water is in the frozen state, and the remaining 10–35% of the initial water is in the sorbed (nonfrozen) state. The shape of the pores, the pore-size distribution, and pore connectivity (1–3,6,9–11,13,15,16,21,23,24) of the porous network of the dried layer formed by the sublimation of the frozen water during the primary drying stage depend on the ice crystals that formed during the freezing stage; this dependence is of extreme importance because the parameters that characterize the mass and heat transfer rates in the dried layer are influenced significantly by the porous structure of the dried layer. If the ice crystals are small and discontinuous, then the mass transfer rate of the water vapor in the dried layer could be limited. On the other hand, if large dendritic ice crystals are formed and homogeneous dispersion of the preeutectic and posteutectic frozen solution can be realized, the mass transfer rate of the water vapor in the dried layer could be high, allowing the product to be dried more quickly. Thus, the method and rate of freezing, as well as the shape that contains the solution and the nature of the product, are critical to the course of lyophilization because they affect the drying rate and the quality of the product.

In industrial pharmaceutical freeze-dryers the freezing of the product is usually done in the same plant where the drying occurs. In certain nonsterile pharmaceutical freeze-dryers the freezing of the product is accomplished by spraying liquid nitrogen into the drying chamber containing the product. In sterile pharmaceutical freeze-dryers the freezing stage is realized by contact between cooled plates and product-supporting containers.

The Primary Drying Stage

After the freezing stage, the drying chamber where the product is placed is evacuated and the chamber pressure is reduced down to a value that allows the sublimation of solvent (water) to take place. When the water molecules sublime and enter the vapor phase, they also keep with them a significant amount of the latent heat of sublimation (2,840 kJ/kg ice), and thus, the temperature of the frozen product is again reduced. If there is no heat supplied to the product by a heat source, then the vapor pressure of the water at the temperature of the product reaches the same value as that of the partial pressure of the water vapor in the drying chamber; therefore, the system reaches equilibrium and no additional water sublimation from the product occurs. In order to have continuous sublimation of water from the product, the latent heat of sublimation must be provided by a heat source. The heat is usually supplied by conduction, convection, and/or radiation; conduction is realized by contact between heated plates and product-supporting containers.

The amount of heat that can be supplied to the product can not be increased freely because there are certain limiting conditions that have to be satisfied during the primary drying stage. One of the constraints has to do with the maximum temperature that the dried product (freeze-dried layer in Fig. 1) can tolerate without *(1)* loss of bioactivity, *(2)* color change, *(3)* the possibility of degradative chemical and biochemical reactions, and *(4)* structural deformation in the dried layer. The maximum temperature that the dried product can tolerate without suffering any of these deleterious effects is denoted, for a given product, by T_{scor} (T_{scor} is often called, by convention, the temperature of the scorch point of the dried product). Another constraint has to do with the maximum temperature the frozen layer can tolerate and still remain frozen. If the material has a eutectic form and if the temperature of the lowest eutectic point is exceeded during the primary drying stage, then melting in the frozen layer (Fig. 1) can occur. Melting at the sublimation interface or any melting that occurs in the frozen layer can give rise to gross material faults such as puffing, shrinking, and structural topologies filled with liquid solution. When melting has occurred at some point in the frozen layer, then the solvent at that point cannot be removed by sublimation, and therefore, there is process failure (Fig. 1). There has also been, at least, a loss in structural stability. If the material is in a glass form and if the minimum freezing temperature is exceeded during the primary drying stage, then the phenomenon of collapse can occur, with a loss of rigidity in the solid matrix. Again in this case, there is a process failure because the water can no longer be removed from the frozen layer only by sublimation, and there has also been, at least, a loss in structural stability. The structural stability of a material relates to its ability to go through the freeze-drying process without change in size, porous structure, and shape. The maximum allowable temperature in the frozen layer is determined by both structural stability and product stability (e.g., product bioactivity) factors; that is, the maximum value of the temperature in the frozen layer during the primary drying stage must be such that the

drying process is conducted without loss of product properties (e.g., bioactivity) and structural stability. Sometimes the product stability factors are related to structural stability factors (as in melting). There are systems in which the product stability factors do not depend on structural stability factors, as is the case for many vaccines, viruses, and bacteria where the temperature of the frozen layer during the primary drying stage must be kept well below the melting temperature so that there is a good level of bioactivity and organism survival after drying. In general, product stability is related to the temperature of the frozen layer during the primary drying stage. The maximum allowable temperature that the frozen layer can tolerate without suffering melting, puffing, shrinking, collapse, and loss of product property or stability, is denoted, for a given product, by T_m (T_m is often called, by convention, the melting temperature of the sublimation interface or of the frozen layer).

The water vapor produced by the sublimation of the frozen water in the frozen layer and by the desorption of sorbed (nonfrozen) water in the dried layer during the primary drying stage travels by diffusion and convective flow through the porous structure of the dried layer and enters the drying chamber. (It should be noted that most of the water removed during the primary drying stage is produced by sublimation of the frozen water in the frozen layer.) The water vapor must be continuously removed from the drying chamber in order to maintain nonequilibrium conditions for the drying process in the system. This is usually accomplished by fitting a refrigerated trap (called an ice condenser) between the drying chamber and the vacuum pump; the water vapor is collected on the cooled surface of the condenser in the form of ice. The time at which there is no more frozen layer is taken to represent the end of the primary drying stage.

Secondary Drying Stage

The secondary drying stage involves the removal of water that did not freeze (sorbed or bound water). In an ideal freeze-drying process, the secondary drying stage starts at the end of the primary drying stage. The word *ideal* is used here to suggest that in an ideal freeze-drying process only frozen water should be removed during the primary drying stage, whereas the sorbed water should be removed during the secondary drying stage. But, as just discussed, in real freeze-drying systems a small amount of sorbed water could be removed by desorption from the dried layer of the product during the primary drying stage, and thus, there could be some secondary drying in the dried layer of the product during the primary drying stage. In real freeze-drying processes the secondary drying stage is considered to start when all the ice has been removed by sublimation (end of primary drying stage). It is then considered that, during the secondary drying stage, most of the water that did not freeze (bound water) is removed. The bound moisture is present due to mechanisms of *(1)* physical adsorption, *(2)* chemical adsorption, and *(3)* water of crystallization. Although the amount of bound water is about 10–35% of the total moisture content (65–90% of the total moisture could be free water that was frozen and then removed by sublimation during the primary drying stage), its effect on the drying rate and overall drying time is very significant, and the time that it takes to remove the sorbed water could be as long or longer than the time that is required for the removal of the free water.

The bound water is removed by heating the product under vacuum. But, as in the case of primary drying, the amount of heat that can be supplied to the product can not be increased freely because there are certain constraints that have to be satisfied during the secondary drying stage. The constraints have to do with the moisture content and the temperature of the product; these two variables influence the structural stability, as well as the product stability, during and after drying. For structural stability, the same phenomena, as in the case of the primary drying stage, have to be considered: collapse, melting (if the temperature is increased at constant moisture), or dissolution (if moisture is increased at constant temperature) of the solid matrix can occur. Product stability (e.g., bioactivity) is a function of both moisture content and temperature in the sample, and during secondary drying the moisture concentration and temperature in the sample could vary widely with location and time. This implies that the potential for product alteration to occur in the sample will vary with time and location. The moisture concentration profile is related to the temperature profile in the dried layer, and thus, the moisture content in the sample cannot be controlled independently. Because many products are temperature sensitive, it is usual to control product stability by limiting the temperature during the secondary drying process, and the final moisture content is checked before the end of the cycle (6,10,25).

In the secondary drying stage, the bound water is removed by heating the product under vacuum, and the heat is usually supplied to the product by conduction, convection, and/or radiation. The following product temperatures are usually employed: *(1)* below 0 °C for vaccines, *(2)* between 10 °C and 35 °C for heat-sensitive products, and *(3)* 50 °C or more for less heat sensitive products.

The residual moisture content in the dried material at the end of the secondary drying stage, as well as the temperature at which the dried material is kept in storage, are critical factors in determining product stability during storage life. Some vaccines remain stable for many years when they are stored at −20 °C, but a significant loss of titer can be found after one year if they are stored at 37 °C (26). Furthermore, certain vaccines such as live rubella and measles can be damaged by overdrying (final moisture content of about 2% is required for best titer retention), whereas other materials such as chemotherapeutics and antibiotics must be dried to a residual moisture content as low as 0.1% for best results.

FREEZE-DRYING EQUIPMENT

Freeze-drying equipment for pharmaceuticals is able to perform the three stages of the freeze-drying process: *(1)* freezing of the pharmaceutical product (usually a pharmaceutical aqueous solution), *(2)* primary drying of the product by removing the frozen solvent by sublimation,

and *(3)* secondary drying of the product by removing the sorbed, or bound, water (solvent that did not freeze).

All three stages of the freeze-drying process are energy consuming. The freezing stage involves freezing of the wet product. Because this is normally considered one of the preparatory steps before the actual freeze-drying, we concentrate here on the other two steps, which take place in the freeze-drying chamber (2,5,12). The primary and secondary drying stages involve the controlled supply of heat to the product to cover requirements for the sublimation and desorption processes, respectively. The removal from the freeze-drying chamber of the vast volumes of water vapor released during the sublimation and desorption processes always consumes the largest amount of energy. The efficiency of water vapor removal (the vapor trap system) therefore has a decisive effect on the total energy consumption of the freeze-drying plant.

The vapor trap is placed in an ice condenser chamber communicating with the freeze-drying chamber. The water vapor condenses to ice on its refrigerated surfaces.

When in operation, the efficiency of the vapor trap is shown by a small total temperature difference, ΔT, between the saturation temperature for water vapor at the pressure in the freeze-drying cabinet and the evaporation temperature of the refrigerant (Fig. 2). This total temperature difference, ΔT, results mainly from the following three resistances:

1. Pressure difference, ΔP; equivalent to the pressure drop caused by the resistances to the vapor flow from the freeze-drying chamber to the cold surfaces of the vapor trap placed in the ice condenser chamber.
2. The temperature difference ΔT_{ice} through the layer of ice on the cold surface.
3. The temperature difference ΔT_{refr} between the cold surface and the evaporating refrigerant.

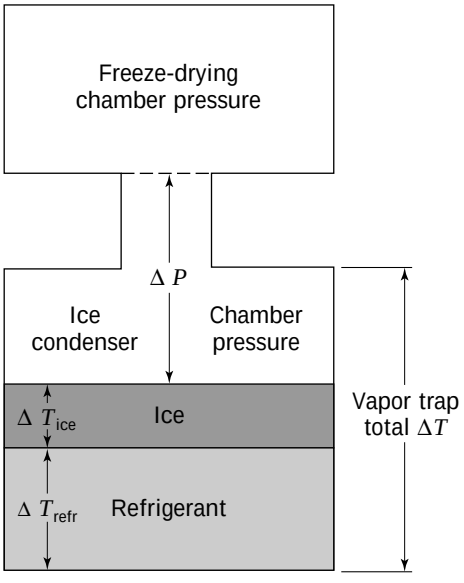


Figure 2. Graphic presentation of the variables ΔP, ΔT_{ice}, ΔT_{refr}, and ΔT.

For an efficient vapor trap it is necessary to have a combination of a large cross-sectional area for the vapor flow (low ΔP), an efficient deicing system (low ΔT_{ice}), and an efficient refrigerating system (low ΔT_{refr}). A less efficient vapor trap means a higher ΔT, thus demanding a lower evaporation temperature of the refrigerating plant to maintain the required vacuum in the freeze-drying cabinet. A lower evaporation temperature means higher operation costs. In this temperature range, an evaporation temperature 10 °C lower means a 50% increased in energy consumption.

The principal problem in the freeze-drying of pharmaceutical solutions is to operate in sterile conditions. In modern plants, the internal sterilization of the equipment is usually made with pressurized steam at 121 °C or more; in old plants, sterilization is realized with the use of certain proprietary sanitizing agents (27,28). A pharmaceutical freeze-dryer must be designed according to current good manufacturing practice (cGMP) in order to be validated before and during use.

When evaluating industrial freeze-drying plants, the following characteristics are thus of prime importance:

1. Operation reliability
2. Ease, safety, and quality of process control
3. Product losses
4. Vapor trap efficiency
5. Internal sanitization
6. cGMP requirements

In the following sections, some of the commonly used types of pilot and industrial freeze-drying equipment for pharmaceuticals are presented, and their most important technical features are discussed.

Pilot Freeze-dryers

The freeze-drying of pharmaceutical products can be performed in the freeze-dryer as: *(1)* bulk solution freeze-drying in trays, or *(2)* freeze-drying in vials.

Freeze-drying pilot units appropriate for use in the pharmaceutical industries and in the laboratory are in high demand because they are used to explore possibilities for the preservation of labile products, especially those of biological origin. These units are portable and of convenient size for developmental work on freeze-dried products in laboratories and factories around the world. A large number of designs incorporate self-contained facilities for refrigeration, heating, and vacuum pumping, and they can freeze-dry batches consisting of 1 to 10 kg of frozen product. Because of the large variety of pilot freeze-dryers that are employed in industries and laboratories, and because of the limitation space we have for description, a pilot freeze-dryer is presented here whose characteristics are very close to the characteristics of industrial large-scale lyophilizers.

A schematic diagram of the pilot unit (Criofarma model C8-2), is shown in Figure 3a. The unit consists of *(1)* a freezing fluid system (R404A) that can be routed to the heat exchanger section of the condenser or into the refrig-

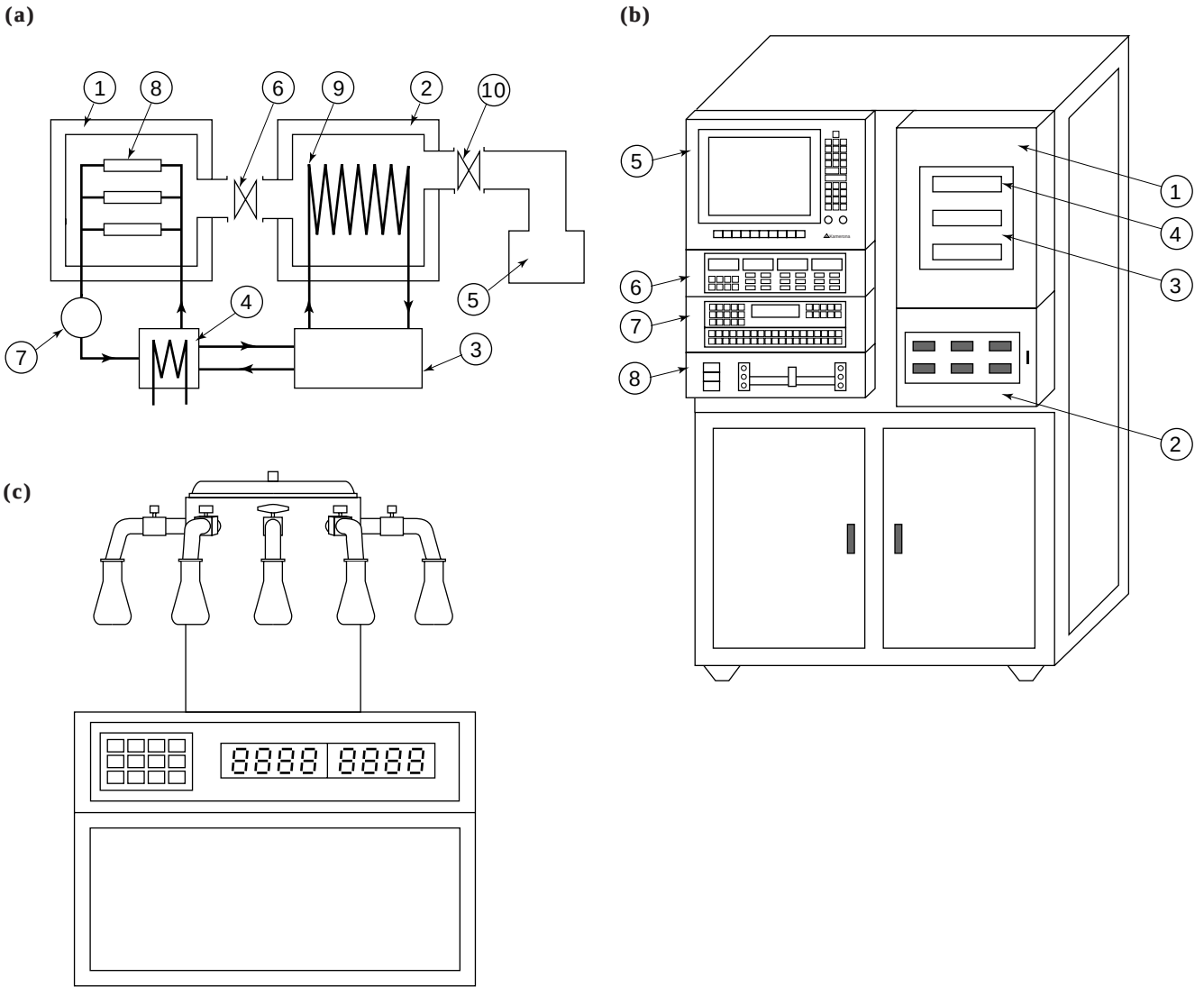


Figure 3. Pilot freeze-dryer (Criofarma model C8-2). (**a**) Diagram of a pilot freeze-dryer: 1, drying chamber; 2, ice condenser chamber; 3, refrigeration unit; 4, cooling/heating system for the plates; 5, vacuum unit; 6, isolation butterfly valve; 7, silicon oil pump; 8, cooling/heating plate; 9, refrigerated coil; 10, condenser vacuum valve. (**b**) Frontal view of a pilot freeze-dryer: 1, drying chamber; 2, ice condenser chamber; 3, cooling/heating plate; 4, inspection window; 5, computer system; 6, vacuum indicator and regulator; 7, temperature control panel; 8, printer. (**c**) Frontal view of a laboratory freeze-dryer (Criofarma model Criolab-8). *Source*: Courtesy of Criofarma.

eration coils for product freezing, *(2)* a heating circuit (silicon oil is the heating fluid) for plate heating and defrosting of the condenser, and *(3)* a vacuum system for evacuating air from the apparatus before and during drying.

The rectangular drying chamber shown in Figure 3b is mounted on top of the condenser; the dimensions are 0.45 m $\times$ 0.45 m, and 0.5 m deep. Viewing windows are incorporated in the drying and condensation sections. The refrigeration and vacuum systems are in the internal part of the apparatus, whose complete dimensions are 1.1 m $\times$ 0.8 m, and 2.0 m high.

For pilot use, the freeze-dryer offers full control of the process variables and is able to achieve conditions of pressure and temperature beyond the limits of production units. The shelf and ice condenser temperatures of the pilot unit can be -55 °C and -75 °C, respectively, and the pressure in the drying chamber can be as low as 1 Pa or less. The pilot freeze-dryer has a control panel fully accessoried with instruments that record and display *(1)* the temperature inside the product, *(2)* the temperature on the plates, *(3)* the temperature of the coils of the condenser, *(4)* the pressure in the drying chamber, *(5)* the pressure in the vacuum unit, and *(6)* the pressure in the condenser section. The use of a personal computer in this pilot unit, with programmable temperature during the freeze-drying cycle and programmable input–output logic in the different freeze-drying stages, offers a wide variety of drying cycles and the capability for extensive data acquisition, so

process optimization can be examined and studied without the risk and cost of investigating a full-size production freeze-dryer. Furthermore, the use of a computer is important to make use of manometric temperature measurement methodology (29) in the primary drying, which is based upon an analysis of the rate of pressure increase in the drying chamber when the valve separating the drying chamber from the ice condenser chamber is periodically quickly closed for a brief period of time. A mathematical analysis of the pressure rise data for a given apparatus gives the average temperature of the product without placing temperature sensors in the product and, therefore, without risk of sterility compromise.

Figure 3c shows a small pilot freeze-dryer, suitable for freeze-drying a few vials (Criofarma model Criolab-8).

Industrial Freeze-Dryers

The objective in a freeze-drying process is to convert most of the water into ice in the freezing stage, remove the ice by direct sublimation in the primary drying stage, and finally remove most of the unfrozen water in the secondary drying stage by desorption. In the industrial freeze-dryer it is possible to perform all three stages of the freeze-drying process. Most of the industrial freeze-dryers in operation are of the vacuum-batch type with freeze-drying of the product in trays. There are two main subtypes, depending on the type of condenser used. In one, the condenser plates are alongside the tray-heater assembly and in the same chamber; in the other the ice condenser is in a separate chamber joined to the first, generally, by a wide butterfly valve. This latter type of plant is always used in pharmaceutical industries. Because of the wide variety and complexity of the problems associated with the production of pharmaceuticals by freeze-drying, in the following paragraphs the principal features of an industrial tray freeze-dryer for pharmaceuticals are presented.

The principal problem in the freeze-drying of pharmaceutical solutions is to operate in sterile conditions. The pharmaceutical solution may be sterile filtered immediately before introduction into a vial or a tray. Also the location of the plant must be able to guarantee sterile conditions during the filling, charging before drying, and discharging after drying of the pharmaceutical product. This is realized by facing the drying chamber door in a wall that separates the sterile room from the machine, or nonsterile, room. In the plant, this separation is accomplished with an isolation valve that separates the ice condenser from the drying chamber; this valve is also able to permit *(1)* the pressure rise test at the end of the freezing-drying cycle, *(2)* the simultaneous discharging/loading of the product and condenser defrosting, and *(3)* the reduction of cross-contamination between batches to a minimum. All the internal parts of the freeze-dryer are of stainless steel type AISI 316L with a finished surface of 0, 1–0, 4 rugosity or less. In modern plants, internal sterilization of the equipment is usually done with pressurized steam at 121 °C or more; in old plants, sterilization is realized with the use of hydrogen peroxide (27) or certain proprietary sanitizing agents (28).

The product containers (vials or bottles loaded on stainless steel trays) are usually sterilized in a separate unit before filling and charging in the freeze-dryer. These operations require the presence of people in the sterile room, with consequent handling of the containers and possible contamination of the batch. For this reason, the human presence in the sterile room is usually reduced the minimum necessary. A new freeze-dryer plant concept has been developed in order to reduce the risk of product contamination. The plant, as shown in Figure 4 (Criofarma model C292-8S), has two full doors. In the first (usually closed) there is a small door for loading the product before drying; the second full door, located opposite to the small door, is for discharging the product after drying. The condenser could be placed on the same floor, or on a floor that is below the floor where the drying chamber is located. The shelves of the freeze-dryer are lowered to the bottom of the drying chamber and are then lifted one by one to a position in line with the loading machine. The charging of the product is made under a laminar flow of sterile air; the small door is opened only for each plate loading and is then immediately closed. If the product is unstable and must be frozen within a short time after it is filled into its container, then it is possible to load trays of product onto the precooled shelves a half plate at a time. When the product container is a bottle, as shown in Figure 4, it usually has a silicon plug, which is partially inserted into the bottle; the solvent vapor leaves the container from the free space between the inserted portion of the plug and the container. After drying and before product discharge, the bottles are stoppered in the drying chamber with the plugs, which are now fully introduced into the bottles. The stoppering operation is done in vacuum conditions, or at atmospheric pressure by breaking the drying chamber vacuum with sterile nitrogen, which prevents successive oxidation of the product; the latter method is most often employed in practice. The silicon plug in the stoppered bottles provides a protection from contamination, and it may be possible to discharge the product in a less-sterile environment from the full door of the freeze-dryer in only one operation. The entire process may be fully automated as the bottles are removed from the filling machine; the disadvantage of automation is that the loading time of the freeze-dryer may become as large as the time it usually takes to complete the filling step of the operation, and this could reduce the theoretical freeze-dryer production for a large installation.

Pharmaceutical freeze-dryers are very often used to produce raw materials such as ampicillin, cloxacillin, and cefazolin (usually as sodium salt), or other specialty materials such as collagen. In these systems, the product is usually charged on stainless steel or polyethylene film trays, and the plant is usually a medium or large unit with a loading surface varying from 20 m^2 to 60 m^2. If the product to be freeze-dried is not particularly unstable (e.g., collagen) and can withstand a delay of some hours between being filled into its tray and being frozen, then one can usually accumulate the trays of product on a loading trolley. When the loading trolley is filled, it is placed in front of the freeze-dryer, and the trays are automatically pushed on the shelves without sliding contact (in order to avoid particle generation), in a single operation. This system is advantageous because it permits maximum utilization of the freeze-dryer; the trolley may be loaded ahead of time,

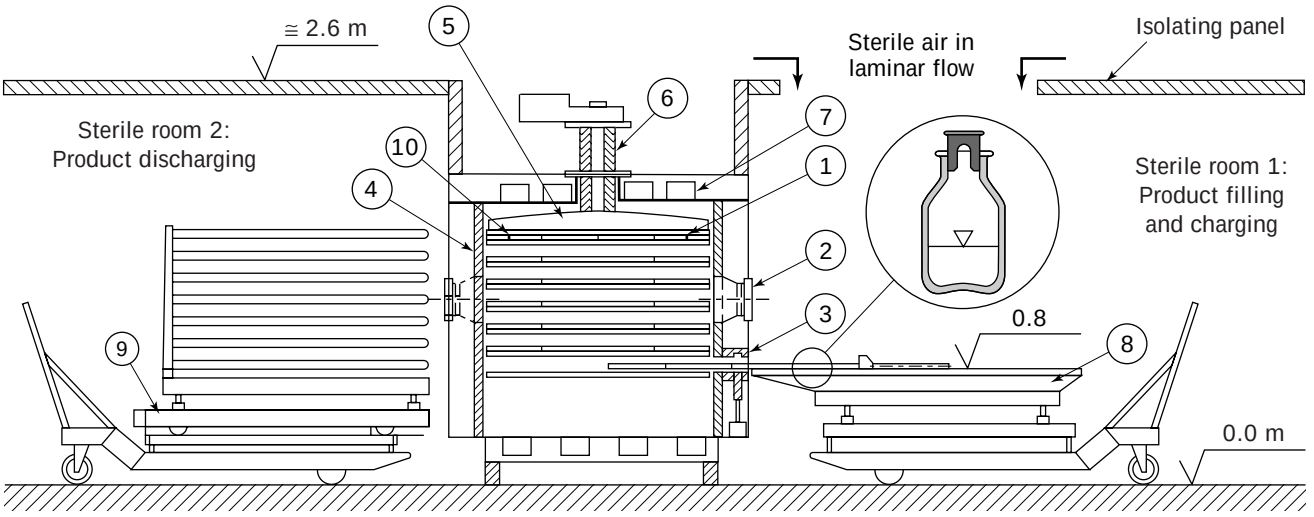


Figure 4. Layout of an industrial freeze-dryer with stoppering device (Criofarma Model C292-8S): 1, drying chamber; 2, inspection window; 3, automatic small door opening; 4, full door; 5, hydraulic press for stoppering the bottles after drying; 6, stainless steel elbows for double sterile condition inside the stoppering plug; 7, reenforcing member and cooling coils after steam sterilization; 8, loading device; 9, discharging device; 10, loaded shelves. *Source*: Courtesy of Criofarma.

so that loading into the freeze-dryer can be carried out in a few minutes. If the product is not stable in the liquid state (e.g., ampicillin sodium salt) and must be frozen within a short time after its preparation, it is common to charge the empty trays on the precooled shelves and then to fill the trays, so that the freezing step is very quick and can proceed during the whole loading operation. This approach is also advantageous because it reduces the freeze-drying cycle time; this happens because the cooling phase starts at the same time as the loading phase, with a consequent reduction in the total time of these two steps.

The freeze-drying system used in the pharmaceutical industry basically consist of the following six units:

1. *Drying chamber*. Holds the product to be dried on shelves an provides an evacuated space for freeze-drying.
2. *Ice condenser chamber*. Collects and disposes of the water vapor produced during drying stages.
3. Shelf heating and cooling unit. Permits the heat transfer required for the freeze-drying cycle.
4. *Vacuum pumping unit*. Evacuates both the drying chamber and the ice condenser chamber to operating pressure and removes noncondensables such as air.
5. *Refrigeration unit*. Provides refrigeration for the chamber shelves during the freezing stage, and ice condenser plates for the primary and secondary drying stages.
6. *Instruments and control panel*. Indicate, record, and control shelf, product, and condenser temperature, as well as drying chamber, ice condenser chamber, and vacuum unit pressure.

A typical freeze-drying equipment lay-out (Criofarma model C292-8S) is showed in Figure 5.

The user of a pharmaceutical freeze-dryer must observe current cGMP for processes and equipment to be validated before and during use. As is the case in any validation effort, the protocol includes sections for installation, operation, and process. Each section of the protocol needs to be comprehensive and sufficiently detailed. The objectives of the validation program are to adequately demonstrate the operation's ability to support the process and to develop documentation demonstrating that control of the unit's operation can be maintained. To develop and implement a protocol for lyophilization, careful consideration should be given to the installation qualification (IQ) and operational qualification (OQ) sections. The IQ section is the foundation of the protocol. Documentation of a general description, specifications, the support utility services, the equipment subsystems, and the installation of the lyophilization system should be addressed within the IQ section. This portion of the protocol is similar to that of any operation being validated. As part of the installation, an initial calibration should be performed to ensure that the instrumentation is in working condition after shipment and to ensure the accuracy of any data collected during start-up of the system. Calibration procedures and results also should be documented.

The OQ section can begin after successful completion of the installation qualification portion of the validation protocol. The objective of the operational qualification is to assure that the system performance is adequate to support the process for which the system is intended. The performance of the system needs to be compared either to the original equipment manufacturer's specification or to the process parameter requirements as noted in the acceptance criteria. Tests should be performed on shelf temperature control, condenser cool-down, vacuum pumping rates, vacuum integrity, sublimation and condensation rates, vacuum level control, steam-in-place (SIP) system,

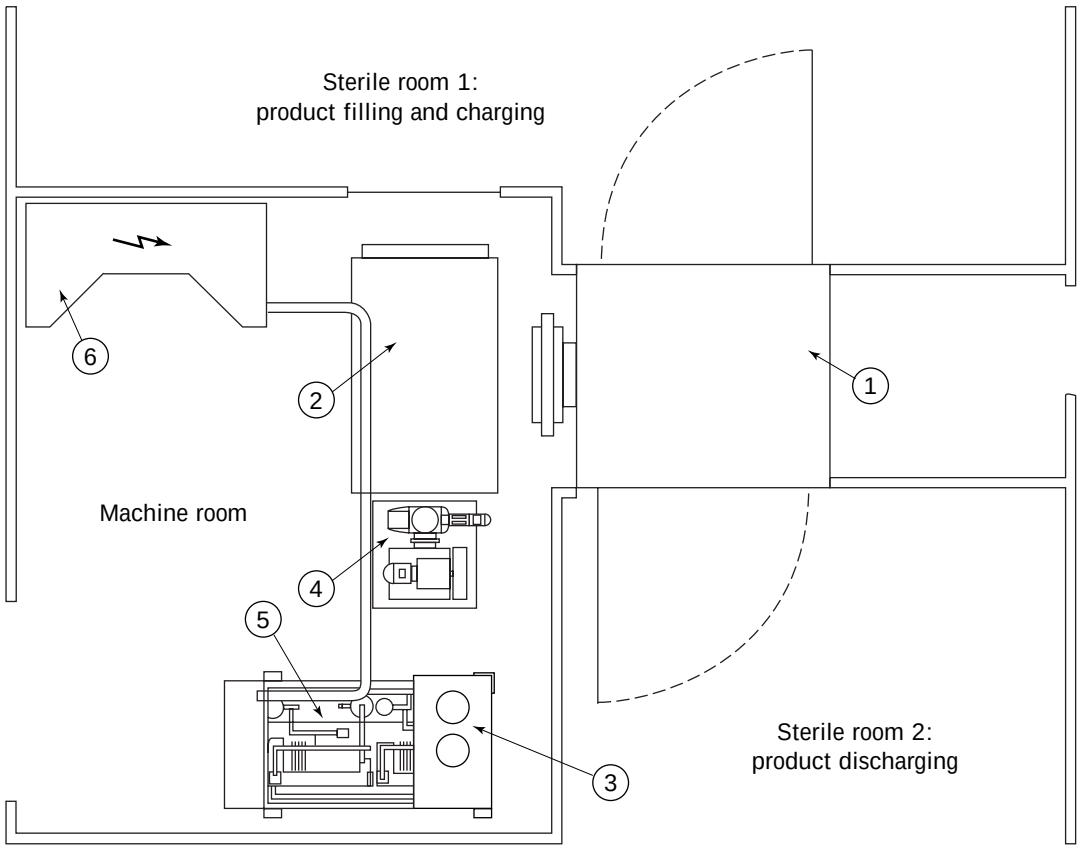


Figure 5. Lay-out of an industrial double-door freeze-dryer (Criofarma model C292-8S): 1, drying chamber with double full doors; 2, ice condenser chamber; 3, shelf heating and cooling unit; 4, vacuum pumping unit; 5, refrigeration unit; 6, instrument and control panel. Courtesy of Criofarma.

clean-in-place (CIP) system (cleaning the inside part of the plant with flushing steam or sterile water), and control hardware and software. The tests for validation of control hardware and software basically require suitable computer hardware and computer programs that perform consistently within preestablished operational limits so that analysis of the effects of possible failures can be carried out. If the tests are conducted in a manner that is in compliance with the procedures and that satisfies the acceptance criteria, the system will perform adequately at levels required to support the process.

The process validation section can be done as concurrent validation or as retrospective validation. Concurrent validation is applicable in situations in which the process data are collected and analyzed for a product that is to be marketed. Retrospective validation, which is particularly applicable to validation of lyophilization, is achieved through analysis of historical batch records. In this method, the data are gathered from records and then evaluated and analyzed. It is important to note that the validation of lyophilization processes is a sophisticated area that requires an understanding of the equipment, the process, and the product if a validation program for the process is to be implemented successfully.

Useful information for cGMP compliance, process, and computer system validation, can be found in Refs. 30–37.

FREEZE-DRYING OF PHARMACEUTICALS

Freeze-drying of pharmaceuticals can be divided according to the kind of product containers used to perform the process. It is possible to have freeze-drying of pharmaceutical products on trays, and freeze-drying of pharmaceutical products in vials. The different amount and kind of product processed (usually raw materials are freeze-dried on trays, whereas vaccines or drugs are freeze-dried in vials) and the different geometry of the container are very important relative to the process parameters and drying rates.

In Figure 1, a material being freeze-dried in a tray is shown. The thickness of the sides and bottom of the tray, as well as the material from which the tray is made, are most often in practice such that the resistance of the tray to heat transfer can be considered to be negligible (1–2,5–6,8–10, 38–39). Heat q_I can be supplied to the surface of the dried layer by conduction, convection, and/or radiation from the gas phase; this heat is then transferred by conduction to the frozen layer. Heat q_{II} is supplied by a heating plate and is conducted through the bottom of the tray and through the frozen material to reach the sublimation interface or plane. The magnitude of the amount of heat q_{III} in the vertical sides of the tray is much smaller (1–2,5–6,8–10,38–39) than that of q_I or q_{II}; q_{III} represents the amount of heat transferred between the environment in

the drying chamber and the vertical sides of the tray. Because the contribution of q_{III} is rather negligible when compared with the contributions of q_I and q_{II}, the contribution of q_{III} to the drying rate will not be further considered (1–2,8–9). The terms N_W and N_t in Figure 1 represent the mass flux of water vapor and the total mass flux, respectively, in the dried layer. The total mass flux is equal to the sum of the mass fluxes of water vapor and inert gas, $N_t = N_W + N_{in}$, where N_{in} denotes the mass flux of the inert gas.

In Figure 6, a typical lyophilizing vial (40) is shown during freeze-drying. The energy of sublimation of ice and desorption of sorbed water is supplied to the vial through radiation from the two shelf surfaces above and below the vial, conduction from vial–vial and vial–shelf contacts, and gaseous transfer along the bottom and side. The thickness of the pharmaceutical product charged into the vial usually is equivalent to the diameter of the vial. In pharmaceutical freeze-drying in vials, therefore, the heat transferred between the environment in the drying chamber and the surface of the vertical sides of the vial cannot be considered negligible (as in the case of pharmaceutical freeze-drying on trays) when compared with the contribution of the heat transferred between the dryer shelf and the surface of the bottom of the vial. For this reason the sublimation interface at time $t = 0$ is planar (as in the case of pharmaceutical freeze-drying on trays), but after the lyophilization process progresses, the moving interface acquires a parabolic shape because of energy entering along the sides of the vial (3,6,11). The traditionally accepted course of the moving sublimation interface is approximately 0-1-2-3-4. Experimental results (41–42) suggest that the dried-product mass transfer resistance is the dominant resistance, accounting for 82–92% of the overall mass transfer resistance, even when small (13-mm) closures are used. Closure and drying chamber mass transfer resistances are estimated (42) to be of the same order of magnitude, and in a well-designed freeze-dryer the chamber resistance (1–3,6) should be less than the resistance of the closure. A simple but satisfactory expression for estimating the resistance of the closure can be found in Ref. 42.

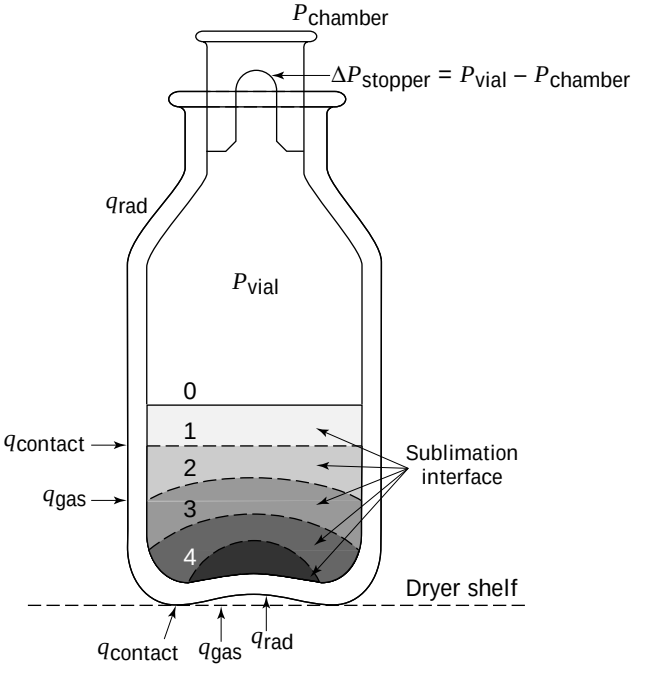


Figure 6. Modes of energy transfer and traditional view of movement of the sublimation interface in a vial during freeze-drying.

The goal of the pharmaceutical process designer and of the processor is to provide an economical drying process that gives reliably uniform and high product quality (1–4,18–20,43–44). A knowledge of the basic phenomena and mechanisms involved in freeze-drying is essential. A qualitative description and a mathematical model of the freeze-drying process on trays can be found in Refs. 1–3, 6, 9–10, 13, 15–16, and 39; the model can be used to analyze rates of freeze drying of pharmaceuticals on trays (1–6,8–11,13–16,19,45–48). Qualitative description and a mathematical model of the freeze drying process in vials can be found in Refs. 3, 6, 9, 11, and 49.

TEMPERATURE AND PRESSURE CONSTRAINTS IN FREEZE-DRYING

A point that should be stressed is that, in any freeze-drying process, it will be desirable to fix the design and operating conditions such that the process is not rate limited by external resistances to either heat or mass transfer. The internal heat and mass transfer resistances are characteristic of the material being dried, but the external resistances are characteristic of the freeze-dryer. The design conditions of the equipment refer to having appropriate capacities for the vacuum pump, water vapor condenser, and heaters, and that the spacings between trays are such that the external heat and mass transfer resistances are not significant.

It is important to note at this point that, because a batch of a pharmaceutical product in an industrial freeze-dryer can easily be worth significant amounts of money, it is of paramount importance that the units and the control systems always operate under conditions at which there is insignificant loss in the quality of the product being freeze-dried. For this purpose, the freeze-dryer usually has one additional refrigeration and/or vacuum unit in a standby condition, and the control instrumentation is designed in such a way that the control policies can be implemented either automatically or by manual override.

In a well-designed freeze-dryer, the external resistances should not be the controlling factor in determining the drying rate (1–4,8–16). The internal heat and mass transfer resistances control the drying rate, and therefore, the freeze-drying process should be conducted within certain temperature and pressure limits in order to avoid product damage during the process.

Two temperature constraints (limits) may possibly be reached during the primary drying stage. First, the surface temperature of the dried layer must not become too high because of the risk of thermal damage; in fact the surface temperature of the dried layer must be kept below the scorch temperature, T_{scor}, of the material being dried. Second, the temperature of the interface, T_X, must be kept

well below the melting temperature, T_m. If the outer surface temperature limit (T_{scor}) is encountered first as the surface temperature of the dried layer is raised, the process is considered to be heat transfer controlled; to further increase the drying rate, the thermal conductivity, k_{Ie}, of the dried layer must be raised (50,51). Many commercial freeze-drying processes are heat transfer controlled (1–6). If the melting point temperature, T_m, is encountered first, then the process is considered to be mass transfer limited, and in order to increase the drying rate, the effective diffusivity of water vapor in the dried layer, $D_{win,e}$, and the total mass flux, N_t, must be raised (an increase in N_t implies that the convective velocity of the vapor in the pores of the dried layer is increased); the values of $D_{win,e}$ and N_t could be raised by decreasing the pressure in the drying chamber (1,2,43). Because of the risks of lost bioactivity or structural and chemical damage of the product, the frozen layer temperature must be maintained below the melting point, which may in some cases be 10 °C or more below the melting point of ice (1–6, 8–17). Typical ice temperatures existing in the freeze-drying of pharmaceuticals under conditions in which the total pressure was primarily due to water vapor and the heat transfer took place via the dried (I) and frozen (II) layers are shown in Table 1.

The end of the primary drying stage occurs (1–3,5,6,10) when the position of the moving sublimation interface is at $X = L$; this condition implies that at the end of the primary drying stage there is no frozen (II) layer, and therefore, there is no (moving) sublimation interface.

During the secondary drying stage, the temperature everywhere in the dried (I) layer must be kept below the value of the scorch temperature, T_{scor}.

The pressure constraints (limits) are more complex to define because of the pressure effect on both the heat transfer (pressure has effect on the thermal conductivity, k_{Ie}, of the dried layer) and the mass transfer (pressure has effect on the effective diffusivity of water vapor in the dried layer, $D_{win,e}$, and the total mass flux, N_t).

The effective thermal conductivity, k_{Ie}, in the dried material (51) has been found to vary significantly with the total pressure and with the type of gas present. At very low pressures the thermal conductivity reaches a lower asymptotic value independent of the surrounding gas. This asymptotic conductivity reflects the geometric structure of the solid matrix itself, with no contribution from the gas in the voids of the material because the gas pressure is so low. At high pressures the thermal conductivity levels out again at a higher asymptotic value. This higher asymptote is characteristic of the heterogeneous matrix composed of solid material and the gas in the voids. Consequently, the high-pressure thermal conductivity is dependent upon the nature of the gas present and increases as the thermal conductivity of the gas increases and, hence, as the molecular weight of the gas decreases. When the thermal conductivity attains the high-pressure asymptotic value, the mean free path of the gas molecules within the void spaces of the dried layer has become substantially less than the dimensions of the void spaces. During the transition in thermal conductivity from the low-pressure asymptote to the high-pressure asymptote, the mean free path of the gas molecules rivals the void space dimensions in magnitude, but once the mean free path is reduced to the point at which the gas phase within the solid matrix obeys simple kinetic theory, the thermal conductivity stops rising. This reflects the fact that the thermal conductivity of a gas obeying simple kinetic theory is independent of the pressure. The transition in thermal conductivity between asymptotes usually occurs between 0.1 and 100 mmHg, which includes the pressures characteristic of freeze-drying processes. The pressure range over which the transition in thermal conductivity between asymptotes occurs is characteristic of the pore-size distribution of the void spaces within the freeze-dried material (52). A smaller pore dimension means that the gas must achieve a higher pressure in order for the mean free path of the gas to become comparable to the pore spacing and, hence, means that the transition between asymptotes will occur at higher pressures. Because fast freezing before freeze-drying leads to smaller pore spacing after freeze-drying (53,54), it follows that faster freezing should lead to lower thermal conductivities at a given pressure. If a freeze-drying process is rate limited by internal heat transfer, the rate of freeze-drying for fast-frozen material should then be less than that of a slowly frozen material (45,46). Thermal conductivity data for freeze-dryed substances and pharmaceuticals are reported in Ref. 2.

The pressure effect on the mass transfer is explained by the fact that the effective diffusivity (2,6) of water vapor in the dried layer, $D_{win,e}$, and the total mass flux, N_t, are a function of the total pressure, P, in the drying chamber (43). Elaborate expressions of $D_{win,e}$ and $D_{inw,e}$ can be found in Refs. 2, 6, 9, 11, 15, and 16. In all cases the magnitude of $D_{win,e}$ and $D_{inw,e}$ decrease if the total pressure, P, is increased. If the value of $D_{win,e}$ is decreased because of increased total pressure, then (2,6) the value of the effective pore diffusivity, $D_{win,e}$, decreases as the total pressure is

Table 1. Frozen Layer and Maximum Dry Surface Temperatures in a Typical Freeze-Drying Operation Conducted with Heat Input through the Dry and Frozen Layers

Pharmaceutical material	Chamber pressure (mm Hg)	Maximum surface temperature (°C)	Frozen-layer temperature (°C)
Ampicillin sodium salt	0.15	40	−24
Cephazolin sodium salt	0.15	40	−25
Cloxacillin sodium salt	0.20	40	−20
Collagen	0.30	70	−20
Fructose-1,6-diphosphate	0.05	30	−45
Lactobacillus	0.10	20	−25
Rubella vaccine	0.05	10	−25

increased. Therefore, if the total pressure in the drying chamber is increased, then the effective diffusivity, $D_{win,e}$, decreases, and thus, the diffusional mass flux of water vapor in the dried layer decreases. Furthermore, when the total pressure in the drying chamber is increased, the gradient of the total pressure, $\partial P/\partial x$, in the dried layer is reduced, and this decreases the convective velocity, v_p, and the total mass flux, N_t. Because the freeze-drying process will become internal mass transfer controlled above a certain pressure ($D_{win,e}$ and N_t decrease with increasing pressure, and k_{Ie} increases with pressure), the highest rate under mass transfer control will occur at the pressure of transition from heat transfer control to mass transfer control, and the attainable drying rate will decrease at higher pressures.

In general, operating conditions in freeze-drying pharmaceuticals include maximum surface temperatures of 20–40 °C and chamber pressures of 0.1–1 mmHg. Freeze drying of biological specimens, vaccines, and microorganisms is usually conducted with maximum surface temperatures of -20 to $+20$ °C and chamber pressures below 0.1 mmHg.

There has been considerable interest in investigating the factors affecting the batch time of freeze-dryers because this variable is most amenable to control, and efforts have bene made to minimize the batch time (2,5,10,13,15,16,45–47).

The heat variables q_I and q_{II} from the energy sources, and the drying chamber pressure P_{ch}, $P_{ch} \cong P^0 = P^0_w + P^0_{in}$ ($P_{ch} = P^0$ when the external mass transfer resistance is insignificant, as it would be the case with a well-designed freeze-dryer) are natural control variables. It should be emphasized at this point that (as the equations of the mathematical models in Ref. 2 indicate) the effects on the heat and mass transfer rates resulting from changes in the values of q_I, q_{II}, and P_{ch} are coupled. The variable P^0_w is taken to represent the water vapor pressure in the drying chamber (the external mass transfer resistance is taken to be insignificant), and its value is determined by the design and the operational temperature of the ice condenser. Thus, P_{ch} may be changed by changes in P^0_w (P^0_w could be changed by changes in the temperature of the ice condenser), and by increasing or decreasing P^0_{in}. Therefore, changes in the temperature of the ice condenser affect the pressure, P_{ch}, (through P^0_w) in the drying chamber, and thus, the mass transfer rate in the dried layer.

Liapis and Litchfield (13) performed a quasi-steady-state analysis for a system where $q_I \neq 0$ and $q_{II} = 0$ and obtained general guidelines about the optimal control policy at the beginning of the drying process (when neither of the state constraints is active), as well as during operation, when the process may be heat and/or mass transfer limited (10,13).

Millman et al. (10) studied the freeze-drying of skim milk under various operational policies that included the case where $q_I \neq 0$ and $q_{II} \neq 0$. They found that the control policy that produced the shortest primary drying stage was also the policy that provided the shortest overall drying time. Their results show that at least 80% of the heat used during the primary drying stage was transferred through the frozen layer of the sample. They also showed that the type of criterion used in terminating the secondary drying stage is of extreme importance, especially for samples of large thicknesses because it may lead to an undesirable sorbed (bound) water profile that may deteriorate the quality of the dried product.

CONCLUSION

The evolution of freeze-drying in the past 40 years indicates that this separation process (unit operation) is a convenient method for drying those decomposable products (mostly pharmaceuticals) that cannot be stabilized in any other way (e.g., plasma, vaccines, antibiotics, sera, and growth hormones).

The economics of the process indicate that freeze-drying can be suitable for high-value products with specific biological and/or physicochemical properties. The highest cost advantages would be obtained from the processing of concentrated solutions of expensive materials; in this respect, freeze-drying could represent a viable alternative to filtration and crystallization. It is certain that freeze-drying has a future, and it is likely that this will essentially be in the fields of biological sciences and biotechnology. Its potential evolution is still great and will, of course, depend upon progress in basic research (4) and upon the level of creativity in the design and operating conditions of plants and instruments.

NOMENCLATURE

$D_{inw,e}$	Effective pore diffusivity of a binary mixture of inert gas and water vapor in the dried layer, m^2/sec
$D_{win,e}$	Effective pore diffusivity of a binary mixture of water vapor and inert gas in the dried layer, m^2/sec
k_{Ie}	Effective thermal conductivity in the dried layer, kW/(m K)
k_{II}	Thermal conductivity in the frozen layer, kW/(m K)
L	Sample thickness, m
M_{in}	Molecular weight of inert gas, kg/kg mol
M_w	Molecular weight of water vapor, kg/kg mol
N_{in}	Mass flux of inert gas in the dried layer, kg/(m^2 s)
N_t	Total mass flux in the dried layer ($N_t = N_{in} + N_w$), kg/(m^2 s)
N_w	Mass flux of water vapor in the dried layer, kg/(m^2 s)
p^0_{in}	Partial pressure of inert gas at $x = 0$, N/m^2
p^0_w	Partial pressure of water vapor at $x = 0$, N/m^2
p_{wX}	Partial pressure of water vapor in equilibrium with the sublimation front ($p_{wX} = g(T_X)$), N/m^2
P	Total pressure ($P = p_{in} + p_w$) in the dried layer, N/m^2

p^0	Total pressure at $x = 0$, N/m^2
p_{ch}	Total pressure in the drying chamber, N/m^2
$P_{chamber}$	Total pressure in the drying chamber, N/m^2
P_{vial}	Vial pressure, N/m^2
q_I	Heat flux at $x = 0$, kW/m^2
q_{II}	Heat flux at the bottom of the tray, kW/m^2
q_{III}	Heat flux at the sides of the tray, kW/m^2
$q_{contact}$	Conduction heat flux, kW/m^2
q_{gas}	Heat flux due to gaseous transfer, kW/m^2
q_{rad}	Radiation heat flux, kW/m^2
t	Time, s
T_I	Temperature in the dried layer, K
T_{II}	Temperature in the frozen layer, K
T_m	Melting temperature, K
T_{scor}	Scorch temperature, K
T_X	Temperature of the sublimation front, K
v_p	Convective velocity of the binary mixture of water vapor and inert gas in the porous dried layer, m/sec
X	Position of sublimation front (interface), m

Greek Symbols

ΔP	Pressure drop, N/m^2
ΔT	Total temperature difference in Figure 2, K
ΔT_{ice}	Temperature difference through the layer of ice on the cold surface, K
ΔT_{refr}	Temperature difference between the cold surface and the evaporating refrigerant, K

Subscripts

I	Dried layer
II	Frozen layer

BIBLIOGRAPHY

1. A.I. Liapis and R. Bruttini, *Sep. Technol.* **4**, 144–155 (1994).
2. A.I. Liapis and R. Bruttini, in A.S. Mujumdar ed., *Handbook of Industrial Drying* 2nd ed., Marcel Dekker, New York and Basel, 1995, pp. 309–343.
3. A.I. Liapis and R. Bruttini, *Drying Technol.* **13**, 43–72 (1995).
4. A.I. Liapis, M.J. Pikal, and R. Bruttini, *Drying Technol.* **14**, 1265–1300 (1996).
5. J.D. Mellor, *Fundamentals of Freeze-Drying*, Academic Press, London, 1978.
6. A.I. Liapis and R. Bruttini, in I. Turner and A.S. Mujumdar eds., *Mathematical Modelling and Numerical Techniques in Drying Technology*, Marcel Dekker, New York and Basel, 1997, pp. 481–535.
7. S.A. Goldblith, L. Rey, and W.W. Rothmayr, *Freeze-Drying and Advanced Food Technology*, Academic Press, London, 1975.
8. R. Bruttini, G. Rovero, and G. Baldi, *Chem. Eng. J.* **45**, B67–B77 (1991).
9. R. Bruttini, Ph.D. Thesis, University of Missouri, Rolla, Mo., 1994.
10. M.J. Millman, A.I. Liapis, and J.M. Marchello, *AIChE J.* **31**, 1594–1604 (1985).
11. M.M. Tang, A.I. Liapis, and J.M. Marchello, in A.S. Mujumdar ed., *Proceedings of the Fifth International Drying Symposium*, vol. 1, Hemisphere Publishing, New York, 1986, pp. 57–65.
12. C.J. King, 1971, *Freeze-Drying of Food*, Chemical Rubber Co. Press, Cleveland, Ohio, 1971.
13. A.I. Liapis and R.J. Litchfield, *Chem. Eng. Sci.* **34**, 975–981 (1979).
14. M.J. Pikal, S. Shah, M.L. Roy, and R. Putman, *Int. J. Pharm.* **60**, 203–217 (1990).
15. R.J. Litchfield and A.I. Liapis, *Chem. Eng. Sci.* **34**, 1085–1090 (1979).
16. R.J. Litchfield and A.I. Liapis, *Chem. Eng. Sci.* **37**, 45–55 (1982).
17. J.W. Snowman, in Avshalom Mizrahi ed., *Downstream Processes: Equipment and Techniques*, Alan R. Liss, New York, 1988, pp. 315–351.
18. M.J. Pikal, in L. Cleland and R. Langer ed., *Formulation and Delivery of Proteins and Peptides*, ACS Symposium Series 567, American Chemical Society, Washington, D.C., 1994, pp. 120–133.
19. C.D. Holland and A.I. Liapis, *Computer Methods for Solving Dynamic Separation Problems*, McGraw-Hill, New York, 1983.
20. P.A. Belter, E.L. Cussler, and W.S. Hu, *Bioseparations: Downstream Processing for Biotechnology*, Wiley-Interscience, New York, 1988.
21. R.J. Litchfield, F.A. Farhadpour, and A.. Liapis, *Chem. Eng. Sci.* **36**, 1233–1238 (1981).
22. R.G. Livesey and T.W.G. Rowe, *J. Parenteral Sci. Technol.* **41**, 169–171 (1987).
23. J.H. Petropoulos, J.K. Petrou, and A.I. Liapis, *Ind. Eng. Chem. Res.* **30**, 1281–1289 (1991).
24. J.H. Petropoulos, A.I. Liapis, N.P. Kolliopoulos, J.K. Petrou, and N.K. Kanellopoulos, *Bioseparation* **1**, 60–69 (1990).
25. M.L. Roy and M.J. Pikal, *J. Parenteral Sci. Technol.* **43**, 60–66 (1989).
26. C.C. Freyrichs and C.N. Herbert, *J. Biol. Stand.* **2**(1), 59 (1974).
27. J.W. Johnson, J.F. Arnold, S.L. Nail, and E. Renzi, *J. Parenteral Sci. Technol.* **46**, 215 (1992).
28. A. Bardat, R. Schmitthaeusler, and E. Renzi, *PDA J. Pharm. Sci. Technol.* **50**(2), 83–88 (1996).
29. N. Milton, S.L. Nail, M.L. Roy, and M.J. Pikal, *Pharm. Res.* **10**(S-170), 6174 (1993).
30. J.Y. Lee, *Pharmaceutical Technology*, October, 54–60 (1988).
31. U.S. Federal Drug Administration, *Lyophilization of Parenterals, Inspection Technical Guide*, U.S. Department of Health and Human Services, Washington, D.C., 1986.
32. U.S. Food and Drug Administration, *Guideline on General Principles of Process Validation*, U.S. Department of Health and Human Services, Washington D.C., 1987.
33. U.S. Food and Drug Administration, *Guide to Inspection of Bulk Pharmaceutical Chemical Manufacturing*, U.S. Department of Health and Human Services, Washington D.C., 1987.
34. E.H. Trappler, *Pharmaceutical Technology*, January, 56–60 (1989).
35. K.G. Chapman and J.R. Harris, *Pharmaceutical Technology International*, May/June, 54–58 (1989).
36. J.S. Alford, and F.L. Cline, *Pharm. Technol.* **14**(9), 88 (1990).
37. P.M. Masterson, *Pharmaceutical Technology*, January, 48–54 (1989).
38. A.I. Liapis and R. J. Litchfield, *Comput. Chem. Eng.* **3**, 615–621 (1979).

39. A.I. Liapis and J.M. Marchello, in J.C. Ashworth ed., *Proceedings of the Third International Drying Syposium*, vol. 2, Drying Research Limited, Wolverhampton, England, 1982, pp. 479–486.

40. M.D. Cise, *AIChE Summer National Meeting*, Cleveland, Ohio, 1982, paper 22d.

41. M.J. Pikal, S. Shah, D. Senior, and J.E. Lang, *J. Pharm. Sci.*, **72**, 635–650 (1983).

42. M.J. Pikal, M.L. Roy, and S. Shah, *J. Pharm. Sci.* **73**, 1224–1237 (1984).

43. C.J. Geankoplis, *Transport Processes and Unit Operations*, Allyn and Bacon, Boston, 1983.

44. A.I. Liapis, *Sep. Purif. Methods*, **19**(2), 133 (1990).

45. M.J. Millman, A.I. Liapis, and J.M. Marchello, *J. Food Technol.* **19**, 725 (1984).

46. M.J. Millman, A.I. Liapis, and J.M. Marchello, *J. Food Technol.* **20**, 541 (1985).

47. M.J. Millman, Ph.D. Thesis, University of Missouri, Rolla, Mo., 1994.

48. H. Sadikoglu and A.I. Liapis, in C. Strumillo and Z. Pakowski ed., *Drying 96*, vol. B, Marcel Dekker, New York and Basel, 1996, p. 1284.

49. J.L. Duda, M.F. Malone, R.H. Notter, and J.S. Vrentas, *Int. J. Heat Mass Transfer* **18**, 901–910 (1975).

50. A.V. Luikov, A.G. Shashkov, L.L. Vasiliev, and Y.E. Fraiman, *Int. J. Heat Mass Transfer* **11**, 117 (1968).

51. L. Marcussen, in T. Ashworth and D.R. Smith eds., *Thermal Conductivity*, vol. 18, Plenum Press, New York, 1985, pp. 585–598.

52. C.-J. King, W.K. Lam, and O.C. Sandall, *Food Technol.* **22**, 1302 (1968).

53. C.G. Haugh, C.S. Huber, W.J. Stadelman, and R.M. Peart, *Trans. ASAE* **11**, 877 (1968).

54. G. Lusk, M. Karel, and S.A. Golblith, *Food Technol.* **19**, 620 (1965).

See also CULTURE COLLECTIONS; GOOD MANUFACTURING PRACTICE (GMP) AND GOOD INDUSTRIAL LARGE SCALE PRACTICE (GLSP).

β-GALACTOSIDASE, ENZYMOLOGY

TORU NAKAYAMA
Tohoku University
Sendai, Japan

TERUO AMACHI
Kyoto University
Kyoto, Japan

KEY WORDS

Functional modification of bioactive compounds
Galactooligosaccharides
β-Galactosidase
Galactosyllactose
Glycosylhydrolase family
Lactase
Lactose hydrolysis
Microbial enzyme
Transgalactosylation
Reaction mechanism

OUTLINE

INTRODUCTION

β-Galactosidase (lactase, EC 3.2.1.23) catalyzes the hydrolysis and transgalactosylation reactions of β-D-galactopyranosides, such as lactose. The enzyme occurs in a wide variety of organisms including microorganisms, plants, and animals. The application of β-galactosidase in bioprocess technology has been achieved exclusively with microbial enzymes, which have long been used for the hydrolysis of lactose for increasing the digestibility of milk or for improving the functional properties of dairy products. During the past decade, another potential application of the enzyme has also been developed; the β-galactosidase-catalyzed transgalactosylation has proved to be useful for structural and functional modifications of food materials, medicines, and other biologically active compounds. In this article, we survey the enzymology and application in bioprocess technology of the microbial β-galactosidases with special emphasis on application of their transgalactosylation.

BRIEF SURVEY OF β-GALACTOSIDASE ENZYMOLOGY

Microbial Distribution and Purification

β-Galactosidase occurs in a variety of microorganisms including yeasts, molds, bacteria, actinomycetes, and archaebacteria (Table 1). β-Galactosidase is known as a marker enzyme for coliform bacteria, which are indicators of the fecal pollution of water.

Microbial β-galactosidases have been purified by a combination of several conventional techniques, such as salting-out fractionation, ion exchange, gel filtration, hydroxyapatite, and hydrophobic interaction chromatographies. The affinity chromatography of β-galactosidase uses *p*-aminophenyl- (or *p*-aminobenzyl-)β-D-thiogalactopyranoside agarose (1), which is commercially available and has been used for the purification of the *Escherichia coli lacZ* β-galactosidase, its recombinant derivatives, and the other β-galactosidases (2).

Molecular Properties, Substrates, and Inhibitors

The specificity and molecular properties of the microbial β-galactosidases differ significantly with the source of the enzyme. For example, the *Saccharopolyspora rectivirgula* β-galactosidase is highly specific for β-D-galactosides (3), whereas the *Sulfolobus solfataricus* enzyme shows broad substrate specificity and can efficiently hydrolyze β-glycosides other than β-galactosides as well (4). The catalytic competence and specificity of transgalactosylation also depend on the enzyme (see following). The reported native molecular weights and subunit structures of β-

Table 1. Examples of β-Galactosidase-Producing Microorganisms

Yeasts

Candida pseudotropicalis, Cryptococcus laurentii, Kluyveromyces (Saccharomyces) lactis, K. fragilis, K. marxianus, Lipomyces sp., *Torulopsis sphaerica, T. versalitis*

Molds

Altenaria altenata, Aspergillus awamori*, A. cellulosae*, A. foetidus*, A. nidulans, A. niger, A. oryzae, A. phoenicis*, A. terreus,* A. wentii*, Aspergillus* sp., *Chaetomium globosum*, C. cochlioides*, C. funicola*, C. thermophile* var. *coprophile*, Fusarium maniliforme, F. oxysporum* var. *lini, Geotricum candida*, Humicola grisea* var. *thermoidea*, H. lanuginosa*, Macrophomina phaseoli, Malbranchea pulchella* var. *sulfurea*, Mucor miehei*, M. pusillus*, Mucor* sp.**, Mucor muccedo*, M. javanicus*, Neurospora crassa, Paecilomyces varioti, Penicillium* sp., *P. multicolor, P. canescens, P. citrinum, P. luteum*, P. chrysogenum*, P. frequentans*, P. cycropium*, P. toxidarium*, P. glaucum*, P. notatum*, P. roqueforti*, Phycomyces blakeleeanus, Rhizopus acidus*, R. niveus*, R. nigricans*, R. delemar*, R. javanicus*, R. formosciensis*, R. chinensis*, Scopulariopsis* sp., *Sclerotium tuliparum, Spicaria* sp., *Sporotrichum* sp.**, S. thermophile*, Sterigmatomyces elviae, Thermomyces lanuginosus, Torula thermophila*, Trichoderma viride**

Basidiomycetes

*Corticium rolfsii, Culvularia inaequalis, Pycnoporus cinnabarinus, Sporobolomyces singulalis**

Bacteria

Gram-negative

Aeromonas cavie, A. formicans, Agrobacterium rediobacter, Bacteroides polypragmatus, Buttiauxella agrestis, Enterobacter* (*Aerobacter*) *cloacae, Escherichia coli, Fibrobacter succinogenes, Klebsiella pneumoniae, Rhizobium meliloti, R. trifolii, Shigella dysenteriae*, Thermotoga maritima, Thermus* sp., *Treponema phagedenis, Xanthomonas campestris, X. manihotis*

Gram-positive

Arthrobacter sp., *Bacillus acidocaldarius, B. circulans, B. coagulans, B. macerans, B. megaterium, B. subtilis, B. stearothermophilus,* alkalophilic *Bacillus, Bifidobacterium* sp., *B. bifidum*, B. longum, Clostridium acetobutylicum, Corynebacterium murisepticum, Lactobacillus delbruckii* subsp. *bulgaricus, L. bulgaricus, L. casei, L. helveticus, L. murinus, L. plantarum, L. sake, Lactococcus lactis, Leuconostoc citrovorum, L. lactis, Streptococcus salivarius* subsp. *thermophilus, S. thermophilus, S. (Diplococcus) pneumoniae, Thermoanaerobacter* sp., *Thermoanaerobacterium (Clostridium) thermosulfurigenes*

Actinomycetes

Actinomyces viscosus, Nocardia sp., *Saccharopolyspora rectivirgula, Streptomyces lividans*, S. venezuelae*, S. violaceus**

Archaebacteria

Caldariella acidophila, Sulfolobus solfataricus, Pyrococcus woesei, Haloferax alicantei

Note: Microorganisms listed include not only those whose β-galactosidase is characterized or sequenced but also those that have been only shown to produce β-galactosidase(s) based on the observed hydrolytic and transgalactosylation activities (indicated by an asterisk; May 1998).

galactosidases range, respectively, from 19,000 to 630,000 (the *Kluyveromyces lactis* enzyme [5]) and from monomer (e.g., the *S. rectivirgula* enzyme [3]) to heterooctamer (the *E. coli ebg* enzyme [6]). The bacterial β-galactosidases generally show optimum pHs at neutral regions whereas most mold enzymes show them at acidic regions, which in some cases reach pH 2 (7). Some β-galactosidases of the *lacZ* family require essential monovalent and divalent ions for their activities; the *S. rectivirgula* β-galactosidase is a multimetal enzyme that requires multiple divalent ions for its maximum thermostability and activity (8). However, some other β-galactosidases (e.g., the *Rhizobium meliloti* enzyme [9]) do not show such metal ion requirements. These observations strongly suggest the structural diversity of the microbial β-galactosidases.

Many chromogenic, fluorogenic, and luminogenic substrates that are specific for β-galactosidase have been designed (Table 2). D-Galactose, D-galactosylamines, β-D-thiogalactosides, and their analogs and derivatives serve as specific inhibitors of some microbial β-galactosidases (Table 3).

Structural Classification

Henrissat has compared all the available sequences of glycosyl hydrolases using hydrophobic cluster analysis and classified these enzymes into families (48). Many microbial β-galactosidases have thus far been cloned and sequenced and have been classified, on the basis of sequence similarities, into at least four categories according to Henrissat's classification (Table 4).

Glycosyl Hydrolase Family 2. β-Galactosidases belonging to this glycosyl hydrolase family, typified by the *E. coli lacZ* β-galactosidase, consist of a large subunit protein of approximately 1,000 amino acids and show very high sequence similarities to the animal β-glucuronidases, suggesting that the β-galactosidases of family 2 have a close evolutionary relationship with the β-glucuronidases (48). Many β-galactosidases of this family require monovalent and divalent metal ions for maximum activity (e.g., the *E. coli lacZ, Kluyveromyces lactis, Streptococcus thermophilus,* and *Saccharopolyspora rectivirgula* enzymes).

Table 2. Chromogenic, Fluorogenic, and Luminogenic Substrates for β-Galactosidase Gal, β-D-Galactopyranoside

A. Chromogenic Substrates

Substrate	Reported molar absorption coefficient of chromophore[a], M^{-1} cm^{-1}	Reference
2-Nitrophenyl-Gal (ONPG)	$\Delta\epsilon_{420}$ = 3000 (pH 7.0)	10
	$\Delta\epsilon_{410}$ = 2100 (pH 7.0)	11
	$\Delta\epsilon_{410}$ = 4500 (pH 8.6)	11
	$\Delta\epsilon_{410}$ = 4600 (pH 9.2)	11
4-Nitrophenyl-Gal (PNPG)	$\Delta\epsilon_{347,iso}$ = 290.4[b]	12
	$\Delta\epsilon_{405}$ = 13400 (pH 7.2)	3
	$\Delta\epsilon_{405}$ = 8900 (pH 7.0)	11
	$\Delta\epsilon_{405}$ = 18300 (pH 8.6)	11
3,4-Dinitrophenyl-Gal	$\Delta\epsilon_{400}$ = 15924 (pH 7.0[c])	12
2-Chloro-4-nitrophenyl-Gal		13
X-Gal[d]	Blue insoluble product	Suitable for histochemical identification
VBzTM-Gal[e]	Red product, $\Delta\epsilon_{520}$ = 55000	Suitable for histochemical identification; Refs. 14 and 15
VLM-Gal[f] and derivative	$\Delta\epsilon_{545}$ = 32500 (VLM-Gal)	15
VQM-Gal[g] and derivative	$\Delta\epsilon_{520}$ = 42500 (VQM-Gal)	15
6-Bromo-2-naphthyl-Gal	Product visualized by subsequent reactions	Suitable for activity staining in the gel; Refs. 16 and 17

B. Fluorogenic Substrates

Substrate	Excitation (nm)	Emission (nm)	Reference
FMG[h], FDG[i]	485	530	18
4-Hydroxy-4-methylcoumarin Gal	365	450	19
7-Hydroxycoumarine-4-acetate Gal	370	455	19

C. Substrates for Luminescent Determination

Substrate	Comments	Reference
D-Luciferin-*O*-Gal	Bioluminogenic	20
AMPGD[j]	Chemiluminogenic	21
Lumi-Gal 530™	Chemiluminogenic	22

[a]The value depends on assay conditions used. See reference for further details.
[b]Isosbestic point.
[c]Calculated from the reported equation: $\Delta\epsilon_{400} = (1.637 \pm 0.007) \times 10^4/(1 + [H^+]/10^{-5.45\pm0.02})$.
[d]5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside.
[e](2-(2-(4-(β-D-Galactopyranosyloxy)-3-methoxyphenyl)-vinyl)-3-methylbenzothiazolium toluene-4-sulfonate).
[f]4-[2-(4-β-D-galactopuyranosyloxy-3-methoxyphenyl)-vinyl]-1-methylquinolinium iodide.
[g]2-[2-(4-β-D-galactopuyranosyloxy-3-methoxyphenyl)-vinyl]-1-methylquinolinium iodide.
[h]Fluorescein mono-β-D-galactoside.
[i]Fluorescein di-β-D-galactoside.
[j](3-(4-Methoxyspiro[1,2-dioxetane-3,2′-tricyclo[3.2.1.1^{3,7}[decane]-4-yl)phenyl-β-D-galactopyranoside.

Glycosyl Hydrolase Family 1. The family 1 β-galactosidases include the enzyme of an archaeon, *Sulfolobus solfataricus* (48). This enzyme is essentially a β-glycosidase that can efficiently hydrolyze β-galactosides as well as the other β-glycosides (4). The *S. solfataricus* enzyme, consisting of 489 amino acid residues (49), differs in subunit size from the family 2 of enzymes and exhibits no sequence similarity to that family. Interestingly, the lactase/phlorizin hydrolase, a mammalian β-galactosidase playing a central role in lactose digestion in the small intestine, belongs to family 1 (48). Thus, it might be possible that family 2 β-galactosidases evolved to the animal β-glucuronidases whereas family 1 enzymes evolved to the digestive lactase enzymes functioning in the mammalian intestines. A comparison of stereo- and primary structures suggested that, along with β-glucosidases, cellulases, xylanases, and glycanases, the β-galactosidases belonging to families 1 and 2 form a superfamily, the 4/7 superfamily, which is characteristic of a common eightfold β/α architecture with two conserved, catalytically important glutamates near the C-terminal ends of β-strands four and seven (50).

The *lacG* β-Galactosidases. Strains of *Arthrobacter* sp. and *Bacillus stearothermophilus* produce β-galactosidases belonging to a recently proposed β-galactosidase family, the *lacG* family, which shows no sequence similarity to enzyme families 1 and 2 (51).

Glycosyl Hydrolase Family 35. *Xanthomonas manihotis*, which is a gram-negative phytopathogenic bacterium, *Arthrobacter* sp., *Bacillus circulans*, and *Aspergillus niger* produce β-galactosidases with strong sequence similarities to animal β-galactosidases belonging to family 35 of glycosyl hydrolases (52–54). Domain 1 of the *Xanthomonas*

Table 3. Specific β-Galactosidase Inhibitors

Inhibitor	Enzyme[a]	Type[b]	K_i and other properties[c]	Reference
D-Galactose	1	C	34 mM	23
	2	C	2.32 mM	3
	3	C	350 mM	24
	4	N	40 mM	25
D-Galactal	1	C	14 μM	26
D-Galactosamine	2	C	2.7 mM	3
2-Amino-D-galactopyranose[d]	1	C	1 mM	27
Galactobiose [β-D-Gal*p*-(1,4)-D-Gal]	1	C	0.1 mM	28
β-D-Thiogalactosides	1	C	Isopropyl-, 0.085 mM	29
	2	C	Methyl-, 12 mM Isopropyl-, 15 mM	3
β-D-Galactopyranosyl trimethylammonium bromide[d]	1	C	1.4 mM	30
C-(β-D-Galactopyranosylmethyl) amine[d] and derivatives	1	C	7.8 μM	31
Furanoses[d]	1		L-Ribose, 0.03 mM D-Lyxose, 0.09 mM	23
D-Galactonolactones[d]	1		D-Galactonolactone, 0.25 mM	23
Amino sugars and amino alcohols[d]	1			27
Galactostatin and derivatives[d]	1	C	4 nM	32
Galactose-type imino sugar[d]	4		4 nM	33
1,5-Dideoxy-1,5-imino-D-galactitol[d]	1		13 μM	34
5-Amino-5-deoxy-D-galactopyranoside	1		45 nM	34
2′-Amino-2′-deoxymethyl β-lactoside	1			35
N-Bromoacetyl-β-D-galactopyranosylamine	1	IRR		36
β-D-Galactopyranosylmethyl-*p*-nitrophenyltriazene	4	A		37
(1/2,5,6)-2-(3-Azibutylthio)-5,6-epoxy-3-cyclohexen-1-*ol*	1	A	29 mM	38
Conduritol C *cis*-epoxide	1	A		39
2-Deoxy-2-fluoro-β-D-galactopyranosyl fluoride	1	A	K_i, 1.3 mM; k_i, 3.2 min^{-1}	40
	4	A	K_i, 5.4 mM; k_i, 2.5 min^{-1}	
	5	A	K_i, 1.3 mM; k_i, 0.8 min^{-1}	
2′,4′-Dinitrophenyl 2-deoxy-2-fluoro-β-D-galactopyranoside	1	A	K_i, 0.71 mM; k_i, 2.4 min^{-1}	41
Diazomethyl β-D-galactopyranosyl ketone	4	A	K_i, 30 mM; k_i, 0.56 min^{-1}	42
2-Nitro-1-(4,5-dimethoxy-2-nitrophenyl) ethyl	1	IRR	Photoreversible thiol label	43
Phytic acid	1	U	3.46 mM	44
L-Histidine				45
Calystegin B1	6	C	1.6 μM	46
Isoflavones and their glycosides	6	C/N	2′,3′,4′,7-Tetrahydroxy-isoflavone, 26 μM	47
	6	C/N	Genistein 4′,7-di-*O*-α-L-rhamnoside, 4 μM	

[a]Sources of enzymes: 1, *Escherichia coli (lacZ)*; 2, *Saccharopolyspora rectivirgula*; 3, *Streptococcus thermophilus*; 4, *Aspergillus oryzae*; 5, *A. niger*; and 6, bovine liver. For the assay conditions, see References.
[b]Type of inhibition: C, competitive; N, noncompetitive; U, uncompetitive; IRR, irreversible; A, affinity labeling reagent.
[c]K_i, Inhibition constant; k_i, inactivation rate constant.
[d]Observed strong inhibitions by these compounds argue for the occurrence of a half-chair cationic transition state during catalysis.

manihotis β-galactosidase shows similarity to regions within catalytic domains of cellulases and xylanases belonging to family 10 of glycosyl hydrolases.

It should be mentioned that the strain of *Arthrobacter* sp. can produce β-galactosidases of different enzyme families, that is, families 1 and 35 as well as the *lacG* family (51,52,55). Cumulative knowledge of amino acid sequences and stereostructures of the microbial β-galactosidases will provide important information for further insights into the classification and evolutionary and structure–function relationships of the microbial β-galactosidases.

MOLECULAR PROPERTIES AND CURRENT STATUS IN BIOPROCESS TECHNOLOGY OF SOME MICROBIAL β-GALACTOSIDASES

Although many β-galactosidase-producing microorganisms are known (see Table 1), only a limited number of microorganisms have been selected as sources of the enzymes used in industry. Most of these are of fungal origin, such as *Kluyveromyces lactis*, *K. fragilis*, *Aspergillus niger*, and *A. oryzae*. These fungi were chosen mainly because they can inexpensively produce the β-galactosidase and are generally recognized as safe (GRAS) as food additives. Recent extensive studies led to discovering many other fungal and bacterial enzymes that are applicable in bioprocess technology. In this section, we describe the molecular properties and current status in bioprocess technology of some microbial β-galactosidases (Table 5).

Fungal Enzymes

The *Kluyveromyces lactis*, *K. fragilis*, *Aspergillus niger*, and *A. oryzae* β-galactosidases, the commercially available GRAS enzymes, are currently used in industry for manufacturing milk and dairy products (see following). The *Pae-*

Table 4. List of GenBank Accession Numbers of the Microbial β-Galactosidase Sequences (May 1998) with Possible Assignment to the Glycosyl Hydrolase Families

Microorganism (enzyme)	GenBank accession number	Possible assignment[a]
Aspergillus niger (*lacA*)	A00968, L06037 S37150	35
Actinobacillus pleuropneumoniae (*lacZ*)	U62625	2
Arthrobacter sp. (*lacZ*)	U12334	2
	U78028	35
	U17417	*lacG*
Bacillus circulans (*bgaA*)	L03424	*lacG*
(*bgaB*)	L03425	*lacG* or 35
(*bgaC*)	D88750	35
B. stearothermophilus	M13466	*lacG*
Clostridium acetobutylicum (*cbgA*)	M35107	2
C. perfringens (*pbg*)	D49537	
C. (Thermoanaerobacter) thermosulfurigenes (*lacZ*)	M57579	2
Enterobacter cloacae	D42077	2
Escherichia coli (*lacZ*)	V00296	2
(*ebg*)	X52031	2
Kluyveromyces lactis (*LAC4*)	M84410	2
Lactobacillus acidophilus (*lacL*)	AB004867	2
(*lacM*)	AB004868	2
L. delbrueckii subsp. *bulgaricus*	M23530	2
Lactococcus lactis (*lacZ*)	U60828	2
Leuconostoc lactis	M92281	2
Pyrococcus woesei	AF043283	1
Rhizobium meliloti (*lacZ*)	L20757	
Saccharopolyspora rectivirgula	D86429	2
Streptococcus thermophilus (*lacZ*)	M63636	2
Sulfolobus solfataricus	M34696, X15950, X15372	1
Thermotoga maritima	U08186	2
Thermus sp.	Z93773	*lacG*
Xanthomonas manihotis (*Bga*)	L35444	35

[a]For details, see text.

cilomyces varioti β-galactosidase is a thermostable extracellular enzyme with a pH optimum at the acidic region and is potentially applicable to the processing of acid whey. The *A. oryzae, Cryptococcus laurentii*, and *Sterigmatomyces elviae* enzymes exhibit high transgalactosylation activities and are used for the production of the "galactooligosaccharides," which are derived from lactose (see following).

***Kluyveromyces lactis*.** Among the fungal enzymes listed in Table 5, biochemical and genetic studies have been the most intensive with the *K. lactis* enzyme. The *K. lactis* β-galactosidase is encoded by the *LAC4* gene, whose regulation involves the galactose/lactose induction and catabolite repression systems (64). The *K. lactis* enzyme consists of a subunit protein with an M_r of 117,618 (64) and is probably a dimer judging from its sedimentation coefficient under nondenaturing conditions (56), although higher aggregates are also known. The deduced amino acid sequence of the β-galactosidase is very similar to that of the *E. coli lacZ* enzyme; thus, the enzyme belongs to the family 2. However, Maxilact, a commercial preparation of the *K. lactis* enzyme, was shown to contain four enzyme forms of β-galactosidase, whose native molecular weights were estimated to be 19,000, 41,000, 550,000, and 630,000, respectively (5). The biochemical relevance of these observations to the reported subunit size (M_r, 117,618) and structure remains to be clarified.

***Aspergillus niger*.** *A. niger* produces three forms of β-galactosidases with M_r of 124,000, 150,000, and 173,000, respectively, which can be separated by ion exchange, gel filtration, and hydrophobic interaction chromatographies (57). These isoforms share very similar isoelectric points, amino acid compositions, and kinetic parameters. The observed multiplicity seems to vary with the culture conditions and is probably related to differences in the carbohydrate content. The amino acid sequence of the *A. niger* β-galactosidase exhibits similarity to animal β-galactosidases belonging to the glycosyl hydrolase family 35 (65) (see Table 4).

***Crypotococcus laurentii* and *Strerigmatomyces elviae*.** *C. laurentii* and *S. elviae* have been found to produce β-galactosidases that have high galactosyl transfer activities (2,62). The *C. laurentii* cells entrapped in the calcium alginate gels are currently used for industrial-scale production of 4′-galactosyllactose-containing "galactooligosaccharides" (see following).

The *C. laurentii* and *S. elviae* β-galactosidases share the following similar characteristics: *(1)* occurrence in the cell wall fractions of the yeast cells; *(2)* homodimer of a subunit protein with an M_r of 86,000–100,000; *(3)* optimum pH for hydrolysis reaction at pH 4.3–5.0; *(4)* high transgalactosylation activity that yields 4′-galactosyllactose as a main product from lactose; and *(5)* strong inhibition by Hg^{2+} but

Table 5. Molecular Properties of Some Microbial β-Galactosidases of Biotechnological Interest

Enzyme	Native M_r[a] (subunit structure)	Optimum pH[b]	Transfer activity[c]	Metal ion requirements	Localization[d]	Comments
			Fungal enzymes			
Kluyveromyces lactis	nr (probably dimeric)	$7.1^{lac,onpg}$	+	Mn^{2+}/Na^{+}	I	Subunit M_r, 117, 618; Refs. 56 and 57
K. fragilis	203,000 (nr)	6.8^{onpg}	+	Co^{2+}, Mg^{2+}, Mn^{2+}/K^{+}	I	Ref. 58
Aspergillus niger	124,000–173,000 (nr)	3.0^{lac}	+	nr	I	Three multiple forms; Ref. 57
A. oryzae	105,000 (homodimer)	4.5^{lac}	++	no	E	Refs. 59 and 60
Paecilomyces varioti	94,000 (nr)	3.5	+	no	E	Thermostable enzyme; Ref. 61
Sterigmatomyces elviae	170,000 (homodimer)	$4.5–5.0^{onpg}$	++	no	I (cell wall)	Thermostable enzyme; Ref. 2
Cryptococcus laurentii	200,000 (homodimer)	4.3^{onpg}	++	no	I (cell wall)	Ref. 62
			Bacterial enzymes			
Escherichia coli	465,000 (homotetramer)	7.0^{onpg}	+	Mg^{2+}, Mn^{2+}/Na^{+}	I	The *lacZ* enzyme
Bacillus circulans	160,000 (monomer)	6.1^{lac}	++	no	I	Isozyme 2; Ref. 63
Saccharopolyspora rectivirgula	145,000 (monomer)	7.1^{lac}	++	Mn^{2+}/Na^{+}	E	Thermostable enzyme; Ref. 3
Streptococcus thermophilus	282,000 (probably dimeric)	7.0	++	Mg^{2+}/K^{+}, Na^{+}	I	Ref. 24

Note: nr, not reported; no, no requirement.
[a]Molecular weight under nondenaturing conditions.
[b]Optimum pH for hydrolysis of lactose[lac] or 2-nitrophenyl-β-D-galactopyranoside[onpg].
[c]+, Transgalactosylation activity has been reported; ++, very high transgalactosylation activity has been reported.
[d]I, intracellular, E, extracellular.

not by *p*-chloromercuribenzoic acid. However, the *S. elviae* enzyme is distinct from the *C. laurentii* enzyme in terms of thermostability; the *S. elviae* enzyme is stable up to 80 °C for 1 h although *S. elviae* is a mesophile; in contrast, the *C. laurentii* enzyme is inactivated after incubation at 58 °C for 10 min. The *S. elviae* enzyme exhibits sevenfold higher β-D-fucosidase activity than β-D-galactosidase activity (2).

Bacterial and Actinomycetous Enzymes

Escherichia coli. The *E. coli lacZ* β-galactosidase is the best characterized β-galactosidase; its three-dimensional structure was already known (66). The regulation mechanism of the transcription of the *lacZ* β-galactosidase gene was established as the operon model. The extensive structural and mechanistic studies of the *E. coli lacZ* enzyme have provided the basis for the detailed insights into β-galactosidase catalysis. However, the usefulness of this enzyme in bioprocess technology has been examined only on a laboratory scale; the *E.coli lacZ* enzyme stops short of making a contribution in the food industry, mainly because of the unacceptability of the bacterium in uses related to food. Rather, the *E. coli lacZ* β-galactosidase and its gene are essential tools in genetic engineering, immunochemical, and molecular biology studies.

The *E. coli lacZ* β-galactosidase is a tetrameric protein consisting of identical subunits of a 1,023-amino acid polypeptide chain, which folds into five sequential domains with an extended segment at the amino terminus (66). The enzyme has four active sites per tetramer, each of which is made up of elements from two different subunits. All the catalytically important amino acid residues are located in domain 3 of the subunit. The enzyme requires the essential divalent and monovalent cations for its maximum activity, but the role of these metal activators in the enzyme mechanism is not yet known in detail. The proposed mechanism of the hydrolysis and transgalactosylation reactions catalyzed by the *E. coli lacZ* β-galactosidase are reviewed in another section. The "α-complementation" phenomenon would be worth noting as a unique characteristic of this enzyme (67,68). A fragment of amino-terminal peptide, which consists of approximately 50 amino acids encoded by a plasmid, is capable of intraallelic (α-) complementation with a defective form of the β-galactosidase encoded by the host. The α-complementation has been used for designing plasmid vectors that permit histochemical identification of recombinant clones in genetic engineering studies.

The *E. coli* cells produce a second type of β-galactosidase, the *ebg* β-galactosidase ("*ebg*" denotes "evolved β-galactosidase"), which was found as a gene product that confers a lactose-positive phenotype to $lacZ^{-}$ mutants. The enzyme is encoded by two distinct structural genes of the *ebg* operon, *ebgA* and *ebgC* genes that encode polypeptide chains of M_r 110,000 and 20,000, respectively (6). The sequence of the *ebgA* gene shows 50% nucleotide identity with that of the *lacZ* β-galactosidase gene, indicating that

the *ebg* β-galactosidase has a strong evolutionary relationship with the *lacZ* β-galactosidase (69). The active form of the *ebg* β-galactosidase is an $\alpha_4\beta_4$ heterooctamer, which is a 1:1 complex of the *ebgA* and *ebgC* gene products (6). Because the wild-type *ebg* β-galactosidase is too catalytically feeble to allow the *E. coli* cells to grow on lactose, spontaneous mutations must occur in the *ebg* β-galactosidase genes before growth can be sustained; when the *lacZ*$^-$ mutants are placed under selection pressure with lactose as a sole carbon source, strains producing mutant *ebg* enzyme with enhanced catalytic competence are selected. Thus, although the *ebg* β-galactosidase has made no significant contribution to bioprocess technology to date, the *ebg* system has been extensively studied as an interesting model for the biology of acquisitive evolution and the chemistry of the evolution of catalytic function.

Bacillus circulans. The commercial preparation of *B. circulans* β-galactosidase contains two isozymes, isozymes 1 and 2, which are reported to be monomeric with M_r of 240,000 and 160,000, respectively (63). The enzyme preparation could be applicable to enzymatic hydrolysis of lactose in skim milk (70). The isozymes are different from each other in terms of substrate specificity and the ability to catalyze transgalactosylation. Isozyme 2 has an extremely high transgalactosylation activity; the reaction of the enzyme with 4.56% (w/v) lactose yields the transgalactosylation products (galactooligosaccharides) consisting of di-, tri-, tetra-, and pentasaccharides, and the amount of the oligosaccharides reaches the maximum, which is 41% (w/w) of the total sugar in the reaction mixture. In contrast, isozyme 1 produces oligosaccharides whose amount is only 6% (w/w) of the total sugar at maximum under the same conditions. Thus, isozyme 2 of *B. circulans* β-galactosidase is used for the production of "galactooligosaccharides" as well as for the functional modification of food materials, medicines, and other bioactive compounds by transgalactosylation. Two β-galactosidase-encoding genes, *bgaA* and *bgaB*, have been identified in *B. circulans* (see Table 4) and are potential candidates that code for these isozymes. Relevance of these genes to the isozymes, however, still remains to be established. Recently, *B. circulans* has also been found to have a third β-galactosidase, *bgaC*, which is an *exo*-β-1,3-galactosidase belonging to glycosyl hydrolase family 35 (53). The commercial preparation of *B. circulans* β-galactosidase does not contain the activity of this third β-galactosidase (53). The recombinant *bgaC* β-galactosidase also showed transgalactosylation activity, which is useful for the synthesis of β-1,3-linked galactosyl oligosaccharides (71).

Streptococcus thermophilus. *S. thermophilus* is one of the dairy lactic streptococci that are used as starter cultures for a variety of industrial dairy fermentations such as those during the yogurt-making processes. The primary function of these dairy lactic streptococci is the rapid conversion of lactose into lactic acid, an essential process in dairy fermentation (72). Two distinct biochemical pathways involved in lactose utilization by the lactic streptococci have been established. One of these pathways uses the lactose permease system for the transport of unmodified lactose, which is then hydrolyzed by an intracellular β-galactosidase. The other pathway is mediated by the phosphoenolpyruvate-dependent phosphotransferase system, in which lactose is phosphorylated during translocation and the resultant lactose-6-phosphate is hydrolyzed by phospho-β-galactosidase (EC 3.2.1.85). In *S. thermophilus* cells, the lactose permease system plays a central role in lactose transport across the bacterial cell membrane, and the β-galactosidase functions as a key enzyme in the metabolism of the internalized lactose in dairy lactic fermentation (72). The *S. thermophilus* β-galactosidase is promising for industrial hydrolysis of milk lactose, judging from its thermostability and established safety as a food additive. Also, the enzyme is used as a catalyst for the industrial production of "galactooligosaccharides" from lactose because of its high transgalactosylation activity.

The *S. thermophilus* β-galactosidase consists of a subunit protein of M_r 116,860 (73). However, its quaternary structure is not yet firmly established. The enzyme exists in three distinct forms, whose apparent native molecular weights are 204,000, 186,000, and 282,000, and one of which is converted to another in a time-, temperature-, and concentration-dependent manner (24). The native molecular weight of 500,000–600,000 has also been reported (72). Structural analysis of the *S. thermophilus* β-galactosidase genes showed that the β-galactosidase is a family 2 enzyme with 48%, 35%, and 32.5% amino acid sequence identity to the *Lactobacillus bulgaricus*, *Escherichia coli*, and *Klebsiella pneumoniae* enzymes, respectively. The enzyme requires monovalent and divalent metal ions for its maximum activity, and the specificity of the metal activation varies with the substrate (24). The enzyme shows high transgalactosylation activity to produce a variety of disaccharides as main products from lactose (74), in contrast to the *Aaspergillus oryzae* β-galactosidase, for example, producing tri-, tetra-, and pentasaccharides as main transfer products from lactose.

Saccharopolyspora rectivirgula. A strain of *S. rectivirgula*, a thermophilic actinomycete, produces an extracellular thermostable β-galactosidase with high transgalactosylation activity. The *S. rectivirgula* β-galactosidase is a monomeric enzyme with a molecular weight of 145,000 and $S^\circ_{20,w}$ of 7.1*S* (3). When the enzyme reacts with 1.75 M lactose at 70 °C and pH 7.0 for 22 h, it yields oligosaccharides in the maximum yield (other than lactose) of 41% (w/w). The enzyme is stable at pH 7.2 up to 60 °C (for 4 h in the presence of 10 μM $MnCl_2$) or 70 °C (for 22 h in the presence of 1.75 M lactose and 10 μM $MnCl_2$). Thus the enzyme is applicable to an immobilized enzyme system at high temperatures (>60 °C) for efficient production of the oligosaccharides from lactose. This monomeric enzyme has eight specific binding sites for divalent metal ions. These sites are classified as follows: a very tight (class I) site for Ca^{2+}, three tight (class II) sites consisting of two Ca^{2+}-specific sites (class II_{Ca}) and one Mn^{2+}-specific site (class II_{Mn}; K_d for Mn^{2+}, 2.0 nM), and four loose (class III) sites for Mn^{2+} (K_d, 1.2 μM) and Mg^{2+} (K_d, 2 μM). The class II Mn^{2+} is catalytically important in maintaining the native structure essential for activity. Occupation of class III sites by Mg^{2+} or Mn^{2+} is of physiological importance to attain suf-

ficient thermostability by which this extracellular β-galactosidase remains active for a prolonged time at elevated temperatures, as was observed during the growth of *S. rectivirgula* (8).

MECHANISM OF THE *ESCHERICHIA COLI lacZ* β-GALACTOSIDASE-CATALYZED REACTIONS

The mechanism of the hydrolysis and transgalactosylation reactions has been the most intensively studied with the *Escherichia coli lacZ* β-galactosidase. Figure 1 shows the proposed kinetic mechanism of the *E. coli lacZ* β-galactosidase (75). In this mechanism, hydrolysis and transgalactosylation occur at the same catalytic site and share a common galactosyl enzyme intermediate. Extensive kinetic studies have also suggested the presence of two distinct subsites near the active site of the *lacZ* enzyme, the "galactose site" and "glucose site," that respectively bind galactose and glucose portions of the lactose molecule (76,77). The glucose site shows affinities to a variety of nucleophilic compounds, such as sugars, alcohols, and thiols. The presence of such a glucose site is important in the kinetic behaviors of hydrolysis and transgalactosylation of lactose.

Kinetic Mechanism of Reactions with ONPG

The kinetic mechanism of the *Escerichia coli lacZ* β-galactosidase-catalyzed reactions can be most simply shown by the reaction with 2-nitrophenyl β-D-galactopyranoside (ONPG), which is a synthetic substrate whose aglycone serves as a good leaving group (75). The ONPG (termed S in Fig. 1) binds to the enzyme to form a Michaelis complex ES (step a). The release of the aglycone, 2-nitrophenolate, occurs very rapidly (k_2 = 2,100 s^{-1} at 25 °C and pH 7.0) to yield a galactosyl enzyme intermediate (step b), which partitions into subsequent hydrolysis (c) and transgalactosylation (g, h). For the hydrolysis reaction, the water molecule reacts with the galactosyl enzyme intermediate with k_3 of 1,200 s^{-1}, yielding galactose and free enzyme (degalactosylation; step c). For the transgalactosylation, a galactosyl acceptor (termed Nu) present in the reaction mixture binds to a glucose site of the enzyme with an affinity of K_i'' to form a ternary complex (step g), followed by the formation of the galactosyl adduct (Gal-Nu; step h). In the transgalactosylation reaction, the apparent rate of the 2-nitrophenol release is increased, decreased, or unchanged with increasing acceptor concentrations, depending on the nature of the acceptor and the relative magnitude of k_2, k_3, and k_6. These observations have been kinetically analyzed to estimate the K_i'' and k_6 values of a variety of acceptors, which are the measures of how well they bind at the glucose site and how readily they react to form a galactosyl adduct, respectively. Mathematical considerations of such analyses are described by Deschavanne et al. (76) and Huber et al. (77).

Kinetic Mechanism of Reactions with Lactose

The reaction with lactose, a natural substrate, proceeds with a mechanism that is somewhat different from that of the ONPG reaction. Steps d, e, and f, which are shown with shadowing in Figure 1, must be considered instead of step b, because glucose, a product of lactose hydrolysis, has an affinity (K_i') to the glucose site, in contrast to 2-nitrophenolate, which has no affinity to the enzyme (75); released glucose can reversibly bind to form a ternary complex (step e), and then reacts with galactose to yield allolactose. During the lactose hydrolysis, allolactose is only transiently formed because it is also a substrate of the enzyme. The reported rate constants (at 25 °C, pH 7.0) are k_4 = 60 s^{-1}, k_5 = 380 s^{-1}, and k_3 = 1,200 s^{-1} (75). When high concentrations of lactose are reacted, lactose also serves as a galactosyl acceptor (Nu), and the transfer product is galactosyllactose. The galactosyllactose can, in turn, serve as a galactosyl acceptor to produce tetrasaccharides. In this manner, a mixture of oligosaccharides consisting of di-, tri-, and higher saccharides is produced from lactose.

Role of Active Site Residues in Catalysis

Extensive chemical modification, affinity labeling, and site-directed mutagenesis studies have been carried out to find functional amino acid residues at the active site of the *Escherichia coli lacZ* β-galactosidase; these have shown that Glu461, Glu537, and Tyr503 are important for en-

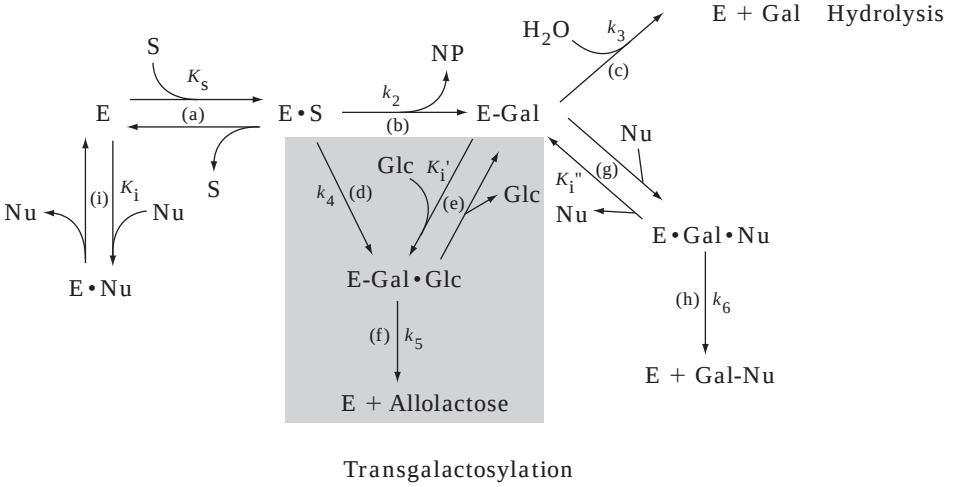


Figure 1. Proposed kinetic mechanism of hydrolysis and transgalactosylation reactions catalyzed by the *E. coli lacZ* β-galactosidase. *E*, enzyme; *S*, substrate; *NP*, 2-nitrophenol; *Gal*, galactose; *Glc*, glucose; *Nu*, galactosyl acceptor. Shadowed portion (steps *d*, *e*, *f*) indicates the pathway that must be considered, instead of step *b*, for the mechanism of the reaction with lactose. This mechanism also shows that the Nu also binds to free enzyme to form a deadend complex whose dissociation constant is *Ki* (step *i*).

zyme activity. All these catalytically important residues are located in the third domain of the enzyme subunit and, indeed, are shown to be located at the active site by X-ray crystallographic studies (66). However, some argument still exists on the details of the role(s) of each active site residue. Here we briefly describe the currently proposed roles of these amino acid residues in β-galactosidase catalysis (Fig. 2).

The glycosidic oxygen atom of the enzyme-bound substrate undergoes protonation, which facilitates aglycon departure. Intensive kinetic studies with site-directed mutants (17,78) suggest that the Glu461 serves as a Brønsted general acid–base catalyst at this step, although some evidence also argues for the role of Tyr503 being the general acid–base catalyst (79,80). The elimination of the aglycon probably produces an oxocarbonium ion intermediate that is planar at its anomeric end. This reaction is assumed because the planar and positively charged analogs of galactose serve as strong inhibitors of the *E. coli lacZ* β-galactosidase (75) (see Table 3). Affinity-labeling studies provided evidence suggesting that Glu537 is a nucleophile that attacks the anomeric carbon of the cationic intermediate (41), yielding a galactosyl enzyme in which the galactose moiety is covalently bound to the enzyme through an ester linkage. The water (or acceptor) then reacts with the covalent intermediate to produce galactose (or galactosyl adduct) and free enzyme. Although the enzyme requires monovalent and divalent ions for its activity, the role of the metal activators is not yet known. The bound Mg^{2+} (1 atom/active site; K_d, 0.65 μM [25]) has been suggested to act as an electrophilic catalyst that accelerates aglycone departure (12), whereas there is alternative evidence suggesting that it is only important in maintaining the proper active site structure (30,75). The Glu416, His418, and Glu461 are ligands to the bound Mg^{2+} ion (66).

APPLICATIONS OF β-GALACTOSIDASE-CATALYZED REACTIONS IN BIOPROCESS TECHNOLOGY

Hydrolysis

Cow milk generally contains 4.5–5.0% (w/v) lactose, whose low digestibility, low solubility, and low sweetness pose the following problems, the so-called lactose problems, in the milk and dairy industries (81).

1. Lactose intolerance is widespread in the world's adult population, with 70% being unable to metabolize large quantities of lactose in milk because of lack of intestinal lactase.
2. The high lactose content in nonfermented milk products, such as ice cream and condensed milk, often causes lactose crystallization during preservation. The crystals must not be greater than 10 μm in length, or they would make the products sandy.
3. The whey and whey permeate, which are the lactose-containing by-products from cheese manufacturing, have posed a serious problem as environmental pollutants, although these are potentially good carbohydrate sources for foods and animal feeds (82).

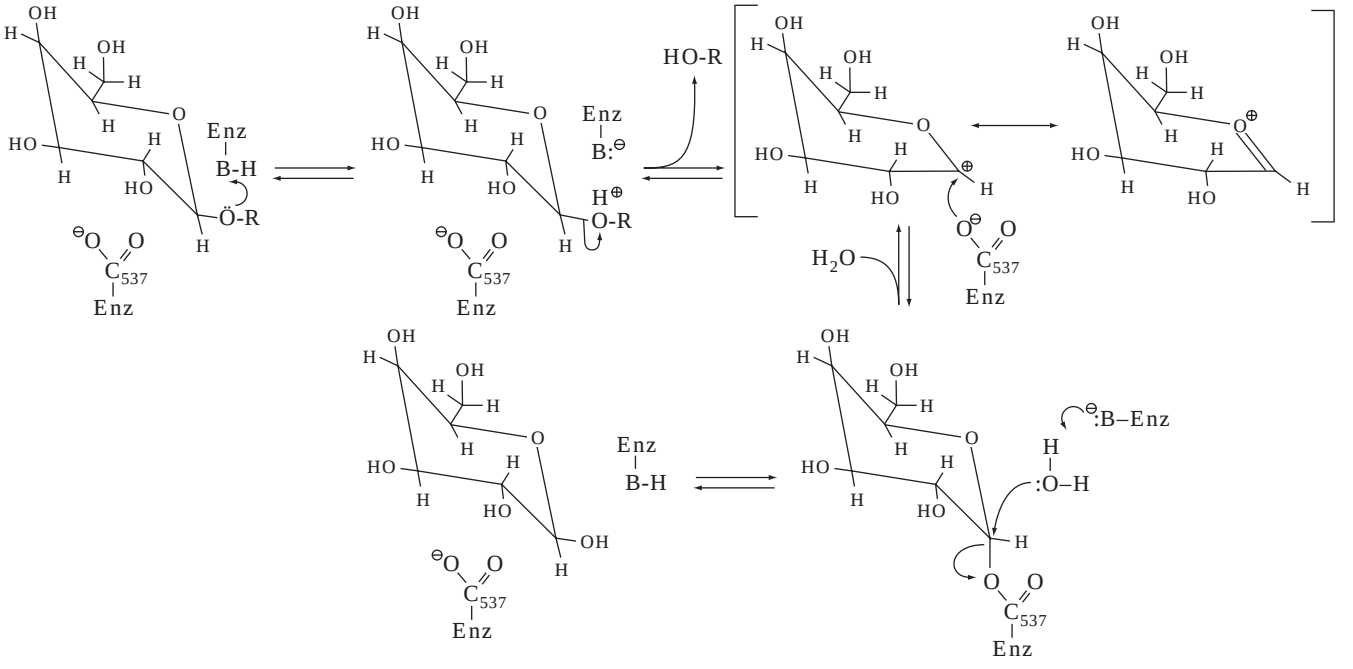


Figure 2. Catalytic mechanism of the *E. coli lacZ* β-galactosidase. Glu461 and Tyr503 are the candidates for the Brønsted general acid/base catalyst (Enz-BH/Enz-B:). A carboxyl group shown with "537" is the Glu537 that is suggested to act as a nucleophile during catalysis. The possible resonance of the planar cationic intermediate is shown with *square brackets*.

All these problems can be circumvented by decomposing lactose in the products by the action of β-galactosidase (83,84). For the purpose of alleviating lactose maldigestion, low-lactose milk is currently produced in industry by the continuous or batchwise treatment of milk with fungal β-galactosidases. Fungal β-galactosidase tablets have also been devised for this purpose and are commercially available (85,86). Enzymatic lactose hydrolysis can be considered to be a unit operation that can be integrated into the transformation of a wide range of whey and whey permeates (87). Hydrolysis of lactose produces a sugar mixture, glucose plus galactose, that is four to five times sweeter and three to four times more soluble than the parent disaccharide. The sugar mixture can serve as a condensed sugar syrup that is used as a sweetener or intermediate food product for manufacturing bakery and confectionary products, ice cream, and other specialities, permitting the efficient utilization of wheys and whey permeates (87).

Application of Immobilized Enzyme Technology in Industrial Lactose Hydrolysis

It is important to achieve industrial lactose hydrolysis with minimum cost. Immobilization of β-galactosidases is a means to reduce the cost of industrial lactose hydrolysis. Methods of immobilization and the properties of the immobilized β-galactosidases have been summarized (84). Baret (87) pointed out the following practical aspects in attaining the best performance of the immobilized enzyme system for the processing of whey and its permeates: (*1*) the nature of the substrate wheys and permeates, (*2*) the design of the reactor system, and (*3*) the operating strategy. The key objective of these considerations is to maximize the amount of hydrolyzed lactose per unit weight of catalyst, in other words, to maximize the half-life of the system. Semiindustrial-scale investigations were carried out to examine these aspects using immobilized *Aspergillus niger* and *A. oryzae* β-galactosidases that are covalently bound on a porous silica carrier:

1. Omission of demineralization of permeates had no adverse effect on the performance and stability of the systems. The need for demineralization is related to the application of the end product. However, controlling the levels of suspended solids and colloids in wheys was specifically important during operations in a fixed-bed reactor.
2. Among the fixed-bed, fluidized-bed, and stirred-tank reactors, the fixed-bed reactor gave the best kinetic results, in which the minimum reactor volume was necessary.
3. When no particular care is taken, the efficiency of the immobilized enzyme system immediately decreases because of deposition of material within the bed, the formation of a coating around the particles, microbial contamination, channeling, plugging, and other reasons. Operating strategy must therefore be considered to minimize these problems. The strategy required includes the removal of solids or colloidal substances in substrates, sanitation of substrates, introduction of the cycling of cleaning/sanitation procedures of the reactor during the operation, and programming the operating temperatures. These strategies greatly enhanced the activity and stability of the reactors.

It should be emphasized that the prevention against microbial contamination during operation is of specific importance for the processing of wheys and their permeates, which are nutritionally rich and excellent growth media for microorganisms. Methods for sanitizing the substrates, reactor, and other possible sources of microbial contamination were evaluated. Dilute acetic acid, which has been commonly used in the laboratory for disinfection of the immobilized enzyme matrices, served as only a poor sanitizer under industrial conditions. The substituted diethylenetriamines were effective for the disinfection of the immobilized enzyme matrices without causing loss of enzyme activity and have been used in large-scale operations.

The technology obtained from these large-scale examinations was transferred to fully industrial-scale operations; a plant processing 20,000 L/day of sweet whey has thus been operated, in which the hydrolysis reactor can process 1,000 L/h of nondemineralized acidified sweet whey using 40 kg of the immobilized β-galactosidase matrices (87). The processed products have been used as ingredients in the food industries.

Production of Lactose-Related Oligosaccharides by Transgalactosylation

It was shown in the 1950s that oligosaccharides are formed during the hydrolysis of lactose in milk with the fungal β-galactosidases (88,89). These oligosaccharides, the so-called galactooligosaccharides (GO), which are the products of the intermolecular transgalactosylation of lactose, have long been regarded as unwanted by-products in lactose hydrolysis (90). The *Escherichia coli lacZ* and *ebg* β-galactosidases were also shown to have transgalactosylation activity; their physiological significance has attracted the attention of biochemists because one of the transgalactosylated products, allolactose, serves as an inducer for the lactose operon of the bacterium (91).

During the past decade, GO has been shown to have some beneficial effects in maintaining human health (see later) and thus has been developed as a food material with health-giving functions. For the efficient production of GO from lactose, many attempts have been made to find microorganisms producing the GO or the β-galactosidase with high transgalactosylation activity. To date, a large number of such microorganisms have been found, among which *Bacillus circulans* (63), *Streptococcus thermophilus* (74), *Saccharopolyspora rectivirgula* (3), *Aspergillus oryzae* (92), *Cryptococcus laurentii* (62), and *Sterigmatomyces elviae* (2) produce enzymes with particularly high transgalactosylation activities. Also, the cells of yeasts, such as *Lipomyces* sp. (93), and those of the lactic acid bacteria, such as *Bifidobacterium bifidum* (94), can produce GO abundantly from lactose. Table 6 shows the identified structures of GO obtained with several enzymes and microbial cells. The structures of GO depend on the enzymes or microorganisms used as catalysts. The *B. circulans, S. rectivirgula, C.*

Table 6. Structures of Oligosaccharides Produced from Lactose by the Transgalactosylation Catalyzed by Microbial β-Galactosidases and Cells

Structure	Enzyme									Microorganism		
	1	2	3	4	5	6	7	8	9	10	11	12
β-D-Gal*p*-(1→6)-D-Glc (allolactose))	●	●			●	●	●		●	●		●
β-D-Gal*p*-(1→3)-D-Glc						●	●		●	●		●
β-D-Gal*p*-(1→2)-D-Glc						●	●		●			
β-D-Gal*p*-(1→6)-D-Gal		●			●	●	●			●		●
β-D-Gal*p*-(1→3)-D-Gal						●	●					
β-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→4)-D-Glc (6′-galactosyllactose)								●		●		
β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→4)-D-Glc (3′-galactosyllactose)								●			●	●
β-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-D-Glc (4′-galactosyllactose)			●	●					●		●	
β-D-Gal*p*-(1→6)[β-D-Gal*p*-(1→4)]-D-Glc					●			●	●			●
β-D-Gal*p*-(1→2)[β-D-Gal*p*-(1→4)]-D-Glc									●			
β-D-Gal*p*-(1→2)[β-D-Gal*p*-(1→6)]-D-Glc									●			●
β-D-Gal*p*-(1→3)[β-D-Gal*p*-(1→6)]-D-Glc									●			
β-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→6)-D-Glc									●			
β-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→6)-D-Glc					●							
β-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→3)-D-Glc									●			
β-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→2)-D-Glc									●			
β-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→6)-D-Gal					●							
β-D-Gal*p*(1→6)-β-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→4)-D-Glc					●			●				
β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→4)-D-Glc								●				
β-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→4)-D-Glc								●				
β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→4)-D-Glc												●
β-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→4)-D-Glc								●				
β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→4)-D-Glc												●
β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→4)-D-Glc												●
β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Galp-(1→3)-β-D-Gal*p*-(1→4)-D-Glc												●

Note: Closed circles indicate that the corresponding sugar has been identified in the transfer products. Enzymes are 1, *Escherichia coli* (Ref. 95); 2, *Streptococcus thermophilus* (Ref. 74); 3, *Sterigmatomyces elviae* (Ref. 2); 4, *Cryptococcus laurentii* (Ref. 62); 5, *Kluyveromyces lactis* (Ref. 96); 6, *K. fragilis* (Refs. 97,98); 7, *Aspergillus niger* (Ref. 98); 8, *A. oryzae* (Ref. 92); 9, *Bacillus circulans* (Ref. 99). Microorganisms are 10, *Penicillium chrysogenum* (Ref. 100); 11, *Chaetomium globosum* (Ref. 101); 12, *Bifidobacterium bifidum* (Ref. 94).

laurentii, S. elviae, and *A. oryzae* enzymes as well as *B. bifidum* cells produce from lactose the oligosaccharides with a general structure of Gal-$(Gal)_n$-Glc* as main products, where Gal and Glc* indicate the galactose and reducing terminal glucose residues and n the integer ($n \geq 1$), whereas the *S. thermophilus* and *E. coli lacZ* enzymes mainly produce disaccharides (Gal-Glc*) (Table 6).

Currently, GO is industrially produced by means of the *A. oryzae* and *S. thermophilus* β-galactosidases (102). In the reported system, lactose (80%, w/v) is first reacted with the *A. oryzae* enzyme at 65 °C for 4 h (103). The content of GO in the reaction mixture reaches 31% of the total sugar. The remaining lactose is further converted to the other disaccharidal transfer products by means of the *S. thermophilus* β-galactosidase (104). Industrial production of 4′-galactosyllactose is also achieved with *C. laurentii* cells entrapped in calcium alginate gels, which is used batchwise, yielding 4′-galactosyllactose whose content reaches 35% of total sugar (105). The immobilized cells can be used repeatedly, more than 30 times.

The lactose-related oligosaccharides that can be produced by the action of β-galactosidase also include the lactitol oligosaccharides, which are produced by the *A. oryzae* β-galactosidase-catalyzed transgalactosylation to lactitol [β-galactopyranosyl-(1→4)-D-glucitol], a reduced form of lactose (106).

Functional Properties of Lactose-Related Oligosaccharides

GO is of low sweetness with excellent taste quality and is stable under acidic conditions during food processing. Thus, the oligosaccharides serve as food materials potentially applicable to a wide variety of food manufacturing and cooking uses. GO also has beneficial physiological effects on human health, as described in the following subsections.

GO Is a Functional Sweetener. Ingested GO passes through the human small intestine without being digested; GO is therefore a low-calorie sweetener. *Streptococcus mutans*, a causative agent for dental caries, cannot metabolize GO or produce from GO the insoluble glucan that is responsible for dental plaque formation. Thus, GO is also a sweetener beneficial for prevention of dental caries.

GO Can Improve Intestinal and Fecal Conditions. The observation that GO can improve intestinal and fecal conditions has been explained in terms of the selective promotion by GO of bifidobacterial growth in human colons. The most dominant indigenous bacteria in the neonatal gut are the bifidobacteria, whose predominance in the gut has been suggested to cause beneficial effects in maintaining human health; these bacteria, as well as their metabolites,

provide protection from infection and facilitate the normal functions of the gut. However, the bifidobacterial population in the human intestine decreases with aging processes, and the etiological relevance of this fact to the increased incidence of intestinal diseases in the middle- and advanced-aged generations has been implicated. Many in vitro experiments showed that GO specifically promotes the growth of the intestinal bifidobacteria and lactobacilli (e.g., see Refs. 2, 99 and 107). Ito et al. (108) further examined the in vivo effects of the ingestion of transgalactosylated disaccharides (TD), which were derived from lactose by the action of *Streptococcus thermophilus* β-galactosidase, on fecal microflora and their metabolism with 12 middle-aged male volunteers. The results showed that TD ingestion resulted in a specific increase in the fecal populations of bifidobacteria and lactobacilli, as well as a decrease in those of the genera *Bacteroides* and *Candida*. It also lowered fecal pH and caused a decrease in fecal concentrations of the products of amino acid catabolism, such as isobutyric and isovaleric acids, ammonia, *p*-cresol, and indole. These results suggest that the ingested TD reached the colon and selectively promoted the growth of bifidobacteria and lactobacilli to depress the microbial decomposition of amino acids. The GO containing di-, tri-, tetra-, or higher oligosaccharides showed a similar growth promotion of the fecal bifidobacteria and affected the metabolism of fecal microflora (109). It should be noted that the lactitol oligosaccharides also showed a selective promotion of bifidobacterial growth and improved the metabolism of fecal microflora (110).

Effect on Intestinal Absorption of Ca^{2+} and Lipid Metabolism. Osteoporosis, a loss of bone mass causing an increased risk of bone fracture, is associated with the aging processes and in particular with menopause in women; estrogen deficiency resulting from menopause results in more rapid bone resorption than formation, resulting in a net loss of bone mineral density. Estrogen deficiency also causes hyperlipemia, which produces a higher risk of coronary heart disease. The effects of GO ingestion on calcium absorption, prevention of bone loss, and total serum cholesterol content were examined (111,112). Ovariectomized Wistar rats fed on a diet containing GO (5% in place of sucrose in the control diet) absorbed calcium more efficiently than those on the control diet, and the bone ash weight and tibia calcium content of the GO-ingesting rats were significantly higher than those of the controls. GO also lowered the total concentration of serum cholesterol. It is suggested that the ingested GO is metabolized in the large intestine by intestinal bacteria to enhance the production of volatile acids, such as propionate, which facilitates intestinal calcium absorption and lowers the total serum cholesterol. These observations argue for possible prevention by GO against osteoporosis and hyperlipemia caused by postmenopausal estrogen deficiency.

Functional Modification of Various Compounds by Transgalactosylation

Glycosylation is a powerful means for structural and functional modification of medicines and many other bioactive compounds; it enhances their solubilities, physicochemical stabilities, biological half-lives, membrane permeabilities, and intestinal absorption and also improves their taste qualities. In the biosynthesis of glycosides and oligosaccharides, sugar nucleotides, such as UDP-glucose, serve as substrates for enzymes that catalyze glycosidic linkage formation. However, the use of these systems in biotransformations has been limited by the restricted availability of the corresponding enzymes and substrates. An alternative is to use transglycosylation catalyzed by glycosidases; there exists a variety of glycosidases with different specificities, and their substrates are inexpensively available. Because the transfer product is always a substrate for the glycosidase, the success of the procedure as a preparative method in part depends on the rate of hydrolysis of the product being slower than that of the glycosyl donor. In practice, these conditions can be attained rather readily. Thus, there has been a steadily growing interest in the glycosidases for glycoside and oligosaccharide synthesis. In this context, β-galactosidases with high transgalactosylation activities have also been used for obtaining the galactosyl adducts of sucrose, cyclodextrins, oligosaccharides, glycosides, amino acids, peptides, alcohols, vitamins, and other bioactive compounds (Table 7). We here describe two interesting examples of synthesis of galactosides using β-galactosidases in further detail.

Synthesis of β-Galactosylthiamin. Thiamine (vitamin B_1) occurs in its phosphorylated forms in nature and plays important roles in sugar metabolism as a cofactor. 5′-*O*-(β-Galactopyranosyl)-thiamine has been produced by the *Aspergillus oryzae* β-galactosidase-catalyzed transgalactosylation using ONPG as a galactosyl donor (121). In contrast to thiamine having a specific odor and strong tongue-pricking taste, the β-galactosylthiamine obtained was odorless and mildly sweet with no such stimulative taste, suggesting its usefulness as a food additive and medicine.

One-Pot Synthesis of Sialyl Oligosaccharides Using a Multienzyme System. Sialyl oligosaccharides are important recognition elements on the surface of cells and serve as ligands of some cell adhesion molecules. Herrmann et al. described a new efficient synthesis of α(2,6)sialyl *N*-acetyllactosamine and analogs by combined use of the *Bacillus circulans* β-galactosidase and pig liver α(2,6)sialyltransferase (116). In this system, the enzymes work sequentially in one pot to form a sialyl *N*-acetyllactosamine. CMP-sialic acid, a substrate for the sialyltransferase, is generated in situ with a coupled recycling enzyme system. The β-galactosidease-catalyzed transgalactosylation proceeds irreversibly because the galactosyl adduct is irreversibly sialylated with the sialyltransferase to form a sialyl *N*-acetyllactosamine, which is no longer a substrate for the β-galactosidase.

CONCLUDING REMARKS

To explore more efficient applications of β-galactosidase in bioprocess technology, efforts are still being made to dis-

Table 7. Synthesis of Galactosyl Adducts by Means of β-Galactosidase-Catalyzed Transgalactosylation

Galactosyl acceptor	Product	Galactosyl donor[a]	Catalyst[b]	Reference
	Sugars, glycosides, and derivatives			
Sucrose	Isoraffinose	1	1	113
Branched cyclodextrins	Mono- and di-β-Gal products at their side chains	1	2, 3, 4	114
Maltopentaose	4^5-*O*-β-Gal-maltopentaose	1	2	115
GlcNAc	Sialyl oligosaccharides	1	2	116
GlcNAc and derivatives	β-Gal-(1→4 or 1→6)-β-GlcNAc-pNP	1	2	117
GlcNAc and derivatives	β-Gal-(1→3)-β-GlcNAc	1	1, 5	118
GalNAc	β-Gal(1→3)-β-GalNAc	1	1, 5	119
α-Gal-OMe	β-Gal-(1→3 or 1→6)-α-Gal-OMe	2	1	120
1-β-Ara-cytosine	3′-*O*-β-Gal- and 3′-*O*-β-Gal-(1→)-β-Gal-1-β-Ara-cytosines	1	6	121
Rubsoside	13-*O*-β-Glc-19-*O*-[β-Gal-(1→4 or 1→6)β-Glc]- and 13-*O*-[β-Gal-(1→4)-β-Glc]-19-*O*-β-Glc-steviols	1	1, 2	122
β-Glc-pNP	β-Gal-(1→3, 1→4, or 1→6)-β-Glc-pNP	1	2	122
Moranoline	β-Gal-(1→2, 1→4, or 1→6)-moranoline	1	2	123
β-GlcNAc (1-2 or 1-6)-Man	β-Gal-(1→4)-β-GlcNAc-(1→2 or 1→6)-Man	1	2, 7	124
	Amino acids and peptides			
N-Protected-L-serine methyl esters	3-*O*-β-Gal-L-Ser derivatives	1	1	125
N-Protected dipeptide esters	β-(1→3)-digalactosyl peptides	1	1	126
	Alcohols			
Alcohols	Alkyl β-Gals	1	1	127,128
2-Fluoroethanol	2-Fluoroethyl-β-Gal and derivatives	1	8	129
Allyl alcohol	Allyl β-Gal	1	9	130
	Vitamins and other bioactive compounds			
Thiamine	5′-*O*-β-Gal-thiamin	3	3	121
Kojic acid	Galactosylkojic acid	1	2	132
Oximes	β-Gal-oxime derivatives	1, 3	3	133

Note: Gal, D-Galactopyranosyl or D-galactopyranose; Glc, D-glucopyranosyl; Ara, D-arabinofuranosyl; GlcNac, *N*-acetyl-D-glucosamine; Man, D-mannopyranose; Me, methyl; Et, ethyl; pNP, *p*-nitrophenyl. (See also Table 6 for the galactooligosaccharide synthesis, in which lactose serves as both galactosyl donor and acceptor.)

[a]1, lactose; 2, PNPG; 3, ONPG.

[b]1, *Escherichia coli* (*lacZ*); 2, *Bacillus circulans*; 3, *Aspergillus oryzae*; 4, *Penicillium multicolor*; 5, bovine; 6, *Sporobolomyces singularis*; 7, *Diplococcus pneumoniae*; 8, *Kluyveromyces lactis*; 9, *Streptococcus thermophilus*.

cover a novel microbial β-galactosidase in nature. In addition, attempts are being made through genetic and protein engineering approaches to improve the specificity and enhance the stability of known β-galactosidases. These attempts, along with basic mechanistic and structural studies, will further enhance our understanding of the molecular mechanism and the structure–function relationship of β-galactosidase.

BIBLIOGRAPHY

1. E. Steers, Jr., P. Cuatrecasas, and H.B. Pollard, *J. Biol. Chem.* **246**, 196–200 (1971).
2. N. Onishi and T. Tanaka, *Appl. Environ. Microbiol.* **61**, 4026–4030 (1995).
3. M. Nakao, M. Harada, Y. Kodama, T. Nakayama, Y. Shibano, and T. Amachi, *Appl. Microbiol. Biotechnol.* **40**, 657–663 (1994).
4. D. Grogan, *Appl. Environ. Microbiol.* **57**, 1644–1649 (1991).
5. A. Mbuyi Kalala, A.G. Schnek, and J. Leonis, *Eur. J. Biochem.* **178**, 437–443 (1988).
6. A.C. Elliott, S. K, M.L. Sinnott, P.J. Smith, J. Bommuswamy, Z. Guo, B.G. Hall, and Y. Zhang, *Biochem. J.* **282**, 155–164 (1992).
7. K. Sakaguchi and T. Yamaguchi, *J. Ferment. Technol.* **51**, 750–752 (1973).
8. M. Harada, M. Inohara, M. Nakao, T. Nakayama, A. Kakudo, Y. Shibano, and T. Amachi, *J. Biol. Chem.* **269**, 22021–22026 (1994).
9. M. Leahy, S. Fanning, F. O'Gara, and D. Sheehan, *Biochem. Soc. Trans.* **19**, 19S (1991).
10. I.S. Dunn and P.A. Jennings, *Prot. Eng.* **5**, 441–446 (1992).
11. J.P. Richard, R.E. Huber, S. Lin, C. Heo, and T.L. Amyes, *Biochemistry* **35**, 12377–12386 (1996).
12. T. Selwood and M.L. Sinnott, *Biochem. J.* **268**, 317–323 (1990).
13. D.R. Hwang and M.E. Scott, *Bioorg. Chem.* **21**, 284–293 (1993).

14. B.W. Bainbridge, N. Mathias, R.G. Price, A.C. Richardson, J. Sandhu, and B.V. Smith, *FEMS Microbiol. Lett.* **80**, 319–324 (1991).
15. I. Pocsi, S.A. Taylor, A.C. Richardson, B.V. Smith, and R.G. Price, *Biochim. Biophys. Acta* **1163**, 54–60 (1993).
16. R.P. Erickson and E. Steers, Jr., *J. Bacteriol.* **102**, 79–84 (1970).
17. I.S. Dunn and P.A. Jennings, *Prot. Eng.* **4**, 205–213 (1990).
18. Z. Huang, *Biochemistry* **30**, 8530–8540 (1991).
19. N. Bagget, M.A. Case, P.R. Darby, and C.J. Gray, *Carbohydr. Res.* **197**, 295–301 (1990).
20. R. Geiger, E. Schneider, K. Wallenfels, and W. Miska, *Biol. Chem. Hoppe-Seyler* **373**, 1187–1191 (1992).
21. V.K. Jain and I.T. Magrath, *Anal. Biochem.* **199**, 119–124 (1991).
22. E.G. Beale, E.A. Deeb, R.S. Handley, H. Akhavan-Tafti, and A.P. Schaap, *BioTechniques* **12**, 320–323 (1992).
23. R.E. Huber and R.L. Brockbank, *Biochemistry* **26**, 1526–1531 (1987).
24. J. Smart and B. Richardson, *Appl. Microbiol. Biotechnol.* **26**, 177–185 (1987).
25. H. Shukla and M. Chaplin, *Enzyme Microb. Technol.* **15**, 297–299 (1993).
26. D.F. Wentworth and R. Wolfenden, *Biochemistry* **13**, 4715–4720 (1974).
27. R.E. Huber and M.T. Gaunt, *Can. J. Biochem.* **60**, 608–612 (1982).
28. H. Nakano, S. Takenishi, and Y. Watanabe, *Agric. Biol. Chem.* **51**, 2267–2269 (1987).
29. C.G. Cupples, J.H. Miller, and R.E. Huber, *J. Biol. Chem.* **265**, 5512–5518 (1990).
30. G.S. Case, M.L. Sinnott, and J.-P. Tenu, *Biochem. J.* **133**, 99–104 (1973).
31. J.N. BeMiller, R.J. Gilson, R.W. Myers, M.M. Santoro, and M.P. Yadav, *Carbohydr. Res.* **250**, 93–100 (1993).
32. Y. Miyake and M. Ebata, *Agric. Biol. Chem.* **52**, 1649–1654 (1988).
33. Y. Ichikawa and Y. Igarashi, *Tetrahedron Lett.* **36**, 4585–4586 (1995).
34. G. Legler and S. Pohl, *Carbohydr. Res.* **155**, 119–129 (1986).
35. K. Bock and K. Adelhorst, *Carbohydr. Res.* **202**, 131–149 (1990).
36. J. Yariv, K.J. Wolson, J. Hildesheim, and S. Blumberg, *FEBS Lett.* **15**, 24–26 (1971).
37. T. Mega, T. Nishijima, and T. Ikenaka, *J. Biochem.* **107**, 641–644 (1990).
38. R.E. Huber, *Carbohydr. Res.* **214**, 35–41 (1991).
39. M. Herrchen and G. Legler, *Eur. J. Biochem.* **138**, 527–531 (1984).
40. S.G. Withers, K. Rupitz, and I.P. Street, *J. Biol. Chem.* **263**, 7929–7932 (1988).
41. J.C. Gebler, R. Aebersold, and S.G. Withers, *J. Biol. Chem.* **267**, 11126–11130 (1992).
42. J.N. BeMiller, R.J. Gilson, R.W. Myers, and M.M. Santoro, *Carbohydr. Res.* **250**, 101–112 (1993).
43. R. Golan, U. Zehavi, M. Naim, A. Patchornik, and P. Smirnoff, *Biochim. Biophys. Acta* **1293**, 238–242 (1996).
44. J. Inagawa, I. Kiyosawa, and T. Nagasawa, *Agric. Biol. Chem.* **51**, 3027–3032 (1987).
45. R.A. Field, A.H. Haines, E.J. Chrystal, and M.C. Luszniak, *Biochem. J.* **274**, 885–889 (1991).
46. N. Asano, A. Kato, K. Oseki, H. Kizu, and K. Matsui, *Eur. J. Biochem.* **229**, 369–376 (1995).
47. K. Nishiyama, S. Esaki, I. Deguchi, N. Sugiyama, and S. Kamiya, *Biosci. Biotechnol. Biochem.* **57**, 107–114 (1993).
48. B. Henrissat, *Biochem. J.* **280**, 309–316 (1991).
49. M.V. Cubellis, C.C. Rozzo, P. Motecucchi, and M. Rossi, *Gene* **94**, 89–94 (1990).
50. J. Jenkins, L. Lo-Leggio, G. Harris, and R. Pickersgill, *FEBS Lett.* **362**, 281–285 (1995).
51. K.R. Gutshall, D.E. Trimbur, J.J. Kasmir, and J.E. Brenchley, *J. Bacteriol.* **177**, 1981–1988 (1995).
52. K. Gutshall, K. Wang, and J.E. Brenchley, *J. Bacteriol.* **179**, 3064–3067 (1997).
53. Y. Ito and T. Sasaki, *Biosci. Biotechnol. Biochem.* **61**, 1270–1276 (1997).
54. C.H. Taron, J.S. Benner, L.J. Hornstra, and E.P. Guthrie, *Glycobiology* **5**, 603–610 (1995).
55. D.E. Trimber, K.R. Gutshall, P. Prema, and J.E. Brenchley, *Appl. Environ. Microbiol.* **60**, 4544–4552 (1994).
56. R.C. Dickson, L.R. Dickson, and J.S. Markin, *J. Bacteriol.* **137**, 51–61 (1979).
57. F. Widmer and J.-L. Leuba, *Eur. J. Biochem.* **100**, 559–567 (1979).
58. T. Uwajima, H. Yagi, and O. Terada, *Agric. Biol. Chem.* **36**, 570–577 (1972).
59. Y. Tanaka, A. Kagamiishi, A. Kiuchi, and T. Horiuchi, *J. Biochem.* **77**, 241–247 (1975).
60. T. Mega and Y. Matsushima, *J. Biochem.* **85**, 335–341 (1979).
61. C.T. Kelly, M.R. O'Mahoney, and W.M. Fogarty, *Appl. Microbiol. Biotechnol.* **27**, 383–388 (1988).
62. K. Ohtsuka, A. Tanoh, O. Ozawa, T. Kanematsu, T. Uchida, and R. Shinke, *J. Ferment. Bioeng.* **70**, 301–307 (1990).
63. Z. Mozaffar, K. Nakanishi, R. Matsuno, and T. Kamikubo, *Agric. Biol. Chem.* **48**, 3053–3061 (1984).
64. O. Poch, H. L'Hote, V. Dallery, F. Debeaux, R. Fleer, and R. Sodoyer, *Gene* **118**, 55–63 (1992).
65. V. Kumar, S. Ramakrishnan, T. Teeri, J. Knowles, and B. Hartley, *Biotechnology* **10**, 82–85 (1992).
66. R.H. Jacobson, X.-J. Zhang, R.F. Dubose, and B.W. Matthews, *Nature* **369**, 761 (1994).
67. I. Zabin, *Mol. Cell. Biochem.* **49**, 87–96 (1982).
68. A. Ullman, *Bioessays* **14**, 201–205 (1992).
69. H.W. Stokes, P.W. Betts, and B.G. Hall, *Mol. Biol. Evol.* **2**, 469–477 (1985).
70. K. Nakanishi, R. Matsuno, K. Torii, K. Yamamoto, and T. Kamikubo, *Enzyme Microb. Technol.* **5**, 115–120 (1983).
71. H. Fujimoto, M. Miyasato, Y. Ito, T. Sasaki, and K. Ajisaka, *Glycoconj. J.* **15**, 155–160 (1998).
72. W.M. de Vos and G. Simons, *Biochimie (Paris)* **70**, 461–473 (1988).
73. C.J. Schroeder, C. Robert, G. Lenzen, L.L. McKay, and A. Mercenier, *J. Gen. Microbiol.* **137**, 369–380 (1991).
74. N.A. Greenberg and R.R. Mahoney, *Food Chem.* **10**, 195–204 (1983).
75. R.E. Huber, M.N. Gupta, and S.K. Khare, *Int. J. Biochem.* **26**, 309–318 (1994).

76. P.J. Deschavanne, O.M. Viratelle, and J.M. Yon, *J. Biol. Chem.* **253**, 833–837 (1978).
77. R.E. Huber, M.T. Gaunt, and K.L. Hurlburt, *Arch. Biochem. Biophys.* **234**, 151–160 (1984).
78. J.P. Richard, R.E. Huber, C. Heo, T.L. Amyes, and S. Lin, *Biochemistry* **35**, 12387–12401 (1996).
79. M. Ring, I.M. Armitage, and R.E. Huber, *Biochem. Biophys. Res. Commun.* **131**, 675–680 (1985).
80. M. Ring and R.E. Huber, *Arch. Biochem. Biophys.* **283**, 342–350 (1990).
81. T.P. Shukla, *CRC Crit. Rev. Food Technol.*, 325–356 (1975).
82. G. Moulin and P. Galzy, *Biotechnol. Gen. Eng. Rev.* **2**, 347–374 (1984).
83. V. Gekas and M. Lopez-Leiva, *Process Biochem.* **2**, 2–12 (1985).
84. N.A. Greenberg and R.R. Mahoney, *Process Biochem.* 2–8 (February/March 1981).
85. M.S. Medow, K.D. Thek, L.J. Newman, S. Berezin, M.S. Glassman, and S.M. Schwarz, *Am. J. Dis. Child.* **144**, 1261–1264 (1990).
86. A.E. Klingerman, *Am. J. Clin. Nutr.* **51**, 890–892 (1990).
87. J.L. Baret, *Methods Enzymol.* **136**, 411–423 (1987).
88. M. Aronson, *Arch. Biochem. Biophys.* **39**, 370–378 (1952).
89. J.H. Pazur, *Science* **117**, 355–356 (1953).
90. J.E. Prenosil, E. Stuker, and J.R. Bourne, *Biotechnol. Bioeng.* **30**, 1019–1025 (1987).
91. B.G. Hall, *J. Bacteriol.* **150**, 132–140 (1982).
92. T. Toba, A. Yokota, and S. Adachi, *Food Chem.* **16**, 147–162 (1985).
93. M. Dombou, H. Nakajima, and T. Komaki, *Tokkaihei*, 1–112978 (1989).
94. V. Dumortier, J. Montreuil, and S. Bouquelet, *Carbohydr. Res.* **201**, 115–123 (1990).
95. R.E. Huber, G. Kurz, and K. Wallenfels, *Biochemistry* **15**, 1994–2001 (1976).
96. N. Asp, A. Burvall, A. Dahlqvist, P. Hallgren, and A. Lundblad, *Food Chem.* **5**, 147–153 (1980).
97. J.H. Pazur, C.L. Tipton, T. Budovich, and J.M. Marsh, *J. Am. Chem. Soc.* **80**, 119–121 (1958).
98. T. Toba and S. Adachi, *J. Dairy Sci.* **61**, 33–38 (1978).
99. S. Yanahira, T. Kobayashi, T. Suguri, M. Nakakoshi, S. Miura, H. Ishikawa, and I. Nakajima, *Biosci. Biotechnol. Biochem.* **59**, 1021–1026 (1995).
100. A. Ballio and S. Russi, *J. Chromatogr.* **4**, 117–129 (1960).
101. P.A.J. Gorin, R.H. Haskin, and D.W.S. Westlake, *Can. J. Chem.* **44**, 2083–2086 (1966).
102. K. Matsumoto, Y. Kobayashi, S. Ueyama, T. Watanabe, R. Tanaka, T. Kan, A. Kuroda, in I. Karube ed., *Japanese Technology Reviews*, Gordon and Breach, New York, 1992, pp. 90–106.
103. K. Matsumoto, Tokkaishou, 63–109789 (1988).
104. K. Matsumoto, *Denpun Kagaku* **36**, 123–129 (1989).
105. O. Ozawa, *Food Chemicals* **6**, 75–81 (1988).
106. S. Yanahira, T. Suguri, T. Yakabe, Y. Ikeuchi, G. Hanagata, and E. Deya, *Carbohydr. Res.* **232**, 151–159 (1992).
107. M. Dombou, H. Yamamoto, H. Nakajima, K. Tomita, and T. Komaki, *Denpun Kagaku* **38**, 365–367 (1991).
108. M. Ito, M. Kimura, Y. Deguchi, A. Miyamori-Watanabe, T. Yajima, and T. Kan, *J. Nutr. Sci. Vitaminol. (Tokyo)* **39**, 279–288 (1993).
109. M. Ito, Y. Deguchi, K. Matsumoto, M. Kimura, N. Onodera, and T. Yajima, *J. Nutr. Sci. Vitaminol. (Tokyo)* **39**, 635–640 (1993).
110. S. Yanahira, M. Morita, S. Aoe, T. Suguri, I. Nakajima, and E. Deya, *J. Nutr. Sci. Vitaminol. (Tokyo)* **41**, 83–94 (1995).
111. O. Chonan and M. Watanuki, *J. Nutr. Sci. Vitaminol. (Tokyo)* **41**, 95–104 (1995).
112. O. Chonan, K. Matsumoto, and M. Watanuki, *Biosci. Biotechnol. Biochem.* **59**, 236–239 (1995).
113. K. Suyama, S. Adachi, T. Toba, T. Sohma, C.-J. Hwang, and T. Itoh, *Agric. Biol. Chem.* **50**, 2069–2075 (1986).
114. S. Kitahata, K. Fujita, Y. Takagi, K. Hara, H. Hashimoto, T. Tanimoto, and K. Koizumi, *Biosci. Biotechnol. Biochem.* **56**, 242–245 (1992).
115. S. Kitahata, K. Fujita, K. Hara, and H. Hashimoto, *Agric. Biol. Chem.* **55**, 2433–2434 (1991).
116. G.F. Herrmann, Y. Ichikawa, C. Wandrey, F.C.A. Gaeta, J.C. Paulson, and C.-H. Wong, *Tetrahedron Lett.* **34**, 3091–3094 (1993).
117. T. Usui, S. Kubota, and H. Ohi, *Carbohydr. Res.* **244**, 315–323 (1993).
118. L. Hedbys, E. Johansson, K. Mosbach, P.O. Larsson, A. Gunnarsson, S. Svensson, and H. Lonn, *Glycoconjugates* **6**, 161–168 (1989).
119. L. Hedbys, E. Johansson, K. Mosbach, and P.O. Larsson, *Carbohydr. Res.* **186**, 217–223 (1989).
120. V. Kery, S. Kucar, M. Matulova, and J. Haplova, *Carbohydr. Res.* **209**, 83–87 (1991).
121. Y. Suzuki and K. Uchida, *Biosci. Biotechnol. Biochem.* **58**, 639–643 (1994).
122. S. Kitahata, H. Ishikawa, T. Miyata, and O. Tanaka, *Agric. Biol. Chem.* **53**, 2923–2928 (1989).
123. T. Murata, S. Akimoto, M. Horimoto, and T. Usui, *Biosci. Biotechnol. Biochem.* **61**, 1118–1120 (1997).
124. M. Kojima, T. Seto, Y. Kyotani, H. Ogawa, S. Kitagawa, K. Mori, S. Maruo, T. Ohgi, and Y. Ezure, *Biosci. Biotechnol. Biochem.* **60**, 694–696 (1996).
125. H. Fujimoto, M. Isomura, T. Miyazaki, I. Matsuo, R. Walton, T. Sakakibara, and K. Ajisaka *Glycoconj. J.* **14**, 75–80 (1997).
126. D. Cantacuzene and S. Attal, *Carbohydr. Res.* **211**, 327–331 (1991).
127. S. Attal, S. Bay, and D. Cantacuzene, *Tetrahedron* **48**, 9251–9260 (1992).
128. D.H. Crout, D.A. MacManus, and P. Critchley, *J. Chem. Soc. Perkin Trans.* **1**, 1865–1868 (1990).
129. D.E. Stevenson, R.A. Stanley, and R.H. Furneaux, *Biotechnol. Bioeng.* **42**, 657–666 (1993).
130. D.E. Stevenson, A.D. Woolhouse, R.H. Furneaux, D. Batcheler, and C.T. Eason, *Carbohydr. Res.* **256**, 185–188 (1994).
131. D.E. Stevenson and R.H. Furneaux, *Carbohydr. Res.* **284**, 279–283 (1996).
132. M.A. Hassan, F. Ismail, S. Yamamoto, H. Yamada, and K. Nakanishi, *Biosci. Biotechnol. Biochem.* **59**, 543–545 (1995).
133. M. Pozo and V. Gotor, *J. Chem. Soc. Perkin Trans.* **9**, 1001–1002 (1993).

See also PROTEIN EXPRESSION, SOLUBLE.

GAS HOLD-UP

ANDREAS LÜBBERT
Martin Luther University
Wittenberg, Germany

KEY WORDS

Aeration
Bubble size
Fermenter design
Foam problems
Gas bubbles
Multiphase flow
Oxygen transfer
Specific interfacial area
Void fraction

OUTLINE

BASIC DEFINITIONS

Cultivation media in biotechnology are basically multiphase fluid systems with a continuous liquid phase in which cells and gas bubbles are dispersed. In such systems, the gas holdup ϵ is defined as the fraction of the gas phase of the total volume V_D of the dispersion considered.

$$\epsilon = \frac{V_G}{V_D} \tag{1}$$

V_G is the total gas volume within V_D (i.e., the gas holdup is a volume-related quantity). When considering the liquid phase, the gas bubbles are void areas. Thus, the gas holdup is sometimes called the void fraction.

When the entire culture within a bioreactor is considered, we then more precisely refer to the corresponding holdup as the total or global gas holdup. However, we may also speak about spatial distributions of the gas phase, particularly in systems where the gas phase is not quasi-homogeneously distributed. We then prefer to consider local gas holdups ϵ_l, which are defined by the differential quotient

$$\epsilon_l = \frac{dV_G}{dV_D} \tag{2}$$

Usually we assume that we can consider the dispersion to be a quasi-homogeneous fluid. This is a reasonable assumption when discussing technical reactor sizes, where the bubbles are usually very small as compared to the reactor dimensions. However, when we discuss effects on small scales, we should keep in mind that in gas–liquid dispersions the gas phase is composed of individual bubbles of volume V_j. It does not make sense to consider gas holdups when looking for the effects on scales in the order of the mean bubble size. On such small scales, the gas–liquid two-phase flow system must be treated in a discontinuous way. Hence, discussions about the gas holdup in bioreactors are more or less restricted to large scales as compared to the bubble size. In most practical cases we think of global gas holdups when discussing dispersions within bioreactors.

When we consider the individual bubbles of volume V_j, we may rewrite the total gas holdup as

$$\epsilon = \frac{V_G}{V_D} = \frac{\Sigma V_j}{V_D} \tag{3}$$

where the sum is drawn over all bubbles in the entire volume V_D of the dispersion. The volume V_j characterizes the bubble size. However, it is often more convenient to use diameter instead, since the diameter can most easily be estimated by visual inspection of the dispersion. Unfortunately, the bubble diameter is not well defined, since different bubbles of the same volume might have different forms. This problem has been circumvented by defining the so-called equivalent diameter d_j, which is the diameter of a sphere with the bubble's volume V_j. We can then write the gas volume V_G as

$$V_G = \Sigma V_j = \frac{\pi}{6} \Sigma d_j^3 \tag{4}$$

Bubbles have several functions in bioreactors, the most important being the transport of a gaseous reactant from the dispersed gas phase across the bubble surfaces into the liquid phase. In this respect, the transport cross section A available for this mass transfer is of primary importance. A is the total interfacial area between the gas and the liquid phase. This is, apart from the top surface of the dispersion, the sum of all bubble surfaces

$$A = \Sigma A_j \approx \pi \Sigma d_j^2 \tag{5}$$

where A_j is the surface area of the jth bubble. In biochemical reaction engineering, one prefers to discuss the transport cross section as a specific, and in this particular case a volume-related, quantity. This specific interfacial area a is defined as

$$a = \frac{A}{V_D} = \frac{\Sigma A_j}{V_D} \approx \frac{\pi \Sigma d_j^2}{V_D} \tag{6}$$

With these definitions we are able to relate the gas holdup to the quantities of primary interest in biochemical reaction engineering. Starting from equation 1, we obtain

$$\epsilon = \frac{V_G}{V_D} = \frac{AV_G}{V_D A} = \frac{A}{V_D} \frac{\frac{\pi}{6} \Sigma d_j^3}{\pi \Sigma d_j^2} = a \frac{\frac{\pi}{6} \Sigma d_j^3}{\pi \Sigma d_j^2} \tag{7}$$

Because the bubbles in biotechnologically relevant dis-

persions are not of uniform size, such expressions are of little practical use. One is more interested in characterizing the bubbles by some mean bubble diameter. In light of the preceding discussion it is convenient, as proposed by Sauter, to take the quantity

$$d_s = \frac{\Sigma d_j^3}{\Sigma d_j^2} \tag{8}$$

as such a mean bubble diameter. The sums are drawn over all bubbles j. In the literature, d_s is referred to as the Sauter diameter.

With the Sauter diameter d_s, we obtain a simple very important practical relation between the relevant gas-phase characteristics of a bubble dispersion:

$$\epsilon = a\,\frac{d_s}{6} \quad \text{or} \quad a = \frac{6\epsilon}{d_s} \tag{9}$$

which essentially states that the important specific interfacial area a is proportional to the gas holdup ϵ, the proportionality constant being dependent in a most simple way from the Sauter mean bubble diameter d_s. In other words, the specific interfacial area can increase by increasing the gas holdup, and is more effective, when the mean bubble size is kept smaller.

Assuming that the volume V_L of the liquid phase of a culture within a bioreactor is given, the higher the gas holdup ϵ, the larger the total volume V_D of the dispersion. Usually, in discussions about gas-phase behavior within bioreactors, the solid phase, in particular the cells, can be attributed to the liquid phase. In reactors where the reactor horizontal cross-sectional area A_{RCS} is uniform with respect to the height H_D of the dispersion, we can write the gas holdup as a function of the height H_D of the fully aerated liquid dispersion and the height H_L of the liquid without aeration, i.e., when there are no gas bubbles present.

$$\epsilon = \frac{V_G}{V_D} = \frac{V_D - V_L}{V_D} = \frac{(H_D - H_L)A_{RCS}}{H_D A_{RCS}} = \frac{(H_D - H_L)}{H_D} \tag{10}$$

This relative enhancement of the culture level in the reactor by aeration, or simply by the presence of gas bubbles, is of practical importance, since, as discussed later, one must monitor this level during the operation of the bioreactor. In cases where it is possible to interrupt the gas throughput through the reactor and to remove the bubbles, this equation may be used to measure the total gas holdup. If such a degassing is possible, this is the most simple measurement of the total gas holdup in dispersions.

From a global view, a higher gas holdup of the dispersion leads to a lower mean density ρ_D of the dispersion. When the gas phases can be assumed to be quasi-homogeneously distributed, the density ρ_D can be represented as a linear function of the gas holdup ϵ:

$$\rho_D = (1 - \epsilon)\rho_L \tag{11}$$

where ρ_L is the density of the liquid phase. This has an immediate consequence for the static pressure at a particular height H within the reactor. Since the static pressure drop Δp over a height interval ΔH is dependent on the density of the dispersion, it is possible to directly relate the static pressure drop and the gas holdup

$$\Delta p = \rho_D g \Delta H \tag{12}$$

$$\epsilon = 1 - \frac{\Delta p}{\rho_L g \Delta H} \tag{13}$$

which is a relationship that is often used to measure the gas holdup in multiphase reactors. Here, g is the gravitational acceleration.

The gas holdup can also be considered in terms of the aeration rate. The aeration rate is often represented by the gas throughput Q_G relative to the cross-sectional area A_{RCS} of the reactor. This quantity is referred to as the superficial gas velocity w_{sg}, since it formally has the dimension of a velocity

$$w_{sg} = \frac{Q_G}{A_{RCS}} \tag{14}$$

When the mean bubble velocity in the laboratory coordinate system is w_{bl}, then with

$$w_{bl} = \frac{H_D}{\tau_B} \tag{15}$$

where τ_B is the mean residence time of the bubbles in the reactor, and

$$Q_G = \frac{V_G}{\tau_B} \tag{16}$$

we obtain

$$\epsilon = \frac{V_G}{V_D} = \frac{V_G}{A_{RCS}H_D} = \frac{V_G\tau_B}{\tau_B A_{RCS}H_D} = \frac{w_{sg}\tau_B}{V_D} = \frac{w_{sg}}{w_{bl}} \tag{17}$$

In other words, the gas holdup is proportional to the superficial gas velocity, the proportionality constant being the reciprocal value of the mean bubble rise velocity in the laboratory coordinate system. The superficial gas velocity is the variable that can be manipulated. The effective bubble velocity w_{bl}, however, depends on the fluid-dynamic properties within the reactor, as well of the broth rheology. Equation 17 has some practical consequences, as discussed later.

BIOTECHNOLOGICAL RELEVANCE OF THE GAS HOLDUP

Of all the properties that determine the performance of a bioreactor with respect to a biochemical conversion process, it is the gas–liquid mass transfer capacity or capability that is the most important. This mass transfer capacity is primarily determined by the specific gas–liquid interfacial area a that the bioreactor is able to provide. As discussed, this quantity is primarily related to the gas

holdup. When there is no gas holdup, then there is no mass transfer.

The mass transfer rate is usually described by simple transport equation analogues to Fick's first or Ohm's law, where it is assumed that the mass current through a region that represents a resistance for the mass transfer is drawn by some driving force. In the gas–liquid mass transfer between gas bubbles and the bulk liquid, the essential mass transfer resistance is assumed to be the liquid-side boundary layer at the physical bubble surface, and the driving force is the difference between the dissolved gas concentrations at both sides of this boundary layer. In this simple picture, the mass transfer resistance is characterized by the product $k_L a$, where a is the specific interfacial area, which by equation 9 is proportional to the gas holdup. The mass transfer coefficient k_L is defined by this simple transport law. In the particular case of the oxygen transfer from a bubble into its surrounding liquid, the transport resistance is essentially the liquid-side boundary layer around the bubble. At the physical interfacial area the liquid is assumed to be saturated with oxygen (oxygen concentration O^*), while on the bulk liquid side we assume the mean dissolved oxygen concentration O present. The oxygen transfer rate OUR can then be written as

$$OUR = k_L a \Delta O \propto \epsilon \tag{18}$$

where $\Delta Q = O^* - O$ is the mean driving force.

Although practically most important, the mass transfer effects that are of interest in biotechnology are not restricted to oxygen gas–liquid mass transfer. The transfer of CO_2 is also of importance, particularly in anaerobic systems. Beer production may serve as an important example of this transfer. Here, however, the mass transfer goes in reverse (i.e., CO_2 is mainly transferred from the liquid to the gas bubbles).

The straightforward way to increase the gas holdup ε, and thus the mass transfer rate, is to increase the volume of gas sparged per time unit ($Q_G = dV_G/dt$) into the dispersion. Considering the specific interfacial area a, it would be preferable to use a sparger that produces smaller bubbles or to apply strong mechanical agitation of the dispersion to redisperse larger bubbles. By equation 9, a smaller mean bubble diameter d_s leads to a higher specific mass transfer area a. These simple arguments, however, have their clear practical limits, as the following discussion will show.

The quality of the gas holdup may be defined by the size of the bubbles, characterized by the mean (Sauter) diameter d_s and by some surface properties of the bubbles. The size of a bubble determines its rise velocity with respect to the surrounding bulk liquid. When this relative velocity decreases, the bubble more closely follows the liquid motion. Since the liquid phase is operated in a batch mode within most of the practically used bioreactors, the residence time of the bubbles within the bioreactor becomes higher with smaller bubble size. Very small bubbles may even stay within the bioreactor for a very long time before leaving the dispersion at its top surface. Additionally, the escape of the bubbles from the dispersion may be delayed or even suppressed by surface active molecules dissolved in the liquid phase, which stabilize the boundary layer. Both effects increase the gas holdup, often in an undesirable way.

Long residence times of very small bubbles are undesirable when they exceed the mean time spans during which the bubbles are able to positively contribute to the oxygen transfer. To get an idea of the time constant of the mass transfer, consider that single bubbles with an equivalent diameter of 5 mm can be assumed to have a mass transfer time constant of roughly 5 min. Consequently, when such bubbles stay within the reactor for periods much longer than 5 min, they no longer actively contribute to mass transfer. They must then be regarded as dead volume, which essentially reduces the effective reactor volume (i.e., it reduces the productivity).

When we are dealing with very long gas residence times, the dispersion accumulates a great amount of gas, and we consider foamy behavior of the dispersion. Such situations are extremely undesirable, since with continuing gas supply, the gas holdup becomes larger and larger until the headspace of the reactor becomes full and the dispersion enters the gas vent line. There it is usually stopped at the off-gas filter. This filter quickly becomes clogged by the material dissolved or dispersed in the liquid (e.g., the cells or dissolved proteins). Since the gas throughput is usually controlled, the reactor pressure will be enhanced to hold the throughput. The pressure may quickly rise, and eventually some part of the reactor, most often the tubing by which substrate or base is supplied, bursts. To prevent such accidents, in general, the gas holdup in a bioreactor must be kept under control. This is most often done by measuring the height of the dispersion with some level sensor and by reducing the gas holdup with chemical additives (antifoam agents) when the maximal level is reached.

Antifoam agents increase the coalescence behavior of the gas bubbles, thus leading to a larger bubble size. Larger bubbles are much more easily disengaging from the dispersion, so that the gas holdup drops. Soya oil or silicone-based solutions most often serve as antifoam agents. However, their addition to biological cultures must be well dosed. One argument for keeping the amount of antifoam agents within the medium low is that these molecules, as they are surface active, most often negatively influence many downstream processing procedures. Also, the addition of too much antifoam agent may drastically reduce the gas holdup in a bioreactor. This discussion shows that, with respect to mass transfer, there are competing requirements. Thus, the addition of antifoam must be optimized to obtain a maximum performance of the mass transfer in the reactor while avoiding the problems with foam.

There is another problem associated with high gas holdups or foam formation. It is well known that microbial cells are in some sense surface active as well; that is, they easily attach themselves to bubble surfaces and are then transported upward toward the surface of the dispersion. This flotation effect is a negative influence on reactor performance when a stable foam layer develops on top of the dispersion. Then, a significant amount of the biomass might become captured in the foam layer. In particular, in

the beginning of a cultivation (after inoculation) when the biomass concentration is low, the biomass can become nearly quantitatively trapped in the foam layer. Then, consequently, one does not find a significant biomass concentration in the bulk. After addition of an antifoam agent, the biomass appears back in the bulk. Hence, it is necessary to avoid these kinds of gas holdup effects by initially adding a sufficient, but not too large, amount of antifoam agent to the medium before inoculation.

While the small-scale shear flow that appears in the surroundings of the bubbles when they rise in the liquid does not have any negative effect on microbes, there are some problems with aerated cultures of shear-sensitive mammalian cells. It was observed that some cell lines die when their cultures are aerated with air bubbles, while they grow without problems when they are indirectly supplied with oxygen (e.g., by means of oxygen diffusion through silicone tubes submerged in the culture medium). The current explanation of this phenomenon is that it is not the shear flow around the rising bubbles that damages the cells but the sudden disruption of the liquid film that appears immediately before the gas bubble is disengaged into the reactor's headspace. When a cell happens to be within this bursting film, then it is subject to extreme shear rates that will most probably destroy the cell membrane. Similar effects are also to be expected during bubble coalescence or redispersion effects, since in such situations the new bubble surfaces are formed very quickly. Thus, very high local velocity gradients appear on the scale of the diameters of these cells.

Another important function of the dispersed gas phase in a bioreactor is mixing. In bubble column bioreactors, the bubbles are the only agitators. The individual bubble agitates the liquid medium locally; however, when the total gas holdup is high enough, eddies with diameters up to the reactor diameter are induced within the dispersion. With higher gas holdups, excellent mixing can be obtained, at least in dispersions with a lower liquid-phase viscosity. Mixing by means of the rising gas bubbles is not only important in bubble columns; even in industrial-scale stirred tank bioreactors, a considerable part of mixing is performed by the gas phase, and this effect is proportional to the gas holdup. This point will be discussed in more detail in the next section.

FLUID–DYNAMIC DETAILS

Practically, the gas holdup within a biotechnical cultivation broth is first of all a function of the aeration rate Q_G. Bioreactors are usually aerated with compressed air, which is dispersed into the culture medium by means of a sparger. The mass transfer via the top surface of the culture is most often comparably small and can thus be neglected. In stirred tank reactors, gas can be drawn into the dispersion by an impeller. This is the case when the impeller is operated in an unbaffled vessel where whirling occurs. However, such operational modes are generally avoided in technical bioreactors. Hence, when there is no aeration, the gas holdup is negligible.

In mechanically stirred bioreactors, which are used in most practical cases in biochemical engineering, a second influence parameter gains importance. This is the mechanical power P_D transferred into the dispersion. For this case of practical importance, we obtain

$$\epsilon \propto \left(\frac{P_D}{V_D}\right)^{0.47} w_{sg}^{0.53} \tag{19}$$

as an experimental or engineering correlation, where both (P_D/V_D) and w_{sg} are power-related quantities. The quantity w_{sg} is related to the power P_c released by the rising bubbles into the dispersion. This power was originally supplied by the air compressor against the static pressure of the liquid at the sparger location.

$$P_c = Q_G \Delta p = w_{sg} W \tag{20}$$

where Δp is the pressure difference between the surface of the dispersion and the position of the sparger and W is the weight of the dispersion above the sparger (i.e., essentially the weight of the liquid phase). The point to note is that the mechanical power pneumatically introduced into the dispersion by the gas is proportional to the superficial gas velocity w_{sg}. In bubble column bioreactors this is the only form of mechanical power input.

What is the reason for this power dependency of the gas holdup? With the mechanical power, two fluid–mechanical effects relevant to the gas holdup appear:

- The first effect is that with higher energy inputs, the velocities of the circulatory fluid flows within the reactors become more intense. Hence, a larger number of the bubbles within the flow are recirculated. Thus, the mean residence time increases and, consequently, by equation 17, so does the gas holdup. In particular, intensive eddy motions with high local flow velocities may trap some smaller bubbles within the dispersion for some time.
- The second effect supports the first one. At higher power input levels, the chaotic eddy flow motions lead to higher shear rates in the liquid phase on the scale of the bubble size. This reduces the mean bubble diameter, as characterized by the Sauter diameter, due to redispersion. Consequently, even more bubbles closely follow the liquid motions and thus the gas holdup increases.

In stirred tank reactors, one also induces a global circulation of the dispersion by means of the agitator to keep the mean gas residence time high and thus to increase the gas holdup and the specific interfacial area a or mass transfer coefficient $k_L a$.

To obtain a high enough gas holdup or specific interfacial area a in dispersions that are not foaming, one might come to the conclusion that the increase of the superficial gas velocity will help to solve the problem. It is true that with higher gas throughputs we usually increase the gas holdup; however, this has its limits, at least with respect to its quality characterized by the mean bubble diameter d_s or the specific interfacial area a.

Two competing processes come into play:

1. With higher gas throughputs we increase the energy flow into the liquid-phase motion, thus the dispersion of gas bubbles or, better, the redispersion by shear rates increases.
2. On the other hand, with higher holdups the number density of the bubbles increases and consequently the number of collisions between bubbles. This results in an increased coalescence probability and hence an increase in the mean bubble size. It is simple to understand that the coalescence probability is directly proportional to the probability that the bubbles meet each other, and this is proportional to the gas holdup. Larger bubbles most probably disengage from the dispersion. In this way the gas holdup is affected.

Consequently, the mean bubble diameter depends on the relative rates of coalescence and redispersion. Both are critically dependent on the composition of the medium, particularly, the concentrations of surface active agents. It is unlikely that this problem can be treated quantitatively in the near future, at least not for a real biochemical cultivation broth, where the composition with respect to the surface active components varies with time in a complex, virtually unpredictable way. The current knowledge is essentially of qualitative nature.

According to equation 17 we can increase the gas holdup with the superficial gas velocity, provided the residence time of the gas phase or the corresponding mean bubble velocity w_{bl} relative to the reactor coordinates remains practically constant. At higher gas throughput, the circulation flows in the reactor are intensified and, hence, so is the mean residence time of the bubbles. This leads to a lower mean bubble-rise velocity and an increased holdup. When the holdup increases too much, the bubble coalescence probability exceeds the redispersion probability. Consequently, the bubble size increases. At higher viscosity the mean bubble-rise velocity is lower, and, according to equation 17, large gas holdups and larger bubbles appear at even lower superficial velocities.

On the other hand, in dispersions with liquid phases of higher viscosity, one observes drift of the gas holdup toward higher values with a long time constant. This effect can be understood from the fact that due to the smaller slip velocity between bubble and bulk liquid, the probability of small bubbles escaping from the dispersion at its top surface becomes lower. Thus, very small bubbles accumulate in the dispersion, leading to an increased holdup. The bubble size distributions tail toward small equivalent bubble diameter values.

The effective bubble rise velocity w_{bl} does not sensitively depend on the superficial gas velocity w_{sg} for bubbles in the relevant Sauter diameter region of some millimeters. Hence, with larger superficial gas velocity w_{sg} the gas holdup will rise according to equation 17. Then, coalescence appears to a considerable extent. Big bubbles appear, and the dispersion becomes quite inhomogeneous with respect to bubble size. At very high superficial gas velocities the reactor will become flooded with gas.

At small gas holdups, the interfacial area that can be produced in a stirred tank reactor is essentially independent of the agitator type, and practically only dependent on the effectively introduced mechanical power. Higher mechanical power inputs by means of the agitator lead to a more intense redispersion of coalesced bubbles; the mean bubble size becomes smaller and the gas holdup higher. However, it is not possible to increase the gas holdup to arbitrarily high values, since agitators are restricted in the amount of gas they are able to disperse. When a critical gas throughput is exceeded, they become flooded. An essential criterion for an agitation system is the maximal gas throughput it can disperse. In other words, with respect to gas holdup and mass transfer considerations, it is straightforward to primarily distinguish different agitators by their ability to disperse more or less gas.

MEASUREMENT TECHNIQUES

To control the gas holdup, at least in the sense of keeping it below a critical value, it must be measured. Many techniques have been proposed in the literature; however, only a few of them can routinely be applied in biotechnical production practice.

The most convenient method to estimate the gas holdup in large reactors is to measure static pressure differences Δp between sites in the reactor at different heights and to relate Δp to the gas holdup ϵ with equation 12. The liquid-phase density ρ_L can be estimated reliably in simple density measurements. Some pressure sensors on the market are sensitive and also robust enough for measurements in production reactors.

The most straightforward way to proceed is to use equation 10 and register the height H_D of the surface of the dispersion. In any case, it is necessary to realize when it exceeds a limit H_{Dmax}. This can be done by various level meters or sensors. Since in most real bioprocesses the oxygen supply cannot be switched off and it is not possible to remove the bubbles, the volume of liquid phase must be determined indirectly. The most simple way is to make use of the culture weight, which is often measured even at production reactors and some estimate on the density of the nongas phase of the dispersion.

The most simple level monitor is a conductivity probe, which makes use of the fact that biotechnical culture media are electrically conductive. Hence, a single electrode placed at the critical height in the headspace of the bioreactor, connected with some counterelectrode in the dispersion (e.g., the wall of the reactor) via a voltage source, would suffice. Only when the dispersion reaches the electrode does the current loop close. The flowing electric current then indicates that the gas holdup reached its maximal value.

The same measurement principle can be used to measure local gas holdups. Then, however, the probe must be pointlike (i.e., small as compared to the bubble size). Whenever the point is within a bubble, the current flow between the electrodes is interrupted. The ratio of the sum of the time intervals Δt_j when the current is interrupted to the total measurement time t_{tm} is exactly the local gas

holdup when the total measurement time t_{tm} was not too short. All other pointlike sensors that can distinguish between gas and liquid phases can be used in the same way. One example is the total light reflection probe.

There are other probes for local gas hold measurements with larger measurement volumes, such as capacitance sensors, which measure the local gas hold-up by introducing a small capacitor into the dispersion, so that the dispersion forms the dielectric. The dielectric constant is dependent on the gas holdup within this capacitor. After some calibration, this sensor can be used to monitor the local gas holdup, provided it can be made certain that the holdup between the plates of the capacitor is representative of the holdup of its surroundings.

CONCLUSION

In bioreactors, air is dispersed primarily to provide a sufficiently large specific interfacial area a for oxygen mass transfer into the liquid phase of the culture medium. Although a, or its product with the mass transfer coefficient k_L, the k_La value, is the key quantity from the point of view of allowing a high oxygen transfer rate, the gas holdup is a very important quantity that must be kept under control in practically all bioreactors. Too little gas holdup does not allow for a sufficient mass transfer; too much gas holdup leads to problems with the reactor operation, in particular when due to foam formation. The gas holdup can be controlled by the gas throughput and, in stirred tanks, also by the stirrer speed. However, when foam is formed, it might become necessary to destroy it. Mechanical foam destroyers are applied in a few cases; however, in most bioreactors, foam is combated by chemical antifoam agents. Since the gas holdup and hence the total volume of the dispersion may become too high when foam is being formed, most reactors are equipped with a level control. This impressively underlines the importance attributed to the gas holdup in biotechnology.

See also FERMENTATION MONITORING, DESIGN AND OPTIMIZATION; FERMENTER DESIGN; MASS TRANSFER.

GENE TRANSFER, GRAM-POSITIVE BACTERIA

IRINA BAGYAN
University of Connecticut Health Center
Farmington, Connecticut

SIMON CUTTING
Royal Holloway University of London
Egham, Surrey, U.K.

KEY WORDS

Bacillus subtilis
Conjugation
DNA-mediated transformation
Plasmids
Protoplast transfection
Transduction

OUTLINE

INTRODUCTION

A variety of methods are available for gene transfer in gram-positive bacteria, and the choice of technique depends on the nature of the experiment, recipient strain, and donor DNA. For naturally transformable species such as *Bacillus*, transformation of competent cells is usually used for the introduction of chromosomal or plasmid DNA. For strains lacking satisfactory levels of natural competence methods such as polyethylene glycol–mediated protoplast transformation have been successfully used, although these are currently being replaced by the more straightforward technique of electroporation. In addition, transduction and conjugation can be used for DNA transfer in gram-positive bacteria. We outline here the major principles of each mode of gene transfer and list particular applications with different gram-positive examples.

TRANSFORMATION

Transformation is defined as the transfer of free exogenous DNA into a cell to permanently alter its heredity. Four modes of transformation exist in gram-positive bacteria: (*1*) transformation of naturally competent bacteria; (*2*) transformation of cells with artificially induced competence, (*3*) polyethylene glycol–mediated protoplast transformation, and (*4*) electrotransformation (electroporation).

Transformation of Naturally Competent Bacteria

Transformation of naturally competent bacteria involves the active uptake by a cell of free extracellular DNA and the heritable incorporation of its genetic information. Natural transformation was first detected in *Streptococcus pneumoniae*, where the ability of this bacterium to be transformed by exogenous DNA was used to prove that DNA is the genetic material (1).

Natural competence is a physiologically and genetically determined property of a particular strain. This physiolog-

ical state is dependent on the expression of specific genes whose products provide the necessary state for competence and subsequent DNA uptake. *Bacillus subtilis* is the most intensively studied naturally transformable species. Almost 40 genes (*com* genes) required for competence have now been identified in *B. subtilis*. About half of them regulate the development of the competent state, and the remainder are involved in DNA uptake and the transformation process (late *com* genes).

In *B. subtilis*, the competent state develops postexponentially during the transition from exponential growth to stationary phase. Culture density is one of the physiological conditions controlling the development of competence. As cells grow to high density, two extracellular peptide pheromones accumulate in the culture medium; they cause increased transcription of the *comS* gene. Synthesis of ComS, in turn, leads to activation of the ComK transcription factor, which directs expression of the late *com* genes (2–4). In addition to these cell density signals, there are also nutritional signals that control activation of ComK (5). Similar mechanisms exist in *Streptococcus*: in *S. pneumoniae*, competence develops in exponentially growing cells at a critical cell density (10^7–10^8 cells/mL) (6). Here, cellular secretion of a small 17-residue peptide is responsible for inducing cells to develop competence (7,8). In *B. subtilis*, only 1–10% of the cell population becomes competent whereas in *Streptococcus* spp. (e.g., *S. pneumoniae* and *S. sanguis*) up to 100% of the cells in a culture become competent (9).

DNA uptake mechanisms in *B. subtilis* and *S. pneumoniae* are also similar (for review, see Refs. 10 and 11). In brief, double-stranded DNA noncovalently binds to the surface of a competent cell (approximately 50 [*B. subtilis*] and 30–80 [*S. pneumoniae*] DNA-binding sites are used per competent cell). The transfer of DNA into the cell initiates with the generation of double-strand breaks. One strand is degraded while the other enters the cell (12); thus, DNA enters in a single-stranded form. In the case of transformation by homologous DNA, single-stranded DNA can integrate into the recipient chromosome (or resident plasmid) by homologous recombination. In *B. subtilis*, if plasmid DNA contains regions homologous to the recipient genome, the efficiency of transformation will be greater than 10^6 transformants per microgram of DNA. In the case of transformation by a nonhomologous plasmid, following entry of single-stranded plasmid DNA the plasmid must first resume a double-stranded form to replicate. If the transforming plasmid is monomeric, two complementary strands should independently enter the cell to produce a double-stranded plasmid. For this reason transformation by plasmid monomers is very inefficient in *B. subtilis*. Most plasmid preparations from gram-positive (or gram-negative) bacteria contain plasmid multimers and can yield transformants in *B. subtilis* at a level of 10^3 to 10^5 transformants per microgram of DNA. Plasmid DNA isolated from Rec^- bacteria contains a much lower amount of multimers and will lead to 100 to 1000-fold lower transformation frequencies.

To bypass the requirement for plasmid multimers in *B. subtilis*, a so-called plasmid marker rescue transformation system has been developed. This system involves the uptake of donor DNA by competent cells and the subsequent recombination with a partially homologous resident plasmid (13–16). A simple procedure can be used for the preparation of competent cells of *S. pneumoniae* (17). Methods for transformation of other naturally competent streptococci have also been developed, for example, *S. sanguis* Challis (18), *S. milleri* (19), and *S. mutans* (20).

For *B. subtilis*, the most commonly used method of preparation of competent cells remains the classical two-step procedure of Dubnau and Davidoff-Abelson (21). Laboratory techniques for *Bacillus* transformation (as well as other molecular biological methods for *Bacillus*) can be found in Harwood and Cutting's work (22). Transformation of naturally competent cells is the method of choice for introduction of DNA into *B. subtilis*. Applications include allele replacement and the introduction of chromosomal markers and plasmids. Transformation has been used for fine structure genetic mapping in *B. subtilis*. This method, however, is limited by the length of DNA taken up by a competent cell. The average size of transforming DNA (about 30 kb) limits the recombinational distance between two markers and hence the genetic distance that can be established. Nonsaturating concentrations of transforming DNA (less than 50 ng/mL in *B. subtilis*) should be used in mapping experiments to avoid congression (i.e., the transformation of the same competent cell by two genetically unlinked markers as a result of independent DNA uptake events). However, congression can be used as a strain construction tool, in that a genetic marker can be introduced into a recipient chromosome by selection for an unlinked marker (22). For this purpose, however, much higher DNA concentrations (e.g., 1–5 μg/mL in *B. subtilis*) must be used (to get 2–3% congression).

Transformation Using Artificially Induced Competence

Some bacterial species can be artificially manipulated to allow the uptake of exogenous DNA. This so-called chemically induced competence can be achieved typically by treatment of cells with calcium or some other cation and is used in the standard *Escherichia coli* transformation protocol. This is the method of choice for transformation of *E. coli*. Among gram-positive bacteria, *Staphylococcus aureus* has been transformed by this technique (23,24). In this procedure, cells collected at early exponential growth phase are washed with a NaCl solution, suspended in a $CaCl_2$-containing buffer, and incubated with DNA. The transformation efficiency has been shown to be dependent on the pH of the buffer, the presence of divalent cations, and the growth phase of the culture. Transformation efficiency is generally higher in nuclease-minus mutants (25). Also, efficient development of artificial competence requires that the cells are lysogenic for phage φ11 (24). Certain phage-expressed products are able to enhance competence; for example, both phage tails (26) and the product of early gene 31 (24) have been reported to enhance competence. Efficiencies of 10^4 transformants per microgram of DNA can be achieved with this system. The main

advantage of the artificial competence procedure for *S. aureus* is that it is effective for transformation of chromosomal DNA whereas protoplast transformation is not. This technique has been used extensively for genetic mapping where linkages representing large distances (100–150 kb) of the chromosome can be defined by the uptake of very large chromosomal fragments by the cell (27).

Protoplast Transformation

Protoplasts are generated by treatment of bacteria with muralytic enzymes (lysozyme or other, depending on the species) to degrade the peptidoglycan cell wall in the presence of an osmotic stabilizer. When exposed to high concentrations of polyethylene glycol (PEG) in the presence of DNA, protoplasts take up the DNA and if allowed to recover on suitably rich media can regenerate the cell wall. Addition of PEG can also cause protoplast fusion.

Use of protoplast techniques has advantages and disadvantages. Protoplasts allow the uptake of double-stranded plasmid DNA, enabling plasmid monomers to be used. However, protoplasts are very fragile, burst easily, and are extremely sensitive to traces of detergent on glassware. In addition, the method is generally considered laborious; protoplast regeneration can be extremely slow (2–6 days) and occurs at a low frequency. Finally, complete media are required, precluding the direct selection for prototrophic markers. Before the introduction of electroporation methods, protoplast transformation had been widely used for transformation of gram-positive species that are not naturally competent.

Protoplast transformation techniques for gram-positive bacteria were first developed for *Streptomyces* and *Bacillus*. PEG-induced transformation of *Streptomyces* protoplasts by plasmid DNA was first reported in 1978 (28) and for *B. subtilis* in 1979 (29). The *B. subtilis* protoplast transformation technique is still used for many *Bacillus* species, such as *B. anthracis* (30), *B. amyloliquefaciens* (31), *B. stearothermophilus* (32,33), and *B. licheniformis* (34).

In *Streptomyces* spp., transformation and transfection of protoplasts are important tools for gene transfer, being very efficient, even with large constructs up to at least 40 kb, and more than 10^7 transformants per microgram of DNA can be obtained, as has been described for *S. lividans* and *S. coelicolor* (35). Numerous protoplast transformation techniques have been developed for other *Streptomyces* species: *S. peucetius* (36), 10^6 transformants/μg DNA; *S. tendae* (37), 10^6 to 10^7 transformants/μg DNA; *S. clavuligerus* (38), 10^5 transformants/μg DNA; *S. venezuelae* (39), 10^7 transformants/μg DNA; *S. niveus* (40), 10^6 transformants/μg DNA; and *S. lincolnensis* (41), 10^3 transformants/μg DNA. *Streptomyces* protoplasts are more easily manipulated than those of other bacterial genera and exhibit reproducibly high levels of regeneration. Indeed, this ease of use has enabled protoplast fusion to be used not only for strain construction but also for genetic mapping in *Streptomyces* (35).

Protocols have been developed for protoplast transformation in a wide range of other gram-positive species. PEG-mediated protoplast transformation methods (giving up to 10^6 transformants/μg DNA) are available for different Actinomycetes such as *Micromonospora olivasterospora* (42), *Amycolatopsis (Nocardia) orientalis* (43), and *Rhodococcus* spp. (44). A method for PEG-mediated transformation of the mycelium has been developed for *Amycolatopsis (Nocardia) mediterranei* that yields 10^6 transformants/μg DNA without the need for protoplast isolation (45).

Staphylococcus protoplasts prepared with lysostaphin and stabilized with sucrose and Mg^{2+} are currently being used for PEG-mediated plasmid transformation (46) and protoplast fusion (47). Protoplast fusion has also been used for chromosomal mapping in *Staphylococcus aureus* (48) and also for plasmid transfer. There has also been reported a method for PEG-mediated plasmid transformation of a temperature-sensitive peptidoglycan-deficient mutant of *S. aureus*, giving transformation frequencies comparable with those obtained with protoplasts prepared with lysostaphin treatment (49). PEG-mediated protoplast transformation has been developed for different streptococci, including *S. lactis* (50), *S. faecalis* (51,52), *S. faecium* (53), *S. thermophilus* (54), and *S. pyogenes* (55).

Early techniques used protoplast transformation for introduction of plasmid DNA into mycobacteria (56). Electroporation techniques have subsequently proved to be more effective, although an efficient method for transformation of *Mycobacterium smegmatis* protoplasts has been described (57). As with mycobacteria, electroporation has now replaced protoplast transformation techniques in the lactic acid bacteria and *Listeria* spp. (58–60). PEG-mediated protoplast transformation was the first method of transformation reported for *Clostridium perfringens* (61,62). Before the development of electroporation techniques, protoplast transformation was used extensively for the introduction of DNA into *Clostridium acetobutylicum* (63,64), the most recent method reported giving up to 10^6 transformants/μg DNA (65).

Electroporation

Application of a high-voltage electric field for a very short period of time (microseconds) causes transient structural changes in lipid bilayers and biological membranes, leading to reversible pore formation (electroporation) and allowing the passive influx of DNA molecules. Electroporation as a tool for cell transformation was first used in the early 1980s for the transformation of animal cells (66). Subsequently, the technique was utilized for prokaryotes (e.g., for transformation of *Bacillus cereus* protoplasts [67]).

Transformation of intact bacterial cells by electroporation was first reported in 1987 for *Lactococcus lactis* (68) and *Lactobacillus casei* (69). A number of reports on electroporation of different gram-positive bacteria appeared shortly thereafter: *Enterococcus faecalis* (70), *Streptococcus thermophilus* (71), *Mycobacterium smegmatis*, and BCG (72), *Clostridium perfringens* (73), and a number of other gram-positive species—*Lactobacillus acidophilus*, *Bacillus cereus*, *Leuconostoc* spp., and *Listeria* spp. (74).

The first attempts to transform gram-positive bacteria using electroporation were not very efficient, and for a large number of gram-positive species the technique has

been extensively refined. The most important parameters affecting electroporation efficiency appear to be field strength and duration and shape of the pulse (75). Other parameters involve the cell culture growth conditions (e.g., media composition and phase of cell growth at the time of harvesting for electroporation) and transformation conditions (e.g., final concentration of cells and DNA in the transformation mixture, ionic strength, and the pH of electroporation buffer). Optimal conditions of course, will differ from strain to strain.

Some strains can give frequencies of 10^7 to 10^8 transformants/μg DNA, and in these strains library construction is a distinct possibility. Most strains, however, give frequencies suitable for the exchange of shuttle vectors between strains. However, frequency of electrotransformation of gram-positive bacteria is generally lower than that observed with *E. coli*. Possibly, the dense composition of the gram-positive cell wall presents a more complex barrier to entry of DNA. Several protocols were developed to overcome this problem, such as incorporation of low concentrations of the cell wall synthesis inhibitor penicillin-G in the growth medium substantially increased the efficiency of transformation of *Listeria monocytogenes* (to 10^6 transformants/μg DNA) (76). Gentle treatment of lactococcal cells with muralytic enzymes before electroporation (77) and pretreatment of *Clostridium perfringens* cells with a peptidase lysostaphin (78) also raised the efficiency of electrotransformation. In the last two cases, muralytic enzymes partially digested the cell wall, providing access of plasmid DNA to the cell membrane pores subsequently created by electroporation.

Electroporation has now become the method of choice for the transformation of a wide range of gram-positive bacteria. For example, in *Clostridium* electroporation has superseded the standard method of protoplast transformation used previously. Electroporation methods have been reported for *Clostridium perfringens* (10^7 transformants/μg DNA) (79–81), *C. beijerinckii* (82), *C. acetobutylicum* (83), and *Clostridium thermosaccharolyticum* (155). Electroporation is also a method of choice for lactic acid bacteria (*Lactobacillus, Lactococcus*) and has replaced the more time-consuming and frequently unreliable protoplast transformation techniques (84–87).

In *Mycobacterium smegmatis*, electroporation is relatively efficient, yielding little more than 10^5 transformants/μg DNA (88). In other mycobacterial species similar yields have been reported: *Mycobacterium intracellulare* (10^6 transformants/μg DNA) (89), *M. vaccae* (10^3–10^5 transformants/μg DNA) (90), and *M. bovis* BCG (91,92). It was shown that elevation of electroporation temperature to 37 °C improves the efficiency of transformation of slow-growing *Mycobacterium* spp. (93).

For *Listeria*, electroporation yields as high as 10^6 transformants/μg DNA have been reported for *L. monocytogenes, L. innocua*, and *L. ivanovii*, and somewhat lower yields for *L. seeligeri* (94).

Efficient electroporation techniques have also been developed for some actinomycetes, for example, *Actinomyces viscosus* and *Actinomyces naeslundii* (10^2–10^7 transformants/μg DNA) (95), *Brevibacterium flavum* and *Corynebacterium glutamicum* (up to 5×10^7 transformants/μg DNA) (96), *Brevibacterium methylicum* (10^5 transformants/μg DNA) (97), and *Rhodococcus fascians* (10^5 transformants/μg DNA) (98). A method for electroporation of *Streptomyces parvulus* and *S. vinaceus* has been reported (10^5 transformants/μg DNA) (99).

Electroporation has also been used for transformation in the *Streptococcus* species: *Streptococcus pyogenes* (10^7 transformants/μg DNA) (100), *S. thermophilus* (7×10^5 transformants/μg DNA) (101), *Enterococcus faecalis* (10^5 transformants/μg DNA) (102), *S. parasanguis* (103), *S. agalactiae* (10^4 transformants/μg DNA) (104), and *S. suis* type 2 (10^7 transformants/μg DNA) (105). Employed later than for other gram-positive bacteria, electroporation techniques have been developed for the staphylococci, such as *Staphylococcus aureus* (4×10^8 transformants/μg DNA) (106), *S. staphylolyticus*, and *S. carnosus* (10^5 transformants/μg DNA) (107).

Electroporation has been applied to many *Bacillus* species, such as *Bacillus thuringiensis* (108), *B. amyloliquefaciens* (109), *B. cereus* (110), *B. subtilis* (111), and *B. anthraecis* (112). However, electroporation is relatively inefficient for the *Bacillus* species and transformation frequencies are low (no more than 10^5 transformants/μg DNA) for most of the species including *B. subtilis* (111,113,114). For *B. subtilis* the maximum electrotransformation frequency reported was 9×10^4 transformants/μg DNA (115). *B. thuringiensis*, however, is an exception, with transformation frequencies up to 10^7 transformants/μg DNA being reported (116). Recently, an efficient method for electroporation of *B. brevis* cells has been developed. It gives 10^7 up to 10^8 transformants/μg DNA (117). Electroporation techniques have also been reported for transformation of *Leuconostoc paramesenteroides* (118) and species of Bifidobacterium (119). Electroporation is rapidly becoming the method of choice for the majority of gram-positive bacteria.

TRANSDUCTION

Gene transfer mediated by a bacteriophage is termed transduction; two types of transduction can be defined, generalized and specialized.

In the process of generalized transduction, any gene has a more or less equal probability of being transferred. During phage infection of the donor bacterium some phage heads incorporate a "headful" of bacterial DNA rather than phage DNA, and upon lysis the bacteria release phage particles carrying bacterial DNA. The phage particle adsorbs to the recipient bacterium and injects the bacterial DNA fragment, which can recombine into the recipient genome by homologous recombination.

During specialized transduction ("restricted transduction"), only genes adjacent to the lysogenic prophage can be transduced as a result of erroneous excision of the prophage. During excision, adjacent bacterial DNA can be excised together with the prophage (it is also possible for some prophage DNA to remain unexcised). After packaging the newly formed phage particle contains excised bacterial DNA, which can be transferred to other bacteria during the life cycle of the virus.

Generalized Transduction

Several transducing phages have been described for *B. subtilis*. The generalized transducing phage PBS1 is the most widely used for genetic analysis. The phage is pseudolysogenic, however, and although it generates turbid plaques on a plate seeded with *B. subtilis* it does not enter a true lysogenic state. PBS1 transduction is used for long-range genetic mapping because a single transducing particle encompasses more than 200 kb (about 10% of the host genome) and can be used to obtain an approximate position of a gene on the chromosome. Following this, DNA-mediated transformation of naturally competent cells is used to obtain a more accurate position of the gene of interest. For rapid mapping of genes, a kit of nine auxotrophic strains has been created (120). Together, these strains contain 22 auxotrophic markers spaced at relatively even intervals around the chromosome. A PBS1-transducing lysate is prepared from the strain containing the mutation to be mapped (which should have the wild-type allele of the auxotrophic markers carried by the kit strains) and used to infect each kit strain with selection for Aux^+ and screening for the mutant phenotype. An alternative and perhaps easier approach is to use the large collection of "silent" chromosomal insertions of Tn*917* (121). Here, a transducing lysate is made from the Tn*917* insertion strain and used to infect the strain with the unknown mutation, followed by selection for MLS resistance (encoded by the transposon) and screening for correction of the mutant phenotype. For a review of genetic mapping in *Bacillus*, see Ref. 122.

A PBS1 lysate prepared from the donor strain (donor lysate) is used to infect a motile recipient strain. PBS1 infects cells by adsorption to the flagella, so it can only infect motile cells. The procedures for preparation of phage lysates and generalized transduction can be found in Refs. 122 and 123.

Generalized transduction is a useful tool for genetic analysis of closely linked markers in *Staphylococcus aureus* (124,125). Transduction is limited by the size of the phage genome (45 kb for φ11); linkages representing large (100–150 kb) distances are being mapped by transformation of competent cells. Phage-based gene transfer systems serve as efficient tools of gene transfer in mycobacteria. Generalized transduction by the phage 13 is used for marker transfer to *Mycobacterium smegmatis* (126). For protocols, see Ref. 127.

Generalized transducing streptomycete bacteriophage FP43 has been used to transduce plasmid pRHB101 into species of *Streptomyces, Streptoverticillium*, and *Saccharapolyspora* (128–131). pRHB101 contains a 7.7-kb segment of FP43 DNA (*hft* segment) cloned into plasmid PlJ702 (130). The transducing particles contain linear concatemers of pRHB101, which form circular monomers in the transductants.

Specialized Transduction

Two specialized transducing phages have been extensively used as cloning vectors in *B. subtilis*: φ105 and SPβ (132). Both phages were designed for cloning *B. subtilis* genes by *trans* complementation (133). Cloning systems based on the phage φ105 have been extensively reviewed (133).

SPβ is a large temperate phage (126 kb) that can tolerate up to 10 kb of extra DNA. It has been used for complementation studies (123) and for transferring *lacZ* fusions between strains (134). Methods have been developed for integration of foreign DNA into the SPβ prophage by homologous recombination, after which DNA can be transferred to other strains by transduction. Procedures for preparation of SPβ lysates and SPβ transductions can be found in Ref. 123.

Phage vectors derived from the temperate phage φC31 have been developed for *Streptomyces*. Integration may involve *att* sites of the phage and chromosome, or, for the vectors in which the *att* site has been deleted, the phage can integrate by homologous recombination, either using a preexisting prophage to form a double lysogen or at the chromosomal site homologous to a DNA fragment cloned into the vector (35).

CONJUGATION

Conjugation requires cell–cell contact and provides a useful method for the interspecies and intergenera plasmid transfer within gram-positive bacteria, especially for strains lacking efficient transformation systems. Conjugative plasmids are very common among gram-positive bacteria. Usually, the conjugative plasmid itself is transferred, and in addition small plasmids can also be transferred. Some gram-positive conjugational plasmids are known to promote transfer of mobilizable plasmids across genus barriers. The conjugative streptococcal plasmid pAMβ1 is efficient in the mobilization of small plasmids in streptococci and lactococci (135). Large conjugative *Bacillus* plasmids, such as pLS20, can mobilize pUB110 for interspecies transfer (136). The majority of *Streptomyces* plasmids are self-transmissible. Conjugational crosses are useful means of strain construction in *Streptomyces* (35). Conjugative *Staphylococcus aureus* plasmids can efficiently mobilize those small plasmids that can form relaxation complexes (137). The most efficient conjugation system exists in *Enterococcus faecalis*, termed *pheromone-mediated plasmid exchange* (138,139). Here, plasmid-free enterococci excrete short oligopeptide pheromones (clump-inducing agents) that induce plasmid-containing cells to express cell surface proteins (aggregation substance) that allow the donor to bind to a receptor on the surface of the recipient and promote cell–cell contact. The plasmid is then transferred to the recipient cell and represses further pheromone synthesis. This transfer system is extremely efficient, even in matings conducted in liquid. Other conjugation systems are not so efficient and can promote DNA transfer only when the cells are forced into close contact by being placed on a filter or an agar surface.

Rare examples of conjugative mobilization of chromosomal markers exist in gram-positive bacteria. Conjugation in *Streptomyces* has been used for mapping of chromosomal genes (35,140). A conjugative transposon Tn*925* from *E. faecalis* not only can mediate its own transfer in *B. subtilis* but also can mobilize the transfer of chromoso-

mal genes (141). Until recently, it was believed that none of the gram-positive conjugation systems are capable of Hfr-like transfer of chromosomal genes. However, such a mechanism was recently proposed to occur in mycobacterium smegmatis (142).

Conjugative transposons (the elements transposing from a DNA molecule in one cell to a DNA molecule in another cell by a conjugational mechanism but that are incapable of autonomous replication) were first isolated from streptococcal strains associated with human disease. The most carefully studied are Tn*916* from *E. faecalis* (143), Tn*3701* from *Streptococcus pyogenes* (144), and Tn*1545* (145) and Tn*5253* from *S. pneumoniae* (146). Conjugative transposons are useful shuttle vectors, especially when the gene to be transferred has to be maintained at the same copy number as the chromosome. The cloned DNA can be inserted into Tn916 at a *Bst*X1 site without disrupting the function of the transposon (147). The chimeric transposon is then introduced into the gram-positive cell by electroporation or protoplast transformation (the method of transfer should not lead to the rearrangement of large incoming DNA molecules). The chimeric transposon will then insert into the host chromosome. If it subsequently proves problematic to transform DNA directly into the strain of interest, an intermediate host (e.g., *B. subtilis*) can be used. This host is transformed with DNA of the *E. coli* plasmid containing Tn*916* with the gene of interest cloned into it, and the strain will then act as a conjugative donor to any other gram-positive strain. Thus, the chimeric transposon will be transferred from the intermediate host to the strain of interest by conjugational mating on a solid surface (agar or filter). As an example of its recent application, Tn*916* was reported to be transferred to *Leuconostoc lactis* LM2301 (148), which opens the possibility of genetic studies in this *Leuconostoc* species for which no efficient conjugal or transformation system has been described previously. A series of Tn*916*-based shuttle vectors for analysis of transcriptional regulation in gram-positive bacteria has also been developed (149).

Conjugative transfer can occur even between gram-negative (*E. coli*) and gram-positive bacteria and can be exploited for transfer to gram-positive bacteria of the DNA cloned in *E. coli*. The first shuttle vector of this type (pAT187) was constructed by Trieu-Cuot et al. (150). It contains replication functions for both *E. coli* (from pBR322) and gram-positive bacteria (from pAMβ1) and an origin of transfer (*oriT*) from the gram-negative broad host range plasmid pRK2. Donor strains of *E. coli* must also harbor pRK2 or some other IncP-type plasmid or chromosomally inserted DNA fragments carrying IncP *tra* functions. Subsequently, other versatile mobilizable vectors have been developed (151–154). Such shuttle vectors can be transferred in filter matings to a wide variety of gram-positive bacteria, including *Bacillus*, *Clostridium*, *Streptomyces*, *Enterococcus*, *Streptococcus*, *Staphylococcus*, *Listeria*, *Nocardia*, and *Rhodococcus* (150,151,155,156). Improved procedures for conjugative transfer from *E. coli* to *Staphylococcus aureus* and *Listeria monocytogenes* have also been developed involving subinhibitory concentrations of cell wall synthesis inhibitors (penicillins) in the mating medium to increase transfer frequency (157). Conjugative transfer from *E. coli* to *Mycobacterium smegmatis* has also been described (158,159).

Conjugation has proved to be extremely valuable to gram-positive strains that lack a natural competence system and have proved refractory to protoplast transformation and electroporation.

BIBLIOGRAPHY

1. O.T. Avery, C.M. MacLeod, and M. McCarty, *J. Exp. Med.* **79**, 137–159 (1944).
2. R. Magnuson, J. Solomon, and A.D. Grossman, *Cell* **77**, 207–216 (1994).
3. J.M. Solomon, R. Magnuson, A. Srivastava, and A.D. Grossman, *Genes Dev.* **9**, 547–558 (1995).
4. D. van Sinderen, A. Luttinger, L. Kong, D. Dubnau, G. Venema, and L. Hamoen, *Mol. Microbiol.* **15**, 455–462 (1995).
5. D. Dubnau, *Microbiol. Rev.* **55**, 395–424 (1991).
6. A. Tomasz and R.D. Hotchkiss, *Proc. Natl. Acad. Sci. U.S.A.* **51**, 480–487 (1964).
7. A. Tomasz and J.L. Mosser, *Proc. Natl. Acad. Sci. U.S.A.* **55**, 58–66 (1966).
8. L.S. Havarstein, G. Coomaraswamy, and D.A. Morrison, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11144–11149 (1995).
9. H.O. Smith and D.B. Danner, *Annu. Rev. Biochem.* **50**, 41–68 (1981).
10. M.G. Lorenz, and W. Wackernagel, *Microbiol. Rev.* **58**, 563–602 (1994).
11. D. Dubnau, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 555–584.
12. V. Mejean and J.-P. Ceaverys, *Biol. Chem.* **268**, 5594–5599 (1993).
13. S. Contente and D. Dubnau, *Plasmid* **2**, 555–571 (1979).
14. Y. Weinrauch and D. Dubnau, *J. Bacteriol.* **154**, 1077–1087 (1983).
15. Y. Weinrauch and D. Dubnau, *J. Bacteriol.* **169**, 1205–1211 (1987).
16. P. Haima, S. Bron, and G. Venema, *Mol. Gen. Genet.* **223**, 185–191 (1990).
17. R.D. Porter and W.R. Guild, *J. Bacteriol.* **97**, 1033–1035 (1969).
18. G. Pozzi, R.A. Musmanno, P.M. Lievens, M.R. Oggioni, P. Plevani, and R. Manganelli, *Res. Microbiol.* **141**, 659–670 (1990).
19. A.E. Jacob, W.A. Horton, and D.B. Drucker, *Microbios* **60**, 167–175 (1989).
20. G.R. Shah and P.W. Caufield, *Anal. Biochem.* **214**, 343–346 (1993).
21. D. Dubnau and R. Davidoff-Abelson, *J. Mol. Biol.* **56**, 209–221 (1972).
22. C.R. Harwood and S.M. Cutting eds., *Molecular Biological Methods for Bacillus*, Wiley, New York, 1990.
23. M. Lindberg, J.-E. Sjostrom, and T. Johanson, *J. Bacteriol.* **109**, 844–847 (1972).
24. L. Rudin, J.-E. Sjostrom, M. Lindberg, and L. Philipson, *J. Bacteriol.* **118**, 155–164 (1974).
25. G.S. Omenn and J. Friedman, *J. Bacteriol.* **101**, 921–924 (1970).

26. M.P. Jackson, J. DeSena, J. Lednicky, B. Mepherson, R. Haile, R.G. Garrison, and M. Pogolsky, *Infect. Immun.* **39**, 939–947 (1983).
27. P.A. Pattee and D.S. Neveln, *J. Bacteriol.* **124**, 201–211 (1975).
28. M.J. Bibb, J.M. Ward, and D.A. Hopwood, *Nature (Lond.)* **274**, 398–400 (1978).
29. S. Chang and S.N. Cohen, *Mol. Gen. Genet.* **168**, 111–115 (1979).
30. S.-I. Mahino, C. Sasakawa, I. Uchida, N. Terakado, and M. Yoshikawa, *FEMS Microbiol. Lett.* **44**, 45–48 (1987).
31. J. Vehmaanpera, *FEMS Microbiol. Lett.* **49**, 101–105 (1988).
32. T. Imanaka, M. Fujii, I. Aramori, and S. Aiba, *J. Bacteriol.* **149**, 824–830 (1982).
33. M. Zhang, H. Nakai, and T. Imanaka, *Appl. Environ. Microbiol.* **54**, 3162–3164 (1988).
34. K.K. Jensen and M.F. Hulett, *J. Gen. Microbiol.* **135**, 2283–2287 (1989).
35. D.A. Hopwood, M.J. Bibb, K.K. Chater, T. Kieser, C.J. Bruton, H.M. Kieser, D.J. Lydiate, C.P. Smith, J.M. Ward, and H. Schrempf, *Genetic Manipulation of* Streptomyces: *A Laboratory Manual*, John Innes, Norwich, U.K. 1985.
36. J.S. Lampel and W.R. Strohl, *Appl. Environ. Microbiol.* **51**, 126–131 (1986).
37. P. Engel, *Appl. Environ. Microbiol.* **53**, 1–3 (1987).
38. M. Garcia-Dominguez, J.F. Martin, B. Mahro, A.L. Demain, and P. Liras, *Appl. Environ. Microbiol.* **53**, 1376–1381 (1987).
39. J. Anne, L. Van Mellaert, and H. Eyssen, *Appl. Microbiol. Biotechnol.* **32**, 431–435 (1990).
40. H.A. Hussain and D.A. Ritchie, *J. Appl. Bacteriol.* **71**, 422–427 (1991).
41. Z. Jandova and P. Tichy, *Folia Microbiol. (Praha)* **37**, 181–187 (1992).
42. M. Hasegawa, *Gene* **115**, 85–91 (1992).
43. P. Matsushima, M.A. McHenney, and R.H. Baltz, *J. Bacteriol.* **169**, 2298–2300 (1987).
44. M.E. Singer and W.R. Finnerty, *J. Bacteriol.* **170**, 638–645 (1988).
45. J. Madon and R. Hutter, *J. Bacteriol.* **173**, 6325–6331 (1991).
46. J. Polak and R.P. Novick, *Plasmid* **7**, 152–162 (1982).
47. P. Schaeffer, B. Cami, and R.D. Hotchkiss, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 2151–2155 (1976).
48. M.L. Stahl and P.A. Pattee, *J. Bacteriol.* **154**, 395–405 (1983).
49. D.T. Nieuwlandt and P.A. Pattee, *J. Bacteriol.* **171**, 4906–4913 (1989).
50. M.A. Sanders and M.A. Nicholson, *Appl. Environ. Microbiol.* **53**, 1730–1736 (1987).
51. M.D. Smith, *J. Bacteriol.* **162**, 92–97 (1985).
52. R. Wirth, F.Y. An, and D.B. Clewell, *J. Bacteriol.* **165**, 831–836 (1986).
53. W. Chen, K. Ohmiga, and S. Shimizu, *Appl. Environ. Microbiol.* **52**, 612–616 (1986).
54. A. Mercenier, C. Robert, D.A. Romero, I. Castellino, P. Slos, and Y. Lemoine, *Biochimie* (*Paris*) **70**, 567–577 (1988).
55. D.M. Catt and J.D. Jollick, *Microbios* **68**, 189–207 (1991).
56. W.R. Jacobs, Jr., M. Tuckman, and B.R. Bloom, *Nature (Lond.)* **327**, 532–535 (1987).
57. S.A. Naser, C.M. McCarthy, G.B. Smith, and A.K. Tupponce, *Curr. Microbiol.* **27**, 153–156 (1993).
58. B.M. Chassy, *FEMS Microbiol. Rev.* **46**, 297–312 (1987).
59. K. Watanaba, Y. Kakita, Y. Nakashima, and F. Miake, *Biosci. Biotechnol. Biochem.* **56**, 1859–1862 (1992).
60. M.F. Vicente, F. Baquero, and J.C. Perez-Diaz, *Plasmid* **18**, 89–92 (1987).
61. D.L. Heefner, C.H. Squires, R.J. Evans, B.J. Kopp, and M.J. Yarus, *J. Bacteriol.* **159**, 460–464 (1984).
62. C.H. Squires, D.L. Heefner, R.J. Evans, B.J. Kopp, and M.J. Yarus, *J. Bacteriol.* **159**, 465–471 (1984).
63. L. Podvin, G.R. Reysset, J. Hubert, and M. Sebald, in J.M. Hardie and S.P. Boriello eds., *Anaerobes Today*, Wiley, Chichester, U.K., 1988, pp. 135–140.
64. G. Reysset, J. Hubert, L. Podvin, and M. Sebald, *Biotechnol. Technol.* **2**, 199–204 (1988).
65. H. Azeddoug, J. Hubert, and G. Reysset, *J. Gen. Microbiol.* **138**, 1371–1378 (1992).
66. E. Neumann, M. Schaefer-Ridder, Y. Wang, and P.H. Hofschneider, *EMBO J.* **1**, 841–845 (1982).
67. N. Shivarova, W. Forster, H.-E. Jacob, R. Grigorova, *Z. Allg. Mikrobiol.* **23**, 595–599 (1983).
68. S.K. Harlander, in J.J. Ferretti and R. Curtiss, eds., *Streptococcal Genetics*, American Society for Microbiology, Washington, D.C., 1987, pp. 229–233.
69. B.M. Chassy and J.L. Flickinger, *FEMS Microbiol. Lett.* **44**, 173–177 (1987).
70. S. Fiedler and R. Wirth, *Anal. Biochem.* **170**, 38–44 (1988).
71. G.A. Somkuti and D.H. Steinberg, *Biochimie* (*Paris*) **70**, 579–585 (1988).
72. S.B. Snapper, L. Lugosi, A. Jekkel, R.E. Melton, T. Kieser, B.R. Bloom, and W.R. Jacobs, Jr., *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6987–6991 (1988).
73. S.P. Allen and H.P. Blaschek, *Appl. Environ. Microbiol.* **54**, 2322–2324 (1988).
74. J.B. Luchansky, P.M. Muriana, and T.R. Klaenhammer, *Mol. Microbiol.* **2**, 637–646 (1988).
75. A. Mercenier and B. Chassy, *Biochimie* (*Paris*) **70**, 503–517 (1988).
76. S.F. Park and G.S.A.B. Stewart, *Gene* **94**, 129–132 (1990).
77. I.B. Powell, M.G. Achen, A.J. Hillier, and B.E. Davidson, *Appl. Environ. Microbiol.* **54**, 655–660 (1988).
78. P.T. Scott and J.I. Rood, *Gene* **82**, 327–333 (1989).
79. S.P. Allen and H.P. Blaschek, *FEMS Microbiol. Lett.* **58**, 217–220 (1990).
80. J.T. Rood and S.T. Cole, *Microbiol. Rev.* **55**, 621–648 (1991).
81. M.K. Phillips-Jones, *FEMS Microbiol. Lett.* **54**, 221–226 (1990).
82. G.A. Birrer, W.R. Chesbro, and R.M. Zsigray, *Appl. Microbiol. Biotechnol.* **41**, 32–38 (1994).
83. S.Y. Lee, L.D. Mermelstein, G.N. Bennett, and E.T. Papoutsakis, *Ann. N.Y. Acad. Sci.* **665**, 39–51 (1992).
84. T.W. Aukrust, M.B. Brurberg, and I.F. Nes, *Methods Mol. Biol.* **47**, 201–208 (1995).
85. D.A. McIntyre and S.K. Harlander, *Appl. Environ. Microbiol.* **55**, 2621–2626 (1989).
86. D.A. McIntyre and S.K. Harlander, *Appl. Environ. Microbiol.* **55**, 604–610 (1989).
87. C.M. Rush, L.M. Hafner, and P. Timms, *J. Med. Microbiol.* **41**, 272–278 (1994).
88. S.B. Snapper, R.E. Melton, S. Mustafa, T. Kieser, and W.R. Jacobs, Jr., *Mol. Microbiol.* **4**, 1911–1919 (1990).
89. B.I. Marklund, D.P. Speert, and R.W. Stokes, *J. Bacteriol.* **177**, 6100–6105 (1995).

90. T.R. Garbe, J. Barathi, S. Barnini, Y. Zhang, C. Abou-Zeid, D. Tang, R. Mukherjee, and D.B. Joung, *Microbiology* **140**, 133–138 (1994).
91. A. Aldovini, R.H. Husson, and R.A. Young, *J. Bacteriol.* **175**, 7282–7289 (1993).
92. S.N. Cho, J.H. Hwang, S. Park, Y. Chong, S.K. Kim, C.Y. Song, and J.D. Kim, *Yonsei Med. J.* **39**, 141–147 (1998).
93. B.J. Wards and D.M. Collins, *FEMS Microbiol. Lett.* **145**, 101–105 (1996).
94. J.E. Alexander, P.W. Andrew, D. Jones, and I.S. Roberts, *Lett. Appl. Microbiol.* **10**, 179–181 (1990).
95. M.K. Yeung and C.S. Kozelsky, *J. Bacteriol.* **176**, 4173–4176 (1994).
96. A.A. Vertes, M. Inui, M. Kobayashi, Y. Kurusu, and H. Yukawa, *Res. Microbiol.* **144**, 181–185 (1993).
97. J. Nesvera, J. Hochmannova, M. Patek, A. Stroglova, and V. Becvarova, *Appl. Microbiol. Biotechnol.* **40**, 864–866 (1994).
98. J. Desomer, P. Dhaese, and M. Van Montagu, *Appl. Environ. Microbiol.* **56**, 2818–2825 (1990).
99. C. Mazy-Servais, D. Baczkowski, and J. Dusart, *FEMS Microbiol. Lett.* **151**, 135–138 (1997).
100. D. Simon and J.J. Ferretti, *FEMS Microbiol. Lett.* **66**, 219–224 (1991).
101. G.A. Somkuti, D.K. Solaiman, T.L. Johnson, and D.H. Steinberg, *Biotechnol. Appl. Biochem.* **13**, 238–245 (1991).
102. G.M. Dunny, L.N. Lee, and D.J. LeBlanc, *Appl. Environ. Microbiol.* **57**, 1194–1201 (1991).
103. J.C. Fenno, A. Shaikh, and P. Fives-Taylor, *Gene* **130**, 81–90 (1993).
104. M.L. Ricci, R. Manganelli, C. Berneri, G. Orefici, and G. Pozzi, *FEMS Microbiol. Lett.* **119**, 47–52 (1994).
105. H.E. Smith, H.J. Wisselink, U. Vecht, A.L. Gielkans, and M.A. Smits, *Microbiology* **141**, 181–188 (1995).
106. S. Schenk and R.A. Laddaga, *FEMS Microbiol. Lett.* **73**, 133–138 (1992).
107. J. Augustin and F. Gotz, *FEMS Microbiol. Lett.* **54**, 203–207 (1990).
108. L. Masson, G. Prefontaine, and R. Brousseau, *FEMS Microbiol. Lett.* **51**, 273–277 (1989).
109. J. Vehmaanpera, *FEMS Microbiol. Lett.* **52**, 165–169 (1989).
110. B.H. Belliveau and J.T. Trevors, *Appl. Environ. Microbiol.* **55**, 1649–1652 (1989).
111. P. Bridigi, E. De Rossi, M.L. Bertarini, G. Riccardi, and D. Matteuzzsi, *FEMS Microbiol. Lett.* **57**, 135–138 (1990).
112. C.P. Quinn and B.N. Dancer, *J. Gen. Microbiol.* **136**, 1211–1215 (1990).
113. H. Kusaoke, Y. Hayashi, Y. Kadowaki, and H. Kimoto, *Agric. Biol. Chem.* **53**, 2441–2446 (1989).
114. I.R. McDonald, P.W. Riley, R.J. Sharp, and A.J. McCarthy, *J. Appl. Bacteriol.* **79**, 213–218 (1995).
115. M. Ohse, K. Takahashi, Y. Kadowaki, and H. Kusaoke, *Biosci. Biotechnol. Biochem.* **59**, 1433–1437 (1995).
116. W. Schurter, M. Geiser, and D. Mathe, *Mol. Gen. Genet.* **218**, 177–181 (1989).
117. A. Okamoto, A. Kosugi, Y. Koizumi, F. Yanagida, and S. Udaka, *Biosci. Biotechnol. Biochem.* **6**, 202–203 (1997).
118. S. David, G. Simons, and W.M. De Vos, *Appl. Environ. Microbiol.* **55**, 1483–1489 (1989).
119. A. Argnani, R.J. Leer, N. van Luijk, and P.H. Pouwels, *Microbiology* **142**, 109–114 (1996).
120. R.A. Dedonder, J.A. Lepesant, J. Lepesant-Kejzlarova, A. Billault, M. Steinmetz, and F. Kunst, *Appl. Environ. Microbiol.* **33**, 989 (1977).
121. M.A. Vandeyar and S.A. Zahler, *J. Bacteriol.* **167**, 530–534 (1986).
122. S.M. Cutting and V. Azevedo, in K.W. Adolph ed., *Methods in Molecular Genetics*, Academic Press, Orlando, Fla., 1995, pp. 323–338.
123. S.M. Cutting and Vander Horn, in C.R. Harwood and S.M. Cutting eds., *Molecular Biological Methods for* Bacillus, Wiley, Chichester, U.K., 1990, pp. 27–74.
124. H.L. Ritz and J.N. Baldvin, *Proc. Soc. Exp. Biol. Med.* **110**, 667–671 (1962).
125. C.J. Schroeder and P.A. Pattee, *J. Bacteriol.* **157**, 533–537 (1984).
126. C.V. Sunder Raj and T. Ramakrishnan, *Nature (Lond.)* **228**, 280–281 (1970).
127. N.D. Connell, *Methods Cell Biol.* **45**, 107–125 (1994).
128. R.H. Baltz and M.A. McHenney, in C.L. Hershberger, S.W. Queener, and G. Hegeman eds., *Genetics and Molecular Biology of Industrial Microorganisms*, American Society for Microbiology, Washington, D.C., 1989, pp. 163–167.
129. P. Matsushima, M.A. McHenney, and R.H. Baltz, *J. Bacteriol.* **171**, 3080–3084 (1989).
130. M.A. McHenney and R.H. Baltz, *J. Bacteriol.* **170**, 2276–2282 (1988).
131. M.A. McHenney and R.H. Baltz, *J. Antibiot. (Tokyo)* **42**, 1725–1727 (1989).
132. S.A. Zahler, in D. Dubnau ed., *The Molecular Biology of the Bacilli*, Academic Press, New York, 1982, pp. 269–305.
133. J. Errington, in C.R. Harwood and S.M. Cutting eds., *Molecular Biological Methods for* Bacillus, Wiley, Chichester, U.K., 1990, pp. 175–220.
134. P. Youngman, in C.R. Harwood and S.M. Cutting eds., *Molecular Biological Methods for* Bacillus, Wiley, Chichester, U.K., 1990, pp. 221–266.
135. D. van der Lelie, H.A.B. Wosten, S. Bron, L. Oskam, and G. Venema, *J. Bacteriol.* **172**, 47–52 (1990).
136. T.M. Koehler and C.B. Thorne, *J. Bacteriol.* **169**, 5271–5278 (1987).
137. S.J. Projan, and G.L. Archer, *J. Bacteriol.* **171**, 1841–1845 (1989).
138. G.M. Dunny, B. Brown, and D.B. Clewell, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3479–3483 (1978).
139. D. Clewell, *Microbiol. Rev.* **45**, 409–436 (1981).
140. D.A. Hopwood, in J.R. Norris and D.W. Ribbons eds., *Methods in Microbiology*, vol. 7B, Academic Press, London, 1972, pp. 29–158.
141. O.R. Torres, R.Z. Korman, S.A. Zahler, and G.M. Dunny, *Mol. Gen. Genet.* **225**, 395–400 (1991).
142. L.M. Parsons, C.S. Jankowski, and K.M. Derbyshire, *Mol. Microbiol.* **28**, 571–582 (1998).
143. A.E. Franke and D.B. Clewell, *J. Bacteriol.* **145**, 494–502 (1981).
144. C. Le Bougwenec, G. De Cespedes, and T. Horaud, *J. Bacteriol.* **170**, 3930–3936 (1988).
145. P. Courvalin and C. Carlier, *Mol. Gen. Genet.* **205**, 291–297 (1986).
146. P. Ayoubi, A.O. Kilic, and M.N. Vijayakumar, *J. Bacteriol.* **173**, 1617–1622 (1991).
147. M. Norgren, M.G. Caparon, and J.R. Scott, *Infect. Immun.* **57**, 3846–3850 (1989).

148. M. Zuniga, I. Pardo, and S. Ferrer, *FEMS Microbiol. Lett.* **135**, 179–185 (1996).
149. R.I. Geist, N. Okada, and M.G. Caparon, *J. Bacteriol.* **175**, 7561–7570 (1993).
150. P. Trieu-Cuot, C. Carlier, P. Martin, and P. Courvalin, *FEMS Microbiol. Lett.* **48**, 289–294 (1987).
151. D.R. Williams, D.I. Young, and M. Young, *J. Gen. Microbiol.* **136**, 819–826 (1990).
152. M. Bierman, R. Logan, K. O'Brien, E.T. Seno, R.N. Rao, and B.E. Schoner, *Gene* **116**, 43–49 (1992).
153. T. Smokvina, P. Mazodier, F. Boccard, C.J. Thompson, and M. Guerineau, *Gene* **94**, 53–59 (1990).
154. A. Schafer, A. Tauch, W. Jager, J. Kalinowski, G. Thierbach, and A. Puhler, *Gene* **145**, 69–73 (1994).
155. P. Mazodier, R. Petter, and C. Thompson, *J. Bacteriol.* **171**, 3583–3585 (1989).
156. T. Voeykova, L. Emelyanova, V. Tabakov, and N. Mkrtumyan, *FEMS Microbiol. Lett.* **162**, 47–52 (1998).
157. P. Trieu-Cuot, E. Derlot, and P. Courvalin, *FEMS Microbiol. Lett.* **109**, 19–23 (1993).
158. R. Lazrag, S. Clavel-Seres, H.L. David, and D. Roulland-Dussoix, *FEMS Microbiol. Lett.* **69**, 135–138 (1990).
159. E.P. Gormley and J. Davies, *J. Bacteriol.* **173**, 6705–6708 (1991).

See also ENERGY METABOLISM, MICROBIAL AND ANIMAL CELLS; ENZYMES, DIRECTED EVOLUTION; METABOLITES, PRIMARY AND SECONDARY; MUTAGENESIS; PROTEIN EXPRESSION, SOLUBLE.

GENETIC INSTABILITY

JOHN CULLUM
MATTHIAS REDENBACH
LB Genetik, University of Kaiserslautern
Kaiserslautern, Germany
DASLAV HRANUELI
PLIVA d.d., Research Institute
Zagreb, Croatia

KEY WORDS

Mutation
Scale-up
Selection
Strain degeneration

OUTLINE

INTRODUCTION

The description *genetic instability* is used in slightly different ways by different authors. Some would consider it synonymous with a high mutation rate. This can affect all genes in the organism or only particular genes and may occur only in particular growth conditions. Ideally the mutation rate should be measured as the number of new mutants arising per genome per generation. In practice it is often difficult or impossible to measure the mutation rate per generation. For example the presence of multiple genomes in fragments of a mycelial organism often prevents quantitation. In many cases the proportion of mutants in a population is measured, and this quantity depends on both the mutation rate and the selective forces acting on the population. For most organisms the frequency of spontaneous mutants for a particular gene in the population is under 10^{-4}. An operational definition of genetic instability is that the frequency of mutants is considerably higher than this. In *Streptomyces*, genes are considered unstable if spontaneous mutants are present at 0.1% per spore or higher (1). Defining genetic instability in terms of the number of mutants in the population is useful for industrial microorganisms because the mutations that can cause problems during fermentation are those that accumulate to significant levels in the population. Genetic instability occurs in brewer's yeast (*Saccharomyces cerevisiae*). In this case, *petite* mutants, which are defective in mitochondrial function, occur at high frequency (about 1% of colonies). However, the mutants grow much slower than wild type and do not cause any problems in fermentations. Such genetic instability is not very important for industrial applications. In contrast, many industrial strains undergo a rather ill-defined process of "strain degeneration" in which undesirable properties such as reduced production level accumulate. This is normally countered by careful strain maintenance and testing. Strain degeneration likely reflects the fact that desirable properties for production (e.g., high levels of secondary metabolites) probably have a negative influence on the ability of the strain to compete, so mutants outgrow it. This will happen even at low mutation rates because of the presence of strong selective forces.

Little literature is available about genetic instability in industrial strains. This reflects that such data can be valuable proprietary knowledge, and the fact that genetic instability is often ignored as a phenomenon during fermentation development. One important industrial organism in which some studies have been carried out is *Streptomyces rimosus*, the oxytetracycline producer (2). Some high producers are genetically unstable, and the instability can be recognized by the frequent appearance of morphological mutants. Two classes of mutants with undesirable properties have been recognized: those in class I retain resis-

tance to the antibiotic but have much lower production, whereas those in class II have also lost resistance. Because high levels of oxytetracycline are produced during industrial fermentation, class II mutants are irrelevant to the production process; they are killed by the antibiotic. However, class I mutants survive, and competition experiments show that they outgrow the parent production strain. If care is taken with inoculum preparation it is possible to reduce problems from genetic instability to a manageable level. Experience shows that the way in which the strain is grown can have large effects on genetic instability. Thus, if the strain is propagated via spores rather than as mycelial fragments, then a much higher frequency of mutants occurs. This is shown in Figure 1 in which plates are shown from cultures propagated as mycelial fragments and compared with those from spores. The dark brown colonies are the high producers; it can be seen that a variety of lighter variants occur—most of these are class I variants. Another important factor is the medium. More instability is seen on malt agar than on soya-mannitol agar. A third factor is growth temperature: higher temperatures (e.g., 37 °C) lead to much higher levels of genetic instability, compared with lower temperatures (e.g., 28 °C). Thus, careful choice of growth conditions can be used to control instability.

MUTATION AND SELECTION

The proportion of mutants in the population is governed both by the frequency of generation of mutants and by selection effects acting on the mutants. In general in batch processes both of these parameters will vary throughout the process, making simple models such as have been used for chemostat populations invalid. If the population is small compared with the mutation rate, then the frequency of mutations will be governed by the time needed to generate the first mutant and there will be a lot of variation between parallel cultures. In most industrial fermentations the cultures should be large enough to make stochastic effects unimportant. However, large fluctuations in yield in small cultures (e.g., at the shake flask stage) might point to genetic instability being present. Experimental studies with chemostats show that mutants usually displace the initial population within 100 generations. The biological basis for this behavior is that strains are not well adapted to growth in the chemostat, so very strong selection pressures result. This tends to limit the usefulness of continuous culture for many industrial applications despite the many desirable features for plant utilization. Unless the selected properties in the continuous culture conditions are the same as those desired, genetic instability will cause problems. One area in which this may be true is in biomass production. In some cases continuous culture can be used with recycling of biomass. In this case there is little growth, so the number of generations remains relatively small.

Usually during scale-up of batch processes one hopes to alter the physiological conditions between the large and small fermenters as little as possible. Thus, scale-up can be considered, at a first approximation, as just growing the culture for more generations. For example, scaling up from 1 L to 200,000 L corresponds to about 18 more generations. This increase can lead to genetic instability becoming a problem. An example of genetic instability occurring during scale-up was provided by the bicyclomycin fermentation using *Streptomyces sapporonensis* (3). There was a drastic reduction in yield compared with shake flasks on scaling up to 4,000 L in a pilot plant, and this was shown to be due to the accumulation of low-producing mutants. The mutants could be recognized on the basis of colony morphology because they were deficient in aerial mycelium production. There is very little information in the literature on the frequency of genetic instability causing problems during scale-up. This is partly because it is not gen-

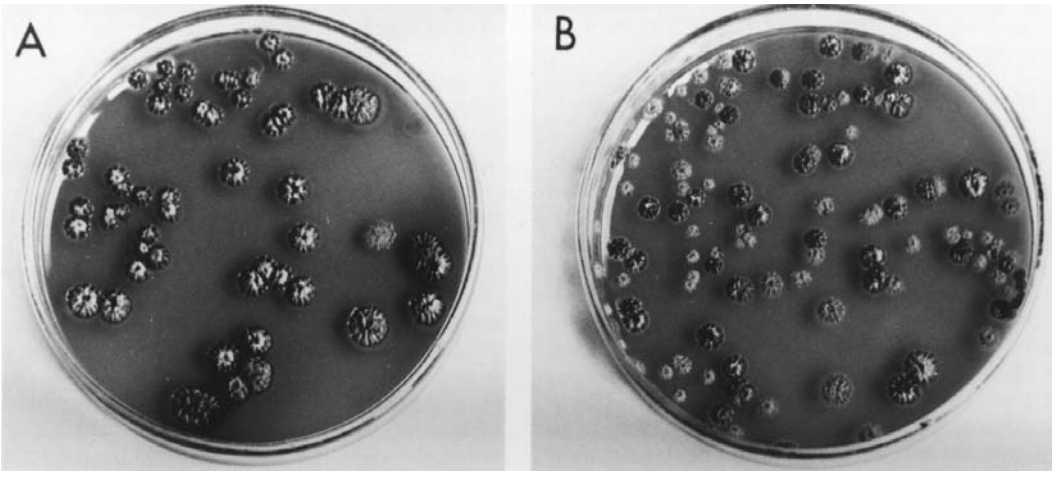


Figure 1. Colonies of an *S. rimosus* strain selected for high oxytetracycline production. (**A**) was propagated via spores on agar whereas (**B**) was propagated as mycelial fragments in liquid culture. The pigmentation of the colonies correlates with their oxytetracycline production. *Source*: Figure modified from Ref. 2 by kind permission of the American Society for Microbiology.

erally recognized as a problem and it is unusual to do genetic studies on the population of cells in a fermenter. We suspect that genetic instability is not uncommon in industrial fermentations, but that the empirical optimization of yield will result in growth conditions that reduce the instability to acceptable levels. However, this would have the consequence that the conditions will probably be suboptimal for product yield by the nonmutated cells. This suggests that it would be valuable to monitor the proportion of mutants so as to provide an extra parameter for the fermentation development. As most large scale fermentations are carried out in a multistage process with the inoculum being introduced to a series of three or more fermenters, on increasing scale it would be possible, for instance, to try to optimize the genetic stability in the smaller fermenters and then optimize the yield in the final production fermenter. This sort of strategy depends on methods for detecting mutants, which will be discussed later.

A genetic approach to increased stability can also be taken. This will to some extent occur automatically in a strain selection program because obviously unstable strains will be rejected. A second approach that has been used in some cases is to carry out genetic crosses between unstable high producers and stable low producers to obtain high producers with improved stability. In principle it should be possible to approach the problem in a directed rational way. If the mechanism for generation of the mutants is known, it ought to be possible to obtain derivatives with a lower mutation rate. However, this will usually involve some disruption of the DNA processing in the cell and is likely to lead to lower viability and production. A second approach is to introduce a system that selects against mutants. Thus, if mutation is caused by deletions it ought to be possible to move a selectable gene into the deletogenic region and thus ensure that the mutants will have a growth disadvantage (4). In many cases the strategies to combat instability involve manipulation of the selection effects rather than the mutation rate.

The use of immobilized cells or biomass recycling should greatly reduce the potential for genetic instability because little growth will occur. However, even in these cases there may be high mutagenesis rates in a subpopulation of stationary-phase cells (see later) and differential growth of a subpopulation. This has been demonstrated in colonies on agar, but the role of genetic instability in industrial applications using nongrowing cells is not yet known.

BACTERIA

Many different types of mutations can occur, and the different classes often involve different molecular mechanisms. Mutations can be classified into point mutations or DNA rearrangements. Point mutations were originally defined as mutations that map as a single point in genetic crosses. In molecular terms they are usually taken as mutations that affect a single base or a small number of bases. In contrast, DNA rearrangements affect a larger number of bases. One type of point mutation is a base change in which a single base in the DNA is changed to another base. In a *transition*, a purine is substituted for another purine (i.e., A → G or G → A), whereas a *transversion* substitutes a purine by a pyrimidine (i.e., A → C, A → T, G → C, G → G). The other important type of point mutation is the addition or deletion of a small number of bases in the DNA. These are often referred to as frameshift mutations because the addition or deletion of one or two bases in a protein coding area causes a shift of the reading frame such that all distal codons will be read wrongly. Important DNA rearrangements in bacteria are deletions (loss of a DNA segment), insertions (insertion of a segment of DNA into another sequence), and inversions (a segment of DNA is inverted with respect to the neighboring sequences). The effect of any mutation depends on its location. In bacteria the genes are usually closely packed with little space in between. This means that most point mutations will occur in genes or their control regions and therefore have a genetic effect. If deletions or insertions affect genes, they will usually result in inactivation. In some cases insertions can carry promoters and activate neighboring genes after insertion. Inversions will have a clear cut phenotypic effect only if the inversion end points lie in genes or operons. However, as the transcription direction of genes with respect to the replication origin is highly nonrandom it is to be expected that many inversions will have subtle effects on the growth of strains. A major problem in understanding genetic changes that might occur in genetic instability is that the functions of about two-thirds of the genes in a microorganism are unknown. Thus even if a particular mutation in a specific gene can be associated with genetic instability, it may be difficult to understand why.

The mutation frequency in a strain may be altered by genetical or environmental factors. Mutator mutations can lead to a greatly enhanced rate of mutagenesis. Many mutator mutations lie in genes involved in DNA repair and may increase mutation rates up to 1,000-fold. Most mutator strains grow more slowly than their nonmutator parent strains. However, despite this, they win against nonmutator strains in chemostat competition experiments. This demonstrates the importance of mutation in conditions with high selection and indicates that inappropriate selection strategies could lead to enrichment of mutator strains. In *E. coli* and *Salmonella* a subpopulation of cells in stationary phase can enter a hypermutable state. This leads to the creation of new mutants in the absence of growth and can produce mutants that can grow and be enriched in the absence of growth of other cells in the population. It is likely that such mutagenic states exist in most microorganisms. It is thus important to culture and store industrial microorganisms in conditions that minimize the chance to enter such a state.

In some cases there are programmed mutation events that switch between expression of different sets of genes. One well known example is that of phase variation in *Salmonella* strains, in which a site-specific recombination system causes inversion of a segment of DNA. The invertible segment carries a promoter that reads out of it and can activate a neighboring operon; inversion leads to switching off this operon. In *Salmonella* this leads to a switch in the expression between two flagellar antigens that probably helps to defeat the host's immune system, but similar sys-

tems might control other functions in other bacteria. In some cases it is known that certain regions of the chromosome are particularly subject to instability and may not be suitable sites for introducing foreign DNA during strain construction. The replication terminator region of the *E. coli* chromosome is subject to a high frequency of DNA rearrangements. The end regions of the linear chromosomes of *Streptomyces* are also subject to genetic instability (5). This can lead to large deletions. (Deletions of over 1 Mb of the 8-Mb chromosome can occur without very dramatic phenotypic effects!) There can also be tandem DNA amplification of chromosomal sequences: the sequences can vary from 2 to 250 kb in size, and the copy number can be over 100. The amplifications may be accompanied by deletion of neighboring chromosomal sequences up to the chromosome end. These unstable regions may affect secondary metabolite production (Table 1). The mutations are usually deleterious for growth, compared with wild-type strains, but may give an advantage in comparison with high-production parents. The mutants can often be recognized by changes in colony morphology.

It may be difficult to detect mutations in bacteria because they may not show very obvious morphological or auxotrophy properties, and it is likely that the phenomenon of genetic instability is often not recognized. If the changes involve large-scale genome rearrangements, then pulsed-field gel electrophoresis (PFGE) may prove useful. PFGE can separate linear DNA fragments in size ranges up to about 10 Mb (bacterial genomes range in size from about 0.6 to 10 Mb). It is often possible to find rarely cutting enzymes that restrict the whole genome into 5–30 fragments that can be conveniently separated by PFGE. The gels can also be used for Southern blotting to detect specific DNA sequences. PFGE can be used to screen strains for DNA rearrangements (deletions, duplications, and sometimes inversions).

Phage infection can be a major problem in some industrial fermentations. In addition to lytic phages that kill all the cells, there are many lysogenic phages that maintain themselves within the host cells either by integrating into the host chromosome or by establishing a plasmid state. Lysogenic phages exist within some industrial strains (e.g., the phages RP2 and RP3 of *S. rimosus*, [6]) and may cause changes in the behavior of a strain. Some phages (e.g., phage Mu in *E. coli*) can integrate into random sites in the chromosome and inactivate genes, in other words, they are transposable elements).

Table 1. Genetic Instability Affecting Secondary Metabolite Production in *Streptomyces*

Species	Product	Effect
S. rimosus	Oxytetracycline	Deletion or amplification of genes
S. fradiae	Tylosin	Amplifications and deletions in biosynthesis cluster
S. griseus	Streptomycin	Deletion of A-factor gene, loss of streptomycin synthesis
S. ambofaciens	Spiromycin	DNA amplification leads to loss of production

It is possible that some changes observed in bacteria are of an epigenetic nature, that is, they are stable changes that do not depend on changes in the DNA sequence. One candidate for such an effect is the mycelial fragment size in *Streptomyces* fermentations. In many species fast growth is associated with larger fragment sizes. If a continuous culture in a chemostat is started from fast-growing mycelium, then large fragments are stably maintained at a variety of dilution rates. However, if the chemostat is started from a slow-growing culture entering stationary phase, then a small fragment size is maintained over a range of dilution rates. It is possible to switch from small fragments to large fragments by a short period of growth at high dilution rate, and from small fragments to large fragments by a short period at a low dilution rate. At dilution rates between these critical rates the form depends on the history of the culture and remains stable for many generations (7).

PLASMIDS

Plasmids are common in most groups of bacteria. Most plasmids are circular DNA molecules, but linear plasmids are present in some species. Natural plasmids may be of significance in some industrial strains, but the most important role of plasmids is for constructing recombinant strains. Natural plasmids come in a large range of sizes (2–1,000 kb) and copy numbers (1–500/chromosome) and have been used to construct vectors with different copy numbers for different applications. Two important classes of genetic instability that can affect plasmids are segregational instability (plasmid loss), and structural instability (DNA rearrangements in the plasmid). Low copy number plasmids have systems to ensure that plasmid copies are distributed to both daughter cells on division. Many natural plasmids carry conjugation systems that would reintroduce the plasmid into any member of the population that had lost it. However, biological containment considerations and the desire to obtain small vectors make this undesirable for recombinant plasmids. It is likely that many high copy number plasmids do not need an active system to segregate them to daughter cells but instead rely on the low statistical chance of producing plasmid-free progeny. (If there are N copies of a plasmid at cell division then the chance of producing a plasmid-free cell is 2^{-N}; for $N = 20$ this is 10^{-6}.) However, the situation is complicated if plasmid multimers are produced (i.e., circles containing two or more copies of the plasmid as a direct repeat). The copy number control determines the number of replication origins, so a plasmid dimer would have only half the number of molecules per cell as a monomer. Because multimers are generated by homologous recombination systems, plasmids often have site-specific recombination systems that break down multimers into monomers and result in a higher segregational stability (e.g., the *xer* system of ColE1 [8]). It is thus important to ensure that the stability-enhancing systems of natural plasmids are retained during vector construction; many

common laboratory cloning vectors have lost such systems because they were constructed with minimal-size replicons. Choice of fermentation conditions can make a large difference in the segregational stability of plasmids. The choice of limiting nutrient in chemostat cultures can radically affect plasmid loss. Choice of host may also have a significant effect; at present this can be tested only empirically because the basis for differences is unknown. The stability of plasmids is often affected by the growth rate of cultures. There are cases in which plasmids are more stable at high growth rates. Inocula for fermenters are typically prepared in several stages involving increasing volumes. An improvement in stability can be obtained by avoiding slow growth during inoculum preparation (i.e., transferring to a larger volume before the culture enters stationary state).

The replication systems of plasmids may differ significantly from that of the chromosome (9), and this may lead to a significantly higher frequency of mutation including DNA rearrangements. Some plasmids have a rolling-circle replication mode with single-stranded DNA intermediates. This system is more prone to DNA rearrangements than other replication systems. DNA rearrangements can include deletions, insertions, duplications, and inversions. The chromosomal replication control system is a positive-control system in which every replication origin in the cell is activated once in the cell cycle, and all origins are activated almost simultaneously. In contrast most plasmids have negative-control systems, which initiate replication at an origin at times spread through the cell cycle. A random pool mechanism functions, which means that replication may be initiated on a particular copy of the origin several times in a cell cycle, while another copy is not replicated. This means that any mutated copy of a plasmid with a small advantage for selection for replication can rapidly replace a parent plasmid. DNA rearrangements in plasmids are usually easily detected using restriction digests and agarose gel electrophoresis.

The selection of segregational or structural mutants can lead to rapid overgrowth of the original population. A plasmid-carrying strain may grow more slowly than one that has lost the plasmid, but this is by no means universally true for natural plasmids. However, it is likely that most constructs for industrial use will negatively affect the growth of the host strain. Although larger derivatives of a plasmid are often less stable than small derivatives, it is not clear that this is mainly due to the size per se. It seems much more the case that the presence of promoters on the plasmid can lead to selection against the insert. In some plasmids, transcription from insert promoters into the vector can lead to stability problems, possibly by interfering with vector replication. Therefore it has become common practice to insert transcriptional terminators into vectors to prevent read-through from the insert into the vector; this was one key step in constructing successful vectors for *Bacillus*. It is also desirable only to transcribe the cloned gene when product is required because this often reduces selection against the insert. Various inducible promoter systems have been developed, such as the recombinant *tac* promoter, which combines the desirable strength of the *trp* operon promoter with the control of the *lac* promoter (induction with IPTG). Other promoters are the lambda P_R and P_L promoters, which can be controlled by temperature by using a temperature-sensitive repressor. Another possibility is to use a promoter from phage T7 that can be transcribed only when the special T7 RNA polymerase is expressed. By using a combination of plasmid stability systems and control of the transcription of cloned genes it is possible to achieve good results in *E. coli*. However, even in *E. coli* there is a need for considerable empirical work, and many surprises will occur. In other bacterial systems the state of knowledge about plasmid stability is considerably worse, and a good portion of luck is necessary to achieve success.

YEAST AND FUNGI

Many of the same considerations apply to yeast and fungi as to bacteria. However, the more complex eukaryotic nature of these organisms leads to some further mechanisms for generating mutants. Eukaryotes have a number of linear chromosomes. Many yeast and fungi are haploid, but there are also diploid and dikaryotic (i.e., two separate haploid nuclei in the same cytoplasm) strains. There can be changes in the number of chromosomes (either in the number of copies of the genome present, "ploidy"; or in the loss or gain of copies of single chromosomes, "aneuploidy"). There can also be transfer of parts of chromosomes to other chromosomes (translocation). Most yeast and fungi also have sexual systems in which meiosis takes place. Such sexual cycles are usually associated with stress conditions (e.g., nutrient stress) and may be avoidable by suitable strain maintenance conditions. Meiosis can lead to changes in the genetic make up of cultures when diploid or dikaryotic organisms are used. Although many of the present important industrial strains are haploid, this may become more of a problem in the future as a wider range of organisms is used in biotechnology. Sexual systems usually involve the interaction of two different mating types; some organisms have only two sexes, but many fungi have several mating types. The mating types may be able to change: the best-characterized example of this is the mating-type switch in *S. cerevisiae* in which it is possible for a strain to switch backward and forward between the two mating types, a and α. There are strains that switch at high frequencies (homothallic strains) due to the presence of a gene, *HOM*, encoding an endonuclease, whereas other strains switch only at low frequency (heterothallic). The ability to switch between mating types can thus be seen as a sort of genetic instability and may have consequences for fermentation if one mating type is selected under fermenter conditions.

Transposable elements are also present in yeast and fungi. One class that has not been found in bacteria the retrotransposons, which resemble retroviruses. They transpose via an RNA intermediate and transcribe into neighboring genes. The example of *Ty1* in yeast shows that this can activate genes by providing a new promoter. This ability to activate as well as inactivate genes could be relevant to genetic stability. Such elements are present in many copies in the cell (>30 copies), which makes it un-

likely that they can be removed by genetic manipulation. Apart from the ability to change gene expression patterns by transposition, it is possible that mutations in control proteins will result in new patterns. The *areA* gene in *Aspergillus nidulans* controls ammonium repression for a large number of genes involved in the catabolism of various nonfavored nitrogen sources (10). Many mutations in *areA* are known that differentially change the pattern of control (i.e., some genes are silenced and some are overexpressed). In one case a duplication within the *areA* gene resulted in another gene becoming subject to *areA* control. This mutational flexibility would allow easy evolution of new expression patterns and could contribute to genetic instability.

Another feature of eukaryotes that makes them genetically more complex than bacteria is the possession of mitochondria, which also have their own genome (mtDNA). In *S. cerevisiae* there are very frequent deletions in the mtDNA that cause a loss of mitochondrial function. The organisms can still grow, but the slower growth results in small colonies (which is why they are called *petite* mutants). These mutants account for 1% of colonies or more, depending on the strain. Another genetic phenomenon caused by mtDNA is senescence (11) in the fungus *Podospora anserina*. This species cannot grow indefinitely vegetatively and, after a certain amount of mycelial growth, death occurs. The senescence is caused by the accumulation of a high number of copies of a plasmid formed by excision from the mtDNA. Such senescence phenomena may cause problems in the maintenance of industrial strains.

Problems of genetic instability can also arise in culture collections. For example, it is by no means uncommon for the secondary metabolite spectrum of a filamentous fungus to change radically if it is examined with the same strain from a culture collection after a gap of several years. Although this problem can arise with any organism, many bacteria can be kept in nongrowing states (e.g., lyophilization, storage at -70 °C in glycerol), which limits the scope for selection to act. However, most filamentous fungi cannot be maintained in this way. In view of the number of different strains that must be maintained, there is probably no real answer to the problem. It is likely that adaptive mutation occurs in most microorganisms as a response to certain stresses and might be one cause for such strain degeneration. Experience can guide the choice of media and storage conditions to try to minimize these effects.

Epigenetic effects can also be important in yeast and fungi. The DNA of most eukaryotes contains some 5-methyl cytosine, which is formed by methylation after DNA replication; this usually occurs at CG dinucleotides, and so the complementary strand can also be methylated. If a methylated sequence is replicated, two hemimethylated daughter molecules (i.e., each methylated in only one strand) are produced. Because DNA methylase enzymes usually prefer such hemimethylated substrates this results in an inheritance of the methylation patterns. Methylation of genes is associated with their nonexpression, but it is not clear how the de novo methylation arises. In some cases the introduction of foreign genes into fungi can result in their inactivation by methylation. The reason why this occurs is not known. In some cases methylation seems to be targeted at duplicate copies of a gene in an organism, that is, introduction of an extra copy of a gene will result in inactivation of either the introduced copy or the original copy. The methylation may also lead to mutation. 5-Methyl cytosine can be spontaneously deaminated to produce thymine and, in the absence of repair, this gives a C → T transition. This can result in a "RIP effect" in some fungi in which a very high mutation rate is observed after introduction of an extra copy of a gene such that one copy is inactivated by multiple C → T transitions (12). A further epigenetic effect that has been observed in *S. cerevisiae* is the existence of prions (13). These are proteins of altered structure that can catalyze the change of copies of normal structure to the altered structure. This can function as a sort of "infective" process and is a potential source of instability.

ANIMAL AND PLANT CELLS

Animal cells have been used for the growth of viruses for the production of vaccines and in more recent years for the production of recombinant proteins. Plant cells are not yet used on an industrial scale, but there has been interest in this possibility. Smaller-scale cultivation of animal cells has been used extensively in laboratories. Most primary cell cultures have a limited life span in the sense that they can divide only a limited number of times. However, immortalized cell lines, which do not have this limitation, are of most interest for large-scale culture. These are tumour cell lines that have often changed considerably in their properties as compared with the original cells from which they derive. Cultured animal cells are often genetically unstable and undergo frequent karyotype changes. These include polyploidy (changes in the number of chromosome sets from diploidy to triploidy and tetraploidy), aneuploidy (loss or gain of copies of particular chromosomes), and frequent chromosomal rearrangements (deletions, inversions, and translocations).

Chinese hamster ovary (CHO) cells are an immortalized cell line that has been used to produce recombinant proteins at a 10,000-L scale (14). They possess good growth properties for large-scale culture and pose less safety problems with respect to viruses than cell lines more closely related to humans. An important strategy to obtain increased production of recombinant products is to work with high copy numbers of the introduced genes. This can be done by using a cell line that lacks dihydrofolate reductase (DHFR) and transforming with the required gene coupled to the DHFR gene. The antibiotic methotrexate inhibits DHFR, and selection for growth on increasing levels of DHFR results in the amplification of the DHFR gene and coupled sequences. The amplifications can have a complex structure and be either extrachromosomal or integrated in the chromosome. Continued selection usually results in integrated copies, but the positions of integration vary more or less randomly between independent clones.

Control of stability requires the careful construction of master cell banks and working cell banks and the avoidance of too many generations of growth. Various methods

can be used to check genetic stability. One method is to use Southern blots. However, this does not show heterogeneity in the population. A better method is to use karyotype-based methods. In particular, fluorescent in situ hybridization (FISH) can locate the amplified sequences and show heterogeneity. If chromosome-specific probes are used it is possible to monitor the copy number of a chromosome in interphase nuclei. For instance, Figure 2 shows a normal diploid cell and a cell carrying five copies of a chromosome. If metaphase cells are used, details of chromosome rearrangements can be monitored with appropriate probes. It would normally be necessary to carry out FISH and measurements of productivity at several times during growing cultures. Flow cytometry can be used to show variations in DNA content in a population of cells that can indicate the extent of polyploidy and, to some extent, aneuploidy.

The fact that the amplifications are selected with methotrexate and that lower producers might be expected to overgrow the best producers would suggest that stability would be enhanced if methotrexate were present in the medium. However, there are data showing an increased instability with methotrexate (15). One explanation is that methotrexate leads to a reduction of thymidine synthesis, which could allow more dUTP incorporation into the DNA. The resulting repair and strand cutting might stimulate instability. In fact an experiment with a heterogeneous population of clones in the absence of methotrexate suggested that such cultures may be surprisingly stable in the absence of selection.

VIRUSES

Viruses are propagated for vaccine manufacture. This includes the use of live attenuated viruses and killed viruses. Various culture systems are used, such as cell cultures and chicken eggs. Viruses show a bewildering range of different genome structures (single- and double-stranded DNA or RNA) with specialized replication methods. Virus replication may be much less accurate than host replication (especially for RNA viruses) and thus show a high mutation rate. Genetic instability has vital biological safety considerations because a reversion of attenuated viruses to higher pathogenicity or changes that reduce the efficiency of killing in killed-virus vaccines could be disastrous. Any release of viruses from the cells used to culture the virus could also have serious implications. The main weapon against these problems is to reduce the number of growth cycles as much as possible. It is also important to monitor the cultures and product.

CONSEQUENCES FOR GOOD MANUFACTURING PRACTICE

Good manufacturing practice aims to guarantee reproducibility of production runs and avoid product contamination. The details are dependent on the organism and product, but there are some general considerations:

1. A master cell bank (MCB) is derived from a single colony or single cell. It consists of many replicates kept in the most stable form possible (e.g., lyophilized or frozen at $-70\ ^{\circ}C$ or in liquid nitrogen). Each ampoule is opened only once. It is important to make sufficient replicates because generation of a new MCB will require an application to the regulatory authority. It is also usual to store the material at more than one location in case of destruction (e.g., freezer problems, fire).

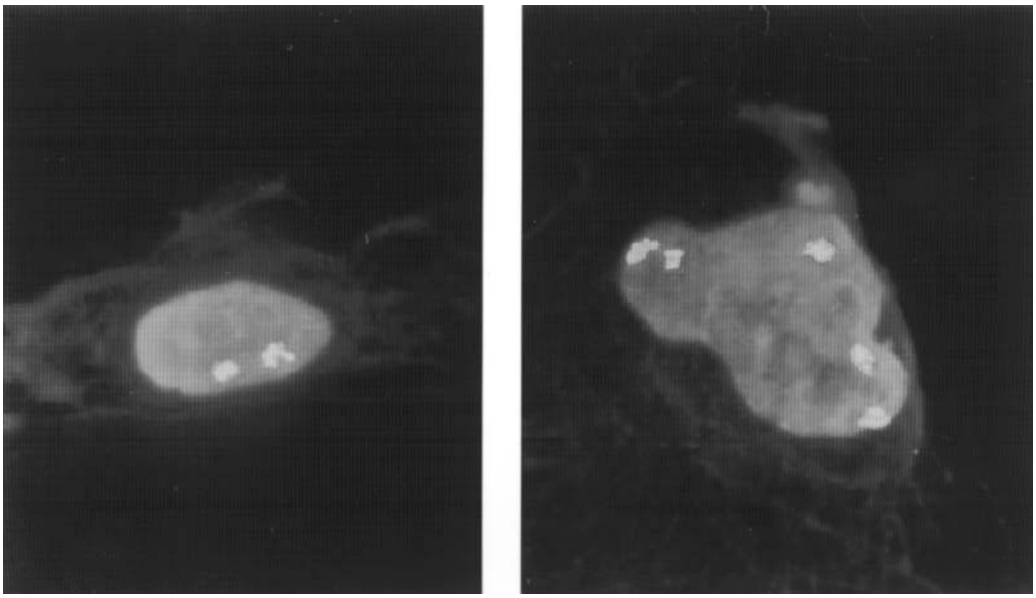


Figure 2. Fluorescent in situ hybridization (FISH) of human fibroblast cells with a probe specific for repetitive sequences of chromosome 1. One interphase nucleus shows the expected two signals for a diploid cell whereas the other nucleus is aberrant with five copies of the chromosome. *Source*: Photograph kindly provided by Dr. Harry Scherthan, Department of Human Biology, University of Kaiserslautern.

2. The working cell bank (WCB) is derived from the MCB and is used to inoculate a production batch.
3. Detailed records of both MCBs and WCBs must be kept, and all possible efforts must be made to limit the number of growth cycles (number of passages) to minimize the effects of genetic instability.
4. Both MCBs and WCBs must be rigorously tested. The tests for MCBs will be particularly exhaustive and depend on the process. A genotypic characterization using various DNA tests is usual (e.g., DNA fingerprinting to confirm the strain, and testing the structure of recombinant DNA constructs). There will also be phenotypic tests and tests that normal production levels occur. There may be tests for contamination (e.g., viruses in cell culture).

These procedures aim to confirm the reproducibility of the starter cultures and the production process. As long as the build up of any variants is fairly reproducible, genetic instability during the fermentation will not conflict with these aims but will lead to a reduction in production compared with the ideal case. However, any evidence for genetic instability and the occurrence of heterogeneity during production fermentations is likely to lead to problems with the regulatory authorities. It is likely that this also contributes to the reluctance to consider genetic instability in many fermentation sections of companies.

BIBLIOGRAPHY

1. J. Cullum, J. Altenbuchner, F. Flett, and W. Piendl, *Gen. Eng. Rev.* **4**, 59–78 (1986).
2. B. Gravius, T. Bezmalinović, D. Hranueli, and Cullum J, *Appl. Environ. Microbiol.* **59**, 2220–2228 (1993).
3. T. Miyoshi, M. Iseki, T. Konomi, and H. Imanaka, *J. Antibiot.* **33**, 480–487 (1980).
4. P. Kaiser, F. Flett, and J. Cullum, *Bio / Technology* **10**, 570–573 (1992).
5. M. Redenbach, F. Flett, W. Piendl, I. Glocker, U. Rauland, O. Wafzig, R. Kliem, P. Leblond, and J. Cullum, *Mol. Gen. Genet.* **241**, 255–262 (1993).
6. H. Rausch, M. Vešligaj, D. Počta, G. Biuković, J. Pigac, J. Cullum, H. Schmieger, and D. Hranueli, *J. Gen. Microbiol.* **139**, 2517–2524 (1993).
7. D. Noack, *J. Basic Microbiol.* **28**, 101–106 (1988).
8. S.D. Colloms, P. Sykora, G. Szatmari, and D.J. Sherratt, *J. Bacteriol.* **172**, 6973–6980 (1990).
9. G. del Solar, R. Giralso, M.J. Ruiz-Echevarría, M. Espinosa, and R. Díaz-Orejas, *Microbiol. Mol. Biol. Rev.* **62**, 434–464 (1998).
10. N.M. Crawford and H.N. Arst, Jr., *Ann. Rev. Genet.* **27**, 115–146 (1993).
11. A.J.F. Griffiths, *Ann. Rev. Genet.* **26**, 351–372 (1992).
12. E.U. Selker, *Ann. Rev. Genet.* **24**, 579–613 (1990).
13. R.B. Wickner, *Ann. Rev. Genet.* **30**, 109–139 (1996).
14. W.R. Arathoon and J.R. Birch, *Science* **232**, 1390–1395 (1986).
15. F.M. Wurm. in H. Hauser and R. Wagner eds., *Mammalian Cell Biotechnology in Protein Production*, Walter de Gruyter, Berlin, 1997, pp. 87–118.

GLUCOSIDASES

RYUICHI MATSUMO
Kyoto University
Kyoto, Japan

KEY WORDS

Alkyl glucoside
Condensation
Reverse hydrolysis
Surfactant
Transfer reaction
Two-phase system

OUTLINE

INTRODUCTION

β-Glucosidase has the potential for practical use in the saccharification of cellulosic materials and the synthesis of aryl or alkyl glucosides in a nonaqueous reaction medium. The latter topic, especially synthesis in an aqueous–organic two-phase system, has been described in detail.

β-Glucosidase (*β*-D-glucoside glucohydrolase, EC 3.2.1.21) catalyzes the cleavage of the *β*-glucosidic ether bond of aryl and alkyl glucosides. Aryl glucosides are, in general, better substrates than alkyl glucoside. *β*-Glucosidase from some origins can also hydrolyze *β*-galactosides and *β*-fucosides. Its hydrolytic potential in practical use would be as a component of cellulases for saccharification of cellulose.

The enzyme also has a synthetic potential for aryl or alkyl glucosides through reverse hydrolysis (condensation) or transfer reactions. Research concerning *β*-glucosidase-catalyzed synthesis of glucosides in a nonaqueous phase has recently been active, although its industrial application has not been reported yet.

HYDROLYTIC REACTIONS

Saccharification of Cellulose

β-Glucosidase is one of the key enzymes for saccharification of cellulose. The enzymatic saccharification proceeds through synergistic actions of *endo*-glycanase, *exo*-glycanase, and *β*-glucosides, and *β*-glucosidase catalyzes

the last step of the saccarification, hydrolysis of cellobiose to glucose. β-Glucosidase favors shorter cellooligosaccharides, especially cellobiose.

β-Glucosidase activity of cellulase from *Trichoderma reesei*, which is one of the most promising microorganisms for cellulose saccharification on an industrial scale, is still weak, so addition of β-glucosidase from other origins for the cellulose saccharification process has been attempted, resulting in a significant increase in cellulose conversion (1). Reinforcement of β-glucosidase activity of *T. reesei* cellulase by gene technology has also been investigated (2).

Hydrolysis of Alkyl Glucoside

Alkyl glucosides, which are substrates for β-glucosidase, have an amphiphilic structure with hydrophilic and hydrophobic moieties and are a group of nonionic surfactants. The phase diagram for a ternary mixture of octyl β-D-glucoside–buffer(water)–octanol was prepared, and it included the regions of micellar solution, microemulsion, liquid crystalline phase, and an undefined structure (3). In the aqueous monophasic system, octyl β-D-glucoside exists as monomers at concentrations below its critical micellar concentration (CMC), and it forms aggregates at concentrations above the CMC, while the monomer concentration remains at the CMC. Figure 1 shows the initial reaction rates for β-glucosidase-catalyzed hydrolysis of octyl β-D-glucoside at various substrate concentrations (4). For the lower substrate concentrations (< ca. 20 mmol/L), the relation between the reaction rate and substrate concentration followed a classical Michaelis–Menten scheme, but the reaction rate remained constant at the higher substrate concentrations. As shown in the insert of Fig. 1, the relation could be divided into two regions, and two lines intersected at 19 mmol/L, which corresponded to the CMC of the system. This indicates that octyl β-D-glucoside monomer is the actual substrate for the enzyme.

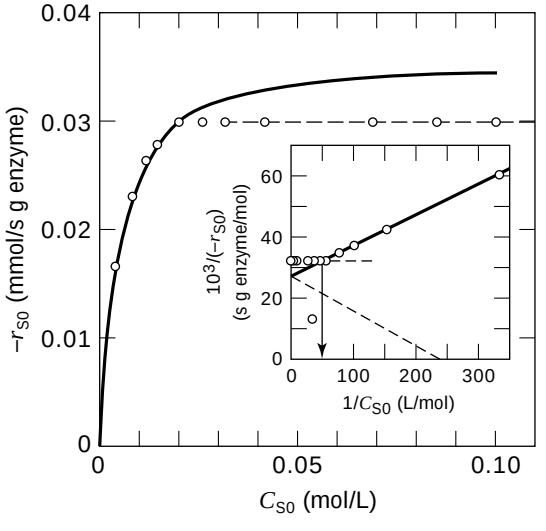


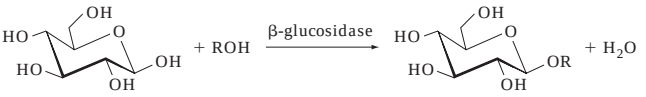
Figure 1. β-Glucosidase-catalyzed hydrolysis of octyl β-D-glucoside. *Source*: Adapted from Ref. 3.

SYNTHESIS OF ALKYL GLUCOSIDES

β-Glucosidase-catalyzed synthesis of aryl or alkyl glucosides in a nonaqueous system has been an active area of research, as have lipase-catalyzed acylation and proteinase-catalyzed peptide synthesis.

Chemical synthesis of alkyl glucosides is well developed and used at the commercial level. However, the processes are complicated by the presence of identically reactive hydroxyl groups and by the toxicity of some of the catalysts used. β-Glucosidase is effective for synthesis through transglycosylation (5,6) or condensation (7–14). Although the reaction rate of condensation is slower and hence a longer time is required to reach equilibrium compared with transglycosylation, condensation has some advantages, such as cheaper substrates and easier purification of the product. Alkyl glucoside synthesis through transglycosylation is kinetically controlled, but the yield of alkyl glucoside through transglycosylation was almost equal to that through condensation (5). On the basis of substrate costs, condensation would be a choice for alkyl glucoside synthesis.

Equation 1 shows the alkyl glucoside synthesis through β-glucosidase-catalyzed condensation between glucose and alcohol.



(1)

where ROH indicates aliphatic alcohol. Thermodynamically, hydrolysis predominates over synthesis in an aqueous reaction medium. One approach to shift the equilibrium is to add a water-miscible solvent, which decreases the thermodynamic water activity and thereby favors synthesis. An alternative approach is to add a water-immiscible solvent. In this case, a hydrophobic product is extracted to the organic phase so that more product is synthesized.

One-Phase System

Aryl and alkyl glucosides were synthesized in a mixture of acetonitrile and water through condensation using β-glucosidase from almonds (7). The synthetic reaction was entirely regio- and stereoselective. The enzyme had a pronounced affinity for the alcohols containing a phenyl group. The ratio of water to acetonitrile was a crucial factor for both the product yield and the enzyme stability, a 9:1 (v/v) acetonitrile–water mixture being found to be the best. It was also reported that the enzyme could synthesize alkyl β-D-galactosides.

Two-Phase System

In the synthesis of alkyl glucosides in an aqueous–organic two-phase system, one of the substrates, alcohol, itself is used as an organic phase. The aqueous phase includes glucose and free (9,10,12,13) or immobilized (5,8,11) β-glucosidase. This synthesis has mostly been used for alkyl glucoside synthesis. The water activity a_w or the water

content in the system plays an important role in both the reaction rate and the equilibrium yield in alkyl glucoside synthesis (5,8–11). An a_w of at least 0.67 was needed for synthesis of octyl β-D-glucoside. The reaction rate increased, but the equilibrium yield decreased with increasing a_w. A high yield of octyl glucoside in a short time was realized by controlling the a_w during the condensation reaction (11). The optimum water content was higher for the synthesis of glucoside with a shorter alkyl residue (8,10). Physicochemical understanding for these results have not been established yet.

Figure 2 shows the transient changes in the yield of alkyl β-D-glucosides, Y_P, through β-glucosidase-catalyzed condensation between glucose and alcohols at 60 °C in the two-phase system (13). The volume ratio of organic to aqueous phase was fixed at 16.7. The length of the alkyl chain significantly affected both the equilibrium yield and the reaction rate. The longer the alkyl chain, the lower were both the yield and the rate. The figure also shows that the enzyme could catalyze the condensation between glucose and secondary alcohols, although the yield of 2-alkyl glucoside was lower than that of 1-alkyl glucoside.

In the two-phase system, the product partitioned to the organic phase. Because it is known (15) that β-glucosidase catalyzes the condensation of glucose molecules to form oligosaccharides, mainly β-glucobioses, at higher glucose concentrations, the undesirable condensation would also occur in the system as well as the intended condensation between glucose and alcohol in the system. Figure 3 summarizes phenomena assumed to occur in the two-phase system for alkyl glucoside synthesis. Under the following assumptions, *(1)* the condensations occur only in the aqueous phase containing the enzyme; *(2)* the product P, alkyl glucoside, can partition to the organic phase with a partition coefficient P_{app}; *(3)* the glucose, G, and produced β-glucobioses, D, locate only in the aqueous phase due to their high hydrophilicity; *(4)* the aqueous phase is always saturated with an alcohol, A, that is, the concentration of the alcohol in the aqueous phase, $C_{A,aq}$, is given by its solubility in water, and *(5)* the water concentration, C_W, is kept constant during the reaction, equations for estimation of the equilibrium yields of alkyl glucoside, Y_P, and β-glucobioses, Y_D, through β-glucosidase-catalyzed condensation in the two-phase system have been derived (13):

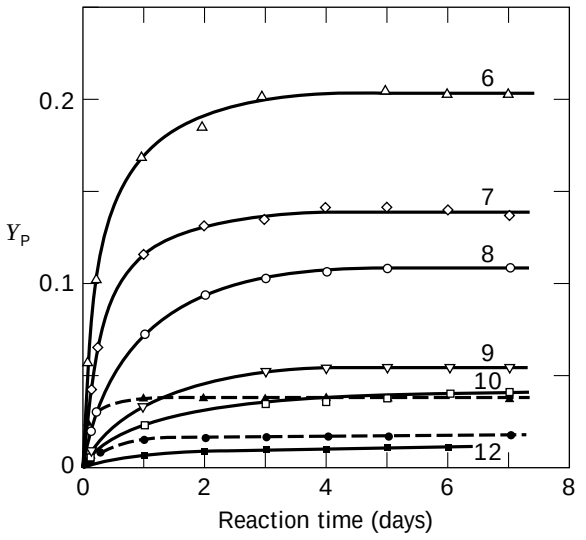


Figure 2. Kinetics for the synthesis of alkyl β-D-glucosides through β-glucosidase-catalyzed condensation of glucose and various alcohols in a two-phase system. The numbers represent the alkyl chain lengths of glucosides produced. The open symbols represent 1-alkyl β-D-glucosides, and the closed ones 2-alkyl β-D-glucosides.

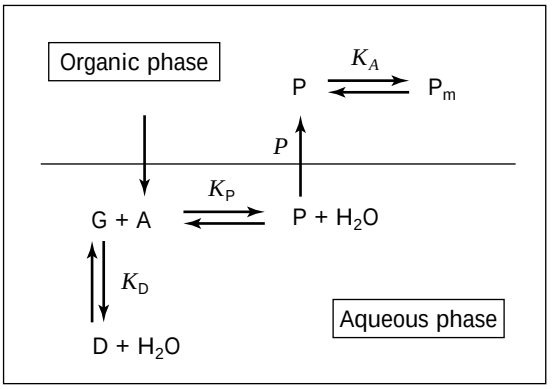


Figure 3. Reaction scheme supposed to occur in the synthesis of alkyl β-D-glucoside through β-glucosidase-catalyzed condensation in an aqueous–organic two-phase system. A, D, G, P, and Pm denote alcohol, β-glucobioses, glucose, alkyl glucoside, and the *m*-meric aggregate of alkyl glucoside, respectively. K_A is the association constant for the *m*-meric aggregate, K_D is the equilibrium constant for the formation of β-glucobioses, K_P is the equilibrium constant for the alkyl glucoside formation, and P is the partition coefficient of the monomeric alkyl glucoside.

$$Y_P = \frac{X}{4K_D C_{G0}/C_W}\left\{\left[(1+X)^2 + \frac{8K_D C_{G0}}{C_W}\right]^{1/2} - (1+X)\right\} \quad (2)$$

$$Y_D = 1 - \frac{C_W(1+X)}{4K_D C_{G0}} \cdot \left\{\left[(1+X)^2 + \frac{8K_D C_{G0}}{C_W}\right]^{1/2} - (1+X)\right\} \quad (3)$$

where

$$X = (1 + \alpha P_{app})K_P C_{P,aq}/C_W \quad (4)$$

C_{G0} is the initial glucose concentration in the aqueous phase, $C_{P,aq}$ is the alkyl glucoside concentration in the aqueous phase, α is the volume ratio of aqueous to organic phase, and K_P and K_D are the equilibrium constants for the alkyl glucoside and β-glucobiose formation, respectively. The water concentration can be calculated by $C_W = -6.43C_{G0} + 54.58$, where both C_W and C_{G0} are in units of moles per liter.

The equilibrium constant for formation of β-glucobioses, K_D, was determined to be 1.17. The equilibrium constant for formation of 1-alkyl β-D-glucosides with various alkyl chains, K_P, was evaluated for the experimentally observed

Y_P values according to equations 2 and 4 (Fig. 4). The K_P value was independent of alkyl chain length and was 1.90. It was also shown that alkyl glucoside with any alkyl chain formed a dimeric aggregate in the alcohol phase with an association constant K_A of 4.2 L/mol, which was independent of alkyl chain length, and that the partition coefficient of the monomeric alkyl glucoside, P, increased exponentially with the alkyl chain length, n ($\ln P = 1.08n - 5.17$). The apparent partition coefficient, P_{app}, can be calculated by

$$P_{app} = P(1 + 2K_A P C_{P,aq}) \quad (5)$$

Using these equations and parameters, we can calculate the equilibrium yields, Y_P and Y_D, under any conditions. The volume ratio of aqueous to organic phase, α, is an important factor affecting the yields. At higher α values, the equilibrium yields of alkyl glucosides are higher, while the equilibrium yields of β-glucobioses become lower. These results indicate that higher α values are preferable for alkyl glucoside synthesis. However, the concentrations of alkyl glucosides in their corresponding alcohol phase, $C_{P,org}$, decrease with an increase in α. The initial glucose concentration, C_{G0}, is another important factor. The yield, Y_D, of β-glucobioses, undesirable products in alkyl glucoside synthesis, largely increases at higher C_{G0}. Although the yield of alkyl glucoside, Y_P, also increases with increasing C_{G0}, the extent is not as great as that for the yield of β-glucobioses.

A Heterogeneous System

Alkyl β-D-glucoside could be synthesized in a heterogeneous system where solid glucose and β-glucosidase are added in a buffer-saturated alcohol (14). In such a system, a very viscous phase appeared, and it could be easily separated from the alcohol phase after the condensation reaction. Because the enzyme in the viscous phase was expected to be still active, the phase was repeatedly used for the synthesis of alkyl glucoside. Figure 5 shows 1-octyl β-D-glucoside synthesis by repeated use of the viscous phase. The enzyme was added only in the first run, and no further addition was made in subsequent batches. In each batch, glucose was added in an amount equal to that of the product in the previous batch. In this system, the overall yield of alkyl glucoside could be significantly improved without immobilization of the enzyme. The reason for the increase in the overall yield by the consecutive synthesis can be explained by

$$Y_{P,n} = \frac{\sum_{i=1}^{n} (VC_{P,i})}{VC_{G0} + \sum_{i=1}^{n} (VC_{P,i})} = \frac{\sum_{i=1}^{n} C_{P,i}}{C_{G0} + \sum_{i=1}^{n} C_{P,i}} \quad (6)$$

where $Y_{P,n}$ is the overall yield of alkyl glucoside at the end of nth batch, $C_{P,i}$ is the product concentration at the end of ith batch, C_{G0} is the initial glucose concentration in the first batch, based on the total volume of reaction medium, V, which is nearly equal to the volume of the alcohol phase, and n is the number of cycles. Equation 6 indicates that the overall yield, $Y_{P,n}$, asymptotically approaches unity when n increases.

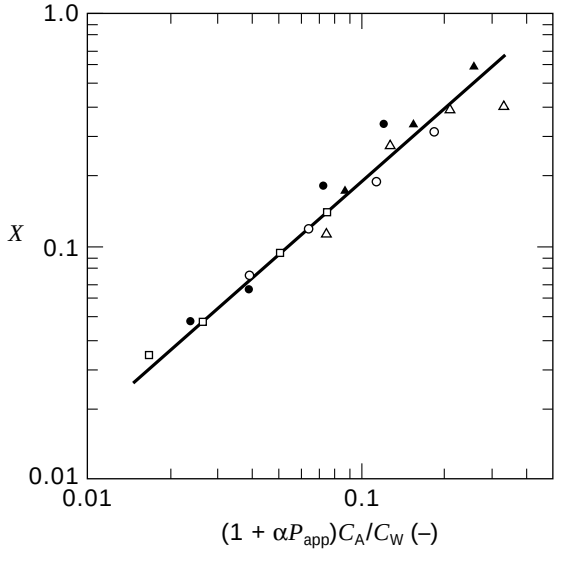


Figure 4. Estimation of the equilibrium constant, K_P, for the formation of 1-alkyl β-D-glucosides at 60 °C. X in the ordinate is defined by equation 4. The X values that satisfied the experimental equilibrium yields of alkyl glucosides in equation 2 were determined numerically. The K_P value was determined according to equation 4. The keys are the same as those in Figure 2.

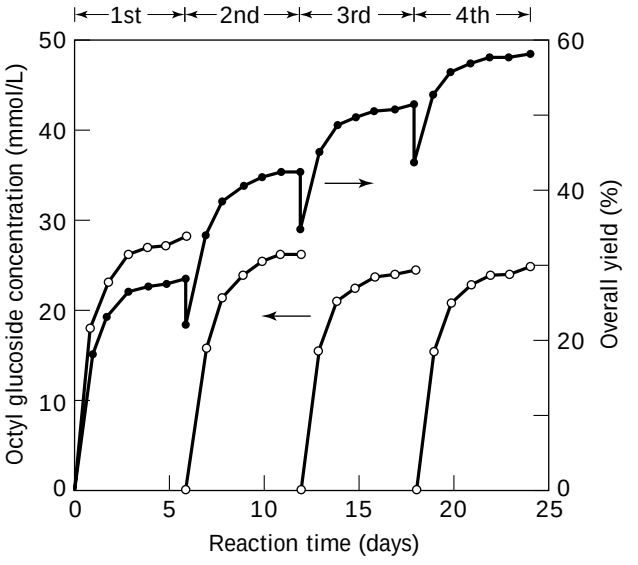


Figure 5. Repeated use of β-glucosidase for the synthesis of 1-octyl β-D-glucoside in the heterogeneous system. The enzyme was added only in the first batch, and no further addition was made in subsequent batches.

BIBLIOGRAPHY

1. B.C. Stockton, D.J. Mitchell, K. Grohmann, and M.E. Himmel, *Biotechnol. Lett.* **13**, 57–62 (1991).

2. C. Barnett, R.M. Berka, and T. Fowler, *Bio/Technology* **9**, 562–567 (1991).
3. J. Chopineau, D. Thomas, and M.-D. Legoy, *Eur. J. Biochem.* **183**, 459–463 (1989).
4. J. Chopineau, M. Ollivon, and M.-D. Legoy, *Ann. N.Y. Acad. Sci.* **672**, 566–571 (1995).
5. E.M. Vulfson, R. Patel, J.E. Beecher, A.T. Andrews, and B.A. Law, *Enzyme Microb. Technol.* **12**, 950–954 (1990).
6. N. Mitsuo, H. Takeichi, and T. Satoh, *Chem. Pharm. Bull.* **32**, 1183–1187 (1984).
7. G. Vic, J. Biton, D.L. Beller, J.-M. Michel, and D. Thomas. *Biotechnol. Bioeng.* **46**, 109–116 (1995).
8. E.N. Vulfson, R. Patel, and B.A. Law, *Biotechnol. Lett.* **12**, 397–402 (1990).
9. V. Laroute and R.M. Willemot, *Biotechnol. Lett.* **14**, 167–174 (1992).
10. Z. Chahid, D. Montet, M. Pina, and J. Graille, *Biotechnol. Lett.* **14**, 281–284 (1992).
11. G. Ljunger, P. Adlercreutz, and B. Mattiasson: *Enzyme Microb. Technol.* **16**, 751–755 (1994).
12. Z. Chahid, D. Montet, M. Pina, F. Bonnot, and J. Graille, *Biotechnol. Lett.* **16**, 795–800 (1994).
13. C. Panintrarux, S. Adachi, Y. Araki, Y. Kimura, and R. Matsuno, *Enzyme Microb. Technol.* **17**, 32–40 (1995).
14. C. Panintrarux, S. Adachi, and R. Matsuno, *J. Mol. Cat. B: Enzymatic* **1**, 165–172 (1996).
15. K. Ajisaka, H. Nishida, and H. Fujimoto, *Biotechnol. Lett.* **9**, 243–248 (1987).

GLUTAMIC ACID. See AMINO ACIDS, GLUTAMATE; AMINO ACIDS, PRODUCTION PROCESSES; GLUTAMIC ACID PRODUCING MICROORGANISMS; PRODUCTION OF L-GLUTAMIC ACID.

GLUTAMIC ACID PRODUCING MICROORGANISMS

SHUKUO KINOSHITA
Kyowa Hakko Kogyo Co., Ltd.
Tokyo, Japan

KEY WORDS

Amino acids
Cell membrane
Corynebacterium
Corynebacterium glutamicum
Glutamic acid
Morphology
Taxonomy

OUTLINE

INTRODUCTION

Amino acid fermentation is now the key industry supplying natural amino acids on an industrial scale. The discovery of a glutamic acid producing bacterium by Kinoshita et al. in 1956 was the start of the subsequent development of amino acid production by regulation of biosynthetic metabolism. At that time, a two-step L-glutamic acid production process through reductive amination of α-ketoglutaric acid had been studied. Instead, screening of glutamic acid producing microorganisms was carried out using a glutamic acid requiring strain of *Leuconostoc mesenteroides* to directly produce glutamic acid by fermentation from carbohydrate, and glutamic acid producing bacteria were isolated from nature and some habitats with high frequency. The glutamic acid producing bacterium first discovered was reported as *Micrococcus glutamicus* in 1958 (1), and other glutamic acid producing coryneform species, all patent strains, were subsequently isolated by many scientists. They were subjected to experiments, and industrial production began with various new species of the genera *Brevibacterium*, *Corynebacterium*, *Microbacterium*, and *Arthrobacter* (Table 1).

Table 1. Glutamic Acid Producing Bacteria

Strain	Isolator
Corynebacterium glutamicum ATCC12032	Kyowa Hakko
Brevibacterium divaricatum ATCC14020	Wei-chuan Foods
Corynebacterium lilium ATCC15990	International Mineral & Chemistry
Corynebacterium callunae ATCC15991	International Mineral & Chemistry
Brevibacterium saccharolyticum ATCC14066	Ajinomoto
Brevibacterium fravum ATCC14067	Ajinomoto
Brevibacterium lactofermentum ATCC13869	Ajinomoto
Brevibacterium roseum ATCC13825	Ajinomoto
Corynebacterium herculis ATCC13868	Herculis Powder
Brevibacterium thiogenitalis ATCC19240	Takeda
Microbacterium ammoniaphilum ATCC15354	Asahi Kasei

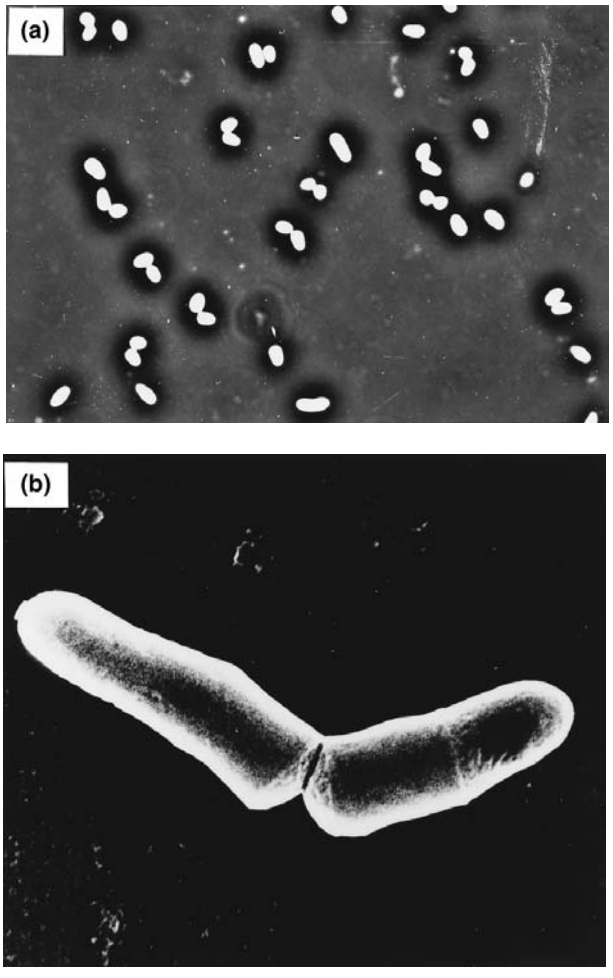


Figure 1. Morphologies of *Corynebacterium glutamicum* ATCC 13032^{T}: (**a**) Phase-contrast micrograph; (**b**) scanning electron micrograph.

CHANGES IN IDEAS REGARDING CORYNEFORM BACTERIA AND THE GENUS *CORYNEBACTERIUM* IN *BERGEY'S MANUAL*

The genus *Corynebacterium* was proposed by Lehmann and Neumann in 1896 (2) to accommodate the diphtheria bacillus. Thereafter, various isolated strains not only of pathogenic bacteria for animal and plants but also of saprophytic bacteria were identified as members of the genus *Corynebacterium*. In this way, many bacterial species with diverse properties were simultaneously classified into this genus. Under the circumstances, Jehsen (3) introduced a generic term of *coryneform* to clarify this genus definition. The word *coryneform* was coined by Oerskov in 1923 (4) to explain the morphology of ray fungi. Jensen used the word *coryneform* to describe one of the more characteristic morphological features, V formation (Fig. 1), particularly on cell division of aerobic and anaerobic gram-positive, nonacid-fast, pleomorphic, nonbranching, rod-shaped, nonsporogenous bacilli. However, it was difficult to distinguish between coryneform bacteria and strains with more or less

Table 2. Taxonomic Position of the Genus *Corynebacterium* and Related Genera in *Bergey's Manual*

Bergey's Manual Edition (year)	Position of the genus *Corynebacterium* and related taxa
Determinative Bacteriology Fifth (1939)	Order II. Actinomycetales Family I. Mycobacteriaceae Genus I. *Corynebacterium* Genus II. *Mycobacterium* Family II. Actinomycetaceae Genus I. *Leptotrichia* Genus II. *Erysipelothrix* Genus III. *Proactinomyces* Genus IV. *Actinomyces*
Determinative Bacteriology Sixth (1948)	Order I. Eubacteriales Family VIII. Corynebacteriaceae Genus I. *Corynebacterium* Genus II. *Listeria* Genus III. *Erysipelothrix*
Determinative Bacteriology Seventh (1957)	Order IV. Eubacteriales Family XII. Corynebacteriaceae Genus I. *Corynebacterium* Genus II. *Listeria* Genus III. *Erysipelothrix* Genus IV. *Microbacterium* Genus V. *Cellulomonas* Genus VI. *Arthrobacter*
Determinative Bacteriology Eighth (1974)	Part 17. Actinomycetes and related organisms Coryneform Group of Bacteria Genus I. *Corynebacterium* a. Human and animal parasites and pathogens b. Plant pathogenic corynebacteria c. Nonpathogenic corynebacteria Genus II. *Arthrobacter* Genera incertae sedis *Brevibacterium* *Microbacterium* Genus III. *Cellulomonas* Genus IV. *Kurthia*
Systematic Bacteriology Vol. 2 (1986)	Section 15. Irregular, nonsporing, gram-positive rods Genus *Corynebacterium* Plant pathogenic species of Corynebacterium Genus *Gardnerella* Genus *Arcanobacterium* Genus *Arthrobacter* Genus *Brevibacterium* Genus *Curtobacterium* Genus *Caseobacter* Genus *Microbacterium* Genus *Aureobacterium* Genus *Cellulomonas* Genus *Agromyces* Genus *Arachnia* Genus *Rothia* Genus *Propionibacterium* Genus *Eubacterium* Genus *Acetobacterium* Genus *Lachnospira* Genus *Butyrivibrio* Genus *Thermoanaerobacter* Genus *Actinomyces* Genus *Bifidobacterium*

Table 2. Taxonomic Position of the Genus *Corynebacterium* and Related Genera in *Bergey's Manual* (*continued*)

Determinative Bacteriology Ninth (1994)	Group 20. Irregular, nonsporing, gram-positive rods
	Genus *Acetobacterium*
	Genus *Acetogenium*
	Genus *Actinomyces*
	Genus *Acromicrobium*
	Genus *Agromyces*
	Genus *Arachnia*
	Genus *Arcanobacterium*
	Genus *Arthrobacter*
	Genus *Aureobacterium*
	Genus *Bifidobacterium*
	Genus *Brachybacterium*
	Genus *Brevibacterium*
	Genus *Butyrivibrio*
	Genus *Caseobacter*
	Genus *Cellulomonas*
	Genus *Clavibacter*
	Genus *Coriobacterium*
	Genus *Corynebacterium*
	Genus *Curtobacterium*
	Genus *Eubacterium*
	Genus *Exiguobacterium*
	Genus *Falcivibrio*
	Genus *Gardnerella*
	Genus *Jonesia*
	Genus *Lachnospira*
	Genus *Microbacterium*
	Genus *Mobiluncus*
	Genus *Pimerobacter*
	Genus *Propionibacterium*
	Genus *Rarobacter*
	Genus *Rothia*
	Genus *Rubrobacter*
	Genus *Sphaerobacter*
	Genus *Terrabacter*
	Genus *Thermoanaerobacter*

similar morphological characters, and it was not possible to clearly classify them. In the fifth edition of *Bergey's Manual of Determinative Bacteriology* (Table 2), the genus *Corynebacterium* belonged to the family Mycobacteriaceae of the order Actinomycetales (1939). However, in the sixth edition (1948), the family Corynebacteriaceae was created under the order Eubacterial, and this family included the genera *Corynebacterium*, *Listeria*, and *Erysipelothrix*. Furthermore, in the seventh edition (1957), the genera *Microbacterium*, *Cellulomonas*, and *Arthrobacter* were included in the family Corynebacteriaceae, in addition to the previously mentioned three genera.

The generic term *coryneform group of bacteria* first appeared in Part 17, "Actinomycetes and Related Organisms," in the eighth edition (1974), and the genera *Corynebacterium*, *Arthrobacter*, *Cellulomonas*, and tentatively, the genus *Kurthia* were included in this group. In addition, the genera *Brevibacterium* and *Microbacterium* were included, though as *incertae sedis*. Most of these genera were heterogeneous in composition. The genera *Listeria* and *Erysipelothrix*, which were not coryneform bacteria, were moved to the gram-positive, asporogenous, rod-shaped bacteria group in the eighth edition and treated as genera of uncertain affiliation. The classification and identification of coryneform bacteria have long been difficult and confusing. As a result, new isolated strains were identified in any genera as new species. However, with the application of modern taxonomic methods, such as cell wall analysis, nucleic acid base composition, and DNA relatedness, the relationships of coryneform bacteria and traditionally associated taxa had been clarified (Table 3). The definition of the genus *Corynebacterium* should now be restricted to those microorganisms that contain *meso*-A_2pm, acetylated muramic acid, mycolic acid with ca. 22 to 38 carbon atoms, and arabinose and galactose in the cell wall. In addition, these species have dihydrogenated menaquinones with either eight or nine isoprene units. In this respect, the genus

Table 3. Chemotaxonomic Characteristics of Coryneform Bacteria

	Peptidoglycan					
Genus	Diamino acid	Type	Major menaquinone	Mycolic acid	Acyl type	GC content (%)
Corynebacterium	meso-A_2pm	A1γ				
	L-Lys	A3α, A4α, B1α	8, 9,	+, −	Glycolyl	52–76
	D-Orn	B2β	10, 11, 12		Acetyl	
	L-DAB	B2γ				
Brevibacterium	meso-A_2pm	A1γ	7			
	LL-A_2pm	A3γ	8, 9, 10	+, −	Glycolyl	54–73
	L-Lys	A3α, A4α, B1β	11, 12		Acetyl	
	D-Orn	B2β				
Microbacterium	meso-A_2pm	A1γ				
	L-Lys	B1α, B1β	8, 9,	+, −	Glycolyl	58–73
	D-Orn	B2β	11, 12		Acetyl	
Cellulomonas	L-Orn	A4β	9	−	Acetyl	66–73
Arthrobacter	meso-A_2pm	A1γ				
	LL-A_2pm	A3γ	8, 9,	+, −	Glycolyl	54–76
	L-Lys	A3α, A4α	12, 13, 14		Acetyl	
	D-Orn	B2β				

Table 4. Reclassification of Coryneform Bacteria and Related Taxa

Peptidoglycan						
Diamino acid	Type	Acyl type	Major menaquinone	Mycolic acid	GC content (mol%)	Genera
meso-A_2pm	A1γ	Acetyl	9(H2), 8(H2)	C22–C38	51–65	*Corynebacterium*
meso-A_2pm	A1γ	Glycolyl	8(H2)	C34–C54	63–72	*Rhodococcus*
meso-A_2pm	A1γ	Glycolyl	9(H2)	C44–C60	60–65	*Gordona*
meso-A_2pm	A1γ	Glycolyl	8(H4)	C49–C69	64–72	*Nocardia*
meso-A_2pm	A1γ	Glycolyl	8(H2)	C60–C90	66–73	*Mycobacterium*
meso-A_2pm	A2γ		9	C64–C78	67–72	*Tsukamurella*
meso-A_2pm	A1γ	Acetyl	8(H2), 7(H2)	—	63–66	*Brevibacterium*
LL-A_2pm	A3γ	Acetyl	8(H4)	—	68–73	*Pimelobacter*
LL-A_2pm	A3γ	Acetyl	8(H4)	—	66–68	*Nocardioides*
LL-A_2pm	A3γ		8(H4)	—	72	*Terrabacter*
L-Lys	A3α, A4α	Acetyl	9(H2), 9, 8	—	59–69	*Arthrobacter*
L-Lys	B1α, B1β	Glycolyl	11, 12	—	69–75	*Microbacterium*
L-Lys	A3α	Acetyl	7	—	53–56	*Exigunobacterium*
L-Lys	A3α	Acetyl	8	—	67	*Rubrobacter*
L-Lys	A3α	Acetyl	9(H4)	—	52	*Arcanobacterium*
L-Orn	A4β	Acetyl	9(H4)	—	71–76	*Cellulomonas*
D-Orn	B2β	Acetyl	9	—	68–73	*Curtobacterium*
D-Orn	B2β	Glycolyl	11, 12, 13	—	66–70	*Aureobacterium*
L-DAB	B2γ	Acetyl	9, 10	—	63–78	*Clavibacter*
L-DAB	B2γ		10	—	60–63	*Rathayibacter*

Table 5. G+C contents of DNA and DNA–DNA Relatedness between *Corynebacterium glutamicum* ATCC13032^T and Other Glutamate Producing Bacteria

Strain	G+C content (mol%)	DNA relatedness to KY9002^T (%)		
		Ref. 10	Ref. 11	Ref. 12
Corynebacterium glutamicum ATCC13032^T	53.4	100	100	100
Corynebacterium lilium ATCC15990^T	52.9		93.3	
Brevibacterium divaricatum ATCC14020^T	54.4	93.4	77.9	
"*Brevibacterium flavum*" ATCC14067	54.6		87.2	
"*Brevibacterium lactofermentum*" ATCC13665	54.4	102.1	82.7	
"*Brevibacterium immariophilum*" ATCC14068	54.6	76.2		105
"*Brevibacterium roseum*" ATCC13825	54.6	73.5		108
"*Brevibacterium saccharolyticum*" ATCC14066	54.3	77.3		109
"*Brevibacterium ammoniagenes*" ATCC13745	nd	75.6		
"*Corynebacterium acetoacidophilum*" ATCC13870	52.9	76.2		87
"*Brevibacterium thiogenitalis*" ATCC19240	54.3			106
"*Corynebacterium herculis*" ATCC13868	54.2			96
"*Microbacterium ammoniaphilum*" ATCC15354	54.4	78.4		
Arthrobacter sp. NCIB942	nd			95
Brevibacterium sp. ATCC14902	nd			93
Microbacterium sp. ATCC 15283	nd			104
Corynebacterium callunae ATCC15991^T	51.2		36.7	47
Corynebacterium ammoniagenes ATCC6871^T	54.6			17

Brevibacterium was distinguished from the genus *Corynebacterium* in that it does not contain mycolic acid and polysaccharide components such as arabinose and galactose. Furthermore, the genera *Arthrobacter* and *Microbacterium* were defined as a group of microorganisms that are additionally different in the diamino acid and major menaquinone (Table 4). As a result, the genus *Corynebacterium*, defined in chemotaxonomic characteristics, includes most animal pathogenic and parasitic species belonging to section a, and some saprophytic species (e.g., *Corynebacterium glutamicum*) belonging to section c in the eighth edition. On the other hand, all plant pathogens and many saprophytic species belonging to section b were excluded from this genus. *Bergey's Manual* was newly started in 1986 with a new title, *Bergey's Manual of Systematic Bacteriology*, in which the genus *Corynebacterium* and the related genera are included in section 15, and, consequently, the term *coryneform bacteria* disappeared from the index.

TAXONOMIC POSITION OF GLUTAMIC ACID PRODUCING BACTERIA

The strain isolated as accumulating a remarkable amount of glutamic acid was initially named *Micrococcus glutam-*

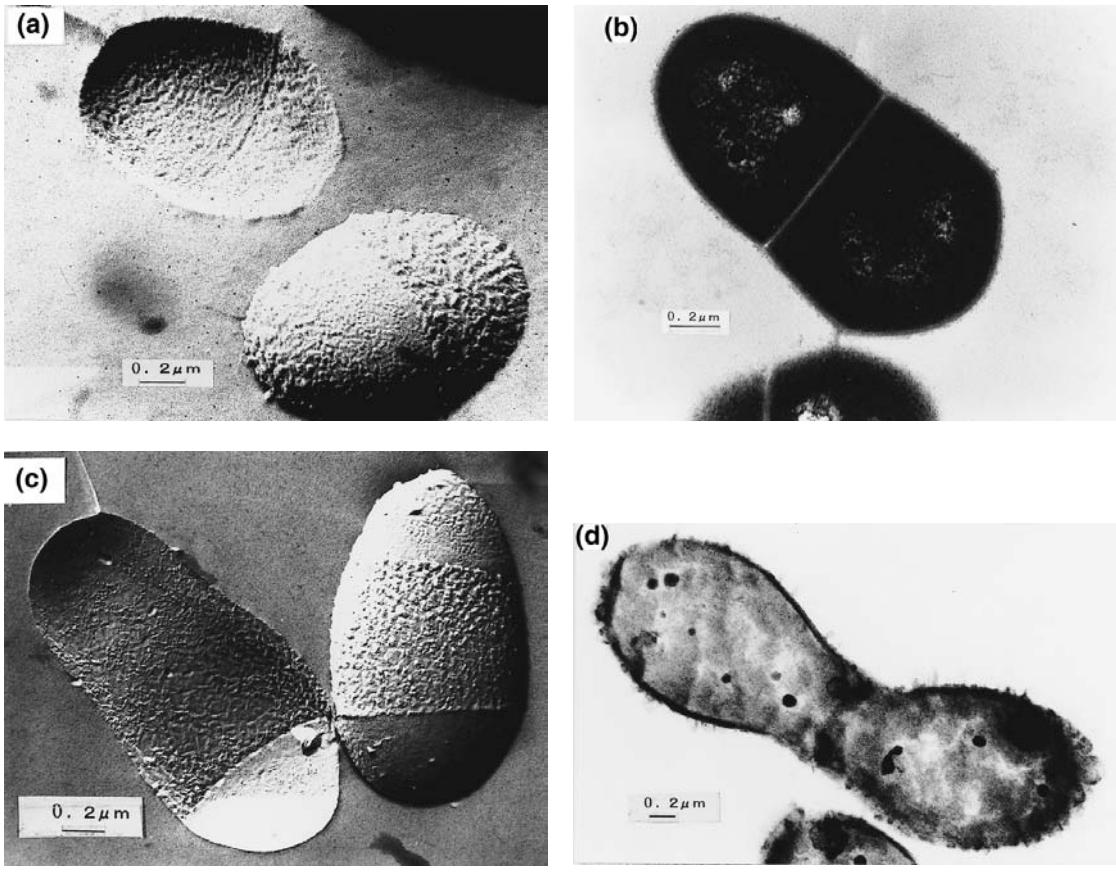


Figure 2. Electron micrographs of *Corynebacterium glutamicum* ATCC 13032^T in biotin-sufficient and biotin-deficient cultivation: (**a**) freeze-fracturing electron micrograph of the biotin-sufficient cells, (**b**) cross-sectioning electron micrograph of the biotin-sufficient cells, (**c**) freeze-fracturing electron micrograph of the biotin-deficient cells, (**d**) cross-sectioning electron micrograph of the biotin-deficient cells.

icus nov. sp. by Kinoshita et al. (1), with ATCC 13032 as a type strain. This strain was later revealed to belong to the genus *Corynebacterium* based on chemotaxonomic characteristics and reclassified to *C. glutamicum* by Abe et al. (6). On the other hand, many other glutamic acid producing species, patent strains, reported as different species than *C. glutamicum*, were found to be so-called coryneform bacteria. They were also found to have very similar chemotaxonomic properties based on studies by Keddie (7), Bousfield (8), and Yamada and Komagata (9), and were revealed to be bacteria belonging to *Corynebacterium* in the strict sense. Furthermore, as shown in Table 5, Komatsu and Kaneko (10) showed in 1980 that eight strains, *C. acetoacidophilum* ATCC 13870, *Brevibacterium lactofermentum* ATCC 13655, *B. divaricatum* 14020^T, *B. saccharolyticum* ATCC 14066, *B. roseum* ATCC 13825, *B. immariophilum* ATCC 14068, *B. ammoniagenes* ATCC 13745, and *Microbacterium ammoniaphilum* ATCC 15354, exhibited 73 to 102% DNA binding to the type strain of *C. glutamicum*. However, as no measures had been taken to make taxonomical changes for these strains, they were subjected to studies using these names. In 1991, Liebl et al. (11) showed that five strains, *B. divaricatum* 14020^T, *B. flavum* ATCC 14067, *B. lactofermentum* ATCC 13665 and ATCC 13869, and *C. lilium* ATCC 15991^T were closely related to the type strain of *C. glutamicum* based on the high levels of DNA relatedness, and were thus transferred to the same species as *C. glutamicum*. *C. acetoacidophilum*, *C. acetoglutamicum*, and *C. acetophilum* have been reported to produce remarkable amounts of glutamic acid from acetic acid as a carbon source other than carbohydrate. While *C. acetoacidophilum* ATCC 13870 was confirmed to be genomically similar to *C. glutamicum* (Table 5), the taxonomic positions of the other two strains are unknown. Furthermore, hydrocarbon-assimilating bacteria *Corynebacterium hydrocarboclastus* ATCC 15592, which produce a remarkable amount of glutamic acid from hydrocarbons, exhibited more than 70% DNA binding to the type strain of *Rhodococcus erythropolis* (not shown).

MORPHOLOGICAL AND PHYSIOLOGICAL CHANGES OF GLUTAMIC ACID PRODUCING BACTERIA

C. glutamicum has some characteristic features, including that it undergoes V-formation (Fig. 1) by snapping division

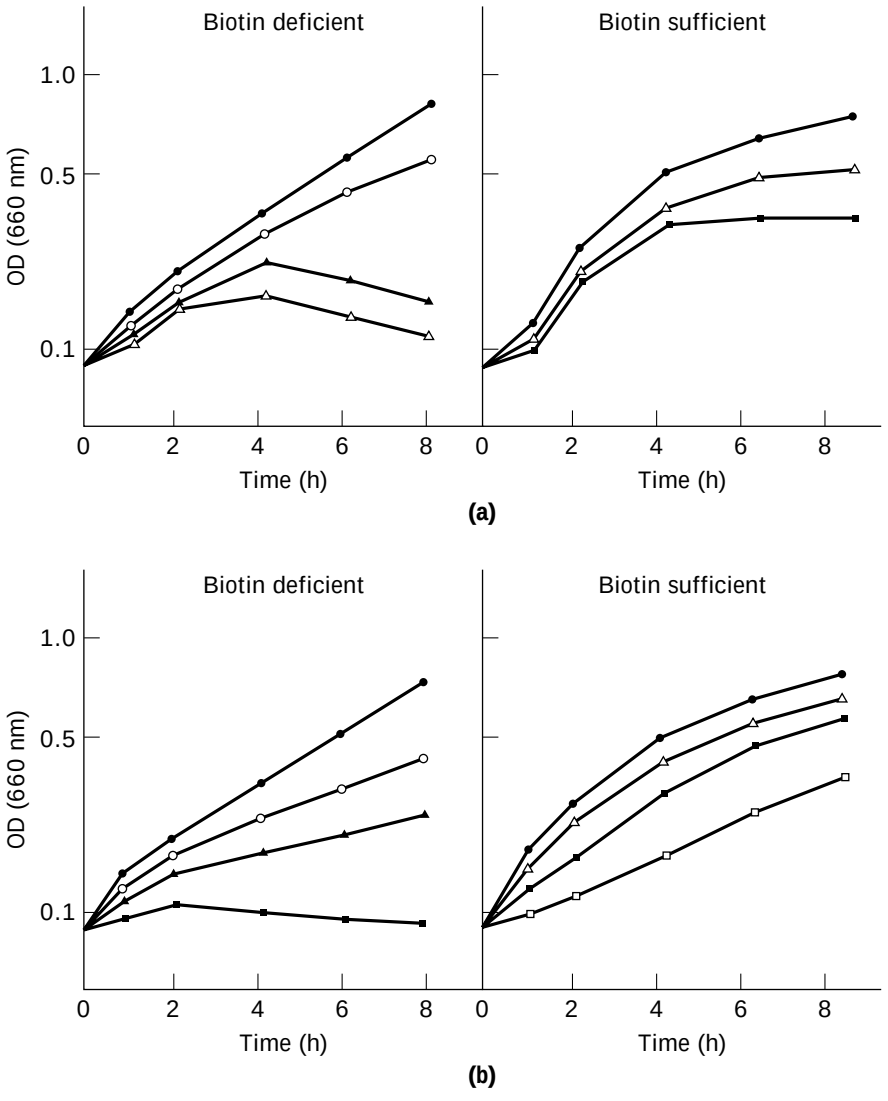


Figure 3. Effect of lysozyme and SDS on growth of *Corynebacterium glutamicum* ATCC 13032^T in biotin-sufficient and biotin-deficient cultivation: (**a**) Effect of lysozyme on the growth of *C. glutamicum* ATCC13032. Cells were inoculated into media supplemented with 0 μg/mL (●), 12.5 μg/mL (○), 25 μg/mL (▲), 50 μg/mL (△), and 100 μg/mL (□) of lysozyme; (**b**) effect of SDS on the growth of *C. glutamicum* ATCC13032. Cells were inoculated into media supplemented with 0 μg/mL (●), 10 μg/mL (○), 20 μg/mL (▲), 30 μg/mL (△), 40 μg/mL (■), and 70 μg/mL (□) of SDS.

upon cell division, and requires biotin for growth. However, glutamic acid production was observed only in a very narrow range of biotin content, that is, at suboptimal concentration (2.5–5.0 μg/mL) for cell growth. In 1962, Somerson et al. (13) found that the addition of penicillin was very effective, even in the presence of excessive amounts of biotin. Therefore, it is widely known that in order to accumulate a remarkable amount of glutamic acid extracellularly, low levels of biotin concentration, or penicillin treatment at high levels of biotin concentration, are required. However, the detail of their mechanisms is not known, although there have been reports attributing the accumulation of glutamic acid due to the change in the membrane permeability. At a low concentration of biotin, cells swelled and elongated. According to electron microscopic observation, outer layers changed grossly and lost ordered structures, and many blebs were observed at low concentrations of biotin (Fig. 2). It was observed that glutamic acid producing bacteria have an increased sensitivity to lyzozyme or SDS treatment when cultured at low concentrations of biotin (Fig. 3).

BIBLIOGRAPHY

1. S. Kinoshita, K. Nakayama, and S. Akita, *Bull. Agric. Chem. S. Jpn.* **22**, 176 (1958).
2. K.B. Lehmann, and R. Neumann, *Atlas und Grundriss der Bakteriologie unt Lehrbuch der speciellen bakteriologischen Diagnostik*, 1st ed., J. F. Lehmann, Munchen, 1896.
3. H.L. Jensen, *Annu. Rev. Microbiol.* **6**, 77–90 (1952).
4. J. Oerskov, *Investigations into the Morphology of the Ray Fungi*, Levin and Munksgaard, Copenhagen, 1923.
5. *Bergey's Manual of Determinative Bacteriology*, Williams & Wilkins, Baltimore, Md., 1974.
6. S. Abe, K. Takayama, and S. Kinoshita, *J. Gen. Appl. Microbiol.* **13**, 279–301 (1967).
7. R.M. Keddie and G.L. Cure, *J. Appl. Bacteriol.* **42**, 229–252 (1977).
8. I.J. Bousfield, *J. Gen. Microbiol.* **71**, 441–455 (1972).
9. K. Yamada and K. Komagata, *J. Gen. Appl. Microbiol.* **18**, 399–416 (1972).
10. Y. Komatsu and T. Kaneko, *Rep. Ferment. Res. Inst.* **55**, 1–5 (1980).

11. W. Lieble, M. Ehmann, W. Ludwig, and K.H. Schleifer, *Int. J. Syst. Bacteriol.* **41**, 255–260 (1991).
12. K. Ochiai, M. Suzuki, and K. Ando, *Nippon Nogeikagaku Kaishi* **72**, suppl. 2A2, p. 6 (1998).
13. U.S. Pat. 3,080,297 (March 5, 1963), N.L. Somerson and T. Phillips (to Merck & Co., Inc.).

See also AMINO ACIDS, PRODUCTION PROCESSES; ENERGY METABOLISM, MICROBIAL AND ANIMAL CELLS; FERMENTATION MONITORING, DESIGN AND OPTIMIZATION; METABOLITES, PRIMARY AND SECONDARY.

GLYCOSYLATION OF RECOMBINANT PROTEINS

DAVID C. JAMES
KYM N. BAKER
University of Kent
Canterbury, U.K.

KEY WORDS

Glycan
Glycosylation
Mass spectrometry
Oligosaccharide
Posttranslational modification
Recombinant protein

OUTLINE

INTRODUCTION

Recombinant proteins produced by animal cell expression systems can be subject to numerous posttranslational modifications that can result in a structurally heterogeneous product. Glycosylation is the most extensive of all posttranslational modifications, and for proteins destined for therapeutic use, this source of molecular diversity may be of particular significance. For example, distinct protein glycoforms may possess different properties such as biological activity, receptor binding, stability, immunogenicity, and clearance rate in vivo (1). The carbohydrate components of glycoproteins can also play crucial roles in protein folding, oligomer assembly, and secretion processes. Recombinant protein heterogeneity therefore represents an important issue for the biopharmaceutical industry and regulatory authorities, with regard to product characterization and definition, consistency, potency, and purity (2–4). Authorities such as the U.S. Food and Drug Administration are now demanding increasingly sophisticated carbohydrate analyses as part of the "well-characterized" product or process validation. Therefore, within the biotechnology/biopharmaceutical industries there is now considerable interest in technologies enabling recombinant glycoprotein analysis and control of glycoprotein heterogeneity.

Although there has been much progress in the field of glycobiology in recent years, with a plethora of published information on biological properties of carbohydrates (5,6), methods of glycan analysis (7–10), cellular biosynthetic pathways, and the molecular cell biology of the key enzymes involved (11–18), this article highlights key bioprocess factors that may influence the glycosylation of a recombinant protein produced by an animal expression system.

Oligosaccharide Structures Present on Recombinant Glycoproteins

Although there is considerable diversity in the types of protein–carbohydrate conjugates that occur naturally, recombinant glycoproteins produced by eukaryotic expression systems are mainly associated with N-linked glycans (N-glycosidic bond to the amide group of an asparagine residue within the consensus sequence Asn-X-Ser/Thr) or O-linked glycans (O-glycosidic bond to the hydroxyl side group of either a serine or threonine residue). Carbohydrates are also components of the glycophosphatidylinositol (GPI) anchor used to secure some proteins to cell membranes.

The molecular diversity of recombinant glycoproteins arises from two major sources: variable glycosylation site occupancy, also called macroheterogeneity (i.e., the presence or absence of glycans), and microheterogeneity (variation in carbohydrate structures at individual glycosylation sites). In the case of the former, the presence of sequence motifs or amino acids necessary for glycosylation within the polypeptide sequence does not guarantee their glycosylation. For example, at the cellular level N-glycosylation site occupancy at Asn-X-Ser/Thr sequons may be related to several factors including oligosacchar-

yldolichol substrate availability, oligosaccharyl transferase activity, rate of protein synthesis, competition with protein folding, and other cotranslational events (19,20). Protein structural or polypeptide sequence factors may also be significant; for example, the target glycosylation sequon Asn-X-Thr has been found to be 40 times more susceptible to glycosylation by oligosaccharyltransferase than the Asn-X-Ser sequence (21).

Recombinant glycoproteins should most appropriately be considered as populations of individual "glycoforms," as illustrated for recombinant IFN-γ produced by Chinese hamster ovary (CHO) cells in Figure 1. In general terms, the relative abundance and distribution of protein glycoforms within a glycoprotein population are a function of the following:

1. The host cell and expression system employed
2. The interaction of the host cell and recombinant glycoprotein product with the bioprocess environment

N-Linked Glycans

All mature N-linked glycans contain a common trimannosyl core sugar structure, a pentasaccharide consisting of three mannose (Man) residues and two *N*-acetylglucosamine (GlcNAc) residues (Manα1-6(Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAc-(Asn)). This can form part of simple oligomannose structures or be extensively modified by the addition of other GlcNAc, galactose (Gal), fucose (Fuc), and sialic acid (*N*-acetyl- or *N*-glycolylneuraminic acids; NeuAc or NeuGc) residues. N-linked chains can be classed into three main subgroups: high-mannose type, complex type, and hybrid type (Fig. 2). High-mannose-type glycans contain α-linked mannose chains bound to both arms of the trimannosyl core. Complex-type chains contain no extra mannose residues and have α-linked GlcNAc residues bound to the trimannosyl core, with Gal and NeuAc residues added subsequently. Commonly, there may be α1-6- or α1-3-linked Fuc residues added to the core (Asn-linked) GlcNAc. Hybrid-type chains contain both α-linked Man residues bound to the α1-6-linked Man arm of the trimannosyl core, while complex-type chains are bound to the α1-3-linked Man arm. However, numerous possible variations are commonly found. Following are the basic sources of *N*-glycan heterogeneity:

1. The degree of branching of complex type chains (number of antennae). For example, the complex chain illustrated in Figure 2 contains two branching chains from the trimannosyl core and is thus termed a biantennary-complex-type chain. Typically, complex N-linked glycans may be binantennary, triantennary, or tetraantennary.

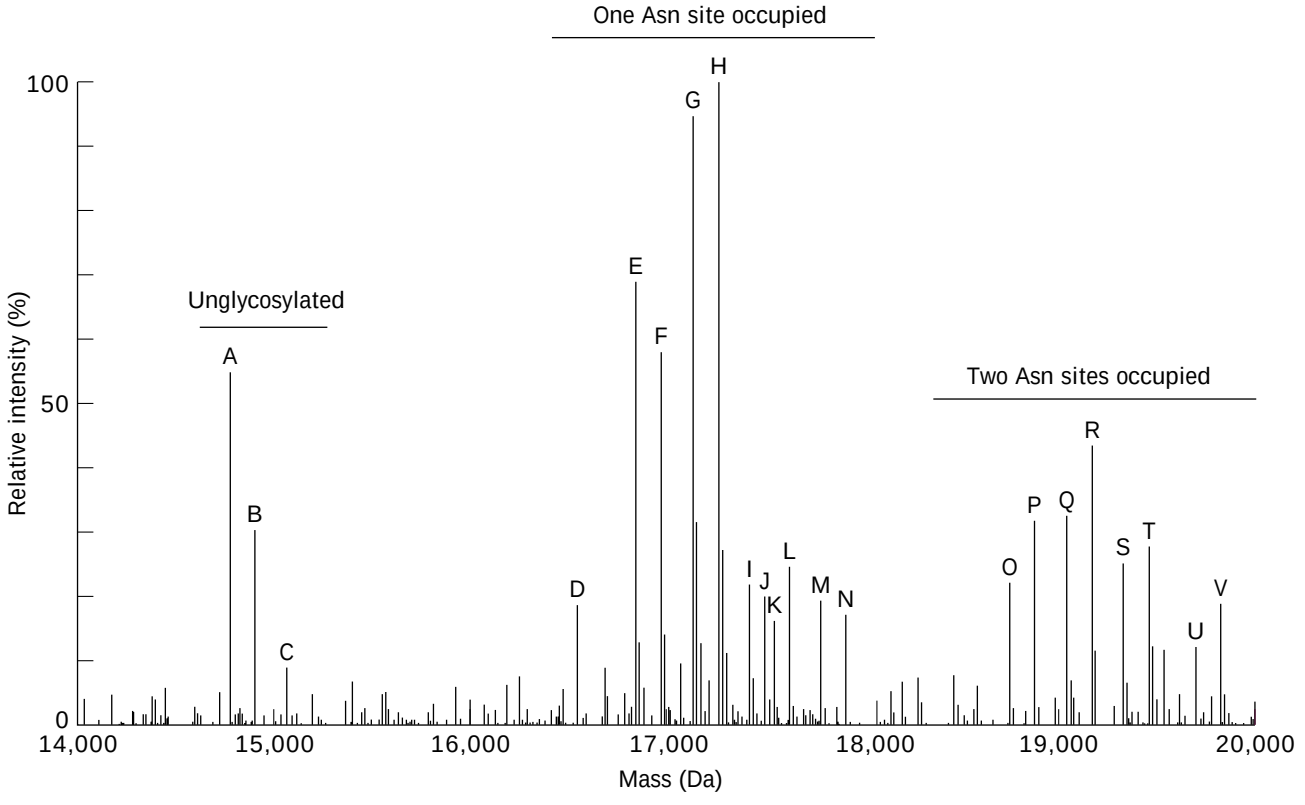


Figure 1. The heterogeneity of recombinant human IFN-γ produced by Chinese hamster ovary cells. Each component of the deconvoluted electrospray mass spectrum represents a different IFN-γ protein varying in number of glycosylation sites occupied, structure of *N*-glycans at individual sites, and C-terminal proteolytic clipping. *Source*: From Ref. 22.

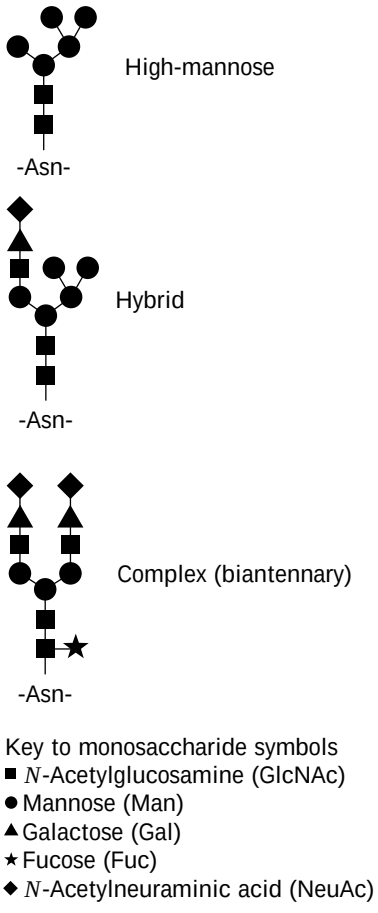


Figure 2. Common *N*-glycan structures found on glycoproteins.

2. The presence or absence of terminal sialic acid residues linked to antennae.
3. Variation in the type of sialic acid; N-acetyl and N-glycolyl variants are the most common.
4. The presence or absence of a 'bisecting' GlcNAc attached to the trimannosyl core β-linked mannose (which prevents branching)
5. The presence or absence of terminal poly-*N*-acetyllactosamine repeats (Galβ1-4GlcNAcβ1-3-).
6. Unusually, sulfation at GlcNAc residues or phosphorylation at Man residues.

O-Linked Glycans

O-linked glycosylation is extensive in structural proteins such as proteoglycans. Although the majority of O-linked glycans are bound to either a serine or threonine residue core, small glycan structures can also be O-linked to the side chain of hydroxylysine or hydroxyproline. *O*-Glycans may occur as a single *N*-acetylgalactosamine (GalNAc) residue, disaccharides NeuAcα2-6GalNAc/Galβ1-3GalNAc, or larger glycans of up to 10 residues attached to the core GalNAc containing Gal, Fuc, and GlcNAc residues (Fig. 3).

Glycan Biosynthesis

The synthesis and elongation of glycan structures is an ordered, rather than random, process that begins in the rough endoplasmic reticulum and is completed in the *trans*-Golgi network. In the case of *N*-glycans, efficient transport of recombinant glycoprotein through the constitutive secretory pathway of eukaryotic cell expression systems is a prerequisite for high-level production of recombinant protein and modification of N-linked oligosaccharides. In the case of the latter, the ordered cellular localization and activity of glycosidase and glycosyltransferase enzymes define the sequence and structure of *N*-glycans associated with the secreted protein (11,14–16,23–25). This dynamic process is mediated by the trafficking of secretory vesicles between distinct subcellular compartments and by molecular interactions within those compartments (26).

N-glycosylation may itself be necessary for efficient secretion of some proteins (27,28). Recent studies have shown that secreted or membrane-targeted proteins become concentrated in an intermediate compartment (IC) between the endoplasmic reticulum (ER) and the *cis*-Golgi (29,31). Although the nature of this compartment remains obscure at present and may not be solely a vesicular transport system but a microtubule-directed system (32), distinct molecular markers of the IC, such as the membrane-bound mannose-binding lectin ERGIC53 (28,33,34) or homologs such as p58 (35), which cycles between the ER, IC, and *cis*-Golgi (36), may function to specifically regulate the sorting of glycoproteins.

As shown in Figure 4, biosynthesis of N-linked glycans begins with the cotranslational en bloc transfer of a large, high-mannose oligosaccharide terminating in glucose (Glc) residues ($Glc_3Man_9GlcNAc_2$) from a dolichol pyrophosphate–linked precursor to an appropriate Asn residue in the nascent polypeptide chain. This process, catalyzed by the oligosaccharyltransferase (OST) complex, occurs in the lumen of the ER. This is followed by the subsequent removal of the three terminal glucose residues via an α-glucosidase (steps 1–3, Fig. 4), a quality-control process that has been shown to play an important part in the correct folding and assembly of some secreted proteins through interactions with the molecular chaperones calnexin and calreticulin (37,38).

Subsequent trimming of mannose residues by ER and Golgi-specific α-mannosidases is followed by addition of terminal GlcNAc, Gal, and sialic acid and core fuc residues by glycosyltransferases utilizing nucleotide-sugar substrates imported into the lumen of Golgi compartments by specific transporters (antiporters) (39). Control of glycosylation reactions is complex and relies upon a multiplicity of molecular, cellular, and physiological interactions such as relative distribution of enzyme activities, substrate availability, and rate of protein transport (25). In some circumstances, local protein structure may serve to modify glycan processing. For example, it is well documented that terminal galactosylation and sialylation of IgG *N*-glycans in the Fc region at Asn_{297} is restricted by steric hindrance (40,41). However, in general, key control points that contribute to *N*-glycan heterogeneity are those catalyzed by the following:

1. ER and Golgi mannosidase I, GlcNAc transferase I (conversion of high-mannose to complex-type glycans)

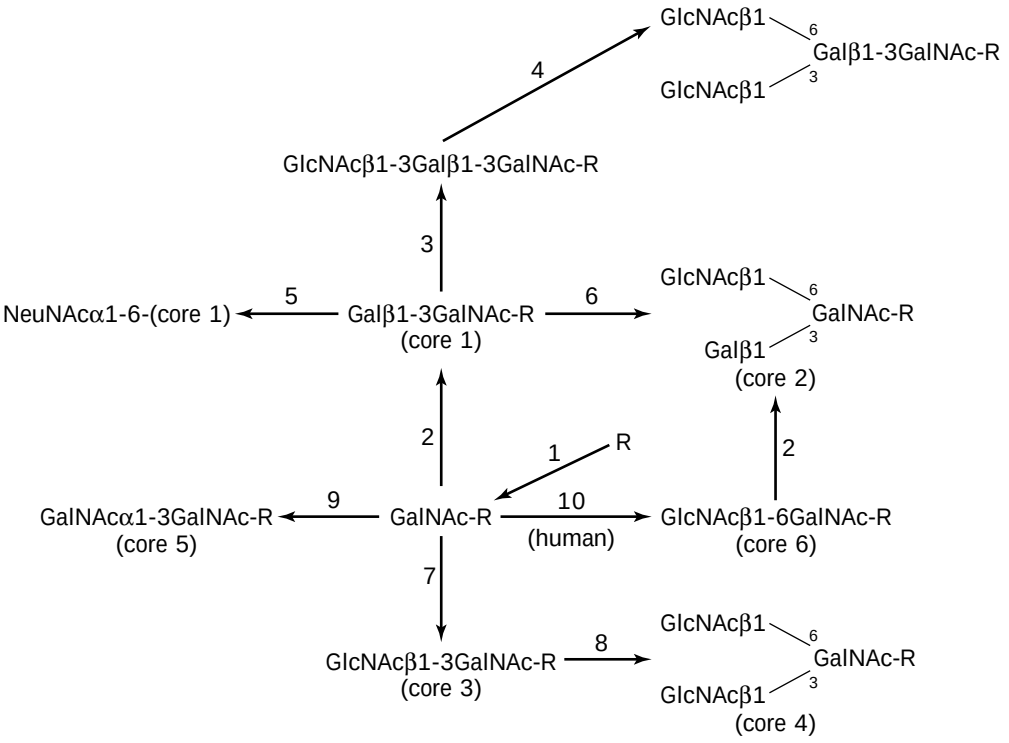


Figure 3. Biosynthesis of *O*-glycans. R represents Ser or Thr.

2. GlcNAc transferases II, IV, V (degree of branching of *N*-glycans)
3. GlcNAc transferase III (bisecting GlcNAc)
4. α-2,3 and α-2,6 sialyltransferases (degree of sialylation and linkage)
5. CMP-NeuAc-4-hydroxylase (conversion of NeuAc to NeuGc)

Mathematical modeling of this complex process has recently been attempted to permit comparison of strategies to obtain specific protein glycoform distributions (20,42).

The biosynthesis of O-linked glycans is similar to that of *N*-glycans and undergoes similar competition and regulation of its glycosyltransferases. It has been well reviewed by Schachter and Brockhausen (43), and four of the basic core structures and known biosynthetic pathways are outlined in Figure 3. Although many of the glycosyltransferases involved in *N*-glycan processing have been identified and cloned, most of the *O*-glycan transferases, with the exception of the core 2 GlcNAc-transferase, have not yet been cloned. This is in part due to the low percentage of glycans present on recombinant proteins that are O-linked.

GLYCOPROTEIN ANALYSIS

Underpinning our ability to control protein glycosylation is the necessity to monitor and in some cases completely characterize the recombinant glycoprotein product. More routine profiling of glycoform populations may also be useful to ensure reproducible product quality, batch-to-batch consistency, and product stability. Recent advances in analytical technology now permit relatively rapid and detailed assessment of glycoprotein and glycan structure by a variety of different methodologies, yielding information of varying complexity (9,44).

The choice of analytical strategy employed for characterization of a recombinant glycoprotein will depend on factors such as level of structural information required (e.g., site-specific glycosylation, oligosaccharide structure or monosaccharide composition, etc.), sample composition, and critically, the amount of protein available for analysis. Although some techniques, such as mass spectrometry, are now proving indispensable for many applications, it is a combination of complimentary techniques that will often provide the total analytical solution (45,46). The principal techniques currently employed are highlighted in the following sections.

Mass Spectrometry

Increasingly prevalent among current analytical strategies are mass spectrometric methods. Since the early 1980s, the emergence of "soft" ionization methods for biopolymer characterization by mass spectrometry (MS), such as fast atom bombardment (FAB), electrospray ionization (ESI), and matrix-assisted laser desorption/ionization (MALDI), has provided a formidable weapon in the armory of the analytical biochemist. There are several excellent reviews describing the application of this technology for protein and peptide analysis in detail (7,47–49).

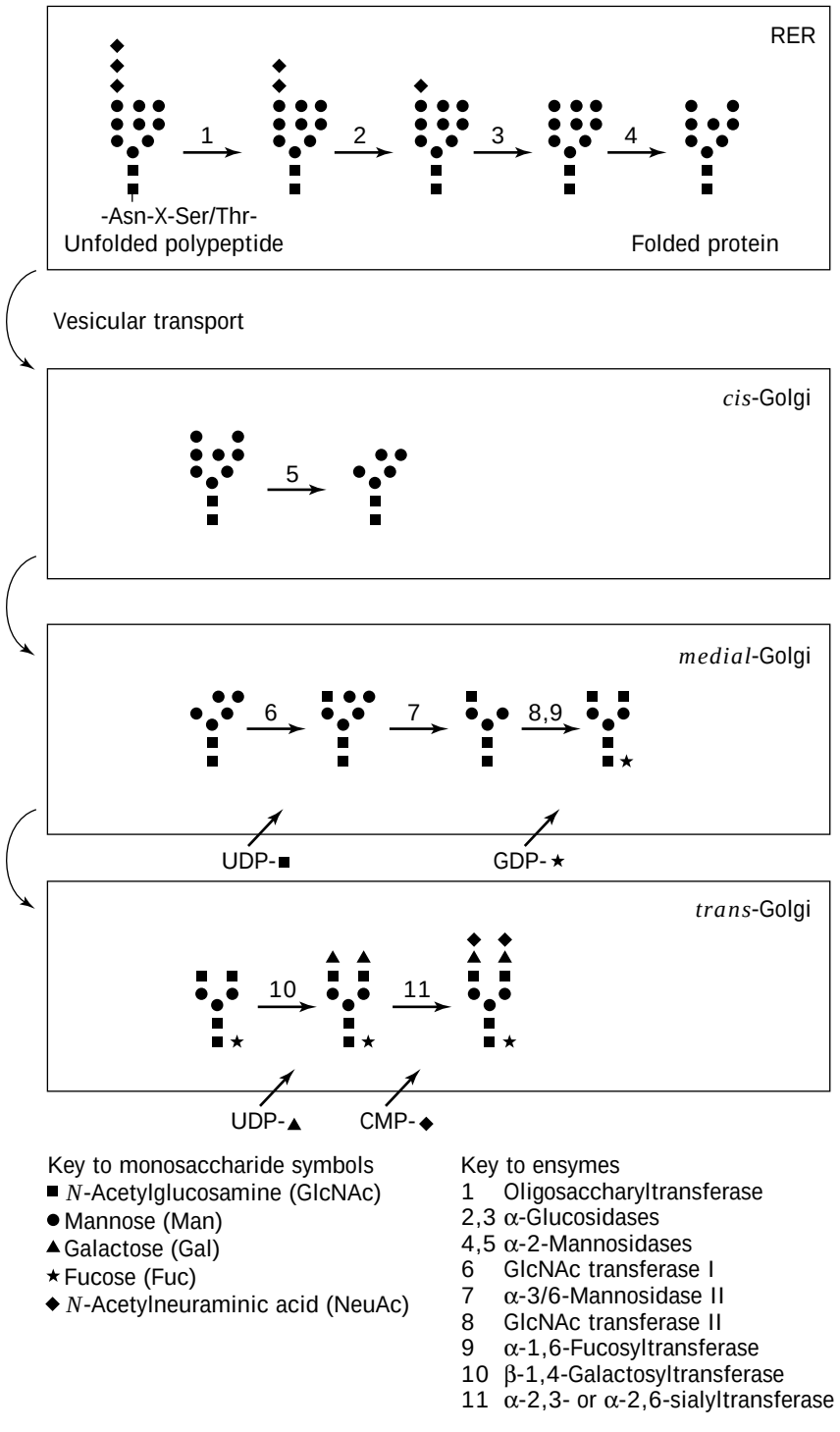


Figure 4. Biosynthesis of *N*-glycans. Schematic diagram illustrating the major enzymatic steps leading to the synthesis of a typical complex biantennary glycan.

Fast Atom Bombardment Mass Spectroscopy. FAB-MS and its variant, liquid secondary ion MS (LSI-MS) have an upper mass limit of ~15 and ~25 kDa, respectively, and thus can be used only to directly characterize smaller glycoconjugates such as glycopeptides (50,51). Because molecules tend to fragment in the ion source, this technique has been employed successfully to provide detailed sequence, linkage, and branching pattern information for N-linked glycans (52); however, derivatization of the carbohydrates by permethylation, peracetylation, or periodate oxidation is frequently necessary to obtain strong signals and provide linkage information. This technique is not generally amenable to rapid analysis of mixtures of glycopeptides or oligosaccharides, and successful operation of instrumentation and interpretation of spectra requires substantial specialist training. It is reviewed in detail by Peter-Katalinic (53).

Matrix-Assisted Laser Desorption/Ionization Mass Spectroscopy. MALDI-MS is a relatively simple technique. Biopolymers (theoretically up to 300 kDa) are cocrystallized on a metal surface in the presence of a molar excess (typically >5000:1) of a low molecular weight, strongly UV absorbing matrix. Biopolymer and matrix ions are desorbed by a short laser pulses (typically at 337 nm, 5 ns), and the gas phase ions are electrostatically directed into a time-of-flight (TOF) mass analyzer. Analyte resolution (width of a peak at half height divided by the mass/charge (m/z) ratio) using MALDI TOF-MS (typically <500) is $\pm 0.1\%$ for continuous ion extraction MALDI or >0.01% for MALDI with delayed ion extraction (54). This technique can be used to identify and characterize heterogeneous samples of glycopeptides (55–57), and released oligosaccharides (enzymatically or chemically) (58–61), and for the simultaneous sequencing of oligosaccharides in conjunction with exoglycosidase digestion (62,63). Recent reports highlight the use of MALDI-MS to rapidly monitor the site-specific glycosylation of a recombinant glycoprotein by interfacing it with HPLC-based peptide mapping (64). There are few examples of MALDI-MS glycoform profiling using whole glycoproteins, a result of the extensive heterogeneity of natural and recombinant glycoproteins. However, glycoproteins of less than 20 kDa, with perhaps only a single occupied glycosylation site, may be amenable to more detailed analysis. For example, variably sialylated forms of CHO-cell-derived recombinant interleukin 4 (~17 kDa) have been observed by Tsarbopoulos et al. (65).

Electrospray Ionization Mass Spectroscopy. ESI, typically in combination with a single- or triple-quadrupole mass analyzer, is capable of determining the molecular weight of whole recombinant glycoproteins or glycopeptides up to 100 kDa, with a mass accuracy of $\pm 0.01\%$ at a resolution routinely in excess of 1,000. Ionization of biopolymers in volatile solvents results in the production of multiply charged molecular ions (averaging 1 proton/kDa), and conditions are sufficiently mild that noncovalent interactions can often be observed because covalent bonds are not disrupted. The resulting mass spectrum of multiply charged molecular ions is then "deconvoluted" to yield a single zero-charge peak. The advantage of ESI-MS is that it may be coupled on-line with liquid chromatography (LC) or capillary electrophoresis (CE) to analyze complex samples such as proteolytic digests at low picomole to femtomole sensitivity.

The increased resolution and softer ionization afforded by ESI-MS makes this technique the most attractive method for direct analysis of recombinant glycoprotein populations because individual glycoforms can often be resolved. For example, heterogeneous glycoform populations of recombinant human interferon γ have been analyzed by James et al. (22) (Fig. 1). ESI-MS interfaced with LC has proved to be the method of choice for direct, sensitive analysis of protein digests. Selective, on-line identification of glycopeptides within these complex mixtures can be achieved by the use of operating conditions that enhance the formation of oxonium fragment ions (e.g., $HexNAc^+ = 204$ Da), such as collision-induced dissociation (66). Isolated glycopeptides may be analyzed in detail by ESI-MS directly (67,68) or in combination with exoglycosidase digestions to provide a more extensive site-specific structural analysis of N- and O-linked oligosaccharides (69). Although ESI-MS may be interfaced with CE for analysis of glycoproteins and glycopeptides, some CE separation protocols rely on acidic buffers, which are compatible with ESI-MS but unsuitable for separation of sialylated glycoconjugates (70).

Methods for glycoconjugate analysis by MALDI-MS and ESI-MS have been reviewed in detail by Burlingame (8).

HPLC

Liquid chromatographic techniques have been extensively employed to separate underivatized and derivatized oligosaccharides released enzymatically or chemically (e.g., by hydrazinolysis [71]) from glycoproteins. These approaches can most conveniently be described as being based on oligosaccharide size or charge.

Low-resolution separations based on hydrodynamic volume, such as those employing gel permmeation chromatography (e.g., Bio-Gel P4), were utilized in combination with limited exoglycosidase digestion of oligosaccharides to provide structural information (72). However, this approach has largely been superceded by the development of high-resolution normal-phase separations using amide-based media, in combination with simple procedures to fluorescently label oligosaccharides at the reducing terminus (73,74). This technology may also be used in combination with exoglycosidase sequencing and for analysis of glycoproteins in polyacrylamide gels (63,75).

Charge-based separations, distinguishing between differently sialylated glycans, have typically involved resolution of borohydride-reduced oligosaccharides by high-pH anion exchange chromatography with pulsed amperometric detection (HPAE-PAD) (76), although effective separation of fluorophore-derivatized glycans has also been demonstrated with amine-based columns now available commercially (73).

Capillary Electrophoresis

CE is proving itself to be a versatile analytical tool for (glyco)protein analysis both by itself (77,78) and in combination with other analytical techniques such as mass spectrometry (70,79). There are now several examples of glycoprotein analysis by CE, from analysis of intact glycoproteins (80–83), glycopeptides (84) and released oligosaccharides (85–87), to individual monosaccharides (88). These examples employ a variety of separation mechanisms (e.g., capillary zone electrophoresis, micellar electrokinetic capillary chromatography, capillary isoelectric-focusing) designed to enhance the resolution of the specific analyte(s) in question.

Gel Electrophoresis of Carbohydrates

Separation of released oligosaccharides in polyacrylamide gels after derivatization with fluorescent labels at the reducing terminus provides a simple, inexpensive method for profiling oligosaccharide populations. Limited structural information is obtained in combination with exoglycosidase digestions (89,90).

CONTROLLING GLYCOSYLATION

Given the ability to detect and analyze glycoproteins in detail, how can recombinant protein glycosylation be controlled? There are five main options currently available:

1. Choice of host cell
2. Genetic engineering of the host cell
3. Genetic engineering of the recombinant glycoprotein
4. Control of the host cell environment
5. Remodeling of protein glycans in vitro

Choice of Host Cell

Different host cell types are known to vary in their ability to glycosylate recombinant proteins. As there is now a profusion of expression systems available (mammalian, insect, or yeast cells in culture; transgenic animals and plants), initial choice of expression system is crucial and will, to a large extent, govern the type of glycosylation pattern obtained. Common bacterial expression systems such as *Escherichia coli* have no capacity to glycosylate proteins in either N- or O-linked conformations.

Yeast. Most yeast strains tend to hypermannosylate (the addition of a large number of mannose residues to the core trimannosyl oligosaccharide) recombinant proteins. We may infer that this modification would speed clearance of the glycoprotein from the bloodstream by hepatic and macrophage mannose receptors (91,92), and it has been shown to reduce the efficacy of recombinant hepatitis B vaccine (93). It is possible to use mutant yeast strains (e.g. mnn-9 or the temperature-sensitive ngd-29) that do not hypermannosylate recombinant proteins. In these strains, N-glycosylation is confined to core oligosaccharide residues containing up to eight mannose residues (93,94).

Insect Cell Lines. The first steps in *N*-glycan biosynthesis appear to be highly conserved throughout all eukaryotic cells, including insect cells, such as those commonly derived from the lepidopteran insect *Spodoptera frugiperda* (*Sf* 9). For example, Kawar et al. (95) recently demonstrated that *Sf* 9 cells encode and express an α1,2-mannosidase, similar to other eukaryotic enzymes, that releases mannose from Man_9, $GlcNAc_2$. In addition, an α1,2-mannosidase has been isolated from virus-infected *Sf* 21 cells; it acts preferentially on Man-6 oligosaccharides (96,97). The authors suggest that this is a pivotal step in glycan processing that determines whether complex or hybrid structures are formed.

However, numerous studies have shown that baculovirus-infected insect cells differ markedly from mammalian cells in subsequent glycan processing capability. These perceived limitations in glycosylation have hindered the wide acceptance of insect cell systems for large-scale production of human therapeutics. Whereas mammalian cells are generally capable of complex-type *N*-glycan processing, insect cells produce proteins with truncated glycans, notably trimannosyl core structures ($Man_3GlcNac_2$) plus or minus fucose (55,98–100). The molecular basis for the accumulation of trimannosyl *N*-glycans on proteins secreted by some insect cells (notably *Sf*-derived cell lines) has recently been elucidated. After initial transfer of *N*-acetylglucosamine onto Man-5 by a GlcNAc transferase I identified in insect cells (101,102), subsequent removal of mannose and *N*-acetylglucosamine by a Golgi-specific α-mannosidase II (103,104) and β-*N*-acetylglucosaminidase (105–107) occurs, yielding the trimannosyl core structure. In some insect cell lines, specifically those derived from *Estigmena acrea*, β-*N*-acetylglucosaminidase activity is absent, and consequently, *N*-glycans bearing terminal *N*-acetylglucosamine residues are present (107,108). Further *N*-glycan processing is cell line specific. For example, several studies report the inability of *Spodoptera*- (*Sf* 9), *Estigmena*-, and *Trichoplusia*-derived cell lines to sialylate *N*-glycans (100,108–110). However, *Mamestra brassicae* and *Manduca sexta* cells have been reported as being capable of sialylating the *N*-glycans of recombinant human plasminogen (111). This group also reported that *Sf*21 cells were capable of synthesizing complex *N*-glycans (112), but that this was dependent upon time postinfection (113)—complex *N*-glycans appearing at later stages of the infection cycle.

Mammalian Cell Lines. Although phylogenetically more similar to humans and generally capable of complex-type oligosaccharide synthesis, commonly used cell lines derived from rodent tissues maintain some key differences in glycan processing.

Firstly, murine cell lines (e.g., hybridomas, mouse–human heterohybridomas, C127 cells, and NS0 cells) express the enzyme α-1,3-galactosyltransferase (α-1,3-GalT), which leads to the synthesis of glycans terminating in Galα1,3-Galβ1,4-GlcNAc motifs (114,115). As more than 1% of human serum IgG is directed against this epitope (116,117), its association with a therapeutic protein would be undesirable, resulting in rapid clearance. In humans, apes, old world monkeys, and CHO cell lines the α-1,3-GalT gene has become inactivated through frameshift mutations (118,119).

Secondly, in contrast to human cells, murine cell lines have a tendency to confer *N*-glycolylneuraminic acid (NeuGc), an oncofetal antigen and immunogenic derivative of *N*-acetylneuraminic acid (NeuAc). The hydroxylase enzyme that converts CMP-NeuAc to CMP-NeuGc is cytosolic (120) and has recently been cloned (121). NeuGc has been found to be associated with the oligosaccharides of an IgM produced by a human–mouse hybridoma, where it was twice as prevalent as NeuAc (122). Similarly, comparison of glycans associated with a chimeric IgG produced by CHO and NS0 cells revealed modification of NS0-derived glycans with NeuGc, and not NeuAc, which was linked to enhanced rate of plasma clearance (123). Proteins produced by CHO cells have a low NeuGc content, typically 1–2% of total sialic acids (124), and are generally not immunogenic (125), although higher levels of substitution with NeuGc may elicit an immune response (126). CHO and baby hamster kidney (BHK) cell lines lack the functional α-2,6-sialyltransferase (ST) enzyme found in human and murine cells and synthesize exclusively α-2,3-linked terminal sialic acids. This alteration has been addressed

by transfection of CHO and BHK cells with an α-2,6-ST gene, as discussed later.

Lastly, although not of immunological significance, another variable *N*-glycan modification that may be associated with recombinant proteins derived from CHO cells is the addition of poly-*N*-acetyllactosamine repeats (Galβ1-4GlcNAcβ1-3-) to terminal Gal residues on the antennae of complex oligosaccharides (126–129).

Transgenic Animals. There are very few studies describing the glycosylation of recombinant proteins produced in the mammary gland of transgenic animals. The few reports that exist indicate that glycan processing is altered. James et al. (54) compared the site-specific glycosylation of recombinant human interferon-γ (IFN-γ) produced by CHO cells, *Sf* 9 insect cells, and the mammary gland of transgenic mice. In the latter case, *N*-glycans at one site (Asn_{97}) were predominantly truncated or of the high-mannose type, whereas *N*-glycans at the other site (Asn_{25}) were of the complex, sialylated variety similar to those associated with IFN-γ from CHO cells.

Similar data has recently been reported for recombinant antithrombin produced in transgenic goat milk, where one of three N-linked sites (Asn_{155}) in the transgenic-derived protein was unaccountably associated with high-mannose-type glycans, not complex glycans as in the natural human protein (130). Furthermore, the recombinant antithrombin contained a GalNAc for Gal substitution on some *N*-glycans, was highly fucosylated and less sialylated than natural human antithrombin, and contained both NeuAc and NeuGc.

The total extent of glycosylation of a recombinant protein produced by transgenic animals may also differ from that of its natural counterpart. A recombinant erythropoietin fusion protein produced in transgenic mouse and rabbits was less glycosylated overall than the molecule produced in CHO cells (131). In addition, recombinant human bile salt–stimulated lipase (BSSL) derived from transgenic mouse milk exhibited a significantly lower degree of O-glycosylation than was found for the native enzyme or for recombinant CHO or C127 cell–derived BSSL (132). In both cases, however, biological activity was not compromised.

Transgenic Plants. Transgenic plants are now being evaluated as a viable option for therapeutic protein production (133–136). However, not surprisingly, *N*-glycan processing in plant cells is substantially different from that obtained in mammalian cell systems. *N*-Glycans universally lack terminal sialic acid residues, and potentially allergenic residues such as core α1-3-linked fucose (137), xylose or β1-2-mannose residues have been detected. Recombinant erythropoietin (EPO) produced in transgenic tobacco cells was found to have no biological activity in vivo (138), possibly because of its lack of sialylation and consequently high clearance rate.

Genetic Engineering of the Host Cell

As already outlined, the glycosylation of recombinant glycoproteins is generally host cell specific, being largely dependent on the complement of specific glycosyltransferases available for glycan processing. Thus, the majority of efforts to augment glycan processing in various host cells by genetic means has involved introduction of a cloned heterologous glycosyltransferase with a desirable processing activity that is absent in the parental host cell. It should be stated, however, that successful introduction of a new biosynthetic activity into a host cell by this route depends on the preexisting capability of the cell to synthesize the requisite nucleotide sugar substrate and the ability to transport it into the Golgi lumen using a specific nucleotide sugar transporter. For this reason, our ability to engineer the protein glycosylation capabilities of yeast and insect cells is restricted at present.

Potential targets for genetic alteration of *N*-glycans of recombinant proteins produced by mammalian cells are summarized in Table 1. In the case of CHO and BHK cells, which both lack a functional α-2,6-sialyltransferase (α-2,6-ST), terminal sialylation of *N*-glycans has received the most attention. This is primarily for two reasons, *(1)* the potential for modulation of the pharmacokinetic behavior of a therapeutic protein product: the relationship between terminal sialylation of *N*-glycans and rate of protein clearance from the bloodstream (mediated by hepatic asialoglycoprotein receptors) is now well documented (149–151), and *(2)* the production of recombinant proteins with "authentic" humanlike *N*-glycans consisting of a mixture of α2-3- and α2-6-linked sialic acids. However, experiments involving cotransfection of cells with cloned α-2,6-ST cDNA have met with variable success: although α2-6-linked sialic acids have been shown to occur in all cases, the absolute level (mole per mole) of recombinant protein sialylation was either not determined (152), unaffected (140), or increased in some transfectants but not others (153). A soluble form of the α-2,6-ST expressed in BHK cells did not modify a model secreted recombinant glycoprotein, indicating that functional activity is dependent upon correct localization of the heterologous transferase (154).

Recent reports highlight the potential use of CHO cells expressing heterologous fucosyltransferases for the production of bioactive oligosaccharides, specifically Golgi membrane and secreted forms of fucosyltransferase III for Lewis (a) synthesis (146), and fucosyltransferase VI for sialyl Lewis X synthesis (145). In the latter case, expression of the introduced glycosyltransferase was controlled using antisense technology.

There have also been recent attempts to engineer baculovirus-infected insect cell expression systems for improved recombinant protein glycosylation. This has been carried out directly, for example, via the stable expression of β1-4-galactosyltransferase in *Sf* 9 insect cells (147), which resulted in the presence of corresponding terminal galactose residues on glycans associated with human tissue plasminogen activator secreted after recombinant baculovirus infection. Alternatively, the glycosylation of proteins produced in insect cells may be improved by the use of baculovirus expression vectors employing the immediate-early promoter element, permitting the production of recombinant glycoproteins with more extensively processed N-linked oligosaccharides (155).

Table 1. Targets for Genetic Engineering of Recombinant Protein Glycosylation in Animal Cells

Protein[a]	Purpose	Function	Reference
	Mammalian cells		
α2,6-Sialyltransferase ↑	Increase α2,6-sialylation	Increase half-life, authenticity	140
Cytosolic sialidase ↓	Eliminate extracellular digestion of *N*-glycans	Increase half-life	141
CMP-Neu5Ac hydroxylase ↓	Eliminate NeuGc	Reduce protein antigenicity	121,142
GlcNAc transferase III ↑	Add bisecting GlcNAc	Enhance bioactivity	143
α1,3-Gal transferase ↓	Eliminate Gal α1,3-Gal	Reduce protein antigenicity	144
Fuc transferase VI ↑	Synthesize sialyl Lewis X	Synthesise biologically active oligosaccharides	145
Fuc transferase III ↑	Synthesize Lewis (a)	Synthesise biologically active oligosaccharides	146
	Insect cells		
β1,4 Gal transferase ↑	Increase glycan complexity	Mimic mammalian proteins	147
β1,2 GlcNAc transferase I ↑	Increase glycan complexity	Mimic mammalian proteins	107
	Transgenic animals		
α1,2-Fuc transferase ↑	Increase α1,2-fucosylation	Synthesise biologically active oligosaccharides	148

[a]A desired up- or downregulation of the gene is indicated by an arrow.

There is only one example where the transgenic animal mammary gland has been engineered for altered glycan processing. Prieto et al. (148) expressed human α-1,2-fucosyltransferase (α-1,2-Fuc T), which generates the H blood group antigen, in transgenic mice. Milk from these animals contained soluble α-1,2-Fuc T, as well as 2′-fucosyllactose (Fucα1-2Galβ1-4Glc) and modified glycoproteins containing the H antigen.

Genetic Engineering of the Recombinant Glycoprotein

There is now a well-documented link between local protein structure and glycosylation at a particular site (5), in other words, glycosylation of a protein produced in a single cell type can be different at individual glycosylation sites. This may arise due to steric hindrance of glycosyltransferases, as in the case of the Fc *N*-glycans of IgG (40,41), or other peptide–carbohydrate interactions. Thus, by utilizing protein engineering technology such as site-directed mutagenesis, the possibility arises that glycosylation sites associated with undesirable glycans (e.g., high-mannose) may be deleted, new sites may be created, or local protein structure may be redesigned to alter glycan processing at a glycosylation site. Two current examples are highlighted: engineered forms of tissue plasminogen activator (tPA) and IgG.

Recombinant human tPA produced by CHO cells has three potential N-glycosylation sites at Asn_{117}, Asn_{184} (variably occupied), and Asn_{448} (156), and one O-linked site at Thr_{61}, which is associated with a single fucose residue. *N*-Glycans at Asn_{117} are invariably of the high-mannose type, in contrast to the complex-type *N*-glycans at the other Asn sites, and are responsible for promoting the clearance of tPA from circulation via interactions with macrophage and hepatic mannose receptors (157). Thus, second-generation tPA molecules, TNK-tPA, have been mutated to remove Asn_{117} (N117Q), with the introduction of a new glycosylation site, Asn_{103} (T103N). The latter mutation restores fibrin binding activity to the wild-type tPA molecule. Significantly, the TNK-tPA variant exhibits complex-type *N*-glycans at Asn_{103} and has a 10-fold increase in its serum half-life (158–160).

Human IgG is glycosylated at Asn_{297} in the Fc region, predominantly with a mixture of variably galactosylated, minimally sialylated biantennary oligosaccharides. Numerous studies have shown that variations in the structure of these glycans can serve to modulate antibody effector functions (40,161) via noncovalent interactions between carbohydrate and local amino acid residues and Fc receptors (41,162). Lund et al. (161) mutated a number of amino acid residues proximal to the Asn_{297} of a recombinant IgG3 produced in CHO cells; specifically, residues 241, 243, 264, 265, and 301 were replaced with alanine in each case. These mutations resulted in increased galactosylation and sialylation relative to the wild-type oligosaccharide chains, although there was a reduction in antibody effector functions in some cases.

Control of the Host Cell Environment

For most recombinant proteins produced by mammalian cells in culture, the cell environment itself can contribute to product molecular heterogeneity. The following factors may all influence the structural diversity of glycans associated with recombinant therapeutic proteins or antibodies:

1. Serum
2. Ammonium
3. Extracellular glycosidases
4. Method of cell culture

Serum. In light of recent bovine spongiform encephalitis epidemics and the desire for control of a fully defined cell culture environment, there is a trend in industry to remove serum from culture media. However, the presence or absence of serum has been shown to affect protein gly-

cosylation. For an IgG1 monoclonal antibody produced by hybridomas, Patel et al. (163) demonstrated that ascites-derived protein differed from serum-free-culture-derived material only with respect to the content of sialic acid. IgG1 derived from culture in serum-containing media had an intermediate sialic acid content and a lower incidence of outer-arm galactosylation than the other two preparations. However, for antibodies produced by CHO cells, *N*-glycan galactosylation may be improved in serum-containing medium (164). For other types of recombinant protein, serum content of the medium may also be significant. Adaptation of BHK-21 cells producing a recombinant mutant IL-2 from a serum-containing to a serum-free environment resulted in changes to *N*-glycan antennarity, sialylation, sialylation, and fucosylation, as well as an increase in the overall degree of protein glycosylation (165).

Ammonium. Increases in the concentration of the toxic catabolite ammonia during cell culture have been linked to various changes in protein glycosylation. Borys et al. (166) demonstrated that exogenous additions of ammonium (as ammonium chloride) up to 9 mM decreased the overall glycosylation of recombinant mouse placental lactogen 1 produced by CHO cells in a pH-dependent manner, this effect being most pronounced at elevated pH (8.0). Subsequently, it was shown that ammonium ion concentrations above 2 mM specifically inhibited the sialylation of recombinant granulocyte-macrophage colony-stimulating factor produced by CHO cells (167). This effect was thought to be mediated by an ammonium-induced elevation of the acidic *trans*-Golgi compartment pH, resulting in reduced sialyltransferase activity in vivo, a concept originally explored by Thorens and Vassalli (168).

In parallel with this work, it was demonstrated that an artificial increase in ammonium ion concentration induced an accumulation of UDP-*N*-acetylhexosamines in mammalian cells, with a concomitant decline in growth rate (169); in fact the "U ratio" (UTP:UDP-*N*-acetylhexosamine) was shown to be a sensitive indicator of cell growth rate. Subsequently, increases in the antennarity of *N*-glycans associated with a human mutant IL-2 produced by BHK-21 cells were also correlated with an increased cellular concentration of UDP-*N*-acetylhexosamines, as artificially induced by exogenous glutamine (7.5 mM) or ammonium (15 mM) (170). In another report, similar increases in the cellular UDP-*N*-acetylhexosamine pool were induced by incubation of BHK cells with glutamine and uridine (171), which increased *N*-glycan antennarity but not sialylation. This suggests that the effect of ammonium on sialylation alone is due to another mechanism—probably changes in *trans*-Golgi pH, as discussed above. In BHK cells, it is now likely that ammonium ions elevate cellular UDP-*N*-acetylhexosamine levels by direct incorporation of ammonium into fructose-6-phosphate to yield a synthetic precursor of UDP-GlcNAc, glucosamine-6-phosphate; the reaction being catalyzed by glucosamine-6-phosphate isomerase (172). Presumably UDP-GlcNAc transport into the Golgi is stimulated, increasing the availability of substrate for resident GlcNAc transferases III and IV involved in *N*-glycan branching.

In practice, therefore, excessive ammonium concentrations in cell culture should be avoided. Alternatively, use of the glutamine synthetase (GS) expression system (173) avoids this issue. Ammonia is utilized enzymatically by the transfected selectable marker, GS, and does not therefore accumulate during cell culture.

Extracellular Glycosidases. Sialidase, β-galactosidase, β-hexosaminidase, and fucosidase have been identified in cell culture supernatant from CHO and other mammalian cell lines (174–177). At neutral pH the most active is sialidase, which has been purified from CHO cell culture supernatant (178). Gu et al. (179) have demonstrated that sialylation of recombinant human IFN-γ produced by CHO cells decreased in parallel with loss in cell viability in batch culture. Therefore elimination of the sialidase from CHO cells constitutes a viable target for cell engineering (141). However, the fucosidase purified from CHO cell culture supernatant could not release fucose from the recombinant CHO cell–produced glycoproteins gp120 or soluble CD4 with the Fucα1-6GlcNAc linkage, or from human serum α(1)-acid glycoprotein with the Fucα1-3GlcNAc linkage (180).

Method of Cell Culture. The type of culture used to produce a recombinant protein affects its glycosylation. For example, during batch and fed-batch cultures of antibodies produced by NS0 cells, and during recombinant human IFN-γ production by CHO cells, there was a progressive increase in the proportion of high-mannose (predominantly $Man_5GlcNAc_2$) and truncated *N*-glycans associated with the recombinant protein (181,182). The increased appearance of high-mannose glycans suggests a change in the glycan processing capability of cells as a function of time in culture (in this case presumably affecting GlcNAc transferase I activity, by as yet unknown mechanisms). This may also be true for truncated *N*-glycans (lacking terminal monosaccharides), although at present it is more likely that these derive from the action of extracellular glycosidases, as previously mentioned.

These considerations emphasize the necessity to harvest product from batch cultures prior to significant cell lysis. Furthermore, As a means of ensuring reproducible product quality, batch-to-batch consistency, and stability, deleterious changes to recombinant glycoprotein product heterogeneity may be averted by implementation of appropriate monitoring technologies capable of rapid quantitative analyses of intact glycoprotein populations, enabling control of bioreactor operations (e.g., to avoid late harvest of undersialylated product).

Few reports describe the effect of long-term continuous culture on the processing and modification of complex recombinant proteins. Goldman et al. (183) have shown quantitatively that both long-term fluidized-bed perfusion and batch cultures of CHO cells produce recombinant human IFN-γ with a consistent and highly comparable degree of glycosylation, monitored as glycosylation site occupancy and level of *N*-glycan sialylation. However, as previously mentioned, the sialylation of recombinant IFN-γ produced in stirred-tank batch culture declined significantly during

the stationary and death phase. Differences in the sialylation of recombinant human tissue kallikrein produced by CHO cells from microcarrier culture and from a serum-free suspension cell recycle process have been reported (130), in which a reduced degree of *N*-glycan sialylation was observed with the microcarrier process.

Other Factors. Other factors may affect glycosylation of recombinant glycoproteins produced by animal cells in culture to a limited extent, although such effects are likely to be host cell and recombinant protein specific. For example, mild hypoxia and elevated pCO_2 have minimal effects on the glycosylation of CHO cell–derived tPA (184,185). In perfusion culture of BHK-21 cells producing a recombinant mutant IL-2, nutrient limitations (glucose, amino acids, dO_2) led to short-term changes in macroheterogeneity, but microheterogenity was largely unchanged (165).

Oligosaccharide Synthesis In Vitro

Advances in the chemical synthesis of oligosaccharides, notably using cloned glycosyltransferases, now permit the synthesis or modification of *N*-glycan-type structures (186,187). Thus, postharvest remodeling of *N*-glycans on purified proteins is now possible (e.g., to fully sialylate *N*-glycans for improved protein half-life in vivo (188), although cost constraints may limit the commercial application of this approach.

SUMMARY

The glycosylation of a potential recombinant therapeutic protein will contribute to a set of product characteristics relating to product efficacy, including bioactivity and/or pharmacokinetic behavior, although these will be protein specific. Initial choice of expression system is of crucial significance and will define the structure of oligosaccharides associated with individual protein glycoforms, although other factors implicit in large-scale bioprocess development may also be significant. At present, the most realistic goal may be a consistent, reproducible, humanlike (albeit heterogeneous) glycosylation pattern. However, an ideal production process would perhaps result in a single, defined protein glycoform, rather than a mixture of glycoform variants, each with varying characteristics. This may be considered as analogous to chiral separation of small-molecule drugs, where there is a necessity for optically pure molecules for medical use.

Such control of recombinant protein glycosylation may possibly be achieved at a number of levels: *(1)* genetic control of glycan processing in the host cell, *(2)* engineering recombinant proteins for a defined glycosylation status, *(3)* control of cell physiology in a bioreactor environment for enhanced production of specific glycoforms, and *(4)* downstream processing to selectively recover specific glycoforms. Incorporation of any of these developments into a bioprocess must involve a consideration of clinical efficacy and potency, technical feasibility, cost of goods, and regulatory requirements.

Central to these issues is our ability to quantitively monitor and completely characterize heterogeneous biological products. Glycoproteins can now be analyzed by a diverse array of techniques, yielding structural information of varying complexity. Recent advances in analytical methodology permit relatively rapid, detailed assessment of glycan structure, glycosylation site-occupancy, and particularly sialylation (extent, type of sialic acid, linkage). More routine fingerprinting of glycoform populations may also be useful to ensure reproducible product quality, batch-to-batch consistency, and stability.

BIBLIOGRAPHY

1. N. Jenkins and E.M.A. Curling, *Enzyme Microb. Technol.* **16**, 354–364 (1994).
2. M.J. Geisow, *Trends Biotechnol.* **10**, 333–335 (1992).
3. N. Jenkins, R.B. Parekh, and D.C. James, *Nat. Biotechnol.* **14**, 975–981 (1996).
4. D.T.Y. Liu, *Trends Biotechnol.* **10**, 114–120 (1992).
5. P.M. Rudd, and R.A. Dwek, *Crit. Rev. Biochem. Mol. Biol.* **32**, 1–100 (1997).
6. A. Varki, *Glycobiology* **3**, 97–130 (1993).
7. J.S. Andersen, B. Svensson, and P. Roepstorff, *Nat. Biotechnol.* **14**, 449–457 (1996).
8. A.L. Burlingame, *Curr. Opin. Biotechnol.* **7**, 4–10 (1996).
9. R.A. Dwek, C.J. Edge, D.J. Harvey, M.R. Wormald, and R.B. Parekh, *Ann. Rev. Biochem.* **62**, 65–100 (1993).
10. D.C. James, *Cytotechnology* **22**, 17–24 (1996).
11. C. Abeijon and C.B. Hirschberg, *Trends Biochem. Sci.* **17**, 32–36 (1992).
12. M.C. Field and L.J. Wainwright, *Glycobiology* **5**, 463–472 (1995).
13. C.B. Hirschberg, C. Abeijon, K. Yanagisawa, E. Mandon, A. Orellana, P.W. Robbins, M. Ishihara, Z. Wei, and S. Swiedler, *Glycoconjugate J.* **10**, 310–310 (1993).
14. D.H. Joziasse, *Glycobiology* **2**, 271–277 (1992).
15. R. Kleene and E.G. Berger, *Biochim. Biophys. Acta* **1154**, 283–325 (1993).
16. C. Rabouille, N. Hui, F. Hunte, R. Kieckbusch, E.G. Berger, G. Warren, and T. Nilsson, *J. Cell Sci.* **108**, 1617–1627 (1995).
17. P. Stanley, and E. Ioffe, *FASEB J.* **9**, 1436–1444 (1995).
18. D.H. Vandeneijnden and D.H. Joziasse, *Curr. Opin. Struct. Biol.* **3**, 711–721 (1993).
19. M. Shelikoff, A.J. Sinskey, and G. Stephanopoulos, *Cytotechnology* **15**, 195–208 (1994).
20. M. Shelikoff, A.J. Sinskey, and G. Stephanopoulos, *Biotechnol. Bioeng.* **50**, 73–90 (1996).
21. Y. Gavel and G. Vonheijne, *Prot. Eng.* **3**, 433–442 (1990).
22. D.C. James, M.H. Goldman, M. Hoare, N. Jenkins, R.W.A. Oliver, B.N. Green, and R.B. Freedman, *Protein. Sci.* **5**, 331–340 (1996).
23. K.J. Colley, *Glycobiology* **7**, 1–13 (1997).
24. R. Kornfeld and S. Kornfeld, *Ann. Rev. Biochem.* **54**, 631–664 (1985).
25. A. Varki, *Trends Cell Biol.* **8**, 34–40 (1998).
26. H.R.B. Pelham, *Cell Struct. Funct.* **21**, 413–419 (1996).

27. S. Cottet, and B. Corthesy, *Eur. J. Biochem.* **246**, 23–31 (1997).
28. K. Fiedler, and K. Simons, *Cell* **81**, 309–312 (1995).
29. W.E. Balch, J.M. McCaffery, H. Plutner, and M.G. Farquhar, *Cell* **76**, 841–852 (1994).
30. J. Krijnselocker, M. Ericsson, P.J.M. Rottier, and G. Griffiths, *J. Cell Biol.* **124**, 55–70 (1994).
31. M. Mizuno and S.J. Singer, *Proc. Natl. Acad. Sci.* U.S.A. **90**, 5732–5736 (1993).
32. J.F. Presley, N.B. Cole, T.A. Schroer, K. Hirschberg, K.J.M. Zaal, and J. LippincottSchwartz, *Nature* **389**, 81–85 (1997).
33. C. Arar, V. Carpentier, J.P. Lecaer, M. Monsigny, A. Legrand, and A.C. Roche, *J. Biol. Chem.* **270**, 3551–3553 (1995).
34. C. Itin, A.C. Roche, M. Monsigny, and H.P. Hauri, *Mol. Biol. Cell* **7**, 483–493 (1996).
35. U. Lahtinen, U. Hellman, C. Wernstedt, J. Saraste, and R.F. Pettersson, *J. Biol. Chem.* **271**, 4031–4037 (1996).
36. C. Itin, M. Foguet, F. Kappeler, J. Klumperman, and H.P. Hauri, *Biochem. Soc. Trans.* **23**, 541–544 (1995).
37. C. Hammond and A. Helenius, *Curr. Opin. Cell Biol.* **7**, 523–529 (1995).
38. A. Helenius, E.S. Trombetta, D.N. Hebert, and J.F. Simons, *Trends Cell Biol.* **7**, 193–200 (1997).
39. C.B. Hirschberg, *Glycobiology* **7**, 169–171 (1997).
40. R. Jefferis and J. Lund, *Chem. Immunol.* **65**, 111–128 (1997).
41. M.R. Wormald, P.M. Rudd, D.J. Harvey, S.C. Chang, I.G. Scragg, and R.A. Dwek, *Biochemistry* **36**, 1370–1380 (1997).
42. P. Umana and J.E. Bailey, *Biotechnol. Bioeng.* **55**, 890–908 (1997).
43. H. Schachter and I. Brockhausen, in H.J. Allen and E.C. Kisailus eds., *Glycoconjugates, Composition, Structure and Function*, Marcel Dekker, New York, pp. 263–332.
44. P. Jackson and J.T. Gallagher, *A Laboratory Guide to Glycoconjugate Analysis* Berkhauser Verlag, Basel, 1997.
45. M.E. Hemling, M.A. Mentzer, C. Capiau, and S.A. Carr, in A.L. Burlingame and S.A. Carr eds., *Mass Spectrometry in the Biological Sciences*, Humana Press, Totowa, N.J., 1996, pp. 307–331.
46. G.D. Roberts, W.P. Johnson, S. Burman, K.R. Anumula, and S.A. Carr, *Anal. Chem.* **67**, 3613–3625 (1995).
47. P. Roepstorff, K.H. Schram, J.S. Andersen, K. Rafn, T. Baldursson, J. Kroll, K. Poulsen, J. Knudsen, and K. Kristiansen, *Mol. Biotechnol.* **4**, 1–12 (1995).
48. G. Siuzdak, *Proc. Nat. Acad. Sci. U.S.A.* **92**, 646–646 (1995).
49. J.T. Stults, *Curr. Opin. Struct. Biol.* **5**, 691–698 (1995).
50. A. Dell, K.H. Khoo, M. Panico, R.A. McDowell, A.T. Etienne, A.J. Reason, and H.R. Morris, in M. Fukuda and A. Kobata eds., *Glycobiology: A 222Practical Approach*, IRL Press, Oxford, 1994, pp. 187.
51. A. Dell, A.J. Reason, K.H. Khoo, M. Panico, R.A. McDowell, and H.R. Morris, *Methods Enzymol.* **230**, 108–132 (1994).
52. A. Dell, H.R. Morris, R.L. Easton, M. Panico, M. Patankar, S. Oehninger, R. Koistinen, H. Koistinen, M. Seppala, and G.F. Clark, *J. Biol. Chem.* **270**, 24116–24126 (1995).
53. J. Peter-Katalinic, *Mass Spectrom. Rev.* **13**, 77–98 (1994).
54. M.L. Vestal, P. Juhasz, and S.A. Martin, *Rapid Commun. Mass Spectrom.* **9**, 1044–1050 (1995).
55. D.C. James, R.B. Freedman, M. Hoare, O.W. Ogonah, B.C. Rooney, O.A. Larionov, V.N. Dobrovolsky, O.V. Lagutin, and N. Jenkins, *Bio / Technology* **13**, 592–596 (1995).
56. D.J. Kroon, J. Freedy, D.J. Burinsky, and B. Sharma, *J. Pharm. Biomed. Anal.* **13**, 1049–1054 (1995).
57. C.W. Sutton, J.A. Oneill, and J.S. Cottrell, *Anal. Biochem.* **218**, 34–46 (1994).
58. D.J. Harvey, *J. Chromatogr. A* **720**, 429–446 (1996).
59. D.I. Papac, J.B. Briggs, E.T. Chin, and A.J.S. Jones, *Glycobiology* **8**, 445–454 (1998).
60. D.I. Papac, E.T. Chin, J.B. Briggs, and A.J.S. Jones, *Glycobiology* **6**, 101–101 (1996).
61. D.I. Papac, A. Wong, and A.J.S. Jones, *Anal. Chem.* **68**, 3215–3223 (1996).
62. B. Kuster, T.J.P. Naven, and D.J. Harvey, *J. Mass Spectrom.* **31**, 1131–1140 (1996).
63. P.M. Rudd, G.R. Guile, B. Kuster, D.J. Harvey, G. Opdenakker, and R.A. Dwek, *Nature* **388**, 205–207 (1997).
64. B.J. Harmon, X.J. Gu, and D.I.C. Wang, *Anal. Chem.* **68**, 1465–1473 (1996).
65. A. Tsarbopoulos, M. Karas, K. Strupat, B.N. Pramanlk, T.L. Nagabhushan, and F. Hillenkamp, *Anal. Chem.* **66**, 2062–2070 (1994).
66. S.A. Carr, M.J. Huddleston, and M.F. Bean, *Prot. Sci.* **2**, 183–196 (1993).
67. E.A. Kragten, A.A. Bergwerff, J. Vanoostrum, D.R. Muller, and W.J. Richter, *J. Mass Spectrom.* **30**, 1679–1686 (1995).
68. R.S. Rush, P.L. Derby, D.M. Smith, C. Merry, G. Rogers, M.F. Rohde, and V. Katta, *Anal. Chem.* **67**, 1442–1452 (1995).
69. P.A. Schindler, C.A. Settineri, X. Collet, C.J. Fielding, and A.L. Burlingame, *Protein Sci.* **4**, 791–803 (1995).
70. J.F. Kelly, S.J. Locke, L. Ramaley, and P. Thibault, *J. Chromatogr. A* **720**, 409–427 (1996).
71. T.P. Patel, and R.B. Parekh, *Methods Enzymol.* **230**, 57–66 (1994).
72. C.J. Edge, T.W. Rademacher, M.R. Wormald, R.B. Parekh, T.D. Butters, D.R. Wing, and R.A. Dwek, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6338–6342 (1992).
73. J.C. Bigge, T.P. Patel, J.A. Bruce, P.N. Goulding, S.M. Charles, and R.B. Parekh, *Anal. Biochem.* **230**, 229–238 (1995).
74. G.R. Guile, P.M. Rudd, D.R. Wing, S.B. Prime, and R.A. Dwek, *Anal. Biochem.* **240**, 210–226 (1996).
75. P.M. Rudd, and R.A. Dwek, *Curr. Opin. Biotechnol.* **8**, 488–497 (1997).
76. M.R. Hardy and R.R. Townsend, *Methods Enzymol.* **230**, 208–225 (1994).
77. M.V. Novotny, *Methods Enzymol.* **271**, 319–347 (1996).
78. G. Teshima and S.L. Wu, *Methods Enzymol.* **271**, 264–293 (1996).
79. D. Figeys, A. Ducret, and R. Aebersold, *J. Chromatogr. A* **763**, 295–306 (1997).
80. S. Hoffstetterkuhn, G. Alt, and R. Kuhn, *Electrophoresis* **17**, 418–422 (1996).
81. D.C. James, R.B. Freedman, M. Hoare, and N. Jenkins, *Anal. Biochem.* **222**, 315–322 (1994).
82. K.G. Moorhouse, C.A. Rickel, and A.B. Chen, *Electrophoresis* **17**, 423–430 (1996).
83. D.E. Morbeck, B.J. Madden, and D.J. McCormick, *J. Chromatogr. A* **680**, 217–224 (1994).

84. R.S. Rush, P.L. Derby, T.W. Strickland, and M.F. Rohde, *Anal. Chem.* **65**, 1834–1842 (1993).
85. K. Kakehi and S. Honda, *J. Chromatogr. A* **720**, 377–393 (1996).
86. G. Okafo, L. Burrow, S.A. Carr, G.D. Roberts, W. Johnson, and P. Camilleri, *Anal. Chem.* **68**, 4424–4430 (1996).
87. P.L. Weber and S.M. Lunte, *Electrophoresis* **17**, 302–309 (1996).
88. A. Guttman, *J. Chromatogr. A* **763**, 271–277 (1997).
89. P. Jackson, *Mol. Biotechnol.* **5**, 101–123 (1996).
90. P. Jackson, *Anal. Biochem.* **216**, 243–252 (1994).
91. K.K.S. Ng, K. Drickamer, and W.I. Weis, *J. Biol. Chem.* **271**, 663–674 (1996).
92. M.E. Taylor and K. Drickamer, *J. Biol. Chem.* **268**, 399–404 (1993).
93. P.J. Kniskern, A. Hagopian, P. Burke, L.D. Schulz, D.L. Montgomery, W.M. Hurni, C.Y. Ip, C.A. Schulman, R.Z. Maigetter, D.E. Wampler, D. Kubek, R.D. Sitrin, D.J. West, R.W. Ellis, and W.J. Miller, *Vaccine* **12**, 1021–1025 (1994).
94. L. Lehle, A. Eiden, K. Lehnert, A. Haselbeck, and E. Kopetzki, *FEBS Lett.* **370**, 41–45 (1995).
95. Z. Kawar, A. Herscovics, and D.L. Jarvis, *Glycobiology* **7**, 433–443 (1997).
96. D.J. Davidson, R.K. Bretthauer, and F.J. Castellino, *Biochemistry* **30**, 9811–9815 (1991).
97. J.X. Ren, R.K. Bretthauer, and F.J. Castellino, *Biochemistry* **34**, 2489–2495 (1995).
98. H.D. Klenk, *Cytotechnology* **20**, 139–144 (1996).
99. E. Kretzschmar, R. Geyer, and H.D. Klenk, *Biol. Chem.* Hoppe-Seyler **375**, 323–327 (1994).
100. M. Manneberg, A. Friedlein, H. Kurth, H.W. Lahm, and M. Fountoulakis, *Protein Sci.* **3**, 30–38 (1994).
101. F. Altmann, G. Kornfeld, T. Dalik, E. Staudacher, and J. Glossl, *Glycobiology* **3**, 619–625 (1993).
102. V. Kubelka, F. Altmann, G. Kornfeld, and L. Marz, *Arch. Biochem. Biophys.* **308**, 148–157 (1994).
103. F. Altmann and L. Marz, *Glycoconjugate Journal* **12**, 150–155 (1995).
104. D.L. Jarvis, D.A. Bohlmeyer, Y.F. Liao, K.K. Lomax, R.K. Merkle, C. Weinkauf, and K.W. Moremen, *Glycobiology* **7**, 113–127 (1997).
105. F. Altmann, H. Schwihla, E. Staudacher, J. Glossl, and L. Marz, *J. Biol. Chem.* **270**, 17344–17349 (1995).
106. I. Vandie, A. Vantetering, H. Bakker, D.H. Vandeneijnden, and D.H. Joziasse, *Glycobiology* **6**, 157–164 (1996).
107. R. Wagner, H. Geyer, R. Geyer, and H.D. Klenk, *J. Virol.* **70**, 4103–4109 (1996).
108. O.W. Ogonah, R.B. Freedman, N. Jenkins, K. Patel, and B.C. Rooney, *Bio / Technology* **14**, 197–202 (1996).
109. T.A. Hsu, N. Takahashi, Y. Tsukamoto, K. Kato, I. Shimada, K. Masuda, E.M. Whiteley, J.Q. Fan, Y.C. Lee, and M.J. Betenbaugh, *J. Biol. Chem.* **272**, 9062–9070 (1997).
110. P.C. Kulakosky, M.L. Shuler, and H.A. Wood, *In Vitro Cell. Dev. Biol. Animal* **34**, 101–108 (1998).
111. D.J. Davidson and F.J. Castellino, *Biochemistry* **30**, 6689–6696 (1991).
112. D.J. Davidson, M.J. Fraser, and F.J. Castellino, *Biochemistry* **29**, 5584–5590 (1990).
113. D.J. Davidson and F.J. Castellino, *Biochemistry* **30**, 6167–6174 (1991).
114. C.A.K. Borrebaeck, A.C. Malmborg, and M. Ohlin, *Immunol. Today* **14**, 477–479 (1993).
115. D.M. Sheeley, B.M. Merrill, and L.C.E. Taylor, *Anal. Biochem.* **247**, 102–110 (1997).
116. R.M. Hamadeh, G.A. Jarvis, U. Galili, R.E. Mandrell, P. Zhou, and J.M. Griffiss, *J. Clin. Invest.* **89**, 1223–1235 (1992).
117. J.A. Lavecchio, A.D. Dunne, and A.S.B. Edge, *Transplantation* **60**, 841–847 (1995).
118. R.D. Larsen, C.A. Riveramarrero, L.K. Ernst, R.D. Cummings, and J.B. Lowe, *J. Biol. Chem.* **265**, 7055–7061 (1990).
119. D.F. Smith, R.D. Larsen, S. Mattox, J.B. Lowe, and R.D. Cummings, *J. Biol. Chem.* **265**, 6225–6234 (1990).
120. E.A. Muchmore, M. Milewski, A. Varki, and S. Diaz, *J. Biol. Chem.* **264**, 20216–20223 (1989).
121. T. Kawano, S. Koyama, H. Takematsu, Y. Kozutsumi, H. Kawasaki, S. Kawashima, T. Kawasaki, and A. Suzuki, *J. Biol. Chem.* **270**, 16458–16463 (1995).
122. T.J. Monica, S.B. Williams, C.F. Goochee, and B.L. Maiorella, *Glycobiology* **5**, 175–185 (1995).
123. J. Marzowski, W.C. Wang, A.R. Flesher, and H.V. Raff, *Biotechnol. Bioeng.* **46**, 399–407 (1995).
124. C.H. Hokke, A.A. Bergwerff, G.W.K. Vandedem, J. Vanoostrum, J.P. Kamerling, and J.F.G. Vliegenthart, *FEBS Lett.* **275**, 9–14 (1990).
125. A. Noguchi, C.J. Mukuria, E. Suzuki, and M. Naiki, *Nephron* **72**, 599–603 (1996).
126. A. Noguchi, C.J. Mukuria, E. Suzuki, and M. Naiki, *J. Biochem.* **117**, 59–62 (1995).
127. K.Y. Do and R.D. Cummings, *J. Biol. Chem.* **268**, 22028–22035 (1993).
128. K.Y. Do, D.F. Smith, and R.D. Cummings, *Biochem. Biophys. Res. Commun.* **173**, 1123–1128 (1990).
129. M. Hummel, H.C. Hedrich, and A. Hasilik, *Eur. J. Biochem.* **245**, 428–433 (1997).
130. E. Watson, B. Shah, L. Leiderman, Y.R. Hsu, S. Karkare, H.S. Lu, and F.K. Lin, *Biotechnol. Prog.* **10**, 39–44 (1994).
131. T. Edmunds, S.M. VanPatten, J. Pollock, E. Hanson, R. Bernasconi, E. Higgins, P. Manavalan, C. Ziomek, H. Meade, J.M. McPherson, and E.S. Cole, *Blood* **91**, 4561–4571 (1998).
132. V.P. Korhonen, M. Tolvanen, J.M. Hyttinen, M. UusiOukari, R. Sinervirta, L. Alhonen, M. Jauhiainen, O.A. Janne, and J. Janne, *Eur. J. Biochem.* **245**, 482–489 (1997).
133. M. Stromqvist, J. Tornell, M. Edlund, A. Edlund, T. Johansson, K. Lindgren, L. Lundberg, and L. Hansson, *Transgenic Res.* **5**, 475–485 (1996).
134. J.K.C. Ma, A. Hiatt, M. Hein, N.D. Vine, F. Wang, P. Stabila, C. Vandolleweerd, K. Mostov, and T. Lehner, *Science* **268**, 716–719 (1995).
135. L. Miele, *Trends Biotechnol.* **15**, 45–50 (1997).
136. M.D. Smith and B.R. Glick, *Food Technol. Biotechnol.* **35**, 183–191 (1997).
137. G.C. Whitelam, *J. Sci. Food Agric.* **68**, 1–9 (1995).
138. F. Altmann, V. Tretter, V. Kubelka, E. Staudacher, L. Marz, and W.M. Becker, *Glycoconjugate Journal* **10**, 301–301 (1993).
139. S. Matsumoto, K. Ikura, M. Ueda, and R. Sasaki, *Plant Mol. Biol.* **27**, 1163–1172 (1995).

140. E. Grabenhorst, A. Hoffmann, M. Nimtz, G. Zettlmeissl, and H.S. Conradt, *Eur. J. Biochem.* **232**, 718–725 (1995).

141. J. Ferrari, R. Harris, and T.G. Warner, *Glycobiology* **4**, 367–373 (1994).

142. A. Gregoire, A. Visvikis, A. Marc, and J.-L. Goergen, in O.-W. Merten, P. Perrin, and B. Griffiths eds., *New Developments and New Applications in Animal Cell Technology: Proceedings of the 15th ESACT Meeting*, 1998, pp. 191–193.

143. A.R. Sburlati, P. Umana, E.G.P. Prati, and J.E. Bailey, *Biotech. Prog.* **14**, 189–192 (1998).

144. J.P. Soulillou, *Biodrugs* **8**, 29–32 (1997).

145. E.G.P. Prati, P. Scheidegger, A.R. Sburlati, and J.E. Bailey, *Biotechnol. Bioeng.* **59**, 445–450 (1998).

146. J. Costa, E. Grabenhorst, M. Nimtz, and H.S. Conradt, *J. Biol. Chem.* **272**, 11613–11621 (1997).

147. J.R. Hollister, J.H. Shaper, and D.L. Jarvis, *Glycobiology* **8**, 473–480 (1998).

148. P.A. Prieto, P. Mukerji, B. Kelder, R. Erney, D. Gonzalez, J.S. Yun, D.F. Smith, K.W. Moremen, C. Nardelli, M. Pierce, Y.S. Li, X. Chen, T.E. Wagner, R.D. Cummings, and J.J. Kopchick, *J. Biol. Chem.* **270**, 29515–29519 (1995).

149. E.S. Cole, E.H. Nichols, L. Poisson, M.L. Harnois, and D.J. Livingston, *Fibrinolysis* **7**, 15–22 (1993).

150. A.R. Flesher, J. Marzowski, W.C. Wang, and H.V. Raff, *Biotechnol. Bioeng.* **47**, 405–405 (1995).

151. B. Smedsrod, and M. Einarsson, *Thromb. Haemostasis* **63**, 60–66 (1990).

152. S.L. Minch, P.T. Kallio, and J.E. Bailey, *Biotechnol. Prog.* **11**, 348–351 (1995).

153. L. Monaco, A. Marc, A. EonDuval, G. Acerbis, G. Distefano, D. Lamotte, J.M. Engasser, M. Soria, and N. Jenkins, *Cytotechnology* **22**, 197–203 (1996).

154. E. Grabenhorst, J. Costa, and H.S. Conradt, *Mol. Biol. Cell* **7**, 2583–2583 (1996).

155. D.L. Jarvis and E.E. Finn, *Nat. Biotechnol.* **14**, 1288–1292 (1996).

156. M.W. Spellman, L.J. Basa, C.K. Leonard, J.A. Chakel, J.V. Oconnor, S. Wilson, and H. Vanhalbeek, *J. Biol. Chem.* **264**, 14100–14111 (1989).

157. M. Otter, M.M. Barrettbergshoeff, and D.C. Rijken, *J. Biol. Chem.* **266**, 13931–13935 (1991).

158. C.R. Benedict, C.J. Refino, B.A. Keyt, R. Pakala, N.F. Paoni, G.R. Thomas, and W.F. Bennett, *Circulation* **92**, 3032–3040 (1995).

159. B.A. Keyt, N.F. Paoni, C.J. Refino, L. Berleau, H. Nguyen, A. Chow, J. Lai, L. Pena, C. Pater, J. Ogez, T. Etcheverry, D. Botstein, and W.F. Bennett, *Proc. Nat. Acad. Sci. U.S.A.* **91**, 3670–3674 (1994).

160. N.F. Paoni, B.A. Keyt, C.J. Refino, A.M. Chow, H.V. Nguyen, L.T. Berleau, J. Badillo, L.C. Pena, K. Brady, F.M. Wurm, J. Ogez, and W.F. Bennett, *Thromb. Haemostasis* **70**, 307–312 (1993).

161. J. Lund, N. Takahashi, J.D. Pound, M. Goodall, and R. Jefferis, *J. Immunol.* **157**, 4963–4969 (1996).

162. R. Jefferis, J. Lund, and M. Goodall, *Immunol. Lett.* **44**, 111–117 (1995).

163. T.P. Patel, R.B. Parekh, B.J. Moellering, and C.P. Prior, *Biochem. J.* **285**, 839–845 (1992).

164. M.R. Lifely, C. Hale, S. Boyce, M.J. Keen, and J. Phillips, *Glycobiology* **5**, 813–822 (1995).

165. M. Gawlitzek, H.S. Conradt, and R. Wagner, *Biotechnol. Bioeng.* **46**, 536–544 (1995).

166. M.C. Borys, D.I.H. Linzer, and E.T. Papoutsakis, *Biotechnol. Bioeng.* **43**, 505–514 (1994).

167. D.C. Andersen and C.F. Goochee, *Biotechnol. Bioeng.* **47**, 96–105 (1995).

168. B. Thorens and P. Vassalli, *Nature* **321**, 618–620 (1986).

169. T. Ryll, U. Valley, and R. Wagner, *Biotechnol. Bioeng.* **44**, 184–193 (1994).

170. M. Gawlitzek, U. Valley, and R. Wagner, *Biotechnol. Bioeng.* **57**, 518–528 (1998).

171. S.I. Grammatikos, U. Valley, M. Nimtz, H.S. Conradt, and R. Wagner, *Biotechnol. Prog.* **14**, 410–419 (1998).

172. A. Cayli, M. Wirth, and R. Wagner, in W. Merten, P. Perrin, and B. Griffiths eds., *New Developments and New Applications in Animal Cell Technology: Proceedings of the 15th ESACT Meeting*, Kluwer Academic, Dordrecht, 1998, pp. 157–160.

173. C.R. Bebbington, G. Renner, S. Thompson, D. King, D. Abrams, and G.T. Yarranton, *Bio / Technology* **10**, 169–175 (1992).

174. M.J. Gramer and C.F. Goochee, *Biotechnol. Bioeng.* **43**, 423–428 (1994).

175. M.J. Gramer, C.F. Goochee, V.Y. Chock, D.T. Brousseau, and M.B. Sliwkowski, *Bio / Technology* **13**, 692–698 (1995).

176. E. Munzert, R. Heidemann, H. Buntemeyer, J. Lehmann, and J. Muthing, *Biotechnol. Bioeng.* **56**, 441–448 (1997).

177. E. Munzert, J. Muthing, H. Buntemeyer, and J. Lehmann, *Biotechnol. Prog.* **12**, 559–563 (1996).

178. T.G. Warner, J. Chang, J. Ferrari, R. Harris, T. McNerney, G. Bennett, J. Burnier, and M.B. Sliwkowski, *Glycobiology* **3**, 455–463 (1993).

179. X.J. Gu, B.J. Harmon, and D.I.C. Wang, *Biotechnol. Bioeng.* **55**, 390–398 (1997).

180. M.J. Gramer, D.V. Schaffer, M.B. Sliwkowski, and C.F. Goochee, *Glycobiology* **4**, 611–616 (1994).

181. A.D. Hooker, M.H. Goldman, N.H. Markham, D.C. James, A.P. Ison, A.T. Bull, P.G. Strange, I. Salmon, A.J. Baines, and N. Jenkins, *Biotechnol. Bioeng.* **48**, 639–648 (1995).

182. D.K. Robinson, C.P. Chan, C.Y. Ip, P.K. Tsai, J. Tung, T.C. Seamans, A.B. Lenny, D.K. Lee, J. Irwin, and M. Silberklang, *Biotechnol. Bioeng.* **44**, 727–735 (1994).

183. M.H. Goldman, D.C. James, M.H. Rendall, A.P. Ison, M. Hoare, and A.T. Bull, *Biotechnol. Bioeng.*, **60**, 596–607 (1998).

184. R. Kimura and W.M. Miller, *Biotechnol. Prog.* **13**, 311–317 (1997).

185. A.A. Lin, R. Kimura, and W.M. Miller, *Biotechnol. Bioeng.* **42**, 339–350 (1993).

186. G.F. Herrmann, Y. Ichikawa, C. Wandrey, F.C.A. Gaeta, J.C. Paulson, and C.H. Wong, *Tetrahedron Lett.* **34**, 3091–3094 (1993).

187. Y. Nakahara, S. Shibayama, Y. Nakahara, and T. Ogawa, *Carbohydr. Res.* **280**, 67–84 (1996).

188. M. Gilbert, R. Bayer, A.M. Cunningham, S. Defrees, Y.H. Gao, D.C. Watson, N.M. Young, and W.W. Wakarchuk, *Nat. Biotechnol.* **16**, 769–772 (1998).

See also CHINESE HAMSTER OVARY CELLS, RECOMBINANT PROTEIN PRODUCTION; PROTEIN GLYCOLYSATION.

GLYOXYLATE BYPASS, REGULATION

DAVID C. LAPORTE
STEPHEN P. MILLER
SATINDER K. SINGH
University of Minnesota
Minneapolis, Minnesota

KEY WORDS

Bypass
Glyoxylate
Isocitrate dehydrogenase
Isocitrate lyase
Kinase
Krebs' cycle
Phosphatase
Phosphorylation

OUTLINE

INTRODUCTION

Overview

The glyoxylate bypass of the bacterium *Escherichia coli* has been used as a model system for the study of metabolic regulation. *E. coli* is an ideal organism for these studies because of the powerful genetic tools that can be used to achieve experimental manipulations that are not possible in many other systems. The long-term goal of these projects is to extract fundamental principles of metabolic control from a detailed analysis of how each regulatory mechanism functions and how these mechanisms interact to control the entire process.

The Glyoxylate Bypass

Adaptation of *E. coli* to growth on acetate or fatty acids requires the induction of the enzymes of the glyoxylate bypass: isocitrate lyase and malate synthase (1–4). This pathway is essential for growth under these conditions because it prevents the net loss of the entering carbon as CO_2 in the Krebs' cycle (Fig. 1) (5,6).

The glyoxylate bypass is regulated at two levels. Coarse control of the glyoxylate bypass is achieved by regulating the transcription of *aceBAK*. The *aceBAK* operon encodes malate synthase, isocitrate lyase, and isocitrate dehydrogenase (IDH) kinase/phosphatase, respectively. Once induced, the flow of isocitrate through the bypass is regulated by the phosphorylation of IDH, the Krebs' cycle enzyme that competes with isocitrate lyase. This phosphorylation cycle is catalyzed by a bifunctional kinase/phosphatase. Both of these regulatory mechanisms are discussed later.

REGULATION OF TRANSCRIPTION

Regulation of the Glyoxylate Bypass Operon

The regulatory circuits that control *aceBAK* are illustrated in Figure 2.

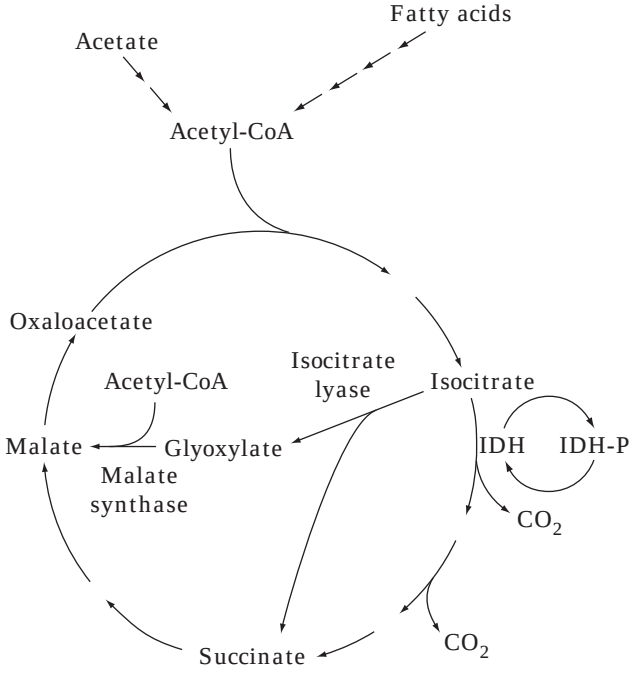


Figure 1. The Krebs' cycle and the glyoxylate bypass. The glyoxylate bypass is catalyzed by isocitrate lyase and malate synthase. The flux through this pathway is regulated by phosphorylation of isocitrate dehydrogenase (IDH).

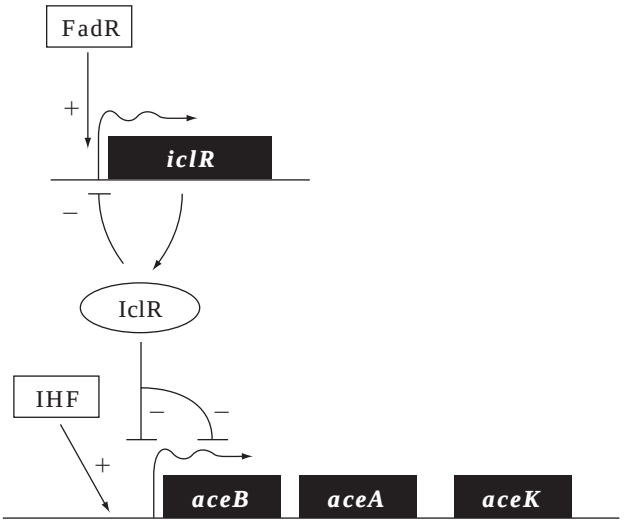


Figure 2. Regulation of the glyoxylate bypass operon. *aceBAK* encodes malate synthase, isocitrate lyase, and IDH kinase/phosphatase, respectively. This operon is repressed by IclR and activated by IHF. The gene encoding IclR is repressed by its own product and activated by FadR.

The *aceBAK* operon appears to respond to metabolic stress. Full expression is observed only during growth on acetate or fatty acids. However, the presence of these substrates is not sufficient for induction because the operon is repressed by the simultaneous presence of preferred carbon sources such as glucose or pyruvate (1–4). Poor carbon sources produce a partial induction of *aceBAK* even though the glyoxylate bypass is not required for growth on these substrates. This partial induction appears to be a response to general metabolic stress because there is a roughly linear relationship between the culture doubling time and the extent of induction when the different growth conditions are compared. Finally, partial induction of the operon occurs upon growth into the stationary phase on rich media (6).

The *aceBAK* operon is transcribed from a single, σ70-type promoter (3). The activity of this promoter is controlled by a repressor protein, IclR, which binds to a site that overlaps the "−35" box of the promoter (3,7,8). Preliminary evidence indicates that IclR represses transcription by competing with RNA polymerase (unpublished). The mechanism that regulates IclR activity has not been determined.

A Remote Operator for IclR

We have recently identified a remote site for IclR. This operator lies within the coding region of *aceB*. Preliminary evidence indicates that binding of IclR to this site does not directly repress *aceBAK* expression. Rather, this site appears to act cooperatively with the promoter-proximal operator, a site that does repress this operon (unpublished). Our working model proposes that this cooperativity results from simultaneous binding of IclR at both sites, with looping out of the intervening DNA. This type of cooperativity between remote sites has been most extensively characterized for the *lac* operon but appears to occur in other operons as well (9–11).

The Role of Integration Host Factor

Integration host factor (IHF) is a sequence-specific DNA binding protein that binds to many sites in the *E. coli* chromosome. This histonelike protein has been implicated in a wide variety of processes including plasmid replication, lambda phage recombination, and transcriptional activation and repression (12–14).

We have identified a binding site for IHF upstream of the *aceBAK* promoter. IHF activates *aceBAK* expression under inducing conditions by preventing IclR repression. In contrast, IHF does not appear to affect IclR activity under repressing conditions (15). The carbon-source specificity of the IHF effect amplifies regulation of *aceBAK*.

IHF can affect transcription by a variety of mechanisms. The most clearly established mechanism is one in which IHF binds to a site well removed from the promoter, bending the DNA at an angle of about 140° and so bringing a remote activator site near the promoter. Similar mechanisms have been proposed for IHF in processes as diverse as plasmid replication and recombination (12–14,16–18).

IHF may also act by direct interactions with other proteins. For example, IHF has been suggested to activate transcription at three promoters by direct contact with RNA polymerase, although this remains to be established (13,14). IHF can also repress transcription by binding to sites that overlap a promoter. In these systems, IHF probably acts by competing with RNA polymerase (19).

IHF plays a novel role in the regulation of *aceBAK* expression. In their recent review, Goosen and Van de Putte have suggested that IHF may not participate in transcriptional regulation (13). This is clearly not true for *aceBAK*, which requires IHF for full carbon-source sensitivity. However, the mechanism by which IHF plays this role must still be determined.

Our working model proposes that IHF and IclR bind independently near the *aceBAK* promoter under repressing conditions (Fig. 3). Under inducing conditions, we suspect that IclR undergoes a structural change that causes the binding of IHF and IclR to become competitive. IclR repression is relieved under these conditions because the IHF site is occupied.

Regulation of *iclR* Expression

The gene encoding IclR is subject to repression by its own product. IclR achieves this autorepression by binding to an operator that overlaps the promoter of *iclR*. This provides a feedback control mechanism that helps maintain a constant level of IclR (20).

The expression of *iclR* is also regulated by FadR. FadR was initially identified because it represses the genes encoding the enzymes of fatty acid degradation (21). It was subsequently shown to activate the transcription of *fabA*, a gene whose product participates in unsaturated fatty acid biosynthesis (22). FadR activates *iclR* expression and may provide a link between *aceBAK* and fatty acids, a carbon source that also requires the glyoxylate bypass (L. Gui, A. Sunnarborg, and D.C. LaPorte, unpublished, 1996).

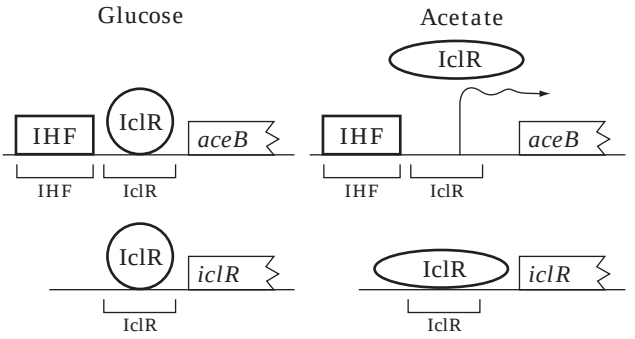


Figure 3. Model for the role of IHF in regulation of IclR. IclR binds to sites that overlap the promoters of *aceBAK* and *iclR*. During growth on glucose (repressing conditions), IHF and IclR can bind simultaneously near the *aceBAK* promoter. Adaptation to acetate (inducing) produces a conformational change in IclR. This conformational change prevents IclR from binding when IHF is bound but has no effect on IclR binding in the absence of IHF. Repression of *aceBAK* is released on acetate because of the presence of an IHF site. Repression of *iclR* is constitutive because this promoter does not include an IHF site.

Differential Control by IclR

Perhaps the most unusual feature of this system is the ability of IclR to differentially regulate *aceBAK* and *iclR*. Whereas IclR-mediated repression of *aceBAK* undergoes a striking response to the carbon source, repression of *iclR* is nearly constant. This difference is probably an advantage to the cell because induction of *iclR* expression on acetate would oppose induction of *aceBAK*. IHF, which is required for the carbon-source sensitivity of IclR, appears to be responsible for differential control by this repressor. *aceBAK* responds to the carbon source because it has an associated site for IHF (see earlier). In contrast, *iclR* does not respond to the carbon source because it does not have a site for IHF (15).

Another Role for IclR?

In addition to the regulation of *aceBAK* and *iclR*, IclR also appears to regulate *acs*, the gene encoding acetyl coenzyme A synthetase in *E. coli*. Mutations in *iclR* reduced consumption of acetate mediated by Acs (23). However, the precise role of IclR in regulating *acs* remains uncertain.

IclR Is a Member of a Unique Protein Family

IclR has lent its name to a family of transcriptional regulatory proteins. This family is characterized by about 25% identity throughout their sequences, with particularly high homology in a region near the N-terminus. This is the region where we identified sequences that appear to match the consensus for helix-turn-helix DNA-binding motifs (24). An "IclR signature" in the C-terminal regions of these proteins has also been defined in the PROSITE database (25).

The IclR family is distinct from other families of transcriptional regulators in bacteria. It is broadly distributed, currently including 19 members found in 12 different Gram-positive and Gram-negative bacterial species; in some cases, the function of the protein is known. Others were identified during genomic sequencing projects, and their functions remain to be identified. The proteins of known function regulate a wide variety of catabolic genes. IclR also fits this pattern because the glyoxylate bypass is essential for the catabolism of acetate (24,26–31 and the NCBI database at *www.ncbi.nlm.nih.gov*).

REGULATION OF PROTEIN PHOSPHORYLATION

Fine Control: The Role of IDH Phosphorylation

Fine control of the glyoxylate bypass is achieved by the phosphorylation of IDH, the Krebs' cycle enzyme that competes with isocitrate lyase. This phosphorylation cycle is catalyzed by a bifunctional protein, IDH kinase/phosphatase (32–34).

Structural Basis of the Regulation of IDH

IDH is a member of a large family of dehydrogenases. This family, comprising at least 17 proteins, includes members from prokaryotes, yeast, and humans. Although some of these enzymes act on isocitrate, others act on isopropyl malate or tartarate. Members of this family exhibit extensive sequence homology, particularly among residues that have been implicated in substrate binding and catalysis for *E. coli* IDH. These enzymes lack the Rossman fold, which is an important structural feature of many other dehydrogenases (35–40).

The phosphorylated form of IDH is completely inactive (33). Inactivation results from the phosphorylation of Ser-113, an active site residue that forms a hydrogen bond with isocitrate in dephospho-IDH (41,42). Phosphorylation has little effect on the three-dimensional structure of IDH. Rather, it blocks isocitrate binding primarily by electrostatic and steric repulsion (43–46).

The mechanism by which phosphorylation controls IDH is a striking contrast with regulatory phosphorylation of glycogen phosphorylase, the only other phosphoprotein whose structure has been extensively characterized. The phosphorylation site of glycogen phosphorylase is remote from the active site. Phosphorylation of this enzyme activates catalysis by inducing a conformational change (47,48).

Structure and Mechanism of IDH Kinase/Phosphatase

IDH kinase/phosphatase is expressed from a single gene and appears to have a single active site that catalyzes the kinase and phosphatase reactions (4,33,34,49–51). This active site also appears to catalyze an ATPase activity that is more efficient then either the kinase or phosphatase reaction (52). Our working model proposes that the IDH phosphatase reaction results from the kinase back reaction tightly coupled to ATP hydrolysis (reviewed in Ref. 34) (Fig. 4).

Regulation of the IDH Kinase/Phosphatase

The IDH phosphorylation cycle is controlled by a wide variety of metabolites. Examples include isocitrate, NADP, 3-

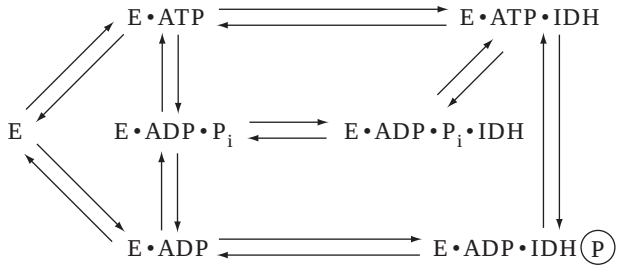


Figure 4. Model for the catalytic mechanism of IDH kinase/phosphatase. This model proposes that the kinase and phosphatase reactions occur at the same active site. The intrinsic ATPase activity of IDH kinase/phosphatase, which occurs even in the absence of IDH, also occurs at this site. The phosphatase reaction results from the back reaction of IDH kinase coupled to the hydrolysis of ATP. E, IDH kinase/phosphatase; IDH, isocitrate dehydrogenase; IDH-P, the phosphorylated form of IDH.

phosphoglycerate, pyruvate, and AMP. There appear to be two distinct mechanisms by which these ligands regulate IDH kinase/phosphatase (Fig. 5) (4,34,50,53–55).

Access to the phosphorylation site of IDH is controlled by the substrates of this enzyme. When these substrates bind to IDH, they induce a conformational change, narrowing the active site cleft and burying the site of phosphorylation. Because isocitrate binds only to dephospho-IDH, this ligand inhibits only IDH kinase. Isocitrate activates IDH phosphatase by preventing product inhibition by isocitrate (56; S.P. Miller, E.J. Karschnia, A.M. Dean, and D.C. LaPorte, unpublished, 1998).

A second mechanism for regulating this cycle results from a conformational equilibrium in IDH kinase/phosphatase. The equilibrium between the kinase and phosphatase conformations is controlled by metabolites that bind preferentially to the phosphatase conformation. Preferential binding activates this activity and inhibits IDH kinase (S.P. Miller, E.J. Karschnia, C. Romfo, and D.C. LaPorte, unpublished, 1998).

CONTROL OF METABOLISM

Growth on Glucose

The glyoxylate bypass serves no useful purpose during growth on carbon sources such as glucose or pyruvate. Under these conditions, *aceBAK* is repressed by IclR. Although a basal level of IDH kinase/phosphatase is present, the conditions favor the dephosphorylation of IDH. Pyruvate, which activates IDH phosphatase and inhibits IDH kinase, appears to be the primary regulator of this phosphorylation cycle under these conditions (6,34).

Growth on Acetate

Adaptation to growth on acetate requires the induction of the glyoxylate bypass operon. The mechanisms responsible for transcriptional regulation have already been discussed.

During steady-state growth on acetate, about 75% of the IDH is converted to the phosphorylated form. Because phospho-IDH is completely inactive, partial phosphorylation of IDH inhibits the Krebs' cycle and so forces isocitrate through the glyoxylate bypass (54,57). Under these conditions, isocitrate must be partitioned between the Krebs' cycle for the production of energy and the glyoxylate bypass for the production of metabolic intermediates. The primary regulators of IDH kinase/phosphatase during growth on acetate appear to be isocitrate, 3-phosphoglycerate, and AMP. Isocitrate and 3-phosphoglycerate act as indicators of the levels of metabolites, whereas AMP reflects the energy needs of the cell (34).

Metabolic Transitions

The flux through the glyoxylate bypass responds dramatically to the culture conditions. For example, addition of glucose to a culture growing on acetate virtually eliminates flux through this pathway (57,58). Inhibition of the bypass is probably advantageous because, when glucose is available, the glyoxylate bypass is no longer required to replenish metabolites used for biosynthesis. Inhibition of the bypass results, in part, from the rapid dephosphorylation of IDH, increasing this enzyme's activity and thus diverting isocitrate through the Krebs' cycle (54,55). IDH kinase/phosphatase catalyzes the rate-limiting step in this transition (55). Long-term adaptation to the carbon source results from control of *aceBAK* transcription.

The response of the IDH phosphorylation cycle during these metabolic transitions is amplified by two mechanisms. Zero-order ultrasensitivity results when one of the modification enzymes (e.g., kinase or phosphatase) is saturated by its protein substrate. Multistep effects, which can either enhance or reduce sensitivity, result when an effector controls more than one step in a protein modification cascade (53). The glyoxylate bypass provided the first experimental demonstrations of these effects. Both phenomena have since been found in other systems.

Study of the glyoxylate bypass led to the discovery of another mechanism for sensitivity enhancement: the branch point effect. This effect results from the competition between enzymes for a limiting amount of a common substrate. During growth on acetate, isocitrate levels are nearly saturating for IDH but are within the pseudo-first-order range of isocitrate lyase. A consequence of this difference in saturation is that the flux carried by isocitrate lyase is ultrasensitive to the phosphorylation state of IDH. Both the theoretical and experimental studies responsible for identifying this phenomena were the products of studies on the glyoxylate bypass (54).

FUTURE DIRECTIONS

The glyoxylate bypass has proven to be a very useful model for metabolic control. The basic control mechanisms have all been identified, with the important exception of the mechanism that controls the IclR repressor. We also have considerable understanding of metabolite levels, flux rates, and how these factors are controlled. These studies have led to the identification of several fundamental principles of metabolic control.

Although the basic control processes have been identified, much remains to be learned concerning their mechanisms of action. How, for example, does IHF control the carbon-source sensitivity of IcIR? Are the models that we

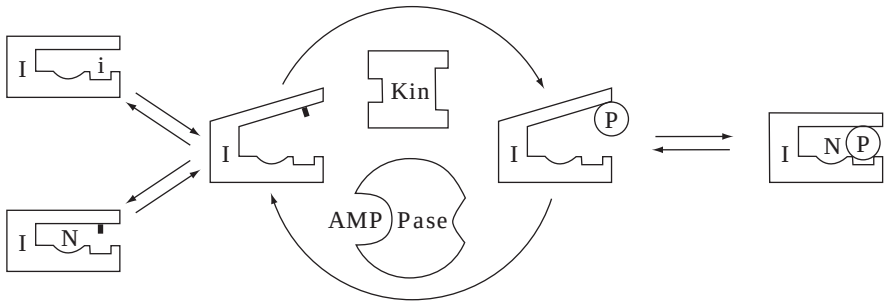


Figure 5. Model for the regulation of the IDH phosphorylation cycle. This model proposes that IDH kinase/phosphatase can exist in two conformations, one favoring the kinase reaction (Kin) and the other favoring the phosphatase (Pase). This conformational equilibrium is proposed to be controlled by metabolites such as AMP, which bind preferentially to the phosphatase form. IDH (I) can exist in either the open or closed conformation, depending on whether the substrates isocitrate (i) or $NADP^+$/NADPH (N) are bound. The open conformation is a substrate for IDH kinase/phosphatase, whereas the closed conformation is not. The site of phosphorylation, Ser-113, is indicated with a tick mark for dephospho-IDH. Phosphoserine-113 is indicated by Ⓟ.

proposed for the catalytic mechanism and regulation of IDH kinase/phosphatase correct and, if so, what are their structural bases? As so often happens in biology, the more we know, the more we realize how much we do not know.

BIBLIOGRAPHY

1. C.B. Brice and H.L. Kornberg, *J. Bacteriol.* **96**, 2185–2186 (1968).
2. S.R. Maloy and W.D. Nunn, *J. Bacteriol.* **149**, 173–180 (1982).
3. T. Chung, D.J. Klumpp, and D.C. LaPorte, *J. Bacteriol.* **170**, 386–392 (1988).
4. D.C. LaPorte, P.E. Thorsness, and D.E. Koshland, Jr. *J. Biol. Chem.* **260**, 10563–10568 (1985).
5. H.L. Kornberg and N.B. Madsen, *Biochim. Biophys. Acta* **24**, 651–653 (1951).
6. H.L. Kornberg, *Biochem. J.* **99**, 1–11 (1966).
7. B. Pan, I. Unnikrishnan, and D.C. LaPorte, *J. Bacteriol.* **178**, 4704 (1996).
8. D. Negre, J.-C. Cortay, A. Galinier, P. Sauve, and A.J. Cozzone, *J. Mol. Biol.* **228**, 23–29 (1992).
9. R. Schleif, *Ann. Rev. Biochem.* **61**, 199–223 (1992).
10. A.M. Friedman, T.O. Fischmann, and T.A. Steitz, *Science* **268**, 1721–7 (1995).
11. A.E. Chakerian and K.S. Matthews, *Mol. Microbiol.* **6**, 963–8 (1992).
12. D.I. Friedman, *Cell* **55**, 545–554 (1988).
13. N. Goosen and P. van de Putte, *Mol. Microbiol.* **16**, 1–7 (1995).
14. M. Freundlich, N. Ramani, E. Mathew, A. Sirko, and P. Tsui, *Mol. Microbiol.* **6**, 2557–2563 (1992).
15. E. Resnik, B. Pan, N. Ramani, M. Freundlich, and D.C. LaPorte, *J. Bacteriol.* **178**, 2715– (1996).
16. S.D. Goodman, S.C. Nicholson, and H.A. Nash, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11910–11914 (1992).
17. L. Moitoso de Vargas, S. Kim, and A. Landy, *Science* **244**, 1457–1461 (1989).
18. T.T. Stenzel, P. Patel, and D. Bastia, *Cell* **49**, 709–717 (1987).
19. P. Tsui, L. Huang, and M. Freundlich, *J. Bacteriol.* **173**, 5800–5807 (1991).
20. L. Gui, A.R. Sunnarborg, B. Pan, and D.C. Laporte, *J. Bacteriol.* **178**, 321– (1996).
21. W.D. Nunn, *Microbiol. Rev.* **50**, 179–192 (1986).
22. M.F. Henry and J.E.J. Cronan, *J. Biol. Chem.* **222**, 843–849 (1991).
23. S. Shin, S.G. Song, D.S. Lee, J.G. Pan, and C. Park, *FEMS Microbiol. Lett.* **146**, 103–108 (1997).
24. A. Sunnarborg, D. Klumpp, T. Chung, and D.C. LaPorte, *J. Bacteriol.* **172**, 2642–2649 (1990).
25. A. Bairoch, P. Bucher, and K. Hofmann, *Nucl. Acids Res.* **25**, 217–221 (1997).
26. A.A. DiMarco, B. Averhoff, and L.N. Ornston, *J. Bacteriol.* **175**, 4499–4506 (1993).
27. A. Ferrandez, J.L. Garcia, and E. Diaz, *J. Bacteriol.* **179**, 2573–2581 (1997).
28. M.A. Prieto, E. Diaz, and J.L. Garcia, *J. Bacteriol.* **178**, 111–120 (1996).
29. C.P. Smith and K.F. Chater, *J. Mol. Biol.* **204**, 569–580 (1988).
30. M.R. Barnes, W.A. Duetz, and P.A. Williams, *J. Bacteriol.* **179**, 6145–6153 (1997).
31. C.S. Harwood, N.N. Nichols, M.K. Kim, J.L. Ditty, and R. Parales, *J. Bacteriol.* **176**, 6479–6488 (1994).
32. M. Garnak and H.C. Reeves, *J. Biol. Chem.* **254**, 7915–7920 (1979).
33. D.C. LaPorte and D.E. Koshland, Jr., *Nature* **300**, 458–460 (1982).
34. D.C. LaPorte, *J. Cell. Biochem.* **51**, 14–18 (1993).
35. J.R. Cupp and L. McAlister-Henn, *J. Biol. Chem.* **267**, 16417–16423 (1992).
36. Y.O. Kim, I.U. Oh, H.S. Park, J. Jeng, B.J. Song, and T.L. Huh, *Biochem. J.* **308**, 63–68 (1995).
37. A. Ishii, M. Suzuki, T. Sahara, Y. Takada, S. Sasaki, and N. Fukunaga, *J. Bacteriol.* **175**, 6873–6880 (1993).
38. P.A. Tipton and B.S. Beecher, *Arch. Biochem. Biophys.* **313**, 15–21 (1994).
39. K. Miyazaki and T. Oshima, *Protein Eng.* **7**, 401–403 (1994).
40. R. Chen, A. Greer, and A.M. Dean, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11666–11670 (1995).

41. J.H. Hurley, A.M. Dean, D.E.J. Koshland, and R.M. Stroud, *Biochemistry* **30**, 8671–8678 (1991).
42. B.L. Stoddard, A. Dean, and D.E.J. Koshland, *Biochemistry* **32**, 9310–9316 (1993).
43. A.M. Dean, M.H. Lee, and D.E.J. Koshland, *J. Biol. Chem.* **264**, 20482–20486 (1989).
44. P.E. Thorsness and D.E.J. Koshland, *J. Biol. Chem.* **262**, 10422–10425 (1987).
45. A.M. Dean and D.E.J. Koshland, *Science* **249**, 1044–1046 (1990).
46. J.H. Hurley, A.M. Dean, P.E. Thorsness, D.E.J. Koshland, and R.M. Stroud *J. Biol. Chem.* **265**, 3599–3602 (1990).
47. M.F. Browner and R.J. Fletterick, *Trends Biochem. Sci.* **17**, 66–71 (1992).
48. S.R. Sprang, K.R. Acharya, E.J. Goldsmith, D.I. Stuart, K. Varvill, R.J. Fletterick, N.B. Madsen, and L.N. Johnson, *Nature* **336**, 215–221 (1988).
49. D.C. LaPorte and T. Chung, *J. Biol. Chem.* **260**, 15291–15297 (1985).
50. D.C. LaPorte, C.S. Stueland, and T. Ikeda, *Biochimie* **71**, 1051–1057 (1989).
51. C.S. Stueland, T.P. Ikeda, and D.C. LaPorte, *J. Biol. Chem.* **264**, 13775–13779 (1989).
52. C.S. Stueland, K.R. Eck, K.T. Stieglbauer, and D.C. LaPorte, *J. Biol. Chem.* **262**, 16095–16099 (1987).
53. D.C. LaPorte and D.E. Koshland, Jr., *Nature* **305**, 286–290 (1983).
54. D.C. LaPorte, K. Walsh, and D.E. Koshland, Jr., *J. Biol. Chem.* **259**, 14068–14075 (1984).
55. C.S. Stueland, K. Gorden, and D.C. LaPorte, *J. Biol. Chem.* **263**, 19475–19479 (1988).
56. J. Finer-Moore, S.E. Tsutakawa, D.B. Cherbavaz, D.C. LaPorte, D.E. Koshland, Jr., and R.M. Stroud, *Biochemistry* (1997).
57. K. Walsh and D.E.J. Koshland, *J. Biol. Chem.* **260**, 8430–8437 (1985).
58. K. Walsh and D.E.J. Koshland, *J. Biol. Chem.* **259**, 9646–9654 (1984).

GOOD MANUFACTURING PRACTICE (GMP) AND GOOD INDUSTRIAL LARGE SCALE PRACTICE (GLSP)

BETH H. JUNKER
Merck Research Laboratories
Rahway, New Jersey

KEY WORDS

Biologics
Drugs
GLSP
GMP
Nonconformances
Performance testing
Qualification
Regulatory guidelines
Validation

OUTLINE

INTRODUCTION

Validation is the process of "establishing documented evidence that a system will do what it purports to do," a simple definition of validation given by the U.S. Food and Drug Administration (FDA) in 1977 (1). The progression of an organization with respect to its validation status has been described by a Validation Capability/Maturity Model (2). It starts with the "validation unaware," then shifts to the "validation aware" once knowledge of the theory is gained. The organization is then "validation active" when it is forced to participate by regulations and customer demand. It moves to the highest level of "validation enthusiast" when it possesses practical experience, recognizes benefits and limitations, and encourages others.

An overview of validation has been presented as a complement to distinct limits in the range of product testing. Validation was developed because of product failures that were not detected as a result of product testing (3). One specific event was a sepsis outbreak caused by improperly autoclaved vials not detected despite sterility testing of a statistically significant number of vials (4). Other authors have compiled an early history of good management practice (GMP) for bulk pharmaceutical drugs (5), early writings on interpretation of validation guidelines (6), and a history of validation in the United States (1). The approaches to issues in biological and pharmaceutical manufacturing change as better analytical techniques evolve, improvements in equipment design and operation are made, and pilot approaches are tested and proven in time.

Bulk pharmaceutical chemicals [BPCs] are made by chemical synthesis, fermentation (recombinant or secondary metabolite), and/or enzymatic reactions. For BPCs, FDA guidelines clearly state that the level of GMP rigor can be modulated based on the position of the step in the overall process (early steps requiring less depth) and the specific product (for some products, early steps might require more depth). A specific evaluation of the FDA's approach to BPCs has been documented with the emphasis placed on determining at what step focused GMP attention

should be given (7). Although the discussion involves a specific processing step, the accompanying issue is the validation of the equipment, portion of the facility, and operating/testing personnel being used for that processing step.

Biologics are products made from living organisms that are complex proteins, viruses, antigens, or antibodies. Production may use recombinant cells, animal or insect cells, plants, or live animals. The guidelines for biologics GMPs include a history of GMPs for biologics (8), a series of 1989 papers by former FDA inspectors (9–11), and a description of differences in validation between conventional pharmaceuticals and biologics (12). Commentary concerning how the guidelines for drug inspections in effect at the time (13) applied to inspections of biopharmaceutical manufacturing facilities by the FDA also was published with clear emphasis on purification, formulation, filling, and lyophilization (14). These papers provided insight as to what the FDA was looking for regarding biologics facility design, validation, and operation. Although these guidelines now are almost a decade old, the areas cited still should be addressed, but in a manner consistent with current GMP guidelines where changes have evolved. As a result of these writings, specific guidelines for biotech facilities were developed (15–18).

There are clear differences in approaches to GMPs between bulk finishing (tableting, granulation, sterile filling) and bulk synthesis (fermentation, chromatography) for both biologics and drugs (19). The focus areas are (*1*) raw material variability, specifically natural sources from cornsteep liquor to fetal calf serum, and the potential for uncontrolled contaminants; (*2*) product tracking and yield, specifically use of activity assays rather than mass assays, and the resulting inability to conduct material balances; (*3*) step yields for which it is difficult to account for step activity losses; and (*4*) process complexity encompassing several processing steps (19). Recent progress in defining a well-characterized biopharmaceutical utilizing advanced analytical techniques (20) has assisted in providing regulatory consistency based on product characteristics rather than whether the product was approved by Center for Drug Evaluation and Research (CDER) or Center for Biologics Evaluation and Research (CBER) (21). It has also permitted processing equipment, procedure, and manufacturing site changes to occur late in the clinical development of a product (22).

Direct and indirect information is available to companies about the scope and nature of validation for both GMP and good industrial large-scale practice (GLSP). Direct statements are present in the Code of Federal Regulations (CFR), specifically Sections 21 CFR 210 and 21 CFR 211, which focus on drugs and pharmaceuticals, and Section 21 CFR 600, which adds additional regulations relevant for biologics. These chapters discuss outcomes rather than methods to achieve them and consequently are subject to interpretation by each company for each type of process. This thematic approach might be analogous to varied interpretations of Bible verses by individuals based on their individual lifetime experiences and goals with no single interpretation universally applicable. Interpretations of the regulations can also be found in various published "Points to Consider" and "Guidelines" documents authored by FDA officials, often with substantial input from industry (23–27). CDER and CBER executive secretariat staff maintain a list of guidelines and provide them readily on request. Public conference presentations, panel discussions, and publications by FDA inspectors also offer interpretations, but, as each situation can be different, it can be difficult for inspectors to speak specifically (or answer a quick question) and for audience members to directly apply the approaches presented. The FDA publication entitled *Human Drug cGMP Notes* provides some guidance regarding specific policy questions. Finally, the major indirect method for companies to obtain information is by reviewing public information about 483 violations and other reports from inspections of other companies. These summaries (obtainable from the "The Pink Sheet" and "The Gold Sheet" published by F-D-C Reports) usually contain two- to three-sentence quotes out of situational context and thus are subject to possible misinterpretation and misapplication. Perhaps the most valuable information concerning "reduction to practice" arises from individual companies who present and publish their specific approaches to GMP and GLSP. Implementation has been a challenge for companies as they evolve from research to a development focus; specifically of note is an early paper published by Biogen outlining these issues for interferon (28).

FACILITY REGULATORY GUIDELINES

Good Manufacturing Practices

The following of current good manufacturing practice (cGMP) regulations, as well as validation execution, is not a complete set of rules but rather a philosophy (29) that is largely dependent on the process and the company. The goal of validation is to satisfy Section 211.42 of 21CFR, which states that the facility and equipment "must be appropriate and suitable for the intended function" (30). Although some may insist that all equipment used in a process must be validated, many others have debated about just how much validation is sufficient. It is most efficient to judge the need for additional validation through scientific justification for the knowledge demonstrated (31). Nevertheless, there can be a "keeping up with the Joneses" philosophy both within and among companies that can lead to wasteful testing.

Validation activities can form a significant percentage of the time and money expenditure in a new pharmaceutical plant. Initially validation was essentially a repeat of installation, construction, and start-up testing and was considered a redundant task whose goal was merely to ensure contractors performed their jobs correctly (30). Often key senior management may not fully appreciate validation scope, benefits, limitations, and intensity. This lack of appreciation is most detrimental if the plant cannot attain full start-up until commissioning is complete; delayed construction and start-up schedules then compress validation, timelines (29). Appropriate validation testing will expose nonconformances and deficiencies in design, construction, and operation (29). Such revelations can be troublesome, particularly if they erupt in waves, but they can be helpful

overall if the end user is permitted to evaluate the problem and to choose to accept operation with the limitation present. Thus, it may not be necessary to resolve every variation or discrepancy with a modification.

The validation master plan is essentially the validation specification and can be as important as the plant specification itself. It is a structured, detailed work plan that, although approved initially by the various subgroups involved, should be a working document linking all aspects of validation from start to finish (29). It might contain a description of the facility, especially its purpose, function, and limitations. A preliminary equipment and utility list, possibly stating what type of testing is required for each item as well as listing preliminary acceptance criteria, is useful for identifying needs for validation test equipment as well as constructing preliminary manpower estimates. Applicable policies and standard operating procedures (SOPs), both approved and those yet to be written, can be referenced, along with floor plans marked with clean/used equipment, raw material/product, and personnel flow patterns. Finally, a sample protocol might be drafted at this point to permit agreement on a common template.

There are several segments and types of validation testing, each of which encompasses various activities depending on the practices of the facility. Grouping of individual activities within these sections is less important than completing the required testing to ensure proper and consistent equipment operation for the processes (31). Duplication of testing among the various validation phases also should be avoided (31). The major categories of validation testing are briefly described next. Most of the initial validation phases lead to subsequent proof of product or process specific reliability.

The design qualification (DQ) focuses on a GMP review to identify potential design weaknesses that may impact qualification/validation before ordering equipment and constructing the plant. One approach is to conduct this GMP review using guidewords, such as *contamination*, *containment*, *validation*, *instrumentation*, *cleaning*, and *sterilization*, to examine each aspect of the design (32). This exercise is best done by representatives of the unit (either internal or outside the company) that is going to be conducting the validation. Some project leaders recommend a different unit conduct the validation than executed the design or construction to maintain appropriate checks and balances. Others believe that "design/build" strategies that encompass validation give a single-source responsibility to a single contractor and avoid fingerpointing when errors are uncovered (33). Potential GMP omissions that are uncovered during this exercise then can be evaluated by the validation/design team. Appropriate solutions may involve additional procedures, extended validation testing, or design modification.

Installation qualification (IQ) provides documented evidence that the equipment was both built and installed correctly according to design criteria and manufacturers' recommendations. Its primary focus is on documentation collection and review including manuals, purchase orders, pressure test data, passivation logs, spare parts lists, SOPs, hazard operability (HAZOP) results, calibrations and loop checks, and as-built piping and instrument diagrams (P&ID)/isometric drawings. IQ ensures that the ordered equipment was received and is safe to operate. Recommendations for the scope and collection of these validation commissioning documents (VCDs) have been documented (34). There is a balance between requesting sufficient documentation and cataloging of excess, redundant, and unnecessary documents (29). Results of factory acceptance testing (FAT) as well as site acceptance testing (SAT) also can be included in the IQ. Typically, the IQ is essentially a documentation review (30) with some not willing even to power up the equipment until the IQ has been executed, reviewed, or even approved. Such constraints might hamper timely identification and resolution of equipment problems, however.

Operational qualification (OQ) provides documented evidence that the equipment and plant operate correctly and have the sufficient range of capabilities over the expected span of operational parameters (29). OQ testing centers on functional tests of equipment features including alarm and interlock verification. It includes draft SOP suitability testing, some aspects of start-up testing, and overall system capacity checks for utilities (30). Often IQ, and in some cases OQ, can be performed by mechanical contractors, and portions of these studies might be performed at the vendor site before shipment of the equipment. The OQ testing needs to be completed satisfactorily to test the limits of the system at a minimum on the aspects on which later performance qualification (PQ) testing will be done.

PQ provides documented evidence that the plant repeatedly can produce product within established specifications and quality attributes (29). It involves extensive testing of the process as well as the process equipment. This testing encompasses many types of qualification and validation, including steam-in-place (SIP), clean-in-place (CIP), sterilizing and decontamination autoclave load patterns, glasswasher load patterns, environmental testing under processing conditions, and analysis of utility quality. It also includes process validation with example protocol strategies available (35). Triplicate runs generally are utilized to show reproducibility but can become extensive when combined with upper and lower processing boundary condition values (1). The concept of a worst-case scenario is used to minimize testing, but the worst case should be within the boundaries of normal operating conditions previously validated for critical parameters and not necessarily an "edge-of-failure" value (1). Before or at least concurrent with performance qualification, key analytical methods utilized should be validated (36).

During facility operation, substantial additional validation effort is required. This effort includes ongoing monitoring of utility and environmental quality, periodic (typically annual) ongoing validation challenges, and revalidation or equivalency testing prompted by equipment or process modification. Appropriate pages and previous acceptance criteria from original protocols might be used in lieu of authoring new protocols in some instances. In other cases, ongoing validation protocols might be reused provided they are written in a generic fashion referencing umbrella policy documents or SOPs for specific details subject to alteration. Once the scope of the ongoing validation

effort is documented and agreed on, the impact of the planned workload on both equipment downtime and manpower can be evaluated.

Commissioning, the proper documentation of facility construction and installation, is an attractive alternative to validation for noncritical utilities that are not in product contact, such as cooling fluids or instrument air. Key aspects and strategies for commissioning have been documented (37). Every activity performed by construction, installation, or start-up personnel is documented although a formal qualification protocol may not be followed. For full effectiveness, commissioning should include challenge and capacity testing (30). Interestingly, one author notes that when noncritical "commissioned" utilities fail or are problematic, their critical importance becomes obvious (30).

The extent of validation is notably dependent on the degree to which in-process controls and end-product testing can ensure quality and safety. Validation thus is most extensive for biologics products (complex therapeutic proteins and vaccines) and less extensive for well-characterized proteins and secondary metabolites (21,38). References are available concerning general validation issues (38), potential validation problems and pitfalls (29,39), process validation (6), process validation life cycle approaches (3), and validation streamlining (40). The details of validation stages, including example regulatory internal and external references, sample systems to be examined, and typical schedules also are available for a model case (41). In addition, a benchmarking survey on current practices in the validation of aseptic processing as well as investigation and additional testing when results are unsatisfactory is available; although it is oriented toward aseptic filling, its approaches might be common to prior aseptic processing steps (42).

Validation protocols should cross-reference but not include items subject to change, such as equipment data sheets and equipment specifications, so that as data sheets are updated by engineering, the protocol is automatically updated (29,43). Protocols may be written and even approved before installation of the equipment, especially when the facility operations group already is familiar with the item. For computer equipment, a computer system validation and system life cycle approach has been adopted (44).

Good Industrial Large-Scale Practices

GLSP (or GILSP) started as the title to a section added to the Guidelines for Research Involving Recombinant DNA Molecules in 1991 (45–47). The GLSP level of containment was recommended for large-scale research or production of biological materials that have an extended history of large-scale safe use (48). The major difference between GLSP and BL1-LS was the reduction of the requirement for a closed system with inactivation of organisms, elimination of the need for exhaust gas treatment, and control of aerosols by procedure not physical devices (47). It referred to a level of physical containment, less rigorous than BL1-LS, which was recommended for recombinant organisms, composed of a nonpathogenic host cell free of adventitious agents and possessing a history of safe use and a well-characterized vector/insert free of harmful sequences. The resulting recombinant organism should be safe to humans and the environment; specifically, the inserted rDNA should not make disease-causing substances more resistant to drugs. The new GLSP section in these guidelines required (47): (*1*) institutional procedures for control of production of recombinant organisms (i.e., an evaluation that health and safety issues are addressed); (*2*) SOPs and training to ensure proper culture handling and maintenance of clean workplaces; (*3*) suitable protective clothing and facility/equipment containment; (*4*) proper handling of environmental discharges; (*5*) aerosol control when necessary; and (*6*) spill containment procedures for large losses. Additional guidelines focus on good lab practices, such as omitting smoking, mouth pipetting, and eating/drinking, as well as the establishment of internal accident reporting procedures.

An outline of BL1 through BL3 issues is available (47) in which the LS designation is added for large scale, typically greater than 10 L, to the corresponding laboratory designation. For BL1-LS, a closed system to reduce the escape of organisms is required, with a validated inactivation procedure before removing material from the closed system. Aerosol production is minimized through design (e.g., sealed sampling devices with filters). Systems are sterilized before maintenance, and an emergency spill plan is in effect. BL2-LS requires all BL1-LS requirements as well as prevention of aerosol generation, enclosure of the closed system within a secondary containment device (e.g., biosafety cabinet), and use of rotating mechanical seals to prevent leakage. Devices should be installed to monitor containment integrity during operation. The closed system should be tested for organism release with a nonrecombinant organism before use and the system should be permanently identified with a biohazard symbol. For BL3-LS, all BL1-LS and BL2-LS requirements are applicable as well as operating the production equipment at the lowest possible internal pressure to ensure containment integrity. The area surrounding the closed system must contain airlock access, readily decontaminated surfaces, sealed penetrations to permit decontamination by gas and reduce cross-contamination into other areas, and hand-washing facilities with valves that are not hand controlled. Finally, there should be controlled airflow into the facility and negative pressure with respect to surrounding areas.

Generally speaking, GLSP guidelines cover National Institutes of Health (NIH) requirements for working with recombinant DNA cultures; however cGMP guidelines may require product protection at the higher BL2 or BL3 levels for some aspects of the process. Specifically negative pressure can be required for BL2 to protect personnel while positive pressure can be required for GMP to protect product. This conflict has been resolved in one instance using inlet HEPA filtration and higher-pressure surrounding areas to provide a barrier as well as physical segregation of live virus areas (49).

A general outline of safety and containment issues and practices at various levels is available that includes examples of monitoring of organism release and suggestions for what equipment and where on the equipment to monitor (50). A goal distinction can be made spanning the range

from product to personnel protection (50): "low" avoids contamination of product cultures, "minimal" targets typical laboratory levels attainable when good laboratory practices are utilized, and "prevent" aims for no release in the air, on surfaces, or on personnel.

MAJOR FOCUSES OF FACILITY PERFORMANCE VALIDATION

Steam-in-Place (SIP)

Strategies. Typically, SIP refers to vessel sterilization, most commonly that of fermenters and associated tankage. It can also refer to sterilization of other items such as dead-end filters in their housings, transfer lines, and cross-flow microfiltration skids. The basic sequence for a SIP procedure involves air removal, steam injection, condensate removal, and the breaking of vacuum with sterile air (51). Typically, the vessel contents are heated with venting until 95 °C, at which time backpressure is applied with continued jacket heating until the vessel temperature is slightly above 100 °C. Steam then can be introduced through various internal ports, such as the sparger line, by turning on associated steam supplies to these ports, creating a forward flow of steam into the vessel. Alternatively, internal ports, such as addition ports without subsurface dip tubes, can be opened to external traps or cracked valves creating a backward flow of steam out from the vessel. Vessel backpressure may be applied by passing the exhaust through a trap or through a backpressure control valve that permits fine adjustments of the sterilization hold temperature. To assure proper equipment operation, backpressure transmitter values should be compared against vessel pressure gauge readings and should be in agreement with expected values for the sterilization hold temperature. An outline of example sterilization procedures for bioreactor, media sterilization filters, and inoculum transfer lines is available (52). In addition, a useful reference describing sterilization theory and practice has been published (53).

The extent of sterilization is calculated using the parameter F_0, the magnitude of which is directly influenced by the area under the sterilization hold temperature versus time curve. The definition and examples of the use of F_0 in the design of sterilization cycles is available (54,55). The F_0 value is a measure of the equivalent kill of a test microorganism, typically *Bacillus stearothermophilus* spores for saturated steam applications, at the vessel sterilization temperature relative to that at 121 °C. Although commonly calculated including only those temperatures at or above 121 °C, calculation beginning at lower temperatures below 121 °C can be appropriate for larger vessels, heat-sensitive media, or difficult-to-heat loads. Accounting for this heat-up time can substantially augment the value of F_0, although the precise amount of this augmentation is complicated by the need to measure D-values for *B. stearothermophilus* spores in the medium under study at various temperatures. A target F_0 thus may be calculated on the basis of temperature mapping results from the system under study and the D-value of spore suspension in the medium to be sterilized.

For most SIP systems, the expected temperature and time required for sterilization are known or can be estimated based on prior experience of the operations group. Thus, production cycle safety factors need not be excessive and might be reduced to as little as a 5-min extension of the hold time or a 1 °C rise in sterilization hold temperature. Conversely, for transfer line sterilizations where F_0 accumulation rises to above 100 min within 5 min post-steam introduction, it can be beneficial to extend the production cycle sterilization time to 20–30 min.

Effectiveness Studies. Effectiveness of SIP cycles is evaluated overall by the rate of foreign contamination occurring in the specific vessel for the specific process over time. Because inoculated cultures might overgrow low-level contaminants, a media challenge, preferably with the same medium to be used during processing, is often conducted. Often a sterility medium is substituted for the process medium, but care should be taken that the medium ingredients are of similar quality as typical raw materials, that current CIP procedures can remove them adequately, and that the vapor pressure and solids content of the medium reflect the problematic aspects of the simulated process. A media hold test typically lasts 7–10 days and uses a growth temperature of 37 °C for favoring bacterial contaminants but a lower temperature of 25–30 °C to favor fungal contaminants. A compromise temperature of 33–35 °C might be appropriate if only one set of tests is desired.

For definitive proof of sterilization performance, a spore-spiked media challenge can be undertaken. Spore challenges, containing 10^{-6} spores, are used in the form of strips for nonsubmerged locations and ampoules for submerged locations. When large particles are involved, spores can be mixed into gels typical of particle sizes, sterilized, then minced to expose spores during sterility testing of the challenges. Successful kill of these challenges provides a sterility assurance of at least a 10^{-6} log reduction of spores. Challenges are placed, along with thermocouples, at the hardest-to-heat locations (i.e., those with the lowest accumulated F_0) based on the temperature distribution study results. Examples of challenged locations include (*1*) condensate return lines for agitator seal, sample valve, addition assemblies, and bottom valve, (*2*) sparger and vent air filters, and (*3*) liquid and headspace of vessel. Typically, the side of the equipment not exposed to direct steam is the side challenged. A summary of spore challenge methodologies is available (56,57); many of these methodologies have been established by the food and pharmaceutical industries.

During processing, sterile samples are obtained from individual vessels on average once per day depending on the facility practices. Sampling might be more frequent if the vessel contents are to be transferred to other sterilized fermenters in the near future. Sampling might be less frequent if containment (either product, personnel, or environment) is an issue. Samples are evaluated for culture purity for an inoculated cultivation and for sterility for an uninoculated cultivation. Aliquots of samples can be subcultured in nutrient-rich broth, incubated at 37 °C to promote growth of bacterial contaminants, then streaked on a nutrient-rich agar and reincubated at 37 °C. Subcultures

can be examined microscopically using Gram stain techniques. In some cases, aliquots can be placed into phenol red–containing medium and any color change noted (typically yellow if contamination is present). Loopfuls of samples can be directly plated onto nutrient agar conductive to bacterial or fungal growth and incubated at 37 °C or a lower temperature of about 25 °C, respectively. False positives on sterility testing have been noted at 0.25% (58). Automated sterility analysis may reduce the number of false positives as well as speed up testing by incorporating fluorescence or other color change endpoints based on metabolic end products (58).

SIP sterilization of hydrophobic sparger or vent filters can be complicated by insufficient air removal. Specifically, dual filters positioned in series on an SIP system may not attain sufficient air removal or condensate drainage due to suboptimal placement of steam condensate piping. Filter integrity testing after the production SIP cycle (which may have the hold time extended by adding a safety factor relative to the fractional test cycle) is required to ensure SIP did not damage the filter. Although in some cases the filters might be heat damaged if the sterilization cycle is unusually long or hot, typically damage is caused by a large pressure differential across the filter, often caused by suboptimal cooldown procedures poststerilization that might not be reviewed during SIP cycle development (59).

Sterilizing filters also may need to be integrity tested after batch use depending on the GMP issues associated with the type of product. Recording of filter lot numbers as well as integrity testing of filters after use can be an important aid in investigating contamination. Post-use integrity testing can be complicated by the difficulty of removing certain materials from the filter that can alter the bubble point or diffusion rate. These issues are best reviewed on a case-by-case basis with the filter manufacturer, who may have measured bubble points or diffusion points for different solutions in addition to those for water. Pre-use integrity testing often is conducted if the vendor testing certification is not sufficient for process use, although passing results do not ensure subsequent filter integrity throughout the batch. Dual series filters can be used to minimize the need for pre-use integrity testing but sterilization strategies need to be adapted to accommodate this arrangement. Post-use integrity testing can be done by removing the filter from its housing, removing the entire housing from the equipment with the filter intact, or testing the filter while located on the equipment with appropriate tie-ins. The latter two methods provide proof of proper installation of the filter if that is required.

Investigations of Failed Results

Most SIP cycle deficiencies can be corrected via SOP (most desired) or equipment modifications (less desired). Equipment modifications might be difficult to accomplish for an automatically sterilized vessel. It should be confirmed that the act of testing is not inducing failures before implementing corrections, however. Investigation strategies for SIP failures might also apply to autoclave load pattern testing failures.

Failures during SIP validation testing might be caused by test procedure faults rather than proposed SIP cycle deficiencies. The biological indicator or thermocouple may be blocking steam/condensate flow, particularly if large spore strips are used in small process lines. Because spore suspensions and strips are incubated in nutrient medium and inspected visually for cloudiness daily, the contaminant cultured should be identified to genus and species if possible. Contamination from handling can result in growth of *Staphylococcus* instead of *B. stearothermophilus* (60). Contamination after testing by contact with unkilled spores from positive controls or from bags used to enclose spore vials might also occur.

The most frequent causes of contamination during fermenter operation (61) include construction materials, mechanical seals, valves, poor or overly complex design, operator error, instrumentation (e.g., calibration, valve failure), process air, transfer/feed lines, contaminated inoculum, incorrect batch or continuous media sterilization, and inadequate procedures. A mechanical checklist for returning a fermenter to service after an initial or repeated contamination can be useful. Factors important in fermenter contamination (62,63) as well as design issues and sterility risks (64) have been documented.

Incorrect specification of procedures often can be responsible for contaminations. Specifically, direct charges of powders through the vessel manway or handhole need to be carefully rinsed into the tank to avoid accumulation of dry powders on the tank upper side wall. Homogenization of clumpy hydrophobic media powders should be considered and sieving of large particles from protein meals (e.g., cornmeal) should be conducted. Media that foam on heat up during sterilization (e.g., medium that contains tomato paste and soy flour) should be avoided as these excessively rich laboratory media, developed for screening and initial growth of a wide variety of organisms, may not be necessary for growth of the desired culture. Use of modified production tank medium for the seed tank stage(s) should be explored instead of using laboratory shake flask seed medium (65).

A monthly contamination tracking system can be used to present trends in the most probable cause of contamination. These might be classified as (*1*) a mechanical problem involving an equipment failure such as leaky valve, leaky mechanical seal, gasket, probe breakage, or incorrect temperature calibration; (*2*) an operational problem in which SOP procedures were not followed; (*3*) a cross-contamination in which contamination detection was not timely and material transferred into a sterilized tank subsequently was found to be contaminated; and (*4*) a process problem in which the procedures specified were not optimal such as incorrect batch sheet instructions concerning sterilization hold time or temperature. These monthly reports should be circulated to the operations group at all management levels and mechanical staff, as well as reviewed by quality personnel.

Contaminating organisms should be identified both in a preliminary fashion by microscopic observation and Gram stain and more definitively using commercial biochemical testing kits or fatty acid analyses. A record of the types of contaminants as well as ages and processes should be maintained and evaluated at least yearly for trends. Yearly contaminants can be classified into categories by

process, age of contamination (e.g., after sterilization [AS], after inoculation [AI], <48 h, 48–120 h or >120 h), and type of organism (cocci, gram-positive, gram-negative rod, fungal, or miscellaneous). The sterility investigation experience of the facility then can be utilized to determine the likely areas where the contamination may have entered. Example guidelines suggest that (*1*) gram-positive rods might result from incomplete media sterilization, improper inlet air filter installation, or air filter integrity failure; (*2*) gram-negative rods might result from incomplete water sterilization (perhaps of filter-sterilized nutrient additions), jacket leakage into the vessel, or water present in the inlet air to the vessel; (*3*) cocci might result from human contact during either laboratory seed preparation or seed fermentor inoculation; and (*4*) fungal contaminants, although rare, might result from improper inlet air sterilization as well as inadequate cleaning of cracks/crevices from fermenter ports.

Clean-in-Place (CIP)

Strategies. The goal of cleaning validation is to satisfy 21 CFR 211.67 in which it is stated that written equipment cleaning procedures should be established and validated. It should be demonstrated that cleaning procedures consistently reduce product and cleaning agent residuals to predetermined levels that are acceptable for the processing steps. CIP was first developed by the dairy industry (66), and this industry has conducted substantial research into effective strategies and appropriate CIP skid design. One major decision important for CIP skid design (67) is the issue of reuse of acid–base cleaning solutions, which must be weighed against concerns of cross-contamination and resoiling of equipment. A second major decision is the degree of dismantling and manual cleaning operationally acceptable for each piece of equipment. The amount of time expected to elapse before initiation of cleaning and the length of time equipment can sit idle before recleaning need to be established on the basis of reasonable expectations of operating practices as well as process requirements.

Key factors in cleaning effectiveness are the discharge of cleaning solutions at their operating pH (if acceptable to site environmental officials) to avoid precipitation of soil, adequate rinsing of cleaning solutions through all internals, adequate contact time of cleaning agent/rinses, adequate velocity for contact of cleaning agent/rinses, maintenance of optimal temperature during cleaning solution hold time or recirculation time, and proper concentration of cleaning solution. The drainability of equipment is another key factor in ensuring cleaning effectiveness. Common problem areas include minor dead legs of a few milliliters that are not rinsed through during the cleaning process, circulation pumps that are not completely drainable, shadowed areas in which impingement of a spray is prevented by vessel internals, and crevices from bolts and fittings. In cases where failed batches are pasteurized or sterilized in the vessel, dilution of the batch with water can assist subsequent cleaning efforts.

Often there are multiple points of entry of CIP liquid into a vessel (e.g., spray balls, air inlet, and vent and transfer lines) and liquid is returned by the reactor drain (67). Spray balls can be fixed or rotating, removable during processing or remaining in batch contact. High pressures and flow rates can be required for proper impingement for best cleaning effectiveness leading to heavy utility loads, particularly for the water rinses. Although in many applications two spray balls are located at the vessel top, additional spray balls can be aimed at hard-to-reach locations such as impeller undersides or behind baffles. The acceptable level of flooding of the vessel needs to be determined as the filling of the vessel to at least the lower impeller may enhance cleaning of lower baffles and bottom mechanical portion of the agitator (67).

A typical cleaning scenario (68) may include several steps depending on the product requirements. These steps are an initial thorough rinse with lower-quality warm water to remove loose soil (particularly at the batch level), a caustic-based detergent solution to hydrolyze biological materials, a rinse with deionized water, an acid-based detergent solution to remove traces of mineral salts, a second rinse with deionized water, and final rinses with higher-purity water. In some cases, a water boilout (heating of water to 80–90 °C) is conducted before the final water rinses. The acid-based detergent typically is not necessary for fermentation applications, especially if deionized water used for cleaning, and can be done on an as-needed basis with 0.5 wt% nitric acid (67). Dilute phosphoric acid and sodium hydroxide can be used in lieu of acid/caustic-based detergents for some applications, although the presence of a surfactant can greatly enhance soil removal. Finally, an empty vessel sterilization, with condensate flowing out from the bottom valve, can aid in soil removal.

For cleaning of nonvessels, the ability to document effectiveness becomes increasingly more difficult. Membrane filters often are restricted to single product use primarily because it is difficult to ensure total removal of the prior product without destructive testing. Filters can be reused for the same product, but sanitization/sterilization procedures are of concern to minimize unwanted bioburden. For homogenizers and cell disrupters, cleaning and visual inspection of the head are likely to require substantial manual disassembly and rebuilding.

In the case of chromatography columns, significant resin expense suggests reuse is especially desirable for multiple products, but product-specific testing for removal is exceedingly difficult. More commonly, chromatographic media are dedicated to one product but equipment (e.g., columns, piping, pumps) is used for multiple products (69). Chromatographic media have been reused for multiple cycles (more than 100) of the same product, but validation of number of multiple uses is required (70). Specifically, the media and equipment must be compatible with validated cleaning processes and consideration must be given to which substances bind reversibly and irreversibly to the resin (69). Cleaning and sanitization can occupy about 20% of each cycle period, which encourages the use of larger versus smaller columns (69). Blank runs are conducted followed by a reduced-cycle cleaning run and evaluation of the contaminant breakthrough potential (69). Finally, quantification of the functional stability of the resin for

specific cleaning agents can be accomplished by measuring retention times of various test proteins (69).

Effectiveness Studies. Hardest-to-clean locations should be identified based on prior experience with the existing vessel, initial cleaning studies of a new vessel, or the best judgment of the processing group based on similar vessel designs. The criteria for sample site location should be documented. Some examples of common locations include the air sparger, underside of impellers, batch level, sample valve, baffles, lifting holes for shaft, fermenter dome, Ingold ports/plugs, threaded fittings, and bolted internals and piping dead legs on addition ports. The use of worst-case scenarios, while a popular approach, may lead to overly rigorous cleaning that is operationally burdensome for many nonworse-case processes although the amount of testing can be reduced for a multiuse facility.

Visual inspection, while subjective, is the most powerful asset available for the prompt analysis of cleaning effectiveness. Training of inspectors properly concerning where and how to look can greatly improve the likelihood of identification of problem areas. Chrome mirrors with extended arms, powerful flashlights, and clean white cloths can be utilized to aid inspection. Tank imperfections should be documented ahead of time to avoid mistaking tank welds, external coiling coil welds, or tank intrinsic stains as cleaning failures. Unfortunately, visual inspection may not be possible for some pieces of equipment without excessive mechanical dismantling (e.g., disk-stack centrifuges, packed chromatography columns). In some cases, the limit of visual detection has been determined using spiked coupons of typical process soils. For example, the visual detection limit was 100 μg/ 25-cm^2 swab area for an active pharmaceutical ingredient isolated powder (71).

The sampling of rinse water indirectly evaluates equipment cleanliness, but it assumes that the residual soil is soluble and thus totally removed in the rinse water. Sometimes rinse water samples are taken throughout the cleaning process so the progression of cleaning can be evaluated; this testing can be detrimental to sensitive assay equipment such as total organic carbon (TOC) analyzers that are not designed for excessively soiled samples. Solution pH can be especially useful to indirectly detect cleaning agent residuals because most cleaning agents are concentrated acids or bases in addition to carbon-based surfactants (Table 1). In addition, a 100 ppm solution of CIP100 corresponds to a TOC of 4.8 ppm (66); other results are available for other concentrations (Table 1).

Direct sampling of the vessel interior is accomplished by a consistent surface swabbing procedure, typically a 25-cm_2 area (5-cm $\times$ 5-cm square area or equivalent). Recoveries depend on the type of contaminant and type of surface being swabbed (72). Specifically, recovery data for a purified recombinant DNA protein (containing phosphate and lactic acid at pH 4.5) varied from 89% for stainless steel to 62% for polypropylene to 55% for glass (72). Recovery data for stainless steel varied from 67% for whole-cell homogenate to 82% for formulated protein (containing mannitol, phosphate, and lactic acid at pH 4.5) and 105% for purified protein (72). Other recovery data for three types of media using different types of materials (e.g., stainless steel, Nalgene, Teflon, polypropylene, Pyrex) are available (Table 2).

Table 1. Ability of Various Dilutions of Cleaning Agents CIP100 and CIP200 to Pass USP Chemical Tests

Solution	pH	Conductivity (μmho/cm)	TOC (ppm)
	CIP100		
1% (10,000 ppm)	12.6	1,927	25.7
0.05% (5,000 ppm)	11.4	581	13.4
0.01% (1,000 ppm)	10.7	123	3.5
0.001% (100 ppm)	9.4	8.6	0.57
0.0001% (10 ppm)	6.5	2.9	0.16
	CIP200		
1% (10,000 ppm)	2.0	7,530	—
0.05% (5,000 ppm)	2.7	872	9.7
0.01% (1,000 ppm)	3.2	179	2.6
0.001% (100 ppm)	4.1	21.4	0.41
0.0001% (10 ppm)	5.1	2.1	0.15

Rinse water and swab samples can be tested for TOC, protein content, specific product, residual cleaning agent, and overall rinse water quality roughly equivalent to that of the water source used for cleaning (e.g., pH, conductivity, TOC, bioburden, limulus amoebocyte lysate [LAL]). Product-specific testing can include ELISA, SDS-PAGE, and other protein-specific assays. TOC analysis, which detects all carbon-based residuals, is becoming a low-cost, rapid-turnaround, low-level detection method for applications when a product-specific test is not required. A description of the method and analyzer and its limitations is available (66). The sample TOC content must be within the range of instrument calibration for accurate results. For equipment that is to be used wet and without subsequent sterilization or sanitization, a bioburden study might be beneficial using both rinse water and swabbing tests. Such a study should be done after the equipment has been sitting idle for a period of time greater than or equal to permitted idle periods during processing. Procedures have been documented for bioburden (73) and LAL (74) testing.

A survey of sampling methods (specifically visual, rinse water, and swab) and assay methods used for early and later phase processing was conducted for multiple companies as well as details of methods in place at Chiron presented (75). Case studies of cleaning validation have been published by Regeneron (76) and Immunex (77). As processing progresses into downstream isolation steps, methods for analysis must have greater product specificity (66), but for upstream steps nonspecific assays such as TOC are likely to be sufficient. In some cases this requirement adds substantial cost and sampling burdens. Acceptance criteria are affected by the sensitivity of analytical equipment in part as well as the amount of residue tolerable in the equipment for the specific processing step. Guidelines for carryover levels of product have been established for specific products by some companies such as Eli Lilly (71).

Cleanability studies using different types of materials can be conducted. The general procedure is to submerge a

Table 2. Recovery Study Results (Fractional Recovery) for Media from Three Processes Tested in Absence of cells

Material of coupon	Recombinant yeast/chemically defined medium 1	Recombinant *E. coli*/chemically defined medium 2	Mammalian cell culture/Williams E medium
Stainless steel	0.86	0.80	0.92
Nalgene	0.28	0.58	1.07
Teflon	0.35	0.65	1.11
Polypropylene	0.36	0.86	1.05
Pyrex	0.26	0.61	0.76

coupon in the soil matrix, evaporate to dryness, and then process it through cleaning solutions/rinses either in stagnant trays or using impingement (Table 3). Each surface is then visually evaluated and swab samples submitted for TOC testing. Dye removal studies can be conducted, but these may not be representative because of different rinsing characteristics (67). Challenge studies also can be developed to document log reduction of TOC, bioburden (based on indicator organisms), or endotoxin by spreading a known amount of material on the equipment. Usually more material is spiked than typically is present as a contaminant so as to have a quantifiable amount remaining at the conclusion of the simulated cleaning process. It is not clear that soil removal is linear between these two rather disparate concentration ranges.

Validation of a backup cleaning agent is recommended in case material from the primary supplier becomes unavailable. This assignment might be accomplished by demonstration of equivalency during the initial validation or as part of a later changeover program. Testing might include cleanability studies on test coupons and test equipment using worst-case scenarios or conducting equivalency runs with a new agent in actual processing equipment.

Investigations of Failed Results. Investigation of failed cleanout procedures generally center on procedural issues rather than testing interferences as in the case of SIP testing. Common deficiencies include insufficient hold time or hold temperature for cleaning agent, inadequate contact of cleaning agent or rinse water with all vessel inner surfaces and internals, improper amount of cleaning agent used, and inappropriate cleaning pretreatment steps. Sufficient time must be devoted to training and clarification of procedures with operating staff. For example, the equipment must be properly vented to allow complete rinsing. Because the rinse water quality depends on both the cleaning effectiveness and the quality of the source water, proper water system operation also may impact the investigation.

Laboratory Autoclave Load Pattern Testing

Load pattern testing should be conducted for typical loads used for laboratory-sterilized materials that are to be transferred to sterilized process equipment. Items to be sterilized should be placed in the autoclave such that they are not touching and there is adequate room for steam to surround each object. Temperature distribution testing of the load should be conducted. Even if the accumulated F_0 appears satisfactory, spore challenges should still be pursued because spores will be killed efficiently only if moist heat is present (78). Assuming both the temperature distributions and spore challenge test results are adequate, a production run should be performed with the actual material to be sterilized. As reasonable, each element in the load should be tested for sterility either through a test usage for equipment or by incubating the liquid load and then testing sterility either through a test usage for equipment or by incubating the liquid load and then testing sterility samples. This production load sterility testing is particularly important to test equipment wrapping closures, ability of personnel to execute validated load patterns, and the successful sterilization of a solids-containing medium.

For dry materials, use of the prevac cycle is most desirable to remove air from the equipment promptly. If a prevac cycle is not available or not considered reliable due to equipment failure or fear of disrupting product closures, then a small amount of water (10 mL in a 4-L bottle) may have to be added to increase the amount of steam generated within the bottle. Sterilization of filter manifolds for nutrient addition is complicated by the fact that hydrophobic filters may not be wet enough to generate sufficient steam for spore kill. This is particularly an issue with autoclaving of capsule filters in series because of poor heat transfer through the plastic filter body and poor steam penetration through the tubing ends. Dual filters in series need to be tested carefully, especially when more than one set is used in a manifolded arrangement that limits steam

Table 3. Cleanability Study Results: Residual TOC after Cleaning of Dried Medium to Point of Visual Cleanliness up to a 10-min rise via Impingement

Material of coupon	Recombinant yeast/chemically defined medium 1	Recombinant *E. coli*/chemically defined medium 2	Mammalian cell culture/Williams E medium
Stainless steel	0.10	0.13	0.08
Nalgene	0.16	0.08	0.07
Teflon	0.09	0.09	0.09
Polypropylene	0.10	0.10	0.08
Pyrex	0.01	0.14	0.10

penetration. After each production cycle run of a validated autoclave load pattern, any filters in the load need to be integrity tested to ensure that no damage has occurred during autoclaving. A complete reference on integrity testing of membrane filters is available (79).

For liquid cycles, liquid heat-up times are greatly affected by the container size, fill volume, type of liquid, and material of construction of the container. There is a conflict between safety directives to remove glass from processing areas by replacing glass with plastic containers and the substantially longer heat-up times observed with plastic (e.g., Nalgene) carboys because of their low heat transfer coefficient. Decontamination of liquids or equipment should be conducted in a different autoclave unit than the sterilizing autoclave. Decontamination load patterns also may undergo load pattern validation, although the heat sensitivity of the individual organism should be considered during validation. Although worst-case studies using *Bacillus stearothermophilus* may be appropriate, typical BL1 or BL2 organisms or viruses may exhibit a substantially lower heat sensitivity, thus reducing the necessary decontamination load cycle time. This potential for reduction may become particularly important for larger liquid volume loads.

Generally, the vacuum break filter is steam sterilized separately before the autoclave load. Periodically, after a specific number of autoclave loads have been sterilized, the filter sterilization cycle is run again. Because most vacuum break filters are not maintained under positive air pressure between the conclusion of the filter sterilization cycle and the beginning of the autoclave chamber cycle, then this filter does not necessarily remain sterile once the autoclave door is opened.

Utilities

Periodic monitoring of utility quality can be a critical factor to ensure product quality. Monitoring focuses on water, steam, and air systems in product contact. Frequencies can vary from daily for water systems to monthly for clean steam systems. In a few applications, sampling just before, during, and/or after system use is warranted. Action limits should be established on the basis of product quality requirements, while alert limits should be established based on historical data (if available) or a comfortable safety margin. Because the alert limit level does not impact product quality but rather serves to warn of the possibility for a quality problem to occur, it should be set considering input from the operations group regarding prior or expected system behavior.

By far the major testing focus is on water systems because impurity levels can become concentrated by the large volumes used in processing. Testing may include USP chemical analysis, bioburden, and endotoxin depending on the application. Although sampling of water systems is complicated by multiple use points, in many cases sampling may be rotated among use points so long as each use point is sampled on a regular basis. Typically, at least one bioburden and one endotoxin sample are taken daily from at least one point of use. Recent changes to the United States Pharmacopeia (USP) affect purified water and water for injection (WFI) testing regimens (80). The goal of these modifications was to reduce required USP chemical testing to conductivity, pH, and TOC (66) analysis. Before this change, numerous chemical tests were required, some of which were qualitative in nature with end points that were difficult to interpret. The impact of these new USP chemical testing procedures needs to be understood so it can be evaluated for existing water systems (81). Newly available on-line TOC, pH, and resistivity meters can ensure proper system operation according to these new USP chemical testing guidelines.

Bioburden evaluation is conducted to determine both the number and type of organism present. Water systems are subject to the formation of biofilms in which bacteria attach themselves to piping crevices or cracks. They create a polysaccharide extracellular matrix that promotes long-term attachment of both the original and subsequently adhering bacteria. References on biofilm formation mechanisms and control are available (82–84). Like weeds on a lawn, once present biofilms are difficult to remove so it is best not to permit them to form initially.

LAL analysis for endotoxin, while primarily a test conducted on clean steam and WFI systems, can also be helpful for monitoring operation of USP purified water systems. In fact, LAL and TOC and bioburden results may be examined together to investigate excursions. All material present in water that is quantified as TOC is not necessarily endotoxin because the endotoxin assay is specific for lipopolysaccharide. All endotoxin should be detected in the less-specific TOC assay, however. Using the *E. coli* endotoxin standard, TOC levels were measured for 0.25, 0.5, and 1.0 EU/mL solutions and a linear relationship was obtained:

$$\text{TOC (mg/L)} = 72.2\ \text{LAL (EU/ml)} - 5.6 (r^2 = 0.99) \quad (1)$$

Consequently, for an endotoxin limit of 0.25 EU/ml (a typical action limit for WFI systems), the TOC measured in the same sample might be as high as 14 mg/L (14 ppm). Thus, TOC analysis can be a good indicator of rising endotoxin levels and may serve as confirmation of endotoxin assay results. Similarly, a rise in bioburden from 1 to 100 cfu/mL may not coincide with a noticeable rise in TOC because 1 to 100 cfu/mL corresponds to a range of 10^{-4} to 10^{-6} ppm TOC substantially lower than usually detected (assuming 10^{-12} g dcw/cell and 50% carbon content).

Controlled Environments

Controlled environments ensure that there is documented control of air quality that is appropriate for each processing step. Establishment of classified areas should be examined carefully to ensure that classification at the proposed level is reasonable while the area is in operation. Specifically, certain areas that produce powders, mist, or steam may not be able to achieve class 100,000 conditions if samples were taken during truly worst-case dynamic operation. Such areas might include rooms containing glasswashers, autoclaves, fermenters, or raw material subdivision operations.

A comprehensive monitoring program should be developed to include items such as viable airborne particles,

nonviable airborne particles, pressure differentials, airflow direction, temperature, relative humidity, and surface microbial contamination present on staff, equipment, work surfaces, walls, doors, floors, and curtains (85,86). Monitoring should be conducted using viable and nonviable air samplers, contact plates, surface swabs, and settling plates. A regular monitoring schedule should be instituted, with some tests taken during simulated or actual dynamic conditions (while work is occurring in the area). When bringing an area in and out of GMP service, initially there might be an intensive amount of monitoring, perhaps 3–5 consecutive days, to show consistency of air quality before initiation of GMP processing. A summary of environmental monitoring program characteristics, including additional references, is available (87).

Evaluation criteria for a laminar flow hood or biosafety cabinet are dependent on the critical nature of the hood, specifically whether there is a subsequent sterilizing step downstream for a sterile product. Monitoring during actual operation sometimes is considered using either exposure plates or an air sampler. If monitoring is conducted directly within the work area, then it may interfere with the technician's ability to properly manipulate materials. Conversely, if monitoring is moved to occur in a corner, then testing is not actually sampling air exposed to product. Motorized air samplers interfere with desired airflow patterns in a biosafety cabinet or laminar flow hood by creating air currents and noise that can be a distraction to the technician. Settling plates give a qualitative outcome because a defined amount of air is not sampled (85). These plates may selectively attract larger particles and may need to be exposed for a far longer time than the length of the actual manipulation (85).

Action and alert limits should be set based on the nature of the step, the ability of subsequent steps to remove or propagate contaminants, and the product exposure risk (open versus closed system) (85). Although a typical alert limit is 80% of the action limit, the alert limit may be set lower to ensure a comfortable quality safety margin. Data should be trended over a year to detect seasonal variations. Often data do not monotonically increase toward the action or alert limit but rather spike to higher levels because of a system disturbance.

For action and/or alert limit excursions, identification of the microorganism class can be useful (e.g., gram-positive or gram-negative rod, cocci, fungal) but identification down to genus and species can be expensive and burdensome. Such detailed identification may be useful if documented evidence is needed to support negligible cross-contamination of cultures during open transfers of multiple processes in the same hood. For any action limit excursion, there is a need to evaluate what processes may have been affected and what was the potential quality impact. Typically, such an event is considered a processing abnormality, and the resulting impact is evaluated by operational personnel in conjunction with relevant quality groups.

Corrective actions implemented in response to an excursion might include improved cleaning regimens, rotation of disinfectants (88), altered sanitization frequencies, monitoring of ultraviolet light intensities, and revised gowning procedures for the area. Testing of disinfectant effectiveness (e.g., disinfectant type and contact time) against a challenge organism also might be useful in troubleshooting reasons for elevated bioburden levels (Table 4). Documented retraining or simple observation of operating personnel using the area, sampling personnel, or sample analysis personnel also may be useful. Additional testing to determine if the action limit value is a one-time-only or reproducible occurrence is key to the investigation. Such testing may consist of repeated sampling of the same locations or increased sampling locations vertically and horizontally within the area to localize the problem. Often quality groups may not support additional sampling because of retesting concerns, but such sampling may be the only way to effectively and scientifically troubleshoot a problem. If the additional sampling plan is defined before execution, then randomness can be minimized. The initial action or alert limit value should not be discarded; rather, all the data should be considered together in evaluating the most probable reason for the excursion.

Constant Temperature Units

Validation testing of constant temperature units centers on documentation of uniform temperature distribution and required humidity level throughout the chamber. Testing is relatively straightforward and inexpensive. Although the requirements for uniformity depend on the type of unit and its particular application, typical examples of set-point deviations might include ± 1.0 °C for an incubator, ± 2.0 °C for a refrigerator, and ± 10 °C for a low-temperature freezer. Testing with the unit loaded can identify problem areas such as reduced air circulation in a low-temperature freezer and excessive heating from a rotary shaker motor in an incubator cabinet or room. Portions of the unit not satisfying the testing requirements can be excluded from use of the process unit if repairs are not made.

Product Changeover

The purpose of a documented product changeover procedure is to ensure that equipment, personnel, and facility

Table 4. Disinfectant Effectiveness Against Microbial Challenge

Disinfectant (dilution ratio, v/v)	1 min	5 min	10 min	45 min
Staphylococcus epidermis				
Klenzyme (1:64)	520	480	425	455
CIP100 (1:1280)	420	490	505	375
CIP100 (1:11)	345	345	295	245
Vesphene (1:128)	<10	<10	0	0
LpH (1:256)	25	<10	0	0
3M Compublend (1:2)	<10	<10	0	0
Aspergillus niger				
Klenzyme (1:64)	48	43	38	46
CIP100 (1:1280)	57	55	46	47
CIP100 (1:11)	34	26	30	27
Vesphene (1:128)	0	0	0	0
LpH (1:256)	0	0	0	0
3M Compublend (1:2)	0	0	0	0

are ready for the next product (89). The major expense of time, money, and effort is expended in preparing the equipment. Equipment cleaning is evaluated using equivalency testing to ensure that existing procedures can clean the next product, with revalidation recommended if major changes in procedures are required. Disposables that may be difficult to clean, such as tubing, filters, or gaskets, may be replaced, although the mechanical support necessary to replace every diaphragm in product contact may be excessive. An evaluation of existing equipment SIP/CIP cycles, autoclave, and glasswasher load patterns, as well as manual equipment-washing procedures, to determine sufficiency for the new process should be conducted. This review should be followed by a cataloging of any changes in equipment, particularly those necessitating revalidation, that might be needed for the next process. The facility preparation spans from the removal of raw materials and equipment used in the previous process from the area to updating or revision of batch records and SOPs. Personnel need to be trained for the new process using batch sheets, newly developed load patterns, SIP or CIP cycles, and altered SOPs.

Surface testing of equipment and the facility might be conducted to document the removal of prior product, particularly for areas in which open transfers are to be conducted. Surface testing after the first run of the new process (typically a practice or engineering run) might be conducted to document the ability of existing procedures to acceptably clean a new product from the area. This minimized testing may only be appropriate if (*1*) similarity of processing steps and their components (e.g., media, host cell organism, contaminating organisms) can be scientifically justified, with the ability to do this successfully varying according to the processing step; (*2*) similarity with respect to TOC recoveries either by comparison testing or scientific justification is documented; and (*3*) similarity with respect to cleanability by comparison testing or scientific justification is documented.

INVESTIGATIONS OF NONCONFORMANCES

There is a tension between facility productivity and possibility of nonconformance. The results of nonconformance can be ranked according to their effect on productivity, schedule, manpower, maintenance, and capital or retrofit costs. As written in the BPC guide for inspections (90), the recommended starting point for inspectors is to review product failures (and by extension failed batches), excursions in monitoring results, and failed validation testing. It is thus of utmost importance to determine the most probable cause of failure, complete any necessary investigation reports, and document any resulting corrective actions. Guidelines for investigation of testing and validation failures are available (60,91–93). Pressure to obtain a most probable or assignable cause should not be imposed, and the cause should be permitted to be listed as "unknown" if truly no one cause emerges as most likely during the investigation. A higher-than-normal percentage of unknown assignments then might prompt reevaluation of investigative or testing procedures and eventually aid in obtaining the real trends in assignable causes for the facility. Failures believed to be caused by poor equipment design, inadequate procedures, or suboptimal process development should precipitate the appropriate alterations, which should be evaluated according to established change control procedures. Investigations are always ongoing, and, although hypotheses can be formed quickly within the short window permitted for an investigation, the accuracy of any hypothesis can be tested and revised in the weeks and months that follow. Recognition of the scientific nature of any investigation of a nonconformance should be featured prominently in any investigation SOPs.

BIBLIOGRAPHY

1. K. Chapman, *Pharm. Technol.* **15**, 82–96 (1991).
2. C. Tayler, *Pharm. Eng.* **16**, 50–58 (1996).
3. D. Gold, *PDA J Pharm. Sci. Technol.* **50**, 55–60 (1996).
4. C. Bruch, *Bull. Parenter. Drug Assoc.* **28**, 105–121 (1974).
5. D. Larkin, *Pharm. Eng.* **9**, 27–30 (1989).
6. Y. Chiu, *Pharm. Technol.* **12**, 132–138 (1988).
7. C. Martinez, *Pharm. Eng.* **14**, 8–14 (1994).
8. M. Beatrice, *Pharm. Eng.* **10**, 29–35 (1991).
9. D. Hill and M. Beatrice, *Pharm. Eng.* **9**, 35–41 (1989).
10. D. Hill and M. Beatrice, *Biopharm. (Jilove)* **2**, 20–26 (1989).
11. D. Hill and M. Beatrice, *Biopharm. (Jilove)* **2**, 28–32 (1989).
12. T. Schoemaker, *Pharm. Eng.* **11**, 23–27 (1991).
13. *Guide to Inspection of Bulk Pharmaceutical Chemical Manufacturing*, CDER, FDA, Washington, D.C., November 1987.
14. H. Avallone, *Pharm. Eng.* **9**, 40–48 (1989).
15. *Interferon Test Procedures: Points to Consider in the Production and Testing of Interferon Intended for Investigational Use in Humans*, Office of Biologics, FDA, Washington, D.C., 1983.
16. *Supplement to the Points to Consider in the Production and Testing of New Drugs and Biologicals Produced by Recombinant DNA Technology: Nucleic Acid Characterization and Genetic Stability*, CBER, FDA, Washington, D.C., April 1992.
17. *Points to Consider in the Characterization of Cell Lines Used to Produce Biologicals*, CBER, FDA, Washington, D.C., May 1993.
18. *Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products for Human Use*, CBER, FDA, Washington, D.C., 1994.
19. S. Fitzpatrick, A. Ma'ayan, and J. Wagget, *Chem. Eng. Prog.* **86**, 26–31 (1990).
20. M. Stern, *Gen. Eng. News* **18**, 10, 32 (1998).
21. L. Little, *BioPharm*, **10**, 8–9 (1997).
22. T. Orr, *Gen. Eng. News* **18**, 12–14 (1998).
23. *Guideline on Sterile Drug Products Produced by Aseptic Processing*, Center for Drugs and Biologics and Office of Regulatory Affairs, FDA, Washington, D.C., June 1987.
24. *Guideline on the Preparation of Investigational New Drug Products*, CDER, FDA, Washington, D.C., March 1991.
25. *Guideline on the General Principles of Process Validation*, Center for Drugs and Biologics, FDA, Washington, D.C., May 1987.
26. *Guideline for Inspections of Biotechnology Manufacturing Facilities*, FDA, Washington, D.C., 1991.
27. *Guide to Inspections of Validation of Cleaning Processes*, FDA, Washington, D.C., 1993.

28. W. Kelley, *Pharm. Technol.* **10**, 48–54 (1986).
29. P. James, *Pharm. Eng.* **18**, 72–82 (1998).
30. L. Angelucci, *Pharm. Eng.* **18**, 40–44 (1998).
31. G. Sofer, J. McEntire, and J. Akers, *Biopharm (Jilove)* **8**, 60–63 (1995).
32. N. Fletcher, *Pharm. Eng.* **9**, 23–27 (1989).
33. A. Signore, *Pharm. Eng.* **18**, 72–82 (1998).
34. D. Tisak and R. Koster, *J. Valida. Technol.* **3**, 394–397 (1997).
35. C. DeSain, *Biopharm (Jilove)* **5**, 22–24 (1992).
36. C. DeSain, *Biopharm (Jilove)* **5**, 30 (1992).
37. W. Wheeler, *Pharm. Eng.* **4**, 48–56 (1994).
38. R.G. Werner, H. Langlouis-Gau, F. Walz, H. Allgaier, and H. Hoffman, *Arzneim.-Forsch.* **38**, 855–862 (1988).
39. J. Akers, J. McEntire, and G. Sofer, *Biopharm (Jilova)* **7**, 8–11 (1994).
40. ISPE San Francisco/Bay Area Chapter, *Pharm. Eng.* **18**, 8–24 (1998).
41. J. Adamson, *Pharm. Eng.* **12**, 16–22 (1992).
42. J. Agalloco and J. Akers, *J. Pharm. Sci. Technol.* **51**, S1–S23 (1997).
43. J. Bauers and J. Hargroves, *Pharm. Eng.* **16**, 36–42 (1996).
44. J. Alford and F. Cline, *Biopharm (Jilova)* **3**, 32–36 (1990).
45. Guidelines for research involving recombinant DNA molecules: notice, *Fed. Regist.* **51**, 16958–16985 (May 7, 1986).
46. Recombinant DNA research: actions under the guidelines; Notice, *Fed. Regist.* **56**, 33174–33183 (July 18, 1991).
47. J. Van Houten and D.O. Fleming, *J. Ind. Microbiol.* **11**, 209–215 (1993).
48. J. Van Houten, in M. Ladisch and A. Bose eds., *Harnessing Biotechnology for the 21st Century: Proceedings of the Ninth International Biotechnology Symposium*, American Chemical Society, Washington, D.C., 1992, pp. 415–418.
49. M. Hamers, *Biotechnology* **11**, 561–570 (1993).
50. G. Muijs, in C. Collins and A. Beale eds., *Safety in Industrial Microbiology and Biotechnology*, Butterworth-Heinemann, Oxford, 1992, pp. 214–238.
51. A. Shahidi, R. Torregrossa, and Y. Zelmanovich, *Pharm. Eng.* **15**, 72–83 (1995).
52. Y. Chisti, *Chem. Eng. Prog.* **88**, 80–85 (1992).
53. F. Bader, in A. Demain and N. Solomon eds., *Manual of Industrial Microbiology*, American Society for Microbiology, Washington, D.C., 1986, pp. 345–362.
54. M. Akers, I. Attia, and K. Avis, *Pharm. Technol. Int.* **1**, 45–49 (1978).
55. M. Enzinger, A. DeVisser, and M. Townsend, *Pharm. Technol. Int.* **1**, 41–43 (1978).
56. M. Enzinger, M. Goodsir, M. Korozynski, F. Parham, and M. Schneier, Technical Monograph No. 1, Parental Drug Association, Inc., Philadelphia, Pa., (1978).
57. *Guidance for Industry for Submission Documentation for Sterilization Process Validation*, CDER, FDA, Washington, D.C., November 1994.
58. M. Reisman, *Pharm. Cosm. Qual.* 22 (1998).
59. D. Berman, T. Myers, and C. Suggy, *J. Parenter. Sci. Technol.* **40**, 119–121 (1986).
60. R. Borghese, *Biopharm (Jilova)* **6**, 28–32 (1993).
61. R. Manfredini, L. Golzi Saporiti, and V. Cavallera, *Proc. Conf. Chem. Biotechnol. Biol. Active Nat. Prod.* 1981, pp. 320–391.
62. C. Perkowski, *Biopharm (Jilova)*, 62–65 (1987).
63. C. Perkowski, *J. Parenter. Sci. Technol.* **44**, 113–117 (1990).
64. M.C. Sharma and A.K. Gurtu, in S. Neidleman and A. Laskin eds., *Advances in Applied Microbiology*, vol. 39, Academic Press, New York, 1993, pp. 1–27.
65. B. Junker, Z. Mann, P. Gailliot, K. Byrne, and J. Wilson, *Biotechnol. Bioeng.* **60**, 580–588 (1998).
66. K. Jenkins, A. Vanderwielen, J. Armstrong, L. Leonard, G. Murphy, and N. Piros, *PDA J. Pharm. Sci. Technol.* **50**, 6–15 (1996).
67. Y. Christi and M. Moo-Young, *J. Ind. Microbiol.* **13**, 201–207 (1994).
68. R. Wisiniewski and C. Burman, in A. Lubiniecki and S. Vargo eds., *Regulatory Practice for Biopharmaceutical Production*, Wiley-Liss, New York, 1994, pp. 407–445.
69. N. Adner and G. Sofer, *Biopharm (Jilova)* **7(3)**, 16–21 (1994).
70. R. Seely, H. Wight, H. Fry, S. Rudge, and G. Slaff, *Biopharm (Jilova)* **7**, 41–48 (1994).
71. G.L. Fourman and M.V. Mullen, *Pharm. Technol.* **17**, 54–60 (1993).
72. S. Lombardo, P. Inampudi, A. Scotton, G. Ruezinsky, R. Rupp, and S. Nigam, *Biotechnol. Bioeng.* **48**, 513–519 (1995).
73. C. Amador, P. Deschenes, M. Dorminy, B. Dudley, B. Gordon, C. Lampe, R. McGarrah, J. Prulello, and R. Wild, *J. Parenter. Sci. Technol.* **44**, 324–331 (1990).
74. D. Held, R. Mehigh, C. Wooge, S. Crump, and W. Kappel, *Biopharm (Jilova)* **10**, 32–37 (1997).
75. P. McArthur and M. Vasilevsky, *Pharm. Eng.* **15**, 24–31 (1995).
76. P. Inampudi, S. Lombardo, G. Ruezinsky, T. Baltrus, J. Dugger, P. Remsen, R. Rupp, and S. Nigam, *Ann. N.Y. Acad. Sci.* **782**, 363–374 (1996).
77. J. Geigert, R. Klinke, K. Carter, and A. Vahratian, *PDA J. Pharm. Sci. Technol.* **48**, 236–240 (1994).
78. T. Macek, in G. Phillips and W. Miller eds., *Industrial Sterilization*, Duke Univ. Press, Durham, 1972, pp. 19–34.
79. J. Brantley and J. Martin, *Gen. Eng. News* **17**, 24, 40 (1997).
80. USP-NF Fifth Supplement, 3442–3444, 3464–3467, 3482–3483, 3547–3555 (1996).
81. W. Collentro, *Ultrapure Water* **11**, 22–29 (1994).
82. F. Riedewald, *Pharm. Eng.* **17**, 8–20 (1997).
83. M. Mittelman, *Microcontamination* **3**, 51–55, 70 (1985).
84. J. Lukanich, *Chem. Process.* **61**, 78 (1998).
85. N. Roscioli, C. Renshaw, A. Gilbert, C. Kerry, and P. Probst, *Biopharm (Jilova)* **9**, 32–40 (1996).
86. J. Barta, *Biopharm (Jilova)* **11**, 84 (1998).
87. B. Immel, *Biopharm (Jilova)* **11**, 46–49 (1998).
88. D. Conner and M. Eckman, *Pharm. Technol.* **16**, 148–160 (1992).
89. J. Odum, *Pharm. Eng.* **15**, 8–20 (1995).
90. *Guide to Inspections of Bulk Pharmaceutical Chemicals*, CDER, May 1994.
91. S. Kuwahara, *Biopharm (Jilova)* **9**, 24–29 (1996).
92. S. Kuwahara, *Biopharm (Jilova)* **9**, 40–45 (1996).
93. J. Sandberg, *Med. Device Diagn. Ind.* **17**, 102–106 (1995).

See also ANTIBODY PURIFICATION; ASTM STANDARDS FOR BIOTECHNOLOGY; CLEANING, CLEANING VALIDATION.

HEATING, VENTILATION, AND AIR-CONDITIONING

DENNIS DOBIE
Fluor Daniel
Marlton, New Jersey

KEY WORDS

Air changes
cGMP
Classification
Containment
Control system
Ductwork
HEPA
Room seals
Space pressurization
Terminal control units

OUTLINE

INTRODUCTION

Biological products are derived from living sources, which presents special challenges for a heating, ventilation, and air-conditioning (HVAC) system serving a facility that processes and handles these products. The most important challenge is to prevent microbial contamination, which to a large measure is controlled by the HVAC system by utilizing classified spaces, providing containment, and main-

taining pressurization. Biological facilities are subject to federal regulation. The Food and Drug Administration's (FDA's) Center for Biologic Evaluation and Research (CEBR) is the primary regulating agency in the United States and its principal concern is the prevention of cross-contamination between different products produced in a multiproduct facility (1). Bioprocess procedures are generally conducted in environments that are cleaner than standard buildings. The HVAC system assumes a large part of the responsibility in maintaining these clean environments. Air used to condition a space must be highly filtered, properly directed, and able to maintain pressure relationships within and between adjacent spaces.

HVAC DESIGN PROCESS

Initial Effort

The HVAC design process begins with meetings with process engineers, architects, and representatives from the owner or facility user. The process and instrument diagrams (P&IDs) are reviewed, and a general understanding of the process is conveyed to all interested parties. Operation of the facility is reviewed, and any plans for future additions or modifications are discussed.

Basis of Design

After the initial meeting, a written basis of design is produced that describes the regulations and codes that will govern the design. Spaces are defined by function, and temperature and humidity requirements are determined. Room classifications are listed and adjacency of spaces and pressure relationships are documented. Any unusual or unique facility requirements must also be designed into the HVAC system at this time, such as emergency backup or redundancy for HVAC systems. This is also the stage of the design process during which alternate studies are conducted to compare options for the HVAC system. The cost of a backup or redundant HVAC supply system may be compared with the cost of product loss or experiment interruption should temperatures or airflow go out of control or specification. Heat recovery from exhaust systems and thermal storage are examples of other potential areas for study (2).

Airflow diagrams are produced that show areas served by a particular air handling system including supply, return, exhaust, and transfer air between spaces. The basis of design also describes major equipment to be used and the level of quality of components and construction material.

Cost Estimate

From the basis of design and a plan of the area, a factored estimate can be developed for budgeting purposes. Preparation of the estimate usually includes a dollar per cubic feet per minute (CFM) quote or a budget quote from vendors, as well as previous history to determine air handling unit costs. Estimates are also made for ductwork, as derived from pounds per CFM history and historic cost data. Costs for fans, high-efficiency particulate air (HEPA) filters, terminal boxes, and building management systems are also included. The level of effort expended to produce the basis of design will affect the accuracy of the estimate. Estimates from a basis of design can range from 15 to 25% accuracy. The estimate at this stage of the project is usually developed for capital appropriation.

Document Preparation

After owner or user approval of the basis of design, final plans and specifications are prepared using manual or computer-aided design tools for preparation of plans and sections. The utilization of three-dimensional design systems for document preparation is becoming increasingly popular. In a three-dimensional system, an electronic model is built that includes all disciplines, including HVAC, architectural, structural, electrical, piping, equipment, and so on. From the model interference, checks can be made that indicate physical conflict. These conflicts are resolved in the design process, avoiding potentially expensive field modifications and possible project delays.

Bidding and Award

When plans and specifications are completed they are sent to HVAC contractors for bid and eventual award. Design and build is also becoming a very popular option, in which design and construction are performed by a single entity with potential cost and schedule savings (3). When awarding a contract the engineer usually meets with the contractor to review the bid for completeness and technical accuracy, and to review the contractor's history on previous projects performing similar work. Since this is specialized HVAC work, it is necessary to have a contractor with sufficient resources and similar experience to avoid delays and potential rework that could increase costs.

REGULATION AND CODE CONSIDERATION

Code Compliance

Codes affect separation of air systems, gowning of people, airflow rates, exhaust requirements, construction materials, and many other aspects of an HVAC system for bioprocess facilities. If codes are not met, the result can be a danger to people working in the facility, products that do not meet specifications, and failure to license the facility.

Classified Spaces

Clean rooms provide clean air in aseptic spaces and are defined or classified in the United States by Federal Standard 209E of September 1992. The European Economic Community (EEC) published guidelines called "Guide to Good Manufacturing Practice for Medical Products in the European Community," which are more stringent than U.S. FDA regulations (4). Room classifications are established by measurement of the number of particles 0.5 μm and larger that are contained in 1 ft^3 of sampled air. Generally class 100 to 100,000 rooms are used in the bioprocess industry. Rooms may be classified as clean at class 1 or 10 for other applications, particularly in the microchip indus-

try. Table 1 (derived from Federal Standard 209E) shows the air cleanliness classes.

EEC guidelines classify spaces by alpha character, as indicated in Table 2, and Table 3 shows the U.S. and EEC comparison.

Laboratory Spaces

Laboratory spaces are associated with bioprocess facilities. These classifications range from BL-1 through BL-4, with BL-1 being the least hazardous. BL-1 compares to a school science laboratory, while BL-4 labs must provide ultimate containment and require glove boxes and double HEPA filtration. These laboratory classifications were formulated in the United States by the National Institutes of Health (NIH) and are published as guidelines in the *Federal Register*. The guidelines contain requirements for filtration of air, recirculation of air, containment, and biosafety cabinet requirements. There are various classifications of biosafety cabinets, depending on the level of containment required. The American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Inc. (ASHRAE) guide should be consulted and owner requirements considered when selecting the required class and type of cabinet (5).

Other Regulations

Other regulations and guidelines may include those published by insuring agencies and corporate or plant guidelines formulated from specific industry experience and current good manufacturing practices (cGMPs) as published in the *U.S. Code of Federal Regulations* parts 210 and 211. The cGMPs apply to manufacture, holding, or processing of drugs to ensure safety, purity, and quality of the product. cGMPs regulate space temperature, and humidity and require a pressure differential between adjacent rooms of different cleanliness levels and a minimum air change rate of 20 air changes per hour. The regulation also addresses class 100 requirements at filling lines and microbiological testing sites, in addition to laminar flow control equipment parameters and testing at fill lines. Good practice usually places class 100 areas in class 10,000 backgrounds, which have between 35 and 60 air changes per hour in the space, depending on the activity and particulate generation within the space. Walls, floors, and ceilings for cGMP areas are to be constructed of smooth, cleanable surfaces, impervious to sanitizing solutions and resistant to chipping, flaking, and oxidizing. Horizontal ducts, pipes, or crevices that foster dust accumulation and cannot be easily cleaned are not permitted.

Figure 1 is a sample graphic depiction of an air classification drawing that is prepared to communicate to the design team and reviewing agencies the classification of each area. This is usually a full-size drawing (scale: 1/8 ft = 1 ft) and uses various shading symbols to identify each classified area.

TEMPERATURE, HUMIDITY, AND AIRFLOW

Space Temperature

Generally areas are designed to provide room temperatures from 67 and 77 °F (19 and 25 °C) with a control point of 72 °F (22 °C). Control of humidity from 40 to 55% is necessary for personal comfort, to prevent corrosion, to control microbial growth, and to reduce the possibility of static electricity. Calculations are made to determine internal heat gains from lights, people, and equipment in addition to transmission gains from adjacent spaces to determine if supply airflow quantities are adequate to provide design space temperature conditions. These calculations are compared with airflow quantities required by classification to establish the minimum air required to satisfy both the space cooling load requirements and air cleanliness classification. Fan heat from recirculating fans can also be a large heat contributor in clean spaces.

Heat-loss calculations must also be made to determine heat loss through walls, roof, and floor. It may be necessary to add additional heat to a reheat coil to overcome these losses. No credit should be taken for process heat gain in this calculation, since the process could be dormant and the space would still need to be maintained at proper temperature. Lower space temperatures may be required where people are very heavily gowned and would be uncomfortable at "normal" room conditions. Humans are generally the largest contributors to airborne particulate in clean spaces; personal discomfort may cause perspiration, which significantly adds to the generation of airborne particulate. To achieve colder room conditions, the temperature of the supply air to the rooms may have to be colder than the normal 50–55 °F (10–13 °C), resulting in a greater dehumidification demand on the air-conditioning system. The amount of dehumidification required can be deter-

Table 1. Airbone Particulate Cleanliness Classes

Class		Measured particle size (μm)									
		0.1		0.2		0.3		0.5		5	
SI	English	m^3	ft^3	m^3	ft^3	m^3	ft^3	m^3	ft^3	m^3	ft^3
M 1.5	1	1,240	35	265	7.5	106	3	35.5	1	—	—
M 2.5	10	12,400	350	2,650	75	1,060	30	353	10	—	—
M 3.5	100	—	—	26,500	750	10,600	300	3,530	100	—	—
M 4.5	1,000	—	—	—	—	—	—	35,300	1,000	247	7
M 5.5	10,000	—	—	—	—	—	—	353,000	10,000	2,470	70
M 6.5	100,000	—	—	—	—	—	—	3,530,000	100,000	24,700	700

Note: Limits listed are maximum concentrations of particle sizes equal to and larger than the measured particle sizes shown.

Table 2. Classification of Clean Areas Defined by the European Economic Community

Grade	Maximum permitted number of particles per m^3 equal to or above 0.5 μm	Maximum permitted number of viable microorganisms per m^3
A	3,500	<1
	Laminar airflow work station	
B	35,000	5
C	350,000	100
D	3,500,000	500

- Laminar airflow speed: 0.3 m/s for vertical flow 0.45 m/s for horizontal flow.
- Air changes per hour for grades B, C, and D should be higher than 20.
- U.S. Federal Standard 209E corresponding particle counts are
  - Class 100 (Grades A and B)
  - Class 10,000 (Grade C)
  - Class 100,000 (Grade D)
- It is accepted that sometimes conformity with particle counts may not be met at point of fill due to generation of particles or droplets from the product itself.

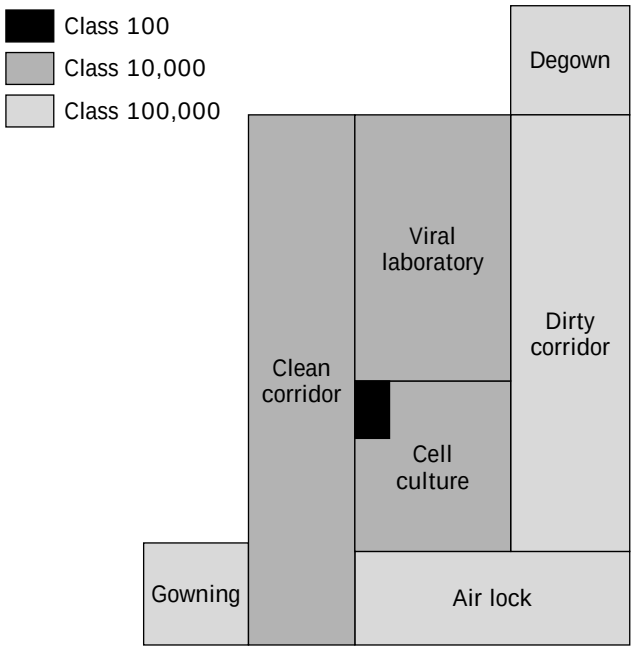


Figure 1. Air Classification Drawing.

mined by use of a standard psychrometric chart by plotting room, mixed air, outdoor air, and coil conditions. Formulas to determine cooling coil loads are available from most coil manufacturers and air handling unit vendors (6). Cooling coil duty in British thermal units per hour (Btuh; *W*) is determined from the enthalpy difference between entering and exiting cooling coil air dry and wet bulb conditions.

Space Humidity

In some cases where products sensitive to moisture are handled, the room relative humidity requirement may be as low as 15 to 20% and may require the use of chemical dehumidifiers. It is necessary to establish this criteria early in the design process to allow selection of dehumid-

Table 3. Air Cleanliness Classification for Aseptically Produced Products

Operation	Parameters	United States	European Economic Community
Nonsterilized product or container (U.S.: controlled area)	Clean room type	Class 100,000	Grade C (Class 10,000)
	Maximum particle size 0.5 μm or above	100,000/ft^3 (3,531,000/m^3)	350,000/m^3 (9,900/ft^3)
	Maximum viable organisms	2.5/ft^3 (88.3/m^3)	100/m^3 (2.8/ft^3)
	Airflow rate	Minimum 20 air changes/h	Higher than 20 air changes/h
	Space pressurization	0.05 in water	Positive
Sterilized product or container (U.S.: critical area)	Clean room type	Class 100 environment in class 10,000 background	Grade A (class 100) environment in grade B (class 100) background
	Maximum particle size 0.5 μm or above	100/ft^3 (3,531/m^3)	3,500/m^3 (99.1/ft^3)
	Maximum viable organisms	0.1/ft^3 (3.53/m^3)	Class 100 room with grade A area: 1/m^3 (0.03/ft^3) Grade B background 5/m^3 (0.14/ft^3)
	Airflow rate	90 ft/min ± 20% (45.7 m/s)	Grade A: laminar flow work station Vertical: 0.3 m/s (59 ft/min) Horizontal: 0.45 m/s (88.6 ft/min) Grade B: higher than 20 air changes/h
	Space pressurization	0.05 in water	Positive

ification equipment and add allowances to the construction-cost budget estimate. The lower the required temperatures and humidity, the more sophisticated the HVAC equipment becomes, resulting in higher start-up and operating costs. Range of temperature control also affects the cost of the control and HVAC delivery system. The closer the space temperature tolerances are, the more sensitive the controls must be to react. To achieve close temperature tolerances, the air distribution system must be more extensive and elaborate to prevent hot or cold spots. An HVAC room balance, shown in Table 4, is a sample of a preliminary design document that indicates room temperature and humidity design conditions, in addition to dimensions. Table 4 also lists room classification air changes per hour, supply, exhaust, return, and transfer air requirements. This information is useful in communicating conditions to the design team and reviewing agencies. The table should be kept up to date, because it is useful to installing contractors and the construction supervision team as a check on ultimate system performance. When the project is near completion this information is usually transferred to the construction documents for ease of validation and agency approval.

Calculation of Air Changes

Table 4 indicates a typical range of air change rates generally used to achieve the desired room cleanliness classifications and to meet federal and local regulations. These air change rates vary widely in actual practice due to the level of activity, number and type of particulate generators in a room (such as people and equipment), and room size and quality of air distribution. It is generally best to use historic data to establish airflows, which is usually done with significant input from the owner based on past experience or preference. There is nothing sacred about an air change rate as long as minimum airflow rates required by code are maintained. The goal is to achieve desired particulate cleanliness levels and stay at or above a 20 air changes/h minimum.

To calculate airflow requirements from room air change rates, the following formula should be used:

$$\text{air flow (CFM)} = \frac{\text{room volume (ft}^3\text{)} \cdot \text{Number of air changes/h}}{60}$$

or (metrically)

$$\text{air flow (m}^3\text{/h)} = \frac{\text{room volume (m}^3\text{)} \cdot \text{Number of air changes/h}}{60}$$

Adjacency

In the early stages of design, the HVAC engineer must work with the architect, process engineer, user, and layout planner, if one is involved, to establish adjacency of spaces. It is suggested that a computerized drafting system with block layouts be used to allow the frequent manipulation that is required to arrive at an optimum arrangement. During this exercise the layout is primarily determined by process function and spaces are classified as clean, dirty, or contained. A contained space is used to contain a product or material that could be spilled into an occupied space or find its way into the building air system. From an HVAC standpoint it is desirable to keep similarly classified areas as physically close to each other as possible so they can be connected to the same air handling system, thereby minimizing duct runs, cost, and air system complexity. It is also imperative that spaces be arranged to allow people to move around without disrupting the cleanliness or containment of the spaces.

Ultimately, as the layout develops, the suites or groups of spaces will be established and groupings of rooms can be made for assignment to separate air handling systems. It is not desirable to mix dirty and clean systems or suites that may allow the possibility of cross-contamination from one suite to another. Leaks can develop in a filter, or some source of contamination could find its way through the air

Table 4. HVAC Room Balance Sheet

Room name and number	Room temperature	Relative humidity (%)	Area (ft^2)	Room height (ft)	Air changes per hour	Hood, general exhaust	Required return CFM	Total supply CFM	Transfer air	
									Out	In
Viral lab #1—100 (BL2, class 10K)	68	50	544	9	40	1,230	1,984	3,264	170	120
C/C lab #1—101 (BL2, class 10K)	68	50	453	9	40	345	2,323	2,718	170	120
Gowning 102 (BL2, class 10K)	68	50	125	9	40	0	385	750	365	0
Viral lab #2—103 (BL2, class 10K)	68	50	544	9	40	1,230	1,984	3,264	170	120
Clean corridor 104 (BL2, Class 10K)	68	50	1,300	9	40	0	6,905	7,800	1,320	425
Dirty corridor 105 (BL2, class 100K)	68	50	700	9	20	0	3,300	2,100	0	1,250
Degown airlock 106 (BL2, class 100K)	68	50	80	9	20	0	480	240	0	240

supply or return systems, providing a source for cross-contamination.

PRESSURIZATION

Requirement and Theory

Pressurization of spaces is required to keep products from being contaminated by particulate or to protect people from contact with harmful substances by physical means or inhalation. Where very severe potential exposure exists, air packs and breathing apparatus are employed.

Pressurization can best be explained by thinking of air leaking from a balloon. Pressure is built up inside the balloon and air escapes from a small leak in the balloon. If the hole stays the same size, a greater pressure differential between the inside and outside of the balloon will allow the air to escape at a higher velocity, which can be termed a higher level of pressurization. As the air in the balloon is depleted, the pressure differential drops, the rate of air escaping diminishes, and the velocity decreases. The same principle applies in bioprocess facilities where rooms are made as tight as possible to contain air, allowing the air handling system to build up or reduce pressure in rooms. These conditions are maintained by airflow, or leakage, between adjacent spaces caused by pressure differentials created by the air supply, return, and exhaust systems. Where major demarcations of pressure are required, air locks are used. These are small rooms with controlled airflows acting as barriers between spaces. Standard 209E recommends maintaining pressures of 0.05 in water (12 Pa) between adjacent spaces with doors closed. During facility operation as doors are opened the design differential is greatly reduced, but air must continue to flow from the higher to lower pressure space, even though at a reduced flow rate. To maintain a differential of 0.05 in water (12 Pa), a velocity of approximately 900 ft/min (4.7 m/s) should be maintained through all openings or leaks in the room, such as cracks around door openings. In theory the actual required velocity is less, but in actual practice it is prudent to use 900 ft/min. Table 5 gives the velocity pressures airflows from 0.01 to 0.20 in water (2.4–48 Pa), which can be used in approximating airflow between rooms. Limits on pressure drops must be set during the design (usually 0.05 in water) and observed. The pressure differential exerts a force on the door that can be calculated. If the force is too great (0.15 in water/36 Pa), the door may not close fully or may be difficult to open. This is particularly important in large complex facilities where many levels of pressurization may be required. Many facilities now use sliding doors, and it is essential that the seals be carefully designed to allow minimum leakage and proper containment or pressurization.

Table 5. Velocity Pressure to Velocity Conversion: Standard Air

VP	V	VP	V	VP	V	VP	V
0.01	401	0.06	981	0.11	1328	0.16	1602
0.02	566	0.07	1060	0.12	1387	0.17	1651
0.03	694	0.08	1133	0.13	1444	0.18	1699
0.04	801	0.09	1201	0.14	1499	0.19	1746
0.05	896	0.10	1266	0.15	1551	0.20	1791

Note: VP, velocity pressure; V, velocity.

Pressurization Diagram

Figure 2 indicates levels of higher to lower pressurization with arrows showing the direction of airflow, (leakage) between the spaces. In actual practice the calculated airflow quantities are approximate since in constructing a facility it is very difficult to build everything absolutely airtight. A very small unexpected leak or a seal that is not as tight as expected in design can affect the pressurization and airflow required to maintain the design value of differential. Some flexibility must be built into the air handling system capacity and controls to compensate for these variations.

One technique to establish pressurization in separate rooms is to have a common reference for the control sensors. This common reference is usually a mechanical room or an interstitial space where the pressure is not affected by changes in the controlled spaces. It may be necessary to build an elaborate zero pressurization chamber to house the common sensor to negate the effects of wind or cross-currents. Referencing spaces to each other can lead to problems if doors are frequently opened and upsets in pressure occur. The controls may never settle down and tend to constantly modulate, a problem termed *hunting*.

Room Seals and Doors

In most facilities the openings around the doors between rooms are where leaks occur due to pressure differentials between rooms. In making rooms tight any room openings

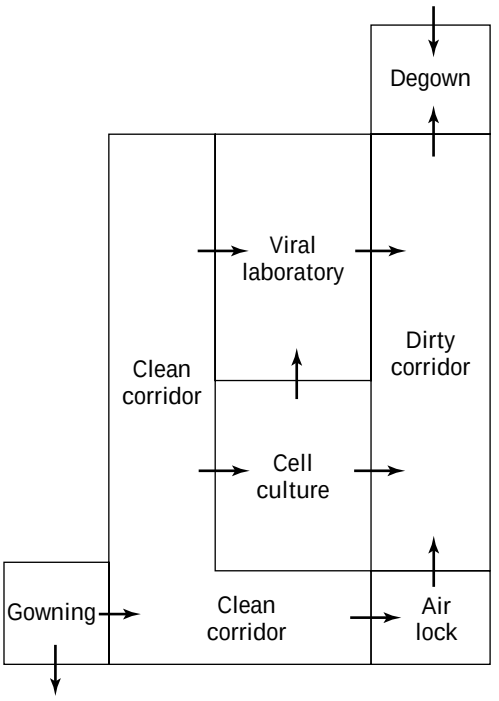


Figure 2. Pressurization Diagram.

must be sealed with a proper sealant that will not promote growth of organisms and can be easily cleaned. Areas to be sealed include ceiling tiles, lighting fixtures, pipe penetrations, telephone outlet penetrations, and any cracks or openings that appear in the structure. A typical door would have the following dimensions and crack area at the perimeter: door size, 3 ft (0.915 m) wide by 7 ft (2.41 m) high; cracks at top and sides, 1/8 in (3.2 mm) with an undercut of 1/4 in (6 mm). The calculated area around the door is equal to 0.24 ft^2 (0.024 m^2). To achieve 0.05 in water (12 Pa) pressure differential across the door, approximately 215 CFM (375 m^2/h) of airflow through the cracks is required. Door seals around the top and sides are usually made of closed cell neoprene and should generally be used to reduce the crack area. To reduce the undercut, a drop-type seal, which is commercially available, should be used. The drop type is preferred to a wipe type, since it will not mar or leave residue on the floor. Air used for pressurization must be accounted for in system calculations. Air through cracks or openings is accounted for as transfer air and shown in the HVAC room balance table.

Of particular importance is exhaust air from equipment and hoods that may be on or off at different times during occupied periods. These variations must be dynamically compensated for to maintain room pressurization. To maintain the required balance, numerous systems are employed using manual and automatic dampers, constant and variable volume air control boxes, and elaborate airflow sensing devices. These components are combined with control systems and sensing devices to ensure that the room pressurization is maintained.

AIR HANDLING SYSTEMS

Constant Volume Systems

After block layouts, adjacency layouts, room design criteria, and pressurization levels have been established, the number and type of air handling systems can be selected. Air handling systems most commonly used are constant and variable volume with and without reheat coils for control. For controlled spaces the most reliable system is a constant volume system with terminal reheat (CVRH). This system is not used in less critical or commercial applications. In a terminal reheat system the air leaving the cooling coil is set at a fixed temperature, and the terminal reheat responds to a space thermostat, turning on heat to satisfy the load. This can waste energy, since air is cooled and then reheated. Reheat systems require energy to both simultaneously heat and cool the air, as can easily be demonstrated on a psychrometric chart. Many energy codes prohibit this practice for comfort applications, however, where close control of temperature and humidity is required for process areas the energy conservation requirement is waived. The advantages of reheat systems are that humidity is always controlled (since dehumidification always takes place at the cooling coil) and each space or zone that needs temperature control can easily be accommodated by adding a reheat coil and thermostat. Another advantage of the CVRH system is that airflow is constant, which makes balancing and pressurization easier to maintain. A reheat system is probably the simplest and easiest of all systems to understand and maintain. Figure 3 is a typical depiction of an airflow diagram for a CVRH system.

Variable Air Volume Systems

A variable air volume (VAV) system is generally used in administrative areas and some storage spaces where pressure control is not critical, humidity control is not essential, and some variations in space temperature can be tolerated. The VAV system works by delivering a constant-temperature air supply to spaces with reductions in airflow as cooling loads diminish. This reduction in air results in a reduced cooling effect, allowing the airflow or cooling to match the actual room load, which provides the design space temperature condition. This eliminates the energy used for reheat and saves fan energy, because the total amount of air moved is reduced. Some form of perimeter heating must be supplied for spaces with exterior walls or large roof heat losses. The perimeter heating can be baseboard radiation or some form of air heating using heating coils. Finned radiation or convection heating devices should not be used in clean spaces, since they are not easily cleaned and allow places for unwanted particulate buildup. Combinations of systems can be used, especially if variable quantities of supply and exhaust air are required for fume hoods or intermittent exhausts.

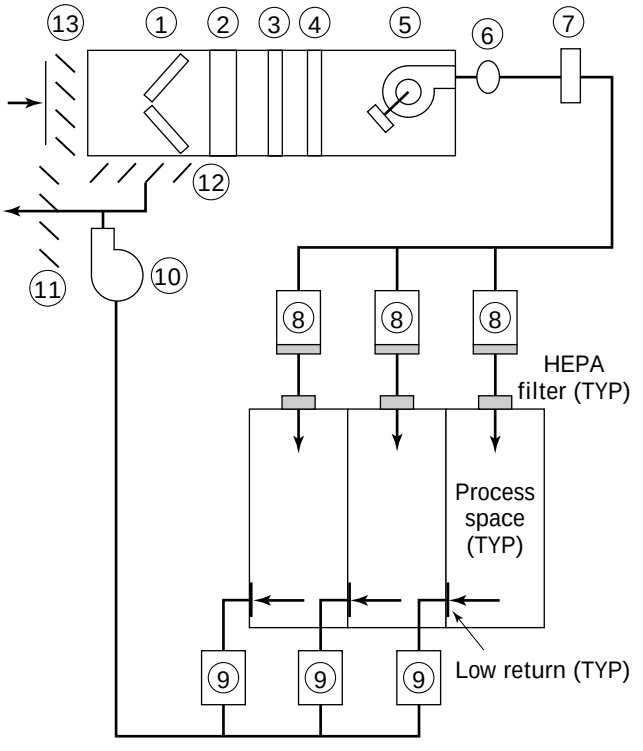


Figure 3. Airflow diagram: 1, 35% filters; 2, 85% filters; 3, preheat coil; 4, cooling coil; 5, fan section; 6, humidifier; 7, air monitoring device; 8, constant volume reheat box; 9, constant volume box; 10, return fan; 11, relief air damper; 12, return air damper; 13, outdoor air damper.

Recirculation Systems

In bioprocess facilities large quantities of air may be required to promote unidirectional flow and air cleanliness. This is particularly true in a class 100 space. In many cases the large quantities of air exceed the requirements for cooling, so it is desirable and possible to recirculate air within the space and only pass enough air through the air handling unit to perform the heating or cooling. In many cases the recirculating systems are supplied as a separate package installation. Numerous manufacturers provide these packages, which include supports, fans, and filters. Some vendors provide fans that are self-contained with each 2-ft by 4-ft filter module.

Airflow Sheets

The airflow sheets should be developed on full-size drawings using computer drafting. They should list each space and show air quantities supplied, returned, and exhausted from each space. They also must show air transferred into and out of spaces, and, while quantities should be shown, they will probably require field modification to attain pressurization. The airflow sheet becomes a useful tool for transfer of information to the owner or user, for agency reviews, for transmission of information to HVAC designers, and for other engineering disciplines. These documents are also invaluable to construction contractors and for system checking by construction managers and balancing contractors. Airflow sheets provide a pictorial description of each air system and show how the elements comprising the system are related.

DEHUMIDIFIERS

Need for Dehumidification

In many cases low relative humidity is required, such as in drying or powder areas where hygroscopic materials are handled and would absorb unwanted moisture from the air. A chemical dehumidifier may be required to obtain the desired low humidity conditions. The need for a dehumidifier can be determined by utilizing a psychrometric chart, which will establish a leaving-coil dew point. Normal chilled water is supplied at 42–44 °F (5.6–6.7 °C) for use in cooling coils. Depending on the number of coil rows and fin spacing, a minimum dew point of about 50–52 °F (10–11.1 °C) can be obtained. This results in a minimum room relative humidity of approximately 50% at 70 °F (21.1 °C). To obtain lower relative humidity conditions chemical dehumidification is employed.

Chemical Dehumidification

Chemical dehumidifiers are commercially available air handling units that contain a sorbent material (desiccant) that can be a solid or liquid. Wet dehumidifiers use absorbents that change physically during the process. Lithium salt solutions are generally used to remove moisture from conditioned air and are then regenerated by heat, usually using a steam heat exchanger. Dry dehumidifiers use adsorbents that do not experience phase changes during the process. Silica gel and activated alumna are generally used. A rotating wheel is commonly used to remove moisture from the conditioned air. The wheel is regenerated by passing heated outdoor air over the wheel to dry it out. Steam or electric coils are usually employed for regeneration. Depending on the amount of dehumidification required and the amount of outdoor air (usually with a high moisture content), it may be best to combine the dehumidifier with a conventional air handling unit and only dehumidify a small portion of the air or just the outdoor air. The dehumidifier has a high initial cost compared with a conventional air handling unit. The size should be optimized to do only the required duty with an appropriate safety factor. Knowledgeable vendors in this specialized area should be consulted to find the best combination of dehumidification equipment, system arrangement, and control for the application. These systems also require considerable physical space, energy consumption, and service—important criteria to be considered in system selection.

HUMIDIFIERS

Air handling systems that bring in quantities of outdoor air for makeup require the addition of moisture to the airstream to provide desired space humidity levels. There are many commercially available devices for providing humidity to air, but the most commonly used are steam grid humidifiers. It is important to use clean steam (7), not plant steam, which may contain boiler chemicals and impurities from deteriorating piping and equipment. Humidifiers are controlled by modulation of a steam valve at the humidifier, and include a chamber to prevent condensation and water droplets in the duct. The valve is controlled by a signal located in the return or exhaust airstream or in a room humidstat. A high-limit stat is placed in the duct downstream of the humidifier to override the controlling stat and prevent condensation in the duct. Placement of the humidifier in the duct is critical and must follow the manufacturer's recommendations to prevent condensation and provide proper dispersion space. Duct downstream of the humidifier is usually constructed of stainless steel to prevent rust or corrosion should water be present.

AIR HANDLING UNIT APPLICATION

Air Handling Unit Construction

Conventional air handling units consist of filters, coils, and fans in a metal casing, with an insulation liner applied to the inside of the casing. For bioprocess applications the unit casing must be a sandwich of metal with insulation between the metal sheets to provide a smooth, cleanable interior surface that does not foster the growth of organisms. Chilled water or propylene glycol solutions are generally used for cooling and dehumidification. Direct expansion refrigerant, in which the refrigerant is in the air unit coil, may be used, but these systems are less reliable than chilled water or glycol and are more difficult to control in the narrow air temperature ranges required. Units should

contain good access doors, view ports, electrical convenience outlets, and interior lighting for maintenance. The casings should be tightly sealed and designed for pressures that are higher than commercial applications due to the generally high system air pressures required for bioprocess applications. All sealants and lubricants exposed to the airstream should be food grade to minimize the chance of air contamination. Units designated as draw-through have the coils on the suction side of the fan. Blow-through units have the coils on the discharge side of the fan and have the advantage of a filter downstream of the coils, reducing potential contamination of the supply duct system. On blow-through units an air distribution plate must be installed to properly distribute air evenly over the filter and coils. It should be noted that if the unit contains a HEPA filter and the air is ducted directly to the clean space without a terminal HEPA filter, the duct should be constructed of a clean, smooth, nonflaking material, usually stainless steel.

Air Handling Unit Location

It is best to locate units indoors for ease of service and to extend the life of the units. They should also be located as close as possible to the main rooms they are serving to minimize larger and longer duct runs. Airflow measuring devices and humidifiers must also be installed, and unit location must allow enough straight run of duct for installation and proper operation of these devices. Location of outdoor air-inlet louvers must be carefully considered. Intakes are generally located on the building sidewall high off the ground to minimize dust intake. Intakes should also be away from truck docks or parking lots, where undesirable fumes and particulate are generated. In locating inlets the prevailing winds should also be considered, and any nearby exhausts or fume concentrations should be avoided to prevent recirculation of exhaust air back into the supply system.

RETURN AND EXHAUST FAN SELECTION AND LOCATIONS

Return Fans

Return and exhaust fans are integral parts of the air handling systems and also must be shown on airflow diagrams. Return fans are used on systems with long duct returns or that have return system pressure drops greater than 0.5 in water (120 Pa). This allows proper total system balance and minimizes suction pressure required from the supply fan. If a return fan is not used, the capacity of the supply fan can be overextended and it may be difficult to limit and properly control the amount of outside air being admitted to the unit. Outside air fluctuations are also more susceptible to exterior wind conditions. Return fans are also needed when required to provide a negative pressure in rooms that require containment. Return fans can be of standard centrifugal type or an in-line type, which works nicely for installation directly into return ducts in crowded equipment rooms. Return fans may also be required to handle varying quantities of air or provide a constant flow of air at varying pressure conditions. To achieve these conditions some form of damper control, inlet vane, or variable frequency drive motor control is generally used.

Exhaust Fans

Building exhausts are generally collected and ducted to exhaust fans in groups or clusters. Exhaust fans should be located as near to the building discharge as possible since this keeps the duct under a negative pressure and any leaks will be into the duct, and not contaminated air from the duct into an occupied space or mechanical room. For this reason roof locations of fans are preferred, even though this may make service difficult in severe weather conditions. When fans are located in mechanical rooms or interstitial spaces, it is essential to tightly seal the discharge duct before it exits the building in a roof vent or wall louver. Roof penetrations should be kept to a minimum to prevent leaks. Fumes and toxic exhausts should be extended through the roof and terminated well above the roof line in a suitable stackhead. Designs for stackheads are found in most duct manuals and are used to direct exhaust vertically at high velocities to enhance dispersion of fumes and prevent water from finding its way into the exhaust system (8). Extremely toxic or dangerous active biological agents may require HEPA filtration or other treatment, such as incineration, before exhaust to the atmosphere.

Redundancy

Another important consideration occurs when return or exhaust fans are used as part of the containment system. Should the fans be essential to maintaining containment, it may be desirable to have a backup fan or redundant system. This is essential if loss of containment can be harmful to humans or would result in an expensive loss of product. Airflow switches, which give a warning in case of fan system failures, are also desirable options for critical systems. The airflow sensing method to prove flow is preferred to an electrical motor indication since the motor could be running with a broken fan belt and the operator would not know that the fan is not moving air.

HEPA FILTERS

HEPA Filter Description

HEPA filters are used in the final cleaning stages of air to remove very fine particles. Rough and intermediate filters that are not as efficient as HEPAs are used to remove larger particulate matter. This prolongs the life of the more expensive HEPAs, significantly reducing the cost of filter element replacement. A HEPA filter, by definition, removes 99.97% of particles of 0.3-μm size. HEPA-type filters called ultraparticulate arrestences (ULPAs) can be even more efficient and are commercially available at a greater cost. Pressure drop across filters is rated for nominal flow at new clean conditions and is usually about 1 in (25.4 mm), and the systems are designed to change out filters at approximately 2 in (50.8 mm) or double the clean ratings, although filters can be operated at higher pressure drops, which some owners prefer. The higher operating pressures

result in a greater use of electrical energy to power the fan to overcome the increased resistance that it must operate against.

Location

HEPA filters are usually located downstream of heating and cooling coils in air handling systems since coils are potential sources of contamination. The most popular HEPA filter location is in the room ceiling using standard laminar flow outlets nominally 24 in by 48 in (0.6 m by 1.2 m). These outlets contain manual control dampers, testing sample ports, a diffusion panel, and a HEPA filter element. They are commercially available in permanent or throwaway units where the entire assembly is disconnected from the flexible duct supply terminal and replaced when the filter is dirty. Sealing of filters to frames is an installation problem and is best solved by using a filter frame with a gel-like seal into which the filter fits. This material does not support biological growth, and these filters are offered by several manufacturers.

Bag-In Bag-Out Housings

Where danger of exposure to filter changing maintenance personnel exists, filters are placed in special enclosures called bag-in bag-out housings. These units are quite expensive and employ elaborate sealing methods and a double bag arrangement that allows changing of the filter element(s) without ever exposing the changer to the dirty filter or its collected contaminant.

Laminar Flow Workstations

HEPAs are also used in class 100 workstations in single or double filtration of air to provide sufficient cleaning before unidirectional (laminar) air distribution to the work space. These workstations are commercially available in horizontal or vertical flow patterns generally recirculating within the clean space. Filters are commercially available in many standard sizes and come with many frame arrangements with different methods used to form a seal between the filter element and the frame.

TERMINAL AIR CONTROL DEVICES

In most bioprocess applications the volume of airflow must be controlled to ensure proper air changes and provide air for pressurization. The choices to control airflow are (*1*) a variable flow responding to some pressure or temperature actuated signal or (*2*) constant volume to provide a uniform quantity of air under varying upstream pressure conditions. Many commercial devices, usually called boxes, or terminal control units, are available that utilize a variety of air dampening designs—from a simple blade type to pneumatic bladder type and conical type. Each type has its own application and associated cost implications. The published damper design and test data for flow and repeatability should be studied against the suspected operation and matched to the proper application. Dampers are controlled from pneumatically controlled operators or electric motors responding to electric or electronic signals. Pneumatic dampers generally react more quickly to changing airflow requirements. Dampers and controls should be selected for linearity of response, repeatability, and percentage of error when compared with desired flow. The damper is generally housed in a device called a volume control box, which contains deflecting baffles and flow-measuring or pressure-sensing devices that are used for flow control. Accuracy of flow control boxes, combined with the controlling system, must be aggregated to ensure that the actual flow will be within acceptable limits to control pressurization. The accuracy of control devices can vary between 5 and 10% of maximum flow. This built-in inaccuracy must be calculated when establishing flow differential rates. For commercial applications, control boxes are lined with insulating acoustical material, which should be avoided on boxes used for critical or controlled areas to avoid dirt catchers or host locations for unwanted growth.

Some vendors offer complete control box systems, including controls, supply and return boxes, and a system to allow tracking. *Tracking* is a term used to describe when the control system monitors the supply flow and automatically adjusts the return or exhaust flow to maintain a fixed CFM differential between supply and return or exhaust. This maintains the space pressure differential.

AIR TERMINAL OUTLETS

Supply Terminals

In clean spaces the desired distribution of air is unidirectional or acts as a piston of air that reduces turbulence and eddy currents in the airflow. This piston of air carries particulate from the ceiling to the floor return and helps to prevent airborne particulate matter from recirculating and contaminating the work space. In most cases it is desirable to recirculate air within a space through a filter since the return air has less particulate than typical outdoor air and does not require extensive heating and cooling. Air terminals should be selected of materials that are nonflaking, nonoxidizing, and are easily wiped clean.

Location

Placement of outlets should be above work areas since the cleanest air should be introduced at these spaces. If the product at the workstation is hazardous to humans the air should be supplied at the operator's back and pulled into the work area to ensure that flow is away from the operator and any particulate will be captured in the airstream. Containment velocity at the inlet of a biosafe hood or enclosure should be about 100 ft/min (0.508 m/s). Airflow near hoods or enclosures should avoid any air disturbance that could cause eddy currents or air turbulence. These currents will upset capture patterns of the hood and make verification of hood face velocity difficult. Many of the cabinet manufacturers have extensively tested capture velocities and provide excellent data on the actual operating conditions. When locating terminals and enclosures it is important to consider movement of personnel within the space and identify frequently opened doors, which also tend to upset airflow patterns. Should the operator pose a threat to the

product, the airflow pattern should be reversed, with the air supplied behind the product and blowing past the product to avoid contamination from the operator.

Return Terminals

Return terminals are also an important consideration and are generally located low in the walls for clean rooms. In class 10,000 to 100,000 rooms low cleanable wall registers are generally used. In cleaner areas low return wall systems, termed *air walls*, are used. The air wall is an almost continuous opening at the base of the wall with the air ducted up in the wall system and collected for return to the air handling system. Air wall inlets are generally located not more than 15 ft (4.5 m) in plain view from a supply terminal to reduce the likelihood of turbulence. Figure 4 shows a typical return air wall detail.

DUCT MATERIALS, PRESSURE, AND CLEANABILITY

Materials of Construction

Unlined galvanized steel ductwork is used in rectangular, round, and elliptical (or flat oval) configurations for the majority of the systems. Because galvanized duct can flake off or rust, it should not be used downstream of the HEPA filters to avoid contamination from the duct system itself. When the HEPA filter is located upstream of the room terminal and a long run of duct is present, the material of choice for the duct is stainless steel, but this is expensive and its use should be minimized. Many systems may also be fumigated or cleaned in place, and the duct material chosen should not be affected by the cleaning agent.

Duct Pressures

Ductwork systems in bioprocess facilities tend to have higher system pressure requirements than commercial systems due to extensive use of filters, volume control devices, and physically complex arrangements. The duct system pressures must be calculated and clearly stated on the contract documents to allow the fabricator to provide the proper metal thickness and construction methods for the required system pressures (9). System pressures will also change as the system is operated with filters that get dirty or space pressure conditions that vary. Duct systems must allow for these pressure fluctuations, and the fans may require speed controls, inlet vanes, or variable pitch blades to match the varying flow and pressure conditions.

Cleanability

Cleanability of duct systems is important to ensure that if an installed system gets dirty or contaminated it can be cleaned. In the design stage care must be taken to locate access doors in the duct, where they can be easily reached without compromising a process or violating a controlled space. It is best to minimize entry of facility system maintenance personnel into the clean spaces. All sealed duct shipped to the site should have only end seals broken, and then quickly resealed, during final installation. In very critical applications the duct is factory cleaned and sealed before shipment to the site. This step removes the oil and other contaminants present during duct construction but is expensive. It may be difficult to find sheet-metal fabricators willing to do this work, since they are not always set up for such procedures and it makes shipping the completed product difficult.

SYSTEM OPERATING PROCEDURES

Operating Procedure

When the airflow sheets are completed it is appropriate to finalize the system operating procedure. This is usually a written document stating how systems are turned on and off, and it gives a description of their frequency of operation. For clean spaces general practice is to operate the air system 15–30 min for purging before beginning production or use of the space. In some cases it is necessary to operate a system continuously to ensure that pressurization or containment is always maintained. When analyzing system operation it is important to assess how an unexpected system shutdown or flow reduction will affect an adjacent sys-

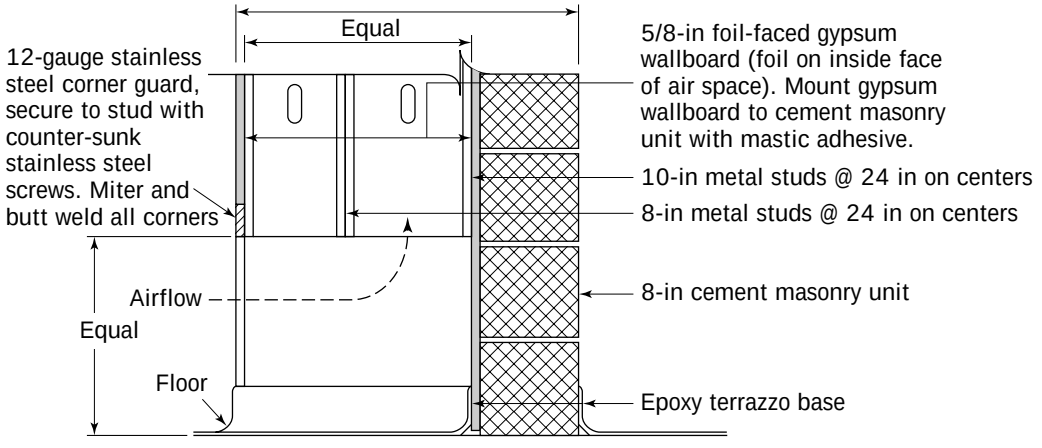


Figure 4. Return air wall detail.

tem or areas. Exhaust fans and fume hoods that may have intermittent or varying flows should be clearly noted in the operating procedure. It is essential that laboratory and production personnel clearly understand and follow the established procedures; otherwise, the air system may go out of balance if an exhaust or return system is inadvertently shut down.

System Alarms and Emergency Measures

Alarms that sound to indicate loss of pressurization are valuable features and essential in the HVAC design of critical areas. At this stage of design, procedures should be developed for the air system to accommodate a product spill or accident in a contained space. The ramifications of a spill on the air system, controlled space, and adjacent operations must be evaluated. Cleanup procedures could include fumigation of the air system, which would require operation of a relief connection to the ductwork for venting the fumigant.

Operations Review

The entire step of establishing operating procedures should ask as many what if questions as possible. System operation should ensure that the HVAC systems can maintain cleanliness and containment requirements. These conditions must be maintained while operating as close to normal as possible, as well as during periods of unexpected problems. Generally the reviewing agencies ask many of these questions; therefore, it is essential that the design team be fully prepared with thoughtful, responsive answers.

EMERGENCY ELECTRICAL POWER

An essential step in the HVAC design process is coordination with the electrical design team. Motor lists for HVAC equipment must be prepared and reviewed with the electrical design team. The need for motors designated for emergency power, variable speed, reduced voltage starting, or other special characteristics must be communicated to the electrical designers early in the design process. The sizing of the emergency generator can be greatly affected by motors required on emergency power from the HVAC system. Fans, equipment, or sensing devices that require interlocks must also be picked up by the electrical designers. The motor list must be kept up to date from project inception through commissioning. The motor list is useful for a reviewing agency, a valuable tool in training plant operators, and a great aid in understanding the HVAC system.

BUILDING CONTROL AND AUTOMATION SYSTEMS

Sequence of Operation

The automatic control system that controls and monitors the HVAC system is called by many names: the automatic temperature control system (ATC), the energy management and control system (EMCS), the building automation system (BAS), or the building management system (BMS). The first element in the design of the system is the development of a sequence of operation, which is a written description of the HVAC and related systems operation. A separate sequence is usually written for each air handling system, describing the complete operation of the system from control of coils and humidifiers to control of the room temperature and humidity. Starting and stopping of the air handling unit fans is outlined, along with interlocking of exhaust or return fans in relation to the main air system fan operation. Generally all fans operate at the same time, which is necessary to maintain pressurization. The sequence also addresses abnormal occurrences such as a smoke detection alarm or failure of an exhaust fan. The sequence describes what happens to system components during an abnormal occurrence. It may be necessary to shut a supply fan down if a major exhaust fan should fail to prevent or minimize the loss of pressurization. The sequence also describes any energy management strategies to be included in the system, such as a night temperature setback or reduced ventilation and exhaust rates during unoccupied periods.

System Architecture

The control system of choice for major facilities, and even for some small systems, is a direct digital control (DDC) system. Most major control system vendors and many of the smaller vendors offer DDC systems that are similar but contain internal differences. The systems are computer based and have the ability to communicate within and outside the system by coded digital signals. System architecture refers to the major components of the DDC system and their interrelationship. The architecture is developed by determining what components are initially required, what may be required in the future, and how the system may expand as additional requirements are added. Figure 5 shows a typical DDC system architecture.

Points List

After the sequence of operation is completed and the airflow diagrams are defined, the next step is to develop the alarm, control, and monitoring points list. This is an all-inclusive list of points that are to be connected to the DDC system. There are two major types of points: digital and analog. A digital point is simpler, generally less expensive, and works on a simple on–off or contact principle. Digital points are used to start and stop fans, indicate an on–off condition, or anything that requires only a single contact. An analog point is used to measure variables such as temperature, pressure, and flow rate. These points generally use 4- to 20-mA signals that provide varying signals in response to the parameter measured. The electronic signals used by the BAS may be transduced from variable pneumatic or pressure signals. The points list should include analog control points such as cooling coil valves and room temperatures. Monitoring points can be digital or an-

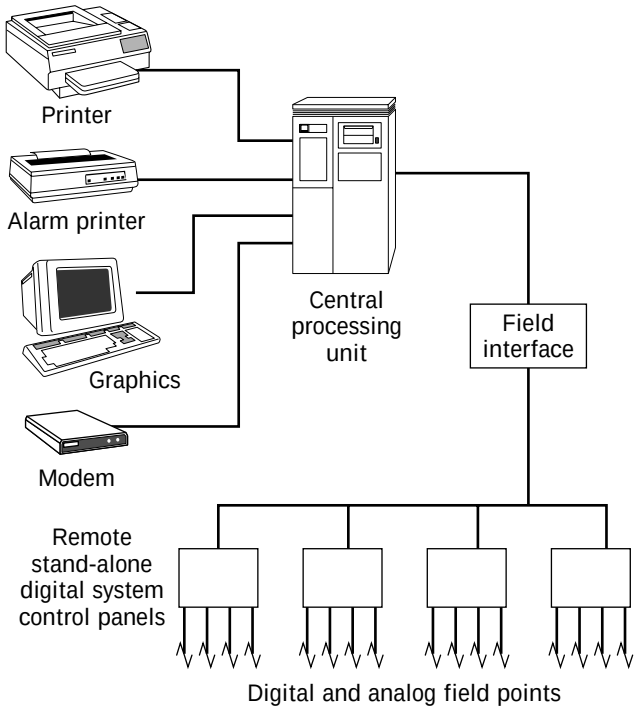


Figure 5. Direct Digital Control System Architecture.

alog, and can include fan run, room temperature indication, damper position, and room pressure indication. Alarm points can be either digital or analog and can include smoke detection in an air handling unit system, high or low environmental chamber temperature, high room humidity, or loss of room pressurization. Sensors located in clean spaces should be unobtrusive and installed in a cGMP acceptable manner.

Estimate of System Cost

The automatic control and monitoring system is a major cost element in the overall HVAC system for a bioprocess facility. After the points list is developed a good estimate can be made for the system cost. Several estimating numbers can be used in providing an educated guess of the cost, with a general range from a low of $500/point to as high as $1,200/point. Larger systems have lower dollar-per-point costs because the cost of the computers, monitors, and printers are diluted over many points. Smaller systems have higher costs because this dilution effect does not come into play. Systems with a higher percentage of analog points over digital points also tend to raise dollar-per-point costs. Control wiring is also a significant component of the BAS cost, and long runs of wiring tend to raise the dollar-per-point cost. When a floor plan is completed, the system architecture is established, and the points list is completed, this information should be given to one or more DDC vendors for a budget estimate. Vendors will make estimates based on the information provided, take off wiring runs, and suggest panel locations in arriving at the budget cost. At the completion of this exercise the cost of this major element can be locked in. Should the cost be greater than desired some points can be eliminated or the scope of the system reduced to bring the cost within the budget. When a system is bid it is recommended that component unit pricing be obtained from the vendor to lock in future costs should additions or modifications be desired. It is essential that vendors give classes and instructions on their systems before final turnover to the owner's operators.

Future System Flexibility

The trend is that after a DDC system is installed the number of points grows as the user becomes familiar with the system. Additional points may be added for maintenance, such as filter change out pressure devices. Run time logs set up in the system software are useful for lubrication schedules, belt checks, and so on. This requires software to log in and integrate data and to print out notices at preselected intervals of run time. Another useful feature is the phone modem, which would allow an operator to call in from a remote location and check on the system operation. This feature can save a trip to the facility if an alarm situation develops. Someone at the plant, such as the security guard who received the alarm, could place a call to the operator at home. The operator could dial into the BMS by modem and possibly analyze the problem by reading some key parameters. He or she could then determine the severity of the problem and either ignore it until the next shift or instruct someone on-site to take corrective action. Even though the original budget may not allow all features that are ultimately desired, it is good to look toward the future in anticipation of system expansion and ensure that the purchased system is capable of desired expansion.

TESTING, BALANCING, AND CLEANING

Testing and balancing for bioprocess systems is more critical than for commercial systems due to the very complex systems and higher pressures encountered. Associations of testing and balancing agencies publish excellent standard forms and procedures that are used in balancing air and water systems. Air balancing consists of setting terminal devices and air-moving equipment to deliver the proper quantities of air. For bioprocess facilities the additional requirement of establishing pressure differentials between adjacent spaces must also be set. These differentials are obtained by adjusting airflows, smoke tests, taking pressure readings, and setting controls. This effort can take some time as each facility is different and each room has different leakage characteristics that affect pressurization. As part of the balancing it may be discovered that duct systems or rooms are not as tight as designed and may require additional sealing to obtain the required pressure differentials. Recall that airflows shown on drawings are design values and generally require minor adjustment to achieve the required pressure differentials. A simple solution to many pressurization problems is to keep increasing outdoor air to the system. This can lead to problems if

design values are exceeded, with heating or cooling coils not meeting this need, resulting in off-design room temperatures and humidity levels. The air handling unit coils will use available cooling capacity to condition excessive quantities of outdoor air, resulting in room supply temperature higher than as designed.

In these cases the rooms or systems should be tightened to get closer to design values. The optimum time to balance is when few construction workers or facility personnel are in the spaces. The balance should be done with all doors closed, since opening and closing result in system pressure upsets and make balancing difficult. Before the system is turned over to the owner, the rooms and ductwork system must be thoroughly cleaned to allow certification of room cleanliness conditions. During construction the work areas should be broom-cleaned after each day's work and open duct systems kept closed. If followed during construction, these procedures will result in an easier final cleanup before turnover.

VALIDATION

When a bioprocess facility is to be validated, the validating agency will peruse the HVAC documentation and should communicate with the design engineers to establish the validation protocol as it relates to the HVAC system. If the design is proper, the contractor has properly installed the system, and the components perform as specified, the systems should be easily validatable. The validator will follow a master plan and protocols to verify the actual system installation and operation against design values and intent (10). The physical parameters reported by the BAS system will be verified by measurements using calibrated instruments to verify accuracy. Airflow measurements will be made at supply and exhaust outlets, as well as traverses across the face of hoods, to verify proper flows and capture patterns. Filters will be physically tested or challenged to uncover leaks in the filter media or at seals between the holding frame and elements. Any leaks will be field corrected. Samples of air will also be taken by the validator to ensure that the cleanliness requirement for the controlled space is actually maintained by the installed operating air system. If problems occur they are usually in the form of pressurization deficiencies or filter leaks. Even after the mechanical contractor has completed his or her work, a hole for a pipe could be cut in a room and not properly sealed, resulting in a change in room pressurization. Filter elements may not be properly sealed with the frame, resulting in leaks, or small holes may be discovered in the filter element as a result of handling or damage during installation. Should any of these problems occur, and are not easily resolvable by the validator, it may be necessary for the HVAC contractor and engineer to lend assistance in making the system valid. Another important criterion at this stage is the calibration of airflow monitors, which should be installed in supply and return systems to critical areas to ensure that airflow rates are as designed. Temperature and humidity sensors at critical areas should also be checked for accuracy at this time by actually reading space conditions and checking against values reported by the BMS.

SUMMARY

The HVAC system is an essential component in the successful operation of a bioprocess facility and must be given a proper amount of attention during the planning and design stage. The space requirements for air walls, control boxes, extensive duct runs, and large air handling units must be established early in the design process to prevent unsatisfactory compromises later on. The owner and operators must work closely with the architect and engineer design team to agree on the HVAC system philosophy and the proposed operation of the system. The amount of energy management and control, as well as system redundancy, are among important economic decisions that should be agreed on during this design process. Compromises made during the design development should make economic sense, be mutually beneficial, and be understood by affected members of the project team. It is best from a cost perspective to resolve HVAC issues as early in the design process as possible to avoid design delays and more costly changes later in the design process. This avoids delays in the design process and minimizes the effect of changes on other engineering disciplines. All parties involved must believe in the HVAC system philosophy to avoid costly and time-consuming retrofits during construction or at commissioning, validation, or after the facility begins operating.

BIBLIOGRAPHY

1. A.J. Shahidi, R. Torregrossa, and Y. Zelmanovich, *Pharm. Eng.* **5**, 72–82 (1995).
2. B.K. Lyderson, N.A. D'Elia, and K.L. Nelson, *Bioprocess Engineering Systems: Equipment and Facilities*, Wiley, New York, 1994, pp. 643–668.
3. J. Olsztynski, *PM Eng.* **2**, 20–21 (1996).
4. M. delValle, *Pharm. Eng.* **15**, 14–22 (1995).
5. *Technical Committee 9.10, HVAC Applications, ASHRAE Handbook*, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Inc., Atlanta, Ga. 1995, pp. 13.6–13.8.
6. Carrier Corporation, *System Design Manual*, vol. 1, *Load Estimating*, Carrier Air Conditioning Company, Syracuse, N.Y., 1972, pp. I-115–I-151.
7. B.K. Lyderson, N.A. D'Elia, and K.L. Nelson, *Bioprocess Engineering Systems: Equipment and Facilities*, Wiley, New York, 1994, pp. 588–593.
8. *Committee on Industrial Ventilation, Industrial Ventilation*, American Conference of Governmental Hygienists, Inc., Lansing, Mich., 1992, pp. 5–53.
9. *Technical Committee 5.2, Fundamentals, ASHRAE Handbook*, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Inc., Atlanta, Ga. 1993, pp. 32.1–32.40.
10. B.K. Lyderson, N.A. D'Elia and K.L. Nelson, *Bioprocess Engineering Systems: Equipment and Facilities*, Wiley, New York, 1994, pp. 2, 747–781.

HEMICELLULASES

LIISA VIIKARI
MAIJA TENKANEN
KAISA POUTANEN
VTT Biotechnology and Food Research
Espoo, Finland

KEY WORDS

Baking
Bleaching
Feed
Food
Hemicellulases
Hemicellulose
Mannan
Pulping
Xylan
Xylanases

OUTLINE

INTRODUCTION

During the past decade, the industrial use of hemicellulases, especially xylanases, has grown remarkably. Hemicellulases form a group of enzymes that degrade different hemicelluloses in annual plants and woody materials. Originally, hemicellulases were studied for use in the total hydrolysis of waste liquors containing hemicelluloses or for fractions derived from lignocellulosics. Today, hemicellulases are used for partial or restricted hydrolysis of fibrous materials containing hemicelluloses. Thus, hemicellulases have found uses in processing of hemicellulose-containing plant materials in the food, feed, and pulp and paper industries. The market for hemicellulases has been growing since the beginning of the 1990s due to a better understanding of the role of hemicelluloses in processing of the raw materials and to increased knowledge of these enzymes. Several commercial hemicellulases with different properties are now available from enzyme producers.

Hemicelluloses are natural components in cereals, and their properties and performance in food and feed raw materials can be improved by the action of enzymes. Exogenous enzymes offer alternatives to chemical additives in baking, where the application of microbial enzymes is expected to increase. Most of the added enzyme types are already present at low levels in the grain raw material. Commercial enzyme preparations principally consist of the same enzyme types but are usually of microbial origin. Enzymes in food and feed applications may also be used for upgrading the quality of the raw material. In the pulp and paper industry, the partial hydrolysis of fiber hemicelluloses by hemicellulases leads to improved processes and decreased chemical consumption (i.e., cleaner technology). The role of hemicellulases has been clearly demonstrated in improving the bleachability of pulps, whereas in other applications their impact is less significant.

Most of the enzymes studied in applications have been hemicellulase mixtures. For a successful application of hemicellulases a good understanding of both the substrate and the enzyme complex is needed. A reason for the sometimes fluctuating results on the effects of enzymes in various applications is that the preparations tested may contain various activities, including enzymes other than hemicellulases. Commercial enzyme preparations may vary considerably in composition and ratio of various activities, depending on the source.

This chapter highlights progress in the use of hemicellulases and their properties and exploitation in different areas (e.g., food, feed, pulp and paper) where the applications are commercially used.

STRUCTURE AND COMPOSITION OF HEMICELLULOSES

Hemicelluloses are heteropolysaccharides consisting of several polymers and varying in their monosaccharide composition, glycosidic linkage composition, substitution pattern, and degree of polymerization (1). They are often reported to be chemically associated (e.g., cross-linked) to other polysaccharides, proteins, or lignin. The structural characteristics affect physical properties, such as solubility, viscosity, crystallinity, and porosity and thus the enzymatic degradability.

The definition of the term *hemicellulose* is not very clear. It mainly refers to plant cell wall polysaccharides that can be extracted by alkaline solutions. Some hemicelluloses are also soluble in water. Hemicelluloses are usually classified according to the main sugar residue in the backbone, for example, xylans (D-xylose), mannans (D-mannose), galactans (D-galactose), and glucans (D-glucose), the first two being the main groups of hemicelluloses. Depending on the origins, hemicelluloses are substituted, to varying degrees, by side groups (1). In cereals in particular, the term *pentosan* is often used as a synonym for *hemicelluloses*, especially for xylans. Hemicelluloses present in cereals and woody materials are discussed in the following. Pectic substances consisting of polygalacturonic acid, arabinans, galactans, and arabino-

galactans, and noncellulosic β-glucans, which sometimes are included in hemicelluloses, are not considered in this review.

Hemicelluloses in Cereals

In cereals, hemicelluloses consist of arabinoxylans, arabinogalactans, and xyloglucans, of which arabinoxylans are most abundant. The content of arabinoxylans in cereals is reported to be 1.4–8.5%, with water-soluble fractions contributing 1.5–4.3%. Wheat and rye brans have the highest content of arabinoxylans. The main component of aleurone cell walls of cereals is arabinoxylan, and the cell walls of grain endosperms also contain significant amounts of arabinoxylan.

The arabinoxylans of annual plants have a linear backbone of (1,4)-β-D-xylopyranosyl units, which is mainly substituted with terminal α-L-arabinofuranosyl substituents attached to carbon 2 and/or carbon 3 of the xylosyl residues (1). Thus, xylans may be mono- or double-substituted by arabinose. The xylan backbone often contains other substituents as well, such as 4-*O*-methyl-α-D-glucuronic, acetic, or ferulic acids. The ratio of arabinose to xylose may be used as a simple indicator of the degree of substitution in arabinoxylans and varies between 0.54 and 0.70. The structure of arabinoxylans in annual plants varies among different species and parts. Thus, even arabinoxylans in aleurone and endosperm have different structures.

Hemicelluloses in Woody Materials

Wood is composed mainly of cellulose, hemicellulose, and lignin. Depending on the wood species, about 20–30% of dry weight of wood is hemicellulose. Hemicelluloses are associated with cellulose and lignin in woods. Not only the relative amounts but also the chemical composition of these polysaccharides in softwood and hardwood vary. Other polysaccharides, such as pectic compounds, are also present in wood in minor quantities. The distribution of carbohydrates in the wood fibers varies depending on the wood species and growing conditions (2). In general, in softwoods the xylan content of the innermost cell wall layer (the warty layer), is very high, but otherwise the xylan is relatively uniformly distributed throughout the fiber walls. In hardwoods, the outermost layers of fiber walls are rich in xylan. In softwoods the glucomannan content increases steadily from the outer to inner cell walls.

In wood, the two most common hemicelluloses are xylans and glucomannans. The hardwoods contain mainly xylan, whereas the amount of glucomannan in softwoods is approximately double compared with that of xylan. Hardwood xylan is composed of β-D-xylopyranosyl units, which may contain 4-*O*-methyl-α-D-glucuronic acid and acetyl side groups. 4-*O*-methylglucuronic acid is linked to the xylan backbone by *O*-(1,2) glycosidic bonds, and the acetic acid is esterified at the carbon 2 and/or 3 hydroxyl group. The softwood xylan is arabino-4-*O*-methylglucuronoxylan, in which the xylan backbone is substituted at carbon 2 and 3 with 4-*O*-methyl-α-D-glucuronic acid and α-L-arabinofuranosyl residues, respectively. Softwood galactoglucomannan has a backbone of β-(1,4)-linked β-D-glucopyranosyl and β-D-mannopyranosyl units, which are partially substituted by α-D-galactopyranosyl and acetyl groups (2).

HEMICELLULASES

Due to the complex structure of hemicelluloses, several different enzymes are needed for their enzymatic degradation and modification (Fig. 1). The two main glycanases that depolymerize the hemicellulose backbone are *endo*-1,4-β-D-xylanase and *endo*-1,4-β-D-mannanase. Small oligosaccharides are further hydrolyzed by 1,4-β-D-xylosidase, 1,4-β-D-mannosidase, and 1,4-β-D-glucosidase. The side groups are removed by α-L-arabinosidase, α-D-glucuronidase, and α-D-galactosidase. Esterified side groups are liberated by acetyl xylan esterase and acetyl glucomannan esterase (3).

Hemicellulases are produced by many species of bacteria and fungi, as well as by several plants. Many microorganisms produce a multiple pattern of these enzymes to degrade the plant materials efficiently. After identification of the most desirable enzymes for a potential application, the production of the enzyme can be improved by genetic engineering. Today, most commercial hemicellulase preparations are produced by genetically modified *Trichoderma* or *Aspergillus* strains.

Most microbial hemicellulases studied are active in slightly acidic conditions between pH 4 and 6 and at temperatures below 70 °C. More thermophilic, acidic, and alkalophilic hemicellulases are of great interest in many applications. Xylanases that are stable and function efficiently at high temperatures can be produced by thermophilic bacteria especially. Several xylanase genes, encoding proteins active at temperatures from 75 °C up to 95 °C (pH 6–8), have been isolated. The most thermophilic xylanases so far described are produced by species of the extremely thermophilic bacterium *Thermotoga* (4). Only a few ther-

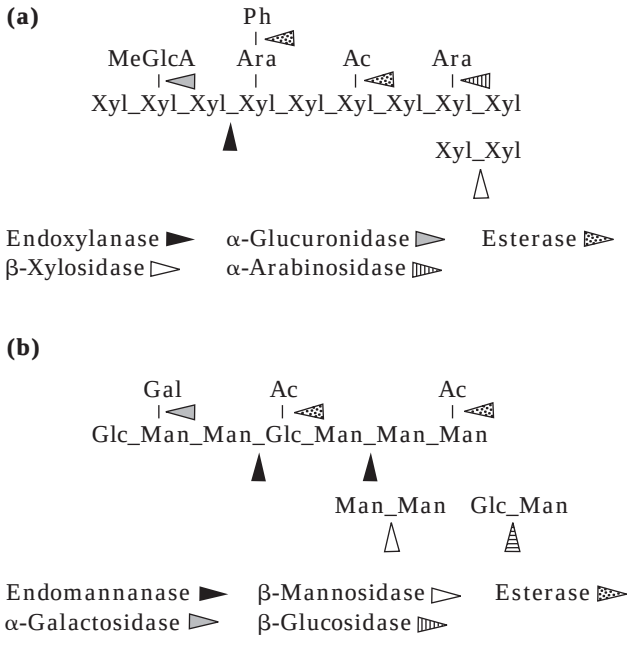


Figure 1. Enzymes participating in the hydrolysis of xylans (**a**) and glucomannans (**b**): Ph, phenolic groups; Ac, acetyl; Ara, arabinose; MeGlcA, methyl-glucuronic acid; Xyl, xylose, Gal, galactose; Glc, glucose; Man, mannose.

mophilic mannanases have so far been characterized (5,6). Xylanases and mannanases with alkaline pH optima have been detected especially in alkalophilic *Bacillus* species.

Xylanases

Since xylan is the most abundant hemicellulose in most raw materials, a major part of the published work on hemicellulases deals with the properties and mode of action, as well as applications of xylanases (7–9). Endoxylanases (1,4-*β*-D-xylan xylanohydrolases, EC 3.2.1.8) catalyze the random hydrolysis of 1,4-*β*-D-xylosidic linkages in xylans (Fig. 1). Most xylanases are rather small proteins (molecular mass around 20 kDa) with a basic isoelectric point (pl 8–10). Another group of xylanases has also been identified. These xylanases, are somewhat larger (molecular mass over 40 kDa) with acidic isoelectric point (pl 3–5). This grouping of xylanases is in consonance with the classification of glycanases based of hydrophobic cluster and sequence analysis (10). The xylanases belonging to the two identified xylanase groups (glycosyl hydrolase families 10 and 11, formerly F and G) differ from each other also with respect to their catalytic properties (11). Xylanases with high Mr/low pl (family 10) seem to exhibit greater catalytic versatility than the low Mr/high pl xylanases (family 11) and thus they are, for example, able to more efficiently hydrolyze highly substituted xylans.

The three-dimensional structures of different low molecular mass xylanases belonging to family 11 have been determined (12–14). These enzymes have very similar structures; they are ellipsoidal, well-packed molecules having diameters between 30 and 45 Å, and do not contain any separate substrate-binding domain. Some bacterial xylanases, however, have been found to contain either a cellulose-binding domain (15) or a xylan-binding domain (16). Only one crystal structure of family 10 xylanases has been solved (17). Figure 2 shows the three-dimensional structures of a catalytic domain of the xylanase A (family 10) of *Streptomyces lividans* and the xylanase II (family 11) of *Trichoderma reesei*.

Most of the xylanases characterized act randomly and are able to hydrolyze different types of xylans, showing only differences in the spectrum of end products. The main products formed from the hydrolysis of xylans are xylobiose, xylotriose, and substituted oligomers of two to four xylosyl residues. The chain length and the structure of the substituted products depend on the mode of action of the individual xylanase (11). Some xylanases, however, have a rather strict substrate specificity. A unique xylanase that requires a glucuronic acid substituent in the xylan backbone is produced by *Bacillus subtilis* (18). On the other hand, xylanases that require at least 24 unsubstituted xylose residues for action are produced by *Talaromyces emersonii* (19).

Mannanases

Endomannanases (1,4-*β*-D-mannan mannanohydrolase, EC 3.2.1.78) catalyze the random hydrolysis of *β*-D-1,4 mannopyranosyl linkages within the main chain of mannans and various polysaccharides consisting mainly of mannose, such as glucomannans, galactomannans, and galactoglucomannans (Fig. 1). Mannanases are generally larger proteins than xylanases (Mr 30–90 kDa) and have acidic isoelectric points. They also seem to be a more heterogeneous group of enzymes than xylanases (i.e., no clear groups based on biochemical properties have been identified). According to protein sequences, the known mannanases belong to three different glycosyl hydrolase families (families 5, 26, and 44) (10).

The mannanase of *T. reesei*, has been found to have a multidomain structure similar to several cellulolytic enzymes, that is, the protein contains a catalytic core domain that is separated by a linker from a cellulose-binding domain (20). The first three-dimensional structure of mannanase was recently published (21).

The main hydrolysis products from galactomannans and glucomannans are mannobiose, mannotriose, and various mixed oligosaccharides. The hydrolysis yield is dependent on the degree of substitution, as well as on the distribution of the substituents (22). The hydrolysis of glucomannans is also affected by the glucose to mannose ratio. Some mannanases are able to hydrolyze not only the *β*-1,4-bond between two mannose units but also the bond between the adjacent glucose and mannose units (23,24). Several endoglucanases are also able to hydrolyze galactoglucomannans.

Other Hemicellulases

Enzymes needed for further hydrolysis of the short oligomeric compounds produced by endoenzymes from hemicelluloses are *β*-xylosidase (1,4-*β*-D-xyloside xylohydrolase, EC 3.2.1.37), *β*-mannosidase (1,4-*β*-D-mannoside mannohydrolase, EC 3.1.1.25), and *β*-glucosidase (Fig. 1). *β*-Xylosidases and *β*-mannosidase catalyze the hydrolysis of xylo- and manno-oligosaccharides, respectively, by removing successive xylose or mannose residues from the nonreducing termini. Exoglycanases are generally larger proteins than endoglycanases, with molecular weights above 100 kDa, and they are often built up by two or more subunits.

Side groups that are still attached to oligosaccharides after the hydrolysis of xylans and mannans by xylanase or mannanase, respectively, restrict the action of *β*-xylosidase and *β*-mannosidase. The hydrolysis stops at a substituted sugar unit. The *β*-xylosidase of *T. reesei* is unable to hydrolyze the linkage adjacent to a substituted xylose unit carrying a 1,3-linked *α*-L-arabinofuranosyl substituent, whereas the linkage is easily hydrolyzed if the substituent is 1,2-linked *α*-L-arabinofuranosyl or 4-*O*-methyl-*α*-D-glucuronic acid (25). Similarly, the 1,6-linked *α*-D-galactopyranosyl side group restricts the action of *β*-mannosidase of *A. niger*, which is able to only slowly hydrolyze the linkage adjacent to the substituted mannose unit.

The side groups connected to xylan and glucomannan main chains are removed by *α*-glucuronidase, *α*-arabinosidase (*α*-L-arabinofuranoside arabinofuranohydrolase, EC 3.2.1.55), and *α*-D-galactosidase (*α*-D-galactoside galactohydrolase, EC 3.2.1.22) (Fig. 1). Acetyl substituents bound to hemicellulose are removed by esterases. There are clearly different types of side group cleaving enzymes. Some of them are able to hydrolyze only substituted short-chain oligomers, which first must be produced by the backbone depolymerizing endoenzymes (xylanases and man-

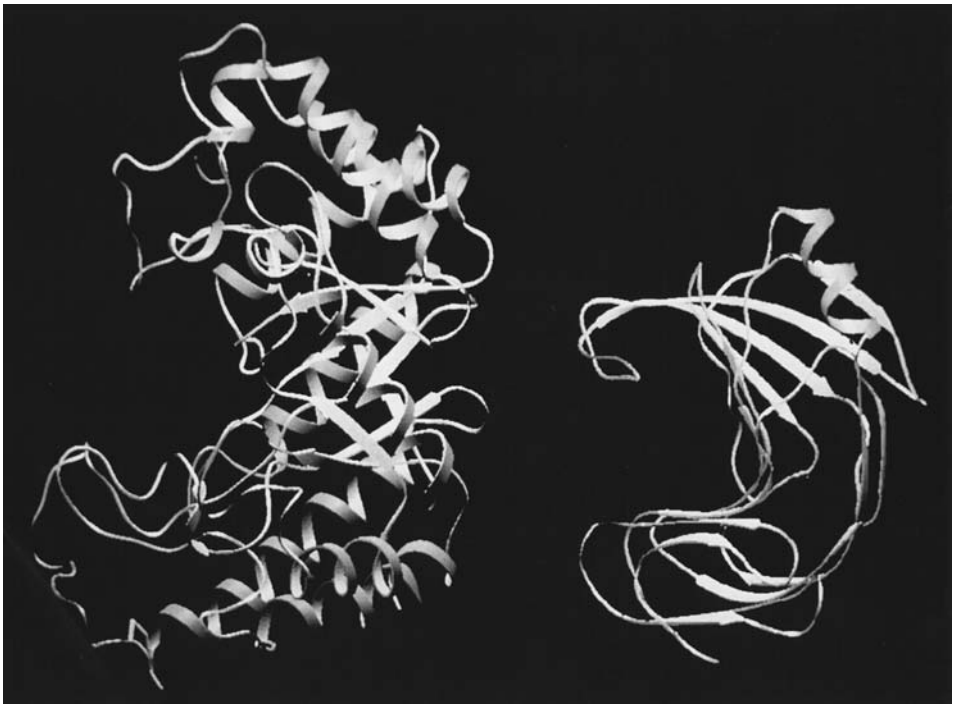


Figure 2. The three-dimensional structures of the catalytic domain of the xylanase A of *Streptomyces lividans* and the xylanase II from *Trichoderma reesei*. *Source*: Courtesy of Professor Juha Rouvinen.

nanases). Others are capable of also attacking intact polymeric substrates. Most accessory enzymes of the latter type, however, prefer oligomeric substrates. The synergism between different hemicellulolytic enzymes is also observed by the accelerated action of endoglycanases in the presence of accessory enzymes. Some accessory enzymes, such as an arabinosidase and an esterase of *Pseudomonas fluorescens* subs. *cellulosa* and an acetyl xylan esterase of *Trichoderma reesei* have been found to contain a multidomain structure with a separate cellulose binding domain (26).

Evaluation of Hemicellulases

Enzyme activity is usually determined using isolated natural substrates. Isolation of a substrate from complex raw materials, however, often leads to chemical and structural modification of the substrate. Thus, in hemicelluloses, some of the side groups may be removed during chemical extraction, or the average degree of polymerization may change. Also, the physical status of the substrate in the activity assay may be different from that in the practical application. Thus, the solubility of the substrate in the activity assay is usually high, whereas in fiber-bound matrix the substrate is usually insoluble and/or may be also chemically linked to other components. Although the substrates in activity assays are not identical to the practical raw materials, they are chosen to correspond to the conditions under which the enzyme will be used. The main advantage and justification for the use of isolated substrates is that they give reproducible assay figures. Furthermore, they are more readily available and consistent and give the opportunity to compare different types of enzymes at various places.

It is well recognized that comparison of enzyme activities is complicated by the utilization of different analysis methods (27). Hence, in practical applications, the action of enzymes must be compared on the actual substrate. In these tests, the liberation of sugars and/or oligosaccharides into solution is the direct measurement of the action of the enzyme. This is usually carried out either by reducing sugar analysis or by HPLC. Several indirect methods to correlate the activity with the actual effect in the practical use, such as viscosity determination, have been developed. However, the most reliable method for understanding the effects of different enzymes is to compare the enzymes and dosages in the final application.

APPLICATION OF HEMICELLULASES

The major applications of hemicellulases in the food, feed, and pulp and paper industries are described in Table 1.

Food

Enzymes have long been recognized as useful tools for improving the processing behavior or properties of cereal foods, and their use is continuously increasing. The main enzymes used are amylases acting on starch, but hemicel-

Table 1. Benefits of Using Hemicellulases in Different Industrial Areas

Application area	Effect	Benefit
Food	Dough rheology	Improved handling
	Loaf volume	Quality and yield
	Staling rate	Improved storage
Feed	Nutritive value	Cost savings
	Feed conversion rate	Cost savings
Pulp and paper	Bleachability	Chemical savings, quality

lulases have also been used in cereal processing. The main area in which hemicellulases are used is baking, where dough handling and bread quality can be improved by the action of enzymes.

Hemicellulases in Cereal Processing. Endogenous and added enzymes both have an important effect on the quality of cereal foods, especially in baking. Although enzymes act on molecular level, they can induce remarkable changes in both the microstructure and the functional properties of cereal foods. Progress has been made recently in the analysis of enzymes present in flours. The presence of endogenous enzymes has been suggested to be the cause of the variation in liquid loss among wheat doughs made from different cultivars (28). Differences in α-arabinofuranosidase, β-xylosidase, and endoxylanase levels have been detected in various European wheat varieties. Although present only at low levels, these enzymes may contribute to the bread-making performance of flours (29).

Enzymes have been shown to differ in their abilities to both solubilize insoluble substrates and further hydrolyze the released fragments. The chemical composition, architecture, and physical properties of the surrounding tissue of the cereal cell walls are all of importance if selective solubilization of the polysaccharides is attempted. Physical inaccessibility of the substrate may retard or hinder the reactions, and several enzymes may be needed. As compared with xylanases, less is known about the other enzymes acting on the hemicelluloses. The role of ferulic acid in cereals is interesting. Ferulic acid is esterified to the arabinose side groups of arabinoxylans, and it has been suggested that it plays an important role in the architecture of the cell walls and in the cross-linking of arabinoxylan molecules. Very recently, it was reported that ferulic acid esterase influenced the solubilization of β-glucans and arabinoxylans from barley cell walls (30).

Baking. Flour, water, yeast, and salt are the basic components of bread and other baked foods. It is, however, common practice to improve the processing of dough and/or the quality of the baked products by adding other ingredients. Process criteria are especially related to dough handling (machinability) and process yield (water retention, proofing times). The product quality is mostly evaluated on the basis of appearance (volume, color) and eating quality (crumb elasticy, flavor) (31). Various bread improving components, such as emulsifiers, oxidizing agents, and enzymes, have been shown to influence bread quality, especially when lower-quality flours with a low gluten content are used.

A small amount of arabinoxylans (pentosans) is present in wheat and rye flour. Insoluble arabinoxylan especially hinder the development of wheat gluten, which is vital for the formation of the loaf structure. White flour contains 2.5–3%, whole-meal wheat flour about 5%, and rye as much as 8% arabinoxylan, which has an important role in the water binding of dough. Arabinoxylans can bind nearly 10 times their weight of water and may account for approximately one-third of the water-binding capacity of dough. By controlled enzymatic hydrolysis of xylans, the dough becomes easier to handle and the resulting bread has a larger loaf volume and an improved crumb structure (32).

To understand the reactions of hemicellulases, several reports have concentrated on studying the mechanisms of action of xylanases and other hemicellulases. The release of high molecular weight water-soluble arabinoxylans from water-inextractable cell wall material has been shown. The specific viscosities of extracts from xylanase-treated doughs were 40–65% higher than those from untreated dough, because the average apparent viscosity of the extracted arabinoxylans remained high (33).

It has been suggested that the effect of xylanase on bread volume improvement can be ascribed to redistribution of water from the xylan phase to the gluten (continuous phase). Water distribution is of crucial importance in the dough and bread. The transfer of water from one component to another, due to the decreased water-binding capacity of a partially hydrolyzed polymer, is suggested to be a major reason behind changed functional properties (34). The increase in gluten volume fraction gives the gluten more extensibility, which eventually results in better oven spring (31).

Xylanases have also been found to improve wheat dough fermentation stability (35), counterbalance the variable flour quality in wheat bread baking (33), retard staling, and improve the texture of high-fiber bread (36). They are also useful in the manufacture of low-fat biscuits (37).

Animal Nutrition and Feed

A part of the energy content in many cereals is locked up in the form of nonstarch polysaccharides, undigestible for several animals. Therefore, selected enzymes, such as hemicellulases, can be added to break down these polymers, leading to increased metabolizable energy. In addition, degradation of hemicelluloses often liberates starch and protein, masked by the cell walls in the cereals, increasing their utilization. In some cereals, a large part of arabinoxylans is soluble and causes high viscosity in the small intestine of some animal species. Addition of xylanases decreases viscosity and improves feed utilization.

Hemicellulases in Animal Nutrition. Animal feed is predominantly composed of plant material, and cereals form the major part of diets for both poultry and pigs. Some of these raw materials, such as rye and barley, are known to be less suitable than maize and wheat. This is emphasized by their gross energy and metabolizable energy in the target species. Thus barley, with a gross energy similar to that of wheat, has, in practice, an approximately 10% lower energy availability than wheat (38). This has previously limited its value as a feed component as compared with the more digestible materials. The decreased nutritive value has been associated with the presence of arabinoxylans in rye and wheat. Feed utilization and digestion can often be increased by the addition of external enzymes to the feed.

Endogenous enzymes play a crucial part in an animal's own digestion of feed components. External enzymes added directly to feeds act as supplements to the normal digestive enzymes already found in animals. The external enzymes carry out their work on the feed materials present in the animal's digestive tract. They can therefore be con-

sidered as an extension of the animal's own digestive enzyme system, especially in the case of young animals whose digestive systems are not fully mature. Enzymes can be added to animal feeds either to degrade soluble fiber, which has an antinutritional effect, or to supplement the animal's own digestive enzymes. In the latter case, supplementation of the feed with amylase or protease is the current practice. β-Glucanases are also used, especially for the hydrolysis of barley β-glucans.

The content of arabinoxylan, β-glucan, and other nonstarch polysaccharides in feed raw materials does not remain constant, and varies with variety and also with the climatic conditions under which the raw material is grown. Prediction of the enzyme response has prove to be complicated. Different methods, such as viscosity measurements or determination of arabinoxylan level, have been studied as possible predictors of the nutritive quality with varying success (39). At present, however, no rapid predictive test exists that can be applied to raw material evaluation, and in vivo methods are still the best.

The desired pH optimum of the enzyme preparations are around 6–6.5, prevailing in the small intestine, the area in which the majority of the enzyme action of xylanases is believed to occur. The stability of the enzyme both in processing and in feed is important. Enzymes must be resistant to the heat processing of the feeds and also to the transition in the digestive system to the point of action. In practical feed-mill conditions, the addition of steam in pelleting is responsible for the loss of enzyme activity. Commercial preparations may be protected by being absorbed to a carrier or coated with a steam-resistant material that hinders access of the steam to the enzyme product. The intrinsic thermal resistance of an enzyme may thus be improved by a protective system. As the majority of xylanases used in feed applications have very similar intrinsic thermal stabilities, the thermal stability of different products is mainly a function of the efficiency of the carrier and/or coating technology used to protect the enzyme product. The loss of activity can be to some extent avoided by using dried enzymes, which are applied after pelleting (38).

Poultry Feed. Wheat has been found to have a variable energy content, depending on local growth conditions. While there has been found to be no correlation between the level of total nonstarch polysaccharides in wheat, a significant negative correlation between the level of water-soluble nonstarch polysaccharides and the average metabolizable energy (AME) has been demonstrated. It has been observed that supplementation of a wheat-based broiler diet with xylanase preparations improved the AME value of the wheat used by 7–10% (40). An increased release of proteins and nonstarch carbohydrates from the cells has been verified by chemical and microscopic analysis (41). The addition of xylanase enzymes can therefore be used to improve the nutritive value of wheat-based diets. This is now common practice in Western Europe and other areas of the world where wheat is available as a feed cereal.

Enzyme supplementation of wheat-containing diets is, however, cost sensitive, and the dosages used may not always be those that give the maximum effect. In practice, addition levels vary from 300 to 500 g of enzyme preparation per ton of feed for broilers (42). Too-high enzyme dosages have, in some cases, led to an imbalance between protein and energy.

Due to the inherent variations in biological systems, small fluctuations in results from place to place have been observed (38). This should be contrasted with the more traditional enzyme applications, where the response is much more predictable and can even be mathematically modeled. It has been observed that wheat with a lower AME than average responds more to the enzyme supplementation, thus introducing another source of variation in the response. Despite this, however, xylanases are now routinely used in wheat-containing diets for broilers, and considerable efforts are being made to understand their mode of action and predict their effects. Similar results have also been obtained with rye-based diets (43). The effects obtained with commercial enzymes may sometimes be due to the action of several enzymes, as the commercial preparations often contain several polysaccharide-degrading activities.

As compared with broilers, laying hens and turkeys are more mature animals with a more developed digestive tract, better able to cope with the nonstarch polysaccharides in cereals and vegetable sources. It has been observed that the advantage of the extra energy liberated by the enzyme supplementation of the feed can be gained by changing the feeding strategy. As the enzyme effect has been found to decrease during the growth of the birds, aging, a feeding strategy where enzyme dose is reduced as the feeding program continues, may provide a more cost-effective route for improvement of feed efficiency and performance.

Pig Feed. Pig feeds containing very high levels of wheat are also less common, as the formulation of pig feeds tends to be multicereal based. According to some experiments, however, increasing levels of xylanase supplementation, up to 600 g of the commercial enzyme per ton of feed (containing 20–35% barley and 30–35% wheat), resulted in improvement of performance, as measured by feed conversion ratio and live weight gain (38).

Pig feeds normally contain more by-products than poultry feeds because of the overall higher digestive efficiency of the pig compared with poultry. These by-products can also be upgraded by enzyme supplementation and, in some instances, offer higher potentials than the materials from which they are derived. The presence of noncereal feed raw materials in the complex diet (e.g., rapeseed meal, sugar beet pulp, sunflower meal) has been found to diminish the enzyme effect, as the substrates present were not degradable by the xylanase used. This implies that attention must be paid to the raw material composition of pig diets if the maximum enzyme effect is to be achieved (38).

Pulp and Paper

In the pulp and paper industry, hemicellulases are mainly used to improve the bleachability of pulps. In this application, xylanases are the most important group of enzymes, although mannanases may also facilitate the bleaching of certain types of pulps (44). In addition, xylan-

ases have been present in enzyme preparations used to improve the drainage (water removal) or deinking of recycled fibers. The mechanisms in these applications are, however, not clear, and they are obviously not based solely on the action of hemicellulases.

Pulping and Bleaching Processes. The main aim in chemical pulping is to remove lignin and to separate the wood fibers from each other to render them suitable for the paper-making process. In the pulping process, the lignified middle lamella located between the wood fibers is solubilized by various chemicals. Today, the predominant pulping method is the kraft process, where the cooking liquors are incinerated and the cooking chemicals recycled.

In bleaching, the primary goal is to remove the low amount of residual lignin present in the pulp after cooking without decreasing the molecular weight of cellulose. Lignin in unbleached pulps typically represents only about 1% of the dry weight. During pulping, however, lignin is chemically modified and condensed, resulting in poorly degradable structures. Cooking and bleaching are separate process phases, differing from each other with respect to the selectivity of the chemicals used. In the bleaching processes, lignin is sequentially degraded and extracted in several phases. Bleaching sequences are generally composed of at least five phases. Previously, the bleaching of chemical pulps was carried out with elemental chlorine and chlorine dioxide. In Europe today pulp is bleached by elemental chlorine free (ECF) or totally chlorine free (TCF) bleaching processes, where oxygen, ozone, or peroxide is used.

Extensive modifications of hemicelluloses take place during pulping processes. During conventional kraft cooking, part of the hemicelluloses is first solubilized in the cooking liquor. At the start of the kraft process xylan in wood is partly solubilized by the alkaline cooking liquid and many of the side groups are cleaved off. It has recently been observed that the majority of the 4-*O*-methylglucuronic acid side groups are converted in the early phases of the kraft cook to hexenuronic acid (45). In the later phases of the process, when the alkalinity of the cooking liquor decreases, part of the solubilized xylan is relocated onto the cellulose fibers (46). Although glucomannan is the main hemicellulose in softwood, the bulk of glucomannans are dissolved and degraded during kraft pulping. Reprecipitation of glucomannan has not been reported to take place to the same extent as xylans. Thus, the relative amount of xylan is increased in pine kraft pulp as compared to pine wood (2).

In addition to xylan, lignin is also partially readsorbed on the fibers. Lignin has been reported to be linked to hemicelluloses, forming lignin–carbohydrate complexes (LCC). Furthermore, hemicelluloses seem to physically restrict the passage of high molecular mass lignin out of the pulp fiber cell wall, and thus the removal of hemicelluloses, especially xylan, can be expected to enhance the extractability of residual lignin from pulps.

Modification of Fibers by Hemicellulases. The effect of hemicellulases in bleaching is based on the modification of pulp hemicelluloses, enhancing the removal of lignin in chemical bleaching. It has been proposed that the action of xylanases is due to the partial hydrolysis of reprecipitated xylan or to removal of xylan from the lignin-carbohydrate (LC) complexes. Both mechanisms would lead to enhanced diffusion of entrapped lignin from the fiber wall. In softwood kraft fibers, removal of xylan by xylanases was found to uncover lignin (47). Thus, it can be expected that removal of xylan improves the extractability of lignin by exposing lignin surfaces.

The partial hydrolysis of xylan or glucomannan may also degrade and improve the extractability of LCC. Both softwood and hardwood kraft pulps have been reported to contain LCC in which carbohydrates and lignin may be connected to each other by ether or glycosidic linkages (48). However, no direct evidence for the type of linkage(s) existing between hemicelluloses and lignin has yet been presented. The action of xylanases on both reprecipitated and LC xylan in enhancing bleachability suggests that it is probably not only the type but also the location of the xylan that is important in the mechanism of xylanase-aided bleaching.

Hemicellulases in Bleaching Processes. The use of xylanases in different bleaching sequences uniformly leads to a reduction in chemical consumption (8,44). The benefits obtained by enzymes are dependent on the type of bleaching sequence used, as well as on the residual lignin content (measured as the kappa number) of the pulp. Originally, xylanases were studied to reduce the consumption of elemental chlorine. Later, xylanases were combined with various ECF and TCF bleaching sequences to improve the otherwise lower brightness of pulp or to decrease the bleaching costs. In chlorine bleaching, an average reduction of 25% in active chlorine consumption in prebleaching, or a reduction of about 15% in total chlorine consumption has been reported both in laboratory scale and in mill trials. As a result, the concentration of chlorinated compounds, measured as adsorbable organic halogen (AOX), in the bleaching effluent during mill trials was reduced by 15–20% (8). Today, xylanases are industrially used both in ECF and TCF sequences. In ECF sequences the enzymatic step is often implemented due to the limiting chlorine dioxide production capacity. The use of enzymes allows bleaching to higher brightness values when chlorine gas is not used. In TCF sequences, the advantage of the enzymatic step is due to improved brightness, maintenance of fiber strength, and savings in bleaching costs.

The amount of enzyme needed for the bleaching is also a key parameter with respect to both enzyme cost and yield loss. Generally, it seems that although the hydrolysis (the solubilization of carbohydrates) increases as a function of the enzyme dosage used, no further benefits can be obtained to the bleachability after a certain limit (49). Thus, to maximize the positive effect of the enzyme on the pulp kappa number and brightness, and simultaneously to minimize the yield loss, laboratory-scale experiments are needed to optimize the enzyme dosage.

Enzymatic pretreatment has been shown to be fully compatible with existing industrial equipment, which is a considerable advantage of this method, especially when

Table 2. Cost Savings with Xylanases in Bleaching

Chemical savings	
Chlorine dioxide	$4 /ton
Costs	
Enzyme	$1.5 /ton
Acid	$0–0.2 /ton
Net savings	$2.3–2.5 /ton pulp

Source: From Roehm Enzymes, Finland.

compared with some other competing technologies. Xylanases are typically mixed with water before being added to the unbleached pulp by a shower bar and allowed to react in the high-density storage tank for 1 or 2 h before the subsequent chemical bleaching steps. Presently a number of mills in North America and Scandinavia use enzymes continuously.

The approximate price of xylanase in 1997 was less than $2/ton pulp. Estimations for the capital cost of enzyme delivery and pH adjustment varied from $10,000 to $1,000,000 in 1995. The potential economic benefits of enzyme bleaching are significant to the pulp and paper industry. Calculation of the relative economic benefits in an ECF sequence reveals that the reduction of the chlorine dioxide consumption leads to savings of at least $2/ton pulp (Table 2). The costs of oxygen-based chemicals (ozone, peroxide) are even higher and the respective savings even more pronounced.

FUTURE OUTLOOK

Today, hemicellulases are used in three major industrial sectors: in food as baking aids, in feed for increasing the nutritional value, and in the pulp and paper industry for saving chemical costs. Most applications are based on the action of xylanases, and several commercial xylanase preparations with different properties are available. The future developments of hemicellulases aim at more efficient and stable enzymes. Obviously, enzymes with different pH and higher-temperature optima and stability are desired for most applications. Thermal stability has been identified as one of the key properties in all industrial sectors. The use of monocomponent enzyme products, alone or in optimized combinations for targeted use, will also become more common. Tailored side group cleaving enzymes will be available, in addition to the present major endoglycanases. Already most of the industrial enzymes are produced by strains developed by genetic engineering. These techniques are expected to lead to more cost-efficient enzyme products.

BIBLIOGRAPHY

1. G.O. Aspinall, *Adv. Carbohydr. Chem.* **14**, 429–468 (1959).
2. E. Sjöström, *Wood Chemistry: Fundamentals and Application*, 2nd ed., Academic Press, San Diego, Calif., 1993.
3. R.F.H. Dekker, in T. Higuchi ed., *Biosynthesis and Biodegradation of Wood Components*, Academic Press, Orlando, Fla., 1985, pp. 505–533.
4. H.D. Simpson, U.R. Haufler, and R.M. Daniel, *Biochem. J.* **277**, 413–417 (1991).
5. P.A. Bicho, T.A. Clark, K. Mackie, H.W. Morgan, and R.M. Daniel, *Appl. Microbiol. Biotechnol.* **36**, 337–343 (1991).
6. G.D. Duffaud, C.M. McCutchen, P. Leduc, K.N. Parker, and R.M. Kelly, *Appl. Environ. Microbiol.* **63**, 169–177 (1997).
7. M.P. Coughlan and G.P. Hazlewood, *Biotechnol. Appl. Biochem.* **17**, 259–289 (1993).
8. L. Viikari, A. Kantelinen, J. Sundquist, and M. Linko, *FEMS Microbiol. Rev.* **13**, 335–350 (1994).
9. A. Sunna and G. Antranikian, *Crit. Rev. Biotechnol.* **17**, 39–67 (1997).
10. B. Henrissat and A. Bairoch, *Biochem. J.* **293**, 781–788 (1993).
11. P. Biely, M. VrŠanská, M. Tenkanen, and D. Kluepfel, *J. Biotechnol.* **57**, 151–166 (1997).
12. H. Okada, *Adv. Prot. Res.* **12**, 81 (1989).
13. A. Törrönen, A. Harkki, and J. Rouvinen, *EMBO J.* **13**, 2493 (1994).
14. A. Törrönen and J. Rouvinen, *Biochemistry* **34**, 847 (1995).
15. K. Sakka, Y. Kojima, T. Kondo, S. Karita, K. Ohmiya, and K. Shimada, *Biosci. Biotech. Biochem.* **57**, 273–277 (1993).
16. F. Shareck, C. Roy, M. Yaguchi, R. Morosoli, and D. Kluepfel, *Gene* **107**, 75 (1991).
17. U. Derewenda, L. Swenson, R. Green, Y. Wei, R. Morosoli, F. Shareck, D. Kluepfel, and Z.S. Derewenda, *J. Biol. Chem.* **269**, 20811–20814 (1994).
18. K. Nishitani and D.J. Nevins, *J. Biol. Chem.* **266**, 6539 (1991).
19. M.G. Tuohy, J. Puls, M. Clayssens, M. VrŠanská, and M.P. Coughlan, *Biochem. J.* **290**, 515 (1993).
20. H. Stålbrand, A. Saloheimo, J. Vehmaanperä, B. Henrissat, and M. Penttilä, *Appl. Environ. Microbiol.* **61**, 1090–1097 (1995).
21. M. Hilge, S.M. Gloor, W. Rypniewski, O. Sauer, T.D. Heightman, W. Zimmerman, K. Winterhalter, and K. Piontek, *Structure* **6**, 1433–1444 (1998).
22. B.V. McCleary, in G.F. Leatham and M.E. Himmel eds., *Enzymes in Biomass Conversion*, ACS Symp. Ser. 460, 1991, American Chemical Society, Washington, D.C., pp. 437–449.
23. I. Kusakabe, G.G. Park, N. Kumita, T. Yasui, and K. Murakami, *Agric. Biol. Chem.* **52**, 519 (1988).
24. M. Tenkanen, M. Makkonen, M. Perttula, L. Viikari, and A. Teleman, *J. Biotechnol.* **57**, 191–204 (1997).
25. M. Tenkanen, E. Luonteri, and A. Teleman, *FEBS Lett.* **399**, 303–306 (1996).
26. G.P. Hazlewood and H.J. Gilbert, in J. Visser, G. Beldman, M.A. Kusters-vanSomeren, and A.G.J. Voragen eds., *Xylans and Xylanases*, Elsevier, Amsterdam, 1992, pp. 259–273.
27. M.J. Bailey, P. Biely, and K. Poutanen, *J. Biotech.* **23**, 257–270 (1992).
28. G. Beldman, D. Osuga, and J.R. Whitaker, *J. Cereal Sci.* **23**, 169–180 (1996).
29. G. Cleemput, W. Bleukx, M. van Oort, M. Hessing, and J.A. Delcourt, *J. Cereal Sci.* **22**, 139–145 (1995).
30. J. Moore, C.W. Bamforth, P.A. Kroon, B. Bartolomé, and G. Williamson, *Biotechnol. Lett.* **18**, 1423–1426 (1996).
31. J. Maat, M. Roza, J. Verbakel, H. Stam, M.J. Santos da Silva, M. Bosse, M.R. Egmond, M.L.D. Hagemans, R.F.M. v. Gorcom, J.G.M. Hessing, C.A.M.J.J. v.d. Hondel, and C. v. Rotterdam, in J. Visser, G. Beldman, M.A. Kusters-van Someren, and A.G.J. Voragen eds., *Xylans and Xylanases*, Elsevier, Amsterdam, 1992, pp. 349–360.

32. H. Gruppen, F.J.M. Kormelink, and A.G.J. Voragen, *J. Cereal Sci.* **18**, 129–143 (1993).
33. X. Rouau, M.L. El-Hayek, and D. Moreau, *J. Cereal Sci.* **19**, 259–272 (1994).
34. K. Poutanen, *Trends Food Sci. Technol.* **8**, 300–306 (1997).
35. B. Sprössler, *Getreide. Mehl Brot* **50**, 281–283 (1996).
36. T. Laurikainen, H. Härkönen, K. Autio, and K. Poutanen, *J. Sci. Food. Agric.* **76**, 239–249 (1998).
37. U.S. Pat. 5,108,764 (April 28, 1992) S.A.S. Craig, L.R. Beehler, R.T. Deihl, H. Levine, A.M. Magliacano, P.R. Mathewson, M.S. Otterburn, L. Slade, and P. Verduin (to Nabisco Brands, Inc.).
38. W.D. Cowan, in T. Godfrey and S. West eds., *Industrial Enzymology*, 2nd ed., Macmillan, London, 1996, pp. 69–101.
39. M. Choct, G. Annison, and R.P. Trimble, *Proc. Austr. Poultry Sci. Symp.* **5**, 78 (1993).
40. G. Annison, *J. Agric. Food Chem.* **39**, 1252–1256 (1991).
41. A. Tervilä-Wilo, T. Parkkonen, A. Morgan, M. Hopeakoski-Nurminen, K. Poutanen, P. Heikkinen, and K. Autio, *J. Cereal Sci.* **24**, 215–225 (1996).
42. B. Carr, M. Lessire, T.H. Nguyen and M. Larbier, *Proceedings of the 19th World's Poultry Congress*, Amsterdam, 1992, pp. 411–415.
43. M.R. Bedford and H.L. Classen, in J. Visser, G. Beldman, M.A. Kusters-van Someren, and A.G.J. Voragen eds., *Xylans and Xylanases*, Elsevier, Amsterdam, 1992, pp. 361–370.
44. A. Suurnäkki, M. Tenkanen, J. Buchert, and L. Viikari, *Adv. Biochem. Eng.* **57**, 261–287 (1997).
45. A. Teleman, V. Harjunpää, M. Tenkanen, J. Buchert, T. Hausalo, T. Drakenberg, and T. Vuorinen, *Carbohydr. Res.* **272**, 55–71 (1995).
46. S. Yllner and B. Enström, *Svensk Papperstidn.* **59**, 229–232 (1956).
47. J. Buchert, G. Carlsson, L. Viikari and G. Ström, *Holzforschung* **50**, 69–70 (1996).
48. G. Gellerstedt and E.-L. Lindfors, *Proc. Int. Pulp Bleaching Conf., vol. 1*, SPCI, Stockholm, 1991, p. 73.
49. J. Buchert, M. Ranua, A. Kantelinen, and L. Viikari, *Appl. Microbiol. Biotechnol.* **37**, 825–829 (1992).

See also ENZYMES, PULP AND PAPER PROCESSING; HEMICELLULOSE CONVERSION.

HEMICELLULOSE CONVERSION

G.T. TSAO
NINGJUN CAO
C.S. GONG
Purdue University
West Lafayette, Indiana

KEY WORDS

2,3-Butanediol
Ethanol
Hemicellulose
Pretreatment
Xylitol
Xylose

OUTLINE

INTRODUCTION

Hemicellulose is a plant cell wall polysaccharide that is localized mostly in cell wall middle lamella. It is in close association with cellulose, lignin, and pectic materials. This close association contributes to the rigidity and flexibility of the plant cell wall. Hemicellulose is the third-most abundant polymer in nature. In some plants, it comprises up to 40% of the total dry material. Unlike cellulose, hemicellulose exhibits a wide diversity in both structure and constitution. The degree of polymerization of hemicellulose is usually less than 200. The type and amount of hemicellulose vary widely, depending upon plant material, type of tissue, stage of growth, and so on. For those reasons, it is difficult to obtain a typical sugar composition of a typical hemicellulose.

Hemicelluloses are heteropolysaccharides that are composed of various hexoses (e.g., glucose, mannose, and galactose), pentoses (D-xylose and L-arabinose), uronic acids, acetic acid, and other minor sugars. Thus, by definition, hemicelluloses are short branched-chain heteropolysaccharides of mixed hexosans and pentosans that are easily hydrolyzed (1). The most common form of hemicellulose is composed of xylose polymer (xylan) that is found in large quantity in annual plants and deciduous trees and in smaller amounts in conifers. Xylan of grasses and cereals is generally characterized by the presence of L-arabinose, which is linked as a single unit side chain to a D-xylose backbone. Hydrolysis of hemicellulose in annual plants, agricultural wastes, and hardwood yields glucose, D-xylose, L-arabinose, and other minor sugars. Table 1 summarizes the amounts of hemicellulose in different plant materials; Table 2 summarizes the sugar composition of selected biomass materials after the hydrolysis.

HEMICELLULOSE HYDROLYSIS

Hemicellulose can be hydrolyzed to its sugar constituents either enzymatically or chemically. Chemical hydrolysis of hemicellulose is much easier to accomplish than the hydrolysis of cellulose due to the heterogeneous nature of hemicellulose, its chemical composition, and its low degree of polymerization. Most acids are good hydrolytic agents. The most common method of acid hydrolysis uses dilute mineral acids. Frequently during acid hydrolysis, xylose is

Table 1. Biomass Constiuents

Type of material	Hemicellulose (%)	Cellulose (%)	Lignins (%)
	Monocotyledons		
Stems	25 ~ 50	25 ~ 40	10 ~ 30
Leaves	80 ~ 85	15 ~ 20	–
Fibers	5 ~ 20	80 ~ 95	–
	Woods		
Hardwood (angiosperms)	24 ~ 40	40 ~ 55	18 ~ 25
Softwood (gymnosperms)	25 ~ 35	45 ~ 50	25 ~ 35
	Papers		
Newspaper	25 ~ 40	40 ~ 55	18 ~ 30
Wastepaper	10 ~ 20	60 ~ 70	5 ~ 10
Waste fibers	20 ~ 30	60 ~ 80	2 ~ 10

Source: From Ref. 2.

Table 2. Hemicellulose Neutral Carbohydrate Content of Agriculture Residues

	% of total sugars			
Plant Residues	Xylose	Arabinose	Glucose	Others
Corn residues				
Cobs	65.1	9.6	25.3	–
Leaves	59	9.4	29.1	2.5
Stalks	70.5	9.0	14.5	5.9
Husks	53.5	12.3	32.6	1.6
Pith	71.5	9.8	15.7	3
Fibers	63.8	6.6	26.8	2.8
Wheat straw	57.9	9.1	28.1	5
Soybean				
Stalks and leaves	59.9	6.6	6.1	27.4
Hulls	26.6	12.7	21	39.7
Sunflower				
Stalks	60.6	2.2	32.6	4.6
Pith	10.7	11.8	63.5	14
Flax Straw	64.6	12.8	1.2	21.4
Sweet clover hays	49.3	21.9	8.9	9.9
Peanut hulls	46.3	5	46.6	2.1
Sugarcane bagasse	59.5	14.5	26	–

Source: From Ref. 3.

degraded rapidly to furfural and condensation byproducts. Some degradation products are inhibitory to microorganisms. The toxic effect of those chemicals on cell growth and ethanol formation of *Pichia stipitis* and *Saccharomyces cerevisiae* are shown in Tables 3 and 4 (4). The use of a very dilute acid (less than 0.2% acid by weight), shorter reaction time, a lower reaction temperature, and rapid removal of hydrolytic agents are preferred to prevent the decomposition of sugars, particularly xylose. In recent years, a great deal of research has focused on the dilute acid hydrolysis of agricultural residues and wood products to obtain hemicellulose sugars (5,6).

BIOCONVERSION OF HEMICELLULOSE SUGARS

When cellulosic materials are subjected to hydrolysis, a mixture of monosaccharides is produced. The predominant sugars released are glucose, xylose, and to a lesser extent, arabinose, galactose, and mannose. When microorganisms are exposed to this sugar mixture, the phenomena of diauxic growth and differential rates of sugar utilization are often observed. The differential utilization of sugars by a given microorganism is dictated by sugar uptake, utilization rates, and by the degree of catabolite repression, which affects specific enzyme activity and biosynthesis. As a rule, microorganisms prefer glucose over galactose, followed by xylose and arabinose.

Because hemicellulose is abundant in nature and renewable, extensive research has been undertaken to convert hemicellulose-derived carbohydrates, particularly xylose, into useful products. In this article, the emphasis is on the utilization of hemicellulose-derived xylose into ethanol, xylitol, and 2,3-butanediol. Other products such as solvents (acetone and butanol) (7) and single-cell proteins (8) have been reviewed.

Table 3. The Effect of Potential Inhibitors on *Pichia stipitis* Growth and Ethanol Production from Xylose

Compound[a]	Concentration (g/L)	% Inhibition (Growth)	% Inhibition (EtOH)
Furfural	2	99	95
HMF	5	98.6	91.4
Acetate	10	36	69
HBA	0.75	70	84
SGA	0.75	62	80
Vanillin	0.5	88	89

Source: From Ref. 9.
[a]HMF, hydroxymethylfurfural; HBA, hydroxybenzaldehyde; SGA, siringaldehyde.

Table 4. The Effects of Potential Inhibitors on *S. cerevisiae* Growth and Ethanol Production from Glucose

Compound[a]	Concentration (g/L)	% Inhibition (Growth)	% Inhibition (EtOH)
Furfural	2	90	89
HMF	5	89	95
Acetate	10	48	27
HBA	0.75	53	37
SGA	0.75	61	54
Vanillin	0.5	51	30

Source: From Ref. 9.
[a]HMF, hydroxymethylfurfural; HBA, hydroxybenzaldehyde; SGA, siringaldhyde.

ETHANOL

Ethanol is the most studied metabolic product from xylose by fermentation. Ethanol can be used as a "neat" fuel substitute for gasoline. It can also be used as an oxygenated fuel extender by converting ethanol to ethyl tertiary-butyl ether (ETBE). Using ethanol as fuel and fuel extender provides energy security and diversity, improves global competitiveness, reduces trade deficits, revitalizes the agricultural industry, and promotes energy independence. It also improves the environment by improving urban air quality and by reducing the threat of global warming. Currently the production of ethanol from biomass is not competitive. Many research efforts were made to achieve the economical production of ethanol from biomass (for review see Ref. 9). In order to increase ethanol production efficiency from biomass, full and efficient utilization of hemicellulose carbohydrates is essential.

Conversion of Xylose to Ethanol by Yeasts

Yeasts are facultative organisms that have the ability to produce energy for their own use from suitable organic compounds under either aerobic or anaerobic conditions. Under aerobic conditions, sugars are metabolized to carbon dioxide and water for the production of energy and cell constituents. Under anaerobic conditions, most sugars are metabolized to ethanol by a process known as alcohol fermentation. Although many facultatively fermentative yeasts utilize xylose as carbon source for growth, the ability of those yeasts to produce ethanol from xylose is limited (10).

Yeast strains that utilize xylose often produce xylitol from xylose extracellularly as a normal metabolic activity, but only a few can produce significant quantities of ethanol. The prominent strains that produce ethanol from xylose include *Kluveromyces maxianus* (11), *Pachysolen tannophilus*, *Candida shihatae*, and *Pichia stipitis*. However, the efficient production of ethanol from xylose is limited by the regulation of dissolved oxygen as well as by the imbalance of cofactors in the metabolic pathway during xylose utilization. In recent years, there have been many efforts carried out to improve yeast strains to produce ethanol from xylose more efficiently.

Metabolic Pathway. Ethanol fermentation by yeasts from xylose can be divided into four distinctive steps. The first step is the reduction of xylose to xylitol, mediated by NADPH/NADH-linked xylose reductase (XR). This is followed by the oxidation of xylitol to xylulose, mediated by NAD-linked xylitol dehydrogenase (XDH). Xylulose-5-phosphate, the key intermediate, is generated from the phosphorylation of xylulose by xylulose kinase. Xylulose is then channeled into the pentose phosphate pathway for further metabolism (Fig. 1). In contrast, a great majority of bacteria convert xylose into its ketoisomer, xylulose, mediated by xylose isomerase (commonly known as glucose isomerase) as the initial step in xylose metabolism.

Ethanol Yield and Carbon Balance. Starting with 6 molecules of xylose, 10 molecules of ethanol can be produced using the combination of two pathways (Fig. 1). The net equation for the reactions leading to the production of ethanol from xylose is

$$6 \text{ xylose to } 10 \text{ ethanol} + 10\ CO_2$$

These values are the theoretical maximum yields (same as for glucose fermentation) that represent the maximum value, which can be achieved only if no sugar is assimilated to cell mass or oxidized through the TCA cycle. Of the total carbon contained in xylose, two-thirds goes to ethanol, and

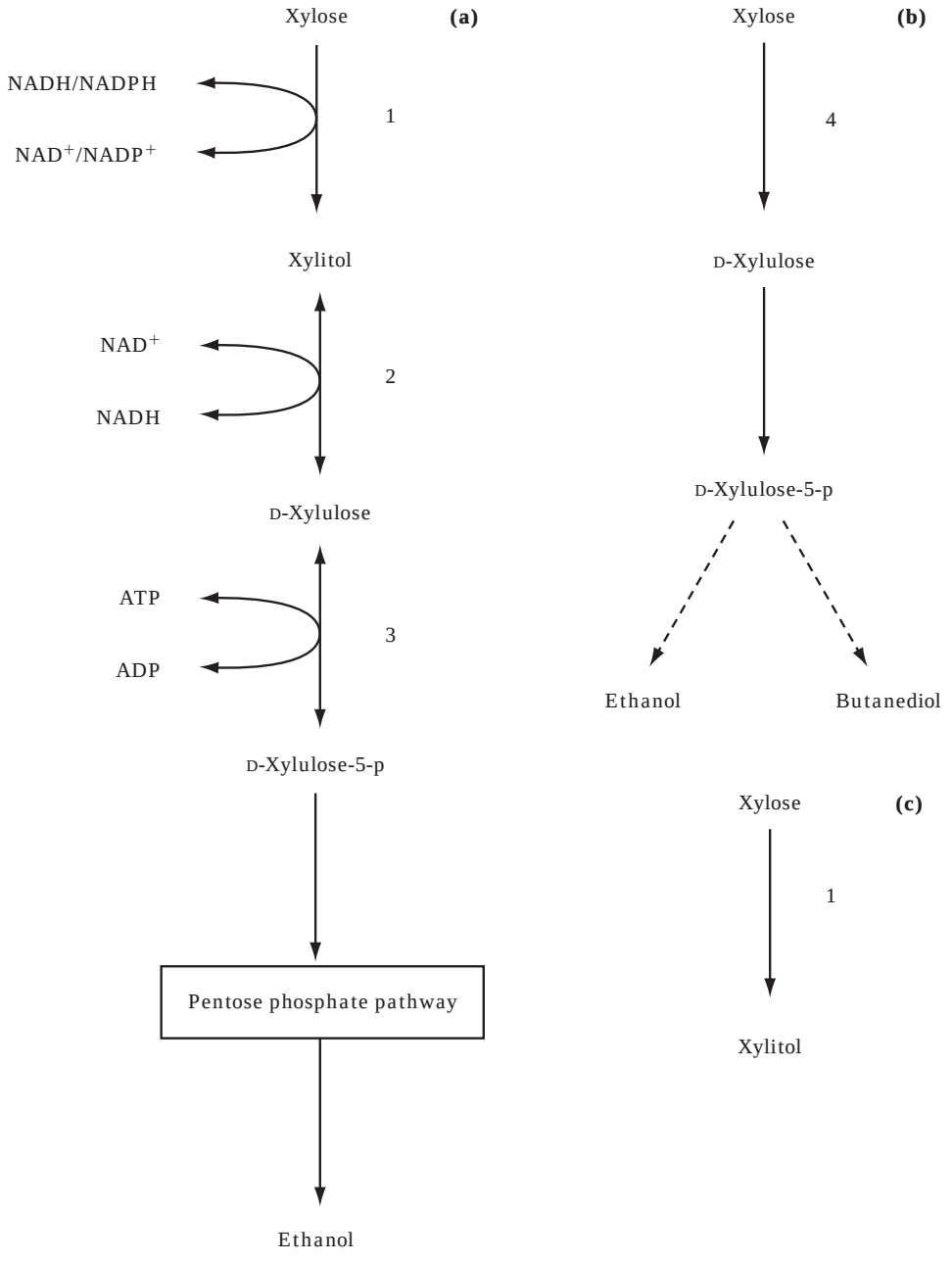


Figure 1. Abbreviated pathways for xylose utilization by microorganisms. (**a**) Yeast pathway; (**b**) bacterial pathway; (**c**) yeast xylitol pathway. (1) Xylose reductase; (2) xylitol dehydrogenase; (3) xylulokinase; (4) xylose isomerase.

the other one-third is lost as CO_2. On a weight basis, the yield of ethanol from both xylose and glucose is 51%.

Yeasts That Ferment Xylose to Ethanol. Following are the general characteristics of yeasts that ferment xylose to ethanol.

1. Oxygen requirement (12).
   a. Oxygen is required for the efficient uptake of xylose.
   b. Under low dissolved oxygen (DO) conditions, the electron transport system is not able to oxidize NADH efficiently, this causes the imbalance of $NADH/NAD^+$ that leads to the accumulation of xylitol.
   c. An increase in DO will enhance cell growth and xylose fermentation.
   d. When oxygen is in excess, TCA cycle activity is enhanced, which results in excessive cell growth and causes the reassimilation of ethanol produced.
   e. The reassimilation of ethanol leads to the accumulation of acetaldehyde and acetic acid.
   f. The optimal DO gives the lowest xylitol accumulation with the highest ethanol yield.
2. Enzyme cofactor imbalance
   a. Xylose reductase and xylitol dehydrogenase have different cofactor requirements.

   b. Different nitrogen sources (organic vs. inorganic) affects the level of xylitol dehydrogenase activity (13).
   c. Different DOs affect the balance of cofactors.
3. Low ethanol tolerance compared with that in *S. cerevisiae*
   a. Fermentation is inhibited at an ethanol concentration of 42 to 45 g/L.
4. Xylitol, glycerol, arabitol, and acetate are the common byproducts.
5. Higher sensitivity to potential fermentation inhibitors than glucose-fermenting yeast strains during glucose fermentation (Tables 3 and 4).
6. Specific ethanol productivity from xylose is at least one order of magnitude lower than that in glucose fermentation by *S. cerevisiae*.

Natural Yeast Strains. *P. tannophilus*, *C. shehatae*, and strains of *P. stipitis* are three of the most intensively studied naturally occurring yeast species that produce a significant quantity of ethanol from xylose.

Pachysolen tannophilus. *P. tannophilus* is an unusual yeast that was originally isolated by Wickerham (14) from wood sulfite liquor. This yeast was the first naturally occurring yeast species that was found to produce a significant quantity of ethanol from xylose (15). More detailed studies of this yeast were reported by Slininger et al. (16). As described in the literature, ethanol production from xylose by *P. tannophilus* exhibited the following characteristics:

1. A small amount of oxygen was required for optimal ethanol production.
2. A lag phase of about 20 h was observed before the accumulation of ethanol.
3. Ethanol yield was reported to be 0.34 g ethanol per g xylose consumed (68% of the theoretical value).
4. Optimal temperature for cell growth and ethanol production was 32 °C.
5. Optimal pH for cell growth was 2.5–4.5, and optimal pH for ethanol production was 2.5.
6. Ethanol is reassimilated under high DO conditions.
7. Ethanol production was subjected to substrate (55 g/L xylose) and product (30 g/L ethanol) inhibition.
8. Xylitol was accumulated during fermentation.

Candida shehatae. The ability of *C. shehatae* to produce ethanol from xylose was first reported by du Preez and van der Walt (17) in 1983. Since then, many studies have been conducted relating to the properties of this yeast. As described in the literature (18–20), ethanol production from xylose by *C. shehatae* exhibited the following characteristics:

1. A small amount of DO is required for maximal ethanol production.
2. Cell growth was inhibited at an ethanol concentration of 37.5 g/L, and cell viability was significantly lower in the presence of up to 50 g/L ethanol.
3. Glucose and mannose are the preferred substrate in a sugar mixture.
4. Ethanol yield can be as high as 84% of the theoretical value (0.43 g ethanol per g xylose consumed).
5. The maximum ethanol production rate was 0.48 g/g/h at 50 g/L xylose.
6. Xylitol was accumulated under very limited DO conditions due to the accumulation of NADH.
7. Under high DO conditions, the ethanol produced was reassimilated.

Pichia stipitis. *P. stipitis* is the most effective natural yeast for the conversion of xylose to ethanol. This yeast species shares many characteristics with its close relative *C. shehatae*. Toivola et al. (21) performed a systemic screening program with type strains of some 200 yeast species and identified *P. stipitis* as one of the yeast species that produces ethanol from xylose. There are many studies that explored the properties of this yeast in relation to its oxygen requirement, ethanol tolerance, enzyme cofactors balance, and so on. As described in the literature (22,23), ethanol production from xylose by *P. stipitis* exhibited the following characteristics:

1. Trace oxygen was required to sustain cell growth and maintenance.
2. *P. stipitis* has a higher ethanol (64 g/L) tolerance than other xylose-fermenting yeasts.
3. Compared with *C. shehatae*, *P. stipitis* has a lower specific ethanol productivity but a higher ethanol yield.
4. Up to 57 g/L of ethanol can be accumulated by *P. stipitis*.
5. Xylitol production by *P. stipitis* was lower than other xylose-fermenting yeasts.
6. The optimum temperature for ethanol production was 25 °C.
7. The optimum pH range for growth and fermentation was 4–7.
8. Ethanol was reassimilated at high DO levels, and acetate was formed when ethanol was metabolized.

Genetically Modified Yeasts. Attempts to modify the xylose-fermenting pathway in *S. cerevisiae* using the xylose isomerase gene from various bacterial sources have not been successful (24) in spite of previous reports showing the ability of glucose-fermenting yeast strains, *S. cerevisiae* (25) and *Schizosaccharomyces pombe* (26), to produce ethanol from xylulose or from xylose in the presence of xylose isomerase (glucose isomerase) (27) (Fig. 2). Other approaches of improving yeast strains through genetic recombination have met with some encouraging results.

Saccharomyces cerevisiae. Kotter and Ciriacy (28) studied xylose utilization of a *S. cerevisiae* transformant that expressed two key enzymes (xylose reductase and xylitol dehydrogenase) derived from *P. stipitis* in xylose metabo-

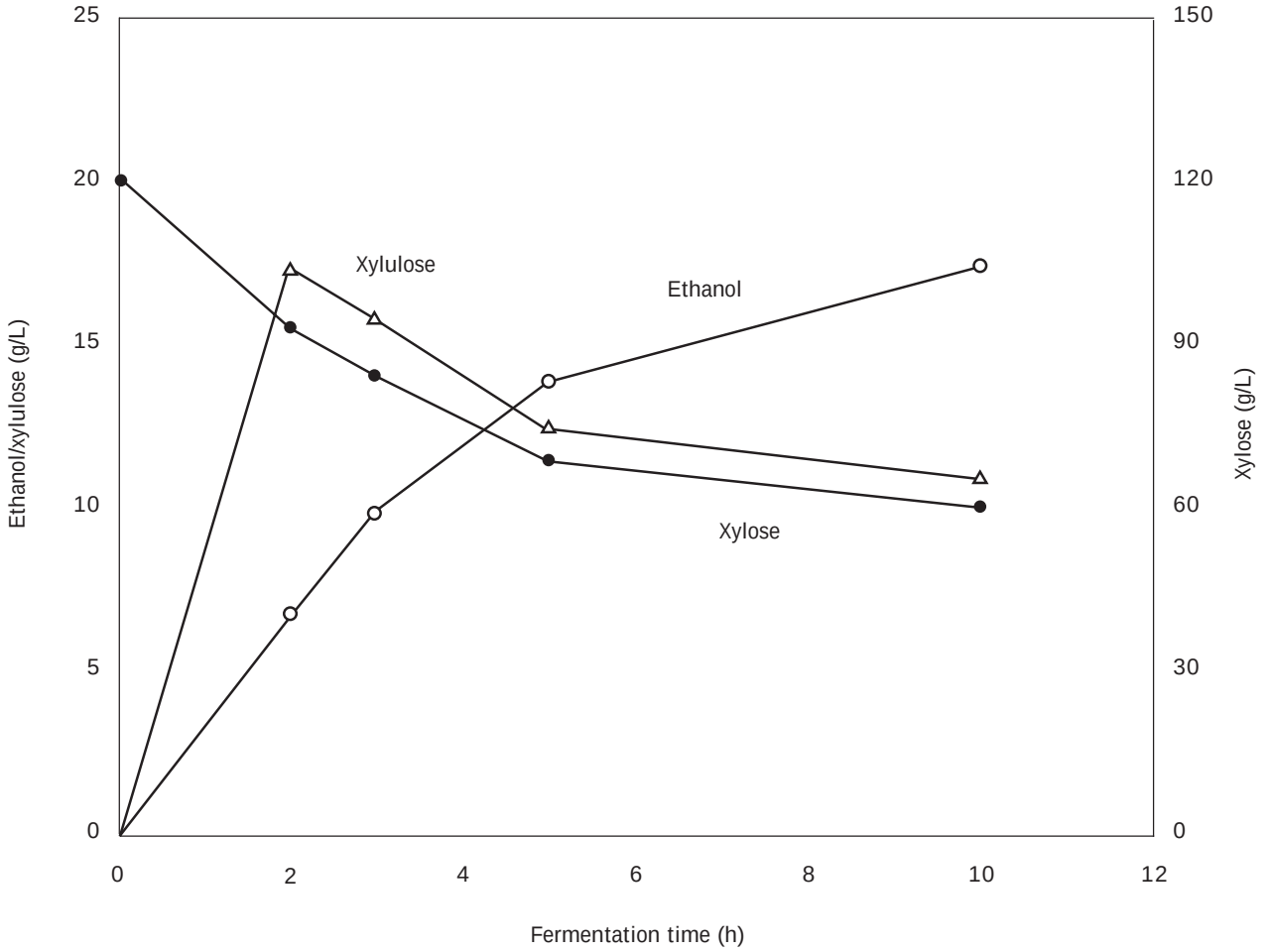


Figure 2. Production of ethanol from xylose by Baker's yeast in the presence of xylose isomerase. *Source*: From Ref. 27.

lism. Under fermentative conditions, the transformant converts only half of available xylose to xylitol and ethanol. The acquired ability to ferment xylose was interpreted to be a result of the duel cofactor dependence of XR and the generation of NADPH by the pentose phosphate pathway. The limitation of xylose utilization in the transformant was likely caused by an insufficient capacity of the non-oxidative pentose pathway, as indicated by the accumulation of sedoheptulose-7-phosphate and the absence of fructose-1,6-diphosphate and pyruvate accumulation.

Recombinant Yeast Saccharomyces 1400 (pLNH33). The recombinant yeast strain *Saccharomyces* 1400 (pLNH33) was developed using the high-ethanol-tolerance *Saccharomyces* yeast 1400 (29) as the host and cloned with the xylose reductase and xylitol dehydrogenase genes from *P. stipitis*, and the xylulokinase gene from *S. cerevisiae*. *Saccharomyces* 1400 (pLNH33) was shown to ferment both sugars in a 1:1 mixture of glucose (52.8 g/L) and xylose (56.3 g/L) (Fig. 3) to ethanol under microaerobic condition in a relatively high yield (84% of theoretical value) (30). A final ethanol concentration obtained was 50 g/L after 48 h of incubation. This recombinant was also shown to produce ethanol from glucose and xylose that were derived from corn fiber during simultaneous saccharification and fermentation process in the presence of a fungal cellulase (31).

Conversion of Xylose to Ethanol by Bacteria

A wide range of bacterial species utilizes D-xylose and L-arabinose as carbon and energy sources. In most case, the direct isomerization of aldopentoses to their corresponding ketoses is the first step in pentose metabolism. For example, D-xylose is converted into D-xylulose, and L-arabinose is converted into L-ribulose. L-Ribulose can be converted into D-xylulose by epimerase. D-xylulose is the key intermediate for further metabolism. For this reason, most bacterial species can readily utilize L-arabinose as well.

Bacterial species, particularly those belonging to *Klebsiellae and Erwinia*, and *Escherichia coli*, are well known for their ability to metabolize hexoses and pentoses to produce either neutral compounds (butanediol, acetoin, and ethanol) or mixed acids and ethanol: Research on the production of ethanol from pentoses by bacteria has revolved around the improvement of such bacteria through genetic recombination.

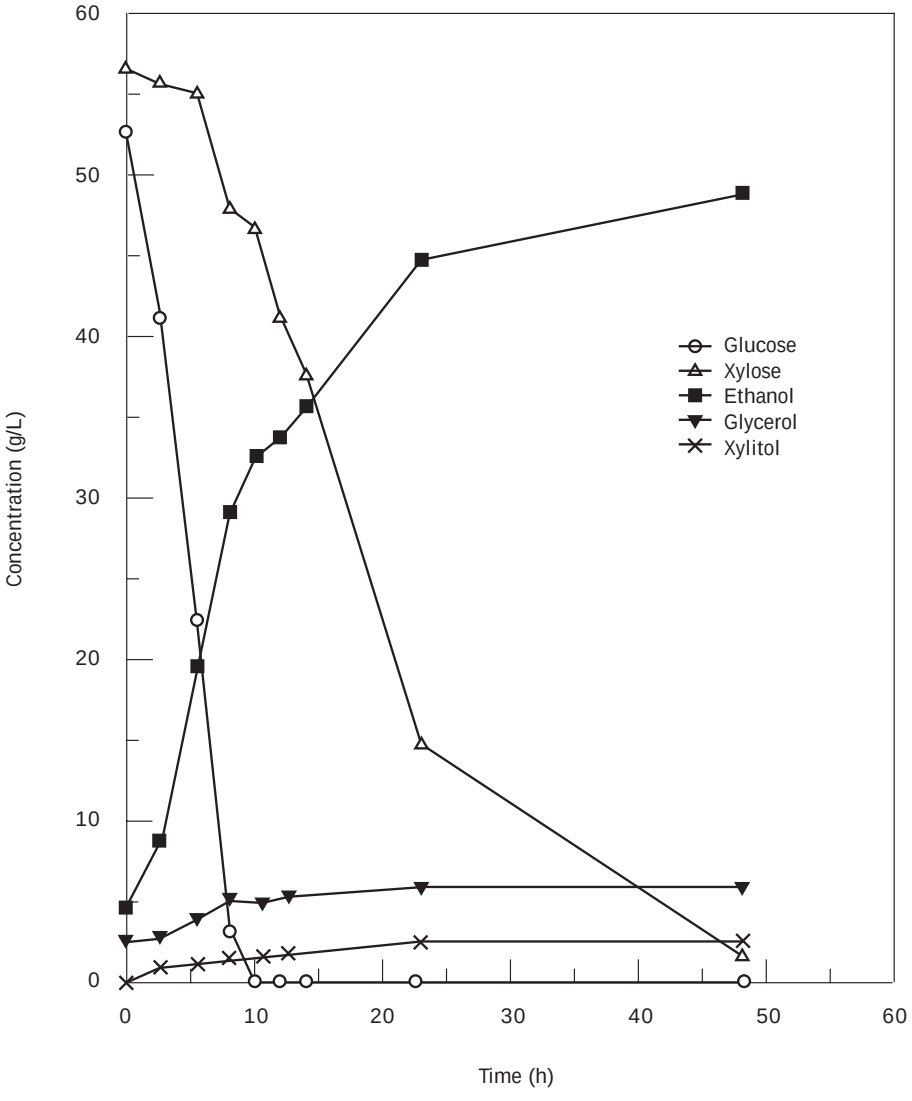


Figure 3. Fermentation of a glucose–xylose mixture by recombinant yeast, *Saccharomyces* 1400 (pLNH33). *Source*: From Ref. 30.

***Klebsiella oxytoca* M5A1.** Pyruvate decarboxylase and alcohol dehydrogenase genes encoding the ethanol pathway from *Zymomonas mobilis* were transferred into *K. oxytoca* (32) The transformant, *K. oxytoca* M5A1, was able to divert pyruvate from the normal fermentative pathway to ethanol production. *K. oxytoca* M5A1 was able to produce ethanol from both glucose and xylose in excess of 40 g/L with an efficiency of 0.48 g of ethanol per g xylose and 0.5 g of ethanol per g glucose. The maximal volumetric productivity for both sugars was similar (ca. 2 g/L).

***Klebsiella planticola*.** Multicopy plasmids containing the *Z. mobilis* pyruvate decarboxylase gene were inserted into *K. planticola*. The transformant was shown to produce 31.6 g/L ethanol from a mixture of sugars (79.6 g/L) that included xylose (33). The yield of acids (formate, acetate, lactate) and butanediol was reduced significantly.

***Escherichia coli*.** Pyruvate decarboxylase of *Z. mobilis* was transferred into *E. coli* (34). The transformant was able to produce 39.2 g/L ethanol from 80 g/L xylose with an indicated yield of 96% of the theoretical value. The maximal volumetric productivity was 0.7 g ethanol per liter per hour in batch fermentation (35).

***Zymomonas mobilis*.** Two operons encoding xylose assimilation enzymes (xylose isomerase and xylulokinase) and pentose phosphate pathway enzymes (transaldolase, and transketolase) were constructed and introduced into *Z. mobilis* (36). The transformant, CP4 (pZB5), was able to not only grow but also ferment xylose anaerobically into ethanol. This transformant also acquired the ability to ferment glucose and xylose simultaneously with a slower rate of xylose than glucose utilization. Anaerobic fermentation using equal amounts of glucose (25 g/L) and xylose (25 g/L) in mixture gave an ethanol yield of 25 g/L ethanol with 95% of the theoretical yield within 30 h.

Conversion of Xylose to Ethanol by Mycelial Fungi

Several filamentous fungi belonging to the genera *Fusarium*, *Mucor*, *Rhizopus*, and *Monilia* are known to ferment

both glucose and xylose to ethanol (37). The early work on xylose metabolism in *Fusarium lini* indicated the potential for ethanol production from xylose by this fungus (38). Overall, the slow rate of ethanol production from xylose by fungi is undesirable because it would add to capital costs in industrial application. However, because additional ethanol can be obtained from xylose, the use of mold for this purpose is an option.

XYLITOL

Xylitol is a five-carbon sugar alcohol that occurs naturally in small quantities (less than 0.9 g/100g) in many fruits and vegetables that constitute part of human diet. It is the only so-called second-generation polyol sweetener that is allowed to have specific health claims in some world markets. Unfortunately, it is one of the most expensive polyol sweeteners. Availability and cost of production are the obstacles impeding the increased use of xylitol. On an industrial scale, xylitol is produced through chemical reduction of xylose derived from hemicellulosic hydrolyzate. Birch hydrolyzate is the most common substrate. The chemical process is expensive because of the high temperature and high pressure required for hydrogenation of xylose. Furthermore, extensive steps for separation and purification add to the cost.

Xylitol can be formed as a metabolic intermediary of D-xylose utilization by a wide range of microorganisms. It is converted from xylose either by NADPH-dependent aldehyde reductase, or from xylulose by NADH-dependent xylitol dehydrogenase (Fig. 1). Many natural yeast strains have the ability to produce a significant quantity of xylitol from xylose extracellularly as a normal metabolic activity. The prominent strains that produce xylitol include *Candida* sp. (39), *Candida boidinii* (40), *Candida guillermondii* (41), *Candida parapsilosis* (42), *Candida tropicalis* (43), and *Debaryomyces hansenii* (44). However, xylose is an expensive substrate for xylitol production. Recent advances in obtaining xylose-rich hemicellulose hydrolyzates from lignocellulosic materials can provide ample substrate for the biological production of xylitol inexpensively. As a result, xylitol can be potentially produced from such materials as an option for effective utilization of lignocellulosic biomass.

Usually the weight yield of xylitol produced from xylose by naturally occurring yeast species is around 75 to 80%. Two of the reasons for the relatively low yield are the growth of cells and the need for energy maintenance. Ideally, xylitol overaccumulation would occur if the pathway from xylitol to xylulose was cut off, but this is unrealistic because of the redox balance. The next option is to reduce the activity of xylitol dehydrogenase. Therefore, the mutation–selection technique, based on differential xylose–xylitol utilization, can be applied to obtain high-xylitol-producing organisms. Based on this principle, a mutant strain from a proficient xylose-utilizing *Candida* yeast was selected; it had xylitol weight yields as high as 90% from xylose (45).

The production of xylitol from hemicellulosic materials has been considered as an alternative approach in the utilization of the xylose fraction of biomass substrates and the production liquid of fuels and chemical feedstocks. Over the years, many studies have been conducted utilizing the hemicellulose portion of agricultural residues for xylitol production. The materials studied include sugarcane bagasse (39,46), *Eucalyptus* wood (47), and rice straw (48).

2,3-BUTANEDIOL

Butanediol (butylene glycol) is a colorless and odorless liquid with a high boiling point (180–184 °C) and a low freezing point (60 °C). The heating value of butanediol (27,198 J/g) is similar to that of ethanol (29,055 J/g) and methanol (22,081 J/g). It can be dehydrated to methylethyl ketone (MEK) and used as an octane booster for gasoline or as high-grade aviation fuel. MEK can be dehydrated to 1,3-butadiene and dimerized to styrene. Therefore, butanediol has a diverse industrial usage particularly as a polymeric feedstock in addition to its use for manufacturing butadiene or antifreeze (49).

Among many isomers, 2,3-butanediol is the only one that can be produced by microorganisms. Bacterial species of *Klebsielleae* are known to metabolize hexoses and pentoses derived from hemicellulose to produce 2,3-butanediol, acetoin, and ethanol as metabolic products. Other groups of enterobacteria such as *Erwinia* also utilize hemicellulose-derived carbohydrates to produce mixed acids and 2,3-butanediol. Research on the production of butanediol and its immediate precursor, acetoin, from renewable biomass has revolved around the microorganisms *Klebsiella oxytoca (K. pneumonia)* and *Enterobacter aerogenes (Aerobacter aerogenes)*.

Butanediol-producing bacteria utilize all the major sugars (hexoses and pentoses) to produce butanediol in high yield and high concentration (up to 10%, w/v) under optimum conditions (e.g., DO, pH, etc.) (49,50). Three enzymes, α-acetolactate synthetase (ALS), α-acetolactate decarboxylase (ALDC), and acetoin reductase (AR), are directly responsible for the accumulation of 1 mole of butanediol from 2 moles of pyruvate (51). In *K. terrigena* and *E. aerogenes*, these enzymes are encoded by the gene *budB* (ALS), *budA* (ALDC), and *budC* (AR), respectively. The three genes are organized in an operon and activated by transcriptional products of the *budR* gene (52). The expression of the operon is optimal under anaerobic conditions at low pH environments and in the presence of a low concentration of acetate. The metabolic pathway leading to butanediol formation is shown in Figure 4.

DO has a profound effect on butanediol production and end-products distribution. Oxygen supply is important because it determines the flow of carbon source via the respiratory pathway versus the butanediol-producing fermentative pathway. The role of oxygen supply on biomass and butanediol accumulations has been studied extensively. Butanediol yield may be optimized by minimizing the available oxygen. However, with a small oxygen transfer rate, the cell yield will be low, causing the total reaction rate to slow down. In the presence of excessive oxygen, metabolism leads to the oxidation of NADH to NAD^+. A high NAD^+/NADH ratio leads to excessive biomass accu-

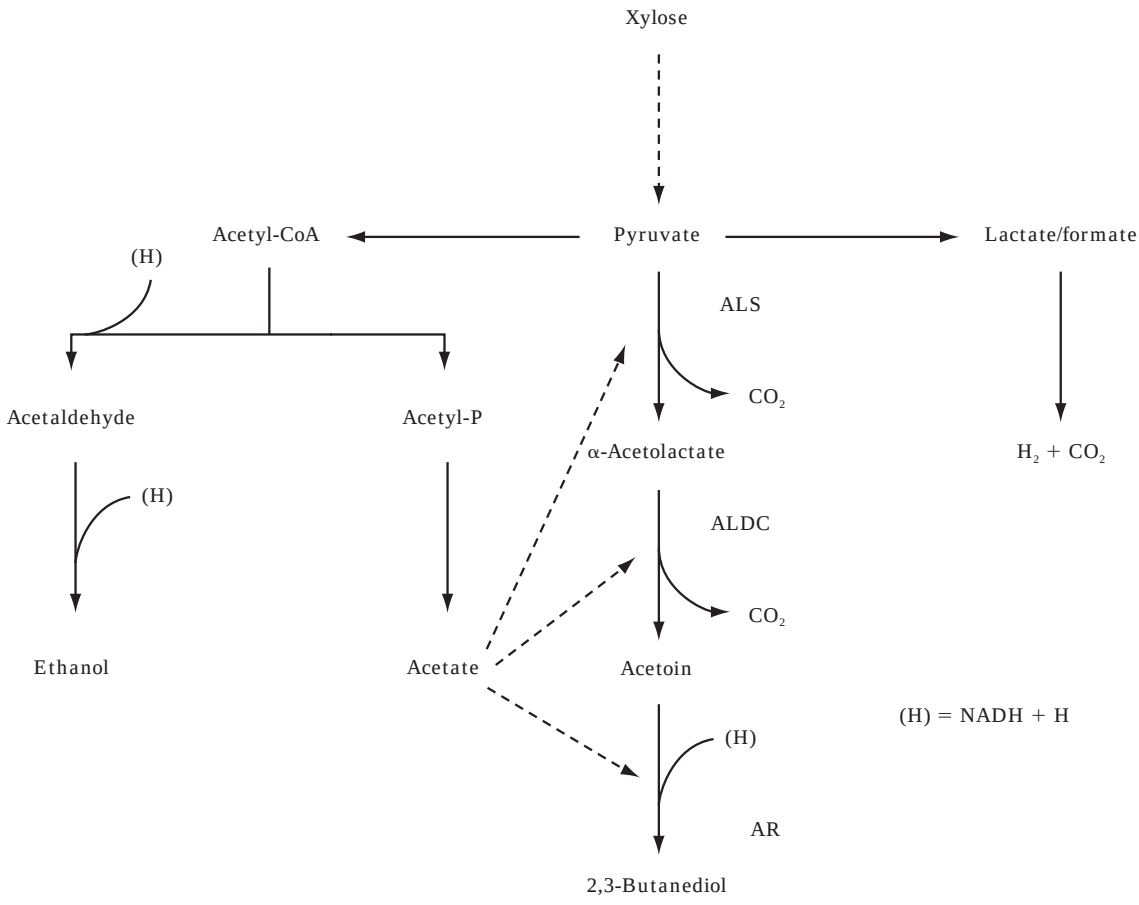


Figure 4. The metabolic pathway leading to the production of 2,3-butanediol and the role of acetic acid. ALS, α-acetolactate synthetase; ALDC, α-acetolactate decarboxylase; AR, acetoin reductase.

mulation. There is a general agreement among the investigators that the oxygen supply rate is perhaps the most critical criterion for the performance of butanediol fermentation (see Ref. 53 for detail).

In addition to DO, the pH environment also exerts great influence on the product distribution and yield. In enterobacteria, the type and ratio of fermentation products formed depend on two biochemical constraints: the need to maintain redox balance, and the requirement for maintaining the pH of the medium in a physiological range. Therefore, in a different pH environment, different relative amounts of end products are formed, resulting in an increased proportion of neutral compounds and decreased pH of the medium (53). Under a neutral or high-pH environment, bacteria tend to utilize a carbon source for cell growth and organic acid production. It has been shown that when the pH is lower than 6, over 50% of the pyruvate was channeled into butanediol production. Under this condition, enzymes of the butanediol pathway can constitute as much as 2.5% of the total protein in *E. aerogenes* (54).

Zeng et al. (55), investigated the effect of pH on growth and product formation from glucose by *E. aerogenes* in a continuous culture operation and found the pH range of 5.5–6.5 is optimal for butanediol and acetoin production. A similar pH optimum was also observed in *K. oxytoca*. In general, the biomass concentration increases steadily with increased pH. At high pH, butanediol concentration decreases and there is a concomitant increase of acetic acid production. Acetic acid has a dual role in the regulation of butanediol formation. It serves as activator for butanediol accumulation at low concentration. At a concentration of 10 g/L or higher, it inhibits butanediol production (56). The strength of acetic acid inhibition depends on the concentration of its undissociated form, HAc. and the concentration of HAc is in turn determined by the pH.

The production of butanediol from lignocellulosic materials has been considered as an alternate approach in the conversion of biomass substrates to liquid fuels and chemical feedstocks. Over the years, there have been many studies utilizing agricultural residues for butanediol production. The materials studied include hardwood hemicellulose fraction (57,58), wheat and barley straws (59), corn stover (59), and corn cobs (60).

BIBLIOGRAPHY

1. T.E. Timell, *Adv. Carbohyd. Chem.* **19**, 247–253 (1964).
2. E.B. Cowling and T.K. Kirk, *Biotechnol. Bioeng. Symp.* **6**, 95–123 (1976).

3. C.S. Gong, L.F. Chen, M.C. Flickinger, and G.T. Tsao, *Adv. Biochem. Eng. Biotechnol.* **20**, 93–118 (1981).
4. J.P. Delgenes, R. Moletta, and J.M. Navarro, *Enzyme Microb. Technol.* **19**, 220–225 (1996).
5. R. Torget, P. Werdene, M. Himmel, and K. Grohman, *Appl. Biochem. Biotechnol.* **24/25**, 115–126 (1990).
6. D.J. Schnell, R. Torget, A. Power, P.J. Walter, K. Grohman, and N.D. Hinman, *Appl. Biochem. Biotechnol.* **28/29**, 87–97 (1991).
7. E.T. Papoutsakis, in O. Kitani and C.W. Hall eds., *Biomass Handbook*, Gordon and Breach Science Publishers, New York, 1989, pp. 271–286.
8. J.H. Litchfield, in M. Moo-Young ed., *Biomass Conversion Technology: Principles and Practice*, Pergamon Press, Oxford, 1987, pp. 113–122.
9. L. Olsson and B. Hahn-Hagerdal, *Enzyme Microb. Technol.* **18**, 312–331 (1996).
10. J.A. Barnett, *Adv. Charbohydr. Chem. Biochem.* **32**, 125–234 (1976).
11. A. Margaritis and P. Bajpai, *Appl. Environ. Microbiol.* **44**, 1039–1041 (1982).
12. J.M. Laplace, J.P. Delgenes, R. Moletta, and J.M. Navarro, *Appl. Microbiol. Biotechnol.* **36**, 158–162 (1991).
13. S. Palnitkar and A. Lachke, *Can. J. Microbiol.* **38**, 258–260 (1992).
14. L.J. Wickerham, in J. Lodder ed., *The Yeasts*, North Holland, Amsterdam, 1971, p. 448–454.
15. H. Schneider, P.Y. Wang, Y.K. Chan, and R. Maleszka, *Biotechnol. Lett.* **3**, 89–92 (1981).
16. P.J. Slininger, R.J. Bothast, J.E. van Cauwenberge, and C.P. Kurtzman, *Biotechnol. Bioeng.* **24**, 371–384 (1982).
17. J.C. du Preez and J.P. van der Walt, *Biotechnol. Lett.* **5**, 357–362 (1983).
18. J.C. du Preez, M. Bosch, and B.A. Prior, *Appl. Microbiol. Biotechnol.* **25**, 521–525 (1987).
19. J.C. du Preez, B. van Driessel, and B.A. Prior, *Appl. Microbiol. Biotechnol.* **30**, 53–58 (1989).
20. J.R. Kastner, M. Ahmad, W.J. Jones, and R.S. Roberts. *Biotechnol. Bioeng.* **40**, 1282–1285 (1992).
21. A. Toivola, D. Yarrow, E. van den Bosch, J.P. van Dijken, and W.A. Scheffers, *Appl. Environ. Microbiol.* **47**, 1221–1223 (1984).
22. J.C. DuPreez, B. van Driessel, and B.A. Prior, *Appl. Microbiol. Biotechnol.* **30**, 53–58 (1989).
23. P.J. Slininger, L.E. Branstrator, R.J. Bothast, W.R. Okos, and M.R. Ladisch, *Biotechnol. Bioeng.* **37**, 973–960 (1991).
24. R. Amore, M. Wilhelm, and C.P. Hollenberg, *Appl. Microbiol. Biotechnol.* **30**, 351–357 (1989).
25. C.S. Gong, L.F. Chen, M.C. Flickinger, L.C. Chiang, and G.T. Tsao, *Appl. Environ. Microbiol.* **41**, 430–436 (1981).
26. P.Y. Wang, C. Shopsis, and H. Schneider, *Biochem. Biophys. Res. Commun.* **94**, 248–254 (1980).
27. L.C. Chiang, C.S. Gong, L.F. Chen, and G.T. Tsao, *Appl. Environ. Microbiol.* **42**, 284–289 (1981).
28. P. Kotter and M. Ciriacy, *Appl. Microbiol. Biotechnol.* **38**, 776–783 (1993).
29. M.T. D'Amore, C.J. Panchal, J. Russell, and G.G. Stewart, *Crit. Rev. Biotechnol.* **9**, 287–305 (1990).
30. M.S. Krishnan, N.W.Y. Ho, and G.T. Tsao, in B.C. Saha, and J. Woodward eds., *Fuels and Chemicals from Biomass*, ACS Symposium Series 666, American Chemical Society, Washington, D.C., 1997, pp. 74–92.
31. M. Moniruzzaman, B.S. Dien, C.D. Skory, Z.D. Chen, R.B. Hespell, N.W.Y. Ho, B.E. Dale, and R.J. Bothast, *World J. Microbiol. Biotechnol.* **13**, 341–346 (1997).
32. K. Ohta, D.S. Beall, J.P. Mejia, K.T. Shanmugam, and L.O. Ingram, *Appl. Environ. Microbiol.* **57**, 2810–2815 (1991).
33. J.S. Tolan and R.K. Finn, *Appl. Environ. Microbiol.* **53**, 2039–2044 (1987).
34. F. Alterthum and L.O. Ingram, *Appl. Environ. Microbiol.* **55**, 1943–1948 (1989).
35. K. Ohta, F. Alterthum, and L.O. Ingram, *Appl. Environ. Microbiol.* **56**, 463–465 (1990).
36. M. Zhang, C. Eddy, K. Deanda, M. Finkelstein, and S. Picataggio, *Science* **267**, 240–243 (1995).
37. P. Ueng and C.S. Gong, *Enzyme Microb. Technol.* **4**, 169–171 (1982).
38. M.G. White and J.J. Williams, *Biochem. J.* **22**, 583–591 (1928).
39. L.F. Chen and C.S. Gong, *J. Food Sci.* **50**, 226–228 (1985).
40. E. Winkelhausen, P. Pittman, S. Kuzmanova, and T.W. Jeffries, *Biotechnol. Lett.* **18**, 753–758 (1996).
41. V. Meyrial, J.P. Delgenes, R. Moletta, and J.M. Navarro, *Biotechnol. Lett.* **13**, 281–286 (1991).
42. V. Nolleau, L. Preziosi-Belloy, J.P. Delgenes, and J.M. Navarro, *Curr. Microbiol.* **27**, 191–197 (1993).
43. H. Horitsu, Y. Yahashi, K. Takamizawa, K. Kawai, T. Suzuki, and N. Watanable, *Biotechnol. Bioeng.* **40**, 1085–1091 (1992).
44. J.C. Roseiro, M.A. Peito, F.M. Girio, and T. Amaral-Collaço, *Arch. Microbiol.* **156**, 484–490 (1991).
45. J. Lu, L.B. Tsai, C.S. Gong, G.T. Tsao, *Biotechnol. Lett.* **17**, 167–170 (1995).
46. M.J. Pfeifer, S.S. Silva, M.G.A. Felipe, I.C. Roberto, and I.M. Mancilha, *Appl. Biochem. Biotechnol.* **57/58**, 423–430 (1996).
47. J.C. Parajo, H. Dominguez, and J.M. Dominguez, *Bioprocess Eng.* **13**, 125–131 (1995).
48. I.C. Roberto, M.G.A. Felipe, I.M. Mancilha, M. Vitola, S. Sato, and S.S. Silva, *Bioresource Technol.* **51**, 255–257 (1995).
49. S.K. Garg and A. Jain, *Biosource Technol.* **51**, 103–109 (1995).
50. N.B. Jansen, and G.T. Tsao, *Adv. Biochem. Eng. Biotechnol.* **27**, 85–99 (1983).
51. R.J. Magee and N. Kosaric, *Adv. Appl. Microbiol.* **32**, 89–161 (1987).
52. D. Mayer, V. Schlensog, and A. Bock, *J. Bacteriol.* **177**, 5261–5269 (1995).
53. I.R. Booth, *Microbiol. Rev.* **49**, 356–378 (1985).
54. L. Johansen, K. Bryn, and F.C.J. Stormer, *J. Bacteriol.* **123**, 1124–1130 (1975).
55. A.P. Zeng, H. Biebl, and W.D. Deckwer, *Appl. Microbiol. Biotechnol.* **33**, 485–489 (1990).
56. E.K.C. Yu and J.N. Saddler, *Appl. Environ. Microbiol.* **44**, 777–784 (1982).
57. F.R. Frazer, and T.A. McCaskey, *Enzyme Microb. Technol.* **13**, 110–115 (1991).
58. B.P. Grover, S.K. Garg, and J. Verma, *World J. Microbiol. Biotechnol.* **6**, 328–332 (1990).
59. E.K.C. Yu, L. Deschatelets, N. Levitin, and J.N. Saddler, *Biotechnol. Lett*, **6**, 611–614 (1984).
60. N. Cao, Y. Xia, C.S. Gong, and G.T. Tsao, *Appl. Biochem. Biotechnol.* **63/65**, 129–139 (1997).

HUMAN AND PRIMATE CELL LINES

SALLY WARBURTON
ECACC
Wiltshire, United Kingdom

KEY WORDS

Adenovirus
Bioreactor
Carcinoma
Cytokine
Diploid
Glycosylation
Immortalization
Oncogene
Transformation
Vaccine

OUTLINE

INTRODUCTION

Before the advent of monoclonal antibodies and genetically manipulated cell lines, the major use of primate and human cells was as a substrate for viral vaccine manufacture (e.g., polio, rubella, and measles in human diploid cells MRC 5 or WI 38 [1]).

It was not until the mid 1970s that it was considered that continuous cell lines may also prove useful for vaccine production. At the same time, there was a great deal of interest in the use of interferon in clinical studies. The major source was from primary human leukocyte cultures, but the amount produced was very limited.

Exploitation of animal cell lines in the field of biotechnology as opposed to bacteria, filamentous fungi, and yeast has occurred relatively recently because of complex culture media, slow population doubling times, and low productivity levels of cells. The ability of animal cells to glycosylate proteins in an active form, unlike bacteria, has led to new developments in manufacturing, purification, and recombinant DNA technology in the engineering of cells for production of therapeutic proteins. Strict safety standards must be set for any biological manufacturing process involving eukaryotic protein production, research into the adaptation of suitable cell lines, the latest bioreactor development, immortalization methods, and cell products are discussed in this article.

HUMAN CELL LINES

Human Diploid Cells

In the 1950s there was an urgent need to produce a vaccine to combat polio, which led to the use of primary monkey kidney cells for growing the virus. This was successful until it was discovered after a large amount of vaccine had been released that monkeys were infected with wild simian viruses (e.g., SV40) (2).

The unreliability of primary cell culture led to the development of human diploid cells, such as MRC 5 and WI 38 (1), and human fetal lung fibroblasts, which were not infected by SV 40 and are now the preferred substrate.

Human diploid cells are particularly susceptible to viruses pathogenic to humans (e.g., polio, echo, arbo, rhino, measles, rubella, and coxsackie A9) and are very valuable for their isolation and development of vaccines (1–3).

These two cell lines are nontumorigenic, essential when considering using human diploid cells for live virus vaccines in humans (4); have a finite life span (because they have no malignant characteristics) (5); and begin to senesce after passage (5). These cells must be cryopreserved to prevent cross-contamination, genetic drift, microbial contamination, and maintenance of low passage stocks (6–9).

MRC 5 and WI 38 are free of latent and oncogenic viruses and are genetically stable; therefore, they are the permitted cell lines used for producing products such as vaccines for human use (5), although MRC 5 is preferred because of its higher replication rate and better tolerance to environmental changes (1).

Human Carcinomas

Useful in the study of cancer and drug therapy, human carcinomas are also useful models for the study of normal and pathophysiological events in vivo (10). Two human colon cell lines particularly useful are CaCo 2 and HT 29 because they undergo enterocyte differentiation in vitro (11). They have been widely used for studies on the barrier properties of the human intestine for drug transport and the cytoprotective activity of potential drugs and their pharmaceutical formulations (12).

A rotating wall vessel (RVW) designed to simulate cell cultivation conditions on earth in the microgravity of space has been set up using these cell lines (13). Reduced shear forces allow the cells to grow in three-dimensional form, to differentiate and to mimic the in vivo tissues. This type of bioreactor could provide an environment for cell cultures, enabling novel cell aggregate configurations that may be

usefully applied in the field of pathogenesis of infection and tissue engineering.

PRIMATE CELL LINES

Vero

The Vero cell line was established in 1962 to overcome the problem of using primary monkey kidney cells that were infected with the SV 40 virus (14–16). The use of primary tissue for poliovirus vaccine became more difficult because of international agreements on the conservation of wild animals. The Institute Merieux in France extensively tested the Vero cell line, demonstrating that Vero cells show no tumorigenicity, contain no aberrant viruses, and sustain efficient proliferation of various vaccine viruses (16). As a result, killed polio vaccine using Vero cells instead of primary monkey kidney cells has been available in France since 1984. Production is carried out on microcarriers in large-scale bioreactors.

Vero is a classical cell line that is still used as a capable substrate for virus reproduction of a wide variety of completely unrelated viruses, such as arboviruses, hemorrhagic fever viruses, rubella, measles, herpes, and growth of rickettsiae. (Unfortunately, Vero cells do not culture wild-type influenza viruses, which still need to be cultured on primary monkey kidney cells for identification.) Vero cells have been one of the most powerful basic resources for the entire field of animal virology in the past quarter of a century and will be a major cell substrate for virology and biotechnology in the years to come.

Other primate cell lines, such as COS 1 and COS 7 African green monkey kidney cells transformed with the SV 40 virus, have been found along with Vero cells to express HIV-1-like particles once transfected (17). This important development offers an alternative to the use of live virus vectors for the production and evaluation of AIDS vaccines based on noninfectious particles.

Table 1 summarizes some of the versatile human and primate cell lines that are currently being used in the biotechnology industry and the products they have to offer.

IMMORTALIZATION

Primary cell lines can be maintained in vitro for a limited period. As cells approach senescence (34), they are unable to synthesize DNA. They reach a crisis point where cell division ceases and the cells die, this is known as apoptosis, or programmed cell death.

The ability to develop immortalized cells in vitro is a powerful tool for the biological investigator. The immortalization of cells establishes continuous cell lines that play an important role in the study of the biology of cell growth, differentiation, and senescence. Large-scale culture of immortal cell lines enables the production of large quantities of DNA or any cell product for research or for the pharmaceutical industry.

The development of the biotechnology industry and the requirement for eukaryotic cells in the post-translational processing of protein products of cells has placed a great deal of pressure on the scientific and regulatory communities to provide continuous cell lines from specific cell lineages that are safe to use and with phenotypic properties relevant to their tissue of origin.

Table 2 provides the most up-to-date examples of the in vitro methods used in the immortalization of primary cultures of various cell types.

BIOREACTORS AND PRODUCTS

Bioreactors are used to obtain the maximum quantity of protein required. Fermenters are used for cell lines that grow in suspension or for high-density production of cells (e.g., HIV-1 and HIV-2 in human T lymphoblastoid cells). Fixed-bed bioreactors are advantageous because they provide close cell-to-cell contact, which is important for virus transfer (18). Porous glass spheres in fixed-bed systems retain nonadherent cells and provide large quantities of HIV.

For attached cell lines (e.g., 293, human embryo kidney) (29), solid-support bioreactors have been developed either by a stirred-tank reactor using microcarriers (55) or hollow-fiber culture systems (56). For example, the polio vaccine is prepared in Vero cells grown on microcarriers in 1,000-L tanks (16).

Medium replacement is important when growing cells on a solid support or in suspension to sustain growth and protein production. A constant supply of growth factors must be supplied while reducing buildup of inhibitory metabolites such as lactic acid and ammonia (29). Serum-free formulations must be considered when dealing with extracellular protein purification from mammalian cells. Other environmental aspects to be considered for efficient management of bioreactors are low shear forces, high transfer rates of oxygen and carbon dioxide, and a controlled microenvironment.

Examples of new bioreactors are the Technomouse, which combines the principle of nutrient and oxygen perfusion via silicone membranes and organlike cultivation of human cells (long-term primary tissue culture is possible) (56). The VERAX process (57) uses a fluidized bed of macroporous collagen carriers, which is particularly suitable for the production of secreted recombinant products, such as Aujeszky herpesvirus in Vero cells.

Products

Interferon was the first cytokine to be made from cultured animal cells, and it is now prepared routinely at more than 95% purity, showing that scale-up to 10,000 L is possible and economically viable (26). Originally produced from human white blood cells, only small amounts could be obtained. NAMALWA, a transformed B-lymphoblastoid cell line, produces plentiful amounts of interferon and is safe to use on humans. The final product (Wellferon) is of high purity, correctly glycosylated, and contains at least six distinct α-interferons, unlike that produced from *Escherichia coli* by recombinant DNA technology (27).

Fibronectin, an effective factor for the growth of anchorage-dependant cells, is produced by HUH 6 (Table 1), a human hepatoblastoma cell line. At more than 90%

Table 1. Human and Primate Cell Lines

Cell line name	Cell line type	Product	Use	Reference
U 937 THP 1 JM, C 8166 JHAN	Human T-cell lymphoblast	Supports growth of HIV-1 and HIV-2	Vaccine production and research	18–22
RPMI 2650	Human carcinoma	Growth factors autocrine-stimulator, TGF-α, and TFG-β	Regulation of cell growth and productivity in vitro of human epithelial cells	23
HUH clone 5	Human hepatoblastoma	Fibronectin	Attachment factor in serum-free media	24
NAWALWA	Human lymphoblastoid	IPSF, IIα	Improves immuno-globulin productivity from human hybridomas	25
NAMALWA	Human lymphoblast	Interferon α when infected with sendai virus and sodium butyrate	Treatment of hairy cell leukemia potentiates cytotoxic T cell and natural killer cell response	26,27
293	Human embryo kidney	Supports growth of adenovirus vectors that carry reporter genes into neuronal cells Acetylcholin, esterase	Gene therapy in cystic fibrosis Structure and function analysis	28
293 transformed with human parathyroid calcium receptor	Human embryo kidney	Recombinant protein (similar to human plasma factor X)	Structural studies and clinical evaluation and growth of viruses	29–31
HepG 2 PLC/PRF/5 Hep3B	Human hepatocellular carcinomas	Human plasma proteins, e.g., transferrin, plasminogen, α-fetoprotein, albumin (PLC/PRF/5-albumin not produced)	Experimental models for investigation of plasma protein biosynthesis and synthesis of hepatitis B surface antigen HbsAg (not HepG2)	32,33
Vero	African green monkey kidney	Aujeszky virus and oral polio virus	Vaccine production	16

TFG = transforming growth factor.

purity, its biological activity is the same as commercial fibronectin from hamster kidney cells, BHK 21 (24), but only 2 L of HUH-6 will produce up to 27 mg of protein.

SAFETY CONSIDERATIONS

Any cell line used for the propagation of a specific product for the pharmaceutical industry must be vigorously controlled. Fully authentic cell banks provide that safety measure and are stored in liquid nitrogen for indefinite periods. We only need to recall the first cell line to be isolated, HeLa in 1950, deemed to be the single most important tool of biomedical research in the past 50 years. It was sent all over the world, resulting in mass cross-contamination and millions of dollars of medical research being lost.

Guidelines have been set for the acceptability of using human diploid cells for the production of live virus vaccines in humans (58) to ensure that there are no extraneous microbial contaminants or transforming factors (59,60). The increasing use of animal cells for the production of recombinant proteins and vaccines means that standardization in this field is ever more important to ensure the provision of safe and reliable therapeutic diagnostic reagents (61,62).

A further major concern is the long-term risk of malignancy represented by heterogeneous contaminating DNA, especially if it were to contain potentially oncogenic sequences (e.g., hepatitis B vaccine produced by recombinant DNA techniques). The World Health Organization (WHO) has published strict guidelines on the use of cell lines for vaccine production (63).

FUTURE TRENDS

Gene Therapy

Recently, treatment of human genetic disorders has been mediated by gene transfer technology, especially for monogenic disorders such as cystic fibrosis. Research has been carried out into the feasibility of using a defective nonreplicating recombinant adenovirus vector to transfer a reporter gene (e.g., β-galactosidase) into neuronal cells (28). Human embryonal kidney cells 293 were used as a packaging cell and cotransfected with the vector, which could then be inserted into cystic fibrosis patients. Gene therapy replaces the specific gene sequence into a target cell (28).

If this is successful, this may lead to an improvement in controlling monogenic diseases in which only a minority of cells are required for genetic alteration in order to replace a deficient enzyme.

Bioartificial Liver

Liver failure is the major cause of mortality in the United States, with 30,000 deaths per year. There are no artificial devices available, transplants are effective but donors are limited.

Table 2. In Vitro Methods

Cell line type	Method of transformation	Uses	References
Human umbilical vein endothelial cells	Low-level radiation, oncogenes, SV40, spontaneous somatic cell hybridization with a human osteosarcoma	Von Willebrand factor, prostacyclin (e.g., ECV304 and EA.hy926)	35–39
HPEC A1, human placental cells	Plasmic pRNS1 lipofection	Endothelial characteristics after 80 cell divisions	40–43
Human B lymphocytes	Epstein-Barr virus	Express normal B-cell markers, e.g., HLA class I and II antigens, tissue typing, Genetic research	44,45
Human dermal fibroblasts	SV40 T antigen	Oncogenic transformation, aging, senescence	48
Monkey kidney epithelium 47/17/3, 42/17/8, ori-11/7	Genetic engineering, oncogenes, polyoma large T, H-ras	Alternative to primary cell culture	49–51
Human embryo kidney 293	Sheared human adenovirus 5 DNA	Sensitive to human adenoviruses; isolation of transformation defective host range mutants of adenovirus type 5 and titration of adeno type 21	52,53
293N3S	Produced by passaging 293 in nude mice	Large-scale production of viruses in bioreactors	54

A temporary support to support biotransformation and detoxification functions is required to sustain patients waiting for a donor or until the liver has repaired itself.

Development of a hollow fiber reactor using hepatocyte entrapment with rabbit and rat models has been successful, and scale-up is now being developed for human clinical trials (64).

CONCLUSION

The products and cell lines of human and primate origin discussed in this article have contributed to the general acceptance of using animal cells as the substrate for production of therapeutic proteins.

Many proteins are now obtained by expressing the gene concerned in Chinese hamster ovary cells by recombinant DNA technology, but human and primate cell lines are used as well. Many proteins will always be made more simply in bacteria, yeast, or insect cells and in the future by chemical synthesis. But until then, where glycosylation of proteins is necessary and chemical refolding of bacterial products is not always possible for full pharmaceutical activity, mammalian cells will still be used, especially with the constant improvement in protein-free and serum-free formulations and improvement in higher cell densities through bioreactor research.

BIBLIOGRAPHY

1. J.P. Jacobs, C.M. Jones, and J.P. Baille, *Nature* **227**, 168–170 (1970).
2. L. Hayflick, S.A. Plotkin, T.W. Norton, and H. Koprowski, *Am. J. Hyg.* **75**, 240 (1962).
3. D. Ikie, *Prog. Immunobiol.* **2**, 305 (1965).
4. J.P. Jacobs, D.I. Magrath, A.J. Garrett, and G.C. Schild, *J. Biol. Standardisation* **9**, 331–342 (1981).
5. J.B. Griffiths, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, Academic Press, London, 1985, pp. 3–11.
6. A. Doyle, C.B. Morris, and W.J. Armitage, *Cryopreservation of Animal Cells: Upstream Processes, Equipment and Techniques*, Liss, New York, 1988, pp. 1–17.
7. G.N. Stacey, *Folia Microbiol* **42**, 113–116 (1997).
8. A. Doyle, B.J. Bolton, *Animal Cell Technology*, New York, Wiley, pp. 244–271.
9. R.J. Hay, Y.A. Reld, P.R. McClintock, T.R. Chen, and M.L. Macey, *J. Cell. Biochem.* **Suppl. 24**, 107–130 (1996).
10. P. Arturson, *Crit. Rev. Therapeutic Drug Carrier Systems* **8**, 305–330 (1991).
11. M. Pinto, M.D. Appay, and P. Simon-Assman, *Biol. Cell.* **44**, 193–196 (1982).
12. J. Fogh, W.C. Wright, and J.D. Loveless, *J. Natl. Cancer. Inst.* **58**, 209–214 (1977).
13. J.R. Latigo, T.P. Battle, R. Cook, P. Freemantle, and G.N. Stacey, in M.J. Corrondo, B. Griffiths, and L.P. Moreira eds., *From Vaccines to Genetic Medicine*, ESACT, Kluwer, London, 1996, pp. 385–389.
14. P.D. Minor, in R.E. Spiers ed., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, ESACT, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 741–748.
15. K. Shaw and N. Nathanson, *Am. J. Epidemiol.* **103**, 1–12 (1976).
16. B.J. Montagnon, B. Fanget and J.C. Vincent-Falquet, in B. Simizu and T. Terasima eds., *Vero Cells: Origin, Properties and Biochemical Applications*, Chiba Univ., Tokyo, 1988, pp. 181–210.
17. Haynes, *AIDS Res. Hum. Retroviruses* **7**, 17–27 (1991).
18. J.B. Clark and J.B. Griffiths, in R.E. Spiers ed., *Production of Biologicals from Animal Cells in Culture*, ESACT, Butterworth-Heinemann, Oxford, U.K., 1990, pp. 519–522.
19. C. Sundstrom and K. Nilsson, *Int. J. Cancer* **17**, 565–577 (1976).
20. P. Ralph, M. Moore, and K. Nilsson, *J. Exp. Med.* **143**, 1528–1533 (1976).
21. H.S. Koren, S.J. Anderson, and J.W. Larrick, *Nature* **279**, 328–331 (1979).
22. D.G. Fischer, M.C. Pike, H.S. Koren, and R. Snyderman, *J. Immunol.* **125**, 463–465 (1980).

23. M. Clynes and M. Dooley, in R.E. Spiers ed., *Production of Biologicals from Animal Cells in culture*, ESACT, Butterworth-Heinemann, Oxford, U.K., 1990, pp. 519–522.
24. K. Nagamine, M. Shiraishi, Z. Kong, K. Shinohara, and H. Murakami, in R.E. Spiers ed., *Production of Biologicals from Animal Cells in culture*, ESACT, Butterworth-Heinemann, Oxford, U.K., 1990, pp. 167–169.
25. H. Murakami, in R.E. Spiers ed., *Production of Biologicals from Animal Cells in culture*, ESACT, Butterworth-Heinemann, Oxford, U.K., 1990, pp. 735–740.
26. N.B. Finter, in R.E. Spiers ed., *Production of Biologicals from Animal Cells in culture*, ESACT, Butterworth-Heinemann, Oxford, U.K., 1990, pp. 3–12.
27. R.G. Werner, in R.E. Spiers ed., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, ESACT, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 369–375.
28. S.D. Sivasubramaniam and G.N. Stacey, in M.J.T. Corrondo, B. Griffiths, and L.P. Moreira eds., *From Vaccines to Genetic Medicine*, ESACT, Kluwer, London, 1996, pp. 385–389.
29. J. Shiloach, J. Kaufman, L. Trinh, and C. Kemp, in M.J.T. Corrondo, B. Griffiths, and L.P. Moreira eds., *From Vaccines to Genetic Medicine*, ESACT, Kluwer, London, 1996.
30. F.L. Graham, *J. Gen. Virol.* **68**, 937–940 (1987).
31. A.E. Rudolf, M.P. Mullane, R. Porche-Sorbet, and J.P. Miletich, *Protein Express. Purif.* **10**, 373–378 (1997).
32. B.B. Knowles, C.C. Howe, and D.P. Aden, *Science* **209**, 497–499 (1980).
33. D.P. Aden, A. Fogel, I. Damjanor, S. Plotkin, and B.B. Knowles, *Nature* (*London*) **282**, 615 (1979).
34. L. Hayflick and P.S. Moorhead, *Exp. Cell. Res.* **25**, 585 (1961).
35. N.A. Punchard, D. Watson, R. Thompson, and M. Shaw, in R.I. Freshney and M.G. Freshney eds., *Culture of Immortalized Cells*, Wiley, New York, 1996, pp. 203–238.
36. K. Takahasi and Y. Sawasaki, *In Vitro Cell Dev. Biol.* **26**, 265 (1990).
37. K. Takahasi and Y. Sawasaki, *In Vitro Cell Dev. Biol.* **27A**, 766 (1991).
38. K. Takahasi and Y. Sawasaki, *In Vitro Cell Dev. Biol.* **28A**, 380 (1992).
39. O. Hohenarter, C. Schmatz, and H. Katinger, in M.J. Carrondo, J.B. Griffiths, and J.L.P. Moreira eds., *Animal Cell Technology: From Vaccines to Genetic Medicine*, ESACT, Kluwer, The Netherlands, 1997, pp. 9–11.
40. P. Litzkas, K.K. Jha, and H.L. Ozer, *Mol. Cell. Biol.* **4**, 2549–2552 (1984).
41. M. Teifel, L.T. Heine, S. Milbredt, and P. Friedl, *Eur. J. Biochem*. 1996, submitted for publication.
42. M. Schutz and P. Friedl, *Eur. J. Cell. Biol.* 1996, submitted for publication.
43. M. Schultz and P. Friedl, in M.J. Corrando, J.B. Griffiths, and J.L.P. Moreira eds., *Animal Cell Technology: From Vaccines to Genetic Medicine*, ESACT, Kluwer, The Netherlands, 1997, pp. 749–753.
44. A.H. Davies, F.J. Evans, A.B. Rickinson, J. Huddleston, and A.N. Emery, in R.E. Spiers ed., *Production of Biologicals from Animal Cells in Culture*, ESACT, Butterworth-Heinemann, Oxford, U.K., 1990.
45. B.J. Bolton and N.K. Spur, in R.I. Freshney and M.G. Freshney eds., *Culture of Immortalised Cells*, Wiley, New York, 1996, pp. 283–298.
46. C. Macdonald and B. Bolton, in A. Doyle and J.B. Griffiths eds., *Mammalian Cell Culture*, Wiley, W. Sussex, U.K., 1997, pp. 154–167.
47. M.A. Epstein and B.G. Achong, *The Epstein Barr Virus*, William Heinemann, London, 1986, pp. 2–11.
48. L.V. Mayne, T.N.C. Price, K. Moorwood, and J.F. Burke, in R.I. Freshney and M.G. Freshney eds., *Culture of Immortalized Cells*, Wiley, New York, 1996, pp. 77–94.
49. H. Land, L.F. Parada, and R.A. Weinberge, *Nature* **304**, 596–598 (1983).
50. M. Rassoulzadegan, A. Cowrie, A. Carr, F. Cuzin, R. Kamen, and N. Glaichenhaus, *Nature* **300**, 713–716 (1982).
51. J. Clarke, C. Macdonald, U. Kreuzburg-Duffy, and J.B. Griffiths, in R. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, ESACT, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 50–51.
52. F.L. Graham, J. Smiley, W.C. Russell, and R. Naim, *J. Gen. Virol.* **36**, 59–72 (1977).
53. T. Harrison, F. Graham, and J. Williams, *Virology* **77**, 319–329 (1977).
54. F.L. Graham, *J. Gen. Virol.* **68**, 937–940 (1987).
55. S. Reuveny, in A.S. Lubiniecki ed., *Large Scale Mammalian Cell Culture Technology*, Dekker, New York, 1990, pp. 271–341.
56. T. Omasa, *Biotechnol. Bioeng.* **48**, 673–680 (1995).
57. T. Marique, D. Malarme, P. Stragier, and J. Werenne, in R.E. Spiers ed., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, ESCAT, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 273–277.
58. J.P. Jacobs, D.I. Magrath, A.J. Garrett, and G.C. Schild, *J. Biol. Standard.* **9**, 331–342 (1981).
59. R.J. Hay, Y.A. Reid, P.R. McClintock, T.R. Chen, and M.L. Macey, *J. Cell. Biochem.* **Suppl. 24**, 107–130 (1996).
60. R.J. Hay, *Anal. Biochem.* **171**, 225–237 (1988).
61. G.N. Stacey, *Folia. Microbiol.* **42**, 113–116 (1997).
62. A. Doyle and B.J. Bolton, in Wiley, London, 1992, pp. 243–271.
63. J. Petricciani, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 3, 1988, pp. 14–24.
64. M.V. Peshwa, S.L. Nyberg, F.J. Wu, B. Amiot, F.B. Cerra, and W.S. Hu, in R.E. Spiers ed., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, ESACT, Butterworth-Heinmann, Oxford, U.K., 1994, pp. 273–277.
65. A. Nagel, E. Effenberger, S. Koch, L. Lubbe, and U. Marx, in R.E. Spiers ed., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, ESACT, Butterworth-Heinmann, Oxford, U.K., 1994, pp. 296–304.

HYBRIDOMA, ANTIBODY PRODUCTION

FRANCISCO J. CASTILLO
Berlex Biosciences
Richmond, California

KEY WORDS

Antibody
Ascites
Bioreactor
Hybridoma
MAb

Monoclonal
Myeloma

OUTLINE

INTRODUCTION

The creation of hybridomas capable of producing monoclonal antibodies (MAbs) was accomplished by Koller and Milstein in 1975 (1). Two decades later, the number of MAbs used in research, purification, and immunoassay applications is large and has sustained continued growth (2–5). The medical applications for MAbs have not lived up to expectations, lagging behind other biologicals, because only a relatively small number have received approval by the FDA (6–8) (Fig. 1), four for therapeutic and five for in vivo diagnostic use (Table 1). However, this could change rapidly in the near future. In 1996, of the 12 biologics approved by the FDA, four were MAbs (9). Two MAbs were approved in 1997 (10), one of which (Zenapax) was the first approved humanized antibody. Three Biologic License Applications (BLAs) for MAbs were also filed in 1997 (11–13), and one was filed in May 1998 (14), for a total of six under review by FDA at that time. In addition, of the 350 biopharmaceutical products in development, more than 70 are MAbs, and several are in pivotal clinical trials near completion (15–18) (Fig. 2). Thus, the number of fillings of BLAs with the FDA and international regulatory agencies for MAb-based products could see significant growth in the future, to be followed by corresponding approvals.

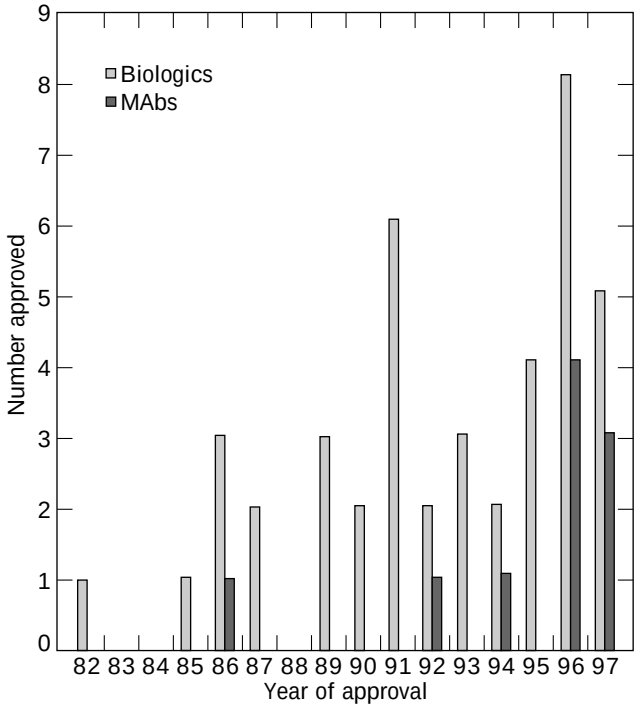


Figure 1. MAbs and other biologics approved by the FDA.

MAbs are routinely produced in vivo and in vitro. For in vivo production, mice (mostly BALB/c) are the preferred species. The abdominal cavity of pristane-primed mice is used to incubate injected hybridoma cells, which grow and secrete MAb that concentrates in the ascitic fluid. The MAb-rich ascites is collected by repeated needle aspiration (tapping) of the animals (19,20).

With the advent of mammalian cell culture scale-up for the production of recombinant proteins during the 1980s (21–26), concomitant with the high expectations regarding the potential uses and demand for MAbs (27–31), the cultivation of hybridoma cells was intensively studied. This lead to the development of a variety of bioreactors, culture media, and culture processes (31–44), many of which, when properly combined and used, easily generate required quantities of high-quality material at relatively low cost.

Genetically engineered MAbs, such as chimeric, humanized, MAb fragments and fusion proteins, are produced in mammalian and insect cells, microorganisms, and transgenic animals and plants (45–57).

Production of human MAbs using human cells still presents important challenges, although progress has been made (58–60). More significant advances have been made with the use of bacteriophage display technique and the development of transgenic mice, which express the human genes for antibody production (61,62). These animals when inoculated with antigens produce only human antibodies. Their esplenocytes can be isolated and fused with murine myelomas to produce stable hybridomas that produce human MAbs, although containing murine glycosylations instead of human. The time and procedures needed to generate these hybridomas are the same as for conventional murine hybridomas.

As clearly seen in Table 1, very different processes for MAb production have been made to work in cost-effective, validated fashion. The choice will depend on the intended application, scale required, personnel experience, and available infrastructure. With the available technology, even with the present limited level of understanding of murine hybridoma physiology, setting up a culture process for MAb production should be a relatively straightforward exercise.

HYBRIDOMA GROWTH AND MAb PRODUCTION

The performance of a process involving cultivation of hybridomas or any other cells will be the result of the com-

Table 1. Monoclonal Antibodies Approved by the FDA (by May 1998)

Product	Company	Indication	Approval date	Production method
Orthoclone OKT®3	Ortho Biotech, (Raritan, NJ)	Reversal of acute kidney transplant rejection	June 1986	Ascites (63)
		Reversal of heart and liver transplant rejection	June 1993	Ascites
OncoScint CR/OV	Cytogen (Princeton, NJ)	Detection, staging, and follow-up of colorectal and ovarian cancers	December 1992	Airlift fermentation
ReoPro	Centocor (Malvern, PA)	Antiplatelet prevention of blood clots in the setting of high-risk PTCA	December 1994	Extended perfusion (spin-filter) (44)
	Eli Lilly (Indianapolis, IN)	Prevention of cardiac ischemic complications in patients undergoing PTCA	November 1997	
CEA-Scan	Immunomedics (Morris Plains, NJ)	Metastatic colorectal cancer imaging	June 1996	Ascites
MyoScint	Centocor (Malvern, PA)	Myocardial necrosis imaging agent	July 1996	Extended perfusion (spin filter)
Verluma	Boehringer lingelheim (Ridgefield, CT) NeoRx (Seattle, WA)	Small cell lung cancer imaging agent	August 1996	
ProstaScint	Cytogen (Princeton, NJ)	Recurrent prostate cancer imaging agent	October 1996	Extended perfusion (hollow fiber) (64)
Rituxan	Idec Pharmaceuticals (San Diego, CA) Genentech, Inc. (S. San Francisco, CA)	Low-grade non-Hodgkin's lymphoma	November 1997	Suspension culture
Zenapax	Hoffman-La Roche, Inc. (Nutley, NJ)	Prevention of acute graft rejection in kidney transplant	December 1997	Suspension culture

Source: Based on Ref. 9.
Note: PTCA = percutaneous transluminal coronary angioplasty

bination of three main factors: the cell line, the culture medium, and the bioreactor conditions.

The Hybridoma Cell

Preparation of Hybridomas. Hybridomas are prepared by fusing splenocytes (antibody-producing spleen B lymphocytes) from antigen-treated animals and myelomas (1). Splenocytes do not propagate indefinitely in vitro but myelomas do, so when both types of cells are fused, the hybridomas formed are able to propagate and produce the antibodies. Many detailed protocols for preparation of hybridomas have been described in the literature (2,65–69).

There are two preferred and widely used murine myelomas that descend from a transplantable murine (BALB/c) plasmacytoma, MOPC 21, that was adapted to in vitro culture and renamed P3K (70). These two myelomas are P3-X63-Ag8.653 (71) and NSO/1 (65), and both have lost the ability to produce the heavy and light chains of the original MOPC 21 antibody (IgG1). A third nonproducing cell, Sp2/0-Ag14 (72), was obtained through several steps (73) and is itself a hybridoma derived from the fusion of myeloma P3-X63-Ag8X and splenocytes from a BALB/c mouse (74). The use of any of these cells guarantees that the antibodies produced by hybridomas derived from them will correspond only to those specified by the genes from the splenocytes used in the fusion. Monoclonality, however, will only be ensured when the proper procedures are used to clone and subclone the hybrids produced (68,75–79). Cloning and subcloning are also required to select stable clones that retain the antibody class and the specific productivity (see "Hybridoma Stability").

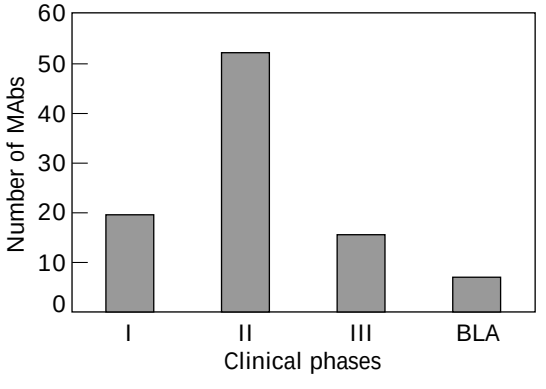


Figure 2. Number of MAbs on test at different stages of clinical trials.

Preparation and Characterization of Hybridoma Banks. After clone selection, it is important to ensure the availability of homogeneous and healthy cell stocks, and this is accomplished by preparing cell banks. For industrial applications, cell banks are prepared at several stages: research, master cell bank (MCB), manufacturer's working cell bank (MWCB), and end-of-production cells (EPC).

MCBs are intended for the preparation of MWCBs, which in turn, are prepared to ensure a supply of homogeneous cell seed for manufacturing purposes and for the life of the product, and EPC banks are for quality control and validation purposes (80,81). The successful preparation and maintenance of cell banks requires very healthy, exponentially growing cells, the use of fresh medium and appropriate quality and quantity of cryopreservant (usually, dimethylsulfoxide [DMSO]) at time of freeze, and storage in monitored liquid nitrogen freezers. Detailed procedures and protocols for preparation, freezing, and storage of hybridoma stock banks are well described in the literature (82–84). Characterization of the banks is required to ensure identity, consistent performance, and lack of contaminating adventitious agents. The characterization includes recoverability of the cells upon thawing, identification of cell species, growth and production kinetics of the cells, and testing for the presence of endogenous and adventitious infectious agents (Table 2).

Hybridoma Stability. Somatic variation (antibody class shift) resulting in structural MAb variants produced by hybridomas have been reported (85–87). This phenomenon occurs as a normal evolutionary event of the humoral immune response, but once stable hybridoma clones are selected its frequency is low. The most frequent instability problem in hybridomas seems to be a decrease in MAb yield that correlates with the emergence of nonproducing cells (88–94). Nonproducing hybridomas arise with the loss of chromosomes 12 for the heavy and 16 and 6 for the light, λ, and κ chains, respectively (95–98). Loss of chromosomes occurs with higher frequency shortly after fusion, so it is very important to clone and subclone to ensure the selection of stable clones. Interspecies hybrids such as murine/human or murine/rabbit have increased chromosome segregation compared to murine/murine, which results in more frequent instability problems (35,99).

For extended cultivation processes, it is important to evaluate the performance of the cells beyond the intended incubation time and to compare the quantity and quality of the MAb produced during the runs and to prepare EPC banks as mentioned above (100).

The Culture Medium

The culture medium has to provide, at a minimum, an energy source and all nutritional requirements of the cells.

Table 2. Characterization of Hybridoma Banks

Test	MCB	MWCB	EPC
Sterility	+	+	+
Mycoplasma	+	+	+
Growth and production kinetics	+	+	+
Authenticity	+	+	+
Species-specific viruses	+	−	−
Retroviruses	+[a]	−	+[a]
Adventitious viruses in vitro assay	+	−	+
Adventitious viruses in vivo assay	+	−	+

Source: Ref. 80 and 81.
[a]Retrovirus testing is not required for murine hybridomas (81).

Energy Sources. Hybridoma cells use glucose and glutamine as the preferred sources of energy. Evidence suggests that glucose and glutamine use by mammalian cells in culture is under reciprocal regulation (101–103). Hybridomas are derived from tumor cells (myelomas) and as such metabolize glucose mainly through glycolysis (104). Many hybridomas generate lactate in quasi-stoichiometric levels compared to the glucose used; the lactate accumulates in the medium, and in later stages of the culture it can be used by the cells (105,106). Glucose utilization rates range from less than 1 to more than 10 pmol/cell/day (43); the sugar is required by hybridomas and growth does not occur in its absence.

Glutamine is also essential for growth and maintenance of cultures, and its utilization rate, which ranges from less than 1 to more than 5 pmol/cell/day (43), depends on the concentration. Glutamine addition has been shown to increase the longevity of cultures, which translates into higher MAb yields (107–109). Recent studies have shown that glucose and glutamine consumption can be redirected to minimize lactate and ammonia production by using a rational medium design and a feeding strategy. Application of such approaches in fed-batch culture has led to very high levels of cells and MAb (110).

Pyruvate is both produced and consumed by hybridomas at different rates (line dependent), but its addition to the medium is not required for growth or MAb production.

Nutritional Requirements

Amino Acids. Mouse cells, like other animal cells, are unable to synthesize certain (essential) amino acids, and these have to be provided in the medium. Nonessential amino acids can be synthesized by the hybridomas, but are frequent components of some media formulations (i.e., Ham's F-12, and others [111]). Addition of these amino acids results in lower energy requirements by the cells because anabolic amino acid metabolism is eliminated. Depending on the cell line, this supplementation may also result in higher specific growth rates.

Lipids. Lipids are essential structural elements of all cell membranes, and one type, the phospholipids, are the most abundant components of membranes (112). Starting from ethanolamine and fatty acids, hybridomas synthesize phosphatidyl ethanolamine (113), from which they can then derive other phospholipids (114,115). Hybridomas respond differently to lipid supplements, and specific effects on cell and MAb yields have been reported (116).

Vitamins and Inorganic Ions. These components are required by hybridomas and mammalian cells in general. Sufficient amounts are included in most media formulations (111), and limitations are not frequently observed under normal non-high-density growth conditions.

Oxygen. Hybridoma specific oxygen consumption rates have been reported to range from 1 to 40 pg/cell/h (117–119). Oxygen is one of the main limiting factors in the maintenance of viable populations in high-density perfusion cultures. In most instances, hybridoma cells will be unaffected by a wide range of oxygen concentrations. However, high oxygen concentrations (i.e., 100% or more of air saturation) can inhibit growth rates and induce cell lysis

in later culture stages, presumably through the production of damaging oxygen radicals. In some hybridomas, the optimal concentrations of dissolved oxygen for growth and MAb production are different (118,120).

Carbon Dioxide. Most hybridomas will initiate growth well in an air atmosphere without additional CO_2. However, CO_2 produced by the cells has a significant effect on the culture pH. Because of accumulation of CO_2 and lactic acid, the pH of batch cultures without base addition may drop to values below 6.5.

Basal Media. Most media used in mammalian cell culture are based on the balanced salt solution. For hybridoma work, Dulbecco's Modified Eagles Medium (DMEM) and a 1:1 mixture of DMEM and Ham's F-12 (111) are commonly used and contain the nutritional requirements discussed. These formulations have been modified often to increase cell and MAb yields.

Supplements. A universal supplement used in cell culture is animal serum. Serum contains nutrients and many factors that directly or indirectly stimulate hybridoma growth.

Basal media supplemented with serum (usually 10% or less fetal bovine serum [FBS]) will allow exuberant growth of murine hybridomas in either stationary (i.e., T flasks) or suspension (shake flasks or spinners), and cultures in simple batch mode without optimization will yield, in most cases, from 1 to 3×10^6 cells/mL and from 10 to over 100 mg/L of MAb. If needed, serum depleted of immunoglobulins can be used to minimize their contamination of the product (121).

Serum can be replaced by combinations of supplements such as insulin, transferrin, ethanolamine, and selenium, which are employed in serum-free formulations (122,123). Proteins and other contaminants are reduced when utilizing serum-free formulations and eliminated with protein-free formulations. Studies on serum-free and protein-free media are abundant in the literature (122–129), and several convenient commercial formulations are available (Table 3). Addition of basic proteins and spermine, a low molecular weight polyamine, to serum-free media has been found to stimulate IgM production by a human hybridoma line (130).

Most murine hybridomas adapt readily to one or more of these media, with growth kinetics, cell, and product yields often as good or better than those in serum-containing media. In addition, when done properly, the cells respond well to freezing in serum-free and protein-free media (123,128,130), as shown in Table 4. Additional advantages in using serum-free and protein-free media include the reduced risk for introducing adventitious agents of animal origin, such as retroviruses and prions, which is a significant concern if the MAb product is intended for therapeutic applications (131–133).

Table 3. Commercially Available Serum-Free and Protein-Free Media for Hybridoma Growth and Production of MAbs

Media	Protein content (mg/L)	Availability as liquid (L) or powder (P)	Suppliers
UltraDOMA	30	L	Bio-Whittaker[a]
UltraDOMA-PF	0	L, P	Bio-Whittaker
Hybridoma-SFM	20	L	Gibco BRL[b]
PFHM-II	0	L, P	Gibco BRL
CCM1	2100	L, P	HyClone[c]
SFX mab	low	L	HyClone
PFmab	0	L	HyClone
HB 101	780	L, P	Irvine Scientific[d]
HB 102	780	L, P	Irvine Scientific
HB 104	780	L, P	Irvine Scientific
HB PRO	1	L, P	Irvine Scientific
HB GRO	50	L, P	Irvine Scientific
Ex-Cell 610	11	L, P	JRH Biosciences[e]
Ex-Cell 620	11	L, P	JRH Biosciences
QBSF-51	45	L	Sigma[f]
QBSF-52	45	L	Sigma
QBSF-53	408	L	Sigma
QBSF-55	65	L	Sigma
QBSF-56	428	L	Sigma
Hybri-Max	0	L	Sigma

[a]Bio-Whittaker, Inc., 8830 Biggs Ford Rd., Walkersville, MD, 21793; U.S. tel. 800-638-8147/301-898-7025; fax 301-845-8338; international fax 301-845-8291.
[b]Gibco BRL, Life Technologies, Inc., P.O. Box 9418, Gaithersburgh, MD, 20898; U.S. tel. 800-828-6686; fax 800-331-2286; international fax 301-258-8238.
[c]HyClone Laboratories, Inc., 1725 South HyClone Rd., Logan, UT 84321-6212, U.S. tel. 800-492-5663; fax 800-533-9450; international tel. 801-753-4584; International Fax 801-753-4589.
[d]Irvine Scientific, 2511 Daimler Street, Santa Clara, CA 92705-5588; U.S. tel. 800-577-6097/800-437-5706; fax 714-261-6522.
[e]JRH Biosciences, 13804 W. 107th Street, Lenexa, KS 66215-0848, U.S. tel. 800-255-6032; international tel. 913-469-5580; fax 913-469-5584.
[f]Sigma Chemical Company, P.O. Box 14508, St Louis, MO 63178-9916; U.S. tel. 800-325-3010/800-848-7791; fax. 800-325-5052; international tel. 314-771-5750; international fax. 314-771-5757.

Bioreactors

A large variety of in vitro systems are used to grow hybridomas, including multiwell plates, Petri dishes, T flasks, multiplate containers, hollow fiber cartridges, roller bottles, shake flasks, spinner flasks, airlift fermentors, and stirred-tank fermentors (30–44). The bioreactors provide the vessel and the physicochemical environment, which includes, among the most important factors, pH, temperature, agitation, gas exchange, redox potential, and osmolarity.

Effect of pH. Hybridomas can tolerate pH extremes (i.e., 6.4 to 8) if adapted gradually. Uptake and utilization of nutrients are affected by pH; for instance, glucose and glutamine utilization rates increase when the pH is alkaline (Fig. 3). Cultures evolve toward acidity because of the production of lactic acid and CO_2. To control the pH, sodium bicarbonate is frequently used as a buffering agent and is initially added to media in concentrations of grams per liter. Organic buffers such as 2-[4-(2-hydroxyethyl)-1-piperazinyl] ethanesulfonic acid (HEPES), 2-morpholinoethanesulfonic acid (MES), 3-morpholino-propanesulfonic acid (MOPS) or others (134) are also used at concentrations of 10 mM or higher.

Table 4. Growth and MAb Production by Three Hybridomas Cultivated in 2-L Spinner Flasks Using Serum-Free and Protein-Free Media

	Line A		Line B		Line C	
Media	SF[a]	PF	SF	PF	SF	PF
Maximum viable cell concentration ($\times 10^6$/mL)	1.2	1.5	1.9	1.9	1.6	1.5
Specific growth rate (h^{-1})	0.038	0.022	0.04	0.043	0.041	0.033
Doubling time (h)	18.3	31.5	17.3	16.1	16.9	21
Maximum MAb concentration (mg/L)	74	108	62	76	90	106

Source: F.J. Castillo and J. Thrift, unpublished results.

Note: All hybridomas were derived from BALB/c splenocytes. Line A was derived from fusion with P3-X63Ag8.653 myeloma (71) and produces IgM. Lines B and C were derived from fusion with P3NS1/1-Ag4-1 myeloma (65) and produce IgG1 and IgG2a, respectively.

[a]Data correspond to averages obtained in 8 and 6 consecutive passages, respectively in serum-free (SF) medium containing bovine transferrin and insulin and protein-free (PF) medium containing chelated iron to replace transferrin. The adapted cells had been frozen in corresponding media formulations supplemented only with 10% DMSO as cryoprotectant, and were grown as described upon thawing.

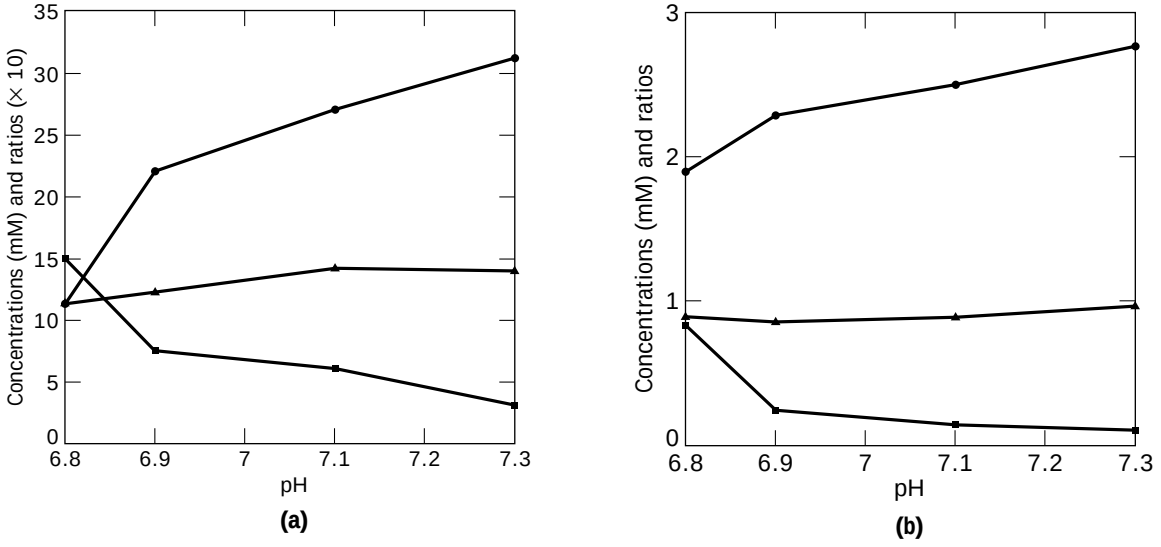


Figure 3. Effect of pH on the utilization of glucose and glutamine and the production of lactate and ammonia by line XMMLY-B. The hybridoma was grown in chemostat culture at a dilution rate of 0.017 h^{-1}. Data were collected at each pH after reaching steady states. (**a**) Glucose concentration (■), lactate concentration (●), ratio of lactate produced to glucose utilized (▲). (**b**) Glutamine concentration (■), ammonia concentration (●), ratio of ammonia produced to glutamine utilized (▲).

Maximum specific growth rates are usually observed at pH values between 6.8 and 7.3, but growth and MAb production are frequently optimal at different pH values. A production process may consist of a growth phase in which the pH is controlled at a certain value, followed by a production phase where the pH is adjusted to the value that maximizes product secretion.

The quality of the secreted MAb can be affected by the pH in the medium either by facilitating chemical changes or, when low, by providing a favorable environment for the action of released acidic proteases (106,135–137). Thus, the effect of pH on growth, productivity, MAb yields, and MAb quality should be determined to maximize yields and product consistency.

Effect of Temperature. Optimal temperature for myeloma and hybridoma growth is 37 °C, and the cells recover from excursions to slightly higher temperatures. Lower temperatures cause decreases in growth and MAb production (138), but the cells are not affected and recover promptly when placed again at 37 °C. Myelomas and hybridomas can be maintained at room temperature for several days, which is convenient for shipments of cells in culture. For permanent storage a liquid nitrogen freezer is essential. A properly frozen hybridoma stock, upon thawing and placed in culture with fresh medium, will usually resume growth within 48 hours.

Effect of Agitation and Gas Bubbles. Concerns about shear sensitivity of mammalian cells prompted the development of airlift fermenters for use with mammalian cells (139) and hybridoma growth (140,141) that are used for commercial production of MAbs (Table 1). However, experience has shown that each line is different (142), shear sensitivity depends on culture conditions (143–145), most hybridomas are not especially sensitive to shear stress in

culture, and some can stand very harsh agitation. Bubble-cell interactions cause damage, but shear-protective agents such as the polyol Pluronic F-68, which is frequently added to cell cultures, can minimize the damaging effects of bubble disengagement (144–145).

Redox Potential. Although the oxidation–reduction potential (redox potential) is routinely measured on-line in microbial fermentations, the same is not frequently done in cell cultures. Little understanding of correlations between the redox potential on growth or MAb production exist (146), but optimal values may be defined and applied in the future.

Effect of Osmolarity. Hybridomas, like other mammalian cells, grow uninhibited in osmolalities ranging from 260 to 300 mOsmol/kg H_2O (approximate osmolarities of 260 to 300 mmol/L [147]) and commercial basal media are designed to have these values by adjusting the NaCl content. High osmolarity has been reported to inhibit growth rate and cell yields (148), but also to induce higher MAb secretion at later stages of cultivation (149–153). Some amino acids and analogs act as osmoprotective agents for hybridomas and can even increase specific antibody productivity (154).

Growth Kinetics

Requirements for hybridoma growth are a viable inoculum and proper extracellular environment, including essential nutrients and suitable physicochemical conditions.

When all these conditions are met, hybridomas, like all other eukaryotic cells in culture, have kinetics of growth similar to microbial cells and also obey the exponential growth law (155,156). Hybridomas are not anchorage dependent and grow as well in nonagitated cultures, such as multiwell plates, Petri dishes, T flasks and hollow fiber bioreactors as they do in agitated containers, including roller bottles, shake flasks, spinner flasks, or fermenters. Hybridoma cells grow monodispersed, and in batch cultures, in most media, they reach concentrations of several million per milliliter, with viabilities of well over 90% throughout the growth phase.

When growth data are plotted as viable cell counts against time, a typical batch growth curve of hybridoma may contain the six phases observed in microbial culture: lag, accelerating growth, exponential growth, decelerating growth, stationary, and decline (156). A lag period is generally observed when starting cultures using cells thawed from frozen stocks. Upon thawing, the cells go through an initial period of adaptation, but once exponential growth is attained no lag period will occur upon subculturing provided that the cells are passaged to fresh medium while still growing exponentially. If a culture is maintained beyond the exponential phase, a lag often is observed upon passage, and such cells are said to have been overgrown. The longer cells are maintained in the spent medium beyond the exponential phase, the more likely that a lag period will occur in the next culture (157). Ultimately, there will be a point of overgrowth past which all cells will die either by apoptosis or by necrosis (158), induced by starvation or toxic metabolite buildup (159–164), and the culture can no longer be recovered.

During exponential growth, doubling times for most hybridomas are between 14 and 23 h, corresponding to maximum specific growth rates ranging from 0.05 to 0.03 h^{-1}.

Hybridomas have been grown in chemostat cultures at dilution rates ranging from critical (Dc), which is equivalent to maximum specific growth rate (156), to less than 10% of Dc. Viability decreases with dilution rate (Fig. 4**a**), and a common observation with all hybridomas is that an intrinsic minimum specific growth rate of approximately 0.02 h^{-1} exists for all of them (164,165) (Fig. 4**b**). A minimum specific growth rate was also reported for other cells and attributed to nutritional limitations (166). Death of hybridomas at low dilution rates may be mostly by apoptosis induced by nutritional limitation(s) and reduced protein synthesis. A practical implication of this phenomenon in a continuous culture process is that the fraction of dead cells will increase with decreases in the dilution rate (purge rate in high-density perfusions [156]). The ratio of viable/dead cells will have effects on MAb concentrations and homogeneity and on the longevity of production runs, all of which impact product yields.

MAb Production Kinetics

MAbs are produced by viable cells (167). Patterns observed for MAb production by hybridomas are the same as those seen with microbial cells for product formation (168), and most can be defined either as growth-associated, partly growth-associated, non-growth-associated, or negatively growth-associated (169–178).

Figure 5 shows the growth and MAb production of two hybridomas in batch cultures. One clearly has growth-associated production (**a**), and the second (**b**) has a non-growth-associated pattern with more that 90% of the MAb produced after growth has ceased.

Specific productivities of several hybridomas at different dilution rates in continuous chemostat cultures are shown in Figure 6. The pattern of line XMMEN-A was non-growth-associated, whereas that of all other hybridomas tested in the group was growth-associated. The specific pattern is very important in defining the productivity profile of continuous cultures, and depending on the system and its mode of operation, the bleed or harvest rate will have to be adjusted for maximum productivity (156). In high-density cultures, productivity can also be affected by other factors, external or autocrine (179), which may not be obvious or predictable in low-density cultures and that need to be evaluated under operation conditions.

The validity and usefulness of applying the microbial growth and non-growth-associated concepts to the production of proteins by animal cells, including hybridomas, has been questioned (180). However, it is clear that for two cell lines with different productivity profiles (as shown in Fig. 5), the optimal harvest times will be very different, and this will have a significant impact on the corresponding operations.

MAb Production Processes

In Vitro. MAb yields by most hybridomas in simple batch cultures without optimization are in the 10 to 100

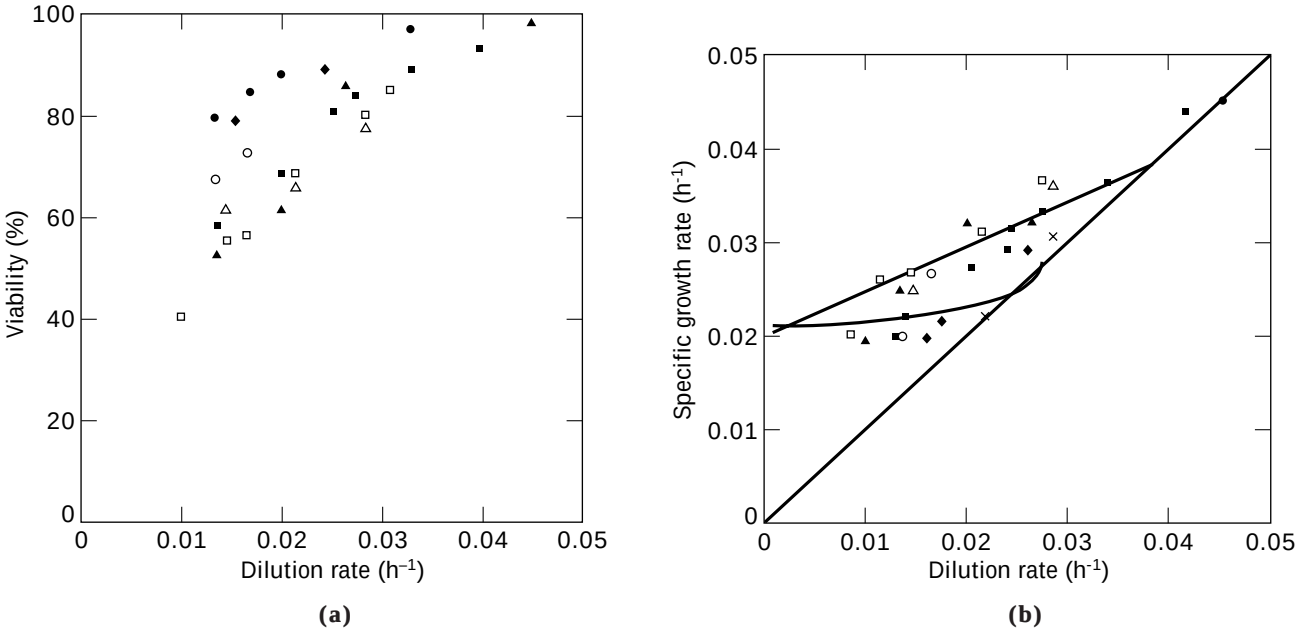


Figure 4. Changes in viability (**a**) and specific growth rates (**b**) with dilution rate in chemostat culture of several hybridomas. Cell lines (Table 5): XMMEN-A (○), XMMME-C (●), XMMLY-B (△), B-17 (▲), XMMCO-E (□), XMMLY-F (■), XMMME-D (x), XMMID-G (◆).

mg/L range. Production of milligram quantities can be easily accomplished using T flasks, spinner flasks, shaker flasks, or roller bottles. Gram quantities require either more of these containers and more labor or the use of larger vessels such as fermenters. Productive batch cultures need a minimum of conditions, which include an exponentially growing inoculum, a buffered medium with an initial pH adjusted between 7.2 and 7.5, and the temperature of incubation regulated at 37 ± 0.5 °C. Growth kinetics, specific MAb productivities, and other characteristics of several hybridomas grown with these conditions in batch suspension cultures and in different vessels are complied in Table 5.

Fed-batch culture of hybridoma growth continues to evolve, and very high MAb yields, in the order of grams per liter, have been reported recently (110,181–184). This may become the process of choice in future commercial operations.

High-density continuous processes can be done at many scales with available preassembled equipment. The systems vary in their complexity, instrumentation, and requirements for infrastructural support. Some are easy to set up and operate, but others require special training and experience. All demand attention in order to be used routinely in successful ways, but they are reliable and productive when understood and operated properly (39–44).

In Vivo. Production in ascites has the advantage of high concentration, and it is widely used. It is commonly used in the production of small quantities of MAb for research and diagnostic purposes; two of the approved MAbs for medical use are made using ascites technology (Table 1). Contract suppliers can provide ascites at different scales from a few milligrams to kilograms per year. Facilities designed for production of up to kilogram quantities of MAbs under cGMP are operated for contract manufacturing by companies such as Charles River in Massachusetts and North Carolina and Genzyme Transgenics. Although clearly not the preferred method of production for therapeutic MAbs, as explicitly stated by the Committee for Proprietary Medicinal Products (185), validated and cost-effective ascites processes exist, and several MAbs in advanced clinical trials are made in ascites.

Emerging in vivo processes involve the use of transgenic animals. Reports of high yields in the milk of goats and other animals coupled with effective separation and purification techniques may make this a competitive technology for certain applications (186). However, for regulatory and strategic reasons, it is advisable to adopt in vitro processes for production of pharmaceutical-grade MAbs. The use of animals for production of MAbs for therapeutic applications will have to deal with, among other complications, the growing concerns of potential bovine spongiform encephalopathy or other prion contaminations (133, 187,188).

Production of antibodies in plants has also been done (50) and may represent a viable alternative for some applications.

MAb Recovery and Purification

MAbs are secreted by hybridomas, and when using protein-free media, the cultures contain high MAb/total protein ratios. In high-density, hollow-fiber perfusion cultures, MAb levels corresponding to more than 40% of the total protein have been reported (43). For some applications, a cell-free harvest of this relative composition may

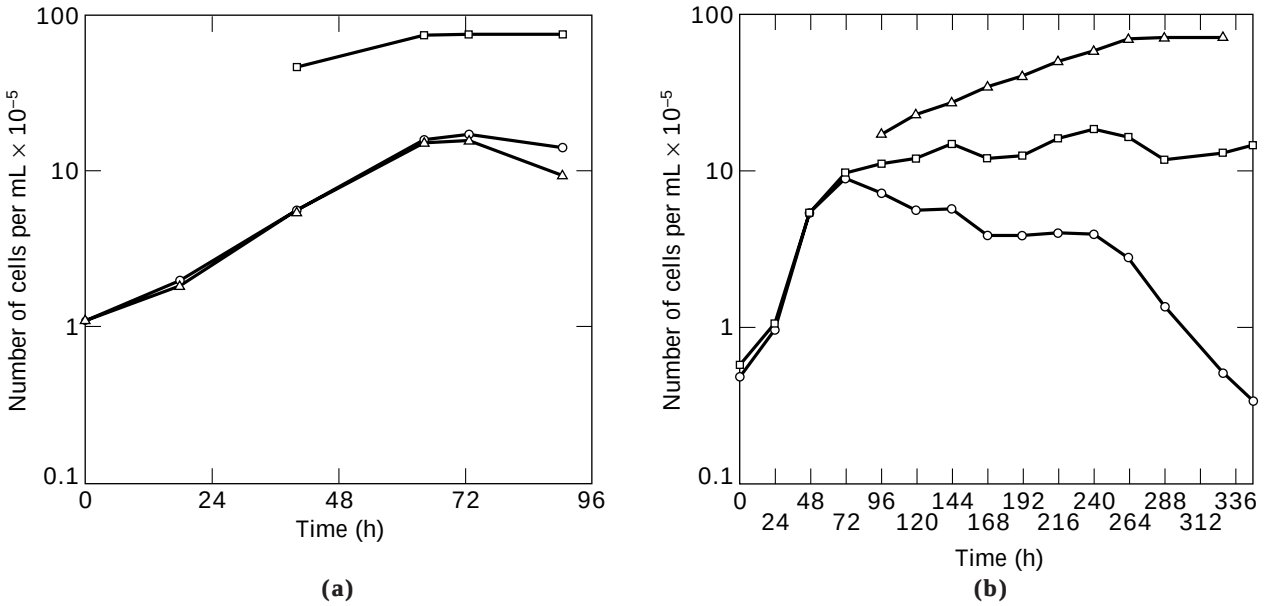


Figure 5. Growth curves and MAb production by two hybridoma lines (**a**, XMMME-C, and **b**, XMMEN-A) with growth-associated and non-growth-associated productivity patterns. Total (○), viable (△), and MAb concentration per mg/L (□).

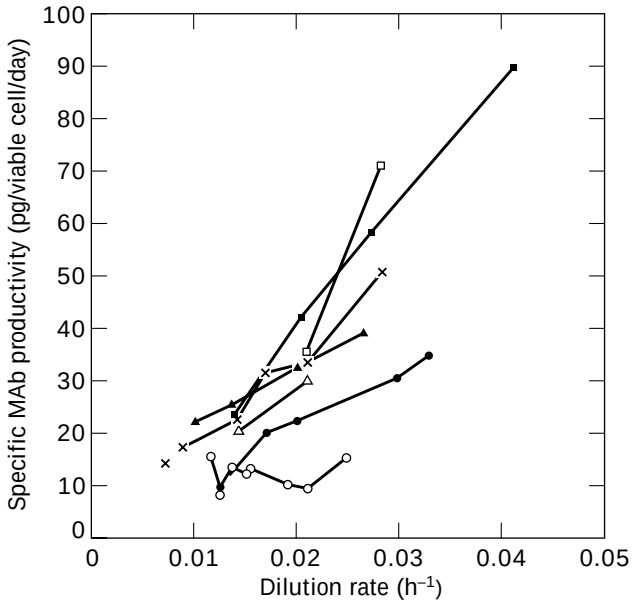


Figure 6. Growth-associated and non-growth-associated productivity patterns of several hybridoma cells grown in continuous chemostat cultures. Symbols same as in Figure 4.

be used as such without further processing to remove other proteins present (although contaminating cellular proteases may degrade the MAb upon prolonged incubation).

MAb recovery traditionally starts with cell removal by centrifugation or filtration. New processes aim at obviating separation steps by applying the entire culture to a fluidized bed of chromatography beads (189). Purification of MAbs involves usually two to three steps, which may include combinations of precipitation or column processes such as ion-exchange separation; affinity separations (protein A or dye ligand); or hydroxylapatite, hydrophobic, mixmode, and gel filtration (190–192)

PROCESS CHANGES

Process development is a multidisciplinary endeavor aimed at maximizing yields and minimizing costs while maintaining product consistency. Traditionally, there has been reluctance in industry to publish results on the development of processes for biopharmaceuticals, including those for MAbs (34). This reluctance reflects the need to maintain advantage over competitors.

Significant advances have taken place in the understanding of hybridoma physiology and in the development of highly productive MAb processes. Through the years, a wealth of information from academia and industry became available in publications and through other sources. The development and implementation of an MAb production process benefits today from this information, and systematic approaches to follow have been described (34,193).

Many approved biologicals are manufactured by obsolete or suboptimal processes. For instance, OKT3, the first therapeutic MAb approved, is made with ascites technology. It was introduced into the market 12 years ago and continues to be manufactured by the same process. There are several reasons for this apparent conservatism in the biopharmaceutical industry.

First, there is the pressure to expedite manufacturing for introduction of a biological product into the market. A recent example of this pressure was that confronted by Chiron Corporation for the production of Betaseron (recombinant interferon-β). After approval of the product in

Table 5. Batch Culture of Hybridoma Lines in Suspension

Line	Parental myeloma[a]	MAb type	Bioreactor	Maximum specific growth rate (h^{-1})	Doubling time (h)	Specific MAb productivity[b] (pg/viable cell/day)		
						A	B	C
XMMEN-A	P3-X63Ag8.653	IgM	2-L spinner	0.05	14	15	13	N/D
XMMEN-A	P3-X63Ag8.653	IgM	40-L stirred fermenter	0.05	14	20	15	19
XMMEN-A	P3-X63Ag8.653	IgM	40-L airlift fermenter	0.018	37	24	15	18
A-72	P3-X63Ag8.653	IgM	2-L spinner	0.05	14	15	14	15
A-78	P3-X63Ag8.653	IgM	2-L spinner	0.05	14	13	14	13
XMMLY-B	P3NSI/1-Ag4-1	IgG1	2-L spinner	0.043	16	52	12	4
B-17	P3NSI/1-Ag4-1	IgG1	2-L spinner	0.033	21	76	10	N/D
B-17	P3NSI/1-Ag4-1	IgG1	40-L stirred fermenter	0.023	29	>80	20	22
XMMME-C	P3-X63/Ag8	IgG2a	2-L spinner	0.043	16	40	2	N/D
XMMME-D	P3NSI/1-Ag4-1	IgG1	2-L spinner	0.038	18	90	5	N/D
XMMCO-E	P3NSI/1-Ag4-1	IgG2b	2-L spinner	0.035	20	90	5	N/D
XMMLY-F	P3NSI/1-Ag4-1	IgG2a	2-L spinner	0.046	15	70–90	50–75	40
XMMID-G	Sp2/0	IgG2a	2-L spinner	0.035	20	100-12	10	15
XMMID-G	Sp2/0	IgG2a	40-L stirred fermenter	0.036	19	100-15	10	10

Note: N/D = no data.
[a]All splenocytes used in fusions were of BALB/c origin.
[b]A = during exponential growth; B = at peak viable cell density; C = during stationary phase.

July 1993 for the treatment of relapsing-remitting multiple sclerosis (194), the company had an initial shortage of manufacturing capacity. It was forced to establish a computerized lottery system to supply between 12,000 and 20,000 patients the first year until capacity was brought up in line with the demand (195). Obviously, with such pressing needs, there is little incentive to introduce process changes that may produce further delays.

A second and very important cause has been the difficulty of introducing process changes after pivotal phase III clinical testing or after commercialization of a product. Regulatory approval of process changes could only be obtained after the results of extensive and costly testing in toxicology and efficacy provided guarantees that the changes were not altering significant properties of the product (196). Thus, in several instances, introduction of process changes has either been abandoned or postponed as long as the difficulties and costs implied could not be offset by the potential gains.

More recently, the established dogma in the production of biopharmaceuticals that "the process defines the product" has lost its absolutism. This has been an evolutionary process guided by caution. New powerful techniques are being used for characterization of biologics, and experience gained with the clinical use of these products has helped redefine the relevance of microheterogeneity and the potential problems with contaminants such as residual DNA, proteins, and endogenous retroviruses (197–198). As a result, regulatory agencies have gradually adopted a less-stringent view on the inalterability of the process and the dedication of facilities and have shifted toward a more flexible position based on the ability to demonstrate equivalence of the products made by different processes and at different sites. New regulations and guidelines encourage process improvements when capable of producing equivalent product (199–201). However, it should be kept in mind that culture conditions may significantly affect the properties of MAbs (202,203), and efforts should be made to identify these effects as early as possible in order to guide the selection of the most advantageous technology for a process (204).

CONCLUSIONS

With the present accessibility to relevant, practical information, availability of hybridoma-specific culture media, alternative bioreactor technology, and multiple purification methods, in most cases a simple MAb production process can be easily implemented.

BIBLIOGRAPHY

1. G. Kohler and C. Milstein, *Nature* **256**, 495–497 (1975).
2. E. Harlow and D. Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York, 1988.
3. O.W. Merten, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 4, Academic Press, London, 1990, pp. 257–315.
4. M.J. Perry, in J.R. Birch and E.S. Lennox eds., *Monoclonal Antibodies: Principles and Applications*, Wiley-Liss, New York, 1995, pp. 107–120.
5. C.R. Hill, in J.R. Birch and E.S. Lennox eds., *Monoclonal Antibodies: Principles and Applications*, Wiley-Liss, New York, 1995, pp. 121–136.
6. T.A. Waldmann, *Science* **252**, 1657–1662 (1991).
7. K.J. Morrow, *Gen. Eng. News.* **15**, 1–24 (1995).
8. D.A. Scheinberg and P.B. Chapman, in J.R. Birch and E.S. Lennox, *Monoclonal Antibodies: Principles and Applications*, Wiley-Liss, New York, 1995, pp. 45–105.
9. *Approved Biotechnology Drugs and Vaccines*, Pharmaceutical Research and Manufacturers of America, 1998.
10. A.F. Holmer, *New Drug Approvals in 1997*, doc. #3020, Pharmaceutical Research and Manufacturers of America, 1998.

11. *Neoprobe plans response to FDA requests for additional information under RIGscan ® CR49 BLA marketing application*, 1997, *http: / /www.neoprobe.com*
12. D. Strickland, *BioWorld Today* **8**, 1–2 (1997).
13. R. Osborne, *BioWorld Today* **8**, 1–2 (1997).
14. R. Osborne, *BioWorld Today* **9**, 2 (1998).
15. A.F. Holmer, *350 Biotechnology Medicines in Development*, doc. #3004, Pharmaceutical Research and Manufacturers of America, 1998.
16. K.J. Morrow, *Gen. Eng. News.* **17**, 1–21 (1997).
17. L. Seachrist, *BioWorld Today*, **8**, 1–3 (1997).
18. A. DePalma, *Gen. Eng. News.* **18**, 18–30 (1998).
19. J.P. Chandler, in S.S. Seaver ed., *Commercial Production of Monoclonal Antibodies: A Guide for Scale-Up*, Dekker, New York, 1987, pp. 75–92.
20. N.J. Hoogenraad and C.J. Wraight, *Methods Enzymol.* **121**, 375–381 (1986).
21. B.K. Lydersen ed. *Large Scale Cell Culture Technology*, Hanser Publishers, Munich, 1987.
22. A. Prokop and M.Z. Rosenberg, in A. Fiechter ed., *Advances in Biochemical Engineering*, vol. 39, Springer-Verlag, Berlin, 1989, pp. 29–71.
23. B. Griffiths, in R.I. Freshney ed., *Animal Cell Culture: A Practical Approach*, IRL Press, Oxford, U.K., 1989, pp. 33–69.
24. A.S. Lubiniecki ed. *Large Scale Mammalian Cell Culture Technology*, Dekker, New York, 1990.
25. R. Spier, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 5, Academic Press, London, 1992, pp. 1–46.
26. J.R. Birch and S.J. Froud, *Biologicals* **22**, 127–133 (1994).
27. D.E. Yelton, D.H. Margulies, B. Diamond, and M.D. Scharff, in R.H. Kennett, T.J. McKearn, and K.B. Bechtol eds., *Monoclonal Antibodies Hybridomas: A New Dimension in Biological Analyses*, Plenum Press, New York, 1980, pp. 3–17.
28. J.A. Roth ed. *Monoclonal Antibodies in Cancer: Advances in Diagnosis and Treatment*, Futura, Mount Kisco, N.Y., 1986.
29. A. Mizrahi ed. *Monoclonal Antibodies: Production and Application. Advances in Biotechnological Processes*, vol 11, Liss, New York, 1989.
30. *Proc. Int. Symp. on Monoclonal Antibodies for Therapy, Prevention and In Vivo Diagnosis of Human Disease, Utrech. Dev. Biol. Standard.*, vol. 71, 1989.
31. K. James, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 4, Academic Press, London, 1990, pp. 205–253.
32. S.S. Seaver ed., *Commercial Production of Monoclonal Antibodies: A Guide for Scale-Up*, Dekker, New York, 1987.
33. A. Mizrrahi ed. *Upstream Processes: Equipments and Techniques, Advances in Biotechnological Processes*, vol 7, Liss, New York, 1988.
34. C. Harbour, J.P. Barford, and K.S. Low, in A. Flechter ed., *Advances in Biochemical Engineering / Biotechnology*, vol. 37, Springer-Verlag, Berlin, 1988, pp. 1–40.
35. O.W. Merten, in A.O.A. Miller ed., *Advanced Research on Animal Cell Technology*, Kluwer, Dordrecht, The Netherlands, 1989, pp. 367–398.
36. C.P. Marquis, C. Harbour, J.P. Barford, and K.S. Low, *Cytotechnology* **4**, 69–76 (1990).
37. J.R. Birch, K. Lambert, R. Boraston, P.W. Thompson, and M.A. Boss, *Biotech'85* 1–3 (1985).
38. D.H. Randerson, *J. Biotechnol.* **2**, 241–255 (1985).
39. K. Nilsson, W. Scheirer, H.W.D. Katinger, and K. Mosbach, *Methods Enzymol.* **121**, 352–360 (1986).
40. S.F. de St. Groth, *Methods Enzymol.* **121**, 360–375 (1986).
41. J. Shevitz, S. Reuveny, T.L. LaPorte, and G.H. Cho, in A. Mizrahi ed., *Monoclonal Antibodies: Production and Applications*, Liss, New York, 1989, pp. 81–105.
42. S. Reuveny and A. Lazar, in A. Mizrahi ed., *Monoclonal Antibodies: Production and Applications*, Liss, New York, 1989, pp. 45–80.
43. F.J. Castillo, L.J. Mullen, J.C. Thrift, and B.C. Grant, *Ann. N.Y. Acad. Sci.* **665**, 72–80 (1992).
44. Y.M. Deo, M.D. Mahadevan, and R. Fuchs. *Biotechnol. Prog.* **12**, 57–64 (1996).
45. V.T. Oi and S.L. Morrison, *Biotechniques* **4**, 214–221 (1986).
46. M. Better, P.C. Chang, R.R. Robinson, and A.H. Horwitz, *Science* **240**, 1041–1043 (1988).
47. M. Better and A.H. Horwitz, *Methods Enzymol.* **178**, 476–496 (1989).
48. M. Sollazzo, R. Billetta, and M. Zanetti, in J.M. Cruse and R.E. Lewis eds., *The Year in Immunology 1989–90. Molecules and Cells of Immunity*, Karger, Basel, 1990, pp. 162–184.
49. A. Hiatt, R. Cafferkey, and K. Bowdish, *Nature* **342**, 76–78 (1989).
50. J. Hodgson, *Biotechnology* **9**, 421–425 (1991).
51. A. Wright, S.U. Shin, and S.L. Morrison, *Crit. Rev. Immunol.* **12**, 125–168 (1992).
52. J.J. Trill, A.R. Shatzmann, and S. Ganguly, *Curr. Opinion Biotechnol.* **6**, 553–560 (1995).
53. J.R. Birch and E.S. Lennox, *Monoclonal Antibodies, Principles and Applications*, Wiley-Liss, New York, 1995.
54. R. Irving and P. Hudson, in H. Zola ed., *Monoclonal Antibodies: The Second Generation*, Bios Scientific Publishers, Oxford, U.K., 1995, pp. 119–139.
55. D.M. Fishwild, S.L. O'Donnell, T. Bengoechea, D.V. Hudson, F. Harding, S.L. Bernhard, D. Jones, R.M. Kay, K.M. Higgins, S.R. Schramm, and N. Lonberg, *Nat. Biotechnol.* **14**, 845–851 (1996).
56. S. Paul, *Antibody Engineering Protocols. Methods in Molecular Biology*, vol. 51, Humana Press, Totowa, N.J., 1995.
57. M.W. Young, W.B. Okita, E.M. Brown, and J.M. Curling, *BioPharm* **10**, 34–38 (1997).
58. G. Winter and C. Milstein, *Nature* **349**, 293–299 (1991).
59. D.M. Knight, M. McDonough, M.A. Moore, D. Abercrombie, R. Siegel, and J. Ghrayeb, *Hum. Antibod. Hybridomas* **3**, 129–136 (1992).
60. T. Lindl, *Cytotechnology* **21**, 183–193 (1996).
61. M. Neuberger and M. Bruggemann, *Nature* **386**, 25–26 (1997).
62. R. Sherman-Gold, *Genet. Eng. News* **17**, 4–35 (1997).
63. W.W. Lime, in A.S. Lubiniecki and S.A. Vargo eds., *Regulatory Practice for Biopharmaceutical Production*, Wiley-Liss, New York, 1994, pp. 235–253.
64. S.L. Danheiser, *Gen. Eng. News* **16**, 1–36 (1996).
65. G. Galfre and C. Milstein, *Methods in Enzymol.* **73**, 3–46 (1981).
66. J.J. Langone and H.V. Vunakis, *Methods Enzymol.* **92**, (1983).
67. A. Osterhaus and F. Uytdehaag, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 2, Academic Press, London, 1985, pp. 49–69.

68. J.W. Goding, *Monoclonal Antibodies: Principles and Practices*, Academic Press, London, 1986.
69. W.C. Davis, *Monoclonal Antibody Protocols. Methods in Molecular Biology*, vol. 45, Humana Press, Totowa, N.J., 1995.
70. K. Horibata and A.W. Harris, *Exp. Cell Res.* **60**, 61–77 (1970).
71. J.F. Kearney, A. Radbruch, B. Liesegang, and K. Rajewsky. *J. Immunol.* **123**, 1548–1550 (1979).
72. M. Shulman, C.D. Wilde, and G. Kohler, *Nature* **276**, 269–270 (1978).
73. G. Kholer and C. Milstein, *Eur. J. Immunol.* **6**, 511–519 (1976).
74. G. Kholer and C. Milstein, *Nature* **256**, 495–497 (1975).
75. J.M. Davis, *Methods Enzymol.* **121**, 307–322 (1986).
76. C.I. Civin and M.L. Banquerigo, *Methods Enzymol.* **121**, 322–327 (1986).
77. M.B. Rittenberg, A. Buenafe, and M. Brown, *Methods Enzymol.* **121**, 327–331 (1986).
78. A. Boeye, *Methods Enzymol.* **121**, 332–340 (1986).
79. X. Li, K. Abdi, and S.J. Mentzer, *Hybridoma* **11**, 645–652 (1992).
80. Points to Consider in the Characterization of Cell Lines Used to Produce Biologicals, FDA, Washington, D.C., 1993.
81. Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products for Human Use, FDA, Washington, D.C., 1997.
82. F.J. Castillo, in J.A. Asenjo and J.C. Merchuk, *Bioreactor System Design*, Dekker, New York, 1995, pp. 13–45.
83. C.B. Morris, in J.G. Day and M.R. McLellan eds., *Methods in Molecular Biology*, vol. 38, Humana Press, Totowa, N.J., 1995, pp. 179–187.
84. R.J. Hay, in J.C. Hunter-Cevera and A. Belt eds., *Maintaining Cultures for Biotechnology and Industry*, Academic Press, San Diego, Calif., 1996, pp. 161–178.
85. D.E. Yelton, W.D. Cook, and M.D. Scharff, *Transplant. Proc.*, **12**, 439–443 (1980).
86. P. Thammana and M.D. Scharff, *Eur. J. Immunol.* **13**, 614–619 (1983).
87. G. Spira, H.L. Aguila, and M.D. Scharff, *J. Immunol.* **140**, 2675–2680 (1988).
88. F. Melchers, M. Potter, and N.L. Warner, in F. Melchers, M. Potter, and N. Warner eds., *Lymphocyte Hybridomas: Current Topics in Microbiology and Immunology*, vol. 81, Springer-Verlag, Berlin, 1978, pp. IX–XXIII.
89. G.L. Altshuler, R. Dilwith, J. Sowek, and G. Belfort, *Biotechnol. Bioeng. Symp. No. 17*, 1986, pp. 725–736.
90. K.K. Frame and W-S. Hu, *Biotechnol. Bioeng.* **35**, 496–476 (1990).
91. S.S. Ozturk and B.O. Palsson, *Hybridoma* **9**, 167–175 (1990).
92. G.M. Lee, A. Varma, B.O. Palson, *Biotechnol. Prog.* **7**, 72–75 (1991).
93. A-P. Zeng, *Biotechnol. Bioeng.* **50**, 238–247 (1996).
94. R. Hubbard, in A. Wiseman ed., *Topics in Enzyme and Fermentation Biotechnology 7*, Ellis Horwood, New York, 1983, pp. 196–263.
95. P. D'Eustachio, D. Pravtcheva, K. Marcu, and F.H. Ruddle, *J. Exp. Med.* **151**, 1545–1560 (1980).
96. D. Swan, P. D'Eustachio, L. Leinwand, J. Seidman, D. Keithley, and F.H. Ruddle, *Proc. Natl. Acad. Sci. USA* **76**, 2735–2739 (1979).
97. T. Meo, H.J. Johnson, V. Beechey, J. Andrews, J. Peters, and A.G. Searle, *Proc. Nat. Acad. Sci. USA* **77**, 550–553 (1980).
98. P. D'Eustachio, A.L. Bothwell, T.K. Takaro, D. Baltimore, and F.H. Ruddle, *J. Exp. Med.* **153**, 793–800 (1981).
99. R. Westerwoudt, in J.J. *Methods Enzymol.* **121**, 3–18 (1986).
100. F.J. Castillo, L.J. Mullen, B.C. Grant, J. DeLeon, J.C. Thrift, L.W. Chang, J.M. Irving, and D.J. Burke, *Dev. Biol. Stand.* **83**, 55–64 (1994).
101. L. Reitzer, B. Wice, and D. Kennell, *J. Biol. Chem.* **254**, 2669–2676 (1979).
102. H. Zielke, C. Zielke, and P. Ozand, *Fed. Proc.* **43**, 121–125 (1984).
103. D. Barngrover, J. Thomas, and M. Thilly, *J. Cell. Sci.* **78**, 173–189 (1985).
104. A. Fiechter and F.K. Gmunder, in A. Fiechter ed., *Advances in Biochemical Engineering/Biotechnology*, vol. 39, Springer-Verlag, Berlin, 1989, pp. 1–28.
105. K. Low and C. Harbour, *Dev. Biol. Stand.* **60**, 73–79 (1985).
106. D. Velez, S. Reuveny, L. Miller, and J. Macmillan, *J. Immunol. Methods* **86**, 45–52 (1986).
107. O.T. Ramirez and R. Mutharasan, *Biotechnol. Bioeng.* **36**, 839–848 (1990).
108. Y.J. Schenider and A. Lavoix, *J. Immunol. Methods* **129**, 251–268 (1990).
109. M.C. Flickinger, N.K. Goebel, T. Bibile, and S. Boyce-Jacino, *J. Biotechnol.* **22**, 201–226 (1992).
110. L. Xie and D.I. Wang, *Biotechnol. Bioeng.* **51**, 725–729 (1996).
111. R.G. Ham and W.L. McKeehan, *Methods Enzymol.* **58**, 44–93 (1979).
112. G. van Meer, *Ann. Rev. Cell. Biol.* **5**, 247–275 (1989).
113. H. Murakami, H. Masui, G.H. Sato, N. Sueoka, T.P. Chow and T. Kanosueoka, *Proc. Nat. Acad. Sci. USA* **79**, 1158–1162 (1982).
114. J. Darnell, H. Lodish, and D. Baltimore, *Molecular Cell Biology*, Scientific American Books, New York, 1986.
115. B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J.D. Watson, *Molecular Biology of the Cell*, Garland Publishing, New York, 1994.
116. S. Savonniere, N. Zeghari, L. Miccoli, S. Muller, M. Maugras, and M. Donner, *J. Biotechnol.* **48**, 161–173 (1996).
117. R. Boraston, P.W. Thompson, S. Graland, and J.R. Birch, *Dev. Biol. Stand.* **55**, 103–111 (1984).
118. W.M. Miller, C.R. Wilke, and H.W. Blanch, *J. Cell. Physiol.* **132**, 524–530 (1987).
119. D. Wolpart, D. Kirwan, and J. Gainer, *Biotechnol. Bioeng.* **36**, 630–635 (1990).
120. H.A. Phillips, J.M. Scharer, N.C. Bols, and M. Moo-Young, *Biotechnol. Lett.* **9**, 745–750 (1987).
121. P.A. Underwood, *Methods Enzymol.* **121**, 301–306 (1986).
122. H. Murakami, in M. Umeda, H. Koyama, J. Minowada, and M. Oishi, eds., *Biotechnology of Mammalian Cells*, Japan Scientific Societies Press/Springer-Verlag, Berlin, 1987, pp. 119–136.
123. H. Murakami, in A. Mizrahi ed., *Monoclonal Antibodies: Production and Applications*, Liss, New York, 1989, pp. 107–141.
124. F.J. Darfler, *In Vitro Cell. Dev. Biol.* **26**, 769–778 (1990).
125. E. Mariani, A.R. Mariani, M.C.G. Monaco, E. Lalli, M. Vitale, and A. Facchini, *J. Immunol. Methods* **145**, 175–183 (1991).
126. H.R. Maurer, *Cytotechnology* **5**, 1 (1991).
127. R.J. Ertola, A.M. Giulietti, and F.J. Castillo, in J.A. Asenjo and J.C. Merchuck eds., *Bioreactor System Design*, Dekker, New York, 1995, pp. 89–137.

128. M.J. Keen and T.W. Steward, *Cytotechnology* **17**, 203–211 (1995).
129. Y.M. Qi, P.F. Greenfield, and S. Reid, *Cytotechnology* **21**, 95–109 (1996).
130. Y. Miyazaki, S. Nishimoto, T. Sasaki, and T. Sugahara, *Cytotechnology* **26**, 111–118 (1998).
131. *Proc. Int. Meeting on Transmissible Spongiform Encephalopathies-Impact on Animal and Human Health, Heidelberg, Dev. Biol. Standard* vol. 80, 1993.
132. D.M. Ascher, *Transmissible Spongiform Encephalopathies: Towards Risk Assessment and Management*, FDA Update, CBER, FDA, Washington, D.C., 1997.
133. German Federal Ministry of Health, *Guidelines on Safety Measures in Connection with Medicinal Products Containing Body Materials Obtained from Cattle, Sheep, or Goats for Minimizing the Risk of Transmission of BSE and Scrapie*, German Federal Bulletin no. 40, Feb. 26, 1994.
134. K. Nagira, M. Hayashida, M. Shiga, K. Sasamoto, K. Kina, K. Osada, T. Sugahara, and H. Murakami, *Cytotechnology* **17**, 117–125 (1995).
135. D.W. Karl, M. Donovan, and M.C. Flickinger, *Cytotechnology* **3**, 157–169 (1990).
136. R.B. Kratje, W. Lind, and R. Wagner, *J. Biotechnol.* **32**, 107–125 (1994).
137. G.K. Sureshkumar and R. Mutharasan, *Biotechnol. Bioeng.* **37**, 292–295 (1991).
138. H.W.D. Katinger, W. Sheirer, and E. Kromer, *Ger. Chem. Eng.* **2**, 31–38 (1979).
139. J.R. Birch, P.W. Thompson, K. Lambert, and R. Boraston, ACS Meeting, Philadelphia, 1984.
140. J.R. Birch, R. Boraston, and L. Wood, *Trends Biotechnol.* **3**, 162–166 (1985).
141. J.F. Petersen, L.V. McIntire, and E.T. Papoutsakis, *Biotechnol. Prog.* **6**, 114–120 (1990).
142. A. Handa-Corrigan, in R.E. Spier and J.B. Griffiths eds. *Animal Cell Biotechnology*, vol. 4, Academic Press, London, 1990, pp. 123–132.
143. R.E. Spier and J.P. Whiteside, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 4, Academic Press, London, 1990, pp. 134–148.
144. I. Abu-Reesh and F. Kargi, *Enzyme Microb. Technol.* **13**, 913–919 (1991).
145. J.D. Michaels, A.K. Mallik, and E.T. Papoutsakis, *Biotechnol. Bioeng.* **51**, 399–409 (1996).
146. S.B. Pluschkell and M.C. Flickinger, *Cytotechnology* **19**, 11–26 (1996).
147. I. Reich, R. Schnaare, and E.T. Sugita, in J.E. Hoover ed., *Remington: The Science and Practice of Pharmacy*, Philadelphia College of Pharmacy and Science, Philadelphia, 1995, pp. 613–627.
148. K. Oyaas, T.M. Berg, O. Bakke, and D.W. Levine, in R.E. Spier, J.B. Griffiths, J. Stephanne, and P.J. Crooy eds., *Advances in Animal Cell Biology and Technology for Bioprocesses*, Butterworth, London, 1989, pp. 212–220.
149. S.K.W. Oh, P. Vig, F. Chua, W.K. Teo, M.G.S. Yap, *Biotechnol. Bioeng.* **42**, 601–610 (1993).
150. S.Y. Park and G.M. Lee, *Biotechnol. Bioeng.* **48**, 699–705 (1995).
151. X. Yang, G.W. Oehlert, and M.C. Flickinger, *Biotechnol. Bioeng.* **50**, 184–196 (1996).
152. S.K.W. Oh, F.K.F. Chua, and A.B.H. Choo, *Biotechnol. Bioeng.* **46**, 525–535 (1995).
153. J.S. Ryu and G.M. Lee, *Biotechnol. Bioeng.* **55**, 565–570 (1997).
154. K. Oyaas, T.E. Ellingsen, N. Dyrset, and D.W. Levine, *Cytotechnology* **17**, 143–152 (1995).
155. J.R. Birch and S.J. Pirt, *J. Cell Sci.* **7**, 661–670 (1970).
156. S.J. Pirt, *Principles of Microbe and Cell Cultivation*, Blackwell Scientific Publications, Oxford, U.K., 1975.
157. A. Martial, M. Dardenne, J.M. Engasser, and A. Marc, *Cytotechnology* **5**, 165–171 (1991).
158. G. Majno and I. Joris, *Am. J. Pathol.* **146**, 3–15 (1995).
159. R.P. Sing, M. Al-Rubeal, C.D. Gregory, and A.N. Emery, *Biotechnol. Bioeng.* **44**, 720–726 (1994).
160. F. Franek, *Biotechnol. Bioeng.* **45**, 86–90 (1995).
161. S. Mercille and B. Masiie, *Biotechnol. Bioeng.* **44**, 1140–1154 (1994).
162. F. Franek and K. Chladkova-Sramkova, *Cytotechnology* **18**, 113–117 (1995).
163. F. Franek and K. Sramkova, *Cytotechnology* **21**, 81–89 (1996).
164. D.E. Martens, E.M. Sipkeme, C.D. de Goojier, E.C. Beuvery, and J. Tramper, *Biotechnol. Bioeng.* **48**, 49–65 (1995).
165. R. Boraston, P. Thompson, K.S. Garland, and J. Birch, *Dev. Biol. Stand.* **55**, 103–111 (1984).
166. M.G. Tovey and D. Brouty-Boye, *Exp. Cell. Res.* **118**, 383–388 (1979).
167. J.M. Savinell, G.M. Lee, and B.O. Palsson, *Bioprocess Eng.* **4**, 231–234 (1989).
168. D.I.C. Wang, C.L. Cooney, A.L. Demain, P. Dunnill, A.E. Humphrey, and M.D. Lilly, *Fermentation and Enzyme Technology*, Wiley, New York, 1979.
169. O.W. Merten, S. Reiter, G. Himmler, W. Scheirer, and K. Katinger, *Dev. Biol. Stand.* **60**, 219–227 (1985).
170. O.W. Merten, G.E. Palfi, G. Klement, and F. Steindl, in R.E. Spier and J.B. Griffiths eds., *Modern Approaches to Animal Cell Technology*, Butterworth, Somerset, U.K., 1987, pp. 381–396.
171. O.W. Merten, *Cytotechnology* **1**, 113–121 (1988).
172. E. Suzuki and D.F. Ollis, *Biotechnol. Prog.* **6**, 231–236 (1990).
173. M.C. Flickinger, N.K. Goebel, and M.A. Bohn, *Bioprocess Eng.* **5**, 155–164 (1990).
174. O.W. Merten, H. Keller, L. Cabanie, M. Leno, and M. Hardefelt, *Cytotechnology* **4**, 77–89 (1990).
175. P.M. Hayter, N.F. Kirkby, and R.E. Spier, *Enzyme Microb. Technol.* **14**, 454–461 (1992).
176. O.W. Merten, D. Moeurs, H. Keller, M. Leno, G.E. Palfi, L. Cabanie, and E. Couve, *Biotechnol. Bioeng.* **44**, 753–764 (1994).
177. M. Harigae, M. Matsumura, and H. Kataoka, *J. Biotechnol.* **34**, 227–235 (1994).
178. G.G. Banik and C.A. Heath, *Biotechnol. Bioeng.* **48**, 289–300 (1995).
179. W.R. Kidwell, *Bio / Technology*, **7**, 462–463 (1989).
180. J.P. Barford, P.J. Phillips, C.P. Marquis, and C. Harbour, *Cytotechnology* **21**, 133–148 (1996).
181. T.A. Bibila and D.K. Robinson, *Biotechnol. Prog.* **11**, 1–13 (1995).
182. H.P.J. Bonarius, V. Hatzimanikatis, K.P.H. Meesters, C.D. de Gooijer, G. Schmid, and J. Tramper, *Biotechnol. Bioeng.* **50**, 299–318 (1996).
183. G.S. Oh, T. Izuishi, T. Inoue, W.S. Hu, and T. Yoshida, *J. Ferment. Bioeng.* **81**, 329–336 (1996).

184. R.K. Biener, W. Waldraff, W. Noe, J. Haas, M. Horwaldt, and E.D. Gilles, *Ann. N.Y. Acad. Sci.* **782**, 272–285 (1996).
185. Commission of the European Communities, *Tibtech* **6**, G5–G8 (1988).
186. *Production of Recombinant Antibodies in the Milk of Transgenic Animals*, Genzyme Transgenics Corp., Framingham, Mass., 1996.
187. J. Hodgson, *Bio / Technology* **8**, 990 (1990).
188. R.H. Kimberlin, *Dev. Biol. Stand.* **75**, 75–82 (1991).
189. J. Thommes, M. Halfar, S. Lenz, and A.R. Kula, *Biotechnol. Bioeng.* **45**, 205–211 (1995).
190. T.C. Ransohoff and H.L. Levine, in R. Seetharam and S.K. Sharma eds., *Purification and Analysis of Recombinant Proteins*, Dekker, New York, 1991, pp. 213–235.
191. P. Gagnon, *Purification Tools for Monoclonal Antibodies*, Validated Biosystems, Tucson, Ariz., 1996.
192. P. Gagnon, *Gen. Eng. News.* **17**, 6–33 (1997).
193. H. Murakami, *Cytotechnology* **3**, 3–8 (1990).
194. *Gen. Eng. News.* **13**, 1–25 (1993).
195. *Wall Street Journal*, **11**, 1, 8 (Sept. 2, 1993).
196. C.N. Durfor and C.L. Scribner, *Ann. N.Y. Acad. Sci.* **665**, 356–363 (1992).
197. D.E. Wierenga, J. Cogan, and J.C. Petricciani, *Biologicals* **23**, 221–224 (1995).
198. J.C. Petricciani and F.N. Horaud, *Biologicals* **23**, 233–238 (1995).
199. G. Schaffner, M. Haase, and S. Giess, *Biologicals*, **23**, 253–259 (1995).
200. *FDA Guidance Concerning Demonstration of Comparability of Human Biological Products, Including Therapeutic Biotechnology-derived Products*, FDA, Washington, D.C., April 1996.
201. M.A. Schenerman and K. Phillips, *BioPharmaceuticals* **10**, 20–72 (1997).
202. B.J. Moellering, J.L. Tedesco, R.R. Townsend, M.R. Hardy, R.W. Scott, and C.P. Prior, *BioPharmaceuticals* **3**, 30–38 (1990).
203. B.L. Maiorella, J. Winkelhake, J. Young, B. Moyer, R. Bauer, M. Hora, J. Andys, J. Thomson, T. Patel, and R. Parekh, *Bio / Technology* **11**, 387–391 (1993).
204. L.C. Chen, *Gen. Eng. News* **17**, 1–10 (1997).

See also AFFINITY FUSIONS, GENE EXPRESSION ; AMMONIA TOXICITY, ANIMAL CELLS; ANTIBODY PURIFICATION; CULTURE MEDIA, ANIMAL CELL, LARGE SCALE PRODUCTION.

HYPOXIA, EFFECTS ON ANIMAL CELLS

DIANE L. HEVEHAN
WILLIAM M. MILLER
Northwestern University
Evanston, Illinois

KEY WORDS

Antioxidants
Gene expression
Metabolism
O_2-sensing
pO_2
Reactive oxygen species
Stress proteins

OUTLINE

INTRODUCTION

Oxygen (O_2) plays a primary role in cell metabolism and viability. By serving as the terminal electron acceptor in oxidative phosphorylation, O_2 is important for energy production. Oxygenation is a major consideration in the design of bioreactors for cultivating cells. Although mammalian cells do not consume O_2 rapidly (the specific O_2 uptake rate, q_{O_2}, is typically $0.05–0.5 \times 10^{-9}$ mmol cell^{-1} h^{-1}) (1,2), O_2 can be a limiting factor in part because of its poor solubility in aqueous media. Oxygen demand by cells is regulated by environmental conditions such as O_2 tension (pO_2), pH, cell type, cell density, and nutrient depletion, all of which may vary throughout the culture period. To accommodate increasing O_2 demands as cultures increase in cell density and proliferative state, a number of bioreactor configurations have been developed (3): agitation, use of microcarriers for adherent cells, and sparging can effectively enhance the driving force for O_2 transfer.

Although much attention has been focused on overcoming extracellular resistances to O_2 transfer, intracellular events further complicate the development of optimal oxygenation strategies. During aerobic metabolism, the reduction of molecular O_2 to water may yield three potentially damaging reactive oxygen species (ROS): the superoxide anion (O_2^-), the hydroxyl radical (OH·), and hydrogen peroxide (H_2O_2). ROS are also produced independently of respiration by O_2-consuming reactions in the cytosol (4) as well as in the culture medium. ROS can damage lipids, proteins, sugars, and DNA. Such damage, collectively known as oxidative stress, can lead to damage of the cell membrane and induction of cell death if it is not counteracted by radical-scavenging defense systems. This dual nature of O_2 should be considered when developing mam-

malian cell culture strategies, and provides an impetus to use controlled hypoxic conditions to optimize reactor productivity (5). The extent of hypoxia can be characterized by the extent to which q_{O_2} is decreased relative to that for cells without O_2 limitation. Moderate hypoxia not only will decrease q_{O_2} but may also lower ROS concentrations and increase metabolic efficiency (6). Insights into the physiological responses of cells to hypoxia will facilitate the use of pO_2 as a manipulatable parameter during cell culture, as well as provide increased understanding of the consequences of an interruption in the O_2 supply.

It would not be possible for a cell to respond to O_2 deprivation and O_2 toxicity without an inherent ability to sense alterations in pO_2. Indeed, cells have evolved complex regulatory mechanisms to ensure metabolic efficiency and protection against oxidative damage. As the O_2 supply changes, cells in vivo and in vitro adapt through multiple physiological responses, ranging from altered energy metabolism and production of antioxidants to changes in growth and differentiation. O_2 deprivation elicits homeostatic responses on the systemic, local, and intracellular levels. Adaptation at the systemic level involves a systemwide increase in O_2 delivery to the tissues through regulation of breathing, red blood cell mass (erythropoiesis), and vasodilation. Locally, the tissue response to hypoxia entails proliferation of capillaries (angiogenesis). Adaptation to O_2 deprivation at the cellular level involves a shift from oxidative phosphorylation and gluconeogenesis to anaerobic glycolysis, as well as the expression of stress-specific proteins. Whether at the organismal or cellular level, many of the effects noted here are mediated through a cascade of events that are initiated by the sensing of pO_2 and propagated by a signal which ultimately leads to hypoxic regulation of specific gene expression. The proteins coded for by these genes facilitate cellular survival at low pO_2 and act to restore normal pO_2.

This article focuses on the current understanding of O_2-dependent responses of animal cells, particularly those exposed to hypoxic conditions. The effects of pO_2 on cells in a wide range of culture systems are addressed first. Subsequently, the molecular mechanisms by which hypoxia mediates its effects—O_2-sensing, signal transduction, and gene expression—are examined. Mechanisms used by cells in vivo have provided the groundwork for uncovering many of the mechanistic details that underlie the hypoxic responses of cultured cells. Although different cells exhibit similar coping mechanisms during periods of O_2 deprivation, their specific adaptation to hypoxia is individually tailored according to cell type and function, as well as the extent of hypoxia.

CELL RESPONSES TO pO_2 AND HYPOXIA

Upon first encountering O_2 deprivation, cells respond rapidly by altering several properties, such as reducing proliferation, shifting to anaerobic glycolysis, and inhibiting protein synthesis. Survival under sustained hypoxia requires a second phase of adaptive responses, which are characterized by extensive cellular and metabolic reprogramming. The selective expression of stress proteins and other specific proteins underlies these secondary processes. Upregulated expression of genes involved in glycolysis and downregulated expression of those involved in gluconeogenesis or the TCA cycle is a prime example of this adaptive mechanism. Immediate and adaptive responses to changes in pO_2 are discussed in the following sections.

Growth, Viability, and Differentiation

The effects of dissolved oxygen concentration (DO) on cell growth in batch culture have been evaluated by many investigators (7–12). The DO value for maximal cell concentration in batch culture is typically about 50% of air saturation (50% DO) but may vary considerably with cell type. Lower values of DO, 5–35% (13) and 0.5% (6), have been reported for maximal steady-state viable cell concentrations in continuous hybridoma culture. Besides differences between cell lines, poor batch growth at low DO could result from added stress from low initial cell concentrations, glucose depletion, or insufficient time for the cells to adapt to low DO.

Steady-state cell viability has been shown to increase as the DO is decreased to 0.1–0.25% DO in continuous culture (6,13). Enhanced viability at lower DO may reflect less oxidative stress imposed on the cells. The greater production of ROS associated with progressively greater DO could damage cellular components and induce cell death. Mammalian cells are able to tolerate the presence of ROS because of the action of antioxidants. However, damage ensues when antioxidant defenses are overwhelmed by ROS. Indeed, in a murine hybridoma culture, increased DNA fragmentation correlated with increasingly hyperoxic culture conditions from 200% to 476% DO (14). Although maximum cell growth in batch culture usually occurs below 100% DO, some cell lines have been adapted to grow at near-maximal rates at DO values as high as 200% DO (15). CHO cells have been adapted to 380% DO (16) and, more recently, a murine hybridoma was adapted for growth at 150% DO by gradually exposing the cells to higher DO levels in continuous culture (17). In both cases, resistance to O_2 toxicity has been attributed to increased activity of antioxidant enzymes during the adaptation process.

An obvious way to avoid oxidative stress is to reduce the amount of O_2 available to the cells, as was demonstrated for hybridomas (6). Low pO_2 stimulates the proliferation of other cell types, particularly fibroblasts seeded at low density (18–20). The proliferation rate and lifetime (i.e., number of population doublings before senescence) of WI-38 and IMR-90 lung embryo fibroblast cells were increased by reducing the gas-phase O_2 concentration from 20% to 2% (19). Based on these findings, Falanga and Kirsner (21) set out to establish fibroblast cultures from single cells, which normally fail to proliferate under 20% O_2. Low pO_2 stimulated and maintained single-cell proliferation, which is consistent with fibroblast proliferation during hypoxic situations in vivo (e.g., during wound healing).

Increased clonal growth of cells under O_2 concentrations less than that of the ambient atmosphere suggest that pO_2 may play a role in directing developmental processes. In other words, pO_2 in tissues may regulate not only cell growth and death but also cell differentiation. This hy-

pothesis and the putative pO_2 gradients that exist within the bone marrow microenvironment provide the rationale for investigating the effects of pO_2 in hematopoietic cell culture. Earlier work showed that a decrease in the gas-phase O_2 tension from 20% to 5% O_2, which better approximates the in vivo environment, enhances the size and number of hematopoietic colonies originating from precursor cells in semisolid culture (22–25). The documented effects of pO_2 in liquid hematopoietic cultures have been controversial, with optimal production reported under 20% O_2 for murine bone marrow cultures (26) and 5% O_2 for cord blood mononuclear cell cultures (27,28). For cord blood mononuclear cell cultures, Smith and Broxmeyer (29) reported comparable cell numbers for mature cells of various lineages under 5% and 20% O_2. However, for cultures of $CD34^+$ peripheral blood cells, Laluppa (30) obtained greater production of megakaryocytes and erythroid cells under 20% O_2, while granulocyte production was greater under 5% O_2. In contrast, most reports are consistent in that low pO_2 is superior for the production and maintenance of progenitor cells of all lineages (27–31). As a result, progenitor cell and mature cell production may be regulated by pO_2 in an inverse manner (30,31). Thus, while low pO_2 may be favorable during the early part of cell culture, higher pO_2 later in culture may enhance the differentiation of certain lineages.

Oxygen also plays a role in the development of other tissues. For example, placental epithelial stem cells tend to proliferate when cultured under 5% O_2 and to differentiate under 20% (32). The use of three-dimensional tissue culture models has facilitated the characterization of embryoid bodies (EBs) obtained from embryonic stem cells. EBs were found to grow equally well and remain highly viable when cultured under 20% or 1% O_2 for the first 5 days of EB differentiation (33). In contrast, embryonic stem cell plating efficiency and differentiation into the hematopoietic lineage are enhanced under 5–7% O_2 (34,35).

Antioxidants and ROS

Mammalian cells have evolved a line of defense to limit the damage caused by ROS generated during aerobic metabolism. This antioxidant defense system includes glutathione, vitamin E, and enzymes such as superoxide dismutase (SOD), catalase, glutathione-*S*-transferase (GST), and glutathione peroxidase (GPX). In addition, extracellular antioxidants such as albumin, urate, and billirubin also protect cells against oxidative damage.

High pO_2 in culture may increase the production of ROS to levels that overwhelm inherent antioxidant defenses. The associated detrimental effects can be overcome by stepwise adaptation to increases in pO_2. Adaptation of hybridoma cells to 100% DO from 10% was associated with an increase in the specific activities of the antioxidant enzymes GPX, GST, and SOD (17). Similarly, adaptation of CHO cells to growth under 99% O_2 was associated with increased specific activities of SOD, catalase, and GPX (16). Thus, enhanced tolerance likely is caused in part by upregulation of the synthesis of these enzymes. For example, a *cis*-acting regulatory sequence that is responsive to ROS, the antioxidant responsive element (ARE), has been found to activate the genes encoding the enzymes GST and NAD(P)H:quinone reductase (36). However, the activity of some antioxidant enzymes may decrease again above 100% DO (17), possibly through inactivation by ROS.

Decreased colony formation at higher pO_2 in semisolid culture suggests that hematopoietic cells are susceptible to oxidative stress. Indeed, addition of antioxidants such as monothioglycerol and β-mercaptoethanol to high-pO_2 cultures yields colonies similar in size and number to those in low-pO_2 cultures (22,23,37,38). H_2O_2 and superoxide concentrations have been found to increase throughout the duration of bone marrow cell cultures (39). Progenitor cell production was increased by adding catalase, an H_2O_2 scavenger, or mannitol, an hydroxyl radical scavenger. In contrast, even though superoxide levels were decreased, the effects of SOD addition were negligible.

Rather than adding antioxidants or upregulating their synthesis through adaptation, an alternate strategy to reduce ROS damage is to decrease ROS generation by operating at low pO_2. Although low pO_2 can be beneficial in promoting proliferation, as already stated, and suppressing apoptosis (see following), it may actually sensitize cells to injury during reoxygenation. Involvement of ROS in reoxygenation injury has been demonstrated by decreased damage in tissues treated with antioxidants during reoxygenation (40). ROS involvement is also shown by increased glutathione levels during reoxygenation of CHO cells from severe hypoxia or anoxia (41). Glutathione levels increase more extensively after more severe hypoxic exposure and increase to a much greater extent after an abrupt increase in pO_2 than during gradual reoxygenation. Sensitization to oxidative damage could not be attributed to a decrease in glutathione levels during hypoxia. In fact, there is evidence to contradict a contribution of decreased antioxidant levels to sensitization to oxidative stress. Some antioxidant enzymes, including SOD, GST, and catalase, are part of a family of stress-related proteins that may be induced during hypoxia (42–44). Upregulated expression of these enzymes during O_2 deprivation may anticipate their role of detoxifying ROS produced during reoxygenation (45).

It has been suggested that intracellular ROS are mediators of apoptosis and may even trigger cell death if present at elevated levels (46). In support of this hypothesis, the features of cells undergoing apoptosis closely resemble those of cells stressed by oxidative damage in terms of their morphology, physiology, biochemistry, and gene expression (47). Intracellular ROS may induce cell death in several ways, from direct damage of DNA to protein oxidation. If oxidants induce apoptosis, a natural extension would be that antioxidants inhibit apoptosis. Indeed, antioxidants such as *N*-acetylcysteine (48), *N*-acetylcarnitine (49), idebenone (50), and trolox (analog of vitamin E) (51) have been shown to inhibit apoptosis. Antioxidant protection against cell death is not confined only to direct oxidant-induced apoptosis but also to receptor- and growth factor-mediated forms of apoptosis. This suggests that oxidation may be a common mechanism in the progression of many different forms of apoptosis. Studies correlating oxidative damage with the induction of apoptosis have led to the

identification of another potential antioxidant, the Bcl-2 protein. The antioxidant role of Bcl-2 is suggested by its ability to protect cells from H_2O_2-induced and menadione (a quinone compound that generates O_2^-)-induced oxidative death (52,53). Overexpression of the Bcl-2 protein has also been shown to suppress apoptosis triggered by glucocorticoids, irradiation, chemotherapeutic agents, growth factor deprivation, and tumor necrosis factor alpha (TNF-α) (47,54). For these reasons, Bcl-2 is thought to inhibit cell death by decreasing the generation of ROS, which would in turn inhibit intracellular oxidation reactions critical to progression of the apoptotic program. However, Bcl-2 may have other functions that are unrelated to its putative antioxidant properties. Under hypoxic conditions with low levels of ROS, Bcl-2 is still able to protect cells from apoptotic death (55–57).

Stress-Related Proteins

Redirection of protein synthesis in response to hypoxia is part of a more general mammalian cell response to stresses such as temperature elevation and substrate deprivation. Such adverse conditions elicit expression of specific sets of stress proteins, namely, heat shock proteins (HSPs), glucose-regulated proteins (GRPs), oxygen-regulated proteins (ORPs), and hypoxia-associated proteins (HAPs). Stress proteins play an important role in adaptation to a new environment, as well as increased resistance to the same or other stresses. Which sets of proteins are upregulated varies with the cell type and the type of stress involved. However, recruitment of a specific family of proteins is not restricted to a single stimulus, as would be implied by its name.

The best-characterized of the stress proteins are HSPs, in particular the HSP70 family, the members of which are produced by many cell types as a protective response to high temperature, glucose deprivation, hypoxia, ischemia, the presence of heavy metal salts, and reoxygenation (58). These proteins function as chaperones that decrease protein denaturation and facilitate protein folding, processing, and trafficking. They are believed to help regulate biological processes including embryogenesis, differentiation, growth, and metabolism (59). GRPs have been less extensively characterized, but like HSPs, their response is not stress specific in regards either to their stimuli for upregulation or to their protection against other stresses (60). Investigations into stress protein induction in response to hypoxia and ischemia have implicated the upregulation of HSPs, GRPs, and ORPs. It should be noted that many of the ORPs have in fact been identified as a subset of the GRPs (61).

The effects of O_2 deprivation and subsequent reoxygenation on the potential coordinated regulation of HSPs and GRPs have been examined in CHO cells (61). The data indicate that while GRP induction is a sustained response to prolonged O_2 deprivation, HSP induction is only transient. Furthermore, when cells are reintroduced to an O_2-containing environment, the GRPs are repressed whereas the HSPs are again temporarily induced. Recently, an antioxidant role for both the constitutive and inducible forms of HSP70 has been confirmed (62). Overexpression of the HSP70s in rat cardiomyocytes confers a marked increase in tolerance to oxidative challenges including exposure to H_2O_2, hydroxyl radical, menadione, and hypoxia and reoxygenation. This finding is consistent with reports that oxidative stress induces expression of a number of HSPs, and that preexposure to mild hyperthermia or moderate oxidative stress significantly increases resistance of lymphocytes challenged with high doses of ROS (63). Interestingly, HSP32 (heme oxygenase-1) is also expressed in response to prolonged ($\sim$12 h) hypoxia (64).

Redirected protein synthesis is further demonstrated by the selective expression of ORP150 in cultured rat astrocytes exposed to hypoxia, but not in cells exposed to heat shock, H_2O_2, Co^{2+}, 2-deoxyglucose, or tunicamycin (65). Interestingly, the human analog of ORP150 is expressed in response to a wider range of conditions—exposure to hypoxia, 2-deoxyglucose, or tunicamycin, but not to heat shock—in the HeLa and human astrocytoma U373 cell lines (66). The wider expression range may be a characteristic of established cell lines because ORP150 induction is also restricted to hypoxia in human mononuclear phagocytes. As discussed earlier, many stress proteins upregulated by exposure to hypoxia are also involved in the response to other stresses. The term HAPs has been coined to describe a unique set of endothelial cell stress proteins that are distinct from HSPs, GRPs, and most ORPs. Like ORP150 in primary cells, these HAPs are preferentially upregulated by exposure to hypoxia, while other stresses such as heat and glucose deprivation do not induce them (67). Moreover, HAPs do not offer any cross-protection against other stresses. The glycolytic enzyme, GAPDH, has been identified as a HAP. GAPDH expression in the nuclear fraction of endothelial cells is upregulated mostly at the transcriptional level, and suggests a nonglycolytic role for this enzyme (68,69). In contrast, the GAPDH mRNA content in rat cardiomyocytes remains unchanged during hypoxia (64).

Although the molecular mechanisms underlying the regulation of stress proteins by hypoxia and reoxygenation are still unclear, it is known that their expression is, at least in part, regulated at the transcriptional level. Thus far, most insight into these regulatory mechanisms has been gained from studies of the HSP genes. Transient exposure of cultured myogenic cells to hypoxia has been shown to transcriptionally activate the HSP70 gene through binding of a nuclear transcription factor, heat shock transcription factor (HSTF), to distinct *cis*-acting regions, heat shock elements (HSEs), within the 5′ promoter (70). Furthermore, it was found that heat shock and hypoxia induce the HSP70 gene through the same regulatory HSTF–HSE complex. However, only one of two functional HSEs is required for heat shock inducibility, whereas both are needed for hypoxic induction (71). The activation of HSTF is highly dependent on cellular redox potential. Thiol-reducing agents inhibit HSTF activity and the expression of HSP genes in HeLa cells challenged by heat or other stresses (72), which is consistent with a report that the antioxidant *N*-acetyl-L-cysteine decreases the induction of HSP32 mRNA and protein in rat cardiomyocytes exposed to hypoxia (64).

Cell Metabolism

O_2 is a substrate for ATP generation and a modulator of the oxidation–reduction potential of cell culture media. Thus, pO_2 is a major determinant of the metabolic profile of a cell, affecting both nutrient consumption (oxygen, glucose, and glutamine) and waste production (lactate and ammonia) (73). Steady-state metabolic rates, including q_{O_2}, are not greatly affected by pO_2 between 10% and 100% DO (15–150 mm Hg). pO_2 dependence above 10 mm Hg is reflected primarily by changes in the oxidative phosphorylation ratios: cytoplasmic [ATP]/[ADP][P_i] and intramitochondrial [NAD^+]/[NADH] (74). As pO_2 and the respiration rate decrease, the rate of ATP synthesis declines. The resulting excess in ATP consumption over synthesis decreases the [ATP]/[ADP][P_i] ratio, which stimulates respiration (74). If pO_2 is held steady at the lower value, the [ATP]/[ADP][P_i] ratio will decline until a new steady state is attained. The rapid turnover of ATP allows these changes to occur within a few seconds and, hence, respiration appears unaffected by changes in pO_2. Moreover, a decrease in pO_2 may also lower the [NAD^+]/[NADH] ratio, which also enhances respiration (74). However, below a pO_2 of about 10 mm Hg, changes in the [ATP]/[ADP][P_i] and [NAD^+]/[NADH] ratios are not sufficient to maintain baseline respiration rates (74), and steady-state q_{O_2} decreases with pO_2 in a Michaelis–Menten-type manner with $K_{O_2} < 1$ mm Hg for hybridomas (6,13). Prolonged exposure to severe hypoxia (80% decrease in q_{O_2}) or anoxia has been shown to suppress CHO-K1 cell O_2 uptake, as evidenced by a transiently (~24 h) lower value of q_{O_2} at the same pO_2 value during reoxygenation (41).

To compensate for reduced aerobic metabolism during O_2 deprivation, the flux through glycolysis increases, with most of the glucose being converted to lactate. This phenomenon, referred to as the Pasteur effect, has been observed by many investigators (5,9,10,13,73,75). Glutamine metabolism is also O_2 dependent. Glutamine is an essential nitrogen and carbon source for the biosynthesis of metabolic intermediates and for energy production via the TCA cycle; ammonia and alanine are its major nitrogen-containing by-products. When excess O_2 is available, cultweed cells typically obtain most of their metabolic energy from glutamine oxidation (13). In contrast, when O_2 is limiting, cells favor glucose over glutamine for energy production. Because glucose metabolism is diverted from providing metabolic intermediates to providing energy at low pO_2, glutamine is then recruited to a greater extent for cell mass production. This is illustrated by an increase in glutamine utilization and a diminishing yield of NADH from glutamine at very low DO (6,13). The greater glutamine consumption rate also makes more nitrogen available for biosynthesis, which is consistent with higher specific production rates for ammonia and alanine (13,73). Similarly, Ozturk and Palsson (13) reported a large increase in the specific consumption rate of most amino acids under anoxic conditions in continuous hybridoma culture.

Interestingly, the metabolic alterations observed under hyperoxic conditions may resemble those occurring under hypoxia. A shift to anaerobic metabolism has been reported in hyperoxic cultures of CHO cells (76), hybridomas (17), WI-38 cells (9), and human lymphocytes (77). Hyperoxia has also been reported to increase the glutamine consumption rate (17). At first thought, it is counterintuitive that there would be an increasing reliance on glycolysis at high DO. However, the increased glucose consumption rate may result from O_2 toxicity. Enhanced glycolysis may compensate for impaired mitochondrial function under hyperoxia, as well as provide additional energy to prevent and repair oxidative damage.

Although an increase in anaerobic metabolism to make up for energy deficits is useful for surviving O_2 limitation, the greatest defense against hypoxia would be a suppression of energy demand. This strategy has been perfected by cells from the highly anoxia-tolerant aquatic turtle. In response to anoxia, ATP demand is decreased by 90% in aquatic turtle hepatocytes and by 50% in brain cortical cells (78).

Protein Synthesis

Protein turnover accounts for a large fraction of the ATP demand during normoxia (78). Therefore, inhibition of protein synthesis and degradation under O_2-limiting conditions would conserve energy and contribute to maintaining the ATP supply–demand balance. Indeed, an important feature and one of the first signs of hypoxia in many cell types is a rapid and dramatic decline in overall protein biosynthesis (78–81), which has been attributed to a block in translation (78,80). Protein synthesis rapidly recovers to control levels on reoxygenation. Protein degradation is increased by ~30% under hypoxia in NHIK 3025 cervical carcinoma cells (79), but is decreased by 94% in aquatic turtle hepatocytes (78). The rapid reduction of protein synthesis that occurs when O_2 is limited does not appear to be in response to blockage of the electron transport chain or to the associated change in redox state, but rather has been attributed to the presence of an O_2 sensor that senses the lack of molecular O_2 or O_2 by-products under hypoxic conditions.

Most reports on the effects of DO on protein production by cultured cells have been limited to antibody production by hybridomas and lymphocytes. The effects of pO_2 on the specific antibody production rate (q_{Ab}) are cell line specific, and the optimal DO level for cell growth is often different than that for antibody production. Over the range from 1% to 100% DO, q_{Ab} may be greater at low DO (7,75), greater at intermediate DO (6), or largely unaffected by DO (13,17,82,83). q_{Ab} may be decreased below 1% DO (6,84). However, other investigators have reported no change in q_{Ab} from anoxic exposure in continuous or perfusion culture (13,85). It must be emphasized that these conclusions are limited to the conditions tested. For example, although q_{Ab} for the CC9C10 hybridoma cell line was similar for 10–100% DO, q_{Ab} was 50% greater at 125% DO (17). Also, cells grown in serum-free media (SFM) appear to be more sensitive to low DO. Three hybridoma cell lines cultured in SFM exhibited a 50–85% lower q_{Ab} for DO $\leq$ 7.5%, while the same cell lines grown with serum showed no decrease in q_{Ab} even at 3% DO (82).

Much less has been reported regarding the effects of pO_2 or hypoxia on recombinant protein production. The limited

data available suggest that q_P is generally greater at higher DO, but the DO range over which the transition occurs is cell line dependent. Chotigeat et al. reported an increase in steady-state follicle-stimulating hormone (FSH) production with DO by CHO cells in perfused microcarrier culture; q_{FSH} was threefold greater at 60–90% DO than at 10% DO (86). Similarly, Wang et al. reported 50–240% greater q_{EPO} by L-929 cells at 75% DO in batch microcarrier culture as compared to that at 7.5–30% DO, with a larger increase in q_{EPO} at 75% DO for cells cultured in SFM (87). In contrast, q_{tPA} by BHK cells was unchanged at 7.5–75% DO in batch microcarrier culture and then increased with DO up to 225–300% DO for cells in serum-containing and serum-free media, although q_{tPA} did fall off at 420% DO (12). Cell growth was inhibited for DO $>$ 75% of air saturation, so that the maximum tPA titer was obtained at 225% and 150% DO for cells in serum-containing and serum-free media, respectively. At the other end of the DO spectrum, Lin et al. examined the effects of O_2 deprivation on tPA production by CHO cells in perfused microcarrier culture (5). q_{tPA} did not decrease from that at 33% DO during exposure to mild hypoxia (40–60% decrease in q_{O_2}), but was affected by more severe O_2 deprivation. q_{tPA} decreased by 80% during the first day of severe hypoxia (80% decrease in q_{O_2}), and then recovered to control levels during the next 2 days of severe hypoxia. Upon reoxygenation, q_{tPA} transiently (3 days) increased to twice the control value. During anoxic exposure, q_{tPA} gradually decreased by 80% over 2 days and then remained at that value even for 2 days after reoxygenation (5).

The attraction in using mammalian cells for recombinant protein expression lies with their ability to carry out complex posttranslational processing, which is often an important determinant of a protein's activity in vivo. Few studies have addressed the effect of pO_2 or hypoxia on protein glycosylation. Regoeczi et al. (88) examined the in vivo glycosylation of serum proteins obtained from rats placed in a hypobaric chamber at 380 mm Hg (80 mm Hg pO_2). The only significant difference in transferrin glycosylation was a greater proportion of molecules fucosylated (fucosylation index) in hypoxic vs. control animals (28% vs. 24%; $p = .02$). In contrast, the IgG fucosylation index was lower in hypoxic animals (10% vs. 20%; $p = .012$). Chotigeat et al. (86) examined the effects of DO on the isoelectric focusing (IEF) gel profile of FSH produced by CHO cells in perfused microcarrier culture. FSH acidity increased with DO, as evidenced by an increase in the fraction of FSH isoforms with pI $\leq$ 4.5 from 65% at 10% DO to 88% at 90% DO. The shift of FSH isoforms to lower pI was attributed to more extensive sialylation at higher DO values, which is consistent with a fivefold greater sialyltransferase activity in cells at 90% DO vs. that at 10% DO. Lin et al. (5) examined the specific activity and gel electrophoresis profiles of tPA produced by slowly growing CHO cells in perfused microcarrier culture. Exposure to hypoxia or anoxia and subsequent reoxygenation did not significantly affect tPA specific activity or site occupancy (fraction of tPA molecules glycosylated at three sites [68 kDa] vs. two sites [65 kDa]). However, during reoxygenation from anoxia, an extra tPA band appeared at 70 kDa. This band, which may have resulted from disrupted signal peptide cleavage or altered glycosylation, disappeared after 3 days at 33% DO.

Growth Factors and Their Receptors

The effects of O_2 on hematopoietic cell growth and differentiation may be caused by differences in the production of growth factors or their receptors with pO_2. Oxygen-regulated production of the cytokine erythropoietin (EPO) and of several angiogenic growth factors has been well established. There is evidence for specific effects of pO_2 on individual cell types. For example, macrophages produce cytokines in an O_2-dependent manner (89), while lymphocytes and macrophages that are cocultured under 5% O_2 cooperate to produce a factor that stimulates proliferation of erythroid precursor cells (90). There is also evidence that cord blood mononuclear cells produce more interleukin 6 (IL-6) under 5% compared to 20% O_2 (91).

Changes in growth factor receptor expression in response to low pO_2 may exhibit different patterns than those for the respective ligands. For example, while low pO_2 upregulates the transcription and synthesis of transforming growth factor β_1 (TGF-β) in dermal fibroblasts, culture under 2% O_2 for 24 h decreases receptor binding of [^{125}I]TGF-β by 65% and TGF-β type II receptor mRNA levels by 90% (92). Fibroblasts cultured under 2% O_2 also showed decreased receptor binding of [^{125}I]EGF (epidermal growth factor) that was reversible after 12 h in culture under 20% O_2 (92). For both receptors, decreased ligand binding was in large part the result of a lower number of binding sites. Exposure to hypoxia has also been shown to reversibly decrease [^{125}I]EGF binding to and the number of available EGF receptors (EGFR) on hepatocytes, without affecting total EGFR protein content (93). Interestingly, there was no effect of hypoxia on [^{125}I]insulin binding.

MOLECULAR MECHANISMS MEDIATING OXYGEN EFFECTS

The metabolic and cellular changes that accompany hypoxia are primarily the result of differential gene expression. The mechanisms underlying these gene-based responses involve hypoxia sensing, signal transduction, and ultimately expression of the appropriate gene.

Oxygen Sensing

Despite the advances being made toward identifying O_2-regulated cellular functions, the nature of the O_2 sensor itself and the path of the signal from the sensor to the machinery that controls protein expression and function remain unclear. Among the molecules that can bind to and react with molecular O_2 are heme proteins. By analogy to hemoglobin, heme proteins are thought to change conformational state between oxygenated and deoxygenated forms, thus regulating their function depending on pO_2. Involvement of a ferroheme protein in the O_2-sensing process, originally suggested by Goldberg et al. (94), is supported by several lines of evidence. Although much of the insight into mechanisms of O_2-sensing stems from studies

on EPO-producing hepatocytes and on carotid body type I cells, it seems that a heme-based O_2 sensor system is present in most cell types (95).

The response to hypoxia can be mimicked, both in magnitude and direction, by exposure to cobaltous ions or iron-chelating agents (96) or blocked by carbon monoxide, hallmarks of the involvement of a heme-containing protein. Expression of hypoxia-inducible genes (Table 1) is generally enhanced in cells treated with desferrioxamine (DF), an iron chelator. Iron chelators prevent synthesis of functional heme proteins by inhibiting the incorporation of ferrous atoms into the heme molecule. Cells with a nonfunctioning O_2 sensor are unable to detect O_2 present in the environment and are likely to interpret their state as hypoxic. Furthermore, because iron catalyzes the conversion of H_2O_2 to $OH\cdot$ and OH^- via the Fenton reaction, iron chelators may significantly alter levels of chemical messengers in the O_2-sensing pathway and lead to an hypoxic-like response (97). The heme protein hypothesis is further substantiated by experiments in which cobalt and nickel substitute for iron in the heme moiety. These transition metals exhibit a characteristic low binding affinity for O_2, thereby locking the heme protein in the deoxy conformation and sending an hypoxic signal. All the hypoxia-induced genes examined have been shown to be stimulated under such conditions (Table 1), while hypoxia-inhibited genes are also inhibited by Co^{2+} and DF. Conversely, the effects of hypoxia can be abrogated by using carbon monoxide (CO), which substitutes for O_2, to lock heme in the oxy conformation (94). Studies on expression of the phosphoenolpyruvate carboxykinase (PCK) gene in primary rat hepatocytes (139) further confirm the CO–O_2 sensor relationship, although in this case it acts conversely: PCK gene expression is inhibited under hypoxia, but is stimulated under elevated pO_2 or in the presence of CO (Table 1). Parallel effects of CO on O_2-sensing are evident in the carotid body, where ventilation is regulated (142). The current through K^+ channels in type I cells is rapidly downregulated by lowering pO_2. Addition of CO to hypoxic type I cells reverses the downregulation of K^+ currents. Collectively, these data implicate a heme protein that changes conformation on binding O_2 or CO as the primary O_2 sensor.

It is likely that the O_2 sensor is a membrane-bound, multisubunit cytochrome *b* protein that binds O_2 and reduces it to ROS such as superoxide and H_2O_2, which then in turn leads to downstream signaling of gene transcription. Numerous findings have contributed to this hypothesis. Absorption spectra of hepatoma cells and carotid body cells coincide with those of type *b*-like cytochromes that undergo a shift in conformation in response to changing pO_2 (143,144). Among this group of *b*-cytochromes are those that are part of the NADPH–oxidase enzyme complex, which generates H_2O_2 in an O_2-dependent manner. During hypoxia, the H_2O_2-generating activity of the oxidase would be inhibited, leading to lower H_2O_2 levels (4). In confirmation, the rate of H_2O_2 production by hepatoma cells decreased as pO_2 decreased (145). A regulatory role for H_2O_2 in the coupling of environmental pO_2 to intracellular events has been demonstrated: adding exogenous H_2O_2 to hypoxic cultures negated the expected increase in EPO production, while adding antioxidants or catalase to high-pO_2 cultures (to reduce H_2O_2) increased EPO mRNA levels (146). H_2O_2 may have several modes of action in transducing the pO_2 signal to the transcriptional machinery in O_2-responsive genes. Thus, it seems likely that changes in environmental pO_2 that regulate cellular functions are sensed by an H_2O_2-generating heme protein. Recent findings, however, indicate that hypoxic induction of genes for vascular endothelial growth factor (VEGF), aldolase, and tyrosine hydroxylase (TH) does not involve NADPH oxidase (103,147). In particular, in PC-12 cells DF and Co^{2+} do not interfere with H_2O_2 generation by the O_2 sensor, but rather affect TH gene expression by a mechanism further downstream from H_2O_2 formation (103).

O_2-sensing also appears to involve the regulation of nucleotide cyclases, heme-containing proteins that catalyze the conversion of intracellular ATP/GTP to cyclic AMP/GMP (148). Hypoxia decreases cAMP and cGMP levels, which in turn may alter stimulation of protein phosphorylation pathways. It is not known whether these nucleotide cyclases actually sense extracellular pO_2 levels, similar to the mechanism proposed for oxidases, or if they are regulated by a signal generated further downstream from the O_2 sensor. For example, H_2O_2 can affect activation of guanylate cyclase and alter cGMP levels (4). In this way, during hypoxia, lower concentrations of H_2O_2 could contribute to modifying the phosphorylation status of certain proteins.

Contrary to these observations, it is also possible that heme proteins may not be involved in the hypoxic regulation of certain genes, such as endothelial GAPDH (69). If a heme protein is involved, it possesses different structural or functional properties from those involved in other hypoxia-responsive systems as regulation of endothelial GAPDH (e.g., by CO) differs from that of EPO or VEGF (see Table 1).

Signal Transduction

Protein Phosphorylation. A decrease in H_2O_2 concentration could transduce the hypoxic signal through several different pathways, most likely involving protein phosphorylation and redox potential. Protein phosphorylation is an important intracellular transduction mechanism because of its regulation of several transcription factors. A heme protein in the bacterium *Rhizobium melliloti* senses hypoxia and activates its kinase domain, leading to phosphorylation of transcription factors and subsequent gene expression (149). In eukaryotes, NADPH oxidase, the putative O_2 sensor that operates in some cell types, becomes phosphorylated on activation (150), suggesting that the dependence of O_2-sensing on a phosphorylation cascade is conserved between bacteria and mammals. Tyrosine phosphorylation has been implicated in the hypoxic activation of NF-κB DNA-binding activity (151) and in the transcriptional and posttranscriptional induction of the VEGF gene (152,153). The DNA-binding activity of hypoxia-inducible factor-1 (HIF-1; see next section) is also dependent on its phosphorylation status. Treatment with phosphatase or with kinase inhibitors prevented hypoxia-inducible binding and EPO gene expression, demonstrating the requirement for HIF-1 to be in a phosphorylated form (154).

Table 1. Oxygen-Regulated Gene Expression

Gene	Stimulus				Regulation		References
	Hypoxia	DF	Co^{2+}	CO reversal[a]	Hypoxia-inducible DNA element	mRNA stability element	
Miscellaneous							
Erythropoietin (EPO)	+	+	+	Y	HIF-1	P	94,97–102
Tyrosine hydroxylase (TH)	+	+	+		HIF-1	Y	103–105
Transferrin (TF)	+				HIF-1		106
Inducible nitic oxide synthase	+				HIF-1		107
Heme oxygenase 1	+				HIF-1		108
Xanthine oxidase I	+						109
Retrotransposon VL30	+				HIF-1		110
Hypoxia-inducible factor-1α (HIF-1α)	+	+	+				111
Hypoxia-inducible factor-1β (HIF-1β)	+	+	+				111
Interleukin 2 (IL-2)	+						112
Interleukin 4 (IL-4)	+						112
Interferon-γ (IFN-γ)	+						112
Ornithine decarboxylase	+						113
Interleukin 10 (IL-10)	−						112
Angiogenic growth factors							
Vascular endothelial growth factor (VEGF)	+	+	+	Y	HIF-1	Y	96,114–118
Platelet-derived growth factor-β (PDGF-β)	+	+	+				96,119,120
Transforming growth factor-β (TGF-β)	+						121
Acidic/basic fibroblast growth factor (a/bFGF)	+		+				120
Tumor necrosis factor-α (TNF-α)	+				NF-IL-6 (P)		122
Endothelin	+						123
Interleukin 6 (IL-6)	+				NF-IL-6		112,124,125
Interleukin 1 (IL-1)	+						126
Interleukin 8 (IL-8)	+				NF-IL-6 (P)		124,127
Placental growth factor (PLGF)	−	−	−				96
Glucose transporters							
GLUT-1	+	+	+	Y	HIF-1	Y	128–131
GLUT-2	−	−	−				129
GLUT-3	+	+	+				129
Metabolic Enzymes							
Lactate dehydrogenase-A (LDH-A)	+	+	+		HIF-1		129,132–135
Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	+	+	+	N			69
Aldolase A	+	+	+		HIF-1		129,135,136
Enolase 1	+				HIF-1		135,136
Phosphoglycerate kinase 1 (PGK-1)	+	+	+		HIF-1		132,136
Phosphofructokinase L (PFK-L)	+	+	+		HIF-1		132,136
Pyruvate kinase	+						136,137
Hexokinase	+						137
Triose phosphate isomerase	+						138
Phosphoenolpyruvate carboxykinase (PCK)	−		−	Y			139,140
Tyrosine aminotransferase	−						141

Note: DF, desferrioxamine; +, induced gene expression; −, suppressed gene expression; Y, yes; N, no; P, probable, but not proven; HIF-1, hypoxia-inducible factor-1-binding motif; NF-IL-6, nuclear factor–interleukin-6-binding motif.
[a]Reversal of hypoxic effects by carbon monoxide (CO).

Redox Potential. Although hypoxia-mediated enhancement of DNA-binding activity of several transcription factors could reflect changes in the phosphorylation state of these proteins, it could also reflect changes in the cellular redox state. Cellular redox potential is governed in part by the ratio of reduced (SH) to oxidized (S–S) thiols that results from the surrounding pO_2. Under normoxic conditions, the equilibrium is maintained by H_2O_2 either through direct oxidation of reduced glutathione (GSH) to oxidized glutathione (GSSG) or through formation of the hydroxyl radical, which can oxidize glutathione and protein thiol groups (4). Under hypoxic conditions, the de-

creasing H_2O_2 concentration is accompanied by a shift in the equilibrium toward the reduced forms of glutathione and protein thiols. This shift in redox state can regulate the activity of K^+ channels and the DNA-binding activity of transcription factors (4).

Several studies have coupled changes in transcription factor activity with perturbations in the redox state of the cell. Transcription factors AP-1, NF-κB, and p53, as well as protein factors such as heat shock proteins (HSPs), are redox sensitive. An ubiquitous nuclear protein called Ref-1 serves as a redox factor that stimulates AP-1 DNA-binding activity through single conserved cysteine residues in the c-Fos and c-Jun proteins that comprise the AP-1 transcription complex (155). AP-1 activity is enhanced by antioxidants and reducing agents (155,156), whereas it is inhibited by H_2O_2 (157) and sulfhydryl blockers (155). Furthermore, data suggest that the c-*fos* and c-*jun* family of genes are regulated at the transcriptional level as well, with exposure to hypoxia inducing mRNA expression (114). However, in these latter studies AP-1 activity was not assessed. In contrast to AP-1, NF-κB activity is induced by oxidative stress: H_2O_2 can activate NF-κB, while treatment with thiol reagents or other antioxidants abolishes its activity (156,158). A conflicting report has shown enhanced NF-κB activity under hypoxic conditions (151). Like AP-1, it seems that NF-κB also requires reduction of specific cysteine residues in its DNA-binding domain, since its association with DNA in vitro is inhibited by alkylating or oxidizing agents (159). However, there is currently no evidence linking modification of NF-κB thiol groups to regulation of its biological activity. This fact demonstrates the complex nature of regulation of cellular signaling through redox chemistry.

The functions of HSPs are multifold: they protect cells from stress from heat, ROS, and sulfhydryl reagents because they aid protein folding, processing, and transport. Expression of HSPs is dependent on binding of heat shock transcription factors (HSFs) to their promoter region. The relevance of the cellular redox state in the heat shock response is demonstrated by the inhibition of HSF activity and HSP gene expression by thiol reducing agents (72). The induction of HSP expression with reoxygenation is consistent with these findings (61).

The role of redox status on the basic helix-loop-helix (bHLH) group of proteins, a broad group of transcription factors that share common helix-loop-helix regions required for dimerization, is also of considerable interest. Binding and activation of USF, a bHLH transcription factor, proceeds only when two cysteine residues in the HLH dimer interface are reduced (160). This mechanism also implicates redox regulation in the assembly of other transcription factors, in particular of the activation of HIF-1, a heterodimer consisting of two bHLH transcription factors, the HIF-1α and HIF-1β subunits (111).

Of the factors mentioned previously, HIF-1 is probably the most relevant to low-pO_2 studies because it is an ubiquitous transcription factor (see Table 1) whose function is restricted to the hypoxic regulation of a broad range of physiologically relevant genes (161–163). HIF-1 DNA-binding activity and the protein levels of HIF-1α and HIF-1β each increase exponentially as cells are exposed to decreasing pO_2 in vitro within the physiological range. A half-maximal response was seen between 1.5% and 2% O_2 in the gas phase, with a maximal response at 0.5% O_2 (124). Maintaining the proper redox ratio is essential for the hypoxic activation of HIF-1. Changing the redox state of cells through treatment with oxidizing agents impairs hypoxia signaling mechanisms by disrupting expression of HIF-1α and HIF-1 DNA-binding activity (164). HIF-1 must remain in the reduced state to enable binding to DNA. No cysteine residues were found in the basic DNA-binding domain of either HIF-1α or HIF-1β. However, one cysteine was found in the HLH region of HIF-1β and several in the PAS (Per-ARNT-Sim; named for the domain found in the *Drosophila* transcription factors Period and Single-minded and the mammalian aryl hydrocarbon receptor nuclear translocator protein) domain of each subunit (164). Interaction between the HLH and PAS domains is necessary for heterodimerization of HIF-1α and HIF-1β and for subsequent DNA binding. The redox-sensitive activation of HIF-1 has since been attributed primarily to the HIF-1α subunit (165).

Another factor that has been shown to activate transcription of specific target genes in hypoxic cells is C/EBPβ. C/EBPβ shows a structural feature akin to the AP-1 complex with a cysteine residue similarly positioned as in c-Fos and c-Jun that is critical for DNA-binding activity (124,155). C/EBPβ protein binds to an hypoxia-inducible element called the NF-IL-6 site in a tissue-specific manner (124,125). Although this site is distinct from the HIF-1-binding motif, the cellular processes leading to gene transcription may still be related. Given the analogy between the actions of HIF-1 and C/EBPβ and the implication that redox potential plays a role in both HIF-1- and C/EBPβ-mediated hypoxia-inducing mechanisms, the NF-IL-6 site becomes another important locus for induced gene expression under O_2 deprivation. The presence of tissue-specific factors regulates whether transcription can be driven by HIF-1 or C/EBPβ.

In addition to increased transcription rates, gene expression during hypoxia is also regulated by the stability of mRNA. Redox may also contribute to this posttranscriptional mode of gene regulation. Czyzyk-Krzeska (4) reported that binding of hypoxia-inducible protein factors to tyrosine hydroxylase mRNA stability elements requires that they be in the reduced state.

Thus, posttranslational signaling mechanisms that involve both phosphorylation and sulfhydryl group modifications contribute to the transcriptional and posttranscriptional response to hypoxia. HIF-1 activation best exemplifies these modes of regulation. In addition, activation of HIF-1 may depend on ligand binding. Association of HSP90 with HIF-1α allows this subunit to change conformation, thereby enabling dimerization with HIF-1β and subsequent DNA binding of the HIF-1 complex (166). Considering this information, it now becomes logical to extend the previous hypothesis suggesting that the signaling cascade begins with a heme protein O_2 sensor and is propagated by an intermediate, such as H_2O_2. Intermediate molecules are likely to target certain transcription factors, such as HIF-1, and regulate their binding to the respective genes.

Gene Expression

Transcriptional Regulation. The identification of the aforementioned transcription factors has facilitated our understanding of O_2 regulation of genes and consequently advanced our understanding of the physiological adaptation to hypoxia. Despite a significant reduction in total mRNA synthesis when cells are deprived of O_2, transcription of several mammalian genes that encode for various hormones, cytokines, and enzymes is markedly increased. In particular, low pO_2 induces expression of genes controlling erythropoiesis, ventilation, angiogenesis, and energy metabolism. Furthermore, the commonality among the regulatory elements present in these genes suggest that they are part of a widespread O_2-sensing system. Although beyond the scope of this article, cellular adaptation to the other extreme, hyperoxia, also involves the upregulation of certain genes which, in this case, lead to detoxification of ROS (36).

EPO. Progress in understanding O_2-dependent gene expression has mostly stemmed from characterizing hypoxia-induced production of erythropoietin (EPO) because the effects are more dramatic for this cytokine than for the other O_2-sensitive genes studied thus far. EPO is the key hormone in erythropoiesis: synthesized primarily in the adult kidney and fetal liver, it travels to hematopoietic tissues where it binds to its receptor on erythroid progenitor cells, protecting them from apoptosis and allowing them to proliferate and differentiate into mature red blood cells. Under the hypoxic stress of anemia or reduced blood oxygenation, EPO production is as much as 1000 fold greater than normoxic levels (167). By increasing the number of red blood cells, EPO enhances the O_2-carrying capacity of the blood. Studies on the transcriptional control of the EPO gene have been greatly facilitated by the use of the cultured human fetal hepatoma cells Hep3B and HepG2 (168). Increased EPO production under hypoxia in these model systems mirrors the in vivo response. There is a close correlation between the level of hypoxia and the production rates of EPO protein and EPO mRNA in these cells.

Regulation of transcription rate involves interactions between *trans*-acting elements (transcription factors) and *cis*-acting elements within the promoters and enhancers of regulated genes. The best-studied mechanism of hypoxic regulation is the binding of HIF-1 to DNA, and is best exemplified by regulation of the EPO gene. The *cis* element most critical to hypoxic induction of EPO transcription is located in the 3′-flanking region of the gene, 120 bp downstream from the polyadenylation site. The 3′ enhancer includes three sequences that contribute to the hypoxic response: (*1*) a 5′ portion containing the highly conserved HIF-1-binding site, (*2*) the middle portion, which does not bind specific proteins but is necessary for the hypoxic induction of transcription, and (*3*) a 3′ portion containing highly conserved nuclear receptor-binding half-sites that resemble known steroid and thyroid hormone response elements. A variety of hormones act by binding to nuclear receptors. In particular, the binding of a specific nuclear receptor, orphan nuclear factor 4, to the EPO enhancer amplifies the transcriptional activation by HIF-1, demonstrates a marked impact on the hypoxic induction of the EPO gene and contributes to tissue specificity of this response (169).

The importance of the 3′ enhancer had previously misled some to believe that it was solely responsible for hypoxic induction of the EPO gene. However, transfection studies with the enhancer revealed that mRNA synthesis of a reporter gene was only induced about 10 fold, which is well below the induced level of endogenous EPO mRNA (170). This discrepancy was resolved through identification of O_2-sensitive elements within the promoter region, 5′ to the EPO gene (171). The 3′ enhancer and 5′ promoter act synergistically to achieve at least a 40-fold stimulation of transcription of the EPO gene under hypoxia, commensurate with endogenous EPO mRNA levels. Similar to the enhancer region, the EPO promoter contains nuclear hormone response half-sites. Alone, they do not contribute to hypoxic induction, but they are needed to amplify the induction signal. It has also been proposed that the EPO gene may be negatively regulated by higher pO_2. Other elements in the EPO promoter including GATA sites and ribonucleoprotein-binding sites may contribute to the inhibition of a constitutive activator protein and lead to repressed transcription under normoxia (167). During hypoxia, when the negative regulators are reduced, they could be displaced from their corresponding DNA-binding sites and allow transcription to proceed. This finding is consistent with studies that suggest that under high pO_2 the O_2 sensor generates greater amounts of ROS that lead to the inhibition of EPO gene expression (146).

More recently, other transcription factors have been implicated in transducing the signal from the HIF-1 site to the transcriptional initiator. The non-DNA-binding transcriptional activator P300/CBP binds to HIF-1α and enhances hypoxic induction by increasing the response of the EPO enhancer (172). P300/CBP is a multifunctional protein that, in addition to regulating several tissue-specific enhancers, plays a role in controlling the cell cycle and differentiation pathways. The HIF-1-binding site within the EPO enhancer also binds constitutively to a smaller protein complex consisting of the activating transcription factor-1 (ATF-1) and the cAMP-responsive element-1 (CREB-1) (173). During hypoxia, these two transcription factors are displaced by HIF-1. Thus, HIF-1-dependent induction of gene expression involves interactions between different transcription factors and sequences. It is likely that these interactions direct the specificity of the hypoxic response, that is, which genes or cells are affected.

Several findings have led to the conclusion that HIF-1-mediated induction of gene expression is not restricted to the EPO gene or to cells originating from the kidney or liver. Transfection experiments in many different cell lines using the EPO 3′ enhancer coupled to reporter genes have demonstrated the universality of this O_2-sensing and signal transduction mechanism (174). Factors with similar affinity for specific DNA sequences within the enhancer were activated regardless of whether or not the cells were EPO producing (98,175). Furthermore, mutated EPO enhancers, which are known to prevent HIF-1 binding, abolished hypoxic induction of mRNA synthesis of the reporter genes. O_2 regulation of other genes in an HIF-1-dependent

fashion is confirmed by the observation that the key regulatory sites for many other O_2-regulated genes contain the HIF-1-binding consensus sequence (see Table 1) with the core motif CGTG (161). Among those to be discussed next are the genes encoding TH, VEGF, GLUT-1, and several glycolytic enzymes.

TH. The most immediate response to hypoxia at the whole-body level is the regulation of ventilation. The sensitivity of type I cells in the carotid body to a decrease in pO_2 is greater than in other O_2-sensitive cells: inhibition of conductance through K^+ channels, membrane depolarization, increase in intracellular calcium, and release of neurotransmitters occur rapidly. Neurotransmitters convey the hypoxic signal to the sinus nerve. Hypoxia stimulates the protein activity, as well as the fast induction of gene expression, of tyrosine hydroxylase (TH), the rate-limiting enzyme in the synthesis of the neurotransmitter dopamine. The molecular mechanisms by which hypoxia induces increased TH gene expression have mostly been explored in PC-12 (pheochromocytoma cell line) cells, which exhibit similar O_2-dependent characteristics as the scarce type I primary cells (104). A decrease of O_2 levels (to 15% O_2 in the gas phase) produces an hypoxic effect. When PC-12 cells are exposed to 10% O_2, TH mRNA levels reach a maximum within 6 h and are sustained at a high level for up to 53 h thereafter (4). The initial increase in TH mRNA is primarily attributed to a rapid induction of transcription, whereas maintenance of TH mRNA during long-term exposure to hypoxia involves mechanisms that enhance mRNA stability (see later section).

The *cis* elements required for hypoxic induction of TH are located in a promoter region containing HIF-1, AP-1, and AP-2 sites (105). As for the EPO gene, binding of the HIF-1 complex to its site is required for hypoxic induction, but other interactions are also involved. Binding of c-Fos/c-Jun heterodimers to the AP-1 site within the TH promoter was shown to increase in hypoxic cells (105). Mutation of the AP-1 site, which prevented binding of these transcription factors, inhibited hypoxic induction. Another similarity to regulation of the EPO gene is the possibility that nuclear receptors may also play a role in cooperation with HIF-1 for TH (176). Moreover, as with EPO mRNA, decreases in H_2O_2 levels caused by the addition of antioxidants or catalase induced TH mRNA expression. This suggests that H_2O_2 serves as an intermediate during the hypoxic regulation of both the TH and EPO genes.

Angiogenic Growth Factors. Angiogenesis is a fundamental process not only in development but also in wound healing, recovery from ischemic injury, and the pathogenesis of tumor growth. The common denominator among these processes is the hypoxic microenvironment in which they usually occur. The formation of new capillaries increases blood flow to the tissues, allowing delivery of an adequate supply of O_2. Control of vascular functions and blood vessel proliferation under hypoxic conditions has been associated with a group of O_2-regulated genes that encode several growth factors important for stimulating growth and migration to sites of ischemic injury (see Table 1). Some of these genes that have been reported to be inducible by hypoxia share homologous sequences and similar regulatory mechanisms to those operating in the EPO and TH genes.

Of the angiogenic cytokines, VEGF, which is believed to be the central mediator of blood vessel formation, has garnered the most attention. VEGF is produced by many different cell types and tissues, and its expression is significantly amplified by hypoxia. Hypoxic induction of VEGF mRNA has been demonstrated both in vivo and in cultured cells (115). The preponderance of data suggest that VEGF regulation employs similar O_2-sensing, signal transduction, and gene expression mechanisms as both EPO and TH (Table 1). Transfection experiments using reporter genes localized the *cis*-acting element containing the well-conserved HIF-1-binding site to a 28-bp sequence, 900 bp upstream of the transcriptional start site (116,177). Minchenko et al. (178) previously reported that two O_2-responsive elements existed in the 5′ and 3′ regions flanking the VEGF gene. AP-1- and AP-2-binding sites were also identified within the VEGF promoter (114,177).

The C/EBPβ DNA-binding factor has been found to mediate IL-6 induction by hypoxia in endothelial cells and vascular smooth muscle cells (124,125). Transfection of IL-6 promoter–reporter gene constructs revealed the NF-IL-6-binding site as the hypoxia-inducible element, suggesting that a similar mechanism may account for hypoxic induction of other genes that contain NF-IL-6 sites, such as TNF-α and IL-8.

Genes Involved in Energy Metabolism. To maintain viability and function, cells have carefully choreographed strategies of nutrient utilization. During periods of O_2 deprivation, when the O_2-dependent tricarboxylic acid cycle and oxidative phosphorylation are impaired, cells adapt by generating ATP primarily through anaerobic glycolysis. Because the glycolytic pathway produces much less ATP than oxidative metabolism, the glycolytic capacity of the cells must be enhanced under hypoxic conditions. The rate of glycolysis can be increased by transporting more glucose into the cell, by upregulating the activity and synthesis of glycolytic enzymes, and by downregulating expression of gluconeogenic enzymes that consume large amounts of ATP. As discussed next, hypoxia governs expression of the genes involved in all three modes of adaptation.

In the first mode, glucose uptake increases to compensate for lower ATP yields. The expression of glucose transporters has been found to increase in response to several environmental stimuli. Most notably, hypoxia induces insulin-independent glucose transporter 1 (GLUT-1) protein and mRNA both in vivo and in various cell lines (179). The induction of GLUT-1 contributes to sustained energy levels under hypoxic stress in muscle cells, endothelial cells, adipocytes, HeLa cells, and HT 1080 (fibrosarcoma) cells. Increases in glucose uptake and lactate production are consistent with increases in glucose transport (128,180). It is likely that hypoxic regulation of the GLUT-1 gene, whose enhancer contains two phorbol ester response elements and a cAMP response element, depends on synergy between multiple *cis*-acting elements and transcription factors (95). In addition to GLUT-1, its isoform GLUT-3 is also induced by hypoxia in the HepG2 hepatoma cell line, whereas GLUT-2 expression is inhibited (129). The isoform-specific regulation of glucose transport coin-

cides with the decrease in the apparent K_m for glucose observed under hypoxia: GLUT-3, which demonstrates the most dramatic induction by hypoxia, also has the lowest apparent K_m for glucose (128,180). Increased steady-state levels of GLUT-1 mRNA in response to hypoxia are not only the result of increased transcription; GLUT-1 is also subject to regulation through mechanisms that stabilize mRNA.

The remaining two modes by which pO_2 alters glucose metabolism involve the regulation of a number of glycolytic and gluconeogenic enzymes. As expected, the response to hypoxia is opposite in direction depending on the pathway with which the enzyme is associated (see Table 1). In hypoxic surroundings, increased expression of rate-limiting glycolytic enzymes would shift metabolism toward glycolysis. However, in a well-oxygenated environment in selected tissues, enhanced expression of PCK would favor gluconeogenesis.

In general, the three groups of genes—those for glucose transporters, glycolytic enzymes, and gluconeogenic enzymes—exhibit patterns of expression in response to hypoxia that are consistent with their response to transition metal ions (e.g., cobalt) or iron chelators (e.g., DF), as has been shown for other O_2-regulated genes (Table 1). Some of these genes possess *cis*-acting elements, which convey this response and cross-compete with the EPO enhancer for binding to HIF-1 (132). This suggests that hypoxic regulation of EPO and certain metabolic genes are linked by a similar O_2-sensing and signal transduction mechanism. Indeed, the HIF-1-binding consensus sequence has been found in several of these genes (Table 1).

Because induction of LDH-A activity by hypoxia is one of the greatest among the glycolytic enzymes, its hypoxia-inducible elements have been studied in detail. Expression of the mouse *LDH-A* gene in response to hypoxia is driven by cooperation among three domains, which are arranged similarly to those observed in the EPO 3′ enhancer. There is (*1*) an HIF-1-binding site, which is required but not sufficient for hypoxic induction of gene expression, (*2*) an adjacent sequence, which is located in an analogous position as a critical region in the EPO enhancer, and (*3*) a cAMP response element (CRE) (133). At a minimum, in transfection experiments the HIF-1 site must be accompanied by either of the other two domains to confer hypoxic inducibility. Furthermore, the CRE-binding protein is highly homologous to P300, which binds to HIF-1α (see earlier discussion). Despite the substantial induction of LDH-A mRNA by hypoxia, no effect on its isozyme LDH-B was observed (129). Therefore, the distinctive pattern of response to hypoxia by the genes involved in energy metabolism is both isoform- and isozyme-specific, as in the cases of GLUT-1 (and GLUT-3) and LDH-A, respectively. Finally, hypoxic induction of LDH and GAPDH, which are used as a measure of viability and as an internal control for electrophoretic purposes, respectively, may lead to the misinterpretation of results collected in the laboratory if the differential effects of pO_2 on their expression are not considered.

Posttranscriptional Regulation. In addition to the mRNA transcription rate, the decay rate should not be overlooked; it, too, is an important contributing factor in increasing gene expression during hypoxia. A potential advantage associated with this bimodal regulation of gene expression is conservation of energy. Initially, hypoxia increases specific mRNAs through transcriptional induction, but, during longer exposure times, mechanisms conferring mRNA stability may take over. This phenomenon is illustrated by the increase in TH mRNA stability during long-term hypoxia, which allows maintenance of TH mRNA levels without the energy demand that accompanies synthesis of new mRNA molecules (104). Studying posttranscriptional regulation by hypoxia is complicated by the fact that many inhibitors of transcription nonspecifically stabilize mRNA. For this reason, although hypoxic regulation of EPO mRNA stability is expected, it has yet to be confirmed. In contrast, using nonstabilizing transcription blockers and other techniques, hypoxic regulation at the posttranscriptional level has been demonstrated for VEGF and GLUT-1 mRNAs (152), in addition to TH mRNA (104).

Both stability and instability elements may be involved in the regulation of steady-state EPO, TH, VEGF, and GLUT-1 mRNA levels. These *cis*-acting sequences within mRNA could direct secondary structure, act as binding sites for *trans*-acting regulatory proteins, or serve as target sites of nucleases (4). The hypoxia-regulated TH mRNA stability element has been identified as a pyrimidine-rich tract in the 3′ untranslated region (UTR) termed the hypoxia-inducible protein-binding sequence (HIPBS). Oxygen regulation by protein binding to this site was confirmed first by mutations that prevent protein binding and consequently decrease TH mRNA half-life and second by the ability of HIPBS to confer stability on reporter mRNA. The HIPBS closely resembles sequences in the 3′ UTR of both EPO and VEGF mRNAs (153,181). Moreover, similar or the same protein factors as those that bind to HIPBS in TH mRNA also bind to the homologous sequences in EPO and VEGF mRNAs (4). The most common determinants of mRNA stability in mammalian cells are AU-rich elements (182). In addition to three HIPBS motifs, the 3′ UTR region of VEGF mRNA contains multiple AU-rich stabilizing and destabilizing elements that also appear to mediate stability by forming hypoxia-regulated complexes (111,155). Thus, in combination with transcriptional upregulation, the presence of O_2-regulated mRNA stability and instability elements could account for enhanced gene expression under hypoxic conditions.

ACKNOWLEDGMENTS

Supported by NSF Grant BES-9402030. DLH was supported in part by Carcinogenesis Training Grant (NIH CA09560).

BIBLIOGRAPHY

1. W.F. McLimans, L.E. Blumenson, and K.V. Tunnah, *Biotechnol. Bioeng.* **10**, 741 (1968).
2. R.J. Fleischaker and A.J. Sinskey, *Eur. J. Appl. Microbiol. Biotechnol.* **12**, 193 (1981).
3. M.C. Borys and E.T. Papoutsakis, in J.B. Griffiths, A. Doyle, and D.G. Newell, eds., *Protocols in Cell and Tissue Culture*, Wiley, Chichester, 1993, p. 1.

4. M.F. Czyzyk-Krzeska, *Respir. Physiol.* **110**, 99 (1997).
5. A.A. Lin, R. Kimura, and W.M. Miller, *Biotechnol. Bioeng.* **42**, 339 (1993).
6. W.M. Miller, C.R. Wilke, and H.W. Blanch, *J. Cell. Physiol.* **132**, 524 (1987).
7. A. Mizrahi, G.V. Vosseller, Y. Yagi, and G.E. Moore, *Proc. Soc. Exp. Biol. Med.* **139**, 118 (1972).
8. S. Reuveney, D. Velez, J.D. Macmillan, and L. Miller, *J. Immunol. Methods* **86**, 53 (1986).
9. A.K. Balin, D.B.P. Goodman, H. Rasmussen, and V.J. Cristofalo, *J. Cell. Physiol.* **89**, 235 (1976).
10. D.G. Kilburn, M.D. Lilly, D.A. Self, and F.C. Webb, *J. Cell. Sci.* **4**, 25 (1969).
11. G.W. Taylor, J.P. Kondig, S.C. Nagle, Jr., and K. Higuchi, *Appl. Microbiol.* **21**, 928 (1971).
12. J. Wang, H. Honda, H. Watanabe, and T. Kobayashi, *J. Ferment. Bioeng.* **79**, 579 (1995).
13. S.S. Ozturk and B.Ø. Palsson, *Biotechnol. Prog.* **6**, 437 (1990).
14. M.A. Cacciuttolo, L. Trinh, J.A. Lumpkin, and G. Rao, *Free Radical Biol. Med.* **14**, 267 (1993).
15. A.R. Oller, C.W. Buser, M.A. Tyo, and W.G. Thilly, *J. Cell. Sci.* **14**, 43 (1989).
16. P. van der Valk, J.J.P. Gille, A.B. Oostra, E.W. Roubos, T. Sminia, and H. Joene, *Cell Tissue Res.* **239**, 61 (1985).
17. C.C.H. Jan, A. Petch, N. Huzel, and M. Butler, *Biotechnol. Bioeng.* **54**, 153 (1997).
18. W.G. Taylor, A. Richter, J. Evans, and K.K. Sanford, *Exp. Cell. Res.* **86**, 152 (1974).
19. L. Packer and C. Fuehr, *Nature* (*Lond.*) **267**, 423 (1977).
20. A.K. Balin, A.J. Fisher, and D.M. Carter, *J. Exp. Med.* **160**, 152 (1984).
21. V. Falanga and R.S. Kirsner, *J. Cell. Physiol.* **154**, 506 (1993).
22. J. Katahira and H. Mizoguchi, *Int. J. Cell Cloning* **5**, 420 (1987).
23. I.N. Rich and B. Kubanek, *Br. J. Haematol.* **52**, 588 (1982).
24. H.E. Broxmeyer, S. Cooper, and T. Gabig, *Ann. N.Y. Acad. Sci.* **554**, 177 (1989).
25. H.E. Broxmeyer, S. Cooper, L. Lu, M.E. Miller, C.D. Langenfeld, and P. Ralph, *Blood* **76**, 323 (1990).
26. M.G. Cipolleschi, P.D. Sbarba, and M. Olivotto, *Blood* **82**, 2031 (1993).
27. M.R. Koller, J.G. Bender, W.M. Miller, and E.T. Papoutsakis, *Exp. Hematol.* **20**, 264 (1992).
28. M.R. Koller, J.G. Bender, E.T. Papoutsakis, and W.M. Miller, *Ann. N.Y. Acad. Sci.* **665**, 105 (1992).
29. S. Smith and H.E. Broxmeyer, *Br. J. Haematol.* **63**, 29 (1986).
30. J.A. LaIuppa, E.T. Papoutsakis, and W.M. Miller, *Exp. Hematol.* **26**, 835 (1998).
31. M.G. Cipolleschi, G. D'Ippolito, P.A. Bernabei, R. Caporale, R. Nannini, M. Mariani, M. Fabbiani, P. Rossi-Ferrini, M. Olivotto, and P.D. Sbarba, *Exp. Hematol.* **25**, 1187 (1997).
32. O. Genbacev, Y. Zhou, J.W. Ludlow, and S.J. Fisher, *Science* **277**, 1669 (1997).
33. M. Gassmann, J. Fandrey, S. Bichet, M. Wartenberg, H.H. Marti, C. Bauer, R.H. Wenger, and H. Acker, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2867 (1996).
34. M.V. Wiles, *Methods Enzymol.* **225**, 900 (1993).
35. A.J. Potocnik, P.J. Nielsen, and K. Eichmann, *EMBO J.* **13**, 5274 (1994).
36. T.H. Rushmore, M.R. Morton, and C.B. Pickett, *J. Biol. Chem.* **266**, 11632 (1991).
37. T.R. Bradley, G.S. Hodgson, and M. Rosendaal, *J. Cell Physiol.* **97**, 522 (1978).
38. Y. Ishikawa and T. Ito, *Eur. J. Haematol.* **40**, 126 (1988).
39. R.C. Meagher, A.J. Salvado, and D.G. Wright, *Blood* **72**, 273 (1988).
40. H. de Groot and M. Brecht, *Biol. Chem.* **372**, 35 (1991).
41. A.A. Lin and W.M. Miller, *Biotechnol. Bioeng.* **40**, 505 (1992).
42. S. Hannah, K. Mecklenburgh, I. Rahman, G.J. Bellingan, A. Greening, C. Haslett, and E.R. Chilvers, *FEBS Lett.* **372**, 233 (1995).
43. M. Hermes-Lima and K.B. Storey, *Am. J. Physiol.* **265**, R646 (1993).
44. G.J. Moffat, A.W. McLaren, and C.R. Wolff, *J. Biol. Chem.* **269**, 16397 (1994).
45. D.K. Das, R.M. Engelman, and Y. Kimura, *Cardiovasc. Res.* **27**, 578 (1993).
46. C. Richter, V. Gogvadze, R. Laffranchi, R. Schlapbach, M. Schweizer, M. Suter, P. Walter, and M. Yaffee, *Biochim. Biophys. Acta* **1271**, 67 (1995).
47. C.M. Payne, C. Bernstein, and H. Bernstein, *Leuk. Lymphoma* **19**, 43 (1995).
48. W. Malorni, R. Rivabene, M.T. Santini, and G. Donelli, *FEBS Lett.* **327**, 75 (1993).
49. N. Ramakrishnan, D.E. McClain, and G.N. Catravas, *Int. J. Radiat. Biol.* **63**, 693 (1993).
50. R.R. Ratan, T.H. Murphy, and J.M. Baraban, *J. Neurochem.* **62**, 376 (1994).
51. V.J. Forrest, Y. Kang, D.E. McClain, D.H. Robinson, and N. Ramakrishnan, *Free Radical Biol. Med.* **16**, 675 (1994).
52. D.M. Hockenberry, Z.N. Oltvai, X. Yin, C.L. Milliman, and S.J. Korsmeyer, *Cell* **75**, 241 (1993).
53. S.J. Korsmeyer, X. Yin, Z.N. Oltvai, D.J. Veis-Novack, and G.P. Linette, *Biochim. Biophys. Acta* **1271**, 63 (1995).
54. A.F.G. Slater, C.S.I. Nobel, and S. Orrenius, *Biochim. Biophys. Acta* **1271**, 59 (1995).
55. S. Shimizu, Y. Eguchi, H. Kosaka, W. Kamiike, H. Matsuda, and Y. Tsujimoto, *Nature* (*Lond.*) **374**, 811 (1995).
56. M.D. Jacobson and M.C. Raff, *Nature* (*Lond.*) **374**, 814 (1995).
57. R.J. Muschel, E.J. Bernahard, L. Garza, W.G. McKenna, and C.J. Koch, *Cancer Res.* **55**, 995 (1995).
58. L.W. Browder, M. Pollock, R.W. Nickells, J.J. Heikkila, and R.S. Winning, in M. DiBerardino and L.D. Etkin, eds., *Developmental Biology: A Comprehensive Synthesis. Genomic Adaptability in Cell Specialization*, Plenum Press, New York, 1989, p. 97.
59. J.R. Subjeck and T. Shyy, *Am. J. Physiol.* **250**, C1 (1986).
60. A.S. Lee, *Curr. Opin. Cell Biol.* **4**, 267 (1992).
61. J.J. Sciandra, J.R. Subjeck, and C.S. Hughes, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4843 (1984).
62. K.Y. Chong, C.C. Lai, S. Lille, C. Chang, and C.Y. Su, *J. Mol. Cell Cardiol.* **30**, 599 (1998).
63. M. Marini, F. Frabetti, D. Musiani, and C. Franceschi, *Int. J. Radiat. Biol.* **70**, 337 (1996).
64. D.R. Borger and D.A. Essig, *Am. J. Physiol.* **274**, H965 (1998).
65. K. Kuwabara, M. Matsumoto, J. Ikeda, O. Hori, S. Ogawa, Y. Maeda, K. Kitagawa, N. Imuta, T. Kinotshita, D.M. Stern, H. Yanagi, and T. Kamada, *J. Biol. Chem.* **271**, 5025 (1996).

66. J. Ikeda, S. Kaneda, K. Kuwabara, S. Ogawa, T. Kobayashi, M. Matsumoto, T. Yura, and H. Yanagi, *Biochem. Biophys. Res. Commun.* **230**, 94 (1997).

67. M. Tucci, R.J. McDonald, R. Aaronson, K.K. Graven, and H.W. Farber, *Am. J. Physiol.* **271**, L341 (1996).

68. K.K. Graven and H.W. Farber, *New Horiz.* **3**, 208 (1995).

69. K.K. Graven, R.J. McDonald, and H.W. Farber, *Am. J. Phys.* **274**, C347 (1998).

70. I.J. Benjamin, B. Kroger, and R.S. Williams, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6263 (1990).

71. R. Mestril, S. Chi, M.R. Sayen, and W.H. Dillmann, *Biochem. J.* **298**, 561 (1994).

72. L.E. Huang, H. Zhang, S.W. Bae, and A.Y.C. Liu, *J. Biol. Chem.* **269**, 30718 (1994).

73. W.M. Miller, C.R. Wilke, and H.W. Blanch, *Bioprocess Eng.* **3**, 103 (1988).

74. D.F. Wilson, *Experientia* (*Basil*) **46**, 1160 (1990).

75. C. Zupke, A.J. Sinskey, and G. Stephanopoulos, *Appl. Microbiol. Biotechnol.* **44**, 27 (1995).

76. P. van der Valk, J.J.P. Gille, L.H.W. van der Plas, J.F. Jongkind, A. Verkerk, A.W.T. Konings, and H. Joenje, *Free Radical Biol. Med.* **4**, 345 (1988).

77. A.L. Sagone, Jr., *Am. J. Hematol.* **18**, 269 (1985).

78. P.W. Hochachka, L.T. Buck, C.J. Doll, and S.C. Land, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9493 (1996).

79. E.O. Pettersen, N.O. Juul, and O.W. Ronning, *Cancer Res.* **46**, 4346 (1986).

80. S. Tinton, Q. Tran-Nguyen, and P. Buc-Calderon, *J. Biochem.* **249**, 121 (1997).

81. S.M. Kraggerud, J.A. Sandvik, and E.O. Pettersen, *Anticancer Res.* **15**, 683 (1995).

82. T. Ogawa, M. Kamihira, H. Yoshida, S. Iijima, and T. Kobayashi, *J. Ferment. Bioeng.* **74**, 372 (1992).

83. S.S. Ozturk and B.Ø. Palsson, *Biotechnol. Prog.* **7**, 481 (1991).

84. S. Mercille and B. Massie, *Cytotechnology* **15**, 117 (1994).

85. Y. Shi, D.D.Y. Ryu, and S.H. Park, *Biotechnol. Bioeng.* **42**, 430 (1993).

86. W. Chotigeat, Y. Watanapokasin, S. Mahler, and P.P. Gray, *Cytotechnology* **15**, 217 (1994).

87. J. Wang, M. Hata, Y.S. Park, S. Iijima, and T. Kobayashi, *J. Ferment. Bioeng.* **78**, 321 (1994).

88. E. Regoeczi, M. Kay, P.A. Chindemi, O. Zaimi, and K.L. Suyama, *Biochem. Cell Biol.* **69**, 239 (1991).

89. I.N. Rich, *Exp. Hematol.* **14**, 746 (1986).

90. R. Pennathur-Das and L. Levitt, *Blood* **69**, 899 (1987).

91. M.R. Koller, J.G. Bender, E.T. Papoutsakis, and W.M. Miller, *Blood* **80**, 403 (1992).

92. V. Falanga, H. Takagi, P.I. Ceballos, and J.B. Pardes, *Exp. Cell Res.* **213**, 80 (1994).

93. T. Hirose, H. Terajima, A. Yamauchi, K. Kinoshita, K. Furuke, T. Gomi, Y. Kawai, S. Tsuyuki, Y. Nakamura, I. Ikai, T. Taniguchi, T. Inamoto, and Y. Yamaoka, *J. Hepatol.* **27**, 1081 (1997).

94. M.A. Goldberg, S.P. Dunning, and H.F. Bunn, *Science* **242**, 1412 (1988).

95. H.F. Bunn and R.O. Poyton, *Physiol. Rev.* **76**, 839 (1996).

96. J.M. Gleadle, B.L. Ebert, J.D. Firth, and P.J. Ratcliffe, *Am. J. Physiol.* **268**, C1362 (1995).

97. V.T. HO and H.F. Bunn, *Biochem. Biophys. Res. Commun.* **223**, 175 (1996).

98. G.L. Wang and G.L. Semenza, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4304 (1993).

99. I.J. Rondon, S.B. Scandurro, R.B. Wilson, and B.S. Beckman, *FEBS Lett.* **359**, 267 (1995).

100. M. Goldberg, C.C. Gaut, and H.F. Bunn, *Blood* **77**, 271 (1991).

101. G.L. Wang and G.L. Semenza, *Blood* **82**, 3610 (1993).

102. B. Jiang, G.L. Semenza, C. Bauer, and H.H. Marti, *Am. J. Physiol.* **271**, C1172 (1996).

103. S.L. Kroll and M.F. Czyzyk-Krzeska, *Am. J. Physiol.* **274**, C167 (1998).

104. M.F. Czyzyk-Krzeska, B.A. Furnari, E.E. Lawson, and D.E. Millhorn, *J. Biol. Chem.* **269**, 760 (1994).

105. M.L. Norris and D.E. Millhorn, *J. Biol. Chem.* **270**, 23774 (1995).

106. A. Rolfs, I. Kvietikova, M. Gassmann, and R.H. Wenger, *J. Biol. Chem.* **272**, 20055 (1997).

107. G. Melillo, T. Musso, A. Sica, L.S. Taylor, G.W. Cox, and L. Varesio, *J. Exp. Med.* **182**, 1683 (1995).

108. P.J. Lee, B.-H. Jiang, B.Y. Chin, N.V. Iyer, J. Alam, G.L. Semenza, and A.M.K. Choi, *J. Biol. Chem.* **272**, 5275 (1997).

109. L.S. Terada, D.M. Guidot, J.A. Leff, I.R. Willingham, M.E. Hanley, D. Piermattei, and J.E. Repine, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 3362 (1992).

110. S.D. Estes, D.L. Stoler, and G.R. Anderson, *J. Virol.* **69**, 6335 (1995).

111. G.L. Wang, B. Jiang, E.A. Rue, and G.L. Semenza, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5510 (1995).

112. A. Naldini, F. Carraro, S. Silvestri, and V. Bocci, *J. Cell. Physiol.* **173**, 335 (1997).

113. L.D. Longo, S. Packianathan, J.A. McQueary, R.B. Stagg, C.V. Byus, and C.D. Chain, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 692 (1993).

114. M.A. Goldberg and T.J. Schneider, *J. Biol. Chem.* **269**, 4359 (1994).

115. A. Minchenko, T. Bauer, S. Salceda, and J. Caro, *Lab. Invest.* **71**, 374 (1994).

116. Y. Liu, S.R. Cox, T. Morita, and S. Kourembanas, *Circ. Res.* **77**, 638 (1995).

117. K.P. Claffey, S. Shih, A. Mullen, S. Dziennis, J.L. Cusik, K.R. Abrams, S.W. Lee, and M. Detmar, *Mol. Biol. Cell* **9**, 469 (1998).

118. J.A. Forsythe, B.-H. Jiang, N.V. Iyer, F. Agani, S.W. Leung, R.D. Koos, and G.L. Semenza, *Mol. Cell. Biol.* **16**, 4604 (1996).

119. S. Kourembanas, R.L. Hannan, and D.V. Faller, *J. Clin. Invest.* **86**, 670 (1990).

120. K. Kuwabara, S. Ogawa, M. Matsumoto, S. Koga, M. Clauss, D.J. Pinsky, P. Lyn, J. Leavy, L. Witte, J. Joseph-Silverstein, M.B. Furie, G. Torcia, F. Cozzolino, T. Kamada, and D.M. Stern, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4606 (1995).

121. V. Falanga, S.W. Qian, D. Danielpour, M.H. Katz, A.B. Roberts, and M.B. Sporn, *J. Invest. Dermatol.* **97**, 634 (1991).

122. P. Ghezzi, C.A. Dinarello, M. Bianchi, M.E. Rosandich, J.E. Repine, and C.W. White, *Cytokine* **3**, 189 (1991).

123. S. Kourembanas, P.A. Marsden, L.P. McQuillan, and D.V. Faller, *J. Clin. Invest.* **88**, 1054 (1991).

124. S.-F. Yan, I. Tritto, D. Pinsky, H. Liao, J. Huang, G. Fuller, J. Brett, L. May, and D. Stern, *J. Biol. Chem.* **270**, 11463 (1995).

125. S.-F. Yan, Y.S. Zou, M. Mendelsohn, Y. Gao, Y. Naka, S.D. Yan, D. Pinsky, and D. Stern, *J. Biol. Chem.* **272**, 4287 (1997).

126. R.S. Shreeniwas, S. Koga, M. Karakurum, D. Pinsky, E. Kaiser, J. Brett, B.A. Wolitzky, C. Norton, J. Plocinski, W. Benjamin, D.K. Burns, A. Goldstein, and D. Stern, *J. Clin. Invest.* **90**, 2333 (1992).

127. M. Karakurum, R. Shreeniwas, J. Chen, D. Pinsky, S.-D. Yan, M. Anderson, K. Sunouchi, J. Major, T. Hamilton, K. Kuwabara, A. Rot, R. Nowygrod, and D. Stern, *J. Clin. Invest.* **93**, 1564–1570 (1994).

128. J.D. Loike, L. Cao, J. Brett, S. Ogawa, S.C. Silverstein, and D. Stern, *Am. J. Physiol.* **263**, C326 (1992).

129. B.L. Ebert, J.M. Gleadle, J.F. O'Rourke, S.M. Bartlett, J. Poulton, and P.J. Ratcliffe, *Biochem. J.* **313**, 809 (1996).

130. I. Stein, M. Neeman, D. Shweiki, A. Itin, and E. Keshet, *Mol. Cell. Biol.* **15**, 5363 (1995).

131. B.L. Ebert, J.D. Firth, and P.J. Ratcliffe, *J. Biol. Chem.* **270**, 29083 (1995).

132. J.D. Firth, B.L. Ebert, C.W. Pugh, and P.J. Ratcliffe, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6496 (1994).

133. J.D. Firth, B.L. Ebert, and P.J. Ratcliffe, *J. Biol. Chem.* **270**, 21021 (1995).

134. H.H. Marti, H.H. Jung, J. Pfeilschifter, and C. Bauer, *Eur. J. Physiol.* **429**, 216 (1994).

135. G.L. Semenza, B.-H. Jiang, S.W. Leung, R. Passantino, J.-P. Concorde, P. Maire, and A. Giallongo, *J. Biol. Chem.* **271**, 32525 (1996).

136. G.L. Semenza, P.H. Roth, H.-M. Fang, and G.L. Wang, *J. Biol. Chem.* **269**, 23757 (1994).

137. D. Wolfle and K. Jungermann, *Eur. J. Biochem.* **151**, 299 (1985).

138. K.A. Webster, *Mol. Cell. Biochem.* **77**, 19 (1987).

139. T. Kietzmann, H. Schmidt, K. Unthan-Fechner, I. Probst, and K. Jungemann, *Biochem. Biophys. Res. Commun.* **2**, 792 (1993).

140. T. Kietzmann, H. Schmidt, I. Probst, and K. Jungermann, *FEBS Lett.* **311**, 251 (1992).

141. M. Nauck, D. Wolfle, N. Katz, and K. Jungermann, *Eur. J. Biochem.* **119**, 657 (1981).

142. J.R. López-López and C. González, *FEBS Lett.* **299**, 251 (1992).

143. A. Gorlach, G. Holtermann, W. Jelkmann, J.T. Hancock, S.A. Jones, O.T.G. Jones, and H. Acker, *Biochem. J.* **290**, 771 (1993).

144. A.R. Cross, L. Henderson, O.T.G. Jones, M.A. Delpiano, J. Hentschel, and H. Acker, *Biochem. J.* **272**, 743 (1990).

145. J. Fandrey, S. Frede, and W. Jelkmann, in M.C. Hogan, ed., *Oxygen Transport to Tissue,* vol. XVI, Plenum Press, New York, 1994, p. 591.

146. J. Fandrey, S. Frede, and W. Jelkmann, *Biochem. J.* **303**, 507 (1994).

147. R.H. Wenger, H.H. Marti, S. Schuerer-Maly, I. Kvietikova, C. Bauer, M. Gassmann, and F.E. Maly, *Blood* **87**, 756 (1996).

148. C.T. Taylor, S.J. Lisco, C.S. Awtrey, and S.P. Colgan, *J. Pharmacol Exp. Ther.* **284**, 568 (1998).

149. M.A. Gilles-Gonzalez, G. Gonzalez, M.F. Perutz, L. Kiger, M.C. Marden, and C. Poyart, *Biochemistry* **33**, 8067 (1994).

150. S.J. Chanock, J.E. Benna, R.M. Smith, and B.M. Babior, *J. Biol. Chem.* **269**, 24519 (1994).

151. A.C. Koong, E.Y. Chen, and A.J. Giaccia, *Cancer Res.* **54**, 1425 (1994).

152. D. Mukhopadhyay, L. Tsiokas, X.-M. Zhou, D. Foster, J.S. Brugge, and V.P. Sukhatme, *Nature* (*Lond.*) **375**, 577 (1995).

153. A.P. Levy, N.S. Levy, and M.A. Goldberg, *J. Biol. Chem.* **271**, 2746 (1996).

154. G.L. Wang, B. Jiang, and G.L. Semenza, *Biochem. Biophys. Res. Commun.* **216**, 669 (1995).

155. C. Abate, L. Patel, F.J. Rauscher, III, and T. Curran, *Science* **249**, 1157 (1990).

156. H. Schenk, M. Klein, W. Erdbrugger, W. Droge, and O.K. Schulze, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1672 (1994).

157. M. Meyer, R. Schreck, and P.A. Baeuerle, *EMBO J.* **12**, 2005 (1993).

158. R. Schreck, P. Rieber, and P.A. Baeuerle, *EMBO J.* **10**, 2247 (1991).

159. M.B. Toledano and W.J. Leonard, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 4328 (1991).

160. P. Pognonec, H. Kato, and R.G. Roeder, *J. Biol. Chem.* **267**, 24563 (1992).

161. R.H. Wenger and M. Gassmann, *Biol. Chem.* **378**, 609 (1997).

162. K. Guillemin and M.A. Krasnow, *Cell* **89**, 9 (1997).

163. N.V. Iyer, L.E. Kotch, F. Agani, S.W. Leung, E. Laughner, R.H. Wenger, M. Gassmann, J.D. Gearhart, A.M. Lawler, A.Y. Yu, and G.L. Semenza, *Genes Dev.* **12**, 149 (1998).

164. G.L. Wang, B.-H. Jiang, and G.L. Semenza, *Biochem. Biophys. Res. Commun.* **212**, 550 (1995).

165. L.E. Huang, Z. Arany, D.M. Livingston, and H.F. Bunn, *J. Biol. Chem.* **271**, 32253 (1996).

166. P.J. Kallio, I. Pongratz, K. Gradin, J. McGuire, and L. Poellinger, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5667 (1997).

167. J. Fandrey, *Respir. Physiol.* **101**, 1 (1995).

168. M.A. Goldberg, G.A. Glass, J.M. Cunningham, and H.F. Bunn, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7972 (1987).

169. D.L. Galson, T. Tsuchiya, D.S. Tendler, L.E. Huang, Y. Ren, T. Ogura, and H.F. Bunn, *Mol. Cell. Biol.* **15**, 2135 (1995).

170. V.A. Ho, A. Acquaviva, E. Duh, and H.F. Bunn, *J. Biol. Chem.* **270**, 10084 (1995).

171. K.L. Blanchard, A.M. Acquaviva, D.L. Galson, and H.F. Bunn, *Mol. Cell. Biol.* **12**, 5373 (1992).

172. Z. Arany, L.E. Huang, R. Eckner, S. Bhattacharya, C. Jiang, M.A. Goldberg, and H.F. Bunn, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12969 (1996).

173. I. Kvietikova, R.H. Wenger, H.H. Marti, and M. Gassman, *Nucleic Acids Res.* **23**, 4550 (1995).

174. P.H. Maxwell, M.K. Osmond, C.W. Pugh, and P.J. Ratcliffe, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 2423 (1993).

175. C.W. Pugh, B.L. Ebert, O. Ebrahim, and P.J. Ratcliffe, *Biochim. Biophys. Acta* **1217**, 297 (1994).

176. L.H. Fossom, C.R. Sterling, and A.W. Tank, *Mol. Pharmacol.* **42**, 898 (1992).

177. A.P. Levy, N.S. Levy, S. Wegner, and M.A. Goldberg, *J. Biol. Chem.* **270**, 13333 (1995).

178. A.S. Minchenko, S. Salceda, T. Bauer, and J. Caro, *Cell. Mol. Biol. Res.* **40**, 35 (1994).

179. N. Bashan, E. Burdett, H.S. Hundal, and A. Klip, *Am. J. Physiol.* **262**, C682 (1992).

180. M. Mueckler, *Eur. J. Biochem.* **219**, 713 (1994).

181. I.J. Rondon, L.A. Macmillan, B.S. Beckman, M.A. Goldberg, T. Schneider, H.F. Bunn, and J. Malter, *J. Biol. Chem.* **266**, 16594 (1991).

182. C. Chen and A. Shyu, *Trends Biol. Sci.* **20**, 465 (1995).

See also MASS TRANSFER.

INOCULUM PREPARATION

CRAIG J.L. GERSHATER
S. B. Pharmaceuticals
Harlow, Essex
England

KEY WORDS

Culture assessment
Fermentation development
Inoculum development
Microbial inoculum
Seed stage fermentation
Stock culture

OUTLINE

INTRODUCTION

Inoculum preparation is the part of the fermentation development process that secures the microbial genotype for experimental and/or production purposes by providing a viable biomass capable of high productivity. Microbial processes generally consist of two distinct elements: a biotic factor the biomass and an abiotic factor (the growth environment). Hence, in applied microbiology terms

$$\text{pheotype} = \text{genotype} + \text{environment}$$

where the phenotype represents the desired attribute (accumulation of biomass or a valuable product of that biomass), the genotype represents the collection of metabolically useful genes and associated gene products, and the environment provides a system (principally a fermenter) of control over gene expression, that is, growth and product formation. At some point in the fermentation development program, the biotic and abiotic elements must be linked; this will be defined as the supply of consistent biomass to the fermenter plus growth medium. The growth and productive phases of fermentation are, of course, the subject of intense investigatory work; much has been written about the strict control requirements of the fermenter system, including the response of the microorganism to the nutritional and environmental factors. However, the requirements for the inoculum and the influence the inoculum can have on the outcome of the fermentation has received only scant attention (in terms of literature citations). A detailed assessment of the effect of the inoculum on microbial productivity can be justified, particularly in the case of nonrecombinant microorganisms, on the basis that commercially useful microbes will be isolated directly or indirectly from an ecosystem very remote from the fermentation laboratory, where the influences on growth are unknown.

Microbial Ecology

The microbial metabolic attribute to be exploited has evolved as a result of environmental influences unknown to the fermentation scientist. There is a case to be made for the basic ecology of the microorganism to be understood prior to culture preservation and inoculum preparation. A simple classification of organisms isolated from soil provides a loose context against which culture maintenance and development may be defined; soil organisms may be divided into two basic categories (1,2), the attributes of which are summarized in the following:

1. Indigenous, or autochthonous species, which have lower reproductive rates and are specialized for a particular habitat or more complex/intractable substrate. These are known as K strategists and include ascomycetes, basidiomycetes, and some actinomycetes.
2. Invaders, or allochthonous species, which have high reproductive rates and tend to function as pioneer species utilizing simple soluble substrates. The zymogenous fungi within this group include the sugar-fungi *Mucor, Rhizopus, Trichoderma, Penicillium*, and *Aspergillus*; organisms such as these may be activated by the presence of fresh substrate (breaking spore dormancy). These types of organisms, sometimes termed r strategists, possess good competitive abilities, which (along with some K strategists) include the ability to produce toxins (antibiotics), which is of obvious commercial interest.

Cognizance of the background ecology of the microorganism may permit a more appropriate culture storage and recovery method to be devised to maximize the chances of the organism successfully colonizing the next stage in the fermentation process.

Inoculum Development Conundrum

Process optimization must endeavor to replicate those factors responsible for the expression of the desired attribute in the totally artificial environment of monoculture; nutrient sufficiency (feast); and relatively stable physical parameters, such as temperature, pH, and oxygen availability. Although relatively few studies on the influence of inoculum on final stage fermentations have been made, it is certain that the inoculum preparation stage of most fermentations will affect the expression of the desired phenotype, whether qualitatively, quantitatively, or temporally. Hence, inoculum development is faced with a fundamental problem; the factors tested in seed stage or inoculum preparation investigations can only be assessed retrospectively in various high-productivity, final stage fermentations, remote from the seed stage influence. The conclusions drawn from inoculum development studies must deconvolute effects due to seed from those due to final stage fermentations. This article therefore seeks to explore how inoculum preparation can be incorporated into standard fermentation development programs.

Factors Critical to Inoculum Preparation

To adequately define the essential requirements for inoculum preparation, it is necessary to identify the elements that make up the process:

- Culture storage
- Stock culture
- Culture assessment
- Inoculum development
- Inoculation criteria

These elements will now be discussed and placed in the context of an overall inoculum strategy for fermentation development.

CULTURE STORAGE

Culture storage is only successful if the subsequent growth and productivity of the microorganism is consistent with expectations. Culture storage methodologies have been described in several authoritative texts (3–7) and are reviewed only briefly here.

Subculture

Depending on how cultures are received for fermentation development, some growth stage on agar will generally be required. The method of subculture relies on transferring colonies from one solid (depleted) medium to a fresh medium. This method is only suitable for short-term storage. Each transfer to fresh medium provides a comparatively nutrient replete environment, adaptation to which may select for an altered phenotypic expression. To minimize this possibility, minimal media are selected and the numbers of subcultures carried out are kept to a minimum. Minimal media agars are prepared in test tubes or screw-top Universal bottles, with the agars slanted to maximize surface area. Colonies from a master plate can be streaked onto the agar slants, allowed to grow for a minimum period, and then stored for a few weeks.

Desiccation

Removal or immobilization of water from the culture prevents metabolic activity, which is the key requirement of preservation. Drying methods are suitable for spore-forming fungi and streptomycetes; the principle is based on providing a solid, inert surface onto which spore suspensions can be dried. The methods differ in terms of the solid support that is used, which can include sterile soil, silica gel, dry sand, filter-paper disks, and various types of inert beads (glass, plastic, or porcelain). A suspension of spores (or possibly vegetative culture) is added to the sterile inert support, and after aseptic stoppering of the tube or vessel the mixture is allowed to dry at 25 °C. This method provides a mechanism to store, typically, spore suspensions for perhaps months, but it is generally not recommended for long-term storage of process critical stock cultures.

Freeze-Drying (Lyophilization)

Freeze-drying removes water from a frozen culture by sublimation under a vacuum. It is suitable for the long-term preservation of many types of culture and is particularly suitable for the preservation of vegetatively grown cells. Bacteria, while frequently stored by lyophilization, may be susceptible to damage by random mutation (7,8). The benefits of freeze-drying cultures is that the lyophials are relatively easy to store, transport, and identify. The cultures may be stable for up to 10 years. The disadvantages include the need to purchase a relatively expensive freeze-drier and the risk of mutation occurring in stored bacteria, which may be minimized by storing the lyophials in the dark at 5 °C.

Freezing

Freezing the culture effectively deprives the cell of water and slows all metabolic and chemical activity. Intracellular water deprivation arises as a result of osmotic pressure induced at temperatures between −5 °C and −15 °C, the temperature at which ice forms in extracellular fluid but the intracellular space remains unfrozen (9). Under these circumstances there is solute enrichment of the extracellular fluid, and water leaves the cell by osmosis, freezes outside, and the volume fraction of the cell is thus reduced. The rate of freezing will affect the size distribution of the ice crystals and with slow cooling will tend to result in considerable water loss from the cell and large extracellular crystals; rapid freezing (−200 °C/min, obtained by plunging the sample in liquid nitrogen) will result in smaller extracellular ice crystals and some intracellular

ice crystal formation. It is necessary to determine the correct freezing rate for the particular cell under investigation. The properties and composition of the growth medium qualitatively affect the success of cryopreservation. Recovery from cryopreservation may also be critical due to an effect known as recrystallization, which refers to small crystals rearranging themselves into potentially more damaging larger crystals, leading to cell rupture. This may typically happen at temperatures between −100 °C and −20 °C. The potentially damaging effects are minimized by rapid thawing.

To minimize the general effect of ice crystal damage, high-viscosity cryopreservants are often added to the cell suspension (in the case of vegetative culture storage); these may be of two basic types, intracellular (membrane penetrating) or extracellular (nonmembrane penetrating). The intracellular agents used include dimethylsulfoxide (DMSO) and glycerol. Glycerol (added at 5–10% v/v to the final culture dispensing fluid) has the effect of reducing lesions due to osmotic shock by reducing the size of ice crystals. Extracellular cryopreservants include various carbohydrates, such as sucrose, mannitol, dextrin, and starch; 20% powdered skim milk and 12% sucrose are widely used.

The principal advantage of freezing as a method of culture storage is the relatively long duration of culture stability. The lower the temperature, the longer the duration. At −20 °C it is reasonable to consider that microbial cultures will remain stable for 1 to 2 years; at −80 °C, cultures should remain stable for 5 years or more, indefinite storage is obtained at temperatures of less than −80 °C using liquid nitrogen. The vapor phase of liquid nitrogen is −156 °C, and the liquid phase is −196 °C; extreme care must be exercised when recovering culture vials from the liquid phase, because leakage of liquid nitrogen into vials may result in explosion of the sealed ampoules due to gas expansion at elevated temperatures. The main disadvantages of freezing for culture storage are the relatively high capital costs for compressor freezers (particularly those for less than −80 °C) and secure purchasing and management procedures for maintaining liquid nitrogen stocks and reliability of electrical supplies. In addition, frozen ampoules are not easily transported.

Fungal culture storage may require a combination of all the methods described earlier for long-term storage, dependent on the morphological characteristics of the strain in question; for example, aconidial strains may be impossible to preserve by lyophilization (10), but may be stored on silica gel under liquid nitrogen. Methods appropriate to bacterial culture storage using both drying and liquid nitrogen have been reviewed (11).

STOCK CULTURE

A fermentation development or production process involving the culturing of commercially significant microorganisms demands that a robust system be in place for stock culture generation and maintenance. Variability of stock culture performance cannot be tolerated because culture variation masks the influence of nutritional or environmental factors within a development program.

What Is a Stock Culture?

The stock culture represents the current microbial investment in a fermentation program, which is manifested in both qualitative and quantitative terms. Qualitatively, the stock culture defines the current genotypic status of a commercially significant microorganism (i.e., its product formation capacity in a defined environment). Quantitatively, the stock culture is represented by a carefully controlled holding of cultivar, which will be drawn on for development and production purposes. At some critical point this holding must be replenished with equivalent stock culture or be replaced by an improved stock. Awareness of the microbial ecology of the culture may be important given that the cultivar will be required to propagate within a new environment (the ecological niche represented by fresh fermentation medium) by rapid colonization. If, for example, the culture of interest is considered to be an allochthonous species (r strategist) capable of colonizing fresh medium from spore inoculum, then the stock culture should be evaluated as a spore-based stock culture.

Management of the microbial culture resources is the most important facet of a microbiological bioprocess enterprise. Each bioprocess organization defines its own protocols to secure the microbial investment; however, the principal objective must be to ensure that the bioprocess enterprise is never placed at risk due to stock culture mismanagement.

Stock Culture Generation

At various stages in a microbial development program it will be necessary to generate stock culture. The method chosen is dictated in part by the preservation protocol appropriate for the fermentation program. In my laboratory, with a relatively high turnover in new cultures for fermentation development, vegetative stock cultures are often generated and stored at −80 °C. The stock generation model for cryopreservation is shown in Figure 1.

The model in Figure 1 assumes that the culture to be laid down to stock has been evaluated or in some other way nominated for long-term storage. Hence, prior to long-term culture storage, working stocks of the same culture will have been available. Working stock cultures may similarly be prepared using the model described earlier, but perhaps with far fewer vials, or they may be preserved using one of the other methods described previously.

The generation of fresh stocks may be dictated by several factors that may or may not be interrelated. As the stock culture holding is depleted, it must be replenished with fresh stocks with performance criteria as close to the existing stocks as possible. Stock culture management protocols determine at what point this should be; for critical bioprocesses (main line production) this may be initiated when 60% to 70% of the existing stock has been utilized. With less critical bioprocesses, stock culture replenishment may commence when 70% to 80% of the existing stock is depleted. Fresh stocks are also required when there is qualitative change in the culture component of the bioprocess. This may result from a deliberate attempt to improve the culture by strain improvement or culture reselection; alternatively, it may result from an inexplicable,

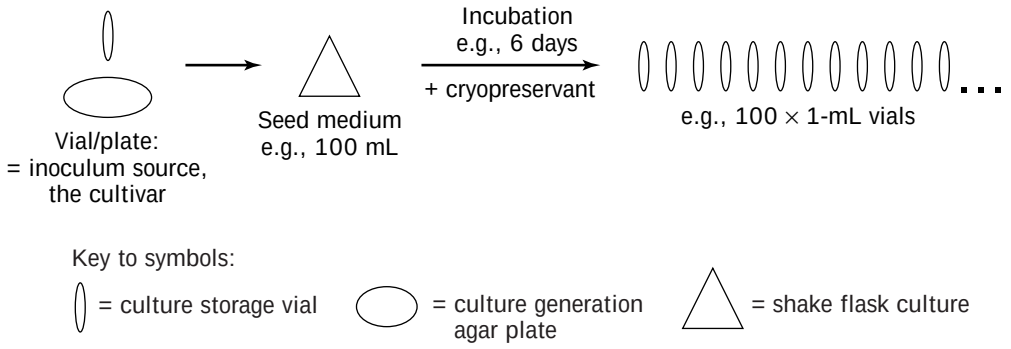


Figure 1. Stock culture generation models. Inoculum source for the first-stock generation is usually culture grown on agar plates; this is typically a plug (agar plugs are aseptically cut through the colony center) inoculated to liquid seed medium. The seed medium provenance, selected for vegetative growth, resides within relevant literature or, in the case of screening derived cultures, from original cultivation protocols. The incubation conditions for the culture are determined from the literature or from preexisting protocols. Dispensing of adequately grown vegetative culture to long-term culture vials is carried out under strict aseptic techniques; the number of vials dispensed to must be restricted to prevent the heterogeneous distribution of culture between first and last vials in the series.

irreversible deterioration in culture performance. If fresh stocks are to be generated from an improved strain, then the working stocks evaluated as part of the strain improvement process serve as the progenitor to the stock culture generation model previously described. If, however, the need to generate new stocks arises from severe culture deterioration, then a previous generation stock must be used for current stock culture generation. This deterioration may be accounted for by a reduction in culture viability, culture productivity performance, or microbial or viral contamination.

As new stock cultures are prepared, a family tree of related strains will be generated. To replenish stocks from an existing stock culture, a vial of the parental working stock can be used to generate a new daughter stock. Within the stock culture model each time a new generation stock culture is prepared after testing, a portion of that stock holding must be set aside for permanent retention. In my laboratory, at each stage in the stock maintenance program a portion of the existing stock is nominated as the master or "crown–jewel" stock and is held inviolable for process operations but represents a temporal marker in the history of a culture development program. To keep track of culture usage and family history, a suitable nomenclature must be devised. We use a simple decimal notation system in which the integer before the decimal point represents the current generation number and the figure after the decimal point represents the sibling stock number for the current generation.

Eventually even master or crown-jewel stocks may become depleted or obsolete, but within the active lifetime of a bioprocess, the master stock can only be drawn on in the event of a catastrophic failure of the main working stocks due to either an inexplicable loss of viability or performance or mechanical failure of the entire culture preservation system. To provide additional protection to the culture storage facility, a decision may be made to preserve split stocks not just in separate freezers but also on different sites, if possible. This type of stock culture management can be determined within an organization's disaster planning protocols.

CULTURE ASSESSMENT

The generation of new stocks imposes a profound influence on the bioprocess, and performance criteria must be reestablished at each stage in the culture maintenance program. As with any biological system, the inherent variability of the culture must be assessed prior to any development or production processes being committed to the new stock culture. The elements of culture maintenance that must be in place before an assessment of culture variability can be carried out include the following:

- Adequate numbers of culture vials (previously found to be monoculture)
- A suitable seed culture fermentation protocol
- A final or production phase fermentation protocol
- Suitable response variables to the final stage fermentation
- Validated assays for quantitatively determining the levels of the response variables
- Full supporting documentation to all of the elements mentioned

Culture variability must be addressed directly by the use of suitable statistical evaluation procedures. These procedures are largely based on a system of replication at each stage of the evaluation so as to obtain sufficient data describing the spread of variability and the contribution to total variation by each critical stage in the evaluation protocol. If the culture is being generated and evaluated as part of a good management practice (GMP) process, the stringency of these requirements may be greater. Figure 2

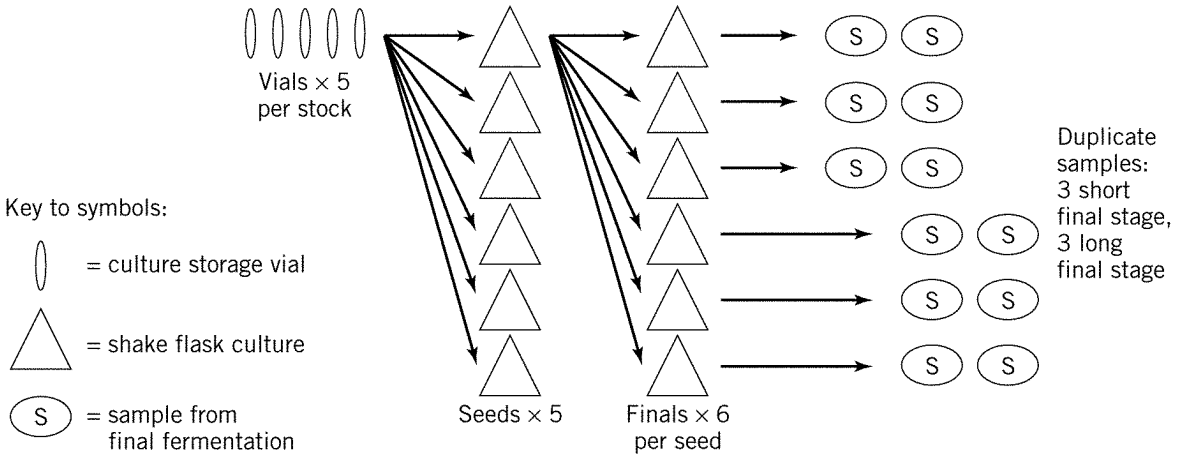


Figure 2. Culture variability testing model. To ensure adequate data is generated for a statistical analysis of variance to be undertaken, appropriate numbers of elements at each stage in the procedure must be selected. In this (model) case, five vials of stock culture are each used to inoculate each of five seed-media shake flasks. Each seed culture, after incubation to generate adequate biomass, is used to inoculate each of six final stage fermentations and will be assessed for reproducibility of growth. Half of the final stage shake flasks (three fermentation flasks) are incubated for a relatively short time and then duplicate sampled; the remaining half are incubated for a relatively longer time and then duplicate sampled. Each sample can be the source of duplicate or triplicate assays to enhance the analysis of variance (not shown in the model).

broadly defines the type of testing that may be carried out using a newly generated or existing stock.

This model permits replication sufficient to explore the contribution to the total variation of each of the stages capable of being a source of variation. This is sometimes known as a nested design, and data arising are analyzed using standard analysis of variance (ANOVA) procedures. Not only are data generated that capture the inherent variation associated with vial-to-vial variation, but the relative robustness of the remaining system can also be evaluated. Data from this kind of trial may be summarized in the form of statistical tables, but perhaps of more use would be data summarized using graphical plots. Figure 3 represents the actual data obtained for a reproducibility trial exploring the vial-to-vial variability of a fungal culture producing an antibiotic.

INOCULUM DEVELOPMENT

As stated earlier, the key to successful inoculum preparation lies in the outcome in the final stage of the fermentation, hence, evaluation of seed protocols is always retrospective. The factors that can influence the inoculum and therefore the final-stage productivity are multifarious. This chapter has already identified certain key functions that may or may not influence the development of a successful seed culture, such as, microbial ecology, culture storage, and variability. In addition, some of the factors that may need to be considered in developing a inoculum preparation strategy include:

- Seed medium
- Seed incubation conditions
- Microbial growth dynamics
- Optimum transfer criteria (qualitative inoculum variable)
- Volume of inoculum on transfer (biomass concentration, quantitative inoculum variable)
- Number of seed-stage passages
- Assessing inoculum development

Several of these influences have been investigated in the past, but often in isolation from one another. The recognition that inoculum development can play a key role in fermentation development programs was recognized over 20 years ago; in a remark attributed to Hockenhull but reported by Calam (12), he stated that “once a fermentation had started it can be made worse but not better.” Despite this recognition, the bioprocess industry (high-added-value products), appears to have been very relaxed about setting resources aside to investigate the influence of inoculum on final stage fermentation productivities. This is perhaps somewhat in contrast to the fermentation food industries, where starter culture methods have been developed to provide high quality and consistent product.

In the dairy industry, factors such as correct carbon and nitrogen sources, vitamin and mineral supplements, and the addition of phage inhibitory agents, antioxidants, and neutralizers have all been considered important in defining the correct formulation for bulk starter media for lactic acid bacteria (13) in cheese production.

In the case of alcohol fermentation, relatively little work appears to have been done to investigate the influence of inoculum on alcohol yields. When reported, the studies are often limited to just one aspect of the inoculum process, for example, that inoculum levels may vary greatly with *Sac-*

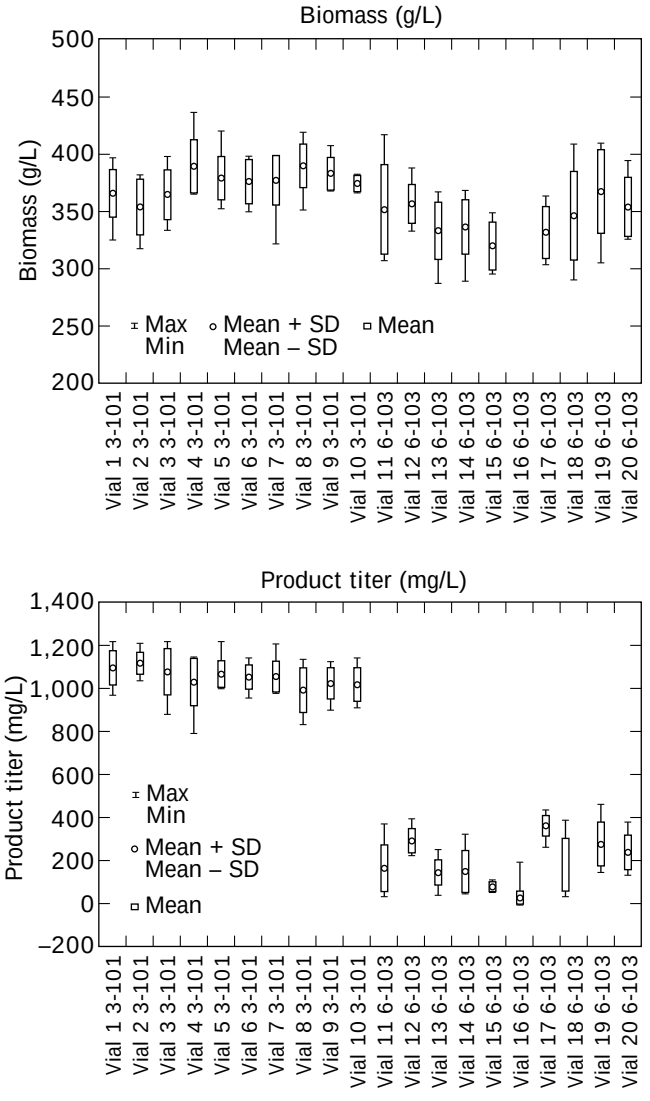


Figure 3. Reproducibility data from an antibiotic-producing culture. A fungal culture subject to a fermentation development program generated a number of reisolates that were tested in a reproducibility trial similar to one described in the model (Fig. 2). The two reisolates were called 3-101 and 6-103. Preliminary testing in single shake-flask experiments had indicated that although 3-101 was superior to 6-103, the difference was not great, therefore, working stocks of both cultures were prepared. In terms of biomass the two cultures appear very similar, although the variability of culture 6-103 was greater than that of 3-101. Biomass data was unavailable for vial 16, but this did not prevent an ANOVA of the data. In terms of product titer, 3-101 was clearly far superior to 6-103, both in terms of overall titre as well as vial-to-vial variation. The ANOVA for these experiments revealed that vial-to-vial variation was the greatest contributor to variation of 6-103 performance, and that factor tended to mask the true culture productivity.

charomyces cerevisiae, at between 0.05 and 60% (14). More recent studies with yeast fermentations have attempted to define the effect of inoculum on growth and ethanol production (15), including investigating at which point in the microbial growth curve inoculation should take place—in this case exponentially growing cells were favored. In addition, these studies were carried out while keeping factors such as medium formulation, medium sterilization, and culture growth conditions all constant, while varying inoculum size, under these circumstances inoculum levels of 10% were found to generate maximum biomass and, hence, ethanol.

Fermentations devoted to solvent formation have also been investigated to determine the influence of inoculum. The production of acetone, butanol, and ethanol by anaerobic fermentation using *Clostridium butylicum* was found to be adversely affected by repeated subculturing. This can be partly ameliorated by incorporating heat shocking and subculturing into an inoculum development program wherein the effect of heat shocking the culture to 75 °C for 1.5 min may reduce the degenerative effects of subculturing (16). Again, more recent studies have confirmed the deleterious effect of repeated subculture with saccharolytic clostridia (17). Using *Clostridium acetobutylicum* it was recognized that there would be an abrupt decrease in solventogenesis after 10–12 successive transfers incorporating heat shock; further, after 25 transfers, the culture failed to form spores, and after heat shock failed to grow at all. The dramatic decrease in solventogenesis on subculture was found to be associated with deficiencies in acetate/butyrate CoA transferase and acetoacetate decarboxylase activity. The addition of exogenous acetic and butyric acids maintained solventogenesis to at least the 40th transfer.

The influence of inoculum development for fungal cultures has been investigated. The production of γ-linolenic acid (GLA) by shake flask culture of the sporulation-deficient strain of *Cunninghamella echinulta* was found to be profoundly affected by the use of fragmented mycelia, increasing biomass and GLA titre (18).

Most investigations into the influence of inoculum involving fungal and filamentous bacterial cultures have been carried out for the production of various bioactive metabolites at the high-value-added end of the bioprocess industry, some of the issues of inoculum preparation here will be now considered in more detail.

Multiple Growth Cycles

The generation of inoculum for large-scale bioactive metabolite production by fermentation generally necessitates a buildup of active biomass through multiple growth cycles. This is illustrated in Figure 4 for a typical 3,000-L pilot-plant fermentation.

In developing the multiple-growth cycle seed train, several experimental objectives must be met. The first of these is the seed medium for culture growth. There are virtually no references to seed stage medium development in the literature (19,20), however, some general principles do apply. The seed medium formulation should seek to minimize nutrient shift effects while supporting good vegetative growth over a short time scale. The overriding proviso for seed medium formulation is that high productivity is supported in the final stage medium. In the case of bioactive metabolite fermentations, the seed stage fermentation should not result in production of the metabolite of inter-

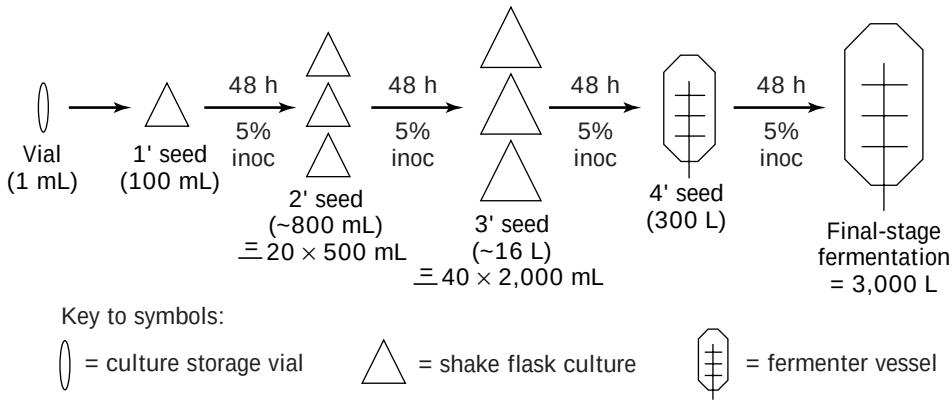


Figure 4. Typical multiple growth cycle fermentation seed train. The illustration indicates the logistics in preparing the inoculum for a 3,000-L final stage fermentation. In the example given, the seed train is defined by having a quaternary seed-stage inoculate the final-stage fermentation. The incubation times defined for each stage of the fermentation are designed to yield optimal inoculum to the next stage in the process. Throughout the seed train, the shock of culture transfer to nutrient feast is minimized by maintaining a relatively short residence in each seed fermentation, thus minimizing the risk of transferring from famine to feast at each stage. Inoculum transfer volumes impose logistic constraints on the seed train such that at the tertiary seed stage 40 2-L shake flasks must be inoculated. This represents a risk to the system in terms of contamination and culture heterogeneity, but small-scale fermenters can be used to overcome these problems.

est. Bioactive metabolites may be growth inhibitory to the producing strain (and hence are secondary metabolites), and their presence in the seed stage broth may complicate or inhibit the final stage product accretion. Sterilization control profiles for the seed stage medium, just as for the final stage medium, should ensure that Maillard reactions (browning) adversely affecting protein and glucose availability are minimized. Sterilization cycles should also be designed such that vitamin destruction is minimized and so the formulation of the seed medium will seek to prevent trace element precipitation (e.g., phosphate and magnesium combinations and iron complexes formed on heating or pH shifts).

The overall inoculation protocol seeks to define a quality standard for the inoculum. Earlier studies had identified that larger volumes of inocula of *Trichoderma reesei* (in the range 0.5 to 5.0% v/v) ended to reduce subsequent lag phases (21). These sorts of studies were extended (22) using *Penicillium chrysogenum* for penicillin production; it had been established that the spore concentration in shake-flask seed culture was critical to growth and antibiotic production. Above 10^4–10^5 spores/mL, growth of the production cultures was good, critical enzyme levels were high, and penicillin production was good. Below 10^3 spores/mL, the fermentations were suboptimal in every respect. In further studies (23), *Penicillium* spores were seeded at different concentrations into seed medium, and samples were removed at regular intervals and examined for growth on agar for numbers of viable centers. At some critical point in the fermentation, the number of viable growth centers fell by up to 85%, presumably due to aggregation and increased branching frequency in the hyphae. These observations have been confirmed by image-analysis work (24), which demonstrates a morphology shift from dense pellet formations to more dispersed forms with inocula above 10^5 spores/mL. These data tend to confirm the notion of a quality curve for inoculum development, as illustrated in Figure 5.

INOCULATION/SEED TRANSFER CRITERIA

Multistage inoculum transfers and the success criteria associated with those transfers must be determined experimentally. Each stage in the process can be examined individually, but at some point the seed train protocol must

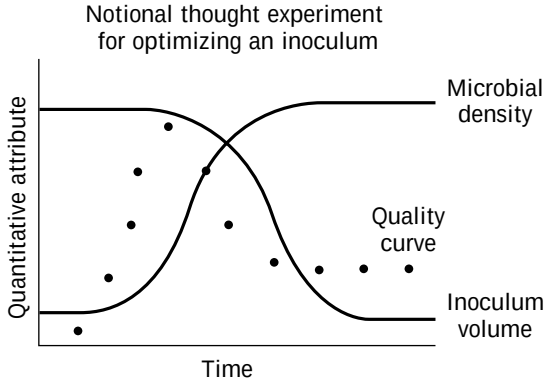


Figure 5. Notional quality curve for a seed stage fermentation. The thought experiment described suggests that there may be a compromise between inoculum volume and the qualitative attributes associated with the inoculum. Determining the notional quality curve requires an investigation into the optimization of the desired response variable in the subsequent fermentation stage. This may be achieved by carrying out timed transfers and measuring the physiological status of the organism at each time point.

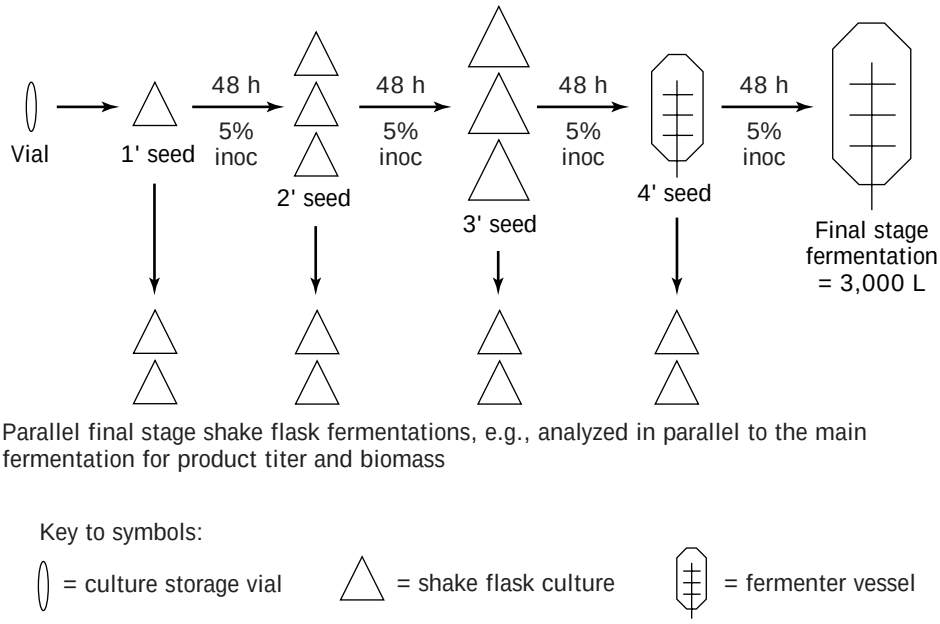


Figure 6. Inoculation–seed transfer experimental design. Final stage shake flask fermentations running in parallel to the main seed train, inoculated from seed stage broth generated in the inoculum train. Principal objective with this type of design is to compare productivity between each final stage fermentation to help determine if there are effects due to passage or some other seed incubation condition.

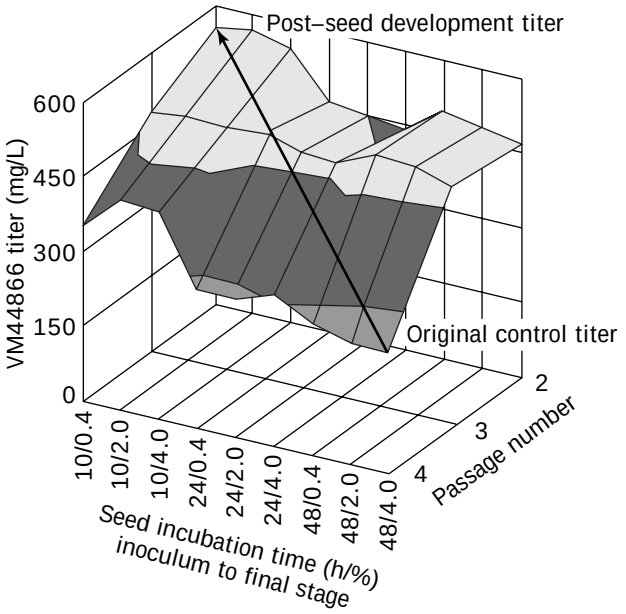


Figure 7. Response surface plot for milbemycin production under the influence of seed stage development factors. The response surface plot shows the final stage titers of milbemycin obtained under the seed inoculum conditions indicated by the three-dimensional plots. These data show that percentage inoculum levels into the final stages were not a key factor. However, passage number was significant, with quaternary seed stages giving rise to the lowest final stage titers. Similarly, juvenile seed was optimal for all conditions tested.

be established. As already indicated, seed stage assays may be expensive in terms of numbers and time required for full analysis, and care must exercised to ensure that seed stage data correlate with final stage fermentation productivity. The use of complex media may impose additional variation on a multistage process, and wherever additional variation is suspected, greater replication in the data-gathering phase will be required.

The principal physiological markers that can be used for seed transfer criteria include:

- Biomass density or an analyte reflecting that variable (ATP)
- Gross morphology attributes (particularly for filamentous bacteria and fungi)
- Respiratory gas analysis from aerated fermenters (carbon dioxide evolution rate)
- Macromolecular titre (DNA, RNA)
- Final stage productivity

Final stage productivity can only be determined by inoculating final-stage fermentations from each seed stage and analyzing for productivity at each stage. The format for this type of experiment is shown in Figure 6.

INOCULUM PREPARATION STUDIES

There are very few studies reported in the literature examining the complete inoculum preparation process from initial inoculation to final stage productivity. *Streptosporangium* cultures that are capable of producing useful secondary metabolites, including chloramphenicol (25), dihydrosinefungin (26), and endothelin-converting enzyme inhibitors (27), have been investigated for various inoculum stage conditions (28). The factors examined with respect to the inoculum stage of *Streptosporangium* cultures include inoculum storage, media sources, and inoculum size on two antibiotic fermentations. Three culture preservation methods were examined: spores, slant tubes, and frozen agar plugs. When each of these sources were tested in different media and analyzed for growth, slant transfer tubes were found to be the most effective for maintaining high growth concentrations in the final stage. Examining

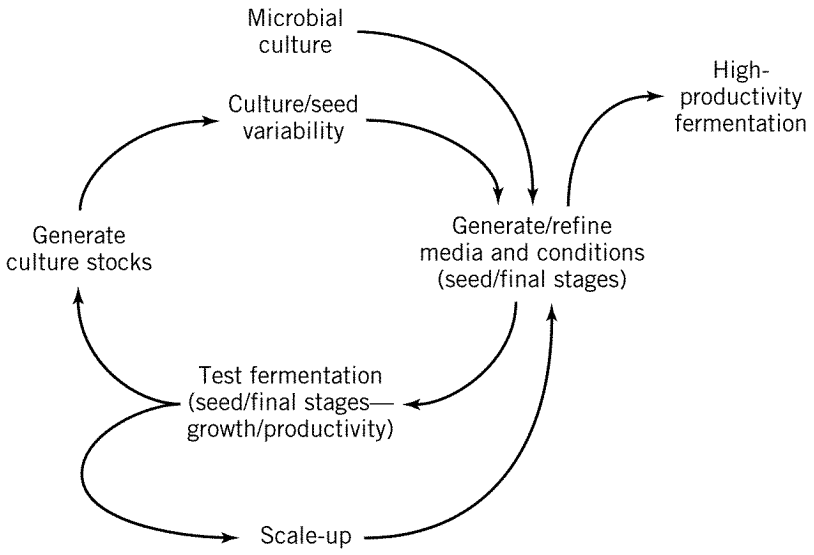


Figure 8. Fermentation development model incorporating seed stage investigations. The model shown assumes the process starts with receipt of the microbial culture. Seed and final stage media are identified initially. Fermentations are carried out to assess growth, and if adequate to the purpose, then culture stocks may be generated. These stocks must be tested for variability before further studies can commence. Seed and final stage fermentation conditions can then be refined with the newly generated and tested stock culture. The fermentation development process continues until performance criteria for the system are met, the process is scaled up, and if necessary further refined. Finally a high-productivity fermentation can be defined using a well-characterized and genotypically secure process.

incubation conditions with various *Streptosporangium* cultures indicated that aeration and growth phase of the inoculum may affect the time course of the final stage fermentation, particularly if that fermentation stage is relatively short. These influences may be less pronounced if longer final stage fermentations are required.

The improved production of milbemycin (a macrolide with anthelminthic properties) from *Streptomyces hygroscopicus* as a result of intensive seed stage development has been reported (29). In this work an attempt was made to examine several interacting factors affecting seed stage development and subsequent final stage production of milbemycin. Time-course fermentation profiles of the seed stage revealed that a major component of the seed medium, dextrin, was not utilized, and the other major carbohydrate component, glucose, could be reduced fourfold without affecting the productivity of the final stage fermentation. Aside from medium composition, the other seed development factors investigated included:

- Age of culture at transfer
- Temperature of seed-stage incubation
- Inoculum level
- Passage number of seed stages

Inoculum age was found to have a profound influence on final stage titers, with more juvenile seeds (10 h as compared to 48 h for the control) proving beneficial to milbemycin production. Inoculum levels could be halved from 4% v/v without loss of productivity. Passage number was found to be critical in seed media containing dextrin, as these seed fermentations showed a marked decrease in final stage titers in the tertiary and quaternary seed stages. Studies investigating the combined effects of seed incubation times, percentage inoculum into each seed stage or the final stage fermentations, and passage number were undertaken. Figure 7 shows the response surface plot summary of that work, where by examining the various seed stage factors, overall the titer of the final stage fermentation was increased fourfold.

INOCULUM DEVELOPMENT SUMMARY

This chapter has attempted to describe the various influences that may be examined when trying to define the optimal inoculum-development protocols. As indicated from the literature, this part of the fermentation process has been largely neglected, and yet the evidence indicates (from the relatively few studies undertaken) that inoculum preparation may have a profound effect on the outcome of a fermentation development process. Figure 8 describes how inoculum development may be incorporated into the overall fermentation development strategy for a bioprocess enterprise.

BIBLIOGRAPHY

1. R. Campbell, *Microbial Ecology*, 2nd ed., Blackwell Scientific Publications 1983, pp. 82, 90.
2. M. Alexander, *Introduction to Soil Microbiology*, 2nd ed., Wiley, New York, 1977, pp. 16–17.
3. L.T. Chang and R.P. Elander, in A.L. Demain and N.A. Solomon eds., *Manual of Industrial Microbiology and Biotechnology*, American Society of Microbiology, Washington, D.C., 1986, pp. 49–55.
4. J.J.S. Snell, in B.E. Kirsop and J.J.S. Snell eds., *Maintenance of Microorganisms: A Manual of Laboratory Methods*, Academic Press, Inc, London, 1984, pp. 11–21.
5. D. Smith and A.H.S. Onions, *The Preservation and Maintenance of Living Fungi*, Commonwealth Mycological Institute, Kew, U.K., 1983, pp. 5–30.
6. A. Dietz and B.W. Churchill, in M. Moo-Young ed., *Comprehensive Biotechnology: The Principles, Applications, and Regulations of Biotechnology in Industry, Agriculture, and Medicine*, vol. 1, Pergamon Press, Oxford, 1985, pp. 37–49.
7. A. Dietz, in H.-J. Rehm and G. Reed eds., *Biotechnology: A Comprehensive Treatise in 8 Volumes*, vol. 1, Verlag Chemie, Weinheim, 1981, pp. 411–434.
8. M.J. Ashwood-Smith and E. Grant, *Cryobiology* **13**, 206–213 (1976).
9. C. Herbert, P. Roux-Salembien, and J.-M. Egly, *Biofutur* **128**, 3–11 (1993).

10. C. Wilson, *Fungal Genet. Newsletter* **33**, 47–48 (1986).
11. I.J. Bousfield, in B.E. Kirsop, and J.J.S. Snell eds., *Maintenance of Microorganisms: A Manual of Laboratory Methods*, Academic Press, Inc, London, 1984, pp. 63–68.
12. C.T. Calam, *Process Biochem*. 7–12 (April 1976).
13. W.E. Whitehead, J.W. Ayers, and W.E. Sandine, *J. Dairy Sci.* **76**, 2344–2353 (1993).
14. P. Strehaino, M. Mota, and G. Goma, *Biotechnol. Lett.*, **5**, 135–140 (1983).
15. G. Birol and K. Ozergin-Ulgen, *Turk. J. Chem*. **19**, 250–257 (1995).
16. J.R. Gapes, V.F. Larsen, and I.S. Maddox, *J. Appl. Bact.* **55**, 363–265 (1983).
17. O. Assobhei, A. El Kanouni, M. Ismaili, M. Loutfi, and H. Petitdemange, *J. Ferment. Bioeng.* **85**, 209–212 (1998).
18. H.-C. Chen and T.-M. Liu, *Enz. Microbiol. Technol.* **21**, 137–142 (1997).
19. G.R. Hunt and R.W. Stieber, in A.L. Demain and N.A. Solomon eds., *Manual of Industrial Microbiology and Biotechnology*, American Society of Microbiology, Washington, D.C., 1986, pp. 32–40.
20. S. John Pirt, *Principles of Microbial and Cell Cultivation*, Blackwell Scientific, Oxford, 1975, pp. 118, 196–198.
21. D.E. Brown and M.A. ZainudeEn, *Biotechnol. Bioeng.* **20**, 1045–1061 (1978).
22. G.M. Smith and C.T. Calam, *Biotechnol. Lett.* **2**, 261–266 (1980).
23. C.T. Calam and G.M. Smith, *FEMS Microbiol. Lett.*, **10**, 231–234 (1981).
24. K.T. Tucker and C.R. Thomas, *Biotechnol. Lett.*, **14**, 1071–1074 (1992).
25. A. Tamura, I. Takeda, S. Naruto, and Y. Yoshimura, *J. Antibiot.* **24**, 270 (1971).
26. R. Cooper, P. Das, C. Federbush, R. Mierzwa, M. Patel, P. Birendra, and I. Truumees, *J. Indust. Microbiol.* **5**, 1–8 (1990).
27. Y. Tsurumi, N. Ohata, T. Iwamoto, S. Kyoto, and S. Okuhara, *J. Antibiot.* **47**, 619–630 (1994).
28. G. Platas, J. Collado, H. Martinez, M. Arrese, P. Pelaez, and M.T. Diez, *Microbiologia SEM* **13**, 193–200 (1997).
29. S.R.C. Warr, C.J.L. Gershater, and S.J. Box, *J. Indust. Microbiol.* **16**, 295–300 (1996).

INSECT CELL CULTURE, PROTEIN EXPRESSION

JAMES VAUGHN
ARS, USDA
Beltsville, Maryland

KEY WORDS

Baculoviruses
Expression vectors
Glycosylation
Insect cell lines
Insect glycoproteins
Nutrient requirements
Oxygen requirements
Protein processing
rDNA
Recombinant proteins

OUTLINE

INTRODUCTION

Biotechnology has made feasible the production of many complex proteins of importance in human or animal medicine in cultures of single cells, either microbial or animal. Animal cells (eukaryotic cells) can accomplish the post-translation modifications of these protein products that are essential for their biological function. For this reason, they are often preferred for production to the more easily grown microbial cells (prokaryotic cells), which cannot perform these modifications. Among the several types of eukaryotic cells available as hosts for gene expression, insect cells have some advantages that make them the host of choice. They can be subcultured indefinitely without being "transformed" eliminating the possible risk of the presence of cancer-promoting substances. Insect cells do not require a suitable substrate for attachment, but grow readily in suspension, making them easier to grow in commercial-scale reactors. They are not likely to contain viruses that would present a hazard to humans or other warm-blooded animals. The vector used to transfer the foreign gene to the cell, a baculovirus, is derived from a group of viruses that infect only arthropods and therefore have no risk for higher animals. This virus contains several very active promoters and so the yields of expressed protein can be very high, as much as 20–250 times that of mammalian cells. The baculovirus–insect cell system has been used to express approximately 500 gene products in insect cells, and 95% of the expressed recombinant proteins have been biologically active (1). Methods for scaling up production to 100-L bioreactors have been developed and successfully tested. In this article, the baculovirus system and the expression of heterologous genes in insect cells are reviewed with emphasis on optimizing the expression and scale-up of production systems. For detailed reviews of baculovirus genetics and the construction of vectors, readers are referred to articles by Frieson and Miller, (2) Luckow and Summers (3), and Kool and Vlak (4).

BACULOVIRUSES

Classification

The viruses used for the construction of expression vectors are pathogens of insects, and the early interest in them was their potential for the biocontrol of insect pests. They are classified in the family *Baculoviridae*. Baculoviruses have rod-shaped nucleocapsids containing a single molecule of circular, supercoiled, double-stranded DNA and are enveloped. The family contains two genera, the genus *Nucleopolyhedrovirus* (NPV) and the genus *Granulovirus* (GV) (5). Figure 1 illustrates the structure of typical members of each of these genera. At this time the expression vectors have been constructed using two species of the genus *Nucleopolyhedrovirus*, *Autographa californica* NPV, abbreviated AcMNPV, and *Bombyx mori* NPV (BmNPV). The NPV are characterized by the formation in the nucleus of infected cells of paracrystalline, proteinaceous occlusion bodies (OB) in which the virions are occluded. The occluded virion, known as the occlusion-derived virus (ODV), may contain a single nucleocapsid (SNPV) or multiple numbers (MNPV). This virion, protected from the environment by the OB, is the means by which the virus is spread horizontally in nature. A second virion phenotype is also produced during replication. This virion is primarily a single nuclear capsid in a loose-fitting envelope obtained as the nucleocapsid buds from the plasma membrane and is referred to as budded virus (BV) or extracellular virus (ECV) (Fig. 2). This envelope contains terminal peplomers consisting of a single viral glycoprotein inserted into the plasma membrane during viral replication. Dispersion of the virus among internal body tissues is accomplished by the virions of this second phenotype, which is the phenotype used to infect cells in culture.

Members of this family of viruses have never been found in animals outside the Arthropoda. Within the arthropods, these viruses predominantly infect insects and within the insects the majority have been found in Lepidoptera, the moths and butterflies. Other than insects, the virus has been found in diseased shrimp, prawns, and crabs. The host range of a virus is usually limited to closely related species within a genus or a family. However, AcMNPV is the most notable exception in that it has been reported to infect species in 10 families (6).

Replication

Infection begins when a permissive host consumes food contaminated with the viral occlusion bodies. The proteinaceous bodies are dissolved by the alkaline contents of the insect midgut, and the released virions invade the cells of the midgut epithelium by fusion of the viral envelops to the plasma membrane of the cells. Following fusion, the released nucleocapsids move through the cytoplasm to the cell nucleus and gain entry through the nuclear pores (7,8). Inside the nucleus, the nucleocapsids are uncoated and replication begins.

The synthesis of new viral proteins occurs in a cascade-like manner. In the baculoviruses the proteins are divided temporally into immediate early (α), early (β), late (γ), and hyperexpressed late (δ). No viral proteins associated with the nucleocapsid are required for the initiation of mRNA production (transcription) or early protein synthesis (translation). Immediate early genes are transcribed by the host RNA polymerase II, and transcription begins very soon after infection. The presence of α-gene products initiates the transcription of β genes. The synthesis of the α and β proteins usually occurs within 3–6 h after infection and has been defined as the time up to the beginning of viral DNA synthesis.

At around 6 h post infection, the infected cells exhibit hypertrophy of the nucleus and the accumulation of a granular chromatin known as the virogenic stroma. This stage coincides with the beginning of viral DNA synthesis, believed to be directed by an α-amanitin-resistant RNA polymerase, and the transcription of the late or γ genes. The γ polypeptides include most of the structural proteins of the virions. Assembly of virions begins in the virogenic stroma about 12–18 h post infection. The nucleocapsids assembled during this period bud from the cell nucleus and are transported to the plasma membrane at sites where the viral protein, gp 64, has been inserted. Here the nucleocapsid buds from the cell and acquires its envelope.

Beginning about 18 h post infection, the δ genes, such as those for polyhedrin and p10, are transcribed and were originally referred to as very late genes. These two genes are transcribed until cell death and now are referred to as "hyperexpressed" genes. Also at this time, the nucleocapsids no longer bud from the cell nucleus, but rather become enveloped within the nucleus and are enclosed in the polyhedrin, forming the occlusion bodies or polyhedra that give these viruses their name, "nuclear polyhedrosis virus" (Fig. 3). Because the ODV are not needed to infect cell cultures and the polyhedrin is synthesized in large amounts, this gene site was the first to be used in constructing expression vectors. Additional information on the Baculoviridae may be obtained in *The Biology of Baculoviruses*, Vols. I and II (9), and in the review of Adams et al. (6).

THE CULTURE SYSTEM

Choosing a Cell Line

Two cell lines, the Sf-21AE (10) and Sf-9, a clone of the Sf-21AE (11), from the ovaries of the fall armyworm *Spodoptera frugiperda* have been widely used in studies with vectors constructed in the AcMNPV. Vectors constructed in the BmNPV are replicated in the Bm5 cell line (12). The BmNPV infects only the silkworm and thus far has been limited to cell lines from that insect. The AcMNPV infects 32 insect species (13) with cell lines available from many of them (Table 1). In addition to lines from several different species, there are different lines from a variety of tissue sources, such as embryo, ovary, blood cells, and fat body. These represent a wealth of potentially valuable host material in which foreign genes could be expressed. The Sf-9 cell line is used for the majority of studies, principally because it is available from commercial sources or from the American Type Culture Collection; it grows well in suspension, and many of the commercial cell culture media were developed using that line or its uncloned parent, Sf-21AE, to evaluate cell growth.

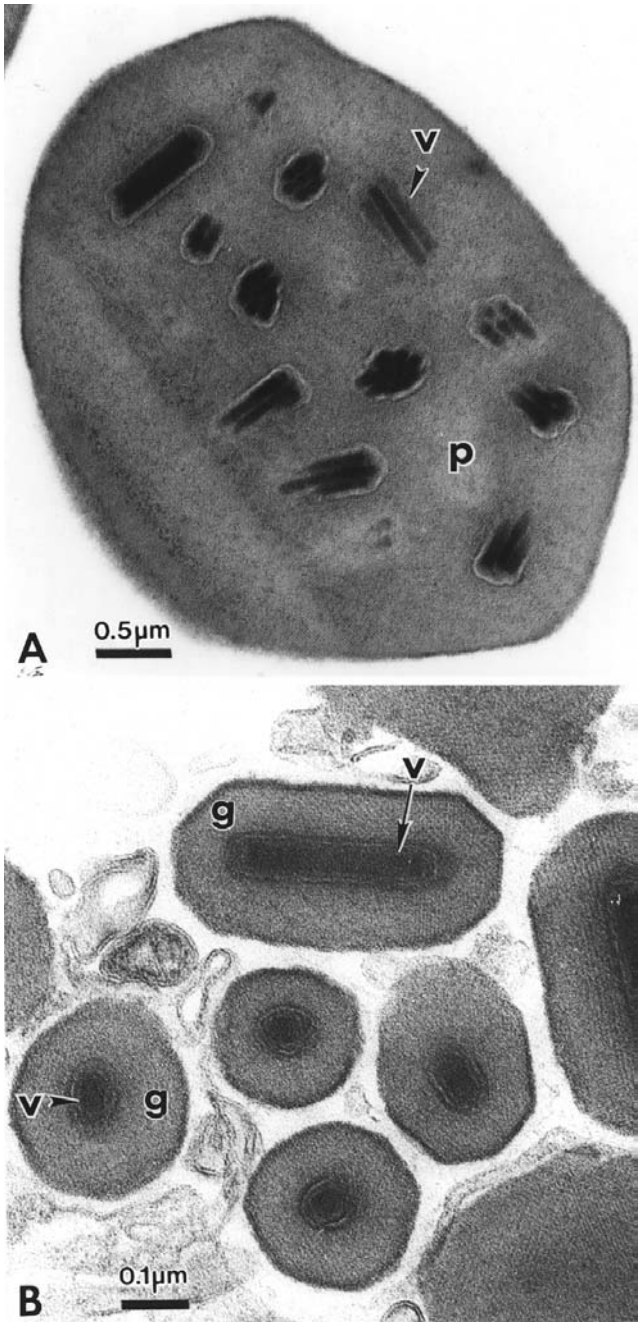


Figure 1. (**a**) Electron micrograph of a section of a polyhedra of a virus from the genus *Nucleopolyhedrovirus*. (**b**) Electron photomicrograph of granules of a virus from the genus *Granulovirus*. *p*, polyhedrin matrix; *v*, occluded virus particles; *g*, granulin matrix.

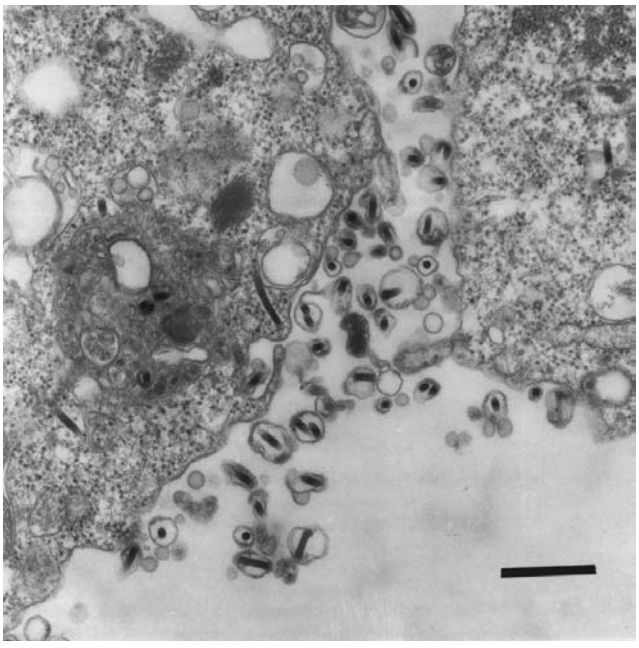


Figure 2. Release of budded virus of the gypsy moth NPV from infected gypsy moth ovarian cell lines. *Bar* 0.5 μm.

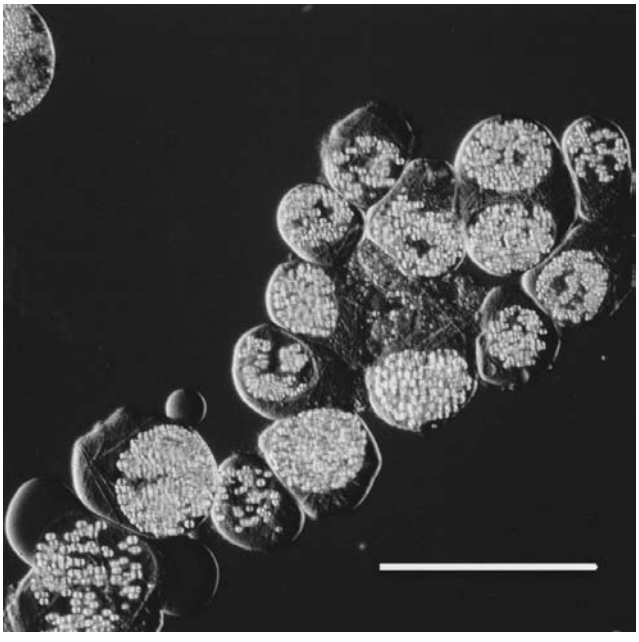


Figure 3. Formation of polyhedra in the nuclei of gypsy moth fat body cells, IPLB-LdFB, infected with LdMNPV. Viewed with Nomarski differential interference contrast optics. *Bar* 100 μm.

For ease in scaling up from laboratory studies to commercial production, cells that grow well in suspension without clumping, have relatively short population doubling times, and achieve high saturation densities are the most desirable. In media containing low levels of the surfactant Pluronic F-68, cells of the two spodopteran lines grow well in suspension in vessels with a capacity of at least 30 L. Population doubling times have been in the range of 18–24 h, and saturation densities of 5–8 $\times$ 10^6 cells/mL were achieved in batch systems. Saturation density was reached in 6–8 days depending on the level of inoculum used. The Tn-368 cell line from the cabbage looper, *Trichoplusia ni* (16), and the line Mb-0503 from the cabbage moth, *Mamestra brassicae* (17), also grow well in sus-

Table 1. Insect Cell Lines Susceptible to ACMNPV

Cell line	Tissue of origin	Species of origin	Reference
IPLB-Sf-21	Pupal ovary	*Spodoptera frugiperda*	10
IPLB-Sf-21AE, clone SF-9	Pupal ovary	*Spodoptera frugiperda*	11
IZD-MB0503	Blood cells	*Mamestra brassicae*	14
NIAS-MaBr-85	Larval fat body	*Mamestra brassicae*	14
Tn-368	Adult ovary	*Trichoplusia ni*	14
IPLB-TN-RR	Embryo	*Trichoplusia ni*	15
BTI-TN5B1-4	Embryo	*Trichoplusia ni*	15
IPLB-LD64BA	Pupal ovary	*Lymantria dispar*	14
IPLB-LdElt	Embryo	*Lymantria dispar*	15
UFL-AG-286	Embryo	*Anticarsia gemmatalis*	15
BCIRL-HZ-AM3	Pupal ovary	*Helicoverpa zea*	15
BCIRL-HV-AM1	Pupal ovary	*Heliothis virescens*	14
UCR-SE-1	Neonate larvae	*Spodoptera exigua*	15
NIAS-LeSe-11	Larval fat body	*Leucania separata*	15
IPRI-MD-108	Blood cells	Malocosoma disstria	14
FPMI-MS-5	Neonate larvae	*Manduca sexta*	15
IPRI-CF-1	Neonate larvae	*Choristoneura fumiferana*	15
IAFEs-1	Ovary	*Euxoa scandens*	15

pension and have been used by several investigators for the expression of genes in AcMNPV.

A recently developed cell line, TN-5B1-4 from *T. ni* (18), grows attached to a substrate and may be a line of choice in the production of expressed gene products secreted from the cells or where profusion is required for optimum yields. However, the cells are not obligatory substrate dependent, and substrains have been adapted to culture in suspension systems. These cells have been patented and are also available from commercial sources.

Product Yield. The levels of recombinant proteins produced in these few cell lines range from extraordinarily high, 1,000 mg/L, to less than 5 mg/L. The choice of the cell line used in production appears to be an important reason for the wide range in yields. Hink and his colleagues were among the first to examine the effect of the host cell line on product yield (19). Recombinant AcMNPVs expressing either the gene for pseudorabies virus glycoprotein (gp50T), human plasminogen (HPg), or β-galactosidase (β-gal) were tested in 23 different cell lines. The lines were selected to enable comparisons among lines from insects with different levels of susceptibility to AcMNPV and cloned lines with their parent line as well as the level of expression of each of the three genes in a cell line.

As might have been expected, no single line produced the highest yield from all three genes. Susceptibility of the insect for the wild-type virus was not a guide for selecting cell lines. Generally high yields corresponded to a high percentage of wild-type infectivity, but a high percentage of infectivity did not always result in high yields of expressed protein. *M. brassicae* is reported to be very susceptible to AcMNPV; however, four of the six lines from this insect had low sensitivity to the wild-type virus. One of these four cell lines did show expression of a recombinant gene. One line of the six had 78% of the cells infected by the wild-type virus and produced the second highest levels of β-gal but only moderate levels of HPg. The yields of expressed protein from two *T. ni* cell lines, with an almost fourfold difference in the percentage of cells infected by wild-type virus, corresponded to percent of infected cells except for gp50T, which was produced equally by both cell lines. Two lines from embryos of the gypsy moth, an insect not susceptible to the AcMNPV, were 70–78% infected with wild-type virus. Both produced moderate yields of HPg and gp50T. One, LdElta, produced more β-gal than Sf-21AE but not Sf-9. Betenbaugh et al. (20) also compared the yields of β-gal from Sf-9 and LdElta and found that LdElta produced slightly more β-gal activity on a per milliliter basis and on a per milligram of protein basis. LdElta also produced 38% more of the rotavirus protein VP4 than did Sf-9. Results of Hink et al. (19) were also mixed when the infectivity of cloned lines was compared with that of the parent uncloned line. The two clones, Sf-9 and Cl-15, of the *S. frugiperda* line Sf-21AE were several times more sensitive to the wild-type virus. However, the parent Sf-21 produced higher yields of HPg and β-gal than either clone. The clone of the *Spodoptera exigua* line, although equally susceptible to the wild-type virus, did not produce as high yields of any of the recombinant proteins as did the parent.

The foregoing studies were done in suspension culture, but similar comparisons have been made in static cultures in which differences in the effect of cell contact inhibition among cell lines must be taken into account. Cell growth in static cultures is often inhibited when the cell density becomes such that the cells are in contact with each other. Wickham et al. (21) screened eight cell lines in static cultures for the production of β-gal. In five of the eight cell lines, the effect of contact inhibition was moderate to severe, causing a three-to sixfold decrease in yield when the cells were infected at confluency. The production of BV also declined with increasing cell density. The cell lines Sf-21AE and Sf-9 were the least affected. These cells tend to attach lightly to a substrate and continue to grow even after confluency is reached. The cell line BTI-TN-5B1-4 was most affected.

When the density effects were eliminated by inoculating each cell line at an appropriate cell density, the

BTI-Tn-5B1-4 and the BTI-TnM cell lines produced 216 and 236 international units (IU) of β-gal/mL culture, respectively, compared to 97.2 and 104 for Sf-9 and Sf-21AE, respectively. Yields from the other cell lines tested were between these values. On a per cell basis, 5B1-4 and the *M. brassicae* line Mb0507 produced more than 1,000 IU per million cells and Sf-9 produced only 204 IU per million cells. The amount of β-gal produced per microgram of cellular protein by 5B1-4 and Tn M cells was 4.5 to 5.0 times greater than for Sf-21AE and Sf-9 cells.

The β-gal produced by the vector cell systems tested was retained within the cells. To determine if expression in a vector cell system in which the product was secreted into the surrounding medium was different, Wickham et al. (22) studied the expression of a gene for secreted alkaline phosphatase (SEAP) using the same eight cell lines as hosts. The cells were infected at a multiplicity of infection (MOI) of 10 and the cultures were harvested after 3 days. The 5B1-4 cells produced 5 to 10 times more SEAP per milliliter than any of the other lines, and 77% of the SEAP was secreted. The Mb0503 cell line produced the least SEAP per milliliter, but 85% was secreted. In Sf-9 cells only 59% of the SEAP was secreted. If secreted products are important factors in production or recovery of expressed proteins, then the choice of cell line should take this into consideration.

Posttranslational Processing. The capability of extensive protein processing by insect cells is particularly important in the preparation of human and animal vaccines. However, knowledge of the biosynthesis and catabolism of glycoprotein glycans from invertebrates, in particular from insects, has been described as "at best fragmentary" (23). The use of the baculovirus expression system to produce foreign proteins resulted in expanded interest in the posttranslational processing of proteins by insects. Cultured insect cells can perform O-linked and N-linked glycosylation, phosphorylation, mryistoylation, and palmitylation, folding, or cleavage of certain proteins to their active state. Contrary to early reports, some insect glycoproteins do contain sialic acid, and cultured cells can process nascent glycoproteins to sialylated, complex structures (23). Native proteins from the normal host animal differ slightly from those proteins produced in mammalian cells from other animals. Thus, proteins processed in insect cells would be expected to differ from the native protein. Glycoproteins expressed in Sf-9 cells are usually smaller than their native counterpart, which results from differences in glycosylation because the unglycosylated molecules are the same size as the native unglycosylated molecules. However, most of the more than 500 gene products expressed in insect cells that undergo co- and posttranslation processing are very similar in biological activity or antigenicity to the native protein. Glycosylation and other posttranslational modifications of foreign proteins by cultured insect cells are reviewed by Marz et al. (23) and Luckow (24).

Results from baculovirus–cell studies are inconsistent if not contradictory. Kuroda et al. (25) showed that a fowl plague virus hemagglutinin expressed in Sf-9 cells lacked galactose and sialic acid. However, a human plasminogen expressed in Sf-21AE cells contained complex type sialic acid-containing side chains normally found in this protein (26). This was the first documentation that these complex type sialic acid side chains could be produced in proteins from insect cells. As Sf-9 is a cloned line derived from Sf-21AE it is not clear what caused the differences. In studies where O-glycosylation is required, results are similarly unclear. The native human insulin receptor expressed in Sf-9 was not detectably glycosylated (27), however, the pseudorabies viral protein, gp50, was glycosylated (28).

Studies on the phosphorylation of proteins expressed in Sf-9 cells reveal considerable variation in the levels and sites involved. In his review of the subject, Jarvis (29) noted that in his work the simian virus 40 large tumor antigen (SVT ag) expressed in Sf-9 cells was phosphorylated at different sites than in mammalian cells, but that others had reported no difference in the phosphorylation patterns of the $p40^x$ protein in Sf-9 and in human lymphoid cells. Jarvis also noted that in a study by Hoss et al. (30) those phosphorylation sites in SVT ag protein recognized by cytoplasmic kinases were better utilized than sites recognized by nuclear kinases. The SVT ag protein and several other proteins were also acetylated in Sf-9 cells. The only processing pathway that seems to be missing in the spodopteran cells, Sf-21AE and Sf-9, is one for α-amidation (29).

Jarvis summarized his review by saying that the available results indicate that lepidopteran cells can adequately process foreign proteins with the exception of the α-amidation and some cleavages at sites containing basic amino acids. This conclusion is based upon studies with only one cell line and a single clone of the line. The differences already demonstrated in the capabilities of the parent line and its clone in glycosylation should provide an inducement to test additional cell lines. For example, cell lines from the insect fat body are more difficult to culture, but this organ is very important for the processing of metabolites in the insect. Baculoviruses replicate very efficiently in fat body cells, and lines from this tissue might very well contain processing pathways that are even better than those in the two lines used thus far.

Choosing a Medium

The first insect cell lines were developed by Grace (31) using a medium he developed. This medium became the standard for the culture of cells from lepidopterous insects, and most media developed later were based on it. One of the first modifications of Grace's medium was the TNM-FH medium of Hink (16), who added lactalbumin hydrolysate and yeast extract to the Grace formulation. Hink used this medium to establish the Tn-368 line from the cabbage looper. This was followed by a medium developed by Gardiner and Stockdate (32) that contained tryptose broth. This medium, originally called BML-TC/10, is sold today by several suppliers under the name "TC-100." A more defined modification of the Grace medium, IPL-41, was developed in 1981 by Weiss and colleagues (33). Several trace metals and vitamin B_{12} were added to the Grace formulation. Maltose replaced some of the sucrose, and the levels of several amino acids were changed. All these media are sold by several media suppliers either as dry powders or

as liquids. Generally all the media require supplementation with 5–10% fetal bovine serum (FBS). When the media are used in suspension cultures, methyl cellulose or Pluronic F-68 is added. Most of the commonly used lepidopteran cell lines will grow well in any of these media, and a cell line seems to adapt easily from one to another.

The requirement for FBS presents some problems when insect cells are to be used for the production of gene expression products. It is expensive, the quality varies from lot to lot, and the supply is sometimes unreliable. There is also the possibility of introducing bovine viruses or mycoplasmas into the medium with the serum. Finally, the serum proteins complicate the downstream recovery and purification of the expressed protein. For these reasons, there has been a great interest in the development of protein-free media. The IPL-41 medium was the starting point for the development of most of the protein-free medium. Maiorella et al. (34) developed the first such formulation. The FBS was omitted and the IPL-41 was supplemented with 0.4% ultrafiltered yeast extract and a lipid emulsion containing cod liver oil, fatty acid methyl esters, cholesterol, and α-tocopherol (added as an antioxidant and free radical scavenger).

The population doubling times and the maximum cell densities of Sf-9 cells grown in this medium were equal to those of the cells in serum-supplemented IPL-41. Yields of macrophage colony-stimulating factor from cells infected with a baculovirus containing the gene for this protein were equivalent in both the IPL-41-FBS-supplemented medium and the serum-free medium. There are a number of protein-free media available from commercial supplies that are, presumably, similar in composition to the medium of Maiorella et al., and the yields of foreign gene products from cells grown in these media generally have been comparable to yields in FBS-supplemented media.

Reporting on the development of the serum-free medium SF900, Weiss et al. (35) showed that Sf-9 cells reached higher maximum cell densities and had shorter population doubling times on the serum-free medium than on Grace's medium plus 10% FBS. Similar results were obtained with the TN-368 cell line, where the final cell density was almost twofold higher in the serum-free medium compared to growth in Grace's + FBS. The production of β-gal in serum-free medium was 152% of that in Grace's medium + serum. A maximum of 500,000 units/mL was produced in 6 days in shaker flask cultures.

The growth of Sf-21AE cells in stirred tank bioreactors was compared using serum-supplemented medium (TNM-FH, TC-100, TC-100 + 2.5 g/L tryptose phosphate broth, IPL-41, and IPL-41 + 2.5 g/L tryptose phosphate broth) and serum-free medium (Ex-CELL 400, EX-CELL 401, and the Sf900) (36). Similar final cell densities and population doubling times were obtained with the TC-100, the IPL-41, and the serum-free media. The additional supplementation of either TC-100 or IPL-41 with tryptose phosphate broth gave superior growth compared to all other media.

Hink and colleagues (19) compared yields of gp50T, HPg, and β-gal from the cell lines Sf-21AE, Sf-9, TN-368, and EAA in the serum-free medium of Maiorella et al. or FBS-supplemented TNM-FH, IPL-41, and TC-100. The results did not consistently favor either serum-supplemented or serum-free media. For example, Sf-9 cells produced more HPg in serum-free medium but more β-gal in serum-supplemented medium. Results with Sf-21AE were reversed—more HPg with serum-supplemented and more β-gal with serum-free medium. These studies were done in static culture systems with the cells attached to the growth surface.

When EX-CELL 400 was used as the serum-free medium, the β-gal produced by either Sf-21AE, Sf-9, or TN-368 was one-eighth to one-third lower than the levels produced by these cells in TC-100 + FBS (37). These studies were also done in static cultures. Comparison of the replication of wild-type virus in the two media indicted very similar levels of BV produced in either serum-supplemented or serum-free medium. Similar numbers of polyhedra were produced in each test system and the specific activity, as determined by bioassay, was the same as or higher in the serum-free medium than in the serum-supplemented medium.

Although it is clear that the choice of medium is an important factor in optimizing either the replication of wild-type baculoviruses or the expression of recombinant gene products, the available data give little guidance in choosing a medium. The currently available protein-free media appear to provide adequate nutritional support for the production of recombinant gene products. With the advantages such media provide in the downstream harvesting of the gene product, they are the most desirable. As the physiology and metabolism of insect cells in culture become better understood, improved media formulations will certainly be developed. At this time, testing various combinations of cell lines and media is the best approach to optimizing yields.

Preparing a Recombinant Baculovirus

Recombinant baculoviruses, because of the large size of their genome, are prepared using a multistep process to insert the foreign gene into the viral genome. Beginning with a bacterial plasmid, the first step is to select or construct a suitable transfer vector and insert the desired foreign gene. Then, susceptible insect cells are simultaneously infected with the transfer vector and wild-type AcMNPV DNA (transfection). In the infected cell, the two DNAs recombine and produce the new recombinant virus. The progeny from this virus must then be isolated and separated from the wild-type virus progeny produced in cells infected only by the parental DNA. This step is typically accomplished by infecting a monolayer of cells and reisolating the virus from well-isolated foci or plaques. Usually, three cycles of plaque purification are required to ensure the complete separation of the recombinant and wild-type viruses. Recombinant viruses have been prepared most often from the AcMNPV and less frequently from BmNPV, but the same or similar methods should be useful in preparing recombinants from other NPVs. For excellent detailed discussions of preparing recombinant baculoviruses, see Refs. 24, 38.

Selecting a Transfer Vector or Plasmid. The transfer vectors have three features in common: (*1*) they are bacterial

plasmids that have an origin of replication and an antibiotic-resistant marker to aid in selection when the plasmid is amplified in *Escherichia coli*; (*2*) the plasmid also has a suitable viral promoter and a suitable restriction enzyme site for cloning downstream from the promoter site; and (*3*) the cloning site is flanked by native AcMNPV DNA to promote homologous recombination. When the polyhedrin promoter is used in constructing the vector, the nature and the amount of the 5′ upstream sequence of the gene in the native DNA has an important effect on the expression of the foreign gene (39,40). Transfer vectors used with some baculovirus are of three types: (*1*) vectors to express a single foreign gene, (*2*) coexpression vectors that contain the gene for a marker protein, such as the *Escherichia coli LacZ* gene, and a gene for the desired protein, and (*3*) dual expression vectors containing genes for two foreign proteins. Many of these vectors are described in the literature and can be purchased from commercial sources (e.g., Clontech, Palo Alto, Calif.; PharMingen, San Diego, Calif.; Invitogen, San Diego, Calif.; Life Sciences Inc, Gaithersburg, Md.; Stratagene, San Diego, Calif.). Descriptions of several baculovirus vectors, some commercially available, are given by López-Ferber et al. (38).

The first, and perhaps most commonly used, cloning site was the polyhedrin gene site. This gene is not essential for maintenance of the virus in cell culture and is abundantly expressed in the infected cell, often to as much as 50% of the total protein produced in an infected cell. Another nonessential gene site often used is the p10 gene, which codes for a fibrillar protein with a relative molecular weight of 10,000. It also is expressed in the late phase of virus replication in copious quantities. These two gene promoters have been used together to construct vectors for multiple gene expression by a number of researchers (e.g., 41,42).

Transfection. Transfection is the next critical step in the development of a recombinant baculovirus. Procedures used should be designed to optimize the recombination between the vector and the viral DNA. The commonly used coprecipitation of DNA with calcium phosphate requires inexpensive reagents that are simple to prepare. The calcium ions and DNA are mixed together and allowed to form a fine precipitate, which when combined with cells will bind to the cell surface and facilitate the transfer of the DNA into the cell. DEAE-dextran, poly-L-lysine polyomithine, and polybrene have also been used to transfect insect cells. The efficiency may sometimes be improved by treating the cells with dimethylsulfoxide or glycerol to osmotically shock them. Typically, only 0.1–1.0% of the progeny viruses are recombinants, requiring that large numbers of plaques must be screened to locate the few containing recombinant viruses. Cationic lipids can also be used in transfection. One such preparation, Lipofectin, is commercially available. A procedure for using this material was described by Trotter and Wood (43), who typically obtained virus populations containing 15% recombinants with the Lipofectin procedure using the supercoiled form of DNA.

Kitts et al. (44) reported that the proportion of recombinant progeny viruses can be increased to about 25% by using linearized viral DNA instead of the circular DNA. DNAs that can be linearized at a specific point are prepared by introducing a unique enzyme recognition site. The vector has sufficient sequences from both sides of the restriction site on the viral DNA to facilitate homologous recombination. Recombination closes the cut and restores circularity to the viral DNA, thereby permitting replication, as it is believed that linear DNA cannot be replicated. In this system, the number of progeny viruses is greatly reduced, making the location of recombinant virus much easier.

PRODUCTION PROTOCOLS

Choosing a Bioreactor

The insect cell lines Tn-368, Sf-21AE, and Sf-9 do not require attachment to a substrate to grow, and thus suspension culture systems are favored for the production of recombinant gene products. In all suspension systems the cells must be kept uniformly distributed throughout the medium to provide equal access to nutrients and prevent the localized buildup of toxic metabolites. The dissipation of the energy required to maintain this needed homogeneity can cause significant cell damage. The larger the system, the more of a problem will be damage resulting from mechanical agitation. Animal cells and insect cells in particular, because of their relatively large size and the lack of protective cell wall, are more susceptible to mechanical damage than microorganisms. A number of systems have been developed to overcome this problem and provide a means for the economic production of insect viruses and gene expression products.

Laboratory Scale. For research purposes, shake flasks or spinner flasks are usually the systems of choice. Shake flask culture systems are simple, require no expensive equipment, and provide an excellent environment for growth of insect cells. Neutra et al. (45) reported that shake flasks with working volumes up to 120–150 mL supported growth of the Sf-9 cells to a final density of 5×10^6 cells/mL with a doubling time of 20 h. The cultures were incubated on an orbital shaker at 100 rpm. Working volumes of 25–45% of the flask volume were found to be most satisfactory. Cultures infected with an engineered baculovirus containing the *E. coli LacZ* marker gene produced about 600 mg/L of β-gal.

Spinner flasks also are easy to use and require little expensive equipment. They have the added advantage that the culture vessels used can be easily modified to add aeration equipment or the incorporation of electrodes for monitoring pH or other parameters. This permits the use of larger working volumes and allows the quantitative monitoring of some culture conditions. Although there are several different designs for culture vessels, the basic design is one in which a magnetic bar or paddle is suspended in the growth chamber and driven by an external magnet. Standard spinner flasks are available with fluid volumes from 25 mL to 36 L. Many designs have side arms that provide access without removing the cap. Larger vessels

have ports in the lids for pH and O_2 probes, medium addition and removal, and gas inlets.

Working volumes in spinner flasks range from 32% to 80% of capacity and rotation speeds of the stirring bar are from 50 to 100 rpm. In working volumes less than 500 mL, oxygen transfer from the headspace of the flask into the agitated culture medium is sufficient to meet the requirements of the insect cells (46). Cameron et al. (47) found that with agitation of 100 rpm oxygen levels became limiting at working volumes of 1000 mL. Therefore, in these small cultures sparging is not needed and the only shear forces affecting the cells result from the agitation of the medium. Eberhard and Schügerl (48) demonstrated the importance of agitation speed in designing culture systems. In cultures with a working volume of 100 mL in 1000-mL flasks at an agitation speed of 35 rpm, cells settled to the bottom and growth was reduced as a result of localized depletion of oxygen and nutrients. Final cell densities increased with increased agitation speed, up to 100 rpm, until oxygen levels were no longer growth limiting. Kioukia et al. (49) compared the growth of Sf-9 cells in 100-mL spin flasks, 50-mL working volume, stirred at speeds of 80, 120, and 150 rpm. Optimum cell growth before infection was at 120 rpm. Viral infection was enhanced at 150 rpm but, as in the noninfected cultures, cell damage occurred. It appears that 100 to 120 rpm is adequate for nonaerated spinner cultures if there is sufficient headspace in the vessel. Culture volumes greater than 1 L require aeration and thus present additional problems in bioreactor design and operation.

Scale-Up. Once the need for culture volumes exceeds 1 L, providing sufficient oxygen for cell growth and virus replication becomes a serious problem. Hink and Strauss (50) reported that the dissolved oxygen (DO) level in a 2-L bioreactor decreased from 100% to 2% in 24 h during the growth of the Tn-368 cells. The amount of sparging required to maintain the DO level at 100% caused excessive medium foaming and vacuolation of the cells. Infected cells consume oxygen at almost twice the rate of uninfected cells, thus increasing the supply problem. Tramper et al. (51) attempted to identify the sources of shear stress and to establish the level at which Sf-9 cells were damaged. They found that the critical shear stress at which cell viability began to decline was 1–4 N m^{-2}. This result agreed with the data of Hink and Strauss for the TN-368 cells. They found, as did Hink and Strauss, that cell death occurred at stirrer speeds greater than 220 rpm. They also observed that cell survival in the foam was very low and that the number of viable cells decreased with every bubble burst. Their conclusion was that the dispersion of gas in the medium by means of stirring and sparging sufficient to provide the needed oxygen would result in a larger death rate than the growth rate of the cells and that membrane bioreactors or oxygen supplied through semipermeable tubing might be required to solve this problem.

The effects of hydrodynamic forces on insect cells in culture were reviewed by Chalmers (52), and his conclusions are discussed here briefly. Cells in sparged, agitated culture systems are subject to shear forces associated with gas bubbles under three conditions: in the region where the bubbles are injected into the medium, during the free rise of the bubbles to the surface, and at the air—medium interface at the top of the vessel. The preponderance of data reviewed by Chalmers indicated that most of the cell death occurred at the air–medium interface. During the rise of the bubbles through the medium, cells became attached to a bubble and death occurred when the bubble ruptured at the surface. According to studies by Trinh et al. (53), on the average 1,200 Sf-9 cells were killed per bubble burst when suspended in TNM-FH medium without protective additives. Only 292 cells killed were present on the bubble film; the remainder was believed to be in the thin layer of medium surrounding the bubble. No cells were killed when the medium contained 0.1% Pluronic F-68.

Protective additives are important components of media for insect cells grown in suspension culture systems. Vaughn (54) used 0.15% methyl cellulose for the first culture of lepidopteran cells in spinners. Hink and Strauss (50) found that 0.3% methyl cellulose allowed them to increase the impeller speed from 100 to 220 rpm in a 4-L stirred tank fermentor. However, the most successful protectant has been Pluronic F-68. Maiorella et al. (34) and Murhammer and Goochee (55) used it for the culture of Sf-9 cells in an agitated, sparged bioreactor and in an airlift bioreactor. Murhammer and Goochee reported that population doubling times and final cell densities were not affected by concentrations up to 0.5% Pluronic F-68. Maximum viability of 95–99% was achieved with 0.1%. This concentration also was used by Maiorella et al. and is the concentration most frequently used for insect cell cultures.

Murhammer and Goochee suggested that Pluronic F-68 protected cells by interacting with the cell membrane and that this interaction was the basis for the protective mechanism. Photographs presented by Chalmers (52) show that in the presence of 0.2% Pluronic F-68, no cells attach to gas bubbles in the culture medium. Therefore, much of the killing of cells when the bubbles rupture at the air–medium interface was avoided. The mechanism that prevents the cell–bubble interaction is not known, but effects have been quite striking. The development of scale-up protocols for the growth of insect cells on a commercially feasible scale now is concentrated on the use of the standard stirred tank reactors or airlift fermentors rather than the development of specialized vessels.

Sparged Stirred Vessels. The sparged, stirred tank fermentors used extensively for many types of microbial fermentations can be operated inexpensively and their scale-up is straightforward. Glass fermentors with working volumes of a few liters to stainless steel tanks with capacity up to 8,000 L have been used for a number of years for the production of biologicals in mammalian cells. The most notable examples are the production of foot-and-mouth virus vaccine, interferon, and monoclonal antibodies. These vessels, used to culture animal cells, are round bottomed and fitted with agitators having marine propellers or one or more Rushton-type turbine stirrers. The stirrers may be driven directly or by external magnetic drives. To protect the cells from contact with air bubbles and the damage that occurs when those bubbles break, some specialized aeration systems have been designed in which the air is

introduced into the medium stream inside rotating stainless steel mesh cages that exclude the cells.

With the discovery of the protective effect of Pluronic F-68, most insect cell culture is now done in the more typical sparged, stirred tank. Overton and Kost (56) reported the use of 36-L stirred vessels to produce human synovial phospholipase A_2 in Sf-9 cells. The vessels were fitted with a ring sparger with 2-mm holes, and DO level was maintained at 50% by oxygen sparging. The expression level of purified product from this system ranged between 1.0 and 1.5 mg/L and was comparable to that from shake flasks. Silberklang et al. (57) have described the scale-up of a stirred tank system of 70 L to produce antistasin, an anticoagulant protein from a leech, in Sf-9 cells. They used an open-pipe (5-mm-diameter), sparged Chemap fermentor equipped with hydrofoil impellers. The key to their success in scale-up, they considered, was the use of a strain of the Sf-9 cells that had been well adapted to suspension growth by serial passage in small spinner cultures for 6–8 weeks before freezing the cells in the master cell bank.

Airlift Bioreactors. Airlift bioreactors have been used for some time to culture fragile cells such as hybridoma cells. These vessels consist of two tubes, one inside the other. Air is sparged in the bottom of one tube, the riser. As the air bubbles rise to the top of the riser the difference in density between the medium in that tube and the second tube, the downcomer, results in circulation of the medium and cells without the need for mechanical stirring. Possible damage to fragile cells from the turbulence and resulting shear forces in the eddies at the ends of the stirrer blades is avoided. Insect cells were considered to be more sensitive to these forces than most other animal cells and so many investigators studied their growth in the airlift bioreactor.

The airlift bioreactors have been modified for use with attachment-dependent cells such as the BTI-TN-5B1-4 line by adding glass beads or other materials to the downcomer tube, which is modified to retain them. The beads form a packed bed to which the cells attach and are then bathed in oxygenated medium. This method avoids the shear forces the cells encounter when the beads are maintained in suspension in stirred tanks. A model system describing the effects of such parameters as the ratio of riser area to downcomer area, airflow rate, and DO to the growth of the 5B1-4 cells has been developed by Kumar and Shuler (58).

Maiorella et al. (34) used an airlift bioreactor in their early studies on the scale-up of Sf-9 cells for recombinant protein production. Pluronic F-68 was added to the medium to protect the cells from damage by the bursting bubbles. It also limited the buildup of foam so that no additional antifoam was needed. Bubble size was controlled at 0.5–1.0 cm in diameter by selection of a sparger with the appropriate orifice diameter. A sparging rate of 420–1260 mL/min was used as this provided an adequate supply of oxygen but did not cause excessive turbulence or foaming in a 21-L vessel. Cell growth was equal to that in 100-mL shake flasks, and the production of recombinant human macrophage-stimulating factor reached 40 mg/L in 4 days.

Rice et al al. (59) compared the growth and production of recombinant protein by Sf-9 cells in an 8-L stirred tank vessel and in a 6-L airlift vessel and found that the cell growth was almost identical in the two systems. Population doubling time in the airlift was 22.5 h and in the stirred tank 21.9 h. Maximum density of the Sf-9 cells in the airlift was 4.94×10^6 cells/mL compared to 4.91×10^6 in the stirred tank. Yields, specific activity of the recombinant human protein kinase, were also virtually identical. They also considered the relevant costs of each system. In the laboratory the cost of providing temperature control, DO control, and other operating expenses were about equal. However, Rice and his colleagues estimated that the cost of scaling up the airlift system to 36 or 45 L would be 98% more than that of scaling up the stirred tank system to the same level. Other examples of the use of these systems for the production of recombinant proteins from insect cells can be found in the proceedings of a workshop on the subject (60).

Establishing Culture Conditions

Sterility. All animal cells must be grown under sterile conditions. The very rapid growth rate of most microorganisms compared to that of animal cells means that even the presence of small numbers of microorganisms at the beginning of the fermentation will result in the overgrowth of the animal cells by the microorganisms. This will cause the loss of the fermentation or at least a significant reduction in the yield and additional problems in product recovery.

Insect cell culture media generally contain heat-labile components and, therefore, must be filter sterilized. Liquid media obtained from commercial media suppliers are sterile and the sterility need only be maintained by use of proper microbiological technique. If the media are prepared on site, pressure filtration, using compressed air or nitrogen, through a filter stack consisting of a depth prefilter, a 0.45-μm membrane filter, and finally a 0.22- or 0.01-μm membrane filter held in a stainless steel filter holder, is recommended. In the preparation of small batches of medium the filtered medium is usually collected aseptically into sterile vessels and held refrigerated until needed. In this situation it is feasible to run sterility checks before the batch of medium is used. When large volumes of medium are required for a single production run, it may be most practical to prepare and sterilize the medium just before use. Clarifying filters need not be sterile and may be separate from the sterilizing filters. Sterilizing filters should be located as near to the fermentors as practical and should be tested for integrity before use. Venders provide methods and equipment for this test.

Small fermentors can be assembled and sterilized by autoclaving, assuming an autoclave large enough to hold them is available. All parts of the system must be thoroughly cleaned and sanitized using one of several detergents sold for this purpose. To prepare the vessel for autoclaving, all access ports should either be fitted with the correct tube or probe or plugged. Oxygen and pH probes should be prepared according to instructions provided by the manufacturers. Glass connectors should be secured, the tips plugged with a nonabsorbent material such as glass wool and covered with aluminum foil. The vessel should be filled with water or phosphate-buffered saline to the working level. The sterilizing time will depend on the

size of the vessel. The vessel must be vented at all times and should not be removed from the autoclave until the vessel has cooled somewhat to avoid drawing nonsterile air into it. The vessel can then be closed, placed in a sterile hood, if possible, and allowed to cool completely. The water is then aseptically removed and the culture medium added.

Fermentation vessels too large or too heavy to be placed in an autoclave are steam sterilized in place. Proper design of the system is very important to achieve sterilization of such equipment. Steam vapor must contact all the parts of the system to be sterilized. Steam must enter the vessel in such a manner as to replace all the cold air, and the resulting condensate must be adequately drained from the system. Pockets and crevices in the system must be avoided or minimized. Perhaps most important, standard operating procedures must be developed and the operating personnel well trained in those procedures. Sterilization is discussed in more detail by Bergman (61).

Oxygen Supply. Insect cells require oxygen to grow and will rapidly deplete the dissolved oxygen in the culture medium. Oxygen availability is critical to optimum yields of both wild-type virus and recombinant protein production. Determining a suitable level of dissolved oxygen and developing ways to maintain that level have been the subjects of considerable research. The oxygen utilization rate (OUR) for insect cells, $3.5–9.9 \times 10^{-17}$ mol cell^{-1} s^{-1}, is slightly higher than that of many other animal cells. Moreover, the rate increases to nearly double by 12–18 h following virus infection and then declines to near that of uninfected cells. This period of maximum OUR corresponds to the time of γ-gene expression and the production of virions.

The infected insect cells are also more fragile than uninfected cells. Thus, simply increasing the sparging rate or the agitation may be counterproductive. One possible alternative may be lowering the incubation temperature from 27 °C to 25 °C, which increased the oxygen solubility and lowered the OUR of Sf-9 cells by 26% without a reduction in yield of recombinant β-gal (62). However, DO levels between 20% and 60% will support optimum cell growth and virus replication and have been achieved in most bioreactors with sparging raters that do not cause significant cell damage. See the review by Taticek et al. (63).

Feeding Strategies. Bioreactors may be operated in several ways related to the way nutrients are supplied to the cells. Batch, fed-batch, semicontinuous, and perfusion methods have all been tested for growing insect cells. Stirred tank bioreactors operated in a batch mode are the most common microbiological fermentations. They are simple, inexpensive, and easy to scale up. In this system medium is added to the bioreactor and inoculated with the desired cells and then run until cell growth or product production stops. During this time the cell nutrients are gradually depleted and metabolic waste products accumulate in the tank. pH may change as a result of the accumulation of such waste products as ammonia or lactic acid. One of these changes can eventually become growth limiting. However, a baculovirus infection ultimately results in death and lysis of the cell, and this may terminate the fermentation before any of the metabolic factors become limiting. Therefore, the batch method, using a carefully designed starting medium, may provide optimum yields with the least cost.

Additional cell growth or product production may be obtained with the fed-batch system, if either nutrient depletion or waste accumulation has terminated cell function before the lysis of infected cells. Limited studies on the accumulation of ammonia in cultures of insect cells indicates that ammonia does not accumulate to toxic levels unless the glucose is depleted (64,65). The addition of ammonia in the form of NH_4Cl or $(NH_4)_2SO_4$ up to a concentration of 10 mM was not inhibitory to cells of either the Sf-9 or BTI-EAA lines. Growth did slow somewhat at a concentration of 30 mM, indicating that these cell lines are considerably less sensitive to ammonia than many mammalian cells (64). If adequate concentrations of glucose were present, the concentration of ammonia found in both these studies did not approach these levels (64,65). Similarly, insect cells produce only small amounts of lactic acid during growth in an adequately oxygenated culture. Bédard et al. found that lactate was consumed by Sf-9 cells at the end of the growth period (64). They attribute the difference between lactate production in mammalian and insect cells to the presence of a complete TCA cycle in insect cells but not in many mammalian cells.

The advantage of the fed-batch system appears to be from the replacement of depleted nutrients. Feeding can be done simply by adding additional complete medium to the culture. Lindsay and Betenbaugh (66) found that the resuspension of high cell density cultures of Sf-9 cells in fresh medium at the time of infection resulted in a fourfold increase in yield of the recombinant β-gal. However, this is somewhat inefficient because many of the components in the medium are not severely depleted during a normal growth cycle; for example, it is unlikely that any vitamins or inorganic salts become growth limiting. If the growth-limiting nutrients can be identified, only they need to be supplemented.

Few of the amino acids in a medium are depleted more than 50%. The need for amino acids varies among insect cell lines. Mitsuhashi (67) determined the amino acid requirements of cell lines from six different insect species. Only five, cystine, histidine, isoleucine, methionine, and threonine, were essential for all six cell lines. In the study of Bédard et al. (64) only methionine, of these five, was consumed by Sf-9 and BTI-EAA cells. However, the nonessential amino acids glutamine, glutamate, and aspartate were extensively depleted. No amino acid was depleted to growth-limiting levels.

Wang et al. (68) found that the specific yield of a recombinant protein from Sf-9 cells was increased 100% by feeding glucose and glutamine to high cell density cultures immediately after infection. The amount added returned the levels to approximately those in fresh medium. Lee and Park (69) demonstrated that the addition of Yeastolate to a culture of Sf-9 cells in stationary growth phase would stimulate further growth and increase the final cell density about 3.5 fold, but no information on the critical components of the Yeastolate was given.

Semicontinuous culture is another way to maintain suitable levels of oxygen and nutrients to support optimum virus replication or recombinant protein production in high cell density cultures. In these systems, cells are grown in one vessel and when high cell densities are reached, a portion of the cell suspension is transferred to one or more new vessels, fed with fresh medium, and inoculated with virus for production. Fresh medium is also added to the cell culture vessel to continue the production of cells. At the time the virus or recombinant protein is harvested some medium may be left in the production vessel to inoculate the next batch of cells. Such systems have been used with TN-368 cells (50) and with Sf-9 cells (70) for the production of wild-type AcMNPV. A similar two-stage system was used with *B. mori* BM-5 cells for the production of recombinant proteins (71).

Perfusion culture systems have supported the highest final cell densities, $>10^7$ cells/mL, achieved with insect cells. In perfusion cultures the fresh medium flows through the culture vessel at a constant rate and all the cells are retained in the vessel during the entire process. This method requires that the cells be removed from the effluent either by filtration or entrapment, as in hollow fibers. This method reduces the amount of headspace needed in a fermentor because the medium can be oxygenated before entering the fermentor. Cells are exposed to a steady-state environment, and higher cell densities per unit volume are possible. The perfusion is either "open," in which the medium passes through the bioreactor and is collected, or "closed" in which the medium is reused after oxygen has been added and, if necessary, the pH adjusted.

Deutschmann and Jäger (36) were able to maintain Sf-21AE in exponential growth for 16 days in a perfused 1.4-L stirred tank. The maximum cell density achieved was 5.5×10^7 viable cells/mL. Perfusion was started after 72 h of cultivation and maintained at a constant rate for the duration of the fermentation. Oxygen was provided through a polypropylene membrane, 5 m/L working volume. As the cell density increased, available oxygen was increased by increasing pressure in the vessel and finally by aerating with pure oxygen. When the cell density reached the maximum, DO could not be maintained, in part because of membrane fouling, and limited oxygen was given as the reason growth stopped.

Weiss and Vaughn (72) produced wild-type AcMNPV in Sf-21AE cells in a perfusion system using a 3-L Moduculture vessel equipped with Biospin filters. Cultures with a cell density of 1×10^5 cells/mL were inoculated at a MOI of 0.005 plaque-forming units (PFU)/mL. Medium was perfused in the open mode at 600 mL/day for 6 days. A 98% infection resulted with a yield of 1×10^8 polyhedra/mL. A total of 6.6 L of spent medium containing 2.4×10^9 BV/mL was produced, a 10-fold increase over the yield from batch spinner cultures.

Caron et al. (73) established Sf-9 cells in a 4-L stirred tank perfusion system using a custom-designed tangential flow filter to separate cells and medium and reoxygenated the used medium, eliminating the need for sparging the culture. The culture was incubated in the batch mode until the cell density reached 3×10^6 cells/mL. Perfusion then began at a rate of 1.4–2.8 mL/min. When the culture reached a cell density of about 1×10^7 cells/mL, the volume was reduced to 1 L, the cells infected with the recombinant virus, and serum-free medium added to return the volume to 4 L. During the production phase the culture was perfused with medium containing only 0.5% or 1.0% serum instead of 10% FBS. Cell density increased from 4×10^6 to 15×10^6 cells/mL in 3 to 4 days. However, the average protein production in this system was reported to be only 76% of that in shaker flask cultures. The lower level was believed to result from the serum level during the infection stage and not from the conditions of culture during the production phase.

Infection Strategies

Passage Effects. As described earlier, baculoviruses produce two different virions during their replication: the budded virus (BV), released from the cell by budding through the plasma membrane, and the virion enclosed in the polyhedra, the occlusion-derived virus (ODV). The ODV released from the occlusion body under alkaline conditions, pH above 9.5, is highly infectious through the gut of a target insect, but is infectious to cells in culture only at extremely high concentrations and therefore is of little use in cell culture studies. The BV are very infectious to cell cultures and are the standard means of infecting cells with either wild-type or recombinant viruses. Continued passage in insect cell culture of a baculovirus using the BV results in the gradual decline in polyhedra yield and infectivity of the polyhedra produced (74).

To avoid these passage effects, it is standard practice to use only low-passage virus in vitro studies. A master bank of virus inoculum is prepared, usually from the blood of insects fed polyhedra, and stored frozen. To prepare this material, susceptible insects in the midlarval stage of development are fed a dose of polyhedra sufficient to kill 90% of them. When the infected larvae begin to show signs of disease, they are bled by cutting a rear leg and collecting the blood in ice-cold cell culture medium. The collection in cold medium is required to prevent enzymatic oxidation of phenols to toxic quinones in the hemolymph upon exposure to air. In addition, antioxidants may be added to the medium. The diluted blood is centrifuged at low speed to remove the blood cells and the supernatant filtered through a sterile 0.45-μm membrane filter. The filtered blood is divided into small volumes and frozen at -80 °C. This blood is used to prepare the subsequent passages in cell culture. Generally a large amount of the third-passage in vitro virus is prepared and becomes the working stock. With the AcMNPV, this allows the buildup of seed inoculum and a production run of several viral replications before significant passage effects develop. Not all viruses will remain stable for this many in vitro passages. Slavicek et al. (75) found that the wild-type gypsy moth NPV was not stable for more than one in vitro passage. However, they were able to select a strain with greater stability for cell culture studies.

Wickham et al. (76) demonstrated the generation and accumulation of defective interfering (DI) particles within high in vitro passage baculovirus. DI particles build up in in vitro culture systems as a result of continued passage of the virus at high MOIs. These particles, identified mi-

croscopically by their shorter length, were lacking large sections of their DNA. Typical of DI particles, they caused large reductions in the amount of recombinant protein produced in cells infected with MOIs of more than 0.01 plaque-forming unit/cell. To avoid the accumulation of defective particles the cells must be inoculated with dilute virus suspensions, low MOI. In a system where the residual culture is left in the production vessels to inoculate incoming cells, the level of infectious virus gradually increases with each run, and with each passage the proportion of defected particles also increases. To delay the buildup of DI particles, De Gooijer et al. (77) proposed a mathematical model for the infection of insect cells with NPV in a semicontinuously or continuously operated two-stage reactor. Such a reactor was assembled and the model tested as a semicontinuous system (78). The system consisted of two 1- or 2-L reactors connected in sequence. Cells were grown in the first reactor and used to inoculate the second, the production reactor. The culture in the production vessel was established with 300 mL of cell suspension and additional cell suspension was added for the next 3 days until the working volume of 800 mL was reached. The production reactor was then operated as a batch reactor for 1 day. Production of β-gal was 35 U/mL for 7 runs compared to an average of 22 U/mL when operated in a simple batch mode for 12 consecutive runs of 5 days each. As predicted in the model, the amount of infectious virus produced began to decline at this point in each system, as did the production of β-gal, as the result of the buildup of DI particles.

MOI and Cell Density Effects. The MOI used may also affect the total yield of virus or recombinant protein. Dougherty et al. (79) found that with a MOI of 5, 98% of the Tn-368 cells were infected in 24 h and at a MOI of 20, 99% were infected. Brown and Faulkner (80) found that in monolayer cultures of Tn-368 cells the yield of wild-type virus was constant for increasing MOIs from 0.001 to 1.0, increased with increasing MOI until 10, and then began to decline. They suggested that at higher MOIs at which cells were infected by several particles that damage to the cell membrane caused the decrease in virus yield. Licari and Bailey (81) reported that in monolayer cultures of Sf-9 cells infected in early log phase of growth that the yield of β-gal was constant for MOIs between 0.1 and 10, but lower at 0.01 and 100. However, when monolayers were inoculated in the late log phase the yield increased with increased MOI from 0.1 to 100. It is not clear why MOIs greater than that required to infect 100% of the cells should increase yield. In suspension culture, studies to optimize the production of recombinant protein from Sf-9 cells demonstrated that MOI did not affect yield and could be chosen on the basis of technical needs or economic objectives without sacrificing yield (80). A very high MOI of 580 plaque-forming units/mL to infect high density cell cultures, 4–5 $\times$ 10^6 cells/mL, combined with medium replacement was used for optimum yields in an airlift fermentor (82). To obtain sufficient virus to inoculate a large bioreactor at these high MOIs requires concentration of culture supernatants, which may not be economical unless the increase in yield is substantial.

The time of infection also has significant effect on the yield of recombinant protein obtained. Vaughn (83) demonstrated that monolayer cultures that had reached confluency at the time of inoculation produced little or no progeny virus. This was later confirmed by Wood et al. (84), whose data suggested that contact inhibition was the cause. The optimal time for infection is related to the MOI used (81,85). It is a very narrow period and is dependent on the cell density and the growth stage of the cells at the time of inoculation. Maiorella et al. (34) found that the yield of recombinant protein was reduced by 60% if cultures with high cell densities were infected in the stationary phase rather than the log phase. King et al. (86) obtained similar results but also noted that in their studies only 50% of the cells were infected at the MOI used. They implied that this was because the uninfected cells continued to grow and depleted the nutrients. Once infected with a baculovirus, insect cells do not divide. In cultures with low cell densities inoculated with a low MOI, the uninfected cells will continue to divide and the progeny cells become infected by viruses produced from viral replication in the cells infected with the original inoculum. The cycle will continue until all cells are infected or until the nutrients are depleted. This is the common practice in batch culture systems. An advantage is that less seed virus must be produced outside the bioreactor. However, there is less control over the passage of virus, and the time required to complete the fermentation may be extended. If the protein produced is degraded by enzymes released from lysing cells, yields may be lower under these conditions even if the amount produced is high.

ACKNOWLEDGMENTS

My thanks to Dr. Jean R. Adams (retired) of this laboratory for kindly providing the electron micrographs in Figures 1 and 2 and to Dr. Dwight E. Lynn of this laboratory for the photograph used in Figure 3.

BIBLIOGRAPHY

1. R.M. Patterson, J.K. Selkirk, and B.A. Merrick, *Environ. Health Perspect.* **103**, 756–759 (1995).
2. P.D. Frieson and L.K. Miller, in W. Doerfler and P. Boehm, eds., *The Molecular Biology of Baculoviruses*, Springer-Verlag, Berlin, 1986, pp. 31–49.
3. V.A. Luckow and M.D. Summers, *Biotechnology* **6**, 47–55 (1988).
4. M. Kool and J.M. Viak, *Arch. Virol.* **130**, 1–16 (1993).
5. L.E. Volkman, G.W. Blissard, B.A. Keddie, R. Possee, and D.A. Theilmann, in F.A. Murphy, C.M. Fauquet, D.H.L. Bishop, S.A. Ghabrial, A.W. Jarvis, G.P. Martelli, M.A. Mayo, and M.D. Summers, eds., *Virus Taxonomy: Sixth Report of the International Committee on Taxonomy of Viruses*, Springer-Verlag, New York, 1995, pp. 104–113.
6. J.R. Adams, J.T. McClintock, and J.A. Couch, in J.R. Adams and J.R. Bonami, eds., *Atlas of Invertebrate Viruses*, CRC Press, Boca Raton, 1991, pp. 87–226.
7. R.R. Granados and K.A. Lawler, *Virology* **108**, 297–308 (1981).
8. D.A. Hammer, T.J. Wickman, M.L. Shuler, H.A. Wood, and R.R. Granados, in M.L. Shuler, H.A. Wood, R.R. Granados, and D.A. Hammer, eds., *Baculovirus Expression Systems and Biopesticides*, Wiley-Liss, New York, 1995, pp. 103–119.
9. R.R. Granados and B.A. Federici, eds., *The Biology of the Baculoviruses*, vols. I, II, CRC Press, Boca Raton, 1986.

10. J.L. Vaughn, R.H. Goodwin, G.J. Tompkins, and P. McCawley, *In Vitro* **13**, 213–217 (1977).
11. M.D. Summers and G.G. Smith, Texas Agriculture Experiment Station Bulletin No. 1555, 1987.
12. T.D.C. Grace, *Nature* (*Lond.*) **216**, 613 (1967).
13. A. Gröner, in R.R. Granados and B.A. Federici, eds, *The Biology of Baculoviruses*, vol. I, *Biological Properties and Molecular Biology*, CRC Press, Boca Raton, 1986, pp. 177–202.
14. W.F. Hink and D.R. Bezanson, in E. Kurstak, ed., *Techniques in Setting Up and Maintenance of Tissue and Cell Cultures*, Elsevier Biomedical, County Clare, Ireland, 1985, pp. C111/1–C111/30.
15. W.F. Hink and R.L. Hall, in J. Mitsuhashi, ed., *Invertebrate Cell System Applications*, vol. II, CRC Press, Boca Raton, 1989, pp. 269–293.
16. W.F. Hink, *Nature (Lond.)* **226**, 466–467 (1970).
17. H.G. Miltenburger, P. David, U. Mahr, and W. Zipp, *Z. Angew. Entomol.* **82**, 306–323 (1977).
18. R.R. Granados, L. Guoxun, A.C.G. Derksen, and K.A. McKenna, *J. Invertebr. Pathol.* **64**, 260–266 (1994).
19. W.F. Hink, D.R. Thomsen, A.L. Meyer, D.J. Davidson, and F.J. Castellino, *Biotechnol. Prog.* **7**, 9–14 (1991).
20. M.J. Betenbaugh, L. Balog, and P.-S. Lee, *Biotechnol. Prog.* **7**, 462–467 (1991).
21. T.J. Wickham, T. Davis, R.R. Granados, M.L. Shuler, and H.A. Wood, *Biotechnol. Prog.* **8**, 391–396 (1992).
22. T.R. Davis, T.J. Wickham, K.A. McKenna, R.R. Granados, M.L. Shuler, and H.A. Wood, *In Vitro Cell. Dev. Biol.* **29A**, 388–390 (1993).
23. L. Marz, F. Altmann, E. Staudacher, and V. Kubelka, in J. Montreuil, J.F.G. Vliegenthart, and H. Schachter eds., *Glycoproteins*, Elsevier, Amsterdam, 1995, pp. 543–563.
24. V.A. Luckow, in M.L. Shuler, H.A. Wood, R.R. Granados, and D.A. Hammer eds., *Baculovirus Expression Systems and Biopesticides*, Wiley-Liss, New York, 1995, pp. 51–90.
25. K. Kuroda, H. Geyer, R. Geyer, W. Doerfler, and H.-D. Klenk, *Virology* **174**, 418–429.
26. D.J. Davidson, M.J. Fraser, and F.J. Castellino, *Biochemistry* **29**, 5548–5590 (1990).
27. J. Sissom and L. Ellis, *Biochem. J.* **261**, 119–126 (1989).
28. D.R. Thomsen, L.E. Post, and A.P. Elhammer, *J. Cell. Biochem.* **43**, 67–79 (1990).
29. D.L. Jarvis, in M.F.A. Goosen, A.J. Daugulis, and P. Faulkner, eds., *Insect Cell Culture Engineering*, Dekker, New York, 1993, pp. 195–219.
30. A. Hoss, I. Moarefi, K.-H. Scheidtmann, L. Cisek, J.L. Corden, I. Dornreiter, A.K. Arthur, and E. Fanning, *J. Virol.* **64**, 4799–4807 (1990).
31. T.D.C. Grace, *Nature (Lond.)* **195**, 788–789 (1962).
32. G.R. Gardiner and H. Stockdale, *J. Invertebr. Pathol.* **25**, 363–370 (1975).
33. S.A. Weiss, G.C. Smith, S.S. Kalter, and J.L. Vaughn, *In Vitro* **17**, 495–502 (1981).
34. B. Maiorella, D. Inlow, A. Shauger, and D. Harano, *Biotechnology* **6**, 1406–1410 (1988).
35. S.A. Weiss, W.G. Whitford, and G. Godwin, in M.J. Fraser, ed., *Proceedings, Eighth International Conference on Invertebrate and Fish Tissue Culture*, Tissue Culture Association, Columbia, Md., 1991, pp. 153–159.
36. S.M. Deutschmann and V. Jäger, *Enzyme Microb. Technol.* **16**, 506–512 (1994).
37. P. Wang, R.R. Granados, and M.L. Shuler, *J. Invertebr. Pathol.* **59**, 46–53 (1992).
38. M. López-Ferber, W.P. Sisk, and R.D. Possee, in C.D. Richardson, ed., *Baculovirus Expression Protocols*, Humana Press, Totowa, N.J., 1995, pp. 25–63.
39. V.A. Luckhow and M.D. Summers, *Virology* **167**, 56–71 (1988).
40. Y. Matsuura, R.D. Possee, H.A. Overton, and D.H.L. Bishop, *J. Gen. Virol.* **68**, 1233–1250 (1987).
41. T.J. French and P. Roy, *J. Virol.* **64**, 1530–1536 (1990).
42. U. Weyer and R.D. Possee, *J. Gen. Virol.* **72**, 2967–2974 (1991).
43. K.M. Trotter and H.A. Wood, in C.D. Richardson, ed., *Baculovirus Expression Vectors*, Humana Press, Totowa, N.J., 1995, pp. 97–105.
44. P.A. Kitts, M.D. Ayers, and R.D. Possee, *Nucleic Acids Res.* **18**, 5667–5672 (1990).
45. R. Neutra, B.-Z. Levi, and Y. Shoham, *Appl. Microbiol. Biotechnol.* **37**, 74–78 (1992).
46. J.L. Vaughn and R.H. Goodwin, in J.A. Romerger, ed., *Beltsville Symposium on Agricultural Research, vol. I, Virology in Agriculture*, Alleheld Osmun, Monclair, N.J., 1977, pp. 109–116.
47. I.R. Cameron, R.D. Possee, and D.H.L. Bishop, *Trends Biotechnol.* **7**, 66–70 (1989).
48. U. Eberhard and K. Schügerl, *Dev. Biol. Stand.* **66**, 325–330 (1985).
49. N. Kioukia, M. Al-Rubeal, Z. Zang, A.N. Emery, A.W. Nienow, and C.R. Thomas, *Biotechnol. Lett.* **17**, 7–12 (1995).
50. W.F. Hink and E.M. Strauss, in E. Kurstak, K. Maramorosch, and A. Dübendorfer, eds., *Invertebrate Systems In Vitro*, Elsevier/North Holland, Amsterdam, 1980, pp. 27–33.
51. J. Tramper, J.B. Williams, and D. Joustra, *Enzyme Microb. Technol.* **8**, 33–36 (1986).
52. J.J. Chalmers, in M.L. Shuler, H.A. Wood, R.R. Granados, and D.A. Hammer, eds., *Baculovirus Expression Systems and Biopesticides*, Wiley-Liss, New York, 1995, pp. 175–204.
53. K. Trinh, M. Garcia-Briones, F. Hink, and J.J. Chalmers, *Biotechnol. Bioeng.* **43**, 37–45 (1994).
54. J.L. Vaughn, in *Proc. Second Int. Colloq. Invertebr. Tissue Cult., Tremezzo, Como*, Instituto Lombardo, Milan, 1967, pp. 119–125.
55. D.W. Murhammer and C.F. Goochee, *Biotechnology* **6**, 1411–1418 (1988).
56. L.K. Overton and T.A. Kost, in M.L. Shuler, H.A. Wood, R.R. Granados, and D.A. Hammer, eds., *Baculovirus Expression Systems and Biopesticides*, Wiley-Liss, New York, 1995, pp. 233–242.
57. M. Silberklang, K. Ramasubramanyan, S.L. Gould, A.B. Lenny, T.C. Seamen, S. Wang, G.R. Hunt, B. Junker, K.E. Mazina, M.R. Tota, O. Palyha, and D. Jain, in M.L. Shuler, H.A. Wood, R.R. Granados, and D.A. Hammer, eds., *Baculovirus Expression Systems and Biopesticides*, Wiley-Liss, New York, 1995, pp. 205–231.
58. A. Kumar and M.L. Shuler, *Biotechnol. Prog.* **11**, 412–419 (1995).
59. J.W. Rice, N.B. Ranki, T.M. Gurganus, C.M. Marr, J.B. Barna, M.M. Walters, and D.J. Burns, *Biotechniques* **15**, 1052–1059 (1993).
60. J.M. Vlak, E.J. Schaeger, and A.R. Bernard, *Baculovirus and Recombinant Protein Production Processes*, Editiones Roche, Basel, Switzerland, 1992.
61. D.G. Bergmann, in A.S. Lubinlechi, ed., *Large-Scale Mammalian Cell Culture Technology*, Dekker, New York, 1990, pp. 417–449.

62. S. Reuveny, Y.J. Kim, C.W. Kemp, and J. Shiloach, *Appl. Microbiol. Biotechnol.* **38**, 619–623 (1993).
63. R.A. Taticek, D.A. Hammer, and M.L. Shuler, in M.L. Shuler, H.A. Wood, R.R. Granados, and D.A. Hammer, eds., *Baculovirus Expression Systems and Biopesticides*, Wiley-Liss, New York, 1995, pp. 131–174.
64. C. Bédard, R. Tom, and A. Kamin, *Biotechnol. Prog.* **9**, 615–624 (1993).
65. L. Öhman, J. Ljunggren, and L. Häggström, *Appl. Microbiol. Biotechnol.* **43**, 1006–1013 (1995).
66. D.A. Lindsay and M.J. Betenbaugh, *Biotechnol. Bioeng.* **39**, 614–618 (1992).
67. J. Mitsuhashi, in J. Mitsuhashi, ed., *Invertebrate Cell System Applications*, vol. I, CRC Press, Boca Raton, 1989, pp. 3–20.
68. M.-Y. Wang, S. Kwong, and W.E. Bentley, *Biotechnol. Prog.* **9**, 355–361 (1993).
69. S.H. Lee and T.H. Park, *Biotechnol. Lett.* **16**, 327–332 (1994).
70. M. Klöppinger, G. Fertig, E. Fraune, and H.G. Miltenburger, *Cytotechnology* **4**, 271–278 (1990).
71. J.L. Zhang, N. Kalogerakis, L.A. Behie, and K. Iatrou, *Biotechnol. Bioeng.* **42**, 357–366 (1993).
72. S.A. Weiss and J.L. Vaughn, in R.R. Granados and B.A. Federici, eds., *The Biology of Baculoviruses*, vol. II, CRC Press, Boca Raton, 1986, pp. 63–87.
73. A.W. Caron, R.L. Tom, A.A. Kamen and B. Massie, *Biotechnol. Bioeng.* **43**, 881–891 (1994).
74. E.A. MacKinnon, J.F. Henderson, D.B. Stoltz, and P. Faulkner, *J. Ultrastruct. Res.* **49**, 419–435, (1974).
75. J.M. Slavicek, M.J. Mercer, M.E. Kelly, and N. Hayes-Plazolles, *J. Invertebr. Pathol.* **67**, 153–160 (1996).
76. T.J. Wickham, T. Davis, R.R. Granados, D.A. Hammer, M.L. Shuler, and H.A. Wood, *Biotechnol. Lett.* **13**, 483–488 (1991).
77. C.D. de Gooijer, R.H.M. Koken, F.L.J. van Lier, M. Kool, J.M. Vlak, and J. Tramper, *Biotechnol. Bioeng.* **40**, 537–548 (1992).
78. F.L.J. van Lier, J.P.T.W. van den Hombergh, C.D. de Gooijer, M.M. den Boer, J.M. Vlak, and J. Tramper, *Enzyme Microb. Technol.* **18**, 460–466 (1996).
79. E.M. Dougherty, R.M. Weiner, J.L. Vaughn, and C.F. Reichelderfer, *Appl. Environment. Microbiol.* **41**, 1166–1172 (1981).
80. M. Brown and P. Faulkner, *J. Invertebr. Pathol.* **26**, 251–257 (1975).
81. P. Licari and J.E. Bailey, *Biotechnol. Bioeng.* **37**, 238–246 (1991).
82. J.E. Lazarte, P.-F. Tosi, and C. Nicolau, *Biotechnol. Bioeng.* **40**, 214–217 (1992).
83. J.L. Vaughn, *J. Invertebr. Pathol.* **28**, 233–237 (1976).
84. H.A. Wood, L.B. Johnston, and J.P. Burand, *Virology* **119**, 245–254 (1982).
85. J.F. Power, S. Reid, K.M. Radford, P.F. Greenfield, and L.K. Nielsen, *Biotechnol. Bioeng.* **44**, 710–719 (1994).
86. G.A. King, A.J. Daugulis, P. Faulkner, and M.F.A. Goosen, *Biotechnol. Prog.* **8**, 567–571 (1992).

See also AMMONIA TOXICITY, ANIMAL CELLS; CHINESE HAMSTER OVARY CELLS, RECOMBINANT PROTEIN PRODUCTION; CULTURE MEDIA, ANIMAL CELL, LARGE-SCALE PRODUCTION ; EXPRESSION SYSTEMS, *E. COLI*; EXPRESSION SYSTEMS, MAMMALIAN CELLS; GLYCOSYLATION OF RECOMBINANT PROTEINS; INSECT CELLS AND LARVAE, GENE EXPRESSION SYSTEMS; PROTEIN EXPRESSION, SOLUBLE.

INSECT CELLS AND LARVAE, GENE EXPRESSION SYSTEMS

MICHAEL J. BETENBAUGH
ERIC AILOR
ERIK M. WHITELEY
Johns Hopkins University
Baltimore, Maryland
TSU-AN HSU
Vaccine Pharmaceutical R&D, Merck & Co., Inc.
West Point, Pennsylvania

KEY WORDS

Baculovirus
Bioreactor
Cell culture
Expression system
Glycosylation
Insect
Larva
Recombinant DNA

OUTLINE

INTRODUCTION

Insects and other invertebrates include more than half of the animal species that have been identified on the planet. A number of these invertebrates are valuable to the clothing industry where they are used in the production of silk. From an agricultural perspective, however, insects are a major pest of crops and other plant life. Developing methods to control the proliferation of insects has been a continuing aim of the agricultural and forest service industries. As with other animal species, these hosts are capable of maintaining and transmitting viral infections. One approach that has been considered for many years to control insect proliferation is the application of insect-specific viruses. As a result, a substantial body of knowledge has evolved over the past 20 years concerning the structure and function of insect viruses. Through these biological advances, it is now possible to alter the genetic content of these insect-specific viruses and incorporate heterologous genes from other organisms. The development of genetically altered insect viruses has had a tremendous impact on the development of enhanced viral biopesticides and on the production of recombinant DNA products of biotechnological interest. Baculoviruses such as the *Autographica californica* nuclear polyhedrosis virus (AcMNPV or AcNPV) and the *Bombyx mori* nuclear polyhedrosis virus (BmNPV) can be modified to incorporate genes from mammalian, insect, plant, and viral sources, among others. Insect cells or insect larvae are then infected with these altered viruses to generate recombinant proteins for applications such as structure and function studies, diagnostics, and vaccines. To generate the large quantities of recombinant protein and virus needed for biotechnological and biopesticide applications, efficient and scaleable production processes have been developed. These processes must be particularly sensitive to the characteristics of the baculovirus expression vector system (BES or BEVS), which includes both a cell growth and infection stage. Various insect cell hosts, viral vectors, and expression methodologies are being explored along with alternative media formulations and different bioreactor operating strategies. These efforts are examined in the following sections.

INSECT CELL LINES

Insect cell lines have been established from nearly 100 different species within the following seven orders of Arthropoda: Lepidoptera, Diptera, Homoptera, Hymenoptera, Orthoptera, Hemiptera, and Coleoptera. These lines are isolated from eggs or embryonic, larval, or adult tissues such as hemocytes, ovaries, imaginal discs, testes, midgut, and fat body using methods that are now well known (1). The procedures typically involve the lysis of tissue followed by suspension in a rich serum-containing medium in tissue culture flasks. The nutrients are replaced periodically, and the cultures are monitored until cell division is observed. The cells are then subcultured after one to several weeks to obtain a new cell line. Eventually, the cell lines are screened for desired properties such as short doubling times, virus susceptibility, recombinant gene expression, or posttranslational processing. Clonal isolates can be obtained that have better growth or gene expression properties, as was the case in the development of the Sf-9 clone from the established cell line (IPLB-Sf-21) of the fall armyworm, *Spodoptera frugiperda*. Both of these cell lines are now widely used for the replication of AcMNPV and production of recombinant proteins.

The screening of established cell lines led to the identification of a *Trichoplusia ni* (cabbage looper) cell line (BTI-TN5B1-4), with the trade name High Five™ (Invitrogen, Inc.), which often provides recombinant protein yields superior to those of Sf-9 and Sf-21 (2). Originally an attachment-dependent cell line, High Five cells have been adapted to suspension growth in many laboratories. Adaptation to suspension is aided by the addition of 50–2,000 U/cm^3 heparin to prevent clumping, but the heparin should be removed prior to viral infection and expression of products (3,4). Shown in Figure 1 is a micrograph of healthy and AcMNPV-infected Sf-9 and High Five cells. Commercial serum-free media have been developed for High Five, Sf-21, and Sf-9 cell lines. Because the Sf-9 cell line produces higher levels of the extracellular virus, this cell line is often used for virus production, whereas the High Five are used for heterologous gene expression. Although High Five and Sf are the cell lines of choice for many applications, recombinant protein production from AcMNPV has been achieved in a number of other cell lines such as those isolated from *Heliothis*, *Lymantria*, *Plutella*, and *Trichoplusia* (5,6). A smaller variety of cell lines from *Bombyx mori* (silkworm) and *Lymantria dispar* (gypsy moth) are available for infection with BmNPV and *Lymantria dispar* nuclear polyhedrosis virus (LdNPV), respectively.

Cell lines are usually cultured at the laboratory scale in polystyrene culture flasks or roller bottles. Those cell lines that are not anchorage dependent can be grown in shaker flasks or small suspension-culture flasks. Cells are monitored for number density and passaged one to two times per week by subculturing into fresh medium. Attachment-dependent cells must be dislodged by agitation, scraping, or enzymatic treatment. Following a lag phase, the cells grow logarithmically with doubling times between 0.5 and 2 days. Maximum cell densities are typically in the range of 10^6 cells/mL but may rise as high 10^7 cells/mL in some cases. Cell size is in the range of 10 to 20 μm, and culture temperatures are typically between 25 and 30 °C. Stocks of cells can be cryopreserved in liquid nitrogen using medium and dimethylsulfoxide (DMSO). Recovery is achieved by a rapid thaw and can be enhanced by using a serum-bearing medium prior to freezing the cells.

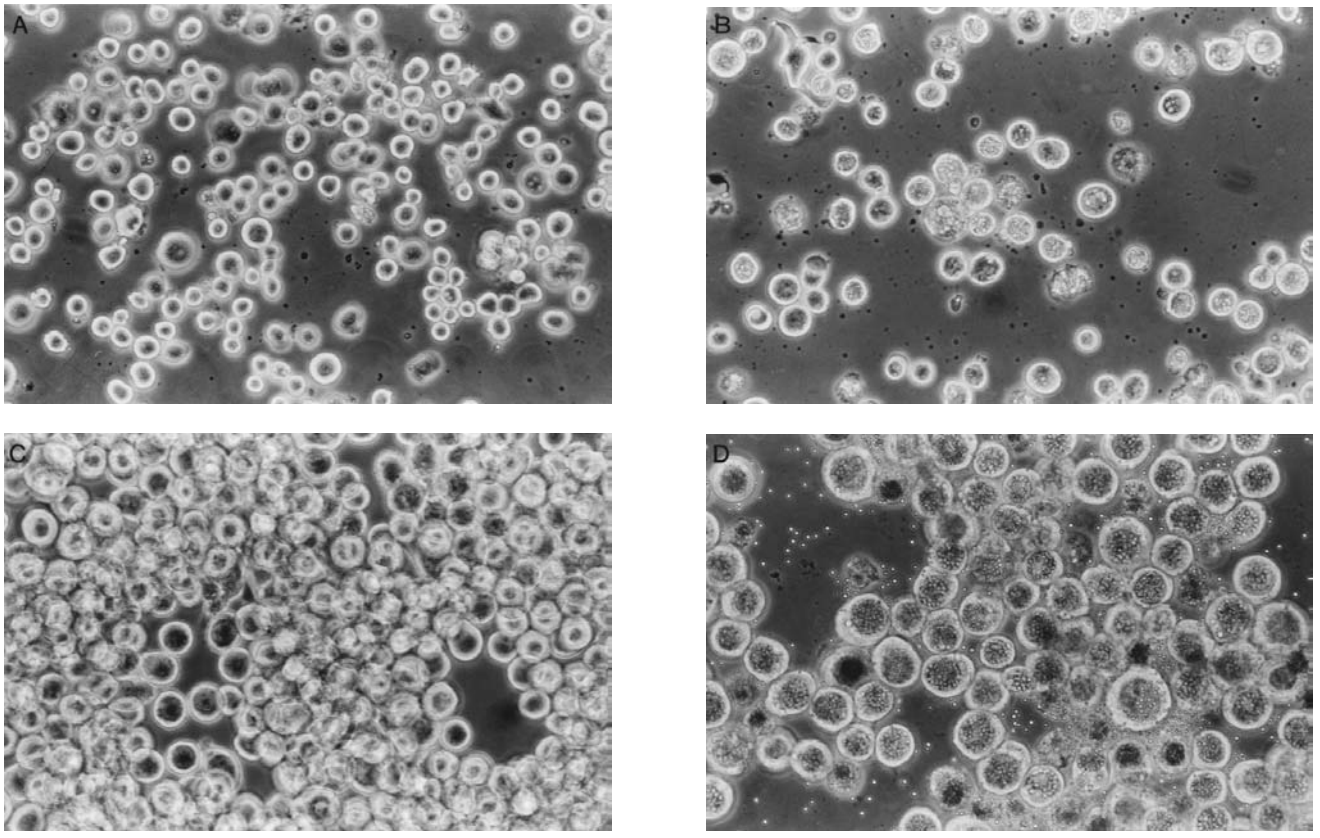


Figure 1. Insect cell morphology. (**a**) Healthy Sf-9 cells grown in insect culture medium. (**b**) Sf-9 infected with wild-type AcMNPV baculovirus. (**c**) Healthy BTI-TN5B1-4. (**d**) BTI-TN5B1-4 infected with wild-type AcMNPV baculovirus. Photographs taken at 32×.

VIRUSES

The most widely studied and utilized insect viruses are the baculoviruses, a helper-independent eukaryotic DNA vector that can infect a wide range of lepidopteran (butterflies and moths) insects and insect cells. "Baculo" refers to the rod-shaped capsid that is characteristic of these viruses, which are 40–50 nm in diameter and 200–400 nm in length. The two most widely studied baculoviruses, AcMNPV and BmNPV, contain DNAs that are approximately 130 kb in length. The relatively large size of the virus allows for the encapsulation of large pieces of foreign DNA in addition to the essential viral DNA.

The infection of a single cell by the baculovirus is illustrated in Figure 2. Baculoviruses are deposited in two forms, budded and occluded, in the infected cells (7,8). At the initial stages of infection, membrane-enveloped viruses bud from the plasma membrane of an infected host 10–14 h after infection in a form commonly referred to as extracellular virus (EV), nonoccluded virus (NOV), or budded virus (BV). A second form of the virus, the occluded virus (OV), forms in the nucleus of an infected cell beginning at 20–24 h after viral infection. This intracellular virus is enveloped and encased within a crystalline viral protein matrix called polyhedrin. These protein crystals, called occlusion bodies, polyhedra, or polyhedral inclusion bodies (PIBs), may be as large as 5 μm in diameter and contain many virions (Fig. 3). The OV encased within the polyhedrin protein matrix may contain multiple nucleocapsids within a single envelope for multiply embedded nuclear polyhedrosis viruses (MNPVs), or a single nucleocapsid per envelope for single nuclear polyhedrosis viruses (SNPVs). In addition to nuclear polyhedrosis virus (NPV), a second type of occluded baculovirus, the granulosis virus (GV), includes one or, very rarely, two virions embedded in a much smaller protein matrix called a capsule or granule. Most baculoviruses used as biopesticides and all baculoviruses used in biotechnology are NPVs, although a few GVs have shown potential for future applications.

It is the OV that is responsible for transmission of the virus from insect to insect, as shown in the baculovirus life cycle depicted in Figure 4 (8). After death and decomposition of an infected insect, the polyhedra are deposited on vegetation where they are ingested by other feeding insect larvae. Transported with food to the midgut, the polyhedra are dissolved in the alkaline environment, subsequently releasing virions into the lumen of the midgut. These viruses then fuse to the epithelial cells where they cause an infection in the midgut tissue. The infected cells then release BV, which are responsible for secondary infection of other tissues by spreading through the hemolymph and trachea. The infected larva continues to feed for 5–7 days until it becomes listless. The larval cuticle melanizes, and the muscular structure dissolves in a process called melt-

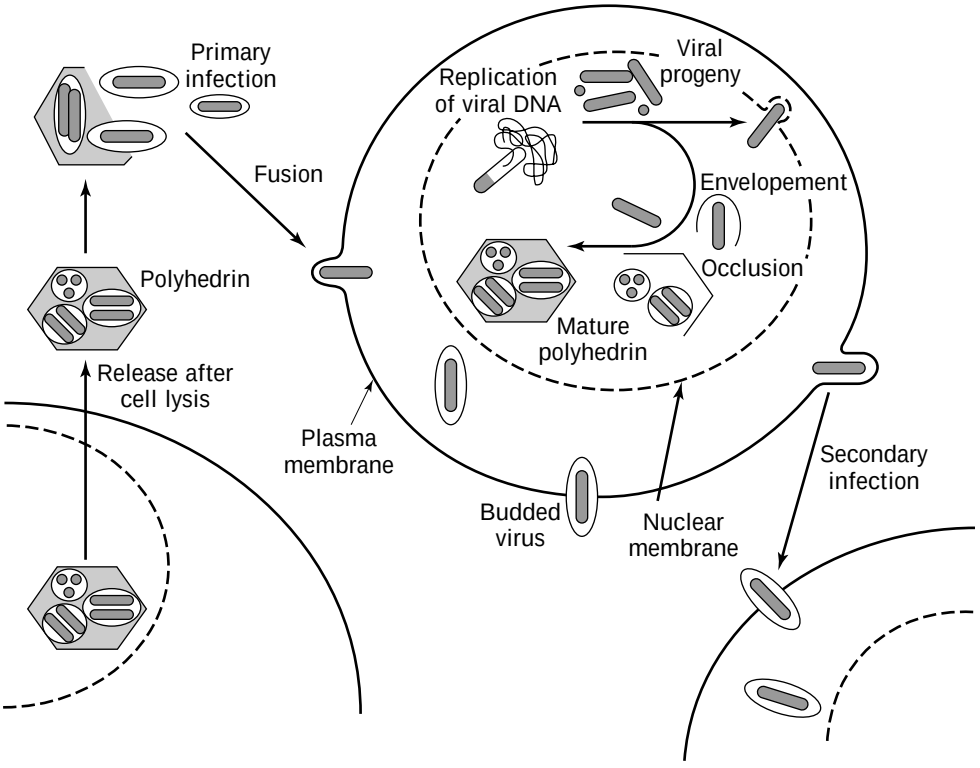


Figure 2. Single-cell replication life cycle of the natural baculovirus, AcMNPV. Primary infection by occluded virus occurs upon disintegration of the polyhedrin matrix and subsequent fusion to the plasma membrane. Budded virus produced within the cell proceed to initiate the secondary infection process. *Source*: Adapted from figure appearing in *Baculovirus Expression Systems and Biopesticides*, M.L. Shuler, H.A. Wood, R.R. Granados, and D.A. Hammer, Wiley-Liss, 1995. Reprinted by permission of John Wiley & Sons, Inc.

ing or wilting. The larva eventually ruptures, releasing about 25% of its dry weight in the form of PIBs, which are spread over vegetation for consumption by other insects and reinitiation of the infection cycle.

BACULOVIRUS GENE EXPRESSION

Baculovirus gene expression of viruses such as AcMNPV occurs over four phases: immediate early (α), early (β), late (γ), and very late (δ). During the immediate early phase, viral genes are expressed using host cell factors. The immediate early gene products then control the expression of the early genes, whose products are responsible for cytoskeletal rearrangement and nucleus enlargement. Extensive viral DNA replication occurs about 5–7 h after infection, followed by expression of the late genes. The late phase extends from 6 until 20–24 h postinfection and involves the assembly of nucleocapsids, the transport and release of budded virus (BV), and the degradation of the host genome. The formation of occluded viruses occurs during the very late phase, beginning about 20–24 h postinfection. The occluded viral particles can accumulate for 3–5 days until the infected cells lyse.

Two genes abundantly expressed during the very late phase are the polyhedrin and p10 genes. The product of the polyhedrin gene (*polh*) is the 29-kDa protein that forms the crystalline matrix of the occlusion particles protecting encased virus nucleocapsids. The product of the p10 gene is a 10-kDa protein that forms large fibrous arrays in the nucleus and occasionally in the cytoplasm. In 1983, it was observed that viruses containing deletions in the polyhedrin gene were stable and could be replicated in cells in culture (9). Though the cells infected with these viruses lack the characteristic occlusion particles, production of BV was unaffected. Furthermore, these BV could be collected from the cell culture media and used to infect other cells in order to passage the altered virus. Because the polyhedrin gene is nonessential for BV infection and propagation in cell culture, the construction of a recombinant baculovirus can be accomplished by replacing the polyhedrin gene with a foreign gene such as interferon-β (10). In this way, the foreign gene, now under the control of the very strong polyhedrin promoter, is expressed, instead of the polyhedrin gene at the very late phase of AcMNPV infection. Similarly, it was observed that disruption of the p10 gene resulted in disruption of the fibrous matrix, but the formation of BV was not affected. Subsequently, recombinant baculovirus constructs containing foreign genes inserted under the control of the p10 promoter were used to produce heterologous proteins in place of the p10 protein (11). Because the polyhedrin and p10 are very strong late promoters, with as much as 50% of the total stained protein consisting of polyhedrin (12), these promoters have

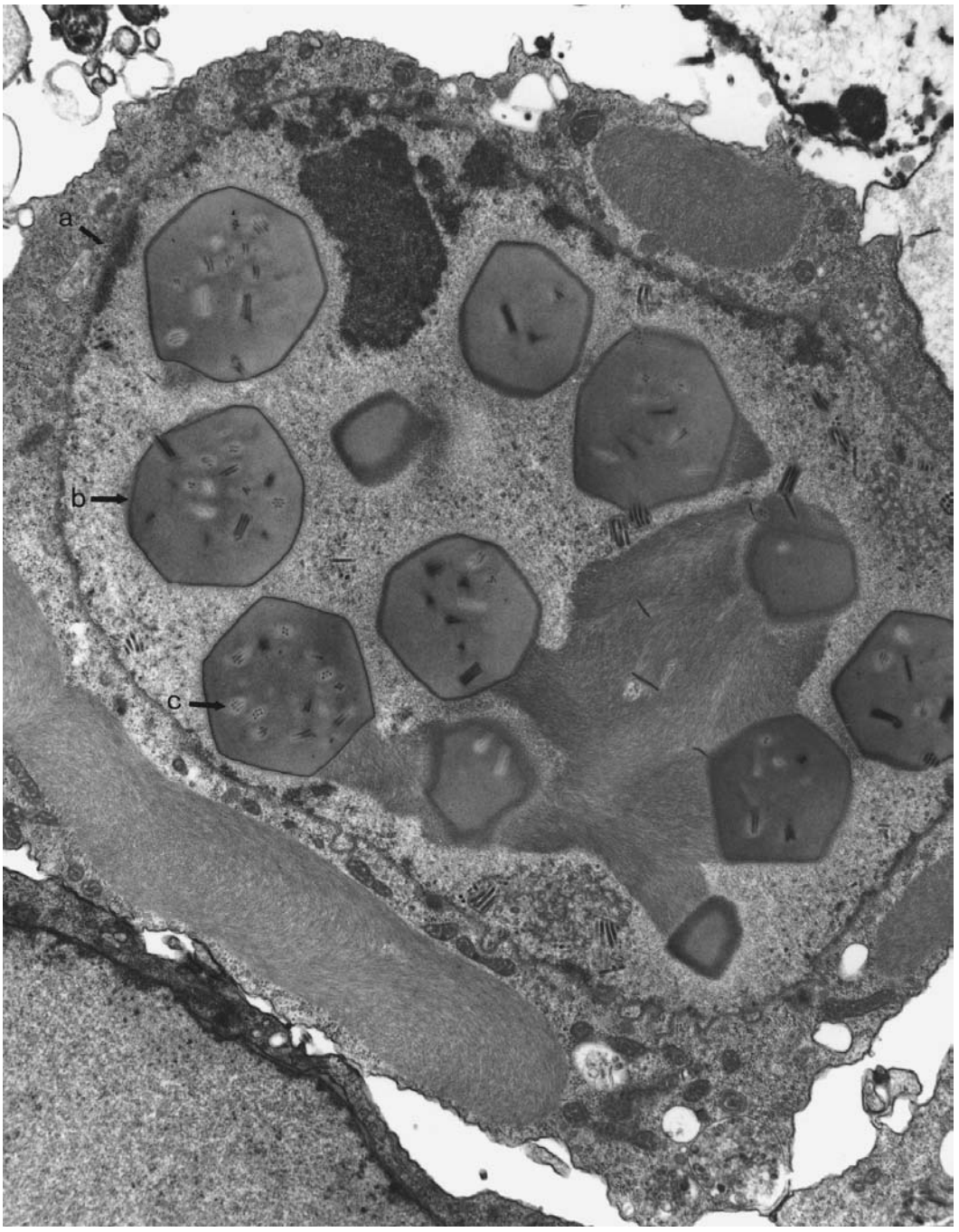


Figure 3. Electron micrograph of AcMNPV infection in *S. frugiperda*. (a) Nuclear membrane, (b) polyhedrin inclusion bodies, and (c) virions at 48 h postinfection (magnification 15,000×). *Source*: Micrograph courtesy of G. Williams.

the potential for very high rates of foreign gene expression. One disadvantage of this system is that production of heterologous protein is accomplished late in the viral infection, at which time host cell function is deteriorating. Furthermore, the insect host eventually succumbs to the infection, thus requiring a new batch of insect cells to be grown and infected each time recombinant protein is to be produced.

In an effort to avoid the problems associated with insect cell death due to virus infection, a number of vectors have been developed that allow the stable integration of foreign DNA into insect cell hosts. Stable expression of heterologous proteins using these integrated vectors has been achieved in several insect cell lines including Sf-9, *Drosophila* S2, and lines from *Aedes albopictus* (13). However, expression levels are in most cases many times lower than

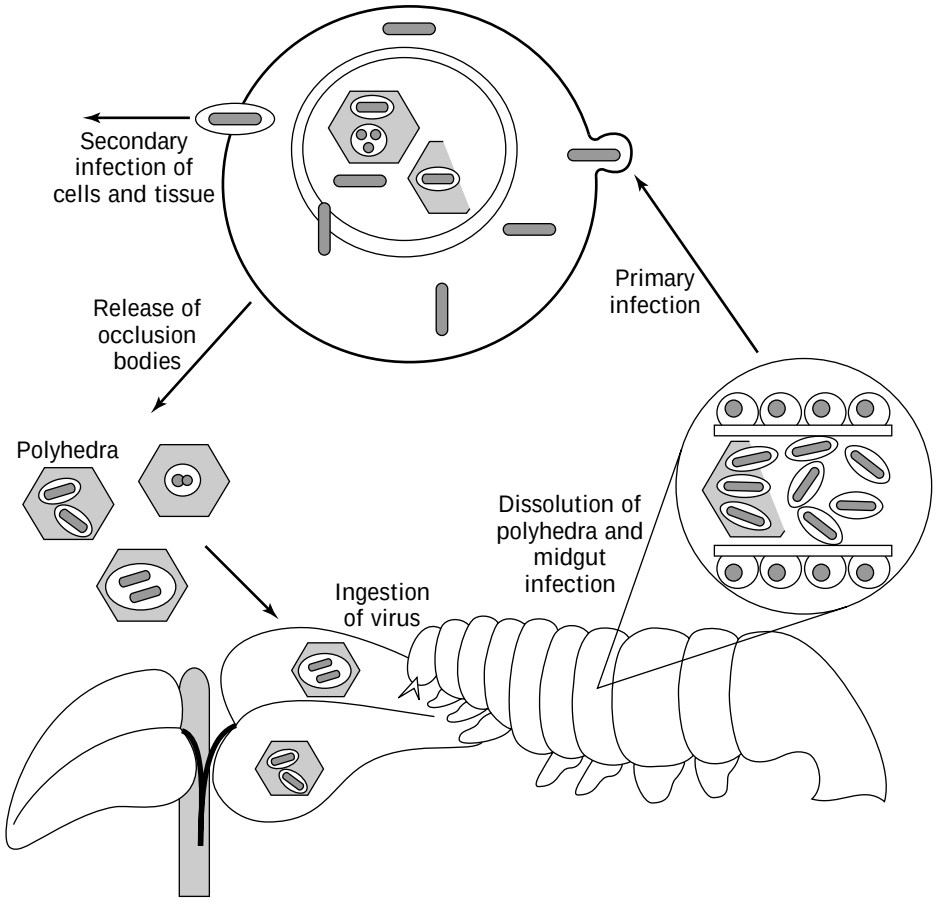


Figure 4. Life cycle of the natural baculovirus, AcMNPV. *Source*: Adapted from figure appearing in *Baculovirus Expression Systems and Biopesticides*, M.L. Shuler, H.A. Wood, R.R. Granados, and D.A. Hammer, Wiley-Liss, 1995. Reprinted by permission of John Wiley & Sons, Inc.

with baculovirus, so these systems have yet to approach the popularity of BES. Stable expression systems are likely to find niche applications when lower expression is acceptable in return for desirable host processing functions.

CONSTRUCTION OF RECOMBINANT BACULOVIRUSES

More than 500 different baculoviruses of the Baculoviridae family are currently known. For use as a vector to deliver foreign genes into insect cells, the most extensively used is AcMNPV (14). BmNPV (15) has also been used but to a much lesser extent, and work is currently in progress to develop expression systems based on other baculoviruses including LdNPV and granulosis viruses. Techniques for manipulating AcMNPV to generate recombinant baculoviruses is discussed here because this subtype is the most widely investigated and the procedures are similar for many other baculoviruses. AcMNPV is a large, double-stranded, circular DNA virus for which the complete nucleotide sequence has been determined (16). The large size of the baculovirus DNA prohibits the typical recombinant DNA manipulations used for cloning with bacterial plasmids. Consequently, an alternative approach involving plasmid transfer vectors and allelic replacement (homologous recombination) was developed and remains the most popular method for foreign gene insertion into the baculovirus. Many improvements have been made for generating recombinant baculoviruses, but almost all are based on homologous recombination.

The first step in allelic replacement involves the manipulation of a transfer plasmid in *Escherichia coli* by cloning and subcloning techniques to insert a foreign gene downstream from a baculovirus promoter, typically those for polyhedrin or p10. The promoter and gene are also flanked by baculovirus DNA derived from a nonessential locus such as polyhedrin or p10 genes. The recombinant transfer plasmid is then mixed with native baculovirus DNA, and the mixture is transfected into insect cells using calcium phosphate or cationic lipids. A small fraction of the baculoviruses will undergo homologous recombination with the transfer plasmid to create a recombinant baculovirus containing a heterologous gene under the control of the appropriate promoter (Fig. 5). The recombinant baculovirus constructs may be identified by examining for the absence or presence of a particular baculovirus or plasmid function due to the insertion of the foreign gene.

The initial screening technique involved identification of plaques of insect cells containing occlusion-minus (occ^-) infections as a result of the deletion of the polyhedrin gene from the original baculovirus DNA. Occlusion-positive plaques include an opaque light refractile characteristic that can be visually distinguished from the characteristics of occ^- plaques under the microscope. However, picking an occ^- clone is a formidable task for inexperienced researchers because young nonrecombinant plaques often show no

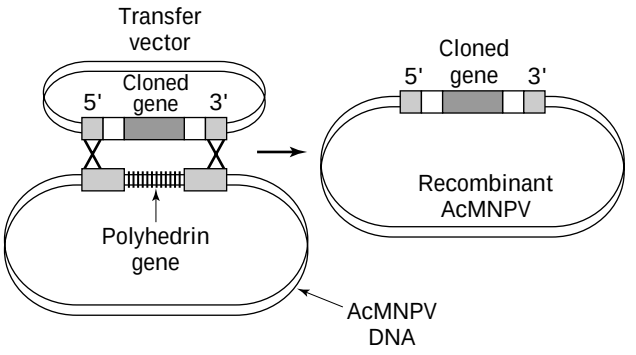


Figure 5. Homologous recombination. A baculovirus gene locus (usually p10 or polyhedrin) is targeted for homologous recombination using a transfer vector that includes flanking homologous sequences and, in some cases, encoding a selectable marker. The recombination event (represented here by X's) results in the replacement of sequences flanked by the homologous areas with the sequence of the foreign gene.

sign of "occlusion," and recombination frequencies can be as low as 0.1%. These difficulties resulted in the time consuming and tedious isolation of recombinant baculovirus through multiple rounds of plaque purification. The first modification of the system was the use of the bacterial gene *lacZ* to produce chromogenic recombinant plaques that enhance plaque purification (17,18). Significant progress in isolating recombinant baculovirus was achieved by derivatizing the circular, double-stranded wild-type baculovirus into a linear form through the introduction of a unique restriction site (Bsu36I) at the polyhedrin locus (19). A higher proportion of recombinant viruses was obtained when the linearized baculovirus genome was cotransfected with a transfer vector, because the infectivity of the linearized virus alone was substantially reduced below wild-type levels. More recently, a faster approach that can generate pure recombinant baculovirus in 7–10 days was reported (20). This technique uses a novel baculovirus shuttle vector (bacmid) that can replicate in *E. coli* and infect susceptible lepidopteran insects. The homologous recombination event and screening for the recombined bacmid occur entirely in *E. coli*. A recombinant baculovirus is then generated by introducing the composite bacmid into insect cells in culture and harvesting the virus from the medium. This approach eliminates the need for plaque purification because only the recombinant baculovirus is generated during the initial infection.

Other expression vectors are available that provide for the expression of up to four genes in the same baculovirus construct. These constructs may include one or more polyhedrin, p10, or heat shock promoters to control expression of the foreign genes. Infection of insect cells with several viruses, each containing multiple genes, can provide simultaneous expression of numerous proteins in one insect cell. Constructs containing baculovirus promoters such as IE1 and core that are activated earlier in the infection are also available (21). Specialized vectors that include histidine tags, which help facilitate protein purification, as well as gp64 or melittin fusions, for enhancing protein secretion, are also available. For a laboratory not experienced with BEVS, it is advisable to employ the linear baculovirus genome with a lethal deletion and selectable marker or the *E. coli* bacmid shuttle vector as the initial method for generating recombinants. Many kits are now commercially available that include the transfer vectors and baculovirus DNA for generating recombinant baculovirus constructs.

POSTTRANSLATIONAL PROCESSING

Because insects are eukaryotic organisms, the derived cell lines are capable of performing many of the posttranslational processes including secretion, folding, and glycosylation observed in higher organisms. In fact, insect cells are often the expression system of choice for the rapid generation of functional eukaryotic proteins. However, the posttranslational processing may not be identical to that performed in the native environment. It may be worthwhile to compare the protein expressed in the insect cells with the native protein to determine if significant differences in posttranslational processing exist that will detract from subsequent structural or functional analysis. Some of the posttranslational processing events that are observed in insect cells are described in the following sections.

Glycosylation

A majority of secreted and membrane proteins are glycosylated, and the attached oligosaccharides can play critical roles in cellular targeting, structural stability, resistance to proteolysis, immunogenicity, and circulatory half-life. The oligosaccharide residues are added to the protein at specific amino acids during processing in the endoplasmic reticulum (ER). Subsequently, these oligosaccharides are trimmed and modified during processing in the ER and Golgi compartments. Both N-linked and O-linked glycosylation have been observed in insect cells.

During N-glycosylation in both insect and mammalian cells an initial $Glc_3Man_9GlcNAc_2$ group is added to the Asn residue of the Asn-X-Ser/Thr sequence. Glucose and mannose residues are then trimmed from the $Glc_3Man_9GlcNAc_2$ group. However, the level of trimming and subsequent modifications are likely to be different in insect cells than those observed in mammalian cells. Recent evidence indicates that these differences may depend significantly on the particular insect cell line used. Many initial studies of *N*-glycans from heterologous glycoproteins obtained from Sf-9 or Sf-21 cells indicated the presence of only high-mannose type ($Man_{9\text{-}5}GlcNAc_2$) or short, truncated, paucimannosidic ($Man_{3\text{-}2}GlcNAc_2$) oligosaccharides, sometimes including an α(1,6)-linked fucose (22,23). In contrast, studies with recombinant human plasminogen indicated that insect cells could synthesize complex N-linked oligosaccharides (24). Several more recent studies on heterologous glycoproteins have indicated that certain insect cell lines can synthesize hybrid and complex oligosaccharides. Analysis of *N*-glycans of heterologous interferon-γ secreted from *Estigmena acrea* (Ea-4) confirmed the presence of complex, hybrid, paucimannosidic, and high-mannose N-linked oligosaccharides (25). In con-

trast, *Spodoptera frugiperda* (Sf-9) cells were observed to produce only hybrid, paucimannosidic, and high-mannose *N*-glycans. A complete description of the *N*-glycans for intracellular and secreted IgG_{2a} from High Five cells indicated the secretion of complex, hybrid, and paucimannosidic glycoforms (26), along with the intracellular retention of glycoforms that were principally high mannose. A comparison of the *N*-glycans of IgG from High Five cells with those attached to an IgG_{3a} from the mammalian NS/0 cells reveals significant similarities and some differences in the glycosylation properties of the two cell lines (Table 1). Based on the intracellular and extracellular *N*-glycans of IgG, a putative pathway for oligosaccharide processing in *T. ni* cells has been proposed (Fig. 6). The initial mammalian ER pathway, including trimming of the terminal glucose and mannose residues, seems operative. Insect cell processing diverges from mammalian cell processing in the pathways leading to the paucimannosidic structures, with the critical step likely to be the cleavage of the GlcNAc from the α-(1,3) branch by a possible Golgi β-*N*-acetylglucosaminidase (28). Other significant differences in insect cell processing are the α-(1-3)-linked addition of fucose, also observed in the processing pathway of plants, and the absence of any attached sialic acids (25,26,29). Recently, glycoprotein processing has been manipulated in insect cells by the expression of glycosylation pathway enzymes in order to generate glycoforms even more similar to those produced by mammalian cells (21,30).

O-glycosylation, which involves the addition of carbohydrates to hydroxylated amino acids such as serine or threonine residues, has been observed for several recombinant proteins produced by baculovirus-infected insect cells. A pseudorabies viral glycoprotein, gp50, and extracellular domains of the respiratory syncytial virus protein were found to contain the attached monosaccharide GalNAc and lesser amounts of the disaccharide Galβ1-3GalNAc (31). In mammalian cells, only the disaccharide form was observed in the presence or absence of attached sialic acids. Activity levels of the relevant glycosyltransferases in the insect cells and mammalian cells supported the particular distribution patterns observed.

Other Posttranslational Processes

Many proteins expressed in insect cells are directed to the appropriate cellular compartment including the nucleus, cytoplasm, organelles, plasma membrane, or extracellular space. For secreted proteins, a short (15–40 amino acids) signal peptide is present for translocating polypeptide from the cytoplasm into the endoplasmic reticulum (ER). This signal peptide, whether from mammalian, plant, or yeast sources, is usually cleaved correctly by the signal peptidases in insect cells. However, signal peptide processing is not always efficient, especially for peptides from plants or other distant eukaryotes. In fact, a complete absence of processing has been observed in at least one case (32). Replacement of the original signal peptide with another signal peptide from a protein such as melittin of honeybee venom (33) or the gp67 and EGT proteins (7) from AcMNPV can help to alleviate this limitation in some cases. Sometimes, a substantial fraction of proteins destined for secretion will be retained intracellularly in incorrectly folded forms or as aggregates. Coexpression of chaperones or folding enzymes can reduce aggregation and enhance secretion of recombinant proteins that rely on such folding assistance factors (34). Other posttranslational processing steps including phosphorylation, fatty acid acylation, α-amidation, and N-terminal amidation have also been observed in insect cells or insect larvae.

APPLICATIONS

Baculovirus-infected insect cells often provide high levels of expression of heterologous proteins, leading to the use of this system across a broad spectrum of applications and fields. BES has been used for the production of proteins for structure and function studies, vaccines, multisubunit assemblies, and diagnostic reagents (35). Furthermore, insect cells have been exploited as tools for examining the role of proteins in physiological processes. The application of BES products as therapeutics is also under consideration, though it is not currently a major application of the technology. Finally, the use of baculoviruses as biopesticides, the original motivation for studying these viruses, represents a major thrust of the technology (4,36). Wild-type baculoviruses have been used as efficacious pest control agents for many years but have a limited host range and can take between 5 and 10 days to debilitate insect larvae. A number of genetic modifications, mostly centered on replacement of polyhedrin or p10 genes with heterologous insect toxins, have been implemented to lower viral killing time and reduce the duration of pest feeding. Some of the heterologous genes that have been considered are listed in Table 2. A major concern in the development of baculovirus pesticides is the potential safety of these agents. However, extensive studies have concluded that baculoviruses have a very low range of infectivity and will not infect beneficial insects or mammals. Inclusion of a foreign gene under the control of a baculovirus promoter appears to present little danger as well.

LARGE-SCALE PROCESSING: CELL CULTURE OR LARVAL PRODUCTION

Structure and functional characterization studies and diagnostic reagents often require product levels beyond the scale of small shaker flasks. Biopesticides will require huge amounts of the final product for use in field trials. The production of these large quantities of heterologous proteins or baculoviruses will require the use of large-scale cell culture or insect larvae. Insect cell culture provides an environment in which the process parameters can be carefully controlled and monitored (4,7). It also allows for the utilization of a single cell line with more uniform posttranslational processing than can be achieved in larvae. Purification of proteins from cell culture medium, especially serum-free medium, is usually easier than from larvae of infected insects. Larvae will include significant levels of other insect proteins, and the presence of larval proteases can lead to a degraded final product. In contrast, the use of larvae may provide substantial cost advantages

Table 1. Comparison of Glycosylation Patterns of IgG Produced in Insect Cells (*Trichoplusia ni*) and Mammalian Cells (NS/0)

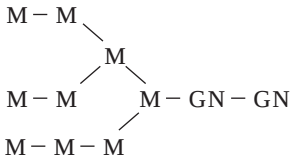
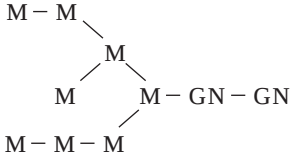
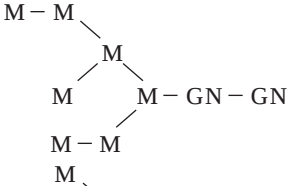
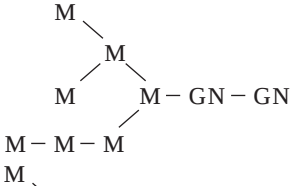
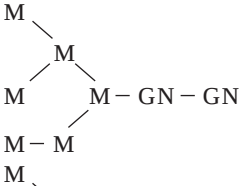
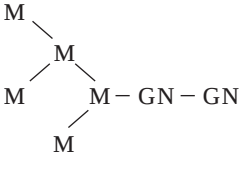
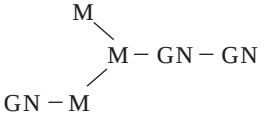
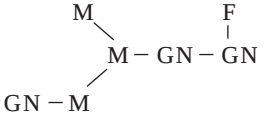
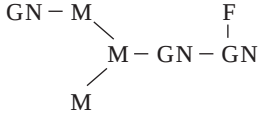
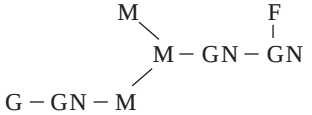
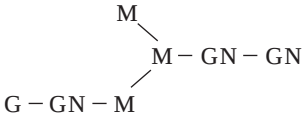
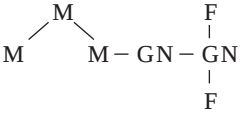
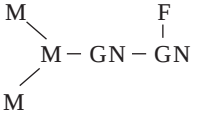
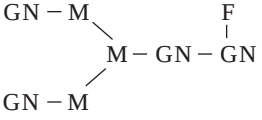
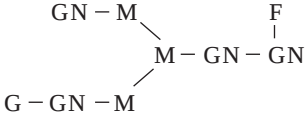
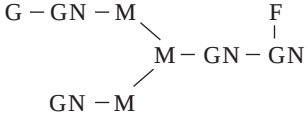
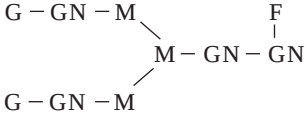
	Insect IgG lysate (mol %)	Insect IgG supernatant (mol %)	Humanized murine IgG NS/0 (mol %)
High-mannose structures			
	3.0	–	–
	12.6	–	1
	9.9	–	2
	19.2	–	2
	2.6	–	7
	2.0	–	23
Hybrid structures			
	1.2	1.6	2
	14.2		7
	0.3	2.7	–
	–	–	4

Table 1. Comparison of Glycosylation Patterns of IgG Produced in Insect Cells (*Trichoplusia ni*) and Mammalian Cells (NS/0) (*continued*)

	Insect IgG lysate (mol %)	Insect IgG supernatant (mol %)	Humanized murine IgG NS/0 (mol %)
	–	–	2
Low-mannose structures			
	11.8	18.0	–
	9.5	16.7	–
Complex structures			
	5.6	16.0	22
	8.1	14.2	9
	–	5.3	10
	–	–	5

Source: Data for *T. ni* from Ref. 26; data for mammalian cells from Ref. 27.

for the production of recombinant proteins. The cost of commercial media used in cell culture can become prohibitive for large volumes, and a significant capital investment is required for purchasing large bioreactors. Some posttranslational processing events such as glycosylation, secretion, and carboxy-terminal amidation also may be more efficient in larvae (37). These considerations may warrant the consideration of insect larvae as an alternative method for the production of recombinant proteins.

Cell Culture Processes

Increasing the volume and mass of production from a small 100-mL spinner flask to a large 100-L bioreactor is not a simple linear process (13). Many of the process parameters, especially the delivery of O_2 and mixing, change when production is moved to a large bioreactor. Even so, large-scale production of proteins in BES has been achieved in 100-L bioreactors while maintaining cell growth kinetics and productivity. Issues that are relevant to cell culture scale-up are addressed next.

Bioreactors: Stirred-Tank versus Airlift

Both airlift and agitated–sparged (stirred-tank) bioreactors have been successfully used for the production of recombinant proteins and virus in larger scales. Stirred-tank bioreactors typically consist of glass, carbonate, or stainless steel vessels containing a marine or paddle impeller driven by an overhead motor (Fig. 7). Temperature, oxygen, and pH can be controlled. Temperature control can be maintained by immersion in a water jacket, using circulating cooling and heating coils, or by enclosure in a controlled incubator. Stirred-tank bioreactors typically require the injection of air or oxygen into the bioreactor at a position below the impeller. Air can also be supplied to the head space in order to limit the amount of oxygen sparged directly into the bioreactor. An alternative to the stirred-

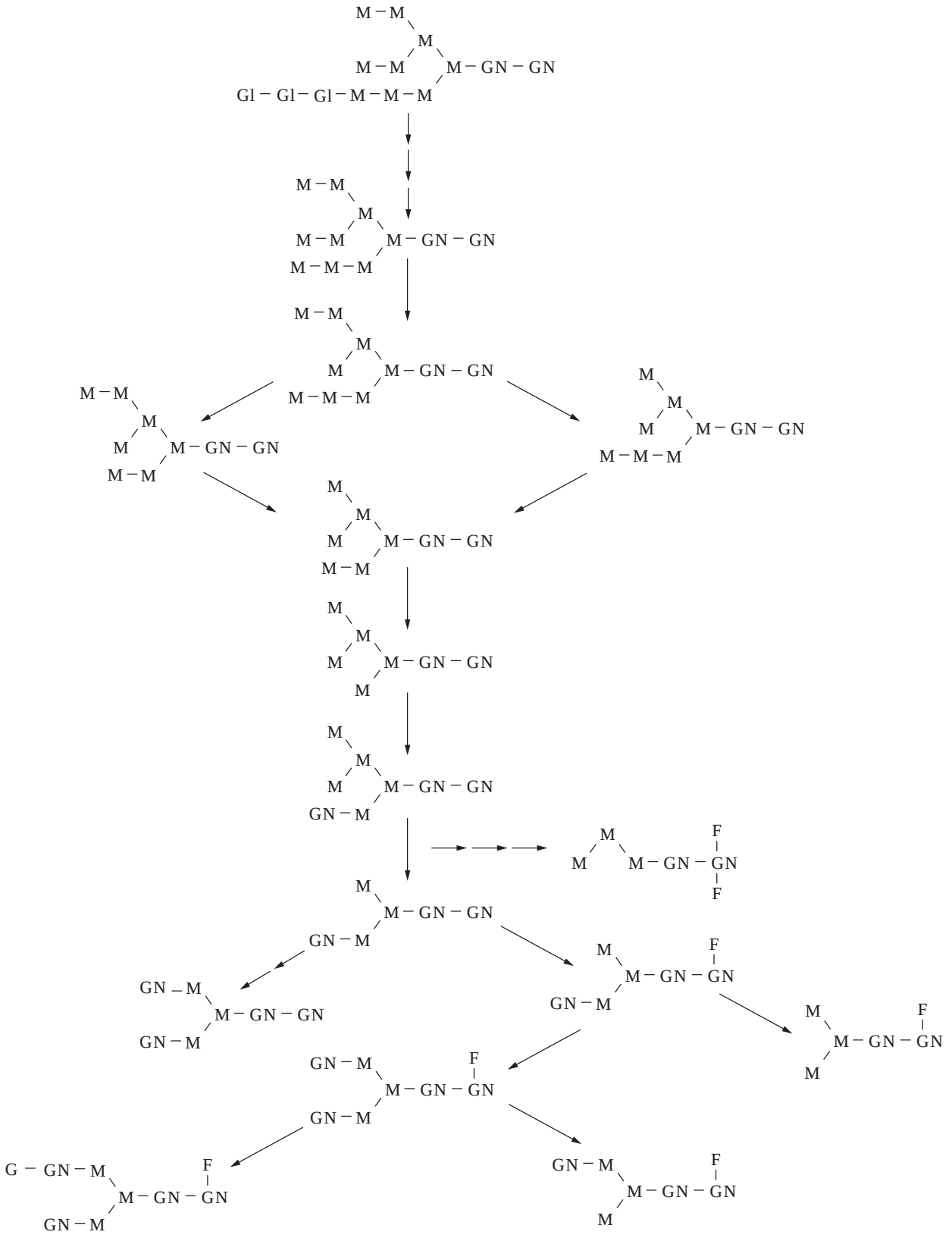


Figure 6. *N*-glycan processing pathway in *T. ni* insect cells. The oligosaccharide chain consists of the following saccharide units: F, fucose; Gl, glucose; GN, *N*-acetylglucosylamine; M, mannose. (Figure shown as published in *Differential N-Glycan Patterns of Secreted and Intracellular IgG Produced in Trichopulsia ni Cells*, T.A. Hsu et al., *The Journal of Biological Chemistry*, 1997. Reprinted by permission of The American Society for Biochemistry & Molecular Biology.

Table 2. Genes Utilized to Increase the Insecticidal Effects of Recombinant Baculoviruses

Name of gene	Effect(s) of expressed gene on host insect	% increase in insect killing
Diurectic hormone	Reduction of hemolymph volume	20
Juvenile hormone esterase	Feeding cessation and blackening	—
Bacillus thuringiensis toxin	Feeding cessation	$(2 \times LD_{50})^a$
Scorpion toxin (AaIT)	Paralysis, body tremors, and dorsal arching	30–40
Mite toxin (Txp-1)	Paralysis	30–40
Wasp toxin (antigen 5)	Premature melanization and low weight gain	—
AcNPV deletion mutant lacking ecdysteroid UDP-glucosyltransferase gene (egt)	Reduces larval feeding times	20% reduction in $ST_{50}{}^b$

[a] LD_{50}, the dose necessary to kill 50% of the larval population.
[b] ST_{50}, the time necessary to kill 50% of the larval population.

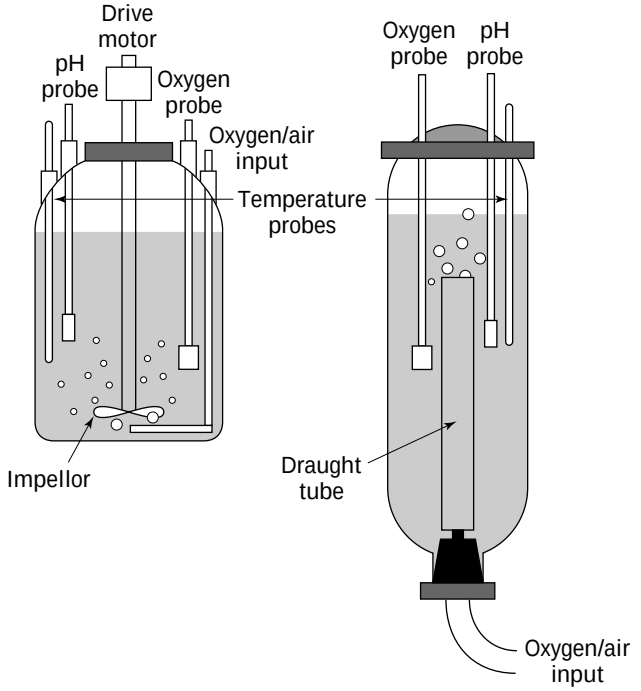


Figure 7. Configuration of stirred-tank and airlift reactors.

tank bioreactor is the airlift bioreactor in which liquid circulation through the bioreactor is driven by the lifting force created by injection of air or oxygen through the bottom of the bioreactor (Figure 7). This bioreactor includes an inner draft tube that carries the air/oxygen and liquid medium up through the bioreactor to the surface of the medium in order to set up circular flow of medium within the bioreactor.

A side-by-side comparison of a 6-L airlift vessel and an 8-L stirred-tank vessel (5.5-L working volume) containing Sf-9 cells in serum-free medium showed similar growth rates, cell viabilities, and cell densities throughout the cell growth period (38). Doubling times were about 22 h in both cases, and the growth curves were similar to those achieved in 100-mL spinner flasks, demonstrating the scaleability of each approach. Most importantly, production of a recombinant protein (kinase C-η) was virtually identical for both vessels in terms of protein yield and activity. Consequently, the principal consideration for choosing a production vessel is likely to depend on availability and cost. Stirred-tank vessels of many different sizes are readily available, whereas airlift scale-up requires that a constant height-to-circumference ratio be maintained in order to allow airlift function. This can lead to very tall vessels for large bioreactor volumes. The stirred-tank fermentation system can be implemented without a commercial vendor, if necessary, by purchasing a large motor and vessel, whereas an airlift system is usually purchased as a single unit from a commercial source.

Though stirred-tank and airlift are the principal bioreactor configurations used, some variations in these two types are available. Clumping of attachment-dependent cell lines in suspension, such as with High Five, is common, so bioreactor configurations that allow attachment-dependent cell growth and infection have been developed (3,39). These attached cell lines are typically grown in roller bottles or on supporting particles such as glass beads or microcarriers. The supporting particles are then incorporated into immobilized perfusion or airlift bioreactor configurations. In addition, bioreactors in which oxygen is supplied through membrane tubing have been utilized, though not extensively for large-scale cell culture (40). As the size of the bioreactor increases, the amount of tubing required for efficient mass transport increases significantly.

Process Monitoring Techniques

Cell growth and infection in insect cell bioreactors are usually monitored by following total and viable cell numbers by means of trypan blue staining and counting with a hemacytometer. Cell viability typically drops after several days after baculovirus infection of an insect cell culture. Given the labor-intensive and invasive nature of this assay, other techniques have been considered for monitoring cell growth and infection, both off-line and on-line (40). Coulter™ particle counters and optical density (OD) measurements taken using a fiber optic probe can speed the cell counting process considerably during the cell growth period. Alternatively, growth can be monitored on-line through indirect measurements of cellular respiration such as oxygen uptake rate (OUR) or CO_2 production rates (CPR). Monitoring the infection process is much more challenging, though several techniques are under development to supplement trypan blue staining. OUR can be used to follow the infection process because of the transient in-

crease in oxygen consumption that occurs during the initial stages of the baculovirus infection. Other tools that have been used to monitor the infection process include epifluorescence microscopy of live and dead cells using fluorescence stains, as well as flow cytometric monitoring of DNA to track virus production (41). Fluorometric assays can be used to examine protein production directly in stained cells and will supplement conventional bioassays and SDS-PAGE analysis.

Oxygen Supply

One of the critical nutrients that must be supplied to an insect cell culture during the growth and infection stages is oxygen. Oxygen consumption rates per cell range from 10^{-17} to 10^{-18} mol/cell s and can increase by 30% or more during the initial stages of infection. Absence of sufficient aeration and oxygen supply during the infection stage can inhibit the viral infection and lower recombinant protein yields several fold (42). Consequently, it is important to monitor the dissolved oxygen level in large-scale insect cell bioreactors using commercially available probes. Some researchers have suggested that dissolved oxygen tension (DOT) be maintained at about 50% saturation to reduce the required oxygen sparging that is potentially lethal to insect cells (43). However, it is more important to ensure that the DOT always remain above 20% air saturation to prevent any limitations during the infection period.

Cell Damage During Sparging

Oxygen supply to the bioreactors is typically provided through sparging of oxygen or air into the bottom of a stirred-tank or airlift bioreactor. The oxygen is transferred from the gas to the liquid as the bubbles rise up the bioreactor. In order to provide sufficient mass transfer area, the air is bubbled through small openings. Impeller rotation can enhance bubble breakup further to increase the surface area available for mass transfer. The resulting bubble size is typically 1–5 mm in diameter, and the Reynolds number of a rising bubble can vary from 100 to 1,000 in conventional bioreactors (44). It is now known that the sparging of oxygen or air can cause significant damage to insect cells as well as mammalian cells in suspension (45–47). This cell damage is principally due to the adhesion of cells to bursting gas bubbles at the gas–liquid interface. In contrast, agitation is likely to be less of a factor in causing cell damage in many sparged bioreactor environments. Suspended cells can often tolerate high agitation rates as long as the gas–liquid interface is removed (48,49). Cell damage from agitation may occur in some cases due to the formation of Kolmogorov eddies on a length scale comparable to the size of a cell (48,49).

High-speed video microscopy studies have been used to observe directly cell–bubble interactions during the bubble rise through an insect cell culture medium (46,47). Cells attach independently or in clumps to a rising bubble and are carried upward through the fluid (Fig. 8). These cells then accumulate in high concentrations in the foam layer at the surface of the medium. The rupture of bubbles at the gas–liquid interface leads to the formation of upward and downward liquid jets that generate large shear

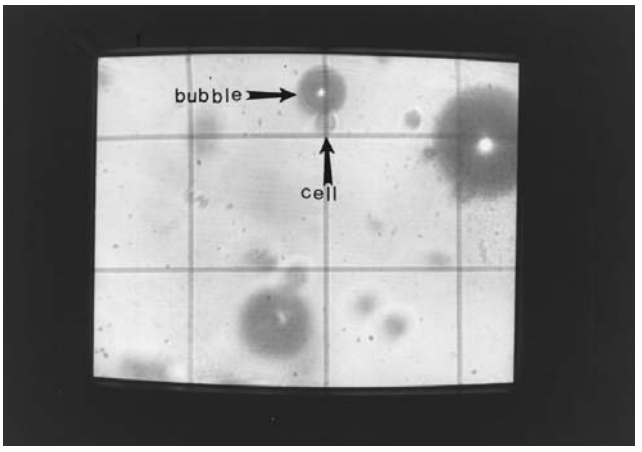


Figure 8. Insect cell attachment to bubbles rising through culture medium. Grid spacing is 100 μm. *Source*: Photograph courtesy of J. Chalmers.

stresses sufficient to damage or kill insect cells attached to the bubbles. The cell concentration in the jets is higher than in the medium, and more than 75% of these cells are killed by bubble rupture and jet formation. Approximately 1,200 Sf-9 cells can be killed per bubble rupture (44).

Much of the cell damage due to bubble rupture can be prevented by the addition of surfactant polymers such as Pluronic F-68 at concentrations in the range of 0.1 to 0.2% (46,50–52). The addition of Pluronic F-68 prevents the adhesion of cells to bubbles and eliminates the damage to cells caused by bubble breakup at the liquid surface. The polymeric surfactants may prevent cell adherence to bubbles by lowering the interfacial tension at the gas–liquid interface. Most commercial serum-free cell culture media for insect and mammalian cells include Pluronic F-68, although other surfactants such as methylcellulose, serum, carboxymethylcellulose, bovine serum albumin, and dextrans have been considered as well. Unfortunately, only a limited selection of surfactants can be used because the interaction of the surfactants with the cells may affect membrane permeability and cause cell lysis (44).

Media Formulation

Media formulations are derived from elements originally identified in the insect hemolymph and include salts, vitamins, amino acids, sugars, organic acids, and trace elements (53,54). A number of proprietary commercial serum-free media are available for insect cell growth, including Express 5 and Sf900II from Gibco-BRL, and Excell 401 and 405 from JRH Biosciences. These media are derivations of noncommercial media such as Grace's, TNM-FH, TC-100, and IPL-41 that were supplemented with 5–20% fetal calf serum and used for many years in the cultivation of insect cells. The serum-free media are supplemented with lipids (cholesterol, α-tocopherol acetate, cod liver oil fatty acid methyl esters, β-sitosterol, and stigmasterol), yeast extract (yeastolate), and Pluronic F-68 in addition to other proprietary alterations (55,56). Antibiotics such as penicillin G and streptomycin are often added to the medium during cell culture.

Media sugars, including glucose, maltose, and fructose, can be partially or completely consumed during insect cell growth and recombinant protein production. Also, a significant fraction of the amino acids, including glutamate, glutamine, serine, asparagine, threonine, aspartate, and arginine, can be consumed during insect cell growth. However, much of this consumption may be devoted to energy use by the cells rather than the production of biomass (54,55). In contrast, a significant increase in the level of alanine is realized during Sf-9 cell growth experiments in IPL-41 medium; the amino acid may represent a sink for ammonia released during amino acid catabolism. Other amino acids remain relatively balanced throughout the insect cell growth experiment. Because the level of amino acids present in IPL-41 medium is far in excess of that required for biomass production, media alterations that reduce these amino acid levels may lead to more efficient nutrient utilization (56). Also, a more detailed characterization of the yeast extract components critical to cell growth may eventually result in the development of a truly defined medium for insect cell culture (53,56). The consumption rates of these nutrient components may change slightly but not significantly following baculovirus infection (43).

Cell Culture By-products

A significant amount of carbon dioxide is exhausted due to nutrient oxidation. Metabolites including lactate and other organic acids (pyruvate, α-ketoglutarate, fumarate) can also accumulate during growth and infection. The production of lactate from Sf-9 cells is generally lower than that observed with mammalian cells and poses no inhibitory effect on growth, although, lactate levels can increase significantly following DO limitations (57). The metabolites secreted in cell culture may change significantly depending on the cell line. Ammonia may be produced or consumed, depending on the particular cell type and cell culture conditions. In fact, much higher levels of lactate and/or ammonia are generated by *T. ni* and *B. mori* cells when compared to Sf-9 cells (43,58). In addition to toxic metabolites, the accumulation of proteases has been noted during the late viral infection stage and may be correlated with recombinant protein degradation (43). The most common approach to prevent protein degradation is to include a cocktail of protease inhibitors with the viral infection medium and during the protein purification process.

Most insect culture media are not buffered with carbon dioxide, so there is no need for a CO_2 environment for growth. The optimal pH for growth of most insect cells is between 6.2 and 6.3, though the pH of Sf-9 cultures in an uncontrolled medium can reach a minimum value between 5.9 and 6.1 (40). Because pH variations in the medium will affect both the growth and infection stages (59), pH control is suggested for optimal productivity in large-scale cell cultures. Osmolarity is another variable that should be considered, although, varying the osmolarity from 350 to 385 mosm/kg did not alter the insect cell growth rate more than 10% (60). Most media osmolarities range from 320 to 375 mosm/kg.

Batch Production Techniques

Recombinant baculovirus and heterologous proteins have been produced using batch, fed-batch, and continuous production modes. However, batch production of recombinant proteins remains the method of choice for most users. Typical processing of insect cells involves the batch cultivation of cells in a small suspension vessel, followed by a gradual scale-up of the cultures into progressively larger bioreactors. The cells are grown to a maximum cell density in one vessel and then pumped through steam-sterilized lines to increasingly larger vessels containing fresh medium. The optimal growth temperature is usually between 27 and 29 °C, with cell doubling times between 17 and 24 h, depending on the cell line and medium utilized. Some reports have recorded that High Five cells have a shorter doubling time (13–17 h) than Sf-9 cells (20–24 h) (58). In the final production-scale vessel, the cells are infected with virus during the exponential growth phase, and the virus or recombinant protein is collected at the conclusion of the infection stage (typically between 48 and 100 h postinfection).

Virus Stocks, Multiplicity of Infection, and Time of Infection

To initiate the infection, an aliquot of virus is added to the cell culture broth. Virus stocks are made from a single viral isolate obtained by plaque assay or end-point dilution. The plaque assay involves the infection of a monolayer of cells with a low concentration of virus, followed by an overlay of agarose-containing medium. Eventually, an isolated region of infected cells, called a plaque, is identified, and an agarose plug is collected at this plaque location for generating viral stocks. With the end-point dilution method, the tissue culture medium is diluted to the point at which any infection must result from a single infectious virus. Large stocks of the virus are obtained by infecting a monolayer or suspension cell culture with the agarose plug (from a plaque) or tissue culture medium (from the end-point dilution). The cell culture medium is then collected after 4–5 days, however, it may be necessary to repeat the infection several times in order to generate a viral stock of sufficient titer. The viral titer is obtained by determining the number of plaque-forming units (PFU) per milliliter, using the plaque assay, or by finding the 50% tissue culture–infectious dose ($TCID_{50}$) per milliliter using end-point dilution. The $TCID_{50}$ dose per milliliter is then multiplied by 0.69 to determine the number of plaque-forming units per milliliter. Several master viral stocks are frozen at −70 °C for long-term storage, with one master stock kept at 4 °C in order to generate working viral stocks.

Cells are then infected at a desired multiplicity of infection (MOI), which represents the number of plaque-forming units per viable cell, from the working viral stock. In small shaker-flask cultures, cells are typically infected at an MOI greater than 1 during the mid to late exponential phase of growth in order to generate a synchronous infection. Infecting the cells at the highest possible cell density and cell viability while still in the exponential growth phase provides high volumetric productivities. However, an efficient infection is also dependent on the presence of adequate nutrient levels in the medium. If the

cells are infected at a point near the end of the exponential phase or during the stationary phase, insufficient substrates will be available for the viral infection, and low recombinant product yields will result (42).

In larger bioreactors, infection at a high MOI (>1) may not be feasible or desirable due to the requirements for large amounts of virus stock. If cells are infected at a lower MOI, then the infection of the cells will not be synchronous, and some insect cell growth will occur prior to secondary infection by BV generated during the initial infection (60). If the MOI is too low or cell density too high, the cells will grow to the late exponential or stationary phase prior to complete infection, and product yields will be lower. Consequently, it is important to optimize both the time of infection (TOI) and MOI to ensure that all cells are infected by virus and that sufficient substrates remain for virus and/or protein production (61). This optimization is facilitated by the use of mathematical models that consider substrate consumption during both cell growth and viral infection (62,63). With these models, protein yields can be achieved using MOIs as low as 0.0001 that are equivalent to those obtained with infections at a MOI of 5 (61). Infections at low MOI and low cell density may be very effective for cell lines such as Sf-9, but the approach may not be as successful for cell lines such as TN-5-B1-4, which produce much lower levels of extracellular virus.

Fed-Batch, Semibatch, and Sequential-Batch Production

It is desirable to grow the cells to a high cell density to maximize the productive capacity of the bioreactor. Studies have shown that Sf-9 cell concentrations can reach 3–5 $\times$ 10^6 cells/mL in batch culture for a typical medium. Supplementation of glucose, glutamine, and/or yeastolate can be used to increase the maximum cell density by at least 50% and enhance recombinant productivity. These nutrients can be added continuously in a fed-batch mode or in a single dose near the end of the initial growth period (Fig. 9). As an alternative, fresh medium can be fed continuously in a perfusion mode to replace the spent medium exiting through a hollow fiber device (0.4–0.6 vol/day) to provide cell densities up to 10 $\times$ 10^6 cells/mL of bioreactor volume (64). However, the perfusion of fresh medium results in a slightly lower cell mass yield on a per milliliter of total medium basis (4.1 $\times$ 10^6 cells/mL total medium) as compared with nutrient supplementation (6.7 $\times$ 10^6 cells/mL total medium). Another approach that has been used to increase cell density is the replacement of 80% of the medium for several consecutive days near the end of the growth stage. Using the same hollow fiber technology (0.2-μm pore size), most of the spent media is removed, and then fresh medium is added to the bioreactor. This approach can be used to double cell densities up to 11.6 $\times$ 10^6 cells/mL reactor volume (64). Even greater enhancements in cell density can be obtained when nutrient supplementation of glucose, glutamine, and yeastolate is combined with medium replacement (15.8 $\times$ 10^6 cells/mL reactor volume) (64). A similar nutrient supplementation strategy is used to enhance the yields of recombinant proteins during the infection stage because the recombinant product yield is highly dependent on the nutrient levels. One approach that is used to ensure high product yields is to resuspend the cells in fresh medium prior to the infection, or else to add fresh medium to the bioreactor at the beginning and/or during the infection stage (42). The supplementation of limiting nutrients such as glucose, glutamine, and yeastolate can also be used to enhance the infection process (4,65–67), eliminating the costly replacement of the entire medium.

Another semibatch operating strategy used for production of recombinant baculoviruses and heterologous proteins is the utilization of separate growth and infection reactors (Fig. 10) (68). An initial cell growth reactor is followed by single or multiple sequential infection vessels in series. After each cycle, the contents from one infection vessel is pumped to a following infection vessel while the contents of the last infection vessel are harvested for product. The first infection vessel is then filled with medium and cells from the growth reactor and inoculated with virus. The cell growth reactor is provided with fresh medium, and the growth cycle repeated. Some of the cells from the growth stage may remain in the reactor to serve as an inoculum for the next growth cycle (69,70), and a portion of the contents of the first infection reactor can be left to infect a new batch of cells. However, modeling studies have indicated that the productivity of this sequential batch system is expected to decline after a number of cycles due to the buildup of mutant viruses known as defective interfering particles (DIPs) (see later section) (71). The formation of these mutants is due to the "passage" effect (see later section). The recombinant protein productivity of the second and subsequent infection stages must be monitored closely for the build-up of these particles, and the second infection stage must be completely emptied and inoculated with a fresh sample of virus in order to eliminate the DIPs. Semicontinuous operation has also been achieved by maintaining a first-growth reactor in a continuous operating mode and then gradually filling a second reactor (72). Following the filling of the second reactor, virus is added to initiate a batch infection in the second stage. At the end of the infection process, the second stage is emptied in part or completely, and the filling and infection cycle repeated. The medium in the infection reactors may also be supplemented or replaced by separating the cells from spent medium by settling, membrane filtration, or ultracentrifugation. Since the nutrients are often not consumed completely in the infection stage(s), it may not be necessary to replace all the spent media. These sequential-batch processes can be achieved with processing times dictated by the exhaustion of nutrients, decline of cell viability, and accumulation of recombinant proteins.

Continuous Production Systems

The release of BV from infected insect cells into the cell culture medium has led to the design and evaluation of continuous-flow stirred-tank reactors for the production of polyhedra and recombinant proteins. The utilization of a continuous system could reduce production costs by 50% when compared with a batch system (73). Batch systems have the disadvantage of long down times as a result of cleaning and sterilization after each run. Continuous sys-

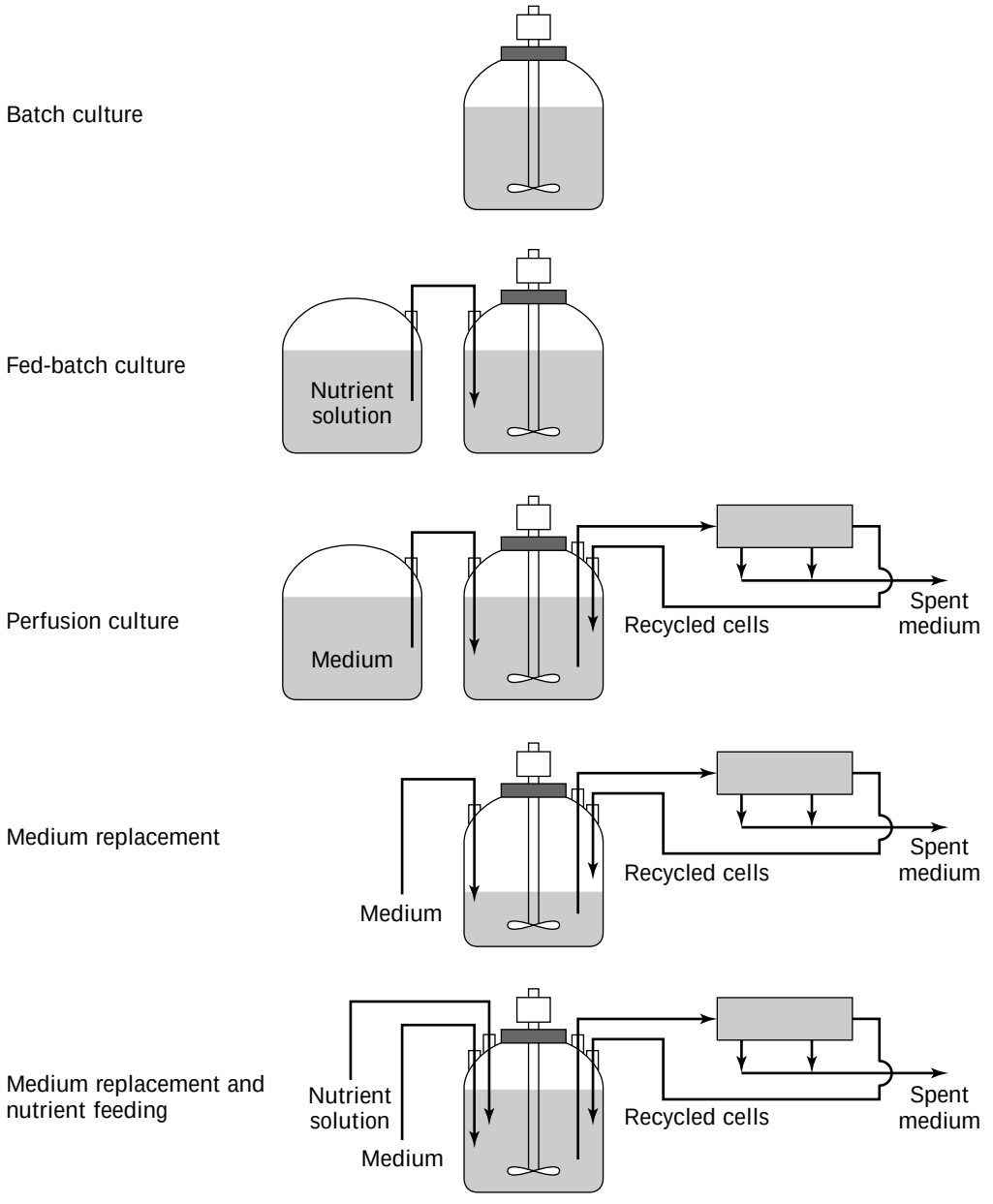


Figure 9. Various reactor configurations for batch and fed-batch production strategies.

tems typically consist of sequential two-stage continuous-flow bioreactors in which the first stage is used for cell growth and the second stage is used for virus production and heterologous gene expression (Fig. 10). The effluent of cells from the first stage feeds the second infection stage, and the effluent of the second stage includes the infected cells and associated baculoviruses and/or recombinant product. The BV generated in the second stage is used to infect the feed from the first reactor. In one study (74), a 60-h mean residence time was maintained in each stage, and production of polyhedra and recombinant protein was sustained for approximately 4 weeks. Because about 200 BVs are produced per cell, there is a gradual accumulation of virus in the second reactor, followed by a severe decline in productivity after 4 weeks due to the formation of DIPs. A model that considers the presence of infectious virus (I-BV), defective virus (DIP), and abortive virus (A-BV) effectively predicted the accumulation of DIPs and the decline of recombinant protein production for a two-stage continuous-infection process (71).

Thus, a continuous reactor system is not likely to be feasible for extended production of baculovirus and recombinant protein. However, such a set-up may be used for production over shorter periods such as a 30-day cycle prior to the accumulation of DIPs (71). This reactor set-up would be analogous to a continuous catalytic reactor that experiences gradual catalyst deactivation during its operation and has to be periodically recharged with new catalyst. This could be accomplished by a complete shut down and restart with fresh virus or else by providing a secondary

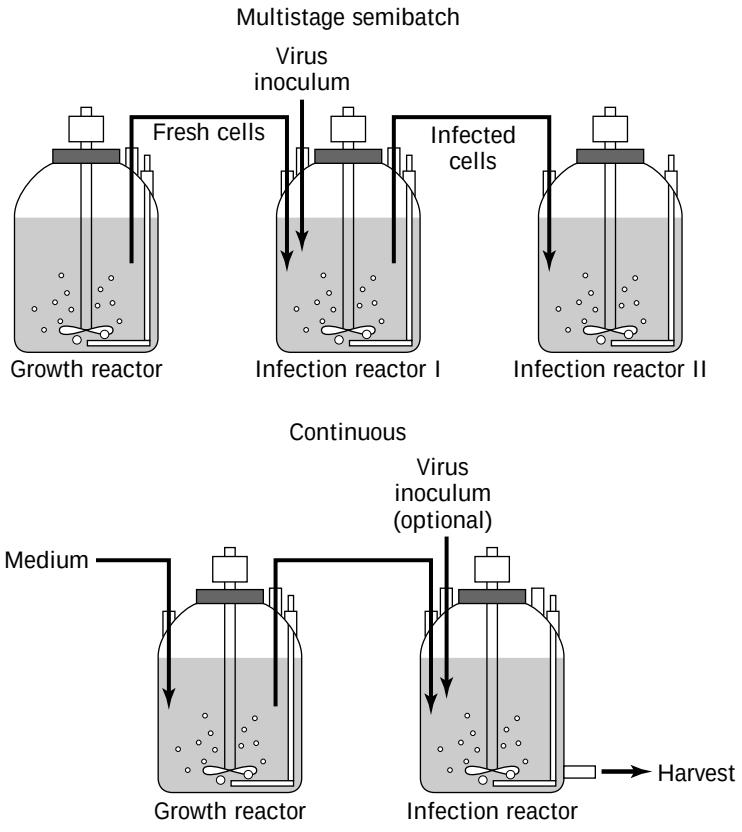


Figure 10. Multistage reactor systems. Multistage reactor systems used for the semibatch and continuous production of recombinant proteins in insect cell culture using Ac-MNPV. The initial reactor is primarily used for cell growth, whereas subsequent reactors contain the insect cells infected with the virus.

feed of fresh virus to the second infection reactor. A continuous set-up provides for larger processing volumes of insect cell culture and recombinant product than does batch operation, although the yield of recombinant protein in the effluent may be low compared with batch or fed-batch operation. A recent innovation in the continuous culture process has been the sequential combination of a continuous stirred-tank reactor (CSTR) for growth, followed by a tubular reactor for infection. The tubular reactor provides for the continuous introduction of baculovirus at higher MOIs without the negative passage effects present in a two-stage CSTR system (75).

Passage Effects and Mutant Virus Formation

The continuous passage of viruses in cell culture often leads to the formation of mutant viruses such as DIPs (76,77). The DIPs will replicate more rapidly than the normal-size viruses, and the concentration of DIPs can rise to as high as 60% of the total virus stock (76,77). These DIPs average about 200–230 nm in size, as compared with the standard 330-nm-sized virions (77), and can lack up to 40% of the baculovirus genome due to spontaneous deletions during passaging (76). The deleted segment often includes the polyhedrin gene and at least one other gene necessary for virus replication. Consequently, DIP mutants cannot be plaque purified because they replicate only in cells that are also infected by a normal virus. To avoid the passage effect, viral passage number should be monitored, and a low passage master viral titer stock used to generate working viral titer stocks. If contamination with DIPs is suspected, then a plaque assay can be performed to reisolate the infectious virus. Also, low MOI infections can be used to minimize DIP formation, but this approach results in asynchronous infections.

Other mutant viruses, such as ones that result in the generation of a few polyhedra (FP), have also been observed. The FP mutants substantially lower the number of polyhedra produced per infected cell and reduce the number of occluded viruses generated (78). These FP mutants may have a selective advantage in cell culture by producing more budded virus at the expense of polyhedra. Unlike DIPs, FP mutants can be plaque purified. Once again, formation of FP mutants can be minimized by working with low-passage viral titers and utilizing a master viral titer stock. Because FP mutants do not include significant virus levels in the polyhedra, insect-to-insect transmission will not include FP mutants, which are not capable of spreading through polyhedra infections.

Larval Production Systems

Based on productivity comparisons with polyhedra, approximately 100–200 larvae will provide a final protein yield approximately equal to 1 L of cells for an intracellular recombinant protein (7). The use of larvae eliminates the need for high capital investments in large-scale bioreactors but requires the development of systems for insect rearing and infection. Furthermore, purification strategies must be implemented to isolate the proteins from insect parts or

hemolymph. Typical species considered include *Spodoptera frugiperda, Trichoplusia ni, Heliothis virescens, Manduca sexta,* and *Hyalaphora cecropia* for AcMNPV infections, and *Bombyx mori* for BmNPV infections (7,37). Lepidopteran eggs are available from commercial suppliers and hatch 5 days to 1 week after incubation at 25–27 °C in a humidified environment under a controlled light cycle. Artificial diets are also available from commercial sources or may be prepared from components in the laboratory. Constituents typically include pinto beans, Brewer's yeast, ascorbic acid, wheat germ, sorbic acid, 14% auromycin, and a vitamin mix. The ingredients are combined with agar and then provided to larvae with periodic replacement to ensure a moist supply of nutrients. It is important to infect the insects at the appropriate developmental stage to provide the maximum yield of baculovirus or heterologous protein. The insects typically molt five or six times as larvae with the last stage being the larval–pupal molt. The intermolt period is the stadium or instar. Optimal infections are typically achieved one or two days after molting into the last larval instar. Infections are accomplished by feeding the insects with polyhedra inclusion bodies (PIBs) or by injecting budded virus (BV) directly into the hemolymph. For recombinant baculoviruses that are occlusion negative, injection of the budded virus is the required method of infection. A small number of mock-infected insects should also be maintained to check for possible contamination problems and to provide a comparison of infected and uninfected larvae. Shown in Figure 11a is a photograph of an infected *T. ni* larva expressing green fluorescent protein as compared with an uninfected larva (79). Figure 11b shows massive expression of *β*-galactosidase in a *T. ni* larva when infected with AcMNPV containing the heterologous *lacZ* gene under the control of the p10 promoter. At 60 h postinfection the infection has cleared the midgut and entered the fat body. The best time to collect product from infected insects is generally prior to death. For secreted proteins, the product is collected by bleeding the insects to collect hemolymph, whereas nonsecreted proteins are isolated by homogenization of whole insects. Unfortunately, homogenization can result in the release of proteases; these can be minimized by removal of the insect gut before processing (7,37). Alternatively, a particular metabolically active tissue such as the fat body can be removed from the insect and used as a source of the desired protein (7,37).

CONCLUSION

The use of the baculovirus–insect cell system for heterologous gene expression and recombinant biopesticides has grown dramatically since its first application in 1983. Commercially available vectors that provide for the rapid selection and screening of recombinant viruses has made this technology accessible to most biotechnology laboratories. The development of serum-free media, the isolation of alternative insect cell lines, and the implementation of efficient bioreactor operating strategies have increased productivities and lowered costs. As a result, insect cells are often the method of choice for rapid and high-level production of functional eukaryotic proteins. The use of BES is likely to increase even more with the growth of genomics and its associated need for functional characterization of unknown proteins. However, the application of insect cell-derived proteins as pharmaceuticals will continue to be slow due to the different posttranslational modifications in insect cells as compared with mammalian cells. Future efforts will be directed at enhancing posttranslational processing functions and improving protein secretion from

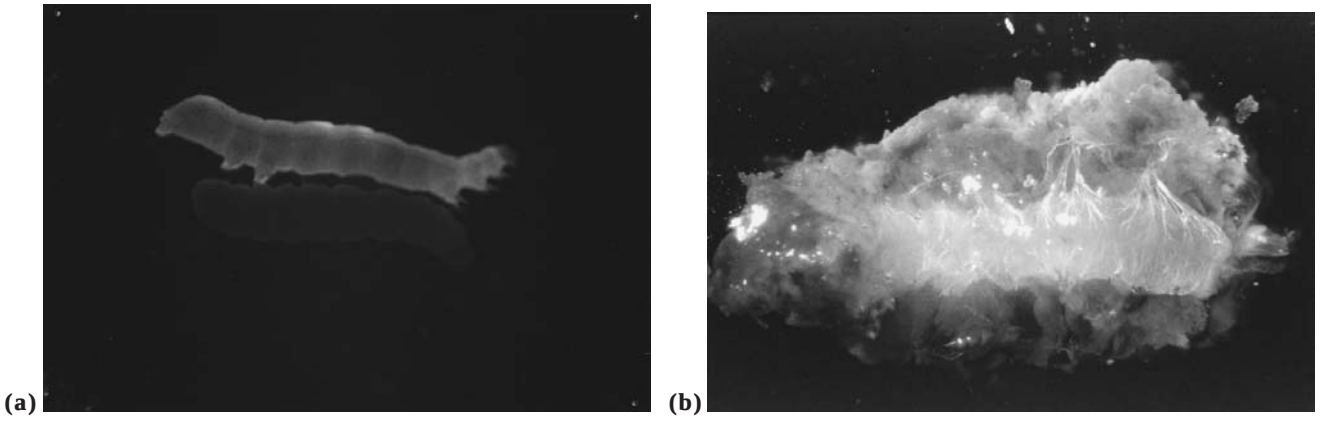


Figure 11. (**a**) Photo of *T. ni* larvae sitting atop a laboratory UV transilluminator table. The upper larva has been infected (72 h) with a recombinant AcMNPV baculovirus that contains the gene for green fluorescent protein (GFPuv) placed under control of the polyhedrin promoter. The brightness is due to the massive overexpression of GFPuv, which has been shown to reach 26% of the total larvae protein. The lower larva (nonglowing) is the uninfected control. *Source*: Photograph courtesy of W. Bentley. (**b**) Micrograph showing AcMNPV-*β*-galactosidase infection of *T. ni* insect larva at 60 h postinfection. Staining represents *β*-galactosidase expressed in insect cell tissues. Infection has cleared the midgut and spread to the fat body. *Source*: Micrograph courtesy of G. Williams.

current and novel insect cell hosts. Other promising areas include the use of insect larvae and the development of stable expression systems such as *Drosophila* S2 cells (80). With these advances, insects and insect cells are likely to remain a popular choice for the manufacture of recombinant DNA products.

ACKNOWLEDGMENTS

The authors are grateful to Greg Williams, Jeffrey Chalmers, and William Bentley for graciously providing unpublished photographs for this review. Special thanks to Mike Owens for technical input and assistance.

BIBLIOGRAPHY

1. D.E. Lynn, *Tissue Culture Methods* **12**, 23–29 (1989).
2. T.R. Davis, T.J. Wickham, K.A. McKenna, R.R. Granados, M.L. Shuler, and H.A. Wood, *In Vitro Cell. Dev. Biol.* **29A**, 388–390 (1993).
3. R.A. Taticek, D.A. Hammer, and M.L. Shuler, in M.L. Shuler, H.A. Wood, R.R. Granados, and D.A. Hammer eds., *Baculovirus Expression Systems and Biopesticides*, Wiley-Liss, New York, 1995, pp. 131–174.
4. D.W. Murhammer, *Appl. Biochem. Biotechnol.* **59**, 199–220 (1996).
5. M.J. Betenbaugh, L. Balog, and P-S. Lee, *Biotechnol. Prog.* **7**, 462–467 (1991).
6. W.F. Hink, D.R. Thomsen, D.J. Davidson, A.L. Meyer, and F.J. Castellino, *Biotechnol. Prog.* **7**, 9–14 (1991).
7. D.R. O'Reilly, L.K. Miller, and V.A. Luckow, *Baculovirus Expression Vectors: A Laboratory Manual*, W.H. Freeman and Company, New York, 1992.
8. M.L. Shuler, D.A. Hammer, R.R. Granados, and H.A. Wood, in M.L. Shuler, H.A. Wood, R.R. Granados, and D.A. Hammer eds., *Baculovirus Expression Systems and Biopesticides*, Wiley-Liss, New York, 1995, pp. 1–11.
9. G.E. Smith, M.J. Fraser, and M.D. Summers, *J. Virol.* **46**, 584–593 (1983).
10. G.E. Smith, M.D. Summers, and M.J. Fraser, *Mol. Cell. Biol.* **3**, 2156–2165 (1983).
11. J.M. Vlak, A. Schouten, M. Usmany, G.J. Belsham, E.C. Klinge-Roode, A.J. Maule, J.W.M. Van Lent, and D. Zuidema, *Virology* **179**, 312–320 (1990).
12. V.A. Luckow and M.D. Summers, *Biotechnology* **6**, 47–55 (1988).
13. R.M. Patterson, J.K. Selkirk, A.B. Merreck, *Environ. Health Perspect.* **103**, 756–759 (1995).
14. M.D. Summers and G.E. Smith, *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*, Texas Agricultural Experiment Station Bulletin no. 1555, 1987.
15. S. Maeda, T. Kawai, M. Obinata, H. Fujiwara, T. Horiuchi, Y. Saeki, Y. Sato, and M. Furusawa, *Nature* **315**, 592–594 (1985).
16. M.D. Ayres, S.C. Howard, J. Kuzio, M. Lopez-Feriber, and R.D. Possee, *Virology* **202**, 586–605 (1994).
17. G.D. Pennock, C. Shoemaker, and L.K. Miller, *Mol. Cell. Biol.* **4**, 399–406 (1984).
18. J. Vialard, M. Lalumiere, T. Vernet, D. Briedis, G. Alkhatib, D. Henning, D. Levin, and C. Richardson, *J. Virol.* **64**, 37–50 (1990).
19. P.A. Kitts and R.D. Possee, *Biotechniques* **14**, 810–817 (1993).
20. V.A. Luckow, S.C. Lee, G.F. Barry, and P.O. Olins, *J. Virol.* **67**, 4566–4579 (1993).
21. D.L. Jarvis and E.E. Finn, *Nat. Biotechnol.* **14**, 1288–1292 (1996).
22. L. Marz, F. Altmann, E. Staudacher, and V. Kubelka in H.S.J. Montreuil and J.F.G. Vliegenthart eds., *Glycoproteins*, Elsevier Science B.V., 1995, pp. 543–563.
23. F. Altmann, *Trends Glycosci. Glycotechnol.* **40**, 101–114 (1996).
24. D.J. Davidson, M.J. Fraser, and F.J. Castellino, *Biochemistry* **29**, 5584–5590 (1990).
25. O.W. Ogonah, R.B. Freedman, N. Jenkins, K. Patel, and B.C. Rooney, *Biotechnology* **14**, 197–202 (1996).
26. T.-A. Hsu, N. Takahashi, Y. Tsukamoto, K. Kato, I. Shimada, K. Masuda, E.M. Whiteley, J-Q. Fan, Y.C. Lee, and M.J. Betenbaugh, *J. Biol. Chem.* **272**, 9062–9070 (1997).
27. C.C. Yu Ip, W.J. Miller, M. Silberklang, G.E., Mark, R.W. Ellis, L. Huang, J. Glushka, H. Van Halbeek, J. Zhu, and J.A. Alhadeff, *Arch. Biochem. Biophys.* **308**, 387–399 (1994).
28. F. Altmann, H. Schwihla, E. Staudacher, J. Gloosl, and L. Marz, *J. Biol. Chem.* **270**, 17344–17349 (1995).
29. D.L. Jarvis and E.E. Finn, *Virology* **212**, 500–511 (1995).
30. R. Wagner, S. Liedtke, E. Kreztschmar, H. Geyer, R. Geyer, and H-D. Klenk, *Glycobiology* **6**, 165–175 (1996).
31. D.R. Thomson, L.E. Post, and A.P. Elhammer, *J. Cell. Biochem.* **43**, 67–79 (1990).
32. L.C. Iacono-Connors, C.S. Schmaljohn, and J.M. Dalrymple, *Infect. Immun.* **58**, 366–372 (1990).
33. D.C. Tessier, D.Y. Thomas, H.E. Khouri, F. Laliberte, and T. Vernet, *Gene* **98**, 177–183 (1991).
34. M.J. Betenbaugh, E. Ailor, E. Whiteley, P. Hinderliter, and T-A. Hsu, *Cytotechnology* **20**, 149–159 (1996).
35. M. Kidd, V.C. Emery, *Appl. Biochem. Biotechnol.* **42**, 137–159 (1993).
36. S. Maeda, *Curr. Opin. Biotechnol.* **6**, 313–319 (1995).
37. S. Maeda, *Ann. Rev. Entomol.* **34**, 351–372 (1989).
38. J.W. Rice, N.B. Rankl, T.M. Guranuss, C.M. Marr, J.B. Barna, M.M. Walters, and D.J. Burns, *BioTechniques* **15**, 1052–1059 (1993).
39. G.A. King, J. Kuzio, A.J. Daugulis, P. Faulkner, D. Bayly, and M.F.A Goosen, *Biotechnol. Lett.* **10**, 683–688 (1988).
40. S.N. Agathos, *Cytotechnology* **20**, 173–189 (1996).
41. M. Al Rubeai, N. Kioukia, and A.N. Emery, *Monitoring of Insect-Baculovirus Culture Processes by Flow Cytometry*, Cell Culture Engineering IV, San Diego, 1994.
42. D.A. Lindsay and M.J. Betenbaugh, *Biotechnol. Bioeng.* **39**, 614–618 (1992).
43. M.Y. Wang, T.R. Pulliam, M. Valle, V.N. Vakharia, and W.E. Bentley, *J. Biotechnol.* **46**, 243–254 (1996).
44. J.J. Chalmers, in M.L. Shuler, H.A. Wood, R.R. Granados, and D.A. Hammer eds., *Baculovirus Expression Systems and Biopesticides*, Wiley-Liss, New York, 1995, pp. 175–204.
45. J. Tramper, J.B. Williams, and D. Joustra, *Enzyme Microb. Technol.* **8**, 33–36 (1986).
46. A. Handa, A.N. Emery, and R.E. Spier, *Dev. Biol. Standard.* **66**, 241–253 (1987).
47. F. Bavarian, L.S. Fan, and J.J. Chalmers, *Biotechnol. Prog.* **7**, 140–150 (1991).
48. S.K.W. Oh, A.W. Nienow, M. Al-Rubeai, and A.N. Emery, *J. Biotechnol.* **12**, 45–62 (1989).

49. K.T. Kunas and E.T. Papoutsakis, *Biotechnol. Bioeng.* **36**, 476–483 (1990).
50. A. Handa-Corrigan, A.N. Emery, and R.E. Spier, *Enzyme Microb. Technol.* **11**, 230–235 (1989).
51. D.W. Murhammer and C.F. Goochee, *Biotechnology* **6**, 1411–1418 (1988).
52. D.W. Murhammer, *Appl. Biochem. Biotechnol.* **31**, 283–310 (1991).
53. M.J. Fraser, *Curr. Topics Microbiol. Immunol.* **158**, 131–172 (1992).
54. M. Drews, T. Paalme, and R. Vilu, *J. Biotechnol.* **40**, 187–198 (1995).
55. J.P. Ferrance, A. Goel, and M.M. Ataai, *Biotechnol. Bioeng.* **42**, 697–707 (1993).
56. R. Bhatia, G. Jesionowski, J. Ferrance and M.M. Ataai, *Cytotechnology* **20** (1996).
57. W.E. Bentley, M.-Y. Wang, and V. Vakharia, *Ann. N. Y. Acad. Sci.* **745**, 336–359 (1994).
58. T. Sugiura and E. Amman, *Biotechnol. Bioeng.* **51**, 494–499 (1996).
59. P. Licari, and J.E. Bailey, *Biotechnol. Bioeng.* **37**, 238–246 (1990).
60. J. Zhang, N. Kalogerakis, L.A. Behie, and K. Iatrou, *J. Biotechnol.* **33**, 249–258 (1994a).
61. K.T.K. Wong, C.H. Peter, P.F. Greenfield, S. Reid, and L.K. Nielsen, *Biotechnol. Bioeng.* **49**, 659–666 (1995).
62. P. Licari and J.E. Bailey, *Biotechnol. Bioeng.* **39**, 432–441 (1992).
63. J.F. Power, S. Reid, K.M. Radford, P.F. Greenfield, and L.K. Nielson, *Biotechnol. Bioeng.* **44**, 710–719 (1994).
64. R.E. Spier, J.B. Griffiths, W. Berthold, *Animal Cell Technol.* 493–503.
65. B. Nguyen, K. Jarnagin, S. Williams, H. Chan, and J. Barnett, *J. Biotechnol.* **31**, 205–217 (1993).
66. S. Reuveny, Y.J. Kim, C.W. Kemp, and J. Shiloach, *Biotechnol. Bioeng.* **42**, 235–239 (1993).
67. M.-Y. Wang, S. Kwong, and W.E. Bentley, *Biotechnol. Prog.* **9**, 355–361 (1993).
68. W.F. Hink and E.M. Strauss, in E. Kurstak, K. Maramorosch, and A. Dubendorfer eds., *Invertebrate Systems In Vitro*, Elsevier/North-Holland Biomedical Press, Amsterdam, 1980, pp. 27–33.
69. J. Zhang, N. Kalogerakis, L.A. Behie, and K. Iatrou, *Biotechnol. Bioeng.* **42**, 357–366 (1993).
70. J. Zhang, N. Kalogerakis, L.A. Behie, and K. Iatrou, *Biotechnol. Prog.* **10**, 636–643 (1994).
71. C.D. de Gooijer, R.H.M. Koken, F.L.J. van Lier, M. Kool, J.M. Vlak, and J.A. Tramper, *Biotechnol. Bioeng.* **40**, 537–548 (1992).
72. F.L.J. van Lier, J.P.T.W. van den Hombergh, C.D. de Gooijer, M.M. den Boer, J.M. Vlak, and J. Tramper, *Enzyme Microb. Technol.* **18**, 460–466 (1996).
73. J. Tramper and J.M. Vlak, *Ann. N. Y. Acad. Sci.* **469**, 279–288 (1986).
74. F.L.J. van Lier, van G.C.F. Duijnhoven, M.M.J.C.M. de Vaan, J.M. Vlak, and J. Tramper, *Biotechnol. Prog.* **10**, 60–64 (1994).
75. Y.C. Hu, M.Y. Wang, and W.E. Bentley, *Cytotechnology*, **24**, 143–152.
76. M. Kool, J.W. Voncken, F.L.J. van Lier, J. Tramper, and J.M. Vlak, *Virology* **183**, 739–746 (1991).
77. T.J. Wickham, T. Davis, R.R. Granados, D.A. Hammer, M.L. Shuler, and H.A. Wood, *Biotechnol. Lett.* **13**, 483–488 (1991).
78. M.J. Fraser, G.E. Smith, and M.D. Summers, *J. Virol.* **47**, 287–300 (1983).
79. H.J. Cha, M.Q. Pham, G. Rao, and W.E. Bentley, *Biotechnol. Bioeng.* **56**, 238–243 (1997).
80. J.S. Culp, H. Johansen, B. Hellmig, J. Beck, T.J. Matthews, A. Delers, and M. Rosenburg, *Biotechnology* **9**, 173–177 (1991).

See also AMMONIA TOXICITY, ANIMAL CELLS; CHINESE HAMSTER OVARY CELLS, RECOMBINANT PROTEIN PRODUCTION; CULTURE MEDIA, ANIMAL CELL, LARGE SCALE PRODUCTION; EXPRESSION SYSTEMS, *E. COLI*; EXPRESSION SYSTEMS, MAMMALIAN CELLS; GENE TRANSFER, GRAM POSITIVE BACTERIA; GLYCOSYLATION OF RECOMBINANT PROTEINS; INSECT CELL CULTURE, PROTEIN EXPRESSION; PHENYLALANINE.

INSECTICIDES, MICROBIAL PRODUCTION

SHUJI SENDA
TETSO OMATA
YASUO NINOMIYA
Nitto Denko Company
OsakaJapan

KEY WORDS

Asymmetric synthesis
Bond formation
Chiral
Enantiomers
Esterifcation
Hydrolysis
Lipase
Oxidation
Pheromone
Transesterification

OUTLINE

INTRODUCTION

Recently, the use of chiral chemicals in insect pest control has increased. Typically, these chiral chemicals are synthetic pyrethroid insecticides and pheromones, originally derived from natural compounds. The stereochemistry of chiral and bioactive products influences their bioactivity (1). In the case of Japanese beetle and gypsy moth pheromones, only one enantiomer is bioactive, but its antipode inhibits pheromone activity. In the case of olive fruit fly pheromone, one enantiomer is active in male insects, while the other is active in female insects. Thus, it is very important that chiral and bioactive chemicals are synthesized with high optical purity.

Methods of synthesizing chiral chemicals are as follows:

1. Use of natural chiral materials as starting substrates
2. Optical resolution
3. Asymmetric synthesis

Biochemical optical resolution and asymmetric synthesis using enzymes, microorganisms, or tissue cells have been actively researched, because these methods can produce chiral compounds with high optical purity. They have therefore contributed to many useful reviews and reports(2–6).

In this article we discuss recent chiral insecticide synthesis, especially of insect pheromones using microbial processes. We have classified these processes by their reaction types.

HYDROLYSIS REACTIONS

Lipase-Catalyzed Hydrolysis

In this field, enzymatic optical resolution using lipase is the most useful method studied by many researchers. Lipase has the following characteristics:

1. It is available at industrial production scale for a relatively low price.
2. It does not usually require coenzymes for catalytic reactions.
3. It has a remarkably wide substrate specificity.
4. It is stable in organic solvents.
5. It can catalyze both hydrolysis and esterification.

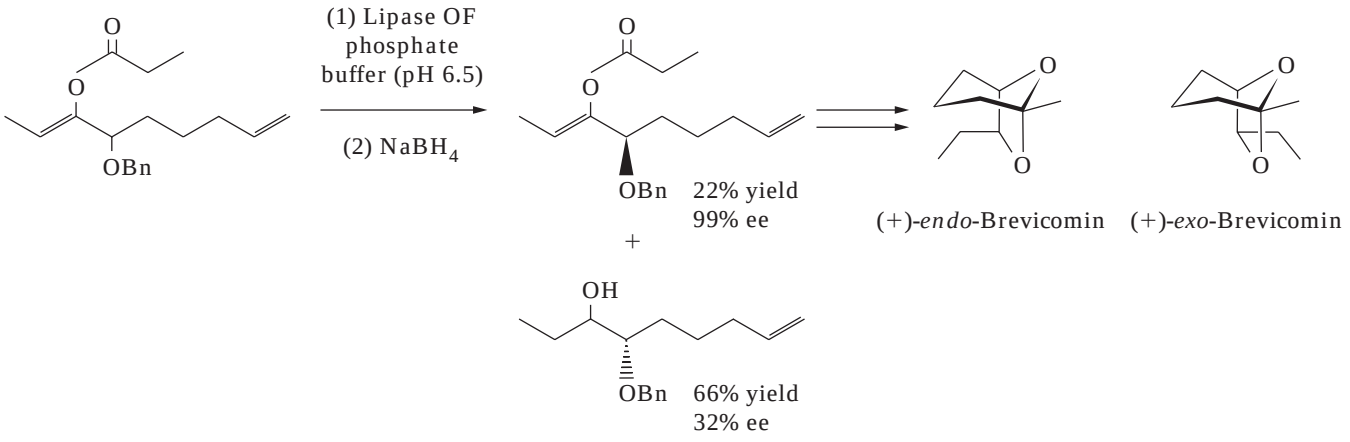
It is for these reasons that the use of lipase has been widely reported.

In the case of hydrolysis using lipase, optically pure (+)-*endo* and (+)-*exo*-brevicomin, the attractant pheromones of the bark beetle, were synthesized in short steps, starting from the α-hydroxy ketone derivative, (*R*)-4-benzyloxy-8-nonen-3-one, which was prepared via enzyme-mediated kinetic resolution of racemic 4-benzyloxy-2,8-nonadien-3-yl propanoate using Lipase OF (from *Candida cylindracea*, Meito Sangyo Co.) in phosphate buffer. The reaction was stopped at 22% yield to obtain high enantiomeric excess. The unused alcohol, (3*RS*,4*S*)-4-benzyloxy-8-nonen-3-ol, was then reused by conversion to ketone (7) (Scheme 1).

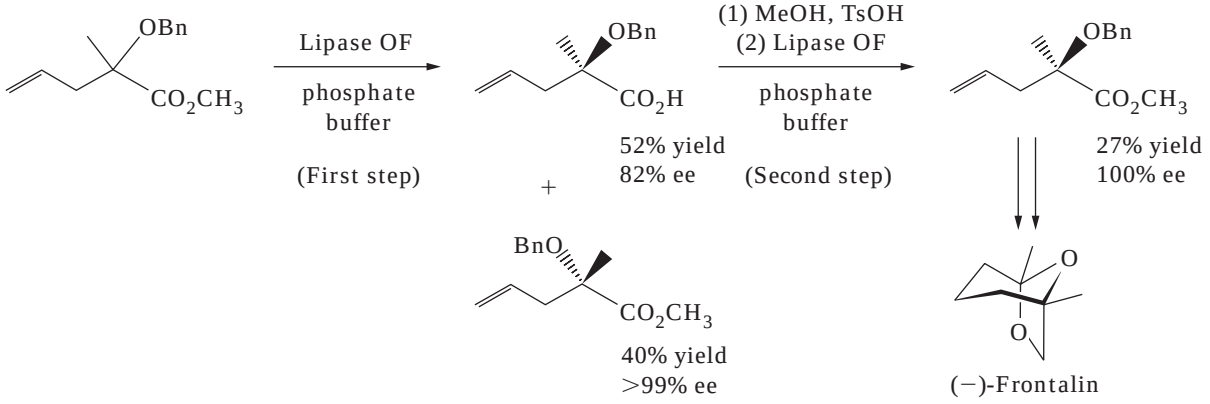
The enantioselective hydrolysis of tertiary α-benzyloxy esters [R-C(CH_3)(OCH_2Ph)CO_2CH_3] using the same lipase (Lipase OF) gave various enantiometrically enriched esters and acids on a preparative scale. The positive aspects of this reaction are (*1*) the racemic substrates are readily obtainable, (*2*) the method of the hydrolytic reaction is quite simple, and (*3*) the functional group of the product is adequately protected. Because the optical purity of carboxylic acid was low after the first reaction, a second reaction was needed to obtain an optically pure carboxylic acid. We call these reactions "two-step resolutions." Using the enantiometrically enriched carboxylic acid obtained above, (−)-frontalin was efficiently synthesized in nine steps to demonstrate the utility of this method (8) (Scheme 2).

Although discrimination of two enantiomers of a secondary alcohol whose hydroxy group is located in the middle of a long hydrocarbon chain is rather difficult, a two-step resolution of 6-methyl-2-heptyn-4-yl acetate using lipase (Lipase A, Amano Pharmaceutical Co.) was accomplished. The first step was enzymatic hydrolysis of racemic acetate. The second step of the reactionary alcohol was acetylation and enzymatic hydrolysis of this ester, with enzymatic hydrolysis and chemical hydrolysis of the anreactionary ester. These enantiomerically enriched alcohols were converted in both enantiomers of (*E*)-6-methyl-2-hepten-4-ol, the aggregation pheromone of the American palm weevil (9) (Scheme 3).

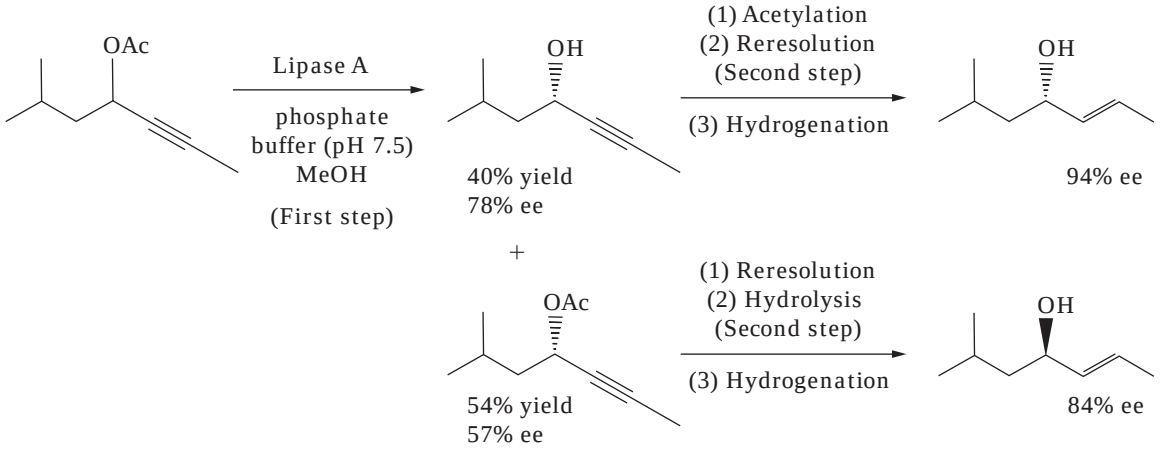
Naoshima et al. showed lipase (Lipase PS from *Pseudomonas cepacia*, Amano Pharmaceutical Co.) catalyzed hydrolysis of the acetate of racemic alkan-2-ol and alkan-3-ol carried out in an acetone–water solvent system. This system led to (*R*)-alcohol with 80–95% ee at 26–35% yield and (*S*)-acetate with 32–69% ee at 40–64% yield. They discussed enantioselectivity of enzymatic reactions using the E-value (enantiomeric ratio) developed by Chen et al. (10). The reaction with a large E-value is more enantioselective. In these hydrolysis reactions in acetone–water solvents, E-values were larger than in other solvent systems (e.g., water and methanol, water and ethylene glycol, water and diethyl ether, and water and hexane). In the case of hydrolysis of dodecan-2-yl acetate, the E-value in the acetone–water solvent was 65 compared to 6–25 in other solvent systems. Similarly, in the case of hydrolysis of dodecan-3-yl acetate, the E-value in the acetone–water solvent was 11, whereas in other solvent systems it was 5. The two-step resolution gave enantiomerically enriched alcohols as well. The increased enantioselectivity observed in the acetone–water reaction system was utilized in the synthesis of optically active natural products possessing alkan-2-ol, alken-2-ol, or alken-3-ol skeletons, such as (2*R*,6*R*,10*R*)-6,10,14-trimethylpentadecan-2-ol (the sex



Scheme 1. Hydrolysis of an enol ester (7).



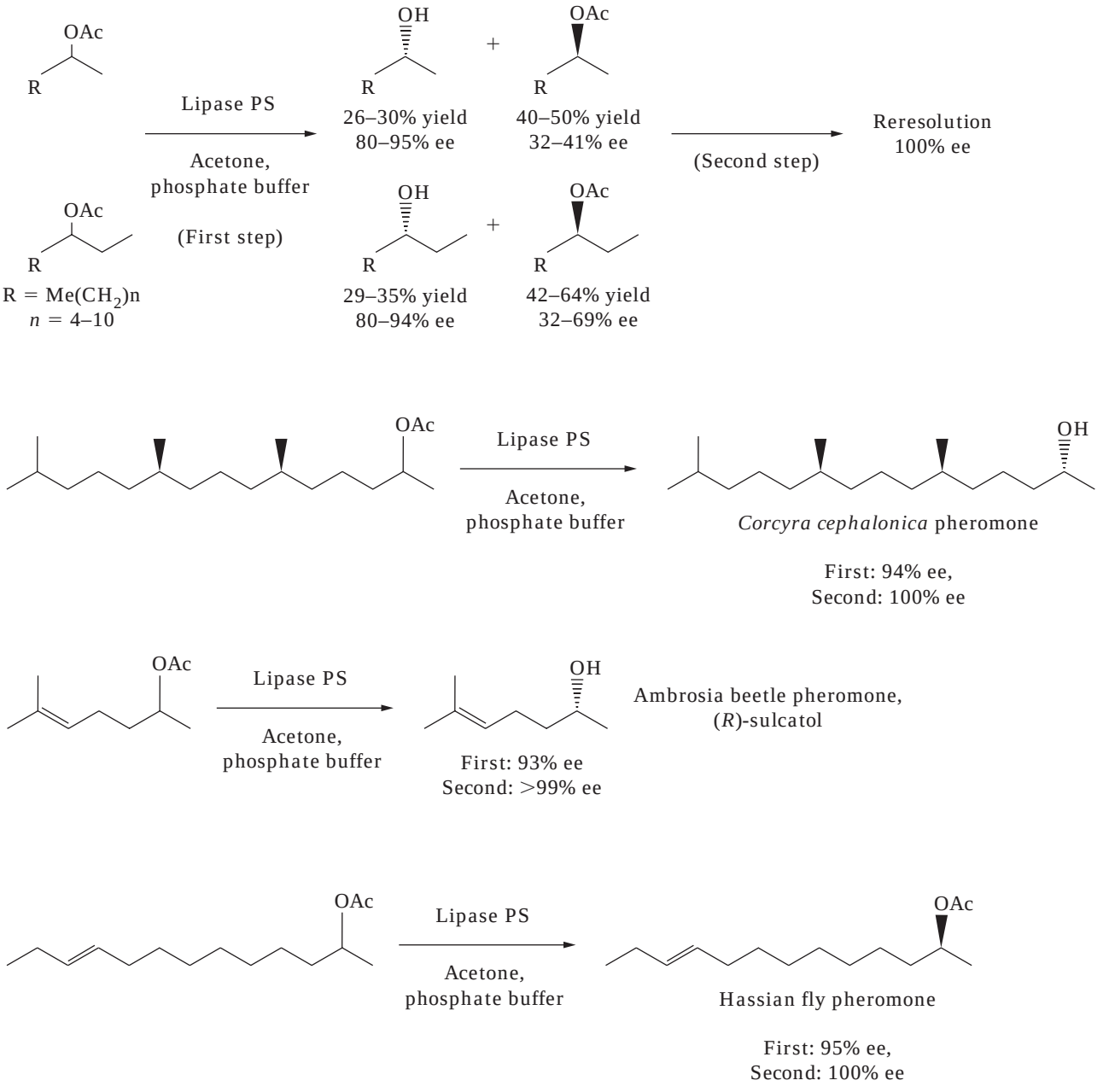
Scheme 2. Hydrolysis of α-benzyloxy acid ester (8).



Scheme 3. Hydrolysis of acetyloxy acetylene (9).

pheromone of *Corcyra cephalonica*), (*S*)-sulcatol (the aggregation pheromone of ambrosia beetles), (*S*,*E*)-10-tridecen-2-yl acetate (the sex pheromone of the hessian fly), and 3-octanol (pheromone of ants) (11–13) (Scheme 4).

Though the optical resolution of racemic starting material is a convenient method for obtaining both enantiomers, the yield of the desired enantiomer cannot exceed 50%. To overcome the low yield, an elaborate procedure, the asymmetric hydrolysis of *meso* substrate, has been developed, an example of which follows. Grandjean et al (14) synthesized optically pure seaweed pheromones dictyopterene A and C′. A key intermediate in this synthesis is



Scheme 4. Hydrolysis of 2- and 3-acetoxyhydrocarbons (11–13).

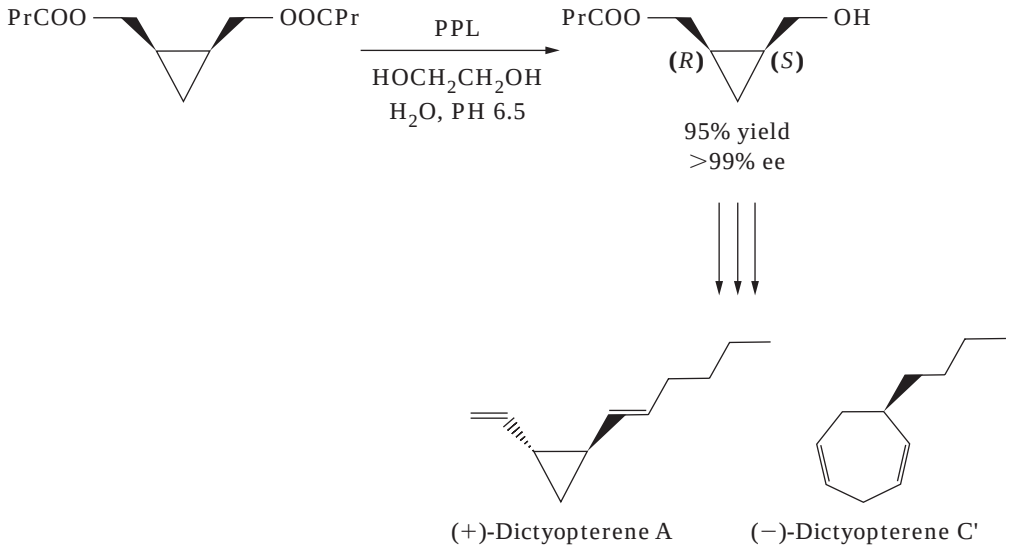
cis-(1*S*,2*R*)-1-hydroxymethyl-2-butyryloxymethylcyclopropane (>99% ee), which was easily obtained from the enantioselective hydrolysis of the *meso* diester, *cis*-1,2-bis(butyryloxymethyl)cyclopropane, with the pig pancreatic lipase (PPL, Sigma) in 95% yield (14) (Scheme 5).

PPL-catalyzed enantioselective hydrolysis of the *meso* epoxydiol diester *cis*-2,3-epoxybutan-1,4-yl diacetate yielded (2*S*,3*R*)-4-acetoxy-2,3-epoxybutan-1-ol. The yield of this enzymatic hydrolysis was 71% (90% ee). This enantiomerically enriched epoxy alcohol was converted into two naturally occurring epoxides, disparlure (gypsy moth pheromone) and ruby tiger moth pheromone (15) (Scheme 6).

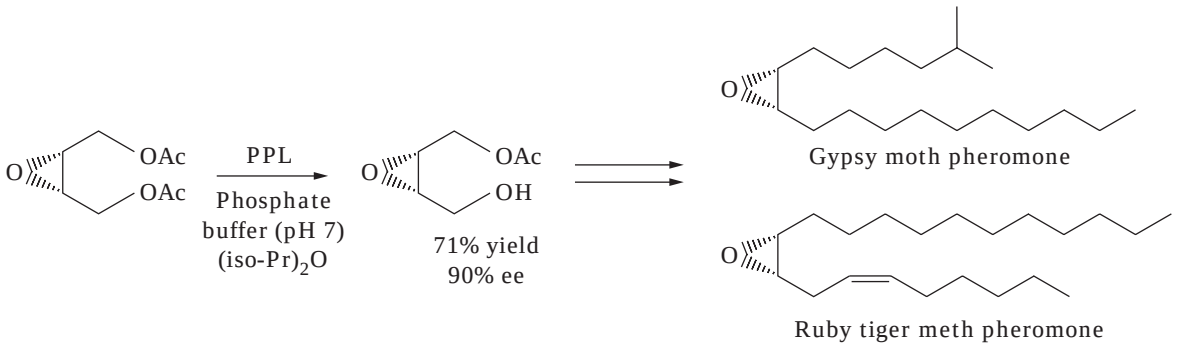
Immobilized lipase was examined for easy separation of product from catalyst. Naoshima et al. synthesized (2*S*,6*S*,10*S*)-6,10,14-trimethylpentadecan-2-ol, the sex pheromone of *Corcyra cephalonica*, in high enantiomerically pure form by the two-step resolution of the 2,2,2-trichloroethyl carbonate of (2*RS*,6*S*,10*S*)-6,10,14-trimethylpentadecan-2-ol with immobilized *Pseudomonas fluorescens* lipase (Amano Pharmaceutical Co.) on glass beads. Catalyst and products were separated by filtration (16).

Other Enzymes as Catalysts for Enantioselective Hydrolysis

Baker's yeast was used as a catalyst for enantioselective hydrolysis. Glanzer et al. prepared (*R*)- and (*S*)-1-alkyn-3-



Scheme 5. Hydrolysis of *meso*-cyclopropanediol diester (14).

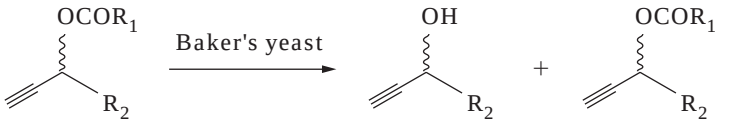


Scheme 6. Hydrolysis of *meso*-epoxydiol diester (15).

ols via enantioselective hydrolysis of their racemic esters. While lipase (AY30, SAM II, AK, P), fermenting microorganisms (*Rhizopus*), and freeze-dried microorganisms (*Brevibacterium*) failed in discriminating between enantiomers, lyophilized cells of baker's yeast gave (*S*)-alkyn-3-ols and their corresponding (*R*)-esters with >90% ee (17) (Scheme 7).

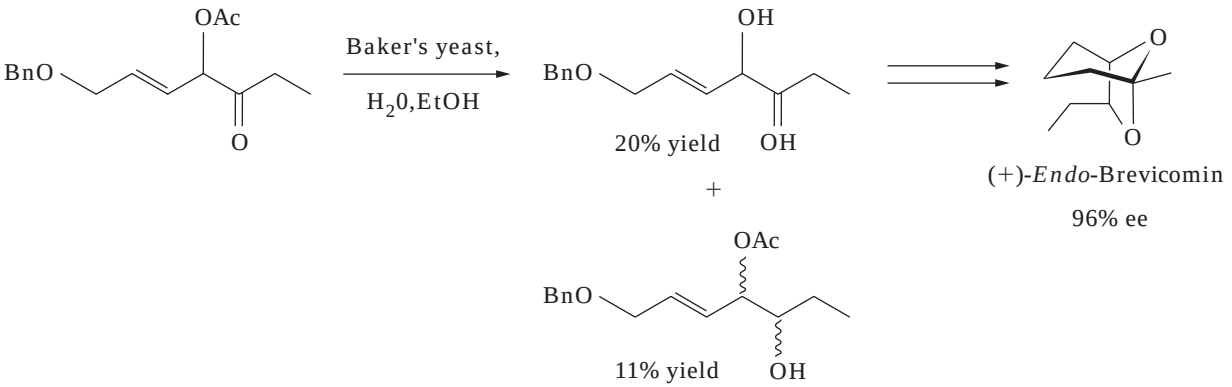
1-Benzyloxy-5-oxo-2-hepten-4-yl acetate was transformed to (4*R*,5*S*)-1-benzyloxy-2-hepten-4,5-diol in fermenting yeast. Prior to the reduction, the acetate was hydrolyzed. Although the high enantiomeric excess was observed for the diol, the hydroxy acetate arising from the direct reduction of acetoxy ketone was not in a diastereomerically or enantiomerically pure state. Thus, the diol arose from hydrolysis reduction. (4*R*,5*S*)-1-Benzyloxy-2-hepten-4,5-diol was transformed into (+)-*endo*-brevicomin with high optical purity (96.5% ee) (18) (Scheme 8).

In one particular case using microbial processes, hydrolysis and the opening of an epoxide occurred at the same time. Treatment of a diastereomeric mixture of 5,6-epoxy-6-methyl-2-heptyl pentanoate with crude immobilized enzyme preparation derived from *Rhodococcus* sp. (Novo SP 409) or whole lyophilized cells of *Rhodococcus erythropolis* (NCIB 11540) in buffer solution led to (2*R*,5*R*)-2-(1-hydroxy-1-methylethyl)-5-methyltetrahydrofuran (pityol, pheromone of the elm bark beetle) and its dia-



$R_1 = CH_3, C_3H_7, CH_2Cl$ 70–96% ee 14 to >95% ee
$R_2 = C_8H_{17}, C_8H_{15}, C_3H_6COOEt$

Scheme 7. Hydrolysis of 1-alkyn-3-yl ester using baker's yeast (17).



Scheme 8. Hydrolysis and reduction of ketoacetate using baker's yeast (18).

stereomer. Because pityol was not formed from 5,6-epoxy-6-methyl-2-heptanol in absence of a biocatalyst, it was concluded that this epoxide-opening reaction was triggered by an enzyme (19) (Scheme 9).

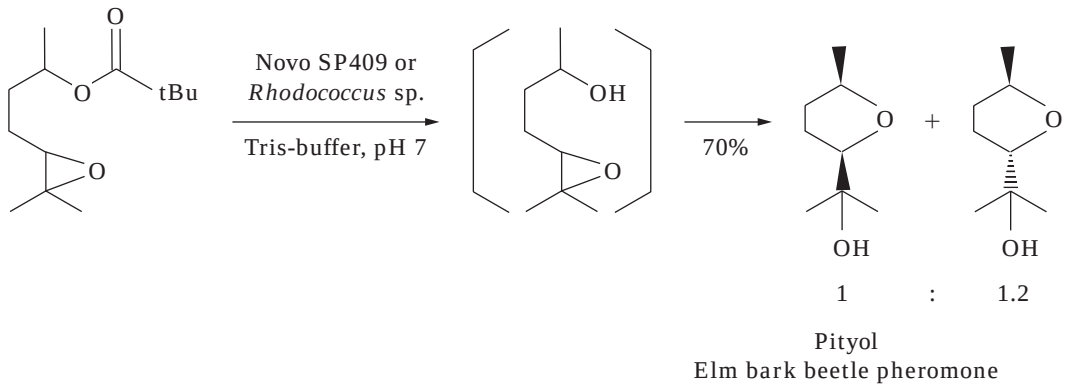
ESTERIFICATION AND TRANSESTERIFICATION REACTIONS

Lipase as Esterification and Transesterification Catalyst

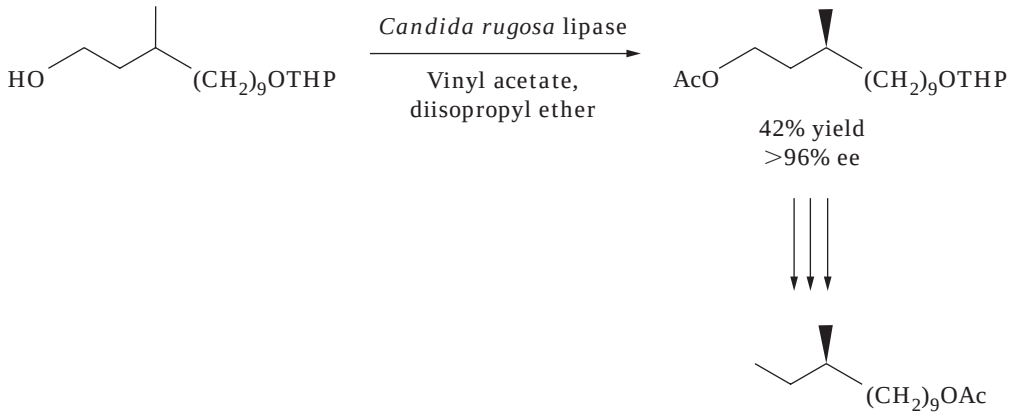
In many cases, esterification and transesterification using lipase are performed in organic solvent (20). This has important implications for their application, as water is not the ideal reaction medium. In preparative organic synthesis, many compounds dissolve better in organic solvents than in water, and indeed some compounds are unstable in water. Lipase is uniquely stable in organic solvent. 12-Tetrahydropyranyloxy-3-methyldodecanol was resolved via a *Candida rugosa* lipase-catalyzed acetylation with vinyl acetate in diisopropyl ether. (*S*)-enantiomer was acetylated at >96% ee. Some chiral insect pheromones were synthesized from this material (21) (Scheme 10).

An intramolecular version of esterification and transesterification produces lactones. For example, Sugar et al. (22) synthesized the Japanese beetle pheromone (*R*,*Z*)-5-(1-decenyl) oxacyclopentan-2-one by means of the enzymatic enantioselective lactonization of racemic methyl 4-hydroxy-5-tetradecynoate, using porcine pancreatic lipase (PPL) in polar solvent (diethyl ether) at a somewhat low temperature (10–15 °C). The desired (*R*)-lactone was obtained (43% yield, 82% ee). After repetition of the enzymatic lactonization, as shown in Scheme 11, lactone with higher enantiomeric excess (97%) was obtained. The natural enantiomer of Japanese beetle pheromone was prepared from this compound by semihydrogenation.

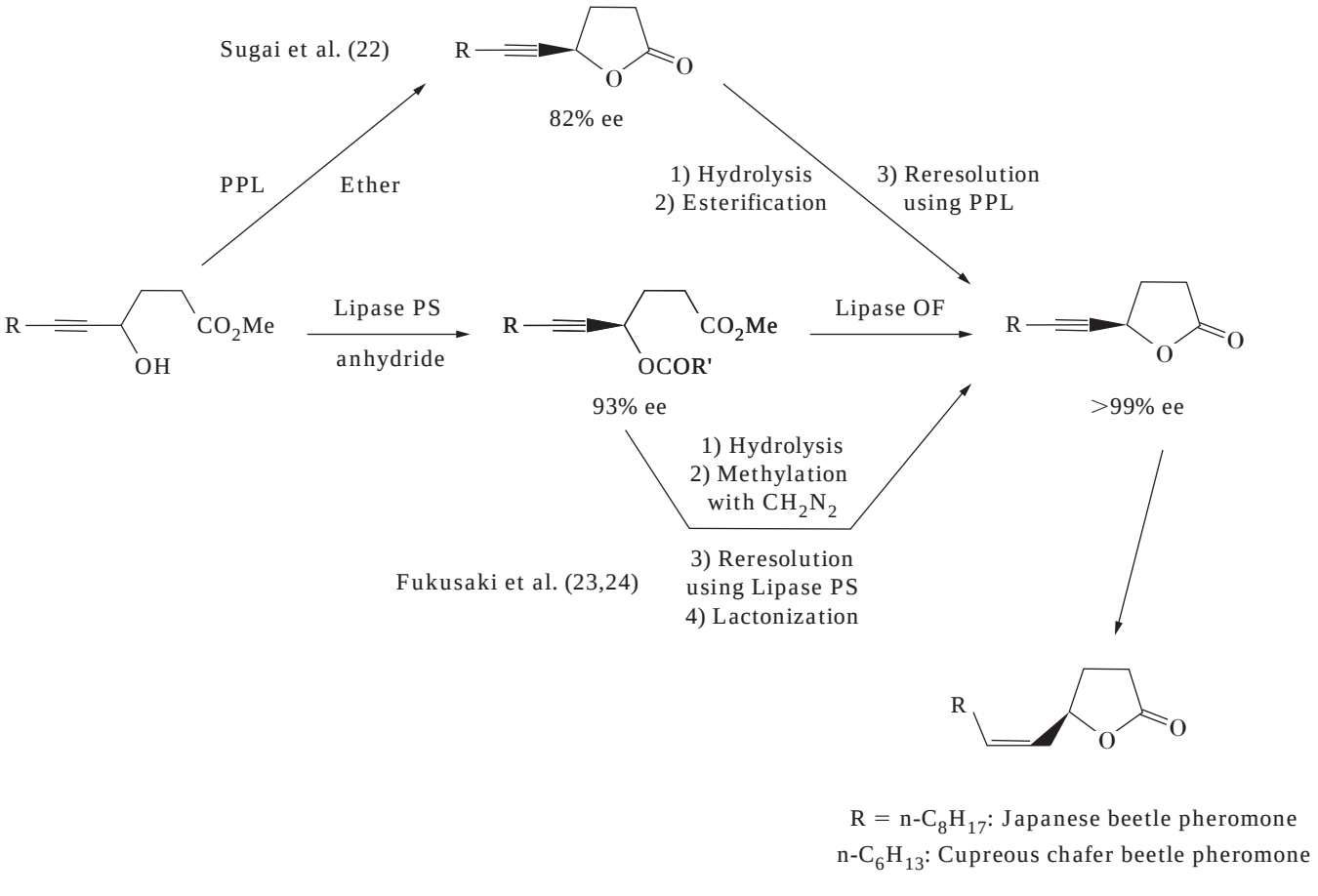
Fukusaki et al (23) also synthesized the Japanese beetle pheromone by lipase-catalyzed enantioselective acylation and lactonization. An acylation of methyl 4-hydroxy-5-tetradecynoate with succinic anhydride was accomplished by Lipase PS (from *Pseudomonas* sp., Amano Pharmaceutical Co.) in diisopropyl ether. This partially enantiomerically enriched diester was subsequently converted to (*R*)-5-(1-decynyl)oxacyclopentan-2-one (over 99% ee) by lipase-catalyzed enantioselective lactonization with Lipase OF (from *Candida cylindracea*, Meito Sangyo Co.) in phosphate buffer (23). It should be emphasized that even in aqueous solution, a lactonization (not hydrolysis) took place. The sex pheromone of the cupreous chafer beetle, (*R*,*Z*)-5-(1-octeny)oxacyclopentan-2-one, was synthesized in a similar manner. Using lipase-catalyzed enantioselective acylation by Lipase PS (Amano Pharmaceutical Co.) in diisopropyl ether, (*R*)-enantiomer of methyl 4-hydroxy-



Scheme 9. Enzyme-triggered opening of epoxide (19).



Scheme 10. Transesterification using lipase (21).



Scheme 11. Lactonization using lipase.

5-dodecynoate was preferentially converted to give (*R*)-acyloxy ester (93% ee). Further enantiomeric purification was carried out by repetition of the enzymatic reaction (24). Both Sugai et al. and Fukusaki et al. employed E-values to evaluate the enantioselectivity of the enzymatic reaction. The enantioselectivity of lactonization by lipase was not so high. The E-value of the reaction obtained by Sugai was 20, and that by Fukusaki was 39. However, the enantioselectivity of acylation by lipase was high. The E-value of an acylation of methyl 4-hydroxy-5-tetradecynoate with succinic anhydride by Lipase PS was 115 (Scheme 11).

Henkel et al. (25) synthesized (5*R*,6*S*) and (5*S*,6*R*)-6-acetoxyhexadecan-5-olide (the major component of mosquito oviposision attractant pheromone) using an enantioselective lipase-catalyzed lactonization and the sub-

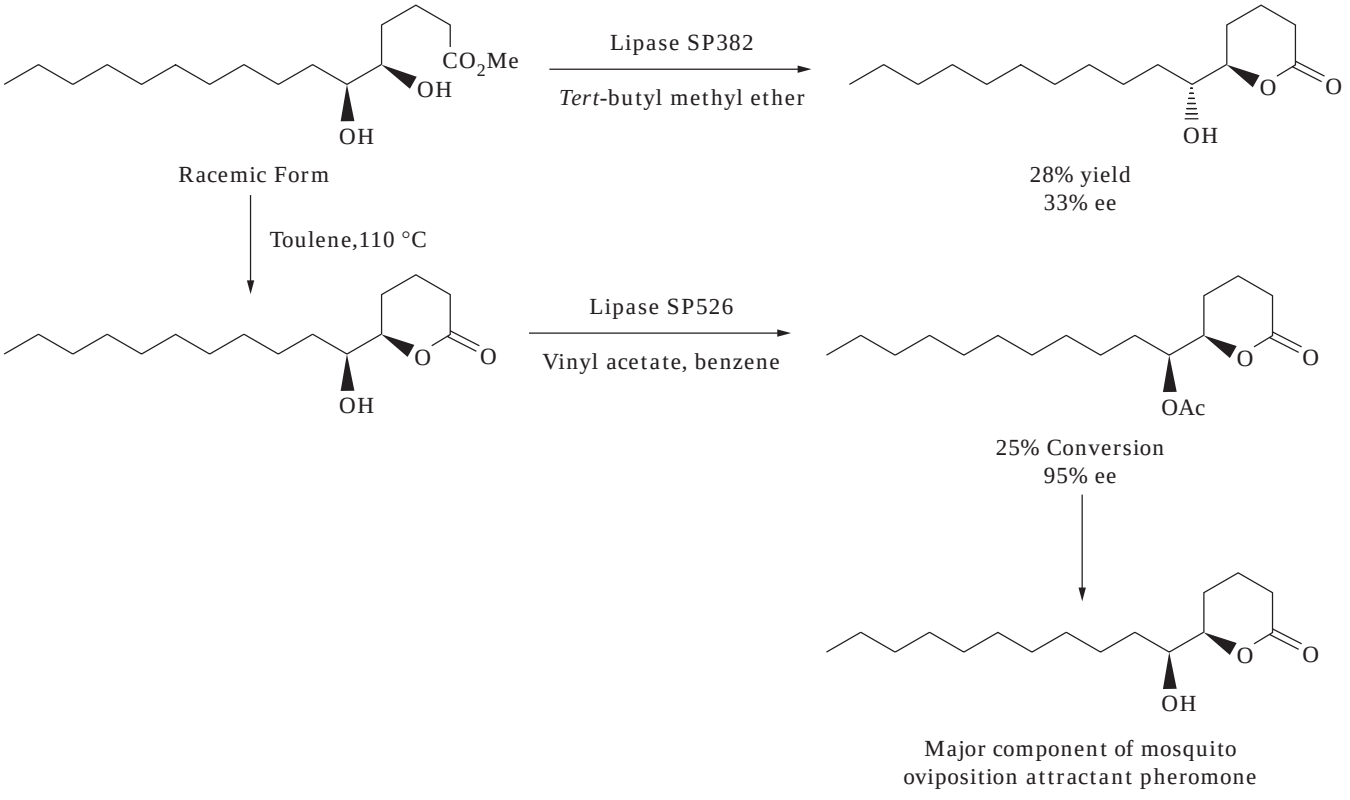
sequent acetylation with vinyl acetate (Lipase SP 382 and SP 526, Novo Nordisk A/S). Acetylation was achieved at high temperatures (55–70 °C). In this case, the enantioselectivity of lactonization was also low and the selectivity of acetylation was comparatively high (E-value of 53) (Scheme 12).

Enzymatic resolution using γ- and δ-hydroxysulfone for syntheses of γ- and δ-lactone was attempted, rather than using hydroxycarbonic acid or ester. Jacobs et al. (26) demonstrated the resolution of racemic hydroxysulfone using a lipase (Lipase PS-30). *R*-(+)-4-hexanolide (the sex pheromone of the female dermestid beetle *Trogoderma glabrum*) and (2*R*,5*S*)-2-methyl-5-hexanolide (the sex pheromone of the carpenter bee) were synthesized via an intramolecular acylation of the carbonates derived from these sulfone intermediates, which provided a new and simple route for syntheses of chiral γ- and δ-lactones. However, the enantioselectivity of this enzymatic reaction was not very high (E-values of 32 for 3-hydroxysulfone and 27 for 4-hydroxysulfone) (Scheme 13).

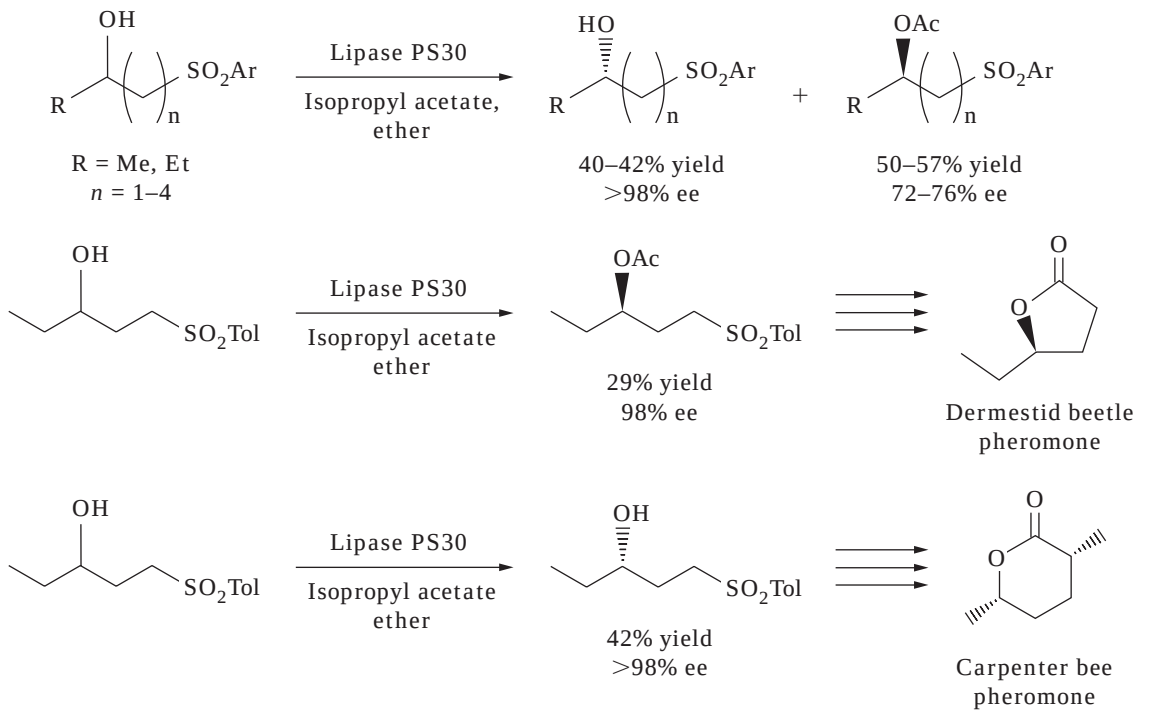
Enzymatic reactions were also efficiently combined with chemically mediated diastereomeric control to set up the enantiomerically enriched building blocks with defined multiple chiral centers. Kim et al. (27) efficiently resolved *anti*-1,2-diol by using lipase PS-catalyzed transesterification and achieving the total synthesis of the bark beetle pheromone (+)- and (−)-*endo*-brevicomin, starting from resolved diols.

Enzymatic differentiation of enantiotopic groups in *meso* compounds (asymmetrization or desymmetrization) provided enantiomerically pure compounds with multiple stereogenic centers in high yield. Mori et al. (28) synthesized (2*R*,3*S*)-2,3-(isopropylidendioxy)-4-hydroxypentan-1-yl acetate, the key intermediate for the synthesis of (+)-*endo*-brevicomin. *Meso*-2,3-isopropylidendioxy-1,4-butanediol was converted to the optically active monoacetate with vinyl acetate and lipase AK (from *Pseudomonas*, Amano Pharmaceutical Co.) in high chemical and optical yield (96% yield, 98.5% ee). The enzymatic esterification of *meso*-2,6-dimethyl-1,7-heptanediol with isopropenyl acetate in the presence of *Pseudomonas cepacia* lipase (Lipase PS30, Amano Pharmaceutical Co.) in organic medium provided the enantiomerically enriched monoester with high enantiomeric excess. In this case, the chemical yield was not very high, but it was more than 50%. This compound was used as the synthetic intermediate of the insect pheromone tribolure (29) (Scheme 14).

To solve the problem of unsatisfactory enantiomeric excess of the product in the course of a kinetic resolution, a sequential use of hydrolysis and acylation with lipase has been demonstrated in order to secure further enantiomerically enriched products. This is one pattern of a two-step resolution. Naoshima et al. (30) synthesized (*E*)-1-ethyl-5-methyl-4-heptenyl acetate, the aggregation pheromone of *Cathartus quadricollis* (square-necked grain beetle) using a two-step lipase reaction. The first reaction was hydrolysis of (*E*)-1-ethyl-5-methyl-4-heptenyl acetate with Lipase PS in a phosphate buffer and acetone. The second reaction was transesterification of the product, (1*R*,4*E*)-alcohol (98% ee) with porcine pancreas lipase



Scheme 12. δ-Lactonization using lipase (25).



Scheme 13. Enzymatic resolution of hydroxysulfones (26).

(PPL, Sigma) in isopropenyl acetate and ether (30). The second example is also shown in Scheme 15. Fukusaki et al. (31) achieved a yield of enantiomerically pure epoxy alcohol using lipase-catalyzed acylation and hydrolysis (Pancreatin F, Amano Pharmaceutical Co.). In this case, the same lipase immobilized on celite was used in both lipase-catalyzed reactions. These reactions were performed at 5 °C. However, the enantioselectivity of these reactions was not very high: The E-value was 23 for acylation and 20 for hydrolysis. Hence, the recrystallization of corresponding 3,5-dinitrobenzoate of this epoxy alcohol was performed to get an enantiomerically enriched epoxy alcohol. This is the intermediate of disparlure (the gypsy moth pheromone).

(+)-*Exo*, (+)-*endo*-brevicomin, and (−)-frontalin were synthesized via lipase-catalyzed acylation of oxygen heterocyclic structure, 2-(1-hydroxypropyl)-6-methyl-3,4-dihydro-2H-pyran and 2-hydroxymethyl-2,6-dimethyl-3,4-dihydro-2H-pyran. Three enzymatic kinetic resolutions were combined in frontalin synthesis to acquire a high enantiomeric excess of the product. The first reaction was hydrolysis of 2-methoxycarbonyl-2,6-dimethyl-3,4-dihydro-2H-pyran with protease; the second reaction was transesterification with Lipase MY (Meitou Sangyo Co.), and the third reaction was hydrolysis with Lipase P. As a result, enantiomeric excess of the pheromone reached 91% (32) (Scheme 15).

REDUCTIONS

Baker's Yeast as a Catalyst for the Reduction

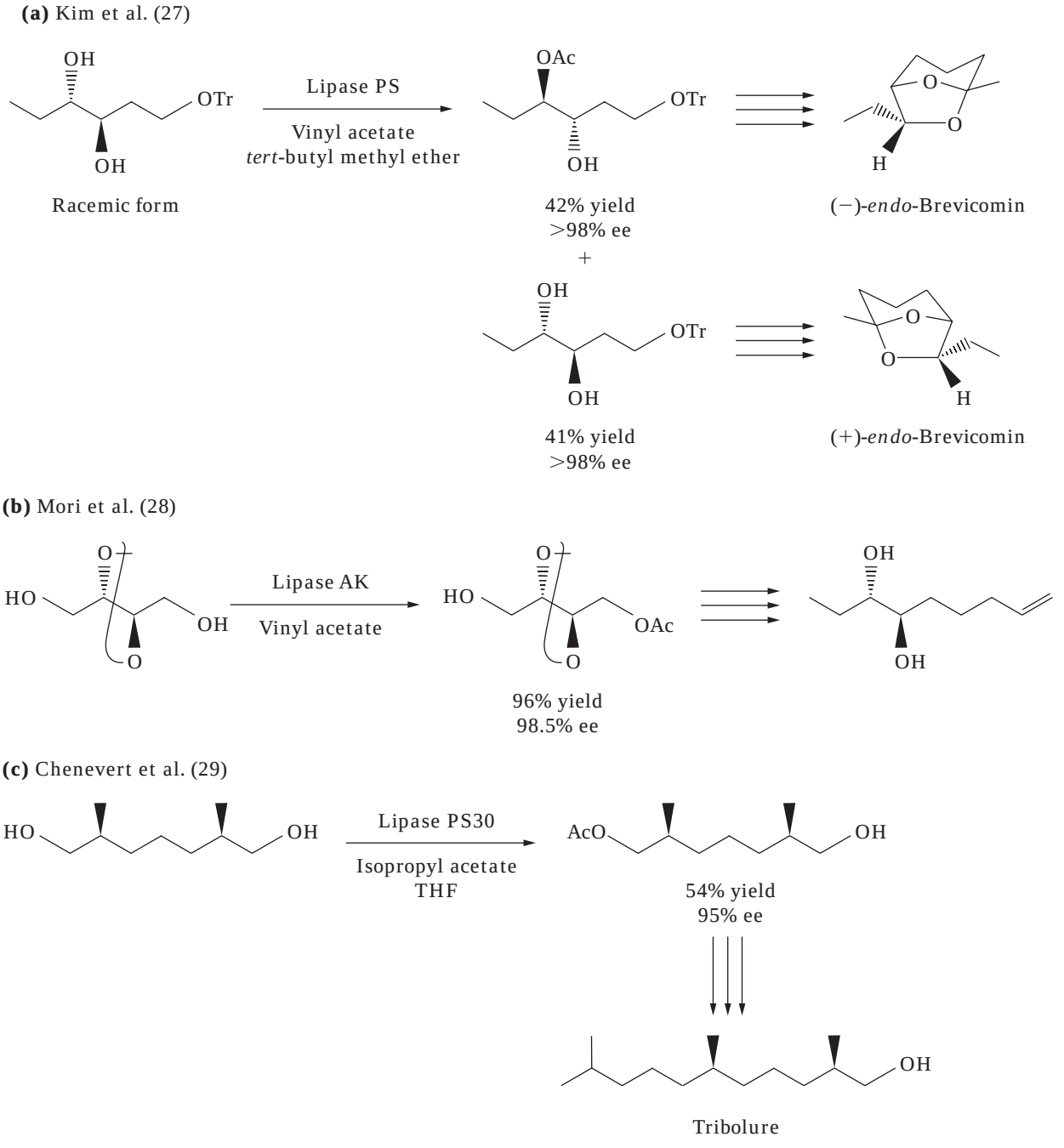
The reduction of ketones is one of the most established microbial processes and has been extensively investigated since the 1960s, when pioneering studies revealed the stereochemistry of the reduction in a general sense (33,34). Alcohol with a high enantiomeric excess can be obtained using this method.

Acetoacetylated Meldrum's acid (1,3-dioxan-4,6-dione) was found to be reduced enantioselectively (>99% ee) by baker's yeast to the corresponding (*S*)-alcohol. Sato et al. (35) converted this (*S*)-alcohol to *cis*-3,6-dimethyl tetrahydropyran-2-one, a major component of pheromonal secretion of the male carpenter bee (*Xylocopa*).

3-Oxo-6-heptenoic acid salt was reduced to the enantiomerically enriched hydroxy acid (>99% ee) by baker's yeast as a key step. This enantiomerically enriched hydroxy acid was further converted to (*S*)-4-hexanolide, the pheromone antipode of the store grain beetle (36) (Scheme 16).

(*S*)-4-(Phenylthio)-2-pentanol was sufficiently produced with high enantiomeric excess by enzymatic reduction of the corresponding ketone using baker's yeast. Using this chiral compound, highly enantiomerically enriched insect pheromonal compounds having spiroketal structures were synthesized (37) (Scheme 17).

(2*R*,5*S*)-Pityol (the pheromone of the bark beetle, *Pityophthorus pityographus*) was synthesized using two highly stereoselective biocatalytic reactions. 6-Methyl-2-oxo-5-heptene was reduced to (*S*)-alcohol (sulcatol, the pheromone of *Gnathotrichus sulcatus*) (98.5% ee) by baker's yeast. (*S*)-Sulcatol was converted to urethane by way of phenylisocyanate, and urethane then underwent epoxydation by the fungus *Aspergillus niger*. Treatment of this epoxide major product with sodium hydroxide led to formation of (2*R*,5*S*)-pityol (100% ee) (38) (Scheme 18).

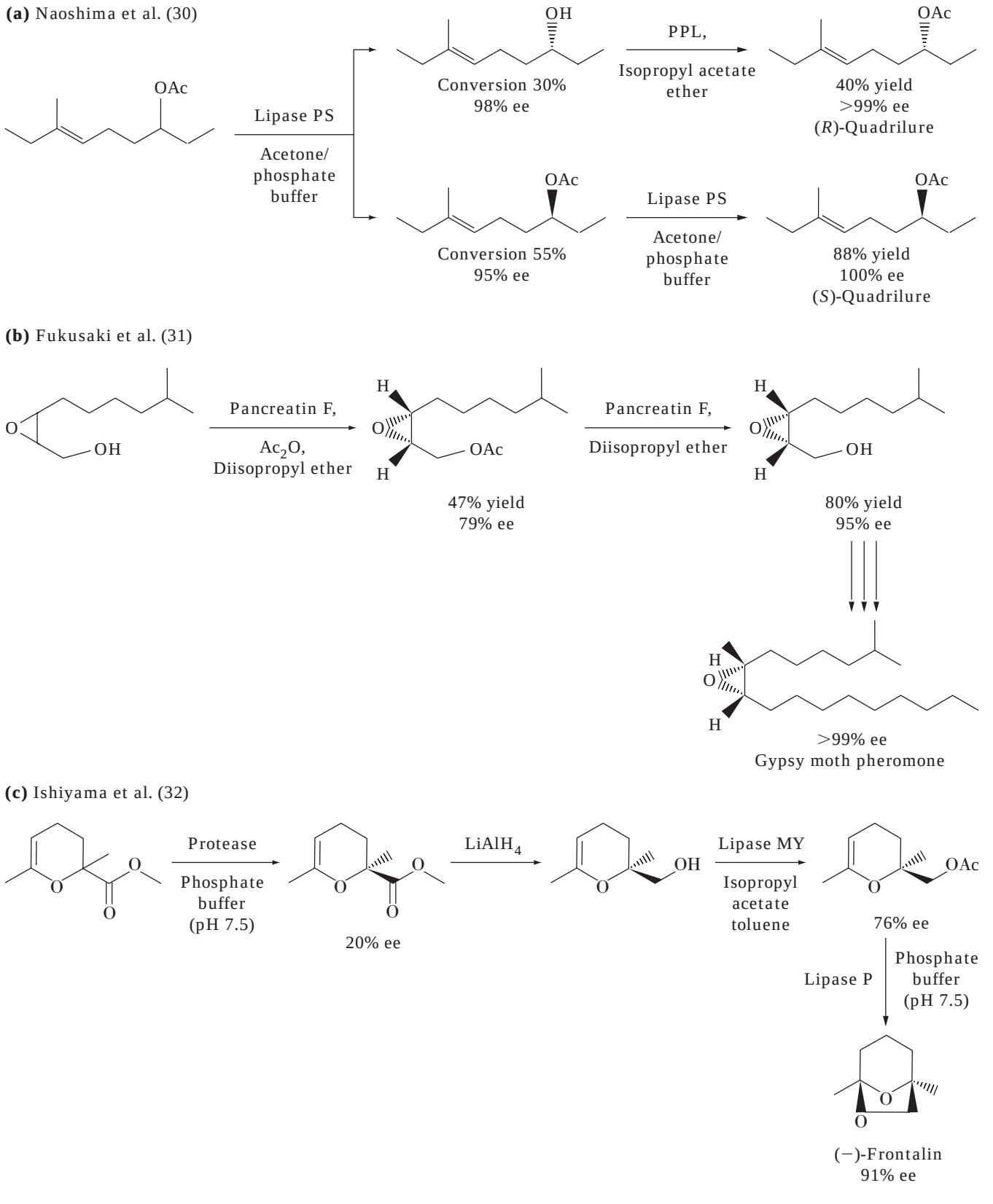


Scheme 14. Enzymatic resolution of the compounds with multistereogenic center.

Other Enzymes as Catalysts for the Reduction

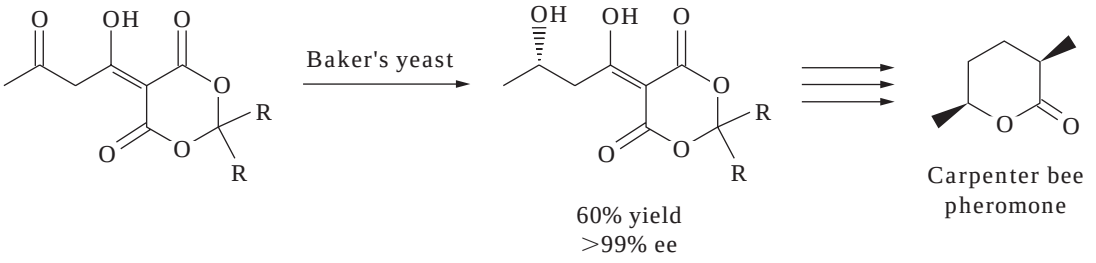
A series of studies using *Thermoanaerobium brockii* alcohol dehydrogenase (TBADH) have been reported. Highly enantioselective reduction of various methyl- and ethyl-ketones having a wide range of functional groups, such as methyl ester, cyano, ethyl ether, phenyl, and chloride as well as double and triple bonds, is carried out using TBADH as a catalyst. Reduction of most of the substrates yields alcohol with an (*S*)-configuration, arising from highly selective hydride attack at the *re* face of the carbonyl groups. (*S*)-(+)-(*Z*)-Tetradec-5-en-13-olide, one of several synergistic aggregation pheromones produced by the male flat grain beetle was prepared from (*S*)-(+)-methyl-8-hydroxynonanoate with a high enantiomeric excess (>99%) in a six-step synthesis (39).

(*S*)-(+)-Methyl 8-hydroxynonanoate, which was obtained via enantioselective reduction of methyl 8-oxononanoate with TBADH, was converted to (*S*)-(+)-ferrulactone II, the aggregation pheromone of the flat grain beetle (40).

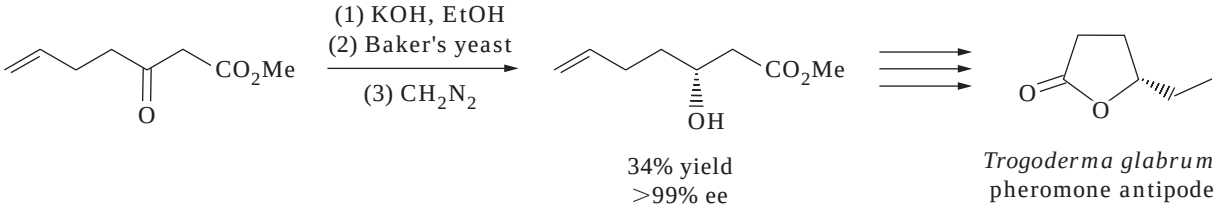


Scheme 15. Hydrolysis and esterification in one synthesis.

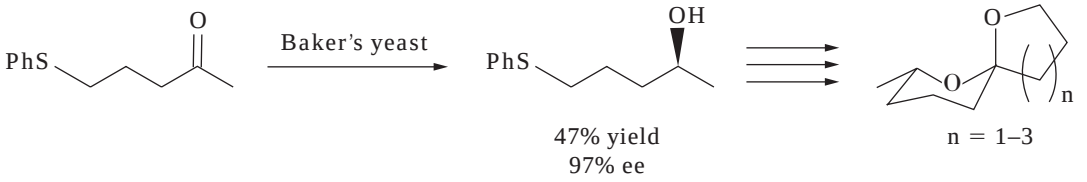
(a) Sato et al. (35)



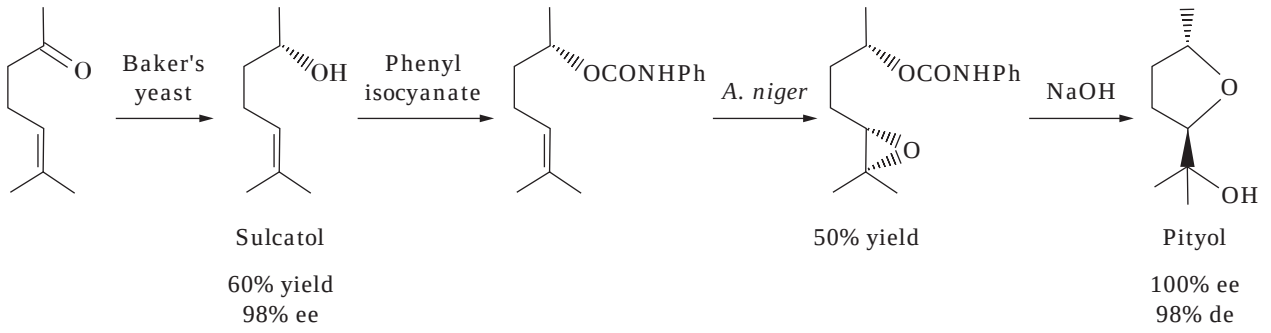
(b) Ghosh et al. (36)



Scheme 16. Reduction using baker's yeast.



Scheme 17. Reduction of ketosulfide using baker's yeast (37).



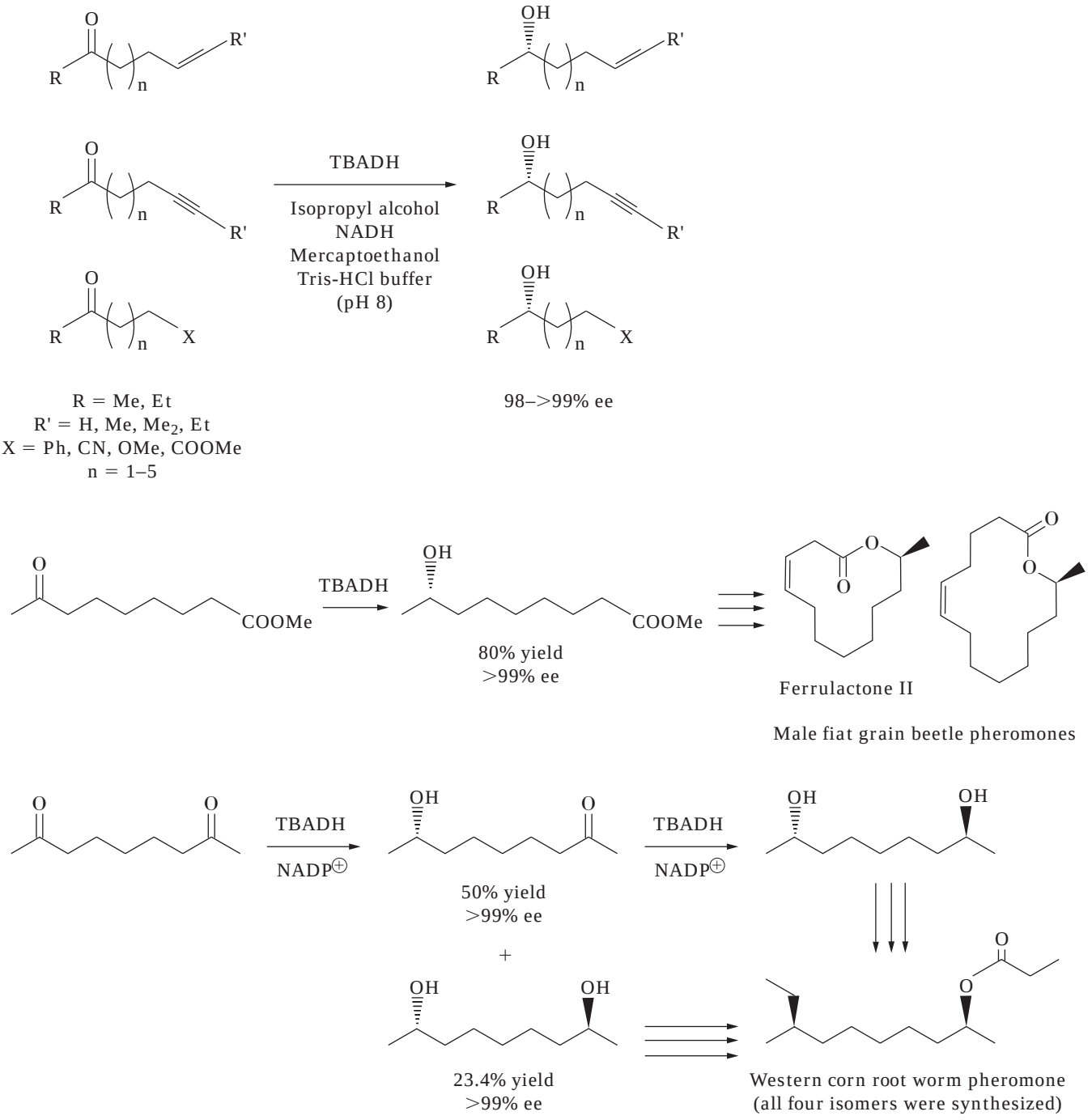
Scheme 18. Reduction and epoxydation using microbiological methods (38).

(2*S*,8*S*)-2,8-Dihydroxynonane, which was characterized by synthetically useful C2 symmetry, was also obtained by the TBADH-catalyzed reaction of nonan-2,8-diol. Keinan et al. (41) synthesized four isomers of 8-methyldec-2-yl propanoate (the western corn rootworm sex pheromone) using this single chiral building block (Scheme 19).

OXIDATION REACTIONS

Microbial Baeyer–Villiger oxidations are important reactions in providing enantiomerically enriched lactones. Asymmetric synthesis of 5-hexadecanolide (oriental hornet pheromone) using a biological Baeyer–Villiger reaction was achieved by two strains of *Acinetobacter*. This allows a three-step synthesis of (*S*)-(−)-5-hexadecanolide, as well as direct access to (*R*)-(+)-2-undecylcyclopentanone, a precursor for chemical synthesis of (*R*)-(+)-5-hexadecanolide. These products were obtained with optical purities of 74 and 95% (42).

Lebreton et al. (43) performed novel microbiological Baeyer–Villiger oxidation using fungi. Racemic cyclobutanone derivatives were oxidized to give an enantiomerically enriched γ-lactone with high enantiomeric excess us-



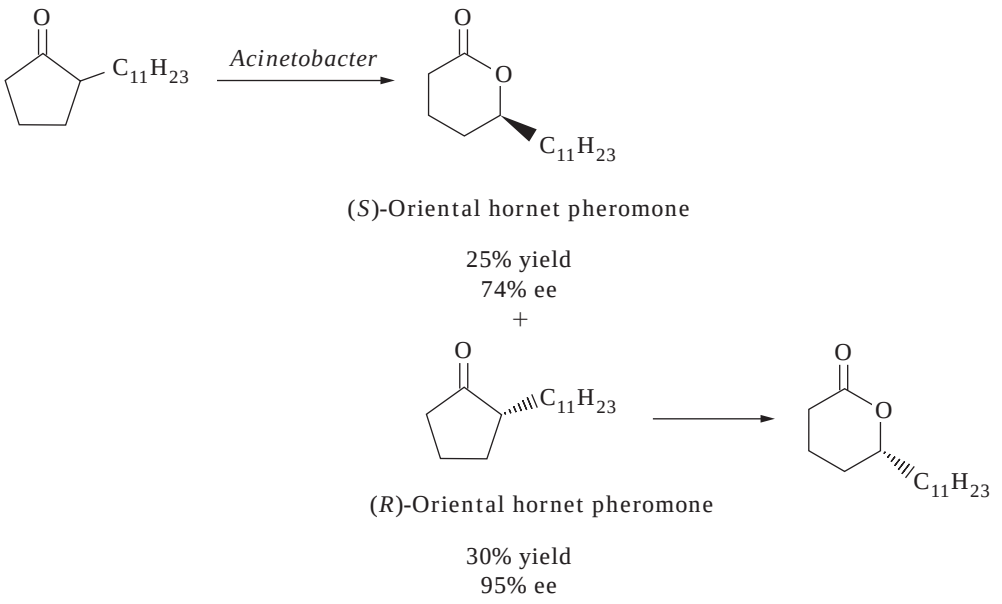
Scheme 19. Reduction using TBADH (39–41).

ing *Cunninghamella echinurata*. They synthesized (+)-multifidine and (+)-viridiene, the major components of a brown algae pheromone (Scheme 20).

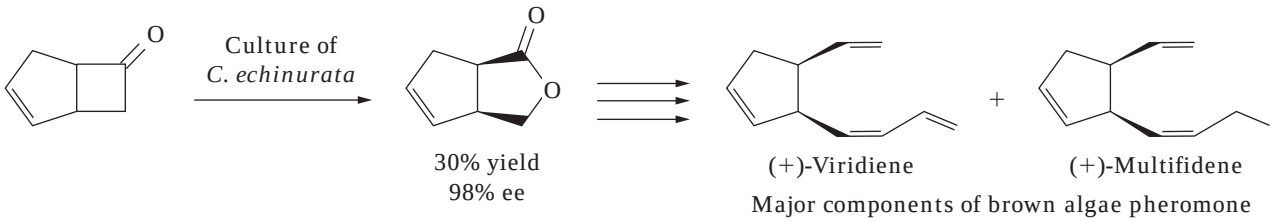
C—C BOND FORMATION REACTIONS

Two enzymatic C—C bond formation reactions of carbohydrate-related compounds were demonstrated in pheromone synthesis. The first example is fructose 1,6-diphosphate (FDP) aldolase from rabbit mussle (RAMA). This aldolase was used in the synthesis of (+)-(1*S*,5*R*,7*S*)-*exo*-brevicomin. This enzyme accepts 5-oxohexanal, as well as the corresponding 5-dithioacetal, as its substrate, and catalyzes the stereospecific aldol reaction of dihydroxyacetone phosphate (DHAP) and aldehyde (80–90% yield). After removal of the phosphate group from the aldol adduct, the product was converted into the desired bicyclic system. (+)-*exo*-Brevicomin was obtained by a deoxygenation of the side chain (44).

(a) Alphand et al. (42)



(b) Lebreton et al. (43)



Scheme 20. Microbiological Baeyer–Villiger oxidation.

Myles et al (45) reported that the commercially available enzyme preparation of transketolase catalyzed the condensation of β-hydroxypyruvic acid and 2-hydroxybutylaldehyde to furnish the enantiomerically enriched dihydroxyketone with vicinal diol of *D-threo* configuration. The remarkable advantage in this is the stereoselective formation of vicinal diols of a desirable configuration, even when starting from a racemic 2-hydroxyaldehyde. Indeed, transketolase accepts only the *D*-enantiomer of 2-hydroxyaldehyde and produces the *threo* isomer of the product with very high diastereoselectivity. This dihydroxyketone with *D-threo* configuration was transformed into (+)-*exo*-brevicomin in six steps (Scheme 21).

CONCLUSION

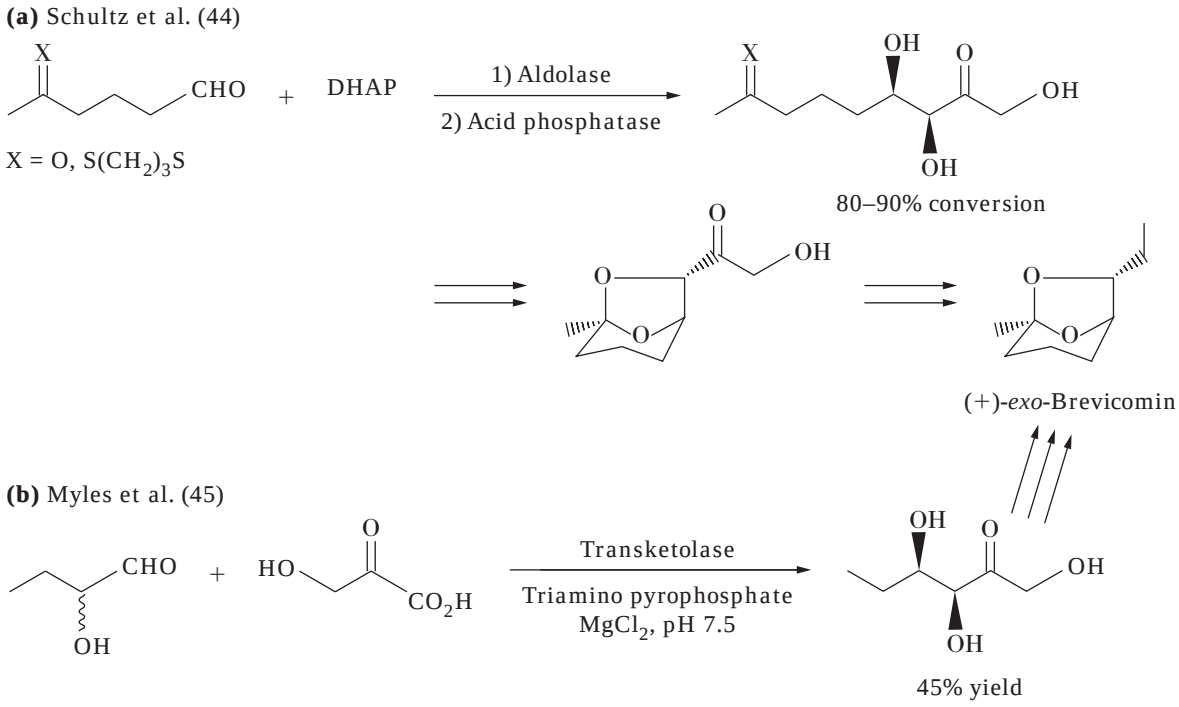
In this article we have introduced about 35 chiral pheromone syntheses using microbial processes. Many microbial processes have been developed in this field, and some of these are already in use. Combination of organic syntheses with microbial processes will provide efficient routes for the enantioselective synthesis of bioactive natural products.

Materials of natural product origin are believed to be more benign to the environment than artificially synthesized compounds. As such, the social demand for reducing artificial insecticide use to protect the environment has promoted development of the use of bioactive natural products, as well as products with structures similar to those of natural products for use as insecticides. It has also promoted development of synthesis technologies of these compounds. Improved technology using microbial reactions allows these chiral compounds to be produced with high chemical and optical purity at a lower price, and clears the way to utilize naturally derived materials as biological insecticides. These advanced insecticides, like synthetic chiral pesticides and pheromones (typical pheromones in this article are listed in the next section), reduce impact on the environment and improve pest control systems.

We did not discuss chiral pyrethroid synthesis in this article, although practical preparations of chiral pyrethroid intermediates using microbial processes have been studied and used in industrial processes (46,47).

PHEROMONE LIST

Ambrosia beetle (*Gnathotrichus sulcatus*) pheromone[sulcatol] (11,38)



Scheme 21. Enzyme-mediated C—C bond formation.

American palm weevil (*Rhynchophorus palmarum*) pheromone (9)

Bark beetle pheromone (*Dryocoetes autographus*, *Dendroctonus frontalis*) [endo-brevicomin] (7,18,27,28,32)

Bark beetle pheromone (*Dendroctonus bremicomis*) [exo-brevicomin] (7,32,44,45)

Brown algae pheromone (*Cutleria mulutifida*, *Chorda tomentosa*) [multifidene] (43)

Brown algae pheromone (*Desmarestia aculeata* and *viridis*) [viridiene] (43)

Carpenter bee (*Xylocopa hirutissima*) pheromone (26,35)

Cupreous chafer beetle (*Anomala cuprea*) pheromone (24)

Dermestid beetle (*Trogoderma glabrum*) pheromone (26)

Elm bark beetle (*Pityophthorus pityographus*) pheromone [pityol] (19,38)

Flat grain beetle (*Cryptolestes ferrugineus*) pheromone [ferrulactone] (39,40)

Flour beetle (*Tribolium confusum* and *castaneum*) pheromone [tribolure] (29)

Gypsy moth pheromone (*Lymantria dispar*) [disparlure] (15,31)

Hessian fly (*Mayetiola destructor*) pheromone (12)

Japanese beetle (*Popillia japonica*) pheromone [japonilure] (22,23)

Oriental hornet (*Vespa orientalis*) pheromone (42)

Ruby tiger moth (*Phragmatobia fuliginosa*) pheromone (15)

Mosquito (*Culex pipients fatigans*) pheromone (25)

Seaweed pheromone [dictyopterene A, C′] (14)

Southern pine beetle (*Dendroctonus frontalis*) pheromone [frontalin] (8,32)

Square-necked

Square-necked grain beetle (*Cathartus quadricollis*) pheromone [quadrilure] (30)

Store grain beetle (*Trogoderma glabrum*) pheromone (36)

Western corn rootworm (*Diabrotica virgifera*) pheromone (41)

BIBLIOGRAPHY

1. K. Mori, in J. ApSimon ed., *The Total Synthesis of Natural Products*, vol. 9, Wiley, New York, 1992, pp. 1–534.
2. J.B. Jones, *Tetrehedron*, **42**, 3351–3403 (1986).
3. K. Faber, *Biotransformations in Organic Chemistry*, 3rd ed., Springer-Verlag, Berlin, 1997.
4. C.-H. Wong and G.M. Whitesides, *Enzymes in Synthetic Organic Chemistry*, Pergamon, Oxford, 1994.
5. K. Drauz and H. Waldmann eds., *Enzyme Catalysis in Organic Synthesis: A Comprehensive Handbook*, VCH, Weinheim, Germany, 1995.
6. K. Mori, *Synlett*, 1097–1109 (1995).
7. K. Matsumoto, N. Suzuki, and H. Ohta, *Tetrahedron Lett.* **31**, 7163–7166 (1990).
8. T. Sugai, H. Kakeya, and H. Ohta, *J. Org. Chem.* **55**, 4643–4647 (1990).
9. K. Mori and K. Ishigami, *Liebigs Ann. Chem.*, 1195–1198 (1992).

10. C.S. Chen, Y. Fujimoto, G. Girdankas, and C.J. Sih, *J. Am. Chem. Soc.* **104**, 7294–7299 (1982).
11. Y. Naoshima, M. Kamezawa, H. Tachibana, Y. Munakata, T. Fujita, K. Kihara, and T. Raku, *J. Chem. Soc. Perkin Trans. 1*, 557–561 (1993).
12. M. Kamezawa, H. Tachibana, T. Ohtani, and Y. Naoshima, *Biosci. Biotechnol. Biochem*. **57**, 1962–1963 (1993).
13. M. Kamezawa, T. Raku, H. Tachibana, T. Ohtani, and Y. Naoshima, *Biosci. Biotechnol. Biochem*. **58**, 598–599 (1994).
14. D. Grandjean, P. Pale, and J. Chuche, *Tetrahedron* **47**, 1215–1230 (1991).
15. J.-L. Brevet and K. Mori, *Synthesis*, 1007–1012 (1992).
16. Y. Naoshima, Y. Munakata, S. Yoshida, and A. Funai, *J. Chem. Soc. Perkin Trans. 1*, 549–553 (1991).
17. B.I. Glanzer, K. Konigsberger, B. Berger, K. Faber, and H. Griengi, *Chem. Phys. Lipids* **54**, 43–48 (1990).
18. G. Pedrocchi-Fantoni and S. Servi, *J. Chem. Soc. Perkin Trans. 1*, 1764–1765 (1991).
19. M. Mishitz, A. Hackinger, I. Francesconi, and K. Faber, *Tetrahedron* **50**, 8661–8664 (1994).
20. A. Klibanov, *Acc. Chem. Res.* **23**, 114–120 (1990).
21. S. Sankaranarayanan, A. Sharma, B.A. Kulkarni, and S. Chattopadhyay, *J. Org. Chem*. **60**, 4251–4254 (1995).
22. T. Sugai, S. Ohsawa, H. Yamada, and H. Ohta, *Synthesis*, 1112–1114 (1990).
23. E. Fukusaki, S. Senda, Y. Nakazono, and T. Omata, *Tetrahedron* **47**, 6223–6230 (1991).
24. E. Fukusaki, S. Senda, Y. Nakazono, and T. Omata, *Biosci. Biotechnol. Biochem*. **56**, 1160–1161 (1992).
25. B. Henkel, A. Kunath, and H. Schick, *Liebigs Ann. Chem*., 921–923 (1995).
26. H.K. Jacobs, B.H. Mueller, and A.S. Gopalan, *Tetrahedron* **41**, 8891–8898 (1992).
27. M.-J. Kim, G.-B. Choi, J.-Y. Kim, and H.-J. Kim, *Tetrahedron Lett.* **36**, 6253–6256 (1995).
28. K. Mori and H. Kiyota, *Liebigs. Ann. Chem*., 989–992 (1992).
29. R. Chenevert and M. Desjardins, *J. Org. Chem*. **61**, 1219–1222 (1996).
30. Y. Naoshima, Y. Munakata, and A. Funai, *Biosci. Biotechnol. Biochem*. **56**, 999–1000 (1992).
31. E. Fukusaki, S. Senda, Y. Nakozono, H. Yuasa, and T. Omata, *J. Ferment. Bioeng.* **73**, 280–283 (1992).
32. J. Ishiyama, M. Toi, C. Ichikawa, S. Oi, and Y. Inoue, *J. Chem. Soc. Jap*., 1162–1166 (1993).
33. V. Prelog, *Pure Appl. Chem*. **9**, 119 (1964).
34. R. MacLeod, H. Prosser, L. Fikentscher, J. Lanyi, and H.S. Mosher, *Biochemistry* **3**, 838 (1964).
35. M. Sato, J. Sakaki, Y. Sugita, T. Nakano, and C. Kaneko, *Tetrahedron Lett.* **31**, 7463–7466 (1990).
36. S.K. Ghosh, S. Chattopadhyay, and V.R. Mamdapur, *Tetrahedron* **47**, 3089–3094 (1991).
37. H. Liu and T. Cohen, *J. Org. Chem*. **60**, 2022–2025 (1995).
38. A. Archelas and R. Furstoss, *Tetrahedron Lett.* **33**, 5241–5242 (1992).
39. E. Keinan, K.K. Seth, R. Lamed, R. Ghirlando, and S.P. Singh, *Biocatalysis* **3**, 57–71 (1990).
40. E. Keinan, S.C. Sinha, and S.P. Singh, *Tetrahedron* **47**, 4631–4638 (1991).
41. E. Keinan, S.C. Sinha, and A. Sinha-Bagchi, *J. Org. Chem*. **57**, 3631–3636 (1992).
42. V. Alphand, A. Archelas, and R. Furstoss, *J. Org. Chem*. **55**, 347–350 (1990).
43. J. Lebreton, V. Alphand, and R. Furstoss, *Tetrahedron Lett.* **37**, 1011–1014 (1996).
44. M. Schultz, H. Waldmann, W. Vogt, and H. Kunz, *Tetrahedron Lett.* **31**, 867–868 (1990).
45. D.C. Myles, P.J. Andrulis, III, and G.M. Whitesides, *Tetrahedron Lett.* **32**, 4835–4838 (1991).
46. H. Danda, T. Nagatomi, A. Maehara, and T. Uemura, *Tetrahedron* **47**, 8701 (1991).
47. S. Masutomo, A. Inoue, K. Kumagai, R. Murai, and S. Mitsuda, *Biosci. Biotechnol. Biochem*. **59**, 720 (1995).

INSULIN, PURIFICATION

INGER MOLLERUP
STEEN WEBER JENSEN
PER LARSEN
OLE SCHOU
LEO SNEL
Novo Nordisk A/S
Bagsvaerd, Denmark

KEY WORDS

Analysis
Characterization
Insulin
Process
Purification

OUTLINE

INTRODUCTION

Until 1982 all insulin preparations available were porcine or bovine insulins prepared by extraction of pancreas. In 1982, Novo launched semisynthetic human insulin, where porcine insulin was enzymatically converted into human insulin; later the same year Eli Lilly and Company launched human insulin of recombinant DNA origin, and

Novo launched human insulin of recombinant origin in 1988. From Eli Lilly and Company two different production methods have been reported. The first method was the expression in *Escherichia coli* of insulin A- and B-chains in two separate systems (1,2), the two chains being isolated and purified separately and later combined chemically, followed by final purification procedures. The second method reported another expression system, also in *E. coli*, where proinsulin was expressed using a tryptophan promotor with a methionine linkage to proinsulin. This linkage was cleaved by cyanogen bromide, followed by folding and formation of the correct disulphide bonds (3,2), and finalized by enzymatic removal of the C-peptide.

This article describes the manufacturing process and analysis of human insulin produced in *Saccharomyces cerevisiae*. A single-chain insulin precursor is expressed and secreted into the medium. After initial recovery, the connecting peptide is cleaved enzymatically in a transpeptidation reaction by the same technology as employed for enzymatic conversion of porcine into human insulin (4). The methods for production of human insulin by recombinant DNA technology have been reviewed by Ladisch (5). The clinical use of human insulin has been reviewed by Heinemann (6), giving an overview of the clinical development resulting from the transition from animal to human insulin and the characteristics of different types of formulation and their impact on blood glucose control.

MANUFACTURING PROCESS

Introduction

In the development of a purification process for insulin, a number of targets must be met. For a product such as insulin to be injected daily, the purity, measured by reverse-phase high-performance liquid chromatography (RP-HPLC), must be very high, and the content of host cell proteins and DNA must be very low. The production takes place on a large scale because the world consumption of insulin is on the order of tons per year. Therefore, the purification processes need to be both robust and performed at a very large scale; the implications for process design were reviewed by Prouty (7). Also the purification process must be adapted to the actual fermentation process.

The insulin production strategies, based on either *E. coli* or yeast technology, include both chemical and enzymatic reactions at different stages of the downstream processing scheme. This leads to the formation of a number of insulin derivatives closely related to insulin and consequently difficult to remove by traditional purification methods. Process-scale RP-HPLC, being a chromatographic separation technique with great separation potential, has been employed for this purpose by Novo Nordisk A/S and other companies (8).

Insulin forms crystals readily (9). This aptitude is utilized throughout the process, both for the purpose of isolating insulin from process streams and for removal of impurities that do not cocrystallize with insulin.

Throughout the process, insulin purity and concentration are monitored by RP-HPLC (see the section "Purity") because this technique has proven to be applicable even in very impure process streams such as the fermentation broth (10).

Expression and Fermentation

Expression and secretion of insulin are achieved by cloning transformants of *S. cerevisiae*. This organism was chosen as host for a number of reasons (11): yeast secretes the insulin precursor into the fermentation broth, making cell disruption unnecessary, the disulfide bonds are established correctly, yeast is an organism easily grown in large fermentors, and yeast is an organism producing no toxic substances.

Investigations of the expression of proinsulin along with a series of insulin precursors where the C-peptide was exchanged by a short peptide (12,13) showed expression of proinsulin in yeast to be too low to be of industrial interest, whereas in *E. coli* the expression is sufficient (1,2). The influence of the composition of the connecting peptide (either amino acid B29 or B30 to A1) on different single-chain insulin precursors containing 1–5 amino acids in the peptide was evaluated. A single-chain precursor with a 3–amino acid connecting peptide between B29 and A1 was chosen for commercial production by Novo Nordisk A/S.

Large-scale fermentations are performed in suitable fermentors (11). The main fermentor is inoculated at first as a fed-batch process; and after 3 days the culture is grown as a continuous fermentation in a simple medium containing yeast extract as its main nitrogen source. The insulin precursor is recovered from the fermentation broth on a continuous basis.

Recovery and Purification

Recovery. The objective of the recovery procedure is to isolate the insulin precursor from the fermentation liquid and obtain a stable intermediate product for further processing. The recovery procedure is outlined in Figure 1, and a view of the recovery plant is shown in Figure 2.

Initially, cell removal is performed by centrifugation of the fermentation broth in disc stack centrifuges. The supernatant is passed through a cation exchange column at low pH, under which conditions insulin is retained on the column. Elution of the insulin precursor takes place by increasing the pH. The eluate is filtered to remove the last traces of yeast particles, whereby the need for containment further downstream is avoided. The product is isolated from the eluate by crystallization, the crystals being removed by centrifugation or filtration. After redissolving the crystals, impurities (mainly host cell proteins) are precipitated at a pH slightly above neutral by addition of ethanol to approximately 60% v/v; at this ethanol concentration insulin is very soluble. The precipitate is removed by filtration, and a second crystallization of insulin results in a product of more than 90% purity, as measured by RP-HPLC, and suited for intermediate storage at $-20\,^{\circ}C$. This intermediate product still contains impurities from the fermentation broth (e.g., yeast proteins and DNA).

Enzymatic Conversion. The use of enzymatic processes in insulin manufacturing was initiated in 1979, when conversion of porcine insulin into human insulin catalyzed by

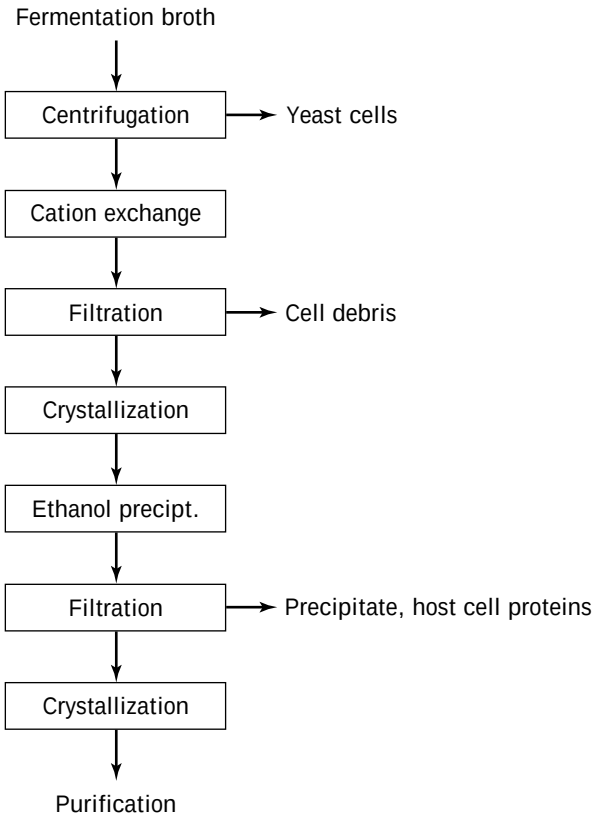


Figure 1. Flow chart of recovery process.

Figure 2. Insulin production plant.

trypsin was demonstrated, resulting in the launch of a semisynthetic human insulin in 1982 (4). Porcine and human insulin differ by only a single amino acid: the C-terminal amino acid (B30) of the B-chain is alanine in porcine insulin and threonine in human insulin. The conversion of porcine into human insulin is achieved by an enzymatic process in which trypsin is used in an organic/aqueous environment to cleave the B-chain after lysine at B29 and to insert threonine instead of alanine (14).

The enzymatic process may be performed as a one-step process, a transpeptidation, or as a two-step process in which the formation of the des-(B30) insulin takes place, and then threonine is coupled to the des-(B30) insulin in the form of a threonine ester. The two-step process is used if a chromatographic purification of the intermediate product des-(B30) insulin is required, but the one-step process is the most simple to perform. In the first industrial human insulin product, launched by Novo in 1982, the conversion was carried out as a one-step process. This enzymatic process utilizing trypsin also proved useful for conversion of single-chain insulin precursors produced by recombinant DNA technology (12–14).

Another enzyme that can be used for the conversion processes is *Achromobacter lyticus* protease (4,13), which compared to trypsin has the benefit of being lysine specific, thus avoiding the risk of cleavage after arginine in position B22; such cleavage may occur when using trypsin. Cleavage sites are shown in Figure 3.

Purification. The purification process of human insulin by recombinant DNA technology from yeast is shown in Figure 4.

After transpeptidation, the crude insulin ester is purified by RP-HPLC at low pH. The purification step is performed on octadecyl-subsituted silica, and the eluent is an ethanol-containing ammonium sulfate buffer at pH 3.0. After chromatography, the human insulin ester is isolated by crystallization. This step removes host cell proteins and trypsin as well as some by-products formed during the transpeptidation reaction (e.g., cyclic des-(B30) insulin). To keep trypsin inactive during the purification step, it is essential that the processing be performed under acid conditions. The purity of the resulting product is typically 90–92%, as measured by RP-HPLC.

After the RP-HPLC process, the human insulin ester is subjected to anion exchange chromatography in a neutral acetate buffer at room temperature. This purification serves the primary purpose of removing any remaining des-(B30) insulin, which is formed as a by-product during the transpeptidation reaction. Des-(B30) insulin is a component with physical properties very similar to those of human insulin. The separation is, therefore, based on the difference in charge between insulin ester and des-(B30) insulin. In addition, this step also reduces a number of other differently charged insulin-related derivatives to a very low level, and host cell proteins are no longer detectable after this step. After collection of the eluate from the column, the product is isolated by crystallization, typically resulting in more than 94% purity, as measured by RP-HPLC.

Hydrolysis. Subsequently, in order to remove the ester group, the human insulin ester is subjected to hydrolysis, resulting in the formation of human insulin. In the choice

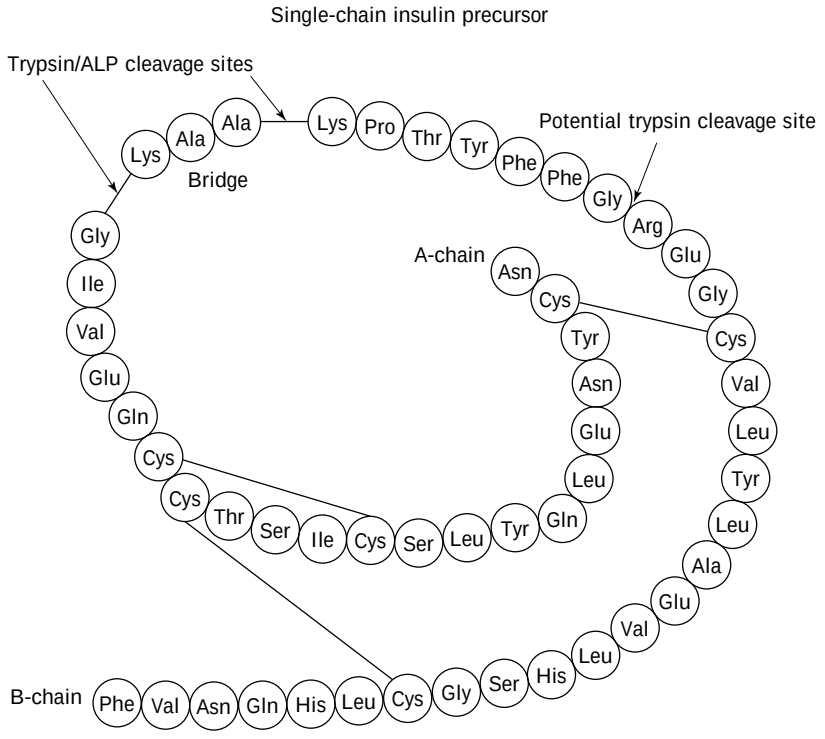


Figure 3. Human insulin precursor. Cleavage sites for trypsin and *Achromobacter lyticus* protease are indicated.

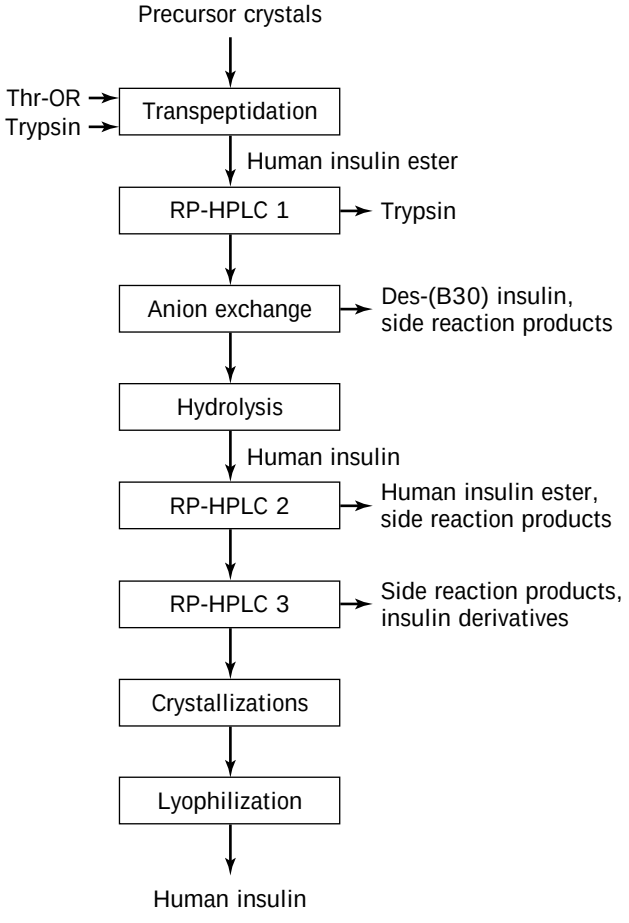


Figure 4. The purification process.

of parameters for the hydrolysis, the relative rates of the desired reaction and the by-product formation must be taken into consideration.

Among the products from the side reactions are cyclic by-products (as single-chain B30-A1 human insulin) and by-products resulting from disulfide-bridge interchange. The parameters finally chosen to balance these different effects resulted in a hydrolysis process performed at high pH and low temperature. It is important to control both pH and temperature very closely because too high pH or temperature may lead to destruction of the disulfide bonds.

At the end of the hydrolysis, the reaction mixture is acidified, and the product is isolated by crystallisation. The purity of this product is typically more than 94%, as measured by RP-HPLC.

Final Purification. After hydrolysis the human insulin must be purified to comply with the specifications for the bulk product. The product contains several impurities such as small amounts of human insulin ester, partially cleaved precursor components, single-chain human insulin, and deamidated insulin. Reverse-phase chromatography at low pH, traditionally used for the purification of peptides, does not suffice to ensure the required purity. This can, however, be achieved by addition of an RP-HPLC step at neutral pH. This purification step (RP-HPLC 2 in Fig. 4) is performed on octadecyl-subsituted silica, and the neutral eluent is based on ethanol and potassium chloride buffered with Tris. After chromatography the human insulin is isolated by crystallization. Primarily this step removes the human insulin ester and reduces the partially cleaved precursor components. The purity of the product is typically more than 99.0%, as measured by RP-HPLC.

The final purification step is also based on reverse-phase chromatography and is performed on octadecyl-subsituted silica at low pH, the eluent being an ethanol-containing ammonium sulfate buffer at pH 3.0 (RP-HPLC 3, Fig. 4). After chromatography the human insulin is isolated by crystallization. Primarily this step removes the single-chain human insulin and reduces the remaining insulin by-products to insignificant levels. The purity of the product is typically more than 99.5%, as measured by RP-HPLC.

Bulk Formulation. After the final chromatographic purification, the product must be conditioned for the subsequent drug product formulation. The target is a product that is stable when stored and with no components unsuitable for the following formulation. Thus residual salts and buffers from the chromatographic steps must be reduced.

During bulk formulation the product must be treated carefully to eliminate decomposition. The main risk factors when handling the product after the final chromatographic purification are formation of aggregates (fibrillation) and deamidation of the insulin molecule (15). Aggregation can be the result of stressing the molecule (i.e., agitation), whereas deamidation primarily results from prolonged exposure to acid conditions (formation of A21 desamido insulin).

Crystallization is used for the final treatment of both animal and human insulins, followed by freeze-drying. Crystallizations are carried out in the presence of zinc; after the last crystallization the crystals are washed to reduce any excess content of zinc. To a wide extent the crystallization processes of insulin bulk production are based upon the development work done in the 1950s (9).

For special drug types and for other purposes where a zinc-free product is needed, alternatives to the crystallization processes can be used. These may include isoelectric precipitation, gel filtration chromatography, or ultra- and diafiltration.

The freeze-drying procedure results in a bulk product with a water content below 10.0% (loss on drying); it is very stable when stored frozen. Storage as a bulk solution is possible, but in the case of insulin it is not feasible due to the demand for very large bulk quantities.

ANALYSIS OF RECOMBINANT HUMAN INSULIN

Introduction

Human insulin is a well defined and well characterized protein, and recombinant human insulin is chemically, physically, and biologically equivalent to both pancreatic human insulin and semisynthetic human insulin (4,16). As purification methods continue to improve, so does the analytical methodology to demonstrate purity and integrity throughout the manufacturing process. The requirement to show process and batch consistency demands an array of analytical techniques to show any presence of adventitious agents, host cell contaminants, process agents, and insulin-related by-products. Moreover, changes of the protein structure during the manufacturing process may influence the biological activity of the insulin product.

Identification

The primary structure of human insulin can be verified by amino acid composition analysis, amino acid sequence analysis, determination of electrophoretic mobility, measurement of retention on RP-HPLC, measurement of molecular mass, and peptide fingerprinting. The first three methods are all well established but of limited specificity.

The need to distinguish between the various insulin products during the manufacturing process was the driving force in the development of RP-HPLC, the separation principle being based on hydrophobic differences (17,18). The RP-HPLC method of choice is based on C18 stationary phases and mobile phases containing acetonitrile or methanol buffers at pH between 2 and 8. The high selectivity and resolving power allow the separation of various insulin species, precursors, and insulin derivatives differing by only one amino acid (Fig. 5).

Determination of molecular mass by mass spectrometry is being established in quality control laboratories as a precise method of identity. Equipped with an electrospray ionization source, the single- or triple-quadropole instruments can determine the mass of insulin within 0.5 Da and clearly distinguish between deamidated insulin and human insulin, which have a difference of only one mass unit.

Fingerprinting or peptide mapping is an excellent tool to confirm the correct amino acid sequence as well as the proper formation of disulfide bonds. The peptide map is generated by fragmentation by an enzyme (preferably *Staphylococcus aureus* V8 protease) cleaving at a specific residue (e.g., glutamic acid), thus giving four fragments easily separated by RP-HPLC. The peptide map of a sample is then compared to the map of an insulin reference standard. The patterns should be identical (Fig. 6). Each fragment peak on the reference map must be characterized at least once by amino acid sequence analysis and mass spectrometry.

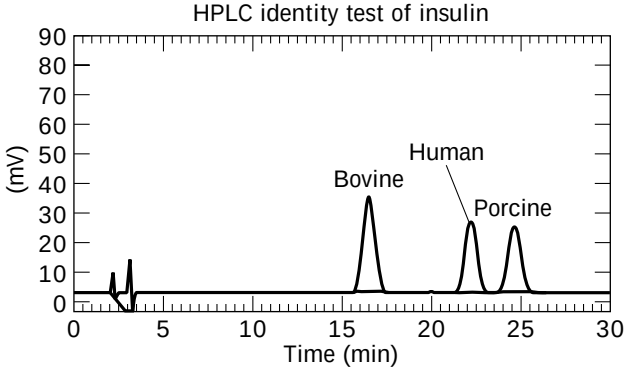


Figure 5. RP-HPLC chromatogram showing the separation of different insulin species differing by only one amino acid. Column: Lichrosorb RP-18, 5 μm; solvent A: 0.2 M sodium sulfate, 0.04 M phosphoric acid in 10% acetonitrile, pH adjusted to 2.5 with ethanolamine; solvent B: 50% acetonitrile; flow rate: 1.0 mL/min; temperature: 40 °C; 50 μg of sample was injected and eluted isocratically at 49% (solvent A) 51% (solvent B).

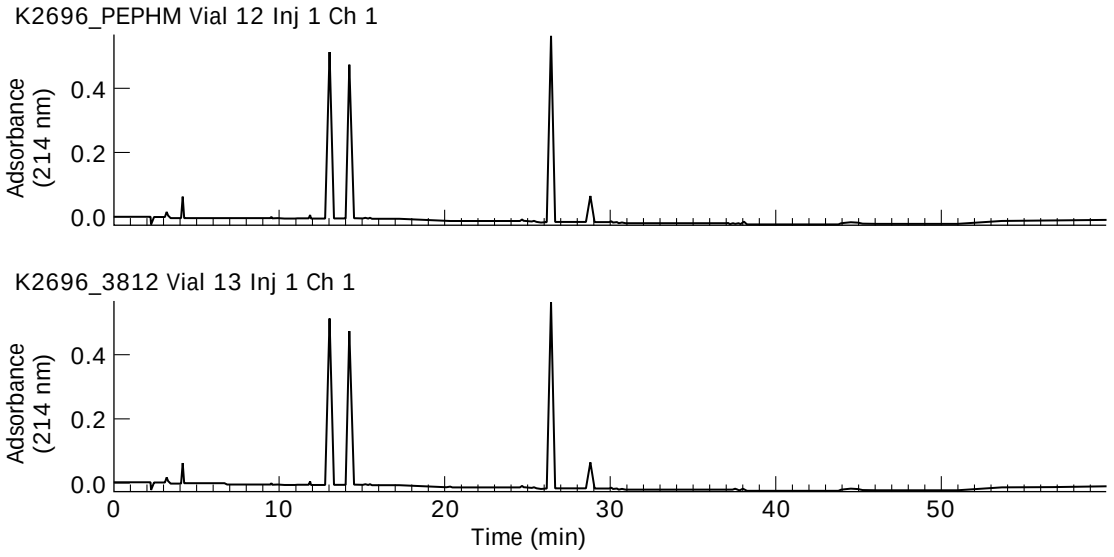


Figure 6. Peptide map of insulin sample (bottom panel) compared to a human insulin reference standard (top panel). Column: Lichorsorb RP-18, 5 μm; solvent A: 0.2 M sodium sulfate, 0.04 M phosphoric acid in 10% acetonitrile, pH adjusted to 2.5 with ethanolamine; solvent B: 50% acetonitrile; flow rate: 1.0 mL/min; temperature: 40 °C; 25 μg of sample was cleared prior to injection and eluted with a gradient.

The secondary and tertiary structures of human insulin can be assessed by nuclear magnetic resonance, circular dichroism, X-ray crystallography, and measurement of biological activity. Among these methods, only bioactivity is measured on a routine basis; the other tests are performed once to fully characterize human insulin.

Purity

Biological products in general require a special assessment to demonstrate purity, safety, and efficacy. The problems associated with this are the following:

- The need to establish batch-to-batch reproducibility because of the biological nature of early stages of the manufacturing process
- The need for chemical and biological tests to assess purity and potency where these can be affected by variation in starting materials, manufacturing process, and product stability
- The need to avoid adventitious agents arising from the starting materials and/or the downstream processes; these can be microbiological or associated with nonmicrobiological contaminants

Routine quality control testing should be designed to take into account the specific nature of the product. Therefore, specifications and methods of analysis are developed for each product. It is generally acknowledged that the final product cannot be completely characterized, and that in-process controls are necessary to prove its consistent quality.

Microbiological Impurities. Each manufacturing step is performed under conditions designed to keep the bioburden low and avoid endotoxins in the recombinant human insulin. Whereas there is an absolute requirement for sterility in the formulated product, a result of below 100 viable counts/g is acceptable for the drug substance. The absence of endotoxins is confirmed by either the Limulus amoebocyt lysate (LAL) test or by a biological test in rabbits. Process validation documents complete removal of potential adventitious virus or other agent.

Chemical Impurities

Contaminants. Host cell proteins, DNA, and process-related chemicals are all impurities that need to be controlled, preferably by validation studies or by in-process/specification testing. Sensitive assays should be developed to detect any residual levels of yeast proteins, trypsin used for the transpeptidation, or DNA to the parts per million limit. For DNA detection the Threshold® DNA assay is much preferred to former hybridization assays, mainly due to its higher sensitivity but also to its robustness. The host cell contaminant assay is based on antibodies generated against the yeast proteins from the production strain lacking the insulin gene. Thus antibodies are raised against proteins ranging in molecular weights from 10 kDa to 100 kDa and thus also covering proteins the size of insulin. Table 1 shows the reduction of DNA and host cell proteins during the purification process.

Absence of trypsin can be shown by a trypsin assay (ELISA or activity), and extensive step reduction studies have proven a more than 20 log reduction of the trypsin level.

Related Impurities. Insulin-related impurities are defined as derivatives generated by deamidation, dimerization, polymerization, peptide cleavage, and disulfide exchange (15), or as intermediate products from the

Table 1. Purity of Human Insulin or Human Insulin Precursor During Downstream Processing as Indicated by RP-HPLC Analysis (%) and Test for Residual Levels of Host Cell Proteins and DNA

	Purity tests		
Process step	HPLC (%)	Yeast proteins (ppm)	DNA (ppm)
1. Yeast fermentation broth			
2. Isolation of precursor	~20	>10,000	1,000
3. Purification of precursor	~90	1,000	<3
4. Conversion to Insulin ester	~85		
5. Purification	~94	<1	<0.5
6. Hydrolysis to insulin	~94		
7. Preparative HPLC	>99.0	<1	<0.5
8. Monocomponent Insulin	>99.0	<1	<0.5

manufacturing process (e.g., insulin precursor, insulin ester).

During human insulin manufacture the purity at different process steps is assessed by RP-HPLC. RP-HPLC on alkylsilane supports has become the method of choice due to the hydrophobic differences between insulin and the insulin-related impurities (10,18,19). Separation on ion exchange HPLC resin has also been reported (18) and compared with RP-HPLC, but the latter is preferred for in-process as well as specification analyses. Figure 7 shows the purity analyses on four different purification process steps.

Deamidation. Human insulin contains six potential deamidation sites: the A-chain residues A5(Gln), A15(Gln), A18(Gln), and A21(Asn), and the B-chain residues B3(Asn) and B4(Gln). The deamidation in acid and neutral environment has been described (15), and it has been shown that the desamido insulins retain almost full biological activity. In order to separate all deamidated forms of insulin, complementary methods might be used on uncoated capillaries (e.g., RP-HPLC and HPCE [20]). The elution order of the various desamido forms can hardly be predicted because it depends on the selectivity of the stationary phase.

Aggregation. Human insulin forms covalent as well as noncovalent dimers and polymers (15). The association to dimers and hexamers (noncovalent) is reversible, and at certain conditions the insulin monomer linearizes and forms fibrils. These fibrils are unwanted reaction products of low solubility and can be avoided by proper manufacturing precautions (15). The covalent insulin dimers and polymers form during handling of insulin solutions and during storage at elevated temperatures.

The method used to quantitate the extent of dimerization/polymerization is gel permeation chromatography. The column support is a silica-based diol, and the mobile phase should contain a high concentration of acetic acid in order to maintain the dissociation state of the analyte.

Assay

The potency of human insulin has been determined by biological assays such as the mouse convulsion assay and the rabbit blood glucose test. RP-HPLC assay data compared with bioassays seemed to justify a chemical assay to express the potency, and such methods are now described in the pharmacopoeias. The resolution of the RP-HPLC assay allows a quantitation of the human insulin content in the presence of even substancial amounts of related substances. Furthermore, the RP-HPLC assay offers superior precision with a standard deviation of only about 1%. This is particularly valuable for the determination of the human insulin content in a sample and for monitoring stability studies on bulk drug substance and drug product.

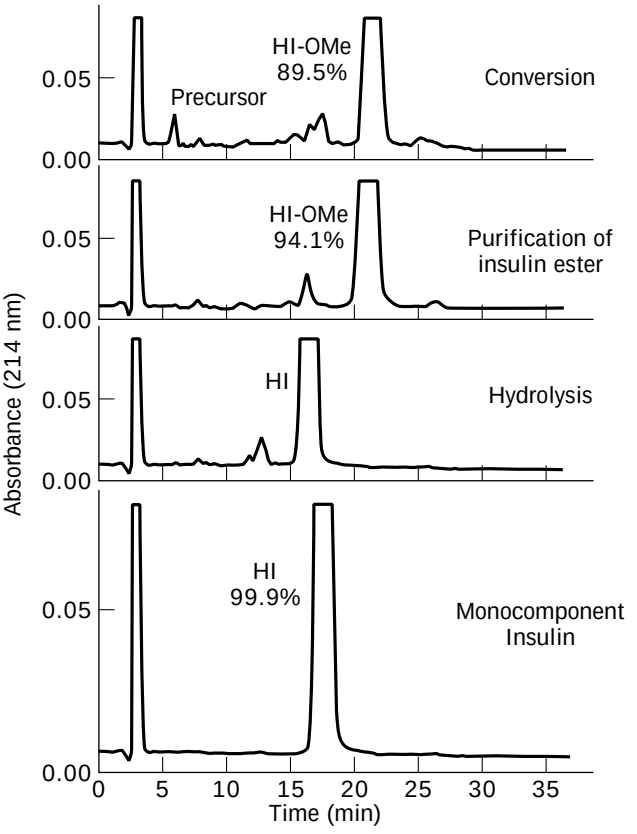


Figure 7. Purification steps after conversion of insulin precursor: RP-HPLC chromatogram of in-process samples from the purification process. Column: Lichrosorb RP-18, 5 μm, 4 $\times$ 250 mm; solvent A: 0.2 M sodium sulfate, 0.04 M phosphoric acid in 10% acetonitrile, pH adjusted to 2.5 with ethanolamine; solvent B: 50% acetonitrile; flow rate: 1.0 mL/min; temperature: 40 °C; 50 μg of sample was injected and eluted isocratically at 49% (solution A) 51% (solution B).

STABILITY

Human insulin manufactured by the method just described is very stable when stored frozen, protected from light. Hardly any change of quality can be detected even after several years of storage. However, when stored at between 2 and 8 °C, which is the prescribed storage temperature, insulin preparations will gradually degrade while forming neutral deamidation products and dimers (15). Therefore, the shelf life of insulin preparations is limited to 2–3 years.

BIBLIOGRAPHY

1. R.E. Chance, J.A. Hoffmann, E.P. Kroeff, M.G. Johnson, E.W. Schirmer, W.W. Bromer, M.J. Ross, and R. Wetzel, in D.H. Rich and E. Gross eds., *Peptides: Synthesis, Structure, Function*, Pierce Chemical Company, 1981, pp. 721–728.
2. B.H. Frank and R.E. Chance, *Münch. med. Wschr.* **125**, 14–20 (1983).
3. B.H. Frank, J.M. Pettee, R.E. Zimmerman, and P.J. Burck, in D.H. Rich and E. Gross eds., *Peptides: Synthesis, Structure, Function*, Pierce Chemical Company, 1981, pp. 729–738.
4. J. Markussen, U. Damgaard, K.H. Jørgensen, E. Rasmussen, L. Snel, L. Thim, and H.O. Voigt, in *Hormone Drugs. Proceedings of the FDA-USP Workshop on Drug and Reference Standards for Insulins, Somatropins, and Thyroid-Axis Hormones*, 1982, pp. 116–126.
5. M.R. Ladisch and K.L. Kohlmann, *Biotechnol. Prog.* **8**, 469–478 (1992).
6. L. Heinemann and B. Richter, *Diabetes Care* **16**, 90–100 (1993).
7. W.F. Prouty, in Y.H. Chiu and J.L. Gueriguian eds., *Drug Biotechnology Regulation*, Marcel Dekker, New York, 1991, pp. 221–262.
8. E.P. Kroeff, R.A. Owens, E.L. Campbell, R.D. Johnson, and H.I. Marks, *J. Chromatogr.* **461**, 45–61 (1989).
9. J. Schlichtkrull, *Insulin Crystals: Chemical and Biological Studies on Insulin Crystals and Insulin Zinc Suspensions*, Munksgaard, Copenhagen, 1958.
10. L. Snel, U. Damgaard, and I. Mollerup, *Chromatographia* **24**, 329–332 (1987).
11. I.V. Diers, E. Rasmussen, P.H. Larsen, and I.-L. Kjærsig, in Y.H. Chiu and J.L. Gueriguian eds., *Drug Biotechnology Regulation*, Marcel Dekker, New York, 1991, pp. 166–176.
12. L. Thim, L. Snel, K. Norris, and M.T. Hansen, in C.L. Hershberger, S.W. Queener, and G. Hegeman eds., *Genetics and Molecular Biology of Industrial Microorganisms*, American Society for Microbiology, Washington D.C., 1989, pp. 322–328.
13. T. Christensen, H. Dalbøge, and L. Snel, in Y.H. Chiu and J.L. Gueriguian eds., *Drug Biotechnology Regulation*, Marcel Dekker, New York, 1991, pp. 206–221.
14. J. Markussen, *Human Insulin by Trypsic Transpeptidation of Porcine Insulin and Biosynthetic Precursors*, MTP Press, Lancaster, U.K., 1987.
15. J. Brange, *Stability of Insulin: Studies on the Physical and Chemical Stability of Insulin in Pharmaceutical Formulation*, Kluwer Academic, Boston, 1994.
16. R.E. Chance, E.P. Kroeff, and J.A. Hoffmann, in *Insulins, Growth Hormone and Recombinant DNA Technology*, Raven, New York, 1981, pp. 71–85.
17. U. Damgaard and J. Markussen, *Horm. Metab. Res.* **11**, 580–581 (1979).
18. B.S. Welinder and S. Linde, in W.S. Hancock ed., *CRC Handbook of HPLC for the Separation of Amino Acids, Peptides and Proteins*, CRC Press, 1984, pp. 357–362.
19. A. McLead and S.P. Wood, *J. Chromatogr.* **285**, 319–331 (1984).
20. G. Mandrup, *J. Chromatogr.* **604**, 267–281 (1992).

L-ISOLEUCINE

MASAYUKI INUI
MASATO TERASAWA
HIDEAKI YUKAWA
Mitsubishi Chemical Corporation
Ibaraki, Japan

KEY WORDS

ilv Genes
L-Isoleucine
LCR process
L-Leucine
L-Valine

OUTLINE

INTRODUCTION

L-Isoleucine, L-valine, and L-leucine are termed branched-chain amino acids, because the chemical structures and physical characteristics of these amino acids are similar. L-Isoleucine belongs to the aspartic acid group of amino acids, which also contains L-threonine, L-lysine, and L-methionine, because four of six carbon atoms in these amino acids are derived from aspartic acid. On the other hand, L-valine and L-leucine belong to pyruvate group of amino acids, which also include L-alanine, because all carbon atoms in these two amino acids are derived from pyruvate. However, these three amino acids are synthesized by parallel biosynthetic pathways.

L-Isoleucine is one of the essential amino acids and therefore of importance for the pharmaceutical and feed industries. Particularly, the primary use of this amino acid in pharmaceuticals is as part of injectable solutions of amino acids. Because L-isoleucine possesses two asymmetric carbon atoms, there exist four optical isomers, L-, D-, L-allo-, and D-alloisoleucine. Therefore, it is difficult to produce and separate pure L-isoleucine at low cost by chemical synthetic methods, and research has been initiated to produce L-isoleucine by fermentative production methods. Although isolation of microorganisms that excrete amino acids in nature has been attempted and a number of strains that can excrete L-valine and L-leucine have been found, none have been detected that can excrete L-isoleucine.

In the early 1950s, Umbarger et al. studied the isoleucine biosynthetic pathway and showed that the precursor of L-isoleucine could be not only L-threonine, but also D-threonine and α-ketobutyric acid. Based on these results, the production of L-isoleucine using D-threonine and α-

ketobutyric acid as a precursor by fermentation was reported in 1962 (1).

In the late 1960s, the genetic determinants of feedback inhibition of amino acid biosynthesis were elucidated, and it was subsequently proved that feedback inhibition of amino acid biosynthesis was released in mutants resistant to the corresponding amino acid analogue (2). The cumulative result of these studies is that L-isoleucine is now produced on an industrial scale by fermentation.

This article will review the biosynthetic pathway and related enzymes and genes for L-isoleucine synthesis and will also describe microbial production of L-isoleucine.

BIOSYNTHETIC PATHWAY OF ISOLEUCINE AND ITS REGULATION

Elucidation of the pathways to L-isoleucine and L-valine began with the analysis of auxotrophic mutants of *Escherichia coli* and *Salmonella typhimurium*. These mutants usually require both L-isoleucine and L-valine for growth, less frequently only L-isoleucine, and very rarely only L-valine. The phenomenon was finally understood when it was found that the two amino acids are produced by a series of reactions catalyzed by the same enzymes (Fig. 1). At the first step of the pathway to L-isoleucine, α-ketobutyrate reacts with active acetoaldehyde derived from pyruvate by the action of acetohydroxy acid synthase (AHAS), whereas in case of L-valine, two molecules of pyruvate are condensed. The next three steps are catalyzed in parallel by acetohydroxy acid isomeroreductase (AHAIR), dihydroxy acid dehydrase (DH), and transaminase B (TrB). Threonine deaminase (TD) catalyses the deamination of threonine to yield α-ketobutyrate. L-Leucine branches from the L-valine route at the α-ketoisovalerate step.

As found in several biosynthetic pathways, feedback inhibition and repression function in the regulation of the isoleucine-valine biosynthesis pathway (Table 1).

The activity of TD is strongly inhibited by L-isoleucine in most microorganisms. Because it is a competitive inhibitor with L-threonine, the degree of inhibition depends on the concentration of L-threonine. However, it was found that in *Rhodospirillum rubrum*, the inhibition by L-isoleucine of TD is very weak. Although it was reported that in *Pseudomonas aeruginosa* the TD gene is constitutively expressed, the TD gene is generally repressed. In *Neurospora crassa*, TD is repressed by only L-isoleucine, whereas in other organisms, TD is repressed by the presence of L-isoleucine, L-valine, or L-leucine as a multivalent repression (Table 1).

In *E. coli* and *S. typhimurium*, three different isozymes have been demonstrated for AHAS. The activities of two, AHAS I and AHAS III, are inhibited by L-valine. Inhibition by L-isoleucine of AHAS I and AHAS III are not strong compared with L-valine inhibition, and L-leucine does not inhibit them. AHAS I genes (*ilvBN*) are repressed by the presence of both L-valine and L-leucine as a multivalent repression, and AHAS III genes (*ilvIH*) are repressed by L-leucine. Although another isozyme, AHAS II, does not show sensitivity to L-valine, AHAS II genes (*ilvGM*) are repressed by the presence of L-isoleucine, L-valine, and L-leucine. In the strain *E. coli* K-12, lack of AHAS II activity is caused by a natural frameshift mutation in the *ilvG* gene for AHAS II. On the other hand, an active AHAS II (*ilvGM* product) is present in *E. coli* B and W and *S. typhimurium*, and these strains are therefore able to grow in a minimal medium containing L-valine.

In *Bacillus subtilis*, *P. aeruginosa*, *Saccharomyces cerevisiae*, and *N. crassa*, some mutants that require L-isoleucine and L-valine show no detectable AHAS activity, suggesting that only one AHAS enzyme may exist in these organisms. However, no genetic evidence concerning the presence or absence of AHAS isozymes has been presented to date. In *Brevibacterium flavum*, gene disruption studies show that only one AHAS gene exists in this organism (3). AHAS enzymes from these organisms are also inhibited by L-valine.

In addition, the biosynthetic pathway leading to the branched-chain amino acids in plants is of interest for herbicide development. AHAS is the target of a variety of herbicides (4,5), and inhibition of AHAIR also was demonstrated to result in some herbicidal effects (6). Furthermore, it is noteworthy that in *Rhizobium meliloti*, the AHAIR gene is required for nodulation (7).

Thus, because (*1*) the common enzymes, which show end-product inhibition and repression, function in the parallel branched-chain amino acids biosynthetic pathway; (*2*) the pathway is located downstream of biosynthetic flows of essential amino acids such as L-threonine, L-lysine, and L-methionine; and (*3*) the central metabolite for production of energy and glucogenesis is linked to this pathway, it presumed that the regulation of this pathway is complicated. Although there are a number of reports analyzing the isoleucine-valine biosynthetic pathway, it is still unclear in *E. coli* and even less so in other organisms.

BIOSYNTHETIC GENES OF ISOLEUCINE

The nucleotide sequences of the *ilv* genes from several organisms have been reported. AHAS enzyme, which is the first enzyme in the parallel isoleucine-valine biosynthetic pathway, is composed of a large and a small subunit in an $\alpha 2\beta 2$ structure in several organisms. It has been suggested that for all isozymes, the small subunit is important for maximal catalytic activity of the large subunit and contributes importantly to the regulation of the activity of the two valine-resistant isozymes in *E. coli*. Three sets of genes for AHAS isozymes have been isolated and sequenced from *E. coli*. These are the *ilvBN* genes for AHAS I, the *ilvGM* genes for AHAS II, and the *ilvIH* genes for AHAS III (8–10). AHAS (*ilvBN*) genes from *Bacillus subtilis* (11), *Brevibacterium flavum* (*ilvLS* genes) (3), *Caulobacter crescentus* (12), *Corynebacterium glutamicum* (13), *Lactococcus lactis* (14), *Mycobacterium avium* (15), and *Streptomyces avermitilis* (16) have been isolated, and all contain genes for a large and a small subunit. However, the presence of AHAS isozymes in these organisms has not been reported. In contrast, the large subunit genes of the AHAS isozymes of the cyanobacterium *Spirulina plantensis* (17) are not adjacent to their corresponding small subunit genes. This is similar to the gene organization observed in *Arabidopsis*

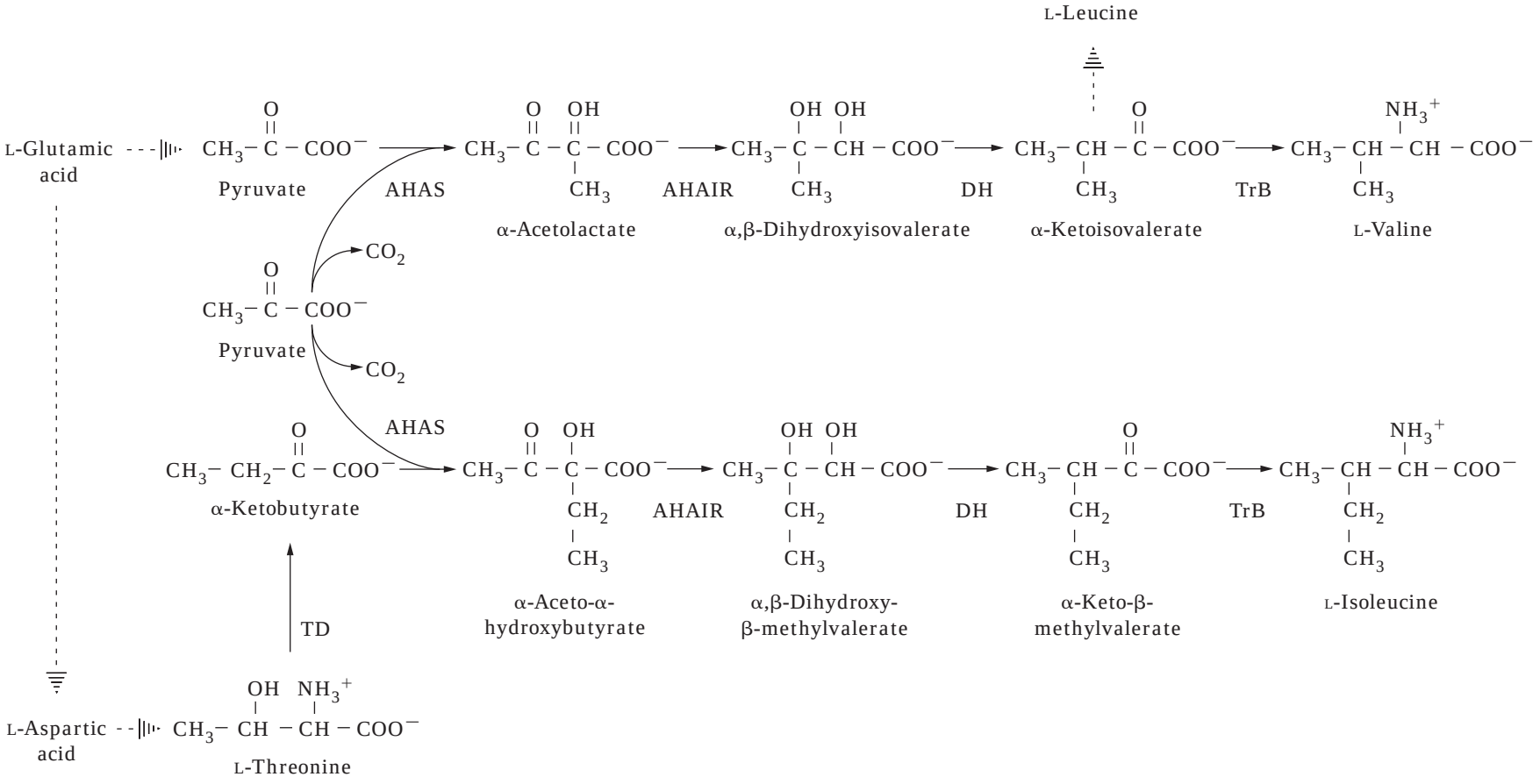


Figure 1. Biosynthesis of isoleucine and valine. TD = threonine deaminase; AHAS = acetohydroxy acid synthases; AHAIR = acetohydroxy acid isomeroreductase; DH = dihydroxy acid dehydrase; TrB = transaminase B.

Table 1. Feedback Inhibition and Repression of Key Enzymes for Branched-Chain Amino Acids Biosynthesis Pathway

	Threonine deaminase		Acetohydroxy acid synthase	
Strains	Inhibition	Repression	Inhibition	Repression
E. coli	ile	ile + val + leu	I; val (ile)	I; val + leu
S. typhimurium			II; insensitive	II; ile + val + leu
			III; val (ile)	III; leu
B. subtilis	ile	ile + leu + (val)	val (ile)	val + leu
P. aeruginosa	ile	constitutive	val (ile, leu)	ile + val + (leu)
S. cerevisiae	ile	ile + val + (leu)	val	ile + val + leu
N. crassa	ile	ile	val (ile)	val

Note: ile = L-isoleucine; val = L-valine; leu = L-leucine.
[a]I, II, and III show isozymes of AHAS.

thaliana (18), *Nicotina tabacum* (5), *Brassica napus* (19), yeasts (20), and *Synechococcus* PCC 7942 (21).

AHAIR (*ilvC*) genes, whose enzyme catalyzes the second reaction in the parallel isoleucine-valine biosynthetic pathway, have been isolated from *Br. flavum* (22), *Synechocystis* sp. strain PCC 6803 (23), *B. subtilis* (11), *L. lactis* (14), *Rhizobium meliloti* (7), *E. coli* (24), *S. cerevisiae* (*ILV5* gene) (25), *C. glutamicum* (13), *St. avermitilis* (15), *A. thaliana* (26), *M. avium* (15), and spinach *Spinacia oleracea* L. (27), respectively. In *E. coli* and *S. typhimurium*, it was demonstrated that expression of the *ilvC* genes is induced by the substrates of the AHAIR enzyme, 2-aceto-2-hydroxy-butyrate and 2-acetolactate (28). This substrate induction of *ilvC* expression is mediated by a positive activator encoded by the *ilvY* gene (29). The *ilvC* gene is linked to *ilvY*, but its transcription is divergent. In *R. meliloti*, nucleotide sequence analysis of the *ilvC* region revealed the presence of a short open reading frame, ORF2, adjacent to the *ilvC* coding region but transcribed in the opposite direction. It was proposed that this ORF2 coded for a putative regulatory protein, but this protein has no homology with the *E. coli ilvY* gene product (7). However, in other bacteria such as *Br. flavum, Synechocystis* sp., *B. subtilis, L. lactis, S. cerevisiae*, and spinach, no open reading frames homologous to the *E. coli ilvY* gene or to the *R. meliloti* ORF2 were detected downstream of *ilvC*.

DH (*ilvD*) genes were cloned from *E. coli* (30), *L. lactis* (14), *B. subtilis* (31), and *Ca. crescentus* (32), respectively. In *Ca. crescentus*, the *ilvR* gene, which is transcribed divergently from the *ilvD* gene and which positively regulates the expression of the *ilvD* gene, was found upstream of the *ilvD* region (32). Amino acid sequence comparison revealed that the IlvR protein is a member of the LysR family of transcriptional regulators. So far, in other organisms, such a regulator gene has not been reported.

Although the TrB (*ilvE*) gene was isolated from *E. coli* (30), there has been no report of the corresponding gene from other organisms.

Interestingly, the *ilv* genes in several bacteria are organized in clusters scattered throughout the chromosome. In *E. coli*, the *ilv* genes are organized into three clusters, namely *ilvGMEDAYC, ilvIH*, and *ilvBN* (8,10,24,30). In *L. lactis*, the *ilvDBNCA* genes have been shown to cluster in one region (14). In *B. subtilis* (23), *Br. flavum* (3,22), *C. gultamicum* (13), *M. avium* (15), and *St. avermitilis* (16), the *ilvBNC* genes are also found in one cluster.

PRODUCTION OF L-ISOLEUCINE

There have been a number of reports regarding the production of L-isoleucine by fermentation. As described earlier, it was found that feedback inhibition in the isoleucine-valine biosynthesis could be released by isolation of a mutant resistant to the corresponding amino acid analogue. Attempts to use the analogue-resistant mutant for production of L-isoleucine were then started.

Kisumi et al. reported in *Serratia marcescens* (strain GIHVLAr2795) that a mutant resistant to both isoleucine hydroxamate and α-aminobutyric acid produced 12 g/L of L-isoleucine in 72 h in a medium containing glucose as a carbon source by fermentation (33). Isoleucine hydroxamate was described as an isoleucine antagonist in *Serratia marcescens* (34), and α-aminobutyric acid was known as a valine analogue in *E. coli* (35). This strain had increased activities of TD, AHAS, aspartokinase, and homoserine-dehydrogenase compared with the parental strain and lacked both feedback inhibition and repression of TD.

Komatsubara et al. reported construction of a threonine-producing strain by the transduction of an isoleucine-producing strain of *Serratia marcescens* (36). First, strain GIHVLr6426 was isolated as an isoleucine-hydroxamate-resistant mutant from wild-type *Serratia marcescens*. This mutant had high activities of TD and AHAS enzymes, both of which were insensitive to feedback inhibition. Genetic analysis indicated that strain GIHVLr6426 carried two regulatory mutations located in the *ilv* region. One, designated *ilvA2*(Fr), leads to desensitization of the TD enzyme, and another, designated *ihr-1*, leads to constitutive synthesis of TD and AHAS enzymes (36). Previously, a threonine-producing strain of *Serratia marcescens*, T-693, which carried four regulatory mutations for threonine biosynthesis and had an increased activity of AHAS, was isolated (37). To construct an isoleucine-producing strain, the strain T-693 was used as a recipient in transduction. Strain T-803 was constructed by transferring the two mutations, *ilvA2*(Fr) and *ihr-1*, carried by strain GIHVLr6426 into strain T-693. The resulting strain had the six regulatory mutations, which are *lysCl*(Fr), *thr*A_1*l*(Fr), *thr*A_2*1*(Fr), *hnr-1*, *thr-1*, and *ilvA2*(Fr), for threonine and isoleucine biosynthesis and produced about 25 g/L of L-isoleucine in 96 h in a medium containing sucrose and urea by fermentation (36).

Shiio et al. reported L-isoleucine production from glucose by α-amino-β-hydroxyvaleric acid (AHV) resistant mutants of *Br. flavum* (38). Previously, it had been demonstrated that in *Br. flavum*, growth inhibition by AHV, an analogue of threonine, was reversed not only by L-threonine but also partially by L-isoleucine, and that threonine-producing mutants, which were isolated as the analogue-resistant mutants, accumulated a small amount of L-isoleucine (39). It was also reported that the AHV resistant strain ARI-129 of *Br. flavum* produced 11 g/L of L-isoleucine in 2 days by fermentation (38). Although homoserine dehydrogenase, which is one of threonine biosynthetic enzymes, from the AHV resistant strain ARI-129 was insensitive to feedback inhibition by L-isoleucine and L-threonine, no difference was observed in the TD enzyme between the parent strain and the AHV-resistant strain ARI-129. In general, because feedback inhibition of TD enzyme by L-isoleucine is competitive with respect to L-threonine, the degree of the inhibition by L-isoleucine decreases as the concentration of L-threonine increases (40). Therefore, the authors presumed that the AHV-resistant mutant ARI-129 also contained a mutation affecting permeability to L-threonine, and L-isoleucine was overproduced because of an increase in the intracellular level of L-threonine, even though the TD enzyme was not released from feedback inhibition by L-isoleucine. But no experimental evidence relative to the mechanism described above has yet been presented. Furthermore, *O*-methyl-L-threonine (OMT), an analogue of isoleucine, inhibits the growth of the ARI-129 strain and this inhibition can be reversed by addition of L-isoleucine. Hence, OMT-resistant mutants derived from the strain ARI-129 could be isolated (38). The OMT-resistant mutant, strain AORI-126, produced 14.5 g/L of L-isoleucine by fermentation (38). Specific activity of TD enzyme from the AORI-126 strain increased about twofold higher than that of the parent strain and the ARI-129 strain, whereas the degree of inhibition of the enzyme by L-isoleucine was the same among three strains.

Ikeda et al. reported L-isoleucine production from glucose by ethionine-resistant mutants of *Br. flavum* (41). Two types of L-isoleucine-producing mutants, strains 168 and 14083, were isolated from L-threonine producers by their enhanced resistance to ethionine. The L-isoleucine productivity of strains 168 and 14083 were 9.9 g/L and 11.3 g/L in 2 days by fermentation, respectively (41). The 168 mutant was shown to increase the specific activity of the TD enzyme, whereas the effect on 14083 was based on the insensitivity of the TD enzyme to feedback inhibition by L-isoleucine. However, because ethionine is an analogue of methionine, the reason that some ethionine-resistant mutants have the insensitivity of the TD enzyme is mysterious. Moreover, Ikeda et al. showed that, when the optimum fermentation conditions were controlled, including the concentration of acetate and ammonium, the L-isoleucine productivity of the 14083 strain was 33.5 g/L after 77 h of cultivation (42).

Eggeling et al. reported that *Corynebaterium glutamicum* ATCC 13032 produces about 13 g/L of L-isoleucine from a synthetic precursor of 200-mM α-ketobutyric acid in 119 h by fermentation and also produces up to about 19 g/L of L-isoleucine in 100 h in fed-batch cultures (43).

Kase et al. reported that RAM-83, a mutant derived from a threonine-producing strain of *Corynebacterium glutamicum*, when successively endowed with resistivity to such substances as thiaisoleucine, ethionine, 4-azaleucine, and α-aminobutyric acid, produces 9.7 g/L of L-isoleucine with a medium containing 10% molasses as sugar by fermentation (44). Although the thioisoleucine and α-aminobutyric-acid resistant mutations resulted in a significant increase in L-isoleucine productivity, it was not shown which enzyme was released from feedback inhibition by L-isoleucine (44).

Thus, several fermentation processes for the microbial production of L-isoleucine have been developed and, in fact, this amino acid is now produced on an industrial scale. However, the fermentation processes still have some disadvantages because of high carbon consumption during cell growth, a low conversion rate from the carbon source to the final product, and high concentrations of by-products. Because chemical characteristics of branched-chain amino acids such as L-isoleucine, L-leucine, and L-valine are similar, it is very difficult to separate them from each other. Therefore, it is preferable that content of L-leucine and L-valine in the broth should be low to facilitate the selective removal and purification of L-isoleucine.

L-ISOLEUCINE PRODUCTION BY LIVING CELL REACTION PROCESS

We have devised a new process: the living cell reaction (LCR) process using native immobilization cells of *Br. flavum* MJ-233 for L-isoleucine and L-valine production (45,46). This process differs from the usual fermentation method in that nongrowing viable cells are used, and the production of L-isoleucine and L-valine is carried out under conditions of repressed cell division and growth. A minimal medium lacking the essential growth factor, biotin, was used as the reaction mixture in the LCR process. For isoleucine production, when ethanol was the energy source and α-ketobutyric acid (α-KB) was the precursor, the productivity of L-isoleucine was 26 g/L/day. The content of L-isoleucine in total amino acids produced in the mixture was 97%, and the molecular yield from α-KB was 95% (46). This is expected to result in a simplified purification process. Moreover, L-isoleucine could be formed at high yield for several runs by recycling the same cells. However, two unnatural amino acids, which were economically difficult to remove from L-isoleucine by ordinary chromatographic separation, were found as by-products in the LCR process for L-isoleucine production. These were identified as norvaline (Nva) and *O*-ethylhomoserine (*O*-EH) (47). Kisumi et al. reported previously that Nva was formed from threonine or α-aminobutyric acid by *Serratia marcescens* in the fermentation broth of L-isoleucine, and they suggested that Nva formation was closely related to leucine biosynthesis (48,49). We reported that Nva formation was depressed by using a leucine auxotrophic *B. flavum* mutant in the LCR process (47). It was also shown that, in *Br. flavum*, Nva formation was closely related to leucine biosynthesis. On

the other hand, Murooka and Harada reported that *O*-EH was specifically formed from ethanol by ethanol-assimilating bacteria (50). Also, Murooka et al. reported that *O*-EH was synthesized from ethanol and *O*-acetyl-homoserine, which is an intermediate of methionine biosynthesis in *Corynebacterium* and *Saccharomyces* (51). We reported that *O*-EH formation was repressed by addition of L-methionine to the reaction mixture in the LCR process using *Br. flavum* (47). However, the homoserine-*O*-acetyltransferase of *Br. flavum* was not subject to either inhibition or repression by addition of L-methionine, and the *O*-EH-forming enzyme, which converts *O*-acetylhomoserine to *O*-EH, was speculated to be repressed by L-methionine (47).

In addition, the LCR process could also be applied to L-valine production from glucose at high yield. In this case, the rate of production of L-valine was 35 g/L/day and the molecular yield from glucose was 80% (46).

Although cloning and expression of *ilv* genes from various organisms has been reported, there is so far no report regarding L-isoleucine production using a recombinant strain that increased or improved isoleucine biosynthetic enzymes by such genetic engineering techniques.

BIBLIOGRAPHY

1. M. Kisumi, *J. Biochem.* **52**, 390–399 (1962).
2. E.A. Adelberg, *J. Bacteriol.* **76**, 326 (1958).
3. M. Inui, A. A. Vertès, M. Kobayashi, Y. Kurusu, and H. Yukawa, *DNA Sequence* **3**, 303–310 (1993).
4. J.V. Schloss and A. Aulabaugh, *Z. Naturforsh., C* **45**, 544–551 (1990).
5. B.J. Mazur, C. F. Chui, and J. K. Smith, *Plant Physiol.* **85**, 1110–1117 (1987).
6. A. Aulabaugh and J. V. Schloss, *Biochemistry* **29**, 2824–2830 (1990).
7. O.M. Aguilar and D.H. Grasso, *J. Bacteriol.* **173**, 7756–7764 (1991).
8. R.C. Wek, C.A. Hauser, and G.W. Hatfield, *Nucleic Acids Res.* **13**, 3995–4010 (1985).
9. R.P. Lawther, D.H. Calhoun, C.W. Adams, C.A. Hauser, J. Gray, and G.W. Hatfield, *Proc. Natl. Sci. USA* **78**, 922–925 (1981).
10. C.H. Squires, M. DeFelice, S.R. Wessler, and J.M. Calvo, *J. Bacteriol.* **147**, 797–804 (1981).
11. M.A. Vandeyar, P.B. Vander Horn, J.A. Rafael, J.A. Grandoni, and S.A. Zahler, unpublished.
12. J. Tarleton, J. Malakooti, and B. Ely, *J. Bacteriol.* **176**, 3765–3774 (1994).
13. C. Keilhauer, L. Eggeling, and H. Sahm, *J. Bacteriol.* **175**, 5595–5603 (1993).
14. J.J. Godon, M.C. Chopin, and S.D. Ehrlich, *J. Bacteriol.* **174**, 6580–6589 (1992).
15. L. Gusberti, R. Cantoni, E. De Rossi, M. Branzoni, and G. Riccardi, *Gene* **177**, 83–85 (1996).
16. E. De Rossi, R. Leva, L. Gusberti, P.L. Manachini, and G. Riccardi, *Gene* **166**, 127–132 (1995).
17. A. Milano, E. De Rossi, E. Zanaria, L. Barbierato, O. Ciferri, and G. Riccardi, *J. Gen. Microbiol.* **138**, 1399–1408 (1992).
18. K. Sathasivan, G.W. Haughn, and N. Murai, *Nucleic Acids Res.* **18**, 2188 (1990).
19. P.A. Wiersma, M.G. Schmiemann, J.A. Condie, W.L. Crosby, and M.M. Moloney, *Mol. Gen. Genet.* **219**, 413–420 (1989).
20. S.C. Falco, K.D. Dumas, and K.J. Livak, *Nucleic Acids Res.* **13**, 4011–4027 (1985).
21. D. Friedberg and J. Seijffers, *Z. Naturforsc., C* **45**, 538–543 (1990).
22. M. Inui, A.A. Vertès, M. Kobayashi, Y. Kurusu, and H. Yukawa, *DNA Sequence* **4**, 95–103 (1993).
23. S. Rieble and S.I. Beale, *J. Bacteriol.* **174**, 7910–7918 (1992).
24. R.C. Wek and G.W. Hatfield, *J. Biol. Chem.* **261**, 2441–2450 (1986).
25. J.G.L. Petersen and S. Holmberg, *Nucleic Acids Res.* **14**, 9631–9651 (1986).
26. R. Dumas, G. Curien, R.T. DeRose, and R. Douce, *Biochem. J.* **294**, 821–828 (1993).
27. R. Dumas, M. Lebrun, and R. Douce, *Biochem. J.* **277**, 496–475 (1991).
28. S.M. Arfin, B. Ratzkin, and H.E. Umbarger, *Biochem. Biophys. Res. Commun.* **37**, 902–908 (1969).
29. M.D. Watson, J. Wild, and H.E. Umbarger, *J. Bacteriol.* **139**, 1014–1020 (1979).
30. R.P. Lawther, R.C. Wek, J.M. Lopes, R. Pereira, B.E. Taillon, and G.W. Hatfield, *Nucleic Acids Res.* **15**, 2137–2155 (1987).
31. F. Kunst et al. *Nature* **390**, 249–256 (1997).
32. J. Malakooti and B. Ely, *J. Bacteriol.* **176**, 1275–1281 (1994).
33. M. Kisumi, S. Komatsumara, and I. Chibata, *Appl. Environ. Microbiol.* **34**, 647–653 (1977).
34. M. Kisumi, S. Komatsubara, M. Sugiura, and I. Chibata, *J. Bacteriol.* **107**, 741–745 (1971).
35. T. Ramarkrishnan and E.A. Adelberg, *J. Bacteriol.* **87**, 566–573 (1964).
36. S. Komatsubara, M. Kisumi, and I. Chibata, *J. Gen. Microbiol.* **119**, 51–61 (1980).
37. S. Komatsubara, M. Kisumi, and I. Chibata, *Appl. Environ. Microbiol.* **38**, 1045–1051 (1979).
38. I. Shiio, A. Sasaki, S. Nakamori, and K. Sano, *Agric. Biol. Chem.* **37**, 2053–2061 (1973).
39. I. Shiio and S. Nakamori, *Agric. Biol. Chem.* **34**, 448–456 (1970).
40. R. Miyajima and I. Shiio, *J. Biochem.* **71**, 951–960 (1972).
41. S. Ikeda, I. Fujita, and F. Yoshinaga, *Agric. Biol. Chem.* **40**, 511–516 (1976).
42. S. Ikeda, I. Fujita, and F. Yoshinaga, *Agric. Biol. Chem.* **40**, 517–522 (1976).
43. I. Eggeling, C. Cordes, L. Eggeling, and H. Sahm, *Appl. Microbiol. Biotechnol.* **25**, 346–351 (1987).
44. H. Kase and K. Nakayama, *Agric. Biol. Chem.* **41**, 109–116 (1977).
45. M. Terasawa, N. Kakinuma, K, Shikata, and H. Yukawa, *Process Biochem.* **21**, 196–199 (1986).
46. M. Terasawa, M. Inui, M. Goto, K. Shikata, M. Imanari, and H. Yukawa, *J. Microbiol. Biotechnol.* **5**, 289–294 (1990).
47. M. Terasawa, M. Inui, M. Goto, Y. Kurusu, and H. Yukawa, *Appl. Microbiol. Biotechnol.* **35**, 348–351 (1991).
48. M. Kisumi, J. Kato, S. Komatsubara, and I. Chibata, *Appl. Microbiol.* **21**, 569–574 (1971).
49. M. Kisumi, S. Komatsubara, M. Sugiura, and I. Chibata, *J. Gen. Microbiol.* **69**, 291–297 (1971).
50. Y. Murooka and T. Harada, *J. Bacteriol.* **96**, 314–317 (1968).
51. Y. Murooka, K. Kakihara, T. Miwa, K. Seto, and T. Harada, *J. Bacteriol.* **130**, 62–73 (1977).

KINETICS, ENZYMES

ATHEL CORNISH-BOWDEN
National Center for Scientific Research
Marseilles, France

KEY WORDS

Enzymes
Inhibition
Kinetics
Manipulation of enzyme activity
Multienzyme systems

OUTLINE

INTRODUCTION

The kinetic behavior of enzymes has been studied in detail for a century, beginning with the classic work of Henri (1) and Michaelis and Menten (2). The objectives have been threefold: to gain an understanding of the mechanisms of enzyme action; to illuminate the physiological roles of enzyme-catalyzed reactions; and, mainly in recent years, to manipulate enzyme properties for biotechnological ends. Experimental practice has been overwhelmingly dominated by the first of these aims; most experiments have been designed as if shedding light on the mechanism was the principal, or even the only, objective. However, although much valuable information has been obtained in this way, there are some important aspects of enzyme function that are obscured when working mainly with isolated enzymes in conditions far removed from those that exist in the cell. For this reason the latter part of this chapter will be devoted to a discussion of enzymes as they behave in complex mixtures.

MICHAELIS–MENTEN KINETICS

The starting point for any discussion of enzyme kinetics is the Michaelis–Menten equation, which expresses the initial rate v of a reaction at a concentration a of the substrate transformed in a reaction catalyzed by an enzyme at total concentration e_0:

$$v = \frac{k_0 e_0 a}{K_m + a} = \frac{Va}{K_m + a} \quad (1)$$

The parameters are k_0, the catalytic constant, and K_m, the Michaelis constant. The form shown in the middle is more fundamental than that on the right, but the second form, in which $k_0 e_0$ is written as the limiting rate V, is often used because the enzyme concentration in meaningful units is often not known. V is the limit that v approaches as the enzyme becomes saturated (that is, when a becomes very large) and K_m is the value of a at which $v = 0.5V$ (that is, at which the rate is half-maximal). The ratio k_0/K_m is called the specificity constant and given the symbol k_A: It is a more fundamental constant than K_m in the analysis of enzyme mechanisms (i.e., it has a simpler mechanistic meaning), but the equation is usually written in terms of V (or k_0) and K_m nonetheless.

The principal assumption implied by equation 1 is that the rate of the reverse reaction is negligible: This may be because the reaction is irreversible for practical purposes, but even with reversible reactions the reverse reaction can be made negligible by measuring the rate in the absence of products and by extrapolating the rate back to zero time, that is, by estimating the initial rate at a time when no products have accumulated. Even if there is no significant reverse reaction, products can still affect the rate of the forward reaction because product inhibition (discussed later) is a common phenomenon.

Another assumption is that the reaction is allowed sufficient time to reach a steady state because all reactions pass through an initial acceleration phase known as the transient state. This phase is normally very brief (a few milliseconds), and in practice enzymes are usually studied under steady-state conditions. Note, however, that this is made experimentally possible by working with extremely small enzyme concentrations compared with those that may exist in the cell. The use of very low enzyme concentrations has two important consequences: First, it normally means that the enzyme concentration can be neglected in comparison with the substrate concentration, and second it makes the steady-state rate sufficiently slow to be easily measured and the steady-state phase long enough to be meaningful. If high enzyme concentrations were used, the transient state would be as brief as before, but the steady state would also be very brief, so that there would be no period in which one could adequately treat the rate as constant.

Equation 1 may be derived from the following model

$$A + E \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} EA \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} EP \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} E + P \qquad (2)$$

which assumes that the reaction passes through an enzyme–substrate complex EA, which undergoes catalytic transformation to an enzyme–product complex EP, which then breaks down to form products. Although real enzyme mechanisms may be more complicated than this, every reaction passes through steps of substrate binding, chemical transformation, and product release. In simple introductory treatments the second and third steps are often treated as a single step, but conceptually they are clearly distinct. In reactions of more than one substrate, the steps do not necessarily occur in the order one might guess; that is, some products may be released before all substrates have bound, but the general principle that any reaction involves the same three kinds of step remains valid.

The Michaelis–Menten parameters can be defined as follows in terms of the rate constants shown in equation 2:

$$k_0 = \frac{k_2k_3}{k_{-2} + k_2 + k_3};\ k_A = \frac{k_1k_2k_3}{k_{-1}k_{-2} + k_{-1}k_3 + k_2k_3};$$
$$K_m = \frac{k_{-1}k_{-2} + k_{-1}k_3 + k_2k_3}{k_1(k_{-2} + k_2 + k_3)} \qquad (3)$$

Note that none of the three parameters has a simple transparent meaning. The interpretations commonly attributed to them depend on additional simplifying assumptions that are not always correct. For example, K_m is often said to be equal to the equilibrium constant k_{-1}/k_1 for dissociation of A from EA, but the expression in equation 3 does not take this form unless k_2 is very small. As there is no good reason for k_2 to be small, and indeed ideas of evolutionary optimization of enzyme function lead one to expect the opposite when the enzyme is acting on its natural physiological substrate, it follows that K_m should not, in general, be regarded as a measure of the equilibrium dissociation constant.

Despite these difficulties in providing a detailed mechanistic meaning to K_m, it does provide a measure of the tightness of substrate binding in the steady state, as it is quite correct to take a/K_m as equal to the ratio of the sum of concentrations of all enzyme complexes (i.e., both EA and EP) over the concentration of free enzyme. Similarly k_0 provides a valid measure of the capacity of the enzyme–substrate complex to react to give products, even if it cannot be interpreted as the rate constant for a unique step in the mechanism. The reason for the term *specificity constant* for k_A, that is, its relationship to enzyme specificity, will become clear once inhibition has been considered.

Graphical Analysis

Equation 1 defines a hyperbolic dependence of rate on substrate concentration, as illustrated in Figure 1. The initial steep rise in v as a increases from zero is rapidly transformed into the phenomenon of saturation, whereby further increases in a produce smaller and smaller increases in v, which approaches, but does not reach or exceed, the limiting rate V. The rectangular hyperbola makes this type of plot inconvenient for estimating the values of the kinetic parameters (because the line does not approach the saturation limit closely enough at reasonable values of a to allow direct measurement of V). For this purpose, therefore, it is usual to transform equation 1 into one of the following three forms, which underlie the three straight-line plots illustrated in Figures 2 through 4:

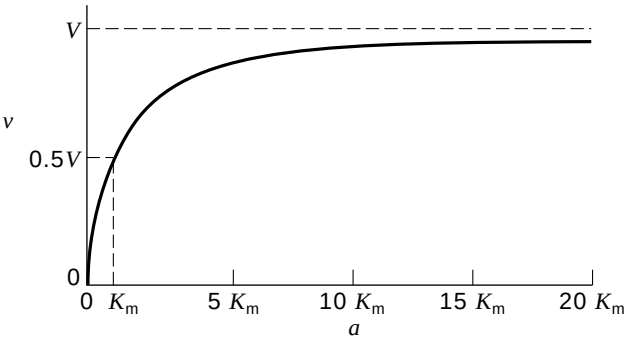


Figure 1. Michaelis–Menten dependence of rate on substrate concentration. The curve is a rectangular hyperbola through the origin, approaching a limit of $v = V$ at saturation. The rate is 0.5 V at a substrate concentration equal to the Michaelis constant, K_m, but note that the approach to the limit is slow, so that, for example, even at $a = 10K_m$ the rate is still nearly 10% less than V.

$$\frac{1}{v} = \frac{1}{V} + \frac{K_m}{V}\frac{1}{a} \qquad (4)$$

$$\frac{a}{v} = \frac{K_m}{V} + \frac{1}{V}a \qquad (5)$$

$$v = V - K_m\frac{v}{a} \qquad (6)$$

The double-reciprocal plot, illustrated in Figure 2 and based on equation 4, is the mostly widely used, but it is also the least satisfactory because it distorts the effect of experimental error to such an extent that it is difficult to form any visual judgement of where the best line should be drawn. The other two plots are better, and the plot of v

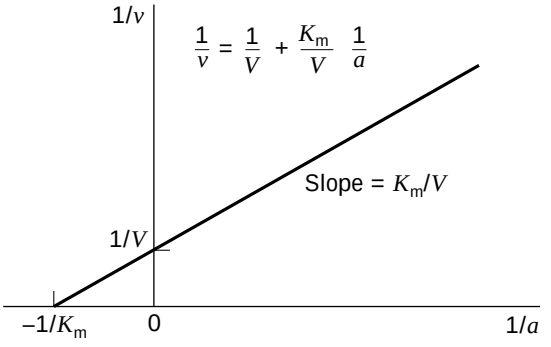


Figure 2. The double-reciprocal plot. This is the most widely used method of plotting the Michaelis–Menten equation as a straight line. However, the severe distortion of any experimental errors in the original data causes it to give a misleading impression.

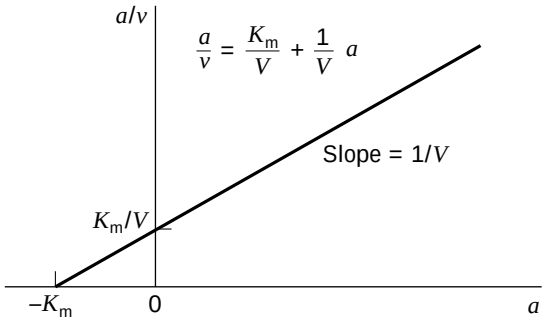


Figure 3. The plot of a/v against a. This alternative to the plot shown in Figure 2 produces much less distortion of the experimental error.

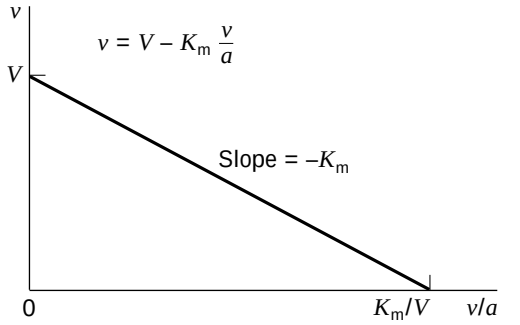


Figure 4. The plot of v against v/a. The third way of plotting the Michaelis–Menten equation as a straight line also avoids the error-distorting property of the plot shown in Figure 2, and maps the entire range of observable rates (from 0 to V) onto a finite range of paper. This is a desirable property because it makes it impossible to disguise deficiencies in the experimental design.

against v/a (Fig. 4, equation 6) has the particular advantage that the entire observable range of v values, from 0 to V, is mapped onto a finite range of graph; this makes it easy to judge by eye if an experiment has been well designed. On the other hand, it has the disadvantage that v, normally the less reliable measurement, contributes to both coordinates, and errors in v cause deviations along lines through the origin, rather than parallel with one or the other axis.

In modern practice it is usually best to regard these plots as for illustration purposes only, and to use suitable computer programs for the actual parameter estimation. For this purpose it is not sufficient just to apply unweighted linear regression to the straight-line plots, as this suffers from the same statistical distortions as the plots themselves. Full treatment would require more space than is available here, but see Ref. 3. The following two equations for calculating best-fit values of K_m and V give satisfactory results if the v values have uniform coefficient of variation (uniform standard deviation expressed as a percentage), as is usually at least approximately correct:

$$K_m = \frac{\Sigma v^2 \Sigma(v/a) - \Sigma(v^2/a)\Sigma v}{\Sigma(v^2/a^2)\Sigma v - \Sigma(v^2/a)\Sigma(v/a)} \quad (7)$$

$$V = \frac{\Sigma(v^2/a^2)\Sigma v^2 - [\Sigma(v^2/a)]^2}{\Sigma(v^2/a^2)\Sigma v - \Sigma(v^2/a)\Sigma(v/a)} \quad (8)$$

Each summation is made over all observations.

Two-Substrate Reactions

Enzymes that catalyze reactions of a single substrate are only a small minority of all the enzymes known, but the Michaelis–Menten equation remains useful for examining the kinetics of the more common case of a reaction with two substrates and (often but not necessarily) two products, because such a reaction normally obeys Michaelis–Menten kinetics when only one substrate concentration is varied at a time. This is illustrated by the following typical equation for such a reaction:

$$v = \frac{k_0 e_0 ab}{K_{iA}K_{mB} + K_{mB}a + K_{mA}b + ab} \quad (9)$$

Although at first sight this appears quite different from equation 1, it can be arranged in the same form if one of the two substrate concentrations, for example b, is treated as a constant:

$$v = \frac{\left(\frac{k_0 b}{K_{mB} + b}\right)e_0 a}{\frac{K_{iA}K_{mB} + K_{mA}b}{K_{mB} + b} + a} \quad (10)$$

The two fractions in this equation can be regarded as "apparent" values of the Michaelis–Menten parameters for A, that is, the equation can be written as

$$v = \frac{k_0^{app} e_0 a}{K_{mA}^{app} + a} \quad (11)$$

with

$$k_0^{app} = \frac{k_0 b}{K_{mB} + b};\ k_A^{app} = \frac{(k_0/K_{mA})b}{(K_{iA}K_{mB}/K_{mA}) + b};$$
$$K_{mA}^{app} = \frac{K_{iA}K_{mB} + K_{mA}b}{K_{mB} + b} \quad (12)$$

Notice that the expressions for the apparent values of k_0 and k_A are both individually of Michaelis–Menten form with respect to b, whereas that for the apparent value of K_m is more complicated: This behavior is quite typical and is one of the reasons why k_0 and k_A should regarded as more fundamental parameters than K_m. More generally, the concept of apparent parameters pervades the analysis of simple cases in steady-state enzyme kinetics, being important for the study of reactions with multiple substrates and inhibition.

INHIBITION AND ACTIVATION

Inhibition

For most enzyme-catalyzed reactions, molecules exist that resemble the substrate closely enough to bind to the en-

zyme, but not closely enough to undergo a chemical reaction. Such a molecule is known as a competitive inhibitor and causes competitive inhibition, characterized by a rate equation of the following form:

$$v = \frac{k_0 e_0 a}{K_m(1 + i/K_{ic}) + a} \tag{13}$$

in which i is the concentration of the inhibitor and K_{ic} is the competitive inhibition constant. (The qualification "competitive" and the second subscript c are usually omitted if only this simplest kind of inhibition is being considered.)

Inhibitors can interfere with catalysis as well as with substrate binding. In the simplest case, an inhibitory term affects the variable term in the denominator of the Michaelis–Menten equation, instead of the constant term:

$$v = \frac{k_0 e_0 a}{K_m + a(1 + i/K_{iu})} \tag{14}$$

This is called uncompetitive inhibition, and the inhibition constant K_{iu} is the uncompetitive inhibition constant. This is important as a limiting case of inhibition, but in its pure form it is not at all common. Much more often one has mixed inhibition, when both competitive and uncompetitive effects occur simultaneously:

$$v = \frac{k_0 e_0 a}{K_m(1 + i/K_{ic}) + a(1 + i/K_{iu})} \tag{15}$$

There is no particular reason for the two inhibition constants K_{ic} and K_{iu} to be equal, and most of the mechanisms one might propose to account for mixed inhibition lead one to expect them to be different, yet the case where $K_{ic} = K_{iu}$ is often given an undeserved prominence in discussions of inhibition, largely because experiments done many years ago suggested that it was a more common phenomenon than it is. This is called noncompetitive inhibition and its rate equation is the same as equation 15, but with both K_{ic} and K_{iu} written simply as K_i.

All of these kinds of inhibition are conveniently discussed in terms of apparent Michaelis–Menten parameters. In the general case (equation 15), these are as follows:

$$k_0^{app} = \frac{k_0}{1 + i/K_{iu}}; \; k_A^{app} = \frac{k_A}{1 + i/K_{ic}}; \; K_m^{app} = \frac{K_m(1 + i/K_{ic})}{1 + i/K_{iu}} \tag{16}$$

Note that the first two expressions have the same form, and both simplify to independence of i in the event that one or other inhibition term is negligible. The expression for the apparent value of K_m is more complicated, especially when one considers how it varies with the different types of inhibition: It increases with the concentration of a competitive inhibitor, it decreases as the concentration of an uncompetitive inhibitor increases, it may change in either direction as the concentration of a mixed inhibitor increases, or it is independent of inhibitor concentration if the inhibition is noncompetitive. In general it is simplest to regard k_A as the parameter affected by competitive inhibition, negligibly so when the competitive component is negligible, k_0 as the parameter affected by uncompetitive inhibition, negligibly so when the uncompetitive component is negligible, and K_m just as the ratio of the two, so $K_m = k_0/k_A$.

The effects of the different kinds of inhibition on the common plots as illustrated in Figures 2 through 4 follows naturally from equation 16. Any competitive effect affects the apparent value of k_A, hence, it increases the slope of the plot of $1/v$ against $1/a$ (Fig. 2), it increases the ordinate intercept of the plot of a/v against a (Fig. 3), and it decreases the abscissa intercept of the plot of v against v/a (Fig. 4). Conversely, any uncompetitive effect increases the ordinate intercept of the plot of $1/v$ against $1/a$, increases the slope of the plot of a/v against a, and decreases the ordinate intercept of the plot of v against v/a. When both components of the inhibition are present, both kinds of effects occur. As an illustration we may consider just one example, the effect of competitive inhibition on the plot of $1/v$ against $1/a$: Plots made at various different inhibitor concentrations produce a family of straight lines intersecting on the ordinate axis, as shown in Figure 5, the lack of effect on the ordinate intercept being a direct consequence of the lack of effect on the apparent value of V.

Specificity

Specificity is the most fundamentally important property of enzymes. Although one is often impressed by the catalytic effectiveness of enzymes, accelerating a reaction is not, in reality, difficult: Heating the reaction mixture in a sealed tube is an efficient way of accelerating virtually any reaction, essentially without limit. What is difficult is to accelerate one selected reaction without at the same time accelerating a great mass of unwanted reactions. What is important about an enzyme, therefore, is not that it is an excellent catalyst for a small set of reactions, but that it is an extremely bad catalyst—virtually without any activity—for all other reactions. In other words, the essential properties of enzymes are that they act under very mild conditions and are highly specific.

In the past, specificity was often assessed by comparing the kinetic parameters for different reactions measured in

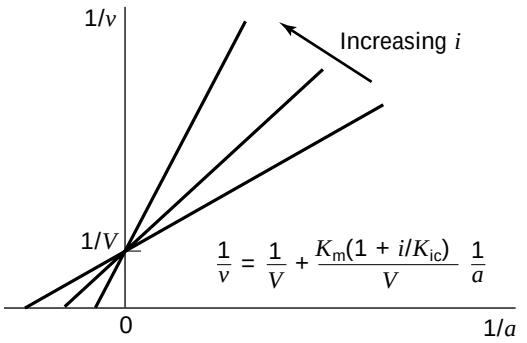


Figure 5. Effect of competitive inhibition on the double-reciprocal plot.

isolation from one another, which led to sterile arguments as to whether specificity was best measured in terms of k_0, K_m, or some combination of the two. This type of argument was resolved once it was realized that the only meaningful way of defining specificity is as a property of an enzyme that allows it to discriminate between substrates that are mixed together. The simplest way to consider this is with a model similar to that for competitive inhibition, except that one assumes that both molecules are capable of reacting. The equation for reaction of one substrate A in the presence of a competing substrate A′ follows an equation similar to that for competitive inhibition (equation 13):

$$v = \frac{k_0 e_0 a}{K_m(1 + a'/K'_m) + a} \tag{17}$$

with the inhibitor concentration replaced by a', the concentration of A′, and the inhibition constant by K'_m, the Michaelis constant for the reaction of A′ considered in isolation. The rate of reaction of A′ is given by the same equation with an obvious transposition of symbols:

$$v' = \frac{k'_0 e_0 a'}{K'_m(1 + a/K_m) + a'} \tag{18}$$

It can then be seen that the ratio of rates is the ratio of substrate concentrations multiplied by the ratio of specificity constants:

$$\frac{v}{v'} = \frac{k_0/K_m}{k'_0/K'_m}\frac{a}{a'} = \frac{k_A a}{k'_A a'} \tag{19}$$

This result, which is still less well known than its importance merits, is the reason for the term *specificity constant*. Note that although inspection of equation 1 suggests that k_A is no more than the parameter that defines the rate at very low substrate concentrations, no assumption about the magnitudes of the concentrations was made in arriving at equation 19. It is thus valid at all concentrations, and the specificity constant measures specificity at all concentrations, not just low ones.

Activation

Activation is the opposite from inhibition, in which a reaction proceeds more rapidly in the presence of a particular molecule than in its absence. It is less common than inhibition, and discussion is complicated by the fact that a variety of quite different phenomena have been termed "activation." The most important of these is a confusion between true activation and the case where the activator is really a component of the substrate. Numerous ATP-handling enzymes are said to be activated by magnesium ions, when in reality the complex MgATP is the true substrate, that is, the species that reacts with the enzyme. Other metal ions, such as the zinc in a number of enzymes, may be true activators as they bind to the enzyme itself and confer catalytic properties on it.

Another misuse of the term activation relates to coenzymes such as NAD in many dehydrogenases: Alcohol dehydrogenase, for example, may be said to be activated by NAD, but although consideration of the metabolic pathway in which the reaction occurs may make it convenient to regard ethanol as the substrate and NAD as the coenzyme, this is meaningless when the reaction is considered in isolation. So far as alcohol dehydrogenase is concerned, it catalyzes a reaction that requires two substrates, ethanol and oxidized NAD; the reaction will not proceed unless both are present, and neither has any more reason to be called the substrate than the other.

When such improper uses of the term are excluded, there remain a number of enzymes for which the true inverse of inhibition occurs. In the simplest cases the equations are just the inverse of inhibition equations, with terms of the form i/K_i replaced by ones of the form K_x/x (for an activator X with activation constant K_x). However, the simplest cases constitute a smaller proportion of the whole than they do for inhibition. This is because whereas most inhibitors inhibit completely, in the sense that enzyme species with inhibitor bound retain no activity as long as the inhibitor remains bound, many enzymes subject to activation retain some activity in the absence of the activator. As a result, full analysis of activation is often more complicated than it is for inhibition, but this will not be discussed further here.

Irreversible Inhibition

The types of inhibition considered so far are examples of reversible inhibition; the inhibitor binds reversibly and catalytic activity returns when the inhibitor is released. Irreversible inhibition also occurs, in which the inhibitor either binds so tightly that for practical purposes it cannot be removed, or reacts with the enzyme and converts it irreversibly to a form that has no catalytic activity. These two cases are conceptually different, and the former is more correctly called tight-binding inhibition, rather than irreversible inhibition. However, they are not easy to distinguish in practice, and have similar practical effects and, hence, similar practical uses.

Although irreversible inhibition has played a smaller part than reversible inhibition in the academic study of enzyme mechanisms, it has far greater industrial and pharmacological importance. This is because competitive inhibitors, the most common kind of reversible inhibitors, are almost completely ineffective in complete physiological systems, for reasons to be considered shortly. By contrast, whenever irreversible (or tight-binding) inhibition occurs in a physiological system, it can be expected to have profound effects. Many toxic and pharmacologically active substances owe their effects to irreversible inhibition.

Both tight-binding and irreversible inhibition manifest themselves in ways that allow them to be confused with noncompetitive inhibition, as in equation 15 with the two inhibition constants equal. This is because the practical effect of irreversible inhibition is not on any of the kinetic parameters in the Michaelis–Menten equation, but on e_0, the total enzyme concentration. However, a decrease in e_0 can be confused with a decrease in the apparent value of k_0, as they occur as a product in equation 1. Although uncompetitive inhibition affects k_0, it does so without affecting k_A, and thus also changes K_m. Decreasing k_0 without

affecting K_m, similar to what one would observe if e_0 decreases, is a definition of noncompetitive inhibition.

Inhibitory Effects in Metabolic Systems

Competitive and uncompetitive inhibition are sufficiently similar in their effects in artificial experiments on isolated enzymes that they are often not distinguished, and an uncompetitive component in mixed inhibition often passes unnoticed. Many inhibitors are described in the literature as competitive inhibitors in the absence of any real evidence of the type of inhibition. This sort of confusion can easily lead to the entirely false idea that they are similar in their effects in systems where the inhibited enzyme is mixed with other enzymes and catalyzes a step in the middle of a pathway.

In a typical experiment in vitro, one decides the concentrations of the various components in advance and measures the rate that results; however, this is very different from what happens in the cell. To a first approximation, an enzyme catalyzing a step in the middle of a pathway must transform its substrate at the rate at which it arrives, that is, within certain limits it has little or no effect on the rate of its reaction, but instead determines the concentrations of the metabolites around it. (This is an oversimplification, but is useful for discussion.)

It is useful therefore to transform equations 13 and 14 into expressions for a in terms of i:

$$a = \frac{vK_m(1 + i/K_{ic})}{k_0e_0 - v} \tag{20}$$

$$a = \frac{vK_m}{k_0e_0 - v(1 + i/K_{iu})} \tag{21}$$

However similar equations 13 and 14 may seem, their transformed versions are drastically different. Equation 20 shows a linear dependence of a on i, which means that increasing i can never result in uncontrolled increases in a. This is illustrated in Fig. 6a. Even at an inhibition concentration equal to the inhibition constant, the substrate concentration is only doubled. By contrast, the curve defined by equation 21 is a rectangular hyperbola (Fig. 6b) that produces a steep and uncontrolled rise in substrate concentration at quite moderate inhibitor concentrations (4). The point is that in competitive inhibition, rises in substrate and inhibitor concentrations oppose one another—not only does the inhibitor compete with the substrate, but equally, the substrate competes with the inhibitor. In uncompetitive inhibition, however, these effects potentiate one another.

It follows that although it is relatively easy to find molecules that will act as competitive inhibitors, it is also largely useless as a strategy for designing pesticides or drugs because it is correspondingly easy for the organism to counteract the effect of the inhibition. To produce major metabolic effects one needs uncompetitive inhibitors, irreversible inhibitors, or tight-binding inhibitors: None of these are as easy to produce as weakly binding competitive inhibitors, but they are far more effective.

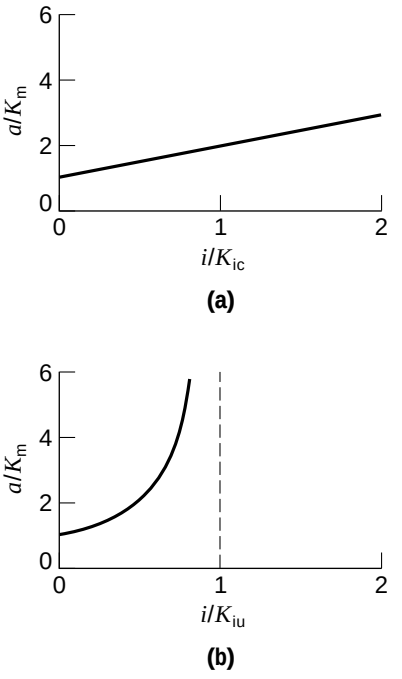


Figure 6. Effects of **(a)** competitive and **(b)** uncompetitive inhibition on the concentration of substrate in a constant-rate system. Both curves are drawn for the case of $a = K_m$ in the absence of inhibitor. Note that both kinds of inhibitor have quantitatively equal effects at very low concentrations, but the initial slope is maintained indefinitely if the inhibition is competitive, whereas it rapidly becomes infinite if the inhibition is uncompetitive.

NON-MICHAELIS–MENTEN BEHAVIOR

All of the cases considered so far can be regarded as generalizations of the Michaelis–Menten equation (equation 1). However, although many enzymes do behave in this way, at least as a first approximation, there are some important exceptions. It is simple to calculate from equation 1 that if $a = K_m/9$ then $v = 0.1V$ and if $a = 9K_m$ then $v = 0.9V$; in other words, spanning the 10–90% range of available rates requires an 81-fold range of substrate concentrations, almost two orders of magnitude. Similar calculations may be done with any of the equations of the Michaelis–Menten type for additional substrates, inhibitors, or activators. Their implication is that as long as enzymes follow Michaelis–Menten kinetics, their rates cannot be adequately varied by manipulating concentrations of substrates, for example, because effective regulation will often require sensitivity to small changes in signals—certainly changes much smaller than two orders of magnitude.

A second difficulty arises from the fact that inhibition of the types considered commonly derives from structural similarities between inhibitors and substrates or products, whereas there is no reason to expect the molecules needed for metabolic signals to resemble the substrates or products of the enzymes that need to respond to the signals. In reality, the concentration of the end product of a pathway often serves as such a signal: too low, and the pathway

needs to be activated; too high, and it needs to be inhibited. It is often found, therefore, that the enzyme that catalyzes the first committed step of a pathway, that is, the first step after a branch point, in the branch that leads to the end product in question, is inhibited by that end product.

For inhibition of this kind to be possible, the enzyme must have a specific binding site for the end product, independent of the binding sites for substrates and products. Such a site is called an allosteric site, and the phenomenon is called allosteric inhibition. Because the need for it often coincides with the need for higher sensitivity than is provided by Michaelis–Menten kinetics and the common kinds of inhibition, allosteric inhibition is often cooperative: This means that the equations that define it are more complicated than those considered above, allowing, for example, a change from 10% to 90% inhibited over a concentration range much smaller than 81-fold, and typically less than 10-fold (though rarely, if ever, less than 3-fold) (Fig. 7).

The fact that allosteric and cooperative behavior are often found together has led many authors to treat the two terms as synonymous, a tendency encouraged by the fact that in one of the most widely accepted models of non-Michaelis–Menten behavior, that of Monod, Wyman, and Changeux (5), both result from the same structural properties attributed to the enzyme. Nonetheless, most careful authors consider the two properties to be conceptually distinct and not necessarily occurring together, so the two terms should be considered distinct as well.

KINETICS OF MULTIENZYME SYSTEMS

As noted in the introduction, nearly all kinetic studies of enzymes have been carried out using isolated enzymes, and although this has been very valuable for arriving at a good understanding of the nature of enzyme catalysis, it is quite inadequate as a guide to how systems of enzymes will behave. One cannot assume that the flux through a metabolic pathway is a property of a unique enzyme catalyzing the rate-limiting step, and that the properties of the pathway as a whole can be deduced from studies of the kinetics of this one enzyme in isolation.

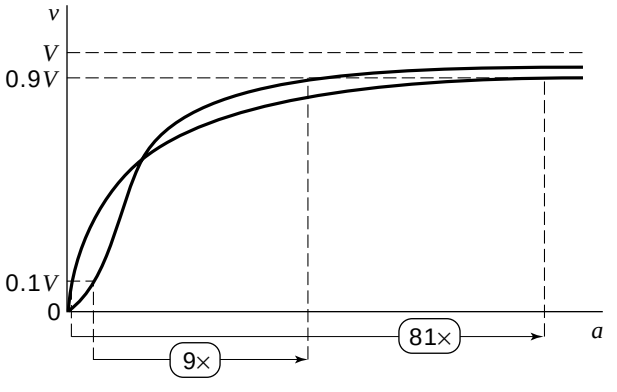


Figure 7. Non-Michaelis–Menten kinetics. For an enzyme obeying the Michaelis–Menten equation (cf. Fig. 1), an 81-fold increase in substrate concentration is needed to bring the rate from 10% to 90% of V. If the enzyme shows positive cooperativity the curve typically becomes sigmoid (S-shaped), and this range of substrate concentrations is decreased (to nine-fold in the example, but in strongly cooperative cases it can be as small as three-fold).

Space does not permit a full analysis of this subject, which has developed from a classic paper by Kacser and Burns (6), but it should suffice to examine an example of a pathway in which biotechnologists have attempted to increase the flux by identifying the rate-limiting enzyme and using genetic manipulation to increase its activity. Tryptophan biosynthesis in yeast provides such an example, tryptophan being a commercially valuable metabolite for which increased production would be very desirable. The tryptophan pathway consists of five enzymes, and in the classical model any one of these could be the "key" enzyme catalyzing the rate-limiting step. However, when the activity of each of these enzymes was increased in turn, either singly or at the same time as others in the pathway, the results were trivial: Increases of enzyme activity of 20-fold or greater produced flux increases of perhaps 30%. Only when all five enzyme activities were increased (or all but one, apparently unimportant, activity) was there a substantial increase in flux, which even then was much smaller than the increase in enzyme activity (7).

The object here is not to analyze in detail why manipulation of tryptophan biosynthesis did not produce the desired results, but to use it to illustrate the point that the whole approach is misconceived. One cannot treat the kinetic behavior of systems of enzymes as if it were determined by the properties of a single component. Moreover, abundant evidence exists to show that all organisms have evolved regulatory mechanisms to control metabolic fluxes and concentrations to satisfy their own requirements. Artificially trying to force more activity in a pathway by increasing the activities of certain enzymes simply stimulates the regulatory mechanisms to resist. Tryptophan biosynthesis is just one example of a general result, and similar efforts in other pathways—for example, increased alcohol production in yeast and increased starch production in potatoes—have produced similar results.

The essential point is that flux control is not a property of a single enzyme in a pathway, but is shared among all the enzymes in the pathway. Strictly speaking, it is shared among all the enzymes in the cell, but it is usually safe to assume that enzymes catalyzing reactions remote from the pathway of interest have very little effect on the flux through it, so to a first approximation one can consider flux control to be shared among the enzymes of the pathway.

To express this idea in quantitative terms, one can define a flux control coefficient for any enzyme by the following equation:

$$C_i^J = \frac{\partial \ln J}{\partial p} \Big/ \frac{\partial \ln v_i}{\partial p} \qquad (22)$$

This definition compares the effect on the flux J through a pathway of some perturbation of the activity of the ith enzyme, represented by a change in the parameter p, with the effect the same perturbation would have on the rate v_i of the same reaction if it were considered in isolation. The

identity of the parameter p does not have to be specified, because as long as it affects only one enzyme, the control coefficient defined by equation 22 is independent of the manner in which the flux and isolated rate are perturbed. However, to make the definition more concrete, consider the case where p is the logarithm of the enzyme concentration, e_i. As most reactions are considered under conditions where the rate is proportional to the enzyme concentration, it will often be true to write $d\ln v_i = d\ln e_i$, so that the denominator in equation 22 has a value of unity and the whole equation simplifies to the following:

$$C_i^J = \frac{\partial \ln J}{\partial \ln e_i} \tag{23}$$

This equation is less abstract and simpler to understand than equation 22, and in the past was often used as a primary definition of a control coefficient. However, this is not recommended, first because it is not always true that the isolated rate is proportional to the enzyme concentration, and second because equation 23 can give the false impression that control coefficients are concerned only with effects brought about by changes in enzyme concentration.

Flux control coefficients are measures of how much the flux through a pathway is dependent on the activities of the individual enzymes. Mathematical analysis shows that they satisfy a property called the summation relationship:

$$\Sigma C_i^J = 1 \tag{24}$$

The interpretation of this relationship would be straightforward if all flux control coefficients were always positive, but in reality negative flux control coefficients also occur; for example, if a pathway contains a branchpoint, the enzymes in one branch normally have negative flux control coefficients for flux through the other. Nonetheless, in practice, negative flux control coefficients are nearly always small in magnitude, so although the sum of positive flux control coefficients may be greater than 1, it is rarely much greater than 1, so that the idea of sharing flux control among all the enzymes of a system is reasonably accurate. It follows that we should not expect any enzyme to have complete control over flux, and the closer to complete control any enzyme approaches, the less control all of the others have, taken together.

Moreover, a flux control coefficient is not a constant, but tends to decrease when the activity of an enzyme is increased. In other words, even if one can identify one enzyme that has a large proportion of the total flux control, increasing its activity will tend to decrease that proportion, so that the amount the flux can be increased by increasing the activity of one enzyme will always be small. Once this is understood, the failure to increase the flux to tryptophan (in the example mentioned) or to achieve significant flux increases (in other similar examples), ceases to be a mysterious result, but is simply what was to be expected.

FURTHER READING

The topic of this article is referred to in all textbooks of biochemistry, but unfortunately the treatment is nearly always very superficial and potentially misleading. Since Mahler and Cordes (8), there has been no serious attempt to cover enzyme kinetics in a general biochemistry text. More specialized textbooks were abundant in the past, but have become much less so. There are, however, two recent ones (9,10) that may be consulted for more information about most of the points discussed here, as well as topics such as the pH- and temperature-dependence of enzyme-catalyzed reactions, which require more detail than can be covered adequately in a short article. Another source is the book by Segel (11), which was originally published in 1975, but has been reissued recently.

Fast reactions are not dealt with in this article, and as they require both techniques and methods of analysis different from those of steady-state kinetics, it is best to refer to a specialized source, such as Hiromi (12), who provides a thorough treatment. Two recent books deal with the kinetics of multienzyme systems, one (13) at a very readable level, the other (14) much more advanced. This is also covered much more briefly in two of the books mentioned earlier (9,10).

BIBLIOGRAPHY

1. V. Henri, *Lois Générales de l'Action des Diastases*, Hermann, Paris, 1903.
2. L. Michaelis and M.L. Menten, *Biochem. Z.* **49**, 333–369 (1913).
3. A. Cornish-Bowden, *Analysis of Enzyme Kinetic Data*, Oxford University Press, Oxford, 1995.
4. A. Cornish-Bowden, *FEBS Lett.* **203**, 3–6 (1986).
5. J. Monod, J. Wyman and J.-P. Changeux, *J. Molec. Biol.* **12**, 88–118 (1965).
6. H. Kacser and J.A. Burns, *Symp. Soc. Exp. Biol.* **27**, 65–104 (1973).
7. P. Niederberger, R. Prasad, G. Miozzari, and H. Kacser, *Biochem. J.* **287**, 473–479 (1992).
8. H.R. Mahler and E.H. Cordes, *Biological Chemistry*, 2nd ed., Harper & Row, New York, 1972.
9. A.R. Schulz, *Enzyme Kinetics: From Diastase to Multi-Enzyme Systems*, Cambridge University Press, Cambridge, 1994.
10. A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, 2nd ed., Portland Press, London, 1995.
11. I.H. Segel, *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*, Wiley, New York, 1993.
12. K. Hiromi, *Kinetics of Fast Enzyme Reactions: Theory and Practice*, Kodansha, Tokyo 1979.
13. D. Fell, *Understanding the Control of Metabolism*, Portland Press, London, 1997.
14. R. Heinrich and S. Schuster, *The Regulation of Cellular Systems*, Chapman & Hall, New York, 1996.

KINETICS, MICROBIAL GROWTH

NICOLAI S. PANIKOV
Institute of Microbiology, Russian Academy of Sciences
Moscow, Russian Federation

KEY WORDS

Cell size distribution
Colonies
Energy and conserved substrates
Growth models
Macrostoichiometry
Maintenance
Microstoichiometry
Physiological state
Steady-state and transient dynamics
Yield

OUTLINE

INTRODUCTION

Kinetics (Greek *κινετικοσ*, forcing to move) is a branch of natural science that deals with the *rates and mechanisms* of any processes—physical, chemical, or biological. Kinetic studies in microbiology cover all dynamic manifestations of microbial life: growth itself, survival and death, product formation, adaptations, mutations, cell cycles, environmental effects, and biological interactions. Kinetics provides a theoretical framework for optimal design in biotechnologies based on fermentation and enzyme catalysis, as well as on employment of outdoor activity of natural microbial populations (wastewater treatment, soil bioremediation, etc.)

Contrary to simple rates measurements, kinetic studies require the perception of the underlying basic *mechanisms* of studied processes. We will define mechanistic studies as those that interpret some complex process as an interplay of several simpler reactions, for example, cell growth can be explained through activity of enzymes and microbial community dynamics can be interpreted through behavior of individual cells and populations. Ideally, mechanistic studies infer the coupling of experimental measurements with analysis of simulating mathematical models. The models formalize postulated mechanisms, so that the comparison of observations and the model's predictions allows one to discard an incorrect hypotheses.

The quantitative studies in microbiology often involve the assessment of *growth stoichiometry*. Stoichiometry [Greek *στοικηειον*, element] is the quantitative relationship between reactants and products in a chemical reaction. In microbiology, stoichiometry stands for a quantitative relationship between substrates and products of microbial processes, including biomass formation (the consequence of complying with mass and energy conservation laws). In practical terms, kinetic and stoichiometry are tightly linked to each other, but stoichiometry mainly addresses problems of a static nature (how much? in what proportion?), whereas kinetics considers the dynamics questions (at what rate? by which mechanism?).

GROWTH STOICHIOMETRY

Macrostoichiometry of Microbial Growth

By analogy to simple chemical reactions, we can represent growth as a conversion of a number of substrates (medium components) into cell mass and products. Growth of aerobic heterotrophic microorganisms can be approximated by the following stoichiometric equation (substrates = biomass + products) (1,2):

$$CH_mO_1 + a_1NH_3 + a_2HPO_4^{2-} + a_3K^+ + \ldots + bO_2 = YCH_pO_nN_qP_oK_v \ldots + a_4CO_2 + a_5H_2O \quad (1)$$

Here, microbial biomass is empirically expressed by the gross formula $CH_pO_nN_qP_oK_v \ldots$, for example, if some av-

erage microbial cell contains per dry cell weight (%) C 46, H 7.5, O 31, N 11; and P, 1.3, then the biomass formula is $CH_{1.9}O_{0.5}N_{0.2}P_{0.01}$. The stoichiometric quotients a_1–a_5 . . . , b and Y (biomass yield) specify quantities of substrate and products of microbial growth. If we know biomass yield and gross formulas of all substrates and products, then quotients a_1–a_6 . . . are easily calculated from conservation conditions. There are at least two such conditions. First, the mass of each element (C, H, O, N, P, K, . . .) on the left side of equation 1 should be equal to that on the right side (mass balance). Second, if ionized substances are involved, we should take into account the balance of charges to satisfy the condition of electroneutrality. Table 1 demonstrates some examples of stoichiometric growth reactions relevant to biotechnology.

Described formalism is useful as a first step in biotechnological studies aimed at planning and optimizing microbial growth. It estimates how much nutrient should be supplied to the fermenter to obtain the required amount of biomass or target product. However, it should be absolutely clear that stoichiometric equations like equation 1 are no more than an approximation to reality. The most severe deviation stems from the fact that unlike chemical reagents, microbial cells are characterized by changeable composition, and stoichiometric coefficients are not true constants. One task of contemporary microbial stoichiometry is to find out the functional relationships between stoichiometric parameters and internal (physiological) and external (environmental) factors.

Growth Yield: Catabolic and Conserved Substrates

The growth yield is one of the main stoichiometric parameters. It is defined as follows

$$Y = -\frac{\delta x}{\delta s} \approx -\frac{\Delta x}{\Delta s} \qquad (2)$$

where Δx is the increase in microbial biomass consequent on utilization of the amount Δs of substrate, and δx and δs are respective infinitely small increments. Rigorous definition of Y as derivative $\delta x/\delta s$ stems from the fact that Y can vary in time, the negative sign being introduced because x and s vary in opposite senses. Sometimes, it is used as the reciprocal of Y: $\alpha = 1/Y$, which is called the *economic coefficient*. It expresses explicitly the nutrient requirements for growth: how many mass units of a particular substrate should be consumed to produce one unit mass of cell material.

The growth efficiency depends generally on the partitioning of consumed element between new cell biomass and extracellular products. The mass balance (total element consumed = amount incorporated into cell plus amount incorporated into extracellular products) is as follows:

$$\delta E_s = \delta E_x + \delta E_p \qquad (3)$$

There are two groups of substrates for microbial growth: (*1*) *catabolic substrates*, which are sources of energy; and (*2*) *anabolic* or *conserved substrates*, which are sources of

Table 1. Selected Macrostoichiometric Equations Describing Growth of Microorganisms with Different Types of Energy Generation

Microbial process	Substrates		Biomass		Products	Stoichiometric parameters	
Heterotrophic growth and by-product formation	$CH_mO_l + a_1O_2 + a_2NH_3$	=	$YCH_pO_nN_q$	+	$Y_pCH_rO_sN_t + a_3CO_2 + a_4H_2O$	$a_1 = 0.5(Yn + Y's - l + a_4) + a_3$ $a_2 = Yq + Y't, a_3 = 1 - Y - Y'$ $a_4 = 0.5[m + Y(3q - p) + Y'(3t - r)]$	A
Phototrophic growth (algae or plant cells)	$CO_2 + a_1H_2O + a_2NH_3$	=	$YCH_pO_nN_q$	+	$a_3CH_rO_sN_t + a_4O_2$	$a_1 = 0.5[Y(p - 3n) + (r - 3t)(1 - Y)]$ $a_2 = Y(n - t) + t, a_3 = 1 - Y$ $a_4 = 0.5(2 + a_1 - Yn - a_3s)$	B
Methanogenesis[a]	$CO_2 + a_1H_2 + a_2NH_3$	=	$YCH_pO_nN_q$		$a_3CH_4 + a_4H_2O$	$a_1 = 4 - 0.5Y(4 + 2n + 3q - p)$ $a_2 = Yq, a_3 = 1 - Y$ $a_4 = 2 - Yn$	C
Nitrification[b]	$YCO_2 + a_1O_2 + a_2NH_4^+$	=	$YCH_pO_nN_q$	+	$a_3NO_3^- + a_4H^+ + a_5H_2O$	$a_1 = 2Y^{-1} - Y[1 + 0.25(p - 2n - 3q)]$ $a_2 = Y^{-1} + Yq, a_3 = Y^{-1}$ $a_4 = 2Y^{-1} + Yq, a_5 = Y^{-1} - 0.5Y(p - 3q)$	D

[a] Particular example of H_2-utilizing methanogenic bacteria.
[b] The total growth balance for two phases of nitrification: oxidation of ammonium to nitrite and subsequent oxidation of nitrite to nitrate. Y is growth yield of bacterial mass per mass unit of oxidized N.

biogenic elements forming cellular material. Catabolic substrates include H_2 for lithotrophic hydrogen bacteria, NH_4^+ and NO_2^- for nitrifying bacteria, S^0 for sulfur-oxidizing bacteria, oxidizable or fermentable organic substances for heterotrophic bacteria and fungi, and so on. Their consumption is accompanied by oxidation and dissipation of chemical substances into extracellular waste products that are no longer reusable as an energy source* (H_2O, NO_3^-, SO_4^{2-}, CO_2, etc.) The anabolic substrates after uptake are incorporated into de novo synthesized cell components, being conserved in biomass (that is why they are called conserved). Contrary to catabolic substrates, they can be reused (e.g., after cell lysis to be taken up by survived cells). The conserved substrates include nearly all the noncarbon sources of biogenic elements (N, P, K, Mg, Fe, and trace elements), CO_2 for autotrophs, and the indispensable amino acids and growth factors. Most catabolic substrates are used also as a source of biogenic elements. We can assess both these components separately in terms of respective yields, Y_E (biomass yield per mass unit of oxidized substrate) and Y_A (biomass yield per mass unit of assimilated substrate), from the experimentally measured yield Y. For C substrate, equation 4 can be specified as follows (total carbon consumed equals C incorporated into cell plus C oxidized to CO_2 to provide energy plus C incorporated into by-products):

$$\delta C_s = \delta C_x + \delta C_{CO_2} + \delta C_P \tag{4}$$

Let us neglect the last term δC_P (by assuming that extracellular by-products can be reused and functionally are equivalent to C substrate) and divide the substrate balance by δC_x, which is the amount of biomass C produced, then:

$$\frac{1}{Y} = 1 + \frac{1}{Y_E} \tag{5}$$

where Y = g biomass C g^{-1} substrate C and Y_E = g biomass C g^{-1} CO_2-C

$$\frac{1}{Y} = \frac{\sigma_x}{\sigma_s} + \frac{12}{Y_E \sigma_s} \tag{6}$$

where Y = g CDW g^{-1} substrate, Y_E = g CDW $mmol^{-1}$ CO_2, and σ_x and σ_s are fractions of carbon in biomass and substrate, respectively. For example, if total measurable yield Y is 0.6 g biomass C g^{-1} glucose C, it means that from each g of consumed C, 0.6 g is incorporated into biomass (assimilated), and 0.4 g is dissimilated (oxidized to CO_2), then $Y_E = 0.6/0.4 = 1.5$. To calculate oxygen demand for aerobic growth (or biomass yield on O_2) we have a balance (oxygen required to produce 1 g CDW equals oxygen required to burn substrate consumed to produce 1 g CDW minus oxygen required to burn 1 g CDW):

*Fermentation products such as acetate, ethanol, butyrate, and H_2 seem to be an exception because they do contain reusable oxidation potential, but it is not available under anaerobic conditions supervising fermentation.

$$1/Y_{O_2} = A/Y - B \tag{7}$$

where A and B are constants estimated from stoichiometry of their respective combustion reactions (see equation 10 later), for example, the value of A is 33.33 mmol O_2 g^{-1} glucose and B is about 42 mmol O_2 g^{-1} CDW. The relationship between biomass yields on O_2 and CO_2 is derived from comparison of equations 6 and 7:

$$\frac{Y_{CO_2}}{Y_{O_2}} = \frac{A - BY}{\sigma_s - \sigma_x Y} \tag{8}$$

Now we will go back to the general substrate balance (equation 3) and derive an expression for conserved substrate. Again, we neglect term δE_P (because extracellular products are assumed to be reusable) and divide the balance by δx, which is the amount of biomass produced:

$$\frac{1}{Y} = \frac{\delta E_x}{\delta x} \approx \sigma_x \tag{9}$$

where σ_x is the intracellular content of element incorporated into biomass from consumed substrate. Sometimes σ_x is called the *cell quota*. The values $1/Y$ and σ_x are not identical although they have the same dimension (e.g., milligram N per gram biomass) and very close numeric value. The reciprocal $1/Y$ is characterizing the *process* (the expenditure of conserved substrate to synthesize biomass unit), whereas σ_x is an index of cell composition (the content of intracellular N per biomass unit). Formally, $1/Y$ is equal to the σ_x value of an infinitely small increment of cell biomass, and σ_x is the averaged value for entire cell. Notice that although σ_x is a slow and $1/Y$ is a rapid variable, their numerical values are exactly the same for balanced steady-state growth and can differ considerably during transients.

Yield Variation as Dependent on the Chemical Nature of Organic Substrates

In this section, we will discuss why biomass yield varies when microorganisms are grown on different C substrates. This problem was best solved within the framework of the *theory of mass and energy balance* (TMEB) (3). Evidently, the fraction of C in dry biomass is almost constant. By contrast, the content of carbon in utilized substrates, σ_s, and energetic quality of substrate vary over a broad range (e.g., compare methane versus oxalic acid). To characterize substrate and biomass by a single common measure, TMEB uses an index of *degree of carbon reduction*, γ related to the internal energy of organic compounds. The heat liberated by biological or chemical oxidation is proportional to oxygen uptake or equally to the number of electrons gained by oxygen from oxidized substrates, according to Payne's term *available electrons* (*ae*) (4). The heat production from an oxidation reaction averages at 27 kcal per *ae* equivalent. A carbon reduction degree, γ is defined as the number of *ae* per one carbon atom. Its numeric value can be determined from the stoichiometry of the oxidation reaction:

$$CH_pO_nN_q + bO_2 = CO_2 + 0.5(p - 3q)H_2O + qNH_3 \tag{10}$$

$$\gamma = 4b = 4 + p - 2n - 3q \tag{11}$$

The *ae* balance for equation 1 can be written as

$$\gamma_s + b(-4) = Y\gamma_x + Y_P\gamma_p \tag{12}$$

where γ_s, γ_x, and γ_p are the carbon reduction degree of, respectively, substrate, biomass, and extracellular product. Dividing both sides of equation 1 by γ_s we obtain the relationship delineating the *ae* distribution between oxygen (*ae* used for respiration), biomass, and the intracellular product:

$$\frac{4b}{\gamma_s} + \frac{Y\gamma_x}{\gamma_s} + \frac{Y_P\gamma_p}{\gamma_s} = 1 \tag{13}$$

The second term in this equation is the fraction of *ae* transferred to biomass from utilized substrate, termed the *energetic growth yield*.

$$\eta = Y_C\gamma_x/\gamma_s \tag{14}$$

The third term designates that fraction of total substrate internal energy that is transferred to the product. It is called the *energetic product yield*

$$\zeta = Y_P\gamma_p/\gamma_s \tag{15}$$

Energetic yield η is related to other stoichiometric parameters as follows:

$$\eta = Y\sigma_x\gamma_x/(\sigma_s\gamma_s)$$

$$\eta = Y_C\gamma_x/\gamma_s$$

where Y is g CDW/g substrate and Y_C is g CDW-C/g substrate C.

The advantage of using η is that it varies within a much smaller range than other yield expressions. At one and the same efficiency of energy utilization (η), the conventional biomass C yield Y_C is proportional to substrate reduction degree γ_s and, for example, it is four times higher on glucose ($\gamma_s = 4$) than on oxalate ($\gamma_s = 1$), 0.48 and 0.12 g C g^{-1} C, respectively (assuming $\eta = 0.5$ and $\gamma_x = 4.2$). The energetic growth yield η is more or less constant (0.5 to 0.7) for substrates with $\gamma_s \leq 4.2$ (4.2 corresponds to average reduction degree of microbial biomass), and it declines at higher γ_s.

The attractiveness of macrostoichiometry and TMEB is that all growth coefficients are interrelated and could be measured from any available components of the culture mass balance. For example, if you cannot record microbial growth by conventional routine as dry weight biomass (because of presence of solids in broth liquid), you may still calculate it from N or O_2 uptake, CO_2 evolution, pH titration rate, and so on.

Variations in Yield from Energy Source, Maintenance Requirements

To multiply and grow cells requires energy, but the opposite is not true: cells do not require growth to spend energy. Sometimes catabolic machinery is entirely wasteful (respiration without cell growth) and always at least some minor part of energy consumption is diverted from growth. To account for this phenomenon, it was postulated that microbes and cells require energy not only for growth but also for other *maintenance purposes*. Certain specific maintenance functions recognized now are turnover of cell material, osmotic work to maintain concentration gradients between the cell and its exterior, and cell motility.

According to conventional definition of maintenance (5), the balance of energy source is total energy source consumed equals consumption for cell growth plus consumption for maintenance:

$$\delta S_E = \delta S_G + \delta S_M \tag{16}$$

Let us divide it by δx, the amount of biomass produced, then

$$\frac{1}{Y} = \frac{\delta S_G}{\delta x} + \frac{\delta S_M}{\delta x} = \frac{1}{Y^{max}} + \frac{m}{\mu} \tag{17}$$

Here, $Y^{max} = \delta x/\delta S_G$ is true growth yield, that is, yield under imaginary conditions of maintenance being zero. The maintenance coefficient, m, is introduced as the specific (i.e., expressed per unit of biomass) rate of energy consumption for maintenance functions: $m = (1/x)(\delta S_M/\delta t)$. The ratio m/μ on the right side of equation 17 was derived as follows: $m/\mu = [(1/x)(\delta S_M/\delta t)]/[(1/x)(\delta x/\delta t)] = \delta S_M/\delta x$.

If we divide equation 16 by $x\delta t$ (note that the second term is $\delta S_G/(x\delta t) = [\delta x/(x\delta t)]/[\delta x/\delta S_G] = \mu/Y^{max}$), then we have:

$$q = \mu/Y^{max} + m \tag{18}$$

where q is specific rate of energy source consumption, $q = (1/x)(\delta S_E/\delta t)$.

It should be noticed that Y^{max} is a parameter, but not the yield of a real culture that always has some nonzero maintenance requirements. It is a very common mistake in the application of the maintenance concept to a particular organism: to take the real measured Y value and pick up from literature some average m coefficient. The correct way would be either to borrow concurrently two parameters Y^{max} and m or to treat actually observed Y as a variable that is altered along with specific growth rate μ according to equation 17:

$$Y = \frac{\mu Y^{max}}{\mu + mY^{max}} \tag{19}$$

There is another way to formulate maintenance requirements by stating that the net growth of cells μ is the difference between true growth (μ_{true}) and endogenous decay of cellular components (specific rate, a):

$$\mu = \mu_{true} - a$$

$$\mu_{true} = \mu + a$$

Then, for the rate of energy source uptake, we have

$$\frac{\mu x}{Y} = \frac{\mu_{true} x}{Y^{max}} = \frac{(\mu + a)x}{Y^{max}}$$

or

$$\frac{1}{Y} = \frac{1}{Y^{max}} + \frac{a}{\mu Y^{max}} \qquad (20)$$

Comparing equations 17 and 20, we see that $a = mY^{max}$.

Experimental Determination of *m*

To practically determine the maintenance coefficient, the microorganisms are grown in chemostat culture limited by energy sources at several dilution rates D (numerically D is equal to specific growth μ if steady state is achieved). At each D, we have to measure steady-state biomass $\tilde{x}$ and at least one of the following quantities (*1*) residual substrate, $\tilde{s}$ to calculate $Y = \tilde{x}/(s_0 - \tilde{s})$; and (*2*) the rate of respective energy-yielding process, such as respiration rate, v_{resp}, from O_2 uptake or CO_2 production rates to calculate specific metabolic activity, $q = v_{resp}/\tilde{x}$. These data are fitted to equations 17 or 18, m and Y^{max} being found as nonlinear regression parameters. An example is presented in Figure 1. Most available experimental data do obey this relationship. However, considerable deviation occurs at very low growth rates usually attained in chemostat with biomass retention or in dialysis culture. The experimental Y values for slowly growing cells are higher than predicted by equations 17 and 18 (see inset on Figure 1). The explanation is very simple: the maintenance coefficient varies in response to nutritional status and could not be taken as an absolute constant; under substrate deficiency, the cells adjust their maintenance requirements to lower values by reducing turnover rate, osmotic work, and motility (2).

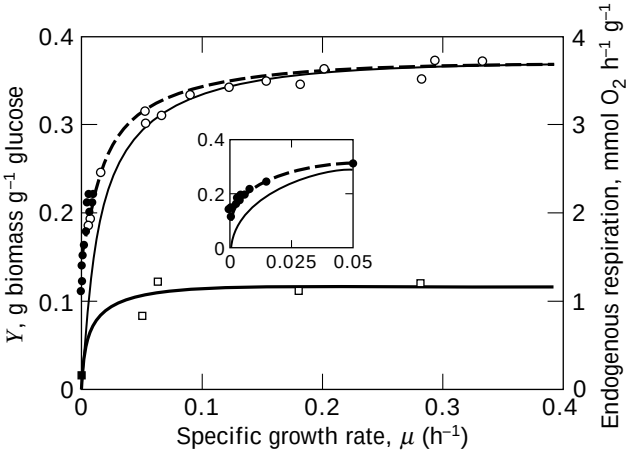


Figure 1. Variation of growth yield (*circles*) and endogenous respiration (*squares*) as dependent on specific growth rate in chemostat (*open symbols*) and continuous dialysis culture (*closed symbols*). Solid curves were calculated from the synthetic chemostat model (2). The dotted curve was derived from the Pirt-Herbert model (equations 17 to 20), which predicts quite well intensive growth but fails in the region of extremely low growth rates (see inset).

The described experimental technique is indirect because it is based on measurements of μ-dependent Y variation rather than m itself, and there are some assumptions needed to be confirmed (e.g., that m is constant and that maintenance requirements are the only reason of Y variation). However, some components of maintenance requirements are available for direct estimation. In particular, we can assess the total turnover rate of cellular material a which is one of the main components of maintenance requirements (equation 20). The principal cell constituents that are turned over are proteins, nucleic acids, and cell wall polymers. The turnover rate is very close to *endogenous respiration*, which is the oxidation of those compounds produced from the turnover (breakdown) of cellular macrocomponents. Accurate measurements of endogenous respiration need to be made under normal growing conditions. It is known that the simple removal of cells from nutrient broth by filtration with subsequent washing and incubation in buffer renders strong stress and may alter the normal turnover rate (6). To avoid artifacts, we can use a label-substitution technique (Fig. 2). The chemostat culture is fed alternately from two bottles containing unlabeled and labeled $^{14}C(U)$ substrate respectively. The $^{14}CO_2$ evolution rate is recorded after switching to unlabeled substrate, when the main source of $^{14}CO_2$ are cell components. The calculated a value was found to be rather high, accounting for the major part of total maintenance determined by the indirect method (2). The endogenous respiration declined at the low growth rate (Fig. 1), indicating that under starving conditions, self-adjustment of the maintenance requirement occurs mainly as a reduction in the turnover rate of macromolecules.

Maintenance Requirements and Wasteful Catabolism

The described concept of maintenance requirements was the subject of severe criticism (8). One of the strongest arguments against it was an apparent increase in Y^{max} observed in chemostat cultures limited by P, N, and other conserved substrates under conditions of energy excess. To preserve the constancy of the true yield, Pirt (9) had to modify equation 18 in the following way:

$$q = \mu/Y^{max} + m + m'(1 - \mu/\mu_m) \qquad (21)$$

where $m'(1 - \mu/\mu_m)$ is the second μ-dependent component of maintenance energy that operates under excess of energy substrate.

However, it is better to differentiate maintenance requirements *sensu stricto* as those more or less a minor component of the cell energy budget that is observed under energy-limitation and wasteful use of catabolic substrate under energy excess. In physiological terms, these two groups of nonproductive catabolic reactions are completely different. The first reactions are mainly responsible for compensation of turned-over macromolecules and therefore belong to the category of regular primary catabolism. The catabolic reactions of the second group include excretion into environment of partly oxidized substances (*overflow metabolism*), uncoupling of respiration from ATP generation by metabolic inhibitors, functioning of futile cycles,

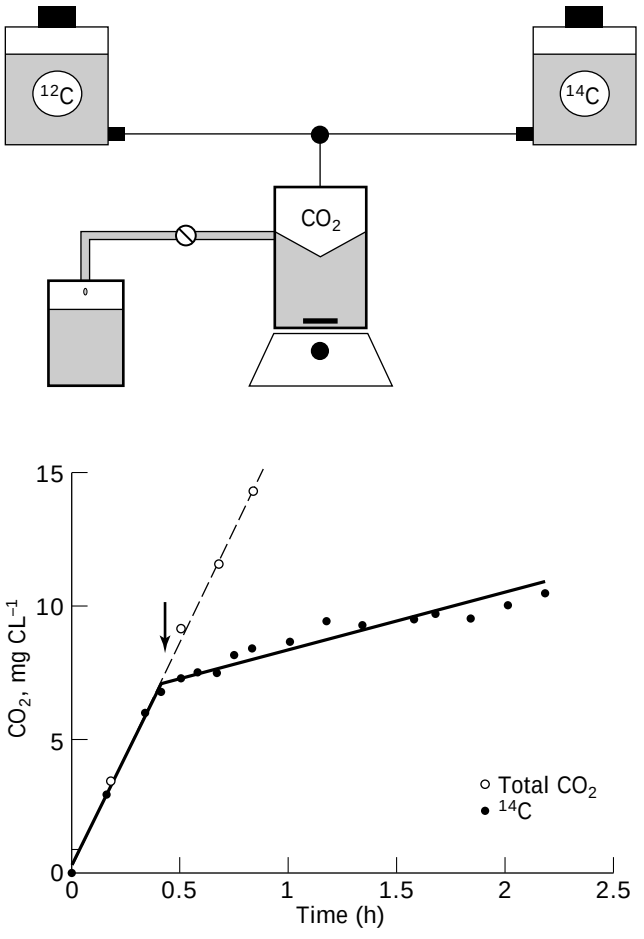


Figure 2. Label substitution technique for determination of turnover rate of cell macromolecular constituents. *Top*, experimental setup including two medium reservoirs containing ^{14}C- and ^{12}C-glucose pumped into a cultivation vessel through a two-way valve. *Bottom*, example of $^{14}CO_2$ evolution dynamics before and after (*arrow*) switching of medium feed from ^{14}C to ^{12}C-glucose, glucose-limited culture of *Pseudomonas fluorescens* 1472, D = 0.08 h^{-1} (7).

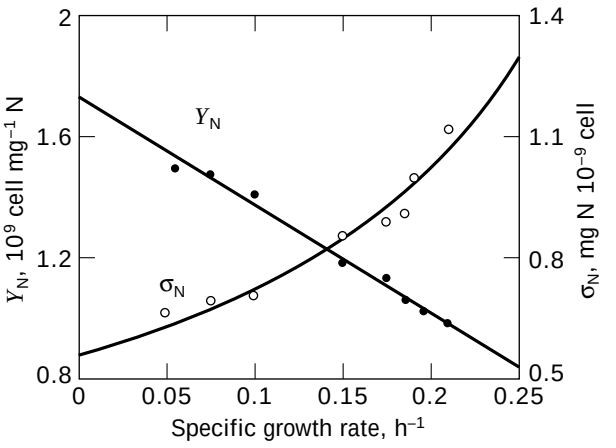


Figure 3. Relationship between stoichiometric parameters Y and s and specific growth rates of *Chlorella vulgaris* grown in chemostat culture limited by nitrogen source (10). The curves are calculated using equations 23 and 24.

or substrate oxidation through alternative oxidases without ATP generation. These and related phenomena take place in chemostat culture limited by conserved substrates (opposite to limitation by energy source) as well as during lag phase of batch culture started from starving inoculum. We will discuss the mathematical formulation of these phenomena in the section devoted to growth kinetics.

Variation in Biomass Yield from Conserved Substrates

Yield on conserved substrates varies mainly as a result of alterations in biomass chemical composition expressed by parameter σ_s, the intracellular content of deficient element or cell quota (see equation 9). For most of known cases, the content σ_s increases parallel to growth acceleration (Fig. 3). As yield and cell quota are inversely related to each other (equation 9), then Y values decrease with growth rate. The physiological mechanisms of this variation are as follows. The intensive growth requires higher *internal concentration* of some conserved limiting substrates that preserve their chemical identity after uptake (K^+, Mg^{2+}, vitamins). Other conserved substrates (sources of P, N, S, etc.) are incorporated into macromolecular cell constituents (mainly nucleic acids and proteins) whose *intracellular content* also should be kept high at high growth rate. Both types of changes in cellular composition are manifested as σ increase, and both of them require additional maintenance energy (to maintain concentration gradient or compensate turnover of macromolecules). The observed μ-dependent variation in σ is therefore a compromise between biosynthetic requirements and energy conservation that is attained because of optimal metabolic control of cell performance. However, it would be erroneous to consider μ as truly independent variable setting up chemical composition of cells. In fact, both μ and σ are functions of one common independent variable, the limiting substrate concentration in the medium, s. For steady-state chemostat culture we have:

$$\mu = \frac{q}{\sigma} = \frac{1}{\sigma}\frac{Q\tilde{s}}{K_s + \tilde{s}}$$

$$\sigma = \sigma_0 + \frac{(\sigma_m - \sigma_0)\tilde{s}}{K_\sigma + \tilde{s}} \qquad (22)$$

where μ is specific growth rate, q is specific substrate uptake rates; σ_0 and σ_m are, respectively, lower and upper limits of σ variation; low limit $\sigma \to \sigma_0$ is attained when $\tilde{s} \to 0$ and upper limits $\sigma \to \sigma_m$-when $\tilde{s} \to \infty$. By excluding $\tilde{s}$ from both these equations we arrive at following relationship between σ and μ:

$$\mu = \mu_m \frac{\sigma_m(\sigma - \sigma_0)}{\sigma[\lambda(\sigma_m - \sigma_0) + \sigma - \sigma_0]}$$

$$\lambda = \frac{K_s}{K_\sigma} \qquad (23)$$

Under realistic assumption $\lambda \approx 1$ ($K_s \approx K_\sigma$) we have

$$\sigma = \frac{\sigma_0}{1 - (1 - \sigma_0/\sigma_m)\mu/\mu_m}$$

$$Y = Y_m - (Y_m - Y_0)\frac{\mu}{\mu_m} \quad (24)$$

where Y_m and Y_0 are, respectively, upper and lower limits of yield variation ($Y_m = 1/\sigma_0$, $Y_0 = 1/\sigma_m$). As we can see, the linear relationship between Y and μ is normally observed in chemostat culture (Fig. 3).

Microscopic Approach in Studies of Growth Stoichiometry

Equations 1 to 24 exemplify the *macroscopic approach* in studying of microbial growth stoichiometry. Its typical features are the use of gross formulas for biomass and metabolic products, evaluation of total mass balance for chemical elements (C, N, P), and formal description of microbial growth as a single-step conversion of substrate(s) into biomass. By contrast, the *microscopic approach* focuses on the much more complex real metabolic reactions and attempts to account for a limited but still quite large number of individual metabolic intermediates. The final aim of this approach is to organize the biochemical information into a consistent picture of microbial metabolism at the level of entire cell.

The microscopic approach has become possible by virtue of advancements in biochemistry, which has succeeded in establishing a sufficiently full picture of metabolic processes in certain microorganisms. The pioneering work in this area was done by Bauchop and Elsden (11), who were able to sum up the balance of ATP for fermenting microorganisms. As a result, a relation was established between the biomass yield (a macroscopic quantity) and the number of generated ATP moles (a stoichiometric characteristic of real catabolic reactions):

$$Y_{ATP} = MY_E/n \quad (25)$$

where n = mol of ATP made available to the organism by the metabolism of one mole of energy source, and M = molecular weight (g) of energy source. The following example illustrates the Y_{ATP} calculation: if biomass yield of some organisms aerobically grown on glucose is 0.52 g CDW/g, then Y_E is 1.49 g CDW per g of oxidized glucose (calculated from equation 6) or $Y_E = 1.49 \times 180 = 268$ g CDW/mol (180 is glucose molecular weight); assuming that P/O = 2 (that is, 2 mol of ATP produced per atom oxygen taken up) and that 2 ATP mol are produced via glycolysis (substrate phosphorylation) we arrived at $n = 2 + 12 \times 2 = 26$ and $Y_{ATP} = 268/26 = 10.3$ g CDW/mol ATP. Careful determination of n and Y_{ATP} is possible only for anaerobic growth of fermenting microorganisms generating ATP via substrate phosphorylation. The mean value tends to be around 10.5 g CDW/mol ATP (Fig. 4). For aerobic growth, we need to make assumptions on the P/O ratio. As soon as the respiratory chain of bacteria differ widely for various organisms and growth conditions, this assumption can never be reliable. To avoid this obstacle, an interesting approach was proposed (1): microbial culture is grown in a chemostat limited by two carbon-containing energy sources, their ratio is varied while the total carbon feed rate is kept constant; yield measurements should allow one to determine both parameters (P/O and Y_{ATP}) independently by multiple linear regression.

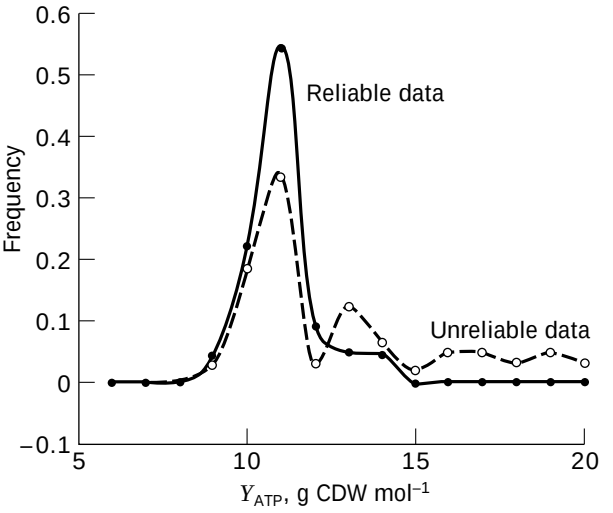


Figure 4. Frequency distribution of experimentally measured values of Y_{ATP} at different degrees of creditability. The reliable data refer to studies of anaerobic growth with direct measurements of fermentation products (plotted from data base in Ref. 12. Note that these data are normally distributed with mean value 10.55, whereas all data display considerable skewness.

Today, microstoichiometry is quickly progressing as so-called metabolic balancing. Cell growth is viewed as a set of transport and intracellular metabolic reactions known for some particular organisms. As a rule, the produced metabolic networks are composed of a combination of true stoichiometric equations for individual metabolites and empirical gross equations (Table 2). The amount of such equations vary in different models from 20 to 30 to more than 100. For example, van Gulik and Heijnen (13) describe yeast growth by a set of more than 90 reactions including glycolysis and the citric acid cycle (14); PEP phosphotransferase; pentose phosphate pathway (6); glyoxylate shunt (2); oxidative phosphorylation (4); CO_2 interaction with THF (3); transport of inorganic P, NH_4^+, SO_4^{2-}, acetate, lactate, pyruvate, glucose, gluconate, succinate, and citrate (totally 10 transport reactions); amino acid synthesis (15) and polymerization (2); nucleotide synthesis (9); RNA synthesis; ATP consumption for maintenance; fatty acids synthesis (2); formation of glycogen and polysaccharides; and finally, the biomass formation from proteins, polysaccharides, RNA, fatty acids, and glycerol.

For each compound, i, involved in a metabolic system, a mass balance can be defined:

$$\frac{dC_i}{dt} = r_{Ai} + \Phi_i \quad (26)$$

where C_i is concentration of ith compound, r_{Ai}, and Φ_i denote the net rates of, respectively, i chemical conversion and transport over the boundaries of bioreactor (fluxes of

Table 2. Metabolic Networks

Reaction	Equation	Stoichiometric Equation	Empirical Gross Equation
Glycolysis reaction	Glucose + ATP → glucose-6-P + ADP + H	x	
	Glucose-6-P → fructose-6-P	x	
	0.5 Fructose-6-P + 0.5 ATP → glyceraldehyde-3-P + 0.5 ADP + 0.5 H	x	
Oxidative phosphorylation	NADH + $0.5O_2$ + d_1ADP + d_2Pi + (1 + d_1) H → (1 + d_1)H_2O + NAD + d_1ATP	x	
Biomass formation	a_1Proteins + $a_2$polysaccharides + a_3RNA + $a_4$lipids → biomass		x

CO_2, O_2, nutrients, cells, and products). Most metabolic balancing equations are applied to steady-state growth (which means that no intracellular accumulation of metabolites occurs). In such cases, the differential equations like equation 26 are reduced to linear algebraic ones. Besides, an extensive use of matrix calculus is customarily made to obtain a concise notation. The problem of experimental support of such model is especially important (13). The degree of freedom, *df*, of the resulting system of linear equations is equal to the total number of unknown rates (both intracellular and exchange reactions) minus the total number of linear equations. To resolve the system, *df* rates have to be measured, and then the system is fully *determined*. If the number of measured rates is greater than *df*, then the system is *overdetermined*, and the redundancy of the data can be used for statistical analysis and error minimization. However, it is much more typical to have an *underdetermined* system when the sum of measured rates is less than *df*. In this case, the number of possible solutions is infinite unless additional constraints are applied (e.g., maximization of biomass yield, minimization of energy expenditure) to find the one and only one solution by the linear optimization technique.

In most studies, flux estimates are obtained using measurements of substrate consumption and product formation. This approach has proved to be efficient in some particular biotechnological cases, such as when only specific pathways need to be considered (16) or if the contribution of flux for cellular growth is weak, as with mammalian or hybridoma cells (17). The more complex microbial systems are turned out to be seriously underdetermined. In such cases, the application of metabolic balancing requires the use of one or another maneuvers: (*1*) to lump together several sets of reactions (18); (*2*) to utilize data from in vitro enzyme assays; (*3*) to make assumptions on numeric values of some stoichiometric growth parameters, such as Y_{ATP}, P/O, and H^+/e ratios, which are the subject of controversial debates (13).

However, the best solution would be to get direct experimental data on in vivo flux and resolve the system. Isotopic tracers are one of the best candidates for such a purpose. We will illustrate this point by describing a recently published work (19). This novel approach is based on the analytical power of ^{1}H-detected ^{13}C nuclear magnetic resonance. *Corynebacterium glutamicum* was grown in chemostat culture continuously fed with [1-^{13}C]-glucose; when steady state was established, the cells were harvested and hydrolyzed and the amino acids were separated by ion-exchange chromatography and analyzed by NMR spectroscopy. NMR provides data on ^{13}C enrichment at each specified carbon position of amino acid. Because metabolic pathways for amino acid synthesis are exactly defined, then the entire central metabolism can be assessed for in vivo fluxes, including determination of the forward and back rates of bidirectional reactions. In *C. glutamicum*, the flux through the pentose phosphate pathway turned out to be 66.4% (relative to glucose input flux 1.49 mmol g^{-1} CDW h^{-1}); the entry into tricarboxylic acid cycle, 62.2%, and the contribution of the succinylase pathway to lysine synthesis, 13.7%. The total net flux of the anaplerotic reactions (carboxylation of PEP/pyruvate into oxaloacetate/malate) was quantitated as 38%, the true forward flux of C3 → C4 being 68.6% (1.8 times of 38%) and a back flux of C4 → C3 being 30.6% (0.8 times of 38%) (19). The metabolic balancing proved to be very promising and useful to identify metabolic constraints for intensive synthesis (overproduction) of products such as amino acids. On the other hand, this approach still is restricted to steady-state and balanced growth and is not able to cope with complex dynamic behavior of microorganisms (transient growth, changes in biomass composition).

BASIC PRINCIPLES OF GROWTH KINETICS

Kinetics of Chemical and Enzyme Reactions

We need to introduce some basic principles of kinetic analysis of chemical and enzymatic reactions. Quantitative description and understanding of microbial growth dynamics and kinetics are impossible without some elementary knowledge in underlying scientific disciplines. Enzymatic and chemical reactions play an essential role in biotechnology, which is one of the most important fields in industrial development.

Order and Molecularity of Chemical Reactions. The molecularity of any chemical reaction is defined by the number of molecules that are altered in the reaction (Table 3). The order is a description of the number of concentration terms multipled together in the rate equation (Table 4). Hence, in a first-order reaction, the rate is proportional to one concentration of reactant; in a second-order reaction, it is proportional to two concentrations or to the square of one concentration. For a simple single-step reaction, the order is generally the same as the molecularity. For a complex reactions involving a sequence of unimolecular and bimolecular steps, the molecularity is not the same as its order. Reactions of molecularity greater than 2 are com-

Table 3. Molecularity of Chemical Reaction

Molecularity of reaction	Reactants		Product
Unimolecular or monomolecular	S	→	P
Bimolecular	S + S	→	P
or	$S_1 + S_2$	→	P
Trimolecular or termolecular	S + S + S	→	P
or	$S_1 + S_1 + S_2$	→	P
or	$S_1 + S_2 + S_3$	→	P

mon, but reactions of order greater than 2 are very rare. For instance, a trimolecular reaction, such as A + B + C → P, as a rule proceeds through two elementary steps, A + B → X and X + C → P, each of which are of the second or first order. Very often, bimolecular reactions between S_1 and S_2 occur under the condition that their respective concentrations $s_2 \gg s_1$ (e.g., if the second reactant S_2 is solvent), then we have a *pseudo first-order* reaction. Some reactions are observed to be of the zero order, that is, the rate appears to be constant, independent of the concentration of reactant. This is a characteristic feature of catalyzed reactions and occurs if reactant is present in such large excess that the full potential of catalyst is realized.

Dimensions of Rate Constants. Knowledge of dimensions is very useful to check the correctness of derived kinetic equations: the left- and right-hand sides of an equation must always have the same dimensions. This general rule is applicable to all mathematical models (not only in chemical kinetics). It is incorrect to add or subtract terms of different dimensions, although you may multiply or divide them. For example, if expression " . . . $(1 + s)$" occurs in an equation, where s has dimension (concentration), then either equation is incorrect, or the 1 is a concentration that happens to have a numerical value of 1 unit. The operation rising to power is allowed for only simple dimensionless numbers, for example, expression $e^{2.5t}$, where t is time, is correct only if 2.5 has dimension $(\text{time})^{-1}$. The comparison of velocities of two reactions does make any sense only for kinetic terms of the same dimension. If the kinetic order of inspected reactions is different, then we have to equalize the respective rate of expressions; for example, the second-order rate constant k $(\text{time})^{-1}$ $(\text{concentration})^{-1}$ should be multiplied by instant concentration of reactant s to be compared with the first-order rate constant having dimension $(\text{time})^{-1}$. The dimensions of the zero-, first-, and second-order rate constants are shown in Table 4.

Reaction Dynamics. If rate constant and so-called initial conditions (concentrations of reactants at zero-time [s_{01}, s_{02}, . . .]) are known, then it is possible to calculate the time course of reactions either in terms of dynamics of residual reactant concentration, $s(t)$ or product accumulation, $p(t)$. For this purpose, we have to integrate a differential equation under specified initial conditions (see results of integration in Table 4). The dynamics of $s(t)$ are linear in the case of a zero-order reaction and hyperbolic in the case of the first and second order. The difference between the last two dynamic curves can be made visible with a semilogarithmic plot of log(concentration) versus time; it should became linear for the first-order reaction and remain to be curvilinear for the reaction of higher order.

Reaction Half-Time. The reaction half-time ($t_{0.5}$) is a very popular kinetic parameter, especially among biologists. It is easily calculated from integral equations by putting $s = s_0/2$ when $t = t_{0.5}$. A unique feature of the first-order reaction is the constancy of $t_{0.5}$ independently of the initial reactant concentration s_0. However, the half-time of other reactions does depend on s_0; it increases for zero-order reactions and decreases for the second-order reactions with increase in s_0. Thus, it is not recommended to use half-time as a parameter or estimator for reactions other than first-order reactions.

Determination of the Order of Reaction and Numeric Values of Kinetic Constants. If the reaction has an order n and rate constant k, then the reaction rate v and reactant concentration s are related by the equation

Table 4. Kinetic Order of Chemical Reactions

Reaction order	Differential equation	Dimension of rate constant	Dynamics of residual reactant	Dynamics of product accumulation	Reaction half-time[a]
Zero	$\frac{ds}{dt} = -k$	$(\text{conc.})(\text{time})^{-1}$	$s(t) = s_0 - kt$	$p(t) = p_0 + kt$	$t_{0.5} = \frac{s_0}{2k}$
First	$\frac{ds}{dt} = -ks$	$(\text{time})^{-1}$	$s(t) = s_0 \exp(-kt)$	$p(t) = s_0[1 - \exp(-kt)]$	$t_{0.5} = \frac{\ln 2}{k}$
Second	$\frac{ds}{dt} = -ks^2$	$(\text{conc.})^{-1}(\text{time})^{-1}$	$s(t) = \frac{s_0}{1 + ks_0t}$	$p(t) = \frac{s_0^2kt}{2(1 + s_0kt)}$	$t_{0.5} = \frac{1}{ks_0}$ [b]
Second	$\frac{dp}{dt} = -ks_1s_2$	$(\text{conc.})^{-1}(\text{time})^{-1}$	$s_1(t) = s_{01} - p(t)$ $s_2(t) = s_{02} - p(t)$	$p(t) = \frac{s_{01}s_{02}(1 - \exp[(s_{02} - s_{01})kt])}{s_{01} - s_{02}\exp[(s_{02} - s_{01})kt]}$	$t'_{0.5} = 1/ks_{01}$ $t''_{0.5} = 1/ks_{02}$
Pseudo-first $s_{01} \gg s_{02}$	$\frac{dp}{dt} = -ks_1s_2$	$(\text{conc.})^{-1}(\text{time})^{-1}$	$s_2(t) = s_{02}\exp(-ks_{01}t)$ $s_1(t) \approx s_{01}$	$p = s_{02}[1 - \exp(-ks_{01}t)]$	

Note: t = time; s = reactant concentration; p = product concentration; s_0 = reactant concentration at $t = 0$.

[a]The half-time of the reaction is the time required for half-completion.

[b]The half-time is defined for the reagent that has lower initial concentration and is depleted first.

$$v = -ds/dt = dp/dt = ks^n \quad (27)$$

The simplest way to find both unknown values (n and k) would be to measure reaction rate v at several concentrations of reactant. Then a plot of log(rate) against log(concentration) gives a straight line with a slope equal to n and intercept equal to log(k): $\log v = \log k + n \log s$. If there are several reactants, then it is useful to know the order in respect to each one. For this purpose, you need to have several experiments with variation of each reactant concentration while keeping the other concentrations constant.

The most frequent goal is the determination of the first-order rate constants, first because many reactions do obey first-order conditions in respect to each reactant and second because it is possible to carry out many reaction under pseudo first-order conditions. There are some special methods to determine the values of the first-order rate constants from the experimental curves of product formation or substrate depletion dynamics. Some of them were designed to improve the accuracy of k determination when the initial reactant concentration (s_0) or final value of product concentration (p_∞) were not known (methods of Guggenheim, Kezdy-Swinbourne, and others, see details in Ref. 20). All are based on plotting the experimental points in a some sophisticated manner to convert the original curves to a straight lines. Today, these methods are replaced by computer-aided nonlinear regression, which is much more convenient and precise and, contrary to graphic methods, allows for more rigorous estimate of confidence limits of measured quantities.

Derivation of Basic Kinetic Equations for Enzymatic Reactions. Contrary to simple chemical reactions, enzyme-catalyzed reactions proceed through *reversible* formation of the dynamic *enzyme-substrate complex* (ESC). The word *reversible* is essential because the ESC can be decomposed into free enzyme E and product P or dissociate back to E and substrate S. There are many ways to simulate mathematically the mechanism of the enzymatic reaction. We will consider here equilibrium, steady-state, and general non-steady-state approaches.

Equilibrium Approach. This approach was used by Michaelis and Menten (21) to describe the effect of sucrose concentration on invertase activity. They assumed that the first step of ESC formation is so rapid that could be represented by an *equilibrium constant* K_s:

$$E \overset{e}{+} S \overset{s}{\leftrightarrow} ES \overset{K_s}{\rightarrow} E \overset{x}{+} P \overset{k_2}{\rightarrow} \overset{p}{} \quad (28)$$

where e, s, x, and p denote the concentrations of free enzyme, substrate, FSC, and product, respectively. The equilibrium constant, K_s, is defined as $K_s = es/x$. The instantaneous concentrations s and e are not directly measurable, but they could be expressed in terms of the initial, measured concentrations, e_0 and s_0, using mass-balance relationships:

$$e_0 = e + x \quad s_0 = s + x$$
$$x \ll s$$
$$s \approx s_0 \quad (29)$$

The overall reaction rate, v, is a simple first-order reaction with rate constant k_2:

$$v = dp/dt = -ds/dt = k_2 x \quad (30)$$

The x value can be found from expression for K_s and mass-balance conditions (equation 29): $x = e_0 s/(K_s + s)$. Substitution of this expression into equation 30 finally produces

$$v = \frac{k_2 e_0 s}{K_s + s} \quad (31)$$

Steady-State Approximation. This approach was applied by Briggs and Haldane (22) for the following scheme of enzymatic reaction

$$E \overset{e}{+} S \underset{k_{-1}}{\overset{s \;\; k_1}{\rightleftarrows}} ES \overset{x \;\; k_2}{\rightarrow} E \overset{p}{+} P \quad (32)$$

This scheme implies the reversibility of ESC formation instead of much more restrictive equilibrium postulate. However, still there was steady-state assumption in respect to ESC formation:

$$dx/dt = k_{+1}(e_0 - x)s - k_{-1}x - k_{+2}x \approx 0 \quad (33)$$

Therefore, $x = k_{+1}e_0 s/(k_{+1}s + k_{-1} + k_{+2})$, and substitution of this expression into equation 30 gives

$$v = -\frac{ds}{dt} = k_{+2}x = \frac{k_{+1}k_{+2}e_0 s}{k_{+1}s + k_{-1} + k_{+2}}$$
$$= \frac{k_{+2}e_0 s}{\frac{k_{-1} + k_{+2}}{k_{+1}} + s} = \frac{Vs}{K_m + s} \quad (34)$$

Equation 34 is what usually called the *Michaelis-Menten equation*, the fundamental equation of enzymatic kinetics. It contains two parameters, K_m, the *Michaelis constant*, and V, the *maximum velocity*. V is the rate of reaction that would occur under full substrate saturation of an enzyme's active sites ($s \gg K_m$). In reality, the V value can be never attained, because extremely high substrate concentrations inhibit enzymes (see later text). It is clear also that V is not a fundamental property of enzyme because it depends on e_0, the enzyme concentration. More advantageous as a specific enzyme characteristic is the *catalytic constant* or *turnover number*, k_{cat}, which is V/e_0. For equation 32, k_{cat} is identical with k_{+2}, but in general the more noncommittal notation k_{cat} is preferable (e.g., k_{cat} may differ from k_{+2} if the product formation from ESC is a reversible reaction). Numerically, the Michaelis constant, K_m, is the substrate concentration that provides half the maximal reaction rate. Contrary to this simple practical definition, the mechanistic interpretation of K_m is not so lucid. Sometimes

K_m is interpreted as a substrate binding constant, K_s, assuming that $k_{+2} \ll k_{-1}$. This is a very dubious assumption. There are very few enzymes for which the individual values of k_{+2} and k_{-1} are known. Ironically, the best-studied examples (e.g., horseradish peroxidase) present just opposite case: $k_{+2} \gg k_{-1}$. It is important that many specific mechanisms (not necessarily equations 28 and 32) generate the same steady-state rate equation 34. However, the particular expression for K_m should be different for each individual case.

Although undefined in mechanistic terms, the parameters K_m and V are very useful at the first steps of kinetic studies:

1. The use of equation 34 allows the expression of the complex effects with simpler terms.
2. K_m and V are helpful as predictive parameters to design a valid enzyme assay. One practical recommendation is to keep substrate concentration in incubation mixture at the level of 10 K_m or higher.
3. Equation 34 permits one to obtain at least rough estimate of in vivo enzyme activity provided the internal substrate concentrations are known.

General Non-Steady-State Approach. Equation 32 contains four variables (s, p, e, and x) constrained by two mass-balance equations ($e_0 = e + x$ and $s_0 = s + p + x$). Therefore, it is enough to integrate just two differential equations (e.g., equations 33 and 34) to characterize the whole system. A typical example of a numeric solution is shown in Figure 5. We can see that enzyme-catalyzed reaction proceeds through three definite phases well separated on the time scale. Each phase is now safe to analyze with simpler mathematical expressions because any assumptions could be tested against the full exact solution.

1. The first transient phase occurs before the steady-state concentration of ESC is reached. It occupies less than 10^{-3} s for most of enzymes. During this phase, the substrate concentration remains fairly constant ($s \approx s_0$), allowing for the following analytical solution:

$$x(t) = \frac{k_{+1}e_0s_0}{A}[1 - e^{-At}], \quad v(t) = \frac{Vs}{K_m + c}[1 - e^{-At}]$$
$$A = k_{+1}s_0 + k_{-1} + k_{+2} \tag{35}$$

As compared with equation 34, it contains a relaxation term $[1 - \exp(-At)]$ that is very large when t is small, but decays to zero as t increases above $\tau = 1/A$. Accordingly, the rate of enzymatic reaction is initially zero but increases rapidly to the steady-state value as the exponential term decays. Because the relaxation term contains s_0, then the experimentally observed delay depends on substrate concentration. It allows for direct determination of individual rate constants k_{+1} and $(k_{-1} + k_{+2})$ by techniques such as in stopped-flow apparatus (23).

2. The second steady-state phase is characterized by the constancy of reaction rate due to the exact balance between the rates of ESC formation and breakdown ($dx/dt = 0$) while the substrate concentration remains close to the initial value s_0. This phase proceeds for at least several seconds. As a rule, it is enough to measure the *initial velocity* of enzymatic reactions unaffected by substrate depletion or product accumulation. Steady-state kinetics is the most popular research domain; it is the most accessible and developed and provides the most kinetic data. However, it fails to determine individual rate constants as fully as the transient and relaxation approaches do. Steady-state equations were derived earlier (equation 34).

3. The third phase is characterized by considerable changes in the concentrations of substrate and product. Thus, we can no longer assume that $s = s_0$, and we should integrate equations 32 and 33. However, very good precision provides the quasi-steady-state approximation $ds/dt \gg dx/dt \approx 0$. Then, we arrive at a relatively simple differential equation that is solved analytically:

$$\frac{ds}{dt} = -\frac{Vs(t)}{K_m + s(t)} \tag{36}$$

For initial conditions $s = s_0$, $p = 0$, $t = 0$, we have

$$K_m \ln\frac{s_0}{s} + s_0 - s = Vt \tag{37}$$

Equation 37 is called the *integrated Michaelis-Menten equation*. It remains to be valid not only for initial rate measurements but for any point within the reaction progress curve.

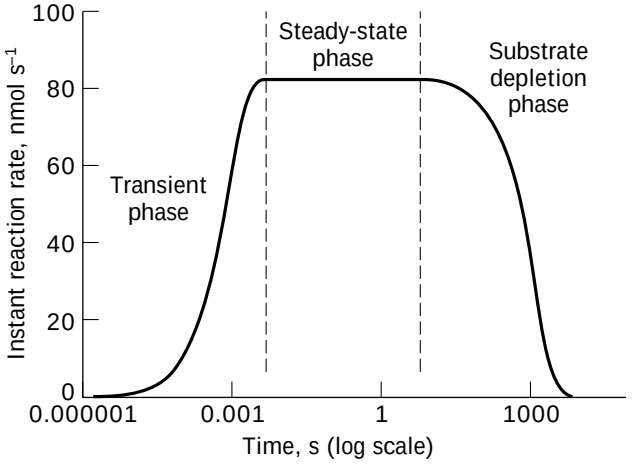


Figure 5. Time course of reaction proceeding by the Michaelis-Menten mechanism. Numeric integration of equations 33 and 34: $s_0 = 10^{-4}$M, $e_0 = 10^{-8}$M, $k_{+1} = 10^6$M^{-1}s^{-1}, $k_{-1} = 10^3$ s^{-1}, $k_{+2} = 10^2$ s^{-1}.

Experimental Determination of Kinetic Parameters of the Michaelis-Menten Equation. Until recently, most enzyme kinetic experiments have been analyzed by means of one of linear plots in Table 5. Linear plots are used to examine an agreement of experimental data with equation 34 as well as to determine numeric values of parameters V and K_m from slopes and intercepts. Today, this graphic ap-

Table 5. Linear Plots

$\frac{1}{v} = \frac{1}{V} + \frac{K_m}{V}\frac{1}{s}$	Lineweaver-Burk or double-reciprocal plot
$\frac{s}{v} = \frac{K_m}{V} + \frac{s}{V}$	Langmuir-Hanes plot
$v = V - K_m\frac{v}{s}$	Eadie-Hofstee plot
$V = v + \frac{v}{s}K_m$	Direct linear plot (20)

proach seems to be too cumbersome as compared with much more efficient computer routines. The main objection is also that any linear transformation introduces some statistical bias (ironically, the highest bias is attributed to the most popular double-reciprocal plot!). In addition, manual line drawing is very arbitrary, so obtained parameters could not be assessed statistically for confidence limits. However, plotting of original or linearized data is useful as an illustration and the first sketchy estimate of enzymatic parameters. For definitive work, it is advisable to avoid all plots and to use statistical analyses instead.

Reversible Enzymatic Reactions. The majority of biochemical reactions are reversible. To account for this feature, the Michaelis-Menten mechanism can be modified as follows:

$$\underset{e_0-x-y}{E} + \underset{s}{S} \underset{k_{-2}}{\overset{k_{+1}}{\rightleftarrows}} \underset{x}{ES} \underset{k_{-2}}{\overset{k_{+2}}{\rightleftarrows}} \underset{y}{EP} \underset{k_{-3}}{\overset{k_{+3}}{\rightleftarrows}} \underset{e_0-x-y}{E} + \underset{p}{P} \qquad (38)$$

Contrary to the basic scheme in equation 32, there are two intermediates, one of which is the normal ESC and another is the *enzyme-product complex* (EP). The substrate, *S*, and product, *P*, can interconvert to each other. Application of the steady-state approach to this scheme results in the following equation:

$$\begin{aligned} v &= \frac{V^f s/K_m^s + V^r p/K_m^p}{1 + s/K_m^s + p/K_m^p} \\ V^f &= \frac{k_{+2}k_{+3}e_0}{k_{-2} + k_{+2} + k_{+3}} \\ K_m^s &= \frac{k_{-1}k_{-2} + k_{-1}k_{+3} + k_{+2}k_{+3}}{k_{+1}(k_{-2} + k_{+2} + k_{+3})} \\ V^r &= \frac{k_{-1}k_{-2}e_0}{k_{-1} + k_{-2} + k_{+2}} \\ K_m^p &= \frac{k_{-1}k_{-2} + k_{-1}k_{+3} + k_{+2}k_{+3}}{k_{-3}(k_{-1} + k_{-2} + k_{+2})} \end{aligned} \qquad (39)$$

where superscripts *f* and *s* denotes parameters of forward reaction, and *r* and *p* are indicators of reverse reaction.

When a reaction is at equilibrium, the net velocity must be zero and, consequently, if s_∞ and p_∞ are the equilibrium values of *s* and *p*, it follows from equation 39 that equilibrium constant *K* is expressed via kinetic parameters (the *Haldane relationship*):

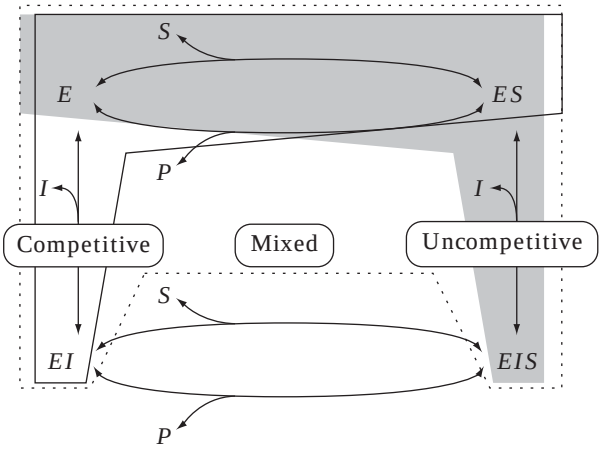
$$K = p_\infty/s_\infty = V^f K_m^p/V^r K_m^s \qquad (40)$$

Equation 39 implies that the rate must decrease as the product accumulates, even if the decrease in substrate concentration is negligible. Thus, reversibility is closely associated with and requires the *product inhibition*. In many essentially irreversible reactions (e.g., invertase-catalyzed hydrolysis of sucrose), product inhibition is also significant. It can be explained by equation 38 with only the second step being irreversible. In such a case, the accumulation of product causes the enzyme to be sequestered because the EP complex and rate equation are as follows (compare with equation 39):

$$v = \frac{V^f s/K_m^s}{1 + s/K_m^s + pK_m^p} = \frac{V^f s}{K_m^s(1 + pK_m^p) + s} \qquad (41)$$

Inhibitors and Activators. Compounds that reduce the rate of enzyme-catalyzed reactions when they are added to the reaction mixture are called *inhibitors*. Just opposite effects (increase of reaction rate) are caused by *activators*. Both of them belong to the category of *modifiers*. We will concentrate here on inhibitors as having more practical application and even more specifically on *reversible* inhibitors, which form various dynamic complexes with enzymes. The *irreversible* inhibitors are known as catalytic poisons (many heavy metals, such as mercury); their binding to enzyme molecules reduces activity to zero, leaving no room for quantitative analysis.

The reversible inhibitors can form complexes with free enzymes or enzyme-substrate complexes as shown in the scheme of Botts and Morales (20):



There are three major simple (or linear) inhibition mechanisms, which can be generated from this scheme by omitting some of the six involved:

- *Competitive* inhibition if *EIS* is missing (inhibitor binds to the same site on the enzyme as the substrate forming abortive nonproductive complex *EI*).
- *Uncompetitive* if *EI* is missing and *EIS* occurs as a dead-end complex; it implies that the inhibitor-binding site becomes available only after the substrate has bound.

- *Mixed* if *EI* and *EIS* both occur but are not interconvertible (the complex *EIS* does not break down to products; this situation frequently occurs when the inhibitor is reaction product). The particular case of mixed inhibition is *pure noncompetitive inhibition*, which takes place if two inhibitor dissociation constants (for *EI* and *EIS*) are exactly matched to each other.

For all types of inhibition, the Michaelis-Menten equation remains valid: under constant inhibitor concentration, *i*, the *v-s* relationship is of the same hyperbolic type as predicted by equation 34, the only difference is that apparent values of K_m and V are now more or less simple functions of *i* (Table 6).

Some parameters (in boldface in Table 6) are not affected by inhibitors, whereas others are changed: competitive inhibitors increase K_m, pure noncompetitive inhibitors reduce V, uncompetitive inhibitors decrease at the same degree both V and K_m, leaving first-order rate constant V/K_m to be unchanged. In mixed inhibition, there are no unchanged parameters.

One interesting case is so-called substrate inhibition. It occurs when two substrates are bound to the same active site on the enzyme, forming nonproductive triple complex ES_2:

$$E + S \overset{K_{ss}}{\leftrightarrow} ES \overset{k_2}{\rightarrow} E + P$$
$$\qquad\qquad K_{SS} \uparrow\downarrow$$
$$\qquad\qquad\quad ES_2$$

then, the enzyme rate (24) is:

$$v = \frac{Vs}{K_s + s + s^2/K_{ss}} \tag{42}$$

where K_{ss} is the dissociation constant for complex ES_2.

Cooperativity. Many enzymes respond to changes in metabolite concentrations (substrates, modifiers) with much higher sensitivity as compared with predictions from the classic hyperbolic equations 34 to 42. This property is generally known as *cooperativity*, because it is thought to arise from cooperation between the active sites of the polymeric enzymes. As a rule, such enzymes consist of several *subunits* and display so-called sigmoid or S-shaped dependence of rate on substrate concentration. Many cooperative enzymes (but not all!) have *active sites* binding substrate and *allosteric sites* binding effectors. There are *homotropic* and *heterotropic* cooperative effects caused by interactions between, respectively, identical and different ligands (e.g., substrate and an allosteric effector).

To explain the cooperativity and associated sigmoid kinetics, a number of models have been suggested. The earliest one is the Hill equation, which was originally designed to describe the S-shaped curve of oxygen binding to hemoglobin. It was assumed that each protein molecule E binds n molecules of ligands S in a single step, an amount of other possible forms (ES_{n-1}, ES_{n-2}, . . . ES) being negligible:

$$E + nS \overset{K_h}{\leftrightarrow} ES_n$$

where K_h is the respective equilibrium constant (Hill and colleagues described equilibrium in terms of association constant, but for the sake of uniformity we will adhere to the previous formalisms, keeping in mind the dissociation constant, $K_h = [E][S]^n/[ES_n]$). The fractional saturation of protein (enzyme or hemoglobin). Θ is given as

$$\Theta = \frac{\text{Number of occupied binding sites}}{\text{Total number of binding sites}} = \frac{[ES_n]}{[E] + [ES_n]} = \frac{[S]^n}{K_h + [S]^n} \tag{43}$$

Equation 43 can be rearranged into

$$\frac{\Theta}{1 - \Theta} = \frac{[S]^n}{K_h}, \log\left(\frac{\Theta}{1 - \Theta}\right) = -\log K_h + n \log[S] \tag{44}$$

A plot of $\log \Theta/(1 - \Theta)$ against $\log [S]$ is known as the *Hill plot* and should be a straight line of slope n (*so*). This equation is used to fit the experimental binding and kinetic data displaying a sigmoid shape. When plotting kinetic measurements, it is assumed that $\Theta = v/V$, and maximum velocity V should be known from independent measurements taken at saturation substrate concentration. However, the results of fitting should be interpreted with care. First, the Hill equation is empirical and generally provides good agreement only in the Θ range 0.1 to 0.9 (the discrepancy at extreme s is probably caused by neglecting of other forms of ESC apart from ES_n). Second, parameter n (*Hill coefficient*) could not be interpreted as the number of subunits in the fully associated protein, rather it is an index of cooperativity.

Monod et al. (25) proposed a general model explaining cooperativity and allosteric phenomena within a simple set of postulates:

1. Each subunit of enzyme can exist in two different conformations, designated *R* (relaxed) and *T* (tense).

Table 6. Types of Inhibition

Type of inhibition	V^{app}	V^{app}/K_m^{app}	k_m^{app}
Competitive	$\boldsymbol{V}$	$(V/K_m)/(1 + i/K_i)^a$	$K_m(1 + i/K_i)$
Uncompetitive	$V/(1 + i/K_i')$	$\boldsymbol{V/K_m}$	$K_m(1 + i/K_i'))$
Mixed	$V/(1 + i/K_i')$	$(V/K_m)/(1 + i/K_i)$	$K_m(1 + i/K_i)/(1 + i/K_i')$
Pure noncompetitive	$V/(1 + i/K_i')$	$(V/K_m)/(1 + i/K_i)$	$\boldsymbol{K_m}$

2. All subunits of a molecule must occupy the same conformation at any time (e.g., for tetrameric enzyme only R_4 or T_4 are permitted, not R_3T or R_2T_2, etc.).
3. The two states are in equilibrium with constant $L = [R_4]/[T_4]$.
4. The affinity of ligand to subunit depends on the conformation state: $K_R = [R][S]/[RS]$, $K_T = [T][S]/[TS]$, $c = K_R/K_T$.

The general equilibrium solution of this scheme is rather bulky, so we demonstrate one particular case of tetrameric protein, if $c = 0$ (i.e., $K_R \ll K_T$, S binds only to the R state), then we have

$$\Theta = \frac{(1 + s/K_R)^3 s/K_R}{L + (1 + s/K_R)^4} \tag{45}$$

At high s when $s/K_R \gg L$, we can neglect the term L in the denominator, and the entire expression is converted to no more than the simple hyperbolic Langmuir isotherm. At low s, the contribution of L is considerable, so the saturation curve rises slowly from the origin displaying the S-shape.

There are other models describing the mechanisms of cooperativity and allosteric effects (20), for example, the sequential model of Koshland and colleagues and the association–dissociation model of Freiden.

Effects of pH. Every enzyme contains a large number of acidic and basic groups. Some of them are either fully deprotonated (aspartate, glutamate) or fully protonated (arginine, lysine). However several groups with pK_a 5 to 9 (imidazole group of histidine, sulphydryl group of cysteine) do change their ionization state when pH is varied. Assume as a first approximation that enzyme is represented as a dibasic acid H_2E and only a singly ionized complex, HES^-, is able to react to give products:

$$\begin{array}{ccccccc} & & H_2E & & H_2ES & & \\ & & \Updownarrow K_1^E & & \Updownarrow K_1^{ES} & & \\ S & + & HE^- & \underset{k_{+1}}{\rightleftarrows} & HES^- & \xrightarrow{k_{+2}} & HE^- + P \\ & & \Updownarrow K_2^E & & \Updownarrow K_2^{ES} & & \\ & & E^{2-} & & ES^{2-} & & \end{array}$$

Then the reaction rate is dependent on H^+ concentration, h, as follows (20):

$$\begin{aligned} v &= \frac{Vs}{K_m + s} \\ V &= \frac{\tilde{V}}{\left(\frac{h}{K_1^{ES}} + 1 + \frac{K_2^{ES}}{h}\right)} \\ \frac{V}{K_m} &= \frac{\tilde{V}/\tilde{K}_m}{\left(\frac{h}{K_1^E} + 1 + \frac{K_2^E}{h}\right)} \end{aligned} \tag{46}$$

where $\tilde{V}$ and $\tilde{K}_m$ are the pH-corrected constants. It is clear that pH-dependent variation of V reflects the ionization of ESC, whereas V/K_m reflects the ionization of the free enzyme. The pH effects on K_m are more complicated, being affected by both. Generally, the dependence of enzyme activity on pH is described by equation 43 as a bell-shaped curve with maximum at $1/2(pK_1 + pK_2)$.

Effects of Temperature. In chemical kinetics, the dependence of reaction rate on temperature is explained by the *transition-state theory* developed by Eyring in 1930 to 1935. It is based on the use of thermodynamics and principles of quantum mechanics. The reaction proceeds through a continuum of energy states and must surpass the state of maximum energy, when transient activated complex is formed. Then the dependence of reaction rate constant k on absolute temperature, T, is expressed as follows:

$$\frac{d \ln k}{dT} = \frac{\Delta H^* + RT}{RT^2} \tag{47}$$

where R is the gas constant and ΔH^* is the enthalpy of activated complex formation. The classic *Arrhenius* equation may be obtained from equation 47 under a simplified condition $\Delta H^* + RT \approx \Delta H = E_a$ (where E_a is *activation energy*). Most often, the Arrhenius equation is used in its integrated form:

$$\ln k = \ln A - E_a/RT$$

or

$$k = A \exp(-E_a/RT) \tag{48}$$

where A is the integration constant, interpreted as the frequency of collisions of reacting molecules. Apart from mechanistic derivations, there are a number of empirical expressions relating k and Celsius temperature T_c. The most popular is the exponential formula:

$$\ln k = \ln A + \alpha T_c$$

or

$$k = A \exp(\alpha T_c) \tag{49}$$

where α is the empirical constant related to the widely used temperature coefficient $Q_{10} = \exp(10 * \alpha)$.

All presented mathematical expressions predict exponential or almost exponential increases of chemical reaction rates with temperature. However, enzymatic reactions deviate from this relationship at high temperature because of thermal denaturation of enzymes. Assume that denaturation is reversible with equilibrium constant $K_T = [E']/[E]$, where E represents active enzyme molecules and E' represents inactive molecules. Then, the combination of equation 48 with the van't Hoff relationship for K_T ($-RT \ln K_T = \Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$) results in

$$v = \frac{A \exp(-E_a/RT)}{1 + \exp(\Delta S^\circ/R - \Delta H^\circ/RT)} \tag{50}$$

where ΔG°, ΔS°, and ΔH° are the standard Gibbs free en-

ergy, enthalpy, and entropy of denaturation reaction, respectively, and v is the observed rate of biological process. This equation produces a curve with single maximum and fits to most of the available experimental data on temperature-dependent variations of enzymatic activity.

Sometimes denaturation is irreversible, and there are no possibilities for simple mathematical expressions, because temperature effects depends on the exposing time. However, numeric solutions of respective differential equations can still be used.

Simple Models of Microbial and Cell Growth

This section deals with simple unstructured models. These models mainly ignore any changes in cell quality (biochemical composition, spectrum of enzymatic activity, etc.) induced by environmental factors.

Main State Variables and Growth Parameters. There are two types of growth models, *deterministic* and *stochastic*. The former describe clear determined and regular processes. The latter deal with random or stochastic processes. The main variables used in deterministic models are the same as in chemical kinetics, concentrations of biomass, substrates, and products, and stochastic models consider instead the probabilities, frequency distributions, variance, and so on. For example, a stochastic model can consider the probability of a single bacteria cell dividing under specified environmental conditions. Although any real-life biotechnological process has both deterministic and stochastic components, most useful growth models are strictly deterministic. In this section, we will concentrate on this type of model, and the stochastic counterparts will be considered only in "Cell Cycle".

The major state variables of deterministic models describing cell growth are described in Table 7. The concentration of biomass and cell number are related to each other by simple formula:

$$x = m \times N \tag{51}$$

where conversion factor m_1 is the average dry mass of a single cell. In some studies, it is save to assume m to be a constant and use both variables x and N as equivalent measures of growth. However, the average size and mass of single cells vary depending on the nature of studied organisms and environmental conditions (see "Cell Cycle"); therefore, it is advisable to make a selection. The biomass, x, has obvious advantage in studies aimed at understanding or control of mass flows, and cell number, N, is preferred in population studies when, for example, mutation or plasmid transfer is an essential factor controlling the efficiency of the biotechnological process.

The choice of method to determine biomass or cell number depends on many factors (2,5). Today, the preference should be given to those techniques that allow exact and automated measurements (Table 8). The most advanced analytical methodology is now available for automatic recording of gaseous or volatile substrates, intermediates, and end products, such as methane, CO_2, O_2, H_2, volatile fatty acids, alcohols, and other fermentation products (IR-analyzers, mass spectrometry, gas chromatography, NMR, etc.).

The *primary* state variables that are measured directly in cell culture are usually recalculated into the *secondary* growth characteristics: gross growth rate, dx/dt; specific growth rate, $\mu = (dx/dt)/x$; degree of multiplication, x/x_0; biomass doubling time, $t_d = \ln 2/\mu$; growth yield, $Y = -dx/ds$; product yield, $Y_{p/s} = -dp/ds$, $Y_{p/x} = -dp/dx$; specific rate of substrate consumption, $q_s = (ds/dt)/x = \mu/Y$; specific rate of product formation, $q_p = (dp/dt)/x = \mu Y_{p/x} = \mu(Y_{p/x}/Y)$. Most of the listed secondary growth characteristics are called *specific* rates expressed as a first-time derivatives of measured variable *per unit of cell mass*. The specific rates are not sensitive to variations in total cell biomass, so they can be considered as analogous to enzyme activity (q_s, q_p) in most kinetic derivations. The specific growth rate, μ, may be viewed as the activity of an autocatalytic enzyme producing itself. The specific growth rate, μ, measured from biomass dynamic, x, can differ from that estimated from the increase in cell number, N (denoted as μ_N). It follows from equation 51 that

$$\mu = \frac{1}{x}\frac{dx}{dt} = \frac{1}{x}\frac{d(N \times m)}{dt} = \frac{1}{N \times m}\left(N\frac{dm}{dt} + m\frac{dN}{dt}\right) = \frac{1}{m}\frac{dm}{dt} + \frac{1}{N}\frac{dN}{dt} = \mu_{cell} + \mu_N \tag{52}$$

If mass of single cell m is constant, then $\mu_{cell} = 0$ and $\mu = \mu_N$. Otherwise, we have to take into account m variation.

Validity of Exponential Growth Low. One of the earliest postulates in microbial kinetics is that under optimal non-

Table 7. Major State Variables of Deterministic Models

Variable	Notation	Dimension (examples)
Concentration of cell biomass	x	g CDW L^{-1} of cultural liquid
Cell number[a]	N	10^9 cell mL^{-1} of cultural liquid
Single-cell mass[a]	m	g CDW per cell
Mycelium length[b]	L	meter L^{-1} of cultural liquid
Tips number[b]	n	10^6 tips mL^{-1} of cultural liquid
Concentration of limiting substrate	s	g L^{-1} of cultural liquid
Concentration of product	p	g L^{-1} of cultural liquid or g g^{-1} of CDW

[a]For unicellar organisms (bacteria, yeasts).
[b]For filamentous organisms (fungi, actinomycetes).

Table 8. Methods Used for Determination of Microbial Biomass

Method	Measuring principle	DL	Advantage	Disadvantages
Dry mass	Mass of separated and dried solids	50	Provides direct unconditional estimate	Interference from dead cells and noncell solids
Wet mass	Mass of separated material	50	Simplicity, quickness	The variation of wet biomass bulk density
Wet biovolume	Linear dimension of pelleted cells or colony	100	Simplicity, quickness	The variation of wet biomass bulk density
Particulate organic carbon	CO_2 after cell separation and combustion	1.0	High precision and sensitivity, provides direct estimate of mass	Interference from dead cells and noncell solids
Biuret proteins	Colorimetric reaction of peptide bonds	1.0	High uniformity	Variation of protein-to-biomass ratio, possible extracellular accumulation
Folin-Ciocalteu protein	Colorimetric reaction of tyrosine and tryptophan	0.1	High sensitivity	Variation of protein-to-biomass ratio and amino acid composition of cell protein, possible extracellular accumulation
DNA	Colorimetric estimation of deoxyribose	1.0	High specificity, constancy of the DNA cellular content	Possible extracellular accumulation
ATP	Bioluminescence assay	10^{-5}	High sensitivity and specificity	Variation of the intracellular ATP pool
Fatty acids	Gas or liquid chromatography, colorimetric reaction	1.0	High specificity, allows identification of composition of mixed culture	Variation of the intracellular content, possible extracellular accumulation
Metabolio potential	Rate of added substrate uptake or product formation	10^{-3}–10	High specificity, quickness	Variation in conversion factor from measured rate to biomass
Opacity	Light scattering	0.1	Simplicity, quicknerss, automation	Interference from noncell solids, cell aggregation, wall growth
Electrical measurements	Conductivity	1	Simplicity, quickness, automation	Variation of conversion factor, interference from electrolytes in the medium
Manual microscopy	Cell count, measuring of hypha length	10^{-5}	Low cost, sensitivity, allows visual assay of biomorphological structure	Time-consuming, low precision
Image analysis	Computer-aided count	10^{-5}	Combined benefit of manual and automated quantification, speed, generation of size distribution	High cost
Coulter counter	Automated count and sizing	10^{-5}	Quickness, automation, generation of size distribution	Interference from noncell solids
Plating	Growing of the colonies on Petri dish	10^{-5}	Low cost, high sensitivity	Time consuming, low precision

Note: DL = detection limit, the minimum CDW required for an estimation with an error <2% (mg).

restrictive conditions (nutrient media containing all essential components at nonlimiting concentrations, absence of inhibitors, adequate physicochemical parameters, perfect mixing), the increase in biomass (dx) during an infinitely small time interval (dt) is proportional to this time interval and the instant biomass concentration (x), that is

$$dx = \mu x dt$$

or

$$dx/dt = \mu x \tag{53}$$

where μ is the proportionality coefficient. If μ is constant, integration of equation 53 gives

$$\ln x = \ln x_0 + \mu t \Rightarrow x = x_0 \exp(\mu t) \tag{54}$$

where x_0 is biomass at zero time or *inoculum size*.

According to early views, true exponential growth occurs only in the case of symmetrically dividing bacteria with equal probability of subsequent division for the mother and daughter cells. Organisms with asymmetrical multiplication (budding yeasts) were thought to obey the exponential low only approximately, whereas the growth of filamentous organisms (fungi, streptomycetes) was deviated considerably. However, direct measurements performed from 1930 to 1960 (5) revealed that exponential growth does occur in all prokaryotes and eukaryotes independent of their biomorphological features, including protozoa, fungi, and homogeneously cultivated plant and

animal cells. The deviation is observed only as a result of a growth-associated change in the environment. Wang and Koch (26) made very precise measurements of μ dynamics by growing *Escherichia coli* directly in an aerated cuvette of the computer-linked double-beam spectrophotometer. They found temporary slowdowns in the complex peptone plus beef extract media attributed to diauxie phenomena (e.g., depletion of some preferred oligopeptides) and a gradual increase of the μ value during sequential subculture in succinate minimal medium. Long-term cultivation of microorganisms in continuous turbidostat or pH-stat culture (14) revealed that the increase in μ is mainly the result of autoselection of mutants having higher maximum specific growth under specified cultivation conditions (see "Population Dynamics").

Early Views of Cell Growth. The derivation of the exponential growth equation was done for binary dividing bacteria, based on the geometric progression 2, 2^2, 2^3 . . . (27):

$$N = N_0 2^n \tag{55}$$

where n is the number of divisions, $n = t/t_d$.

The growth dynamics were viewed as a succession of distinct phases approximated with empirical formulas (15):

$$N = N_0 \exp[\mu F(t) t] \tag{56}$$

where function $F(t)$ undergoes stepwise changes as shown in Table 9 ($1.56 \leq n \leq 2.7$). The alternative way to describe growth dynamics was to use the logistic equation borrowed from demographic studies (28):

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right)$$

or

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{x_m}\right) \tag{57}$$

where r is the net growth rate (the difference between true growth and death rates) and K or x_m are the upper limits of, respectively, population density or biomass (K is used preferentially in ecological literature, being called the carrying capacity of the ecosystem). This equation simulates S-shaped growth dynamics frequently observed in natural ecosystems and cell cultures.

Table 9. Stepwise Changes of Function $F(t)$

Initial stationary phase	$F(t) = 0$
Lag phase	$F(t) = t^{n-1}$
Logarithmic growth phase	$F(t) = 1$
Negative growth acceleration phase	$F(t) = t^{t-1}$
Maximum stationary phase	$F(t) = 0$
Accelerating death phase	$F(t) = t^{n-1}$
Logarithmic decrease phase	$F(t) = 1$

A similar (but not identical) equation of cell biomass dynamics would be produced if we assumed that the limiting substrate is consumed according to first-order kinetics and is converted to biomass with a fixed yield factor:

$$Y = (x - x_0)/(s_0 - s) \tag{58}$$

then

$$\frac{ds}{dt} = -ks$$

$$s = \frac{x_0 + Ys_0 - x}{Y} = \frac{x_m - x}{Y} \Rightarrow \frac{dx}{dt}$$

$$= -Y\frac{ds}{dt} = kx_m\left(1 - \frac{x}{x_m}\right) \tag{59}$$

Monod's Model. This model still is very popular because of its elegant simplicity. It played an important role in the history of microbial growth kinetics as a first encouraging example when the theoretical development based on mathematical formalisms came before novel experimental designs. As compared with simple exponential equation 53, this model (29) takes into account the mass-conservation condition linking substrate uptake to biomass formation through constant yield factor Y (equation 2) and the dependence of the specific growth rate on limiting substrate concentration:

$$\frac{dx}{dt} = \mu(s)x = \mu_m \frac{s}{K_s + s} x$$

$$\frac{ds}{dt} = -\frac{1}{Y}\frac{dx}{dt} \tag{60}$$

By substitution of s by x through Y (equation 58), we can reduce equation 60 to a single equation having the analytical solution:

$$\mu_m t = (1 + P)\ln(x/x_0) - P\ln(Q - x/x_0) + P\ln(Q - 1) \tag{61}$$

where $P = YK_s/x_m$, $Q = x_m/x_0$, $x_m = Ys_0 + x_0$, and s_0 and x_0 are, respectively, s and x at $t = 0$.

Equation 61 contains three parameters, Y, μ_m, and K_s that can be thought of as passport data for a particular organism (e.g., for *E. coli* grown on glucose at 30 °C, $Y = 0.23$, $\mu_m = 1.35\ h^{-1}$, and $K_s = 4\ mg\ L^{-1}$) and describe the S-shaped growth dynamics of a batch culture. The initial conditions, s_0 and x_0, are set by the experimenter, who selects the inoculation dose and medium composition. Thus, knowing the values of all these entities, the growth dynamics can be calculated *before the experiment*. Moreover, this model was used as a basis for development of the *chemostat theory* to understand and predict the behavior of microbial cultures in continuous-flow bioreactors before the respective hardware was constructed. For this purpose, original equation 60 was modified as follows (30):

$$\frac{dx}{dt} = (\mu - D)x$$

$$\mu = \mu_m \frac{s}{K_s + s} x$$

$$\frac{ds}{dt} = D(s_r - s) - \frac{1}{Y}\mu x$$

$$D = \frac{F}{V} \tag{62}$$

where s_r is the concentration of limiting substrate in the fed-fresh medium delivered with pump from reservoir, F is the pumping rate (cm^3/h), V is culture volume (cm^3), and D is the dilution rate defined as the ratio $D = F/V$, h^{-1}.

The mathematical analysis of equation 62 allows us to make the following conclusions on the growth kinetics of continuous culture:

1. The described open system attains stable *steady state* when $dx/dt = 0$, $ds/dt = 0$, and variables x and s are not changed with time (denoting steady-state concentrations $\tilde{x}$ and $\tilde{s}$, respectively). From the first equation in equation 60, it follows that $\mu = D$, that is, the specific rate of microbial growth is equal to the dilution rate, which is under the full control of the experimenter. From the second equation, we find

$$\tilde{x} = Y(s_0 - \tilde{s}) \tag{63}$$

   As compared with Y expression for batch culture (equation 58) the formula in equation 63 no longer contains the term, x_0, the inoculum size. Thus, a steady-state culture is not dependent on past history, and its properties are determined solely by current cultivation parameters.
2. The dilution rates permitting stable microbial growth (i.e., those D, at which $\tilde{x} > 0$) are confined between 0 and washout point or *critical dilution rate* D_c

$$D_c = \mu_m s_r/(K_s + s_r) \tag{64}$$

3. The specific growth rate μ does not depend on either s_0 or x and is governed solely by the substrate concentration in the cultural liquid. On the other hand, the indirect effect of s_0 on μ may be evaluated from the dependence of $\tilde{x}$ on D, as soon as $\tilde{x}$ and $\tilde{s}$ are linked by the conservation condition in equation 62.

The most important biological implication of the chemostat theory was the discovery of the fact that microorganisms can grow endlessly with any rate between 0 and μ_m. Such substrate-limited growth is nevertheless exponential, because two subsequent acts of cell division will be separated by a constant time interval. Earlier, substrate limitation had been observed only transiently at the end of the exponential phase of batch growth. The most important biotechnological implication of the chemostat model was the concept of *controlled cell biosynthesis*, which implies the purposeful manipulation of microbial culture to optimize the productivity by such tools as changing the flow rate and the composition of medium in continuous-flow bioreactors.

Derivation of the Monod Equation from Mechanistic Considerations. The equation relating μ and s is known as the *Monod equation*:

$$\mu = \mu_m \frac{s}{K_s + s} \tag{65}$$

It was introduced by the author as a purely empirical relationship resembling adsorption isotherm. However, many microbiologists did view the Monod equation as something having deep inherent meaning rather than as just an empirical formula. Below, we shall outline a number of attempts to deduce this equation logically from the conjectured growth mechanisms.

1. *The bottle-neck concept*. There is an obvious similarity between the Monod and Michaelis-Menten equations. A theoretical substantiation for this similarity can be the bottleneck postulate originally put forward by Blackman (31). According to this postulate, the growth rate of a cell is determined by a single *enzymatic master reaction*. Microbial metabolism is symbolized as a unidirectional chain of reactions of substrate S conversion to cell biomass X' via some hypothetical intermediates P_1, P_2 . . . P_n (32):

$$S + X' \xrightarrow{k_1} P_1 \xrightarrow{k_2} P_2 \rightarrow \ldots \xrightarrow{k_n} P_n \xrightarrow{k_{n+1}} 2X' \tag{66}$$

Assuming steady state in respect to intermediate concentration p_i, $dp_i/dt = 0$, we arrive at the Monod equation, composite parameters μ_m and K_s being expressed via elementary kinetic constants of individual enzymatic reactions:

$$\frac{dx}{dt} = \mu_m \frac{s}{K_s + s} x$$

$$\mu_m = \frac{1}{\sum_{i=2}^{n+1} 1/k_i}$$

$$k_s = \frac{1}{k_1 \sum_{i=2}^{n+1} 1/k_i} \tag{67}$$

where $x = x' + p_1 + \ldots p_n$. The bottleneck idea can now be formulated as follows. If one of the constants $k_j \ll k_i$, $i = 2, \ldots, n+1$, $i \neq j$, then $\mu_m = k_j$ and $K_s = k_j/k_1$, and so the jth enzymatic reaction is the master reaction.

The obvious shortcomings of the reaction scheme in equation 66 are the unjustified oversimplification of the cell metabolism (which is a network, rather than a simple unidirectional sequence of reactions) and the lack of a definite interpretation of variables x' and p in real biochemical terms. An evaluation of the characteristic times of major intracellular metabolic reactions showed that there could be two possible bottlenecks: uptake of limiting sub-

strate (processes of active transport, transphosphorylation of sugars etc.) (33) and the formation of intracellular protein-synthesizing structures such as rRNA and the ribosomal proteins (34).

2. *Stochastic considerations*. Microbial populations were assumed to consist of active (which utilize the substrate) and inactive cells (35). The probability of a transition from the inactive state to the active one was supposed to be proportional to s, whereas the probability of the reverse transition was s independent. From these assumptions, a relationship between μ and s can be inferred, similar to equation 65.

3. *Derivations based on mechanistic considerations of protein synthesis (36)*. The rate of protein synthesis, dp/dt, is supposed to be determined by rRNA concentration R and by the size of the amino acids pool, A, $dp/dt = kxRA/(K_A + A)$, k = constant. Other conditions are defined as $R = R_0 + (R_m - R_0)\mu/\mu_m$, $A = bs$, and $P = p/x$, where R_m and R_0 are, respectively, the upper and lower limits of R variation, and P and b are constants. For a steady-state chemostat culture, $dp/dt = 0$ and $dx/dt = (1/P)(dp/dt) = 0$, then

$$\mu = (kR_m/P)s/(R_mK_A/\beta r_0 + s) = \mu_m s/(K_s + s) \quad (68)$$

Here again, the Monod equation is derived through a consideration of underlying intracellular processes. Needless to say, none of the cited derivations are free from criticism. The Monod equation remains empirical. Any attempts to provide the mechanistic interpretation of this equation inevitably lead to much more complicated mathematical expressions (see late section on structured models).

Biological Meaning and Experimental Determination of Growth Parameters K_s and μ_m. The parameter μ_m, maximal specific growth rate, has very lucid biological meaning: it is the upper limit of μ variation on specified nutrient medium. It could not be attained in reality because of its asymptotic nature: $\mu \to \mu_m$ as $s \to \infty$. However, in practice μ_m is achieved if $s \gg K_s$. It should be remembered that μ_m is not absolute maximum of growth rate, because it depends on the nature of the limiting substrate. For example, *E. coli* has μ_m above 2.5 h^{-1} on complex beef-extract medium and below 1.0 h^{-1} on minimal medium with succinate. Saturation constant K_s could be defined functionally as such concentration of limiting substrate that provides a specific growth rate equal to 0.5 of μ_m. We can say also that K_s is a measure cell affinity to substrate: the lower is K_s the better the organism is adapted to consume substrate from diluted solution. Any other definitions are speculative, e.g., K_s interpretation as the dissociation constant of ESC of the cellular enzyme involved in the first step of substrate conversion.

There are several ways to determine numeric values of K_s and μ_m:

1. *Batch culture, differential form of Monod equation*. The biomass dynamics, $x(t)$, are followed from the start of the exponential phase until the complete consumption of the limiting substrate and the attainment of the maximal cell density, x_m. The residual substrate concentration, s is calculated from mass balance (equation 58); yield is calculated as $(x_m - x_0)/s_0$; the s_0 value should be known) and corresponding $\mu(t)$ values from the slopes $d(\ln x)/dt$. Finally, parameters K_s and μ_m (equation 65) are to be found graphically or from nonlinear regression as in the case of the Michaelis-Menten equation (see later text).

2. *Batch culture, ignoring the substrate uptake*. The inoculum size, x_0, and the duration of experiment are chosen to minimize the uptake of added substrate $s \approx s_0$ (37). Typically, bacterial cell density should be of the order 10^{-5} or less, and it is measured by such sensitive instruments as the Coulter counter. The quasi-steady-state growth rate is measured at several s_0 values and then fitted to equation 65 as described before. In numerous determinations made at high cell densities (when we can no longer neglect substrate consumption), it was observed that dependence of μ during exponential growth phase on s_0 is formally described by the Monod equation with s replaced by s_0 (38). The explanation of these results could be made only by more complex structured models; however, the described procedure is not appropriate for K_s and μ_m determination.

3. *Batch culture, integral form of equation 61*. The S-shaped curve of biomass dynamics is fitted directly to equation 61 through either preliminary linearization or nonlinear regression (preferential). Sometimes (e.g., when cells are grown in opaque media) it is more convenient to follow the dynamics of residual substrate $s(t)$ or product formation $p(t)$, for example, CO_2 evolution or dynamics of O_2 uptake (respiration). In these cases, we can integrate the set of differential equation 60 in terms of $s(t)$ or $p(t)$ dynamics, taking into account mass–balance relationships in equations 7, 8, and 58 and similar equations.

4. *Steady-state chemostat culture*. The chemostat provides the opportunity for estimations under steady-state condition, which is commonly believed to be more reliable. By running an experiment at different dilution rates, D, the corresponding $\tilde{s}$ values may be measured and the dependence of $\mu = D$ on $\tilde{s}$ obtained; in principle, this can be done as accurately and carefully as desired. Such an approach, however, may and frequently does encounter serious technical problems because of the high affinity to limiting substrate of some microorganisms. There is a need to (*1*) select highly sensitive analytical techniques to measure extremely low residual concentrations of particular substances; (*2*) develop instant sampling procedures to minimize substrate loss, and (*3*) eliminate apparatus-related artifacts such as nonperfect mixing and fluctuations in nutrient medium supply. This can be accomplished by use of radiolabeled substrate, and other tips are covered in specific experimental works (39,40). However, the most essential objection to this method is that organisms at various steady states do change their kinetic properties, which is not accounted by Monod's model (see "Structured Models").

5. *Non-steady-state chemostat culture*. The measurements are made during short-term experiments started by addition of different amounts of limiting substrate to a steady-state chemostat culture. Then, substrate uptake or respiration rates are recorded in perturbed culture until

the new steady state is established (2). The opportunity of continuous culture is that the physiological state of cells just before perturbation is well defined and reproducible. On the other hand, these experiments do provide data on affinity to substrate (K_s) and maximal rates of respiration (e.g., Q_{CO_2} or uptake Q_s, which are related but are not identical with μ_m):

$$\mu_m = Y_{p/x}Q_{CO_2} = Y_{s/x}Q_s \tag{69}$$

where stoichiometric parameters $Y_{p/x} = dp/dx$ and $Y_{s/x} = ds/dx$ should be determined in independent experiments. Immediate assay of μ_m in such an experimental setup is possible as follows: the steady state is perturbed by setting the dilution rate by 20 to 100% higher than the critical value D_c (equation 64). The washout dynamics are followed and μ_m is determined from approximate relationship (5): $x \approx x(0)\exp[(\mu_{m-D})t]$, where $x(0)$ is biomass before perturbation. More rigorous estimate of μ_m could be provided if such washout experiments were made at several (at least two) input substrate concentrations s_0 to account for the fact that s_r is still not infinity.

Modification of the Monod Equation. It was found that not all experimental data could be reasonably well fitted by equation 65. The best-fit hyperbola often passed above experimental points at small s and below them at large s. A better fit was claimed to be provided by using the following, entirely empirical, equations:

$$\mu = \mu_m[1 - \exp(-Ks)] \tag{70}$$

$$\mu = \mu_m s^n/(K_s + s^n) \tag{71}$$

$$\mu = \mu_m s/(K_s x + s) \tag{72}$$

The preference of the first expression, equation 70 (41), is questionable. An expansion of Monod equation by addition of the third parameter, n in equation 72 (42), or introduction of the second variable, x in equation 71 (43), does improve the approximation capability of kinetic equations. We can even provide the mechanistic basis for this improvement. Thus, the Moser equation (equation 71) is similar to the Hill equation in enzymology (equation 43), indicating the cooperativity effects in performance of some master reaction of cellular metabolisms. The Contois equation (equation 72) could be interpreted in terms of growth autoinhibition by-products, because under the realistic assumptions ($x \gg x_0$, $xY_{p/x}/K_p \gg 1$ and product p formation coupled with growth), the apparent saturation constant K_s^{app} is almost proportional to accumulated biomass x:

$$K_s^{app} = K_s(1 + p/K_p) = K_s[1 + Y_{p/x}(x - x_0)/K_p] \approx K_s x \tag{73}$$

An Account of Maintenance Requirements. Powell (44) assumed that substrate uptake rate q obeys the Michaelis-Menten kinetics; then, from mass–balance equation 18 it follows that

$$q_s = Q_s/(K_s + s) = \mu/Y^m + m \tag{74}$$

If μ_m is defined as $\mu_m = YQ_s$ (5), then

$$\mu = \mu_m s/(K_s + s) - mY^m \tag{75}$$

If we define $\mu_m = Y^m(Q_s - m)$ (2), then

$$\mu = \mu_m(s - s^*)/(K_s + s)$$
$$s^* = K_s m Y^m/\mu_m \tag{76}$$

Equations 75 and 76 both predict the occurrence of threshold substrate concentration s^*, below which growth is impossible. It yields a substantially better fit to experimental data. The difference is that according to equation 75, the parameter μ_m is the specific growth rate under imaginary conditions of zero maintenance requirements, whereas equation 76 implies the traditional definition of μ_m as the specific growth rate under substrate excess: $\mu \to \mu_m$ when $s \to \infty$.

Account for Substrate Leakage. Mathematically identical to equation 76, a modification of the Monod equation was proposed for the case of conserved substrates. However, the biological meaning is entirely different: if the specific leakage rate is assumed to be constant, then a decrease in s down to some threshold value s^* will lead to the counterbalance of the two reverse processes (uptake and leakage), so that the net consumption of the limiting substrate will be zero.

Account of Inhibitory Effects. A few valuable refinements of the Monod equation were borrowed from enzymology; most often they were noncompetitive and substrate inhibition. The former inhibition mechanism is by the so-called Monod-Ierusalimsky equation:

$$\mu = \mu_m \frac{s}{K_s + s} \times \frac{K_p}{K_p + p} \tag{77}$$

The growth retardation with an excess of such substrates as phenol, methanol, and ethanol is described by an analogue of Haldane's equation (45):

$$\mu = \mu_m \frac{s}{K_s + s + s^2/K_{ss}} \tag{78}$$

Equation 78 simulates a single-peak curve, and so the same μ value may be obtained at two different s, one in the substrate-limiting range, $d\mu/dt > 0$ (stable), and the other in the substrate inhibition range, $d\mu/dt < 0$ (unstable). A sustainable maintenance of a population under conditions of substrate inhibition is possible either in the second stage of a two-stage chemostat or in the case of plentiful wall growth in a conventional chemostat (46).

Account of Diffusion Effects. We will present one example of such models (44). The basic assumption is that substrate is taken up by an enzyme that obeys Michaelis kinetics and is localized on the inner side of cell membrane. The actual substrate concentration around the enzyme-active centers is smaller than in the solution, because of a limited diffusion rate. By applying a simplified Laplace equation, it was found eventually that

$$\mu = \frac{\mu_m(K_s + L + s)}{2L}\left(1 - \sqrt{1 - \frac{4Ls}{(K_s + L + s)^2}}\right)$$
$$\approx \mu_m \frac{s}{K_s + L + s} \quad (79)$$

where L is a factor determined by membrane permeability and by the maximum rate of the enzymatic reaction.

Structured Models

The unstructured models (such as Monod's model or its modifications) are able to predict and describe only simplest manifestation of growth phenomena. Sometimes it is declared that Monod-type models are able to describe only *balanced* and *steady-state* growth. The analysis of more complicated unbalanced and non-steady-state growth requires formulation of structured models.

Definitions. *Balanced growth* was defined by Campbell (47) as a proportional increase in the amounts of all cell components, in other words, balanced growth produces cells of the same quality without any variation of composition. The terms *steady state* and *non steady state* stem from chemical and enzyme kinetics. The first one refers to such a regime when the reaction rate remains constant because of an exact balance between formation and breakdown of intermediary products such as the enzyme–substrate complex. In microbial culture, the growth is called steady state if *specific* rather than *total rates* remain constant. In an open system, such as a chemostat, both total (dx/dt, ds/dt) and specific rates ($\mu = (1/x)dx/dt$, $q_s = (1/x)ds/dt$) tend to have constant steady-state values. A closed system, such as a batch culture, should be considered under steady state only during the exponential phase when μ and q_s are constant. The linear growth with a constant total rate ($dx/dt = \mu x =$ constant) is characterized by monotonously declining μ, and is not steady-state growth. However, under some conditions (such as in dialysis culture) it may attain *quasi steady state*, when $ds/dt \approx 0$.

The non-steady-state kinetic regimes take place before establishment of steady state or after its perturbation. In enzyme kinetics, non-steady-state measurements are taken in the millisecond range of time scale. In microbial cultures, similar non-steady-state transient and perturbation processes advance much more slowly, typically during several hours and days. An example is a transient process in the chemostat induced by changes in D or s_r (fed-substrate concentration). In such growth, μ, q_s, q_p, and other metabolic rates exhibit continuous variation in time. The attractiveness of non-steady-state studies for microbiology and biotechnology is obvious:

- They allow a wider range of hypotheses to be tested and yield much more data on the studied objects.
- They have higher practical value; in biotechnology, steady-state operation is the exception rather than the common routine because of unavoidable disturbances in cultivation conditions.
- They provide additional tools for optimal regulation of cell performance in the bioreactor, because purposeful non-steady-state growth may display greater efficiency and higher productivity (48).

The notions of steady-state and balanced growth are close but not identical. The first is more strict; steady-state growth has to be balanced by definition (otherwise some of the specific rates responsible for synthesis of the changed cell component should vary). On the other hand, the balanced growth can be for some time nonsteady, perhaps during the late exponential phase of batch culture when μ declines while cell composition remains unchanged. During long-term experiments, non-steady-state growth leads inevitably to a change of cell composition; it becomes unbalanced. For heterogeneous populations, the situation may be more complicated. For example, the growth in the second stage of a two-stage chemostat attains a steady state, and the biomass and residual substrate concentration are constant. However, such growth is not balanced, because cells delivered from the first stage differ in their properties and composition as compared with the bulk of cells in the second stage. This situation was termed the *transient steady state* (49).

By *structured*, we mean mathematical models describing growth-associated changes in microbial cell composition. It includes mass–balance equations for all assigned intracellular components. Their concentrations can be expressed either per unit volume of fermenter vessel (c_1, c_2, . . . , c_n), or per unit cell mass (C_1, C_2, . . . , C_n), and hence $C_i = c_i/x$. The mass–balance equations can be written as follows

$$\sum_{i=1}^{n} c_i = x$$
$$\sum_{i=1}^{n} C_i = 1 \quad (80)$$

For each variable C_i, a differential equation is written that takes into account all sources, r_+, and sinks, r_-, as well as its dilution from cell mass expansion (growth)

$$\frac{dC_i}{dt} = r_+(s, C_1, \ldots, C_n) - r - (s, C_1, \ldots, C_n) - \mu C_i \quad (81)$$

The simplest structured models with n no more than 2 or 3 are called two- or three-compartment models. For example, a model (50) incorporated two compartments: nucleic acids and proteins combined with other active cell components. The model variables also included concentrations of the limiting substrate and the inhibitor. Compared to Monod's model, the proposed set of four equations was able to account for a much wider range of dynamic patterns. Specifically, it simulated D-dependent changes in the cell composition (chemostat) and all known growth phases of batch culture from the lag to decline. However, the choice of variables in this model was more or less arbitrary, so it should be regarded more as a bright illustration rather than a research tool.

During the past decade, much more realistic structured models based on biochemical data have been developed.

The simulation model of *E. coli* growth (51) contains the following dynamic variables: glucose and NH_4^+, as exosubstrates; CO_2 and acetate, as products excreted into the medium; amino acids; ribonucleotides; deoxyribonucleotides; monomeric precursors of cell wall components; rRNA and tRNA; nonprotein polymeric components; glycogen; guanosinetetraphosphate; enzymes transforming ribonucleotides into deoxyribonucleotides; ATP; NAD(H); and protons. Altogether, the dynamic model amounts to a system of 21 differential and 14 algebraic equations. An even more complicated model simulating growth of *Bacillus subtilis* (52) is the set of 39 nonlinear and coupled differential equations containing nearly 200 parameters! These models are able to simulate particular growth features such as changes in cell sizes, shape, and composition as well as the *D*-dependent variations in replication time brought about by the shifts in glucose concentration. However, the predictive capability of such an intricate dynamic model should be still estimated as modest as compared with invested modeling efforts; they are still nothing more than a "caricature parody" of microbial biochemistry and are too complex to be studied by conventional mathematical tools (stability analysis, parameters identification, etc.) or to be used in biotechnological applications. The best choice of a mathematical model lies, apparently, midway between unstructured and highly structured models outlined here. One of the known compromises has been found through attempts to express quantitatively the *cell physiological state*.

Physiological State of Chemostat Culture. The term was coined by Malek (53) without giving a clear definition. The impetus for the development of the concept of a physiological state was the evidence on changes in the chemical composition of microorganisms as dependent on dilution rate *D* and medium composition in chemostat culture (54). It has been found that some of the studied parameters remained constant (content of cell DNA and carbon) while others exhibited regular *D*-dependent variations, either an increase (RNA content, cell sizes) or a decrease (the content of reserved polysaccharides). Those properties that were *D* dependent were recognized as components of the *vector of the physiological state*.

Powell (44) combined and put on a quantitative mathematical footing three notions, which were beforehand separated and cloudy: (1) physiological state, (2) past history, and (3) non-steady-state growth kinetics of microbial culture. The specific rate of substrate uptake, q_s, was presented as a product

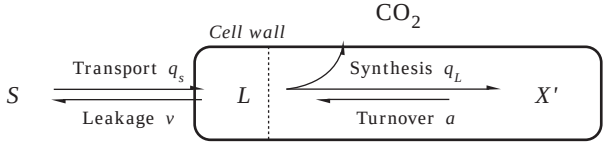
$$q_s(s) = q'_a(s)S(s) \tag{82}$$

where $S(s)$ is a simple saturation function (e.g., a Michaelis hyperbola), $S(s) = s/(K_s + s)$, and $q'_a(s)$ is a function associated with the microbial physiological state. The instant value of $q'_a(s)$ is determined by way in which s varies until the given moment (τ h ago), effects of later events contributing more than earlier ones. Transient processes are influenced by the past history of the culture in the following manner. Suppose that steady-state growth of a chemostat culture is upset and the residual substrate concentration jumps from $s(0)$ to $s(1)$. Immediately, q_s will increase from $Q(0)S[s(0)]$ to $Q(0)S[s(1)]$. If no further changes in $s(1)$ occur, then Q will also eventually attain a new steady-state value equal to $q'_a[s(1)]$. In essence, Q is the potential metabolic activity, that is, the specific rate of a key metabolic process measured just at the moment of relief from substrate limitation. The transient Q dynamics are described as net change equals production minus dilution caused by cell growth:

$$\frac{dQ}{dt} = r(Q, s) - \mu(Q, s)Q \tag{83}$$

Substance Q may, in reality, be represented by a single enzyme or multienzymatic complexes as well as by ribosomes or other cell components occupying the bottleneck position. Monod's model, supplemented by equation 83, made possible at least a qualitative understanding of chemostat-transient processes triggered by a D switch (55).

The *synthetic chemostat model (SCM)* (2) combines Powell's ideas with the routine of conventional structured models. This model is similar to the cybernetic model (56,57). The basic SCM interprets microbial growth as a conversion of exosubstrate S into cell macromolecules X' via a pool of intermediates L:



Macromolecular cell components are susceptible to degradation (turnover), and monomeric metabolites can escape into environment because of leakage. Contrary to simple chemical catalysts, the composition of end product X' is not uniform and varies in response to a changing environment because of the adaptive nature of microbial metabolism. At the heart of the SCM are the solution of the problem, how to cope with these variations, and how to describe adaptive changes in cell composition by relative simple models.

All macromolecular cell constituents are divided into two groups: primary cell constituents necessary for intensive growth (P components), and components needed for cell survival under any kinds of growth restriction (U components). The characteristic examples of P components are ribosomes (rRNA and ribosomal proteins) and enzymes of the primary metabolic pathways. Their intracellular content increases parallel to an increase in μ. The contribution of U components to cell biomass decreases with growth acceleration. Examples are enzymes of the secondary metabolism, protective pigments, reserved substances, and transport systems of high affinity.

The analysis of available data as well as mass-conservation conditions allowed the formulation of several rules of variation of cell components taking place because of optimal control of cell biosynthesis:

1. Amounts of P and U components expressed as a fraction of total cell mass (P and U, gram per gram of

biomass respectively) vary within the upper (P^{max}, U^{max}) and low (P^{min}, U^{min}) limits, the latter being the constitutive part.

2. An increase of one individual P component is accompanied by increase of other P components.
3. The total enlargement of P_{sum} is accompanied by corresponding decrease of U_{sum} and vice versa.
4. The P/U ratio is controlled by the limiting substrate concentration in an environment.

These rules are translated into mathematical terms as follows:

$$\frac{P_1 - P_1^{min}}{P_1^{max} - P_1^{min}} \approx \cdots \approx \frac{P_n - P_n^{min}}{P_n^{max} - P_n^{min}} = r$$

$$\frac{U_1 - U_1^{min}}{U_1^{max} - U_1^{min}} \approx \cdots \approx \frac{U_m - U_m^{min}}{U_m^{max} - U_m^{min}} = 1 - r$$

$$\tilde{r} = \frac{\tilde{s}}{K_r + \tilde{s}} \tag{84}$$

where variable r (index of physiological state) is already scalar (not vector!) function controlled by environmental factors (e.g., concentration of the limiting substrate). The r value in steady-state chemostat culture changes from zero (in culture at almost zero growth rate, when $\tilde{s} \to 0$ and all P components come down to low limits) to 1.0 (in unlimited culture, when $\tilde{s} \to \infty$ and P components attain maximum). During transients caused by sudden s changes, an instant r value goes toward new steady state (compare with equation 82):

$$\frac{dr}{dt} = \mu\left(\frac{s}{K_r + s} - r\right) \tag{85}$$

The introduction of the r variable greatly simplifies the use of structured models because the adaptive variation of cell composition (and metabolic activity, which is determined by the intracellular content of particular enzymes) now could be expressed via one single master variable r. For example, the specific rate of substrate uptake q_s is defined as:

$$q_s = r\frac{Qs}{K_s + s} + (1 - r)\frac{Q's}{K'_s + s} \tag{86}$$

where the first and the second terms on the right side stand, respectively, for low (P component) and high (U component) affinity of transport system. Similar r-dependent expressions are derived for other reactions (q_L, v, a) and stoichiometric parameters.

Predictive and clarifying capabilities of SCM turned out to be higher than more complex structured models. Contrary to all known chemostat models, SCM provides adequate simulation of D-dependent variation of the microbial physiological state. Under energy- and C-limited growth, it was expressed as an increase of apparent K_s, potential uptake and respiration rates, maintenance ration, and turnover parallel to increase of D. Under N limitation, a similar trend was complemented by considerable decrease of Y_N due to alteration of cell composition in favor of N-rich P components (Fig. 6). SCM adequately describes the transient growth caused by shift-up in chemostat culture. The phenomena of overshoot in substrate concentration and undershoot in biomass during transient growth are explained by slow adjustment of cell composition (RNA content, respiratory activity) to new growth conditions (Fig. 7). Batch culture limited by carbon and energy source was simulated by SCM on the whole from inoculation to death stage; conventional growth phases (lag, exponential, deceleration, stationary, decline) were generated automatically without setting any specific conditions (Fig. 8). It is important that SCM realistically describes and predicts not only net growth but also the survival dynamics of starving cells after substrate depletion. During this phase,

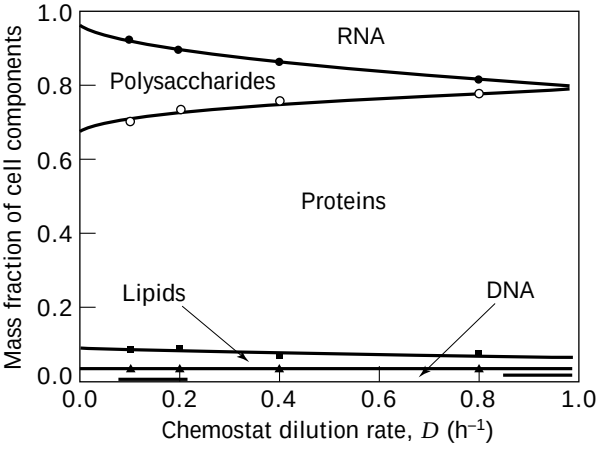


Figure 6. Simulation of D-dependent changes in cell composition of *Aerobacter aerogenes* grown in NH_4^+-limited chemostat culture. The curves are calculated from SCM (2), and the original experimental data are from Ref. 58.

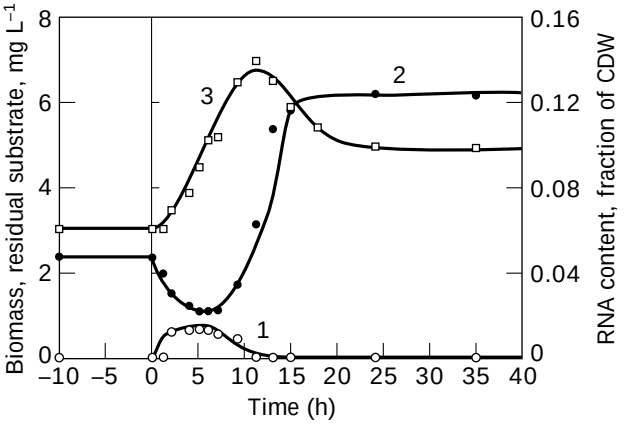


Figure 7. SCM simulation of transient growth induced by change of chemostat dilution rate. Residual glycerol concentration (*1*), biomass (*2*), and cell RNA content (*3*). At $t = 0$, dilution rate was shifted from 0.004 to 0.24 h^{-1}. *Source*: Redrawn from Ref. 2; the original data (59) are for chemostat culture of *Aerobacter aerogenes* limited by glycerol.

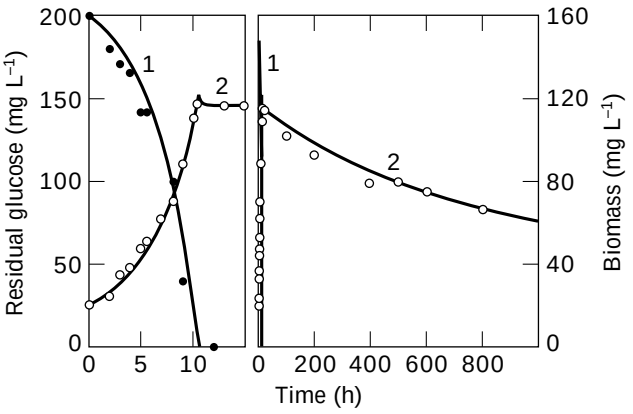


Figure 8. SCM simulation of complete dynamic curve of batch culture; residual glucose (*1*) and biomass (*2*) of yeasts *Debaryomyces vanrijiae* (old name *D. formicarius*) grown on glucose-mineral medium (2). Note that contrary to old empirical models (equation 56), all growth phases are reproduced automatically without specifying preset time ranges.

surviving bacteria sustain very slow cryptic growth at the expense of cell turnover and *L* leakage. The rate of decline in biomass gradually decreases as the result of buildup of some U components (parallel to the decrease of such *P* components as RNA and CN-sensitive respiration enzymes). Batch culture limited by conserved substrate (the source of N, P, Mg, Fe, or other) has an interesting feature: the growth is not stopped after substrate depletion but proceeds at an even higher rate. SCM explains this phenomenon by the partitioning of deficient elements between mother and daughter cells.

Cell Cycle

The term *cell cycle* is used to designate the regularly repeated sequence of events that occur between consecutive *cell divisions*, for example, the formation of two identical daughter cells from one cell, which takes place in most known bacteria. Equivalent cell cycle events include *budding* (most yeasts and budding bacteria), *branching of hypha* (filamentous organisms), *fragmentation* (some coccoid and corineform bacteria). The majority of available data have been obtained for rod-shaped bacteria (*E. coli, Salmonella typhimurium, Bacillus cereus*) under steady-state growth conditions when the cell cycle consists of three distinct phases:

1. The growth of the newborn cell without chromosome replication from the initial mass m_0 until some critical initiation mass m^*
2. DNA replication (C period)
3. Separation of daughter cells (D period)

Relationship between Cell Size and Specific Growth Rate. The periods C and D are constant; in *E. coli*, they occupy about 40 and 20 min, respectively, independent of environmental conditions. However, the duration of phase 1 depends on cultivation conditions; the time of the single-cell mass increase from m_0 to m^* is inversely related to the specific growth rate μ. It is this difference that yields μ-dependent variations of cell sizes, because faster-growing cells produce for the same τ = C + D time a larger cell mass than do slower-growing cells. It became clear from the following simple algebra.

The steady-state growth of an individual cell proceeds exponentially throughout the entire cell cycle [$m = m_0 \cdot \exp(\mu t)$]. At the time of the second part of the cycle, it takes exactly τ min for the cell to enlarge from critical mass m^* to $2m_0$. Hence,

$$2m_0 = m^* \times \exp(\mu\tau) \Rightarrow m_0 = 0.5m^* \times \exp(\mu\tau) \quad (87)$$

Equation 87 remains valid for steady-state culture at any μ, which can be varied from 0 to μ_m. Assuming that cell critical mass m^* is constant and does not depend on growth rate, we see that equation 87 predicts a μ-dependent variation mass of newborn cells m_0 and hence the average mass and size of the bacterial population. This finding is supported by numerous experimental studies from the beginning of this century (60) that displayed the positive correlation between cell size and growth rate. Because the term $\mu\tau$ is rather small ($\mu\tau \leq \mu t_d = \ln 2$), it is difficult to notice the curvature of the experimental curve within the measurement error. Thus, the relationship between the average cell size $\bar{m}$ ($\bar{m} \approx m_0 \times 2 \times \ln 2$ for rod-shaped bacteria; see derivation in equation 88) and μ is given as an empirical linear approximation:

$$\bar{m} = \bar{m}(0) + k\mu \quad (88)$$

The regression parameters $\bar{m}(0)$ and k have a meaningful biological interpretation: y intercept, $\bar{m}(0) \approx 2 \ln 2 \times 0.5m^* = (\ln 2) \times m^*$ is equal to approximately 69% of the cell critical mass m^* initiating chromosome replication, and slope $k = [\exp(\mu_m\tau) - 1]/\mu_m \approx \tau$ is close to the duration of C + D periods.

The postulate on the constancy of the critical mass, m^* was derived from the observation that the cell has accurate control over its size at division and poorer control over its age at division (61). Recently, accurate measurements with flow cytometry (62) revealed that m^* is inversely related to the specific growth rate; slowly growing cells tend to initiate DNA replication at a slightly higher critical mass as compared with intensively growing cells. However, we may safely assume that m^* variation is much smaller than the variation of cell mass during the cell cycle: $dm^*/dt \ll dm/dt$.

Simulation of Cell Cycle by Simple Deterministic Structured Model. In biochemical terms, it is difficult to envisage how cell mass per se could determine when to initiate replication. A more likely candidate is some mass-related parameter, such as intracellular concentration of some signal metabolite like guanosine tetraphosphate or an *initiator protein*, according to the popular model proposed by Helmstetter and Cooper (63,64). It was postulated that the initiation of DNA replication is triggered by a threshold intracellular concentration of this protein V^*; this protein is synthesized at a rate proportional to the total growth

rate and requires exactly one mass doubling time to reach its threshold concentration again. This mechanism is translated into mathematical form of a structured model such as SCM as follows (2):

$$\frac{dV}{dt} = \Theta(\mu + \mathrm{a}) - \mu V$$

$$\frac{d\lambda}{dt} = \zeta \begin{cases} \text{if } V \geq V^* & \text{then } \zeta = \zeta_1 > 0 \text{ else } \zeta = 0 \quad \text{(initiation)} \\ \text{if } \lambda \geq \lambda^* & \text{then } \lambda = \lambda/2,\ V = 0,\ m = m/2 \quad \text{(division)} \end{cases} \tag{89}$$

where Θ is the fractional contribution of protein V to total cell synthesis. The V content is an intrinsically transient entity; even during steady-state growth, it continuously changes between zero (Helmstetter and Cooper postulated the annihilation of the initiator protein after every replication cycle) to an upper-threshold value, V^*, which is less than the potential steady-state level, $\Theta(\mu + a)/\mu$. The second variable, λ, imitates the replicating chromosome; it sets up the discontinuity associated with cell division.

Analysis of coupled equation set 89 reveals that this model is stable toward perturbations. Suppose that by chance, the content of initiator protein has risen to some abnormally high level. The immediate result would be several more frequent cell divisions, with smooth reversal to a normal multiplication pace. Similar events take place under the opposite situation of V deficiency; several divisions are delayed, resulting in production of abnormally long cells, but then steady state is restored. The negative feedback mechanism that brings things back into line is based on the dynamic nature of variable V; it is characterized by a unique, stable steady state that is approached from different initial conditions. It may be easily shown that the described model adequately simulates various morphological effects exhibited during non-steady-state growth (Table 10).

Table 10. Morphological Effects during Non-Steady-State Growth

Effects	Explanation
The longer lag phase in batch culture when growth is surveyed by cell count rather than biomass measurements	The partial synchronization of cell division is delayed as compared with mass growth until attainment of the value $2\bar{m}$
The accumulation of enlarged cells during transition from lag to exponential growth phase	The μ–m relationship (equation 87) combined with the transient misbalance in V synthesis (equation 89)
The formation of dwarf cells in starving population	The small cells are produced during very slow cryptic growth of surviving organisms
The accumulation of division potential if division is blocked by inhibitor, then after block release all missing division takes place in quick sequence	This phenomenon is simulated by equation 89 and explained by overproduction of initiator protein in the presence of inhibitor

Statistical Analysis of the Population Distributions. Equation 87 to 89 were derived for average cells in the population. To embrace the variability of sizes, we have to analyze the *frequency distributions*. If certain conditions are met—the culture is fully desynchronized, cells grow according to some deterministic low, all divide into only two identical daughter cells, and there is no cell elimination—then age distribution (61) is given by:

$$\varphi(\mathrm{a})da = 2\mu e^{-\mu \mathrm{a}} da\colon 0 \leq \mathrm{a} \leq \ln 2/\mu = \mathrm{t_d} \tag{90}$$

where μ is specific growth rate, t_d is mean doubling time, a is the age since birth, and $\varphi(a)\,da$ is the frequency of cells whose ages are between a and $a + da$. Assuming that cells grow exponentially between divisions, then the frequency of mass distribution is

$$\varphi(m)dm = 2m_0/m^2 dm\colon m_0 \leq m \leq 2m_0 \tag{91}$$

where m_0 is the mass of newborn cell and $\varphi(m)\,dm$ is the frequency of cells whose masses are between m and $m + dm$. The mean cell size is $2m_0 \times \ln 2$ (calculated as an integral of $m \times \varphi(m)dm$). If cell growth between two consecutive divisions is linear (65), then

$$\varphi(m)dm = 4 \times \ln 2/\mathrm{m_0} \cdot \exp(-m \times \ln 2/m_0)dm\colon m_0 \leq m \leq 2m_0 \tag{92}$$

Equations 90 and 91 are called *canonical* age or mass distributions to emphasize that they are an idealized form applicable when cell division takes place at a precise size. Assuming that momentary distribution of size at division of individual cells is *normal* and *random* (not correlated with other cell cycle events), we can obtain computer-simulated curves for any fixed level of noise expressed as the coefficient of variation (CV). As shown in Figure 9, random variations of size at division tend to round the corners of the canonical distribution.

Another source of cell size variation can be nonequal separation of mother cells into two daughter cells. It is characterized by the K distribution, which is the distribu-

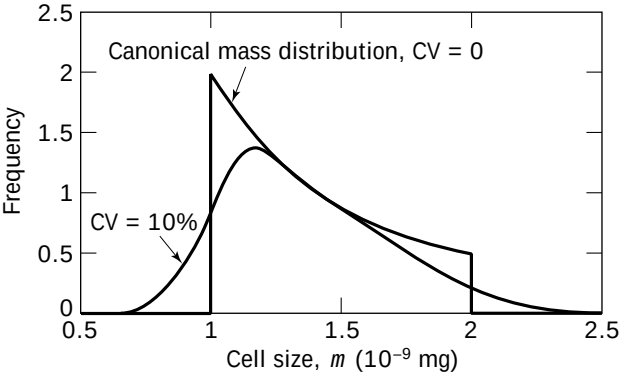


Figure 9. Distribution of cell sizes (as single-cell mass) for canonical cases where there is no variation in the size at division and for the case of normal distribution of size at division with CV = 10% (see details in Ref. 61).

tion of the ratio of daughter-cell length to mother-cell length. The average values absolutely necessary equals 0.5, but CV is at best about 4% (e.g., for well-behaved *E. coli* strains) and can attain rather high values for other organisms. The deviation of observed versus predicted size distribution may be caused also by cell death and different kinds of cell pathology (abnormally long or dwarf cells).

Collins and Richmond (66) introduced an entirely different approach based on the use of three distributions:

$$v_m = \mu\{2\int_0^m \Psi dm - \int_0^m \varphi(m)dm - \int_0^m \lambda(m)dm\}/\lambda(m) \tag{93}$$

where v_x is the growth rate of cells of size m, $\lambda(m)$ is the extent of population distribution, Ψ is the momentary distribution of dividing cells, and $\varphi(m)$ is the momentary distribution of newborn cells. This equation allows the calculation of the mean growth rate of cells of particular size class. Instead of this analytical method, Koch in his work with Shaechter (61) proposed the synthetic approach. Starting from the set of specific postulates of the cell multiplication mechanism (linear or exponential growth of cell mass between divisions, kind of control, the evenness of the division), they derived the size or age distribution, which then was compared with the observations.

The Growth Low. Some believe that there should be a general low of cell growth that can be discovered by sensitive methods of analysis. The rate of biomass growth throughout the cell cycle was hypothesized to be linear, bilinear, exponential, double-exponential, and so on. Two limiting cases have generally been considered: the exponential and linear growth models proposed, respectively, by Cooper and Kubitschek (57). To differentiate between these two mechanisms, three classes of experiments have been used:

1. Size measurements of individual cells growing in the microculture by use of phase-contrast microscopy or recently developed confocal scanning light microscopy combined with image analysis.
2. Pulse-chase labeling of cells with their subsequent separation into different phases of the cell cycle. Most frequently, labeled uracil and leucine are used as precursors of RNA and protein synthesis, respectively. There are two major sources of errors in this approach: poor resolving power of separation methods and artifacts associated with effects of exogenous compounds on intracellular fluxes (feedback inhibition, pool expansion, label dilution, etc.). One of the best options for separation is probably the "baby machine," based on the membrane elution principle (67). To minimize the second source of error, the mutants blocked in the synthesis of the probe compound can be used.
3. Analysis of the frequency distributions of steady-state populations (see equations 91 and 92). However, the resolving power of this approach is rather low because linear and exponential models produce similar patterns.

Most of the obtained results are in better agreement with the exponential growth model rather than with the linear model. However, there are some serious doubts about whether there is a unique simple mathematical growth low describing bacterial growth during the division cycle. Cooper (67) proposed distinguishing three categories of cell components that are synthesized with a unique pattern:

1. Cytoplasm (proteins, RNA and ribosomes, small molecules) that is accumulated exponentially.
2. Cell DNA that is replicated in a linear fashion as a sequence of constant and zero rates.
3. Cell surface composed of peptiodoglycan and membranes that are synthesized exponentially during most parts of the cell cycle, but immediately before cell division the synthesis accelerates to accomplish new pole formation.

Thus, the growth pattern of the whole cell is the sum of these three patterns. Because the cytoplasm is the major constituent (up to 80% of CDW), then the growth of the cell should be approximately exponential.

Population Dynamics (Mutations, Autoselection, Plasmid Transfer)

Description of Mutation and Autoselection. The continuous culture turned out to be a very efficient tool to study mutation and autoselection (67). Let N be the total cell concentration, M the concentration of mutants, μ the specific growth rate of the main nonmutated part of the cell population, and η the specific growth rate of neutral mutants, then

$$\frac{dM}{dt} = \lambda\mu N + \eta M - DM$$
$$\frac{dN}{dt} = \mu N - \lambda\mu N - DN \tag{94}$$

where λ is the mutation rate expressed as the ratio of the numbers of mutants to total number of cells formed. If $\lambda \ll 1$ and $\mu = \eta$, in the steady state, we obtain $\mu = \eta = D$, and

$$\frac{dM}{dt} = \lambda ND$$

or

$$M = M_0 + \lambda DNt \tag{95}$$

If $\eta > \mu$, then the original strain will be displaced by the mutant, otherwise, if $\eta < \mu$, M will tend to a lower limit $M^* = \lambda N/(1 - \eta/D)$.

Experimental studies of phage-resistant mutants in a tryptophan-limited chemostat culture of *E. coli* showed that the period of linear M increase in accordance with equation 95 was fairly short. Every 20 to 100 generations there was an abrupt fall in the number of mutants, after which the linear growth resumed at the same rate (67).

The observed saw-tooth dynamics in M were explained by Moser (43) as a combined effect of mutation and selection. The original wild clone gives not a single but a whole array of mutations with subsequent reversions. Let us denote the total cell population in a chemostat culture as N, which is the sum of all subpopulations, including original and emerging variants, $N = \Sigma N_i$. All possible transitions between variants are given by the matrix, $\lambda_{i\to j}(j = 1, \ldots, n; i = 1, \ldots, n; j \neq i)$. Then, the chemostat model takes the following form:

$$\frac{dN}{dt} = \sum_{j=1}^{n} \frac{dN_j}{dt} = \bar{\mu}(s)N - DN$$

$$\bar{\mu} = \frac{1}{N}\sum_{j=1}^{n} \mu_j N_j$$

$$\frac{dN_j}{dt} = \mu_j(s)N_j - DN_j - \sum_{i\neq j}^{n} \lambda_{j\to i}N_j + \sum_{j\neq i}^{n} \lambda_{i\to j}N_i$$

$$\frac{ds}{dt} = D(s_0 - s) - \sum_{j=1}^{n} \mu_j(s)N_j/Y_j$$

$$\mu_j(s) = \mu_m \frac{s}{K_{sj} + s} \tag{96}$$

Every drop in the saw-tooth dynamics of neutral mutants detected by their phage resistance can be interpreted as the appearance of other types of spontaneous mutant with higher growth capabilities. In a chemostat culture, such mutants overcompete and displace all other cells by virtue of their higher affinity to limiting substrate (a decreased K_s value). Let such a mutant be denoted by the subscript k and $\bar{\mu}(s)$ be the average growth rate of all other subpopulations. The selection pressure for this mutant, σ, is given by

$$\frac{d(M_k/N)}{dt} = \sigma \frac{M_k}{N} = [\mu_k(s) - \bar{\mu}(s)] \frac{M_k}{N} \tag{97}$$

If $K_{sk} < K_{sj}(j \neq k)$, then $\sigma > 0$ until a new equilibrium is established. In this process, the original cells will be displaced by the mutants and the growth-limiting substrate concentration will decrease from $\bar{s}_1 = DK_{sj}/(\mu_m - D)$ to $\bar{s}_2 = DK_{sk}/(\mu_m - D)$, and the culture density will rise by $Y(\bar{s}_1 - \bar{s}_2)$.

Autoselection in Turbidostat and pH Auxostat. The affinity to substrate was not always the only driving force of selection outcome. A number of instructive examples were reviewed by Pechurkin (14).

1. In turbidostat and pH auxostat, autoselection is in favor of mutants with higher maximum specific growth rates, $\sigma = \mu_{mk} - \mu_{mj} > 0$. Because the population density is kept constant instrumentally and the dilution rate is allowed to vary, then autoselection results in the increase in D from μ_{mj} to μ_{mk}.

2. Mutation toward a higher growth efficiency, $Y_k > Y_j$, will lead to the same result as an increase in μ_m: $\sigma = \mu_k - \mu_j = (\mu_{mk} - \mu_{mj})s/(K_{sj} + s)$, as soon as $\mu_k = q_s Y_k$ and $\mu_j = q_s Y_j$.

3. Growth of a mutant with a higher resistance to inhibitory metabolic products can be described by equation 77 with $K_{pk} > K_p$. Under selection pressure $\sigma = \mu_k - \mu = \mu_m(s/K_s + s)[K_{pk}/(K_{pk} + p) - K_p/(K_p + p)]$, the original population will be completely displaced, and the product concentration will reach a higher steady-state level, $\bar{p} = \mu_m K_{pk}\bar{s}/(K_s + \bar{s})D - K_{pk}$.

4. Growth of mutants resistant to an antibiotic will not be affected by competition if the respective antibiotic is continuously supplied: $\sigma = \mu_k(s)$, because $\mu = 0$ for all other forms. The dynamics of the total population will be governed by the initial density of the mutant.

5. A mutation resulting in enhanced adhesion to fermenter walls will lead to accumulation of slow growing cells (because the adhesion prevents washout); eventually we have $\sigma = \mu_k + D - \mu_j$.

Extrachromosomal Cell Elements. The present-day hot spot in microbial population genetics is the study of extrachromosomal cell elements (ECE) such as plasmids, phages, transposons, and insertion elements. Normally, ECEs do not carry genes, absolutely essential for growth, but they are capable of fast replication, surpassing the chromosome DNA in the number of copies. R plasmids are responsible for bacterial growth in the presence of antibiotics, but under normal conditions (with no antibiotics present), their synthesis becomes too heavy a burden for the host cell, which is manifested in a decreased growth rate. Among the more than 100 R factors studied, about a quarter were found to increase the bacterial generation time by 15% (68). For this reason, plasmid-bearing strains are unable to compete with plasmid-free populations, although there are a few exceptions. Thus, colicin-positive cells carrying respective plasmids are able to withstand the competition with faster-growing plasmid-free strains by virtue of antagonistic inhibition. In recent years, rather intricate and detailed dynamic models of autoselection have been proposed that take into account the ECE-related effects, including the transfer of ECEs within the population, their segregation loss, changes in μ arising from the ECEs carriage, and so on. Some of these models have biotechnological and medical applications (69).

MICROBIAL GROWTH AS DEPENDENT ON CULTIVATION SYSTEMS

The array of laboratory cultivation systems that define the dynamic patterns of microbial growth is summarized in Table 11. Microbial growth patterns are distinguished by three features:

- Regime of substrate supply (1, continuous supply; and 2, single-term addition)
- Elimination of growing microorganisms (α, significant; β, absent)
- Magnitude of spatial gradients (a, homogeneous systems; b, heterogenous systems)

Each specific cultivation procedure can be represented by a point inside a cube with the axes 1 to 2, α to β, and a to

Table 11. Matrix of Cultivation Techniques

	Substrate input			
	Continuous (1)		Single-term (2)	
Spatial organization	Cell eliminated (α)	No elimination (β)	Cell eliminated (a)	No elimination (b)
Homogeneous (a)	$1a\alpha$ Chemostat Turbidostat pH-auxostat	$1a\beta$ Fed-batch culture Dialysis culture Retentostat	$2a\alpha$ Continuous culture with substrate pulses Phased culture	$2a\beta$ Simple batch culture
Heterogeneous (b)	$1b\alpha$ Plug-flow (tubular culture) Colonies	$1b\beta$ Column packed with microbe attached	Forbidden combination	

b. The $2b$ combination is logically forbidden because any spatial segregation results in protracted substrate utilization and so transforms a batch process into a continuous one. The dynamics of microbial growth in any type of cultivation system can be described by the following mass–balance equations:

$$\frac{ds}{dt} = F - G(s) - \mu(s)x/Y - mx$$
$$\frac{dx}{dt} = V - H(x)x + \mu(s)x \qquad (98)$$

where F is the substrate input rate; $G(s)$ is the rate of unused substrate removal from cultivation vessel (washout, leaching, evaporation, etc.); V is the rate of microbial biomass input, which may be a single-term inoculation or continuous delivery of cells to the fermenter (specially designed or unintentional, e.g., contamination); and $H(x)$ is the rate of microbial elimination, such as washout, death, grazing, or lysis. The rest of the notation is conventional. To simulate microbial growth dynamics in a particular homogeneous system, one has to make the following selection: $F(t) = 0$, $s_0 > 0$ for a category 2 (batch culture), $F(t) > 0$ for category 1 (continuous cultivations); $H(t) = 0$ for systems retaining cell biomass (dialysis, fed-batch, simple batch, column with immobilized cells), and $H(t) > 0$ when cell elimination occurs (chemostat, turbidostat, phased culture, etc.)

Spatially heterogeneous systems can be simulated either by partial derivatives or compartmental models (e.g., the total biomass x of a microbial colony may be represented as a sum of the peripheral and central components).

1aα—Homogeneous Continuous Culture (Continuous-Flow Fermenters with Complete Mixing)

There are two subgroups within this type of cultivation technique. In the first one, steady-state growth is maintained naturally by the microbial culture itself. Self-regulation is performed through negative feedback that originates from the dependence of the growth rate on substrate concentration (chemostat) or on temperature (caloristat). In the second group, electronic devices are used for the automatic adjustment of dilution rate to the instantaneous growth rate of the microbial culture. Electronic control is based on the sensing of cell density or growth-linked products, that is, optical density (turbidostat), culture liquid viscosity (viscostat), CO_2 concentration in output air, culture pH (pH auxostat), and so on.

In theory, the steady-state growth may be established in chemostat between 0 and μ_m, but in practical terms neither very low nor very high values are attainable because of the long time needed to reach the steady state in the first case and the risk of culture washout in the second. The second group of continuous techniques (turbidostat, pH auxostat, viscostat, etc.) are capable of maintaining steady growth at high s when either $\mu \to \mu_m$ or under substrate inhibition, when $d\mu/ds < 0$. There is also a potpourri mixed technique known as the *bistat* (70), which combines a chemostat and a pH auxostat. The mass–balance equations for the chemostat and its modifications have already been given (equations 62 to 64). In a simple, complete mixing cultivator, all cells have an equal probability of being washed out, hence $\mu = D$. If there is substantial wall growth, biomass retention, or feedback, then $\mu < D$; this difference increases with the extent of biomass retention in the fermentation vessel. In terms of our scheme, such cultivation systems correspond to points on the edge $1a\alpha$ to $1a\beta$.

1aβ—Continuous Cultivation without Cell Washout

This group embraces cultures with batch or continuous dialysis, fed-batch culture (FBC), and batch culture with a supply of limiting substrate via the gas phase (gases and volatile compounds). It also includes the chemostat with complete biomass feedback by means of filtration (71). The latter fermenters are less reliable in practical terms as compared with dialysis culture, because the membrane filters are quickly plugged with cells. The limiting substrate is supplied into the dialysis culture through a semipermeable membrane and, in the case of a gaseous nutrient, through the gas–liquid interface. In both cases, mass transfer is reasonably well described by Fick's law. The culture volume remains fixed in all systems, with the exception of a FBC. In a FBC, a constant nutrient feed F provides a linear increase of culture volume V during the cultivation span; the dilution rate $D = F/V$ is decreased hyperbolically. The great advantage of these cultivation techniques for biotechnology is that they provide the possibility of realizing very slow continuous growth accompanied with derepression of synthesis of many secondary metabolites. With a constant limiting substrate flux, Fs_0,

the absence of cell washout means that at each subsequent moment an equal ration is shared by an increasing microbial biomass, and, as a result, μ eventually falls down to negligible values or even to zero (the maintenance state). Unlike the chemostat, no true steady state is established in this case, but when the substrate is virtually depleted, we have $ds/dt \approx 0$ and the system reaches a quasi steady state (5). If a quasi steady state approximation is found to be sufficient, then extremely slow continuous growth can be obtained after a reasonable period of time, perhaps a few weeks, as compared to the several months needed in a chemostat.

2aα—Continuous Cultivation with a Discontinuous Supply of Limiting Substrate

Suppose we have a simple chemostat culture fed by nutrient medium lacking just one essential component. This component is added as a small volume of concentrated solution at regular and sufficiently large time intervals, Dt. Then

$$\frac{ds}{dt} = D[s_0(t) - s] - q(s)x$$

$$\frac{dx}{dt} = \mu(s)x - Dx$$

$$s_0(t) = \begin{cases} A > 0, \text{ when } t = i\Delta t,\ i = 0, 1, \ldots, n \\ 0, \text{ otherwise} \end{cases} \qquad (99)$$

This cultivation method was originally used to obtain synchronized cell division (72). A continuous phased culture (73) is a repeated simple batch culture, that is, at regular intervals, Δt, half of the culture volume is withdrawn and the fermenter is refilled with an equal volume of fresh medium. Under such conditions, repeated batchwise growth proceeds with biomass increasing cyclically from x_0 to $2x_0$. Obviously, the growth dynamics are governed by the time interval between consecutive substrate additions Δt. If $\Delta t = \ln 2/\mu_m$, a sawtooth nonlimited growth takes place as in a turbidostat. With $\Delta t < \ln 2/\mu_m$ the culture should be washed out and, when $\Delta t > \ln 2/\mu_m$, the maximal attainable biomass decreases with increasing Δt because of endogenous biomass decomposition and waste respiration during the lag phase.

2aβ—Simple Batch Culture

Cultivation begins at the initial limiting substrate concentration, s_0, and inoculum size, x_0. The biomass reaches its maximum, x_m, when the limiting substrate is depleted ($s = 0$) and then declines even in the absence of exogenous elimination, so that $x \to 0$ as $t \to \infty$. Description of batch dynamics has been given earlier. Specified growth phases are described by simple nonstructured models (equations 56, 61, and 73), and entire dynamics are described by SCM and other structured models (equations 84 to 86; Fig. 8).

1bα—Plug-Flow (Tubular) Culture

The inoculum and the medium are mixed on entry into a long reactor tube, and the culture flows in the tube at a constant velocity without mixing. In each small element of culture liquid, moving along the spatial coordinate z at a linear velocity $f = F/A$ (where F is flow rate, cm^3/h, and A is the cross-sectional area, cm^3), growth of biomass proceeds as in a simple batch culture. To account for the culture movement per se, we have to pass from ordinary to partial derivatives and replace dx/dt by $\partial x/\partial t + f\partial x/\partial z$ (along-the flow-growth rate). Allowing for some dispersion of the moving front by diffusion, we can write

$$\frac{\partial s}{\partial t} + f\frac{\partial s}{\partial z} = D_s\frac{\partial^2 s}{\partial z^2} - q(s)x$$

$$\frac{\partial x}{\partial t} + f\frac{\partial x}{\partial z} = D_x\frac{\partial^2 x}{\partial z^2} + \mu(s)x \qquad (100)$$

where D_s and D_x are the diffusion coefficients of the substrate and microbial cells, respectively.

1bβ—Continuous-Flow Reactors with Microbes Attached

The nutrient solution is pumped through a column filled with adsorbent material and is utilized as it moves by growing immobilized cells. The mass–balance equations for a packed column are obtained from equation 100 by simplifying the equation for x,

$$\frac{\partial s}{\partial t} + f\frac{\partial s}{\partial z} = D_s\frac{\partial^2 s}{\partial z^2} - q(s)x(z)$$

$$\frac{\partial x}{\partial t} = \mu(s)x = Y[q(s) - m]x \qquad (101)$$

Bacterial cells accumulate faster on the top of the column because of larger $q(s)$ values, and as a result, a distinctive spatial biomass distribution develops in the form of a hyperbolic decrease of x with column depth. Growth does not reach steady state with respect to x until cell elimination becomes well expressed (the effects of inhibitory products, endogenous cell decomposition, and leaching).

Colonies

Besides continuous-flow columns, other heterogeneous systems are widely used. Especially popular is plating on solid media made from natural or synthetic gels (agar, PAAG, silica gel, synthetic alumosilicates, etc.) as well as on some porous materials (sand, glass beads, glass fiber). Impregnated with nutrient solution, such materials are used to grow microorganisms in the form of colonies or lawns. At first glance, it is tempting to consider these techniques as analogies of a simple batch culture with single-term substrate input (type 2bβ in Table 11). However, a closer look at the mechanism of colony growth reveals a greater resemblance to chemostat culture! Here we will outline our considerations.

The spread of a colony over a solid substrate, for example, a layer of agar, proceeds by the growth of only a peripheral zone with biomass, x_p (Fig. 10). Then

$$\frac{ds}{dt} = -q(s)x$$

$$\frac{dx}{dt} = \mu_w x_p - ax \qquad (102)$$

where a is the specific decay rate of cell (mycelium) com-

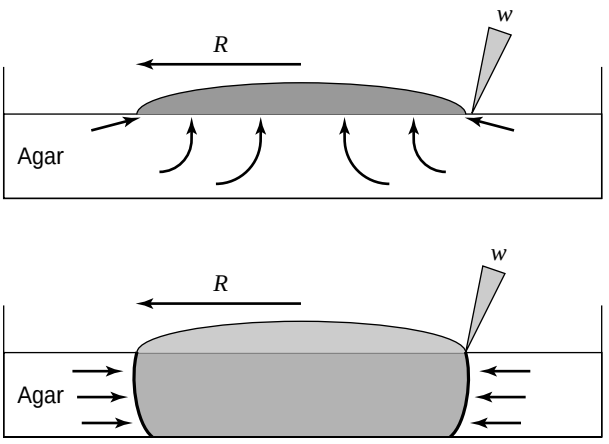


Figure 10. Schematic illustration of the difference between the growth modes bacterial (*top*) and fungi (*bottom*) colonies. Note that unicellular organisms (*bacteria*) do not penetrate into the agar layer as opposed to filamentous fungi. *Arrows* indicate directions of substrate diffusion across concentration gradients. *R* and *w* are, respectively, radius and width of the peripheral zone of the colony.

ponents. Straightforward geometrical analysis shows that the linear spread rate for a regularly shaped colony (a cylinder, a sphere, or a strip) of size R is given by the following relation (2,5):

$$\frac{dR}{dt} = \mu_w w = K_r$$

or

$$R = R_0 + K_r t \qquad (103)$$

where K_r is the colony linear expansion rate, mm/h; μ_w is the microbial specific growth rate within the peripheral zone; and w is the zone width.

In the case of unicellular organisms (e.g., bacteria, yeasts) that are incapable of penetrating into the gel's matrix, the substrate is available through passive diffusion to the peripheral zone from the underlying gel layer. As the colony grows, this flux diminishes, K_r decreases, and eventually the growth stops altogether. The filamentous organisms (fungi and actinomycetes) are able to propagate both on the surface and in the depth of gel. As a result, their growth is not completely dominated by diffusion effects, and the colony front advances at a faster rate than substrate is depleted in the frontier zone. The colony spreads at a constant radial rate, K_r, until the Petri dish is filled or the agar deteriorates from dryness. (In the case of rich nutrient media, there are also effects of self-inhibition by metabolic products [see equations 76 and 77].)

Thus, the colony growth is (*1*) continuous, (*2*) substrate-limited, (*3*) directed, and (*4*) spatially ordered. Properties 1 and 2 suggest a strong similarity between growth of a colony (especially a fungal one) and that of a chemostat culture. If so, a colony's peripheral zone is analogous to the cell culture in the fermenter, and its central part is comparable with the waste cell suspension that is discharged into the product bottle. A steady running of a pump, delivering medium at a rate, F (cm^3/h), corresponds to the active and uniform substrate utilization by the growing mycelium at a rate $K_r A$ (cm^3/h, A is the area of colony–medium interface). Finally, both cultivation systems are characterized by the elimination of propagating biomass, that is, by the expulsion of grown cells (mycelium) from further active growth (either by washing out or by the motion of the colony front). In both cases, the elimination rate is equal to biomass growth in the active compartment. The essential deviation from the chemostat lies in properties 3 and 4. Spatial differentiation of hypha and the direction of colony expansion are governed by spatial gradients of limiting substrate and, possibly, some metabolic products. Steepness of gradients is partly diminished by the effects of metabolic translocation along hypha over distances of the order of w.

It can be concluded from the previous discussion that colony growth belongs to class $1b\alpha$, and not to $2b\beta$. In general, it is very likely that the $2b$ combination is an empty, logically forbidden combination, because a single-term momentary input of substrate may only be realized in a homogeneous system. Any spatial segregation whatever will actually prolong the consumption of substrate and, therefore, transform a batch process into a continuous one. For this reason, growth on any insoluble substrate (lignocellulose and other plant polymers, oil droplets, grains of sulfur, etc.) should always be treated as a continuous process. The solid-phase fermentation can also not be anything but continuous, whether new portions of substrate are added to the reactor or not. (Obviously, this applies to the growth mechanism itself and not to the engineering operation.)

BIBLIOGRAPHY

1. A.A. Esener, J.A. Roels, and N.W.F. Kossen, *Biotechnol. Bioeng.* **25**, 2803–2841 (1983).
2. N.S. Panikov, *Microbial Growth Kinetics*, Chapman & Hall, London, 1995.
3. I.G. Minkevich and V.K. Eroshin, *Uspekhi Sovrem. Biol.* **82**, 103–116 (1976).
4. W.R. Mayberry, G.J. Prochazka, and W.J. Payne, *Appl. Microbiol.* **15**, 1332–1338 (1967).
5. IS.J. Pirt, *Principles of Microbe and Cell Cultivation*, Blackwell Scientific Publishers, Oxford, U.K., 1975.
6. A.G. Marr, E.H. Nilson, and D.J. Clark, *Ann. N.Y. Acad. Sci.* **102**, 536–548 (1963).
7. A.G. Dorofeyev and N.S. Panikov, *Mikrobiologiya–Engl. Tr.* **64**, 601–609 (1994).
8. D.W. Tempest and O.M. Neijssel, *Annu. Rev. Microbiol.* **38**, 459–486 (1984).
9. S.J. Pirt, *Arch. Microbiol.* **133**, 300–302 (1982).
10. N. Panikov and J.S. Pirt, *J. Gen. Microbiol.* **108**, 295–303 (1978).
11. T. Bauchop and S.R. Elsden, *J. Gen. Microbiol.* **23**, 457–469 (1960).
12. V.N. Ivanov, *Energetics of Microbial Growth*, Naukova Dumka, Kiev, 1981 (in Russian).
13. W.M. van Gulik, and J.J. Heijnen, *Biotechnol. Bioeng.* **48**, 681–698 (1996).

14. N.S. Pechurkin, *Populational Microbiology*, Novosibirsk, Nauka 1978, p. 277.
15. R.E. Buchanan, *J. Infect. Dis.* **23**, 109–125 (1918).
16. G. Stephanopoulos, and J.J. Vallino, *Science* **252**, 1675–1680 (1991).
17. Y. Kashiwaya, K. Sato, N. Tsuchiya, S. Thomas, D.A. Fell, R.L. Veech, and J.V. Passonneau, *J. Biol. Chem.* **269**, 25502–25514 (1994).
18. J.J. Vallino, and G. Stephanopoulos, *Biotechnol. Bioeng.* **41**, 633–646 (1993).
19. A. Marx, A.A. de Graaf, W. Wiechert, L. Eggeling, and H. Sahm, *Biotechnol. Bioeng.* **49**, 111–119 (1996).
20. A. Cornish-Bowden, *Principles of Enzyme Kinetics*. Butterworths, London-Boston, 1976.
21. N. Michaelis and M.I. Menten, *Biochem. Z.* **49**, 333–369 (1913).
22. G.E. Briggs and J.B.S. Haldane, *Biochem. J.* **19**, 338–339 (1925).
23. H. Gutfreund, *Discuss. Faraday Soc.* **20**, 167–173 (1955).
24. J.B.S. Haldane, *Enzymes*. Longmans Green, London, 1930.
25. J. Monod, J. Wyman, and J. Changeux, *J. Mol. Biol.* **12**, 88–118 (1965).
26. C.H. Wang and A.L. Koch, *Bacteriol.* **136**, 969–975 (1978).
27. O. Rahn, *Physiology of Bacteria*, P. Blackiston's, Philadelphia, 1932.
28. P.F. Verhulst, *Correspondence Mathematique et Physique* **10**(1), 113–121 (1838).
29. J. Monod, *Recherches sur la Croissance des Cultures Bacteriennes*, Paris, Hermann and Co, 1942.
30. J. Monod, *Ann. Inst. Pasteur* **79**, 390–410 (1950).
31. F.F. Blackman, *Ann. Bot.–London* **19**, 281–293 (1905).
32. S.D. Varfolomeyev, in S.D. Varfolomeyev ed., *Modern Problems of Biokinetics*, Moscow University Publishers, Moscow, 1987, pp. 6–77.
33. J. Monod, *Ann. Rev. Microbiol.* **3**, 371–394 (1949).
34. N.D. Jerusalimsky, in E.O. Powell, C.G.T. Evans, R.E. Strange, and D.W. Tempest eds., *Continuous Cultivation of Microorganisms*, H.M. Stationery Office, Salisbury, U.K., 1967, pp. 23–33.
35. I.A. Shvytov, *Zhurnal Obshchei Biologii* **35**, 904–910 (1974).
36. J.M. Romanovsky, N.V. Stepanova, and D.S. Chemavski, *Mathematical Biophysics*, Moscow, Nauka, 1984.
37. T.E. Shehata and A.G. Marr, *J. Bacteriol.* **107**, 210–216 (1971).
38. A.F. Guady Jr., A. Obayashi, and E.T. Gaudy, *Appl. Microbiol.* **22**, 1041–1047 (1971).
39. D.K. Button, *Microbiol. Rev.* **49**, 270–297 (1985).
40. B.R. Robertson and D.K. Button, *Bacteriology* **138**, 884–895 (1979).
41. G. Teissier, *Ann. Physiol. Physicochem. Biol.* **12**, 527–586 (1956).
42. D.E. Contois, *J. Gen. Microbiol.* **21**, 40–49 (1959).
43. H. Moser, *Carnegie Inst. Wash. Pub.* **614** (1958).
44. E.O. Powell, in E.O. Powell, C.G.T. Evans, R.E. Strange, and D.W. Tempest eds., *Continuous Cultivation of Microorganisms*, H.M. Stationery Office, Salisbury, U.K., 1967, pp. 34–55.
45. J.F. Andrews, *Biotechnol. Bioeng.* **10**, 707–723 (1968).
46. U. Pawlowsky, S.A. Howell, and C.T. Chi, *Biotechnol. Bioeng.* **15**, 905–916 (1973).
47. A. Campbell, *Bacteriol. Rev.* **21**, 263–272 (1957).
48. C.L. Cooney, *Ann. N.Y. Acad. Sci.* **326**, 295–314 (1979).
49. E.O. Powell, C.G.T. Evans, R.E. Strange, and D.W. Tempest eds., *Continuous Cultivation of Microorganisms*, H.M. Stationery Office, Salisbury, U.K., 1967, pp. 259–261.
50. D. Ramkrishna, A.G. Fredrickson, and H.M. Tsuchiya, *Biotech. Bioeng.* **9**, 129–170 (1967).
51. M.M. Domach, S.K. Leung, R.E. Cahn, G.G. Cacks, and M.L. Shuler, *Biotech. Bioeng.*, **26**, 203–216 (1984).
52. J.W. Jeong, J. Snay, and M.M. Ataai, *Biotechnol. Bioeng.* **35**, 160–184 (1990).
53. I. Malek, in I. Malek ed., *Continuous Cultivation of Microorganisms: A symposium*, Publishing House of Czech. Acad. Sci., 1958, pp. 11–28.
54. D. Herbert, *Symp. Soc. Gen. Microbiol.* **11**, 391–416 (1961).
55. P. Agrawal, H.C. Lim, and D. Ramkrishna, *J. Chem. Tech. Biotechnol.* **33B**, 155–163 (1983).
56. D.S. Kompala, D. Ramkrishna, N.B. Jansen, and G.T. Tsao, *Biotechnol. Bioeng.* **28**, 1044–1055 (1986).
57. B.G. Turner, D. Ramkrishna, N.B. Jansen, *Biotechnol. Bioeng.* **34**, 252–261 (1989).
58. D.W. Tempest, J.R. Hunter, and J. Sykes, *J. Gen. Microbiol.* **39**, 355–366 (1965).
59. D.W. Tempest, D. Herbert, and P.J. Phipps, in E.O. Powell, C.G.T. Evans, R.E. Strange, and D.W. Tempest eds., *Continuous Cultivation of Microorganisms*, H.M. Stationery Office Salisbury, U.K., 1967, pp. 240–253.
60. A.T. Henrici, *Morphologic Variation and the Rate of Growth of Bacteria*, Charles C. Thomas, Springfield, Ill., 1928, p. 194.
61. A.L. Koch, *Crit. Rev. Microbiol.* **19**, 17–42 (1973).
62. S. Wold, K. Skarstad, H.B. Steen, T. Stokke, E. Boye, *EMBO J.* **13**, 2097–2102 (1994).
63. C.E. Helmstetter, S. Cooper, O. Pierucci, and E. Revela, *Cold Springs Harb. Symp. Quant. Biol.* **33**, 809–822 (1968).
64. S. Cooper, *Res. Microbiol.* **141**, 17–29 (1990).
65. A.L. Koch, *J. Gen. Microbiol.* **45**, 409–417 (1966).
66. J.F. Collins, and M.H. Richmond, *J. Gen. Microbiol.* **28**, 15–33 (1962).
67. A. Novick, in G. Tunevall ed., *Recent Progress in Microbiology*, Blackwell Scientific Publications, Oxford, U.K., 1959, pp. 403–415.
68. J. Adams, T. Kinney, S. Thompson, L. Rubin, and R.B. Helling, *Genetics* **91**, 627–637 (1979).
69. R. Freter, R.R. Freter, and H. Brickner, *Infect. Immunol.* **39**, 60–84 (1983).
70. I.G. Minkevich, A.Y. Krynitskaya, and V.K. Eroshin, *Mikrobiologiya (Engl. Tr.)* **56**, 254–258 (1987).
71. W.R. Chesbro, T. Evans, and R. Eifert, *J. Bacteriol.* **139**, 625–638 (1979).
72. B.C. Goodwin, *Eur. J. Biochem.* **10**, 511–514 (1969).
73. P.S.S. Dawson, *Can. J. Microbiol.* **11**, 893–903 (1965).

See also MICROBIAL GROWTH MEASUREMENT, METHODS.

LACCASE

FENG XU
Novo Nordisk Biotech
Davis, California

KEY WORDS

Application
Biocatalysis
Cloning
Expression
Inhibition
Kinetics
Laccase
Properties
Structure

OUTLINE

MOLECULAR PROPERTIES, ENZYMOLOGY, AND PREPARATION OF LACCASE

Detection, Purification, and Cloning of Laccase

Laccase (EC 1.10.3.2) is a family of multicopper oxidases that catalyze the oxidation of a range of inorganic and aromatic substances (particularly phenols) with the concomitant reduction of O_2 to water. This enzyme has been detected in various plant species, including lacquer, mango, mung bean, peach, pine, prune, and sycamore (1–6), one bacterium, (*Azospirillum lipoferum*) (7), a dozen insects of genera that include *Bombyx, Calliphora, Diploptera, Drosophila, Lucilia, Manduca, Musca, Oryctes, Papilio, Phormia, Rhodnius, Sarcophaga, Schistocerca,* and *Tenebrio* (8–10), as well as more than 40 different fungi, including *Agaricus, Antrodiella, Armillaria, Aspergillus, Bjerkandera, Botrytis, Ceriporiopsis, Cerrena, Chaetomium, Coprinus, Cryphonectria, Cryptococcus, Curvularia, Cyathus, Geotrichum, Daedalea, Fomes, Fusarium, Halosarpheia, Lactarius, Lentinus, Monocillium, Myceliophthora, Neurospora, Penicillium, Phanerochaete, Phellinus, Phlebia, Pholiota, Pleurotus, Podospora, Pycnoporus, Pyricularia, Rhizoctonia, Rigidoporus, Schizophyllum, Sclerotium, Scytalidium, Sporotrichum, Stagonospora, Thermoascus, Trametes,* and *Trichoderma* (11–42). It has been postulated that laccase is involved in various cellular and microbial activities (43–51). Recent studies on the physiological function of laccase include those on plant cell wall biosynthesis (52), phytopathogenesis (53), wood material degradation and humification (32,54), insect sclerotization (8,10,55,56), bacterial melanization (7), and melanin-related virulence for human (57).

To date, thorough purification and characterization have been made on laccases from more than 20 different organisms (Table 1). The primary structure has been determined for laccases from 14 organisms, including *Acer pseudoplatanus* (58), *Agaricus bisporus* (11), *Aspergillus nidulans* (77), *Coprinus cinereus* (78), *Cryptococcus neoformans* (57), *Myceliophthora thermophila* (23), *Neurospora crassa* (64), *Pleurotus ostreatus* (36), *Rhizoctonia solani* (26), *Scytalidium thermophilum* (30), *Trametes (Coriolus) hirsutus* (79), *Trametes versicolor* (80), *Trametes villosa* (33), and a basidiomycete designated as PM1 (25). Extensive characterization of laccase has been carried out in the past decades, and the work accumulated by 1993 has been comprehensively reviewed (43–51). An attempt is made in this article to review recent progress, with emphasis on the potential industrial and medicinal applications of laccase, made during the last decade.

Molecular Properties of Laccase

Most laccases studied are extracellular proteins, although intracellular laccases have been detected in several fungi and insects (8,10,17,57). In general, fungal laccases have isoelectric points (pI) ranging from 3 to 7 whereas plant laccase pI ranges to 9 (Table 1). Under nondenaturing conditions, gel filtration data suggest that at least some fungal laccases have quaternary structure (8,11,15,23,26,30, 33,69,81). On SDS-PAGE, most laccases show mobilities corresponding to molecular weight (MW) of 60–100 kDa, of which 10–50% may be attributed to glycosylation (Table 1). Available data indicate that mannose is one of the major components of the carbohydrates attached to laccase (1,26,33,36,57,60,69,81,82). Glycosylation in laccase is suggested to play a role in secretion, proteolytic susceptibility, activity, copper retention, and thermal stability (60,69,82,83). Purified laccases exhibit a characteristic

Table 1. Molecular Properties of Some Laccases

Species	MW, kDa	pI	Glycosylation (%)	References
Acer pseudoplatanus	66[a]		40	1,58
Agaricus bisporus	65[a]	3.7	15	11
Armillaria mellea I, II	59	4.1, 3.3	15	13
Aspergillus nidulan	80		12	59
Botrytis cinerea	74	4.0	49	60
Cerrena maxima	67			46
Cerrena unicolor	66			14
Ceriporiopsis subvermispora I, II	71, 68	3.4, 4.8	15	61
Chaetomium thermophile	36			15
Cryphonectria parasitica	77		24	17
Cryptococcus neoformans	75[a]		6	57
Fomes annosus	73			21
Junghuhnia separabilima	58–62	3.5		62
Lactarius piperatus	67			63
Lentinus edodes	66			47
Manduca sexta	70–90			8
Monocillium indicum	72			22
Myceliophthora thermophila	85[a]	4.2	14	23
Neurospora crassa	64[a]	6.8	12	64
Phanerochaete chrysosporium	47			24
Phellinus noxius	70			65
Phlebia radiata	64	3.5	2	66
Phlebia tremellosa	67	3.3		62
Pholiota mutabilis	64			21
Pinus taeda	90	9	22	4
Pleurotus ostreatus I, II	64, 59[a]	2.9	13,4	36,67
PM1	64[a]	3.6	6.5	25,68
Podospora anserina	70–80	4.9	23	21,69
Pycnoporus cinnabarinus	77	3.7	9	35
Pycnoporus coccineus	70	3.5	8	70
Rhizoctonia pratocola	78			21
Rhizoctonia solani	66[a]	7.5	11	26
Rhus vernicifera	110	8.6	45	71,72
Rigidoporus lignosus I, II	53, 52	3.8, 3.3		62,65
Schizophyllum commune	64			28
Scytalidium thermophilum	80[a]	5.1	20	30
Trametes anisoporus	58			46
Trametes consors	58			73
Trametes hirsutus	55[a]			46
Trametes sanguinea	62	3.5	9.1	74
Trametes versicolor I, II	67[a], 70	3.2, 5.6	14	75,76
Trametes villosa I, III	63[a]	3.5, 6.5	0.5	33
Trichoderma	72			34

[a]DNA sequence-deduced MW: *Acer pseudoplatanus*, 60; *Agaricus bisporus*, 58; *Cryptococcus neoformans*, 66; *Myceliophthora thermophila*, 61; *Neurospora crassa*, 61; *Pleurotus ostreatus*, 54; PM1, 53; *Rhizoctonia solani*, 55; *Scytalidium thermophilum*, 63; *Trametes hirsutus*, 54; *Trametes versicolor*, 55; *Trametes villosa*, 55.

blue appearance from their electronic absorption around 600 nm. Typical UV-visible spectra of laccase (at resting state) show two maxima around 280 and 600 nm and one shoulder near 330 nm. The ratio of the absorbance at 280 nm to that at 600 nm is generally 14 to 30, and the ratio of the absorbance at 330 nm to that at 600 nm is 0.5 to 2 (26,33,43–51,84).

In the holoenzyme form, most laccases have four copper atoms per monomer (25,26,33,43–51,57,61,67, 70,72–74,85), although the laccase from *Phlebia* was reported to have two coppers and one pyrroloquinoline quinone as prosthetic groups (86). Among the four copper atoms, one belongs to the "blue" type 1 copper site that has a strong electronic absorption around 600 nm and a characteristic electron paramagnetic resonance (EPR) signal; one belongs to the type 2 copper site that has a characteristic EPR signal, and two belong to the strongly coupled type 3 site that has a weak UV absorption around 330 nm and is EPR active only when perturbed by strong anion binding. It is postulated that the type 2 and type 3 sites constitute a trinuclear copper cluster center. The redox potentials of the active coppers measured for several fungal and plant laccases are in the range of 0.4–0.8 V versus normal hydrogen electrode (43–48,84). The redox potential

of the type 1 site in several plant and fungal laccases has been measured as a function of pH and showed a total potential change of 0.1 V over the pH range 3–9 (71,87).

The primary structures of more than 20 laccases (from more than 14 organisms) have been determined by protein or DNA sequencing. Analyses of the primary structure indicate the occurrence of N-terminal posttranslational processing in which the signal (pre-) peptide is cleaved during the secretion of an extracellular laccase. The maturation of some laccases also includes either N-terminal processing of a propeptide or blocking of the N-terminus (23,30). A C-terminal propeptide is processed to form mature *Neurospora crassa* laccase as well (64). Within the highly conserved copper-binding domains, the putative copper-ligating residues (10 histidines and 1 cysteine) are strictly conserved. Preliminary crystallization of two *Trametes* laccases has been reported (74). Extensive spectroscopic characterizations and primary sequence comparison (43–48,88) strongly suggest that, in laccase, the structures of the copper sites and polypeptide backbone are quite similar to those of ascorbate oxidase and ceruloplasmin as determined by X-ray crystallography (89,90).

Enzymology of Laccase

The enzymology of two representative laccases, *Rhus vernicifera* and *Trametes versicolor* laccases, has been extensively studied (43–51). Laccase can catalyze the oxidation of various reducing substances such as inorganic or organic metal complexes, anilines, thiols, and especially phenols. Although laccase strongly prefers O_2 as its oxidizing substrate (K_m at submillimolar levels), it usually has low specificity toward reducing substrates. In general, the optimal pH of laccase activity depends on the type of substrate (87). For phenols, the optimal pH can range from 3 to 7 for fungal laccases and up to 9 for plant laccases. The optimal pH, K_m, and k_{cat} for some laccases are listed in Table 2.

It is believed that laccase catalysis involves (*1*) reduction of the type 1 copper by reducing substrate, (*2*) internal electron transfer from the type 1 copper to the type 2 and type 3 copper, and (*3*) reduction of O_2 to water at the type 2 and type 3 copper site (43–51). The oxidation of a reducing substrate by laccase typically involves the loss of a single electron and the formation of a free (cation) radical. The radical is in general unstable and may undergo further laccase-catalyzed oxidation (e.g., to form quinone from phenol) or nonenzymatic reactions (e.g., hydration, disproportion, or polymerization). The electron transfer from substrate to type 1 copper is probably controlled by redox potential difference (95). A lower oxidation potential of substrate or a higher redox potential of laccase (type 1 site) often results in a higher rate for substrate oxidation. It seems that the binding pocket of reducing substrate (or the type 1 copper site) is quite shallow and has limited steric effect on simple phenol substrate (78,95). In contrast, the O_2-binding pocket (or the type 2 and type 3 copper sites) appears to restrict the access of oxidizing agents other than O_2. Activation of O_2 likely involves chemical bond formation on the trinuclear copper cluster. Under turnover conditions, the rate-limiting step may be the oxidation of substrate, whereas under transient or anaerobic conditions, the internal electron transfer step may be rate limiting (43–51).

Laccase can be inhibited by various reagents (43–51). Small anions such as halides [excluding iodide (96)], azide, cyanide, and hydroxide bind to the type 2 and type 3 copper, resulting in an interruption of the internal electron transfer and activity inhibition. Other inhibitors include metal ions (e.g., Hg^{2+}), fatty acids, sulfhydryl reagents, hydroxyglycine, Kojic acid, desferal, and cationic quaternary ammonium detergents (15,18,21,25,61,97–99). Laccase is generally more stable at alkaline pH than at acidic pH (74,84). Under identical conditions, thermophilic fungal laccases are in general more thermostable than mesophilic laccases (84).

Preparation of Laccase

In general plant laccases are purified from sap or tissue extracts, whereas extracellular fungal laccases are purified from medium culture (fermentation broth) of the selected organism. Certain aromatic or phenolic inducers (such as ferulic acid, xylidine, *p*-anisidine, or anthraquinone) can be used in some cases to boost the yield of an inducible laccase from the wild-type strains (21,27,35,76,100). Recent progress in applying genetic engineering technologies may be an alternative route to industrial laccase production (26,33,79,84,101–104). Various protein purification techniques are frequently employed in purifying laccase. Typical purification protocols involve ultrafiltration, ion exchange, gel filtration, hydrophobic interaction, or other electrophoretic and chromatographic techniques. Affinity chromatography using a phenolic group as ligand can increase purification efficiency (91). The purity of a laccase preparation is often measured by SDS-PAGE and by the ratio of the absorbance at 280 nm to that at 600 nm. Laccase transcription or expression can be detected by DNA–RNA hybridization or by immunochemical assay with antilaccase antibody. Laccase activity can be detected with characteristic chromogenic substrates, such as syringaldazine and 2,2′-azino-bis-(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS) are often used. ABTS is very suitable to screening because its one-electron-oxidation product is soluble in water, stable, and intensely green. Currently, several wild-type fungal laccases, including those from *Pyricularia oryzae, Trametes (Coriolus) consors, Myceliophthora thermophila, Trametes hirsutus,* and *Trametes versicolor,* are commercially available.

APPLICATIONS OF LACCASE

Industrial Biocatalysis

Delignification of Lignocellulosics by Laccase. Conventional methods of delignifying or decolorizing paper pulp involve either chlorine- or oxygen-based chemical oxidants (e.g., ClO_2 and O_3). Although very effective, these methods have serious drawbacks such as disposal of chlorinated by-products or loss of cellulose fiber strength. Microbial or enzymatic delignification systems that overcome these draw-

Table 2. Kinetic Parameters of Some Laccases

Species	Substrate	pH	K_m, mM	k_{cat}/min	References
Armillaria mellea	*p*-phenylenediamine	3.5	1.7		13
Botrytis cinerea	2,6-Dimethylphenol	3.5	0.1	466,000	60
Ceriporiopsis subvermispora I, II	Guaiacol	3, 5	1.6, 0.44	3,100, 4,000	61
Cerrena unicolor	Syringaldazine	5.5			14
Cryphonectria parasitica	2,6-Dimethylphenol	2.5		2,000	17
Curvularia	Guaiacol	5	0.75		18
Fomes annosus	Syringaldazine	4.6			21
Fomes fomentarius	Syringaldazine	4.6	0.28		91
Lactarius piperatus	Catechol	5.6			63
Lentinus edodes	Syringaldazine	5.0			21
Manduca sexta	2-Methyhydroquinone	6.0	0.02	1,100	8
Monocillium indicum	*o*-Dianisidine	3.0	0.025		22
Myceliophthora thermophila	Syringaldazine	7	0.01	370	84
Phanerochaete chrysosporium	Syringaldazine	5.0			21
Phellinus noxius	Guaiacol	4.0			65
Pholiota mutabilis	Syringaldazine	5.2			21
Pinus taeda	Coniferyl alcohol	5.9			4
Pleurotus ostreatus I, II	Syringaldazine	6.5		192, 107	67
PM1	Guaiacol	4.5	0.5		68
Podospora anserina	Dopamine	5.5, 7.5	3.3		81
Pycnoporus cinnabarinus	Guaiacol	4	0.75		35,92
Pycnoporus coccineus	Catecol	4.5			70
Rhizoctonia praticola	2,6-Dimethylphenol	6.8	0.26		21
Rhizoctonia solani	Syringaldazine	7	0.017	1,900	84
Rhus vernicifera	Syringaldazine	9	0.043	600	84
Rigidoporus lignosus	Syringaldazine	6.0		1200	93
Sarcophaga bullata	Dopamine	4.5			10
Scytalidium thermophilum	Syringaldazine	7	0.001	290	84
Sporotrichum pulverulentum	Syringaldazine	5.0			21
Trametes anceps	Syringaldazine	5.3		140	94
Trametes hirsutus	Guaiacol		0.75		92
Trametes sanguinea	Dimethylphenylenediamine	5			74
Trametes versicolor I, II	Catechol	4.6	0.02, 0.09	52, 110	85
Trametes villosa I	Syringaldazine	5	0.058	2700	84

backs and can be easily adapted to current pulp production lines are attractive alternatives.

It has long been postulated that laccase participates in natural delignification (for recent reviews, see Refs. 66,76,105,106). Various laccases have been shown capable of degrading either natural lignin (wood chips or Kraft pulp) or synthetic lignin models (32,66,76,105–112). These laccases could either directly oxidize phenolic components exposed in lignin or, in the presence of proper redox mediator, indirectly oxidize heterogeneous phenolic and nonphenolic (especially methoxybenzene) components (35,112–114). As a result of laccase oxidation, (cationic) radicals could be generated in lignin, resulting in subsequent aliphatic or aromatic C–C bond cleavage and depolymerization.

As a "biopulping" agent, laccase could be applied to wood chips before pulping. The pretreatment with laccase could partially degrade lignin and loosen lignin structure so that the pulping can be done more efficiently. In one example, the use of a laccase-producing *Antrodiella* strain in a biopulping pretreatment step resulted in 26% energy savings in the subsequent mechanical pulping step (12). As a "biobleaching" agent, laccase could be applied (as an alternative to conventional chemical oxidants) to pulp. The residual lignin in pulp could be degraded by laccase, resulting in pulp decolorization. In one example, a *Trametes versicolor* laccase, under the mediation of *N*-hydroxy compounds, delignified Kraft pulp effectively in pilot-plant scale (115,116).

Cross-Linking of Polysaccharides by Laccase Polymerized saccharides are useful materials for producing composites such as particle board and liner board. However, most chemicals used in conventional polymerization, such as urea, formaldehyde, isocyanate, and petrochemical resins, are hazardous. Based on its lignin oxidation property, laccase could be used to replace these chemicals and serve as a "biogluing" agent (117). To initiate or enhance the cross-linking (gluing) efficiency, laccase could be used in three ways. The first is to directly oxidize wood particles or pulp to generate more radicals for cross-linking (117). The second is to functionalize wood particles or pulp with small compounds (such as aromatic, carboxyl, isocyanate, and acrylamide substances) that act as cross-linking agents (118). The third is to transform isolated lignin (often a by-product from pulping), starch, phenolic polysaccharide, or protein into radical-rich and nontoxic adhesive materials useful for wood composites (119). Such applications of laccase not only could replace toxic or expensive chemical ad-

hesives but also could transform wastes such as lignin from the paper industry into value-added product.

Waste Detoxification and Decontamination by Laccase. Laccase has been used to oxidatively detoxify or remove various aromatic xenobiotics and pollutants found in industrial waste and contaminated soil or water (103,120,121). The laccase catalysis could result in either direct degradation or polymerization/immobilization. Reported examples of direct degradation by laccase includes dechlorination of chlorophenols (122–124), cleavage of aromatic ring (125), mineralization of polycyclic aromatic hydrocarbons (126), decolorization of pulp or cotton mill effluent (37,127–129), and bleaching of textile dyes (38,115,116,130). The processes include polymerization among pollutants themselves or copolymerization with other nontoxic substances (such as humic materials). Polymerized pollutants often become insoluble or immobilized, thus facilitating easy removal by such means as adsorption, sedimentation, or filtration (131,132).

Textile Dye Transformation by Laccase. Laccase has been reported to prevent back-staining of dyed or printed textiles. As part of the washing solution, laccase could quickly bleach released dye stuff, thus resulting in the reduction of processing time, energy, and water needed to achieve satisfactory quality of the textile (133).

Laccase-catalyzed textile dye-bleaching may also be useful in finishing dyed cotton fabric. Replacing conventional chemical oxidants (e.g., hypochlorite), a laccase-based system has been shown capable of bleaching indigo dye in denim and achieving various bleached appearances on the fabric (134,135).

Another interesting laccase application in this field is oxidative transformation and consequent coupling of dye precursors to the collagen matrix in hides (136). Under laccase catalysis, soluble dye precursors could be adsorbed, oxidized, and polymerized to give the desired tanning effect. The process could improve dyeing efficiency, reduce cost (by using inexpensive precursors), or provide improved hide characteristics.

Medical and Personal Care Applications of Laccase

Poison ivy dermatitis (resulting from skin contact with poison ivy, poison oak, poison sumac, and the like) is caused mainly by urushiol, which is a catechol-derivative toxin. Oxidized urushiol (an *o*-quinone derivative), however, is nontoxic. Laccase has been shown to oxidize, polymerize, and detoxify urushiol, thus reducing the effect of poison ivy dermatitis (137).

Another potential laccase application in the field is laccase-based generation of iodine in situ. Laccase can oxidize iodide to produce iodine (96), a reagent widely used as a disinfectant. The application of a laccase–iodide salt binary iodine-generating system (for sterilization) may have several advantages over the direct iodine application. First, the iodide salt is more stable and much safer than I_2 in terms of storage, transport, and handling. Second, the release of iodine from a laccase–iodide system could be easily controlled (by means such as adjusting laccase concentration). The system may be used in various industrial, medical, domestic, and personal care applications such as sterilization of drinking water and swimming pools as well as disinfection of minor wounds (96,121,138).

Current hair dyeing or waving processes often involve oxidative or additive chemicals that either have unpleasant odors, are irritant (or even harmful) to tissues, or are difficult to handle. A laccase-based system may overcome these drawbacks by replacing harsh chemicals and operating at milder conditions (in terms of pH and solvent). Laccase-catalyzed oxidation, transformation, and cross-linking of various precursors (mostly phenols and anilines) have been reported to result in satisfactory hair dyeing or waving (139–143). In addition to providing an easier-to-handle hair care procedure, a laccase-based system may also improve or complement the cosmetic effect (in terms of color type, shade, and compatibility with hair type) achieved by conventional chemical methods.

Biosensor and Diagnostic Applications of Laccase

Detection of Phenol, Aniline, O_2, and Other Substances. Laccase catalysis (coupled with various physical instruments) could be useful as biosensors for detecting O_2 and a wide variety of reducing substrates (especially phenols or anilines). Two types of laccase-based O_2 sensor are widely used. One type monitors visible spectral changes (at 600 nm) of laccase resulting from the reoxidation of the type 1 copper (I) in laccase by O_2 (144). Another type monitors current or voltage change from a modified oxygen electrode on which O_2 reduction is enhanced under the electrocatalysis of immobilized laccase (46). For detecting phenols, anilines, or other reducing substrates, three types of laccase-based sensors have been reported (46,145,146). One type detects the photometric change resulted from the oxidation of a chromogenic substrate, one type monitors the O_2 concentration change that is coupled to the substrate oxidation, and one type uses an electrode that replaces O_2 as the acceptor for the electrons flown from the substrate (through laccase). For these applications, laccase is either immobilized or free in solution, and the coupled physical converter is either optic, amperometric, or piezo effect in nature or a field-effect transitor or thermister, etc (46).

Laccase-Assisted Enzymatic and Immunochemical Assays. Laccase catalysis can be used to assay other enzymes. In these assays, either the enzyme of interest catalyzes the production of a compound whose subsequent oxidation by laccase generates detectable physical change, or a product from a laccase catalysis (whose production is accompanied by a detectable physical change) is quenched by the activity of the enzyme of interest. The strategy has been applied to assay various enzymes including amylase, (alanine, cystine, or leucine-specific) aminopeptidases, alkaline phosphatase, angiotensin I converting enzyme, arylamidase, cellobiose oxidase, chymotrypsin, glucosidase, γ-glutamyl transpeptidase, kallikrein, plasmin, thrombin, and X factor (as a coagulant factor) (147–151).

Laccase, that is covalently conjugated to an antibody or antigen can be used as a marker enzyme for immuno-

chemical assay. In this application, binding of the antibody (or antigen) to its immunological counterpart is detected by localized laccase activity on a gel or a blot membrane, much like the conventional peroxidase- or phosphatase-assisted immunochemical assays (46). Under certain conditions, the antibody–antigen binding impairs the function of conjugated laccase, thus allowing immunochemical detection through modulation of laccase activity (43,152).

Organic and Medicinal Syntheses Catalyzed by Laccase

Laccase catalysis can be used to synthesize various functional organic compounds, which include polymers with specific mechanical, electrical, or optical properties (68, 153,154), textile dyes (155), cosmetic pigments (156,157), flavor agents (158), aromatic aldehydes (159), pesticides (160), and heterocyclic compounds (such as phenoxazine) (92). The reported applications involve laccase-catalyzed oxidation of aromatic compounds that result in either polymerization (68,92,153–157), transformation (159), or depolymerization (158). Laccase can also be used to synthesize several complex medicinal agents, which include triazolo(benzo)thiadiazines (a group of antiinflammatory, analgesic, central stimulant/depressant antisecretory and sedative depressant agents) (161), vinblastine (a cytostatic, antitumor drug) (46), penicillin X dimer (162), cephalosporin antibiotics (163), and dimerized vindoline (a plant alkaloid for treating neoplastic diseases) (164). Laccase can be used as well for oxidative deprotection in peptide synthesis (46). In comparison with counterpart chemical synthetic methods, a laccase-catalyzed oxidation in general requires milder conditions (in terms of temperature, pH, solvent nature), may improve regioselectivity, is easier to control, and is more environmentally friendly.

Other Miscellaneous Applications of Laccase

Beverage and Food Treatment. Browning and haze formation during the processing and storage of juice or wine are major problems for the industry. It is believed that phenolic compounds are involved in this process. Conventionally, undesirable phenolics are adsorbed and removed by various fining agents (e.g., gelatin, bentonite), which usually have low specificity, may affect color or aroma, and can pose disposal problems. Laccase may be used to remove or modify problematic phenolic saccharides and to improve clarity, color appearance, flavor, aroma, taste, and stability in fruit juice or fermented alcohol beverages (165). Following treatment by laccase, oxidized and polymerized (or precipitated) unwanted phenolic substances could be removed by silicate fining or filtration.

Laccase activity has been suspected to play an undesirable role in the processing of certain foods or beverages. Excessive or uncontrolled laccase activity, along with the action of other phenol oxidases, could cause browning of canned foods or vegetables, deteriorate the quality of wine or juice, or participate in various fruit pathogenesis (53). In these cases, inhibition of laccase activity has economic importance. Potential inhibitors include ferulic acid (166), oxalic acid (167), and cucurbitacin (168). In the case of food or beverage canning, elimination of O_2 can also be used.

Laccase may be applied to certain processings that enhance or modify the color appearance of a food or beverage. One interesting case involves ripe olive processing in which laccase replaces the conventional lye solution and oxidatively polymerizes various phenolics (such as oleuropein) in the olive (169), resulting in color darkening and debittering.

Desulfurization of Fossil Fuels. To reduce the harmful emission of sulfur-containing chemicals from the combustion of fossil fuels (petroleum, liquefied coal, natural gas), desulfurization is necessary during post combustion or, especially, precombustion processing of fuels. Conventional chemical and physical desulfurization methods (such as oxidizing and solubilizing inorganic pyrite sulfur by ferric salt or chlorine) often have low efficiency toward organic sulfur compounds, require extreme conditions (high temperature, high-pressure, and corrosion-compatible equipment), or require high infrastructure investment and maintenance costs. As an alternative to chemical methods, enzymatic desulfurization could be achieved under milder conditions, command less infrastructure cost, and remove organic sulfur more effectively. An enzymatic system capable of desulfurizing coal has been demonstrated (170) in which inorganic and organic sulfur compounds in coal slurry are oxidized by laccase, solubilized in water, and separated from solid (desulfurized) coal particles.

OVERALL REMARKS

Preferred Application Modes

Immobilization. Although native laccase functions in solution, preferred forms of laccase in various industrial or diagnostic applications should be immobilized or supported by nonliquid matrices. These include encapsulation in gel (hydrogel, alginate, silica gel), entrapment in micelles, or immobilization (either covalently or noncovalently) on solid materials (clay, glass, zeolite, active carbon, etc.) (91,132,165,171–176). In general, proper immobilization or encapsulation does not much alter the specific activity of laccase. Immobilized laccase may be suited for nonaqueous applications. Immobilized laccase could be easily recovered and recycled, thus reducing laccase dosage (or cost), making the treatment suitable to continuous operations, and avoiding the need to inactive residual laccase after treatment.

Mediation by Small Redox Compounds. For applications that target lignocellulosics or other insoluble materials, a laccase catalysis is heterogeneous. To enhance or extend laccase catalysis, small redox mediators may be used. Suitable mediators, such as 2,2′-azino-bis-(3-ethylbenzthiazoline-6-sulfonic acid) (112,114), phenoxazine (92, 117,149,160,177), *N*-hydroxy/oxides (115), Mn(III)-pyrophosphate (178), and Remazol blue (179), could significantly enhance catalysis by effectively shuttling electrons between laccase and the insoluble substrate, as well as enabling laccase to indirectly oxidize compounds that are normally not substrate for laccase.

Synergy with Other Enzymes. Laccase rarely acts alone and often requires synergy from other enzymes in complex natural processes (such as wood degradation). To better mimic nature, other enzymes should be tested to enhance laccase catalysis. For instance, flavin-containing oxidoreductases (180–182), xylanase (183), and manganese peroxidase (66,93,178) could be used together with laccase in delignifying pulp. These enzymes may provide a range of beneficial synergies for laccase catalysis, such as modifying targeted substances into better laccase substrates, or eliminating those products that are inhibitors of laccase catalysis.

Advantages of Applying Laccase

Live microorganisms that produce laccase or other active enzymes have been tested with several of the applications previously mentioned: these include *Thiobacillus* and *Sulfolobus* bacteria in coal desulfurization (170) and white-rot fungi *Trametes* and *Pleurotus* in delignification (112). In comparison with individual enzymes such as laccase, live microorganism cells could provide an array of enzymes capable of acting in synergy and with high efficiency. However, applying purified or enriched laccase preparation may hold several advantages over live microbial cells (120). Laccase could (*1*) be insensitive to many compounds or microorganisms that are toxic for or prey on the selected microbes; (*2*) be compatible with a wide range of environmental conditions (no need for specific nutrients or optimal growth conditions that may differ from optimal operational conditions); (*3*) be superior at a low contaminant concentration; (*4*) have controlled specificity; (*5*) be effective in terms of diffusion, permeability, and mobility; or (*6*) be able to accomplish the task in a shorter time.

Prospect of Laccase Applications

As an oxidase, laccase may be used in many industrial and medicinal applications where oxidation or oxidative derivatization is involved. Currently, investigations are focused on laccase-based biooxidation, biotransformation, biosensor, and enzymatic synthesis. The applications of laccase discussed here have all been satisfactorily carried out in either laboratory or pilot plant scale. However, an important factor that prevents large-scale laccase application as applied for pulp bleaching (115) is the unrealistically high dosage of laccase (and mediator), which could potentially be reduced by further optimizing the catalysis.

To overcome this limiting factor, one needs to better understand the mechanism for the reaction of interest. In the case of applying laccase to pulp, for example, one needs to further elucidate how glycolignin is degraded by white-rot fungi (and their secreted enzymes). In addition to studying the natural processes, protein engineering is probably also needed to optimize catalysis of laccase in various applications. For instance, developing a laccase that has better activity at alkaline pH or high temperature is of importance for pulp application. Protein engineering requires a breakthrough in obtaining three-dimensional structure information (78) and elucidating key structure–function correlations such as those that govern the redox potential and pH activity profile in laccase.

ACKNOWLEDGMENTS

I thank Alan V. Klotz, Stephen H. Brown, Debbie S. Yaver, Glenn E. Nedwin, Ejner B. Jensen, Ture Damus, Tomas T. Hansen, Claus Felby, and Donald Boyce of Novo Nordisk for critical reading and helpful suggestions.

BIBLIOGRAPHY

1. A. Driouich, A.-C. Laine, B. Vian, and L. Faye, *Plant. J.* **2**, 13–24 (1992).
2. S.P. Robinson, B.R. Loveys, and E.K. Chacko, *Aust. J. Plant Physiol.* **20**, 99–107 (1993).
3. A. Chabanet, A.M. Catesson, and R. Goldberg, *Phytochemistry* **33**, 759–763 (1993).
4. W. Bao, D.M. O'Malley, R. Whetten, and R.R. Sederoff, *Science* **260**, 672–674 (1993).
5. L. Dijkstra and J.R.L. Walker, *J. Sci. Food. Agric.* **54**, 229–234 (1991).
6. E. Lehman, E. Harel, and A.M. Mayer, *Phytochemistry* **13**, 1713–1717 (1974).
7. D. Faure, M.L. Boullant, and R. Bally, *Appl. Environ. Microbiol.* **60**, 3412–3415 (1994).
8. B.R. Thomas, M. Yonekura, T.D. Morgan, T.H. Czapla, T.L. Hopkins, and K.J. Kramer, *Insect Biochem.* **19**, 611–622 (1989).
9. M. Sugumaran and K. Nellaiappan, *Arch. Insect Biochem. Physiol.* **15**, 165–181 (1990).
10. F.M. Barrett, *Can. J. Zool.* **65**, 1158–1166 (1987).
11. C.R. Perry, M. Smith, C.H. Britnell, D.A. Wood, and C.F. Thurston, *J. Gen. Microbiol.* **139**, 1209–1218 (1993).
12. R. Patel, G.D. Thakker, and K.K. Rao, *J. Biotechnol.* **36**, 19–23 (1994).
13. A.U. Rehman and C.F. Thurston, *J. Gen. Microbiol.* **138**, 1251–1257 (1992).
14. E.G. Bekker, S.D. Petrova, O.V. Ermolova, V.I. Elisashvili, and A.P. Sinitsyn, *Biokhimiya* **55**, 2019–2024 (1990).
15. T. Ishigami, Y. Hirose, and Y. Yamada, *J. Gen. Appl. Microbiol.* **34**, 401–407 (1988).
16. P. Vnenchak and M.N. Schwalb, *Mycol. Res.* **93**, 546–548 (1989).
17. D. Rigling and N.K. Van Alfen, *Appl. Environ. Microbiol.* **59**, 3634–3639 (1993).
18. U.C. Banerjee and R.M. Vohra, *Folia Microbiol.* **36**, 343–346 (1991).
19. M.J.R. Shannon, and R. Bartha, *Appl. Environ. Microbiol.* **54**, 1719–1723 (1988).
20. D.S. Arora and D.K. Sandhu, *Enzyme Microb. Technol.* **7**, 405–408 (1985).
21. J.-M. Bollag and A. Leonowicz, *Appl. Environ. Microbiol.* **48**, 849–854 (1984).
22. G.D. Thakker, C.S. Evans, and K.K. Rao, *Appl. Microb. Biotechnol.* **37**, 321–323 (1992).
23. International Pat. WO 9533836 A1 (December 14, 1995), R.M. Berka, S.H. Brown, F. Xu, P. Schneider, D.A. Aaslyng, and K.M. Oxenbøll (to Novo Nordisk Biotech).
24. C. Srinivasan, T.M. D'Souza, K. Boominathan, and C.A. Reddy, *Appl. Environ. Microbiol.* **61**, 4274–4277 (1995).
25. P.M. Coll, C. Tabernero, R. Santamaria, and P. Perez, *Appl. Environ. Microbiol.* **59**, 4129–4135 (1993).
26. J.A. Wahleithner, F. Xu, K.M. Brown, S.H. Brown, E.J. Golightly, T. Halkier, S. Kauppinen, A. Pederson, and P. Schneider, *Curr. Genet.* **29**, 395–403 (1996).

27. S. Ding, C. Wang, and J.A. Buswell, *Xianweisu Kexue Yu Jishu* **2**, 36–46 (1994).
28. O.M.H. De Vries, W.H.C.F. Kooistra, and J.G.H. Wessels, *J. Gen. Microbiol.* **132**, 2817–2826 (1986).
29. I. Chet and A. Huttermann, *FEMS Microbiol. Lett.* **14**, 211–215 (1982).
30. International Pat. WO 9533837 A1 (December 14, 1995), R.M. Berka, F. Xu, and S.A. Thompson (to Novo Nordisk Biotech).
31. M. Bergbauer and C. Eggert, *Can. J. Microbiol.* **40**, 192–197 (1994).
32. A. Machuca and N. Duran, *Appl. Biochem. Biotechnol.* **43**, 37–44 (1993).
33. D.S. Yaver, F. Xu, E.J. Golightly, K.M. Brown, S.H. Brown, M.W. Rey, P. Schneider, T. Halkier, K. Mondorf, and H. Dalbøge, *Appl. Environ. Microbiol.* **62**, 834–841 (1996).
34. A. Assavanig, B. Amornkitticharoen, N. Ekpaisal, V. Meevootisom, and T.W. Flegel, *Appl. Microbiol. Biotechnol.* **38**, 198–202 (1992).
35. C. Eggert, U. Temp, and K.-E. Eriksson, *Appl. Environ. Microbiol.* **62**, 1151–1158 (1996).
36. P. Giardina, V. Aurilia, R. Cannio, L. Marzullo, A. Amoresano, R. Siciliano, P. Pucci, and G. Sannia, *Eur. J. Biochem.* **235**, 508–515 (1996).
37. C. Raghukumar, D. Chandramohan, F.C. Michel, and C.A. Reddy, *Biotechnol. Lett.* **18**, 105–106 (1996).
38. K. Vasdev, R.C. Kuhad, and R.K. Saxena, *Curr. Microbiol.* **30**, 269–272 (1995).
39. M. Svabova, J. Zemek, J. Augustin, L. Kuniak, and L. Marvanova, *Folia Microbiol.* **25**, 233–237 (1980).
40. A. Rodriguez, M.A. Falcon, A. Carnicero, F. Perestelo, G. De la Fuente, and J. Trojanowski, *Appl. Microbiol. Technol.* **45**, 399–403 (1996).
41. P. Ander and K.-E. Eriksson, *Arch. Microbiol.* **109**, 1–8 (1976).
42. J. Zhou, F. Zhang, and M. Kuwahara, *Acta Microbiol. Sin.* **33**, 387–391 (1993).
43. A. Messerschmidt, *Adv. Inorg. Chem.* **40**, 121–185 (1994).
44. B. Reinhammar and B.G. Malmström, in T.G. Spiro, ed., *Copper Proteins*, Wiley, New York, 1981, pp. 109–149.
45. A.M. Mayer, *Phytochemistry* **26**, 11–20 (1987).
46. A.I. Yaropolov, O.V. Skorobogat'ko, S.S. Vartanov, and S.D. Varfolomeyev, *Appl. Biochem. Biotechnol.* **49**, 257–280 (1994).
47. C.F. Thurston, *Microbiology* **140**, 19–26 (1994).
48. F.S. Sariaslani, *Crit. Rev. Biotechnol.* **9**, 171–257 (1987).
49. J.E. Coleman, in J.R. Whitaker, ed., *Food-Related Enzymes*, American Chemical Society, Washington, D.C., 1973, pp. 267–304.
50. J.A. Fee, *Struct. Bonding* **23**, 2–60 (1975).
51. E.I. Solomon, M.J. Baldwin, and M.D. Lowery, *Chem. Rev.* **92**, 521–542 (1992).
52. D.M. O'Malley, R. Whetten, W. Bao, C.-L. Chen, and R. Sederoff, *Plant J.* **4**, 751–757 (1993).
53. M. Sbaghi, P. Jeandet, R. Bessis, and P. Leroux, *Plant Pathol.* **45**, 139–144 (1996).
54. T. Katase and J.-M. Bollag, *Soil Sci.* **151**, 291–296 (1991).
55. M. Miessner, O. Crescenzi, A. Napolitano, G. Prota, S.O. Andersen, and M.G. Peter, *Helv. Chim. Acta* **74**, 1205–1212 (1991).
56. S.O. Andersen, in G.A. Kerkut and L.I. Gilbert, eds., *Comprehensive Insect Physiology, Biochemistry, and Pharmacology*, vol. 3, Pergamon Press, Oxford, 1985, pp. 59–74.
57. P.R. Williamson, *J. Bacteriol.* **176**, 656–664 (1994).
58. P.R. LaFayette, K.-E.L. Eriksson, and J.F.D. Dean, *Plant Physiol.* **107**, 667–668 (1995).
59. M.B. Kurtz and S.P. Champe, *J. Bacteriol.* **151**, 1338–1345 (1982).
60. D. Slomczynski, J.P. Nakas, and S.W. Tanenbaum, *Appl. Environ. Microbiol.* **61**, 907–912 (1995).
61. Y. Fukushima and T.K. Kirk, *Appl. Environ. Microbiol.* **61**, 872–876 (1995).
62. T. Vares, O. Niemenmaa, and A. Hatakka, *Appl. Environ. Microbiol.* **60**, 569–575 (1994).
63. H. Iwasaki, T. Matsubara, and T. Mori, *J. Biochem. (Tokyo)* **61**, 814–816 (1967).
64. U.A. Germann, G. Muller, P.E. Hunziker, and K. Lerch, *J. Biol. Chem.* **263**, 885–896 (1988).
65. J.P. Geiger, B. Rio, D. Nandris, and M. Nicole, *Appl. Biochem. Biotechnol.* **12**, 121–133 (1986).
66. A. Hatakka, *FEMS Microbiol. Rev.* **13**, 125–135 (1994).
67. H.-D. Youn, K.-J. Kim, J.-S. Maeng, Y.-H. Han, I.-B. Jeong, G. Jeong, S.-O. Kang, and Y.C. Hah, *Microbiology* **141**, 393–398 (1995).
68. P.M. Coll, J.M. Fernandez-Abalos, J.R. Villanuva, R. Santamaria, and P. Perez, *Appl. Environ. Microbiol.* **59**, 2607–2613 (1993).
69. W. Minuth, M. Klischies, and K. Esser, *Eur. J. Biochem.* **90**, 73–82 (1978).
70. Y. Oda, K. Adachi, I. Aita, M. Ito, Y. Aso, and H. Igarashi, *Agric. Biol. Chem.* **55**, 1393–1395 (1991).
71. T. Nakamura, *Biochim. Biophys. Acta* **30**, 44–52 (1958).
72. B. Reinhammar, *Biochim. Biophys. Acta* **205**, 35–47 (1970).
73. T. Sakurai, *J. Inorg. Biochem.* **41**, 277–281 (1991).
74. Y. Nishzawa, K. Nakabayashi, and E. Shinagawa, *J. Ferment. Bioeng.* **80**, 91–93 (1995).
75. M. Jonsson, E. Pettersson, and B. Reinhammar, *Acta. Chim. Scand.* **22**, 2135–2140 (1968).
76. R. Bourbonnais, M.G. Paice, I.D. Reid, P. Lanthier, and M. Yaguchi, *Appl. Environ. Microbiol.* **61**, 1876–1880 (1995).
77. R. Aramayo and W.E. Timberlake, *Nucleic Acids Res.* **18**, 3415 (1990).
78. V. Ducros, A.M. Brzozowski, K.S. Wilson, S.H. Brown, P. Østergaard, P. Schneider, D.S. Yaver, A.H. Pedersen, and G.J. Davies, *Nature Structural Biol.* **5**, 310–316 (1998).
79. Y. Kojima, Y. Tsukuda, Y. Kawai, A. Tsukamoto, J. Sugiura, M. Sakaino, and Y. Kita, *J. Biol. Chem.* **265**, 15224–15230 (1990).
80. L. Jönsson, K. Sjöström, I. Häggström, and P.O. Nyman, *Biochim. Biophys. Acta* **1251**, 210–215 (1995).
81. H. Molitoris and K. Esser, *Arch. Mikrobiol.* **72**, 267–296 (1970).
82. A. Yoshitake, Y. Katayama, M. Nakamura, Y. Iimura, S. Kawai, and N. Morohoshi, *J. Gen. Microbiol.* **139**, 179–185 (1993).
83. M.T. Graziani, L. Antonilli, P. Sganga, G. Citro, B. Mondovi, and M.A. Rosei, *Biochem. Int.* **21**, 1113–1124 (1990).
84. F. Xu, W. Shin, S.H. Brown, J.A. Wahleitner, U.M. Sundaram, and E.I. Solomon, *Biochim. Biophys. Acta* **1292**, 303–311 (1996).
85. M.A. Pickard and D.W.S. Westlake, *Can. J. Biochem.* **48**, 1351–1358 (1970).
86. E. Karhunen, M.-L. Niku-Paavola, L. Viikari, T. Haltia, R.A. Van der Meer, and J.A. Duine, *FEBS Lett.* **267**, 6–8 (1990).
87. F. Xu, *J. Biol. Chem.* **272**, 924–928 (1997).

88. A. Messerschmidt and R. Huber, *Eur. J. Biochem*. **187**, 341–352 (1990).
89. A. Messerschmidt, R. Ladenstein, R. Huber, M. Bolognesi, L. Avigliano, R. Petruzzelli, A. Rossi, and A. Finazzi-Agro, *J. Mol. Biol*. **224**, 179–205 (1992).
90. I. Zaitseva, V. Zaitsev, C. Card, K. Moshkov, B. Bax, A. Ralph, and P. Lindley, *J. Biol. Inorg. Chem*. **1**, 15–23 (1996).
91. J. Rogalsky, M. Wojtas-Wasilewska, R. Apalovic, and A. Leonowicz, *Biotechnol. Bioeng.* **37**, 770–777 (1991).
92. C. Eggert, U. Temp, J.F.D. Dean, and E.-K.L. Eriksson, *FEBS Lett.* **376**, 202–206 (1995).
93. H. Galliano, G. Gas, J.L. Seris, and A.M. Boudet, *Enzyme Microb. Technol.* **13**, 478–482 (1991).
94. R.J. Petroski, W. Peczynska-Czoch, and J.P. Rosazza, *Appl. Environ. Microbiol.* **40**, 1003–1006 (1980).
95. F. Xu, *Biochemistry* **35**, 7608–7614 (1996).
96. F. Xu, *Appl. Biochem. Biotechnol.* **59**, 221–230 (1996).
97. T. Kreuter, A. Steudel, and H. Pickert, *Acta Biotechnol.* **11**, 81–83 (1991).
98. S. Murao, Y. Hinode, E. Matsumura, A. Numata, K. Kawai, H. Ohishi, H. Jin, H. Oyama, and T. Shin, *Biosci. Biotechnol. Biochem*. **56**, 987–988 (1992).
99. M.C. De Pinto and A.R. Bacelo, *Phytochemistry* **42**, 283–286 (1996).
100. F. Sollai, N. Curreli, M.C. Porcu, A. Rescigno, A.C. Rinaldi, A. Rinaldi, P. Rossino, G. Soddu, and E. Sanjust, *Biochem. Arch*. **12**, 7–12 (1996).
101. C.R. Perry, S.E. Matcham, D.A. Wood, and C.F. Thurston, *J. Gen. Microbiol.* **139**, 171–178 (1993).
102. B. Schilling, R.M. Linden, U. Kupper, and K. Lerch, *Curr. Genet.* **22**, 197–203 (1992).
103. International Pat. WO9201046 (January 23, 1992), M. Saloheimo, M.-L. Niku-Paavola, M. Penttila, J. Knowles, and A. Kantelinen (to Valtion Teknillinen Tutkimuskeskus).
104. Germany Pat. DE4141901 A1 (June 24, 1993), K. Crabbe and A. Cordes.
105. P. Bajpai and P.K. Bajpai, *J. Biotechnol.* **33**, 211–220 (1994).
106. V.I. Elisashvili, *Microbiology* **62**, 480–487 (1993).
107. A.M. Bonnen, L.H. Anton, and A.B. Orth, *Appl. Environ. Microbiol.* **60**, 960–965 (1994).
108. C. Ruttimann-Johnson, L. Salas, R. Vicuna, and T.K. Kirk, *Appl. Environ. Microbiol.* **59**, 1792–1797 (1993).
109. Y. Iimura, Y. Katayama, S. Kawai, and N. Morohoshi, *Biosci. Biotechnol. Biochem*. **59**, 903–905 (1995).
110. K. Ruel, K. Ambert, and J.-P. Joseleau, FEMS *Microbiol. Rev.* **13**, 241–254 (1994).
111. R. Kondo, T. Iimori, H. Imamura, and T. Nishida, *J. Biotechnol.* **13**, 181–188 (1990).
112. M.G. Paice, R. Bourbonnais, I.D. Reid, F.S. Archibald, and L. Jurasek, *J. Pulp Paper Sci.* **21**, J280–J284 (1995).
113. International Pat. WO 9209741 (June 11, 1992), M. Vaheri and O. Piirainen (to Enso-Gutzeit Oy).
114. A. Muheim, A. Fiechter, P.J. Harvey, and H.E. Shoemaker, *Holzforschung* **46**, 121–126 (1992).
115. International Pat. WO9429425 (December 22, 1994), H.-P. Call (to Fitzner).
116. Germany Pat. DE4137761 A1 (November 19, 1992), H.-P. Call.
117. International Pat. WO 9603546 A1 (February 8, 1996), L.S. Pedersen and C. Felby (to Novo Nordisk A/S).
118. O. Milstein, A. Huttermann, R. Frund, and H.-D. Ludemann, *Appl. Microbiol. Biotechnol.* **40**, 760–767 (1994).
119. A. Haars, A. Kharazipour, H. Zanker, and A. Huttermann, *ACS Symp. Ser.* **385**, 126–134 (1989).
120. P. Nannipieri and J.-M. Bollag, *J. Environ. Qual.* **20**, 510–517 (1991).
121. U.S. Pat. 4478683 (October 23, 1984), S.A. Orndorff (to Westvaco Corp.).
122. J. Dec and J.-M. Bollag, *Environ. Sci. Technol.* **29**, 657–663 (1995).
123. L. Roy-Arcand and F.S. Archibald, *Enzyme Microb. Technol.* **13**, 194–203 (1991).
124. Y. Iimura, P. Hartikainen, and K. Tatsumi, *Appl. Microbiol. Biotechnol.* **45**, 434–439 (1996).
125. S. Kawai, T. Umezawa, M. Shimada, and T. Higuchi, *FEBS Lett.* **236**, 309–311 (1988).
126. L. Bezalel, Y. Hadar, and C.A. Cerniglia, *Appl. Environ. Microbiol.* **62**, 292–295 (1996).
127. P. Manzanares, S. Fajardo, and C. Martin, *J. Biotechnol.* **43**, 125–132 (1995).
128. S. Davis and R.G. Burns, *Appl. Microb. Biotechnol.* **32**, 721–726 (1990).
129. U.S. Pat. 5407577 (April 18, 1995), N.P. Nghiem (to Nalco Chemical Co.).
130. M. Chivukula and V. Renganathan, *Appl. Environ. Microbiol.* **61**, 4374–4377 (1995).
131. J.C. Roper, J.M. Sarkar, J. Dec, and J.-M. Bollag, *Water Res.* **29**, 2720–2724 (1995).
132. C. Crecchio, P. Ruggiero, and M.D.R. Pizzigallo, *Biotechnol. Bioeng.* **48**, 585–591 (1995).
133. International Pat. 9218687 (October 29, 1992), G. Pedersen and M. Schmidt (to Novo Nordisk A/S).
134. International Pat. 9612845 A1 (May 2, 1996), A.H. Pedersen and J.V. Kierulff (to Novo Nordisk A/S).
135. International Pat. 9612846 A1 (May 2, 1996), A.H. Pedersen and J.V. Kierulff (to Novo Nordisk A/S).
136. International Pat. WO 9312259 (June 24, 1993), N.H. Sørensen, K. Ingvorsen, and A. Møllgaard (to Novo Nordisk A/S).
137. U.S. Pat. 4259318 (March 31, 1981), N.V. Duhe and D.L. Hendrix (to the Regents of the University of Houston).
138. International Pat. WO9616165 A1 (May 30, 1996), F. Xu (to Novo Nordisk Biotech).
139. Y. Tsujino, Y. Yokoo, and K. Sakato, *J. Soc. Cosm. Chem*. **42**, 273–282 (1991).
140. European Pat. EP 504005 A1 (September 16, 1992), M. Roure, P. Delattre, and H. Froger (to Perma).
141. French Pat. FR 2694018 A1 (January 28, 1994), R. Martin, A. Jumino, C. Dubief, G. Rosenbaum, and M.-P. Audousset (to L'Oreal SA).
142. Japenese Pat. JP9177813 A2 (April 3, 1991) R. Saruno.
143. Germany Pat. DE2155359 (May 25, 1972) E. Zeffren and J.F. Sullivan (to Procter and Gamble Co.).
144. A.E. Gardiol, R.J. Hernandez, B. Reinhammar, and B.R. Harte, *Enzyme Microb. Technol.* **18**, 347–352 (1996).
145. N. Zouari, J.-L. Romette, and D. Thomas, *Biotechnol. Tech.* **8**, 503–508 (1994).
146. A.L. Ghindilis, A. Makower, C.G. Bauer, F.F. Bier, and F.W. Scheller, *Anal. Chim. Acta* **304**, 25–31 (1995).
147. S. Murao, M. Arai, N. Tanaka, H. Ishikawa, K. Matsumoto, and S. Watanabe, *Agric. Biol. Chem*. **49**, 981–985 (1985).

148. H. Agematu, T. Nakashima, N. Shibamoto, T. Yoshioka, T. Shin, and S. Murao, *J. Ferment. Bioeng.* **77**, 479–482 (1994).

149. B.P. Roy and F. Archibald, *Anal. Biochem.* **216**, 291–298 (1994).

150. F.F. Bier, E. Ehrentreich-Föster, C.G. Bauer, and F.W. Scheller, *Fresenius J. Anal. Chem.* **354**, 861–865 (1996).

151. U.S. Pat. 5196312 (March 23, 1993), A. Miike and T. Tatano (to Kyowa Medix Co. Ltd.).

152. Germany Pat. DE4314417 A1 (November 10, 1994), F. Scheller, U. Wollenberger, and A. Makower (to Byk Gulden Italia).

153. J.A. Akkara, M. Ayyagari, F. Bruno, L. Samuelson, V.T. John, C. Karayigitoglu, S. Tripathy, K.A. Marx, D.V.G.L.N. Rao, and D.L. Kaplan, *Biomimetics* **2**, 331–339 (1994).

154. R. Ikeda, H. Uyama, and S. Kobayashi, *Macromolecules* **29**, 3053–3054 (1996).

155. W.L. Baker, K. Sabapathy, M. Vibat, and G. Lonergan, *Enzyme Microb. Technol.* **18**, 90–94 (1996).

156. French Pat. FR2694021 (January 28, 1994), A. Junino, and R. Martin (to l'Oreal SA).

157. W. Cai, R. Martin, B. Lemaure, J.-L. Leuba, and V. Petiard, *Plant Physiol. Biochem.* **31**, 441–445 (1993).

158. B. Falconnier, C. Lapierre, I. Lesage-Meessen, G. Yonnet, P. Brunerie, B. Colonna-Ceccaldi, G. Corrieu, and M. Asther, *J. Biotechnol.* **37**, 123–132 (1994).

159. A. Potthast, T. Rosenau, C.-L. Chen, and J.S. Graftzel, *J. Org. Chem.* **60**, 4320–4321 (1995).

160. F.S. Sariaslani, J.M. Beale, Jr., and J.P. Rosazza, *J. Nat. Prod.* **47**, 692–697 (1984).

161. U.T. Bhalerao, C. Muralikrishna, and B.R. Rani, *Tetrahedron* **50**, 4019–4024 (1994).

162. H. Agematu, T. Tsuchida, K. Kominato, N. Shibamoto, T. Yoshioka, H. Nishida, and R. Okamoto, *J. Antibiot. (Tokyo)* **46**, 141–148 (1993).

163. H. Agematu, K. Kominato, N. Shibamoto, T. Yoshioka, H. Nishida, R. Okamoto, T. Shin, and S. Murao, *Biosci. Biotechnol. Biochem.* **57**, 1387–1388 (1993).

164. F.S. Sariaslani, F.M. Eckenrode, J.M. Beale, Jr., and J.P. Rosazza, *J. Med. Chem.* **27**, 749–754 (1984).

165. A. Zamorani, P. Spettoli, A. Lante, A. Crapisi, and G. Pasini, *Ital. J. Food Sci.* **Spec. Issue**, 195–200 (1995).

166. F.O. Asiegbu, A. Paterson, and J.E. Smith, *World J. Microbiol. Biotechnol.* **12**, 16–21 (1996).

167. C.B.S. Tong, K.B. Hicks, S.F. Osman, A.T. Hotchkiss, Jr., and R.M. Haines, *J. Agric. Food Chem.* **43**, 592–597 (1995).

168. A. Viterbo, R.C. Staples, B. Yagen, and A.M. Mayer, *Phytochemistry* **35**, 1137–1142 (1994).

169. *Debittering and Darkening of Olives by Laccase Treatment*, vol. 378, Novo Nordisk A/S Res. Discl., Novo Nordisk A/S, Bagsvœrd, Denmark, 1995, p. P676 (no. 37828).

170. U.S. Pat. 5094668 (March 10, 1992), E.E. Kern, W.M. Menger, D.A. Odelson, A.S. Sinsky, D.L. Wise, and D.J. Trantolo (to Houston Industries Inc.).

171. O. Milstein, B. Nicklas, and A. Huttermann, *Appl. Microbiol. Biotechnol.* **31**, 70–74 (1989).

172. S. Davis and R.G. Burns, *Appl. Microbiol. Biotechnol.* **37**, 474–479 (1992).

173. A.V. Vakurov, A.K. Gladilin, A.V. Levashov, and Y. Khmelnitsky, *Biotechnol. Lett.* **16**, 175–178 (1994).

174. S. Cliffe, M.S. Fawer, G. Maier, K. Takata, and G. Ritter, *J. Agric. Food Chem.* **42**, 1824–1828 (1994).

175. G. Palmieri, P. Giardina, B. Desiderio, L. Marzullo, M. Giamberini, and G. Sannia, *Enzyme Microb. Technol.* **16**, 151–158 (1994).

176. U.S. Pat. 5443975 (Aug. 22, 1995), V. Capuano and S. Cervelli (to Consigtio Nazionale delle Richerche IT and Ente per le Nuove Tecnologie l'Energia e l'Ambiente ENEA IT).

177. International Pat. WO 9501426 A1 (January 12, 1995), P. Schneider and A.H. Pedersen (to Novo Nordisk A/S).

178. F. Archibald and B. Roy, *Appl. Environ. Microbiol.* **58**, 1496–1499 (1992).

179. R. Bourbonnais and M.G. Paice, *FEBS Lett.* **267**, 99–102 (1990).

180. G. Daniel, J. Volc, and E. Kubatova, *Appl. Environ. Microbiol.* **60**, 2524–2532 (1994).

181. L. Marzullo, R. Cannio, P. Giardina, M.T. Santini, and G. Sannia, *J. Biol. Chem.* **270**, 3823–3827 (1995).

182. P. Ander, C. Mishra, R.L. Farrel, and K.-E.L. Eriksson, *J. Biotechnol.* **13**, 189–198 (1990).

183. M.-L. Niku-Paavola, M. Ranua, A. Suurnakki, and A. Kantelinen, *Biosource Technol.* **50**, 73–77 (1994).

LACTONES, BIOCATALYTIC SYNTHESIS

TAKESHI SUGAI
HIROMICHI OHTA
Keio University
Yokohama, Japan

KEY WORDS

Asymmetric reduction
Baeyer–Villger reaction
Esterase
Hydrolysis
Kinetic resolution
Lipase
Optically active compounds
Transesterification

OUTLINE

INTRODUCTION

The most important feature of biocatalyst-mediated reaction is the chemo-, regio-, and enantioselectivity. The recent increasing demands for enantiomerically enriched compounds, and hopefully pure single enantiomers, in the field of pharmaceuticals and agrochemicals (1) promoted the development of excellent chemoenzymatic procedures (2). These achievements have been well reviewed in books and selected papers (3). In this article, we focus on the recent biocatalytic procedures (mostly from 1990 to 1997) for the synthesis of enantiomerically enriched forms of lactones.

Lactones play important roles as the starting materials and intermediates of pharmaceuticals, agrochemicals, flavors, and other fine chemicals, as well as macromolecules with industrial importance. They are derived from many kinds of precursors, essentially by means of C—O bond formation (esterification) between hydroxyl and carboxyl groups or insertion of an oxygen atom on the C—C bond adjacent to carbonyl groups (Scheme 1).

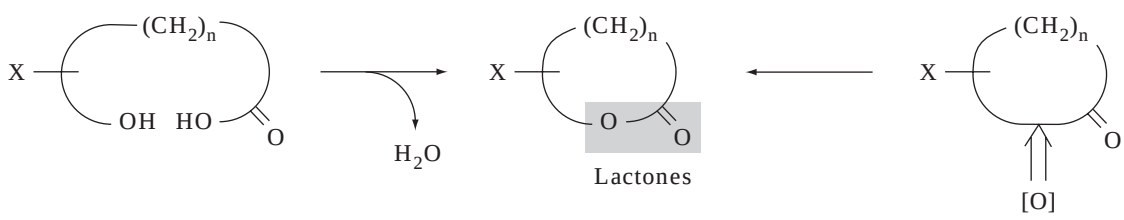
In the following sections, several approaches are described: *(1)* hydrolysis of ester functionality of corresponding racemic lactones; *(2)* lactonization of hydroxy acids and their derivatives; *(3)* reduction of carbonyl precursors; *(4)* oxidation of diol precursors; *(5)* Baeyer–Villiger type oxidation of ketones; and *(6)* separation of isomers by the enzymatic action on other functionalities (Y) that are appended to the molecules (Scheme 2). In addition, an approach based on the biodegradation of lipids and aromatics *(7)* is briefly discussed. Finally, preparation of some lactones of special industrial interests *(8)* is emphasized in detail.

HYDROLYSIS OF ESTER FUNCTIONALITY OF CORRESPONDING RACEMIC LACTONES

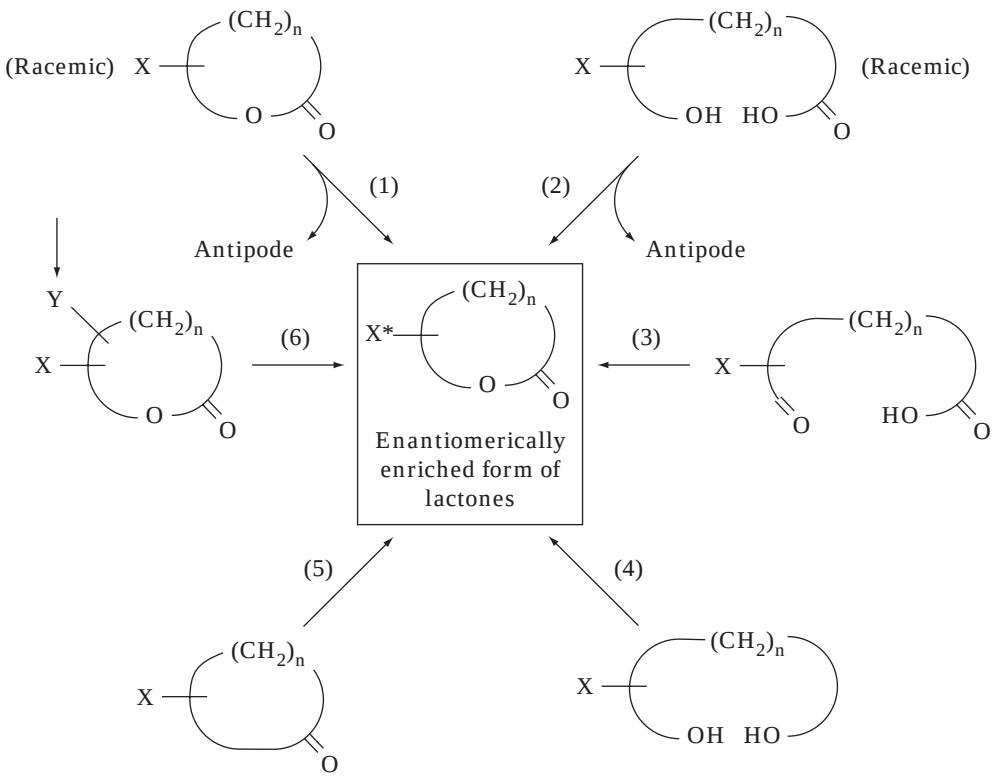
The simplest approach for the preparation of enantiomerically enriched forms of lactones is the optical resolution (separation of enantiomers) of the racemic mixture, which is readily synthesized in a chemical manner and often available from commercial sources. A representative example is illustrated in Scheme 3 (4). The rate of hydrolysis of (*S*)-enantiomer is higher than that of the (*R*)-form and, accordingly, a "kinetic" resolution occurs to give enantiomerically enriched forms of the products, (*R*)-lactone and (*S*)-hydroxy acid.

Pig liver esterase (PLE, as Sigma L 8285, for example), horse liver esterase (HLE, Sigma L 9627), and lipases of different origins involving pig pancreas lipase (PPL, Sigma, L 3126) are available for this purpose. A number of related five- and six-membered lactones have been resolved in a similar manner (4–7). Bicyclic compounds are well accepted as the substrates (8). Moreover, this method can be applied to substrates with rather small four-membered rings (9–11), and medium- and large-membered rings (12). Examples cited here are listed in Figure 1. It should be noted, however, that the absolute configuration as well as the enantiomeric excess (*e.e.*) depends upon both the structure of substrates and the kind of enzymes. It was also reported that the addition of calcium chloride affected the enantioselectivity of the reaction in Scheme (4).

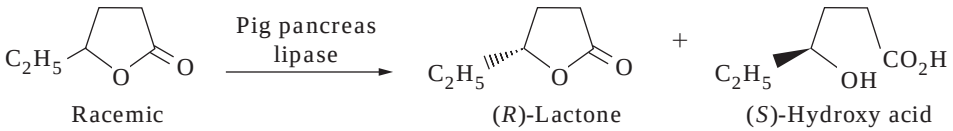
As for the estimation of the enantiomeric excess, there have been developed gas chromatographic analyses with chiral stationary phase such as cyclodextrin derivatives (Lipodex E from Marcherey Nagel, CP-Chirasil-DEX CB from GL Science, Chiraldex BP from Tokyo Kasei, and others) (6,13,14). The chiral stationary phase of liquid chromatography (e.g., ChiraSpher, Merck) can be used for the same purpose (15). In the case that no good separation of enantiomers on the chromatogram is observed, other methods can be utilized, such as the conversion of lactones to the corresponding diols by the action of methyllithium and subsequent nuclear magnetic resonance (NMR) measurement in the presence of chiral NMR shift reagent (16). Another alternative is the analysis of diastereomeric derivatives on GLC or HPLC. For this purpose, phenylethyl urethanes (15), α-acetyllactates (15), and α-methoxy-α-trifluoromethylphenyl acetate (MTPA) esters (17) derived from the alcohol(s) are available. Another factor to note is the separation of the products and recovered starting materials. Although the products are essentially acidic while the starting materials are neutral, sometimes it is seriously troublesome, as in the case when these two materials show similar properties with regard to volatility, water solubility, and chromatographic behavior. The hydrolysis often spontaneously and reversibly proceeds in nonbiocatalytic conditions. The relactonization of the resulting



Scheme 1. Lactones and their precursors.



Scheme 2. Biocatalytic approaches to lactones. For the explanation of each approach, see text.



Scheme 3. Lipase-catalyzed hydrolysis of a racemic lactone.

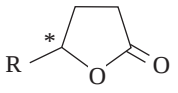
hydroxy acids to the lactones would bring about the lowering of the enantiomeric excess of the "kinetically resolved" lactones.

LACTONIZATION OF HYDROXY ACIDS AND THEIR DERIVATIVES

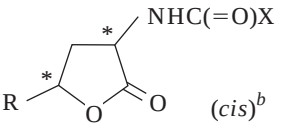
Hydrolytic enzymes, especially lipases, also catalyze the esterification of hydroxy acids with a concomitant dehydration as well as transesterification of hydroxy esters. Lipase-catalyzed formation of five-membered lactone (γ-lactone) from γ-hydroxy carboxylates is a highly favored process. Kinetic resolution of the racemic hydroxy ester was achieved using PPL-catalyzed lactonization in an organic solvent, such as diethyl ether (Scheme 5, Fig. 2) (7,15,18–23). As mentioned, spontaneous lactonization is the competing reaction that lowers the "apparent" enantioselectivity of the reaction. For example, the rate of lactonization of (*S*)-isomer of methyl 3-hydroxypentanoate is 26 times faster than that of the (*R*)-isomer at the start of the reaction (18). After a prolonged incubation [more than 70% conversion for the (*S*)-isomer and 25% for the (*R*-isomer)], however, both of the rates gradually decrease and are almost the same. Only five-membered ring formation occurred from γ,δ-dihydroxy carboxylates (23). Another interesting difunctional substrate is prochiral γ-hydroxypimelate (5-hydroxynonanedioate). Lipase cleanly discriminates between the two enantiotopic groups to give a single enantiomer (>95% *ee*) of lactone (19).

The situation is rather complex in the case of the formation of six-membered (and larger) ring lactones. The competitive reaction, oligomerization of the starting material with an intermolecular reaction, occurs. The preference of intramolecular versus intermolecular reaction depends on the structure of substrates. The reactivity of the esterification/transesterification is greatly affected by the stereochemical environment of the hydroxyl group that concerns the reaction. In this sense, enantiomers of the substrates (right and wrong enantiomer in Scheme 6) are kinetically resolved when they react as acyl acceptors toward the preformed acyl donor (acyl–enzyme complex), resulting in the enantiomerically enriched form of the lactones. At the initial stage, the acyl–enzyme complex and small oligomeric intermediates form from both right and wrong isomers; however, intermediates, including the wrong isomer, lower their own reactivities when they want to work as acyl acceptors in the further oligomer chain elongation, as well as the lactonization. As all the reactions

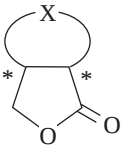
Five-membered ring lactones



R^a = C_2H_5, C_3H_7, C_5H_{11}, C_7H_{15}

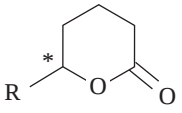


R = H, CH_3, C_6H_5, –CH=CH_2
X = C_6H_5, OCH_3 (*cis*)[b]

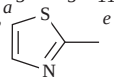


X^c = CH_2, $(CH_2)_2$, $(CH_2)_3$, $(CH_2)_4$, –CH_2CH=CHCH_2–

Six-membered ring lactones

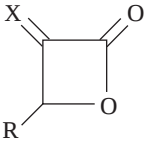


R = CH_3,[d] C_5H_{11},[a] C_7H_{15},[a]



[e]

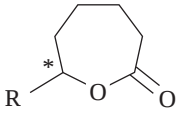
Four-membered ring lactones



X = H_2, R = CH_3,[f] C_3H_7,[f] $(CH_3)_2CH$,[f] CF_2CI[g]

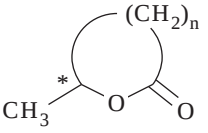
X = CH_2, R = CH_3,[h] $(CH_3)_2CH$[h]

Seven-membered ring lactones



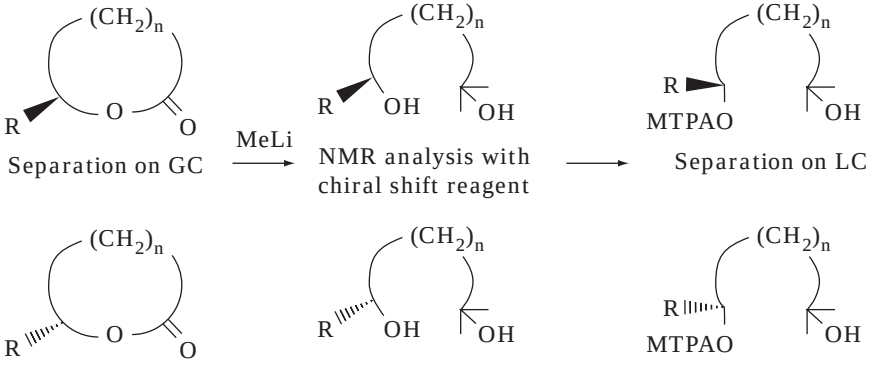
R = CH_3,[d,i] C_2H_5,[i] C_3H_7,[a,i] C_4H_9,[i] C_5H_{11},[a,i] C_6H_{13},[a] C_7H_{15},[i] C_8H_{17},[i]

Eight-membered ring lactones and larger

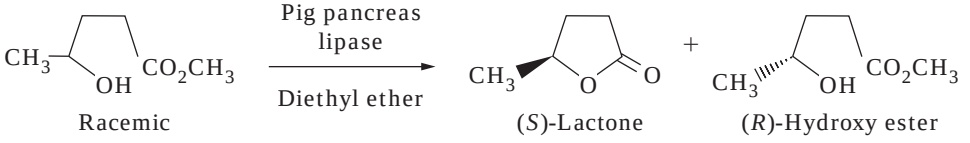


n = 5 – 7[d]

Figure 1. Enantiomerically enriched lactones by enzyme-catalyzed hydrolysis. a, Ref. 4; b, Ref. 5; c, Ref. 8; d, Ref. 12; e, Ref. 7; f, Ref. 9; g, Ref. 10; h, Ref. 11; i, Ref. 6.

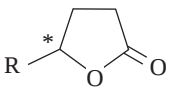


Scheme 4. Determination of enantiomeric excesses of lactones.



Scheme 5. Lipase-catalyzed lactonization of a racemic γ-hydroxy ester.

Five-membered ring lactonization



Aliphatic

R = CH_3,[a,b,c] C_2H_5,[b,c] C_3H_7,[c] C_4H_9,[c] C_5H_{11},[c] C_6H_{13},[b,c] C_7H_{15},[c] C_8H_{17},[b,c] C_8H_{17}—C≡C–[d]

Aromatic

R = C_6H_5,[a,b] *p*-$CH_3C_6H_4$,[b] *p*-$CH_3OC_6H_4$,[b] *p*-BrC_6H_4,[b] C_6H_5–C≡C–[e]

Functionalized

R = $CH_3O_2C(CH_2)_2$–,[b] $C_2H_5O_2C(CH_2)_2$–,[b] C_6H_5,$CH_2O_2C(CH_2)_2$–,[b] $(CH_3)_3CO_2C(CH_2)$–,[f] $HOCH_2$–,[g]

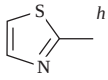
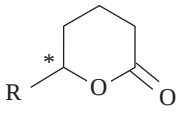


Figure 2. Enantiomerically enriched five-membered lactones by lipase-catalyzed transesterification. a, Ref. 18; b, Ref. 19; c, Ref. 15; d, Ref. 20; e, Ref. 21; f, Ref. 22; g, Ref. 23; h, Ref. 7.

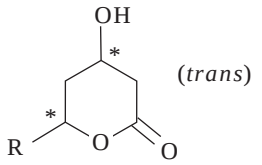
work in a reversible manner, the right enantiomer contained in right–wrong mixed fragments goes back to the starting material, and the enriched right enantiomer of lactone thus becomes further concentrated. That lactones themselves serve as the biocatalytic precursors of polyesters (19) also supports this reversible mechanism.

Unsubstituted methyl 5-hydroxypentanoate itself undergoes an exclusive oligomerization by the lipase-catalyzed reaction. In contrast, when a substituent was introduced to the δ- (5-) position on methyl 5-hydroxypentanoate, the desired lactonization proceeded (25), and this is well explained by Scheme 6. A number of δ-hydroxyalkanoates were applied as the substrates for lipases involving that from pig pancreas (15,25–27), *Candida antarctica* (Novo SP328) (26), and *Pseudomonas* (Amano P) (28), as illustrated in Figure 3. To improve the enantioselectivity of the reaction, an elaborated condition was required by means of the immobilization of enzyme (28).

Six-membered ring lactonization



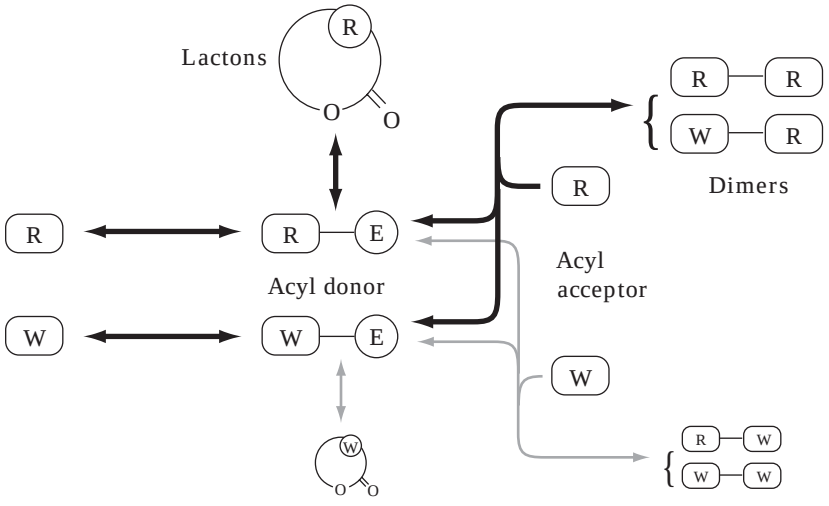
R = CH_3,[a,b] C_2H_5,[b] C_3H_7,[b] C_4H_9,[b] C_5H_{11},[b] C_6H_{13},[b] C_7H_{15},[b] C_8H_{17},[b] $C_{11}H_{23}$,[c]



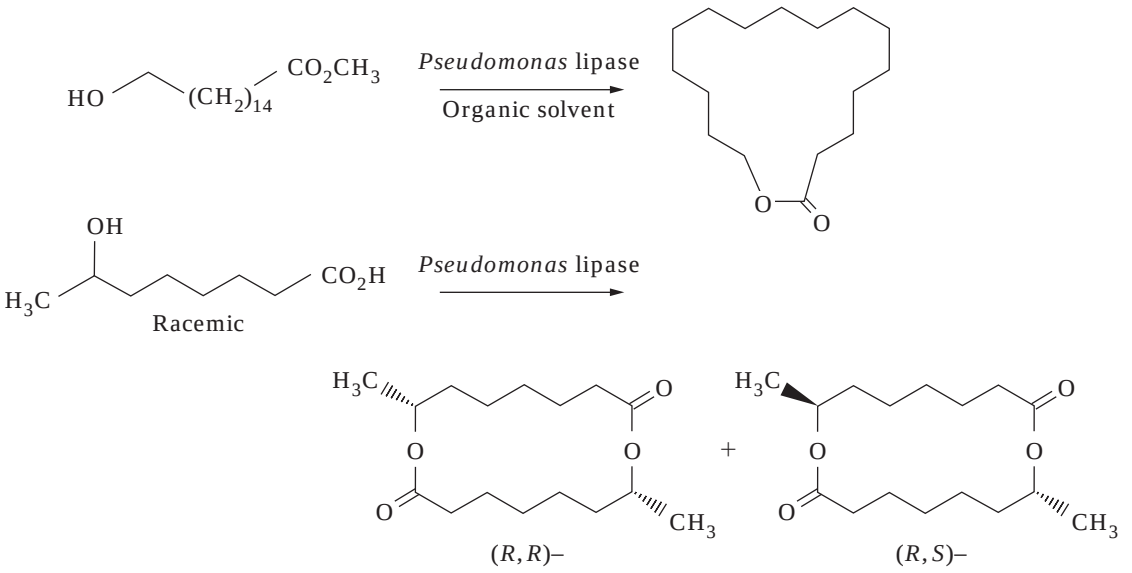
R = C_5H_{11},[d] $C_6H_5(CH_2)_2$–,[d] C_6H_5CH=CH– (*trans*)[e]

Figure 3. Enantiomerically enriched lactones by enzyme-catalyzed transesterification. a, Ref. 25; b, Ref. 15; c, Ref. 28; d, Ref. 27; e, Ref. 26.

Another characteristic feature of lipase-catalyzed transesterification is the ability of the "macrolactonization" in an organic solvent. The first reported example is illustrated in Scheme 7, the formation of 17-membered lactone from methyl 16-hydroxyhexadecanoate catalyzed by lipases from *Pseudomonas* sp. (P from Nagase), *Pseudomonas fluorescens* (P from Amano), and porcine pancreas (Sigma) (29). The yield was increased to 80% using a nonpolar solvent, such as cyclohexane and *n*-heptane. Further attempts to improve the rate of lactonization and the yield of lactone were made successful by adding an irreversible process to the reaction, such as adsorption of methanol, which comes from transesterification on molecular sieves (30,31) or the use of vinyl ester as the substrate (32). The competitive reaction is the formation of dimeric, trimeric, and higher oligomeric lactones, as well as linear oligomers. The distribution of the products (monomeric, dimeric, and so on) depends on the structure (30,32–33), as well as the concentration of substrates (29). Increasing availability of lipases, such as those from *Pseudomonas* sp. (AK from Amano), *Candida cylindracea* (Meito OF-360), *C. antarctica* (Novo SP525), and those already cited made this protocol attractive, and indeed, some of them worked well for the syntheses of naturally occurring macrolactones and diolides (31,34–36). Although the lactonization essentially



Scheme 6. Lipase-catalyzed lactonization and oligomerization; R, right enantiomer; W, wrong enantiomer.



Scheme 7. Lipase-catalyzed macrolactonization.

proceeds in an enantioselective manner (29–30,34–37), both of two diastereomeric products (*R,R*)-isomer and [*R,S* (*meso*)]-isomer formed in the course of dimeric lactone synthesis from racemic substrate (Scheme 7) (32,37). This result is easily deduced from the discussion of the mechanism of lipase-catalyzed lactonization as in the previous section (cf. 32). An approach used to obtain an enantiomerically enriched form of substrate in advance was effective to avoid this situation (31), examples cited in this section are listed in Figure 4.

REDUCTION OF CARBONYL PRECURSORS

The reduction of γ- (4-) and δ- (5-)-oxoacids provides the corresponding hydroxy acids (7,13,14,23,38–44); however, in most cases, the products are isolated as lactones with an intramolecular dehydration during the workup procedure. Many kinds of microorganisms that reduce carbonyl groups are available, involving commercial baker's yeast. The change of microorganism gives a striking contrast to the stereochemical outcome, as illustrated in Scheme 8. (*R*)-γ-Valerolactone was obtained by the reduction of levulinic acid with a yeast, *Yamadazyma farinosa* IFO 10896 (*Pichia farinosa* IAM 4682) (13), while baker's yeast–mediated reduction of ethyl levulinate (4-oxopentanoate) provides ethyl (*S*)-4-hydroxypentanoate, a precursor of (*S*)-enantiomer of lactone (23,37). In the latter case, the hydroxy ester was lactonized by the action of PPL (discussed earlier) with a further enhancement of enantiomeric excess (23). A similar approach could be applied for the synthesis of (*R*)-γ-caprolactone (14). Another interesting combination of hydrolytic enzymes and the yeast-mediated reduction was reported in the synthesis of a functionalized lactone, which is elaborated as a versatile starting material for natural product synthesis. Selective hydrolysis of one of the two alkoxycarbonyl groups by means of *Pseudomonas diminuta* IFO 13181 provided the substrate for the subsequent reduction (39).

Yeast-mediated reduction is not limited to a carbonyl group; the C═C double bonds, of α,β-unsaturated lactones were reduced to the enantiomerically enriched forms of saturated products (Scheme 9) (45). Examples cited in this section are listed in Figure 5.

OXIDATION OF DIOL PRECURSORS

Enantiomerically enriched forms of lactones are available based on the selective oxidation of two enantiotopically different primary alcohols. The formation of lactone proceeds in two steps: (*1*) the selective oxidation of the primary hydroxyl group and the hemiacetal (lactol) formation of the resulting hydroxy aldehyde, and (*2*) the subsequent oxidation of secondary hydroxyl group of lactol, as illustrated in Scheme 10. The most studied enzyme so far is horse liver alcohol dehydrogenase (HLADH, for example, Sigma A 6128 and Boehringer No. 102733) (46). This enzyme usually catalyzes both oxidation steps; however, when the resulting lactol is a poor substrate for this enzyme, another chemical step for oxidation to the desired lactone under mild conditions is required separately, such as the use of silver carbonate. HLADH works with the aid of oxidized form of the cofactor, NAD, which is recycled with another oxidant, flavin mononucleotide sodium salt (FMN), due to the expense. A number of lactones in enantiomerically enriched form have been reported (46–51), as illustrated in Figure 6, and some of the products were developed as the starting materials for natural product syntheses (49,51).

Based on the extensive studies on the substrate specificity and the stereochemistry of the products, as well as the information of the structure of the enzyme itself, a cubic-space section model has been established by Jones (52,53). This is an empirical rule for stereochemical preference in regard to the enantiotopic two hydroxyl groups

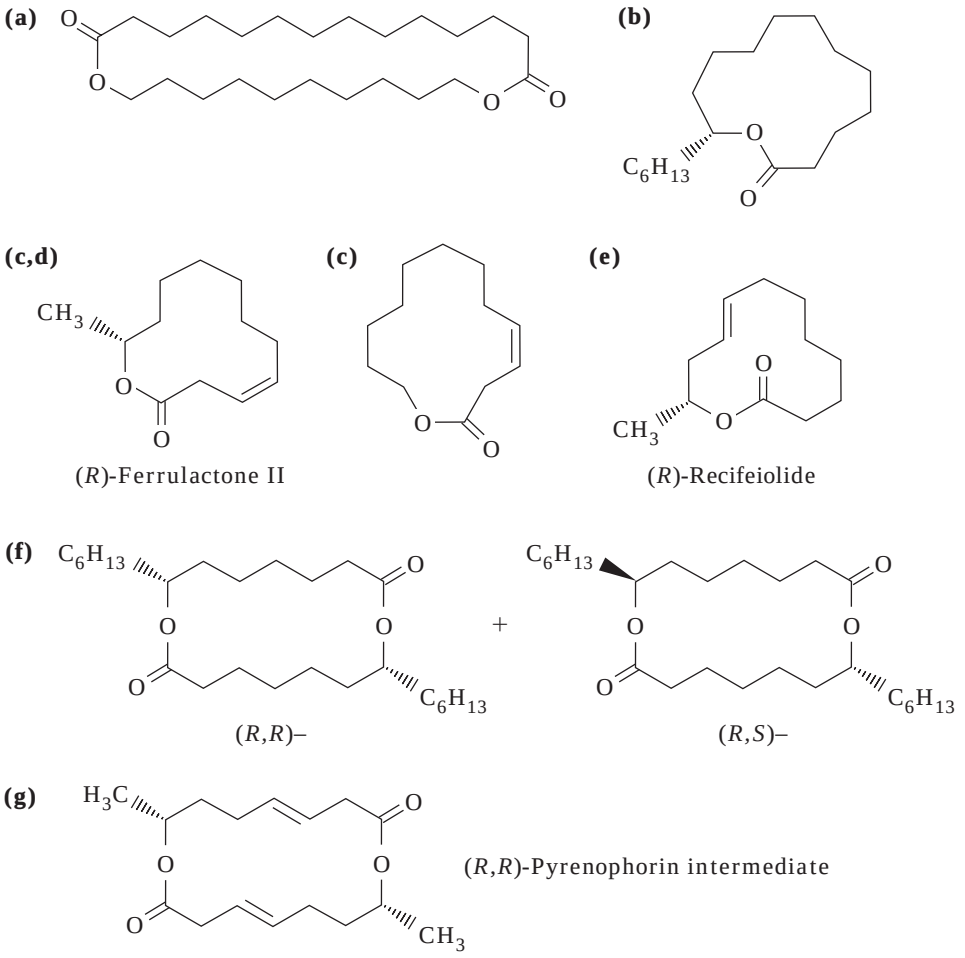
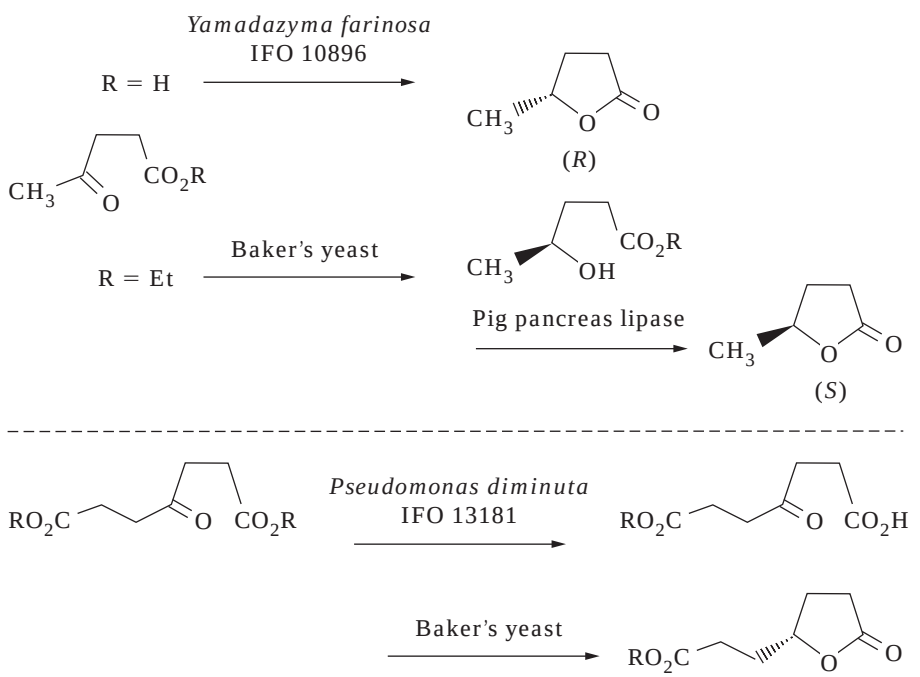


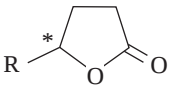
Figure 4. Macrocyclic lactones and diolides by lipase-catalyzed cyclization. a, Ref. 33; b, Ref. 30; c, Ref. 34; d, Ref. 36; e, Ref. 35; f, Ref. 37; g, Ref. 31.



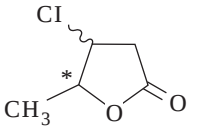
Scheme 8. Yeast-mediated reduction of γ-oxoacids and esters.

Scheme 9. Yeast-mediated reduction of C═C double bond.

Five-membered ring lactones

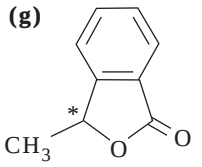


R = CH_3,[a,b,c] C_2H_5,[d] C_3H_7,[d] C_4H_9,[d] C_5H_{11},[d] C_8H_{17},[d,e] $C_{11}H_{23}$,[d] $C_{13}H_{27}$,[d]

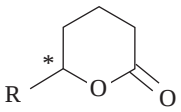


(Diastereomeric mixture)[f]

(g)



Six-membered ring lactones



R = CH_3,[h] C_2H_5,[d] C_3H_7,[d] C_4H_9,[d] C_5H_{11},[d] C_8H_{17},[d] $C_{11}H_{23}$,[d,e] $C_{13}H_{27}$,[d] $CH_2{=}CH(CH_2)_2{-}$[i]

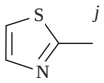
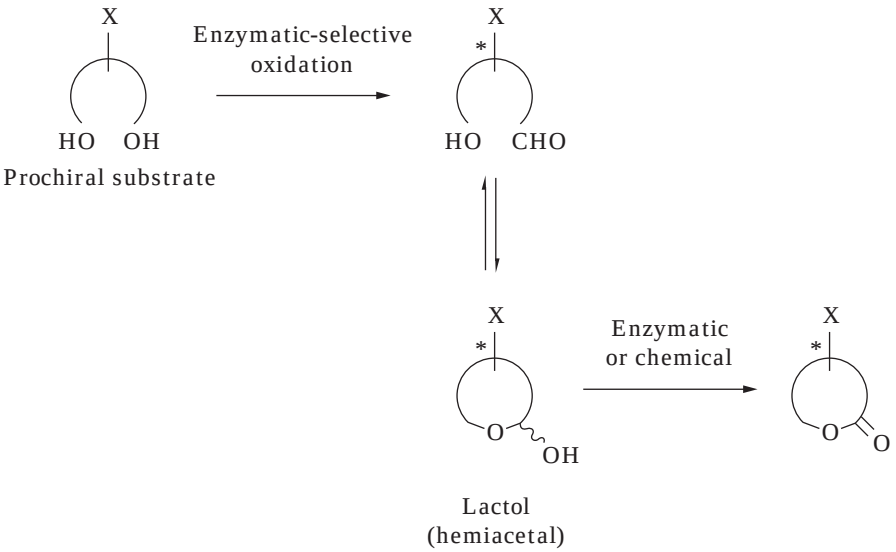


Figure 5. Enantiomerically enriched lactones by yeast-mediated reduction of oxoacids or esters. a, Ref. 13; b, Ref. 38; c, Ref. 23; d, Ref. 40; e, Ref. 41; f, Ref. 42; g, Ref. 44; h, Ref. 14; i, Ref. 43; j, Ref. 7.



Scheme 10. HLADH catalyzed enantiotopic group-selective oxidation of diols.

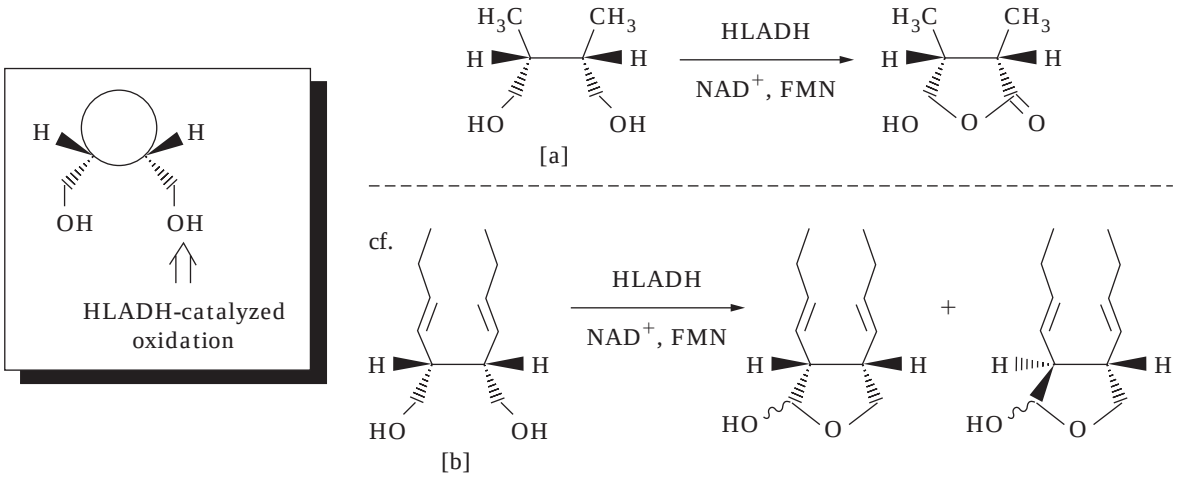
and, accordingly, a way of predicting the absolute stereochemistry of the products. Scheme 11 summarizes the conclusion of stereochemical preference in the oxidation. Indeed, the formation of (3*S*,4*R*)-2,3-dimethyl-γ-butyrolactone (4-butenolide) from *meso*-2,3-dimethyl-1,4-butanediol [a] is well explained by the model (50). An open-chain substrate with considerable conformational flexibility suitably fits to the space of the enzyme. However, there are always some exceptions. Even in the case of very closely related substrate [b] (Scheme 11), the product was a complex mixture of isomeric lactols caused by a nonenzymatic epimerization at the α-position of aldehyde intermediate (51). The gross result seems to be that a primary alcohol located in the opposite site was preferentially oxidized. Similar reversed selectivity was observed in another bicyclic substrate (54). It might be possible that the epimerization of intermediates and the equilibrated oxidoreductive condition caused the integration of the secondary products in

R^a = CH_3, C_2H_5

X = CH_2,[b] $C(CH_3)_2$,[b] $(CH_2)_2$,[b] $(CH_2)_3$,[b] $(CH_2)_4$,[b] $-CH_2SCH_2-$,[a] $-CH_2CH=CHCH_2-$[b]

R^c = CH_3, CH_2CH_3, $CH_2CH_2CH_3$, $CH(CH_3)_2$, cyclo-C_6H_{11}, C_6H_5, $CH_2C_6H_5$, $(CH_2)_2C_6H_5$

Figure 6. Enantiomerically enriched lactones by HLADH-mediated oxidation. a, Ref. 50; b, Ref. 48; c, Ref. 47.

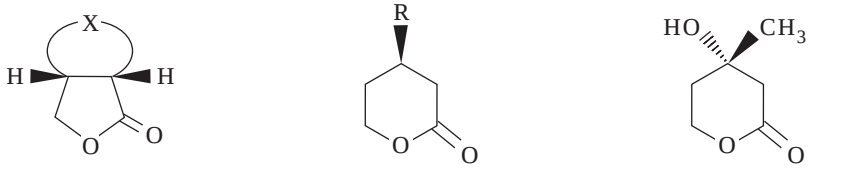


Scheme 11. An empirical rule for the specificity of HLADH catalyzed oxidation. For [a] and [b] see text.

these two examples. Regardless, we should pay considerable attention to the application of this rule to determine the absolute configuration of the molecule with an unknown stereochemistry.

The selective oxidation of diols has also been reported by means of whole cells of microorganisms involving *Gluconobacter roseus* (55) and *Nocardia corallina* (56) (Fig. 7). It should be noted that when whole cells of a microorganism are used, it is difficult to simply ascribe the origin of the enantiomerically enriched form of lactones to the enantioselectivity of the alcohol dehydrogenases. Sih noted the possibility of the assimilation of natural enantiomer in a study where *Flavobacterium oxydans* oxidized 3-methylpentane-1,3,5-triol to give an unnatural enantiomer of mevalonolactone (57). A related example is that *Rhodococcus rhodochrous* IFO 15564–mediated "enantioselective" hydrolysis with discrimination of the remote asymmetric center of hydroxy nitriles, afforded an enantiomerically enriched form of lactones (58). However, another report pointed out the chance of enantioselective oxidation, as well as the further catabolic degradation of the hydroxy acid, which is a possible intermediate of the metabolic pathway by the same microorganism (59).

Enzyme-mediated asymmetrization of prochiral and *meso* compounds, coupled with the subsequent chemical reactions, widely became accepted in synthetic organic



X^a = $(CH_2)_4$, $-CH_2CH=CHCH_2-$,

R^b = CH_3, $CH(CH_3)_2$, C_6H_5

(*S*)-mevalonolactone[b,c]

Figure 7. Enantiomerically enriched lactones by whole cell–mediated enantioselective oxidation. a, Ref. 56; b, Ref. 55; c, Ref. 57.

chemistry. Those examples of applications related to synthesizing enantiomerically enriched forms of lactones by use of hydrolytic or oxidoreductive enzymes are cited in selected articles and reviews (60).

BAEYER–VILLIGER TYPE OXIDATION OF KETONES

Enzyme-catalyzed Baeyer–Villiger type oxidation by means of cycloalkanone monooxygenases (61) have come to the attention of organic chemists. Purified enzymes (cyclohexanone and cyclopentanone monooxygenases and related enzymes) from microorganisms have been utilized for this purpose by combining the recycling system of cofactors (for example, FAD-NADP) with glucose-6-phosphate dehydrogenase and glucose-6-phosphate (Scheme 12). The use of the whole cell system was also effective to avoid the set up of the above recycling systems. Recent studies on the structural elucidation of enzymes (62), as well as the extensive investigation on the structure of substrates and products, made it possible to clarify the origin of stereoselectivity of the enzyme-catalyzed oxidation, and a model for predicting the active site has also been proposed (63).

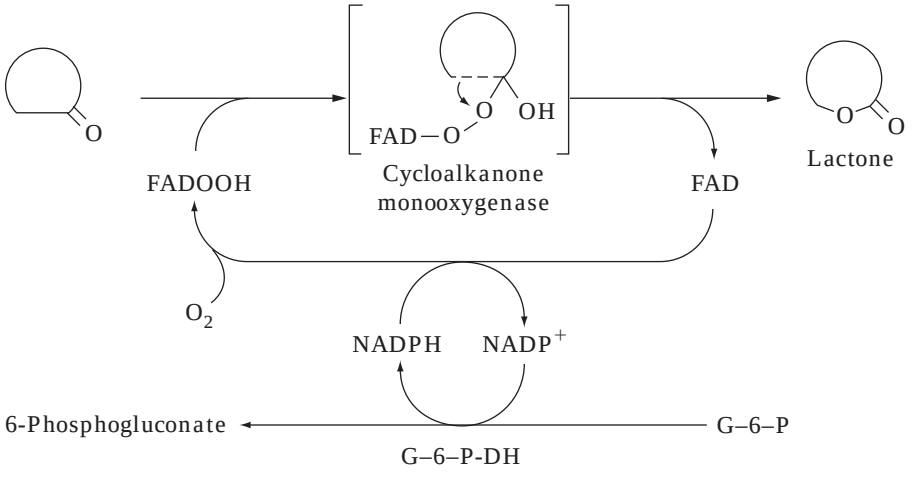
Substituted cyclohexanones are good substrates for cyclohexanone monooxygenase from *Acinetobacter* NCIB 9871. It was very interesting that although there are equal chances for the migration for two C—C bonds in the relationship of enantiotopic groups, the enzyme-catalyzed reaction proceeds in a quite stereoselective manner, as illustrated in Scheme 13 (64–66). The examples of kinetic resolution of racemic ketones via enantioselective Baeyer–Villiger reaction were also reported. Enzymes from *Pseudomonas putida* NCIMB 10007 (67), *Acinetobacter* NCIB 9871 (68), and *Pseudomonas* sp. NCIMB 9872 (69) were effective in preferentially oxidizing one enantiomer of substrate to give lactones, while leaving the other enantiomer intact as the ketone form (Scheme 13). The unreacted enantiomer is readily converted to the corresponding lactone by chemical Baeyer–Villiger oxidation, such as a treatment with peracid with complete retention of the absolute configuration. Because of the high selectivity and wide substrate specificity of *Acinetobacter* enzyme, an approach toward a "designer yeast" has been reported (70). When a strain of baker's yeast was engineered to express cyclohexanone monooxygenase, the reaction became unbelievably easy. Only the incubation of substrate with a stored yeast cells under frozen condition in an appropriate nutrient broth, for example, YEP-galactose, worked well to carry out the Baeyer–Villiger oxidation. Problems that had been caused from the insolubility of the substrates were cleanly solved by adding β-cyclodextrin. The major side reaction is the reduction of substrate to the corresponding secondary alcohol. Lactones available from enantioselective Baeyer–Villiger oxidation are listed in Figure 8.

In the case of some bicyclic substrates, the kinetic resolution brought about a not simple outcome. The more substituted C—C bond of ketone had migrated on one enantiomer to give a "normal" lactone, while an "abnormal" lactone came from the other enantiomer by migration of less substituted C—C bond of ketone, as described in Scheme 14 (71,72). The resulted lactones involving carbocyclic and heterocyclic rings were shown in Figure 9.

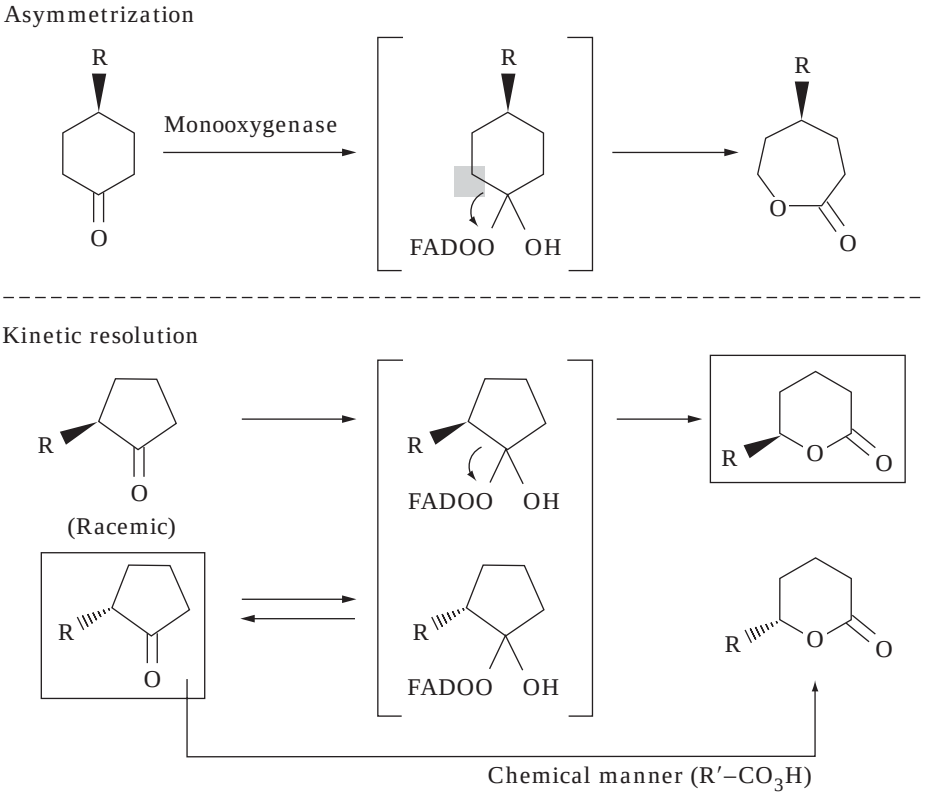
The other application of this monooxynegase is the addition or insertion of an "oxygen" atom on a sulfur atom (from sulfides to sulfoxides), and on benzylic methylene groups (74). As the latter example, a *P. putida*–catalyzed formation of phthalide from alkyl substituted benzoic acid on the *ortho* position has been reported (44).

SEPARATION OF ISOMERS BY THE ENZYMATIC ACTION ON OTHER FUNCTIONALITIES THAT ARE APPENDED TO THE MOLECULES

So far, the biochemical reactions directly concerned with the lactone carbonyl moiety have been reviewed. However, there are many natural and synthetic lactones with multiple functional groups. In this section, miscellaneous approaches for preparing enantiomerically enriched forms by utilizing the reaction on other suffixed functional groups are reviewed. The first approach is the kinetic resolution by use of hydrolytic enzymes. Racemic mixtures of the lactones with primary alcohols were effectively resolved by way of a lipase-mediated enantioselective acylation with acetic anhydride. Unreactive enantiomers could readily be separated by a chromatographic purification, while the



Scheme 12. Monooxygenase-mediated Baeyer–Villiger oxidation.



Scheme 13. Stereocourse of enzyme-catalyzed Baeyer–Villiger oxidation.

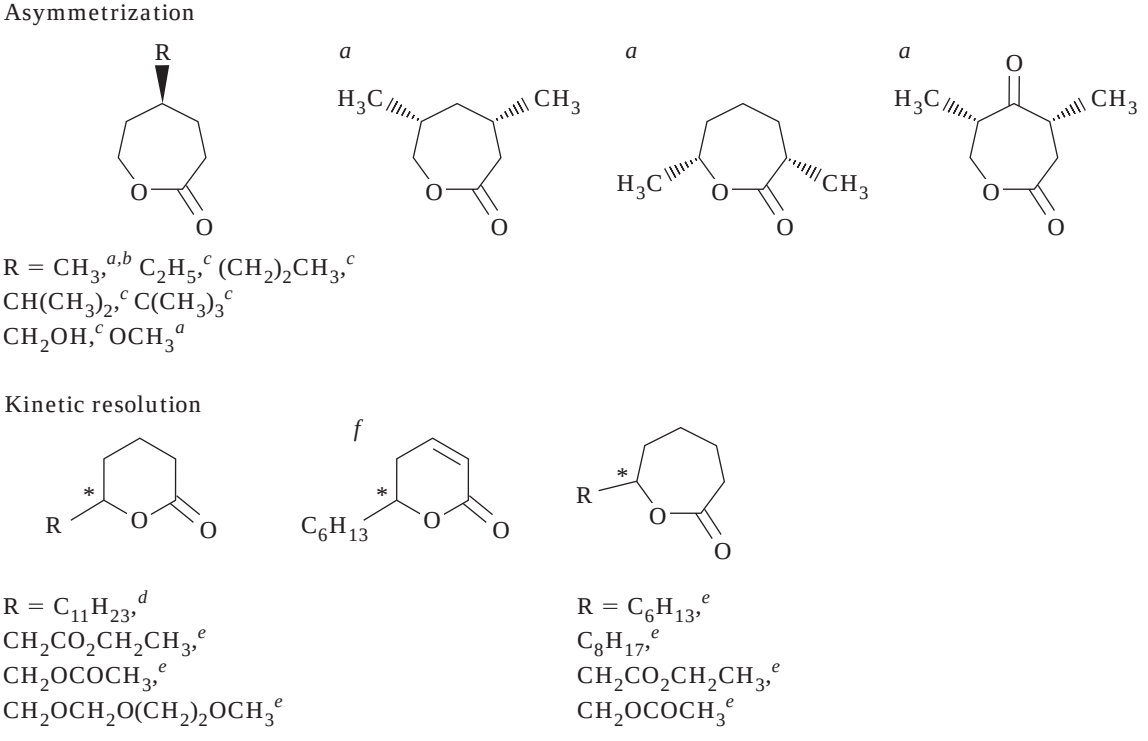
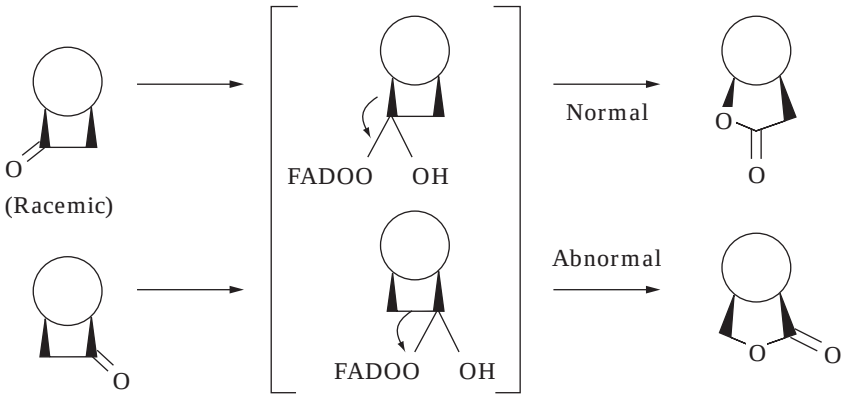
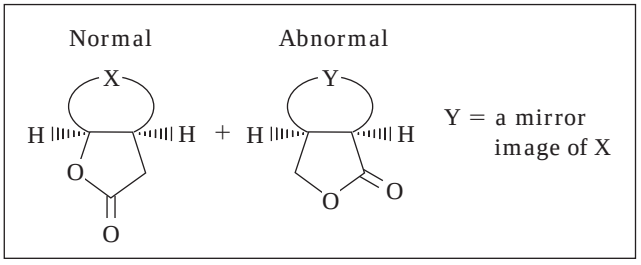


Figure 8. Enantiomerically enriched lactones by Baeyer–Villiger oxidation. a, Ref. 64; b, Ref. 66; c, Ref. 65; d, Ref. 68; e, Ref. 67; f, Ref. 69.



Scheme 14. Normal and abnormal products by enzyme-mediated Baeyer–Villiger oxidation.



Carbocyclic
X = $-(CH_2)_3-$,[a]
$-CH_2CH=CH-$,[a]
$-(CH_2)_4-$,[a]
$-(CH_2)_2CH=CH-$,[a]
$-CH_2CH=CHCH_2-$,[a]
$-(CH_2)_2CR^1R^2-$[b]
(R^1, R^2 = H, Br, OH, $OCH_2C_6H_5$)

Heterocyclic[c]
X = $-O(CH_2)_2-$,
$-CH_2OCH_2-$,
$-(CH_2)_2O-$,
$-O(CH_2)_3-$,
$-(CH_2)_3O-$

Figure 9. Enantiomerically enriched lactones by Baeyer–Villiger oxidation as normal and abnormal products. a, Ref. 71; b, Ref. 73; c, Ref. 72.

opposite enantiomer was converted to the corresponding acetate. Hydrolysis of the corresponding acetate was elaborated by the use of lipase-catalyzed neutral condition to avoid the racemization of the product under either acidic or basic conditions, which is caused by an intramolecular transesterification (Scheme 15) (75). A similar approach on the asymmetrization of a lactone diol with a *meso* structure has been reported (76).

In contrast, there has been reported another interesting result utilizing an inherent nature of the substrate, that in which the stereochemistry of hemiacetal position is prone to racemize, even under very mild conditions. An enantioselective acylation of one enantiomer with a concomitant spontaneous racemization enabled the dynamic kinetic resolution of a lactone (Scheme 16) (77). Ester and amide functionalities are also available as the clues for the lipase-catalyzed kinetic resolution (78,79).

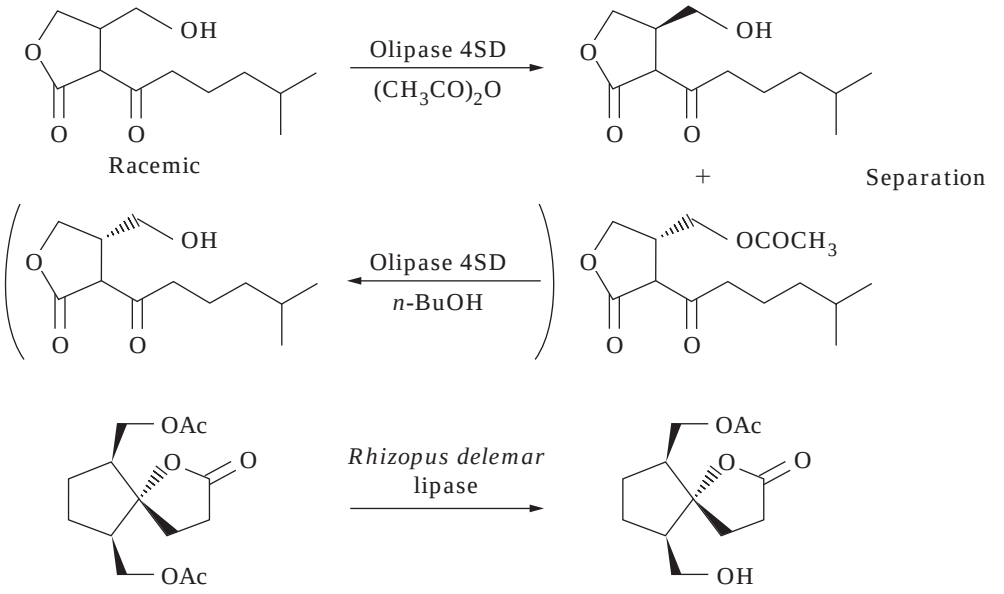
The second approach is the separation of diastereomeric products. For example, the chromatographic separation of diastereomeric lactones produced by a baker's yeast mediated reduction of the racemic substrate, as illustrated in Scheme 17, made the separation of enantiomers possible, in regard to the chiral center already present in the substrate (80). Needless to say, it is desirable that the enantioface selectivity in the reduction of carbonyl group is as high as possible. Indeed, the whole cell–mediated reduction of the racemic substrates in most cases accompanies the kinetic resolution to give rather complex results (81) and, accordingly, efforts for the screening of suitable strain of microorganism have been reported (82).

APPROACHES BASED ON THE BIODEGRADATION OF LIPIDS AND AROMATICS

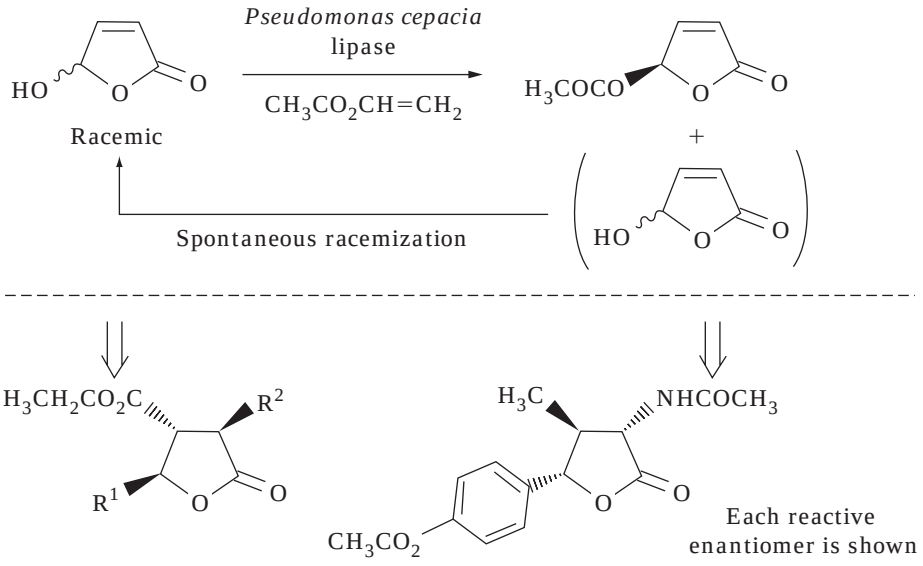
Naturally occurring flavoring lactones themselves are in most cases the oxidative degradation products of unsaturated fatty acids. Lactone metabolites, as illustrated in Figure 10, were reported in the microbial incubation of natural products containing fatty acids, such as hydroxy unsaturated fatty acids (83,84), vernoic acid (85), and castor oil (86). An aromatic substrate was also converted to the partial degradation product. (+)-Muconolactone was prepared from a racemic mixture of mandelic acid by the action of *P. putida* mutant PRS 2912. Since this microorganism possesses a mandelic acid racemase, both enantiomers of the starting material are used as the substrate. The origin of chirality of the muconolactone is the step utilizing a muconate cycloisomerase. As this mutant strain lacks a muconate isomerase, which works as the key step in the further metabolic pathway, muconolactone was obtained in the incubation broth in as high as 97% yield (Scheme 18) (87).

PREPARATION OF SOME LACTONES OF SPECIAL INTEREST

(*R*,*Z*)-(−)-tetradec-5-en-4-olide is the sex pheromone of Japanese beetle and is of great industrial value as the new substituent of pesticides. Because any trace of contaminating (*S*)-enantiomer strongly inhibits the pheromone activity, the synthesis of highly enantiomerically enriched form had been secured. Although approaches based on the lipase-catalyzed lactonization toward (*R*)-tetradec-5-yn-4-



Scheme 15. Lipase-catalyzed kinetic resolution by the action on primary alcohol moiety.

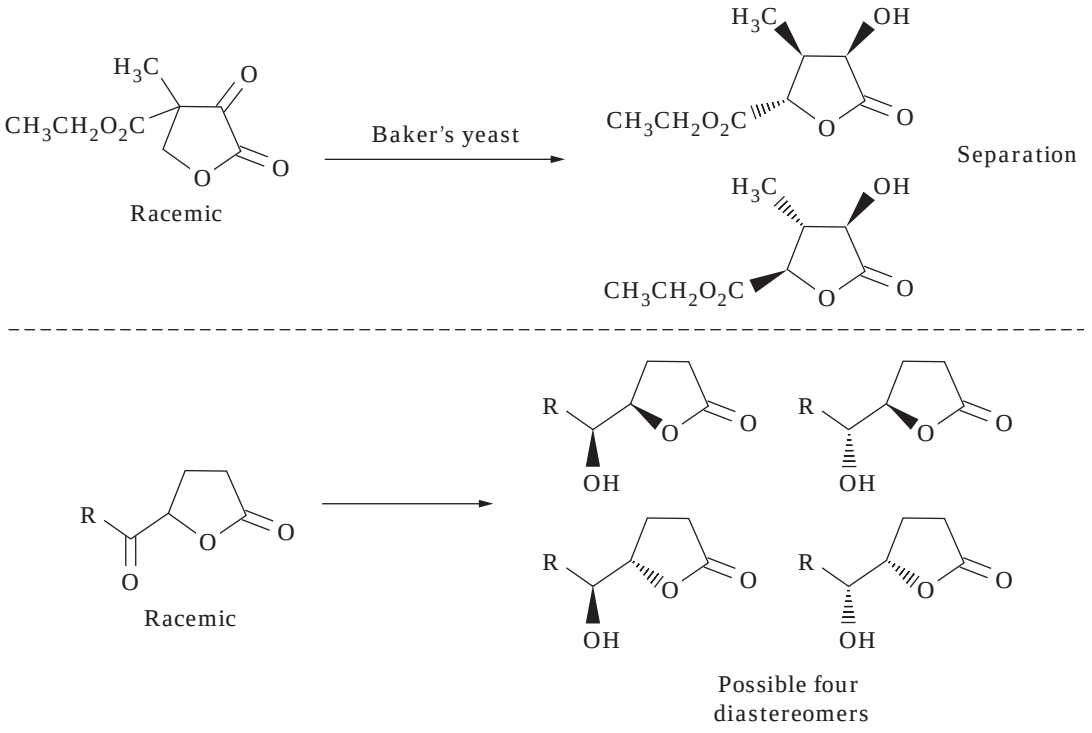


Scheme 16. Kinetic resolution of lactones by the action of enzymes on other functionalities.

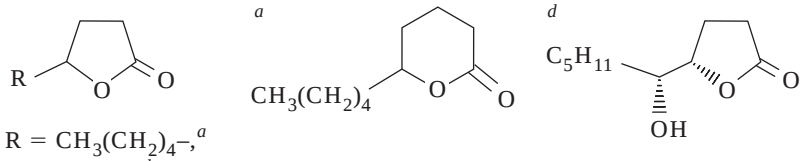
$R^1 = CH_3, C_5H_{11}$
$R^2 = H, CH_3$

olide, a direct precursor for the pheromone, have been reported, the enantioselectivity of the enzymatic step was not satisfactory (20,23). The problem was solved by successive two-lipase-catalyzed resolution, as illustrated in Scheme 19 (88). Treatment of racemic alkynyl alcohol with *Pseudomonas cepacia* lipase (Amano PS) and succinic anhydride afforded the desired (*R*)-isomer of 94% ee as a hemisuccinate. The further enhancement of the enantiomeric excess was achieved by the hydrolysis of isolated product with *Candida cylindracea* lipase (Meito OF). The concomitant lactonization proceeded, thus the (*R*)-lactone was obtained in the state of more than 99% ee. Semihydrogenation afforded the pheromone of nearly pure enantiomer.

D-Pantolactone has a great importance as the starting material in medicinal and nutritional industries. So far, various approaches were devoted to develop efficient enzyme-mediated preparations. Enzymes that can hydrolyze its racemate in an enantioselective fashion were screened, and a lactonohydrolase from *Fusarium oxysporum* was discovered (89). Immobilized cells of this microorganism worked well, even at a surprisingly high concentration of substrate (350 g/L) and D-pantolactone was obtained in 90–97% ee (Scheme 20). The immobilized mi-

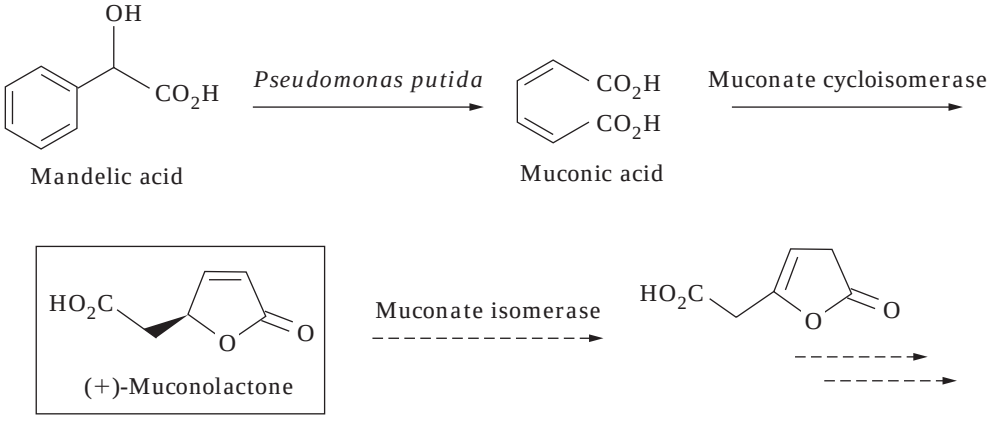


Scheme 17. Enantioface-selective reduction and the separation of resulting diastereomers.

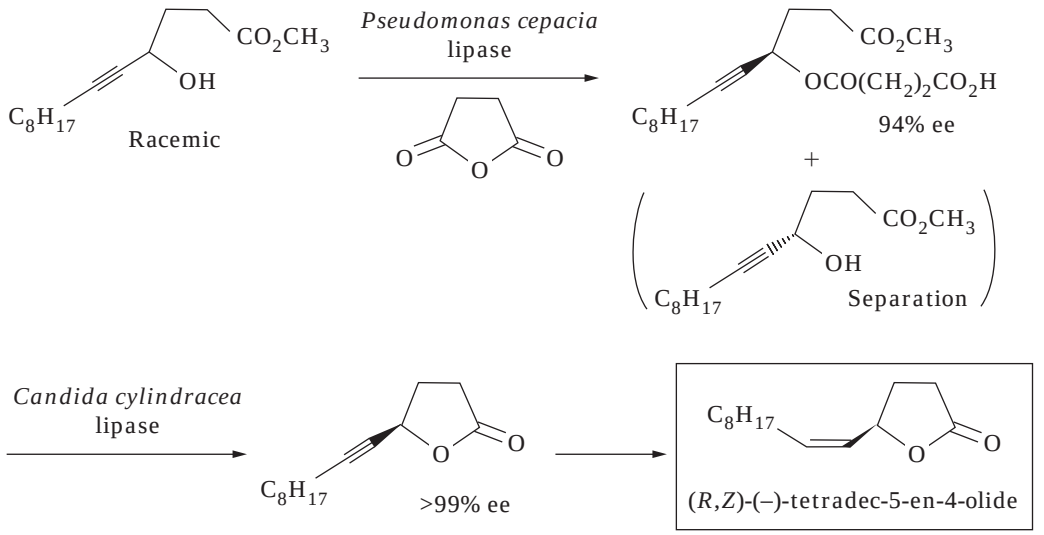


R = $CH_3(CH_2)_4$–,[a]
$CH_3(CH_2)_5$–,[b]
$CH_3CH_2CH{=}CH(CH_2)_2$–,[c]
$CH_3(CH_2)_7$–[a]

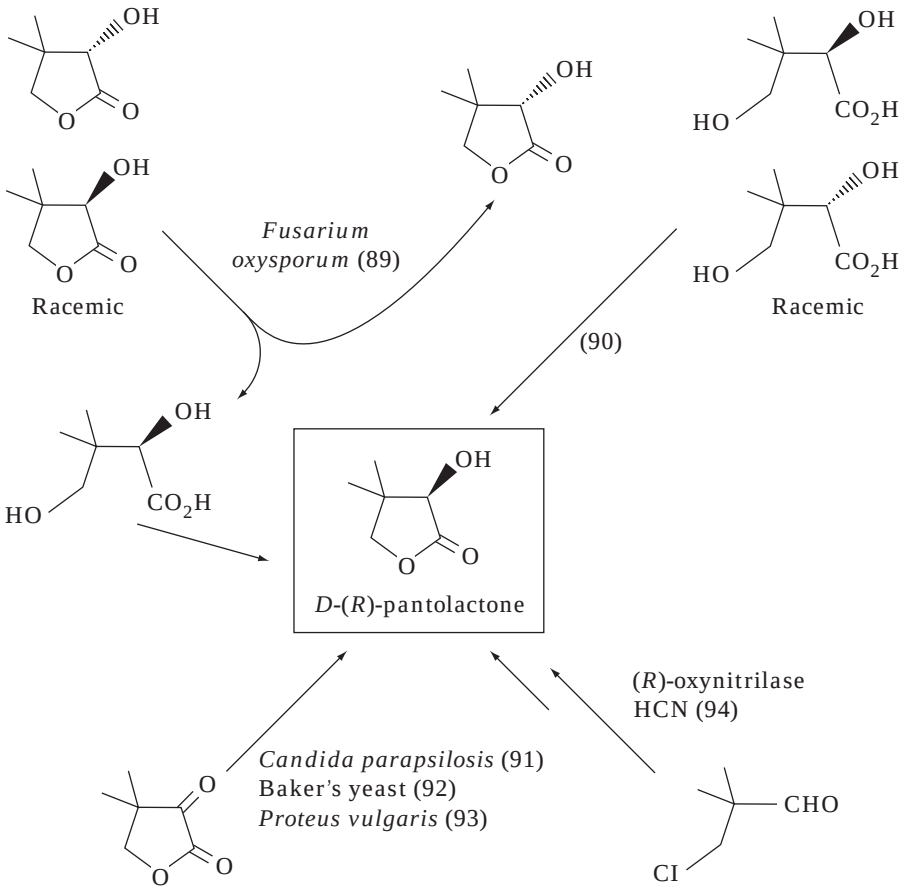
Figure 10. Enantiomerically enriched lactones as the microbial metabolites from lipids. a, Ref. 83; b, Ref. 86; c, Ref. 84; d, Ref. 85.



Scheme 18. Production of muconolactone from mandelic acid by *Pseudomonas putida*.



Scheme 19. (*R*,*Z*)-(−)-tetradec-5-en-4-olide, the Japanese beetle pheromone.



Scheme 20. Biocalytic approach to D-pantolactone.

croorganisms survived after the repetition of reaction for more than 180 days. The reverse reaction, a selective lactonization of D-pantoic acid to D-pantolactone in the equimolar mixture with L-pantoic acid worked well by the same enzyme (90).

Syntheses of D-pantolactone from ketopantolactone by microbial reduction have also been reported. After screening of microorganisms, *Candida parapsilosis* was selected as the most effective strain in terms of enantioselectivity (91). On the other hand, in the use of conventional baker's yeast, *β*-cyclodextrin as an additive was reported to enhance the enantioselectivity of the reduction (92). Asymmetric synthesis from the ketone precursor is indeed attractive to avoid the recycle of the unwanted isomer via

racemization. An interesting approach toward an enzyme-based bioreactor with a reductase from *Proteus vulgaris* and *Proteus mirabilis* has been reported (93). As this enzyme works with very wide range of synthetic electron mediators, all of the electrochemical cells (cathodes), formate, and hydrogen are available as electron donors.

The third approach is the use of (*R*)-oxynitrilase (94), which is different from the other approaches from the standpoint of the construction of the carbon skeleton of the target molecule. There is the advantage that one carbon is directly introduced as the cyano group with stereoselective C—C bond formation from an easily accessible precursor, hydroxypivalaldehyde derivative.

CONCLUSIONS

As discussed, there are versatile methods for preparing lactones by biocatalytic means. The choice depends on (*1*) the nature of the substrate, such as solubility in water and organic solvent, and volatility; (*2*) the ease of purification of the products from byproducts, unreacted starting materials, and reagents; (*3*) the accessibility of precursors, as well as the enzymes and microorganisms; and (*4*) the scale of the reaction to be performed.

ACKNOWLEDGMENT

The authors thank Professor Yasuhiro Yamada of Department of Biotechnology, Osaka University for his suggestion of preparing this manuscript as well as kind discussion.

BIBLIOGRAPHY

1. S.C. Stinton, *Chem. Eng. News* **72**, 38–72 (1994); **73**, 44–74 (1995); **75**, 38–70 (1997); K. Mori, *Chem. Commun.* 1153–1158 (1997).
2. S.C. Stinton, *Chem. Eng. News* **76**, 15–24 (1998); M. McCoy, *Chem. Eng. News* **76**, 13–19 (1998).
3. C.-H. Wong and G.M. Whitesides, *Enzymes in Synthetic Organic Chemistry*, Pergamon, New York, 1994; K. Drauz and H. Waldmann, *Enzyme Catalysis in Organic Synthesis*, VCH, Weinheim, 1995; K. Faber, *Biotransformations in Organic Chemistry*, 3rd ed. Springer, Berlin, 1997; J.-M. Fang and C.-H. Wong, *Synlett* 393–402 (1994); P. Besse and H. Veschambre, *Tetrahedron* **50**, 8885–8927 (1994); R.O. Duthaler, *Tetrahedron* **50**, 1539–1650 (1994); F. Theil, *Chem. Rev.* **95**, 2203–2227 (1995); C.-H. Wong, R.L. Halcomb, Y. Ichikawa, and T. Kajimoto, *Angew. Chem. Int. Ed. Engl.* **34**, 412–432 (1995); E. Schoffers, A. Golebiowski, and C.R. Johnson, *Tetrahedron* **52**, 3769–3826 (1996); T. Sugai, T. Yamazaki, M. Yokoyama, and H. Ohta, *Biosci. Biotechnol. Biochem.* **61**, 1419–1427 (1997); H. Ohta, *Bull. Chem. Soc. Jpn.* **70**, 2895–2911 (1997); R.D. Schmid and R. Verger, *Angew. Chem. Int. Ed. Engl.* **37**, 1608–1633 (1998).
4. L. Blanco, E. Guibé-Jampel, and G. Rousseau, *Tetrahedron Lett.* **29**, 1915–1918 (1988).
5. A.L. Gutman, K. Zuobi, and E. Guibé-Jampel, *Tetrahedron Lett.* **31**, 2037–2038 (1990).
6. R. Fellous, L. Lizzani-Cuvelier, M.A. Loiseau, and E. Sassy, *Tetrahedron: Asymmetry* **5**, 343–346 (1994).
7. G. Fantin, M. Fogagnolo, A. Medici, P. Pedrini, S. Poli, F. Gardini, and E. Guerzoni, *Tetrahedron: Asymmetry* **3**, 107–114 (1992).
8. E. Guibé-Jampel, G. Rousseau, and L. Blanco, *Tetrahedron Lett.* **30**, 67–68 (1989).
9. Y. Koichi, K. Suginaka, and Y. Yamamoto, *J. Chem. Soc., Perkin Trans. 1*, 1645–1646 (1995).
10. T. Ito, M. Shimizu, and T. Fujisawa, *Tetrahedron* **54**, 5523–5530 (1998).
11. W. Adam, P. Groer, and C.R. Saha-Möller, *Tetrahedron: Asymmetry* **8**, 833–836 (1997).
12. E. Fouque and G. Rousseau, *Synthesis*, 661–666 (1989).
13. H. Ikeda, E. Sato, T. Sugai, and H. Ohta, *Tetrahedron* **52**, 8113–8122 (1996).
14. K. Hamada, T. Akeboshi, H. Ikeda, T. Sugai, and H. Ohta, *Synlett*, 983–985 (1997).
15. M. Huffer and P. Schreier, *Tetrahedron: Asymmetry* **2**, 1157–1164 (1991).
16. I.J. Jakovac and J.B. Jones, *J. Org. Chem.* **44**, 2165–2168 (1979).
17. K. Mori, H. Mori, and T. Sugai, *Tetrahedron* **41**, 919–925 (1985).
18. A.L. Gutman, K. Zuobi, and A. Boltansky, *Tetrahedron Lett.* **28**, 3861–3864 (1987).
19. A.L. Gutman, K. Zuobi, and T. Bravdo, *J. Org. Chem.* **55**, 3546–3552 (1990).
20. T. Sugai, S. Ohsawa, H. Yamada, and H. Ohta, *Synthesis*, 1112–1114 (1990).
21. T. Sugai, H. Noguchi, and H. Ohta, *Biosci. Biotechnol. Biochem.* **56**, 122–126 (1992).
22. M.K. Sibi and J.A. Gaboury, *Tetrahedron Lett.* **33**, 5681–5684 (1992).
23. S.K. Taylor, R.F. Atkinson, E.P. Almli, M.D. Carr, T.J. van Huis, and M.R. Whittaker, *Tetrahedron: Asymmetry* **6**, 157–164 (1995).
24. H. Uyama, K. Takeya, and S. Kobayashi, *Bull. Chem. Soc. Jpn.* **68**, 56–61 (1995).
25. A.L. Gutman, D. Oren, A. Boltanski, and T. Bravdo, *Tetrahedron Lett.* **28**, 5367–5368 (1987).
26. B. Henkel, A. Kunath, and H. Schick, *Liebigs Ann. Chem.*, 809–811 (1992).
27. C. Bonini, P. Pucci, R. Racioppi, and L. Viggiani, *Tetrahedron: Asymmetry* **3**, 29–32 (1992).
28. H. Yamada, T. Sugai, H. Ohta, and S. Yoshikawa, *Agric. Biol. Chem.* **54**, 1579–1580 (1990).
29. A. Makita, T. Nihira, and Y. Yamada, *Tetrahedron Lett.* **28**, 805–808 (1987).
30. H. Yamada, S. Ohsawa, T. Sugai, and H. Ohta, *Chem. Lett.*, 1775–1776 (1989).
31. T. Sugai, O. Katoh, and H. Ohta, *Tetrahedron* **51**, 11987–11988 (1995).
32. M. Lobell and M.P. Schneider, *Tetrahedron: Asymmetry* **4**, 1027–1030 (1993).
33. Z.-W. Guo and C.J. Sih, *J. Am. Chem. Soc.* **110**, 1999–2001 (1988).
34. K. Mori and H. Tomioka, *Liebigs Ann. Chem.*, 1011–1017 (1992).
35. N. Mochizuki, H. Yamada, T. Sugai, and H. Ohta, *BioMed. Chem.* **1**, 71–75 (1993).
36. A.S. Pawar, S. Sankaranarayanan, and S. Chattopadhyay, *Tetrahedron: Asymmetry* **6**, 2219–2226 (1995).

37. T.K. Ngooi, A. Scilimati, Z.-W. Guo, and C.J. Sih, *J. Org. Chem.* **54**, 911–914 (1989).
38. A. Manzocchi, R. Casati, A. Fiecchi, and E. Santaniello, *J. Chem. Soc., Perkin Trans. 1*, 2753–2757 (1987).
39. F. Moriuchi, H. Muroi, and H. Aibe, *Chem. Lett.*, 1141–1144 (1987).
40. M. Utaka, H. Watabu, and A. Takeda, *J. Org. Chem.* **52**, 4363–4368 (1987).
41. Y. Naoshima, H. Hasegawa, and T. Saeki, *Agric. Biol. Chem.* **51**, 3417–3419 (1987).
42. S. Tsuboi, J. Sakamoto, T. Kawano, M. Utaka, and A. Takeda, *J. Org. Chem.* **56**, 7177–7179 (1991).
43. P. Mazur and K. Nakanishi, *J. Org. Chem.* **57**, 1047–1051 (1992).
44. T. Kitayama, *Tetrahedron: Asymmetry* **8**, 3765–3774 (1997).
45. K. Takabe, H. Hiyoshi, H. Sawada, M. Tanaka, A. Miyazaki, T. Yamada, T. Katagiri, and H. Yoda, *Tetrahedron: Asymmetry* **3**, 1399–1400 (1992).
46. J.B. Jones, C.J. Sih, and D. Perlman, *Applications of Biochemical Systems in Organic Chemistry*, Wiley, New York, 1976, pp. 236–401.
47. J.B. Jones and K.P. Lok, *Can. J. Chem.* **57**, 1025–1032 (1979).
48. I.J. Jakovac, H.B. Goodbrand, K.P. Lok, and J.B. Jones, *J. Am. Chem. Soc.* **104**, 4659–4665 (1982).
49. J.B. Jones, M.A.W. Finch, and I.J. Jakovac, *Can. J. Chem.* **60**, 2007–2011 (1982).
50. G.S.Y. Ng, L.-C. Yuan, I.J. Jakovac, and J.B. Jones, *Tetrahedron* **40**, 1235–1243 (1984).
51. K. Mori, M. Amaike, and H. Watanabe, *Liebigs Ann. Chem.*, 1287–1294 (1993).
52. J.B. Jones and I.J. Jakovac, *Can. J. Chem.* **60**, 19–28 (1982).
53. J.B. Jones, *Tetrahedron* **42**, 3351–3403 (1986).
54. B. Mohar, A. Stimac, and J. Kobe, *Tetrahedron: Asymmetry* **5**, 863–878 (1994).
55. H. Ohta, H. Tetsukawa, and N. Noto, *J. Org. Chem.* **47**, 2400–2404 (1982).
56. H. Luna, K. Prasad, and O. Repic, *Tetrahedron: Asymmetry* **5**, 303–306 (1994).
57. F.-C. Huang, L.F.H. Lee, R.S.D. Mittal, P.R. Ravikumar, J.A. Chan, C.J. Sih, E. Caspi, and C.R. Eck, *J. Am. Chem. Soc.* **97**, 4144–4145 (1975).
58. S.K. Taylor, N.H. Chmiel, L.J. Simons, and J.R. Vyvyan, *J. Org. Chem.* **61**, 9084–9085 (1996).
59. Y. Ohtsuka, O. Katoh, T. Sugai, and H. Ohta, *Bull. Chem. Soc. Jpn.* **70**, 483–491 (1997).
60. C.-S. Chen, Y. Fujimoto, and C.J. Sih, *J. Am. Chem. Soc.* **103**, 3580–3582 (1981); D.W. Brooks and J.T. Palmer, *Tetrahedron Lett.* **24**, 3059–3062 (1983); S. Kobayashi, K. Kamiyama, T. Iimori, and M. Ohno, *Tetrahedron Lett.* **25**, 2557–2560 (1984); Y. Yamamoto, M. Iwasa, S. Sawada, and J. Oda, *Agric. Biol. Chem.* **54**, 3269–3274 (1990); M. Lautens, S. Ma, and A. Yee, *Tetrahedron Lett.* **36**, 4184–4188 (1995); H. Kaku, M. Tanaka, Y. Norimine, Y. Miyashita, H. Suemune, and K. Sakai, *Tetrahedron: Asymmetry* **8**, 195–201 (1997), L.-M. Zhu and M.C. Tedford, *Tetrahedron* **46**, 6587–6601 (1990); C. Tamm, *Pure. Appl. Chem.* **64**, 1187–1191 (1992).
61. N.A. Donoghue, D.B. Norris, and P.W. Trudgill, *Eur. J. Biochem.* **63**, 175–192 (1976); J.M. Schwab, and W.-B. Li, *J. Am. Chem. Soc.* **105**, 4800–4808 (1983); C.T. Walsh and Y.-C.J. Chen, *Angew. Chem. Int. Ed. Engl.* **27**, 333–343 (1988); S.M. Roberts and P.W.H. Wan, *J. Mol. Catal. B: Enz.* **4**, 111–136 (1998).
62. D.R. Kelly, C.J. Knowles, J.G. Mahdi, M.A. Wright, I.N. Taylor, S.M. Roberts, P.W.H. Wan, G. Grogan, S. Pedragosa-Moreau, and A.J. Willetts, *Chem. Commun.*, 2333–2334 (1996); D.R. Kelly, *Tetrahedron: Asymmetry* **7**, 1149–1152 (1996).
63. G. Ottolina, G. Carrea, S. Colonna, and A. Rückemann, *Tetrahedron: Asymmetry* **7**, 1123–1136 (1996).
64. M.J. Taschner and D.J. Black, *J. Am. Chem. Soc.* **110**, 6892–6893 (1988).
65. M.J. Taschner, D.J. Black, and Q.-Z. Chen, *Tetrahedron: Asymmetry* **4**, 1387–1390 (1993).
66. S. Rissom, U. Schwarz-Linek, M. Vogel, V.I. Tishkov, and U. Kragl, *Tetrahedron: Asymmetry* **8**, 2523–2526 (1997).
67. B. Adger, M.T. Bes, G. Grogan, R. McCague, S. Pedragosa-Moreau, S.M. Roberts, R. Villa, P.W.H. Wan, and A.J. Willetts, *J. Chem. Soc., Chem. Commun.*, 1563–1564 (1995).
68. V. Alphand, A. Archelas, and R. Furstoss, *J. Org. Chem.* **55**, 347–350 (1990).
69. M.T. Bes, R. Villa, S.M. Roberts, P.W.H. Wan, and A. Willetts, *J. Mol. Catal. B: Enz.* **1**, 127–134 (1996).
70. J.D. Stewart, K.W. Reed, C.A. Martinez, J. Zhu, G. Chen, and M.M. Kayser, *J. Am. Chem. Soc.* **120**, 3541–3548 (1998).
71. V.A. Alphand and R. Furstoss, *J. Org. Chem.* **57**, 1306–1309 (1992).
72. F. Petit and R. Furstoss, *Tetrahedron: Asymmetry* **4**, 1341–1352 (1993).
73. K. Königsberger and H. Griengl, *BioMed. Chem.* **2**, 595–604 (1994).
74. S. Colonna, N. Gaggero, P. Pasta, and G. Ottolina, *Chem. Commun.*, 2303–2308 (1996); H.L. Holland, L.J. Allen, M.J. Chernishenko, M. Diez, A. Kohl, J. Ozog, and J.-X. Gu, *J. Mol. Catal. B: Enz.* **3**, 303–309 (1997).
75. K. Mizuno, S. Sakuda, T. Nihira, and Y. Yamada, *Tetrahedron* **50**, 10849–10858 (1994).
76. T. Fujita, M. Tanaka, Y. Norimine, H. Suemune, and K. Sakai, *J. Org. Chem.* **62**, 3824–3830 (1997).
77. H. van der Deen, A.D. Cuiper, R.P. Hof, A. van Oeveren, B.L. Feringa, and R.M. Kellogg, *J. Am. Chem. Soc.* **118**, 3801–3803 (1996); M. van den Heuvel, A.D. Cuiper, H. van der Deen, R.M. Kellogg, and B.L. Feringa, *Tetrahedron Lett.* **38**, 1655–1658 (1997).
78. S. Drioli, F. Felluga, C. Forzato, P. Nitti, and G. Pitacco, *J. Chem. Soc., Chem. Commun.* 1289–1290 (1996); S. Drioli, F. Felluga, T. Fortato, P. Nitti, G. Pitacco, and E. Valentin, *J. Mol. Catal. B: Enz.* **3**, 203–207 (1997), S. Drioli, F. Felluga, C. Forzato, P. Nitti, G. Pitacco, and E. Valentin, *J. Org. Chem.* **63**, 2385–2388 (1998).
79. H. Kapeller, W.G. Jary, W. Hayden, and H. Griengl, *Tetrahedron: Asymmetry* **8**, 245–251 (1997).
80. G. Fronza, C. Fuganti, P. Grasselli, L. Malpezzi, and A. Mele, *J. Org. Chem.* **59**, 3487–3489 (1994).
81. G. Fronza, C. Fuganti, P. Grasselli, R. Pulido-Fernandez, S. Servi, A. Tagliani, and M. Terreni, *Tetrahedron* **47**, 9247–9252 (1991).
82. G. Fantin, M. Fogagnolo, P. Giovannini, A. Medici, E. Pagnotta, P. Pedrini, and A. Trincone, *Tetrahedron: Asymmetry* **5**, 1631–1634 (1994).
83. R. Cardillo, G. Fronza, C. Fuganti, P. Grasselli, A. Mele, D. Pizzi, G. Allegrone, M. Barbeni, and A. Pisciotta, *J. Org. Chem.* **56**, 5237–5241 (1991).
84. C. Fuganti, P. Grasselli, G. Zucchi, S. Allegrone, and A. Barbeni, *BioMed. Chem. Lett.* **3**, 2777–2780 (1993).

85. W. Albrecht and R. Tressl, *Tetrahedron: Asymmetry* **4**, 1391–1396 (1993).
86. S.-J. Lin, S.-L. Lee, and C.-C. Chou, *J. Ferment. Bioeng.* **82**, 42–45 (1996).
87. D.W. Ribbons and A.G. Sutherland, *Tetrahedron* **50**, 3587–3594 (1994).
88. E. Fukusaki, S. Senda, Y. Nakazono, and T. Omata, *Tetrahedron* **47**, 6223–6230 (1991); E. Fukusaki, S. Senda, Y. Nakazono, and T. Omata, *Biosci. Biotechnol. Biochem.* **56**, 1160–1161 (1992); E. Fukusaki and S. Satoda, *J. Mol. Catal. B: Enz.* **2**, 257–269 (1997).
89. S. Shimizu, M. Kataoka, K. Shimizu, M. Hirakata, K. Sakamoto, and H. Yamda, *Eur. J. Biochem.* **209**, 383–390 (1992); M. Kataoka, K. Shimizu, K. Sakamoto, H. Yamada, and S. Shimizu, *Appl. Microbiol. Biotechnol.* **44**, 333–338 (1995).
90. M. Kataoka, M. Hirakata, K. Sakamoto, H. Yamada, and S. Shimizu, *Enzyme Microb. Technol.* **19**, 307–310 (1996).
91. H. Hata, S. Shimizu, S. Hattori, and H. Yamada, *J. Org. Chem.* **55**, 4377–4380 (1990).
92. K. Nakamura, S. Kondo, and A. Ohno, *BioMed. Chem.* **2**, 433–438 (1994).
93. R. Eck and H. Simon, *Tetrahedron: Asymmetry* **5**, 1419–1422 (1994).
94. F. Effenberger, J. Eichhorn, and J. Roos, *Tetrahedron: Asymmetry* **6**, 271–282 (1995); F. Effenberger, *Angew. Chem. Int. Ed. Engl.* **33**, 1555–1556 (1994); H. Griengl, A. Hickel, D.V. Johnson, C. Kratky, M. Schmidt, and H. Scwab, *Chem. Commun.*, 1933–1940 (1997).

See also BIOCATALYSIS DATABASES.

LACTONOHYDROLASE

SAKAYU SHIMIZU
MICHIHIKO KATAOKA
Kyoto University
Kyoto, Japan

KEY WORDS

Aldonate lactone
Aldonolactonase
Esterase
Gluconolactonase
Lactonase
Lipase
Optical resolution
Pantoic acid
Pantolactone
Stereospecific synthesis

OUTLINE

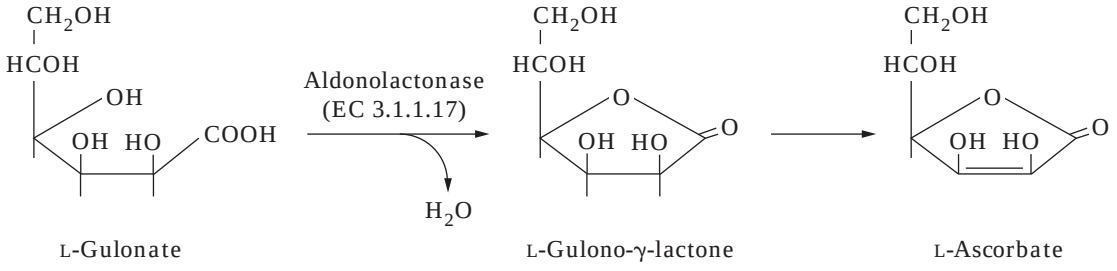
INTRODUCTION

The enzymes catalyzing the reversible or irreversible hydrolysis of carboxylic esters to the corresponding carboxylic acids and alcohols (EC 3.1.1.) play an important role in the biosynthesis and biodegradation of various compounds. In particular, the functions and structures of lipase (EC 3.1.1.3) and esterase (EC 3.1.1.1) have been well analyzed (1–5). Lactonohydrolases (lactonases) also catalyze the hydrolysis of carboxylic esters, but act on the intramolecular ester bonds of lactone compounds. They are involved in the synthesis and degradation of various lactone compounds in vivo. For example, gluconolactonase (aldonolactonase; EC 3.1.1.17) was suggested to participate in the formation of L-gulono-γ-lactone from L-gulonate in L-ascorbate biosynthesis (6). Although the activities of lactonohydrolases have been detected in various organisms (6–13), there has not been a detailed study on their enzyme chemistry so far.

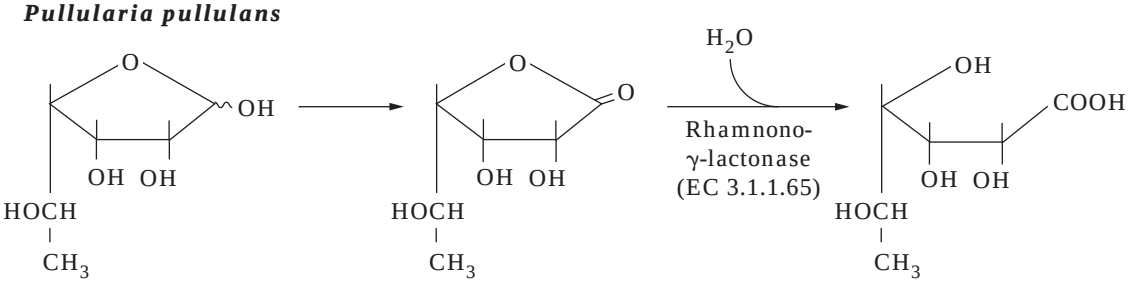
DIVERSITY OF LACTONOHYDROLASES

Lactonohydrolases reversibly catalyze the hydrolysis of lactone compounds into hydroxy acids, that is, they mediate the interconversion between the lactone and acid forms of hydroxy carboxylic acids. Gluconolactonase (aldonolactonase; EC 3.1.1.17), L-arabinonolactonase (EC 3.1.1.15), and D-arabinonolactonase (EC 3.1.1.30) specifically hydrolyze aldonate lactones to the respective aldonic acids. The activities of these enzymes have been detected in various organisms (6–13). As to the roles of these enzymes in vivo, it has been suggested that gluconolactonase participates in the formation of L-gulono-γ-lactone from L-gulonate in L-ascorbate biosynthesis (Fig. 1) (6). L-Rhamnonolactonase (EC 3.1.1.65) and D-xylonolactonase (EC 3.1.1.68) of *Pullularia pullulans* and *Pseudomonas fragi*, respectively, are involved in the oxidative degradation of L-rhamnose and D-xylose, respectively (Fig. 1) (14,15). In addition to these enzymes, pyridoxolactonase (EC 3.1.1.27), dihydrocoumarin hydrolase (EC 3.1.1.35), limonin-D-ring-lactonase (EC 3.1.1.36) and deoxylimonate A-ring-lactonase (EC 3.1.1.46), and actinomycin lactonase (EC 3.1.1.39) have been reported to be the enzymes

Animal



Pullularia pullulans



Pseudomonas fragi

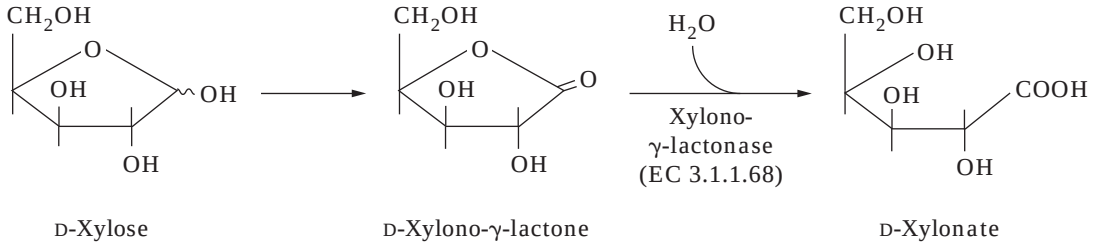


Figure 1. Lactonohydrolase reactions involved in ascorbate biosynthesis and the degradation of L-rhamnose and D-xylose.

Table 1. Properties of the Lactonohydrolases from *F. oxysporum* and *B. protophormiae*

Origin	*F. oxysporum*		*B. protophormiae*	
Native M_r	125,000		62,000	
Subunit M_r	60,000		26,500	
Number of subunits	2		2	
Substrate specificity	K_m (mM)	V_{max} (μmol/min/mg)	K_m (mM)	V_{max} (μmol/min/mg)
D-(−)-Pantolactone	120	653	—	—
L-(+)-Pantolactone	—	—	3.59	13.7
D-Galactono-γ-lactone	3.6	1,440	—	—
Dihydrocoumarin	6.3	2,800	2.56×10^3	7.74×10^4
pH stability	5.0–10		>5.0	
Optimum temperature	50 °C		45 °C	
Thermal stability	<50 °C		<50 °C	
Inhibitors	L-(+)-pantolactone, chelating reagents		Chelating reagents	

Table 2. Substrate Specificity of the Lactonohydrolase from *F. oxysporum*

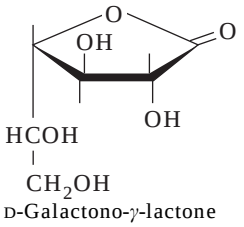
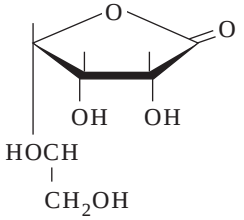
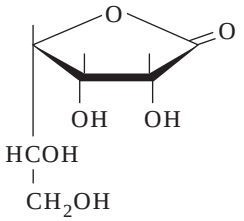
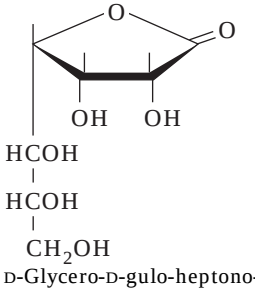
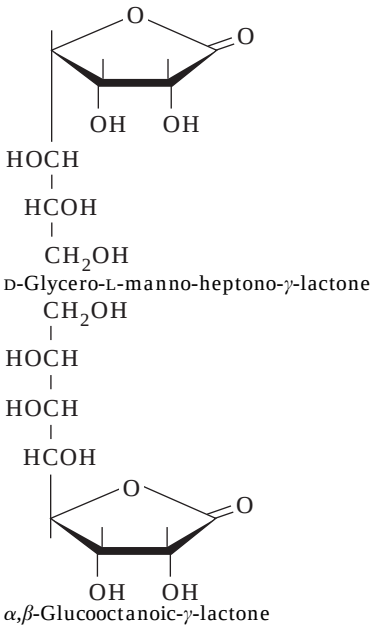
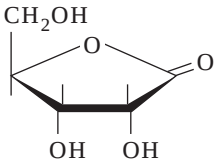
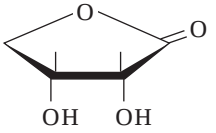
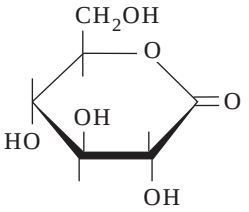
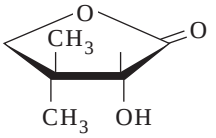
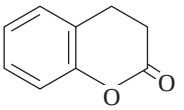
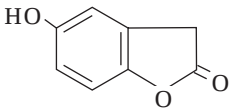
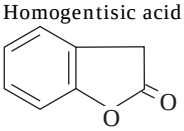
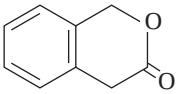
	Relative hydrolyzing activity (%)	K_m (mM)
Substrate (150 mM)		
 D-Galactono-γ-lactone	100	3.6
 L-Mannono-γ-lactone	128	23
 D-Gulono-γ-lactone	167	29
 D-Glycero-D-gulo-heptono-γ-lactone	59.4	118
 D-Glycero-L-manno-heptono-γ-lactone	132	167
α,β-Glucooctanoic-γ-lactone	16.4	—

Table 2. Substrate Specificity of the Lactonohydrolase from *F. oxysporum* (*continued*)

	Relative hydrolyzing activity (%)	K_m (mM)
 D-Ribono-γ-lactone	39.0	2.5
 D-Erythrono-γ-lactone	117	377
 D-Glucono-δ-lactone	107	—
 D-Pantolactone	48.3	120
Substrate (2.5 mM)		
 Dihydrocoumarin	100	6.3
 Homogentisic acid lactone	16.6	2.5
 2-Coumaramone	16.4	8.7
 3-Isochromanone	0.58	4.4

involved in the degradation pathways for vitamin B-6 (16), aromatic compounds (17,18), limonoids (19,20), and actinomycin (21), respectively. As for molecular biochemical studies on lactonohydrolases, only the gene encoding gluconolactonase from *Zymomonas mobilis* has been cloned and sequenced (22), and there has been no report on the structure of any other lactonohydrolase.

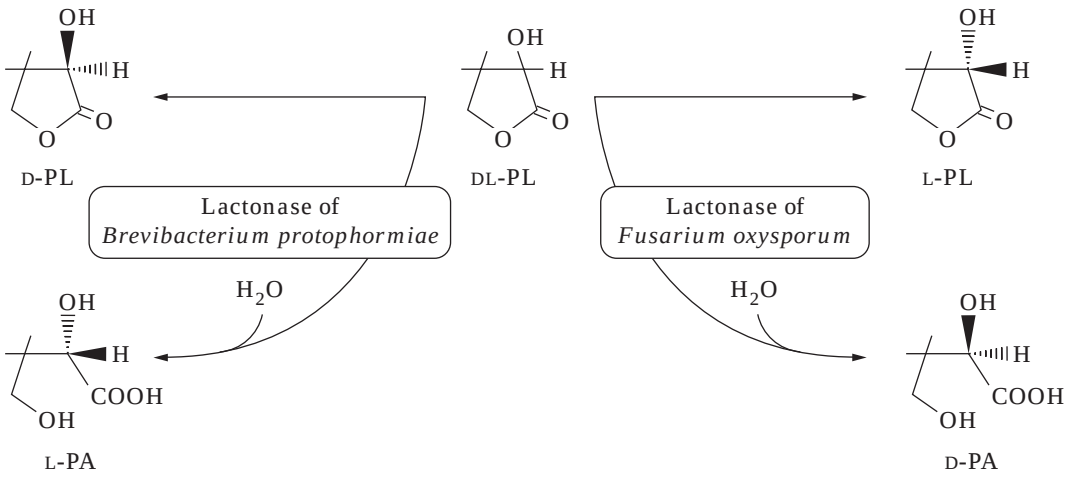


Figure 2. Novel microbial lactonohydrolases catalyzing the stereospecific hydrolysis of pantolactone. PL = pantolactone; PA = pantoic acid.

NOVEL MICROBIAL LACTONOHYDROLASES

During the course of studies on the microbial production of chiral intermediates for D-pantothenate (23–25), Shimizu et al. found that several microorganisms, such as *Fusarium*, *Brevibacterium*, and so on, produce a novel enzyme that catalyzes the hydrolysis of aldonate or aromatic lactones (26,27).

The enzyme isolated from cells of *Fusarium oxysporum* has a relative molecular mass of 125 kDa and a subunit molecular mass of 60 kDa. The enzyme thus appears to be a dimer consisting of identical subunits. The enzyme contains about 1 mol of calcium per subunit and 15.4% (w/w) glucose equivalent of carbohydrate. Calcium seems to be necessary for both enzyme activity and stability. The enzyme carries three kinds of N-linked high-mannose-type sugar chains at the 28th, 106th, 179th, and 277th Asn residues. The carbohydrate moiety is an essential participant in the stabilization of the enzyme. Table 1 summarizes the properties of the enzyme (26). The enzyme hydrolyzes aldonate lactones, such as D-galactono-γ-lactone, L-mannono-γ-lactone, D-gulono-γ-lactone, and D-glucono-δ-lactone, stereospecifically, but the corresponding enantiomers (i.e., L-galactono-γ-lactone, D-mannono-γ-lactone, L-gulono-γ-lactone, and L-glucono-δ-lactone, respectively) do not serve as substrates. All these substrates have downward hydroxy groups at their 2-position carbons in common, according to Haworth's projection system, (Table 2), whereas their enantiomers have upward hydroxy groups at their 2-position carbons. The enzyme can also catalyze the asymmetric hydrolysis of D-(−)-pantolactone. For every substrate, the reverse reaction (i.e., lactonization of the aldonic acid) was observed. Aromatic lactones, such as dihydrocoumarin, homogentisic acid lactone, 2-coumaranone, and 3-isochromanone, were also effective substrates. They are good substrates for dihydrocoumarin hydrolase (EC 3.1.1.35) and have no hydroxy group at their 2-position carbons (17). With regard to the reverse reactions, no detectable lactone compounds were formed.

A similar kind of lactonohydrolase was found in *Brevibacterium protophormiae* (Table 1) (27). This bacterial enzyme only hydrolyzes L-(+)-pantolactone; D-(−)-pantolactone does not act as a substrate. Various kinds of aromatic lactones, such as dihydrocoumarin, homogentisic acid lactone and so on, are also hydrolyzed, but aldonate lactones are not hydrolyzed. The relative molecular mass of the native enzyme is 62 kDa, and its subunit molecular mass is 26.5 kDa.

APPLICATION OF MICROBIAL LACTONOHYDROLASE TO THE OPTICAL RESOLUTION OF RACEMIC PANTOLACTONE

Calcium D-(+)-pantothenate is mainly used as an additive for animal feeds and for various pharmaceutical products. At present, the commercial production of D-(+)-pantothenate depends exclusively on chemical synthesis. The chemical synthesis process involves reactions yielding racemic pantolactone from isobutyraldehyde, formaldehyde, and cyanide; optical resolution of the racemic pantolactone to D-(−)-pantolactone; and condensation of the D-(−)-pantolactone with β-alanine. Drawbacks of this chemical process are the troublesome resolution of the racemic pantolactone, for which the use of an expensive alkaloid as a resolving agent is unavoidable. Recently, Kataoka et al. showed that the microbial lactonohydrolase could be used as a catalyst for the resolution of racemic pantolactone (28–30).

The principle of the optical resolution of racemic pantolactone is shown in Figure 2. If racemic pantolactone is used as a substrate for the hydrolysis by the stereospecific lactonohydrolase, only the D-(−)- or L-(+)-pantolactone might be converted to the D-(−)- or L-(+)-pantoic acid, and the L-(+)- or D-(−)-enantiomer might remain intact, respectively. Consequently, the racemic mixture could be resolved into D-(−)-pantoic acid and L-(+)-pantolactone, or D-(−)-pantolactone and L-(+)-pantoic acid. In the case of

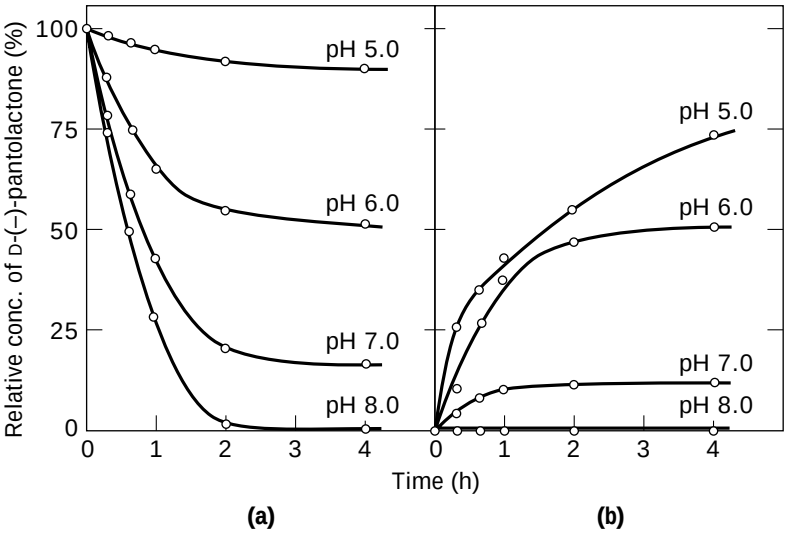


Figure 3. Effect of pH on the reaction equilibrium of the *Fusarium* lactonohydrolase. The enzymatic hydrolysis of D-(−)-pantolactone (**a**) and enzymatic lactonization of D-(−)-pantoic acid (**b**) were carried out at the pHs indicated.

L-(+)-pantolactone-specific lactonohydrolase, the optical purity of the remaining D-(−)-pantolactone might be low, except when the hydrolysis of L-(+)-pantolactone is complete. On the other hand, using the D-(−)-pantolactone-specific lactonohydrolase, D-(−)-pantoic acid of high optical purity could be constantly obtained irrespective of the hydrolysis yield. Therefore, the enzymatic resolution of racemic pantolactone with the D-(−)-pantolactone-specific lactonohydrolase was investigated (28–30).

The distribution of the D-(−)-pantolactone-specific lactonohydrolase is narrow; only filamentous fungi in three genera (i.e., *Fusarium*, *Gibberella*, and *Cylindrocarpon*) show high hydrolytic activity. The D-(−)-pantolactone hydrolysis reaction catalyzed by the purified *Fusarium* lactonohydrolase favored alkaline pH, and the D-(−)-pantoic acid lactonization reaction favored rather acidic pH. At pH 7.0, about 80% of the substrate, pantolactone, was present as pantoic acid, and about 75% of the pantoic acid added was present as pantolactone at pH 5.0. The pantoic acid/pantolactone ratio in the reaction mixture was nearly 1:1 at pH 6.0 (Fig. 3) (26).

Optical Resolution of Racemic Pantolactone with the Microbial Lactonohydrolase

Washed cells of *F. oxysporum* as well as the purified lactonohydrolase catalyze well the stereospecific hydrolysis of D-(−)-pantolactone. As shown in Figure 4, D-(−)-pantolactone was stereospecifically hydrolyzed to D-(−)-pantoic acid with a substrate concentration of 700 mg/mL. The formation of L-(+)-pantoic acid was barely detected. After 24 h, the total concentration of pantolactone hydrolyzed in the reaction mixture reached 322 mg/mL, with an optical purity of 96% ee. After the removal of L-(+)-pantolactone from the reaction mixture by solvent extraction, D-(−)-pantoate remaining could be easily converted to D-(−)-pantolactone by heating under acidic conditions.

Stereoselective Production of D-(−)-Pantolactone from Racemic Pantoic Acid through Microbial Stereoselective Lactonization

As described earlier, an alkaline pH favors the hydrolysis of D-(−)-pantolactone, and a rather acidic pH favors the lactonization of D-(−)-pantoate. At pH 5.0, the ratio of D-(−)-pantoate and D-(−)-pantolactone after the reaction is 1:3, whereas that at pH 7.0 is 4:1. These results strongly suggest that this stereospecific lactonization reaction might be used for the asymmetric production of D-(−)-pantolactone from racemic pantoate.

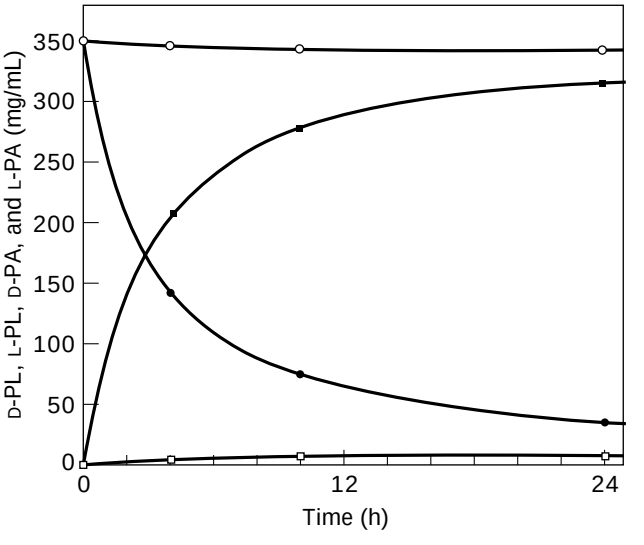


Figure 4. Optical resolution of racemic pantolactone with washed cells of *F. oxysporum*. The concentrations of pantoic acid given are those of pantolactone. ●, D-(−)-pantolactone (D-PL); ○, L-(+)-pantolactone (L-PL); ■, D-(−)-pantoic acid (D-PA); □, L-(+)-pantoic acid (L-PA).

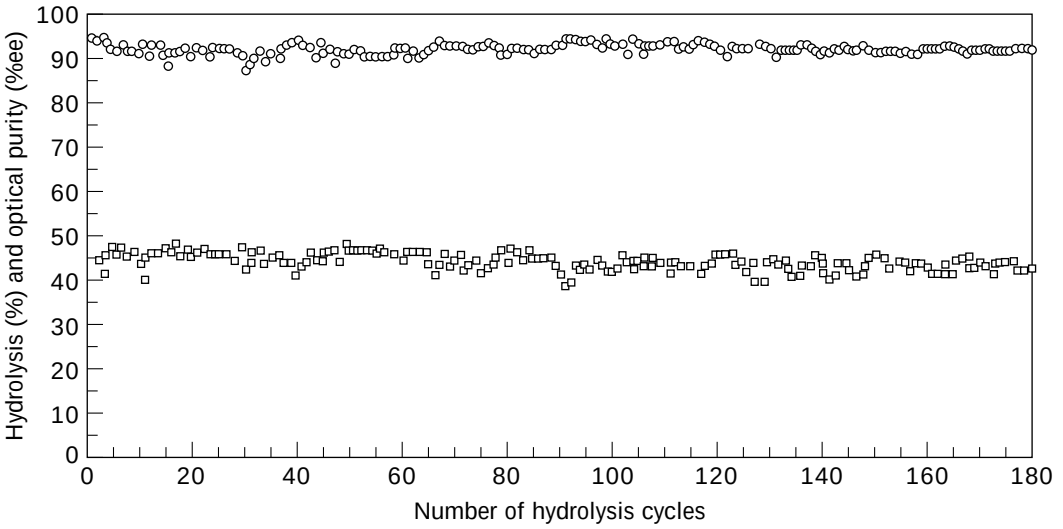


Figure 5. Stereospecific hydrolysis of pantolactone by *F. oxysporum* mycelia entrapped in a calcium alginate gel. Here, 280 L of immobilized mycelia (containing 15.2 kg of wet cells) was incubated with 350 L of an aqueous DL-pantolactone solution (350 g/L) at 30 °C for 21 h. The pH of the mixture was automatically controlled at 6.8 to 7.2 with 15 M NH_4OH. All the immobilized mycelia, after filtration from the reaction mixture, were used for the subsequent reaction. The hydrolysis reaction was carried out 180 times. ○ = optical purity of D-pantoate; □ = hydrolysis rate.

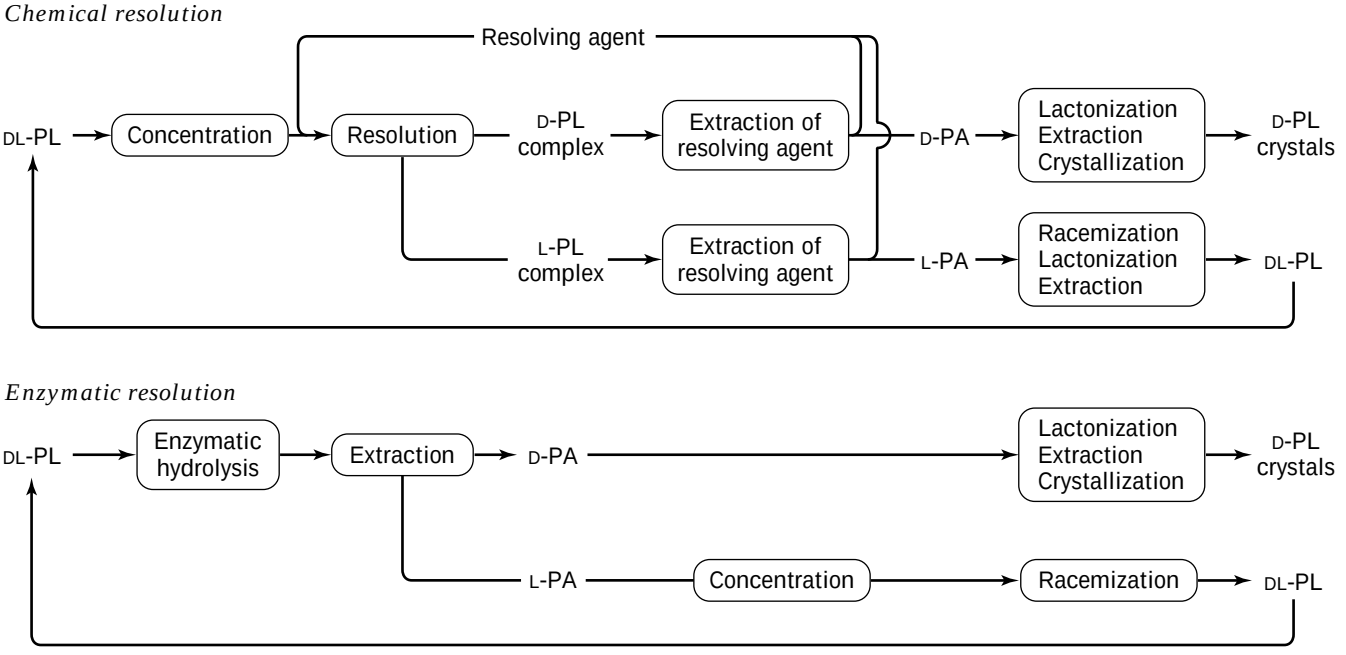


Figure 6. Comparison of the enzymatic and conventional chemical resolution processes for racemic pantolactone.

When 385-mM racemic pantoic acid (50 mg/mL as racemic pantolactone) was incubated with washed cells of *F. oxysporum* at 30 °C for 14 h at pH 4.5 to 5, 40% of the pantoate was lactonized, with an enantioselectivity of 83% ee for the D-(−)-isomer. In an organic solvent-water two-phase system for which an aqueous solution of 1,540 mM racemic pantoic acid (200 mg/mL as racemic pantolactone) containing 50 mM $CaCl_2$ and cells was incubated with ethyl acetate at 30 °C for 20 h at pH 4.5 to 5, D-(−)-pantoate in the racemic mixture was almost stoichiometrically lactonized to D-(−)-pantolactone (90% ee). This improvement of the two-phase reaction system was suggested to be the result of a shift of the reaction equilibrium during the lactonization on extraction of the reaction product, i.e., D-(−)-pantolactone moved into the organic phase.

LARGE-SCALE RESOLUTION OF RACEMIC PANTOLACTONE WITH IMMOBILIZED CELLS OF *F. OXYSPORUM*

Practical hydrolysis of the D-(−)-isomer in a racemic mixture is carried out using immobilized mycelia of *F. oxysporum* as the catalyst. A stable catalyst with high hydrolytic activity can be prepared by entrapping the fungal mycelia in a calcium alginate gel. When the immobilized mycelia were incubated in a reaction mixture containing 350 g/L of racemic pantolactone for 21 h at 30 °C under automatic pH control conditions (pH 6.8 to 7.2), 90 to 95% of the D-(−)-enantiomer was hydrolyzed (optical purity, 90 to 97% ee). After repeated reactions for 180 times (i.e., 180 days), the immobilized mycelia retained more than 90% of their initial activity (Fig. 5).

The overall process for the present enzymatic resolution is compared with the conventional chemical process in Figure 6. The enzymatic process avoids several tedious steps that are necessary in chemical resolution and thus is highly advantageous for practical purposes. There have been several reports of the application of enzymatic asymmetric hydrolysis to the optical resolution of pantolactone (31,32). In these cases, esterified substrates, such as *O*-acetyl or *O*-formyl pantolactone, and lipases were used as starting materials and catalysts, respectively. Because the lactonohydrolase of *F. oxysporum* hydrolyzes the intramolecular ester bond of pantolactone, it is not necessary to modify the substrate, pantolactone. This is one of the practical advantages of this enzyme.

BIBLIOGRAPHY

1. J.D. Schrag, Y. Li, S. Wu, and M. Cygler, *Nature* **351**, 761–764 (1991).
2. F.K. Winkler, A. D'Arcy, and W. Hunziker, *Nature* **343**, 771–774 (1990).
3. H.D. Beer, G. Wohlfahrt, J.E.G. McCarthy, D. Schomburg, and R.D. Schmid, *Protein Eng.* **9**, 507–517 (1996).
4. J. Uppenberg, N. Ohrner, M. Norin, K. Hult, G.J. Kleywegt, S. Patkar, V. Waagen, T. Anthonsen, and T.A. Jones, *Biochemistry* **34**, 16838–16851 (1995).
5. M. Cygler, J.D. Schrag, J.L. Sussman, M. Harel, I. Silman, M.K. Gentry, and B.P. Doctor, *Protein Sci.* **2**, 366–382 (1993).
6. C. Bublitz and A.L. Lehninger, *Biochim. Biophys. Acta* **47**, 288–297 (1961).
7. J. Winkelman and A.L. Lehninger, *J. Biol. Chem.* **233**, 794–799 (1958).
8. K. Yamada, *J. Biochem.* **46**, 361–372 (1959).
9. K. Yamada, S. Ishikawa, and N. Shimazono, *Biochim. Biophys. Acta* **32**, 253–255 (1959).
10. M. Kawada, H. Takiguchi, Y. Kagawa, K. Suzuki, and N. Shimazono, *J. Biochem.* **51**, 405–415 (1962).
11. F. Hucho and K. Wallenfels, *Biochim. Biophys. Acta* **276**, 176–179 (1972).
12. R. Weinberg and M. Doudoroff, *J. Biol. Chem.* **217**, 607–624 (1955).
13. N.J. Palleroni and M. Doudoroff, *J. Bacteriol.* **74**, 180–185 (1957).
14. L.U. Rigo, L.R. Marechal, M.M. Vieira, and L.A. Veiga, *Can. J. Microbiol.* **31**, 817–822 (1985).
15. J. Buchert and L. Viikari, *Appl. Microbiol. Biotechnol.* **27**, 333–336 (1988).
16. Y.-J. Jong and E.E. Snell, *J. Biol. Chem.* **261**, 15112–15114 (1986).
17. T. Kosuge and E.E. Conn, *J. Biol. Chem.* **237**, 1653–1656 (1962).
18. M. Casella, M. Grifoll, J.M. Bayona, and A.M. Solanas, *Appl. Environ. Microbiol.* **63**, 819–826 (1997).
19. S. Hasegawa, R.D. Bennett, and C.P. Verdon, *Phytochemistry* **19**, 1445–1447 (1980).
20. M.T. Merino, L. Humanes, J.M. Roldàn, J. Diez, and A. Lopez-Ruiz, *Biotechnol. Lett.* **18**, 1175–1180 (1996).
21. C.T. Hou and D. Perlman, *J. Biol. Chem.* **245**, 1289–1295 (1970).
22. V. Kanagasundaram and R. Scopes, *Biochim. Biophys. Acta* **1171**, 198–200 (1992).
23. H. Yamada and S. Shimizu, *Angew. Chem. Int. Ed. Engl.* **27**, 622–642 (1988).
24. S. Shimizu and H. Yamada, in E.J. Vandamme ed., *Biotechnology of Vitamins, Pigments, and Growth Factors*, Elsevier Science Publishers, London, 1989, pp. 199–219.
25. S. Shimizu and H. Yamada, in T.O. Baldwin, F.M. Raushel, and A.I. Scott eds., *Chemical Aspects of Enzyme Biotechnology*, Plenum Press, New York, 1990, pp. 151–163.
26. S. Shimizu, M. Kataoka, K. Shimizu, M. Hirakata, K. Sakamoto, and H. Yamada, *Eur. J. Biochem.* **209**, 383–390 (1992).
27. M. Kataoka, J. Nomura, K. Nose, and S. Shimizu, *Nippon Nogeikagaku Kaishi* **70**, 374 (1996).
28. M. Kataoka, K. Shimizu, K. Sakamoto, H. Yamada, and S. Shimizu, *Appl. Microbiol. Biotechnol.* **43**, 974–977 (1995).
29. M. Kataoka, K. Shimizu, K. Sakamoto, H. Yamada, and S. Shimizu, *Appl. Microbiol. Biotechnol.* **44**, 333–338 (1995).
30. M. Kataoka, K. Shimizu, K. Sakamoto, H. Yamada, and S. Shimizu, *Enzyme Microb. Technol.* **19**, 307–310 (1996).
31. H.S. Bevinakatti and R.V. Newadkar, *Biotechnol. Lett.* **11**, 785–788 (1989).
32. B.I. Glänzer, K. Faber, and H. Griengl, *Enzyme Microb. Technol.* **10**, 689–690 (1988).

See also BIOCATALYSIS DATABASES.

LYSINE. See AMINO ACIDS, PRODUCTION PROCESSES.

MALATE, D-MALATE

MAKOTO UEDA
YASUHISA ASANO
HIDEAKI YAMADA
Mitsubishi Chemical Co.
Yokohama, Japan

KEY WORDS

D-Malate
Maleate
Maleate hydratase

OUTLINE

INTRODUCTION

Optically active malate is a key intermediate in the synthesis of pharmaceuticals and agrochemicals, as well as a resolving agent for racemic compounds. L-Malate can be obtained economically from natural resources or enzymatically from fumarate with the enzyme fumarase (EC 4.2.1.2). However, the unnatural stereoisomer, D-malate, is expensive. Several methods for synthesizing D-malate have been described. The latest biological method for the production of D-malate is the stereoselective hydration of maleate through the reaction of maleate hydratase (D-malate hydrolyase, EC 4.2.1.31) derived from microorganisms. Maleate hydratase catalyzes the hydration of maleate to D-malate through *trans* addition to the double bond. In early studies on maleate hydratase, Sacks and Jensen demonstrated that an enzyme mixture from corn kernels catalyzed the conversion of maleate to malate, although they did not mention its optical purity (1). Subsequently, several workers detected the enzyme in mammalian kidneys (2) and microorganisms (3). Since 1992, the enzyme have been used in investigations on D-malate production in microbial reactions (4–9). In this chapter, D-malate production involving microbial maleate hydratase and the characteristics of the enzyme are described. Many applications of D-malate are found in the synthesis of intermediates for pharmaceuticals; however, the commercial production of D-malate has not yet been realized.

METHODS OF D-MALATE PRODUCTION

Chemical and Physical Methods

The chiral pool refers to relatively inexpensive, readily available, optically active natural products. L(+)-Tartarate is a natural product of which about 50,000 tons are purified from agol for wine production per year. Gao and Sharpless reported D-malate synthesis from L(+)-tartarate with a 55% yield (10). Amino acids are valuable as chiral synthons. The nitrous deamination of α-amino acids is a well-known process, which was claimed to proceed with retention of amino acid configuration. Henrot et al. synthesized the L- or D-form of malic acid from aspartic acid (11). Asymmetric hydrogenation has been an important advance in organic synthesis in recent years. Noyori and Takasago Inc. have demonstrated the successful commercialization of rhodium-BINAP complexes as asymmetric hydrogenation catalysts. D-malate is formed via a critical bond-breaking process from the sodium salt of epoxysuccinic acid with rhodium phosphine complexes as the catalyst (12). Optically active α-hydroxy acids, such as malate and citramalate, are derived from optically pure β-lactones, which are asymmetrically synthesized from ketene and chloral using quinine or quinidine as a catalyst. These β-lactones are available on a commercial scale from Lonza Ltd. (13).

Classical resolution is most often used to produce optically active materials. It is based on the interaction of a racemic product with an optically active resolving agent, yielding two diastereometric derivatives. The resultant diastereoisomeric salts can usually be separated by selective crystallization. Aliphatic or aromatic amines are effective for D-malate resolution (14–16).

Biological Methods

Recently, biological methods for the production of D-malate have been reported. The most traditional means of producing optical isomers is microbiological resolution, in which a microorganism only degrades the opposite isomer. D-Malate can be produced from a malate racemic mixture using microorganisms that assimilate L-malate (17). Several microorganisms can specifically assimilate L-malate, and D-malate of high optical purity (100%) was obtained from *Acinetobacter Iwoffi* culture broth (18). The enzymatic resolution of DL-malate can be achieved with fumarase and aspartase, resulting in the conversion of the DL-malate to D-malate and L-Asp (19). (*R*)-Dimethyl malate was obtained from a racemic mixture by hydrolysis with pig liver

esterase (93% ee at 0 °C in 20% aqueous methanol) (20) or *Rhizopus* lipase (99% ee) (21). Santaniello et al. reported that *Lactobacillus fermentum* ILC G18D produced (*R*)-diethyl malate through the enantioselective reduction of sodium diethyl oxalacetate, the enantioselectivity being 80% (20).

D-MALATE PRODUCTION FROM MALEATE

Maleate

Maleic anhydride, which is industrially produced as a raw material for unsaturated polyesters, succinic anhydride, γ-butyrolactone, 1,4-butanediol, tetrahydrofuran, fumarate, DL-malate, and so on, is an inexpensive chemical. Several industrial processes have been developed for the manufacture of maleic anhydride. Since 1974, *n*-butane has been increasingly used in place of benzene as the raw material for the production of maleic anhydride. Fluidized catalysts containing vanadium and phosphorus oxides are used for the oxidation of *n*-butane to maleic anhydride. The *n*-butane method is less expensive and nontoxic. Therefore, for maleic anhydride production in the United States, there has been a change to methods involving *n*-butane instead of benzene. In the case of Mitsubishi Chemical Co., the manufacturing process for maleic anhydride was changed to a *n*-butane method with a fluid-bed reactor from the butene method in 1984. Annually, manufacturers produce 630,000 tons of maleic anhydride from *n*-butane, 248,000 tons still being produced from benzene (22). Maleic anhydride can be used for the maleate hydratase reaction as maleate in an aqueous medium.

Maleate is not a natural compound; however, it is an intermediate in the pathway for the catabolism of nicotinate (23) and gentisate (24). Several investigators have studied the use of maleate as a carbon source to characterize the metabolism of microorganisms. Perry and Edward observed the growth of *Pseudomonas fluorescens* ATCC11150 on a medium containing maleate as the sole carbon and energy source, and deduced the metabolic pathway from the cultivation profile (25). They discussed the possibility of economical L-malate production from maleate using maleate *cis–trans* isomerase (EC 5.2.1.1.) and fumarase from a microorganism that grows on maleate. Maleate *cis–trans* isomerase was first identified by Behrman and Stanier in *Pseudomonas* sp. (26). Many researchers have investigated the characteristics of this enzyme. Kimura et al. demonstrated that *Alcaligenes* sp. T501 showed high activity as to the enzymatic conversion of maleate to L-malate *via* fumarate (27). L-Aspartic acid is widely used as a food additive, as a raw material for aspartame, and in medicines. In 1973, the continuous production of L-aspartic acid from fumarate by means of an immobilized microbial cell system was commercialized by Tanabe Seiyaku. L-Aspartic acid was also enzymatically synthesized from maleate and ammonia, using maleate isomerase and aspartase. Takamura et al. showed the enzymatic production of L-aspartic acid, for which maleate was employed instead of fumaric acid as the starting material, with a cell suspension of *Alcaligenes faecalis* 5-24 through the actions of both maleate isomerase and aspartase (28).

Screening of D-Malate Producing Strains

In general, many maleate-assimilating bacteria are known to metabolize maleate to L-malate *via* fumarate (23). Rahatekar et al. (29) have already reported a *Pseudomonas* strain that exhibits both maleate hydratase and maleate *cis–trans* isomerase activity (Fig. 1). Malease is found in a wide variety of organisms, including bacteria, yeasts, fungi (30), plants (1), and mammals (31). Therefore, microorganisms that have a powerful D-malate producing enzyme without a maleate isomerase were necessary to produce D-malate. From among 440 strains of maleate-assimilating bacteria isolated from soil, 82 bacteria that could produce more than 5.0 g/L malate in 24 h were obtained. The ratio of D-malate to the total amount of malate produced by these 82 microorganisms was measured. Fig. 2 shows that only one strain produced L-malate stereospecifically, and the products of 32 strains were mixtures of D- and L-malate. Fifty strains that specifically produced D-malate were obtained and the optical purity of their products was above 90%. Among these strains, *Arthrobacter* sp. MCI2612 showed high maleate hydatase activity, and the malate produced was 100% D-type (7). This result suggests that only maleate hydratase, and not maleate isomerase, is responsible for the maleate metabolism in this strain.

Cultivation and Reaction Conditions for the Production of D-Malate

Asano et al. (7) studied the conditions for the cultivation of *Arthrobacter* sp. MCI2612, and D-malate production by the maleate hydratase with resting cells. As shown in Table 1, the D-malate producing activity was found not to be inducible, but the addition of maleate, D-malate, and citraconate to the medium increased the activity. Maleate was a suitable inducer for the formation of D-malate producing activity. The addition of organic solvents and surfactants for D-malate production was effective. Triton X-100, Triton N-101, and tergitol NPX remarkably increased the activity; the initial reaction rate was about 10-fold that without detergents, and Triton N-101 (1% of the reaction mixture) gave the optimum production of D-malate.

Production of D-Malate by Intact Cells

Since maleate concentrations above 5% inhibit D-malate production, the reaction was carried out at 36 °C for 20 h, keeping the concentration of maleate below 5% through successive feeding of maleate. The time course for D-malate synthesis is shown in Figure 3; 87 g/L of D-malate was produced after 20 h reaction. The molar yield of D-malate was 72%.

Within the last few years, there have been several reports on D-malate production by microorganisms from maleate. Table 2 summarizes the profiles of the microbial production of D-malate (4–7,32,33). The maleate hydratase of all strains except one is inducible. These enzymes are induced by structurally different compounds, such as maleate, D-malate, and 3-hydroxybenzoate, which are the substrate, product, and precursor, respectively, of maleate. The enzyme in *Ustilago sphaerogena* S402 is constitutive (5). van der Welf et al. produced a high concentration of

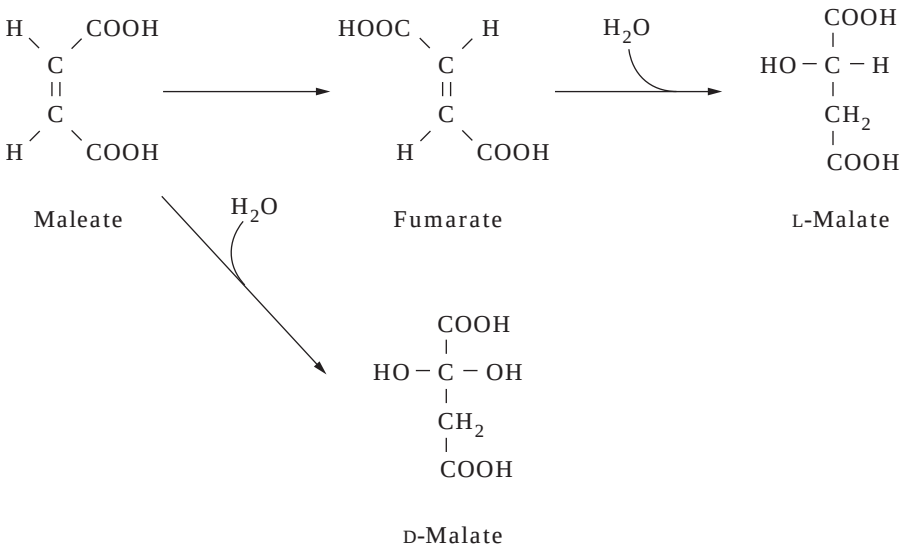


Figure 1. Enzymatic reaction scheme of maleate.

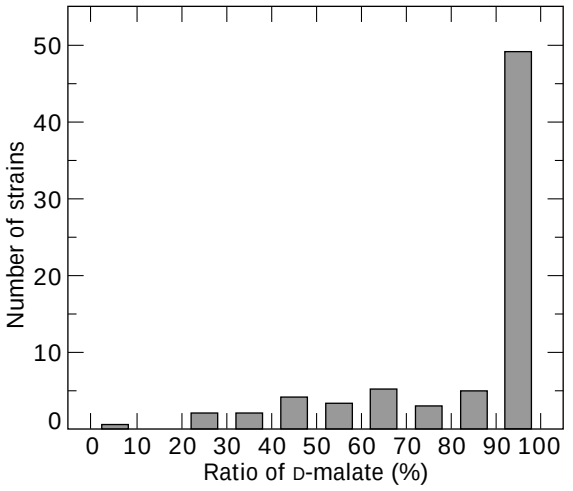


Figure 2. Ratio of D-malate in total malate produced by various strains.

malate in a crystal-liquid two-phase system with Ca-maleate as the substrate (33). The advantages of this method are that no substrate or product inhibition occurs during the reaction and that the reaction equilibrium proceeds towards the product (34). The high enantioexcess of D-malate produced suggest that these microorganisms have no maleate isomerase or that it is not induced during cultivation. The reaction temperature was rather low compared with that for L-malate production from fumarate (50–55 °C) (35). Screening for heat-resistant maleate hydratases will lead to improvement of productivity and prevent contamination that may lead to a decrease in the optical purity.

Comparison of Manufacturing Processes for Malate

DL-Malic acid is used in the food industry, for example as an acidulant in juice and in processed foodstuffs. About 25,000 tons of DL-malate is produced per year through the chemical hydration of maleic anhydride. L-Malate is used for blood infusion and as a resolving agent in organic synthesis. L-Malate can be produced through natural resources, fermentation, and enzymatic processes. An enzymatic process using fumarase is the most efficient. As *Brevibacterium flavum* was found to show high enzyme activity in the case of cells immobilized with κ-carrageenan, Tanabe Seiyaku produces L-malate from fumarate using an immobilized biocatalyst on an industrial scale (35). Ta-

Table 1. Effect of Inducers on D-Malate Producing Activity on *Arthrobacter* sp.

Inducer	(g/L)	Cultivation: Final pH	Cultivation: Growth (OD[a])	Initial rate of D-malate production (mg/mL/h)
None	—	5.8	18.3	1.09
Maleic acid	1	5.7	18.3	1.73
	2	5.9	18.4	2.19
	3	5.9	18.9	2.39
	4	6.0	19.6	2.22
	5	5.8	19.3	2.26
Fumaric acid	2	6.6	19.4	0.31
D-Malic acid	2	6.5	22.0	1.60
L-Malic acid	2	6.5	20.2	0.66
Citraconic acid	2	6.6	20.0	1.47
Mesaconic acid	2	6.5	20.0	0.59
DL-*cis*-epoxysuccinic acid	2	5.8	19.0	0.85
DL-*trans*-epoxysuccinic acid	2	6.0	19.1	0.31
L(+)-Tartaric acid	2	7.0	19.7	0.27
Nicotinic acid	1	6.7	19.5	0.71
2,5-Dihydroxybenzoic acid	2	6.7	20.7	0.75
Crotonic acid	0.5	6.6	9.58	0.12
L-Ascorbic acid	2	6.1	18.3	0.18
o-Phthalic acid	2	5.9	19.6	0.58

Note: The composition of culture medium was 10 g sorbitol, 2 g NH_4NO_3 3 g KH_2PO_4, 10.5 g K_2HPO_4, 250 mg $MgSO_4 \cdot 7H_2O$, 10 mg $FeSO_4 \cdot 7H_2O$, 10 mg $MnSO_4 \cdot 4–6H_2O$, 5 g yeast extract, 2 g L-proline, and inducers in 1,000 mL of tap water.

[a]OD660, optical density at 660 nm.

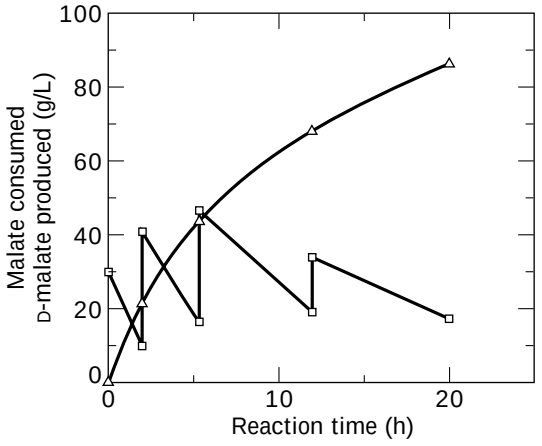


Figure 3. Time course of conversion of maleate to D-malate. Cells were grown at 28 °C in the culture medium, which was the same as in Table 1 except that it contained 3 g maleate/L. Triton N-101 (1%) was added to the reaction mixture. The reaction was carried out at 37 °C. Symbols: ○, D-malate produced; ■, residual maleate. After 2, 6, and 12 h of incubation, maleate 2Na was fed to the reaction mixture.

ble 3 compares the D-malate, L-malate, and DL-malate manufacturing processes. These two enzymatic reactions proceed under mild conditions of pressure, temperature, and pH values.

PURIFICATION AND CHARACTERIZATION OF MALEATE HYDRATASE

The properties of maleate hydratases from a wide variety of organisms have been reported. England et al. (31) demonstrated that an enzyme mixture from rabbit kidney catalyzed the hydration of maleate to D-malate by *trans* addition to the double bond. Britten et al. reported that partially purified maleate hydratase from rabbit kidney had an absolute requirement for chloride ions (36). A crude extract from a pseudomonad produced D-malate from maleate; the optical purity of D-malate was 99.7% (24). Dreyer (37) reported that maleate hydratase from rabbit kidney was a monomeric protein with a molecular mass of 68,000 and required activation by ferrous ion and Na_2S.

Ueda et al. (38) purified the maleate hydratase of *Arthrobacter* sp. MCI2612. It has a molecular mass of about 90,000 and consists of two subunits with molecular masses of 58,000 and 28,000. The purified enzyme was incubated with sulfhydryl compounds and Fe^{2+}. L-Cysteine was found to be the most effective activator, followed by dithiothreitol, thioglycolate, reduced glutathione, and 2-mercaptoethanol, in the presence of Fe^{2+}. Several hydratases, such as citraconase (39), mesaconase (40), L-tartaric dehydratase (41), aconitase (42), and fumarase (43), are known to require ferrous ion and sulfhydryl compounds for their activities. Fumarase is a well-characterized enzyme among these hydratases. A fumarase in aerobically grown *Escherichia coli* required activation by incubation with

Table 2. Comparison of the D-Malate Production by Various Microorganisms

Strain	*Arthrobacter* sp. MCI-2612 (7)	*Pseudomonas fluorescens* IFO3081 (6)	*Clostridium formicoaceticum* DSM94 (32)	*Ustilago sphaerogena* S402 (5)	*Pseudomonas pseudoalcaligenes* NCIMB9877 (4,33)
Inducer	Maleate, D-Malate, Citraconate	Citraconate	NE	None (constitutive)	3-Hydroxybenzoate, Gentisate
pH	7.0	7.0	7.5	NE	7.5
Temperature (°C)	36	26	37	30	30
Substrate inhibition	Yes	NE	NE	None (up to 100 g/l)	Yes
Product inhibition	Yes	NE	NE	NE	Yes
Treatment of cells	Triton N-101	None	None	Triton X-100	Triton X-100
Productivity	87 g/L/20 h	6.94 g/L/24 h	1.87 g/L/18 h	40.8 g/L/48 h	215 g/L/48 h
Reaction volume culture volume	1	3	NE	2.5–3.3	NE
Optical purity (% c.c.)	100	99	>99	98.9	99.96
Reaction yield (mole %)	72	NE	>98	49.6	99.4

NE, not examined.

Table 3. Comparison of Manufacturing Processes of Malate

	L-Malate	DL-Malate	D-Malate
Raw material	Fumarate	Maleic anhydride	Maleate
Catalyst	Fumarase	None	Maleate hydratase
Reaction temperature (°C)	30–50	160–200	30
Reaction pressure (atm)	Atmospheric pressure	5–20	Atmospheric pressure
pH	6–8	2–4	7–8
Solvent	Water	Water	Water

Table 4. Kinetic Parameters of Substrates of Maleate Hydratase of *Arthrobacter* sp.

Substrate	K_m (mM)	V_{max} (μmol/mg/min)
Maleate	3.85	16.5
2-Chloromaleate	2.74	25.8
2-Bromomaleate	3.27	8.26
Citraconate	4.55	13.2

Note: The initial velocity was measured at 30 °C in a reaction mixture containing 20 mM Tris-HCl buffer (pH 7.5), an appropriate amount of purified enzyme, and various concentrations of substrate.

Fe^{2+} and a thiol, and was suggested to have an Fe-S cluster (44). Thus, the maleate hydratase from *Arthrobacter* sp. is suggested to be a member of the hydratases containing an Fe-S cluster represented by fumarase, because the feature of the activation is similar to that reported for fumarase. The reaction product of the purified maleate hydratase is D-malate and the optical purity is 100% D-type.

The substrate specificity and kinetic properties of this reaction were also examined. As shown in Table 4, chloromaleate served as a better substrate than the natural substrate, maleate. The rates of hydration of these substrates followed: chloromaleate > maleate > citraconate > bromomaleate (Table 4). The reaction products from citraconate, chloromaleate, and bromomaleate are (*R*)-(−)-

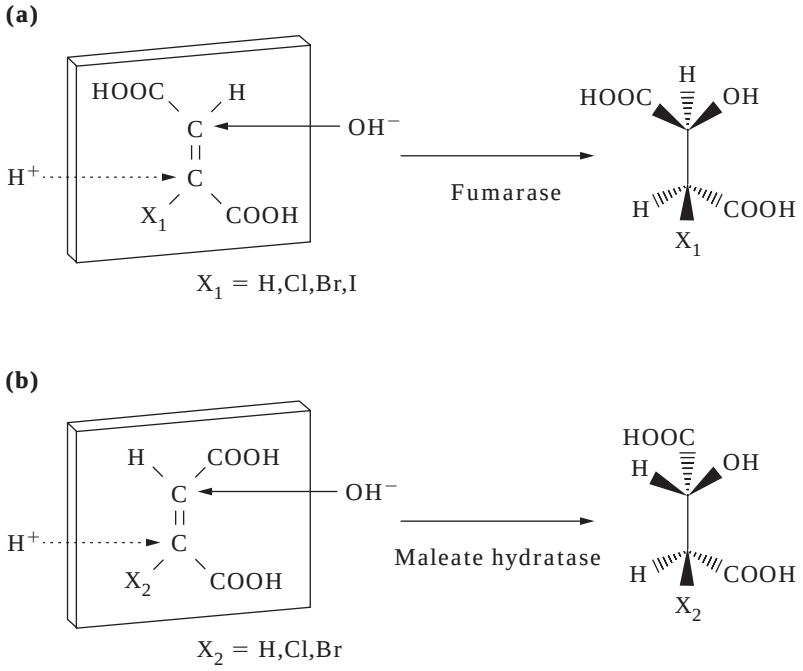


Figure 4. Stereochemistry of the reaction catalyzed by fumarase (**a**) compared to that catalyzed by maleate hydratase (**b**).

Table 5. Some Properties of Animal and Bacterial Maleate Hydratases

Origin	Rabbit kidney (37)	*Pseudomonas pseudoalcaligenes* NCIMB9867 (47,48)	*Arthrobacter* sp. MCI-2612 (38,49)
M.W. (kDa)	68 (Sephacryl S-300)	24,000, 57,000 (SDS) 89,000 (Gel filtration)	58,000, 28,000 (SDS)
Optimum pH	NT	8	8.5
Optimum temperature (°C)	NT	45	45
Activating agents	Fe^{2+}, Na_2S	None	Fe^{2+}, L-Cysteine
Specific activity (U/mg protein)	317	108	6.55
Another substrate	None	Citraconate 2,3-Dimethylmaleate	Citraconate 2-Chloromaleate 2-Bromomaleate
K_m (mM)	NT	0.35 for maleate 0.20 for citraconate	3.85 for Maleate 2.74 for 2-Chloromaleate 3.27 for 2-Bromomaleate 4.55 for Citraconate

NT, not tested.

citramalate, (−)-*erythro*((2*S*, 3*S*)-3-)-chloromalate and (−)-*erythro*((2*S*, 3*S*)-3-)-bromomalate, respectively. The enzyme did not act on dimethyl maleate, acetylenedicarboxylate, fumarate, mesaconate, *cis*-epoxysuccinate, and *trans*-epoxysuccinate. Consequently, it was possible to determine that the bacterial enzyme-catalyzed hydration reaction also involved *trans* addition of the hydrogen and hydroxyl ion to maleate (Fig. 4a).

Fumarase (fumarate hydratase, EC 4.2.1.2), which catalyzes the reversible hydration of fumarate to L-malate, has a broad substrate specificity (45). Many dicarboxylic acids, such as acetylenedicarboxylate, epoxysuccinate, and tartarate, as well as α-monosubstituted derivatives of fumarate, such as chlorofumarate, bromofumarate, and iodofumarate (Fig. 4b), have been reported to serve as substrates for fumarase from swine heart muscle (46). Although acetylenedicarboxylate, epoxysuccinate, and tartarate did not serve as substrates for the maleate hydratase from *Arthrobacter* sp., these dicarboxylic acids were found to be competitive inhibitors.

Table 5 summarizes the characteristics of the maleate hydratases purified from microorganisms and rabbit kidney. Although the maleate hydratase derived from rabbit kidney exhibited absolute specificity for maleate, the hydration of chloromaleate, bromomaleate, and citraconate, in addition to that of maleate, was observed for the maleate hydratase from *Arthrobacter* sp. (38,49). The purified *P. pseudoalcaligenes* enzyme also catalyzes the hydration of citraconate and 2,3-dimethylmaleate (47,48). The enzyme from rabbit kidney requires Fe^{2+} and a thiol as reducing agents for its activity. In contrast, the maleate hydratase from *P. pseudoalcaligenes* does not require metals or other cofactors. Thus, it appears that maleate hydratases can be classified into two classes of carboxylic-acid hydratases, either containing or not containing on iron-sulfur cluster, of which fumarase is probably the most thoroughly investigated (50).

BIBLIOGRAPHY

1. W. Sacks and C.O. Jensen, *J. Biol. Chem.* **192**, 231–236 (1951).
2. J.V. Taggart, S. Angielski, and H. Morell, *Biochim. Biophys. Acta.* **58**, 144–156 (1962).
3. D.J. Hopper, P.J. Chapman, and S. Dagley, *Biochem. J.* **110**, 798–800 (1968).
4. M.J. van der Werf, W.J.J. van den Tweel, and S. Hartmans, *Appl. Environ. Microbiol.* **58**, 2854–2860 (1992).
5. T. Nakajima, S. Manzen, T. Shigeno, and T. Nakahara, *Biosci. Biotechnol. Biochem.* **57**, 490–491 (1993).
6. K. Nakayama and M. Ushijima, *Biotechnol. Lett.* **15**, 271–276 (1993).
7. Y. Asano, M. Ueda, and H. Yamada, *Appl. Environ. Microbiol.* **59**, 1110–1113 (1993).
8. Japanese Pat. 97-275991 (October 28, 1997), I. Oonishi, T. Tanba, and K. Yokozeki (to Ajinomoto Co., Japan).
9. Japanese Pat. 98-52289 (February 24, 1998), S. Yasuda, S. Mukoyama, and K. Sakano (to Nippon Shokubai Kagaku Kogyo Co., Japan).
10. Y. Gao and K.B. Sharpless, *J. Am. Chem. Soc.* **110**, 7538–7539 (1988).
11. S. Henrot, M. Larcheveque, and Y. Petit, *Synth. Commun.* **16**, 183–190 (1986).
12. A.S.C. Chan and J.P. Coleman, *J. Chem. Soc. Chem. Commun.* **104**, 166–168 (1982).
13. H. Wynberg and E.G.J. Staring, *J. Am. Chem. Soc.* **104**, 166–168 (1982).
14. Japanese Pat. 57056439 (April 5, 1982), T. Katsura and H. Yamatika (to Sumitomo Chemical Co., Japan).
15. K. Saigo, M. Kai, N. Yonezawa, and M. Hasegawa, *Synthesis* **2**, 214–216 (1985).
16. T. Shiraiwa, Y. Sado, M. Inoue, K. Sakamoto, H. Miyazaki, and H. Kurokawa, *Bull. Chem. Soc. Jpn.* **61**, 899–903 (1988).
17. U.S. Pat. 4,912,042 (March 27, 1990), K. E. Laumen (to Eastman Kodak Co., USA).
18. M. Miyama and K. Nakayama, *Biotechnol. Lett.* **15**, 23–28 (1993).
19. Japanese Pat. 91-136331 (June 7, 1991), M. Terasawa, S. Nara, H. Yamagata, and H. Yugawa (to Mitsubishi Chemical Co., Japan).
20. E. Santaniello, P. Ferraboschi, P. Grisenti, F. Aragozzini, and E. Maconi, *J. Chem. Soc. Perkin, Trans. 1* **3**, 601–605 (1991).
21. K. Ushio, K. Nakagawa, K. Nakagawa, and K. Watanabe, *Biotechnol. Lett.* **14**, 795–800 (1992).
22. F. Cavani and F. Trifiro, *Chem. Tech.* **24**, 18–25 (1994).
23. E.J. Behrman and R.Y. Stanier, *J. Biol. Chem.* **228**, 923–945 (1957).
24. D.J. Hopper, P.J. Chapman, and S. Dagley, *Biochem. J.* **110**, 798–899 (1968).
25. J.T. Perry and V.H. Edward, *Appl. Microbiol.* **20**, 710–714 (1970).
26. E.J. Behrman and R.Y. Stanier, *J. Biol. Chem.* **228**, 923 (1957).
27. T. Kimura, Y. Kawabata, and E. Sato, *Agric. Biol. Chem.* **50**, 89–94 (1986).
28. Y. Takamura, I. Kitamura, M. Iida, K. Kono, and A. Ozaki, *Agric. Biol. Chem.* **30**, 338–344 (1966).
29. H.I. Rahatekar, F.S. Maskati, S.S. Subramanian, and M.R.R. Rao, *Indian J. Biochem.* **5**, 143–144 (1968).
30. U.S. Pat. 5,270,190 (December 14, 1993), K. Nakayama and Y. Kobayashi (to Bior Inc., Tokyo, Japan).
31. S. England, J.S. Britten, and I. Listowsky, *J. Biol. Chem.* **242**, 2255–2259 (1967).
32. E. Richard and S. Helmut, *Tetrahedron* **50**, 13641–13654 (1994).
33. M.J. van der Werf, S. Hartmans, and W.J.J. van den Tweel, *Enzyme Microb. Technol.* **17**, 430–436 (1995).
34. K. Kitahara, S. Fukui, and M. Misawa, *Nihon Nougei Kagaku Kaishi* **33**, 528–531 (1959).
35. I. Chibata, T. Tosa, and T. Shibatani, A.N. Collins, G.N. Sheldrake, and J. Crosby eds., in *Chirality in Industry*, Wiley, Chichester, 1992, pp. 351–370.
36. J.S. Britten, H. Morell, and J.V. Taggart, *Biochim. Biophys. Acta.* **185**, 220–227 (1969).
37. J.L. Dreyer, *Eur. J. Biochem.* **150**, 145–154 (1985).
38. M. Ueda, Y. Asano, and H. Yamada, *Biosci. Biotechnol. Biochem.* **57**, 1545–1548 (1993).
39. S.S. Subramanian and M.R.R. Rao, *J. Biol. Chem.* **243**, 2367–2372 (1968).
40. S. Suzuki, T. Osumi, and H. Katsumi, *J. Biochem.* **81**, 1917–1925 (1977).

41. R.E. Hurlbert and W.B. Jakoby, *J. Biol. Chem.* **240**, 2772–2777 (1965).
42. M.C. Kennedy, M.H. Emptage, J.L. Dreyer, and H. Beinert, *J. Biol. Chem.* **258**, 11098–11105 (1983).
43. S.K. Reaney, S.J. Bungard, and J.R. Guest, *J. Gen. Microbiol.* **139**, 403–416 (1993).
44. N. Yumoto and M. Tokushige, *Biochem. Biophys. Res. Commun.* **153**, 1236–1243 (1988).
45. P.D. Boyer, *The Enzyme*, 3rd ed., Academic Press, New York, 1971, pp. 539–568.
46. J.W. Teipel, G.M. Hass, and R.L. Hill, *J. Biol. Chem.* **243**, 5684–5694 (1968).
47. M.J. van der Werf, W.J.J. van den Tweel, and S. Hartmans, *Appl. Environ. Microbiol.* **59**, 2823–2829 (1993).
48. M.J. van der Werf, W.J.J. van den Tweel, and S. Hartmans, *Eur. J. Biochem.* **217**, 1011–1017 (1993).
49. M. Ueda, H. Yamada, and Y. Asano, *Appl. Microbiol. Biotechnol.* **41**, 215–218 (1994).
50. Y. Ueda, N. Yumoto, M. Tokushige, K. Fukui, and H. Ohya-Nishiguchi, *J. Biochem.* **109**, 728–733 (1991).

See also ENERGY METABOLISM, MICROBIAL AND ANIMAL CELLS; METABOLITES, PRIMARY AND SECONDARY.

MALIC ACID, PRODUCTION BY FUMARASE

TADASHI SATO
TETSUYA TOSA
Tanabe Seiyaku Co. Ltd.
Osaka, Japan

KEY WORDS

Fumarase
Immobilized *Brevibacterium ammoniagenes*
Immobilized *Brevibacterium flavum*
Immobilized cells
L-Malic acid

OUTLINE

L-Malic acid plays a very important metabolic role in living cells. It has been used to treat hyperammonemia and liver disfunction, and it is component of amino acid infusions. Also, if it becomes possible to supply large amounts of L-malic acid at low cost, L-malic acid—the natural form—is expected to be used instead of DL-malic acid—the racemic form—in the food industry as a food additive.

Existing preparation methods for L-malic acid include isolation from natural fruit juice, the optical resolution of synthetic DL-form, fermentation, and others. However, for industrial purposes, the acid has been produced from fumaric acid by the action of fumarase as follows.

$$\underset{\text{Fumaric acid}}{HOOCCH{=}CHCOOH} + H_2O \overset{\text{Fumarase}}{\rightleftarrows} \underset{\text{L-Malic acid}}{HOOCCH_2CH(OH)COOH}$$

Many kinds of microbial cells having high fumarase activity have been successfully immobilized. The immobilized cells having high enzyme activity were obtained by the polyacrylamide gel method. Moreover, the immobilized cells were stable, and a column packed with them could be used for continuous enzyme reaction for long periods.

The industrial production of L-malic acid from fumaric acid has been performed by an enzymatic batch process using the broth of *Lactobacillus brevis* (1), but a continuous reaction using immobilized fumarase of immobilized cells with fumarase activity should be more economical.

Marconi et al. (2) reported that fumarase isolated from microbial cells can be efficiently immobilized on cellulose triacetate, making it possible to develop an economically attractive method for producing L-malic acid. However, no industrial process using immobilized fumarase has been brought into operation.

Several microorganisms having high fumarase activity were immobilized by the polyacrylamide gel method, and their activities were compared (3). *Brevibacterium ammoniagenes* was found to be the most active before and after immobilization.

In 1974, Tanabe Seiyaku Co., Ltd., introduced a process of continuous production of L-malic acid from fumaric acid using immobilized *B. ammoniagenes*. Since then, the process has been modified in several major and minor respects to create the current system, which is 20 to 25 times more efficient than the original.

Yamamoto et al. (3) investigated the industrial application of immobilized microbial cells with fumarase activity and developed a method using *B. ammoniagenes* immobilized with polyacrylamide gel.

When the immobilized cells were used for the production of L-malic acid from fumaric acid, succinic acid was formed as a by-product, and the separation of succinic acid from L-malic acid was technically very difficult. Accordingly, for the industrial production of pure L-malic acid in high yield, it was necessary to prevent the formation of succinic acid during the enzyme reaction. Thus, various treatments of intact or immobilized cells were performed. Heat treatment, autolysis, and freezing-thawing were not effective for the prevention of succinic acid formation. However, treatment with substrate solution containing deoxycholic acid or bile acid, which was commonly used as a solubilizer of membrane-bound enzymes, was found to be very effective for the suppression of succinic acid formation, as shown in Table 1. Further, this treatment also markedly enhanced the formation of L-malic acid. This effect was attributed to an increase of membrane perme-

Table 1. Effect of Detergent Treatment on Formation of L-Malic Acid and Succinic Acid

Treatment		Formation of L-malic acid	Formation of succinic acid
Detergent	Concentration (%)	(μmol/h/mL gel)	(mol% L-malic acid)
No addition		75	2.5–5.0
CPC	0.02	340	2.5–5.0
CPC	0.16	225	1.0–2.5
SLS	0.02	445	1.0–2.5
SL-10	0.02	90	2.5–5.0
Triton X-100	0.20	395	>5.0
Bile acid	0.20	425	<0.2
Bile extract	0.20	485	<0.2
Deoxycholic acid	0.20	480	<0.2

Note: CPC = cetylpyridium chloride; SLS = sodium lauryl sulfate.

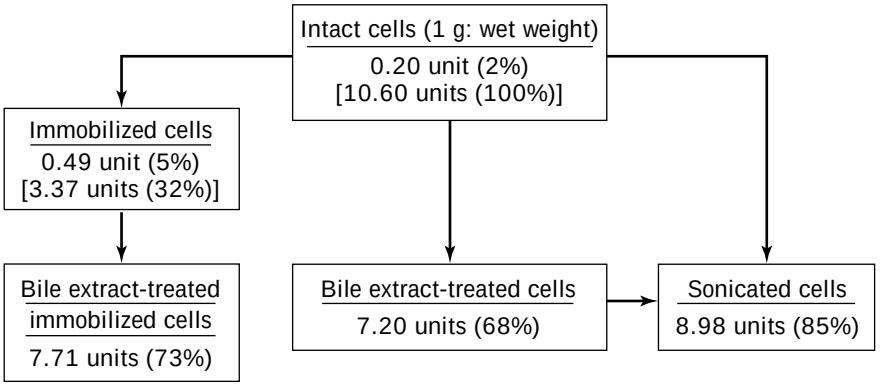


Figure 1. Comparison of the fumarase activities of various enzyme preparations per unit of intact cells. One unit is defined as activity producing 1 mmol of L-malic acid at 37 °C in 1 h. Square brackets show activity in the presence of 0.02% cetylpyridium chloride; percentages are values relative to the activity (100%) of intact cells in the presence of cetylpyridium chloride.

Table 2. Comparison of Productivities of *Brevibacterium ammoniagenes* and *Brevibacterium flavum* Immobilized with Polyacrylamide and with Carrageenan for Production of L-Malic Acid

	B. ammoniagenes			*B. flavum*		
Immobilization method	Activity (unit/g cells)	Half-life at 37 °C (day)	Relative productivity	Activity (unit/g cells)	Half-life at 37 °C (day)	Relative productivity
Polyacrylamide	5,800 (60%)	53	100	6,680 (34%)	94	204
Carrageenan	5,800 (60%)	75	142	9,920 (51%)	160	516

Note: Activity after treatment with bile extract. Productivity = $\int_0^t E_0 \exp(-\kappa d \times t)\mathrm{d}t$, where E_0 = initial activity, κd = decay constant, and t = operational period.

ability to the substrate or product resulting from bile acid treatment. Bile extract is suitable for industrial purposes because it is commercially available at a low price. Various conditions for treatment with bile extract were studied, and the most effective procedure was to stand the immobilized cells in 1 M sodium fumarate (pH 7.5) containing 0.3% bile extract at 37 °C for 20 h. The yield of fumarase activity of immobilized cells obtained in this way is shown in Figure 1.

Enzymic properties of the immobilized cells treated with bile extract were studied and compared with those of the native enzyme. The optimum temperature and optimum pH of the enzyme reaction were almost same as those of the native enzyme, namely 60 °C and pH 7.0 to 7.5, respectively. The equilibrium point of the enzyme reaction was scarcely affected by immobilization or by the reaction pH, but tended to shift toward L-malic acid formation with decreasing reaction temperature.

Conditions for the continuous enzyme reaction using a column packed with the immobilized cells were investigated. It was found that when 1 M sodium fumarate (pH 7.5) was passed through the column at 37 °C at a space velocity of below 0.23 (residence time in the column, more than 4.3 h), the reaction reached equilibrium with about 80% conversion of fumaric acid to L-malic acid. Fumaric acid was removed from the effluent by acidification, and L-malic acid was separated in about 70% yield from fumaric acid.

For industrialization of this system, chemical engineering studies were carried out, and a continuous reaction sys-

tem using immobilized cells was designed. L-Malic acid of high purity has been industrially produced at Tanabe Seiyaku by this process since 1974.

To improve L-malic acid productivity, the authors investigated immobilization methods that could be carried out under mild conditions without denaturation of fumarase, and they also screened again to find microbial cells with a higher fumarase activity than *B. ammoniagenes*.

Among many synthetic and natural polymers tested as matrixes for entrapping microbial cells, κ-carrageenan was found to be best (4).

Brevibacterium flavum was found to have higher enzyme activity after immobilization with carrageenan as compared with the former *B. ammoniagenes*. For industrial purposes, the stability was also studied, and the productivity was compared, as shown in Table 2. When the productivity of immobilized *B. ammoniagenes* with polyacrylamide was taken as 100, that of immobilized *B. flavum* with carrageenan was 516. Therefore, it is evident that also in this case the carrageenan method is more advantageous than the conventional polyacrylamide method. So, we changed the polyacrylamide method to the carrageenan method in 1977. This new method gives us satisfactory results from an industrial point of view.

In other studies, *Thermus rufens nov* sp., a thermophilic bacterium, was immobilized ionically on Duolite for production of L-malic acid from fumaric acid (5).

L-Malic acid is one of the members of tricarboxylic acid cycle. It is easy to be metabolized in a living body and is accompanied by generation of ATP, which is the energy source of a living body.

L-Malic acid is also converted to L-aspartic acid via oxalacetic acid. In this process, ammonia is metabolized with coupling to the L-glutamic acid–α-ketoglutaric acid system. Further, L-aspartic acid condenses with L-citrulline to form arginosuccinate and enters into the urea cycle. Ammonia is also incorporated when L-citrulline is formed from L-ornithine.

L-Malic acid is effective in detoxication of ammonia by introducing ammonia into the urea cycle. Its detoxicative action is cumulative in combination with L-arginine (6–8).

In addition to the role of L-malic acid mentioned in this article, it has been reported that L-malic acid is effective for lowering blood cholesterol (9), preventing atherosclerosis (10), and lowering blood pressure (11). The diuretic effect of L-malic acid was also reported (12).

BIBLIOGRAPHY

1. K. Kitahara, S. Fukui, and M. Misawa, *J. Gen. Appl. Microbiol.* **6**, 108–113 (1960).
2. W. Marconi, F. Morisi, and R. Mosti, *Agric. Biol. Chem.* **39**, 1323–1324 (1975).
3. K. Yamamoto, T. Tosa, K. Yamashita, and I. Chibata, *Eur. J. Appl. Microbiol.* **3**, 169–183 (1976).
4. T. Tosa, T. Sato, T. Mori, K. Yamamoto, I. Takata, Y. Nishida, and I. Chibata, *Biotechnol. Bioeng.* **21**, 1697–1709 (1979).
5. Y. Ado, T. Kawamoto, I. Masunaga, T. Takayama, S. Takasawa, and K. Kimura, *Enzyme Eng.* **6**, 303–304 (1982).
6. R. Normann, *Dtsch. Med. Wschr.* **94**, 230–233 (1969).
7. I. Szam, J. Szenter, I. Hagedues-Wein, and A. Vass, *Intern. J. Clin. Pharmacol. Ther. Toxicol.* **6**, 260–265 (1972).
8. A. Tissort-Favre and R. Brette, *Therapie* **25**, 629–631 (1970).
9. D.A. Kaidin, *Patol. Fiziol. Eksp. Terapiya.* **12**, 72–74 (1968).
10. D.A. Kaidin, *Byul. Esksp. Biol. Med.* **66**, 29–31 (1969).
11. R. Nose, *Shikoku Igakukaishi* **22**, 525–529 (1966).
12. J.B. Brueckener, *Anaesthesist* **19**, 219–223 (1970).

MAMMALIAN CELL BIOREACTORS

MARIA J. GUARDIA
WEI-SHOU HU
University of Minnesota
Minneapolis, Minnesota

KEY WORDS

Airlift
Bioreactors
Cell culture
Hollow fiber
Immobilized cells
Mammalian cells
Microcarriers
Perfusion
Roller bottle
Stirred-tank

OUTLINE

INTRODUCTION

Before the advent of recombinant DNA technology, industrial mammalian cell culture and animal cell culture were generally used primarily in the production of viral vaccines, using primary cells and normal diploid cell strains, and biologics, using isolated transformed cells. The arrival of rDNA technology and the expression of heterologous proteins in mammalian cells completely transformed cell

culture technology. Except for the production of alternated or inactivated viral vaccines, virtually all mammalian cell culture processes now employ continuous cell lines, most of which have been adapted to grow in suspension, without a surface support. The adoption of the growth of suspension cells as the typical vehicle for the production of complex molecules requiring extensive posttranslational modification also transformed the concept of bioreactor selection.

Much research activity and many commercial efforts were devoted to developing bioreactors for mammalian cell culture in the 1980s. The anticipation was that the potentially detrimental effect of the air sparging of mammalian cells would require a newly designed bioreactor to accommodate the manufacturing needs of the new mammalian cell–based products. Many of these efforts have since declined, and the bioreactors used for industrial cell culture converge on a few basic configurations. This article focuses only on the main types used currently for industrial cell culture. Those interested in the historical aspects of bioreactor development are referred to other monographs and review articles. It should be noted that many of the bioreactors developed in the 1980s, although not commonly used, might find specialized applications in other areas, such as tissue engineering and cellular therapy.

BASIC REACTOR KINETIC

Mammalian cell bioreactors are generally categorized similarly to chemical reactors according to their mixing characteristics. It is instructive to review two ideal reactors: well-mixed stirred tank and plug-flow (tubular) reactor. In an ideal well-mixed bioreactor, the mixing is assumed to be intense enough that the fluid is homogeneous through the reactor. The mathematical description of the ideal continuous flow stirred tank reactor is described by the following first-order differential equation:

$$\frac{d(VC_A)}{dt} = F_i C_{Ai} - F_o C_{Ao} + r_A V \qquad (1)$$

where V is the culture volume in the bioreactor, C_A is the concentration of nutrient or product A, t is time, F is the flow rate, and r_A is the volumetric consumption rate of nutrient or production rate of product A.

In an ideal stirred tank reactor, there is no flow bypass and no shunt of substrate from inlet to outlet and no dead zones or clumps of undissolved solid substrate floating around. The addition of a substrate is through feeding instantaneously distributed throughout the entire reactor, and when gas sparging is employed the agitator provides intimately mixed gas and liquid.

The basic model for the tubular reactor (such as hollow fiber and ceramic systems to be described later) is that the liquid phase moves as a plug flow, meaning that there is no variation of axial velocity over the cross section. The mass balance for component A in a volume element $S\partial z$ that described an ideal plug-flow reactor is the following:

$$\frac{\partial C_A}{\partial t} = -v_z \frac{\partial C_A}{\partial z} + r_A \qquad (2)$$

where v_z is the linear velocity in the z direction and S is the cross-sectional area. At any given moment of time or at steady state (i.e., cell concentration and cellular activities are not changing with time), the equation becomes

$$\frac{\partial c_A}{\partial z} = \frac{r_A \cdot S}{F} \qquad (3)$$

which describes changes of concentration of A along the direction of fluid flow. It is clear that the nutrient concentration will decrease from the inlet to the distal end of the reactor, whereas the metabolite concentration will increase. The length of the reactor is limited because eventually nutrient depletion or metabolite accumulation inhibits growth and metabolism.

These ideal cases of completely mixed tanks or plug-flow tubular reactors are situations that can be approximated in small-scale laboratory conditions. The conditions in larger-scale process reactors deviate significantly from these ideal conditions.

In a well-mixed bioreactor, there are no concentration gradients in either the gas or the liquid phase. In other words, none of the chemical species or cells is segregated in the reactor. The other extreme of mixing is total segregation, where there is no interaction between different volume elements in the bioreactor. An ideal plug-flow reactor is assumed to be under conditions of total segregation. Most bioreactor systems have a mixing pattern between the two extremes and are under partially segregated conditions.

In general, laboratory and small pilot plant bioreactors are used for process development and optimization. The fluid-mixing characteristics are rather sensitive to the scale of the reactor. Furthermore, plug-flow bioreactors are intrinsically more difficult to scale up than mixing vessels, because the concentration gradient of essential nutrients, oxygen in particular, inevitably becomes limiting in the downstream region of the reactor. In considering the selection of bioreactors for mammalian cell cultures, the mixing characteristics and their relationship to scale-up have to be kept in mind.

OPERATION MODES

Mammalian cell cultures are operated in batch, fed-batch, and continuous modes. Batch culture is used most frequently. All nutrients are added at the beginning of a production cycle, except for oxygen, which is provided continuously during the culture through surface aeration, membrane creation, or direct gas sparging. Using a conventional medium, the typical densities achieved in batch culture are rather low, about $2–4 \times 10^6$ cells/mL. In many cases, batch operation tends to be run as extended batch processes. In this operation mode, spent medium with or without cells may be harvested periodically and replaced by an equal quantity of fresh medium. It allows a production process to be extended to a couple of weeks, with in-

creased productivity for the reactor. It is possible to control the inhibitory products accumulated in the culture with this technique.

Another possibility to prolong the batch processes is to use fed-batch operation. In this process, certain nutrients such as amino acids, glucose, and some other micronutrients are added continuously or intermittently to increase the cell concentration and to prolong the product-formation period. In most cases bulk salts (such as NaCl) are not added to the feed to avoid increasing the osmolality. The final product concentration is substantially higher than that in a batch culture.

Continuous chemostat operation, used primarily for laboratory-scale bioreactors, has found limited application in production-scale systems. In this operation, spent medium is continuously replaced with an equal volume of fresh medium and therefore continuously discharging cells. The cell concentration in the bioreactor usually is at a low level similar to that in batch cultures. A major drawback of a simple continuous culture is thus the low cell concentration and the corresponding low throughput of the system. To increase the cell concentration and the reactor throughput, a perfusion culture is more frequently used in industrial continuous operations. In a perfusion culture, a high medium flow rate (one much higher than the cell growth rate, which would wash out cells in a simple continuous culture) is used, but a cell retention device is used to prevent most of the cells from leaving the reactor. The cell concentration in the reactor is up to 10 times higher than that for a simple continuous culture.

When a plug-flow type of bioreactor is used, usually the medium is continuously recirculated through the bioreactor. The recirculating medium carries oxygen to the reactor to sustain cell viability and growth. However, the system is not a continuous reactor in a strict sense. Although the flow of recirculating medium is continuous, the medium itself is replaced periodically for nutrient replenishment and product harvest. It is more similar to repetitive-batch types of bioreactors.

MIXING TANK BIOREACTORS

Cell Support for Mixing Vessels

In the past decade, most of the cell culture processes developed have been for the production of recombinant proteins and antibodies; the majority employ suspension cells. However, despite the general trend of using suspension cells, anchorage-dependent or anchorage-preferred cells are still widely used for the manufacture of some products, especially viral vaccines. Many cell types used in tissue engineering applications are also anchorage dependent. To cultivate these cells in a mixing vessel, some form of anchoring support is necessary. By employing such support, the cells can be cultivated in a relatively homogeneous environment with appropriate environmental control, be infected for virus production, or be induced for product formation.

Microcarriers. The use of microcarriers for cell culture was first demonstrated by Van Wezel (1). The basic concept is to allow cells to attach to the surface of small suspended beads so that conventional stirred tank bioreactors can be used for cell cultivation. To ease suspending these cell-laden microcarriers, their diameter and density are usually in the range of 100–300 μm and 1.02–1.05 g/cm^3, respectively. This diameter range also gives a good growth surface area per reactor volume. Even at a moderate microcarrier concentration, in the range of 8–15% culture volume being occupied by microcarriers, a significantly larger surface area per reactor volume can be achieved than that in roller bottles.

Although smaller microcarriers less than 100 μm in diameter can provide even more surface area, they are not used often. Most anchorage-dependent cells do not develop their normal morphology and in many cases do not multiply well on surfaces with an excessively high curvature. These cells do not grow well on small microcarriers. On the other hand, some cells, especially transformed cells, multiply well even though they are attached to small microcarriers and do not develop a well-spread-out morphology. These cells, after attaching to the small microcarriers, agglomerate to form aggregates and continue to grow to high density. The small microcarriers, usually with a diameter of about 50 μm, serve as nuclei for the initiation of aggregate formation.

The microcarriers can be made of many different materials, including dextran, gelatin (2), polystyrene (3,4), glass (5), and cellulose (6). Not all of these are available commercially. In general, in addition to having a wettable surface, the backbone materials of microcarriers often need to be chemically modified to improve cell attachment. One of the most widely used microcarriers, the dextran-based ones, are derivatized with charged molecules or collagen. Polystyrene microcarriers have also been coated with collagen or other adhesion molecules for better performance.

An advantage for the industrial-scale culture of anchorage-dependent cells is the ease of separating cells from the culture medium. Many microcarrier cultures require medium exchanges during cultivation to remove accumulating lactate, ammonia, and other metabolites and to replenish nutrients. In many cases, the cell-laden microcarriers are simply allowed to settle while the spent medium is withdrawn and replenished. In large-scale operations, continuous or semicontinuous perfusion is used more frequently. Such perfusion can be accomplished by withdrawing the medium through a coarse screen that allows the medium to pass through but retains microcarriers in the reactor. Many cells have been grown on microcarriers, including fibroblast, kidney, epithelial, hepatocyte, neuroblastoma, and endothelial cells from various species. Overall, microcarrier culture is a convenient laboratory and research tool, and it has the advantage of being amenable to large-scale production if a large quantity of product is needed.

Macroporous Microcarriers. A variant of microcarriers is one with large pores of tens of micrometers in its interior. The void space inside allows cells to be cultivated not only on the external surface but also internally. Cells in the interior are less susceptible to mechanical damage caused by

agitation and gas sparging. Of course, cells in the interior of microcarriers are more likely to be subject to oxygen limitation due to the long diffusional distance, especially because most macroporous microcarriers have a larger diameter (500 μm to a couple of millimeters). Macroporous microcarriers are made of different materials, including gelatin, collagen, and plastic. Many cell lines have been successfully grown on macroporous microcarriers including Vero, HepG2, CHO and 293 cells. The final cell concentration achieved tends to be higher than that obtained with an equivalent volumetric concentration of conventional microcarriers. But in some cases the growth kinetics are a bit slower because the penetration of cells into the interior may be slowed or even retarded by restrictive opening of these pores.

Aggregate Culture. Some transformed cells that normally attach to or spread and grow on a surface can form aggregates when cultivated in shaker flasks or in stirred vessels (7). Different methods have been used to induce aggregate formation for cells. Aggregation can be promoted by manipulating the calcium concentration (8) in conjunction with the agitation rate. Aggregate cultures have advantages similar to microcarrier cultures. They can be cultivated in conventional stirred tank reactors with environmental control. They can also be allowed to settle relatively rapidly by stopping agitation, which readily permits medium replenishment.

For many cell types, one limitation for application of the aggregate culture is that the rate of aggregate formation is slow. One way to induce aggregate formation is to add microspheres to the cell suspension to allow for rapid agglomeration to the microspheres (9). If the aggregate diameters become too large, necrotic centers can be formed as a result of transport limitation. The aggregate size may influence the viral infection kinetic and yield for vaccine formation (10). Many cell lines, including BHK, CHO, 293, and ST cells, have been cultivated as aggregates with sizes ranging between 90 and 400 μm, without the formation of necrotic centers (9,11).

Cell Entrapment. Another method of cell cultivation that enables the use of a stirred tank reactor is cell entrapment. This technique entails entrapping cells in a polymeric matrix to form spheres. The spheres are then either coated with another polymeric film to control the crossing of molecules according to their size—commonly referred to as microencapsulation—or cultivated as they are. These beads are suspended in medium to allow cell growth inside. One of the polymers most used for cell entrapment is calcium alginate. Cell entrapment in calcium alginate is accomplished by preparing a cell suspension in sodium alginate and adding it, in a dropwise fashion, into a solution of calcium chloride. Calcium cross-links alginate, instantaneously forming beads that contain entrapped cells. Alginate may be coated with polylysine for increased mechanical and chemical stability, but such treatment decreases the molecular weight cutoff and prohibits large molecular weight proteins in the medium from reaching the cells. To prepare hollow spheres, the alginate gel inside a bead coated with a polylysine is liquefied through treatment with a calcium chelator such as EDTA or citrate.

The diameter of these beads is often on the order of millimeters. The mechanical strength of the gel that constitutes the beads is relatively weak. Large-scale application using such techniques is not easy. On the other hand, the microcapsule provides a means of immunoisolation of transplanted cells or tissues (sometimes referred to as artificial cells) and could be suitable for some tissue engineering applications. Cell entrapment has been applied to hybridomas (12) and small clusters of adherent cells (13). It has also found applications in the cultivation of differentiated cells, such as insulin-secreting cells. For tissue engineering applications using differentiated cells, the enclosing membrane around the cells and the hydrogel must be biocompatible (14). The membrane used in microencapsulation is semipermeable to allow sufficient diffusion and transport of low molecular weight molecules important for cell survival, such as oxygen, nutrients, and metabolites. For optimum transport across the membrane, the microcapsules should have a uniform wall thickness. Ideally, the semipermeable membrane prevents the passage of high molecular weight proteins, such as immunoglobulins, to allow for product to accumulate inside the microcapsule.

Because the cells are inside the capsules, they are protected from hydrodynamic damage. The culture can be stirred at faster rates than conventional culture, which improves nutrient, metabolite, and gas exchange. Entrapped cells may be cultured in stirred tank bioreactors, fixed-bed reactors, or fluidized-bed reactors. An airlift reactor could be used as well, provided that the beads are small and not significantly denser than the medium.

Stirred Tank Bioreactors

The cells respond in different ways when they are exposed to fluid forces. Usually the morphology of the cell and the cell metabolism are affected by fluid stress. The fluid mechanical consequences in the cells on microcarriers depend on the type of cells and the quality of cell attachment. Cell damage from the interaction between microcarriers and turbulent eddies may be severe when the Kolmogorov length scale for the smallest turbulent eddies becomes comparable to the diameter of the microcarriers (15). The microcarrier–microcarrier collisions are also expected to increase in frequency and severity in the presence of eddies of size equal to the average interparticle distance. Suspended cells are generally less prone to hydrodynamic damage than cells on microcarriers. Two principal types of fermentors—stirred tank and airlift bioreactors—are most commonly used for cell cultivation.

Stirred tanks, or conventional fermentors, have been widely used for culturing suspension cells since the 1960s. With the use of microcarriers, whether conventional or macroporous, adherent cells can also be cultured in a stirred tank. For cell entrapment in hydrogel, the use of stirred tanks is limited to laboratory scale, because the preparation of a large quantity of cell-entrapped beads can be a daunting task, and the risk of mechanical damage caused by agitation increases with scale.

The basic configuration of stirred tank bioreactors for mammalian cell culture is similar to that of microbial fer-

mentors. A major difference is that the aspect ratio (the height-to-diameter ratio) is usually smaller in mammalian cell culture bioreactors. The power input per unit volume of bioreactor is also substantially lower in mammalian cell culture bioreactors. Although the Rushton-type impeller is the norm in microbial fermentors, mammalian cell culture fermentors mostly employ marine-type impellers (16). As in microbial fermentors, an axial impeller is often mounted at the top of the bioreactor to drive the liquid downward in large-scale cell culture bioreactors. These differences reflect the different purposes of agitation in microbial fermentation and cell culture. In microbial fermentation agitation is needed at a higher power input to disperse air bubbles and to increase oxygen transfer efficiency, whereas in mammalian cell culture the primary purpose of agitation is to keep cells or microcarriers suspended in nutrient medium relatively uniformly. In general, the mixing time in a mammalian cell culture bioreactor is substantially longer than that in a microbial fermentor with similar scale. The oxygen transfer capacity in a cell culture bioreactor is also substantially lower than that in a microbial fermentor. However, the typical oxygen demand in a mammalian cell culture is also much lower (10 to 50 times) than that in microbial fermentation.

Because of the relatively low oxygen demand in cell culture, surface aeration is often sufficient to supply enough oxygen for optimal growth in small-scale bioreactors. In fact, surface aeration is the primary means of supplying oxygen in almost all laboratory-scale bioreactors. If necessary, enriched oxygen is used to increase the oxygen transfer from the gas phase to the culture fluid.

However, as the volume of the reactor increases, the area of gas–liquid interface per culture volume decreases, and surface aeration alone becomes inadequate. To supply oxygen, silicone rubber tubing is often installed inside the reactor to allow air or oxygen passing through it to supply oxygen without direct gas sparging. In addition to silicone rubber tubing, other materials with high oxygen permeability, such as porous polypropylene, can also be used. The oxygen transfer tubing need not necessarily be used as a single long piece; many different configurations for the installation of the oxygen transfer tubing can be employed. This method of supplying oxygen provides a convenient, relatively trouble-free way of sustaining cell growth at a reasonably high cell density. Applying such a technique in a large-scale operation could be very cumbersome. Nevertheless, it has been used in bioreactors of up to hundreds of liters in volume.

The most efficient way to supply oxygen, especially in large-scale reactors, is the direct sparging of air- or oxygen-enriched air. Direct sparging, though, can damage cells. Numerous studies have tried to clarify the mechanism of cell damage caused by direct gas sparging (17). The large amount of energy released as the gas bubbles emerge from the liquid surface appears to be responsible for its most damaging effects. These consequences can be partially alleviated by the addition of polyols in the culture medium. In fact, it is standard practice to include Pluronic F-68® in a low-protein or protein-free medium for suspension cultures.

Airlift Bioreactor

A variant of the bubble column reactor with internal circulation loops is used to improve the performance of conventional bubble column reactors. In airlift column reactors, internal liquid circulation is achieved by sparging only part of the reactor with air. The sparged section has a lower effective density than the bubble-free section, and the difference in hydrostatic pressure between the two sections induces the liquid circulation upward in the sparged section (riser) and downward in the bubble-free section (downcomer). The loop has the advantages of permitting high-efficiency mass transfer and improving the flow and mixing properties in the vessel.

These reactors are characterized by low costs, mainly because of their simple mechanical configuration. Considerable backmixing in both gas and liquid phases, high-pressure drop, and bubble coalescence can be disadvantages, in some cases. An additional benefit is the low energy requirement compared with stirred tank reactors. Although simple in construction, sound design is critical for optimal hydrodynamic behavior. Nevertheless, the flow in bubble column reactors is relatively well defined compared with that in stirred tank reactors.

Airlift bioreactors for cell cultivation are considered to be low-shear devices because there is no mechanical agitation. Also, the direct air sparging may not cause excessive cell damage, at least under normal cultivation conditions. The flow regime depends on the sparger used and the flow rate. Various types of spargers are used to provide different bubble sizes.

Airlift reactors have been used successfully with suspension cultures of BHK 21, human lymphoblastoid, CHO, hybridomas, and insect cells.

Stirred Tank in Perfusion Culture

The stirred tank bioreactors can be operated in batch, fed-batch, or continuous mode. Using typical culture medium, the cell concentration achieved in batch and continuous cultures is rather low, largely because of the high conversion ratio of key nutrients (glucose and glutamine) to waste and growth inhibitory metabolites, lactate, and ammonia. One method of increasing cell concentration in continuous culture is to retain cells inside the bioreactor while perfusing medium through. With cell retention, a 10-fold increase in cell concentration can be achieved. An effective cell retention device should permit easy and rapid separation of cells from spent medium and allow for long-term operation. Those continuous cultures with cell retention are typically referred to as perfusion cultures.

A cell retention device may be physically positioned inside or outside of the bioreactor vessel. The devices that have been used include settling tanks (18), external centrifuges (19), external membrane filtration (20), internal microfiltration (21), and rotating wire cage bioreactors (22). All of these, or similar devices, have been shown to be effective at laboratory scale. At industrial scale, bioreactors of up to hundreds of liters, centrifuges, settling devices, and spin filters appear to be top choices. The employment of a centrifuge or a settling device necessitates a recirculation loop to return the concentrated cell stream

back to the reactor while discarding the purged cells along with the other stream. The spin filter, on the other hand, is an internal cell retention device. This device is a rotating cylindrical screen mounted on the agitator shaft. It can be mounted to a shaft rotated by a top motor while a bottom motor drives the impeller. The filter screen has an opening ranging from 25 to 60 μm. The average cell diameter is 10–15 μm; however, the cell density inside the cage is much lower than outside the cage. The greatest difficulty encountered with the spin filter is the control of flux and cell retention efficiency.

NONMIXING TANK BIOREACTORS

Roller Bottle

Roller bottles are cylindrical screw-capped bottles mostly made of disposable plastic; reusable glass bottles are still used occasionally. Each bottle is typically about 1 to 1.5 L in total volume. A bottle is usually filled with 0.1 to 0.3 L culture medium for cell cultivation. A stack of bottles is placed on a roller, and the bottles rotate on the roller rack at 1 to 4 rpm and are incubated in an incubator or an incubation room. Roller bottles are used for the cultivation of both suspension cells and anchored cells. However, for suspension cells, it is usually more convenient to grow the cells in a stirred vessel, and roller bottles are only used on a small scale when convenience dictates this selection of cultivation methods. Anchorage-dependent or anchorage-preferred cells are frequently cultivated in roller bottles, both in the laboratory and in manufacturing of biologics. These cells attach to the inner wall of the bottle. Initially, they cover only part of the surface of the inner wall after inoculation. The cell layer is bathed in a liquid medium and is alternately exposed to medium and gaseous nutrients as the bottle rotates. As cells grow, they cover the entire surface and reach confluence. Normal diploid cells then reach contact inhibition and stop growing until they are detached by protease treatment and inoculated into a larger number of roller bottles. Many continuous cell lines or tumorous (transformed) cells can continue to grow to form multiple layers of cells after reaching confluence if sufficient nutrients are provided. Some variations of roller bottles are available to increase the cell growth surface area in a bottle. Spiral films and waffles on the inner wall of roller bottles have been introduced and are sometimes used.

For small-scale operations, roller bottles provide many advantages for the cultivation of adherent cells. It is relatively inexpensive to set up. Often it allows for a rapid change of throughput in response to need. Furthermore, replacing cell growth medium with one designed for product formation is rather straightforward. It is particularly useful when the serum-containing medium needs to be replaced by a serum-free or protein-free medium for the production of secreted proteins or viruses. The transparent glass or plastic wall allows visual or microscopic examination of the culture status. Microbial contaminated bottles can be readily spotted and discarded before they pool together with contaminated ones.

However, the drawbacks of roller bottles are numerous for large-scale production of biologics. On-line environmental monitoring and control are virtually impossible or at least impractical. The aseptic bottle handling for inoculation, trypsinization for cell detachment and expansion, medium exchange, and product harvest, require extensive well-trained labor to ensure a high success rate. Each batch can easily involve hundreds or even thousands of bottles, and the large number of manual steps involved makes the risk of microbial contamination rather high. Despite these significant drawbacks, roller bottles are still widely used in the production of viral vaccines, often because the products involved were approved by regulatory agencies before more recent cultivation technologies became more robust and widely accepted. In some cases, roller bottles were selected over other bioreactors because the process was easy to implement compared with those based on other bioreactors. One notable example is the production of erythroprotein (EPO) using recombinant Chinese hamster ovary cells. Recent improvements include developing a fully automatic and continuous cell culture for a large number of roller bottles to make industrial-scale production by mammalian cell culture (23).

Hollow Fiber

The use of hollow fiber reactors for cultivation of mammalian cells dates back to the early 1970s (24). A hollow fiber system can be used for anchorage-dependent and suspension cells. It consists of a bundle of capillary fibers sealed inside a cylindrical tube. The basic configuration is similar to the hollow fiber cartridge used in kidney dialysis. The hollow fiber, in most cases, consists of supporting polymeric porous materials for mechanical strength and a thin layer of membrane that provides selective passage of molecules depending on their size. In most cases, an ultrafiltration membrane is used (25). The molecular weight cutoff (MWCO) of the membrane differs according to the application, ranging from a few thousand to a hundred thousand daltons. The ultrafiltration membrane prevents free diffusion of secreted product molecules from passing through the membrane and allows them to accumulate in the extracapillary space to a high concentration. The culture media is usually pumped through the fiber lumen, and cells grow in the extracapillary space, or the shell side. Supply of low molecular weight nutrients to the cells and removal of waste products occur by diffusive transport across the membrane between the lumen and the shell spaces. Although the use of microfiltration hollow fiber membranes for cell culture is infrequent, it finds application in various research uses for studying metabolism and for the cultivation of anchorage-dependent or highly aggregated cells for which a convective flow of medium through the extracapillary space to bathe cells in medium is desired.

Table 1 lists some of the types of hollow fiber membrane that have been used in cell culture and kidney dialysis. Most hollow fiber systems used for cell culture employ fibers of approximately 250–350 μm in outer diameter. The interfiber distance is at least as large. A factor limiting the potential of scaling up a hollow fiber system is the diffu-

Table 1. Membrane Materials Used in Hollow Fiber Reactor and Kidney Dialysis

Membrane material	Cell type
Cellulose nitrate or cellulose acetate	Liver cells (26)
Cellulose acetate	Hybridoma (27)
	Hepatocytes (28,29)
Polysulfone	Hybridoma (30,31,32)
	Chinese hamster lung fibroblast (33)
	MDCK (Madin-Darby Canine Kidney) (34)
	Hepatocytes (35)
Polymethyl-methacrylate	Kidney (36)
Cuprophane or polyacrylonitrile	Kidney (37)
Acrylic copolymer	SV3T3 cells (38)

sional resistance. In the radial direction, the nutrients (including oxygen) diffuse through the hollow fiber wall and membrane to reach the site of cell consumption. In the axial direction, the medium entrance region is exposed to fluid with higher oxygen, whereas the exit end inevitably sees a lower concentration of oxygen. Oxygen is of primary concern because it is usually the first nutrient to become growth limiting. To ensure ample oxygen supply, the recirculating medium is typically passed through an oxygenator (e.g., one used in surgery operations) before it enters the hollow fiber system. A typical setup of a hollow fiber system therefore includes not only a constant-temperature incubator to house the hollow fiber bioreactors but also pH control, oxygenation and oxygen monitoring systems, and a medium reservoir for recirculating medium. In addition, a small to medium reservoir is frequently used to supply high molecular weight nutrient supplements to the extracapillary space directly. A harvest reservoir is used to collect the product containing fluid from the extracapillary space. In a hollow fiber bioreactor, a nonuniform growth of cell mass and product accumulation is often seen. This growth is usually caused by the small convective flow in the shell side, the difficulty of inoculating cells uniformly, and settling due to gravity.

One technique to improve oxygen supply is the use of fiber or silicon tubing for oxygen supply into the shell side of the reactor. However, the use of mixed fibers in the bioreactor makes the bioreactor harder to manufacture. Another technique is to reverse the flow direction periodically to inverse the axial regions exposed to high and low oxygen concentrations.

Despite its limitation in scaling up, the hollow fiber system is a proven technology relatively easy to apply to processes that require a modest amount of product. In some tissue engineering and cellular therapy applications, the product is often patient specific and the quantity of cells needed is relatively small. A hollow fiber system could be ideal for those applications.

Ceramic Matrix

The ceramic bioreactor, to some extent, can be considered a variant of the hollow fiber system. It consists of a cylindrical ceramic core with many channels that pass longitudinally through the ceramic material. Cells inoculated in the channels adhere to the material or become entrapped in the pores of the ceramic. As in a hollow fiber system, ceramic reactors are supported by medium perfusion loops. Cell culture medium is pumped through the longitudinal channels in the ceramic cores from a medium reservoir in a recirculating loop configuration. Fresh medium is fed into the system, and harvested culture fluid is removed to the medium reservoir. Unlike the hollow fiber system, no membrane separates the cells and bulk medium. Product is secreted directly into the bulk medium. Essentially, the ceramic bioreactor can be used to conveniently replace a large number of roller bottles. As in hollow fiber systems, an oxygen concentration gradient develops along the axial direction and limits the length (i.e., the scale) of the reactor.

CONCLUDING REMARKS

The bioreactors described in this article represent those used widely in the cultivation of mammalian cells. They can be operated in batch, fed-batch, or continuous mode. Regardless of which operating mode a process is based on, it invariably involves a transient period in which the cell concentration is increasing. Even in a continuous perfusion culture, the concentrations of cells, nutrients, and product are often not at steady state. On-line sensors for environmental factors and control of important variables, including pH, temperature, dissolved oxygen concentration, and medium perfusion rate, are thus commonly employed. The sensors and control strategies used for mammalian cell cultivation are basically identical to those used in microbial fermentation. However, mammalian cell cultures are mostly used for the production of more complex molecules or viruses, for which the quality of the product is less definable than the smaller molecules (such as primary or secondary metabolites in microbial fermentation). They are thus more prone to change caused by fluctuations in environmental factors. A key consideration in selecting a mammalian cell culture bioreactor for a particular process is therefore the controllability of process variables and the consistency in reactor performance. Ultimately a good manufacturing process must employ a reactor that can be scaled up and controlled to produce the necessary quantity of product with the highest quality.

BIBLIOGRAPHY

1. A.L. Van Wezel, *Nature* **216**, 64–65 (1967).
2. T.-Y. Tao, G.-Y. Ji, and W.-S. Hu, *J. Biotechnol.* **6**, 9–12 (1987).
3. A. Zuhlke, B. Roder, H. Widdecke, and J. Klein, *J. Biomater. Sci. Polymer Edn.* **5**, 65–78 (1993).
4. B. Roder, A. Zuhlke, H. Widdecke, and J. Klein, *J. Biomater. Sci. Polymer Edn.* **5**, 79–88 (1993).
5. J. Varani, M. Dame, J. Fediske, T.F. Beals, and W. Hillegas, *J. Biol. Stand.* **13**, 67–76 (1985).
6. S. Reuveny, L. Silberstein, A. Shahar, E. Freeman, and A. Mizrahi, *Dev. Biol. Stand.* **50**, 115–123 (1982).

7. W. Tolbert, M.M. Hitt, and J. Feder, *In Vitro* **16**, 486–490 (1980).
8. M.V. Peshwa, Y.-S. Kyung, D.B. McClure, and W.-S. Hu, *Biotechnol. Bioeng.* **41**, 179–187 (1993).
9. S. Goetghebeur and W.S. Hu, *Appl. Microbiol. Biotechnol.* **34**, 735–741 (1991).
10. C.M. Perusich, S. Goetghebeur, and W.-S. Hu, *Biotechnol. Tech.* **5**, 145–148 (1991).
11. J.L. Moreira, P.M. Alves, J.M. Rodriguez, P.E. Cruz, J.G. Aunins, and M.J.T. Carrondo, *Ann. N.Y. Acad. Sci.* **745**, 122–133 (1994).
12. K. Nilsson, W. Scheirer, O.H. Horton, L. Osfberg, E. Liedhl, H.W.D. Katinger, and K. Mosbach, *Nature* **302**, 629–630 (1983).
13. K.K. Papas, I. Constantinidis, and A. Sambanis, *Cytotechnology* **13**, 1–12 (1993).
14. A.C. Jen, M.C. Wake, and A.G. Mikos, *Biotechnol. Bioeng.* **50**, 357–364 (1996).
15. M.S. Croughan, J.-F. Hamel, and D.I.C. Wang, *Biotechnol. Bioeng.* **29**, 130–141 (1987).
16. W.-S. Hu and T.C. Dodge, *Biotechnol. Prog.* **1**, 209–215 (1985).
17. J.J. Chalmers and F. Bavarian, *Biotechnol. Prog.* **70**, 151–158 (1991).
18. K. Kitano, Y. Shintani, Y. Ichimori, K. Tsukamoto, S. Sasai, and M. Kida, *Appl. Microbiol. Biotechnol.* **24**, 282–286 (1986).
19. M. Tokashiki, T. Arai, K. Hamamoto, and K. Ishimaru, *Cytotechnology* **3**, 239–244 (1990).
20. Y. Takazawa, M. Tokashiki, K. Hamamoto, and H. Murakami, *Cytotechnology* **1**, 171–178 (1990).
21. T.C. Seamans and W.-S. Hu, *J. Ferment. Bioeng.* **70**, 241–245 (1990).
22. V.M. Yabannavar, V. Singh, and N.V. Connelly, *Biotechnol. Bioeng.* **43**, 159–164 (1994).
23. R. Kunitake, A. Suzuki, H. Ichihashi, S. Matsuda, O. Hirai, and K. Morimoto, *J. Biotechnol.* **52**, 289–294 (1997).
24. R.A. Knazek, P.M. Gullino, P.O. Kohler, and R.L. Dedrick, *Science* **178**, 65–67 (1972).
25. T.L. Evans and R.A. Miller, *BioTechniques* **6**, 766–767 (1988).
26. J. Rozga, F. Williams, M.-S. Ro, P.F. Neuzil, T.D. Giorgo, G. Backfisch, A.D. Moscroni, R. Hakim, and A.A. Demetriou, *Hepatology* **17**, 258–265 (1993).
27. R.J. Gillies, P.G. Scherer, N. Raghunand, L.S. Okerlund, R. Martinez-Zagrilan, L. Hesterberg, and B.E. Dale, *Magn. Reson. Med.* **18**, 181–192 (1991).
28. N.L. Sussman, G.T. Gislasson, C.A. Conlin, and J.H. Kelly, *Artif. Organs* **18**, 390–396 (1994).
29. V. Dixit, E. Piskin, M. Arthur, A. Denizli, S. Tuncel, E. Denkbas, and G. Gitnick, *Cell Transplant* **1**, 391–399 (1992).
30. G.L. Altshuler, D.H. Dziewulski, J.A. Sovek, and G. Belfort, *Biotechnol. Bioeng.* **28**, 646–658 (1986).
31. J.M. Piret, D.A. Devens, and C.L. Cooney, *Can. J. Chem. Eng.* **69**, 421–428 (1991).
32. D. Patankar and T. Oolman, *Biotechnol. Bioeng.* **36**, 97–103 (1990).
33. J.P. Tharakan and P.C. Chau, *Biotechnol. Bioeng.* **28**, 329–342 (1986).
34. H.D. Humes, M. Mackay, A.J. Funke, and D.A. Buffington, *Am. J. Kydney Dis.* **30**, S58–S31 (1997).
35. H.O. Jauregui, C.-P. Mullon, D. Trenkler, S. Naik, H. Santangini, P. Press, T.E. Muller, and B.A. Salomon, *Hepatology* **21**, 460–469 (1995).
36. R.M. Hakim, R.L. Wingard, and R.A. Parker, *N. Engl. J. Med.* **331**, 1338–1342 (1994).
37. M. Schiffl, S.M. Lang, A. Konig, T. Strasser, M.C. Haider, and E. Held, *Lancet* **344**, 570–572 (1994).
38. K. Ku, M.J. Kuo, J. Delente, B.S. Wildi, and J. Feder, *Biotechnol. Bioeng.* **213**, 79–95 (1981).

See also SUSPENSION CULTURE, ANIMAL CELLS.

MAMMALIAN CELL CULTURE REACTORS, SCALE-UP

J. BRYAN GRIFFITHS
Scientific Consultancy & Publishing
Salisbury, United Kingdom

KEY WORDS

Anchorage-dependent cells
Bioreactors
Cell culture processes
Cell products
Mammalian cell culture
Process parameters
Scale-up
Suspension cells

OUTLINE

INTRODUCTION

Although mammalian cells have been grown in vitro since before 1907 (1) the factor that gave impetus to their current widespread laboratory and industrial use was the discovery by Enders (2) that human pathogenic viruses could be grown in cell cultures. Following this demonstration by Enders in 1949 it took only 5 years for the first cell-based vaccine to be licensed for clinical use (Salk polio vaccine in primary monkey kidney cells in 1954). This opened up 20 years of continuous development of human and veterinary viral vaccines and created the need for industrial-scale cell culture processes. The most effective large-scale process developed during this period was for foot-and-mouth disease virus (FMDV) based on suspension culture of baby hamster kidney (BHK) cells (3). Developments in processes for human vaccines were less dramatic because of the need to use biologically safe cell lines. This meant that the use of human diploid cell lines, such as WI-38 (4) and MRC-5 (5), because of their normality (i.e., they behaved as cells in vivo without tumorgenic transformation), only grew attached to a substrate (anchorage dependent) and only reached low cell densities. The wide range of animal cell reactors available is partly due to the dual development of systems for anchorage-dependent cells (ADCs) and free-suspension cells (6,7). ADCs were grown in small culture vessels, and a production batch constituted hundreds or even thousands of replicate cultures (i.e., a multiple-batch process) (8). The need for a unit batch process (one large culture vessel), such as the fermentors used for suspension cells, saw the development of a wide range of novel culture reactors (6,7), but the significant breakthrough came with the microcarrier system (9). This procedure, developed in 1967 by van Wezel (9), allowed cells to grow attached to small (200 μm) spheres that were stirred in a large tank fermentor analogous to suspension cells. This was the first successful large-scale unit process for ADCs.

Vaccines were the dominant product until the 1970s, but changes in regulatory and licensing procedures then allowed cells from sources other than normal tissues to be used for human medicinal products. This came about during the development of an interferon process using a cancer cell line, Namalva, by Wellcome (10). This company's pioneering work established the safety criteria, and thus acceptance, for using nonnormal (heteroploid, transformed, or tumor-derived cell lines) and the feasibility of scaling up an industrial cell culture process to 8,000 L. Cell culture then entered a new, or modern, phase in which a wide variety of cell products (Table 1) is produced from a range of cell types (Table 2). Two of these products can be highlighted as being significant milestones. First, the production of monoclonal antibodies from hybridoma cells (the fusion of a normal antibody producing cell and a hemopoitic cancer cell) (12) has given rise to hundreds of new products. Second, the development of recombinant tissue plasminogen activator (tPA) by Genentech (14) gave rise to the first genetically engineered clinical product from cell cultures. Currently applications are widening to include the cell itself as a product in tissue engineering, organ replacement, and gene therapy (15).

In this article the description of cell culture reactors is given against this historical evolution of cultures from small-scale virus production through scale-up to the more than 10,000-L unit processes that now play a dominant role in the health biotechnology industry, with hundreds of products either already available in the clinic or going through phase I through III clinical trials. The emphasis is on scale-up because of the relatively low biomass productivity of animal cells compared with bacteria, the low product expression rates (one clinical dose is often equivalent to 0.1 to 100 L of culture supernatant), and the need to introduce more efficient and economical industrial processes.

BACKGROUND

Mammalian Cells

Major differences between single mammalian cells and microorganisms should be noted because they influence, and put constrictions on, the culture process. Cells do not have a cell wall but a thin cell membrane (a phospholipid bilayer) that mediates communication between the cell and the environment. They are thus very sensitive to environmental stimuli, such as pH, nutrients, oxygen, regulatory molecules, and hydrodynamic forces. Cell shape and morphology are controlled by cytoskeletal elements attached to the membrane and running through the viscous cytoplasmic gel. The degree of polymerization of the microtubular network regulates the shape, whether spherical or elongated, flattened or attached. Mammalian cells are on the order of 12–20 μm in diameter (i.e., 1,000-fold greater volume than microorganisms), which means diffusion into the cytoplasm is greatly reduced in the spherical state. When flattened, membrane folding and ruffling give a much greater surface area. Also, mammalian cells are far more fastidious in their nutrient requirements, with critical optimal chemical and physical environmental parameters. The consequence of these factors is that cells need more gentle stirring and aeration procedures than bacteria, which greatly reduces mass transfer and homogeneous mixing in the culture. This need for more moderate stirring and aeration procedures, together with greater sensitivity to nonoptimum or fluctuating culture conditions, slower growth rate (14–22 h average doubling time), and vulnerability to microorganism contamination, means that the culture system has to be extremely well controlled and contained. Some key issues in setting up successful cultures are as follows:

- Cells should be inoculated from the late logarithmic phase rather than stationary cultures.
- Cells should be prepared for inoculation as quickly as possible and the time they are exposed to trypsin, acidity, or low temperatures limited during harvesting.
- The new culture medium should be at the required temperature (37 °C), pH (usually 7.1–7.3), and so forth before inoculating.
- The minimum inoculation density ($1–2 \times 10^5$/mL or 2×10^4/cm^2) should always be exceeded.

Table 1. Mammalian Cell Products

	Native products
Human vaccines	Polio (1954), measles (1963), rabies (1964), mumps (1969)
Veterinary vaccines	Foot-and-mouth disease vaccine, rabies, Marek's, pseudorabies, BVD, louping ill, bluetongue, avian influenza, canine distemper (examples)
Interferon	Refs. 10 and 56
	Engineered products
Monoclonal antibodies	OKT3/Orthoclone (1987), Centoxin (1990), Reopro (1994), Myoscint (1989), Oncoscint (1990)
Tissue-type plasminogen activator	Activase/actilyse (1987)
Erythropoeitin (EPO)	Epogen/Procrit/Eprex (1989), Epogin/Recormon (1990)
Human growth hormone	Saizen (1989)
HBsAg	GenHevac B Pasteur (1989), HBGamma (1990)
Interferon	Roferon (1991)
Colony stimulating factor (G-CSF)	Granocyte (1991), Neupogen (1991)
Blood factor VIII	Recombinate (1992), Kogenate (1993)
Dnase I	Pulmozyme (1993)
Glucocerebrosiduse	Cerezyme (1994)
Follicle-stimulating hormone	Gonal-F (1995)
	rDNA products in development/clinical trial
HIV vaccines (gp120, pg160, CD4)	
Herpes simplex vaccines (gB, gD)	
Chimeric monoclonal antibodies (her2, CD4, TNFa, CD20, Cd18, TAC, leukointegrin, CF54, RSV)	
Others (TSH, TNF, M-CSF, IL-6, IL-1)	
In vitro diagnostic monoclonal antibodies (1–200)	
Tissue engineering and replacement (e.g., skin, artificial kidneys)	

Scale-Up

Principles. Normally a culture process begins with an ampoule, stored at −196 °C, of about 5 million cells put into 10 mL of medium, which can then be expanded up to a final volume of 10–20,000 L at 2 million cells per milliliter (i.e., a 1 to 2 million–fold scale-up). This expansion is achieved by a series of ever-increasing culture volumes (the seed train), each step bringing about a 5- to 10-fold volume increase and taking 4 to 5 days. Thus, before the final production culture is initiated, cells will have been passaged at least five times over 25 days. This investment in time and resources shows how important it is to keep sterility by strict attention to all aspects of the process (equipment, reagents, and handling procedures). The scale-up process involves many factors, including engineering (especially to maintain sterility and reliability), physiological (supply of oxygen and nutrients and critical maintenance of the correct environmental conditions), monitoring and control (biosensors, performance measurements, and computer control), logistics, and compliance with pharmaceutical licensing and health and safety regulations. The aim of such scale-up is to provide more cells and more cell product in as efficient and cost-effective a manner as possible while complying with all regulatory and health and safety criteria. This emphasizes the importance of moving from a multiple-unit (many replicate bottles, roller cultures, spinner flasks, etc.) to a high-volume-unit process. Such a shift allows economy in manpower (reduction in repetitive steps), space (expensive hot rooms), and time and an increased efficiency by being able to reproducibly control environmental factors, such as pH and oxygen, in a unit process. The wide range of alternative culture systems that are or have been available (6,7,16) is due (*1*) the need to find processes for both suspension and ADC cells and (*2*) various solutions to overcoming limiting factors in scale-up (17).

Limiting Factors in Scale-Up. The first problem in moving from a small, multiple unit process to a large unit process is usually oxygen limitation. To overcome this problem and supply oxygen without the damaging effects of air bubbles from sparging and the high stirring speeds needed to effect efficient mass transfer, bioreactors have been developed that oxygenate through a membrane. This bubble-free aeration has been achieved by means of long lengths of tubing in the bioreactor (18–20) or external medium loops through hollow fiber oxygenators. A related factor is mixing, and low-shear modifications to impellers and the airlift bioreactor concept (21) have been introduced. A problem for ADCs is the limitation on surface area during scale-up. A huge range of reactors have been developed (multiple plates, spirals, ceramic matrices, glass beads), but the most successful method has been the microcarrier (9,22). The large fermentor systems used for suspension cells thus became available for ADCs because the small microcarriers with cells attached behaved analogously to a suspension cell.

The next problem to be overcome is nutrient limitation and toxic metabolite buildup. Although media changes can be used (especially for attached cells), the most efficient means of overcoming these problems is continuous medium perfusion. To perfuse suspension cells cells and waste

Table 2. Typically Used Cells in Biotechnology

Cell line	Derivation	Application
CHO	Chinese hamster ovary	Good product glycosylation
CHO dhfr	Mutant for genetic amplification	Principal substrate for recombinant products
BHK21	Syrian hamster	Viruses/veterinary vaccines
Vero	African green monkey	Viral vaccines
HeLa	Human cervical adenocarcinoma	Polio vaccine
MRC-5/WI-38	Human fetal lung fibroblasts	Human vaccines
MDCK	Cocker spaniel kidney	Viral vaccines
Namalwa	Human B lymphoblastoid	Interferon
COS 1/COS 7	African green monkey kidney	Recombinant products
NS0/NS1 and Sp2/0-Ag14	Myeloma	Fusion partners for monoclonal antibody production immunoglobulin
J558L	Mouse Balb/C myeloma	Secretes immunoglobulin A
NIH/3T3	Swiss mouse embryo	Sarcoma virus
GH3	Rat pituitary tumor	Hormone studies
293	Adenovirus transformed HEK	Gene therapy
X CRE/Y CRIP	Mouse NIH3/Moloney leukemia	Retrovirus packaging
	Insect	
Sf9	Spodoptera frugipeda	Baculoviruses
Schneider-2	Drosophila melangogaster	Recombinant products

Source: Ref. 12.

medium have to be physically separated; this separation has been done by means of spin filters, external loops to filters or centrifuges, or gravitational settling devices (18). The use of perfusion resulted in the development of many immobilization procedures to prevent cell washout (e.g., encapsulation, hollow fiber devices, membrane reactors, gels, fibers, and porous microcarriers). Additional benefits to these developments were the provision of huge unit surface areas for cell attachment and the ability to grow cells at 50- to 100-fold higher densities than previously achieved in free suspension. Thus these developments mean that manufacturers have a choice in scale-up routes (6).

Scale-Up Strategies. Some scale-up strategies are as follows (22–24):

- High-volume (to 10–20,000 L) tank fermentor systems with conventional cell densities of 1–2 $\times$ 10^6/mL
- Low-volume (<1 L) but very-high-density (1–2 $\times$ 10^8/mL) heterogeneous systems (e.g., hollow fiber bioreactors)
- Intermediate systems using spin filters to increase cell density to 1–2 $\times$ 10^7/mL but to a volumetric scale of only 500 L
- High-density (1 $\times$ 10^8/mL) scaleable systems (to 200 L) based on porous microcarriers

These systems are summarized conceptually in Figure 1. It should be emphasized at this stage that there is no ideal system, but the advantages of high-density systems, summarized in Table 3, show why a choice has to be made. Immobilized systems, such as porous microcarriers, also have the advantage of being equally suitable for suspension and anchorage-dependant cells (the Universal System?), protecting from shear, and having a three-dimensional structure that enhances productivity. Basically, the low-volume, high-density systems fill an important niche in manufacturing capability (e.g., gram amounts of monoclonal antibodies), whereas the high-volume systems are the most widely used scale-up route for pharmaceutical products.

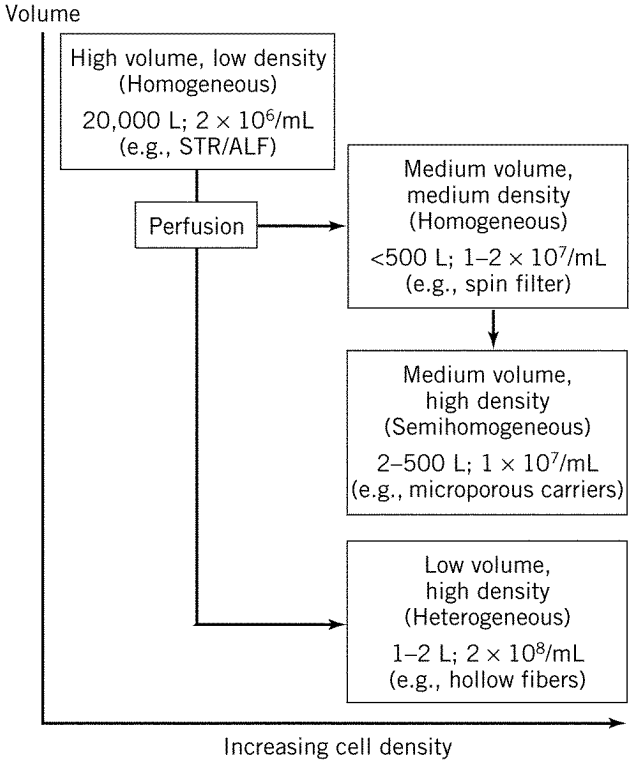


Figure 1. Scale-up strategies and the evolution of high cell density culture systems. STR, stirred-tank reactor; ALF, air-lift fermenter.

Table 3. Comparison of High- and Low-Cell-Density Culture Systems

Advantages of high cell density
Smaller reactor volume and therefore reduced need for high-quality laboratory space
More concentrated product
Reduced requirement for expensive serum and growth factors
Long run lengths (50–100 days) and thus less downtime and seed preparation
Daily or continuous harvesting
Shorter residence time for products (and exposure to preteolytic degradation)
Cell debris washout
Disadvantages of high cell density
Volumetric scale-up more difficult and complex
Mass transfer problems leading to heterogeneity
Sterilization problems because some systems cannot be heat sterilized
Higher degree of process control needed (direct cell measurement impossible)
More validation and regulatory conditions for product licensing needed
Perceived as experimental and therefore not accepted as readily as traditional methods

Culture Modes

Suspension culture is the preferred method for scale-up because it is easier and cheaper to scale up, requires less space, can monitor cell growth, and can more easily control environmental parameters. However, many cell lines have higher specific productivity when attached to the substrate. In the urge to move to suspension systems the following advantages of anchorage-dependent systems should not be overlooked:

- They facilitate complete medium changes, which are often needed to wash out serum before adding serum-free production medium, for example.
- They are easier to perfuse and thus achieve higher cell densities and healthier cells.
- Media-to-cell ratios are more easily changed during an experiment, thus increasing product titre.

Microcarrier culture, especially porous microcarrier (25), has most of the advantages of both suspension and anchorage-dependent systems.

Different types of culture strategy can be employed; they are defined as follows:

Batch culture. Cells grow from seed to final density over a 4- to 7-day period and are then harvested (typical monoclonal antibody yield is 6 mg/L per batch run)

Fed-batch culture. Additional media or medium components are added to increase culture volume and density of cells as the culture progresses.

Semicontinuous batch culture. A batch culture is partially harvested (e.g., 70%), topped with fresh medium, allowed to grow up again, and reharvested. Typically three to four harvests can be made, although usually with diminishing returns at each harvest.

Continuous-flow (chemostat) culture. Because it is completely homogeneous for long periods of time while cells are in a steady state, this is an extremely useful tool for physiological studies but not very economical for production (by definition, cells are never at maximum density).

Continuous-perfusion culture. This type of culture differs from chemostat in that all nutrients are kept in excess and cells are retained within the bioreactor. This is the most productive culture system, with typical monoclonal antibody yields of 10 mg/L per day for 50–100 days.

There is no definitive answer as to the best system to use; the answer depends on the nature of the cell and product, the quantity of product, downstream processing capability, licensing regulations, and so forth. However, a rough guide to relative costs for producing monoclonal antibodies by perfusion, continuous-flow, and batch culture is 1:2:3.5 (26).

REACTORS FOR ADCS

Basic Culture Units

Tissue culture flasks and tubes giving surface areas of 5–200 cm^2 are the familiar stock items of any culture laboratory. The largest stationary flask routinely used is the Roux bottle (or disposable plastic equivalent), which gives a surface area of 175–200 cm^2 (depending on make and type), needs 100–150 mL medium, and utilizes 750–1,000 cm^3 of storage space. Such a vessel will yield 2×10^7 diploid, or 10^8 heteroploid cells; thus, to produce a modest 10^{10} cells, more than a hundred replicate cultures are needed (i.e., manipulations have to be repeated 100 times) and 100 L culture space is required. The need to move to larger units is obvious and is approached by increasing the surface-area-to-volume ratio. The first step in scale-up usually involves a change from stationary flasks to roller bottles (i.e., a dynamic system). Roller bottles can be up to 1,750 cm^2, use 350 mL medium, and have a volume of 2.5 L (i.e., a ninefold increase in surface area but only a threefold increase in medium and total volume). This is brought about by using the total internal surface area for cell growth and more efficient aeration as the cells move in and out of the culture fluid. However, this procedure has been used industrially for viral vaccines, veterinary vaccines in multiples of 28,000 (8), interferons, and EPO. It has a place in modern processes largely because it is automatable with robotic systems, such as the Cellmate (Automation Partnership, Royston, Herts, U.K.) and that described by Kunitake et al. (27). All the routine manipulations of cell seeding, media changing, bottle gassing, cell sheet rinsing, trypsinization, and cell collection by scraping can be carried out automatically, and reproducibly, with very precise volumes. Examples of products produced by this method are Varivax (Merck varicella vaccine) and Saizen (Serono recombinant human growth hormone).

Many modifications have been made to the roller bottle to increase the surface-area-to-volume ratio. Although some of the following adaptations are no longer commercially available, they are included to stimulate further innovation:

- *Disks*. A series of vertical disks placed along the horizontal axis of the roller bottle (28)
- *Spira-Cell Multi-Surface Roller Bottle (Bibby Sterlin Ltd.)*. A spiral polystyrene cartridge on which cells grow on both sides of the film, available in 2,000, 4,500, and 6,000 cm^2 sizes (29).
- *Glass tubes*. A roller bottle packed with a cluster of small glass tubes arranged in parallel and separated by silicone spacer rings, available in 1,000, 10,000, and 15,000 cm^2 sizes (Bellco-Corbeil Culture System) (30). The initial concept was developed by Chemap as the Gyrogen (31)
- *Extended surface area roller bottles (ESRB)*. A polystyrene roller bottle with a ribbed, or corrugated, surface that doubles the area of an 850-cm^2 bottle to 1,700 cm^2 (Bibby Sterlin Ltd.).
- *ImmobaSil*. Silicone rubber matrix–coated roller bottles that increase the surface area severalfold (Ashby Scientific Ltd., Coalville, U.K.).

Also, to increase cell productivity, roller bottle perfusion systems (32) with specially modified swivel caps (Bellco Autoharvester, New Brunswick Roller Apparatus) can be used. The roller bottle is a well-established technique and is still widely used in both the research laboratory and the industrial plant. Some cell lines (particularly epithelial) may not be as successfully grown in roller bottles because of streaking, clumping, or inadequate spreading over the total surface (i.e., nonlocomotory cell lines) as in stationary bottles. An alternative scale-up route is to use multisurface plate stationary systems.

Multisurface Plate Units

There are two commercially available examples of multiple-layered polystyrene plates stacked within a polystyrene box. In the Cell Factory (A/S NUNC, Roskilde, Denmark), the trays are 335 × 205 mm (600 cm^2) and can be obtained in multiples of 1, 2, 10, and 40. The largest, therefore, has a surface area of 24,000 cm^2 and needs 8 L medium (i.e., it is equivalent to 14 large roller bottles but requires half the incubation space and no ancillary roller equipment). However, for convenience of handling, especially large numbers, the Cell Factory Manipulator is advantageous because it can be operated robotically.

Although the CellCube (Costar, Cambridge, Mass., USA) (33) also has parallel polystyrene trays, it is a modular closed-looped perfusion system, including an oxygenator, pumps, and a system controller (pH, O_2, level control). The unit is very compact, with the trays being only 1 mm apart; thus, the smallest unit of 21,250 cm^2 is less than 5 L total volume (1.25 L medium). Additional units of 42,500 cm^2 (2.5 L medium) and 85,000 cm^2 (5 L medium) are available, and four units can be run in parallel in the system, giving 340,000 cm^2 growth area.

High-Volume Units

High-volume units are true unit process systems, analogous in volume and performance to suspension cell fermentor vessel processes. Examples include (*1*) the multisurface propagator (34), with vertically stacked circular stainless steel plates on a central revolving shaft scaled up to 200 L (200,000 cm^2), and the titanium plate reactor (35); (*2*) the plate heat exchanger (36,37); (*3*) glass spheres (38–41); (*4*) microcarriers (9,42,43); (*5*) porous microcarriers (44–46); and (*6*) stainless steel wire coils (48).

Glass Bead Culture. Packed beds of 3- to 5-mm glass spheres through which medium is continuously perfused has a demonstrated scale-up to 100 L (49). Spheres of 3-mm diameter pack sufficiently tightly to prevent the bed from shifting but allow sufficient medium flow up the column so that fast flow rates, which would cause shear damage, are not needed. The physical properties of glass spheres are given in Table 4. Medium can be circulated by pump or by airlift (to give better oxygenation) (50). Although 3-mm beads are normally used to maximize the available surface area, the use of 5-mm spheres can actually give a higher total cell yield, despite the reduced surface area per unit volume, because of better physiochemical conditions (6,39). Scale-up has vertical limitations in that it is difficult to maintain a uniform plug flow without channeling. For inoculation, the use of a perforated tube placed centrally for the height of the bed allows even distribution of cells. Also, alternating the direction of medium flow on a daily basis increases homogeneity and thus scale-up potential. In conclusion, glass sphere packed beds constitute a simple system (that minimizes moving parts and the risk of mechanical failure) with an inexpensive and reusable substrate capable of considerable radial and reasonable vertical scale-up to operate beds of volumes more than 200 L using 5-mm spheres. The disadvantages are that it is a low-intensity system (spheres have the minimum surface area per unit volume) and is limited to secreted products because it is difficult to harvest cells from the bed (i.e., it is ideal for long-term continuous cultures rather than batch cultures, as demonstrated by Brown et al. [51]). Experimental protocols for using glass sphere technology have been published (52).

Table 4. Properties of Solid Glass Sphere Fixed-Bed Reactors

Glass sphere (diameter)	3 mm	5 mm
Weight, g	0.0375	
Surface area, cm^2/kg	7,546	4,570
Volume, mL/kg	990	625
Void fraction	0.46	0.40
Void volume, mL/kg	455	250
Channel size, mm^2	0.32	1.00
Vero cell yields		
$10^5/cm^2$	0.78	2.5/3.1[a]
10^8/kg	4.00	8.0/9.9
10^6/mL	1.6	3.2/4.0

[a]Yield in 1 kg/10 kg bed reactors.

Microcarrier Culture. Microcarriers are small particles, usually spheres, 100 to 300 μm in size, that are suspended in stirred culture medium. The concept was initiated in 1967 (9); however, the chromatographic grade of dextran (Sephadex A-50) then used proved unsuitable for consistent and reliable growth. After considerable developmental work, Pharmacia produced a range of suitable microcarriers—the Cytodex series. The first industrial process based on microcarriers was for FMDV (43,53). Subsequently a wide range of microcarriers based on gelatin, collagen, polystyrene, glass, cellulose, polyacrylamide, and silica have been manufactured to meet all situations (23,54) (Table 5). The key criteria were to get the surface chemically and electrostatically correct for cell attachment, spreading, and growth. The power of the method is exemplified by the following data: 1 g Cytodex = 6,000 cm^2, which at 2 g/L = 12,000 cm^2/L (i.e., eight large roller bottles). Scale-up to 4,000 L (equivalent to 3,200 roller bottles) has been achieved in an environmentally controlled and optimized process. This method opened up industrial production opportunities (Table 6) and allowed research laboratories to easily produce substantial quantities of developmental products.

Experimental considerations and principles (54) are as follows:

- Because cells differ in their attachment requirements, a range of microcarriers should be assessed.
- Microcarriers vary in their density (1.01 to 1.06), mechanical robustness, hardness, and the concentration at which they can be used. These factors have implications for the choice of process (e.g., stirring system, need for rapid sedimentation due to frequent medium changes, and total unit surface area required: Cytodex 6,000 cm^2/g, glass or polystyrene 300 cm^2/g).
- Culture equipment is critical. Small-scale culture requires the use of specially designed spinner flasks rather than the conventional magnetic bar and siliconized glass to prevent microcarrier attachment.
- Culture initiation conditions are critical and should include a high-quality single-cell inoculum and a prewarmed and stabilized culture medium. Inoculation should not be undertaken below the minimum cell number per bead (usually five, optimally seven), and a nutritionally supplemented medium (e.g., with nonessential amino acids) might be necessary.
- Instead of counting by the trypan blue technique, after trypsinization the nuclei-counting method (57), should be used.

Table 5. Microcarriers

Name	Manufacturer	Material	cm^2/g
	Solid microcarriers		
Bioglas	Solo Hill Engineering	Glass/latex	350
Bioplas	Solo Hill Engineering	Polystyrene	350
Biosilon	Nunc	Polystyrene	255
Cytodex 1, 2	Pharmacia	Dextran	6,000
Cytodex 3	Pharmacia	Collagen	4,600
Dormacell	Pfeifer & Langen	Dextran	7,000
Gelibead	Hazelton Laboratories	Gelatin	3,800
Micarcel G	Reactifs IBF	Polyacrylamide	5,000
Microdex	Dextran Products	Dextran	250
			Diameter (μg)
	Microporous microcarriers		
Cellsnow	Kirin Ltd	Cellulose	800–1,000
Cytocell	Pharmacia	Cellulose	180–210
Cultispher	Hyclone	Gelatin	170–270
Cytoline 1,2	Pharmacia	Polyethylene	1,200–1,500
ImmobaSil	Ashby Scientific	Silicone rubber	1,000
Siran	Schott Glaswerke	Glass	400–5,000

Table 6. Industrial Processes Using Microcarrier Technology

Product	Cell line	Scale (L)	Company	Reference
Foot-and-mouth disease vaccine	Pig kidney	235	Merrieux	50
Polio and rabies vaccines	PKMC	500	RIVM (Bilthoven)	55
Polio and rabies vaccines	Vero	1,000	Merrieux	43
Human IFN-β	HDC	50	Sclavo	11
Human IFN-β	HDC	4,000	Toray	
Human growth hormone	Vero	100	UNSW (Australia)	56
Tissue-type plasmingon activator	Bowes melanoma	40	RIVM (Bilthoven)	42

- Some microcarriers can be reused, and it may be possible to recolonize microcarriers in situ during scale-up (58).
- Enzyme harvesting (e.g., trypsin) can be difficult or damaging, so consideration should be given to using a microcarrier that can be enzymically digested (e.g., by dextranase or collagenase), leaving the cells intact.
- Microcarriers, such as Cytodex, should be used between 1 and 3 g/L unless a perfusion system (e.g., using a spin filter) is employed, with which concentrations of 15 g/L are feasible (39).
- Inoculation in a reduced volume assists attachment and should be followed by topping up to the final working volume after 4–8 h.
- An advantage of microcarriers is the ease with which cell samples can be taken and analyzed.
- As the culture progresses, partial medium changes will have to be carried out and the stirring speed increased from 40–60 rpm to 75–90 rpm to prevent microcarrier aggregation.

In conclusion, microcarrier culture is the most versatile, reliable, and well-characterized procedure for unit volume scale-up of ADCs. It has had widespread use for industrial processes (Table 6), as well as many developmental uses (59), has been scaled up to 4,000 L, and has the potential for process intensification by perfusion with spin filters or by the use of microporous microcarriers.

REACTORS FOR SUSPENSION CELLS

Laboratory Scale

The basic culture unit for suspension cells is the spinner flask, so called because it has a magnetic bar operated by standing the unit on a magnetic stirrer. Initially a bar magnet was placed on the floor of the glass vessel, but spinner flasks now have the magnet suspended just above the bottom of the culture on a stirrer shaft fitted to the bottle top to prevent cells from being crushed. Side arms are usually fitted to allow gassing with CO_2/O_2 through a filter and for sampling. There are many variations in the design of spinner flasks, most affecting the size, number, and positioning of the stirrer bar (6,23). The spinner flask is usually glass, with a silicone- or Teflon-coated magnet and is available in sizes from 50 mL to 20 L, although 10 L should be considered the maximum practical size. The quality of the magnet stirrer (usually multibased to take at least four replicates) is extremely important because it must maintain a stable rpm between 30 and 300, an accurate tachometer, ability to restart after power interruptions at the set speed, a surface that does not become hot, and reliable operation for long periods in a 37 °C environment. It is advisable to siliconize the culture vessel and add medium supplements, such as Pluronic F-68 (polyglycol) (BASF, Wyandot) at 0.1% to protect against mechanical damage, especially if low serum concentrations are used. Carboxymethyl cellulose (CMC, 15–20 cps) can also be used to protect cells (0.1%) and antifoam (6 ppm) if foaming is a problem (usually when serum is present at 5–10%). Stirring speeds for suspension cells are usually within the range of 100–250 rpm (depending on cell type and vessel geometry), and a culture will expand from 1–2 $\times$ 10^5/mL to 1–2 $\times$ 10^6/mL in 4 to 5 days before needing either harvesting or a medium change. Cell clumping is often a problem but can be reduced by using low Mg^{2+} and Ca^{2+} ion media. Conventional spinner flasks are available from a wide range of laboratory suppliers. The range includes ones with large-bladed paddles attached to the magnet. They are primarily designed for microcarrier culture to give better mixing at the lower speeds necessary. Modified spinner cultures include a radial stirring action (Techne) (60), a floating impeller (Techne BR-06 Bioreactor), the Superspinner (Braun Melsungen AG), CellSpin (Tecnomara AG, Switzerland), and a dual overhead drive system to allow perfusion (Bellco). Experimental details for using spinner flasks have been published (61).

Scale-Up

Suspension culture is the preferred method for scaling up cell cultures because it is easier to volumetrically increase the size of a single fermentor vessel than the specialized units used for ADCs. Although some cell lines will not grow in suspension (e.g., human diploid cell lines WI-38 and MRC-5), many lines can be adapted to grow in free suspension by recognized procedures (62).

Several physical factors must be satisfied for successful scale-up. Good mixing is essential for homogeneity and efficient mass transfer, but during scale-up the power needed to produce these conditions can cause problems. The energy generated at the tip of the stirrer blade is a limiting factor because it gives rise to damaging shear forces (created by fluctuating liquid velocities in turbulent areas); because mixing efficiency increases with turbulence, a compromise has to be reached to minimize cell damage. Thus cell cultures use large impellers running at relatively low speeds. Magnetic bars used in spinner flasks give only radial mixing, with no lift or turbulence, and scale-up means a move to marine (not turbine) impellers. Alternative systems developed to avoid stirring are as follows:

- *Airlift fermentor.* The airlift fermentor uses the bubble column principle to both agitate and aerate a culture (21,28,63,64). Air bubbles are introduced into the bottom of a culture vessel (aspect ratio 8–12:1) and rise up an inner draft tube. Aerated medium has a lower density than nonaerated medium so the medium rises through the draft tube and circulates down the outside of the vessel (upflow and downflow approximately equal volumes). The amount of energy needed is very low and shear forces are absent; thus it is an ideal method for fragile cells. Airflow rates of about 300 mL min^{-1} are used. Vessels are available from 2 L upward, with 2,000 L reactors used in industrial manufacturing processes. However, scale-up is more or less linear, and a 90-L vessel requires 4 m headroom; special manufacturing suites are thus needed for large-scale operations.

- *Vibro-mixer.* The Vibro-mixer (65) is a nonrotating agitator that produces a stirring effect by a vertical reciprocating motion with a path of 0.1–3 mm at a frequency to 50 Hz. The mixing disk is fixed horizontally to the agitator shaft, and cone-shaped holes in the disk cause a pumping action as the shaft vibrates. This is a low energy system with reduced shear forces and foaming.
- *Celligen fermentor.* This fermentor (66,67) is available from New Brunswick Scientific and is designed so that the impeller acts as a fluid pump and aerator. Macrocirculation mixing is generated by the hollow central shaft filled with three rotating horizontal jet tubes in a low-shear bulk movement of cells and medium (cell-lift effect).

The other principal physical factor to be accommodated is aeration. Sparging can cause cellular damage unless carried out at very low rates. General guidelines are to use large bubbles (1–3 mm diameter) and low flow rates (5–10 cm^3/L/min), which means relatively inefficient mass transfer, especially at the low stirring rates used. Even though cell bioreactors have an aspect ratio of less than 2:1 to maximize surface diffusion, a means of increasing oxygen supply during scale-up is essential. Strategies for increasing the oxygen supply have included a surface aerator (68), multiple nozzle injection on the medium surface (69), caged aerator (70), spin exchange aerator (37), external and internal loop oxygenator (71), membrane tubing arranged in the vessel (19), and stirring system (20). The use of closed perfusion loops through external reservoirs is also an efficient means of oxygenation as long as the perfusion rate is fast enough (6), which requires a very efficient spin filter or other cell separation device in the culture.

Scale-up from laboratory spinner cultures to fermentors (1–10,000 L) has the following consequences:

- Change from glass to stainless steel vessels
- Change from mobile (autoclavable) to in situ system connected to steam, seed, medium holding and downstream harvesting vessels, and so forth
- Requirement for water jacket, or internal temperature control
- Greater sophistication in environmental control systems and engineering to preserve integrity

Animal cell bioreactors were adapted from the microbial fermentation tanks with only minor modifications:

- Suitable impeller (e.g., marine)
- No baffles and a curved base for better mixing at low speeds
- Water jacket rather than immersion heater type of temperature control (to avoid localized heating at low stirring speeds)
- Top-driven motor stirrer so that cells cannot become entangled between moving parts
- Mirror finish to internal surfaces to reduce mechanical damage and cell attachment

Large-Scale Bioreactors

Large-scale bioreactors are based on either the fermentor tank or the airlift principle. Stirred tanks have been operated at 8–10,000 L (14) for the production of interferon and tPA and airlift at 2,000 L principally for monoclonal antibodies. The preference is for the tank fermentor because this is a well-tried and familiar production method using conventional production plant facilities that have been refined over the past 30 years. However, most processes run at scales between 50 and 500 L. Airlift bioreactors are considered more flexible in that a greater range of cells (including extremely fastidious and fragile types) can be grown in them. However, scale-up is very dependent on tall buildings. A modification of the single-cell suspension system is to grow cells as aggregates or flocculants brought about by the addition of slightly positively charged polymers (e.g., poly-L-histidine) (72). Flocculants facilitate process operations such as filtration, centrifugation, perfusion, and sedimentation, and the microenvironment created within the cell aggregate can increase specific productivity (73).

Scale-up increases the complexity of equipment and the process and thus gives rise to greater risks of system failure, loss of integrity (sterility), and potential release of hazardous materials. Because the reason for scale-up is the production of therapeutic biological products, conditions have to comply with current good manufacturing practice (cGMP) standards acceptable to the relevant licensing authorities. Containment of the process to prevent ingress of contaminating factors and release of materials that may affect the process worker and environment are thus key factors in these processes. Microbiological containment of small-scale systems (below 20 L) in class III microbiological cabinets is possible (74), and larger-scale fermentation plants can be contained to P3 standards at least up to the 150-L scale (75).

The scale-up chain to the final production vessel is governed by the split ratio possible with the particular cell line at each subculture and is usually between 1:3 and 1:5. The chain starts with an ampoule of cells stored at -196 °C (liquid nitrogen) in a master cell bank (MCB) that has been fully characterized for biological contaminants and cell line stability (76). All manipulations are carried out according to standard operating procedures to obtain reproducibility between runs and to generate the necessary quality control data for acceptance of the final product.

The engineering complexity and considerable resource investment in money and time to build and operate a large-scale production process means that many new products are still brought to market using replicate small-scale systems. This provides for quicker market entry but requires complete reregistration of the product if the product is later moved to a large-scale production process. The experimental procedures for scaling up a stirred bioreactor process have been described (77).

HIGH-CELL-DENSITY BIOREACTORS

The alternative strategies for scale-up were outlined earlier. In this section the processes that support cells at sig-

nificantly greater densities than the classical 2–3 $\times$ 10^6/mL level are described. In the human body cells in tissues are on the order of 2–3 $\times$ 10^9 cells/cm^3, and this has always been a target for in vitro systems. However, 2 $\times$ 10^8 cells/mL has so far been the practical limit of density scale-up; even so, this is a 100-fold increase over the conventional low-density methods described so far. The two critical factors needed to achieve higher unit cell density are (*1*) perfusion to supply nutrients and remove waste products (78) and (*2*) cell immobilization to perfuse media at a fast enough rate without damaging or washing the cells out of the culture. Many devices have been developed to meet these requirements (6,7,18,22) (Table 7), of which hollow fiber bioreactors, spin filter perfusion, and microporous microcarriers are the most successful in terms of acceptance and use.

Hollow Fiber Bioreactors

Use of hollow fiber bioreactors was pioneered by Knazek et al. (95) using ultrafiltration capillary fibers. They can be considered analogous to a blood vascular system because the fibers selectively (by molecular weight cutoff) allow passage of macromolecules through the spongy fiber wall (60 μm) as the medium flows continuously through the lumen (200 μm in diameter). Cells are kept in the extracapillary space. A unit consists of thousands of fibers potted at either end in a cylindrical housing and capable of supporting both ADCs and suspension cells at 1–2 $\times$ 10^8/mL. This concept has been widely developed, with many commercial units available, and has been used particularly for producing monoclonal antibodies. The principal limiting factor in its exploitation is the difficulty in scaling up beyond very small volumes (25 mL). The problem is due to transmembrane flux being directed back into the lumen for the latter half of the flow path (6,112). Methods of overcoming this gradient problem include the insertion of contraflow and aeration fibers (112), cyclical flushing of medium between the lumen and extracapillary compartment (113), use of flat-bed rather than cylindrical filtration cartridges (91), and use of a radial flow reactor (114). The concept of using pressure differentials to simulate in vivo arterial and venous flow is used in the Acusyst System (Cellex Biosciences, Inc.). This system is well documented (113), including experimental details (115), and is probably the most used and successful application of hollow fiber technology.

To allow ADCs to attach, special fibers can be used (e.g., Amicon 3S1000) to give a huge surface area (1-L cartridge has 1 m^2 surface area and supports 10^{10} suspension cells or 2 $\times$ 10^9 ADCs). Thus this method gives a very large surface-area-to-volume ratio, allows continuous removal of waste products and supply of nutrients, supports high cell densities with a tissuelike architecture, and allows a concentrated product to be harvested. The disadvantages are diffusional limitations that cause culture inhomogeneity and cell necrosis, process control complexity, and difficulty in sterilization.

Table 7. Devices Used for Immobilizing Animal Cells to Permit Medium Perfusion

Device	Reference
Gravitational settling of cells in medium outflow	79–81
Spin filter	5,66,69,81–85
Centrifuge circulation loop	86,87
Filtration membrane	
Dialysis	88,90
Hollow fibers	91–95
Membranes	96,97
Encapsulation	98–100
Entrapment	
Textured ceramics	101
Tubular flow reactor	102
Cellulose or polyurethane sponges	103,104
Fibers	105,106
Polyester (Fibra-cel)	107
Microporous microcarriers	46,47,108–111

Spin Filters

Because all static filters within a culture eventually become blocked, the development of a rotating filter that creates a boundary effect, thus delaying cell attachment and filter clogging, has been a successful step forward. The principle was introduced by Himmelfarb in 1969 (82) using a 3-μm stainless steel mesh rotated at 300 rpm. Subsequently there have been many variations (reviewed in Ref. 6): polymers (84), ceramics, microfibers, porcelain, immersed cages, baskets semiimmersed in the culture or attached to the stirrer shaft, separate rotating shaft (to increase rpm), filter within a draft tube and so forth. They provide a simple technical solution to scaling up process intensity within well-established stirred fermentors. They are limited in that clogging eventually occurs and in the perfusion rate that can be maintained. However, they are particularly useful for microcarriers because a far larger mesh can be used, thus reducing the onset of filter clogging and permitting cell densities of 2 $\times$ 10^7/mL to be attained.

Microporous Microcarriers

Microcarrier culture has proven to be the most effective scale-up method for ADCs, despite its limitations (critical procedures, low surface-area-to-volume ratio of a sphere). Attempts were made to increase the surface area by surface indenting techniques but were not successful until the fabrication of porous particles. The initial particle was the Verax microsphere (47,108), which was 500 μm in diameter and manufactured from bovine collagen; the interconning channels of 20–40 μm diameter gave an internal open volume of 80% of the sphere. They were designed for fluidized-bed reactors, and thus their specific gravity was artificially increased to 1.7 with the addition of small metallic weights incorporated into the matrix. The spheres were fluidized at 75 cm/min upward flow, and cell densities in excess of 10^8 per milliliter intrasphere volume were achieved (equivalent to 4 $\times$ 10^7/mL in the bioreactor). The sphere matrix provided a huge surface area for attachment of ADCs but was equally suitable for suspension cells (which were entrapped in the pores). The system was scaled up to 24 L routinely and 200 L in the company's production unit and used for more than 85 different cell lines producing many

Table 8. Advantages of Microporous Microcarriers

Unit cell density 50- to 100-fold higher than that of free suspension
Suitable for both attached and suspended cells
Can be used in fluidized- and fixed-bed reactors and stirred suspension
In situ 100- to 250-fold seed expansion eliminates seed chain steps
Diffusion into a sphere efficient (30% diameter penetration = 70% of volume)
Cells protected from shear
Specific cell productivity increased three- to five-fold
Three-dimensional structure easily derivitized for specialized cells
Capable of long-term (>100 days) culture with continuous harvesting
Scale-up potential by comparison with analogous systems (microcarrier to 4,000 L)

cell products (including tPA, pro-urokinase [proUK], EPO, interferon, interleukin, factor VIII, and various immunoglobulins).

Verax was initially a closed product in that you had to invest in the complete system. Thus other microporous beads were quickly developed (Table 5). Alternative systems are also available such as the Pharmacia Cytopilot using Cytoline porous microcarriers (110). Microporous microcarriers that could be used in stirred rather than fluidized bioreactors are preferred and now available (Cellsnow and ImmobaSil). ImmobaSil (116) is of particular value because it is extremely permeable to oxygen, a robust, nonanimal product safe from bovine contaminants that can be used in all culture modes. Because the concentration of spheres is far less, the total cell mass is greatly reduced in stirred systems. However, it allows manufacturers to use conventional stirred equipment and increases cell density in the microcarrier system from 2×10^6/mL to in excess of 5×10^7/mL (using spin filter perfusion and aeration).

An alternative to using stirred or fluidized microporous carriers is use of a fixed bed of porous glass spheres (e.g., Siran) (46,109). Fixed-bed reactors of solid glass spheres have many advantages (111), and the disadvantage of low cell density can be overcome by using 5-mm Siran spheres. These spheres have surface areas of 75 m^2/L with interconning pores and channels of 60–300 μm and can be stacked in 5-L beds and perfused at 5 linear cm/min to provide a high feed rate without washing out or damaging the cells (protected from shear within the particles). The versatility and usefulness of this technique has been demonstrated for a range of suspension and anchorage-dependent cells, resulting in a 10-fold higher productivity over equivalent systems (18,117–119). A characteristic of all microporous carrier systems is that cell specific productivity is always higher (18,117), presumably due to the favorable environment of cells packed together in almost tissuelike density. Details on experimental methods have been published; fluidized-bed (120), fixed-bed (121), and commercial fixed-bed reactors are available (Meredos GmbH, Bovenden).

The advantages of porous carrier culture are summarized in Table 8. This technology is truly universal in being

Table 9. Necessary Conditions for the Acceptance of Scaleable Immobilized Bioreactors

Insufficient quantities produced by other methods
Product only produced by immobilized or specialized cells that need three-dimensional culture
End product price only attainable with scale-up economies possible with these systems
Increased production plant capacity when expansion is no longer possible
New company or product beginning registration process
Sufficient data and experience accumulated to give manufacturing confidence

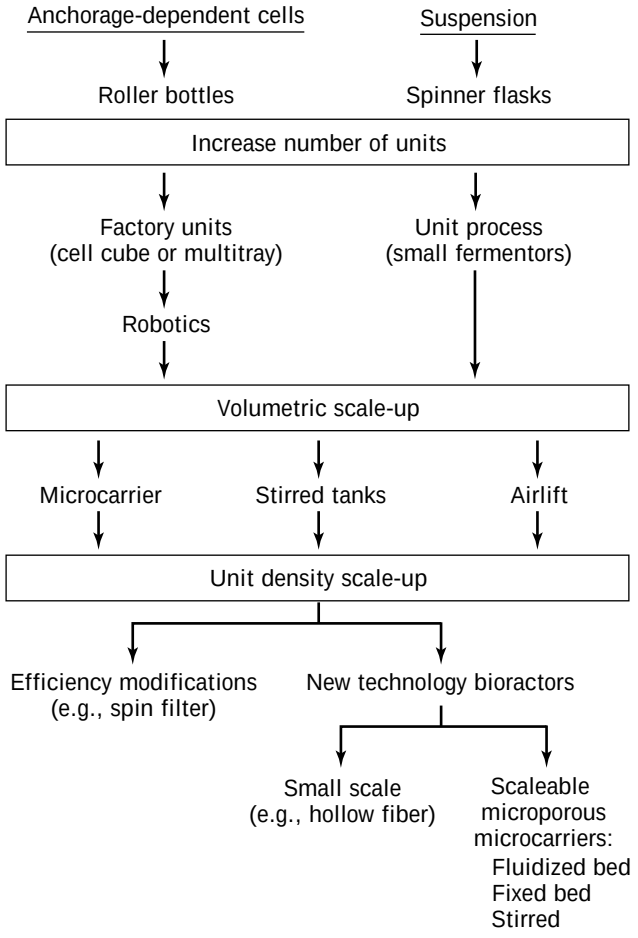


Figure 2. Scale-up: the decision process.

equally suitable for suspension cells and ADCs and flexible in the range of bioreactor systems it can be in. It is the only truly scaleable high-cell-density system available and provides a microenvironment that stimulates cell product expression. As the need for high-quantity biopharmaceutical manufacture increases, this technology will become increasingly widespread in application and use (Table 9).

COMPARISONS, CONCLUSIONS, AND FUTURE DEVELOPMENTS

Given the range of culture strategies and options mentioned in this article, how does one choose the system best

Scale (L)	System	Area (cm^2) (batch size)	Cells per batch (cells/L)
	Roller	1,750 (× 100)	2×10^{10} (2×10^{8})
	Multitray	24,000 (× 10)	3×10^{10} (3×10^{9})
	Cell cube	85,000 (× 4)	4×10^{10} (9×10^{9})
1	Hollow fiber	20,000	10^{10} (10^{11})
20	Microporous microcarrier	2.5×10^{6}	10^{12} (5×10^{10})
100	Glass sphere	10^{6}	10^{11} (10^{9})
200	Stacked plate	2×10^{5}	2×10^{10} (1×10^{8})
500	Microcarrier (with spin filter)	2×10^{10}	10^{13} (5×10^{10})
2,000	Airlift		4×10^{12} (2×10^{9})
4,000	Microcarrier	5×10^{6}	8×10^{12} (2×10^{9})
10,000	Stirred tank		2×10^{13} (2×10^{9})

Figure 3. Comparison of culture systems indicating volumetric size, unit, and system cell yields.

suited to the problem at hand? In Figure 2, a flow sheet of the options is given by which a product manager can reach such a decision based on the cell line, the product type, the time table, the available plant and resources, the manpower, and the annual product requirement. The question of whether to take the high-volume or high-density option is based on the considerations summarized in Tables 3 and 8 and the range of total cell densities that various bioreactors will support (Fig. 3). The system chosen will be the one that fulfills the defined need at the minimum commercial risk, because there is no one ideal system to meet all needs.

The future of animal cell biotechnology as a production vehicle for therapeutic products is as an essential component in a range of biological processes. Recombinant microorganisms will be the first method of choice for simple, nonglycosylated products without complex three-dimensional configurations necessary for bioactivity. Transgenic production may take over for products needed in 50 to 100-kg quantities. However, society may demand that these products be produced even at greater considerable cost in culture systems to reduce the exploitation of animals; this will be the driving factor for high-volume high-density systems based on microporous microcarriers. Finally, engineering solutions (122) to scale-up are not the only means of increasing unit productivity and economy of cell manufacturing processes. Considerable advances are being made in increasing cell-specific productivity from a modest 0.2 mg/10^6 cells/day to nearer in vivo levels of 7 mg/10^6 cells/day by cell engineering (recombinant control of metabolic processes). Advances are also being made in development of media and the environment to regulate cell growth, metabolism, and product expression in simpler (cheaper) media. There are also considerable improvements to be made in downstream scale and efficiency, where often only 20% of the product survives the purification process.

BIBLIOGRAPHY

1. R.G. Harrison, *Proc. Soc. Exp. Biol. Med.* **4**, 140–143 (1907).
2. J.F. Enders, T.H. Weller, and F.C. Robbins, *Science* **109**, 85–87 (1949).
3. P.J. Radlett, T.W.F. Pay, and A.J.M. Garland, *Dev. Biol. Stand.* **60**, 163–170 (1985).
4. L. Hayflick and P.S. Moorhead, *Exp. Cell Res.* **25**, 585–621 (1961).
5. J.P. Jacobs, *Nature (London)* **227**, 168–170 (1970).
6. J.B. Griffiths, in R.E. Speir and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 3, Academic Press, London, 1988, pp. 179–220.
7. A. Prokop and M.Z. Rosenberg, *Advances in Biochemical Engineering / Biotechnology* **39**, 29–71 (1989).
8. G.F. Panina, *Animal Cell Biotechnology*, vol. 1, Academic Press, London, 1985, pp. 211–242.
9. A.L. van Wezel, *Nature (London)* **216**, 64–65 (1967).
10. K.F. Pullen, M.D. Johnston, A.W. Phillips, G.D. Ball, and N.B. Finter, *Dev. Biol. Stand.* **60**, 175–178 (1985).
11. M. Morandi and A. Valeri, *Advances in Biochemical Engineering / Biotechnology* **37**, 57–72 (1988).
12. B. Clough, in A. Doyle and J.B. Griffiths eds., *Cell and Tissue Culture: Laboratory Procedures for Biotechnologists*, Wiley, Chichester, U.K., 1998, pp. 5–17.
13. G. Kohler and C. Milstein, *Nature (London)* **256**, 495–497 (1975).
14. A. Lubiniecki, R. Arathoon, G. Polastri, J. Thomas, M. Wiebe, R. Garnick, A. Jones, R. van Reis, and S. Builder, in R.E. Speir, J.B. Griffiths, J. Stephenne, and P.J. Crooy eds., *Animal Cell Biology and Technology for Bioprocesses*, Butterworths, Guildford, U.K., 1989, pp. 442–451.
15. R.P. Lanza, J.L. Hayes, and W.L. Chick, *Nature Biotechnology* **14**, 1107–1110 (1996).
16. M.W. Glacken, R.J. Fleischaker, and A.J. Sinskey, *Trends Biotechnol.* **1**, 102–108 (1983).

17. J.B. Griffiths, *Cytotechnology* **3**, 109–116 (1990).
18. J.B. Griffiths, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 4, Academic Press, London, 1990, pp. 149–166.
19. W. Kuhlmann, *Dev. Biol. Stand.* **66**, 263–268 (1987).
20. J. Lehmann, G.W. Piehl, and R. Schulz, *Dev. Biol. Stand.* **66**, 227–240 (1987).
21. J.C. Merchuk and M.H. Siegal, *J. Chem. Technol. Biotechnol.* **41**, 105–120 (1988).
22. J.B. Griffiths, *Biotechnol.* **10**, 30–32 (1991).
23. J.B. Griffiths, in R.I. Freshney ed., *Animal Cell Culture: A Practical Approach*, IRL Press, Oxford, 1992, pp. 47–93.
24. J.B. Griffiths, *Ann. N.Y. Acad. Sci.* **665**, 364–370 (1992).
25. D. Looby and J.B. Griffiths, *Trends Biotechnol.* **8**, 204–209 (1990).
26. J.B. Griffiths, *J. Biotechnol.* **22**, 21–30 (1992).
27. R. Kunitake, A. Suzuki, H. Ichihashi, S. Matsuda, O. Hirai, and K. Morimto, *J. Biotechnol.* **52**, 289–294 (1997).
28. H.G. MacMorine, G.D. Laurence, W. Parisius, and N.B. Cucakovich, *Dev. Biol. Stand.* **37**, 139–142 (1977).
29. W. House, M. Shearer, and N.G. Maroudas, *Exp. Cell Res.* **71**, 293–296 (1972).
30. M. Corbeil, M. Trundel, and P. Payment, *J. Clin. Microbiol.* **10**, 91–95 (1979).
31. H.C. Girard, M. Sutcu, H. Erdem, and I. Gurhan, *Biotechnol. Bioeng.* **22**, 477–493 (1980).
32. P.F. Kruse, L.N. Keen, and W.L. Whittle, *In Vitro* **6**, 75–81 (1970).
33. I.A. Beeksma and R. Kompier, in E.C. Beuvery, J.B. Griffiths, and W.P. Zeijlemaker eds., *Animal Cell Technology: Developments towards the 21st Century*, Kluwer Academic, Dordrecht, 1995, pp. 661–663.
34. J.B. Schleicher and R.E. Weiss, *Biotechnol. Bioeng.* **10**, 617–624 (1968).
35. O. Molin and C.G. Heden, *Prog. Immunobiol. Stand.* **3**, 106–110 (1969).
36. C. Burbidge and I.K. Dacey, *Dev. Biol. Stand.* **55**, 255–259 (1984).
37. J.B. Griffiths, D.R. Cameron, and D. Looby, *Dev. Biol. Stand.* **66**, 331–338 (1987).
38. C. Burbidge, *Dev. Biol. Stand.* **46**, 169–172 (1980).
39. D. Looby and J.B. Griffiths, in R.E. Spier and J.B. Griffiths eds., *Modern Approaches to Animal Cell Technology*, Butterworths, Guildford, U.K., 1987, pp. 342–352.
40. O.T. Ramirez and R. Mutharasan, *Biotechnol. Bioeng.* **33**, 1072–1076 (1989).
41. J.H. Robinson, P.M. Butlin, and R.C. Imrie, *Dev. Biol. Stand.* **46**, 173–181 (1980).
42. A.L. van Wezel, *Dev. Biol. Stand.* **55**, 3–9 (1984).
43. B. Montagnon, J.C. Vincent-Falquet, and B. Fanget, *Dev. Biol. Stand.* **55**, 37–42 (1984).
44. K. Nilsson, F. Buzsaky, and K. Mosbach, *Biotechnology* **4**, 989–990 (1986).
45. D. Looby and J.B. Griffiths, *Trends Biotechnol.* **8**, 204–209 (1990).
46. D. Looby and J.B. Griffiths, *Cytotechnology* **1**, 339–346 (1988).
47. P.W. Runstadler and S.R. Cernek, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 3, Academic Press, London, 1988, pp. 305–320.
48. W. Noe, R. Bux, W. Berthold, and W. Werz, *Cytotechnology* **15**, 169–176 (1994).
49. J.P. Whiteside and R.E. Spier, *Biotechnol. Bioeng.* **23**, 551–556 (1981).
50. Y. Chisti and M. Moo-Young, *Trans. Inst. Chem. Eng.* **72C**, 91–94 (1994).
51. P.C. Brown, C. Figueroa, M.A. Costello, and S.M. Maciukas, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 3, Academic Press, London, 1988, pp. 251–262.
52. D. Looby, in A. Doyle, J.B. Griffiths, and D. Newell eds., *Cell and Tissue Culture: Laboratory Procedures*, Wiley, Chichester, U.K., 1993, p. 28D:2.
53. B. Meignier, H. Mougeot, and H. Favre, *Dev. Biol. Stand.* **46**, 249–256 (1980).
54. J.B. Griffiths, in A. Doyle, J.B. Griffiths, and D. Newell eds., *Cell and Tissue Culture: Laboratory Procedures*, Wiley, Chichester, 1993, p. 28C:1.1.
55. A.L. van Wezel, *Dev. Biol. Stand.* **46**, 151–158 (1980).
56. J.S. Friedman, *Biotechnology* **7**, 359–364 (1989).
57. K.K. Sanford, W.R. Earle, and V.J. Evans, *J. Natl. Cancer Inst.* **11**, 773–795 (1951).
58. C.L. Crespi and W.G. Thilly, *Biotechnol. Bioeng.* **23**, 983–993 (1981).
59. B.H. Junker, F. Wu, S. Wang, J. Waterbury, G. Hunt, J. Hennessey, J. Aunins, J. Lewis, M. Silberklang, and B.C. Buckland, *Cytotechnology* **9**, 173–188 (1992).
60. N.A. de Bruyne, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 3, Academic Press, London, 1988, pp. 142–176.
61. J.B. Griffiths, in A. Doyle, J.B. Griffiths, and D. Newell eds., *Cell and Tissue Culture, Laboratory Procedures*, Wiley, Chichester, U.K., 1993, p. 28A:2.3.
62. P.B. Capstick, A.J. Garland, R.C. Masters, and W.G. Chapman, *Exp. Cell Res.* **44**, 119–129 (1963).
63. H.W.D. Katinger, W. Scheirer, and E. Kromer, *Ger. Chem. Eng.* **2**, 31–38 (1979).
64. M. Rhodes, S. Gardner, and D. Broad, in C.S. Ho and D.I.C. Wang eds., *Animal Cell Bioreactors*, Butterworth-Heinemann, Boston, Mass., 1991, pp. 253–268.
65. H.C. Girard, G. Okay, and Y. Kilvilcin, *Bulletin of the Office of International Epizoology* **79**, 805–822 (1973).
66. S. Reuveny, Z.-B. Zheng, and L. Epstein, *Am. Biotechnol. Lab.* 28–36 (January/February 1986).
67. M. Johnson, G.A.C. Chavrie, and H. Archambault, *Biotechnol. Bioeng.* **35**, 43–49 (1990).
68. W.-S. Hu and D.I.C. Wang, in W.G. Thilly ed., *Mammalian Cell Technology*, Butterworths, Boston, Mass., 1986, pp. 167–197.
69. A.L. van Wezel, *J. Chem. Technol. Biotechnol.* **32**, 318–323 (1982).
70. J.P. Whiteside, S. Farmer, and R.E. Spier, *Dev. Biol. Stand.* **60**, 283–290 (1985).
71. C. Leist, H.P. Meyer, and A. Fiechter, *J. Biotechnol.* **4**, 235–246 (1986).
72. J.G. Aunins and D.I.C. Wang, *Biotechnol. Bioeng.* **34**, 629–638 (1989).
73. A. Racher, J.L. Moreira, P.M. Alves, M. Wirth, U.H. Weidle, H. Hauser, M.J.T. Carrondo, and J.B. Griffiths, *Appl. Microbiol. Biotechnol.* **40**, 851–856 (1994).
74. P. Hambleton and J. Melling, in P. Hambleton, J. Melling, and T.T. Salusbury eds., *Biosafety in Industrial Biotechnology*, Blackie Academic, London, 1994, pp. 129–148.
75. P. Hambleton, J.B. Griffiths, D.R. Cameron, and J. Melling, *J. Chem. Technol. Biotechnol.* **50**, 167–186 (1991).

76. A. Doyle and J.B. Griffiths, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 5, Academic Press, London, 1992, pp. 75–96.
77. A.B. Lang and U. Schuerch, in A. Doyle, J.B. Griffiths, and D. Newell eds., *Cell and Tissue Culture, Laboratory Procedures*, Wiley, Chichester, U.K., 1993, p. 28B:1.1.
78. J.B. Griffiths, D. Looby, and A.J. Racher, *Cytotechnology* **9**, 3–9 (1992).
79. S. Sato K. Kawamura, and N. Fujiyoshi, *J. Tissue Culture Methods* **8**, 167–171 (1983).
80. M. Butler, T. Inamura, J. Thomas, and W.G. Thilly, *J. Cell Sci.* **61**, 351–363 (1983).
81. J. Feder and W.R. Tolbert, *Sci. Am.* **248**, 24–31 (1983).
82. P. Himmelfarb, P.S. Thayer, and H.E. Martin, *Science* **64**, 555–557 (1969).
83. R. Varecka and W. Scheirer, *Dev. Biol. Stand.* **66**, 269–272 (1987).
84. G.C. Avgerinos, D. Drapeau, J.S. Socolow, J. Mao, K. Hsiao, and R.J. Broeze, *Biotechnology* **8**, 54–58 (1990).
85. E. Favre and T. Thaler, *Cytotechnology* **9**, 11–19 (1992).
86. V. Jager, in R.E. Spier, J.B. Griffiths, and C. MacDonald eds., *Animal Cell Biotechnology: Developments, Processes, and Products*, Butterworth-Heinemann, Oxford, 1992, pp. 397–402.
87. K.O. Gronvik, H. Frieburg, and U. Malmstran, in R.E. Spier, J.B. Griffiths, J. Stephenne, and P.J. Crooy eds., *Advances in Animal Cell Biology and Technology for Bioprocesses*, Butterworths, Guildford, U.K., 1989, pp. 336–344.
88. G.B. Gori, *Appl. Microbiol.* **13**, 93–97 (1965).
89. A. Bohmarn, R. Portner, J. Schmieding, V. Kasche, and H. Markl, *Cytotechnology* **9**, 51–57 (1992).
90. H. Buntemeyer, C. Bohme, and J. Lehmann, *Cytotechnology* **15**, 243–251 (1994).
91. K. Ku, M.J. Kuo, J. Delente, B.S. Wildi, and J. Feder, *Biotechnol. Bioeng.* **23**, 79–95 (1981).
92. G. Klement, W. Scheirer, and H.W.D. Katinger, *Dev. Biol. Stand.* **66**, 221–226 (1987).
93. G.L. Altshuler, D.M. Dziewulski, J.A. Sowek, and G. Belfort, *Biotechnol. Bioeng.* **28**, 646–658 (1986).
94. P. Knight, *Biotechnology* **7**, 459–461 (1989).
95. A. Knazek, P.M. Gallino, P.O. Kohler, and R.L. Dendrick, *Science* **178**, 65–67 (1972).
96. F.W. Falkenberg, H. Weichert, M. Krane, I. Bartels, M. Palme, H.-O. Nagels, and H. Fiebig, *J. Immunol. Methods* **179**, 13–29 (1995).
97. Tecnomouse, Integra Biosciences, AG, CH-8304 Wallisellen.
98. K. Nilsson, *Trends Biotechnol.* **5**, 73–78 (1987).
99. R.G. Duff, *Trends Biotechnol.* **3**, 167–170 (1985).
100. A. Tinto, P. Vila, C. Casca, and F. Godia, in M.J.T. Carrondo, J.B. Griffiths, and J.L.P. Moreira eds., *Animal Cell Biotechnology: From Vaccines to Genetic Medicine*, Kluwer Academic, Dordrecht, 1997, pp. 423–428.
101. G.J. Berg, *Dev. Biol. Stand.* **60**, 297–307 (1985).
102. H.W.D. Katinger, *Dev. Biol. Stand.* **66**, 195–209 (1987).
103. A. Lazar, L. Silberstein, A. Mizrahi, and S. Reuveny, *Cytotechnology* **1**, 333–338 (1988).
104. H.A. Phillips, J.M. Scharer, N.C. Bols, and H. Moo-Young, *Cytotechnology* **9**, 29–40 (1992).
105. H. Yamaji, H. Fukuada, Y. Nojima, and C. Webb, *Appl. Microbiol. Biotechnol.* **30**, 609–613 (1989).
106. J. Litwin, *Dev. Biol. Stand.* **60**, 237–242 (1985).
107. A. Kadouri, G. Wang, W. Zhang, C. Jacklin, and D. Friedman, in M.J.T. Carrondo, J.B. Griffiths, and J.L.P. Moreira eds., *Animal Cell Biotechnology: From Vaccines to Genetic Medicine*, Kluwer Academic, Dordrecht, 1997, pp. 307–312.
108. M.W. Young and R.C. Dean, *Biotechnology* **5**, 835–837 (1987).
109. S. Pak and G. Stephanopoulas, *Biotechnol. Bioeng.* **41**, 25–34 (1993).
110. G. Klima, G. Kreismayr, D. Muller, C. Schmatz, S. Wiederkum, W. Steinfellner, A. Assadian, G. Lhota, G. Bluml, O. Doblhoff-Dier, and H. Katinger, in M.J.T. Carrondo, J.B. Griffiths, and J.L.P. Moreira eds., *Animal Cell Technology: From Vaccines to Genetic Medicines*, Kluwer Academic, Dordrecht, 1997, pp. 447–454.
111. R.E. Spier, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 1, Academic Press, London, 1985, pp. 243–263.
112. J.P. Tharakan and P.C. Chau, *Biotechnol. Bioeng.* **28**, 1064–1071 (1986).
113. M.L. Wolf and M.D. Hirschel, in R.E. Spier and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 6, Academic Press, London, 1994, pp. 237–258.
114. J.P. Tharakan and P.C. Chau, *Biotechnol. Bioeng.* **29**, 657–671 (1987).
115. J.A.J. Hanak and J.M. Davis, in A. Doyle, J.B. Griffiths, and D. Newell eds., *Cell and Tissue Culture: Laboratory Procedures*, Wiley, Chichester, U.K. 1993, pp. 28D:3.1–16.
116. D. Looby, A.J. Racher, J.P. Fuller, and J.B. Griffiths, in E.C. Beuvery, J.B. Griffiths, and W.P. Zeijlemaker eds., *Animal Cell Technology: Developments Towards the 21st Century*, Kluwer Academic, Dordrecht, 1995, pp. 783–786.
117. J.B. Griffiths and A.J. Racher, *Cytotechnology* **15**, 3–9 (1994).
118. A.J. Racher and J.B. Griffiths, *Cytotechnology* **13**, 125–131 (1993).
119. A.J. Racher, D. Looby, and J.B. Griffiths, *J. Biotechnol.* **15**, 129–146 (1990).
120. D. Looby, in A. Doyle, J.B. Griffiths, and D. Newell eds., *Cell and Tissue Culture: Laboratory Procedures*, Wiley, Chichester, U.K., 1993, pp. 28D:1.1–16.
121. D. Looby, in A. Doyle, J.B. Griffiths, and D. Newell eds., *Cell and Tissue Culture: Laboratory Procedures*, Wiley, Chichester, U.K., 1993, pp. 28D:2.1–13.
122. S. Ozturk, *Cytotechnology* **22**, 3–16 (1996).

MASS TRANSFER

YUSUF CHISTI
University of Almería
Almería, Spain

KEY WORDS

Bioreactors
Diffusion coefficient
Fluidized beds
Gas–liquid mass transfer
Liquid–liquid mass transfer
Oxygen
Packed beds
Reynolds number
Solid–liquid mass transfer

OUTLINE

INTRODUCTION

Mass transfer phenomena impact upon all facets of bioprocessing. Transport effects often determine the productivity of bioreactors and the downstream recovery operations. Gas–liquid mass transfer problems arise during supply of oxygen from a gas phase to liquid culture and during removal of metabolically generated carbon dioxide from a culture fluid to a gas phase. Similarly, gas–liquid mass transfer issues must be faced again during recovery operations such as distillation. Liquid–liquid mass transfer occurs when oxygen is supplied through liquid carriers such as perfluorocarbons, during liquid–liquid extraction, and during degradation of water immiscible liquid substrates. Solid–liquid mass transfer problems are faced during recovery by adsorption, during chromatographic separations, and in operations such as crystallization. Performance of solid–phase biocatalysts such as immobilized cells and enzymes is often limited by solid–liquid mass transfer. Solid–liquid mass transfer effects influence the workings of membrane separations such as microfiltration and ultrafiltration. Transport within solid particles or intraparticle mass transfer becomes limiting in certain cases. Gas–solid transport is seen during some drying situations. Some of these transport issues are the focus of this article.

Ultimately, the transport of a solute through any fluid or space is governed by the molecular diffusivity or the diffusion coefficient of the solute in the fluid or solution. How diffusivity is affected by factors such as temperature and viscosity and the estimation of diffusivities in various situations are discussed next.

DIFFUSION COEFFICIENT OR DIFFUSIVITY

For diffusion of a solute through a solvent, the diffusion coefficient depends on temperature, the type of solvent and its viscosity, and the concentration of solute in solution. Diffusivities in liquids and gases generally increase with temperature. Liquid-phase diffusivities are little affected by pressure, but in gases, diffusivities decline as pressure increases. Typically, liquid-phase diffusivities are much lower—only about a hundredth—compared to those in gases. Diffusion coefficients of common solute–solvent combinations are available in handbooks. Tables 1 and 2 provide some bioprocess-relevant data. Methods of measuring diffusion coefficients have been described by Geankoplis (1) and Pratt (2). When diffusion coefficients are not available, they can be estimated.

Table 1. Diffusivities of Some Common Solutes in Dilute Liquids

Solute	Solvent	Temperature (°C)	D_L ($\times 10^9$ m^2 s^{-1})
Acetic acid	Water	25	1.26
Acetic acid	Benzene	25	2.09
Acetone	Water	25	1.28
Benzoic acid	Water	25	1.21
Carbon dioxide	Water	20	1.50
Carbon dioxide	Water	25	2.00
Ethanol	Water	25	1.24
Glucose	Water	20	0.60
Glycerol	Water	20	0.83
Methanol	Water	15	1.26
Nitrogen	Water	20	1.64
Oxygen	Water	20	1.80
Oxygen	Water	25	2.41
Phenol	Water	20	0.84
Sucrose	Water	20	0.45
Urea	Water	25	1.38
Water	Ethanol	25	1.13

Table 2. Diffusivities of Some Globular Proteins in Dilute Aqueous Solutions

Protein	Molecular weight (kg kmol^{-1})	Temperature (°C)	D_L ($\times 10^{11}$ m^2 s^{-1})
Bovine serum albumin	67 500	25	6.81
Urease	482 700	25	4.01
Human serum albumin	72 300	20	5.93
Human fibrinogen	339 700	20	1.98

Diffusivities of small solutes diffusing in dilute liquids can be estimated using the Wilke-Chang (3) equation

$$D_L = 1.173 \times 10^{-16} \frac{(\chi M_L)^{0.5} T}{\mu_L V_M^{0.6}} \qquad (1)$$

where M_L and μ_L are the molecular weight and the viscosity of the solvent, respectively; T is the absolute temperature, V_M is the molar volume of the solute at its boiling point, and χ is the association parameter, a measure of polar interactions among molecules, of the solvent. Association parameters of some common solvents are noted in Table 3. The molar volumes are given in handbooks, or, for organic solutes, they are easily calculated by the group contribution method (1,2). Equation 1 is suitable for small solutes with $V_M \leq 0.500$ m^3 kmol^{-1}. When molar volumes are larger, as with globular proteins, the Stokes-Einstein equation may be used to estimate diffusivities; thus

Table 3. Association Parameters for Some Solvents

Solvent	Association parameter, ξ
Water	2.6
Methanol	1.9
Ethanol	1.5
Benzene	1.0
Ether	1.0
Heptane	1.0
Other unassociated solvents	1.0

Source: Based on Wilke and Chang (3).

$$D_L = \frac{9.96 \times 10^{-16} T}{\mu_L V_M^{1/3}} \quad (2)$$

Larger molecules diffuse more slowly than smaller ones. For example, the diffusivities of globular proteins (Table 2) are much lower than those of smaller solutes listed in Table 1.

Small solutes such as sugar or oxygen diffusing in solutions of proteins move more slowly than in water. Diffusion coefficient D_P of a small solute in a solution of globular protein can be approximated (1) as

$$D_P = D_L(1 - 1.81 \times 10^{-3} C) \quad (3)$$

where D_L is the diffusivity of the solute in water and C (kg m^{-3}) is the concentration of protein in solution. Equation 3 is suitable only when the diffusing solute does not associate with the protein molecule. When a solute associates with the protein, a part of the observed diffusivity of solute is due to diffusion of the protein molecules and a part is due to diffusion of unassociated solute.

Small molecules diffusing through gels such as agar, gelatin, and alginate diffuse more slowly than in water. The diffusion coefficient declines with increasing concentration of the gelling agent; however, in typically formulated gels, diffusivities are only 20 to 50% lower than in water. Inorganic salts and electrolytes that dissociate in water actually diffuse as molecules in the absence of electric fields. Diffusivities of electrolytes may be estimated using methods discussed by Pratt (2).

GAS–LIQUID MASS TRANSFER

Many industrial fermentations require oxygen. Oxygen supply must at least equal the oxygen demand or the fermentation will be oxygen limited. Oxygen demand is especially difficult to meet in viscous fermentation broths and in broths containing a large concentration of the oxygen-consuming cells. In large fermenters, oxygen transfer becomes difficult when demand exceeds 4 to 5 kg m^{-3} h^{-1} (4). Above a critical concentration of dissolved oxygen, the amount of oxygen does not influence the specific growth rate. The critical dissolved oxygen level depends on the microorganism, the culture temperature, and the substrate being oxidized. The higher the critical dissolved oxygen value, the greater the likelihood that oxygen transfer will become limiting (4). Under typically used culture conditions, fungi such as *Penicillium chrysogenum* and *Aspergillus oryzae* have a critical dissolved oxygen value of about 3.2×10^{-4} kg m^{-3}. For baker's yeast and *Escherichia coli*, the critical dissolved oxygen values are 6.4×10^{-5} and 12.8×10^{-5} kg m^{-3}, respectively (4). For comparison, the dissolved oxygen concentration in an air-saturated sterile medium is typically around 7×10^{-3} kg m^{-3}; hence, in terms of percent of air saturation, the minimum acceptable dissolved oxygen level in a fungal fermentation is about 5%.

Oxygen uptake rate of mammalian cells is typically $(0.8 - 8.0) \times 10^{-15}$ kg O_2 $cell^{-1}$ h^{-1}. In suspension culture of animal cells, cell concentrations typically reach $(1 - 1.2) \times 10^6$ cells mL^{-1}; concentrations up to 1.2×10^7 cells mL^{-1} are attained in perfusion culture, and even higher values occur in immobilized cell bioreactors. The critical dissolved oxygen concentration for animal cells tends to be low, for example, 0.5% of air saturation or 0.035 kg m^{-3}.

The transport route of a soluble gas from the gas to the liquid phase is schematically illustrated in Figure 1. The gas may be imagined as passing through stagnant gas and liquid films on either side of the gas–liquid interface. The resistance to transfer is localized within the stagnant films and the gas–liquid interface itself is assumed to offer no resistance to mass transfer. Mass transfer in the films occurs solely by diffusion; therefore, at steady state, linear concentration profiles exist in the films (Fig. 1). The transport flux (J) of the diffusing species is related to the concentration gradient (ΔC) in the film and to the film thickness (δ) as follows:

$$J = \frac{D}{\delta} \Delta C \quad (4)$$

In equation 4, D is the diffusivity of the transferring component in the film. The ratio D/δ is known as the mass transfer coefficient, k.

At steady state, the flux of the diffusing component (or the rate of transfer per unit cross-sectional area) is the same through the gas and the liquid films; thus, equation 4 may be written for each of the two films:

$$J = k_G(C_G - C_{Gi}) \quad (5)$$

$$= k_L(C_{Li} - C_L) \quad (6)$$

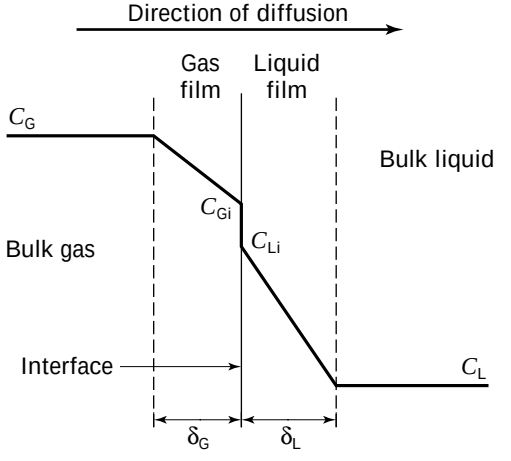


Figure 1. Steady-state dissolved oxygen concentration profile in the vicinity of the gas–liquid interface.

where k_G and k_L are the gas and the liquid film mass transfer coefficients, respectively. The overall mass flux from the gas to the liquid phase may be written as follows:

$$J = K_L(C^* - C_L) \tag{7}$$

where C^* is the saturation concentration (i.e., the maximum possible value) of the diffusing component in the liquid in equilibrium with the gas phase, and K_L is the overall mass transfer coefficient based on the liquid film. The saturation concentration C^* in the liquid is related to the gas phase concentration (C_G) of the diffusing component by Henry's law:

$$C_G = HC^* \tag{8}$$

where H is the dimensionless Henry's constant (5). Using equations 5 to 8 and the knowledge that $C_{Gi} = H \times C_{Li}$ (i.e., equilibrium exists at the interface because the interface offers no resistance to mass transfer), the various mass transfer coefficients can be shown to be related as follows (5):

$$\frac{1}{K_L} = \frac{1}{k_L} + \frac{1}{Hk_G} \tag{9}$$

For a sparingly soluble gas such as oxygen, H is much greater than unity ($H \approx 30$ for oxygen in water); furthermore, the coefficient k_G is typically larger than k_L because the gas–phase diffusivities are vastly greater than those in liquids (e.g., $D_{oxygen/air} = 10^4\, D_{oxygen/water}$ at 20 °C). Consequently, the second term on the right side of equation 9 is negligible relative to the first, and the equation becomes

$$\frac{1}{K_L} \approx \frac{1}{k_L} \tag{10}$$

This implies that essentially all resistance to interfacial mass transfer of a sparingly soluble gas is localized in the liquid film at the interface, that is, the mass transfer is liquid-film controlled (5).

Because the transport flux J equals the rate of transfer per unit gas–liquid interfacial area, equation 7 may be expressed in terms of rate as follows

$$\frac{dC_L}{dt} = K_L a_L(C^* - C_L) \tag{11}$$

where C_L is the dissolved gas concentration at time t, and a_L is the gas–liquid interfacial area per unit volume of the liquid.

Volumetric Gas–Liquid Oxygen Transfer Coefficient

In view of approximation 10, the overall volumetric mass transfer coefficient $K_L a_L$ for gas–liquid mass transfer is often expressed as $k_L a_L$. The overall volumetric gas–liquid mass transfer coefficient $k_L a_L$ and the concentration driving force ($C^* - C_L$) for mass transfer must be known if the rate of supply of oxygen or other soluble gas is to be quantified (5). The mass transfer driving force, which is easily determined, depends on temperature and pressure. The coefficient $k_L a_L$ depends on the properties of the fluid, the hydrodynamic regime, and the configuration of the bioreactor. The factors influencing $k_L a_L$ and mass transfer are summarized in Table 4. Prediction of $k_L a_L$ is an important part of bioreactor design. The individual terms in $k_L a_L$ are difficult to measure directly, but the product is relatively easily measured (5).

Usually, the overall volumetric gas–liquid mass transfer coefficient is expressed in terms of the liquid volume in the fermenter; however, sometimes the basis of expression has been the gas–liquid dispersion volume (5). An identical basis is necessary for comparing mass transfer capabilities, and the coefficient expressed in one form may be converted to the other if the overall gas holdup (ϵ_G) is known:

$$k_L a_L = \frac{k_L a_D}{1 - \varepsilon_G} \tag{12}$$

Some typical $k_L a_L$ values seen in various types of bioprocesses are noted in Table 5.

As noted earlier, the rate of diffusive transport of a component through a unit area perpendicular to the direction of transport can be expressed as mass flux or molar flux—kilogram or kilomole of component transferred per unit area per unit time. The flux J is proportional to a mass transfer driving force and a mass transfer coefficient K_L, or $J = K_L \times$ (driving force). Mass transfer is driven by differences in chemical potential. A species diffuses from regions of high potential to those of low potential until the difference has been eliminated. In practice, the driving force for mass transfer may be expressed as a concentration difference, partial pressure difference, or mole fraction difference. For example, when oxygen-free water is

Table 4. Factors Influencing Gas–Liquid Mass Transfer

Temperature
Pressure
Diffusivity
Viscosity
Density
Surface tension
Presence of surfactants and ions
Ionic strength
Concentration of solids
Hydrophobicity of solids
Morphology of solids
Shear rate or power input
Aeration velocity
pH
Geometry of the bioreactor
Flow parameters of non-Newtonian fluids

Table 5. Typical $k_L a_L$ Values Seen in Bioprocesses

Process	$k_L a_L$ (s^{-1})
Fungal fermentations	10^{-2}
Bacterial and yeast fermentations	10^{-1}
Wastewater treatment	3×10^{-3}

Source: Data from Ref. 5.

brought in contact with air, oxygen diffuses from air to water until the water becomes saturated. At this point, the oxygen concentrations in the two phases are in dynamic equilibrium, but there is no net transfer. Note that the concentrations of oxygen in water and air are not equal, but the chemical potentials are. Depending on how the mass flux J and the driving force are expressed, the mass transfer coefficient may have different numerical values and different units, as noted in Table 6 for a few cases. A coefficient expressed in one form may be easily converted to a different one.

Measurement of k_La_L. Several methods are available for measuring k_La_L values (1,5–7). Here only a few of the more useful and common methods will be noted.

Sulfite Oxidation Method. The sulfite oxidation method has been widely used, but it is suited to measurements only in physical systems (no live microorganisms) such as water or some polymer solutions. The method is subject to too many interferences (5,6) and, therefore, it is not suitable for use in most culture media or in nonviable fermentation broths. Moreover, because of high ionic strength, the sulfite oxidation medium is strongly noncoalescing and rheologically quite different from fermentation systems. Consequently, the sulfite oxidation measurements are not readily translated to biologically relevant systems.

The sulfite oxidation method relies on oxidation of sulfite to sulfate using dissolved oxygen:

$$SO_3^{2-} + 1/2O_2 \xrightarrow{\text{Catalyst}} SO_4^{2-}$$

The reaction is catalyzed by copper (II) and cobalt (II) as well as other metal ions. The rate of sulfite oxidation is a measure of the rate of oxygen transfer to the liquid. The overall reaction generates H^+, and pH control is necessary because the reaction rate is pH dependent. Also, the temperature needs to be controlled carefully to ensure constant reaction kinetics. Impurities such as iron, manganese, and other transition metals in the parts per million range affect the reaction.

For k_La_L measurements, the reaction conditions should be such that the reaction occurs in the bulk liquid and not at the gas–liquid interface. Thus, Cu (II) catalysis is used with a sufficiently high sulfite concentration that the oxidation rate is independent of the sulfite level. The rate of sulfite consumption is given by

$$-\frac{d[SO_3^{2-}]}{dt} = k_La_L(C^* - C_L) \qquad (13)$$

where C^* and C_L are the saturation concentration of dissolved oxygen and the actual instantaneous concentration at time t. C_L is usually close to zero. The sulfite consumption rate, which should remain constant, is followed by quenching the samples taken at various times with excess iodine and back titration of the residual iodine with thiosulfate (5).

Another oxygen-consuming reaction is the oxidation of dithionite ($S_2O_4^{2-}$) in alkaline solutions; however, in view of uncertain kinetics and other factors, the usefulness of that reaction for k_La_L measurements remains questionable.

Chemical Absorption of Carbon Dioxide. This method is similar in principle to the sulfite oxidation procedure. Carbon dioxide is absorbed into a mildly alkaline or suitably buffered solution where it is converted to other species. Chemical absorption of carbon dioxide into Na_2CO_3–$NaHCO_3$ solution is sufficiently slow for use in k_La_L determinations when the ratio of carbonate to bicarbonate concentrations is between 3 and 5 (5). Other reaction conditions need to be controlled to ensure that the reaction rate does not enhance mass transfer. The method has limitations similar to those of the sulfite oxidation technique.

The Hydrazine Method. The hydrazine method makes use of the reaction

$$\underset{\text{Hydrazine}}{N_2H_4} + O_2 \xrightarrow{\text{Catalyst}} N_2 + H_2O$$

The reaction is best carried out at pH 11 to 12 using copper sulfate catalyst (5). A steady flow of hydrazine into an aerated reactor is used to maintain a constant concentration of dissolved oxygen. The amount of hydrazine consumed per unit time equals the oxygen absorption rate. The reaction does not produce any ionic species; hence, the ionic strength of solution can be kept lower than with the sulfite oxidation or the carbon dioxide absorption methods.

Oxygen Balance Method. This technique depends on measurement of mass flow rates of aeration gas into and out of the fermenter. The mass fraction of oxygen in the inlet and exhaust gas streams must also be determined (mass spectrometer, paramagnetic oxygen analyzer) as well as the steady state dissolved oxygen concentration in the broth (dissolved oxygen electrode). The k_La_L is obtained from the oxygen balance:

Table 6. Possible Ways of Expressing Mass Flux and Mass Transfer Coefficient

Flux, J	Driving force	Units of K_L
$kg\ s^{-1}\ m^{-2}$	Concentration difference, $kg\ m^{-3}$	$m\ s^{-1}$
$kg\ s^{-1}m^{-2}$	Mole fraction difference	$kg\ s^{-1}m^{-2}$
$kg\ s^{-1}m^{-2}$	Partial pressure difference	$kg\ s^{-1}m^{-2}$
$kmol\ s^{-1}m^{-2}$	Concentration difference, $kg\ m^{-3}$	$m\ s^{-1}$
$kmol\ s^{-1}m^{-2}$	Mole fraction difference	$kmol\ s^{-1}m^{-2}$
$kmol\ s^{-1}m^{-2}$	Partial pressure difference	$kmol\ s^{-1}m^{-2}$
$m^3m^{-2}s^{-1}$ (for transfer of a gas)	Partial pressure difference	$m\ s^{-1}$

$$F(x_i - x_o) = V_L k_L a_L (C^* - C_L) \quad (14)$$

where F is the mass flow rate of gas, x_i and x_o are the mass fractions of oxygen in the gas (i = inlet, o = outlet), and V_L is the volume of the broth in the fermenter. Equation 14 assumes no evaporation and it does not correct for carbon dioxide production; however, the necessary corrections can be easily incorporated. For constant volume continuous culture, equation 14 needs to be further modified to account for oxygen entering and leaving the fermenter via the liquid streams.

The oxygen balance method requires accurate measurements of gas flow rates and oxygen concentrations. The oxygen mass fraction in the inlet and outlet gas must differ measurably (i.e., the oxygen consumption rate in the fermenter should be relatively large) for reliable $k_L a_L$ measurements. Under suitable conditions, the method is very reliable, and it does not disturb the fermentation by interrupting the air supply (7).

Enzymatic Oxidation Method. Oxygen transfer may be followed by the formation of gluconic acid via oxidation with oxygen in presence of the enzyme glucose oxidase. The rate of absorption of oxygen is calculated from the rate of consumption of alkali that is needed to neutralize the acid. The liquid-phase oxygen concentration is determined with a dissolved oxygen electrode. The cost of the enzyme restricts the method to small-scale applications (5). The procedure applies only to physical systems.

The Dynamic Gassing-In Method. In a fermentation at steady state, the oxygen supply rate exactly matches the rate of consumption by the microorganisms and the concentration of dissolved oxygen does not change:

$$\frac{dC_L}{dt} = k_L a_L (C^* - C_L) - q_{O_2} X = 0 \quad (15)$$

where q_{O_2} is the specific oxygen consumption rate, X is the concentration of viable biomass, $q_{O_2} X$ is the consumption term, and $k_L a_L (C^* - C_L)$ is the oxygen supply or mass transfer term. If now the oxygen supply is interrupted, the mass transfer term in equation 15 becomes zero and, as shown in Figure 2, the dissolved oxygen concentration declines linearly with time (7). The slope of the C_L versus t line during the unaerated period provides $q_{O_2} \times X$. After a short period, before the oxygen concentration has dropped to the critical level, the oxygen supply must be resumed. Now, the dissolved oxygen level rises (Fig. 2). During the rise period, the dissolved oxygen concentration as a function of time is given by

$$C_L = C^* - \frac{1}{k_L a_L}\left(q_{O_2} X + \frac{dC_L}{dt}\right) \quad (16)$$

which is a rearranged form of equation 15. Thus, $k_L a_L$ can be calculated from the slope of a plot of

$$C_L \text{ versus } \left(q_{O_2} X + \frac{dC_L}{dt}\right)$$

This procedure requires interrupting the oxygen flow to the fermenter, and the dissolved oxygen electrode needs to be capable of a rapid response.

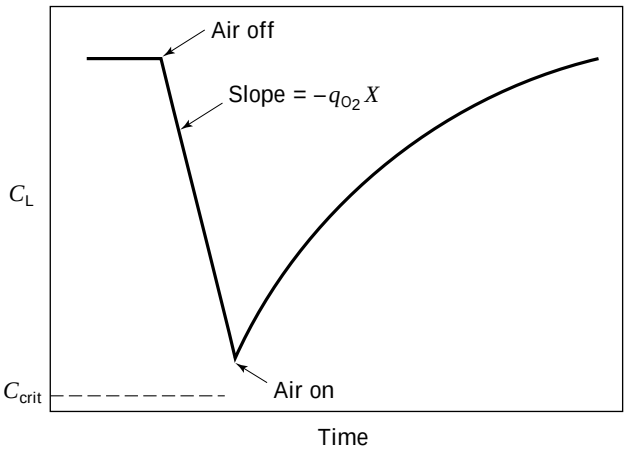


Figure 2. Dissolved oxygen concentration as a function of time during dynamic determination of the overall volumetric gas–liquid mass transfer coefficient in a fermentation.

The gassing-in method is commonly applied also to physical systems (5). In such cases the dissolved oxygen in the liquid is first reduced to near zero by bubbling with nitrogen. The nitrogen flow is then stopped, the bubbles are allowed to leave the fluid, and the system is oxygenated by bubbling with air. The concentration of the dissolved oxygen in the liquid is followed as a function of time, and the $k_L a_L$ is calculated as the slope of the linear equation

$$\ln\left(\frac{C^* - C_{L0}}{C^* - C_L}\right) = k_L a_L t \quad (17)$$

where C^* is the saturation concentration of dissolved oxygen, C_{L0} is the initial dissolved oxygen concentration at time t_0 when a hydrodynamic steady state has been reestablished upon commencement of aeration, and C_L is the dissolved oxygen concentration at any time t (5). Equation 17 is obtained by integration of equation 15, assuming a well-mixed liquid (i.e., uniform C_L), no change in gas phase composition (i.e., a constant C^*), and an absence of oxygen consumption (i.e., $q_{O_2} = 0$) in the physical system.

Correct evaluation of $k_L a_L$ by this method requires attention to several factors, including the response dynamics of the dissolved oxygen electrode and the nature of mixing in the gas and the liquid phases. In large intensely mixed fermenters, a well-mixed liquid phase may still be assumed, but the gas phase composition changes significantly from the inlet to exhaust. The gas is generally in plug flow and, thus, a log mean concentration driving force $(C^* - C_L)_{LM}$ should be used in equation 15. The driving force is

$$(C^* - C_L)_{LM} = \frac{C^*_{in} - C^*_{out}}{\ln\left(\frac{C^*_{in} - C_L}{C^*_{out} - C_L}\right)} \quad (18)$$

where C^*_{in} and C^*_{out} are the saturation concentrations in

equilibrium with the gas phase composition at the fermenter inlet and outlet, respectively.

Generally, the nature of mixing in the gas phase is inconsequential in k_La_L determinations so long as $1/k_La_L$ is less than the residence time of the gas phase in the liquid. The mean residence time t_R of the gas phase in bubble columns and stirred tanks can be estimated (5) using

$$t_R = \frac{h_L\varepsilon_G}{(1 - \varepsilon_G)U_G} \tag{19}$$

where ϵ_G is the overall gas holdup. Equation 19 may be applied also to risers of airlift reactors when there is little or no gas in the downcomer. In the latter case, the U_G should be replaced with U_{Gr} (the superficial gas velocity based on the riser cross-section), and the gas holdup should be replaced with ϵ_{Gr}. When there is gas in the downcomer, equation 19 may still be used because the downcomer gas holdup acts mostly as a dead volume.

The response characteristics of the dissolved oxygen electrode have little influence on the dynamic determination of k_La_L if the response time (63% of full response) is less than or equal to $1/k_La_L$. Consult Chisti (5) for additional details. Other variations of the gassing-in technique exist.

Estimation of k_La_L for Gases Other Than Oxygen. Most of the available data and correlations for k_La_L were obtained for mass transfer of oxygen. The overall volumetric mass transfer coefficient for transfer of other gases can be estimated using the oxygen transfer data and the diffusion coefficients for oxygen and the other gas (5); thus

$$(k_La_L)_{gas} = \frac{D_{gas}}{D_{oxygen}}(k_La_L)_{oxygen} \tag{20}$$

or

$$(k_La_L)_{gas} = \sqrt{\frac{D_{gas}}{D_{oxygen}}}(k_La_L)_{oxygen} \tag{21}$$

Equation 20 applies when the transfer can be ascribed to the film penetration mechanism of the two-film theory (5). Equation 21 is applicable when the surface renewal is the predominant mechanism of mass transfer (5). The difference between the two correction factors is small. The diffusivities used for the correction should be at the same temperature as the k_La_L data. The correction is satisfactory only for transfer in identical fluids at identical hydrodynamic conditions. Data on carbon dioxide mass transfer have been reviewed by Ho and Shanahan (8).

Solubility of Oxygen. Information on solubility of oxygen is needed for estimating the concentration difference driving force for mass transfer. Calculation of the saturation concentration C^* of dissolved oxygen (equation 8) requires a knowledge of the Henry's law constant, H. The dimensionless H for oxygen in pure water may be calculated as a function of temperature T (°C) using the equation

$$H = -9281.4 \times 10^{-3} + 4030.9 \times 10^{-3}T - 1403.5 \times 10^{-4}T^2 + 178.8 \times 10^{-5}T^3 \tag{22}$$

Equation 22 provides almost exact values of H over the range 20 to 30 °C. Oxygen solubility in water at a pressure of 1 atm can be calculated with the equation

$$C^* = 2178.549 \times 10^{-3} - 549.304 \times 10^{-4}T - 848.781 \times 10^{-6}T^2 - 483.785 \times 10^{-8}T^3 \tag{23}$$

where the temperature T is in °C and C^* is in mmol L^{-1}. Equation 23 applies over 0 to 40 °C. In addition to temperature, other factors influence oxygen solubility and the H value. Solubility declines with increasing ionic strength of solution. Solubility depends also on the types of ions in solution. For example, for the same molar concentrations of sodium chloride and hydrochloric acid in different samples of pure water, the oxygen solubility is lower in the salt solution. Solubility data for oxygen and other gases are available in handbooks. At 37 °C, the saturation dissolved oxygen concentration in a typical animal cell culture medium with serum is about 6.5×10^{-3} kg m^{-3}.

LIQUID–LIQUID MASS TRANSFER

Transfer of a solute from an immiscible liquid dispersed into another liquid (the continuous phase) can be analyzed using the two-film approach discussed for gas–liquid mass transfer. Now, the overall mass transfer coefficient K_L and the film coefficients for the dispersed (k_{Ld}) and the continuous (k_{Lc}) phases are related as follows:

$$\frac{1}{K_L} = \frac{1}{k_{Lc}} + \frac{K_p}{k_{Ld}} \tag{24}$$

In equation 24, K_p is the partition coefficient of the solute between the two phases. The partition coefficient defines the equilibrium distribution of the solute between the phases; thus

$$K_p = \frac{C_{Lc}}{C_{Ld}} \tag{25}$$

where C_{Lc} and C_{Ld} are the equilibrium concentrations of the solute in the continuous and the dispersed phases, respectively. Again, as for gas–liquid mass transfer, the interface between the phases is assumed to offer no resistance to mass transfer; hence, the equilibrium relationship determines the solute concentrations at the interface.

The overall volumetric mass transfer coefficient for liquid–liquid mass transfer is K_La_{Lc} where the specific interfacial area (a_{Lc}) is based on the volume (V_{Lc}) of the continuous phase. The rate of mass transfer from the dispersed to the continuous phase is given by

$$\frac{dC_{Lc}}{dt} = K_La_{Lc}(C^* - C_{Lc}) \tag{26}$$

where C^* is the saturation concentration of the solute in

the continuous phase in equilibrium with the dispersed phase.

SOLID–LIQUID MASS TRANSFER

Mass transfer to or from suspended solids is important in many bioprocess situations. Solid–liquid mass transfer may be rate limiting in immobilized enzyme and cell reactors, during dissolution of solids, in adsorption, and in other cases. The rate of mass transfer to or from the suspended particle depends on the solid–liquid mass transfer coefficient (k_L), the total solid–liquid interfacial area (A_s), and the concentration driving force; thus

$$V_L \frac{dC_L}{dt} = k_L A_s (C^* - C_L) \tag{27}$$

where C_L is the concentration of the transferring component in the liquid at time t, C^* is the saturation concentration (or solubility) of the transferring component, and V_L is the volume of the suspending liquid. Equation 27 is written for transfer from the solid to the liquid. By analogy with the film model of mass transfer, the coefficient k_L conceptually equals the diffusivity of the solute divided by the thickness of the stagnant liquid film at the solid–liquid interface; k_L is needed for quantifying the rate of mass transfer. Methods for estimating the solid–liquid or solid–gas (also gas–liquid if liquid is adsorbed as a very thin film on the surface of the solid) mass transfer coefficient k_L in various situations are discussed in a later section of this monograph.

MASS TRANSFER BEHAVIOR

General Aspects

Mass transport by diffusion and forced convection are influenced by the following variables: the mass transfer coefficient k_L, the molecular diffusivity of the transferring component D_L, the fluid density ρ_L and viscosity μ_L, a characteristic length d, and the velocity of flow U_L. Mass transfer by natural convection is influenced also by the gravitational acceleration g and the density difference $\Delta\rho$ between phases. These variables can be grouped into the following principal dimensionless groups or numbers:

$$Re \text{ (Reynolds number)} = \frac{\text{Inertial force}}{\text{Viscous force}} = \frac{\rho_L U_L d}{\mu_L} \tag{28}$$

$$Sh \text{ (Sherwood number)} = \frac{\text{Total mass transfer}}{\text{Diffusive mass transfer}} = \frac{k_L d}{D_L} \tag{29}$$

$$Sc \text{ (Schmidt number)} = \frac{\text{Momentum diffusivity}}{\text{Mass diffusivity}} = \frac{\mu_L}{\rho_L D_L} \tag{30}$$

$$Gr \text{ (Grashof number)} = \frac{(\text{Inertial force})(\text{buoyancy force})}{(\text{Viscous force})^2} = \frac{d^3 \rho_L \Delta\rho g}{\mu_L^2} \tag{31}$$

$$Fr \text{ (Froude number)} = \frac{\text{Inertial force}}{\text{Gravitation force}} = \frac{U_L^2}{gd} \tag{32}$$

The above-noted dimensionless numbers represent ratios of various factors that may play a role in a given situation. The Grashof number is important in situations where density-difference-driven natural convection affects mass transfer. The Reynolds number is used in situations where forced convection is the predominant influence on transport processes. The Froude number is useful in describing gravity-influenced flows. The Schmidt number, ratio of momentum diffusivity to mass diffusivity, is a measure of relative effectiveness of mass and momentum transfer. Some of those dimensionless numbers may be expressed in other forms, depending on the situation. A few possible alternatives are noted in Table 7. In addition to the dimensionless groups discussed above, other groups relevant to mass transfer and related fluid mechanics are encountered in the literature. Some of the more common ones and their physical significance are noted in Table 8.

Grouping the many variables that affect mass transfer into a few dimensionless numbers greatly simplifies the process development experimentation and data correlation for predicting the behavior of a system. For example, for fluids flowing in straight pipes, the flow pattern changes from laminar to turbulent at a Reynolds number value of about 2,300. This occurs irrespective of which specific fluid—air, water, milk, or thick glucose syrup—is being used so long as the combination of the fluid density, viscosity, flow velocity, and pipe diameter yields a Reynolds number value that exceeds 2,300. Equations for correlating the mass transfer behavior of systems are often expressed in terms of dimensionless groups.

Behavior of Dispersions

A dispersion consists of one or more phases dispersed in a continuous phase. The continuous phase may be a gas or liquid. The dispersed phase may be liquid droplets, solid particles, or gas bubbles. The magnitude of mass transfer coefficient inside and outside fluid particles depends on whether the particles behave as ridged spheres or are mobile. Small drops and bubbles behave as noncirculating ridged spheres. Larger bubbles and drops have internal fluid circulation and mobile interfaces because of relative motion between the particle and the surrounding fluid. Bubbles are immobile if the dimensionless diameter d^* is 10 or below. The interface is nearly always mobile if d^* exceeds 50. Here, the dimensionless diameter number is

$$d^* = d_B \left(\frac{\mu_L^2}{\rho_L g (\rho_L - \rho_G)} \right)^{-1/3} \tag{32}$$

as defined by Wesselingh (5). Using this criterion, in air and water, bubbles smaller than 0.5 mm will always be ridged, whereas bubbles larger than about 2.5 mm will have mobile interfaces (5).

In certain cases, drops and bubbles may experience additional interfacial motions that are a consequence of mass transfer itself. Because the local mass transfer coefficient around a particle may be different at different locations,

Table 7. Alternative Methods of Expressing Dimensionless Groups

Group	Alternative expressions
Reynolds number, *Re*	$\frac{\rho_L U_L d}{\mu_L}$ (in pipes and channels)
	$\frac{\rho_L N d_i^2}{\mu_L}$ (stirred tanks)
	$\frac{\rho_L U_L L}{\mu_L}$ (flow past a plate)
	$\frac{\rho_L U_L d_p}{(1-\phi)\mu_L}$ (flow past particles in packed beds)
Sherwood number, *Sh*	$\frac{k_L d}{D_L}$ (d may be diameter of a particle, flow channel, plate, etc.)
Froude number, *Fr*	$\frac{U_L^2}{g h_L}$, or $\frac{U_G^2}{g h_L}$ (in bubble columns and airlift fermenters)
	$\frac{N d_i^2}{g h_L}$, or $\frac{N^2 d_i}{g}$ (in stirred tanks)
Peclet number, *Pe*	$\frac{U_L h_L}{D_L}$, or $\frac{U_B d_B}{D_L}$ (in bubble columns and airlift fermenters)
	$\frac{N d_i^2}{D_L}$ (in stirred tanks)

Table 8. Other Dimensionless Groups Relevant to Mass Transfer and Related Fluid Mechanics

Group	Definition	Physical significance
Bond number, *Bo*	$\frac{g d_p^2 \Delta\rho}{\sigma_L}$	$\frac{\text{Gravity force}}{\text{Surface tension force}}$
Peclet number, *Pe*	$\frac{U_L L}{D_L}$	$\frac{\text{Bulk mass transport}}{\text{Diffusional mass transport}}$
Poiseuille number, *Ps*	$\frac{\mu_L U_p}{\rho_L g d_p^2 \Delta\rho}$	$\frac{\text{Viscous force}}{\text{Gravity force}}$
Power number, *Po*	$\frac{P}{\rho_L N^3 d_i^5}$	
Rayleigh number, *Ra*	$Gr \times Sc$	
Stanton number, *St*	$\frac{Sh}{Re \times Sc}$	
Weber number, *We*	$\left(\frac{U_p^2 \rho_L d_p}{\sigma_L}\right)^{1/2}$	$\frac{\text{Inertial force}}{\text{Surface tension force}}$

the local mass transfer rates may be different. In this situation, uneven transfer of a solute such as acetone or ethanol will generate solutions of different local concentrations and surface tensions. Adjacent films of different surface tension lead to violent motions: regions of high surface tension contract, and the film ruptures in regions of low surface tension. This effect is the well-known Marangoni effect, and it can significantly enhance mass transfer, especially in liquid–liquid extraction processes.

For suspended solids, increasing agitation beyond that needed to achieve complete suspension has little impact on mass transfer coefficient. The coefficient is influenced by relative velocity between the phases, which is determined by density differences. In gas–liquid dispersions, too, the k_L is not substantially affected by increasing agitation, but the overall volumetric mass transfer coefficient ($k_L a_L$) increases because of the increase in gas–liquid interfacial area.

For freely suspended particles and bubbles, mass transfer to or from the bubble or particle obeys the general relationship

$$Sh = f(Gr, Sc) \tag{33}$$

For just suspended small particles ($d_p < 0.6$ mm) in stirred tanks, Calderbank and Moo-Young (9) recommended the equation

$$k_L = \frac{2D_L}{d_p} + 0.31\, Sc^{-2/3}\left(\frac{\Delta\rho_L \mu_L g}{\rho_L^2}\right)^{1/3} \tag{34}$$

which is suitable for suspensions as well as gas–liquid dispersions of noncirculating bubbles ($d_p < 0.6$ mm). For larger circulating bubbles and liquid drops ($d_p > 2.5$ mm) in just-suspended state, the recommended equation is

$$k_L = 0.42 Sc^{-0.5}\left(\frac{\Delta\rho_L \mu_L g}{\rho_L^2}\right)^{1/3} \tag{35}$$

For bubbles in the 0.6 to 2.5-mm range, the k_L may be estimated as a linear function of bubble diameter.

Large spherical cap bubbles frequently occur in gas–liquid dispersions, especially when the viscosity of liquid exceeds about 7×10^{-2} Pa s. Calderbank and Lochiel (10) correlated mass transfer from such bubbles using the equation

$$Sh = \frac{1.79(3R_b^2 + 4)^{2/3}}{R_b^2 + 4}(Re \times Sc)^{1/2} \tag{36}$$

where R_b is the ratio of bubble width to bubble height. For spherical caps, R_b is about 3.5; hence, equation 36 becomes

$$Sh = 1.31(Re \times Sc)^{1/2} \quad (37)$$

Mass transfer to or from a particle suspended in a stagnant fluid occurs solely by diffusion. For a single spherical particle surrounded by a stagnant medium, the theoretical minimum value of the transfer coefficient is given by $Sh = 2$. For a single particle in the creeping flow regime (Re < 0.1), the specific relationship is

$$Sh = 0.39(Gr \times Sc)^{1/3} \quad (38)$$

Equation 38 applies to particles with ridged interfaces, and this includes noncirculating small bubbles ($Re < 0.1$) in fermentation broths. Thus, compared to the case for a single noncirculating bubble (equation 38), the mass transfer coefficient in swarms of ridged bubbles (equation 34) is about 20% lower.

For power inputs greater than needed for just suspension, Calderbank and Moo-Young (9) have established the following equation in stirred tanks

$$k_L = 0.13\left(\frac{(P_G/V_L)\mu_L}{\rho_L^2}\right)^{1/4} Sc^{-2/3} \quad (39)$$

where P_G/V_L is the power input per unit volume.

Transfer coefficient correlations generally require a knowledge of liquid properties such as viscosity and density. This is usually not a problem when dealing with Newtonian liquids, but difficulties arise with slurries and non-Newtonian media. Slurries of biomass solids may be treated usually as pseudohomogeneous fluids (5). Depending on the amount of suspended solids, slurries may behave as Newtonian or non-Newtonian power law fluids. For small amounts of spherical solid particles, Newtonian behavior is commonly observed and in this case the viscosity of the slurry is independent of shear rate. The Newtonian viscosity may be estimated using equations of Einstein and Vand (7). Another suitable equation is that of Thomas:

$$\mu_{SL} = \mu_L(1 + 2.5\varphi_S + 10.05\varphi_s^2 + 0.00273e^{-16.6\varphi_S}) \quad (40)$$

where φ_S is the volume fraction of suspended solids and μ_L is the viscosity of the suspending fluid.

When the slurry behaves as a non-Newtonian power law fluid, the apparent viscosity depends on shear rate; thus

$$\mu_{ap} = K\gamma^{n-1} \quad (41)$$

where K is the consistency index or thickness of the fluid and n is its flow behavior index. The parameter γ is the shear rate. The shear rate is difficult to define in most realistic configurations of bioreactors under the usual operational conditions; nevertheless, the following expressions are commonly used in estimating shear rates.

In *bubble columns and airlift devices*,

$$\gamma = \alpha U_G \quad (42)$$

where the constant α has been specified variously as 1,000, 2,800, 5,000 m^{-1}, and so on (11). Equation 42 has been applied also to airlift reactors, but that use is incorrect; for airlift reactors, the superficial gas velocity based on the cross-sectional area of the riser should be used in expressions such as equation 42 as recommended by Chisti (5). Depending on the constant used, equation 42 provides wildly different values of the supposed shear rate; hence, its use is not generally favored and it has been severely criticized (5,11,12). Another expression for shear rate in airlift reactors is

$$\gamma = 3.26 - 3.51 \times 10^2 U_{Gr} + 1.48 \times 10^4 U_{Gr}^2 \quad (43)$$

which was developed for $0.004 < U_{Gr}$ (m s^{-1}) < 0.06 (13). Equation 43 also has significant problems, as discussed elsewhere (12,14). Alternative approaches for estimating shear rates have been propounded by Molina Grima et al. (12).

In *stirred vessels*, the mean shear rate in impeller agitated tanks is usually given as

$$\gamma = k_i N \quad (44)$$

where k_i is an impeller-dependent constant (7). Some typical k_i values are 11 to 13 for six-bladed turbines, 10 to 13 for paddles, ~10 for propellers, and ~30 for helical ribbon impellers. Again, as with pneumatically agitated bioreactors, much of the discrepancy among various predictive correlations that rely on a shear-rate-dependent apparent viscosity has been associated with the use of equations such as equation 44 (15).

The apparent viscosity of many fermentation fluids declines with increasing shear rate, that is, $n < 1$ (equation 41), and these fluids are known as shear thinning, or pseudoplastic. Because shear rate is not easily defined in typical bioreactors (5,11,12,15), correlations using solids holdup directly are preferred to those relying on a poorly established apparent viscosity of the slurry (5). As with viscosity, the density of a pseudohomogeneous slurry may be related to phase holdups as follows

$$\rho_{SL} = \rho_L(1 - \varphi_S) + \rho_S\varphi_S \quad (45)$$

In some viscous fermentations, the principal resistance to oxygen transfer may be in the bulk fluid or at the cell–liquid interface and not at the gas–liquid interface. In such cases, dilution of the broth with water and increased agitation may improve the transfer rate a little. In broths of filamentous fungi, promoting pelletlike growth as opposed to the usual pulplike morphology may substantially reduce the broth viscosity and improve mass transfer. Additional details on transport fundamentals of dispersions have been noted by Moo-Young and Blanch (16). Similar discussions with a focus on gas–liquid mass transfer in non-Newtonian media are given in the work by Oolman and Blanch (17). Carbon dioxide mass transfer in bioreactors has been treated in depth by Ho and Shanahan (8).

Gas–Liquid Mass Transfer

Either bubble-free aeration of the liquid surface in a bioreactor may be used to provide the needed oxygen or the

aeration gas mixture may be bubbled or sparged through the body of the fluid. These schemes are discussed separately below.

Surface Aeration. Oxygen transfer through the surface of a fluid is useful in relatively small-scale devices or in microaerophilic processes and certain wastewater treatment applications (18). Surface aeration is especially common in early stages of culture, in shake flasks, tissue culture flasks, roller bottles, and spinner flasks. Understanding and characterization of surface aeration in small-scale devices is essential for process scale-up that minimally attempts to reproduce on larger scale the culture performance attained in the laboratory. Surface aeration is used also in relatively small production bioreactors for growing cultures with extremely low oxygen requirements. Surface aeration in the principal types of laboratory and production culture devices is detailed in the following sections.

Laboratory Culture Devices

1. *Shake flasks.* In 500-mL Erlenmeyer flasks placed on reciprocating and rotary shaker platforms, Yamada et al. (19) measured volumetric rates of oxygen transfer in sulfite solutions and in *Acetobacter suboxydans* fermentations for converting sorbitol to sorbose. The rates were expressed in terms of the volumetric oxygen transfer coefficients through the flask closure (K_P) and across the gas–liquid interface (K_S):

$$\begin{matrix}\text{Volumetric rate}\\ \text{of oxygen transfer}\\ (\text{mol } O_2/\text{mL h})\end{matrix} = \frac{1}{\left(\frac{1}{K_P}+\frac{1}{K_S}\right)}\frac{1}{V_L}(p_{atm}-p_L) \quad (46)$$

In equation 46, V_L (mL) is the volume of liquid in the flask; p_{atm} (atm) and p_L (atm) are the partial pressures of oxygen in the atmosphere outside the flask and in the liquid in the flask, respectively. The units of K_P and K_S are mol O_2 atm^{-1} h^{-1}. The mass transfer coefficient K_P through the closure declined with increasing weight (3 to 7 g) of cotton used to form the plug. The mean K_P value for cotton plugs was 2.87×10^{-2} mol O_2 atm^{-1} h^{-1}. Values for polyurethane foam and silicone foam plugs were a little lower. Compared to open flasks, cotton plugs restricted oxygen transfer. The K_S values in sulfite oxidation solutions agitated at 110 to 140 rpm ranged over $(2.04 \text{ to } 4.63) \times 10^{-2}$ mol O_2 atm^{-1} h^{-1}. Under similar conditions in *Acetobacter suboxydans* fermentations, the K_S values were lower—only 50 to 60% of those obtained with sulfite oxidation. Various types of baffles and indentation in the flasks enhanced K_S by about 50% relative to the base case.

For mass transfer from liquid surface in unbaffled Erlenmeyer shake flasks, the following equation has been recommended (20)

$$k_L a_L\left(\frac{\mu_L}{\rho_L g^2}\right)^{1/3} = 0.5\left(\frac{\mu_L}{\rho_L D_L}\right)^{-1/2}\left(\frac{d_F}{V_L}\right)^{8/9}\left(\frac{\mu_L^2}{\rho_L^2 d_F^3 g}\right)^{8/27} \cdot\left(\frac{N^2\sqrt{\hat{e}d_F}}{g}\right)^{0.5} \quad (47)$$

where d_F is the maximum diameter of the flask, N is the speed of rotation, and $\hat{e}$ is the eccentricity of the shaker platform. Equation 47 applies to animal cell culture media when the liquid volume in the flask is 50 to 200 mL.

2. *T flasks.* Tissue culture flasks or T flasks are commonly used in initial stages of animal cell culture. During typical culture conditions, the cells utilize the oxygen in the liquid layer, but the concentration of oxygen in the gas phase is little affected irrespective of whether the flask is fully closed or the cap is cracked open (21). Clearly, diffusion through the liquid layer controls mass transfer; thus, the shallower the liquid depth the better. The oxygen transfer resistance of the deposited cell layer at the bottom of the flask is insignificant compared to the resistance through the bulk liquid layer (21). The observed oxygen transfer rates through the liquid layer tend to be greater than if the transfer were purely by diffusion. This suggests that convection and micromixing (e.g., during sampling, opening or closing of incubator door) significantly contribute to transfer. In typical culture, the dissolved oxygen declines linearly with time until, by about 60 h, the dissolved oxygen level is reduced to below 10% of air saturation (21). Because T flasks are typically inoculated and transferred every other day, the observed decline in oxygen may be of little real consequence; however, if the culture period is to exceed 60 h and dissolved oxygen levels of 10% or higher are wanted, then once or twice daily gentle agitation of T flasks is recommended.

3. *Roller bottles.* Roller bottles are commonly used for initial stages of cell culture. Typically about 0.1 m in diameter and about 0.25 m tall, plastic roller bottles are filled to about a third of their height with culture fluid. The bottles are laid on their sides in a roller cabinet and rotated at 1 to 2 rpm. Roller bottles are used to culture both adherent cells as well as nonanchorage-dependent lines. The interfacial area for oxygen transfer is the surface of the culture fluid and the liquid film on walls that continuously enters and leaves the pool of liquid. Mechanized and automatic handling of a large number of roller bottles have extended this culture technique to commercial production scale in a few cases, but this kind of processing has a very limited scope. The surface aeration capability of roller bottles, corrected for differences in surface-to-volume ratio, is somewhere between that of Erlenmeyer-style shake flasks under typical conditions and the stationary T flasks.

4. *Spinner flasks.* Aunins et al. (22) reported $k_L a_L$ measurements (surface aeration) in 500-mL Corning spinner flasks (Fig. 3) filled to a depth of 0.08 m with Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 5% (v/v) fetal calf serum and 1 kg m^{-3} EDTA. The data were obtained at 37 °C and followed the equation

$$Sh = 1.08Re^{0.78} \quad (48)$$

where the Sherwood number is based on the vessel diameter and the Reynolds number is based on the impeller diameter. The measurements spanned the impeller speed range 25 to 150 rpm, impeller diameters of 0.0525 m and 0.078 m, and impeller Reynolds number of 1,500 to 20,000. Locating the impeller less than 0.01 m below the liquid surface dramatically enhanced the $k_L a_L$. At 50 rpm, with

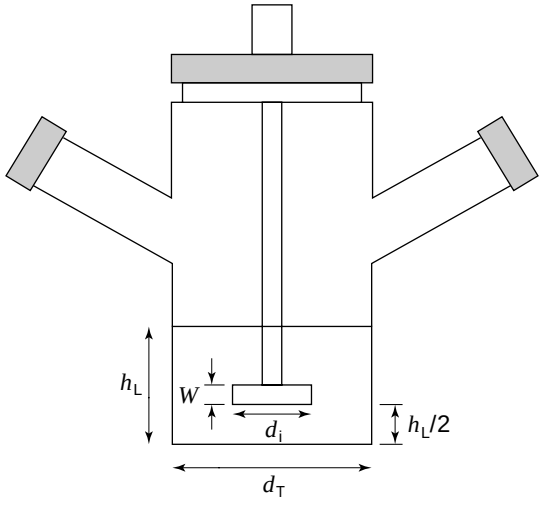


Figure 3. Geometry of 500-mL Corning spinner flasks: $h_L = 0.08$ m; $d_T = 0.096$ m; $d_i = 0.078$ m ($W = 0.025$ m), or $d_i = 0.053$ m ($W = 0.019$ m).

0.078 m diameter paddle impeller placed at the liquid surface, the k_La_L value was about 9.72×10^{-4} s^{-1}. Moving the impeller to 0.01 m below the surface provided a k_La_L value of about 4.17×10^{-4} s^{-1}. Lowering the impeller further into the fluid did not affect the k_La_L significantly.

5. *Stirred vessels*. Lavery and Nienow (23) reported k_La_L measurements in water, RPMI 1640 basal cell culture medium, and medium supplemented with 5% (v/v) fetal calf serum in a small, unbaffled, spherical cell culture vessel (1.5 L, static surface area of liquid $= 2.23 \times 10^{-2}$ m^2) stirred with one or two three-bladed propellers ($d_i = 0.060$ m, located 0.035 m apart). The lower impeller was positioned 0.003 m from the bottom of the vessel flask and the upper was 0.002 m below the surface of the liquid. The agitation speeds were 1.6 to 5.8 s^{-1} (100 to 350 rpm). Air was sparged (100 mL min^{-1}) either in the liquid under the lower propeller or only through the headspace (i.e., surface aeration). The submerged aeration k_La_L values generally agreed with predictions of Van't Riet's correlation for nonionic solutions (i.e., equation 57). Relative to measurements in water, the serum and the basal medium had little effect on k_La_L values. The k_L for surface aeration was little affected by the number of impellers (whether one or two). The k_La_L values for surface aeration were about 75% of those for submerged aeration in the reactor used. Addition of silicone antifoam (6 ppm) reduced the k_La_L values by about 50%. In the basal medium with serum and the antifoam, the k_L values for surface aeration (no vortex) were $(1.18 \text{ to } 3.54) \times 10^{-5}$ m s^{-1}. The k_La_L values for submerged aeration were $(2.8 \text{ to } 8.5) \times 10^{-4}$ s^{-1}.

6. *Other devices*. For absorption of oxygen at free surface of sodium sulfite solution in cylindrical containers placed on an orbital shaker platform, the volumetric mass transfer coefficient has been correlated (24) as follows:

$$k_La_L = 6 \times 10^{-5}(P/V_L)^{0.4} d_T^{-0.25} h_L^{-0.6} \tag{49}$$

Equation 49 is independent of gas flow rate in the headspace; it was determined for $20 \le P/V_L$ (W m^{-3}) ≤ 500 and $0.5 \le h_L/d_T \le 1.5$ ($d_T = 0.12$ or 0.15 m). The agitation speed of the shaker platform varied over 1.2 to 3.3 s^{-1}, and the moving vessel described a diameter of 0.01 to 0.04 m.

Larger Systems

1. *Stirred tanks*. Multibladed disk turbines located at the surface provide superior surface aeration relative to other types of impellers (25). To prevent entrainment of bubbles, the turbines must be operated (25) such that

$$Nd_i \leq 0.11(d_i/d_T)^{-0.2} \tag{50}$$

As the tank diameter increases, the ability of a given impeller to agitate the entire liquid surface declines if the agitation speed is to remain below the bubble entrainment limit. The optimum impeller size for surface aeration (25) is given by

$$\frac{d_i}{d_T} \approx 0.5 \tag{51}$$

Surface aeration is suited only to small reactors because k_La_L declines rapidly with increasing tank volume of surface aerated bioreactors. For subsurface impellers operated so that there is no entrainment of gas, the mass transfer coefficient k_L has been correlated with the impeller Reynolds number as follows (20):

$$Sh = 1.4Re_i^{0.76} \tag{52}$$

which applies to waterlike media.

The contribution of the free surface to aeration in conventionally stirred, baffled tanks declines as the scale of operation increases. The liquid-phase mass transfer coefficient at the free surface (no vortex) in such tanks can be estimated using the correlations summarized in Table 9. A preferred correlation (27) is

$$k_L = 0.138\, Sc^{-2/3}\left(\frac{4\mu_L Po N^3 d_i^5}{\pi d_T^2 h_L \rho_L}\right)^{1/4} \tag{53}$$

where Po is the power number ($Po = P/(\rho_L N^3 d_i^5)$); d_T and d_i are the diameters of the tank and the impeller, respectively; h_L is the height of liquid in the tank; N is the impeller rotational speed, and P is the power input. Equation 53 applies to Newtonian media without antifoams or surfactants. Added surfactants, fatty acids, and proteins reduce k_L relative to values in clean liquids.

2. *Vortex aeration*. In unbaffled stirred tanks, agitation with a centrally located impeller creates a vortex that draws closer to the impeller as the rate of agitation increases (Fig. 4). In a vortex-aerated industrial fermenter (240 L) used in producing diphtheria vaccine, the following equation has been reported (18)

$$k_La_L = 7.33 \times 10^{-4} + 1.36 \times 10^{-4} Q_G \tag{54}$$

where Q_G is the surface aeration rate (10 to 45 normal L min^{-1}). The equation was developed in aqueous sodium

Table 9. Correlations for Liquid-Phase Mass Transfer Coefficient at Free Surface in Baffled Stirred Tanks (No Vortex)

$k_L = 5.11 \times 10^{-3}\left[\frac{11^{n-1}K}{\rho_L}\left(\frac{3n+1}{4n}\right)^n\right]^{-0.426} D_L^{0.5} d_i^{0.852} N^{1.352-0.426n}$	Perez and Sandall (26)
$k_L = 0.322 Sc^{-2/3} N^{0.7} d_i^{0.4} (\mu_L/\rho_L)^{0.3}$	Hikita and Ishikawa (27)
$k_L = 0.0256 Sc^{-1/2} \frac{Po^{1/3} N d_i^2}{(\pi d_T^2 h_L/4)^{1/3}}$	Farritor and Hughmark (28)
$k_L = \left\{0.432\left(\frac{d_i}{d_T}\right) - 0.13\right\}\left(\frac{E\rho_L}{\mu_L}\right)^{1/3} D_L^{1/2}$	Bin (29)
$\frac{k_L d_i}{D_L} = 0.04\left(\frac{N d_i^2 \rho_L}{\mu_L}\right)\left(\frac{N^2 d_i^3 \rho_L}{\sigma_L}\right)^{1/2}\left(\frac{\mu_L}{\rho_L D_L}\right)^{1/2}$	Kataoka and Miyauchi (27)

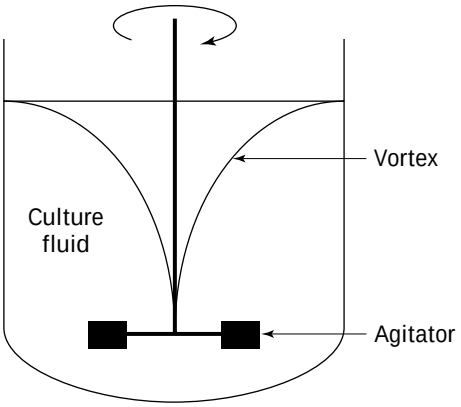


Figure 4. Vortex aeration in a stirred fermenter.

chloride (2.5 kg m^{-3}) and in the medium used in culturing the diphtheria bacterium. The tank was agitated with a six-bladed Rushton-type turbine (without disk), and the agitation rate was 300 to 800 rpm. The working aspect ratio was 0.74, which is substantially lower than values that are typically used in sparged microbial fermenters. Over the entire range of agitation rates used, the vortex was fully developed and reached the eye of the impeller (18). The conditions used in establishing equation 54 were identical to those used in commercial processes for producing the vaccine (18).

3. *Wetted-wall columns*. Mass transfer coefficient in liquid films in wetted-wall falling film columns can be estimated (2) with the equation

$$k_L = 9.0 Re^{-0.40} Sc^{-0.67} U_L \qquad (55)$$

where the *Sc* and *Re* are based on properties of the liquid. Depth of the film and the mean film velocity (U_L) are used to calculate the Reynolds number. Equation 55 is appropriate for $Re < 2{,}000$ and $Sc < 100$. Another useful equation is

$$Sh = 0.023 Re^{0.8} Sc^{1/3} \qquad (56)$$

Other correlations suited to highly turbulent films ($Re >$ 4,000) are available (30).

Submerged Geration. Submerged aeration in which an oxygen-containing gas is sparged or bubbled through the culture fluid is the norm in large-scale processes, including animal cell culture. Typically, submerged cultivation is carried out in stirred tanks, bubble columns, and airlift bioreactors. Stirred vessels are relatively more common, but the number of airlift devices continues to increase. Bubble columns are used less frequently (14). Typically, sparged aerated bubble columns and stirred tanks have aspect ratios between 3 and 4. Airlift devices generally have higher aspect ratios, usually greater than 6. Stirred vessels used in culturing animal cells are usually shorter, with aspect ratios less than 2. Microbial fermentations are aerated at high rates, with superficial aeration velocities of up to 0.1 m s^{-1} in airlift vessels and somewhat lower maximum values in bubble columns. Aeration rates are substantially lower in stirred tanks to prevent flooding of the impeller. A flooded impeller is a poor mixer and gas disperser.

Sparged aeration can potentially damage sensitive animal cells (12,31–34) and cells of certain cyanobacteria. Nevertheless, sparged aeration is successfully used quite widely in commercial culture of animal cells (35–37); however, the superficial aeration velocity in cell culture vessels is always low, usually well below 2.5×10^{-3} m s^{-1}. Supplementing the culture medium with serum, albumin, viscosity enhancers such as carboxymethyl cellulose, or surfactants such as Pluronic F68 is known to suppress the damaging effects of sparged aeration in cell culture (12,38).

As a rule, serum supplementation of basal animal cell culture media has little effect on mass transfer under conditions that are typically used. Properties of serum-supplemented media are usually quite similar to those of serum-free basal media. At 37 °C the surface tension of a typical cell culture medium with serum is about 0.060 to 0.064 N m^{-1} as compared to 0.071 N m^{-1} for water, and the viscosity is about 0.75×10^{-3} Pa s^{-1}. This viscosity value is satisfactory for media containing 5 to 10% serum. In submerged aeration, serum-containing media tend to foam, especially if a fine-pore sparger is used for aeration (35); however, fine-pore aeration is not the norm.

Stirred Fermenters. The gas–liquid volumetric mass transfer coefficient in stirred vessels has generally shown good correlation with agitation power input and the superficial gas velocity; thus

$$k_L a_L = \alpha (P_G/V_L)^{\beta} U_G^{\gamma} \quad (57)$$

where P_G/V_L (W m^{-3}) is the gassed power input and U_G (m s^{-1}) is the superficial aeration velocity. Generally, $0.4 < \beta < 1$ and $0 < \gamma < 0.7$ (6). For pure water, α, β, and γ values are 2.6×10^{-2} ($s^{\beta+\gamma-1}m^{3\beta-\gamma}J^{-\beta}$), 0.4, and 0.5, respectively, when $500 < P_G/V_L$ (W m^{-3}) $< 10{,}000$ (6). These values have been shown to apply over a wide scale: $0.002 \leq V_L$ (m^3) ≤ 2.6. In strongly ionic aqueous solutions, the α, β, and γ values become 2.0×10^{-3} ($s^{\beta+\gamma-1}m^{3\beta-\gamma}J^{-\beta}$), 0.7, and 0.2, respectively, when $500 < P_G/V_L$ (W m^{-3}) $< 10{,}000$ and $0.002 \leq V_L$ (m^3) ≤ 4.4 (6). In pure water and ionic aqueous media, the volumetric mass transfer coefficient depends only on the specific power input irrespective of the type of stirrer—whether turbines, propellers, paddles, or self-sucking agitators—used for agitation. The number of impellers does not matter and neither does the location so long as it is within the usual ranges. Similarly, the height of liquid and the type of sparger have negligible influence under representative operating conditions (6). The parameters values noted above assume an absence of flooding. Maximum aeration rates that still ensure unflooded operation depend on the geometry of the impeller, as discussed by Nienow (39).

For single-impeller stirred air–water system, Bakker et al. (40) have recommended the following values of the constants in equation 57: $\beta = \gamma = 0.6$ and $\alpha = 0.015 \pm 0.005$. The parameters noted apply only to the nonflooded state. At high gas flow rates, concave-bladed gas dispersion impellers (Fig. 5) yield at least 20% higher $k_L a_L$ than obtainable with Rushton turbines under the same conditions (40). The constants in equation 57 can be strongly affected by fluid properties, presence of surfactants, and insoluble oils. For fluids other than water and for multiimpeller systems, Bakker et al. (40) recommended measuring α, β, and γ in a geometrically similar small-scale stirred fermenter. This data can be used in equation 57 for scale-up.

For fermentations of *Endomyces* sp. in stirred fermenters, α, β, and γ values in equation 57 have been reported to be 4.015×10^{-2} ($s^{\beta+\gamma-1}m^{3\beta-\gamma}J^{-\beta}$), 0.33, and 0.56, respectively, over a broad scale ($0.03 \leq V_L$ (m^3) ≤ 50). In a 5-L stirred fermenter operated with high-density culture of a recombinant *Escherichia coli*, Shin et al. (41) reported the following equation for $k_L a_L$

$$k_L a_L = 0.0195(P_G/V_L)^{0.55} U_G^{0.64}(1 + 2.12X + 0.20X^2)^{-0.25} \quad (58)$$

where X (g L^{-1}) is the biomass concentration. The X values ranged over 0 to 75 g L^{-1}, and the $k_L a_L$ values during culture varied over (10 to 25) $\times 10^{-3}$ s^{-1}.

In stirred-tank bioreactors for animal cell culture the type of impeller used does not affect $k_L a_L$ so long as mixing is sufficient. For power inputs that are typical (<25 W m^{-3}), better mixing is attained with relatively large ($d_i/d_T \sim 0.5$ to 0.6), slow moving, axial flow impellers. For typically used conditions, the $k_L a_L$ should be calculated as for bubble columns (equation 60) because for impeller power inputs less than about 40 W m^{-3}, the impeller has little effect on $k_L a_L$. Stirred-tank design considerations for animal cell culture have been discussed by Chisti (35) and others (25).

Some of the main correlations for estimating gas–liquid volumetric mass transfer coefficient in stirred microbial fermenters are summarized in Table 10, which includes equations suitable for relatively viscous power law broths. Methods for estimating the power input in mechanically agitated fermenters have been detailed by Chisti and Moo-Young (7) and Nienow (39).

Bubble Columns. For a given fluid in bubble columns, the $k_L a_L$ is affected mainly by the superficial gas velocity, which also determines the specific power input (5). The specific power input can be calculated using the equation

$$\frac{P_G}{V_L} = \rho_L g U_G \quad (59)$$

where V_L is the volume of the liquid or slurry and ρ_L is its density. Properties of the fluid also affect $k_L a_L$. In waterlike fluids, the aspect ratio of bubble columns does not affect $k_L a_L$ so long as the column diameter exceeds about 0.1 m (5). Similarly, liquid superficial velocities do not affect $k_L a_L$ provided $U_L \leq 0.1$ m s^{-1}.

For power inputs that are typically used in microbial culture, the $k_L a_L$ in relatively coalescing media is little influenced by the design of the sparger because the turbulence in the fluid controls the bubble size. In strongly ionic low-viscosity media, bubbles do not coalesce easily, and the bubble size depends on the characteristics of the gas sparger. Porous plate spargers and those with smaller holes produce smaller bubbles and higher $k_L a_L$ values; however, small bubbles are rapidly exhausted of oxygen and have long residence times in the fluid. Small bubbles are produced also under highly turbulent conditions irrespective of the sparger hole size. Such bubbles are especially persistent in viscous media because the viscous drag prevents them from leaving the fluid. Persistent microbubbles are seen in broths of *Streptomyces*, mycelial fungi, and certain fermentations that produce extracellular polymers such as xanthan.

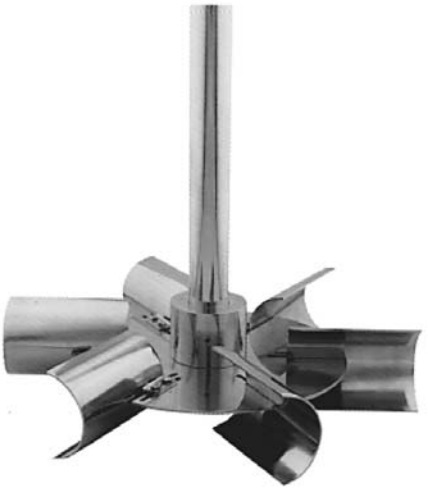


Figure 5. A concave-blade impeller for dispersing gases in liquids. *Source*: Courtesy of Chemineer, Inc.

Table 10. Mass Transfer Correlations for Stirred Tanks

Correlation	Ranges
1. Taguchi and Yoshia (42) $\frac{k_L a_L}{N} = 0.113\left(\frac{d_T^2 h_L}{W_i(d_t - W)}\right)^{1.437}(Nt_m)^{-1.087}(d_i/d_T)^{1.021}$	Pseudoplastic paper pulp slurries (1.6% w/v pulp); disk turbine stirred vessel
2. Yagi and Yoshida (43) $\frac{k_L a_L d_i^2}{D_L} = 0.06\left(\frac{d_i^2 N\rho_L}{\mu_L}\right)^{1.5}\left(\frac{d_i N^2}{g}\right)^{0.19}\left(\frac{\mu_L}{\rho_L D_L}\right)^{0.5}\left(\frac{\mu_L U_G}{\sigma_L}\right)^{0.6} \times \left(\frac{Nd_i}{U_G}\right)^{0.32}[1 + 2\sqrt{\lambda N}]^{-0.67}$ For non-Newtonian fluids use apparent viscosity instead of μ_L. λ is the characteristic time or relaxation time of viscoelastic fluids; $\lambda = 0$ for nonviscoelastics	Viscous Newtonian and non-Newtonian liquids, including viscoelastic fluids
3. Perez and Sandall (26) $\frac{k_L a_L d_i^2}{D_L} = 21.2\left(\frac{\rho_L N d_i^2}{\mu_{ap}}\right)^{1.11}\left(\frac{\mu_{ap}}{\rho_L D_L}\right)^{0.5}\left(\frac{U_G d_i}{\sigma_L}\right)^{0.447}\left(\frac{\mu_G}{\mu_{ap}}\right)^{0.694}$ $\mu_{ap} = K(11N)^{n-1}\left(\frac{3n + 1}{4n}\right)^n$	Carbon dioxide absorption in water and non-Newtonian power law solutions
4. Kawase and Moo-Young (44) $k_L a_L = 0.675\sqrt{D_L}\frac{\rho_L^{3/5}E^{9+4n/10(1+n)}}{(K/\rho_L)^{1/2(1+n)}\sigma_L^{3/5}}\left(\frac{U_G}{U_B}\right)^{0.5}\left(\frac{\mu_{ap}}{\mu_w}\right)^{-0.25}$	Newtonian and non-Newtonian power law fluids
5. Niebelshütz (15) $k_L a_L\left(\frac{V_L}{Q_G}\right)Sc^{0.3} = 0.103\left(\frac{P_G}{Q_G\rho_L(g\mu_L/\rho_L)^{2/3}}\right)^{0.53}$	Non-Newtonian power law fluids such as broths of *Aspergillus niger*, *Penicillium chrysogenum*, and xanthan fermentation; $0.035 \le V_L(m^3) \le 0.070$, $0.02 \le \mu_L(Pa\ s) \le 3.1$; μ_L was calculated at shear rate $= 10^4\ s^{-1}$
6. Reuss (15) $\frac{k_L a_L}{U_G}\left(\frac{\mu_L^2}{g\rho_L^2}\right)^{1/3}Sc^{0.3} = 5.5 \times 10^{-4}\left(\frac{P_G}{V_L U_G \rho_L g}\right)^{0.7}$	Newtonian broths of *Saccharomyces cerevisiae* and gluconic acid, glucose solutions, $\mu_L \le 12 \times 10^{-3}$ Pa s, $0.05 \le V_L(m^3) \le 1.9$
7. Chandrasekharan and Calderbank (45) $k_L a_L = \left(\frac{0.0248}{d_T^4}\right)\left(\frac{P_G}{V_L}\right)^{0.551}Q_G^{0.551/\sqrt{d_T}}$	Air and water

In coalescence-inhibiting media typical of animal cell culture, the $k_L a_L$ in bubble columns at the very low superficial gas velocities that are normally used depends on the sparger hole diameter; the $k_L a_L$ declines as the sparger hole diameter increases over the range of 0.01 to 2 mm (25). Relative to values in clean fluids, surfactants and proteins reduce $k_L a_L$, but the reduction is greater in finer dispersions (i.e., in small, ridged bubbles) generated by small-hole spargers than in dispersions of larger, circulating bubbles. For aeration rates relevant to cell culture, the $k_L a_L$ for water may be calculated (25) with the equation

$$\frac{k_L a_L}{U_G}\left(\frac{\mu_L^2}{\rho_L^2 g}\right)^{1/3} = \alpha[(\mu_L g/\rho_L)^{-1/3}U_G]^{-\beta} \qquad (60)$$

The values of α and β for use with various sparger hole sizes are noted in Table 11. The $k_L a_L$ values in culture media are only about 50% of those obtained with equation 60 for water. Equation 60 was developed in bubble columns with aspect ratios $1 \le (h_L/d_T) \le 3$ when the specific power inputs were $3 \le P_G/V_L\ (W\ m^{-3}) \le 38$.

Although the overall $k_L a_L$ in tall bioreactors may be little affected by the height, the overall mass transfer rates tend to be greater in taller systems because the mass transfer driving force increases with increasing hydrostatic pressure. Mass transfer in bubble columns is discussed further by Shah et al. (46), Chisti (5,14), Deckwer (47), and Deckwer and Schumpe (48). The various correlations available for estimating $k_L a_L$ or k_L in bubble columns are summarized in Table 12. Those correlations apply to the relatively high gas flow rates that are typical of microbial fermentations. The correlation of Akita and Yoshida (51) is recommended for viscosities less than about 21×10^{-3} Pa s. The correlation of Kawase et al. (54) (Table 12) has also agreed well with independent measurements (59).

Airlift Bioreactors. Because of superior oxygen transfer performance, airlift bioreactors have given excellent productivities with yeasts, bacteria, and filamentous fungi (5,14). Processes that produce highly viscous broths, including several biopolymer-producing fermentations, have been successfully carried out in airlift devices (14). Similarly, many hybridoma cultures and plant cell suspensions have given good results. As a general rule, the volumetric productivity of airlift bioreactors equals or betters that of conventional stirred tanks (14). Typically, this level of performance is achieved at substantially lower power input than in stirred vessels. Furthermore, the probability of me-

Table 11. Values of α and β for Use in Equation 60

Sparger	d_H (mm)	α	β	Range
Frit	0.01	6.0×10^{-5}	0.4	$0.001 \leq U_G(\mu_L g/\rho_L)^{-1/3} \leq 0.1$
	0.02	6.3×10^{-5}	0.2	
Perforated pipe	0.5	4.2×10^{-5}	0.2	$0.01 \leq U_G(\mu_L g/\rho_L)^{-1/3} \leq 0.2$
	2.0	2.2×10^{-5}	0.17	

Source: Data from Ref. 25.

chanical failure and likelihood of loss of sterility are lower with airlift bioreactors. Airlift devices outperform bubble columns in suspending solids. Relative to bubble columns, the airlift reactors have a better-defined flow pattern; airlift systems are better than bubble columns in heat-transfer capabilities (14). The operational range of airlift bioreactors tends to be generally broader than similar bubble columns. Unlike in bubble columns, the overall height h_L of liquid in airlift reactors affects k_La_L because the height strongly influences the liquid circulation velocity and, hence, the gas holdup (14).

Many equations have been developed to relate the apparent viscosity of non-Newtonian media to the k_La_L in airlift reactors. Most authors agree that k_La_L generally declines with increasing apparent viscosity; however, there is great discrepancy on the precise magnitude of the viscosity effect (5,14). The exponent on the apparent viscosity term has been given variously as -0.255 (60), -0.66, -0.89, and other values (see Ref. 5). Such discrepancies are associated in part with the questionable practice of defining shear rate as an arbitrary linear function of the gas velocity (5,11). The viscosity-associated decline in k_La_L is especially pronounced for viscosities exceeding about 2.5×10^{-2} Pa s. In highly viscous non-Newtonian fermentations, the use of a supplementary axial flow impeller in the draft tube of airlift reactors is known to enhance fermentation performance.

Table 13 lists the correlations available for estimating the gas–liquid mass transfer coefficient in airlift bioreactors. These correlations are suitable only for initial estimates because the hydrodynamic and transport behavior of airlift devices is quite sensitive to the type of fluid, the reactor geometry, the distribution of gas holdup in the various zones, and the induced liquid circulation rate. Because of the numerous influences, a superior approach to prediction of mass transport for a more reliable design and scale-up relies on data measured at small scale. This method is discussed in the next section.

A General Method of Predicting k_La_L for Scale-up. Although various correlations can be used to directly estimate k_La_L, the reliability of such estimates is often quite poor. A more reliable prediction of k_La_L for process scale-up is based on methodology that combines fundamental principles with small-scale experimentation. This proved approach is especially suited to prediction of k_La_L in complex media. Thus, based on purely analytical reasoning, the volumetric mass transfer coefficient (k_La_L), gas holdup (ϵ_G), the Sauter mean bubble diameter (d_B), and the true mass transfer coefficient (k_L), have been related (5) as follows

$$\frac{k_L}{d_B} = \frac{k_La_L(1 - \varepsilon_G)}{6\varepsilon_G} \tag{61}$$

Calculations of the k_L/d_B ratio from the measured k_La_L and gas holdup have shown that in bubble columns and airlift devices this ratio is constant for a given fluid irrespective of the aeration rate (5,66). The value of k_L/d_B depends on the properties of the fluid. Thus, in slurries of mycelia-like cellulose fiber particles, k_L/d_B is affected by the concentration of solids (Fig. 6). For aqueous salt solution (0.15 M sodium chloride) with or without suspended cellulose fiber, the k_L/d_B ratio (5) has been expressed as

$$\frac{k_L}{d_B} = 5.63 \times 10^{-5}\left(\frac{g\rho_L^2\sigma_L D_L}{\mu_L^3}\right)^{1/2} e^{-0.131C_S^2} \tag{62}$$

where C_S is the concentration of solids (% w/v). Equation 62 applies to airlifts as well as bubble columns. Chisti's (5) observations of constant k_L/d_B ratio in pneumatically agitated reactors have also been confirmed by others (14).

The observed constancy of k_L/d_B provides a reliable method for predicting k_La_L in large-scale bioreactors from data easily measured in a small-scale model unit (5). This proved technique consists of the following steps:

1. Measure the gas holdup and the k_La_L in a small-model reactor with the fluid of interest over the range of gas flow rates that are of interest.
2. Use the data from step 1 to calculate the constant k_L/d_B ratio (equation 61).
3. Measure or calculate the riser gas holdup (ϵ_{Gr}) in the full-scale reactor (see Refs. 5 and 14).
4. Calculate the k_La_L value for the large reactor: for bubble columns

$$k_La_L = \frac{k_L}{d_B}\frac{\varepsilon_G}{(1 - \varepsilon_G)} \tag{63}$$

and for airlift bioreactors

$$k_La_L = \frac{k_L}{d_B}\frac{\varepsilon_{Gr}}{(1 - \varepsilon_{Gr})}\frac{A_r}{A_r + A_d} \tag{64}$$

In equations 63 and 64, the gas holdup and the geometric parameters A_r and A_d are for the large reactor; the k_L/d_B value is from step 2. The rationale for equations 63 and 64 has been discussed by Chisti (5).

Table 12. Gas–Liquid Mass Transfer Correlations for Bubble Columns

Correlation	Ranges
1. Chisti (5) $k_L a_L = 2.39 \times 10^{-4}(P_G/V_L)^{0.86}$	Air and water, $d_T \geq 0.1$ m, $200 \leq P_G/V_L(\text{W m}^{-3}) \leq 1{,}000$, h_L/d_T up to 25
2. Heijnen and Van't Riet (49) $k_L a_L = 0.32 U_G^{0.7}$	Air and water at 20 °C, $0.08 < d_T(\text{m}) < 11.6$, $0 < U_G(\text{m s}^{-1}) < 0.3$, $0.3 < h_L(\text{m}) < 21$; perforated pipe or perforated plate spargers producing 4–6-mm bubbles
3. Fair (50) $k_L a_D = 3.31\left(\frac{D_L \varepsilon_G}{d_B^2}\right)\left(\frac{\mu_L}{\rho_L D_L}\right)^{1/3}\left(\frac{d_B \rho_L U_G}{\mu_L}\right)^{0.5}$	Newtonian fluids
4. Akita and Yoshida (51) $\frac{k_L a_D d_T^2}{D_L} = 0.6\left(\frac{\mu_L}{\rho_L D_L}\right)^{0.5}\left(\frac{g d_T^2 \rho_L}{\sigma_L}\right)^{0.62}\left(\frac{g d_T^3 \rho_L^2}{\mu_L^2}\right)^{0.31} \varepsilon_G^{1.1}$	Air and water and other Newtonian fluids; $\mu_L \leq 21 \times 10^{-3}$ Pa s, U_G and U_L up to 0.4 m s^{-1}
5. Akita and Yoshida (51) $\frac{k_L d_B}{D_L} = 0.5\left(\frac{\mu_L}{\rho_L D_L}\right)^{0.5}\left(\frac{g d_B^3 \rho_L^2}{\mu_L^2}\right)^{0.25}\left(\frac{g d_B^2 \rho_L}{\sigma_L}\right)^{3/8}$	Air and water and other Newtonian fluids, $\mu_L \leq 21 \times 10^{-3}$ Pa s, U_G and U_L up to 0.4 m s^{-1}
6. Chisti (5) $k_L = 5.63 \times 10^{-5}\left(\frac{g \rho_L^2 D_L \sigma_L}{\mu_L^3}\right)^{0.5} d_B e^{-0.131 C_S^2}$	Air and water, and slurries of mycelia-like fibrous solids (paper pulp) in salt solutions; bubble and churn turbulent flow; $0 \leq C_S(\text{w/v \%}) \leq 3$
7. Hughmark (52) $\frac{k_L d_B}{D_L} = 2 + 0.0187\left[\left(\frac{\mu_L}{\rho_L D_L}\right)^{0.339}\left(\frac{d_B U_G \rho_L}{\mu_L}\right)^{0.484}\left(\frac{d_B g^{1/3}}{D_L^{2/3}}\right)^{0.072}\right]^{1.61}$	Newtonian fluids, $0.0009 \leq \mu_L(\text{Pa s}) \leq 0.152$
8. Nakanoh and Yoshida (53) $\frac{k_L a_D d_T^2}{D_L} = 0.09\left(\frac{\mu_L}{\rho_L D_L}\right)^{0.5}\left(\frac{g d_T^2 \rho_L}{\sigma_L}\right)^{0.75}\left(\frac{g d_T^3 \rho_L^2}{\mu_L^2}\right)^{0.39}\left(\frac{U_G^2}{g d_T}\right)$ $\times \left(1 + \chi\left(\frac{U_B \lambda}{d_B}\right)\right)^{-0.45}$ $\chi = 0$ (inelastic fluids) or 0.133 (elastic fluids). For non-Newtonian power law fluids $\mu_L = \mu_{ap} = K(5{,}000 U_G)^{n-1}$	Newtonian, non-Newtonian, and viscoelastic fluids, $0.005 \leq \mu_L(\text{Pa s}) \leq 0.06$, $U_G < 0.1$ m s^{-1}
9. Schumpe et al. (54) $\frac{k_L a_L d_T^2}{D_L} = 0.021\left(\frac{\mu_L}{\rho_L D_L}\right)^{0.5}\left(\frac{g d_T^2 \rho_L}{\sigma_L}\right)^{0.21}\left(\frac{g d_T^3 \rho_L^2}{\mu_L^2}\right)^{0.60}\left(\frac{U_G}{\sqrt{g d_T}}\right)^{0.49}$ μ_L is the apparent viscosity of the power law fluid	Non-Newtonian media
10. Kawase et al. (55) $\frac{k_L a_L d_T^2}{D_L} = 0.555\left(\frac{K d_T^{1-n}}{\rho_L D_L U_G^{1-n}}\right)^{0.5}\left(\frac{g d_T^2 \rho_L}{\sigma_L}\right)^{3/5}\left(\frac{d_T^n U_G^{2-n}}{K/\rho_L}\right)$ $\times \left(\frac{U_G^2}{g d_T}\right)\frac{(2+n)(2n-3)}{20(1+n)^2}\varepsilon_G^{0.5}$	Newtonian and non-Newtonian power law fluids. For Newtonian fluids $K = \mu_L$ and $n = 1$
11. Hikita et al. (56) $k_L a_L = \frac{14.9 g f}{U_G}\left(\frac{U_G \mu_L}{\sigma_L}\right)^{1.76}\left(\frac{\mu_L^4 g}{\rho_L \sigma_L^3}\right)^{-0.248}\left(\frac{\mu_G}{\mu_L}\right)^{0.243}\left(\frac{\mu_L}{\rho_L D_L}\right)^{-0.604}$ f is 1.0 for nonelectrolytes; $f = 10^{0.0681 I}$ for ionic strength $I < 1.0$ kg ion m^{-3}; and $f = 1.114 \times 10^{0.021}$ for $I > 1.0$ kg ion m^{-3}	Low viscosity Newtonian media: water and various gases (air, oxygen, carbon dioxide, hydrogen, methane), sucrose solution, aqueous electrolytes, alcohols, $0.0008 \leq \mu_L(\text{Pa s}) \leq 0.011$, $0.042 \leq U_G(\text{m s}^{-1}) \leq 0.38$
12. Deckwer et al. (57) $k_L a_D = 3.15 \times 10^{-3} U_G^{0.59}[K(5{,}000 U_G)^{n-1}]^{-0.84}$	Non-Newtonian power law fluids, slug flow ($U_G \geq 0.02$ m s^{-1})
13. Godbole et al. (58) $k_L a_D = 8.35 \times 10^{-4} U_G^{0.44}[K(5{,}000 U_G)^{n-1}]^{-1.01}$	Newtonian power law fluids

Figure 7 compares the measured $k_L a_L$ values with prediction of the foregoing procedure for 11 reactors, including bubble columns and external- and internal-loop airlift devices. The comparison spans waterlike media as well as pseudoplastic slurries of cellulose fiber. In all cases, the predictions agree remarkably well with the measurements (Fig. 7). Note that the Solka Floc cellulose fiber slurries in aqueous salt solutions used in Figures 6 and 7 were meant to simulate the fermentation broths of mycelial fungi and filamentous bacteria (5,66). Such broths are usually

Table 13. Gas–Liquid Volumetric Mass Transfer Coefficient Correlations for Airlift Bioreactors

Configuration	Equation	Ranges
Concentric-tube internal-loop vessels (annulus sparged)	$\frac{k_L a_D \sigma_L}{D_L g \rho_L} = 2.25\left(\frac{\mu_L}{\rho_L D_L}\right)^{0.500}\left(\frac{\rho_L \sigma_L^3}{g\mu_L^4}\right)^{0.136} \times \left(\frac{d_H}{d_T}\right)^{-0.0905} \varepsilon_G^{1.26}$ Average estimation error was 12% for 175 measurements Koide et al. (61)	Newtonian media, $3.71 \times 10^2 \le \mu_L/\rho_L D_L \le 6.00 \times 10^4$, $1.18 \times 10^6 \le \rho_L \sigma_L^3/g\mu_L^4 \le 5.93 \times 10^{10}$, $0.471 \le d_i/d_T \le 0.743$, $7.14 \times 10^{-3} \le d_H/d_o \le 2.86 \times 10^{-2}$, $0.0302 \le \varepsilon_G \le 0.305$, $6 \le$ aspect ratio ≤ 15, $0.52 \le A_d/A_r \le 1.23$
Concentric-tube internal-loop vessels (draft-tube sparged)	$\frac{k_L a_D d_T^2}{D_L} = 2.66\left(\frac{\mu_L}{\rho_L D_L}\right)^{0.500}\left(\frac{g d_T^2 \rho_L}{\sigma_L}\right)^{0.715}\left(\frac{g d_T^3 \rho_L^2}{\mu_L^2}\right)^{0.251} \times \left(\frac{d_i}{d_T}\right)^{-0.429} \varepsilon_G^{1.34}$ Koide et al. (62)	Newtonian media, $\rho_L = 997 - 1182$ kg m^{-3}, $\mu_L = (0.894 - 17.0) \times 10^{-3}$ Pa s, $\sigma_L = (51.7 - 73.0) \times 10^{-3}$ N m^{-1}, $D_L = (0.145 - 2.42) \times 10^{-9}$ m^2 s^{-1}, $A_r/A_d = 0.3$–0.8, aspect ratio = 6–15
External-loop reactor	$\frac{k_L a_L L}{U_{Lr}} = 14.5\left(\frac{U_{Gr}}{U_{Lr}}\right)^{0.83} Pe^{-0.6}$ Pe value is for the entire loop Verlaan et al. (63)	Air and water in 0.165 m^3 reactor, aspect ratio = 16 (based on riser), $A_r/A_d = 4$, no gas in downcomer, $0.01 \le U_{Gr}$(m s^{-1}) ≤ 0.14, $Pe = U_L L/E_L$, $Pe = 40$–60 (increased with gas flow rate)
A bubble column and a draft-tube sparged concentric-tube airlift vessel	$\frac{k_L a_L d_T^2}{D_L} = 0.018\left(\frac{K(2{,}800U_G)^{n-1}}{\rho_L D_L}\right)^{0.5}\left(\frac{g d_T^2 \rho_L}{\sigma L}\right)^{0.20} \times \left(\frac{U_G}{\sqrt{g d_T}}\right)^{0.51}\left(\frac{g d_T^3 \rho_L^2}{[K(2{,}800U_G)^{n-1}]^2}\right)^{0.62} \times (1 + 0.12\text{Wi})^{-1}$ Suh et al. (64)	Non-Newtonian xanthan fermentation broths of the obligate aerobic bacterium *Xanthomonas campestris*. Bubble column (0.05 m^3, aspect ratio $\approx$ 19); airlift bioreactor (1.2 m^3, $A_r/A_d = 1$, aspect ratio $\approx$ 12)
Concentric-tube internal-loop (draft-tube sparged)	$k_L a_L = 3.43 \times 10^{-2} U_{Gr}^{0.524} \mu_{ap}^{-0.255}$ μ_{ap} was calculated by assuming the shear rate to equal $5{,}000 \times U_{Gr}$ for $U_{Gr} \ge 0.04$ m s^{-1}, or $1{,}000 \times (U_{Gr})^{0.5}$ for $U_{Gr} < 0.04$ m s^{-1} Li et al. (60)	Aqueous carboxymethyl cellulose ($K = 0.286$–11.5 Pa s^n; $n = 0.441$–0.617), 0.055 m^3 vessel, $A_r/A_d = 0.618$, aspect ratio $\sim$ 26. $0.020 \le \mu_{ap}$ (Pa s) ≤ 0.85
Concentric-tube internal-loop reactors (draft-tube sparged)	$\frac{k_L a_D d_T^2}{D_L} = 2.66\left(\frac{\mu_L}{\rho_L D_L}\right)^{0.500}\left(\frac{g d_T^2 \rho_L}{\sigma_L}\right)^{0.715}\left(\frac{g d_T^3 \rho_L^2}{\mu_L^2}\right)^{0.251} \times \left(\frac{d_i}{d_T}\right)^{-0.429} \varepsilon_G^{1.34}\left(1 + 0.099\left(\frac{C_S}{\rho_S}\right)^{0.069}\right) \times \left(\frac{\rho_L \sigma_L^3}{g\mu_L^4}\right)^{0.023}\left(\frac{U_t}{U_G}\right)^{0.046}\Bigg)^{-1}$ Average estimation error was within 17% for 383 measurements Koide et al. (62)	Suspensions of glass or bronze spheres, $3.71 \times 10^2 \le \mu_L/\rho_L D_L \le 9.92 \times 10^4$, $1.36 \times 10^3 \le g d_T^2 \rho_L/\sigma_L \le 1.22 \times 10^4$, $1.29 \times 10^8 \le g d_T^3 \rho_L^2/\mu_L^2 \le 1.26 \times 10^{11}$, $0.471 \le d_i/d_T \le 0.743$, $3.99 \times 10^{-2} \le \varepsilon_G \le 2.73 \times 10^{-1}$, $1.69 \times 10^{-11} \le g\mu_L^4/\rho_L \sigma_L^3 \le 2.55 \times 10^{-6}$, $0 \le C_S/\rho_S \le 8.00 \times 10^{-2}$, $1.17 \times 10^{-2} \le U_t/U_G \le 0.844$, $A_r/A_d = 0.3$–0.8, aspect ratio = 6–15
Concentric-tube internal-loop reactors (draft-tube sparged)	$\frac{k_L a_D d_T^2}{D_L} = \frac{4.04}{1 + 2\varphi_S^{1.3}}\left(\frac{\mu_L}{\rho_L D_L}\right)^{0.5}\left(\frac{g d_T^2 \rho_L}{\sigma_L}\right)^{0.67}\left(\frac{g d_T^3 \rho_L^2}{\mu_L^2}\right)^{0.26} \times \left(\frac{d_i}{d_T}\right)^{-0.047} \varepsilon_G^{1.34}$ Average error of estimation was 14% for 260 data; U_G is based on the diameter d_T of the outer column; the $k_L a_D$ values in water and salt solutions were similar Koide et al. (65)	Suspensions of relatively low-density calcium alginate beads in water, aqueous glycerol, and aqueous salt solutions (0.10 or 0.27 kmol m^{-3} barium chloride; 0.4 kmol m^{-3} sodium sulfate), 0–20% v/v solids, 1.88–3.98-mm bead diameter, $A_r/A_d = 0.3$–1.2, aspect ratios = 6–16, $\mu_L = (0.894$–$12.5) \times 10^{-3}$ Pa s, D_L of oxygen $= (0.194$–$2.42) \times 10^{-9}$ m^2 s^{-1}, $3.71 \times 10^2 \le \mu_L/D_L\rho_L \le 5.52 \times 10^4$, $2.66 \times 10^3 \le \rho_L d_T^2 g/\sigma_L \le 1.22 \times 10^4$, $2.35 \times 10^8 \le \rho_L^2 d_T^3 g/\mu_L^2 \le 3.29 \times 10^{11}$, $0.471 \le d_i/d_T \le 0.743$, $1.69 \times 10^{-11} \le Mo \le 6.67 \times 10^{-7}$, $3.79 \times 10^{-2} \le \varepsilon_G \le 2.24 \times 10^{-1}$, $0 \le \varphi_S \le 0.2$; where the Morton number is $Mo = \frac{g(\rho_L - \rho_G)}{\sigma_L^3 \rho_L^2} K^4\left(\frac{8U_{Lr}}{d_r}\right)^{4(n-1)}\left(\frac{3n+1}{4n}\right)^{4n}$

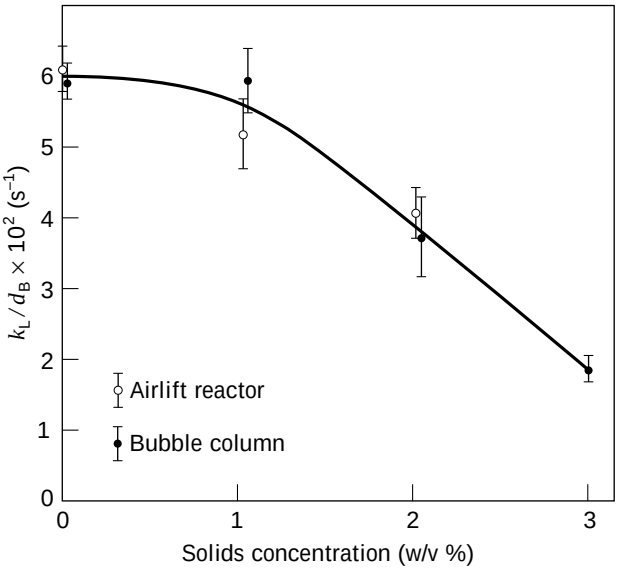


Figure 6. Effect of concentration of mycelia-like cellulose fiber particles on the k_L/d_B ratio in aqueous salt solution (0.15 M sodium chloride) (5).

pseudoplastic. The non-Newtonian flow properties of cellulose fiber suspensions are quite close to those of the typical fungal fermentation broths (5).

For the low aeration rates typical of animal cell culture, the k_La_L in an internal-loop draft tube airlift vessel is almost identical to that in an equivalent bubble column; hence, equation 60 can be used for an estimation. Although the gas is sparged only in the draft tube or in the annulus of an airlift bioreactor, for estimation of k_La_L with equation 60 the superficial gas velocity should be calculated using the cross-sectional area of the outer column. For conditions typical of animal cell culture, whether the gas is sparged in the draft tube or the annulus does not affect k_La_L so long as the outer-to-inner tube diameter ratio is 1.6 to 2.1 and $2 \leq h_L/d_T \leq 5$. Large concentric draft-tube types of airlift bioreactors are used in commercial culture of animal cells (36,37). Various aspects of design of airlift bioreactors have been exhaustively treated in the literature (5,14). A few specifics are discussed in the following sections.

Axial Dissolved Oxygen Profiles. As clearly demonstrated (5), the steady-state dissolved oxygen concentration varies axially in tall bubble columns and airlift reactors even in the absence of an oxygen-consuming reaction. The exact nature of the concentration profile depends on the state of mixing in the reactor. During culture of baker's yeast in a draft-tube sparged concentric tube internal-loop airlift reactor ($A_r/A_d = 0.8$; aspect ratio ≈ 10; 0.25 m^3), Russell et al. (67) observed that the steady-state dissolved oxygen concentration in the riser and the downcomer increased with increasing aeration rate. The dissolved oxygen levels were always several-fold higher in the downcomer than in the riser (67). In the riser, the dissolved oxygen levels tended to be 30% or less of air saturation. The oxygen concentration did not vary radially either in the riser or in the downcomer. Similarly, no significant variations were observed axially in the downcomer (67).

Using the annulus-sparged mode of the same vessel that was used by Russell et al. (67), Pollard et al. (68) found that the axial-dissolved oxygen profiles in the riser showed

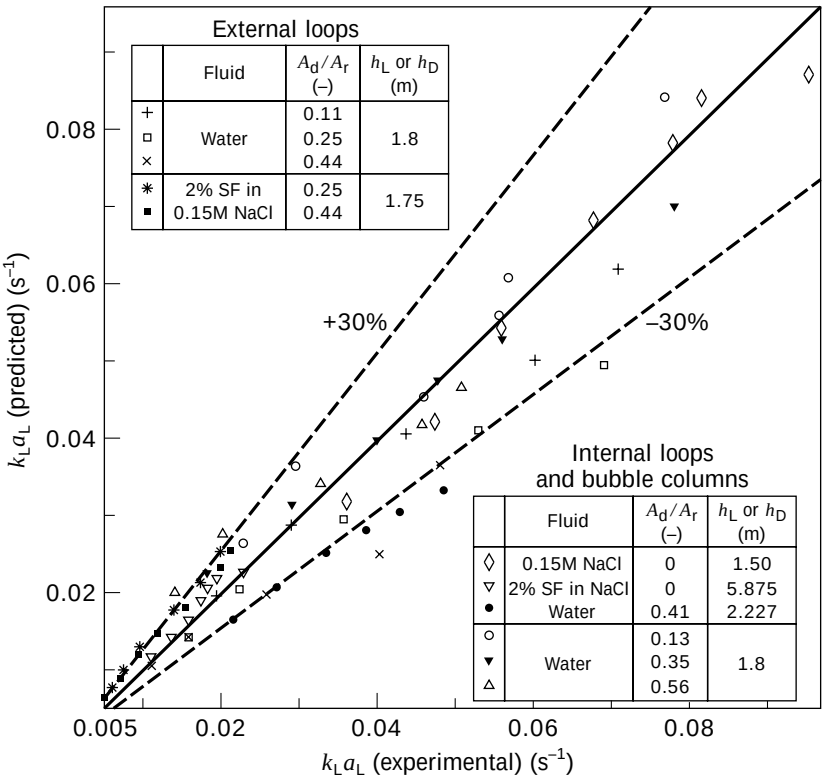


Figure 7. Predicted versus measured overall volumetric gas–liquid mass transfer coefficient in bubble columns, external-loop and internal-loop airlift devices in Solka Floc (SF) cellulose fiber slurries and solids-free media (5). The solid line represents exact prediction.

a constant zero level of oxygen over a broad range of superficial gas velocities (0.036 to 0.136 m s^{-1}) during baker's yeast fermentations. Surprisingly, the dissolved oxygen concentration increased down the downcomer, peaking close to the bottom of the draft tube and declining sharply to zero as the fluid moved into the riser. These anomalous results seemed to indicate better oxygen transfer in the downcomer relative to the riser (68). In view of the observed profile, the cells experienced extreme dissolved oxygen heterogeneity in the reactor; however, there seemed to be no effect on the culture performance. Installation and operation of a marine propeller (diameter = 0.16 m; 1000 rpm) near the bottom of the draft tube, combined with a high level of aeration (0.358 m^3 min^{-1}), enhanced liquid circulation and oxygen transfer so much so that the dissolved oxygen level did not drop below 35% of saturation anywhere in the reactor.

Contributions of Various Zones to Overall k_La_L. Many airlift reactors consist of two hydrodynamically distinct zones, the riser and the downcomer, that contribute differently to the observed mass transfer (5). Sometimes a third zone, the gas–liquid separator (69), may also exist with its own contribution toward mass transfer.

In airlift bioreactors that do not have gas–liquid separators, the measured k_La_L consists of contributions from the riser and the downcomer zones; thus

$$(k_La_L)_{\text{measured}} = \frac{A_r(k_La_L)_r + A_d(k_La_L)_d}{A_r + A_d} \tag{65}$$

Equation 65 is an analytical expression that relates the measured k_La_L to the separate, hypothetical contributions of the riser and the downcomer (5). Equation 65 leads to the general conclusion that for any airlift device in which gas–liquid mass transfer is confined to the riser (i.e., no gas in the downcomer), the actual volumetric mass transfer coefficient in the riser would be larger than the apparent measured value; the difference being dependent on the A_r/A_d ratio. Equation 65 applies to reactors in which the riser and the downcomer have approximately equal heights. In cases in which a significant portion of the liquid resides in the head zone, equation 65 should be modified to account for that volume (14). Generally, $(k_La_L)_d < (k_La_L)_r$.

Based on data obtained in several concentric draft-tube airlift reactors that were sparged either in the draft tube or the annulus (aspect ratios = 6 to 11; U_G = 0.019 to 0.12 m s^{-1}), Koide et al. (70) suggested that in air and water the contribution of the downcomer zone to gas–liquid mass transfer was "relatively large." Furthermore, in liquids such as aqueous salt solutions in which the bubbles are smaller, the riser and the downcomer were claimed to contribute to the total mass transfer in proportion to the fraction of the total volume that those zones occupied (70). These observations in relatively small reactors are not entirely consistent with other findings (5,14).

Mass Transfer Enhancement by Static Mixers. Installation of static mixer elements in airlift and bubble column bioreactors is known to enhance the gas–liquid mass transfer coefficient relative to mixer-free operation (14,71). The enhancement is mostly the result of an increase in the gas–liquid interfacial area that is brought about by the bubble-breaking action of the mixer elements. Improved productivity of cephalosporin C by *Cephalosporium acremonium* was observed (72) in an airlift reactor when static mixers were installed.

The mass-transfer-enhancing effect of static mixers is most pronounced in viscous fluids that give rise to large spherical cap bubbles (71); less dramatic, but still large improvements are observed with waterlike media. Generally, for a given fluid, the mass transfer coefficient and the gas velocity in the riser of a given airlift reactor may be correlated in the form

$$k_La_L = \alpha U_{Gr}^{\beta} \tag{66}$$

For a given fluid, the exponent β in equation 66 is not affected by the presence of static mixers in the reactor; however, the mixers significantly enhance the value of α (71). The improvement in α depends mainly on the consistency index (K) or the thickness of the fluid; generally, the thicker the fluid, the greater the effect of static mixers (71). The approximate relationship among α, α_M (the α-value in presence of static mixers), and K has been established (71) to be

$$\frac{\alpha_M}{\alpha} = 4.43K^{0.12} \tag{67}$$

Equation 67 was determined for approximately 10^4-fold variation in K. The flow behavior index n (varied between 0.5 and 1.0) had no direct impact on enhancement of k_La_L by the mixers (71). Depending on the fluid, the static mixers were found to enhance k_La_L by 30 to 500% relative to mixer-free operation (71). The constants in equation 67 are expected to depend on the type of static mixer used.

Installation of mixing elements in airlift bioreactors slows down the circulation of fluid; hence, fewest possible mixing elements should be used to satisfy the mass transfer needs while limiting the loss of liquid circulation. The optimal interval between elements depends on the type of fluid. The more coalescing the fluid, the closer must the mixers be positioned for maximum performance. There is an upper limit to the viscosity that may be used in airlift reactors with static mixers (71). In excessively viscous media, just a few mixing elements can slow down the liquid so much so that the reactor stagnates.

Reciprocating Jet Fermenter. A reciprocating plate or reciprocating jet fermenter has been used in a few laboratory studies. The device is similar to a pulsed extraction column. It consists of a cylindrical tank containing a central shaft that supports a frame of evenly spaced perforated plates or disks. The plates are mounted horizontally relative to the axis of the shaft. The plates or baffles remain clear of the fermenter walls; the interdisk space and the diameter of perforation varies. Typically, disks are placed 0.07 to 0.10 m apart and the perforations, 0.06 to 0.08 m in diameter, occupy a large proportion of the disk surface. The stack of disks oscillates up and down through the fluid. The liquid moves through the perforations as jets that change direction as the movement of stack reverses. The frequency of oscillation is usually a few cycles per second,

and the vertical movement is small. Brauer reported on these devices in the early 1980s, and more gas–liquid mass transfer data have since appeared (73–75). The k_La_L in reciprocating jet reactors increases with increasing frequency and amplitude of oscillation of the baffles. The power input also increases. The k_La_L generally follows the equation

$$k_La_L = \alpha(P/V_L)^{\beta}U_G^{\gamma} \tag{68}$$

where α, β, and γ depend on the bioreactor geometry, fluid properties, and the power input range. Some typical values for those parameters are noted in Table 14. Commercial utility of reciprocating plate fermenters remains to be proved.

Tubular Loop Fermenters. Tubular loop fermenters use concurrent flow of gas and liquid in a pipe. The liquid is recirculated by a pump. This design is not practical except for low-volume fermentations or unusual situations such as seen in the tubular solar receivers of photobioreactors used in culturing microalgae. For a tubular loop (0.025-m tube diameter, 22-m length), Ziegler et al. (76) recommended the equation

$$k_La_L = 5.321 \times 10^{-5}(P_G/V_L)^{0.9}U_G^{0.2} + 8.333 \times 10^{-3} \tag{69}$$

Equation 69 was obtained using the sulfite oxidation system.

Fluidized Beds. Gas–liquid–solid fluidized bed bioreactors are seen in immobilized cell culture and wastewater treatment processes. These reactors usually consist of a cylindrical vessel containing a suspension of immobilized biocatalyst. The solids are kept in suspension by continuous flow of a liquid and gas. A solid–liquid separator, usually a sedimentation device, located near the top of the reactor prevents washout of solids. Essential design features of gas–liquid–solid fluidized bed reactors are discussed by Muroyama and Fan (77). The gas–liquid mass transfer coefficients in three-phase fluidized beds may be estimated using the correlation:

$$\frac{k_La_Ld_p^2}{D_L} = \alpha\left(\frac{d_pU_t\rho_L}{\mu_L}\right)^{3.3}\left(\frac{\mu_LU_G}{\sigma_L}\right)^{0.7} \tag{70}$$

$\alpha = 2.33 \times 10^{-5}$ when $Re_t < 2{,}000$; $\alpha = 3.946 \times 10^{-7}$ when $Re_t > 2{,}000$; where U_t is the terminal setting velocity of the particle and Re_t is the Reynolds number based on the terminal settling velocity. Equation 70 was developed for carbon dioxide absorption in water containing porcelain beads, sand, lead shot, or iron shot. The terminal settling velocity of the solids ranged over 0.12 to 0.815 m s^{-1}. Other parameters were $0 \le U_G$ (m s^{-1}) ≤ 0.1; $0.05 \le U_L$ (m s^{-1}) ≤ 0.172; $1.06 \le d_p$ (mm) ≤ 6.84; and $2{,}400 \le \rho_S$ (kg m^{-3}) $\le 11{,}180$.

Another correlation for gas–liquid mass transfer in fluidized beds was proposed by Nguyen-Tien et al. (78):

$$k_La_L = 0.394\left(1 - \frac{\varepsilon_S}{0.58}\right)U_G^{0.67} \tag{71}$$

where ϵ_S is the volume fraction of solids. For three-phase circulating bed fermenters, Loh et al. (79) have recommended the following equation

$$k_La_L = 1.4 \times 10^{-4}\left(\frac{P}{V_L}\right)^{0.91}\frac{(1 - 2.5\varepsilon_S)^{0.95}}{(1 - \varepsilon_S)^{4.3}} \tag{72}$$

Because k_La_L is generally not too sensitive to the liquid flow rate, the equations developed for slurry bubble columns may also be applied to fluidized beds. Consult the work by Muroyama and Fan (77) for further details on gas–liquid mass transfer in fluidized beds.

Other Factors Affecting k_La_L

1. *Surfactants and antifoam agents.* Relative to clean systems, addition of surfactants may reduce or enhance k_La_L. The specific effect and its magnitude depend on the type of surfactant, its concentration, and the nature of the broth. The coalescence promotion effect of some surfactants reduces the specific interfacial area a_L and the k_La_L. Addition of surface-active agents such as ethanol to water generally increases a_L and k_La_L. Surfactants such as sodium dodecyl sulfate (SDS) or sodium lauryl sulfate (SLS) accumulate at the gas–liquid interface and usually cause a reduction in k_L.

2. *Temperature.* The overall volumetric gas–liquid mass transfer coefficient k_La_L generally increases with increasing temperature. This is mainly because the diffusion coefficient increases with temperature. Also, the viscosity of the liquid declines with increasing temperature; hence, for a given energy input, the film thickness declines and

Table 14. The Parameters in Equation 68 for Pulsed Baffle Fermenters

Fluid	α ($s^{\beta+\gamma-1}m^{3\beta-\gamma}J^{-\beta}$)	β (−)	γ (−)	Ranges
Water	4.86	0.364	1.55	$23 \le P/V_L$(W m^{-3}) $\le 3{,}620$ $0.0049 \le U_G$(m s^{-1}) ≤ 0.0099 $h_L/d_T \approx 40$
Water	0.281	0.632	1.138	$35 \le P/V_L$(W m^{-3}) $\le 6{,}600$ $0.0005 \le U_G$(m s^{-1}) ≤ 0.004 $h_L/d_T \approx 1.87$
Water–glycerol (1:1 by vol)	0.0039	0.735	0.668	$35 \le P/V_L$(W m^{-3}) $\le 6{,}600$ $0.0005 \le U_G$(m s^{-1}) ≤ 0.004 $h_L/d_T \approx 1.87$
Aqueous yeast slurry	4.8	0.268	1.184	$d_T = 0.05$ m

the interfacial area may increase slightly. The k_La_L value at any temperature T (°C) may be calculated from the k_La_L at 20 °C using the equation:

$$(k_La_L)_T = (k_La_L)_{20°C}\theta^{(T-20)°C} \quad (73)$$

where θ is 1.200 to 1.024.

3. *Suspended solids.* How and how much suspended solids affect k_La_L depends on several factors, including concentration of solids, the density difference between the solid and the suspending fluid, the particle diameter, the operating conditions of the reactor (i.e., the power input), and the hydrophobicity of the solids. Up to 15% (w/w) of particles smaller than about 50 μm have little effect on k_La_L; however, much smaller amounts of larger particles can reduce k_La_L significantly. Very small concentrations of relatively high-density solids may actually enhance k_La_L a little. Increasing amounts of low-density filamentous or mycelial biomass rapidly increase the apparent viscosity of the slurry, and the k_La_L may decline sharply (5). The effect of such solids on k_La_L is illustrated in Figure 8 for paper pulp fibers that simulate filamentous mycelial biomass.

In a draft-tube sparged concentric draft-tube airlift reactor for potential application to microbial desulfurization

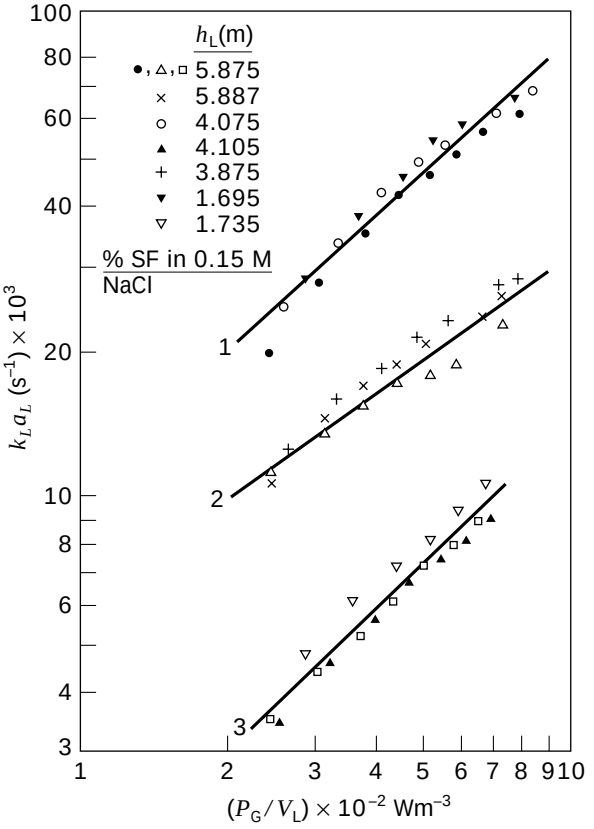


Figure 8. The effect of Solka Floc (SF) cellulose pulp fibers on gas–liquid mass transfer coefficient in bubble columns at various static slurry heights h_L. The fibers were suspended in 0.15 M sodium chloride in tap water (5). The k_La_L is shown as a function of the specific power input.

of coal, the mass transfer coefficient data in a simulated basal salt medium were reported by Smith and Skidmore (80). The medium contained a total of 1.9 kg m^{-3} of various salts in distilled water. Pulverized coal particles (74 × 10^{-6} m particle diameter; density = 1,415 kg m^{-3}) were suspended in the fluid to a concentration of 0 to 40% wt (equivalent to 0 to 0.315 volume fraction of solids in gas-free medium). Over the temperature range 303 to 345 K, the k_La_L-values were strongly enhanced by increasing temperature. Increasing concentration of solids over 0 to 5% wt (equivalent to φ_S of 0 to 0.035) increased the k_La_L values slightly, but further increase in solids loading strongly lowered the mass transfer coefficient (80). This phenomenon was explained as being due to penetration of fine solid particles in the stagnant liquid film around gas bubbles, reducing the effective diffusion length (80). However, as the concentration of solids increased beyond 5% wt, the solids caused a blocking effect that reduced the effective mass transfer area (5,80). In the region of enhanced mass transfer ($0 \le \varphi_S < 0.035$), the k_La_L was correlated (80) with

$$k_La_L = (5.7 - 144.8\varphi_S + 5048.3\varphi_S^2) \times 10^4 U_{Gr}^{1.55} e^{-2990/T} \quad (74)$$

whereas in the blocking region ($0.035 \le \varphi_S \le 0.315$), the following equation (80) was obtained

$$k_La_L = (2.6 - 16.8\varphi_S + 29.4\varphi_S^2) \times 10^6 U_{Gr}^{1.45} e^{-4130/T} \quad (75)$$

Equations 74 and 75 agreed with the data to within about ±20%. These equations applied over the U_{Gr} range (0.154 to 1.39) × 10^{-1} m s^{-1}, when the A_r/A_d and the aspect ratios were 1.3 and about 5, respectively.

Mass transfer measurements in suspensions of agar-filled soft polyurethane foam particles (d_p = 3 mm; ρ_p = 1030 kg m^{-3}) in uninoculated penicillin culture medium (μ_L = 1.9 × 10^{-1} Pa s) in draft tube sparged concentric tube airlift reactors have been reported (81). Increasing volume fraction of solids over the 0 to 0.1 range enhanced the volumetric mass transfer coefficient by 15 to 20% over solids-free operation (81). Further increase in solids loading to 40% (v/v) caused a decline in k_La_L values. The mass transfer coefficient values in airlift reactors were up to threefold higher than in comparable fluidized beds.

In suspensions of relatively low-density calcium alginate beads in water and other Newtonian fluids (0 to 20% v/v solids, 1.88- to 3.98-mm bead diameter), in draft-tube sparged concentric-tube airlift reactors, Koide et al. (65) noted that the mass transfer coefficient declined with increasing concentration of solids, but was not affected by the size of the particles.

Oxygen Transfer in Wastewater Treatment Processes. The typically used bioreactor in activated sludge treatment of wastewater consists of a relatively shallow (~4 m deep) rectangular basin aerated by sparging through perforated pipes or diffusers located at the bottom of the tank. The oxygen transfer capability of such systems is quite limited; hence, biodegradation is slow. Usually only about 0.5 to 2 kg of oxygen can be transferred per kilowatt-hour of energy spent. The oxygen transfer capabilities of other conven-

tional aeration systems have been detailed by Winkler (82).

Faster degradation of pollutants can be achieved with low-volume high-rate oxygen transfer systems. One such technology is the deep shaft reactor based on the airlift principle. This advanced activated sludge process relies on high hydrostatic pressure in a deep airlift column to significantly enhance oxygen transfer. In comparison with conventional processes, oxygen transfer rates are up to 10-fold greater (82). The transfer rate at peak load is about 1 kg oxygen m^{-3} h^{-1} (14). Several factors combine to yield this high level of performance, including long gas–liquid contact times and intense turbulence in the circulating fluid with Reynolds numbers of the order of 10^5 or higher (82). The shaft is 30 to 220 m deep (82), 0.5 to 10 m in diameter, and partitioned vertically into a riser and a downcomer. Air is injected into the downcomer, about 20 to 40 m below the surface (82), except during start-up when the riser is aerated (Fig. 9). To ensure that air bubbles move down the downcomer, a superficial liquid velocity of 1 to 2 m s^{-1} must be generated in the downcomer zone. Because the gas is not recirculated, the downcomer region above the sparger is free of bubbles. Oxygen transfer efficiencies of 3 to 5.5 kg oxygen/kWh can be attained (82). Up to 90% of the oxygen in the air is used up.

For deep shaft plants, Winkler (82) cites a BOD loading of 0.9 kg BOD/kg sludge solids per day. A retention time of about 1.5 h has been mentioned for 92% BOD removal (82). Volumetric BOD removal is of the order of 3.7 to 6.6 kg BOD m^{-3} d^{-1}, which is generally associated with high-rate treatment processes. Sewage is treated without primary sedimentation; only preliminary degritting is needed (82). Sedimentation of grit at the bottom of the shaft is prevented by ensuring that the flow velocity at the bottom exceeds 1 m s^{-1}. See Chisti (14) for additional details.

In addition to the deep shaft, a biotower configuration is increasingly being used in wastewater treatment. The biotower units consist of a relatively shallow (18 to 20 m) above-ground pool of liquid with or without multiple draft tubes (downcomers). The towers handle streams with 2 to 12 kg m^{-3} chemical oxygen demand. The oxygen transfer efficiency ranges over 1.2 to 3.8 kg oxygen/kWh (83). The biomass sludge produced rises to the top with the bubbles and is separated in an integral settling zone. Chisti (14) provides further details.

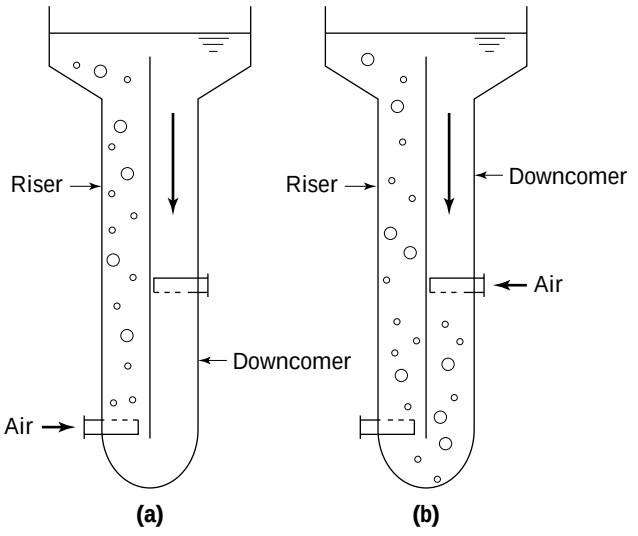


Figure 9. Deep-shaft airlift reactor: (**a**) Aeration in riser during start-up. (**b**) Aeration in downcomer during normal operation. In (**b**), no gas is being injected in the riser; all the gas bubbles in the riser result from circulation from the downcomer.

Liquid–Liquid Mass Transfer

Liquid–liquid dispersions are encountered in solvent extraction and in fermentations of hydrocarbons or other water-immiscible liquid substrates. Liquid–liquid dispersions occur also in extractive fermentations where a product is continuously extracted into a water immiscible phase. Similarly, oxygen supply using water immiscible perfluorocarbons and other fluids utilizes liquid–liquid mass transfer. Whereas in gas–liquid dispersions only the liquid film around the bubble is the principal resistance to mass transfer, in liquid–liquid dispersions the film inside the dispersed drops also affects the transport rate. Tables 15 and 16 list the correlations that are useful in calculating the mass transfer coefficients in the dispersed and continuous phases. Noncirculating small drops can be treated as ridged particles and solid–liquid mass transfer correlations (see later sections) developed for suspended spherical solids can be used. Presence of surfactants often renders drops nonmobile, and again, solid-sphere correlations apply. For larger, mobile, or oscillating droplets, both the continuous phase and the dispersed phase mass transfer coefficients are greater than for solid spheres.

Perfluorocarbons and Oxygen Vectors. Perfluorocarbons are water immiscible liquids that dissolve 10 to 20 times more oxygen than does water. These fluids can be used for bubble-free oxygenation and removal of carbon dioxide in animal cell culture. Another potential application is stripping of inhibitory oxygen produced via photosynthesis in cultures of microalgae. Perfluorocarbons are biologically inert, and suitably selected ones are nontoxic to animal cells and microorganisms. Indeed, emulsified perfluorocarbons have been used to supplement blood to improve oxygen supply to human patients. In microbial culture, per-

Table 15. Mass Transfer Coefficient Inside Mobile or Oscillating Drops in Liquid–Liquid Dispersions

Correlation	Range
$Sh \approx 16.7$	$Re < 50$
$Sh = 0.320 Re^{0.68}\left(\frac{\sigma^3 \rho_C^2}{g\mu_C^4 \Delta\rho}\right)^{0.10}\left(\frac{4D_D t_C}{d_D^2}\right)^{-0.14}$	$Re > 150$–200
$Sh = 0.32 Re^{0.63} Sc^{0.50}\left(1 + \frac{\mu_D}{\mu_C}\right)^{-0.5}$	$10 \le Re \le 1{,}000$ spherical drops
$Sh = 7.5 \times 10^{-5} Re_C^{2.0} Sc^{0.56}\left(1 + \frac{\mu_D}{\mu_C}\right)^{-0.5}$	$100 \le Re_C \le 1{,}500$ larger oblate drops

Note: Re in these equations is based on drop diameter and relative velocity between phases; all other properties are for the drop phase unless otherwise noted. The density and viscosity in Re_C are for the continuous phase. The subscripts C and D refer to continuous and dispersed phases, respectively.

Table 16. Continuous Phase Mass Transfer Coefficient Correlations for Liquid–Liquid Dispersions

Small, noncirculating drops

$Sh = 2 + 0.79Re^{0.5}Sc^{0.33}$
$Sh = 2 + 0.76Re^{0.5}Sc^{0.33}$
$Sh = 0.562Re^{0.5}Sc^{0.33}$

Larger, mobile drops

$$Sh = \frac{2}{\sqrt{\pi}}[1 - Re^{-0.5}(2.89 + 2.15(\mu_D/\mu_C)^{0.64})]^{0.5}(Re \times Sc)^{0.5}$$

when $Re < 120$, $(\mu_D/\mu_C) \leq 2$, and $0 \leq (\rho_D/\rho_C) \leq 4$

$Sh = -126 + 1.8Re^{0.50}Sc^{0.33}$
when $8 \leq Re \leq 800$

$Sh = -178 + 3.62Re^{0.50}Sc^{0.33}$

Note: *Re* in these equations is based on drop diameter and relative velocity between phases; all other properties are for the continuous phase unless otherwise noted. The subscripts *C* and *D* refer to continuous and dispersed phases, respectively.
Source: Based on Pratt (2) and Schügerl (84).

fluorocarbon concentrations as low as 10% v/v have significantly enhanced culture performance.

One scheme for bubble-free oxygenation using perfluorocarbons is illustrated in Figure 10. A separate vessel is used to aerate the organic phase and strip the carbon dioxide. The oxygen-enriched liquid is pumped to the fermenter and sprayed into the culture medium. As the droplets settle to the bottom in the relatively quiescent environment of animal cell bioreactors, the oxygen transfers to the aqueous phase. The oxygen-depleted perfluorocarbon from the bottom of the fermenter is recycled to the external aerator. Perfluorocarbons and other oxygen vectors enhance oxygen transfer performance also when added to an aerated reactor. The oils extract more oxygen from the gas phase and effectively increase oxygen-culture contact time. In agitated continuous culture of *Escherichia coli* operated at a dilution rate of 0.275 h^{-1} and 35.5 °C, an emulsified perfluorocarbon added to the feed at 50% volume fraction could provide 0.17 kg O_2 $m^{-3}h^{-1}$ when the biomass concentration was 0.74 kg m^{-3} (85). The same system removed 0.23 kg $m^{-3}h^{-1}$ of carbon dioxide.

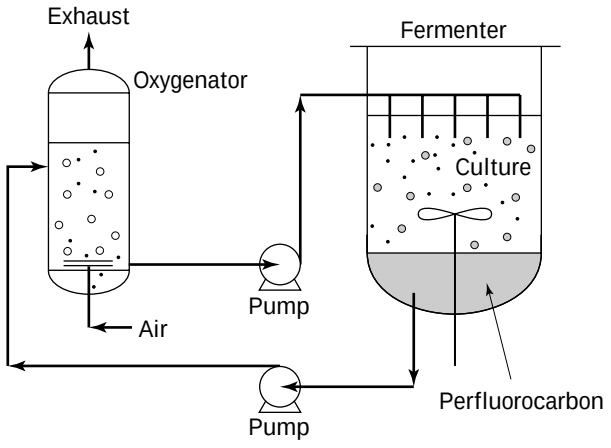


Figure 10. Bubble-free aeration with perfluorocarbon oxygen vector in a cell culture bioreactor.

With animal cells, clumping has been reported at the interface between the culture medium and the perfluorocarbon layer at the bottom of the fermenter when the agitator is placed in the bottom layer. Cells generally do not adhere to perfluorocarbon droplets. Other oxygen carriers have been tested, including hemoglobin solutions and erythrocytes, hydrocarbons such as *n*-hexadecane and *n*-dodecane, silicone oil emulsions, and soybean oil. Relevant properties of two types of oxygen vectors are noted in Table 17.

Aqueous Two-Phase Extractions. Aqueous two-phase extraction is used in bioprocessing to recover and purify bioactive proteins that are susceptible to denaturation during traditional solvent extraction that uses hydrophobic solvents. Save et al. (86) measured mass transfer coefficients of cytochrome-c, amyloglucosidase, and *β*-galactosidase in aqueous two-phase systems in a stirred cell with a flat interface between phases. In the poly(ethylene glycol)-dextran system, the controlling resistance to mass transfer occurred in the heavy phase (the dextran-rich phase). This was correlated as

$$k_L = 3.9 \times 10^{17} D_L^{1.8} E^{0.42} \mu_L^{0.78} \quad (76)$$

For poly(ethylene glycol)-sodium sulfate systems, the controlling resistance lay in the light phase (PEG-rich phase), and the mass transfer coefficient (86) was correlated as

$$k_L = 0.17 D_L^{0.33} E^{0.38} \quad (77)$$

In equations 76 and 77, μ_L is the viscosity of the phase, and D_L is the estimated diffusivity of the protein in that phase. The agitation power input per unit phase mass is represented by E. The E values ranged over ~(2 to 32) $\times 10^{-4}$ W kg^{-1}.

Solid–Liquid and Gas–Solid Mass Transfer

Dissolution or vaporization of a solid phase into a fluid or crystallization from a liquid are problems of solid–fluid mass transfer. Other similar situations occur in adsorption/desorption of material and during reaction catalyzed by a solid-phase biocatalyst. Common examples of gas–solid mass transfer are seen in various kinds of drying

Table 17. Properties of Some Oxygen Vectors

Property	*n*-Dodecane	Perfluorocarbon
Oxygen solubility (mg L^{-1})	59.4	118.0
Density (kg m^{-3})	743	1,750
Interfacial tension (liquid–air) (N m^{-1})	24.6×10^{-3}	17.8
Interfacial tension (liquid–water) (N m^{-1})	32.9×10^{-3}	25.0
Boiling point (°C)	214	206

Note: Oxygen solubility data at 35 °C; all other data at 30 °C.

operations. Mass transfer may occur to or from a fluid flowing past a stationary solid, or the solid may be suspended in the fluid. These cases are treated separately in the following sections.

Flow Past a Solid Surface. Material transport to or from a fluid moving past a stationary solid is encountered in packed beds of biocatalysts, chromatographic media, and solid adsorbents such as activated carbon. Other occurrences are in membrane processes, solid-state fermentations (87), surfaces of biofilms, capillaries containing surface-immobilized enzymes, and biofilters for treating gaseous effluents. The rate of mass transfer in flow past a solid is controlled by the flow rate of the fluid, the geometry of the solid surface, the diffusivity of the solute, and the physical properties of the flowing fluid. Suitable correlations for estimating the mass transfer coefficient for flow past solids of various geometries (flat plate, sphere, pipe, etc.) are listed in Table 18. The correlations given are for smooth surfaces. Predictive equations such as those in Table 18 and others in this monograph are usually accurate to only about $\pm 50\%$ of the actual value. Some more important bioprocess-relevant cases of mass transfer in flow past stationary solids are discussed in greater depth in the following sections.

Packed Beds. External mass transfer to or from spherical solid particles in packed beds has been correlated with the following equations

$$Sh = \frac{0.4584}{\varphi} Re^{0.5931} Sc^{1/3} \tag{78}$$

for gases (89) when $10 \leq Re \leq 10{,}000$;

$$Sh = \frac{1.09}{\varphi} Re^{1/3} Sc^{1/3} \tag{79}$$

for liquids (1) when $0.0016 \leq Re \leq 55$ and $165 \leq Sc \leq 70{,}600$; and

$$Sh = \frac{0.250}{\varphi} Re^{0.69} Sc^{1/3} \tag{80}$$

for liquids when $55 \leq Re \leq 15{,}000$ and $165 \leq Sc \leq 10{,}690$ (1). The Reynolds number in these equations is based on the particle diameter, the superficial fluid velocity is based on the total cross-section of the bed, and φ is the voidage of the packed bed. For nonspherical particles, these equations should be corrected; the Reynolds number should be calculated using the diameter of a sphere having the same surface area as the particle.

Aeration through Polymer Tubing. Bubble-free gas exchange through polymer tubing has been used in animal cell culture and culture of microalgae. In the latter case, the gas inside the tube is carbon dioxide, which is used by photosynthetic cultures to generate carbohydrates and cell mass. For animal cell culture, the tube may be supplied with air, pure oxygen, or an atmosphere containing 5% (v/v) carbon dioxide for pH control.

Either microporous or nonporous tubing may be used for oxygenation of cultures having low oxygen demands. Microporous tubing is made of polytetrafluoroethylene or polypropylene, both of which are strongly hydrophobic. Micropores, 0.03 to 3.5 μm in diameter, occupy between 30 and 75% of the surface of the tube. The tubing is fairly ridged, with typical outer diameters of 2.5 to 4 mm and a wall thickness of about 0.5 mm. The pressure inside the tubing cannot exceed the bubble point pressure or the gas will issue through the pores as bubbles. The bubble point pressure tends to be low, of the order of 10 mbar. So long as the pores are gas filled, they do not pose a significant resistance to mass transfer (20), which occurs only through the liquid film held outside the pores. Microporous tubing generally provides a better mass transfer performance than the nonporous silicone tubing if the pores remain unwetted. Pore wetting can be a problem especially during prolonged use.

In homogeneous nonporous silicone tubing, also known as solution-diffusion tubing, the oxygen from inside the tube transfers to the outside by diffusion through the silicone tube wall; the transfer rate is quite slow compared to unwetted microporous tubing. The typical dimensions of silicone tubing are inside diameter = 1.8 mm and outside diameter = 3.2 mm, or inside diameter = 3 mm and outside diameter = 4.6 mm. Silicone tubing may be internally pressurized. Because of stretching and internal gas pressure, the surface area of the installed tubing differs from that in the relaxed state. The entire surface of the tubing is effective in mass transfer, except when the tubing has been reinforced. Solubilities and diffusivities of different gases are different in silicone; hence, transport is selective (20).

The mass transfer coefficient k_L outside the microporous tubing in agitated tanks has been correlated (25) as

$$Sh = (7.8 + 0.0021 Re^{1.2}) Sc^{1/6} \tag{81}$$

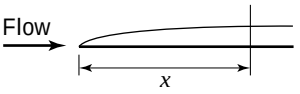
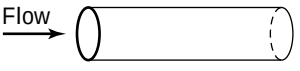
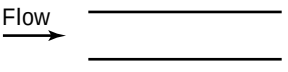
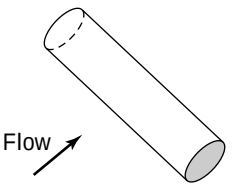
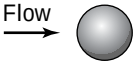
for $250 < Re < 6{,}000$ and $200 < Sc < 500$. The Reynolds number in equation 81 is based on the impeller tip speed and the outer diameter d_o of the tube; thus

$$Re = \frac{\pi N \rho_L d_i d_o}{\mu_L} \tag{82}$$

The Sherwood number is based on d_o and oxygen diffusivity in the liquid. The d_i in equation 82 is the impeller diameter. Equation 81 was developed in a cell culture vessel having the polymer tubing coiled into a cylindrical draft-tube configuration (Fig. 11). A two-bladed anchor impeller located inside the draft tube was used for agitation. For any given Reynolds number, the anchor impeller yielded higher values of Sherwood number relative to propellers and screw impellers. Equation 81 is for protein-free medium. For microporous tubes, the k_L values in protein-containing media are about 20% lower than in water. Presence of protein does not seem to affect the mass transfer coefficient of nonporous tubing (25). Equation 81 may be used also for homogeneous silicone tubing.

Depending on the porosity of tubing and its diameter, 1.8 to 3.0 m of polymer tubing is needed per liter of culture volume to meet the $k_L a_L$ demands (25). This length of tubing is close to technical limits of accommodation in a given

Table 18. Relationships for Estimating Solid–Liquid Mass Transfer Coefficients

Flow geometry	Laminar flow	Flow transition	Turbulent flow
1. Flow parallel to a flat plate 	$Sh_x = 0.332Re_x^{1/2}Sc^{1/3}(Sc \geq 0.5)$ $Sh = 0.66Re^{1/2}Sc^{1/3}(Sc \geq 0.5)$ $Sh_x = 0.565Re_x^{1/2}Sc^{1/2}(Sc \leq 0.025)$ Subscript x indicates *local* value a distance x (characteristic length) from the leading edge of plate; mean coefficient is given by the second equation where the whole length of the plate is used to calculate Re	$Re_x \approx 5 \times 10^5$	$Sh_x = 0.0296Re^{0.8}Sc^{0.6}$ $(10.0 \geq Sc \geq 0.5)$
2. Fully developed flow in straight pipe 	$Sh = 3.65 + \dfrac{0.0668(d/L)Re \times Sc}{1 + 0.04[(d/L)Re \times Sc]^{2/3}}$ Use hydraulic diameter d_h for noncircular section: $d_h = \dfrac{4 \times \text{flow area}}{\text{Wetted perimeter}}$	$Re \approx 2{,}300$	$Sh = 0.023Re^{0.8}Sc^{0.33}$ $(Sc \geq 0.5)$ $Sh = 0.0102Re^{9/10}Sc^{1/3}$ $(Sc \geq 10^3;\ Re \geq 3{,}000)$
3. Fully developed flow between parallel plates 	$Sh = 7.54 + 0.0234\dfrac{Re \times Sc}{(L/d)}$	$Re \approx 2800$	$Sh = 0.023Re^{0.8}Sc^{0.33}$ $(Sc > 0.5)$
4. Flow across a cylinder 	$Sh = 0.43 + \alpha Re^{\beta}Sc^{0.31}$ $1{,}000 \leq Re \leq 4{,}000$ $(\alpha = 0.53, \beta = 0.5)$ $4{,}000 \leq Re < 40{,}000$ $(\alpha = 0.193, \beta = 0.618)$ $Sh = 0.600Re^{0.513}Sc^{1/3}$ $0.6 \leq Sc \leq 2.6$ (gases) $1{,}000 \leq Sc \leq 3{,}000$ (liquids) $50 \leq Re \leq 50{,}000$	$Re \approx 40{,}000$	$Sh = 0.43 + 0.0265Re^{0.8}Sc^{0.31}$
5. Flow across a sphere 	$Sh = 2 + 0.37Re^{0.6}Sc^{0.33}$ $Sh = 2 + 0.552Re^{0.53}Sc^{1/3}$ (gases) $0.6 \leq Sc \leq 2.7$ $1 \leq Re \leq 48{,}000$ $Sh = 2 + 0.95Re^{0.50}Sc^{1/3}$ (liquids) $2 \leq Re \leq 2{,}000$ $Sh = 2 + 0.347Re^{0.62}Sc^{1/3}$ (liquids) $2000 \leq Re \leq 17{,}000$	$Re \approx 150{,}000$	
6. Flow through a packed bed of spheres	$St \times Sc^{2/3} = 1.625Re_d^{-1/2}$ $15 < Re_d < 120$ $Re_d = \dfrac{\rho_L U_L d_{hi}}{\mu_L}$, where $d_{hi} = \dfrac{6(1 - \varphi) \times \text{flow area}}{\text{Wetted perimeter}}$ where φ is the void fraction and St is the Stanton number: $St = \dfrac{Sh}{Sc \times Re} = \dfrac{K_L}{U_L}$	$Re_d = 120$	$St \times Sc^{2/3} = 0.687Re_d^{-0.327}$ $120 < Re_d < 2{,}000$

Source: Based on Refs. 1 and 88.

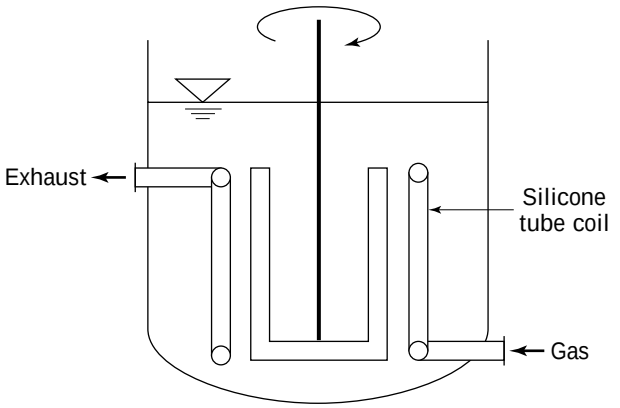


Figure 11. Oxygen supply via silicone tubing.

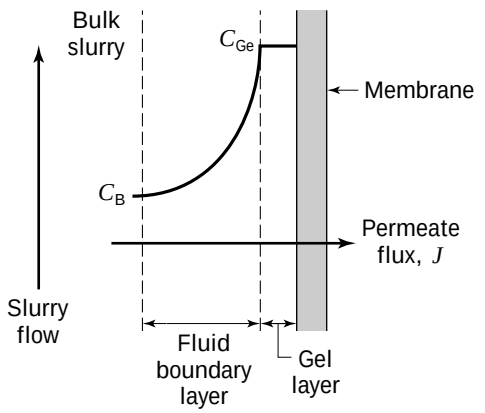


Figure 12. Steady-state solute concentration profile in ultrafiltration and microfiltration processes.

volume; therefore, aeration through polymer tubes is useful only for specific cases. The tubing is susceptible to fouling and requires periodic replacement. Anchorage-dependent cells sometimes attach to the surface of silicone tubing; hence, oxygen transfer to bulk fluid is prevented.

For transfer of carbon dioxide through nonporous silicone membranes and tubes (wall thickness 1.5 to 2.0 mm) the diffusion coefficient (D) has been reported to be 1.92 ± 0.14 m^2 min^{-1} at 25 °C (90). For mass transfer through such a tube to a highly agitated culture volume V_L, the transfer rate can be written as

$$\frac{dC_L}{dt} = \frac{AD(C^* - C_L)}{V_L \delta_w} \tag{83}$$

where A is the surface area of the tube, δ_w is its wall thickness, C_L is the instantaneous concentration of the diffusing gas in the liquid phase, and C^* is the saturation concentration of the diffusing component in a liquid sample that is in equilibrium with a gas phase having the same composition as in the silicone tube; C^* may be calculated using Henry's law. Note that equation 83 assumes a constant composition of the gas within the tube (i.e., no diffusion of other dissolved gases from the liquid into the tube). This assumption may be valid for relatively short tubes and high flow rates within the tubes. Aeration through polymer tubing is discussed further by Aunins and Henzler (20).

Mass Transfer Effects in Membrane Processes. Membrane filtrations, particularly microfiltration and ultrafiltration, are commonly used in bioprocessing to separate cells, particles, microemulsions, and macromolecules. Generally, a cross-flow scheme is used in which the fluid being filtered flows parallel to the membrane but perpendicular to the direction of the permeate flux. The turbulence generated by the flow improves mass transfer at the membrane surface; consequently, the buildup of a solute layer or gel layer on the surface of the membrane is reduced to ensure relatively high permeate or filtrate flux through the membrane. At steady state, a solute concentration profile develops on the upstream side of the membrane, as shown in Figure 12. The permeate flux J is related to the concentration C_{Ge} of the solute in the gel layer, the concentration C_B in the bulk fluid, and the mass transfer coefficient k_L as follows:

$$J = k_L \ln \frac{C_{Ge}}{C_B} \tag{84}$$

The mass transfer coefficient depends on the Reynolds number on the slurry side of the membrane as follows

$$Sh = \alpha Re^{\beta} Sc^{1/3} \tag{85}$$

where the Sh and Re are based on the hydraulic diameter d_h of the flow channel. During ultrafiltration, β is 0.5 in laminar flow and about 1.0 in turbulent flow. In cross-flow microfiltration of particles and cells, β equals 0.8 in laminar flow, but increases to about 1.3 in turbulent flow (91). The α value is 0.023 in turbulent flow. Other expressions for the Sherwood number are

$$Sh = 1.62\left(Re \times Sc \frac{d_h}{L}\right)^{0.33} \tag{86}$$

$$Sh = 1.86\left(Re \times Sc \frac{d_h}{L}\right)^{0.33} \tag{87}$$

and

$$Sh = 0.023 Re^{0.875} Sc^{0.25} \tag{88}$$

Equations 86 and 87 are for laminar flow in tubes and channels, respectively; equation 88 is for turbulent flow. For fully developed laminar flow in ultrafiltration, k_L may be related to shear rate γ by the Porter equation:

$$k_L = 0.816 \gamma^{0.33} D_L^{0.67} L^{-0.33} \tag{89}$$

where L is the length of the flow channel and the shear rate depends on the channel geometry:

$$\gamma = 8U_L/d$$

for tubes with diameter d, and

$$\gamma = 6U_L/h$$

for rectangular channels of height h.

The cross-flow velocity is the principal operating variable for enhancing the performance of a given filtration membrane module. The optimal cross-flow velocity depends on the product and the configuration of the filtration module. For tubular microfiltration membranes with ~5.5-mm inner diameter, the optimal cross-flow velocity is about 2.5 to 5 m s^{-1} (92). The inner diameter of membrane tubes is usually 4 to 20 mm. Hollow fibers have much smaller diameters at 0.5 to 2 mm. Membranes are often deployed as flat sheets in a plate-and-frame configuration in which the height of the flow channel is 0.5 to 1.5 mm (92).

Other methods for improving mass transfer in membrane processes include operation at higher temperature (i.e., higher diffusivity and lower viscosity), addition of large inert particles to the feed to agitate the gel layer, and use of pulsating flow. Module design may also be modified to enhance turbulence in the flow channel; hence, turbulence promoters such as static mixers (e.g., wire screens) may be used in tubes and channels, or membranes may be formed into a corrugated configuration. Another method of mass transfer enhancement is the use of dynamic filtration systems with rotating membranes or agitators placed in close proximity to the membrane (7).

In both ultrafiltration and microfiltration, the mass transfer coefficient tends to be quite small because of the small diffusivities of the cells and macromolecules. Formation of the gel layer, also known as concentration polarization, reduces the permeate flux in microfiltration to only about 5% of the pure water flux. Unlike microfiltration and ultrafiltration, pervaporation processes use nonporous homogeneous membranes. Typically, the solute flux is low and the mass transfer coefficient k_L is relatively large in view of the higher diffusivities of small solutes such as ethanol. For additional information on mass transfer in membrane processes, see works by Mulder (93), Ho and Sirkar (94), and Brindle and Stephenson (95).

Mass Transfer at Rough Surfaces. Compared to smooth surfaces, mass transfer from a geometrically similar rough surface is generally higher for otherwise identical conditions. Roughness induces an earlier transition to turbulence in flow past a surface. The effect of surface roughness on mass transfer coefficients in Newtonian fluids has been correlated by Kawase et al. (96) using the equation

$$Sh = 0.0093(e/d)^{0.15}Re \times Sc^{0.5}(1.11 + 0.44Sc^{-1/3} - 0.70Sc^{-1/6}) \quad (90)$$

where e is the absolute roughness (mean height of projections from the surface) and d is the pipe or channel diameter. The ratio e/d is known as relative roughness. Equation 90 is for the ranges $5 \times 10^3 < Re < 5 \times 10^5$, $5 < Sc < 400$, and $0.003 \leq e/d \leq 0.056$.

Fluid–Solid Slurries. For mass transfer to or from suspended solids, energy dissipation in the vicinity of the particles is generally assumed to control mass transfer in an isotropically turbulent field. Under such conditions, the small energy-dissipating eddies are independent of the nature of the bulk flow; the properties of those eddies depend only on energy dissipation rate per unit mass of fluid, E. In general, the liquid-film mass transfer coefficient k_L at the solid–liquid interface is expressed (97) in terms of E as follows:

$$Sh = \frac{k_L d_p}{D_L} = 2 + c\left(\frac{E^{1/3}\rho_L d_p^{4/3}}{\mu_L}\right)^a Sc^b \quad (91)$$

where $(E^{1/3}\rho_L d_p^{4/3}/\mu_L)$ may be regarded as the energy dissipation Reynolds number (98). Notice that in a quiescent environment, purely diffusive mass transfer occurs, and equation 91 reverts to $Sh = 2$ as expected from theoretical considerations. Irrespective of the reactor configuration, solid–liquid mass transfer in slurries is not significantly influenced by presence of gas bubbles. Solid–liquid mass transfer cases for the principal types of bioreactors are discussed in the following sections.

Stirred Tanks. Use the Calderbank and Moo-Young (9) equation 34 noted earlier or the Liepe and Möckel equations recommended by Hempel (99):

$$Sh = 2 + 0.67\left(\frac{P}{V_L}\frac{\rho_L^2 d_p^4}{\mu_L^3}\right)^{1/8} Sc^{1/3} \quad (92)$$

which is valid when

$$d_p < 5.2\left(\frac{\Delta\rho}{\rho_L}\right)^{-5/6}\left(\frac{\mu_L^3 V_L}{P\rho_L^2}\right)^{1/4} \quad (93)$$

When the particle diameter exceeds that given by inequality 93, the recommended equation is

$$Sh = 2 + 0.35\left(\frac{\Delta\rho}{\rho_L}\right)^{1/3}\left(\frac{P}{V_L}\frac{\rho_L^2 d_p^4}{\mu_L^3}\right)^{2/3} Sc^{1/3} \quad (94)$$

Bubble Columns. The recommended correlation (100) is

$$\frac{k_L d_p}{D_L} = 2 + 0.545Sc^{1/3}\left(\frac{g d_p^4 \rho_L^3 U_G}{\mu_L^3}\right)^{0.264} \quad (95)$$

It is suitable for $137 \leq Sc \leq 50{,}000$.

Airlift Bioreactors. For typical operating conditions, the solid–liquid mass transfer coefficient in airlift bioreactors is insensitive to the gas flow rate until such a point when increasing the superficial gas velocity ceases to produce much increase in circulation of the slurry. If the aeration rate is increased further, the k_L increases rapidly (14,97). Pneumatic energy imparted to the slurry initially produces bulk circulation of fluid; only when circulation no longer improves is the energy dissipated in microeddies that penetrate to the vicinity of the solid–liquid interface and enhance k_L. Slight variations in solids concentration (e.g., 0.2 to 4% w/w) do not affect k_L in airlift reactors (97,101).

Under similar conditions, solid–liquid mass transfer in external-loop airlift reactors exceeds that in stirred tanks and bubble columns; airlifts perform marginally better than fluidized beds (14). Performance of draft-tube internal-loop airlift devices also exceeds that of bubble col-

umns. The solid–liquid mass transfer coefficient in both external- and internal-loop types of airlift reactors may be increased further by using static mixers. In draft-tube sparged reactors, Kenics-type twisted-ribbon static mixers placed in the draft tube have been shown to enhance the mass transfer coefficient by about 20% relative to operation without the mixers (14).

The various correlations developed for solid–liquid mass transfer coefficient in internal- and external-loop types of airlift reactors are noted in Table 19. These correlations require a knowledge of the energy dissipation rate per unit mass of fluid; E is calculated using the equation

$$E = \frac{gU_{Gr}}{1 + \frac{A_d}{A_r}} \tag{96}$$

where U_{Gr} is the superficial gas velocity in the riser (5,66).

Fluidized Beds. The mass transfer coefficient for spherical particles suspended in a gas or liquid fluidized bed can be estimated using

$$Sh = \frac{0.4548 Re^{0.5931} Sc^{1/3}}{\varphi} \tag{97}$$

where φ is the void fraction of the bed. Equation 97 is suitable for $10 \leq Re \leq 4{,}000$. Another correlation applicable only to liquid fluidized beds is

$$Sh = \frac{1.1068 Re^{0.28} Sc^{1/3}}{\varphi} \tag{98}$$

which is suitable for $1 \leq Re \leq 10$.

External Mass Transfer and Heterogeneous Bioreaction. At steady state, the rate of mass transfer of a substrate being consumed at the surface of a nonporous particle equals the rate of consumption; thus

$$k_L a_s (C_B - C_{Sb}) = R \tag{99}$$

where a_s is the solid–liquid interfacial area per unit liquid volume, R is the overall rate of reaction, and C_B and C_{Sb} are substrate concentrations in the bulk fluid and at the solid's surface, respectively. Because R is a function of the substrate concentration C_{Sb}, R tends to a maximum when the concentration at the interface approaches that in the bulk fluid, i.e., the mass transfer rate or the $k_L a_s$ is large. In this situation, the rate of reaction is maximum, and it is controlled by the intrinsic kinetics and not by mass transfer effects. At the other extreme, when the mass transfer rate is low compared to the reaction rate, the substrate concentration at the interface approaches zero, and the reaction is mass transfer controlled. Because now $C_{Sb} = 0$, the reaction rate is

$$R = k_L a_s C_B \tag{100}$$

Thus, the observed rate of a mass transfer controlled reaction will be influenced by changes in C_B, k_L, or the specific solid–liquid interfacial area a_s. Indeed, observing influence of these parameters on the reaction rate provides methods for determining if the reaction is mass transfer limited. In addition, because the activation energy for mass transfer is much smaller than that of biochemical reactions, the observed rate of a mass transfer controlled reaction is not as sensitive to temperature changes as the rate of reaction when it is not limited by transport effects. In a batch-stirred vessel, an increase in the rate of a solid–phase catalyzed reaction with increasing agitation implies a mass transfer limited reaction. Similarly, in a packed-bed reactor, the rate of a mass transfer controlled reaction will increase as the flow rate is increased, but the contact time is kept unchanged (i.e., the depth of the bed is increased proportionately).

The solid–liquid mass transfer correlations given earlier in this article may be used to estimate k_L in bubble columns, packed beds, fluidized beds, and other geometries of bioreactors. External mass transfer in immobilized enzyme systems has been discussed in greater detail by Goldstein (104).

Table 19. Solid–Liquid Mass Transfer Coefficient in Airlift Reactors

Reactor configuration	Correlation	Ranges
Internal-loop reactors (draft-tube sparged)	$Sh = 2 + 0.064\left(\frac{E\rho_L^3 d_p^4}{\mu_L^3}\right)^{0.165} Sc^{0.45}$ Kushalkar and Pangarkar (102)	0.5% wt benzoic acid granules, $d_p = 0.55$–3 mm, in air and water; $A_r/A_d = 0.17 - 1.29$, $U_G = 0.08 - 0.35$ m s^{-1}
Internal-loop reactors (draft-tube sparged)	$Sh = 2 + 1.01\left(\frac{E\rho_L^3 d_p^4}{\mu_L^3}\right)^{0.173} Sc^{0.33}$ Goto et al. (101)	Amberlyst 15 ion exchange resin, $d_p = 0.55$–0.92 mm, suspended in dilute aqueous sodium hydroxide; $A_r/A_d = 0.1 - 1.4$
Internal-loop reactors with static mixers (draft-tube sparged)	$Sh = 2 + 1.68\left(\frac{E\rho_L^3 d_p^4}{\mu_L^3}\right)^{0.133} Sc^{0.33}$ Gaspillo and Goto (103)	Kenics-type twisted-ribbon static mixers were in draft tube
External-loop reactors	$Sh = 2 + 0.48\left(\frac{E^{1/3}\rho_L d_p^{4/3}}{\mu_L}\right)^{0.72} Sc^{1/3}$ Mao et al. (97)	Benzoic acid coated particles, $d_p = 3.8$ mm, $\rho_s \approx 1{,}080$ kg m^{-3}, suspended in water

Intraparticle Mass Transfer. When a substrate is being consumed or a product is being produced inside a porous immobilization matrix or in homogeneous gels, an internal substrate (and product) concentration profile exists within the matrix at steady state. In microporous particles, in the absence of convective transport, the main internal mass transfer parameter is the effective diffusivity D_e in the particle. The effective diffusivity is related to the liquid- or gas-phase diffusivity in the pores as follows

$$D_e = \frac{p}{\tau} D_L \tag{101}$$

In equation 101, p is the porosity, and τ is the tortuosity of the immobilization matrix. Intraparticle mass transfer is encountered in immobilized enzyme and cell particles, in chromatographic media, and during leaching of solubles from inert solids.

In macroporous particles, changes in local fluid velocity outside the particle create pressure fluctuations within the larger pores, and this leads to some agitation of the fluid held in the pores. In this situation, the apparent effective diffusivity ($D_{e\,ap}$) is greater than that predicted by equation 101. For example, the apparent effective diffusivity of oxygen in pellets of *Aspergillus niger* has been reported to be $\sim 10^{-8}$ m^2 s^{-1} or about 10-fold greater than in water, which suggests a level of fluid movement within the pellet, possibly because of turbulence-induced elastic structural deformations of the loose floc.

For porous particles in packed beds, the apparent effective diffusivity has been observed to depend on the effective diffusivity (no convection) and the Peclet number; thus

$$D_{e\,ap} = \frac{D_e}{\frac{3}{Pe}\left(\frac{1}{\tanh Pe} - \frac{1}{Pe}\right)} \tag{102}$$

where the Peclet number is given as

$$Pe = \frac{d_p U_L}{(1 - \varphi) D_e} \tag{103}$$

In equation 103, d_p is the diameter of the particle, U_L is the superficial fluid velocity through the bed, and φ is its void fraction. More complete treatments of intraparticle mass transfer have been provided by Chisti and Moo-Young (7), Radovich (105), Willaert et al. (106), Pitcher (107), and Bailey and Ollis (108).

NOMENCLATURE

A	Surface area of tube (m^2)
A_d	Cross-sectional area of downcomer (m^2)
A_r	Cross-sectional area of riser (m^2)
A_s	Total solid–liquid interfacial area (m^2)
a_D	Gas–liquid interfacial area per unit volume of dispersion (m^{-1})
a_L	Gas–liquid interfacial area per unit liquid volume (m^{-1})
a_{Lc}	Liquid–liquid interfacial area per unit volume of continuous phase (m^{-1})
a_s	Solid–liquid interfacial area per unit liquid volume (m^{-1})
BOD	Biochemical oxygen demand
Bo	Bond number (–)
C	Concentration (kg m^{-3})
C^*	Saturation concentration of transferring gas or solute in liquid (kg m^{-3})
ΔC	Concentration gradient (kg m^{-3})
C_B	Solute or substrate concentration in bulk liquid (kg m^{-3})
C_G	Concentration in the gas phase (kg m^{-3})
C_{Ge}	Solute concentration in gel layer (kg m^{-3})
C_{Gi}	Interfacial concentration of the diffusing component in the gas phase (kg m^{-3})
C_L	Instantaneous concentration of transferring component in liquid (kg m^{-3})
C_{L0}	Initial dissolved oxygen concentration at time t_0 (kg m^{-3})
C_{Lc}	Equilibrium concentration of solute in continuous phase (kg m^{-3})
C_{Ld}	Equilibrium concentration of solute in dispersed phase (kg m^{-3})
C_{Li}	Interfacial concentration of the diffusing component in the liquid phase (kg m^{-3})
C_S	Concentration of solids in slurry (% w/v or kg m^{-3})
C_{Sb}	Substrate concentration at solid–liquid interface (kg m^{-3})
C^*_{in}	Saturation concentration of oxygen in equilibrium with ingoing gas (kg m^{-3})
C^*_{out}	Saturation concentration of oxygen in equilibrium with the exhaust gas (kg m^{-3})
D	Diffusivity (m^2 s^{-1})
D_D	Diffusivity of dispersed phase (m^2 s^{-1})
D_e	Effective diffusivity in particle (m^2 s^{-1})
$D_{e\,ap}$	Apparent effective diffusivity in particle (m^2 s^{-1})
D_{gas}	Diffusivity of gas in liquid (m^2 s^{-1})
D_L	Diffusivity of gas or solute in liquid (m^2 s^{-1})
DMEM	Dulbecco's modified Eagle's medium
D_{oxygen}	Diffusivity of gas in liquid (m^2 s^{-1})
D_P	Diffusivity of small solute in protein solution (m^2 s^{-1})
d	Characteristic length dimension (m)
d^*	Diameter number defined by equation 32 (–)
d_B	Sauter mean bubble diameter (m)
d_D	Diameter of drop (m)
d_F	Maximum diameter of flask (m)
d_H	Diameter of sparger hole (m)
d_h	Hydraulic diameter (m)
d_{hi}	Hydraulic diameter in packed bed as defined in Table 18 (m)

d_i Diameter of impeller (m)
d_o Outer diameter of tubing (m)
d_p Diameter of particle (m)
d_r Diameter of riser (m)
d_T Diameter of tank or column (m)
E Energy dissipation rate per unit mass (W kg^{-1})
EDTA Ethylenediaminetetraacetic acid
E_L Overall axial dispersion coefficient of liquid (m^2 s^{-1})
e Absolute roughness (m)
$\hat{e}$ Eccentricity of shaker platform (−)
F Mass flow rate of gas (kg s^{-1})
Fr Froude number (−)
f Parameter (−)
Gr Grashof number (−)
g Gravitational acceleration ($m \cdot s^{-2}$)
H Dimensionless Henry's law constant (−)
h Height of channel (m)
h_D Height of dispersion (m)
h_L Height of gas-free liquid (m)
I Ionic strength (kg ion m^{-3})
J Mass flux (kg m^{-2} s^{-1} or kmol m^{-2} s^{-1}) or permeate flux (m s^{-1})
K Consistency index of fluid (Pa s^n)
K_L Overall mass transfer coefficient based on liquid film (m s^{-1})
K_P Mass transfer coefficient for plug (mol O_2 atm^{-1} h^{-1})
K_p Partition coefficient of solute defined by equation 25 (−)
K_S Mass transfer coefficient at surface (mol O_2 atm^{-1} h^{-1})
k Mass transfer coefficient (m s^{-1})
k_G Gas film mass transfer coefficient (m s^{-1})
k_i Impeller-dependent constant in equation 44 (−)
k_L Liquid film mass transfer coefficient (m s^{-1})
k_La_L Overall gas–liquid volumetric mass transfer coefficient (s^{-1})
$(k_La_L)_{rord}$ Overall volumetric mass transfer coefficient in riser or downcomer (s^{-1})
k_{Lc} Liquid film mass transfer coefficient in continuous phase (m s^{-1})
k_{Ld} Liquid film mass transfer coefficient in dispersed phase (m s^{-1})
L Length of circulation loop, pipe, plate, or flow channel (m)
l Length of impeller blade (m)
M_L Molecular weight of liquid (kg $kmol^{-1}$)
Mo Morton number (−)
N Rotational speed (s^{-1})
n Flow behavior index (−)
P Power input (W)
PEG Poly(ethylene glycol)
Pe Peclet number (−)
P_G Gassed power input (W)
Po Power number (−)
Ps Poiseuille number (−)
p Porosity of particle (−)
p_{atm} Partial pressure of oxygen in the atmosphere outside flask (atm)
p_L Partial pressure of oxygen in liquid in flask (atm)
Q_G Volume flow rate of gas (m^3 s^{-1})
q_{O_2} Specific oxygen consumption rate (s^{-1})
R Overall rate of reaction (kg m^{-3} s^{-1})
Ra Rayleigh number (−)
R_b Ratio of bubble diameter to its width (−)
Re Reynolds number (−)
Re_C Reynolds number for the continuous phase (−)
Re_d Reynolds number based on d_{hi} as in Table 18 (−)
Re_i Reynolds number based on impeller (−)
Re_t Reynolds number based on terminal settling velocity of particle (−)
Re_x Local value of Reynolds number at distance x from leading edge (−)
RPMI Cell culture medium
Sc Schmidt number (−)
SDS Sodium dodecyl sulfate
Sh Sherwood number (−)
Sh_x Local value of Sherwood number at distance x from leading edge (−)
SLS Sodium lauryl sulfate
St Stanton number (−)
T Temperature (°C) or absolute temperature (K)
t Instantaneous time (s)
t_C Contact time of droplets (s)
t_m Mixing time (s)
t_R Residence time of gas in liquid (s)
U_B Bubble rise velocity (m s^{-1})
U_G Superficial gas velocity based on outer column cross-sectional area (m s^{-1})
U_{Gr} Superficial gas velocity in the riser (m s^{-1})
U_L Mean superficial liquid velocity in the reactor or pipe (m s^{-1})
U_{Lr} Superficial liquid velocity in the riser (m s^{-1})
U_p Velocity of particle (m s^{-1})
U_t Terminal settling velocity of a single particle in liquid (m s^{-1})
V_L Volume of liquid or slurry (m^3)
V_{Lc} Volume of continuous phase (m^3)
V_M Molar volume of solute at its boiling point (m^3 $kmol^{-1}$)
W Width of impeller blade (m)
We Weber number (−)

Wi	Weissenberg number (−)
X	Viable biomass concentration (kg m^{-3})
x	Distance (m)
x_i	Mass fraction of oxygen in inlet gas stream (−)
x_o	Mass fraction of oxygen in exhaust gas (−)

Greek Symbols

α	Parameter; parameter in equation 66 (m$^{-\beta}$ s$^{-(\beta+1)}$)
α_M	Value of α (equation 66) in presence of static mixers (m$^{-\beta}$ s$^{-(\beta+1)}$)
β	Parameter (−)
γ	Parameter or shear rate (s^{-1})
δ	Film thickness (m)
δ_G	Thickness of gas film (m)
δ_L	Thickness of liquid film (m)
δ_w	Wall thickness of tube (m)
ϵ_G	Overall gas holdup (−)
ϵ_{Gr}	Gas holdup in the riser (−)
ϵ_S	Volume fraction or holdup of solids in three-phase systems (−)
θ	Parameter in equation 73 (−)
λ	Characteristic time or relaxation time of viscoelastic fluid (s)
μ_{ap}	Apparent viscosity of non-Newtonian fluid (Pa s)
μ_C	Viscosity of the continuous phase (Pa s)
μ_D	Viscosity of the dispersed phase (Pa s)
μ_G	Viscosity of the gas (Pa s)
μ_L	Viscosity of the liquid (Pa s)
μ_{SL}	Viscosity of the slurry (Pa s)
μ_w	Viscosity of water (Pa s)
χ	Association parameter (−)
π	Pi (−)
ρ_C	Density of continuous phase (kg m^{-3})
ρ_D	Density of dispersed phase (kg m^{-3})
ρ_G	Density of gas (kg m^{-3})
ρ_L	Density of the liquid (kg m^{-3})
ρ_p	Density of the particles (kg m^{-3})
ρ_S	Density of solids (kg m^{-3})
ρ_{SL}	Density of slurry (kg m^{-3})
$\Delta\rho$	Density difference between phases (kg m^{-3})
σ	Interfacial tension (liquid–liquid) (kg s^{-2})
σ_L	Interfacial tension (kg s^{-2})
τ	Tortuosity of pores (−)
φ	Void fraction of the packed bed (−)
φ_S	Volume fraction of solids in gas-free slurry (−)
χ	Parameter (−)

BIBLIOGRAPHY

1. C.J. Geankoplis, *Transport Processes and Unit Operations*, 3rd ed., Prentice Hall, Englewood Cliffs, 1993.
2. H.R.C. Pratt, in T.C. Lo, M.H.I. Baird, and C. Hanson eds., *Handbook of Solvent Extraction*, Wiley, New York, 1983, pp. 91–123.
3. C.R. Wilke and P. Chang, *AIChE J.* **1**, 264–270 (1955).
4. Y. Chisti, in R. Robinson, C. Batt, and P. Patel eds., *Encyclopedia of Food Microbiology*, Academic Press, London, in press.
5. Y. Chisti, *Airlift Bioreactors*, Elsevier, London, 1989, p. 350.
6. K. Van't Riet, *Ind. Eng. Chem. Process Des. Dev.* **18**, 357–363 (1979).
7. Y. Chisti and M. Moo-Young, in V. Moses and R.E. Cape eds., *Biotechnology: The Science and the Business*, Harwood Academic Publishers, New York, 1991, pp. 167–209.
8. C.S. Ho and J.F. Shanahan, *CRC Crit. Rev. Biotechnol.* **4**, 185–252 (1986).
9. P.H. Calderbank and M.B. Moo-Young, *Chem. Eng. Sci.* **16**, 39–54 (1961).
10. P.H. Calderbank and A.C. Lochiel, *Chem. Eng. Sci.* **19**, 485–503 (1964).
11. Y. Chisti and M. Moo-Young, *Biotechnol. Bioeng.* **34**, 1391–1392 (1989).
12. E. Molina Grima, Y. Chisti, and M. Moo-Young, *J. Biotechnol.* **54**, 195–210 (1997).
13. L.K. Shi, J.P. Riba, and H. Angelino, *Chem. Eng. Commun.* **89**, 25–35 (1990).
14. Y. Chisti, *Appl. Mech. Rev.* **51**, 33–112 (1998).
15. M. Reuss, in H.-J. Rehm and G. Reed eds., *Biotechnology*, vol. 3, 2nd ed., VCH, Weinheim, 1993, pp. 185–217.
16. M. Moo-Young and H.W. Blanch, *Adv. Biochem. Eng.* **19**, 1–69 (1981).
17. T. Oolman and H.W. Blanch, *CRC Crit. Rev. Biotechnol.* **4**, 133–184 (1986).
18. Y. Chisti and M. Moo-Young, *J. Chem. Technol. Biotechnol.* **58**, 331–336 (1993).
19. S. Yamada, M. Wada, and I. Chibata, *J. Ferment. Technol.* **56**, 20–28 (1978).
20. J.G. Aunins and H.-J. Henzler, in H.-J. Rehm and G. Reed eds., *Biotechnology*, vol. 3, 2nd ed. VCH, Weinheim, 1993, pp. 219–281.
21. L. Randers-Eichhorn, R.A. Bartlett, D.D. Frey, and G. Rao, *Biotechnol. Bioeng.* **51**, 466–478 (1996).
22. J.G. Aunins, B.A. Woodson Jr., T.K. Hale, and D.I.C. Wang, *Biotechnol. Bioeng.* **34**, 1127–1132 (1989).
23. M. Lavery and A.W. Nienow, *Biotechnol. Bioeng.* **30**, 368–373 (1987).
24. Y. Kato, S. Hiraoka, Y. Tada, K. Sato, and T. Ohishi, *J. Chem. Eng. Jpn*, **30**, 362–365 (1997).
25. H.-J. Henzler and D.J. Kauling, *Bioprocess Eng.* **9**, 61–75 (1993).
26. J.F. Perez and O.C. Sandall, *AIChE J.* **20**, 770–775 (1974).
27. Y. Kawase and M. Moo-Young, *Trans. Inst. Chem. Eng.* **68**, 189–194 (1990).
28. R.E. Farritor and G.A. Hughmark, *AIChE J.* **20**, 1027–1028 (1974).
29. K.A. Bin, *Chem. Eng. Commun.* **31**, 155–183 (1984).
30. C.H.E. Nielsen, S. Kiil, H.W. Thomsen, and K. Dam-Johansen, *Chem. Eng. Sci.* **53**, 495–503 (1998).
31. E.T. Papoutsakis, *Trends Biotechnol.* **9**, 427–437 (1992).
32. A. Handa, A.N. Emery, and R.E. Spier, *Devel. Biol. Stand.* **66**, 241–253 (1987).

33. A. Handa-Corrigan, A.N. Emery, and R.E. Spier, *Enzyme Microb. Technol.* **11**, 230–235 (1989).
34. M.S. Croughan, J.-F. Hamel, and D.I.C. Wang, *Biotechnol. Bioeng.* **29**, 130–141 (1987).
35. Y. Chisti, *Bioprocess. Eng.* **9**, 191–196 (1993).
36. W.R. Arathoon and J.R. Birch, *Science* **232**, 1390–1395 (1986).
37. J.R. Birch, K. Lambert, P.W. Thompson, A.C. Kenney, and L.A. Wood, in B.K. Lydersen ed., *Large Scale Cell Culture Technology*, Hanser Publishers, New York, 1987, pp. 1–20.
38. Z. Zhang, Y. Chisti, and M. Moo-Young, *J. Biotechnol.* **43**, 33–40 (1995).
39. A.W. Nienow, *Appl. Mech. Rev.* **51**, 3–32 (1998).
40. A. Bakker, J.M. Smith, and K.J. Myers, *Chem. Eng.* **101**, 98–104 (1994).
41. C.S. Shin, M.S. Hong, and J. Lee, *Biotechnol. Technol.* **10**, 679–682 (1996).
42. H. Taguchi and F. Yoshida, *J. Ferment. Technol.* **46**, 814 (1968).
43. H. Yagi and F. Yoshida, *Ind. Eng. Chem. Process Des. Develop.* **14**, 488–493 (1975).
44. Y. Kawase and M. Moo-Young, *Chem. Eng. Res. Des.* **66**, 284–288 (1988).
45. K. Chandrasekharan and P.H. Calderbank, *Chem. Eng. Sci.* **36**, 819–823 (1981).
46. Y.T. Shah, B.G. Kelkar, S.P. Godbole, and W.-D. Deckwer, *AIChE J.* **28**, 353–379 (1982).
47. W.-D. Deckwer, *Bubble Column Reactors*, Wiley, New York, 1992.
48. W.-D. Deckwer and A. Schumpe, *Chem. Eng. Sci.* **48**, 889–911 (1993).
49. J.J. Heijnen and K. Van't Riet, *Chem. Eng. J.* **28**, B21–B42 (1984).
50. J.R. Fair, *Chem. Eng.* **74**, 67 (1967).
51. K. Akita and F. Yoshida, *Ind. Eng. Chem. Process Des. Develop.* **12**, 76–80 (1973).
52. G.A. Hughmark, *Ind. Eng. Chem. Process Des. Develop.* **6**, 218–220 (1967).
53. M. Nakanoh and F. Yoshida, *Ind. Eng. Chem. Process Des. Develop.* **19**, 190–195 (1980).
54. A. Schumpe, C. Singh, and W.-D. Deckwer, *Chem.-Ing.-Technol.* **57**, 988–989 (1985).
55. Y. Kawase, B. Halard, and M. Moo-Young, *Chem. Eng. Sci.* **42**, 1609–1617 (1987).
56. H. Hikita, S. Asai, K. Tanigawa, K. Segawa, and M. Kitao, *Chem. Eng. J.* **22**, 61–69 (1981).
57. W.-D. Deckwer, K. Nguyen-Tien, A. Schumpe, and Y. Serpemen, *Biotechnol. Bioeng.* **24**, 461–481 (1982).
58. S.P. Godbole, A. Schumpe, Y.T. Shah, and N.L. Carr, *AIChE J.* **30**, 213–220 (1984).
59. J. Dudley, *Water Res.* **29**, 1129–1138 (1995).
60. G.-Q. Li, S.-Z. Yang, Z.-L. Cai, and J.-Y. Chen, *Chem. Eng. J.* **56**, B101–B107 (1995).
61. K. Koide, H. Sato, and S. Iwamoto, *J. Chem. Eng. Jpn.* **16**, 407–413 (1983).
62. K. Koide, K. Horibe, H. Kawabata, and S. Ito, *J. Chem. Eng. Jpn.* **18**, 248–254 (1985).
63. P. Verlaan, J.-C. Vos, and K. van't Riet, *J. Chem. Technol. Biotechnol.* **45**, 181–190 (1989).
64. I.-S. Suh, A. Schumpe, and W.-D. Deckwer, *Biotechnol. Bioeng.* **39**, 85–94 (1992).
65. K. Koide, K. Shibata, H. Ito, S.Y. Kim, and K. Ohtaguchi, *J. Chem. Eng. Jpn.* **25**, 11–16 (1992).
66. Y. Chisti and M. Moo-Young, *Chem. Eng. Commun.* **60**, 195–242 (1987).
67. A.B. Russell, C.R. Thomas, and M.D. Lilly, *Bioprocess. Eng.* **12**, 71–79 (1995).
68. D.J. Pollard, A.P. Ison, P.A. Shamlou, and M.D. Lilly, in E. Galindo and O.T. Ramírez eds., *Advances in Bioprocess Engineering*, Kluwer, Dordrecht, The Netherlands, 1994, pp. 163–170.
69. Y. Chisti and M. Moo-Young, *Chem. Eng. Progress*, **89**, 38–45 (1993).
70. K. Koide, K. Horibe, H. Kitaguchi, and N. Suzuki, *J. Chem. Eng. Jpn.* **17**, 547–549 (1984).
71. Y. Chisti, M. Kasper, and M. Moo-Young, *Can. J. Chem. Eng.* **68**, 45–50 (1990).
72. W. Zhou, K. Holzhaur-Rieger, and K. Schügerl, *J. Biotechnol.* **28**, 165–177 (1993).
73. J. Audet, J. Thibault, and A. LeDuy, *Biotechnol. Bioeng.* **52**, 507–517 (1996).
74. M. Lounes, J. Audet, J. Thibault, and A. LeDuy, *Bioprocess Eng.* **13**, 1–11 (1995).
75. X. Ni, S. Gao, and D.W. Pritchard, *Biotechnol. Bioeng.* **45**, 165–175 (1995).
76. H. Ziegler, I.J. Dunn, and J.R. Bourne, *Biotechnol. Bioeng.* **22**, 1613–1635 (1980).
77. K. Muroyama and L.-S. Fan, *AIChE J.* **31**, 1–34 (1985).
78. K. Nguyen-Tien, A.N. Patwari, A. Schumpe, and W.-D. Deckwer, *AIChE J.* **31**, 194–201 (1985).
79. V.Y. Loh, S.R. Richards, and P. Richmond, *Proceedings: Int. Conference on Bioreactor Fluid Dynamics*, 1986, 15–17, April BHRA, Cambridge, England.
80. B.C. Smith and D.R. Skidmore, *Biotechnol. Bioeng.* **35**, 483–491 (1990).
81. D.G. Karamanev, T. Nagamune, and I. Endo, *Chem. Eng. Sci.* **47**, 3581–3588 (1992).
82. M.A. Winkler, *Biological Treatment of Waste-Water*, Ellis Horwood, Chichester, U.K., 1981.
83. W. Sittig, *J. Chem. Technol. Biotechnol.* **32**, 47–58 (1982).
84. K. Schügerl, *Solvent Extraction in Biotechnology: Recovery of Primary and Secondary Metabolites*, Springer-Verlag, New York, 1994.
85. S. Martin, P. Soucaille, and J.-S. Condoret, *Bioprocess Eng.* **13**, 293–300 (1995).
86. S.S. Save, R.B. Desai, S.B. Sawant, and J.B. Joshi, *Trans. Inst. Chem. Eng.* **74C**, 171–177 (1996).
87. Y. Chisti, in M.C. Flickinger and S.W. Drew, eds., *Encyclopedia of Bioprocess Technology*, Wiley, New York, 1999, in press.
88. D.K. Edwards, V.E. Denny, and A.F. Mills, *Transfer Processes*, Holt, Rinehart and Winston, New York, 1973.
89. P.N. Dwivedi and S.N. Upadhyay, *Ind. Eng. Chem. Process Des. Dev.* **16**, 157–165 (1977).
90. Y.-K. Lee and H.-K. Hing, *Appl. Microbiol. Biotechnol.* **31**, 298–301 (1989).
91. R.S. Tutunjian, *Biotechnology* **3**, 615–626 (1985).
92. S. Ripperger, *Chem. Eng. Technol.* **11**, 17–25 (1988).
93. M.H.V. Mulder, in R.D. Noble and S.A. Stern eds., *Membrane Separations Technology: Principles and Applications*, Elsevier, Amsterdam, 1995, pp. 45–84.

94. W.S.W. Ho and K.K. Sirkar eds., *Membrane Handbook*, Van Nostrand Reinhold, New York, 1992.
95. K. Brindle and T. Stephenson, *Biotechnol. Bioeng.* **49**, 601–610 (1996).
96. Y. Kawase, A.V. Shenoy, and K. Wakabayashi, *Can. J. Chem. Eng.* **72**, 798–804 (1994).
97. H.H. Mao, Y. Chisti, and M. Moo-Young, *Chem. Eng. Commun.* **113**, 1–13 (1992).
98. W.-D. Deckwer and A. Schumpe, *Ger. Chem. Eng.* **7**, 168–177 (1984).
99. D.C. Hempel, in R.K. Finn, P. Präve, M. Schlingmann, W. Crueger, K. Esser, R. Thauer, F. Wagner eds., *Biotechnology Focus 1: Fundamentals, Applications, Information*, Hanser, New York, 1988, pp. 51–94.
100. P. Sänger and W.-D. Deckwer, *Chem. Eng. J.* **22**, 179–186 (1981).
101. S. Goto, Y. Matsumoto, and P. Gaspillo, *Chem. Eng. Commun.* **85**, 181–191 (1989).
102. K.B. Kushalkar and V.G. Pangarkar, *Chem. Eng. Sci.* **49**, 139–144 (1994).
103. P.D. Gaspillo and S. Goto, *J. Chem. Eng. Jpn.* **24**, 680–682 (1991).
104. L. Goldstein, *Methods Enzymol.* **44**, 397–443 (1976).
105. J.M. Radovich, *Enzyme Microbial. Technol.* **7**, 2–10 (1985).
106. R.G. Willaert, G.V. Baron, and L. De Backer eds., *Immobilised Living Cell Systems: Modelling and Experimental Methods*, Wiley, Chichester, 1996.
107. W.H. Pitcher Jr., *Adv. Biochem. Eng.* **10**, 1–26 (1978).
108. J.E. Bailey and D.F. Ollis, *Biochemical Engineering Fundamentals*, 2nd ed., McGraw-Hill, New York, 1986, pp. 204–220.

See also BIOFILMS, MICROBIAL; BIOREACTORS, AIR-LIFT REACTORS; BIOREACTORS, CONTINUOUS STIRRED-TANK REACTORS; BIOREACTORS, FLUIDIZED-BED; GAS HOLD-UP; HYPOXIA, EFFECTS ON ANIMAL CELLS; SCALE-UP, STIRRED-TANK REACTORS; SOLID STATE FERMENTATION, MICROBIAL GROWTH KINETICS; TRANSFER PHENOMENA IN MULTIPHASE SYSTEMS IN MIXING VESSELS.

MEDIA COMPOSITION, MICROBIAL, LABORATORY SCALE

ROSALIE J. COTE
Becton Dickinson Microbiology Systems
Sparks, Maryland

KEY WORDS

Agar
Buffers
Carbon and energy sources
Detoxifiers
Energy
Growth factors
Oxidation–reduction potential
pH
Protein hydrolysates (peptones)
Sterilization
Water

OUTLINE

INTRODUCTION

The increased understanding of microbial biochemistry in recent decades, along with major advances in genetic and molecular techniques, has resulted in an expanding diversity of microorganisms entering the biotechnology arena. Be it novel uses for well-studied organisms, such as the commercial production of restriction endonucleases from many common bacteria and yeasts, or the potential for development of new applications from new species such as the unidentified hydrogen-producing bacterium FOX-1 (1), microorganisms, themselves, are integral components in the overall aspects of biotechnology from discovery, through research and development, to production.

For those biotechnologists not formally trained in classical microbial physiology, the too-prevalent assumption is that microorganisms function similarly to the purified re-

agents used in a well-documented chemical reaction. In reality, microorganisms (even authenticated and characterized strains) are heterogeneous conglomerations of multiple biochemical pathways with the potential for working antagonistically or synergistically, depending upon the environment in which the microorganism resides. For those strains under study in the laboratory or being potentially harnessed for bioprocess use, the microbial environment is the culture medium.

This article surveys elements of microbial media and other considerations of microbial nutrition that impact upon the general growth and metabolic activity of microorganisms. Although specific references are given for many useful compendia of microbial media formulations (2–7), the purpose here is not to further cite well-documented media and buffer formulations. The intent of this article and its specific examples is to illustrate general principles of controlled media design, formulation, and preparation for the cultivation of microorganisms.

ESSENTIAL NUTRITIONAL REQUIREMENTS

Water

Depending upon the physiological state of the organism, the chemical composition of bacteria is 70–90% water. As a polar solvent, water dissolves other compounds of similar chemical nature, which also happen to be the biologically important molecules such as amino acids, carbohydrates, and nucleic acids. The polarity of water helps to permit the hydrogen bonding necessary in the formation of the final working configuration of a cell's large macromolecules. As a polar solvent, water chemically forces nonpolar molecules such as lipids into hydrophobic aggregates that are integral in the formation cell membranes. The presence of water is critical to the existence of microorganisms. Under laboratory conditions, microorganisms are cultivated with liquid and solid media. In both situations, the adequacy and purity of water used in media preparation can drastically effect the performance and/or sustained viability of an organism.

Because the availability of high-quality, purified water has become increasingly affordable for even small laboratories, the use of tap water or marginally distilled water for the preparation of media and reagents is not a recommended option. For most municipalities, tap water is seasonably variable with respect to its chemical and microbial content. Hard water ladened with calcium and magnesium will precipitate many media formulations, with deleterious results. The terms *deionized* or *distilled* water with respect to water purity are merely euphemistic unless actual determinations of residual ion and organic levels are known. Water passed through spent deionization beds or boiled and condensed through a scale-clogged distillation system is hardly purified. At minimum, water used in the preparation of microbial media and reagents for general laboratory purposes should meet U.S.P. specifications for purified water; other organizations such as the American Society for Testing and Materials (ASTM) and the National Committee for Clinical Laboratory Standards (NCCLS) have similar criteria with respect to the conductivity, resistivity, and total organic content for Type II water (8). Point of use deionization systems available from many manufacturers (e.g., Barnstead/Thermolyne Corporation, Dubuque, Ia. and Millipore Corporation, Bedford, Mass.) easily produce Type I, 18 MΩ water. Some systems include ultrafiltration to reduce pyrogen levels. Purified water for highly regulated applications and processes (e.g., biopharmaceuticals) must meet U.S.P. specifications for water for injection (9). Water, unless sterilized, should not be held for any length of time in storage containers such as bottles or carboys or allowed to stand in uncirculated piping systems (e.g., dead-legs) because there is a good probability that the water will become populated with bacterial biofilms that grow on container and piping walls. Bacterial contamination will organically pollute water purity in addition to increasing endotoxin and pyrogens to unacceptable levels for certain procedures including product purifications or animal injections. Note that pyrogens are lipopolysaccharides from bacterial membranes. With respect to purified water, they indicate bacterial contamination. Their presence, of course, in bacterial cultures is a mute point. Use of plastics for water storage containers may be a consideration because plasticizer leaching can cause inhibition of sensitive organisms.

Energy

All microorganisms require an exogenous source of energy. The biochemical mechanisms involved with this aspect of metabolism are diverse. *Lithotrophic* organisms obtain energy from inorganic sources. Lithotrophs fall into two metabolic categories with respect to energy procurement. *Photolithotrophic bacteria* and *cyanobacteria* obtain energy from light in conjunction with, respectively, reduced sulfur (e.g., hydrogen sulfide) or water, to release energy for biosynthetic functions. Variations in microbial chlorophyll molecules create adsorption spectra that, in nature, limit a given species to a particular photic zone. Under laboratory conditions, the light requirements necessary for cultivation reflect this consideration. For example, the green and purple sulfur bacteria, such as *Chlorobium* and *Chromatium*, with adsorption peaks above 750 nm, are cultivated under the red wavelengths of tungsten light. Cultures, however, should not be placed too close to an incandescent light source because the heat emitted from the bulbs can quickly raise the temperature of the culture vessel beyond the upper growth limits for an organism. Cyanobacteria (e.g., *Anabaena*) adsorb blue wavelengths of light below 700 nm. This requirement can be satisfied by cultivating the bacteria under fluorescent tubes that emit light at this range of the visible spectrum. *Chemolithotrophic* organisms utilize inorganic compounds for energy. Examples within this second form of energy procurement are *Sulfolobus* spp., which oxidize elemental sulfur to sulfuric acid, and *Nitrobacter* spp., which biochemically exploit the oxidation of nitrites to nitrates. Numerous new genera and species of prokaryotic chemolithotrophs that obtain energy from various inorganic electron donors have been isolated in recent years, including *Archaeoglobus* spp., *Methanococcus* spp., and *Pyrodictium* spp. (10). Many obligate lithotrophs are inhibited by organically rich en-

vironments and are best cultivated under chemically defined media conditions.

Organotrophic organisms obtain energy more conventionally from organic compounds. As with the lithotrophs, organotrophs many be either phototrophic or chemotrophic. *Photoorganotrophs* are represented in nature by the anaerobically cultivated purple nonsulfur bacteria (e.g., *Rhodospirillum*), which utilize light in conjunction with organic compounds such as malate, pyruvate, or yeast extract, rather than O_2, as electron acceptors to release energy for growth and nutrition. *Chemoorganotrophs* encompass a vast number of organisms from all the microbial groups including archaea and bacteria, fungi, protozoans, and nonphotosynthesizing eukaryotic cells that obtain energy from the oxidation or other dissimilation of organic compounds.

Major Chemical Elements (C, H, N, O)

Unlike energy-procuring mechanisms, which vary widely between microbial genera and species, the anabolic or biosynthetic requirements of eukaryotic and prokaryotic cells are remarkably similar. Molecular mechanisms for intracellular communication and structural components share much chemistry whether they are bacterial *E. coli* cells or fungal *Aspergillus nidulans* cells. Whereas individual strains may vary in their ability to manufacture basic cell components and metabolites due to specific mutations at points within biochemical pathways, all known organisms share the same basic cellular needs for carbohydrates, lipids, proteins, and molecular combinations thereof. The elements carbon, hydrogen, nitrogen, and oxygen can be regarded as major essential elements of nutrition because they contribute in significant percentages to the biomass of cells. These major elements may be required in the gaseous form, as organically fixed compounds, or both. Iron, magnesium, phosphorous, potassium, sulfur, and other elements, although critically important to sustained growth and metabolic activity, are often termed minor or trace elements in that they collectively represent less than 2% of cellular dry weight.

Carbon. Nutritionally, carbon appears in cellular metabolism in two forms: carbon dioxide gas (CO_2) and organic molecules. CO_2 is required by all organisms for a number of biochemical functions including tricarboxylic acid (TCA) cycling and de novo fatty acid synthesis from acetyl-CoA. Although much of the carbon dioxide is metabolically recycled throughout the cell's biochemical pathways, an initial exogenous source of carbon dioxide is mandatory. Autotrophic organisms readily assimilate atmospheric carbon dioxide; yet, also, all heterotrophic organisms have some ability to fix this compound, although at a much-reduced capacity. Microbial cultures will sometimes experience long lag periods when subcultured at low-volume inoculum into fresh media. This phenomenon is often attributable to suboptimal CO_2 tension, which the culture slowly corrects with cellular respiration before any significant increase in biomass can occur. The term *CO_2 sparking* refers to the sudden increase in bacterial growth observed when a freshly inoculated culture is briefly sparged with the gas. An enriched atmosphere of 5% CO_2 is used for the cultivation of clinical (particularly for fresh isolates) bacterial strains. CO_2 incubators are most efficient for multipurpose cultivations, but the anaerobic gas-pack CO_2-generating systems offered by BBL (Becton Dickinson Microbiology Systems, Cockeysville, Md.) or Difco (Difco Laboratories, Detroit, Mich.) are recommended for small volume or infrequent use. Depending upon the absolute requirement, capnophilic (CO_2-requiring) bacteria can also be cultivated in jars in which inoculated test tubes or agar plates are placed along with a burning candle. The candle jar is then tightly sealed while the candle continues to burn until free oxygen in the jar is consumed. The resulting atmosphere achieves approximately 2% CO_2. Both the anaerobic gas-pack systems and candle jars have draw backs with respect to increased risk of culture contamination. Respiring cultures give off significant amounts of water, which raises the humidity within the sealed jars. As the jars themselves are normally not sterile at time of culture incubation, any environmental bacterial or fungal contaminant within the jar will proliferate in the additional moisture. Capillary action of excess water at the culture vessel openings (particularly petri dishes) can draw contaminating organisms or their spores into the culture vessel.

Autotrophs can use carbon dioxide gas as a sole source of carbon for biosynthesis. Although autotrophy is often regarded by nonbacteriologists as synonymous with photosynthesis, autotrophy and phototrophy are quite independent metabolic functions. For example, methanogenic archaea (e.g., *Methanobacterium* sp.), while nonphotosynthetic, are capable of utilizing carbon dioxide and hydrogen as sole sources of carbon and energy. Yet, as noted elsewhere, purple nonsulfur bacteria are photosynthetic but rely upon organic acids for metabolic reducing power and, ultimately, as the carbon source for biosynthetic functions. Whereas autotrophy is widely represented among prokaryotes, the ability to utilize CO_2 as a sole carbon source is not a general characteristic of fungi. Among filamentous fungi, *Cephalosporium* spp. and *Fusarium* spp. have been documented to be able to do so under long-term cultivation (11).

Heterotrophs obtain carbon principally from organic molecules such as carbohydrates, amino acids, or lipids. For heterotrophs, a compound utilizable for its carbon often also functions as an energy source. Although glucose is widely assimilated by many heterotrophic microorganisms, it is not necessarily the universal carbon source for culture medium supplementation. Many bacteria, whether their characteristics are genuswide or species-specific, show little or no stimulation by glucose and preferentially utilize amino or other organic acids. The marine genus *Oceanospirillum* falls into this category as do many individual species within the diverse genus *Clostridium*. In an extreme example, *C. aminovalericum* (ATCC 13725) appears to be limited to δ-aminovaleric acid as its singular assimilable source of both carbon and energy.

Some organisms can alternate between autotrophy and heterotrophy, depending upon environmental and nutritional conditions. For these cultures, biomass vigor and yield are most often greater with heterotrophic rather than

autotrophic conditions; however, secondary metabolites produced can vary considerably between the two conditions.

Hydrogen. Certain archaea and bacteria can grow by utilizing atmospheric hydrogen (H_2) as an electron donor and, depending upon the organism, assorted compounds including carbon dioxide, oxygen, nitrate, or sulfate as terminal electron acceptors (10). The hydrogen-oxidizing genus *Hydrogenophaga* (12) is facultative with respect to energy metabolism in that it functions chemoorganotrophically in a nutritionally heterotrophic aerobic environment by metabolizing various amino acids or sugars for sources of carbon and energy. However, in an atmosphere containing 60% H_2, 10% CO_2, 25% N_2, and 5% O_2, the organisms switch to chemolithoautotrophic pathways. *Methanobacterium thermoautotrophicum*, a strictly anaerobic archeaebacteria, metabolizes hydrogen and carbon dioxide as sole energy and carbon sources to yield methane and water (methanogenesis). For the anaerobic methanogens as well as the aerobic hydrogen bacteria, conditions favoring autotrophy require a high (60–80%) concentration of hydrogen for cultivation. Under all conditions, work with compressed hydrogen gas should be conducted with adherence to pertinent laboratory safety practices. Lower concentrations of H_2 (3–10%) are frequently included in gas mixtures for the cultivation of anaerobic bacteria in gas-generating systems and glove boxes where, in the presence of a palladium catalyst, trace amounts of inhibitory molecular oxygen are chemically reduced to water. The final form of hydrogen required for nutritional purposes is the molecularly combined element required by all organotrophic cells. Combined hydrogen originates from the catabolic metabolism of organic compounds assimilated by a cell as carbon and energy sources.

Oxygen. As already noted for combined hydrogen, the bulk of combined oxygen within cells originates from the complex nutritional compounds, including water, assimilated by the cell. If a medium provides enough matter to satisfy the carbon and energy requirements of a culture, then the combined oxygen requirements will also be met. However, molecular oxygen (O_2) is an important, often growth-limiting, variable in microbial nutrition. O_2 is the primary terminal electron acceptor for many aerobic and facultatively anaerobic organisms. In oxygen-depleted or diminished environments, cells with no other biochemical mechanisms for exploiting alternative electron acceptors stop metabolizing and cease multiplying. Facultative organisms switch to fermentative pathways. In either situation, the biochemical activity of the cells, as well as the secondary metabolites produced, is significantly altered by oxygen starvation. Be it large-volume fermentors or small-volume chemostats or Fernbach flasks, the need for O_2 sparging in bulk cultivation vessels is regarded as a basic cultivation requirement for aerobic organisms (13). Be that as it may, one should also be aware that significant O_2 gradients quickly form in liquid test tube cultures, particularly for those cultures that form surface mats or pellicles of growth. Various micrometabolies can occur within the cell population in a standing culture along the O_2 gradients from surface to bottom of the test tube that might, if unrecognized, cause undue problems in scale-up development. Cultures with strong O_2-requirements (even for those that do not form pellicles) may not grow to very high density in a test tube because of the limited liquid–air interface. Better liquid growth of these test tube cultures is effected by incubating the tube at a slant to increase the surface area of the culture, or by shaking.

Nitrogen. Cells can assimilate nitrogen by two methods. Nitrogen fixation is the ability to acquire elemental nitrogen from the air and to molecularly combine the element into chemical forms utilizable for biosynthetic processes. The bacterial genera *Azotobacter*, *Bradyrhizobium*, and *Rhizobium* are most recognized for this characteristic, although some facultative anaerobes as well as anaerobic bacterial species also have the capacity. Fungi are not recognized as competent nitrogen fixers. Nitrogenase activity is induced with cultivation on nitrogen-free media. Growth on chemically defined "nitrogen-free" media might presumptively select for potential nitrogen fixers; however, trace amounts of nitrogen from agars, chemicals, and glassware are present in sufficient concentration to permit effective scavenging and growth (albeit, less than optimal) by many non-nitrogen-fixing organisms. Conclusive nitrogen fixation is best determined by the acetylene reduction assay (14).

The second, and most widely utilized, microbial mechanism for nitrogen procurement is via combined nitrogen assimilation from inorganic or organic molecules. Media formulations satisfy combined nitrogen requirements either with inorganic salts [e.g., NH_4Cl; NH_4NO_3 or KNO_3; $(NH_4)_2SO_4$] or from organic compounds (e.g., amino acids). The ability to utilize inorganic salts as a nitrogen source varies widely even within species. The chemolithotrophs as a group effectively utilize inorganic ammonium or nitrate salts as sole sources for nitrogen. For other organisms, the inability to grow under chemically defined medium conditions with ammonium or nitrate salts as a sole nitrogen source is due not so much to the inability to assimilate the ion directly from the medium or to incorporate it as the glutamine precursor to all subsequent biosynthetic nitrogenous pathways, but to the specific need for additional nitrogenous growth factors such as purines, pyrimidines, amino acids, or small peptides. These kinds of nitrogenous nutritional requirements are discussed in greater detail in the section "Growth Factors".

Nitrate is not a preferentially assimilated or utilized source of nutritional nitrogen for many bacteria or fungi; concentrations as low as 0.2% are inhibitory to the growth of some bacteria. However, the use of nitrate in energy metabolism as a terminal electron acceptor is a widely used mechanism among bacteria and is termed dissimilatory nitrate reduction or denitrification. The biochemical pathways involved with nitrate assimilation for cellular syntheses and for dissimilatory nitrate reduction are different. Nitrate reduction occurs with either aerobic or anaerobic bacteria. Strains of *Clostridium perfringens*, for example, utilize nitrate as an electron sink, which causes a shift from butyrate to acetate fermentation and an increase in growth yield (15). All recognized species in the extremely

halophilic genus *Haloarcula* are capable of using nitrate as a sole nitrogen source and, in addition, are strong denitrifiers with the ability to readily tolerate and reduce 3 M concentrations of KNO_3 under either complex or chemically defined media conditions; aerobic growth of haloarculas with liquid media lacking nitrate supplementation is markedly limited compared to aerobic growth in media supplemented with 0.2% KNO_3 (R. J. Cote, unpublished data, 1995).

Minor and Microelements (Fe, Mg, K, P, S, and others)

Trace metal function and speciation have been extensively studied in bacteria and fungi (16,17). Nutritionally essential metal ions function metabolically as enzyme cofactors in membrane transport and in structural components. Microbial response to trace metals is detectable at extremely low concentrations (i.e., microgram or nanogram quantities, depending upon the metal). However, the bioavailability of a particular metal is markedly dependent upon proper speciation of the ion as it occurs in the culture medium (18). Depending on interaction with other medium components or methods of sterilization, metal valences can change to nonusable forms, or the ions can irreversibly bind as insoluble precipitates. Although problems with metal-binding phenomena are often directed toward the nutritionally essential metals, the potential for significant and irreversible binding of ions to media components must also be taken into consideration in studies of bacterial resistance to heavy metals. Reports of bacterial resistance to silver and selenium (19,20) could not be verified when the metals were added to culture media in a manner that eliminated their binding to peptones (R. J. Cote and R. L. Gherna, unpublished data, 1994). Similar observations made regarding media influence and metal sensitization of *Enterobacter cloacae* to silver sulfadiazine (AgSu) (21) are explained by heavy-metal binding under complex media conditions. Tilton and Rosenberg (22) discuss the issue of heavy-metal sequestration in a study of the effects of agar on silver inhibition.

Microelement supplementation is not normally required in media formulated with chemically complex components such as yeast extract or protein hydrolysate because the essential metals are naturally present in these materials in sufficient concentration to provide optimum culture growth and performance. Well-designed, chemically defined formulations include calcium, iron, magnesium, manganese, phosphorous (as phosphate), potassium, sodium, and sulfur (as sulfate) in the form of inorganic salts. Other micronutrients may be required in the medium, depending upon the nature of the metal function under study or, in the case of general culture propagation, the purity of the water and cleanliness of the glassware used to prepare the medium. Understanding of cell physiology, as well as any physical interactions within the culture system, is a basic requirement when working with nutritional trace metals. For example, although boric acid is often included in trace metal mixes, a boron deficiency is practically noninducible for cultures grown in borosilicate glass vessels because enough of the element leaches from the glass to satisfy nutritional needs. Table 1 lists metal mix formulations that are useful for chemically defined cultivation of numerous microorganisms. As noted elsewhere in this article, the use of tap water to satisfy trace element requirements is not recommended because regional and seasonal variability in water purity can introduce considerable ionic variation in a medium, with frequently deleterious results.

Table 1. Trace Metal Mixes for Bacteriological Media

Metal mix A: Balch's trace element solution

Nitrilotriacetic acid	1.5 g
$MgSO_4 \cdot 7H_2O$	3.0 g
$MnSO_4 \cdot H_2O$	0.5 g
NaCl	1.0 g
$FeSO_4 \cdot 7H_2O$	0.1 g
$CoCl_2 \cdot 6H_2O$	0.1 g
$CaCl_2$	0.1 g
$ZnSO_4 \cdot 7H_2O$	0.1 g
$CuSO_4 \cdot 5H_2O$	10.0 mg
$AlK(SO_4)_2 \cdot 12H_2O$	10.0 mg
H_3BO_3	10.0 mg
$Na_2MoO_4 \cdot 2H_2O$	10.0 mg
Distilled water	to 1.0 L

Suspend nitrilotriacetic acid in about 500 mL water and adjust to pH 6.5 with KOH pellets to dissolve. Add remaining salts one at a time. Bring final volume to 1.0 L with distilled water. Use 10.0 ml trace element solution per 1 L medium.

Metal mix B: modified Balch's trace element solution

To 1.0 L metal mix A add	
$NiSO_4 \cdot 6H_2O$	30.0 mg
Na_2SeO_3	20.0 mg
$Na_2WO_4 \cdot 2H_2O$	20.0 mg

These metals are included for the cultivation of methanogenic bacteria under chemically defined conditions.

Source: From Ref. 23, p. 158.

The whitish haze or precipitate that develops in some media upon standing or after autoclaving is frequently caused by imbalanced metal compositions. Divalent and trivalent cations (i.e., Ca^{2+}, Fe^{2+}, Fe^{3+}, and Mg^{2+}) will spontaneously bind with OH^- and PO_4^{2-} groups in the medium to form insoluble hydroxides and phosphate complexes. The condition is most evident at neutral to alkaline pH values and is intensified by heat, either boiling or autoclaving. Similarly, precipitation caused by metal binding to proteinaceous material in the medium and darkening of chemically defined agar media are also frequently encountered in poorly designed or manipulated formulations. All of these adverse metal interactions are generally regarded as deleterious and growth limiting. Depending upon the composition of the medium, these kinds of chemical interactions can be eliminated by sterile filtration or separate autoclaving of inorganic salts and any phosphate buffers, and separate autoclaving of agar.

The use of chelating agents can also help to selectively maintain the solubility of polyvalent cations. Care must be taken when using powerful chelators such as ethylenediaminetetraacetic acid (EDTA) or nitrilotriacetic acid

(NTA) because overchelation also depletes bioavailability of the metals. For microbial media, 0.5–2.0 mg/L EDTA or 15–150 mg/L NTA are reasonable concentrations. Mild chelation can be achieved with the use of specific amino acids (e.g., glycine, histidine) or carboxylic acids (acetate, citrate, succinate, tartrate). However, their use and concentration must be evaluated with respect to the function of the medium and the physiology of the microbial culture. For example, carboxylic acids are readily utilized carbon and energy sources for many microorganisms, and glycine is inhibitory at 1% concentration to many bacteria.

Under iron-deficient growth conditions, some bacteria produce and secrete chelators called siderophores to improve transport of ionic iron into the cell (24,25). Siderophores have been found in a number of diverse genera including *Aureobacterium (Arthrobacter) flavescens* (26), *E. coli*, *Klebsiella* (27), *Pseudomonas*, and *Mycobacterium avium* subsp. paratuberculosis (28). Other organisms rely upon the hemin molecule in blood as an easily assimilable iron chelate. A growth factor requirement for hemin is common among medically important bacteria. Hemin requirement (the X-factor) is diagnostic in the identification of *Haemophilus influenzae*. *Porphyromonas gingivalis* (29) and *Vibrio vulnificus* biotypes 1 and 2 require iron in the form hemin for growth and virulence (29,30). Whereas a number of classical media formulations (e.g., GC Agar) include crude hemoglobin as the source of the iron porphyrin molecule, use of hemin, instead, will produce a medium free of the heavy debris associated with hemoglobin. Hemin is used at a concentration of 0.5 mg/L medium. To improve solubility, initially suspend the compound in 1–2 mL 1 N NaOH, then add the hemin-NaOH solution to the bulk of the medium.

Many of the metals listed here that have been historically categorized as needed by microorganisms only in trace amounts have, in recent years, come to be evaluated on a different metabolic scale. Sulfur serves as an energy source for numerous species of bacteria, and thiosulfate seems to be utilized for energy by some fungi. While the role of sulfate as the terminal metabolic electron acceptor for the many anaerobic sulfate-reducing bacteria has been long recognized, new genera are currently being isolated that possess the ability to utilize one or more biologically important metals [(Fe (III), Mn (IV)], U (VI), and others] as terminal electron acceptors in biochemical dissimilatory metal reduction pathways (31). Because the function of terminal electron acceptor is not recyclable within the cell and bears a heavy metabolic load, the amounts of metal ion needed per liter for efficient energy transfer are orders of magnitude greater than those needed to satisfy trace element nutrition. One to four grams per liter is a guideline, depending upon the specific metabolism of the organism as well as the molecular weight and form of the metal salt selected for use.

Growth Factors

Given less than a dozen essential chemical building blocks and a source of metabolic energy to power the cellular machinery, some organisms, such as the wild-type *E. coli*, can synthesize every molecule needed for cell maintenance, growth, and reproduction. Many other microorganisms (even strains within the same species) are auxotrophic in that they have lost the ability to synthesize one or more key organic compounds essential to their physiology. These missing compounds are called growth factors and must be present in the culture medium if cell growth is to be vigorous, metabolically productive, and sustainable.

L-Amino Acids. L-Amino acids are common growth factor requirements for prokaryotes and eukaryotes. Many bacterial genera (e.g., *Lactobacillus* spp.) have multiple amino acid requirements. The requirement may be absolute or conditional and influenced by the presence and concentration of other amino acids (32–34). Similarly, a phenotypic amino acid requirement may disappear with the addition of an appropriate synthetic pathway intermediate. For example, a tryptophan "requirement" may actually be caused by a metabolic block earlier in the tryptophan pathway and might be satisfied by one of the earlier pathway intermediates (i.e., aspartic acid or indole). Because amino acids may share a common transport system, an excess of one amino acid can result in an increased need for a competing amino acid. Amino acid supplementation is not required in complex media containing peptones or other protein hydrolysates. In chemically defined media, the addition of 20–50 mg/L amino acid stimulates prokaryotic growth (100–200 mg/L for yeast), although adjustment of the concentrations is necessary for media optimization with multiple amino acid-requiring strains to chemically balance the formulation. The most notable exception is the requirement for glutamic acid, which, as precursor to all other amino acid biosyntheses, must be added at 200–500 mg/L to achieve optimal growth of a bacterial culture. Amino acids, except glutamine, are stable with autoclave sterilization at the working concentrations used in media. Glutamine in aqueous solution will spontaneously break down with time to ammonium and pyrrolidone carboxylic acid residues. This reaction occurs more rapidly at elevated temperatures (37 °C and above). The use of glutamine peptides in a medium can eliminate this problem, however, they are expensive for large-scale use. Concentrated stock solutions of amino acids may not be soluble, but neutralization of the stock solution with NaOH or KOH will increase solubility with the formation of amino acid salts for many of these compounds. Stock solutions of cysteine will autooxidize to cystine unless stored anaerobically (e.g., under N_2).

For many years, the perceived need by bacteria for combined amino acids (peptides or proteins) was satisfied by the argument that peptide transport utilized different systems than did amino acids; thus, the need for a specific peptide was actually an alternative cellular method for obtaining a specific amino acid contained within that peptide via a biochemically noncompetitive mechanism. However, increased knowledge of the role of peptides in protein trafficking suggests that peptide auxotrophies can affect cellular physiology and communication (35). Successful laboratory cultivation of *Thermoplasma acidophilum*, having no other known growth factor requirements, is dependent upon specific lots of powdered yeast extract containing a methanol-soluble growth factor, which purification re-

vealed to be a small 1,000-MW peptide (36). Similarly, recent laboratory cultivation of *Treponema maltophilum* also included the requirement for a methanol-soluble growth factor from yeast extract (37). The treponeme requirement has not been further purified. Finally, successful cultivation of the protozoan *Entamoeba histolytica* in complex media enriched with peptone, vitamins, and serum is nevertheless dependent upon an unknown growth factor found only in certain lots of powdered yeast extract (38). Whether the elusive yeast extract growth factor is the same for all these metabolically dissimilar organisms or whether the relatively gentle methods used in the manufacture of yeast extract simply yield a large pool of assorted intact peptides possibly required by some microorganisms poses an interesting question.

Whole Proteins. Whole proteins (e.g., albumin, blood, egg, meat, serum) are not known to be essential requirements for microbial growth, although they are an integral medium component for many nutritionally fastidious microorganisms. Clinical isolates of mycobacteria have been historically cultivated with Lowenstein–Jensen medium, containing whole egg and inspissated (coagulated by heating at 80–85 °C for 45 min) to form a solid matrix (39). Most mycobacteria, however, can be effectively cultured with Dubos medium or Middlebrook 7H9/or 7H10 media, in which the egg protein is replaced by the more defined components bovine serum albumin, catalase, dextrose, and oleic acid (40). Variations of chopped meat media (2) were for many years the recommended formulations for cultivation of pathogenic clostridia and similar bacteria, yet many of these strains give equal cell yield with the more defined, and more easily prepared, Reinforced Clostridial medium (40), although this basal medium frequently requires the addition of hemin and/or vitamin K_1 to satisfy the growth factor requirements of some strains; additionally, 0.5% cellobiose added to reinforced clostridial medium will enhance the growth of cellulolytic clostridia.

Defibrinated rabbit or sheep blood (5–10% v/v) is a common addition to media used for the cultivation of microorganisms pathogenic to animals. Human blood is not recommended for use for microbial culture work for health safety reasons and also because outdated blood obtained from banks is normally citrated. Citrated blood of any type should not be used in culture media because the anticoagulant may have negative effects on culture growth or hemolytic reactions. Although diagnostic hemolytic reactions differ between rabbit and sheep bloods, the ability to support microbial growth is generally the same. Human or sheep bloods, unless chocolated, are not used for the cultivation of *Haemophilus* spp. or other organisms requiring NAD because these fluids contain enzymes that inactivate the growth factor. Chocolate blood agar is prepared by heating the medium containing blood to 70–80 °C for 15 min. Purchase of whole animal blood should be from a reputable supplier and pooled from animals on an antibiotic-free diet. Well-collected and processed defibrinated blood is free from clots and cell lysis. Although commercial blood suppliers carefully collect their products, each lot of blood should be checked for sterility prior to use. Not all types of bacterial contaminants found in blood will grow well in either of the U.S.P.-recommended sterility test media (i.e., Fluid Thioglycollate and Soybean–Casein Digest). To broaden the screening procedure, the battery of media should be enlarged to include Brain Heart Infusion (BBL 11059 or Difco 0037) to enhance the recovery of somewhat fastidious contaminants, as well as HTYE (Trypticase peptone (BBL), 0.5%; yeast extract 0.2%; HEPES, acid, 0.4%; adjust to pH 7.2 $\pm$ 0.1 prior to sterilization at 121 °C for 15 min), which has wide versatility in the ability to cultivate diverse environmental heterotrophic organisms that do not grow well with conventional, general bacterial growth media. Whole blood should be stored under refrigeration (not frozen) and should be incorporated into media as soon as possible to help stabilize the fluid. Whole blood can not be stored longer than 2–3 weeks without significant deterioration. However, when stored at 4–8 °C, media containing blood can, for most media purposes, have a 4-month shelf life.

Serum (5–10% v/v) is an alternative to blood for the cultivation of pathogens. Horse, fetal, or newborn bovine sera are most frequently used for the cultivation of bacteria, whereas adult bovine serum is specified for the cultivation of many pathogenic protozoans. Sera vary in their intrinsic enzyme activity. For example, horse serum degrades glutamine, a necessary component in many animal cell line media formulations, and as already noted for chocolatizing whole blood, sheep and human blood can inactivate NAD. Preparation instructions for many media formulations call for serum heat inactivation (incubation of the serum at 56 °C for 30 min) to inactivate the protein complement. This temperature may not be high enough to eliminate adverse serum enzyme activity. Conversely, heat inactivation may not be a universally advantageous specification for all microbial growth and performance. A researcher should investigate types of sera as well as the effects of heat inactivation upon the performance of a microbial culture.

Whole proteins and their hydrolysates supply an eclectic and undefined mix of nutrients, detoxifiers, and other molecules that slowly release or absorb metabolites that may be required for cell nutrition in small amounts but quickly become toxic at elevated levels (e.g., fatty acids). Proteins hydrolysates supply a myriad of nutritional growth factors and help control physical parameters such as pH and osmolarity.

Be that as it may, the use of proteins and their derivatives in media formulations is becoming problematical with respect to biotechnology for two reasons: (*1*) Because many of the products of interest produced by microorganisms are proteins or substances with chemical affinity to proteins, presence of a complex proteinaceous medium component can significantly complicate a product purification process. (*2*) Biotechnology applications under regulatory compliance are facing increasing scrutiny about the potential presence of unidentified adventitious agents in media components. The international concern with bovine spongiform encephalopathy (BSE) is real and pervasive throughout the field of biotechnology, ranging from final product, to the type and source of medium components used in the preparation of a frozen or lyophilized master cell bank used in the manufacture of a product.

The European Commission decision 97/534/EC bans the importation of specified risk materials (SRM) into the European Union member countries (including Australia, Canada, and New Zealand). U.S. biotechnology companies wishing to deal with the European market must be able to provide a declaration to customs agents that assures their product does not contain, nor is derived from, the specified SRM as defined in 97/534/EC. The specified risk material, in this instance, is defined as "the skull, including the brain and eyes, tonsils and spinal cord of bovine animals over 12 months, ovine and caprine animals which are over 12 months or have a permanent incisor tooth erupted through the gum, and the spleen of ovine and caprine animals" (41).

Many of the peptone hydrolysates used in biotechnology are of animal (bovine) origin and are manufactured from "leftovers" of the slaughterhouse trade. As large-scale animal processing is not a clean operation, SRM contamination of raw material used in the manufacture of peptones is probable, particularly where source material is not directly traceable to government-regulated, clean herds.

The degree of concern with animal peptones with respect to 97/534/EC varies with industry and product. For example, from a regulatory perspective, the manufacture of biopharmaceuticals, especially direct injectables, demands high scrutiny for presence and elimination of any SRM. But the use of an animal peptone in a medium used solely for clinical diagnostic purposes (e.g., a peptone used clinically in MacConkey agar to differentiate *E. coli* from other normal intestinal flora), might be of lesser concern because the peptone never comes in direct contact with animals or humans at risk to BSE or similar prion agents. Acid-hydrolyzed animal peptones are generally of lower risk than enzymatic hydrolysates because the acid and heat associated with acid hydrolysis inactivates viral agents. Ultrafiltration of peptone solutions and sera has been proposed as a mechanism for eliminating possible BSE contaminants as well as reducing endotoxin levels. In any event, the concern with BSE and related animal/human pathogenic agents has caused an escalation of documentation requirements for the tracking of animal-derived media components. Certificates of origin for sera, meat, and milk peptones can, and should, be obtained from reputable vendors of these materials. Although certificates of analysis are now routinely supplied by raw material vendors to the biotechnology industry, certificates of origin are currently customer driven and are generally supplied only as requested.

BSE has fostered much current interest in the development of nonanimal peptones. With yeast, soy, or wheat as the source protein, some of these "veggie peptones" have interesting nutritional attributes but, nevertheless, still suffer the drawback of elevated intrinsic carbohydrate levels characteristic of plant peptones.

Difficulties in medium optimization and performance currently preclude wide use of serum-free or SRM-free media with animal cell lines. However, design of SRM-free media for most bacteria and yeast are feasible and cost-effective.

Vitamins. Vitamins are biochemical catalysts and appear in cellular metabolism as coenzymes or as the prosthetic group of enzymes. Vitamins elicit microbial growth response at very low levels, at nanogram concentrations for vitamins such as biotin and cyanocobalamin (vitamin B_{12}). Although many microorganisms are capable of synthesizing all of the required vitamins, many others are auxotrophic for one or more of these growth factors. The genus *Lactobacillus* characteristically has multiple vitamin requirements, and several strains within the group were widely used for many years in protocols for vitamin assays until replaced by instrumentation-based methodologies. The U.S.P. 23 still retains *Lactobacillus delbrueckii* subsp lactis (*L. leichmannii*) (ATCC 7830) as the assay organism for vitamin B_{12}. Marine bacteria frequently require B_{12}, whereas thiamine and biotin are necessary for the growth of many mycelial fungi. Menadione or its more active form, vitamin K_1, is a supplement for the cultivation of some clinical anaerobic bacteria. As noted with amino acids, a shift in medium composition can spare the need for a specific vitamin. *E. coli* (ATCC 10799; NCIB 8134) requires vitamin B_{12} when cultivated with a chemically defined minimal medium; when grown in the presence of 20 μg/mL methionine, the vitamin requirement is eliminated (42). Vitamin auxotrophies can appear as a need for the vitamin precursor, the intact vitamin, or as conjugates in coenzyme form. Koser noted, many years ago, that vitamin and amino acid auxotrophies in thermophilic bacilli are temperature variable (43). NAD and cocarboxylase (thiamine pyrophosphate) are coenzymes required by many pathogenic bacteria.

All of the vitamins are water soluble and heat stable at working concentrations used in media; coenzymes should be filter sterilized, as should concentrated vitamin stock solutions. Most vitamin stock solutions remain water soluble in significant concentrations, although biotin, folic acid, nicotinic acid, and riboflavin require neutralization with NaOH to retain solubility of the concentrate. Vitamin K and thioctic acid stock solutions are ethanol soluble. Ascorbic acid, used in microbial cultivation as an antioxidant, has a short life (about 3 weeks) in media; use of ascorbate salts rather than the free acid will improve stability. Table 2 lists a complete vitamin mix for the cultivation of bacteria and fungi. Yeast extract at a concentration of

Table 2. Balch's Vitamin Solution

p-Aminobenzoic acid	5.0 mg
Folic acid	2.0 mg
Biotin	2.0 mg
Nicotinic acid	5.0 mg
Calcium panothenate	5.0 mg
Riboflavin	5.0 mg
Thiamine HCl	5.0 mg
Pyridoxine HCl	10.0 mg
Cyanocobalamin	100.0 mg
Thioctic acid	5.0 mg
Distilled water	1.0 L

Mix all vitamins in distilled water. Adjust to pH 7.0, if necessary, to dissolve. Filter sterilize and store in brown bottle with refrigeration.

Source: From Ref. 44, p. 160.

0.1–0.5% is used as an undefined vitamin supplement for microbial growth in complex media; however, concentrations of yeast extract as low as 0.05% can elicit growth stimulation by organic growth factors, which may or may not be vitamins.

Nucleic Acid Precursors. Nucleic acid precursors are often required by nutritionally fastidious organisms when cultivated in defined or semidefined media. In some instances, the purine or pyrimidine will satisfy the requirement; more commonly, an organism needs preformed nucleosides or nucleotides. Organisms with these types of synthetic deficiencies will show a growth response to 5–50 μg/mL filter-sterilized additive.

Lipids. Lipids or their precursors are required by diverse organisms. Some pneumococci have a unique requirement for choline as a lipid precursor. *i*-Inositol is a structural component of some phospholipids. Inositol-deficient strains of *Neurospora crassa* and *Saccharomyces cerevisiae* have been used in studies of phospholipids and cell membrane structure. Fatty acids stimulate the growth of some fungi; certain mutants of *N. crassa* are auxotrophic for fatty acid synthesis (45). Rumen bacteria require short-chain fatty acids for growth. Many mycobacteria have a generalized requirement for exogenously supplied lipid. Various media formulations for nonglycerophobic mycobacteria include 2.0 mL/L glycerol, 0.05 g/L oleic acid, or 0.5 g/L Tween 80 to satisfy this requirement. Most of the bacterial mollicutes (anaeroplasmas, mycoplasmas, and spiroplasmas) require sterols. Cholesterol is a prevalent component in semidefined media formulations for these organisms, whereas fetal calf or horse serum supplies the growth factor in general-use complex mycoplasma media. *Eubacterium coprostanoligenes* (ATCC 51222) and *Eubacterium* sp. (ATCC 21408) represent a novel group of bacteria that use cholesterol and similar sterols as terminal electron acceptors (46,47). The former organism does not require cholesterol but metabolizes and requires phosphatidylcholine for growth, whereas the latter requires cholesterol as well as an alkenyl ether lipid for growth. These bacteria have been used in probiotic studies on decreasing the level of cholesterol in the digestive tract. All the lipids just mentioned are stable to autoclave temperatures. Short-chain fatty acids are water soluble; higher molecular weight lipids are soluble in less polar solvents such as ethanol or chloroform. In the presence of other media components at alkaline pH, some high-weight lipids can form growth-limiting soapy complexes. Separate sterilization of the fatty compounds helps to alleviate the problem.

Polyamines. Polyamines (e.g., putrescine, spermidine, spermine) are formed via the arginine pathway and have undefined functions in the cell. Research on polyamines has been done principally in vitro where the compounds bind to and neutralize charges on DNA and RNA. Although high concentrations of Mg^{2+} exert the same general ionic neutralization effect, only polyamines have the ability to influence correct tRNA conformation. Much of the cellular polyamine appears bound to ribosomes. In *E. coli*, putrescine and spermidine account for half the total product of the arginine pathway (48). Polyamine supplements as growth stimulants appear in media for the cultivation of *Ureaplasma urealyticum* (2p.530), *Haemophtlus parainfluenzae* (49), *Veillonella atypica* (50), and other bacteria. Where complex media are used, there is generally a sufficient polyamine content in yeast extract to satisfy the requirements of deficient strains.

Unique Growth Factors. Any organism has the possibility to, either spontaneously or through human intervention, develop mutations in biosynthetic pathways. Thus, unusual or strain-specific growth factors are always a consideration in the development of media for "unculturable" organisms. Deliberate mutation for the inability to synthesize the cell wall component diaminopimelic acid was engineered into early recombinant host strains of *E. coli* as an attenuation factor to reduce the possibility of survival outside of laboratory cultivation. Until recently, *Bacteroides forsythus* had only been culturable in cocultivation with other oral anaerobic bacteria, until its requirement for *N*-acetylmuramic acid was recognized (51). As noted earlier for the L-amino acids, the inability to synthesize functional signal and transport peptides may be limiting factors in the protein chemistry of some organisms.

PHYSICAL PARAMETERS

pH

pH is a strong selection factor for microorganisms. Although most bacteria have a wide pH range ($\pm$3.0 pH units) at which measurable growth occurs, optimum growth is limited to a narrow range. Minimum and maximum pH values can often be distinct and sharply defined. Members of the genus *Bacillus* will grow at pH 6.8 but not at pH 6.5. Extreme halophiles with a biomass yield optimum at pH 7.4 continue to grow slowly at pH 7.0 and exhibit little or no growth below pH 6.7. Eukaryotic and prokaryotic organisms shift their biochemistry according to the pH of their environment and thus change cellular characteristics as well as metabolic by-products. Such changes in cultivation parameters can be consciously directed toward a particular application such as secondary metabolite production. Unfortunately, the effects of pH shifts are too often not critically examined, particularly at the small-scale cultivation stage. For example, in an unbuffered test tube culture of anaerobic carbohydrate-fermenting clostridia, acid production can quickly drop the medium to below pH 4.0, well outside the pH minimum of 6.5 necessary for significant growth. Either the culture will sporulate (if medium conditions are conducive to the process) or viability quickly diminishes as acid concentration increases. Effective media design must be constructed around the pH at which the medium is developed and will be used. Radical adjustment to media naturally posed at a particular pH too often leads to precipitation of components or other problems with performance. Inclusion of buffers, as well as the types of buffers used, are important considerations in media design.

Incubation Temperature

Incubation temperature influences microbial metabolism both with respect to the rates at which cellular processes run and the rates at which nutrients are assimilated. Strains of *Serratia marcesans* produce pink-pigmented colonies at 26 °C but are colorless or slow to produce pigment at 37 °C. Temperature effects may sometimes be subtle. Bacteria alter the fatty acid compositions of their phospholipids according to temperature. The proportion of low-melting-point fatty acids increases as growth temperature decreases (52,53). Organisms may display fairly broad temperature tolerances within a range, but as with pH, optimum growth range is fairly narrow. Psychrophiles characteristically exhibit efficient metabolic rates at temperatures below 15 °C; the obligate psychrophilic bacterial species *Colwellia psychroerythrus* has a temperature optimum around 10 °C and lyses above 20 °C (54). Mesophiles are active between 20 °C and 40 °C. Free-living environmental strains are more tolerant of wider temperature ranges than are pathogenic strains that effectively function within only a few degrees of the body temperature of their hosts. Thermophilic microorganisms exist at temperatures above 40 °C. Although some fungi (e.g., *Mucor miehet* and *Sporotrichum thermophile*) can grow at temperatures as high as 50 °C, extreme thermophily, characterized by the ability to metabolize at temperatures above 70 °C, is restricted to members of the archaea and, to a lesser extent, the bacteria.

Oxidation–Reduction Potential

Oxidation–reduction (redox) potential of the culture environment determines the ability of a microorganism to initiate growth. Redox values of media and respiring cultures are measured with a pH meter set to the millivolt (mV) mode and a platinum electrode. Aerobic organisms require positive millivolt conditions. Obligate microaerophiles require 0 to −100 mV; clinical anaerobes less than −100 mV; obligate and extreme anaerobes less −200 mV. Actively metabolizing anaerobic cultures can drop redox levels below −300 mV. Media for obligate microaerophiles and anaerobes include reducing agents to lower the redox potential to a range optimal for a culture inoculum to be able to initiate metabolic activity. Subsequent cellular growth will maintain an overall redox condition reflecting the varying metabolic activities of the culture. Organic matter such as meat particles, yeast extract, and reducing sugars exert reducing effects in a medium. Filling a culture vessel almost to capacity with medium or adding 0.2% agar to the medium to retard diffusion of O_2 will be adequate to culture less fastidious microaerophiles. Anaerobic organisms with more stringent redox requirements need one or more chemical reducing agents included in the growth medium. Frequently used reducing agents, listed in terms of approximate redox potential in millivolts, include sodium thioglycollate (−100 mV), cysteine hydrochloride (−210 mV), titanium (III) citrate (−480 mV), sodium sulfide (−571 mV), and sodium dithionite (−600 mV). The overall redox level of a medium containing any of these reducing agents will, of course, be dependent upon the chemical interactions of all media components as well as the reducing agent. Media chemically reduced with any of these compounds should not be allowed to reoxidize, because the chemical structure of the reducing agents can be irreversibly altered by oxidation and lose effectiveness; in some instances, the oxidized forms of reducing agents can be toxic to cells (55). For some strict anaerobes cultivated under chemically defined conditions, it is best to add the reducing agent just prior to inoculation because reduced conditions are difficult to maintain under prolonged media storage. Ljungdahl and Wiegel (56), Breznak and Costilow (57), and Balch et al. (58) provide excellent detailed instructions for the cultivation of strict anaerobes.

Detoxifiers

Detoxifiers are required for the cultivation of some fastidious microorganisms. In the natural environment, metabolic by-products released by some organisms are enzymatically neutralized or utilized by other organisms for nutritional or other biochemical needs before the compounds accumulate to toxic levels. In vitro, for a pure culture lacking the appropriate enzymatic mechanisms to neutralize its own metabolic wastes, these metabolites can build to inhibitory concentrations before significant culture growth has occurred. The inability to break down hydrogen peroxide by organisms lacking catalase, and strain sensitivity to fatty acid accumulation are two common problems associated with metabolite accumulation in closed culture conditions. To cultivate these types of microorganisms, some media require the addition of detoxifiers, of which most are absorbents. Frequently utilized detoxifiers include blood/serum, bovine serum albumin (BSA, fraction V), catalase, and charcoal. More recently, liposomes have been utilized to slowly release lipid growth factors to deficient organisms, thus eliminating the toxic effects that a full "dose" of a medium lipid supplement would have for these organisms (59). Similarly, cylodextrins have been shown to effectively chelate and then slowly release potentially toxic, yet essential, nutrients for some clinical bacteria under semidefined growth conditions (60,61).

MEDIA DESIGN AND COMPOSITION

Optimal growth and performance of a microbial culture is most influenced by the design and composition of its growth medium. There is no one "best" medium for a particular application; however, there exist many examples of poorly designed and/or poorly manipulated media, or well-designed media being used for an inappropriate purpose, in which a microorganism manages to survive in spite of the medium and not because of it. Effective media design comes from an understanding of an organism's biochemistry and physiology, as well as the nutritional ecology of the environment from which the organism was isolated. Although isolation media are selective and frequently nutritionally suboptimal to limit the growth of unwanted organisms, general growth or maintenance media should provide conditions to sustain a culture at optimal morphology and physiology through repeated subcultures. Under microscopic examination, bacteria grown under favorable conditions do not have a misshapen or shriveled

appearance; cell inclusions are generally minimal or absent; and cell lysis and/or spheroplast development is minimal or absent. At logarithmic phase of growth, motility is detectable for motile strains. Depending upon its application, a medium may be further defined in terms of chemical composition or optimized to increase biomass or cell product yield. An organism exhibiting optimal growth in a nutritionally balanced maintenance medium may not give high yields of a much-wanted metabolic by-product. Thus, medium optimization is a process directed toward a specific application. With so many nutritional and physical variables impacting upon microbial physiology, attempts at media optimization can become unnecessarily expensive, time-consuming, and largely unproductive when blindly tackled one variable at a time. Statistical methods using modified factorial approaches have been developed to aid in systematic media optimization for large-scale development (62), but for general laboratory purposes, knowledge and understanding of an organism's metabolic pathways and physical growth parameters can indicate specific areas in which quick design improvements to an existing formulation can effect significant results.

Media formulations may be chemically defined, in which discrete compound identity and concentration of all components are known, or chemically undefined, in which ingredients include one or more chemically complex substances such as protein hydrolysates and extracts. The choice of defined or undefined medium is dependent upon its application. Chemically defined media are useful in biochemical or metabolic studies of organisms. These types of formulations are also sometimes used in large-scale bacterial or yeast fermentations to produce products which further require stringent purifications such as those from recombinant DNA technology. General laboratory growth media for heterotrophic microorganisms as well as industrial media for large-scale production of metabolites (e.g., antibiotics, alcohols and other solvents) are often chemically complex. Scale-up of any medium formulation must include considerations of availability, cost, performance and lot-to-lot reproducibility of components as balanced against the value of the product. Chemically defined media can be expensive, particularly if the ingredients include a lengthy list of growth factors.

Microorganisms can transport only small molecules across their membranes. Large molecules (e.g., DNA, starch, proteins) can be nutritionally utilized by a microorganism only if it can synthesize and secrete the appropriate extracellular enzymes to digest large molecules into subunits small enough for assimilation into the cell. Because this type of enzymatic activity is absent from many organisms, in vitro microbial cultivation relies upon predigested macromolecules in the form of hydrolysates, peptones, and extracts and simple carbon compounds as nutritional substrates.

Protein Hydrolysates (Peptones)

Protein hydrolysates supply a mixed source of nitrogen and growth factors in the form of amino acids, nucleic acid fractions, peptides, and salts to a complex growth medium. For organisms possessing the metabolic capacity, protein hydrolysates can also serve as complex carbon and energy sources. Industrial large-scale fermentations using bacteria and fungi effectively use by-products from other industrial processes as proteinaceous material for media, including cottonseed, yeast, and soy. Although these substrates are chemically undefined, published analyses of their nutritional content greatly help in determining the appropriateness of a product for inclusion in a medium (63). Manufacturers of refined media and media components offer water-soluble products with more controlled nutritional composition. Table 3 lists and briefly describes the extracts and hydrolysates frequently encountered in microbiological laboratory media. Meat, milk, and soy commonly serve as starting material for the manufacture of hydrolysates. Digestion of these proteins can be by acid, alkali, or enzymatic hydrolysis. The source protein and method and degree of hydrolysis determine the nutritional characteristics of a hydrolysate. Prolonged hydrolysis releases free amino acids; more gentle hydrolytic processes yield mixtures of amino acids and peptides. The strong conditions of acid or alkali hydrolysis destroy much intrinsic vitamin content of a raw protein, as well as some amino acids such as tryptophan, and reduce the levels of serine and threonine. Acid-hydrolyzed protein hydrolysates are high in salt content formed during the neutralization step; however, some manufacturers feature acid-hydrolyzed peptones with additional processing to reduce the amount of salt [e.g., Casamino Acids (Difco 0230)]. Enzymatic digests contain free amino acids as well as small peptides while maintaining the vitamin content of the original source material. Enzymatic casein digests, reflecting the amino acid composition of the starting material, are high in tryptophan and low in cystine. Peptones from plant proteins [i.e., Soytone (Difco); Phytone (BBL)] contain intrinsic carbohydrate and, while conducive to rapid heterotrophic growth, can cause quick culture decline due to substantial metabolic acid production. Because of the carbohydrate level, plant peptones are also unsuitable in basal media for fermentation or sole carbon source utilization studies. For routine growth, peptones are normally added to media at concentrations of 0.5–1.0%. Growth of microorganisms with diverse and multiple nitrogenous auxotrophies (e.g., the lactobacilli) is enhanced by elevated peptone concentrations as well as nutritional supplementation with peptones of mixed sources and degrees of hydrolysis.

Carbon and Energy Sources

Microorganisms other than autotrophs require organic compounds as a source of carbon and energy. These can include simple sugars, complex carbohydrates, alcohols, amino and other organic acids, and short-chain lipids. Some organisms (e.g., *Aspergillus, Pseudomonas*) possess the complex metabolic machinery that permits numerous classes of organic compounds for carbon/energy substrates. Other taxa exploit specific classes of molecules metabolically ignored by or toxic to general heterotrophs. Methylotrophs (*Methylobacterium* spp., *Methylomonas* spp., *Methylophaga* spp., and others) can extract energy from C_1 compounds such as methanol and formaldehyde, whereas

Table 3. Hydrolysates and Extracts

Hydrolysis method	Source	Product/manufacturer[a]	Characteristics
Acid	Casein	Acidicase (BBL) Casamino acids (Difco) Hy-Case (Sheffield)	Free amino acids; deficient in tryptophan and cystine; high intrinsic salt concentration unless designated salt-free
Enzymatic	Casein	Casitone (Difco) NZ-Amine (Sheffield) Trypticase (BBL)	Free amino acids and small peptides; high levels of tryptophan; low in carbohydrates
Enzymatic	Meat	Myosate (BBL) Thiotone (BBL) Primatone (Sheffield) Peptone (Difco) Proteose peptone (Difco)	Mixtures of amino acids and peptides; those designated proteose are hydrolyzed for a high peptide content; Thiotone high in sulfur
Enzymatic	Plant	Phytone (BBL) Soytone (Difco) Hy-Soy (Sheffield)	High in carbohydrates and vitamins
Mixed	Yeast/casein	Biosate (BBL)	Combines amino acids as well as vitamins lost during protein hydrolysis of peptones
	Casein/meat	Polypeptone (BBL)	
Autolysis	Yeast	Yeast extract (BBL; Difco; Sheffield)	High in amino acids, peptides, water-soluble vitamins, and carbohydrates

[a]Items listed are representative of of a manufacturer's product line. Consult manufacturers' technical manuals for more information, Refs. 39, 40, and 64.

the metabolic diversity of the bacterial genus *Pseudomonas* has allowed selection of strains that can utilize the long-chained hydrocarbons in crude oil or attack the degradation-resistant chemical rings of cresols and phenols as sole carbon/energy sources. Assorted fungal genera (*Cladosporium*, *Candida*, and others) have been found as actively growing contaminants in aircraft and diesel fuels (65). The type and concentration of a carbon substrate present in a medium will affect overall growth rates, morphology, and metabolites of a culture. Sole carbon utilization profiles are used in identification matrices for prokaryotes and yeasts.

Carbohydrates and related compounds are preferential carbon sources for many genera of microbes. Hexoses are most diversely utilized, followed by the disaccharides, trisaccharides, smaller polysaccharides (starch and glycogen), and pentoses. Glycerol, mannitol, and sorbitol are commonly utilized alcohols. Under laboratory conditions with defined media and mixed sugars at normal substrate levels (e.g., 2–10 g/L), microorganisms often exhibit diauxic growth, in which a culture first utilizes the carbon compound that supports the greatest rate of growth, with concurrent catabolic repression of other substrates (66–68). Enzyme systems then shift accordingly as the organism switches to and consumes each less-preferential carbon compound. This results in a biphasic growth curve. However, under carbon-limited (e.g., microgram to milligram per liter concentrations) chemostat conditions that mimic the low-nutrient conditions of the natural environment, a number of organisms use the components of carbon mixtures simultaneously (69).

Proteolytic organisms have the capacity to utilize L-amino acids and related compounds as sole carbon/energy sources. This metabolic characteristic is more common among bacteria than fungi and varies widely between species; identification keys for the genus *Clostridium* use saccharolytic versus proteolytic metabolisms for species differentiation. As a group, the helical/vibroid Gram-negative bacteria (e.g., *Aquaspirillum*, *Azospirillum*, *Camplylobacter*, *Oceanospirillum*) do not ferment or oxidize carbohydrates but satisfy carbon/energy requirements with amino acid catabolism. *Legionella* spp. also use amino acids as carbon/energy sources. Filamentous fungi vary in their ability to utilize amino acids as effective carbon sources; none, however are limited to nitrogenous sources of carbon. Glutamic acid, alanine, arginine, and proline are frequently utilized by many prokaryotes and yeasts. The aromatic rings of the amino acids tryptophan, phenylalanine, and tyrosine, as well as the purine and pyrimidine rings of the nitrogenous bases, are not readily attacked by many organisms. The ability to do so is used for taxonomic speciation. Degradation of any of these ring compounds, however, is not presumptive of an organism's ability to further use the breakdown products as a sole source of carbon/energy; in many cases, the compound is not so metabolized. Organisms that catabolize amino acids frequently utilize organic acids in the same fashion (e.g., *Oceanospirillum*). In addition, many carbohydrate-utilizing organisms are also capable of utilizing organic acids. For example, citrate and propionate utilizations are classical tests in the identification of the carbohydrate-metabolizing genus *Bacillus*. When used as sole carbon and energy sources, amino and organic acids may not be soluble at required substrate levels (2–5 g/L); however, neutralization with KOH or NaOH will form soluble salts of these compounds.

Some organisms have the capacity to utilize fatty acids and other nonpolar molecules as carbon and energy sources. In some instances, the requirement is absolute. The bacterial genus *Leptospira* requires long-chain fatty acids (70), which are supplied in complex growth media by rabbit serum. The anaerobic sulfate-reducing *Desulfovibrio* sp. (DSM 2056) oxidizes the fatty acids butyrate through stearate (71).

Simple sugars and other carbon compounds can undergo chemical changes when autoclaved with other media

components. Amino acids and organic acids, though heat stable at growth factor levels, may degrade at substrate concentrations. The characteristic darkening of autoclaved media containing glucose and peptones, called the Maillard reaction, is caused by an interaction of amino acids with the glycosidic hydroxyl group of the sugar. Darkening of autoclaved media can also be caused by the interaction of mono- and disaccharides, particularly glucose, with phosphates; the reaction is more intense at alkaline pH. Carlsson et al. (72) demonstrated that the glucose–phosphate reaction produces hydrogen peroxide at levels toxic to catalase-negative organisms. These types of adverse chemical reactions between nutrient components compromise the performance of a medium and are to be avoided either by autoclaving or filter sterilizing the reactive compounds separately from the balance of the medium. Not all commonly used carbon substrates can withstand autoclave sterilization. In particular, arabinose, fructose, lactose, pyruvate, ribose, and xylose should be filter sterilized. Light-weight alcohols and similar solvents are volatile at autoclave temperatures. Autoclaving of heat-stable alcohols can be done in tightly stoppered containers; alternatively, these compounds can be sterilized with appropriate solvent-resistant membrane filters. Long-chain hydrocarbons (e.g., oils and waxes) are hydrophobic and are not sterilizable by autoclaving because the sterilant (steam) cannot adequately penetrate the material. A good chemical reference text, such as the *Merck Index* (73), is invaluable for determining the heat stability, solubility, and other important characteristics of many chemical components used in media formulations.

Extracts

Spray-dried powders derived from infusions of meat, yeast, or other chemically complex proteinaceous materials provide a nutritionally useful but undefined source of water-soluble growth factors and sugars. Because they are not subject to the same harsh conditions as those used in the manufacture of peptones, meat or yeast extracts can supply those nutrients lost during the peptone hydrolysis (e.g., vitamins and certain amino acids). Extracts, in themselves, do not contain sufficient concentrations of essential nitrogen or carbon compounds to effectively function as the sole source of these nutrients in a medium. Meat extracts include appreciable amounts of glycolic and lactic acids, as well as creatine. Yeast extract is high in carbohydrates and, therefore, interferes with fermentation studies. The nutritional properties of the fresh yeast extract solution used in many media formulations for the cultivation of mycoplasmas is not replaceable by powdered yeast extract. Malt extract used in some media for fungi is composed primarily of carbohydrates (glucose, fructose, sucrose, maltose, dextrins) and small amounts of nitrogenous matter. Powdered extracts are generally incorporated in media at a concentration of 0.5–1.0%.

Aqueous extracts and infusions of various materials have historically been used to supply undefined nutritional factors in laboratory media. Concentrated liquid extracts generally are used at 10–20% v/v concentration as a medium supplement, whereas infusions will constitute the entire liquid base of a medium to which additional nutrients are added. As already noted, a filtered, aqueous solution prepared from autolyzed baker's yeast (150 g yeast/L) is a necessary ingredient in media for mycoplasmas and other mollicutes. The fresh yeast extract solution is then filter sterilized and incorporated into the mycoplasma media at 10% v/v concentration. Lactic acid bacteria from milk or milk products have been cultivated with infusions prepared from whole milk, whereas tomato juice or other fruit juice extracts have been used at 10–20% v/v in media for lactic acid bacteria isolated from plants. Potato infusions (200–300 g potatoes/L) serve as a nutritional base in many media for the cultivation of bacterial and fungal plant pathogens. Soil extracts prepared from fertile, chemically uncontaminated soils have been used to induce sporulation in many bacteria. Rumen fluid, containing partially digested nutrients from the first compartment of a cow's stomach, is a supplement in media for the cultivation of rumen microorganisms. Rumen fluid contains an undefined mixture of fatty acids, carbohydrates, peptides, vitamins, and metabolites. As rumen fluid is most commonly added to and autoclaved with basal medium components, any heat-labile nutrients are destroyed. Rumen fluid is obtained from fistulated dairy cows and can often be procured gratis from universities with departments of animal or dairy nutrition. Rumen fluid may or may not contain significant amounts of solid material, depending upon the length of time the material has been in the rumen. Centrifugation will effectively remove the bulk of debris; however, a series of filtrations down through 0.2 μm is necessary to remove dead cells of indigenous microflora that can confuse microscopic examinations of the organism being cultured. A defined mixture of fatty acids (Table 4) can effectively replace the need for rumen fluid for many bacteria; however, as with most chemically complex, undefined supplements, cell yield is frequently better when it is included in the medium.

Buffers

pH buffers are molecules that, in solution, resist fluctuations in hydrogen ion concentration. Weak acids, bases, and their corresponding salts in partial dissociation function as pH buffers. Buffering capacity is greatest when the

Table 4. Volatile Fatty Mix for Rumen Bacteria and Fungi

Acetic acid	17.0 mL
Propionic acid	6.0 mL
N-Butyric acid	4.0 mL
N-Valeric acid	1.0 mL
Isovaleric acid	1.0 mL
Isobutyric acid	1.0 mL
DL-α-Methylbutyric acid	1.0 mL

Pipette the given volumes of fatty acids into a container that has a tight-fitting closure and mix well. The objectionable odor of the free fatty acids can be reduced by neutralizing the pH of this solution with the addition of NaOH pellets. Use 3.1 mL of the fatty acid mix per liter of medium.

Source: From Ref. 44, p. 162.

compound is at 50% dissociation. The pH at which this occurs is equal to the pK_a value of the compound, and greatest buffering capacity falls between ± 1 pH unit of the pK_a. Thus, no one buffer can be used across the entire pH range of 1 to 14 (74). Some acids are zwitterions (compounds that can chemically behave as acids or bases) and have more than one pK_a. In an extreme example, glycine when titrated against HCl has pK_a 2.3; titrated against NaOH, pK_a 9.6. Other zwitterions include citric acid (pK_{a1} 3.06, pK_{a2} 4.74, pK_{a3} 5.40), succinic acid (pK_{a1} 4.19, pK_{a2} 5.57), and the Good buffers (see later). The perfect media buffer would be (*1*) completely effective at poising a desired pH throughout the metabolic flux of a culture's growth cycle, as well as any other chemical or environmental changes that may affect pH (e.g., metal ion interaction, temperature), (*2*) resistant to microbial enzymatic activity, (*3*) chemically inert with respect to other medium components, and (*4*) inexpensive. The perfect buffer does not exist.

Lacking perfect buffers, in vitro microbial cultivation relies, for better or worse, upon a standard array of buffering systems. Nevertheless, the choice of buffer must not be made haphazardly. No discussion of buffers could be complete without N. E. Good's comment: "It is impossible even to guess how many exploratory experiments have failed, how many reaction rates have been depressed, and how many processes have been distorted because of imperfections of the buffers employed" (75). The capacity to buffer well at a particular pH is only one of many criteria in the determination of an appropriate buffer for a particular application. Chemical interaction with other media components or culture metabolites will compromise both the buffering capacity as well as the nutritional balance of a medium. The effectiveness of a microbially metabolized buffer diminishes with culture growth. The pH of some buffers can vary, sometimes considerably, with concentration or temperature. Synthetic buffers, while very effective for bench-top laboratory needs, may become prohibitively expensive upon scale-up.

Natural Buffering Agents. Natural buffering agents are those compounds most frequently encountered in microbiological media; they are also the buffers subject to the greatest drawbacks in their use. Carboxylic acids (e.g., acetate, citrate, succinate) are effective buffers for the lower ranges (<pH 6) of biological activity. These compounds are metabolically catabolized for carbon and energy by many microorganisms. Also, at the concentrations needed for buffering, carboxylic acids can become cation chelators. The mixtures of amino acids in protein hydrolysates and extracts have significant buffering capacity. Thus, peptone-based complex media frequently do not require additional buffering agents when used at or about the initial pH of the medium. Many peptones are adjusted during manufacture to pH 6.8–7.0. Peptone solutions are stable to sterilization at 121 °C; however, they adversely interact at elevated temperatures with other medium components such as glucose, phosphates, and polyvalent cations. In addition, because peptones are nutrients, buffering capacity diminishes or shifts as the amino acids are microbially consumed.

Sodium bicarbonate–carbon dioxide buffers can maintain pH in the range 6.0–8.0. However, when used in conjunction with intrinsically buffered media, bicarbonate buffers have a working range of pH 6.8–7.2. Despite the necessity for closed systems to retain the gaseous CO_2 phase, bicarbonate buffers are widely used in the cultivation of anaerobic bacteria and eukaryotic animal cell lines. Bicarbonate buffers are nutritionally consumed. Also, carbon dioxide has minimal solubility in aqueous solutions, which somewhat limits its buffering capacity. Solid calcium carbonate (powdered) is used in various media for strongly acidifying organisms such as *Acetobacter, Clostridium*, and *Lactobacillus*. Ethanol-tolerant, acid-producing acetobactors can be isolated and differentiated from mixed-culture fermentations using a carbonate-ethanol medium (yeast extract, 3%; calcium carbonate, 2%; agar, 2%; filter-sterilized ethanol, 2%). Strongly acidifying acetobacter isolates able to grow in the presence of ethanol will also dissolve the carbonate and form a zone of clearing around the colony.

Although widely used in media formulations at pH 6–8, phosphate buffers suffer from a great number of drawbacks. Unless a medium is carefully designed and prepared with respect to metal speciation and concentration, phosphates will adversely bind to and precipitate polyvalent cations and proteinaceous material at neutral to alkaline pH. This chemistry creates two problems: (*1*) buffering capacity is diminished as cations complex with the phosphate anion, and (*2*) cations bound as precipitates are unavailable as nutrients. These unwanted chemical interactions are intensified by heat. The problem can be reduced by autoclaving the phosphate buffer separately from the balance of the medium and then combining both solutions when cooled to room temperature after autoclaving. The addition of a chelator such as EDTA may retard trace metal precipitation; however, overchelation will also restrict the availability of trace metals from the nutrient pool. Above pH 7.2, with poorly designed media formulations, precipitation can become spontaneous even at room temperature; the reaction may be immediate or may slowly develop over time and is often misinterpreted as bacterial contamination. Balanced phosphate-buffered salines (e.g., Dulbecco's, Hank's) containing calcium and magnesium salts are filter sterilized to avoid precipitation of the salts. When autoclaved together, phosphate and glucose will also adversely react at neutral pH and above; trace metals will intensify the reaction. Because soluble phosphates are nutrients, the effectiveness of these buffers can diminish with culture metabolism. Finally, phosphate buffers exhibit a concentration effect characterized by a rise of 0.2–0.3 pH unit with each 10-fold dilution.

Synthetic Buffers. Synthetic buffers were developed to alleviate the many problems encountered with the earlier generations of natural buffers. Tris [Tris(hydroymethyl)-aminomethane; pK_a 8.3 at 20 °C] was the first organic compound widely utilized for its improved buffer characteristics. Tris and Tris HCl can be combined for buffering between pH 7.0 and 9.0, with strongest activity between 7.2 and 8.0. Tris can also be paired with other acids such as acetate, borate, or phosphate for specific buffering ap-

plications. Tris does not precipitate calcium ions and helps to maintain the solubility of magnesium salts. Tris is, however, a chemically reactive amine and can penetrate cell membranes; some microorganisms (e.g., *Sporocytophaga myxococcoides* ATCC 10010) have shown growth inhibition by the compound. Tris exhibits marked concentration and temperature effects that can limit its effectiveness for certain biological purposes. Tris buffers fall approximately 0.1 pH unit per 10-fold dilution. At 0.05 M, a Tris buffer at 25 °C will be pH 7.20; at 5 °C, pH 7.76; at 37 °C, pH 6.91. Therefore, when used in culture media, Tris buffers should be constructed with the incubation temperature of the culture in mind and not upon the pH of the buffered medium at room temperature.

During the 1960s Good et al. (75) developed a number of zwitterionic compounds for use in biological systems to include as many as possible of the following desired buffer characteristics: (*1*) freely soluble in aqueous solutions to permit preparations of concentrated stock solutions, (*2*) buffering optima within the range of most biological systems, (*3*) biochemically nonreactive and chemically stable, (*4*) inert to microbial or other enzymatic attack, (*5*) minimal temperature and ionic effects upon buffer pH, and to further enhance their versatility to include postculture analytical procedures, (*6*) insoluble in organic solvents and (*7*) transparent to the ultraviolet wavelengths used in many biological spectrophotometric assays. Collectively these compounds have come to be known as the Good buffers, of which, HEPES (pK_a 7.5 at 25 °C) is most commonly recognized due to its prevalence in media formulations for the cultivation of animal cell lines. Other Good buffers useful for microbial media formulations include ACES (pK_a 6.8 at 25 °C), which is a component in the recommended cultivation medium for *Legionella*; buffered charcoal yeast extract (BCYE) (2p.518); MES (pK_a 6.1 at 25 °C), PIPES (Pk_a 6.8 at 25 °C), and TES (pK_a 7.5 at 25 °C).

HEPES has been implicated in the phototoxicity of RPMI medium for the cultivation of eukaryotic cells; however the effect was eliminated when sodium pyruvate was added to the culture medium (76). Other Good buffers have also shown various types of inhibition in certain cell lines (77). As with any buffering system or other media component, the Good buffers should be evaluated for applicability against the microorganism under study.

Biological buffers for cell growth purposes are commonly used at 0.01–0.05 M; chemical effectiveness of the buffer, initial pH of the medium without the buffer, production of acid or basic metabolites during culture growth, and enzymatic or other reagent optima are considerations in determining appropriate buffer ion concentration. Buffer tables can be found in numerous publications. Lacking a table, buffer solutions can be prepared with either of the two following methods and a calibrated pH meter:

1. Prepare a 2× solution of the buffer compound. Using HCl or NaOH, titrate the solution to the desired pH. Bring the solution to 1× volume with deionized or distilled water.
2. Prepare equimolar solutions of the acidic member and the basic member of the buffer pair. The molarity of these solutions is that of the final concentration needed for the working buffer. Choosing the solution of the buffer pair with a pH closest to that needed for the working buffer, titrate the solution to the desired pH with the equimolar solution prepared with the remaining component of the buffer pair.

Gelling Agents

Many microbial applications, such as strain isolation and purification, necessitate the use of solid media. Growth of aerobic organisms, in addition to specific morphologies, are often enhanced on solid media providing unrestricted gas exchange. The different gelling agents available to microbiologists each have attributes and drawbacks that one must consider before choosing a particular gelling agent to include in a medium formulation. Although agar is the most frequently encountered gelling agent in microbial culture media, it must be properly used in order to maintain the chemical and nutritional integrity of a formulation. Many chemolithotrophic organisms are inhibited by organics and will not grow on media solidified with agar. Conversely, a number of marine bacteria are agarolytic and can quickly dissolve or soften this support matrix. Other gelling agents, including carrageenan and Gelrite, are alternatives to agar, but each has its own chemical peculiarities that must be understood before they can be effectively used.

Agar. Agar is a polysaccharide complex extracted from a number of species of red algae (*Gelidium*, *Gracilaria*, and others) found in the Pacific and Indian Oceans and the Japan Sea. The agar complex can be separated into (*1*) agarose, a neutral gelling fraction, and (*2*) agaropectin, a sulfated, nongelling fraction. Depending upon the purity and gel strength of the agar, and also upon the chemical composition of the medium in which the agar is used, a molten 1.5% agar solution will set at temperatures between 32 and 39 °C. However, agar media with high concentrations of salt (e.g., those for the extreme halophiles, which contain >18% NaCl and 0.5% each of $MgSO_4$ and KCl) gel quickly at temperatures above 55 °C. The gelling capacity of agar is thermally reversible, however, repeated remelting of agar-containing media is not recommended because the chemical stability of many nutrients cannot withstand repeated exposure to elevated temperatures. Agar is subject to acid hydrolysis, which destroys the gelling capacity of the polysaccharide; the phenomenon is intensified by heat. Agar-containing media autoclaved at less than pH 5.0 will not gel when cooled. If a low-pH agar is required, it is advisable to prepare and sterilize the medium in two separate solutions: For 1 L of low-pH agar medium, (*1*) heat and dissolve agar, with stirring, in 500 mL water; (*2*) dissolve remaining medium components in 500 mL water; (*3*) sterilize each solution separately at 121 °C; (*4*) cool each solution to approximately 50 °C; (*5*) aseptically combine both solutions, mix well, and dispense into containers as required.

Agar products vary with respect to nitrogenous and inorganic salt contaminants. A number of manufacturers feature agars purified though repeated water and solvent washes. Some agar products contain significant amounts

of metals that can bind to and precipitate with phosphates in the medium, resulting in diminished growth yield with the solid medium as compared with that of the corresponding liquid medium. Similarly, autoclaving any agar in the same solution with the carefully balanced nutrient salts of a chemically defined medium, particularly at neutral to alkaline pH, will frequently give rise to a pronounced darkening or haze that can compromise the effectiveness of the medium. Under chemically defined conditions, component integrity is best preserved by autoclaving a purified agar separately from the autoclaved or filter-sterilized chemical portion of the medium, as already outlined for preparation of a low-pH agar medium. The use of electrophoretic grades of agarose would be more suitable than purified agars for gels in which a minimal level of trace metal contamination is required.

As with all gels, agar exhibits syneresis in that it will squeeze out solute as it contracts. The small puddle of liquid at the bottom of a freshly prepared agar slant is an example of this phenomenon. The amount of syneresis is dependent upon the concentration of the agar in the gel. Agar media can also lose water through desiccation during storage. Once water has been eliminated from the agar gel, it will not reenter the matrix unless the gel is remelted; however, melting of solid media for the purpose of rehydration not recommended because performance of the medium will be markedly reduced.

Agar is normally used at concentration from 1.0 to 2.0% for solid media; at 0.075 to 0.2% to create a viscosity that retards oxygen diffusion in media such as fluid thioglycollate medium U.S.P. and the classical oxidation–fermentation (O-F) medium of Hugh and Leifson (78); and at 0.5 to 0.75% for soft, top-layer agars in media for bacteriophage propagation.

Agarose. Agarose, the purified gelling fraction of agar, is used in some formulations for organisms that are inhibited by agar. Purified agarose products, particularly those grades used for electrophoretic work, are low in inorganic and organic contaminant and, therefore, react less with the phosphates present in a medium. Gel clarity is greater than for agar. Low gelling temperature agaroses melt above 65 °C and form gels at less than 30 °C. Agarose will form solid gels at 0.75–1.0%.

Gellan Gum. Gelrite, a gellan gum, is a heteropolysaccharide composed of glucuronic acid, rhamnose, and glucose, produced by bacterial fermentation, and commercialized as a gelling agent by Kelco (Merck & Co., Inc., Kelco Division, San Diego, Calif.). Sigma Chemical Co. (St. Louis, Mo.) distributes the material under the trade name Phytagel. Under either name, this gellan gum is somewhat difficult to work with for routine media solidification purposes, but nevertheless, its unique characteristics make the material the most practical solution to some unusual microbial cultivation requirements. Gelrite/Phytagel is a high-melting-temperature (~70 °C) gel, which makes it useful for the cultivation of thermophilic bacteria and fungi at 45–50 °C, temperatures at which agar shows considerable syneresis and softening. The gel is generally thermally reversible, although under some conditions, such as the elevated soluble salt concentrations of marine media, gelling can become permanent. Gelrite/Phytagel is subject to acid hydrolysis when autoclaved at low pH. The gellan is resistant to microbial attack by most organisms and so proves useful in the isolation of organisms such as the marine cytophagas, a group in which many strains characteristically possess the ability to degrade other biopolymer gels including agar, algin, carrageenan, and gelatin. Many chemolithotrophic organisms (e.g., *Nitrobacter*, spp. and *Thiobacillus ferroxidans*), although inhibited by agar and other organic gelling agents, can be grown successfully on a solid medium using the gellan. Full gelling capacity of this polysaccharide is cation dependent, and media formulations may need amending to include at least 0.1% $MgSO_4 \cdot 7H_2O$, or other magnesium salt to permit sufficient gelling. Under chemically defined media conditions, further chemical adjustments to other components may be necessary to retain chemical balance because of a modification in cation level. Recommended concentrations for Gelrite/Phytagel are in the 0.1–0.2% range, however experimentation will probably be needed to select the appropriate concentration for one's own needs. Table 5 provides a media formulation that has been successfully used as a solid medium for the cultivation of laboratory strains of *T. ferroxidans* (ATCC 21834 and others); the gel is rigid enough for the isolation and picking of colonies, and growth

Table 5. Solid Medium for *Thiobacillus ferroxidans*

Solution A	
$(NH_4)_2SO_4$	0.08 g
$MgSO_4 \cdot 7H_2O$	0.20 g
K_2HPO_4	0.04 g
Trace elements (see Table 1, mix A)	0.5 mL
Distilled water	40.0 mL
Adjust Solution A to pH 2.3 with concentrated H_2SO_4. Filter sterilize.	
Solution B	
Phytagel (Sigma P-8169)	0.5 g
Distilled water	40.0 mL
With constant stirring, heat Phytagel to boiling to dissolve. Sterilize at 121 °C for 15 min. Cool to no less than 80 °C.	
Solution C	
$FeSO_4 \cdot 7H_2O$	2.0 g
Distilled water	20.0 mL
Quickly dissolve ferrous sulfate with constant stirring. Immediately filter sterilize.	
Complete medium	
Warm solution A to about 50 °C. Add solution C to solution A and immediately combine with solution B. Mix well and pour plates. As this medium must be poured at a rather high temperature to avoid premature gelling of the Phytagel, it will be necessary to dry the plates before use by leaving the lids ajar in a sterile environment or by incubating them at 37 °C overnight to eliminate excess condensation.	

is rapid (3–5 days) at 20–30 °C. Preparation methods for this particular formulation illustrate many of the topics discussed in this article with respect to treatment of chemically incompatible salts, acid hydrolysis, and heat instability of media components.

Alginic Acid and κ-Carrageenan. Alginic acid (79) and κ-carrageenan (80,81) have been used microbiologically for specialized gelling needs. Gelling efficiency of alginate is Ca^{2+} dependent, whereas κ-carrageenan is K^+ or Na^+ dependent. Since all of these substances are easily degraded by assorted bacteria and fungi, their use as gelling agents is restricted. Microbial degradation of these substances is used for taxonomic purposes.

Silica Gel. A final solid substrate for the laboratory cultivation of obligate chemolithotrophic organisms is silica gel, first used by Weintraub and Hanks (82) in 1936 for the isolation of ammonia-oxidizing bacteria. Many authors have since attempted improvements on the tedious method (83–86). Preparation of any quantity of plates is laborious. Briefly, sodium silicate is reacted with strong HCl to produce silicic acid, NaCl and almost immediate gelling. Nutrient salts may be included in the formulation, or the surface of the silica gel may be later flooded with nutrients, with some absorption. Because the amount of NaCl is high (~8–10%), the gel may first need clearing of the salt, with repeated washings with water prior to flooding with a nutrient solution. Silica gels are translucent and rigid with a high degree of syneresis. The gels are autoclavable, with a tendency for cracking, and are not reversible. Table 6 provides simplified instructions for silica gel preparation adapted from Ref. 82.

Table 6. Simplified Silica Gel Plates

Stock solutions

Solution 1	
Concentrated HCl (~37%)	14.5 mL
Distilled water	to 100.0 mL
Solution 2	
$Na_2SiO_3 \cdot 9H_2O$	20.0 g
Distilled water	100.0 mL

Use slight heat and constant stirring to dissolve sodium silicate. Filter out any particulates or colloidal material with Whatman #1 paper.

Preparation of silica gel plate

1. Volumes given are for one 20 × 100 mm petri dish. Gelling will occur within 10–15 s of mixing solutions. Be prepared to work quickly.
2. Add 14.5 mL solution 1 to a 100-mL beaker and stir continuously. (A magnetic stirrer will work best.)
3. Quickly add 19.0 mL solution 2 into the beaker and mix well.
4. As soon as effervescence stops, quickly pour contents of beaker into a 20 × 100 mm petri dish. Use a glass dish if subsequent autoclaving is required.
5. Initial gelling will occur almost immediately. Gel strength will increase slightly over a half-hour period, at which time the gel may be washed to remove excess NaCl and then flooded with nutrient solution.

MEDIA STERILIZATION

With few exceptions, media and buffers used for microbial cultivation must be sterilized as soon as possible after preparation to limit proliferation of environmental contaminants that will alter the nutritional characteristics of the medium either by consumption of nutrients or by production of metabolites. Even nonnutritive buffers or trace metal stock solutions can give rise to unwanted fungal growth if stored unsterilized for any length of time at room or refrigerator temperatures. Sterile media and buffers are necessary for culture propagation; the presence of microbial contaminants, even at a low level, can invalidate any study or production process. The physical stress of sterilization, however, can compromise the nutritional and physiological characteristics of a medium or buffer. Consequently, one must choose a sterilization method that has the least negative impact upon the material being sterilized. Although steam sterilization had historically been the only form of sterilization accessible by many laboratories, the widespread availability and diversity of reasonably priced and reliable disposable filtration units offer more appropriate choices for media sterilization, even for small laboratories.

Steam Sterilization (Autoclaving)

Steam remains the most widely used sterilization method for rugged, heat-stable media and has the advantage of often being the terminal step in media preparation. With no additional manipulation of an uninoculated medium poststerilization, possibility of inadvertent introduction of contamination is reduced. The term *autoclaving* is often used synonymously for the definitional procedure: "sterilize at 121 °C for 15 minutes." This phrase *must* be applied in context to the volume of the medium being sterilized as well as to the efficiency of the autoclave being used. The precise meaning of the phrase infers that the contents of each container within an autoclave cycle must achieve an uninterrupted exposure to 121 °C for 15 min; it does not mean that the autoclave itself was set for that temperature and time. From a practical laboratory standpoint, this means that small volumes (e.g., test tubes) will require the minimal amount of autoclaving time, whereas large volumes (e.g., one or two flasks) will require additional autoclaving time. Fully loaded autoclaves may require longer run times. Under laboratory autoclave conditions, a 30–35 min sterilization cycle for 0.5 L liquid medium in a 1-L flask is a practical guideline; agar media that have been allowed to solidify before sterilization may require longer autoclave times. Large-scale fermentation volumes can require hours to sterilize; media processed in poorly maintained autoclaves might never be adequately sterilized. Where mechanical agitation of a medium is not possible during the autoclave cycle, maximum container volumes of media should not exceed the ratio of >1 L media in a 2-L vessel because the prolonged autoclave times required for heat conduction to raise the innermost portions of large volumes of media to 121 °C will subject the outer portions of the media to extreme overheating.

Autoclaves and their cycles should be validated for media sterilization. Ideally, this would include replicate au-

toclave run data from multiple thermocouple probes spaced strategically within the autoclave chamber and load configurations. While this technological expense is beyond the budget of many small research laboratories, quantified preparations of *Bacillus stearothermophilus* sterilization indicators, such as the Verify Biological Indicators (Steris Corporation, Mentor, Ohio), can indicate general "121 °C, 15 min" parameters. Chemical indicator strips or labels are not reliable indicators of sterilization and should be used only as visual aids to indicate that a particular item has been run through an autoclave cycle. Under the most elemental quality control conditions, media samples should be held at least 2 weeks at 26 °C and 37 °C, and also at conditions and temperatures conducive to the growth of the types of environmental bioburden normally encountered during media preparation and sterilization, to ensure that latent contamination does not subsequently proliferate. Unusual media formulations (e.g., extremes in pH, chemolithotrophic conditions, unusual energy sources) that may not allow proliferation of common microorganisms may require inoculation of lot samples into more conventional heterotrophic media (e.g., tryptic soy broth) to encourage the growth of common environmental contaminants.

Steam at pressures below 15 psi (121 °C) are used for sterilization of many media components not stable at 121 °C. Skim milk (used as a cryoprotectant in lyophilization and as a substrate in phenotypic characterization) is best sterilized at 116 °C for 20 min; the lower pressure reduces carmelization of indigenous lactose. Powdered sulfur (S^0) used as an energy source for microorganisms can be treated with Tyndallization (100 °C for 30 min for 3 consecutive days) to avoid melting the element into a solid mass that physically limits its availability to organisms. Tyndallization is also used as a method of bioburden reduction (it is *not* a sterilization process) for other extremely heat-labile media. In theory, vegetative cells are killed by the first day's boiling; quickly germinating spores are killed the second day; by the third day, late-germinating spores are killed, leaving a relatively clean preparation.

Filter Sterilization

Heat-labile media components should be filter sterilized. Any chemically balanced liquid buffer or medium can also be filter sterilized to ensure that chemical integrity will not be compromised by the elevated heat and pressure of steam sterilization. Membrane filters, rather than depth filters, are preferred for sterilization purposes because their uniform pore size permits absolute retention of particles down to a specified pore diameter when the membranes are stored, sterilized, and used according to manufacturer's instruction. Depending upon the polymer composition, membrane filters are available in porosities from 10 μm down through nominal molecular weight limit used for ultrafiltration. For general microbiological purposes, 0.2-μm filtration is used for sterilization of media and buffers because this porosity effectively traps many vegetative bacteria and spores. However, in serum-based media, because mycoplasmas are not always held back at 0.2 μm, filtration through a stack of three 0.2-μm membranes rather than a single membrane will increase the probability of retention of these common serum contaminants. In addition, membranes are now available in 0.1-μm porosity that can retain smaller-sized microbial contamination, but flow rates are slower and rate of membrane clogging increases as porosity decreases. In any event, the filtration method selected for sterilization should be challenged against the smallest microbial contaminants known to be present in the working environment and production stream (87). As pore size decreases to submicron diameters, there is the possibility of holding back large molecular weight nutritional components and aggregates.

Water-soluble hydrophilic microbial media, stock solutions, and buffers are most commonly filter sterilized with membranes composed of cellulose acetate or nitrate esters. Fastidious cultures may require initial washing of the membranes with hot water to remove residual organics prior to media filtration; filters and membranes specified for tissue-culture applications are also recommended for sensitive organisms. Nonaqueous media components such as ethanolic concentrates of antibiotics or dimethyl sulfoxide (DMSO) must be filter sterilized with hydrophobic polymer membranes designed to withstand organic solvents. Nylon and Teflon membranes are suitable for many organic-solvent based biological applications. These types of membranes can require initial washing with a weak nonpolar solvent such as methanol to condition them for filtration. Some membranes, such as Nylon, are protein binding; if specific proteins are key media components, choose a membrane specifically identified as low protein binding.

Membrane filters have limited loading capacity and will quickly clog. Consequently, effective membrane filtration with any large volume, especially with particulate-laden preparations, requires sequential filtration from larger through smaller porosities to clarify the medium of debris, as much as possible, prior to the final filtration step. Low-speed centrifugation may also be used to first clarify crude media preparations. Glass fiber depth filters are efficient prefilters for many media applications because they have great loading capacity; however, they do not always eliminate the need for sequential membrane clarification steps prior to sterilization. Sterile filtration can be a slow process. It may be necessary, for critical applications, to perform the entire operation in a cold room to retard environmental microbial or other denaturing activity upon the raw material.

Positive-pressure filtration with an inert gas such as nitrogen is used in most large-scale filtration applications. Positive pressure reduces foaming of the filtrate and possible denaturation of any essential protein components. For critical and regulatory applications, the gas supply must also be sterile filtered before it reaches the medium reservoir. In many laboratories, filter sterilization with vacuum is the only available option. With vacuum, foaming of proteinaceous material will be best reduced with sequential prefiltration through larger membrane porosities as well as with minimal vacuum applied to the process. For most bacterial applications, serum foaming is not a deleterious nutritional problem. However, do note that me-

dia foaming can be a serious source of environmental contamination if the foam is not trapped and collected before passing into the laboratory vacuum line.

Disposable, presterilized filtration systems are offered by a number of manufacturers (e.g., Falcon, Corning, Gelman, Millipore, Nalge) in small- to large-volume capacities. Capsule filters are available for small to medium inline batch processing (20–60 L). Filter/storage systems in the 200-mL to 1-L sizes offer the ease of vacuum-operated membrane sterilization and storage with minimal preparation and clean up. Syringe filters are most used for volumes of less than 50 mL; however some types can sterilize volumes up to 1 L without clogging.

Dry Heat Sterilization

Some materials with microbiological applications require dry heat sterilization. Whereas autoclaving is suitable only for heat-stable aqueous material, and filter sterilization is amenable to those materials capable of passage through membranes, dry heat is used for sterilization of nonaqueous, heat-stable material such as oils and waxes. It is also used for glassware or equipment depyrogenation and enzyme (RNase) elimination in many production as well as research protocols. A minimal standard for dry heat sterilization is 350 °C for 30 min, or 121 °C for 6 h (88). As with autoclaving, the time can vary considerably with the efficiency of the sterilization oven as well as with the volume and kind of material being sterilized. Sterilization oven validation is required for many production process. However, for basic laboratory needs, biological dry-heat spore indicators (*Bacillus stearothermophilus* and *B. subtilis*) such as Spore-O-Chex (PyMaH 00190) will give a better indication of actual oven operation than will chemical indicator strips that show only that a particular temperature has been reached, but not for how long, nor if any temperature interruptions occurred. Oils and waxes are best heat sterilized in shallow volumes in large containers to permit rapid and thorough heat penetration of the material because temperature lag is much longer with nonaqueous than with aqueous materials.

MEDIA STORAGE

Prepared microbial media should stored refrigerated at 4–8 °C and in the dark. Containers should be tightly sealed or, in the case of petri dishes, packaged in plastic bags to retard evaporation. Desiccated agar media are characterized by surface cracking and pulling of the agar away from container walls. Water evaporation from liquid or solid media concentrates solutes, which can result in permanent precipitation of some components, alters concentrations of critical nutritional factors or substrates, and increases osmolarity to possibly growth-limiting levels. The ability to freeze media is entirely dependent upon the formulation. As water freezes and concentrates the solutes, poorly soluble chemicals, such as tyrosine, or chemicals, such as phosphates, prone to complexing may not go back into solution when the medium is subsequently thawed. Phototoxicity of media can develop when light activates the glass or plastic containers, as well as media components, and initiates formation of free radicals; aerobic conditions intensify these deleterious reactions as do the blue wavelengths of light emitted by fluorescent tubes (89). The effects of light and free radical formation with respect to tissue culture media have been well studied. In one such mechanism, light at 360 nm and 450 nm activates riboflavin, resulting in a transfer of energy to tryptophan or tyrosine and the formation of cytotoxic free-radical products (90). Similar hydrogen peroxide formation in RPMI medium has been shown to be intensified by the presence of HEPES buffer. With bacteria, an inhibitory effect of light on growth-supporting properties of eosin methylene blue agar has been demonstrated (91). It is recommended that tissue culture media and media for those microorganisms sensitive to the presence of hydrogen peroxide or free radicals (e.g., anaerobic bacteria) be stored and cultured in the dark or in yellow wrappers to reduce the inhibitory effects of light on the media and growth of the organisms.

BIBLIOGRAPHY

1. R.A. Sanford, J.W. Urbance, and J.M. Tiedje, in *Abstracts of the Annual Meeting of the American Society for Microbiology, 1996*, American Society for Microbiology, Washington, D.C., 1996, abstract O-66, p. 365.
2. P. Pienta, J. Tang, and R. Cote eds., *American Type Culture Collection Catalogue of Bacteria and Bacteriophages*, 19th ed., American Type Culture Collection, Rockville, Md., 1996.
3. S.C. Jong, F. Dugan, and M.J. Edwards eds., *American Type Culture Collection Catalogue of Filamentous Fungi*, 19th ed., American Type Culture Collection, Rockville, Md., 1996.
4. *DSM Catalogue of Strains 1993*, Deutsche Samlung von Microorganismen und Zellkulturen GmbH, Braunscheig, Germany, 1993.
5. E.Y. Bridson, in M. Rechcigl, Jr. ed., *CRC Handbook Series in Nutrition and Food, Section G, Diets, Culture Media, Food Supplements*, vol. III, CRC Press, Cleveland, Ohio, 1978, pp. 91–281.
6. R.M. Atlas, *Handbook of Microbiological Media*, CRC Press, Boca Raton, Fla., 1993.
7. P. Gerhardt, R.G.E. Murray, W.A. Wood, and N.R. Krieg eds., *Methods for General and Molecular Bacteriology*, American Society for Microbiology, Washington, D.C., 1994.
8. *C3-A2 Preparation and Testing of Reagent Water in the Clinical Laboratory*, 2nd ed., National Committee on Clinical Laboratory Standards, Villanova, Pa., 1991.
9. *The U.S. Pharmacopoeia XXIII / The National Formulary XVIII*, U.S. Pharmacopoeial Convention, Rockville, Md., 1995, pp. 1635–1636.
10. P. Schönheit and T. Schäfer, *World J. Microbiol. Biotechnol.* **11**, 26–57 (1995).
11. C.J. Microcha and J.E. Devay, *Can. J. Microbiol.* **17**, 1373–1378 (1971).
12. A. Willems, J. Busse, M. Goor, B. Pot, E. Falsen, E. Jantzen, B. Hoste, M. Gillis, K. Kersters, G. Auling, and J. De Ley, *Int. J. Syst. Bacteriol.* **39**, 319–333 (1989).
13. W. Crueger and A. Crueger, Thomas D. Brock ed., *Biotechnology: A Textbook of Industrial Microbiology*, English edition, Sinauer Associates, Sunderland, Mass., 1982, pp. 73–80.

14. R.M. Smibert and N.R. Krieg, in P. Gerhardt, R.G.E. Murray, W.A. Wood, and N.R. Krieg eds., *Methods for General and Molecular Bacteriology*, American Society for Microbiology, Washington, D.C., 1994, pp. 607–654.
15. S.M. Hasan and J.B. Hall, *Can. J. Microbiol.* **87**, 120–128 (1975).
16. S.H. Hutner, L. Provasoli, A. Shatz, and C.P. Haskins, *Proc. Am. Philos. Soc.* **94**, 152–170 (1950).
17. M.O. Garraway and R.C. Evans, *Fungal Nutrition and Physiology*, Wiley, New York, 1984, pp. 124–170.
18. M.N. Hughes and R.K. Poole, *J. Gen. Microbiol.* **137**, 725–734 (1991).
19. P. Kaur and D.V. Vadehra, *Antimicrob. Agents Chemother.* **29**, 165–167 (1986).
20. C. Lindblow-Kull, A. Schrift, and R.L. Gherna, *Appl. Environ. Microbiol.* **44**, 737–743 (1982).
21. H.S. Rosenbranz, J.E. Coward, T.J. Wjodkowski, and H.S. Carr, *Antimicrob. Agents Chemother.* **5**, 199–201 (1974).
22. R.C. Tilton and B. Rosenberg, *Appl. Environ. Microbiol.* **35**, 1116–1120 (1978).
23. R.J. Cote and R.L. Gherna, in P. Gerhardt, R.G.E. Murray, W.A. Wood and N.R. Krieg eds., *Methods for General and Molecular Bacteriology*, American Society for Microbiology, Washington, D.C., 1994.
24. J.H. Crosa, *Annu. Rev. Microbiol.* **38**, 69–89 (1984).
25. E. Griffiths, in J.J. Bullen and E. Griffiths eds., *Iron and Infection, Molecular and Clinical Aspects*, Wiley, Chichester, U.K., 1987, pp. 69–137.
26. M.A.F. Jalal, R. Mocharla, C.L. Barnes, M.B. Hossain, D.R. Powell, D.L. Eng-Wimot, S.L. Grayson, B.A. Benson, and D. Van Der Helm, *J. Bacteriol.* **158**, 683–688 (1984).
27. R.D. Perry and C.L. San Clemente, *J. Bacteriol.* **140**, 129–1132 (1979).
28. M.-F. Thorel, M. Krichevsky, and V.V. Levy-Frebault, *Int. J. Syst. Bacteriol.* **40**, 254–260 (1990).
29. T.E. Bramanti and S.C. Holt, *J. Bacteriol.* **173**, 7330–7339 (1991).
30. B. Fouz, R. Mazoy, M.L. Lemos, M.J. DelOlmo and C. Amaro, *Appl. Environ. Microbiol.* **62**, 2806–2810 (1996).
31. D.R. Lovley, *Annu. Rev. Microbiol.* **47**, 263–290 (1993).
32. M. DeFelice, M. Levinthal, M. Iaccarino, and J. Guardiola, *Microbiol. Rev.* **43**, 42–58 (1979).
33. M.K. Winston and J.K. Bhattacharjee, *J. Bacteriol.* **152**, 874–879 (1982).
34. L.A. Collins and R.L. Thune, *Appl. Environ. Microbiol.* **62**, 848–852 (1996).
35. M.H. Saier, Jr., M.J. Fagan, C. Hoischen, and J. Reizer, in A.L. Sonenshein, J.A. Hoch, and R. Losick eds., Bacillus subtilis *and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 133–156.
36. P.F. Smith, T.A. Langworthy, and M.R. Smith, *J. Bacteriol.* **124**, 884–892 (1975).
37. C. Wyss, B.K. Choi, P. Schüpbach, B. Guggenheim, and U.B. Göbel, *Int. J. Syst. Bacteriol.* **46**, 745–752 (1996).
38. T.A. Nerad, ed., *ATCC Catalogue of Protists*, 17th ed., 1991, p. 60.
39. D.A. Power ed., and P.J. McCuen, *Manual of BBL Products and Laboratory Procedures*, 6th ed., Becton Dickinson Microbiology Systems, Cockeyville, Md., 1988, pp. 178–179.
40. *Difco Manual*, 10th ed., Difco Laboratories, Detroit, Mich., 1984, pp. 567–568.
41. F. Fischer, 97/534/EC, *Journal of the European Communities*, Brussels, Belgium, 1997.
42. B.D. Davis and E.S. Mingioli, *J. Bacteriol.* **60**, 17–27 (1950).
43. S.A. Koser, in *Vitamin Requirements of Bacteria and Yeasts*, C. Thomas Publisher, Springfield, Ill. 1968, pp. 379–401.
44. R.J. Cote and R.L. Gherna, in P. Gerhardt, R.G.E. Murray, W.A. Wood, and N.R. Krieg eds., *Methods for General and Molecular Microbiology*. American Society for Microbiology, Washington, D.C., 1994, pp. 155–178.
45. M.O. Garraway and R.C. Evans, *Fungal Nutrition and Physiology*, Wiley, New York, 1984, pp. 171–211.
46. T.A. Freier, D.C. Beitz, L. Li, and P.A. Hartman, *Int. J. System. Bacteriol.* **44**, 137–142 (1994).
47. A.W. Brinkley, A.R. Gottesman, and D.E. Mott, *Appl. Environ. Microbiol.* **40**, 1130–1132 (1980).
48. B.D. Davis, R. Dulbecco, H.N. Eisen, H.S. Ginsberg, and W.B. Wood, Jr., *Microbiology*, 2nd ed., Harper & Row, Hagerstown, Md., 1973, p. 79.
49. E.J. Herbst and E.E. Snell, *J. Biol. Chem.* **176**, 989–990 (1948).
50. M. Rogosa and F.S. Bishop, *J. Bacteriol.* **87**, 574–580 (1964).
51. A.C.R. Tanner, M.A. Listgarten, J.L. Ebersole, and M.N. Strzenko, *Int. J. Syst. Bacteriol.* **36**, 213–221 (1986).
52. J.E. Cronan, Jr., and C.O. Rock, in F.C. Neidhart, J.L. Ingrahan, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger eds., Escherichia coli *and* Salmonella typhimurium: *Cellular and Molecular Biology*, vol. 1, American Society for Microbiology, Washington, D.C., 1987, pp. 474–497.
53. D. DeMendoza, R. Grau, and J.E. Cronan, Jr., in A.L. Sonenshein ed., Bacillus subtilis and *Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*, American Society for Microbiology, Washington, D.C., 1993, pp. 411–421.
54. J.W. Deming, J.K. Somers, W.L. Straube, D.F. Swartz, and M.T. MacDonell, *System. Appl. Microbiol.* **10**, 152–160 (1988).
55. G.K. Nyberg, G.P.D. Granberg, and J. Carlsson, *Appl. Environ. Microbiol.* **38**, 29–34 (1979).
56. L.G. Ljungdahl and J. Wiegel, in A.L. Demain and N.A. Solomon eds., *Manual of Industrial Microbiology and Biotechnology*, American Society for Microbiology, Washington, D.C., 1986, pp. 84–96.
57. J.A. Breznak and R.N. Costilow, in P. Gerhardt, R.G.E. Murray, W.A. Wood, and N.R. Krieg eds., *Methods for General and Molecular Bacteriology*, American Society for Microbiology, Washington, D.C., 1994, pp. 135–154.
58. W.E. Balch, G.E. Fox, L.J. Magrum, C.R. Woese, and R.S. Wolfe, *Microbiol. Rev.* **43**, 260–296 (1979).
59. R.G. Cluss and N.L. Somerson, *Appl. Environ. Microbiol.* **51**, 281–287 (1986).
60. R. Olivieri, M. Bugnoli, D. Amellini, S. Bianciardi, R. Rappuoli, P.F. Bayeli, L. Abate, E. Esposito, L. DeGregario, J. Asis, C. Basagni, and N. Figura, *J. Clin. Microbiol.* **31**, 160–162 (1993).
61. *The Source*, 7, Sigma Chemical, St. Louis, Mo., 1992, pp. 1–2.
62. R. Greasham and E. Inamine, in A.L. Demain and N.A. Solomon eds., *Manual of Industrial Microbiology and Biotechnology*, American Society for Microbiology, Washington, D.C., 1986, pp. 41–48.
63. T.L. Miller and B.W. Churchill, in A.L. Demain and N.A. Solomon eds., *Manual of Industrial Microbiology and Biotechnology*, American Society for Microbiology, Washington, D.C., 1986, pp. 122–136.

64. *Sheffield Series Technical Manual*, Quest International, Norwich, N.Y.

65. R.A. Neihof and C.A. Bailey, *Appl. Environ. Microbiol.* **35**, 698–703 (1978).

66. S.E. George, C.J. Costenbader, and T. Melton, *J. Bacteriol.* **164**, 866–871 (1985).

67. R.K. Bajpai and T.K. Ghose, *Biotechnol. Bioeng.* **20**, 927–935 (1978).

68. D.S. Ucker and E.R. Signer, *J. Bacteriol.* **136**, 1197–1200 (1978).

69. P. Loubiere, E. Gros, V. Paquet, and N.D. Lindley, *J. Gen. Microbiol.* **138**, 979–985 (1992).

70. H. Vogel, *J. Gen. Microbiol.* **26**, 223–230 (1961).

71. H.J. Laanbroek and N. Pfenning, *Arch. Microbiol.* **128**, 330–335 (1981).

72. J. Carlsson, G. Nyberg, and J. Wrethen, *Appl. Environ. Microbiol.* **36**, 223–2239 (1978).

73. S. Budavari, M.J. O'Neil, A. Smith, P.E. Heckelman, and J.F. Kinneary eds., *The Merck Index. An Encyclopedia of Chemicals, Drugs, and Biologicals*, 12th ed., Merck Research Laboratories, Whitehouse Station, N.J., 1996.

74. D.E. Gueffroy ed., *Buffers. A Guide for the Preparation and Use of Buffers in Biological Systems*, Calbiochem-Behring, La Jolla, Calif., 1983.

75. N.E. Good, G.D. Winget, W. Winter, T.N. Connolly, S. Izawa, and R.M.M. Singh, *Biochemistry* **5**, 467–477 (1966).

76. G.T. Spierenburg, F.T.J.J. Oerlemans, J.P.R.M. can Laarhhoven, and C.H.M.M. deBruyn, *Cancer Res.* **44**, 2253–2254 (1984).

77. *Art to Science in Tissue Culture*, vol. 6, Hyclone Laboratories, Ogden, Utah, 1987, pp. 4–5.

78. R. Hugh and E. Leifson, *J. Bacteriol.* **66**, 24–26 (1953).

79. E.N. Adaoha Mbanaso and D.H. Roscoe, *Plants Sci. Lett.* **25**, 61–66 (1982).

80. A.D. Lines, *Appl. Environ. Microbiol.* **34**, 637–639 (1977).

81. I.A. Abbott and F.A. Chapman, *Arch. Microbiol.* **128**, 355–359 (1981).

82. J.H. Hanks and R.L. Weintraub, *J. Bacteriol.* **32**, 639–651 (1936).

83. A.S. Dietz and A.A. Yayanos, *Appl. Environ. Microbiol.* **36**, 966–968 (1978).

84. D.A. Bazlinski and F.A. Rosenberg, *Appl. Environ. Microbiol.* **39**, 934 (1980).

85. H.B. Funk and T.A. Krulwich, *J. Bacteriol.* **88**, 1200–1201 (1964).

86. H. Seki, *Appl. Microbiol.* **26**, 318–320 (1973).

87. K.L. Roche and R.V. Levy, *Biopharm* **5**, 22–33 (1992).

88. J.J. Perkins, *Principles and Methods of Sterilization in the Health Science*, Charles C. Thomas, Publisher, Springfield, Ill., 1976.

89. L.C. Ellis, in *Art to Science in Tissue Culture*, vol. 10, Hyclone Laboratories, Ogdon, Utah, 1991, pp. 1–5.

90. M.E. Hoffman and R. Meheghini, *Photochem. Photobiol.* **29**, 299–303 (1979).

91. R.L. Girolami and J.M. Stamm, *Appl. Environ. Microbiol.* **31**, 141–142 (1976).

MEDIA, ANIMAL CELL CULTURE

S. A. CASNOCHA
R. A. WOLFE
Monsanto Co.
St. Louis, Missouri

KEY WORDS

Animal
Cell culture
Defined medium
Mammalian
Media/medium
Tissue culture

OUTLINE

INTRODUCTION

The aqueous environment to which cultured cells are exposed includes low molecular weight nutrients, hormones, transport proteins, buffers, and reagents intended to either reduce stresses imposed on the cultured cells in the bioreactor environment or stabilize the product. The medium, in many cases, also includes attachment proteins and/or hormones that adsorb to the surface(s) of the bioreactor and mediate interaction with the cultured cells. The cells rapidly modify the environment (both aqueous and substratal) by utilization of nutrients and secretion of proteins and metabolites. In many cases these secreted products can be detrimental to both the cells and the prod-

uct of interest. Thus it is common practice to include in media antioxidants, protease inhibitors, and buffers. The cell-specific rates of medium utilization and metabolite production impact many of the values for process control parameters, which are in turn influenced by the medium design. Media components that destabilize or copurify with the product must, of course, be avoided.

After a few days of culture in the bioreactor, the cells are exposed to a modified or "conditioned" medium, which differs considerably from the initial formulation. The levels of some amino acids such as alanine usually increase, while glutamine levels usually decrease. The metabolism of glutamine results in the secretion of ammonia (which can be cytotoxic at 1–2 mM), and glucose metabolism leads to lactic acid production (which causes medium acidification). Different clonal lines isolated from the same transfection, fusion, or primary isolate usually exhibit differences in growth characteristics and metabolism. This results in differences in the conditioned medium that frequently require "fine tuning" of the production process, initial medium, and supplemental feedstocks. Cells of comparable derivation generally do, however, perform well in similar bioreactor systems.

In developing any process utilizing animal cell culture, all of the aspects of upstream and downstream processing must be also considered. The goal is usually to consistently produce conditioned medium containing a high concentration of an intact, uniformly processed product, which can be readily isolated from other cellular products and medium components. The interrelationship between the cultured cells (and their associated metabolic activities), aqueous environment, substratum, bioreactor design (three-dimensional geometry, materials, mass transfer limitations, "shear stresses", etc.), and product of interest (including the associated downstream processing considerations) must be emphasized. This interrelationship dictates the approach for producing any product through mammalian cell culture technology. In Figure 1 many of the important parameters are grouped within these five categories to illustrate their interrelationships. The composition and range of aqueous environments utilized in mammalian cell culture are presented here, and further information can be found in discussions surrounding the other components of a bioreactor system.

While it is true that many factors of the production process influence medium design, it is, in practice, usually the case that the medium design dictates many aspects of the overall process strategy. In addition, the medium composition defines raw material requirements including their storage and quality control (QC). Accordingly, much of the overall cost of manufacturing the final product can depend on the medium composition and its degree of optimization. Therefore, a common first step toward process design is the definition of the hormonal, nutritional, and substratum requirements of the cell line. In most cases, these are independent of the bioreactor design and process parameters. Evaluation of these data in light of the restrictions defined by the desired product suggest appropriate conditions, bioreactors, processes, and the course for medium optimization. It should be noted that even small modifications in the medium can influence the cell line stability, the product quality/yield, operational parameters of the bioreactor, and downstream processing.

Medium Composition

The aqueous environment to which the cells are exposed results from a combination of many of the factors just discussed. Dissolved gasses and pH are monitored and controlled. Metabolites or other cell products may be either stimulatory or inhibitory to the cultured cells. Almost all of the medium components have both lower and upper limits within which the cells perform optimally, and these ranges are interdependent. The effective concentration range of a given amino acid, for example, is dependent on the concentrations of other amino acids that are taken up by the cells through the same uptake mechanism. The medium and process parameters must therefore be designed such that no component either exceeds or falls below its effective concentration range (both net utilization and biosynthesis of specific amino acids are observed, dependent on the metabolite and conditions). This concept of "metabolic balance" has been clearly expounded in relation to static culture systems (T-flasks) by Ham and his colleagues (1,2).

In practice, an analytical comparison of fresh and conditioned medium provides considerable insight in medium optimization (3). Most of the components routinely included in media are discussed here in detail and can be grouped into four categories: water, low molecular weight nutrients, proteins, and supplements of a nonhormonal/nonnutritional nature. Table 1 lists the components of six common medium formulations that are frequently used as a starting point in medium development.

Scrutiny of these formulations provides some insight into the complexity of media and the evolution of in vitro cell culture technology over the past 40 years. Dulbecco's Modified Eagle's (DME) medium (4–6) was designed for high-density culture of attachment-dependent cells in the presence of serum supplementation, whereas F-12 medium (7) represents an initial round of optimization for clonal (low-density) protein-free growth of CHO cells. The concentration of each component of F-12 was empirically determined through sequential concentration/growth response (C/GR) analyses. MCDB 302 medium was derived from F-12 via further C/GR analyses (8). RPMI medium was developed for suspension culture of lymphocytic cell lines in the presence of serum (9). Iscove's DME (10) and mixtures of DME/F-12 are commonly used to provide all the components necessary for low-density low-protein cultures, while also including high enough levels of nutrients to support mass cultures (11,12). The subsequent analysis of the amino acid and carbohydrate consumption by cells cultured in the specified bioreactor system utilizing a "good mixture" often suggests appropriate modifications of the basal nutrient formula.

Water Quality

It is recommended that water of at least the pharmaceutical "water for injection" (WFI) quality be utilized for mammalian cell culture, and it should be noted that such preparations are not necessarily pure; they simply meet

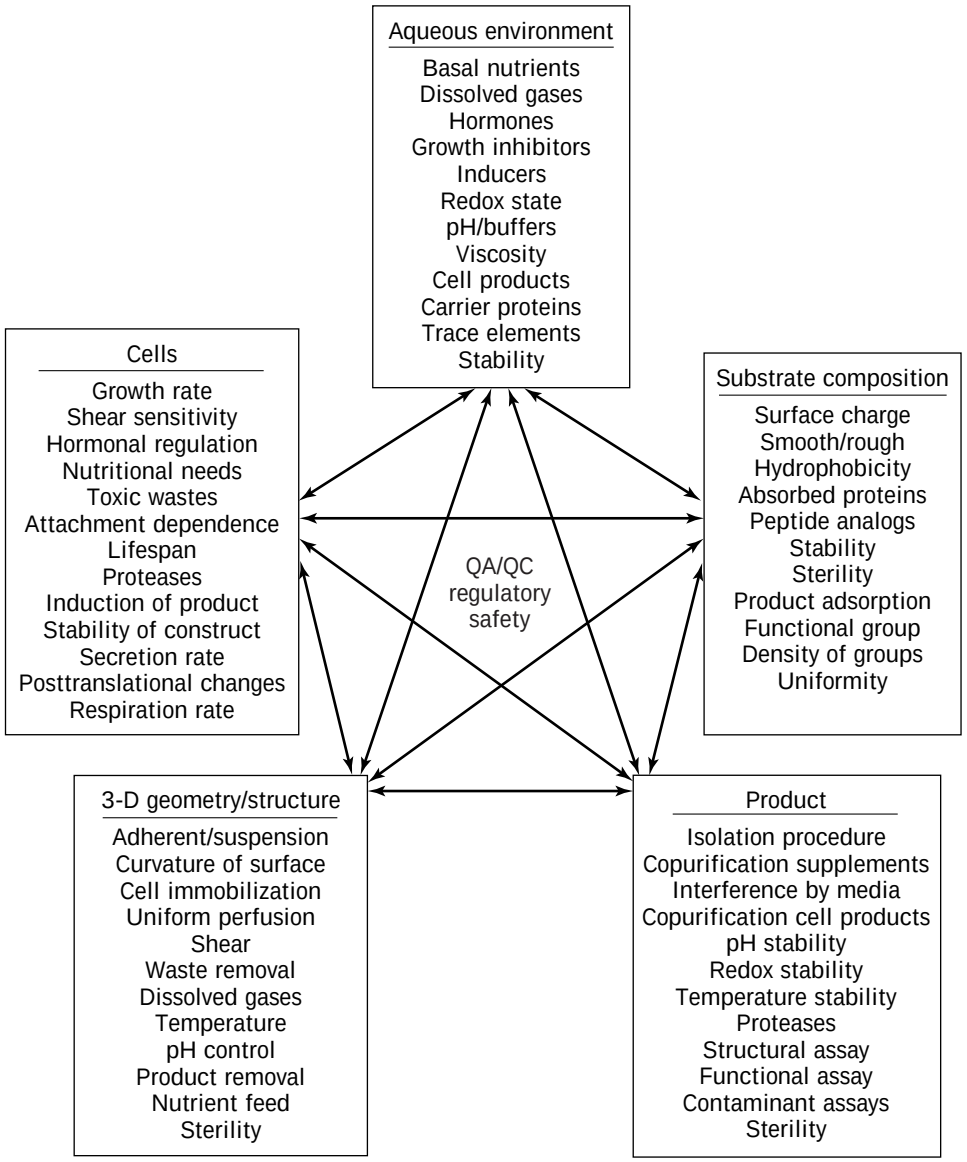


Figure 1. Bioreactor systems: interrelationship of components.

standards for injection. Agents present in water are diluted throughout the organism with injection, but remain in a relatively concentrated form in cell culture.

Water from any natural source is full of contaminants, and mammalian cells are exceedingly sensitive to many compounds that often contaminate the water used for media preparation (13). This is particularly noticeable with "reduced protein" formulations, and contaminants not removed during the purification process or introduced during storage are common causes of aberrant cultures. These contaminants can be either growth stimulatory or growth inhibitory (with the latter being more common). Unpurified water can contain organics (from detergents or degraded biomass), inorganics (such as chlorine and heavy metals), microorganisms (which often result in endotoxin/pyrogen contamination), or particles.

It is recommended that a multiple step purification is used to purify water for cell culture. In the authors' facility, water is filtered, deionized, adsorbed with carbon, purified by reverse osmosis, passed through a Milli-Q™ polishing unit (Millipore Corporation, deionization/carbon adsorption), depyrogenated via ultrafiltration, and then stored at 80 °C in stainless steel tanks until use. Routine maintenance and monitoring of the various stages of the purification stream are performed. Other techniques such as distillation and UV oxidation are commonly utilized.

Low Molecular Weight Nutrients

Bulk ions and trace elements should be considered nutrients because their absorption, secretion, and concentration in a given organism are finely controlled (14). The concentrations of these ions maintain the osmotic pressure of the microenvironment, and their distribution within the tissue (primarily Na^+ and K^+) maintains membrane potential. Therefore, both the absolute and relative amounts of sodium and potassium are important. The transient alterations in the availability of Ca^{2+} and Mg^{2+} mediate

Table 1. Media Formulations Commonly Used in Animal Cell Culture

Compound (g/mol)	DME	F-12	MCDB 302	RPMI 1640	Iscove's DME	Coon's DME:F12
Bulk ions and trace elements						
$CaCl_2$ (anh) (111.00)	1.80×10^{-3}	2.99×10^{-4}	6.00×10^{-4}		1.49×10^{-3}	1.36×10^{-3}
$CuSO_4 \cdot 5H_2O$ (249.68)		9.97×10^{-9}	1.00×10^{-8}			4.01×10^{-9}
$FeSO_4 \cdot 7H_2O$ (278.02)		3.00×10^{-6}	3.00×10^{-4}			1.44×10^{-4}
$Fe(NO_3)_3 \cdot 9H_2O$ (404.00)	2.48×10^{-7}					1.24×10^{-7}
KCl (74.55)	5.37×10^{-3}	3.00×10^{-3}	3.00×10^{-3}	5.37×10^{-3}	4.43×10^{-3}	4.65×10^{-3}
KH_2PO_4 (anh) (136.09)						2.18×10^{-4}
$MgCl_2$ (anh) (95.22)		6.01×10^{-4}				2.60×10^{-4}
$MgSO_4$ (anh) (120.40)	8.11×10^{-4}			4.06×10^{-4}	8.11×10^{-4}	4.85×10^{-4}
NaCl (58.44)	1.10×10^{-1}	1.30×10^{-1}	1.30×10^{-1}	1.03×10^{-1}	7.71×10^{-2}	1.19×10^{-1}
$Na_2HPO_4 \cdot$ (142.1)		1.00×10^{-3}		2.98×10^{-3}		2.33×10^{-4}
$NaH_2PO_4 \cdot 2H_2O$ (156.00)	8.01×10^{-4}				8.01×10^{-4}	2.86×10^{-4}
$ZnSO_4 \cdot 7H_2O$ (287.54)		3.00×10^{-6}	2.87×10^{-6}			2.50×10^{-7}
Buffers and other components						
Carbon Dioxide (CO_2, gas)						
$NaHCO_3$ (84.01)	4.40×10^{-2}	1.40×10^{-2}	1.40×10^{-2}	2.40×10^{-2}	3.60×10^{-2}	3.65×10^{-2}
Essential amino acids						
L-Arg (HCl) (210.70)	3.99×10^{-4}	1.00×10^{-5}	1.00×10^{-3}	9.49×10^{-4}	3.99×10^{-4}	5.49×10^{-4}
L-Cys HCl $\cdot H_2O$ (175.60)	3.56×10^{-4}	2.00×10^{-4}	1.00×10^{-4}	3.71×10^{-4}	5.20×10^{-4}	1.82×10^{-4}
L-Gln (146.10)	4.00×10^{-3}	1.00×10^{-3}	3.00×10^{-3}	2.05×10^{-3}	4.00×10^{-3}	3.00×10^{-3}
L-Hts HCl $\cdot H_2O$ (209.70)	2.00×10^{-4}	1.00×10^{-4}	1.00×10^{-4}	7.15×10^{-5}	2.00×10^{-4}	1.91×10^{-4}
L-Ile (131.20)	8.00×10^{-4}	3.00×10^{-5}	3.00×10^{-5}	3.81×10^{-4}	8.00×10^{-4}	4.31×10^{-4}
L-Leu (131.20)	8.00×10^{-4}	9.98×10^{-5}	1.00×10^{-4}	3.81×10^{-4}	8.00×10^{-4}	4.99×10^{-4}
L-Lys HCl (182.70)	7.99×10^{-4}	2.00×10^{-4}	2.00×10^{-4}	2.19×10^{-4}	7.99×10^{-4}	6.02×10^{-4}
L-Met (149.20)	2.01×10^{-4}	3.00×10^{-5}	3.00×10^{-5}	1.01×10^{-4}	2.01×10^{-4}	1.27×10^{-4}
L-Phe (165.20)	4.00×10^{-4}	3.00×10^{-5}	3.00×10^{-5}	9.08×10^{-5}	4.00×10^{-4}	2.30×10^{-4}
L-Thr (119.10)	7.98×10^{-4}	1.00×10^{-4}	1.00×10^{-4}	1.68×10^{-4}	7.98×10^{-4}	5.00×10^{-4}
L-Trp (204.20)	7.84×10^{-5}	9.99×10^{-6}	9.99×10^{-6}	2.45×10^{-5}	7.84×10^{-5}	4.90×10^{-5}
L-Ttr (diNaSalt) $2H_2O$ (237.20)	4.38×10^{-4}	3.28×10^{-5}	3.31×10^{-5}	1.10×10^{-4}	4.38×10^{-4}	2.49×10^{-4}
L-Val (117.20)	8.00×10^{-4}	9.98×10^{-5}	1.00×10^{-4}	1.71×10^{-4}	8.00×10^{-4}	5.03×10^{-4}
L-Cystine 2HCL (310.80)						1.01×10^{-4}
Nonessential amino acids						
L-Ala (89.09)		9.99×10^{-5}	1.00×10^{-4}		2.81×10^{-4}	1.01×10^{-4}
L-Asn $\cdot H_2O$ (150.10)		1.00×10^{-4}	1.00×10^{-4}	3.33×10^{-4}	5.60×10^{-4}	8.66×10^{-5}
L-Asp (133.10)		9.99×10^{-5}	1.00×10^{-4}	1.50×10^{-4}	2.13×10^{-4}	9.77×10^{-5}
L-Glu (147.10)		1.00×10^{-4}	1.00×10^{-4}	1.36×10^{-4}	5.10×10^{-4}	1.02×10^{-4}
Gly (75.07)	4.00×10^{-4}	9.99×10^{-5}	1.00×10^{-4}	1.33×10^{-4}	4.00×10^{-4}	3.06×10^{-4}
L-Pro (115.10)		3.00×10^{-4}	3.00×10^{-4}	1.74×10^{-4}	3.48×10^{-4}	3.04×10^{-4}
L-Ser (105.10)	4.00×10^{-4}	1.00×10^{-4}	1.00×10^{-4}	2.85×10^{-4}	4.00×10^{-4}	3.04×10^{-4}
Amino acid derivatives						
Putrescine 2HCl (161.10)		9.99×10^{-7}	9.99×10^{-7}			9.31×10^{-7}
Water-soluble vitamins and coenzymes						
Ascorbic acid (176.12)						4.26×10^{-5}
Biotin (244.30)	1.64×10^{-5}	2.99×10^{-6}	3.00×10^{-5}	8.19×10^{-7}	5.32×10^{-6}	1.43×10^{-7}
D-Ca pantothenate (238.30)	1.68×10^{-5}	2.01×10^{-8}	9.99×10^{-7}	1.05×10^{-4}	1.68×10^{-5}	9.44×10^{-6}
Folic acid (441.40)	9.06×10^{-6}	2.95×10^{-6}	3.00×10^{-6}	2.27×10^{-6}	9.06×10^{-6}	5.66×10^{-6}
Niacinamide (122.10)	3.28×10^{-5}	3.03×10^{-7}	3.00×10^{-7}		3.28×10^{-5}	1.65×10^{-5}
Pyridoxal HCl (203.60)					1.96×10^{-5}	9.82×10^{-6}
Pyridoxine HCl (205.60)		3.02×10^{-7}	3.00×10^{-7}	4.86×10^{-6}		1.46×10^{-7}
Riboflavin (376.40)	1.06×10^{-6}	1.01×10^{-7}	9.99×10^{-8}	5.31×10^{-7}	1.06×10^{-6}	5.84×10^{-7}
Thiamine HCl (337.30)	1.19×10^{-5}	1.01×10^{-6}	9.99×10^{-7}	2.96×10^{-6}	1.19×10^{-5}	6.37×10^{-6}
Vitamin B_{12} (1355)		1.00×10^{-06}	1.00×10^{-6}	3.69×10^{-9}	9.59×10^{-5}	5.02×10^{-7}

Table 1. Media Formulations Commonly Used in Animal Cell Culture (*continued*)

Compound (g/mol)	DME	F-12	MCDB 302	RPMI 1640	Iscove's DME	Coon's DME:F12
			Carbohydrates and derivatives			
D-Glucose (180.16)	2.50×10^{-2}	1.00×10^{-2}	8.00×10^{-3}	1.11×10^{-2}	2.50×10^{-2}	1.75×10^{-2}
Na pyruvate (110.00)	1.00×10^{-3}	1.00×10^{-3}	1.00×10^{-3}		1.00×10^{-3}	1.50×10^{-4}
			Purine derivatives			
Hypoxanthine (Na Salt) (146.10)		2.87×10^{-5}	3.70×10^{-5}			1.60×10^{-5}
			Pyrimidines			
Thymidine HCl (337.30)		2.16×10^{-4}	2.16×10^{-6}			1.04×10^{-4}
			Lipids and derivatives			
Choline chloride (139.60)	2.87×10^{-5}	1.00×10^{-4}	1.00×10^{-4}	2.15×10^{-5}	2.87×10^{-5}	6.45×10^{-5}
i-Inositol (180.20)	4.00×10^{-5}	9.99×10^{-5}	1.00×10^{-4}	1.94×10^{-4}	4.00×10^{-5}	6.99×10^{-5}
Linoleic acid (280.40)		2.85×10^{-7}	3.00×10^{-7}			1.43×10^{-7}
Lipoic acid (Thioctic) (206.30)		1.02×10^{-6}	9.99×10^{-7}			4.85×10^{-7}
Total osmolarity (mosm)	367	316	312	293	286	357
Total sodium ion (mM)	156	145	133	116	157	137
Total potassium ion (mM)	5	3	3	5	4	5
Total chloride ion (mM)	120	135	135	109	86	129

transmembrane signaling events, and these ions are also necessary for integrin-mediated cell adhesion. Among other biochemical roles, phosphate and bicarbonate ions buffer the acidity of the microenvironment. Other ions such as selenium, vanadium, copper, zinc, iron, molybdenum, and manganese are necessary as cofactors for certain enzymes. Depending on their concentration, trace elements can be either stimulatory or inhibitory in cell culture, and therefore must be used at optimal concentration.

Trace elements are contained as contaminants in the reagents used to prepare media, and care must be taken to assure that such unknown materials remain consistent throughout a production operation. In order to insure this consistency, it is common practice to "ball mill" large batches of powdered nutrients, and perform QC on random aliquots. The batch is then packaged in appropriately sized containers because the uniformity of the powder decays with time. Preparation of media from these large lots then requires minimal QC (provided a good source of water is utilized). Verification of osmolarity, pH, O_2 and CO_2, levels (blood gas analyzer), and sterility is usually sufficient.

There are two major energy sources utilized by cultured mammalian cells: six-carbon sugars and glutamine (15). Glutamine (2–10 mM) can serve both as a source of energy and carbon. However, ammonia produced from the conversion of glutamine to glutamic acid can be detrimental at concentrations exceeding 2 mM, and cultures of cells such as murine hybridomas tend to build up significant concentrations of this metabolite. Cell lines can sometimes be adapted to tolerate slightly higher levels of ammonia, but success in this endeavor has been limited with levels not exceeding about 10 mM. Glutamine is also unstable in aqueous solution at 4 °C (half-life is approximately 100 days), and ammonia is one of the principal degradants. Thus, long-term storage of media containing this amino acid should be avoided. If necessary, glutamine-free formulations can be supplemented with this amino acid just prior to use. Glucose is the most common carbohydrate (2–20 mM) provided as a source of energy, but many cell lines can be adapted to growth in galactose or other six-carbon sugars (16,17). In this case, a desirable reduction in the rate of lactate production and six-carbon sugar utilization is sometimes observed, presumably achieved via a decrease in the steady-state concentration of glycolytic intermediates. Concomitant reduction in growth rate is often seen. Pyruvate can also function as an energy/carbon source.

Amino acids provide the major nitrogen source in these cultures and have been grouped into two categories as a result of tissue culture experiments: the essential and nonessential amino acids (Table 1). However, most cultured cells perform better in the presence of the nonessential amino acids, and increases in the quantitative requirements for the essential amino acids are observed in cultures deprived of nonessential amino acids. An exception to this classification is the CHO-cell-derived lines that are auxotropic for proline (0.1–1 g/L generally required). Other amino acids such as serine and glycine have been shown to be beneficial supplements for low-density cultures but are relatively unimportant under conditions generally employed in large-scale mammalian cell cultures.

Vitamins have classically been grouped according to their solubility water soluble (B_{12}, biotin, folic acid, niacinamide, pantothenic acid, pyridoxine, riboflavin, and thiamine) or fat soluble (vitamins A, E, K). These compounds are enzymatic cofactors that are necessary for cellular metabolism and are reutilized (not metabolized and/or degraded) by the cells. The fat-soluble vitamins are usu-

ally added to the medium from a concentrated ethanolic stock and are not particularly stable in solution. Vitamin A (a retinoid) can have no effect, stimulate growth, or induce differentiation (dependent on the cell type and culture conditions). Vitamin E serves as an antioxidant, and vitamin K is necessary for the γ-carboxylation of proteins. The water-soluble vitamins function in multiple metabolic roles and are generally necessary components of a medium formulation. Vitamin C, a water-soluble antioxidant, has been demonstrated to be beneficial in culture, particularly with cells that tend to synthesize collagen. It is readily oxidized, and continuous supplementation is advisable if beneficial effects are to be observed (18).

Many cultured cells require supplementation with fatty acids, phospholipids (and/or their precursors), and other hydrophobic compounds such as cholesterol. The quantitative and qualitative requirements vary with cell type, culture conditions, and duration of exposure to the supplement (19). Much is still to be discerned about the interconversion and biosynthesis of lipids by cultured cells, but a few generalizations can be made. *cis*-Unsaturated fatty acids (for example, linoleic, oleic, arachidonic) have been shown to be beneficial for many cell lines, but are relatively insoluble and inherently unstable in aqueous solutions. They are usually supplied as conjugates with delipidated albumin or cyclodextrins, or added as unilamellar vesicles (liposomes). Certain lots of albumin or improperly prepared albumin conjugates can be quite toxic. Liposomes have been utilized to provide cells with both fatty acids and phospholipids such as spingomyelin, phosphotidyl choline, phosphotidyl ethanolamine, and phosphotidyl inositol. This approach is difficult to utilize in large-scale cell cultures; in many cases the need for supplementation with the more complex lipids can be circumvented by providing the cells with precursors such as choline, ethanolamine, and inositol. A few cells such as NS-1-derived hybridomas have a deficiency in cholesterol biosynthesis, and supplementation is usually necessary.

Nucleic acid precursors such as the nucleosides and nucleotides are usually not required in a medium formulation, provided at least one source of pyrimidine and purine is supplied. Therefore, hypoxanthine and thymidine are included in most media. Particular care must be taken with mutant lines lacking enzymes of the biosynthetic pathways or with cultures grown without glucose. In the latter case a source of five-carbon sugars or a small concentration of glucose is advisable. Hybridoma cell lines produced via NS0 (a cell line derived from NS1) fusions do not express functional glutamine synthetase (GS), and therefore this line has been used to produce cell constructs using GS as a selectable marker (20). Recombinant lines produced in this manner grow in glutamine-free medium, but this formulation is usually supplemented with adenosine, guanosine, cytidine, uridine, and elevated levels of thymidine.

Culture in glutamine-free medium results in a system where ammonia does not accumulate to toxic levels, and frequently, higher cell densities can be achieved. However, even after adaptation to glutamine-free medium, the proliferation rate is generally less than that seen in glutamine-containing media. NS0-derived lines usually need to be provided with a source of cholesterol unless considerable "weaning/adaptation" has been performed (21).

Proteins

The first cultures were established in mixtures of plasma and tissue extracts (22–24). More defined and reproducible nutrient media designed in the late 1950s to early 1960s were developed using cultures supplemented with dialyzed serum. We now know that serum provides a host of factors not contained in many basal nutrient formulations, including low molecular weight nutrients (sugars, lipids, vitamins, amino acids, trace elements, etc.), hormones, blood proteins, metabolites, and protective compounds. Significant inter-lot, seasonal, and source variation in serum composition exists.

Because quantitative analysis of all of the serum components is not yet possible (particularly for the proteins, lipids, and lipoprotein complexes), variation in the serum supplement is frequently cited as the principal source of variability between cultures. It is also the root of many problems associated with the production of mammalian tissue culture products. Supplementation with 10% serum results in the contamination of the product-enriched medium with a mixture of hundreds of proteins totaling about 4 g/L. This significantly complicates and increases the expense of downstream processing. In addition, serum is expensive, and the availability of high quality serum has been (and is likely to be) a problem for large-scale efforts.

It is now well accepted that the major role played by serum in cell culture is to provide hormones, growth factors, transport proteins, and attachment factors. Serum also has other beneficial effects in culture. It supplies "generic" protein that coats the system, preventing nonspecific adsorption of essential factors to the bioreactor walls and filters. Enzymes contained in serum modify media components, producing molecules the cells can more readily utilize, and also detoxify other materials. Serum also provides essential nutrients that are not included in the older, less-complex basal media. Recently, the importance of the numerous protease inhibitors present in serum has also become apparent, and several reports of increased product integrity in cultures supplemented with protease inhibitors are present in the literature (tPA/aprotinin). In addition, the physical properties of medium (colloid osmotic pressure, diffusion rates, viscosity) are altered by serum supplementation.

If serum is utilized in a culture system, extensive labor-intensive prescreening of several lots is usually necessary. This screening usually focuses on the growth and productivity of the cell line, and small-scale testing for interference in the product isolation. Long-term growth, cellular morphology, cellular productivity, and the possibility of adventitious contaminants must be examined. Once identified, a large lot of serum is usually purchased. Long-term storage of costly, unstable biologicals such as serum in a monitored freezer facility is expensive. Taken together, these factors have driven the development of the technology to reduce the need for serum supplementation in mammalian cell culture. In fact, with the recent concern about potential prion contamination (25), a major emphasis has

been placed on eliminating all animal-derived components from culture systems designed to produce parenteral pharmaceuticals.

Many of the useful growth promoting agents, attachment factors, carrier proteins, and protease inhibitors found in serum have been identified and characterized by investigators attempting to reduce the level of serum supplementation in their particular culture system by providing their cultured cells with purified molecules (26). In the last few years a significant industry has emerged to market recombinant proteins for health care, research, and cell culture manufacturing. Many proteins (or modified polypeptides exhibiting the same activity in vitro) can be obtained at a reasonable expense if ordered in quantity. In fact, most protein-containing serum-free formulations are significantly less expensive than medium supplemented with 5% serum. Proteins commonly utilized in these media formulations are discussed later and listed in Table 2 (along with other supplements).

Hormones and growth factors bind to cellular receptors and, in the process, signal the cells to perform specific tasks (growth factors are hormones that were discovered in tissue culture). Mammalian cells evolved in the presence of a mixture of these factors, only some of which are well characterized today. The many components of basal cellular physiology are regulated by the concentrations of hormones to which the cells are exposed (27). In addition, a given cell (depending upon the receptors that the cell expresses) can be signaled to grow, die (apoptosis), or differentiate by exposure to different regulatory factors. Hormones such as steroids, insulin, and thyroid hormone regulate basal metabolic activities. In many cases such signals are necessary for continued viability of cells in culture. In fact, nearly every cultured mammalian cell requires stimulation by insulin (or an insulin-like growth factor [IGE]) for survival (insulin, 0.5–10 μg/mL; IGF, 1–10 ng/mL).

Analogs of many of the proteins produced through recombinant DNA technology exhibit activities similar to those of native proteins in vitro (although pharmacokinetics, tissue distribution, and pharmacodynamics may differ in vivo). Some of such analogs are now available from commercial sources and provide the necessary regulatory signal(s) to the cultured cells when present at 1–10 μg/L. The quality and expense of such analogs can be consistent with the requirements of animal cell production systems for parenteral pharmaceuticals. Analogs of IGF and EGF, for example, are marketed for less than $50 per milligram when purchased in quantity.

Transport proteins such as albumin, transferrin, low- and high-density lipoproteins (LDL, HDL), transcobalamin, hemoglobin, and thyroid hormone binding protein bind necessary nutrients and hormones that would in their unbound form be toxic to the cultured cells at high concentrations. These compounds are either slowly released in a controlled manner (fatty acids, thyroid hormone) or are transported into cells via binding-protein-mediated or cell-receptor-mediated processes (Fe^{2+} and transferrin). In "reduced serum" cultures, most cells require supplementation with additional transferrin (0.5–10 μg/mL) unless frequent supplementation of the medium with ferrous citrate/sulfate and iron chelators is performed; some success utilizing hemoglobin as a transport moiety has also been reported. Transport proteins also serve to help detoxify medium by binding contaminants (heavy metals, detergents) and solubilizing essential nutrients that do not readily dissolve in aqueous solutions (cholesterol and other lipids and lipoproteins, and transcobalamin). Lipids, for instance, when added to cultures above the critical micellular concentration rapidly form micelles that can lyse cultured cells. Albumin and serum lipoproteins are frequently used to supply lipids and cholesterol to cultured cells (28). It should be noted that the lipid composition of these supplements can vary considerably between preparations, and care must be taken to select one appropriate for the cell of interest.

In the last decade many cytokines have been identified and are now available from commercial sources. Culture of blood and marrow-derived cells (i.e., for cell-based therapies) is essentially impossible in the absence of the appropriate mixture of cytokines, and the hormonal signals that direct the proliferation and differentiation of the early progenitor (stem) cells are not fully understood. Large-scale cultures of such cells are used to produce, isolate, and characterize new cytokines and have led to several that are being utilized in adjunctive therapy to cancer treatment to speed recovery after chemotherapy or radiation treatment. Reconstitution of the immune system through ex vivo expansion and posttherapy infusion of bone marrow cells (or mobilized peripheral blood progenitor [stem] cells [PBSC]) taken from patients prior to high-dose chemotherapy is currently being tested in clinical trials. The effectiveness of these cultures depends on the development of an appropriate balance of cytokine signals in the culture media (29). Several cytokines that are being examined in such medium formulations are listed in Table 2. The use of other recently discovered cytokines, for example IL-15, which stimulates T-cell proliferation, are likely to become more significant in the near future.

Attachment proteins (fibronectin, laminin, vitronectin) adsorb to culture surfaces and then interact with anchorage-dependent cultured cells via specific receptor molecules. The attachment, spreading, and growth of these cells is abnormal (and/or limited) in the absence of these proteins. These molecules (along with proteoglycans) are secreted by cells in vivo to form the extracellular matrix (ECM) with which the cell and its neighbors interact. The composition of the ECM varies with the tissue and has been shown to regulate certain aspects of cellular physiology (30). Thus, it is sometimes necessary to provide cultured cells with the appropriate molecules in order for the cells to perform specific desired tasks in vitro. These molecules can be either preadsorbed to the surfaces or added to the medium at concentrations of a few micrograms per milliliter during inoculation of anchorage-dependent cell cultures. Precoating microcarriers with serum proteins is also a common practice, and some success utilizing the attachment motifs in synthetic peptides has been achieved.

Synthetic polymers containing multiple copies of attachment motifs such as RGD (Arg-Gly-Asp) in many cases adequately substitute in culture systems for proteins derived from animal sources, and the expense and quality of

Table 2. Proteins/Polypeptides Used in Animal Cell Culture

Supplement	Tissue/cell source	Structure	Functions/activities
		Growth factors	
Nerve growth factor β (NGF-β)	Submaxillary	Two-chain protein, 26 kDa	Cell growth, neuron development, and survival
Epidermal growth factor (EGF)	Submaxillary, tumor-associated macrophages, prostate tumors	Polypeptide, 6.1 kDa	Stimulates broad range of responses including growth
Platelet-derived growth factor (PDGF)	Platelets, ovarian and lung tumors, melanomas	Dimeric glycoprotein, 28.6, 27.6 kDa (AA or BB)	Stimulates broad range of responses including growth
Insulin-like growth factor I (IGF-I)	Liver, breast tumors	Polypeptide, 7.6 kDa	Stimulates growth and insulin-like activity
Insulin-like growth factor II (IGF-II)	Liver	Polypeptide 7.5 kDa	Stimulates growth and insulin-like activity
Fibroblast growth factor (basic) (bFGF)	Brain, hypothalamus, pituitary, kidney	Monomeric protein 16.4 kDa	Stimulates growth of mesodermal cells
Fibroblast growth factor (acidic) (aFGF)	Brain, hypothalamus	Monomeric protein, 15.9 kDa	Stimulates growth of mesodermal cells
Liver cell growth factor	Plasma	Gly-His-Lys	Synergistic with Cu^{2+} to stimulate growth and survival
Transforming growth Factor β (TGF-β)	Kidney, platelets, and renal, breast, and prostate carcinomas	Two-chain protein, 25 kDa	Stimulation or inhibition of growth or differentiation
Transforming growth factor α (TGF-α)	Breast, neuroendocrine, and squamous carcinomas	Two-chain protein, 6.0 kDa	Similar to EGF
Vascular endothelial cell growth factor (VEGF)	Variety of tumors, tumor-associated macrophages	Dimeric protein, 28 kDa	Stimulate growth and survival of microvasculature cells
		Cytokines	
Erythropoietin (EPO)	Kidney	Glycoprotein, 45 kDa	Stimulates growth and differentiation, red cell precursors
Interferon α/β (IFN-α/β)	Fibroblasts, T and B cells, monocytes, macrophages	Monomeric protein, α 16–28 kDa, β 20 kDa	Inhibits growth of B and T cells and some tumors
Interferon γ (IFN-γ)	T cells	Homodimer, 40 kDa	Stimulate NK and T-cell growth, differentiate myeloid cells
Granulocyte colony stimulating factor (G-CSF)	Macrophage	Monomeric protein, 23 kDa	Growth and differentiation of granulocyte lineage
Granulocyte macrophage colony stimulating factor (GM-CSF)	Fibroblasts, T cells	Monomeric protein, 14 kDa	Growth and differentiation of granulocyte/macrophage lineage
Tumor necrosis factor β (TNF-β)	Macrophages	Monomeric, 19 kDa	Stimulates fibroblast growth and activates a variety of cells
Interleukin 1 (IL-1 [α/β])	Macrophages, monocytes, T and B cells, endothelial cells, fibroblasts	Monomer, 17.5/17.3 kDa	Stimulates inflammatory and T-cell growth
Interleukin 2 (IL-2)	T cells	Monomer, 15 kDa	Stimulates T cells, NK cells, and TIL cells
Interleukin 3 (IL-3)	T lymphoma, eosinophils, mast cells, keratinomas	Monomer, 16 kDa	Stimulates growth of megakaryocyte and myelocyte lineages, late stem cells
Interleukin 4 (IL-4)	T cells, stromal cells, mast cells, thymoma	Glycoprotein, 16 kDa	Stimulates B- and T-cell growth, isotype switching (IgG4 and IgE)
Interleukin 5 (IL-5)	T cells, eosinophils, mast cells	Glycoprotein homodimer, 45 kDa	Eosinophil proliferation and differentiation
Interleukin 6 (IL-6)	Monocytes, fibroblastoid cells, T, B, endothelial, and stromal cells, hepatocytes, fibroblasts	Monomeric protein, 16 kDa	Stimulates stem and B-cell growth, hepatic acute phase, hybridoma survival
Interleukin 7 (IL-7)	Stromal cells, spleen cells	Monomeric protein, 17 kDa	Stimulates progenitor B- and T-cell growth

Table 2. Proteins/Polypeptides Used in Animal Cell Culture (*continued*)

Supplement	Tissue/cell source	Structure	Functions/activities
Interleukin 8 (IL-8)	Neutrophils, monocytes, fibroblasts, T cells, endothelial and epithelial cells	Dimeric protein, 8 kDa	Stimulates angiogenesis, neutrophils, and chemotaxis
Interleukin 9 (IL-9)	T cells	Monomeric protein, 14 kDa	Proliferation of T cells and erythroid precursors
Interleukin 10 (IL-10)	B and T cells, monocytes, keratinocytes	Homodimeric protein, 37 kDa	Stimulates B-cell, mast cell, and thymocyte growth
Interleukin 11 (IL-11)	Fibroblasts and marrow stromal cells, osteosarcomas	Dimeric protein, 8 kDa	Stimulates growth of early and late megakaryoblastic and myeloblastic cells
	Attachment/substratal proteins and compounds		
Cell Tak™	Marine mussel	Polyphenolic protein	Effectively "glues" most cells to most surfaces
Collagen types I–IV	Most tissues, distribution specific for type	Multisubunit proteins	Binds and presents adhesion molecules and stimulates attachment and growth
Extracellular matrix (solubilized)	Extracts of EHS mouse sarcoma and human placenta available	Mixture of proteoglycans, collagens, and attachment factors	Promotes attachment, spreading, mitosis, and differentiation of many epithelioid cells
Fibronectin	Plasma	Dimeric protein, 440 kDa	Promotes attachment and spreading via RGD (Arg-Gly-Asp) and other motifs of most cells from mesenchyme
Laminin	Extracellular matrix	Dimeric protein, 900 kDa	Stimulates attachment and growth of cells derived from endoderm and ectoderm; stimulates neutrophil oxidative burst
Pronectin®F	Synthetic polymer	Polypeptide with multiple RGD motifs	Supports attachment of some cell lines (e.g., CHO, 3T3, BHK)
Poly-D-lysine	Synthesis	Polypeptide, polymers, mw > 50,000	Reduces net negative charge of surface-enhancing attachment
Vitronectin	Plasma	Monomeric protein, 70 kDa	Stimulates attachment and growth of many cell types
	Transport proteins		
Albumin	Plasma	Monomeric protein, 68 kDa	Carries fatty acids and trace elements and can detoxify
Ceruloplasmin	Plasma	Protein, 135 kDa	Transports copper
Hemoglobin	Red cells	Quaternary protein, 65 kDa	Binds oxygen, carbon dioxide and can detoxify
High-density lipoprotein	Plasma	Lipid, cholesterol and multiple protein subunits	Transports cholesterol, cholesterol esters, and other lipids
Low-density lipoprotein	Plasma	Particle containing lipid, cholesterol, and protein (Apoprotein B) subunits	Transports cholesterol, cholesterol esters, and other lipids
Transferrin	Plasma	Monomeric protein, 78 kDa	Carries iron and can detoxify

such polymers can also be consistent with the limitations imposed on manufacturing a pharmaceutical protein. Plasma and chemically treated culture surfaces are widely utilized to provide a surface that more closely mimics the ECM than untreated plastic or glass. Both small-scale plasticware and microcarriers that have surfaces with exposed attachment motifs have recently been introduced into the marketplace.

Nonhormonal and Nonnutritional Supplements

Other compounds are frequently included in medium formulations to reduce specific detrimental conditions that

are generated in the artificial microenvironment of the bioreactor, or to facilitate control and monitoring of the system. Also included in this group are antibiotics, which although frequently utilized are not necessary nor recommended. These compounds interfere with cellular metabolism and have been developed by selecting molecules that are more active with microbial enzymes than their analogous mammalian counterparts. Nevertheless, most antibiotics compromise some aspects of mammalian metabolism when present in their effective concentration range, and all are toxic at relatively high dosages. Serum and other supplements can, in some cases, reduce the detrimental effects of antibiotics on cultured cells, and the inhibitory effects of a specific antibiotic are greatly dependent on the cell type. Therefore, if antibiotics must be utilized in a specific culture system, an extensive dose–response analysis should be performed under conditions resembling (as close as possible) those of the bioreactor. Potential copurification with the desired product or interference with downstream processing must also be examined.

In the bioreactor, pH is stabilized with buffers (phosphate and bicarbonate) and usually maintained through control of dissolved CO_2 and base addition. Phosphate has a pK_a of 7.21 and is present in media (2–3 mM) both as a nutrient and as a buffer. The concentration of phosphate is limited by the solubility of calcium phosphate, and higher concentrations of phosphate are practical only in reduced-calcium formulations. The principle pH buffer in mammalian organisms is the bicarbonate–carbon dioxide system, and a similar control strategy is utilized with most bioreactor systems:

$$H_2O + CO_2 \rightarrow H_2CO_3 \rightarrow H^+ + HCO_3^-$$
$$pK' = 6.1 \text{ at } 37\ ^\circ\text{C and } 760 \text{ mmHg}$$

Where $[H_2CO_3]$ (mM) = 0.0308 PCO_2 (mmHg). Applying the Henderson–Hasselbalch equation ($pH = pK' + \log \{[HCO_3^-]/[H_2CO_3] = 6.1 + \log \{(\text{total } CO_2) - 0.0308\ PCO_2\}/0.0308\ PCO_2$) it is apparent that increases in the partial pressure of CO_2 result in decreases of pH. Equilibrium in a bioreactor is essentially dependent on the rate of CO_2 dissolution, diffusion, and bulk mixing because the rate of formation of H_2CO_3 from water and dissolved carbon dioxide is rapid ($T_{1/2}$ = 11.5 s) (31). At pH 7.4 the relative ratios are approximately 20,000:1000:1 (HCO_3^-:CO_2:H_2CO_3).

The amount of bicarbonate present in the medium is chosen to be close to equilibrium conditions for the specified pH operating condition. This amount is usually about 23 mM (similar to circulating blood) and will be at equilibrium with 5–10% CO_2. Bicarbonate in excess of this equilibrium point is used to titrate acid produced by the cultured cells. Carbon dioxide is also a nutrient, and aqueous solutions equilibrated with air at 37 °C contain only 0.21 mm CO_2 (7 μM). Therefore, if air equilibrated cultures are desired, it is recommended to use an additional organic buffer like HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid) and 1–2 mM bicarbonate to maintain adequate concentrations of dissolved CO_2.

Although relatively expensive, HEPES is frequently used as an additional buffer because, unlike bicarbonate, its p*K* falls in the physiological range (7.31 at 37 °C). The pK_a is, however, extremely temperature sensitive (-0.014 per change of 1 °C), and thus a room temperature solution at pH 7.55 would be pH 7.31 when warmed to operating temperature. HEPES is utilized at concentrations ranging from 10 to 50 mM. The concentration of sodium chloride needs to be reduced to compensate for changes in osmolarity at the higher levels, and a reduction in the level of potassium ion should be made to maintain similar Na^+/K^+ ratios. At concentrations of HEPES more than 15 mM the buffer is sometimes growth inhibitory (dependent on cell type via stimulation of the production of detrimental oxygen metabolites).

3-(*N*-Morpholino)propanesulfonic acid (MOPS) (pK_a = 7.20 at 20 °C) is also used for this purpose. The temperature sensitively is less (-0.006 per change of 1 °C), but the pK_a is slightly more acidic (which can be advantageous or detrimental, depending on the cell type).

Phenol red has classically been included in culture media as a pH indicator. There is no compelling reason for its use with instrumented systems and in fact has been shown to have estrogenic activity with cultured cells (C127, MCF-7) and to copurify with proteins (tPA, IFN-γ).

Many protective agents are often added to the medium to protect the cells from damage caused by osmotic gradients, shear, and interaction with air–liquid interfaces (bubble damage). The agents include dextran (30–50 g/L, 80,000 MW); polyethylene glycol (20,000 MW), methyl cellulose (0.1–0.2% w/v), and pluronic polyol F-68 (0.025–0.3%). The latter, a block copolymer of polyoxpropylene and polyoxyethelene of 8,300 MW, has been demonstrated to reduce bubble-induced cell damage and is the most frequently used protective agent (32).

Antioxidants are sometimes beneficial. Superoxide radicals and hydrogen peroxide are generated by normal cellular metabolism and enzymes such as xanthine oxidase present in serum supplements. These compounds can also be generated by photooxidation of compounds such as tryptophan and riboflavin, and their production can be stimulated by several factors such as hyperoxic conditions, general cell stress, and compounds such as HEPES. These highly reactive metabolites can cause damage to both the cultured cells and the cell products. Enzymes (superoxide dismutase, catalase), vitamins (C, E, and β-carotene), carrier proteins (transferrin, hemoglobin), amino acids, trace elements (Se^{2+} as a cofactor for glutathione peroxidase), bilirubin, and reduced glutathione are all antioxidants useful in mammalian cell cultures. All except the enzymes and glutathione are components of serum.

Reducing agents (β-mercaptoethanol) can help to prevent oxidative damage, primarily by restoring the reduced form of glutathione. When added to cultured cells, increases in cystine uptake (presumably by the mixed disulfide) and in antibody secretion rates (hybridomas) have been observed.

Many cells actively synthesize and secrete proteases, and even limited cell death (inherently a part of the microenvironment of any bioreactor) causes the release of intracellular protease into the medium. Therefore, supple-

mentation with protease inhibitors specific to proteases generated in a particular culture system is recommended. Cultures that have been established or maintained in serum-supplemented medium are particularly sensitive due to large quantities of adsorbed plasminogen (which is readily activated to plasmin by a variety of proteases). Ten to 40 kIU of aprotinin will effectively inhibit plasmin (and kallikrein) activity. ϵ-Amino caproic acid also inhibits plasmin at 2–7 mM concentrations (higher levels are somewhat toxic). Several other biocompatible inhibitors are now commercially available at nonprohibitive costs (see Table 3). Pepstatin inhibits acid proteases (at 0.7–1 μg/mL). Leupeptin is a reversible, competitive, polypeptide inhibitor of serine and thiol proteases that shows little toxicity at 5 μg/mL. In addition, both soybean and egg white trypsin inhibitors inhibit serine proteases, are effective between 10 and 100 μg/mL, and cost only a few cents per liter. Pefabloc® SC [4-(2-aminoethyl)-benzenesulfonyl fluoride] is relatively nontoxic (compared with phenylmethylsulfonyl fluoride or diisopropyl fluorophosphate) and blocks serine proteases (trypsinlike and chymotrypsin-like) when present at 1–2 mM concentrations. However, it is hydrolyzed relatively rapidly in medium at pH 7.0 (half life ~6 h) and is less stable in more basic environments ($t_{1/2}$ at pH 7.5 ~ 2 h).

Heparin or heparin sulfates (10–100 μg/mL) can help to reduce cell clumping when a normally adherent cell line is grown in suspension. Heparins also stabilize certain growth factors and enhance their bioactivity (endothelial cell growth factor, bFGF, VEGF). Heparins are crude mixtures available from many vendors and are not likely to be identical. Preparations are available with different nominal molecular weights, and those below 5,000 lack some of the biological activities present in other preparations.

MEDIA SELECTION

The most common goal of large-scale animal cell culture is the production of conditioned media with the highest possible titer of a homogeneous protein for subsequent isolation as a functionally active, pure product. In other cases the product can be an inactive proenzyme, a mixture of proteins from nonrecombinant cells, or even the cells themselves. Media selection is therefore highly dependent on the particular product and cellular system and is influenced by the design and composition of the bioreactor. A medium suitable for production of a specific cell product in a batch system is usually not ideal for producing the same product (with the same cells) in a highly instrumented perfusion system. The impact of medium components on downstream processing should guide medium development.

In vivo cells function differently depending on their environment (consider fibroblasts in wounds, osteoblasts, neutrophils, etc.). A given cell line can be cultured in different medium formulations. Metabolism can be modified by providing specific hormonal signals (e.g., insulin), providing physical signals (e.g., shear stress and endothelial cells), and/or altering the availability of nutrients (e.g., the scavenger pathways for nucleotide biosynthesis in hypoxanthine-aminopterin-thymidine selection). Cultures starved of glucose can utilize glutamine as an energy source, and lipid-depleted medium forces cultures into de novo lipid synthesis. The rate of lactic acid production can frequently be reduced by substituting other sugars for glucose (see the section "Low Molecular Weight Nutrients"), but concomitant reductions in proliferation rate are usually observed. Changes in the fatty acid composition of the medium can alter cell membrane fluidity and result in changes in signal transduction, oxidative damage, and other membrane-associated cellular processes (33). In cultures with a transport phenotype (i.e., differentiated cultures of intestinal origin), media nutrients can enhance the transport function (34).

Cellular metabolism alters in response to changes in the environment, and this process is referred to as *adaptation*. Large changes are usually inconsistent with cell viability and proliferation. Cells are generally adapted through

Table 3. Biocompatible Protease Inhibitors

Inhibitor	gm/mol	mg/L	Properties
Albumin	68,000	1,000	Serves as a nonspecific substrate
Aprotinin	6,511	6.8	Inhibits serine proteases (trypsin, kallikren, 40 U/mL = 6.8 μg/mL); at pH 7–8 aqueous solutions are stable for only 1 week at 4 °C
ϵ-Aminocaproic acid (EACA)	131.2	920	Inhibits plasmin
Egg white trypsin inhibitor	28,000	50	At 10–100 μg/mL inhibits serine proteases (trypsin, plasmin, coagulation factor Xa, kallikrein, no antichymotrypsin activity)
Leupeptin	460	2.0	Reversible competitive inhibitor of serine and thio proteases; plasmin, trypsin, kallikrein, cathepsin b, thrombokinase; 2 γg/mL nontoxic, plasmin inhibition 50% at 8 μg/mL
Ovoinhibitor		2.5	Inhibits tyrpsin and chymotrypsin equally on a weight basis; 5 μg/mL usually nontoxic
Pefabloc® SC	239.5	250	100–1,000 μg/mL inhibits serine proteases (daily addition in batch cultures due to rapid hydrolysis in medium)
Pepstatin	685	0.7	Inhibits acid proteases, especially pepsin, renin, cathepsin D, chymosin (make 2 mM stock in alcohol)
Bovine pancreas trypsin inhibitor		2.0	Inhibits trypsin and chymotrypsin. Usually 2–5 μg/mL, nontoxic
Soybean trypsin inhibitor	20,100	50	At 10–100 μg/mL inhibits serine proteases (trypsin, plasmin, coagulation factor Xa, kallikrein, no antichymotrypsin activity)

gradual changes made to a medium, which is selected based on the constraints of the product, the bioreactor system, and the nutritional requirements of the cell line. Further optimization is then performed. The composition of the final formulation depends on initial medium choice and the order in which components are optimized.

A considerable component of the cost of a product can be the medium cost (either directly or indirectly). Some media components are relatively expensive and in many cases not justifiable by the small incremental increases in product titer (e.g., nucleosides). The addition of some costly components can be supported by enhanced culture performance (e.g., insulin). Many examples of increasing product titer by more than an order of magnitude through medium optimization have been documented, and it is likely that no published formulation is "optimal" for a particular purpose. The potential for optimization can be limited by the availability of time, other resources, and the suitability of the current performance of the culture system.

Optimization is therefore highly dependent on the product and the tissue culture system and can result in reduced-serum formulations, a serum-free medium, a chemically defined formulation, or a combination that involves a medium exchange. In the latter case the cell mass is usually generated in medium supplemented with serum or a reduced-serum formulation in which the cells proliferate rapidly. The production of conditioned medium for product isolation then proceeds in more a defined, generally less expensive, formulation. This strategy also addresses downstream processing concerns by eliminating serum from the production phase. Although impractical in batch mode processes, perfusion systems can be utilized for this approach. A potential drawback of a medium exchange process is a reduction of cell mass with time. Also, cell damage can occur if a washing step is included in the process (resulting in the release of intracellular contents, including proteases, into the conditioned medium).

Reduced Serum and Serum-Free Formulations

The addition of serum or plasma to cell culture media dates back to the early years of cell culture (22–35) and logically follows the attempt to mimic the environment of cells in vivo, which are nourished from interstitial fluids that resemble blood plasma. Serum is the liquid phase remaining after the clotting of whole blood and, as such, is similar in composition to blood plasma (the noncell component of blood), except for the absence of fibrinogen and other clotting factors and increased levels of serotonin. Serum is thus a complex mixture of proteins, vitamins, inorganic salts, amino acids, hormones, lipids, and other components in an aqueous solution. Serum provides cells with hormonal, nutritional, and stromal (attachment) factors (36). In addition, serum proteins can act as carriers for other factors that are necessary for cell viability and function. Serum proteins also help protect cells (and product proteins) from hydrodynamic stress (impeller or bubble damage) and proteases secreted by the cultured cells or released during cell lysis. Serum acts as a sink to bind and detoxify the formulation (e.g., deleterious trace elements).

In spite of all these benefits of serum supplementation of the culture medium, the goal is usually to reduce or eliminate this supplement. Because of its biological nature, there are wide fluctuations in serum composition and ability to support cell growth/function, and consistency between batches is hard to maintain (purchases of serum should be preceded by lot testing to demonstrate its growth-supporting function and product generation with specific cells of interest).

Even small amounts of residual serum can complicate filtration, concentration via precipitation or ultrafiltration, and chromatographic processing steps. Reduction or elimination of serum therefore reduces cost of media and simplifies downstream processing. It also reduces foaming in the bioreactor, decreases the possibility of adding adventitious infectious agents to the culture, increases the consistency of medium preparations while decreasing the complexity of QC, and may help maintain expression levels of recombinant proteins or antibodies. A formulation is considered serum-free if sufficient factors have been added to the media to eliminate the need for whole serum.

Most of the classic formulations were designed to be used with serum and to support relatively low-density small-scale cultures (billion cells per liter), for example Eagle's Minimum Essential medium, Dulbecco's Modified Eagle's medium, and M199 (1). Some formulations were developed specifically for clonal growth, such as F-12 medium, which was tested to support clonal growth and serial propagation of two Chinese hamster cell lines for cultures of less than 10^5 cells/mL (1,7). Therefore, while such formulations support the proliferation of low density cultures, essential nutrients are rapidly depleted by higher-density cultures.

Table 1 compares the composition of many widely used media: DME, Ham's F-12, Coon's DME/F-12, MCDB 302, RPMI 1640, and Iscove's DME. These formulations typically require further optimization for most cell types. A proven strategy is to mix commercially available media to determine appropriate mixtures empirically (36–38). Subsequently, the analysis of consumed metabolites in cell-conditioned media can lead to the fine tuning of a formulation for a particular cell type.

Thus, many factors encourage reduction or elimination of serum supplementation: medium expense, downstream processing, the lot-to-lot variation and incomplete characterization, and the potential for introduction of contaminating agents (i.e., mycoplasma, viral pathogens, or prions). Reduction of serum from 10% to 1% can save approximately $25/L in serum costs, not including the savings associated with growth testing, storage at $-20\ ^{\circ}C$, heat inactivation and filtration, and the costs associated with additional purification steps.

The movement toward serum-free media was championed in a series of publications by Sato and his colleagues in the mid-1970s. They demonstrated that serum could be replaced by mixtures of defined and/or partially defined factors, and that the necessary replacements were qualitatively and quantitatively different, depending on the cell type (38). Most cell lines have been established with medium containing a relatively high level of serum supplementation. The key to reducing or eliminating serum is to determine which components of serum the cells are responding to or utilizing for both growth and product ex-

pression. The cells must then be adapted to an alternative environment, or these components need to be provided independent of the serum supplement. Thus, a first step toward serum-free media is to determine, by weaning (adapting) the cells from the serum-supplemented environment in small increments, the level of supplementation where culture performance is significantly reduced (i.e., changes in proliferation and product expression rates, morphology, etc.).

The next step is supplementation with serum components to restore culture performance. This is a reiterative process. Examples of these supplements include many of the proteins and other compounds listed in Tables 1–3. This process may be accelerated by the use of multivariant analysis such as the Plackett–Burmann method (39). Enrichment of the medium with amino acids which are depleted during culture has been shown to increase growth rates and productivity (40). Depending on the process, other substitutions can be beneficial, for example, lactalbumin hydrolysate, bovine colostrum, or bovine whey (41) instead of albumin, ferric citrate (42), or aurintricarboxylic acid/citrate/EDTA (43) to replace transferrin.

Because cells in reduced serum and serum-free cultures are generally more sensitive to environmental conditions, it may be necessary to optimize media for different reactor configurations and modes of operation (44). This may include the addition of protease inhibitors to counteract the effect of increased cell lysis and/or reduced concentrations of nonspecific proteins, which taken together increase the incidence of destruction of the protein product from cell-released proteases.

Some cultures at higher cell densities no longer require serum due to the presence of beneficial autocrine factors. This may be one of the major factors accounting for the success of culture strategies that involve a switch to serum-free or reduced-serum formulations for the production phase after high cell densities are achieved. Once the cells adapt to conditions more consistent with the desired process, it is useful to prepare a frozen cell bank of adapted cells. Recent reports (45,46) recommend cryopreservation media based on nonnutritional polymeric additives in place of serum.

Chemically Defined Media and Protein-Free Media

Beyond the reduction or elimination of serum is the need for chemically defined media or protein-free media. The feasibility for chemically defined media increases as the critical components necessary for many varied culture systems become defined. Several of these defined systems are utilized in large-scale cultures today. The impetus for chemically defined media partially stems from regulatory considerations because such formulations provide greater lot-to-lot consistency. Similar reasons support the use of recombinant proteins rather than those isolated from animal tissue (with potential contaminating agents). From a regulatory viewpoint, defined formulations help to demonstrate control of all aspects of the synthesis and processing of proteins destined for licensing as pharmaceutical agents. The practicality of this approach will increase as more purified factors become available (at significantly reduced prices) through recombinant technology. Even today, many hormone-supplemented, chemically defined formulations can be prepared at less expense than medium supplemented with 10% serum.

EXAMPLES OF MEDIA

Choosing the most appropriate medium for a particular cell type and desired function is a fundamental decision. There are a wide variety of expression systems for recombinant proteins, and each system has its own requirements, for example, the necessity of proline in CHO cell culture, the ability of NS0 cells to proliferate in glutamine-free media (see the section "Low Molecular Weight Nutrients"), and the requisite fatty acids for BHK proliferation. The challenge is not only to find the medium conditions that enhance cell growth and/or production, but also to choose economical components that will not bring contaminating agents to the media or interfere with purification of the product. Culture conditions for cell lines at low density and small scale typically do not support high-density, large-scale production. As medium requirements have become more stringent (from serum-containing to serum-free to chemically defined), whether from economic, scientific, or regulatory concerns, there are more and more medium options on the market, many of which are proprietary and developed for a single cell type. Therefore, a commercially available cell line is often offered with an "optimized medium formulation." Although this medium may work well for the specific cell line, it is not likely to be the best choice for other cell lines, nor will it be economical.

Two types of media are reviewed next, one for attachment-independent (hybridoma) cells and the other for attachment-dependent (endothelial) cells. Media for hematopoietic cell culture systems have recently been reviewed and are also good examples of the targeted cell population the formulation being defined by (29,47,48).

Hybridoma Media

The media requirements of hybridoma cells have been extensively studied, and many formulations and variations thereof can be found in the literature. Hybridomas are attachment-independent cell lines of mixed hematopoietic origin. Media formulations based on DME, RPMI 1640, IMDM, and DME/F-12 (among others) have been used with serum supplementation. However, the potential for copurification of serum proteins (albumin and antibodies) with hybridoma-produced antibodies is a tremendous incentive to utilize serum-free media. Serum-free formulations of basal media supplemented with insulin, transferrin, ethanolamine, and selenium (ITES) are commonly used, along with the variations and (ITES with β-mercaptoethanol (βITES) or ITES with hydrocortisone (HITES) (49,50). Amino acid supplementation has been shown to enhance cell growth and antibody production; however, full optimization of amino acid supplementation is specific to a particular cell line (40,51–53). A fed-batch bioreactor study that aimed to maintain nutrient concentrations with a process control strategy based on stoichiometric consumption rates and total cell number has lead

to impressive improvements in cell density, culture life, and antibody titer for a specific hybridoma cell line (54). Because ammonia is toxic to hybridoma cultures (55), the use of glutamine-based dipeptides as a substitute for glutamine is being investigated. The dipeptide is stable in storage, preventing the spontaneous degradation of glutamine to ammonia; however during culture glutamine is released by hydrolysis of the dipeptide. In selected cell lines, higher cell yields are observed coincident with a significant reduction of accumulated ammonia (56,57). In a move toward more defined and ultimately protein-free culture media, the use of ferric citrate (50–500 μM) has been demonstrated to be a substitute for transferrin (42,58).

Endothelial Cell Media

Endothelial cells, the cells that line the blood vessels, are a rich source of enzymes of pharmaceutical interest and are used as components of in vitro assays for cardiovascular, inflammation, and oncology research. They are attachment dependent and are typically utilized as primary or low-passage-number (P5–P20) cultures, with a strong dependence on proper culture conditions for retaining function, morphology, and life span. Optimized culture media are specific for origin of the cells, including both the species and the location in the vasculature. Human endothelial cells are the most difficult to grow in culture, but perhaps the most important for pharmaceutical and biotechnology applications. Human umbilical vein endothelial cells (59,60) and human dermal vascular endothelial cells (61) can be cultured in similar medium. A basal medium of M199, DME or MCDB 131 (62–64) is typically used, supplemented with fetal bovine and/or human serum in the range of 2–50%, with a typical value of 10–20%. Because of the primary nature of typical endothelial cultures, it is not unusual, although not completely necessary either, to add both antibiotics (penicillin, streptomycin, neomycin, and/or gentamicin) and antifungal agents (i.e., Fungizone®). Endothelial cells require glutamine, which should be added to the medium no more than one week before feeding to the cells.

Primary cultures at a proper seeding density typically grow without additional growth factors; however, subcultured cells require a mitogen such as fibroblast growth factor (FGF) or endothelial cell growth factor (ECGF), which is potentiated by the glycosaminoglycan heparin (65). The advent of commercially available endothelial cell cultures has led to the availability of commercially prepared complete media. These media provide the advantage of being optimized for individual types of cells; however, a potential disadvantage is cost. Endothelial cells in vivo grow on a basement membrane that contains laminin and type IV collagen. In vitro endothelial cells benefit from treatment of the growing surfaces with an attachment factor such as collagen, gelatin (denatured collagen), fibronectin, or attachment peptides (i.e., the RGD sequence).

MEDIA PREPARATION AND OPTIMIZATION

There are three major forms in which medium can be supplied: bulk powder, media concentrates, and bulk liquid. Each has its own advantages and disadvantages, as described later. The choice of powder, concentrate, or liquid will vary depending on convenience, cost, storage space (or lack thereof), and availability of a reliable water supply. Whatever the starting configuration, quality control measures are required to ensure lot-to-lot consistency and conformity to standards. Some quality control techniques (i.e., amino acid analysis by HPLC) can also be used for media optimization.

Media Preparation from Powder

Typically the most common medium used for large-scale animal cell culture is that prepared in the end-user laboratory from a powdered formulation. The advantages are cost (both in the cost of the media and the required amount of storage space at 4 °C), increased shelf life of powdered media versus liquid media (years versus months, respectively), flexibility of preparation (i.e., it is relatively easy to modify the formulation at the time media is prepared), large lot sizes for custom and standard formulations, control over the type of container used for the final, prepared media (for ease of use in filling and/or perfusion of reactors), and customized quality control (samples can be taken at any step or every step during the process and run in as few or as many assays as required).

The requirements for media preparation from powder are good source of water (see section “Water Quality”), adequate storage at 4 °C for the bulk powder and the reconstituted media, storage at −20 °C or −70 °C for other media components, mixing vessels and agitation or mixing equipment, sterilization equipment (i.e., filters and pumps for submicron filtration, and autoclave or steam-in-place for media holding tanks), and equipment for quality control (i.e., pH meter or blood gas analyzer for pH, vapor pressure osmometer for osmolarity, scale or volumetric containers for volume). Bulk powder (after checking for any discoloration or unusual odor) is added to water (approximately 90% of final volume) and mixed thoroughly. Other ingredients are added one at a time to ensure complete dissolution. Volume and pH adjustments are made and the medium filtered. Quality control checks on pH, osmolarity, and volume should be made prior to sterile filtration, with the caveat that pH usually increases a few tenths during filtration. In addition, a sterility check from the final material pumped through the filter should be performed (37 °C overnight). After the quality control checks are passed, the reconstituted medium can be released for use. Over time glutamine-containing media will build up significant amounts of ammonia, even at 4 °C (see the section “Low Molecular Weight Nutrients”). As an alternative, medium can be prepared glutamine-free, with glutamine supplemented just prior to use or prepared with glutamine-containing dipeptides to extend shelf life.

Powdered medium is produced by the ball milling process, in which ingredients in powdered form are added to a mixing unit along with ceramic particles or balls. The drum-shaped mixing unit is rotated, which results in mixing of the contents. This process has its drawbacks, notably the ball mill batch size before blending is limited to the size of the mixing drum, heat is generated during ball mill-

ing, which can degrade heat sensitive components, and the ceramic balls can chip and break into the medium, resulting in ball mill particles in the packaged product. BioWhittaker has recently introduced a micronization process in which the powdered media is continuously mixed without the use of ceramics. They claim increased lot size to a 6,500-kg blended lot, elimination of the possibility of ceramic chips in the powdered media, quicker production time, and the potential for a completely enclosed automated process.

Media Concentrates

Life Technologies, Inc. has introduced media concentrates, which offer some of the convenience of liquid media while retaining some of the flexibility and cost savings of powdered media. As a substitute for bulk powder, concentrates require less mixing and filtering but similar QC strategies. The advantages lie in reducing the amount of time and effort spent weighing out individual ingredients, increasing the shelf life over complete bulk liquid media (concentrates are segregated into 3–4 component mixtures, each of which is more stable than the complete liquid media itself), and the potential for automated mixing (66). Concentrates, by definition, also require less storage space than bulk liquid media. However, using concentrates to reconstitute full-volume media is not their only or even their best use. Concentrates can be utilized as perfusion supplements, adding back key components to high-density reactors without greatly effecting the volume (67,68). This is ultimately done to increase the productivity of the reactor. However, another reason to add supplements is to provide a minimally required level of nutrients to prevent apoptosis (programmed cell death), which has been shown to occur when even a single essential nutrient is limiting (69,70). Cell death, whether by apoptosis, hydrodynamic damage, or other mechanisms, leads to increased levels of cell-released proteases as well as a reduction in cell density. Nutrient feeding by concentrates can occur along with $1\times$ media perfusion, which, in addition to adding nutrients, dilutes waste products such as lactate and ammonia.

Bulk Liquid Media

Bulk liquid media is the most expensive form of media but the least time-consuming for the end user. The media is mixed, filtered, and QC checked at the vendor, reducing or eliminating the need for the end-user to provide a water supply, mixing and filtration equipment, and QC instruments. However, the extra labor and processing adds to the cost of the media. In addition, shipping costs increase. Another cost to the end-user is storage in the appropriate temperature-controlled environment. A drawback of bulk liquid media has been its availability in bottles (0.5–1 L), which require transfer to other vessels for filling or perfusion of the bioreactors. This concern has been addressed by the sale of liquid in drums and large bags (50–100 L), which can be equipped with custom fittings. However, special handling equipment may be necessary to move these large containers around the end-user facility.

Quality Assurance

The goal of quality assurance is to provide consistency between lots and to provide a level of confidence that the media meets the necessary conditions. Vendor-supplied certificates of analysis for media in powder, concentrate, or liquid form provide a list of components and their mass or concentration. Simple tests, such as pH, osmolarity, sterility, and volume, can provide additional information about reconstituted media and should be routine. Other tests can provide additional information such as ammonia and glutamine (especially for media close to the expiration date) levels, amino acid analysis (for new lots of media to verify formulation), and endotoxin levels (for any media, but especially those containing albumin, which binds endotoxin). Individual components of the media, especially serum and growth factors, should be lot tested with the appropriate cell lines for their ability to support growth.

Optimization

Media for individual cell lines or groups of cell lines can be optimized for cell growth, specific production rates, attachment (in)dependency, cost of media, or other factors critical to a project. This optimization is especially important for larger-scale and/or longer-term projects.

There are two major approaches to optimization of fed-batch reactor schemes: bottom-up and top-down (67). The bottom-up approach refers to adding back components that are determined to be depleted. This approach can range from a simple addition of glutamine or glucose feed, or be based on an extensive study of cellular metabolism (71). The analysis required then can range form a single measurement in a metabolite analyzer (such as a YSI glucose or glutamine analyzer, or a Kodak Biolyzer) to complex analytical methods (i.e., HPLC) for detecting trace compounds. The specificity of a particular media supplement has the advantage of enhancing that feature that is most important for a particular production project, but this specificity also has the disadvantage of not translating to other cell lines or optimization conditions. The optimized formulation may be valid for only a narrow range of cell density and is likely to be cell-line specific.

The top-down approach uses complete media supplementation with either $1\times$ media or concentrates. This is a simple approach that does not require elaborate metabolic studies, but does increase the cost of the actual media supplement, a factor that must be weighed against the time savings in development of a bottom-up type of supplement. Top-down strategies require the identification of at least one key nutrient (i.e., glucose), which is monitored to control the supplementation rate through a stoichiometric or heuristic model.

POSTTRANSLATIONAL MODIFICATION AND MEDIUM COMPOSITION

Proteins, whether produced in vivo or in vitro, undergo posttranslational modifications that ultimately effect the biological properties of the protein. Although some posttranslational modifications occur extracellularly in vivo,

(i.e., activation of the complement system or proteolysis of prothrombin to thrombin during blood coagulation), many posttranslational modifications occur prior to release of the protein from the cell. These modifications include hydroxylation, phosphorylation, sulfation, adenylation, carboxylation, acetylation, methylation, and glycosylation (72). Of these modifications, glycosylation has been most extensively studied in cell culture because the ability to properly glycosylate proteins is an advantage of mammalian recombinant expression over prokaryotic systems. Altered glycosylation patterns may effect clearance rates, specific activity, further processing of the protein, solubility, and immunogenicity. The cell type is a major contributing factor to glycosylation patterns of proteins made in cell culture; however, there is evidence that environmental factors, including cell culture media, can influence glycosylation.

Glycosylation

Oligosaccharides attached to asparagine residues are referred to as N-linked, whereas those attached to serine or threonine hydroxyls are O-linked. More studies have been conducted on N-linked oligosaccharides due to their more complex nature. Glucose starvation has been shown to reduce the incidence of glycosylation at asparagine residues and lead to aberrant oligosaccharide structures (73,74). An individual assessment of the result of such deviant glycosylation on the recombinant protein properties is warranted, and if necessary the cell culture medium should be adjusted to ameliorate the problem. Other cell culture components that have been reported to disrupt glycosylation include hormones, DMSO, acidotropic amines, EDTA, HEPES buffer, the antibiotic tunicamycin, and the ionophore monensin (74). If it is determined that important protein properties are altered by the media composition, media optimization for production of the appropriate glycoform is required.

BIBLIOGRAPHY

1. R.G. Ham and W.L. McKeehan, in S.P. Colowick and N.O. Kaplan eds., *Methods in Enzymology*, vol. 58, Academic Press, Orlando, Fl., 1979, pp. 44–93.
2. R.G. Ham in *Handbook of Experimental Pharmacology*, Springer-Verlag, New York, 1980, pp. 13–88.
3. M.J. Keen and C. Hale, *Cytotechnology* **18**, 207–217 (1996).
4. H. Eagle, *Science* **122**, 501–504 (1955).
5. H. Eagle, *Science* **129**, 432–437 (1959).
6. J.D. Smith, G. Freeman, M. Vogt, and R. Dulbecco, *Virology* **12**, 185–192 (1960).
7. R.G. Ham, *Proc. Natl. Acad. Sci. U.S.A.* **53**, 288–293 (1965).
8. W.L. McKeehan, K.A. McKeehan, S.L. Hammond, and R.G. Ham, *In Vitro* **13**(7), 399–416 (1977).
9. R. Cote ed., *ATCC Media Handbook*, American Type Culture Collection, Rockville, Md., 1984, pp. 81–82.
10. N.N. Iscove and F. Melchers, *J. Exp. Med.* **147**, 923–33 (1978).
11. M. Butler, in W.G. Thilly ed., *Mammalian Cell Technology*, Butterworths, Boston, Mass., 1986, pp. 91–108.
12. K. Higuchi and R.C. Robinson, *In Vitro* **9**(2), 114–121 (1973).
13. C. Waymouth, in C. Waymouth, R.G. Ham, and P.J. Chapple eds., *The Growth Requirements of Vertebrate Cells In Vitro*, Cambridge University Press, New York, 1981, pp. 105–115.
14. R.R. Zombola, R.C. Bearse, and P.A. Kitos, *J. Cell Physiol.* **101**, 57–66 (1979).
15. L.J. Reitzer, B.M. Wice, and D. Kennell, *J. Biol. Chem.* **264**(8), 2669–2676 (1979).
16. A. Leibovitz, *Am. J. Hyg.* **78**, 173–80 (1963).
17. D. Barngrover, J. Thomas, and W.G. Thilly, *J. Cell Sci.* **78**, 173–189 (1985).
18. H. Senoo, Y. Tsukada, T. Sato, R. Hata, *Cell Biol. Int. Rep.* **13**, 197–206 (1989).
19. R.A. Wolfe and G.H. Sato, in D. Sirbasku ed., *Growth of Cells in Hormonally Defined Media*, Cold Spring Harbor Press, New York, 1982, pp. 1075–1088.
20. C.R. Bebbington, G. Renner, S. Thomson, D. King, D. Abrams, and G.T. Yarranton, *Bio/Technology* **10**, 169–175 (1992).
21. M.J. Keen and T.W. Steward, *Cytotechnology* **17**, 203–211 (1996).
22. R.G. Harrison, *Anat. Record* **2**(9), 385–410 (1908).
23. M.R. Lewis and W.H. Lewis, *Anat. Record* **5**(6), 277–293 (1912).
24. A. Carrel, *J. Exp. Med.* **17**, 14–19 (1913).
25. "Mad Cows and Englishmen," *The Economist*, March 30, 1996, pp. 25–27.
26. W.L. McKeehan, D.P. Genereux, and R.G. Ham, *Biochem. Biophys. Res. Comm.* **80**(4), 1013–1021 (1978).
27. W.J. Bettger and W.L. McKeehan, *Physiol. Rev.* **66**(1), 1–35 (1976).
28. H. Polet and H. Spieker-Polet, *J. Exp. Med.* **142**, 949–55 (1975).
29. M.R. Koller and B.O. Paulson, *Biotechnol. Bioeng.* **42**, 909–930 (1993).
30. K.M. Yamada, *J. Biol. Chem.* **266**(20), 12809–12812 (1991).
31. V.B. Mountcastle ed., *Medical Physiology*, 14th ed., C.V. Mosby, St. Louis, 1980.
32. D.W. Murhammer and C.F. Goochee, *Biotechnol. Prog.* **6**, 142–148 (1990).
33. S.I. Grammatikos, P.V. Subbaiah, T.A. Victor, and W.M. Miller, *Cytotechnology* **15**, 31–50 (1994).
34. A.B.R. Thomson and G. Wild, *J. Nutr. Biochem.* **6**, 240–245 (1995).
35. L.E. Baker and A.H. Ebeling, *J. Exp. Med.* **69**, 365–378 (1939).
36. D. Barnes and G. Sato, *Cell* **22**, 649–655 (1980).
37. R.A. Wolfe, G.H. Sato, and D.B. McClure *J. Cell Biol.* **87**, 434–441 (1980).
38. J. Bottenstein, I. Hayashi, S. Hutchings, H. Masui, J. Mather, D.B. McClure, S. Ohasa, A. Rizzino, G. Sato, G. Serrero, R. Wolfe, and R Wu, in W.B. Jakoby and I.H. Pastan eds., *Methods in Enzymology*, 45, Academic, New York, 1979, pp. 94–109.
39. P.M.L. Castro, P.M. Hayter, A.P. Ison, and A.T. Bull, *Appl. Microbiol. Biotechnol.* **38**, 84–90 (1992).
40. M.J. Keen and C. Hale, *Cytotechnology* **18**, 207–217 (1996).
41. J. Capiaumont, S. Ostrovidov, C. Legrand, F. Belleville, and P. Nabet, *In Vitro Cell. Dev. Biol. Animal* **32**, 8–12 (1996).
42. F. Franek and J. Dolnikova, *Cytotechnology* **7**, 33–38 (1991).
43. K. Bertheussen, *Cytotechnology* **11**, 219–231 (1993).
44. H. Buntemeyer, D. Lutkemeyer, and J. Lehmann, *Cytotechnology* **5**, 57–67 (1991).

45. O.-W. Merten, S. Petres, and E. Couve, *Biologicals* **23**, 185–189 (1995).
46. C.D. Severson, *In Vitro Cell. Dev. Biol. Animal* **32**, 321 (1996).
47. C.E. Sandstrom, W.M. Miller, and E.T. Papoutsakis, *Biotechnol. Bioeng.* **43**, 706–733 (1994).
48. T.A. McAdams, W.M. Miller, and E.T. Papoutsakis, *TIBTECH* **14**, 341–349 (1996).
49. M. Butler and H. Jenkins, *J. Biotechnol.* **12**, 97–110 (1989).
50. H. Murakami, H. Masui, G.H. Sato, N. Sueoka, T.P. Chow, and T. Kano-Sueoka, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1158–1162 (1982).
51. D. Duval, C. Demangel, K. Munier-Jolain, S. Miossec, and I. Geahel, *Biotechnol. Bioeng.* **38**, 561–570 (1991).
52. F. Franek, *Biotechnol. Bioeng.* **45**, 86–90 (1995).
53. T.S. Stoll, K. Muhlethaler, U. von Tockar, and I.W. Marison, *J. Biotechnol.* **45**, 111–123 (1996).
54. L. Xie and D.I.C. Wang, *Biotechnol. Bioeng.* **51**, 725–729 (1996).
55. I. Ludemann, R. Portner, and H. Markl, *Cytotechnology* **14**, 11–20 (1994).
56. M. Butler and A. Christie, *Cytotechnology* **15**, 87–94 (1994).
57. A. Christie and M. Butler, *J. Biotechnol.* **37**, 277–290 (1994).
58. Z. Chen, Y. Ke, and Y. Chen, *Cytotechnology* **11**, 169–174 (1993).
59. M.A. Gimbrone, *Prog. Hemostasis Thromb.* **3**, 1–28 (1976).
60. E.A. Jaffe, R.L. Nachman, G.C. Becker, and C.R. Minick, *J. Clin. Invest.* **52**, 2745–2756 (1973).
61. R.M. Marks, M. Czernieki, and R. Penny, *In Vitro Cell. Devel. Biol.* **21**(11) 627–635 (1985).
62. J.B. Warren, in J.B. Warren ed., *The Endothelium: An Introduction to Current Research*, Wiley-Liss, New York, 1990, pp. 263–272.
63. P.A. D'Amore and S.J. Braunhut, in U.S. Ryan ed., *Endothelial Cells*, CRC Press, Boca Raton, Fla., 1988, pp. 13–36.
64. A. Knedler and R.G. Ham, *In Vitro Cell. Devel. Biol.* **23**(7) 481–491 (1987).
65. T. Maciag and W.H. Burgess, in U.S. Ryan ed., *Endothelial Cells*, CRC Press, Boca Raton, Fla., 1988, pp. 3–11.
66. G. Roth, J.M. Kubiak, J.F. Long, and G.M. Schoofs, *BioPharm*, **8**, 31–35 (1995).
67. T.A. Bibila, C.S. Ranucci, K. Glazomitsky, B.C. Buckland, and J.G. Aunins, *Biotechnol. Prog.* **10**, 87–96 (1994).
68. R. Fike, J. Kubiak, P. Price, and D. Jayme, *BioPharm* **6**, 49–54 (1993).
69. F. Franek and K. Chladkova-Sramkova, *Cytotechnology*, **18**, 113–117 (1995).
70. R.P. Singh, M. Al-Rubeai, C.D. Gregory, and A.N. Emery, *Biotechnol. Bioeng.* **44**, 720–726 (1994).
71. L. Xie and D.I.C. Wang, *Biotechnol. Bioeng.* **43**, 1175–1189 (1994).
72. K.-K. Han and A. Martinage, *Int. J. Biochem.* **24**, 19–28 (1991).
73. H. Tachibana, K. Jiyoun, K. Taniguchi, Y. Usho, K. Teruya, K. Osada, Y. Inoue, S. Shirahata, and H. Murakami., *In Vitro Cell. Dev. Biol. Animal* **32**, 178–183 (1996).
74. C.F. Goochee and T. Monica, *Bio/Technology*, **8**, 421–427 (1990).

See also MAMMALIAN CELL BIOREACTORS; MAMMALIAN CELL CULTURE REACTORS, SCALE-UP.

MEDIUM FORMULATION AND DESIGN, *E. COLI* AND *BACILLUS* SPP.

ROSALIE COTE
ROBERT L. GHERNA
American Type Culture Collection
Rockville, Maryland

KEY WORDS

Antibiotics
Catabolite repression
Chemically defined media
Competence
Diauxic growth
Enrichment media
Freezing
Isolation media
Lyophilization
Sporulation
Undefined (complex) media

OUTLINE

INTRODUCTION

Escherichia coli and *Bacillus subtilis* (later to include other species of the genus *Bacillus*) were the first microbial tools used in studies of biochemical pathways and genetics because of their undemanding needs for cultivation as well as their general lack of pathogenicity. Consequently, more is known about the workings of the metabolic processes of these organisms than all others, and this knowledge has

been successfully applied in the era of biotechnology. Whether by human-directed and human-enhanced natural strain selection or by genetic manipulation, these prokaryotic organisms have provided a wealth of molecular information that has much accelerated biotechnology's scientific and industrial growth.

The purpose of this article is not to belabor basic nutritional considerations required to propagate these organisms: most wild-type strains are so metabolically versatile that they will grow in a simple medium containing only 1% peptone. Microbial growth alone, however, is rarely the ultimate goal in scientific pursuit. Strain isolation, enhancement, genetic stability, performance, and reproducibility are all entangled aspects of cultivation for these as well as other microorganisms. Unfortunately, there are no universal road maps to guide a biotechnologist to success in manipulating a microorganism; the best solutions are often strain specific as well as need specific. Thus, this article highlights broad discussions of media considerations involved with the growth and manipulation of these metabolically incredible bacteria.

NUTRITIONAL REQUIREMENTS

E. coli and *Bacillus* spp. have complete anabolic and catabolic biochemical pathways. This metabolic complexity allows the cultivation of these organisms with relatively simple media containing an organic source of energy, the nutritionally essential minerals (N, C, H, O, PO_4^{2-}, SO_4^{2-}), and trace elements (Ca^{2+}, K^+, Na^+, Mg^{2+}, Mn^{2+}, Fe^{2+}, Zn^{2+}, Cu^{2+}, Co^{2+}, Mo^{2+}, BO_3^{3-}). Auxotrophies for amino acids, vitamins, or other growth factors do occur for some strains whether deliberately engineered into a genome or naturally occurring.

These bacterial strains are organotrophs and can utilize heterogenous organic compounds as sources of carbon and energy including alcohols, amino acids, carbohydrates, fatty acids, hydrocarbons, organic acids, and polysaccharides. *E. coli* is facultative and can respire aerobically or anaerobically depending upon the carbon–energy source and other growth parameters. However, secondary metabolites produced by the organism vary tremendously between aerobic and anaerobic conditions. With the exception of the anaerobic species *Bacillus infernus* (DSM 10276) (1), the genus *Bacillus* is strictly aerobic or facultative. Morphological and physiological characteristics can be influenced by media conditions with these metabolically responsive bacteria, and care should be taken to standardize media conditions as well as other growth parameters to ensure reproducibility in research or production.

GENERALIZED MEDIA

Bacteriological media may be either defined or undefined with respect to chemical composition. The type of medium selected is entirely dependent on the ultimate goal of the cultivation process.

Chemically Defined Media

In a chemically defined formulation, each component is discretely identifiable as to chemical identity and concentration. The mineral salt-based formulations used in studies of biochemical pathways or gene function are examples of chemically defined media. Auxotrophs can be screened on this type of mineral medium when it is amended with appropriate concentrations of amino acids, vitamins, or other other growth factors (2). Such formulations can also be modified for selectivity by the incorporation of specific carbon or nitrogen sources, inhibitory compounds, or adjustment to physiochemical parameters (e.g., pH, salinity). Chemically defined media are also employed, where possible, in the large-scale production of high-value recombinant DNA-derived biopharmaceuticals. Because product yield is often low for such types of fermentation protein and peptide products, isolation and purification procedures are rigorous. Use of chemically defined media eliminates interfering proteinaceous matter present in the peptones and extracts of complex formulations that further complicates product recovery.

Chemically defined media are not particularly forgiving in use if poorly designed and balanced or poorly prepared. For example, growth inhibition caused by amino acid antagonism is a well-studied phenomenon observed under chemically defined, mineral media conditions. In particular, valine inhibition of *E. coli* K12 has been studied in detail (3). Chemical imbalance, such as an excess of phosphates used as medium buffering agents, can cause deleterious binding and precipitation of trace element divalent cations (e.g., Ca^{2+}, Fe^{2+}, Mg^{2+}) necessary as enzyme cofactors for key metabolic functions. Use of a trace metal in a nonmetabolically active valence or in medium conditions that alter metal speciation (4) can also lead to chemical imbalance in a mineral medium. Sometimes, trace metal and other growth factor requirements in mineral media can be spared. Iron limitation will induce the production of siderophores (iron-binding transport molecules) in *Bacillus megaterium* (ATCC 19213) as well as other bacteria (5) to increase cellular uptake of the required ion. The requirement for Mn^{2+} for sporulation in *B. subtilis* can be bypassed in mineral media by properly balancing the ratio of carbon sources (glucose to malate) entering the metabolic pathway above and below the manganese-requiring enzyme phosphoglycerate phosphomutase (PGA mutase) (6). The vitamin B_{12} requirement in *E. coli* (ATCC 10799; NCIB 8134) can be spared by high levels of methionine (20 μg/mL) added to a mineral growth medium (7).

For nutritionally fastidious microorganisms, such as many animal and human pathogens, development of a good chemically defined medium with respect to the identification of key nutrients, as well as the important balancing of their concentrations, can be laborious and most often necessitates a lengthy list of growth factors. Fortunately, wild-type *E. coli* and *Bacillus* spp. can generally be cultivated with rather simple chemically defined formulations such as M9 medium (see final section, "Media Formulations") composed of a few mineral salts and an organic carbon source. With M9, those essential trace elements and microelements not deliberately represented in the formulation are spuriously supplied as poorly defined contaminants in the distilled water or from the glassware used to prepare the medium. Neidhardt et al. (8) developed a more nutritionally complete defined mineral medium

suitable for use in rigorous biochemical studies of the enterobacteria.

Undefined (Complex) Media

Undefined media, also called complex media, provide many of the nutritional elements as mixtures of protein hydrolysates (e.g., peptones) or extracts (e.g., yeast extract) in which the exact chemical identity and concentration of essential nutrients, trace elements, and growth factors are not known. The demands of the biotechnology industry have influenced the manufacturers of these raw media components to provide detailed lot analyses of hydrolysates and extracts (9) but, for critical needs, any new lot of peptone or extract must be qualified by performance testing before use.

Complex media formulations, rather than those chemically defined, will generally give greater cell yield for most strains of *E. coli* and *Bacillus* spp. Commonly available laboratory media such as soy bean casein digest broth manufactured under different trade names (e.g., Trypticase Soy Broth [BBL]; Tryptic Soy Broth [Difco]) will support vigorous growth for most strains with growth optima close to neutral pH. For general purposes, a combination of 1.0% peptone, 0.5% yeast extract, and 0.2% utilizable carbon source will adequately satisfy the nutritional needs for these heterotrophic organisms as well as for many auxotropic mutants. Complex nitrogen and carbon substrates (e.g., cottonseed meal, molasses) are frequently used in large-scale fermentation media for the microbial production of enzymes and antibiotics; however, lot-to-lot variation in these raw materials can be significant issues in maintaining consistency in media performance.

Although any number of generalized formulations, either chemically defined or complex, will support growth of these microorganisms, strain-specific characteristics, product desirability, enzyme induction or increased expression, or other investigational needs will necessitate optimization of a medium. Greasham and Inamine (10) highlight many of the considerations and decisions that must be made with respect to optimization. The design of media for *E. coli* and *Bacillus* spp. must also take into account a variety of media-induced phenomena that can affect growth or other biochemical activities of the organisms as well as the quantity and types of secondary metabolites produced.

Catabolite Repression

Catabolite repression of carbon or nitrogen metabolism is frequently encountered in both *E. coli* and *Bacillus* spp., although the exact biochemical mechanisms differ between genera. In general, a readily metabolized substrate (such as glucose) can repress the activity of enzymes acting on similar substrates, thereby affecting the nature of secondary metabolite production. For example, α-amylase activity in *B. subtilis*, normally expressed when the organism is grown on starch, is repressed when glucose is present; bacitracin synthesis in *Bacillus licheniformis* is repressed by glucose but unaffected by citrate (11). In *B. subtilis*, the expression of the histidine-inducible, histidine-degradative pathway (*hut* operon) is repressed by rapidly metabolized carbon compounds including glucose, glycerol, and malate (12) as well as by amino acid mixtures (13). Derepression of the *hut* operon occurs when histidine is added to the medium (14).

Diauxic Growth

Under nonlimiting media conditions with mixed carbon–energy sources, bacteria can exhibit diauxic growth in which substrates are preferentially consumed on the basis of ease of assimilation and catabolism while utilization of other substrates is repressed (15). The metabolic shift from one substrate to another is accompanied by a lag in growth. However, using carbon-limited chemostat cultures of *E. coli* ML30 (DSM 1329), Lendemann et al. (16) demonstrated that this phenomenon does not occur under the low nutrient (e.g., μg/L) concentrations frequently encountered in natural environments. Using combinations of as many as six sugars (glucose, galactose, maltose, ribose, arabinose, and fructose), the authors demonstrated that, regardless of the mixture combination or dilution rate, *E. coli* simultaneously utilized all the sugars.

Effects of Nutrient Fluctuations

With media designed to maximize culture growth, care must be taken to ensure that media parameters do not deteriorate in terms of either nutrient depletion or accumulation of inhibitory metabolites. While the need for maintenance of an overall nitrogen–carbon balance of a medium is generally recognized, the effects of minor nutrient depletion can be subtle, but important. For example, under conditions of methionine deprivation, *E. coli* will substitute norleucine in tRNA (A.S. Lubinecki, personal communication, 1994), a potentially serious anomaly in the production of a biopharmaceutical peptide. Phosphate limitation has been shown to cause the loss of the plasmid pBR322 from *E. coli* (17). With respect to metabolic byproducts, rapidly growing aerobic cultures of *E. coli* accumulate acetate to growth-inhibiting levels. Acetate formation in *E. coli* has been shown to limit expression of recombinant growth hormone (18) as well as to limit induction of human interleukin 2 expression in fermentor cultures (19). Kleman and Strohl (20) have studied acetate metabolism in *E. coli* under chemically defined nutritional conditions.

ENRICHMENT AND ISOLATION MEDIA

Obtaining a pure culture of a specific type of bacterium, particularly for the general heterotrophs discussed here, often requires both general and selective enrichment. When the organism of interest is present only in small numbers or as a nonrobust population of the total microbial flora, an initial general enrichment of the sample may help to enhance the presence of the desired organism. With *E. coli* and *Bacillus* spp., almost any general bacteriological growth medium (e.g., tryptic soy broth) can often support enrichment of the desired organism as well as proliferation of any other general heterotroph in a mixed sample. Under such conditions, selective enrichment tech-

niques can help to increase the numbers of the organism or phenotype of interest while suppressing the growth of nondesirable cells in a mixed population.

In some instances, raw samples may include physiologically stressed organisms that might be best recovered with a slightly less nutritious medium such as R2YA (21), or with HTYE (see final section) developed for cultivation of a broad spectrum of "heterotrophically delicate" bacteria (e.g., *Caulobacter, Cytophaga*) and stressed cells, while also being supportive to the growth of general heterotrophs. Clavero and Beuchat (22) determined that heat-stressed *E. coli* 0157:H7 was best recovered with tryptic soy agar rather than with modified sorbitol MacConkey agar.

If strain selection is intent upon a particular phenotypic characteristic (e.g., utilization of heptane as a sole carbon source) rather than upon taxonomic concern, isolation may be done directly by inoculating sample material into nutritionally complete defined media or into dilute complex media supplemented with a substrate compound of interest. Cultivation of sample material under a physiological parameter (alkalophilic or acidophilic conditions) of interest is also a selective pressure. The judicious incorporation of antibiotics in a medium will also keep many of the unwanted organisms in a mixed culture to low levels. In general, prolonged incubation or multiple transfers of any microbial growth that develop in the selection media will eventually give rise to a culture expressing the desired characteristic. Continuous fermentation methods can be used to the same effect to select for a phenotype from a heterogenous cell population, but often the resulting dominant or selected population phenotype is not necessarily genetically stable once selection factors are relaxed (23).

Isolation is accomplished from an enrichment culture generally by spatially separating the cell of interest from other cells in the enrichment culture. This is most conveniently done by streaking aliquots on agar or other solid gel plates with the aim to have discretely separated bacterial colonies to pick, or by serial dilution in liquid media with the hope that the desired cells, being dominant in the population, will eventually be isolated by increasing dilutions of the original enrichment culture.

Because *E. coli* is an indicator organisms of fecal contamination, many versions of media have been developed to quickly isolate and identify the bacteria in potable and nonpotable waters, food, etc. This organism can be isolated using a variety of commercially available selective media (e.g., brilliant green broth, EC broth, lauryl tryptose broth, MacConkey agar). These media contain bile salts (0.15–0.5%) to inhibit the growth of competative gram-positive as well as some gram-negative organisms in a sample mixture. Dyes such as crystal violet or brilliant green are included in some of these formulations to further selectivity for coliforms. The inclusion of lactose (0.5–1.0%) in these media as a fermentable carbon source helps to differentiate *E. coli* from other bile- and dye-resistant organisms. Lactose fermentation is detected either by the presence of gas in an inverted Durham vial contained in a test tube of liquid medium or, with agar media, by differential colony color caused by acid production from lactose and a chemical pH indicator (such as neutral red) included in the medium. Based on the work of Feng and Hartman (24) and Robison (25), the accurate detection of *E. coli* with the classical selective media can be significantly increased by the inclusion of 4-methylumbelliferyl-β-D-glucuronide (MUG) at 50 mg/L. Most *E. coli* strains are β-glucuronidase positive and cleave the substrate to free the methylumbelliferyl moiety, which can then be detected by fluorescence under 365-nm ultraviolet light. Similarly, 5-bromo-4-chloro-3-indolyl-β-D-glucuronide (X-Gluc) can be incorporated as a chromogenic substrate targeting *E. coli*. Gaudet et al. (26) recently evaluated some of these selective media for their usefulness in enumerating *E. coli* from nonpotable waters and concluded that lauryl tryptose agar supplemented with MUG or X-Gluc gave the best sample recovery.

Isolation of spore-forming bacilli can frequently be accomplished by allowing culture growth in a generalized growth medium to proceed until nutrient depletion induces sporulation (see next section). Deliberate limitation of a key nutrient such as a carbon or nitrogen source can often quicken sporulation. Media selectivity can be introduced for pH parameters, salinity, substrate utilization, etc. as desired. As spores of *Bacillus* spp. are among the most heat-resistant of endospore-forming bacteria, heat can be used as a selective factor to eliminate vegetative cells as well as weakly thermal-resistant spores from a sporulated culture. Exposure of the culture to 70–80 °C for a minimum of 10 min is one general method for *Bacillus* spp. selection. Harvest and purification of the spores from remaining vegetative cells and other debris should be performed with gentle procedures, either enzymatically or by density gradient, because strong mechanical methods such as high-speed centrifugation can damage the spores. A single spore isolation can then be attempted by germinating aliquots of the spore sample using agar streak plates.

SPORULATION MEDIA FOR *BACILLUS* SPP.

Bacillus spp. tend to form spores in any suitable growth medium providing a sufficient incubation period. With complex media formulations containing peptones and yeast extract, the incubation time required for depletion of nutrients and stimulation of sporulation is frequently manipulated by a simple reduction of total nutrients as demonstrated with the laboratory medium, sporulation broth (see final section). Increased rates of sporulation can be achieved with the addition of appropriate cations (Ca^{2+}, Mg^{2+}, Mn^{2+}) at assimilable trace metal (millimolar, mM) concentrations (27). Classical media formulations include these trace elements as poorly defined organic chelates supplied in crude soil extracts as in SEY medium (see final section). Many strains of *Bacillus cereus, B. coagulans, B. megaterium, B. subtilis,* and others sporulate well with media based on potato infusions supplemented with yeast extract and 2.5 g/L glucose (28). The pH of these media should be adjusted to approximately 7.2 for pH-neutral strains; sporulation media for alkalophilic and acidophilic strains should be adjusted to a pH supporting optimal growth. If the optimum pH for the sample is known, adjust the pH of the sporulation medium to that of the habitat where the sample was collected. Highly aerobic strains

may require aeration in liquid cultures to obtain a high yield of spores. Elevated glucose levels in a medium during the stationary phase of culture growth can repress sporulation.

As with all microbial processes, the optimal medium and growth conditions for sporulation is strain specific. Similarly, sporulation conditions for optimum chemical resistance may not be the same as those of temperature resistance (29). One should should evaluate more than one medium to find conditions that provide both a high degree of sporulation as well as adequate resistance. Wild-type *B. subtilis* can sporulate in chemically defined media with glucose as a sole carbon source. In studies on transduction with *Bacillus thuringiensis*, LeCadet (30) used media containing peptones to satisfy nitrogen and amino acid requirements, inorganic salts to supply cations, glucose as a carbon and energy source, and potassium phosphate as a buffer for a high yield of heat-stable spores. Setlow and Kornberg (31) noted that small additions of trace metals and glucose to nutrient agar greatly increased hydrogen peroxide resistance of *B. subtilis* spores. Although spore production in *B. stearothermophilus* and *B. thuringiensis* is readily achieved at commercial scale, successful in vitro sporulation of the insect pathogen *B. popilliae* remains elusive, thereby limiting its prospects in biocontrol.

The important enzymes of commercial interest secreted by *Bacillus* spp., as well as many of the antibiotic secondary metabolites produced by the genus, appear at the onset of sporulation. To maximize product yield, a culture must be able to be held in this physiological state without completing the sporulation process. Development of sporulation-deficient strains rather than media design is the more frequent approach to solving this dilemma.

MOLECULAR BIOLOGY MEDIA

A number of excellent manuals such as those of Sambrook et al. (32) and Miller (33) detail the fundamentals of molecular biology work with *E. coli* and *B. subtilis*. This section discusses only the general media most frequently encountered for the manipulations of these organisms.

General Growth

LB (Luria–Bertani) medium (see final section of this article) is the general workhorse formulation used in many molecular biology protocols involving the maintenance and propagation of genetic bacterial stocks. NZCYM medium and its variations (NZYM, NZM) are also used when nutritionally rich formulations are required such as in the propagation of bacteriophage. Similarly, TB (terrific broth) medium (see final section) supports increased bacterial growth and is used to prepare high-yield plasmid DNA or recombinant products from *E. coli, B. subtilis,* and other organisms. Powdered basal media for many of these formulations are available from Sigma.

λ-Bacteriophage Propagation

Because λ-phage tail fibers bind specifically to the *lamB* (maltoporin; maltose-binding protein) receptor on the outer membrane of *E. coli* K12, LB medium supplemented with filter-sterilized maltose at a concentration of 0.2% will induce the receptor in a bacterial culture and help maximize phage adsorption. Glucose in the medium acts as a catabolite repressor. Divalent cations are also required for phage adsorption and integrity, and LB media for λ can be further supplemented with filter-sterilized 2 mM $MgSO_4$. Adsorption by other phage may be enhanced by calcium salts rather than magnesium. However, for the preparation of λ stocks, particularly when liquid media are used, the medium components are adjusted to reduce the amount of adsorption because the lysed phage remain in prolonged contact with bacterial cell debris and thus readsorption to debris can occur, which will reduce overall phage titer when the lysate is centrifuged to remove the cell debris. Thus, the use of LB or NZYCM supplemented with a low concentration of glucose (0.3%) to slightly repress *lamB* and 0.075 mM $CaCl_2$ as a less efficient divalent cation is an effective way to reduce readsorption problems in the preparation of λ stocks from liquid culture.

Recombinant Detection

LB agar supplemented with 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) provides an easy way to detect bacterial colonies containing plasmid DNA. The method is used for vectors encoding for the operator-proximal portion of the β-galactosidase (*lacZ*) gene and hosts containing the carboxy-terminal portion of the gene. When a normal plasmid is incorporated into a bacterial cell, α-complementation of gene fragments between the vector–host permit the formation of the functional β-galactosidase enzyme. When the chromogenic substrate X-Gal is included in the medium, the Lac^+ bacterial colonies turn blue. Plasmids containing foreign DNA inserted at the polycloning site located within the vector β-galactosidase fragment do not allow α-complementation with the host gene fragment. Thus, a recombinant bacterial colony arising from a cell transformed with a plasmid containing foreign DNA does not develop the blue color associated with β-galactosidase activity on X-Gal; recombinant colonies remain white. X-Gal is generally used at a concentration of 40 mg/L when incorporated into LB agar. X-gal is not heat stable and must be added to LB agar after the medium is autoclaved. X-gal is not water soluble and is normally prepared as a stock solution in dimethylformamide (20 mg X-Gal/1 mL DMF). The stock solution can be sterilized through a solvent-resistant (e.g., nylon) 0.2-μm membrane filter. The stock solution is light sensitive and should be stored frozen.

Some host strains contain the *lac* repressor, which prevents β-galactosidase activity until derepressed. In this capacity, isopropyl-β-D-thiogalactopyranoside (IPTG) is used as an inducer for enzyme activity. IPTG is also used to induce the TAC promoter. Because it is heat labile, IPTG is added to sterile LB medium to a final concentration of 0.5 mM from a filter-sterilized aqueous stock solution.

Competent Cells and Transformation

Although transformation is not a universal trait in bacteria, strains of *Bacillus subtilis* demonstrate the competent

ability to take up extracellular DNA via a complex membrane-mediated, nonnutritional mechanism considered to be a natural means of genetic exchange for these organisms (34). Only a small percentage of cells within a culture possess the capacity for transformation, and the generation of a competent subpopulation is best achieved with a minimal salts liquid medium amended, as necessary, with growth factors specific to the strain. Although the competent state arises with the stationary phase of culture growth, media conditions that induce maximum sporulation are not necessarily those which support a maximum population of transformation-competent cells because the two functions are physiologically unrelated.

For organisms lacking the ability to develop natural competence, or for which the natural occurrence is low (35), preparation of competent cells can be achieved chemically. For example, competent cells of *E. coli* are commonly prepared by growing a culture overnight in LB broth at 37 °C followed by treatment with highly concentrated solutions of $CaCl_2$ to increase membrane permeability to DNA. Be that as it may, the success of generating a high rate of competent cells is strain dependent as well variable with preparation batch.

THE USE OF ANTIBIOTICS IN MEDIA

Antibiotics are incorporated into culture media to eliminate the growth of contaminants or to select for an antibiotic-resistant population of cells. As antibiotic resistance is often plasmid mediated, the incorporation of a specific antibiotic into a medium that selects for plasmid-containing cells helps to prevent the spurious loss of important genetic elements carried on plasmids which can occur with constant culturing or with further genetic or physiological manipulations.

In general, antibiotics are heat labile and must be added aseptically to media from filter-sterilized or other types of cold-sterilized stock solutions. Although some antibiotics, such as penicillin-G, are stable at room temperature in powdered form, chemical degradation of these compounds begins to occur quickly once an antibiotic is in solution. The rate of degradation increases with temperature and, as noted in Table 1, many of the commonly used antibiotics lose their effectiveness within 3 days at the 37 °C incubation temperature frequently used for bacterial cultivations. In addition, tetracycline is light sensitive, and ampicillin is degraded from media by resistant bacteria. Thus, to maintain antibiotic effectiveness, it is more prudent to add the agent from a freshly prepared stock solution at time of culture inoculation rather than to incorporate the compound during the preparation and aseptic fill of the basal growth medium, especially if the medium is to be stored, even under refrigeration, for any prolonged length of time.

Table 1. Frequently Used Water-Soluble Antibiotics Used in Bacterial Media

Antibiotic	Stability at 37 °C (days)	Working concentration
Ampicillin	3	50–100 mg/L
Chloramphenicol	5	5 mg/L
Gentamycin sulfate	5	50 mg/L
Kanamycin sulfate	5	100 mg/L
Penicillin-G	3	100,000 U/L
Streptomycin sulfate	3	100 mg/L
Tetracycline	4	10 mg/L

Note: Chloramphenical and tetracycline are also ethanol soluble.

MEDIA FOR CULTURE PRESERVATION BY FREEZING OR LYOPHILIZATION

Any medium that gives sufficient growth with robust cell appearance and maintains the important characteristics of the strain is a candidate for a medium used in culture preservation. *Bacillus* spp. retain greatest viability when preserved as spores, and thus strains should be cultivated, if possible, under conditions conducive to sporulation. For nonsporulating bacilli and for strains of *E. coli*, cultures should be grown in an appropriate growth medium and harvested at the late logarithmic phase of growth. Genetically engineered strains may require the addition of a selective agent (e.g., antibiotics) to the medium to retain unstable genotypes. Previous preservation studies on plasmid containing strains of *E. coli* and *B. subtilis* indicated that plasmids were best retained with liquid nitrogen preservation. Plasmids in *E. coli* and in *B. subtilis* spores were also not affected by freeze-drying. However, freeze-drying of vegetative cells of *B. subtilis* resulted in loss of cell viability by an order of magnitude and also resulted in the requirement of a lengthy postlyophilization rehydration period before full plasmid-coded antibiotic resistance returned. Thus, the *B. subtilis* culture grown from vegetatively harvested lyophilized cells required an antibiotic-free medium for initial growth (36).

Cells of either bacilli or *E. coli* can be frozen in liquid nitrogen in growth medium supplemented with either 10% glycerol or 5% dimethyl sulfoxide (DMSO) as the cryoprotectant. This is most easily done by adding 20% sterile (autoclaved) glycerol to an equal volume of cells suspended in the culture medium or by adding filter-sterilized DMSO to cells suspended in the culture medium to a final concentration of 5%. If the culture is to be lyophilized (freeze-dried), cells should be harvested from the growth medium and suspended in an appropriate cryoprotectant such as 20% skim milk (sterilized in test tubes at 116 °C for 20 min) or the growth medium to which separately autoclaved 24% sucrose is added to a final concentration of 12%. A more complex cryoprotectant medium is a filter-sterilized reagent composed of trypticase soy broth (BBL 11768); sucrose, 100 g; bovine serum albumin fraction V, 50 g; and distilled or deionized water to 1.0 L.) Actual preservation methods are detailed in a number of publications (37,38).

MEDIA FORMULATIONS

1. HTYE Medium
   Trypticase peptone (BBL 11921), 5.0 g
   Yeast extract, 2.0 g

HEPES, 4.0 g
Distilled water, 1.0 L

Dissolve ingredients one at a time in distilled water. Adjust to pH 7.1 with NaOH. Sterilize by autoclaving at 121 °C for 15 min.

2. SEY Medium
Difco yeast extract, 1.0 g
K_2HPO_4, 0.2 g
Agar, 15.0 g
Soil extract (see below) 1.0 L

Adjust medium for final pH 7.0 and sterilize at 121 °C for 15 min.

Soil extract: Autoclave 500 g air-dried, sifted garden soil of high fertility with 500 ml tap water in a 2-L flask for 30 min. When cool, decant the supernatant, filter through paper, and bring to 500 ml with additional tap water.

3. Sporulation Medium
Yeast extract, 1.0 g
Beef extract, 1.0 g
Tryptose, 2.0 g
$FeSO_4$, 2.0 mg
Glucose, 10.0 g
Agar, 15.0 g
Distilled water, 1.0 L

Adjust medium for final pH 7.2. For broth, eliminate agar and reduce concentration to one-third of the given quantities. Sterilize at 121 °C for 15 min.

4. LB (Luria–Bertani) Medium
Tryptone (Difco 0123), 10.0 g
Yeast extract, 5.0 g
NaCl, 10.0 g
Distilled water, 1.0 L

Sterilize by autoclaving at 121 °C for 15 mins. Final pH 7.0.

5. NZCYM Medium
NZ amine (Sigma C 0626), 10.0 g
NaCl, 5.0 g
Bacto–yeast extract (Difco 0127), 5.0 g
Casamino acids, 1.0 g
$MgSO_4 \cdot 7H_2O$, 2.0 g
Distilled water, 1.0 L

Adjust to pH 7.0 with NaOH. Sterilize at 121 °C for 15 mins. Variations: NZYM (NZYCM without casamino acids); NZM (NZYM without yeast extract).

6. Terrific Broth
Bacto-tryptone (Difco 0123), 12.0 g
Bacto-yeast extract (Difco 0127), 24.0 g
Glycerol, 4.0 mL
Distilled water, 900.0 mL

Sterilize at 121 °C for 15 mins. Cool to room temperature and add sterile buffer solution (as below), 100.0 mL

Sterile buffer solution: KH_2PO_4, 2.31 g; K_2HPO_4, 12.54 g; distilled water, 90.0 ml. Sterilize at 121 °C for 15 min.

7. M9 Medium
Basal medium
Na_2HPO_4, 6.0 g
KH_2PO_4, 3.0 g
NaCl, 0.5 g
NH_4Cl, 1.0 g
Agar (for solid medium), 15.0 g
Distilled water, 990.0 mL

Adjust basal medium to pH 7.4. Autoclave at 121 °C for 15 min and cool to less than 50 °C. Aseptically add the additives listed next. If agar is needed, autoclave mineral salts and agar separately as 2× solutions and aseptically combine after sterilization.

Additives (sterilized by autoclaving as separate solutions at 121 °C for 15 min)
1 M $MgSO_4$, 2.0 mL
20% glucose, 10.0 mL
1 M $CaCl_2$, 0.1 mL

BIBLIOGRAPHY

1. D.R. Boone, Y. Liu, Z. Zhao, D.L. Balkwill, G.R. Drake, T.O. Stevens, and H.C. Aldrich, *Int. J. Syst. Bacteriol.* **45**, 441–448 (1995).
2. R.J. Cote and R.L. Gherna, in P. Gerhardt, R.G.E. Murray, W.A. Wood, and N.R. Krieg eds., *Methods for General and Molecular Bacteriology*, American Society for Microbiology, Washington, D.C., 1994, pp. 155–178.
3. M. DeFelice, M. Levinthal, M. Iaccarino, and J. Guardiola, *Microbiol. Rev.* **43**, 42–58 (1979).
4. M.N. Hughes and R.K. Poole, *J. Gen. Microbiol.* **137**, 725–734 (1991).
5. X. Hu and G.L. Boyer, *Appl. Environ. Microbiol.* **62**, 4044–4048 (1996).
6. N. Vasantha and E. Freese, *J. Gen. Microbiol.* **112**, 329–336 (1979).
7. B.D. Davis and E.S. Mingioli, *J. Bacteriol.* **60**, 17–27 (1950).
8. F.C. Neidhardt, P.L. Bloch, and D.F. Smith, *J. Bacteriol.* **119**, 736–747 (1974).
9. T.L. Miller and B.W. Churchill, in A.L. Demain and N.A. Solomon eds., *Manual of Industrial Microbiology and Biotechnology*, American Society for Microbiology, Washington, D.C., 1986, pp. 122–136.
10. R. Greasham and E. Inamine, in A.L. Demain and N.A. Solomon eds., *Manual of Industrial Microbiology and Biotechnology*, American Society for Microbiology, Washington, D.C., 1986, pp. 41–48.
11. H.I. Havik, *J. Gen. Microbiol.* **84**, 321–326 (1974).
12. L.A. Chastin and B. Magasanik, *J. Biol. Chem.* **243**, 5165–5178 (1968).
13. M.R. Atkinson, L.V. Wray, Jr., and S.H. Fisher, *J. Bacteriol.* **172**, 4758–4765 (1990).
14. M.R. Atkinson, L.V. Wray, Jr., and S.H. Fisher, *J. Bacteriol.* **175**, 4282–4289 (1993).
15. W. Harder and L. Dijkhuizen, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **297**, 459–480 (1982).
16. U. Lendenmann, M. Snozzi, and T. Egli, *Appl. Environ. Microbiol.* **62**, 1493–1499 (1996).
17. S.A. Jones, K. Dearnley, P.M. Bennet, and J. Melling, *Soc. Gen. Microbiol. Q.* **8**, 44–52 (1980).
18. E. Bech-Jensen and S. Carlsen, *Biotechnol. Bioeng.* **36**, 1–11 (1990).
19. H.L. MacDonald and J.O. Neway, *Appl. Environ. Microbiol.* **56**, 640–645 (1990).

20. G.L. Kleman and W.R. Strohl, *Appl. Environ. Microbiol.* **60**, 3952–3958 (1994).
21. D.J. Reasoner and E.E. Geldreich, *Appl. Environ. Microbiol.* **49**, 1–7 (1985).
22. M.R.S. Clavero and L.R. Beuchat, *Appl. Environ. Microbiol.* **61**, 3268–3272 (1995).
23. D.E. Dykhuizen and D.L. Hartl, *Microbiol. Rev.* **47**, 150–168 (1983).
24. P.C.S. Feng and P.A. Hartman, *Appl. Environ. Microbiol.* **43**, 1320–1329 (1982).
25. B.J. Robison, *Appl. Environ. Microbiol.* **48**, 285–288 (1984).
26. I.D. Gaudet, L.Z. Florence, and R.N. Coleman, *Appl. Environ. Microbiol.* **62**, 4032–4035 (1996).
27. J.G. Holt and N.R. Krieg, in P. Gerhardt, R.G.E. Murray, W.A. Wood, and N.R. Krieg eds., *Methods for General and Molecular Bacteriology*, American Society for Microbiology, Washington, D.C., 1994, pp. 179–215.
28. G.W. Gould and Z.J. Ordal, *J. Gen. Microbiol.* **50**, 77–84 (1968).
29. W.M. Waites and C.E. Bayliss, in G.W. Gould and J.E.L. Corry eds., *Microbial Growth and Survival in Extremes of Environment*, Academic Press, London, 1980, pp. 159–172.
30. M.M. LeCadet, M.O. Blondel, and J. Ribier, *J. Gen. Microbiol.* **121**, 203–212 (1980).
31. P. Setlow and A. Kornberg, *J. Bacteriol.* **100**, 1155–1160 (1969).
32. J.E. Sambrook, F. Fritsch, and T. Maniatis eds., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989.
33. H. Miller, in S.L. Berger and A.R. Kimmel eds., *Methods in Enzymology, vol. 152, Guide to Molecular Cloning Techniques*, Academic Press, New York, 1987, pp. 145–170.
34. J. Spizizen, *Proc. Natl. Acad. Sci. USA* **44**, 1072–1078 (1958).
35. B. Baur, K. Hanselmann, W. Schlimme, and B. Jenni, *Appl. Environ. Microbiol.* **62**, 3673–3678 (1996).
36. W.C. Nierman and T. Feldblyum, *Dev. Ind. Microbiol.* **26**, 423–434 (1985).
37. R.L. Gherna, in P. Gerhardt, R.G.E. Murray, W.A. Wood, and N.R. Krieg eds., *Methods for General and Molecular Bacteriology*, American Society for Microbiology, Washington, D.C., 1994. pp. 278–292.
38. F.P. Simione and E.M. Brown, eds., *ATCC Preservation Methods: Freezing and Freeze-Drying*, American Type Culture Collection, Rockville, Md., 1991.

See also CULTURE MEDIA, ANIMAL CELL, LARGE SCALE PRODUCTION; ENERGY METABOLISM, MICROBIAL AND ANIMAL CELLS; FERMENTATION MONITORING, DESIGN AND OPTIMIZATION; MASS TRANSFER; SECONDARY METABOLITE PRODUCTION, ACTINOMYCETES, OTHER THAN STREPTOMYCES; YEAST, BAKER'S.

MEMBRANE CHROMATOGRAPHY

JÖRG THÖMMES
Heinrich-Heine Universität Düsseldorf
Jülich, Germany

KEY WORDS

Adsorptive membrane
Mass transport
Process integration
Protein adsorption
Protein chromatography
Protein purification

OUTLINE

INTRODUCTION

The key to the successful development of a chromatographic procedure for the industrial-scale purification of proteins is to establish a high-productivity method that delivers the product in sufficient purity. Laboratory-scale protocols are usually set up with special regard to resolution, which is maximized by choosing high-performance methods based on small adsorbent particles and high-pressure equipment. On the process scale, however, high throughput seems to be the major goal, so during scale-up the small particles usually have to be replaced by larger ones due to pressure drop limitations at high-flow velocity in large columns. Additionally capacity per unit volume of adsorbent becomes an issue, thus requiring the use of porous adsorbents, which offer a high internal surface area for product adsorption. When using comparatively large porous adsorbents (d_p = 50–200 μm) the process performance is mainly determined by the velocity of protein transport to the binding sites situated on the internal surface. Because large (protein) molecules are characterized by small diffusion coefficients in free solution ($D = 1 - 50$

$\times 10^{-12}$ m^2/s) and because "hindered" transport within a porous structure leads to even lower "effective" diffusion coefficients, protein chromatography in these media has been described as being limited by slow intraparticle transport (1). Large-scale operations therefore seem to be characterized by the compromise between low pressure drop at high flow velocity and slow mass transport employing larger adsorbent particles.

The driving force for new developments in stationary-phase design has been the need to reduce the limitation by intraparticle mass transport. This may be achieved by increasing the influx velocity into the particle by enhancing diffusive transport (2) or by perfusing the stationary phase in order to support (or even replace) diffusive movement of the solutes to the ligands by convection. Perfused particles with very short diffusive pores are one option to accomplish this (3). An even more consequent way is the use of a continuous stationary phase, where the internal surface is completely accessible by convection, thus removing the necessity for diffusive motion. Adsorbents manufactured accordingly may be short in length (L) because the concept of convective solute transport essentially speeds up the sorption process and permits short residence times, t_C ($= L/v$). Sufficient capacity (adsorbent volume) is supplied by units of increased diameter, which has up to now resulted in a membrane type of design characterized by a low L/d ratio. This concept has first been realized employing modified microfiltration membranes of variable design (hollow fiber, flat sheet or membrane stack [4]), therefore the description *membrane chromatography* is justified despite the fact that a microfiltration membrane is not the only possible stationary phase fulfilling the criteria just discussed. The increasing importance of membrane adsorbents is demonstrated by two recent reviews (5,6). In this article the basic ideas behind membrane chromatography are elucidated from a conceptual point of view, supported by a short theoretical analysis. Possible limitations of the technology are discussed, followed by an attempt to position the membrane concept within a train of protein purification methods. Finally applications of adsorptive membranes on small and pilot scale are used to demonstrate the large potential of this technology as a highly productive method of purifying proteins on the process scale.

THE BASIC CONCEPT

Improved Mass Transport

As discussed earlier, membrane adsorbents try to overcome the diffusion limitation present in porous particles by allowing convective transport of solutes to the ligands. This is achieved by attaching ligands to the internal surface of stationary phases, which contain mainly through-pores and only few or no dead-ended pores. The underlying picture is the modification of the internal pore surface of microfiltration membranes with ligands in a way that, by pumping the feed solution through the membrane, the solutes are convectively transported to their binding site. Slow diffusion in dead-ended pores, as is the case with porous adsorbents, is thus avoided. The situation is depicted schematically in Figure 1. Additionally, a continuous stationary phase is used, so packing problems, which are common with particle-based media and necessitate a minimum bed length in order to neutralize packing irregularities, are circumvented. As a result, from these characteristics, adsorbents of low L/d ratio may be built, which stand for high productivity at short residence time due to fast mass transport in a continuous matrix. Prerequisite, however, is that the adsorbate solution be homogeneously distributed over the whole frontal area of the membrane, otherwise flow irregularities will neutralize the advantages mentioned. This is discussed in more detail in the following sections.

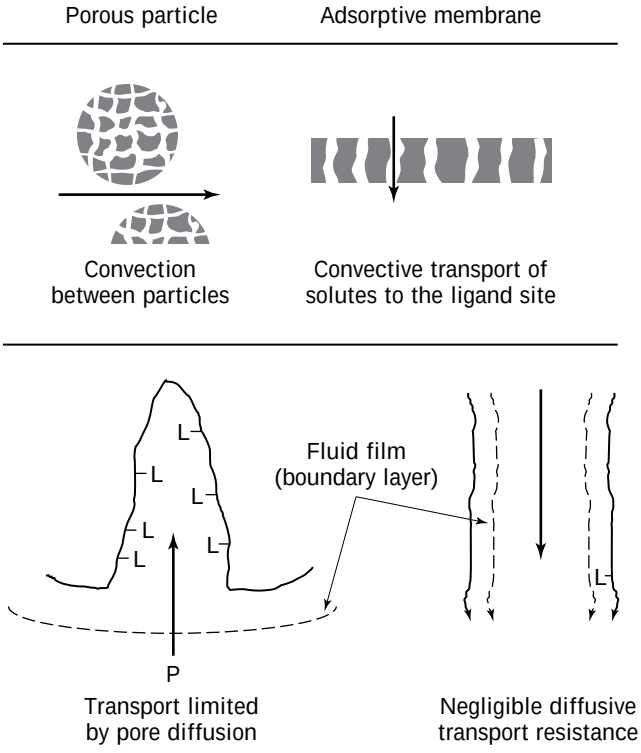


Figure 1. Schematic representation of the major transport characteristics of adsorptive membranes and porous particles.

Geometric Formats

Adsorbents fulfilling the criteria developed may be arranged in different formats. As shown in Figure 2, it is possible to differentiate between four basic concepts: single or stacked flat-sheet membranes (A), hollow-fiber modules (D), radial-flow cartridges containing spiral-wound membranes (C), and cartridges packed with continuous stationary phases in the form of membranes or rods (B). All these arrangements are commercially available; a complete list of available products, however, is difficult to compile because the life-span on the market of some of these products has been limited. The reader is referred to the information provided by Langlotz et al. (7), Roper and Lightfoot (5), and Thömmes and Kula (6).

Single-sheet membranes may be employed in suitable filter holders, as long as sufficient care is taken to equally distribute the process liquid across the frontal membrane

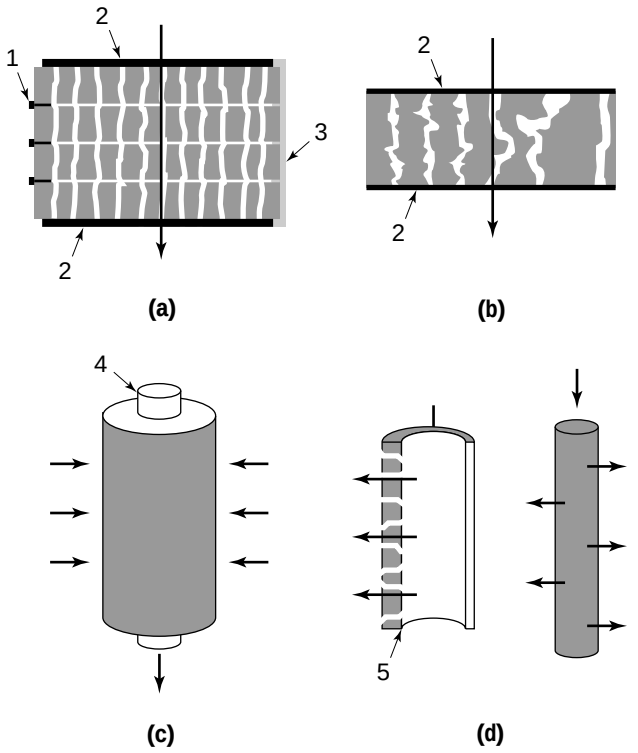


Figure 2. Schematic representation of the four major classes of membrane adsorbents: (**a**), stack of flat sheet membranes; (**b**), continuous cartridge with through-pores; (**c**), radial flow unit with spiral wound membranes; (**d**), hollow-fiber membrane; 1, individual gasket; 2, distributor and collector plate; 3, complete sealing around a membrane stack (e.g., silicone); 4, central tube for liquid collection; 5, cross-cut through an individual hollow-fiber.

area and to efficiently collect the liquid after passage through the membrane. The arrangement of membranes in stacks allows the increase of the capacity of an individual unit as well as an adjustment of the length of the adsorbent bed. The sheets must be separated by proper sealing in order to avoid fluid flow around the membranes instead of through them. This may be achieved by single gaskets or by embedding the stack in a suitable sealing substance (e.g., silicone). Alternatively a continuous adsorbent packing with thoroughgoing pores may be used, thus making the stacking of individual sheets obsolete. Radial-flow cartridges contain an internal tubing for fluid collection, which is surrounded by a spiral-wound membrane sheet. The liquid is transported in a radial direction from outside to inside, thus keeping the column aspect ratio, L/d, small, as is the case in a flat-sheet membrane. The same is true for hollow-fiber modules, where the fluid passes the membrane perpendicular to the influx direction. L is defined here as the wall thickness of a single fiber. These modules may be operated in a dead-end mode with the column outlet sealed, so that all the liquid must pass through the membrane; alternatively a cross-flow mode may be chosen with the possibility to recycle the process solution. The total capacity of a hollow-fiber adsorption module can be adjusted by bundling several fibers in a single module.

Static Capacity

As already discussed, on a preparative scale the capacity per unit volume of adsorbent available for a certain purification task becomes an important issue because adsorbent cost as well as plant space requirements of the whole purification unit have to be taken into account in economic considerations. Because the ligands are attached to the internal surface of stationary phases, their specific surface area, $a(m^2/m^3)$, determines the equilibrium capacity in a first approximation. Champluvier and Kula measured a for nylon membranes ($d_p = 1.2\,\mu m$) as 250 m^2/m^2 of frontal area (8). Assuming an average membrane thickness, L, of 200 μm we obtain a as 1.2×10^6 m^2/m^3. Conventional porous particles for protein chromatography have a specific area, which is significantly larger than this value (e.g., 1.5×10^7 m^2/m^3, our own measurements with the silica-based matrix Bioran CPG, Schott, Germany, 200-μm average particle diameter, 100-nm average pore diameter), so this seems to be a major drawback of membrane-based stationary phases. In fact membrane adsorbents with a monolayer coverage of ligands on their internal pore surface have suffered from reduced static binding capacity, as has been summarized by Roper and Lightfoot in their review (5). An approach to circumvent reduced capacity is the introduction of a three-dimensional binding layer on the pore surface. This is achieved by graft polymerization of monomers containing functional groups, which may be derivatized to protein-binding ligands. The resulting functionalized polymer layer works as a three-dimensional binding space with a gellike structure, which is able to take up adsorbate molecules in multiple layers, as depicted schematically in Figure 3. This technique has been applied to hollow-fiber membranes with great success (9–19) and is also used in commercially available flat-sheet membrane adsorbents (Sartobind, Sartorius, Göttingen, Germany) yielding a capacity comparable to porous adsorbent particles (20). The increase in static capacity for standard proteins compared with (theoretical) monolayer coverage has been reported to be between 4- and 20-fold for hollow-fiber modules (16) and up to 100-fold for flat-sheet membranes (21). Wang et al. compared the static capacity of polymer

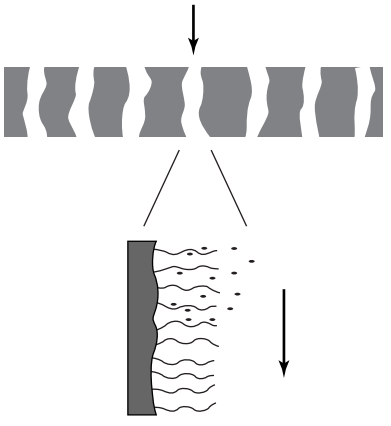


Figure 3. Schematic representation of the modification of the internal pore surface with grafted polymer ligands.

coated membranes for cytochrome *c* to membranes with monolayer ligand coverage, as well as to porous particles, and found the three-dimensional ligand layer to yield significantly higher values (22). However, as is discussed later, the gellike structure of the layer may change the mass transfer characteristics of the membrane adsorbents, a fact that has to be included in the analysis of the sorption performance.

Process Performance

When the suitability of a new adsorbent design for the large-scale purification of proteins is to be estimated and compared with established stationary phases, a benchmarking method must be defined. On the process scale, an important question is, how much product may be obtained within a certain period of time with a given amount of adsorbent? Additionally, the product should have sufficient purity in order to reduce the overall number of process steps in a purification train. Summarizing, the basic questions are the following:

- Which throughput is possible with a given set-up as a result of the permeability of the stationary phase (mechanical limitation)?
- Which throughput is possible with a given set-up as a result of the efficiency of the sorption process (dynamic limitation), in other words, is the static capacity of the adsorbent used well enough under the process conditions chosen?
- Which product purity is achieved with a given set-up under the process conditions chosen (resolution limitation)?

The permeability, B, of a membrane adsorbent is a function of the adsorbent's pore diameter and of the overall porosity of the matrix; typical values may be calculated from pressure-drop data available from the literature. Values ranging from 7×10^{-16} m^2 for continuous phases from methacrylate block polymers with 0.4-μm pore diameter (23), to 2×10^{-15} m^2 for 1.2-μm d_p hollow fibers with grafted polymer ligands (24), and 2.5×10^{-14} m^2 for commercially available stacks of flat-sheet membranes (average $d_p = 15\ \mu$m) (20) may be found. The pressure drop in the adsorption unit may then be calculated for a given set of operating conditions (L, v, η), after equation 1.

$$\Delta p = \frac{v \cdot L \cdot \eta}{B} \tag{1}$$

The dynamics of the sorption process may be determined from frontal adsorption experiments. The deviation of measured breakthrough curves from the ideal step function shows how efficiently the static capacity of the adsorbent is used. Measuring the breakthrough capacity, Q_D, (at 1, 5, or 10% breakthrough) under different experimental conditions (e.g., by varying L and v) allows the definition of the process conditions that make the best use of the static capacity of the adsorbent. In addition, a theoretical analysis will yield characteristic quantities (e.g., numbers of transfer units, plate heights, or transport and dispersion coefficients) that allow a prediction of the process performance according to standard models (21,25–27).

The overall performance, however, must take washing, elution, and regeneration into consideration; this can be expressed as the total number of bed volumes required for all of these operations, α. The product yield is considered by the recovery ratio, R_E. Productivity may be defined as amount of product per unit time and unit adsorbent volume, as shown in equation 2 (21).

$$P = R_E \frac{1}{\left(\frac{Q_D}{C_0} + \alpha\right)} \left(Q_D \cdot \frac{v}{L}\right) \tag{2}$$

Finally the product purity should be addressed by introducing the resolution, R_s, of the membrane-based separation under process conditions. An attempt to optimize R_s will often show that a compromise has to be found between maximizing the throughput at low L/d and maintaining high resolution, because the short bed lengths make adsorptive membranes into typical low-plate-number systems (28). Coffman et al. have shown, however, that most protein separations may be run successfully by employing membrane adsorbents because separation will occur according to an on–off separation mechanism that requires no separation by differential migration in long columns (29).

To summarize, the quality of a membrane adsorption method will be defined by the product purity it delivers at maximized productivity. A significant increase of this quality compared with standard methods using porous particles is anticipated due to the fact that the improved mass transport capability allows high throughput at low L/d at maintained resolution under on–off separation conditions. The factors that may limit the achievement of this ambitious goal are discussed next.

LIMITING FACTORS IN MEMBRANE ADSORPTION PROCESSES

As with all processes based on fixed beds, sorption dynamics of adsorptive membranes may be limited by mixing, mass transport, and the kinetics of the binding reaction. In order to discuss these effects precisely, one must attempt to define them according to the specific details of membrane adsorption units. Kinetic limitations arise from the rate of formation of the adsorbate–ligand complex, which is described by the forward reaction rate constant, k_a. Together with the rate constant of dissociation of the complex, k_d, we obtain the dissociation constant of the binding equilibrium, K_D (k_d/k_a). Further limitations are caused by slow transport of adsorbate molecules from the bulk solution to the binding site (e.g., diffusive motion through the fluid boundary layer above the binding surface). Although the concept of throughgoing pores in membrane adsorbents has been invented to overcome slow diffusion in dead-ended pores, the membrane production process may result in adsorbents with some pores of that kind, so in specific cases pore diffusion has to be considered in the description of sorption dynamics. Additionally, as

already mentioned, the attachment of a three-dimensional adsorbing layer to the internal pore surface must be regarded as a potential mass transport barrier. Mixing in adsorptive processes should be split into three major effects: fickian dispersion, nonideal fluid flow within the adsorbent, and extracolumn mixing (30). The first effect describes the dispersive motion "against" the direction of convective flow, the second effect accounts for irregularities within the adsorbent, leading to uneven fluid flow through the packing, whereas the third summarizes contributions of tubing, extracolumn dead volumes, liquid distributors and collectors, zone broadening, and so forth. Very often these different contributions are lumped together in a single dispersion coefficient, D_{ax}, which is then characteristic for the whole experimental set-up rather than for the adsorbent itself. In the following the respective importance of these limitations is discussed in detail in order to provide a basis for the prediction of the performance of adsorptive membranes in the purification of biomolecules.

Mixing

Roper and Lightfoot (30) analyzed peak broadening in stacked flat-sheet membranes under nonbinding conditions with a specific flow-reversal technique in order to differentiate between fickian dispersion, nonideal fluid flow within the adsorbent, and extracolumn mixing. They used the height equivalent of a theoretical plate (HETP) concept to determine the individual contributions to the overall peak broadening and found the plate heights, h, resulting from fickian dispersion to be 1–5 μm for a range of flow velocity, v, between 0.5 and 4 cm/min. These values were in accordance with theoretical considerations, which predicted h as a function of the dispersion coefficient D_{ax} and of v according to equation 3.

$$h = \frac{2 \cdot D_{ax}}{v} \tag{3}$$

Values for h reported in literature for adsorptive membranes, however, were between 20 and 200 μm and showed a slight increase with linear velocity. This difference could be attributed to extracolumn effects and nonideal fluid flow. Peak broadening by a factor of 1.7 occurred due to extracolumn dispersion; the additional difference between h measured in this work and h in the literature data arose from uneven flow distribution in the stack. Similar results were reported by Gebauer et al. (21), who found the mixing in stacks of commercially available flat-sheet membranes (represented by an overall D_{ax}) to be 10-fold higher compared with an estimation based on literature correlations. This difference was attributed to extracolumn effects as well. Enlarged overall mixing may be a significant influence on the performance of membrane adsorbents. This was demonstrated in a scale-up study on the isolation of human serum albumin from plasma using stacks of ion exchange membranes (20). The increased mixing found with large-scale modules compared with laboratory-scale units led to a more than threefold reduction of the productivity of the process. Josic et al. (31) investigated the influence of liquid distribution at the column inlet of cartridges with a continuous stationary phase of type B (compare Fig. 2). In the case of nonoptimized inlet design, a significant part of the fluid passed through the central region of the unit, thus leaving large parts of the membrane unreachable by convective transport. The uneven flow distribution led to a significant loss of performance of chromatographic separations employing these units. With optimized inlet design, resolution was comparable with established HPLC methods. Reif and Freitag (32) demonstrated the importance of reduced dead volume behind the membrane; a large mixing zone led to deteriorated resolution in the analytical separation of standard proteins.

Nonuniform flow distribution may be caused by variations in the pore size of adsorptive membranes. Suen and Etzel (26) as well as Liu and Fried (33) demonstrated by computer simulations that a wide pore-size distribution may lead to significantly flattened breakthrough curves during frontal protein adsorption and thus to a strongly reduced useful capacity of the adsorbents. Similar effects were shown for variations in membrane thickness. As a consequence, stacking of membranes was recommended to reduce these effects. Kim et al. described the modification of hollow fibers with grafted polymer chains bearing phenylalanine and tryptophan as pseudoaffinity ligands for immunoglobulin adsorption. The deviation of protein breakthrough from ideal behavior with these adsorbents was also attributed to a wide distribution in pore length (12). The distribution in pore diameter in protein A modified-hollow-fiber modules was described by Klein et al. as the reason for incomplete use of the static capacity of these adsorbents in immunoglobulin capture (34). The influence of overall mixing on performance during frontal adsorption was investigated in a theoretical study by Suen and Etzel (26). These authors used the dimensionless column Peclet number, Pe, to quantify dispersion according to equation 4 (v_e is the interstitial flow velocity).

$$Pe = \frac{v_e L}{D_{ax}} \tag{4}$$

Peclet numbers above 40 were described to be sufficient to exclude a deteriorating influence of mixing on the breakthrough behavior during frontal adsorption to membrane stacks. Because Pe is proportional to the bed length, augmented mixing (represented by increased D_{ax}) may be compensated for by increasing L. Similarly Liu and Fried (33) reported a Pe above 25 to be sufficient to neglect dispersion as a possible limitation of the sorption process. It is important to note that these considerations were set up for frontal adsorption of a one-component system; stronger boundary conditions may be required for the separation of a multicomponent system.

To summarize, as long as sufficient care is taken during process design to minimize extracolumn effects as well as nonuniform flow, mixing phenomena are unlikely to be the limiting factors in membrane chromatography. However, it has to be kept in mind, that adsorptive membranes are used in a chromatographic process; therefore standard criteria of filtration module design will not meet the requirements of a sorption process and may possibly lead to a loss of the advantages of the membrane-based process.

Mass Transport

When dealing with porous particles for protein purification, there are two potential sources of transport resistance to adsorption: transport of adsorbate from the bulk of the solution to the external particle surface, and transport within the particle to the binding site. In this case the general perception is that fluid-side transport is an order of magnitude faster than transport within the adsorbent; therefore film transport resistance is often neglected in a simple analysis. Band spreading in frontal analysis, as well as peak broadening during chromatography, then is attributed to slow, often hindered, intraparticle diffusion. The concept of using membranes with throughgoing pores for protein adsorption was supposed to remove the transport resistance by reducing the characteristic length for protein diffusion to the ligands. Brandt et al. (4) provided a general estimation of the importance of mass transport for the efficiency of an adsorptive membrane by relating the characteristic diffusion time in a membrane pore, t_D ($t_D = d_p^2/D$), to the contact time, t_C ($t_C = L/v$), during passage through the membrane. Efficient usage of the static capacity was supposed to take place if $t_C \gg t_D$. Inserting standard values for adsorptive membranes into this equation ($d_p = 3 \times 10^{-6}$ m, $L_{\text{single sheet}} = 2 \times 10^{-4}$ m, $v = 8.3 \times 10^{-4}$ m/s) and assuming an average diffusion coefficient, D, of 5×10^{-11} m^2/s we obtain $t_D = 0.18$ s, whereas t_C even for a single membrane is 0.24 s. Therefore a stack of 10 membranes, which is often found in commercially available modules, easily fulfills the condition, which allows us to neglect diffusive transport resistance. This is in contrast to standard porous particles, where t_D is based on the particle radius and therefore can be estimated as 40.5 s for the same protein in a particle 9×10^{-5} m in diameter.

The absence of dead-ended pores and the concomitant disappearance of the intraparticle transport resistance should make diffusion through the fluid boundary layer above the ligands the only transport step (see Fig. 1) in membrane chromatography. The rate of transport through this layer is dependent on the free diffusion coefficient of the adsorbate as well as on the thickness of the layer. Because interstitial flow velocities in narrow pores of microfiltration membranes are comparably high and boundary layer thickness thus may be low, this transport resistance commonly is regarded as not likely to limit the sorption process. Suen and Etzel (26) demonstrated the validity of this approach for membranes with small pore diameter by comparing the characteristic time needed to diffuse through the fluid boundary layer with the average residence time in a module. They were able to show that even for very small diffusion coefficients (i.e., large protein molecules) the residence time in a correctly designed membrane stack is much higher than the diffusion time through the fluid film, so the importance of this transport resistance is low compared with other possible limitations. In cases of much larger pore diameters these considerations may no longer be valid. Adisaputro et al. (35) measured the dynamics of whey protein uptake to membranes of very large pore diameter (50–300 μm) and found reduced capture efficiency with decreasing residence time. Calculation of the fluid film diffusion time for an average protein showed that, with these large pore diameters, film transport was a likely candidate for the dominant limitation of sorption efficiency.

As mentioned earlier, the introduction of a three-dimensional ligand layer to the pore surface may also introduce a mass transport resistance. Gebauer et al. (21) investigated commercially available ion exchange membranes with respect to their sorption dynamics and found a limitation of sorption efficiency by a diffusion mechanism. Depending on the length and density of the ligand layer a pore or solid diffusion mechanism of protein transport was identified. The breakthrough capacity, however, could be maintained on a very high level for a wide range of flow velocities, a fact that was attributed to convection-aided mass transport into the flexible chains of the ligand layer. Attaching polymer-grafted ligands to the internal pore walls of hollow fibers also yielded stationary phases with three-dimensional adsorbing layers. Here no dependency of breakthrough behavior from residence time in the module was detected, thus supporting the thesis of fast protein influx into the ligand sphere (12,15–17,24).

Mass transport limitations are not as dominant in membrane adsorption as they are in standard protein chromatography using porous particles. The complete neglecting of transport resistance, however, is possible only under certain conditions. Therefore transport in adsorptive membranes should be analyzed carefully during process design based on these stationary phases.

Binding Kinetics

If axial mixing and mass transport may under certain circumstances be excluded as potential limiting factors of sorption efficiency, the kinetics of the protein–ligand interaction becomes the focus of attention. Brandt et al. (4) used a second-order rate expression to estimate the reaction time, t_R, for interaction of BSA with a monoclonal antibody (MAb) affinity ligand. They obtained $t_R = 5$ s, a significantly higher value than the average diffusion time in membrane pores, which was calculated earlier. From this estimation it can be deduced, that slow association kinetics may limit the performance of affinity-membrane adsorbents. A similar situation was predicted by Suen and Etzel (26) from a thorough theoretical analysis, which was later experimentally verified with the pepstatin A/pepsin affinity system (36). Kinetic limitations were also found to be responsible for deviations of protein breakthrough from an ideal step function in dye ligand/dehydrogenase pseudoaffinity systems employing stacks of flat-sheet membranes (37). Briefs and Kula (38) reported the onset of peak broadening due to kinetic limitation with a similar pseudoaffinity interaction from t_C below 20.5 s. For ion exchange interactions, however, no kinetic limitation could be verified within the range of experimental conditions ($t_C = 5.1$ to 20.5 s). An investigation on kinetically limited membrane chromatography was performed by Nachman et al. (39–41) employing hollow-fiber modules with affinity ligands attached to the internal pore surface. In this study it was shown that reducing the residence time in the module in turn reduced the capture efficiency of the adsorbents. In addition, under all experimental conditions, t_C was well

above t_D, which had been calculated based on the membrane geometry and the diffusion coefficient of the adsorbate. Also the capture efficiency was independent from initial concentration, as can be predicted for kinetic limitation under conditions of a rectangular isotherm (21). Kinetically limited adsorption may be exploited in membrane chromatography to separate adsorbate molecules that have identical binding isotherms but different association rate constants. In standard porous adsorbents these pairs cannot be separated due to the fact that slow diffusion in the pores leads to the formation of a local equilibrium between adsorbate in the pore liquid and bound adsorbate. Membrane adsorbents are sensitive to differences in adsorption kinetics, provided neither mass transport nor mixing are "hiding" the kinetic differences. This was theoretically predicted (27) and proven experimentally (36) for an affinity interaction based on stacks of flat-sheet membranes.

Adsorptive membranes offer a rare opportunity to investigate effects of binding kinetics in adsorptive processes. The fact that this is possible underlines the excellent mass transport characteristics of these matrices, which allow very fast protein purification under conditions of protein–ligand interactions as the rate-limiting step in the sorption process. However, this seems to be limited to affinity interactions, which show comparatively slow association rates with ligands attached to planar surfaces due to higher steric requirements. With fast ion exchange interactions, even adsorptive membranes have not been able to demonstrate kinetic effects during protein sorption.

POSITIONING OF MEMBRANE ADSORPTION IN A PURIFICATION TRAIN

Now that adsorptive membranes have been characterized according to their basic properties, where should they be introduced into a standard protein purification scheme? Wheelwright differentiated between initial recovery, low-resolution purification, and high-resolution purification in a downstream process (42). Initial recovery is supposed to deliver a clarified product solution of reduced volume, whereas low-resolution purification should lead to a significant enrichment of the target molecule and to the removal of major contaminants. High-resolution purification refers to the removal of minor impurities with properties similar to those of the desired product, as well as final polishing (e.g., separation of correctly processed molecules from false variants, multimers, etc.). Depending on the design of the adsorption unit, adsorptive membranes can accommodate any of these tasks.

Krause et al. demonstrated that application of dye–ligand affinity membranes in a cross-flow mode allowed the purification of malate dehydrogenase from *E. coli* homogenate (43). A clarified, concentrated, and initially purified product solution was obtained after elution, thus demonstrating the feasibility of membrane adsorption for integration into primary recovery. Additional support for the use of membrane-based adsorption in this early process stage comes from the isolation of proteins from unclarified whey, which was accomplished with large-pore membrane stacks (35).

A second assignment in the area between initial recovery and low-resolution purification is the use of adsorptive membranes for product capture from diluted but already clarified solutions. Good mass transport characteristics together with fast binding kinetics make ion exchange membranes a successful candidate; hydrophobic interaction and pseudoaffinity ligands are also suitable for this task (11,17,18,44–46).

Low-resolution purification is achieved when membranes are used for the separation of proteins in order to initially fractionate clarified product solutions (32,47–50). Sequential arrangement of membrane units for different separation mechanisms allows the development of a multidimensional separation concept (51).

The use of membranes for high-resolution purification was realized in high-performance membrane chromatography (HPMC) (23,31,52). Here the properties of continuous stationary phases of type B (Fig. 2) are used with ion exchange, hydrophobic interaction, and reversed-phase ligands. Affinity (34,53,54), immunoaffinity (40), and receptor affinity membranes (41) also belong to this category, although they may be used early in a purification train in order to introduce a high selectivity step in the very beginning of the downstream process. Fast binding kinetics favor the use of very dilute feed solutions because high-affinity adsorption may be operated with large throughput in membrane devices, even if the feed concentration is on the order of magnitude of the dissociation constant, K_D, of the adsorbate–ligand interaction (39).

The versatility of adsorptive membranes in serving the needs of the different stages of a downstream process is summarized in Figure 4.

OPTIMIZING THE PERFORMANCE OF ADSORPTIVE MEMBRANES

If we define the performance of a membrane adsorption unit according to the parameters described earlier ("Process Performance") as the realization of a desired product purity at maximized productivity, then the optimization task has two main goals: the achievement of a certain resolution and the simultaneous increase of throughput. The problem of maximizing throughput requires information on mechanical stability of the stationary phase, the quality of fluid flow through the unit, and knowledge of the overall kinetics of the sorption process. The latter point may in a first approximation be reduced to the question of residence time within the adsorbent phase. As discussed in the section "Mass Transport", adsorptive membranes are supposed to have superior transport properties even if the static capacity is increased by introducing a three-dimensional ligand layer on the internal membrane surface. This fact in turn reduces the residence time ($t_C = L/v$) that is needed for efficient usage of the available capacity. An indicator of this efficiency often used in protein chromatography is the breakthrough capacity, which increases with residence time and is dependent on the performance of the stationary phase with regard to the overall dynamics of the sorption process, be it limited by transport processes, dispersion, or binding kinetics. The better the

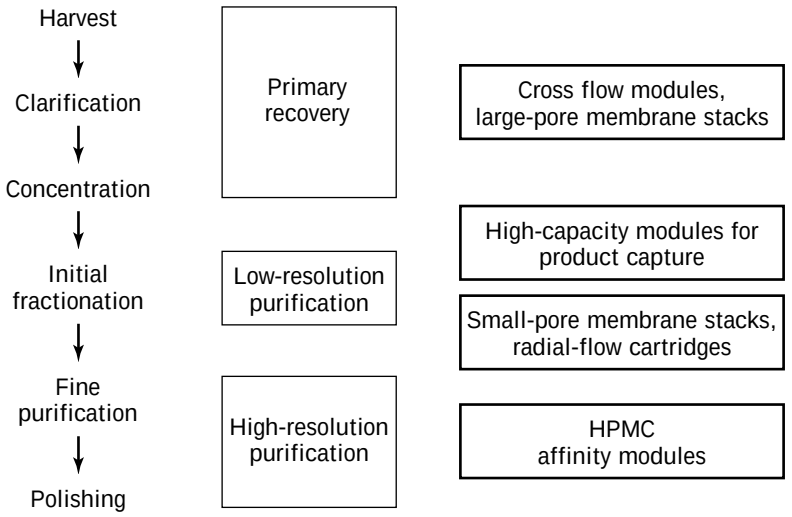


Figure 4. Positioning of membrane adsorbents in a purification train.

dynamic properties of the adsorbent, the shorter the residence time can be without compromising the breakthrough capacity. In the case of adsorptive membranes this allows using adsorbent beds that are short in length (L) and can be operated at high fluid velocity (v). The work of Yamamoto and Sano (55) as well as equation (2) predict the productivity to increase with the breakthrough capacity, and to decrease with t_C. As breakthrough capacity increases with t_C, there is an optimum in t_C with respect to productivity. The advantage of adsorptive membranes is that efficient usage of the static capacity is already obtained at short t_C; therefore a high productivity of membrane-based adsorption methods can be anticipated.

A sophisticated way of predicting the dynamic behavior of membrane adsorbents has been presented by Suen and Etzel (26,27). A numerical solution of the equations describing membrane adsorption was used to analyze the influence of operating parameters on performance. In addition to breakthrough behavior, the separation of two proteins in kinetically limited chromatography on membrane stacks was simulated and proven experimentally. As shown by Gebauer et al. (21) productivity of membrane adsorption systems can also be analyzed based on analytical solutions of the equations describing protein breakthrough, if a single limiting step can be identified for the process under investigation.

Similar considerations are necessary if the dependency of resolution on the operating parameters is estimated. From the fact that membranes may be regarded as low-plate-number systems, a minor performance in the separation of proteins could be predicted. As shown by Coffman et al. (29), however, most protein separations are operated in an on–off mode during elution chromatography. This means that the selectivity of the separation relies on the differences in the interaction between the proteins to be separated and the adsorbent at various concentrations of the eluting agent. Separation by differential migration, which is caused by the different retardation of the individual molecules during their movement through the adsorbent, plays only a minor role. Therefore only a few theoretical stages are needed for the satisfactory separation of proteins by ion exchange, hydrophobic interaction, or reversed-phase methods; the intrinsic separation factor of the different proteins is high enough to justify separation according to an on–off mechanism. Consequently, short membrane adsorbents allow very good separations, and the resolutions obtained are mostly comparable to HPLC separations, where long columns yield very high plate numbers under high-pressure conditions. Coffman et al. reported that about 50 stages is sufficient for most separation tasks, so HETP values between 20 and 200 μm, which have been found for adsorptive membranes, necessitate bed lengths of 10 mm for good resolution in protein separation. Similar findings on the bed length necessary for satisfactory results in protein separation using adsorptive membranes were presented by Dubinina et al. (56). In their analysis the authors demonstrated, that only a short bed is needed if proteins are separated according to an on–off mechanism. In an extended study the influence of gradient steepness, flow velocity, and bed length was investigated, and it was shown that short adsorptive membranes may be used for high throughput separation with good resolution.

Because increasing throughput (either by increasing the load or by reducing t_C) must be suspected to reduce resolution, the same criteria for optimizing L/v apply as in the case of the breakthrough capacity of membrane adsorbents. The good mass transport capabilities then allow the use of high flow rates at retained resolution when using well-designed membrane units. This is supported by the findings of many workers, who reported excellent resolution over a wide range of flow velocities during the separation of proteins with these devices (23,31,32,47,49, 52,57–59).

A major problem of adsorptive membranes (introduced earlier in "Mixing") is a direct consequence of the improved transport properties. Operating at short t_C results in configurations of small column aspect ratio (L/d). Ensuring an even fluid flow through an adsorbent bed of large diameter is a major task during module design for application of

membrane chromatography, especially on a large scale. Part of the problem arises from the difficulty of distributing liquid evenly over a large diameter. In the case of adsorption to particle-shaped stationary phases, uneven distribution of fluid at the column inlet will be neutralized during passage through the comparatively long beds; the lower part of a column will act as an additional distribution unit. In the case of membrane adsorbents this is not possible because the very short bed length, which can be chosen without compromising the performance, is the major advantage of this concept. Thus a very careful design of the distribution unit of large-scale equipment is essential. Up to now, commercially available units have not taken this into consideration, a fact that can be considered as one of the main reasons why membrane adsorption has not lived up to its preparative potential in protein purification. Furthermore the fabrication of the membrane adsorbent itself has to be taken into account. Wide distribution in pore size will lead to a preferential passage of fluid through certain regions of the membrane, resulting in a broad distribution of residence times with the respective consequences for process performance. This was analyzed theoretically by Suen and Etzel (26), who demonstrated that stacks of flat-sheet membranes are suited to achieving a homogenized flow through membranes with a large variance in porosity. It must be pointed out that hardware problems (uniform membranes and proper housing design) presently are the major limitation of a further success of membrane adsorption on the process scale. Until these problems are solved this promising technique will remain the "eternal talent" within the large variety of stationary phases for protein adsorption that are available on the market.

APPLICATIONS

As described earlier, adsorptive membranes may be used at several positions in a purification train. The scope of this section is to discuss applications from each of the positions within a purification scheme, but the selection provided here is not intended to be a complete collection of membrane-based adsorption methods. Therefore the reader is referred to recent review articles for additional applications (5,6,60). The examples were selected in order to test the prediction that processes employing membrane adsorbents have superior productivity at retained resolution.

Primary Recovery

Because microfiltration membranes are suitable for the clarification of culture broth or homogenates, the attachment of ligands to the internal pore surface allows a combination of protein adsorption and solid–liquid separation in a single step. This approach has been realized by Krause et al. (43), who used Cibacron blue–modified microfiltration membranes (Sartobind Blue, Sartorius, Göttingen) for the isolation of malate dehydrogenase directly from unclarified *E. coli* homogenate. The membrane adsorbent had a static capacity similar to commercially available particles with the same ligand and was used in a cross-flow mode with recycling of the filtrate (i.e., a batch adsorption mode). The overall productivity was 1,200,000 U/(L h) and combined clarification, capture, and initial purification in a single step. Data on long-term stability of this method are lacking, however; therefore the influence of traditional shortcomings of microfiltration membranes due to fouling, layer formation, and so forth cannot be judged.

A different field of application is the capture of whey proteins from unfiltered whey, employing large-pore membrane stacks (35) (Productiv S and QM, BPS Separations, Spennymoor, U.K.). Adisaputro et al. replaced the traditional method of batch adsorption of whey proteins with membrane stacks. Based on their results the authors predicted an increased productivity of the membrane method at the short residence times, which become possible due to fast mass transport characteristics. The size of the pores, however, was a point of concern because, on one hand, a large pore size was required to allow the use of unfiltered whey as an initial recovery operation, but on the other hand, too-large pore diameters cause fluid-side boundary layer problems, as already discussed.

Capture

If the use of clarified solutions as a feedstock for adsorptive processes is preferred to the whole-broth scenario just outlined, then the situation after removal of particulate matter is characterized by a large volume of product at a comparatively low concentration. This is especially valid for the production of recombinant proteins, so the primary goal at this point is the reduction of process volume. Usually, a concentration step such as ultrafiltration or precipitation is introduced, but this may, as well, be a perfect task for membrane adsorption. The efficient capture at short residence time is coupled with the selectivity of an adsorption process, so concentration may be combined with a high-productivity initial purification.

A typical process example is the purification of MAb from hybridoma cultures. The product concentration usually does not exceed 1 g/L even with highly efficient cell lines; the range of titers found is instead on the order of a few hundred milligrams per liter. Additionally the growth medium has physiological ionic strength (13–15 mS/cm), which requires diafiltration or dilution of the feedstock in order to allow sufficient capacity if an ion exchange step is chosen as the initial purification method. Ion exchange membrane adsorption has been described as an efficient method to capture MAb directly from clarified broth, possibly preceded by dilution to reduce ionic strength. Jungbauer et al. (61) used Zetaprep QAE radial flow cartridges (LKB, Bromma, Sweden) for this purpose and performed a scale-up study from 50 mL to 930 mL total adsorbent volume. The productivity reported was 0.2 g/(L h); a certain enrichment of the MAb was demonstrated by SDS gel electrophoresis. Cation exchange membrane stacks as well as a pleated module were used by Lütkemeyer et al. for the isolation of IgG from diluted hybridoma supernatant on a laboratory scale (28 mL of adsorbent, feedstock volume 5.5 L) (58). The high static capacity of the adsorbent (50 mg/mL) together with short residence times in the short adsorbent bed yielded a productivity of 75.2 g/(L h),

a value that was 13 times higher compared with the same process performed using a packed bed of porous cation exchange particles. In a later publication a similar process was shown on the pilot scale with undiluted hybridoma supernatant. Even at the high ionic strength, the overall productivity was 29 g/(L h). During the experiments, 100 L of supernatant was processed in 2 h in a unit with a 4-mm bed length (Sartobind S, Sartorius, Göttingen, Germany; $V = 220$ mL); the process volume was reduced 44-fold at 100% recovery (62). A similar approach was chosen by Wang et al. for the purification of a recombinant immunofusion protein secreted by *E. coli* (22). The same stack of cation exchange membranes was used for capture of the target protein from 200 L of clarified and 10-fold-diluted culture broth. The reported productivity was up to 1.5 g/(L h) at 10-fold reduction of process volume. Capture of urease to a hollow fiber, which had been modified to a tentacle type of adsorbent by grafting polymer chains with anion exchange ligands to the internal pore surface, was described by Matoba et al. (17). The productivity (expressed as urease collection rate) found in these investigations was up to 200 g/(L h). These examples show how the efficiency of membrane adsorbents may be exploited for high-productivity capture of target proteins from diluted culture supernatants.

Low-Resolution Purification

Because even during capture processes with adsorptive membranes a separation of the target molecules from impurities is achieved, there is no simple borderline to low-resolution purification. One way of differentiating between the processes is with regard to their main tasks, which are reduction of process volume and isolation of the main component from major impurities. During capture mainly ion exchange membranes are used, but hydrophobic interactions as well as pseudoaffinity ligands may be applied during low-resolution purification. At this point the considerations for optimizing t_C with regard to resolution under high-throughput conditions, which have been discussed, become important.

A typical example of low-resolution purification is the isolation of HSA in a plasma fractionation process based on ion exchange adsorbents. This was first described by Lacoste-Bourgeac for the valorization of fraction IV paste, which originated from the traditional ethanol precipitation during plasma fractionation (63). A series of anion exchange and cation exchange spiral-wound membrane cartridges (Cuno, Cergy-Pointoise, France) was used to recover albumin at 66% yield from the paste. Thus the high content of protein in this fraction, which otherwise was reported to be of only little value, could be exploited. A process for the isolation of HSA employing stacks of flat-sheet membranes was presented by Gebauer et al. and was compared with the well-established ion exchange process based on porous beads (20). Here a series of strong anion and cation exchange steps was used with a productivity of up to 220 g/(L h) on the laboratory scale. This value was six to eight fold higher than the traditional process based on a porous ion exchange matrix (Sepharose FF, Pharmacia, Uppsala, Sweden). The procedure was increased in scale based on commercially available equipment to 110 mL of adsorbing volume; the results, however, showed a drop in productivity of a factor of 3 compared with the small scale. Additionally a significant dilution of the product was observed. The deteriorated performance of the large-scale module could be attributed to overall mixing effects in the module, which originally was a filtration unit and had not been optimized with regard to chromatographic needs. This supports the earlier conclusion that neglective column design will lead to a dramatic loss of performance due to the sensitivity of short-bed adsorbents to extracolumn effects. A series of anion exchange and hydrophobic interaction was presented by Luksa et al. for the purification of recombinant tumor necrosis factor α from clarified *E. coli* extract (64). Comparison of the results obtained with commercially available cartridges of type B (Quick Disk, Knauer, Berlin, Germany) with a conventional method based on porous particles showed a similar product purity at half the process time, thus demonstrating that increased throughput at retained resolution may be obtained with adsorptive membranes. The performance of another commercially available stack of anion exchange membranes (Memsep DEAE, Millipore, Bedford, Mass.) was evaluated by Gerstner et al. (47). The authors were able to separate three component mixtures of standard proteins at a productivity of 58 g/(L h). Reif et al. published the use of immobilized metal ion chromatography on membrane stacks for the purification of a $(His)_6$-tagged recombinant protein (65). In this example it could be shown that, contrary to ion exchange membrane chromatography, the slow association kinetics of the pseudoaffinity interaction may limit the throughput with regard to resolution. Additionally it was shown that step gradients give better results in protein separation on membrane adsorbents than linear gradients—a universal phenomenon that may be due to the fact that the short membrane bed is not sufficient to allow the development of the self-sharpening effects of the linear gradient, which favorably influences the resolution in particle-based chromatography. The separation of whey proteins in strong ion exchange membranes was reported by Splitt et al. (66). In this work the influence of bed length was shown in the separation of the two genetic variants of β-lactoglobulin from whey, which became possible when a minimum bed length of 2 mm was chosen by the coupling of commercially available flat-sheet units. The capture of whey proteins from diluted solutions was achieved at a productivity of 45.5 g/(L h) with more than 60-fold reduction of the feed volume. Champluvier and Kula used Dye ligand membranes to isolate dehydrogenases from clarified yeast extracts. A productivity of 150,000 U/(L h) was found for glucose-6-phosphate dehydrogenase with Cibacron blue immobilized to nylon flat-sheet membranes (37). In this case the throughput was described to be limited by the kinetics of the association reaction. In a sequential arrangement of ion exchange and dye-ligand-modified membranes, formate dehydrogenase was purified with a productivity between 24 and 37 g/(L h). The protein was purified 4.6-fold with 27% overall yield by this the sequence; the eluate from one step was directly applied to the next membrane without intermediate buffer exchange (51). An interesting application was

reported by van Huynh et al. with the isolation of plasmid DNA from crude cell lysates using stacks of flat-sheet membranes (DEAE Memsep, Millipore, Bedford) (67). Because these comparatively large bioproducts provide some challenge to conventional particulate matrices due to their size, membranes with throughgoing pores are an interesting alternative for the processing of these potentially large-scale products (e.g., for gene therapy).

Table 1 summarizes the productivity obtained during the applications described here. This selection from a much larger number of low-resolution purification operations described in the literature shows that membrane adsorbents can be used efficiently at this stage of a downstream process. They also elucidate, that the scale-up of this method must be based on solid considerations with regard to the needs of an adsorption process, otherwise the impressive results on the laboratory-scale with regard to high throughput at retained resolution are not transferable to the process scale.

High-Resolution Purification

Membrane-shaped stationary phases were introduced into high-resolution purification with the concept of high-performance membrane chromatography (HPMC) (23,31, 52,59,68). By choosing continuous stationary phases of a membrane type of geometry, separations could be performed at resolutions comparable to HPLC methods. The improved transport characteristics of the continuous phase allowed higher throughput in spite of the low L/d ratio, which was used to obtain reduced pressure drop along the columns. A thorough analysis of the parameters influencing the performance of HPMC was provided by Tennikova and Svec (23). An interesting result of their investigations was an increased resolution of anion exchange HPMC during the separation of standard proteins, with increasing flow velocity; the separation in hydrophobic interaction chromatography (HIC) mode was not affected. Compared with low-resolution purification employing adsorptive membranes, step gradients yielded better separations than a linear gradient.

Attaching affinity ligands to membranes offers another opportunity to use them in high resolution purification. As shown by Nachman et al. (39–41), affinity membrane adsorbents are well suited for the fast purification of complex molecules from diluted feedstock. By coupling of receptors or antibodies to commercial hollow-fiber modules (Sepracor, Marlboro, Mass.) automatic separations were performed at high productivity (2–6 g/(L h)). Similar results were obtained with protein A–modified hollow-fiber modules from the same manufacturer (34). A dimensionless analysis of IgG adsorption to this matrix revealed a kinetic limitation of the sorption efficiency; breakthrough capacity increased with t_C but was independent from C_0 (69). As an alternative to hollow fibers, flat-sheet membranes were modified with protein A by Langlotz et al. (53) and were used to purify IgG from hybridoma supernatant. In this work it could be demonstrated that high resolution by affinity ligands and high productivity may be obtained at the same time (10.8 g/(L h)).

CONCLUSION

From the existing literature some important features of adsorptive membranes may be extracted, making them a very interesting candidate for stationary phases for the process-scale purification of proteins. Despite this fact, large-scale applications of the technology have not been undertaken. This may partly due to the traditional reluctance to use new matrices for the industrial-scale purification of proteins, where the proof of long-term stability and reproducibility is essential for the success of the downstream process. Additionally, creative solutions for the practical realization of large-scale units with very low L/d have not been developed. At this point the manufacturers have to prove that the demands of an adsorptive process with regard to regular fluid flow, reduced extra column effects, and so on can be served on the process scale. Currently there are only few examples of pilot-scale processes available in the literature; it is hoped that the near future will bring many more successful applications of adsorptive membranes. The technology certainly has potential for process-scale separations, as has been demonstrated by theoretical analysis and laboratory-scale experiments.

Table 1. Some Typical Productivity Values Obtained with Membrane Adsorbents

Ligand	Geometric format	Target protein	Productivity (g/(L h))[a]	Ref.
Q	Stack	Formate dehydrogenase	36	51
Q	Pleated module	Antithrombin III	21.5	58
Q	Stack	HSA	220	20
DEAE	Hollow fiber	Urease	200	17
S	Pleated module	IgG	72.5	58
S	Stack	IgG	29	62
MAb	Hollow fiber	Recombinant interferon α2a	2.7	40
Histidin	Hollow fiber	IgG	6.7	46
Procion yellow	Stack	Pyruvate decarboxylase	40.3	38
Procion red	Stack	Formate dehydrogenase	24	51
Cibacron blue	Cross-flow module	Malate dehydrogenase (from unclarified broth)	1.1	43
Protein A	Hollow fiber	IgG	35	34
Protein A	Stack	IgG	10.8	53

Productivity was calculated from the cited references by calculating the total time of the respective purification cycles (including washing and elution, when available) and relating the amount of product purified to the cycle time and the adsorbent volume used.

NOMENCLATURE

a	Specific surface area (m^2/m^3)
α	Number of bed volumes required for wash, elution and regeneration during chromatographic cycle
B	Permeability of a packing (m^2)
C	Initial concentration (kg/m^3)
D	Diffusion coefficient (m^2/s)
D_{ax}	Axial dispersion coefficient (m^2/s)
d	Adsorbent bed diameter (m)
d_p	Pore diameter (m)
ϵ	Void fraction (−)
η	Dynamic viscosity (kg/(m s))
HETP	Height equivalent of a theoretical plate (m)
k_a	Second-order association rate constant (l/(mol s))
k_d	Dissociation rate constant (s^{-1})
L	Adsorbent bed length (m)
L/d	Column aspect ratio (−)
n	Plate number in HETP concept (−)
P	Productivity (g/(L h))
Pe	Peclet number (−)
Q_{max}	Static capacity (kg/m^3)
Q_D	Breakthrough capacity (kg/m^3)
R_E	Recovery ratio
R_S	Resolution
t_C	Residence time (s)
t_D	Diffusion time (s)
t_R	Reaction time (s)
V	Adsorbent volume (m^3)
v	Flow velocity (m/s)
v_e	Interstitial flow velocity (m/s)

BIBLIOGRAPHY

1. J.C. Janson and T. Peterson, in G. Ganetsos and P. E. Barker ed., *Preparative and Production Scale Chromatography*, Marcel Dekker, New York, 1993, pp. 559–590.
2. E. Boschetti, L. Guerrier, P. Girot, and J. Horvath, *J. Chromatogr. B* **664**, 225–231 (1995).
3. N.B. Afeyan, S.F. Fulton, I. Mazsaroff, and F.E. Regnier, *Bio / Technology* **8**, 203–206 (1990).
4. S. Brandt, R.A. Goffe, S.B. Kessler, J.L. O'Connor, and S.E. Zale, *Bio / Technology* **6**, 779–782 (1988).
5. D.K. Roper and E.N. Lightfoot, *J. Chromatogr. A* **702**, 3–26 (1995).
6. J. Thömmes and M.-R. Kula, *Biotechnol. Progr.* **11**, 357–367 (1995).
7. P. Langlotz, S. Krause, and K.H. Kroner, *Filtrieren und Separieren* **5**, 62–70 (1991).
8. B. Champluvier and M.-R. Kula, *J. Chromatogr.* **539**, 315–325 (1991).
9. K. Saito and M. Ito, *Ind. Eng. Chem. Res.* **28**, 1808–1812 (1989).
10. H. Iwata, K. Saito, and S. Furusaki, *Biotechnol. Progr.* **7**, 412–418 (1991).
11. M. Kim, K. Saito,S. Furusaki, T. Sugo, and I. Ishigaki, *J. Chromatogr.* **586**, 27–33 (1991).
12. M. Kim, K. Saito, S. Furusaki, T. Sato, T. Sugo, and I. Ishigaki, *J. Chromatogr.* **585**, 45–51 (1991).
13. H. Shinano, S. Tsuneda, K. Saito, and S. Furusaki, *Biotechnol. Prog.* **9**, 193–198 (1993).
14. K. Kobayashi, S. Tsuneda, K. Saito, H. Yamagishi, S. Furusaki, and T. Sugo, *J. Membrane Sci.* **76**, 209–218 (1993).
15. S. Tsuneda, H. Shinano, K. Saito, S. Furusaki, and T. Sugo, *Biotechnol. Progr.* **10**, 76–81 (1994).
16. S. Tsuneda, K. Saito, S. Furusaki, and T. Sugo, *J. Chromatogr. A* **689**, 211–218 (1995).
17. S. Matoba, K. Tsuneda, K. Saito, and T. Sugo, *Biotechnology* **13**, 795–797 (1995).
18. N. Kubota, M. Kounosu, K. Saito, K. Sugita, K. Watanabe, and T. Sugo, *J. Chromatogr. A* **718**, 27–34 (1995).
19. S. Tsuneda, K. Saito, T. Sugo, and K. Makuuchi, *Radiat. Phys. Chem.* **2**, 239–245 (1995).
20. K.H. Gebauer, J. Thömmes, and M.-R. Kula, *Biotechnol. Bioeng.* **54**, 181–189 (1997).
21. K.H. Gebauer, J. Thömmes, and M.-R. Kula, *Chem. Eng. Sci.* **52**, 405–419 (1996).
22. W.K. Wang, S.-P. Lei, H.G. Monbouquett, and W.C. McGregor, *Biopharm*, 52–59 (June 1995).
23. T.B. Tennikova and F. Svec, *J. Chromatogr.* **646**, 279–288 (1993).
24. N. Kubota, Y. Konno, S. Miura, K. Saito, K. Sugita, K. Watanabe, and T. Sugo, *Biotechnol. Progr.* **12**, 869–872 (1996).
25. D.D. Frey, R. van de Water, and B.R. Zhang, *J. Chromatogr.* **603**, 43–47 (1992).
26. S.Y. Suen and M.R. Etzel, *Chem. Eng. Sci.* **47**, 1355–1364 (1992).
27. S.Y. Suen, M. Caracotsios, and M.R. Etzel, *Chem. Eng. Sci.* **48**, 1801–1812 (1993).
28. S. Yamamoto, *Recovery of Biological Products VIII*, Tucson, Ariz., October 20–25, 1996.
29. J.L. Coffman, D.K. Roper, and E.N. Lightfoot, *Bioseparation* **4**, 183–200 (1994).
30. D.K. Roper and E.N. Lightfoot, *J. Chromatogr. A* **702**, 69–80 (1995).
31. D. Josic and J. Reusch, *J. Chromatogr.* **590**, 59–76 (1992).
32. O.W. Reif and R. Freitag, *J. Chromatogr. A* **654**, 29–41 (1993).
33. H.-C. Liu and J.R. Fried, *AIChE J.* **40**, 40–48 (1994).
34. E. Klein, E. Eichholz, and D.H. Yeager, *J. Membrane Sci.* **90**, 69–80 (1994).
35. I.A. Adisaputro, Y.-J. Wu, and M.R. Etzel, *J. Liq. Chrom.* **19**, 1437–1450 (1996).
36. S.-Y. Suen and M.R. Etzel, *J. Chromatogr. A* **686**, 179–192 (1994).
37. B. Champluvier and M.R. Kula, *Biotechnol. Bioeng.* **40**, 33–40 (1992).
38. K.G. Briefs and M.R. Kula, *Chem. Eng. Sci.* **47**, 141–149 (1992).
39. M. Nachman, *J. Chromatogr.* **597**, 167–172 (1992).
40. M. Nachman, A.R.M. Azad, and P. Balion, *Biotechnol. Bioeng.* **40**, 564–571 (1992).
41. M. Nachman, A.R. Azad, and P. Balion, *J. Chromatogr.* **597**, 155–166 (1992).
42. S.M. Wheelwright, *J. Biotechnol.* **11**, 89–102 (1989).

43. S. Krause, K.H. Kroner, and W.D. Deckwer, *Biotechnol. Techn.* **5**, 199–204 (1991).
44. S. Le Borgne, M. Graber, and J.S. Condoret, *Bioseparation* **5**, 53–64 (1995).
45. M.H. Heng and C.E. Glatz, *Biotechnol. Bioeng.* **42**, 333–338 (1993).
46. S.M.A. Bueno, K. Haupt, and M.A. Vijayalakshmi, *J. Chromatogr. B* **667**, 57–67 (1995).
47. J.A. Gerstner, R. Hamilton, and S.M. Cramer, *J. Chromatogr.* **596**, 173–180 (1992).
48. W.F. Weinbrenner and M.R. Etzel, *J. Chromatogr. A* **662**, 414–441 (1994).
49. R. Freitag, H. Splitt, and O.-W. Reif, *J. Chromatogr. A* **728**, 129–137 (1996).
50. A. Jungbauer, F. Unterluggauer, F. Steindl, F. Ruker, and H. Katinger, *J. Chromatogr.* **397**, 313–320 (1987).
51. B. Champluvier and M.R. Kula, *Bioseparation* **2**, 343–351 (1992).
52. T.B. Tennikova, M. Bleha, F. Svec, T.V. Almazova, and B.G. Belenkii, *J. Chromatogr.* **555**, 97–107 (1991).
53. P. Langlotz and K.H. Kroner, *J. Chromatogr.* **591**, 107–113 (1992).
54. Z. Kucerova and J. Turkova, *IJBC* **2**, 145–151 (1997).
55. S. Yamamoto and Y. Sano, *J. Chromatogr.* **597**, 173–179 (1992).
56. N.I. Dubinina, O.I. Kurenbin, and T.B. Tennikova, *J. Chromatogr. A* **753**, 217–225 (1996).
57. K.G. Briefs and M.R. Kula, in M. Ladisch and R. Bose eds., *Harnessing Biotechnology for the 21st Century*, American Chemical Society, Washington D.C., 1992, pp. 258–261.
58. R. Luetkemeyer, M. Bretschneider, H. Buentemeyer, and J. Lehmann, *J. Chromatogr.* **639**, 57–66 (1993).
59. F. Svec and T.B. Tennikova, *J. Bioact. Comp. Polymers* **6**, 393–405 (1991).
60. E. Klein, *Affinity Membranes*, Wiley, New York, 1991.
61. A. Jungbauer, F. Unterluggauer, K. Uhl, A. Buchacher, F. Steindl, D. Pettauer, and E. Wenisch, *Biotechnol. Bioeng.* **32**, 326–333 (1988).
62. D. Luetkemeyer, N. Ameskamp, H. Tebbe, K. Bracht, and J. Lehmann, in M.J.T. Carrondo ed., *Animal Cell Technology: From Vaccines to Genetic Medicine*, Kluwer Academic, 1997, pp. 325–329.
63. J.F. Lacoste-Bourgeacq, C. Desneux, and M. Allary, *Chromatographia* **532**, 27–32 (1991).
64. J. Luksa, V. Menart, S. Milicic, B. Kus, V. Gaberc-Porekar, and D. Josic, *J. Chromatogr. A* **661**, 161–168 (1994).
65. O.-W. Reif, V. Nier, U. Bahr, and R. Freitag, *J. Chromatogr.* **664**, 13–25 (1994).
66. H. Splitt, I. Mackenstedt, and R. Freitag, *J. Chromatogr. A* **729**, 87–97 (1996).
67. N. van Huynh, J.C. Motte, J.F. Pilette, M. Decleire, and C. Colson, *Anal. Biochem.* **211**, 61–65 (1993).
68. D. Josic, Y.-P. Lim, A. Strancar, and W. Reutter, *J. Chromatogr. A* **662**, 217–226 (1994).
69. C. Charcosset, Z. Su, S. Karoor, G. Daun, and C.K. Colton, *Biotechnol. Bioeng.* **48**, 415–427 (1995).

See also CHROMATOGRAPHY, COMPUTER-AIDED DESIGN; CHROMATOGRAPHY, ION EXCHANGE; CHROMATOGRAPHY, SIZE EXCLUSION; PROTEIN PURIFICATION, AQUEOUS LIQUID EXTRACTION.

MEMBRANE SEPARATIONS

MANOHAR KALYANPUR
Consultant, Bioseparations and Pharmaceutical Validation
Plaisir, France

KEY WORDS

Asymmetric membranes
Cross-flow filter
Diafiltration
Depyrogenation
Microfiltration
Nanofiltration
Permeate
Retentate
Tangential-flow filter
Ultrafiltration

OUTLINE

INTRODUCTION

The term *filtration* refers to the separation of two or more components from a liquid or gas. Generally, it refers to the separation of solids from either liquids or gases. In membrane separations, one goes beyond the classical separation of solids from the other phases. Membranes can be and are used to separate dissolved solutes in a solution from one another and also to remove the solutes from the liquid in which they are dissolved. In this manner, membranes are used to either purify products or to concentrate them. This is particularly advantageous in biotechnology where the products of biological origin are very often protein-

aceous in nature and are heat labile. The membranes offer a means of separation that occurs at or below ambient temperatures. This gives the biotechnologist a separation or purification technique that provides maximum assurance of product stability. In membrane-based separations, the membrane acts as a barrier to the passage of certain components while others move freely across the membrane. Therefore, the end result of a simple membrane separation is the concentration in the upstream fraction of the components that do not cross the membrane barrier. This fraction is referred to as the *retentate*. The fraction that passes across the membrane is called the *permeate* or *filtrate* and contains the components of the solution in essentially the same concentration as the starting fluid. This is the simple basic *membrane separation*.

THREE MAJOR APPLICATIONS OF MEMBRANE SEPARATIONS

Product Purification

Product purification is the most important application for membrane separations in biotechnology. Membranes are made of different materials, and the biotechnologist needs to make a choice based on the goal of the application and the suitability of the membrane for the application. This article aims at providing the biotechnologist with information that will help select the best membrane and the manner in which to use the membrane to accomplish the separation and purification task easily. Membrane separations are an important part of both the upstream and downstream processing in biotechnology. A typical process can include a number of steps, such as centrifugation followed by membrane-based clarification of the supernatant, one or more chromatographic procedures, final membrane processing steps to accomplish further purification, and finally concentration.

Sterile Filtration

One of the most critical steps in the manufacture of injectable therapeutic agents is the sterilization of the product. The simplest method of sterilizing solutions is the one that uses heat, and autoclaving is a typical example of such methods. However, products of the biotechnology industry are most often thermolabile and therefore cannot be heat sterilized. Membranes provide an excellent alternative. Membrane filters are used for filtering out bacteria and other microorganisms from the final formulations, rendering the products very safe for injection. This step is followed by lyophilization of products that are marketed as solid powders. Those products marketed as liquid formulations are filled straight into ampules, vials, or other appropriate containers after sterile filtration through membrane filters.

Virus Removal by Membrane Filtration

This is the most recent application of membrane separations in the biopharmaceutical industry. A new generation of membranes is today used to remove viruses from biological preparations. Therapeutic agents manufactured from blood, by cell culture processes, or by the most modern method using transgenic animals can contain dangerous viral contaminants from the source materials. These products can be filtered through membrane devices to remove most viral contaminants, including HIV and the viruses known to cause different forms of hepatitis.

There is a point of differentiation here from the previously mentioned sterile filtration in which there is complete or total removal of microbiological contaminants from the solution. In virus filtration, the removal is not complete, especially when the viruses are very small. Therefore, for complete viral safety, the practice is to use membrane filtration as a method complementary to the longer known generic methods of viral inactivation, such as treatment with a combination of solvents and detergents, other chemical inactivating agents, and pasteurization by mild heating for prolonged periods of time.

CLASSIFICATION OF MEMBRANES AND MEMBRANE PROCESSES

The more commonly used membrane processes in the biotechnology industry are microfiltration, ultrafiltration, and reverse osmosis. All three are pressure-driven processes and differ from each other in the size of the components they retain. The membrane itself, under the influence of applied pressure, controls the retention and passage of individual components of the solution being processed.

Microfiltration

Microfiltration processes retain particles in the submicron range that are suspended in the process fluid. The membranes are also used in removing particles and microorganisms from air and gases, as in the fermentation processes in biotechnology and other industries. The particles removed by membrane filtration range in size from 0.1 μm to 10 μm. Conventional filtration processes such as those using pad filters and filter presses in the more conventional pharmaceutical industry are generally capable of retaining particles only above the 10 μm size.

Ultrafiltration

Ultrafiltration membranes generally retains macromolecules or particles from about 0.001 to 0.2 μm in size. By choosing one or more of the available membranes, one can separate larger macromolecules from smaller ones within the above size ranges. For example, separation of large proteins or polysaccharides from smaller molecules is a typical application for ultrafiltration membranes in the biotechnology industry.

Reverse Osmosis

Reverse osmosis is one of the oldest known applications for membrane separations. This is the process where membranes are capable of retaining even smaller molecules than ultrafiltration membranes. Salts and other molecules with a molecular weight below 1,000 Da that are not normally retained by ultrafiltration membranes are retained

by the reverse osmosis membranes. The first use of these membranes was in desalting sea water. The membranes are today used for concentrating small molecules, and ultrafiltration is used for fractionation and purification of macromolecules with molecular weights between 1,000 and 1 million Da. Figure 1 summarizes the use of membrane filters in biotechnology. Table 1 gives examples of products that can be retained by the three types of membranes. In a typical biotechnology process, microfiltration is the first membrane-based step, and it is used to separate suspended material such as cells from a bioreactor and other particulate matter, including cell debris, from the multitude of dissolved substances, which very often includes the specific product of interest expressed by the cell line. In many instances, clarifying filters made of a variety of materials are used for preclarification of the process stream before microfiltration. This usually helps to improve the performance of the microfiltration step and extends the life of the membrane filter. Honig and Schwartz (1) have written an excellent review of this subject. The filtrate or permeate from the microfiltration step contains the desired product elaborated by the process along with other proteins, carbohydrates, and smaller molecules such as salts from the culture medium as well as those produced during the process. To separate the product of interest from these contaminants, the ideal membrane separation step is ultrafiltration. Membranes with the appropriate nominal molecular weight cutoff, a term used to identify and differentiate ultrafiltration membranes from one another, are next used in successive purification steps to facilitate recovery of the desired product of high purity. Quite often, the manufacturing process also includes chromatographic or other suitable process steps in between membrane separations to achieve an overall process advantage in terms of product quality and yield. Figure 2 illustrates how membrane and chromatography procedures are integrated together step-by-step for upstream and downstream processing in the biotechnology industry.

MEMBRANE CHEMISTRY, STRUCTURE, AND FUNCTION

The high-volume microfiltration in the biopharmaceutical industry requires removal of particulate and colloidal substances in an economical and efficient manner. The nature of these particles and the interplay of filtration dynamics contribute to the difficulty of the task. It is important to balance filter performance against process costs to arrive at a system that provides the ideal separation solution at a low operating cost. In general, the filtration process starts out with passing the process stream through a porous medium that entraps the solids in its matrix or retains them on its surface. The materials used in the manufacture of these filters are porous or fibrous, such as paper, felt, woven yarns, and fabrics, and other materials. Depending on the basic filtration mechanisms, particle retention mode and construction characteristics, the different filters used in the biologicals industry are classified as *depth filters*, *surface filters*, and *screen filters*. Figures 3, 4, and 5 show the mechanism of retention of particulates by these three types of filters.

Depth filters consist of fibrous, granular, or sintered materials that produce a random porous structure. When a

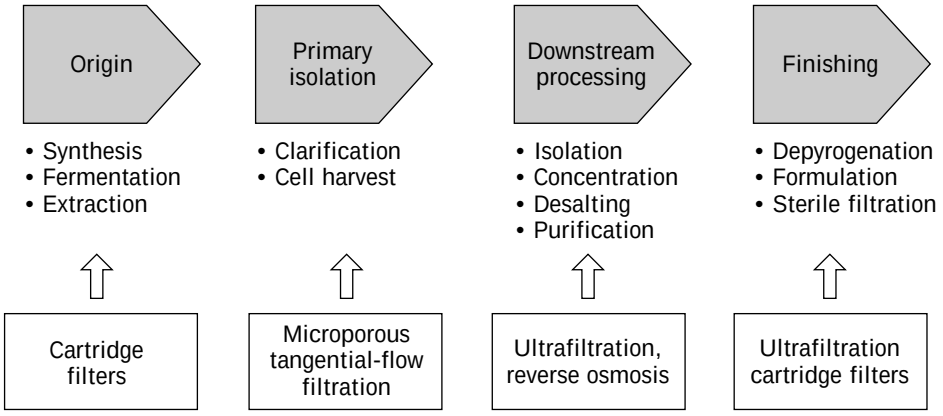


Figure 1. Different steps in downstream processing where membrane separations are used.

Table 1. Products Retained by Microfiltration, Ultrafiltration, and Reverse Osmosis Membranes

Size	Product		Membrane process
100 μm	Pollen	] →	Microfiltration
10 μm	Starch granules, blood cells, bacteria		
1 μm	Smaller organisms		
50,000–100,000 D	Albumin, immunoglobulins, enzymes, other proteins, large oligopeptides	] →	Ultrafiltration
10,000–100,000 D			
1000–10,000 D			
<1000 D	Gluose, small peptides, salts	]→	Reverse osmosis

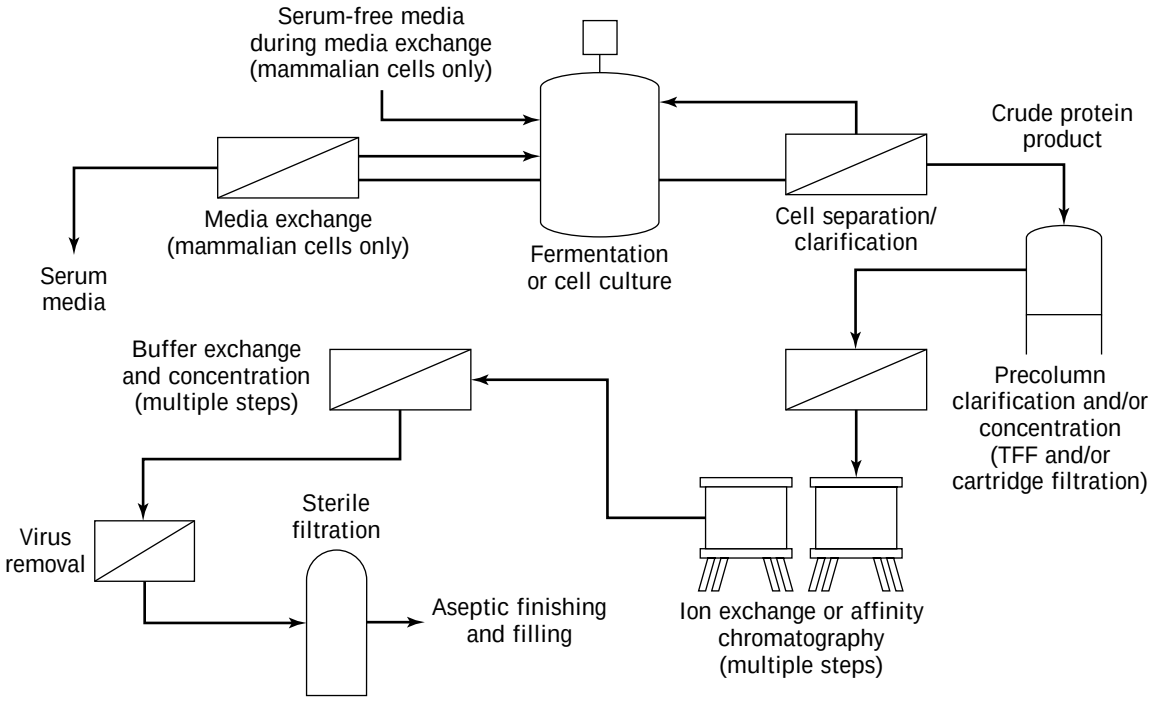


Figure 2. Integration of membrane separations with other techniques in downstream product recovery operations in biotechnology.

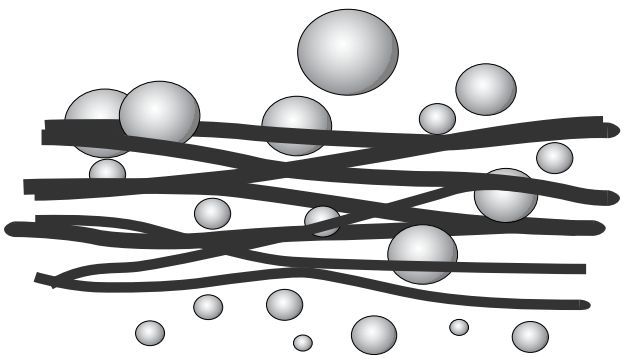


Figure 3. Mechanism of particle retention by depth filters.

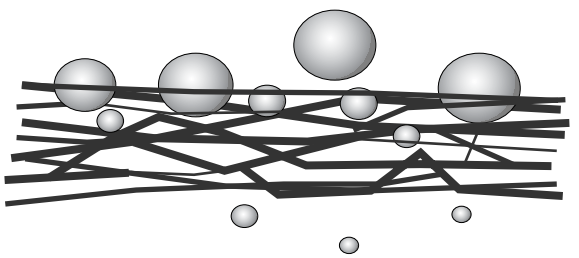


Figure 4. Mechanism of particle retention by surface filters.

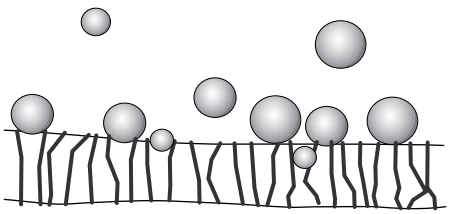


Figure 5. Mechanism of particle retention by screen filters.

fluid passes through a depth filter, particles become entrapped in its tortuous network of flow channels. Particle retention mechanisms are random adsorption and mechanical entrapment throughout the depth of the filter matrix. The filtration media of these filters can be wound cotton, rayon, or polypropylene as well as resin bonded laminates such as cellulose and fiberglass.

Surface filters are made up of several layers of media, such as glass or polymeric microfibers. During filtration, particles larger than the spaces within the filter matrix are retained primarily on the filter surface. Other smaller particles can get entrapped within the matrix. Thus, these filters have the property of both depth and screen filters. The surface filters are generally made of polymeric materials, resin-bonded paper, or fiberglass and paper.

Screen filters have a geometrically regular porous matrix, and particles are retained on its surface primarily by a sieving or size exclusion mechanism. Small particles and

microorganisms larger than the pore size are retained on the surface. The filtration process here is absolute because everything larger than the rated pore size cannot get through the filter. Examples of such filters include cast polymeric membranes used for all critical submicronic and macromolecular separations in the biotechnology industry. The structure of the filter material is more rigid, and the manufacture is much better controlled than for the other types of filters.

Screen filters made of *microfiltration membranes* have a well-defined pore size in the range of 0.1 to 1 μm. Because particles and microorganisms are retained at or near the surface of the membrane, there is very little possibility of the so-called grow through phenomenon taking place when these filters are used for sterile filtration applications in the manufacture of parenteral drugs. This is a truly major application for the microfiltration membranes in the general pharmaceutical and biotechnology industries. They also find use in another biotechnology application, the harvesting of cells from bioreactors.

Ultrafiltration membranes are the second major group of screen filters and are often called *asymmetric* or *skinned* membranes. The skin of the membrane does the job of retaining the dissolved solutes larger than the membrane rating. The skin is supported by a more rigid support layer or backing. Unlike microfiltration membranes, ultrafiltration membranes are not designated by their pore size but according to the molecular weights of the dissolved solutes that the membrane will *retain* or *reject*. This, as mentioned before is called the nominal molecular weight cutoff of the membrane.

METHODS OF MEMBRANE MANUFACTURE

Most membranes are made of organic polymers such as the natural polymer, cellulose, and its derivatives (e.g., cellulose acetate) and certain other synthetic polymers (e.g., polyethylene, polycarbonate, polysulfone) and their derivatives. Some membranes are also made of inorganic materials such as zirconium and aluminium oxide. These are often referred to as *ceramic membranes*.

In the manufacture of polymeric membranes, the organic polymers are dissolved in an appropriate solvent and combined with the nonsolvent or swelling agent. This makes up the membrane casting solution. The membranes are cast on casting machines, which are almost always made by the membrane manufacturers themselves. Most membranes are made in the form of flat sheets, but some are cast on the module support as in the case of tubular module membranes. Some membranes are also made directly in the form of hollow fibers. The membranes are then annealed.

Flat sheet membranes are subsequently incorporated into flat membrane plate and frame modules or made into spirally wound or pleated cartridges. Critical details of membrane manufacture are well-guarded secrets and are proprietery processes developed and owned by membrane manufacturers.

Microfiltration membranes are made to a thickness of about 100 to 120 μm, and the retention of particulate material takes place at the membrane surface or within the first few micrometers of the tortuous length of the membrane pores. *Ultrafiltration membranes* consist of a very thin skin that performs the task of separation of the solutes in solution. This skin is supported by a much thicker support material that contains voids, often referred to as finger voids because of their fingerlike shape. These voids facilitate the rapid flow of the fluid once it gets past the fine skin of the ultrafiltration membrane. *Reverse osmosis membranes*, as per their classical definition, are meant to retain all solutes other than the solvent itself. These were originally developed for sea water desalination to make potable water. This is still by far the largest industrial application for reverse osmosis membranes. These membranes have since been made into somewhat modified versions that can retain smaller molecular weight solutes that are generally not retained by the ultrafiltration membranes with a low nominal molecular weight cutoff. The molecules retained are larger than the salts that can pass more easily through the membrane. These new-generation reverse osmosis membranes find a niche application in the concentration of low molecular weight solutes, such as peptides, antibiotics, and other products in the biopharmaceutical industry. Table 2 summarizes the characteristics of the three membrane processes described.

Track-Etch Membranes

As described earlier, most polymeric membranes contain tortuous pores in a spongy material. The pores of these membranes can have a fairly wide pore size distribution as a result of their method of manufacture. A different type of membrane is made by a two-step process that includes the manufacture of the polymer sheet membrane of a certain thickness followed by bombardment of the membrane sheet with massive energetic nuclei in a beam of ^{235}U fission fragments. These nuclei go through the membrane, leaving tracks that are then the pores formed in the membrane. After the bombardment step, the membrane undergoes further process steps to make it antistatic and to modify its surface characteristics such as either hydrophobicity or hydrophilicity. It is natural and quite logical to assume that such a process of making pores in a membrane would produce a membrane with highly uniform pore diameters and with a sharp nominal molecular weight cutoff. However, in actual practice, the random bombardment can as well result in two or more pores being formed close together, thus giving rise to much larger pores than the individual pores. This problem of overlapping pores is supposedly avoided if the nuclear bombardment takes place at an angle of 29° between the particle beam and the membrane sheet (2). There are also some other disadvantages to these membranes. Their extreme thinness of only about 15 μm makes them difficult to handle. Also, if the bombarding particles fail to penetrate completely through the membrane sheet, the apparent pores in the membrane are not true pores.

Virus Removal Membranes

The biotechnology industry, which has benefitted more from membrane processes than the conventional phar-

Table 2. Characteristics of Membrane Processes

Process	Permeate	Retentate
Microfiltration[a]	Solvent with all dissolved solutes	Particulates in suspension, colloidal material
Ultrafiltration[b]	Water with dissolved solutes of lower molecular weight	Dissolved solutes of higher molecular weight and colloidal material
Reverse osmosis	Solvent, typically water	Solutes

[a]In microfiltrator, particulate retention generally depends on the pore size of the membrane.
[b]In ultrafiltration, the passage or retention of solutes depends on the nominal molecular weight cutoff of the membrane.

maceutical industry, has constantly changing requirements. These are the result of the development of new biologicals by newer and rather very innovative processes as well as the more and more stringent regulatory requirements that call for high product quality and safety. An excellent example here is the need for biologicals to be rendered free of both relevant and adventituous viral contaminants. The membrane manufacturers who have regularly kept up with the demands of the biotechnology industry have come up with special membranes for this application that are being used by the industry to make virus-free biologicals.

HOW MEMBRANE PROCESSES ARE OPERATED

Membrane processes are driven by pressure, which is a great advantage to the biotechnology industry, in which large majority of products are heat labile. The processes also consume less energy because they do not require either heat transfer or heat generation. They call for a pump to run the process fluid through the membrane under pressure, as illustrated in Figure 6. Some of the processes can, however, be run with either compressed air or nitrogen and do not need pumps. Most membrane processes are run at ambient temperature, with few exceptions such as those processes where extremely heat-labile products call for the processes to be run at lower-than-ambient temperatures. Low temperature membrane processing is also common when the process time is long, and there is the accompanying risk of product degradation or microbial contamination of the process fluid. At the other extreme, higher temperatures are used when a fluid is very viscous, and the product of interest is not quite heat sensitive. The higher temperature reduces the viscosity of the process stream and increases the rate of filtration with the overall advantage of significant reduction in process time.

In a well-developed process in the biotechnology industry, it is quite common to find membrane filtrations operating in tandem with one or more chromatographic purification steps or other methods of separation and purification. Further, one may also use more than one membrane separation step (i.e., microfiltration, ultrafiltration, and reverse osmosis) at appropriate points in the downstream process just as different chromatography procedures, such as ion exchange, size exclusion, or affinity methods, are used in a complementary fashion. One single membrane separation step is not always enough to accomplish the end result. Centrifugation is sometimes used before the membrane processes for clarification of fluids containing a high biomass, as in fermentation processes. The supernatant of the centrifugation is then better handled by other methods of purification, including membrane-based processes.

Two Methods of Operation of Membrane Processes

The processes are run according to two distinctly different methods. Figure 7 illustrates these two methods of filtration. The first is the classical method of filtering a solution containing suspended particulates either to recover the particulates themselves, as in the recovery of a product crystallized from a solution, or to obtain a clear filtrate when the product of interest is in the filtrate. This method is known as *dead-end filtration* or *normal-flow filtration*.

When the solution being filtered contains a high concentration of particulates or bioburden, the filter can quickly get clogged because of the accumulation of solids at the surface of the membrane. This usually ends up with the deposition of a boundary layer at the membrane surface, which then acts as a secondary membrane. This prevents free passage of the liquid and eventually the mem-

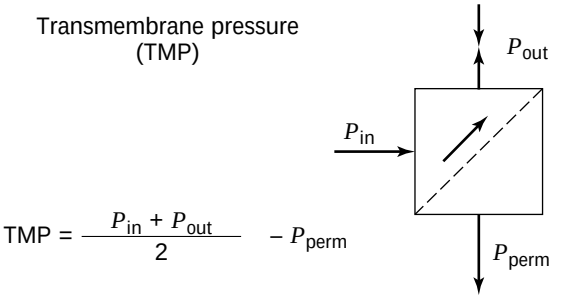


Figure 6. Basic scheme of a tangential-flow filtration process.

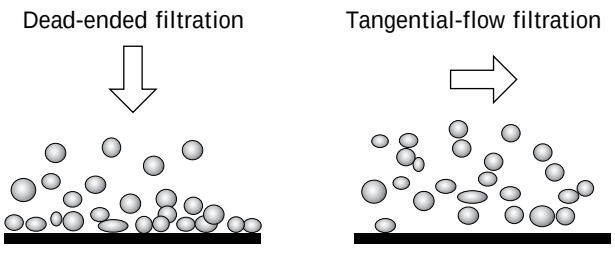


Figure 7. The two methods of using membrane filters.

brane is completely clogged. When this happens, filtration cannot be continued, and the only solution is to stop the process and replace the membrane filter with a new one. Such stoppages are time consuming and often result in loss of valuable product, which affects the economics of the overall process.

A practical and often successful alternative to the classical dead-end filtration is the use of the membrane filters by a method known as *tangential-flow filtration* or *cross-flow filtration*. Figure 6 illustrates the circulation of the complete process fluid in a direction parallel to the membrane surface.

Different types of pumps are used in the process. The advantage of the method is that the particulates that would accumulate at the membrane surface in dead-end filtration are now kept in the body of the fluid. They stay away from the membrane surface, permitting the liquid to flow relatively unhindered through the membrane pores. If the rate of filtration or *flux* drops significantly, the process fluid can be diluted by adding an appropriate pure solvent such as water or a buffer. This procedure helps to recover much of the original flux, and the filtration can be continued as before. The addition of the diluting fluid can take place without stopping the process. This step of *in-process dilution* is referred to as *diafiltration*. A typical tangential flow filtration process scheme is illustrated in Figure 8. If the desired product is in the filtrate, which is now more dilute, the filtrate can be reconcentrated to the desired concentration by either ultrafiltration or reverse osmosis, choosing the appropriate membrane according to the molecular weight of the product of interest. The concentration and further purification can also be performed by a chromatographic method.

Equipment Employed in Membrane Filtration

Membrane filters in different configuration are required for different processes depending on the method of filtration used. The choice of a particular type of filter depends on several considerations, such as process fluid composition; goal of the process, which would either be to recover the particulates or the filtrate; and, of course, the economics of the process.

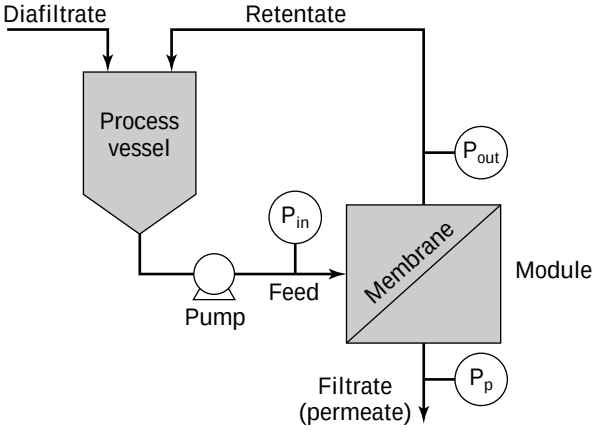


Figure 8. Schematic of a tangential-flow filtration operation (clarification, concentration, and diafiltration).

Equipment for Microfiltration in the Dead-End or Normal-Flow Filtration Mode For clarifying small volumes of solutions as in a research and development environment, small filter disks of different diameters and pore sizes are used. The filter disks are placed in appropriate holders of plastic materials or stainless steel. The stainless steel holders can be autoclaved for sterility before the filtration and can be used over and over again.

These less-expensive filters are also quite useful in determining what type of filters are appropriate for a certain application. Filters are also available in the form of fabricated devices that can be placed at the end of a syringe. Such filters are ideal for filtering volumes from 1 mL to about 100 mL. For applications where large volumes of process fluids need to be filtered, as in an industrial environment, cartridge filters placed in suitable stainless steel holders are ideal. The filters placed in such holders are easy to handle; thus can be autoclaved or steam sterilized and sometimes cleaned for reuse.

Quite often, the particulate level in filtrations is high, and the membrane filters can quickly get clogged. The common solution to this problem is to place in the filtration step prefilters and clarifying filters before the membrane filters. The economic advantage of extending the life of the usually more-expensive final filter is far greater than the cost of the prefilters. In well-designed processes, the filters are used in a sequence as follows:

$$\text{Clarifying filter} \Rightarrow \text{Prefilter} \Rightarrow \text{Membrane filter}$$

Each filter removes the suspended or colloidal material for the process stream to improve the performance of the filter next in line (1).

Equipment for Small-Scale Ultrafiltration, Including Tangential Flow Filtration For concentrating small volumes of 1- to 20-mL protein solutions, centrifugal filtration units (Fig. 9) fitted with ultrafiltration membranes are available. Other devices such as the stirred cell (Fig. 10), which can be used with a wide range of ultrafiltration membranes, are also marketed by filter suppliers. These devices are ideal for processing volumes of a few milliliters to about half a liter. A stirring bar helps to stir the fluid above the membrane and keeps the foulants moving in the body of the solution, reducing filter clogging and increasing filtration rates. Although this device will help to achieve small laboratory-scale separations or protein concentrations, it will not give indications of the membrane area or process conditions to use when the process is eventually scaled up to large volume operations.

Because large-scale ultrafiltration processes are usually conducted in tangential-flow filtration systems, it is better to start with a device that operates in this mode at the research and development laboratory. Small-scale devices for these applications are available from filter suppliers. Devices that hold single sheets of membrane or small filter modules with more membrane area are also available. It is important to proportionately increase the membrane

Figure 9. Small ultrafiltration devices used in centrifuges for small-scale ultrafiltration. *Source*: Photo courtesy Millipore S.A., Molsheim, France.

Figure 10. A bench-top device for small-volume ultrafiltration applications. *Source*: Photo courtesy Millipore S.A., Molsheim, France.

area of the filtration device as the volume of the process fluid increases. When the process is scaled up to the pilot plant level or to full-scale manufacture, large surface area modules and systems are available and offer the convenience of proper scale-up. Membrane filters in the following configurations are available from suppliers: *hollow fiber modules*, *spiral wound modules*, *plate and frame devices using flat sheet membranes*, and *tubular filters*. All these configurations have been developed by individual manufacturers based on sound hydrodynamic principles and rather complex mathematical models for module and system performance. The idea behind the development of each type of device is to improve its performance and ultimately reduce processing costs for the end user. The decision to use a particular device for a given application is usually made by taking into consideration the properties, and more specifically, the composition of the process stream, the goal of the application, and the results of the initial small-scale trials. The impact of the filtration step on the overall process economics is undoubtedly an important factor.

Small devices operated on bench tops are ideal for early studies in the research and development laboratories. Properly performed experiments help to determine which membranes and membrane devices are best suited for a particular application and also for correct optimization of the process. Most filter manufacturers help in proper planning of these studies, which are critical when one needs a quick answer to a bioseparation problem. (Some companies even offer hands-on training for the operators.) The result of the studies is the knowledge of which membrane pore size (in case of microfiltration membranes) or what nominal molecular weight cutoff (in case of ultrafiltration membranes) is best for the application.

Hollow fiber modules are sometimes used in ultrafiltration processes. The active membrane that participates in

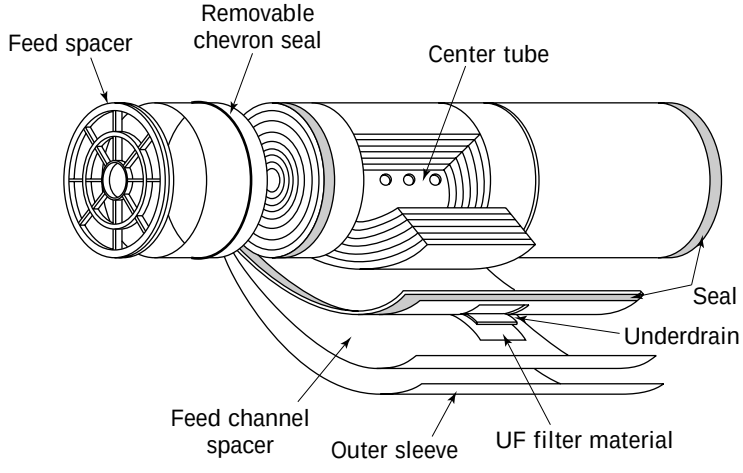


Figure 11. Components of a spiral-wound ultrafiltration membrane module. *Source*: Photo courtesy Millipore S.A., Molsheim, France.

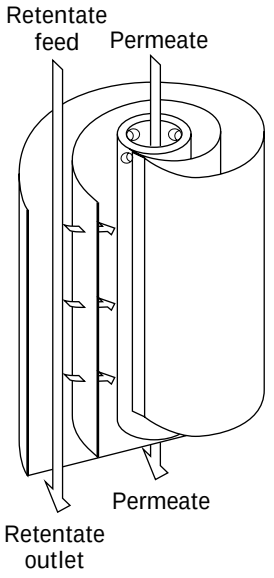


Figure 12. Flow schematic through a spiral wound ultrafiltration module. *Source*: Photo courtesy Millipore S.A., Molsheim, France.

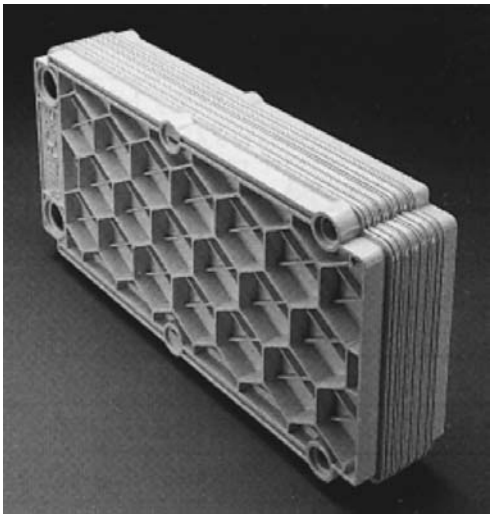


Figure 14. Photograph of a Prostak membrane module. *Source*: Photo courtesy Millipore S.A. Molsheim, France.

the separation process is on the inside of the hollow fiber of a uniform bore diameter. The process fluid is pumped through the hollow fiber, and the solvent and molecules that cross the membrane appear on the outside of the fiber. Bundles of many fibers, from 50 to 3,000 based on the internal diameter, are sealed inside a transparent cylindrical cartridge with the two ends sealed with epoxy sheets. Each cartridge has permeate ports at either end as well as feed and retentate ports. The ports can be connected to the piping of the system in which the module is installed by standard triclover clamps. Hollow fiber modules carrying ultrafiltration membranes of different cutoffs and surface area are available from several suppliers.

Spiral wound modules offer the biotechnology industry the most compact and least expensive of the ultrafiltration devices. The device is constructed of the same flat sheet membranes used in plate and frame devices. Each cartridge is made by wrapping alternate layers of membrane and separator screens concentrically around a hollow core, as shown in Figure 11. The flow schematic in a spiral wound ultrafiltration membrane cartridge manufactured by Millipore Corporation is shown in Figure 12. The product to be processed enters one end of the cartridge under pressure and flows tangentially down the central axis. The

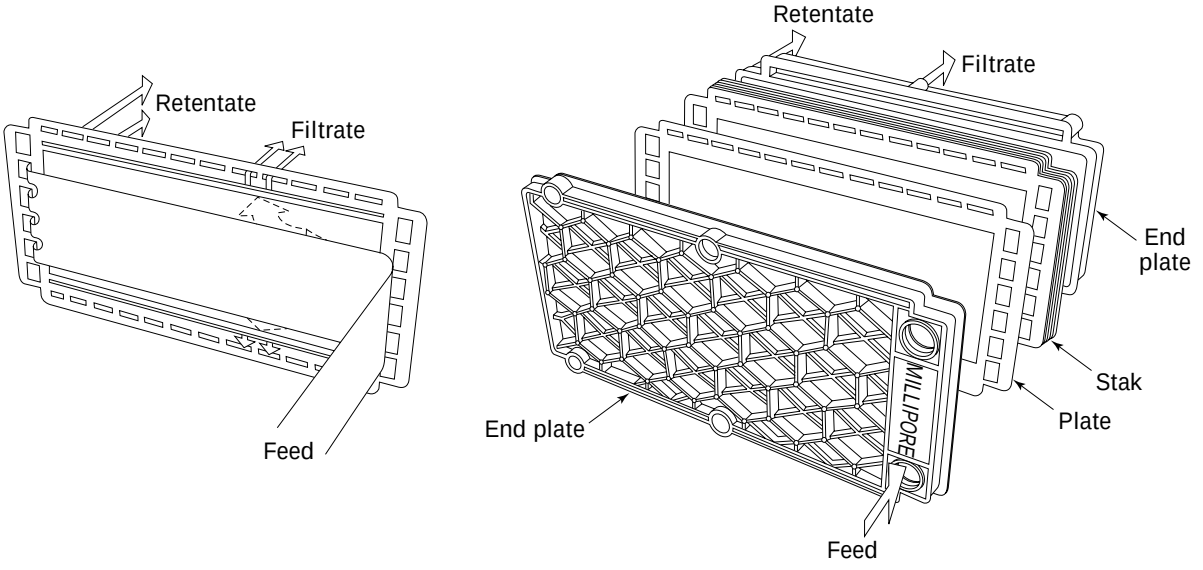


Figure 13. Flow schematic through an individual membrane plate (left) and through the entire Prostak plate and frame membrane module (right). *Source*: Photo courtesy Millipore S.A., Molsheim, France.

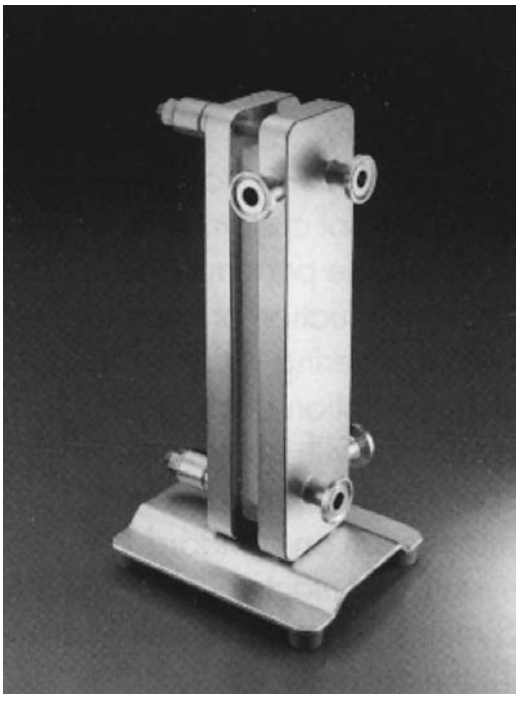


Figure 15. A simple bench-top device for evaluation of membranes in tangential-flow filtration in downstream processes. *Source*: Photo courtesy Millipore S.A., Molsheim, France.

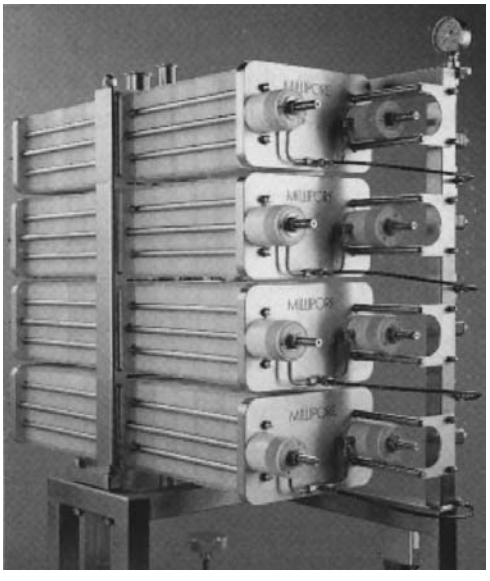


Figure 16. An assembly of Prostak modules giving several square feet of membrane area for use in a large system for bioseparations. *Source*: Photo courtesy Millipore S.A., Molsheim, France.

ultrafiltrate containing salts, water, and molecules not retained by the membrane flows through the membrane into the permeate channel and spirals all the way to the central core of the cartridge. The permeate is continuously removed from the central core while the retentate containing the rejected material exits through the retentate port.

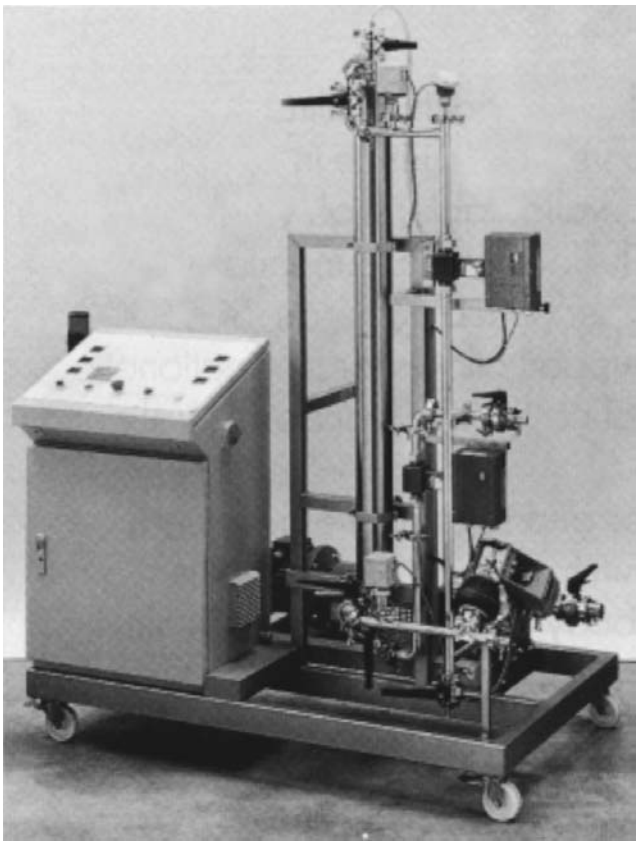


Figure 17. A small system with spiral-wound membrane modules. *Source*: Photo courtesy Millipore S.A., Molsheim, France.

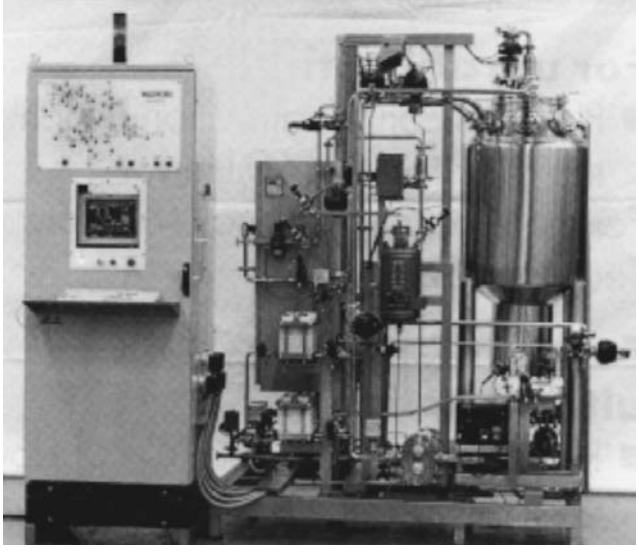


Figure 18. A fully automatic system with several membrane cassettes for protein concentration. *Source*: Photo courtesy Millipore S.A., Molsheim, France.

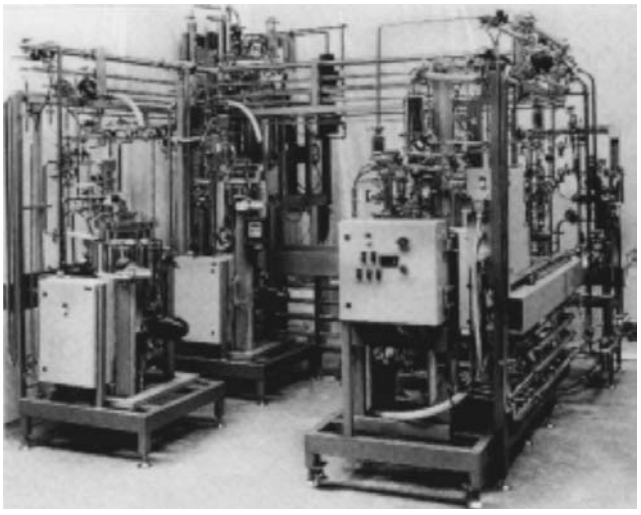


Figure 19. A fully automatic system used in downstream processing. *Source*: Photo courtesy Millipore S.A., Molsheim, France.

Spiral cartridges made with different ultrafiltration membranes with varying cutoffs are available from filter suppliers.

Flat sheet cassettes and *plate and frame modules* are a frequently employed configuration in tangential-flow filtration, and devices offered by suppliers contain microfiltration membranes of different pore sizes and ultrafiltration membranes with several cutoffs. The microfiltration membranes are the same ones offered in cartridges used in the dead-end mode. Small devices that can be operated in the tangential-flow filtration mode with single sheets of membranes or modules are ideal. The work at this scale will help the user to understand problems of filter clogging or membrane polarization and fouling phenomena and how to address the problems rapidly by either choosing another type of membrane or changing the process conditions.

The cassette filter devices are made by stacking and bonding together identical membrane sheets and separator screens together in parallel. Devices with identical membranes contain the same materials of construction. Therefore, development of a process can start with a small minicassette for the small-volume experiment and is eventually scaled up to a several hundred liter process by simply stacking together many larger surface area cassettes in stainless steel or acrylic holders.

As a process moves from the small bench-top experiment in the research laboratory to the floor of the manufacturing plant, the industry needs equipment to process large volumes of process streams. The filter suppliers have responded to this need of the biotechnology and other industries by developing modules that not only carry large membrane areas of several square meters but are extremely robust. These can be used repeatedly after proper cleaning and sanitization for several years. Typical examples of such devices are the Prostak membrane modules available from Millipore Corporation. These are tangential-flow stacked plate devices. The modules are solvent bonded and contain neither adhesives nor compression-dependent membrane seals. The modules containing Millipore's Durapore range of microporous membranes can be steam sterilized dozens of times. Figure 13 shows the flow path of the process fluid in the Prostak modules, which are shown in Figure 14. Because these are available in both the microfiltration and ultrafiltration versions, it is possible to use the same hardware for both steps in a downstream process, giving a significant reduction in capital investment. Most filter suppliers offer either manually operated or fully automatic systems that contain the membrane devices. A small bench-top filter evaluation module in a stainless steel holder is illustrated in Figure 15. Figure 16 shows how a number of plate and frame membrane modules can be put together in several stainless steel holders and plumbed together to form an assembly of membrane modules that can be part of a large process system. Figures 17, 18, and 19 illustrate a variety of typical small and large membrane systems used in the biotechnology industry. The systems carry the different membrane configurations described earlier.

CONCLUSION

The information provided above will give present and potential users of membranes for bioseparations a clear picture of how best to approach the task to achieve the desired end results. The manufacturers of membranes and membrane devices are always ready to give the users the necessary help in choosing the right product for each application. For detailed information on membrane separations, the readers can also refer to numerous review articles and textbooks on the subject, such as the one on ultrafiltration by M. Cheryan (3). Another useful book edited by McGregor (4) includes articles on a variety of specialized topics written by persons practicing the art of membrane separations in biotechnology. The use of microfiltration in the tangential mode in the recovery of intracellular products has been reviewed by Le and Atkinson (5), whereas Short and Webster (6) have authored an excellent article on the use of ultrafiltration in the pharmaceutical industry.

BIBLIOGRAPHY

1. E.S. Honig and P.D. Schwartz, *Filtr., Sep.* Jan/Feb 1997.
2. M.C. Porter, *Handbook of Separation Techniques for Chemical Engineers*, McGraw-Hill, New York, 1979.
3. M. Cheryan, *Ultrafiltration Handbook*, Technomic Publishing Company, Lancaster, Pa., 1986.
4. W.C. McGregor, *Membrane Separations in Biotechnology*, Dekker, New York, 1986.
5. M.S. Le and T. Atkinson, *Process Biotechnol.*, February 1995.
6. J.L. Short and D.W. Webster, *Process Biochem.*, March/April 1982.

See also CELL DISRUPTION AND LYSIS; CHROMATOGRAPHY, SIZE EXCLUSION; MEMBRANE CHROMATOGRAPHY; PROTEIN ULTRAFILTRATION.

MEMBRANE SURFACE LIQUID CULTURE, MICROORGANISMS, FUNGI

KAZUHIRO NAKANISHI
Okayama University
Okayama, Japan

KEY WORDS

Bioreactor
Continuous culture
Cultivation
Kojic acid
Membrane surface liquid culture
Molds
Neutral protease
Porous membrane
Secondary metabolite
Solid-state culture

OUTLINE

INTRODUCTION

Mold, a typical filamentous microorganism, is utilized widely in the field of biotechnology. Although molds are usually cultivated either by submerged liquid cultures and agitation or by solid-substrate cultures, these conventional methods may have various difficulties (1). Membrane surface liquid culture (MSLC) is one of the promising methods that has recently been developed by Nakanishi and his coworkers to solve difficulties associated with conventional cultivation methods (2–7). Although MSLC is very simple in principle, it has advantages of both the conventional submerged liquid culture and solid-substrate culture. Molds cultivated with MSLC sometimes produce metabolites in a larger quantities and are stable against autolysis.

PROBLEMS OF THE CONVENTIONAL CULTIVATION METHODS

Molds are difficult to cultivate because they are obligate aerobes and their growth is characterized by long, thin, branched threads of mycelium. Molds are usually cultivated either by submerged liquid cultures, such as shake-flask cultures and culture using an agitated vessel, or by solid-substrate cultures. However, these conventional methods include several severe problems for cultivation of molds. Because molds are obligate aerobes, it is of crucial importance to keep a sufficiently high dissolved oxygen level in the medium. In a shake-flask culture or a culture using an agitated vessel, molds take on either a pulpy state or a pelletlike morphology (aggregated forms) (8). In a pulpy state, molds may be damaged, suffering from shear stress caused by high agitation of the medium and spurging of air. The culture broth is highly viscous particularly at a high cell concentration, and accordingly, it becomes more difficult to supply oxygen to the cells. Also, the cells existing in the center of a pellet are apt to be autolyzed because of oxygen deficiency. On the other hand, in a solid-substrate culture, molds are grown statically on a solid substrate, such as unprocessed agricultural products or by-products, facing the air (9). Thus, in the solid-substrate culture, molds grow as they do in their natural habitats. Many researchers have reported that enzymes are sometimes produced at higher levels by solid-substrate cultures than by shake-flask or agitated vessel cultures (1–15). These findings might be partly ascribable to the fact that molds that are statically cultivated by the solid-substrate method do not suffer from fluid shear and low oxygen levels, although details of the mechanisms are not known. On the other hand, features of the solid-substrate culture is disadvantageous because the medium is restricted to natural agricultural products and thus control of the medium composition and pH is difficult. Furthermore, the production rate of metabolites is sometimes limited by mass transfer rates inside the solid substrate. Handling of the fine particles of the solid substrate during cultivation and a low yield of the product after cultivation are also serious problems.

PRINCIPLES AND CHARACTERISTICS OF MSLC

In MSLC, molds are grown on the surface of a microporous membrane, one side of which faces the air and the other side of which is in contact with liquid medium (2–7) (Fig. 1.) Although MSLC is very simple in principle, it has the advantages of both the solid-substrate culture and the conventional submerged liquid culture. In MSLC, molds are grown on the surface of a solid material (a porous membrane) facing the air, as in the case of solid-substrate culture. Moreover, MSLC is a static culture without vigorous agitation or shaking of the medium and spurging of air. Thus, the molds do not suffer from fluid shear. However, in contrast to the solid-substrate culture, liquid medium is used in MSLC, as in conventional submerged liquid cultures, which allows easy control of composition and pH during cultivation.

In MSLC, mycelia or hyphae do not pass through the membrane pores and thus do not invade the liquid medium side. If the hyphae should pass through the membrane pores, they would adhere onto the backside of the membrane, which in turn would block the pores and obstruct

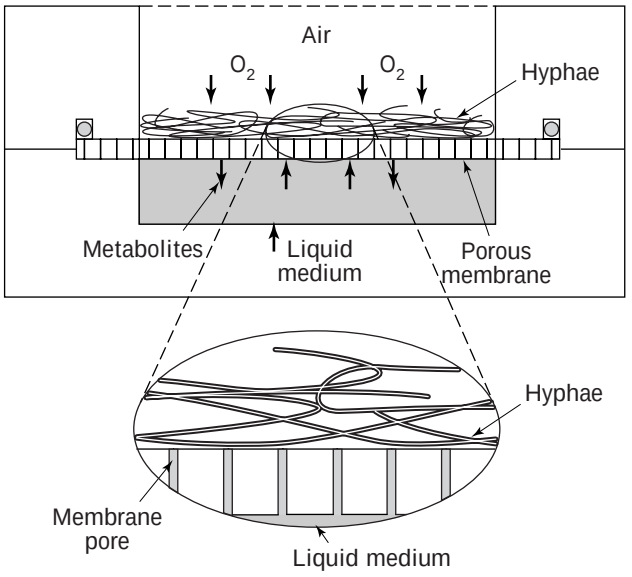


Figure 1. Conceptual scheme for MSLC. A porous membrane is set on the surface of liquid medium on one side and faces humidified air on the opposite side (upper surface). Molds are grown on the upper surface of the membrane, taking oxygen directly from the air and at the same time absorbing the liquid medium existing in the membrane pores. The enlarged figure below shows that the liquid medium rises up through the pores to the upper surface.

the transfer of medium components and metabolites. Thus, it is important to select a porous membrane with appropriate pore size, structure, and material. The membrane pore size should usually be smaller than 0.2–0.4 μm, with no appreciable mass transfer resistance (9). The membrane should be hydrophilic so that liquid medium can reach the top of the membrane pores at the outer surface where molds are grown, as shown in Figure 1. For example, an SE20 membrane made of polysulfone with hydrophilic treatment (Fuji Photo Film Co., Tokyo, Japan) cellulose membranes from Millipore (Bradford, Mass.) having a molecular weight cut-off of 5,000–300,000 could be used. Because the hyphae do not enter the membrane pores, the medium remains clear and uncontaminated with cells, which may contribute to simplification of the downstream processing.

Furthermore, MSLC is energy saving because it does not require shaking, agitation, or spurging.

Table 1 summarizes some major differences among MSLC, solid-substrate culture, and submerged liquid culture.

APPARATUS AND MODES OF CULTIVATION

Apparatus

Various types of apparatus are available for MSLC. For preliminary experiments such as screening of strains for production of a specific metabolite and optimization of medium composition, a plastic dish can be used. Liquid medium that has been autoclaved is poured into a plastic dish and then a porous membrane is set onto the surface of the liquid medium and held in place by appropriate supports, as shown in Figure 2 (4). Spores are inoculated uniformly onto the membrane surface, and cultivation is carried out statically in a thermostat-controlled and humidified aseptic box.

In experiments for the optimization of cultivation conditions, an apparatus made of heat-resistant plastic, such as polycarbonate, equipped with a peristaltic pump and a medium container is available, as shown in Figure 3 (2,4,6). The MSLC apparatus shown in Figure 3 is composed of an upper plate with a circular openings and a lower plate with a cylindrical groove. The membrane is set between the upper and lower plates with an O-ring. Spores are inoculated uniformly onto the membrane surface, and cultivation is allowed to start statically. The liquid medium can be circulated between the groove in the lower plate of the apparatus and an outside container. Medium components and pH can be controlled in the container. The medium in the groove may be gently stirred if necessary.

To conduct MSLC on a large scale, a cylindrical MSLC apparatus, shown in Figure 4 (7), or a plate-type apparatus in which two plates face each other at a short distance are available. In the cylindrical apparatus, the liquid medium is pumped up from a reservoir to the top of the cylindrical membrane and spouted toward the inner surface of the cylinder. The medium runs down along the inner surface of the cylindrical membrane module as in a wetted wall column. The medium is collected into the reservoir and circulated. Pumping should not be constantly, but instead intermittent. In continuous operation, fresh medium is supplied from a medium container and withdrawn using pumps with holding cells on the membrane surface. The entire cultivation system should be placed in a thermostat-controlled aseptic box or room equipped with an aseptic filter. Spores may be inoculated uniformly onto the membrane surface with a paintbrush or sprayed from a nozzle. Again, pH and medium components can be easily controlled in a container placed outside the membrane module.

Table 1. Some Major Differences in the MSLC, Solid-Substrate Culture, and Submerged Liquid Culture

	MSLC	Solid-substrate culture	Submerged liquid culture
Oxygen uptake	Mostly from gaseous phase	Mostly from gaseous phase	Soley from dissolved oxygen
Agitation	Not essential	Not essential	Strongly required
Medium	Liquid medium	Solid substrate	Liquid medium
Growth	Grown on surface of membrane	Penetrated deep into the solid substrate	Grown in a pulpy or pelletlike form
Recovery of extracellular product	Directly from the liquid medium	Extracted with liquid	Recovered after cell separation

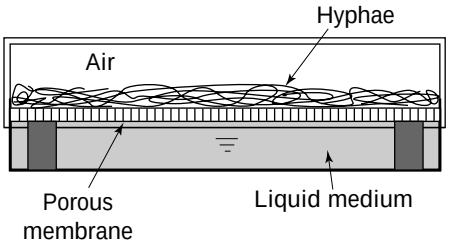


Figure 2. Schematic diagram for an MSLC apparatus using a plastic dish. A porous membrane is set on the surface of liquid medium and held in place by appropriate supports.

Cultivation Modes

Any cultivation mode can be applied to MSLC, including batch, fed-batch, repeated-batch, repeated fed-batch, and continuous. In every mode, molds are cultivated with holding cells on the membrane. In fed-batch MSLC, some components of the medium can be added at appropriate time intervals or added continuously into the medium container during the batch cultivation using the apparatus shown in Figures 3 and 4. In continuous MSLC, fresh medium is continuously fed into the apparatus and withdrawn at the same rate, with an appropriate retention time. The cells are held on the membrane surface during cultivation, in contrast to the ordinary continuous cultivation using an agitated vessel in which the steady state is attained in such a way that the amount of cells grown in per unit time in the fermentor is equivalent to that withdrawn from it.

EXPERIMENTAL RESULTS

Production of Neutral Protease Using Aspergillus oryzae IAM2704

A. oryzae usually produces neutral protease when grown in an appropriate liquid medium such one containing 4 g casein, 15 g glucose, 1 g K_2HPO_4, 0.5 g $MgSO_4 \cdot 7H_2O$, 0.5 g NaCl, and 0.01 g $FeSO_4 \cdot 7H_2O$ in 1 L of ion exchange water, pH 7.0 (3,4). Casein is an inducer for production of protease as well as a nitrogen source. Glucose is the carbon source. Figure 5 shows a comparison of courses of cultivation at 30 °C with respect to changes in glucose concentration, total protease activity in 20 mL of culture medium, and total dry cell weight in shake-flask culture, liquid-surface culture, solid-substrate culture, and MSLC. In all cases, cultivation was started by inoculating spores. In the shake-flask culture, 100-mL shake flasks were used. Liquid-surface culture was carried out statically on the surface of the medium. In the solid-substrate culture, cultivation was carried out on a 1.5% (w/v) agar containing every component of the medium. In MSLC, an SE20 membrane having a nominal pore size of 0.2 μm was used, and culture was carried out statically. A neutral protease activity was expressed as the amount of enzyme determined by the method using 0.6% casein as a substrate. One unit of enzyme activity was defined as the amount of enzyme that solubilizes 1 mg protein at pH 7.0 and 30 °C in 1 min, measured by Lowry's method with bovine serum albumin as a standard (16).

In the shake-flask culture and liquid-surface culture, the amount of protease reached its maximum after several days of cultivation and then rapidly decreased to a low level. The maximum amount of protease produced in both cultures was around 5 U/mL. In the solid-substrate culture on an agar medium, the highest amount of protease produced was about 25 U/mL of agar medium, which was 5 times higher than obtained with the shake-flask culture and liquid-surface culture. Furthermore, the amount of enzyme decreased more gradually after it reached a maximum in the solid-substrate culture. With MSLC, the amount of enzyme produced was higher than 50 U/mL, twice as high as by solid-substrate culture on an agar plate. As shown in Figure 5, both the consumption rate of glucose and the mycelial growth rate were also higher at the initial stage of MSLC than in the shake-flask culture,

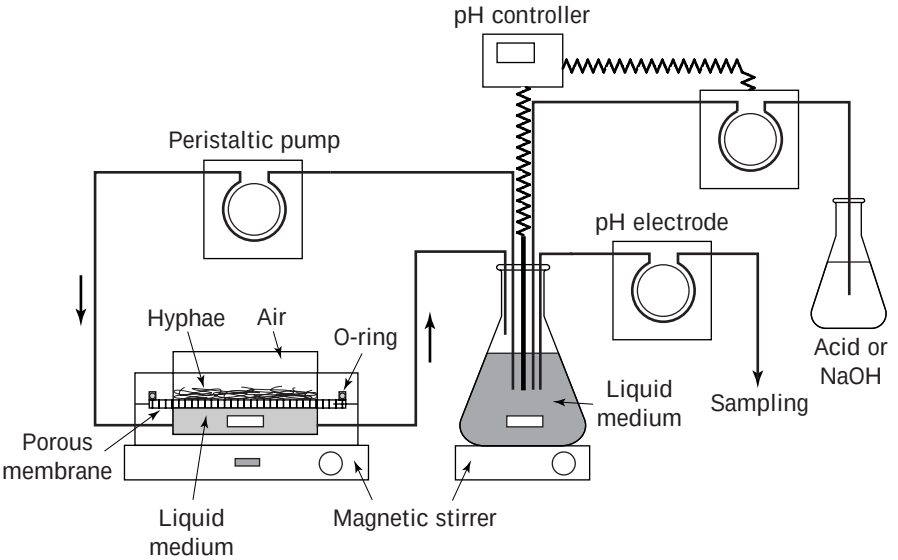


Figure 3. Schematic diagram for a disc-type MSLC system with a medium container, pH controller, and peristaltic pumps (7). The MSLC apparatus is composed of an upper plate with circular openings and a lower plate which is cylindrical, shallow, and contains liquid medium. A porous membrane is set between the upper and lower plate with an O-ring.

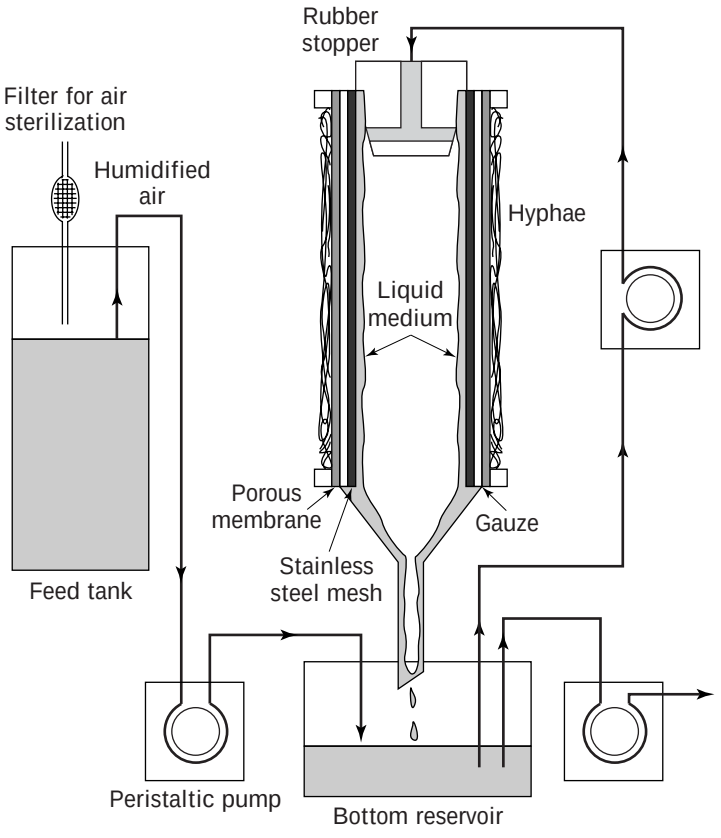


Figure 4. Schematic diagram of an MSLC system using a cylindrical membrane module (8). Liquid medium is intermittently circulated between a cylindrical membrane module and the bottom container. Liquid medium flows down along the inner surface of the module. In continuous MSLC, fresh medium is supplied from the feed tank to a bottom reservoir and withdrawn. The liquid medium is circulated, as in a batch MSLC.

probably because of higher efficiency of oxygen uptake. However, the reason why the enzyme was produced in a larger quantity by MSLC compared with the other cultures may not be due to the larger amount of the molds produced. The highest value for the specific amount of enzyme produced by the shake-flask culture, liquid surface-culture, solid-substrate culture, and MSLC in Figure 3 was around 1.1, 1.2, 4.0, and 6.3 U/mg, respectively.

In MSLC, the effects of the composition and concentration of medium on the production of metabolites were usually different from those in the shake-flask culture. Thus, it is important to optimize the composition and concentration of the medium by using a plastic dish, as shown in Figure 3. In Figure 6, a typical example is shown in which the glucose concentration in the medium for MSLC affects the amount of protease produced in a quite different fashion than seen in shake-flask culture. In MSLC, the maximum amount of protease and specific amount of protease per dry cell weight were increased by lowering the glucose concentration from 1.5% to 0.2%, and in shake-flask culture, the tendency was opposite. It seems that glucose was more effectively utilized to produce the enzyme in the MSLC than in the shake-flask culture.

The molds are stabilized to a greater extent in MSLC. Usually, it is very difficult to conduct a repeated-batch culture or a long-term continuous culture with cell holding in a shake-flask set-up. In MSLC, cultures with cell holding are easy to maintain because of the increased stability of the mold. Keeping the cells on the membrane surface, one can conduct cultures either in a repeated-batch or continuous mode. As shown in Figure 7, repeated batch culture of *A. oryzae* IAM2704 to produce neutral proteases could be effectively carried out by MSLC (5); the liquid medium was exchanged with fresh medium every day except in the first batch. In the first batch, it took a longer time for the molds to germinate and to grow over the whole membrane surface. Once the membrane surface was covered with hyphae, production of the proteases occured satisfactorily. As shown in Figure 7, the cumulative amount of protease increased almost linearly with cultivation time at the initial stage of cultivation. Although the amount of protease produced in each batch gradually decreased, even in the 27th batch it was about 7 times the amount produced in the shake-flask culture.

Production of Kojic Acid Using *A. oryzae* NRRL484

Most molds can produce secondary metabolites, particularly organic acids such as citric acid, erythorbic acid, and kojic acid. *A. oryzae* NRRL484 produced kojic acid in a quite large amount by MSLC (5–7). In the production of kojic acid, molds exhibited behaviors different from those for production of neutral protease. Molds were active in production of kojic acid even when the concentration of carbon and nitrogen sources was in the high range. Figure 8 shows the effect of glucose and yeast extract concentrations on the production of kojic acid using *A. oryzae* NRRL484 in MSLC and shake-flask culture at 30 °C, with culture medium containing 0.75 g $(NH_4)_2SO_4$, 0.25 g K_2HPO_4, 0.25 g $MgSO_4 \cdot 7H_2O$, and 0.5 g of 300 g/L HCl

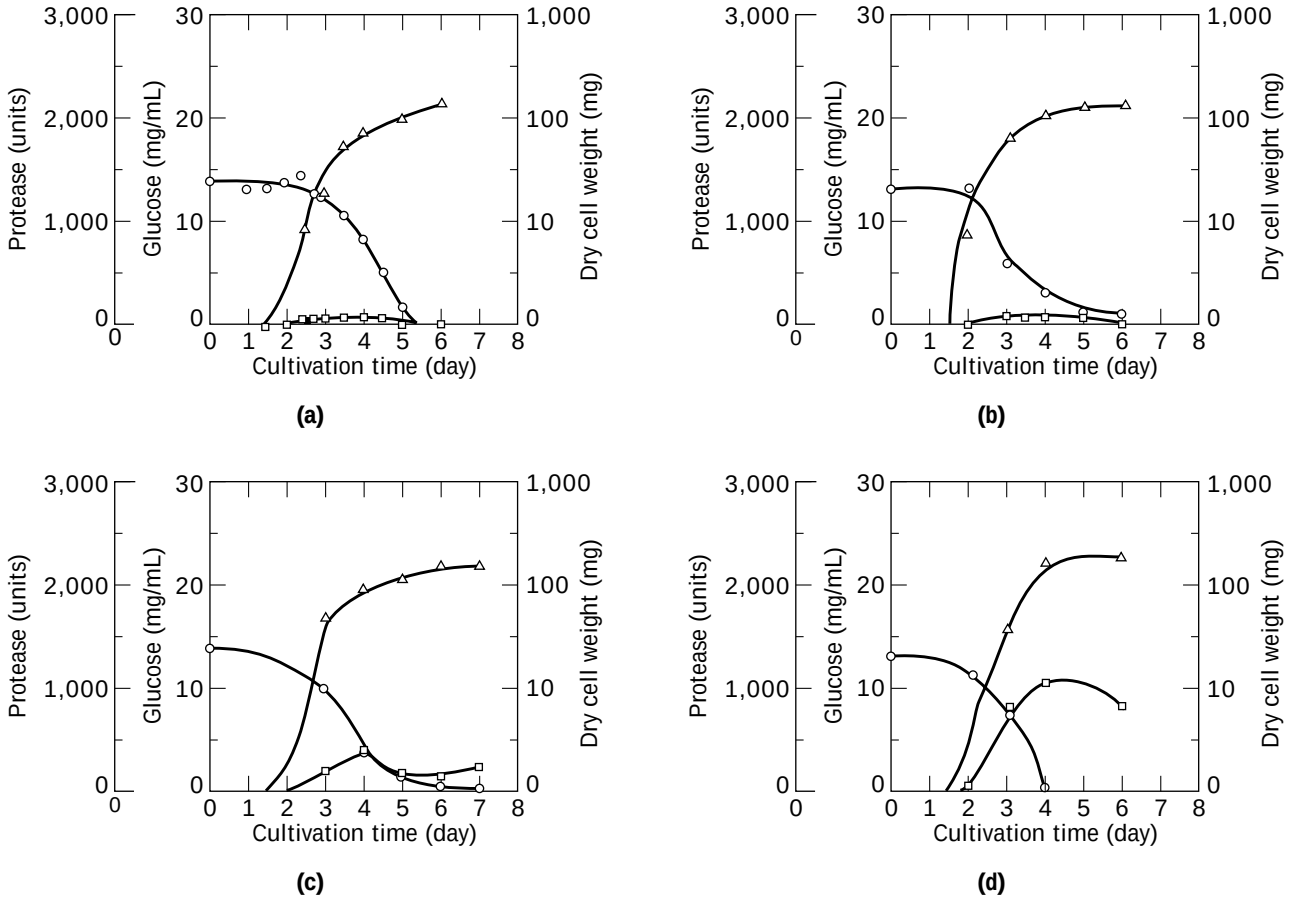


Figure 5. Change of concentrations of dried cell mass (△), glucose (○), and protease (□) during cultivation of *Aspergillus oryzae* IAM2704. (**a**) Shake-flask culture; (**b**) liquid-surface culture; (**c**) solid-substrate culture on agar plate; (**d**) MSLC (4).

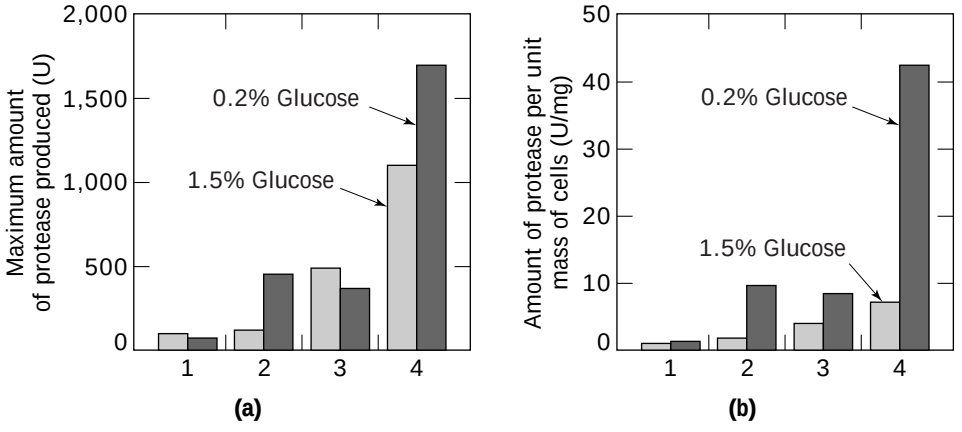


Figure 6. Maximum amount of protease produced (**a**) and specific amount of protease (**b**) during batch cultivation of *Aspergillus oryzae* IAM2704 using 0.2 and 1.5% glucose. 1, Shake-flask culture; 2, liquid-surface culture; 3, solid-substrate culture on agar plate; 4, MSLC (4).

solution containing 20 g/L $SO_4 \cdot 7H_2O$, 20 g/L $FeSO_4 \cdot 7H_2O$, and 30 g/L $ZnSO_4 \cdot 7H_2O$ in 1 L of ion exchanged water, pH 6.0, in addition to 50–400 g/L glucose and 0.2–10 g/L yeast extract. Glucose was a carbon source as well as the substrate for production of kojic acid. Yeast extract was the most suitable nitrogen source. The amount of kojic acid produced can be measured either by Bentley's method (17) or by high-performance liquid chromatography (HPLC). As shown in Figure 8 for a wide range of glucose and yeast extract concentrations, the maximum amount of kojic acid produced by MSLC was higher than by shake-flask culture (7).

Because molds were quite stable in MSLC, repeated fed-batch culture of *A. oryzae* NRRL484 was possible to

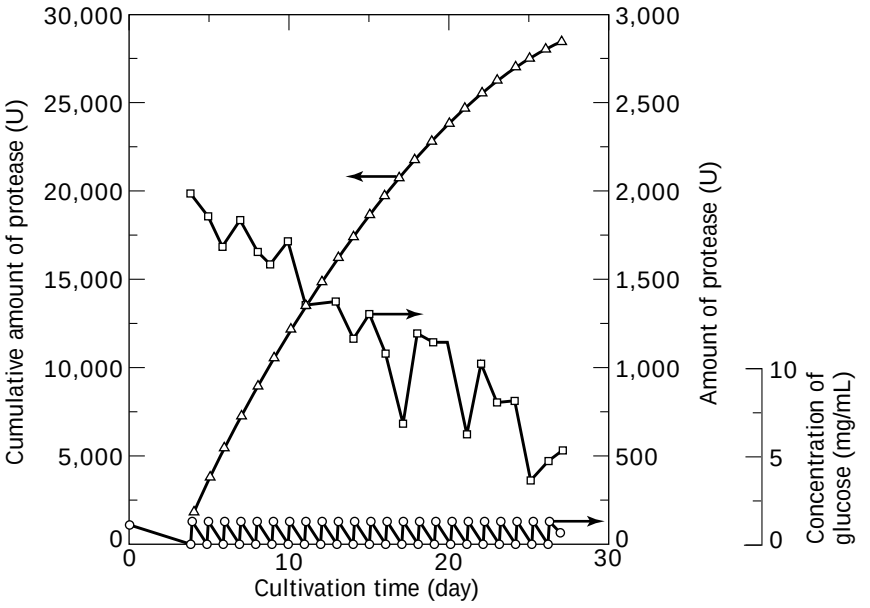


Figure 7. Changes of amount of protease produced in each batch (□), cumulative amount of protease (△), and concentration of glucose (○) during long-term repeated-batch culture by MSLC. The medium containing 0.2% glucose and 0.4% casein was exchanged every 24 h after 4 days of cultivation (first batch).

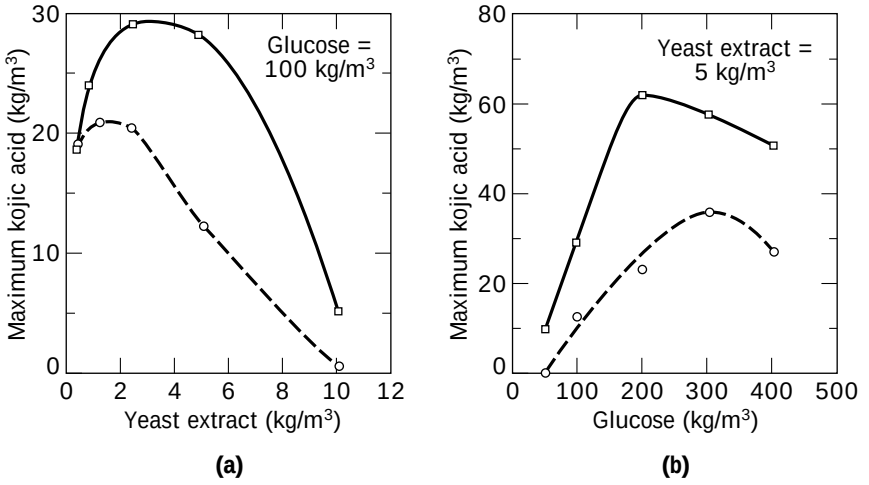


Figure 8. Effects of the concentrations of yeast extract (**a**) and glucose (**b**) on kojic acid production by MSLC (□) and shake-flask culture (○).

produce kojic acid as shown in Figure 9, in which powder glucose was added three times during one batch cultivation and the whole liquid medium was exchanged every 4 days, except the first batch (5). It took a longer time in the first batch until the molds had germinated and grown over the entire membrane surface, as was the case for production of neutral protease. Once the membrane surface was covered with hyphae, there was substantial production of kojic acid. In Figure 9, 12 fed-batch cultures are plotted. For batches 2 through 10, the kojic acid concentration in every batch reached 75 g/L or more, and it precipitated as fine crystals in the medium. In the batches 11 and 12, the production rate was greatly decreased. One possible reason for this decrease is that the amount of mycelia mat formed on the membrane surface increased, which might suppress oxygen transfer into the inside of the mat. Thus, some mycelia that were exposed adjacently on the membrane surface may have been autolyzed.

Continuous production of kojic acid was also possible using *A. oryzae* NRRL484 in a tubular-type MSLC apparatus with cell holding (7). In the cylindrical MSLC apparatus shown in Figure 4, a vertical cylinder 4 cm in diameter and 24 cm in length made of stainless steel mesh is surrounded by a sheet of gauze and then by an SE20 membrane sheet. The top end of the vertical cylindrical membrane module was sealed with a silicon rubber stopper. Typical results for continuous MSLC are shown in Figure 10. In this case, the effective membrane surface area was around 220 cm^2. The volume of medium in the entire system for recycling was 700 mL. At the start of cultivation, about 3 mL of a suspension of spores (5,000/mL) was inoculated uniformly onto the membrane surface using a paintbrush. During cultivation, medium was pumped up from a reservoir at a flow rate of about 6.6 mL/min using a peristaltic pump, and it flowed toward the inner surface of the cylinder through four small holes bored inside the

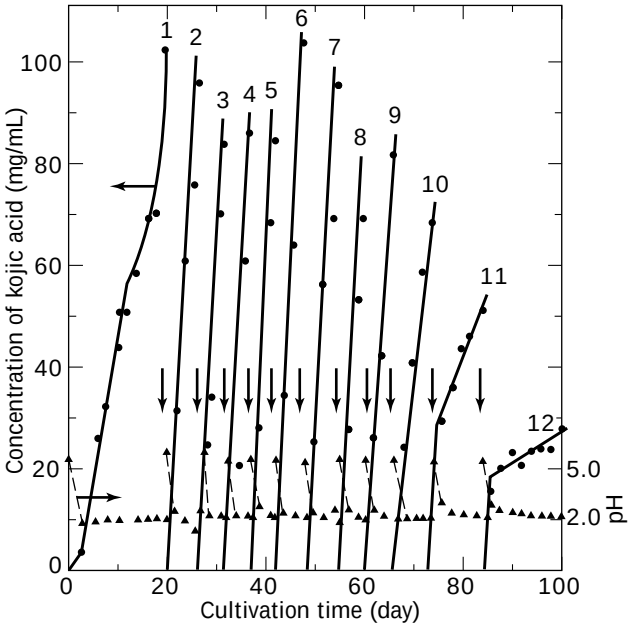


Figure 9. Changes of the concentration of kojic acid produced in each batch (●) and pH (▲) during repeated fed-batch MSLC using the exchange medium containing 0.25% yeast extract and 10% glucose. During each batch culture, powdered glucose was added to a final concentration of about 10% usually twice (batches 1 through 9) and three times (batches 10 and 12). Arrows show the times at which the entire medium was exchanged with fresh medium.

stopper. The pump was driven for 5 min every 15 min using a timer, and thus medium circulation was repeated. Fresh medium was supplied from a medium container and withdrawn using peristaltic pumps. The pumps for medium supply and withdrawal were driven for 1 min every 1 h at a flow rate of 1.25 mL/min. Thus, an average flow rate was 30 mL/day. As shown in Figure 10, after 12 days of cultivation, an almost steady state was attained, although glucose still remained. The steady-state concentration level of kojic acid was in the range of about 40 to 50 g/L for about 65 days. The yield of kojic acid on the basis of glucose consumed was evaluated to be around 70–80%, which is higher than that produced by shake-flask culture (30–40%). The glucose concentration gradually decreased even after 15 days of cultivation, probably because of maintenance metabolism. The fact that long-term continuous MSLC is possible, as shown in Figure 10, may indicate the increased stability of the mold.

PROBLEMS AND FUTURE PERSPECTIVES

In MSLC, molds tend to efficiently produce enzymes and secondary metabolites, and their stability against autolysis seems to be increased. Molds are typical aerobes with a strong oxygen requirement. Molds grow on a solid surface in their natural habitats. Such properties of molds probably favor MSLC, in which molds are grown on the membrane surface, obtaining oxygen directly from the atmosphere while taking liquid medium through the membrane pores with negligibly small mass transfer resistance. In conventional submerged liquid cultures using a shake-flask or an agitated vessel, mold cells are surrounded by the liquid medium, which contains a low amount of dis-

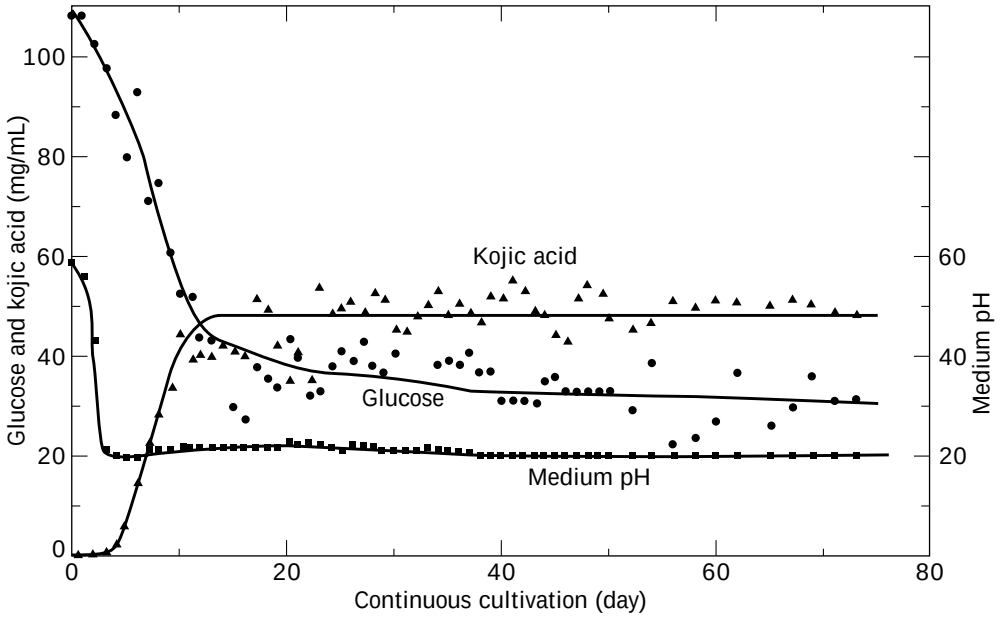


Figure 10. Continuous MSLC using a cylindrical membrane module at 30°C, holding molds on the membrane. About 30 mL of fresh medium was supplied semicontinuously per day, and the same amount was withdrawn while the medium was circulated between the membrane module and bottom container.

solved oxygen; in addition the cells suffer from shear stress, which may promote autolysis. In fact, a long-term operation with cell holding is quite difficult in conventional submerged liquid cultures. On the other hand, MSLC resembles liquid-surface culture (18), in which molds are suspended on the surface of liquid medium without a porous membrane at the interface. Liquid-surface culture is in principle very simple and an energy saving method. Liquid-surface culture was successfully applied to the production of citric acid using molds (19). However, there are some difficulties in conducting liquid-surface culture. In the production of neutral protease by liquid-surface culture using *A. oryzae* IAM2704, the amount of enzyme produced was slightly more than that by shake-flask culture but much less in comparison with that by MSLC. Furthermore, the cell mat is apt to sediment with increasing size. On the other hand, in kojic acid production using *A. oryzae* NRRL484, the cells formed a dense mat on the surface of the liquid medium and did not sediment; however, kojic acid production was less than that obtained by the MSLC. Thus, behaviors of the molds in liquid-surface culture seem to depend on the kinds of mold and cultivation conditions.

Sudo et al. have recently reported (20) that α-amylase is produced at a higher level by solid-substrate culture using rice bran than by agar-plate culture due to better growth of submerged mycelia, which mainly produce the enzyme. Thus, it may be interesting to compare the cells cultivated by MSLC with those obtained from other cultures (such as solid-substrate culture, agar-plate culture, liquid-surface culture, and shake-flask culture) from a stand point of the differences in respiratory activity, the amounts of related enzymes and, also gene expression. It may also be interesting to examine the effect of the gas composition on cultivation of molds by MSLC.

MSLC is an energy-saving process because the amounts of energy required for agitation and aeration are negligible, and downstream processing is simplified because the medium is not contaminated with cells. However, to utilize MSLC industrially some technical problems should be solved, such as those in spore inoculation, removal of the cells, and sterilization of the apparatus. The properties of the membrane are an important factor. The membrane should be reusable, resistant to autoclaving, and show appropriate hydrophilicity.

MSLC is an interesting method for investigating the physiological properties of molds, as well as a promising method for effective production of metabolites using molds.

BIBLIOGRAPHY

1. S. Braun and E.V. Lifshitz, *Trends Biotechnol.* **9**, 153–156 (1995).
2. A. Yasuhara, A. Ogawa, T. Tanaka, T. Sakiyama, and K. Nakanishi, *Biotechnol. Tech.* **8**, 249–254 (1994).
3. A. Yasuhara, A. Ogawa, T. Tanaka, T. Sakiyama, and K. Nakanishi, *Biotechnol. Tech.* **9**, 153–156 (1995).
4. A. Ogawa, A. Yasuhara, T. Tanaka, T. Sakiyama, and K. Nakanishi, *J. Ferment. Bioeng.* **80**, 35–40 (1995).
5. A. Ogawa, A. Yasuhara, T. Tanaka, T. Sakiyama, and K. Nakanishi, *J. Ferment. Bioeng.* **80**, 41–45 (1995).
6. K. Nakanishi, Y. Wakisaka, T. Tanaka, and T. Sakiyama, *Proc. 5th World Congress of Chemical Engineering*, American Institute of Chemical Engineers, New York, 1996, pp. 271–276.
7. Y. Wakisaka, T. Segawa, K. Imamura, T. Sakiyama, and K. Nakanishi, *J. Ferment. Bioeng.* **85**, 488–494 (1998).
8. S. Braun and S.E.V. Lifshitz, *Trends Biotechnol.* **9**, 63–68 (1991).
9. D.A. Mitchell and B.K. Lonsane, in H.W. Doelle, D.A. Mitchell, and E.R. Carlos eds., *Solid Substrate Cultivation*, Elsevier, London, 1992, pp. 1–16.
10. K. Sakaguchi, H. Okazaki, and M. Takeuchi, *Nippon Nougeikagaku Kaishi*, **29**, 349–353 (1955) (in Japanese).
11. T. Fujishima, K. Uchida, and H. Yoshino, *J. Ferment. Technol.* **50**, 724–730 (1972).
12. C.W. Hesseltine, *Process Biochem.* **12**, 29–32 (1977).
13. R. Bajracharya and R.E. Mudgett, *Biotechnol. Bioeng.* **22**, 2219–2235 (1980).
14. C.-F. Lin and H. Iizuka, *Appl. Environ. Microbiol.* **43**, 671–676 (1982).
15. W. Grajek, *Appl. Microbiol. Biotechnol.* **26**, 126–129 (1987).
16. T. Yano, S. Ashida, T. Tachiki, H. Kumagai, and T. Tochikura, *Agric. Biol. Chem.* **55**, 379–385 (1991).
17. M.R. Johns, in H.W. Doelle, D.A. Mitchell, and E.R. Carlos eds., *Solid Substrate Cultivation*, Elsevier, London, 1992, pp. 341–352.
18. R. Bentley, in S.P. Colowick and N.O. Kaplan eds., *Methods in Enzymology*, vol. 3, Academic, New York, 1957, pp. 238–245.
19. S. Ozaki, *Hakko Kyokaishi*, **16**, 7–17 (1958) (in Japanese).
20. S. Sudo, T. Ishikawa, K. Sato, and T. Oba, *J. Ferment. Bioeng.* **77**, 483–489 (1994).

See also BIOREACTORS, AIR-LIFT REACTORS; BIOREACTORS, CONTINUOUS STIRRED-TANK REACTORS; BIOREACTORS, FLUIDIZED-BED; MASS TRANSFER; SOLID STATE FERMENTATION, MICROBIAL GROWTH KINETICS; SOLID SUBSTRATE FERMENTATIONS, ENZYME PRODUCTION, FOOD.

METABOLITES, PRIMARY AND SECONDARY

ARNOLD L. DEMAIN
Massachusetts Institute of Technology
Cambridge, Massachusetts

KEY WORDS

Amino acids
Antibiotics
Biosynthesis
Hybrid antibiotics
Nucleotides
Recombinant DNA
Vitamins

OUTLINE

The versatility of microbial biosynthesis is enormous. Today's products of the microbe vary from the simplicity of

methane and acetic acid to the complexity of vitamin B_{12} and proteins. The importance of the fermentation industry resides in five characteristics of microorganisms: *(1)* a high ratio of surface area to volume, which facilitates the rapid uptake of nutrients required to support high rates of metabolism and biosynthesis; *(2)* the tremendous variety of reactions microorganisms are capable of carrying out; *(3)* a facility to adapt to a large array of different environments, allowing a culture to be transplanted from nature to the laboratory flask or the factory fermentor, where it is capable of growing on inexpensive carbon and nitrogen sources and producing valuable compounds; *(4)* the ease of genetic manipulation, both in vivo and in vitro, to increase production of the products, to modify structures and activities, and to make entirely new products; and *(5)* an ability to make specific enantiomers, usually the active ones, in cases where normal chemical synthesis yields a mixture of active and inactive enantiomers.

Microorganisms are important to us for many reasons, but one of the principal ones is that they produce things of value to us. These may be very large materials such as proteins, nucleic acids, carbohydrate polymers, or even cells, or they can be smaller molecules, which we usually separate into metabolites essential for vegetative growth and those that are inessential (i.e., primary and secondary metabolites, respectively).

The power of the microbial culture in the competitive world of commercial synthesis can be appreciated by the fact that even simple molecules (e.g., L-glutamic acid and L-lysine) are made by fermentation rather than by chemical synthesis. Although a few products have been temporarily lost to chemical synthesis (e.g., solvents such as acetone and butanol), it is obvious that most natural products are made by fermentation technology. Despite the efficiency of the chemical route to riboflavin, much of the production of this compound is carried out by fermentation; chemical processes to vitamin C and steroids still employ microbial bioconversion steps. Most natural products are so complex and contain so many centers of asymmetry that they probably will never be made commercially by chemical synthesis.

Although microbes are extremely good in presenting us with an amazing array of valuable products, they usually produce them only in amounts that they need for their own benefit; thus, they tend not to overproduce their metabolites. Regulatory mechanisms have evolved in microorganisms that enable a strain to avoid excessive production of its metabolites so that it can compete efficiently with other forms of life and survive in nature. Some of the important control mechanisms are induction, feedback regulation, and nutrient (catabolite) regulation. The fermentation microbiologist, however, desires a "wasteful" strain that will overproduce and excrete a particular compound that can be isolated and marketed. During the screening stage, the microbiologist searches for organisms with weak regulatory mechanisms. Once a desired strain is found, a development program is begun to improve titers by modification of culture conditions, mutation, and recombinant DNA technology. The microbiologist is actually modifying the regulatory controls remaining in the original culture so that its inefficiency can be further increased and the microorganism will excrete tremendous amounts of these valuable products into the medium.

The main reason for the use of microorganisms to produce compounds that can otherwise be isolated from plants and animals or synthesized by chemists is the ease of increasing production by environmental and genetic manipulation. Thousandfold increases have been recorded for small metabolites. Of course, the higher the specific level of production, the simpler the job of isolation. Consider the case of *Ashbya gossypii*, which has been forced into making over 40,000 times more riboflavin than it needs, or *Pseudomonas denitrificans*, which produces a 100,000-fold excess of vitamin B_{12}. The original Oxford strain of *Penicillium notatum* produced 5 mg of penicillin per liter; today's strains make over 60,000 mg/L, a figure higher than the dry weight of the cells in the fermentor.

Primary metabolites are the small molecules of all living cells that are intermediates or end products of the pathways of intermediary metabolism, are building blocks for essential macromolecules, or are converted into coenzymes. The most industrially important are the amino acids, nucleotides, vitamins, solvents, and organic acids. Primary metabolites vary in size from hydrogen gas (2 Da) and methane (16 Da) to vitamin B_{12} (1,355 Da). It is not unexpected that amino acids and vitamins are used in human and animal nutrition, that ethanol, acetone, and butanol are used as fuels and solvents, and that citric and acetic acids are used as acidulants. However, many of these general metabolites are used in novel ways: the sodium salts of glutamic, 5′-inosinic, and 5′-guanylic acids as flavor enhancers, sodium gluconate as a sequestering agent to prevent the deposition of soap scums on cleaned surfaces, and fumarate in the manufacture of polyester resins. Organisms that produce such products are fantastic in their degree of overproduction.

For many years, microbial products have been used to enhance the quality, appeal, and availability of food and drink. A list of primary metabolites that have been or are being used in the food and feed industries is shown in Table 1.

About 1 million tons of amino acids are produced annually. In amino acid production, feedback regulation is

Table 1. Some Primary Metabolites Used in the Food and Feed Industries

Class	Examples
Alcohols	Ethanol
Amino acids	Glutamic acid (monosodium glutamate), lysine, threonine, tryptophan, phenylalanine
Nucleotides	5′-Guanylic acid, 5′-inosinic acid
Organic acids	Acetic acid, propionic acid, succinic acid, fumaric acid, lactic acid, malic acid, tartaric acid, citric acid, gluconic acid
Polyols	Glycerol, mannitol, erythritol, xylitol
Polysaccharides	Xanthan, gellan
Sugars	Fructose, sorbose
Vitamins	Riboflavin (B_2), cyanocobalamin (B_{12}), biotin

bypassed by isolating an auxotrophic mutant and partially starving it of its requirement. A second means to bypass feedback regulation is to produce mutants resistant to a toxic analog of the desired metabolite (i.e., an antimetabolite). Combinations of auxotrophic and antimetabolite resistance mutations are common in primary metabolite producing microorganisms. Permeability is very important in the production of L-glutamic acid, the major amino acid of commerce. About 1.2 billion pounds of monosodium glutamate, a potent flavor enhancer, are made annually by fermentation using various species of the genera *Corynebacterium* and *Brevibacterium* (e.g., *Corynebacterium glutamicum*, *Brevibacterium flavum*, and *B. lactofermentum*). Today, these glutamate producers are classified as subspecies of *Corynebacterium glutamicum* (e.g., *C. glutamicum* ssp. *flavum* and *C. glutamicum* ssp. *lactofermentum*). The major route of glutamate production from glucose is via the Embden–Meyerhof pathway and the early steps of the tricarboxylic acid cycle (Fig. 1). α-Ketoglutarate, which would normally be converted to succinyl coenzyme A in the cycle, is instead reductively aminated to glutamate by glutamate dehydrogenase. Molar yields of glutamate from sugar are 50–60%, and broth concentrations reach more than 100 g/L. Normally, glutamic acid overproduction would not occur because of feedback regulation. However, due to a modification of the cell membrane, glutamate is pumped out of the cell, thus allowing its biosynthesis to proceed unabated. The membrane alteration is intentionally effected by biotin limitation (all glutamic acid bacteria are biotin auxotrophs), glycerol limitation of glycerol auxotrophs, oleate limitation of oleate auxotrophs, or addition of penicillin or fatty acid derivatives to exponentially growing cells. Apparently all of these manipulations result in a phospholipid-deficient cytoplasmic membrane. Also required for effective production are a high level of carbon and energy source as well as growth inhibition. The actual excretion is carried out by a specific efflux system involving a carrier that is dependent on membrane potential (1).

The bulk of the cereals consumed in the world are deficient in the essential amino acid L-lysine. Lysine supplementation converts such cereals into balanced food or feed. Lysine is a member of the aspartate family of amino acids. It is produced in bacteria by a branched pathway that also produces methionine, threonine, and isoleucine. This pathway is controlled very tightly in an organism such as *Escherichia coli*; this organism contains three aspartate kinases, each of which is regulated by a different end product. In addition, after each branch point, the initial

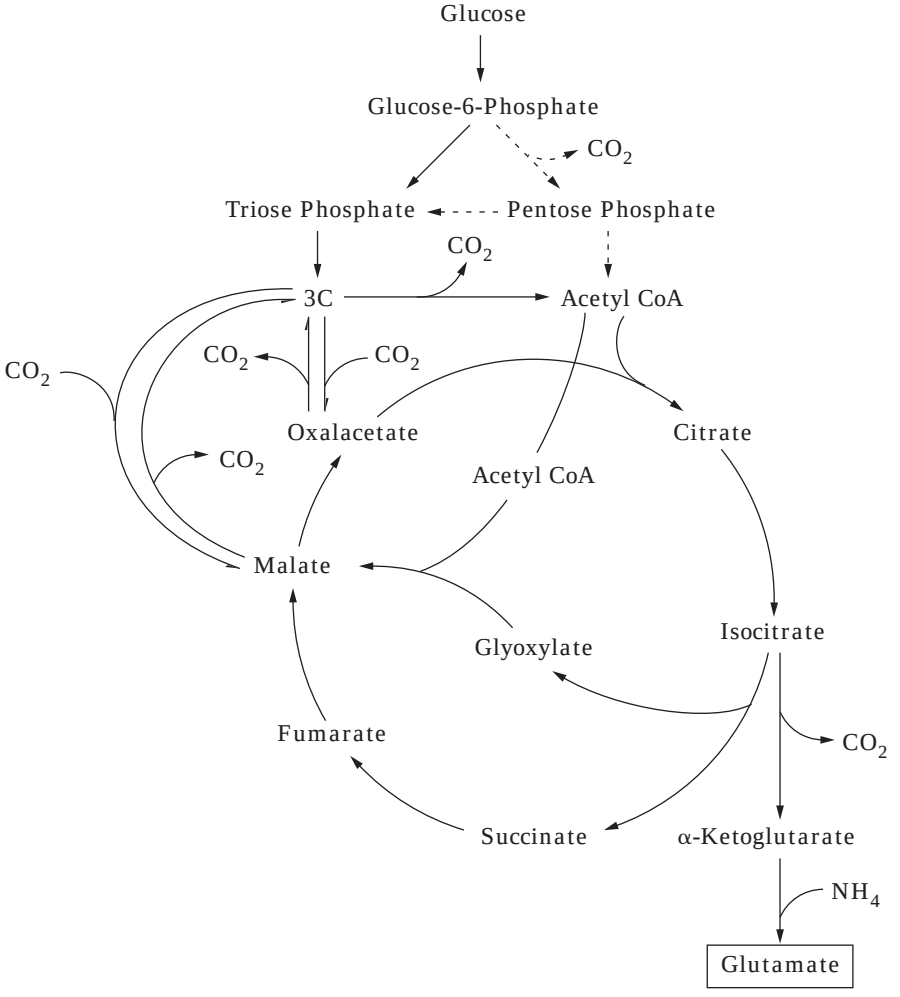


Figure 1. Pathway of glutamic acid production from glucose by *Corynebacterium glutamicum*.

enzymes are inhibited by their respective end products and no overproduction occurs. However, in lysine fermentation organisms (e.g., various mutants of *C. glutamicum* and its relatives), there is a single aspartate kinase, as shown in Figure 2. This aspartate kinase is regulated via concerted feedback inhibition by threonine plus lysine. By genetic removal of homoserine dehydrogenase, a glutamate-producing wild-type *Corynebacterium* is converted into a lysine-overproducing mutant that cannot grow unless methionine and threonine are added to the medium. As long as the threonine supplement is kept low, the intracellular concentration of threonine is limiting and feedback inhibition of aspartate kinase is bypassed. Other enzymes in the aspartate family biosynthetic pathway such as aspartate semialdehyde dehydrogenase and threonine synthetase are unregulated. In addition to the difference in the mode of aspartate kinase feedback inhibition, lysine overproducers differ from *E. coli* in the following ways: (*1*) no feedback repression of aspartate kinase or aspartate semialdehyde dehydrogenase occurs in lysine overproducers;

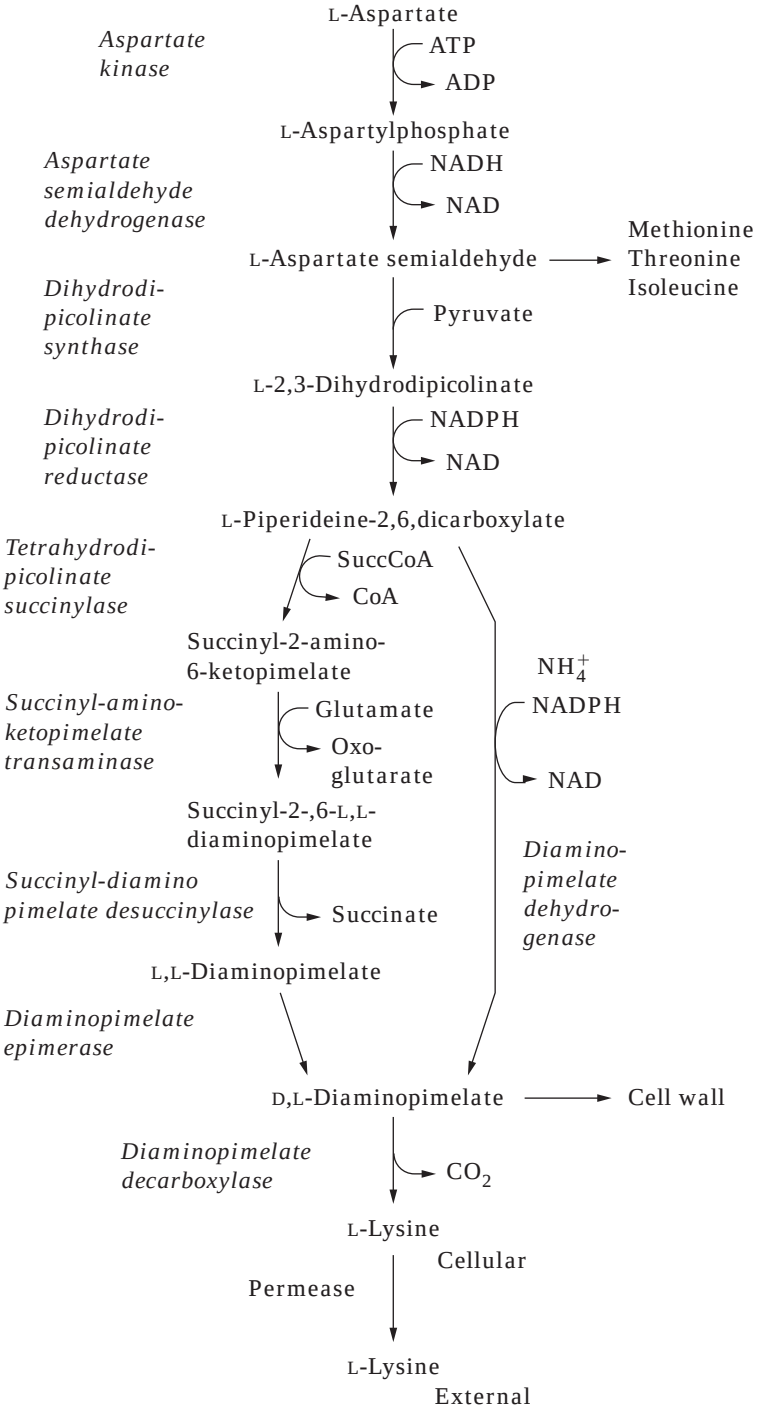


Figure 2. Biosynthesis of L-lysine by *Corynebacterium glutamicum Source:* From Ref. 2.

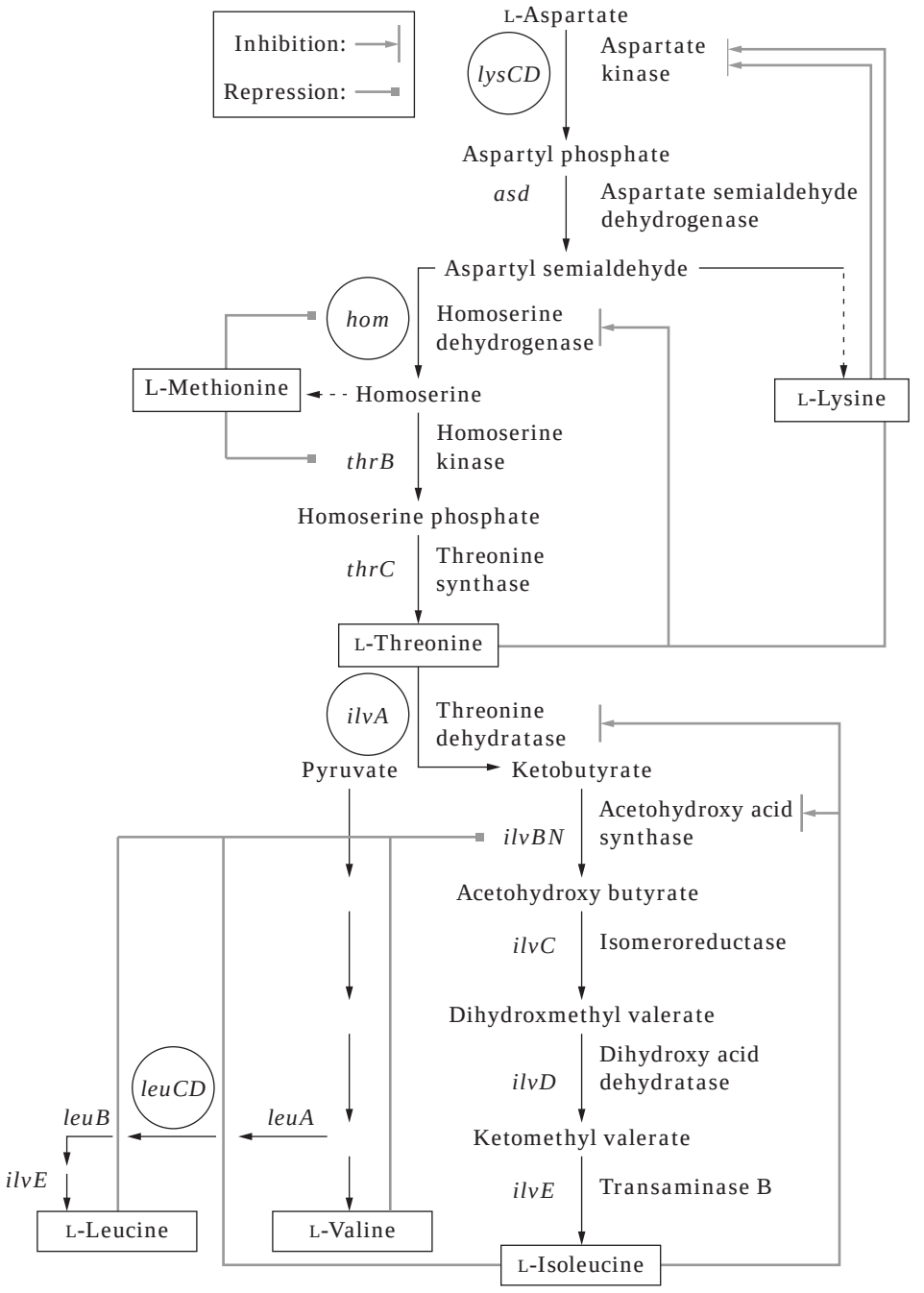


Figure 3. L-Isoleucine biosynthesis in recombinant *Corynebacterium glutamicum* strains. Repression on the DNA level (shaded lines with square ends) and inhibition of enzymes (shaded lines with arrowhead ends). The last four enzymatic steps are the same for L-isoleucine and L-valine biosynthesis. *Source*: Ref. 7.

(*2*) the first and second enzymes of the lysine branch (dihydrodipicolinate synthetase and dihydrodipicolinate reductase) are neither inhibited nor repressed by lysine in lysine overproducers; (*3*) L-lysine decarboxylase is absent in lysine overproducers. The pentose phosphate pathway is of much greater importance in lysine production than in glutamate formation due to the greater need of NADPH for the former path (3). Lysine excretion is by active transport involving a carrier and is driven by membrane potential, the lysine gradient, and the proton gradient. Excretion

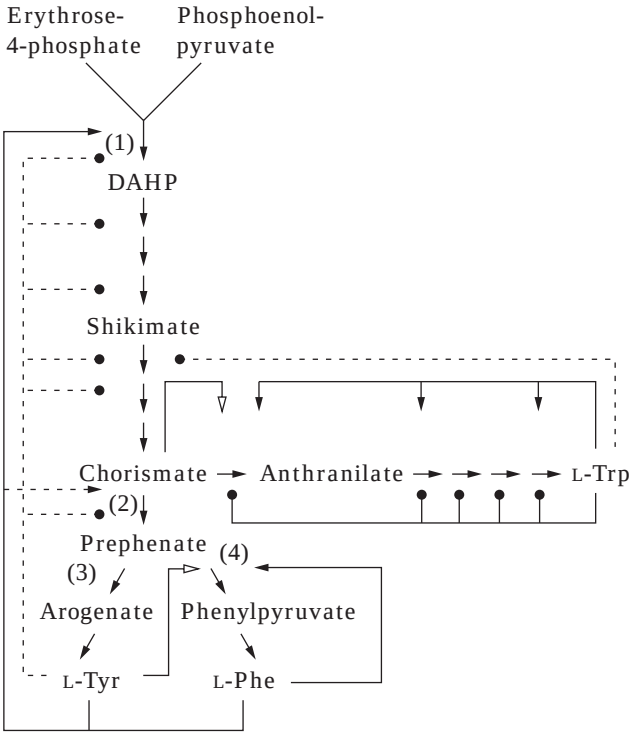


Figure 4. Biosynthetic pathway of aromatic amino acids and their regulatory mechanisms in *Corynebacterium glutamicum* sp. *lactofermentum*. 1, 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) synthase; 2, chorismate mutase; 3, arogenate aminotransferase; 4, prephenate dehydratase; ◀—, inhibition; ◀---, partial inhibition; ◁—, activation; ●—, repression; ●---, weak repression. *Source*: From Ref. 8.

uses a 2 OH/lysine symporter and is catalyzed by a dipetide uptake system dependent on electromotive force, not ATP (4).

Recombinant DNA techniques have made their way into the amino acid production area (5). *E. coli* and *Serratia marcescens* strains have been constructed with plasmids bearing amino acid biosynthetic operons. Plasmid transformation has been accomplished in *C. glutamicum*, so that recombinant DNA technology is now used to improve these commercial amino acid–producing strains. The major manipulations have involved gene cloning to increase the levels of feedback-resistant aspartate kinase and dihydrodipicolinate synthase (6). As a result, lysine industrial production yields 120 g/L and 0.25–0.35 g L-lysine·HCl per gram glucose used (molar yield of 0.25–0.35 mol L-lysine per mole of glucose used). The market for L-lysine·HCl is 360 million pounds per year (6).

Recombinant technology and traditional mutagenesis and selection have been major influences in constructing bacterial strains capable of producing 100 g/L L-threonine, 40 g/L L-isoleucine, 34 g/L L-leucine, 31 g/L L-valine, 28 g/L L-phenylalanine, 55 g/L L-tryptophan, 26 g/L L-tyrosine, 100 g/L L-proline, 65 g/L L-arginine, and 40 g/L L-histidine. The relationship between the lysine biosynthetic pathway and that for threonine, methionine, isoleucine, valine, and leucine in *C. glutamicum* is shown in Figure 3. The pathway to the aromatic amino acids (phenylalanine, tryptophan, and tyrosine) is given in Figure 4.

The commercial interest in nucleotide fermentations is due to the activity of two purine ribonucleoside 5′-monophosphates, namely guanylic acid (GMP) and inosinic acid (IMP), as enhancers of flavor (9,10). Some 2,500 tons of GMP and IMP are produced annually in Japan alone. Three main processes are used: (*1*) hydrolysis of yeast RNA by fungal nuclease to AMP and GMP followed by enzymatic deamination of AMP to IMP; (*2*) fermentative production of the nucleosides inosine and guanosine by *Bacillus subtilis* mutants followed by chemical phosphorylation, and (*3*) direct fermentation of sugar to IMP by *C. glutamicum* mutants plus conversion of guanine to GMP by salvage synthesis using intact cells of *Brevibacterium ammoniagenes*. Titers of IMP by direct fermentation have reached 27 g/L (10).

The pathway of purine ribonucleotide biosynthesis (Fig. 5) is virtually identical in all living systems. In the first specific step, 5′-phosphoribosyl pyrophosphate (PRPP), derived from ATP and ribose-5-phosphate, is aminated with glutamine to form 5′-phosphoribosylamine (PRA). By a series of nine known enzymic steps, PRA is converted to IMP, the first purine ribonucleotide formed. This pivotal intermediate is the precursor of both AMP and GMP. IMP is then converted to adenylosuccinate (S-AMP), aspartate acting as the succinyl donor and GTP as the energy source. Adenylosuccinate is then converted to AMP by adenylosuccinate lyase, a bifunctional enzyme that also catalyzes an earlier step. Inosinic acid is oxidized by IMP dehydrogenase using NAD^+ as a cofactor to form XMP. The latter is aminated by XMP aminase to GMP. In this reaction, ATP is the energy source, and, depending on the organism, the amino donor is either glutamine or ammonia. Both AMP and GMP can be used for synthesis of nucleic acids after being converted to their respective di- and triphosphates. Alternatively, GMP can be reduced to IMP in a reaction catalyzed by GMP reductase (which utilizes NADPH as an electron donor) and AMP can be converted to IMP by deamination. These reactions permit a cyclic regeneration of intermediates common to both purine nucleotides from an accumulated excess of either. All of the purine ribonucleotide interconversions occur via their 5′-isomers.

With such a complex series of biosynthetic reactions, it is essential to the economy of the cell that there be regulation of the rates of the individual reactions. This regulation is accomplished by negative feedback inhibition of enzyme action and repression of enzyme synthesis. Figure 6 shows an example of a typical feedback control system for nucleotide biosynthesis. Usually both AMP and GMP exert feedback repression and inhibition on the common pathway and on their respective branches. An important control site is step 1, catalyzed by PRPP amidotransferase. In some microorganisms, AMP and GMP each inhibit action of the enzyme completely at high concentrations. In other organisms, a cumulative type of regulation is observed. Synthesis of the amidotransferase, as well as many of the other enzymes of the common pathway, is also re-

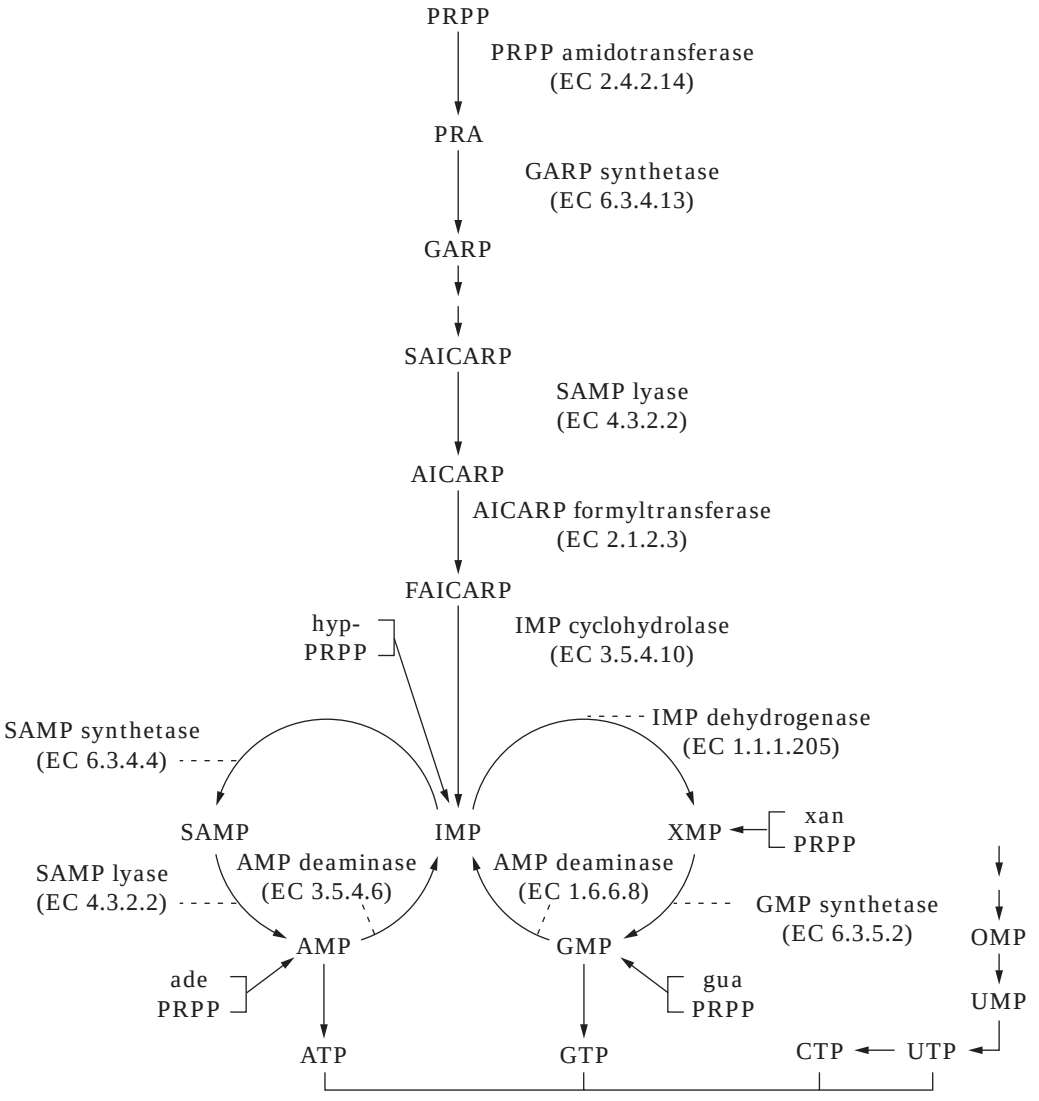


Figure 5. Biosynthetic pathway to purine nucleotides. PRPP, 5-phosphoribosyl-1-pyrophosphate (5-phospho-α-D-ribose-1-diphosphate); PRA, 5-phosphoribosyl-1-amine; GARP, glycineamide ribonucleotide; AICARP, 5-aminoimidazole-4-succinocarboxamide ribonucleotide; FAICARP, 5-formaminoimidazole-4-carboxamide ribonucleotide; SAMP, adenylosuccinate (succinoadenosine 5′-monophosphate). *Source*: From Ref. 10.

pressed by AMP and GMP. Another important control site is IMP dehydrogenase, which is usually inhibited and repressed by GMP. The conversion of IMP to S-AMP (catalyzed by adenylosuccinate synthase) is usually repressed and inhibited by AMP.

The key to effective purine accumulation is the limitation of intracellular AMP and GMP. This limitation is best effected by restricted feeding of purine auxotrophs. Thus, adenine-requiring mutants lacking adenylosuccinate synthetase (see Fig. 7) accumulate hypoxanthine or inosine that results from breakdown of intracellularly accumulated IMP. Because of the strong inhibition and repression of IMP dehydrogenase by GMP, the accumulation does not proceed past IMP into the GMP branch of the pathway. Certain adenine auxotrophs of *Bacillus subtilis* excrete more than 10 g/L inosine. These strains are still subject to GMP repression of enzymes of the common path. To minimize the severity of this regulation, the adenine auxotrophs are further mutated to eliminate IMP dehydrogenase. These adenine–xanthine double auxotrophs show a twofold increase in specific activity of some common-path enzymes and accumulate up to 15 g/L inosine under conditions of limiting adenine and xanthine (or guanosine). Further deregulation is achieved by selection of mutants resistant to purine analogs. Thus, mutants requiring adenine and xanthine and resistant to azaguanine produce more than 20 g/L inosine. Insertional inactivation of the IMP dehydrogenase gene in another *B. subtilis* strain yielded a culture producing 35 g/L inosine (11). Genetic engineering of the inosine monophosphate dehydrogenase

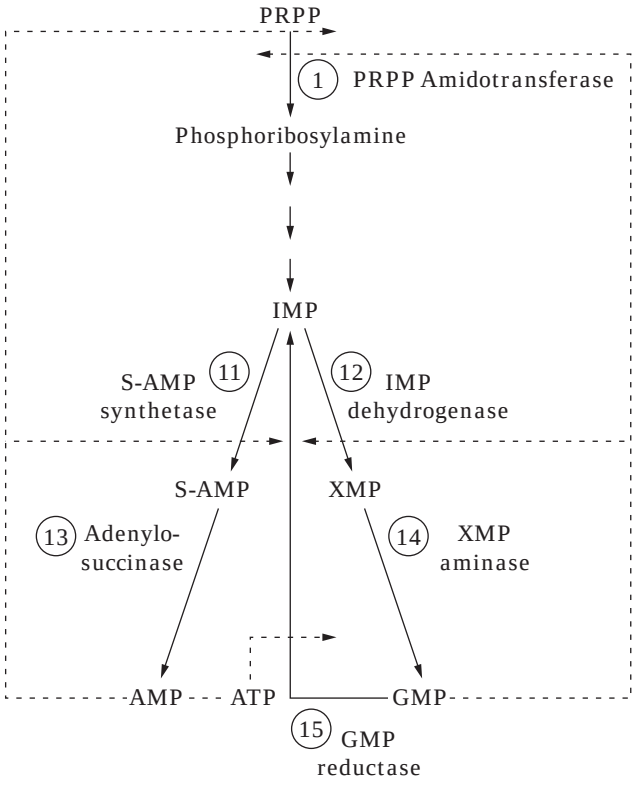


Figure 6. Biosynthesis of 5′-purine nucleotides. Solid lines, pathway steps; dashed lines, feedback inhibition or repression.

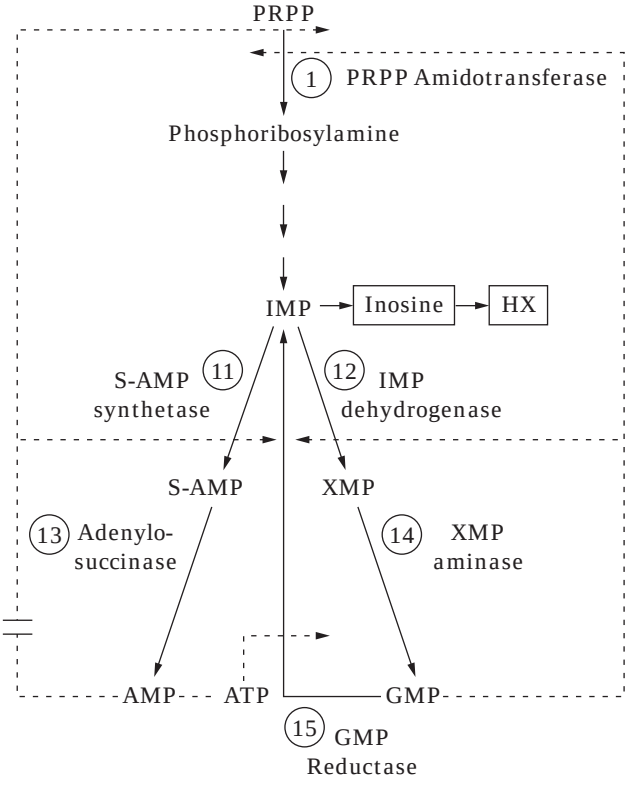


Figure 7. Overproduction of 5′-inosinic acid (5′-IMP) or its degradation products (inosine and hypoxanthine [hx]) by adenine auxotrophs. Solid lines, pathway steps; dashed lines, feedback inhibition or repression; dots, mutated (inactive) enzyme.

gene in a *B. subtilis* strain, which was producing 7 g/L guanosine and 19 g/L inosine, changed production to 20 g/L guanosine and 5 g/L inosine (12). Other *B. subtilis* mutants produce as much as 23 g/L guanosine (10). With regard to pyrimidine production, another recombinant strain of *B. subtilis* produces 18 g/L cytidine, and a mutant lacking homoserine dehydrogenase (which increased the concentration of the precursor aspartate in the cell) produces 30 g/L cytidine (13).

Riboflavin (vitamin B_2: Fig. 8) is produced commercially by both fermentation and chemical synthesis (15). Riboflavin overproducers include two yeastlike molds, *Eremothecium ashbyii* and *Ashbya gossypii*, which synthesize riboflavin in concentrations greater than 20 g/L. Production of riboflavin amounted to 4 million pounds in 1992 (16). A riboflavin overproducer such as *A. gossypii* makes 40,000 times more vitamin than it needs for its own growth. The biochemical key to riboflavin overproduction appears to involve insensitivity to the repressive effects of iron. Ferrous ion severely inhibits riboflavin production by low and moderate overproducers, such as clostridia and *Candida* but has no inhibitory action against *E. ashbyii* and *A. gossypii*. In normal microorganisms, it appears that iron represses almost all of the riboflavin biosynthetic enzymes, whereas riboflavin or a derivative inhibits the first enzyme of the pathway, GTP cyclohydrolase II. New processes using *Candida* sp. or recombinant *B. subtilis* strains that produce 20–30 g/L riboflavin have been developed in recent years.

Vitamin B_{12} (Fig. 9) is industrially produced by *Propionibacterium shermanii* or *Pseudomonas denitrificans* (16,18). Such strains used for industrial production make about 100,000 times more vitamin B_{12} than they need for their own growth. The key to the fermentation is avoidance of feedback repression by vitamin B_{12}. The early stage of the *Propionibacterium shermanii* fermentation is conducted under anaerobic conditions in the absence of the precursor 5,6-dimethylbenzimidazole. These conditions prevent vitamin B_{12} synthesis and allow for the accumulation of the intermediate, cobinamide. Then the culture is aerated and dimethylbenzimidazole is added, converting cobinamide to the vitamin. In the *Pseudomonas denitrificans* fermentation, the entire process is carried out under low oxygen conditions. A high level of oxygen results in an oxidizing intracellular environment that represses formation of the early enzymes of the pathway. The key to *Pseudomonas denitrificans* fermentation is the addition of betaine (19). Vitamin B_{12} overproduction by *Pseudomonas denitrificans* is totally dependent on betaine, but the mechanism of control is unknown. Production of vitamin B_{12} has reached levels of 150 mg/L (18).

In production of biotin, feedback repression is caused by the enzyme protein acetyl-CoA carboxylase biotin holoenzyme synthetase, with biotin 5-adenylate acting as corepressor (20). Strains of *Serratia marcescens* obtained by mutagenesis, selection for resistance to biotin antimetabolites, and molecular cloning produce 600 mg/L in the presence of high concentrations of sulfur and ferrous iron (21).

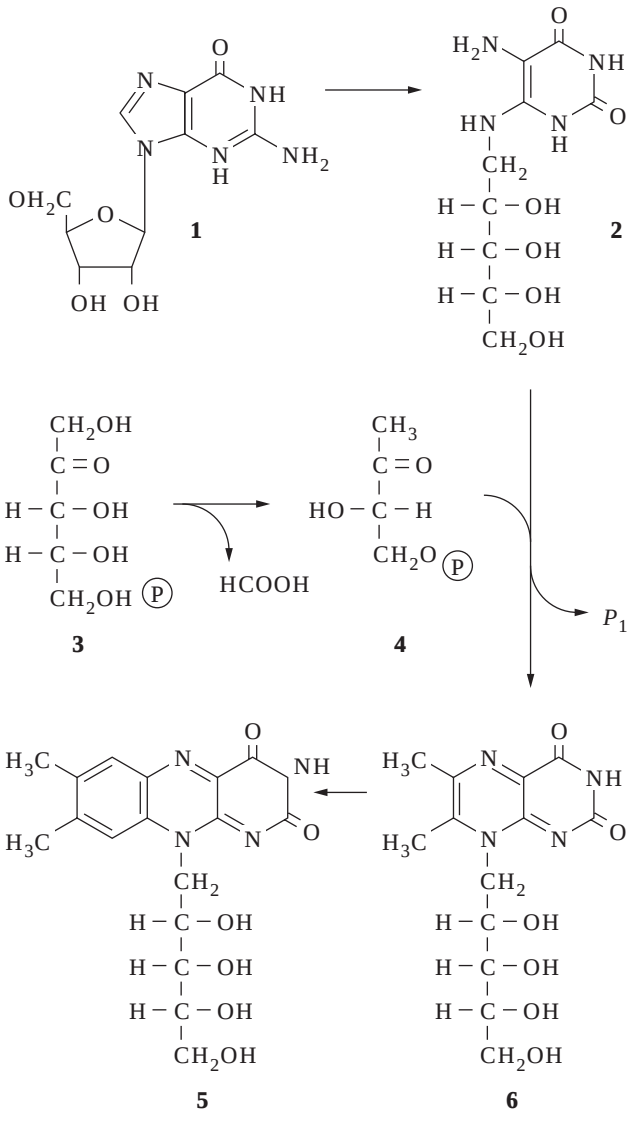


Figure 8. Biosynthesis of riboflavin. 1, guanosine triphosphate; 2, 5-amino-6-ribitylamino-2,4 (1H, 3H)-pyrimidinedione; 3, D-ribulose 5-phosphate; 4, L-3,4, dihydroxy-2-butanone 4-phosphate; 5, riboflavin; 6, 6,7-dimethyl-8-ribityllumazine. *Source*: From Ref. 14.

Such a titer is high enough to economically compete with the traditional chemical process. The pathway is shown in Figure 10.

A novel process for vitamin C (ascorbic acid) synthesis involves the use of a genetically engineered *Erwinia herbicola* strain containing a gene from *Corynebacterium* sp. The engineered organism converts glucose into 2-ketogulonic acid, which can be easily converted by acid or base to ascorbic acid (23). Another process devised independently converts 40 g/L glucose into 20 g/L 2-keto-L--gulonate (24). This process involves cloning of the gene encoding 2,5-diketo-D-gluconate reductase from *Corynebacterium* sp. into *Erwinia citreus*. Plasmid cloning of the genes encoding L-sorbose dehydrogenase and L-sorbosone dehydrogenase from *Gluconobacter oxydans* back into the same organism yielded a strain capable of converting 150 g/L D-sorbitol into 130 g/L 2-keto-L-gulonate (25).

Filamentous fungi have been widely used for the commercial production of organic acids. About 1 billion pounds of citric acid are produced each year. This organic acid is produced via the Embden–Meyerhof pathway and the first step of the tricarboxylic acid cycle (Fig. 11). The major control of the process involves the inhibition of phosphofructokinase by citric acid. The commercial process employs *Aspergillus niger* in media deficient in iron and manganese. Manganese deficiency has two beneficial effects in the citric acid fermentation: *(1)* it leads to high levels of intracellular NH_4, which reverse citric acid inhibition of phosphofructokinase, and *(2)* it brings on the formation of small mycelial pellets, which are the best morphological form for citric acid production. The morphological effect is due to a change in cell wall composition caused by growth in low Mn^{+}. A high level of citric acid production is also associated with a high intracellular concentration of fructose 2,6-biphosphate, an activator of glycolysis (27). Other factors that contribute to high citric acid production are the inhibition of isocitrate dehydrogenase by citric acid and the low pH optimum (1.7–2.0). Higher pH levels (e.g., 3.0) lead to production of oxalic and gluconic acids instead of citric acid. The low pH level inactivates glucose oxidase, which normally would yield gluconic acid (26). In approximately 4 to 5 days, the major portion (80%) of the sugar is converted to citric acid, titers reaching about 100 g/L. Because citric acid is easily assimilated, is palatable, and exhibits low toxicity, it is widely used in the food and pharmaceutical industries. It is employed as an acidifying and flavor-enhancing agent, as an antioxidant for inhibiting rancidity in fats and oils, as a buffer in jams and jellies, and as a stabilizer in a variety of foods. The pharmaceutical industry uses approximately 16% of the available supply of citric acid. Processes have been developed for the production of citric acid by *Candida* yeasts, especially from hydrocarbons. Such yeasts are able to convert *n*-paraffins to citric and isocitric acids in extremely high yields (150–170% on a weight basis). Production of citric acid instead of isocitric acid is favored by selecting mutants that are deficient in the enzyme aconitase. Titers as high as 225 g/L have been reached (26).

Cloning of the aldehyde dehydrogenase gene from *Acetobacter polyoxogenes* on a plasmid vector into *A. aceti* sp. *xylinum* increased the rate of acetic acid production by more than 100% (from 1.8 to 4 g L^{-1} h^{-1}) and titer by 40% (from 68 to 97 g/L) (28).

Ethyl alcohol is a primary metabolite that can be produced by fermentation of a sugar or a polysaccharide that can be depolymerized to a fermentable sugar. Yeasts are preferred for these fermentations, but the species used depends on the substrate employed. *Saccharomyces cerevisiae* is employed for the fermentation of hexoses, whereas *Kluyveromyces fragilis* or *Candida* species may be utilized if lactose or pentoses, respectively, are the substrates. Under optimum conditions, approximately 10 to 12% ethanol by volume is obtained within 5 days. Such a high concentration slows down growth and the fermentation ceases. With special yeasts, the fermentation can be continued to alcohol concentrations of 20% by volume.

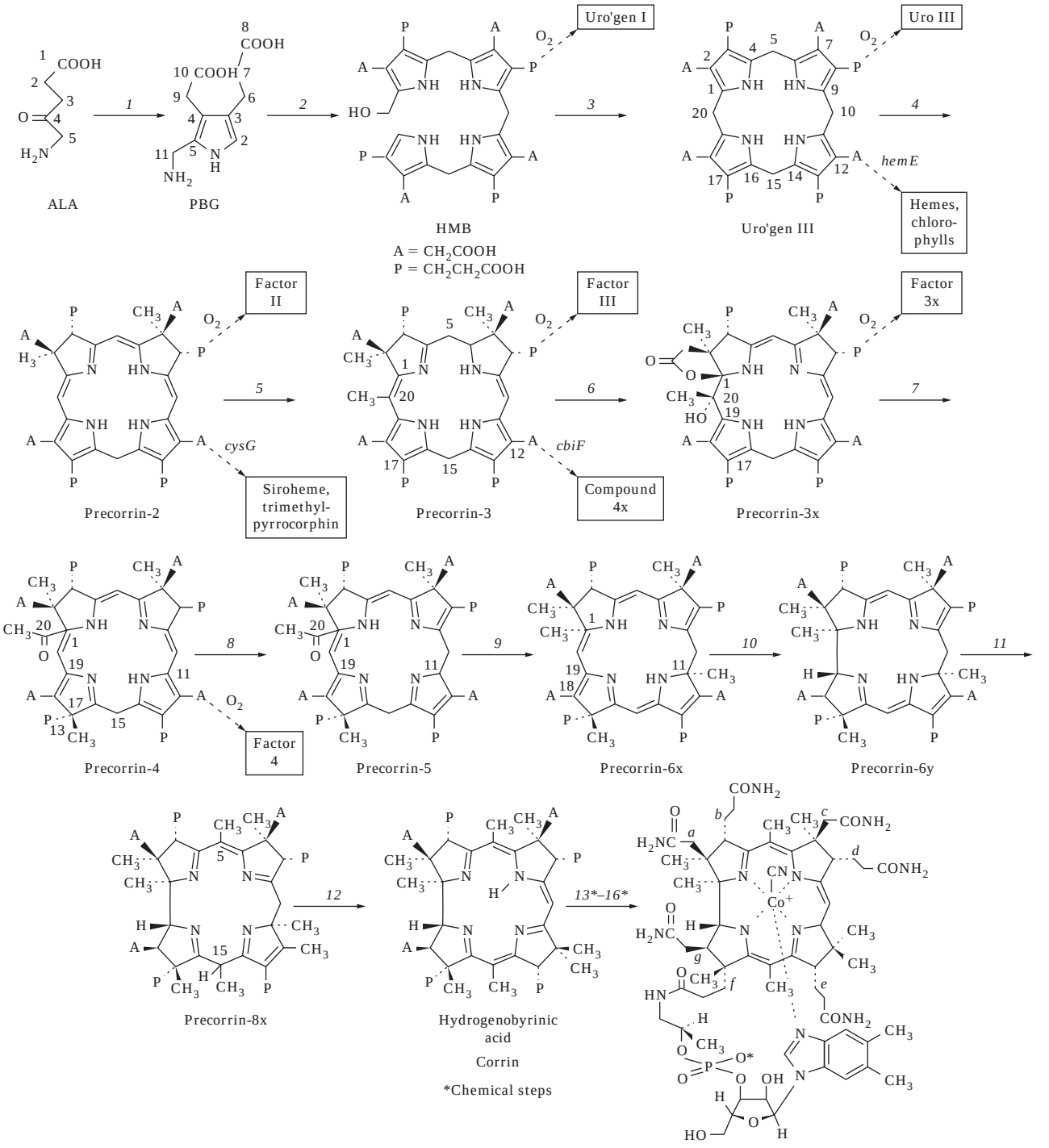


Figure 9. The biosynthetic pathway from δ-aminolevulinic acid to vitamin B_{12}. The enzyme catalyzing each reaction and its corresponding gene are 1, δ-aminolevulinic acid dehydratase (*hemB*); 2, porphobilinogen (PBG) deaminase (*hemC*); 3, uroporphyrinogen III (uro'gen III) synthase (*hemD*); 4, uro'gen III methyltransferase (*cysG*, *cobA*); 5, precorrin 2 methyltransferase (*cobI*); 6, precorrin 3 oxidase (*cobG*); 7, precorrin 3 hydroxylactone methyltransferase (ring contraction) (*cobJ*); 8, precorrin 4 methyltransferase (*cobM*); 9, precorrin 5 methyltransferase (loss of acetic acid) (*cobF*); 10, precorrin 6× reductase (*cobK*); 11, precorrin 6y decarboxylase and methyltransferase (*cobL*); 12, precorrin 8 methylmutase (*cobH*); 13, cobalt insertion; 14, esterification; 15, add nucleotide loop; and 16, ammonolysis. Possible side products are indicated in boxes. Asterisk indicates chemical step. *Source*: From Ref. 17.

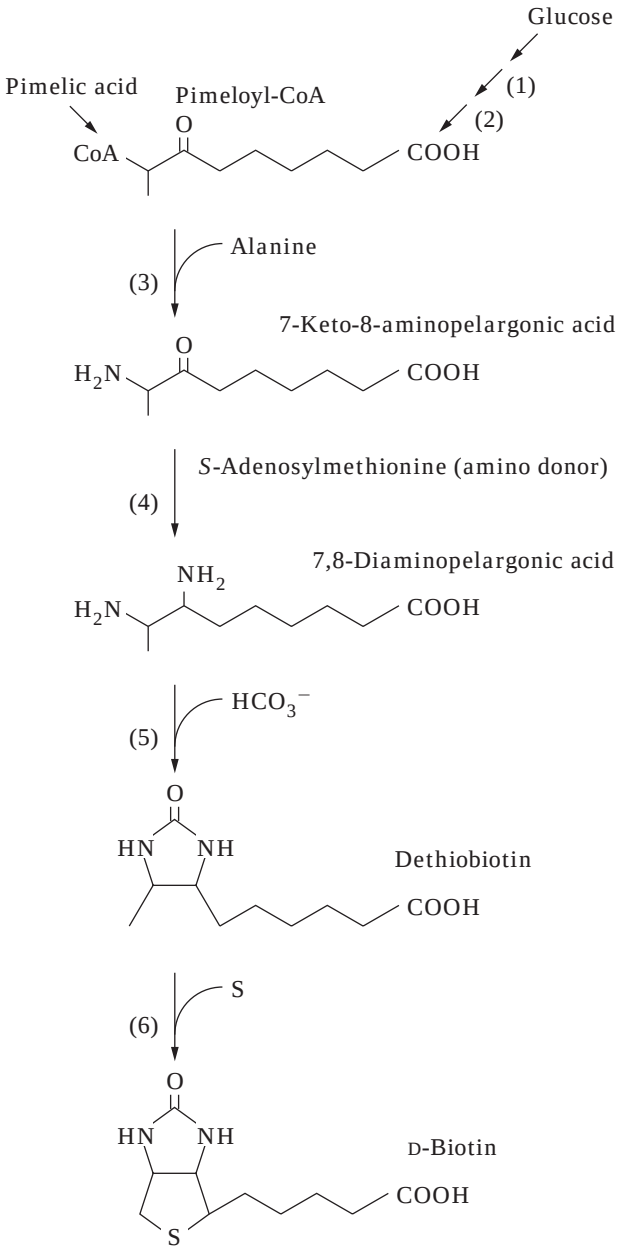


Figure 10. Pathway of biotin biosynthesis in bacteria. 1, *bioH*, block before pimeloyl-CoA; 2, *bioC*, block before pimeloyl-CoA; 3, *bioF*, 7-keto-8-aminopelargonic acid (KAPA) synthase; 4, *bioA*, 7,8-diaminopelargonic acid (DAPA) aminotransferase; 5, *bioD*, dethiobiotin synthase; 6, *bioB*, biotin synthase. *Source*: From Ref. 22.

However, these concentrations are attained only after months or years of fermentation. At present, industrial ethanol is mainly manufactured by fermentation, but there are still considerable amounts made by the petrochemical industry from ethylene. Bacteria such as clostridia and *Zymomonas* are being reexamined for ethanol production after years of neglect. *Clostridium thermocellum*, an anaerobic thermophile, can convert waste cellulose directly to ethanol. Other clostridia produce acetate, lactate, acetone, and butanol and will be utilized as petroleum becomes depleted in the world. Fuel ethanol produced from biomass would provide relief from air pollution caused by the use of gasoline and would not contribute to the greenhouse effect. Because of the elimination of lead from gasoline, ethanol is being substituted as a blend to raise gasoline's octane rating. *E. coli* has been converted into an excellent ethanol producer (43%, v/v) by recombinant DNA technology (29). Alcohol dehydrogenase II and pyruvate decarboxylase genes from *Zymomonas mobilis* were inserted in *E. coli* and became the dominant system for NAD regeneration. Ethanol represents more than 95% of the fermentation products in the genetically engineered strain.

In addition to the multireaction sequences of fermentations, microorganisms are extremely useful in carrying out biotransformation processes in which a compound is converted into a structurally related product by one or a small number of enzymes contained in the cells (30). Bioconverting organisms are known for practically every type of chemical reaction. One of the earliest and most famous is the biotransformation of progesterone to 11 α-hydroxyprogesterone (Fig. 12). The reactions are sterospecific, the ultimate in specificity being exemplified by the steroid bioconversions. This specificity is exploited in the resolution of racemic mixtures when a specific isomer rather than a racemic mixture is desired. Bioconversions are characterized by extremely high yields (i.e., 90 to 100%). Other attributes include mild reaction conditions and the coupling of reactions using a microorganism containing several enzymes working in series. There is a tremendous interest in immobilized cells to carry out such processes. These cells are usually much more stable than either free cells or enzymes and are more economical than immobilized enzymes. Recombinant DNA techniques have been useful in developing new bioconversions. For example, the cloning of the fumarase-encoding gene in *Saccharomyces cerevisiae* improved the bioconversion of malate to fumarate from 2 to 125 g/L in a single manipulation (31). The conversion yield using the constructed strain is near 90%.

Microbially produced secondary metabolites are extremely important to our health and nutrition (32). As a group that includes antibiotics, other medicinals, toxins, pesticides, and animal and plant growth factors, they have tremendous economic importance. In batch or fed-batch culture, secondary metabolites are usually produced after growth has slowed down. They have no function in growth of the producing cultures (although they contribute to survival of producing organisms in nature), are produced by certain restricted taxonomic groups of organisms, and are usually formed as mixtures of closely related members of a chemical family. From 1990 to 1994, more than 1,000 new secondary metabolites were characterized from actinomycetes alone (33).

The best known of the secondary metabolites are the antibiotics (Figs. 13 and 14). This remarkable group of compounds form a heterogeneous assemblage of biologically active molecules with different structures and modes of action (Table 2). In 1996 the antibiotic market involved 160 antibiotics and a $23 billion market (36). Since 1940 we have witnessed a virtual explosion of new and potent antibiotic molecules that have been of great use in medicine, agriculture, and basic research. However, the search

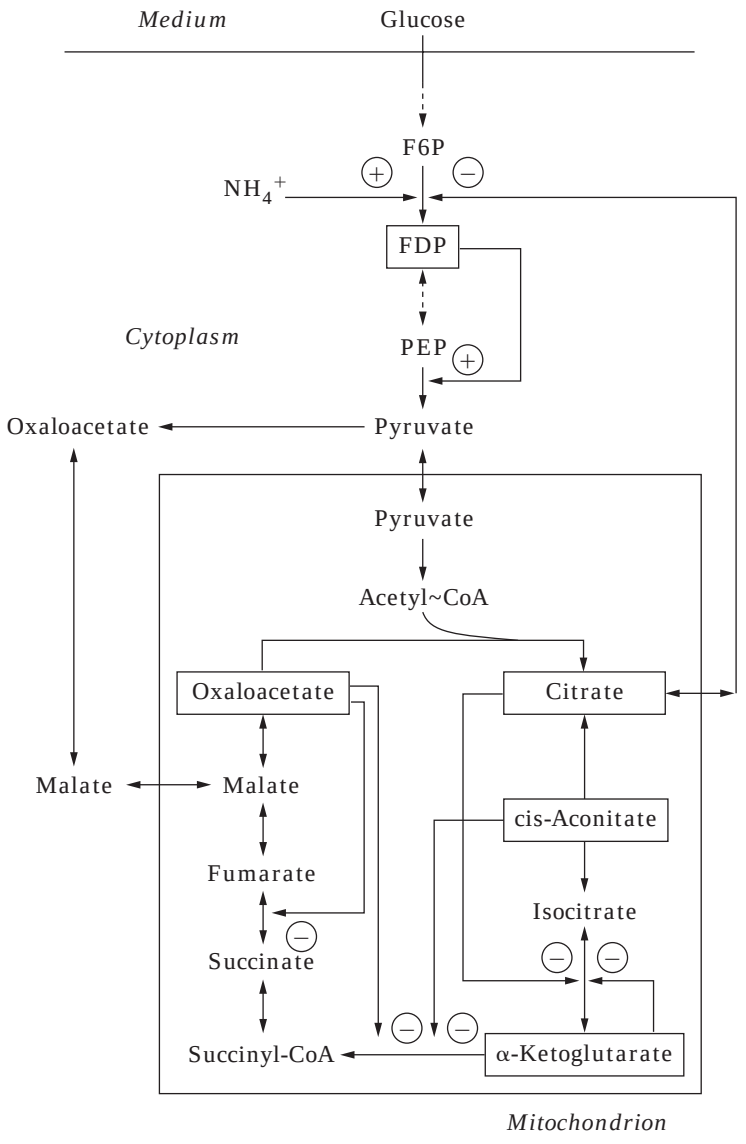


Figure 11. Metabolic scheme of citric acid overproduction. +, activation by the respective metabolite; −, inhibition. *Source*: From Ref. 26.

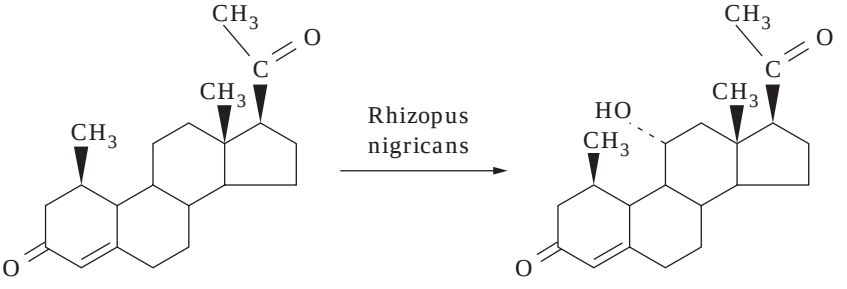


Figure 12. 11α-Hydroxylation of progesterone by *Rhizopus nigricans*.

for new antibiotics continues in order to combat evolving pathogens and naturally resistant bacteria and fungi and previously susceptible microbes that have developed resistance; improve pharmacological properties; combat tumors, viruses, and parasites; and discover safer, more potent, and broader spectrum compounds. About 6,000 antibiotics have been described, 4,000 from actinomycetes alone, and they still are being discovered at a rate of about 300 per year. *Streptomyces griseus* strains produce more than 40 different antibiotics and *Bacillus subtilis* more than 60 different antibiotics. Strains of *S. hygroscopicus* make almost 200 antibiotics. One *Micromonospora* strain

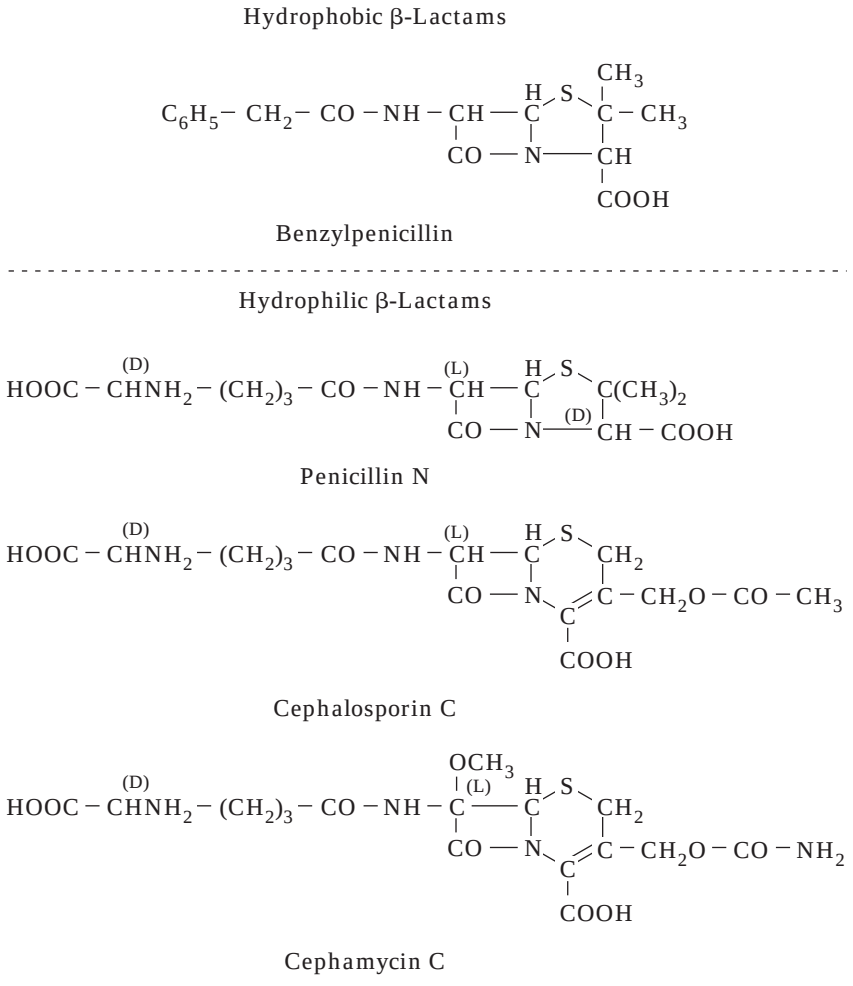


Figure 13. Some *β*-lactam antibiotics. Benzylpenicillin is also known as penicillin G.

can produce 48 aminocyclitol antibiotics. The antibiotics vary in size from small molecules such as cycloserine (102 Da) and bacilysin (270 Da) to polypeptides such as nisin, which contains 34 amino acid residues. They attack virtually every type of microbial activity, including DNA, RNA, and protein synthesis; membrane function; electron transport; sporulation; germination; and many others. In the search for new antibiotics, many of the new products are made by chemists by modification of natural antibiotics; this process is called semisynthesis. By 1974 more than 20,000 semisynthetic penicillins, 4,000 cephalosporins, 2,500 tetracylines, 1,000 rifamycins, 500 kanamycins, and 500 chloramphenicols had been prepared. In 1980 worldwide antibiotic production amounted to about 25,000 tons. Included were 17,000 tons of penicillins, 5,000 tons of tetracyclines, 1,200 tons of cephalosporins, and 800 tons of erythromycins. Antibiotics are used not only chemotherapeutically in human and veterinary medicine but also for growth promotion in farm animals and for the protection of plants.

In nature, secondary metabolites are important for the organisms that produce them, functioning as *(1)* sex hormones; *(2)* ionophores; *(3)* competitive weapons against other bacteria, fungi, amoebae, insects, and plants; *(4)* agents of symbiosis; and *(5)* effectors of differentiation (37). For years, the major pharmaceuticals (such as hypertensive and antiinflammatory agents) used for noninfectious diseases were strictly synthetic products. Similarly, major therapeutics for nonmicrobial parasitic diseases in animals (e.g., coccidiostats and antihelmintics) came from the screening of synthesized compounds followed by molecular modification. Despite the testing of thousands of synthetic compounds, only a few promising structures were uncovered. As new lead compounds became more and more difficult to find, microbial broths filled the void and microbial products increased in importance in treatment of nonmicrobial diseases. Today, microbially produced polyethers (38) such as monensin, lasalocid, and salinomycin dominate the coccidiostat market and are also the chief growth promotants in use for ruminant animals. The avermectins (Fig. 14), another group of streptomycete products with a market of more than $200 million per year, have high activity against helminths and arthropods (39). Indeed, their activity is an order of magnitude greater than previously discovered antihelmintic agents, the vast majority of which were synthetic. Known pharmacological activities of microbial secondary metabolites are shown in Table 3.

Many microbial products with important pharmacological activities were discovered by screening for inhibitors

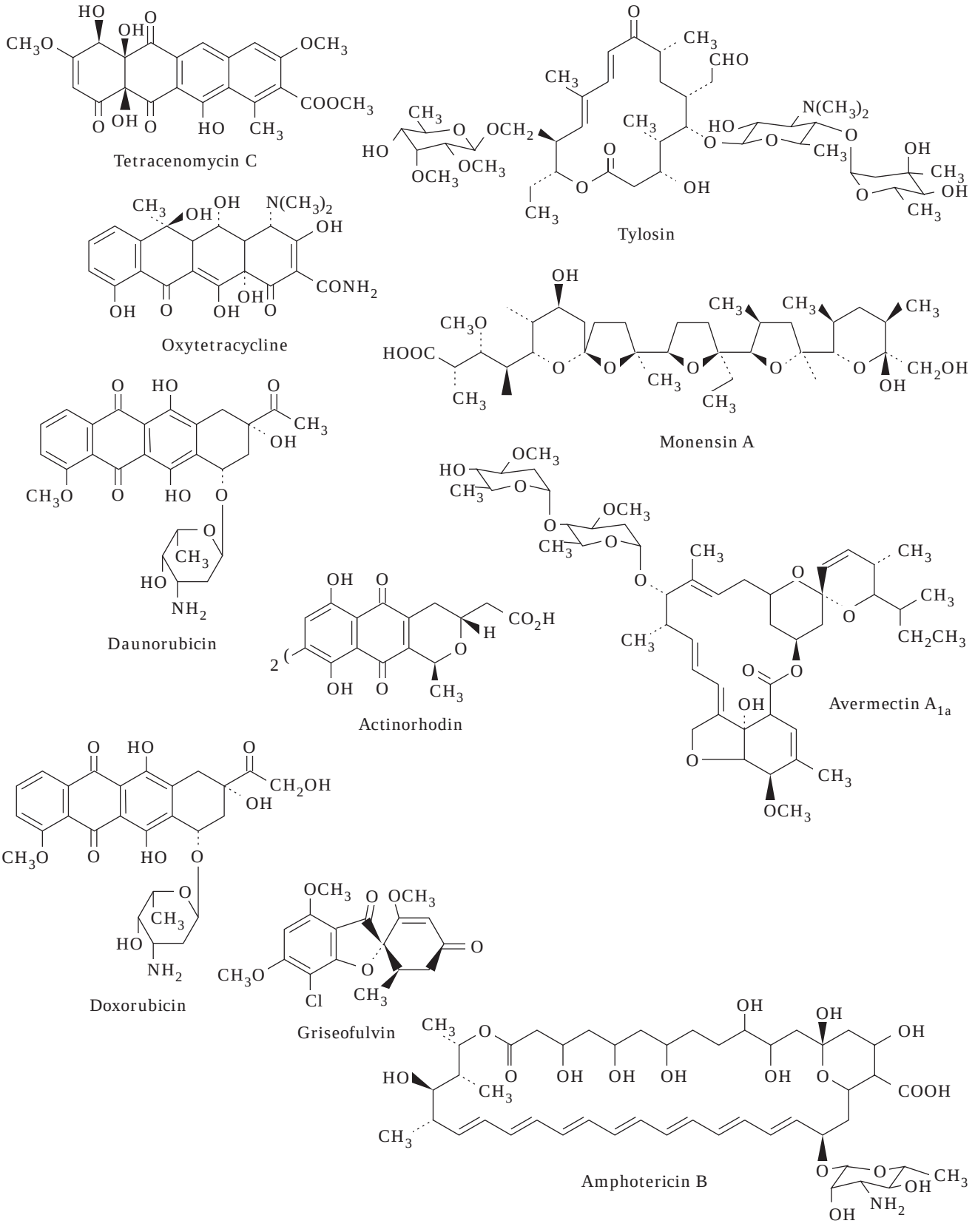


Figure 14. Some secondary metabolites of polyketide orgin. All except avermectin possess antibiotic activity. *Source*: From Refs. 34 and 35.

Table 2. Mode of Action and Structural Type of Selected Antibiotics

Target	Antibiotic	Structure
DNA replication	Bleomycin	Peptide
	Griseofulvin	Condensed aromatic
	Actinomycin	Peptide
Transcription	Rifamycin	Ansamycin
Translation		
70-S ribosomes	Chloramplenicol	Aromatic
	Tetracycline	Polyketide
	Lincomycin	Sugar ester
	Erythromycin	Macrolide
	Streptomycin	Aminocyclitol
70-S and 80-S ribosomes	Puromycin	Nucleoside
	Fusidic acid	Steroid
80-S ribosomes	Cycloheximide	Glutarimide
Cell wall synthesis	Cycloserine	Amino acid
	Bacitracin	Peptide
	Penicillin	β-Lactam
	Vancomycin	Oligosaccharide aromatic
Cell membrane		
Surfactants	Polymyxin	Peptide
	Amphotericin	Polyene
Ionophores		
Channel formers	Linear gramicidin	Peptide
Mobile carriers	Monensin	Polyether

using simple enzymatic assays (40). One huge success has been lovastatin (mevinolin), a fungal product of the compactin group (41) that acts as a cholesterol-lowering agent in animals. Lovastatin is produced by *Aspergillus terreus*. In its hydroxy acid form (mevinolinic acid), mevinolin is a potent competitive inhibitor of 3-hydroxy-3-methylglutaryl-coenzyme A reductase from liver. Another enzyme inhibitor on the market is acarbose, a natural inhibitor of intestinal glucosidase (42), which is produced by an actinomycete of the genus *Actinoplanes*. It decreases hyperglycemia and triglyceride synthesis in adipose tissue, liver, and the intestinal wall of patients suffering from diabetes, obesity, and type IV hyperlipidemia.

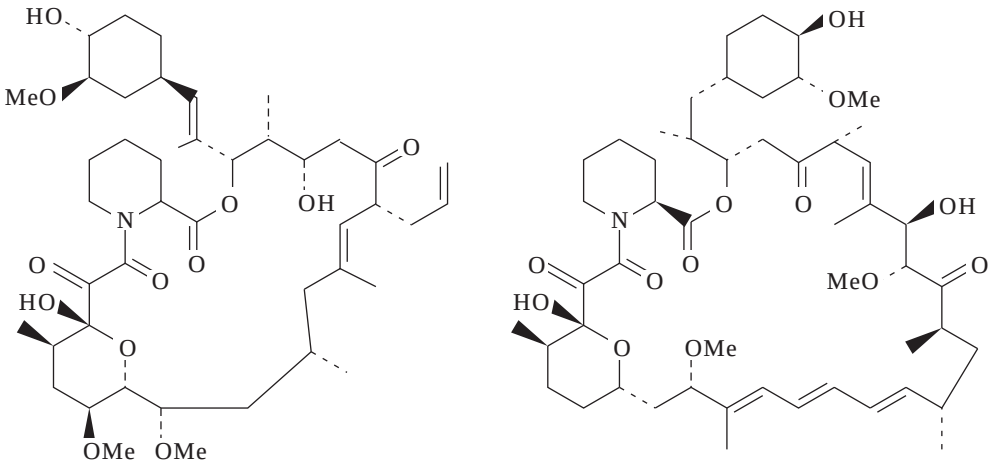
Also in commercial or near-commercial use are biopesticides including fungicides (e.g., kasugamycin, polyoxins), insecticides (*Bacillus thuringiensis* crystal, nikkomycin, spinosyns), herbicides (bialaphos), antihelmintics and coccidiostats as mentioned earlier, ruminant growth promoters (monensin, lasalocid, salinomycin), plant growth regulators (gibberellins), immunosuppressants for organ transplants (cyclosporin A, FK-506, rapamycin [Fig. 15]), anabolic agents in farm animals (zearelanone), uterocontractants (ergot alkaloids), and antitumor agents (doxorubicin, daunorubicin, mitomycin, bleomycin) (43,44). A remarkable aspect of this development is that many of the compounds were first isolated as poor or toxic antibiotics (e.g., monensin, cyclosporin, rapamycin) or as mycotoxins (ergot alkaloids, gibberellins, zearelanone) before they were put to work for our benefit. Following in their footsteps are many secondary metabolites with potent activities that are under investigation right now. Indeed, antibiotic activity is merely the tip of the iceberg, a mere scratching at the surface of the potential of microbial activity.

In batch culture, some secondary metabolite processes exhibit a distinct growth phase (trophophase) followed by a production phase (idiophase). In other fermentations, trophophase and idiophase overlap; the timing depends on the nutritional environment presented to the culture, the growth rate, or both. Secondary metabolite production in nature probably allows organisms to differentiate and to compete effectively with other forms of life. Delay in antibiotic production until after trophophase aids the producing organism, because the microbe is usually sensitive during growth to its own antibiotic. Resistance mechanisms in producers include enzymatic modification of the antibiotic, alteration of the cellular target of the antibiotic, and

Table 3. Some Pharmacological Activities of Microbial Secondary Metabolites

ACTH-like	Diuretic	Erythematous	Leukemogenic
Anabolic	Edematous	Estrogenic	Motility inhibition
Analeptic	Emetic	Fertility enhancing	Nephrotoxic
Anesthetic	Enzyme inhibitory	Hallucinogenic	Neuromuscular blockade
Anorectic	Aldol reductase	Hemolytic	Neurotoxic
Anticoagulant	α-Amylase	Hemostatic	Paralytic
Antidepressive	Amyloglucosidase	Herbicidal	Parasympathomimetic
Antihelminthic	Catechol-*O*-methyltransferase	Hormone releasing	Photosensitizing
Antiinfective	Cholinesterase	Hypersensitizing	Relaxant (smooth muscle)
Antiinflammatory	Cyclic AMP phosphodiesterase	Hypocholesterolemic	Sedative
Antiparasitic	Dopamine β-hydroxylase	Hypoglycemic	Serotonin antagonist
Antispasmodic	Elastase	Hypolipidemic	Spasmolytic
Carcinogenesis inhibition	Esterase	Hypotensive	Telecidal
Coagulative (blood)	Glucosidase	Immunostimulating	Ulcerative
Complement inhibition	Invertase	Inflammatory	Vasodilative
Convulsant	β-Lactamase	Insecticidal	
Dermonecrotic	Protease		
Diabetogenic	Tyrosine hydroxylase		

Note: List does not include antimicrobial, antiviral, and antitumor activities.



FK-506 Rapamycin

Figure 15. Two new immunosuppressants made by streptomycetes.

decreased uptake of excreted antibiotic. The raw materials of secondary metabolism are primary metabolites, which often exert a negative or positive effect on secondary metabolism.

Manipulation of culture media in any development program often involves the testing of hundreds of additives as possible limiting precursors of the desired product. Occasionally, a precursor that increases production of the secondary metabolite is found. The precursor may also direct the fermentation toward the formation of one specific desirable product; this is known as *directed biosynthesis*. Examples of the latter include phenylacetic acid in benzylpenicillin fermentation, specific amino acids in the production of actinomycins and tyrocidins, and substituted benzoic acids in the formation of novobiocins. In many fermentations, however, precursors show no activity because their syntheses are not rate limiting. In such cases, screening of additives has often revealed dramatic effects, both stimulatory and inhibitory, of nonprecursor molecules on the production of secondary metabolites. These effects are usually due to interaction of these compounds with the regulatory mechanisms existing in the fermentation organism.

The regulatory mechanisms prevalent in secondary metabolism are essentially the same as found in primary metabolism (45). Thus chloramphenicol limits its own synthesis by repressing arylamine synthetase, and ergot alkaloids limit their production by inhibiting dimethylallyl transferase; these are the initial enzymes unique to the secondary pathway. Glucose, a sugar usually used rapidly for growth, exerts carbon catabolite repression of secondary metabolite synthesis. Similarly, ammonium ion or rapidly used amino acids exert nitrogen metabolite regulation on secondary metabolism. Inorganic phosphate exerts a strong negative effect in secondary metabolite fermentations. Enzyme induction is evident in the production of alkaloids and antibiotics. In actinomycetes, γ-butyrolactones function as autoinducers of secondary metabolism. Besides the specific types of regulation mentioned earlier, the onset of secondary metabolism is also controlled by growth rate. Generally, secondary metabolites are not produced at growth rates near the maximum specific growth rate. In fact, the idiophase can be shifted into the trophophase when media that support only low rates of growth are used. Antibiotic biosynthesis ends via the decay of antibiotic synthetases or because of feedback inhibition and repression of these enzymes. Because the regulatory mechanisms are genetically determined, mutation has had a major effect on the production of secondary metabolites (46). Indeed, it is the chief factor responsible for the hundred- to thousandfold increases obtained in production of antibiotics from the time of their initial discovery to the present.

Environmental manipulations have proven to be quite useful for bringing on metabolite overproduction. Optimization of fermentation parameters such as temperature, pH, oxygen transfer, and nutrients is of course extremely important. If a process is inducible, inducers are added. They include methionine for cephalosporin C formation, valine for tylosin production, and tryptophan for ergot alkaloid production. Virtually all fermentation processes operate, after the initial growth stage, under conditions of unbalanced growth, and from the start of any development program, the microbiologist must find the key to this unbalanced growth. The key is often a specific nutrient limitation that restricts growth and brings on product accumulation.

Genetics has had a long history of contributing to the production of microbial secondary metabolites, such as antibiotics (47). The tremendous increases in fermentation productivity and the resulting decreases in costs have come about mainly by mutagenesis and screening for higher-producing microbial strains. Mutation has also served to *(1)* shift the proportion of metabolites produced in a fermentation broth to a more favorable distribution, *(2)* elucidate the pathways of secondary metabolism, and *(3)* yield new compounds. The medically useful products demethyltetracycline and doxorubicin (adriamycin) were discovered by simple mutation of the cultures producing

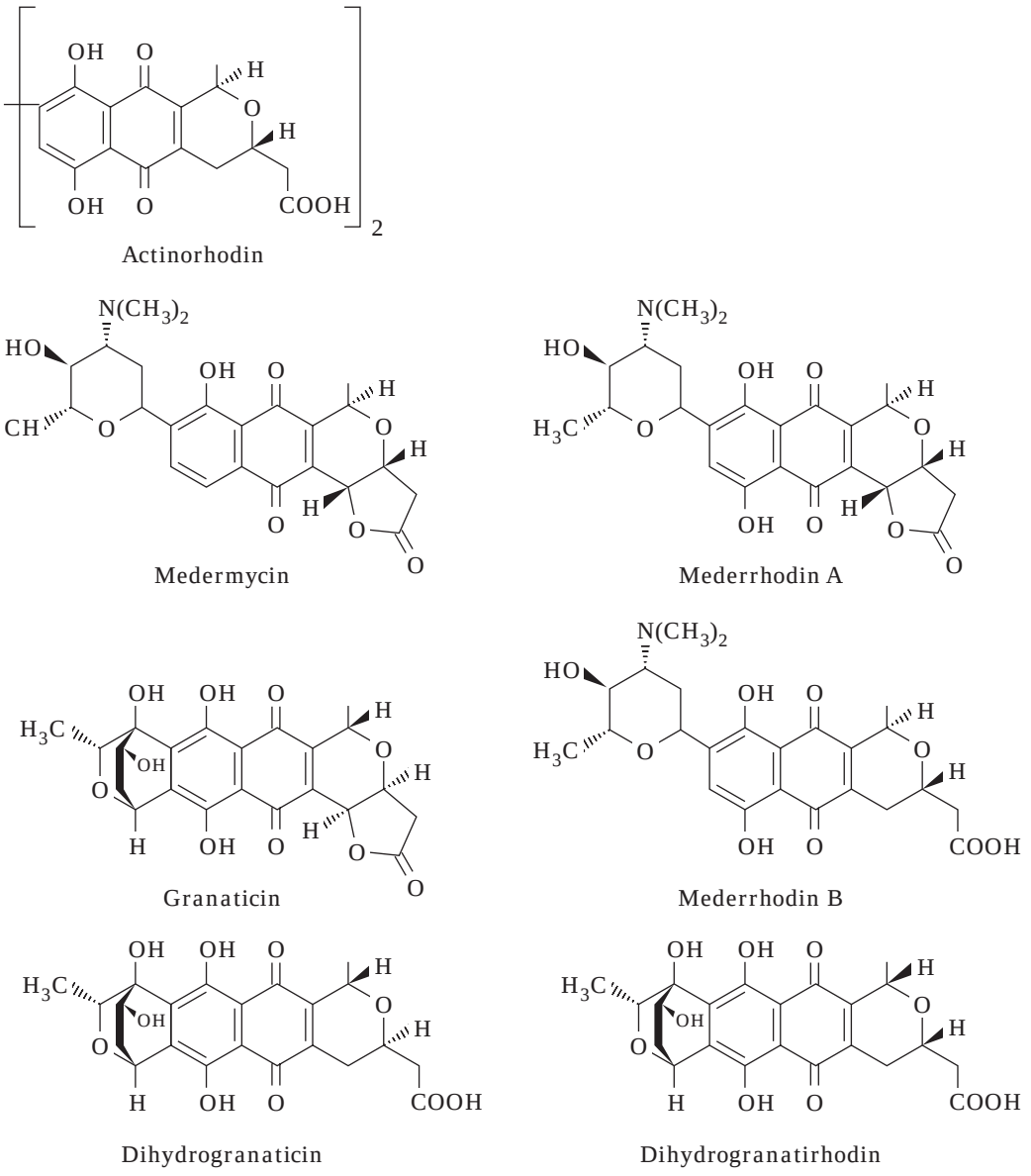


Figure 16. Hybrid antibiotics. Left, structures of the natural isochromane quinone antibiotics: actinorhodin, granaticin, dihydrogranaticin, and medermycin. Right, structures of new hybrid antibiotics: mederrhodin A and B and dihydrogranatirhodin produced by genetically engineered organisms (the absolute stereochemistry of dihydrogranaticin at C-1 and C-3 is tentative). *Source*: From Ref. 52.

tetracycline and daunorubicin (daunomycin), respectively. The technique of mutational biosynthesis has been used for the discovery of many new aminoglycoside, macrolide, and anthracycline antibiotics. It has recently been successfully employed in producing a new commercial avermectin (48).

The use of protoplast fusion has facilitated genetic recombination of different strains from the same or different species to yield new antibiotics. Protoplast fusion has also been useful in elimination of an undesirable component from penicillin broths.

Recombinant DNA technology has been applied to the production of antibiotics (49). Many genes that encode individual enzymes of antibiotic biosynthesis have been cloned and expressed at high levels in heterologous microorganisms. Continued efforts in the application of recombinant DNA technology to fermentations have led to overproduction of limiting enzymes of important biosynthetic pathways, thereby increasing production of the final products. This has been accomplished in the area of cephalosporin biosynthesis (50). In addition, many antibiotic resistance genes from antibiotic-producing organisms have been cloned and expressed. Some antibiotic biosynthetic pathways are encoded by plasmid-borne genes (e.g., methylenomycin A). Even when the antibiotic biosynthetic pathway genes of actinomycetes are chromosomal (the

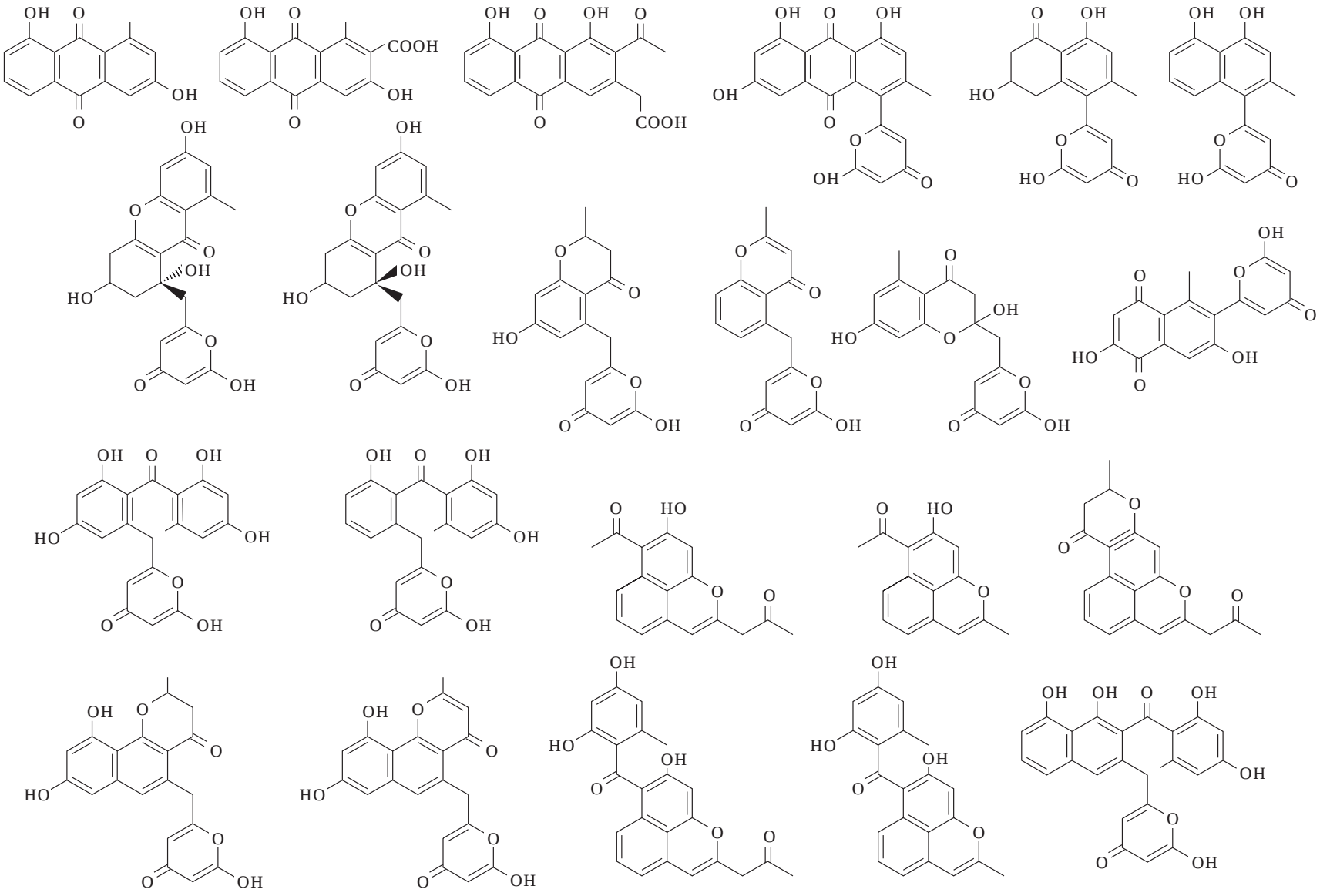


Figure 17. Novel polyketides produced by the actinorhodin (*act*), frenolicin (*fren*), and tetracenomycin (*tcm*) PKS gene clusters. *Source*: From Ref. 47.

usual situation), they are clustered, which facilitates transfer of an entire pathway in a single manipulation. The genes of the actinorhodin pathway, normally clustered on the chromosome of *Streptomyces coelicolor*, were transferred en masse on a plasmid to *S. parvulus* and were expressed in the latter organism. The presence of the bialophos (a herbicidal antibiotic) resistance gene in the bialaphos biosynthetic cluster allowed cloning of the entire 13-enzyme pathway of this herbicidal antibiotic from *S. hygroscopicus* into *S. lividans*. Here, the resistance gene is also one of the biosynthetic genes, coding for demethylphosphinothricin acetyltransferase. Even in fungi, pathway genes can be clustered such as the penicillin genes in *Penicillium* or the aflatoxin genes in *Aspergillus*. However, the genes that code for cephalosporin biosynthesis in *Cephalosporium* are distributed among different chromosomes.

For the discovery of new or modified products, recombinant DNA techniques can be used to introduce genes that code for antibiotic synthetases into producers of other antibiotics or into nonproducing strains to obtain modified or hybrid antibiotics. Gene transfer from a streptomycete strain producing the isochromanequinone antibiotic actinorhodin into strains producing granaticin, dihydrogranaticin, and mederomycin (which are also isochromanequinones) led to the discovery of two new antibiotic derivatives, mederrhodin A and dihydrogranatirhodin (51,52; Fig. 16). A hybrid antibiotic, isovalerylspiramycin, was obtained by introducing the *carE* gene from the carbomycin producer *S. thermotolerans* into the spiramycin producer *S. ambofaciens* (53). These are both 16-membered macrolides, but one of the two sugars in carbomycin is isovalerylmycarose and the corresponding sugar in spiramycin is mycarose. When the acylating gene was cloned into the nonmacrolide producer *Streptomyces lividans*, the recombinant was capable of converting spiramycin to isovalerylspiramycin. Many novel polyketide secondary metabolites have been obtained by cloning DNA fragments from one polyketide producer into various strains of other streptomycetes (54,55). Figure 17 shows some of these novel polyketide molecules.

BIBLIOGRAPHY

1. L. Palmieri, D. Berns, R. Krämer, and M. Eikmanns, *Arch. Microbiol.* **165**, 48–54 (1996).
2. L. Eggeling, *Amino Acids* **6**, 261–272 (1994).
3. K. Sonntag, J. Schwinde, A.A. de Graaf, A. Marx, B.J. Eikmanns, W. Wiechert, and H. Sahm, *Appl. Microbiol. Biotechnol.* **44**, 489–495 (1995).
4. A. Erdmann, B. Well, and R. Krämer, *J. Gen. Microbiol.* **139**, 3115–3122 (1993).
5. M.S.M. Jetten and A.J. Sinskey, *Crit. Rev. Biotechnol.* **15**, 73–103 (1995).
6. H. Sahm, L. Eggeling, B. Eikmanns, and R. Krämer, *FEMS Microbiol. Rev.* **16**, 243–252 (1995).
7. R. Kelle, T. Hermann, D. Weuster-Botz, L. Eggeling, R. Krämer, and C. Wandrey, *J. Biotechnol.* **50**, 123–136 (1996).
8. H. Ito, K. Sato, K. Matsui, K. Sano, H. Enei, and Y. Hirose, *Appl. Microbiol. Biotechnol.* **33**, 190–195 (1990).
9. A.L. Demain, in A.H. Rose ed., *Economic Microbiology, vol. 2: Primary Products of Metabolism*, Academic Press, New York, 1978, pp. 187–208.
10. A. Kuninaka, in H.-J. Rehm, G. Reed, A. Pühler, and P. Stadler eds., *Biotechnology*, vol. 6, 2nd ed., VCH Verlagsgesellschaft, Weinheim, 1996, pp. 561–612.
11. K. Miyagawa, N. Kanzaki, H. Kimura, Y. Sumino, S. Akyama, and Y. Nakao, *Bio / Technology* **7**, 821–824 (1989).
12. K. Miyagawa, H. Kimura, K. Nakahama, M. Kikuchi, M. Doi, S. Akiyama, and Y. Nakao, *Bio / Technology* **4**, 225–228 (1986).
13. S. Asahi, M. Izawa, and M. Doi, *7th International Symposium of the Genetics of Industrial Microorganisms*, 1994, abstract P.17, p. 102.
14. R. Volk and A. Bacher, *J. Biol. Chem.* **266**, 20610–20618 (1991).
15. A.L. Demain, *Annu. Rev. Microbiol.* **26**, 369–388 (1972).
16. E.J. Vandamme, *J. Chem. Technol. Biotechnol.* **53**, 313–327 (1992).
17. C.A. Roessner and A.I. Scott, *Annu. Rev. Microbiol.* **50**, 467–490 (1996).
18. C. Spalla, A. Grein, L. Garofano, and G. Ferni, in E.J. Vandamme ed., *Biotechnology of Vitamins, Pigments, and Growth Factors*, Elsevier, New York, 1989, pp. 257–284.
19. J.P. Kusel, Y.H. Fa, and A.L. Demain, *J. Gen. Microbiol.* **130**, 835–841 (1984).
20. D.F. Barker and A.M. Campbell, *J. Mol. Biol.* **146**, 469–492 (1981).
21. M. Masuda, K. Takahashi, N. Sakurai, K. Yanagiya, S. Komatsubara, and T. Tosa, *Process Biochem.* **30**, 553–562 (1995).
22. N. Sakurai, H. Akatsuka, E. Kawai, Y. Imai, and S. Komatsubara, *Microbiology* **142**, 3295–3303 (1996).
23. M.J. Pramik, *Gen. Eng. News* **2**, 9, 12 (1986).
24. J.F. Grindley, M.A. Payton, H. van de Pol, and K.G. Hardy. *Appl. Environ. Microbiol.* **54**, 1770–1175 (1988).
25. Y. Saito, Y. Ishii, H. Hayashi, Y. Imao, T. Akashi, K. Yoshikawa, Y. Noguchi, S. Soeda, M. Yoshida, M. Niwa, J. Hosoda, and K. Shimomura, *Appl. Environ. Microbiol.* **63**, 454–460 (1997).
26. C.P. Kubicek and M. Röhr, *CRC Crit. Rev. Biotechnol.* **3**, 331–373 (1986).
27. H.J.M. Harmsen, E.M. Kubicek-Pranz, M. Röhr, J. Visser, and C.P. Kubicek, *Appl. Microbiol. Biotechnol.* **37**, 784–788 (1992).
28. M. Fukaya, K. Tayama, T. Tamaki, H. Tagami, H. Okumura, Y. Kawamura, and T. Beppu, *Appl. Environ. Microbiol.* **55**, 171–176 (1989).
29. L.O. Ingram, E. Conway, D.P. Clark, G.W. Sewell, and J.F. Preston, *Appl. Environ. Microbiol.* **53**, 2420–2425 (1987).
30. J.P. Rosazza, *Microbial Transformations of Bioactive Compounds*, vols. 1 and 2, CRC Press, Boca Raton, Fla., 1982.
31. R.J. Neufeld, Y. Peleg, J.S. Rokem, O. Pines, and I. Goldberg, *Enzyme Microb. Technol.* **13**, 991–996 (1991).
32. A.L. Demain and A. Fang, *Actinomycetologia* **9**, 98–117 (1995).
33. J.J. Sanglier, H. Haag, T.A. Huck, and T. Fehr, *Exp. Opin. Invest. Drugs* **5**, 207–223 (1996).
34. C.R. Hutchinson and I. Fujii, *Annu. Rev. Microbiol.* **49**, 201–238 (1995).
35. C.R. Hutchinson, H. Decker, K. Madduri, S.L. Otten, and L. Tang, *Antonie van Leeuwenhoek* **64**, 165–176 (1993).
36. K.S. Brown, *Scientist* **10**, 1, 8, 9 (1996).
37. A.L. Demain, in B. Sutton ed., *A Century of Mycology*, Cambridge Univ. Press, Cambridge, U.K., 1996, pp. 233–254.

38. J.W. Westley, *Adv. Appl. Microbiol.* **22**, 177–223 (1977).
39. W.C. Campbell, M.H. Fisher, E.O. Stapley, G. Albers-Schönberg, and T.A. Jacob, *Science* **221**, 823–828 (1983).
40. H. Umezawa, *Annu. Rev. Microbiol.* **36**, 75–99 (1982).
41. A. Endo, *J. Med. Chem.* **28**, 401–405 (1985).
42. L. Müller, in H.-J. Rehm and G. Reed eds., *Biotechnology*, vol. 4, V.H. Verlagsgesellschaft, Weinheim, 1986, pp. 1–37.
43. A.L. Demain, *Science* **219**, 709–714 (1983).
44. D.J. Chadwick and J. Whelan eds., *Secondary Metabolites: Their Function and Evolution*, Wiley, New York, 1992.
45. A.L. Demain, in W. Kuhn and H.-P. Fiedler eds., *Sekundärmetabolismus bei Mikroorganismen, Beiträge zur Forschung*, Attempto Verlag, Tübingen, 1995, pp. 9–35.
46. A.L. Demain, *Adv. Appl. Microbiol.* **16**, 177–202 (1973).
47. G. Holt and G. Saunders, in A.T. Bull and H. Dalton eds., *Comprehensive Biotechnology*, vol. 1, Pergamon Press, New York, 1985, pp. 51–76.
48. C.D. Denova, R.W. Fedechko, E.W. Hafner, H.A.I. McArthur, M.R. Morgenstern, D.D. Skinner, K. Stutzman-Engwall, R.G. Wax, and W.C. Wernau, *J. Bacteriol.* **177**, 3504–3551 (1995).
49. K.F. Chater, *Bio / Technology* **8**, 115–121 (1990).
50. P.L. Skatrud, A.J. Tietz, T.D. Ingolia, C.A. Cantwell, D.L. Fisher, J.L. Chapman, and S.W. Queener, *Bio / Technology* **7**, 477–485 (1989).
51. D.A. Hopwood, F. Malpartida, H.M. Kieser, H. Ikeda, J. Duncan, I. Fujui, B.A.M. Rudd, H.G. Floss, and S. Omura, *Nature* **314**, 642–644 (1985).
52. H.G. Floss, *TIBTECH* **5**, 111–115 (1987).
53. J.K. Epp, M.L.B. Huber, T. Goodson, and B.E. Schoner, *Gene* **85**, 293–301 (1989).
54. W.R. Strohl, P.L. Bartel, N.C. Connors, C.-B. Zhu, D.C. Dosch, J.M. Beale, Jr., H.G. Floss, K. Stutzman-Engwall, S.L. Otten, and C.R. Hutchinson, in C.L. Hershberger, S.W. Queener, and G. Hegeman eds., *Genetics and Molecular Biology Industrial Microorganisms*, American Society for Microbiology, Washington, D.C., 1989, pp. 68–84.
55. C. Khosla, R. Caren, C.M. Kao, R. McDaniel, and S.-W. Wang, *Biotechnol. Bioeng.* **52**, 122–128 (1996).

See also AMINO ACIDS, PRODUCTION PROCESSES; *ASPERGILLUS*; CEPHALOSPORINS; DEXTRAN, MICROBIAL PRODUCTION METHODS; PANTOTHENIC ACID AND RELATED COMPOUNDS; SECONDARY METABOLITE PRODUCTION, ACTINOMYCETES, OTHER THAN STREPTOMYCES.

METHANOTROPHS

IAN R. MCDONALD
J. COLIN MURRELL
University of Warwick
Coventry, U.K.

KEY WORDS

Methane monooxygenase
Methane monooxygenase gene
Methane oxidation
Methanotroph
Methylotroph
Polymerase chain reaction

OUTLINE

INTRODUCTION

Methane oxidizing bacteria (methanotrophs) are defined as those bacteria that can grow on methane as their sole source of carbon and energy. They are therefore a subgroup of methylotrophs (bacteria that can grow on one-carbon compounds). Methanotrophs are generally regarded as obligate in nature, growing only on methane and usually on methanol but not on multicarbon substrates. Methanotrophs have attracted considerable interest over the past 25 years due to their potential for the production of bulk chemicals, single-cell protein, and use in biotransformations (1–3). Their ability to degrade the groundwater pollutant trichloroethylene and other chlorinated compounds has also been examine in detail (reviewed in Refs. 4 and 5). There is considerable interest in methanotrophs because they appear to be ubiquitous in nature (6–8) and are likely to be a major sink for methane in the natural environment (reviewed in Refs. 9 through 12).

In a pioneering study, Whittenbury and colleagues isolated and characterized a large number of different methanotrophs from soils, sediments, and aquatic environments (6). More recently, Bowman and colleagues (7) isolated more than 100 strains of methanotrophs from many different environments and carried out a detailed chemotaxonomic study on these isolates. This study reinforced the view that these bacteria are widespread in nature and are important in the global cycling of methane in the natural environment.

Methanotrophs are all Gram-negative bacteria and can be classified into two groups. Type I methanotrophs of the genera *Methylomonas*, *Methylobacter*, *Methylomicrobium*, *Methylococcus*, *Methylosphaera*, and *Methylocaldum* utilize the ribulose monophosphate (RuMP) pathway for the assimilation of formaldehyde into cell carbon, possess bundles of intracytoplasmic membranes (Fig. 1), and are members of the γ-subdivision of the purple bacteria (class Proteobacteria). Type II methanotrophs, such as *Methylosinus*, *Methylocystis*, and *Methylosporovibrio*, utilize the serine pathway for formaldehyde fixation, possess intracytoplasmic membranes arranged around the periphery of the cell (Fig. 1), and fall within the α-subdivision of the

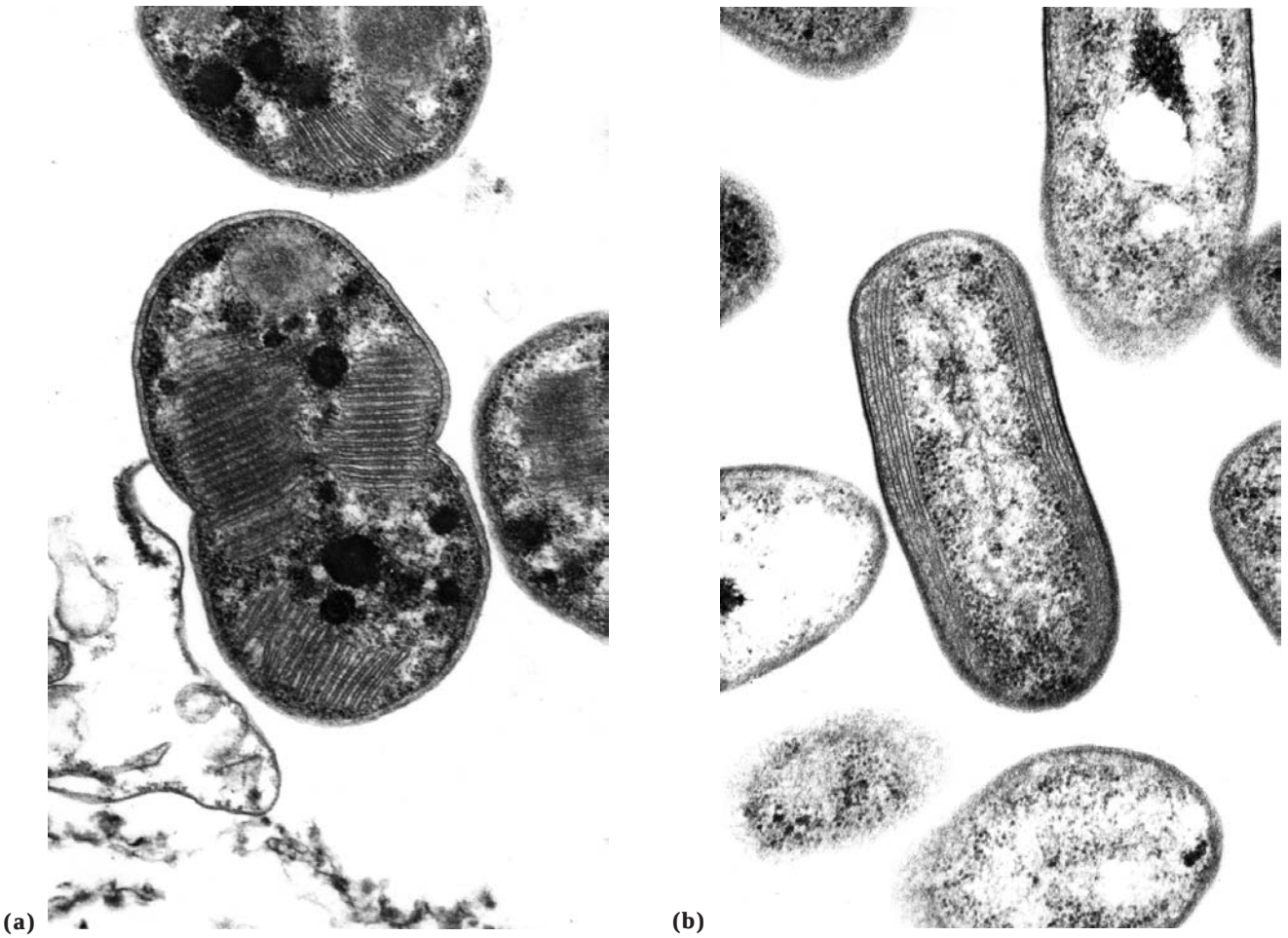


Figure 1. Electron micrographs. (**a**) Section of *Methylomonas methanica* (NCIMB 11130), showing a Type I membrane system. (**b**) Section of *Methylocystis parvus* (NCIMB 11129), showing a Type II membrane system. Magnification, all ×68,800.

Proteobacteria. Some methanotrophs, such as *Methylococcus capsulatus* (Bath), appear to have some properties of both Type I and Type II methanotrophs and are sometimes termed Type X methanotrophs. The taxonomy of methanotrophs has been reviewed in some detail (7,13). Classification schemes, based on pheno- and chemotaxonomic studies have been strengthened as a result of the nucleotide sequencing of both 5S and 16S ribosomal RNA (rRNA) from a large number of methanotrophs and methylotrophs. Type II methanotrophs such as *Methylocystis parvus*, *Methylosinus trichosporium*, and *Methylosporovibrio methanica* appear to cluster in the α-2 subdivision of the class Proteobacteria and form a separate cluster from other serine pathway methylotrophs. Type I methanotrophs such as *Methylomonas methanica*, *Methylomicrobium album*, *Methylomonas rubra*, and *Methylomonas luteus*, together with the Type X methanotroph *Methylococcus capsulatus*, are found in the γ-subdivision of the purple bacteria (reviewed in Refs. 8 and 14). The key features of representative genera of methanotrophs are summarized in Table 1. Properties such as mol% G + C content of DNA, membrane fatty acid composition, nitrogen fixation, and morphological features can be used to discriminate between Type I and Type II methanotrophs.

PHYSIOLOGY AND BIOCHEMISTRY OF METHANOTROPHS

Methanotrophs grow by oxidizing methane using the pathway shown in Figure 2. The key feature of all methanotrophs is their ability to oxidize methane to methanol (see later). Methanol, the oxidation product of methane, is further metabolized to formaldehyde by the pyrolloquinoline quinone (PQQ)-linked enzyme methanol dehydrogenase (MDH). This is an enzyme that is found in all Gram-negative methylotrophic bacteria and has been extensively studied at the biochemical and molecular level (15–17). There are a large number of genes involved in methanol oxidation, and there are complex regulatory networks controlling their expression. The majority of work in this area has been carried out using methanol utilizers rather than methanotrophs (18). The oxidation of methanol yields formaldehyde, a central intermediate in the methane

Table 1. Characterization of the Most Well Studied Methanotrophs

	Methylomonas	Methylobacter	Methylomicrobium	Methylococcus	Methylosinus	Methylocystis
Phylogeny	Gamma Proteobacteria	Gamma Proteobacteria	Gamma Proteobacteria	Gamma Proteobacteria	Alpha Proteobacteria	Alpha Proteobacteria
Membrane type	Type I	Type I	Type I	Type I	Type II	Type II
Major PLFAs	16:1	16:1	16:1	16:1	18:1	18:1
Cell morphology	Rods	Cocci, ellipsoidal, or fat rods	Rods	Cocci or ellipsoidal	Vibriod or pyriform	Cocci, curved rods, or ellipsoidal
Formaldehyde assimilation pathway	RuMP	RuMP	RuMP	RuMP/serine	Serine	Serine
Enzyme type	pMMO	pMMO	pMMO	pMMO/sMMO	pMMO/sMMO	pMMO/sMMO
Mol% G + C content of DNA	50–54	50–54	50–54	62.5	62–63	62–63
Cyst formation	+	+	−	+	−	+
Exospore formation	−	−	−	−	+	−
N-fixation	−	−	−	+	+	+
Carotenoids	+	−	−	−	−	−
Marine representatives	+	+	−	−	−	−
Representative species	*M. aurantiaea, M. fordinarum, M. methanica*	*M. albus, M. marinus, M. whittenburyi*	*M. agile, M. album, M. pelagicum*	*M. capsulatus, M. thermophilus*	*M. trichosporium, M. sporium*	*M. parvus, M. echinoides, M. pyriformis, M. minimus*

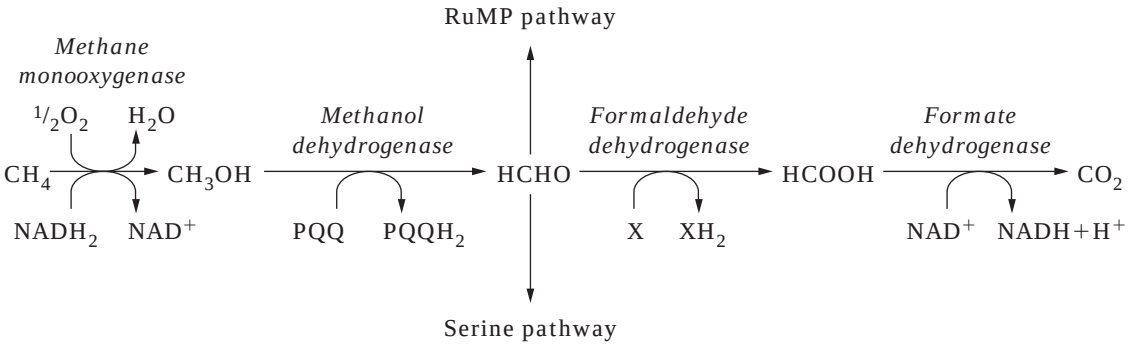


Figure 2. Pathway for the oxidation of methane and the assimilation of formaldehyde.

oxidation pathway. Approximately half of the formaldehyde produced is further oxidized by the enzymes formaldehyde dehydrogenase and formate dehydrogenase to yield carbon dioxide. This results in the generation of reducing power for biosynthesis and the initial oxidation of methane. The CO_2 is not fixed into cell carbon in significant amounts but is lost to the atmosphere. The remainder of the formaldehyde is assimilated into cell carbon by one of two pathways. In Type I methanotrophs, formaldehyde is condensed with ribulose phosphate into hexulose-6-phosphate in the RuMP pathway. The key enzymes of this pathway are hexulose-6-phosphate synthase and hexulose-phosphate isomerase. Subsequent rearrangement reactions result in the net incorporation of glyceraldehyde-3-phosphate into cellular material (Fig. 3). Type II methanotrophs utilize the serine pathway for the incorporation of formaldehyde into the cell. Key enzymes of this pathway include serine hydroxymethyl transferase, hydroxypyruvate reductase, malate thiokinase, and malyl CoA lyase (Fig. 4). The biochemistry of these methylotrophic pathways has been extensively reviewed by Anthony (15,16).

Methane oxidation is carried out by the enzyme methane monooxygenase (MMO). A membrane-bound, particulate methane monooxygenase (pMMO) appears to be present in all methanotrophs grown in the presence of relatively high concentrations of copper ions. Studies on this enzyme have been hampered by its lability in cell-free extracts and difficulties in purification from the membranes of methanotrophs. Biochemical and biophysical studies have shown that pMMO activity is directly related to the copper/total membrane protein ratio in membrane fractions from *Methylococcus capsulatus* (Bath) grown at different copper concentrations. pMMO appears to consist of at least three polypeptides of approximately 46, 23, and 20 kDa, and EPR studies suggest the presence of a number of copper clusters. Recently, the pMMO enzyme complex has been isolated in active form from *M. capsulatus* (Bath) using a variety of detergent treatments (19). An active preparation of pMMO from *Methylococcus* has been reported to contain three subunits (47,27, and 25 kDa) and iron and copper atoms (probably involved in catalysis at the active site) and uses duroquinol as the in vitro electron donor (20). Further characterization of pMMO at the biochemical level is ongoing in a number of research groups.

In some methanotrophs, a second form of MMO, a cytoplasmic, soluble form (sMMO), is synthesized when the growth conditions are such that the culture copper-to-biomass ratio is low (less than 1 μM/g dry weight). This sMMO is structurally and catalytically distinct from the pMMO and has a broad substrate specificity, oxidizing a

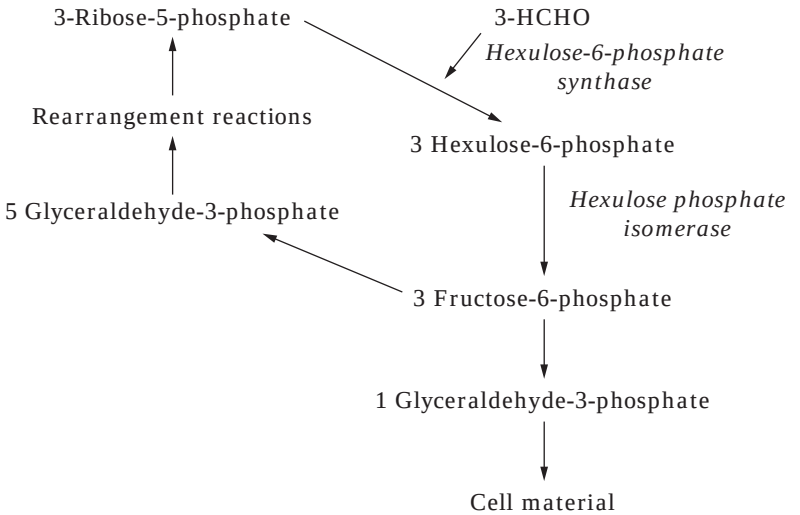


Figure 3. RuMP pathway for formaldehyde fixation.

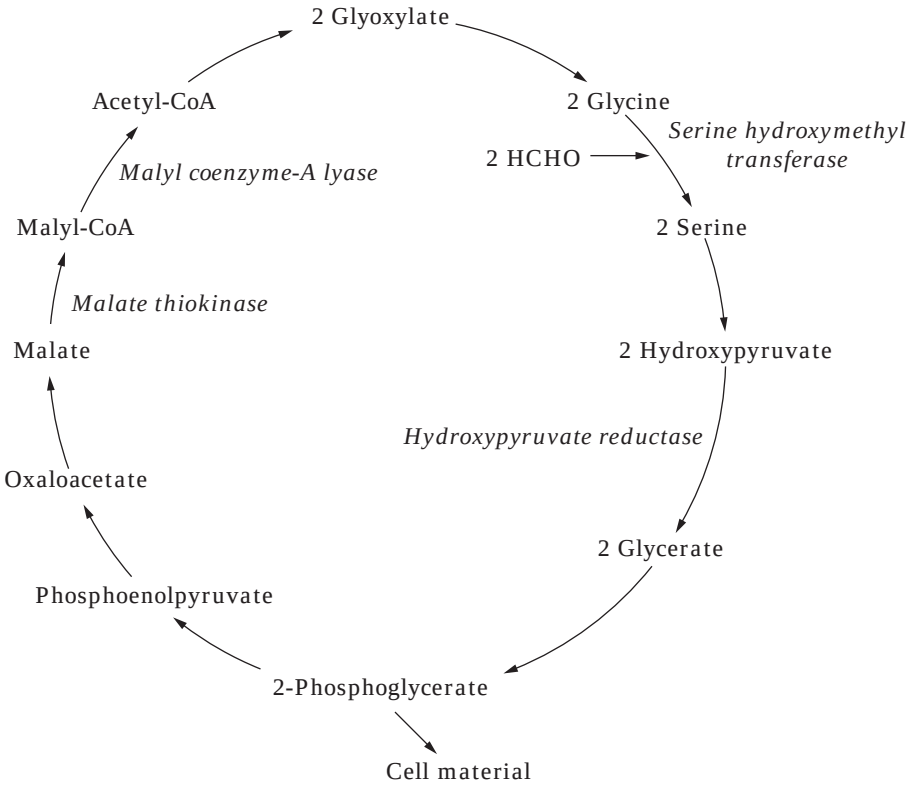


Figure 4. Serine pathway for formaldehyde fixation.

wide range of aliphatic and aromatic compounds. Probably the most extensively characterized sMMOs are those isolated from the Type II methanotroph *Methylosinus trichosporium* OB3b (reviewed in Ref. 21) and the Type X methanotroph *M. capsulatus* (Bath) (reviewed in Ref. 22). sMMO has also been characterized from the Type II strains *Methylosinus sporium* 5 and *Methylocystis* strain M. sMMO has also been reported in the Type I methanotroph *Methylomonas methanica* 68-1.

The sMMO complex of *M. capsulatus* (Bath) and *M. trichosporium* OB3b have very similar properties and both consist of three components: A, B, and C. Protein A is made up of two copies of each of three subunits, α, β, and γ, with molecular masses of 60, 45, and 20 kDa and is the hydroxylase component of the enzyme complex. Protein A contains a binuclear iron-oxo center believed to be the reactive center for catalysis (21–23). Protein B, 16 kDa, is a single polypeptide containing no metal ions or cofactors and functions as a regulatory or coupling protein (21). Protein C is a 39-kDa NADH reductase containing 1 mol each of FAD and a 2Fe2S cluster that accepts electrons from NADH and transfers them to the diiron site of the hydroxylase component of the Protein A component of sMMO. The mechanism of action of sMMO has recently been reviewed (21–23). The X-ray crystal structure of the hydroxylase component of sMMO from *M. capsulatus* (Bath) has been reported, which has further stimulated research on the mechanism of oxidation of methane by this unique enzyme (23).

GENETICS OF METHANOTROPHS

A significant problem in the development of genetic systems for methanotrophs has been the lack of suitable ways of introducing and maintaining heterologous vector DNA in these organisms. Coupled with their relatively slow growth on plates, this has meant that fairly slow progress has been made in this area. Bacteriophages specific for methanotrophs have been isolated from a wide range of environments, including groundwater, fermentors, soils, rumen of cattle, fish, and oil and gas installation waters. They have been characterized in terms of their plaque morphology, ultrastructure, antigenic properties, nucleic acid content, restriction enzyme analysis, and lytic spectrum. *Methylosinus*- and *Methylocystis*-specific phages are particularly prevalent in nature, but few attempts to use them for transduction experiments with methanotrophs have been reported (reviewed in Ref. 24).

A number of large plasmids, ranging from around 50 to 190 kbp have been detected in a number of methanotrophs including *Methylomonas*, *Methylocystis*, *Methylosinus*, and *Methylobacter*. Of 10 representatives examined, only *M. capsulatus* (Bath) did not contain any detectable plasmid DNA (25). Some cross-hybridization with plasmid DNA from different *Methylosinus* species was noted, and a rudimentary restriction endonuclease map of the 55-kbp plasmid from *M. albus* was constructed. However, difficulties in the large-scale isolation of plasmid DNA from these organisms have hampered their development as potential

cloning vectors. There is currently no evidence for the presence of antibiotic or metal resistance genes, nitrogen fixation genes, or methane oxidation genes on these plasmids, but because these plasmids represent "significant potential coding" regions, their presence should not be disregarded in any molecular genetics work with these organisms.

There are no good conventional transformation systems for methanotrophs, and several workers have tried electroporation of plasmids into methanotrophs with little success. Transfer of broad-host-range plasmids into methanotrophs has been the most successful route to gene transfer in these organisms. Lidstrom and colleagues have transferred the Inc P1 broad host range cosmid cloning vector into *Methylomonas*, *Methylocystis*, and *Methylosinus* species at frequencies ranging from 10^{-2} to 10^{-8}. The mobilizing plasmid used in these filter matings was pRK2013 (26). Similar results using similar IncP and IncQ plasmids have been achieved in the authors' laboratory with *M. trichosporium* OB3b and *M. capsulatus* (Bath). Kanamycin and streptomycin appear to be good genetic markers for methanotrophs because the rate of spontaneous mutation for resistance to these antibiotics is very low (less than 10^{-8}). Plasmids carrying kanamycin and streptomycin resistance have been successfully transferred to *M. capsulatus* (Bath) at frequencies of 10^{-3} to 10^{-9} per recipient.

The major obstacles to the development of mutagenesis procedures for the obligate methanotrophs are probably their relatively slow growth on agar plates and the fact that they rely on methane as their sole source of carbon and energy. A consequence of this is that selection of methane oxidation mutants may be difficult. This can usually be overcome because most methanotrophs can be adapted to grow on methanol, thus allowing the selection of an MMO^- phenotype. Conventional mutagenesis procedures using chemical mutagens have not been particularly successful. However, mutants lacking the ability to grow on methane have been isolated using the "suicide-substrate" dichloromethane (DCM). This is cooxidized by MMO to the potentially toxic product carbon monoxide, and therefore methanotrophs growing on methanol agar plates that are DCM resistant are potential methane monooxygenase mutants. Screening of methanotrophs for the presence or absence of a functional sMMO has also been made easier due to the development of a quick, reliable plate assay. This is based on the work of Hanson and colleagues (27) and relies on the conversion of naphthalene to naphthols by soluble MMO but not particulate MMO. The naphthols thus formed can be reacted with tetrazotized *O*-dianisidine to form purple diazo dyes with intense color. This method now allows the assay of soluble MMO activity on plates. The procedure does not kill potential mutants, which can subsequently be recovered and streaked onto fresh agar plates under the appropriate growth conditions (28).

A similar naphthalene assay, coupled with DCM selection, has recently been used to isolate and characterize MMO mutants of *M. trichosporium* OB3b (29). *M. trichosporium* OB3b mutants defective in only the particulate form of MMO were isolated by *(1)* suppressing sMMO expression by the presence of high concentrations of copper in selection plates, *(2)* adding excess methanol with DCM to inhibit possible sMMO-mediated activation of DCM to toxic products in any colonies where sMMO might be expressed, *(3)* adding low concentrations of yeast extract to enhance growth and plating efficiency on methanol, and *(4)* carrying out mutagenesis selection over a long period (5 weeks). Although DCM resistance was unstable in many mutants (a phenomenon observed in other laboratories, including the authors'), mutants that were deficient in pMMO and expressed sMMO constitutively in the presence of elevated copper concentrations (up to 12 μM, which inhibits expression of sMMO in the wild-type organism) were isolated. These mutants had lost the ability to express pMMO activity and concomitantly failed to form the intracytoplasmic membranes characteristic of wild-type *M. trichosporium* OB3b expressing pMMO. Subsequent detailed analysis of these sMMO-constitutive *M. trichosporium* OB3b mutants suggested that either a defect in copper metabolism affecting copper assimilation causes lack of pMMO expression, or that a more general copper-dependent regulatory system that may control pMMO induction, sMMO repression, copper uptake, and production of the intracytoplasmic membranes associated with pMMO has been inactivated (30). Such mutants would clearly be useful for in situ bioremediation processes involving sMMO where the external concentration of copper ions cannot be controlled.

MOLECULAR BIOLOGY OF THE METHANE OXIDATION PATHWAY

Research, primarily in the laboratories of Lidstrom, Hanson, Goodwin, and Harms, has resulted in the identification and cloning of a number of genes involved in methanol oxidation from the Gram-negative methylotrophs *Methylobacterium extorquens* and *Paracoccus denitrificans* (reviewed in Refs. 17, 31, and 32). One of these genes, *mxaF*, from *Methylobacterium*, encoding the large subunit structural gene of methanol dehydrogenase, has been used as a heterologous hybridization probe to identify and clone the corresponding *mxaF* genes from *M. capsulatus* (Bath) and *M. albus* BG8. Methanol oxidation structural genes have also been isolated from the marine methanotroph *Methylomonas* strain A4. Molecular characterization of these genes suggests that they are similar to those that have been extensively characterized in Gram-negative methylotrophs (17), although their regulation may be somewhat different (33).

The genes encoding the sMMO enzyme complex have been cloned from *M. capsulatus* (Bath) and are clustered on the chromosome of this methanotroph. DNA sequencing of this gene cluster has revealed that the genes encoding the α-, β-, and γ-subunits of Protein A (*mmoX*, *Y*, and *Z*), Protein B (*mmoB*), and Protein C (*mmoC*) are all linked on the chromosome. The sMMO gene cluster of *M. capsulatus* (Bath) has been used to probe for sMMO genes in a number of representative strains of methanotrophs. sMMO homologues were detected only in *Methylococcus* and *Methylosinus* strains. sMMO appeared to be absent from the Type II methanotroph *Methylocystis parvus* OBBP or Type I representatives of *Methylomicrobium album*, *Methylomonas*

methanica, *Methylomonas agile*, and *Methylobacter capsulatus*.

Subsequently, sMMO gene probes from *Methylococcus* were used to clone the sMMO gene cluster from *M. trichosporium* OB3b (reviewed in Ref. 34) and *Methylocystis* sp. strain M (35). These gene clusters have also been DNA sequenced, and the genes are arranged in the same order as in *Methylococcus* (Fig. 5). The function of *orfY*, which lies between *mmoZ* and *mmoC*, is unknown at present. Derived polypeptide sequences of sMMO components from all three methanotrophs showed a high degree of identity, highlighting the conserved nature of this enzyme complex (35). Amino acid sequences within the α-subunit of Protein A align well with the four-helix iron coordination bundle of the R2 protein of ribonucleotide reductase and is characteristic of a family of proteins that contain a catalytic carboxylate-bridged diiron center. The diiron center in sMMO is now well defined by X-ray crystal structure analysis (23). The N-terminal region of Protein C also exhibited extensive homology with reductase components of other oxygenases that contain a ferredoxin-like 2Fe-2S center. Protein B, the coupling protein of sMMO, has some sequence identities with other oxygenase-coupling proteins (36), but all other polypeptides of sMMO including the derived sequence of *orfY* appear to be novel.

Expression of functionally active Proteins B and C has been achieved using a T7 polymerase expression system in *Escherichia coli*. Attempts to express the hydroxylase components in *E. coli* in the authors' laboratory have proved unsuccessful, but expression of sMMO genes has recently been achieved using other heterologous hosts including methanotrophs that contain only pMMO (37), which will facilitate further studies on the structure, function, and catalytic mechanism of methane monooxygenase by site-directed mutagenesis.

Use of the cloned sMMO genes has also been made in the construction of sMMO mutants of *M. trichosporium* OB3b. A reverse-genetics approach has been taken using marker exchange mutagenesis, a method first used successfully by Toukdarian and Lidstrom (38) for the construction of nitrogen-fixation mutants of *Methylosinus* sp. strain 6. A DNA fragment containing *mmoX* of OB3b was cloned into pBR325. An *Xho*I fragment internal to the *mmoX* gene was replaced with a kanamycin resistance cassette, and the *mob* fragment was also inserted into pBR325 to facilitate transfer by conjugation of the resulting recombinant plasmid pHM32 from *E. coli* S17-1 to OB3b. Homologous recombination of the mutated *mmoX* gene with the wild-type *mmoX* gene in the chromosome and subsequent loss of the unstable pBR325-based plasmid allowed the isolation of kanamycin-resistant OB3b colonies. These were DCM resistant on methanol-agar plates and grew only in the presence of Cu^{2+} on methane, that is, they had the expected $pMMO^{+}$, $sMMO^{-}$ phenotype (39). These mutants have now been complemented by transfering the sMMO gene cluster from wild-type *M. trichosporium* OB3b back into these mutants by conjugation.

M. trichosporium OB3b is a better organism than *M. capsulatus* (Bath) in which to study the regulation of expression of sMMO because OB3b will also grow well on methanol agar plates. The physiological growth conditions under which OB3b switches between using pMMO and sMMO, in response to the Cu^{2+}-to-biomass ratio are also now well defined and have been the subject of several studies (40,41). Differential expression of sMMO and pMMO is regulated by the amount of copper ions available to the cells; sMMO is expressed at low copper-to-biomass ratios, whereas pMMO is expressed at high copper-to-biomass ratios. The transcriptional regulation of the sMMO gene cluster in the methanotrophic bacterium *M. capsulatus* (Bath) has been investigated (42). Using Northern blot analysis and primer extension experiments, it was shown that the six open reading frames of the sMMO gene cluster are organized as an operon, and the transcripts produced upon expression of this operon have been identified. The synthesis of these transcripts was under the control of a

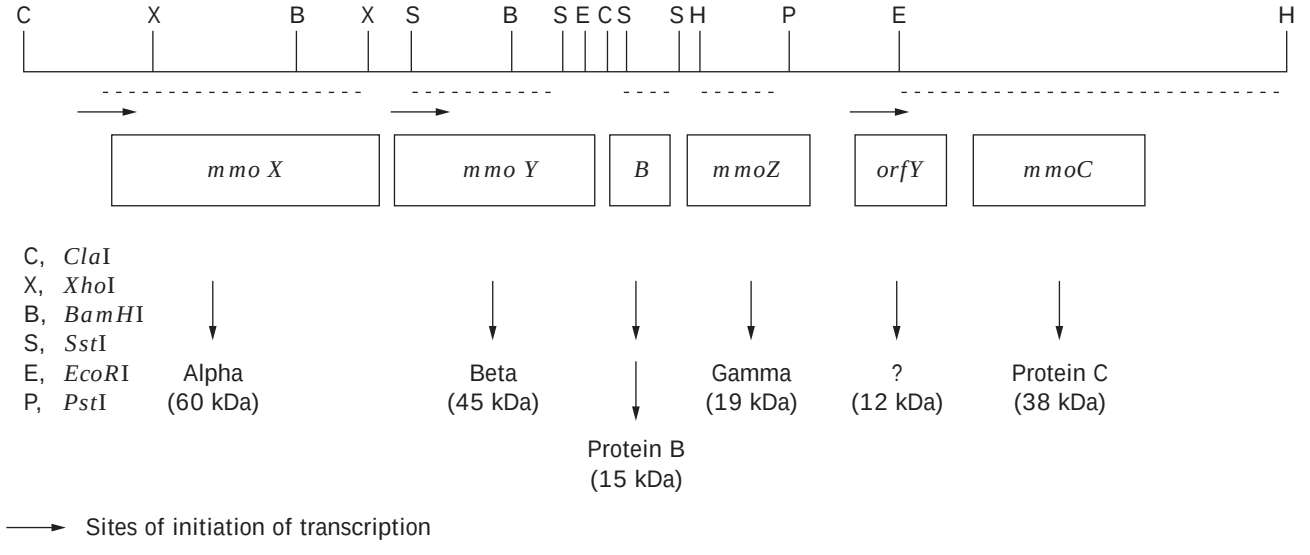


Figure 5. sMMO gene cluster of *Methylosinus trichosporium* OB3b.

single copper-regulated promoter, which has not been precisely defined. In both *M. capsulatus* and *M. trichosporium*, transcription of the sMMO gene cluster is negatively regulated by copper ions. Data suggest that transcription of the *M. trichosporium* OB3b sMMO gene cluster is directed from a σ^{54}-like and a σ^{70}-like promoter. The pMMO (*pmo*) genes of *M. capsulatus* (Bath) are transcribed into a polycistronic mRNA of 3.3 kb. The synthesis of this mRNA was activated by copper ions. Activation of *pmo* transcription by copper ions was concomitant with repression of sMMO gene transcription in both methanotrophs. This suggests that a common regulatory pathway may be involved in the transcriptional switch between sMMO and pMMO gene expression (42,43).

MOLECULAR ECOLOGY OF METHANOTROPHS

The difficulties in using traditional culture-based techniques to study the ecology of methanotrophs have been discussed previously (44). In particular the slow growth of methanotrophs, together with scavenging of nonmethanotrophs on agar plates, has hampered studies.

Phospholipid analysis of natural environment samples or enrichment cultures is one approach to assessing the types of methanotrophs present, because they contain more unusual fatty acids; the technique has been used successfully for the wetland environment (45). The detection and identification of methanotrophs in environmental samples can also be attempted using antibodies either to whole cells of culturable methanotrophs or to key enzymes such as methane monooxygenase or methanol dehydrogenase (reviewed in Ref. 44). However, these methods do not guarantee that all methanotrophs may be detected, and extensive control experiments with large numbers of extant methanotrophs are necessary to validate antibody probes. Estimates of the biomass and activities of methanotrophs can be made by measuring the release of $^{14}CO_2$ and incorporation of $^{14}CH_4$ added to environmental samples. However, this technique needs to be coupled with further experiments designed to analyze bacterial community structure.

Application of molecular biology techniques to methanotrophs has been aided considerably by the sequencing of a number of 16S rRNA genes of methanotrophs and methylotrophs (7,8) and the cloning of several methanotroph-specific genes (34,46). Hanson and colleagues have used 16S rRNA sequence data to examine phylogenetic relationships within genera of methanotrophs and methylotrophs. Type I methanotrophs such as *Methylomonas*, *Methylomicrobium*, and *Methylobacter*, together with the Type X *Methylococcus*, cluster in the γ-subdivision of the class Proteobacteria, whereas Type II methanotrophs *Methylosinus* and *Methylocystis* are found in the α-2 subdivision of the Proteobacteria (8). Hanson and colleagues have used this information to design oligonucleotide probes having a broad specificity for methanotrophs and methylotrophs in the natural environment (47). When coupled to suitable fluors, probes 9α and 10γ were successfully used to differentiate between mixed cultures of Type I and Type II methanotrophs by in situ hybridization (14).

16S rRNA technology has also been used to study the ecology of methanotrophs in blanket bog peat from the north of England. DNA was extracted from several 1-cm sections of a 30-cm-depth peat core covering various depths throughout the core, including the most active methane-oxidizing layer. Total 16s rRNA libraries from each section were screened with 16S rRNA probes specific for the genera *Methylococcus* (Mc 1005), *Methylomonas* (Mm 1007), *Methylobacter* (Mb 1007), and *Methylosinus* (Ms 1020). Results suggested that *Methylosinus* and *Methylococcus* were probably the dominant methanotrophs present in the peat environment. Two clones that hybridized to the *Methylosinus* probe were sequenced. Their 16S rDNA sequences, when analyzed, suggested that they branched closest to the *Methylosinus / Methylocystis* cluster of methanotrophs but were distinct from these sequences (48).

The 16S rRNA data coupled with polymerase chain reaction (PCR) technology have also been used successfully in analyzing methanotrophs in the marine environment. Seawater samples were enriched for methanotrophs by addition of essential nutrients and methane. The change in composition of the bacterial population was then monitored by analysis of 16S rDNA libraries constructed from DNA taken before and after enrichment, using universal eubacterial 16S rRNA primers and PCR. The dominant 16S rRNA sequence that was present in samples after enrichment on methane was recovered from 16S rRNA gene libraries, sequenced, and found to show a close phylogenetic relationship to *Methylomonas*, indicating that novel methanotrophs related to extant *Methylomonas* spp. were present in enrichment cultures (49). Based on the 16S rDNA sequence of this putative marine methanotroph, two oligonucleotides (Mm 650 and Mm 850, Table 1) were synthesized with rhodamine groups attached to their 5′ terminus. These fluorescently labeled probes have been used to determine the response of this marine methanotroph to changes in cultural conditions, thereby allowing optimization of enrichment and isolation methods for this organism (50).

FUNCTIONAL GENE PROBES FOR METHANOTROPHS

Genes unique to methanotrophs include those encoding sMMO and pMMO polypeptides. Another potentially useful marker is *mxaF*, encoding the large subunit of methanol dehydrogenase (the second enzyme of the methane oxidation pathway), which is present in virtually all Gram-negative methylotrophs (51). *mxaF* appears to be highly conserved in all Gram-negative methylotrophs (e.g., methane and methanol utilizers) and is therefore a good indicator of the presence of these organisms in the natural environment. PCR primers designed to specifically amplify a 550-bp fragment of *mxaF* have been successfully used to extend the database of *mxaF* sequences and to identify methylotroph sequences in peat core samples (49,52,53).

The high degree of identity between sMMO genes has enabled the design of PCR primers that specifically amplify each of the five sMMO structural genes (52) such that amplification of each of these genes is now possible from cultured methanotrophs containing sMMO. In addition,

when total DNA from a variety of different freshwater, marine, soil, and peat samples was used as a template with these sMMO-gene-specific primers, a PCR product of the correct (predicted) size was obtained. Subsequent DNA sequencing of a number of these PCR products from peat samples has proved that methanotroph-specific sMMO-encoding DNA had been amplified from environmental DNA samples, and that these genes are very similar but not identical to known sMMO gene sequences. This suggested that there was considerable diversity of methanotrophs in these peat samples, and also that sMMO-containing methanotrophs were present in significant numbers (52). This is further supported by the observation that sMMO-containing methanotrophs have been enriched for and isolated from the same peat core samples, and that 16S rRNA/PCR experiments described earlier also revealed the presence (dominance) of *Methylosinus*- and *Methylococcus*-like organisms.

The PCR technology developed with sMMO genes is useful for studying environments such as peat or contaminated aquifers where methanotrophs containing sMMO might dominate in a copper-depleted environment. However, a better functional gene probe for these organisms would be one based on the pMMO present in all extant methanotrophs. Sequence data on *pmo* and *amo* (the ammonia monooxygenase gene, which codes for a related enzyme found in nitrifying bacteria) has allowed the design of degenerate PCR primers that will specifically amplify a 525-bp internal DNA fragment of *pmoA* or *amoA* from a variety of methanotrophs and nitrifiers. Analysis of the predicted amino acid sequences of these genes from representatives of each of the phylogenetic groups of methanotrophs (α- and γ-Proteobacteria) and ammonia oxidizing nitrifiers (β- and γ-Proteobacteria), revealed strong conservation of both primary and secondary structure. These new sequence data suggest that pMMO and ammonia monooxygenase may be evolutionarily related enzymes, despite their different physiological roles in methanotrophs and ammonia-oxidizing bacteria (54).

These observations are particularly important in the study of the ecological role of these organisms because ammonia oxidizers can also oxidize methane (with MMO also cooxidizing ammonia). The degenerate oligo PCR primers just described have now been used successfully to specifically amplify *pmo/amo* sequences from DNA isolated from a wide variety of environments including seawater, freshwater, agricultural and forest soils, and peat (J.C. Murrell and I.R. McDonald, unpublished data, 1997). Sequencing of these PCR products and examination of derived amino acid sequences of PmoA and AmoA now allows identification of these sequences as those originating from an α- β-, or γ-representative of the methane-oxidizing Proteobacteria (55), and an extensive database of sequences from extant methanotrophs and nitrifiers is also being established, which will aid further molecular ecological studies.

BIOTECHNOLOGY

The unique ability of the MMO enzyme in methanotrophic bacteria to catalyze some reactions of environmental and possibly commercial importance has generated interest in these organisms. The discovery that methanotrophs can degrade halogenated hydrocarbons that are ubiquitous in nature and also toxic pollutants has resulted in research that may lead to greater commercial exploitation (4,5,56).

The most ubiquitous of the groundwater pollutants are trichloroethylene (TCE), 1,2-dichloroethane, and chloroform. Bacteria that utilize TCE or chloroform as a sole carbon and energy source have not been discovered. However, several bacteria that contain nonspecific oxygenases, such as MMO, cause the mineralization of these compounds by cometabolism (cooxidation).

The capacity of methanotrophs to convert halogenated compounds was recognized by Colby et al. (57), who showed that methyl bromide could be used as a model substrate for the MMO reaction. Further work was stimulated by the observation of Wilson and Wilson (58), who demonstrated that methane could stimulate biological TCE degradation in soil. Since then, various pure cultures of methanotrophic bacteria have been found to be capable of cometabolic transformation of halogenated hydrocarbons (27,41,56,59–61). Large differences between various cultures were observed, and high TCE degradation rates were found with *M. trichosporium* OB3b cultivated under conditions of copper limitation and incubated with formate instead of methane as a source of reducing equivalents (60). Cells grown under copper limitation express exclusively the soluble MMO (40,62), and the conclusion that it is this form of the enzyme that is responsible for TCE oxidation was confirmed by in vitro incubations with the purified enzyme (63). Much lower rates of TCE oxidation have been reported for methanotrophs that possess only pMMO (64).

The feasibility of in situ bioremediation by stimulation of the growth of indigenous populations of methanotrophs has been investigated by Semprini and McCarty (65). They reported that biodegradation of TCE occured within an aquifer, at the Moffett Naval Air Station, after it had been supplied with methane and oxygen to stimulate indigenous methane-oxidizing bacteria. The initial results from a further study, at the Savannah River site, indicated that a substantial increase in Type II methanotrophic bacterial populations occurred in some areas of the aquifer, and an increase in the frequency of bacteria that contained DNA with sequences complementary to the *mmoB* gene at different sites was observed (66).

The second biotechnological application of the methanotrophs has been as single-cell protein (SCP). Methane, as natural gas, is available inexpensively and in large amounts; however, as a substrate for SCP production, it has many drawbacks. These include explosive hazards and the low aqueous solubility of methane. Therefore methanol has been the preferred substrate because it is easily derived from methane and highly soluble in water. Methanol is also easily removed from the SCP product in downstream processing. Not surprisingly it was the substrate of choice for the only self-contained SCP business in the United Kingdom, the Pruteen process of ICI (67), which utilized the methylotroph *Methylophilus methylotrophus*. Methanol was also used in two smaller-scale processes by Probion (Hoechst-Uhde, Germany) and Norprotein (Norsk-

Hydro, Sweden), both of which utilized *Methylomonas* species (for reviews see Refs. 2 and 3).

Methanotrophs have also been used for the conversion of propylene to propylene oxide, a high-value product. However, even though high rates of propylene oxidation were obtained, these rates fell dramatically after 3–4 min due to loss of biocatalytic activity (68). sMMO and pMMO both have the potential for cooxidation of a wide range of other compounds. sMMO will cooxidize compounds including alkanes, haloalkanes, alkenes, alicyclics, and aromatics (69); recently pMMO has been shown to cooxidize chlorinated ethenes at yields that are greater than those of sMMO (70). These cooxidations may have future biotechnological applications. Finally the recent characterization of thermophilic/thermotolerant methanotrophs also opens up further avenues for the exploitation of methanotrophs in biotechnological processes (71).

ACKNOWLEDGMENTS

Work in J. C. M.'s laboratory is supported by grants from the NERC, BBSRC, AFRC, and EEC. We thank J. McN. Sieburth for electron micrographs.

BIBLIOGRAPHY

1. M.E. Lidstrom and D.I. Stirling, *Annu. Rev. Microbiol*, **44**, 27–58 (1990).
2. D.J. Leak, in J.C. Murrell and H. Dalton eds., *Methane and Methanol Utilizers*, Plenum, New York, 1992, pp. 245–279.
3. P.J. Large and C.W. Bamforth, *Methylotrophy and Biotechnology*, Longman, Harlow, U.K., 1988.
4. R. Oldenhuis and D.B. Janssen, in J.C. Murrell and D.P. Kelly eds., *Microbial Growth on C_1 Compounds*, Intercept Press, Andover, U.K., 1993, pp. 121–133.
5. L. Alvarez Cohen, in J.C. Murrell and D.P. Kelly eds., *Microbial Growth on C_1 Compounds*, Intercept Press, Andover, U.K., 1993, pp. 337–350.
6. R. Whittenbury, K.C. Phillips, and J.F. Wilkinson, *J. Gen. Microbiol.* **61**, 205–218 (1970).
7. J.P. Bowman, L.I. Sly, P.D. Nicholas, and A.C. Hayward, *Int. J. Syst. Bacteriol.* **43**, 735–753 (1993).
8. R.S. Hanson and T.E. Hanson, *Microbiol. Rev.* **60**, 439–471 (1996).
9. W.S. Reeburgh, in J.C. Murrell and D.P. Kelly eds., *Microbial Growth on C_1 Compounds*, Intercept Press, Andover, U.K., 1993, pp. 1–14.
10. G.M. King, in J.C. Murrell and D.P. Kelly eds., *Microbial Growth on C_1 Compounds*. Intercept Press, Andover, U.K., 1993, pp. 303–313.
11. R. Conrad, *Adv. Microb. Ecol.* **14**, 207–250 (1995).
12. R. Conrad, *Microbiol. Rev.* **60**, 609–640 (1996).
13. P.N. Green, in J.C. Murrell and H. Dalton eds., *Methane and Methanol Utilizers*. Plenum, New York, 1992, pp. 23–84.
14. R.S. Hanson, B.J. Bratina, and G.A. Brusseau, in J.C. Murrell and D.P. Kelly eds., *Microbial Growth on C_1 Compounds*. Intercept Press, Andover, U.K., 1993, pp. 285–302.
15. C. Anthony, *The Biochemistry of Methylotrophs*. Academic, London, 1982.
16. C. Anthony, *Adv. Microb. Physiol.* **27**, 113–210 (1986).
17. M.E. Lidstrom, in J.C. Murrell and H. Dalton eds., *Methane and Methanol Utilizers*, Plenum, New York, 1992, pp. 183–206.
18. P.M. Goodwin and C. Anthony, *Microbiology* **141**, 1051–1064 (1995).
19. H.A.T. Nguyen, M. Zhu, S.J. Elliot, K.H. Nakagawa, B. Hadman, A.M. Costello, T.L. Peeples, B. Wilkinsob, H. Morimoto, P.G. Williams, H.G. Floss, M.E. Lidstrom, K.O. Hodgson, and S.I. Chan, in M.E. Lidstrom and F.R. Tabita eds., *Microbial Growth on C_1 Compounds*, Kluwer Academic, Dordrecht, The Netherlands, 1996, pp. 150–158.
20. J.A. Zahn and A.A. DiSpirito, *J. Bacteriol.* **178**, 1018–1029 (1996).
21. J.D. Lipscomb, *Annu. Rev. Microbiol.* **48**, 371–399 (1994).
22. H. Dalton, P.C. Wilkins, and Y. Jiang, in J.C. Murrell and D.P. Kelly eds., *Microbial Growth on C_1 Compounds*, Intercept Press, Andover, U.K., 1993, pp. 65–80.
23. A.C. Rosenzweig, C.A. Frederick, and S.J. Lippard, in M.E. Lidstrom and F.R. Tabita eds., *Microbial Growth on C_1 Compounds*, Kluwer Academic, Dordrecht, The Netherlands, 1996, pp. 141–149.
24. J.C. Murrell, *FEMS Microbiol. Rev.* **88**, 233–248 (1992).
25. M.E. Lidstrom and A.E. Wopat, *Arch. Microbiol.* **140**, 27–33 (1984).
26. M.E. Lidstrom, A.E. Wopat, D.N. Nunn, and A.E. Toukdarian, in G.S. Omenn and A. Hollaender eds., *Genetic Control of Environmental Pollutants*, Plenum, New York, 1984, pp. 273–279.
27. G.A. Brusseau, H.G. Tsien, R.S. Hanson, and L.P. Wackett, *Biodegradation* **1**, 19–29 (1990).
28. D.W. Graham, D.G. Korich, R.P. Le Blanc, N.A. Sinclair, and R.G. Arnold, *Appl. Environ. Microbiol.* **58**, 2231–2236 (1992).
29. P.A. Phelps, S.K. Agarmal, G.E. Speital, and G. Georgiou, *Appl. Environ. Microbiol.* **58**, 3701–3708 (1992).
30. M.W. Fitch, D.W. Graham, R.G. Arnold, S.K. Agarwal, P. Phelps, G.E. Speitel, and G. Georgiou, *Appl. Environ. Microbiol.* **59**, 2771–2776 (1993).
31. T.M. Barta and R.S. Hanson, *Antonie van Leeuwenhoek* **64**, 109–120 (1993).
32. N. Harms, in J.C. Murrell and D.P. Kelly eds., *Microbial Growth on C_1 Compounds*, Intercept Press, Andover, U.K., 1993, pp. 235–244.
33. D. Waechter-Brulla, A.A. DiSpirito, L.V. Chistoserdova, and M.E. Lidstrom, *J. Bacteriol.* **175**, 3767–3775 (1993).
34. J.C. Murrell, *Biodegradation* **5**, 145–159 (1994).
35. I.R. McDonald, H. Uchiyama, S. Kambe, O. Yagi, and J.C. Murrell, *Appl. Environ. Microbiol.* **63**, 1898–1904 (1997).
36. H. Qian, U. Edlund, J. Powlowski, V. Shingler, and I. Sethson, *Biochemistry* **36**, 495–504 (1997).
37. J.S. Lloyd, A. Bhambra, J.C. Murrell, and H. Dalton, *Eur. J. Biochem.* **248**, 72–79 (1997).
38. A.E. Toukdarian and M.E. Lidstrom, *J. Bacteriol.* **157**, 979–983 (1984).
39. J.C. Murrell, in J.C. Murrell and D.P. Kelly eds., *Microbial Growth on C_1 Compounds*. Intercept Press, Andover, U.K., 1993, pp. 109–120.
40. K.J. Burrows, A. Cornish, D. Scott, and I.J. Higgins, *J. Gen. Microbiol.* **130**, 3327–3333 (1984).
41. H-C. Tsien, G.A. Brusseau, R.S. Hanson, and L.P. Wackett, *Appl. Environ. Microbiol.* **55**, 3155–3161 (1989).
42. A.K. Nielsen, K. Gerdes, H. Degn, and J.C. Murrell, *Microbiology* **142**, 1289–1296 (1996).
43. A.K. Nielsen, K. Gerdes, and J.C. Murrell, *Mol. Microbiol.* **25**, 399–409 (1997).

44. J.C. Murrell and A.J. Holmes, in I. Joint ed., *Molecular Ecology of Aquatic Microbes*, Spriger-Verlag, Berlin, 1995, pp. 365–390.

45. I. Sundh, P. Borga, M. Nilsson, and B.H. Svensson, *FEMS Microbiol. Ecol.* **18**, 103–112 (1995).

46. J.C. Murrell and A.J. Holmes, in M.E. Lidstrom and F.R. Tabita eds., *Microbial Growth on C_1 Compounds*, Kluwer Academic, Dordrecth, The Netherlands, 1996, pp. 133–140.

47. G.A. Brusseau, E.S. Bulygina, and R.S. Hanson, *Appl. Environ. Microbiol.* **60**, 626–636 (1994).

48. I.R. McDonald, G.H. Hall, R.W. Pickup, and J.C. Murrell, *FEMS Microbiol. Ecol.* **21**, 197–211 (1996).

49. A.J. Holmes, N.J.P. Owens, and J.C. Murrell, *Microbiology* **141**, 1947–1955 (1995).

50. A.J. Holmes, N.J.P. Owens, and J.C. Murrell, *J. Exp. Mar. Biol. Ecol.* **203**, 27–38 (1996).

51. M.E. Lidstrom, C. Anthony, F. Biville, F. Gasser, P. Goodwin, R.S. Hanson, and N. Harms, *FEMS Microbiol. Lett.* **177**, 103–106 (1994).

52. I.R. McDonald, E.M. Kenna, and J.C. Murrell, *Appl. Environ. Microbiol.* **61**, 116–121 (1995).

53. I.R. McDonald and J.C. Murrell, *Appl. Environ. Microbiol.* **63**, 3218–3224 (1997).

54. A.J. Holmes, A. Costello, M.E. Lidstrom, and J.C. Murrell, *FEMS Microbiol. Lett.* **132**, 203–208 (1995).

55. I.R. McDonald and J.C. Murrell, *FEMS Microbiol. Lett.* **156**, 205–210 (1997).

56. R. Oldenhuis, R.L.J.M. Vink, D.B. Janssen, and B. Witholt, *Appl. Environ. Microbiol.* **55**, 2819–2826 (1989).

57. J. Colby, H. Dalton, and R. Whittenbury, *Biochem. J.* **151**, 459–462 (1975).

58. J.T. Wilson and B.H. Wilson, *Appl. Environ. Microbiol.* **49**, 242–243 (1985).

59. C.D. Little, A.V. Palumbo, S.E. Herbes, M.E. Lidstrom, R.L. Tyndall, and P.J. Gilmer, *Appl. Environ. Microbiol.* **54**, 951–956 (1988).

60. H. Uchiyama, T. Nakajima, O. Yagi, and T. Tabuchi, *Agric. Biol. Chem.* **53**, 2903–2907 (1989).

61. S.M. Henry and D. Grbic-Galic, *Appl. Environ. Microbiol.* **57**, 236–244 (1991).

62. S.H. Stanley, S.D. Prior, D.J. Leak, and H. Dalton, *Biotechnol. Lett.* **5**, 487–492 (1983).

63. B.G. Fox; J.G. Borneman, L.P. Wackett, and J.D. Lipscomb, *Biochemistry* **29**, 6419–6427 (1990).

64. A.A. DiSpirito, J. Gulledge, J.C. Murrell, A.K. Shiemke, M.E. Lidstrom, and C.L. Krema, *Biodegradation* **2**, 151–164 (1992).

65. L. Semprini and P.L. McCarty, in P.V. Roberts, L. Semprini, G.D. Hopkins, D. Grbic-Galic, P.L. McCarty, and M. Reinhard eds. *In Situ Aquifer Restoration of Chlorinated Aliphatics by Methanotrophic Bacteria*, Center for Environmental Research Information, Cincinnati, 1989, pp. 239–250.

66. J.P. Bowman, L. Jimenez, I. Rosario, T.C. Hazen, and G.S. Sayler, *Appl. Environ. Microbiol.* **59**, 2380–2387 (1993).

67. K.A. Powell and B.L.F. Rodgers, in C.T. Hou ed., *Methylotrophs: Microbiology, Biochemistry and Genetics*, CRC Press, Boca Raton, Fl., 1984, pp. 119–144.

68. S.H. Stanley and H. Dalton, *Biocatalysis* **6**, 163–175 (1992).

69. H. Dalton, *Adv. Appl. Microbiol.* **26**, 71–87 (1980).

70. J.E. Anderson and P.L. McCarty, *Appl. Environ. Microbiol.* **63**, 687–693 (1997).

71. L. Bodrossy, J.C. Murrell, H. Dalton, M. Kalman, L.G. Puskas, and K. Kovacs, *Appl. Environ. Microbiol.* **61**, 3549–3555 (1995).

See also BIOENERGETICS OF MICROBIAL GROWTH.

METHYLOTROPHS, INDUSTRIAL APPLICATIONS

DAVID LEAK
Imperial College of Science, Technology, and Medicine
London, United Kingdom

KEY WORDS

Amino acids
Bioremediation
Biotransformation
Methane
Methanol
Polyhydroxyalkanoates
Polysaccharides
Ribulose monophosphate pathway
Serine pathway
Single-cell protein
Xylulose monophosphate pathway

OUTLINE

INTRODUCTION

Methane and methanol are cheap, readily available carbon sources for microbial growth, and their metabolism involves specialized routes of assimilation and dissimilation. Consequently, considerable work has gone into exploring potential applications of organisms capable of growth and product formation from these substrates, particularly on methanol. Because the attraction of methylotrophs derives primarily from substrate costs, most of the work has been directed toward the higher volume–lower price end of the fermentation market, in particular single-cell protein

(SCP). However, SCP from methanol proved to be uneconomical, although an unexpected benefit arose from the early SCP program, namely the application of methylotrophic yeasts as hosts for expression of recombinant proteins.

The full potential of methanol as a fermentation substrate has yet to be realized. Future applications could arise from the use of methanol to supplant carbohydrates in existing fermentation processes for metabolite (e.g., amino acid and vitamin) production. However, the benefit of lower substrate costs needs to be balanced against the (generally) increased oxygen requirements of methanol- compared with carbohydrate-based processes.

MICROBIOLOGY AND ECOLOGY

Methylotrophs are organisms that grow on carbon substrates more reduced than carbon dioxide but with no carbon–carbon bonds (reduced C_1 substrates). Although some aerobic methylotrophs have been shown to oxidize reduced C_1 substrates to carbon dioxide and fix carbon by typical autotrophic routes (autotrophic methylotrophs), most aerobic methylotrophs assimilate carbon at the oxidation level of formaldehyde (true methylotrophs). The primary metabolism of these organisms is thus fundamentally distinct from that of autotrophs and heterotrophs, and some "obligate" methylotrophs are unable to use heterotrophic substrates containing carbon–carbon bonds as the sole carbon source.

Methane, the most reduced C_1 compound, is a key link between the processes of anaerobic degradation and the aerobic food chain. The activity of anaerobic and aerobic methanotrophic bacteria (i.e., methylotrophs that can use methane as a sole source of carbon and energy) significantly reduces the potential release of methane to the atmosphere arising from the activity of methanogens. Many of the methanogens are also methylotrophic, being able to use substrates such as methanol and formate to drive methanogenesis (methylotrophic acetogens are also known). Methanol arises from methyl esters and methoxy groups found in plant materials (e.g., lignin and pectin) and also through chemical oxidation of methane in the troposphere and its subsequent deposition in rainfall. The ability to utilize methanol as the sole source of carbon and energy is found in representatives of all three taxonomic domains: methanogenic methylotrophs belonging to the Archaea and other anaerobic and aerobic methylotrophic prokaryotes belonging to the eubacteria (with both Gram-negative and Gram-positive representatives); a number of methanol utilizing yeasts (Eucarya) and filamentous fungi have also been isolated. Methane utilizing yeasts have also been described (1), but these have exceptionally long generation times and are currently of no more than academic interest. There is strong evidence for anaerobic methane oxidation associated with sulfate reducing and/or methanogenic activity (2). However, the organisms responsible have been difficult to culture axenically. Hence most of the work on methanotrophic organisms has concentrated on the methane utilizing aerobic (eu)bacteria (2).

In addition to the more obvious reduced C_1 substrates methane and methanol, nature provides a wide range of compounds based on methylated heteroatoms such as methylamines, methyl sulfides, methyl sulfonates, methoxy compounds, and halogenated methanes. Oxidation or hydroxylation of these methyl groups results in the release of formaldehyde, which can feed directly into the dissimilatory and assimilatory pathways of C_1 metabolism. Some of these organisms play an important role in nutrient cycling in the environment and probably contribute to the natural degradation of toxic compounds.

GROWTH ON C_1 COMPOUNDS

Bacterial growth on methane and methanol was first described nearly a century ago. However, the subject lay fairly dormant until the 1960s and 1970s, when research activity driven partly by academic studies on the relationship between methylotrophic and autotrophic biochemistry and, in no small way, by the perceived applications of methylotrophs increased dramatically. The latter originated from the interest in producing SCP (i.e., microbial biomass) from methane or methanol as an animal or human food supplement, a venture in which a number of companies became involved to a greater or lesser extent but that was ultimately shown to be uneconomic. Despite this uncertain start the basic principles on which the choice of C_1 substrates was made remain valid, namely that substrates such as methane and methanol are cheap and because they are highly reduced should give high growth yields. As substrate costs become an increasingly important factor with increasing production volume (and decreasing unit price of the product), the potential applications of methylotrophs are likely to be in bulk products. Alternatively, the enzymes or metabolic sequences unique to methylotrophic metabolism may offer some possibilities not available in heterotrophs. The biotechnological application of methylotrophs can therefore be conveniently categorized under three headings:

1. Processes highly dependent on substrate economics
2. Processes based on C_1-specific enzymes or metabolism
3. Processes or products incidentally associated with C_1 metabolism

Although these subdivisions are not mutually exclusive they are nonetheless useful in explaining the rationale behind the development of certain C_1-based processes.

Metabolic Biochemistry and Energetics

The pathways of carbon assimilation affect both the efficiency of growth and the types of products that might feasibly be produced from methylotrophs. In aerobic (true) methylotrophic bacteria two fundamentally different routes of carbon assimilation have been demonstrated: the ribulose monophosphate (RuMP) pathway and the serine pathway (3,4). The RuMP pathway (Fig. 1) has similarities to the Calvin Benson cycle of CO_2 fixation in the sense that 3 C_1 units (formaldehyde) are added to C5 acceptors (ribulose monophosphate) that subsequently undergo a se-

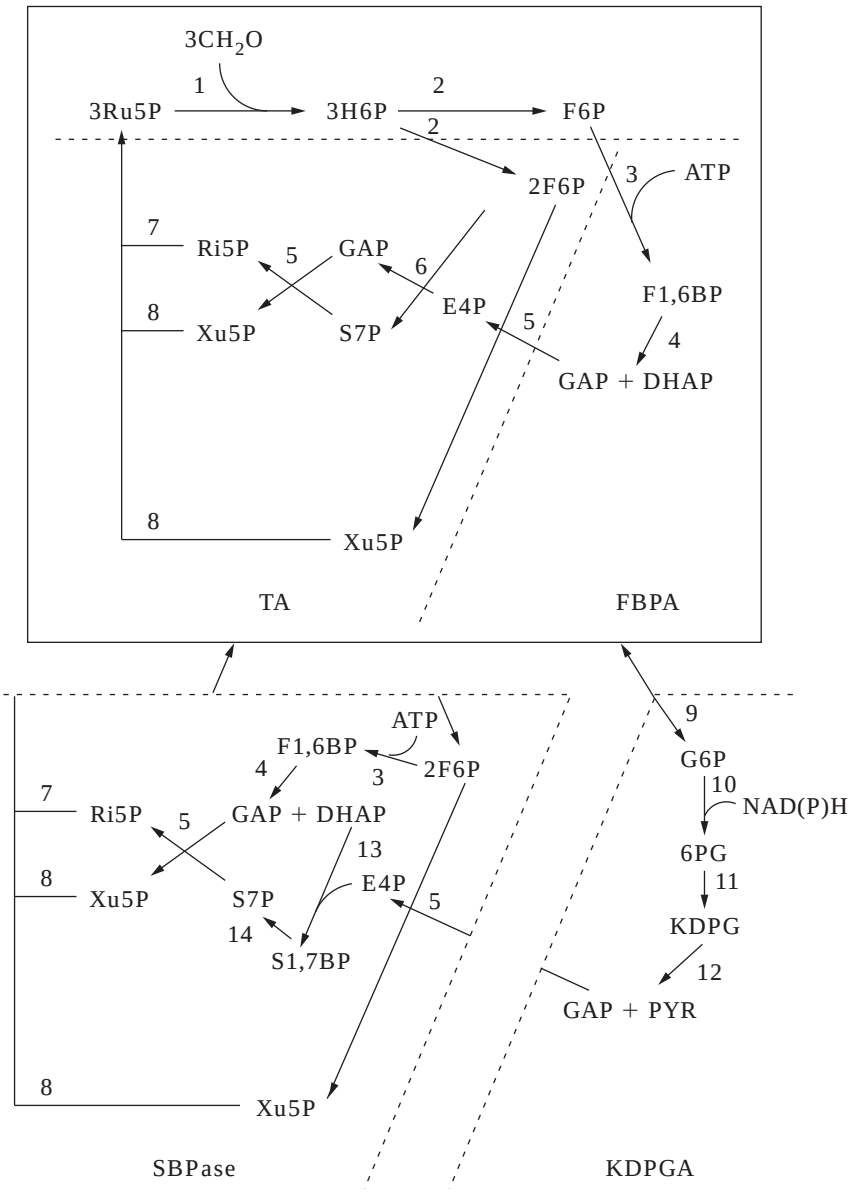


Figure 1. Intermediates and enzymes in the ribulose monophosphate pathway of formaldehyde fixation. Intermediates: Ru5P, ribulose-5-phosphate; H6P, hexulose-6-phosphate; F6P, fructose-6-phosphate; F1, 6BP, fructose-1,6-bisphosphate; DHAP, dihydroxyacetone phosphate; GAP, glyceraldehyde-3-phosphate; E4P, erythrose-4-phosphate; Xu5P, xylulose-5-phosphate; S7P, sedoheptulose-7-phosphate; Ri5P, ribose-5-phosphate; G6P, glucose-6-phosphate; 6PG, 6-phosphogluconate; KDPG, 2-keto-3-deoxy-6-phosphogluconate; PYR, pyruvate; S1,7BP, sedoheptulose-1,7-bisphosphate. Enzymes: 1, H6P synthase; 2, H6P isomerase; 3, phosphofructokinase; 4, F1,6BP aldolase; 5, transketolase; 6, transaldolase (TA); 7, Ri5P isomerase; 8, Pentose phosphate 3-epimerase; 9, phosphoglucose isomerase; 10, G6P dehydrogenase; 11, 6PG dehydratase; 12, KDPG aldolase (KDPGA); 13, S1,7BP aldolase; 14, S1,7BPase (SBPase).

ries of rearrangements yielding a C3 intermediate for cell biosynthesis and regenerating the acceptor molecules. Two variations of both cleavage reaction (fructose-1,6-bisphosphate aldolase [FBPA] and 2-keto-3-deoxy-6-phosphogluconate aldolase [KDPGA]), which generates the C3 biosynthetic intermediate, and the rearrangement stage (transaldolase [TA] and sedoheptulose-1,7-bisphosphatase [SBPase]), which regenerates the acceptor molecule, have been recognized and give four possible combinations. A comparison of the energetics of assimilation based on the production of a common C3 intermediate, pyruvate, shows that the FBPA/TA combination which is found in a number of Gram-positive methylotrophs, is the most efficient (Table 1).

The xylulose monophosphate (XuMP) pathway found in methylotrophic yeasts (5) has similarities to the RuMP pathway, being based on carbohydrate transformations and involving condensation of C_1 and C5 units and subsequent rearrangements based on transaldolases and transketolases (Fig. 2). However, the serine pathway found in bacteria is fundamentally different, involving assimilation of a C_1 unit into serine and subsequent rearrangements based on hydroxy acids and amino acids (Fig. 3). It also involves a CO_2 fixation step. Neither the XuMP nor serine pathway is as energetically efficient as the RuMP pathway but may offer advantages over the RuMP pathway organisms for the accumulation of certain metabolites.

Most of the yeasts, nonmethanotrophic serine pathway, and Gram-positive RuMP pathway organisms are facultative methylotrophs (3) able to use a range of heterotrophic substrates (e.g., carbohydrates). All of the methanotrophs appear to be obligate C_1 utilizers, some being very sensitive to excess methanol; many of the Gram-negative RuMP organisms are restricted facultative methylotrophs, growing poorly on some heterotrophic substrates. For me-

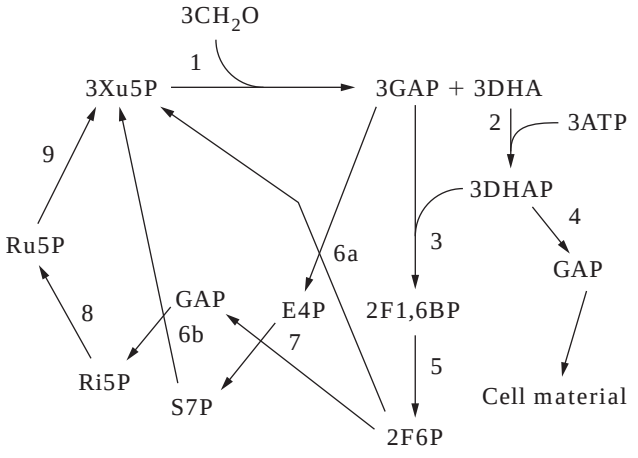


Figure 2. Intermediates and enzymes in the xylulose monophosphate pathway of formaldehyde fixation. Intermediates: Xu5P, xylulose-5-phosphate; GAP, glyceraldehyde-3-phosphate; DHA, dihydroxyacetone; DHAP, dihydroxyacetone phosphate; F1,6BP, fructose-1,6-bisphosphate; F6P, fructose-6-phosphate; E4P, erythrose-4-phosphate; S7P, sedoheptulose-7-phosphate; Ri5P, ribose-5-phosphate; Ru5P, ribulose-5-phosphate. Enzymes: 1, DHA synthase; 2, triokinase; 3, F1,6BP aldolase; 4, triose phosphate isomerase; 5, F1,6BPase; 6, transketolase; 7, transaldolase; 8, Ri5P isomerase; 9, pentose phosphate 3-epimerase.

tabolite overproduction, facultative methylotrophs can offer significant advantages because the ability to grow on heterotrophic substrates allows the isolation of mutants and genetic manipulation of methylotrophic metabolism.

Biomass Production

Despite the demise of C_1-based SCP processes, any future process adopted on the basis of substrate cost will inevitably involve growth of methylotrophs on a large scale. A retrospective consideration of the development of large-scale C_1-based processes is therefore warranted.

Table 1. Energetic Requirements of the Different Variants of the RuMP Cycle for the Production of Pyruvate ($3CH_2O \rightarrow$ Pyruvate), Chosen as a Typical Metabolic Intermediate

RuMP variant	NAD(P)H	ATP
KDPGA/SBPase	+1	−1
KDPGA/TA	+1	0
FBPA/SBPase	+1	0
FBPA/TA	+1	+1

KDGPA, 2-keto-3-deoxy-6-phosphogluconate aldolase; SBPase, sedoheptulose-1,7-bisphosphatase; FBPA, fructose-1,6-bisphosphate aldolase; TA, transaldolase.

Growth Yields and Energetics. A summary pathway for bacterial methane oxidation is shown in Figure 4. The pathway for methanol oxidation in Gram-negative methylotrophic bacteria that do not use methane is essentially the same, with the omission of the first step, although in some organisms formaldehyde is oxidized in a cyclic manner (6) after fixing in the initial steps of the assimilation pathway. Some methylotrophs may not use an NAD-linked formaldehyde dehydrogenase. However, those that do should give the highest growth yields and are therefore more likely to be adopted in a commercial process. $PQQH_2$ in methanol dehydrogenase is reoxidized by passing electrons via a periplasmic cytochrome c_L to the respiratory chain with energy conservation via chemiosmosis yielding approximately 1 ATP/$PQQH_2$ (7). Gram-positive methylotrophs use an NAD-linked methanol dehydrogenase (8). In principle, reoxidation of the resulting NADH via the respiratory chain should yield more ATP than gained from methanol oxidation in Gram-negative bacteria, which should translate into higher growth yields. For a wide range of substrates the yield of cells is proportional to the reduction level of the substrate, because biochemical oxidation yields proportionate amounts of ATP and cell yield is proportional to the ATP yield from substrate oxidation (9). Thus, one might expect that cell yields on methane

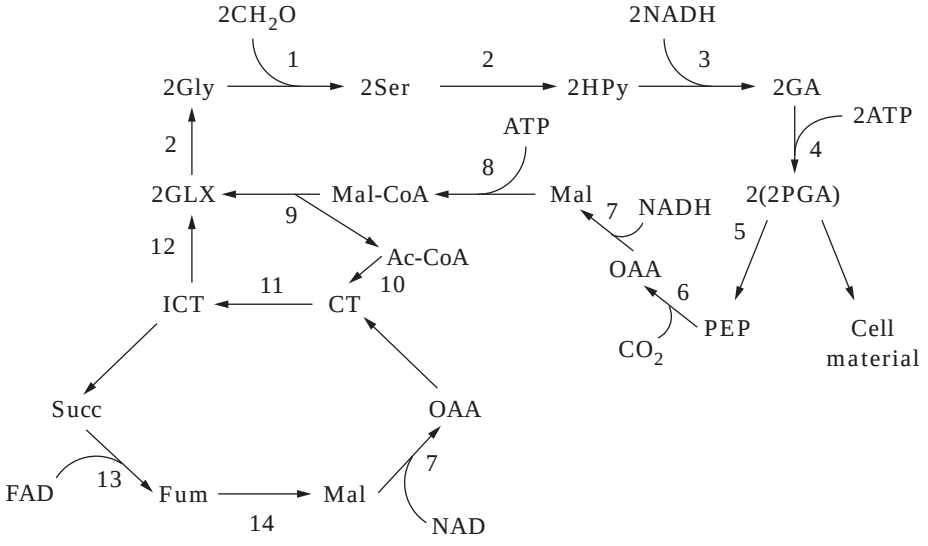


Figure 3. Intermediates and enzymes in the serine pathway of formaldehyde fixation. Intermediates: Gly, glycine; Ser, serine; HPy, hydroxypyruvate; GA, glycerate; 2PGA, 2-phosphoglycerate; PEP, phosphoenolpyruvate; OAA, oxaloacetate; Mal, malate; Mal-CoA, malyl-CoA; Ac-CoA, acetyl-CoA; CT, citrate; ICT, isocitrate; Succ, succinate; Fum, fumarate; GLX, glyoxylate. Enzymes: 1, serine transhydroxymethylase; 2, serine-glyoxylate aminotransferase; 3, hydroxypyruvate reductase; 4, glycerate kinase; 5, enolase; 6, PEP carboxylase; 7, malate dehydrogenase; 8, malate thiokinase; 9, malyl-CoA lyase; 10, citrate synthase; 11, aconitase; 12, isocitrate lyase; 13, succinate dehydrogenase; 14, fumarase.

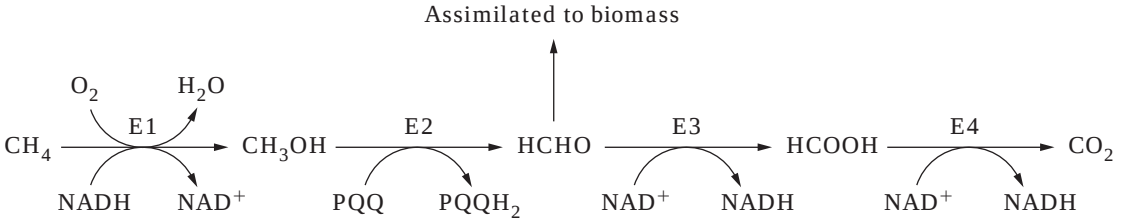


Figure 4. Generalized pathway for methane catabolism in methanotrophic bacteria. E1, methane monooxygenase; E2, methanol dehydrogenase; E3, formaldehyde dehydrogenase; E4, formate dehydrogenase; PQQ, pyrroloquinoline quinone.

would be higher than those on methanol. However, this relationship breaks down for highly reduced substrates; Figure 4 demonstrates the causes for methane oxidation. The biological oxidation of methane requires an energy input in the form of NADH, thus making methanol a potentially higher yielding substrate than methane (this is borne out in practice [10]). A number of other factors also mitigate against the use of methane as a substrate:

1. Flammability
2. Lack of compressibility
3. Low aqueous solubility
4. High oxygen demand (from Fig. 4 it is evident that the minimum oxygen demand would be 1.5 O_2/CH_4) resulting from the operation of methane monooxygenase

It should also be recognized that natural gas contains higher saturated and unsaturated hydrocarbons that can be partially oxidized by methanotrophs, resulting in the accumulation of inhibitory by-products. Therefore, except in circumstances in which the activity of methane monooxygenase (MMO) is required, bulk product formation from methylotrophs is likely to be based on methanol as substrate. Methanol can be derived from methane or a variety of other fossil fuels or renewable energy sources via synthesis gas (Fig. 5). The latter has also been proposed as a substrate for bulk product formation using carboxidotrophic bacteria (11) that use CO as a carbon source. However, despite the higher solubility of CO compared with methane, the problems of handling large volumes of a toxic substrate mitigate against this proposal.

Biomass Production from Methanol. A number of companies developed pilot-scale processes for biomass production from methanol (Table 2), with Imperial Chemical Industries (ICI) in the United Kingdom going on to full-scale

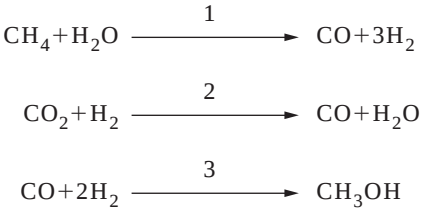


Figure 5. Imperial Chemical Industries process for the catalytic synthesis of methanol from methane 1, 15–20% Ni catalyst on Al_2O_3 or SiO_2, 700–900 °C, 1–25 bar; 2, Ni catalyst; 3, Cu/Zn catalyst, 250–280 °C, 70–110 bar.

production in a 1,500 m^3 continuous bioreactor using a Gram-negative bacterium. This was a significant technical achievement, the largest aseptic continuous culture process ever to operate.

The oxygen demand for growth on methanol (0.6–0.7 O_2/CH_3OH), although significantly less than that for methane, is still high compared with growth of heterotrophs on glucose (approximately 0.4 $O_2/[CH_2O]$). This high oxygen demand is a consequence of the relative inefficiency of energy conservation of periplasmic methanol dehydrogenase. For large-scale processes this inefficiency results in significant heat generation and oxygen transfer problems. The ICI solution to these problems was to use the moderately thermotolerant (t_{opt} = 40 °C) bacterium *Methylophilus methylotrophus* and a high aspect ratio (42 m × 7 m) pressure cycle reactor (17). This design also minimized the number of mechanical seals in the reactor, which were the most likely routes of contamination in long-term continuous operation. Gas transfer and mixing were achieved according to the airlift principle, with pressurized air being introduced at the base of the riser. However, the height of the liquid was such that the oxygen dissolved almost completely as it entered the reactor. To avoid toxicity or nutri-

Table 2. Full- or Pilot-Scale SCP Processes Using Methanol as Substrate

Company	Capacity (m^3)	Product (organism)	Reference
Imperial Chemical Industries PLC, U.K.	1,500	Pruteen *(Methylophilus methylotrophus)*	12
Hoechst-Uhde, Germany	20	Probion *(Methylomonas clara)*	13
Phillips Petroleum Co. (Provesta Corp.) U.S.	25	Provesteen *(Pichia pastoris)*	14
Norsk-Hydro, Sweden	45	Norprotein *(Methylomonas methanolica)*	15
Mitsubishi Gas-Chemical Co., Japan	2	(*Pseudomonas* sp.)	16

ent limitation methanol was introduced through a series of nozzles placed at various points up the riser. To achieve a product of the right composition as an animal feed, the reactor was run at a dilution rate of 0.16–0.19 h^{-1} (μ_{max} = 0.55 h^{-1}) with a cell density of 30 g dry wt L^{-1} and a projected annual production capacity of 50,000–70,000 tons. *M. methylotrophus* uses the KDPGA/TA route of C_1 assimilation (18), which, as discussed earlier, is marginally less efficient than the FBPA/TA route. It was also found to assimilate ammonia solely via the less energetically efficient glutamine synthase–glutamine oxoglutarate aminotransferase (GS/GOGAT) route. By knocking out this route and expressing glutamate dehydrogenase (GDH) from *Escherichia coli* in this organism a 7% improvement in yield from methanol was achieved (19). However, the latter construct was never used for commercial production.

Phillips Petroleum (later Provesta Corporation) in the United States opted for a process using the methylotrophic yeast *Pichia pastoris*, a decision based on the greater public acceptability of yeasts in nutrition. Methanol dissimilation in yeasts occurs by an energetically inefficient route involving an alcohol oxidase and catalase (Fig. 6). Not only is there no energy conservation arising from the conversion of methanol to formaldehyde but the oxidase mechanism involves stoichiometric amounts of oxygen. Although some oxygen is subsequently released through the action of catalase, the net requirement is $0.5O_2/CH_3OH$. Because inefficient energy conservation results in excessive heat generation the low optimum growth temperature (30 °C) was also a disadvantage, as was the long generation time (5.5–6 h). Nevertheless, through medium optimization and bioreactor design the Phillips process managed to achieve oxygen transfer rates of 0.8–1 mol L^{-1} h^{-1}, allowing operation at very high cell densities and producing a continuous slurry containing 125–150 g dry wt L^{-1} cells that could be spray-dried directly without concentration (14).

POTENTIAL APPLICATIONS

Processes Dependent on Substrate Economics

With the exception of SCP three other types of bulk product have been or are currently being considered as products from methylotrophs: poly-β-hydroxyalkanoate–based biodegradable plastics, polysaccharides, and amino acids. Although the driving force for all of these programs is substrate economics, in the last case the mode of carbon assimilation may predispose certain methylotrophs toward production of particular classes of amino acids.

One practical benefit from the research arising from the SCP programs was a detailed understanding of metabolic pathways and the energetics of carbon assimilation and dissimilation in methylotrophs. Using this information, Linton and Niekus (20), Collins (21), and Ackermann and Babel (22) have compared the likely carbon conversion efficiencies and oxygen requirements for conversion of methanol to these three bulk products. Predictions for amino acid production were based on glutamate. For polysaccharides and glutamate the predicted yields (w/w) from glucose and methanol are similar (Table 3), indicating that, for equally high yielding organisms, methanol would offer a cost advantage. However, the higher oxygen demand for methylotrophs is a significant disadvantage. In the case of exopolysaccharides the yield on oxygen is predicted to be 10-fold lower than that on glucose. Because polysaccharide volumetric productivity from glucose is often limited by the efficiency of oxygen transfer into increasingly viscous media the substrate cost advantage is more than offset by this disadvantage. Therefore, exopolysaccharides produced from methanol would have to have some unique characteristics to justify commercial production. A few years ago, Celgene Corporation marketed Poly-54/Methylophilan, an exopolysaccharide produced by continuous culture of *M. viscogenes*. Several uses were envisaged, but of particular note was its high viscosity at 80 °C (higher than xanthan gum), which, unlike xanthan, was maintained for extended periods. This suggested a possible role in enhanced oil recovery and in oil drilling muds. However, production has subsequently ceased (D. Stirling, personal communication, June 1997).

Amino Acid Production. Glutamate production from methanol has a threefold disadvantage compared with glucose in terms of yield from oxygen. Thus, production of glucose or other bulk amino acids from methanol has more realistic prospects than polysaccharide production. Indeed, methanol was being examined as an alternative substrate for glutamate production in the early 1970s. It is interesting to note that yields of up to 33 g/L were obtained in these early studies using metabolically ill-defined Gram-negative methylotrophs (23), which probably used the RuMP pathway. Similar yields have been reported more recently (24) using mutants of the Gram-negative RuMP organism *Methylobacillus glycogenes* (formerly classified as *Pseudomonas insueta* and *Protaminobacter thiaminophagus*). However, Ackermann and Babel have pointed out (22) that, despite their lower growth yields, serine pathway organisms should potentially give higher yields of glu-

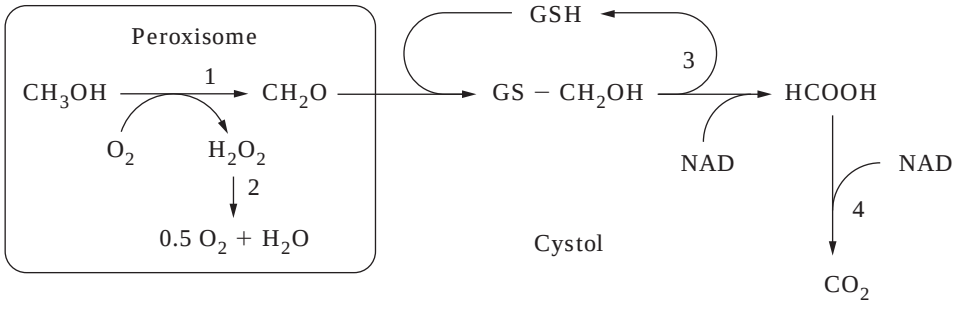


Figure 6. Linear (noncyclic) pathway for methanol dissimilation in methylotrophic yeasts.

Table 3. Predicted Product Yields (Y_s) from Methanol and Glucose, and the Associated Yield from Oxygen (Y_{O_2}).

Product[a]	Carbon source	Substrate requirement	Y_s (g/g)	Y_{O_2} (g/g)
EPS ($C_6H_{10}O_5$)	Methanol	$6.14CH_3OH + 3.21O_2$	0.82	1.58
	Glucose	$1.05\ C_6H_{12}O_6 + 0.3O_2$	0.86	16.3
Glutamate ($C_5H_9O_4N$)	Methanol^R	$6CH_3OH + NH_3 + 4.5O_2$	0.76	1.02
	Methanol^S	$5CH_3OH + NH_3 + 2.5O_2$	0.92	1.84
	Glucose	$C_6H_{12}O_6 + NH_3 + 1.5O_2$	0.82	3.06
PHB ($C_4H_6O_2$)	Methanol	$5.5CH_3OH + 3.75O_2$	0.49	0.72
	Glucose	$C_6H_{12}O_6$	0.48	[b]

Note: Yields are based on the known metabolic pathways between substrate and product and their associated energy requirements; substrate required for biomass formation is not considered. EPS synthesis from methanol is based on the RuMP pathway (20); glutamate synthesis from methanol (20,22) is based on both the RuMP pathway (^{R}) and the serine pathway (^{S}); PHB synthesis from methanol is based on the serine pathway and makes the same assumptions as cited in Ref. 21 regarding efficiency of methanol usage.

[a]EPS, extracellular polysaccharide (hexose); PHB, poly-β-hydroxybutyrate.

[b]Conversion of glucose to PHB generates ATP and NAD(P)H. There is therefore no additional ATP required for PHB synthesis, so the yield based on oxygen usage is, technically, infinite.

tamate from methanol; it should be borne in mind that high yields of amino acids may also depend on the type of organism being used.

Glutamic acid has traditionally been produced from *Corynebacterium* or *Brevibacterium* spp., which, as Gram-positive organisms, appear to be disposed toward this sort of application, partly because of the lack of an outer membrane and partly because of efficient glutamate efflux systems. Although the majority of Gram-positive methylotrophs isolated use the RuMP pathway (3) they may offer a greater possibility for realizing the theoretical yields of glutamate than the Gram-negative organisms previously studied. Furthermore, the isolation of stable mutants from Gram-negative methylotrophs is known to be difficult. Bearing in mind this difficulty and the cooling requirements for large-scale production, Schendel and coworkers (25) have isolated lysine overproducing mutants of a thermotolerant (t_{opt} = 50 °C), methylotrophic *Bacillus* sp. Process modeling and optimization (26) have demonstrated that production is primarily growth associated and optimal productivity should be obtained in a variable-feed, constant-volume fed-batch process. A final concentration of greater than 45 g/L lysine, which also included a significant amount of by-products (mainly glutamate and alanine) has been reported (27). Therefore, there is further room for improvement in control of metabolic flux to lysine. Abouzeid and coworkers (28) are examining the possibility of aromatic amino acid production (e.g., phenylalanine) from the Gram-positive *Amycolatopsis methanolica* (formerly *Nocardia* sp. 239). In the latter case, there is also a strong precedent for starting with a RuMP (or XuMP) pathway organism because erythrose-4-phosphate, a precursor for aromatic amino acid biosynthesis, occupies a central position in the RuMP and XuMP pathways.

Serine and methionine have also been considered for methylotrophic fermentative production. The former is required as a precursor in tryptophan production and is also included in some cosmetics. Its catalytic production from glycine and methanol by methylotrophs was an obvious target because serine hydroxymethyltransferase (SHMT) catalyzes the first step in carbon assimilation by serine pathway organisms and should therefore be expressed at high levels. Mutants of *Hyphomicrobium methylovorum* were reported (29) to produce 34 g/L serine from 100 g/L glycine and 48 g/L methanol, while resting cells of a strain of *Methylobacterium extorquens* produced 54 g/L serine (30). However, considerably higher concentrations have been achieved (31) in an enzyme bioreactor using a recombinant SHMT (from *E. coli*), and it is unlikely that a methylotrophic whole organism process will supplant this concentration. Unlike the previous examples, methionine for animal feeds is currently produced chemically from acrolein and methanethiol. Hence, the economic feasibility of fermentative whole cell biocatalytic production by methylotrophs is more difficult to assess than in the previous cases. Biosynthesis requires L-homoserine (from aspartate), serine, sulfide, and the C_1 carrier 5-methyl-THF, suggesting that serine pathway organisms would also be the obvious candidates. This suggestion was made valid by the studies of Yamada and coworkers on ethionine-resistant mutants of RuMP and serine pathway organisms (32,33). However, none of the yields were commercially significant, and the lack of any recent literature on this topic suggests that it is not being pursued.

Biodegradable Plastics from Methanol. Marlbrough Biopolymers, a subsidiary of Monsanto, now produces a biodegradable plastic under the tradename Biopol by fermentation using *Alcaligenes eutrophus* grown on glucose as the carbon and energy source (34). The plastic is a derivative of poly-β-hydroxybutyrate (PHB), a high molecular weight polyester produced by many bacteria as a carbon and energy storage compound. *A. eutrophus* was selected because PHB can comprise up to 80% of its dry weight under appropriate conditions and also because, when supplied with propionic acid, the organism produces a random copolymer of 3-hydroxybutyrate and 3-hydroxyvalerate, P(3HB-3HV), which is tougher and more flexible than pure PHB. Current production runs at about 1,000 tons per annum. However, it has been predicted that even at a 10-fold higher level the cost of P(3HB-3HV) would be about $5/kg, similar to that of other biodegradable plastics but considerably higher than synthetic plastics derived from petrochemicals (35). Glucose accounts for 35–40% of this price. Therefore, at present, microbially derived plastics are still in the category of niche products.

Two underlying research trends are evident in this area. For truly bulk production of PHB-based plastics at a price similar to synthetics it has been recognized that production in plants such as rapeseed or soybean represents the way forward, and some progress in this area is being made (35). Production in bacteria, due to its higher cost, is more likely to concentrate on niche products. In that light, the evidence that PHB is just one of a family of polyhydroxyalkanoates (PHAs) and that some PHA synthases will accept hydroxyacyl CoAs derived from a wide range of derivatized carboxylic acids supplied to the medium opens up a potentially fruitful avenue for the production of specialist biodegradable polymers (36). The greater ease of physiological and genetic control of metabolic flux in bacteria would favor production by fermentation.

Because substrate costs represent a significant proportion of the selling price of PHAs, substitution with a cheaper substrate could offer some short- and long-term benefits for P(3HB-3HV) and specialist PHA production, respectively. Collins (21) calculated that the maximum theoretical yield from methanol would be 0.18 g/g compared with 0.33 g/g from glucose. Although the price differential between the two substrates would still favor methanol, the cost savings are marginal. In the early stages of Biopol development methanol was apparently considered as a substrate but not pursued because the product was of low molecular weight and extraction was difficult. More recent research by a number of groups has shown that high molecular weight PHB can be produced from methylotrophs and that P(3HB-3HV) can be produced in methylotrophs by supplementing the fermentation with *n*-amyl alcohol or valeric acid (propionic acid was less effective). However, the 66% of dry weight (136 g/L PHB) achieved by Suzuki et al. (37) using computer-controlled fed-batch fermentation is the highest reported yield for any methylotroph. This apparent limit probably reflects a restriction of metabolic flux to the PHB precursors, which suggests that genetic manipulation of flux control mechanisms may be necessary to match the yields currently achieved with *A. eutrophus*. PHB synthesis in methylotrophs appears to be restricted to *Methylobacterium* spp. ("pink-pigmented facultative methylotrophs") and the autotrophic methylotrophs (38). *Methylobacterium* spp. assimilate carbon via the serine pathway and have a complete TCA cycle, unlike restricted facultative methylotrophs that use the RuMP cycle. It is evident that, without a complete TCA cycle, acetyl CoA resulting from the breakdown of PHB would be of limited benefit. Hence, the latter organisms have evolved to accumulate extracellular polysaccharide rather than PHB. Recent work on the cloning from *A. eutrophus* and heterologous expression of genes required for PHB biosynthesis (39) is extending the range of potential strains for PHB production. Thus, RuMP pathway methylotrophs may be considered in the future.

Processes Based on C_1 Metabolism

Although the pathways of carbon assimilation in methylotrophs are specialized, few of the enzymes involved are unique. In the RuMP pathway only the hexulose phosphate synthase and isomerase are peculiar to C_1 metabolism. Therefore, beyond the observation that certain routes of C_1 assimilation might provide a higher flux of precursors for amino acid biosynthesis, for example, the assimilatory enzymes of C_1 metabolism have not been exploited. However, there has been considerable interest in the development of processes based on C_1 dissimilatory enzymes, in particular methane monooxygenase, methanol dehydrogenase, methanol oxidase (from yeasts), and formate dehydrogenase.

Methane Monooxygenase. Two types of MMO have been recognized: a nonheme iron-soluble (cytoplasmic) enzyme (sMMO) and a membrane-associated enzyme (pMMO; p, particulate) that requires copper for activity and has some similarity to ammonia monooxygenase (40). Soluble MMOs from *Methylococcus capsulatus* (Bath) and *Methylosinus trichosporium* OB3b have been extensively studied, and crystal structures of the oxygenase component from both enzymes are available (41). Additionally, the genes encoding sMMO from a number of sources have been cloned, sequenced (42), and expressed in heterologous hosts (43). Both sMMO and pMMO insert oxygen into a range of substrates other than methane, with sMMO having a very low substrate specificity (44,45). Initially this observation led to interest in the use of cells expressing MMO as biocatalysts for synthetic biotransformations, particularly in the conversion of propene to propene oxide. Although a bulk product, the chemical synthesis of propene oxide is not straightforward, and there was considerable interest in the possibility of a one-step biological process that did not yield by-products (46). However, the dual requirements of high catalytic rates and extended biocatalytic lifetimes could not be realized, primarily because the process of propene epoxidation actually inactivated the MMO (47). Other potential applications as a synthetic biocatalyst have also foundered, partly because sMMO has a low stereoselectivity and therefore has little value for the synthesis of chiral chemicals. Another limitation is that with many substrates it exhibits poor regioselectivity, producing, for example, mixtures of alkan-1-ols and alkan-2-ols from the oxidation of straight chain alkanes (44).

More recently the interest in MMO has focused on applications in biotreatment, particularly of the priority pollutant trichloroethene (TCE) and other low molecular weight partially halogenated hydrocarbons (e.g., dichloroethene, vinyl chloride). In fact, MMO is only one of a number of mono- and dioxygenases that have been shown to fortuitously co-oxidize these compounds (48). Cometabolism processes typically yield no benefit to the host organism and, in the case of co-oxidation, represent an energy drain, requiring an additional energy source for continued operation of the oxygenase. In the case of methane oxidizing bacteria, methanol is an obvious energy source, although growth on methanol may have undesirable physiological consequences (see later section). Oxidation of TCE by methane oxidizing bacteria leads to the transient accumulation of TCE epoxide (49), which is chemically unstable. As a consequence, TCE epoxide and some of the derived breakdown products are also highly reactive and can give rise to toxicity. This result has been demonstrated with both purified sMMO (49) and whole cells expressing

sMMO (50). However, it seems that the toxicity can be alleviated by continued supply of reducing power and particularly by the use of mixed rather than axenic cultures (51). In the latter, it is probable that other members of the bacterial consortium assist in the removal of toxic intermediates, thus prolonging the life of the primary biocatalyst.

Studies with pure cultures have demonstrated that oxidation of TCE is fastest with bacteria expressing the sMMO (2). The ability to express sMMO is restricted to certain types of methanotrophs: primarily the type II (serine pathway) and type X bacteria such as *Methylosinus trichosporium* and *Methylococcus capsulatus*, respectively. Furthermore, in these organisms, production of sMMO is regulated by the availability of copper in the medium. With a copper-sufficient medium these organisms preferentially produce the pMMO (which results in a higher growth efficiency). Therefore, to get high rates of TCE oxidation, it is necessary to grow these organisms in a copper-deficient medium. Although this condition may be feasible in laboratory experiments it may be more difficult to achieve in operational biotreatment plants and impossible to control for in situ biotreatment operations. However, with increasing access to the molecular genetics of methanotrophs (42), the generation of stable mutants (e.g., by reverse genetics [52]) unable to synthesize pMMO presents an interesting prospect. Regulatory mutants that constitutively produce sMMO would also circumvent another physiological problem, which is that growth of type II and type X organisms on methanol, even under copper limitation, results in production of pMMO (53) (at very low copper concentrations this is produced as an apo-enzyme). Mutants that constitutively produced sMMO could be supplied with methanol as a growth substrate and also to drive the co-oxidation reaction, without the risk of switching off sMMO synthesis.

Several reports (2) have described the practical applications of methanotrophs for TCE and other chlorinated hydrocarbon degradation, both in controlled bioreactors and in situ. Bioreactor studies have mainly been based on *Methylosinus trichosporium*, which was shown, in laboratory studies using copper-limited growth conditions, to have the highest reported specific activity of TCE degradation of any of the oxygenase-based systems that have been examined. Because of the problems of sMMO expression using methanol as growth substrate, these reactors have usually been supplied with methane to sustain the active biomass and formate to drive the biotransformation. However, with a continuously fed bioreactor, methane and TCE compete for the sMMO, which reduces rates of biodegradation. A practical alternative is the use of sequencing reactors (54) that are alternately supplied with methane and TCE. The feasibility of in situ biotreatment (55) of contaminated aquifers by stimulating indigenous methanotrophic activity has also been investigated. Injection of methane–air mixtures was shown to stimulate growth of methanotrophs, particularly type II methanotrophs, which have the potential to produce sMMO. In one study, degradation of TCE was shown to increase after this treatment, which suggests that these organisms were also expressing sMMO. However, without the ability to control the availability of copper for growth, it is more difficult to optimize expression of sMMO activity and hence TCE biodegradation in in situ treatment.

Whole cells of methanotrophic bacteria have also been considered as methane sensors (56) (using oxygen uptake as the mode of detection) and methane filters (57) for use in the mining industry. However, neither of these potential applications has been developed beyond the laboratory stage.

Methanol Dehydrogenase. Methanol oxidation in Gram-negative aerobic methylotrophs involves a periplasmic methanol dehydrogenase (MDH) that uses pyrroloquinoline quinone (PQQ, or methoxatin) as a cofactor. Reducing equivalents from MDH would normally be passed to a soluble cytochrome c (7) but can also be passed to artificial electron acceptors such as ferrocenes, which are used as mediators in biosensors. Higgins et al. (58) developed a sensor based on MDH that had a rapid response time and useful sensitivity range (1–100 μM). However, MDH oxidizes a number of primary alcohols in addition to methanol. Although the enzyme from *Rhodopseudomonas acidophila* has a higher affinity for ethanol than for other primary alcohols (59), this lack of specificity limits the usefulness of MDH in diagnostics. In contrast, the broad substrate range may present opportunities for use in synthetic biotransformations (60). In Japan there also appears to be a market for PQQ, the MDH cofactor, as a health food (61).

Methanol or Alcohol Oxidase. Methanol oxidase is the enzyme used by eukaryotic methylotrophs to catalyze the first step in methanol metabolism. It produces formaldehyde (from methanol) and hydrogen peroxide and is normally found in association with catalase, which degrades the peroxide (5). As with MDH, methanol oxidase also oxidizes a range of higher alcohols and is therefore often referred to as alcohol oxidase (AOX). In methylotrophic yeasts AOX and catalase are expressed at very high levels and are located in organelles known as peroxisomes. Peroxisomes can be seen in thin sections to occupy most of the intracellular volume. This very high level of expression has been exploited for the development of eukaryotic expression systems, which are discussed later.

As with MDH, AOX has been considered for use in biosensors but with detection based on oxygen uptake rather than through direct or indirect electron transfer reactions (62). Similar to MDH, the enzyme also has a broad substrate range. Methylotrophic yeasts and AOX have also been considered for the production of formaldehyde, where it has been demonstrated that surprisingly high concentrations (0.95 M) could be accumulated by mutants of *Candida boidinii* (63) and for production of hydrogen peroxide. The latter is toxic to most AOX enzymes, but that from *Pichia pastoris* NRRL Y-4290 appears to be remarkably resistant (64). The production of peroxide using methanol oxidase from methylotrophic yeasts has been patented by Unilever, although there is no evidence that it has been put into commercial practice.

Formate Dehydrogenase. The formate dehydrogenases (FDH) from methylotrophic bacteria appear to be surprisingly complex given their apparently simple function.

However, those from methylotrophic yeasts are simpler. The enzyme from *C. boidinii* is a homodimer with subunits of approximately 36 kDa; it is both oxygen insensitive and moderately thermostable (65). Because the end product of formate oxidation is gaseous and can therefore be easily removed from the reaction, the combination of formate and FDH is ideal for reduced nucleotide recycling. The enzyme is produced commercially and has been used with great effect for continuous biotransformation reactions requiring NADH (66).

Applications Fortuitously Associated with C_1 Metabolism

Heterologous Gene Expression in Methylotrophic Yeasts. One of the fortuitous and enduring aspects of the various SCP programs resulted from the discovery that the methanol or alcohol oxidase of the yeast *P. pastoris* was expressed at very high levels (5), essentially from a single chromosomal copy of the AOX gene (there are actually two AOX genes, but *AOX1* is expressed much more strongly than *AOX2*). One of the problems with the use of *Saccharomyces cerevisiae* as a host for heterologous expression of eukaryotic proteins is the lack of a strong, yet tightly regulated, inducible promoter. During growth on methanol, methanol oxidase can comprise up to 30% of the soluble cell protein in *P. pastoris* (67), with the enzyme being localized in specialized microbodies (peroxisomes) together with catalase and dihydroxyacetone synthase, the first enzyme involved in formaldehyde assimilation. As mentioned earlier, these peroxisomes occupy most of the cell contents of methanol grown cells, but when grown on glucose, peroxisomes are virtually absent. A similar situation is found with another well-studied methylotrophic yeast, *Hansenula polymorpha* (68). The strong inducible *AOX1* promoter has been exploited by Provesta Corporation for the development of a powerful expression system that is now commercially available in a variety of forms (69). For production purposes, the system of choice involves cloning into an integrative additive or transplacement vector. Transplacement involves replacement of the *AOX1* locus by the cloned construct giving a strain with an aox1AOX2 phenotype, capable of only slow growth on methanol. Additive integration at the *AOX1* or *HIS4* locus will leave a functional *AOX1*, allowing continued growth on methanol.

Higher levels of heterologous gene expression are usually obtained with the transplacement system. Chromosomal integration yields stable expression systems; however, multiple copies of the expression cassette, usually associated with expression from plasmid vectors, result in higher levels of expression. Scorer and coworkers (70) have shown that, by including a kanamycin resistance gene in the integration cassette and selecting for high levels of resistance to the antibiotic G418, it is possible to select for multicopy integration, which occurs naturally in about 5–10% of transformants. In this way it has been possible to achieve expression levels comprising up to 32% of total cell protein (Table 4); by combining this technique with techniques for high cell density culture developed during the SCP program (71), yields of up to 12 g/L recombinant protein have been recorded. The highest protein yields have been obtained in cytosolic form. However, secreted proteins can be obtained using either native (including prokaryotic) secretion signals, the leader sequences from prepro-α mating factor, or invertase from *S. cerevisiae*. As with *S. cerevisiae, P. pastoris* also carries out some typical eukaryotic posttranslational modification including N- and O-linked glycosylation.

Table 4. A Selection of Heterologous Proteins Produced in *Pichia pastoris* (Pp) and *Hansenula polymorpha* (Hp)

Protein (MW)	Host/system	Yield (g/L)
HBsAg (23)	Pp (FB) + Hp	0.4 (Pp)
Human TNF (17)	Pp (FB + CC)	10
Human EGF (6)	Pp (FB)	0.5
Streptokinase-*Se* (47)	Pp (CC)	0.077
Aprotinin (6.5)	Pp (FB)	0.93
Tetanus tox. frag Cs*Ct*(50)	Pp (FB)	12
Pertactin-*Bp* (69)	Pp (FB)	3
Human PN-2 (6.8)	Pp	1
HIV gp120 (54)	Pp (FB)	1.25
HSA (67)	Pp (FB) + Hp	
Human IGF-1 (7.5)	Pp (FB)	0.5
Human lysozyme (14)	Pp	0.7
Human SOD (16)	Pp	0.75
IL-2	Pp	4
Human tPA	Pp	
Invertase-*Sc* (58)	Pp (FB) + Hp	2.5 (Pp)
Glucose oxidase-*An* (75)	Hp	2.25

Note: These proteins were chosen to illustrate the range of yields obtained with therapeutically important proteins. For a fuller list see Ref. 68. FB, fed batch; CC, continuous culture; *Se, Streptococcus equisimilis; Ct, Clostridium tetani; Bp, Bordetella pertussis; Sc, Saccharomyces cerevisiae; An, Aspergillus niger.*

Expression in *H. polymorpha* (Table 4) offers the same features seen with *P. pastoris*, potentially coupled with more efficient production because the optimum growth temperature is 37–43 °C, reducing the cooling requirements of large-scale processes, and growth on methanol is faster (68). However, because *H. polymorpha* has only one *AOX* gene, transplacement constructs are unable to grow on methanol as the sole carbon source. Usually glucose is used for growth, with methanol added to induce heterologous gene expression at the appropriate time. The inducible formate dehydrogenase (FMD) promoter has also been used.

Recent studies with a number of organisms have started to unravel the processes of peroxisome (and other microbody) proliferation and protein targeting. Given the dramatic increase in peroxisomes associated with methanol induction and the peroxisomal location of AOX, it has been suggested (68) that methylotrophic yeasts may offer an additional advantage for heterologous gene expression: namely, targeting to peroxisomes as a mechanism for compartmentalizing toxic or proteolytically susceptible proteins.

Production of Vitamins and Coenzymes. Although they do not command the high prices of recombinant proteins and other pharmaceutical products, vitamins and coenzymes occupy a price range dictated more by productivity than substrate economics. Nevertheless, various methylotrophs have been found to produce relatively high levels

of potentially valuable products such as cytochrome c, coenzyme Q_{10}, and vitamin B_{12}. Methanol oxidation occurs in the periplasm of aerobic Gram-negative methylotrophic bacteria and links directly to the respiratory chain via a soluble cytochrome c (7). Some wild-type strains have been observed to release significant quantities of this cytochrome into the culture medium during growth (71); cyanide- and glycine-resistant mutants accumulate levels higher than those of wild-type strains (72). The high respiratory capacity associated with aerobic methanol oxidation also underlies the ability of some mutant facultative methylotrophs to produce high levels of coenzyme Q_{10}. Both of these products have applications in cardiotherapy, with Q_{10} also finding an increasing market in dietary supplements, but face competition from sources such as spent brewing yeast or bovine or equine heart, neither of which incur specific production costs.

Vitamin B_{12} is produced commercially by fermentation using *Propionibacterium* spp. However, anaerobic methylotrophic methanogens and acetogens have been shown to accumulate high concentrations of corrinoids, including vitamin B_{12}. Despite the low biomass productivity and cell yields of these anaerobes, potential production with these organisms remains an active area of research (73).

CONCLUDING REMARKS

Despite the commercial failure of methanol-based SCP production, the rationale for adopting methylotrophic organisms for processes in which substrate costs form a significant part of overall production costs still holds true. The early work on methylotrophic bacteria was dominated by studies of Gram-negative organisms, and our understanding of the biochemistry of these organisms is now relatively advanced. Work on Gram-positive methylotrophs has lagged behind. However, given that Gram-positive organisms are favored for fermentation processes to produce low molecular weight metabolites such as amino acids (which increasingly fall into the high volume–low price category of products), it is probable that Gram-positive organisms will receive increasing attention. As far back as the early 1970s the potential of methylotrophs for producing such metabolites was appreciated. Given the increasing access to the molecular genetics of methylotrophic bacteria, it would not be surprising to see this area develop.

Substrate costs also underlie the high degree of interest in producing polyhydroxyalkanoates in methylotrophs. However, even if methanol could supplant carbohydrate in the current processes, PHA production costs could still not compete with those of bulk plastics. Thus, at least for the foreseeable future, PHAs are likely to remain a specialized product; as such, the fermentation substrate is likely to be dictated by the nature of the process and product required rather than the opposite.

Finally, it is probable that some of the specialized applications of methylotrophs will continue to develop. The *P. pastoris* expression system appears to be well established, and the more recently developed *H. polymorpha* system could provide a useful complement. Methanotrophic bacteria (or recombinant methane monooxygenase), on the other hand, may yet find an application in biotreatment plants. Certainly the practice of stimulating the proliferation of methanotrophs to boost their cometabolic contribution in natural ecosystems is a rational approach and may offer a cost-effective solution to some problems of environmental contamination.

BIBLIOGRAPHY

1. H.I. Wolfe, in H. Dalton ed., *Microbial Growth on C_1 Compounds*, Heydon, London, 1981, pp. 202–210.
2. R.S. Hanson and T.E. Hanson, *Microbiol. Rev.* **60**, 439–471 (1996).
3. L. Dijkhuizen, P.R. Levering, and G.E. de Vries, in J.C. Murrell and H. Dalton eds., *Methane and Methanol Utilizers*, Plenum Press, New York, 1992, pp. 149–181.
4. C. Anthony, in I. Goldberg and J.S. Rokem eds., *Biology of Methylotrophs*, Butterworth-Heinemann, Stoneham, Mass., 1991, pp. 79–109.
5. W. de Koning and W. Harder, in J.C. Murrell and H. Dalton eds., *Methane and Methanol Utilizers*, Plenum Press, New York, 1992, pp. 207–244.
6. C. Anthony, *Int. J. Biochem.* **24**, 29–39 (1992).
7. C. Anthony, *Adv. Microbiol. Physiol.* **27**, 113–210 (1986).
8. N. Arfman, E.M. Watling, W. Clement, R.J. van Oosterwijk, G.E. de Vries, W. Harder, M.M. Attwood, and L. Dijkhuizen, *Arch. Microbiol.* **152**, 280–288 (1989).
9. J.J. Heijnen and J.A. Roels, *Biotechnol. Bioeng.* **23**, 739–763 (1981).
10. D.J. Leak and H. Dalton, *Appl. Microbiol. Biotechnol.* **23**, 477–481 (1986).
11. E. Williams, J. Colby, G.W. Logan, and C.M. Lyons, in J.D. Stowell, A.J. Beardsmore, C.W. Keevil, and J.R. Woodward eds., *Carbon Substrates in Biotechnology*, IRL Press, Oxford, 1986, pp. 185–201.
12. P.J. Senior and J. Windass, *Biotech. Lett.* **2**, 205–210 (1980).
13. W. Sittig, *Chemtech.* **13**, 606–613 (1983).
14. L.K. Shay, H.R. Hunt, and G.H. Wegner, *J. Ind. Microbiol.* **2**, 79–85 (1987).
15. H. Mogren, *Process Biochem.* **14**, 2–7 (1979).
16. G.L. Solomons, *CRC Crit. Rev. Biotechnol.* **1**, 21–58 (1985).
17. R.B. Vasey and K.A. Powell, in G.E. Russell ed., *Biotechnology and Genetic Engineering Reviews*, vol. 2, Intercept, Newcastle upon Tyne, 1984, pp. 285–311.
18. A.J. Beardsmore, P.N. Aperghis, and J.R. Quayle, *J. Gen. Microbiol.* **128**, 1423–1439 (1982).
19. J.D. Windass, M.J. Worsey, E.M. Pioli, D. Pioli, P.T. Barth, K.T. Atherton, E.C. Dart, D. Byrom, K. Powell, and P.J. Senior, *Nature (London)* **287**, 396–401 (1980).
20. J.D. Linton and H.G.D. Niekus, in H.W. van Verseveld and J.A. Duine eds., *Microbial Growth on C_1 Compounds: Proceedings of the 5th International Symposium*, Martinus Nijhoff, Dordrecht, 1987, pp. 263–271.
21. S.H. Collins, in J.D. Stowell, A.J. Beardsmore, C.W. Keevil, and J.R. Woodward eds., *Carbon Substrates in Biotechnology*, IRL Press, Oxford, 1986, pp. 161–168.
22. J.-U. Ackermann and W. Babel, *J. Basic Microbiol.* **34**, 211–216 (1994).
23. T. Oki, A. Kitai, K. Kuono, and A. Ozaki, *J. Gen. Appl. Microbiol.* **19**, 79–83 (1973).
24. H. Motoyama, H. Anazawa, R. Katsumata, K. Araki, and S. Teshiba, *Biosci. Biotechnol. Biochem.* **57**, 82–87 (1993).

25. F.J. Schendel, C.E. Bremmon, M.C. Flickinger, M. Guettler, and R.S. Hanson, *Appl. Environ. Microbiol.* **56**, 463–470 (1990).
26. G.H. Lee, W. Hur, C.E. Bremmon, and M.C. Flickinger, *Biotechnol. Bioeng.* **49**, 639–653 (1996).
27. R.S. Hanson, R. Dillingham, P. Olson, G.H. Lee, D. Cue, F.J. Schendel, C.E. Bremmon, and M.C. Flickinger, in M.E. Lidstrom and F.R. Tabita eds., *Microbial Growth on C_1 Compounds*, Kluwer Academic, Dordrecht, 1996, pp. 227–236.
28. A. Abouzeid, G. Euverink, G.I. Hessels, R.A. Jensen, and L. Dijkhuizen, *Appl. Environ. Microbiol.* **61**, 1298–1302 (1995).
29. H. Yamada, S.S. Miyazaki, and Y. Izumi, *Agric. Biol. Chem.* **50**, 17–21 (1986).
30. P. Sirirote, T. Yamane, and S. Shimizu, *J. Ferment. Technol.* **64**, 389–396 (1986).
31. D.M. Anderson and H.-Y. Hsiao, in D. Rozell and F. Wagner eds., *Biocatalytic Production of Amino Acids and Derivatives*, Carl Hanser Verlag, Munich, 1992, pp. 22–41.
32. H. Yamada, Y. Morinaga, and Y. Tani, *Agric. Biol. Chem.* **46**, 47–55 (1982).
33. Y. Morinaga, Y. Tani, and H. Yamada, *Agric. Biol. Chem.* **46**, 473–480 (1982).
34. D. Byrom, *Trends Biotechnol.* **5**, 246–250 (1987).
35. Y. Poirer, C. Nawrath, and C. Somerville, *Biotechnology* **13**, 143–150 (1995).
36. S.Y. Lee, *Biotechnol. Bioeng.* **49**, 1–14 (1996).
37. T. Suzuki, T. Yamane, and S. Shimizu, *Appl. Microbiol. Biotechnol.* **23**, 322–329 (1986).
38. W. Babel, *FEMS Microbiol. Rev.* **103**, 141–148 (1992).
39. S. Fidler and D. Dennis, *FEMS Microbiol. Rev.* **103**, 231–236 (1992).
40. H. Dalton, in J.C. Murrell and H. Dalton eds., *Methane and Methanol Utilizers*, Plenum Press, New York, 1992, pp. 85–114.
41. B.J. Wallar and J.D. Lipscomb, *Chem. Rev.* **96**, 2625–2657 (1996).
42. J.C. Murrell, *FEMS Microbiol. Rev.* **88**, 233–248 (1992).
43. D. Jahng, and T.K. Wood, *Appl. Environ. Microbiol.* **60**, 2473–2482 (1994).
44. H. Dalton, *Adv. Appl. Microbiol.* **26**, 71–87 (1980).
45. K.J.A. Burrows, A. Cornish, D. Scott, and I.J. Higgins, *J. Gen. Microbiol.* **130**, 335–392 (1984).
46. C.T. Hou, in C.T. Hou ed., *Methylotrophs: Microbiology, Biochemistry, and Genetics*, CRC Press, Boca Raton, Fla., 1984, pp. 145–166.
47. S.H. Stanley, A.O.L. Richards, M. Suzuki, and H. Dalton, *Biocatalysis* **6**, 177–190 (1992).
48. D.J. Arp, *Curr. Opin. Biotechnol.* **6**, 352–358 (1995).
49. B.G. Fox, J.G. Borneman, L.P. Wackett, and J.D. Lipscomb, *Biochemistry* **29**, 6419–6427 (1990).
50. R. Oldenhuis and D.B. Janssen, in J.C. Murrell and D.P. Kelly eds., *Microbial Growth on C_1 Compounds*, Intercept Press, Andover, U.K., 1993, pp. 121–133.
51. H.L. Chang and L. Alvarez-Cohen, *Biotechnol. Bioeng.* **45**, 440–449 (1995).
52. H. Martin and J.C. Murrell, *FEMS Microbiol. Lett.* **127**, 243–248 (1995).
53. S.D. Prior and H. Dalton, *J. Gen. Microbiol.* **131**, 155–163 (1985).
54. L. Alvarez-Cohen, in J.C. Murrell and D.P. Kelly eds., *Microbial Growth on C_1 Compounds*, Intercept Press, Andover, U.K., 1993, pp. 337–350.
55. J.P. Bowman, L. Jimenez, I. Rosario, T.C. Hazen, and G.S. Sayler, *Appl. Environ. Microbiol.* **59**, 2380–2387 (1993).
56. I. Karube, T. Okada, and S. Suzuki, *Anal. Chim. Acta.* **135**, 61–67 (1982).
57. L.I. Sly, L.J. Bryant, J.M. Cox, and J.M. Anderson, *Appl. Microbiol. Biotechnol.* **39**, 400–404 (1993).
58. I.J. Higgins, W.J. Aston, D.J. Best, A.P.F. Turner, S.G. Jezequel, and H.A.O. Hill, in R.L. Crawford and R.S. Hanson eds., *Microbial Growth on C_1 Compounds: Proceedings of the 4th International Symposium*, ASM, Washington, D.C., 1984, pp. 297–305.
59. C.W. Bamforth and J.R. Quayle, *Biochem. J.* **169**, 677–686 (1978).
60. A. Geerlof, J. Vantol, J.A. Jongejan, and J.A. Duine, *Biosci. Biotechnol. Biochem.* **58**, 1028–1036 (1994).
61. T. Urakami, K. Yashima, H. Kobayashi, A. Yoshida, and C. Itoyoshida, *Appl. Environ. Microbiol.* **58**, 3970–3976 (1992).
62. C. Verduyn, J.P. van Dijken, and W.A. Scheffers, *Biotechnol. Bioeng.* **25**, 1049–1055 (1983).
63. Y. Sakai and Y. Tani, *Agric. Biol. Chem.* **50**, 2615–2620 (1986).
64. T.R. Hopkins and F. Muller, in H.W. van Verseveld and J.A. Duine eds., *Microbial Growth on C_1 Compounds: Proceedings of the 5th International Symposium*, Martinus Nijhoff, Dordrecht, 1987, pp. 150–157.
65. H. Schutte, J. Flossdorf, H. Sahm, and M.-R. Kula, *Eur. J. Biochem.* **62**, 151–160 (1976).
66. R. Wichmann, C. Wandrey, A. Buckmann, and M.-R. Kula, *Biotechnol. Bioeng.* **23**, 2789–2802 (1981).
67. R. Couderc and J. Baratti, *Biotechnol. Bioeng.* **22**, 1155–1173 (1980).
68. K.N. Faber, W. Harder, G. Ab, and M. Veenhuis, *Yeast* **11**, 1331–1344 (1995).
69. J.M. Cregg, T.S. Vedvick, and W.C. Raschke, *Biotechnology* **11**, 905–910 (1993).
70. C.A. Scorer, J.J. Clare, W.R. McCombie, M.A. Romanos, and K. Sreekrishna, *Biotechnology* **12**, 181–184 (1994).
71. A.R. Cross and C. Anthony, *Biochem. J.* **192**, 429–439 (1980).
72. B.D. Yoon, M. Ueno, and Y. Tani, *J. Ferment. Technol.* **65**, 629–634 (1987).
73. K. Inoue, S. Kageyama, K. Miki, T. Morinaga, Y. Kamagata, K. Nakamura, and E. Mikami, *J. Ferment. Bioeng.* **73**, 76–78 (1992).

MICROALGAE, MASS CULTURE METHODS

E. MOLINA GRIMA
University of Almería
Almería, Spain

KEY WORDS

Fluid dynamics
Gas–liquid mass transfer
Light distribution
Microalgae mass culture
Outdoor production
Photosynthetic efficiency
Tubular reactor

OUTLINE

INTRODUCTION

Microalgae have been cultured commercially for about 50 years for applications in secondary wastewater treatment (1) and use as animal feeds (2,3) as fertilizers (4), in production of chemicals (5,6), and in secondary metabolites with pharmaceutical potential (7), as well as other purposes. The early interest in using microalgae as a significant source of protein has waned because of the relatively low biomass productivities. Now the focus of algal mass cultures is on relatively specialized niche applications; nevertheless, progress remains slow relative to biotechnological use of bacteria, yeasts, and animal cells. Cost-effective culture of microalgae requires substantially improved biomass productivity, better utilization of the available light, and more efficient use of carbon dioxide, the principal carbon source.

Open outdoor mass culture systems, such as ponds and raceways, typically operate at low cell densities (0.01 to 0.06% w/v); consequently harvesting is expensive and economic returns are low. In addition, species control is difficult in open systems, and much of the incident radiation is wasted. The low productivity of open systems has prompted the development of enclosed photobioreactors (PBRs) (i.e., transparent tubes or containers in which the culture may be circulated by various devices). Enclosed PBRs allow a more controlled culture environment, including improved temperature control. Successful design of PBRs relies on attaining a proper balance of engineering and biological factors that are interdependent and depend also on the species being cultured. Of course, from a commercial perspective, reliability and the costs of construction and operation are also important. The physical rate processes of concern are the light delivery, the gas–liquid mass transfer, temperature control, fluid dynamics, and mixing. The biological rate processes of concern are photosynthesis, growth, and production of metabolites.

This article details the major types of bioreactors used in mass culture of microalgae. The basic requirements for successful outdoor culture are noted, and the physics and biology of mass cultivation are discussed.

BIOREACTORS FOR MICROALGAE MASS CULTURE

Three microalgae—*Chlorella*, *Dunaliella*, and *Spirulina*—are currently produced commercially on a relatively large scale (approximately 1,000 tons each per annum). The current commercial production systems for these species are mostly open raceway ponds. Many commercially useful algae are difficult to cultivate in open systems because they require a mild culture environment that is highly susceptible to contamination by other algae, bacteria, protozoa, and fungi. The inability to cultivate desirable algal species that contain valuable products (e.g., astaxanthin, aquaculture feeds, polysaccharides, and polyunsaturated fatty acids) in open ponds has increased interest in closed systems. The closed systems in use include (*1*) bag culture, which is widely used for culture of algae for aquaculture; (*2*) alveolar panels and other flat plate reactors of various designs (8–10); (*3*) stirred tank reactors with internal illumination (11,12); (*4*) airlift reactors (13,14); and (*5*) tubular reactors. Tubular reactors may be arranged in a horizontal position relative to the land or they may be inclined (15–22). Some of the major types of photobioreactors are discussed in the following.

Open Raceways

Open raceways, or shallow mixed ponds, for microalgal production were introduced in the 1950s and early 1960s by Oswald and coworkers (23). The principles of design and construction of shallow paddle-stirred raceways for microalgal culture have been reviewed by Dodd (24) and Oswald (25). The pond is set in a meandering configuration. Various designs are used to promote flow and low shear mixing. A simplified open raceway is depicted in Figure 1. Selection of a suitable bottom lining and construction of walls are important to the success of the open pond. The lining may be made of concrete or sheets of plastic and rubber material. Sheets of polyvinyl chloride 1–2 mm thick, are commonly used. The size of commercial ponds varies from 1,000 to 5,000 m^2; one or two paddle wheels per pond are used for stirring. A larger wheel (ca. 2 m in diameter) revolving slowly (e.g., 10 rpm) is preferable to a smaller, faster wheel, which could produce shear damage and foam.

Supply of Carbon Dioxide. The method of supplying carbon dioxide is an important aspect of the raceway design, particularly for species that grow near neutral pH. Several systems have been developed to efficiently supply CO_2 to shallow suspensions. In most cases the gas is supplied as fine bubbles. Because of the shallow depth of culture, the residence time of the bubbles is not sufficient for complete absorption of the CO_2; hence, some gas is lost to the atmosphere. To avoid this, Märkl and Mather (26) used a

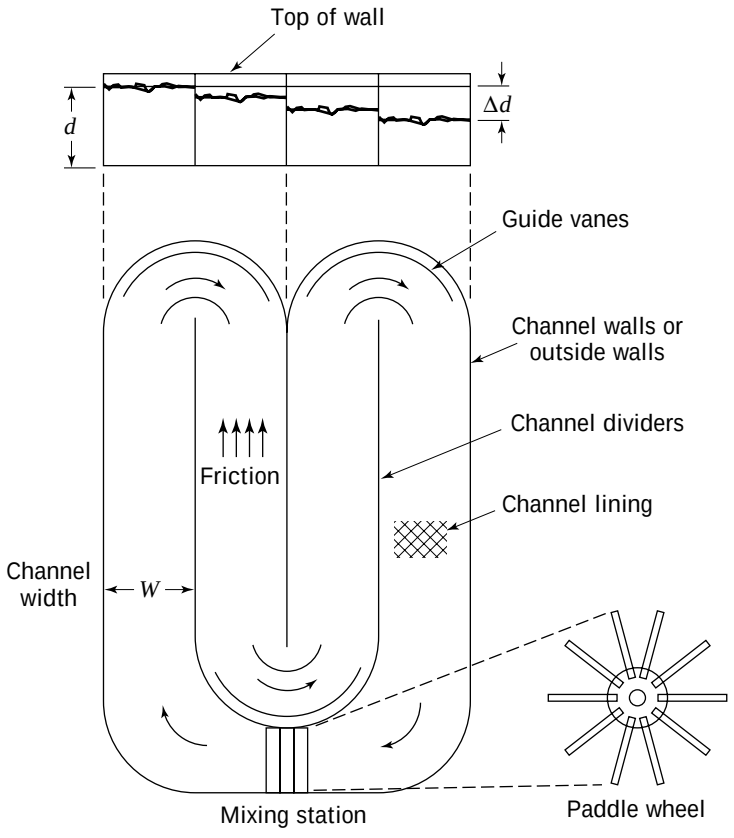


Figure 1. Open raceway with a paddle wheel mixing device. The channel length is the distance traveled by the culture from the discharge side of the paddle wheel to its upstream side. *Source*: Adapted from Oswald (25) and Dodd (24), with permission.

pump to push the algal suspension through a nozzle to a chamber in which the CO_2 was added. High CO_2 transfer rates were obtained. The jet produced an additional flow in the pond. The most effective method currently available for transferring CO_2 to algal cultures is countercurrent carbonation (25), in which the gas is injected as minute bubbles into a column of downcoming culture. The culture velocity is adjusted to that of the small bubbles of CO_2 rising against the current. Using this technique, Laws et al. (27) reported a 70% utilization of the supplied CO_2 compared with 13 to 20% if conventional bubbling was used (28).

Mixing. Mixing is necessary to prevent settling of cells and to avoid thermal stratification of the culture. Properly designed paddle wheels are by far the most efficient and durable pond mixers. The engineering of a raceway equipped with paddle wheels has been comprehensively described by Oswald (25). As shown in Figure 1, the depth, d, of culture in a channel of finite width, w, varies along the channel. The mean channel velocity, v(m s^{-1}) is given by

$$v = \frac{1}{n} R^{2/3} \cdot s^{1/2} \tag{1}$$

where n is the Manning friction factor (s m$^{-1/3}$), R is the hydraulic radius (m), and s is the rate of loss of energy in the channel per unit length. The channel length, L, that corresponds to the assumed change in depth for a given friction factor, hydraulic radius, and velocity (v) is given as (25)

$$L = \frac{\Delta d(dw/(w + 2d))^{4/3}}{v^2 \cdot n^2} \tag{2}$$

The value of n varies according to the relative roughness of the channel. Experimentally determined n values in algae growth channels vary from 0.008 to 0.030 (25). The lower value is for smooth, plastic-lined channels, and the higher one is for relatively rough earthen channels. The channel velocity, v, affects the paddle wheel's power requirements; hence,

$$P = \frac{Q \cdot \rho \cdot \Delta d}{\eta} \tag{3}$$

where P is the power (kW), Q is the volumetric flow rate of culture ($Q = v \cdot d \cdot w$) (m^3 s^{-1}), ρ is the density of culture (kg m^{-3}), Δd is the change in depth (m), and η is the efficiency of the paddle wheel. Because $\Delta d = s \cdot L$ and s is a function of v^2, P increases as the cube of velocity; therefore, velocity should be minimized whenever energy is a major cost factor. Typically, the flow rates range between 15 and 30 cm s^{-1}. Velocities greater than 30 cm s^{-1} result in large values of Δd in long channels and may require high channel walls and higher divider walls. Note that for a finite value of the channel width, w, the permissible mixing channel length, L, and thus the mixable area, $a = L \cdot w$, is strongly dependent on depth, d.

Contamination and temperature are difficult to control in open pond systems. Also, the culture depth cannot be reduced below 12 to 15 cm or a severe reduction of flow and turbulence would occur. Hence, areal volumes tend to be large, about 120 to 150 L/m^2 (29), requiring maintenance of a rather low cell concentration (around 500 mg/L). A low cell concentration in turn increases the cost of harvesting the biomass, and ponds are expensive to maintain. These problems could, in principle, be minimized by photobioreactor designs that employ shallower cultures, maximize cell densities, allow better process control, and minimize contamination. Closed photobioreactors achieve these objectives, as described in the next section.

Enclosed Photobioreactors

PBRs can be defined as algal culture systems that do not allow direct contact of the culture with the atmosphere. PBRs generally may be grouped into two classes: tubular and flat plate. These groups can be further classified into several categories, depending on orientation of tubes, the type of gas exchange system, and the arrangement of the individual growth units. The materials of construction provide additional variations. For example, in the case of flat plate reactors, most work has been carried out with alveolar systems in which the flat plate is divided into many individual cells. Another example is the use of flexible plastic bags. A major factor in the design of these systems is the depth of the culture, which ultimately determines the biomass concentration and productivity. Lee (30) and Borowitzka (31) provided comprehensive summaries of the state of the art of the closed PBRs, none of which are in widespread commercial operation.

Conventional mass culture ponds covered with inexpensive plastic sheets have been used in several algal mass culture projects, but they provide few benefits beyond an increased temperature during early spring or fall in temperate climates. Also, their design and operation are essentially identical to those of the conventional open pond systems described earlier. Here, only tubular reactors capable of operation outdoors with natural sunlight and with high surface-to-volume ratios (>10 m^{-1}) are considered.

In vertical column reactors, bubble columns or airlifts, aeration, and agitation are provided by injection of CO_2-enriched air through a sparger at the bottom of the column. In practice, a vertical column can extend for only a few meters or an expensive structure is needed for support. James and Al-Khars (13) studied the growth and productivity of *Chlorella* and *Nannochlropsis* in a vertical airlift photobioreactor. Contreras Gómez (14) showed that the average annual volumetric productivity of a draught tube airlift was about 1 g L^{-1} per day in culture of *Phaeodactylum tricornutum*. A major advantage of tower reactors is that oxygen does not accumulate in the growth medium because the residence time of the gas phase is quite small in a relatively short reactor. Also, tower bioreactors are potentially easy to keep sterile compared to other tubular systems. A major drawback of any tubular system positioned vertically is that in the summer the reactor is always at a large angle to the sun's rays; hence, the amount of solar energy absorbed during the midday hours is lower than for horizontally placed tubular devices.

The earliest described outdoor tubular reactor was made of 3-cm-diameter glass tubes arranged in a serpentine configuration (32). The tubular loop was immersed in water. The culture was recirculated in the system by a motor-driven pump. The flow velocity was 15 to 30 cm s^{-1}. The culture circulated through a bubble column that was injected with 5% CO_2 at a rate of 1–1.5 cm^3 min^{-1}. The total culture volume was 40 L. The more recent systems are essentially identical to that first design. Thus, Chaumont et al. in France (33) employed tubes placed parallel to ground and Pirt et al. in England (15) used a vertically placed loop. Both systems used pumps to circulate the culture. Because fragile cells can be sensitive to mechanical pumping, the authors recommended the use of airlift devices for both culture recirculation and CO_2 supply.

Although other designs of tubular reactors have been proposed—including the two-plane tubular reactor (16), the near-horizontal tubular reactor (21), the helical bubble reactor (34), the α-tubular reactor (35), and the parallel flow tubular photobioreactor (18)—they are essentially the same as the ones used by Pirt et al. (15) and Chaumont et al. (33).

Figure 2 shows a fully instrumented tubular photobioreactor similar to those described by Chaumont and coworkers. The system shown consists of a vertical external-loop airlift pump that drives the culture fluid through the horizontal tubular solar receiver. The airlift section (riser, downcomer, and degasser) has a height of 3.5 m. The gas-injected riser and the downcomer sections (Fig. 2) are extensions of the solar receiver tube. The solar receiver is made of transparent Plexiglas® tubes (0.05-m internal diameter, 0.005-m wall thickness) joined into a loop configuration by Plexiglas® joints to obtain a total horizontal length of 98.8 m. The solar receiver is submerged ($\sim$0.05 m) in a shallow pool of water that is maintained at 21 $\pm$ 2 °C by cooling with a heat pump when needed. The bottom and inside walls of the pool are painted white to improve reflectance. The surface area of the pool is 21.4 m^2. The total culture volume in the bioreactor is 0.200 m^3.

The exhaust gas from the top of the degasser zone is analyzed for oxygen (paramagnetic analyzer) and carbon dioxide (infrared analyzer). Filter sterilized culture medium is fed in at the degasser zone. A port located on the side of the degasser section is used to continuously harvest the culture (Fig. 2). The temperature, the dissolved oxygen, and the pH measurements are made via sensors located in the degassing zone, in addition to a dissolved oxygen sensor located at the end of the external loop. The sensors are connected to a control unit and a computer for data acquisition.

The air supply point is located at the entrance of the riser, whereas the CO_2 is injected at the entrance of the solar receiver in response to a signal from the pH sensor. The photosynthetically active radiation (PAR) is determined at the ground level and inside the photobioreactor. In the previously described microalgae mass culture facility, the irradiance inside the thermostatic pond was higher than the incident radiation on the surface of the water because of the light reflected by the walls and the bottom of the pool. This phenomenon of radiation enhancement is quantified in terms of the albedo of a system.

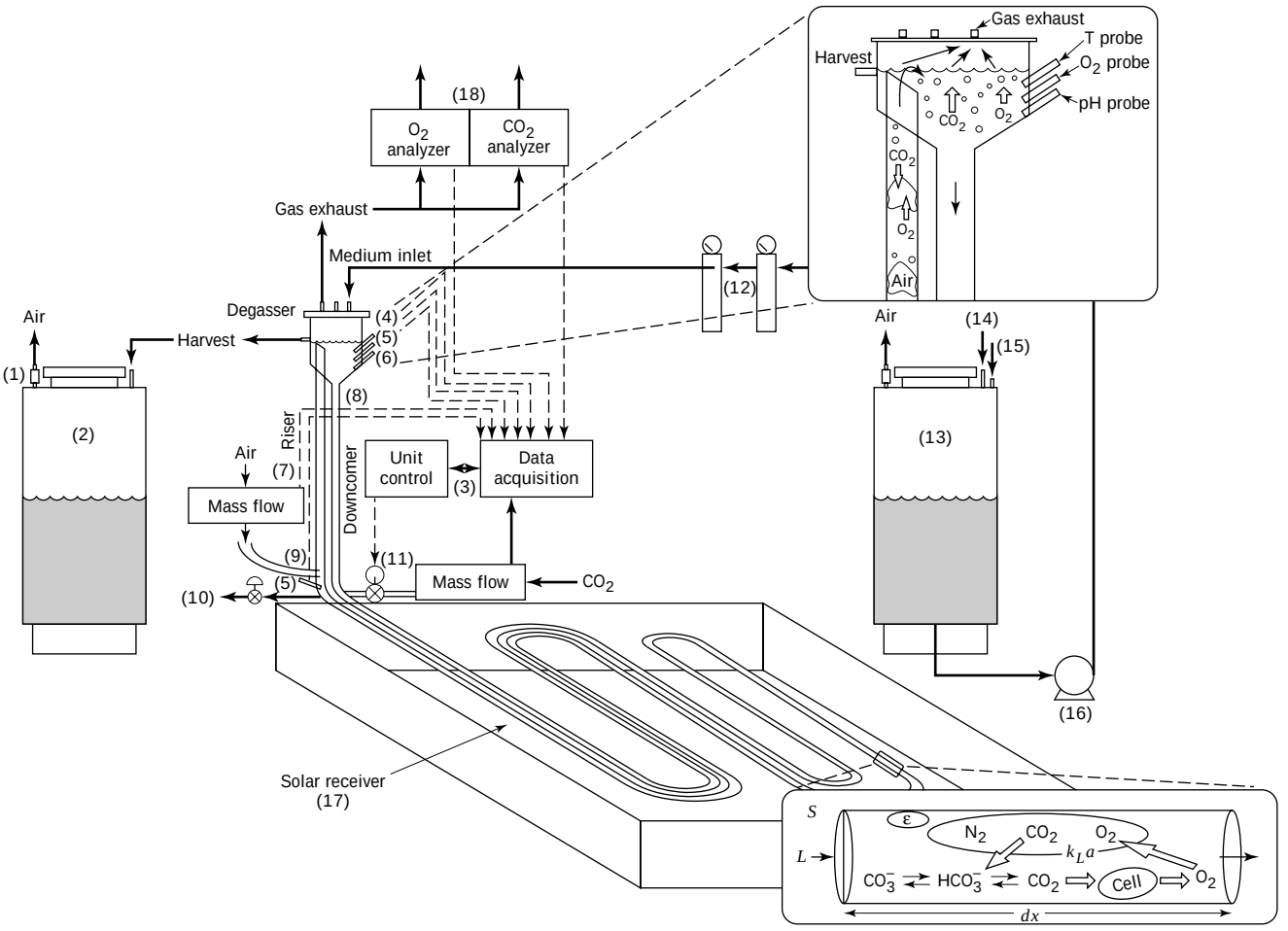


Figure 2. Tubular photobioreactor: (1) air filters, (2) harvest tank, (3) control unit, (4) temperature sensor, (5) dissolved oxygen probe, (6) pH probe, (7) riser, (8) downcomer, (9) air injection, (10) sampler, (11) CO_2 injection, (12) medium sterilization filters, (13) fresh medium, (14) nutrient inlet, (15) seawater inlet, (16) pump, (17) thermostated water pool, (18) gas composition analyzer.

For efficient and reliable large-scale culture of microalgae in PBRs, several operational, environmental, and design variables need to be assessed. A summary of some of the main variables and their impact on the various aspects of a reactor's performance are depicted in Figure 3. Of these, an important design criterion is the efficiency of light utilization, which in turn is related to the productivity of the system. Also important are the costs of construction and operation (36).

As indicated in Figure 3, the on-line data acquisition of dissolved oxygen, pH, and temperature inside the culture allows for rapid characterization of the physiological state of the cells. The measurement of CO_2 and O_2 molar flows at the gas inlet and outlet enables estimation of the mass transfer capability of the system and efficiency of CO_2 utilization. The measurement of the biomass concentration and the product content allows the estimation of the biomass and product productivities. Finally, the biomass productivity, P_b, and the photon flux absorbed per unit of volume, F_{vol}, may allow the assessment of the efficiency of a system in harvesting energy and converting it into biomass.

PHYSICAL FACTORS THAT INFLUENCE BIOMASS PRODUCTIVITY

Biomass productivity in any culture system depends on the degree to which the culture conditions match the requirements of the selected strain. Because mineral nutrient limitation is easily avoided in microalgal mass culture (29,37), light availability inside the photobioreactor and temperature are the main factors that determine productivity. Once the temperature is suitably controlled, light availability becomes the only limiting factor (22,38).

The interrelations among the various factors that influence productivity in outdoor conditions are shown in Figure 4. Productivity is determined by the growth rate, which, for fixed fluid dynamics and temperature, is a function of the light profile within the reactor and the light regime to which the cells are subject. The average irradiance (I_{av}) within the reactor depends on the external irradiance (I_0) on the surface of the reactor and its geometry (22,39,40). I_0 is a function of time, the geographic location, and environmental factors (41). Although temperature is relatively easily controlled, the available solar radiation at

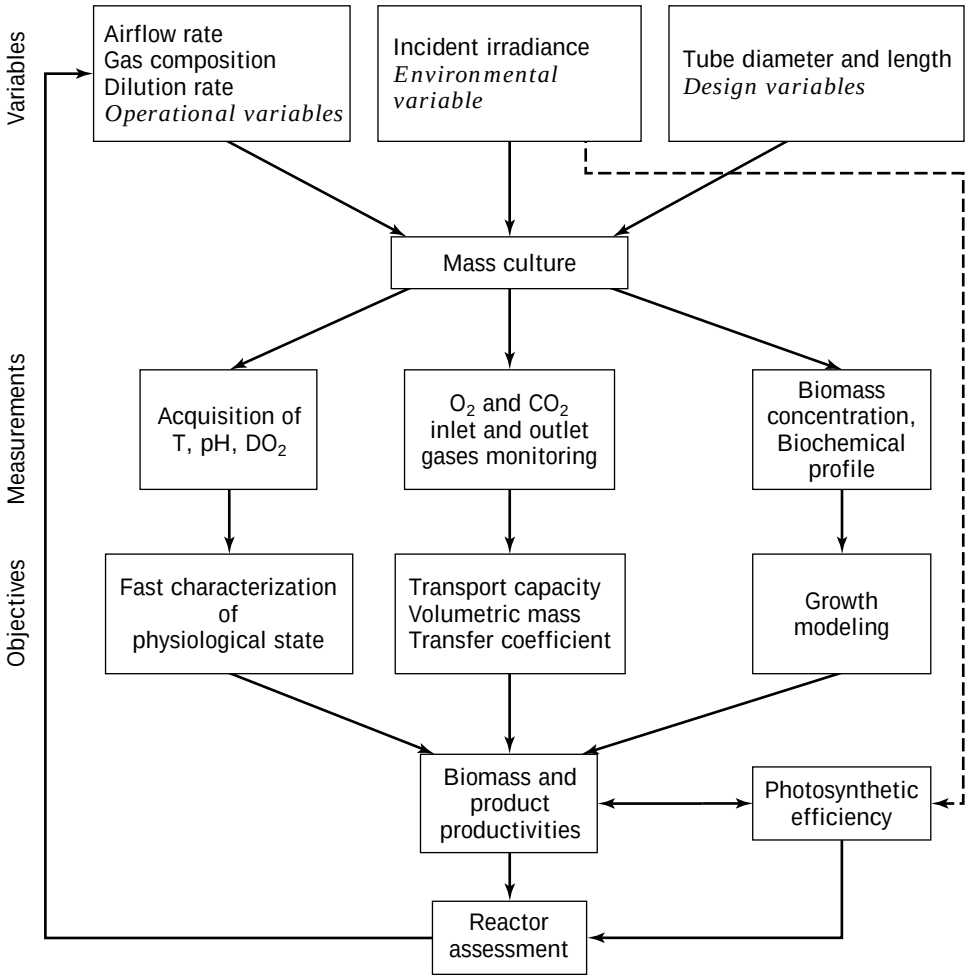


Figure 3. Relationships among bioreactor design and operational variables.

any location cannot be controlled. The self-shading effect of the cells also affects the radiation profile inside the culture (40). In addition, the fluid dynamics inside the reactor also influence the light regime seen by the cells because, depending on the nature and intensity of fluid movement, the cells spend different amounts of time in zones of different illumination (42–44). Fluid mechanics also affect the transport behavior and availability of certain nutrients. In an optimal system where no other factors limit, light availability determines the rate of photosynthesis and productivity. However, excessive light can be harmful and is known to produce a photoinhibitory response (45). The issues relating to light availability are discussed in the following sections for photobioreactors located outdoors.

Light Availability inside a Photobioreactor

For commercial-scale culture, sunlight is usually the preferred source of light because artificial light is too expensive for all but the most valuable products. The effects of natural outdoor light on cultures are not entirely clear because most studies on characterizing the photosynthetic response to light (e.g., 46–48) were conducted under carefully controlled indoor laboratory conditions. The irradiance levels used in such studies were significantly lower than the peak solar irradiance that typically occurs outdoors. Cultures grown outdoors experience large variations in light availability: Within a few hours the irradiance increases from 0 at dawn to more than 2,000 μE $m^{-2}s^{-1}$ at noon (i.e., a rate several times above saturation); within a few additional hours, the irradiance declines to 0 again as evening sets in. Because of this rapid change, the culture may go from being light limited in the morning to being photoinhibited at noon and light limited once again during the late afternoon and evening.

Information concerning light distribution in microalgal mass culture systems is scant, but recently a comprehensive model for light distribution and average irradiance inside a horizontal tubular outdoor photobioreactor has been proposed (40). That model has been further generalized for use with other configurations of photobioreactors (49). The mathematical model determines the distance traveled by an incident ray of light to any point inside the culture and estimates the local irradiance by taking into account the light attenuation due to biomass. The incident total radiation is divided into direct and disperse radiation. The former is characterized by having a direction defined as a function of the position of the sun at any given mo-

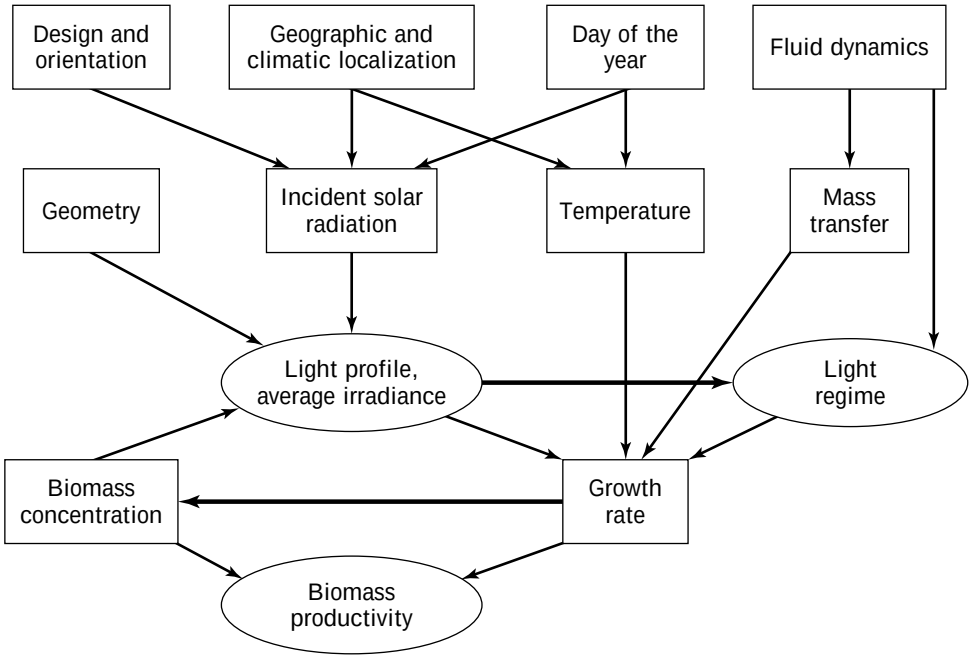


Figure 4. Relationships among parameters that influence biomass productivity in outdoor culture.

ment, and the latter, being reflected radiation, has no specific direction. The model is mathematically complex. For example, the angle of incidence of direct radiation on the photobioreactor surface is a function of five variables: day of year, solar hour, geographic latitude, surface slope, and surface azimuth angle (i.e., the deviation of the projection on a horizontal plane of the normal to the surface from the local meridian, with zero due south, east negative, and west positive) (50,51). Detailed information for estimating the angle of the incidence of the solar radiation, the irradiance impinging on the reactor surface, and the path traveled by the ray of light inside both horizontally and vertically placed tubular reactors has been reported (40,49,52).

Local and Average Irradiance inside the Culture. In a well-mixed microalgal mass culture system, light attenuation by biomass gives rise to a heterogeneous illumination profile inside the culture bulk for which mathematical evaluation is essential to estimate the average irradiance on which the growth of the microalgae depends (53). This attenuation is usually represented by Lambert–Beer's law. The incident radiation on the culture surface may be considered as consisting of the sum of two contributions: direct radiation, which has a direction defined by the position of the sun, and the disperse radiation. The latter is the sum of the diffuse radiation and that reflected by the surface surrounding the reactor.

The local direct irradiance inside the culture may be estimated using Lambert–Beer's law, as described elsewhere (40). The disperse radiation coming from all spatial directions can be determined by an integration over all the angles of incidence (40). Thus, it is possible to determine the light profiles and average irradiance inside the culture. Figure 5 shows the modeled local irradiance at two radial positions in a horizontal or vertical arrangement at 8 and 12 solar hours. As expected, as solar hour changes, curves of local irradiance at the same internal ratio adjust to the sun's position. In the morning (8 solar hours) in the vertical reactor, the side facing the sun had much higher direct irradiance values than the opposite side; this difference was lower in the horizontal reactor. However, at noon, the irradiance distribution in the vertical reactor, unlike the horizontal tube, was practically homogeneous. This profile results because during midday, the contribution of disperse irradiance to the total irradiance is very large in the vertical arrangement. Therefore, the level of irradiance for each internal radius and for any hour of the day is higher in the horizontally placed reactor tube than in the vertically placed reactor tube. The irradiance distribution inside the culture and the patterns of cell movement inside the photobioreactors are also different.

The Light Saturation Constant and the Photoinhibition Phenomenon

Algal growth is related to light intensity. According to Goldman (54), the generalized curve relating algal growth to light intensity has the shape shown in Figure 6. The curve has four main features: (*1*) at some low growth rate, the growth is balanced by decay, and the net growth rate is zero; (*2*) the initial slope of the curve represents the maximum efficiency of growth in response to light; (*3*) there is a saturating light intensity at I_s for which $\mu = \mu_{max}$; (*4*) at some light intensity $I_p > I_s$, the growth is inhibited and $\mu < \mu_{max}$.

Because most algal species become light saturated at a fraction of the peak solar light intensity, much potentially useful solar energy is essentially wasted. The light satu-

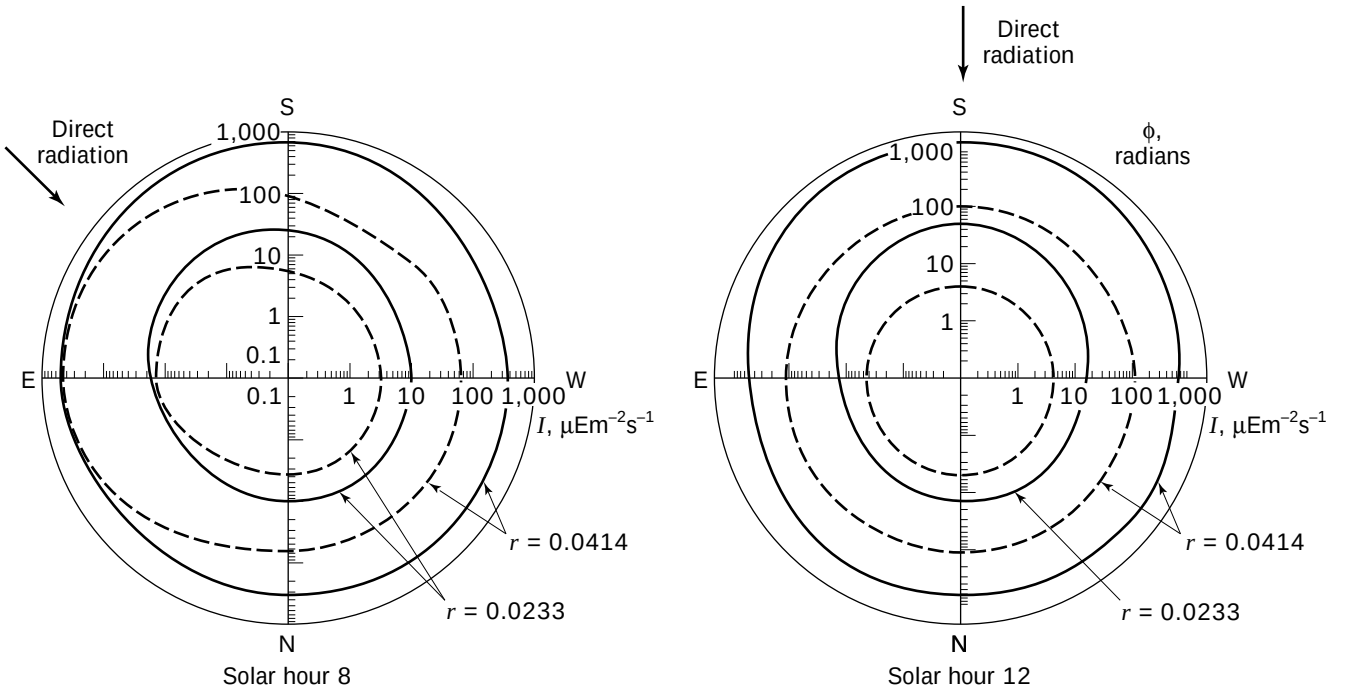


Figure 5. Calculated radial irradiance profiles in the cross section of a bubble column (dashed lines) and a photobioreactor placed parallel to the ground (solid lines) at 8 and 12 solar hours. In both cases the parameters were 0.1-m tube diameter; 1.9 g L^{-1} biomass concentration of *Phaeodactylum tricornutum*; dilution rate of 0.033 h^{-1}; and pigment content, $X_p = 2.21\%$.

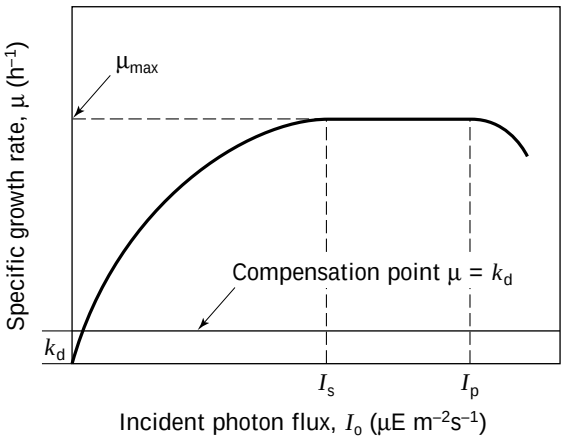


Figure 6. Dependence of the growth rate (μ) on the incident photon flux (I_0). *Source*: From Goldman (55), with permission.

ration constant is a measure of the intrinsic capacity of cells to utilize light energy. A high saturation constant is advantageous because it increases the efficiency of light use and reduces photoinhibition.

Photoinhibition—a decline in the photosynthetic capacity due to light intensities in excess of those required to saturate the photosynthesis (29)—has been studied often in algae. Shading the culture with shad nets is commonly used to protect against inhibitory levels of light during midday hours. The kinetics of photoinhibition and its recovery in *Porphyridium cruentum* have been studied (56). A linear relationship was found between the specific rate of photoinhibition and the specific light absorption rate. The temperature was reported to have a pronounced modifying effect on photoinhibition. At 15 °C, no damage was detected at 2,300 μE $m^{-2}s^{-1}$ light intensity, even after 45 min of illumination. In contrast, at 35 °C, 84% inhibition of the photosynthetic activity was observed within 10 min of exposure. *P. cruentum* cells recovered readily when transferred to low light (90 μE $m^{-2}s^{-1}$) or darkness; the specific rate of recovery was independent of the light intensity to which the cells had been exposed during the inhibitory treatment.

The existence of photoinhibition in continuous outdoor cultures of *Phaeodactylum tricornutum* has also been documented (52). Figure 7 shows how the average irradiance, I_{av}, inside the culture varies with dilution rate, D, regardless of the tube diameter, indicating that the cultures adapt mainly to the average irradiance. At $D = 0.025\ h^{-1}$, the I_{av} increased only slightly with increasing external irradiance, I_0, on the reactor surface. The mean value of I_{av} for the whole range was $84 \pm 9\ \mu E\ m^2\ s^{-1}$. With a smaller tube diameter or higher external irradiance, the biomass concentration increased as expected for self-shaded light-limited growth. However, the slight increase in I_{av} with increasing I_0 points to some photoinhibition by external irradiance, although this is not very extensive due to the high biomass concentrations. At $D = 0.040\ h^{-1}$, I_{av} values are higher ($163 \pm 20\ \mu E\ m^{-2}\ s^{-1}$) than at $D = 0.025\ h^{-1}$ because of the lower steady-state biomass concentrations. The greater light availability gives rise to a higher specific growth rate and photoinhibition because of the greater in-

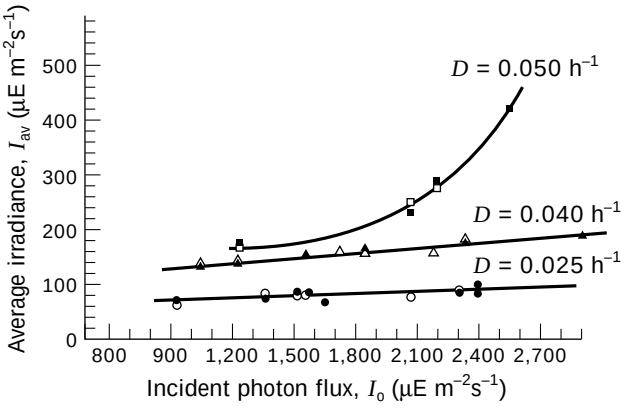


Figure 7. Variation of the average irradiance (I_{av}) inside a tubular photobioreactor during culture of *Phaeodactylum tricornutum* UTEX 640. Data are shown as functions of the mean daily photosynthetic irradiance (I_0) on the reactor surface for various dilution rates and two different tube diameters. Open symbols are for 3-cm tube diameter; filled symbols are for 6-cm tube diameter.

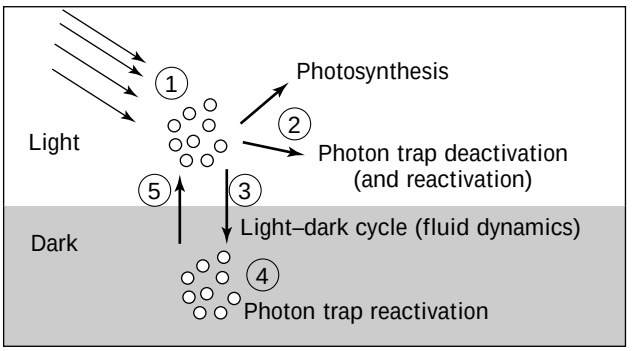


Figure 8. A simplified mechanistic representation of photosynthesis. (1) photon capture → biomass synthesis; (2) photon capture → photon trap deactivation (reactivation); (3) cell transfer to the dark; (4) photon trap reactivation; (5) cells return to the illuminated zone. Photon trap recovery is zero order with respect to unrecovered photon trap, and the recovery time is species dependent. *Source*: From Merchuk et al. (59), with permission.

crease in I_{av} with I_0. At 0.050 h^{-1} dilution rate, the exponential increase in I_{av} with I_0 caused by the low steady-state biomass concentration underlines the existence of photoinhibition; hence, the one-to-one relationship between μ and I_{av} no longer applies.

Photoinhibition is caused by oversaturation of photosystem II, which damages the D1 protein that carries the binding site of the electron carrier (57,58). This photoinhibition effect is quite distinct from that of temperature rise that occurs in uncontrolled systems as a function of the photon flux density (PFD). Whereas a single cell of microalgae cannot be simultaneously photolimited and photoinhibited, in bioreactors photolimited and photoinhibited cell populations may coexist because of variations in photointensity in different zones. At low light intensities the few damaged D1 protein molecules are replaced rapidly and the net damage to the photosynthetic apparatus is negligible. In this situation, a dark period reduces growth rate (photosynthesis) because fewer photons are captured but no gain is obtained from the dark time. In contrast, under conditions of intense illumination, part of the light energy impairs the photosynthetic apparatus. Repair and damage proceed simultaneously, and the observed growth is the sum of the two processes. If a dark period is introduced in this situation, the duration of the photosynthetic period declines, but the damaging period is also reduced; the photon trap repair continues during the dark time. Consequently, during the next light period, a substantially rejuvenated photon trap compensates for the loss in photosynthetic time. Under these circumstances, alternating light and dark periods do not reduce growth, which may in fact be slightly enhanced. Nevertheless, the length of the dark period is important: lengthening the light period beyond an optimal value will produce loss of growth. The optimal or critical dark period is not a fixed quantity; instead it depends on the PFD of the previous light period and the fluid residence time in zones of different irradiance. Figure 8 outlines a mechanism of photosynthesis that takes into account the steps noted in the foregoing (59). Clearly, therefore, the principal problem of designing or choosing a photobioreactor is ensuring, for any species with preestablished photosynthetic characteristics, that the largest possible fraction of cells experiences optimal exposure to light in the largest possible reactor volume.

Biomass Productivity

Many studies on the influence of environmental variables on biomass productivity have been carried out in laboratories (20,32,60,61), but few have been performed under outdoor conditions. Some models have been proposed specifically for outdoor algal ponds (41,62–64). Thus, Incropera and Thomas (41) present a model for determining the mean irradiance in an open pond as a function of climate and geographic location parameters. Starting from estimated irradiance values, a six-parameter growth model is proposed (41). The model does not consider photoinhibition and predicts maximum productivities of about 34 g m^{-2} per day. A five-parameter mathematical model for biomass productivity in outdoor systems has been proposed by Sukenik et al. (62,63). According to this model, the cells adapt to mean irradiance inside the culture; however, photoinhibition is not taken into account, and the results are not valid at low biomass concentration. The model yields productivities of 10.2 to 27.2 g m^{-2} per day. By using mean temperature and irradiance data, Guterman et al. (64) propose a growth model for open systems. The model estimates the variation of culture parameters such as dissolved oxygen, temperature, pH, and biomass concentration through empirical linear relationships and does not consider photoinhibition. The estimated maximum productivity was 38 g m^{-2} per day at irradiances of 2,500 μE m^{-2}s^{-1}. According to these models the maximum productivities are reached at low biomass concentration, when the mutual shading effect is negligible and the culture may be considered light saturated. The estimated productivities from these models have not been experimentally verified. In fact, such cultures cannot be carried out, because

photoinhibition in the middle of the day decreases growth rate and may even lead to culture washout.

A macromodel for estimating the yearlong biomass productivity of microalgal cultures in outdoor tubular photobioreactors has been developed by Acién Fernández et al. (52). In this model the solar irradiance impinging on the reactor surface I_0 is initially determined as a function of the day of the year and the location. Then, by taking into account the geometry of the system, the average irradiance to which the cells are exposed inside the culture, I_{av}, is determined. Finally, the specific growth rate, μ, is correlated with I_{av} as follows:

$$\mu = \frac{\mu_{max} \cdot I_{\alpha v}^{(n_2 + n_3/I_0)}}{\left(I'_k\left(1 + \left(\frac{I_0}{K_I}\right)^{n_1}\right)\right)^{(n_2 + n_3/I_0)} + I_{\alpha v}^{(n_2 + n_3/I_0)}} \quad (4)$$

where I_0 is the irradiance on the surface of the reactor. The characteristic parameters of this equation were obtained by fitting the experimental results to the equation. The parameter values were $\mu_{max} = 0.063\ h^{-1}$, $I'_k = 94.3\ \mu E\ m^{-2}\ s^{-1}$, $K_I = 3426\ \mu E\ m^{-2}\ s^{-1}$, $n_1 = 3.04$, $n_2 = 1.209$, and $n_3 = 514.6$. Equation 4 has been used to simulate the biomass productivity in a continuous outdoor culture of *Phaeodactylum tricornutum* (52). The measured data agreed with the simulations within 20%. The simulated biomass productivity over the annual cycle is shown in Figure 9. The approach described can in principle be extended to other strains. If photoinhibition were not taken into account as in earlier models (41,62–64) the experimental results could be fit to a hyperbolic model that considers only photolimitation (61):

$$\mu = \frac{\mu_{max} I_{\alpha v}^n}{I_k^n + I_{\alpha v}^n} \quad (5)$$

however, the scatter of the data would significantly exceed ±20%. In equation 5 I_k is the affinity cells have for light and n is a fitting parameter of the model.

In view of the observations discussed, several conclusions can be made regarding biomass production by microalgal mass cultures: (*1*) the optimum dilution rate varies over the year as a function of incident radiation and the diameter of the bioreactor tube; (*2*) the larger the tube diameter, the more extensive the photoinhibition effect, although the simultaneous existence of photolimitation and photoinhibition permits growth to be maintained and a quasi steady state is still attained.

Fluid Dynamics and Mixing

To achieve high productivity, algal cultures need to be mixed. The method of mixing (turbulence) is critical because some algae are quite shear sensitive (65,66). When light is the limiting factor for growth, stirring has been suggested as the most practical means by which to attempt to distribute solar energy evenly to all cells in the culture. The stirring is associated with a flashing light effect. The duration for which the cells are exposed to the light and dark zones of the reactor is a function of agitation intensity. The flashing light effect has been intensively investigated by Laws et al. (67). The flashing light effect or agitation has been associated with increased photosynthetic efficiencies. Laws and coworkers (67) introduced a simple technique for utilizing the flashing light effect in shallow but optically dense outdoor algal cultures by inserting an array of small foil segments across the width of a continuous

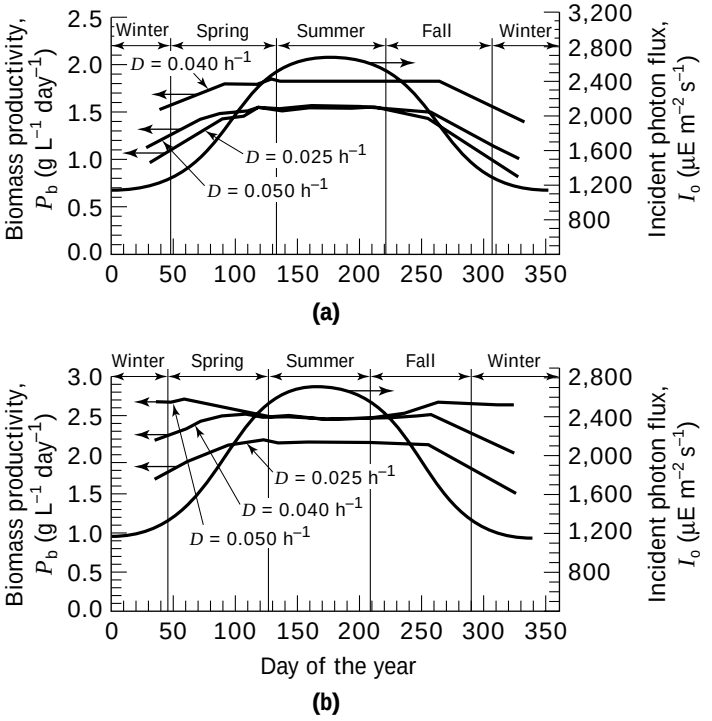


Figure 9. Variation of the estimated biomass productivity with the day of the year for a continuous culture of *Phaeodactylum tricornutum* UTEX 640 in two tubular photobioreactors: (**a**) 0.06-m tube diameter, (**b**) 0.03-m tube diameter. Simulations are shown for three dilution rates.

flume at intervals of 1.2 m along the length of the flume. The wing-shaped foils created a vortex circulation that produced an ordered pattern of vertical stirring. By using these devices, Laws and coworkers obtained photosynthetic efficiencies (based on the PAR) of 9.6 and 7.5, respectively, with and without the foils in place.

In tubular photobioreactors and ponds, higher flow rates generally enhance productivity. Typically, flow rates range between 20 and 30 cm s^{-1}. As noted, the flow can be generated by various methods such as paddle wheels and pumps. In closed tubular photobioreactors the flow is commonly driven by airlift systems, which are generally less damaging to cells than the mechanical devices. The velocity in the tube controls the turbulence and shear rate, determines the mixing behavior and the mass transfer capacity, and greatly influences the axial oxygen concentration profile in the tube. The liquid velocity, U_L, may be estimated by an extension of the well-known and widely tested model developed by Chisti (68); thus,

$$U_L = \left[\frac{g h_D \varepsilon_r}{2f \dfrac{L_{eq}}{D}} \frac{A_r}{A_d}\right]^{0.5} \quad (6)$$

where f is the Fanning friction factor that can be calculated using the Blasius equation,

$$f = 0.0791 \mathrm{Re}^{-0.25} \quad (7)$$

where Re is the Reynolds number, or $(\rho_L U_L D)/\mu$, D is the tube diameter, and ρ_L and μ are the density and the viscosity of the liquid, respectively. As shown in equation 6 the superficial liquid velocity in a tubular photobioreactor is a function of gas holdup in the riser, ϵ_r, the height of the dispersion, h_D, and the ratio between the cross-sectional area of the riser and the downcomer.

Gas–Liquid Mass Transfer

In addition to availability of light, the availability of carbon dioxide (the carbon source) affects the productivity of a photosynthetic culture. The cells combine carbon dioxide and water to generate carbohydrates and cell biomass. Photosynthesis generates oxygen, which is released into the extracellular fluid. Because large amounts of oxygen can inhibit photosynthesis, oxygen must be removed to prevent buildup. Similarly, carbon dioxide must be provided such that availability of carbon does not become a limiting factor. Removing oxygen and supplying carbon dioxide are problems of gas–liquid mass transfer. Therefore, modeling of an algal reactor must consider gas–liquid mass transfer and how it is affected by hydrodynamics.

The mass transfer from the gas phase to the culture suspension in open ponds has been extensively studied (26,69). According to Märkl (70), the critical CO_2 concentration in an algal suspension of *Chlorella vulgaris* in an open pond is very low. A concentration as low as 10^{-6} M assures unlimited photosynthesis. Experiments with the same strain and with different oxygen concentrations in the medium show that the photosynthetic reaction rate is enhanced by 14% when there is almost no oxygen in the algal suspension and reduced by 35% when the medium is saturated with pure oxygen. In technical systems an oxygen content no higher than the saturation concentration with air (21%) is desirable (26). Mathematical expressions for longitudinal profiles of CO_2 and dissolved oxygen in algal cultivation units have been derived by Livansky et al. (71). The mass transfer issues relating to tubular photobioreactors have been treated by Camacho Rubio et al. (72).

Because carbon dioxide tends to be expensive one consideration is minimizing its loss; hence, quantitative transfer of CO_2 from the gas to the liquid phase is aimed at but normally not achieved. The dissolved carbon dioxide is in equilibrium with carbonate and bicarbonate species. These equilibria are pH dependent. Loss of dissolved carbon dioxide due to uptake into algal cells is partly compensated for by regeneration from carbonates and bicarbonates. Consequently, carbon dioxide uptake is accompanied by changes in pH. The relevant equilibrium and the corresponding equilibrium constants (73) that need to be considered are

$$H_2O \leftrightarrow H^+ + OH^- \quad K_w = [OH^-] \cdot [H^+] = 10^{-14} \quad (8)$$

$$CO_2 + H_2O \leftrightarrow H_2CO_3 \leftrightarrow HCO_3^- + H^+ \quad K_1 = [HCO_3^-] \cdot [H^+]/[CO_2] = 10^{-6.381} \quad (9)$$

$$HCO_3^- \leftrightarrow CO_3^{2-} + H^+ \quad K_2 = [CO_3^{2-}] \cdot [H^+]/[HCO_3^-] = 10^{-10.377} \quad (10)$$

As noted (72), models based on mass balances and gas–liquid mass transfer phenomena can be used to predict the various culture variables such as dissolved oxygen concentration, the carbon content within the culture, the composition of the outlet gas, the carbon dioxide losses and requirements, and the pH of the culture. As shown in Figure 10, through any 24-hour period, the simulated data closely agree with the measurements.

Temperature

Many studies in both open ponds and closed reactors have shown that temperature is an important factor affecting productivity. The highest biomass growth rate can be obtained only at the optimal temperature for growth. Although this effect is obvious and well documented under laboratory conditions, the magnitude of the effect on the annual production of biomass outdoors seems not be sufficiently appreciated. The diurnal variation in temperature to which cultures in open raceways are exposed results in temperature-limited growth during a significant fraction of the day. Lee et al. (48) showed that the light yield (g biomass kJ^{-1}) increased some three- to fivefold with increasing temperature in cultures of *Chlorella*, an effect that may be readily observed with other microalgae. The temperature of the culture at night is yet another important factor that affects the net output rate of biomass. The effect of temperature on dark respiration is, of course, very well documented (16,52,53).

Compared with open ponds, enclosed reactors generally warm up faster in the morning, thus reaching the optimum temperature for algal photosynthesis earlier in the day

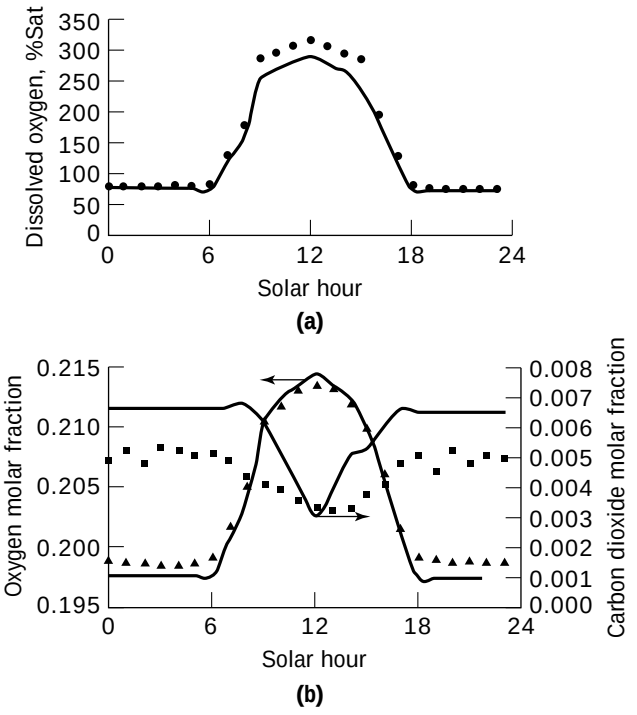


Figure 10. Daily variation of (**a**) the culture dissolved oxygen in the photobioreactor solar receiver tube (●) and (**b**) concentrations of O_2 and the CO_2 in the gas phase coming out of the photobioreactor (■, oxygen molar fraction; ▲, carbon dioxide molar fraction). The continuous lines represent simulations. Data are for a culture of *Porphyridium cruentum*.

(38). Closed reactors, however, can also overheat and thus require some sort of temperature control systems. In closed systems this control is achieved by either incorporating a heat exchanger into the design or evaporative cooling of water sprayed on the surface of the reactor. Economic considerations generally favor evaporative cooling.

PHOTOSYNTHETIC EFFICIENCY

The photosynthetic efficiency (PE) of biomass growth is defined as the energy stored in biomass per unit of light energy absorbed (74). This parameter is of fundamental importance in biomass production, because the microalgal production should aim to achieve the maximum attainable PE so as to produce the most biomass possible per unit land area (open ponds) or volume (enclosed systems). According to Richmond (74), microalgae use only the PAR, which constitutes about 43% of the total solar radiation at the Earth's surface and has an energy input equivalent to that of monochromatic light at a wavelength of 575 nm. Because one Einstein of 575-nm light contains 49.74 kcal of energy, and assuming that free energy of photosynthesis is 114 kcal mol^{-1} of reduced CO_2, the theoretical maximum energy efficiency for the photosynthetic reduction of CO_2 to glucose with white light is $114/(8 \cdot 49.74) = 0.286$, or approximately 29% (75). This estimation assumes that eight photons of light are needed for the requisite transfer of four electrons. Under field conditions, much lower photosynthetic efficiencies have been obtained. The maximum practical photosynthetic efficiency is a problem still to be resolved (76); however, a theoretical photosynthetic yield of 6 to 7% of the total solar energy seems possible (77). A 6–7% efficiency represents a biomass productivity of 35 to 70 g m^{-2} per day depending on the geographic locations. A photosynthetic efficiency of 10% during a period of 122 days has been reported (78). Nevertheless, producing biomass on large scale under outdoor conditions with an efficiency similar to that obtained under controlled laboratory conditions remains a challenge.

Information on algal yields reported over 30 years has been summarized by Goldman (54). Although the early attempts at mass culture produced no better than 10 to 15 g dry wt biomass m^{-2} per day over short periods, today it is not uncommon for sustained yields to reach 20 to 25 g m^{-2} per day. Goldman presented a thorough analysis of maximum yield potential of algae based on the thesis that algal production under the proper conditions is limited only by light (55). Goldman's equation can be written as

$$P_b = \frac{1}{Q_b} \cdot I_s\left(\ln \frac{0.45 \cdot I_0}{I_s} + 1\right) \qquad (11)$$

where P_b is the productivity of the system (g m^{-2} s^{-1}), I_0 is the total irradiance at the culture surface (μE m^{-2} s^{-1}), I_s is the saturating light intensity for which the growth rate is the maximum (i.e., when $\mu = \mu_m$, see Fig. 6) (μE m^{-2} s^{-1}), and Q_b is the heat of combustion of the biomass (μE g^{-1}). The maximum available intensity of sunlight at the Earth's surface is about 2,000 μE m^{-2} s^{-1}. Equation 11 predicts that, for a range of I_s values between 60 and 180 μE m^{-2} s^{-1}, as is typical for microalgae culture, the upper limit of algal yield is 30 to 60 g dry wt m^{-2} per day. According to this model, the factor that most strongly influences the algal yield is I_s. By contrast with the laboratory experiments, most outdoor algal production systems report productivities in the range of 15–30 g m^{-2} per day (79). These figures are far different from those observed in carefully controlled indoor cultures.

The laboratory work reported previously was carried out under conditions of low intensities, at or below saturation for photosynthesis. In practice, when dealing with sunlight, the major problem is not the maximal quantum efficiency but the fraction of absorbed photons that can actually be used in photosynthesis. The problem, well described by Benemann (80), is that if algal photosynthesis saturates at low light intensities, and algal cells have a high extinction coefficient (high pigment content), then more light will be absorbed than can be used in photosynthesis. Depending on the actual saturating and incident light intensities, the extinction coefficient, K_a, and the biomass concentration, the maximum efficiency (regardless of the minimum quantum requirement) will be reduced by about two- to fivefold from that possible at low light levels (80). The only solution to this problem is to increase the light intensity at which photosynthesis saturates, which means reducing the antenna and accessory pigment level of the cell without affecting the rest of the photosynthetic apparatus.

A new procedure for calculating yields in microalgal mass cultures at high incident irradiances has been proposed by Molina Grima et al. (81). This approach is based on an averaged irradiance for assessing the photon flux absorbed by the biomass. The starting point of the procedure (81) is the definition of the quantum yield, Ψ_E. The quantum yield is defined as the amount of biomass generated per unit of radiation (usually photons or Einsteins) absorbed by the culture. Because it represents the ratio of biomass generation to absorbed photon flux, the quantum yield can be calculated by the expression

$$\Psi_E = \frac{P_b}{F_{vol}} \tag{12}$$

where P_b is the volumetric biomass productivity and F_{vol} is the photon flux absorbed per unit volume. The quantum yield can be converted to energy units, commonly denoted as Ψ_{kJ}, by taking into account the average energy of the light used (kJ E^{-1}). The bioenergetic yield, Ψ, quantifies the percentage of light energy that is converted to chemical energy; Ψ can be calculated as the product of Ψ_{kJ} and the biomass heat of combustion, Q_b:

$$\Psi = \Psi_{kJ} Q_b \tag{13}$$

Because P_b in steady-state continuous culture can be calculated as the product of biomass concentration, C_b, and the dilution rate, D, any of the growth yield defined in equations 12 and 13 can be evaluated as long as the absorbed photon flux can be found.

Absorbed Photon Flux

The quantification of the absorbed photon flux is a complicated matter, especially when the flux must be evaluated in geometries other than flat reactors. Many factors such as light scattering effects may lead to a misevaluation of the absorbed flux. In every case the flux needs to be evaluated by direct measurement of the light leaving the photobioreactor (82). The light absorbed can be evaluated by integrating the local volumetric rate of energy absorption over the total reactor volume (83). This integral can be readily obtained from the extinction coefficient of biomass, K_a, and the average irradiance, I_{av}.

$$F_{vol} = I_{\alpha v} K_a C_b \tag{14}$$

Calculation of the absorbed photon flux by the preceding procedure has several advantages: the calculation is independent of the system geometry once I_{av} is known and can thus be applied to any type of photobioreactor as long as I_{av} can be determined. Photon flux losses due to reflection of light off the reactor walls that cannot be neglected (82) are automatically taken into account if the irradiance is measured inside the culture. The validity of equation 14 for assessing the absorbed photon flux has been checked experimentally in a culture of *Isochrysis galbana* (81) and using data published by Lee and Erickson (84) for *Chlorella* in a flat photobioreactor. *Isochrysis galbana* showed a maximum efficiency of light conversion into biomass of 0.6 g E^{-1} at $I_0 = 820$ μE $m^{-2}s^{-1}$ and a dilution rate of 0.030 h^{-1}. The maximum capacity of the biomass to metabolize light was found to be 13.1 μE $g^{-1}s^{-1}$ (81). Above this value a significant drop in system efficiency occurred.

OPERATIONAL CONSIDERATIONS

Successful maintenance of a microalgal culture facility requires continuous monitoring to prevent the development of nonoptimal conditions. As shown in Figure 3, the on-line information typically monitored includes the dissolved oxygen (DO), temperature, and pH. In addition, daily microscopic observations are important for characterizing the physiological state of the culture. Measurement of growth should be routinely accompanied by detailed microscopic evaluations. Any increase in number of contaminating organisms in the culture is a warning signal: the temperature may have deviated from the optimal, or the concentration for a particular nutrient may have gone out of the optimal range. Monitoring of the nitrogen content is another useful indicator of the state of continuous cultures and may give information about the possible depletion in mineral nutrient in the culture. This information can be used as a guideline for adding, in equivalent amounts, the entire formula of the elements making up the growth medium. The concentration of DO is another reliable and sensitive indicator of conditions that relate to growth and productivity. There are many indications that early detection of an inexplicable decrease in DO concentration or a decline in the normal rate of the daily increase of DO in outdoor mass culture serves as a reliable warning signal that the culture is stressed and may quickly deteriorate (74).

Reliability of the microalgal mass culture systems and the ability to easily clean them are other important operational considerations. Closed reactors may experience fouling of the walls by biomass. For continuous cleaning of tubular photobioreactors, Chaumont et al. (85) have developed and patented a system that consists of balls that continuously circulate through the tubes. Such systems are effective for fouling control.

Tubular photobioreactors are also reasonably reliable. They have been used to continuously culture many species including the marine algae *Phaeodactylum tricornutum*, *Tetraselmis chuii*, and *Isochrysis* sp. for periods in excess of 3 months (19,53). These systems are particularly suited to continuous or semicontinuous culture, which permits superior control over the culture conditions and provides a product of consistent quality. As an additional operational advantage, the cell concentration reached is much greater in most closed systems than values achieved in open systems; consequently, the harvesting cost is lower for closed systems. However, the actual cell density needs to be managed carefully to maximize productivity (Fig. 11). At low cell densities, especially in outdoor systems, the algae may be severely photoinhibited, whereas at very high cell densities self-shading can reduce productivity as individual algal cells do not receive an optimal level of light.

In summary, culture management aims to achieve and sustain an optimal state of the culture. For maximum biomass productivity, solar irradiance should be the sole lim-

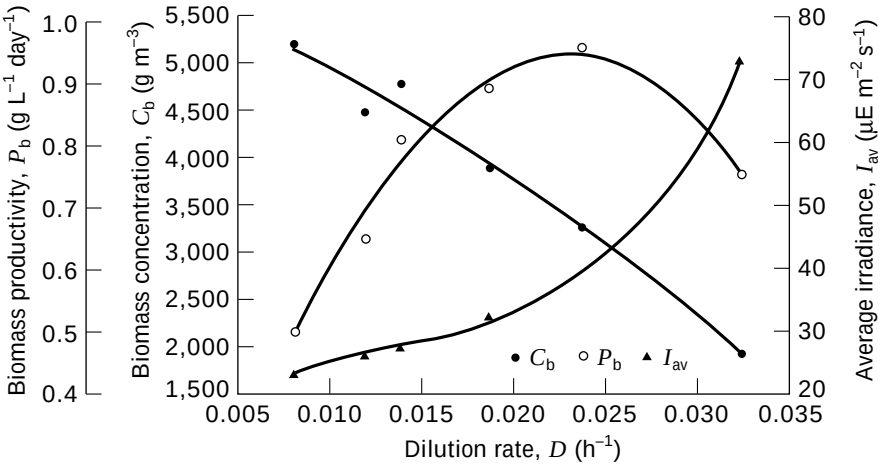


Figure 11. Effect of dilution rate, D, on the steady-state biomass concentration, C_b, the biomass productivity, P_b, and the average irradiance inside a photobioreactor, I_{av}. Data were obtained in an outdoor continuous culture of *P. tricornutum*, in a concentric tube airlift bioreactor. *Source*: From Refs. 14 and 49, with permission.

itation; however, when a metabolite is the desired product, the operational conditions for maximal productivity might be different from those for biomass productivity.

CONCLUDING REMARKS

Microalgae represent an important but little known resource with the genetic potential to provide valuable compounds. Future development of algal biotechnology will depend largely on what unique high-value products microalgae can provide. Algal products or processes must have some unique attribute, either biochemical or engineering, that would make them superior to heterotrophic microbes. Current commercial use of microalgae is restricted to a few niche applications.

Based on a comparison of the major types of algal culture systems (Table 1), the open systems are clearly deficient in several aspects. Despite this deficiency, open raceway systems are in commercial use. A principal advantage of open culture systems is the relatively small capital investment needed. Open systems are simple, but their productivity is low relative to the theoretical maximum. Much improvement in photosynthetic activity could be effected by genetic engineering of the photosynthetic complex to raise the light intensity threshold at which the photosynthetic machinery is light saturated. Also, resistance to photoinhibition may be introduced by genetic modifications.

Enclosed photobioreactors, which allow better control of operational conditions, are better suited to producing high-value products. These bioreactors permit culture of a greater diversity of algae than possible in open ponds. The potential of enclosed outdoor tubular photobioreactors has been demonstrated with high-productivity culture of microalgal species such as *Porphyridium*, *Isochrysis*, and *Phaeodactylum*. Because of better control of culture parameters, enclosed culture allows for manipulation of the biochemical composition of the algal biomass for various purposes (86–89). Although the enclosed culture technology is more capital intensive than the open pond systems, the additional cost may be recoverable. Temperature increase and fouling are the major problems with closed photobioreactors. The choice of a suitable mass culture system depends on several factors including geographic location, the microalgal species, and the metabolite to be produced. The following factors should be considered in making a selection: *(1)* the culture system should permit the culture of several species; *(2)* the culture system should provide uniform illumination and high rates of mass transfer that are attained by means that do not damage the culture; *(3)* the reactor must prevent or minimize fouling, particularly of

Table 1. Comparison of Properties of Different Large-Scale Algal Culture Systems

Reactor type	Mixing	Light utilization efficiency	Temperature control	Gas transfer	Hydrodynamic stress	Species control	Sterility
Unstirred ponds	Very poor	Poor	None	Poor	Very low	Difficult	None
Open raceway	Fair to good	Fair to good	None	poor	Low	Difficult	None
Airlift	Uniform	Good	Excellent	High	Low	Easy	Easily achievable
Bubble column	Fair	Fair	Excellent	Fair	Low	Easy	Easily achievable
Tubular	Almost uniform	Good	Excellent	High	Low to high	Easy	Achievable
Biocoil	Uniform	Excellent	Excellent	Low to high	Low to high	Easy	Achievable

Source: Adapted from Borowitzka (31).

its light transmitting surfaces; and *(4)* the volume of the nonilluminated parts of the reactors should be minimized.

NOMENCLATURE

A_d	Cross-sectional area of the downcomer (m^2)
A_r	Cross-sectional area of the riser (m^2)
C_b	Biomass concentration (g L^{-1} or g m^{-3})
d	Depth channel (m)
f	Fanning friction factor
F_{vol}	Photon flux absorbed in unit volume (μE m^{-3} s^{-1})
h_D	Height at which the degasser is situated (m)
I	Hourly incident photosynthetic radiation (μE m^{-2} s^{-1})
I_{av}	Photosynthetically active hourly average irradiance inside culture (μE m^{-2} s^{-1})
I_0	Solar irradiance impinging on the reactor surface (μE m^{-2} s^{-1})
I_s	Saturating light intensity for which the growth rate is maximum (μE m^{-2} s^{-1})
K_a	Absorption coefficient (m^2 g^{-1})
K_1	First equilibrium constant for bicarbonates buffer (mol m^{-3})
K_2	Second equilibrium constant for bicarbonates buffer (mol m^{-3})
K_W	Dissociation constant of water
L	Channel length (m)
L_{eq}	Equivalent length of the solar receiver (m)
n	Manning friction factor (s $m^{-1/3}$), or characteristic parameter of the hyperbolic growth model
n_1	Characteristic parameter of the photolimitation–photoinhibition model
n_2	Characteristic parameter of the photolimitation–photoinhibition model
n_3	Characteristic parameter of the photolimitation–photoinhibition model
P	Power consumption for the culture flow in the channel (kW)
P_b	Biomass productivity (g m^{-3} h^{-1})
Q_b	Heat of combustion of biomass (μE g^{-1})
S	Cross section of the tube (m^2)
s	Rate of loss of energy in the channel (dimensionless)
U_L	Superficial liquid velocity in the tube (m s^{-1})
v	Mean culture flow velocity channel (m s^{-1})
w	Width channel (m)
ψ	Bioenergetic yield coefficient (%)
ψ_E	Quantum yield (g E^{-1})
ψ_{kJ}	Energetic yield (g kJ^{-1})
η	Efficiency of the paddle wheel (−)

ACKNOWLEDGMENTS

I wish to acknowledge the contribution of all my colleagues of the Marine Biotechnology Group of the University of Almería, who have worked with me on this subject. Thanks are also due to Prof. Y. Chisti for helpful comments and assistance with the final manuscript. Some of the work reported was financially supported by the Plan Andaluz de Investigación (Junta de Andalucía), Comisión Interministerial de Ciencia y Technología (C.I.C.Y.T.), Spain, and the European Union.

BIBLIOGRAPHY

1. W.J. Oswald, in M.A. Borowitzka and L.J. Borowitzka eds., *Micro-Algal Biotechnology*, Cambridge Univ. Press, Cambridge, 1988, pp. 305–328.
2. W.N. De Pauw and G. Persoone, in M.A. Borowitzka and L.J. Borowitzka eds., *Micro-Algal Biotechnology*, Cambridge Univ. Press, Cambridge, 1988, pp. 197–221.
3. J.R. Benemann, *J. Appl. Phycol.* **4**, 233–245 (1992).
4. W.B. Metting, in M.A. Borowitzka and L.J. Borowitzka eds., *Micro-Algal Biotechnology*, Cambridge Univ. Press, Cambridge, 1988, pp. 288–304.
5. D.J. Chapman and K.W. Gollenbeck, in R.C. Cresswell, T.A.V. Rees, and N. Shah eds., *Algal and Cyanobacteria Biotechnology*, Longman, New York, 1989, pp. 1–27.
6. D.M. Calvin and S.E. Taylor, in R.C. Cresswell, T.A.V. Rees, and N. Shah eds., *Algal and Cyanobacteria Biotechnology*, Longman, New York, 1989, pp. 137–160.
7. D.K.W. Glombitza and M. Koch, in R.C. Cresswell, T.A.V. Rees, and N. Shah eds., *Algal and Cyanobacteria Biotechnology*, Longman, New York, 1989, pp. 161–238.
8. M.R. Tredici, P. Carlozzi, G.C. Zitelli, and R. Materassi, *Bioresour. Technol.* **38**, 153–159 (1991).
9. A. Ramos de Ortega and J.C. Roux, *Biomass* **10**, 141–156 (1986).
10. I.A.J. Ratchford and H.J. Fallowfield, *J. Appl. Phycol.* **4**, 1–9 (1992).
11. P. Pohl, M. Kohlhase, and M. Martin, in T. Stadler, J. Mollion, M.C. Karamanos, H. Morvan, and D. Christiaen eds., *Algal Biotechnology*, Elsevier, London, 1988, pp. 209–217.
12. A.K. Hilaly, M.N. Karim, and D. Guyre, *Biotechnol. Bioeng.* **43**, 314–320 (1994).
13. C.M. James and A.M. Al-Khars, *Aquaculture* **87**, 381–393 (1987).
14. A. Contreras Gómez, Ph.D. Thesis, University of Almería, Almería, Spain, 1996, p. 373.
15. S.J. Pirt, Y.K. Lee, M.R. Walach, M.W. Pirt, H.H.M. Balyuzi, and M.J. Bazin, *J. Chem. Technol. Biotechnol.* **33B**, 35–58 (1983).
16. G. Torzillo, P. Carlozzi, B. Pushparaj, E. Montaini, and R. Materassi, *Biotechnol. Bioeng.* **42**, 891–898 (1993).
17. French Pat. 9,115,773 (June 25, 1991), D. Chaumont, P.F. Dos Santos, C. Gudin, D. Assie, and G. Chaintron (to Ussie Inginierie).
18. A. Richmond, S. Boussiba, A. Vonshak, and R. Kopel, *J. Appl. Phycol.* **5**, 327–332 (1993).
19. T. Chrismadha and M.A. Borowitzka, *J. Appl. Phycol.* **6**, 67–74 (1994).
20. E. Molina Grima, F. García Camacho, J.A. Sánchez Pérez, J. Urda Cardona, and J. Fernández Sevilla, *Biotechnol. Appl. Biochem.* **20**, 279–290 (1994).
21. M.R. Tredici and G.C. Zittelli, *Biotechnol. Bioeng.* **57**, 187–197 (1998).
22. Y.K. Lee and C.S. Low, *Biotechnol. Bioeng.* **38**, 995–1000 (1991).

23. W.J. Oswald, in E.F. Gloyna, J.F. Malina, and E.M. Davis eds., *Ponds as Wastewater Treatment Alternative*, Water Resources Symp. No. 9, University of Texas, Austin, Tex 1975.
24. J.C. Dodd, in A. Richmond ed., *Handbook of Microalgal Mass Culture*, CRC Press, Boca Raton, Fla., 1986, pp. 265–283.
25. W.J. Oswald, in M.A. Borowitzka and L.J. Borowitzka eds., *Micro-Algal Biotechnology*, Cambridge Univ. Press, Cambridge, 1988, pp. 357–394.
26. H. Märkl and M. Mather, *Arch. Hydrobiol.* **20**, 85–93 (1985).
27. E.A. Laws, S. Taguchi, J. Hirata, and L. Pang, *Biotechnol. Bioeng.* **28**, 191–197 (1986).
28. P. Heusler, J. Castillo, S. Morino, and V. Vasquez, *Arch. Hydrobiol.* **11**, 254–260 (1978).
29. A. Richmond, in F.E. Round and D.J. Chapman eds., *Progress in Physiological Research*, vol. 7, Biopress, Bristol, 1990, pp. 269–330.
30. Y.K. Lee, *Trends Biotechnol.* **4**, 186–189 (1986).
31. M.A. Borowitzka, *J. Mar. Biotechnol.* **4**, 185–191 (1996).
32. H. Tamiya, E. Hase, K. Shibata, A. Mituya, T. Iwamura, T. Nihei, and T. Sasa, in J.S. Burlew ed, *Algal Culture from Laboratory to Pilot Plant*, Carnegie Institution of Washington, Washington, D.C., 1953, pp. 204–232.
33. D. Chaumont, C. Therpenier, and C. Gudin, in T. Stadler, J. Mollion, M.C. Karamanos, H. Morvan, and D. Christiaen eds., *Algal Biotechnology*, Elsevier, London, 1988, pp. 199–208.
34. European Pat. 0,239,272 (March 6, 1988), L.F. Robinson; A.W. Morrison, and M.R. Bamforth (to Biotechna Ltd.).
35. Y.K. Lee and C.S. Low, *Biotechnol. Bioeng.* **40**, 1119–1122 (1992).
36. M.A. Borowitzka, *J. Appl. Phycol.* **4**, 267–269 (1992).
37. A.D. Little, in J.S. Burlew ed, *Algal Culture from Laboratory to Pilot Plant*, Carnegie Institution of Washington, Washington, D.C., 1953, pp. 235–273.
38. M.R. Tredici and R. Materassi *J. Appl. Phycol.* **4**, 221–231 (1992).
39. H. Quiang, H. Gutterman, and A. Richmond, *Biotechnol. Bioeng.* **51**, 51–60 (1996).
40. F.G. Acién Fernández, F. García Camacho, J.A. Sánchez Pérez, J.M. Fernández Sevilla, and E. Molina Grima, *Biotechnol. Bioeng.* **55**, 701–714 (1997).
41. F.P. Incropera and J.F. Thomas, *Solar Energy* **20**, 157–165 (1978).
42. J.N. Philliphs and J. Myers, *Plant Physiol.* **29**, 152–161 (1954).
43. J.U. Grobellar, *J. Appl. Phycol.* **6**, 331–335 (1994).
44. K.L. Terry, *Biotechnol. Bioeng.* **28**, 988–995 (1986).
45. A. Vonshak, *7th International Conference International Association of Applied Algology: Opportunities from Microalgae*, Knysna, South Africa, April 16–19, 1996, pp. 76–77.
46. R.H. Foy and C.E. Gibson, *Br. Phycol. J.* **17**, 183–193 (1982).
47. L. Van Liere and L.R. Mur, *J. Gen. Microbiol.* **115**, 153–160 (1975).
48. Y.K. Lee, H.M. Tan, and C.S. Hew, *Biotechnol. Bioeng.* **27**, 555–561 (1985).
49. F.G. García Camacho, A. Contreras Gómez, F.G. Acién Fernández, and E. Molina Grima, *Enzyme Microb Technol.* **24**, 164–172 (1999).
50. B.Y.H. Liu and R.C. Jordan, *Solar Energy* **7**, 53–65 (1960).
51. J.A. Duffie and W.A. Beckman, *Solar Engineering of Thermal Processes*, Wiley, New York, 1980, p. 762.
52. F.G. Acién Fernández, F. García Camacho, J.A. Sánchez Perez, J.M. Fernández Sevilla, and E. Molina Grima, *Biotechnol. Bioeng.* **58**, 505–611 (1998).
53. E.M. Molina Grima, F. García Camacho, J.A. Sánchez Pérez, J.A. Fernández Sevilla, F.G. Acién Fernández, and A. Contreras Gómez, *J. Chem. Technol. Biotechnol.* **61**, 167–173 (1994).
54. J.C. Goldman, *Water Res.* **13**, 60–118 (1979).
55. J.C. Goldman, *Water Res.* **13**, 119–160 (1979).
56. Y.K. Lee and A. Vonshak, *Arch. Microbiol.* **150**, 529–533 (1988).
57. J.A.A. Barber, *Trends Biochem. Sci.* **17**, 61–66 (1992).
58. S. Jensen and G. Knutsen, *J. Appl. Phycol.* **5**, 495–504 (1993).
59. J. Merchuk, M. Ronen, S. Giris, S. Arad (Malis), *Biotechnol. Bioeng.* **59**, 705–713 (1998).
60. T.T. Bannister, *Limnol. Oceanogr.* **24**, 79–96 (1979).
61. E. Molina Grima, J.A. Sánchez Pérez, F. García Camacho, J.L. García Sánchez, and D. López Alonso, *Appl. Microbiol. Biotechnol.* **38**, 599–605 (1993).
62. A. Sukenik, P.G. Falkowski, and J. Bennett, *Biotechnol. Bioeng.* **30**, 970–977 (1986).
63. A. Sukenik, R.S. Levy, Y. Levy, P.G. Falkowski, and Z. Dubinsky, *J. Appl. Phycol.* **3**, 191–201 (1991).
64. H. Guterman, A. Vonshak, and S. Ben-Yaakov, *Biotechnol. Bioeng.* **35**, 809–819 (1990).
65. H.J. Silva, T. Cortinas, and R.J. Ertola, *J. Chem. Technol. Biotechnol.* **40**, 41–49 (1987).
66. C. Gudin and D. Chaumont, *Bioresour. Technol.* **38**, 145–151 (1991).
67. E.A. Laws, K.L. Terry, J. Wickman, and M.S. Chalup, *Biotechnol. Bioeng.* **25**, 2319–2335 (1983).
68. Y. Chisti, *Airlift Bioreactors*, Elsevier, London, 1989, p. 343.
69. K. Livansky, B. Prokes, F. Dittrt, and V Benes, *Biotechnol. Bioeng. Symp.* **4**, 513–518 (1973).
70. H. Märkl, *Biotechnol. Bioeng.* **19**, 1851–1862 (1973).
71. K. Livansky, M. Kayan, and P.S. Pilarski, *Algological Studies* **70**, 97–119 (1993).
72. F. Camacho Rubio, F.G. Acién Fernández, F. García Camacho, J.A. Sánchez Pérez, and E. Molina Grima, *Biotechnol. Bioeng.* **62**, 71–86 (1999).
73. K. Livansky, *Algological Studies* **58**, 87–97 (1990).
74. A. Richmond, *Crit. Rev. Microbiol.* **4**, 369–438 (1986).
75. J.A. Bassham, in A. Mitsui, S. Miyachi, A.S. Pietro, and S. Tamura eds., *Biological Solar Energy Conversion*, Academic Press, New York, 1971, pp. 151–165.
76. S.J. Pirt, *Biotechnol. Bioeng.* **25**, 1915–1922 (1983).
77. D.O. Hall, in A. Richmond ed., *Handbook of Microalgal Mass Culture*, CRC Press, Boca Raton, Fla., 1986, pp. 1–24.
78. E.A. Laws, S. Taguchi, J. Hirata, and L. Pang, *Biomass* **16**, 19–32 (1988).
79. J. Weisman, R.P. Goebel, and J. Benemann, *Biotechnol. Bioeng.* **31**, 336–344 (1988).
80. J.R. Benemann, in R.C. Cresswell, T.A.V. Rees, and N. Shah eds., *Algal and Cyanobacteria Biotechnology*, Longman, New York, 1989, pp. 317–337.
81. E.M. Molina Grima, F. García Camacho, J.A. Sánchez Pérez, J.G. Acién Fernández, and J. Fernández Sevilla, *Enzyme Microb. Technol.* **21**, 375–381 (1997).
82. S. Aiba, *Biotechnol. Bioeng.* **25**, 2775–2776 (1983).
83. O.M. Alfano, R.L. Romero, and A.E. Cassano, *Chem. Eng. Sci.* **41**, 421–444 (1986).

84. H.Y. Lee and L.E. Erickson, *Biotechnol. Bioeng.* **29**, 476–481 (1987).
85. French Pat. 9,103,781 (June 25, 1991), D. Chaumont, P. Ferreira dos Santos, and L. Sauze (to Ussie Inginierie).
86. P.S. O'Colla, in R.A. Lewin ed., *Physiology and Biochemistry of Algae*, Academic Press, New York, 1962, pp. 337–356.
87. P. Pohl, in O. Zaborsky ed., *Handbook of Biosolar Resources*, CRC Press, Boca Raton, Fla., 1982, pp. 383–404.
88. C. Gudin and C. Therpenier, *Adv. Biotechnol. Process* **6**, 73–110 (1986).
89. E. Molina Grima, J.M. Fernández Sevilla, J.A. Sánchez Pérez, and F. García Camacho, *J. Biotechnol.* **45**, 59–69 (1996).

MICROBIAL GROWTH. See BIOENERGETICS OF MICROBIAL GROWTH.

MICROBIAL GROWTH MEASUREMENT, METHODS

ARTHUR L. KOCH
Indiana University
Bloomington, Indiana

KEY WORDS

Balanced growth
Biomass estimation
Cell differentiation
Flow cytometry
Particle counts
Phases of the growth cycle
Sampling head gas
Turbidimetry
Unbalanced growth
Viable counts

OUTLINE

INTRODUCTION

In biotechnology, the denominator is biomass, for example, yield is the desired product per unit of biomass. Biomass concentration is one of the major determinants to identify the stage of the culture cycle. It is necessary to understand the controlling points in the physiology of the bacterial growth. Biotechnologists have talked extensively about growth phase and idiophase as the key measures of fermentation technology. There are a number of approaches for biomass measurement, but they are not very good and have significant limitations. The reader should supplement the treatment presented here with a previous article (1) that details a number of procedures and leads back to the earlier literature. Another useful overview is provided by Sonnleitner et al. (2). The former paper comes from the approach of a physiologist, and the latter is from the point of view of biotechnologists. For historical interest, the collected source is by Dawson (3).

Principles

Principles of bacterial growth are discussed in microbial physiology texts. I would particularly recommend that by White (4). Although there are many books on bioprocessing, most do little to discuss biomass measurements. The book by Pirt in 1975 (5) was a milestone in the field of microbe growth for product formation; newer publications are now available (6,7), but they do little for the estimation of biomass.

Definitions of Growth

To measure biomass and bacterial growth rate, precise definitions are needed. Although biomass is fundamental to carrying out technology, we have to consider first the growth process that makes the biomass and examine its

several definitions. The most basic definition of growth is based on the ability of individual cells to multiply, that is, to repetitively initiate and complete cell and organismal division. This definition implies monitoring the increase in total number of discrete bacterial particles. There are three basic ways to do this: by microscopic enumeration of the particles, by electronic enumeration of the particles passing through an orifice (Coulter counter) and by modern flow cytometry. Assessment of particle number would falsely include dead cells and detritus, which would tend to lower estimations of growth rate. The rate would be artificially raised by the progressive dissolution of aggregates of bacteria and the fragmentation of nongrowing filamentous organisms. An increase in cell number need not correlate exactly with an increase in biomass or useful product. Commonly, at the end of an exponential growth phase, cell division overshoots biomass production and the cells become smaller.

A second definition of growth involves determining the increase in colony-forming units (CFU). Because some cells may be dead or dying, this definition of growth may be different from the one based on the detection of discrete particles as a function of time. Although in the long run, the increase in the number of organisms capable of indefinite growth is the only important consideration for the physiologist, this is not so for the biotechnologist, for several reasons. First, for strain purity, each new production run starts afresh from starter cultures; these cultures would had been specially treated much differently than the cells in the actual production run. This care is to avoid contamination and the buildup of unwanted mutants. Second, dead, dying cells and stationary cells may be the productive members of the culture in terms of product formation. This second definition is the reason that colony counting and most-probable-number (MPN) methods of measurement are so important. It must be noted that viable counting methods, which seem so natural to a bacteriologist, are really quite special in that cultures are diluted so highly that individual organisms cannot interact. For example, these methods cannot in principle be applied to obligate sexually reproducing organisms requiring male–female interaction or to colonial organisms such as myxobacteria that in certain conditions need to be part of a large mass of organisms that produces sufficient exoenzymes in order to grow. Even when applied to the prokaryote, there are special restrictions and limitations, for example, CO_2 must be available in sufficient concentration, although this need not be supplied if many organisms are present generating CO_2. This and possible detoxification of the medium can lead to a lower count than would have occurred in the more concentrated actual culture. To illustrate, consider a single bacterium and an isolated mammal, both introduced into an environment without any individuals of their own kind. The former would produce a colony, whereas the latter could only produce a viable colony if the founding individual was a pregnant female.

A third definition of growth (and the practical one) is based on an increase in biomass. Macromolecular synthesis and increased capability for synthesis of cell components are the obvious basis for the measurement of growth by everyone interested in microbes and what they can do. From our point of view, the whole process of chromosome replication and cell division is an essential but minor process that seldom limits growth; what usually limits growth is the ability of enzymatic systems to use available resources to form biomass. Moreover, the restriction of growth of the culture is usually the result of the depletion of resources or degradation of the local environment.

A fourth definition of growth is based on the organisms' action in chemically changing their environments. This category can be considered as simply a different consequence of the increase in biomass. But it allows the rate of the growth process and the biomass production to be estimated in indirect ways. Most important in industry, the gases coming from the fermentation are analyzed, and the consumption of oxygen and production of carbon dioxide are used as effective and efficient means to follow growth. Especially important is the use of mass spectrometer methods, because of their precision and the ability to automate measurements, growth of cultures can be followed usefully and effectively.

Biomass is the result of growth; it has variously different definitions according to purpose. In most biotechnology, biomass is organic material that is cheap and can be converted into fuel, or heat, or structural materials. It is necessary that such sources be self-generating, cheap, and available. Whether biomass is alive or dead is less relevant; the assumption is that it was, or at least once was, alive. If one has a mixture of coal and fresh plant material, the term biomass is fully ambiguous. For product formation there is less ambiguity; microbial biomass requires the expenditure of resources, and useful product may be produced by organisms that may be living or dead according to which of the definitions of growth is used.

Balanced Growth

These four definitions of growth become equivalent under a single circumstance: the state of balanced growth (8). This is certainly an academic concept, but its practicality is clear. An asynchronous culture can be said to be in balanced growth when all intensive properties are constant in time. An intensive property is one that does not matter how big the system is. An extensive property is another term from physical chemistry and refers to those properties of the system that are proportional to the amounts of substances of various kinds in the system. Thus, biomass and cell number are extensive properties of a culture and proportional to the volume of culture considered. These properties increase regularly and exponentially with time during balanced growth. However, temperature or ratios (such as DNA, RNA, or product per cell) are intensive properties that do not change during balanced growth. The application of this physical–chemical principle to bacteriology lies in the thought that if a culture is grown for a long-enough duration (usually with many subcultures) so that the biomass is sparse enough to not alter the environment significantly, sooner or later the organisms will come to achieve the same growth state for any particular constant environment, no matter what the cell's condition was initially—be it log, lag, or the stationary state. Once this balanced growth state has been achieved, if the conditions

remain the same, the culture will remain in balanced growth indefinitely. This would be true of a culture that had grown in a continuous culture or one that was diluted manually periodically, avoiding entry into either lag or the stationary state. However, if a nutrient is exhausted, or if the culture is altered by mutation and selection, or if the physical nature of the environment is changed, then balanced growth would cease and the culture would be in some nonbalanced state of growth.

The criteria given above are too stringent to be fully met with any practicality, but it is readily possible to study cultures that are substantially in balanced growth by growing cultures maintained at low density by dilution and using arbitrarily any one measure to follow growth, for example, biomass assayed turbidimetrically. If the doubling time remains constant over an extended period of time, then it is a good assumption that growth is balanced. Studies of these cultures can lead to an understanding of the properties of individual cells in the idealized state of balanced growth. But they would not bear on the properties of aggregates or clumps of cells or of cells reacting to some calamity or catastrophe in their environment.

Although we may consider any extensive property of a culture in balanced growth, let us first focus on biomass and call it X for tradition's sake. The rate of formation of biomass will be proportional to the amount of biomass. Call the proportionality constant, μ, then:

$$dX/dt = \mu X$$

Through the laws of calculus, this equation can be integrated and a boundary condition ($X = X_0$ when $t = 0$) imposed, yielding the well-known equation for exponential growth:

$$X = X_0 e^{\mu t}$$

However, every cellular substance, or even an extracellular product of the cells, could be substituted for our X. Thus, after growth becomes balanced, the increase of any component or group of components will be exponential in time. If the ratio of one substance to every other is to remain constant, the same proportionality constant must apply for every component. This common value of μ is called the *specific growth rate* or *growth rate constant*. Not only will any substance increase exponentially, but any combination of substances or the rate of change of any substance also will increase exponentially with this same specific growth rate, μ. Practically, this allows the rate of oxygen uptake, heat production, or production of an internal or external metabolite to be used as an index of growth. It will also allow growth to be monitored by the depletion of some medium component or the change in conductivity of the medium.

Intensive properties, such as μ and the ratio of the concentrations of different cell substances, must necessarily remain constant under conditions of balanced growth. In addition, the distribution of cell sizes and the average number of chromosomes will stay constant. An average cell will have a time-independent constant rate of carrying out every cellular process, and newly arisen daughter cells will have a constant probability of being able to form a colony (i.e., the percentage of nonviable cells will remain constant). Therefore, under these special conditions, no matter what measure of growth is used (whether it is particle counting, colony formation, chemical determination of a cell substance, or consumption or excretion of a substance), the same specific growth rate will be obtained and that rate will be constant through time.

Changes in Cell Composition at Different Growth Rates of Balanced Growth

Organisms respond to environmental conditions, both physical and chemical, by altering their own composition. These changes have been well documented for certain enteric bacteria but may occur with all prokaryotes (9). In general, favorable growth conditions mean faster growth, which requires a higher concentration of ribosomes and associated proteins. In terms of gross composition, the most obvious difference comparing a faster culture with a slower culture is a greater RNA content in the former (10). Also, under favorable growth conditions, the cells can lay down reserve materials such as glycogen and poly-β-hydroxybutyric acid. These changes in composition lead to the possible pitfall of the bacteriologist falsely relating one measure of growth to others when comparing growth in different environments.

Unbalanced Growth

Although balanced growth conditions lead to reproducible cultures, much of physiology deals with the responses of organisms to changes in their environment that lead to progressive changes in the organism (i.e., to the study of unbalanced growth). When a stationary culture is inoculated into fresh medium, the properties of organisms change drastically through the course of the batch culture cycle. Though only very well documented for certain enteric bacteria, similar changes probably apply with other prokaryotes. The exact course of changes in composition and morphology depends on the medium and on the age and condition of the inoculum. Culture cycle phenomena have relevance to growth measurements. These phenomena are particularly important in ecological studies, where the conditions under which the organisms grow are critical and, to a large degree, uncontrollable. The changes in characteristics are also involved in response to the fluctuations in natural conditions (11,12). For biotechnology purposes, it is the progression of changes in the so-called idiophase of growth that are paramount.

Noting that a typical bacterial culture generally means one of *Escherichia coli* (13), a typical bacterial culture cycle progresses as follows. When a stationary culture is diluted into rich medium, macromolecular synthesis accelerates. The components of the protein-synthesizing system (i.e., ribosomal proteins and ribosomal RNA) are made first. Only after considerable macromolecular growth has ensued does cell division take place. During this lag phase, the average size of cells increases greatly. When the capacity of the medium to support rapid balanced exponential growth is exceeded, metabolic processes (both intermediate metabolism and macromolecular synthesis) slow down. Different processes slow differentially in a way to

make small, RNA-deficient cells. Finally, the cells die and may eventually lyse.

Pitfalls in Measuring Growth

Errors in measuring bacterial growth fall into four classes. The first major pitfall is the general tendency of most organisms toward clumping or a filamentous habit of growth. This can occur even under mildly toxic conditions with bacteria that ordinarily divide regularly. The second pitfall is the differential viability of injured bacteria under different culture conditions. Repair processes may permit the recovery of viable cells under some, but not other, conditions. A third major pitfall is the possible development of resistant stages. Bacteria known to form resistant stages, such as spores, pose no problem because the controls and measurements to correct for such differentiated forms are well known; but when resistant stages are not suspected, error can arise. The fourth pitfall pertains to the way in which the inoculum is exposed to the new environment. Different results may be obtained if the concentration of an agent is raised gradually, if it is raised discontinuously, or if a high concentration is temporarily presented and then removed or lowered, such as when a concentrated solution of a nutrient is added and the mixing is not rapid. The cell concentration at the time of challenge can be critical. Bacteria may have special ways, sometimes inducible mechanisms or sometimes unknown ones, to protect themselves against toxic agents. Finally, the protection may be dependent on the number of organisms cooperating in detoxification.

Mycelial Habit of Growth

The evaluation of mycelial growth (the first pitfall) is both easier and much more difficult than that of well-behaved organisms that engage in binary fission and then promptly separate. The problems and methods for colonial growth have been discussed by Calam (14) and will be briefly dealt with here. Filtering filamentous cells, with or without drying, is easier than with smaller nonfilamentous cells.

The rate of increase in size of the colony in one, two, or three dimensions depends on the rate of elongation of those terminal hyphae that happen to be growing perpendicular to the surface of the colony, but the mobilization of resources into the mycelial mass depends on the surface area of the colony (15,16). Therefore, particularly with shake cultures, the results depend on the nature and size of inoculated fragments; with large fragments, growth becomes limited at an earlier stage by diffusion of nutrients into the mycelial mass. Thus, the major pitfall with the mycelial habit of growth is that the growth is quickly deviates from exponential growth and depends on the geometry of growth and the nature of the inoculum. In shake cultures, the apparent growth may depend on the shape of the vessel and on the shaking speed, its character (circular or reciprocal), and the distance moved, because all of the above can affect the tendency of the mycelia to break into smaller pieces. An important aspect of biotechnology is controlling the filamentous versus pellet conformations.

Cell Differentiation

The change of enteric bacteria from large, RNA-rich forms in the exponential phase to small, RNA-poor forms in the stationary phase has many of the aspects of differentiation. Bacteriology, however, has much clearer examples of cell differentiation in the cases of transition of rod to coccus, of vegetative bacillus to endospore, and in the formation of exospores, cysts, and buds. The tendency to form filaments in certain circumstances can also be considered a differentiation. These changes and their reversal pose potential pitfalls for all approaches to growth measurement.

Adsorption of Cells to Surfaces

Many bacteria naturally adhere to certain surfaces or can adapt and mutate to achieve a high avidity for solid surfaces, including glass. Many experiments with chemostat culture have failed to achieve their primary goal because the organisms adhered to the vessel walls. It is for these reasons that plastic or Teflon should be used whenever there is long-term contact of the organism with a culture vessel (16) Other approaches to minimizing the effects of growth on vessel walls include the use of large culture volumes in large containers, the use of violent agitation, and the frequent subculture of the bacteria into fresh glassware. These are the conditions in large biofermenters, but not during the production of the starter cultures. In addition, the use of detergents, vegetable oils, silicone coatings, and a high-ionic-strength medium may to some degree alleviate this problem.

This problem is particularly pertinent during dilution of the culture for measurement by highly sensitive means, such as microscopic and plate counts. It is, therefore, given consideration in the discussion of those methods (1,17,18).

DIRECT PARTICLE COUNTS

For many biotechnology purposes, the counting methods that can be used are limited because the growth medium generally contains particulate matter. This does not preclude their use in the development of strains and other phase of biotechnology.

Microscopic Enumeration

Microscopic enumeration is a commonly used technique that is quick and cheap and uses equipment readily available in the bacteriological laboratory. The technique is subject to gross errors, however. These can be overcome to a large degree using improvements suggested by the work of Norris and Powell (19). Precise details are given in my earlier review of this subject that are a further modification of this method (1), and some versatile techniques that have some of the characteristics of viable counts or metabolic measurements are given in the work by Murray et al. (20).

Methods and equipment, some linked to computer programs, are now available to record, scan, and view a microscope field. Software and hardware are available that enable a sample to be placed in a counting chamber viewed by a microscope, and then an electronic count can be made from the image. The computer can distinguish between cells adhering side by side and between a long cell and a pair of sister cells that have completed division but have

not separated. The capability exists to measure the biomass from the density of the pixels since the phase contrast (see later text) is a measure of the biomass. For filamentous organisms, the computer could output the total length of hyphae and convert that into biomass (21,22). Direct enumeration from charge-coupled device (CCD) images with highly intelligent computer filtering will allow distinguishing the bacteria or fungi from the particulate matter in the brew in many cases.

Electronic Enumeration

The Coulter Counter (Coulter Electronics, Hialeah, Fla.), its commercial competitors, and particularly the laboratory-built versions have been important in the development of bacterial physiology in the past 50 years. Such instruments are used routinely in clinical hematology. They are also very useful in the enumeration of nonfilamentous yeasts and protozoa, but not of mycelial or filamentous organisms.

The principle of electronic enumeration is as follows. A fixed volume of diluted cell suspension is forced to flow through a very small orifice connecting two fluid compartments. Electrodes in each compartment are used to measure the electrical resistance of the system. Even though the medium conducts electricity readily, the orifice is so small that its electrical resistance is very high. Consequently, the electrical resistance of the rest of the electrical path is negligible by comparison. When a cell is carried through the orifice, the resistance further increases because the conductivity of the cell is less than that of the medium. This change in resistance is sensed by a measuring circuit and converted into a voltage or current pulse. The counting rates can be very high, but beware of coincidence errors. The pulses are counted by an electronic circuit similar to that used in counting radioactivity. Very small pulses are eliminated by a discriminator circuit. Very high pulses, which might result from dirt or other irrelevant particles, are eliminated by an upper discriminator. In advanced models, the pulses may be analyzed by size and stored in a multichannel analyzer; later, the data can be recovered and plotted in a histogram, and the numbers, mean size, and standard deviation can be calculated. All the data may be collected, and the discrimination against pulses that are too high or too low can be carried out as the data are analyzed. The instruments have some method of forcing an accurately known volume through the orifice during the counting period. This is usually done by displacing the fluid in contact with a mercury column past triggering electrodes that conduct effectively through the mercury, but not through the diluent medium.

The major problem areas in electronic enumeration are three. The first problem concerns the size and shape of the cells. Some bacterial cells are very small (less than 0.4 μm^3), and the resistance pulses produced as they pass through the orifice are comparable to the noise generated by the turbulence that develops in the fluid flowing through the orifice. One can set the discriminator dial on the instrument to reject the turbulence noise, but one loses sample information, particularly about newly divided cells. One can run blanks and subtract the blank values, but blanks are particularly variable for small cell sizes. In addition, a pattern of turbulence can become established, remain for a while, and then be replaced with another pattern. This can cause major interference. Determining the size of cells has many difficulties, including the effect of the path that the particle takes through the orifice and the resistance of the various layers and compartment of the cell. Therefore, this potential use of the instrument is not discussed in this article. Finally, the overall error increases when the blank has large statistical variation. This is not a major problem for biotechnology because currently used microbes are relatively large.

The second major problem for physiological studies in electronic counting results from the failure of cells to separate promptly from each other after cell division. This, and the tendency to form filaments and aggregates, can be minimized by careful choices of the organism and the conditions. Possibly it can be reduced by vortexing or sonicating the sample. It was fortuitous for the advancement of microbial physiology that *E. coli* was chosen for physiological and genetic studies in the 1940s and 1950s because only a relatively small number persists in pairs or chains or in the form of aggregates. There is the related problem of coincidence, which is the passage of more than one cell through the orifice in a short enough time so that a single larger cell is registered by the electronic equipment. This problem can be dealt with by increasing the dilution so that the probability of coincidence is reduced.

The third major problem is technical: the clogging of the orifice in electronic counting. The resistance change is a smaller proportion of the total electrical resistance of the orifice when the orifice diameter is larger. Consequently, for small rod-shaped bacteria such as *E. coli*, orifices with diameters in the range of 12 to 30 μm must be used. The exact choice is a tradeoff between the increased signal strength of small apertures and the increased noise and probability of becoming clogged. Clogging is best prevented by ultrafiltration of all reagents. Alteratively, the diluent can be prepared and allowed to settle for a long time (months) in a siphon bottle so that the particulate-free solution can be withdrawn from above. Additional information about this technique can be found in the works by Drake and Tsuchiya (23) and Kubitschek (24).

Flow Cytometry

Flow cytometry became an extremely powerful method for the studies of many aspects of the biology of eukaryotes two decades ago. The instruments are now common in hospitals and clinical research laboratories. Because prokaryotes are smaller, it is only more recently that the instrumentation and methods have been refined and are only now coming into their own in the study of bacteria. The instruments operate by forming a small-diameter stream of the sample suspension. This is encased in a stream of particulate-free fluid. The tube through which they flow constrict in such a way that the stream of the sample is made still narrower in diameter. This process is called hydrodynamic focusing. The flowing stream is examined with laser light with various frequencies and angles, and the output of the detection circuits is used to detect when a

particle passes through and determine its characteristics. The design permits the analysis of biomass by light-scattering methods and by fluorescent staining of chemical components such as DNA. This will be discussed in "Flow Cytometry." The electronic circuits allow cells to be counted very rapidly. Growth can be followed by the increase in counts in samples counted at intervals. The problems for the application to prokaryotes is in sensitivity and background noise.

The original designs were made for study of problems in immunology. Applications, including microbiology, have been previously described (25–27).

Another quite different instrumental design was developed in Norway (28,29). It depends on flow directly on a microscope slide on an inverted microscope. This instrument uses an inverted microscope essentially operated in the dark-field mode. It uses a high-intensity mercury arc. Now with commercially available equipment, it has been utilized to study a variety of problems. Clever applications have been described (30,31).

The instrument of the first type uses a well-collimated laser source and fiberoptics. Great care is needed to maintain the instrument and its calibration. Routinely, polystyrene beads are mixed with the biological sample in a fixed amount. They have a uniform size and serve to quality control the electronics. By relating the biological count to the bead count, the concentration of biological particles can be determined. These kinds of controls can be used with the second type of instrument, which is cheaper and is now in use in many laboratories for many microbiological purposes. It is also capable of counting at very fast rates.

COLONY COUNTS EQUAL VIABLE COUNTS

Bacteriology really became an experimental science when Herr Dr. Robert Koch listened to Frau Fannie Hesse and developed the agar plate. This allowed not only the cloning of pure strains but also the enumeration of colonies arising from individual viable cells. Many variations have been developed and used in the ensuing century: (*1*) spread plates, (*2*) pour plates, (*3*) thin-layer plates, (*4*) layered plates, and (*5*) membrane filter methods. Although these techniques are very useful in the development of industrial strains, they take too much time and cannot be used on-line.

These techniques are described in many places (1). The pour-plate methods have variations in which the cells are grown in roller tubes or microtubes (32) and are examined with low-power microscopes when the colonies are small (33). There are many individual variations of techniques, sometimes resulting from historical accidents and sometimes resulting from special bacteriological circumstances. Automation of colony counting has put additional special restrictions on techniques, but allows colonies on the Petri dishes to be counted rapidly without operator error.

It must be emphasized again that mycelial organisms pose special problems. To some degree they can be overcome by sonication or shearing. A practical approach is to conduct a pilot study to find the amount of ultrasound that produces the maximal count that optimizes the separation of propagules from each other and minimizes the destruction of them by the physical treatment.

In the future, the development of growth media that is perfectly clear and can be solidified in uniform layers on a new generation of microscope slides will permit the analysis by telometering and computer analysis after only a few hours of incubation. In some cases, observation will be repeated and the difference in the images used to compute the growth rate as well as the initial CFU and viable biomass. Selectivity and identification of special properties with dyes and chemical reactions could be used as well. Again, particulate matter can be a problem, but less severe as the microcolonies grow to be larger than the detritus in the medium.

Spread Plates

The sample is pipetted onto the surface of solidified agar medium in a Petri dish, and then the cells are distributed with a wire, glass, or Teflon spreader. In the spread-plate technique, the growth forms surface colonies. It is reliable when methods to transfer all of a defined volume of cell suspension to the surface of the agar are used. The method is additionally useful because surface colonies are required to produce the proper color responses with many indicator agars that depend on the redox state of the colonies. In many cases, different colors are given from subsurface colonies because the oxygenation is different, and, therefore, the acid production and reducing potential are different than on the surface, and the selectivity is usually much less. (See Murray et al. [20] for recent improvements in technique.)

Pour Plates

An aliquot sample of diluted cells is pipetted into an empty sterile Petri dish; molten, but cool (45 °C), agar medium is poured onto the sample; and the contents are mixed by swirling and then allowed to harden. The pour-plate technique is a standard technique. Its main advantages is that most colonies are subsurface. They are, therefore, small, making it possible for a plate to have many colonies and still be countable. This method has a large dynamic range, and automation can be very useful.

Thin-Layer Plates

The use of thin-layer plates came out of virology. The sample is pipetted into a tube containing a small volume (2.5 to 3.5 mL) of molten but cool (45 °C) soft agar (0.6 to 0.75%) medium. This is then poured onto previously poured and hardened (1.5 to 2% agar) medium in a Petri dish, and the overlay is allowed to harden. This was the technique developed to measure virus plaques in the 1940s.

Layered Plates

Layered plates are like thin-layer plates except that an additional layer of agar medium is poured onto the newly congealed soft agar medium containing the cells so that all colonies are subsurface, but in the same plane. It is very useful because all colonies are subsurface and, therefore,

much smaller and compact. Under the right conditions they can be intensely colored because they are highly anaerobic. Many more colonies may be present, and yet the coincidence by fusion of colonies is small; this means that several thousand colonies per plate can be used and give meaningful results. This approach is especially recommended because the main difficulty with usual surface colony counting method is its lack of dynamic range. Rules have been issued that between 30 and 300 colonies are required for spread plates. The lower limit is set by statistical accuracy, and the upper limit is set by coincidence limitations. This 10-fold range is inconvenient for many purposes, because in many cases one cannot predict the number within a factor of 10 when choosing the dilution factor. The extra care needed to prepare the overlay plates is justified because it allows the counting in a larger range of 30 to 2,000 colonies. A second major advantage is flexibility for nutrient supplementation. Minimal agar can be used to pour the basal layer in a large number of Petri plates for indefinite storage. Stock supplies of the minimal soft agar can also be kept on hand. Then, 10- to 50-fold excesses of needed special nutrients can be added to the aliquots used for the molten top agar. In some cases, dyes and chromogenic substrates can be added as needed to the soft agar to allow screening of the colonies. Toxic substances can be incorporated to measure frequencies of resistant mutants.

Membrane Filter Methods

The diluted cells are filtered onto an appropriate membrane filter, which is then placed on an agar medium plate or onto blotter pads containing liquid medium. Sometimes it is necessary to carefully prewash the membranes and the pads with water or medium and then sterilize them. It is also necessary to supply nutrients at a higher level because they must diffuse to where the cells are located.

Automation

Several articles and books have been devoted to attempts to speed and automate growth measurement (34,35).

Special Situations

Certain organisms are very sensitive to substances present in agar. Meynell and Meynell (36) presented an excellent discussion of these problems. Injured organisms may have additional special requirements, and the entire 26th symposium of the Society for General Microbiology (37) was devoted to these problems. Genetically defective organisms pose individual problems that can be research problems on their own, such as the ability of various repair mutants to form countable colonies (38).

The problem of quantitating the number of organisms in cultures of strict anaerobes is discussed by Holdeman et al. (39) and Hungate (40).

Most Probable Number Method

The most probable number (MPN) method consists of making a number of replicate dilutions in a growth medium and recording the fraction of tubes showing bacterial growth. The tubes exhibiting no growth presumably failed to receive even a single cell that was capable of growth. Because the distribution of such cells must follow a Poisson distribution, the mean number plated at this dilution can be calculated from the formula $P_o = e^{-m}$, where m is the mean number and P_o is the ratio of the number of tubes with no growth to the total number of tubes. The mean number, estimated by $-\ln P_o$, is then simply multiplied by the dilution factor and by the volume inoculated into the growth tube to yield the viable count of the original sample. The MPN method is a very inefficient method from the point of view of statistics, because each tube corresponds to a small fraction of the surface of a Petri dish in a plate count. Consequently, many tubes or wells in a titer plate must be used, or the worker must be prepared to settle for a very approximate answer.

It is important to understand when the MPN method is of advantage. It can be used if there is no way to culture the bacterium on solidified medium, and it is preferred for mixed cultures if the kinetics of growth of the individual types are markedly different. Suppose some cells grow immediately and rapidly and end up making a large colony on solid agar that spreads over and obscures colonies of the organism of interest that form later. The small colony formers may be more numerous but unmeasurable on plates because of the fewer but highly motile or rapidly growing bacteria. Another use of the method is when other organisms that are not of interest are present in the sample, and no selective method is available to prevent their growth. Thus, the method has utility when the bacterium of interest produces some detectable product (e.g., a colored material, specific virus, or antibiotic). Then, even though contaminating organisms may overgrow the culture, the numbers of the bacterium in question can be estimated by the fraction of the tubes that fail to produce the characteristic product. Finally, if agar and other solidifying materials have some factors (such as heavy metals) that may alter the reliability of the count or interfere with the object of the experimental plan, then the MPN method can be used to avoid these difficulties.

Modern developments in laboratory techniques currently available can be used to speed the execution of the MPN method and make it more efficient. Machines are available to fill the wells of plastic trays that have as many as 144 depressions. However, it may be better to use 96-well trays because most apparatus, such as hand-held pipets, currently use them. Scanning devices designed for other purposes can be used to aid in counting the number of wells with no growth. Similarly, automatic and semiautomatic pipets can be used to fill small test tubes. Because these procedures make it possible to examine many more cultures, the classical tables of fixed numbers of tubes and fixed dilution series are obsolete and should be abandoned. A different more flexible approach and method of calculation has been provided (1). It may easily be optimized for the purpose at hand.

DIRECT BIOMASS MEASUREMENTS

Some measure of the bacterial cell mass or numbers of a culture is almost always used as the reference basis for

measurements of cellular metabolic activities, the types of morphological characters, or the amount of chemical constituent; biomass and cell numbers are the two basic independent parameters of bacterial growth. The methods for measuring biomass seem obvious and straightforward, but in fact they are complicated if accuracy if sought. Furthermore, the results may be expressed in different ways and, in some of these ways, the values may be more relative than absolute.

Wet Weight

A nominal wet weight of bacterial cells originally in liquid suspension is obtained by weighing a sample in a tared pan after separation and washing the cells by filtration or centrifugation. In either case, however, diluent is trapped in the interstitial (intercellular) space and contributes to the total weight of the mass. The amount of interstitial diluent may be substantial. A mass of close-packed, rigid spheres contains in its interstices 27% of space. This is independent of sphere size. A mixture of sizes packs more densely, and close-packed bacterial cells may contain an intersticial volume of 5 to 30 %, depending on their shape and amount of deformation. This is a problem that is readily solved if the washing step can be carried out with pure water.

Simple weighing or just measuring the packed volume of cells can be an excellent and rapid method with filamentous organisms or those that grow as pellets. Then filter, wash and weigh or centrifuge, and measure the height of the pellet. Both can be very rapid, and the procedures can be calibrated to correct for the exogenous water or shapes of centrifuge tubes. Because the particulate matter in the medium has different physical properties, it can be possible to separate the cells from the particulates and estimate the cell volume directly. This correction could be established with radioactive or fluorescent dextrans or other molecules that are too big to enter the cells.

Dry Weight

A nominal dry weight (solids content) of bacterial cells originally in a liquid suspension is obtained by drying a measured wet weight or volume in an oven at 105 °C to constant weight. The cells could be washed with water (possibly extracting cell components), or (better) a correction could be made for medium or diluent constituents that are dried along with the cells. Separating the cells by filtration poses particular problems. More problems arise if volatile components of the cells can be lost by oven drying, or if some degradation and volatilization occurs, evidenced by discoloration (particularly if a higher temperature is used). Some regain of moisture occurs during the transferring and weighing process in room atmosphere, so this should be done quickly within a fixed time for all replicate samples, especially if the relative humidity is high. It is best, of course, to use tared weighing vessels that can be sealed after drying.

It is possible that more accurate determination can be made by drying the sample to constant weight in a desiccator vessel with P_2O_5 and under oil-pump vacuum at 80 °C or by lyophilization. Results with the three methods are indistinguishable within 1% for certain cases (41). An excellent discussion of dry weight procedures and errors is given by Mallette (42).

The dry weight of cells may be reexpressed on a wet weight basis (grams of solids per gram of wet cells) or on a wet volume basis (grams of solids per cubic centimeter of wet cells or per cubic centimeter of cell suspension).

Because the drying can be a time-consuming process, and adequate knowledge about the possibility of volatilization of some cell components is not known, in the future, adequate drying quickly at lower temperature could be done in specially designed vacuum ovens. Such procedures could also be calibrated for the residual water content. Often, a practical method is to dry the cells in a microwave oven.

BIOMASS BY LIGHT SCATTERING

Light-scattering methods are the techniques most generally used to follow the growth of pure cultures. The major advantages are that they can be performed quickly and nondestructively. However, they may give information about a quantity not of primary interest to the investigator. Although they can be powerful and useful, they can lead to erroneous results. Under a range of conditions, they give information about macromolecular content (dry weight) and not about the number of cells or the state of hydration.

To understand their use in biomass determination, only the basic principle of Huygens is needed. Electromagnetic radiation interacts with the electronic charges in all matter. When the light energy cannot be absorbed (colorless materials), a light quantum of the same energy must be reirradiated. This light photon may emerge in any direction. This means that all atoms in a physical body serve as secondary sources of light. Those photons that happen to go in the direction of the original wave will stay in phase with the wave arriving directly from the light source, but light reirradiated from different points in the body that go in other directions will differ in phase at an observation point. The phase will differ depending on the distance that the photon must travel throughout the object because light going through matter is slowed. (The degree of slowing is measured by the index of refraction.)

It is refraction that controls how light is bent and focused in large bodies such as prisms, lenses, raindrops, or a pane of glass. In the last case, all the light that is scattered in any direction except straight ahead cancels, leaving only a beam of light going in the original direction. However, it is slightly retarded relative to a light ray not going through the pane; this is the basis causing the index of refraction of substances to be greater than 1. For the present purposes, the deflection of the light into different directions is the basis for the three types of useful techniques outlined next.

Theoretical Aspects

When the particles are very small, the light scattered is given by the Rayleigh equation:

$$\mathbf{I} = \frac{8\pi^4 r^6 \eta_0^4}{R^2 \lambda} \frac{[(\eta/\eta_0)^2 - 1]^2}{[(\eta/\eta_0)^2 + 2]^2} v\mathbf{I}_0 v[1 + \cos^2(\theta)] \quad (1)$$

In this equation, **I** is the intensity of the scattered light at a given direction and distance; r is the radius of the spherical equivalent of the actual particle; η_o is the index of refraction of the suspending medium; η is the index of refraction of the particle; v is the volume illuminated; $\mathbf{I}_o$ is the intensity of the incident light; v is the concentration of particles; θ is the angle of observation; R is the distance of the observer from the sample; and λ is the wavelength of light in vacuo. The factor depends on the angle of observation, θ, relative to the forward direction of the illuminating beam. The $[1 + \cos^2(\theta)]$ factor takes into account the two mutually perpendicular beams of plane-polarized light that comprise unpolarized light. For a given physical setup, most of these factors can be combined into an empirical constant.

In studies of bacteria, the biomass is commonly measured by turbidity. Turbidity is the amount of light scattered away from the forward direction and is thus a measure of the integral of all the scattered light. Similarly, for the flow cytometer applications, **I** must be integrated over the directions that are collected by the instrument's optics. This integral will be designated by **F**. In nephelometry only light scattered at 90° to the primary beam is recorded. This measure is actually a mixture involving on one hand the size shape and orientation of the particles and on the other the inhomogeneities inside the particle.

The range of very small particles where the equation 1 holds is called the Rayleigh domain. For this domain, the actual volume, index of refraction, shape, and orientation of the particles are irrelevant to the total light scatter. Most viruses are sufficiently small for their total scattered light to belong to this domain. For such small particles, the total light scatter integrated over all angles is proportional to the square of the volume, V^2. This results because the Rayleigh equation contains the radius, r, raised to the sixth power, which is equivalent to the volume squared. From the relative index of refraction, $m = \eta/\eta_o$ of the cells and the refractive index increment, the particle dry mass, q, and the wet weight can be calculated as discussed next, but only if all the parameters in equation 1 are known.

For somewhat larger particles with low index of refraction, such as most bacteria, which typically range from 0.3 to 3 μm in linear measure, light scattered from different parts of the particle can interfere both destructively and constructively as a result of phase differences. This defines the Rayleigh-Gans domain and is the relevant case for most microbes. The shape and orientation of the particles become important for this domain, where the particle's radius is comparable to λ', the wavelength of light in the suspending medium which is equal to λ/η_o. For particles in this domain, it is sufficient to take only the constructive and destructive interference into account, whereas phase shifts as the light traverses the particle can be neglected. The light scattering in this domain can be treated relatively simply, and for calculation, equation 1 need only be amended by multiplication with quantities called P functions. These functions are specific for shape, size, and orientation. Thus

$$I_{RG} = I(P(\theta)) \quad (2)$$

where I_{RG} is the light scattered as corrected by the Rayleigh-Gans method. The P functions are given and discussed in a variety of sources (see Ref. 1).

The mathematics of the third, the Mie, domain applies to particles of any size and index of refraction and are computer intensive. The computations can be carried out much more easily if the particles are spherical. For pigment-bearing cells, such as green and blue-green algae, the computations are very elaborate.

Although the Rayleigh-Gans or Mie mathematics apply as well to the simpler cases, we have defined the Rayleigh-Gans domain as the region in which interference, but not phase, must be taken into account, and the Mie domain is where both interference and phase shifts must be taken into account for accuracy. The choice of the appropriate mathematics depends on both the size and the index of refraction of the particle relative to the medium above that of the medium (i.e., on $\eta/\eta_o - 1$).

Bacteria (e.g., *Escherichia coli* growing in minimal medium) belong to the Rayleigh-Gans domain and not the Rayleigh domain because they are large enough to require correction for interference; however, they do not belong to the Mie domain because they have only a small dry mass, being mostly, water, and have an index of refraction only slightly larger than the aqueous growth medium and thus do not elicit phase shifts. Typically the index of refraction relative to the growth medium of many bacteria is 1.04. In any case, a relative index of refraction between 1.00 and 1.05 ensures that most bacteria suspended in their growth medium will fall in the Rayleigh-Gans domain (43–47).

Studies (48) show that dilute suspensions of most bacteria, independently of cell size, have nearly the same apparent absorbance per unit of the dry weight concentrations. Measurements at various angles (49) show that light is collimated mainly in the first few degree from the primary beam and at high angles almost purely measures the inhomogeneity of the bacterial cells.

Very different absorbancies (turbidities) are found per particle or per colony-forming unit with different sizes of bacterial cells. An approximate rule commonly assumed is that the dry weight concentration is directly proportional to the absorbance. This rule applies roughly to both cocci and rods and is a first approximation to a more precise rule, that is, that the light scattered out of the primary beam is proportional to the four-thirds power of the average volume of cells in the culture. Objects smaller than bacteria, such as suspensions of viruses, will not obey these rules. The rules also fail to apply when the suspensions of very small bacteria have a bluish cast. Larger objects (yeasts and filaments and aggregates of bacteria) may still appear cloudy and not necessarily colored, but may not obey the rule. Where the rule applies, the proportionality constant relating the absorbency measurement to the dry weight concentration is the same for any good, well-collimated photometer.

Turbidimetry

Bacterial suspensions are in between the size limits of atoms and objects like window panes. Consequently, most of

the scattered light is deviated only slightly; it is directed almost but not quite in the same direction as was the incident beam. The light scattered from an atom or a very small particle depends inversely on the fourth power of the wavelength of light. For a pane of glass, there is no light left uncancelled in any but the original direction, so the light retains its original hue and the pane appears transparent. If the glass was powdered to the size of bacteria, a suspension of powdered glass or bacteria (like a cloud) would appear white but not transparent nor colored. Because bacterial suspensions appear cloudy or turbid, instruments to measure the phenomenon are called turbidimeters (*turbidus* = confused or disturbed).

The common practice in bacteriology is to use any available colorimeter or spectrophotometer to measure turbidity. Ideally, such instruments measure only the primary beam of light that passes without deviation through the sample and reaches the photocell. Usually, the measurement is made of the light intensity relative to that which reaches the photocell when a sample of the suspending medium has replaced the cell suspension. From what has been said, an ideal photometer must be designed with a narrow beam and a small detector so that only the light scattered in the forward direction reaches the photocell. That is, the instrument must have well-collimated optics. Such an ideal instrument gives larger apparent absorbance values than simple instruments with poorly collimated optics, because in the latter a large percentage of the light scattered by the suspension is still intercepted by the phototube. Therefore, the measuring system responds as if there were less light scattered than is actually deviated from the forward direction.

If the instrument or the size of organism does not meet these criteria, valid measurements in many cases can be made by establishing a standard curve on either a relative or an absolute basis. Difficulty arises only as physical or biological conditions vary with the samples to be compared, since then the experimenter cannot be sure that the same standard curve applies.

These considerations apply only to dilute suspensions. Deviation from Beer's law must be applied after the absorbance exceeds about 0.3 when light of 500 nm or shorter is used. A good discussion of this question is given by Kavenagh (50). Failure to take this deviation into consideration is the most common difficulty encountered in the microbiological literature that involves turbidimetry. One way to avoid this problem is to restrict culture densities and use only dilute cultures for measurement (i.e., keep the absorbance below 0.3). It is usually better to make measurements at higher densities and correct for deviations from Beer's law. The longer the wavelength of light used, the less important the need for the Beer's law correction.

Nephelometry

Although most of the light scattered by bacteria is nearly in the forward direction, instruments that measure the light scattered at 90° from the primary beam have been used to measure bacterial concentrations. Instruments designed to make right-angle measurements of light scattering are called nephelometer (*nephelos* = cloud). Bacterial concentration may be measured with this nephelometric principle in a spectrofluorometer when the dials are set so that the excitation monochromator and the emission monochromator are at the same wavelength.

Nephelometric measurements can be ultrasensitive, so that very small concentrations of bacterial cells can be measured. If it is known which factors are important and whether elongated particles are oriented or randomly distributed in space, then light-scattering measurement over a range of angles can give information about several aspects: the state of aggregation of the protoplasm, such as whether the ribosomes are in polysomes or as monosomes within the living cell; the thickness of the cell envelope (46, 49, 51, 52); and the distribution of cell mass from the center of the cell

Wyatt (52) has developed an apparatus to measure the angular dependency of the light-scattering signal from 30° to 150°. In this range, different organisms, different treatments of a culture, and cultures in different phases of growth give characteristic patterns. These may be of some use as a fingerprint technique for the diagnosis and study of drug action. For the reasons stated above, the light-scattering signals in the high-angle range cannot be used reliably for growth measurement because they are sensitive to details of subcellular structure.

Light-scattering measurements made over a range of angles from 0° to 20° are relatively independent of these three factors, but are more critically dependent on overall size. Measurement in this range could nondestructively yield the biomass concentration, the number concentration, the average axial ratio, and a measure of average biomass distribution around the center of the cells for bacteria in balanced growth.

It may be easily possible to build a device in which an inexpensive neon laser illuminates a flow through a cuvette or a series of cuvettes of different path lengths. It would be so constructed that little stray light from the laser arrives at the detectors. One detector monitors the input beam, and a second monitors the light scattered a few degrees from the direction of the primary beam, say at 4°. A third measures the light scattered at right angles, and software outputs the dry weight biomass and intracellular heterogeneity that is a measure of the physiological state of the cells. It would avoid problems of the failure of Lambert's law by having multiple path lengths and would usually neglect the signals from the path length cuvettes in which deviations from the Beer-Lambert law were important. A variation would have several laser beams. The appropriate choice would depend on the color of the surrounding medium and the size of the microbes. The farther into the infrared that can be used, the less the Beer's law correction is needed, but there may be more interference from medium constituents.

It is possible to put probes into the fermentation apparatus for light scattering, turbidity, and nephelometry that have light sources and light detectors. These methods usually fail because the organisms grow on the optical surfaces. It is possible to remove, sterilize, and reinsert the probes, but this could lead to contamination. There are some ways around the surface-growth problem that may

be effectively implemented in the future to essentially by cleaning the optical surfaces in situ.

Flow Cytometry

For the conversion of forward scatter measurements in flow cytometry, the intensity of light pulse from a passing particle must be converted to volume or wet biomass or dry biomass. Our analysis (53) showed that in the Rayleigh region the biomass is related to the intensity raised to the 0.5 power, and in the Rayleigh-Gans region the power is approximately 0.75 for the usual bacteria and is somewhere in between for the very small bacteria that we are studying. Rules have been developed to establish the exact power. The results are applicable to bacteria and very small bacteria (0.02 to 0.6 μm^3) that are not usually spherical. The physical limitations and the mathematical basis of the calculations have been justified for the application of the proposed formula for practical use in microbial ecology and physiology of extremely small aquatic bacteria as well as the more well-studied forms. Although the method was developed for the small bacteria present in natural populations that are difficult to size by alternative methods, it will also be useful for industrial cultures of bacteria.

The major problem with both types of instruments discussed earlier is to convert the forward scattered light into a value proportional to the biomass of the particle. Although complicated, this problem has been solved for the first type of instrument (53). Consequently, the number of the channel of the multichannel analyzer can be related to biomass through a power law function. As of yet, a reliable way to do the same for the second class of instrument that uses an inverted microscope is not available. This does not prevent the instrument from being a very valuable instrument, but it will be more valuable when accurate sizes can be developed.

It may not be possible to theoretically analyze the light-scattering pulses from the second type of instrument and to convert them into dry weight biomass, but it is not impossible to imagine that something akin to the polystyrene latex beads that have become so useful in medical technology will be developed that have precisely defined size and an index of refraction in the range that includes most bacteria. Then, the second type of instrument will become useful for biomass determinations. As with light-scattering measurements, aggregation, clumping, and failure of cells to separate will remain problems to be coped with.

SAMPLING

Attention must be paid to analyzing an appropriate sample. Although this just takes attention to detail, it is so important that it should be emphasized again and again. During bioprocess technology, it is important to sample from a number of parts of the system. Pooling and mixing and then analyzing are appropriate, but in many cases, analysis of each part individually and the body of data analyzed statistically are efficient and useful.

ACKNOWLEDGMENT

I would like to thank George Hegeman and Matt Hilton for help in preparing this article.

BIBLIOGRAPHY

1. A.L. Koch, in P. Gerhardt ed. *Methods for General and Molecular Bacteriology*, American Society of Microbiology, Washington, D.C., 1994, pp. 248–277.
2. B. Sonnleitner, G. Locher, and A. Fiechter, *J. Biotechnol.* **25**, 5–22 (1992).
3. P.S.S. Dawson ed., *Microbial Growth*, Halsted Press, New York, 1974.
4. D. White, *The Physiology and Biochemistry of Prokaryotes*, Oxford Press, New York, 1995.
5. S.J. Pirt, *Principles of Microbe and Cell Cultivation*, Blackwell Scientific Publications, Oxford, U.K., 1975.
6. S.J. Pirt, *Stoichiometry and Kinetics of Microbial Growth*, Pirtferm, London, 1994.
7. S.J. Pirt, *Bioenergetics of Microbial Growth and Product Formation*, Pirtferm, London, 1994.
8. A. Campbell, *Bacteriol. Rev.* **21**, 263–272 (1957).
9. A.L. Koch, *Bacterial Growth and Form*, Chapman and Hall, New York, 1995.
10. M. Schaechter, O MaalØe, and N.O. Kjeldgaard, *J. Gen. Microbiol.* **19**, 592–606 (1958).
11. A.L. Koch, *Adv. Microbial. Physiol.* **6**, 147–217 (1971).
12. A.L. Koch, *Perspectives Biol. Med.* **20**, 44–63 (1976).
13. A.L. Koch and M. Schaechter, in A.L. Demain and N.A. Solomon eds., *Biology of Industrial Microorganisms*, vol. I, Addison-Wesley, Reading, Mass., 1985, pp. 1–25)
14. C.T. Calam, in J.R. Norris and D.W. Ribbons eds., *Methods in Microbiology*, vol. 1, Academic Press, New York, 1969, pp. 567–591.
15. J.I. Prosser, in N.A.R. Gow and G.M. Gadd eds., *The Growing Fungus*, Chapman and Hall, London, 1994, pp. 301–318.
16. A.L. Koch, *Microbiol. Mol. Biol. Rev.* **61**, 305–318 (1997).
17. F. Kavenagh ed., *Analytical Microbiology*, vol. 1, Academic Press, New York, 1963.
18. F. Kavenagh ed. *Analytical Microbiology*, vol. 2, Academic Press, New York, 1972.
19. K.P. Norris and E.O. Powell, *J.R. Microsc. Soc.* **80**, 107–119 (1961).
20. R.G.E. Murray, R.N. Doetch, and C.F. Robinow, in P. Gerhardt ed. *Methods for General and Molecular Bacteriology*, American Society of Microbiology, Washington, D.C., 1994, pp. 21–41.
21. U. Reichl, R. King, and E.D. Gilles, *Biotechnol. Bioeng.* **39**, 164–170 (1992).
22. G.C. Paul, C.A. Kent, and C.R. Thomas, *Biotechnol. Bioeng.* **42**, 11–23 (1993).
23. J.F. Drake and H.M. Tsuchiya, *Appl. Microbiol.* **26**, 9–13 (1973).
24. H.E. Kubitschek, in J.R. Norris and D.W. Ribbons ed., *Methods in Microbiology*, vol. 1, Academic Press, New York, 1969. pp. 593–610.
25. H.M. Shapiro, *Practical Flow Cytometry*, 3rd ed., Alan R. Liss, New York, 1995.
26. M.R. Melamed, T. Lindmo, and M.L. Mendelsohn, *Flow Cytometry and Sorting*, 2nd ed., Wiley-Liss, New York, 1990, p. 824.

27. D.K. Button and B.R. Robertson, *Cytometry* **10**, 558–563 (1989).
28. H.B. Steen, in M.R. Melamed, T. Lindmo, and M.L. Mendelsohn eds., *Flow Cytometry and Sorting*, 2nd ed., Wiley-Liss, New York, 1990, pp. 11–25.
29. D. Lloyd ed., *Flow Cytometry in Microbiology*, Springer-Verlag, New York, 1993, p. 188.
30. R. Nir, Y. Yisraeli, R. Lamed, and E. Sajar, *Appl. Environ. Microbiol.* **56**, 3861–3866 (1990).
31. R.I. Amann, B.J. Binder, R.J. Olson, S.W. Chisholm, R. Devereux, and D.A. Stahl, *Appl. Environ. Microbiol.* **56**, 1919–1925 (1990).
32. B.V. Perfil'ev and D.R. Gabe, *Capillary Methods of Investigation of Microorganisms*, University of Toronto Press, Toronto, 1969.
33. J.R. Postgate, *Adv. Microb. Physiol.* **1**, 2–23 (1967).
34. L.S. Gall and W. A. Curby, *Instrumental Systems for Microbiological Analysis of Body Fluids*, CRC Press, West Palm Beach, Fla., 1979.
35. C.-G. Heden and T. Illeni, *Automation in Microbiology and Immunology*, Wiley, New York, 1975.
36. G.G. Meynell and E. Meynell, *Theory and Practice in Experimental Bacteriology*, 2nd ed., Cambridge Press, Cambridge, 1970.
37. T.R.G. Gray, *Symp. Soc. Gen. Microbiol.* **26**, 327–364 (1976).
38. F.N. Capaldo and S.D. Barbour, in P.C. Hanawalt and R.B. Setlow eds., *Molecular Mechanisms for Repair of DNA*, part A, Plenum Publishing, New York, 1975, pp. 405–418.
39. L.V. Holdeman, E.P. Cato, and W.E.C. Moore eds., *Anaerobe Laboratory Manual*, 4th ed. Virginia Polytechnic Institute and State University, Blacksburg, 1977.
40. R.E. Hungate in J.R. Norris and D.W. Ribbons eds., *Methods in Microbiology*, vol. 3a, Academic Press, New York, 1971, pp. 117–132.
41. S.H. Black and P. Gerhardt, *J. Bacteriol.* **83**, 960–987 (1962).
42. M.F. Mallette, in J.R. Norris and D.W. Ribbons eds., *Methods in Microbiology*, vol. 1, Academic Press, New York, 1969, pp. 521–566.
43. A.L. Koch, *Biochim. Biophys. Acta* **51**, 429–441 (1961).
44. A.L. Koch, *J. Theor. Biol.* **18**, 133–156 (1968).
45. A.L. Koch, *Anal. Biochem.* **38**, 252–259 (1970).
46. H.H. Gunther and F. Berhgter, *Z. Allg. Mikrobiol.* **11**, 191–197 (1971).
47. A.L. Koch, *Anal. Biochem.* **38**, 252–259 (1970).
48. A.L. Koch, *Biochim. Biophys. Acta* **51**, 429–441 (1961).
49. A.L. Koch and E. Ehrenfeld, *Biochim. Biophys. Acta* **165**, 262–273 (1968).
50. F. Kavenagh, in F. Kavenagh ed., *Analytical Microbiology*, Academic Press, New York, 1972, pp. 43–121.
51. M. Kerker, D.D. Coke, H. Chew, and P.J. McNulty, *J. Opt. Soc. Am.* **68**, 592–560 (1978).
52. P.J. Wyatt, in J.R. Norris and D.W. Ribbons eds., *Methods in Microbiology*, vol. 8, Academic Press, New York, 1973, pp. 183–263.
53. A.L. Koch, B.R. Robertson, and D.K. Button, *J. Microb. Methods* **27**, 49–61 (1996).

See also KINETICS, MICROBIAL GROWTH.

MICROCARRIER CULTURE

MADHUSUDAN V. PESHWA
Dendreon Corp.
Mountain View, California
KATHY WONG
WEI-SHOU HU
University of Minnesota
Minneapolis, Minnesota

KEY WORDS

Anchorage-dependent cells
Bioreactor
Cell aggregate
Cell attachment
Cell immobilization
Mammalian cells
Microcarrier
rDNA protein
Shear damage
Vaccine

OUTLINE

INTRODUCTION

Most recombinant DNA–derived mammalian cell culture products are manufactured using cells grown in suspension. Commonly used cells include Chinese hamster ovary (CHO), baby hamster kidney (BHK), hybridoma, and myeloma cells. Those cells can be cultivated in fermentors similarly to microorganisms. However, the vast majority of cells grown in research laboratories, and many of those used in the manufacturing of viral vaccines for human and veterinary use, are anchorage-dependent or anchorage-preferred cells. This means that they need to attach to a compatible surface to survive and multiply, or, if given a compatible surface, they will readily attach to it and multiply rather than stay in suspension. Frequently used anchorage-dependent cells include MDCK, vero, mouse L, fibroblasts such as MRC-5 and WI-38, and almost all differentiated primary cells from nonblood tissues, such as hepatocytes, neuronal cells, and chondrocytes. These cells are grown on petri dishes, flasks (T-flasks), and roller bottles in research laboratories. Traditionally roller bottles are used in the production of biologics, such as cytokines and viruses, derived from those cells. In a typical production plant of those biologics, it is common to see tens of thousands of roller bottles being used for cell cultivation. The amount of labor requirement for roller bottle handling is daunting. Many of those processes manufacture products that were approved by regulatory agencies long ago. If the products were developed more recently, many, if not most, would have been manufactured with an alternative and improved method.

The need to establish a better method of culturing anchorage-dependent cells in large scale prompted the development of microcarriers. The first microcarrier culture was demonstrated by van Wezel in 1967 with diethylaminoethyl (DEAE) Sephadex A50 particles originally designed as column packing for ion exchange chromatography (1). The first industrial-scale product (inactivated polio vaccine) produced in microcarrier cultures was subsequently introduced by van Wezel in 1972 (2). This new way of culturing anchorage-dependent cells spurred other investigators to discover novel techniques to improve cell attachment to the beads, to test new types of microcarriers, and to increase the cell density. Serum coating of DEAE Sephadex beads (3), electric corona discharge for treatment of plastic (polystyrene) beads, and use of dextran microcarriers with a lower charge density (4) led to improvements in microcarrier technology. Since then, a variety of microcarriers have been developed and much research has provided insights into optimization of culture conditions. Table 1 presents a list of the most commonly used commercial microcarriers. The advantages of microcarrier culture over conventional monolayer culture methods (e.g., roller bottles, stacking flat plates) are numerous. Using microcarriers to provide a surface for cell attachment, a large surface area can be contained in a given reactor volume with a relatively uniform environment. Using a reactor similar to conventional fermentors, environmental monitoring and control become possible, thus opening up the possibility of further process optimization.

Most conventional microcarriers have a diameter of approximately 200–250 μm when suspended in medium and

Table 1. Commercially Available Conventional and Macroporous Microcarriers

Microcarrier type	Supplier
Conventional microcarriers	
Biosilon (polystyrene based)	Nunclon
Cytodex 1 (DEAE dextran)	Pharmacia
Cytodex 2 (DEAE dextran)	Pharmacia
Cytodex 3 (gelatin)	Pharmacia
Glass-, collagen-, and hydroxyapatite-coated plastics	Solohil Engineering
Macroporous microcarriers	
Cultisphere (collagen or gelatin)	Hyclone
Cytoline (polyethylene)	Pharmacia
Cytopore (cellulose)	Pharmacia

a specific density of 1.02–1.03 g/cm^3, slightly higher than that of the medium. This higher density allows a minimum agitation to be used to keep the microcarriers in suspension. It also allows for ready settlement after agitation is turned off, which is ideal for culture harvest or exchange of medium for supporting higher cell growth. A typical culture uses a microcarrier concentration equivalent to a settled bead volume 5–15% of total culture volume. This gives a surface area of about 0.5–1.5 m^2/L culture volume. Each liter of culture is equivalent to 5–150 roller bottles. The potential of increasing cell growth surface area was further advanced by the development of macroporous microcarriers (5). These are highly porous convoluted spheres with internal cavities for cell attachment and growth. Macroporous microcarriers can be used for culturing both anchorage-dependent and suspension cells.

Over the past 30 years, microcarrier culture technology has developed and matured to its present-day status. It allows for an easy way of producing a large quantity of cells both in laboratories and for industrial manufacturing. The ease of cell retention in microcarrier cultures also facilitates the employment of a perfusion system, which increases the maximum achievable cell density significantly. For large-scale cultivation of anchorage-dependent cells, microcarrier culture is still the best method.

CONVENTIONAL MICROCARRIERS

The microcarriers developed in the 1970s were initially based on cross-linked dextran. Soon after their introduction they were used successfully for the production of various vaccines and biologicals (6). Further development led to the use of other materials as microcarrier supports. These materials include polystyrene (7), glass (8), cellulose (9), and gelatin (10). The original dextran-based microcarriers require tedious procedures for detachment of cells (11) for subsequent inoculation, which is not the case for some of the other materials developed (8). Several studies have been directed at the effect of surface charge on cell attachment (12,13) and the kinetics of attachment (12).

Cell Surface Interactions and Attachment Kinetics

A microcarrier culture is typically initiated by trypsinizing cells that are adherent on a surface, either on tissue culture flasks or on microcarriers. Sometimes cells for inoculation are cultivated in suspension, and the production scale is conducted with microcarriers. Inoculation from a suspension culture eliminates the need for trypsinization. Once inoculated into a microcarrier suspension, the cells can either attach to the microcarriers or agglomerate to each other to form aggregates. It is essential that the rate of cell attachment to the microcarriers follows rapid kinetics to prevent them from agglomerating. The need for rapid attachment is even more pronounced in cases of primary cells such as hepatocytes or in case of other adherent cell lines that lose their viability if maintained in suspension over extended periods of time. For single cell populations, several studies have demonstrated that the initial rate of attachment of cells to microcarriers is essentially a first-order process, meaning the rate of increase in attached cells is proportional to the concentration of unattached cells, provided that the surface area available for attachment is not limiting (12). The process of cell attachment to any foreign surface can be viewed as a combination of two separate processes: initial adsorption of cells to the microcarrier surface followed by adhesion molecule–mediated attachment and cell spreading. The affinity of the binding of the cells to the microcarriers depends on the cell line, the microcarrier characteristics, the growth phase of the culture, the medium composition, and the cell loading per microcarrier. Cells initially adsorbed can possibly desorb from the microcarrier surface without really developing cell adhesion; the surface should therefore have sufficient affinity so as to allow the cells to remain attached to develop a firm grip by cell adhesion molecule–mediated interactions.

The initial adsorption of cells onto the microcarrier surface is a physical process and is believed to be facilitated by various attractive or affinitive interactions between the cell surface and the microcarriers. Cells in general fail to attach to unmodified crossed-dextran beads such as Sephadex but attach readily to DEAE Sephadex, although the optimal charge density for cell spreading and growth is less than that in DEAE Sephadex (4). Plastics such as polypropylene and polystyrene, in their native state, have a lower surface energy and are therefore incompatible for cell attachment. When treated with an electric corona discharge, their surface energy rises to that of tissue culture dishes, which enhances their ability for cell attachment. It has been demonstrated that corona discharge leads to the unmasking of carboxylic groups on the polystyrene surface, leading to this increase in surface charge density (14). Also, the attachment of cells has been shown to be affected by the length of the carbon chain on the unmasked surface groups (15). The cell surface is negatively charged due to the presence of negatively charged residues on the side chains of membrane-associated cell surface proteins. The interaction of the cell surface with positively charged surfaces can be attributed to the electrostatic interactions. The attachment of cells to negatively charged surfaces is likely due to Lewis acid type reactions in which short-range electron donor–acceptor interactions lead to coupling. In some cases microcarriers coated with cell adhesion factors, such as fibronectin, collagen, and laminin (16), have also been used to facilitate cell adhesion. It is interesting to note that for many cell types the initial attachment rate to charged microcarriers such as DEAE-modified cross-linked dextran beads (Cytodex-1) is faster than to those coated with collagen, which is an excellent surface for cell adhesion on petri dishes.

Furthermore, although serum provides many factors that promote cell adhesion, the initial attachment to charged microcarriers is often faster in the absence of serum (12). These observations suggest that electrostatic interactions between the surfaces of cells and microcarriers is a dominant factor for initial cell attachment. The provision of adhesion molecules, either covalently bound to microcarriers (16) or in soluble form in medium (17), promotes subsequent development of firm grips of cells on microcarriers and facilitates cell spreading. In many cases, no exogenous adhesion molecules are necessary because

cells synthesize their own. For some cell types the omission of these factors impedes cell spreading despite successful cell attachment to microcarriers (16). It is possible then to coimmobilize charged groups and cell adhesion molecules on the microcarriers to promote both initial physical attachment and subsequent biological events of development of adhesion plaques and cell spreading. Irrespective of the surface charge and the hydrophobicity or hydrophilicity of the microcarriers, it is essential that the surface have good wettability characteristics (18).

It has also been shown that a correlation exists between the exchange capacity of the microcarriers and the rate of cell attachment. A higher exchange capacity of the microcarriers leads to a higher first-order rate constant for cell attachment. However, at physiological pH, a higher exchange capacity is not correlated to a higher charge density. Moreover, a difference in cell spreading is observed at different exchange capacities (12). It appears that higher exchange capacities lead to increased cell attachment, but one must be cautioned that a higher exchange capacity may not lead to optimal growth of cells. The optimum exchange capacities for cell attachment and for cell growth are thus very different, and it is advisable to perform preliminary experiments to determine what exchange capacity is best suited for a particular cell type and a particular microcarrier type.

Inoculum Cell Density

After inoculation, cells attach to microcarriers randomly and the number of cells per microcarrier is distributed over a range. Once the adhesion plaques are developed and spreading occurs, most cells do not detach readily and reattach to other microcarriers. They grow until they cover the surface of the bead. For cells subject to contact inhibition, growth ceases then. Many cell types can form multiple layers, but growth rate often slows down after reaching confluence. Therefore, cells on beads with more cells initially reach confluence and stop growing or slow their growth rate faster while others are still growing. Once this happens the overall growth rate in the culture slows. The initial cell distribution on microcarriers thus affects the overall growth kinetics of microcarrier culture. The narrower the initial cell distribution is, the better the growth kinetics will be. Furthermore, microcarriers that do not receive any cells initially will remain barren. Thus, the initial cell distribution after inoculation also affects the maximum cell concentration attainable in the culture.

The initial random distribution of cells on the microcarriers can be simulated by a Poisson distribution, which predicts the probability of a microcarrier having a certain number of cells given the average number of cells per microcarrier (19). Ideally the average number of cells that attach per bead should be such that the fraction of beads without any attached cells is kept at a minimum. It has been shown that a minimal average of about five to six cells per microcarrier is a necessary inoculum size (19,20). With this inoculum ratio, the percentage of microcarriers without any cells attached to them is less than 1%, as estimated from the Poisson distribution.

Cell Shear

Adequate agitation in microcarrier cultures is critical for providing a homogeneous environment and thus facilitates oxygen transfer. The hydrodynamic forces that result from agitation can be detrimental to cells grown on microcarriers. Cells grown microcarrier culture are more shear sensitive than those grown in suspension. The dominant cell damage mechanism in microcarrier bioreactors is believed to be due to forces generated by the collision of microcarrier beads with each other and by small turbulent eddies, which are motions generated by the random variations of fluid direction and velocity in a turbulent flow field (21). In general, except in the spinner flasks used in laboratories, a pitch-blade or marine-type impeller is used for agitation. The ratio of impeller diameter to tank diameter is typically larger than that used in microbial fermentation. The larger impeller provides a larger capacity in liquid pumping for suspending the microcarriers. Because of the low solubility and mass transfer rate of oxygen in culture medium, external aeration is needed at larger scale. In many cases oxygen is supplied indirectly through membrane or silicone rubber tubing, which has a high permeability for oxygen. Use of such tubing is practical even up to a reactor of hundreds of liters. Direct sparging with air or oxygen is used in large stirred tank bioreactors. It is well known in suspension cultures that shear force from rupturing bubbles in a sparged system is the most dominant cell damaging mechanism. This is also likely the case in microcarrier cultures (22). With direct sparging surface active agents, such as Pluronic F-68, are used to reduce the attachment of cell-laden microcarriers to rising bubbles. Attaching to these bubbles and rising with them at a high velocity is often detrimental to cells attached to microcarriers. Furthermore, it is necessary to employ antifoaming agents. Otherwise, cell-laden microcarriers may rise to the top of the foam layer and accumulate there, eventually losing viability.

MACROPOROUS MICROCARRIERS

The cell concentration achievable in a conventional microcarrier system is limited by the available surface for growth. Increasing the microcarrier concentration beyond a certain limit does not lead to an increase in cell concentration because the increased agitation needed to suspend the microcarrier slurry becomes damaging to cells. Macroporous microcarriers were developed with an aim of circumventing some of these limitations of conventional microcarriers.

Macroporous beads are essentially convoluted spheres with relatively large pores for cell attachment and growth. The commonly used materials are gelatin (5), collagen (23), cellulose, glass, polyethylene (24), and polystyrene. The desirable properties of the materials of construction for macroporous microcarriers were discussed earlier. After inoculation of cultures, the cells attach to the external surface of the microcarriers and subsequently migrate inward to populate the interior of the beads. If large open channels leading to internal pores exist, cells may also populate the interior soon after inoculation. Thus, these macroporous

microcarriers provide extended surface area in the interior and potentially support a higher cell concentration with an equivalent number of beads. Also, the cells in the interior of the macroporous microcarriers are shielded from fluid flow and are hence not susceptible to detrimental shear effects (22). However, there are two areas for potential concern. First, in contrast to conventional microcarriers the surface of the macroporous microcarriers is full of cavities and is highly irregular. This irregularity interrupts the flow pattern in the bulk fluid in the immediate vicinity of the beads and gives rise to local microeddies that may damage the cells attached on the surface. Therefore, to avoid shear damage, once attached cells need to migrate inward at a relatively fast rate or the internal pore must be relatively open to allow for cells be carried into the interior by fluid convection during the cell attachment period after inoculation. The second cause for concern is the potential gradients of nutrients, oxygen, and metabolites along the radius of the microcarrier which may cause growth limitation. As a result, necrotic core could be induced within the beads (24,25). Use of intraparticle convection has been postulated as a method to overcome the formation of heterogeneous environment. Numerous cell types have been successfully cultivated on macroporous microcarriers for purposes of product production and virus production (26). Apart from stirred tank reactors, fluidized-bed reactors (27) are also used for cultivating cells in macroporous microcarriers.

Spatial Distribution of Cells and Cell Viability

To ensure efficient use of macroporous microcarriers the internal pores should be open to the exterior through open channels that allow for cell movement into the interior. Some collagen-based microcarriers have leaflike interiors with open and flat surfaces, and others have open pores interconnected by tortuous channels. The materials that the microcarriers are made of also affect the rate of cell movement. Thus, both the structure and the material may affect the performance of the culture. Ideally, macroporous microcarriers should have very open structures to allow cell penetration into the interior immediately after inoculation. Architectural staining of paraffin-embedded thin sections of cell-laden macroporous microcarriers with hemotoxylin and eosin (H&E) indicate that the distribution of cells in the interior is usually not uniform. Some of the internal cavities are packed to confluency, whereas others are devoid of cells.

Traditional methods of examining the pore structure and cell behavior in the interior require fixation and sectioning, which cause aberation. Furthermore, reconstruction of the three-dimensional image is difficult. By digitizing sectioned macroporous culture in three colors, it is possible to classify pixels as conforming to either background, cytoplasmic, microcarrier matrix, or nuclear parameters, based on a set of classification rules determined by statistical analysis (28). When relating the number of pixels occupied by cellular material to the total number of pixels in the sectioned microcarrier, quantification of cell growth can be generated. To examine the accessibility and geometrical features of the pores, confocal laser scanning microscopy (CLSM) can be used to optically section the beads (29). CLSM requires fluorescent samples. Most microcarriers can be rendered fluorescent by appropriate staining techniques. It is also possible to stain cells with a vital dye that has different excitation wavelengths than the dye used for microcarrier staining. This dual stain allows simultaneous observation of cells and microcarriers. Individual beads can be optically sectioned serially through their depths. A three-dimensional image can be reconstructed by combining the sections for examination of the openness and connectivity of the pores. To further evaluate the geometry of the pores, the solid structure of the beads can be removed from the reconstructed three-dimensional images to allow for direct visualization of the pores. Cell distribution, divisional, and other morphological characteristics can be observed directly. This confocal microscopic technique for direct evaluation of pore structure of macroporous beads is noninvasive and requires minimal sample preparation. It can easily be adapted for in situ observation of cell behaviors and to provide information to optimize the structure of those microcarriers.

MICROSPHERE-INDUCED AGGREGATES

A possible way to increase cell concentration in conventional microcarrier systems is by increasing the microcarrier concentration. However, a potential drawback of this scheme is the large ratio of settled bead volume to the culture volume. This large volume of settled beads calls for sophisticated methods of agitation to keep the microcarriers in suspension. Compared with these systems, the settled volume of cell aggregate with an equivalent cell concentration is relatively small, and the agitation requirements parallel those of simple suspension cultures. Also, some adherent cell lines have been engineered to grow in suspension or as aggregates by processes of adaptation or media adaptation. It has been demonstrated that agglomeration can be induced by lectin, concanavalin A, and wheat germ agglutanin (30). However, these processes are relatively slow, and a significant number of the cells initially inoculated lose their viability because of the duration required for aggregate induction.

Microspheres made of DEAE-derivatized Sephadex beads with diameters of 20–50 μm have been demonstrated to increase the rate of aggregate formation (31). Hence, this process can be thought of as a process similar to nucleation and growth in crystallization. After the initial attachment phase, these cell-laden microspheres come into contact with each other to form larger clumps, which serve as nucleation sites for aggregate formation via growth of these attached cells. Over time, aggregates emerge, each consisting of a few microspheres surrounded by multiple layers of cells.

A variety of cells, including CHO, vero, human embryonic kidney cells (293 cells), and swine testicular cells have been successfully grown on these microspheres (31) for virus and protein C productions (31). The drawback of employing aggregate cultures is that little is known about controlling the size of aggregates. As the aggregate size increases, nutrient limitation and a decreased viability can be expected.

CONCLUSION

Microcarrier processes, because of their ease of operation and control, are ideal for cultivating anchorage-dependent cells in both the laboratory and in manufacturing. The bioreactor operation requirements are virtually identical to those for a typical suspension culture in fermentors except that the upper limit for agitation is lower. Operating modes such as batch, fed batch (32), and perfusion are routinely being used in microcarrier cultures. The prevailing use of microcarriers in industrial scale is for the manufacture of viral vaccines and reagents. Depending on the cell type, serum-free cultivation is also possible. Because most newly developed cell culture processes for rDNA products employ suspension cells, microcarrier culture processes will see only moderate growth in the future. However, with increased efforts in research and product development employing differentiated cells, we may see microcarriers play an increasingly important role in small to moderate scales. This technology—developed three decades ago—is still intriguing to cell culture technologists. Its versatility in supporting the growth of a variety of anchorage-dependent cells ensures it a unique place in bioprocessing.

BIBLIOGRAPHY

1. A.L. van Wezel, *Nature* **216**, 64–65 (1967).
2. A.L. van Wezel, *Progr. Immunobiol. Stand.* **5**, 187–192 (1972).
3. R.E. Spier and J.P.W. Whiteside, *Biotechnol. Bioeng.* **18**, 659–667 (1976).
4. D.W. Levine, D.I.C. Wang, and W.G. Thilly, *Biotechnol. Bioeng.* **21**, 821–845 (1979).
5. K. Nilsson, F. Buzsaky, and K. Mosbach, *Bio/Technology* **4**, 989–990 (1986).
6. B. Meignier, *Dev. Biol. Stand.* **42**, 141–145 (1978).
7. A. Johansson and V. Nielsen, *Dev. Biol. Stand.* **46**, 125–129 (1980).
8. J. Varani, M. Dame, T.F. Beals, and J.A. Wass, *Biotechnol. Bioeng.* **25**, 1359–1372 (1983).
9. S. Reuveny, L. Silberstein, A. Shahar, E. Freeman, and A. Mizrahi, *Dev. Biol. Stand.* **50**, 115–123 (1982).
10. J. Varani, M. Dame, J. Fediske, T.F. Beals, and W. Hillegas, *J. Biol. Stand.* **13**, 67–76 (1985).
11. T.-Y. Tao, G.-Y. Ji, and W.-S. Hu, *Biotechnol. Bioeng.* **32**, 1037–1052 (1988).
12. V.B. Himes and W.-S. Hu, *Biotechnol. Bioeng.* **29**, 1155–1163 (1987).
13. S. Reuveny, A. Mizrahi, M. Kotler, and A. Freeman, *Biotechnol. Bioeng.* **25**, 469–480 (1983).
14. W.S. Ramsey, W. Hertl, E.D. Nowlan, and N.J. Binkowski, *In Vitro* **20**, 802–808 (1984).
15. M.J. Kuo, C.J. Lewis, R.A. Martin, R.E. Miller, R.A. Schoenfeld, J.M. Scheck, and B.S. Wildi, *In Vitro* **17**, 901–906 (1981).
16. J. Varani, S.E.G. Fligiel, D.R. Inman, T.F. Beals, and W.J. Hillegas, *Journal of Biomedical Materials Research* **29**, 993–997 (1995).
17. M. Kan, Y. Miyamoto, S. Sunami, I. Yamane, and M. Umeda, *Cell Structure and Function* **7**, 245–252 (1982).
18. M. Kiremitci and E. Piskin, *Biomat., Art. Cells, Art. Org.* **18**, 599–603 (1990).
19. W.-S. Hu, J. Meier, and D.I.C. Wang, *Biotechnol. Bioeng.* **27**, 585–595 (1985).
20. W.-S. Hu and D.I.C. Wang, *Biotechnol. Bioeng.* **27**, 1466–1476 (1985).
21. E.T. Papoutsakis, *TIBTECH* **9**, 427–437 (1991).
22. D.E. Martens, E.A.A. Nollen, M. Hardeveld, C.A.M. Vander Velden-de Groot, C.D. de Gooijer, E.C. Beuvery, and J. Tramper, *Cytotechnology* **21**, 45–59 (1996).
23. R.C. Dean, Jr., S.B. Karkare, P.G. Phillips, N.G. Ray, and P.W. Runstadler, Jr., in B.J. Lydersen ed., *Large Scale Mammalian Cell Culture Technology*, Hanser, New York, 1987, pp. 145–167.
24. A. Preissmann, R. Wiesmann, R. Buchholz, R.G. Werner, and W. Noe, *Cytotechnology* **24**, 121–134 (1997).
25. M.L. Kennard and J.M., *Biotechnol. Bioeng.* **47**, 550–556 (1995).
26. T.J. Nikolai and W.S. Hu, *Enzymol. Microb. Technol.* **14**, 203–208 (1992).
27. P.W. Runstadler, Jr., A.S. Tung, E.G. Hayman, N.G. Ray, S.G. Sample, and D.E. De Lucia, *Bioprocess Technol.* **10**, 363–391 (1990).
28. D.J. Foran, F. Cahn, and E.F. Eikenberry, *Analytical & Quantitative Cytology & Histology* **13**, 215–222 (1991).
29. S. Bancel and W.S. Hu, *Biotechnol. Prog.* **12**, 398–402 (1996).
30. P. Whur, H. Koppel, C. Urguhart, and D.C. Williams, *J. Cell Sci.* **23**, 193–209 (1977).
31. S. Goetghebeur and W.S. Hu, *Appl. Microbiol. Biotechnol.* **34**, 735–741 (1991).
32. R.Z. Mendonca and C.A. Pereira, *Bioproc. Eng.* **12**, 279–282 (1995).

See also MAMMALIAN CELL BIOREACTORS; MAMMALIAN CELL CULTURE REACTORS, SCALE-UP.

MICROENCAPSULATION

HOLGER HUEBNER
R. BUCCHOLZ
Technical University of Berlin
Berlin, Germany

KEY WORDS

Alginate
Beads
Cellulose sulfate
Coated beads
Encapsulation techniques
Hollow microspheres
Immobilization
Microencapsulation
Polyvinyl alcohol

OUTLINE

INTRODUCTION

The immobilization of active biological substances or living cells has become a universal tool in biotechnology over the past decades. Immobilization can be defined as any procedure that confines substances or cells inside a given system and limits its free diffusion or migration. In biotechnology, the concept of microencapsulation is related to a special immobilization system, where the biological material is confined inside particles, beads, or hollow spheres.

Spherical particles can be considered to be the best-suited immobilization system for use in bioreactors, due to their optimal hydrodynamic properties and abrasion resistance. Because microencapsulation is a very powerful and universal tool, the applications for microencapsulated systems cover a very wide field. To date, it has been used to improve the relatively outdated biotechnological processes in the food industry (1,2), wastewater treatment (3), leaching (4), and environmental detoxification (5). In modern biotechnology it has been applied to cell culture processes for the production of high-value substances such as antibodies (6,7), erythropoetin (8), and the anticancer drug taxol (9), as well as for the production of artificial seeds (10,11) and cryopreservation (12). In medicine, microencapsulated cells have been used as artificial organs to treat diabetes (13,14), as a delivery system for gene therapy (15,16), as an intermediate step in cancer therapy (17), and for the treatment of Parkinson's disease (18,19), to name a few. Furthermore, microencapsulation can serve as a controlled release system for drugs in medicine (20–22) and pesticides in agriculture (23,24).

MICROCAPSULES

Although there are a variety of encapsulation techniques and suitable materials, the resulting microcapsules can be divided into only three main groups: beads, coated beads, and hollow spheres.

Beads

Beads can be produced by cooling liquid drops of gelling agents (gelatine, agarose) under their melting point, thus transforming the polymer solution into a stable cryogel by internal hydrogen bonding, or by chemically or ionically cross-linking polymers to produce a hydrogel. Because bead formation is a simple one-step process, it is relatively easy to develop a large-scale production process (25). The most common bead formation system is alginate cross-linked with calcium ions or any other divalent metal ions.

The fully developed three-dimensional internal network structure enables the beads to withstand extreme mechanical stress. Permeability of hydrogel beads is excellent, and they are often used for culturing cells and microorganisms. The disadvantage is their lack of a real barrier on the surface, so cells can expand during cultivation.

Coated Beads

Providing beads with one or several additional walls leads to the formation of coated beads, which have been developed to overcome the problems associated with the open porous structure of bead surfaces. To produce a membrane, the beads from charged polymers (e.g., alginate, agarose) are treated with a diluted solution of an oppositely charged polyelectrolyte (e.g., poly-L-lysine, poly(ethylene imine), poly(*N*-vinylamine), chitosan) thus forming a simplex membrane on the bead surface. The use of several alternatively charged electrolytes leads to multilayer membranes (26–28). Alternatively, beads can be soaked with a photosensitive substance that cross-links when irradiated (29–31).

A proper coating process produces an additional membrane without incorporation of the material to be immobilized and has the advantage that the membrane properties can be engineered independently with respect to diffusion, molecular weight cut-off, cell retention, or immunoprotection from the internal bead structure and material. Possibly the most well-investigated clinical application of microencapsulation is based on poly-L-lysine-coated calcium alginate beads in the treatment of diabetes (32,33), which has recently been applied to humans (34,35).

Hollow Spheres

Hollow spheres can be produced in a one-step process using two membrane-forming materials that cannot penetrate each other due to diffusional limitations. The cross-linking reaction is then limited to the interface area of the hollow sphere, forming drops and producing a stable membrane around a liquid core. Typical hollow-sphere-forming materials are all polyelectrolyte combinations, for example, cellulose sulfate (CS) and poly(diallyl dimethyl ammonium chloride) (PDADMAC), and CS and poly(ethylene imine) (PEI), able to produce a simplex membrane (36). Another method for the production of hollow spheres is the use of two materials such that the cross-linking material of the drop can diffuse into the surrounding solution. The membrane starts forming at the surface of the droplets and then proceeds outward, creating cell-free membranes. The membrane thickness can be controlled by the electrolyte concentration of the droplet. In principle it is a reversal of the bead-forming process (37). Of course, hollow spheres can also be obtained from a multistep process in which the

core of the coated beads is solubilized. This method has been used for controlled release systems in agriculture and the pharmaceutical industry. It is applied on shear-sensitive compounds such as animal cells in particular (38,39).

Recently simplex membrane systems have gained importance for medical applications such as artificial liver support (40) and gene therapy (41). Considerable progress has been made through a U.S. National Aeronautics and Space Administration study in which over a thousand combinations of polyanions and polycations were tested to identify new polymer candidates that could be suitable for the encapsulation of living cells, especially pancreatic islet cells for treatment of diabetes in humans (42,43).

FUNDAMENTAL METHODS IN MICROENCAPSULATION

The basis of all microencapsulation processes is the formation of a liquid drop followed by gelation (44,45), a cross-linking reaction, or membrane formation. The liquid drops can be obtained from extrusion of a liquid through a small needle or orifice, or from emmulsification of the drop-forming solution in a second immiscible solution by use of dynamic or static mixers. Scale-up ability of the process and a uniform microcapsule size and shape are the main parameters in encapsulation technology (46).

Dropping Methods

A liquid ejected with low velocity from a needle will break into individual drops. If the velocity is increased, drop formation increases until the maximum velocity is reached and the liquid begins to form a jet. The maximum velocity can be calculated from the following formula using the liquid velocity, v, interfacial tension, γ, liquid density, ρ, and the inner needle diameter, d_i:

$$v < 2\ (\gamma/\rho d_i)^{0.5}$$

Simple Dropping. The two main factors affecting drop size are the force of gravity trying to tear the drop from the needle tip, and the resisting product of the interfacial tension and the tip perimeter. Other forces interacting with drop formation are resistance power and inertial force, but these can be neglected in computations. The drop mass can be calculated from the equilibrium of the two main forces with the mass, m, gravity acceleration constant, g, external needle diameter, d_e, interfacial tension, γ, and the liquid density, ρ:

$$mg = \pi d_e \gamma \text{ and } m = \frac{\rho \pi d^3}{6}$$

The drop mass must however be corrected by a factor of 0.85 to get the true mass, because drops stretch out and leave a small portion behind when dropping from the needle. Even if very thin needles are used, it is very difficult to obtain droplets with a diameter of 1 mm or less (47,48).

Dropping with a Superposed Air Jet. The droplet size obtained from simply dropping can be reduced by superposing the drop forming process with an additional air jet. For calculating the droplet mass, the dragging forces of the airflow have to be taken into account. The viscous drag force is the laminar effect of the fluid, and the kinetic energy dissipation term represents the impact of the turbulence on the drag. Assuming that the drop diameter is larger than the tip, that there are no lateral fluid effects, and that the nascent drop has a spherical shape, the mass, m, of the generated droplet can be calculated using the gravity acceleration constant, g, liquid viscosity, μ, fluid velocity, v, droplet diameter, d_D, external needle diameter, d_e, the surface tension, γ, and the density of the fluid, ρ (49):

$$mg = \pi d_e \gamma - (3\pi\mu v d_D + 0.055\ \rho v^2 d_D^2)$$

Depending on the immobilization material used, uniform droplets from 2 to 6 mm can be produced. Droplet production can be scaled up by use of several needles at once and has been successfully performed for the encapsulation of IgG-producing hybridoma cell cultures and yeasts (50). However, the increasing airflow influences the line of flight of the droplets, which becomes important if several adjacent needles are used simultaneously. As long as the superposed airflow is laminar ($Re < 0.3$, Stokes' area), droplets will fall in a perfectly vertical line. For turbulent superposed airflow producing periodical eddies ($1{,}000 < Re < 10^5$, Kármán vortex street), the diversion can be neglected up to $Re < 1{,}080$. Stronger turbulence may cause collisions of the falling drops but can be avoided if the distance between the needles is increased. At any rate, small satellite drops are formed and site distribution increases significantly.

This method uses an airstream that is concentric with respect to the droplet-forming needle. Excentric air sleeve positioning combined with a beveled droplet-forming needle facilitates the generation of very small droplets. The distal end of the needle is beveled at an angle of about 15 to 45°; the beveled end is disposed facing the central axis

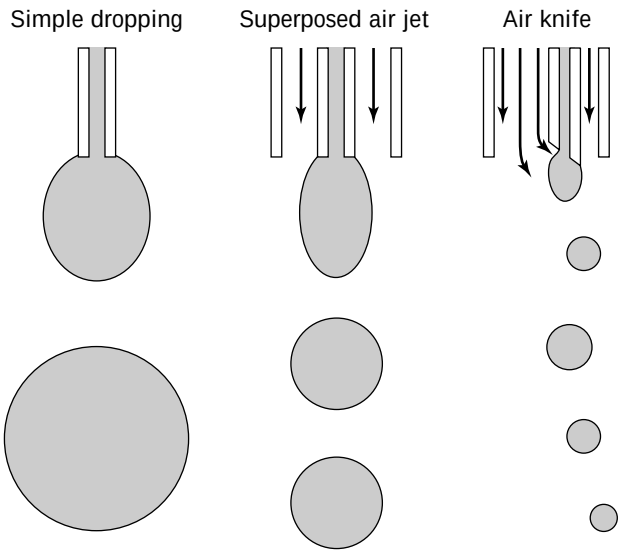


Figure 1. Monoaxial extrusion technologies based on simple dropping.

of the sleeve for the airstream. This arrangement leads to the formation of an air knife that generates droplets of 25–300 μm in diameter. The distribution of size is quite high due to the fact that the airstream not only drags the nascent drops but also breaks up the drop into smaller droplets. The rate of production is limited to 0.1–3.0 mL/min; the preferable limit is 0.3 mL/min (51,52).

Coaxial Dropping with a Superposed Air Jet. All of the systems mentioned so far include the mixing of the microcapsule-forming polymer with the material to be immobilized. This material is distributed evenly in the immobilization matrix and affects the membrane-forming process. In particular, if animal cells are immobilized in hollow spheres from polyelectrolyte complexes, the cells will be incorporated into the membrane. Because these membranes are very thin (1–5 μm) compared with those of animal cells (10–18 μm), several gaps result when these cells disintegrate. Complete cell and product retention or immunprotection cannot therefore be achieved. Coaxial dropping can be used to overcome this problem. This system uses one needle for the membrane-forming polymer solution and a second one, which is fixed in the center of the first, for the cell suspension. Both needles are of the same length. The nascent drop formed at their ends has a core containing the cell suspension and is enveloped with the polymer solution. While falling, the polymer solution completely surrounds the cells that comprise the core. The membrane-forming process in the precipitation bath is not affected. The core may have a volume of up to 25% of the whole drop. The droplet diameter can be varied from 1 to 5 mm. All the equations of the previous section are still valid.

This system has been scaled up to a productivity of 4 L/H by the simultaneous use of 24 concentric needles. To date this system has been successfully used for the immobilization of hybridomas (53), insect cell cultures (54), and the treatment of rheumatoid disease in mice, where microcapsules have been used intraperitoneously in mice containing a hybridoma cell line secreting antirheumatoid monoclonal antibodies (55).

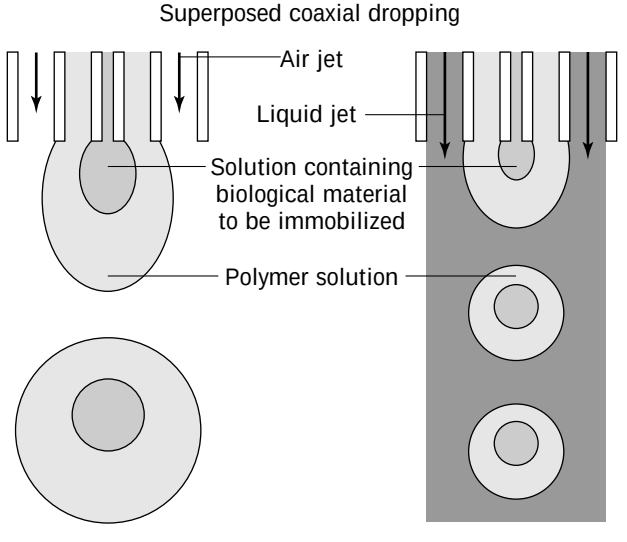


Figure 2. Coaxial extrusion technologies based on simple dropping.

Coaxial Dropping with a Superposed Liquid Jet. Superpositioning is not restricted to the use described in the last section; it can be used with any liquid immiscible with a drop-forming solution along with a precipitation bath. There are two main advantages to this method. Fluids still flow in a laminar fashion at higher velocities as compared with gases, and the viscous drag force is much higher. Both effects enable the production of hollow spheres from polyelectrolytes with a diameter ranging from 500 to 1,000 μm. This is considerably less compared with the previous system. Unfortunately this system is restricted to a productivity of 150 mL/h, and scale-up by use of multiple devices is problematic.

Electrostatic Extrusion. The application of a high static potential between the capillary and collecting solution has been used to improve droplet formation with respect to reducing the droplet size. The electric field force effectively pulls the forming droplet off the tip of the needle at a much lower mass (and hence size) compared with the simple dropping method. A series of small droplets (as small as

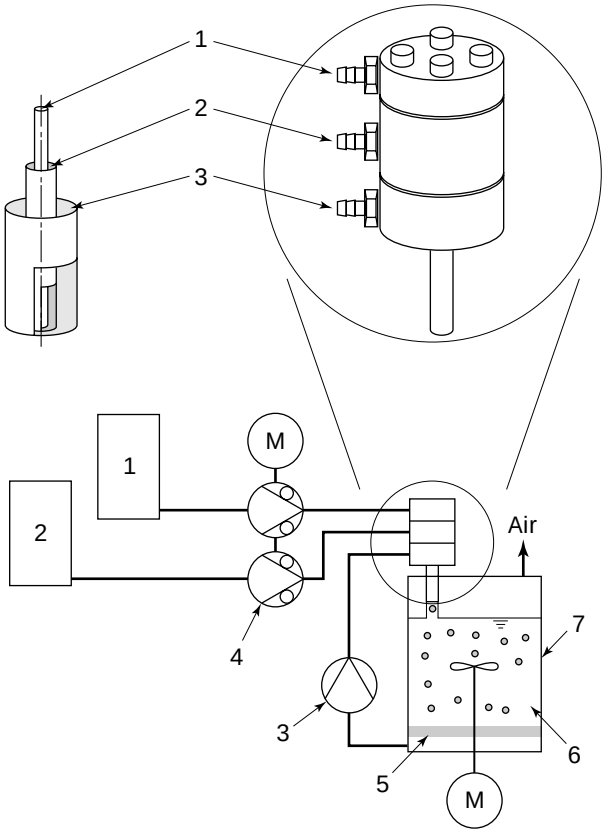


Figure 3. Schematic representation of the coaxial encapsulation apparatus (Torsten Steinau Verfahrenstechnik Berlin): (1) cell suspension, (2) cellulose sulfate solution, (3) inert carrier fluid, (4) multichannel peristaltic pump, (5) sieve for capsule retention, (6) polyDADMAC precipitation bath, and (7) stirred tank reactor.

26 μm in diameter) is generated with a comparably low standard deviation. The average microbead size can be easily adjusted from 2.5 to 0.2 mm by decreasing the concentration of the polymer solution, by using a higher-gauge needle, and by increasing the applied voltage. However there is a limit to how much the size can be reduced. In very high electric fields, droplets are no longer formed, and the solution pours from the needle in a steady stream. Charged molecules accumulate at the nascent drop's surface and counteract the surface tension. According to Lippmann's theory, the size of the drops can be determined with the electrostatic force, Fe, electric permitivity, ε, droplet diameter, d, the applied electrostatic potential, U, surface tension, γ, mass, m, and gravity acceleration constant, g:

$$mg = \pi d\gamma - Fe$$

$$mg = \pi d(\gamma - (k\varepsilon U^2/d))$$

$$d = 1/(6\ m/\pi\rho)^{1/3}$$

The correction factor, k, is difficult to evaluate. It depends on system design and may be a function of the form of the pending drop and of the electrostatic potential. This system has so far been applied to the encapsulation of cells in alginate beads (56,57), pancreatic islets in coated beads to study the influence of diameter and membrane compactness integrity and immunoprotection (58), and to the production of protein-loaded ethylene vinyl acetate beads for controlled-release purposes (59).

Rotating Disk. A liquid can be sprayed into a precipitation bath by using a rotating disk or a dish rotating at a high speed. The liquid for dispersion emerges from an orifice at the center of the axisymmetric rotating disk. The liquid leaves the disk in the form of small jets or ligaments, which form drops prior to falling into the precipitation bath. Droplet diameters, d, can be determined from the surface tension, γ, rotation speed (per second) of the disk, w, disk radius, r, and the liquid density, σ (60,61):

$$d = 0.425\ (\gamma/(w^2 r\sigma))^{0.5}$$

This system can be improved by applying a well-designed wave on the liquid flowing on the spinning disk. The ligaments are broken into uniform droplets with a standard deviation lower than 5%. Typically, a disk of 1 cm in diameter rotating at 2,000 rpm forms about 60 ligaments, which then form droplets of 500 μm at a flow rate of 6 L/h. Equations describing this process are much more complex than those for classical jet rupture. Physical properties of the liquid, design of the rotating disk, and wave frequency and amplitude must be adjusted to achieve correct particle size. This method is currently being used for experimental research and agriculture (62).

Rotating Nozzles. Instead of a rotating disk, a perforated rotating cylinder can be used. This cylinder can be placed in a reactor, with the reactor serving as the precipitation bath and the culturing vessel. Monodisperse and spherical drops are produced by rotating the cylinder at a velocity that hurls the drops directly down into the reactor liquid. The size of the beads is a function of the cylinder's rotation velocity, with a size of 1–3 mm and a deviation of up to 10%. The diameter and number of holes drilled in the cylinder can be varied according to the desired capacity, which can be up to hundreds of liters of alginate beads per hour. The volume, V, of the droplets can be calculated from the hole diameter, d, surface tension, γ, density of the liquid, ρ, gravity acceleration, g, and a correction factor, k, which depends on the shape of the holes:

$$V = (2\pi d)/(k\rho g)$$

The variation in bead diameter, d, for gelled alginate beads can be expressed as a function of the rotation velocity, ω:

$$d = (1/\omega^2)^{1/3}$$

As has been observed for simple dropping, the diameter, d, has to be corrected by a factor of 0.84 to 0.95, depending on the rotation velocity used, because the drop leaves behind a portion of the pendant drop upon breaking. The landing radius of the drops on the liquid surface must be taken into account to avoid splattering of the drops on the reactor walls. The landing radius, r, is dependent on the velocity, v, of the drops when leaving the hurler, the height, h, of the holes above the liquid surface, and the acceleration force, g:

$$r = v(2h/g)^{0.5}$$

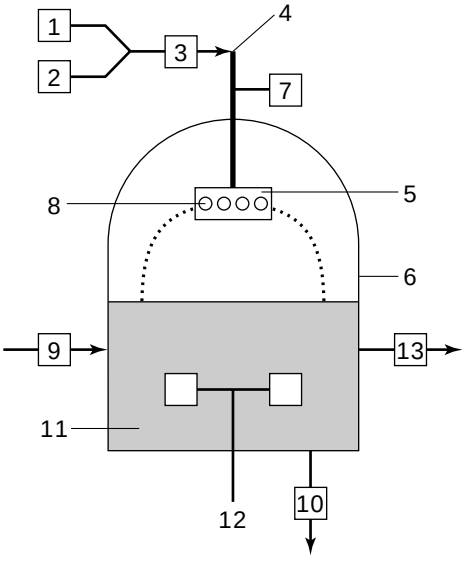


Figure 4. Schematic representation of the Biosphere (Landteknikk A/L): (1) biocatalyst solution, (2) polymer solution, (3) mixing chamber, (4) rigid tube, (5) perforated cylinder, (6) bioreactor, (7) velocity regulation unit, (8) nozzles, (9) feed pipe, (10) product outlet, (11) precipitation bath, (12) stirring device, and (13) reactor outlet.

Liquid-Jet-Based Methods

Dropping methods can be scaled up only by the parallel use of several needles, because drop formation depends on a relatively slow liquid velocity. Hence, dropping methods are limited to about 200 mL/h per needle. To overcome this problem, several methods for producing droplets from liquid jets have been developed, and several liters of microcapsules can be produced. The maximum productivity of these systems is limited due to the fact that the pressure needed to ensure flow increases with the second power of the liquid velocity, and with increasing liquid viscosity and decreasing orifice diameter. Pressure can be calculated using the following equations with pressure, P, resistance coefficient, χ, liquid density, ρ, liquid velocity, v, needle diameter, d, needle length, I, Reynold's number, Re, and the liquid viscosity, η:

$$\Delta P = (\chi \rho v^2/2)(l/d)$$

$$Re = v\rho d/\eta$$

and

$$\chi = 64/Re \text{ for } Re < 2{,}320 \text{ (laminar flow)}$$

or

$$\chi = 0.3164/Re^{0.25} \text{ for } 3{,}000 < Re < 10{,}000 \text{ (turbulent flow)}$$

or

$$\chi = 0.0054 + 0.3964/Re^{0.3} \text{ for } 20{,}000 < Re < 10^6 \text{ (turbulent flow)}$$

Dipping Jet. The dipping jet method is the simplest process for producing a large amount of microcapsules. It depends on the fact that a liquid jet at high velocity disintegrates into many drops when penetrating a fluid. Drop formation depends on irregular dynamic processes and leads to drops with a large size distribution (63,64). This method has been used to immobilize the fungi *Claviceps purpurea* in alginate particles for further production of alcaloids in a gas–solid fluidized-bed reactor. Researchers produced 110-L capsules with a mean diameter of 2.5 mm (65).

Vibrating Jet Breakage. Lord Raleigh influenced liquid jets with tuning forks in the late nineteenth century. He observed that vibration can force a liquid jet to disintegrate into equal drops if the wavelength of the vibration is greater than the perimeter of the liquid jet. He showed that the frequency, f, for maximum instability is related to the jet velocity, v, and the wavelength, λ (66):

$$f = v/\lambda$$

When a laminar-flowing jet is mechanically disturbed at this frequency, uniform drops are formed. The optimum wavelength, λ_{opt}, for breakup can be calculated with the needle diameter, d, dynamic viscosity, η, liquid density, ρ, and the surface tension, σ (67):

$$\lambda_{opt} = 2^{0.5}\,\pi d(1 + (3\eta/(\rho\sigma d)^{0.5}))^{0.5}$$

By considering that the jet will break into liquid cylinders with the wavelength λ, each of which will form uniform droplets, the diameter of the droplets, d, can be calculated:

$$d = (1.5\ d^2\lambda)^{1/3}$$

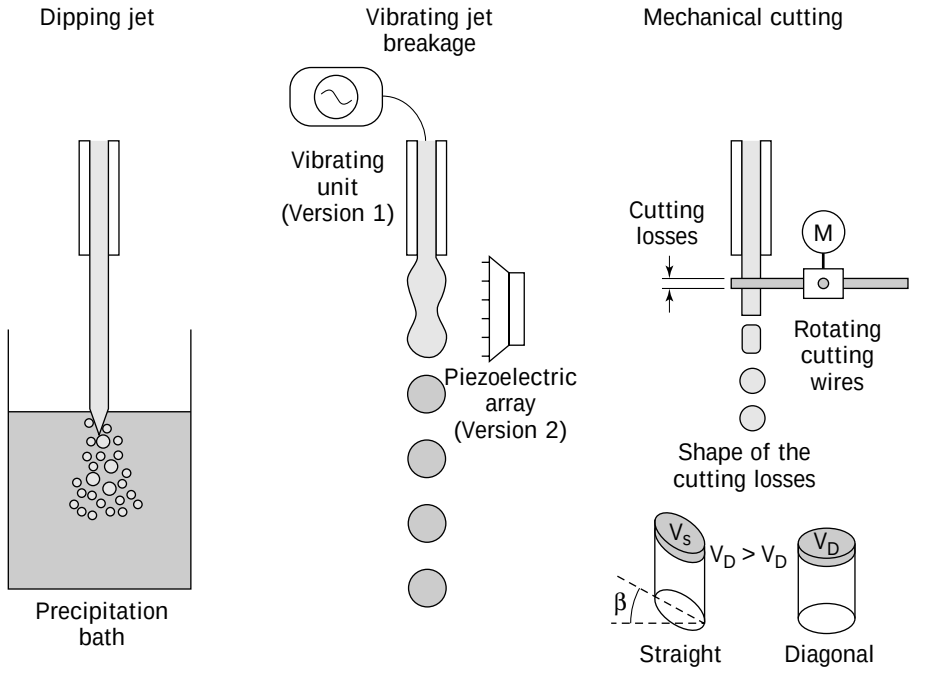


Figure 5. Liquid-jet-based extrusion technologies.

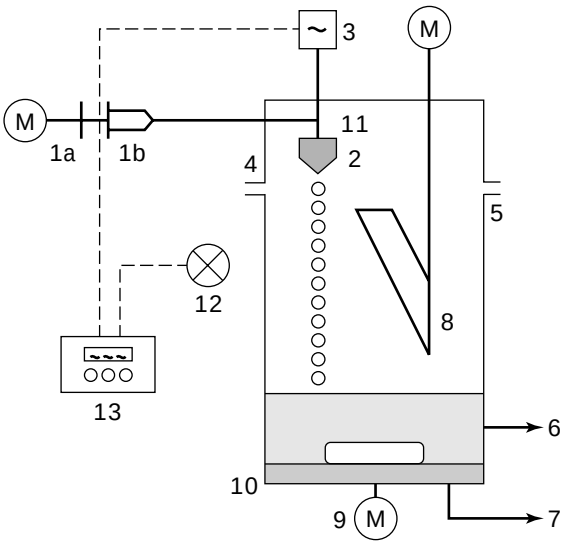


Figure 6. Schematic representation of the Encapsulator AP (INOTECH AG): (1a) syringe pump, (1b) syringe containing polymer solution and cell suspension, (2) vibrating nozzle, (3) vibration unit, (4) pulsation unit, (5) reaction vessel, (6) reactor outlet for capsule harvesting, (7) waste outlet, (8) bypass unit (used to eliminate undesired beads at the beginning and the end of bead formation), (9) magnetic stirrer, (10) sieve, (11) capillary, (12) stroboscope lamp, and (13) frequency generator.

Mechanical Cutting. The principle of the mechanical cutting encapsulation method is to cut a liquid jet with a rotating wire into a series of uniform liquid cylinders behind the cutting wire. Due to surface tension, the liquid cylinders form drops immediately as they fall down into the precipitation bath. The bead build volume, V, of the cylinder can be calculated from the cutting angle, β, which is a function of the liquid velocity, v, and the local velocity of the wire, w, at the cutting point. Other relevant parameters are the thickness of the wire, d_w, the number of revolutions, n, of the cutting wire, the number of wires, z, and the nozzle diameter, d.

$$V = 0.25\ \pi d^2((v/(nz)) - (d_w + d\ \sin\beta)/\cos\beta))$$
$$\text{with } \beta = \arctan(v/w)$$

Because the wire used for cutting cannot be infinitesimally small, and its size is also determined by the material, it slings away about 10–20% of the liquid jet; this can be considered lost material. This cutting loss, V_L, can be determined as follows:

$$V_L = 0.25\ \pi d^2(d_w/\cos\beta + d\ \tan\beta)$$

When the cutting device is arranged at an optimal angle, α, depending on the liquid jet's velocity, the the drop volume is affected, and cutting loss can be reduced further:

$$\alpha = \arcsin(v/w)$$
$$V = 0.25\ \pi d^2((v/nz) - d_w)$$
$$V_L = 0.25\ \pi d^2 d_w$$

The cutting loss can also be reduced by use of a thinner wire for cutting. However, if a stainless steel wire becomes thinner than 0.2 mm, distortion effects of the wire can be observed when using fluids of high viscosity. Therefore, special stabilized wires 0.06 mm in diameter have been designed; in combination with diagonal cutting, they reduce cutting loss to only 3.3% (68,69).

This method is especially suited to produce beads from high-viscosity polymer solutions, with high productivity rates of up to 20 kg/h for 1-mm beads. This method has been used to immobilize a variety of microorganisms for further use in wastewater treatment experiments.

Liquid–Liquid Emulsification

Microcapsules can be obtained from dispersion of one liquid into another followed by a membrane-forming process. Generally, dispersions can be obtained from static and dynamic mixers. However dynamic mixers are preferred for microencapsulation because of the need to stabilize the dispersion until microcapsule formation has taken place, otherwise coalescence leads to increasing drop size.

Dynamic Mixer. It is advantageous to use an impeller that disperses the lighter liquid in both the center and the periphery at the same impeller speed. The optimum impeller diameter, d, depends on the vessel diameter, D, and presence of baffles:

$$d = D/3 \text{ (impeller)}$$

or

$$d = 0.4\ D \text{ (impeller in baffled vessels)}$$

The minimum agitation speed, n, required to disperse a dispersed (index D) liquid into a continuous (index C) liquid by using a four-paddled stirrer centered in a flat-bottomed cylindrical vessel with a liquid height equal to the vessel diameter, D, can be calculated from the liquid densities, ρ, and the viscosity, η_H, of the continuous phase:

$$n = 750\ D^{-2/3}\ (\eta_c/\rho_c)^{1/9}\ ((\rho_c - \rho_D)/\rho_c)^{0.26}$$

The viscosity of the dispersed phase and the interfacial tensions can be neglected (70). Calculating the mean particle size is fairly difficult and depends on type and diameter of the stirrer, agitation rate, and the physical properties of the liquids and their relative amounts (71). Generally, droplet size and size distribution decreases with increasing agitator speed. Droplet sizes ranging from a few micrometers to several hundreds of micrometers have been reported. A variety of correlations have been published relating the Sauter mean diameter, d_{32}, of the droplets to the vessel geometry and the physical properties of the dispersion system. The most well known correlation is derived from the impeller diameter, d_I and speed, n, the Weber number, N_{We}, liquid density, ρ, interfacial tension, η, the dispersed phase volume fraction, φ, and two correlation factors, b and c (72):

$$d_{32}/d_1 = b(1 + c\varphi)\ N_{We}^{-0.6}$$

$$N_{We} = \rho n^2 d_1^3/\eta$$

Due to the considerable internal sheer forces applied during dispersion, this method is especially suited for microencapsulation of inanimated matter (73–75) or viruses and microorganisms (76,77), but it can also be adjusted to the needs of cell cultures (78,79).

SELECTED MICROCAPSULE-FORMING MATERIALS

There are many microcapsule-forming materials available, indicating that there is no such thing as an all-purpose material. Hence, various materials have been proposed as carriers. They are based on either natural polymers, such as the polysaccharides alginate, carrageenan, agarose (80), and chitosan (81,82), and proteins such as gelatin (83,84) and collagen, or synthetic polymers such as polyacrylates, polyurethane and its less-toxic derivates (85), and polyethers. Each material has its pros and cons with respect to physicochemical stability, biodegradability or degradation into units small enough for renal excretion (86,87), and influence on the entrapped cells. Sulfated polyanions, for example, have been successfully used to obtain and maintain stable single-cell suspensions of insect cells (88), whereas these polymers exhibit the opposite effect: the formation of multicellular spheroids on mammalian cells (89). Technical considerations such as simplicity of preparation, availability, and cost also have to be taken into account.

Calcium Alginate

Alginic acid is a polyuronic acid extracted from brown seaweeds and some bacteria, mostly derived from the genus *Pseudomonas*. Chemically, alginates are block copolymers of 1,4-linked β-D-mannuronic acid (M) and α-L-guluronic acid (G). The mannuronate residuals of the bacterial alginate are acetylate to some extent at positions O-2 and/or O-3. The residues, occuring in varying proportions, depend on the source, and the residues are arranged in block patterns comprising homopolymeric regions (MM blocks and GG blocks) linked with alternating regions of heteropolymeric regions (MG blocks). Divalent cations, such as Ca^{2+}, bind preferentially between GG blocks, which appear to have preferred binding sites. Therefore, the composition (20–75% G content) and molecular size of the polymer impacts on the porosity, stability, gel strength, and swelling behavior (90–93).

Interfacial polymerization of alginate and Ca^{2+} is instantaneous, with precipitation of calcium alginate followed by a more gradual gelation of the interior as calcium ions penetrate through the alginate. During formation of gel beads, the biopolymer localized at the bead center moves gradually to the bead surface by attracting the ionotropic cations, which induce gelation. Consequently, the immobilized material flows with the macromolecules to the periphery of the bead, producing a superficial crust. At the same time the bead volume is reduced due to cross-linking of the macromolecules and expulsion of water. The interior of the bead consists of a fluffy core surrounded by a medium gelled zone containing supramolecular structures such as shafts and cavities. These structural inhomogeneities originate from gradients of local concentrations of components during the gelling process or the presence of chemically inert substances such as PEG. They remarkably affect the strength of beads and their permeability for low and high molecular weight components (94,95).

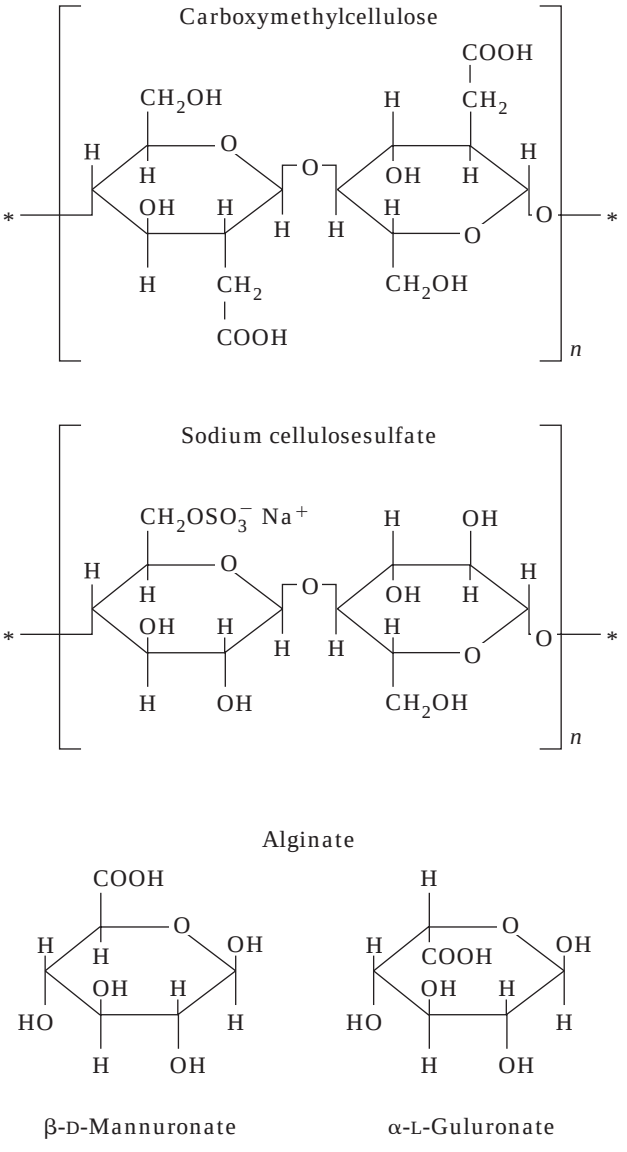


Figure 7. Polysaccharides used for encapsulation (I).

Calcium alginate beads are very stable for long periods and can tolerate a pH range from 3 to 10 at 25 °C. Thermal stability is also good, and beads can withstand incubation temperatures of up to 85 °C at pH 7.0. Destabilization of calcium alginates is promoted by chelating agents, which remove divalent cations. Phosphate, citrate, or EDTA should be avoided. The breaking of beads during fermentation of microorganisms is mainly caused by low initial alginate and/or high initial cell concentration, whereas the length of incubation has only a minor impact (96,97).

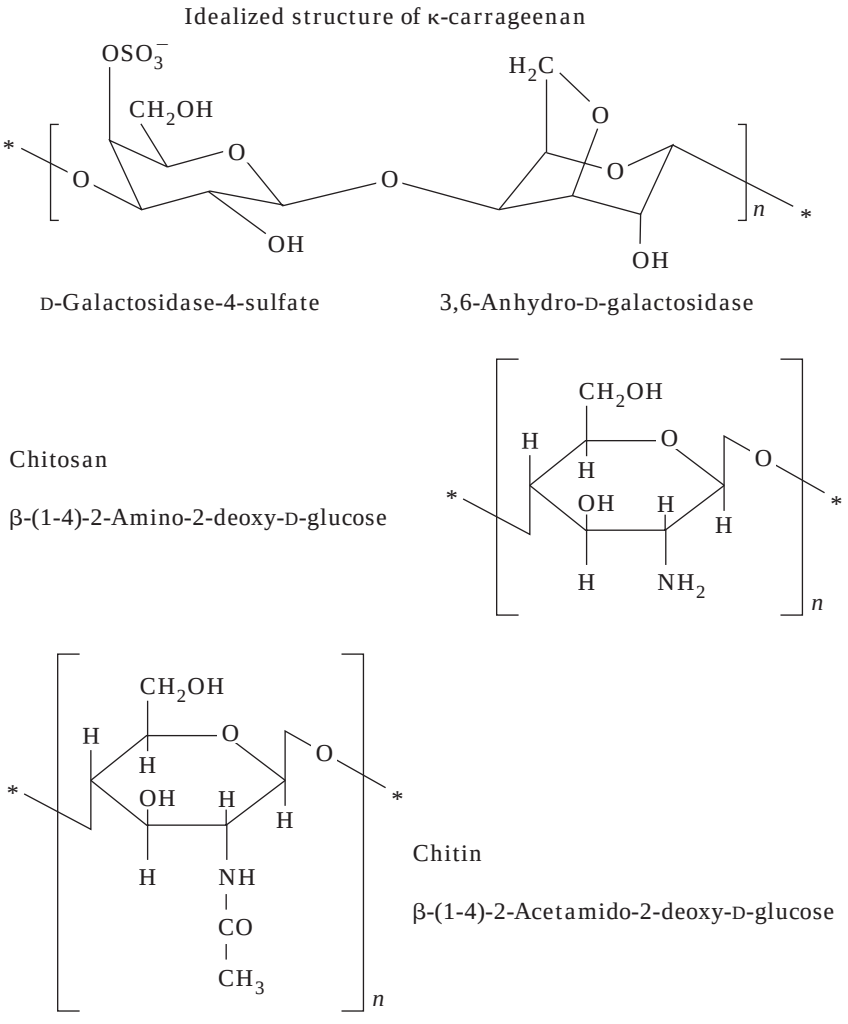


Figure 8. Polysaccharides used for encapsulation (II).

Microencapsulation in alginate beads is an inexpensive, versatile, and nontoxic method. It is used in food process industries to enhance fermentation productivity and cell stability. Costs for downstream processing are reduced by facilitating cell recovery and recycling (98). In the beverage industry, several processes including several different microorganism species have been developed for producing beer (99), wine (100), cider, (101) and ethanol (102).

Coating with Poly-L-Lysine

To prevent cell leakage from alginate beads in general and to provide immobilized cells with a semipermeable immunisolation barrier in particular, the coating of alginate beads with poly-L-lysine (PLL) has been introduced. This system is currently the best investigated and by far the most widespread microencapsulation system used in medicine.

First, viable cells are immobilized within an alginate bead. The bead is washed and transferred to another precipitation bath containing a PLL solution. Because the alginate is a negatively charged polymer, simplex membrane formation with the positively charged PLL is instant, decreasing the molecular weight cut-off of the system to about 100 kDa. Finally, the now-coated capsules are washed again and transferred to another precipitation bath containing a dilute solution of more oppositely charged polymers, with respect to the outer PLL surface. This final polyanion may be PEI (103) or alginate (104). The purpose of the additional coating is to neutralize nonreacted PLL and to mask the PLL that otherwise might provoke unwanted immune or fibrotic response. Because the outmost layer consist of alginate, purification of the alginate prior to use improves biocompatibility. Bioincompatibility is influenced by the chemical composition and, to a greater extent, by individual imperfections of the capsules; it has been clearly shown that fibrotic response against implanted capsules is associated with the presence of PLL around the capsules (105,106).

The stability of the PLL layer is governed mainly by the composition of the alginate, where alginate with a high M content ensures both more rapid binding of the PLL during the coating step and much thicker PLL layers. Because the reaction between the polymers is very fast, the rate-limiting step is the diffusion of PLL into the alginate gel. The free diffusion of the PLL is determined by gel porosity and the relative molecular weight of the polycation. Therefore the PLL layer can be improved through an increased concentration of PLL and by using a lower molecular

weight sample of PLL. However, a certain minimum polymer length is necessary to obtain stable capsules, so optimum PLL molecular weight is believed to be around 18 kDa (107,108).

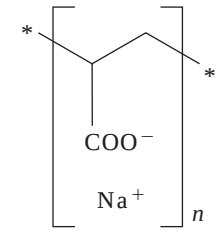
Improper performance of the microcapsules can be caused by a distorted coating process due to the presence of living cells on the surface area of the alginate bead. As previously pointed out, immobilized material has the tendency to be concentrated in the outer areas of alginate beads. The consequence is that small alginate beads are even more liable to improper coating caused by protruding cells than large beads, while cells in large beads are more likely to suffer from mass transport limitations. The additional entrapment of microspheres within a larger sphere can cover undesirable cell protuberances and debris integrated or exposed on the surface of the microsphere with a complete spherical shell of calcium alginate prior to the following coating process (109). Because only a small fraction of the alginate may diffuse throughout the coacervate membrane, the capsules are liable to osmotic swelling under physiological conditions because the cross-linking divalent ions of the bead are exchanged with monovalent ions. This swelling can also lead to capsule breakage. Strategies to reduce swelling include strengthening of the PLL layer by simply letting more polycations bind to the capsule surface, exchanging of a fraction of the calcium ions with ions with higher affinity toward the G blocks such as strontium or barium, using more alternate PLL–alginate layers, and decreasing capsule diameter. PLL-coated alginate capsules have been successfully used to treat diabetes in humans for periods of over one year (110–114).

Polyelectrolytes

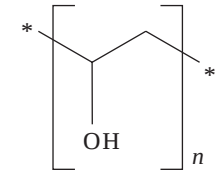
Polyelectrolytes are polymers that contain a charged group in their monomer repeating unit. They have been used as a flocculent in water purification, protein precipitation, and membrane technology. When brought into contact, aqueous solutions of oppositely charged polyelectrolytes immediately form water-insoluble but highly swollen precipitates or membranes. Membrane formation is completed within minutes.

Although there are is a vast number of possible polyanion polycation combinations, only a minority can be used for microencapsulation. For the encapsulation of living cells, it is strongly recommended that a polyanion be used for the inner polymer. This is in contact with the cells during capsule formation. The reason for this is that polycations are cytotoxic, especially those that contain quarternary ammonium groups such as PDADMAC. Polycations interact with the negatively charged phosphate groups of the cytoplasmic membrane, which finally leads to membrane rupture and cell death. This has been extensively studied with liposomes (115). Additionally, polycations form stable complexes with DNA and proteins (116,117). Membrane-bound polycations do not show these effects at all.

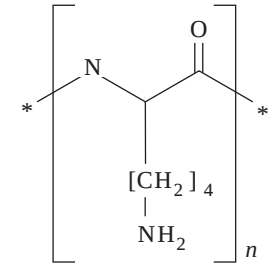
The best-suited polyanions to form a cell suspension containing drop are highly charged polysaccharide derivates with a molar mass in the range of 10^6 Da, with a cyclic backbone and pH-dependent charge. Typical examples are cellulose sulfate, carboxymethyl cellulose, and alginate. The corresponding polycation should contain either permanently charged quarternary amonium groups (PDADMAC) or a tertiary amine. The charge density should be between 50 and 100 mol%. The molecular mass may range from 10^3 to 10^6 Da, with high molecular weight species requiring a higher charge density to compensate for the lower diffusion coefficient (118–121).



Poly(acrylic acid)



Poly(vinyl alcohol)



Poly(L-lysine)

Figure 9. Polymers used for encapsulation.

Microcapsules from CS and PDADMAC exhibit immense mechanical stability up to 7 N per 3-mm capsule (122). About 0.3 N is required for this type of capsule to withstand the stress occurring in a heavily stirred tank reactor. The capsules typically possess a thin but very compact primary layer (a few microns) followed by a less densely packed region joining the inner layers of high porosity. The membrane properties mainly depend on the primary layer; therefore distortion of the membrane-forming reaction leads to microcapsules with very different properties. The molecular weight cut-off can be finely tuned according to the needs of the application. It can be as low as 5,500 Da for complete product recovery (123) or immunoprotection (40,124), in the range of 4.4×10^5 to 2.0×10^6 Da for continuous production of monoclonal antibodies

(53,125), or within the range of several microns to allow free virus passage for some gene therapy (126) and insect cell culture (54) applications.

Poly(Vinyl Alcohol)

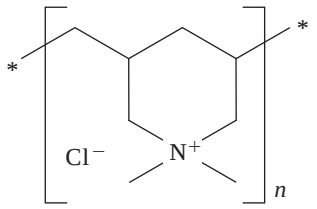
An aqueos poly(vinyl) alcohol (PVA) can be cross-linked through hydrogen bonding by a freeze–thaw process with hydroxyl groups on PVA molecules. The typical freezing temperature is around −20 °C. The mechanical strength of the cryogel greatly depends on the thawing rate, and shear modulus increases with decreasing thawing rate (preferably slower than 0.2 °C/min). During slow thawing, the system resides in the temperature range from −4 °C to −1 °C for a long time, and numerous hydrogen bonds are linked. It is important to consider that such interactions occur relatively slowly due to the high viscosity of the liquid microphase and the low temperature of the frozen system. Stability can be further increased by using a more highly concentrated PVA solution and storage in a frozen state. The ability of the gel to melt depends strongly on the preparation process. Thermoresistance is excellent, and immobilized thermophiles in PVA beads have operated at 65 °C for a period of 1 year and longer (127).

As a result of the formation of ice crystals during freezing, the cryogel has a high micro- and macroporous structure, which is necessary for nonhindered mass transport of any solutes for cultivation purposes. The polymer itself is nontoxic and biocompatible and is not degraded by most enzymes and microbes (128). The gel beads are stable in any biotechnology liquid media; swelling or softening can occur only under harsh conditions such as extreme pH or high concentration of chaotropic substances such as urea. Physiological concentrations of most compounds do not affect the gel or cause it to shrink, which increases stability (129,130). Apart from the freeze–thaw method, PVA can be cross-linked chemically by additions of compounds such as aldehydes or borates to form a hydrogel. Chemical cross-linking creates a gel of a comparable low strength; it can be described as having an appearance and strength similar to common gelatine (131). It is a very suitable method and, with the exception of the cross-linking reaction, nontoxic biomaterial has been studied extensively as membranes in artificial kidneys and as materials for skin replacement and contact with blood (132). It is therefore especially suited for coating beads. Microcapsules from PVL cryogels seem to be especially useful for use in wastewater treatment for removal of nitrogen (133,134), heavy metals (135), and environmental decontamination (136,137).

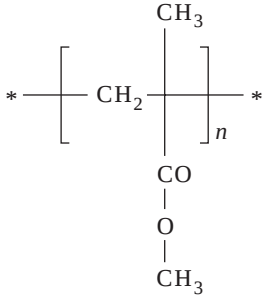
Poly(Methylmethacrylate-Hydoxyethylmethacrylate)

Poly(methylmethacrylate-hydroxyethylmethacrylate) (HEMA-MMA), which is 75–80% HEMA and 25–20% MMA, is a nonionic random copolymer. The polymer is water insoluble, which makes encapsulation more difficult. However water insolubility is considered an advantage because this is presumed to be a key factor leading to greater biocompatibility.

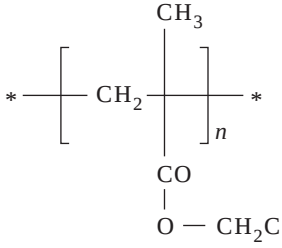
The cell suspension and the polymer are mixed together in a biocompatible solvent such as polyethylene glycol (PEG), and droplets are formed by jet extrusion or emulsification in another liquid. The membrane formation is initiated upon removal of the polymer solvent. This can be achieved easily by transferring the droplets in a phosphate-buffered saline (PBS) precipitation bath. The PEG is also soluble in the PBS solution, thus rapidly depleting the polymer solvent at the interface of the droplets, leading to a dense outer layer that provides all the resistance to the mass transfer. Below the dense layer are macroporous "trabecular" layers, similar to asymmetric phase-intervention membranes. Typically the membrane thickness ranges from 20 to 150 μm, with an average of 60 μm. The molecular weight cut-off of the membrane is above 52 kDa, possibly above 340 kDa, but can be influenced by the polymer composition, whereas the permeability of the membrane increases with higher HEMA content of the polymer and when the solubility parameters of the solvent and the precipitate are both high. Furthermore, the precipitation media influences the permeability of the microcapsule membrane with Tris-buffered glycerol, leading to denser



Poly(diallydimethylammonium chloride)



Poly(methylmethacrylate)



Poly(2-hydroxyethylmethacrylate)

Figure 10. Synthetic polymers used for encapsulation.

membranes that reflect the slower rate of precipitation in PBS or unbuffered glycerol. The membrane-forming process is not absolutely reliable with respect to complete membrane formation, and procedures to screen and remove capsules containing pinholes or flaws might be necessary for certain applications (138–140).

BIBLIOGRAPHY

1. A. Groboillot, D.K. Boadi, D. Poncelet, and R.J. Neufeld, *Crit. Rev. Biotechnol.* **14**(2), 75–107 (1994).
2. Z.M. Sun, D. Poncelet, J. Conway, and R.J. Neufeld, *J. Microencapsulation* **12**(5), 495–504 (1995).
3. V.A. Santos, J. Tramper, and R.H. Wijffels, *Biomater. Art. Cells Immob. Biotech.* **21**(3), 317–322 (1993).
4. Y. Andrès, H.J. MacCordick, and J.C. Hubert, *Appl. Microbiol. Biotechnol.* **44**, 271–276 (1995).
5. N.G. Wilson and G. Bradley, *J. Appl. Bacteriol.* **80**, 99–104 (1996).
6. E. Poscillico, *Biotechnology* **4**, 114–117 (1986).
7. U.S. Pat. 5,116,747 (May 26, 1992), M. Moo-Young, N.C. Bols, S.E. Overgaard, and J.M. Scharer (to University of Waterloo).
8. Y. Shirai, R. Sasaki, K. Hashimoto, H. Hitomi, and H. Chiba, *Appl. Microbiol. Biotechnol.* **29**, 544–549 (1988).
9. M. Seki and S. Furusaki, *Chemtech* **26**(3), 41–45 (1996).
10. K. Redenbaugh, *Synseeds: Applications of Synthetic Seeds to Crop Improvement*, CRC Press, Boca Raton, Fla., 1993, pp. 3–7.
11. J.I. Shigeta, *Biotechnol. Techn.* **9**(10), 771–776 (1995).
12. Y. Bachiri, C. Gazeau, J. Hansz, C. Morisset, and J. Dereuddre, *Plant Cell Tissue and Organ Culture* **43**, 241–248 (1995).
13. F. Lim and A.M. Sun, *Science* **210**, 908–910 (1980).
14. U.S. Pat. 5,578,314 (November 26, 1996), K.C. Cochrum, R.E. Dorian, and S.A. Jemtrud (to the Regents of the University of California).
15. I.T. Tai and A.M. Sun, *FASEB J.* **7**, 1061–1069 (1993).
16. P. Aebischer, N.A.-M. Pochon, B. Heyd, N. Déglon, J.-M. Joseph, A.D. Zurn, E.E. Baetge, J.P. Hammang, M. Goddard, M. Lysaght, F. Kaplan, A.C. Kato, M. Schluep, L. Hirt, F. Porchet, and N. De Tribolet, *Hum. Gene Ther.* **7**, 851–860 (1996).
17. E.B. Reilly, G. Antognetti, J.S. Wesolowski, Jr., and P. Sakorafas, *J. Immunol. Methods* **126**, 273–279 (1990).
18. S.R. Win, P.A. Tresco, B. Zielinski, L.A. Greene, C.B. Jaeger, and P. Aebischer, *Exp. Neurol.* **113**, 322–329 (1991).
19. L. Christenson, D.F. Emerich, and P.R. Sanberg, *Mov. Disord.* **7**, 185–186 (1995).
20. J.K. Embleton and B.J. Tighe, *J. Microencapsulation* **10**(3), 341–351 (1993).
21. K.A. Brown, C.A. Moser, T.J. Speaker, C.A. Khoury, J.E. Kim, and P.A. Offit, *J. Infect. Dis.* **171**, 1334–1338 (1995).
22. D. Chickering, J. Jacob, and E. Mathiowitz, *Biotechnol. Bioeng.* **52**(1), 96–101 (1996).
23. B. Katzbauer, M. Narodoslawsky, and A. Moser, *Bioprocess Eng.* **12**, 173–179 (1995).
24. C. Cokmus and Y.M. Elçin, *Pestic. Sci.* **45**, 351–355 (1995).
25. PCT WO 93/15190 (August 5, 1993), M. Lysberg, I. Storrø, H. Svendsen, O. Eide, and B. Oterholm (to Norske Meierier).
26. U.S. Pat. RE28779 (April 20, 1976), S. Katayama, H. Matsukawa, J. Matsuyama, and M. Yamamoto (to Fuji Foto Film Co., Ltd.).
27. R. Pommersheim, J. Schrezenmeir, and W. Vogt, *Macromol. Chem. Phys.* **195**, 1557–1567 (1994).
28. M.L. Huguet and E. Dellacherie, *Process Biochem.* **31**(8), 745–751 (1996).
29. U.S. Pat. 5,410,016 (April 25, 1995), J.A. Hubbell, C.P. Pathak, A.S. Sawhney, and N.P. Hill (to Board of Regents, the University of Texas System).
30. U.S. Pat. 5,464,932 (November 7, 1995), H.R. Allcock, C.G. Cameron, and D.E. Smith (to the Penn State Research Foundation).
31. U.S. Pat. 5,529,914 (June 25, 1996), J.A. Hubbell, C.P. Pathak, A.S. Sawhney, N.P. Desai, and S.F.A. Hossainy (to the Board of Regents, the University of Texas Systems).
32. U.S. Patent 4,352,883 (October 5, 1978), F. Lim (to Damon Corp.).
33. F. Lim and A. Sun, *Science* **210**, 908–910 (1980).
34. P. Soon-Shiong, P. Sandford, R. Heintz, T. Zheng, Q. Yao, Z. Yao, N. Meredith, M. Schmehl, M. Moloney, and M. Murphey, *Trans. Soc. Biomater.* **356** (1994).
35. P. Soon-Shiong and P. Sandford, *Surg. Technol. Int. IV*, 93–99 (1995).
36. H. Dautzenberg, F. Loth, K. Fechner, B. Mehlis, and K. Pommering, *Macromol. Chem. Suppl.* **9**, 203–210 (1985).
37. P. Spiekermann, K.-D. Vorlop, and J. Klein, *Proc. 4th Eur. Cong. Biotechnol.* **3**, 590–593 (1987).
38. J.E. Tyler, *Bioprocess Technol.* **10**, 343–361 (1990).
39. L.C. Colombi, *Feed Mix* **4**, 12–16 (1995).
40. J. Stange, S. Mitzner, H. Dautzenberg, W. Ramlow, M. Knippel, M. Steiner, B. Ernst, R. Klinkmann, and H. Klinkmann, *Biomater. Art. Cells Immob. Biotechnol.* **21**, 343–352 (1993).
41. R.M. Saller, J. Stange, S. Mitzner, K. von Rombs, P. Hutzler, H. Dautzenberg, W.H. Güntzburg, and B. Salmons, in H. Dautzenberg and D. Poncelet eds., *Bioencapsulation Research Group, International Workshop, Bioencapsulation V: From Fundamentals to Industrial Applications*, Potsdam, Germany, September 23–25, 1996.
42. D. Hunkeler, A. Prokop, S. DiMari, M.A. Haralson, and T.G. Wang, in H. Dautzenberg and D. Poncelet eds., *Bioencapsulation Research Group, International Workshop, Bioencapsulation V: From Fundamentals to Industrial Applications*, Potsdam, Germany, September 23–25, 1996.
43. T. Wang, I. Lacík, M. Brissová, A.V. Anilkumar, A. Prokop, D. Hunkeler, R. Green, K. Shahrokhi, and A. Powers, *Nat. Biotechnol.* **15**, 358–362 (1997).
44. U.S. Pat. 5,053,332 (October 1, 1991), R.B. Cook, R.B. Provonchee, and S. Nochumson.
45. U.S. Pat. 5,160,745 (November 3, 1992), P.P. DeLuca and F. Rypacek (to the University of Kentucky Research Foundation).
46. D. Poncelet, C. Dulieu, and R.J. Neufed, in D. Poncelet and V. Santos eds., *Bioencapsulation Research Group, International Workshop, Bioencapsulation VI: From Fundamentals to Industrial Applications*, Barcelona, Spain, August 30–September 1, 1997.
47. W.D. Harkins and F.E. Brown, *J. Am. Chem. Soc.* **41**, 499 (1919).
48. M.C. Levy and D. Poncelet, *Biofuture* **132**, 16–21 (1994).
49. W. Lane, *Rev. Sci. Instrum.* **24**, 98–101 (1947).
50. U.S. Pat. 4,814,274 (March 21, 1989), T. Shioya and Y. Shinoki (to Snow Brand Milk Products Co., Ltd.).
51. M. Van Dyke, *An Album of Fluid Motion*, Parabolic Press, Stanford, 1982, p. 86.
52. U.S. Pat. 5,521,079 (May 28, 1996), R.E. Dorian and K.C. Cochrum (to the Regents of the University of California).

53. M.G. Cho, *Verfahrenstechnische Auslegung einer Apparatur zur Herstellung mikroverkapselter Biokatalysatoren mit getrennter Zuführung von Katalysatorlösung und Kapselgrundsubstanz,* VDI Fortschrittsberichte, VDI Verlag, Düsseldorf, 1994, pp. 93–102.
54. H. Hübner and R. Buchholz, in K. Maramorosch and J. Mitsuhashi eds., *Invertebrate Cell Culture: Novel Directions and Biotechnology Applications,* Science Publishers, Enfield, N.H., 1997, pp. 123–130.
55. C. Kloth, C. Riese, M. Cho, R. Buchholz, and U. Marx, in E.C. Beuvery, J.B. Griffiths, and W.P. Zeijlemaker eds., *Animal Cell Technology: Developments towards the 21st Century,* Kluwer Academic, Dordrecht, The Netherlands, 1995, pp. 765–769.
56. B. Bugarski, B. Amsden, R.J. Neufeld, D. Poncelet, and M.F.A. Goosen, *Can. J. Chem. Eng.* **72**, 517–521 (1994).
57. B. Bugarski, Q. Li, M.F.A. Goosen, D. Poncelet, R.J. Neufeld, and G. Vunjak, *MchE J.* **40**, 1026–1031 (1994).
58. B.R.S. Hsu, Y.S. Ho, S.H. Fu, Y.Y. Huang, S.C. Chiou, and H.S. Huang, *Transplant. Proc.* **27**(6), 3227–3231 (1995).
59. B.G. Amsden and M.F.A. Goosen, *J. Controlled Release* **43**, 183–196 (1997).
60. R.P. Fraser and P. Eisenklamm, *Trans. Inst. Chem. Eng.* **34**, 294–319 (1956).
61. H. Brauer, *Grundlagen der Einphasen- und Mehrphasenströmung,* Verlag Sauerländer, Frankfurt, 1971, pp. 272–352.
62. Eur. Pat. EP0446134 (September 11, 1991), J.-P. Renaudeaux and J.-M. Chicheportiche (to Government of France, represented by Delegue General pour l'Armement).
63. W. Ohnesorge, *Z. angew. Math. Mech.* **16**(6), 355–358 (1936).
64. R.M. Christiansen and A.N. Hixon, *Ind. Eng. Chem.* **49**(6), 1017–1024 (1957).
65. K. Merkel, *Vergleichende Untersuchungen verschiedener Kultivierungssysteme zur Produktion von Ergotalkaloiden mit* Claviceps purpurea *Dl-nic,* VDI Fortschrittsberichte, VDI Verlag, Düsseldorf, 1996, pp. 70–73.
66. Lord Raleigh, *Proc. Lond. Math. Soc.* **10**(4), 4–13 (1878).
67. C. Weber, *Z. angew. Math.* **11**, 136 (1936).
68. Ger. Pat. DE 4,424,998 (July 20, 1996), K.-D. Vorlop and J. Breford.
69. U. Prüße, B. Fox, M. Kirchhoff, F. Bruske, J. Breford, and K.-D. Vorlop, *Chem. Eng. Technol.* **21**(1), 29–33 (1998).
70. S. Nagata, *Mixing: Principles and applications,* John Wiley & Sons, New York, 1975, pp. 297–315.
71. A. Mersmann and H. Grossmann, *Chem. Ing. Technol.* **52**(8), 621–628 (1980).
72. E.G. Chatzi, A.D. Gavrielidis, and C. Kiparissides, *Ind. Eng. Chem. Res.* **28**, 1704 (1989).
73. U.S. Pat. 3,869,406 (March 4, 1975), H. Matsukawa and K. Saeki (to Fuji Foto Film Co., Ltd.).
74. U.S. Pat. 4,822,535 (April 18, 1989), B. Ekman and A. Lindahl (to Norsk Hydro A.S.).
75. U.S. Pat. 5,271,961 (December 21, 1993), E. Mathiowitz, H. Bernstein, E. Morrel, and K. Schwaller (to Alkermes Controlled Therapeutics, Inc.).
76. U.S. Pat. 5,500,161 (March 19, 1996), A.K. Andrianov and R.S. Langer (to Massachusetts Institute of Technology and Virus Research Institute).
77. U.S. Pat. 4,774,178 (September 27, 1988), P. Egerer, W. Haese, H. Perrey, and G. Schmidt-Kastner (to Bayer Aktiengesellschaft).
78. U.S. Pat. 4,353,888 (October 12, 1982), M.V. Sefton.
79. U.S. Pat. 4,647,536 (March 3, 1987), K. Mosbach and K. Nilsson.
80. A.C. Jen, M.C. Wake, and A.G. Mikos, *Biotechnol. Bioeng.* **50**, 357–364 (1996).
81. S.-K. Kim and C.K. Rha, in G. Skjåk-Bræk, T. Anthonson, and P. Sandford eds., *Chitin and Chitosan,* Elsvier Applied Science, London, 1990, pp. 617–626.
82. C.-H. Chu, H. Kumagai, and K. Nakamura, *J. Appl. Polym. Sci.* **60**, 1041–1047 (1996).
83. E. Esposito, R. Cortesi, and C. Nastruzzi, *Biomaterials* **17**, 2009–2020 (1996).
84. H. Akin and N. Hasirci, *J. Appl. Polym. Sci.* **58**, 95–100 (1995).
85. M.Z.-C. Hu and M. Reeves, *Biotechnol. Prog.* **13**, 60–70 (1997).
86. S. Pulapura and J. Kohn, *J. Biomater. Appl.* **6**, 216–250 (1992).
87. B. Jeong, Y.H. Bae, D.S. Lee, and S.W. Kim, *Nature* **388**, 860–862 (1997).
88. K.U. Dee, M.L. Shuler, and H.A. Wood, *Biotechnol. Bioeng.* **54**, 191–205 (1997).
89. W. Handschack, J. Kopp, and C. Pitra, *Acta Biol. Med. Germ.* **36**, 1429–1434 (1977).
90. G. Skjåk-Bræk, H. Grasdalen, and B. Larsen, *Carbohydr. Res.* **154**, 239–250 (1986).
91. O. Smidsrød, *Faraday Discuss. Chem. Soc.* **57**, 263–274 (1974).
92. A. Martinsen, G. Skjåk-Bræk, and O. Smidsrød, *Biotechnol. Bioeng.* **33**, 79–89 (1989).
93. B.H.A. Rehm and S. Valla, *Appl. Microbiol. Biotechnol.* **48**, 281–288 (1997).
94. U.S. Pat. 5,175,093 (December 29, 1992), D.B. Seifert (to Leight University of Bethlehem).
95. J.E. Nava Saucedo, C. Roisin, and J.-N. Barbotin, in R.H. Wijffels, R.M. Buitelaar, C. Bucke, and J. Tramper eds., *Immobilized Cells: Basics and Applications,* Elsevier Science, B.V., 1996, pp. 39–46.
96. J.C. Ogbonna, Y. Amano, and K. Nakamura, *J. Ferment. Bioeng.* **67**(2), 92–96 (1988).
97. O. Smidsrød and G. Skjåk-Bræk, *Trends Biotechnol.* **8**, 71–78 (1990).
98. A. Grobillot, D.K. Boadi, D. Poncelet, and R.J. Neufeld, *Crit. Rev. Biotechnol.* **14**(2), 75–107 (1994).
99. C.A. Masschelein, D.S. Ryder, and J.-P. Simon, *Crit. Rev. Biotechnol.* **14**(2), 155–177 (1994).
100. C.R. Davis, D.J. Wibowo, T.H. Lee, and G.H. Fleet, *Appl. Environ. Microbiol.* **51**(3), 539–545 (1986).
101. A. Durieux, X. Nicolay, J.-M. Jourdain, C. Depaepe, A. Pietercelie, A.-M. Plaisant, and J.-P. Simon, in D. Poncelet and V. Santos eds., *Bioencapsulation Research Group, International Workshop, Bioencapsulation VI: From Fundamentals to Industrial Applications,* Barcelona, Spain, August 30–September 1, 1997.
102. B.J.M. Hannoun and G. Stephanopoulos, *Biotechnol. Prog.* **6**, 341–345 (1987).
103. U.S. Pat. 4,407,957 (October 4, 1983), F. Lim (to Damon Corporation).
104. U.S. Pat. 4,673,566 (June 16, 1987), M.F.A. Goosen, G.M. O'Shea, and A.M. Sun (to Connaught Laboratories Limited).
105. P. De Vos, B. De Haan, and R. Van Schilfgaarde, *Biomaterials* **18**, 273–278 (1997).

106. P. De Vos, B.J. De Haan, G.H.J. Wolters, J.H. Strubbe, and R. Van Schilfgaarde, *Diabetologica* **40**, 262–270 (1997).
107. M.F.A. Goosen, G.M. O'Shea, H.M. Charapetian, S. Chu, and A.M. Sun, *Biotechnol. Bioeng.* **27**, 146–150 (1984).
108. B. Thu, P. Brunheim, T. Espevik, O. Smidsrød, P. Soon-Shiong, and G. Skjåk-Bræk, *Biomaterials* **17**, 1031–1040 (1996).
109. H. Wong and T.M.S. Chang, *Biomat. Art. Cells Immob. Biotechnol.* **19**(4), 687–697 (1991).
110. H.A. Clayton, R.F.L. James, and N.J.M. London, *Acta Diabetol.* **30**, 181–189 (1993).
111. P. Soon-Shiong and P.A. Sandford, *Surg. Technol. Int. IV* 93–99 (1995).
112. D.E. Awrey, M. Tse, G. Hortelano, and P.L. Chang, *Biotechnol. Bioeng.* **52**, 472–484 (1996).
113. B. Thu, P. Brunheim, T. Espevik, O. Smidrød, P. Soon-Shiong, and G. Skjåk-Bræk, *Biomaterials* **17**, 1069–1079 (1996).
114. P. De Vos, B.J. De Haan, J. Pater, and R. Van Schilfgaarde, *Transplantation* **62**, 893–899 (1996).
115. V.A. Kabanov, A.A. Yaroslavov, and S.A. Sukhishvili, *J. Controlled Release* **39**, 173–189 (1996).
116. V.A. Kabanov, in P. Dubin, J. Bock, R. Davis, D.N. Schulz, and C. Thies eds., *Macromolecular Complexes in Chemistry and Biology*, Springer-Verlag, Berlin, 1994, pp. 151–174.
117. J. Xia and P.L. Dubin, in P. Dubin, J. Bock, R. Davis, D.N. Schulz, and C. Thies eds., *Macromolecular complexes in chemistry and biology*, Springer-Verlag, Berlin, 1994, pp. 247–271.
118. D.J. Burgess, in P. Dubin, J. Bock, R. Davis, D.N. Schulz, and C. Thies eds., *Macromolecular Complexes in Chemistry and Biology*, Springer-Verlag, Berlin, 1994, pp. 285–300.
119. H. Dautzenberg, G. Arnoldt, B. Tiersch, B. Lukanoff, and U. Eckert, *Prog. Colloid. Polym. Sci.* **101**, 149–156 (1996).
120. D. Hunkeler, A. Prokop, S. DiMari, M.A. Haralson, and T.G. Wang, in H. Dautzenberg and D. Poncelet eds., *Bioencapsulation Research Group, International Workshop, Bioencapsulation V: From Fundamentals to Industrial Applications*, Potsdam, Germany, September 23–25, 1996.
121. R. Darvari and V. Hasirci, *J. Microencapsulation* **13**(1), 9–24 (1996).
122. H. Dautzenberg, B. Lukanoff, U. Schuldt, and B. Tiersch, in D. Poncelet and V. Santos eds., *Bioencapsulation Research Group, International Workshop, Bioencapsulation VI: From Fundamentals to Industrial Applications*, Barcelona, Spain, August 30–September 1, 1997.
123. H. Dautzenberg, F. Loth, K. Fechner, B. Mehlis, and K. Pommering, *Makromol. Chem. Suppl.* **9**, 211–217 (1985).
124. K.L. Lafferty, S.E. Prowse, C.J. Simeonovic, and H. Warren, *Annu. Rev. Immunol.* **1**, 143 (1983).
125. J. Mansfels and H. Dautzenberg, in G.F. Bickerstaff ed., *Methods in Biotechnology*, vol. 1, Humana Press, Totowa, N.J., 1997, pp. 309–317.
126. I.T. Tai and A.M. Sun, *FASEB J.* **7**, 1061–1069 (1993).
127. E.I. Rainina, M.A. Pusheva, A.M. Ryabokon, N.P. Bolotina, V.I. Lozinsky, and S.D. Varfolomeyev, *Biotechnol. Appl. Biochem.* **32**, 321–329 (1994).
128. U.S. Patent (March 29, 1988), T. Tanabe and M. Nambu (to Nippon Oil Company, Ltd.).
129. N.A. Pepas and S.R. Stauffer, *J. Controlled Release* **16**, 302–310 (1991).
130. V.I. Lozinsky and A.L. Zubov, in D. Poncelet and V. Santos eds., *Bioencapsulation Research Group, International Workshop, Bioencapsulation VI: From Fundamentals to Industrial Applications*, Barcelona, Spain, August 30–September 1, 1997.
131. P.K. Watler, C.H. Cholakis, and M.V. Sefton, *Biomaterials* **9**, 150–154 (1988).
132. N.A. Peppas, in N.A. Peppas ed., *Hydrogels in Medicine and Pharmacy*, CRC Press, Boca Raton, Fla., 1987.
133. K. Furukawa, *J. Ferment. Bioeng.* **76**(6), 515–520 (1993).
134. M. Jekel, A. Buhr, T. Wilke, and K.-D. Vorlop, in D. Poncelet and V. Santos eds., *Bioencapsulation Research Group, International Workshop, Bioencapsulation VI: From Fundamentals to Industrial Applications*, Barcelona, Spain, August 30–September 1, 1997.
135. P.K. Robinson, A.L. Mak, and M.D. Trevan, *Bioprocess Chem.* 122–127 (August 1986).
136. S.F. Nishino and J.C. Spain, *Environ. Sci. Technol.* **27**, 489–494 (1993).
137. E.I. Rainina, M.A. Pusheva, A.M. Ryabokon, N.P. Bolotina, V.I. Lozinsky, and S.D. Varfolomeyev, *Proceedings of the 1993 ERDEC Scientific Conference on Chemical Defense Research*, 895–900, (1994).
138. C.A. Crooks, J.A. Douglas, R.L. Broughton, and M.V. Sefton, *J. Biomed. Mater. Res.* **24**, 1241–1262 (1990).
139. G.D.M. Wells, M.M. Fisher, and M.V. Sefton, *Biomaterials* **14**(8), 615–620 (1993).
140. H. Uludag, V. Horvath, J.P. Black, and M.V. Sefton, *Biotechnol. Bioeng.* **44**, 1199–1204 (1994).

See also FREEZE DRYING, PHARMACEUTICALS.

MIXED CULTURE

SUTEAKI SHIOYA
Osaka University
Osaka, Japan

KEY WORDS

Activated sludge process
Bacteriocin
Bulking
Competition
Cooperation
Lactobacilli
Mixed culture
Mutualism
Nisin
Yeast

OUTLINE

INTRODUCTION

Most fermented foods are produced by mixed cultures acting on various substrates. Cheese, yogurt, pickles, whiskey, and sake (Japanese rice wine) are some examples of fermented foods (1). Numerous interactions occur between microbial communities, plants, and animals. The field of mixed culture is too broad to be reviewed completely in a short article. Of course, it is our belief that models, simulations, and stability analysis used for mixed culture systems are important aids in planning various experiments and interpreting the results (2). Pioneering works in this direction had been done by Henry Tsuchiya and Arnold Fredrickson. For example, we can see their careful results by referring to their review articles in this field (3,4). However, in spite of their extensive efforts, when using these aids general close agreement with data on the dynamic behavior of a complicated mixed culture has yet to be obtained. Thus we stress here applications of mixed culture and discuss case studies that demonstrate how these techniques can contribute to the design, operation, and control of mixed culture systems.

INTERACTION BETWEEN TWO SPECIES

Interactions between two species can be classified into three types: neutral (0), positive, and negative (5). Moreover, positive interactions are characterized into two interactions in detail: one is necessary for the survival of microorganisms (+*) whereas the other has an enhancing effect but is not obligatory for the life of these microorganisms (+). Negative interactions are also divided into two interactions: one directly affects other microorganisms (−*) and the other indirectly affects them (−).

Combination of these interaction categories results in more than nine groups classified under the following headings: neutralism (0, 0), mutual inhibition competition (−*, −*), competition—resource use type (−, −), amensalism (−, 0), parasitism (+, −), predation (+, −), commensalism (+, 0), proto-cooperation (+, +), mutualism (+*, +*). The first symbol in the parentheses shows the relationship of species 1 to species 2 and the second symbol shows the reverse relationship. The difference between parasitism and predation is only with respect to the size of species 1; that is, in parasitism, the size of species 1 is smaller than that of species 2 whereas in predation species 1 is larger than species 2.

Sometimes, of course, the actual relationship between two microorganisms cannot be classified into these nine groups because the relationship cannot be defined clearly. Nevertheless, the nine-group classification is useful for building a mathematical model and understanding the relationship between two species through simulation because a rough overview of the interaction between two microorganisms is possible based on this classification.

There are several examples of interactions belonging to these nine groups, but here only three examples that have been studied by the author are given. One is a mixed culture system of lactobacilli and propionibacteria in a bioreactor, the interaction of which is classified under proto-cooperation (+, +); the second example is a mixed culture system of lactobacilli and yeast classified as mutualism under a limited initial medium condition; and the last example is that of an activated sludge process consisting not of two microorganisms but of two competing groups of microorganism populations because of the convenience of the stability analysis and application of the result to the operation of the system.

One of the aims of this article is to show how models, simulations, and stability analysis are important aids in interpreting results of laboratory experiments or interacting phenomena and for the operation and control of a mixed culture system. Another aim is to demonstrate that the mixed culture system can be utilized for effective bioproduction. Mixed culture systems are encountered frequently in nature, but artificially constructed mixed culture systems can also contribute to the microbial production of various bioproducts such as bacteriocins.

MIXED CULTURE OF LACTOBACILLI AND PROPIONIBACTERIA

Bacteriocins are peptides produced from bacteria such as lactate-producing microorganisms, that is, lactobacilli, and effective antibiotics. Nisin is a typical bacteriocin effective against gram-positive microorganisms (6). Lactobacilli grow at pH higher than 4.0 and their growth is affected by the lactate produced, that is, product inhibition. Thus, the removal of lactate during fermentation is an effective way of optimizing nisin production, but the method of removal of lactate is a difficult problem to solve to assure optimal nisin production. In this study, we try to remove lactate by coculture of lactobacilli with propionibacteria, which can assimilate lactate efficiently.

Growth and Lactate Fermentation Characteristics of *Lactococcus lactis*

Figure 1 shows a typical result of the anaerobic batch culture of *Lactococcus lactis* subsp. *lactis* ATCC11454, a homolactate-producing bacterium. Without pH control, the pH decreased from 7.0 to 5.0 after 2 h of culture and to 4.5 after 3 h. Cell growth as well as nisin production stopped

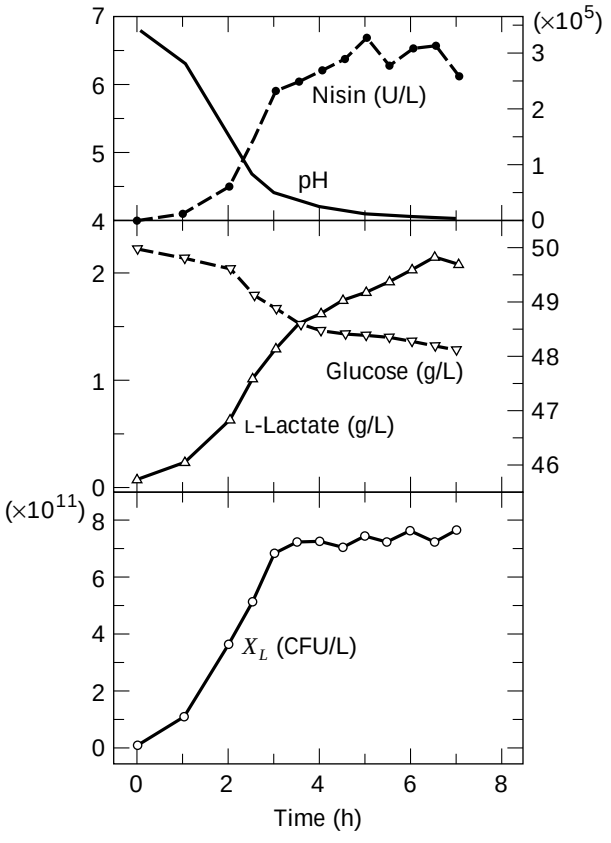


Figure 1. Time course of a batch culture of *Lactococcus lactis* without pH control. *Circles*, cell concentration; *inverted triangles*, glucose concentration; *triangles*, lactate concentration; *solid circles*, nisin concentration.

at 3 h after the start of culture. Lactate was produced at a level of approximately 1.5 g/L and was still produced at a decreased rate until 7 h later. Thus, control of pH to a value greater than 4.5 is essential for achieving a high cell growth rate. Also, nisin production is strictly associated with cell growth. Thus, for effective nisin production, the pH of the medium should be controlled to a value greater than 4.5. Finally, 300 U/mL nisin was produced in this batch culture without pH control.

To improve this low productivity of nisin, pH was controlled and lactate concentration was kept at a fairly low level by diluting the cultivation medium by adding fresh medium. This cultivation method is called repeated dilution culture. Using the method of repeated dilution culture, about 2,000 U/mL of nisin was produced after 10 h of operation. Cell growth, however, stopped at about 15 g/L of lactate. Thus, the control of lactate concentration kept to a level less than 15 g/L is also required for effective continuous production of nisin.

Growth Characteristics of Propionate Bacteria

Propionibacterium freudenreichii IFO12424 was used as a propionate-producing microorganism that can assimilate lactate. Figure 2 shows a typical result of anaerobic batch culture of *P. freudenreichii*. Lactate and glucose were used as carbon sources. From the figure, it can be seen that this

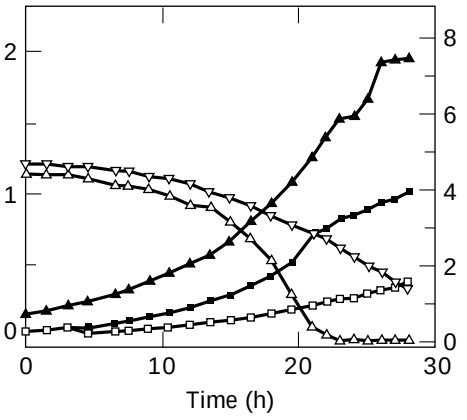


Figure 2. Time course of a batch culture of *Propionibacterium freudenreichii*. *Solid triangles*, cell concentration; *inverted triangles*, glucose concentration; *triangles*, lactate concentration; *open squares*, acetate concentration; *solid squares*, propionate concentration.

microorganism assimilates lactate. However, glucose was also assimilated, which is disadvantageous for the control of lactate concentration because glucose, the main carbon source for *L. lactis*, can be assimilated by *P. freudenreichii*. The glucose is competitively assimilated by both *L. lactis* and *P. freudenreichii*. If the population balance of these microorganisms is destroyed by the deficiency of glucose for *L. lactis* or by accumulation of lactate, stable operation of the mixed culture system will be difficult. Acetate was also produced by *P. freudenreichii*.

Interaction between Two Microorganisms

The interaction between the two microorganisms, *L. lactis* and *P. freudenreichii*, is illustrated in Figure 3. Lactate produced by *L. lactis* can be assimilated by *P. freudenreichii*. Without removing lactate, *L. lactis* cannot grow to a high density because of inhibition by low pH and by lactate itself. In this mixed culture system, the interaction between the two microorganisms is termed proto-cooperation.

Propionate and acetate produced by *P. freudenreichii* contribute somewhat to microbial contaminants in their

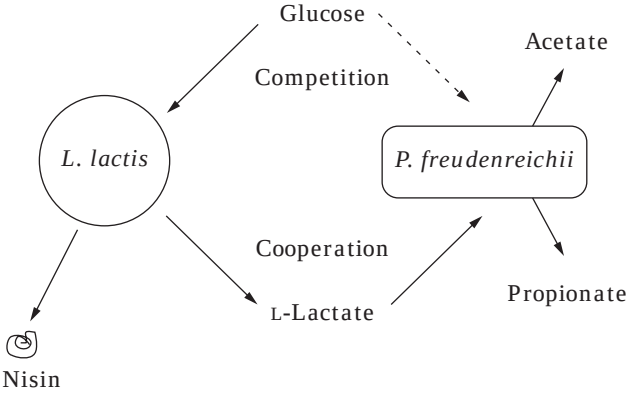


Figure 3. Interaction between *L. lactis* and *P. freudenreichii*.

capacities as antibiotics when they are used for food additives. However, these organic acids also contribute to the decrease in medium pH. Glucose, a carbon source for *L. lactis*, can be assimilated competitively by *P. freudenreichii*. In this sense the interaction between the two microorganisms is classified under competition. Thus, the interaction between the two microorganisms is a mix of cooperation and competition, that is, (±, ±), in glucose-rich medium. However, if we can control the glucose concentration to a level at which it is utilized by *L. lactis* but not by *P. freudenreichii*, the system can operate as a cooperation system. The most important issue is to manipulate or manage the system so as to represent the interaction (+, +), in practice, and to maintain a low supply of glucose, as well as preventing pH decrease of the medium. For this, mathematical modeling and simulation of the system are important tools.

Simulation Model of the Mixed Culture

Based on the kinetic data for cultivation, the following mathematical model was proposed under the assumption that the mixed culture can be described by use of the kinetic data of individual pure cultures.

Dynamics of the mixed culture system can be written by the following equations. To control the growth rate of *L. lactis* at a level comparable to that of *P. freudenreichii* by limiting glucose concentration, a fed-batch culture was employed. The cell growth rate for *L. lactis* in fed-batch culture is written as

$$\frac{dX_L}{dt} = \left(\mu_L - \frac{F}{V}\right)X_L \tag{1}$$

and the rate for *P. freudenreichii* is written as

$$\frac{dX_P}{dt} = \left(\mu_P - \frac{F}{V}\right)X_p \tag{2}$$

where X_i ($i = L,P$) and μ_i are viable cell concentration and specific growth rate, respectively. In these equations, the subscripts L and P represent *L. lactis* and *P. freudenreichii*, respectively, and F and V indicate the flow rate of the feeding medium and the liquid volume in the tank, respectively. Nisin (N, U/mL) production by *L. lactis* is written by

$$\frac{dN}{dt} = \rho_N X_L - \frac{F}{V}N \tag{3}$$

Propionate (P, g/L) production by *P. freudenreichii* is written as

$$\frac{dP}{dt} = \rho_P X_P - \frac{F}{V}P \tag{4}$$

Also, acetate (A, g/L) production by *P. freudenreichii* is written as

$$\frac{dA}{dt} = \rho_A X_P - \frac{F}{V}A \tag{5}$$

where ρ_k is the specific production rate of product k ($k = A,L,N,P$). Lactate production by *L. lactis* and its assimilation by *P. freudenreichii* can be summarized as

$$\frac{dL}{dt} = \rho_L X_L - v_{LP}X_P - \frac{F}{V}L \tag{6}$$

As a result of glucose assimilation by both organisms, the dynamics of glucose concentration (S, g/L) can be written as

$$\frac{dS}{dt} = -v_{SL}X_L - v_{SP}X_P + \frac{F}{V}(S_F - S) \tag{7}$$

where v_{ji} is a specific uptake rate of substrate j by microorganism i ($j = L,S$; $i = L,P$).

The dynamic pH change should be described by a mathematical model in which the following three types of equations are necessary. For example, for lactate concentration, one equation describes the equilibrium between dissociation and association for ionization of lactate such as

$$K_L = \frac{[L^-][H^+]}{[LH]} \tag{8}$$

where K_k is a dissociation constant for product k ($k = A,L,P$).

The second equation describes the total balance between ionized lactate and nonionized lactate $[LH]$ as

$$L_{MT} = [L^-] + [LH] \tag{9}$$

where L_{MT} is the mole-based total lactate concentration. From equation 8, nonionized lactate $[LH]$ is described by other variables, $[H^+]$ and $[L^-]$, and is substituted into equation 9 to give

$$L_{MT} = [L^-] + \frac{[L^-][H^+]}{K_L} = [L^-]\left\{1 + \frac{[H^+]}{K_L}\right\} \tag{10}$$

Then,

$$[L^-] = \frac{L_{MT}}{\left\{1 + \frac{[H^+]}{K_L}\right\}} = \frac{L_{MT}}{(1 + 10^{-pH+pK_L})} \tag{11}$$

where pK_L is defined in the same way as pH and $pK_L = -\log_{10}(K_L)$. In a similar way, ionized propionate $[P^-]$ can be written as

$$[P^-] = \frac{P_{MT}}{(1 + 10^{-kpH+pK_P})} \tag{12}$$

Also, the ionized acetate $[A^-]$ can be written as

$$[A^-] = \frac{A_{MT}}{(1 + 10^{-kpH+pK_A})} \tag{13}$$

These mole-based concentrations can be related to the weight-based concentrations as

$$L_{MT} = \frac{L}{MW_L}, \; P_{MT} = \frac{P}{MW_P} \text{ and } A_{MT} = \frac{A}{MW_A} \quad (14)$$

Here MW_k ($k = A,L,P$) is the molecular weight of chemicals k.

Finally, the following total ion balance should be considered.

$$[H^+] + X_B = [OH^-] + [L^-] + [P^-] + [A^-] + X_A \quad (15)$$

where X_A and X_B are the concentrations of strong acid and base in the medium, respectively. By substituting equations 11, 12, and 13 into equation 15 and using the definition of pH:

$$X_A - X_B = 10^{-pH} - 10^{pH-14} - \frac{L_{MT}}{1 + 10^{-pH+pK_L}} - \frac{P_{MT}}{1 + 10^{-pH+pK_P}} - \frac{A_{MT}}{1 + 10^{-pH+pK_A}} \quad (16)$$

By differentiating this equation with respect to time t:

$$\frac{d(X_A - X_B)}{dt} = \frac{\partial f}{\partial pH}\frac{dpH}{dt} + \frac{\partial f}{\partial L_{MT}}\frac{dL_{MT}}{dt} + \frac{\partial f}{\partial P_{MT}}\frac{dP_{MT}}{dt} + \frac{\partial f}{\partial A_{MT}}\frac{dA_{MT}}{dt} \quad (17)$$

where f represents the right-hand side of equation 16. If we can assume that

$$\frac{d(X_A - X_B)}{dt} = 0 \quad (18)$$

then

$$\frac{dpH}{dt} = \frac{1}{\left(\frac{\partial f}{\partial pH}\right)}\left\{-\left(\frac{\partial f}{\partial L_{MT}}\right)\frac{dL_{MT}}{dt} - \left(\frac{\partial f}{\partial P_{MT}}\right)\frac{dP_{MT}}{dt} - \left(\frac{\partial f}{\partial A_{Mt}}\right)\frac{dA_{MT}}{dt}\right\} \quad (19)$$

Now

$$\frac{\partial f}{\partial L_{MT}} = \frac{-1}{(1 + 10^{-pH+pK_L})}, \frac{\partial f}{\partial P_{MT}} = \frac{-1}{(1 + 10^{-pH+pK_P})}, \frac{\partial f}{\partial A_{MT}} = \frac{-1}{(1 + 10^{-pH+pK_A})} \quad (20)$$

and

$$\frac{\partial f}{\partial pH} = \mathrm{Ln}10\left[-10^{-pH} - 10^{pH-14} - \frac{L_{MT}\cdot 10^{-pH+pK_L}}{(1 + 10^{-pH+pK_L})^2} - \frac{P_{MT}\cdot 10^{-pH+pK_P}}{(1 + 10^{-pH+pK_P})^2} - \frac{A_{MT}\cdot 10^{-pH+pK_A}}{(1 + 10^{-pH+pK_A})^2}\right] \quad (21)$$

can be easily derived. Taking into account that equation 8 is not always valid, a correction term was introduced into the denominator of equation 19 as 1/230, which was determined using experimental data of the mixed culture system and resulted in the following equation (22). Finally, the pH change can be described by

$$\frac{dpH}{dt} = \left\{-\frac{1}{(1 + 10^{-pH+pK'_L})}\frac{dL_{MT}}{dt} - \frac{1}{(1 + 10^{-pH+pK'_P})} \cdot \frac{dP_{MT}}{dt} - \frac{1}{(1 + 10^{-pH+pK_{AL}})}\frac{dA_{MT}}{dt}\right\} \div \left\{\frac{1}{230} + \mathrm{Ln}10\left[10^{-pH} + 10^{pH-14} + \frac{L_{MT}\cdot 10^{-pH+pK_L}}{(1 + 10^{-pH+pK_L})^2} + \frac{P_{MT}\cdot 10^{-pH+pK_P}}{(1 + 10^{-pH+pK_P})^2} + \frac{A_{MT}\cdot 10^{-\&pH+pK_A}}{(1 + 10^{-pH+pKA})^2}\right]\right\} \quad (22)$$

where the dynamic change of L_{MT}, P_{MT}, and A_{MT} can be calculated using equations 4, 5, 6, and 14. Finally, the kinetic parameters such as μ_i, v_{ji}, and ρ_k ($i = L,P$, $j = L,S$, $k = A,L,N,P$) are given as a function of pH as well as substrate concentration, and the dynamic behavior of the fed-batch culture can be simulated using the mathematical model described by equations 1 through 7, 14, and 22.

Control of pH by Adjusting an Appropriate Initial Population

To maintain the culture pH at a value higher than 5.0 for a long period, *P. freudenreichii* should grow at a fast rate so that lactate is assimilated well. By comparing the specific growth rates of both microorganisms, it can be seen that *L. lactis* grows at a much faster rate than *P. freudenreichii*. This result means that lactate produced in correlation with *L. lactis* growth will accumulate in the fermentor if there is no control of the growth of *L. lactis*. For control of the growth rate of *L. lactis*, a fed-batch culture was employed. By manipulating the exponential feeding rate of glucose (in this case, μ_L was set at 0.2 h^{-1}; initial tank volume at 1.5 L, and initial glucose concentration at 2 g/L), the growth of *L. lactis* can be controlled. The mathematical model proposed here can be used for simulation of this exponential fed-batch mixed culture system. The kinetic parameters of *L. lactis* that depend on the pH were determined from the data for the pure batch culture. Other kinetic parameters were also determined from the experimental data.

By changing the initial cell populations of *L. lactis* and *P. freudenreichii*, the dynamic behavior of the mixed culture can also be altered. In this way the most appropriate initial cell populations were selected. After several simulation studies, an appropriate initial ratio of cell concentration of *L. lactis* to that of *P. freudenreichii* was determined as 1:14. The simulation result is shown in Figure 4. Here, to easily compare the simulation and experimental data, the cell concentration of *L. lactis* is represented by cell number whereas that for *P. freudenreichii* is by dry cell weight, because in the experiment the cell concentration of *P. freudenreichii* was calculated by subtraction of the

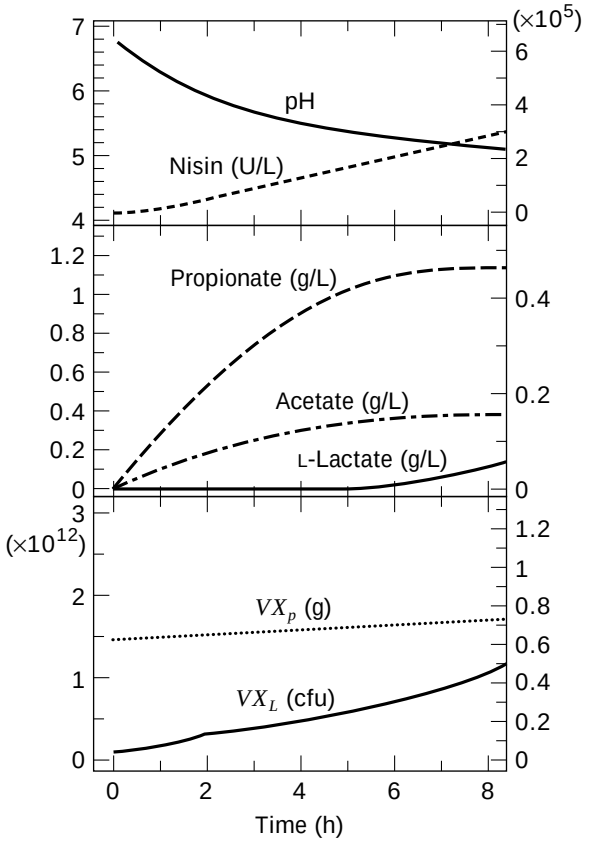


Figure 4. Simulation of a mixed culture for effective nisin production.

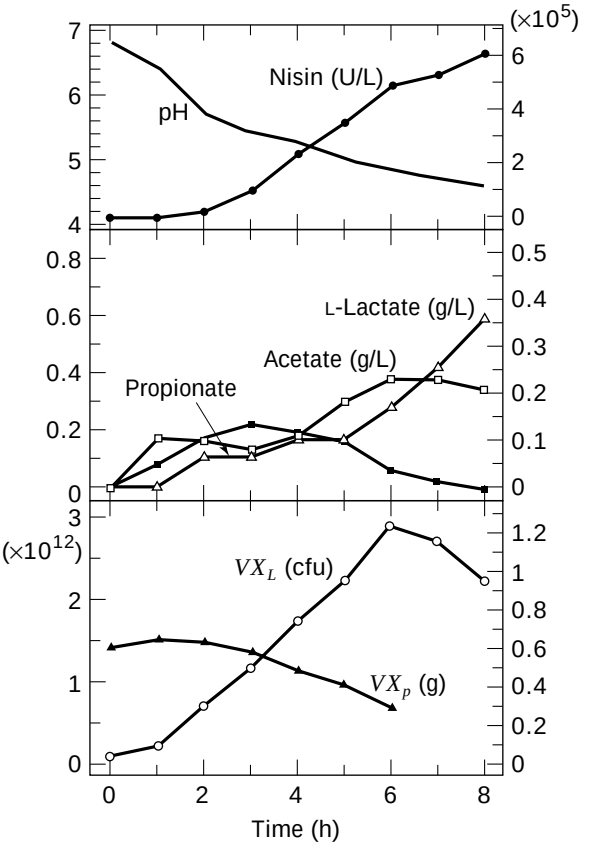


Figure 5. Experimental result of a mixed culture of *L. lactis* and *P. freudenreichii*. *Circles*, cell concentration of *L. lactis*; *solid triangles*, cell concentration of *P. freudenreichii*; *triangles*, lactate concentration; *open squares*, acetate concentration; *solid squares*, propionate concentration; *solid circles*, nisin concentration.

amount of dry cell weight of *L. lactis* converted from the cell number counted by plate colony from the total dry cell weight. It is shown that pH will be maintained at a value higher than 5 for a period of 5 h and that nisin will be produced much more effectively than for a given specific growth rate of *L. lactis*.

Experimental Result for the Mixed Culture

Figure 5 shows an experimental result of the mixed fed-batch culture with the same initial cell concentrations of both microorganisms, *L. lactis* and *P. freudenreichii*, as those in Figure 4. The pH could be controlled until 5 h from the start of cultivation. However, under these conditions *P. freudenreichii* could not grow well, whereas *L. lactis* grew faster than the estimated rate shown in Figure 4. The reason why there is a large difference between dynamic simulation using the mathematical model and experimental data is not clear at this point. However, the simulation model at least can be used for forecasting the dynamics of the mixed culture and for adjusting the initial cell concentrations of both species so as to achieve effective production of nisin by this mixed culture. Finally, the experimental data for the mixed culture show that the final nisin concentration was improved twofold.

In the near future, the process will be improved further after determination of the reason for the difference between the experimental data and the simulation data.

MIXED CULTURE OF LACTOBACILLI AND YEAST

To enhance bacteriocin productivity, a mixed culture system was employed as outlined in the previous section. However, the propionibacteria can assimilate glucose, which in the system described is the main carbon source of lactobacilli. Thus the glucose was competitively assimilated by both microorganisms, which may result in the need to use a fed-batch operation. One improvement that can be made to the system is use a main carbon source of lactobacilli which cannot be assimilated by another microorganism. Only the lactate produced by the lactobacilli should be a carbon source of the second microorganism. From this aspect, maltose was chosen as a carbon source of lactobacilli, and a microorganism that can assimilate lactate but not maltose was selected. From kefir, a type of yogurt, such a microorganism, *Kluyveromyces marxianus*, was isolated. The following mixed culture system was then constructed.

Interaction between *Lactococcus lactis* and *Kluyveromyces marxianus*

Kluyveromyces marxianus MS1, isolated from a type of yogurt (kefir), can assimilate lactate as a carbon source. *K.*

marxianus cannot assimilate starch and maltose directly whereas *L. lactis* can grow on maltose. Whether *L. lactis* could grow and produce nisin under aerobic conditions had to be determined because *K. marxianus* grows under aerobic conditions. The results indicated that there is no significant difference in the productivity of nisin and growth rate of *L. lactis* between aerobic and anaerobic conditions. Thus, the mixed culture system can be operated under aerobic conditions. The interaction of the two microorganisms is illustrated in Figure 6. *L. lactis* can assimilate the maltose and produce lactate, which can be assimilated by the yeast, *K. marxianus*. This cooperative interaction results in a very low or undetectable lactate concentration, which may enhance the growth of *L. lactis* as well as nisin production. Thus, the relationship between these microorganisms can be termed proto-cooperation or mutualism because their relation can be summarized by (+, +*).

Mixed Culture of *L. lactis* and *K. marxianus*

From the characteristics of the interaction, it is expected that a high growth rate of *L. lactis* as well as high productivity of nisin will be achieved. In Figure 7, the experimental data indicated that the system worked as desired because the lactate concentration can be kept at almost zero or at least at an undetectable level and *L. lactis* grew significantly. The monoculture of *L. lactis* using maltose as a carbon source stopped at 7 h because of low pH (Fig. 7a). Under mixed culture, pH could be maintained well around 6.0 (Fig. 7b) even by a manual dissolved oxygen (DO) concentration control. The role of DO control is explained later. Also, *L. lactis* as well as *K. marxianus* could grow at a level fourfold higher than the monoculture of *L. lactis*. The yeast also grew well and the nisin yield reached 1,400 U/mL at the end of fermentation after only 8 h operation (Fig. 7c), similar to the amount obtained in the pH-controlled repeated batch culture, which was fourfold higher than the level in the monoculture of *L. lactis*.

The pH level of the culture medium can be controlled by the lactate uptake rate linked with the growth of *K. marxianus*. On the other hand, the growth of *K. marxianus* can be controlled by the DO level. Then, the DO level can control the pH level in this mixed culture system. Finally, it is clear that a setpoint for the DO control by changing the agitation speed can control pH level. This is a so-called cascade controller. Based on this cascade controller concept, the pH level could be maintained almost completely automatically at a constant level by manipulating the agitation speed of the fermentor. Finally, the result of this mixed culture experiment with cascade control of pH is shown in Figure 8; the product nisin concentration was enhanced more and reached about 4,000 U/mL after 11 h of operation, which was almost 10 fold that by anaerobic monoculture of *L. lactis* without pH control and about twice as much as by anaerobic monoculture of *L. lactis* with pH control by NaOH addition.

The system developed here functions successfully. However, it should be stressed that a simple and reliable method for the measurement of individual cell concentrations should be developed. One way is to use a flow cytometer that can count individual cells if the cell size of each population is significantly different. One example of flow cytometry measurement of the concentration of lactobacilli and yeast in the mixed culture is shown in Figure 9. The abscissa of the figure represents the cell size, and one dot indicates one cell. Thus, densely clustered points means there are many cells. In this case, the lactobacilli cells are small compared with yeast cells. Using the scale of the abscissa, it may be possible to count each population. However, to identify the two populations of microorganisms more effectively, staining of their DNA using a dye such as propidium iodine (PI) also aids in discriminating the yeast and lactobacillus by cell size as well as by DNA content under two-dimensional plotting as shown in the figure. The ordinate indicates the DNA contents represented by strength of PI staining. In Figure 9, the clustering between lactobacilli and yeast cells becomes evident because the two populations separate like two islands.

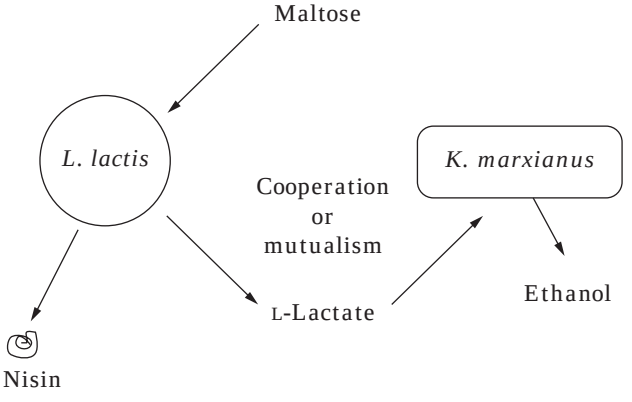


Figure 6. Interaction between *L. lactis* and *Kluyveromyces marxianus*.

DYNAMICS OF AN ACTIVATED SLUDGE PROCESS AS A MIXED CULTURE SYSTEM

An activated sludge process is a typical mixed culture system. The formation of activated sludge flocs as a result of the overgrowth of filamentous microorganisms results in a phenomenon known as bulking, which is a major problem in the operation of the activated sludge process. Wastewater composition, sludge loading, DO concentration, the feeding pattern of the aerator, and pH are some of the factors involved in the bulking phenomenon. The objective of this work is to develop a simple mathematical model describing the bulking phenomenon caused by high sludge loading, including the growth kinetics of microorganisms in an aeration tank in terms of a mixed culture system and the ability of the sludge to settle, and to obtain some means of preventing and controlling bulking. In our experiment, only the sludge loading rate was changed; the other operating conditions such as pH and temperature were kept constant. The DO concentration in the aeration tank was maintained at a high level so as not to induce the bulking phenomenon. The model system selected for examination is a mixed culture of two groups of microorganisms, one having the ability to induce the formation of flocs and the

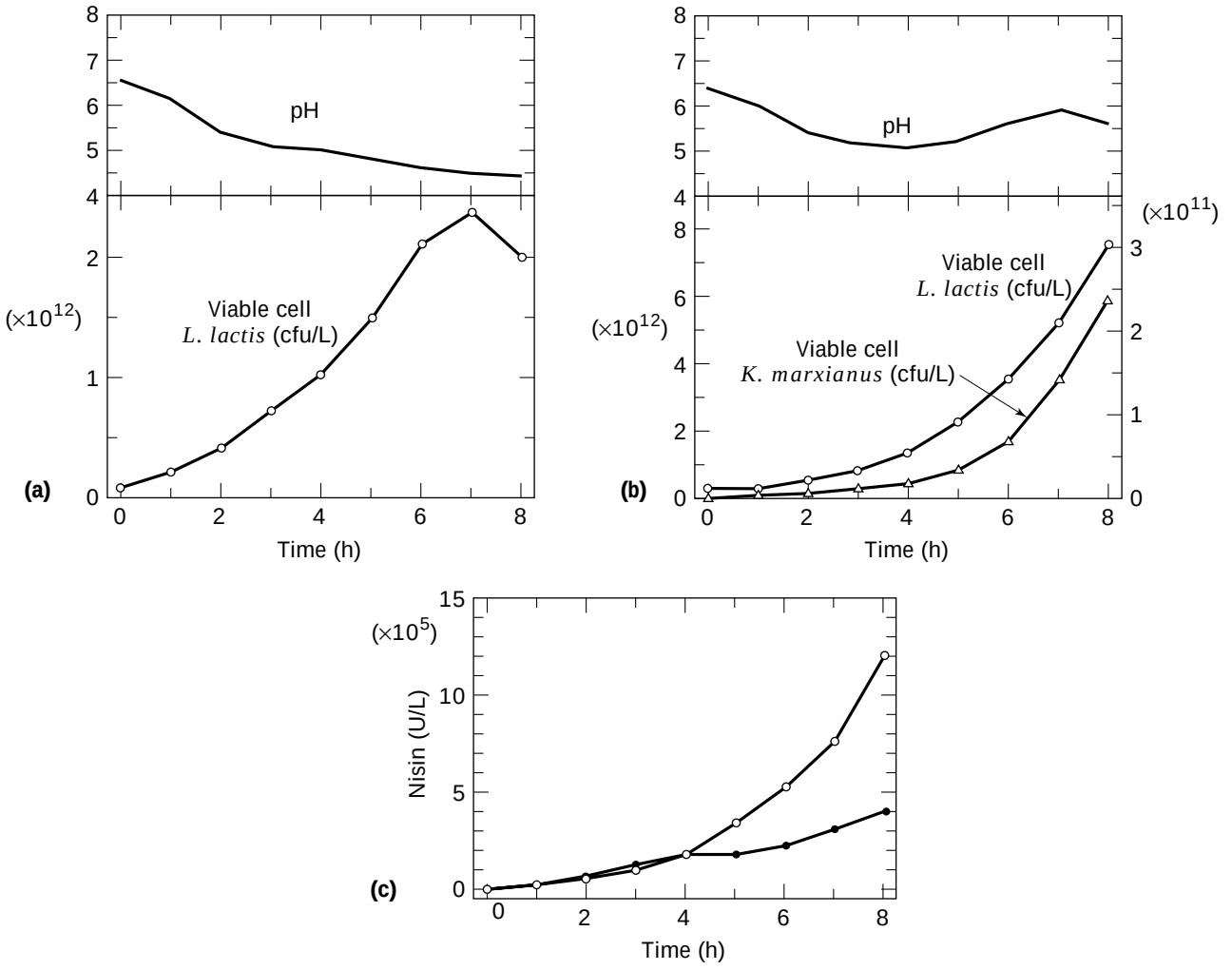


Figure 7. (**a**) Monoculture of *L. lactis* without pH control or pH changes. *Circles*, cell concentration of *L. lactis*. (**b**) Mixed culture of *L. lactis* and *K. marxianus* by manual dissolved oxygen (DO) control. *Circles*, cell concentration of *L. lactis*; *triangles*, cell concentration of *K. marxianus*. (**c**) Production of nisin under a mono- and mixed cultures by manual DO control. *Solid circles*, nisin concentration by anaerobic monoculture without pH control; *open circles*, nisin concentration by (aerobic) mixed culture.

other inducing bulking. The interaction between the two groups is classified under competition with respect to the utilization of the same substrate (7).

A continuous flow biochemical reactor in which a mixed culture of microorganisms is grown very often exhibits multiple steady states. The analysis of the stability of these steady states has been the subject of research. The techniques most frequently employed in determining the stability characteristics of a steady state are based on Liapunov's first method, which consists mainly of a linearization of the differential equations around the steady state and an examination of the eigenvalue of the Jacobian matrix. Here, the dynamic behavior of the activated sludge process, in which competition occurs with respect to assimilation of the substrate, is analyzed based on a mathematical model utilizing phase plane analysis and computer simulation. The results including those pertaining to the conditions that result in the system reaching the bulking steady state are very informative and useful for the practical operation of activated sludge processes.

Kinetics of Substrate Uptake and Cell Growth

The kinetics in two synthetic media, one in which the carbon source is glucose and polypeptone (medium A) and the other in which the carbon source is glucose (medium B), were studied experimentally using batch culture. In this system, the chemical oxygen demand (COD) concentration, S, is considered to be a limiting substrate for describing the kinetics of the process. Both groups of microorganism populations, that is, the floc-forming and the bulking-inducing groups, were cultured by feeding different types of synthetic medium into a batch reactor on a fill-and-draw basis. Two synthetic media were used to purify and main-

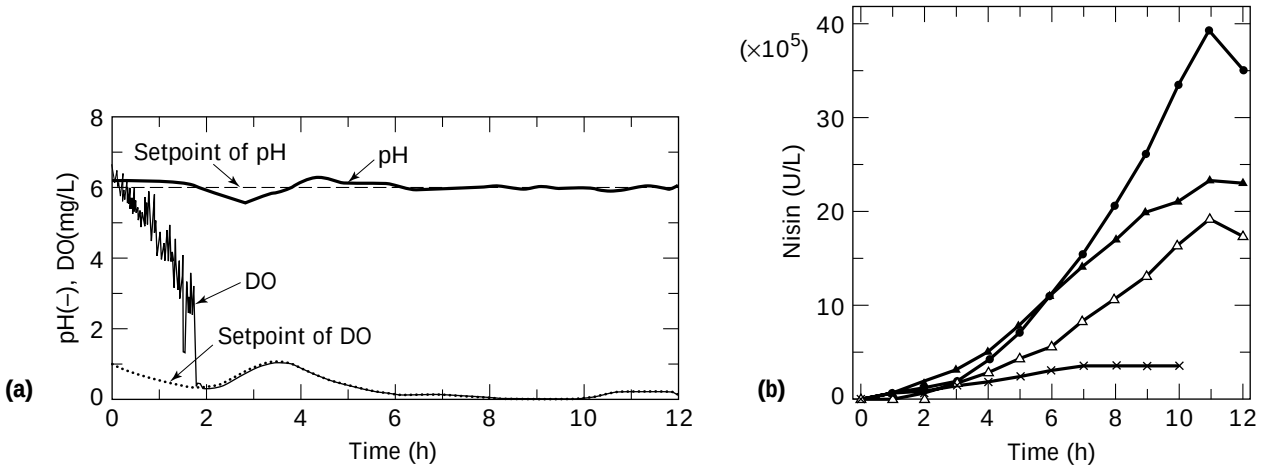


Figure 8. (**a**) pH control by an automatic cascade controller coupled with DO control in the mixed culture system. (**b**) Production of nisin by cascade control of pH compared with other pH control strategies. ×, anaerobic monoculture without pH control; *solid triangles*, anaerobic monoculture with pH control by NaOH addition; *open triangles*, aerobic monoculture with pH control by NaOH addition; *solid circles*, mixed culture with cascade control of pH.

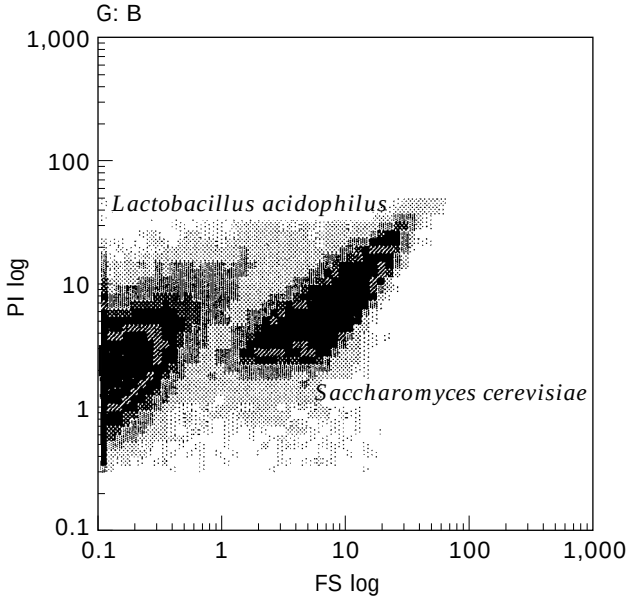


Figure 9. Two-dimensional plot of flow cytometry for a mixed culture system of *Lactobacillus acidophilus* and *Saccharomyces cerevisiae*.

tain each population of microorganisms for more than 2 months. Using medium A in which the initial COD concentration was about 110 mg/L, the sludge settled well and was called floc-forming sludge. On the other hand, using medium B in which the initial COD concentration was about 450 mg/L, the sludge did not settle well, that is, it had a high sludge volume index. The sludge could be kept stable for more than 1 month by the feeding of medium B and was called bulking sludge. From the results of microscopic observation, it was determined that the bulking sludge was composed of filamentous microorganisms such as *Sphaphaeroutilus* sp., while the floc-forming sludge did not contain such microorganisms.

A 1-L jar fermentor equipped with automatic temperature control, an air flow rate indicator, DO sensors, and a magnetic agitator was used. The kinetics of floc-forming and bulking populations were analyzed from the batch experiments using medium A or B with various initial COD concentrations. A typical result (Fig. 10) shows the time courses of floc-forming sludge concentration X_F and the COD concentration S with medium A. From this figure, it was shown that the specific growth rate μ_F and the substrate uptake rate v_F could be represented by a linear relationship with respect to substrate (COD) concentration S and that the yield coefficient of the sludge from COD was constant. Similar results for growth rates and uptake rates as described here were obtained for both types of sludge and for both types of media. The mathematical forms of growth rates and uptake rates thus are as follows:

$$\mu_j = K_j(S - S_j) \tag{23}$$

and

$$v_j = \frac{-1}{Y_j}\mu_j \quad j = F, B \tag{24}$$

where F and B mean floc-forming and bulking sludge, respectively. The solid lines in Figure 10 show the calculated values using the estimated parameters. The kinetic model described by equations 23 and 24 is generally acceptable in the low COD concentration range.

The ability of the activated sludge to settle in the sedimentation vessel was also studied experimentally. An empirical description of the concentration of the sedimenting sludge related to the ratio of X_F/X_B was obtained. The ratio of the concentration of the settled sludge, X_g, to the con-

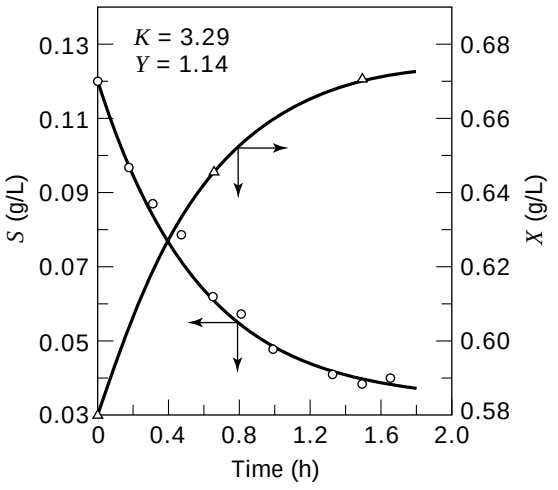


Figure 10. Batch culture of floc-forming sludge using medium A. *Circles*, chemical oxygen demand (COD) concentration; *triangles*, cell concentration.

centration of the sludge to be settled, X, which is denoted by χ, can be represented by

$$\chi = \frac{X_R}{X} = \alpha\left(\frac{X_B}{X_F} + \beta\right)^{-\gamma} + 1 \qquad (25)$$

Interaction between Bulking and Floc-Forming Sludges

It is assumed that there is competitive use of substrate between the bulking and floc-forming sludge. This assumption could be partially verified by the following experiment and computer simulation. When a fill-and-draw-type batch experiment is conducted from the point at which the ratio of the bulking to the floc-forming sludge is very small, the ratio of the bulking sludge to the total sludge increases gradually using medium B, and its ability to settle becomes poorer after more than 4 days from the start of the experiment. This experimental result can be simulated as shown in Figure 11 by the mathematical model given in equations 23 and 24. After 4 or 5 days from the start of the experiment, χ becomes almost 1.0; that is, the state of the sludge approaches bulking, which is reflected by the experimental data.

Mathematical Model of the Activated Sludge Process

For an activated sludge process in a completely mixed aeration tank and sedimentation vessel, a mathematical model of the system is derived. The mathematical model of the activated sludge process considered here is as follows:

$$\frac{dX_j}{dt} = \{(\chi - 1)q_r - q_i + \mu_j\}X_j, \quad j = F, B \qquad (26)$$

$$\frac{dS}{dt} = q_i(S_i - S) - \frac{\mu_F}{Y_F}X_F - \frac{\mu_B}{Y_B}X_B \qquad (27)$$

$$\mu_j = \begin{cases} K_j(S - S_j) & \text{for } S \geq S_j \\ 0 & \text{for } S < S_j, \end{cases} \quad j = F, B \qquad (28)$$

and

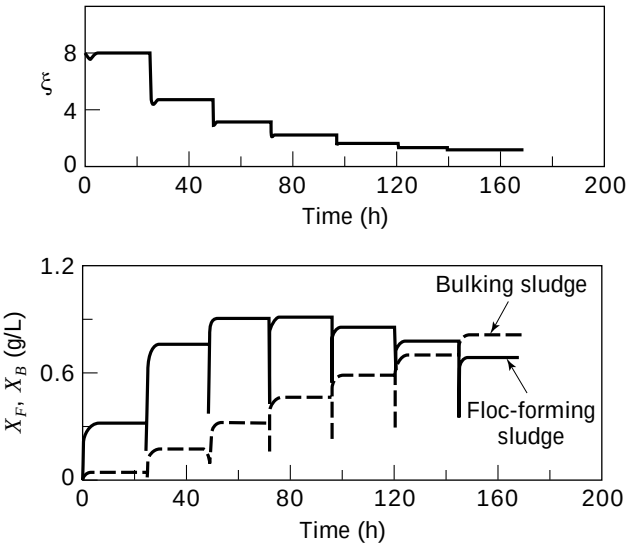


Figure 11. Simulation of fill-and-draw-type batch culture using medium B. After more than 4 days, the concentration ratio χ approached 1.0. After that, the bulking phenomenon was observed.

$$\chi = \alpha\left(\frac{X_B}{X_F} + \beta\right)^{-\gamma} + 1 \qquad (29)$$

where X_F, X_B, and S are the outlet concentration of the floc-forming sludge and the bulking sludge and the COD in the aeration tank, respectively, and S_i is the inlet COD concentration in the system; q_i and q_r are the inlet flow rate and recycled flow rate divided by the tank volume, respectively. Equations 26 through 29 are analyzed to obtain information about the bulking phenomenon caused by high organic loading. One of the important characteristics of process equations is that the growth rates of the bulking and floc-forming sludges intersect (Fig. 12). This result means that the interaction between the two sludges can be classified as competition for assimilation of the same substrate (COD). Using medium B, the relationship between the specific growth rates of both sludges becomes similar to those using medium A.

The intersection in Figure 12 defines the value of the COD concentration S_c at which both types of sludge, bulking and floc-forming, can coexist. As shown in Figure 12, in the region where COD is smaller than S_c, the specific growth rate of the floc-forming sludge is higher than that of the bulking sludge; that is, the floc-forming sludge may become dominant. On the other hand, in the region where COD is larger than S_c, the bulking sludge is dominant.

Another distinguishable feature of the mixed culture system employed here is the existence of the cell recycle stream. Reflecting on this feature, the coexistence of both sludges should be taken into account at a certain dilution rate, although the coexistence state does not occur without the cell recycle stream.

Dynamics of an Activated Sludge Process

Equilibrium Points. We next analyzed the dynamic behavior of equations 26 through 29 when the organic load-

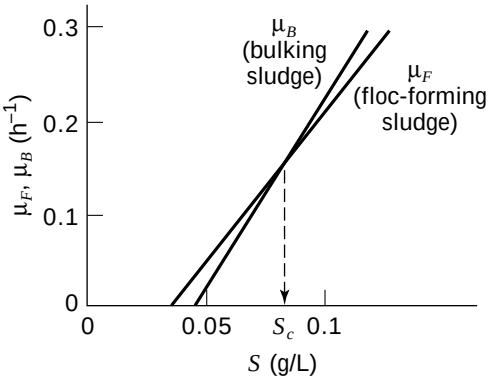


Figure 12. The specific growth rates of floc-forming sludge and bulking sludge using medium A.

ing rates S_i and q_i are changed. As the characteristics of a mixed culture system using medium B are similar to those using medium A, analysis is based on the parameters obtained using medium A. First, the possible steady states should be clarified.

It can be shown that there are, in general, four equilibrium points: normal state E-1 ($X_F \neq 0$, $X_B = 0$); bulking state E-2 ($X_F = 0$, $X_B \neq 0$); coexistence state E-3 ($X_F \neq 0$, $X_B \neq 0$); and washout state E-4 ($X_F = 0$, $X_B = 0$). The normal state, E-1, is the only desirable state for stable operation, and the bulking state, E-2, should be avoided. Figure 13 shows the relationship between these four equilibrium points in the state space. These points are given as intersections of three planes, called "isoclines," and the plane $X_B = 0$ or $X_F = 0$. An isocline is defined here as the curve or the plane along which the first derivative of

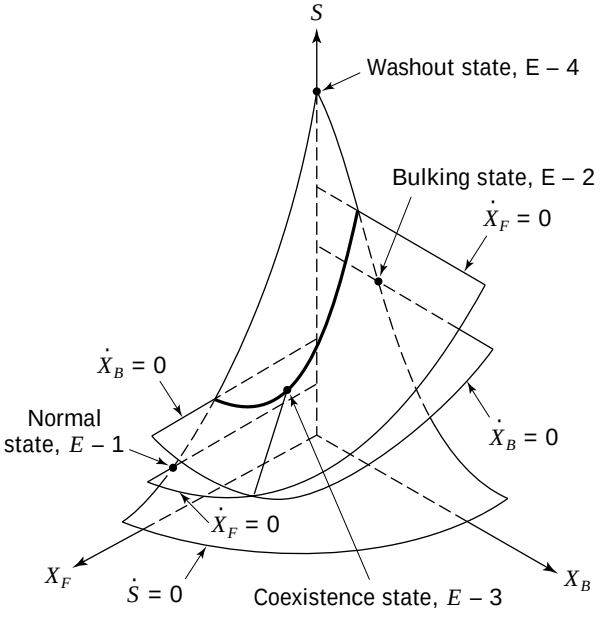


Figure 13. The possible independent equilibrium points, which are shown as intersection of isoclines and $X_F = 0$ or $X_B = 0$. Four possible points are shown as E-1 to E-4.

the state variables with respect to time t is zero excluding $X_B = 0$ and $X_F = 0$. Then, the isoclines are three planes that satisfy $dX_F/dt = 0$, $dX_B/dt = 0$, and $dS/dt = 0$ in equations 26 through 29. From the definition, the intersections of these planes are equilibrium points. For example, point E-1 is given as the intersection of both isoclines of $dX_F/dt = 0$ and $dS/dt = 0$, and the plane $X_B = 0$. The number of possible equilibrium points under a given operating condition is counted using the shapes and interconnections of these isoclines.

The trajectory of the state from a given initial condition to a final state is roughly estimated using these three isoclines because an isocline divides the gradient of a state variable into two regions having negative and positive values. From the gradient of the three state variables, the gradient of the trajectory in the state space is estimated. Finally, the maximum number of possible equilibrium points is four, as shown in Figure 13, and normal state E-1 is the only desirable state for the stable operation of the activated sludge process. The other states should be avoided. To make it possible for normal state E-1 to exist, parameters such as input flow rate q_i or recycled flow rate q_r and input COD concentration S_i must satisfy the following equations:

$$S_F \leq S_e \leq S_i \tag{30}$$

where

$$S_e = \frac{\{q_i - (\chi^* - 1)q_r + K_F S_F\}}{K_F} \tag{31}$$

Here χ^* is the value of χ at $X_B/X_F = 0$. This constraint was deduced from the assumption that normal state E-1 really exists. The same type of analysis can be performed with respect to the other equilibrium points.

Locally Stable Conditions of the Activated Sludge Process. The activated sludge process should be operated under conditions guaranteeing that the equilibrium point E-1 exists and is locally stable. The term locally stable means that the equilibrium point is asymptotically stable in the linearized system. The occurrence of the locally stable condition can be determined by checking the sign of the real part of the eigenvalues of the characteristic equation in the linearized system (8). Stability is guaranteed only in the vicinity of the equilibrium point. Of course, in practical operation, whether the system reaches the state E-1 from a given initial state depends on global stability, based not only on local stability but also on the global dynamic behavior. However, locally stable conditions provide useful information with respect to the stable operation of the activated sludge process. Using the locally stable condition, the domain of stable operation in the operating space consisting of S_i, q_i, and q_r can be obtained.

The locally stable conditions of four equilibrium points were deduced analytically as functions of the operating conditions S_i, q_i, and q_r. We showed that the coexistence state E-3 is unstable under all operating conditions. The locally stable conditions and the existence conditions of the four equilibrium points divide the domain of the operating

conditions into many regions. Such an analysis can be used to identify the existence and stability of an equilibrium point.

Figure 14 shows the region described by q_i, q_r, and S_i in which E-1 exists and is locally stable. Under the locally stable condition at E-1, S_e defined in equation 31 is as follows:

$$S_F \leq S_e \leq S_c \tag{32}$$

where S_c is the intersection of μ_F and μ_B as shown in Figure 12; that is, $\mu_F \geqq \mu_B$ in this region, and the floc-forming sludge becomes dominant. The well-known fact that high organic loading will cause the bulking phenomenon is verified using the model employed here (Fig. 14). This figure can also be utilized to determine the strategy required to obtain locally stable conditions. When the state of both sludges and COD concentrations does not change substantially from E-1, the system will be operated in a stable way as long as the operating condition is maintained in the region indicated in Figure 14. If the flow rate of wastewater q_i increases, the recycled flow rate q_r should be increased, as in Figure 14, to maintain the conditions in this locally stable region.

Figure 15 shows the region in which the bulking equilibrium E-2 is unstable or does not exist. In the domain of the operating conditions indicated in Figure 15, only E-1 is locally stable. Even though the bulking sludge increases slightly, based on the local stability it can be said that the desired operating state will be restored as long as the operating condition is maintained in this region. These conditions coincide with the well-known rule that the sludge loading rate should be kept low. The kind of deviation in the operating and initial conditions that renders the normal state E-2 unstable was clarified.

The locally stable conditions are not sufficient to maintain a stable operation that prevents the occurrence of the bulking phenomenon. This is because in practice locally stable conditions are valid only in the vicinity of equilibrium points. Therefore, the manipulating variables should be changed dynamically to maintain the system in normal state E-1 when disturbances have been introduced. The control problem has been solved (7). Finally, the interaction between the two sludges has been classified as competitive, and a mathematical model has been proposed and simulations as well as a stability analysis have been performed to determine the manner in which to operate the activated sludge process in a stable way.

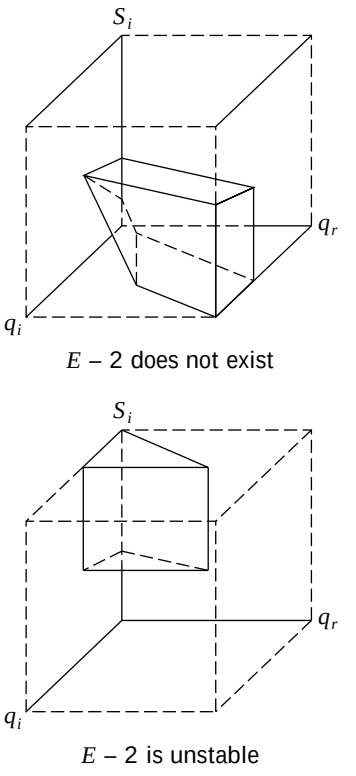


Figure 15. The desirable operating domain so that E-1 is stable and the other equilibrium points do not exist or are unstable.

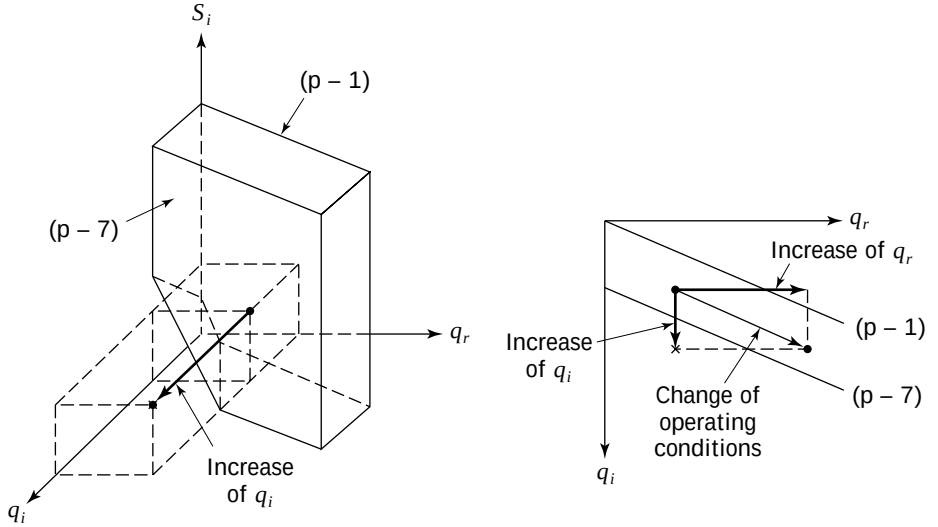


Figure 14. The domain of operating conditions in which E-1 exists and is stable.

MUTAGENESIS

RICHARD H. BALTZ
CognoGen Enterprises
Indianapolis, Indiana

KEY WORDS

Antibiotic biosynthesis
Gene duplication
Mutagenesis
Secondary metabolite
Streptomyces
Transposon

OUTLINE

INTRODUCTION

Mutations are heritable changes in the genome that arise "spontaneously" or are induced by chemical or physical means. Mutational changes include deletions, duplications, translocations (transpositions), and base-pair substitutions. Base-pair substitutions are subdivided into transitions (a purine substituted for a purine, or a pyrimidine substituted for a pyrimidine) and transversions (a purine substituted for a pyrimidine and vise versa). There are two types of transitions (GC to AT and AT to GC), and four types of transversions (GC to CG, GC to TA, AT to CG, and AT to TA). Mutagenesis is an important aspect of fermentation process development for the production of primary metabolites (e.g., amino acids and vitamins) and secondary metabolites (e.g., antibiotics, antitumor agents, antihelmintic agents, and immunomodulators). Mutagenesis can be used to increase the levels of important precursors, cofactors, and enzymes involved in the biosynthesis of the desired product. It can also be used to eliminate competing nonessential biosynthetic pathways or to modify other parameters important in the fermentation process (e.g., culture viscosity, oxygen transfer, culture stability). There are two general approaches to mutagenesis, both of which are important at different stages of process development. The first approach is random chemical or physical mutagenesis. This is an indispensible approach to improve the production of complex secondary metabolites, because a priori it is essentially impossible to deduce what factors will be most important in the fermentation to improve the yield. The key to success with random mutagenesis is to identify procedures that give the optimum yields of mutants exhibiting the desired features. The second type of approach is more directed and can be focused specifically at rate-limiting steps in biosynthesis or at key regulatory elements. This generally requires more sophisticated molecular genetic tools and more knowledge about the genetics and biosynthesis of the product of interest. I discuss some key elements of the random and directed mutagenesis approaches, using secondary metabolite production in actinomycetes as models. Many of the methods and approaches can be adapted for other prokaryotes and for lower eukaryotes that have a haploid sporulation stage.

APPROACHES TO RANDOM BASE-PAIR SUBSTITUTION MUTAGENESIS

Comparative Mutagenesis

A number of chemical mutagens and UV light have been compared for their efficiencies of induction of mutations to spectinomycin resistance in *Streptomyces fradiae* (1,2). Of the mutagens tested, *N*-methyl-*N'*-nitro-*N*-nitrosoguanide (MNNG) was the most efficient, and UV light (UV) was the least efficient. The mutagens ranked in efficiency of mutation induction as follows: MNNG > 4-nitroquinoline-1-oxide (NQO) > methyl methanesulfonate (MMS) > ethyl methanesulfonate (EMS) > hydroxylamine (HA) > UV (Fig. 1). The efficiency of MNNG mutagenesis was enhanced by carrying out mutagenesis in the presence of chloramphenicol (4).

Optimizing Mutagenesis

The objective in optimizing mutagenesis conditions is to obtain the highest frequency of survivors containing single mutations altering secondary metabolite production. Higher levels of mutagenesis are not desirable because most mutations that alter product yields have negative effects (2). Therefore higher levels of mutagenesis will result in the coupling of gain mutations with loss-of-yield mutations. The optimum level of mutagenesis can be identified by statistical analysis of the progeny of mutagenesis. Because mutations, under ideal conditions, will be distributed among individual cells according to the Poisson distribution (2), the optimum condition giving a multiplicity of one mutation affecting secondary metabolite production per cell is obtained when 37% of the survivors of mutagenesis produce control yields (i.e., the null fraction of the Poisson distribution, $P_0 = e^{-m} = 0.37$, where the multiplicity (m) = 1). At this multiplicity, the fraction of cells containing only one mutation is also 37% (i.e., $P_1 = 1 \times e^{-1}/1!$). In *S. fradiae* the only mutagen capable of achieving

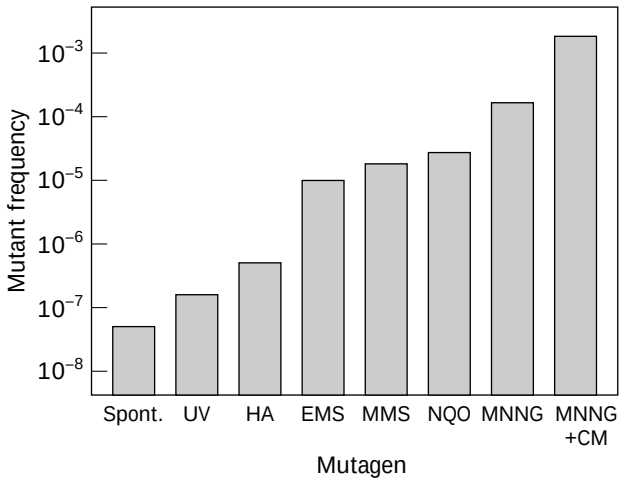


Figure 1. Frequency of spectinomycin-resistant mutants induced by mutagenic agents in *Streptomyces fradiae*. Spont., spontaneous; UV, ultraviolet light; HA, hydroxylamine; EMS, ethyl methanesulfonate; MMS, methyl methanesulfonate; NQO; 4-nitroquinoline-1 oxide; MNNG, *N*-methyl-*N′*-nitro-*N*-nitrosoguanidine; CM, chloramphenicol. *Source*: Data from Refs. 2 and 3.

a mutational multiplicity of as high as 1 is MNNG. However, MNNG plus CM (4) gives much too high levels of mutagenesis to be a useful procedure for random mutagenesis.

Mutagen Specificity and Limitations of Current Protocols

The specificity of base-pair substitutions induced by mutagenic agents has been studied extensively in *Escherichia coli* (5–7). MNNG induces mutations by the transition pathway 99% of the time, and 95% of these are GC to AT transitions (5,7). EMS induces GC to AT transitions almost exclusively, and NQO induces mutations by this pathway about 90% of the time (5). UV induces a wider variety of transition and transversion mutations, but it appears to be too weak to be very useful in actinomycetes. Because MNNG, NQO, and EMS are among the most potent mutagens to choose from for streptomycetes, the spectrum of base-pair substitutions obtainable by chemical mutagenesis is almost completely limited to GC to AT transitions.

Streptomycetes and other actinomycetes have DNA containing very high G + C content (~70%). The G + C content in the third position in codons is generally >90%. Recently, the spectrum of potential amino acid substitutions induced by different transition or transversion mutation pathways at dominant streptomycete codons (those containing G or C in the third position) has been analysed (3). The GC to AT transition pathway cannot induce mutations at phenylalanine-, isoleucine-, asparagine-, or lysine-prevalent codons, nor can it generate codons for alanine, glycine, or proline. Of the other base-pair substitution pathways, the AT to CG transversion pathway is strikingly complementary to the GC to AT transition pathway. Thus AT to CG mutations can occur at phenylalanine-, isoleucine-, asparagine-, or lysine-prevalent codons and can also generate alanine, glycine, and proline codons (3). Twenty-six different amino acid substitutions can be generated by GC to AT transitions at prevalent codons, and 25 can be generated by AT to CG transversions. Importantly, none of the specific amino acid substitutions induced by one pathway can be induced by the other. Therefore, the simplest way to broaden the spectrum of base-pair substitution mutations for yield improvement would be to develop a mutagenic protocol that is specific for AT to CG transversions.

Approach to Expand the Range of Base-Pair Substitution Mutagenesis

No chemical mutagen induces AT to CG transversions as the primary pathway. However, in *E. coli* the *mutT* mutation causes ~1,000-fold enhancement in spontaneous AT to CG transversions (5,8). Because the *mutT* gene is highly conserved in bacteria, it should be present in streptomycetes. A system for reversible gene disruption has been proposed that would allow for the screening of mutants containing spontaneous AT to CG mutations in a *mutT* background, followed by direct selection of $mutT^+$ recombinational revertants that would have normal spontaneous mutation rates (3). The implementation of such a system awaits the cloning of an actinomycete *mutT* gene.

OTHER APPROACHES TO MUTAGENESIS

Gene Duplications by Genetic Engineering

Although gene duplications can arise as spontaneous or induced mutations at low frequencies, the process has not been harnessed as a general technique to improve product yields. In some cases, gene duplications or amplifications have been recognized after the fact as being associated with yield enhancement (9–12). Gene cloning methodologies and chromosomal insertion methods have provided the means to carry out directed gene duplications. These approaches can be used to alleviate rate limitations in secondary metabolite biosynthetic pathways, thus improving product yields and product quality (3,13–15).

The general approach for directed gene duplication is to identify a neutral site in the chromosome where genes can be inserted without changing the product yield or the fermentation properties of the production strain. The neutral site can be cloned, and a gene of interest can be inserted within the neutral site on a plasmid vector. The vector is introduced into the streptomycete host by transformation or by conjugation from *E. coli*, and the gene is inserted into the chromosomal neutral site by a homologous double crossover (Fig. 2), resulting in a gene duplication and loss of the delivery plasmid (13–15). In practice, the individual single crossover can be selected sequentially. The first crossover, which integrates the plasmid into the chromosome, is selected using the antibiotic resistance marker (AR) on the plasmid. This can be accomplished with a temperature-sensitive plasmid replicon that can be readily cured at elevated temperature in the absence of homologous recombination into the chromosome. The second crossover can be screened or selected using a counterselectable marker (CSM), such as the *rpsL* gene (16). In this

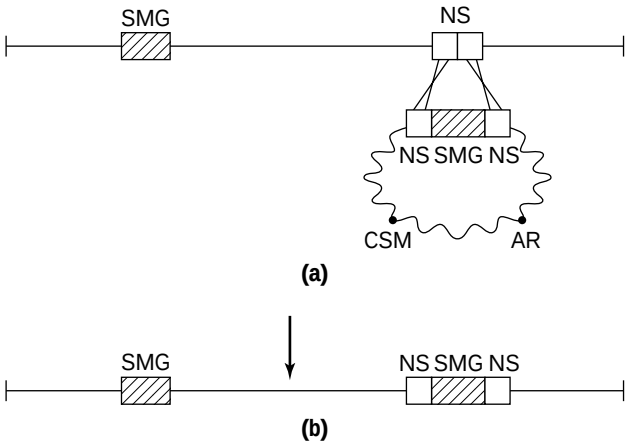


Figure 2. Gene duplication by homologous recombination of a secondary metabolite gene (SMG; crosshatched) into a predetermined neutral site (NS) from a plasmid (wavy line) containing markers for antibiotic resistance (AR) and counterselection (CSM).

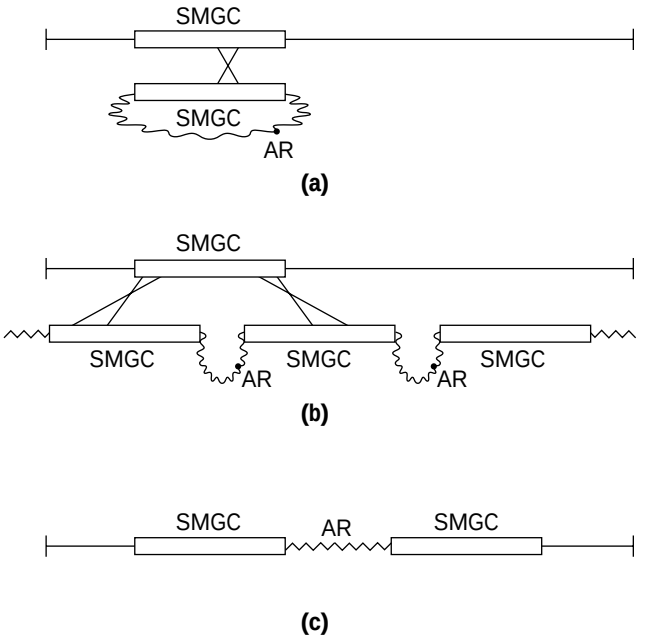


Figure 3. Gene duplication by homologous recombination of plasmid DNA into a secondary metabolite biosynthetic gene cluster (SMGC). A, Single crossover of a circular plasmid containing cloned SMGC into the chromosomal SMGC, selecting for an antibiotic resistance marker (AR); B, double crossover of a linear plasmid concatemer containing cloned SMGC into the chromosomal SMGC; C, recombinant containing SMGC duplication flanking plasmid DNA (wavy line).

case the wild-type *rpsL* gene encoding dominant streptomycin sensitivity is located on the plasmid, and the recessive streptomycin-resistant mutant allele of *rpsL* is located on the chromosome. Thus recombinants that have undergone a second crossover, deleting the plasmid, can be selected for streptomycin resistance at the temperature nonpermissive for plasmid replication. A good example of this neutral site cloning approach was the duplication of the *tylF* gene that encodes a rate-limiting step in the biosynthesis of the antibiotic tylosin in *S. fradiae*. In this case the neutral site was identified by transposon mutagenesis and cloned by selecting for the transposon marker that was expressed in streptomycetes and in *E. coli* (17,18).

Another variation of this method is to introduce the cloned gene into a plasmid that can insert site specifically into a plasmid or phage chromosomal insertion site, providing that the insertion site is neutral (15). A good example of this was the duplication of a gene involved in a rate-limiting step in pristinamycin biosynthesis by inserting a copy of the *snaAB* genes under the transcriptional control of the strong promoter *ermEp** into the plasmid pSAM2 *attB* site (19).

Duplication of much larger blocks of genes (~40 kb) can be accomplished by delivering the genes on a conjugal cosmid cloning vector lacking actinomycete replication functions (15,20). Recombinants are formed by homologous single crossovers of circular plasmids, or double crossovers of linear plasmid concatemers (Fig. 3), duplicating the complete ~40-kb region of interest. In this case, it is not predictable a priori if the insertion will be neutral, but in many cases it is, because ~50% of chromosomal sites in streptomycetes may be neutral (17). With this method of gene cluster duplication, the recombinants retain the plasmid sequences in the chromosome.

These methods also have the advantage that the recombinants are generally stable and lack self-replicating plasmids, which often reduce the productivity of the fermentation (13–15). Furthermore, the neutral site insertion method by double crossover results in a recombinant containing no heterologous DNA, so this method is in effect a directed gene duplication and directed translocation procedure that can give a very high frequency of the desired mutant that might otherwise occur randomly at an exceedingly low frequency, if at all.

Insertional Inactivation of Regulatory Elements by Transposons

Both positive and negative regulatory elements can influence the levels of production of secondary metabolites (21,22). Because transposons can be used to insertionally inactivate genes, they should be useful to disrupt regulatory genes. Disruption of negative regulatory genes should cause enhanced production of the desired products, whereas disruption of positive regulatory genes should cause reduced production. In both cases, the regulatory gene can be isolated by cloning the transposon and the flanking chromosomal DNA. This approach has been facilitated in IS*493*-derived transposons (17,23,24) by using markers that are selectable in streptomycetes and in *E. coli* (e.g., hygromycin resistance or apramycin resistance). In the case of negative regulatory genes, the cloned genes can then be used to make deletion mutants. Cloned positive regulatory genes can be used to make directed gene duplications to enhance production, as discussed in the previous section.

Random Insertion of Promoters by Transposition

The concept of using transposons as portable promoters was first proposed by Kleckner et al. (25). Several transposons are now available for streptomycetes, and those derived from IS*493* have outward reading promoter activity (18). Therefore this concept can now be tested using the natural promoter activity or by cloning stronger or regulatable promoters into a transposon. A preliminary study in the *Streptomyces roseosporus* indicated that Tn*5099* can induce mutations that cause enhanced daptomycin production (26). It has not been determined if the phenotype of these mutants is due to the portable promoter, to the inactivation of negative regulatory elements, or to both mechanisms. The data indicate, however, that transposon mutagenesis could become an important adjunct to robust industrial mutagenesis programs.

Biosynthetic Pathway Amplification

A somewhat more speculative approach to yield enhancement is by the duplication or amplification of the complete biosynthetic pathway. Because many secondary metabolite biosynthetic gene clusters are very large (e.g., 50 to 100 kb), it is not possible to duplicate them using standard cosmid cloning vectors. Recent genetic mapping experiments indicate that a number of secondary metabolite gene clusters are located close to one end of their respective linear chromosomes (3). In *Streptomyces rimosus*, which has a linear chromosome and a linear plasmid (11,27), the oxytetracycline biosynthetic genes are located about 600 kb from one end of the chromosome. One strain with improved oxytetracycline production contained 3 to 4 copies of a 1-Mb chimeric plasmid containing one plasmid end, one chromosomal end, and the complete oxytetracycline gene cluster. The strain also had one copy of the oxytetracycline gene cluster on the chromosome. In terms of classical mutagenesis, this strain harbors a duplication, translocation, and amplification because of the multicopy nature of the chimeric plasmid. This example of secondary metabolite pathway amplification has suggested a more directed homology-driven recombinational approach that might be applied to other secondary metabolite biosynthetic pathways clustered near the ends of linear chromosomes (3). Because linear plasmids are relatively common in streptomycetes (21,28), this could be a fruitful approach to assure that the secondary metabolite biosynthetic enzymes are not limiting, thus allowing other beneficial mutations influencing the flow of precursors into the pathway to be fully expressed phenotypically. Much more work is needed to fully develop this approach, but the potential benefits would appear to justify the additional studies needed to test this concept in other actinomycetes.

BIBLIOGRAPHY

1. J. Stonesifer and R.H. Baltz, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1180–1183 (1985).
2. R.H. Baltz, *Industrial Microbiology and Biotechnology*, American Society of Microbiology, Washington, D.C., 1986, pp. 184–190.
3. R.H. Baltz, *J. Ind. Microbiol. Biotechnol*, **20**, 360–363 (1998).
4. R.H. Baltz and J. Stonesifer, *J. Bacteriol.* **164**, 944–946 (1985).
5. C. Coulondre and J.H. Miller, *J. Mol. Biol.* **117**, 577–606 (1977).
6. J.H. Miller, *Annu. Rev. Genet.* **17**, 215–238 (1983).
7. C.G. Cupples and J.H. Miller, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5345–5349 (1989).
8. R.M. Schaaper, B.I. Bond, and R.G. Fowler, *Mol. Gen. Genet.* **219**, 256–262 (1989).
9. D.J. Smith, J.H. Bull, J. Edwards, and G. Turner, *Mol. Gen. Genet.* **216**, 492–497 (1989).
10. J.L. Barredo, B. Diez, E. Alvarez, and J.F. Martin, *Curr. Genet.* **16**, 453–459 (1989).
11. B. Gravius, D. Glocker, J. Pigac, K. Pandza, D. Hranueli, and J. Cullum, *Microbiology* **140**, 2271–2277 (1994).
12. U. Peschke, H. Schmidt, H.-Z. Zhang, and W. Piepersberg, *Mol. Microbiol.* **16**, 1137–1156 (1995).
13. R.H. Baltz and T.J. Hosted, *Trends Biotechnol.* **14**, 245–250 (1996).
14. R.H. Baltz, *Drugs Pharmaceut. Sci.* **82**, 49–62 (1997).
15. R.H. Baltz, *Trends Microbiol.* **6**, 76–83 (1998).
16. T.J. Hosted and R.H. Baltz, *J. Bacteriol.* **170**, 180–186 (1997).
17. P.J. Solenberg, C.A. Cantwell, A.J. Tietz, D. Mc Gilvray, S.W. Queener, and R.H. Baltz, *Gene* **168**, 67–72 (1996).
18. R.H. Baltz, M.A. McHenney, C.A. Cantwell, S.W. Queener, and P.J. Solenberg, *Antonie van Leeuenhoek*, **71**, 179–187 (1997).
19. G. Sezonov, V. Blanc, N. Bamas-Jacques, A. Friedman, J.-L. Pernodet, and M. Guerineau, *Nat. Biotechnol.* **15**, 349–353 (1997).
20. P. Matsushima, M.C. Broughton, J.R. Turner, and R.H. Baltz, *Gene* **146**, 39–45 (1994).
21. R.H. Baltz, *Bioprocess Technol.* **22**, 309–381 (1995).
22. M. Bibb, *Microbiology* **142**, 1335–1344 (1996).
23. R.H. Baltz, D.H. Hahn, M.A. McHenney, and P.J. Solenberg, *Gene* **115**, 61–65 (1992).
24. P.J. Solenberg and R.H. Baltz, *Gene* **147**, 47–54 (1994).
25. N. Kleckner, J. Roth, and D. Botstein, *J. Mol. Biol.* **116**, 125–159 (1977).
26. M.A. McHenney and R.H. Baltz, *Microbiology* **142**, 2363–2373 (1996).
27. K. Pandza, G. Pfalzer, J. Cullum, and D. Hranueli, *Microbiology* **143**, 1493–1501 (1997).
28. H. Kinashi, *Actinomycetologica* **8**, 87–96 (1994).

See also GENE TRANSFER, GRAM POSITIVE BACTERIA; GLUTAMIC ACID PRODUCING MICROORGANISMS; INSECT CELLS AND LARVAE, GENE EXPRESSION SYSTEMS; METABOLITES, PRIMARY AND SECONDARY; SECONDARY METABOLITE PRODUCTION, ACTINOMYCETES, OTHER THAN STREPTOMYCES.

MYXOBACTERIA

HANS REICHENBACH
Gesellschaft für Biotechnologische Forschung
Braunschweig, Germany

KEY WORDS

Adsorber resin
Antibiotics fermentation
Cellulose degraders
Immobilization
Liquid cultures
Myxobacteria
Myxovirescin

OUTLINE

INTRODUCTION

Myxobacteria are Gram-negative, rod-shaped, strictly aerobic soil bacteria that move by gliding or creeping and are famous for their cell–cell interactions, culminating in fruiting body formation. Until now, no actual technical process has been performed with myxobacteria, but they have already been grown in industrial-scale bioreactors for the production of certain secondary metabolites. Myxobacteria certainly have a potential for technical applications. They could be used as producers of enzymes, for example, proteases to be applied for milk clotting (1,2); enzymes degrading xylan, chitin, cellulose, or yeast cell wall β-glucans (sometimes with unusual cleavage patterns, such as converting starch into trisaccharides) (3); or unusual enzymes for biochemical research such as restriction endonucleases (4,5), reversed transcriptases (6,7), or protein kinases (8). All these possibilities have been very little explored because relatively few scientists have studied myxobacteria in the past, and very few study them at present on a broader scale for such applications. As myxobacteria have efficient export systems for exoenzymes and proteins, *Myxococcus xanthus* was investigated for some time as a host of foreign genes coding for useful proteins (9,10). The advantage in this application was seen in the fact that myxobacteria are nonpathogenic and apparently free of endotoxin. Also, less obvious applications are conceivable. Thus, an anticoagulant, myxalin, has been described from *M. xanthus* that could be useful in medicine (11,12). Lectins have been isolated from myxobacteria and might be further explored (13,14). However, the most promising application of myxobacteria may be in the field of secondary metabolites. So far, about 80 different basic structures have been isolated and elucidated from these organisms, with roughly 350 structural variants (1,15–18). Two compounds, soraphen and epóthilon, proceeded to the stage of an industrial development project. Although soraphen had to be given up as a fungicide for plant protection because of its intolerable side effects (19), epóthilon still is under investigation as a cytotoxin for antitumor application (20).

WHAT KIND OF ORGANISMS ARE MYXOBACTERIA?

Several up-to-date reviews about myxobacteria in general (21,22) and, more specifically, about their isolation, cultivation, preservation, and taxonomy are available (23–25). These facts, therefore, need be only briefly summarized here. Also, citations of the older literature are given in the recent reviews. Only fermentation of myxobacteria is discussed in some detail in this article.

Distinguishing Characteristics of the Myxobacteria

The myxobacterial cells are relatively large, between 3.5 and 12 μm long and from 0.6 to 1 μm wide, depending on the species and, to some extent, on growth conditions. Two morphological types, representatives of two suborders, can readily be distinguished under the microscope: the cells are either slender with tapering ends (*Cystobacterineae*) or are cylindrical with rounded ends; the latter are usually in the shorter size range and often more stout (*Sorangineae*).

The cells move by gliding over or within the substrate. Therefore the colonies spread, sometimes over the whole culture plate within a couple of days, and are named swarm colonies or swarms. Gliding motility is much reduced or totally suppressed on media with a high nutrient content, such as those containing 1% or more peptone. When transferred into a liquid medium, myxobacteria grow at first as a film along the walls of the container or, when shaken, as nodules or flakes suspended in the medium. After several transfers, often homogeneously growing strains can be obtained. To achieve this, transfers should be made from the supernatant, avoiding clumps as much as possible, at short intervals and using small volumes (20 mL in 100-mL flasks). Also, the composition of the medium may play a role. In general, strains belonging to the suborder *Sorangineae* are more recalcitrant than those of the *Cystobacterineae*. For these, weeks and months may be required before a specific strain grows as single cells, and sometimes this stage is never obtained. For this reason, until the 1960s it was generally believed that myxobacteria cannot be grown in liquid media, an opinion that definitely delayed their research.

While gliding bacteria are found in most branches of the bacterial phylogenetic system, two other characteristics of the myxobacteria are exceptional and even unique. Myxobacterial cells seem to communicate with one another with the aim to stay together. This can be seen in the behavior of the swarms. While individual cells can move much faster than the swarm expands, they still keep together, returning whenever they push beyond the swarm edge or waiting for the other cells to catch up. Small swarms may even migrate as a whole, like a pseudoplasmodium. Within the swarm, patterns of veins, rings, and parallel rows of cells in a ripplelike arrangement may develop. The cells in the ripples appear to move back and forth, creating in this way the illusion of oscillating waves, particularly striking when seen in time-lapse films but also recognizable by the patient observer under the microscope. The culmination of intercellular communication is the production of fruiting bodies. Under unfavorable conditions, especially starvation, hundreds of thousands to millions of cells aggregate at certain sites within the swarm. They often mill around for some time on a ring-shaped path and pile up. From such cell masses, fruiting bodies may differentiate. The simplest types are just knobs of cells and soft slime (*Myxococcus*); the most sophisticated ones consist of an excreted slime stalk, which may even be branched, with clusters of sporangioles at the end (*Chondromyces*). These sporangioles consist of a tough wall enclosing hardened slime and cells. In other genera, sporangioles are also found single on a stalk or sitting directly on or within the substrate, often in large, dense clusters. The size of the fruiting bodies varies between 10 and 1,000 μm, and most of them can readily be recognized with the naked eye. They usually are brightly colored (yellow, orange, red, brown, black), and are often produced in large numbers. The biological significance of the fruiting bodies is apparently to guarantee that a new life cycle is started with a community rather than a single cell. Why this strategy would be favorable will be seen when the physiology of the myxobacteria is discussed.

Inside the maturing fruiting body, the vegetative cells convert into myxospores. In the suborder *Cystobacterineae*, the vegetative cells always change their shape dramatically. They shorten and fatten—in the family *Myxococcaceae* they even become spherical—and turn optically refractile. They surround themselves with a capsule, which in most cases, however, is seen only under the electron microscope. In the suborder *Sorangineae*, the shape change is less spectacular; no capsule seems ever to be developed, but the cells also become optically refractile. Myxospores are fully desiccation resistant, in contrast to vegetative cells that are rather sensitive to drying. Temperature resistance is only mildly increased: wet myxospores may survive at 58–60 °C for 30–60 min, just long enough to get rid of contaminants in many cases. Dry myxospores may be heated to 140 °C for 1 h and still germinate. Thus, myxospores clearly are the myxobacteria's survival strategy, and fruiting must have some other significance.

There Are Two Physiological Types of Myxobacteria

All myxobacteria are specialized in the decomposition of biomacromolecules. Most species are bacteriolytic, that is, they excrete a host of hydrolytic enzymes—proteases, nucleases, lipases, polysaccharidases—that decompose living and dead, sometimes only dead, bacteria, yeasts, and other microorganisms, and of course also organic matter coming from higher organisms. One group, the genus *Sorangium* and a second as yet unnamed related genus, decompose cellulose. These organisms seem to attack only dead microorganisms, such as autoclaved *E. coli* and baker's yeast. Because a population obviously is more efficient than a single cell in the decomposition of environmental macromolecules by creating high enzyme levels, minimizing diffusion losses, and maximizing recovery of solubilized monomers, it becomes understandable why myxobacteria developed a social behavior.

Myxobacteria of both physiological types can, as a rule, be cultivated, and most of them in pure culture. A few, seemingly rare myxobacteria have never been grown in culture, mainly the *Haploangium* species, *Polyangium parasiticum* (discovered on the freshwater green alga *Cladophora*), and a few other species that have never been seen again by anybody after their original description from collected samples or crude cultures. A case of obligatory symbiosis has recently been reported for *Chondromyces crocatus* (26,27). Most strains of that species in our collection grow only if a companion is present. As the companion, too, cannot grow without *Chondromyces*, both could be cultivated together in large bioreactors, where they keep themselves automatically in balance (28). The companion could be characterized by 16*S* rRNA sequencing. It belonged, in all our *Chondromyces* strains, to the same species, a new member of the *Flavobacterium* group, regardless from where the *Chondromyces* had been isolated. The intimacy of the relationship is emphasized by the fact that the companion can even be seen within the sporangioles of the *Chondromyces* fruiting body (27), whereas normally the interior of myxobacterial fruiting bodies is free of contaminants for some time after their production.

All myxobacteria, the cellulose degraders included, can be grown on yeast agar, such as VY/2 agar (baker's yeast, 0.5% by fresh weight of yeast cake; $CaCl_2 \cdot 2H_2O$, 0.1%; vitamin B_{12}, 0.5 mg/L; agar, 0.5%; pH adjusted to 7.2 with KOH; autoclaved). On this medium, spreading colonies are obtained; the cultures stay alive for at least 2–3 weeks (30 °C) and often produce fruiting bodies. Fruiting bodies are useful when the strain is to be preserved by lyophilization, because vegetative cells cannot reliably be dried. Further, fruiting bodies are required for classifying an isolate (see following). Fresh isolates usually form fruiting bodies, but during cultivation fruiting body formation usually is lost after three to eight transfers. Fruiting body formation can be maintained, if cultures are started from fruiting body preserves. Those can be prepared by streaking swarm material onto small pieces of filter paper placed on water agar. After incubation for 2–3 days at 30 °C, often fruiting bodies are produced. They are allowed to mature for another 4–6 days. The filter papers are then inserted into sterile medicine bottles and dried in a desiccator under vacuum for a week. From dry filter paper preserves stored at room temperature, fruiting cultures can still be started after 8–20 years. The fruiting cultures must then be used after one to

two transfers to make fresh preserves; otherwise, fruiting may be lost.

The bacteriolytic myxobacteria usually also grow well on peptone-containing media, for example, CY agar (Casitone, Difco, 0.3%; yeast extract, Difco, 0.1%; $CaCl_2 \cdot 2H_2O$, 0.1%; agar 1.5%; pH adjusted with KOH to 7.2; autoclaved). On this medium growth often is heavier, but the swarms stay more compact, are often more slimy, rarely produce fruiting bodies, and die sooner (after 8–10 days) than on VY/2 agar. The reason for early death could be the production of ammonia and the concomitant shift of pH to alkaline conditions. The pH range of myxobacteria is 6.8 to 7.8. Of course, the media could be buffered. A useful buffer is HEPES, which is tolerated without adverse effects up to at least 100 mM. Several bacteriolytic myxobacteria seem to grow only on peptone, proteins, or amino acids and seem not to utilize carbohydrates at all (the genera *Myxococcus*, *Corallococcus*, and *Nannocystis*). However, there may be some unorthodox ways of carbohydrate utilization. Thus, *Myxococcus* degrades yeast cell wall β-glucan, and *Nannocystis* corrodes agar, so why should they leave the valuable decomposition products to other organisms? At least some *Corallococcus* strains decompose starch to tri- and oligosaccharides; those are utilized, while glucose and maltose remain untouched (3). (For more media recipes, see Ref. 24.)

Incidentally, in spite of their name of "slime bacteria," myxobacteria are poor producers of slime on plates and in liquid media. Only rarely colonies are somewhat raised and liquid media become recognizably viscose. Ropy liquids have never been reported. The slime of the swarm sheet, however, may become extremely tough, rubberlike, and almost impossible to divide with the inoculation loop.

For the cellulose degraders, too, alternative media are available (24), for example, filter paper on mineral salts agar with KNO_3 as the only nitrogen source (e.g., ST21 agar), on which they often form huge numbers of fruiting bodies, or mineral salts–glucose agar (e.g, CA2 agar). The organisms seem to prefer NO_3^- to NH_4^+ as the nitrogen source, grow even better in the presence of low peptone concentrations, but often are absolutely dependent on a carbohydrate (e.g., glucose, starch, cellulose) for carbon and energy. All myxobacteria can be grown in liquid media. This aspect is discussed in connection with fermentation in the next section.

Taxonomy and Genetics of the Myxobacteria

The taxonomy of the myxobacteria rests almost exclusively on morphological characteristics. Only very few useful physiological traits have been found; thus, cellulose degradation in *Sorangium* and synthesis of lanosterol in *Polyangium* (29) distinguish the two genera. A bright yellow fluorescence at 366 nm of the swarm colonies caused by the production of phenalamid (30,31) helps to distinguish *Myxococcus stipitatus* from *M. fulvus* when typical fruiting bodies are lacking. While suborders, families, and most genera can usually be identified without problems because the morphology of their vegetative cells, myxospores, and swarm colonies is sufficiently typical and can normally reliably be registered, differentiation of species is often not possible, because it requires typical fruiting bodies and those may be degenerate or completely lacking. In such cases, one can only guess what species one has. Today, molecular taxonomy could be applied to solve the problem (DNA–DNA hybridization).

Presently about 40 different species are recognized, but there certainly are more species in nature, at least another 20 to 30. They are classified in 12 genera, 4 families, and 2 suborders in the order *Myxococcales*. As already mentioned, the suborders can easily be distinguished by the shape of the vegetative cells. As demonstrated by 16*S* rRNA sequencing studies, the order is phylogenetically coherent and belongs to the δ-branch of the Proteobacteria (32). Complete 16*S* rRNA sequences are available for about 40 strains of many different genera and species, but the resulting picture does not much modify the established structure of the order. Clearly, the resolution of the technique is not particularly helpful at the species and genus level.

It should be mentioned that formerly, before it was appreciated that gliding motility is not a useful taxonomic characteristic, the genera *Cytophaga, Sporocytophaga*, the *Cytophaga*-like bacteria, and *Flexibacter* were classified as nonfruiting myxobacteria. In fact, they belong to a phylum of their own and are not related to the myxobacteria at all. This is of some practical importance, because several fish-pathogenic and allergenic species are known in that group, while no pathogenic myxobacteria have ever been reported. Another group, later recognized as a new genus, *Lysobacter*, was even classified as *Myxobacter* and *Sorangium* for some time. Those names can still be found occasionally in the current literature. *Lysobacter* comprises biotechnologically useful strains—suppliers of well-known antibiotics, such as myxin, and of interesting enzymes—but has nothing to do with myxobacteria. Phylogenetically the genus belongs into the γ-branch of the Proteobacteria.

The DNA of myxobacteria has a high G+C content of 66 to 72 mol%. The genomes are large for bacteria, between 9,500 and 10,000 kbp. Physical genome maps are available for *Myxococcus xanthus* (33) and *Stigmatella aurantiaca* (34). Two myxobacteria, *Myxococcus xanthus* and *Stigmatella aurantiaca*, have been analyzed with genetic methods as model systems for questions of morphogenesis, cell signaling, gliding motility, carotenoid synthesis, protein export, etc. (for details, see Refs. 21,22). Several gene transfer techniques have been developed: *E. coli* phage P1 suicide vectors, conjugation with *E. coli*, myxobacterial generalized transducing phages, and electroporation (21). Transposon mutagenesis has successfully been applied to gene tagging (35). In a few cases, those techniques have also been used for other myxobacteria, for example, *Sorangium cellulosum* (36,37), but there is always need of an adaptation.

Myxobacteria Are Very Common Soil Organisms

One might expect that organisms that are so unusual and fascinating as the myxobacteria, yet are so little investigated, are rare and difficult to obtain. The contrary is the case. Myxobacteria are found worldwide in all climate zones and vegetation belts and are very common every-

where. They live in soil, on rotting plant material, on decaying wood, on bark of living and dead trees, and on dung of various herbivores, especially rabbits, hares, deer, goats, and sheep; in general, wherever there is a rich microbial flora. The most diversified myxobacterial communities are found in warm, semiarid areas, like Egypt, the southwestern United States, or northern India, probably because, with their myxospores and fruiting bodies, they are beautifully adapted to such environments. Myxobacteria also seem to live in freshwater and to survive for some time in marine sediments; they can be isolated from marine sediments but are not halotolerant and thus probably cannot live there. They seem not to be able to grow at a low pH, and acid soils are the one habitat free of myxobacteria. Alkalophilic myxobacteria, however, have been obtained from alkaline lakes in Africa. Myxobacterial are typical mesophiles, although psychrophilic strains have been isolated from Antarctic samples.

Isolation and Maintenance of Myxobacteria

The relevant methods of maintenance of these bacteria have relatively recently been described in great detail (24), so a brief summary may suffice here. Because Myxobacteria have a rather generalized metabolism, a highly selective enrichment is not possible. The isolation techniques make use of gliding motility, the ability to attack living bacteria or cellulose, and resistance to certain antibiotics. There are essentially three ways to start enrichment cultures. In the first procedure, materials collected in nature are incubated in a moist chamber at 30 °C. Suitable sources are rotting wood, bark, or the dung of herbivores. To suppress growth of fungi, cycloheximide (50 mg/L) may be added to the water used to wet the source material. The cultures are inspected under a dissecting microscope with incident light at magnifications of 20–80×. Fruiting bodies may form within 2–8 days on the surface of the substrate and can be used to make transfers. From soft slimy *Myxococcus* fruiting bodies, myxospores can be transferred directly to a growth medium such as CY agar by cautiously touching the top of the fruiting body with the sharp tip of a needle on a 1-mL syringe. The young fruiting bodies are normally free of contaminants, so that often a pure culture is obtained in one step. In a similar way, sporangioles can be removed from the ends of the long stalks of *Chondromyces* fruiting bodies. The fruiting bodies of all other myxobacteria cannot be harvested without touching the heavily contaminated substrate and transferring pieces of it together with the fruiting bodies. Therefore, in these cases the fruiting bodies are transferred to streaks of living *E. coli* on water agar plus cycloheximide for further enrichment. The developing swarms usually expand quickly beyond the streak of the feed bacteria, and either uncontaminated sections from the swarm edge or fruiting bodies can be inoculated on growth medium. Often the procedure has to be repeated several times before a pure culture is obtained, which typically takes 2 weeks to 3 months. Sometimes special purification steps have to be included, such as heating fruiting bodies, suspended in water, at 58–60 °C for 10–60 min; shaking fruiting bodies overnight in a rich medium (in which they will not germinate) with high doses of suitable antibiotics; or gassing cultures with ammonia vapors to eliminate soil amoebae.

In the second approach, cultures are started by inoculating circular spots of living *E. coli* (about 2 cm in diameter) on water agar plus cycloheximide with small samples of soil or rotting plant material. After 2–3 days at 30 °C, swarm colonies and soon also fruiting bodies may develop, and transfers and purification steps are performed as described in the first approach. With this method, different types of samples can be processed, and myxobacteria with small, inconspicuous fruiting bodies are also obtained. Incidentally, on such plates often other gliding bacteria such as *Cytophaga*-like bacteria, *Flexibacter*, *Taxeobacter*, and *Herpetosiphon* also develop and can be recognized by their spreading colonies. In the third method, the cellulose degraders are isolated by inoculating filter paper pads on mineral salts plus cycloheximide agar with soil, rotting wood, decaying plant material, or dung of herbivores. After 10–14 days at 30 °C, bright yellow, orange, brown, or black spots around the inoculum indicate growth of *Sorangium* strains. Those usually produce huge numbers of densely packed fruiting bodies, which in fact are responsible for the colored areas. Transfers are made first to small pieces of filter paper on mineral salts agar, then on VY/2 agar. All *Sorangium* strains appear to be resistant to high doses of kanamycin, so that the antibiotic can be used for purification (50 mg/L kanamycin sulfate). Cellulose degraders are more tedious to purify, and it may take 1–2 years before the last contaminant is eliminated. Fortunately, pure strains are usually obtained within 1–3 months. In crude cultures for cellulose degraders, noncellulolytic myxobacteria always also develop. The inoculum, the solubilized cellulose, and other saprophytic bacteria supply sufficient nutrients for their growth. In particular, *Nannocystis*, *Polyangium*, and *Chondromyces* strains are often found in this way. They must, of course, be transferred to *E. coli* streaks on water agar.

Preservation of strains is done best by freezing 0.5- to 1-ml suspensions of cells in a 1% Casitone (Difco) solution at −70 °C in a deep freeze or in liquid nitrogen. The organisms appear to stay alive for many years (about 20 years has been experimentally proven). Freezing at −20 to −25 °C in ordinary household freezers seems not to work. In our experience, myxobacteria survive that at best for a couple of weeks. Drying of fruiting bodies has already been mentioned.

FERMENTATION OF MYXOBACTERIA

Fermentation of myxobacteria at a 5- to 50-m^3 scale has only been performed for the production of secondary metabolites. Data are available from many publications. Process descriptions are also found in a large number of published patents. Consequently, the following discussion centers upon this topic.

General Principles

With homogeneously growing strains, typical growth curves can be obtained. The generation times are long, typically between 4 and 14 h. The faster growing species are

normally those of the suborder *Cystobacterineae*. On the other hand, those species usually show short stationary phases; that is, soon after the end of the log phase is reached, the culture breaks down. This collapse can be really dramatic. Cultures of *Stigmatella aurantiaca* strains in peptone medium without carbohydrate become pitch black within 2–3 h, probably because of the activity of phenol oxidases from lysed cells, and are then dead. Therefore, transfers are best made in the upper log phase, and extended production in the stationary phase is usually not possible. Species of the suborder *Sorangineae*, in contrast, often are under stationary conditions for days and may continue to produce secondary metabolites during this period. With these species it is also advisable to make transfers in the upper log phase, because later the inoculum may result in a very long lag phase or will no longer start growing. Obtainable cell densities typically are 10^8–10^{10} cells/mL depending on the strain and the medium. It should be remembered, however, that myxobacterial cells are larger than those of most other myxobacteria. The biomass thus is often between 4 and 12 g wet weight/L, corresponding to 1–3 g dry weight. As already mentioned, slime production is normally not seen.

Liquid media suitable for bacteriolytic myxobacteria are all based on peptones or proteins, while those for the cellulose degraders can be varied in wider limits. Enzymatically digested casein peptones are the substrates of choice, for example, Difco Casitone or Marcor peptone from casein, tryptically digested. Peptones from meat and acid-hydrolyzed peptones are much less useful or outright inhibitory, possibly because of elevated phosphate concentrations (38) or an imbalance in amino acid composition. Yet some myxobacteria can also be grown on casamino acids (e.g., Ref. 39). Examples of simple media suitable for many bacteriolytic myxobacteria are CAS lm (Casitone, Difco, 1%; $MgSO_4 \cdot 7H_2O$, 0.1%; pH 6.8; autoclaved); if the peptone concentration is too high, which is the case with quite a few myxobacteria, MD1 lm may be used (Casitone, Difco, 0.3%; $CaCl_2 \cdot 2H_2O$, 0.07%; $MgSO_4 \cdot 7H_2O$, 0.2%; cyanocobalamin, 0.5 mg/L; standard trace element solution; pH 7.0; autoclaved). For liquid cultures of *Sorangium* strains, AMB lm (40) could be used (Casitone, Difco, 0.25%; soluble starch, 0.5%; $MgSO_4 \cdot 7H_2O$, 0.05%; K_2HPO_4, 0.025%; pH 6.8; autoclaved). Those organisms also grow, however, on much simpler media containing only KNO_3 as a nitrogen source and glucose as a carbon source in a mineral salts solution (e.g., CK1 and CK6 lm).

Casein peptones contain sufficient phosphate, so that extra phosphate need not be added. Thus, 0.5 mM phosphate was determined in a 0.2% casein peptone solution (38). In general, myxobacteria are rather sensitive to phosphate. Also it may have been realized that the media mentioned contain relatively high concentrations of $MgSO_4$; the optimum is indeed often at 5 to 10 mM. Calcium is normally present in adequate amounts or is not specifically required. In contrast, the cellulose degraders would not grow without an addition of $CaCl_2$ (41). Vitamins are usually not required, but an occasional strain may depend on vitamin B_{12}.

A peptone medium may be complemented with 0.2–1% of a carbohydrate, for example, glucose, sucrose, or starch. The latter is usually readily hydrolyzed. A carbohydrate is a must for most strains of *Sorangium*. Normally cultures of bacteriolytic myxobacteria do not grow faster in the presence of carbohydrate but rather achieve higher cell density. Also, the pH does not rise as fast and as high as in a pure peptone medium. The details of carbohydrate metabolism have not yet been studied for a wide variety of myxobacteria. Some appear not to utilize carbohydrates because they lack essential glycolytic enzymes (hexokinase and pyruvate kinase in *Myxococcus xanthus*; 42). Others utilize only polysaccharides, because they can only ingest tri- and oligosaccharides (*Corallococcus coralloides*; 3). In this case it was shown that the glucose produced within the cell is mainly utilized via the pentose phosphate pathway, probably only to produce ribose and deoxyribose for nucleic acid syntheses. Many bacteriolytic myxobacteria, however, metabolize also mono- and disaccharides under production of acid and radioactive CO_2, if labeled sugars are supplied (e.g., *Stigmatella aurantiaca*; 43). Nothing is known about the biochemistry of sugar metabolism in those organisms. The cellulose degraders utilize sugars very efficiently for energy and all kinds of syntheses, as is suggested by their growth in the minimal media just described. Fully synthetic, defined media have also been reported for several bacteriolytic myxobacteria. They contain various amino acids and can in fact be very simple. *Cystobacter fuscus* and *Cystobacter ferrugineus* strains grow, for example, in HP16 Im with (per liter) L-Gln·Na, 1 g; L-Ile, 40 mg; L-Phe, 40 mg; glucose, 5 g; $MgSO_4 \cdot 7H_2O$, 1 g; standard trace elements; in phosphate buffer, 0.5 mM, pH 6.5. Some strains become vitamin dependent on thiamine or biotin, when grown in this minimal medium, while in peptone media they are not. For most strains, defined media are, however, more complex, containing more amino acids and still other ingredients. Apparently the relative concentrations of amino acids are as important as their kind. Also, every strain may require a medium of its own. Essential amino acids often are the branched-chain and aromatic amino acids. In any case, growth in those defined media is always substantially slowed down, and cell yields are much reduced.

When an organism relies solely on amino acids for nitrogen, carbon and energy, elimination of NH_3/NH_4^+ becomes a problem. This problem is indeed a serious one with bacteriolytic myxobacteria. The intracellular ammonia pools may become very high under such circumstances: in *Myxococcus virescens* 80–140 mM has been measured (44). Although the pH shift by ammonia can easily be controlled in the bioreactor by titration and in shake flasks by buffering with HEPES or MOPS, for example, the regulatory effects of ammonium still may interfere substantially with growth and production of antibiotics. This was demonstrated by pumping the contents of a small bioreactor through a module with hollow fibers made of a hydrophobic membrane. The ammonia in the culture supernatant could thus be reduced from 35–42 mM to 3–7.5 mM. In consequence, the generation time fell from 4 to 2 h, the cell density more than doubled, and myxovirescin production rose from 8 to 115 mg/L (K. Gerth, personal communication). A technical solution of this problem is presently not available.

For large-scale fermentations to be cheaper, technical substrates may be desirable. A wide variety of such substrates have been tested with myxobacteria and several have been found useful: soybean flour and meal (must be defatted), single cell protein (Probion, Hoechst, no longer available), skim milk, maize gluten, corn steep powder, or suspensions of baker's yeast (45). Which substrate should be used in a specific case must be tested. Also, as is typical for such poorly defined materials, considerable differences are observed from batch to batch and from supplier to supplier. Production of secondary metabolites may vary substantially with the particular substrate used (39,45,46). An unexpected discovery was that soybean meal may be free of iron, so that the addition of $Na{\cdot}Fe^{III}$ EDTA (8 mg/L) becomes essential (47).

The temperature range for growth of myxobacteria is rather wide: they still grow in the refrigerator at 6 °C, and quite a few strains still grow at 38 or even 40 °C. The temperature curves are skewed as usual, and the optimum is close to the maximum by 1–3 °C. Useful temperatures for fermentations are much narrower, normally 28–32 °C. At least with respect to secondary metabolites, production often goes down or stops completely at higher temperatures even if growth still is satisfactory. Also, at higher temperatures the generation time may rise (47), and the cultures become more sensitive to unfavorable pH and oxygen conditions and quickly break down as soon as the stationary phase is reached. At temperatures below 28 °C, growth, production, and process performance slow down and become unfavorable.

The pH range of myxobacteria is narrow, as a rule 6.8 to 7.8. Cellulose degraders tolerate a slightly lower pH, down to about 6.2, but with ever more reduced growth rates. If a utilizable carbohydrate is present, the pH drops in unbuffered media but can be controlled by titration with a 10% KOH solution. Without carbohydrate or after carbohydrate is used up, the pH always rises and usually becomes decidedly alkaline (pH 8.0–8.5). This change can be prevented by adding 5% sulfuric acid or 30% acetic acid. In the latter case care must be taken because acetate concentrations exceeding 0.1% may soon become inhibitory for production and growth. In a process, pH can often be stabilized by feeding glucose whenever pH reaches a set limit.

Although myxobacteria are strict aerobes, their oxygen demand is moderate because of slow growth and modest biomass. This requirement may become a problem if the aeration rate cannot be lowered unrestrictedly in a bioreactor. In fermentations for secondary metabolites, the oxygen level may play a crucial but unpredictable role for production. There are examples of a high as well as a low requirement of oxygen for optimal biosynthesis. In Table 1, the aeration schedules of a selection of such fermentations are listed. It should be understood that in no case have strictly standardized and optimized fermentations been performed. Maximum total oxygen consumption of 1.8 g O_2/L during 22 h of fermentation on a single cell protein (Probion) medium has been reported for *Myxococcus fulvus* strain Mx f16 (45).

Special Problems with Secondary Metabolite Production

In principle, the situation with secondary metabolite-producing myxobacteria is not much different from that with other antibiotics-producing bacteria. Yields initially are usually, but not always, low, typically between 0.5 and 20 mg/L. However, wild strains with initial yields of 70 mg/L soraphen or 120 mg/L myxalamid (39,58) have also been isolated. In several cases, yields could be much improved, for example, to nearly 1 g/L with sorangicin or close to 2 g/L with soraphen. In other instances, yields remained low in spite of much effort, as with saframycin Mx (59).

The methods of yield improvement are the usual ones. The influence of medium components can be substantial, as already mentioned. Thus, in fermentations of *Stigmatella aurantiaca* yields varied by a factor of 3, between 3 and 11 mg/L stigmatellin and between 47 and 120 ml/L myxalamid, depending on whether corn steep powder or maize gluten (Zein) was used as the main medium ingredient (39). At the same time growth was rather poor on the Zein medium. An improvement of yield and production stability is almost always achieved if the strain is cloned and if low and nonproducing clones are eliminated. Unfortunately this is less trivial as it seems, because initially all strains grow in clumps when transferred to a liquid medium, and if homogeneous cell suspensions are available a suitable plating medium giving high numbers of single cell colonies must first be found. This can take weeks and months, because normally no colonies are formed on the usual culture media that allow excellent growth when massively inoculated. Also, every strain needs its own plating medium. In developing a plating medium, the kind and concentration of the medium components are crucial. Inclusion of autoclaved old liquid cultures may help (36). Development of single-cell colonies takes 8–12 days for most bacteriolytic myxobacteria and 14–21 days for *Sorangium* strains (30 °C).

As soon as single-cell colonies can be produced, a mutation program becomes feasible. Mutations can be induced in myxobacteria with high yields using UV irradiation or the mutagen NTG (*N*-methyl-*N'*-nitro-*N*-nitrosoguanidine, about 40 μg/mL) (60). When inducing mutations, photoreactivation has to be taken in account (61). The size of the inoculum should be 5–10% (v/v). Lower ratios, down to 1–2%, have been found possible in some cases, but this may be paid for with a longer lag phase and a lower biomass attained at the end of the fermentation. Also, a certain minimal cell density may be necessary to start solubilization of polymeric substrates, such as casein (62). Under these circumstances a substantial decrease of the mean doubling time was observed when the size of the inoculum was increased, for example, with two *Myxococcus xanthus* strains on a casein medium from 15 to 8 h and from 26 to 8.5 h, respectively, when initial cell densities were raised from 10^4 to 10^7 cells/mL (62).

A very useful strategy to improve yields is the addition to the cultures of a neutral adsorber resin such as Amberlite XAD-16 (formerly XAD-1180; Rohm & Haas). All myxobacteria appear to tolerate the resin in the culture without adverse effects on growth. So far, virtually all excreted myxobacterial metabolites have been adsorbed by the resin (46,47,49,52,55,58). Only strictly hydrophilic compounds would remain unadsorbed, but such substances appear not to be produced by the organisms or only

Table 1. Examples of Oxygen Demand during Myxobacterial Fermentations for Secondary Metabolites

Metabolite	Producer[a]	Main components of medium[b]	Volume (L)	Stirrer speed (rpm)	Aeration rate[c] (L L broth^{-1} h^{-1})	Oxygen saturation[d] (%)	Reference
Rhizopodin	Mx s	pep	300	200	6		48
Angiolam	An d	Prob	650	300–450	1.8–2.2	90–20	49
Gephyronic acid	Ar g	Prob/yex/starch	300	200	6		50
Stigmatellin	Sg a	maize/gluten	55	200	1.5–5.4		39
			265	600	2.4–3.9	90–5	
Crocacin	Cm c	Prob/starch	55	200	3.3	90–65	28
Chondramid	Cm c	Prob/starch	90	150	3.3		46
			600	50	8	90–40	
Thiangazol	Pl fu	Prob/starch	65	200	3.3	100–85	51
Phenoxan	Pl spec.	Prob/starch	65	200	3.3	90–50	52
Ripostatin	So ce	Soy/starch/Fruc	300	300	1	100–15	53
Soraphen	So ce	Soy/yex/starch	230	400	6.5		47
			4400	230	4.5		
Epothilon	So ce	Soy/yex/starch	60	250	5		54
			230	350	4.3		54
Tartrolon	So ce	Soy/yex/starch	65	150	2.3		55
Jerangolid	So ce	Soy/yex/starch	60	250	12		56
Nannochelin	Na e	pep	65	200	3.3	90–70	57

[a]Mx s, *Myxococcus stipitatus*; An d, *Angiococcus disciformis*; Ar g, *Archangium gephyra*; Sg a, *Stigmatella aurantiaca*, Cm c, *Chondromyces crocatus*; Pl fu, *Polyangium fumosum*; So ce, *Sorangium cellulosum*; Na e, *Nannocystis exedens*.

[b]pep, casein peptone; yex, yeast extract; Prob, Probion (an experimental single-cell protein of the Hoechst Company); soy, soybean meal defatted; starch is often complemented with glucose.

[c]Stirrer speed is sometimes changed during fermentation to adjust dissolved oxygen.

[d]Dissolved oxygen is sometimes increased or held constant in the later stage of fermentation.

very rarely so. The advantages of the resin are obvious: by removing the metabolites from the supernatant, end product inhibitions are prevented, the adsorbed substances are stabilized and concentrated in a small volume, and harvest is much facilitated. If the strain grows as a homogeneous cell suspension, harvest can be done simply by collecting the resin on a sieve. The metabolites are recovered by eluting the washed resin with methanol or methanol–water mixtures. Sometimes a metabolite may crystallize from the eluate. The regulatory and protective effects of the resin are clearly demonstrated when fermentations are done without it. Almost always the yields decrease dramatically, and variants of the metabolite appear that have not been seen in cultures with resin. Normally an addition of 1–2% (v/v) is sufficient. If yields go up, and depending on the metabolite, higher amounts of resin may have to be added. Freshly bought resin is strongly alkaline and has to be washed until it becomes neutral. The resin is added to the culture medium and autoclaved together with it. Resin has also been used with excellent results in large-scale fermentations (47). There may be a loss of about 10% in the bioreactor from mechanical attrition, but the recovered resin can be regenerated by washing with NaOH solution after elution of the metabolites.

Most myxobacterial substances do not inhibit the producer. There are, however, exceptions. Myxovalargin, a peptide antibiotic that destroys the cell membranes (63), blocks growth of the producer, *Myxococcus fulvus* strain Mx f65, at 6 mg/L (64). Myxobacterial saframycin has a minimal inhibitory concentration of 15 mg/L for the producing strain, *Myxococcus xanthus* Mx x48 (58). In both cases mutants with somewhat increased resistance could be obtained, but removal of the inhibitor by adding adsorber resin to the cultures is an easy solution. As foaming is often a serious problem, an antifoam agent must always be added to myxobacterial fermentations. Silicon antifoams usually have been applied because they do not interfere with the chemical purification steps. For the organisms, however, other antifoam agents are also acceptable.

Feeding of precursors was successful only in a few fermentations. Thus, the addition of 0.3% sodium acetate stimulated production of myxalamid by *Myxococcus xanthus* strain Mx x12 (58). Normally, acetate is not a limiting factor, and its addition may even be inhibitory. This situation appears to be different with propionate. Added propionate is readily incorporated into polyketides containing propionate units and may indeed stimulate their production. Propionate is only tolerated in low concentrations, however, so that real yield improvements require repeated addition of small quantities. We have never observed a stimulation of secondary metabolite production by adding certain amino acids, probably because our peptone and protein media already contain an excess supply. It has been reported however that Ala, Ser, and Gly (0.9 mg/mL each) when given to *Myxococcus xanthus* strain TA result in higher yields of antibiotic TA (identical with myxovirescin). All three amino acids are indeed incorporated into the antibiotic, Ala and Ser serving as precursors for acetate, which seems to be not as readily taken up, while Gly is a building block of the molecule directly (65). The removal of inhibitory components may also be a prerequisite for production. The catechole iron chelator, myxochelin, is only produced when the iron content of the medium is about 10^{-7} M. Production is achieved by removing the iron in the medium by treatment with the chelating resin, Chelex 100 (Bio-Rad), and adding the optimal amount of $FeCl_3$ (66).

While feeding of specific precursors is problematical, fed-batch fermentations may be very useful (e.g., 47,67). Myxobacteria prefer moderate media concentrations: 1% peptone–protein and 0.5–1% carbohydrate are the upper limits allowing unrestricted growth, and optimal concentrations for production may be even lower. As mentioned, carbohydrate metabolism normally leads to acid production, so that feeding strategies can be developed using pH shifts or titration curves, respectively, as indicators. For, as soon as carbohydrate is used up, the metabolism switches to amino acids under ammonia production with a concomitant rise of pH. Sometimes production is directly connected to carbohydrate consumption, for example, that of tartrolon (55) and soraphen (47). Fed-batch fermentations often result in much higher yields; for example, in the case of disorazol a rise from 18 to 292 mg/L was achieved, with a production rate of 21 mg/L per day for 13 days (67). A fully standardized fed-batch process has still to be reported, however.

Often certain medium components have specific effects on production without much affecting growth. The type and concentration of peptone or protein is often critical (28,46,47,56,68). Individual sugars have differential effects on yields of myxalamid and stigmatellin even when produced biomass is comparable (46). Ripostatin synthesis is stimulated when 0.3% fructose and 0.01% pyruvic acid are added (53). *Sorangium cellulosum* strain So ce12 synthesizes concomitantly sorangicin, disorazol, chivosazol, and sorangiolid. In the presence of 0.1 mg/L sodium barbital the production rate of disorazol was reduced from 21 to about 2.5 mg/L per day irrespective of whether barbital was added in early or late log phase (67). Sorangicin production was not impeded.

As is well known for other producers, individual strains of myxobacteria may show a completely different production behavior for the same compound. A comparison of four wild strains producing soraphen revealed that all four synthesized the antifungal compound with reasonable yields at 30 °C in the presence of XAD resin (47). When the temperature was raised to 36 °C, however, one strain doubled soraphen yield while another one reduced it to 25%. When 0.2% casein peptone was added to the XAD culture at 30 °C, the yields of two strains dropped to near zero, while one strain increased its yield by more than 100%. In the absence of XAD the strains decreased their yields by 50% to 100%.

Efforts have been made to optimize myxovirescin production by *Myxococcus virescens* strain Mx v48 in a medium containing 1% casein peptone (69). The antibiotic is only synthesized under good oxygen supply. In normal batch cultures, antibiotic production started at the end of logarithmic growth at about 20 h after inoculation and continued until the stationary phase at about 50 h, that is, production took place in a typical idiophase. At the beginning of the production phase, the oxygen uptake and CO_2 production rates (both about 100 mg L^{-1} h^{-1}) slowed down, and the respiratory quotient, RQ, started to rise, up to 0.8 during the production phase. After 50 h, 3.1 g/L wet biomass was reached. The cell yield was 0.31 g dry biomass/g peptone, or 0.5 g dry biomass/g oxygen. Total oxygen consumption over 55 h was 6.25 g oxygen/L. Production of myxovirescin was very low, the final yield being 0.04 mg/L. By feeding peptone at different rates to a peptone-limited culture (0.16% peptone), it was demonstrated that the peptone consumption rate closely correlated with the myxovirescin production rate. In the batch culture, this rate reached a maximum of 0.4 g peptone g dry weight^{-1} h^{-1} after 8 h, but myxovirescin synthesis started only much later, after 26 h, when the peptone consumption rate was down at about 0.12 g peptone g dry weight^{-1} h^{-1}.

The feeding experiments revealed a very sharp optimum of the specific peptone uptake rate for maximal myxovirescin production. Under the chosen experimental conditions the optimal rate was about 1.8 g peptone g dry weight^{-1} h^{-1} and resulted in a specific myxovirescin production rate of about 0.05 mg myxovirescin g dry weight^{-1} h^{-1}. For optimal production, the specific peptone uptake rate had to be maintained between 0.15 and 0.2 g peptone g dry weight^{-1} h^{-1}. The final yield of myxovirescin still was very low, 2 mg/L. In another experiment under chemostat conditions, the myxovirescin production rate was raised from 0.12 to 0.55 mg L^{-1} h^{-1} under constant biomass by passing the culture supernatant through a vortex chamber containing XAD resin and thus removing myxovirescin continuously from the culture (70). Cell dry mass and antibiotic concentration reached a steady state after 170 h. The latter was reduced from 0.95 to 0.1 mg myxovirescin/L after the vortex chamber had been connected to the bioreactor. The effect on myxovirescin production of removal of NH_3 from the culture has already been mentioned. Incidently, in chemostat experiments myxobacteria turned out to respond very slowly to changed conditions, perhaps because they have a high homeostasis. For the same antibiotic, a scale-up procedure in a 0.5% Casitone (Difco) medium is also reported, although with a different producer, *Myxococcus xanthus* strain TA (71). The best titers were obtained at a relatively low oxygen absorption rate, which in fact was too low to be maintained in the largest bioreactor used. No direct correlation between growth and antibiotic production was seen. Final biomass yield was 3–4 g wet weight/L.

Several studies have reported immobilization of myxobacterial cells. The objectives were to retain a relatively high biomass, to perform chemostat experiments (not possible with a strain growing in flakes and nodules), and to investigate the influence of immobilization on the production rates of enzymes (72–74) and secondary metabolites (38,75). A more theoretical study considered the role of hydrophobic and electrostatic properties of bacterial cell surfaces on adsorbtion to porous SiO_2 (glass) and Al_2O_3 particles (76). This study included *Sorangium cellulosum*, which showed a relatively low zeta potential (-33 mV at pH 7 and in the presence of 1 mM NaCl) comparable to that of *E. coli*, but a very high hydrophobicity in the range typical for gram-positive rather than gram-negative bacteria. The contact angle, the measure of hydrophobicity, increased from 52° at the beginning of the growth curve to 75° at its end. Both values—low zeta potential and high contact angle—are contrary to good adsorption.

Immobilized cells of *Sorangium cellulosum* strain So ce12 were used to study the limitations of sorangicin synthesis (38). As immobilizing agent, calcium alginate (Pro-

tanal LF 20/60, Carroux) was found most suitable. Beads 1 mm in diameter containing 5% biomass (v/w) were prepared and used in a 1.6-L bioreactor with 1 L medium in continuous culture under chemostat conditions. Phosphate (0.23 mM) was used as the limiting factor preventing growing out of the enclosed cells. Higher phosphate concentrations also repressed sorangicin production. Consequently peptone from casein could only be added to such a concentration that its phosphate content did not exceed the optimal phosphate concentration; if that was the case, peptone stimulated sorangicin production. With the applied medium, the optimum dilution rate was 0.016/h, the optimum pH 7.8, and the optimum p_{O_2} 80% saturation (with free cells, 40% were optimal). For sorangicin production, the cells within the beads had to be maintained under growth conditions. Their generation time was 19 h (14 h with free cells). Productivity began to decline after 50 days and came to an end after 75 days when the cells within the beads degenerated. An increase of the biocatalyst wet weight beyond 60 g/L, or a liquid to catalyst ratio of 10:1 (v/v), resulted in a decrease of growth and production, partly because of oxygen limitation. Productivity could be stimulated by the addition of glycerol (9 mM) and pyruvate (5 mM), while the direct precursors, acetate and propionate, were inhibitory. A shift to enzyme activities of the pentose phosphate pathway in the immobilized cells compared with free cells was seen. The maximum volumetric productivity obtained with immobilized cells was 0.8 to 5 mg sorangicin $L^{-1}\,h^{-1}$ and the specific production rate was 0.25 mg sorangicin g biomass^{-1} h^{-1}, compared with 0.17 mg $L^{-1}\,h^{-1}$ and 0.12 mg g dry biomass^{-1} h^{-1}, respectively, with free cells.

In another project, cells of *Cystobacter* and *Corallococcus* strains were immobilized by either adsorbing them to nylon pads or by enclosing them in κ-carrageenan beads (75). The preparations were tested in shake flasks for production of chemically unknown antibiotics. Yields with adsorbed cells were 20–30% and those with enclosed cells 45–49% higher than with free cells which, however, also grew in small pellets. A substantial increase in the excretion of proteolytic enzymes has been reported for *Myxococcus xanthus* cells immobilized in calcium alginate beads (72,73,74). Yield improvements as high as eightfold have been found.

CONCLUSION

Myxobacteria may be at the threshold of biotechnological application as producers of secondary metabolites and of enzymes. No large-scale technical process has as yet been worked out for any one of them, but it already is obvious that they do not behave fundamentally different from other bacteria used in those fields. Solutions of problems met there may well be applicable or adaptable to myxobacteria.

ACKNOWLEDGMENT

I thank my secretary, Ms. Karin Rahn, for patiently and diligently typing the manuscript. I gratefully acknowledge financial support by the Fond der Chemischen Industrie.

BIBLIOGRAPHY

1. J.R. Carias, J. Raingeaud, C. Mazaud, G. Vachon, N. Lucas, Y. Cenatiempo, and R. Julien, *FEBS Lett.* **262**, 97–100 (1990).
2. E. Petit and J.F. Guespin-Michel, *J. Bacteriol.* **174**, 5136–5140 (1992).
3. H. Irschik and H. Reichenbach, *Arch. Microbiol.* **142**, 40–44 (1985).
4. D.W. Morris and J.H. Parish, *Arch. Microbiol.* **108**, 227–230 (1976).
5. H. Mayer and H. Reichenbach, *J. Bacteriol.* **136**, 708–713 (1978).
6. S. Inouye, M.Y. Hsu, S. Eagle, and M. Inouye, *Cell* **56**, 709–717 (1989).
7. S. Inouye, P.J. Herzer, and M. Inouye, *Proc. Natl. Acad. Sci. USA* **87**, 952–945 (1990).
8. W. Zhang, J. Munoz-Dorado, M. Inouye, and S. Inouye, *J. Bacteriol.* **174**, 5450–5453 (1992).
9. A.M. Breton and J.F. Guespin-Michel, *FEMS Microbiol. Lett.* **40**, 183–188 (1987).
10. A.M. Breton, G. Younes, F. van Gijsegem, and J. Guespin-Michel, *J. Biotechnol.* **4**, 303–311 (1986).
11. A. El Akoum, M. Vijayalakshmi, P. Cardon, B. Fournet, M. Sigot, and J.F. Guespin-Michel, *Enzyme Microb. Technol.* **9**, 426–429 (1987).
12. P.J. Masson and J.F. Guespin-Michel, *J. Gen. Microbiol.* **134**, 801–806 (1988).
13. D.R. Nelson, M.G. Cumsky, and D.R. Zusman, *J. Biol. Chem.* **256**, 12589–12595 (1981).
14. M.G. Cumsky and D.R. Zusman, *J. Biol. Chem.* **256**, 12596–12599 (1981).
15. H. Reichenbach and G. Höfle, *Biotechnol. Adv.* **11**, 219–277 (1993).
16. H. Reichenbach and G. Höfle, in M. Dworkin and D. Kaiser eds., *Myxobacteria II*, American Soc. for Microbiology, Washington, D.C., 1993, pp. 347–397.
17. G. Höfle and H. Reichenbach, in W. Kuhn and H.P. Fiedler, eds., *Sekundärmetabolismus bei Mikroorganismen, Beiträge zur Forschung*, Atempto Verlag, Tübingen, 1995, pp. 61–78.
18. H. Reichenbach and G. Höfle, in S. Grabley and R. Thiericke eds., *Drug Discovery from Nature*, Springer-Verlag, Berlin, 1999, pp. 149–179.
19. H. Reichenbach and G. Höfle, in *Scientific Annual Report of the GBF, Gesellschaft für Biotechnologische Forschung*, Braunschweig, 1994, pp. 5–22.
20. H. Reichenbach, in *Scientific Annual Report of the GBF, Gesellschaft für Biotechnologische Forschung*, Braunschweig, 1996, pp. 25–30.
21. M. Dworkin and D. Kaiser, eds., *Myxobacteria II*, American Soc. for Microbiology, Washington, D.C., 1993.
22. M. Dworkin, *Microbiol. Rev.* **60**, 70–102 (1996).
23. H.D. McCurdy, E.R. Brockman, H. Reichenbach, and D. White, in J.T. Staley, M.P. Briant, N. Pfennig, and J.G. Holt eds., *Bergey's Manual of Systematic Bacteriology, vol. 3*, Williams & Wilkins, Baltimore, 1989, pp. 2139–2170.
24. H. Reichenbach and M. Dworkin, in A. Balows, H.G. Trüper, M. Dworkin, W. Harder, and K.H. Schleifer eds., *The Prokaryotes*, 2nd ed., *vol. IV*, Springer-Verlag, New York, 1992, pp. 3416–3487.
25. H. Reichenbach, in J.G. Holt, N.R. Krieg, P.H.A. Sneath, J.T. Staley, and S.T. Williams eds., *Bergey's Manual of Determi-*

native Bacteriology, 9th ed., Williams & Wilkins, Baltimore, 1994, pp. 515–525.

26. C.A. Jacobi, H. Reichenbach, B.J. Tindall, and E. Stackebrandt, *Int. J. Syst. Bacteriol.* **46**, 119–122 (1996).
27. C.A. Jacobi, B. Assmuss, H. Reichenbach, and E. Stackebrandt, *Appl. Environ. Microbiol.* **63**, 719–723 (1997).
28. B. Kunze, R. Jansen, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **47**, 881–886 (1994).
29. B. Zeggel, Doctoral Thesis, TU Braunschweig, 1993.
30. J.R. Lampky and E.R. Brockman, *Int. J. Syst. Bacteriol.* **27**, 161 (1977).
31. W. Trowitzsch-Kienast, E. Forche, V. Wray, H. Reichenbach, E. Jurkiewicz, G. Hunsmann, and G. Höfle, *Liebigs Ann. Chem.* **1992**, 659–664 (1992).
32. L. Shimkets and C. Woese, *Proc. Natl. Acad. Sci. USA* **89**, 9459–9463 (1992).
33. H.W. Chen, A. Kuspa, I.M. Keseler, and L.J. Shimkets, *J. Bacteriol.* **173**, 2109–2115 (1991).
34. B. Neumann, A. Pospiech, and H.U. Schairer, *Mol. Microbiol.* **10**, 1087–1099 (1993).
35. M. Kalos and J. Zissler, *Proc. Natl. Acad. Sci. USA* **87**, 8316–8320 (1990).
36. S. Jaoua, S. Neff, and T. Schupp, *Plasmid* **28**, 157–165 (1992).
37. T. Schupp, C. Toupet, B. Cluzel, S. Neff, S. Hill, J.J. Beck, and J.M. Liggon, *J. Bacteriol.* **177**, 3673–3679 (1995).
38. K. Becker, Doctoral Thesis, TU Braunschweig, 1990.
39. B. Kunze, T. Kemmer, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **37**, 454–461 (1984).
40. S.M. Ringel, R.C. Greenough, S. Roemer, D. Connor, A.L. Gutt, B. Blair, G. Kanter, and M. von Strandtmann, *J. Antibiot.* (*Tokyo*) **30**, 371–375 (1977).
41. P. Coucke and J.P. Voets, *Z. Allg. Mikrobiol.* **7**, 175–182 (1967).
42. B.F. Watson and M. Dworkin, *J. Bacteriol.* **96**, 1465–1473 (1968).
43. K. Gerth and H. Reichenbach, *Arch. Microbiol.* **117**, 173–182 (1978).
44. K. Gerth and H. Reichenbach, *Anal. Biochem.* **152**, 78–82 (1986).
45. K. Gerth, W. Trowitzsch, G. Piehl, R. Schultze, and J. Lehmann, *Appl. Microbiol. Biotechnol.* **19**, 23–28 (1984).
46. B. Kunze, R. Jansen, F. Sasse, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **48**, 1262–1266 (1995).
47. K. Gerth, N. Bedorf, H. Irschik, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **47**, 23–31 (1994).
48. F. Sasse, H. Steinmetz, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **46**, 741–748 (1993).
49. B. Kunze, W. Kohl, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **38**, 1649–1654 (1985).
50. F. Sasse, H. Steinmetz, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **48**, 21–25 (1995).
51. B. Kunze, R. Jansen, L. Pridzun, E. Jurkiewicz, G. Hunsmann, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **46**, 1752–1755 (1993).
52. B. Kunze, R. Jansen, L. Pridzun, E. Jurkiewicz, G. Hunsmann, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **45**, 1549–1552 (1992).
53. H. Irschik, H. Augustiniak, K. Gerth, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **48**, 787–792 (1995).
54. K. Gerth, N. Bedorf, G. Höfle, H. Irschik, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **49**, 560–563 (1996).
55. H. Irschik, D. Schummer, K. Gerth, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **48**, 26–30 (1995).
56. K. Gerth, P. Washausen, G. Höfle, H. Irschik, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **49**, 71–74 (1996).
57. B. Kunze, W. Trowitzsch-Kienast, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **45**, 147–150 (1992).
58. K. Gerth, R. Jansen, G. Reifenstahl, G. Höfle, H. Irschik, B. Kunze, H. Reichenbach, and G. Thierbach, *J. Antibiot.* (*Tokyo*) **36**, 1150–1156 (1983).
59. H. Irschik, W. Trowitzsch-Kienast, K. Gerth, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **41**, 993–998 (1988).
60. K. Grimm, *J. Bacteriol.* **135**, 748–753 (1978).
61. K. Grimm, *Z. Allg. Mikrobiol.* **18**, 399–407 (1978).
62. E. Rosenberg, K.H. Keller, and M. Dworkin, *J. Bacteriol.* **129**, 770–777 (1977).
63. H. Irschik and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **38**, 1237–1245 (1985).
64. H. Irschik, K. Gerth, T. Kemmer, H. Steinmetz, and H. Reichenbach, *J. Antibiot.* (*Tokyo*), **36**, 6–12 (1983).
65. S. Fytlovitch, P.D. Nathan, D. Zafriri, and E. Rosenberg, *J. Antibiot.* (*Tokyo*) **36**, 1525–1530 (1983).
66. B. Kunze, N. Bedorf, W. Kohl, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **42**, 14–17 (1989).
67. H. Irschik, R. Jansen, K. Gerth, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **48**, 31–35 (1995).
68. H. Irschik, R. Jansen, K. Gerth, G. Höfle, and H. Reichenbach, *J. Antibiot.* (*Tokyo*) **40**, 7–13 (1987).
69. J.N. Nigam, J. Lehmann, K. Gerth, H. Piehl, R. Schultze, and W. Trowitzsch, *Appl. Microbiol. Biotechnol.* **19**, 157–160 (1984).
70. V. Hecht, J. Vorlop, H. Kalbitz, K. Gerth, and J. Lehmann, *Biotechnol. Bioeng.* **29**, 222–227 (1987).
71. A. Mizrahi, J. Arnan, G. Miller, Z. Liron, M. Manai, Y. Batus, and E. Rosenberg, *J. Appl. Chem. Biotechnol.* **26**, 160–166 (1976).
72. G. Younes, A.M. Breton, and J. Guespin-Michel, *Appl. Microbiol. Biotechnol.* **25**, 507–512 (1987).
73. G. Younes, J.M. Nicaud, and J. Guespin-Michel, *Appl. Microbiol. Biotechnol.* **19**, 67–69 (1984).
74. J.C. Vuillemard, S. Terré, S. Benoit, and J. Amiot. *Appl. Microbiol. Biotechnol.* **27**, 423–431 (1988).
75. F.N. Yasouri and H.A. Foster, *Biotechnol. Tech.* **5**, 149–152 (1991).
76. C. Krekeler, H. Ziehr, and J. Klein, *Appl. Microbiol. Biotechnol.* **35**, 484–490 (1991).

NITRILE HYDRATASE

YOSHIRO ASHINA
MASARU SUTO
TAKAKAZU ENDO
Nitto Chemical Industry Co., Ltd.
Tokyo, Japan

KEY WORDS

Acrylamide
Bioreactor
Flocculant
Inducer
Metalloenzyme
Nicotinamide
Nitrilase
Nitrile hydratase
Rhodococcus chlororaphis B-23
Rhodococcus rhodochrous J-1
Rhodococcus sp. N-774

OUTLINE

MANUFACTURE OF ACRYLAMIDE

Manufacture of acrylamide is the most important commercial application of nitrile hydratase. Acrylamide is used as a starting material for a variety of polymer products such as flocculant, paper-strength agent, coating materials, and EOR (enhanced oil recovery) polymers. Current world production of acrylamide exceeds 200,000 metric tons per annum (MTA). Nitto Chemical Industry Co., Ltd. (Japan) started commercial production of acrylamide in 1986 using a catalyst based on nitrile hydratase. This was one of the first commercial bioconversion processes that made a commodity chemical. The very first acrylamide process based on hydration of acrylonitrile with sulfuric acid (Fig. 1) was commercialized in 1950, which was replaced by the copper-catalyst process 20 years later. This copper-catalyzed process currently prevails in worldwide acrylamide production. Although the copper process gives an excellent efficiency in comparison with the acid hydration process, it still has some drawbacks, such as formation of impurities that are difficult to remove and adversely influence polymerization reactions. These impurities include ethylene cyanohydrin, β-hydroxypropionamide, and nitrilotrispropionamide, which may interfere with the chain reactions lowering the molecular weight of polymers and deteriorate the stability of the monomer product. Overcoming these limitations was one of the important motivations for the development of the enzymatic process using nitrile hydratase, which would minimize the side reactions caused by high specificity of enzymes and mild conditions under which they work.

Strain N-774

Because microorganisms capable of converting nitrile compounds to corresponding amides had been reported in the 1960s (1) many studies were conducted to find a microorganism with an activity high enough for industrial applications. *Rhodococcus* sp. N-774, discovered in 1980 (2), was the first strain that was put into commercial use to produce acrylamide from acrylonitrile. The nitrile hydratase enzyme is produced in this strain without an inducer. Thiamine is essential for the cell growth. Fe^{2+} and Fe^{3+} remarkably promote production of nitrile hydratase in the cell (3). Table 1 and Figure 2 show the effect of metal ions on the activity and the typical time course of the cell growth and the enzyme activity, respectively. The electron micrograph of the strain N-774 is shown in Figure 3.

Reaction Conditions

According to the semi-batchwise reaction test using immobilized cells of N-774 with a serial addition of acrylonitrile, a reaction temperature of 0 °C is optimum to obtain the product in high concentration, as shown in Figure 4 (4). Enzyme activity of this cell is light-inductive (5). Figure 5 shows the effect of light irradiation on the activity, which should be properly reflected on the reactor design.

Immobilization of the Cell

For mass production of biologically converted products, immobilized cells give wider options in selection and designing of the reactor. The bacteria cells entrapped in a polyacrylamide-based hydrogel are used in the production of acrylamide. To maintain the enzyme activity and improve the physical characteristics of the immobilized cells, the following techniques are used: (*1*) Swelling of the gel is controlled by copolymerization with cationic monomers (6); (*2*) Shearing force applied while mixing the cells with monomers or polymers is carefully controlled; (*3*) polymerization conditions are adjusted to minimize diffusion resistance within the gel, which not only reduces the apparent

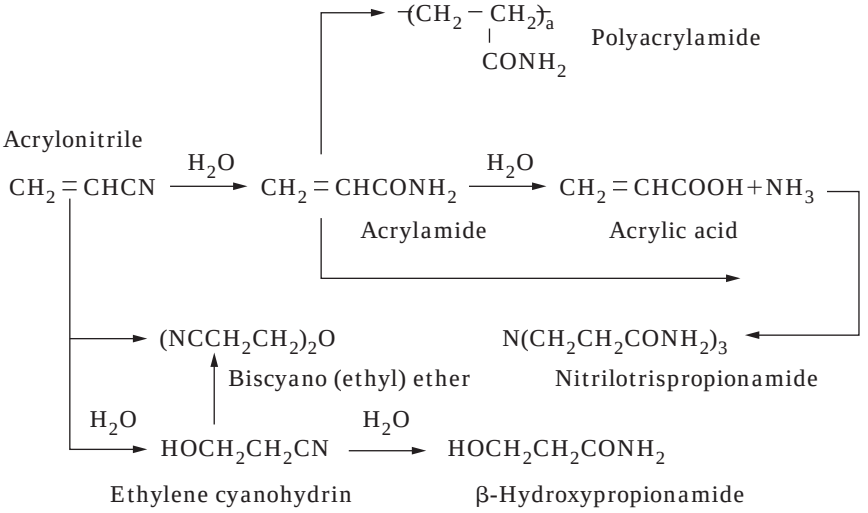


Figure 1. Products of acrylonitrile hydration.

Table 1. Effect of Metal Ions on Activity

Inorganic salts	(mg/L)	Call growth (g/L)	pH (−)	Specific activity (units/mg-dry cell)
$FeSO_4$ $7H_2O$	10	4.30	7.58	50.0
$MnSO_4$ $4H_2O$	10	4.75	6.94	0.70
$NiSO_4$ $6H_2O$	10	4.42	7.53	0.92
$CoSO_4$ $7H_2O$	10	4.34	7.37	1.27
$BaCl_2$ $2H_2O$	10	4.63	7.42	1.16
Na_2MoO_4 $2H_2O$	10	4.05	7.60	0.70
$CuSO_4$ $5H_2O$	2	4.54	7.30	1.24
$ZnSO_4$ $7H_2O$	2	4.38	7.45	0.96
None	—	4.42	7.47	1.05

Note: Basal medium: 10 g glucose, 3 g $(NH_4)_2SO_4$, 0.5 g KH_2PO_4, 0.5 g K_2HPO_4, 0.5 g $MgSO_4$ $7H_2O$, 0.1 g NaCl, 0.1 g $CaCl_2$ $2H_2O$, 5 g casamino acid and 0.002 g thiamine HCl in 1 L of distilled water, pH 7.5. Culture: at 30 °C for 55 hrs by shaking.

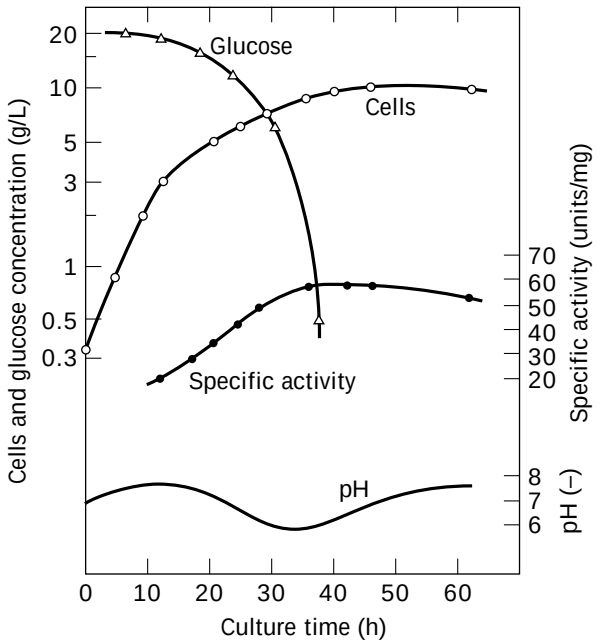


Figure 2. Time courses of cell growth and enzyme production. Culture conditions: medium consisting of 20 g glucose, 10 g peptone, 6 g yeast extract, and 6 g malt extract in 1 L of distilled water, pH, 7.2; aeration, 0.75 v/v/m; temperature, 30 °C; stirring, 350 rpm; culture volume, 13 L in a 20-L jar fermenter.

activity but also degenerates the activity by increasing the amide concentration near the active site (Figs. 6 and 7); and (*4*) Physical stability of the gel is improved by adjusting the polymer composition and polymerization conditions to prevent the cells from escaping the gel. An electron micrograph of the entrapped state of the cell is shown in Figure 8. The immobilization is carried out on a continuous basis at the commercial plant.

Bioreactor

Various types of the reactors—fixed bed, moving bed, fluidized bed, and stirred type—can be used (Membrane type does not seem suitable because of its flow-through resistance). The reactor is designed in consideration of (*1*) inhibition of polymerization; (*2*) replacement of degenerated catalyst; (*3*) control of reaction temperature; (*4*) control of substrate concentration; and (5) irradiation of light.

Comparison of the Processes

In comparison with the copper-catalyst process the biological acrylamide process features (*1*) 100% one-pass conversion; (*2*) no metal ions; (*3*) minimized organic impurities; (*4*) mild reaction conditions; and (*5*) a simplified process. The block diagram of the two processes is shown in Figure 9. Table 2 compares the typical product analysis.

Strain B-23

Pseudomonas chlororaphis B-23, discovered by Yamada et al. of Kyoto University (*7*) (Table 3), potentially gives higher activity than N-774. The tolerance to the product is also higher than that of N-774. Sucrose, glutamic acid, cystine, and proline are used as medium components to grow the enzyme with minimized catabolite repression (8).

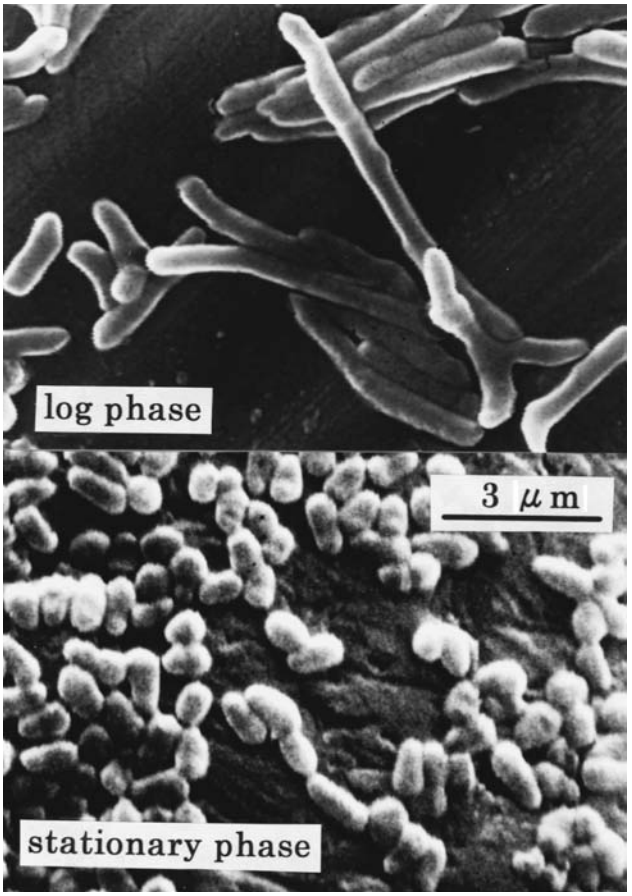


Figure 3. Electron micrograph of strain N-774.

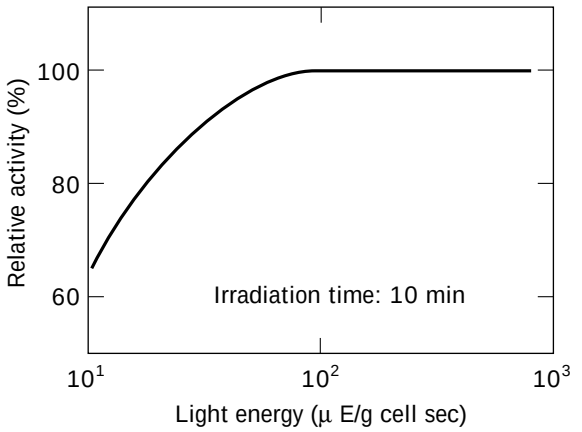


Figure 5. Activation of enzyme by light energy.

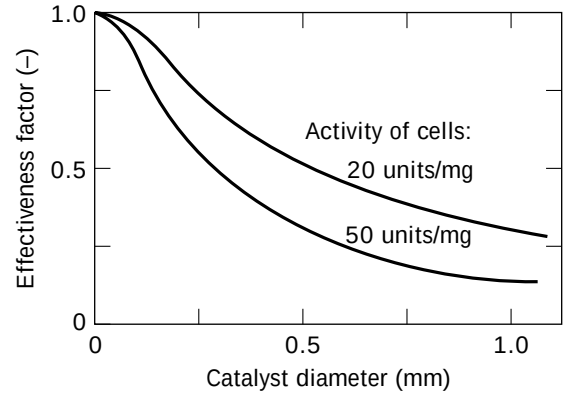


Figure 6. Catalyst particle size and effectiveness factor.

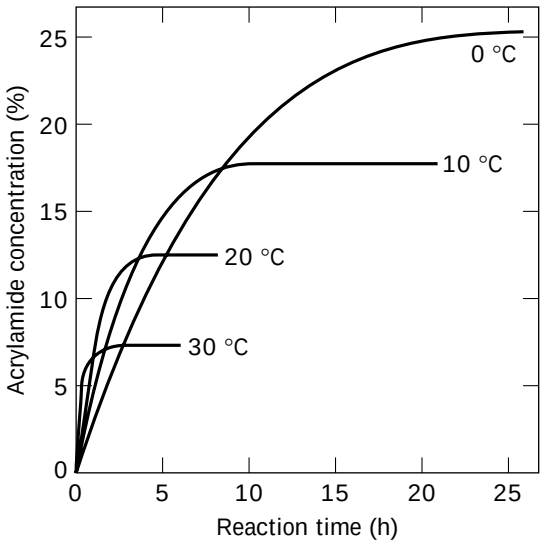


Figure 4. Effect of temperature on acrylamide production. The reaction was carried out in a semi-batchwise fashion in a stirred-tank reactor. Acrylonitrile was added consecutively. Reaction conditions: pH, 8.5; catalyst concentration, 0.5%; acrylonitrile concentration, 2.0%; temperature, as indicated in the figure.

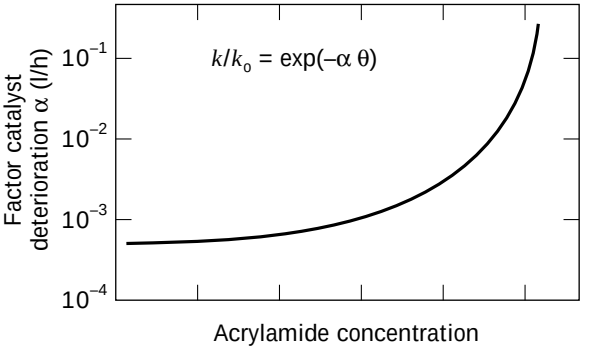


Figure 7. Catalyst deterioration by amide without reaction.

Methacrylamide works remarkably well as an enzyme inducer for this strain. The activity is improved by means of mutation using chemicals such as nitrosoguanidine. The mutant Am-324, shown in Figure 10, gives the activity 3,000 times as high as the original cell from nature. The catalyst based on this mutant, when used in the commercial acrylamide plant, increased the production capacity to 6,000 MTA from 4,000 MTA with the N-774 catalyst.

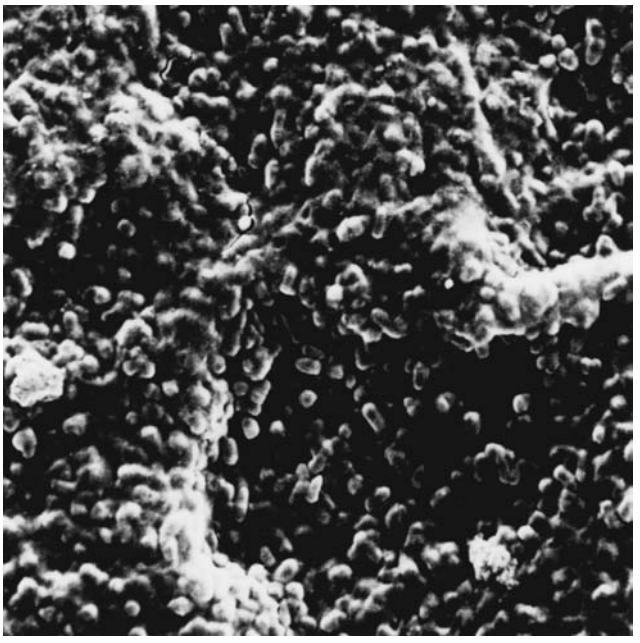


Figure 8. Electron micrograph of immobilized cells.

Table 2. Typical Product Analysis

Impurities	Acrylamide from enzymatic process	Acrylamide from copper-catalyzed process
Acrylonitrile	N.D. (<5)	100 ~ 200
Acrylic acid	<5.0	100 ~ 200
Ethylene cyanohydrin	N.D. (<10)	300 ~ 400
β-Hydroxypropionamide	N.D. (<10)	600 ~ 1000
Nitrilotrispropionamide	N.D. (<10)	~100
Propionamide	N.D. (<10)	20 ~ 50

Note: ppm in 50% aq. solution.

Strain J-1

The enzyme induction mechanism of *Rhodococcus rhodochrous* J-1 was also found by Yamada et al. (9). At the beginning of their discovery, J-1 produced two different kinds of nitrile-converting enzymes (nitrilase and nitrile hydratase), and the reaction activity of both enzymes was very low. They found addition of cobalt ions into the cell-cultured medium improved the nitrile hydratase activity, and further increase of the activity was affected by action of urea or its derivatives as an inducer in the presence of cobalt ions. This strain has very high tolerance to acrylamide. Table 4 shows properties of the nitrile hydratase from the strains N-774, B-23, and J-1. All three enzymes belong to the metalloenzyme group but are different in kind of the coordinated metals to their active sites (i.e., iron for N-774 and B-23 and cobalt for J-1). Nitrile hydratase enzymes having nickel as the coordinated metal have also been found. The protein chemistry of nitrile hydratase, including molecular weight and light inducibility, differs widely among different genuses and species. Reaction performance of the three strains, in laboratory tests and at the commercial plant, are compared in Figure 11 and

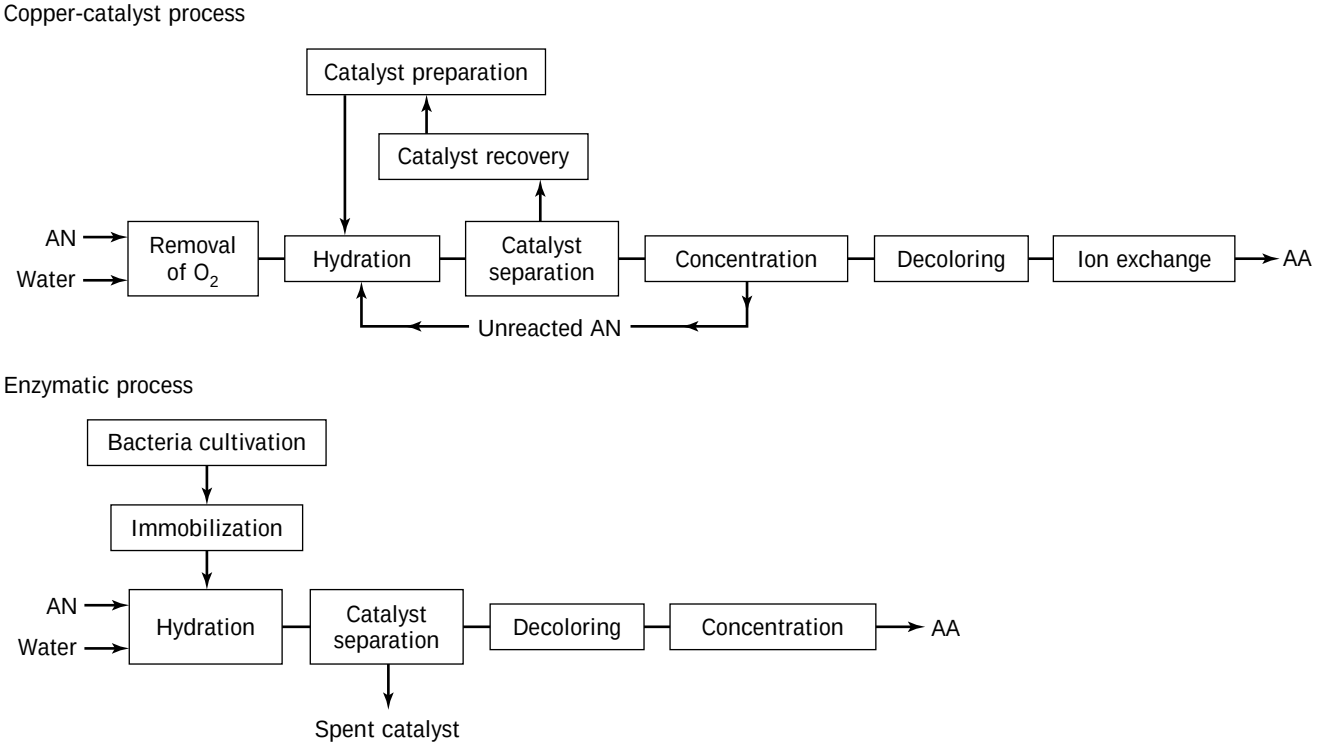


Figure 9. Comparison of enzymatic process with conventional process. AN = acrylonitrile; AA = Acrylamide (50% aqueous solution).

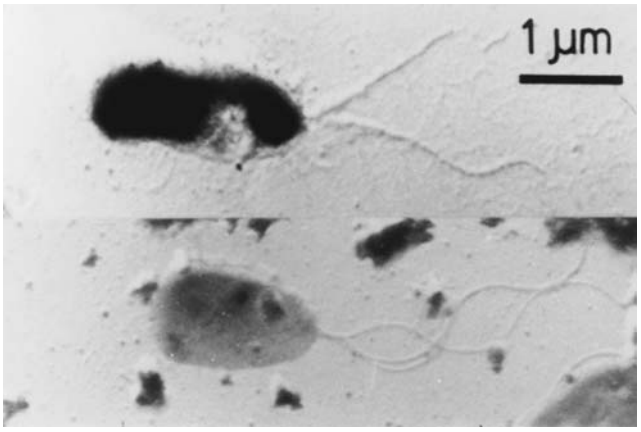


Figure 10. Electron micrograph of strain B-23.

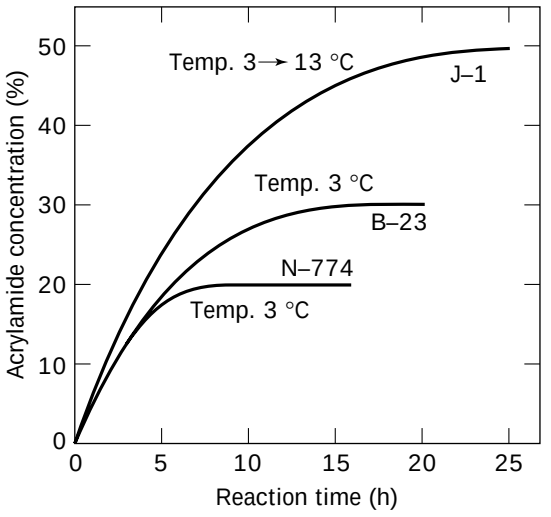


Figure 11. Comparison of reaction performance. The reaction was carried out in a semi-batchwise fashion in a stirred-tank reactor. Acrylonitrile was added consecutively. Reaction conditions: pH, 6.5 ~ 8.5; temperature, as indicated in the figure; acrylonitrile concentration, 2.0%.

Table 5. The optimized J-1 exhibits a magnitude of order higher productivity than the N-774. Production capacity of the acrylamide plant using the J-1 catalyst is by four to five times greater than that of the same plant with the N-774 catalyst.

APPLICATION TO OTHER AMIDES

Another important industrial application of nitrile hydratase is the production of nicotinamide, which is widely used as a vitamin supplement. At the end of 1997, Lonza Ltd. (China) was scheduled to manufacture 3,000 MTA of nicotinamide by the continuous biotransformation of 3-cyanopyridine using immobilized cells of *Rhodococcus rhodochrous*. Considerable quantity of nicotinic acid is inevitably formed as a by-product in alkaline hydrolysis of 3-cyanopyridine. Under the biological route, nicotinic acid and other contaminating by-products are not produced. This route is also competitive with the methylethylpyridine route.

NITRILASE

Nitrile hydratase is one of the two types of the enzymes that catalyze hydration or hydrolysis of nitrile compounds. The other one, nitrilase, converts nitriles to corresponding carboxylic acids. Some of the nitrilases are capable of hydrolyzing nitriles stereospecifically. Optical resolution of α-substituted nitriles using a nitrilase to obtain one of the corresponding acid enantiomers was reported (10). The yield of the proposed process, however, was limited to 50% with the other nitrile enantiomer remaining intact. The biotransformation process to produce *R*-mandelic acid using a *R*-specific nitrilase (equation 1) has been commercialized by Nitto Chemical (11). In this process, the unconverted *S*-nitrile enantiomer is racemized in one-pot equilibrium dissociation-association reactions (equation 2), which eventually gives nearly 100% conversion to *R*-acid enantiomer at an optical purity over 99%. Addition of sodium sulfate in the reaction mixture enhances the productivity sixfold by reversibly forming the benzaldehyde-sulfite complex to promote racemization of *S*-mandellonitrile (equation 3) (12).

$$\text{RS-}C_6H_5CH\ (OH)\ CN \rightarrow \text{S-}C_6H_5CH\ (OH)\ CN + \text{R-}C_6H_5CH\ (OH)\ COOH + NH_3 \quad (1)$$

$$C_6H_5CH\ (OH)\ CN \leftrightarrow C_6H_5CHO + HCN \quad (2)$$

$$C_6H_5CHO + Na_2SO_3 + H_2O \leftrightarrow C_6H_5CH\ (OH)\ SO_3Na + NaOH \quad (3)$$

Although R-specific nitrilase is found in a variety of bacteria, such as *Pseudomonas* sp., *Alcaligenes* sp., *Acineto-*

Table 3. The Development of B-23 Strain Am-324

Strain	Mutation	Medium	SA (U/mg)	BA (U/mL)	Property
Wild	—	M-A	0.72	0.4	Mucilaginous bacterium
↓	—	M-R	66	363	
Mutant I (Am-3)	NTG	M-R	65	465	Unmucilaginous
↓					
Mutant II (Am-324)	NTG	M-R	125	952	High activity
↓	—	Consecutive addition of inducer	141	1260	

Note: NTG = *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine; SA = specific activity, BA = enzyme activity of culture broth.

Table 4. Properties of Nitrile Hydratase from N-774, B-23, and J-1

	Rhodococcus sp N-774	*Pseudomonas chlororaphis* B-23	*Rhodococcus rhodochrous* J-1
Molecular weight	70,000	100,000	520,000
Subunit molecular weight	α: 27,000	25,000	α: 26,000
	β: 27,500		β: 29,000
Metal	Fe	Fe	Co
Activation by light	Positive	Negative	Negative
Tolerance for acrylamide	High	Very high	Very high
Production of organic acid	Very low	Negligible	Low
Optimum temperature [°C]	35	20	
Heat stability	Unstable	Unstable	Stable
Optimum pH	7.7	7.5	
Stable pH	7.0 ~ 8.5	6.0 ~ 7.5	6.0 ~ 8.5
Enzyme formation	Constitutive	Inducible	Inducible

Table 5. Reaction Performance in Commercial Production

	Reaction conditions		Reaction performance	
	pH	7.5 ~ 8.5	Conversion of AN	99.9 %
N-774	Temperature	0 ~ 5 °C	Selectivity of AA	99.9 %
	AN concentration	1.5 ~ 2.0 %	AA concentration at outlet of reactor	20%
	pH	7.5 ~ 8.5	Conversion of AN	99.97 %
B-23	Temperature	0 ~ 5 °C	Selectivity of AA	99.98 %
	AN concentration	1.5 ~ 2.0 %	AA concentration at outlet of reactor	30 %
	pH	6.5 ~ 8.0	Conversion of AN	99.97 %
J-1	Temperature	0 ~ 15 °C;	Selectivity of AA	99.98 %
	AN concentration	1.5 ~ 2.0 %	AA concentration at outlet of reactor	50 %

Note: AN = acrylonitrile, AA = acrylamide.

bacter sp., *Caseobacter* sp., *Nocardia* sp., *Brevibacterium* sp., *Bacillus* sp., *Aureobacterium* sp., and *Rhodococcus* sp. (13), *S*-specific nitrilase has hardly been isolated.

BIBLIOGRAPHY

1. A. Miura et al., *J. Ferment. Technol.* **47**, 631 (1969).
2. Jpn. Pat. 17918 (1981), I. Watanabe et al. (to Nitto Chemical Industry Co., Ltd.).
3. Jpn. Pat. 4987 (1984), I. Watanabe et al. (to Nitto Chemical Industry Co., Ltd.).
4. Jpn. Pat. 38118 (1981), I. Watanabe et al. (to Nitto Chemical Industry Co., Ltd.).
5. Jpn. Pat. 162195 (1986), K. Enomoto et al. (to Nitto Chemical Industry Co., Ltd., Mitsubishi Rayon).
6. Jpn. Pat. 35078 (1983), I. Watanabe et al. (to Nitto Chemical Industry Co., Ltd.).
7. Jpn. Pat. 37951 (1984), H. Yamada et al. (to H. Yamada).
8. Jpn. Pat. 43996 (1986), H. Yamada et al. (to H. Yamada, Nitto Chemical Industry Co., Ltd.); Jpn. Pat. 43997 (1986), H. Yamada et al. (to H. Yamada, Nitto Chemical Industry Co., Ltd.); Jpn. Pat. 43998 (1986), H. Yamada et al. (to H. Yamada, Nitto Chemical Industry Co., Ltd.); Jpn. Pat. 43999 (1986), H. Yamada et al. (to H. Yamada, Nitto Chemical Industry Co., Ltd.).
9. Jpn. Pat. 55148 (1986), H. Yamada et al. (to H. Yamada).
10. Eur. Pat. 0,348,901 B, K. Yamamoto et al. (to Asahi Kasei Kogyo).
11. Eur. Pat. 0,610,048 A, Y. Hashimoto et al. (to Nitto Chemical Industry Co., Ltd.).
12. U.S. Pat. 5,326,702, T. Endo et al. (to Nitto Chemical Industry Co., Ltd.).
13. U.S. Pat. 5,223,416, T. Endo et al. (to Nitto Chemical Industry Co., Ltd.).

OILS, MICROBIAL PRODUCTION

SAKAYU SHIMIZU
JUN OGAWA
Kyoto University
Kyoto, Japan

KEY WORDS

Arachidonic acid
Docosahexaenoic acid
Eicosadienoic acid
Eicosapentaenoic acid
Fatty acid
Linoleic acid
Mortierella
Polyunsaturated fatty acid
Single cell oil
Triacylglycerol

OUTLINE

INTRODUCTION

As are microbial carbohydrates, microbial lipids are a diverse group of compounds with important biochemical functions. The exploitation of microbial oil (triacylglycerol) has been attempted since the 1980s. The term *single cell oil* is used for unique oils produced by microorganisms (1,2). It is probable that single cell oil could compete against plant seed oils or fish oils. Some yeasts and molds have long been known to accumulate very high levels of triacylglycerols. A lipid content in excess of 40% (w/w) is not exceptional, and values of 70% and even 80% have been reported (3). The oil produced by these oleaginous species not only has the same triacylglycerol structure as plant oils but also shows a similar distribution of fatty acids at the three positions: saturated fatty acids are almost totally excluded from the central (*sn*-2) carbon atom.

The fatty acyl groups themselves are the same as those found in plant seed oils or, in the highly polyunsaturated fatty acids, are the same as those found in fish oils or other animal products. Many attempts have been made to produce useful fatty acids or oils containing them by microbial processes. One example is the production of cocoa-butter-equivalent oil by a yeast that abundantly produces stearic acid (4,5). The most successful example occurs in the production of oils containing polyunsaturated fatty acids (PUFAs). Certain microorganisms may also produce prostaglandin precursors, or even prostaglandins themselves from PUFAs, as well as cerebroside lipids and other unusual lipids that are not normally regarded as being of microbial origin (6). Here, microbial PUFA production is discussed mainly as a typical example of a single cell oil.

OUTLINE OF MICROBIAL PUFA PRODUCTION

PUFAs are fatty acids having more than one double bond. Among these, 5(*Z*),8(*Z*),11(*Z*)-eicosatrienoic acid (20:3ω9; Mead acid), dihomo-γ-linolenic acid (20:3ω6; DGLA), arachidonic acid (20:4ω6; AA), and 5(*Z*),8(*Z*),11(*Z*),14(*Z*),17(*Z*)-eicosapentaenoic acid (20:5ω3; EPA) (see Fig. 1 for structures and biosynthetic pathways) are of interest because they are precursors for prostaglandins, thromboxanes, leukotrienes, prostacyclins, and so on (7–9), all of which have hormonelike activities. Possibly as a result of the physiological activities of these compounds, PUFAs exhibit several unique biological activities, such as lowering of plasma cholesterol level and prevention of thrombosis. The inclusion of EPA and 4(*Z*),7(*Z*),10(*Z*),13(*Z*),16(*Z*),19(*Z*)-docosahexaenoic acid (22:6ω3; DHA) as dietary supplements has been recommended for the prevention of heart diseases (10), and γ-linolenic acid (18:3ω6; GLA) has been recommended for the relief of eczema (11). Accordingly, PUFAs are highly important substances in the pharmaceutical, medical, and nutritional fields. Because food sources rich in these PUFAs are limited to a few seed oils that contain GLA and fish oil, which contains EPA and DHA, recent investigations have focused on microorganisms as alternative sources of oils containing these PUFAs.

Microorganisms are potentially very promising lipid sources because of their extremely high growth rates in simple media and the simplicity of their manipulation.

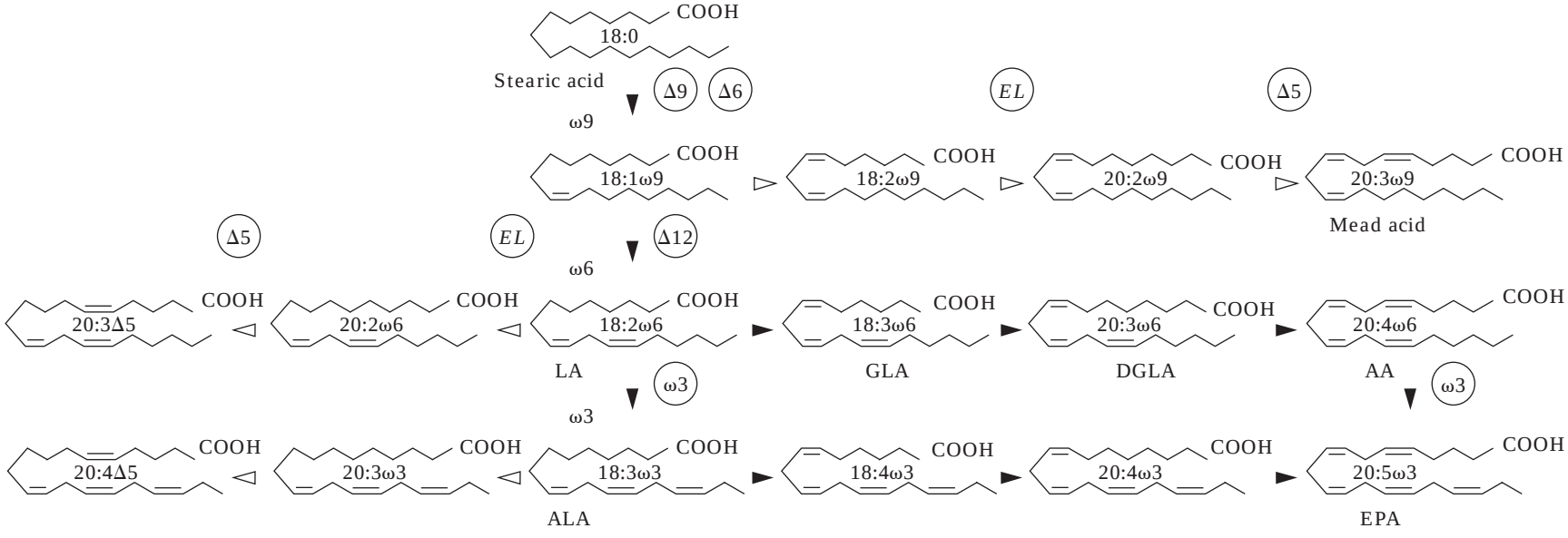


Figure 1. Proposed pathways for polyunsaturated fatty acid (PUFA) biosynthesis in *Mortierella alpina* 1S-4. *Open arrowheads* show bypaths through which *n*-9 fatty acids and nonmethylene-interrupted PUFAs are formed in Mut48 and Mut49, respectively. **Δ***n*, *Δ*n desaturation; *EL*, elongation.

Furthermore, microbial PUFA-containing oils are usually characterized by the absence of other PUFAs, making purification of individual fatty acids from microorganisms an easier task than it is from other sources. A variety of PUFAs have been detected in microorganisms including bacteria, fungi, algae, mosses, and protozoa (Table 1). Erwin and Bloch (12) suggested that lower classes of organisms including microorganisms can be classified into several groups based on their ability to produce PUFAs. Shaw (13) pointed out that some fungi belonging to the Mucorales can accumulate relatively large amounts of GLA in their mycelia. Based on these early observations, several groups have been screening for microorganisms capable of accumulating lipids containing PUFAs so as to obtain more suitable sources for large-scale preparation of important nutritional components.

Suzuki (14,15) found several potent producers of GLA through extensive screening of filamentous fungi. They reported that a strain of *Mortierella isabellina* accumulates about 3 g/L of GLA when grown in a medium containing glucose as the major carbon source. Based on this finding, commercial production of GLA with a *Mortierella* fungus recently commenced in Japan.

Shimizu et al. (16,17) found that several *Mortierella* fungi are potent producers for triacylglycerol containing AA, DGLA, or EPA. A soil isolate strain, *Mortierella alpina* 1S-4, produced AA on cultivation in a medium containing glucose and yeast extract. It also produced EPA when grown at low temperature (20 °C) or when grown in a medium supplemented with α-linolenic acid (18:3ω3; ALA). When grown in the presence of sesame seed oil, the same fungus accumulated DGLA in place of AA because of specific inhibition of Δ5 desaturase by sesamin and related lignan compounds present in the oil. Mutants that were considered to be defective (or to have low activity) in Δ5, Δ6, Δ12, Δ9, and ω3 desaturase were derived from *M. alpina* 1S-4 (Fig. 2) (18,19). Mutant Mut44 and S14 had low Δ5-desaturase activity and accumulated large amounts of DGLA. Mutant Mut48 was completely defective in Δ12 desaturase and produced three *n*-9 PUFAs: 6(*Z*),9(*Z*)-

Table 1. Microbial Polyunsaturated Fatty Acids

Polyenoic acid[a]	Major source	Typical source	Content (%) in total FA
14:2	Green algae	*Chrorella pyrenoidosa*	tr[b]
14:3	Green algae	*Chrorella* sp.	tr
15:2(7,10)	Algae	*Euglena gracilis*	1
15:4(4,7,10,13)	Algae	*Euglena gracilis*	1
16:2(5,9)	Slime molds	*Dictyostelium* sp.	10
16:2(7,10)	Green algae	*Chrorella yulgaria*	3
16:2(9,12)	Diatoms	*Bidulphia sinensis*	28
16:3(4,7,10)	Green algae	*Scenedesmus* sp.	tr[b]
16:3(6,9,12)	Algae (diatoms, etc.)	*Prorocentrum* sp.	12
16:3(7,10,13)	Green algae, *Euglena*, etc.	*Chrorella pyrenoidosa*	8
16:4(4,7,10,13)	Green algae, *Euglena*, etc.	*Chrorella pyrenoidosa*	3
16:4(6,9,12,15)	Diatoms	*Bidulphia sinensis*	9
17:2	Yeasts	*Candida tropicalis*	5
17:2(5,9)	Slime molds	*Dictyostelium* sp.	tr[b]
18:2(5,9)	Slime molds	*Dictyostelium* sp.	tr[b]
18:2(5,11)	Slime molds	*Dictyostelium* sp.	41
18:2(9,12)LA	Many microorganisms	*Agricus bisporus*	86
18:3(6,9,12)GLA	Filamentous fungi, green algae, blue-green algae, red algae, ciliates, etc.	*Mucor genevenis*	32
18:3(9,12,15)ALA	Yeasts, filamentous fungi, basidiomycetes, green algae, brown algae, etc.	*Chrorella* sp.	41
18:4(6,9,12,15)	Green algae, brown algae	*Scenedesmus* sp.	36
19:2(9,12)	Algae	*Euglena gracilis*	tr
19:2(11,14)	Algae	*Euglena gracilis*	tr
19:4(5,8,11,14)	Algae	*Euglena gracilis*	6
20:2(11,14)	Algae (Chrysophyceae, *Euglena*)	*Achanthamoeba* sp.	7
20:3(8,11,14)DGLA	Algae (Chrysophyceae, Xanthophyceae, red algae, *Euglena*, flagellates, etc.)	*Ocheromonas* sp.	5
20:3(11,14,17)	Algae (flagellates, etc.)	*Leishmania* sp.	2
20:4(5,8,11,14)AA	Filamentous fungi, many algae, protozoa	*Conidiobolus* sp.	41
20:4(8,11,14,17)	Algae (Crytomphyceae, brown algae, flagellates, *Euglena*)	*Cryptomonas ovata*	6
20:5(5,8,11,14,17)EPA	Filamentous fungi, marine bacteria, algae	*Monodus* sp.	31
22:4(7,10,13,16)	Algae (Crytomphyceae), ciliates	*Cryptomonas ovata*	2
22:5(4,7,10,13,16)	Algae (Haptophyceae), ciliates	*Critthidia* sp.	14
22:5(7,10,13,16,19)	Algae (Dinophyceae, Chrysophyceae)	*Amiphidinium* sp.	25
22:6(4,7,10,13,16,19)DHA	Algae (Dinophyceae, Haaptomphyceae)	*Exuviella* sp.	16

[a]All fatty acids are *cis*-form. Numbers in parentheses indicate double-bond position.
[b]tr, trace.

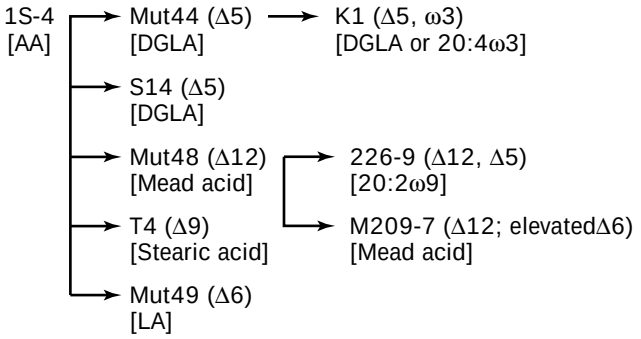


Figure 2. Mutants derived from *M. alpina* 1S-4. Desaturases missing are shown in *parentheses*; major fatty acids produced are shown in *square brackets*.

octadecadienoic acid (18:2ω9), 8(*Z*),11(*Z*)-eicosadienoic acid (20:2ω9), and Mead acid. Two nonmethylene-interrupted PUFAs, that is, 5(*Z*),11(*Z*),14(*Z*)-eicosatrienoic acid (20:3Δ5) and 5(*Z*),11(*Z*),14(*Z*),17(*Z*)-eicosatetraenoic acid (20:4Δ5), and accumulation of linoleic acid (18:2ω6; LA) were found in the Δ6 desaturase-defective mutant Mut 49. Two other mutants, T4 and K1, were considered to have defects in Δ9 and ω3 desaturase, respectively. T4 accumulated a high level of stearic acid (18:0) and K1 did not produce *n*-3 fatty acids, which were usually produced by the parental strain at low temperature. Above all, *M. alpina* 1S-4 and its mutants are new and promising sources of C_{20} PUFAs. In addition, because of intensive investigations done worldwide, all PUFAs listed in Table 2 can now be obtained from a variety of microorganisms.

PRODUCTION OF C_{18} FATTY ACIDS

γ-Linolenic Acid

Suzuki (14,15) reported that several *Mortierella* strains belonging to the subgenus *Micromucor* produce lipids, mainly triacylglycerol rich in GLA, in their mycelia when grown in a liquid medium containing glucose or sugar cane molasses as the major carbon source. These strains grow rapidly in a medium containing a high concentration of glucose (200 g/L) under the conditions of 2 ppm dissolved oxygen, pH 4.0, and 30 °C, yielding a mycelial mass of 40–156 g/L. The lipid content of GLA was 18–41 mg/g dry mycelia, corresponding to a GLA yield of 1.2–3.4 g/L. The purified oil obtained from the mycelia contained myristic acid (0.7%, by weight), palmitic acid (27.2), palmitoleic acid (0.9), stearic acid (5.7), oleic acid (43.9), linoleic acid (12.0), GLA (8.3), arachidic acid (0.6), eicocenoic acid (0.4), behenic acid (0.1), and erucic acid (0.2). The GLA content of the oil was comparable to that of *Oenothera biennis* seed oil. The low linoleic acid content of the oil may be advantageous for obtaining a highly pure GLA preparation with a high recovery. Recently, *Mortierella ramanniana*, *Mucor rouxii*, and *Mucor ambiguus* were evaluated as to their productivities of GLA. *M. ambiguus* immobilized in porous support particles was shown to excrete lipids containing GLA into the culture broth or onto the surface of the cell wall in the presence of a nonionic surfactant.

Linoleic Acid

The mutant Mut49 derived from *M. alpina* 1S-4 was considered to be defective in Δ6 desaturase. This mutant accumulated high amounts of LA with low levels of GLA, DGLA, and AA in its mycelia (20) (Table 3).

Stearic Acid

Mutant T4 derived from *M. alpina* 1S-4 is considered to be defective in Δ9 desaturation (19). Its mycelial fatty acids included 38% stearic acid (18:0), the level being only 5% for the wild type (Table 3). When grown at 24–28 °C, its mycelial lipids included a markedly high level (up to 50 mol% of total mycelial lipids) of free fatty acids, of which about 90 mol% was stearic acid. However, the levels of free fatty acids were markedly decreased with a concomitant increase in the triacylglycerol level when the mutant was grown at 20 °C or lower. When mutant T4 was grown in a culture medium supplemented with LA or linseed oil that included about 60% ALA, decreased levels of free fatty acid and increased levels of triacylglycerol were also found, suggesting that PUFA is needed for the synthesis of triacylglycerol in this mutant.

PRODUCTION OF C_{20} FATTY ACIDS

Arachidonic Acid

Shimizu et al. have assayed the C_{20} PUFA productivity of a wide variety of microorganisms (16,17). Most C_{20} PUFA producers were found to be filamentous fungi belonging to the orders Mucorales and Entomophthorales. Mainly they produce C_{20} PUFAs of the *n*-6 family (i.e., AA and DGLA) together with C_{18} PUFAs of the same family (i.e., GLA). Most of the PUFA-producing isolates from natural sources were found to belong to the genus *Mortierella*. The genus *Mortierella* is subdivided into two subgenera, *Micromucor* and *Mortierella*. Interestingly, all the strains found as C_{20} PUFA producers belong to the subgenus *Mortierella*. Neither the stock cultures nor isolates belonging to the subgenus *Micromucor* showed any detectable accumulation of C_{20} PUFA, although they were good producers of GLA (51). As a result of this screening effort, a soil isolate was obtained that was taxonomically identified as *Mortierella alpina* 1S-4, as a potent producer of AA (16,17). The fungus effectively utilized not only glucose but glycerol, maltose, palmitate, stearate, oleate, *n*-hexadecane, and *n*-octadecane as carbon sources for AA production (Fig. 3).

Using a 10-m^3 fermentor, a mycelial mass of 20 kg/m^3 of medium (dry wt) containing 60% lipid, and 40% of AA in the total fatty acids, was produced by *M. alpina* 1S-4. This fermentation involved intermittent feeding of glucose and 7 days cultivation at 28 °C. The concentration of AA in the harvested mycelia increased to nearly 70% when allowed to stand at room temperature for a further 6 days. The lipids containing AA could be recovered as a triacylglycerol (80–90% recovery) from the mycelia using a seven-step process: *(1)* separation of the mycelia by filtration, *(2)* drying, *(3)* crushing by ball mill, *(4)* extraction of lipids with *n*-hexane, *(5)* removal of insoluble materials by centrifugation, *(6)* decolorization and deodorization with ac-

Table 2. Sources of Polyunsaturated Fatty Acids (PUFAs) and Their Occurrence in Microorganisms

Fatty acid	Representative sources from: Microorganisms	Others
	n-6 series	
18:2ω6	*Mortierella alpina* 1S-4 Mut49 (20)	Soybean, sunflower, evening primrose, borage, black currant
GLA	*Mucor circinelloides* (3), *Mo. isabellina* (14), *Mo. ramanniana* var. *anglulispora* (21) *Cunninghamella echinucalta* (22), *Mu. mucedo* (23), *Cunninghamella elegans* (24), *Rhizopus arrhizus* (25), *Rhizopus* spp. (26) *Thamnidium elegans* (24)	
DGLA	*Mo. alpina* 1S-4 (27), *Mo. alpina* 1S-4 Mut44 (28), *Mo. alpina* 1S-4 S14 (29) *Saprolegnia ferax* (30)	Human milk, animal organs, mosses
AA	*Mo. alpina* 1S-4 and other *Mortierella* subgenus (16,17) *Entomophthora exitalis* (32), *Blastocladiella emersonii* (30) *Conidiobulos nanodes*	Algae (31), porcine liver, moss (33)
DPA(n-6)	*Schyzochytrium* sp. (34)	
	n-3 series	
ALA	Most microorganisms	Linseed, perilla
20:4ω3	*Mo. alpina* 1S-4 S14 (35)	
EPA	*Mortierella* sp. (36), *Pythium* sp. (37), *Shewanella putrefaciens* (38), *Mo. alpina* 1S-4 Mut48 (39) *Phythium irregulare (40)*	Fishes, algae (41)
DHA	*Thraustochytrium aureum* (42) *Schyzochytrium* sp. (34)	Fishes, algae (43)
DPA(n-3)	*Thraustochytrium aureum* (32)	
	n-9 series	
18:2ω9	*Mo. alpina* 1S-4 Mut48 (44)	Animal organs
20:2ω9	*Mo. alpina* 1S-4 Mut48 (44) *Mo. alpina* 1S-4 M226-9 (unpublished observation)	Animal organs
Mead acid	*Mo. alpina* 1S-4 Mut48 (44) *Mo. alpina* 1S-4 M209-7 (45)	Animal organs
	Others	
19:4ω5	*Mortierella* sp. (46)	Mullet
19:5ω2	*Saprolegnia* sp. (47)	Mullet
20:5ω1	*Mo. alpina* 1S-4 (48)	None
20:3Δ5	*Mo. alpina* 1S-4 Mut49 (18)	Plant seeds (49)
20:4Δ5	*Mo. alpina* 1S-4 Mut49 (18)	Plant seeds (50)

AA, arachidonic acid (20:4ω6); ALA, α-linolenic acid (18:3ω3); DGLA, dihomo-γ-linolenic acid (20:3ω6); DHA, 4(*Z*), 7(*Z*), 10(*Z*), 13(*Z*), 16(*Z*), 19(*Z*)-docosahexaenoic acid (22:6ω3); DPA(*n*-3), 7(*Z*), 10(*Z*), 13(*Z*), 16(*Z*, 19(*Z*)-docosapentaenoic acid (22:5ω3); DPA(*n*-6), 4(*Z*), 7(*Z*), 10(*Z*), 13(*Z*), 16(*Z*)-docosapentaenoic acid (22:5ω6); EPA, 5(*Z*), 8(*Z*), 11(*Z*), 14(*Z*), 17(*Z*)-eicosapentaenoic acid (20:5ω3); GLA, γ-linolenic acid (18:3ω6); LA, linoleic acid (18:2ω6); Mead acid, 5(*Z*), 8(*Z*), 11(*Z*)-eicosatrienoic acid (20:3ω9); *Mo., Mortierella; Mu., Mucor*.

tive charcoal, and *(7)* concentration (Fig. 4). If necessary, the AA in the purified oil could be isolated as the methyl or ethyl ester after successive transesterification, liquid–liquid partition chromatography, and high-performance liquid chromatography.

Dihomo-γ-Linolenic Acid

Most AA-producing fungi can accumulate small amounts of DGLA when they are cultivated under the conditions optimized for AA production (52). Because AA is derived from DGLA through the *n*-6 route, all AA-producing fungi potentially should produce large amounts of DGLA if one can block its conversion to AA (Δ5 desaturation).

The first approach to blocking the Δ5 desaturation began with screening for Δ5-desaturase inhibitors. Among various substances tested, sesame oil and peanut oil were found to cause a marked decrease in AA content (27,53) when added to the culture medium. On addition of 3% sesame seed oil to a glucose–yeast extract medium, the production of DGLA with *M. alpina* 1S-4 reached 1.7 g/L, whereas the production of AA was only 0.7 g/L. The effective factors responsible for this phenomenon were isolated from sesame seeds extracts or sesame oil and identified to be lignan compounds, that is, (+)-sesamin, (+)-episesamin, (+)-sesaminol, and (+)-episesaminol (Fig. 5) (54). The results obtained from experiments using either cell-free extracts of *M. alpina* 1S-4 or a rat liver microsomal fraction clearly demonstrated that these lignan compounds specifically inhibited Δ5 desaturase at low concentrations.

Table 3. Comparison of Fatty Acid Compositions of the Fatty Acid Desaturation-Defective Mutants of *Mortierella alpina* 1S-4

	Growth temperature (°C)								
	28						12		
Fatty acid (wt%)	1S-4	Mut48	Mut49	Mut44	S14	T4	Mut49	Mut44	K1
16:0	11.1	9.9	12.7	10.8	16.9	8.2	5.8	10.4	10.7
Stearic acid	5.3	6.1	10.7	7.8	7.5	38.0	10.2	9.8	6.5
18:1ω9	13.7	49.3	15.3	11.1	10.7	5.2	8.7	13.7	21.9
LA	7.4	—[a]	31.8	5.6	2.9	3.9	25.5	4.3	5.5
18:2ω9	—	12.1	—	—	—	—	—	—	—
GLA/20:0[b]	4.6	0.7	1.7	9.6	6.7	11.9	2.3	10.0	11.6
ALA/20:1ω9[c]	1.1	2.7	1.0	1.1	0.6	—	2.7	1.0	1.2
18:4ω3	—	—	—	—	—	—	—	1.3	—
20:2ω6	0.6	—	2.1	1.0	0.5	—	6.1	0.6	0.8
20:2ω9	—	2.1	—	—	—	—	—	—	—
DGLA/22:0[b]	6.5	2.2	2.9	32.8	48.8	1.1	1.7	30.8	32.7
Mead acid	—	9.3	—	—	—	—	—	—	—
20:3Δ5	—	—	7.8	—	—	—	6.4	—	—
AA/20:3ω3[c]	44.4	—	6.8	11.7	0.9	31.7	22.7	7.1	8.1
20:4ω3	—	—	—	—	—	—	—	6.4	—
20:4Δ5	—	—	—	—	—	—	3.2	—	—
24:0	5.2	5.6	7.8	8.5	4.2	—	3.3	1.5	1.0
EPA	—	—	—	—	—	—	1.5	3.1	—

Note: All strains were grown in medium of 2% (w/v) glucose and 1% (w/v) yeast extract (pH 6.0) at either 28 °C for 7 days (28 °C) or at 12 °C for 6 days after preincubation for 1 day at 28 °C (12 °C), except that K1 was grown at 12 °C for 10 days after 1-day cultivation at 28 °C (12 °C).
AA, arachidonic acid (20:4ω6); ALA, α-linolenic acid (18:3ω3); DGLA, dihomo-γ-linolenic acid (20:3ω6); EPA, 5(*Z*), 8(*Z*), 11(*Z*), 14(*Z*), 17(*Z*)-eicosapentaenoic acid (20:5ω3); GLA, γ-linolenic acid (18:3ω6); LA, linoleic acid (18:2ω6); Mead acid, 5(*Z*), 8(*Z*), 11(*Z*)-eicosatrienoic acid (20:3ω9); stearic acid, 18:0.
[a]—, undetectable.
[b]Both were found in all strains except for Mut48, in which GLA and DGLA were not detected. 20:0, arachidic acid; 22:0, behenic acid.
[c]ALA and 11(*Z*), 14(*Z*), 17(*Z*)-eicosatrienoic acid (20:3ω3) were found as an additional fatty acid on growth at 12 °C except for Mut48.

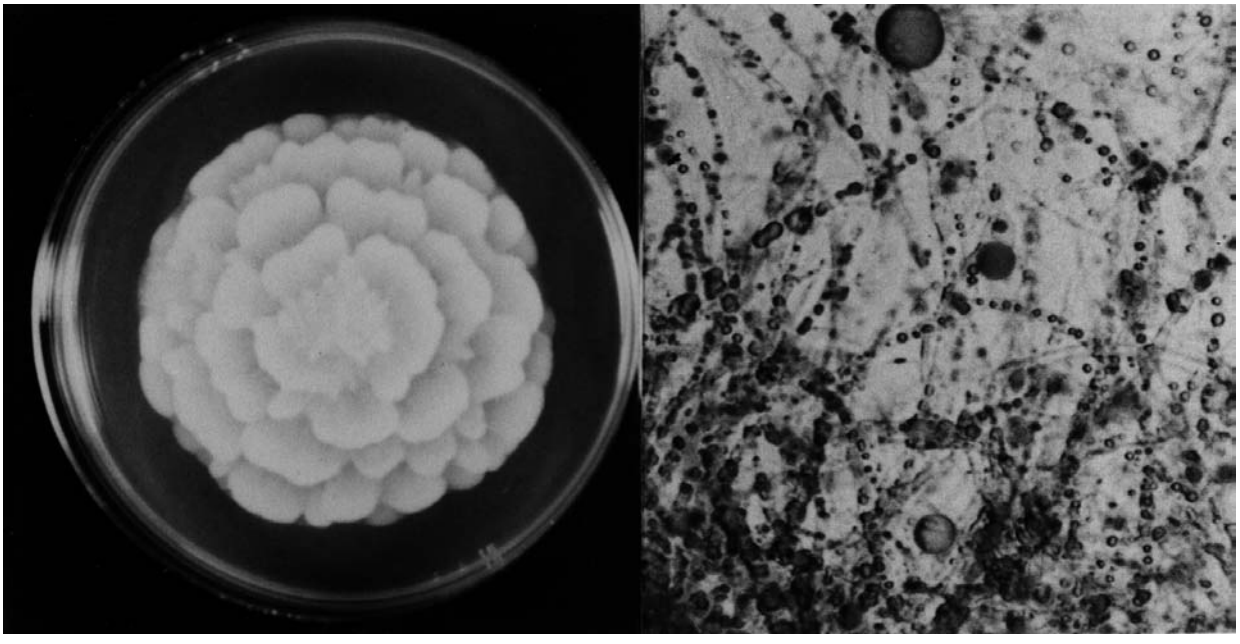


Figure 3. Photomicrographs of *M. alpina* 1S-4 colony grown on potato-dextrose agar (*left*) and oil droplets accumulated in mycelia (*right*).

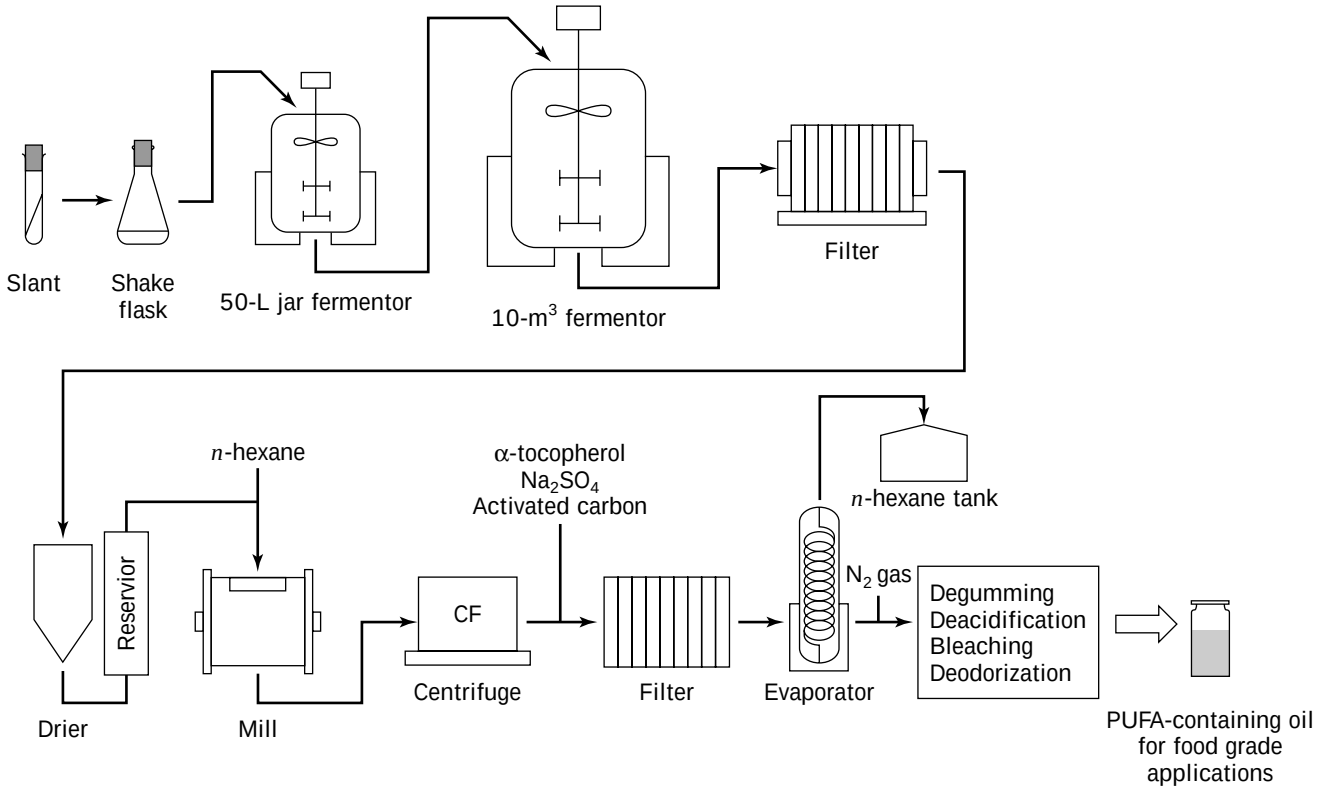


Figure 4. Flowchart for the production of a fungal single cell oil with a high PUFA content.

At these concentrations, none of the Δ6, Δ9, and Δ12 desaturases were inhibited by these lignan compounds. Kinetic analysis showed that sesamin is a noncompetitive inhibitor (K_i = 155 μM for rat liver Δ5 desaturase) (54). (−)-Asarinin and (−)-epiasarinin, the stereoisomers of (+)-episesamin and (+)-sesamin, respectively, which were isolated from a Chinese crude medicine "Saishin" (*Asiasari radix*), also showed specific noncompetitive inhibition of Δ5 desaturase (55). Another desaturase inhibitor, curcumin, was found in turmeric. Curcumin inhibited both Δ5 and Δ6 desaturase in vitro, but it was not as effective as sesamin in stimulation of DGLA production (56,57). Recently, an antioxidant, alkyl gallate (Fig. 5), and the Ca channel blockers nicardipine and nifedipine (Fig. 5) were also shown to inhibit both Δ5 and Δ6 desaturases (58). In a study on optimization of the culture conditions for the production of DGLA by *M. alpina* 1S-4, a medium containing glucose, yeast extract, and the non-oil fraction of sesame oil was found to be suitable. Under optimal conditions in a 50-L fermentor, the fungus produced DGLA at 2.2 g/L (107 mg/g dry mycelia) (Table 4). This value accounted for 23.1% of the total mycelial fatty acids.

The second approach to blocking the Δ5 desaturation began with screening for Δ5-desaturase-defective mutants. Several mutants of this type, which is characterized by a high DGLA level and a reduced AA level, were obtained, and two (Mut44 and S14) of them were studied in detail (28,29). The mycelial fatty acid composition of the two mutants is compared in Table 3. The mutant S14 has a higher DGLA level than Mut44; its AA level is about one-tenth that in Mut44. On cultivation for 6 days at 28 °C in a 10-L fermentor, Mut44 produced 3.2 g DGLA per liter of culture broth (123 mg/g dry mycelia), which accounted for 23.4% of the total mycelial fatty acids (28). The mycelial AA amounted to only 19 mg/g dry mycelia (0.5 g/L culture broth), which accounted for 3.7% of the total mycelial fatty acids (Table 4). The level of AA in Mut44 could be further decreased by growing the fungus in the presence of a mixture of Δ5-desaturase inhibitors, sesamin and episesamin. Using a culture medium supplemented with 0.075 g/L sesamin mixture, the production of DGLA was 2.1 g/L culture broth (16.1% of total fatty acids) and the AA production was 0.2 g/L (1.3%) (59). The DGLA production of S14 was 2.4 g/L culture broth (43.3% of total fatty acids) when grown at 28 °C for 7 days in a 5-L jar fermentor (29). The production of DGLA by using mutants was superior to the aforementioned method of using inhibitors because it did not require inhibitors and the concentration of DGLA in the mycelial oil was higher (Table 4).

By using mutant K1 derived from strain Mut44, which was considered to be defective in the conversion of *n*-6 to *n*-3 fatty acids (ω3 desaturation) (60), a DGLA-rich oil without *n*-3 fatty acids was obtained (Table 3). The mutant could not produce 8(*Z*),11(*Z*),14(*Z*), 17(*Z*)-eicosatetraenoic acid (20:4ω3) or any other *n*-3 fatty acids, of which about 10% was found in its parental strain on cultivation at 12 °C. The growth rate of the mutant was comparable to that of the parental strain when grown at 28 °C, but it became much slower when the mutant was grown at 12 °C, at

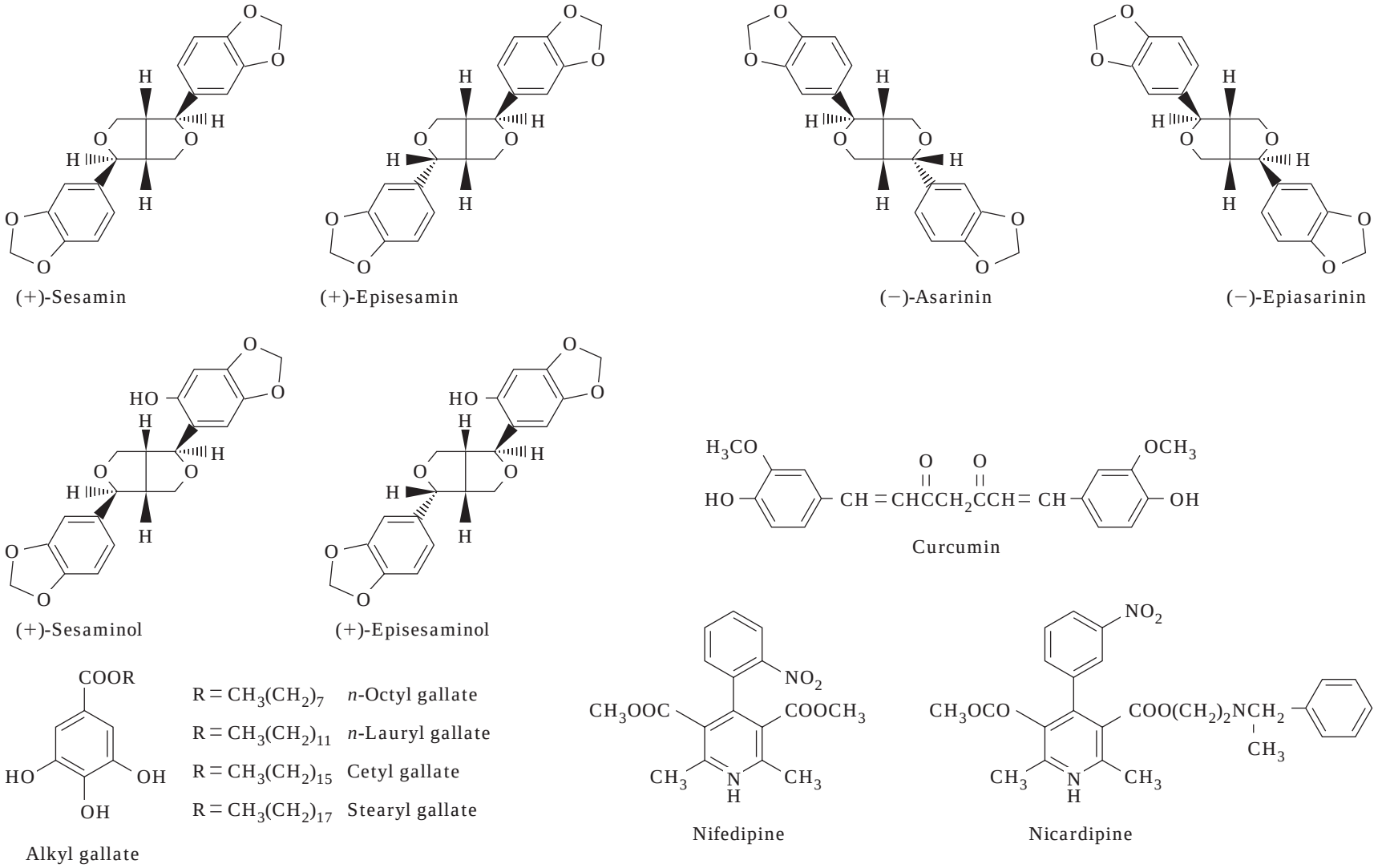


Figure 5. Sesamin and related compounds as desaturase inhibitor.

Table 4. Comparison of DGLA Productivities and Mycelial Fatty Acid Profiles of Some *Mortierella* Strains under Different Culture Conditions

Strain and conditions	Mycelial mass (g/L)	DGLA content (mg/g dry mycelia)	DGLA yield (g/L)	Mycelial fatty acid composition (wt%)			
				LA	GLA	DGLA	AA
			M. alpina 1S-4				
+ Sesame oil	20.3	107	2.2	6.6	4.1	23.1	11.2
			Mutant Mut44				
− Inhibitors[a]	26.0	123	3.2	8.9	6.5	23.4	3.7
+ inhibitors[a]	28.4	74	2.1	10.0	6.8	16.1	1.3
			Mutant S14				
− inhibitors[a]	17.1	140	2.4	4.4	5.8	43.3	1.5

AA, arachidonic acid (20:4ω6); DGLA, dihomo-γ-linolenic acid (20:3ω6); GLA, γ-linolenic acid (18:3ω6); LA, linoleic acid (18:2ω6).
[a]Inhibitors, a mixture of sesamin and episesamin.

which point the lag phase for Mut44 was about 2 days but 5 days for the mutant.

Eicosapentaenoic Acid

Shimizu et al. (36,52) found that lowering the cultivation temperature caused the additional accumulation of EPA by all the AA producers tested. Most of these AA producers grew well at low temperature (6–16 °C) and produced enough mycelia in simple growth media. Experiments with cell-free extracts of *M. alpina* 1S-4 demonstrated that the enzyme that catalyzed the formation of EPA was produced even when the fungus was grown at high temperature, but that the reaction(s) yielding EPA did not take place at high temperature (52). The Δ17 (or ω3) desaturation of AA to EPA was observed in *Saprolegnia parasitica* by Gellerman and Schlenk (61). If this is also the case for *Mortierella* fungi, an enzyme or enzyme system catalyzing the methyl end-directed desaturation of AA (Δ17 (or ω3) desaturation) may be activated by cold treatment and the resultant EPA may be necessary for maintaining the proper membrane fluidity in a low-temperature environment. As shown in Table 5, *M. alpina* 1S-4 produced EPA at levels of more than 0.3 g/L at 12 °C.

It was further demonstrated that several AA-producing *Mortierella* strains accumulated detectable amounts of EPA in their mycelia when grown in media containing ALA (62). This observation suggests that the conversion of ALA to EPA through the *n*-3 route occurs in *Mortierella* fungi as well as in animals. This conversion is independent of the growth temperature, because EPA production takes place even at 28 °C. The ability of the *Mortierella* fungi to convert added ALA to EPA is very promising from a biotechnological point of view because various kinds of easily available natural oils contain ALA. The potential of such natural oils as precursors of EPA was examined. Linseed oil, in which ALA amounts to about 60% of the total fatty acids, was found to be the most suitable for EPA production. Under optimal culture conditions, *M. alpina* 1S-4 converted 3.8% of the ALA in the added linseed oil into EPA, and EPA production reached 0.99 g/L (41 mg/g dry mycelia). This value is 3.3-fold higher than that obtained under low-temperature growth conditions (Table 5). Another advantage of this conversion is that it can be carried out under normal growth temperature conditions (20–30 °C). Under such conditions, the fungal growth is rapid and dense, and the energy costs should be lower than those for low-temperature cultivation.

EPA production was further stimulated when AA-producing fungi were grown in a medium containing linseed oil at low temperature. This phenomenon was suggested to be mainly caused by the low temperature-dependent production of EPA from AA formed through the *n*-6 route, and the conversion of the 18:3ω3 in the added linseed oil to EPA through the n-3 route took place at the same time (63). Stimulation of the *n*-3 route itself at low temperature was also suggested to contribute to the increased EPA production. The amount of EPA accumulated reached 1.88 g/L (66.6 mg/g dry mycelia) on cultivation of *M. alpina* 1S-4 with 3% linseed oil at 12 °C (Table 5).

By using the Δ12 desaturase-defective mutant Mut48 derived from *M. alpina* 1S-4, an EPA-rich oil with a low level of AA was obtained. Like the wild-type strain, Mut48 converted exogenous ALA to EPA (39). On cultivation at 20 °C for 10 days in a 5-L fermentor containing medium supplemented with linseed oil, the production of EPA was about 1 g/L culture broth (64 mg/g dry mycelia), accounting for approximately 20% of the total mycelial fatty acids. The mycelial AA content was 26 mg/g dry mycelia (0.4 g/L), accounting for 7.8% of the total mycelial fatty acids. As shown in Table 5, the EPA concentrations were lower than AA concentrations in the oils produced by all strains, except for Mut48 in which the concentration of EPA was 2.5-fold that of AA because of the Δ12 desaturase defect.

Recently, *Saprolegnia* sp. 28YTF-1, isolated from a freshwater sample, was shown to be a potent producer of EPA. Accompanying production of AA and other *n*-6 PUFAs was markedly low in this organism (64).

8(*Z*),11(*Z*),14(*Z*),17(*Z*)-Eicosatetraenoic Acid

Δ5 Desaturase-defective mutants of *M. alpina* 1S-4 accumulate 8(*Z*),11(*Z*),14(*Z*),17(*Z*)-eicosatetraenoic acid (20:4ω3) when grown with linseed oil. One of the mutants,

Table 5. Comparison of EPA Productivities and Mycelial Fatty Acid Profiles of Some *Mortierella* Strains under Different Culture Conditions

Strain and conditions	Mycelial mass (g/L)	EPA content (mg/g dry mycelia)	EPA yield (g/L)	Mycelial fatty acid composition (wt%)					
				GLA	ALA	DGLA	AA	EPA	Others
M. alpina 1S-4									
12 °C	11.1	27.0	0.30	4.5	—	2.9	63.8	10.9	17.9
28 °C + linseed oil	24.1	41.0	0.99	1.4	29.3	1.5	13.8	9.2	44.8
12 °C + linseed oil	28.2	66.6	1.88	1.6	28.2	3.0	15.7	12.0	39.5
M. alpina 2O-17									
12 °C	17.0	29.0	0.49	3.5	—	2.3	60.0	13.5	20.7
28 °C + linseed oil	32.4	41.5	1.35	0.9	38.5	1.4	12.3	7.1	39.8
M. hygrophila									
12 °C	15.2	20.4	0.31	3.5	—	1.6	7.2	6.7	81.0
Mutant Mut48									
20 °C + linseed oil	15.1	64.0	0.97	0.3	20.3	0.8	7.8	19.5	51.3

AA, arachidonic acid (20:4ω6); ALA, α-linolenic acid (18:3ω3); DGLA, dihomo-γ-linolenic acid (20:3ω6); EPA, 5(*Z*), 8(*Z*), 11(*Z*), 14(*Z*), 17(*Z*)-eicosapentaenoic acid (20:5ω3); GLA, γ-linolenic acid (18:3ω6).

M. alpina S-14, produced about 1.7 g/L (66 mg/g dry mycelia; 11.6 wt% of total mycelial fatty acids) of this PUFA per liter of culture medium (35).

Mead Acid

A Δ12 desaturase-defective mutant, Mut48, derived from *M. alpina* 1S-4 is characterized by a high oleic acid (18:1ω9) level and the absence of *n*-6 PUFA in its mycelia. In contrast to the wild type, this mutant produces three *n*-9 PUFAs, that is, 6(*Z*),9(*Z*)-octadecadienoic acid (18:2ω9), 8(*Z*),11(*Z*)-eicosadienoic acid (20:2ω9), and Mead acid (44). Significantly high levels of these fatty acids were produced on growth of the mutant at low temperature (12–20 °C). On cultivation at 20 °C for 10 days in a 5-L fermentor, the production of Mead acid reached approximately 0.8 g/L (56 mg/g dry mycelia), accounting for 15 wt% of the total mycelial fatty acids. The other major fatty acids were palmitic acid (6%), stearic acid (11%), oleic acid (45%), 18:2ω9 (12%), and 20:2ω9 (3%). About 70 mol% of Mead acid was present in the triacylglycerol and the remainder in the phospholipid fraction, especially in phosphatidylcholine. The formation of Mead acid could occur through the same sequential reactions suggested by Fulco and Mead (65), that is, desaturation 18:1ω9 at the Δ6 position into 18:2ω9 followed by elongation and Δ5 desaturation to Mead acid (see Fig. 1).

Recently, enhanced production of Mead acid was reported by a mutant fungus, *M. alpina* M209-7, derived from Δ12 desaturase-defective Mut48 strain. The production of Mead acid by M209-7 strain was 1.3 times greater than in Mut48, thought to result from its enhanced Δ6 desaturation activity, which was 1.4 times higher than that of Mut48 (45).

8(*Z*),11(*Z*)-Eicosadienoic Acid

A mutant strain, *M. alpina* M226-9, was derived from the Δ12 desaturase-defective mutant *M. alpina* Mut48. In the M226-9 strain, Δ5 and Δ12 desaturases were completely defective and the microorganism could produce large amounts of 8(*Z*),11(*Z*)-eicosadienoic acid (20:2ω9) (1.68 g/L; 101 mg/g dry mycelia). No detectable Mead acid occurred in the resultant fungal oil. The major fatty acids other than 20:2ω9 (17.6%) were palmitic acid (6.1), stearic acid (7.0), oleic acid (44.0), and 6(*Z*),9(*Z*)-octadecadienoic acid (18:2ω9, 17.6). Addition of olive oil, in which oleic acid is abundantly present, to the growth medium enhanced the production of this unique fatty acid (S. Shimizu, unpublished observation, 1998).

Nonmethylene-Interrupted C_{20} PUFAs

The occurrence in plants of two nonmethylene-interrupted PUFAs, that is, 5(*Z*),11(*Z*),14(*Z*)-eicosatrienoic acid (20:3Δ5) and 5(*Z*),11(*Z*),14(*Z*),17(*Z*)-eicosatetraenoic acid (20:4Δ5), has been reported (40,41). Microorganisms could be a good source of nonmethylene-interrupted PUFAs. These nonmethylene-interrupted PUFAs were found in the Δ6 desaturase-defective mutant Mut49 (18). The amount of 20:3Δ5 was greatest (27 mg/g dry mycelia) on growth at 20 °C, and it accounted for about 7% of the total mycelial fatty acids. 20:4Δ5 was detected only when the mutant was grown at a temperature lower than 24 °C or in a culture medium supplemented with either ALA or 20:3ω3. 20:3Δ5 is believed to be derived from LA by two subsequent reactions without Δ6 desaturation, that is, elongation and Δ5 desaturation (see Fig. 1).

PRODUCTION OF C_{22} FATTY ACIDS

Docosahexanoic Acid

Nakahara et al. (66) have reported that a thraustochytrid isolated from the marine environment grew well in conventional seawater medium and accumulated DHA intra-

cellularly. The microorganism was thought to be a strain belonging to *Thraustochytrium / Shizochytrium*, because heterkont zoospores were observed in the early stage of growth in its life cycle. A dry cell weight of 40 g/L, lipid content of 2 g/L, and DHA content of about 35% were obtained in a stirred tank fermentor at 28 °C when grown in a medium that contained 90 g glucose, 10 g polypeptone, and 10 g corn steep liquor in 1 L diluted artificial seawater (one-half concentration of seawater) in 14 days. The lipids extracted were composed mostly of triacylglycerol (53.5% of total lipids). The main fatty acids in the total lipids were 16:0 (44.6%), DHA (34.0%), and 22:5 (8.5%); the EPA content was only 0.4%. The organisms may be an excellent source of DHA.

PRODUCTION OF NOVEL PUFAS

Odd-Numbered PUFAs

The occurrence of straight-chain odd-numbered PUFAs in the lipids of ruminants, fish, and several animals has been reported. Shimizu et al. found that *Mortierella* and several related filamentous fungi convert *n*-alkanes or fatty acids of 15 or 17 carbon atoms to unusual C_{19} PUFAs efficiently (Fig. 6) (46,47,67).

n-1 Fatty Acid

Mortierella alpina 1S-4 accumulates a novel 5(*Z*),8(*Z*),11(*Z*),-14(*Z*),19-eicosapentaenoic acid (20:5ω1) together with additional *n*-1 fatty acids, such as 15-hexadecenoic acid, 17-octadecenoic acid, and 8(*Z*),11(*Z*), 14(*Z*),19-eicosatetraenoic acid (20:4ω1) on growth with 1-hexadecene or 1-octadecene (Fig. 6) (48). Based on the hypothesis that a terminal double bond has no effect on PUFA biosynthesis, it is suggested that the 20:5ω1 is formed through the *n*-6 route. The ω methyl group, which is opposite the double bond end, may be first oxidized to a carboxyl group, and the resultant *n*-1 fatty acid may be introduced to the *n*-6 route.

PRODUCTION OF COCOA-BUTTER-EQUIVALENT YEAST OIL

Cocoa butter, the essential ingredient of chocolate, owes its unique physical properties to the fatty acyl groups on the three available positions of the glycerol moiety, these being palmitoyl, oleoyl, and stearoyl residues, respectively. Thus, an ideal oil substitute, or cocoa butter equivalent, should consist of two-thirds saturated fatty acid (palmitate and stearic acid) and one-third monounsaturated acid (oleic acid; 18:1ω9). When sterculic acid, a known inhibitor of the Δ9 desaturase, was added to growing yeast cultures, stearic acid accumulated to greater than 40% (w/w) of the total fatty acid in the yeast oil. The resulting oil was good for a cocoa butter equivalent (68). Based on this result, a number of oleic acid–requiring yeasts that had a functionally inactive Δ9 desaturase were screened and found to produce considerable amounts of stearic acid good for a cocoa butter equivalent (4). Decreasing the O_2 supply to culture of the yeasts could also increase the stearic acid content of the cell (69). As a result, the large-scale development of a yeast cocoa butter equivalent was established (5).

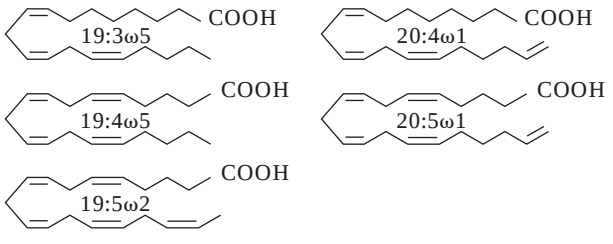


Figure 6. Novel PUFAs produced by *M. alpina* 1S-4 and related filamentous fungi.

CLOSING REMARKS

Filamentous fungi of the genus *Mortierella* are good producers of several useful single cell oils, especially those containing PUFAs such as AA, DGLA, and EPA. Furthermore, mutants of these fungi have expanded abilities for producing new oils and PUFAs or increased productivities for existing ones. A soil isolate, *M. alpina* 1S-4, and its mutants are useful not only as producers of some useful fatty acids and PUFAs such as AA, DGLA, and Mead acid, but they also provide useful information on PUFA biosynthesis. The enzymes, that is, desaturases and elongase, that are involved in the PUFA biosynthesis in this fungus seem to have a wide substrate specificity. For example, Δ5 desaturase can act on odd-numbered fatty acids (i.e., 19:3ω5), PUFA with an ω-terminal double bond (i.e., 20:4ω1), *n*-3 PUFA (i.e., 20:4ω3), *n*-6 PUFA (i.e., DGLA), *n*-9 PUFA (i.e., 20:2ω9), and PUFAs with no C-8 double bond (i.e., 11(*Z*),14(*Z*)-eicosadienoic acid and 11(*Z*),14(*Z*), 17(*Z*)-eicosatrienoic acid). However, the first double bond needs to be inserted at the C-9 position. These mutants may also be used as tools for cloning the desaturase genes that may be introduced into other organisms to produce either transgenic plants or animals that have more desirable fatty acid compositions.

Of longer-term interest is the prospect of using microorganisms to produce the end products of PUFA metabolism found in animals, the prostaglandins, leukotrienes, and thromboxanes, which are extraordinarily expensive to produce by chemical synthetic routes. Some microorganisms appear to have the ability to carry out a cascade of reactions, similar to those found in animals, leading to the formation of hydroxy-PUFAs and perhaps even to the prostaglandins (70). This concept would be a logical extension of the work already accomplished in the production of the PUFAs. The ability of microorganisms to catalyze the formation of products not found in plants, or that are difficult to extract from animals, is also promising. Microorganisms may be a source of these materials, and perhaps also some unusual lipids, ranging from the cerebrosides found in yeasts (6) to the ether lipids of archaebacteria, which approximate the platelet-activating factor found in animals (6).

BIBLIOGRAPHY

1. R.S. Moreton, ed., *Single Cell Oil*, Longman, Harlow, U.K., 1988.
2. D.J. Kyle and C. Ratlege, eds., *Industrial Applications of Single Cell Oils*, American Oil Chemists Soc., Illinois, 1992.
3. C. Ratledge, *Microbial Lipids*, vol. 2, Academic Press, London, 1989, pp. 567–668.
4. R.J. Davis, *Lipid Technol.* **4**, 6–13 (1992).
5. R.J. Davis and J.E. Holdworth, *Adv. Appl. Lipid Res.* **1**, 119–159 (1992).
6. C. Ratledgs, *Trends Biotechnol.* **11**, 278–284 (1993).
7. S. Bergstrom, H. Danielsson, and B. Samuelsson, *Biochim. Biophys. Acta* **90**, 207–210 (1964).
8. D.A. Van Dorp, R.K. Beerthuis, D.H. Nugteren, and H. Vonkeman, *Biochim. Biophys. Acta* **90**, 204–207 (1964).
9. B.A. Jakschik, A.R. Sams, H. Sprecher, and P. Needleman, *Prostaglandins* **20**, 401–410 (1980).
10. P.M. Herold and J.K. Kinsella, *Am. J. Clin. Nutr.* **43**, 566–598 (1986).
11. L. Oxdale, *Chem. Br.* **21**, 813 (1990).
12. J. Erwin and K. Bloch, *Science* **143**, 1006–1012 (1964).
13. R. Shaw, *Biochim. Biophys. Acta* **98**, 230–237 (1965).
14. O. Suzuki, *Hakko to Kogyo* **43**, 1024–1031 (1985).
15. O. Suzuki, *Proc. of the World Conf. on Biotechnology for the Fats and Oils Industry, Hamburg*, American Oil Chemists' Soc., 1988, p. 110.
16. H. Yamada, S. Shimizu, and Y. Shinmen, *Agric. Biol. Chem.* **51**, 785–790 (1987).
17. Y. Shinmen, S. Shimizu, K. Akimoto, H. Kawashima, and H. Yamada, *Appl. Microbiol. Biotechnol.* **31**, 11–16 (1989).
18. S. Jareonkitmongkol, S. Shimizu, and H. Yamada, *J. Gen. Microbiol.* **138**, 997–1002 (1992).
19. S. Jareonkitmongkol, E. Sakuradani, H. Kawashima, and S. Shimizu, *Nippon Nogeikagakukaishi* **67**, 531 (1993).
20. S. Jareonkitmongkol, S. Shimizu, and H. Yamada, *Biochim. Biophys. Acta* **1167**, 137–141 (1993).
21. O. Hiruta, K. Yamamura, H. Takebe, T. Futamura, K. Iinuma, and H. Tanaka, *J. Ferment. Bioeng.* **83**, 79–86 (1997).
22. M. Certik, S. Sereke-Berhan, and J. Sajbidor, *Acta Biotechnol.* **13**, 193–196 (1993).
23. M. Certik, P. Andrasi, and J. Sajbidor, *J. Am. Oil Chem. Soc.* **73**, 357–365 (1996).
24. M. Certik, L. Balteszova, and J. Sajbidor, *Lett. Appl. Microbiol.* **25**, 101–105 (1997).
25. L. Kristofikova, M. Rosenberg, A. Vlnova, J. Sajbidor, and M. Certik, *Folia Microbiol.* **36**, 451–455 (1991).
26. M. Certik and J. Sajbidor, *Microbios* **85**, 151–160 (1996).
27. S. Shimizu, K. Akimoto, H. Kawashima, Y. Shinmen, and H. Yamada, *J. Am. Oil Chem. Soc.* **66**, 237–244 (1989).
28. S. Jareonkitmongkol, H. Kawashima, N. Shirasaka, S. Shimizu, and H. Yamada, *Appl. Environ. Microbiol.* **58**, 2196–2200 (1992).
29. S. Jareonkitmongkol, E. Sakuradani, and S. Shimizu, *Appl. Environ. Microbiol.* **59**, 4300–4304 (1993).
30. D.M. Lösel, in C. Ratledge and S.G. Wilkinson eds., *Microbial Lipids*, Academic Press, London, 1988, pp. 699–806.
31. T.J. Ahern, S. Katoh, and E. Sada, *Biotechnol. Bioeng.* **25**, 1057–1070 (1983).
32. A. Kendrick and C. Ratledge, *Appl. Microbiol. Biotechnol.* **37**, 18–22 (1992).
33. Y. Shinmen, K. Katoh, S. Shimizu, S. Jareonkitmongkol, and H. Yamada, *Phytochemistry* **30**, 3255–3260 (1991).
34. T. Yaguchi, S. Tanaka, T. Yokochi, T. Nakahara, and T. Higashihara, *J. Am. Chem. Soc.* **74**, 1431–1434 (1997).
35. H. Kawashima, N. Kamada, E. Sakuradani, S. Jareonkitmongkol, K. Akimoto, and S. Shimizu, *J. Am. Oil Chem. Soc.* **74**, 455–459 (1997).
36. S. Shimizu, H. Kawashima, Y. Shinmen, K. Akimoto, and H. Yamada, *J. Am. Oil Chem. Soc.* **65**, 1455–1459 (1988).
37. S.R. Gandai and J.D. Weete, *J. Gen. Microbiol.* **137**, 1825–1830 (1991).
38. K. Yazawa, K. Watanabe, C. Ishikawa, K. Kondo, and S. Kimura, in D.J. Kyle and C. Ratledge eds., *Industrial Applications of Single Oils*, American Oil Chemists' Soc., Illinois, 1992, pp. 29–51.
39. S. Jareonkitmongkol, S. Shimizu, and H. Yamada, *J. Am. Oil Chem. Soc.* **70**, 119–123 (1993).
40. D.J. O'Brien, M.J. Kurantz, and R. Kwoczak, *Appl. Microbiol. Biotechnol.* **40**, 211–214 (1993).
41. Z. Cohen and Y.M. Heimer, in D.J. Kyle and C. Ratledge eds., *Industrial Applications of Single Oils*, American Oil Chemists' Soc., Illinois, 1992, pp. 243–273.
42. P.K. Bajpai, P. Bajpai, and O.P. Ward, *J. Am. Oil Chem. Soc.* **68**, 509–514 (1991).
43. D.J. Kyle, V.J. Sicotte, and S.E. Reet, in D.J. Kyle and C. Ratledge eds., *Industrial Applications of Single Oils*, American Oil Chemists' Soc., Illinois, 1992, pp. 287–300.
44. S. Jareonkitmongkol, H. Kawashima, S. Shimizu, and H. Yamada, *J. Am. Oil Chem. Soc.* **69**, 939–944 (1992).
45. H. Kawashima, M. Nishihara, Y. Hirano, N. Kamada, and S. Shimizu, *Appl. Environ. Microbiol.* **63**, 1820–1825 (1997).
46. S. Shimizu, H. Kawashima, K. Akimoto, Y. Shinmen, and H. Yamada, *J. Am. Oil Chem. Soc.* **68**, 254–258 (1991).
47. N. Shirasaka, T. Yokochi, and S. Shimizu, *Biosci. Biotechnol. Biochem.* **59**, 1963–1965 (1995).
48. S. Shimizu, S. Jareonkitmongkol, H. Kawashima, K. Akimoto, and H. Yamada, *Arch. Microbiol.* **156**, 163–166 (1991).
49. C.R. Smith, Jr., R. Kleiman, and I.A. Wolff, *Lipids* **3**, 37–42 (1968).
50. J.L. Gellerman and H. Schlenk, *J. Am. Oil Chem. Soc.* **42**, 504–511 (1965).
51. N. Amano, Y. Shinmen, K. Akimoto, H. Kawashima, T. Amachi, S. Shimizu, and H. Yamada, *Mycotaxon* **44**, 257–265 (1992).
52. S. Shimizu, Y. Shinmen, H. Kawashima, K. Akimoto, and H. Yamada, *Biochem. Biophys. Res. Commun.* **150**, 335–341 (1988).
53. S. Shimizu, K. Akimoto, H. Kawashima, Y. Shinmen, S. Jareonkitmongkol, and H. Yamada, *Agric. Biol. Chem.* **53**, 1437–1438 (1989).
54. S. Shimizu, K. Akimoto, Y. Shinmen, H. Kawashima, M. Sugano, and H. Yamada, *Lipids* **26**, 512–516 (1991).
55. S. Shimizu, H. Kawashima, K. Akimoto, Y. Shinmen, H. Sugano, and H. Yamada, *Phytochemistry* **31**, 757–760 (1992).
56. S. Shimizu, S. Jareonkitmongkol, H. Kawashima, K. Akimoto, and H. Yamada, *Lipids* **27**, 509–512 (1992).
57. H. Kawashima, K. Akimoto, S. Jareonkitmongkol, N. Shirasaka, and S. Shimizu, *Biosci. Biotechnol. Biochem.* **60**, 108–110 (1996).
58. H. Kawashima, K. Akimoto, S. Jareonkitmongkol, N. Shirasaka, and S. Shimizu, *Biosci. Biotechnol. Biochem.* **60**, 1672–1676 (1996).

59. S. Jareonkitmongkol, H. Kawashima, and H. Yamada, *J. Ferment. Bioeng.* **76**, 406–407 (1993).
60. S. Jareonkitmongkol, E. Sakuradani, and S. Shimizu, *Arch. Microbiol.* **161**, 316–319 (1994).
61. J.L. Gellerman and H. Schlenk, *Biochim. Biophys. Acta* **573**, 23–30 (1979).
62. S. Shimizu, H. Kawashima, K. Akimoto, Y. Shinmen, and H. Yamada, *J. Am. Oil Chem. Soc.* **66**, 342–347 (1989).
63. S. Shimizu, H. Kawashima, K. Akimoto, Y. Shinmen, and H. Yamada, *Appl. Microbiol. Biotechnol.* **32**, 1–4 (1989).
64. N. Shirasaka and S. Shimizu, *J. Am. Oil. Chem. Soc.* **72**, 1545–1549 (1995).
65. A.J. Fulco and J.F. Mead, *J. Biol. Chem.* **234**, 1411–1416 (1959).
66. T. Nakahara, T. Yokochi, K. Takeda, K. Higashihara, and O. Suzuki, *Abstracts of 21th World Congress and Exhibition of the International Society for Fat Research*, 13 (1995).
67. S. Shimizu, H. Kawashima, K. Akimoto, Y. Shinmen, and H. Yamada, *Proceedings of Session Lectures and Scientific Presentation on ISF-JOCS World Congress 1988*, vol. 2, Japan Oil Chemists' Soc., 1989, pp. 1102–1106.
68. R.S. Moreton, *Appl. Microbiol. Biotechnol.* **33**, 41–45 (1985).
69. R.J. Davies, J.E. Holdworth, and S.L. Reader, *Appl. Microbiol. Biotechnol.* **33**, 569–573 (1990) [Erratum: **43**, 832–833].
70. J.L.F. Kock and C. Ratledge, *J. Gen. Microbiol.* **139**, 459–464 (1993).

OPINE DEHYDROGENASE, SECONDARY AMINE DICARBOXYLIC ACIDS

YASUO KATO
YASUHISA ASANO
Toyama Prefectural University
Toyama, Japan

KEY WORDS

Amino acid dehydrogenase
Enzymatic synthesis
Enzyme purification and properties
Opine compound
Opine dehydrogenase
Secondary amine dicarboxylic acid

OUTLINE

NATURALLY OCCURRING OPINES

Opines are unusual compounds found in crown gall tumor tissues in plants induced by an infection of *Agrobacterium tumefaciens* (1–6), and in muscle tissue of marine invertebrates (2,3,7–8). They can be categorized as secondary amine dicarboxylic acids (octopine and nopaline families), mannityl opines (agropine), and phosphorylated sugars (agrocinopine). Synthesis of opines in infected plants occurs when the infecting bacterium (i.e., *A. tumefaciens*) transfers a tumor-inciting (Ti) plasmid to the plant, causing tumorigenesis (9). A gene present in this plasmid induces opine synthesis in the plants, while a chromosomal gene encodes for the utilization of the synthesized opine. In the muscle of marine invertebrates, secondary amine dicarboxylic acids are the biochemical equivalent of lactate and represent the terminal products of the glycolytic pathway (10) due to an absence of lactate dehydrogenase (LDH) in their tissues.

SECONDARY AMINE DICARBOXYLIC ACIDS

Of the naturally occurring opines, secondary amine dicarboxylic acids (Fig. 1) are the most extensively studied. These compounds may be regarded as two amino acids linked via a common imino nitrogen, or as secondary amines. One group can be categorized as an octopine family (2,3,11), which has an N-substituted D-alanine moiety, and the other, representing N-substituted D-glutamate derivatives, is the nopaline family (2,3,12). The members of the former family are formed by the reductive condensation reaction between pyruvate and the α-NH_2 group of L-arginine, L-lysine, L-ornithine, L-histidine, and L-methionine, and the corresponding opines are named as octopine (13,14), lysinopine (4,15), octopinic acid (14,16), histopine (6), and methiopine (17), respectively. Octopine is also found in marine invertebrates, together with structurally related alanopine (18), strombine (19), β-alanopine (20), and tauropine (21), which have L-alanine, glycine, β-alanine, and taurine moieties, respectively. Opines of the nopaline family are derived from the condensation between α-ketoglutarate and the α-NH_2 group of L-arginine and L-ornithine, to give the corresponding opines nopaline (22) and nopalinic acid (23), respectively. Saccharopine, N^6-(1,3-L-dicarboxypropyl)-L-lysine, was isolated from brewer's yeast (24) and has the L,L-stereochemistry. In the amino acid pool of *Lactococcus lactis* subsp. lactis (2,3), Thompson et al. reported the existence of N^5-(1-L-carboxyethyl)-L-ornithine and N^5-(1-L-carboxyethyl)-L-lysine, the condensation products between alanine and the ω-NH_2 group of L-ornithine and L-lysine, respectively. Recently, Fushiya et al. (25) isolated four new opines, valinopine,

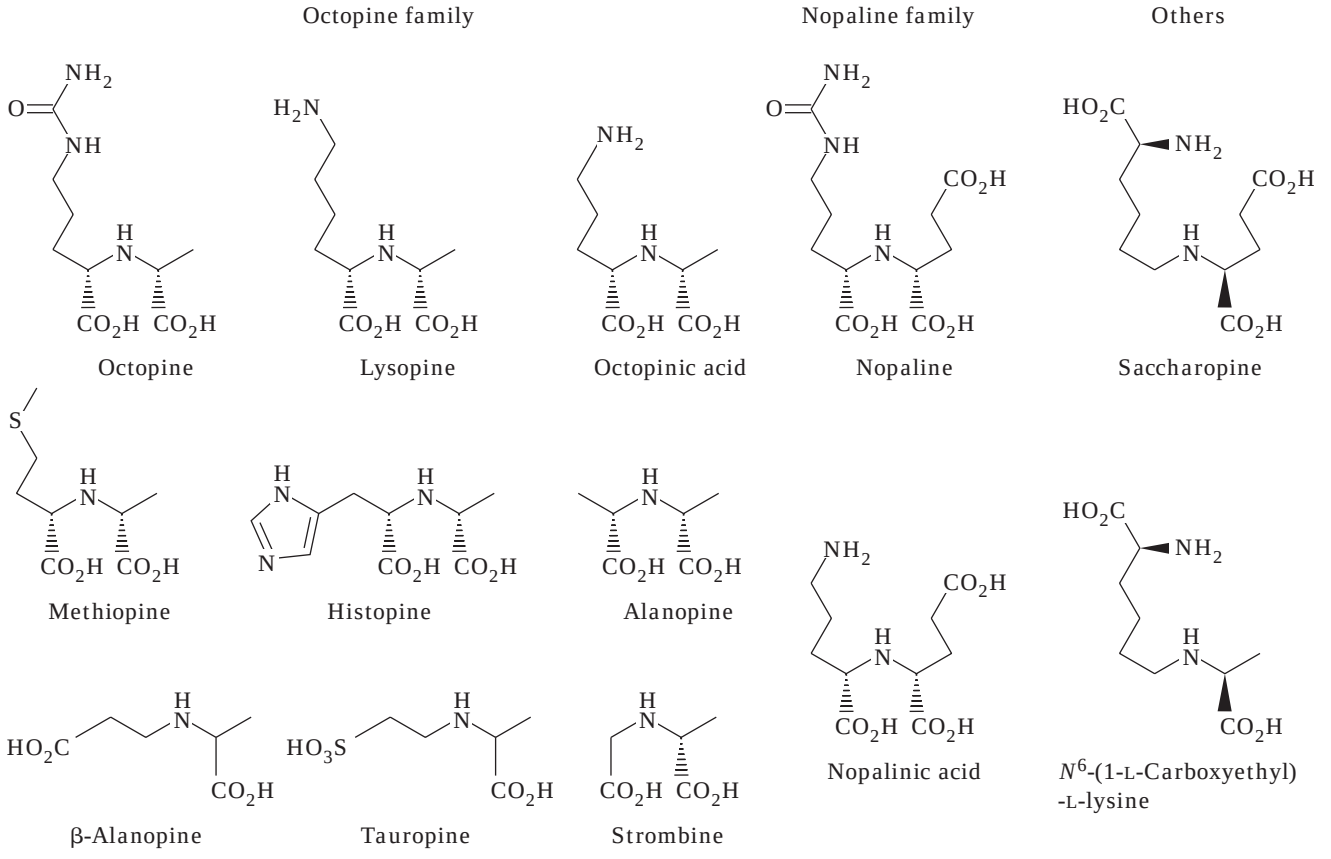


Figure 1. Naturally occurring secondary amine dicarboxylic acids.

epileucinopine, isoleucinopine, and phenylalanopine, which were derived from α-ketoglutarate and L-valine, L-leucine, L-isoleucine, and L-phenylalanine, respectively, from a poisonous mushroom, *Clitocybe acromelalga*.

ENZYMES ACTING ON SECONDARY AMINE DICARBOXYLIC ACIDS

Enzymes that can synthesize secondary amine dicarboxylic acids by the reductive condensation of an α-keto acid with the NH_2-group of an amino acid have been isolated from plants, marine invertebrates, and bacteria (2,3). These enzymes are all NAD(P)H-dependent oxidoreductases and appear to be responsible for the biosynthesis of opines.

Octopine dehydrogenase [N^2-(1-D-carboxyethyl)-L-arginine: $NAD(P)^+$ oxidoreductase, EC 1.5.1.1.1] has been purified from crown gall (26), *Pecten maximus*, and other marine invertebrates (27). The enzyme is a monomeric enzyme and has a molecular weight of 37,000–46,000. The enzyme from crown gall catalyzes the NAD(P)H-dependent condensation reaction between pyruvate and several amino acids such as basic and short-chain neutral amino acids (26), whereas that from *Pecten maximus* exhibits a marked preference for pyruvate, L-arginine, and NADH. The gene encoding the enzyme (*ocs*) has been cloned and sequenced from *A. tumefaciens* (28,29).

Lysopine dehydrogenase [N^2-(1-D-carboxyethyl)-L-lysine: $NADP^+$ oxidoreductase, EC 1.5.1.16] has been purified from crown gall tumor tissues (30). The enzyme catalyzes pyruvate-dependent synthesis of secondary amine dicarboxylic acids of L-lysine, L-arginine, L-ornithine, L-histidine, L-methionine, L-glutamine, and L-citrulline. The enzyme is a monomeric enzyme with a molecular weight of 35,000 and prefers NADP(H) to NAD(H) as the cofactor. This makes it different from the analogous enzyme, octopine dehydrogenase, which utilizes both NAD(H) and NADP(H).

Nopaline dehydrogenase [N^2-(1,3-D-dicarboxypropyl)-L-arginine: $NADP^+$ oxidoreductase, EC 1.5.1.19] catalyzes the α-ketoglutarate-dependent condensation reaction with L-arginine and L-ornithine, producing nopaline and nopalinic acid (ornaline), respectively. The enzyme has been purified from tumor tissue of *Nicotiania tabacum* (31) and sunflower crown gall (32). The enzyme is a tetrameric protein with a subunit molecular weight of 40,000 and requires NADP(H) as a specific cofactor. L-Lysine and L-histidine, good substrates for octopine and lysopine dehydrogenases, do not act as substrates for the enzyme. The nopaline dehydrogenase gene (*nos*) from *A. tumefaciens* has been cloned and sequenced, and the enzyme was expressed in *Escherichia coli* (33,34).

Alanopine dehydrogenase [*meso*-*N*-(1-carboxyethyl)-alanine: NAD^+ oxidoreductase, EC 1.5.1.17] has been purified to homogeneity from adductor muscle of oyster (35)

and partially from the lungworm (36) and the foot muscle of the mussel (37). The monomeric enzyme with a molecular weight of 38,500–46,000 requires NADH as a cofactor and catalyzes the reversible condensation reaction between hydrophilic α-keto acids, such as pyruvate, glyoxylate, α-ketoglutarate, and hydroxypyruvate, and amino acids, such as L-alanine, glycine, L-cysteine, L-serine, and L-threonine.

Strombine dehydrogenase [*N*-(1-carboxymethyl)-D-alanine: NAD^+ oxidoreductase, EC 1.5.1.22] has been purified from *Mytilus edulis* (37). The enzyme (M_r = 38,000) catalyzes the NAD(H)-dependent reversible condensation reaction between glycine and pyruvate. Although the enzyme and alanopine dehydrogenase exhibit similar substrate specificities with respect to glycine and alanine, the enzyme prefers glycine as its amino donor.

Tauropine dehydrogenase [*N*-(2-sulfonylethyl)-D-alanine: NAD^+ oxidoreductase, EC 1.5.1.-] has been purified from *Arabella iricolor* (Polychaeta: Errantia) (38). The monomeric enzyme (M_r = 43,500) catalyzes the NAD(H)-dependent reversible condensation reaction between taurine and hypotaurine, and hydrophilic α-keto acids, such as pyruvate, oxalacetate, α-ketobutyrate, hydroxypyruvate, and glyoxylate.

These $NAD(P)^+$-dependent dehydrogenases appear to share a similar reaction mechanism with the $NAD(P)^+$-dependent amino acid dehydrogenases. It is peculiar that the $NAD(P)^+$-dependent secondary amino acid dehydrogenases will form a D-configuration from α-keto acid in the reductive amine-forming reaction because no $NAD(P)^+$-dependent amino acid dehydrogenase forms a D-amino acid from the ketoanalog in the reductive amination reaction, with $NADP^+$-dependent *meso*-α, ε-diaminopimelate D-dehydrogenase as an exception (39).

Saccharopine dehydrogenase, isolated from yeasts and fungi (2,3), catalyze the reversible oxidoreduction reaction of N^6-(1,3-L-dicarboxypropyl)-L-lysine, and the enzyme gene from *Yarrowia lipolytica* has been cloned and sequenced (40). Thompson purified an NADPH-dependent oxidoreductase responsible for the biosynthesis of both N^5-(L-1-carboxyethyl)-L-ornithine and N^6-(L-1-carboxyethyl)-L-lysine from *Lactococcus lactis* subsp. lactis (2,3). This enzyme, N^5-(L-1-carboxyethyl)-L-ornithine: $NADP^+$ oxidoreductase (EC 1.5.1.24), is a tetramer composed of identical subunits (M_r = 35,000) and mediates the reductive condensation reaction between pyruvate and the ω-NH_2 group of L-ornithine and L-lysine, but not an oxidation reaction of the secondary amine dicarboxylic acids. Recently, the gene encoding for the enzyme was cloned and sequenced, and the enzyme was expressed in *E. coli*. (41). Both enzymes are different from the already-mentioned dehydrogenases because opines synthesized by the enzymes arise via alkylation of the ω-NH_2 group of amino acids, and the stereochemistry of the newly formed asymmetric center is of L-configuration.

OPINE DEHYDROGENASE FROM *ARTHROBACTER* SP. STRAIN 1C

Asano et al. have purified phenylalanine dehydrogenases in crystalline forms or to homogeneity and characterized those from *Sporosarcina ureae* SCRC-R04 (42), *Bacillus sphaericus* SCRC-R79a (42), and *B. badius* IAM 11059 (43). The *pdh* genes have been cloned into plasmids and expressed in *E. coli*, and the *pdh* gene has been sequenced (44,45). They have used phenylalanine dehydrogenase and other amino acid dehydrogenases to synthesize optically pure, natural and unnatural amino acids from their corresponding α-keto acids (46,47). Much attention has been paid to a new amino acid dehydrogenase acting on secondary amine dicarboxylic acid. Comparison of the primary structure of the dehydrogenase with that of the known amino acid dehydrogenases would give a valuable information on the structure and function of this class of enzymes. Furthermore, the reversible nature of the enzyme-catalyzed reaction will make it possible to synthesize new chiral opine derivatives, the chemical synthesis of which is rather laborious.

Isolation of (1-D-Carboxyethyl)-L-Phenylalanine-Assimilating Bacteria

Several opine degraders, such as *A. tumefaciens*, *Pseudomonas* sp., and *Arthrobacter* sp., were isolated from soil and crown gall tumors (48,49). Degradation of opines (i.e., lysopine by *A. tumefaciens*) has been attributed by Jubier (50) to a membrane-bound enzyme complex that is responsible for the oxidation of lysopine, forming lysine and pyruvate.

In an attempt to utilize an NAD^+-dependent dehydrogenase for the stereospecific in vitro enzymatic synthesis of a wide variety of opines, Asano et al. (51,52) attempted to isolate several microorganisms that can grow on chemically synthesized opines. They isolated the bacterial strain YEO8, which utilizes *N*-(1-D-carboxy-2-phenylethyl)-D-phenylalanine as its carbon source from soil and is identified as a glucose-nonfermenting Gram-negative bacteria under CDC group IVc-2 (52). The cell-free extract of the strain contained enzymes that catalyzed the oxidation of the opine into phenylpyruvate and D-phenylalanine, with DCPIP but not NAD(P) as its hydrogen acceptor. They isolated a novel NAD^+-dependent secondary amine dicarboxylic acid dehydrogenase producer, strain 1C, from soil with *N*-(1-D,L-carboxyethyl)-L-phenylalanine as a substrate (51). The taxonomical studies of the strain 1C indicate that it belongs to the genus *Arthrobacter* because it is nonmotile and never fermentative, occurs in rod-coccus shape, its peptidoglycan contains lysine as the diamino acid, and no mycelium formation is observed. *Arthrobacter* strains in the type culture collections did not grow on a medium containing *N*-(1-D-carboxyethyl)-L-phenylalanine as the sole source of carbon.

Purification of Opine Dehydrogenase from Arthrobacter sp. Strain 1C

The enzyme activity found in *Arthrobacter* sp. strain 1C was inducible in the presence of *N*-(1-D,L-carboxyethyl)-L-phenylalanine; only one tenth of total enzyme activity per culture was obtained in the absence of the substrate. The enzyme was purified about 57-fold, with a 36% yield from the cell-free extract of the strain by a procedure involving ammonium sulfate fractionation, and DEAE-Toyopearl,

butyl-Toyopearl, and Sephadex G-200 column chromatographies (53). The enzyme has a molecular weight of about 70,000 and consists of two identical subunits with an M_r of 36,000. Typical properties of the enzyme are shown in Table 1. The enzyme catalyzes a reversible oxidation–reduction reaction of opine-type secondary amine dicarboxylic acids. The enzyme is unique because the only similar microbial enzyme thus far reported (2,3) is N^5-(L-1-carboxyethyl)-L-ornithine: $NADP^+$ oxidoreductase from *L. lactis* subsp. lactis, which catalyzes only the reductive secondary amine forming reaction with NADPH, not the oxidative degradation reaction. The velocity of opine dehydrogenase for the secondary amine–forming reaction is about six times greater than that for the oxidative reaction. In the oxidative deamination reaction, the enzyme is active toward synthetic opines such as *N*-(1-D-carboxyethyl)-L-methionine and *N*-(1-D-carboxyethyl)-L-phenylalanine, and its K_m values for them are 9.0 and 14 mM, respectively. On the other hand, the enzyme does not act on naturally occurring opines, such as saccharopine, D-octopine, and L-allooctopine, or stereoisomers of *N*-(1-D-carboxyethyl)-L-phenylalanine, such as *N*-(1-L-carboxyethyl)-L-phenylalanine, *N*-(1-L-carboxyethyl)-D-phenylalanine, and *N*-(1-D-carboxyethyl)-D-phenylalanine. In the reaction that forms secondary amines using pyruvate as a fixed amino acceptor, the enzyme is active toward short-chain aliphatic L-amino acids and those substituted with acyloxy, phosphonooxy, and halogen groups, as shown in Table 2. Hydrophilic α-keto acids, such as pyruvate, oxaloacetate, glyoxylate, and α-ketobutyrate, are accepted as amino acceptors (Table 3). Based on its results of substrate specificity, the enzyme is tentatively named *N*-(1-D-carboxyethyl)-L-norvaline: NAD^+ oxidoreductase (L-norvaline-forming: EC 1.5.1.-, opine dehydrogenase [ODH]). Other substrates for the enzyme are 3-aminobutyric acid and L-phenylalaninol.

The $NADP^+$-dependent D-lysopine dehydrogenase (31,32) and D-nopaline dehydrogenase (30) catalyze a reversible oxidation–reduction reaction similar to the reaction catalyzed by ODH. However, it is a monomeric protein and the substrate and coenzyme specificities are totally different from those of ODH. In the reductive secondary amine–forming reactions with NADPH and NADH (20% the velocity with NADPH) as cofactors, the enzyme does not accept L-phenylalanine, L-alanine, and L-isoleucine, and so on, which are good substrates for ODH.

Table 1. Properties of Opine Dehydrogenase from *Arthrobacter* sp. Strain 1C

Molecular weight	
SDS	36,000
Gene sequencing	37,935
Molecular structure	Dimer
Maximum absorption wavelength	280 nm
Absorption coefficient	ε_{280} = 75.87 mM^{-1} cm^{-1}
Specific activity[a]	140 U/mg
Optimum pH	8.0 (reduction) 10.0 (oxidation)
Optimum temperature	55 °C
pH stability[b]	pH 5.0 ~ 9.5
Temperature stability[c]	~50 °C
K_m for NAD^+	0.76 mM
Inhibitors	Sulfuhydryl reagents
	Heavy metals

[a]Measured by an oxidation of 10 mM of *N*-(1-D-carboxyethyl)-L-phenylalanine at pH 9.5.
[b]The enzyme activity was measured after incubation at various temperatures and pH 8.0 for 10 min.
[c]The enzyme activity was measured after incubation at 30 °C for 1 h with buffers at various pHs.

Table 2. Substrate Specificity of Opine Dehydrogenase in the Reductive Secondary Amine–Forming Reaction with Pyruvate as a Fixed Substrate

Amino donor[a]	Relative activity (%)[b]	K_m (mM^{-1})
L-Norvaline	100	2.17
L-α-Aminobutyric acid	83.7	20.0
L-Norleucine	72.5	3.72
β-Chloro-L-alanine	53.1	3.19
O-Acetyl-L-serine	35.8	5.94
L-Methionine	23.9	4.10
L-Isoleucine	22.3	6.20
L-Valine	21.9	3.00
L-Phenylalanine	21.7	8.70
L-Leucine	21.2	2.90
L-Alanine	16.0	5.10
O-Phospho-L-serine	15.8	8.69
D,L-α,γ-Diaminopropionic acid	9.67	11.9
L-Phenylglycine	4.40	28.3
D,L-β-Aminobutyric acid	0.683	7.73
O-Phospho-L-threonine	0.640	9.01
L-Phenylalaninol	0.631	46.0

Note: Assayed with fixed concentrations of pyruvate and NADH at 10 and 0.1 mM, respectively.
[a]The following compounds were inert as substrates: D-amino acids, γ-amino-β-hydroxybutyric acid, γ-aminobutyric acid, ε-aminohexanoic acid, L-phenylalanine methylester, L-phenylglycine methylester, L-norvaline methylester, L-alanine methylester, L-phenylalanine amide, L-phenylglycine amide, L-alanine amide, L-isoleucinol, L-valinol, L-alaninol, L-serinol, 1-amino-2,3-propanediol, 1-amino-2-propanol, (2*S*,3*S*)-2-amino-1-phenyl-1,3-propanediol, 3-amino-4-phenyl-4-butanol, ethanolamine, 2-amino-2-methyl-1,3-propanediol, methylamine, ethylamine, isopropylamine, butylamine, benzylamine, 3-aminopentane, 2-phenylethylamine, methoxyamine, benzyloxyamime dimethylamine, hydrazine, phenylhydrazine, and hydroxylamine.
[b]The enzyme activity for L-norvaline was taken as 100%.

Table 3. Substrate Specificity of Opine Dehydrogenase in the Reductive Secondary Amine–Forming Reaction with L-Methionine as a Fixed Substrate

Keto acid[a]	Relative activity (%)[b]	K_m (mM^{-1})
Pyruvate	100	3.0
Oxaloacetate	30	3.5
Glyoxylate	11	8.8
α-Ketobutyrate	11	28
Hydroxypyruvate	6.1	N.D.[c]
α-Ketocaproate	2.1	N.D.[c]

Note: Assayed with fixed concentrations of L-methionine and NADH at 10 and 0.1 mM, respectively.
[a]α-Keto-γ-phenylbutyrate, glyoxal, α-ketoglutarate, phenylpyruvate, and acetone were inert as a substrate.
[b]The enzyme activity for pyruvate was taken as 100%.
[c]Not determined.

Cloning, Nucleotide Sequencing, and Expression of Opine Dehydrogenase

The gene coding for ODH was cloned into plasmid p-Bluescript KS(−), and the nucleotide sequence of the 1,077-bp open reading frame consisting of 359 codons was identified as the *odh* gene coding for the enzyme (54). About 6.6-fold total activity per liter of culture was detected with a feeding of 1 mM IPTG, as compared with the wild-type strain *Arthrobacter* sp. 1C (51). The enzyme was purified and crystallized from the cell-free extracts of the recombinant strain (Fig. 2). The chemical and physical properties of the purified enzyme showed a complete agreement with those of the wild type.

The enzyme (359 amino acids) is similar to the sequence of the 40-kDa protein (27.4% identity among 359 amino acids) of the Ti plasmid of *A. tumefaciens* (55), D-lysopine dehydrogenase (24.6% identity among 358 amino acids) (28) and D-nopaline dehydrogenase (20.8% identity among 168 amino acids) (34) of *A. tumefaciens*, D-nopaline dehydrogenase (21.8% identity among 371 amino acids) of *A. vitis* (56), and phenylalanine dehydrogenases of *B. sphaericus* (21.8% identity among 78 amino acids) (44) and of *B. badius* (25.6% identity among 60 amino acids) (45). There is no apparent homology between ODH and saccharopine dehydrogenase (21) or N^5-(L-1-carboxyethyl)-L-ornithine: $NADP^+$ oxidoreductase (41). The common similarities between the 40-kDa protein, D-lysopine dehydrogenase, and the N-terminal region of ODH and the phenylalanine dehydrogenase (PheDH) are in the glycine-rich nucleotide-binding domain G-X-G-X-X-(G or A), connecting the β-strand with the α-helix in the region of adenine dinucleotide phosphate (ADP)–binding $\beta\alpha\beta$ folds, which is strongly conserved among $NAD(P)^+$-dependent dehydrogenases and FAD-containing oxidoreductases (57).

It can be suggested that ^{34}D of ODH corresponds to ^{214}D of the NAD^+-dependent *B. sphaericus* phenylalanine dehydrogenase, the position of which is considered to be important for the NADH binding discriminating NADPH among $NAD(P)^+$-dependent dehydrogenases (57). The equivalent positions with ^{34}D of the enzyme in the 40-kDa protein and D-lysopine dehydrogenase (which utilizes both $NADP^+$ and NAD^+) are neutral ^{30}S and ^{32}A, respectively, which are found in $NADP^+$-dependent dehydrogenases. There is also a similarity among these enzymes in the conserved ^{238}D-X-X-^{241}R residue, which is thought to be important for the proton-relay mechanism (3). The nucleotide-binding domain, as can be seen in typical lactate dehydrogenases, is located at the N-terminal region of these secondary amine dicarboxylic acid dehydrogenases, whereas other amino acid dehydrogenases have one more domain in the N-terminus important for accommodating the amino acid substrate (58). Octopine and nopaline dehydrogenases show 26% amino acid identity (3), but no similarity was seen with saccharopine dehydrogenase and N^5-(L-1-carboxyethyl)-L-ornithine: $NADP^+$ oxidoreductase. The former enzymes synthesize opines having D,L-configuration and are linked via the α-NH_2 group of amino acids, but the latter ones produce opines with L,L-configuration and are linked via the α- and ε-NH_2 group of amino acids. These differences in substrate specificities reflect the amino acid sequence similarity between these enzymes.

Although N^5-(L-1-carboxyethyl)-L-ornithine: $NADP^+$ oxidoreductase (41) has a similarity with lactate and alanine dehydrogenases (59) at pyruvate-binding sites, including active arginine (which is a catalytically relevant residue), ODH might have a different binding site with pyruvate because the enzyme had no similarity with these enzymes.

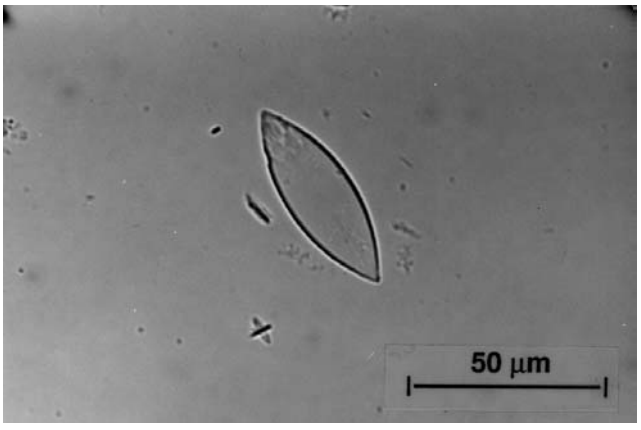


Figure 2. Photomicrograph of crystalline opine dehydrogenase.

Stereoselective Synthesis of Opine-Type Secondary Amine Carboxylic Acids by Opine Dehydrogenase

Opine-type secondary amine dicarboxylic acids are useful chiral intermediates of angiotensin-converting enzyme (ACE) inhibitors such as enalapril and lysinopril (60). They have been chemically synthesized by reductive condensation reactions of α-keto acids or their esters and amino acid derivatives (61,62) or SN_2 reactions of optically active 2-halo- and 2-trifluoromethanesulfonyloxy esters with amino acid derivatives (63,64). However, these methods generally require protection of the functional groups, and the stereoselectivity is not always high. Kato et al. (53) constructed a new system for the synthesis of secondary amine dicarboxylic acids by ODH coupled with NADH regeneration by formate dehydrogenase (FDH) from *Moraxella* sp. C-1 (65), without protection of the substrates (Fig. 3).

As shown in Table 4, several secondary amine dicarboxylic acids have been synthesized from amino acids and their analogues, such as L-methionine, L-isoleucine, L-leucine, L-valine, L-phenylalanine, L-alanine, L-threonine, and L-serine, and α-keto acids, such as glyoxylate, pyruvate, and 2-ketobutyrate, using ODH. A diastereo and enantiomerically pure secondary amine carboxylic acid from (*S*)-phenylalaninol and pyruvate was also quantitatively synthesized. The absolute configuration of the nascent asymmetric center of the opines was of the D-stereochemistry with $> 99.9\%$ ee. This method seems to be the simplest for stereoselectively synthesizing opines without protection of the substrates, because chemical processes generally require five to eight steps.

On the other hand, one-pot synthesis of *N*-[1-D-(carboxyl)ethyl]-L-phenylalanine from phenylpyruvate

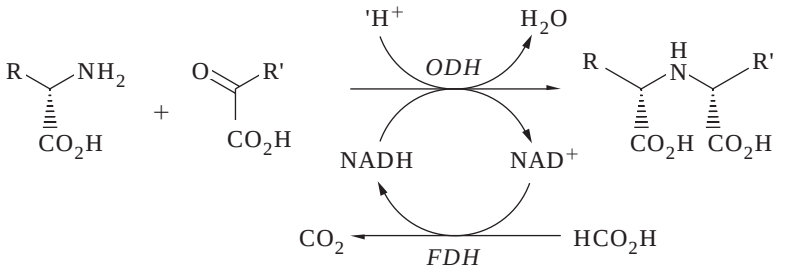


Figure 3. Synthesis of secondary amine dicarboxylic acids from L-amino acids and α-keto acids by opine dehydrogenase with regeneration of NADH by formate dehydrogenase. ODH, opine dehydrogenase from *Arthrobacter* sp. 1C; FDH, formate dehydrogenase from *Moraxella* sp. C-1.

Table 4. Synthesis of Opine-Type Secondary Amine Dicarboxylic Acids from L-Amino Acids and α-Keto Acids Using Opine Dehydrogenase and Formate Dehydrogenase

	α-Keto Acid						
	Glyoxylic acid	Pyruvic acid			2-Ketobutyric acid		
			Purity			Purity	
Amino acid	Yield[a] (%)	Yield[a] (%)	(% d.e.)[b]	(% e.e.)[a]	Yield[a] (%)	(% d.e.)[b]	(% e.e.)[a]
L-Methionine	–	96	>99.9	>99.9	–	–	–
L-Isoleucine	>99	>99	>99.9	>99.9	97	>99.9	>99.9
L-Leucine	>99	>99	>99.9	>99.9	–	–	–
L-Valine	>99	>99	>99.9	>99.9	>99	>99.9	>99.9
L-Phenylalanine	–	95	>99.9	>99.9	–	–	–
L-Alanine	–	98[c]	>99.9	>99.9	–	–	–
L-Threonine	–	>99	>99.9	>99.9	–	–	–
L-Serine	–	97	>99.9	>99.9	–	–	–

[a]Determined by HPLC.
[b]Determined by 400 MHz ^{1}H- and ^{13}C-NMR.
[c]The obtained compound was in the *meso*-form.

and pyruvate by using ODH, FDH, and PheDH from *B. sphaericus* (43,46) is also described, as shown in Figure 4. This reaction can be carried out well because of the substrate specificities of the enzymes thus used; ODH acts on pyruvate but not on phenylpyruvate (51), whereas PheDH acts on phenylpyruvate but not on pyruvate (43). The opine compound was obtained in a yield of 87.5% with high (>99.9%) enantio- and diastereomeric excess.

CONCLUSION

Both secondary amine dicarboxylic acid and amino acid dehydrogenases use NAD(P)H, α-keto acids, and amino donors as substrates and have a common consensus motif in their primary structures. The structural and catalytic similarities of these enzymes suggest that these genes have evolved from a common progenitor. On the other hand, differences have been found in the substrate specificities: *(1)* secondary amine dicarboxylic acid dehydrogenases utilize the α- or ε-amino group of amino acids, but amino acid dehydrogenases use ammonia as amino donors; and *(2)* the asymmetric center newly formed by secondary amine dicarboxylic acid dehydrogenases is of D-stereochemistry, except saccharopine dehydrogenase and N^5-(L-1-carboxyethyl)-L-ornithine: $NADP^+$ oxidoreductase, but that by formed amino acid dehydrogenases is exclusively of L-

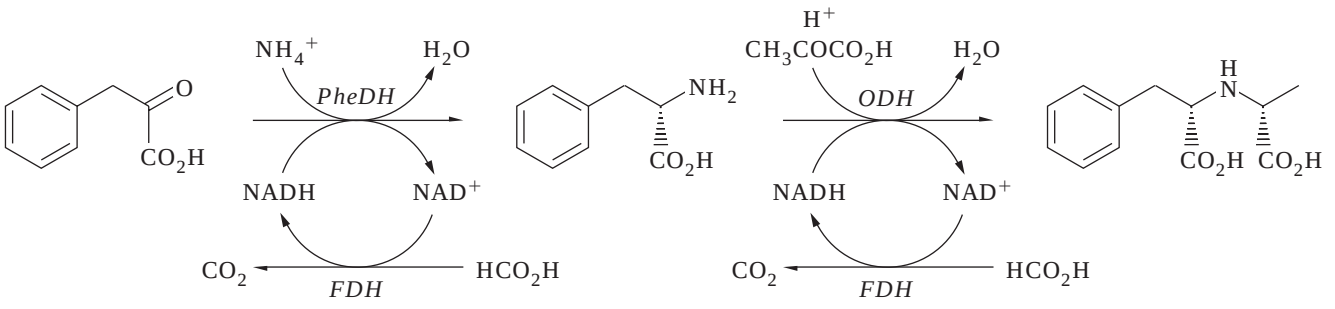


Figure 4. One-pot synthesis of *N*-(1-D-carboxyethyl)-L-phenylalanine from phenylpyruvate and pyruvate by phenylalanine and opine dehydrogenases with NADH regeneration by formate dehydrogenase. PheDH, phenylalanine dehydrogenase from *Bacillus sphaericus* R79a; ODH, opine dehydrogenase from *Arthrobacter* sp. 1C; FDH, formate dehydrogenase from *Moraxella* sp. C-1.

stereochemistry. Among secondary amine dicarboxylic acid dehydrogenases, some differences found in the substrate specificity include octopine, lysopine, and nopaline dehydrogenases characteristically promoting reductive alkylation at the α-NH_2 of basic amino acids; ODH, alanopine, and strombine dehydrogenases acting on that of neutral amino acids; and saccharopine dehydrogenase and N^5-(L-1-carboxyethyl)-L-ornithine: $NADP^+$ oxidoreductase accommodating the ε-NH_2 of basic amino acids. Octopine, lysopine, alanopine, and strombine dehydrogenases, N^5-(L-1-carboxyethyl)-L-ornithine: $NADP^+$ oxidoreductase, and ODH accept pyruvate as an amino acceptor, whereas nopaline and saccharopine dehydrogenases act on α-ketoglutarate. A comparison of the structure of ODH with other amino acid dehydrogenases would provide insight to the catalytic, structural, and evolutionary relationships among them. To determine the high-dimensional structural elucidation of ODH, X-ray crystallographic studies are in progress (66,67) and should provide details of the structure of the catalytic domains of ODH. The findings might help in understanding the mechanism and evolutionary relationship among the amino acid dehydrogenases.

BIBLIOGRAPHY

1. E.W. Nester, M.P. Gordon, R.M. Amasino, and M.F. Yanofsky. *Annu. Rev. Plant Physiol.* **35**, 387–413 (1984).
2. J. Thompson and S.P.F. Miller, *Adv. Enzymol.* **64**, 317–399 (1991).
3. J. Thompson and J.A. Donkarsloot, *Annu. Rev. Biochem.* **61**, 517–557 (1992).
4. W.S. Chilton, J. Tempe, M. Matzke, and M.D. Chilton, *J. Bacteriol.* **158**, 650–658 (1984).
5. E. Szegedi, M. Czako, L. Otten, and C.S. Koncz, *Physiol. Mol. Plant Pathol.* **32**, 237–248 (1988).
6. C.C. Chang, R.K. Japaswal, C.M. Chen, and S.B. Gelvin, *J. Bacteriol.* **171**, 5922–5927 (1989).
7. Y. Dessaux, A. Petit, and J. Tempe, in D.P.S. Verma ed., *Molecular Signals in Plant–Microbe Communications*, CRC Press, Boca Raton, Fla., 1992, pp. 109–136.
8. W.S. Chilton, A.M. Stomp, V. Beringue, H. Bouzar, and V. Vaudequin-Dransart, *Phytochemistry* **40**, 619–628 (1995).
9. A. Petit and J. Tempe, in L. van Floten-Doting, G.P.S. Groot, and T.C. Hall eds., *Molecular Form and Function of the Plant Genome*, Plenum Press, New York, 1984, pp. 625–636.
10. J. Baldwin and A.M. Opie, *Comp. Biochem. Physiol. B* **61**, 85–92 (1978).
11. E. Hack and J.D. Kemp, *Biochem. Biophys. Res. Commun.* **78**, 785–791 (1977).
12. S. Hatanaka, S. Atsumi, K. Furukawa, and Y. Ishida. *Phytochemistry* **21**, 225–227 (1982).
13. N. Izumiya, R. Wade, M. Winitz, M.C. Otey, S.M. Birnbaum, R.J. Koegel, and J.P. Greenstein, *J. Am. Chem. Soc.* **79**, 652–658 (1957).
14. A. Menage and G. Morel, *Coll. R. Acad. Sci.* **259**, 4795–4796 (1964).
15. C. Lioret, *Bull. Soc. Fr. Physiol. Veg.* **2**, 76 (1956).
16. K. Goto, M. Waki, N. Mitsuyasu, Y. Kitajima, and N. Izumiya, *Bull. Chem. Soc. Jpn.* **55**, 261–265 (1982).
17. J.L. Firmin, I.M. Stewart, and K.E. Wilson, *Biochem. J.* **232**, 431–434 (1985).
18. M. Sato, Y. Sato, and Y. Tsuchiya, *Nippon Suisan Gakkaishi*, **43**, 1077–1079 (1977).
19. A.W. Sangster, S.E. Thomas, and N.L. Tingling, *Tetrahedron*, **31**, 1135–1137 (1975).
20. M. Sato, M. Takahara, N. Kanno, Y. Sato, and W.R. Ellington, *Comp. Biochem. Physiol. B*, **88**, 803–806 (1987).
21. M. Kuriyama, *Nature* **192**, 969 (1961).
22. A. Goldmann, D.W. Thomas, and G. Morel, *Coll. R. Acad. Sci.* **268**, 852–854 (1969).
23. J.L. Firmin and R.G. Fenwick, *Phytochemistry* **16**, 761–762 (1977).
24. A. Kjaer and P.O. Larsen, *Acta Chem. Scand.* **15**, 750–759 (1961).
25. S. Fushiya, M. Matsuda, S. Yamada, and S. Nozoe, *Tetrahedron* **52**, 877–886 (1996).
26. E. Hack and J.D. Kemp, *Plant Physiol.* **65**, 949–955 (1980).
27. K.B. Storey and P.R. Dando, *Comp. Biochem. Physiol.* **73B**, 521–528 (1982).
28. H. De Greve, P. Dhaese, J. Seurinck, M. Lemmers, M. van Montagu, and J. Schell, *J. Mol. Appl. Genet.* **1**, 499–511 (1983), NBRF-PIR database accession number A00399.
29. R.F. Barker, K.B. Idler, D.V. Thompson, and J.D. Kemp, *Plant. Mol. Biol.* **2**, 335–350 (1983), NBRF-PIR database accession number S28683.
30. L.A. Otten, D. Vreugdenhil, and R.A. Schilperoort, *Biochim. Biophys. Acta* **485**, 268–277 (1977).
31. A. Goldmann, *Plant Sci. Lett.* **10**, 49–58 (1977).
32. D.W. Sutton, J.D. Kemp, and E. Hack, *Plant Physiol.* **62**, 363–367 (1978).
33. Y. Gafni and M.-D. Chilton, *Gene* **39**, 141–146 (1985).
34. C. Drevet, A.C.M. Brasileired, L. Jouanin, *Plant. Mol. Biol.* **25**, 83–89 (1994), NBRF-PIR database accession number S46512.
35. J.H.A. Fields and P.W. Hochachka, *Eur. J. Biochem.* **114**, 615–621 (1981).
36. B. Siegmund, M. Griesharker, M. Reitze, and E. Zebe, *Comp. Biochem. Physiol.* **82**, 337–345 (1985).
37. P.R. Dando, *Biochem. Soc. Trans.* **9**, 297–298 (1981).
38. N. Kanno, M. Sato, E. Nagashita, and Y. Sato, *Comp. Biochem. Physiol.* **114B**, 409–416 (1996).
39. H. Misono and K. Soda, *J. Biol. Chem.* **255**, 10599–10605 (1980).
40. J.-W. Xuan, P. Fournier, N. Declerk, M. Chasles, and C. Gaillardin, *Mol. Cell. Biol.* **10**, 4795–4806 (1990), NBRF-PIR database accession number A36467.
41. J.A. Donkersloot and J. Thompson, *J. Biol. Chem.* **21**, 12226–12234 (1995), NBRF-PIR database accession number A57499.
42. Y. Asano, A. Nakazawa, and K. Endo, *J. Biol. Chem.* **262**, 10346–10354 (1987).
43. Y. Asano, A. Nakazawa, K. Endo, Y. Hibino, M. Ohmori, N. Numao, and K. Kondo, *Eur. J. Biochem.* **168**, 153–159 (1987).
44. N. Okazaki, Y. Hibino, Y. Asano, M. Ohmori, N. Numao, and K. Kondo, *Gene* **63**, 337–341 (1988), Swiss-Plot database accession number P23307.
45. A. Yamada, T. Dairi, Y. Ohno, X.-L. Huang, and Y. Asano, *Biosci. Biotechnol. Biochem.* **59**, 1994–1995 (1995), NBRF-PIR database accession number JC4328.
46. Y. Asano, A. Yamada, Y. Kato, K. Yamaguchi, Y. Hibino, K. Hirai, and K. Kondo, *J. Org. Chem.* **55**, 5567–5571 (1990).

47. Y. Kato, K. Fukumoto, and Y. Asano, *Appl. Microbiol. Biotechnol.* **39**, 301–304 (1993).
48. G. Tremblay, R. Gagliardo, W.S. Chilton, and P. Dion, *Appl. Environ. Microbiol.* **53**, 1519–1524 (1987).
49. C.J. Beauchamp, J.W. Kloepper, R. Lifshitz, P. Dion, and H. Antoun, *Can. J. Microbiol.* **37**, 158–164 (1990).
50. M.F. Jubier, *FEBS Lett.* **28**, 129–132 (1972).
51. Y. Asano, K. Yamaguchi, and K. Kondo, *J. Bacteriol.* **171**, 4466–4471 (1989).
52. B.R.B.-D. Guzman, Y. Asano, and H. Yamada, *Annu. Rep. Int. Center Coop. Res. Biotechnol. Jpn.* **14**, 167–178 (1991)
53. Y. Kato, H. Yamada, and Y. Asano, *J. Mol. Catal. B: Enzymatic*, **1**, 151–160 (1996).
54. T. Dairi and Y. Asano, *Appl. Environ. Microbiol.* **61**, 3196–3171 (1995).
55. A. Schrell and J. Schröder. *Biochim. Biophys. Acta* **1174**, 303–304 (1993), NBRF-PIR database accession number S38355, PIR protein sequence database, National Biomedical Research Foundation, Washington, URL: *http://www.psc.edu/general/software/packages/nbrf-pir/nbrf.html*.
56. L. Otten and P. Ruffray, NBRF-PIR database accession number S41895, PIR protein sequence database, National Biomedical Research Foundation, Washington, URL: *http://www.psc.edu/general/software/packages/nbrf-pir/nbrf.html.*
57. K.L. Britton, P.J. Barker, P.C. Engel, D.W. Rice, and T.J. Stillman, *J. Mol. Biol.* **234**, 938–945 (1993).
58. D. Garmyn, T. Ferain, N. Bernard, P. Hols, and J. Delcour, *Appl. Environ. Microbiol.* **61**, 266–272 (1995).
59. A.R. Clarke, T. Atkinson, and J.J. Holbrook, *Trends Biochem. Sci.* **14**, 101–105 (1989).
60. M.T. Wu, A.W. Douglas, D.L. Ondeyka, L.G. Payne, T.L. Ikeler, H. Joshua, and A.A. Patchett, *J. Pharm. Sci.* **74**, 352–354 (1985).
61. T.J. Blacklock, R.F. Shuman, J.W. Butcher, W.E. Shearin, Jr., J. Budavari, and V.J. Grenda, *J. Org. Chem.* **53**, 836–844 (1988).
62. G. Iwasaki, R. Kimura, N. Numao, and K. Kondo, *Chem. Lett.* 1691–1694 (1988).
63. G. Iwasaki, R. Kimura, N. Numao, and K. Kondo, *Chem. Pharm. Bull.* **37**, 280–283 (1989).
64. F. Effenberger and U. Burkard, *Liebigs Ann. Chem.* 334–358 (1986).
65. Y. Asano, T. Sekigawa, M. Inukai, and A. Nakazawa, *J. Bacteriol.* **170**, 3189–3193 (1988).
66. K.L. Britton, H.F. Rogers, Y. Asano, T. Dairi, Y. Kato, T.J. Stillman, and D.W. Rice, *Acta Cryst.* **D54**, 124–126 (1998)
67. K.L. Britton, Y. Asano, and D.W. Rice, *Nat. Struc. Biol.* **5**, 593–601 (1998).

OPTICAL RESOLUTION, BIOCATALYSIS

TETSUO OMATA
SHUJI SENDA
YASUO NINOMIYA
Nitto Denko Company
Osaka, Japan

KEY WORDS

Asymmetric microbial degradation
Backer's yeast
Enzymatic hydrolysis
Enzymatic transesterification
E-value
Immobilized microorganisms
Lipase-catalyzing esterification
Microbial oxidation
Prelog's rule
ρ-Ketoesters

OUTLINE

INTRODUCTION

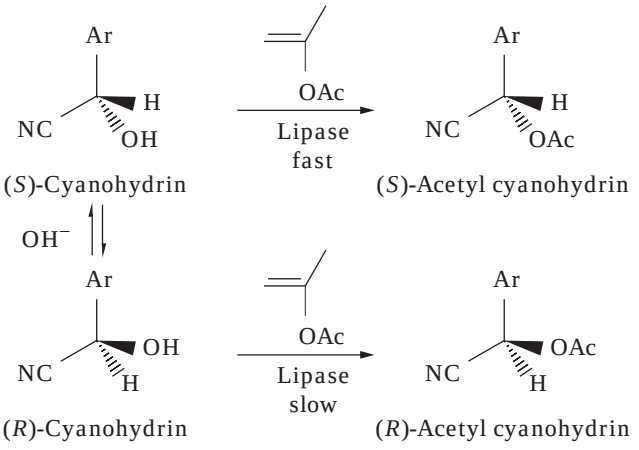
Enzyme proteins are produced by living organisms as biochemical catalysts. A typical feature of biocatalysts is the ability to catalyze stereo- and regiospecific reactions under mild conditions (e.g., neutral pH, normal temperature, and pressure).

Increasingly, enzymes are used in organic synthesis, especially for optical resolution of racemic organic compounds. This article provides an overview of some of the newer optical resolution reactions that use enzymes and microbial cells.

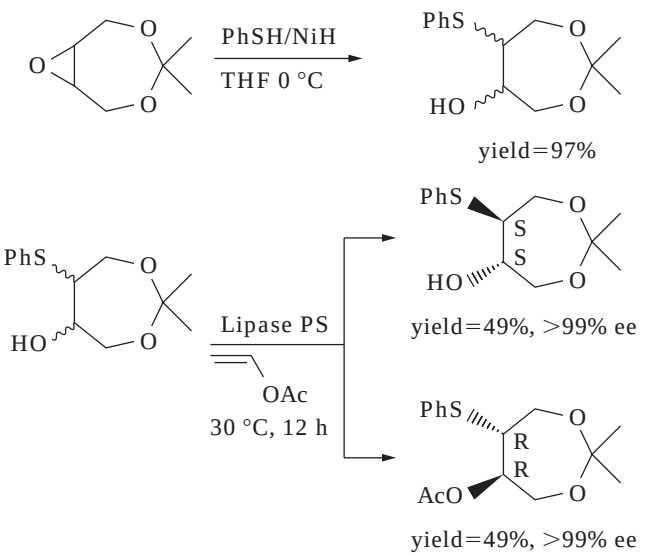
OPTICAL RESOLUTION WITH LIPASE-CATALYZED TRANSESTERIFICATION

Transesterifications catalyzed by lipases in organic solvent are most widely applied for optical resolution of racemic alcohols. Various kinds of racemic alcohols have been successfully resolved by lipase-catalyzed enantioselective transesterification reactions (1,2). Ward reported several examples of optical resolution accompanying the epimerization or racemization of starting materials at high conversion rates (3). Ward further introduced the enzyme-catalyzed acylating reaction of a racemic cyanohydrin under racemization of the residual enantiomer by a basic anion with extremely high conversion rates (Scheme 1).

A method of kinetic resolution of *cis*-1-diethylphosphomethyl-2-hydroxymethyl cyclohexane **1** by lipase has been developed (Scheme 2 and Table 1). The transesterification of (±)-**1** with vinyl acetate in the presence of lipase AK without solvent resulted in (+)-**1** with 99% ee and 41% yield and the corresponding acetate with 93% ee and 35% yield (the E value as reported by Sih was 152) (4).



Scheme 1. Lipase-catalyzed asymmetric acylation of racemic cyanohydrin.



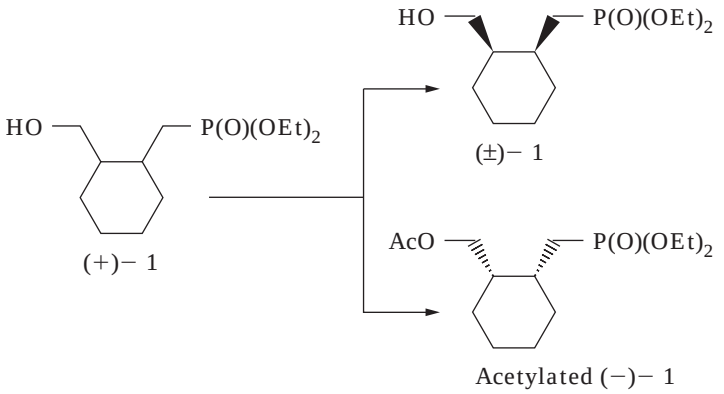
Scheme 3. *meso*-Mediated preparation of optically pure four-carbon triols with transesterification by lipase PS.

Enantioselective synthesis of both enantiomers of 1-arylethanols with a condensed aromatic ring has been achieved through acetylation of the racemic alcohols with vinyl acetate in the presence of lipase from *Pseudomonas aeruginosa* (Toyobo, LIP). The lipase catalyzed the reaction of (±)-(3-phenanthrenyl)ethanol with vinyl acetate at high enantioselectivity (99% ee and conversion; 49% of (±)-substrate). The remaining alcohol had a high optical purity (94% ee). The E value was 100 as determined by Sih's method (5).

Yamada et al. (6) reported an expedient route to optically pure four-carbon diols and triols using a *meso*-symmetric four-carbon precursor and lipase-mediated kinetic resolution as a key step in the transesterification reaction with vinyl acetate in an organic solvent (Scheme 3).

Tanaka et al. (7) reported a facilitated route to optically pure 3-*endo*-hydroxy dicyclopentadiene via lipase-mediated kinetic resolution in a transesterification reaction with vinyl acetate in an organic solvent (Scheme 4).

Chromobacterium viscosum lipase (CVL) catalyzes selectively the acylation of the C-5 hydroxyl of the three stereoisomeric vitamin D A-ring precursors **1a**, **2a**, and **2b** (Fig. 1). It stereoselectively acylated the hydroxyl group at the carbon of the 3-position of the fourth stereoisomer, **1b**,

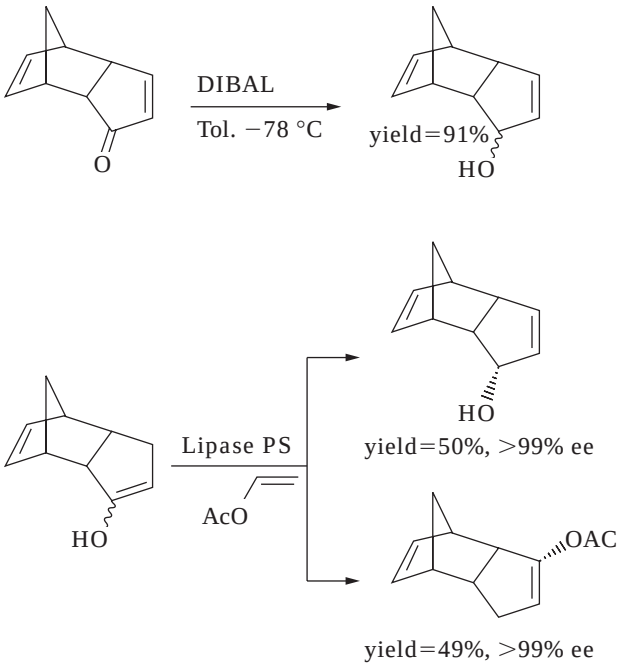


Scheme 2. Transesterification of phosphono alcohol 1 catalyzed lipases.

Table 1. Transesterification of (+)-1 Catalyzed Lipases

	Condition		Alcohol (+)1		Acetyl-(−)1		
Lipase mg/100 mg of 1	solv.	time (h)	Yield (%)	ee (%)	Yield (%)	ee (%)	E value
PS (200)	THF	15	38	60	55	36	3.7
PS (100)	THF	15	60	62	35	97	62
PS (100)	*t*-BuOMe	12	38	14	25	14	1.3
PS (50)	None	52	42	98	54	64	20
AK (50)	None	6.5	41	>99	35	93	152

Note: All reactions were carried out at 37 °C. PS = *Pseudomonas cepacia*, Amano; AK = *Pseudomonas fluoresence*, Amano.



Scheme 4. Preparation route to optically pure 3-*endo*-hydroxydicyclopentadiene.

with vinyl acetate under the same conditions in organic solvent (95% isolated yield, 90% ee) (8).

Izumi et al. (9) reported that *trans*-2,3-dihydro-3-hydroxy-2-phenyl-4H-1-benzopyran-4-one (*trans*-flavanon-3-ol) was resolved to acetate of (2*S*,3*S*)-(−)-*trans*-flavanon-3-ol with 96% ee and (2*R*,3*R*)-(−)-*trans*-flavanon-3-ol by an enzymatic transesterification with vinyl acetate in the presence of *P. capacia* lipase (Scheme 5).

Lundth et al. (10) proposed that enzyme-catalyzed transesterification of racemic alcohol with vinyl acetate by lipase from *P. fluorescens* (PFL) was a reversible process, although it is commonly referred to as irreversible. When racemic *β*-methyl-(2-thiophene)propanol **1** was resolved via a transesterification reaction catalyzed by lipase from

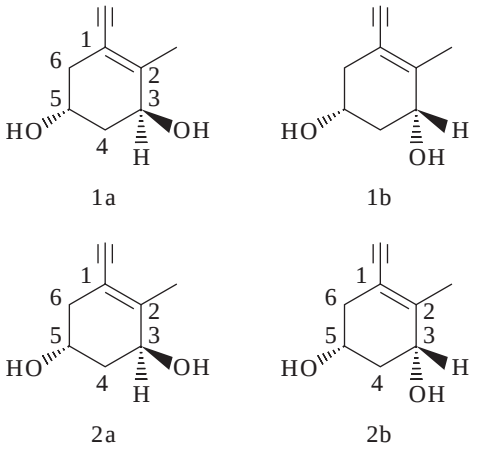
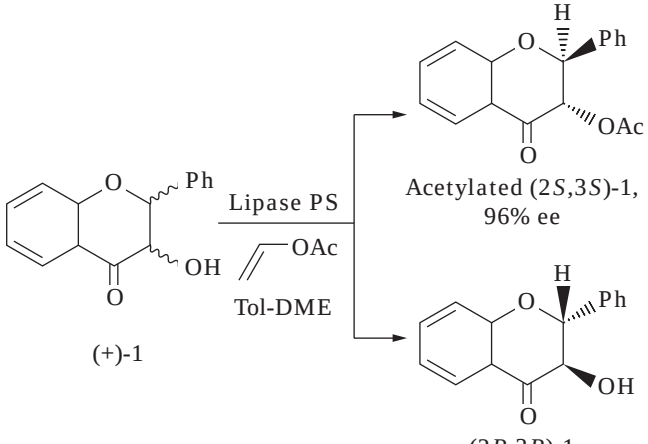


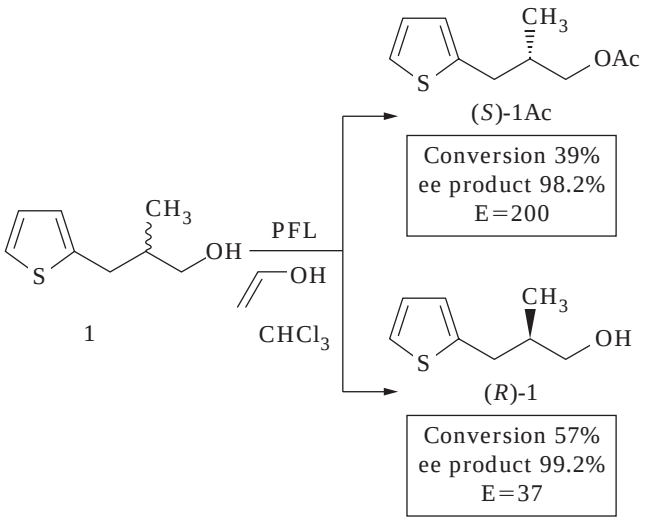
Figure 1. Structure of precursors of vitamin D A-ring.

PFL using an excess vinyl acetate in chloroform, the enantioselectivity was a high E = 200 (conversion, 39%), as calculated from enantioexcess (ee) of (*S*)-ester, (*S*)-**1Ac**. However, based on the ee of the remaining (*R*)-alcohol (*R*)-**1**, assuming irreversibility, and when measured at a conversion rate of 57%, the value E = 37 was obtained. Similarly, when the 5-(1-ethoxyethoxy)-3-pentyn-2-ol **2** was resolved using immobilized lipase B from *Candida antarctica* (Novozym SP435) under similar conditions, the (*R*)-ester product (*R*)-**2Ac** was obtained at 22% conversion with E = 140, compared with 65% conversion and E = 15 for the remaining alcohol (*S*)-**2**. These results were interpreted as the consequence of a reversible process occurring in transesterifications of this type, which were referred as irreversible (Schemes 6 and 7).

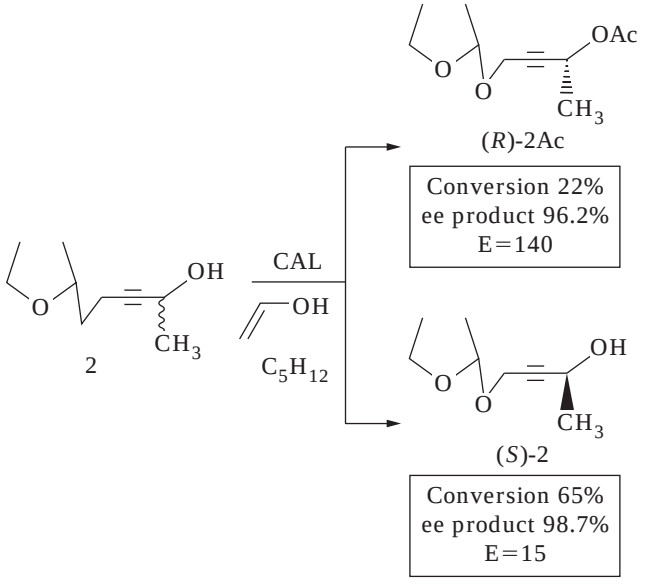
In the total synthesis of *endo*-brevicomin **1**, a pheromone of bark beetles such as *Dendrocutonus frontalis* and *Dryocoetes aoutograghus*, Kim et al. (11) examined *anti*-trityloxy-3,4-diol **2** as a substrate with lipase PS (LPS) from *P. capacia* in transesterification. They found that it was optically resolved by LPS with high enantioselectivity, thus allowing preparation of each enantiomer in optically pure form (>98% ee). The synthesis of optically pure (+)-



Scheme 5. Enzymatic resolution of *trans*-flavanon-3-ol by lipase.



Scheme 6. Transesterification of 2-methyl-3-(2′-thienyl)-1-propanol.



Scheme 7. Transesterification of 5-(1-ethoxyethoxy)-3-pentyn-2-ol by lipase.

endo-brevicomin **1** was achieved from the resolved (−)-2 by transesterification with vinyl acetate using LPS (Scheme 8).

Yamane (12) reported the total synthesis of optically pure (+)-Mitsugashiwa lactone by lipase-catalyzed transesterification under the presence of vinyl acetate in organic solvents (Scheme 9).

OPTICAL RESOLUTION WITH LIPASE-CATALYZING HYDROLYSIS

The stereoselective hydrolysis of 2-(3′-methylphenoxy)propionyl and 2-(2,3-dimethylphenoxy)propionyl esters by carboxy esterase NP was estimated under different experimental conditions to verify whether their optical resolution could be performed using this procedure. The reaction was performed on methyl, ethyl, isopropyl, isobutyl, and benzyl esters of 2-(3-methylphenoxy)propionic and 2-(2,3-dimethylphenoxy)propionic acids in aqueous media. The best substrate for the biocatalytic resolution was the *S* enantiomer for the monomethyl-substituted compounds and the *R* enantiomer for the dimethyl derivative. The extent of 2-(3-methylphenoxy) propionic ester hydrolysis was greater than the 2-(2,3-dimethyl phenoxy) propionic esters. The hydrolysis of isobutyl and benzyl derivatives suffered from poor enantioselectivity and long reaction time; this may be attributable to steric hindrance (caused by the acidic moiety) of the substrates' interaction with the active sites of the enzymes. Changes in pH, temperature, and substrate/enzyme ratio were investigated.

Optimum conditions for these biocatalytic resolutions were obtained when methyl 2-(3-methylphenoxy)-propionate and isopropyl 2-(2,3-dimethyl phenoxy)propionate were hydrolyzed at pH 7.2, 23 °C, and substrate/enzyme ratio 50. The scaling-up to preparative amounts of the enzymatic hydrolysis of these compounds succeeded in producing satisfactory quantities of the corresponding acids and residual esters with a good-enough optical yield (13). An enzymatic method for preparation of the enantiomers of chiral diols **1** and their monoacetate **2** with 100% enantiomeric purity was established using lipase-catalyzed esterification and/or hydrolysis.

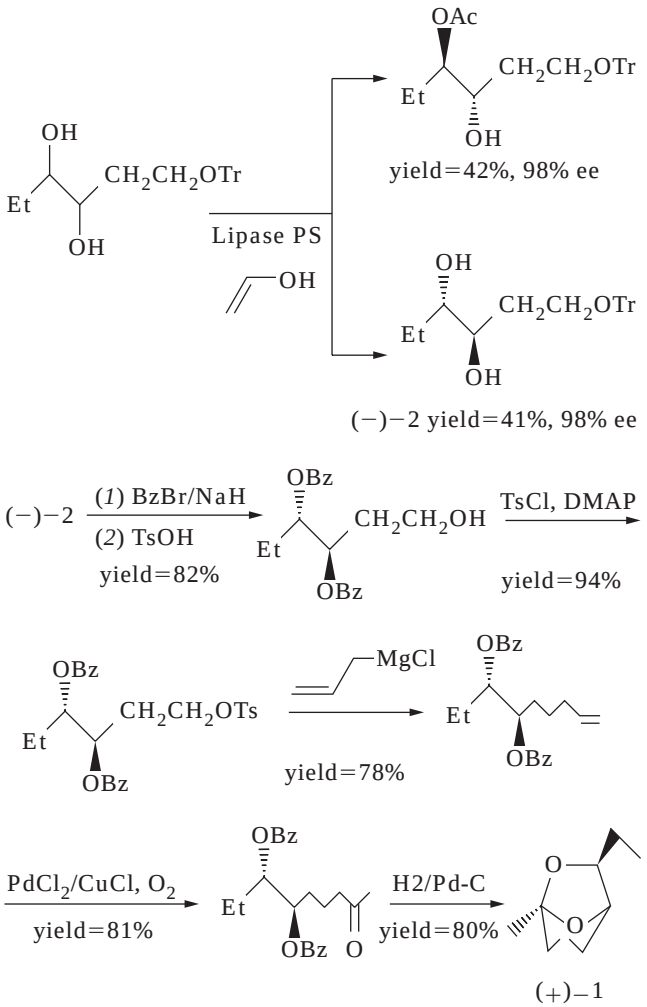
The (*R*)-enantiomers were more susceptible in both acetylation of (±)-**1** and hydrolysis of (±)-**2** catalyzed by lipase OF from *Candida cylindracea* (14) (Scheme 10).

Tan et al. (15) envisioned that thioesters of certain carboxylic acids having the chiral center at the α-carbon might be susceptible to sufficiently acidic deprotonation, resulting in racemization under conditions of enzymatic resolution. Although thioesters have been used as substrates in enzymatic resolution, their α-proton acidity has apparently never been exploited for resolution coupled with substrate racemization. To demonstrate that thioester could be used for racemization coupled with enzymatic resolution, ethyl thioester of α-(phenylthio)propionate **1** was used (Scheme 11).

The corresponding (*R*)-acid **2** was subjected to diastereoselective oxidation for conversion to sulfoxide in the synthesis of a chiral vinyl sulfoxide. Because of the insolubility in water of **1** and most other substrates of interest, it was decided to use a lipase as catalyst in a biphasic system (water and water-immiscible organic solvent) while inducing racemization of the substrate in the organic phase.

A novel enzymatic process for the optical resolution of racemic pantolactone through the stereospecific hydrolysis of D-pantolactone by lactonohydrolase of *Fusarium oxysporum* was reported (16). *F. oxysporum* cells were found to catalyze the stereoselective hydrolysis of the D-enantiomer of racemic pantolactone. With 135g/L of DL-pantolactone as the substrate, 41% of the substrate was hydrolyzed to produce pantoic acid with optical purity of 90% enantiomer excess (for D-pantoic acid) (Scheme 12).

Bommarius et al. (17) reported the synthesis of enantiometrically pure *tert*-leucine, which was prepared using



Scheme 8. Total synthesis of (+)-*endo*-brevicomin via transesterification reaction by lipase PS.

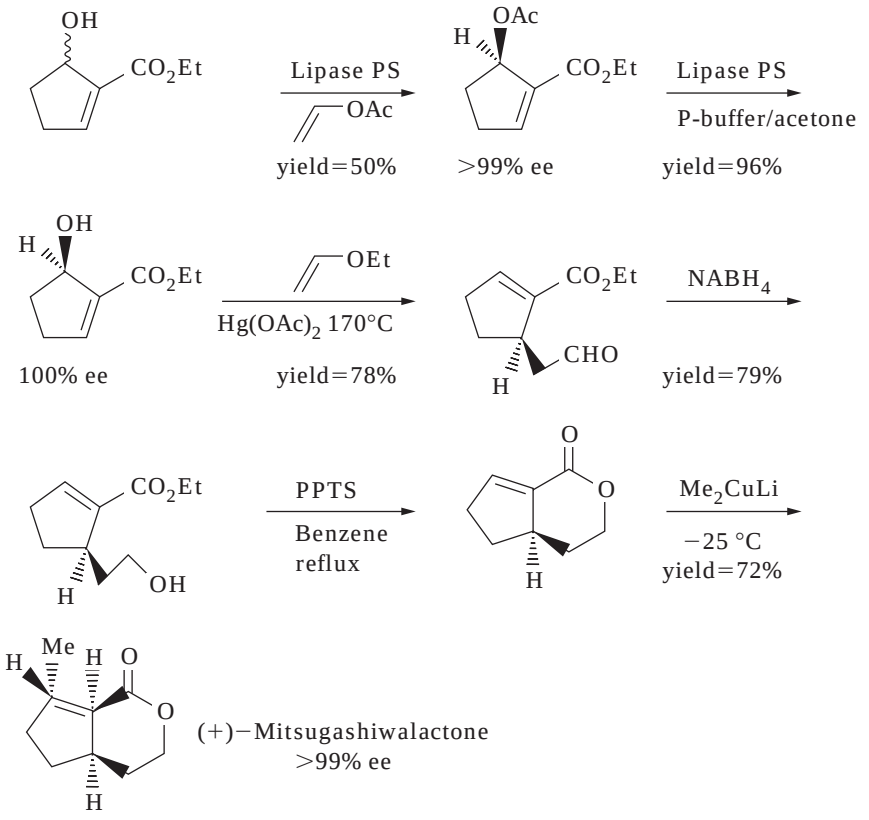
various methods using penicillin-G acylase, hydantoinase, and dehydrogenase.

Owing to its bulky, inflexible, and hydrophobic *tert*-butyl side chain, the use of *tert*-leucine has increased in templates and catalysts in asymmetric synthesis as well as in medicinal peptide compounds. (*S*)-*tert*-Leucine, available through a large-scale enzymatic reductive amination process, has been incorporated into a variety of anti-AIDS and anticancer compounds.

Optically pure inden-1-ol **1** has potential utility as a key starting material for the preparation of HIV-1 protease inhibitors such as L-735,524 and L-754,394. Synthesis of optically active inden-1-ol **1**, however, has been little studied, and only one enantiotopical preparation has been reported so far. Takahashi et al. (18) first examined the kinetic acetylation of racemic inden-1-ol, (±)-**1** with vinyl acetate in organic solvent in the presence of a lipase. However, no practically satisfactory conditions could be established, although partial optical resolution was observed using lipase PS-on-Celite (*Pseudomonas* sp., Amano) (22% ee). They next examined the kinetic deacetylation of racemic 1-acetoxyinden (±)-**2** in a phosphate buffer-acetone solution in the presence of lipases. Of the enzymes tested, lipase PS (*Pseudomonas* sp., Amano) gave the best result. From 10 mg/mmol of the racemic acetate (±)-**2**, 94% ee of (*R*)-inden-1-ol (*R*)-**1** was obtained in 46% yield with recovery of optically pure (*S*)-1-acetoxyinden (*S*)-**2** in 45% yield after 48 h (Scheme 13).

Ibuprofen is a nonsteroidal anti-inflammatory drug with annual sales of about $100 million (U.S.) in 1994. Sales are expected to reach $600 million by the year 2000. The physiological performance of the *S* form of ibuprofen is 28-fold higher than that of the *R* form. The price of the *S* form is three times that of its racemic mixture. Therefore, pharmaceutical companies have devoted much effort to developing production of the *S* form. Ibuprofen can be optically resolved by chemical and enzymatic methods and asymmetric synthesis. Many enzymatic resolutions are reported with lipase-catalyzed hydrolysis or transesterifications. Tai et al. (19) reported that enzymatic hydrolysis of ibuprofen esters produced 99% optically pure (*S*)-(+)-ibuprofen by immobilized lipase MY entrapped in chitosan-*N*-(3-dimethyl aminopropyl)-*N′*-ethylcarbodiimide-glutaraldehyde.

Sulfoxides are valuable chiral auxiliaries as precursors because they carry out carbon–carbon bond formation. As a route to enantiomerically pure sulfoxides, Serreqi et al. (20) examined hydrolase-catalyzed kinetic resolution by



Scheme 9. Total synthesis of Mitsugashiwalactone via lipase catalyzing transesterification with vinyl acetate.

hydrolysis of a pendant acetoxy group. Screening of hydrolases for enantioselective hydrolysis of the acetate in 2-(methylsulfinyl) phenyl acetate (**1**) identified cholesterol esterase (CE) as the most enantioselective enzyme. The enantiomeric ratio, E, ranged from 10 to 25, favoring the (*R*)-configuration at sulfur. Competing chemical hydrolysis of **1** caused a large range in measured E values. A small-scale resolution of (±)-**1** yielded (*S*)-**1** with 99% ee (11% yield) after recrystallization. Changing the methyl substituent to phenyl or *n*-butyl did not significantly change enantioselectivity (E = 10 and 15, respectively), but changing it to a chloromethyl substituent did lower enantioselectivity slightly (E = 5). Changing the phenyl acetate to a naphthyl acetate (2-(phenylsufinyl) acetate versus 1-(phenylsulfinyl)-2-naphthyl acetate) increased the enantioselectivity from 10 to 19 (E value). In all cases, CE favored the (*R*)-sulfoxide. To aid the design of new resolutions with CE, Serreqi et al. proposed an empirical rule that accounts for the observed enantiopreference of CE toward these five compounds of sulfoxides and 15 compounds of other chiral aryl acetates. This empirical rule uses both the size of the substituents and their conformational preferences to predict which enantiomer reacts faster (Fig. 2).

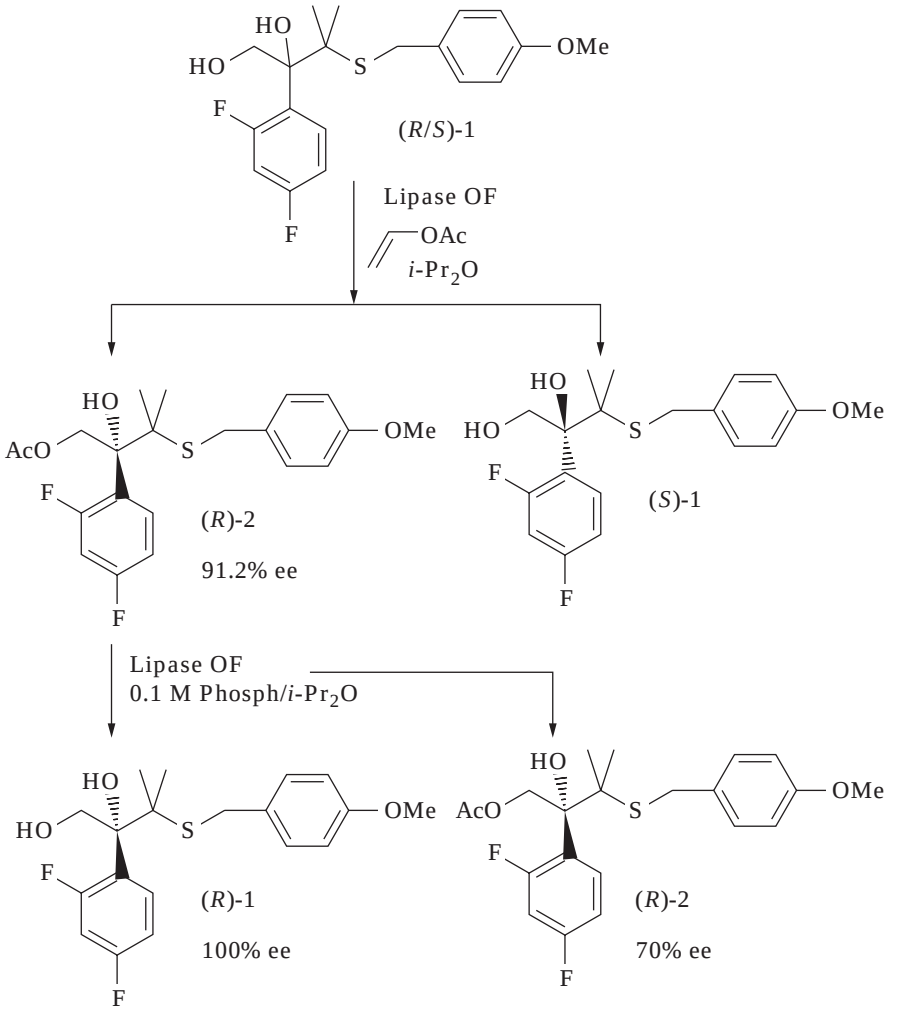
(±)-1-(2-Thienyl)propyl acetate (±)-**1** was optically resolved by *P. capacia* lipase (PCL)–catalyzed hydrolysis to afford (*S*)-(−)-1-(2-thienyl)propyl acetate (*S*)-**1** with >99% ee (21). (*S*)-(−)-1-(2-Thienyl)propyl acetate (*S*)-**1** thus obtained was transformed to (*S*)-(+)-3-octanol (*S*)-**4**, the alarm pheromone of ants, *Crematoster castanea* and *C. liengmei* (Scheme 14).

4-Aryl-1,4-dihydropyridines are novel and potent calcium antagonists; their derivatives have been widely investigated from pharmacological points of view, and some of them have already been used therapeutically as antihypertensive drugs. These agents are known to exhibit their vasodilative activities through affinity with voltage-dependent calcium channels. Different substituents in their compounds lead to chiral derivatives possessing an asymmetric carbon at the 4-position, and the enantiomers have been reported to show different biological activities.

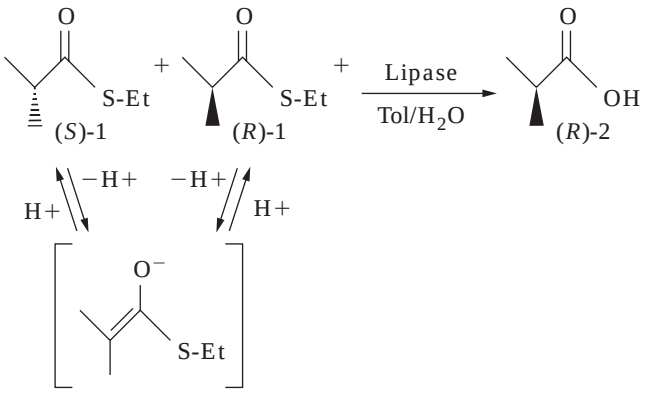
Ebiike et al. (22) synthesized a homochiral 4-aryl-1,4-dihydropyridine-5-phosphate (NZ105) from racemic materials. Enantioselective hydrolysis by lipase AH (*Pseudomonas* sp.) catalyst proceeded smoothly with use of propionyloxymethyl (PROM) ester derivative in organic solvents. They succeeded in converting an optically active carboxylic acid into an optically active calcium antagonist medicine (Scheme 15).

The hydrolysis of methyl D,L-3,3-difluorophenyl alanate **1a** and methyl D,L-3,3-difluoro-2-aminobutanoate **1b** and their *N*-acetyl derivatives **2a** and **2b** by subtilisin (protease) has been studied (23). All derivatives examined were enzymatically resolved to separable mixtures of the corresponding 3,3-difluoro-L-amino acids (**3a** and **3b**) or *N*-acetyl amino acids (**5a** and **5b**) and the unchanged 3,3-difluoro-D-amino acid (**4a** and **4b**) or *N*-acetyl amino acid (**6a** and **6b**). Acidic hydrolysis of methyl 3,3-difluoro-D-phenyl alanate (**4a**) or its *N*-acetyl derivative (**6a**) led to 3,3-difluoro-D-phenyl alanine (**8a**).

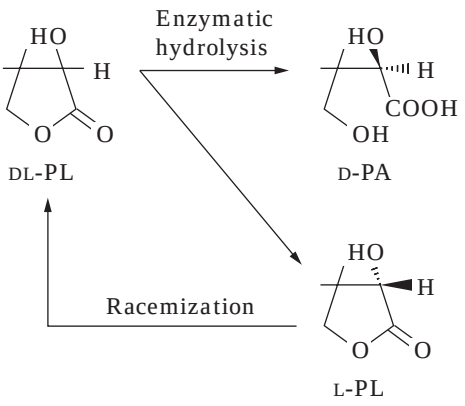
In the same manner, L- and D-2-amino-3,3-difluorobutanoic acids (**7b**) and (**8b**) were prepared starting from (**5b**)



Scheme 10. Enzymatic resolution of sterically crowded diols.



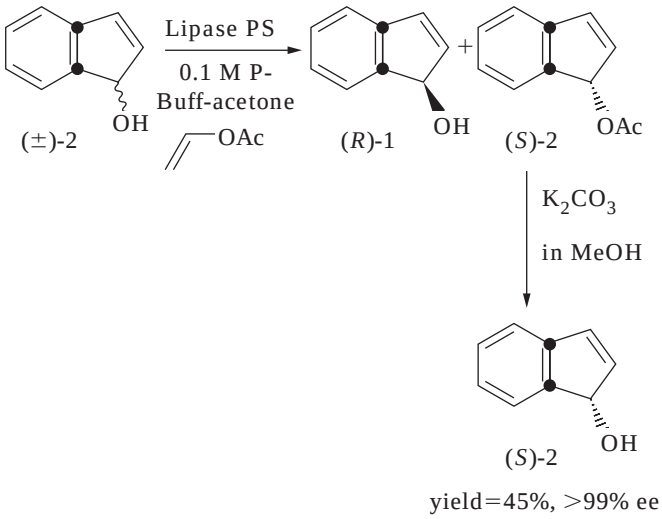
Scheme 11. Enzymatic resolution coupled with substrate racemization using a thioester substrate.



Scheme 12. Schematic presentation of the enzymatic resolution of DL-pantolactone.

and (**6b**). With these methods, unchanged methyl 3-3-difluoro-D-amino ester derivatives showed an ee of >90%, and L-amino acids were estimated to have an ee of >95% (Scheme 16).

Kinetic resolution of 2,2-disubstituted epoxides was accomplished with epoxide hydrolases from bacterial and fungal origin by using lyophilized whole microbial cells (24). In all cases investigated, biocatalytic hydrolysis was shown to occur with retention of configuration at the stereogenic center, leading to 1,2-diols and other results leading to epoxides. The selectivity of the reaction was depen-



Scheme 13. Lipase-mediated resolution of inden-1-ol.

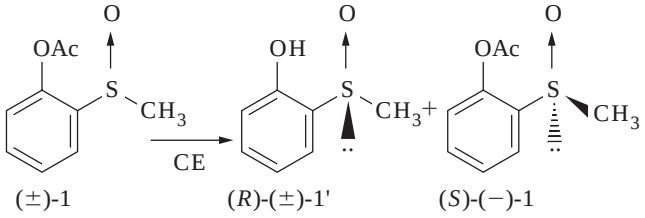


Figure 2. Kinetic resolution of sulfoxide with pendant acetoxy group using cholesterol esterase.

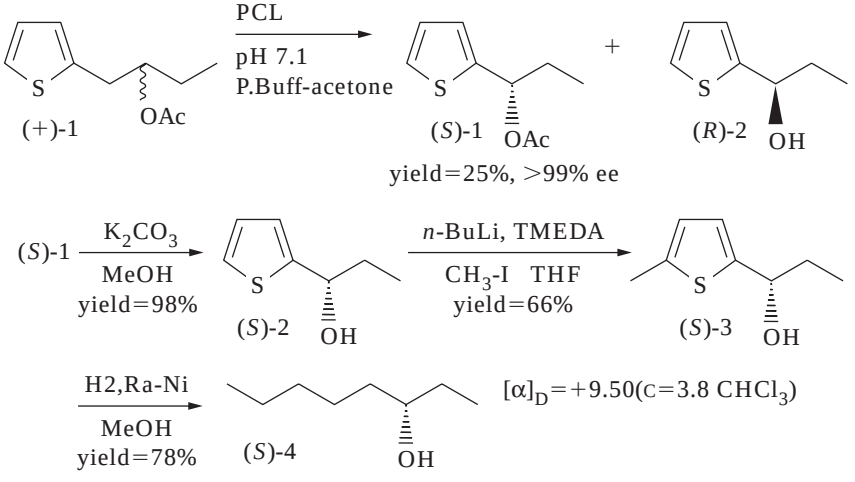
dent on the substrate structure and the strain used. Specifically, in the case of 2,2-disubstituted oxyrane (2-methyl-2-nonyl-1,2-oxyrane), (*S*)-(2-methyl)undecan-1,2-diol was obtained by lyophilized *Rhodococcus* sp. with high stereoselectivity (E >200) (Scheme 17).

Akita et al. (25) reported a highly stereoselective synthesis of a versatile chiral synthon possessing two stereogenic centers (2*S*,3*S*)-**1** (>99% ee) and successful synthesis of Nikkomycin B. (Scheme 18).

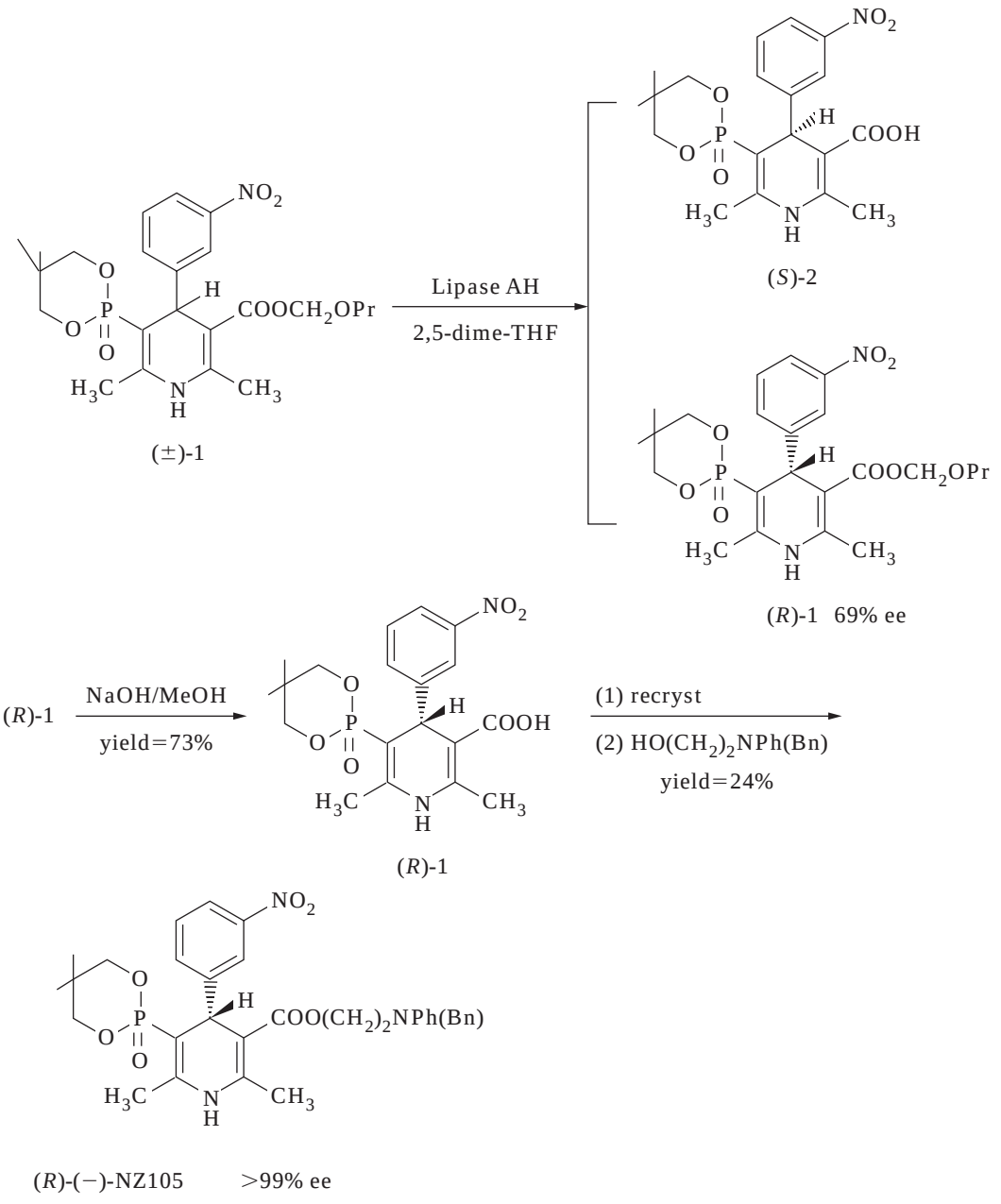
OPTICAL RESOLUTION BY LIPASE-CATALYZING ESTERIFICATION

The reaction for the resolution of (*R*,*S*)-ibuprofen was scaled-up to yield gram quantities of (*S*)-ibuprofen, which was accomplished through a two-stage enantioselective reaction, each stage catalyzed by Novozym 435 (26). In the first reaction, starting from 300 g of racemic ibuprofen, 88.9 g of enantio-enriched (*S*)-ibuprofen with 85% ee was produced. In the subsequent reaction, 76 g of the 85% ee materials were used to produce 38.4 g of (*S*)-ibuprofen with 97.5% ee (Scheme 19).

Synthesis of both enantiomers of fluorophoracantholide **1**, (9*S*)-10-fluorodecan-9-olide (9*S*)-(−)-**1** and (9*S*)-10-fluorodecan-9-olide (9*R*)-(+)-**1**, has been achieved through lipase-catalyzed enantioselective acetylation of methyl 10-fluoro-9-hydroxy decanoate (**2a**) with acetic anhydride in the presence of lipase amano PS (*P. capacia*) (27). The higher homologue, methyl 11-fluoro-10-hydroxyundecanoate (**2b**), has also been optically resolved. These enzymatic resolutions by acetylation of ω-fluorohydroxyalka-



Scheme 14. Chemicoenzymatic synthesis of alarm pheromone of ants, *Crematoster* sp.



Scheme 15. Lipase-catalyzed enantiosynthesis of a calcium antagonist for medication.

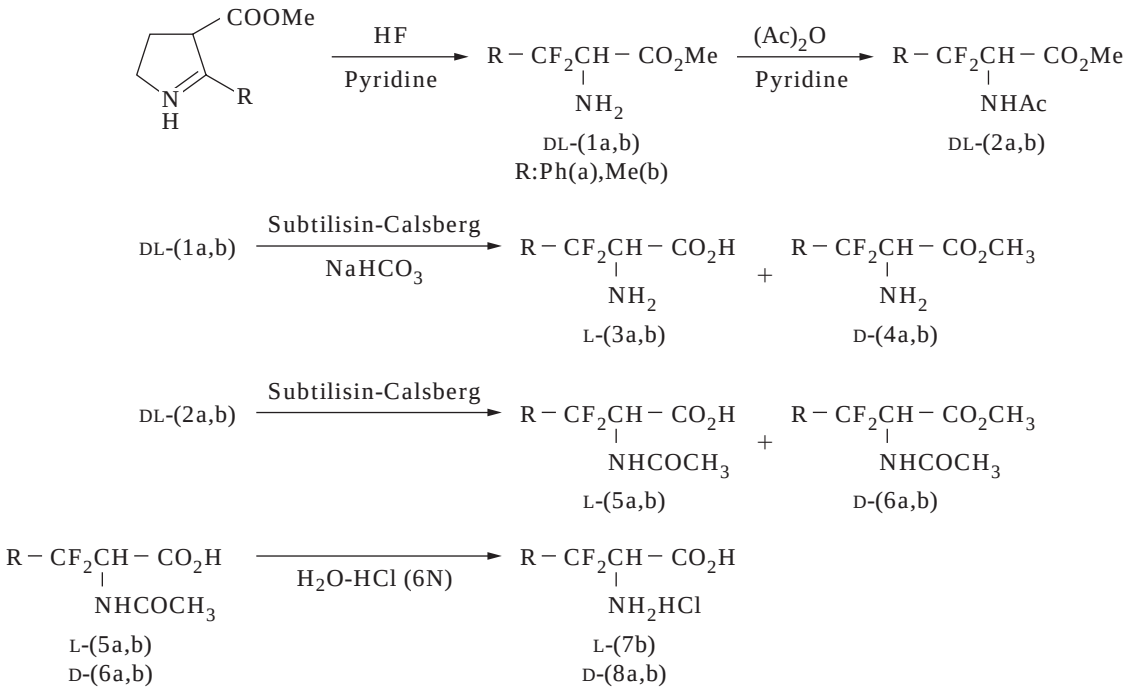
noic acid esters of long-chain fatty acids represent the first lipase-catalyzed enantioselective esterification of *β*-fluoroalcohol (Scheme 20).

Many adrenergic drugs are 2-amino-1-arylethanols possessing a stereogenic center with a secondary alcohol function. It is well known that the two enantiomers of a chiral drug often display different biological activities. Consequently, the enantiomeric purity of such compounds is very important. The drugs of (*R*)-enantiomers of amino ethanol and (*S*)-enantiomers of amino-2-propanol are generally associated with agonist or *β*-blocking activity because of their structural similarity to natural catecholamines.

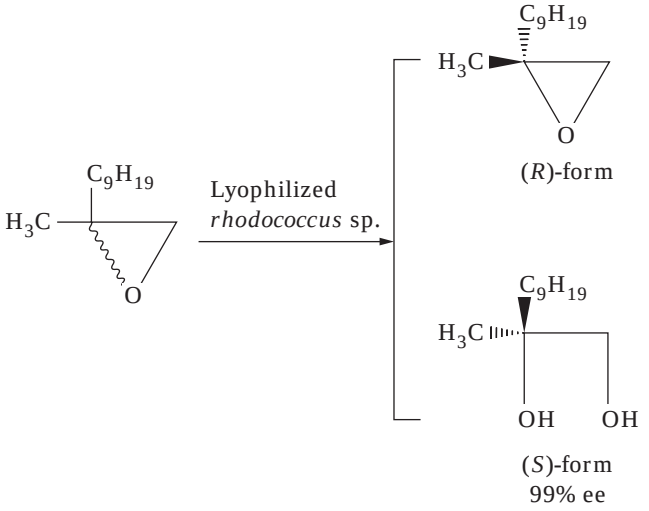
The enantiomers of 2-amino-1-(4-methoxy)phenylethanol were obtained at enantiomeric purity (99% ee) at 50% conversion by *P. capacia* lipase-catalyzed O-acylation of the corresponding amino alcohol (28) (Scheme 21).

The kinetic lactonization of *γ*-hydroxyesters by porcine pancreatic lipase (PPL) has been demonstrated by Sugai et al. (29). The reaction has been used in the synthesis of Japanese beetle pheromone with 78% ee in 29% conversion. Clearly, this chemistry has tremendous potential because of the importance of lactones in biologically active compounds. However, it is restricted to very low yields if high enantiomeric purity is desired.

Taylor et al. prepared the enantio-enriched *γ*-hydroxyesters from corresponding *γ*-ketoesters using Baker's yeast reduction (30). They attempted Baker's yeast reduction of **1** and obtained the corresponding *γ*-hydroxyesters **2** in a



Scheme 16. Enzymatic resolution of fluoro amino acids by subtilisin.



Scheme 17. Resolution of 2,2-disubstituted oxyrane using *Rhodococcus* sp. NCIMB 11216.

yield of 46% with 95% ee. This product **2** (predominantly (*S*)-**2**) was obtained by treating with PPL with 99.7% ee in 61% conversion (Scheme 22).

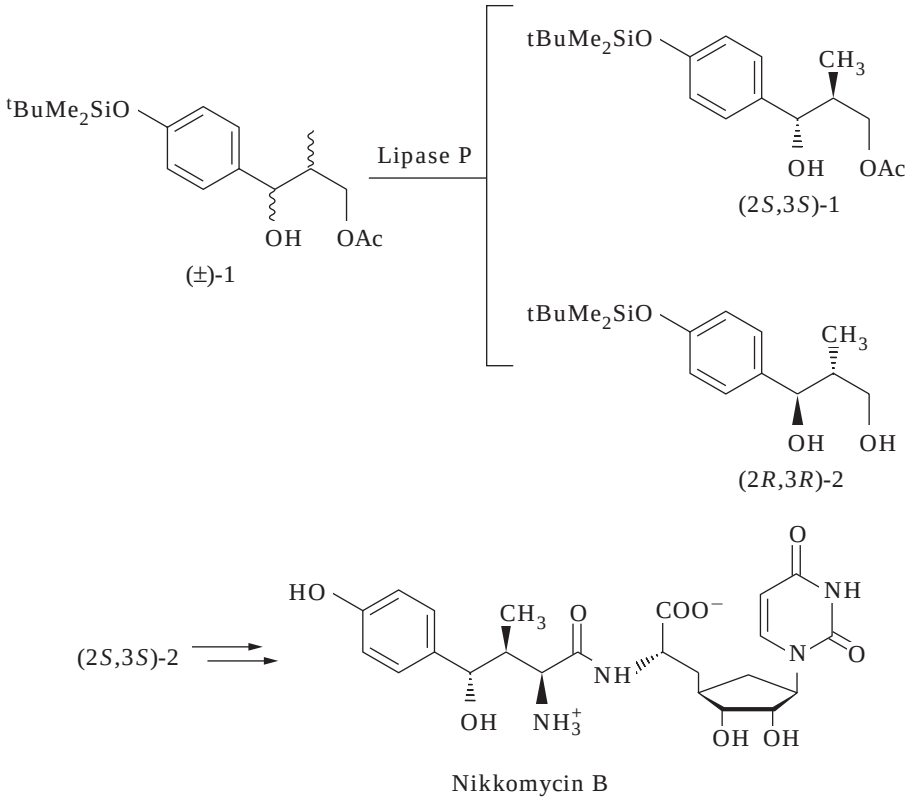
Taylor et al. also reported the high optical purity preparation of another two important lactones, key intermediates used in the synthesis of Japanese beetle pheromone and anti-AIDS drugs, using a complementary, two-stage enantioselective strategy. In this approach, an enzymatic kinetic resolution is performed to enrich the faster-reacting enantiomer in a starting material obtained by an earlier enantioselective step. The results show that conversions, yields, and ee can be dramatically increased by this complementary approach. The authors illustrate how modestly enantio-selective techniques can be combined with kinetic resolution procedures to yield optically pure compounds. The utility of many modestly enantio-selective techniques can be enhanced thereby.

OPTICAL RESOLUTION BY HYDANTOINASE-CATALYZING HYDROLYSIS

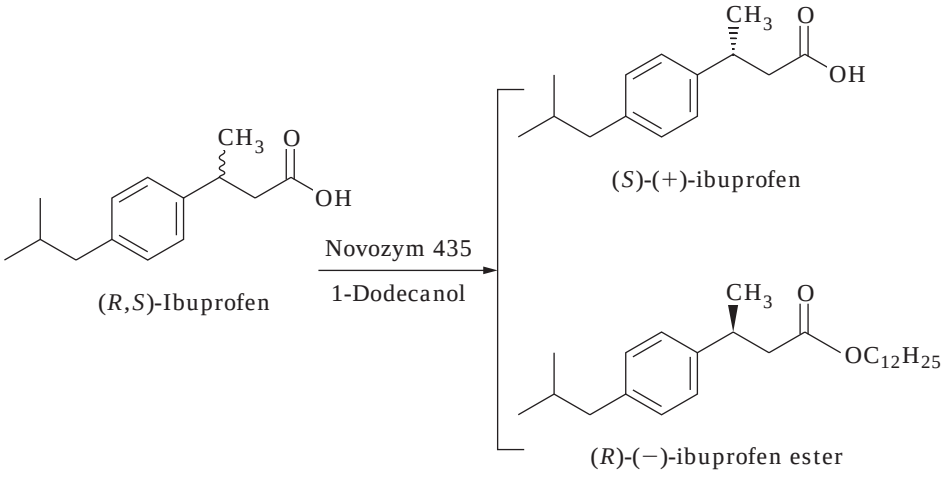
Racemic amino acid hydantoins can be converted into D-amino acids using a system consisting of a D-hydantoinase and a carbamoylase. D-Citrulline can be obtained from corresponding racemic hydantoin with 92% ee in 79% yield (31) (Scheme 23).

Agrobacterium tumefaciens strain 47C expresses an inducible D-hydantoinase that catalyzes the formation of optically pure *N*-carbamoyl D-amino acids from racemic hydantoin precursors. The D-hydantoinase was shown to be active and stable at higher temperature and higher pH values, thus affording favorable bioreaction conditions that result in racemization of DL-hydantoins to D-isomer. The optimal performance (reaction kinetics) was demonstrated at pH 10 and 70 °C. The enzyme exhibited a distinctive substrate specificity and was not activated further by any metal ions. *A. tumefaciens* hidantoinase was most active on 5,6-dihydro uracil and DL-5-methyl hydantoin but only slightly active on DL-benzyl hydantoin (32).

Kim and Kim (33) reported that D-hydantoinase and N-carbamoylase of the bacterium *Agrobacterium* sp I-671 were partially purified to about 90% purity on sodium dodecyl sulfate–polyacrylamide gel electrophoresis and that biochemical properties of the enzymes were characterized.



Scheme 18. A total synthesis of Nikkomycin B based on enzymatic resolution of a primary alcohol.



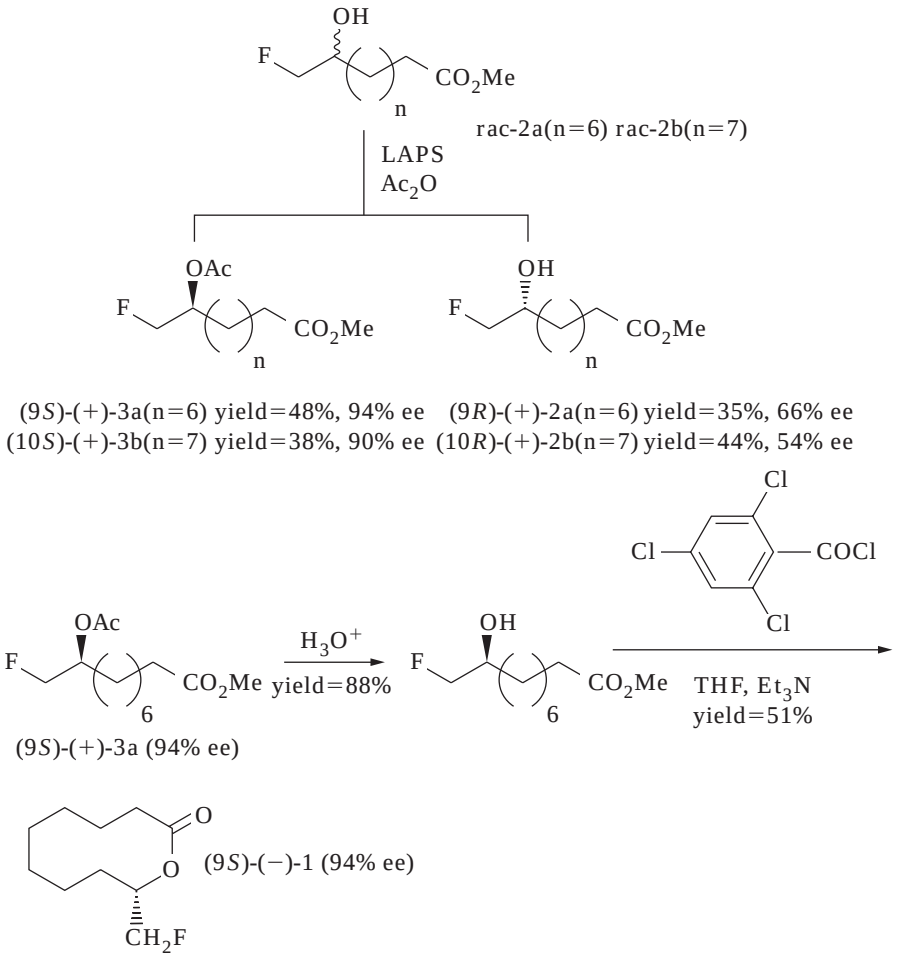
Scheme 19. The enantioselective reaction for the enzymatic resolution of (*R*,*S*)-ibuprofen.

The *N*-carbamoylase was found to be severely inhibited by ammonium ions coproduced with D-*p*-hydroxyphenylglycine (D-HPG). The conversion yield of D-HPG significantly increased on addition of specific adsorbents for ammonium ions. To determine the ratio of D-hydantoinase to *N*-carbamoylase, which best minimizes the accumulation of intermediate (*N*-carbamoyl-D-HPG) in the direct enzymatic production of D-HPG, a sequential reaction was numerically simulated. Simulation results coincided well with experimental data, and the optimal ratio of D-hydantoinase to *N*-carbamoylase was found to be about 1:3 by weight.

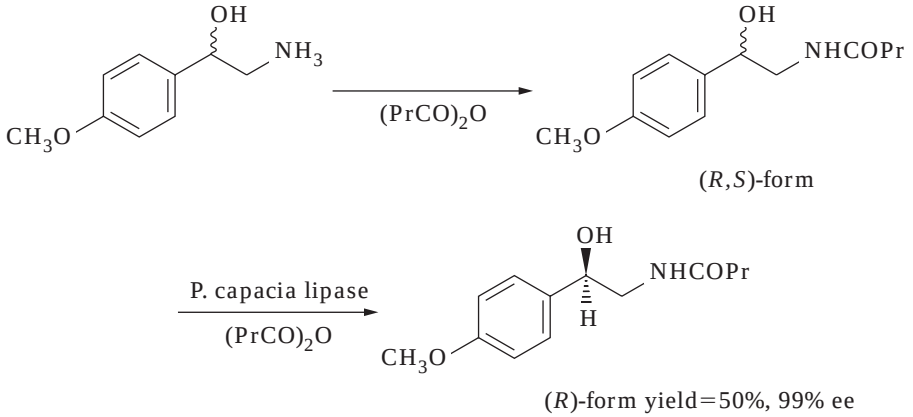
OPTICAL RESOLUTION BY OXIDOREDUCTASE-CATALYZING OXIDATION AND REDUCTION

Lipase-catalyzed kinetic resolution of secondary alcohols via hydrolysis, esterification, and transesterification has been widely used to obtain optical active compounds as described in this article.

Fantin et al. (34) succeeded in achieving the kinetic resolution of heteroaryl ethanols via oxidation both with baker's yeast and *Bacillus stearothermophilus*. These organisms were also used in the kinetic resolution of bicyclic octenol and heptenol with excellent results (35). Although



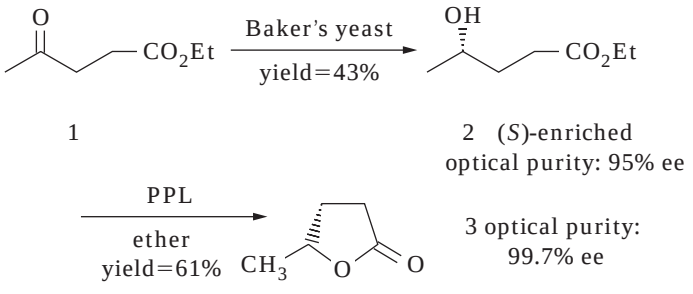
Scheme 20. Synthesis of (9*S*)-(−)-enantiomer of fluorophoracantholide.



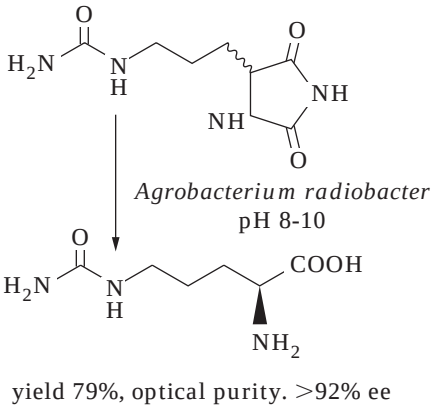
Scheme 21. Synthesis of optically pure precursor of adrenergic drugs by lipase-catalyzed acylation.

hydrolytic enzymes such as lipases and hydrolases are known to retain their activity in organic solvents, dehydrogenase and oxidoreductase are different because their catalytic reactions require the assistance of a stoichiometric amount of coenzymes, which has to be reproduced in order to elaborate the substrates catalytically in the cells. However, the organic solvents often cause serious damage on the hydrophobic cell membranes of the microbe, resulting in a leakage of the contents of the cell. To avoid this problem, the organic solvents were used with immobilized microorganisms because immobilization enhances their stability against denaturation by organic solvents.

Fantin et al. (36) reported that racemic 6-*endo*-bicyclo[3.2.0]hept-2-en-6-ol **1** was oxidized and kinetically resolved in heptane to give (1*S*,5*R*)-bicycloheptenone **3** in 49% conversion with 100% ee. Parallel experiments in water and with immobilized cells in heptane gave lower yields (41%) but equally high enantiomeric excess (99 to 100%



Scheme 22. The kinetic lactonization of *r*-hydroxyesters by porcine pancreatic lipase.



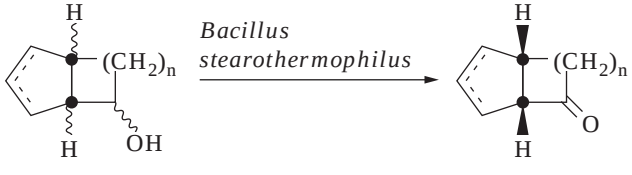
Scheme 23. D-Citrulline preparation by hydantoinase.

ee). Similar results were obtained in the oxidation of the racemic 2-*endo*-bicyclo[3.3.0]oct-7-en-2-ol **2**, but the reaction rate was higher. In fact, after only 4 h in heptane, (1*S*,5*S*)-bicyclo-octenone **4** was obtained in 48% yield and with 98% ee. In water, the yield was lower (37%). Oxidation with *B. stearothermophilus* in heptane is the first example of a reaction with bacteria carried out in organic solvents, and it is as efficient as water. The advantages of this methodology are that the procedure is simpler because the filtration of the cells and concentration of the solvents afford the crude reaction mixture. On the other hand, oxidation with the immobilized cells is very promising for use on a preparative scale (Scheme 24).

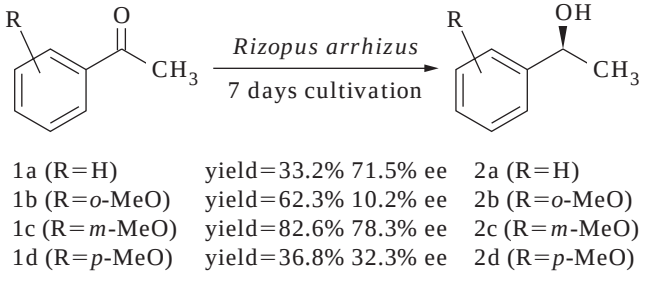
The fungus *Rhizopus arrhizus* has been used for reduction of acetophenone and its *o*-, *m*-, and *p*-methoxy derivatives **1a** to **1d** to the corresponding (*S*)-(−)-alcohols (37). To optimize the reaction time for microbial cell-catalyzed reduction, experiments were carried out with **1a** for different periods of time. Fermentation for 7 days was found to yield the best results. The 1-phenylalkanones **1a** to **1d** were therefore incubated with *R. arrhizus* in modified Czepak Dox medium for 7 days. At the end of fermentation, the transformation products 2a–d were isolated and characterized. All these alcohols were found to have negative specific rotations (Scheme 25).

The (1*R*,2*S*)- and (1*S*,2*R*)-2-methyl-2,3-epoxy-1-phenylpropanols were prepared from α-methylcinnamyl chloride by use of baker's yeast and lipase PS. Takeshita et al. (38) reported the synthesis of (1*R*,2*S*)-2-methyl-2,3-epoxy-1-phenylpropanol **3a** and (1*S*,2*R*)-2-methyl-2,3-epoxy-1-phenylpropanol **3b** starting from α-methylcinnamyl chloride **1** by use of baker's yeast and lipase PS (*Pseudomonas* sp., Amano). At first, when α-methylcinnamyl chloride **1** was fermented with baker's yeast for 36 h at 28 °C, rearrangement occurred to give 2-methyl-1-phenylallyl alcohol **2** in 68% yield. Two pairs of racemic epoxy alcohols, (**3a**, **3b**) and (**3c**, **3d**) (2:1), were separated by silica gel column chromatography (total yield 73%) after epoxidation of **2** with *m*-chloroperbenzoic acid. Enzymatic treatment of epoxy alcohols (**3a**, **3b**) with vinyl acetate in *t*-butyl methyl ether in the presence of lipase PS for 17 h or 24 h at 30 °C gave the optically active acetate **4** and alcohol **3b** in chemical yields of 31 and 46%, respectively. However, the same reaction for racemic epoxy alcohols (**3c**, **3d**) was very sluggish (under the above conditions, esterification proceeded only in 27% chemical yield in 3 weeks, and the optical yield was very low, 56% ee) (Scheme 26).

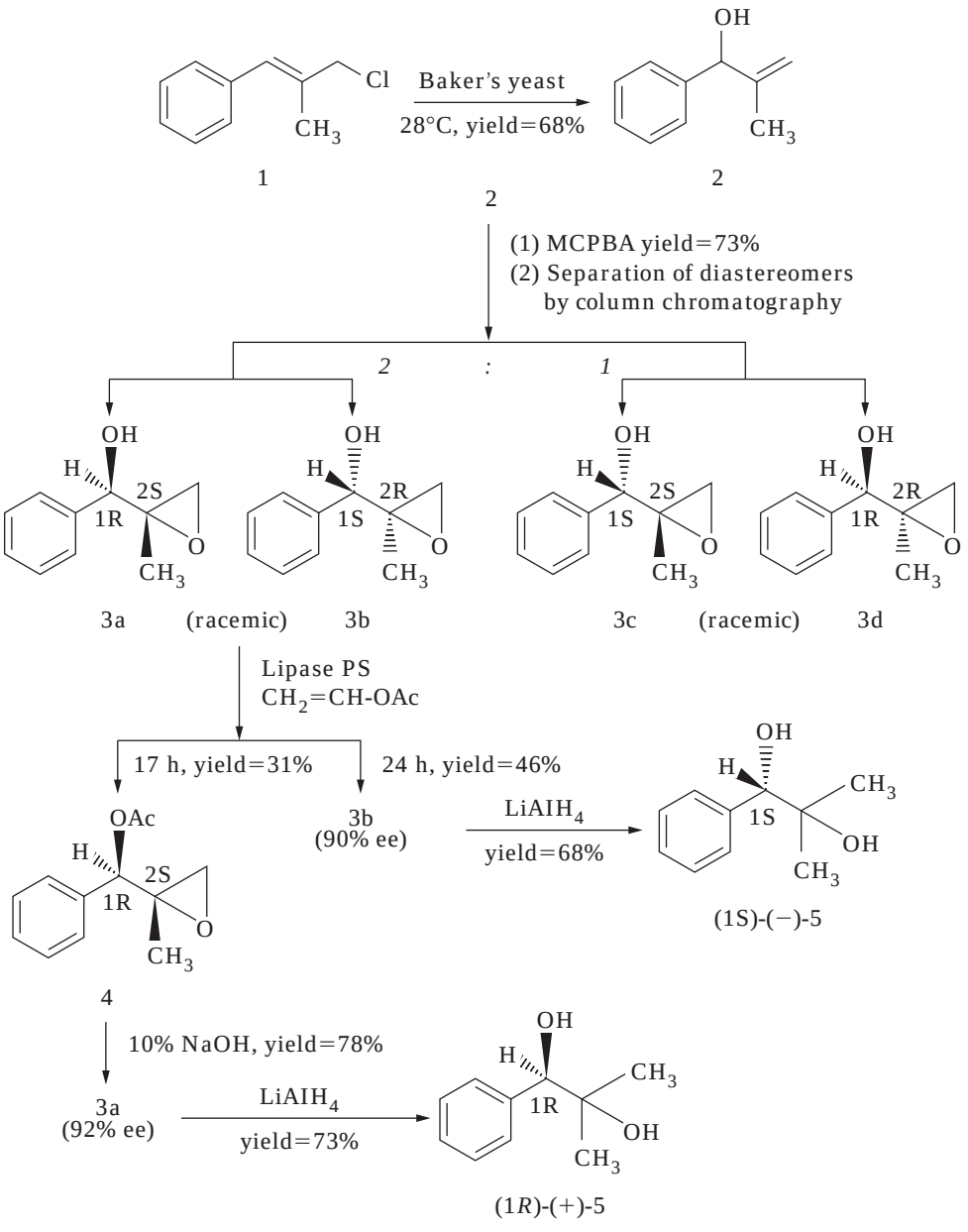
Fantin et al. (39) reported a new method for preparing optically pure secondary alcohols by combining microbial oxidation and reduction. The chiral secondary alcohols were obtained by microbial or enzymatic reductions of the



Scheme 24. Oxidation with *Bacillus stearothermophilus* in heptane.

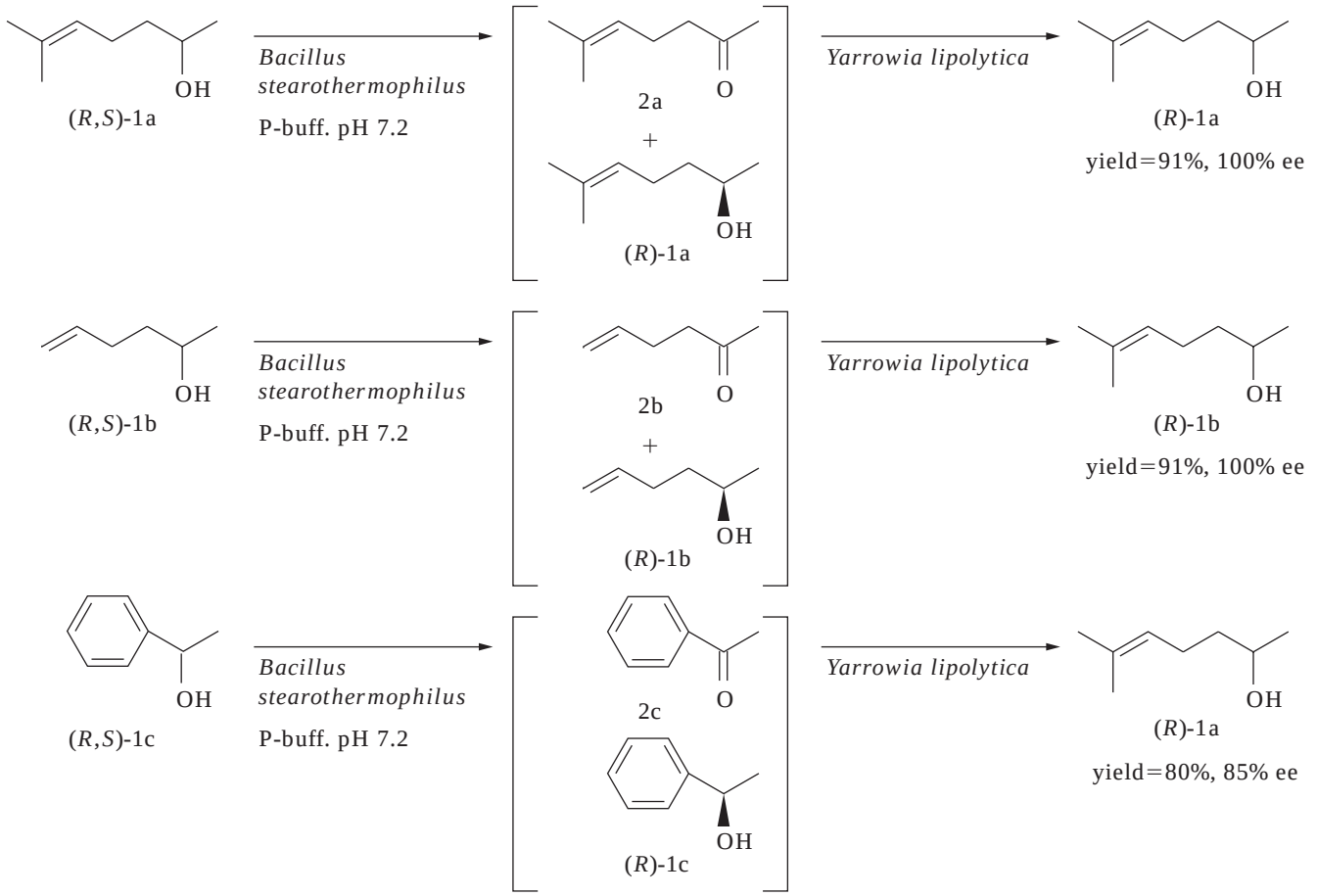


Scheme 25. Reduction of acetophenone derivatives with *Rizopus arrhizus*.



Scheme 26. Enzymatic synthesis of (1*R*,2*S*)- and (1*S*,2*R*)-2-methyl-2,3-epoxy-1-phenylpropanols with baker's yeast and lipase PS.

corresponding ketones followed by kinetic resolution of the racemic alcohols via esterification or hydrolysis of the corresponding ester with lipases. Various ketones were reduced using baker's yeast or purified alcohol dehydrogenases. The reaction with baker's yeast yielded almost exclusively the (*S*)-alcohol, whereas the enzyme produced the hydride either from the *si*-side (i.e., *Lactobacillus kefir* alcohol dehydrogenase) or the *re*-side (i.e., *Thermoanaerobium brockii* alcohol dehydrogenase) of the ketone to give (*R*)- or (*S*)-alcohols, respectively. For most cases, the stereochemical course of the reaction, which mainly depended on the steric requirements of the substrate, could be predicted from the simple model that is generally referred to as Prelog's rule. A majority of the dehydrogenases used for the stereospecific reduction of ketones and a majority of microorganisms follow Prelog's rule, giving (*S*)-alcohols. However, the enzymatic resolution of racemic alcohols with lipases, in which a high enantioselectivity is exhibited, stops at the 50% conversion stage. The authors (39) succeeded in establishing an alternative method for kinetic resolution of secondary alcohols via oxidation with baker's yeast and *B. stearothermophilus*, affording (*R*)-enantiomer in high enantiomeric excess. Moreover, they also reported an anti-Prelog behavior (formation of the (*R*)-alcohol) in microbial reductions of various carbonyl compounds with *Yarrowa lipolytica* strains, a new yeast species. They described a new approach to the high-yield synthesis of homochiral secondary (*R*)-alcohols starting from the corre-



Scheme 27. Optical resolution of the racemic alcohols by a combination of microorganism-catalyzed oxydation and reduction.

sponding racemates. The approach consisted of (*1*) kinetic resolution via oxidation with *B. stearothermophilus* and (*2*) reduction of the obtained oxidation mixture with *Yarrowia lipolytica* (Scheme 27).

The asymmetric microbial degradation of (*S*)-(+)-mandelic acid was investigated in order to develop a practical process for (*R*)-(−)-mandelic acid production from racemic mandelic acid, which is commercially produced on a large scale at low cost using chemical synthesis. Among the 790 culture strains tested, microorganisms belonging to the *Brevibacterium*, *Pseudomonas*, *Rhodococcus*, *Rhodotorula*, *Rhodosporidium*, *Sporobolomyces*, and *Gibberela* genera exhibited high (*S*)-(+)-mandelic acid–degrading activity. *P. polycolor* IFO 3918 was determined to be the best strain and used as a biocatalyst for eliminating the (*S*)-(+)-isomer (40). The maximum rate of (*S*)-(+)-isomer degradation was obtained at 30 °C and pH 7.0. Under these optimal conditions, the (*S*)-(+)-isomer in a racemic mandelic acid 45 g/L mixture was completely degraded within 24 h, with 20 g of (*R*)-(−)-mandelic acid per liter remaining in the reaction mixture. Crystalline (*R*)-(−)-mandelic acid with a chemical purity greater than 99% and optical purity with 99.9% ee was obtained in 35% yield by acidification of the reaction mixture, extraction with ethyl acetate, and subsequent concentration.

CONCLUSION

Optical resolution by enzymes and microorganisms facilitates enantio-selective synthesis of the bioactive chiral compounds and precursors of natural compounds. Combination of organic synthesis with biotransformation should afford efficient optical-selective and regioselective preparation processes for bioactive natural products such as antibiotics, hormones, pheromones, natural fragrances, and various kinds of natural drugs.

BIBLIOGRAPHY

1. C.-S. Chen and C.J. Sih, *Angew. Chem. Int. Ed. Engl.* **28**, 695 (1989).
2. Z.-F. Xie, *Tetrahedron Asymmetry* **2**, 733 (1991).
3. R.S. Ward, *Tetrahedron Asymmetry* **6**, 1475 (1995).
4. T. Yokomatsu, N. Nakabayashi, K. Matsumoto, and S. Shibuya, *Tetrahedron Asymmetry* **6**, 3305 (1995).
5. K. Kato, M. Katayama, S. Fujii, and H. Kimoto, *Biosci. Biotech. Biochem.* **59**, 2178 (1995).
6. O. Yamada and K. Ogasawara, *Synthesis* **10**, 1291 (1995).
7. K. Tanaka and K. Ogasawara, *Synthesis* **10**, 1237 (1995).

8. S. Fernandez, M. Ferrero, and W.H. Okamura, *J. Org. Chem.* **60**, 6057 (1995).
9. T. Izumi and S. Murakami, *J. Hetrocycl. Chem.* **32**, 1125 (1995).
10. M. Lundth, O. Nordin, E. Hedenstroem, and H.-E. Hoegberg, *Tetrahedron Asymmetry* **6**, 2237 (1995).
11. M.-J. Kim, G.-B. Choi, J.-Y. Kim, and H.-J. Kim, *Tetrahedron Lett.* **35**, 6253 (1995).
12. T. Yamane, M. Takahashi, and K. Ogasawara, *Synthesis* **4**, 444 (1995).
13. O. Azzolina, S. Collina, and D. Vercesi, *Farmaco* **50**, 10, 725 (1995).
14. V. Khlebnikov, K. Mori, K. Terashima, Y. Tanaka, and M. Sato, *Chem. Pharm. Bull.* **43**, 1659 (1995).
15. D.S. Tan, M. Gunter, and G. Drueckhammer, *J. Am. Chem. Soc.* **117**, 9093 (1995).
16. M. Kataoka, K. Shimizu, K. Sakamoto, Y. Yamada, and S. Shimizu, *Appl. Microbial. Biotechnol.* **43**, 974 (1995).
17. A.S. Bommarius, M. Schwarm, K. Stingl, M. Kottenharhn, and K. Hethmature, *Tetrahedron Asymmetry* **6**, 2851 (1995).
18. M. Takahashi, R. Koike, and K. Ogasawara, *Chem. Pharm. Bull.* **43**, 1585 (1995).
19. D.-F. Tai, Y.-H. Chao, C.-Y. Huang, J.-M. Luo, Y.-T. Lin, and S.-H. Wu, *J. Chin. Chem. Soc.* **42**, 801 (1995).
20. A.N. Serreqi and R.J. Kazrauskas, *Can. J. Chem.* **73**, 1357 (1995).
21. S.-K. Kang, J.-H. Jeon, T. Yamaguchi, J.-S. Kim, and B.S. Ko, *Tetrahedron Asymmetry* **6**, 2139 (1995).
22. H. Ebiike, Y. Yamazaki, and K. Achiwa, *Chem. Pharm. Bull.* **43**, 1251 (1995).
23. A. Ayi, R. Guedj, and B. Septe, *J. Fluor. Chem.* **73**, 165 (1995).
24. M. Mishitz, W. Kroutil, and K. Faber, *Tetrahedron Asymmetry* **6**, 1261 (1995).
25. H. Akita, C.Y. Chen, and K. Uchida, *Tetrahedron Asymmetry* **6**, 2131 (1995).
26. M. Trani, A. Ducret, P. Pepin, and R. Lorti, *Biotechnology Lett.* **17**, 1095 (1995).
27. A. Sattler and G. Haufe, *Tetrahedron Asymmetry* **6**, 2841 (1995).
28. K. Lundell and L.T. Kanerva, *Tetrahedron Asymmetry* **6**, 2281 (1995).
29. T. Sugai, S. Ohsawa, H. Yamada, and H. Ohta, *Synthesis*, 1112 (1990).
30. S.K. Taylor, F. Richard, R.F. Atkinson, A.P. Almli, M.D. Carr, T.J. Van Huis, and M.R. Whittaker, *Tetrahedron Asymmetry* **6**, 157 (1995).
31. K. Draut, M. Kottenhahn, K. Macrealeas, H. Klenk, and M. Bernd, *Angew. Chem. Int. Ed. Engl.* **30**, 714 (1991).
32. D.R. Durham and J.E. Weber, *Biochem. Biophys. Res. Commun.* **216**, 1095 (1995).
33. G.-J. Kim and H.-S. Kim, *Enzyme Microbial Technol.* **17**, 63 (1995).
34. G. Fantin, M. Fogagnolo, A. Medici, P. Pedrini, S. Poli, and M. Sinigaglia, *Tetrahedron Lett.* **34**, 883 (1993).
35. G. Fantin, M. Fogagnolo, A. Medici, P. Pedrini, S. Poli, and F. Gardini, *Tetrahedron Asymmetry* **4**, 1607 (1993).
36. G. Fantin, M. Fogangnolo, P.P. Giovennini, A. Medici, P. Pedrini, and S. Poli, *Tetrahedron Lett.* **36**, 441 (1995).
37. N.A. Slavi, P.S. Patil, S.R. Udupa, and A. Banerji, *Tetrahedron Asymmetry* **6**, 2287 (1995).
38. M. Takeshita, R. Yaguchi, and Y. Unuma, *Heterocycles* **40**, 967 (1995).
39. G. Fantin, M. Fogagnolo, P. Paolo, M. Alessandro, and P. Pedrini, *Tetrahedron Asymmetry* **6**, 3047 (1995).
40. K. Takahashi, K. Nakamichi, M. Furui, and T. Mori, *J. Ferment. Bioeng.* **79**, 439 (1995).

OPTICAL SENSORS

MAHESH UTTAMLAL
Glasgow Caledonian University
Glasgow, Scotland

DAVID R. WALT
Tufts University
Medford, Massachusetts

KEY WORDS

Arrays
Biosensors
CO_2
Fermentation
Fiber optics
Monitoring
Optical sensors
pH
Sensors

OUTLINE

INTRODUCTION

Many industries such as chemicals, pharmaceuticals, and food require a means of monitoring their production processes. For example, during fermentation, pH, oxygen, carbon dioxide, amino acids, and product concentrations are only a handful of the many parameters that should be monitored to maximize the utilization of reactants and maximize product formation. Most significantly, the lack of suitable sensors in the biotechnology industry is the major obstacle to improve yields through real-time process control. Similarly, monitoring a multitude of analytes during the experimental research and development phase, such as in shaker flasks or microtiter plates, would significantly enhance the ability of biochemical engineers to optimize a process. This article presents the development and application of fiberoptic chemical sensors for bioprocess monitoring.

BIOPROCESS MONITORING

Important Analytes in Bioprocesses

The most common industrial biochemical process is batch aerobic fermentation. Most fermentation processes require control of physical, chemical, biochemical, and biomass parameters. Physical variables such as temperature, pressure, agitation speed, and feed rates are presently measured on-line with conventional sensors that are generally well-developed and reliable. The common chemical measurements of interest in all fermentation broths, from brewing beer to the most modern biotechnological process, include pH, dissolved oxygen, dissolved carbon dioxide, and redox potential. Oxygen and carbon dioxide are also monitored in the reactor vent gases using either infrared or mass spectroscopic analyzers. Next in importance are the biochemical components in the medium essential for harvest decisions, reaction optimization, and model building. These species include those of general interest in all bioreactions, such as glucose, lactate, ammonia, and phosphate, and those of specific interest, such as tetracycline, penicillin, amino acids, and ethanol. Presently these variables are analyzed off-line by removing filtered samples using sterile sampling techniques.

Requirements of Analytical Measurement Systems for Bioprocess Monitoring (1)

Chemical sensors for monitoring can be divided into four broad categories: electrochemical, optical, mass, and thermal. Bioengineers are perhaps most familiar with electrochemical sensors in the guise of pH electrodes. These sensors have been used for decades in the fermentation industry and have many of the attributes of an ideal sensor. They can be sterilized-in-place; they maintain calibration (to a reasonable degree) during the fermentation process; they are robust and can be used multiple times; and they can be interfaced to a controller that automatically adjusts the pH. The pH electrode is almost an ideal sensor from the bioengineer's perspective. Because the sensor is in direct contact with the fermentation broth, it avoids the need to acquire samples, which limits the number of samples that can be measured per unit time, and avoids errors introduced by separating, preparing, handling, transporting, and storing the samples. In addition, there is no time delay in the feedback control loop. The time-delay problem is particularly significant when intervention or corrective action is required to adjust the parameter being measured to a desired level.

By contrast, chemical sensor technology allows a device to dwell in the sampling environment. By responding rapidly and continuously to concentration changes, sensors can generate a vast amount of data. Microprocessors can store and evaluate as much data as the sensors can produce. Sensors avoid all the problems associated with sample acquisition and handling and allow for timely intervention to make appropriate adjustments.

PRINCIPLES OF CHEMICAL SENSORS AND THEIR APPLICATION TO BIOPROCESS MONITORING

In this section, a brief overview of chemical sensors, their construction, and types are given. The reader is referred to other texts (2,3) that treat the subject of chemical sensor technology in greater depth, including design, mathematical treatments, examples, and limitations.

Definition of a Chemical Sensor

A chemical sensor (4) is a device that incorporates a sensing reagent integrated with a signal transducer to afford a system that can measure a target analyte(s) (Fig. 1). The interaction of the target analyte with the sensing reagents generates a signal (electrochemical, thermal, optical) proportional to the concentration of the analyte in the bulk medium. Binding can occur by molecular recognition, such as host–guest, antibody–antigen, or protein-ligand binding, or can involve chemical reactions such as acid–base, redox, or ion exchange.

An ideal chemical sensor must possess certain characteristics for it to be a useful analytical tool in bioprocess monitoring (5). First, the sensor must be specific to the target analyte. This specificity depends on the immobilized

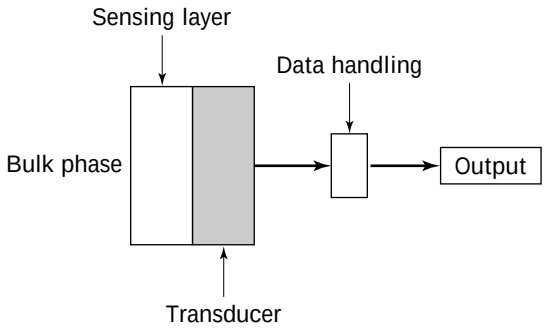


Figure 1. Components of a chemical sensor.

reagents. Biologically derived materials such as enzymes, receptor proteins, and antibodies are extremely valuable in this regard because nature has generated many analyte specific compounds. Second, the signal changes on interaction with the target analyte must be large and be sensitive in the concentration range of interest. If the latter is not possible, then sample pretreatment is required, which leads to uncertainty in the final measurement. Third, the sensor must have a response time consistent with timely intervention. For bioprocesses, response times on the order of a few seconds are sufficient. Finally, the sensor must be stable for the duration of the fermentation process; recalibration is difficult when the sensor is being used on-line or in situ.

Other important factors must be considered for application to bioprocess monitoring, such as the ability of the sensors to withstand autoclaving in situ and to avoid losing sensitivity or calibration during the process. In situ calibration is difficult and seldom practiced.

Types of Chemical Sensors

The interaction of the target analyte with the sensing layer produces a signal when it is interrogated by a suitable transducer. The transducer is usually one of four types: electrochemical, mass, thermal, or optical.

Electrochemical Sensors (6). Many chemical interactions at the sensing layer involve ionic equilibrium or electron transfer. These interactions can be interrogated using an electrochemical transducer such as a metal electrode. There are three main types of electrochemical sensors: potentiometric (symmetric), potentiometric (asymmetric), and amperometric.

Symmetric potentiometric sensors are commonly known as ion-selective electrodes (ISEs) (7) and typically consist of two electrolytes separated by a selective membrane. The potential difference between the internal reference solution and the bulk sample solution is measured with respect to a reference electrode (e.g., Ag/AgCl or saturated calomel electrode).

The classical example of an ISE is the pH electrode. First reported by Cremer (8), it is the best developed sensor and has been commercially available since the 1920s. Symmetric potentiometric devices are also available for many other single and divalent cations (e.g., Na^+, NH_4^+, Ca^{2+}). The pH electrode has been used to prepare CO_2 sensors and other potentiometric biosensors.

Asymmetric potentiometric sensors are different from ISEs in that they do not contain an internal electrolyte solution. In these devices, sometimes referred to as solid-state devices, the sensing membrane is in contact with only one electrolyte (9). This approach has led to the development of ion-selective field effect transistors (ISFETs) (10) and consists of an ion-selective membrane in contact with semiconductor surfaces. The stability of these devices has been limited and matches that of ISEs. Potentiometric sensors are equilibrium devices, and the output is governed by the Nernst equation. The potential is proportional to the log of the activity of the target analyte (11).

In amperometric devices (12), the current at an electrode is measured at a fixed electrode potential. The current generated is proportional to the concentration of the target analyte. The first report of this type of sensor was the Clark oxygen electrode (13). In this device, a layer of buffer solution is trapped on the surface of a platinum electrode by a gas permeable membrane, typically PTFE or silicone rubber. The electrode is polarized to -750 mV versus Ag/AgCl reference electrode. At this potential, any oxygen arriving at the electrode surface is reduced; thus, the magnitude of the resulting current is directly proportional to the external oxygen concentration when mass transport through the membrane is rate limiting.

Mass Sensors (14). The interaction of a target analyte with a sensing layer results in an overall increase in mass at the sensing site. Microbalances can be used to prepare chemical sensors based on the mass increases. Piezoelectric materials (e.g., quartz) are ideal for use in preparing microbalances or piezoelectric sensors. In these devices, the piezoelectric material is used to create and conduct acoustic waves. The crystals are configured into resonant structures that transmit acoustic signals in a frequency-dependent manner. A change in mass of the material on the piezoelectric crystal will cause the resonant frequency to change, which is proportional to the concentration of the analyte. A detailed mathematical treatment of the principles of this work is given elsewhere. Piezoelectric mass detectors have been used to prepare humidity gauges, thin film sensors, and gas detectors. Liquid-phase systems have also been reported for a variety of analytes.

Thermal Sensors (15). Many chemical interactions result in the evolution or absorption of heat. The change in temperature brought about by the interaction can be used to make analytical measurements. The sensing tip is usually a thermocouple or any other temperature probe such as a thermistor. The sensing layer is usually an enzyme that generates heat via catalysis of substrate to product.

Optical Sensors (16). Spectroscopic techniques are well established and are routinely used in analytical laboratories. These techniques interrogate the optical properties of molecules that vary depending on the nature of their environment or change as a consequence of a chemical interaction. Fiberoptic chemical sensors (FOCSs) consist of an indicator molecule immobilized on the distal end or along

the length of an optical fiber. The fiber is used primarily for carrying light information to and from the controlling instrument. A full treatment of FOCSs is given later.

Biosensors (17). Biosensors are devices that incorporate a biological catalyst connected or integrated with a transducer. The biocatalyst catalyzes a chemical reaction, and the transducer converts the biochemical activity into an electrical, optical, or thermal signal. The biological element can be an enzyme, antibody, receptor, tissue, organism, or whole cells.

The first biosensor reported was developed for the measurement of glucose (18). The device consists of a Clark oxygen electrode with the enzyme glucose oxidase immobilized on the external side of the gas-permeable membrane. The device operates by measuring the depletion of oxygen brought about by the oxidation of glucose by glucose oxidase.

The amperometric determination of glucose has been commercialized in the form of test strips for blood glucose measurements by Medisense Inc. In their device, the oxygen dependency of the Clark device is removed by the introduction of an Fe^{2+}/Fe^{3+} redox transport system.

The literature is replete with articles on biosensors, and there are several journals dedicated to this subject.

Sampling Methods (19)

There are two main strategies for on-line fermentation analysis: either the sensor is located outside the reactor and the sample is transported to the sensor (ex situ) or it is placed directly in the fermenter (in situ).

In Situ Monitoring (20). A chemical sensor that meets all the requirements for applications monitoring can be used on-line or in situ. Several assemblies that have been developed to undertake in situ monitoring are shown in Figure 2. In situ monitoring with chemical sensors is rare and generally limited to gas sensing where the transducer surface is protected by a gas-permeable membrane. Sensors that use semipermeable membranes are subject to rapid fouling and poisoning of the sensor surface.

Ex Situ Monitoring (21). The problems associated with in situ monitoring have resulted in the development of ex situ (Fig. 3) systems. In ex situ monitoring, the sample is drawn continuously to a remote site where the analysis is carried out. The procedure removes the need for sensor sterilization, and on-line calibration is easily achieved. In many cases, sample pretreatments are required before analysis, such as filtering, adjustment of the sample pH, temperature, and dilution. The main drawbacks to ex situ monitoring are the need for the use of pumping systems, the possibility of back-contamination at the sampling port, and the effect on the analyte after sample pretreatment.

Application of Chemical Sensors to Bioprocess Monitoring (22)

In the previous section, the requirements of chemical sensors and other analytical systems to bioprocess monitoring were discussed. There are inherent problems with using sensors for bioprocesses. Some of the main difficulties that have to be overcome include the needs to provide aseptic sensors; to prevent microbial fouling, poisoning, and disruptive effects of chemical and biochemical constituents on the sensors; to avoid biochemical and chemical interferences in complex analyte mixtures; to preclude analyte chelation, binding, and buffering effects; and to perform sensor calibration, all while operating under a wide variety of aqueous physiochemical conditions.

Despite the large number of papers, books, and reviews on chemical sensors in the public domain, only a tiny fraction are dedicated to bioprocess monitoring, and very few results have been presented on the successful application of these devices to bioprocess monitoring.

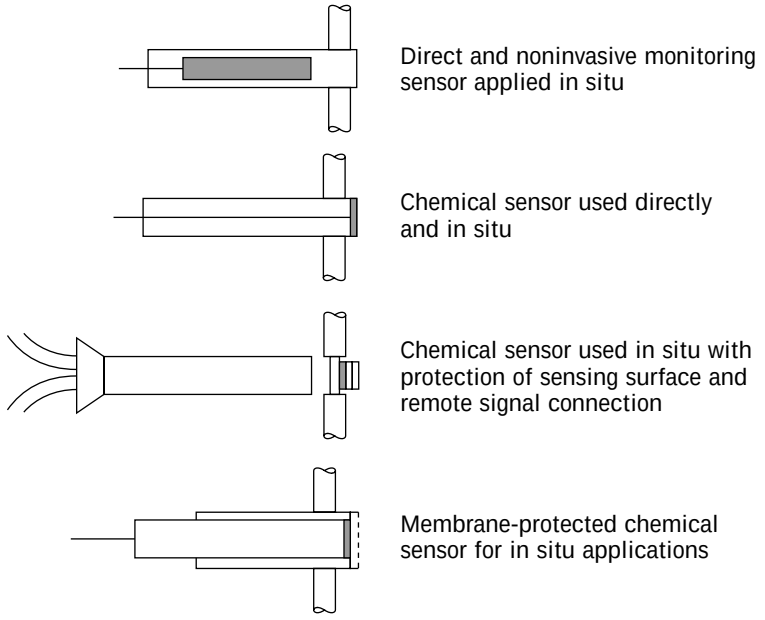
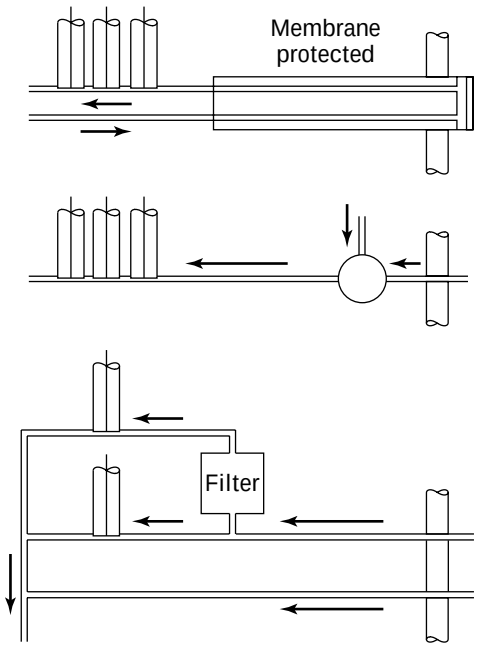


Figure 2. In situ sampling assemblies used for chemical sensors in fermentation monitoring. The sensor is in direct contact with the fermentation sample. Sample pretreatment is not possible with these configurations.



Continuous flow dilution cell and multiprobe used ex situ

Semicontinuous sampling and analysis by principles similar to flow injection analysis

Continuous flow-line sampling of untreated or filtered culture

Figure 3. Ex situ sampling assemblies for using chemical sensors on-line. The sample is drawn from the fermentation vessel to the sensor(s). Sample pretreatment is possible with these configurations.

The first electrochemical sensor developed was the potentiometric pH glass electrode, and it has been modified and implemented for in-line or in situ bioprocess monitoring (23). It is by far the most widely used and the most successful chemical sensor to date. The pH sensor can be autoclaved and is generally considered biocompatible. The sensor is calibrated both before and after the fermentation, and these data are used to correct the measured pH values.

Carbon dioxide sensors based on pH measurements have also been used for in situ bioprocess monitoring (24). Carbon dioxide is measured indirectly by measuring the pH of a bicarbonate buffer solution in equilibrium with carbon dioxide. These types of sensors are based on original work by Severinghaus and Bradley (25).

The Clark oxygen sensor has also been adapted for bioprocess monitoring (26). Steam sterilizable glucose sensors based on oxygen measurements have been reported with lifetimes sufficient to last the course of the fermentation (27).

Commercial Chemical Sensor Systems for Bioprocess Monitoring

The measurement of key analytes in bioprocesses has led to the commercialization of several chemical sensors. The market is dominated by sensors for the most important analytes: pH and dissolved O_2. Table 1 summarizes a few of the commercially available sensors. Sensors for CO_2 are available but used primarily for off-line gas analysis. Most CO_2 gas analysis is performed using nondispersive FT-IR measurements. Optical sensors are not currently available commercially for bioprocess monitoring.

Table 1. Commercial Sensors for Bioprocess Monitoring

Manufacturer	Analyte	Transducer
YSI Inc. (Ohio)	pH	Potentiometric
	Glucose	Enzyme/amperometric
Russell pH Electrode Ltd (Scotland)	pH	Potentiometric
Mettler-Toledo (Greifensee, Switzerland)	pH	Potentiometric
	Dissolved O_2	Amperometric
Phoenix Electrode Inc. (Houston, Texas)	pH	Potentiometric
	Dissolved O_2	Amperometric

FIBEROPTIC CHEMICAL SENSORS

Optical Fibers

Optical fibers (28), also called fiberoptics, are used extensively in the telecommunications industry as a means for conducting information over long distances. Optical fibers are strands of glass or plastic consisting of a core, clad, and protective outer sheath. Generally, optical fibers range in diameter from 50 μm to 1 mm. The transmission of optical signals through an optical fiber is based on the phenomenon of total internal reflection. Figure 4 shows a typical optical fiber made of a core material of refractive index n_1 surrounded by a cladding material with a lower refractive index n_2. Because the index of refraction changes abruptly from the core to the clad, this type of arrangement is described as a step-index fiber. When the incident light entering the fiber strikes the core–clad interface, part of the light will be transmitted and some reflected. If light strikes the interface at an angle greater than the critical angle C as defined by equation 1,

$$\sin C = n_2/n_1 \tag{1}$$

then only light within a characteristic acceptance cone will be introduced into and carried within the fiber, as described by equation 2,

$$\sin\alpha = [(n_1^2 - n_2^2)^{1/2}]/n_0 \tag{2}$$

where sinC is the acceptance cone half angle and n_0 is the

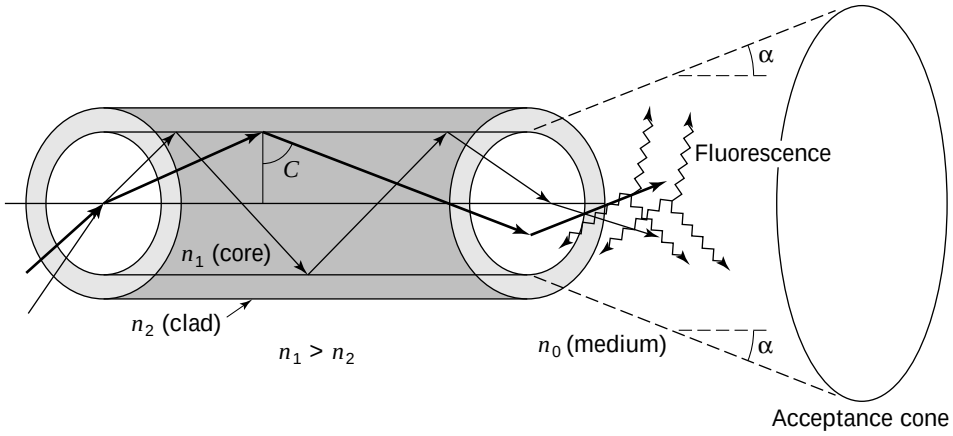


Figure 4. Cross-section of an optical fiber showing core, clad, and optical pathways.

index of refraction of the medium. By convention, collection efficiency is described by the numerical aperture (NA) or f number. As defined by equations 1 and 2, the greater the difference in refractive indices between the core and clad, the larger the acceptance half angle and therefore the more efficient the light collection of the fiber.

Preparation of FOCSs

The preparation of FOCSs requires immobilizing the indicator onto the optical fiber. The indicator is immobilized onto the core of the fiber either on the distal end or along the length of the fiber. The latter approach is used for the preparation of evanescent wave devices (16). In this article, only distal-end FOCSs will be considered because these are dominant and more useful in bioprocess monitoring.

Immobilization Techniques (29). One of the most critical aspects of optical sensors is the method of coupling the various indicating species to the optical fiber, a process known as immobilization. Immobilization techniques vary and are used according to the type of sensor being developed.

Single-Analyte Sensors. Several ways of immobilizing indicating species are shown in Figure 5. Figure 5**a** shows the most rudimentary form of an optical sensor. In this design, the optical fiber is used simply to conduct light to and from an indicator solution that is affixed to the distal end via a semipermeable membrane. Analyte diffuses across the semipermeable membrane and reacts with the indicator. Under ideal circumstances, the indicator is reversible, so that when the concentration of analyte is reduced, it dissociates from the indicator and leaves the solution in contact with the optical fiber. Many different forms of this design exist. Such designs can be extremely robust. This approach suffers from the need to assemble them by hand. Furthermore, they are not reproducible and generally possess long response times.

Figure 5**b** shows a sensor in which the sensing layer is coupled directly to the optical fiber surface. In this approach, a polymer layer containing immobilized indicator is in direct contact with the fiber. This approach is technically more demanding because it requires significant

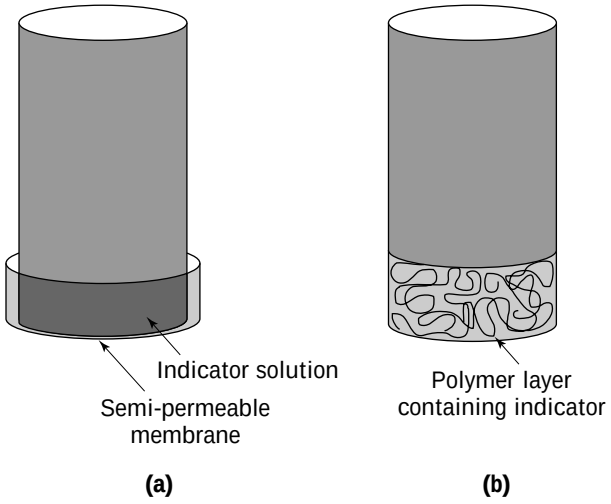


Figure 5. Immobilization methods for the preparation of FOCSs.

surface and polymer chemistry for implementation. First, the optical fiber surface is modified to possess reactive functional groups. Next, a prepolymer solution containing a modified indicator is applied to the optical fiber surface. Polymerization is initiated using either a thermal or photochemical approach. Surface reactive groups on the fiber engage in reaction with the prepolymer and functionalized indicator. In this manner, the indicator is covalently linked to the fibers' distal tip, creating a robust, resilient polymer layer to which the indicator is attached.

Multianalyte Sensors (30). In a typical measurement scenario, a multitude of analytes is often measured. Using conventional single-core optical fibers, it is necessary to use a separate optical sensor for each analyte. With additional sensors, the complexity of the readout instrumentation increases. An alternative approach uses imaging optical fibers (31). These fibers are arrays of coherent fused bundles containing thousands of individual optical fibers, each with its own clad and core, that are capable of transmitting images from one end of the array to the other. Such a transmitted image is shown in Figure 6.

As can be seen in Figure 6, the image is built up pixel by pixel similar to the way the compound eyes of insects

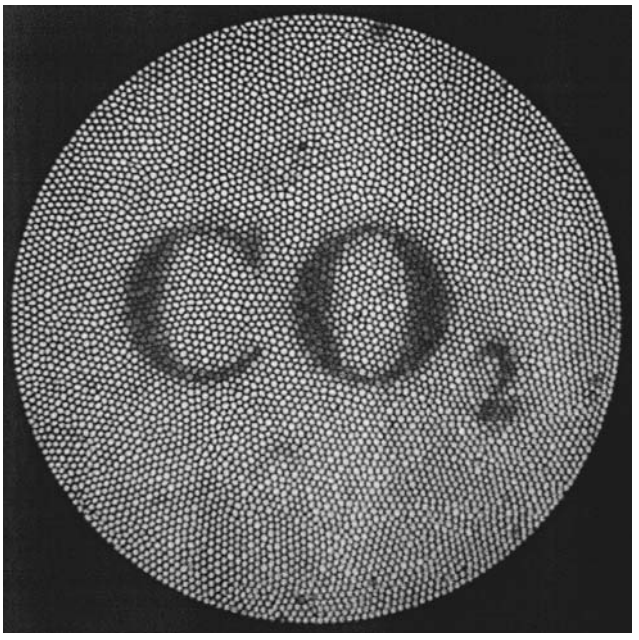


Figure 6. Imaging fiber is comprised of thousands of coherent individual fibers and is able to transmit an image.

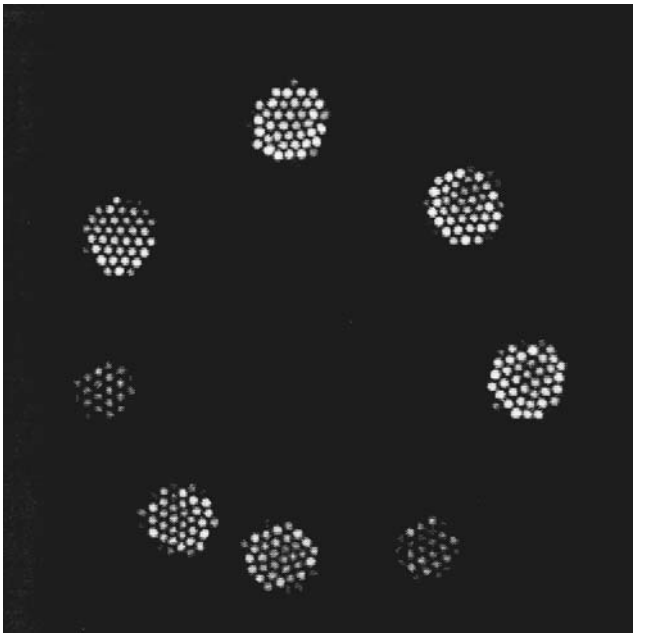


Figure 7. Collected fluoroscence image from a multianalyte imaging sensor shows sensor regions are polymerized over multiple fibers.

operate. This array architecture can be used to create a multianalyte optical sensor (32). In this approach, individual sensors are polymerized in discrete locations on the distal tip of the imaging fiber array using a photodeposition process. Each sensor element conveys an optical signal only through the optical fibers in contact with it. All sensors in the array can be addressed simultaneously. An array detector, such as a CCD camera, can be used to measure the signals from each sensor region. A example of a collected image from a multianalyte sensor array is shown in the Figure 7.

Transduction

There are a myriad of ways in which an optical fiber can be converted into a chemical sensor. In general, one must attach some type of chemical-indicating species to the optical fiber. The chemical-indicating species must provide an optical response that can be measured by the optical fiber.

Absorbance (33). In absorbance-based measurements, the presence of a particular chemical entity is detected by a change in the intensity or color of an indicator or reagent. The indicator is attached to the distal end of the optical fiber. In the presence of analyte, the absorbance increases or decreases in accordance with Beer's law:

$$A = \varepsilon bc \tag{3}$$

where ε is the extinction coefficient, b is the path length, and c is the sample concentration.

Fluorescence (34). An alternative method of detection uses fluorescence. Fluorescence sensors use fluorescent indicators that are more sensitive than absorbance-based sensors. Fluorescent indicators absorb light of one wavelength and emit light at a different wavelength. The advantage of such indicators is that they provide a black-background method because the emission signal arises from a zero background.

Theory of FOCSs

Fluorescence measurements can be made by using either the intensity of fluorescence or the fluorescence lifetime. With intensity measurements, the presence of the analyte correlates directly with the fluorescence intensity signal:

$$I_0/I = 1 + K_s[A] \tag{4}$$

where $K_s = [FA]/[F][A]$ is the binding constant, F is the fluorescence indicator, A is the analyte concentration, I_o is the fluorescence intensity in absence of analyte, and I is the fluorescence intensity in presence of analyte.

Upon binding analyte, the indicator can either increase or decrease its fluorescence intensity. An alternative intensity-based fluorescence measurement results from dynamic fluorescence quenching. This process involves collisional quenching in which the excited state indicating species collides with the analyte during the excited state lifetime. This process can be described by the Stern-Volmer equation:

$$I_0/I = 1 + K_{sv}[A] \tag{5}$$

where K_{sv} is the Stern-Volmer constant and $[A]$ is the analyte concentration.

Fluorescence Lifetime Measurements (35). A second measurement technique involves fluorescence lifetime. A variety of techniques are available for making fluorescence lifetime measurements. For most sensor applications, indicators are still used. The most common technique involves measuring the fluorescence lifetime directly. In this approach, the fluorescent indicator's lifetime is affected by analyte binding. By measuring the fluorescence decay as a function of time, the indicator lifetime can be calculated and correlated to analyte concentration. A significant advantage of this approach is that the background fluorescence does not interfere. Typical background fluorescence decays very rapidly, but most fluorescent indicators have longer lifetimes in the millisecond to nanosecond range. By watching the signal decay, one collects only signals emanating from the indicator. An increasingly popular method for performing fluorescence lifetime measurements is to use a frequency-modulated or phase-angle approach. This approach is based essentially on interferometry in which the delay between the excitation source and the emission signal is measured. The advantage of this technique is it does not require the sophisticated instrumentation necessary to measure fluorescence lifetimes. Fluorescence methods have the advantage of greater sensitivity relative to absorbance-based measurements, with fluorescence lifetime measurements requiring more sophisticated and somewhat more expensive instrumentation.

Ratiometric Measurements (36). With absorbance- or fluorescence-based measurements, it is advantageous to use ratiometric indicators. Ratiometric indicators possess either two excitation or two emission wavelengths that can be measured separately. A hypothetical example of a ratiometric indicator is shown in Figure 8. In this example, the fluorescent indicator has two distinct emission wavelengths. As the analyte concentration increases, the spectral contribution from one of the emission wavelengths increases while the other emission wavelength decreases. Replotting the data in Figure 8 shows that by plotting the ratio versus analyte concentration, the measurement becomes insensitive to the fluorescence intensity of a single peak. Furthermore, the ratiometric measurement is insensitive to changes in light intensity of the excitation source and/or the detector sensitivity. In addition, the measurement is independent of the absolute amount of dye present; dye leaching or photobleaching does not affect the measurement.

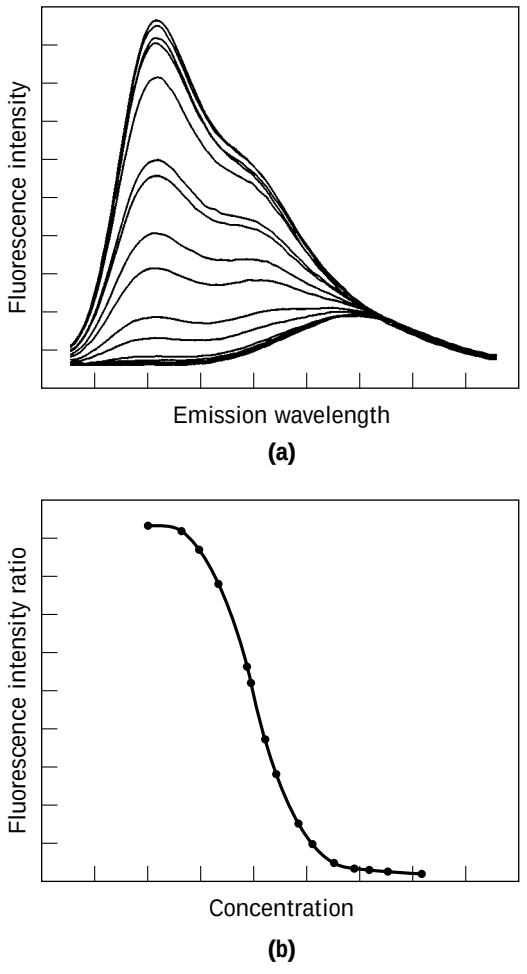


Figure 8. Ratiometric dye with two emission maxima (**a**). Replot of ratio versus analyte concentration is insensitive to instrumental variables (**b**).

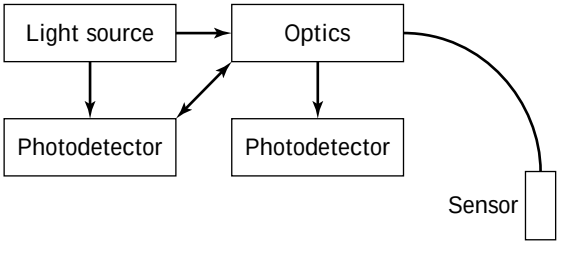


Figure 9. Typical instrumentation design for operating FOCSs.

Instrumentation (37)

The elements required to operate FOCS are shown in Figure 9. Typically, the instrument consists of a light source, optics, photodetector, and controller. The second photodetector is a reference detector that is sometimes necessary when ratiometric and lifetime measurements are not possible.

Light Source. Light sources can be either lasers or lamps, and their choice is dependent on the mode of operation. For example, for fluorescence lifetime measurements, lasers are used because they are good at producing nanosecond excitation light pulses. Table 2 provides a summary of some lasers that are used for optical sensing. In applications where a continuous light source is sufficient, then an arc lamp is often the choice. Lamps generally have relatively broad spectral emission and are much cheaper to purchase and operate than lasers. Figure 10 summarizes the available lamp types.

Optics. The optical components are responsible for focusing, splitting, directing, and selecting light from the

Table 2. Laser Light Sources

Medium	Type	Emitted wavelength (nm)
Ar	Gas	334, 351.1, 363.8, 454.5, 457.9, 465.8, 472.7, 476.5, 488.0, 496.5, 501.7, 514.5, 528.7
Kr	Gas	350.7, 356.4, 406.7, 413.1, 415.4, 468.0, 476.2, 482.5, 520.8, 530.9, 568.2, 647.1, 676.4, 752.5, 799.3
He-Ne	Gas	632.8
He-Cd	Gas	325.0, 441.6
N_2	Gas	337.1
XeF[a]	Gas	351
KrF[a]	Gas	248
ArF[a]	Gas	193
Ruby	Solid	693.4
Nd:YAG	Solid	266, 355, 532
$Pb_{1-x}Cd_xS$[b]	Solid	$2.9 \times 10^3 - 2.6 \times 10^4$
$PbS_{1-x}Se_x$[b]	Solid	$2.9 \times 10^3 - 2.6 \times 10^4$
$Pb_{1-x}Sn_xSe$[b]	Solid	$2.9 \times 10^3 - 2.6 \times 10^4$
$Pb_{1-x}Sn_xTe$[b]	Solid	$2.9 \times 10^3 - 2.6 \times 10^4$
Dye	Liquid	400–460 (Stilbene 420)
	Liquid	450–500 (Coumarin 460)
	Liquid	510–620 (Coumarin 503)
	Liquid	565–620 (Rhodamine 6G)
	Liquid	590–660 (Rhodamine 610)
	Liquid	655–720 (Oxaxine 720)
	Liquid	690–780 (Nile blue 690)

[a]Excimer media.
[b]This is a semiconductor diode laser. Lasing occurs at the *pn* junction. The emitted frequency is altered by altering the concentration of the dopant in the lead.

light source to the sensor and back to the detector. Figure 11 shows two possible configurations for operating FOCSs. Figure 11a is suitable for operating a single fluorescence-based sensor, and Figure 11b is suitable for absorbance-based systems.

Refracting, reflecting, and focusing light is achieved using lenses. Refractive optics include spherical, aspheric, and gradient index lenses, and reflection is accomplished using elliptical, spherical, and parabolic mirrors. Light splitting is achieved using optical couplers, beam splitters, or dichroic filters. Optical wavelength selection is necessary for (*1*) isolating an excitation spectral band, (*2*) selecting an optical signal from stray light, and (*3*) separating the excitation from the emission light. Even with the monochromatic light from a laser, a notch filter is required to avoid scattered or reflected laser light from reaching the detector. Wavelength selection is carried out using optical filters such as absorption and interference filters. It can also be achieved using diffraction gratings.

Photodetectors. The photodetector collects the photons emerging from the optical fiber that carries the information relating to the analytical measurement. There are three types of photodetectors: thermal, photoemissive, and quantum. The most popular types are the photoemissive or photomultiplier tubes (PMTs), which are based on the emission of electrons from a photocathode in a vacuum chamber. When photons of sufficient energy strike the photocathode, it emits an electron from the surface—the photoelectric effect. The electrons are accelerated by a voltage drop across a gap to a secondary emitting electron surface where its numbers are multiplied. The electrons are collected at the anode where they can readily be detected. Available photodetectors are summarized in Figure 12.

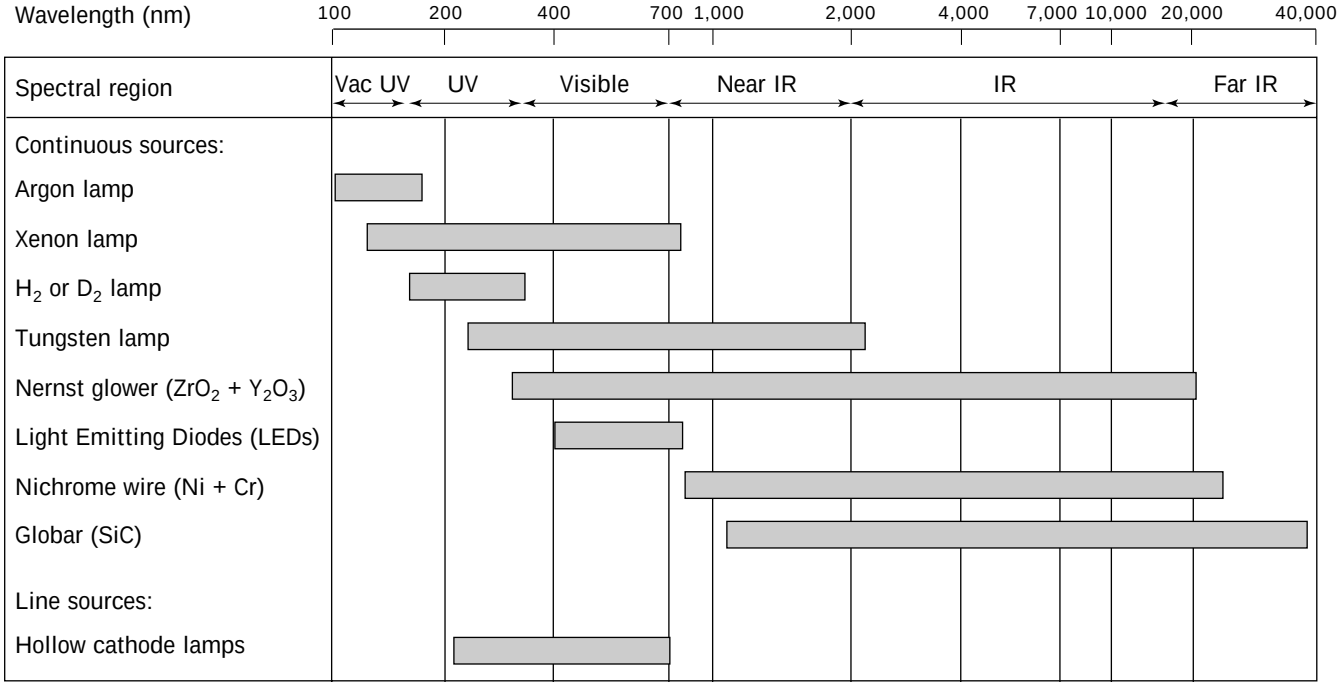


Figure 10. Light sources and their wavelengths.

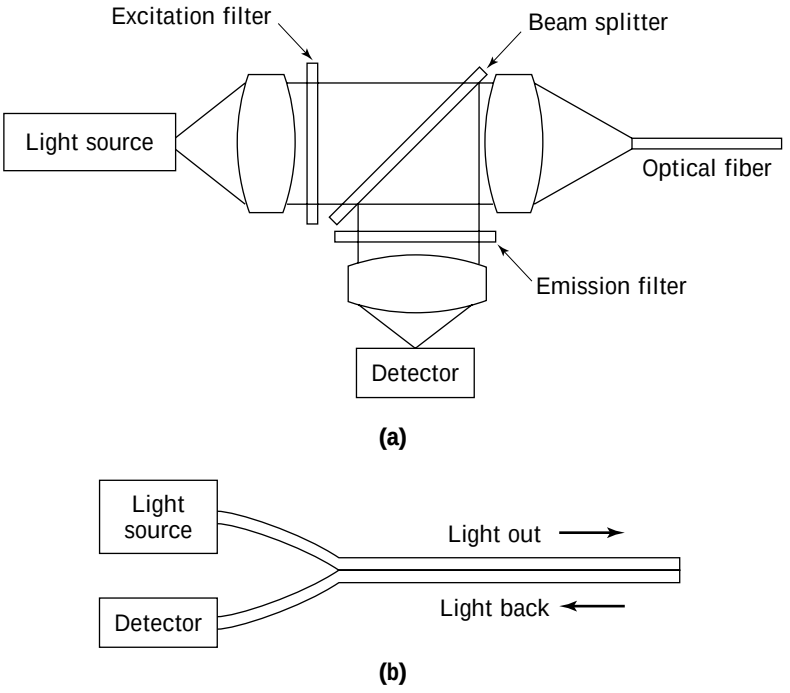


Figure 11. FOCS configurations. (**a**) is suitable for fluorescence-based devices and (**b**) is used for absorbance-based devices.

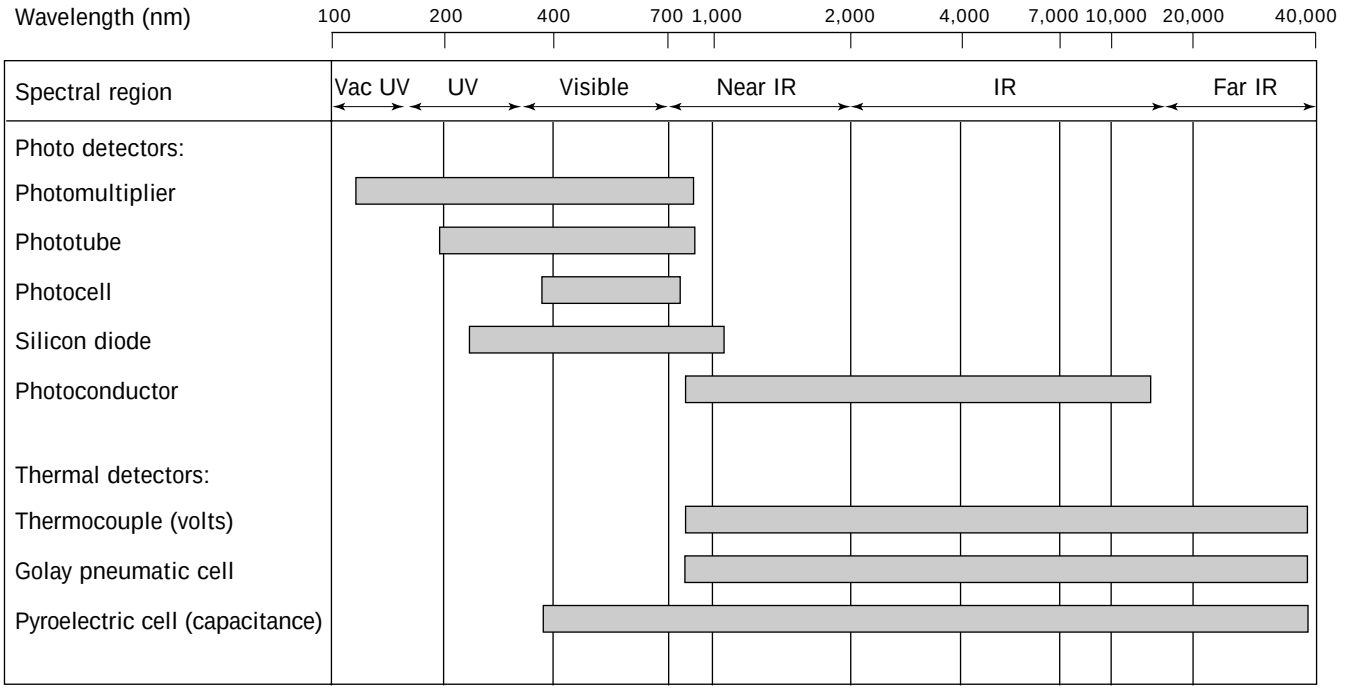


Figure 12. Light detectors and their wavelengths.

FIBEROPTIC SENSORS FOR BIOPROCESS MONITORING

The successful application of FOCSs to bioprocess monitoring has had limited success. The reasons for this lack of success will be discussed later. In this section, optical sensors that have been successfully used in fermentation analysis will be described.

pH Sensors

Measurement of pH. The measurement of pH in bioprocess is routinely carried out and is a good indicator of the state of the process. The spectroscopic determination of pH can be achieved using a variety of absorbance or fluorescent indicators. The particular indicator used depends on

the pH range of interest because most indicators are only sensitive to ±2 pH units around their pK_a. Generally, indicators are selected with a pK_a close to 7 because most fermentations are carried out between pH 5 and 9.

Theory of Spectroscopic pH Measurement (38). The spectroscopic determination solution pH is accomplished by measuring the concentration ratio of the dissociated (In^-) and undissociated (HIn) forms of the indicator. The overall equilibrium can be written $HIn \rightleftharpoons H^+ + In^-$. The acid dissociation constant, K_a, is given by $K_a = [H^+][In^-]/[HIn]$. Provided the total concentration of the indicator remains constant, then

$$pH = pK_a + \log([In]/[HIn]) \quad (6)$$

One form of the indicator (either associated or dissociated) must absorb or fluoresce. In the case where the absorbance of the undissociated indicator is negligible,

$$pH = pK_a + \log\{A/(A_0 - A)\} \quad (7)$$

where A is the absorbance of the dissociated indicator and A_0 is the absorbance when the indicator is completely dissociated. In the case of fluorescence,

$$pH = pK_a + \log\{I/(I_0 - I)\} \quad (8)$$

where I is the fluorescence intensity and I_0 is the intensity when the indicator is completely dissociated. A more optimal situation occurs when both forms of the indicator absorb or fluoresce. In this case, the ratio of the two indicator forms is characteristic of pH (see earlier text).

Bioprocess pH Sensors. The first measurement of fermentation pH using fiberoptics was achieved by dosing the fermentation media with a pH indicator. The optical fiber was placed in situ as a means of delivering and collecting light in order to measure the pH (39). This approach is not optimal because the introduction of a foreign species in a chemically sensitive environment can be detrimental to the outcome of the overall process.

The first example of a FOCS for fermentation pH was described by Agayn and Walt (40), and, to date, it is the only successful report of this type of sensor. The sensor consists of a fluorescein-derivatized hydroxyethyl methacrylate (HEMA) polymer covalently attached to the distal end of an optical fiber. The polymer thickness is approximately 10 μm, and the response time for a step change in pH is less than 5 s. The sensor operates in the ratiometric mode to account for system instability. The sensor can be steam sterilized in situ without any effect on the response characteristics and can be used for at least 2 months with no special storage requirements. The effects of temperature, stirring, and aeration had no significant effect on the sensor response for application to typical fermentation processes. Practical use of the device was demonstrated in the continuous measurement of pH during an *Escherichia coli* fermentation. The ability of this sensor to operate in fermentation media is attributed to the use of the ratiometric method and the use of an acrylate-based hydrogel polymer for the indicator immobilization. Acrylate polymers have been found to be compatible with fermentation and physiological media (41).

CO_2 Sensors

Measurement of CO_2. Cell activity in bioprocesses usually involves the production or uptake of CO_2. The continuous monitoring of both dissolved and off-gas CO_2 are good indicators of fermentation efficiency. Off-gas CO_2 analysis can be achieved using FT-IR analyzers (42) or mass meters, but direct measurement of dissolved CO_2 is not well established. The only chemical sensor developed for CO_2 measurements is the Severinghaus P_{CO2} electrode, as described in "Application of Chemical Sensors to Bioprocess Monitoring." This sensor has not been widely accepted because of extensive drift.

Theory of Spectroscopic CO_2 Measurement (43). The fiberoptic chemical sensor for CO_2 is the optical equivalent of the Severinghaus electrochemical sensor. In this case, the pH electrode is simply replaced by a pH-sensitive optical sensor. The CO_2 sensor consists of a pH sensor covered by a CO_2-sensitive bicarbonate solution. The solution is trapped within the fiber by a gas-permeable membrane, normally PTFE or silicone rubber, which excludes H^+ ions from the solution. The sensor measures the pH of the internal bicarbonate solution in equilibrium with the external CO_2 outside the membrane:

$$CO_{2(ext)} \rightleftharpoons CO_{2(mem)} \rightleftharpoons CO_{2(int)} + H_2O \rightleftharpoons H^+ HCO_3^- \rightleftharpoons 2H^+ + CO_3^{2-}$$

The relationship between the total CO_2 and H^+ is given by

$$[H^+]^3 + N[H^+]^2 - (K_1[CO_2] + K_w)[H^+] - 2K_1K_2[CO_2] = 0 \quad (9)$$

where N is the concentration of sodium bicarbonate in the solution, and K_1 and K_2 are the first and second acid dissociation constants for carbonic acid. CO_2 sensors operate in the pH and CO_2 range where the first and third terms are negligible. K_w is ignored because it is much greater than $K_1[CO_2]$. With these consideration the above equation reduces to

$$[H^+] = K_1[CO_2]/N \quad (10)$$

making the CO_2 concentration directly proportional to the H^+ concentration. A direct consequence of this analysis shows that as the bicarbonate concentration is reduced, the sensor's sensitivity is increased. Unfortunately, this increase in sensitivity is at the expense of a reduced dynamic range.

Bioprocess CO_2 Sensors. The successful measurement of dissolved CO_2 in fermentation media using a FOCS was first reported by Uttamlal and Walt (44). The sensor is based on the Severinghaus P_{CO2} electrode as described earlier. The pH indicator used in the sensor was a dual exci-

tation fluorescent dye, hydroxypyrenetrisulfonic acid (HPTS), and allows the sensor to be operated in the ratiometric mode. The sensor has a reversible dissolved CO_2 dynamic range of 0 to 0.25 atm. The sensor can be autoclaved without affecting calibration. The response time to varying CO_2 step changes is in the range 1 to 13 minutes, in good agreement with studies carried out with the Severinghaus electrode. The temperature effect is negligible in the range 5 to 35 °C. Calibrations performed in fermentation media and clean solutions were alike, making it possible to carry out ex situ calibration. The practical suitability was tested in the monitoring of CO_2 production during the early stages of a beer fermentation.

A second system, reported by Sipior et al. (45), has been developed for off-gas CO_2 analysis and again is based on pH measurement. The sensor operates on changes in fluorescence lifetime in the frequency domain of a pH fluorophore sulforhodamine 101 (SR101) induced by an absorbance-based pH indicator, *m*-cresol purple (MCP). As the pH of the internal solution changes, the absorbance spectrum of MCP changes and alters the degree of interaction with SR101. The change in spectral overlap causes a change in the lifetime of SR101. The sensor calibration is obtained by correlating the phase change as a function of CO_2 concentration. The dynamic working range is 0 to 0.1 atm, and the response time to step changes in CO_2 is a few seconds. The sensor was used to take off-gas CO_2 measurements in an *E. coli* fermentation over a period of 10 h successfully. The lifetime of the sensor is on the order of 2 weeks.

O_2 Sensors

Measurement of O_2. The measurement of dissolved oxygen is important and, like pH and dissolved CO_2, is also an indicator of fermentation efficiency. For example, in beer fermentation, oxygen is required in the early stages for yeast to generate the lipids and sterols in the cell membrane essential for growth. The rate of oxygen uptake has also been shown to be an indicator of yeast vitality (46). Therefore, the importance of dissolved oxygen measurements in beer fermentation is twofold: as a tool to ensure the initial oxygen content is sufficient and to obtain an early indication of fermentation performance.

Theory for Spectroscopic O_2 Measurement (47). The vast majority of fiberoptic oxygen sensors are based on fluorescence quenching of the indicator. The best systems consist of a fluorescent indicator immobilized in a hydrophobic environment (e.g., siloxane). The fluorescence quantum yield is reduced by either static quenching (i.e., the indicator and oxygen associate in the ground state to give a nonfluorescent molecule), or dynamic quenching (i.e., an indicator collides with oxygen, leading to radiation-less deactivation of the excited state). The relationship between the fluorescence intensity (I), the unquenched intensity (I_o), and oxygen concentration is given by

$$I_0/I = 1 + K_{SV}[O_2] \qquad (11)$$

The relationship in equation 11 is known as the Stern-Volmer equation, and it is applicable for both quenching models. If the sensor is operated using lifetime measurements, then I_o/I is replaced by τ_o/τ, where τ_o and τ are the unquenched and quenched lifetimes, respectively.

Bioprocess O_2 Sensors. The successful application of FOCS for O_2 to on-line or in situ bioprocess monitoring has not been demonstrated. However, several systems have been described where the sensors can be steam sterilized and hence are suitable for in situ applications (48,49).

Multianalyte Sensors

The measurement of a single analyte is important, but it reveals only limited information on the state of the fermentation process. A more informed analysis requires the simultaneous measurement of several analytes. Using imaging fiber bundles, Ferguson et al. (50) reported a sensor capable of simultaneously measuring O_2, CO_2, and pH in a beer fermentation.

Figure 13 shows a SEM of the individual sensors that comprise the multianalyte sensor. The working dynamic range of the individual sensors are 0 to 1 atm dissolved O_2 and 0 to 0.1 atm for dissolved CO_2; the pH range is 5.5 to 7.5.

Figure 14 shows the results obtained with the sensor during a beer fermentation. Upon pitching the yeast, O_2 in the wort decreases, pH decreases, and CO_2 increases. These results are as one would predict for such a fermentation and can be used to assess yeast vitality at very early stages of the process.

Optical Biosensors

Carbon and nitrogen sources provide fuel for cell metabolism in fermentation processes. Carbon sources generally consist of glucose and other related sugars, and nitrogen sources include a variety of amino acids. The monitoring of key carbon and nitrogen compounds indicate the state of the fermentation process. The measurement of individ-

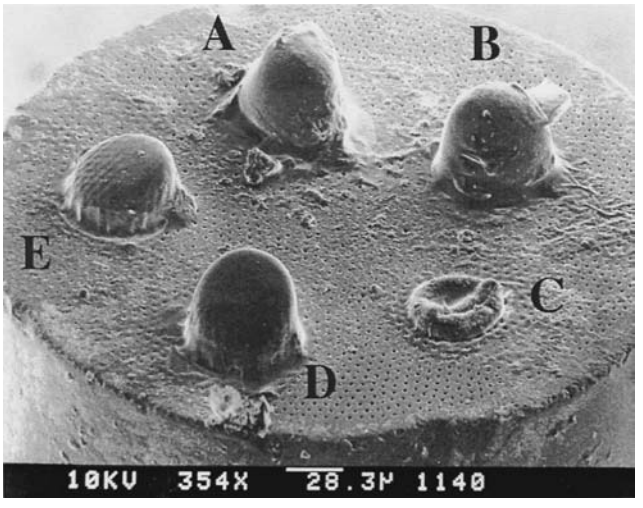


Figure 13. Multianalyte fiberoptic imaging sensor for pH, O_2, and CO_2. Sites *A* and *B* measure CO_2; *C*, pH; and *D* and *E*, O_2.

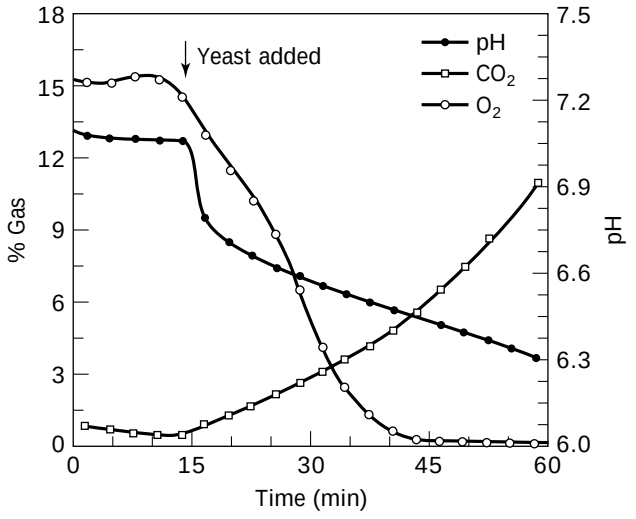


Figure 14. pH, O_2, and CO_2 profile in a beer fermentation using an imaging fiber sensor.

ual sugars and amino acids can be achieved using optical biosensors (51).

Theory of Optical Biosensors. The use of biological species in general and enzymes in particular for analysis is well established, and their incorporation into biosensors was first achieved in the 1950s by Clark and Lyons (18). In optical biosensors, the enzyme is immobilized onto the optical fiber in a suitable matrix. In most biosensors, the reaction of the biological component with the analyte is designed to generate a change in pH, O_2, or CO_2 near the sensor. The measurement of pH, O_2, and CO_2 has been described earlier. To illustrate this principle, glucose analysis can be performed using the enzyme glucose oxidase (GO/FAD):

$$\text{glucose} + \text{GO/FAD} \rightarrow \text{gluconic acid} + \text{GO/FADH}_2 \rightarrow \text{gluconate} + \text{H}^+$$

$$\text{GO/FADH}_2 + \text{O}_2 \rightarrow \text{GO/FAD} + \text{H}_2\text{O}_2$$

By either measuring the production of acid or the depletion of O_2 at the sensor, the glucose concentration can be inferred. It is also possible to measure H_2O_2 production using the indicator luminol; however, this is an irreversible measurement.

Bioprocess Biosensors. Biosensors are very difficult to use in on-line or in situ monitoring because the biological component cannot withstand autoclaving. Therefore, the use of optical biosensors is primarily for ex situ analysis. If in situ measurements are required, alternative sterilization protocols are possible, such as treating the sensor with radiation, ethanol, or antibiotics. After such treatment, the sensor can be inserted aseptically into the fermenter. There are only a few reports of optical biosensors that have been used in fermentation analysis. Short descriptions of three are given next.

Glucose. Glucose and fructose are the main carbon energy sources for cell activity in bioprocesses. Other carbohydrates such as sucrose and maltose are also used as energy sources, but they are broken down first to glucose and fructose by enzymes in the cell (52).

The development of FOCS glucose sensors using the enzyme GO/FAD has been achieved based on O_2 (53), pH (54), and H_2O_2 (55) measurements. These sensors have fast response times and have the dynamic range suitable for fermentation analysis (i.e., 0 to 50 mmol dm^{-3}); however, use in on-line bioprocess monitoring has not yet been achieved.

Amino Acids. Many bioprocesses require a nitrogen source for cell growth (52). Free amino acids, ammonium salts, urea, and gaseous nitrogen are typical of nitrogen-containing compounds used. Pure nitrogen-containing compounds are expensive, so alternatives such as corn steep liquor, yeast extract, peptones, and soy meal are commonly used. These liquors and extracts also contain a carbon source.

The measurement of nitrogen-containing compounds is therefore important in order to achieve a better understanding of the processes taking place and the state of the fermentation, with the particular nitrogen compound depending on the fermentation. An important nitrogen source are amino acids. It is possible to make optical sensors for L-amino acids using L-amino acid oxidase with transduction via the ammonia, ammonium ion, or H_2O_2 produced during biocatalysis:

$$\text{R—CH(NH}_3^+\text{)—CO}_2^- + \text{H}_2\text{O} + \text{O}_2 \xrightarrow{\text{L-Amino acid oxidase}} \text{R—C(=O)—CO}_2^- + \text{NH}_4^+ + \text{H}_2\text{O}_2$$

The sensor responds to many amino acids, and the output corresponds to an aggregate amino acid concentration. Electrochemical devices based on this enzyme have been developed and used in bioprocess monitoring (56).

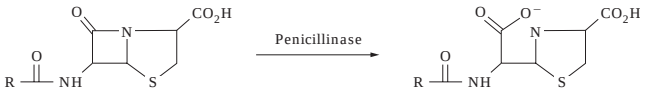
Another example that has been reported is the measurement of urea using the enzyme urease (57):

$$\text{Urea} \xrightarrow{\text{Urease}} 2\text{NH}_4^+ + \text{HCO}_3^-$$

The reaction generates a pH change that can be monitored using a pH sensor.

Penicillins. The discovery of penicillin by Fleming was a major breakthrough in the treatment of infectious diseases. The full-scale production was developed after World War II, and today penicillin and penicillin derivatives are produced in huge quantities (58).

Penicillin production has been monitored using both electrochemical and fiberoptic sensors. Most devices are based on the measurement of pH changes brought about by the hydrolysis of the β-lactam ring by the enzyme penicillinase (59):



The major drawback to this method is that an independent measurement of pH must be made because the solution pH varies during the course of the fermentation. This measurement is not difficult, and pH is generally monitored anyway. A multianalyte FOCS that measures both pH and penicillin has been reported (60). The sensor was used offline to measure penicillin production during a *Penicillium chrysogenum* fermentation (61), and results obtained from the sensor were in close agreement using a standard kit system.

DISCUSSION

The need to measure key analytes in bioprocesses is important for two reasons. First, monitoring will ensure that the process is operating correctly, and second, the information obtained will allow a better understanding of the process that regulates production. To achieve this, new methods of analysis have to be developed. FOCSs, because of their advantages over other analytical systems, are ideally suited for this type of monitoring; however, there are a number of issues that need to be solved before sensors can be implemented in bioprocess monitoring: fouling, drift, in situ calibration, and sterilization are a few of the more important ones.

Fouling

Biofouling occurs when surfaces come into contact with cells and media. When surfaces such as glass, plastic, or metal come into contact with a biological medium, protein adsorption occurs. After protein adsorption, adherent cells can attach to the protein layer. Cell and protein adhesion often result in a relatively thick layer called a biofilm that covers the end of the sensor. Such coverage compromises the measurement capability of the sensor and, in the worst cases, isolates the sensor from the medium. To avoid biofilm formation, it may be necessary to include an antifoulent, such as an antibiotic, that can be placed on the end of the sensor. This material must not interfere with the sensing capabilities of the sensor. Alternatively, it may be possible to place the sensor in a high-shear environment so that cells cannot readily adhere to the sensor surface.

Drift

In cases where internal calibration cannot be used, it may be necessary to account for sensor drift. Drift can be accounted for in several ways. If the properties of the sensor are measured for extended periods of time, it may be possible to observe a consistent change in sensor response as a function of time. For example, if dye photobleaching is a problem, one can measure the photobleaching rate as a function of time and obtain a photobleaching rate constant. This constant can then be used to correct for photobleaching effects during the fermentation process. A second method to account for sensor drift is to perform an in situ calibration in which a sample is taken from the fermentation vessel and the parameter of interest is measured using a conventional chemical analysis. The sensor is then reset to the measured value.

In situ calibration

In situ calibration for sensors being used in situ is difficult to achieve. Doping the fermentation medium with a known amount of analyte can be used, but the implications to the process can be severe. This approach can be a source of contamination, and in addition to other disadvantages, it means that in situ calibration is very rarely undertaken in nonexperimental fermentations.

For sensors that are being used ex situ, the problem is more easily solved. A pumping mechanism with valves can be used to switch between sample and a calibration solution. Sensor lifetime and reliability are increased using this method, but it is not always possible to make measurements in this way.

Sterilization

One of the most vexing issues with respect to sensor implementation is that of sterilization. The most common sterilization protocol in bioprocessing is steam sterilization. Unfortunately, many sensing chemistries are extremely sensitive to the high temperatures required during autoclaving. This problem is of particular concern when dealing with biosensors in which the active enzyme or antibody is denatured during the sterilization process. The only way to obviate this problem is to sterilize the sensor outside the fermentation vessel using a chemical approach, such as soaking the sensor in ethanol or a high concentration of antibiotic and then inserting it aseptically through the appropriate port on the fermentation vessel.

Sensor Lifetime

Another limitation of sensors, particularly biosensors, is their limited lifetime. During the long fermentation runs (e.g., months) of some mammalian cell cultures, sensors and biosensors may lose their activity or sensitivity. In this case, it may be necessary to replace a sensor during the culture. For particularly sensitive cultures in which it would be dangerous or inadvisable to replace a sensor, it may be possible to prolong the lifetime of a sensor by simply decreasing the duty cycle. For example, if photobleaching is a serious problem, one can simply calculate the sensor's total allowable light exposure and then determine how often a reading can be obtained during the fermentation run without compromising the photobleaching sensitivity.

SENSING BIOPROCESSES IN THE FUTURE

The successful monitoring and eventual control of bioprocesses in the future using optical sensors will depend on overcoming the problems discussed earlier and then on developing new molecules and chemical methodologies to undertake the analyses. The following examples of work in

the early stages of development will, we hope, mature to produce ideal optical sensors for bioprocess monitoring.

Multianalyte Sensing

As sensors become more sophisticated, there will be an increasing demand to measure an ever-increasing number of analytes. In addition to the consensus analytes discussed earlier, there are a number of other species that, if monitored, would provide additional information as to the status of a particular fermentation. These analytes include trace minerals such as metal ions, complex or exotic carbon and nitrogen sources, growth regulators, enzymes or other proteins, and a wide variety of products and by-products. As information is gleaned from the Human Genome Initiative, there may be even more need to monitor additional regulators, peptides, and small organic molecules. Furthermore, there will be tremendous value placed on being able to detect specific gene sequences or, more importantly, the expression levels of messenger RNA. In principle, such gene-expression monitoring combined with the conventional and expanded set of analytes discussed above may result in the need to monitor hundreds, if not thousands, of distinct species simultaneously. Although this may seem like a daunting task, tremendous advances are being made presently in the areas of high detection sensitivity, miniaturization, and high-density arrays. Advances in all of these areas may one day lead to the ability to continuously monitor the multitude of analytes necessary to give a complete picture of the conditions and state of a particular fermentation process.

Electronic Noses

All the sensors discussed earlier are based on the conventional one-sensor–one-analyte paradigm in which a selective sensor is designed to respond to a single analyte. As the number of analytes increases, the size and complexity of the sensing system increases proportionately. A recent trend is to construct a sensing system that mimics the olfactory system. The olfactory system relies on a multitude of cross-reactive receptor cells that respond broadly to a wide variety of odorous substances (62). The complex pattern generated by the cells when an odor is presented is relayed to the olfactory bulb and then further transmitted to the higher-order cortical regions of the brain where it is recognized as a particular odor. In a similar vein, electronic noses (63,64) are based on cross-reactive arrays of sensors that differentially respond when exposed to a particular vapor. The generated response pattern is used to train a computational network that then recognizes the particular pattern upon a subsequent vapor exposure. Such a system presents a significant advantage in terms of sensor construction because additional sensors are not required to detect additional analytes; it is sufficient simply to train the network on an additional pattern. Existing manifestations of these systems are expensive and require significant sample preconditioning before sample introduction (65). A limited amount of work has been conducted on the off gases from particular fermentation processes (66). All such systems have only been demonstrated for vapor-phase measurements. The application of such an approach to liquid measurements has not been demonstrated to date. These systems can be trained in a manner more sophisticated than simply detecting the presence or absence of a particular analyte. Such a system can be trained to look at the overall quality or character of the volatile materials in the head space. In principle, such a system may allow the biochemical engineer to train for a desired outcome without necessarily knowing the specific chemical composition of the fermentation broth. For example, it may be possible to train the system on such qualities as maximum product concentration, minimum by-product concentration, off odors, and so on, without specifically training on those particular characteristics. In other words, the sensor may be able to key on surrogate components that are indicative of the desired outcome. This approach has the potential to be a very powerful tool for the biotechnologist and challenges our conventional thinking with regard to chemical monitoring.

Thermostable Enzymes

Thermostable enzymes are isolated from thermophilic microorganisms that exist at elevated temperatures in hot springs and deep ocean vents. Thermostable versions of useful enzymes are presently used in glucose sensors and as reagents for the polymerase chain reaction used in molecular biology. The ability to incorporate thermostable enzymes into biosensors may ultimately enable such sensors to be steam sterilized. Attempts to accomplish such heat stability have met with failure because existing thermostable enzymes still exhibit appreciable denaturation at elevated temperatures.

By isolating and determining the structures of thermostable proteins, researchers are beginning to understand the structural features that lead to enhanced stability. Eventually it may be possible for molecular biologists to engineer enzymes, antibodies, and receptor proteins with extremely high thermal stabilities.

Designed Receptors

An area of significant relevance to sensors is that of molecular recognition. Organic chemists are making great strides in understanding the principles governing how molecules and ions are recognized. For sensors, recognition is only half the story; it is critical that recognition be coupled to signal transaction. Rational molecular design is one approach to developing such receptors. An alternative and increasingly popular approach relies on combinatorial chemistry. In this approach, libraries of large numbers of distinct molecules are created and screened for function. Although most of the focus of combinatorial chemistry has been directed toward drug discovery (67), some recent work has used the approach to search for structures that combine binding with signal transduction (68).

SUMMARY

There are presently a limited number of analytes that can be monitored continuously during a fermentation or cell culture. The bioprocess monitoring scene is changing as

new sensors become available. Optical sensors offer significant advantages because of their small size, internal calibration capability, freedom from electromagnetic interference, and multiplexing ability. As new sensing materials and sensing paradigms become available from the fields of molecular recognition and combinatorial chemistry, thermostable enzymes, microfabrication, electronic noses, and advanced sensors with powerful analytical capabilities will become available. Ultimately, such sensing systems will provide the biotechnologist and bioengineer with timely information regarding the status and composition of a particular fermentation process.

BIBLIOGRAPHY

1. M.T. Riebe and D.J. Eustace, *Anal. Chem.* **62**, 65A–71A (1990).
2. T.E. Edmonds, *Chemical Sensors*, Blackie, Glasgow, 1988.
3. J. Janata, *Principles of Chemical Sensors*, Plenum Press, New York, 1989.
4. R.W. Cattrall, *Chemical Sensors*, Oxford University Press, Oxford, 1997.
5. O.S. Wolfbeis, in J.V. Twork and A.M. Yacynych eds., *Sensors in Bioprocess Control*, Dekker, New York, 1990, pp. 95–125.
6. C. Brett and A.M. Oliveira-Brett, *Electroanalysis*, Oxford University Press, Oxford, 1998.
7. J. Koryta and K. Stulik, *Ion-Selective Electrodes*, Cambridge University Press, Cambridge, 1983.
8. M. Cremer, *Z. Biol.* **47**, 562–608 (1906) (in German).
9. M.J. Madou and S.R. Morrison, *Chemical Sensing with Solid-State Devices*, Academic Press, New York, 1989.
10. B.H. Vanderschoot, P. Bergveld, M. Bos, and L.J. Bousse, *Sens. Actuators* **4**, 267–272 (1983).
11. B.H. Vassos and G.W. Ewing, *Electroanalytical Chemistry*, Wiley, New York, 1983.
12. S.C. Chang, J.R. Stetter, and C.S. Cha, *Talanta* **40**, 461–477 (1993).
13. L.C. Clark, *Trans. Am. Soc. Artific. Internal Organs* **2**, 41–57 (1956).
14. J.F. Alder and J.J. McCallum, *Analyst* **108**, 1169–1189 1983.
15. G.C.M. Meijer, *Sens. Actuators* **10**, 103–125 (1986).
16. S.A. Momin and R. Narayanaswamy, *Optical Fibre Sensors: Current State Review*, UMIST, 1993.
17. E.A.H. Hall, *Biosensors*, Open University Press, Milton Keynes, 1990.
18. L. Clark and C. Lyons, *Ann. N.Y. Acad. Sci.* **102**, 29–45 (1962).
19. P.N. Royce, *Crit. Rev. Biotechnol.* **13**, 117–149 1993.
20. J.B. Callis, D.L. Illman, and B.R. Kowalski, *Anal. Chem.* **59**, 624A–637A (1987).
21. S.O. Enfors and N. Cleland, *Methods Enzymol.* **137**, 298–307 (1988).
22. D.J. Clarke, M.R. Calder, R.J.G. Carr, B.C. Blake-Coleman, S.C. Moody, and T.A. Collidge, *Biosensors* **1**, 213–320 (1985).
23. D.J. Clarke, B.C. Blake-Coleman, M.R. Calder, and A. Burns, *Ion Selec. Elec. Rev.* **4**, 75–133 (1982).
24. E. Puhar, A. Einsels, H. Buhler, and W. Ingold,
25. J.W. Severinghaus and A.F. Bradley, *J. Appl. Physiol.* **13**, 515–520 (1956).
26. Y.H. Lee and G.T. Tsao, *Adv. Biochem. Eng.* **13**, 35–86 (1979).
27. N. Cleland and S.-O. Enfors, *Eur. J. Appl. Microbiol. Biotechnol.* **18**, 141–147 (1983).
28. J.A. Buck, *Fundamentals of Optical Fibers*, Wiley, New York, (1995).
29. G.T. Hermanson, A.K. Mallia, and P.K. Smith, *Immobilized Affinity Ligand Techniques*, Academic Press, Sidcup, U.K., 1992.
30. P. Pantano and D.R. Walt, *Anal. Chem.* **67**, 481A–487A (1995).
31. T. Tsumanuma, K. Tanaka, S. Chigira, K. Sanada, and K. Inada, *SPIE Optical Fibers in Medicine* **906**, 92–96 (1988).
32. S.M. Barnard and D.R. Walt, *Nature* **353**, 338–340 (1991).
33. G.M. Barrow, *Physical Chemistry*, 6th ed., McGraw Hill, London, 1996, pp. 568–572.
34. J.R. Lakowicz, *Topics in Fluorescence Spectroscopy: Volume 1 Techniques*, Plenum Press, New York, 1991.
35. J.R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Plenum Press, New York, 1983.
36. J.E. Jones and R.C. Spooncer, *J. Phys. Ed.* **16**, 1124–1126 (1983).
37. D.N. Modlin and F.P. Milanovich, in O.S. Wolfbeis ed., *Fiber Optic Chemical Sensors and Biosensors Vol I*, CRC Press, Florida, 1991, pp 237–301.
38. T.E. Edmonds, N.J. Flatters, C.F. Jones, and J.N. Miller, *Talanta* **35**, 103–107 (1988).
39. B.H. Junker, D.I.C. Wang, and T.A. Hatton, *Biotechnol. Bioeng.* **32**, 55–63 (1988).
40. V. Agayn and D.R. Walt, *Bio/Technology* **11**, 726–729 (1993).
41. Y. Benmakroha, S. Zhang, and P. Rolfe, *Med. Biol. Eng. Comput.* **33**, 811–821 (1995).
42. R.D. Eckles, J.M. Welles, and K. Petersen, *Measurements & Control* **161**, 83–89 (1993).
43. M.B. Tabacco, M. Uttamlal, M. McAllister, and D.R. Walt, *Anal. Chem.* **71**, 154–161 (1999).
44. M. Uttamlal and D.R. Walt, *Bio/Technology* **13**, 597–601 (1995).
45. J. Sipior, L. Randers-Eichhorn, J. Lakowicz, G.M. Carter, and G. Rao, *Biotechnol. Prog.* **12**, 266–271 (1996).
46. F.L. Peddie, W.J. Simpson, B.V. Kara, S.C. Robertson, and J.R.M. Hammond, *J. Inst. Brew.* **97**, 21–25 (1991).
47. E.R. Caraway, J.N. Demas, B.A. DeGraff, and J.R. Bacon, *Anal. Chem.* **63**, 337–342 (1991).
48. S.B. Bombat, R. Halavanahali, J.R. Lakowicz, G.M. Carter, and G. Rao, *Biotechnol. Bioeng.* **43**, 1139–1145 (1994).
49. B.H. Weigl, A. Holobar, W. Trettnak, T. Klimant, H. Kraus, P. O'Leary, and O.S. Wolfbeis, *J. Biotechnol.* **32**, 127–138 (1994).
50. J.A. Ferguson, B.G. Healey, K.S. Bronk, S.M. Barnard, and D.R. Walt, *Anal. Chim. Acta.* **340**, 123–131 (1997).
51. J.J. Ramsden, *J. Mol. Recog.* **10**, 109–120 (1997).
52. P.F. Stanbury and A. Whitikar, *Principles of Fermentation Technology*, Pergamon Press, Oxford, 1984.
53. B.P.H. Schaffar and O.S. Wolfbeis, *Biosensors Bioelectron.* **5**, 137–148 (1990).
54. W. Trettnak, M.J.P. Leiner, and O.S. Wolfbeis, *Biosensors* **4**, 1519–1523 (1989).
55. F. Schubert, F. Wang, and H. Rinneberg, *Microchim. Acta* **121**, 237–247 (1995).
56. W.J. Albery, J.E.C. Hutchins, and M. Uttamlal, *Appl. Electrochem. A.* **26**, 243–248 (1996).
57. A. Sansubrino and M. Mascini, *Biosensors Bioelectron.* **9**, 207–216 (1994).
58. D. Wilson, *Penicillin in Perspective*, Faber, London, 1976.

59. J.B.K. Nielsen, *Methods Enzymol.* **97**, 153–158 (1983).
60. B.G. Healey and D.R. Walt, *Anal. Chem.* **67**, 4471–4476 (1995).
61. L.H. Christensen, C.M. Henriksen, J. Nielsen, J. Villadsen, and M. Egelmitani, *J. Biotechnol.* **42**, 95–107 (1995).
62. J.S. Kauer, *Trends Neurosci.* **14**, 79–85 (1991).
63. K. Persand and G. Dodd, *Nature* **299**, 352–355 (1982).
64. T.A. Dickinson, J. White, J.S. Kauer, and D.R. Walt, *Nature* **382**, 697–700 (1996).
65. T.A. Dickinson, J. White, J.S. Kauer, and D.R. Walt, *Trends Biotechnol.* **16**, 250–258 (1998).
66. J.W. Gardner, T.C. Pearce, S. Fried, P.N. Bartlett, and N. Blair, *Sens. Actuat. B.* **18–19**, 240–243 (1995).
67. E.M. Gordon, M.A. Gallop, and D.V. Patel, *Acc. Chem. Res.* **29**, 144–154 (1996).
68. W.C. Still, *Acc. Chem. Res.* **29**, 155–163 (1996).

OPTICALLY ACTIVE 1,2-DIOLS, MICROBIAL PRODUCTION BY STEREOINVERSION

JUNZO HASEGAWA
MASAHIRO OGURA
SATORU TSUDA
SHUN-ICHI MAEMOTO
HIDETOSHI KUTSUKI
TAKEHISA OHASHI
Kaneka Corporation
Hyogo, Japan

KEY WORDS

Alcohol dehydrogenase
(*S*)-1,2-Butanediol
Candida parapsilosis
(*R*)-3-Methylthio-1,2-propanediol
(*S*)-4-Methyl-1,2-propanediol
(*S*)-1,2-Pentanediol
(*S*)-1-Phenyl-1,2-ethanediol
(*S*)-3-Phenyl-1,2-propanediol
(*S*)-4-Phenyl-1,2-butanediol
Stereoinversion

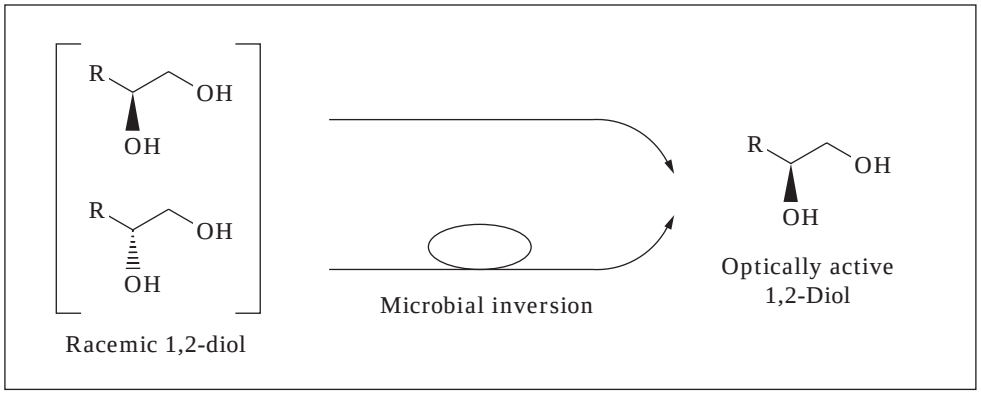
OUTLINE

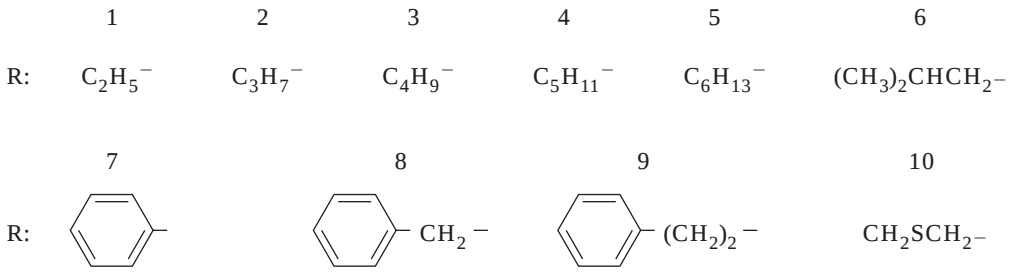
INTRODUCTION

Optically active 1,2-diols are a versatile group as chiral starting materials for the syntheses of pharmaceuticals, agrochemicals, pheromones, liquid crystals, and so on. Several routes of preparation for optically active 1,2-diols have been reported.

1. Microbial reduction of 1-hydroxy-2-ketones (1–5)
2. Chemical reduction of optically active 2-hydroxy acids prepared from 2-keto acids by enzymatic reduction of amino acids (6–11)
3. Optical resolution of racemic 1,2-diols by lipase-catalyzed transesterification in organic media (12)
4. Chemical preparation from naturally occurring products, such as (*S*)-malic acid (2), D-mannitol (2,13,14) or (+)-pulegone (15).

We developed two efficient methods for producing various optically active (*S*)-1,2-diols with high optical purities. One is the organic synthetic method using the Grignard reaction starting from (*R*)-3-chloro-1,2-propanediol as a starting material can be produced easily from the racemate by microbial optical resolution using *Serratia marcescens* (17). The other is the microbial stereoinversion method using racemate as a substrate (18). In the course of the studies on the microbial optical resolution of racemic 1,2-pentanediol, we found a unique microbial stereoinversion consisting of two successive reactions. It is a known fact that in all enzymatic or chemical optical resolution processes, the theoretical yield of one part of enantiomer is maximally 50%. However, (*S*)-enantiomer of 1,2-pentanediol was found higher than the theoretical yield when the racemate was incubated in the culture broth of *Lodderomyces elongisporus* IFO 1676. This result indicated that this microorganism could invert (*R*)-1,2-pentanediol in the racemate to (*S*)-enantiomer. This finding led us to rescreen the microorganisms that could invert (*R*)-1,2-pentanediol to the (*S*)-enantiomer, and we found that a number of microorganisms had this ability, similar to *L. elongisporus* IFO 1676. We applied this microbial stereoinversion to various 1,2-diols by use of *Candida parapsilosis* IFO 0708, and we confirmed that nine kinds of (*S*)-1,2-diols (**1** to **9**) and (*R*)-3-methylthio-1,2-pentanediol (**10**) with high optical purities could be obtained in high yields (Scheme 1).





Scheme 1.

EXPERIMENTAL

Chemicals

Racemic 1,2-heptanediol (**4**), 4-methyl-1,2-pentanediol (**6**), 3-phenyl-1,2-propanediol (**8**), and 4-phenyl-1,2-butanediol (**9**) were prepared from the corresponding 1-alkenes by a method similar to that of Blau et al. (19). Other 1,2-diols used were purchased from Aldrich Chemicals Company. 1-Hydroxy-2-pentanone was prepared from 1,2-pentanediol by microbial oxidation using *Saccharomycopisis lipolytica* IFO 1551; (bp 62 °C, 6 to 7 mmHg), NMR $\delta_H(CDCl_3)$: 0.95 (3H, t, J = 7.5 Hz, CH_3), 1.68 (2H, m, J = 7.3 Hz, -C**H**$_2$-CH_3), 2.44 (2H, t, J = 7.5 Hz, -C**H**$_2$-(C=O)-), 3.43 (1H, broad, OH), 4.29 (2H, s, -C**H**$_2$-OH), MS m/z: 102 (M^+), 71 (M^+ -CH_3O).

Microorganisms and Cultivation

All microorganisms used were stock cultures kept in our laboratory. The medium used contained 4% glucose, 0.3% yeast extract, and 10% (v/v) mineral solution, pH 7.0. The mineral solution consisted of 13% $(NH_4)_2HPO_4$, 7% KH_2PO_4, 0.8% $MgSO_4$ $7H_2O$, 0.1% NaCl, 600 ppm $ZnSO_4$ $7H_2O$, 900 ppm $FeSO_4$ $7H_2O$, 50 ppm $CuSO_4$ $5H_2O$, and 100 ppm $MnSO_4$ $4H_2O$. Microorganisms were cultured in a 500-mL Sakaguchi flask containing 50 mL of the medium at 30 °C for 1 to 2 days with shaking. When the cell grew maximally, the cells were harvested by centrifugation and were used for the reaction.

Screening of Microorganisms Producing (*S*)-1,2-Pentanediol by Stereoinversion

Cells were suspended in 50 mL of 0.1 M potassium phosphate buffer (pH 6.5) containing 250 to 500 mg of racemic 1,2-pentanediol. The reaction was conducted in 500 mL-Sakaguchi flasks that were incubated at 30 °C for 24 to 96 h with shaking. Two milliliters of the reaction mixture was used for GLC analysis of 1,2-pentanediol. The remainder was centrifuged to remove the cells, and the supernatant was concentrated to 3 to 4 mL. 1,2-Pentanediol was recovered by extraction with ethyl acetate (100 mL) and used for measurement of enantiomeric excess by HPLC.

Preparation of Cell-Free Extract and Enzyme Assay

Cells of *C. parapsilosis* IFO 0708, which grew for 24 h under the conditions described earlier, were collected and washed with 10 mM tris-HCl buffer (pH 7.0). Thirty-three grams of washed cells were suspended in 40 mL of the same buffer containing 1 mM EDTA and 0.2 mM phenylmethylsulfonylfluoride, a serine-protease inhibitor, and disrupted with a Braun homogenizer. Cell debris was removed by centrifugation (15,000 rpm × 30 min), and the supernatant was filtered through a millipore membrane (0.22 μm) and dialyzed against 10 mM tris-HCl buffer overnight. The cell-free extract was used for assay of the enzyme activities. The enzyme activities for (*R*) and (*S*)-1,2-pentanediol oxidation were assayed by measuring the increase of optical absorbances at 370 nm caused by the reduction of NAD^+ or $NADP^+$, respectively. The assay mixture in a final volume of 3.0 mL contained 3.0 μmol of $NAD(P)^+$, 0.3 mmol of (*S*) or (*R*)-1,2-pentanediol, 0.3 mmol of sodium carbonate buffer (pH 10), and a suitable amount of the cell-free extract. Incubation was performed at 25 °C in a cuvette of 1-cm light path length. The reduction of 1-hydroxy-2-pentanone was monitored by measuring the decrease of optical absorbency at 340 nm caused by the oxi-

dation of NADH or NADPH. The assay mixture in a final volume of 3.0 mL contained 1.0 μmol of NAD(P)H, 0.3 mmol of 1-hydroxy-2-pentanone, 0.3 mmol of tris-maleate buffer (pH 6.0), and a suitable amount of the cell-free extract. The incubation was carried out under the same conditions as the method for oxidation of 1,2-pentanediol.

Preparative-Scale Reaction and Isolation of Optically Active 1,2-Diols

C. parapsilosis IFO 0708 was cultured in a 5-L jar fermenter containing 3 L of the medium at 30 °C for 23 h (aeration 0.5 vvm, agitation 450 rpm, inoculum size 1.6%). After the cultivation, racemic 1,2-diol (9 to 30g) was added to the culture broth. During the reaction, the pH of the reaction mixture was controlled at 6.5 with 4 N NaOH, and 30 g of glucose was added once a day. After the reaction, cells were removed by centrifugation, and the supernatant was concentrated to about 150 mL under reduced pressure. 1,2-Diols were extracted with 500 mL of ethyl acetate three times. The solvent was removed by evaporation, and the 1,2-diols were purified by distillation or silica gel chromatography (Wako gel G-200, solvent: *n*-hexane/ethyl acetate).

Analytical Methods

1,2-Diols were assayed by GLC using a Hitachi 163 gas chromatograph with a hydrogen flame ionization detector. Two different types of liquid phases were used. The glass column (0.3 × 100 cm) packed with Silicone OV-17 on Chromosorb WAW DMCS (Gasukuro Kogyo, Japan) was used to assay **7,8** and **9**. Flow rate of nitrogen gas was 20 mL/min, and column temperature was 220 °C. The other glass column (0.3 × 100 cm) packed with FAL-M 6%/TENAX GC (Wako Pure Chemical, Japan) was used to assay the other 1,2-diols and 1-hydroxy-2-ketones. Flow rate of nitrogen gas was 35 mL/min, and the column temperature was 175 °C. The samples for GLC analyses were prepared as follows: 2 mL of the reaction mixture was saturated with ammonium sulfate, 1,2-diol was extracted with 4 mL of ethyl acetate, and the organic layer was used for GLC analysis.

The enantiomeric excess of 1,2-diol was measured by HPLC analysis of the corresponding 1-tosylated derivatives except for 1-phenyl-1,2-ethanediol. Preparation and purification of 1-tosylated derivatives were done by the procedures reported (5,20). The columns and the other conditions used are shown in Table 1.

NMR and mass spectra were recorded on a Varian EM-90 spectrometer and Shimadzu QP-1000 GC-Mass spectrometer, respectively. Optical rotation was measured with a Horiba polarimeter SEPA-200. Melting points were measured using a Büchi melting point apparatus.

Table 1. Conditions of HPLC for Measurement of Enantiomeric Excess of 1,2-Diols

1,2-Diol[a]	Column[b]	Solvent[c]	Flow rate (mL/min)	Retention time (min) (*R*)	(*S*)
1	OC	13:1	1.0	42	46
2	OC	13:1	1.0	36	41
3	OC	13:1	1.0	34	39
4	OC	13:1	1.0	30	36
5	OC	13:1	1.0	28	34
6[d]	OC	13:1	0.8	31	36
7	OB	30:1	0.7	33	43
8	OD	13:1	0.8	37	49
9	OD	13:1	0.8	39	51
10	OC[e]	20:1	1.0	78	72

[a]The numbers in this column were corresponding to the ones cited in Scheme 1.
[b]Daicel Chiral Cel (Ø 0.45 × 25 cm).
[c]The ratio of *n*-hexane/2-propanol.
[d]Only this compound was directly analyzed without 1-tosylation.
[e]Two columns were connected and used.

RESULTS

Screening of Microorganisms Producing (*S*)-1,2-Pentanediol by Stereoinversion

At the start of this study, we had planned the production of (*S*)-1,2-diols from racemates by stereoselective degradation of (*R*)-1,2-diols using microorganisms. We screened for suitable microorganisms for this purpose from the stock culture collection in our laboratory and found that many microorganisms degraded (*R*)-1,2-pentanediol more easily than (*S*)-enantiomer. *Lodderomyces elongisporus* IFO 1676 was selected among them and used for large-scale preparation of (*S*)-1,2-pentanediol. In this experiment, we found that the optical purity of residual 1,2-pentanediol reached approximately 100% ee as (*S*)-enantiomer when only 22% of racemic 1,2-pentanediol added in the reaction mixture was degraded. This result suggested that this microorganism could convert (*R*)-1,2-pentanediol to the (*S*)-enantiomer. This observation led us to rescreen the microorganisms that had the ability to produce (*S*)-1,2-pentanediol having high optical purity from the racemate by stereoinversion, and we found that various microorganisms had the similar ability as *L. elongisporus*. They belonged to the genera *Candida*, *Debaryomyces*, *Pichia*, *Stephanoascus*, *Endomyces*, *Fusarium*, *Giberella*, *Gongronella*, and others. Among them, the microorganisms that produced (*S*)-1,2-pentanediol excess of 70% molar yield from the racemate are listed in Table 2. *C. parapsilosis* IFO 0708 was used for further experiments.

Production of (*S*)-1,2-Pentanediol by *C. parapsilosis* IFO 0708

C. parapsilosis IFO 0708 was cultured in a 5-L jar fermenter containing 3 L of medium for 23 h under the conditions described earlier. After then, 75 g of racemic 1,2-pentanediol and 25 g of glucose were added to 2.5 L of the culture broth in a 5-L jar fermenter, and the pH of the reaction mixture was adjusted to 6.5 with 4 N NaOH. The reaction-accompanied cultivation was conducted for 26 h under the same conditions. Figure 1 shows the course of the reaction. Racemic 1,2-pentanediol was converted to (*S*)-1,2-pentanediol with a molar yield of 93% after incubation for 24 h. It was noteworthy that 1-hydroxy-2-pentanone was de-

Table 2. (*S*)-1,2-Pentanediol Production from the Racemate by Microbial Stereoinversion

Microorganism	Incubation period (h)	Residual 1,2-PDO[a] (A) (%)	ee of 1,2-PDO (B) (%)	(*S*)-Isomer yield[b] (C) (%)
Candida maltosa ATCC 20275	48	75	95 (*S*)	73
Candida parapsilosis IFO 0585	24	77	97 (*S*)	76
Candida parapsilosis IFO 0708	24	89	98 (*S*)	88
Candida parapsilosis IFO 1396	24	89	96 (*S*)	87
Candida tropicalis CBS 1926	24	92	91 (*S*)	88
Lodderomyces elongisporus IFO 1676	24	79	93 (*S*)	76
Pichia bovis IFO 0872	24	77	88 (*S*)	72
Gongronella butleri IFO 8080	96	84	71 (*S*)	72

[a]At the start of reaction, 10 mg/mL of racemic 1,2-pentanediol (1,2-PDO) was added except for *G. butleri* (5 mg/mL).
[b](*S*)-Isomer yield (C) = Net yield of (*S*)-isomer from the racemate = A × [1 − (100 − B)/200].

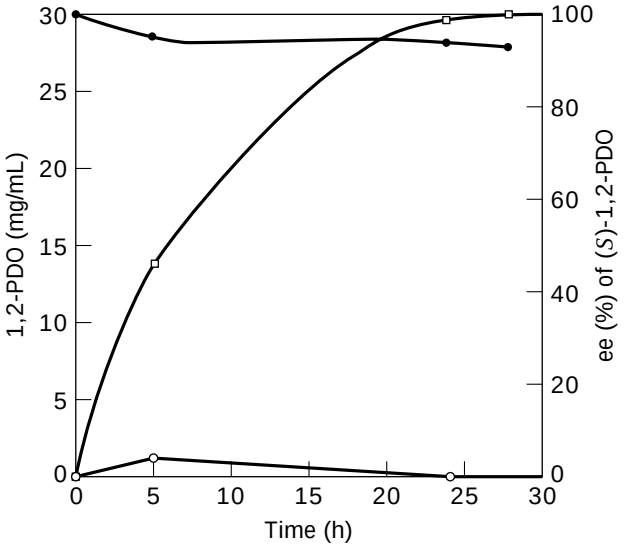


Figure 1. Production of (*S*)-1,2-pentanediol by *C. parapsilosis* IFO 0708. ● = 1,2-pentanediol; ○ = 1-hydroxy-2-pentanone; □ = enantiomeric excess of (*S*)-1,2-pentanediol.

tected in the reaction mixture. After the reaction, the cells were removed by centrifugation, and the supernatant was concentrated to 375 mL under reduced pressure. (*S*)-1,2-Pentanediol was extracted with 750 mL of ethyl acetate. The solvent was dried up by evaporation, and 66 g of yellowish oil was obtained. (*S*)-1,2-Pentanediol was purified by distillation (76 to 77 °C [3 mmHg]), and 60 g of (*S*)-1,2-pentanediol was obtained (yield 80%). GLC purity 99.9%, $[\alpha]_D^{20}$ −17.3° (c = 1, MeOH), −23.2°(c = 1, EtOH), 100% ee, (lit.(14), $[\alpha]_D^{25}$ −16.1°(c = 3, EtOH)), NMR δ_H($CDCl_3$): 0.92 (3H, t, CH_3), 1.13 to 1.83 (4H, m, -$(CH_2)_2$-), 3.27 to 3.92 (3H, m, -C**H**(OH)-C**H**$_2$OH), 4.12 (2H, broad, OH), MS *m*/*z*: 87 (M^+-OH), 73 (M^+-CH_3O), 61 ($C_2H_5O_2^+$).

Mechanism of Stereoinversion of 1,2-Pentanediol in *C. parapsilosis* IFO 0708

To obtain the information on the stereoinversion mechanism, we used the cell-free extract of *C. parapsilosis*. Usually, alcohol dehydrogenase requires NAD^+ or $NADP^+$ as a coenzyme to catalyze the oxidation of alcohols. As the first experiment, we investigated the requirement of coenzyme for the oxidoreduction between (*R*)/(*S*)-1,2-pentanediol and 1-hydroxy-2-pentanone in this microorganism. As shown in Table 3, both NADH and NADPH served as a coenzyme to reduce 1-hydroxy-2-pentanone to 1,2-pentanediol in the cell-free extract. In the reverse reaction, NAD^+ mainly played the coenzyme role in oxidation of (*R*)-1,2-pentanediol, and the reaction rate of the oxidation of (*R*) and (*S*)-1,2-pentanediol using $NADP^+$ as a coenzyme was very slow. This NAD^+-linked (*R*)-specific 1,2-pentanediol dehydrogenase resembles the glycerol dehydrogenase isolated from *Cellulomonas* sp. (5,21) in its stereospecificity and NAD^+ requirement. Sugiura et al. reported that NAD^+-linked glycerol dehydrogenase from *Erwinia aroideae* (22) was inhibited by $CuSO_4$ and $ZnSO_4$. The effects of these metals and *p*-chloromercuribenzoate on both NADH-linked and NADPH-linked 1-hydroxy-2-pentanone reductase activities were tested. It was found that 1 mM $CuSO_4$ and 0.17 mM *p*-chloromercuribenzoate inhibited NADH-linked 1-hydroxy-2-pentanone reductase very strongly. In order to know the stereospecificity of NADPH-linked 1-hydroxy-2-pentanone reductase, 3 mM of $CuSO_4$ was added into the reaction mixture (10 mL) containing 0.50 mmol NADPH, 0.24 mmol of 1-hydroxy-2-pentanone, and the cell-free extract (2 mL). The reduction was carried out for 16 h at 25 °C. The steric configuration of 1,2-pentanediol produced under the above conditions was the

Table 3. Oxidoreduction between 1,2-Pentanediol and 1-Hydroxy-2-Pentanone by the Cell-Free Extract of *C. parapsilosis*[a]

Reaction and substrate	OD at 340 nm[a]	
Reduction (pH 6.0)	NADH	NADPH
1-Hydroxy-2-pentanone	−0.667	−0.232
Oxidation (pH 10.0)	NAD^+	$NADP^+$[b]
(*R*)-1,2-Pentanediol	0.722	0.014
(*S*)-1,2-Pentanediol	0.005	0.018

Note: The conditions of oxidation and reduction were described in the text.
[a]The change of optical density at 340 nm during the first 1 min of the reaction is listed.
[b]In the case of oxidation using $NADP^+$, the reactions were carried out at pH 8.0 using 0.1 M tris-HCl buffer instead of 0.1M sodium carbonate buffer (pH 10.0), because the reduction of $NADP^+$ was not observed at all under these conditions.

(*S*)-form, and its optical purity was very high (Table 4). The activities of NAD^+-linked and $NADP^+$-linked 1,2-pentanediol dehydrogenases at various pHs are shown in Figure 2. Oxidoreduction between 1,2-pentanediol and 1-hydroxy-2-pentanone by NAD^+-linked dehydrogenase is reversible, but reduction of 1-hydroxy-2-pentanone to (*S*)-1,2-pentanediol is practically less reversible. From these observations, the mechanism of the stereoinversion of 1,2-diol in *C. parapsilosis* may be explained as follows. Only (*R*)-1,2-pentanediol in the racemate is oxidized to 1-hydroxy-1,2-pentanone by NAD^+-linked (*R*)-specific dehydrogenase. Then, 1-hydroxy-2-pentanone is reduced to (*S*)-1,2-pentanediol by NADPH-linked (*S*)-specific reductase (Scheme 2).

Table 4. Stereospecificity of NADPH-Linked 1-Hydroxy-2-Pentanone Reductase

		Formed 1,2-pentanediol	
Coenzyme	Inhibitor $CuSO_4$ (mM)	Yield (%)	ee (%)
NADH	0	89	84 (*R*)
	3	0	—
NADPH	0	73	51 (*S*)
	3	31	98 (*S*)

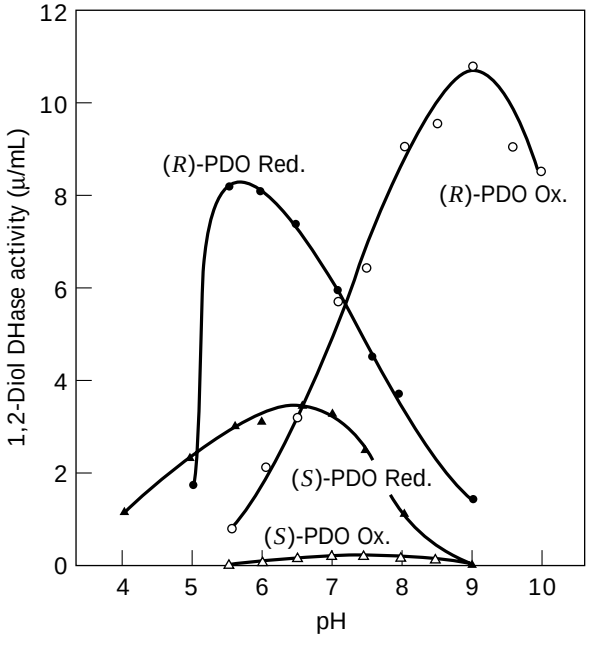
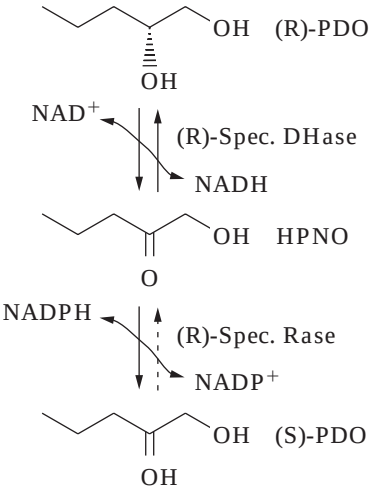


Figure 2. Effect of pH on 1,2-pentanediol oxidoreductase activities. The basic composition of the reaction mixtures was described in the text, and the following buffers were used to maintain the respective pHs: pH 4.0 to 5.5 acetate buffer, pH 5.5 to 6.5 tris-maleate buffer, pH 6.5 to 7.5 phosphate buffer, pH 7.5 to 9.0 tris-HCl buffer, pH 9.0 to 11.0 sodium carbonate buffer. ○ = NAD^+, (*R*)-1,2-pentanediol oxidation; ● = NADH, 1-hydroxy-2-pentanone reduction; △ = $NADP^+$, (*S*)-1,2-pentanediol oxidation; ▲ = NADPH, 1-hydroxy-2-pentanone reduction; (1 unit = 1 μmol of $NAD(P)^+$ reduction or NAD(P)H oxidation per min).



Scheme 2.

Production of Various Optically Active 1,2-Diols by Stereoinversion Using *C. parapsilosis* IFO 0708

It was very beneficial for us to apply this unique microbial reaction to production of various optically active 1,2-diols. Fourteen kinds of racemic 1,2-diols were tested: 1,2-propanediol, 1,2-butanediol (**1**), 1,2-pentanediol (**2**), 1,2-hexanediol (**3**), 1,2-heptanediol (**4**), 1,2-octanediol (**5**), 4-methyl-1,2-pentanediol (**6**), 1-phenyl-1,2-ethanediol (**7**), 3-phenyl-1,2-propanediol (**8**), 4-phenyl-1,2-butanediol (**9**), 3-methylthio-1,2-propanediol (**10**), 3-chloro-1,2-propanediol, 3-methoxy-1,2-propanediol, and 3-phenoxy-1,2-propanediol. Among them, the microbial stereoinversion was observed on 10 kinds of 1,2-diols listed in Table 5. These 1,2-diols had the same absolute configuration on the secondary alcohol. The yields and properties of optically active 1,2-diols obtained were as follows.

(*S*)-1,2-Butanediol (1). (*S*)-Isomer yield 56%, bp 65 °C (3 mmHg), (lit.(6), bp 94 to 96 °C (9 mmHg)), GLC purity 99.9%, $[\alpha]_D^{20}$ −8.57° (c = 1, MeOH), −31.9° (c = 1, EtOH), 79% ee, (lit.(6), $[\alpha]_D^{25}$ −15.35° (c = 2.6, EtOH)). NMR δ $_H(CDCl_3)$: 0.92 (3H, t, CH_3), 1.45 (2H, m, -C**H**$_2$-CH_3), 3.23 ~ 3.83 (3H, m, -C**H**(OH)-C**H**$_2$OH), 4.03 (2H, broad, OH), MS *m*/*z*: 61 ($C_2H_5O_2^+$), 59 (M^+-CH_3O).

(*S*)-1,2-Hexanediol (3). (*S*)-Isomer yield 95%, bp 86 °C (3 mmHg), (lit.(1), bp 110 to 113 °C (6 mmHg)), GLC purity 99.7%, $[\alpha]_D^{20}$ −16.4° (c = 1, MeOH), −22.1° (c = 1, EtOH), 100% ee, (lit.(1), (*R*)-isomer $[\alpha]_D^{22}$ + 15.2° (c = 14, EtOH)), NMR δ $_H(CDCl_3)$: 0.92 (3H, t, CH_3), 1.13 to 1.83 (6H, m, -$(CH_2)_3$-), 3.23 to 4.23 (3H, m, -C**H**(OH)-C**H**$_2$OH), 3.73 (2H, broad, OH), MS *m*/*z*: 87 (M^+-CH_3O), 61 ($C_2H_5O_2^+$).

(*S*)-1,2-Heptanediol (4). (*S*)-Isomer yield 97%, bp 101 °C (4 mmHg), (lit.(2), 67 to 70 °C (0.04 mmHg)), GLC purity 99.6%, $[\alpha]_D^{20}$ − 15.4° (c = 1, MeOH), −20.6° (c = 1, EtOH), 98% ee, (lit.(2), $[\alpha]_D^{22}$ − 16.6° (c = 11.9, EtOH)), NMR δ $_H(CDCl_3)$: 0.87 (3H, t, CH_3), 1.10 to 1.87 (8H, m, -$(CH_2)_4$-), 3.22 to 4.13 (3H, m, -C**H**(OH)-C**H**$_2$OH), 3.75 (2H, broad, OH), MS *m*/*z*: 101 (M^+-CH_3O), 61 ($C_2H_5O_2^+$).

Table 5. Production of Optically Active 1,2-Diols by *C. parapsilosis* IFO 0708

1,2-Diol[a]	Concentration of substrate (g/L)	Reaction period (h)	Residual 1,2-diol (%)	e,e of 1,2-diol (%)	(*S*) or (*R*) isomer yield[b] (%)
1	10	54	62	79 (*S*)	56
2	30	26	93	100 (*S*)	93
3	10	16	95	100 (*S*)	95
4	5.0	4	98	98 (*S*)	97
5	3.0	9	96	97 (*S*)	95
6	3.0	23	97	97 (*S*)	96
7	5.0	30	100	100 (*S*)	100
8	6.3	72	72	82 (*S*)	66
9	6.3	72	100	100 (*S*)	100
10	3.0	54	66	98 (*R*)	65

[a]The numbers in this column correspond to the numbers in Scheme 1.
[b]See Table 2.

(*S*)-1,2-Octanediol (5). (*S*)-Isomer yield 95%, bp 104 °C (3 mmHg), (lit.(15), 100 to 110 °C (0.4 mmHg)), GLC purity 99.8%, $[\alpha]_D^{20}$ − 13.6° (c = 1, MeOH), −19.0° (c = 1, EtOH), 97% ee, (lit.(15), (*R*)-isomer $[\alpha]_D^{20}$ + 17.5° (c = 1.164, EtOH)). NMR δ_H($CDCl_3$): 0.90 (3H, t, CH_3), 1.10 to 1.93 (10H, m, -$(CH_2)_5$-), 3.23 to 3.93 (3H, m, -C**H**(OH)-C**H**$_2$OH), 3.05 (2H, broad, OH), MS *m*/*z*): 115 (M^+-CH_3O), 61 ($C_2H_5O_2^+$).

(*S*)-4-Methyl-1,2-Pentanediol (6). (*S*)-Isomer yield 96%, bp 82 to 83 °C (5 mmHg), (lit.(10), 89 to 91 °C (5 mmHg)), GLC purity 99.4%, $[\alpha]_D^{20}$ − 26.3° (c = 1, MeOH, −31.5° (c = 1, EtOH), (lit.(10) $[\alpha]_D^{23}$ −24.7° (c = 1.84, EtOH), 97% ee, NMR δ_H($CDCl_3$): 0.93 (6H, d, *J* = 6.6Hz, (CH_3)$_2$-CH-), 1.27 (2H, m, CH-C**H**$_2$-), 1.80 (1H, m, (CH_3)$_2$-C**H**-, 3.20 ∼ 4.05 (3H, m, -C**H**(OH)-C**H**$_2$OH), 3.68 (2H, broad, OH), MS *m*/*z*: 117 (M^+ − 1), 87 (M^+-CH_3O), 61 ($C_2H_5O_2^+$), 57 ((CH_3)$_2C_2^+$).

(*S*)-1-Phenyl-1,2-Ethanediol (7). (*S*)-Isomer yield 100%, mp 66 to 67 °C (lit.(7), (*S*)-isomer 63 to 65 °C), GLC purity 99.9%, $[\alpha]_D^{20}$ +45.0° (c = 1, MeOH), +38.6° (c = 1, EtOH), +49.9° (c = 1, acetone), 100% ee, (lit.(4), $[\alpha]_D^{25}$ +51.2° (c = 0.25, acetone), lit(7) (*R*)-isomer $[\alpha]_D^{25}$ −39.7° (c = 4.33, 95% EtOH)), NMR δ_H($CDCl_3$): 2.97 to 3.30 (1H, m, OH), 3.40 to 3.60 (1H, m, OH), 3.63 (2H, d, -CH_2-), 4.75 (1H, m, -C**H**(OH)-), 7.30 (5H, m, C_6H_5-), MS *m*/*z*: 138 (M^+), 120 (M^+-H_2O), 107 (M^+-CH_2O).

(*S*)-3-Phenyl-1,2-Propanediol (8). (*S*)-Isomer yield 66%, mp 47 to 48 °C, bp 140 °C (6 mmHg), (lit.(11), 101 to 102 °C (0.001 mmHg)), GLC purity 99.6%, $[\alpha]_D^{20}$ −28.2° (c = 1, MeOH), −29.5° (c = 1, EtOH), 82% ee, (lit.(11), $[\alpha]_D^{20}$ −36° (c = 1, EtOH)), NMR δ_H($CDCl_3$): 2.68 (2H, d, *J* = 6.6Hz, C_6H_5-C**H**$_2$-), 3.03 (2H, broad, OH), 3.25 to 3.73 (2H, m, -C**H**$_2$OH), 3.73 ∼ 4.10 (1H, m, -C**H**(OH)-), 7.22 (5H, m, C_6H_5-), MS *m*/*z*: 152 (M^+), 134 (M^+-OH), 133 (M^+-H_2O), 91 (C_6H_5-CH_2^+), 61 ($C_2H_5O_2^+$).

(*S*)-4-Phenyl-1,2-Butanediol (9). (*S*)-Isomer yield 100%, mp 36 to 37 °C, bp 142 °C (4 mmHg), (lit.(19), the racemate 162 to 165 °C (0.65 kPa)), GLC purity 98.8%, $[\alpha]_D^{20}$ −29.0° (c = 1, MeOH), −34.1° (c = 1, EtOH), 100% ee, NMR δ_H($CDCl_3$): 1.72 (2H, m, -C**H**$_2$-CH(OH)-), 2.68 (2H, m, C_6H_5-C**H**$_2$-), 3.03 ∼ 4.00 (3H, m, -C**H**(OH)-C**H**$_2$OH), 3.47 (2H, broad, OH), 7.23 (5H, m, C_6H_5-), MS *m*/*z*: 149 (M^+-OH), 148 (M^+-H_2O), 135 (M^+-CH_3O), 105 (C_6H_5-$CH_2CH_2^+$), 61 ($C_2H_5O_2^+$).

(*R*)-3-Methylthio-1,2-Propanediol (10). (*R*)-Isomer yield 65%, bp 101 °C (3 mmHg), GLC purity 99.7%, $[\alpha]_D^{20}$ −12.9° (c = 1, MeOH), −11.9° (c = 1, EtOH), 98% ee, MNR δ_H($CDCl_3$): 2.13 (3H, s, CH_3-), 2.60 (2H, d, *J* = 6.0Hz, -S-CH_2-), 3.18 ∼ 4.32 (3H, m, -C**H**(OH)-C**H**$_2$OH), 3.83 (2H, broad, OH), MS *m*/*z*: 122 (M^+), 104 (M^+-H_20), 61 (CH_3-S-CH_2^+, $C_2H_5O_2^+$).

We tried this microbial stereoinversion for production of optically active 1,2-propanediol, 3-phenoxy-1,2-propanediol, 3-methoxy-1,2-propanediol, and 3-chloro-1,2-propanediol, but these 1,2-diols were not good substrates. In cases of 1,2-propanediol and 3-phenoxy-1,2-propanediol, their (*S*)-isomers were isolated from the reaction mixtures, but the yields and the optical purities of both (*S*)-1,2-diols were not over 50%. With the other 1,2-diols, their racemates were recovered from the reaction mixtures.

In this study, we presented a novel method for production of optically active 1,2-diols using a microbial stereoinversion. This unique bioinversion consisted of two successive reactions via 1-hydroxy-2-ketone, namely the oxidation of (*R*)-1,2-diol to 1-hydroxy-2-ketone by NAD^+-linked (*R*)-specific alcohol dehydrogenase, followed by the reduction of 1-hydroxy-2-ketone to (*S*)-1,2-diol by an NADPH-linked (*S*)-specific reductase. To the best of our knowledge, this is the most convenient procedure for producing optically active 1,2-diols among the reported methods.

BIBLIOGRAPHY

1. P.A. Levene and H.L. Haller, *J. Biol. Chem.* **74**, 343 (1927), *J. Biol. Chem.* **77**, 555 (1928); *J. Biol. Chem.* **79**, 475 (1928).
2. J. Barry and H.B. Kagan, *Synthesis*, 453 (1981).
3. D.D. Ridley and M. Stralow, *J. Chem. Soc., Chem. Commun.*, 400 (1975).

4. H. Kutsuki, I. Sawa, N. Mori, J. Ishizu, J. Hasegawa, M. Shimazaki, T. Ohashi, and K. Watanabe, *Ann. Mtg. Soc. Ferment. Technol. Japan*, 106 (1986).
5. L.G. Lee and G.M. Whitesides, *J. Org. Chem.* **51**, 25 (1986).
6. H.K. Chenault, M-J. Kim, A. Akiyama, T. Miyazawa, E.S. Simon, and G.M. Whitesides, *J. Org. Chem.* **52**, 2608 (1987).
7. J.A. Dale and H.S. Mosher, *J. Org. Chem.* **35**, 4002 (1970).
8. K. Hintzer, B. Koppenhoefer, and V. Shurig, *J. Org. Chem.* **47**, 3850 (1982).
9. N. Sakota, S. Tanaka, K. Okita, and N. Koine, *Nippon Kagaku Zasshi* **91**, 265, (1970).
10. K. Mori, *Tetrahedron Lett.* **26**, 2187 (1975).
11. W. Bergstein, A. Kleemann, and J. Martens, *Synthesis*, 76 (1981).
12. B. Cambou and A.M. Klibanov, *J. Am. Chem. Soc.* **106**, 2687 (1984).
13. T. Sugiyama, H. Sugawara, M. Watanabe, and K. Yamashita, *Agric. Biol. Chem.* **48**, 1841 (1984).
14. J. Mulzer and A. Angermann, *Tetrahedron Lett.* **24**, 2843 (1983).
15. K-Y. Ko, W.J. Frazee, and E.L. Eliel, *Tetrahedron* **40**, 1333 (1984).
16. S. Maemoto, H. Kutsuki, J. Hasegawa, and T. Ohashi, *Ann. Mtg. Agric. Biol. Chem. Soc. of Japan*, 419 (1989).
17. H. Takahashi, S. Tsuda, Y. Nakamura, M. Ogura, T. Shiraishi, and Y. Shimada, *Ann. Mtg. Agric. Biol. Chem. Soc. of Japan* 351 (1989).
18. J. Hasegawa, M. Ogura, S. Tsuda, S. Maemoto, H. Kutsuki, and T. Ohashi, *Agric. Biol. Chem.* **54**, 1819 (1990).
19. K. Blau, V. Voerckel, and L. Willecke, *J. Prakt. Chem.* **328**, 29 (1986).
20. J.P. Guett and N. Spassky, *Bull. Soc. Chim. Fr.* 4217 (1972).
21. H. Yamada, A. Nagao, H. Nishise, and Y. Tani, *Agric. Biol. Chem.* **46**, 2333 (1982).
22. M. Sugiura, T. Oikawa, K. Hirano, H. Shimizu, and F. Hirata, *Chem. Pharm. Bull.* **26**, 716 (1978).

ORGANIC COMPOUNDS, CELLULOSE CONVERSION

C.S. GONG
NINGJUN CAO
G.T. TSAO
Purdue University
West Lafayette, Indiana

KEY WORDS

Cellulase
Cellulose
Ethanol
Hemicellulose
Pretreatment
Simultaneous saccharification and fermentation (SSF)
Xylose

OUTLINE

INTRODUCTION

Cellulose is the most plentiful of all of the naturally occurring organic compounds. It is a high molecular weight linear polysaccharide composed of two residues (β-1,4-linked glucose, known as cellobiose) in repeated units. Complete hydrolysis of cellulose using strong mineral acids yields glucose. The basic building blocks of cellulose are chemically reactive, making it susceptible to chemical modifications. Cellulose has great economical importance; it is processed to produce papers, fibers, and chemical derivatives to yield substances used in the manufacturing of such items as plastics, photographic films, rayon, and so on. Other cellulose derivatives are used as adhesives, explosives, and thickening agents for foods. Figure 1 shows some common uses of cellulose.

Cellulose exists in nature in two principal types: the pectocelluloses such as flax, hemp, and ramie that contain more than 80% cellulose; the lignocelluloses composed mostly of cellulose, hemicellulose, and lignin. As a basic component of plant cell walls, cellulose comprises at least 30% of all plant matters (90% of cotton and up to 50% of wood are cellulose). Table 1 shows the approximate composition of the selected lignocellulosic materials.

Because lignocellulosic materials are renewable and abundant, extensive research in recent years has been undertaken to convert cellulose, particularly agricultural residues, forest by-products, and waste cellulose, into food, chemicals, and fuels. In this article, the emphasis will be on the production of ethanol from lignocellulosic materials. Other products such as solvents (acetone and butanol) (7), 2,3-butanediol (8), and single-cell proteins (9) have been reviewed.

Relationship between Structural Features of the Cell Wall and Cellulolytic Enzymes

Native lignocellulosic materials are heterogeneous, containing cellulose, hemicellulose, lignin, ash, protein, and a

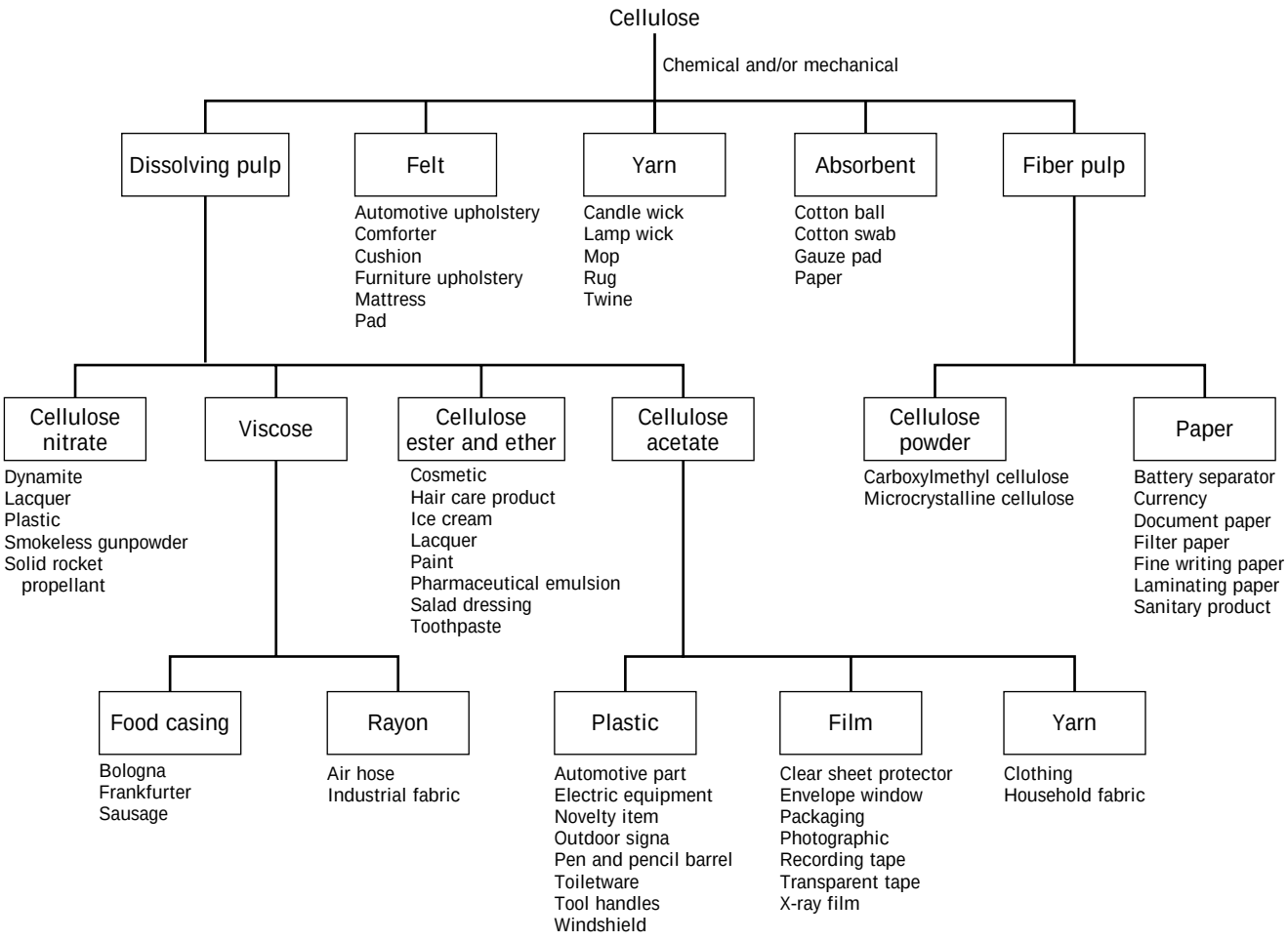


Figure 1. Common uses of cellulose. *Source*: From Ref. 1.

wide array of minor extractives. The most important structural feature in a plant cell wall is the capillary area, whose dimensions and shapes control the accessibility of lignocelluloses by hydrolytic enzymes. The lumen (Fig. 2) is the largest capillary in a plant cell, followed by pit apertures and pores in the pit-membrane, ranging from 200 Å to 10 or more μm in diameter. Most of these capillaries are filled with hemicellulose, pectic materials, and lignin, which form a protective sheath around the cellulose (10,11). Therefore, any treatment of lignocellulosic materials that solubilizes and/or removes lignin and hemicelluloses from the cell wall can enlarge the capillaries. As a result, the accessibility of cellulose to cellulase can be increased.

Molecules of cellulolytic enzymes of various microorganisms are water-soluble proteins of high molecular weight that are 13 $\times$ 79 Å in width and 42 $\times$ 252 Å in length. The relatively large enzyme molecules cannot be diffused into plant cell walls unless pretreatment is conducted to increase the size of the cell wall capillaries.

CELLULOSE PRETREATMENT

Many factors have been identified as influencing the reactivity and digestibility of the cellulose fraction of lignocellulose materials. These factors include lignin and hemicellulose contents, crystallinity of the cellulose, and the porosity of the biomass. Pretreatment of lignocellulosic materials prior to utilization is a necessary element in the biomass-to-ethanol conversion processes. The objective of the pretreatment is to render biomass materials more accessible to either chemical or enzymatic hydrolysis for efficient product generation. There are several goals of the pretreatment.

- Disrupt and remove the lignin sheath
- Remove and separate hemicellulose from cellulose
- Decrease the crystallinity of the cellulose
- Increase the accessible surface area
- Increase the pore size to facilitate the penetration of hydrolysis agents

There have been many processes developed for the treatment of various biomass. In general, agricultural residues are more easily treated than hardwood materials. The most difficult biomass to treat is that derived from softwood because it has a relatively high lignin content. Most processes exploit a variety of mechanisms to render

Table 1. Approximate Composition of Selected Plant Materials and Cellulose Waste

Materials	Cellulose (%)	Hemicellulose (%)	Lignin (%)	Ref.
Softwood				
Spruce	43	26	29	2
Pine	44	26	29	2
Hardwood				
Aspen	46	26	18	3
Black locust	49	21	22	4
Birch	38	27	23	3
Hybrid poplar	43	21	26	
Red oak	38	27	23	3
Silver maple	46	23	21	4
Sycamore	44	22	23	4
Willow	37	23	21	2
Crop residues				
Corn cobs	45	35	15	5
Corn stalks	35	25	35	5
Corn stover	41	21	17	4
Sugar cane bagasse	40	30	20	5
Wheat straw	36	28	29	2
Others				
Alfalfa hay	38	9	14	6
Nut shells	25	25	30	
Cotton	90	5	0	
Cellulose Wastes				
Paper	85	0	15	
Newsprint	61	16	21	2
Sorted refuse	60	20	20	

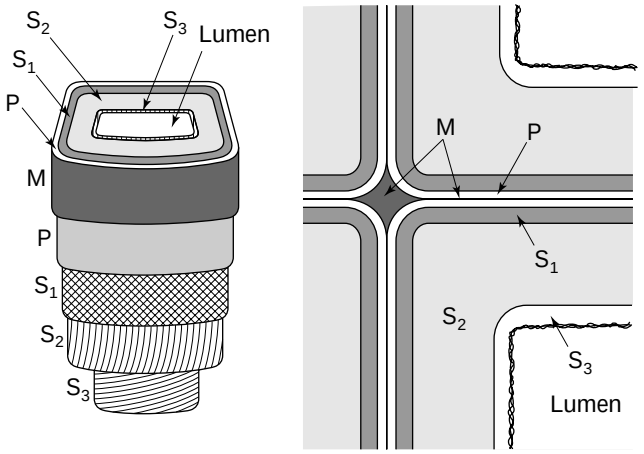


Figure 2. Diagrammatic representation of various layers of the cell wall. The intercellular materials: M, middle lamella; P, primary walls; S, secondary wall with outer (S1), middle (S2), and inner (S3) layers. *Source*: From Ref. 10.

the carbohydrate components of lignocellulosic materials more susceptible to enzymatic hydrolysis and microbial conversion (for detail see refs. 10 and 12).

The majority of the pretreatment methods involve a combination of mechanical size reduction, alkali swelling, acid hydrolysis, steam and other fiber explosion techniques, or exposure to supercritical fluids. The processes involving irradiation, strong mineral acids, cellulose solvents, and concentrated alkali chemicals are not suitable methods for the pretreatment of lignocellulosic materials because of their high cost. Biological methods, such as those applying "white rot" fungi to remove lignin from cellulose-hemicellulose, are also unsuitable due to the length of time they take, unless the lignocellulosic material is intended for single-cell protein (e.g., mushrooms) production (13). Therefore, the chosen pretreatment process must be efficient, cost effective, and environmental friendly.

Many different pretreatment approaches have been designed and tested, and some processes have also been tried in pilot scale. Ideally, the most desirable method of treatment is the dissolution of the solid materials into aqueous substrate. Many cellulose solvents at high enough concentrations can penetrate into the cellulose crystalline structure, causing the dissolution of solid cellulose. The resulting materials can be readily hydrolyzed to glucose by cellulase either in situ, or after the dissolved cellulose has been reprecipitated (14). Although cellulose solvents are powerful-solubilization agents, they are toxic, corrosive, and hazardous, and recovery costs are high. For these reasons, cellulose solvents are generally considered inappropriate for large-scale pretreatment processes.

Alkaline Treatment

Alkaline treatment causes lignocellulosic materials to swell; increased swelling leads to higher susceptibility of cellulose to saccharification. In the presence of alkaline chemicals (e.g., NaOH or NH_3), cellulose, hemicellulose, and lignin bonds can be disrupted, which permits cellulose to swell beyond normal water-swollen stages. Consequently, the pore size, intraparticle porosity, and capillary size are increased. There is also a phase change in the cellulose crystalline structure (15).

Sodium hydroxide is the most commonly used chemical in the treatment of lignocellulose to enhance digestibility for animals. It is a strong swelling agent for cellulose and may also be responsible for changing cellulose crystalline structure from cellulose I to cellulose III (decrystallization). The amount of NaOH used for treatment ranges from 2 to 20%, and the temperature for the treatment ranges from ambient to 120 °C. Under mild conditions (low concentration and low temperature) substrate components remain unchanged. Under harsher conditions, most of the lignin and hemicellulose are solubilized. Consequently, the cellulose fraction is exposed to hydrolysis agents.

Similar to NaOH, ammonia is useful in increasing the in vitro digestibility of lignocellulosic materials. It is also effective in the preparation of substrates for single-cell protein production. Ammonia is one of the most heavily used industrial chemicals. On a weight basis, aqueous ammonia is about one-third the cost of NaOH. In addition, ammonia is easily recoverable from an aqueous mixture because of its high volatility; it is also far less corrosive than mineral acids at high temperature. Increasingly, ammonia has been used in combination with other methods to treat lignocellulosic materials.

Ammonia-Recycled Percolation Process. The ammonia-recycled percolation (APR) process (16) uses ammonia as

the treatment agent. In one example, the reactor was a packed-bed flow-through type (percolation reactor) and was used in a recirculation mode (Fig. 3). The intended biomass material was packed in a pressure vessel with preheated liquid ammonia passing through the substrate. The reaction temperature was 150 °C, and pressures were up to 325 psi. As the name of the process indicates, ammonia was recycled for further treatment of cellulosic materials. The scanning electron microscope examination of treated samples showed an increase in pore size and porosity. The amount of lignin removed was in the range of 23–63%, and hemicellulose removed was in the range of 20–36%.

Alkaline Peroxide Process. Chopped wheat straw was steeped in an aqueous alkaline solution (pH = 11.5) of H_2O_2 at a ratio of 0.25 g/g substrate at 25 °C. Lignin removal was 50%, and hemicellulose solubilized was almost 100%. The results of *Trichoderma* cellulase hydrolysis of cellulose indicated that glucose yield was approaching 100% with a substrate solid content of 5% (17).

Ammonia Freeze Explosion. The ammonia freeze explosion (AFEX) process is the pretreatment method that utilizes steam explosion techniques with ammonia as the chemical reagent (6,18). The intended biomass material is placed in a pressure vessel with liquid ammonia (1:1) and treated similarly as in the steam explosion process but with much lower temperatures (up to 80 °C) and pressures (up to 15 psi). Using this process, ammonia was recovered. Exposure to ammonia in the AFEX process causes swelling and partial decrystallization of the cellulose. It also caused the partial solubilization of lignin and hemicellulose resulting in the exposure of cellulose. This pretreatment method is more effective with agricultural residues (e.g., corn stover and straws) than with woody materials.

Steam Treatment with or without SO_2

Steam explosion of biomass materials has been studied by many investigators (see Ref. 19 for review). Biomass material (woodchips or wheat straw) was placed in a high-pressure stainless steel tube and was exposed to steam under pressures ranging from 250 to 650 psi at 200–240 °C for up to 20 min. The sudden pressure release caused an explosion of biomass material, thereby disrupting the lignin and hemicellulose bonding toward cellulose. The addition of SO_2 enhanced the pretreatment effect and also increased the recovery of cellulose (20–22).

Dilute Acid Pretreatment

The use of dilute acid hydrolysis pretreatment has been studied extensively (4,23). The overall goal of dilute acid pretreatment is to attain a high yield while minimizing the breakdown of hemicellulose sugars into decomposition products. The pretreatment usually involves mineral acids (e.g., HCl and H_2SO_4) at a concentration from 0.3 to 2% (w/w). The temperature for treatment ranged from 120 to 180 °C, and the length of treatment ranged from less than 1 min to 2 h or longer. The objective in dilute acid treatment is to separate hemicellulose from lignin and cellulose. It is effective in hydrolyzing hemicellulose, but a portion of lignin was also solubilized, resulting in the complication of utilization of hemicellulose carbohydrates. Depending on hydrolysis conditions, carbohydrate degradation products also formed, which interfered with microbial activities.

Because the hemicellulose fraction of biomass materials can be separated from lignin and cellulose by dilute acid treatment, cellulose becomes more reactive toward cellulase. Hemicellulose hydrolysis rates vary with acid concentration, temperature, and solid-to-liquid ratio. With most lignocellulosic materials, complete hemicellulose hydrolysis can be achieved in 5–10 min at 160 °C, or 30–60 min at 140 °C. Dilute acid hydrolysis forms the basis of many pretreatment processes; autohydrolysis and steam explosion are two examples that are based on high-temperature-dilute-acid-catalyzed hydrolysis of biomass.

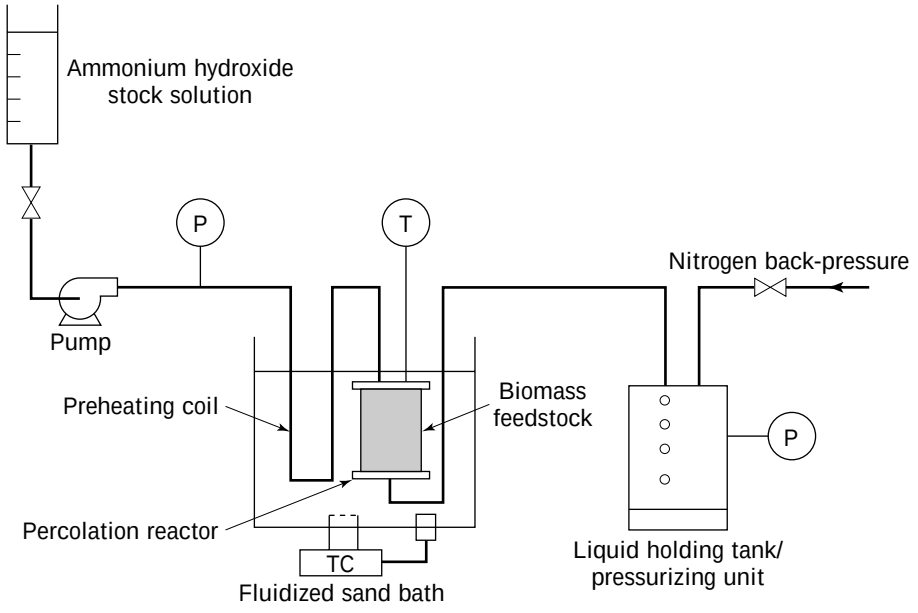


Figure 3. Schematic diagram of the ammonia-recycled percolation process. *Source*: From Ref. 16.

CELLULASE

Unlike chemical hydrolysis, enzymatic hydrolysis of cellulose has the advantage of producing only sugar as product and no sugar degradation byproducts. However, enzymatic hydrolysis of cellulose is a slow process with low yield and a long reaction time. The slow reaction is further complicated by the complex features of biomass materials. Factors affecting the rate of enzymatic hydrolysis of lignocelluloses are the following: the presence of a lignin-hemicellulose shield surrounding cellulose fibers in plant cell walls; the crystallinity of the cellulose; and the relatively small pore sizes in biomass materials that impose mass transfer limitations on the penetration of both microorganisms and the hydrolytic enzymes. Enzyme activity is also subjected to hydrolysis product feedback inhibition. Consequently, some form of pretreatment is needed to expose the cellulose matrix and increase the diffusion of cellulases into the matrix of the substrate.

There are many different microorganisms, including fungi, bacteria, and protozoa, that are capable of producing cellulase enzymes. *Trichoderma reesei* (formerly *T. viride*) is the best-known cellulase-producing microbe; it produces a highly active cellulase complex toward insoluble cellulose. A large number of fungi isolated from rotten cellulosic materials were identified, and their ability to produce cellulase was examined by Reese and Levinson (24).

The extensive research by Reese et al. resulted in the development of the multiple enzymes (C_1 and Cx) concept that could involve in the degradation of crystalline cellulose (25). It also resulted in the use of *Trichoderma reesei* as a source of most investigated active cellulase for the hydrolysis of insoluble cellulose. *Trichoderma* cellulase preparations are available in several different forms from various commercial entities.

Mode of Action of *T. reesei* Cellulase Components

It is widely recognized that true cellulolytic microorganisms produce three basic cellulase components, namely, endocellulase, exocellulase, and cellobiase. The three enzyme components act in concert to hydrolyze crystalline cellulose to its end product, glucose. The components of cellulase enzymes have been studied extensively, and the modes of action of cellulase components have been elucidated (see Refs. 26 and 27 for review). Table 2 summarizes the three major *Trichoderma* cellulase components, the substrates for measurement of enzyme activities, and the mode of action of enzymes.

Crude cellulase activity can be approximated by analyzing the soluble color intensity released from commercially available remazol brilliant blue acid-swollen cellulose (cellulose-azure, Sigma). The standard method (Procedure No. 009) for measurement of overall cellulase activity in terms of filter paper units (FPU) has been provided by National Renewable Energy Laboratory (NREL) (28).

Production of Cellulase

The basic medium for cellulase production from *T. reesei* was developed by Mandels and Reese (29). The composition of the medium is given in the following section. There have been variations in medium composition, especially in the nitrogen source and in the addition of surfactants, such as Tween 80, to reduce the cost of the medium as well as to increase the yields of cellulases. The most common substrates for cellulase production are commercially available materials such as filter paper and Solka-Floc.

Medium. The basic medium has the following composition: $(NH_4)_2SO_4$, 1.4 g/L; KH_2PO_4, 2.0 g/L; urea, 0.3 g/L; $CaCl_2$, 0.3 g/L; $MgSO_4$, 0.3 g/L; trace elements, $FeSO_4 \cdot 7H_2O$, 5.0 mg/L; $MnSO_4 \cdot H_2O$, 1.6 g/L; $ZnSO_4 \cdot 7H_2O$, 1.4 g/L; $CoCl_2$, 2.0 g/L; cellulose, 0.75–1%; protease peptone, 0.075–0.1%; Triton X-100, 0.1–0.2%; initial pH = 5.0–6.0. Rapidly metabolizable carbon sources, such as glycerol or glucose, repress cellulase formation until these carbon sources are depleted by the organism, that is, unless the hypercellulase-producing mutant Rut-C30 is used (30). The Rut-C30 produces cellulase even in the presence of glucose. The addition of small amounts (0.1%) of a readily metabolizable carbon source, such as glycerol, with peptone and cellulose reduces the lag phase for the growth of the organism, consequently increasing cellulase production. The representative cellulase production profile is presented in Figure 4 (31,32).

Supplementation of Cellulases

Commercial cellulase preparations are derived mostly from the *Trichoderma* species (*T. reesei*, and *T. koningii*). *Trichoderma* enzyme preparations are deficient in cellobiase. To enhance the saccharification of cellulose during SSF process, cellobiase derived from *Aspergillus* species is often added. The amounts of cellobiase supplementation are often at 1 unit per FPU cellulase (33).

SIMULTANEOUS SACCHARIFICATION AND FERMENTATION

In general, conversion of cellulosic materials to ethanol involves two stages: *(1)* conversion of cellulose to glucose by either chemical or enzymatic processes, and *(2)* microbial conversion of the resulting glucose to ethanol. In nature, the microbes that are rich in cellulolytic enzymes are usually those with relatively slow growth, such as wood-rotting fungi. The natural selection process has resulted in cellulolytic enzymes that are often highly regulated by the reaction end products. Cellulase activity is subjected to product feedback inhibition (34). When there is a significant concentration of glucose, the enzymatic activity of cellobiase will be affected, which results in the accumulation of cellobiose. The accumulation of cellobiose in turn inhibits its cellulase activities (for details see Ref. 27). To overcome the problems caused by the glucose feedback inhibition, a process known as simultaneous saccharification and fermentation (SSF) was developed.

SSF is a process in which production of ethanol from cellulosic materials is achieved by utilizing cellulose, cellulase, ethanol producing microbes, and nutrients in the same reactor. This process is attractive because the continuous removal of sugars by fermentative organisms allevi-

Table 2. Common Names and Mode of Action of Three Cellulase Components

General category	Other Names	Substrate	Mode of action
β-1,4-Glucan glucanohydrolase (EC.3.2.1.4)	Cx Endoglucanase Endo-β-glucanase CMCase	Acid-swolen cellulose Bacterial cellulose Cellodextrins CMC	Hydrolyze cellulose in random fashion into smaller chains; create reactive site for exocellulase
β-1,4-Glucan cellobiohydrolase (EC.3.2.1.91)	C1 Exocellulase Avicelase Cellobiohydrolase	Filter paper Cellodextrins Bacterial cellulose Acid-swollen cellulose	Cellobiose producer from nonreducing end of the celulose chains
β-1,4-Glucosidase (EC.3.2.1.21)	Cellobiase	Cellobiose	Hydrolyze cellobiose into two glucose molecules

2*Source*: From Refs. 26 and 27.

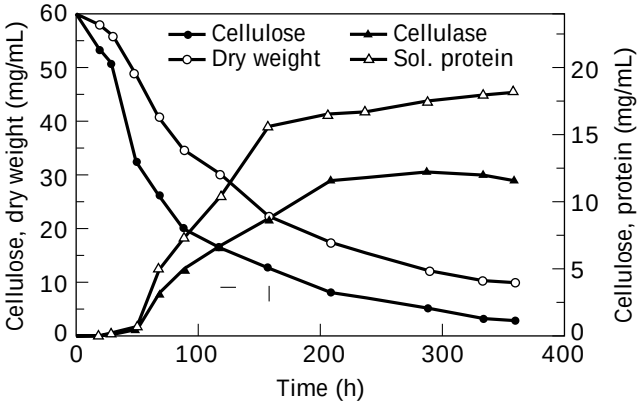


Figure 4. Typical *Trichoderma* cellulase production profile. *Source*: From Ref. 32.

ates the end-product inhibition of saccharification. The process is also simplified because only one reactor is used for SSF. The SSF process for ethanol production from cellulosic materials was reported by Blotkamp et al. (35) and was later tested in the pilot scale (for details see Ref. 36).

The advantages of SSF over the separate saccharification and fermentation of cellulose include the following:

- Reduction of contamination risk
- Lower enzyme loading requirements
- Faster hydrolysis reaction rate
- Higher product yield
- Lower operational expenses

For these reasons, SSF is the preferred process for the bioconversion of the cellulose fraction of lignocellulosic materials to ethanol and other chemicals at the present time. The SSF process involving lignocellulosic material has been reviewed in Ref. 36. There are several factors that are important for an efficient SSF process:

- The physical state of the substrate (particle size and crystallinity of the cellulose)
- Characteristics of the cellulase and cellulase concentration
- Effect of fermentation product on enzyme activity
- Compatibility of saccharification and fermentation temperature
- Characteristics of the yeast strains and yeast concentration
- Maintaining a low level of saccharification products

The nature of the cellulosic substrate is important because the degree of saccharification in the SSF system is often rate limiting with respect to ethanol production. Temperature also determines the economics of ethanol production from cellulose. In most studies, cell densities and enzyme concentration exerted little effect on the amounts of ethanol produced from cellulose under typical SSF conditions (37). In designing an efficient SSF system for the conversion of cellulose to ethanol, the fermentation temperatures should be compatible with operation temperatures, which are generally between 45 and 55 °C. The optimal temperature for the most commonly available cellulase is about 50 °C. Therefore, the use of high-temperature-tolerant microbes is desirable for the application of the SSF process for ethanol production. For the industrial production of ethanol from biomass materials, the use of yeast cultures is preferred over bacteria because the bacterial phage problem can be avoided. In addition, by-product credit can be obtained by marketing spent yeast cells as single-cell protein.

Typical industrial ethanol-producing yeast strains are mesophilic, with an optimal fermentation temperature of 30–37 °C. No true thermophilic yeast species has been reported, and only a few yeast strains that are thermotolerant as well as good ethanol fermentors have been described.

Supraoptimal Temperature Fermentation

In 1986, Anderson et al. (38) studied several thermotolerant yeast strains isolated from sugarcane factories in Australia. Most of the good ethanol-producing yeast strains belong to *Kluyeromyces marxianus* var. maxianus. Some are capable of producing over 6% ethanol (w/v) from glucose at an incubation temperature up to 43 °C after 14 h of fermentation. Over 80% of yeast cells remained viable at the end of fermentation.

Ballesteros et al. (39) studied the high-temperature SSF fermentation of pure cellulose at 42 and 45 °C at cellulase loading of 15 FPU/g substrate. The best results were achieved at 42 °C with a strain of *K. maxianus* and a strain of *K. fragilis*. Both yeast strains produced close to 38 g/L ethanol in 78 h. The results also confirmed the importance of using thermotolerant yeast in SSF processes in order to improve hydrolysis rates and achieve higher ethanol production. There are several possible benefits of using thermotolerant yeast to carry out ethanol fermentation at a supraoptimal temperature:

- Faster rate of saccharification
- Higher rates of sugar production and product (ethanol) formation
- Reduction of contamination by mesophilic yeast and bacteria
- Facilitation of product recovery
- Simplification of cooling during industrial-scale fermentation

However, yeast activities are more sensitive to inhibitory effects of ethanol with increasing temperature.

Improvement of the SSF Process

Typically, the SSF process is carried out in a continuous stirred-tank reactor (CSTR) in batch mode. Under this reaction condition, the fermentation product, ethanol, exerted its effect not only on microbes but also on cellulose hydrolysis activity. To overcome this problem and to improve the efficiency of ethanol production from cellulose, the continuous removal of end product during ethanol production should be beneficial. With this type of process application, the SSF process can be operated in a fed-batch mode. Fed-batch operation is similar to continuous operation except the fermentation broth is retained in the fermentor at all times, whereas the solid substrate is continuously fed into the fermentor.

Different methods of in situ removal of ethanol while it was being formed have been investigated (for review see Ref. 40). The processes include gas stripping, vacuum fermentation, solvent extraction, and the use of membrane bioreactors. Of these alternatives, gas stripping offers advantages in terms of its effectiveness and ease of operation. Ethanol can be recovered from the carrier gas stream by adsorbing onto activated carbon (41), or by condensation from the gas carrier (42). The experiments involving the gas (CO_2) stripping of ethanol during fermentation have been conducted in conventional stirred-tank bioreactors (41,43). Due to poor gas distribution in the stirred tanks, energy consumption for gas input into the fermentor is high. With an airlift fermentor, this drawback can be overcome. The airlift fermentor is the reactor that has a simple structure and provides better gas distribution and uniform air bubble size. The airlift fermentor is also about one-third less energy intensive than stirred tanks (44) because mechanical agitation is replaced by gas upflow to provide the required mixing. Recently, Tsao et al. (45) developed a novel airlift fermentor with a side arm (ALSA). This modified airlift fermentor improves liquid circulation velocities and minimizes the need for antifoaming agents. An ALSA fermentor coupled with gas stripping was used for the removal of ethanol from the fermentation broth during the fermentation of a substrate that contained a high glucose concentration. The ethanol concentration in the broth was maintained at under 4% (w/v) at all times due to the continuous removal of ethanol. The configuration of an ALSA fermentor integrated with a CO_2 stripping system is shown in Figure 5.

The application of this type of integrated fermentation system for high-solid SSF fermentation of biomass materials at an elevated temperature has the potential to overcome ethanol inhibition and temperature incompatibility.

XYLOSE FERMENTATION

Upon the hydrolysis of hemicellulose, xylose becomes the major monosaccharide component in the hemicellulose hydrolysate. Also, in the hydrolysates of lignocellulosic materials, mixtures of glucose and xylose will be expected. Usually, glucose in the hydrolysate can be completely fermented to ethanol within a few hours. However, the complete conversion of xylose to ethanol will take 48–72 h or longer. The nonmatching fermentation rates between glucose and xylose cause difficult design problems. The development of the SSF process has made it possible to convert lignocellulosics to ethanol with cellulase and fermenting microbes in the reactor at the same time. Again, the conversion of cellulose to glucose and finally to ethanol can be completed quickly. Nevertheless, the batch lasts a long time, until the xylose fermentation is completed.

Xylose can be metabolized by bacteria, fungi, or yeast. In bacteria, the initial step of xylose metabolism involves inducible enzymes (xylose transport enzymes, xylose isomerase, and xylulokinase). The direct isomerization of xylose to xylulose is the first step in xylose utilization by bacteria. They produce a variety of end products, including ethanol, from xylose (46). Recently, through advances in genetic engineering, a few bacterial strains have been genetically constructed to efficiently produce ethanol as the major product from xylose (47,48).

In yeast and mycelial fungi, xylose is metabolized via the coupled oxidation–reduction reactions. Xylose reductase is the enzyme involved in the reduction of xylose to xylitol, the sequential enzymatic events through the oxidation of xylitol to xylulose that lead to the utilization of xylose. Many yeast species utilize xylose readily, but ethanol production capability is very limited. Only a few yeast species such as *Pachysolen tannophilus*, *Candida shihatae*, and *Pichia stipitis* are relatively good ethanol producers when xylose is the substrate (49). The production of ethanol from xylose by those three yeast strains has been studied extensively in recent years. Recently, genetically engineered yeast strains have been constructed for more effective conversion of xylose to ethanol (50).

Xylose fermentation by yeast strains is subjected to regulation by dissolved oxygen (51) and by the presence of extraneous materials in the substrate (52). It is also affected by the presence of the fermentation end product,

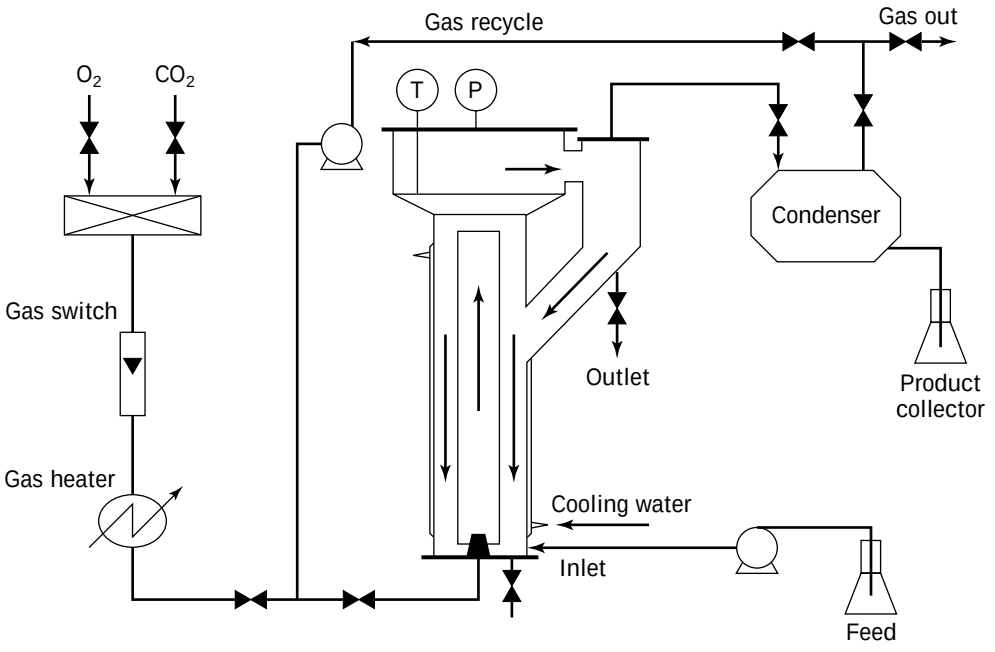


Figure 5. Schematic configuration of ALSA integrated with gas stripping. *Source*: From Ref. 45.

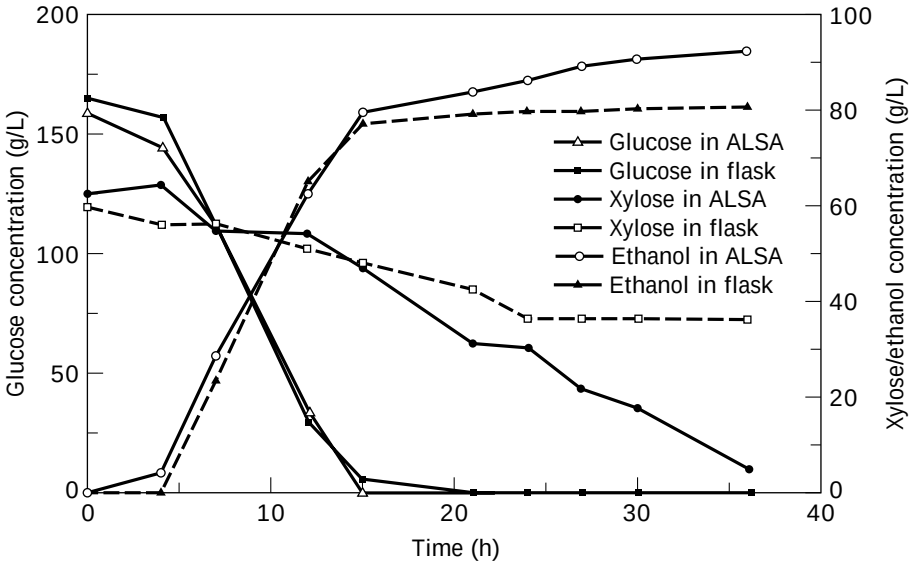


Figure 6. Fermentation of glucose–xylose mixture to ethanol by *Saccharomyces* 1400 (pLHN33) in an ALSA with gas stripping. *Source*: From Ref. 45.

ethanol. Ethanol, at about 45 g/L, will become inhibitory toward the ability of either naturally occurring or genetically modified xylose-fermenting microbes. To overcome the end-product inhibition by ethanol, Tsao et al. (45) used a modified airlift fermentor (ALSA) coupled with a gas-stripping system to remove ethanol continuously from the fermentation broth. Their results (Fig. 6) showed that xylose can be converted into ethanol without the feedback inhibition by ethanol that was encountered under typical stirred-tank fermentation conditions.

Inhibitors

In addition to producing monomer sugar residues, thermochemical processing of biomass is known to produce substances that are inhibitory to microbial growth and ethanol production. For example, furfural derived from pentose sugars, hydroxymethylfurfural, and levulinic acid from hexose sugars are known inhibitors (53). In addition, soluble phenolic compounds derived from lignin degradation and alkaline extractives are also inhibitory toward yeast in low concentration (Table 3). Another potential inhibitor is acetic acid, which can be present in relatively high concentrations (6–15 g/L); it exists as an integral part of hemicellulose. Various procedures have been investigated to minimize the inhibitory effect of the processing by-products. The methods include careful control of the pretreatment process to minimize sugar degradation, using ion exchange resins and activated charcoal to remove

Table 3. The Effects of Potential Inhibitors on Growth and Ethanol Production from Xylose

Compound	Concentration (g/L)	Inhibition (% growth)	Inhibition (% EtOH)
Furfural	2	90	90
HMF[b]	5	92	96
Acetate	10	16	38
HBA[c]	0.75	77	83
SGA[d]	0.75	55	60
Vanillin	0.5	33	53

Source: From Ref. 53.
[a]Percent inhibition of ethanol production from xylose by *Candida shehatae*.
[b]Hydroxymethylfurfural.
[c]Hydroxybenzaldehyde.
[d]Siringaldehyde.

the inhibitors, or the soaking of materials in an alkaline solution prior to hydrolysis.

BIOMASS FRACTIONATION

The rationale behind the fractionation of lignocellulose into lignin, hemicellulose, and cellulose components prior to utilization is as follows:

- Alkaline extractives, soluble lignin, and lignin-derived soluble products are inhibitory to microbial fermentation.
- Hemicellulose is much more susceptible to acid hydrolysis than cellulose.
- Cellulose carbohydrate can be fermented to ethanol much faster than can hemicellulose-derived carbohydrates (e.g., xylose).

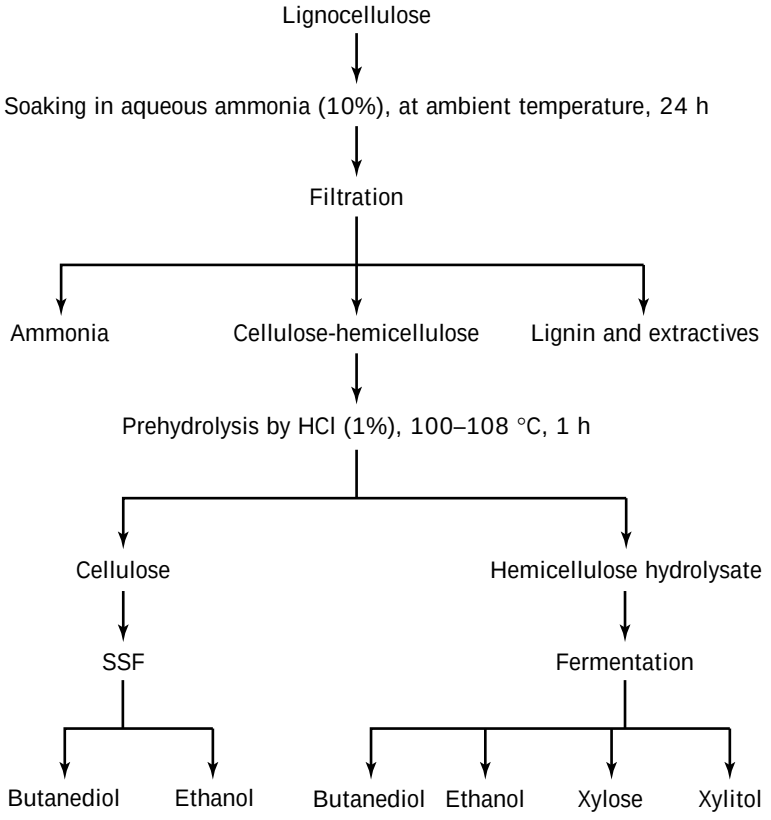


Figure 7. Flow chart of biomass fractionation and potential products. *Source*: From Ref. 55.

Table 4. Composition of Materials after Treatments

Material		Wood Chips			Corn Cobs		
Treatment		Original	After ammonia soaking	After acid hydrolysis	Original	After ammonia soaking	After acid hydrolysis
Composition (%)	Cellulose	42.9	58.09	85.37	44.88	56.2	90.4
	Xylose	22.5	30.37	4.48	32.68	38.5	5.29
	Lignin	25.0	7.11	6.97	7.41	0.85	0.91

Source: From Ref. 55.

The biomass fractionation options and their economical benefits were recently reviewed and discussed by Elander and Hsu (54). The objective of the fractionation is to separate different biomass components and subsequently convert the cellulose and the hemicellulose into separate products. To achieve this objective, mild treatment conditions such as low temperature, low concentration of reagents (or easily recoverable reagents), and low pressures are necessary. Under this biomass component fractionation scheme, cellulose sugar (glucose) and hemicellulose carbohydrates (mostly xylose) can be converted in different reactors to produce different products, if so desired.

Cao et al. (55,56) examined a fractionation option that used corn cobs and aspen woodchips as the substrates. In this biomass fractionation scheme (Fig. 7), the majority of lignin, alkaline extractives, and acetate were solubilized and separated from cellulose and hemicellulose fractions by alkaline treatment. Hemicellulose was thereafter hydrolyzed to its sugar constituents by dilute acids. Hemicellulose carbohydrates were then fermented to ethanol, 2,3-butanediol, or xylitol by xylose-fermenting microbes. The cellulose fraction, after separation from lignin and hemicellulose, was used as the substrate in the SSF process for ethanol production with a thermotolerant yeast strain as the biocatalyst, or it can be bioconverted into other products. Table 4 shows the changes of biomass components after ammonia and dilute acid treatments.

BIBLIOGRAPHY

1. E. Glade, *Industrial Uses of Agricultural Materials*, IUS-5, Economic Research Service, USDA, Washington, D.C., 1995, pp. 21–23.
2. L. Olsson and B. Hahn-Hagerdal, *Enzyme Microb. Technol.* **18**, 312–331 (1996).
3. Y.C.T. Chou, *Biotechnol. Bioeng. Symp.* **17**, 19–32 (1986).
4. R. Torget, P. Walter, M. Himmel, and K. Grohmann, *Appl. Biochem. Biotechnol.* **28/29**, 75–97 (1991).
5. F. Parisi, *Adv. Biochem. Eng. Biotechnol.* **38**, 53–87 (1989).
6. B.E. Dale and M.J. Moreira, *Biotechnol. Bioeng. Symp.* **12**, 31–43 (1983).
7. E.T. Papoutsakis, *Biomass Handbook*, Gordon and Breach Science Publishers, New York, 1989, pp. 271–286.
8. S.K. Garg and A. Jain, *Biosource Technol.* **51**, 103–109 (1995).
9. J.H. Litchfield, in M. Moo-Young ed., *Biomass Conversion Technology: Principles and Practice*, Pergamon, 1987, pp. 113–122.
10. E. Sjöström, *Wood Chemistry: Fundamentals and Applications*, Academic, New York, 1981, pp. 1–20.
11. D.S. Chahal, in D.S. Chahal ed., *Food, Feed and Fuel from Biomass*, Oxford & IBH, New Delhi, 1991, pp. 59–83.
12. J.D. McMillan, in M.E. Himmel, J.O. Baker, R.P. Overend eds., *Enzymic Conversion of Biomass for Fuels Production*, American Chemical Society, Washington, D.C., 1994, pp. 292–324.
13. J.H. Litchfield, *Microbial Technol.* **1**, 93–155 (1979).
14. N.J. Cao, Q. Xu, C.S. Chen, C.S. Gong, and L.F. Chen, *Appl. Biochem. Biotechnol.* **45/46**, 521–530 (1994).
15. M.M. Chang, Y.C.T. Chou, and G.T. Tsao, *Adv. Biochem. Eng.* **20**, 15–42 (1981).
16. H.H. Yoon, Z.W. Wu, and Y.Y. Lee, *Appl. Biochem. Biotechnol.* **51/52**, 5–19 (1995).
17. J.M. Gould, *Biotechnol. Bioeng.* **25**, 46–52 (1984).
18. M.T. Holtzapple, J.-H. Jun, G. Ashok, S.L. Patibanla, and B.E. Dale, *Appl. Biochem. Biotechnol.* **28/29**, 59–74 (1991).
19. Weil, P. Westgate, K. Kohlmann, and M.R. Ladisch, *Enzyme Microb. Technol.* **16**, 1002–1004 (1994).
20. W. Schwald, C. Breuil, H.H. Brownell, M. Chan, and J.N. Saddler, *Appl. Biochem. Biotechnol.* **20/21**, 29–44 (1989).
21. D.J. Schell, R. Torget, A. Power, P.J. Walter, K. Grohmann, and N.D. Hinman, *Appl. Biochem. Biotechnol.* **28/29**, 87–97 (1991).
22. J.E. Carrasco, M.C. Saiz, A. Navarro, P. Soriano, F. Saez, and J.M. Martinez, *Appl. Biochem. Biotechnol.* **45/46**, 23–34 (1994).
23. K. Grohmann, R. Torget, and M. Himmel, *Biotechnol. Bioeng. Symp.* **15**, 59–80 (1985).
24. E.T. Reese, and H.S. Levinson, *Physiol. Plant* **5**, 345–366 (1952).
25. E.T. Reese, S.G.H. Siu, and H.S. Levinson, *J. Bacteriol.* **59**, 485–497 (1950).
26. C.S. Gong, M.R. Ladisch, and G.T. Tsao, *Adv. Chem. Ser.* **181**, 261–287 (1979).
27. D.S. Chahal, P.S. Chahal, and G. Andre, in D.S. Chahal ed., *Food, Feed and Fuel from Biomass*, Oxford & IBH, New Delhi, 1991, pp. 281–312.
28. R. Torget, *Chemical Analysis and Testing Standard Procedures*, Procedure No. 009. NREL, Golden, Col., 1994.
29. M. Mandel and J. Weber, *Adv. Chem. Ser.* **95**, 391–414 (1969).
30. B.S. Montenecourt and D.E. Eveleigh, *Adv. Chem. Ser.* **181**, 289–301 (1979).
31. M. Mandel and D. Sternberg, *J. Ferment. Technol.* **54**, 267–286 (1976).
32. K. Nishizawa, *Biomass Handbook*, Gordon and Breach Science Publishers, New York, 1989, pp. 307–318.
33. D. Ryu and M. Mandel, *Enzyme Microb. Technol.* **2**, 91–102 (1980).
34. C.S. Gong, M.R. Ladisch, and G.T. Tsao, *Biotechnol. Bioeng.* **19**, 959–981 (1977).
35. P.J. Blotkamp, M. Tagagi, M.S. Pemberton, and G.H. Emert, *AIChE Symp. Ser.* **181**, 85–90 (1981).
36. G.H. Emert, in D.S. Chahal ed., *Food, Feed and Fuel From Biomass*, Oxford & IBH, New Delhi, 1991, pp. 313–333.
37. G.P. Philippidis, and T.K. Smith, *Appl. Biochem. Biotechnol.* **51/52**, 117–124 (1995).
38. P.J. Anderson, K. Mcneil, and K. Watson, *Appl. Environ. Microbiol.* **51**, 1314–1320 (1986).
39. I. Ballesteros, J.M. Oliva, M. Ballesteros, and J. Carrasco, *Appl. Biochem. Biotechnol.* **39/40**, 201–211 (1993).
40. P. Cen and G.T. Tsao, *Separation Technol.* **3**, 58–75 (1993).
41. P.K. Walsh, C.P. Liu, M.E. Findley, A.I. Liapis, and D.J. Siehr. *Biotechnol. Bioeng. Symp.* **13**, 629–647 (1983).
42. M.C. Dale, M.R. Okos, and P.C. Wanket, *Biotechnol. Bioeng.* **27**, 943–952 (1985).
43. H.S. Liu and H.W. Hsu. *Chem. Eng. Sci.* **45**, 1289–1299 (1990).
44. M. Trager, G.N. Qazi, U. Onfen, and C.L. Chopra. *J. Ferm. Bioeng.* **68**, 112–116 (1989).
45. G.T. Tsao, N.J. Cao, M.S. Krishnan, J.X. Du, and C.S. Gong, *Proceedings, The Seventh National Bioenergy Conference*, 1996.
46. C.S. Gong, L.F. Chen, M.C. Flickinger, and G.T. Tsao, *Adv. Biochem. Eng.* **20**, 93–118 (1981).

47. F. Alterthum, and L.O. Ingram, *Appl. Environ. Microbiol.* **55**, 1943–1948 (1989).
48. M. Zhang, C. Eddy, K. Deanda, M. Finkelstein, and S. Picataggio, *Science* 267, 240–243 (1995).
49. J.N. Nigam, R.S. Ireland, A. Margaritis, and M.A. Lachance, *Appl. Environ. Microbiol.* **50**, 1486–1489 (1985).
50. PCT Pat. WO95/13362 (Aug. 4, 1995), N.W.Y. Ho and G.T. Tsao (to Purdue University Research Foundation).
51. J.M. Laplace, J.P. Delgenes, R. Moletta, and J.M. Navarro, *Appl. Microbiol. Biotechnol.* **36**, 158–162 (1991).
52. A.V. Tran and R.P. Chambers, *Biotechnol. Lett.* **7**, 841–846 (1985).
53. J.P. Delgenes, R. Moletta, and J.M. Navarro, *Enzyme Microb. Technol.* **19**, 220–225 (1996).
54. R.T. Elander and T. Hsu, *Appl. Biochem. Biotechnol.* **45/46**, 463–478 (1994).
55. N.J. Cao, Y. Huang, C.S. Gong, and G.T. Tsao, *AICHE Series: 5th World Congress of Chemical Engineering*, 1996, pp. 16–21.
56. N.J. Cao, M.S. Krishnan, J.X. Du, C.S. Gong, N.W.Y. Ho, Z.D. Chen, and G.T. Tsao, *Biotechnol. Lett.* **18**, 1013–1018 (1996).

See also ENZYMES, PULP AND PAPER PROCESSING; HEMICELLULASES.

ORGANOSILICON COMPOUNDS

ATSUO TANAKA
TAKUO KAWAMOTO
Kyoto University
Kyoto, Japan

KEY WORDS

Bioconversion
Deacylation
Dehydrogenation
Enantioselectivity
Esterification
Hydrolysis
Oxidoreduction
Regioselectivity
Stereoselectivity
Transesterification

OUTLINE

INTRODUCTION

Silicon, which is the second most abundant element in the Earth's crust, belongs to the same group as carbon, which is one of the most fundamental elements for organisms, but is different from carbon in its characteristics (1). Silicon is less electronegative and has a bigger covalent radius than carbon (Table 1). A bond between silicon and oxygen is stronger than the corresponding bond between carbon and oxygen, whereas its bonds with carbon and hydrogen are weak.

Silicon plays an important role in the biosphere and is present mainly in living organisms, especially at a lower stage of evolutionary development. Silicate bacteria, the simplest algae, and spore plants, in fact, contain very large amounts of silicon (2,3). Recent research has indicated that silicon is also important for higher plants, animals, and humans (4), although its actual biochemical function is at present not fully defined. It seems, however, that natural biochemical processes mainly involve silicon bound to oxygen, whereas organosilicon compounds containing Si–C bonds have not been detected (1–3). Therefore, the effects of organosilicon compounds on organisms and/or enzymes are very interesting.

The explosive growth of organosilicon chemistry has created a growing awareness of its considerable utility to organic chemists and has shown that a wide spectrum of organosilicon compounds often exhibit interesting biological properties (2,3). Recently, pharmaceutical applications of organosilicon chemistry have also been gaining attention (2,3,5–7). Therefore, the production of organosilicon compounds with desired chemical and physical properties by the use of effective biocatalysts is of great importance. Several examples of bioconversion of organosilicon compounds have been summarized by Ryabov (8) and Tanaka and Kawamoto (9).

MICROBIAL REDUCTION OF ORGANOSILICON COMPOUNDS

Stereoselective microbial reduction of organosilicon compounds was first demonstrated by Tacke et al. (10). Analogous to the microbial reduction of alkyl acetoacetates to optically active 3-(*S*)-hydroxybutyrates, the sila analogues **1a** and **1b** could be transformed stereoselectively into the corresponding reduction products (*S*)-**2a** and (*S*)-**2b** by using growing cells of the yeast *Kloeckera corticis* (Fig. 1). This stereoselective reduction could be performed on a preparative scale with good yields. The optical purities obtained were about 80% ee for (*S*)-**2a** and (*S*)-**2b**.

In this case, the carbonyl groups attacked by the microorganism are relatively far away from the silicon atom. However, it was also shown that compounds containing a carbonyl group bound directly to the silicon atom can also be reduced stereoselectively by microorganisms. For example, achiral acetyldimethylphenylsilane (**3**) was transformed stereoselectively into the optically active reduction product (*R*)-**4** (86% ee) by growing and resting cells of the yeast *Trigonopsis variabilis* DSM 70714 (Fig. 2) in a yield of 70% (11). Furthermore, a wide range of microorganisms

Table 1. Physical Properties of the Silicon Atom and the Carbon Atom

Atom	Electronegativity	Bond	Bond length (nm)	Bond dissociation energy (kJ mol^{-1})
C	2.55	C—C	0.153	334
Si	1.90	C—Si	0.189	318

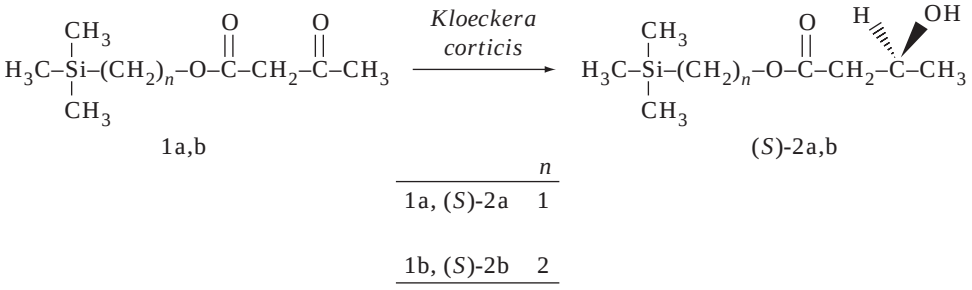


Figure 1. Stereoselective microbial reduction of sila-analogues of alkyl acetoacetates.

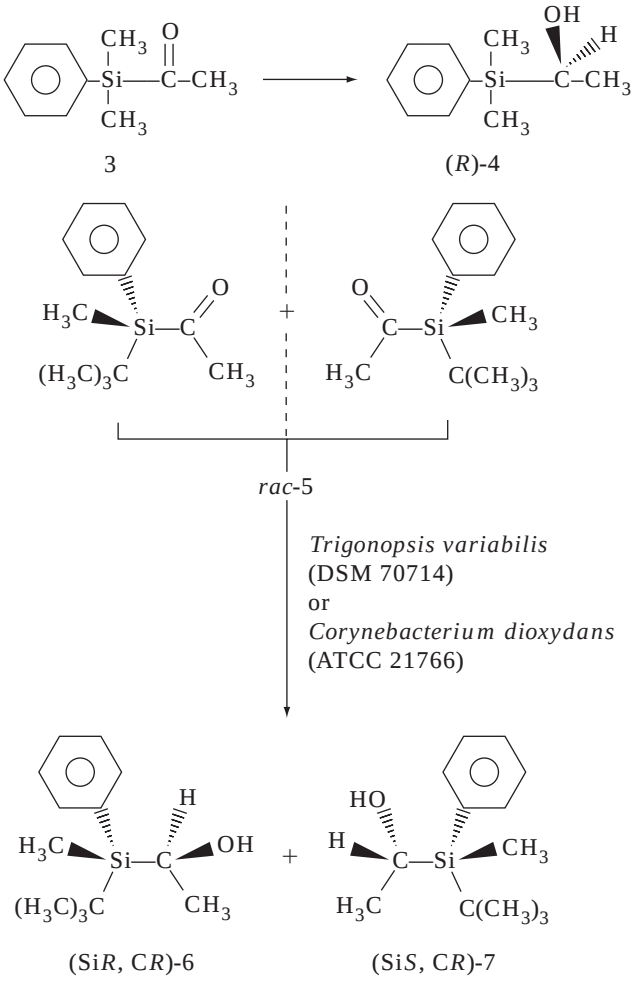
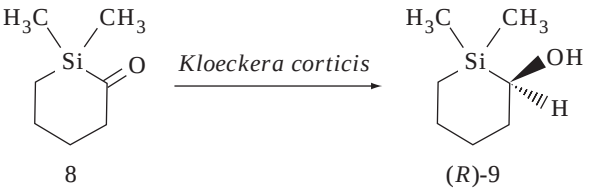


Figure 2. Microbial reduction of organosilicon compounds having a carbonyl group bound directly to the silicon atom.

such as bacteria, yeasts, fungi, and green algae showed the ability to accept **3** as a substrate for stereoselective reduction (12). Among 30 strains of microorganisms tested as resting cells, *T. variabilis* DSM 70714 was found to exhibit the highest specific activity (1.5 mg product g wet $mass^{-1}$ min^{-1}), whereas the highest stereoselectivity (95% ee) was observed with the bacteria *Acinetobacter calcoaceticus* ATCC 31012 and *Corynebacterium dioxydans* ATCC 21766. The same bioconversion could be performed by resting cells of *Saccharomyces cerevisiae* DHW S-3 at a rate of 44–45 μmol L^{-1} min^{-1} with a yield of 40% and optical purity of more than 99% after purification (13). Plant cell suspension cultures of *Synphytum officinale* L. and *Ruta graveolens* L. served as the biocatalysts for the conversion of **3** to (*R*)-**4**, although the yields and optical purities were not satisfactory (*S. officinale*: 81% ee, yield 15%; *R. graveolens*: 60% ee, yield 9%) (14).

It is noteworthy that a carbonyl group bound to the silicon atom can be transformed by various microorganisms, as in the case of conventional ketones, although there are big differences in physical and chemical properties between these compounds. Another example of microbial reduction of organosilicon compounds is the transformation of racemic acetyl(*tert*-butyl)methylphenylsilane (*rac*-**5**) with resting cells of *T. variabilis* DSM 70714 (15) and *C. dioxydans* ATCC 21766 (16). In these cases, enantioselective reduction of both enantiomers of the racemic compounds was observed, resulting in the formation of the diastereomeric optically active silanes (Si*R*,C*R*)-**6** (*T. variabilis*: 97% ee, yield 74%; *C. dioxydans*: 99% ee, yield 20%) and (Si*S*,C*R*)-**7** (*T. variabilis*: 96% ee, yield 70%; *C. dioxydans*: 99% ee, yield 20%). Growing cells of *K. corticis* catalyzed the stereoselective reduction of 1,1-dimethyl-1-silacyclohexan-2-one (**8**) to (*R*)-1,1-dimethyl-1-silacyclohexan-2-ol ((*R*)-**9**) (92% ee, yield 60%) (Fig. 3) (17). Reduction of (+)-(benzyloxymethyl)(*tert*-butyl)methylsilyl methyl ketone ((+)-**10**) to (Si*R*,1*R*)-(+)-1-[(benzyloxymethyl)(*tert*-butyl)methylsilyl]ethanol ((+)-**11**) and its



8 (*R*)-9

Figure 3. Microbial reduction of 1-silacyclohexan-2-one.

(Si*S*,1*R*)-enantiomer ((+)-**12**) was performed by resting free and immobilized cells of *T. variabilis* (>96% ee, yield 67–99%) (Fig. 4) to prepare (*R*)-(−)-**10** and (*S*)-(+)-**10**, the former being the intermediate in the synthesis of (*R*)-(+)-1-phenylethanol (18).

These studies clearly demonstrate that organosilicon compounds could be transformed by microorganisms and that the microbial transformation is useful for the preparation of optically active organosilicon compounds.

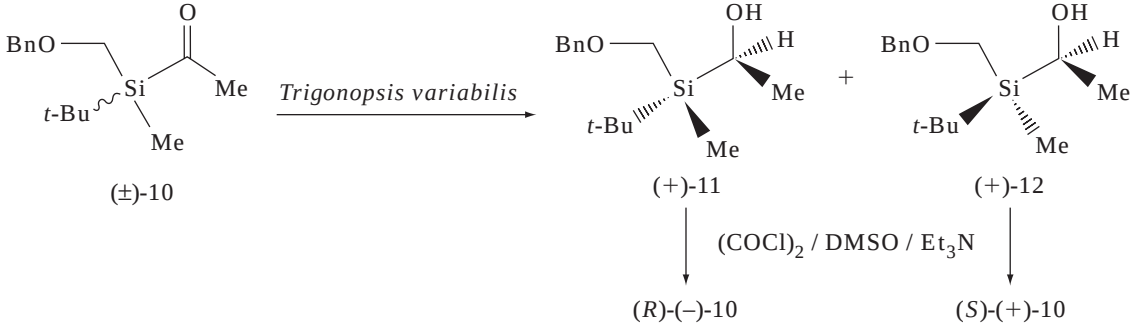
OXIDOREDUCTION OF ORGANOSILICON COMPOUNDS BY ALCOHOL DEHYDROGENASES

Bioconversion of organosilicon compounds by alcohol dehydrogenases (ADHs) is very attractive because of the importance of asymmetric alcohol/carbonyl oxidoreduction in organic chemistry. Although several papers are available on the reduction of organosilicon compounds by microorganisms as mentioned earlier, it is not clear whether ADH can actually catalyze the oxidoreduction of organosilicon compounds in these organisms, and it has not been possible to discuss the effect of the silicon atom on enzymatic reactions with ADH.

It was shown for the first time by Zong et al. (19) that ADH from horse liver (HLADH) could catalyze the dehydrogenation of organosilicon compounds and that the silicon atom in the substrates had unique effects on enzymatic reactions due to specific characteristics of the silicon atom. The bioconversion of three organosilicon compounds with different chain lengths between the silicon atom and the hydroxyl group ($Me_3Si(CH_2)nOH$, n = 1–3) by HLADH was systematically studied in comparison with the corresponding carbon compounds ($Me_3C(CH_2)nOH$, n = 1–2). HLADH, but not yeast alcohol dehydrogenase, catalyzed the dehydrogenation of 2-trimethylsilylethanol (n = 2) **(14a)** and 3-trimethylsilylpropanol (n = 3) **(15a)**. **14a** was a better substrate than both its carbon analogue, 3,3-dimethylbutanol **(14b)** and ethanol. The improved activity of HLADH on **14a** could be accounted for by a lower activation energy of the reaction by HLADH than that with the carbon analogue. In contrast, HLADH showed no activity on trimethylsilylmethanol (n = 1) **(13a)**, whereas it catalyzed the dehydrogenation of the carbon analogue, 2,2-dimethylpropanol **(13b)**. This phenomenon was explained by the β-effect of the silicon atom, that is, stabilization of a negative charge at the α-position and a positive charge at the β-position of the carbon atom from the silicon atom (1). As the silicon atom stabilizes the positive charge of the β-carbon atom, the H^- anion could be removed more easily from the β-carbon atom of **14a**, which is the first step of the dehydrogenation of alcohol catalyzed by ADH. As a result, the activation energy of the dehydrogenation of **14a** by HLADH probably became lower than that of **14b**, and thus **14a** is a better substrate than the carbon counterpart (Fig. 5). Conversely, it is more difficult for the H^- anion to be removed from the α-carbon atom of **13a** because the electron density around the α-carbon of **13a** is higher than that of **13b** owing to a lower electronegativity of the silicon atom (Fig. 5). Sonomoto et al. (20) also confirmed the same phenomena.

It was also demonstrated that HLADH could catalyze the enantioselective dehydrogenation of racemic trimethylsilylpropanols, 1-trimethylsilyl-2-propanol **(16a)** and 2-trimethylsilyl-1-propanol **(18a)** (Fig. 6) in a water–organic solvent two-layer system with coenzyme regeneration by a glutamate dehydrogenase-2-oxoglutarate system (21). Especially, the enantiomeric excess of **16a** remaining at 51% conversion was as high as 99%. In this case, 1-trimethylsilyl-1-propanol **(17a)** did not serve as a substrate. A comparative study with their carbon counterparts (**16b** and **18b**) showed results similar to those of the hydrolases to be described later, that is, substitution of the silicon atom for the carbon atom improved the enantioselectivity of HLADH.

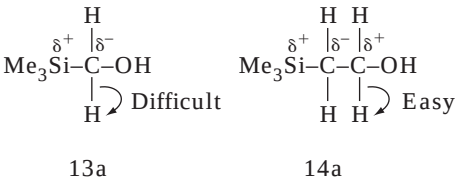
Tsuji et al. (22) found that enantioselective dehydrogenation of primary alcohol β-hydroxyalkylsilane, 2-trimethylsilyl-1-propanol **(18a)**, could be carried out with only a catalytic amount of NAD^+ and HLADH in an aqueous system containing 5% tetrahydrofuran. When **18a** (5



(±)-10 (+)-11 (+)-12

(*R*)-(–)-10 (*S*)-(+)-10

Figure 4. Microbial reduction of a silicon-containing ketone.



13a 14a

Figure 5. Electric effect of the silicon atom on removal of H^- from trimethylsilylalkanols.

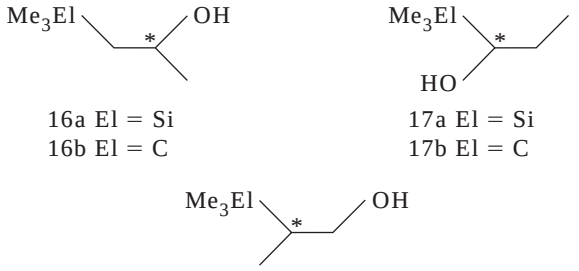


Figure 6. Structures of three isomers of trimethylsilylpropanols and their carbon analogues.

$\times$ 10 μmol) was applied to the HLADH-catalyzed dehydrogenation with a small amount of NAD^+ (5×10^{-1} μmol) in the aqueous system, NAD^+ could be regenerated in situ through reduction of 1-propanal **(20)** formed by spontaneous degradation of the dehydrogenation product, 2-trimethylsilyl-1-propanal **(19)**, in the presence of water (Fig. 7). In this case, conversion ratio reached 75% after 14 h of reaction; the turnover number of NAD^+ was calculated to be about 75, where the optical purity of remaining **18a** was 90% ee. If the amount of NAD^+ was reduced to 5×10^{-2} or 5×10^{-4} μmol (**18a**, 5×10 μmol), the turnover number of NAD^+ reached 640 or 7,200, respectively. This novel reaction system is simple and advantageous because it does not require any other enzyme and substrate for regeneration of NAD^+, and there is no product inhibition due to the spontaneous degradation of the dehydrogenation product. On the other hand, dehydrogenation of the corresponding carbon analogue **(18b)** did not proceed under this condition because the aldehyde formed by dehydrogenation of **18b** with HLADH did not undergo the degradation to provide a substrate for NAD^+ regeneration. These results indicate that application of organosilicon compounds makes it possible to construct an efficient and unique conversion system, which is impossible in the case of conventional carbon compounds. Other primary β-hydroxysilanes having different substituents on the silicon atom or on the stereogenic carbon atom, 2-trimethylsilyl-1-butanol, 2-trimethylsilyl-1-hexanol, and 2-dimethylphenylsilyl-1-butanol, were also found to serve as substrates in enantioselective dehydrogenation by HLADH with this novel NAD^+ regeneration system. In contrast, dehydrogenation of the secondary β-hydroxysilane, 1-trimethylsilyl-2-propanol **(16a)**, was negligible under the same conditions, although HLADH could convert **16a** with conventional coenzyme regeneration (21). The dehydrogenated product of **16a**, 1-trimethylsilyl-2-propanone, was supposed to be degraded by the addition of water into trimethylsilanol and acetone also in this case, but regeneration of NAD^+ did not occur because acetone was not recognized as the substrate by HLADH and so the dehydrogenation reaction did not proceed over the catalytic amount of NAD^+.

Recently, Wong et al. (23) have reported that NADPH-requiring *Lactobacillus kefir* ADH could catalyze the enantioselective carbonyl reduction of ketones having a large variety of side chains, including trimethylsilyl-protected terminal alkyne ketones, and NADPH regeneration in the presence of 2-propanol.

BIOCONVERSION OF ORGANOSILICON COMPOUNDS BY HYDROLASES

Hydrolases such as lipases, esterases, and proteases are the easiest enzymes to handle because hydrolases do not require coenzymes and are readily available in a large quantities. Therefore, they have been extensively used as catalysts in regio- or enantioselective conversion.

Therisod and Klibanov (24) mentioned additionally, in the regioselective acylation of secondary hydroxyl groups of C-6 protected glucose, that modification of the C-6 hydroxyl group of D-glucose with a more bulky *tert*-butyldiphenylsilyl group increased the regioselectivity of *Candida cylindracea* lipase toward the C-2 position, compared with 6-*O*-butylylglucose. It was, however, not clear how reactive the compound modified with the *tert*-butyldiphenylsilyl group was in the lipase-catalyzed acylation. Luyten et al. (25) reported in the study on enantioselective hydrolysis of several functionalized dimethyl malonates **(21a–f)** by pig liver esterase (PLE) (Fig. 8) that a *tert*-butyldimethylsilyl derivative of dimethyl 2-(hydroxymethyl)-2-methylmalo-

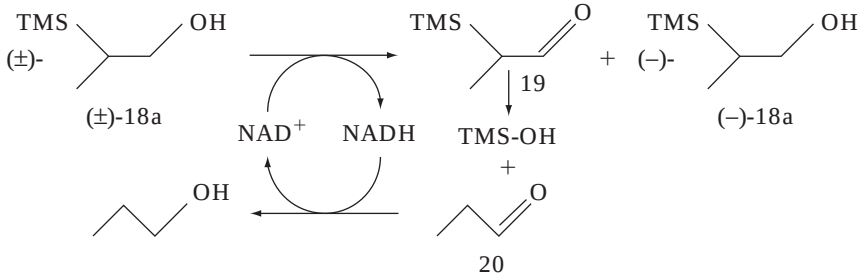
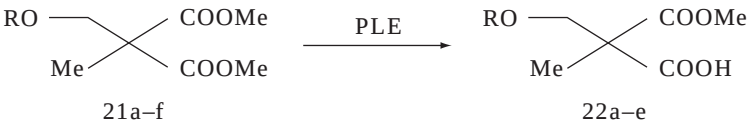


Figure 7. Horse liver alcohol dehydrogenase–catalyzed enantioselective dehydrogenation of 2-trimethylsilyl-1-propanol with in situ coenzyme regeneration. TMS, trimethylsilyl.



21a–f 22a–e

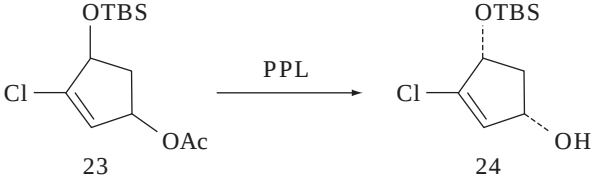
a, R = H
b, R = Me
c, R = (*t*-Bu)(Me)$_2$Si
d, R = $C_6H_5CH_2$
e, R = *t*-Bu
f, R = $CH_3C{=}O$

Figure 8. Enantioselective hydrolysis of functionalized dimethyl malonates by pig liver esterase.

nate **(21c)** gave a high enantioselectivity (95 % ee) but a sluggish reaction. As a result, the yield of the optically pure product ((*R*)-**22c**) was very low (49%). In this study, a (*tert*-butoxy)methyl derivative **(21e)** showed the best results, namely, a high enantioselectivity (96% ee) and a high yield (90%). Mori and Takeuchi (26) prepared enantiomerically pure (1*S*, 4*R*)-4-*tert*-butyldimethylsilyloxy-3-chloro-2-cyclopenten-1-ol **(24)** (ca. 100% ee), which is a key chiral building block for synthesis of punaglandin 4, a chlorinated marine prostanoid, through the asymmetric hydrolysis of the mixture of 3-acetoxy-5-*tert*-butyldimethylsilyloxy-1-chlorocyclopentenes **(23)** with pig pancreatic lipase (PPL) (Fig. 9) in 25% yield.

In the case of the organosilicon compounds already mentioned, functional groups attacked by hydrolases were far away from the silicon atom. Therefore, the effect of the silicon atom itself on enzymatic reactions seems very weak. It is interesting to study the bioconversion of organosilicon compounds whose silicon atom more strongly affects the functional groups to be attacked by biocatalysts. Interesting attempts were made by Therisod (27). He tried to use trimethylsilyl ethers instead of alcohols as novel nucleophiles in hydrolase-catalyzed transesterification (Fig. 10), but failed.

Kawamoto et al. (28) attempted a comparative study by using organosilicon compounds having a different chain length between the hydroxyl group and the silicon atom ($Me_3Si(CH_2)nOH$, n = 0, 1, 2, and 3) and the corresponding carbon compounds ($Me_3C(CH_2)nOH$, n = 0, 1, and 2) as an acyl acceptor in the enantioselective esterification of 2-(4-chlorophenoxy)propanoic acid **(25)**, whose *R*-enantiomer is useful as an herbicide, with lipase OF 360 of *C. cylindracea* (Fig. 11) and discussed the effects of the silicon atom on the enzymatic reaction in connection with the distance between the hydroxyl group and the silicon atom. Trimethylsilylmethanol (n = 1) **(13a)** was found to be a particularly superior substrate, that is, the reaction rate was much higher than that with the corresponding carbon compound **(13b)**, and the enantiomeric excess of the acid remaining was above 95% at about 50% conversion. In the case of conventional substrates such as the carbon analogue and linear-chain alcohols, a high reaction rate was not consistent with a high enantioselectivity. These results indicate that organosilicon compounds may solve the unavoidable and probably inherent problems of enantioselective reactions with conventional substrates. On the other hand, no difference was observed between 2-trimethylsilylethanol (n = 2) **(14a)** and its carbon analogue, 3,3-dimethylbutanol **(14b)**, with respect to enzymatic activity and enantioselectivity. Thus, the silicon atom mimicked the carbon atom for lipase in the case of **14a** but was different from the carbon atom in enhancing the reactivity of **13a**. The phenomena were explained on the basis of the properties of the silicon atom, such as its low electronegativity and big atomic radius compared with those of the carbon atom. A lower electronegativity of the silicon atom results in a higher nucleophilicity of the oxygen atom of **13a** compared with that of the corresponding carbon analogue **(13b)**, and the hydroxyl group of **13a** is less sterically hindered due to the longer bond between Si and C than between C and C in **13b** (Fig. 12). **13a** is, therefore, more easily accessible to the acyl-enzyme intermediate and reacts as an acyl acceptor much faster than the carbon analogue. In the case of **14a**, the favorable effect of the silicon atom mentioned earlier was negligible owing to the presence of a long ethylene group between the silicon atom and the hydroxyl group. **14a**, and probably 3-trimethylsilylpropanol (n = 3) **(15a)**, were consequently regarded as similar substrates to the corresponding carbon compounds by lipase. In spite of the favorable characteristics of the silicon atom, trimethylsilanol (n = 0) did not serve as a substrate similarly to the corresponding tertiary alcohol.



23 24

Figure 9. Enantioselective hydrolysis of a silicon-containing ester. TBS, *t*-butyldimethylsilyl.

$$CH_3(CH_2)_2COOEt + ROSiMe_3 \longrightarrow CH_3(CH_2)_2COOR + EtOSiMe_3$$

Figure 10. Attempt to use trimethylsilyl ethers as novel nucleophiles in hydrolase-catalyzed transesterification. R, hexyl or cyclohexyl.

To our knowledge, this work is the first study in which the effects of the silicon atom in substrates on enzymatic reaction were discussed systematically. Furthermore, this study showed for the first time the ability of organosilicon compounds to break the limit of conventional substrates, owing to their unique characteristics. According to this explanation, trimethylsilylmethylamine ($Me_3SiCH_2NH_2$) can also be expected to be a better acyl acceptor than the cor-

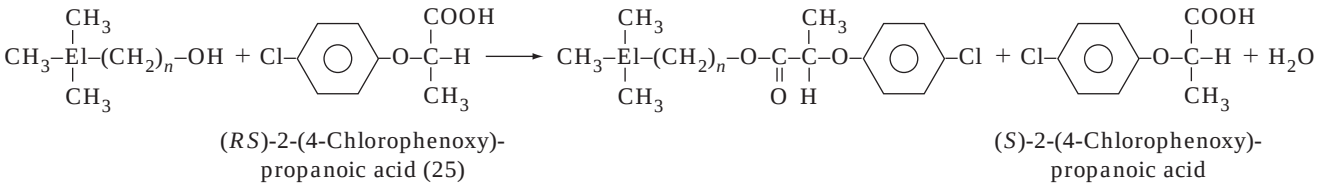


Figure 11. Lipase-catalyzed enantioselective esterification of 2-(4-chlorophenoxy)propanoic acid with silicon-containing alcohols and the corresponding carbon analogues. El, Si or C.

responding carbon compound, 2,2-dimethylpropylamine ($Me_3CCH_2NH_2$), in hydrolase-catalyzed amide bond formation. In fact, trimethylsilylmethylamine was a better substrate for hydrolases such as lipase OF 360 (*C. cylindracea*), lipase KLIP-100 (*Pseudomonas* sp.), lipoprotein lipase Type A (*Pseudomonas* sp.), and cholesterol esterase Type A (*Pseudomonas* sp.) than the corresponding carbon compound in amide bond formation with octanoic acid (R.S. So, Y. Masuda, T. Kawamoto, and A. Tanaka, unpublished results).

In contrast to our results, Tsai and Wei (29) reported that both **13a** and **14a** were effectively applied to the enantioselective esterification of racemic naproxen (2-(6-methoxy-2-naphthyl)propionic acid), a nonsteroidal anti-inflammatory drug, by *C. cylindracea* lipase. However, these alcohols were not good substrates for *Mucor javanicus* lipase. They have developed this esterification reaction system with **13a** (30,31).

Optical resolution of organosilicon compounds having a functional group attached to the carbon atom at the α-position from the silicon atom was also studied, and the enantioselective hydrolysis of racemic 2-acetoxy-1,1-dimethyl-1-silacyclohexane (*rac*-**26**) in an aqueous solution and enantioselective esterification of racemic 1,1-dimethyl-1-silacyclohexan-2-ol (*rac*-**27**) with triacetin in isooctane were successfully carried out by the use of a crude lipase preparation of *C. cylindracea*, leading to (*S*)-**27** (yield 71%, 95% ee) and (*S*)-**26** (yield 92%, 95% ee), respectively (Fig. 13) (32).

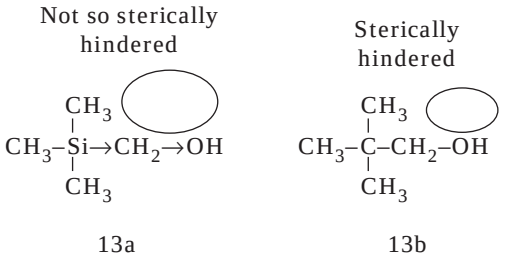


Figure 12. Effect of the silicon atom in trimethylsilylmethanol on esterification.

In general, it is difficult to perform the highly enantioselective conversion of primary alcohols. However, the efficient optical resolution of a primary alcohol, 2-trimethylsilyl-1-propanol **(18a)**, was achieved with 95% ee at 50% conversion by the enantioselective esterification with lipase Saiken 100 from *Rhizopus japonicus* in water-saturated isooctane (33). In the study on the enantioselective esterification of three isomers of trimethylsilylpropanols, 1-trimethylsilyl-2-propanol **(16a)**, 1-trimethylsilyl-1-propanol **(17a)**, and **18a**, and the corresponding carbon analogues (**16b, 17b**, and **18b**) (Fig. 6) with 5-phenylpentanoic acid by hydrolases in water-saturated isooctane, it was indicated that the silicon atom in the substrates generally enhanced the enzyme enantioselectivity. It is also worth noting that the β-hydroxyalkyl-

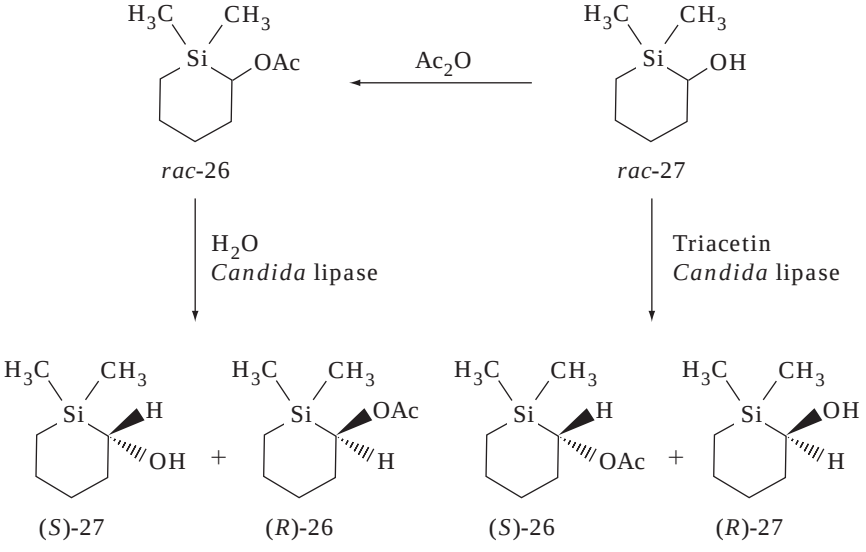
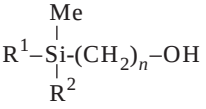


Figure 13. Lipase-catalyzed enantioselective hydrolysis and transesterification of organosilicon compounds whose functional group is attached to the carbon atom at the α-position from the silicon atom.

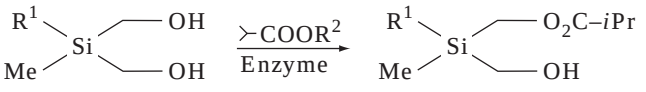
silanes, **16a** and **18a**, were effectively esterified by hydrolases, but they were easily converted to alkenes via *β*-elimination under both acidic and basic conditions (Peterson olefination) (1), and consequently, it is not possible to esterify them by chemical catalysts such as acids. Several enzymes could convert such unstable compounds with enantioselectivity. This fact is a good example showing the effectiveness of introduction of biochemical methods into organosilicon chemistry.

Optical resolution of organosilicon compounds, which have a stereogenic silicon atom, is also an important target for the application of enzymes. Furthermore, it is interesting to find out whether enzymes can recognize the chirality on the silicon atom, and such chiral silanes are expected to be new materials, synthetic reagents (34), or biologically active compounds (2). When PhSi(Me)(Et)CH_2OH **(28)** was esterified with 5-phenylpentanoic acid, various types of hydrolases exhibited a good esterification activity but low enantioselectivity. Of the enzymes examined, a commercial preparation of crude papain gave the best results, that is, a moderate reaction rate and a high enantioselectivity (35). Consequently, the enantioselective esterification of several primary alcohols, hydroxyakylsilanes, having a stereogenic silicon atom **(28–36)** was attempted (Fig. 14) with 5-phenylpentanoic acid by crude papain in water-saturated isooctane to study the effects of chain length between the silicon atom and the hydroxyl group **(28–30)** and the substitution groups on the silicon atom **(28,31–36)** (35). A short methylene chain between the silicon atom and the hydroxyl group and a phenyl substitution on the silicon atom were found to be essential for the high activity and the high enantioselectivity in the crude papain-catalyzed kinetic resolution of these organosilicon compounds. The optical purities of (+)-PhSi(Me)(Et)CH_2OH **(28)**, (+)-PhSi(Me)(Pr)CH_2OH **(31)**, (+)-*p*-Me-PhSi(Me)(Et)CH_2OH **(33)**, and (+)-*p*-F-PhSi(Me)(Et)CH_2OH **(34)**, obtained were 92 (yield 42% from the racemic substrate), 93 (yield 42%), 96 (yield 41%), and 99% ee (yield 41%), respectively.



	R^1	R^2	*n*
28	Ph	Et	1
29	Ph	Et	2
30	Ph	Et	3
31	Ph	Pr	1
32	Ph	n–C_6H_{13}	1
33	p–Me–Ph	Et	1
34	p–F–Ph	Et	1
35	n–C_6H_{13}	Et	1
36	Pr	Et	1

Figure 14. Organosilicon compounds having asymmetric silicon atoms used in the enantioselective esterification with a crude papain.



37a, b

	R^1	R^2	Enzyme
37a	Ph	Me	LCC
		N≍	LCV
37b	n-C_8H_{17}	Me	LCC
		N≍	LCV

Figure 15. Asymmetrization of prochiral organosilicon compounds by lipases. LCC, lipase from *Candida cylindracea*; LCV, lipase from *Chromobacterium viscosum*.

Djerourou and Blanco (36) demonstrated the preparation of silyl-chiral optically active compounds by the asymmetrization of prochiral organosilicon compounds with hydrolases. The prochiral 2-sila-1,3-propanediol derivatives **(37a,b)** were asymmetrized through transesterification with oxime ester or methylisobutyrate by lipase from *C. cylindracea* (LCC) and lipase from *Chromobacterium viscosum* (LCV) (Fig. 15), although the enantiomeric excess of the resulting esters did not exceed 70–76%.

The enantioselective hydrolysis of amide was also useful for the preparation of optically active organosilicon compounds. Optically active (*R*)-(1-aminoethyl)dimethylphenylsilane ((*R*)-**39**) was obtained in 40% yield with an enantiomeric purity of 92% ee by enantioselective hydrolysis of racemic dimethylphenyl[1-(phenylacetamido)-ethyl]silane **(38)** using immobilized penicillin G acylase from *Escherichia coli* 5K pHM 12 (Fig. 16) (37).

PREPARATION OF SILICON-CONTAINING AMINO ACIDS

Optically active nonnatural amino acids are very interesting and useful as the precursors of pharmaceuticals, agricultural chemicals, food ingredients, and so forth. Furthermore, they are also used as chiral auxiliaries for organic synthesis. Therefore, it will be very important to prepare novel nonnatural amino acids, such as silicon-containing ones. These amino acids can be expected to exhibit improved biological activities or new functions because of the characteristics of the silicon atom.

An efficient system has been developed to prepare optically active 3-trimethylsilylalanine (TMS-Ala) **(41)** by kinetic resolution with acylase I from porcine kidney from chemically synthesized *N*-acetyl-D,L-TMS-Ala **(40)** (Fig. 17) (38). Although both acylase I from porcine kidney and that from *Aspergillus melleus* showed activity on this substrate, the porcine enzyme was found to have much higher activity. The optimum pH of the reaction was 7.5, and the addition of 0.5 mM Co^{2+} accelerated the reaction. Optically pure L-**41** (>99% ee) was obtained in 72% yield. Furthermore, highly optically pure D-**41** (96% ee) could also be obtained in 76% yield by chemically hydrolyzing the residual substrate. **41** may be used as an analogue of leucine in the synthesis of bioactive peptides.

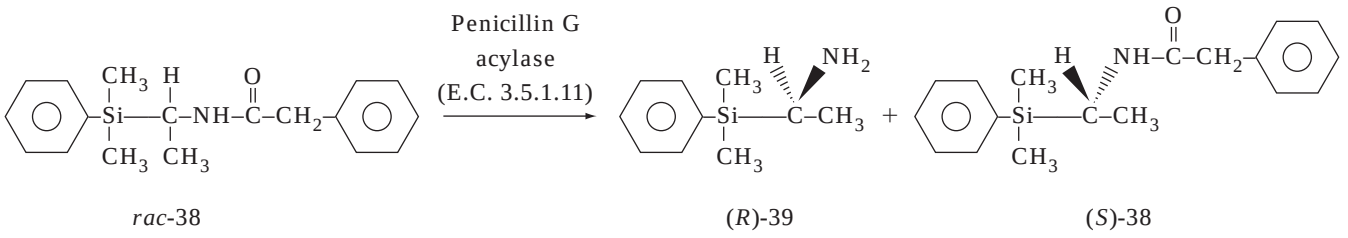


Figure 16. Enantioselective hydrolysis of a silicon-containing amide by penicillin G acylase.

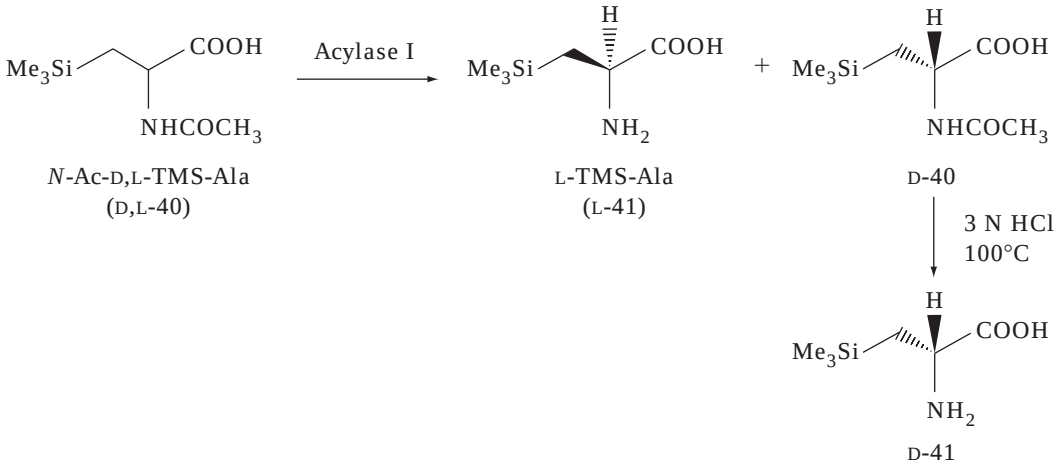


Figure 17. Preparation of optically active trimethylsilylalanine. TMS, trimethylsilyl.

As a homologue of phenylglycine and its analogues, which can be applied as side chains of penicillins and cephalosporins, optically active *p*-trimethylsilylphenylalanine **(44)** was prepared enzymatically. First, an attempt was made to obtain D-**44** from chemically synthesized D,L-5-(*p*-trimethylsilylphenylmethyl)hydantoin **(42)** via *N*-carbamoyl-D-*p*-trimethylsilylphenylalanine **(43)** through hydantoinase reaction, but none microbial strains tested exhibited the satisfactory activity to hydrolyze **42** to **43**. However, several strains, especially *Blastobacter* sp. A17p-4, showed good activity in the hydrolysis of **43** to yield **44**. Therefore, D,L-**43** prepared chemically from D,L-**42** was subjected to enantioselective hydrolysis to obtain optically active **44**. When D,L-**43** was hydrolyzed with cells or cell-free extract of *Blastobacter* sp., optical purity of D-**44** obtained was low because the cells contained not only *N*-carbamoyl-D-amino acid amidohydrolase (DCase) but also *N*-carbamoyl-L-amino acid amidohydrolase (LCase). DCase and LCase in the cell-free extract could be separated by DEAE-Sephacel-column chromatography. The optimum pH for the hydrolysis of D,L-**43** by partially purified DCase was 8.0, and addition of 2.5% *N*,*N*-dimethylformamide was effective to increase the substrate concentration without inactivation of the enzyme. Under the optimized conditions, highly optically pure (98% ee) D-**44** was obtained (39).

This process has been further improved by using the cell-free extract of *Blastobacter* sp. as the source of both DCase and LCase. As DCase of *Blastobacter* sp. had been known to be thermostable, the cell-free extract of this bacterium was treated at 50 °C and pH 7.0 for 40 min to inactivate LCase. After hydrolysis of D,L-**43** with the heat-treated cell-free extract, the product, D-**44**, and the residual substrate, L-**43**, were separated by silica-gel-column chromatography. L-**43** was further subjected to hydrolysis with nontreated cell-free extract as the LCase source after optimization of the reaction conditions (Fig. 18). Thus, optically pure D-**44** (99% ee) and L-**44** (99% ee) were obtained in yields of 80% and 89%, respectively, after purification (40).

CONCLUSION

This article outlines the bioconversion of various organosilicon compounds. In spite of the remarkable development of biotechnology, there have so far been only limited studies on bioconversion of organosilicon compounds. It has, however, been found that biocatalysts such as microorganisms, various types of hydrolases, and alcohol dehydrogenases are able to accept organosilicon compounds as substrates for transformation. Other types of enzymes will also be active on organosilicon compounds, although no experimental results are available at present. Introduction of biochemical methods into organosilicon chemistry has great potential because of the superior properties of biocatalysts, such as reaction selectivity, regioselectivity, enantioselectivity, stereoselectivity, and mild reaction condi-

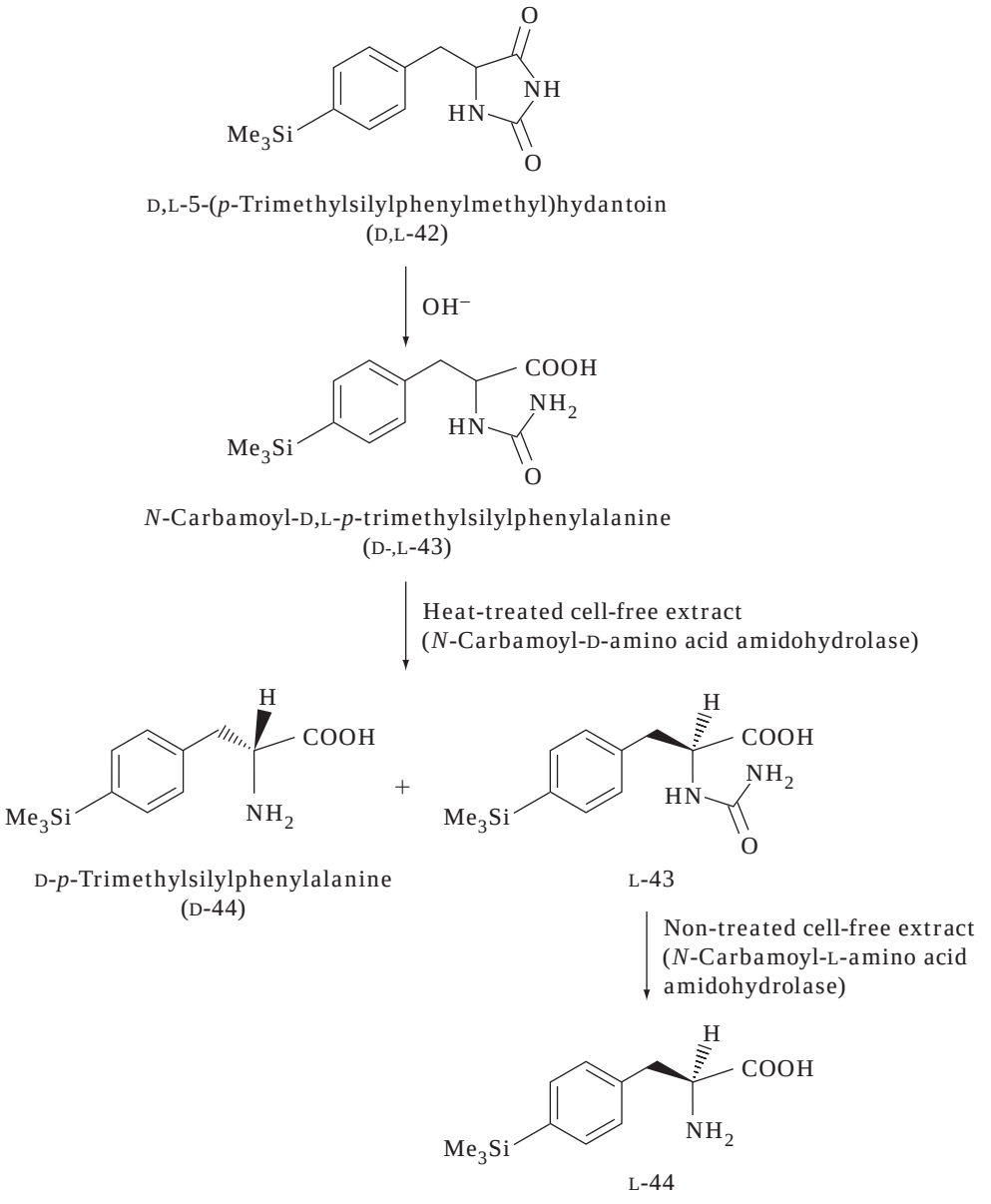


Figure 18. Preparation of optically active *p*-trimethylsilylphenylalanine.

tions. Furthermore, the application of the differences between organosilicon compounds and the corresponding carbon compounds to biochemical processes might make it possible to explore latent abilities of biocatalysts, giving fundamental information about biochemical and biological events. The field of organosilicon biochemistry, a fusion of biochemistry and organosilicon chemistry, will provide great possibilities to biotechnology in the future.

BIBLIOGRAPHY

1. E.W. Colvin, *Silicon in Organic Synthesis*, Butterworths, London, U.K., 1981.
2. R. Tacke and H. Zilch, *Endeavour New Ser.* **10**, 191–196 (1986).
3. A. Rioci, G. Seconi, and M. Taddei, *Chimicaoggi* **7**(9), 15–21 (1989).
4. R.J. Fessenden and J.S. Fessenden, *Adv. Organomet. Chem.* **18**, 275–299 (1980).
5. L.R. Garson and L.K. Kirchner, *J. Pharm. Sci.* **60**, 1113–1127 (1971).
6. C.E. Creamer, *Pharm. Technol.* March, 79–86 (1982).
7. R. Tacke and B. Becker, *Main Group Metal Chem.* **10**, 169–197 (1987).
8. A.D. Ryabov, *Angew. Chem. Int. Ed. Engl.* **30**, 931–941 (1991).
9. A. Tanaka and T. Kawamoto, *Chimicaoggi* **12**, 63–69 (1994).
10. R. Tacke, H. Linoh, B. Stumpf, W.R. Abraham, K. Kieslich, and L. Ernst, *Z. Naturforsch.* **38b**, 616–620 (1983).
11. C. Syldatk, H. Andree, A. Stoffregen, F. Wagner, B. Stumpf, L. Ernst, H. Zilch, and R. Tacke, *Appl. Microbiol. Biotechnol.* **27**, 152–158 (1987).

12. C. Syldatk, A. Stoffregen, F. Wuttke, and R. Tacke, *Biotechnol. Lett.* **10**, 731–736 (1988).

13. L. Fischer, S.A. Wagner, and R. Tacke, *Appl. Microbiol. Biotechnol.* **42**, 671–674 (1995).

14. R. Tacke, S.A. Wagner, S. Brakmann, F. Wuttke, U. Eilert, L. Fischer, and C. Syldatk, *J. Organomet. Chem.* **458**, 13–17 (1993).

15. R. Tacke, K. Fritsche, A. Tafel, and F. Wuttke, *J. Organomet. Chem.* **388**, 47–55 (1990).

16. R. Tacke, S. Brakmann, F. Wuttke, J. Fooladi, C. Syldatk, and D. Schomburg, *J. Organomet. Chem.* **403**, 29–41 (1991).

17. R. Tacke, H. Hengelsberg, H. Zilch, and B. Stumpf, *J. Organomet. Chem.* **379**, 211–216 (1989).

18. P. Huber, S. Bratovanov, S. Bienz, C. Syldatk, and M. Pietzsch, *Tetrahedron: Asymmetry* **7**, 69–78 (1996).

19. M.-H. Zong, T. Fukui, T. Kawamoto, and A. Tanaka, *Appl. Microbiol. Biotechnol.* **36**, 40–43 (1991).

20. K. Sonomoto, H. Oiki, and Y. Kato, *Enzyme Microb. Technol.* **14**, 640–643 (1992).

21. T. Fukui, M.-H. Zong, T. Kawamoto, and A. Tanaka, *Appl. Microbiol. Biotechnol.* **38**, 209–213 (1992).

22. Y. Tsuji, T. Fukui, T. Kawamoto, and A. Tanaka, *Appl. Microbiol. Biotechnol.* **41**, 219–224 (1994).

23. C.W. Bradshaw, W. Hummel, and C.-H. Wong, *J. Org. Chem.* **57**, 1532–1536 (1992).

24. M. Therisod and A.M. Klibanov, *J. Am. Chem. Soc.* **109**, 3977–3981 (1987).

25. M. Luyten, S. Muller, B. Herzog, and R. Keese, *Helv. Chim. Acta* **70**, 1250–1254 (1987).

26. K. Mori and T. Takeuchi, *Tetrahedron* **44**, 333–342 (1988).

27. M. Therisod, *J. Organomet. Chem.* **361**, C8–C10 (1989).

28. T. Kawamoto, K. Sonomoto, and A. Tanaka, *J. Biotechnol.* **18**, 85–92 (1991).

29. S.-W. Tsai and H.-J. Wei, *Enzyme Microb. Technol.* **16**, 328–333 (1994).

30. S.-W. Tsai and H.-J. Wei, *Biotechnol. Bioeng.* **43**, 64–68 (1994).

31. S.-W. Tsai and H.-J. Wei, *Biocatalysis* **11**, 33–45 (1994).

32. K. Fritsche, C. Syldatk, F. Wagner, H. Hengelsberg, and R. Tacke, *Appl. Microbiol. Biotechnol.* **31**, 107–111 (1989).

33. A. Uejima, T. Fukui, E. Fukusaki, T. Omata, T. Kawamoto, K. Sonomoto, and A. Tanaka, *Appl. Microbiol. Biotechnol.* **38**, 482–486 (1993).

34. G.L. Larson and E. Torres, *J. Organomet. Chem.* **239**, 19–27 (1985).

35. T. Fukui, T. Kawamoto, and A. Tanaka, *Tetrahedron: Asymmetry* **5**, 73–82 (1994).

36. A.-H. Djerourou and L. Blanco, *Tetrahedron Lett.* **32**, 6325–6326 (1991).

37. H. Hengelsberg, R. Tacke, K. Fritsche, C. Syldatk, and F. Wagner, *J. Organomet. Chem.* **415**, 39–45 (1991).

38. H. Yamanaka, T. Fukui, T. Kawamoto, and A. Tanaka, *Appl. Microbiol. Biotechnol.* **45**, 51–55 (1996).

39. Y. Tsuji, H. Yamanaka, T. Fukui, T. Kawamoto, and A. Tanaka, *Appl. Microbiol. Biotechnol.* **47**, 114–119 (1997).

40. H. Yamanaka, T. Kawamoto, and A. Tanaka, *J. Ferment. Bioeng.* **84**, 181–184 (1997).

OSMOTIC STRESS, SECRETION RATE

JOON SOO RYU
GYUN MIN LEE
Korea Advanced Institute of Science and Technology
Taejon, Korea

KEY WORDS

Antibody secretion rate
Hybridoma cells
Osmotic pressure

OUTLINE

INTRODUCTION

Osmolality is one of the most important physical properties of culture media. Osmolality is a measure of concentration and is defined as the molality of particles in a solution. Most cell culture media are designed to have an osmolality of 280 to 320 mOsm/kg through balancing the concentration of nutrients and osmolytes (1).

When cells are bathed in a solution of the same osmolality as the cytoplasm, they neither shrink nor swell. Osmotic pressure results when concentrations of dissolved solutes on the inside and on the outside of a cell are different. The effects of osmotic pressure on cells in culture are too complex to be discussed in detail in this article. Here, we focus on the effect of osmotic pressure on hybridoma cells in regard to cell growth and antibody secretion rates. In an effort to improve antibody productivity of hybridoma cells, the effect of osmotic pressure on hybridoma cells has been extensively studied. As a result, the following observations have been made.

1. The response of hybridoma cells to osmotic pressure varies among cell lines (2,3).
2. For most hybridomas, hyperosmotic pressure induced by either sugar or ionized salts suppresses cell growth, but it increases the specific antibody secretion rate (q_{Ab})(2–8).

3. The enhanced q_{Ab} resulting from hyperosmotic pressure is not transient and is maintained in successive passages (9).
4. Hypo-osmotic pressure also suppresses cell growth, but it does not enhance q_{Ab} (10).

Hyperosmotic pressure, which can be induced by addition of cheap salts to media, has been suggested as being an economical solution to increase q_{Ab} in hybridoma cell cultures. Despite the potential of commercial strategies based on hyperosmotic pressure, the use of hyperosmolar medium has not been popular mainly because cell growth is depressed at elevated osmolality. Hence, the enhanced q_{Ab} in a batch culture does not result in a substantial increase in the final antibody concentration (2,4–6). To overcome this drawback, several strategies, such as adaptation of cells to hyperosmotic pressure (5,8), use of osmoprotectants in media (6,7,11), and two-stage culture (12), have been applied for improved antibody production. This article deals with the responses of hybridoma cells to osmotic pressure and its application to hybridoma cell cultures for improved antibody production.

OSMOLALITY

Osmolality is a measure of concentration and is often expressed in units of osmoles. The most commonly used abbreviation, based on milliosmoles per kilogram, is mOsm/kg. One osmole of a substance is equal to 1 g mol divided by the number of particles formed by the dissociation of the molecules. If a solution is composed of a nondissociable solute, the osmolality is simply equal to the molality. For a solution of a dissociable salt, 1 mol solution is n osmole, where n is the number of ions produced per molecule.

In the literature, the terms osmolarity and osmolality are frequently used interchangeably. Osmolality is related to osmolarity in the same manner as molality is related to molarity. Thus, osmolality is a more accurate description of concentration because it is a mass per unit mass measurement and is independent of temperature.

Osmolality values of media cannot be easily calculated because media contain a wide range of compounds as well as degrees of ionization. The osmolality of the media is determined more easily by using either vapor pressure elevation or freezing-point depression. The osmometers operate based on the principle that changes in osmolalities of a solution result in changes in certain properties such as vapor pressure and freezing point. Osmolality values measured by osmometers give reliable and reproducible results with a standard deviation of less than 5%.

MEDIUM OSMOLALITY

Medium osmolality is one of the important factors that is frequently overlooked in cultivating mammalian cells. Cell culture medium is designed to have osmolality in the range of 280 to 320 mOsm/kg, basically to mimic the osmolality of serum at 290 mOsm/kg (1). The major component in maintaining osmotic pressure in cell culture media is NaCl. Other inorganic ions (K^+, Mg^{2+}, Ca^{2+}) and glucose also contribute significantly. On the other hand, larger molecules like proteins play a lesser role, and changes in their concentrations do not affect osmolality significantly.

The measurement of osmolality is a useful control step if one makes up media. It is particularly important to check osmolality if alterations are made in the constitution of the medium. During batch culture, nutrients are consumed and metabolites such as lactate and ammonia are accumulated in the media. However, medium osmolality does not change significantly (9). The osmolality of culture media can change drastically as a result of evaporation when small-scale cultures such as Petri-dish or open-plate cultures are incubated in nonhumidified incubators. In this case, a slightly hypotonic medium may be better to compensate for evaporation during incubation. Osmolality should be also carefully monitored in fed-batch culture where medium concentrates are repeatedly fed to the culture. Addition of medium concentrates results in a substantial increase in medium osmolality, although they do not contain major salts. Thus, the culture longevity, which is important for the design of high-yield processes, is often limited by elevated osmolality caused by repeated nutrient feedings (13,14).

EFFECT OF OSMOTIC PRESSURE ON HYBRIDOMA CELL GROWTH AND ANTIBODY PRODUCTION

Medium osmolality in the range of 280 to 320 mOsm/kg is known to be quite acceptable for most cells; therefore, most commercially available media are designed to have an osmolality of this range (1,15). As the osmolality goes far from this standard range, some obvious changes in cell size, cell shape, nutrient transport, cellular synthesis, specific growth rate (μ), and q_{Ab} occur. Among these changes, the most important changes for cell culturists are probably those in μ and q_{Ab}. The common osmolyte used for altering the medium osmolality is NaCl. As summarized in Table 1, elevated osmolality suppresses cell growth and increases q_{Ab}. The q_{Ab} increased up to 3.2-fold for cells grown above 300 mOsm/kg and up to 435 mOsm/kg. However, because of the depressed cell growth at higher osmolalities, similar final antibody titers were obtained in most cases.

Regardless of serum concentrations, cultivation methods (static and suspension cultures), and basal media (IMDM, DMEM, and RPMI 1640), hyperosmotic pressure substantially enhanced q_{Ab} of S3H5/γ2bA2 hybridoma in batch culture by depressing cell growth, although the degree of enhancement in q_{Ab} was influenced (2,9,10).

Ozturk and Palsson (4) altered the osmolality of the medium by addition of both ionic (NaCl- and phosphate-buffered saline) and nonionic substances (sucrose). Similar data were obtained regardless of the osmolytes used. Thus, the changes in cells in hyperosmolar media described earlier appear to be caused only by medium osmolality.

If the enhanced q_{Ab} at elevated osmolality is not transient, it can be used to maximize antibody productivity in perfusion culture systems where cell growth and production are separated. As shown in Figure 1, the enhanced q_{Ab} of hybridoma at elevated osmolality decreased while the cells were adapting to hyperosmotic pressure in long-term,

Table 1. Effect of Osmolality on Antibody Production in Batch Culture

Cell line	Osmolality (mOsm/kg)	Media	Relative q_{Ab}	Relative antibody titer	References
Hybridoma (S3H5/γ2bA2)	281	IMDM	1.0	1.00	9,15
	390	IMDM	2.5	0.86	
Hybridoma (DB9G8)	281	IMDM	1.0	1.0	2,9,15
	390	IMDM	1.1	0.9	
	440	IMDM	3.2	0.4	
Hybridoma (167.4G5.3)	290	IMDM	1.0	1.0	4
	435	IMDM	>2.0	1.0	
Hybridoma (2HG11)	300	eRDF	—	1.0	5,8
	350	eRDF	—	1.1	
Hybridoma (AB2-143.2)	352	DMEM-HamF12	1.0	—	3
	488	DMEM-HamF12	1.5	—	
Hybridoma (6H11)	330	DMEM	1.0	1.0	6
	510	DMEM	1.7–1.9	0.5	
Recombinant NSO myeloma	270	IMDM	1.0	1.0	13
	400	IMDM	1.6	0.8	

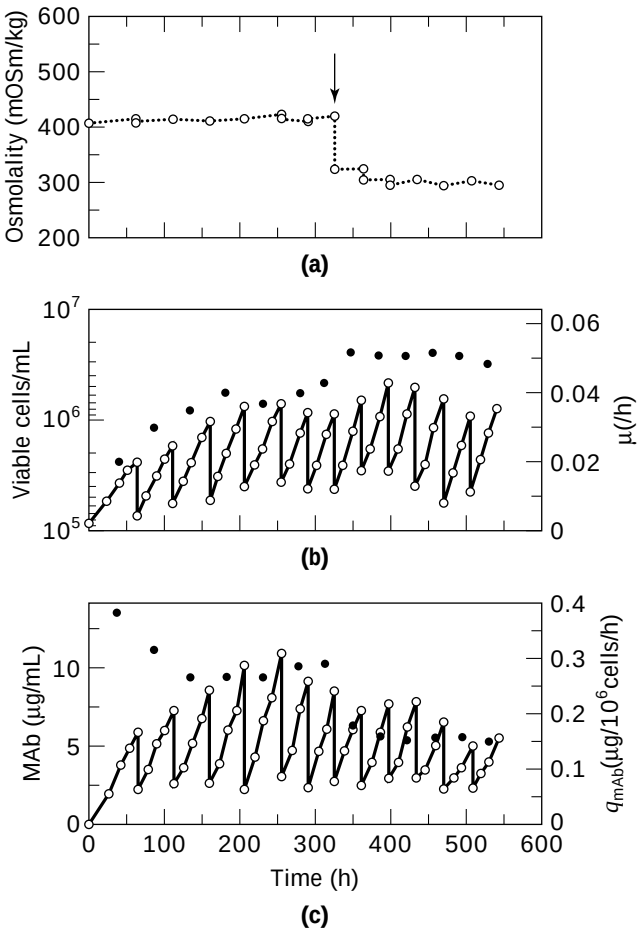


Figure 1. Repeated-fed batch culture in a spinner flask. (**a**) Osmolality, (**b**) viable cell concentration (○) and μ (●); (**c-**) antibody concentration (○) and q_{Ab} (●). The *arrow* indicates the time of switching hyperosmolar medium to standard medium.

repeated-fed-batch culture (9). However, the cells even after adaptation could still maintained the significantly enhanced q_{Ab} when compared with the q_{Ab} of cells obtained from standard medium in the batch culture, indicating the potential of using hyperosmolar medium for improved antibody production in a perfusion culture. In addition, when the hyperosmolar medium switched back to a standard medium, the physiological changes of the cells resulting from hyperosmotic pressure rapidly disappeared. Accordingly, the q_{Ab} and μ in perfusion systems may be controlled simply by changing the osmolality of medium. However, it should be noted that all the hybridomas do not display enhanced q_{Ab} at elevated osmolality. Cell-line specificity will be discussed in a later section.

Although much work has been performed on the response of hybridomas to hyperosmotic pressure, there are, to date, few reported studies on the response of hybridomas to hypoosmotic pressure.

Ryu and Lee reported the response of hybridomas (S3H5/γ2bA2 and DB9G8) to hypo-osmotic pressure resulting from NaCl subtraction (10). As shown in Figures 2 and 3, both hybridomas showed similar responses to hyperosmotic pressure in regard to cell growth and antibody production. The cell growth and antibody production at 276 mOsm/kg were comparable to those at 329 mOsm/kg (standard DMEM). Both cells grew well at 219 mOsm/kg, although their growth and antibody production were slightly decreased. When the osmolality was further decreased to 168 mOsm/kg, the cell growth did not occur. When subjected to hyperosmotic stress, both cells displayed significantly enhanced q_{Ab}. However, the cells subjected to hypoosmotic pressure did not display enhanced q_{Ab}.

Taken together, medium osmolality is one of the important factors that can be easily manipulated by addition or subtraction of inorganic salts or sugars. When subjected to hyperosmotic pressure, most hybridomas display enhanced q_{Ab} by suppressing cell growth. Hypo-osmotic pressure also suppresses cell growth but does not enhance q_{Ab}.

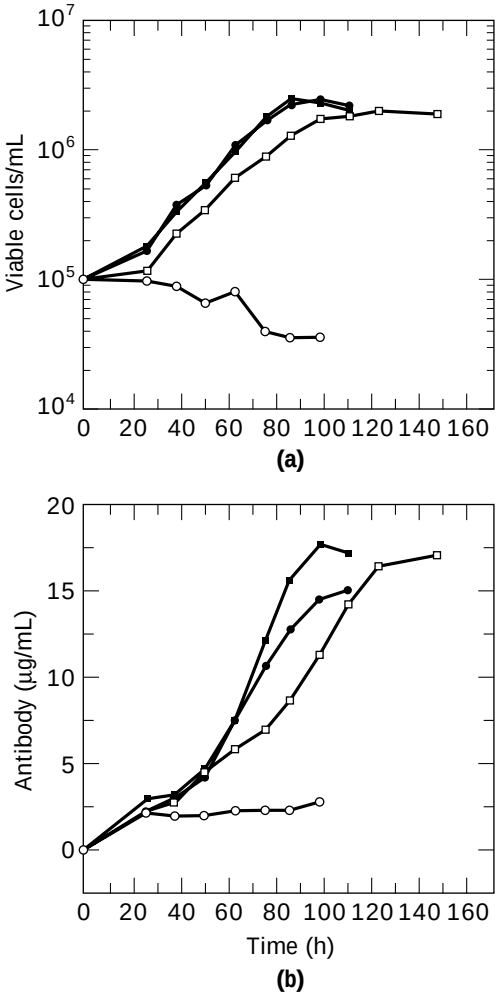


Figure 2. Cell growth (**a**) and antibody production (**b**) during a batch culture of S3H5/γ2bA2 hybridoma cells in hyperosmolar media. Data points: ○, 168; □, 219; ●, 276; ■, 329 mOsm/kg.

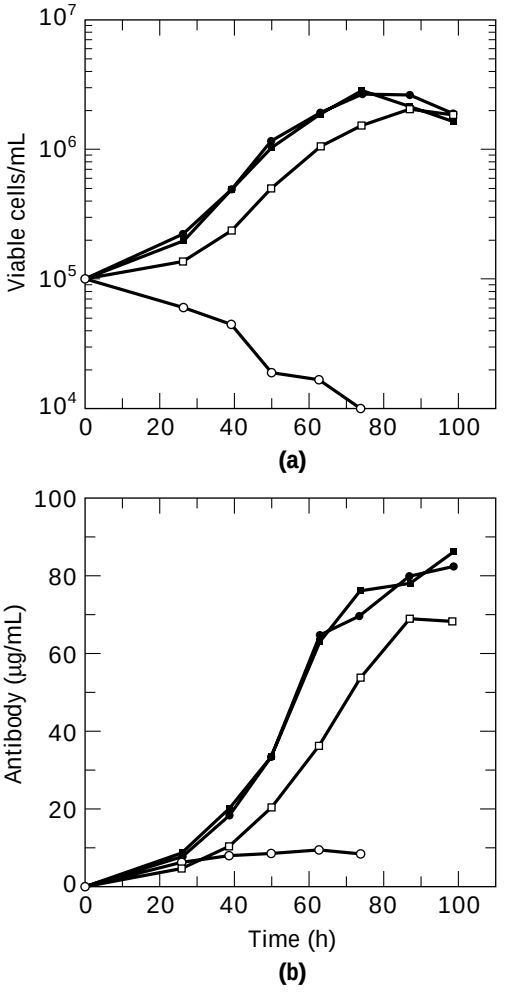


Figure 3. Cell growth (**a**) and antibody production (**b**) during a batch culture of DB9G8 hybridoma cells in hyperosmolar media. Data points: ○, 168; □, 219; ●, 276; ■, 329 mOsm/kg.

Table 2. Summary of Maximum Viable Cell Density, Maximum Antibody Titer, and q_{Ab} at Different Osmolytes

Osmolyte used	Max. viable cell density ($\times 10^6$ cells/mL)	Max. antibody titer (μg/mL)	q_{Ab} (pg/cell/h)
Control	2.92 ± 0.12	156 ± 9	0.75
NaCl	2.76 ± 0.12	299 ± 25	1.46
Sucrose	2.05 ± 0.21	257 ± 13	1.50
KCl	2.35 ± 0.21	198 ± 3	1.08

Source: Data from Oh et al. (8).

Table 3. Comparison of Maximum Viable Cell Density, Maximum Antibody Titer, and μ between Unadapted and Adapted Cells

	Unadapted cells			Adapted cells		
Osmolality (mOsm/kg)	Maximum viable cell density ($\times 10^5$/mL)	Maximum antibody titer (μg/mL)	μ (h^{-1})	Maximum viable cell density ($\times 10^5$/mL)	Maximum antibody titer (μg/mL)	μ (h^{-1})
300	34.6	140	0.046	34.6	140	0.046
350	8.4	150	0.027	31.3	265	0.036
400	7.9	75	0.038	21.3	220	0.033

Source: Data from Oh et al. (5).

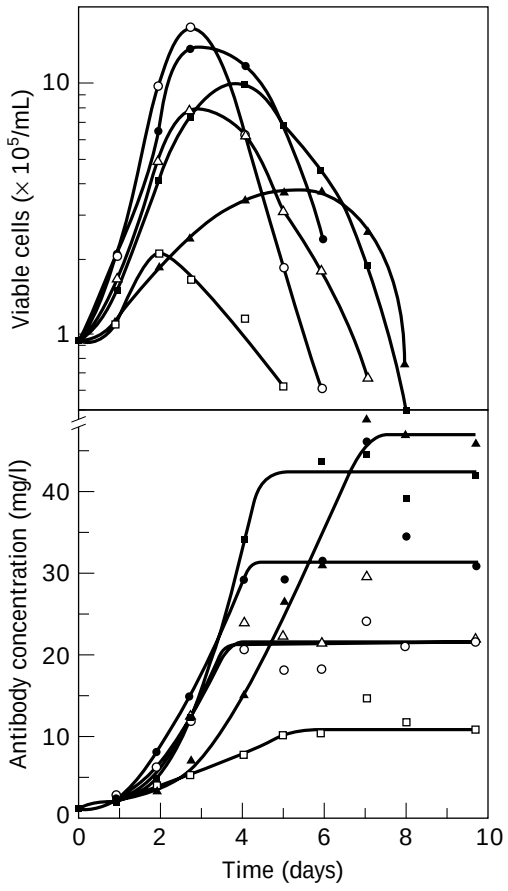


Figure 4. Hybridoma growth and antibody production in NaCl-stressed growth media in the presence and absence of glycine betaine (15 mM) (7). ○, Control, 330 mOsm/kg; △, 60 mM NaCl added, 450 mOsm/kg; ●, 60 mM NaCl and 15 mM glycine betaine added, 465 mOsm/kg; □, 100 mM NaCl added, 510 mOsm/kg; ■, 100 mM NaCl and 15 mM glycine betaine added, 525 mOsm/kg; ▲, 140 mM NaCl and 15 mM glycine betaine added, 610 mOsm/kg.

ENHANCED SPECIFIC ANTIBODY SECRETION RATE OF HYBRIDOMAS RESULTING FROM HYPEROSMOTIC PRESSURE IS CELL-LINE SPECIFIC

As mentioned earlier, some hybridomas do not display enhanced q_{Ab} at elevated osmolality.

Reddy and Miller reported the effects of osmotic pressure on antibody production in hybridoma cells that differed in production kinetics (3). AB2-143.2 cells exhibit non-growth-associated antibody production, whereas IND1 cells exhibit growth-associated production. Both hybridomas displayed enhanced q_{Ab} after abrupt batch osmotic stress. AB2-143.2 cells also increased q_{Ab} in response to gradual osmotic stress in continuous culture. In contrast, IND1 cells decreased q_{Ab} during gradual osmotic stress. It was hypothesized that this difference between cell lines in response to gradual osmotic stress in continuous culture might be related to the fact that antibody production is growth-associated in IND1, but not in AB2-143.2 cells.

Lee and Park (2) reported that two murine hybridomas exhibiting non-growth-associated antibody production (S3H5/γ2bA2 and DB9G8) showed different responses to hypersomotic pressure regarding q_{Ab}, although they showed similar depression of cell growth in hyperosmolar media. The q_{Ab} of S3H5/γ2bA2 hybridoma in a hyperosmolar medium (396 mOsm/kg) was enhanced by approximately 180% when compared with that in a standard medium (283 mOsm/kg), whereas q_{Ab} of DB9G8 hybridoma in the same hyperosmolar medium was enhanced by only 10%. Accordingly, regardless of antibody production kinetics of the cell lines, enhanced q_{Ab} of hybridomas resulting from hyperosmotic pressure appears to be cell-line specific. Thus, the strategy of using hyperosmolar media for improved antibody production cannot be applied for all cell lines.

EXPLANATION FOR ENHANCED Q_{AB} AT ELEVATED OSMOLALITY

Although the detailed mechanism of enhanced q_{Ab} is not clearly understood at the basic cellular level, several explanations based on noticeable changes induced by hyperosmotic pressure can be proposed for possible influences of NaCl on cellular functions and antibody production.

1. The increase in osmotic pressure may enhance the transport of nutrients, in particular, amino acids into the cells (5,8,11). Amino acid transport regulations in animal cells are controlled by both Na^+ ion-dependent and Na^+ ion-independent transport systems (16). Excess Na^+ ions on the outer membrane would increase the transmembrane electrochemical gradient, facilitating the simultaneous flow of ions and amino acids into the cells. Thus, the feeding of excess nutrients into the cell could increase cytoplasmic protein production, especially antibody productivity in hybridomas. The release of additional energy and precursors as a result of increased metabolism is not diverted to growth but to the synthesis and secretion of antibody (5). For hypersomotically stressed hybridomas, amino acid uptake by Na^+-dependent system A increases, whereas the activity of other Na^+-independent systems remains relatively unchanged (8). Proline, alanine, and glutamine, which are transported by this Na^+-dependent system A under hyperosmotic condition, are known to act as osmoprotective agents to compensate for osmotic imbalance of the environment (7,17). This observation would support a model for cell volume swelling and increased q_{Ab} at elevated osmolality. However, this postulate will have to be investigated further because it does not explain the workings of other osmolytes like sucrose or KCl (4,6,8). Table 2 shows that the enhancement of q_{Ab} of hybridomas subjected to hyperosmotic stress is not influenced significantly by the sources of osmotic stress.

2. The increase in osmotic pressure may induce transcriptional activation because of the change of chromatin structure, as is observed in the case of butyrate treatment (8,18–22). Acetylation of histone residue is thought to act as a transcription-stimulating event in many organisms. Unbinding of the histones from DNA results in the disper-

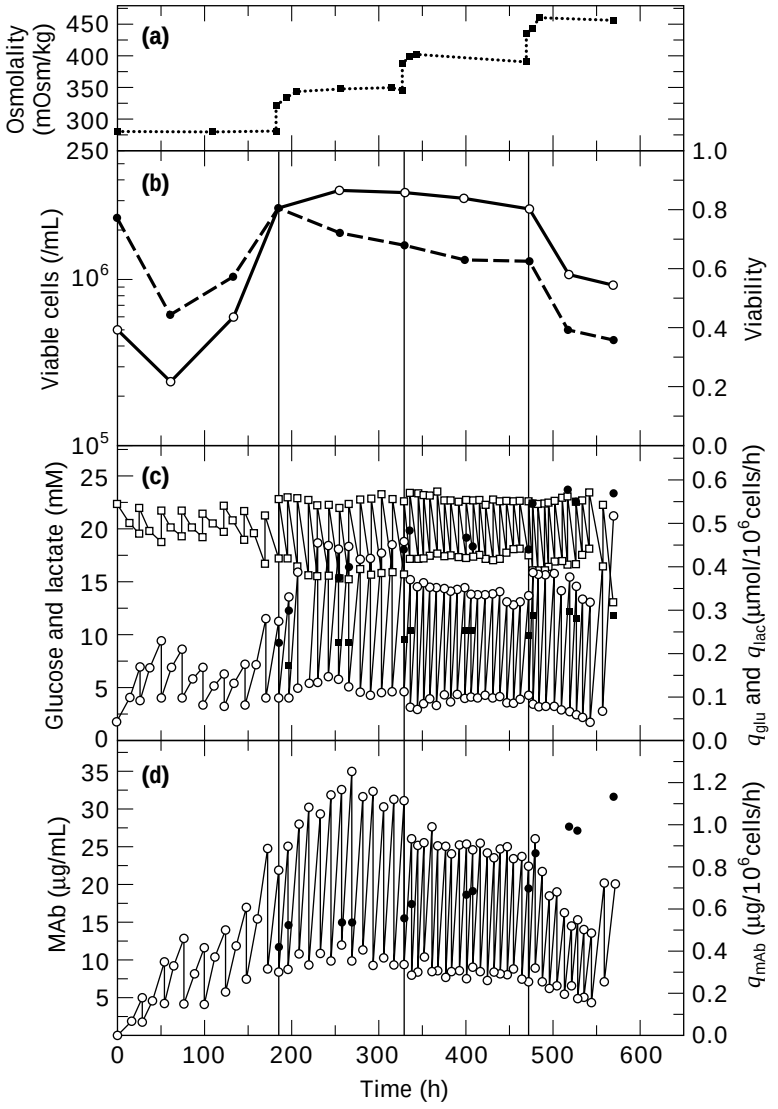


Figure 5. Immobilized cell culture with gradual hyperosmotic stress: (**a**) osmolality; (**b**) viable cell concentration of immobilized cells (○) and viability (●); (**c**) glucose (□) and lactate concentrations (○), q_{glu} (■), and q_{lac} (●); (**d**) antibody concentration (○) and q_{Ab} (●). The *vertical lines* indicate the time of elevating the osmolality of medium.

sal of the chromatin structure, which may expose segments of the chromatin, making them more accessible to RNA polymerase for mRNA transcription (21). In fact, elevated mRNA expression of certain proteins has been reported after treatment of butyrate (18,22–24).

3. The increase in osmotic pressure may stimulate the expression of chaperone-like proteins in the endoplasmic reticulum (ER) such as protein disulfide isomerase (PDI) or glucose-regulating proteins (GRPs) that are involved in protein processing and secretion (25,26).

All these postulates, however, should be confirmed by concrete experimental data, and solid, detailed mechanisms or sequential events that explain increased q_{Ab} will have to be discovered.

APPLICATION OF HYPEROSMOTIC PRESSURE TO ANTIBODY PRODUCTION PROCESSES

Hyperosmotic stress resulting from NaCl addition has been suggested as being an economical solution to increase the q_{Ab} in hybridoma cultures. However, the use of hyperosmolar medium has not been popular probably because of the following reasons. Because the enhanced q_{Ab} of hybridomas resulting from hyperosmotic stress is cell-line specific (2,3), the hyperosmolar medium cannot be used for all cell lines. Furthermore, cell growth is depressed at elevated osmolality. Hence, the enhanced q_{Ab} in batch culture does not result in a substantial increase in the final antibody titer (2,4–6). To overcome these drawbacks, the following strategies have been applied for improved antibody production.

1. Hybridoma cells are first adapted to higher osmolality media to elevate antibody levels (5,8). Before adaptation, 2HG11 hybridoma cells placed in higher osmotic pressures (350 and 400 mOsm/kg) were severely suppressed in growth down to 25% of the control (300 mOsm/kg), although total antibody titers achieved were similar to the control. After 1 week of adaptation to 350 and 400 mOsm/kg media, cell growth was not as drastically suppressed. As a result, considerably higher antibody levels were obtained at these elevated osmolalities (Table 3).

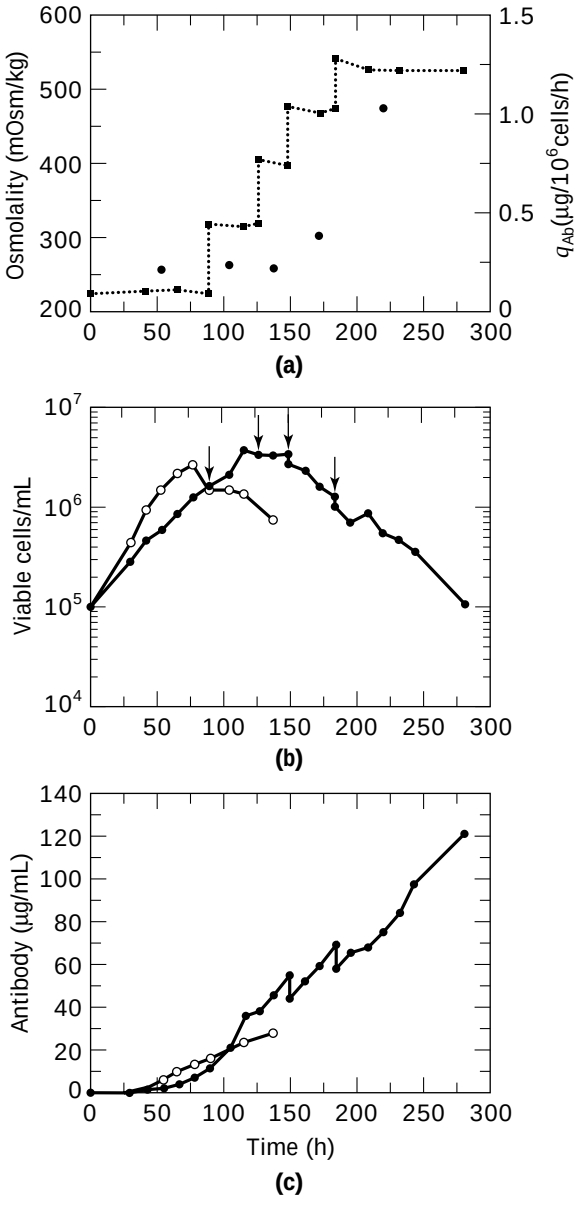


Figure 6. Fed-batch culture using a hypo-osmolar medium as an initial basal medium. (**a**) Changes in medium osmolality (■) and q_{Ab} (●). (**b**) Cell growth in batch (○) and fed-batch (●) cultures. The *arrows* indicate the addition of medium concentrates. (**c**) Antibody production in batch (○) and fed-batch (●) cultures.

2. Osmoprotective compounds, especially glycine betaine, are included in hyperosmolar media (6,7,11). Osmoprotective agents are known to exert osmoprotective effects by being accumulated inside cells to compensate for osmotic imbalance between in and out of the cells. Glycine betaine, sarcosine, proline, and glycine have all been shown to be strong osmoprotective compounds for 6H11 hybridoma cells under various hyperosmotic stress, with glycine betaine being the most effective of the four (6). When glycine betaine was included in the hyperosmolar media, q_{Ab} was increased up to 2.6-fold and maximum antibody titer up to twofold over that obtained in the control culture, as shown in Figure 4. A similar pattern of response was observed when other osmoprotective compounds (sarcosine, proline, glycine) were included in the hyperosmolar media. The improvement of cell growth by use of osmoprotective compounds did not cost increased q_{Ab} obtained under hyperosmotic condition, suggesting that medium osmolality, rather than growth rate, will determine the q_{Ab} by 6H11 hybridoma cells growing in hyperosmolar media. The use of osmoprotective compounds does not influence hyperosmotically enhanced specific metabolic rates. Accordingly, the simultaneous use of hyperosmotic stress and osmoprotective compounds can be a means to increase both q_{Ab} and maximum antibody titer in hybridoma cell culture.

3. The enhanced q_{Ab} resulting from hyperosmotic stress can be used for a higher volumetric antibody productivity in perfusion systems where cell growth and antibody production are separated. In perfusion systems such as hydrogel beads or hollow fibers, the cells maintain minimal growth rates after the cell concentrations are saturated. Thus, after achieving high cell concentration in perfusion systems, the switch of standard medium to hyperosmolar medium may significantly improve the volumetric productivity. A feasibility of using hyperosmolar medium resulting from NaCl addition for improved antibody production in a perfusion system was demonstrated by using calcium-alginate-immobilized S3H5/γ2bA2 hybridoma cells (12). First, immobilized cells were grown in the medium supplemented with 10% serum to achieve high cell concentration rapidly. Second, after the cell concentration reached more than 2×10^6 cells/mL, the medium was switched to the hyperosmolar medium supplemented with 1% serum for economical antibody production. An abrupt increase in osmolality, however, inhibited cell growth, resulting in no increase in volumetric antibody productivity. On the other hand, gradual increase in osmolality allowed further cell growth while maintaining the enhanced q_{Ab} of immobilized cells, as shown in Figure 5. Accordingly, the volumetric antibody productivity was increased by about 40% when compared with that in the control immobilized cell culture. This enhancement of volumetric antibody productivity of immobilized hybridoma cells by applying hyperosmolar stress suggests the potential of using hyperosmolar medium in other perfusion systems for improved antibody production.

APPLICATION OF HYPO-OSMOTIC PRESSURE TO ANTIBODY PRODUCTION PROCESSES

Cell culture longevity in fed-batch culture of hybridomas is often limited by elevated medium osmolality caused by repeated nutrient feedings (13,14). Shotwise feeding of 10× DMEM concentrates elevated the osmolality of medium up to 540 mOsm/kg at the end of fed-batch culture of S3H5/γ2bA2 hybridoma, which is known to be lethal to most hybridomas. S3H5/γ2bA2 hybridoma has been shown to grow without significant growth depression at 219 mOsm/kg in DMEM supplemented with 10% fetal bovine serum. In order to improve culture longevity in fed-batch culture of S3H5/γ2bA2 hybridoma, a hypo-osmolar medium (223 mOsm/kg) was used as an initial basal medium

(27). As shown in Figure 6, the use of hypo-osmolar medium delayed the onset of severe cell death resulting from elevated osmolality and allowed one more addition of 10× DMEM concentrates to the culture. As a result, a final antibody concentration obtained was 121.5 μg/mL, which is approximately 1.5-fold higher compared to fed-batch culture using a standard medium (335 mOsm/kg). When compared to batch culture, a 5.2-fold increase in the final antibody titer was achieved. Taken together, the use of hypo-osmolar medium as an initial medium in fed-batch culture improved culture longevity of S3H5/γ2bA2 hybridoma, resulting in substantial increase in the final antibody titer.

In conclusion, hyperosmotic pressure, which can be induced by addition of cheap salts to media, has been suggested as being an economical solution to increase q_{Ab} in hybridoma cell cultures. The utilization of osmotic pressure for improved antibody production in hybridoma cell cultures will be substantiated in future.

BIBLIOGRAPHY

1. R.I. Freshney, *Culture of Animal Cells: A Manual of Basic Technique*, 3rd ed., Wiley-Liss, New York, 1994, pp. 83–84.
2. G.M. Lee and S.Y. Park, *Biotechnol. Lett.* **17**, 145–150 (1995).
3. S. Reddy and W.M. Miller, *Biotechnol Prog.* **10**, 165–173 (1994).
4. S.S. Ozturk and B.O. Palsson, *Biotechnol. Bioeng.* **37**, 989–993 (1991).
5. S.K.W. Oh, P. Vig, F. Chua, W.K. Teo, and M.G.S. Yap, *Biotechnol. Bioeng.* **42**, 601–610 (1993).
6. K. Øyaas, T.E. Ellingsen, N. Dyrset, and D.W. Levine, *Biotechnol. Bioeng.* **44**, 991–998 (1994).
7. K. Øyaas, T.E. Ellingsen, N. Dyrset, and D.W. Levine, *Biotechnol. Bioeng.* **43**, 77–89 (1994).
8. S.K.W. Oh, F.K.F. Chua, and A.B.H. Choo, *Biotechnol. Bioeng.* **46**, 525–535 (1995).
9. S.Y. Park and G.M. Lee, *Bioproc. Engr.* **13**, 79–86 (1995).
10. J.S. Ryu and G.M. Lee, *Biotechnol. Bioeng.* **55**, 565–570 (1997).
11. K. Øyaas, T.E. Ellingsen, N. Dyrset, and D.W. Levine, *Cytotechnology* **17**, 143–151 (1995).
12. S.Y. Park and G.M. Lee, *Biotechnol. Bioeng.* **48**, 699–705 (1995).
13. T.A. Bibila, C.S. Ranucci, K. Glazomitsky, B.C. Buckland, and J.G. Aunins, *Biotechnol. Prog.* **10**, 87–96 (1994).
14. T.A. Bibila and D.K. Robinson, *Biotechnol. Prog.* **11**, 1–13 (1995).
15. J.S. Ryu and G.M. Lee, *Bioproc. Engr.* **16**, 305–310 (1997).
16. M.A. Shotwell, M.S. Kilberg, and D.L. Oxender, *Biochem. Biophys. Acta.* **737**, 267–284 (1983).
17. L. Silvotti, P.G. Petronini, A. Mazzini, G. Piedimonte, and A.F. Borghetti, *Exp. Cell Res.* **193**, 253–261 (1991).
18. M. Korsaka, T. Nakayama, A. Tsukamoto, and K. Kurokawa, *Cancer Res.* **50**, 3101–3105 (1987).
19. L.C. Boffa, R.J. Gruss, and V.G. Allfrey, *J. Biol. Chem.* **256**, 9612–9621 (1981).
20. A. Chabanas, E. Khoury, P. Goeltz, P. Froussard, R. Gjerset, B. Dod, H. Eisen, and J.J. Lawrence, *J. Mol. Biol.* **183**, 141–151 (1985).
21. M.J. Kruh, *Mol. Cell. Biochem.* **42**, 65–82 (1982).
22. I. Parker, J.B. deHaan, and W. Givers, *J. Biol. Chem.* **261**, 2786–2790 (1986).
23. A.J. Dorner, L.C. Wasley, and R.J. Kaufman, *J. Biol. Chem.* **264**, 20602–20607 (1989).
24. D.P. Palermo, M.E. DeGraaf, K.R. Marotti, E. Rehberg, and L.E. Post, *J. Biotechnol.* **19**, 35–48 (1990).
25. R.B. Freedman, R.E. Brockway, and N. Lambert, *Biochem. Soc. Trans.* **12**, 929–932 (1984).
26. A.S. Lee, *Curr. Opin. Biol.* **4**, 267–273 (1992).
27. J.S. Ryu and G.M. Lee, *Biotechnol. Bioeng.* **62**, 120–123 (1999).

PANTOTHENIC ACID AND RELATED COMPOUNDS

SAKAYU SHIMIZU
MICHIHIKO KATAOKA
Kyoto University
Kyoto, Japan

KEY WORDS

β-Alanine
Coenzyme A
Lactonohydrolase
Pantetheine
Pantoic acid
Pantothenate kinase
4-Phosphopantetheine
4-Phosphopantothenoyl-L-cysteine decarboxylase
4-Phosphopantothenoyl-L-cysteine synthetase

OUTLINE

INTRODUCTION

Pantothenic acid (*R*- or D-(+)-*N*-(2,4-dihydroxy-3,3-dimethyl-1-oxobutyl)-*β*-alanine; for chemical formula, see Table 1) was first isolated in the 1930s from liver and found to be an essential growth factor for yeasts (15). During this period it was also identified independently with the chick antidermatitis factor, the filtrate factor, the chick antipellagra factor, and an essential growth factor for lactic acid bacteria. Later these activities were shown to be identical with those of pantothenic acid. Because the factor could be obtained from a variety of plants and animal tissues, Williams named it pantothenic acid, meaning "from everywhere." The compound is also referred to as vitamin B_5. It was independently synthesized by two groups in 1940.

Pantothenic acid occurs in living organisms of all types, both in a free form and in conjugated forms such as coenzyme A, pantetheine (*Lactobacillus bulgaricus* factor), and 4′-phosphopantetheine (*Acetobacter suboxydans* factor) (see Table 1). The coenzyme form of the vitamin, coenzyme A, was discovered as an essential cofactor for the acetylation of sulfonamide in the liver and the acetylation of choline in the brain by Lipmann and coworkers (16). It has been since identified with "active acetate" and has been found to be essential for a variety of biochemical transacylation reactions. 4′-Phosphopantetheine is also a coenzyme form of pantothenic acid that functions as a prosthetic group of the acyl carrier protein of fatty acid synthetase, citrate-cleaving enzyme, and enzymes involved in the synthesis of peptide antibiotics. These early studies have been reviewed by several authors (15–20).

CHEMISTRY

Pantothenic acid can be obtained as a colorless, viscous oil by drying under high vacuum in P_2O_5. It is an acid and has a marked tendency to absorb water from the air. Under alkaline hydrolysis, it breaks down into *β*-alanine and pantoic acid. The latter readily forms a lactone, D-(−)-pantolactone, in acid solution or on heating. Acid hydrolysis of pantothenic acid gives *β*-alanine and pantolactone. Pantothenic acid is soluble in water, ethyl acetate, dioxane, acetic acid, ether, and amyl alcohol but is insoluble in benzene and chloroform.

The structure of pantothenic acid contains a single asymmetric center, so that it is optically active; only the natural D-(+)-isomer has vitamin activity. The absolute configuration of the vitamin has been defined as *R* (21). The conformation of the vitamin has also been reported (22).

The calcium salt of pantothenic acid, which can be obtained as needle crystals from methanol, is moderately hygroscopic and is rather more stable to heat, air, and light than is the free acid. It is soluble in water and glycerol and slightly soluble in alcohol and acetone. Reviews by Baddiley (23) and Wagner and Folkers (17) summarized early studies on the chemistry of pantothenic acid.

The naturally occurring derivatives of pantothenic acid (see Table 1) can be grouped into three types on the basis of their chemical structures: simple pantothenate derivatives, pantetheine derivatives in which cysteamine (or its analogs) attaches by an amide linkage, and coenzyme A derivatives in which the pantetheine is adenosylylated (Fig. 1).

Pantothenyl alcohol, an alcohol analog of pantothenic acid, is also a pharmaceutically important unnatural derivative. Unnatural analogs of pantothenic acid and coenzyme A have been reviewed by Shimizu (10).

Table 1. Panthothenic Acid and Its Naturally Occurring Derivatives

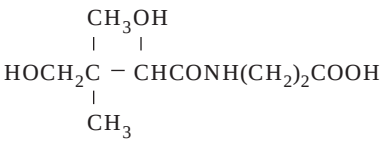
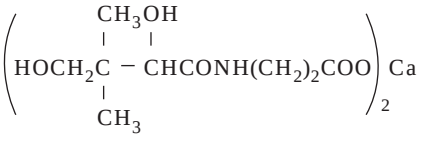
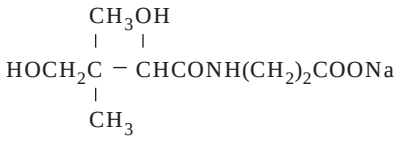
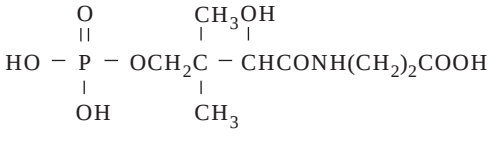
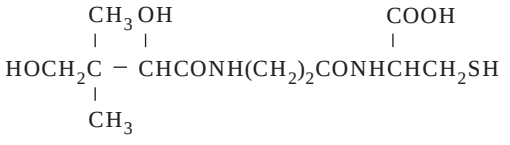
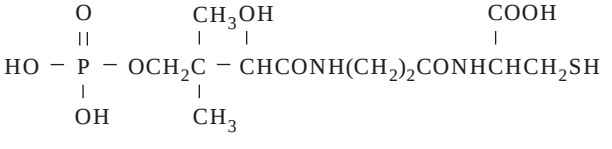
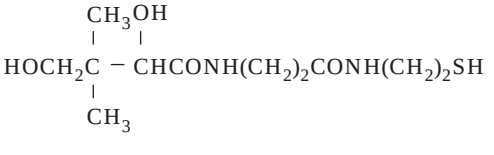
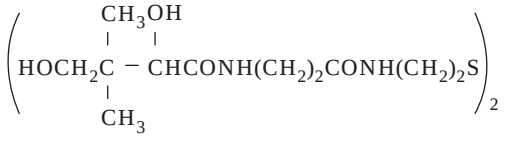
D-Pantothenic acid  $C_9H_{17}NO_5$ MW: 219·23	Unstable, viscous oil. Extremely hydroscopic, easily decomposed by acids, bases, heat. Soluble in water, ethyl acetate, dioxane, glacial acetic acid; moderately soluble in ether, amyl alcohol; insoluble in benzene, chloroform. Solutions are stable between pH 5 and 7. $[\alpha]_D^{25}$ + 37.5°. (From Stiller et al. [1].)
Calcium D-pantothenate  $C_{16}H_{32}CaN_2O_{10}$ MW: 476·53	White needles. Moderately hygroscopic. Soluble in water, glycerol; slightly soluble in alcohol, acetone; insoluble in ether, benzene, chloroform. Decomposed by bases. Solutions are stable between pH 5 and 7. mp 195–196° (dec); $[\alpha]_D^{25}$ + 28.2 °C (c = 5).
Sodium D-pantothenate  $C_9H_{16}NaNO_5$ MW: 241·21	White, hygroscopic crystals. Decomposed by acids and bases. Solutions are stable between pH 5 and 7. mp 122–124° $[\alpha]_D^{25}$ + 27.1° (c = 2). For solubility, see calcium pantothenate.
4′-Phosphopantothenic acid (Ba salt)  $C_9H_{16}NO_8P$ MW: 313·27	Soluble in water; insoluble in ethanol. Unstable to bases. Free acid is unstable. $[\alpha]_D^{24}$ + 9.0° (c = 3.3) (From King and Strong [2] and Okada et al. [3].)
Pantothenoyl-L-cysteine (Ba salt)  $C_{12}H_{22}N_2O_6S$ MW: 322·38	Soluble in water, methanol; moderately soluble in ethanol; insoluble in ether. Unstable to acids and bases. $[\alpha]_D$ −14° (c = 2). (From Ohta et al. [4].)
4′-Phosphopantothenoyl-L-cysteine (Ba salt)  $C_{12}H_{23}N_2O_4PS$ MW: 416·42	Soluble in water; slightly soluble in alcohol. Unstable to acids and bases. Easily oxidized in air. $[\alpha]_D^{20}$ 0° (c = 2). (From Baddiley and Mathias [5] and Nagase [6].)
Pantetheine  $C_{11}H_{22}N_2O_4S$ MW: 278·37	Syrup or glass. Soluble in water; slightly soluble in alcohol; insoluble in ether, benzene, chloroform, ethyl acetate. Unstable to acids and bases. Easily oxidized in air. $[\alpha]_D^{23}$ + 12.2° (c = 3.45). *Lactobacillus bulgaricus* factor. (From Shimizu et al. [7].)

Table 1. Panthothenic Acid and Its Naturally Occurring Derivatives (*continued*)

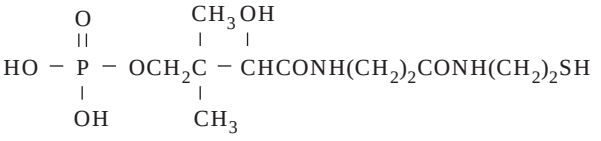
Pantetheine



$C_{22}H_{42}N_4O_5S_2$ MW: 554·72

Disulfide form of pantetheine. Glassy, colorless to light yellow substance. Unstable to acids. $[\alpha]_D^{23}$ + 17.1° (c = 3.2). For solubility and references, see pantetheine.

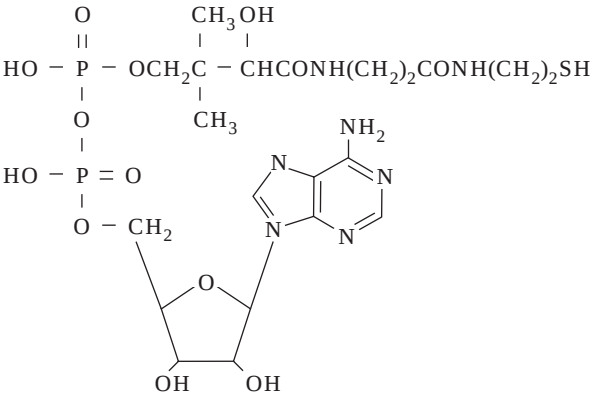
4′-Phosphopantetheine (Ba salt)



$C_{11}H_{23}N_2O_7PS$ MW: 358·35

Soluble in water; slightly soluble in ethanol; insoluble in ether. Unstable to acids and bases. Easily oxidized in air. $[\alpha]_D^{23}$ + 13.3° (c = 2.25). Oxidized form, $[\alpha]_D^{23}$ + 12.2°. *Acetobacter suboxidans* factor. (From Baddiley and Thain [8], Moffatt and Khorana [9], and Nagase [6].)

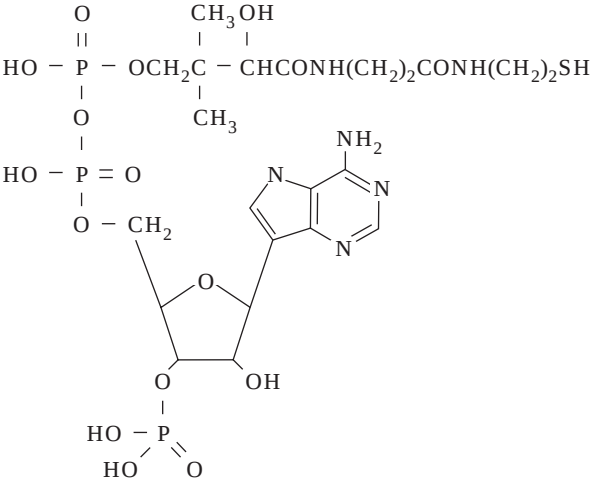
Dephospho-coenzyme A (Li salt)



$C_{21}H_{35}N_7O_{12}P_2S$ MW: 687·56

Soluble in water, methanol; insoluble in acetone. Unstable to acids and bases. For references, see coenzyme A.

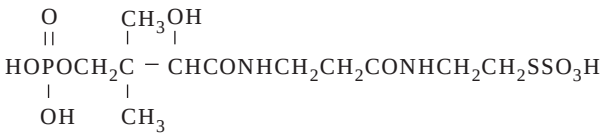
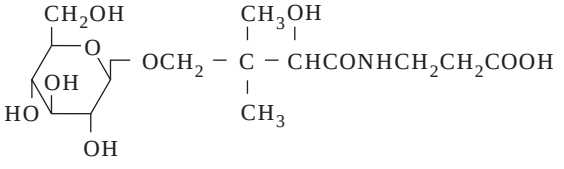
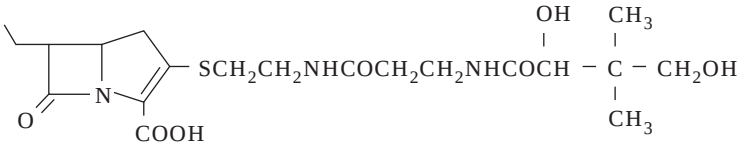
Coenzyme A



$C_{21}H_{36}N_7O_{16}P_3S$ MW: 767·55

Soluble in water; insoluble in ethanol, ether, acetone. Decomposed to pantetheine-2′,4′-cyclic phosphate and 3′,5′-ADP in 1 N NaOH (100°, 2 min). Decomposed to pantetheine-4′-phosphate and adenine in 1 N HCl (100°, 5 min). Easily oxidized in air. (From Moffatt and Khorana [9], Shimizu [10], and Shimizu et al. [11].)

Table 1. Panthothenic Acid and Its Naturally Occurring Derivatives (*continued*)

4′-Phosphopantetheine-*S*-sulfonate (Ca salt)  $C_{11}H_{23}O_{10}N_2S_2P$ MW: 438·41	Soluble in water. $[\alpha]_D^{22}$ + 7.2° (c = 1.95). *Bifidus* factor. (From Yoshioka and Tamura [12].)
4′-*O*-(*β*-glucopyranosyl)-D-pantothenic acid  $C_{44}H_{25}O_{14}N$ MW: 791·68	Soluble in water. $[\alpha]_D^{23}$ −18.2° (c = 1.0). Growth factor of *Leuconostoc*. (From Amachi et al. [13].)
OA-6129A (Na salt)  $C_{20}H_{31}N_3O_7S$ MW: 457·54	Unstable to acids and bases. $[\alpha]_D^{24}$ + 11.6° (c = 1.0). Antibiotic produced by *Streptomyces* sp. OA-6129. (From Yoshioka et al. [14].)

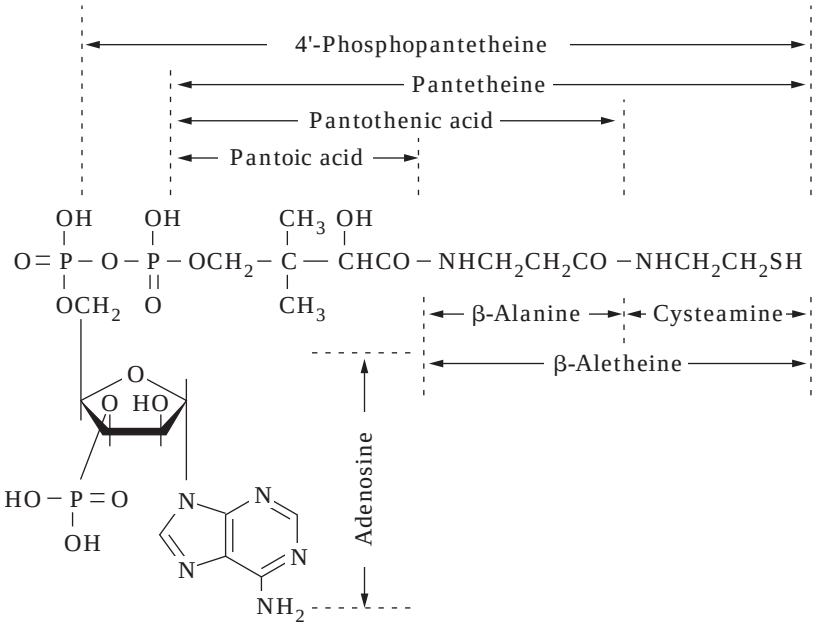


Figure 1. The structure of coenzyme A.

BIOSYNTHESIS

Although pantothenic acid cannot be synthesized by animals, microorganisms and plants are generally able to produce it from the precursors pantoic acid and *β*-alanine through catalysis of the enzyme pantothenate synthetase (EC 6.3.2.1). Animals, plants, and microorganisms can convert pantothenic acid to 4′-phosphopantetheine and coenzyme A, the metabolically active forms of the vitamin. The pathway for the biosynthesis of pantothenic acid and coenzyme A has been elucidated by several workers since the early 1950s and has been reviewed (24–27). The pathway

leading to the vitamin and the coenzymes from common precursors can be summarized as shown in Figures 2 and 3.

β-Alanine

Three routes to β-alanine have been reported. Several microorganisms have been reported to form β-alanine by α-decarboxylation of L-aspartic acid (Fig. 2; reaction 1). Confirmatory evidence for this conversion was provided by Williamson and Brown (28), who purified (to apparent homogeneity), from extracts of *Escherichia coli*, an enzyme that catalyzes the α-decarboxylation of L-aspartic acid to yield β-alanine and CO_2. These investigators also reported that the enzyme is missing in a mutant of *E. coli* that requires either β-alanine or pantothenate as a nutritional factor, but is present in the wild-type strain and in a revertant strain of the mutant. It has also been suggested, on the basis of the observation that mutants of *Salmonella typhimurium* lacking the ability to degrade uracil require *N*-carbamoyl-β-alanine, β-alanine, or pantothenate, as a nutritional factor, that β-alanine is produced by decarboxylation of *N*-carbamoyl-β-alanine formed from uracil (Fig. 2; reactions 2–4) (29). β-Alanine may also be produced by transamination of malonylsemialdehyde produced from propionic acid (Fig. 2; reaction 5 or 6), because enzyme activity catalyzing this conversion has been detected in several microorganisms. However, there have been no further studies concerning this reaction.

Pantoic Acid

The route to pantoic acid from pyruvate (as shown in Fig. 2) has been elucidated mainly in *E. coli* and *Neurospora crassa*. Two enzymes catalyzing the conversion of pyruvate to α-ketoisovalerate (Fig. 2; reactions 7 and 8) in this route are shared by the route for the biosynthesis of the branched chain amino acids. In *E. coli*, two enzyme activities have been detected for the conversion of α-ketoisovalerate to ketopantoic acid (Fig. 2; reaction 9): one is dependent on tetrahydrofolate and the other is not. The physiological significance of tetrahydrofolate-independent activity seemed to be questionable because of its high K_m values for formaldehyde (10 mM) and α-ketoisovalerate (100 mM). Because a mutant lacking tetrahydrofolate-dependent activity requires pantothenate for growth although the same amount of tetrahydrofolate-independent activity is found in the same mutant, concrete evidence is provided to support the theory that the tetrahydrofolate-dependent enzyme is responsible for the ketopantoate needed for the biosynthesis of pantothenate. The tetrahydrofolate-dependent enzyme (i.e., ketopantoate hydroxymethyltransferase, EC 2.1.2.11) has been purified and characterized in some detail. The observation that pantoate, pantothenate, and coenzyme A are all allosteric inhibitors of this enzyme also supports this conclusion (30,31).

The reduction of ketopantoic acid to D-pantoic acid (Fig. 2; reaction 10) is catalyzed by an NADPH-dependent enzyme, ketopantoic acid reductase (EC 1.1.1.169). This enzyme activity has been detected in *Saccharomyces cerevisiae* and *E. coli*. The same reduction is also catalyzed by α-acetohydroxy acid isomeroreductase (EC 1.1.1.1.86), which is the enzyme responsible for the conversion of α-acetolactate to α-ketoisovalerate (Fig. 2; reaction 8) (32). Recently, Shimizu et al. (33) isolated ketopantoic acid reductase in a crystalline form from *Pseudomonas maltophilia* and characterized it in some detail. They also demonstrated that this reductase is the enzyme for D-pantoic acid formation, necessary for the biosynthesis of pantothenic acid, because mutants lacking this enzyme require either D-pantoic acid or pantothenate for growth and the revertants regain this activity.

Coenzyme A

The pathway for the biosynthesis of coenzyme A from pantothenic acid, L-cysteine, and ATP (Fig. 3) was first demonstrated by Brown (34,35) on the basis of his studies with *Proteus morganii* and early observations in the 1950s with pig liver and other organisms. Later, the validity of this pathway was confirmed by Abiko and coworkers, who separated and characterized the enzymes involved in this pathway in rat liver (26).

The first step in this pathway is the phosphorylation of pantothenic acid (Fig. 3; reaction 1) by pantothenate kinase (EC 2.7.1.33). The enzyme has been purified and characterized from rat liver (36) and *Brevibacterium ammoniagenes* (37). Both enzymes undergo allosteric inhibition by coenzyme A, the end product of the pathway. Because such feedback inhibition by coenzyme A is observed only at this step and no other regulation mechanism is known, this inhibition seems to be the most important mechanism for controlling the cellular level of coenzyme A. Pantetheine is also phosphorylated by the same enzyme to yield 4′-phosphopantetheine (Fig. 3; reaction 7), which can be converted to coenzyme A.

The condensation of 4′-phosphopantothenic acid with L-cysteine to yield 4′-phosphopantothenoyl-L-cysteine (Fig. 3; reaction 2) is catalyzed by 4′-phosphopantothenoyl-L-cysteine synthetase (EC 6.3.2.5). The mammalian enzyme requires ATP as energy for the condensation, whereas the bacterial enzyme preferably utilizes CTP (34). 4′-Phosphopantothenoyl-L-cysteine is then decarboxylated to yield 4′-phosphopantetheine (Fig. 3; reaction 3) by 4′-phosphopantothenoyl-L-cysteine decarboxylase (EC 4.1.1.36). The enzyme has been shown to be independent of pyridoxal 5′-phosphate. This step is the only one that does not require ATP among the five steps in the biosynthesis of this coenzyme.

The pyrophosphate linkage formation between 4′-phosphopantetheine and ATP to yield 3′-dephosphocoenzyme A (Fig. 3; reaction 4), and its phosphorylation (Fig. 3; reaction 5) are catalyzed, by dephosphocoenzyme A pyrophosphorylase (EC 2.7.7.3) and dephospho-coenzyme A kinase (EC 2.7.1.24), respectively. In rat liver, both enzymes are present as a complex or a bifunctional enzyme (38). The reversibility of the former reaction may be important for controlling cellular levels of coenzyme A and 4′-phosphopantetheine.

The presence of an alternate route to yield 4′-phosphopantetheine via pantetheine (Fig. 3; reactions 6 and 7) has been suggested in *Acetobacter suboxydans*, *Lactobacillus*

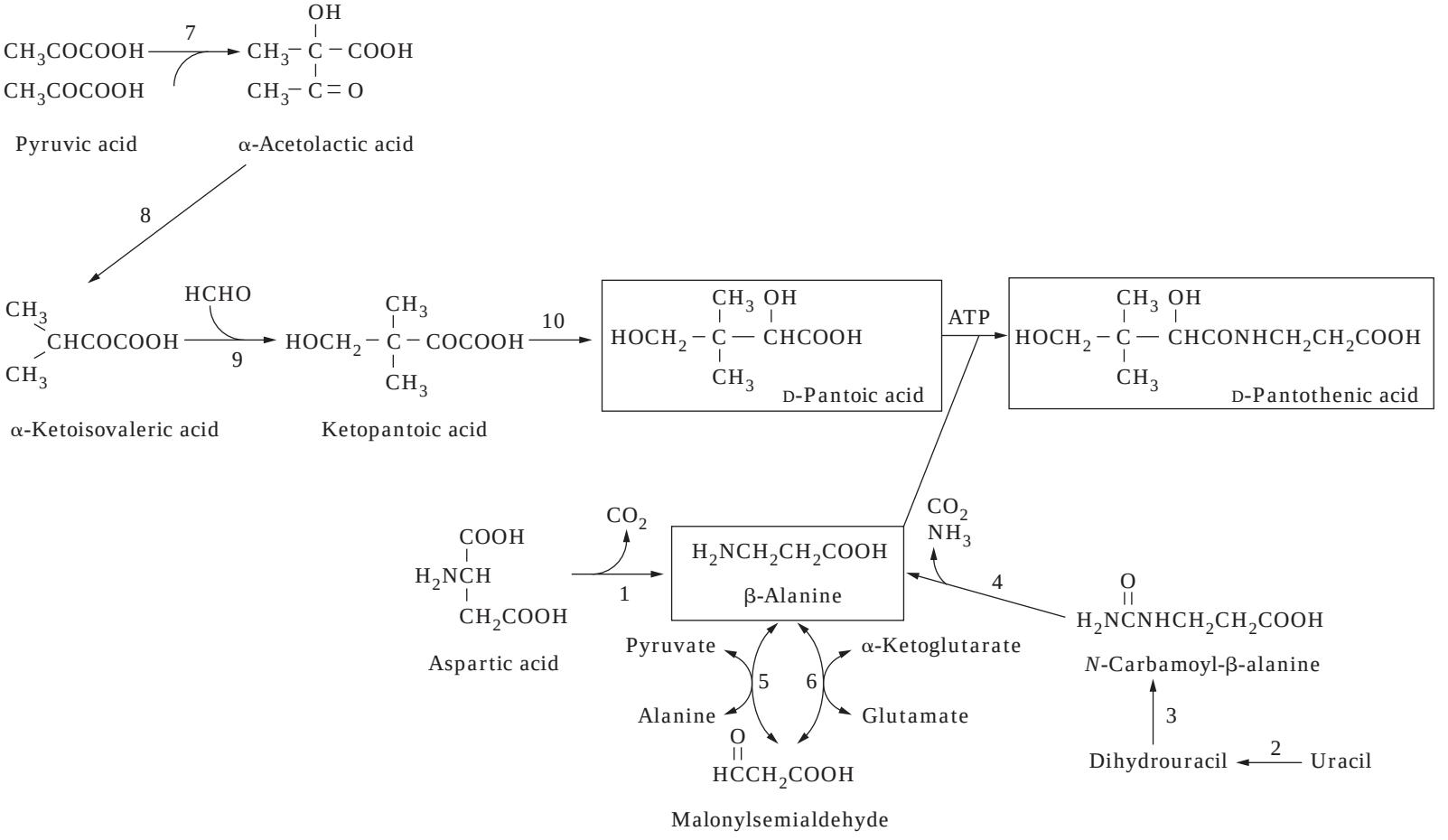


Figure 2. The pathway for the biosynthesis of pantothenic acid.

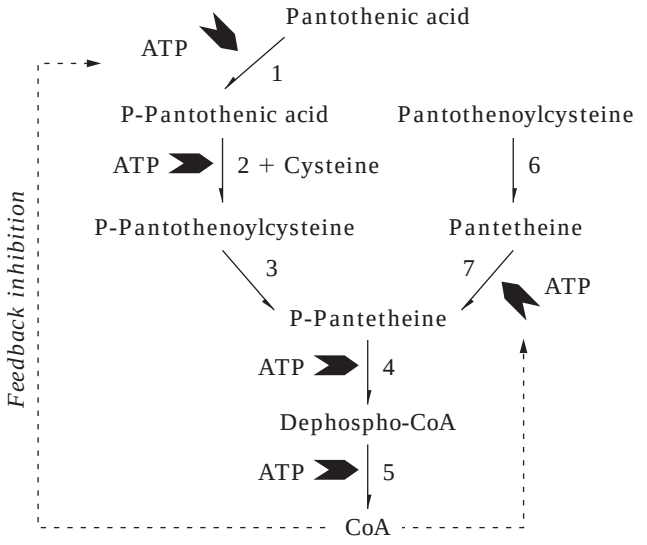


Figure 3. The pathway for the biosynthesis of coenzyme A from pantothenic acid, L-cysteine, and ATP. For chemical structures of each compound, see Table 1. CoA, coenzyme A.

helveticus, and other microorganisms (35), but confirmatory evidence is not available.

Control Mechanisms for the Biosynthesis

The allosteric inhibition of ketopantoic acid hydroxymethyltransferase of *E. coli* by D-pantoic acid, pantothenic acid, or coenzyme A may be involved as a control mechanism in pantothenate biosynthesis (30). On the other hand, such inhibition was not observed in the case of ketopantoic acid reductase of *Pseudomonas maltophilia* (33).

In the pathway to coenzyme A from pantothenic acid, the involvement of the feedback inhibition of pantothenate kinase by coenzyme A and 4′-phosphopantetheine as a control mechanism in the biosynthesis has been demonstrated (36,37,39). Because this inhibition was generally observed regardless of species and the other four steps following this reaction are not significantly inhibited by coenzyme A or 4′-phosphopantetheine, this may be one of the most important mechanisms in the control of cellular levels of coenzyme A. No other mechanism, such as repression, has been observed in either pantothenate or coenzyme A biosynthesis.

Pantetheinase, which specifically degrades pantetheine to pantothenic acid and cysteamine, may also be an important enzyme because coenzyme A can be degraded to pantetheine enzymatically and pantetheine can be reused as a precursor of coenzyme A after phosphorylation by pantothenate kinase. Cellular coenzyme A levels may be influenced by competition between pantetheinase and pantothenate kinase toward their substrate, pantetheine (40).

CHEMICAL AND MICROBIAL PRODUCTION METHODS

Pantothenic Acid

At present, commercial production of pantothenate depends exclusively on chemical synthesis. As outlined in Figure 4, the conventional chemical process involves reactions yielding racemic pantolactone from isobutyraldehyde, formaldehyde, and cyanide, optical resolution of the racemic pantolactone to D-(−)-pantolactone with quinine, quinidine, cinchonidine, brucine, and so on, and condensation of D-(−)-pantolactone with *β*-alanine. This is followed by isolation as the calcium salt and drying to obtain the final product. A problem associated with this chemical process apart from the use of poisonous cyanide is the troublesome resolution of the racemic pantolactone and the racemization of the remaining L-(+)-isomer. Therefore, most of the recent studies in this area have concentrated on development of an efficient method to obtain D-(−)-pantolactone.

Enzymatic resolution of racemic pantolactone can be carried out by specific fungal lactonohydrolases. Shimizu et al. found that many mold strains belonging to the genera *Fusarium*, *Gibberella*, and *Cylindrocarpon* stereospecifically hydrolyze D-(−)-pantolactone to D-(−)-pantoic acid (41,42). If racemic pantolactone is used as a substrate for the hydrolysis reaction by the microbial lactonohydrolase, only the D-(−)-pantolactone might be converted to D-(−)-pantoate and the L-(+)-enantiomer might remain intact. Consequently, the racemic mixture could resolved into D-(−)-pantoate and L-(+)-pantolactone as shown in Figure 5 (42,43). After the removal of L-(+)-pantolactone from the reaction mixture by solvent extraction, etc., remaining D-(−)-pantoate could be easily converted to D-(−)-pantolactone by heating under acidic conditions. The reverse reaction, that is, lactonization of D-(−)-pantoate, might also be possible for the resolution. In this case, D-(−)-pantoate in a racemic mixture of pantoate is specifically lactonized into D-(−)-pantolactone (see Fig. 5) (44). When *Fusarium oxysporum* mycelia were incubated in 700 g/L aqueous solution of racemic pantolactone for 24 h at 30 °C with automatic pH control (pH 6.8–7.2), about 90% of the D-(−)-isomer was hydrolyzed. The resultant D-(−)-pantoic acid in the reaction mixture showed a high optical purity (96% ee), and the coexisting L-(+)-isomer remained without any modification.

Practical hydrolysis of the D-(−)-isomer in a racemic mixture is carried out using immobilized mycelia of *F. oxysporum* as the catalyst. Stable catalyst with high hydrolytic activity can be prepared by entrapping the fungal mycelia in calcium alginate gels. When the immobilized mycelia were incubated in a reaction mixture containing 350 g/L racemic pantolactone for 21 h at 30 °C under automatic pH control (pH 6.8–7.2), 90–95% of the D-(−)-isomer was hydrolyzed (optical purity, 90–97% ee). After reactions repeated 180 times (i.e., for 180 days), the immobilized mycelia retained more than 90% of their initial activity (Fig. 6).

The overall process for the present enzymatic resolution is compared with the conventional chemical process in Figure 7. The enzymatic process allows skipping several tedious steps that are necessary in chemical resolution and is highly advantageous for practical purposes.

Coenzyme A

The production methods for coenzyme A roughly fall into chemical and microbial categories. The chemical methods,

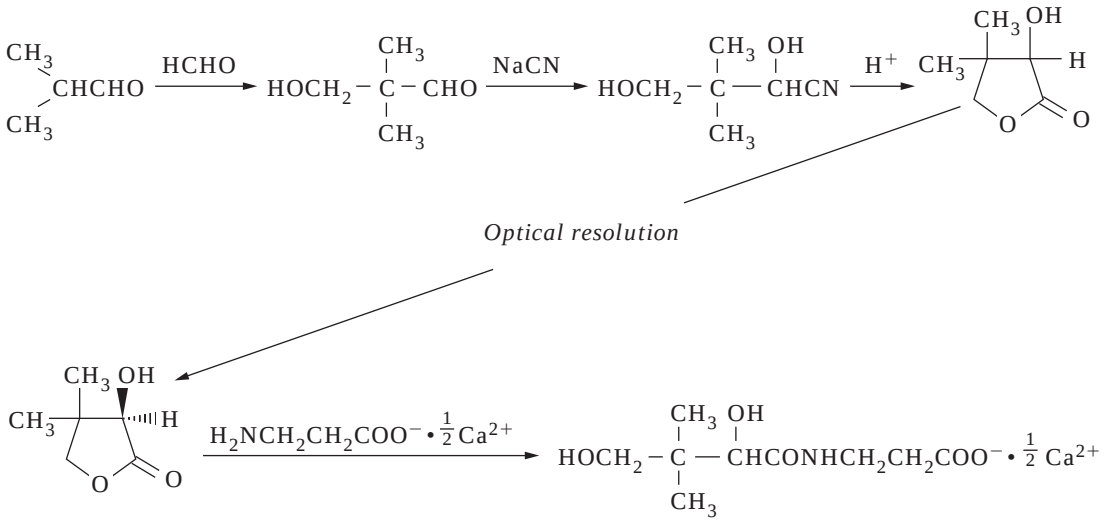


Figure 4. Outline of the chemical synthesis of D-pantothenic acid.

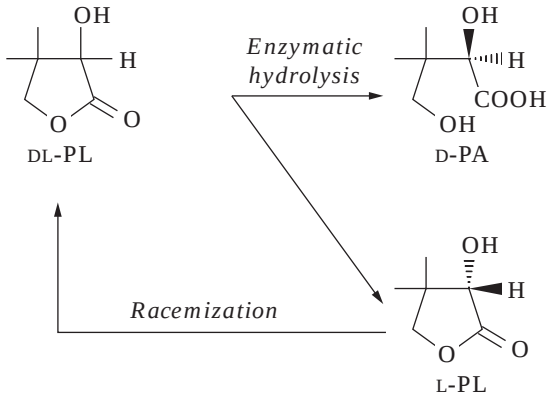


Figure 5. Schematic representation of the enzymatic resolution of racemic pantolactone. PL, pantolactone; PA, pantoate.

which have been reviewed by Shimizu (10) and Mautner (45), are too complex to be practical. Therefore, commercial production is carried out by microbiological methods. Extraction of coenzyme A from yeast cells has been performed since the early 1950s. Cells of baker's or brewer's yeasts, which are relatively rich in coenzyme A, were usually used as the source. Later, an efficient enzymatic method using *Brevibacterium ammoniagenes* cells as the catalyst was developed. These microbial methods have been reviewed by Shimizu and Yamada (46).

A successful enzymatic method using the biosynthetic route of coenzyme A from pantothenic acid, L-cysteine, and ATP was first reported by Ogata et al. (47), who found that *B. ammoniagenes* has all five enzymes necessary for the biosynthesis of coenzyme A in high activities. These three substrates, when added to a reaction mixture containing the bacterial cells, were converted to coenzyme A with a satisfactory yield (2–3 g/L). Ogata et al. also found that the same organism can accumulate coenzyme A directly in the culture medium on addition of pantothenic acid, L-cysteine, and AMP, adenosine, or adenine in the presence of a surfactant, cetylpyridinium chloride, and high levels of glucose (usually 10%), K_2HPO_4, and $MgSO_4 \cdot 7H_2O$. Under optimal conditions, the amount obtained was 5.5 g/L. Most coenzyme A in the medium was present in the disulfide form because of the vigorous shaking during the reaction. After treatment of the culture filtrate with Duolite S-30, charcoal, and Dowex 1 (Cl^-), and reduction of the disulfide, the very pure thiol form was obtained in high yield. The mechanism of this coenzyme A production has been suggested to be that shown in Figure 8 (for details, see Refs. 11 and 48).

To improve product yield further, the mechanism for regulation of biosynthesis was investigated. As described in the preceding section, the biosynthesis of coenzyme A in *B. ammoniagenes* is controlled mainly by the feedback inhibition of pantothenate kinase by coenzyme A. This was the main problem in the practical production, because the overproduced coenzyme A itself stopped the biosynthesis. Two methods to abolish this feedback inhibition have been reported.

A synthetic scheme was investigated in which the reaction is initiated by the condensation of 4′-phosphopantothenic acid and L-cysteine or the transadenosylylation of 4′-phosphopantetheine, because these routes do not involve phosphorylation of pantothenic acid or pantetheine by pantothenate kinase. Replacement of the enzymatic phosphorylation of pantothenate or pantetheine with chemical phosphorylation followed by the enzymatic reaction increased the yield of coenzyme A 10- to 20-fold. Yields from 4′-phosphopantothenic acid and 4′-phosphopantetheine were 33 g/L (molar yield based on ATP; 64.5%) and 115 g/L (100%), respectively (49). This method is applicable to coenzyme A production under ATP-generating conditions (see Fig. 8). 4′-Phosphopantothenic acid (25 g/L), L-cysteine (15 g/L), and AMP (33 g/L), when added to the culture broth of *B. ammoniagenes*, were converted to coenzyme A with a yield of 23 g/L (50).

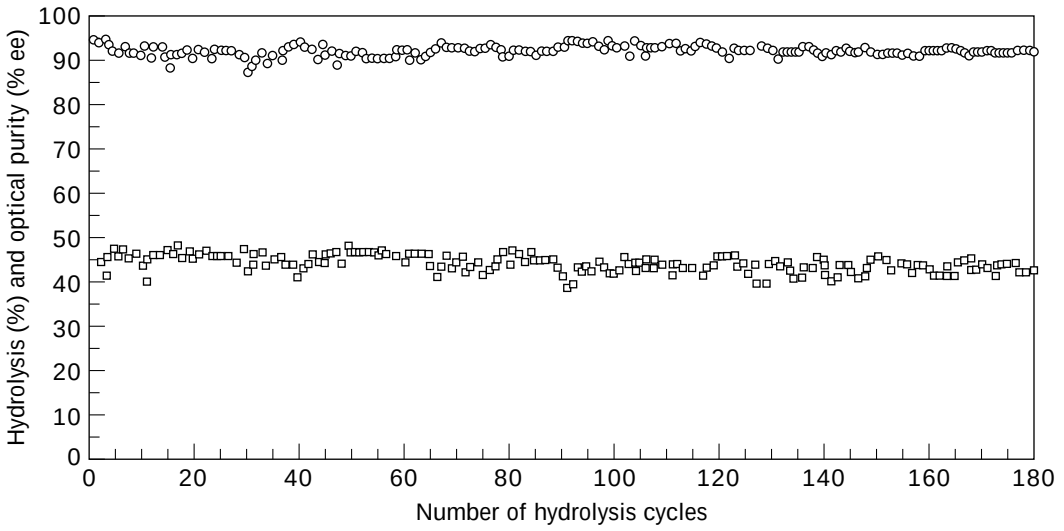


Figure 6. Stereospecific hydrolysis of pantolactone by *Fusarium oxysporum* mycelia entrapped in calcium alginate gels: 280 L immobilized mycelia (containing 15.2 kg wet cells) was incubated with 350 L aqueous racemic pantolactone solution (350 g/L) at 30 °C for 21 h. The pH of the mixture was automatically controlled at 6.8–7.2 with 15 M NH_4OH. All the immobilized mycelia filtrated from the reaction mixture were used for the subsequent reactions. The hydrolysis reactions were carried out 180 times. ○, optical purity for D-(−)-pantoate; □, hydrolysis rate.

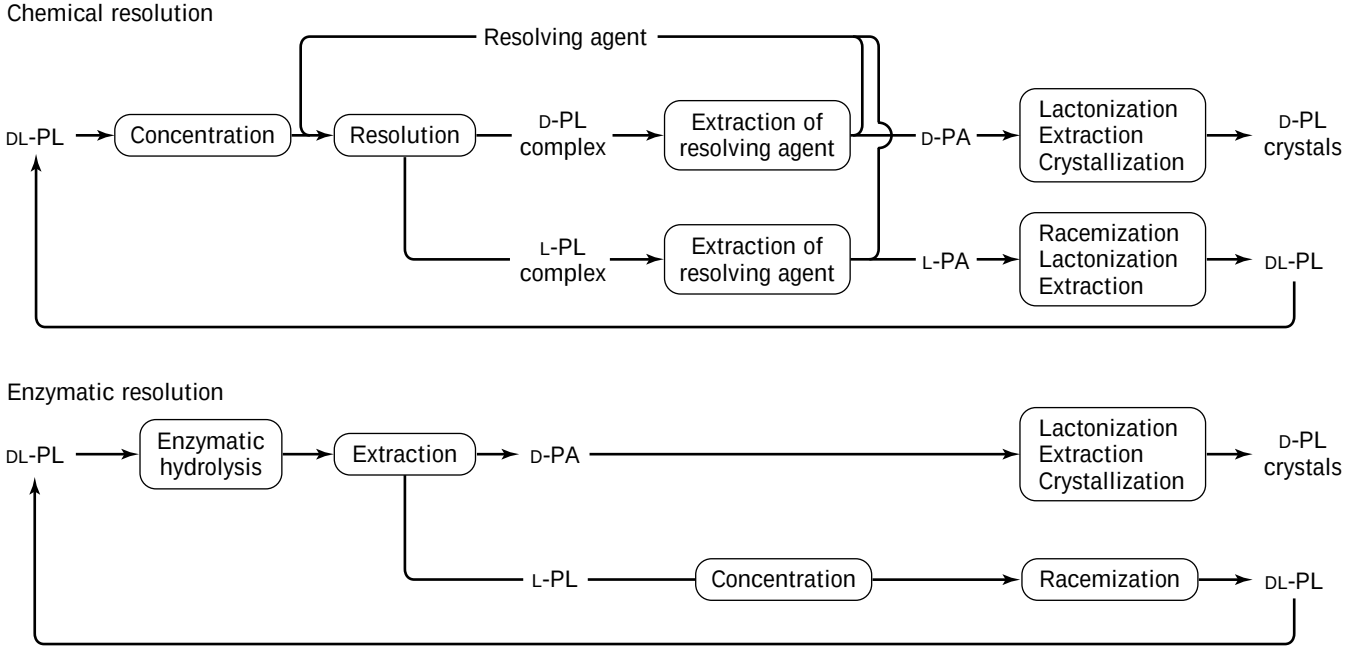


Figure 7. Comparison of enzymatic and conventional chemical resolution processes for racemic pantolactone.

Another way to improve the yield is to use mutants derepressed for the feedback inhibition or those showing elevated pantothenate kinase activity. A mutant of *B. ammoniagenes* that is resistant to oxypantetheine (the corresponding oxygen analog of pantetheine) was found to have an elevated activity of pantothenate kinase. Under ATP-generating conditions, the yields of coenzyme A from pantothenic acid (3.6 g/L), L-cysteine (1.8 g/L), and AMP (6 g/L) or from pantetheine (5 g/L) and AMP (6 g/L) were 9.3 or 11.5 g/L, respectively. These values were about threefold higher than those obtained with the parent strain, and 70–100% of the added AMP was converted to coenzyme A (51) (Fig. 9).

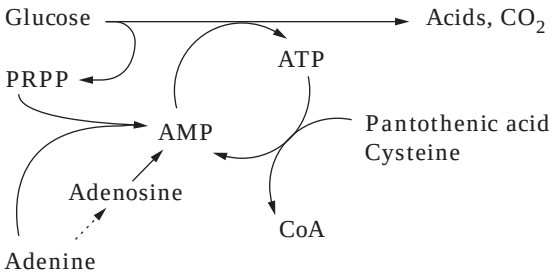


Figure 8. Reaction sequences of coenzyme A production with *Brevibacterium ammoniagenes* under ATP-generating conditions. PRPP, 5-phosphoribosyl-1-pyrophosphate; CoA, coenzyme A.

Continuous production of coenzyme A through five enzymatic steps using gel-entrapped cells of *B. ammoniagenes* has also been reported (48,52).

4′-Phosphopantetheine and Other Intermediates in Coenzyme A Biosynthesis

4′-Phosphopantetheine together with other intermediates in coenzyme A biosynthesis can be effectively synthesized by using *B. ammoniagenes* cells as the catalyst and by modifying the reaction conditions (11) as follows: 4′-phosphopantothenic acid on omission of L-cysteine from the reaction mixture, 4′-phosphopantetheine on addition of CTP or CTP and GTP in place of ATP (because pantothenic acid is phosphorylated in the presence of CTP or GTP as well as ATP and the condensation of 4′-phospho-

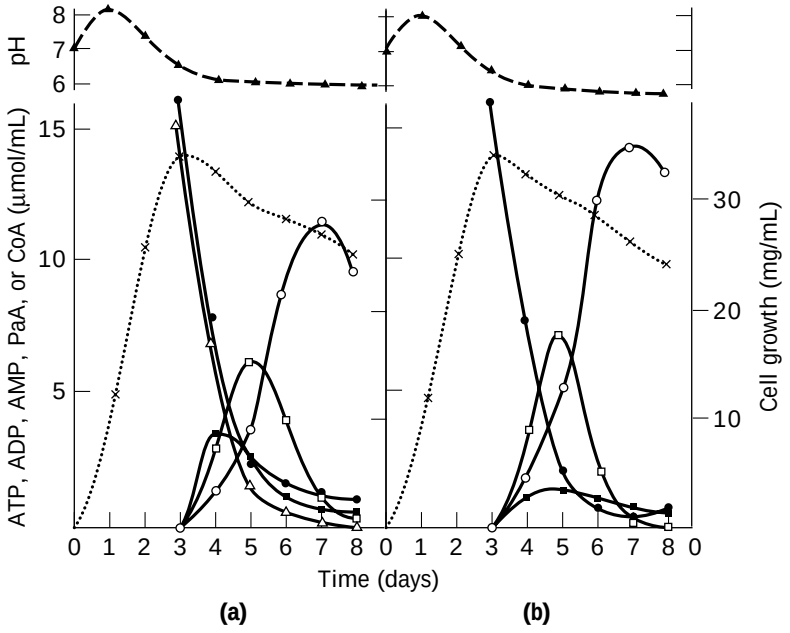


Figure 9. Time course of coenzyme A production by an oxypantetheine-resistant mutant of *Brevibacterium ammoniagenes* under ATP-generating conditions. **(a)** Production from pantothenic acid, L-cysteine, and AMP. **(b)** Production from pantetheine and AMP. ○, coenzyme A (CoA); ●, pantothenic acid (PaA); □, ATP; ■, ADP; △, AMP; ▲, pH; ×, cell growth. For details, see Shimizu et al. (51).

Table 2. Production of the Intermediates in Coenzyme A Biosynthesis by *Brevibacterium ammoniagenes*

Product	Substrates	Nucleotides	Reactions involved[a]	Dried cells	Productivity enzyme source (mg/mL)	
					Culture broth[b]	Immobilized cells
4′-Phosphopantothenic acid	Pantothenic acid	ATP	1	3–4		1.5–2.5
4′-Phosphopantothenic acid	Pantothenic acid	AMP	1		4–5	
4′-Phosphopantetheine	4′-Phosphopantothenic acid and L-cysteine	CTP	2,3	3–4		1.8
4′-Phosphopantetheine	Pantothenic acid and L-cysteine	ITP and CTP	1–3	2–3		0.3
4′-Phosphopantetheine	Pantothenic acid and L-cysteine	GMP and CMP	1–3		3–4	
4′-Phosphopantetheine	Pantetheine	ITP	7	2–3		0.9
4′-Phosphopantetheine	Pantetheine	GMP	7		4–5	
3′-Dephosphocoenzyme A	Pantothenic acid and L-cysteine	ATP	1–5	1–2		

Note: For details, see Shimizu et al. (11,48).

[a]Numbers correspond to those given in Figure 3.

[b]In this case, the nucleoside monophosphates added are presumed to be converted to the corresponding nucleoside triphosphates and used for the reactions in a manner similar to that shown in Figure 8.

pantetheine with nucleoside triphosphate is specific for ATP), and 3′-dephosphocoenzyme A on treatment of the reaction mixture containing coenzyme A with 3′-nucleotidase. The amounts of these intermediates obtained by this method are summarized in Table 2.

ASSAY METHODS

Test microorganisms normally used for the microbiological assay of pantothenic acid are *Lactobacillus plantarum* (ATCC 8014; *L. arabinosus* 17-5), *L. casei* (ATCC 7469), and *Saccharomyces uvarum* (ATCC 9080; *S. carlsbergensis*). *Lactobacillus plantarum* is suitable for determining unconjugated pantothenate in samples. It should be noted that pantetheine, when simultaneously present in a molar ratio to pantothenate of more than 0.5, gives positive errors in the determination. *Saccharomyces uvarum* also shows almost specific growth response to free pantothenate, but *β*-alanine stimulates its growth. Hence, an assay procedure employing this organism is also the one chosen for determining the pantothenic acid that occurs in natural products together with other pantothenate forms. *Lactobacillus casei* responds not only to pantothenate but also several conjugated forms of pantothenate. *Lactobacillus helveticus* (ATCC 12046) and *L. bulgaricus* B1 are recommended for determination of pantetheine (or pantethine) because both these organisms require more than 100 times as much pantothenic acid as pantethine to give the same response. Treatment of samples and details of assay procedures have been described (53).

A sensitive enzymatic assay method using pantothenase has been reported (54), but the enzyme is not commercially available. Chemical and physical methods have also been reported. These are often used in determining pantothenic acid in pharmaceutical products but are not suitable for determination of natural samples because of their low sensitivity.

USE AND ECONOMIC ASPECTS

The current world capacity of calcium pantothenate production and its demand are presumed to be about 4,000 and 3,600–4,000 tons per year, respectively. It is mainly used as an additive to animal feed (about 3,000 tons per year) and as a pharmaceutical product (about 600 tons per year). Pantothenyl alcohol is used as a source of pantothenate activity for pharmaceutical vitamin products. Pantothenyl alcohol itself has no pantothenate activity; in fact, it is a competitive growth inhibitor of several pantothenate-requiring lactic acid bacteria. However, it has been demonstrated to be quantitatively converted to pantothenic acid in the animal body and to be equivalent to pantothenic acid in man.

Pantethine, the disulfide of pantetheine, and coenzyme A are also used as pharmaceutical products in several countries. They have been suggested to be effective in reducing cholesterol level, curing fatty liver, and treating related diseases.

Some sulfonate derivatives of pantetheine or coenzyme A (*Bifidus* factors), such as 4′-phosphopantetheine-*S*-sulfonate, that were originally isolated from carrot roots have been shown to be growth factors of *Bifidobacterium* (12). Addition of the *bifidus* factors to dried milk for infants has been suggested to be useful in improving the quality of the milk. A carbapenem antibiotic, OA-6129A (see Table 1) produced by *Streptomyces* sp. OA-6129, may be an interesting example suggesting a new use of the vitamin as a building block for its synthesis (14).

BIBLIOGRAPHY

1. E.T. Stiller, S.T. Harris, J. Finkelstein, J.C. Keresztesy, and K. Folkers, *J. Am. Chem. Soc.* **62**, 1785–1790 (1940).
2. T.E. King and F.M. Strong, *J. Biol. Chem.* **191**, 515–521 (1951).
3. S. Okada, O. Nagase, and M. Shimizu, *Chem. Pharm. Bull.* **15**, 713–715 (1967).
4. G. Ohta, O. Nagase, Y. Hosokawa, H. Tagawa, and M. Shimizu, *Chem. Pharm. Bull.* **15**, 644–647 (1967).
5. J. Baddiley and A.P. Mathias, *J. Chem. Soc.*, 2803–2812 (1954).
6. O. Nagase, *Chem. Pharm. Bull.* **15**, 648–654 (1967).
7. M. Shimizu, G. Ohta, O. Nagase, S. Okada, and Y. Hosokawa, *Chem. Pharm. Bull.* **13**, 180–188 (1965).
8. J. Baddiley and E.M. Thain, *J. Chem. Soc.*, 1610–1615 (1953).
9. J.G. Moffatt and H.G. Khorana, *J. Am. Chem. Soc.* **83**, 663–675 (1961).
10. M. Shimizu, *Methods Enzymol.* **18A**, 322–328 (1970).
11. S. Shimizu, Y. Tani, and K. Ogata, *Methods Enzymol.* **62**, 236–245 (1979).
12. M. Yoshioka and Z. Tamura, *Chem. Pharm. Bull.* **19**, 178–185 (1971).
13. T. Amachi, S. Iwamoto, and H. Yoshizumi, *Agric. Biol. Chem.* **35**, 1222–1230 (1971).
14. T. Yoshioka, I. Kojima, K. Isshiki, A. Watanabe, Y. Shimauchi, M. Okabe, and Y. Fukagawa, *J. Antibiot.* (*Tokyo*) **36**, 1473–1482 (1983).
15. R.J. Williams, *Adv. Enzymol.* **3**, 253–287 (1943).
16. F. Lipmann, *Fed. Proc.* **12**, 673–715 (1953).
17. A.F. Wagner and K. Folkers, in *Vitamins and Coenzymes*, Interscience, New York, 1964, pp. 93–137.
18. P.R. Vagelos, in P.D. Boyer ed., *The Enzymes*, 3rd ed., vol. 8, Academic Press, New York, 1973, pp. 155–199.
19. E.J. Vandamme, in A. Wiseman ed., *Topics in Enzyme and Fermentation Biotechnology*, vol. 5, Horwood-Wiley, New York, 1981, pp. 185–201.
20. H. Kleinkauf and H. von Döhren, *Trends Biochem. Sci.* **8**, 281–283 (1983).
21. R.K. Hill and T.H. Chan, *Biochem. Biophys. Res. Commun.* **38**, 181–183 (1970).
22. H. Fritz and W. Löwe, *Angew. Chem.* **74**, 751–753 (1962).
23. J. Baddiley, *Adv. Enzymol.* **16**, 1–22 (1955).
24. G.M. Brown and J.J. Reynolds, *Annu. Rev. Biochem.* **32**, 419–462 (1963).
25. G.W. Plaut, C.M. Smith, and W.L. Alworth, *Annu. Rev. Biochem.* **43**, 899–922 (1974).
26. Y. Abiko, in D.M. Greenberg ed., *Metabolic Pathways*, 3rd ed., vol. 7, *Metabolism of Sulfur Compounds*, Academic Press, New York, 1975, pp. 1–25.

27. G.M. Brown and J.M. Williamson, *Adv. Enzymol.* **53**, 345–381 (1982).
28. J.M. Williamson and G.M. Brown, *J. Biol. Chem.* **254**, 8074–8082 (1979).
29. T.P. West, T.W. Traut, M.S. Shanley, and G.A. O'Donovan, *J. Gen. Microbiol.* **131**, 1083–1090 (1985).
30. S.G. Powers and E.E. Snell, *J. Biol. Chem.* **251**, 3786–3793 (1976).
31. J.H. Teller, S.G. Powers, and E.E. Snell, *J. Biol. Chem.* **251**, 3780–3785 (1976).
32. D.A. Primerano and R.O. Burns, *J. Bacteriol.* **153**, 259–269 (1983).
33. S. Shimizu, M. Kataoka, M.C.M. Chung, and H. Yamada, *J. Biol. Chem.* **263**, 12077–12084 (1988).
34. G.M. Brown, *J. Biol. Chem.* **234**, 370–378 (1959).
35. G.M. Brown, *J. Biol. Chem.* **234**, 379–382 (1959).
36. Y. Abiko, S. Ashida, and M. Shimizu, *Biochim. Biophys. Acta* **268**, 364–372 (1972).
37. S. Shimizu, K. Kubo, Y. Tani, and K. Ogata, *Agric. Biol. Chem.* **37**, 2863–2870 (1973).
38. T. Suzuki, Y. Abiko, and M. Shimizu, *J. Biochem.* (*Tokyo*) **62**, 642–649 (1967).
39. D.S. Vallari, S. Jackowski, and C.O. Rock, *J. Biol. Chem.* **262**, 2468–2471 (1987).
40. C.T. Wittwer, D. Burkhard, K. Ririe, R. Rasmussent, J. Brown, B.W. Wyse, and R.G. Hansen, *J. Biol. Chem.* **258**, 9733–9738 (1983).
41. S. Shimizu, M. Kataoka, K. Shimizu, M. Hirakata, K. Sakamoto, and H. Yamada, *Eur. J. Biochem.* **209**, 383–390 (1992).
42. M. Kataoka, K. Shimizu, K. Sakamoto, H. Yamada, and S. Shimizu, *Appl. Microbiol. Biotechnol.* **43**, 974–977 (1995).
43. M. Kataoka, K. Shimizu, K. Sakamoto, H. Yamada, and S. Shimizu, *Appl. Microbiol. Biotechnol.* **44**, 333–338 (1995).
44. M. Kataoka, K. Shimizu, K. Sakamoto, H. Yamada, and S. Shimizu, *Enzyme Microb. Technol.* **19**, 307–310 (1996).
45. H.G. Mautner, *Methods Enzymol.* **18A**, 338–350 (1970).
46. S. Shimizu and H. Yamada, in H.-J. Rehm and G. Reed eds., *Biotechnology*, vol. 4, VCH, Weinheim, Germany, 1986, pp. 159–184.
47. K. Ogata, S. Shimizu, and Y. Tani, *Agric. Biol. Chem.* **34**, 1757–1759 (1970).
48. S. Shimizu, Y. Tani, and H. Yamada, in K. Venkatsubramanian ed., *Immobilized Microbial Cells*, American Chemical Soc., Washington, D.C., 1979, pp. 87–100.
49. S. Shimizu, R. Komaki, Y. Tani, and H. Yamada, *FEBS Lett.* **151**, 303–306 (1983).
50. S. Shimizu and H. Yamada, in *Third European Congress on Biotechnology*, vol. I, Verlag Chemie, Weinheim, Germany, 1984, pp. 401–405.
51. S. Shimizu, A. Esumi, R. Komaki, and H. Yamada, *Appl. Environ. Microbiol.* **48**, 1118–1122 (1984).
52. H. Yamada, S. Shimizu, and Y. Tani, in H.H. Weetall and G. Royer eds., *Enzyme Engineering*, vol. 5, Plenum, New York, 1980, pp. 405–411.
53. O.D. Bird and R.Q. Thompson, in P. Gyorgy and W.N. Pearson eds., *The Vitamins, Chemistry, Physiology, Pathology, Methods*, 2nd ed., vol. 7, Academic Press, New York, 1967, pp. 209–241.
54. R.K. Airas, *Methods Enzymol.* **122**, 33–35 (1986).

See also METABOLITES, PRIMARY AND SECONDARY.

PENICILLIN. See SECONDARY METABOLITES, ANTIBIOTICS.

PEPTIDE

HANS VON DÖHREN
HORST KLEINKAUF
Technical University Berlin
Berlin, Germany

KEY WORDS

Amino acids
Antibiotics
Biosynthesis
Hormones
Incorporation
Nonribosomal
Penicillin
Peptides
Processing
Synthetases

OUTLINE

INTRODUCTION

Peptides may be structurally defined as amino acid–derived compounds containing at least one amide (peptide) bond. Conventionally, a size limit is imposed for polypeptides of defined sequence above the range of 50 to 100 amino acids, which are termed proteins. Peptide classifications so far use either structural features, biological properties, or biosynthetic considerations (1,2). The major group of peptide antibiotics comprises structurally diverse types such as linear, cyclic and multicyclic peptides, peptidolactones, depsipeptides, and peptides modified with diverse nonpeptides moieties including acyl and aryl groups, polyketide or terpenoid-derived building blocks, or carbohydrate decorations (3). In addition, the biological properties of many compounds termed antibiotics have been shown to be diverse, thus extending the classical definition, which implies the inhibition or elimination of a certain type of organism (Table 1) (4).

From the point of view of peptide production, the bio-

Table 1. Selected Peptides of Current Interest

Compound	Structural type	Source	Properties/applications
Aspartame	Dipeptide	Synthetic	Sweetener
Glutathione	Tripeptide	Various	Detoxification and antioxidant compound
TRH	Modified tripeptide	*Homo sapiens* (hypothalamus)	Thyrotropin-releasing hormone, also stimulates prolactin release
Penicillin	Modified tripeptide	*Penicillium chrysogenum*	Antibacterial drug
Cephalosporin	Modified tripeptide	*Acremonium chrysogenum*	Antibacterial drug
Ergotpeptides	Modified tripeptide	*Claviceps* sp.	Uterine relaxant, treatment of parkinsonism, acromegaly, breast cancer
HC-toxin	Cyclotetrapeptide	*Helmintosporium carbonum*	Plant pathogenic (corn)
Enkephalins	Pentapeptides	*Homo sapiens* (brain)	Opioid activity
Actinomycin	Chromophore-attached pentapeptidolactone	*Streptomyces antibioticus*	Anticancer drug
Destruxin	Hexapeptidolactone	*Metarrhizium anisopliae*	Insecticidal peptide
Enniatin	Cyclohexadepsipeptide	*Fusarium* sp.	Antibacterial, antiviral, antihelminthic
Ferrichrome	Cyclohexapeptide	Various fungi	Siderophore
L-365,209	Cyclohexapeptide	Synthetic	Oxytocin antagonist, uterine relaxant
Vancomycins, ristocetins	Modified heptapeptides	Various Actinomycetes	Antibacterial drugs, antiviral compounds
Microcystins	Cycloheptapeptide	*Microcystis* sp. cyanobacteria	Hepatotoxic water contaminant
Lophyrotomin	Aryloctapeptide	*Lophyrotoma interrupta* (sawfly and other insects)	Hepatotoxin
Amanitins	Modified cyclooctapeptides	*Amanita phalloides*	Toxin
Surfactin	Octapeptidolactone	*Bacillus subtilis*	Surfactant, antifungal, antiviral, and antimycoplasma drug
Sandostatin	Modified octapeptide (disulfide)	Synthetic	Somatostatin agonist, treatment of acromegaly and carcinoid syndrome
PF-1022	Cyclooctadepsipeptide	*Mycelia sterilia*	Antihelminthic
Oxytocin	Nonapeptide (disulfide)	*Homo sapiens*	Smooth muscle contraction, principle birth hormone
Vasopressin	Nonapeptide amide(3 disulfides)	*Homo sapiens*	Control of diabetes insipidus
Desmopressin	Nonapeptide amide (disulfide)	Synthetic	Vasopressin analog, treatment of diabetes insipidus
Aureobasidin	Nonapeptidolactone	*Aureobasidium pullulans*	Antifungal drug
Gonadotropin-releasing hormone	Modified decapeptide Amide	*Homo sapiens* (hypothalamus)	Causes release of luteinizing hormone and follicle-stimulating hormone
Buserelin	Modified decapeptide	Synthetic	GnRH analog, treatment of prostate cancer
Tyrocidine	Cyclodecapeptide	*Bacillus brevis*	Topical antibacterial
Gramicidin S	Cyclodecapeptide	*Bacillus brevis*	Topical antibacterial, antihelminthic
Polymyxins	Heptacyclodecapeptide	*Bacillus polymyxa*	Antibacterial drug
Cyclosporin	Cycloundecapeptide	*Tolypocladium niveum*	Immunosuppressant, antifungal
Bacitracin	Modified heptacyclododecapeptide	*Bacillus licheniformis*	Antibacterial drug, proteinase inhibitor
Somatostatin	14-peptide (2 disulfides)	*Homo sapiens* (hypothalamus, gastrointestinal, tract, etc.)	Various gastrointestinal functions (antacid)
Gramicidin (linear)	Pentadecapeptide	*Bacillus brevis*	Antibacterial, ion-channel-forming
Tachyplesins	17- to 18-peptide	*Tachypleus* sp. (horseshoe crab) and related species	Antimicrobial compounds
Alamethicins	Modified nonadecapeptides	*Trichoderma viride*	Antibiotic, ion-channel-forming
Ranalexin	Modified 20-peptide (1 disulfide)	*Rana catesbeiana* (bullfrog, skin)	Antibacterial
Aborycin	Modified 21-peptide (3 disulfides)	*Streptomyces griseoflavus*	Antiviral (HIV-1)
Magainins	21- to 27-peptide	*Xenopus laevis* (skin)	Antimicrobial drugs
Melittin	26-peptide amide	Bee (venom)	Hemolytic compound
Ac-AMP	29-peptide (3 disulfides)	*Amaranthus canditus* (seeds)	Antimicrobial compound

Table 1. Selected Peptides of Current Interest (*continued*)

Compound	Structural type	Source	Properties/applications
Defensins	29- to 32-peptide	Mammalian (various sources)	Antimicrobial compounds
Human neutrophil peptide (HNP-1)	30-peptide	*Homo sapiens* (neutrophils)	Antibacterial, antifungal, cytotoxic
Cecropins	31-peptide	*Sus scrofa* (pig, small intestine)	Antibacterial (Gram negative)
Calcitonin	33-peptide amide	*Homo sapiens*	Treatment of hypercalcemia, osteoporosis
Subtilin	Modified 32-peptide	*Bacillus subtilis*	Antibacterial food preservative
Subtilosin	Modified 33-peptide	*Bacillus subtilis*	Antibacterial
Nisin	Modified 34-peptide	*Streptococcus lactis*	Antimicrobial food preservative
Mj-AMP	36-peptide	*Mirabilis jalapa* (seeds)	Antimicrobial compound
Cecropins (insect)	37- to 41-peptide amides	Various insects	Antibacterial compounds
Corticotropin (ACTH)	39-peptide	*Homo sapiens* (anterior pituitary gland)	Antibacterial compounds Control of glucocorticoid release
Thionins	34- to 47-peptide (3 or 4 disulfides)	Various plant sources	Antibacterial, antifungal
Defensins (plant)	45- to 51-peptide (4 disulfides)	Various plant sources	Antimicrobial compounds
AgaIVC	48-peptide	*Agelopsis aperta* (spider)	Venom, contains a D-Ser in position 48
Insulin	51 amino acids (2 chains, 3 disulfides)	Mammalian	Treatment of type I diabetes
Trefoil peptides	51–56 amino acids	Mammalian	Tissue repair and cell proliferation
Hirudin	65-peptide	Leeches (buccal gland)	Protease inhibitor, anticoagulant, antithrombotic agent

Note: See Ref. 11 for more extensive coverage.

synthetic origin is of primary importance. Although the ribosomal peptide biosynthetic system is an essential constituent of all self-replicating organisms, nonribosomal peptide forming systems, with few exceptions (bacterial cell wall peptides, glutathione), are not essential and are themselves gene products. So ribosomally derived peptides are directly encoded by gene transcripts, whereas nonribosmally derived compounds are the products of biocatalysts or multienzyme systems. To propagate each type of peptide in a cellular environment requires exceedingly different approaches. A decapeptide gene may be defined by a 30-bp reading frame and additional features permitting efficient transcriptional and translational processing. A biocatalyst forming a decapeptide requires approximately 40,000 bp of genetic information, and thus a considerably more complex genetic background (5,6).

Genome sequencing efforts have revealed the absence of such complex biosynthetic systems in most organisms investigated so far. These systems are especially prominent in microbial sources such as Actinomycetes, Bacilli, cyanobacteria, and filamentous fungi and are already being exploited for useful metabolites. Peptides of human origin, acting as hormones, neurotransmitters, growth factors, cytokines, and so forth are involved in the regulation of a variety of functions, including the central nervous system and cardiovascular, gastrointestinal, immunological, reproductive, and growth activities (7–10). Peptide therapeutics have been and are being developed for various treatments (Table 1).

STRUCTURAL FEATURES

The peptides compiled in Table 1 are products of different biosynthetic systems. In most cases, the current state of knowledge permits the prediction of each biosynthetic path from the structural features involved. This information permits the respective biotechnological (fermentation of microbial producers or hosts, cell culture), in vitro enzymatic, or chemical production. The new emerging approaches of combinatorial biosynthesis or synthetic peptide libraries are especially linked to biosynthetic considerations.

Biosynthesis

Peptides of biological origin differ in various structural features (Table 2) but also share the template-mediated assembly from carboxyl-activated amino acid precursors into linear precursor molecules, growing in the carboxy-terminal direction. Depending on the biosynthetic origin, the template is either nucleic acid (ribosomal) or protein (nonribosomal) (5).

Ribosomal Peptide Formation. Ribosomally formed peptides are directly gene encoded, and their mRNA template utilizes the available codons for linear assembly guided by the ribosomal architecture. Each amino acid is activated as aminoacyl-tRNA by a specific aminoacyl-tRNA synthetase; this process includes proofreading steps for structurally related compounds and achieves an accuracy of about 10^{-4} (12). Genes and their respective mRNAs encode polypeptides and proteins up to about 27,000 amino acids (titin), but the average size of translation products is about 300 amino acids.

Amino Acids. The translation process is restricted to the 20 protein amino acids and selenocysteine (13,14). Organisms may differ in their codon usage, and rare codons are thought to function in the control of both the transla-

Table 2. Structural Features of Peptides of Ribosomal and Nonribosomal Origin

	Ribosomal path	Nonribosomal path//multienzymes
Size (amino acids)	No size limitation, 2–20 residues Peptides, 20–50 residues Polypeptides, more than 50 residues Proteins, largest structures up to 30,000 residues	2 to about 50, 4 to 10 dominating; polymers up to about 1,000 residues
Amino acid constituents	21 protein amino acids (including selenocysteine) and modified amino acids	Various types of amino acids including 2-, 3- and 4-amino compounds (more than 300 known)
D-Amino acids	Not more than one, epimerized posttranslationally; in lantibiotics epimerization of several residues by dehydration and thioether formation	Often several, either incorporated directly, or epimerized during synthesis
Non–amino acid constituents in the peptide chain	Acyl residues, amines originating from decarboxylation	Various acyl residues, including aromatic acids, hydroxy acids
Cyclic structures	Rare, frequent disulfide cycles, often with several disulfide linkages; thioether cycles in lantibiotics (lanthionine)	More frequent than linear structures, various peptide bond cyclizations, but also lactones
Side-chain modifications	Hydroxylation, dehydration (Ser, Thr), side-chain cyclization (Cys to thiazoles, Thr to oxazoles, Glu to pyroGlu), frequently glycosylation	N-Methylation, hydroxylation, side-chain cyclization (Cys to thiazole), glycosylation, side-chain cross-linking (aromatic rings), glycosylation
Unusual constituents	Not known	Urea type of peptide bond, phosphoamino acids, amino-modified fatty acids–derived components (lipopeptides), mixed polyketide structures, terpenoid-derived groups
Biosynthesis	Gene can be identified; consider splicing, processing, and posttranslational modification; often prepropeptides detected	Nonribosomal enzyme systems present; peptide families are frequent in the same or related organisms; biosynthesis of rare precursors needed
Sources	Various animals and plants, sometimes bacteria	Mainly bacteria and lower fungi, occasionally plants and insects

tion rates and the folding of the primary intermediates (15,16). Heterologous expression sometimes needs codon adjustment for the intended host. Exploitation of this process to adapt synthetic nonprotein amino acids has been achieved in vitro by making use of nonutilized codons and their respective tRNAs (17). More than 100 mostly nonnatural amino acids have been introduced into peptides by biocatalytic procedures (18,19). Such procedures permit, for example, the direct incorporation of photofunctional amino acids (20). Current work focuses on the introduction of an engineered tRNA/aminoacyl-tRNA synthetase pair into *Escherichia coli*, delivering unnatural amino acids (21). D-Amino acids in ribosomally encoded peptides originate from the respective L-epimers, which are epimerized in chain by specific racemases first isolated from the spider *Agelopsis aperta* (22,23)

Peptide Processing. These primary translation products may undergo a variety of modification reactions. Besides hydrolytic removal of the initiating formyl-methionine residue (24), regions of the pre- or protranslation products may be involved in targeting and transport events, acting as signals recognized by the various proteins or nucleic acid species. Signal sequences may be removed by signal peptidases, and proteolytic processing may even continue; this is well known among hormones. Thus proopiomelanocortin gives rise to different peptides in different tissues, including ACTH, α-MSH, different forms of LPH, and β-endorphin. A rare event, now firmly established, is the joining of peptide fragments, which in analogy to the exon–intron structures of genes have been termed exteins and inteins (25). Cyclization of peptides by peptide bonds is likewise rare (26).

Side-Chain Alterations. Most common alterations besides N-terminal acylation (24) include hydroxylations of proline and lysine residues and the formation of regional specific disulfide bonds. Such processes may involve not only disulfide isomerases (27) but chaperones as well. Multiple disulfide bridges are found in peptides with up to four positionally conserved links. Such conformationally stabilized structures show antimicrobial (defensins, ranalexin, cecropins), antiviral (aborycin), and cytotoxic properties (human neutrophil peptides, HVP), or as the trefoil factor rITF from rat intestine, act as a covalently linked dimer in tissue repair and cell proliferation.

Side-chain modification of cysteine and serine residues to thiazole and oxazole derivatives are found in the gyrase inhibitor microcin and the rhizobial peptide trifolitoxin. A respective transforming enzyme complex has been characterized from the microcin B17 biosynthetic gene cluster (28). Stable thioether links between cysteine and serine or threonine side chains are prominent structural features of the lantibiotics (29–31). The name points to the unusual amino acids lanthionine and methyllanthionine formed upon hydrolysis. The respective biosynthetic clusters contain genes for enzymes catalyzing dehydration of serine or threonine to dehydroalanine and dehydrobutyrate, respec-

tively, which undergo addition with thiols from adjacent cysteines, leading to inversion of the configuration.

In Vitro Applications of the Ribosomal System. Ribosomal peptide synthesis as a complex process involving about 150 proteins and ribonucleic acids and is being studied in detail (32). Functional in vitro systems are available from several organisms including *E. coli*, yeast, wheat germ, and rabbit reticulocytes (33). Such systems are especially useful for the production of antibacterial and cytotoxic peptides, which are not accessable by fermentation (34). Technology for mRNA production or the use of coupled transcription–translation systems introducing plasmids as informational molecules have been worked out (35). So far, peptide synthesis has been restricted to the laboratory scale, but yields up to the milligram per milliliter range have been reported. Efficiencies of such systems can be improved by cofactor regeneration (e.g., ATP). Synthesis of reporter molecules such as photoproteins permit the functional monitoring and the direct detection of, for example, translational inhibitors (36).

Nonribosomal Peptide Formation. Two basic types of peptide-forming systems account for most of the nonribosomal peptide formation structures known so far: single-step and multistep systems (Table 3). Single-step systems add amino acids to activated precursor amino acids or peptides; thus the precursors are directly activated, generally as phosphates by cleavage of ATP to ADP, and the amino acid to be added is in the free soluble state. The limit of performance appears to be the size of the substrates in relation to the enzymes involved. Covalently acting multistep systems add activated precursors to an activated amino acid, thus maintaining the activated state. In analogy to the ribosomal system, these multienzyme systems have no free intermediates, and product release involves catalysis as well (6).

Single-Step Systems. Prominent products of single-step systems are glutathione, coenzyme A, and peptides of the murein sacculus, including the uridin-diphosphomuramyl-N-acetyl precursor and various interpeptide bridges. Some steps involve aminoacyl-tRNA and are thus nucleic acid dependent (37). Enzymes involved can be identified at the gene level by characteristic motifs (38), but the class is heterogenous.

Single-step enzyme systems are obviously suited for the production of repeated sequences and polymers. Such examples include glutathion-related peptides involved in sequenstering metal ions. These have been termed phytochelatins and are induced in yeast or plant cells upon exposure to Cd^{2+}, Cu^{2+}, or Zn^{2+} (39). The compositions vary from (γGlu-Cys)*n*Gly, (γGlu-Cys)*n*Glu, to (γGlu-Cys)*n*, with *n* ranging from 2 to 11 (40,41), and imply dipeptide condensations as a biosynthetic principle. Likewise, cyanophycin, a cyanobacterial storage polypeptide, is composed of the dipeptide units L-Arg-L-Asp. Synthetase and depolymerase have been identified in *Synechocystis* sp. and *Anabaena variabilis*, and the export of the polymerizing activity into *E. coli* has been demonstrated (42). Examples of amino acid polymers include γ-linked glutamate tails of folate (ranging in length from 1 to 11), β-lysine tails of nucleoside antibiotics such as nourseothricin, lysine polymers up to 25 residues in Streptomycetes, and γ-D-glutamyl capsular polymers in *Bacillus anthracis* and other spore-forming Bacilli. The molecular mechanisms of control of peptide and polypeptide chain length are unknown.

Multistep Systems. The covalent catalysis principle found in the multienzyme system closely resembles the ribosomal elongation principle. Equivalents to aminoacyl-tRNA and peptidyl-tRNA are aminoacyl and peptidyl carrier proteins. However, the protein system depends largely on fusing functional domains, which leads to multifunctional structures of impressive sizes. Grouping enzymes according to size, peptide- and polyketide-forming enzymes will occupy all top positions. On the other hand these synthetases will be at the bottom of the list if overall rates are compared. The largest known enzyme, cyclosporin synthetase, contains 41 functional domains fused into a 1.7-MDa protein. It integrates all 40 reactions, leading to a cycloundecapeptide with seven methylated peptide bonds (43). Peptide-forming systems currently under study have been compiled in Table 4.

Activation of Amino Acids and Other Carboxyls. Differing from polyketide systems, activation of amino acid carboxyl groups is an integrated function in peptide synthetases. The activation function is a preselection step in the events controlling fidelity of the peptide chain. Adenylate-forming domains share extensive homologies with acyl-CoA synthetases and the homologous insect luciferases, which likewise form acyl adenylates followed by transfer to the thiol group of CoA. Amino acid selection is controlled with varying selectivity: sometimes there is discrimination of leucine against valine or isoleucine, but occasionally all branched-chain types are accepted. Likewise phenylalanine might be discriminated against in favor of tyrosine and tryptophan, whereas some activation sites are known to accept various aromatic amino acids. Stereoselective discrimination of L- and D-residues is common, and both epimers may be direct precursors. In some cases both epimers are substrate, but the subsequent condensation reaction is stereocontrolled. Adenylates are thought to be stabilized by the interaction of the two subdomains of the adenylate domain. With low efficiency they may be subject

Table 3. Peptide-Forming Enzyme Systems

	Single-step systems	Multistep systems
Types of peptide made	Linear peptides, branched peptides	Linear and cyclic peptides, peptidolactones and depsipeptides; polymers?
Length of peptide made	2 to 5, or polymers	2 to 48 (?)
Activation of carboxyl groups	Phosphate or adenylate	Adenylate
Intermediates	Free intermediates (not clear for polymers)	Intermediates remain enzyme bound

Table 4. Nonribosomal Peptide Synthetase Systems

Peptide	Organism	Structural type	Gene(s) cloned	Enzymology
		Linear		
Bacilysin	*Bacillus subtilis*	P-2-M	(+)	(+)
Anguibactin	*Vibrio anguillarum*	R-P-2-M	+	−
ACV	*Streptomyces clavuligerus*	P-3	+	+
	Nocardia lactamdurans		+	+
	Lysobacter lactamgenus		+	+
	Aspergillus nidulans		+	+
	Penicillium chrysogenum		+	+
	Acremonium chrysogenum			
Bialaphos	*Streptomyces hygroscopicus*	P-3	+	+
Yersiniabactin	*Yersinia enterocolitica*	R-P-3-M	+	+
Phaseolotoxin	*Pseudomonas syringae* pv. ph.	P-4-M	(+)	−
Ardacin	*Kibdelosporangium aridum*	P-7-M	(+)	−
Chloroeremomycin	*Amycolatopsis orientalis*	P-7-M	+	−
Pyoverdin	*Pseudomonas aeruginosa*	R-P-8-M	+	−
Bleomycin	*Streptomyces verticillius*	P-8-M	+	−
Alamethicin	*Trichoderma viride*	R-P-19-M	−	+
		Cyclopeptides		
Enterobactin	*Escherichia coli*	P-C-E-3	+	+
HC-toxin	*Cochliobolus carbonum*	C-4	+	+
Tentoxin	*Alternaria alternata*	C-4	−	+
Echinocandin	*Aspergillus nidulans*	R-C-6	−	(+)
Microcystin	*Microcystis aeruginosa*	C-7	(+)	−
Iturin	*Bacillus subtilis*	C-8	(+)	−
Gramicidin S	*Bacillus brevis*	C-$(P\text{-}5)_2$	+	+
Tyrocidin	*Bacillus brevis*	C-10	+	+
Cyclosporin	*Tolypocladium niveum*	C-11-M	+	+
Mycobacillin	*Bacillus subtilis*	C-13	−	+
		Lactones		
Actinomycin	*Streptomyces chrysomallus*	R-$(L\text{-}5)_2$-M	+	+
Destruxin	*Metarhizium anisopliae*	L-6	+	(+)
Etamycin	*Streptomyces griseus*	R-L-7	−	(+)
Surfactin	*Bacillus subtilis*	L-8	+	+
Quinomycin	*Streptomyces echinatus*	$(R\text{-}P\text{-}4)_2$	−	+
R106	*Aureobasidium pullulans*	L-9	−	(+)
Syringomycin	*Pseudomonas syringae*	R-L-9	(+)	−
		Syringostatin		
SDZ90-215	*Septoria* sp.	L-10	−	+
		Depsipeptides		
Enniatin	*Fusarium* sp.	C-$(P_2)_3$-M	+	+
Beauvericin	*Beauveria bassiana*	C-$(P_2)_3$-M	−	+
PF1022	*Mycelia sterilia*	C-$(P_2)_4$-M	−	(+)
		Branched polypeptides		
Fengycin	*Bacillus subtilis*	P-10-C-7	+	−
Polymyxin	*Bacillus polymyxa*	P-10-C-7	−	+
Bacitracin	*Bacillus licheniformis*	P-12-C-7	+	+
Nosiheptide	*Streptomyces actuosus*	R-P-13-C-10-M	−	(+)
Thiostrepton	*Streptomyces laurentii*	R-P-17-C-10-M	−	(+)
		Branched peptidolactones		
Lysobactin	*Lysobacter* sp.	P-11-L-9	(+)	+
A21798A	*Streptomyces roseosporus*	R-P-13-L-10	(+)	+
A54145	*Streptomyces fradiae*	R-P-13-L-10	−	+
Tolaasin	*Pseudomonas tolaasii*	R-P-18-L-5	(+)	−

Note: Abbreviations: P, peptide; C, cyclopeptide; L, lactone; E, ester; R, acyl; M, modified. The structural types are defined by the number of amino, imino, or hydroxy acids in the precursor chain. The ring sizes of cyclic structures are indicated in the number following C, L, or E, defining the type of ring closure.

to aminolysis, leading to dipeptide formation without stereoselectivity (44).

Acyl and Peptidyl Carrier Domains. Adjacent to each adenylate domain is the carrier domain, homologous to the well-known acyl carrier protein of fatty acid biosynthesis. Carrier domains are posttranslationally modified with 4′-phosphopantetheine. This modification is catalyzed by specific CoA-4′-phosphopantetheine-protein transferases (45,46). The terminal cysteamine thiol of the cofactor is the acceptor of the aminoacyl moiety of the adenylate, with release of AMP. Thiol-bound intermediates can be isolated and characterized. This system property has led to the name thiotemplate mechanism.

Condensation Domains. The directed condensation of either two aminoacyl intermediates (initiation) or a peptidyl and an aminoacyl intermediate (elongation) is thought to take place at a condensation domain, interacting with the two adjacent carrier domains. This reaction is strictly stereospecific. If both epimers are present as intermediates because of the action of a reversible epimerase in the multienzyme, only the respective D-isomer is processed. The condensation domains are highly similar in structure to epimerization domains, catalyzing the conversion of thioester-attached aminoacyl or peptidylcarboxyl groups. Special types of condensation domains have been identified in systems forming thiazolidine-containing peptides such as bacitracin or yersiniabactin (47). These perform the cysteine side-chain modifications as well.

Termination Reactions. Termination of synthesis may proceed by hydrolysis (thioesterase required) or various modes of cyclization, leading to cyclic or branched cyclic peptides, peptidolactones, or cyclodepsipeptides. Flexibility of the process permits the processing of hydroxy acids as well, leading to depsipeptides, or the repeated use of templates, leading to dimers, trimers, or tetramers. Known processes are the dimerization of two pentapeptides in gramicidin S formation, the trimerization of dihydroxybenzoylseryl- or hydroxyisovaleryl-N-methylvalyl intermediates in enterobactin or enniatin synthesis, respectively, or the suspected tetramerization in the cases of bassianolide and PF1022. As in polyketide biosynthesis, repositioning of a terminating thioesterase domain may promote the release of intermediates (48). Both integrated and and associated thioesterase-like proteins are suspected to function in termination reactions (49).

Side Reactions. Two types of side reactions have been observed so far: the aminolysis of aminoacyladenylates as reactive mixed anhydrides, leading to dipeptides (44), or reactions of thioester intermediates. As in solid-phase peptide synthesis, certain steric conditions favor abortive cyclization reactions, especially of proline-containing or N-methylated dipeptides and ornithyl side chains. Less expected have been out-of-sequence dipeptide formations in ACV syntetase, actinomycin synthetase II (50), and cyclosporin synthetase. Such side reactions are promoted when lack of substrates prevents the completion of the entire catalytic cycle.

Reprogramming of Peptide Synthetases. The modular arrangement of most polycondensation pathways implies reprogramming by modular exchange. In the peptide field, functional hybrid synthetases have been constructed by module exchange between the gramicidin S synthetase system, ACV synthetase, and the surfactin synthetase system (51,52). At present the often low catalytic efficiencies of the constructs present a problem difficult to approach without the knowledge of the detailed spatial structures involved.

In Vitro Applications of Peptide Synthetases (53). The in vitro studies of peptide synthetases permitted an understanding of the catalytic principles involved. Small amounts of peptides have been produced in a variety systems (Table 3). Applying cyclosporin synthetase for preparative work, more than 100 cycosporin analogs have been prepared (43). Generally no more than two structural alterations have been found to be tolerated. Peptide synthetases obviously do not match the performance of chemical methods with regard to structural analog utilization because additional constraints are involved, including substrate selection and protein stability. However, peptide synthetases combine a number of advantages, including stereoselectivity and efficient cyclization. Another feature of special significance is the target-oriented product profile, which provides essential clues in structure–function work.

PEPTIDE PRODUCTION

Sources and Screening

Healing experiences and later ecological considerations have led to the concept of microbial warfare and eventually to the isolation of the first peptide antibiotics penicillin, tyrothricin (a mixture of linear gramicidin and tyrocidine), and gramicidin S. Although investigations of this type are continuing with the still-large reservoir of rare and difficult-to-cultivate microorganisms, new fields have been emerging that consider the defensive compounds from higher organisms. Pioneered in the fields of amphibian and insect defense strategies (54–57), antimicrobial peptides keep being discovered in mammalian and plant sources (58–61). In the field of secondary metabolites, microbial sources are generally exploited by fermentational procedures. Products of plants and animals have been investigated largely from chemical synthesis, but production can be achieved by recombinant DNA approaches (62,63).

Besides functional and chemical screening of extracts, genetic screening is now emerging as a tool to identify sources of presumably active natural products. The increasing knowledge of biosynthetic enzyme systems permits the probing of DNA isolates for the presence of peptide synthetase genes (64) or related modifying enzymes.

Fermentation. Submersed fermentation accounts for the majority of peptides produced from bacteria and filamentous fungi. Research has shifted from engineering approaches to detailed studies of the molecular mechanisms involved. Thus the productivity of cultures is increased by empirical media optimization, tight parameter control of the growth and production processes (O_2, CO_2, redox potential, cell morphology), and random and defined mutagenesis operations. The complexity of metabolite production is far from being understood.

The fermentation of β-lactam antibiotics represents the best-studied system (65). Penicillin biosynthesis is catalyzed by three enzymes: δ-(L-α-aminoadipyl)-l-cysteinyl-D-valine synthetase (ACVS), isopenicillin-N-synthase (IPNS), and acyltransferase (AAT). The respective genes are are combined in a biosynthetic cluster, which has been found to be amplified in production strains (66). Thus the gene copy number appears to be a major factor of productivity. However, production strains with a single copy of the cluster have also been selected as efficient producers. The regulation of expression is being currently studied, which affords the detailed understanding of the complex promoters involved (67). In this respect the utilization of autoregulators to induce the biosynthesis of specific metabolites is a fascinating tool already applied in peptide fermentation in Actinomycetes (68).

In addition, subcellular localization plays a significant role. Although IPNS is soluble and located in the cytosol, ACVS seems only partially soluble, and AAT is located in a peroxisome-like compartment, termed the microbody. The substrate transfer between ACVS and IPNS may involve complex formation and channeling, but the transfer of isopenicillin-N into the microbody, and the export of penicillins out of the microbody and through the cellular envelope, have not been studied so far. The role of molecular transport processes involved in biosynthetic processes is largely unknown. Export processes of bioactive compounds, especially peptides, seem also to be essential for self-protection, and thus critical for an intended overproduction. Peptide export proteins have been detected, and compounds such as cyclosporin are known to affect export systems including multi-drug-resistance proteins. Production processes utilize both product excretion and retainment in the microbial cells. Gramicidin S can accumulate up to 20% peptide within the cellular dry mass of the producer *Bacillus brevis*. Fermentation techniques to avoid feedback problems utilize solid-state approaches (69), microfiltration (70), or aqueous two-phase systems (71).

Finally, efficient substrate concentrations play a significant role in peptide fermentation. In special cases direct feeding of amino acid precursors may shift the product profile to a certain analog (72). Considering however, that the price of the penicillin precursor amino acid cystein exceeds that of the product, improvement focuses on the precursor biosynthetic systems, including uptake of basic nutrients such as glucose, ammonium, and sulfate.

Recombinant DNA Approaches in Peptide Production

The introduction of rDNA techniques permits the efficient production of various peptides independent of organ or animal sources, and thus free of viral contaminants. Current limits are related to certain posttranslational modification events, if relevant, which still have to be evaluated. An amidating enzyme system has been established (73), whereas specific glycosylation systems are still unavailable. So current production routes employ, for example, yeast or insect cell cultures that often contain similar modification systems. Prominent examples of peptides currently produced include insulin (74) and hirudin, but the vast majority of products are proteins.

Enzymatic Procedures in Peptide Production

Introduced more than 50 years ago, the reversal of proteinase reactions has been widely used in peptide synthesis (75). Strategies exist for regio- and stereospecific synthesis of peptide bonds (76,77), and new trends consider heterogeneous media (78) or even frozen aqueous solutions (79). Generally all procedures permit the construction of a single, specified peptide bond, but fragment condensations and protein modifications have been established in peptide production as well. Substrates to be converted generally need protection groups to avoid side reactions.

BIBLIOGRAPHY

1. T. Wieland and M. Bodanszky, *The World of Peptides*, Springer-Verlag, Berlin, 1991.
2. E. Gross, J. Meienhofer, and S. Udenfriend eds., *The Peptides: Analysis, Synthesis, Biology*, vols. I–IX, Academic, New York, 1979–1987.
3. H. Kleinkauf and H. von Döhren eds., *Peptide Antibiotics*, de Gruyter, Berlin, 1990.
4. W. Strohl ed., *Biotechnology of Antibiotics*, Marcel Dekker, New York, 1997.
5. H. Kleinkauf and H. von Döhren, *Eur. J. Biochem.* **236**, 335–351 (1996).
6. H. von Döhren, U. Keller, J. Vater, and R. Zocher, *Chem. Rev.* **97**, 2675–2705 (1997).
7. R.C. Hider and D. Barlow, eds., *Polypeptide and Protein Drugs: Production, Characterization and Formulation*, Ellis Horwood, Chichester, 1991.
8. A. Negro-Vilar and P.M. Conn eds., *Peptide Hormones: Effects and Mechanisms of Action*, vols. I–III, CRC Press, Boca Raton, Fla., 1988.
9. W.V. Williams and D.B. Weiner eds., *Biologically Active Peptides: Design, Synthesis and Utilization*, Technomic, Lancaster, Pa., 1993.
10. D.J. Ward, *Peptide Pharmaceuticals: Approaches to the Design of Novel Drugs*, Open University Press, Buckingham, England, 1991.
11. H. Kleinkauf and H. von Döhren, in H.-J. Rehm and G. Reed eds., *Biotechnology*, 2nd ed., vol. 7, Verlag Chemie, Weinheim, 1997, pp. 277–322.
12. J.G. Arnez and D. Moras, *Trends Biochem. Sci.* **22**, 211–216 (1997).
13. T.C. Stadtman, *Annu. Rev. Biochem.* **65**, 83–100 (1996).
14. S.C. Low and M.J. Berry, *Trends Biochem. Sci.* **21**, 203–208 (1996).
15. X. Xia, *Genetics* **144**, 1309–1320 (1996).
16. A.A. Adzubei, I.A. Adzubei, I.A. Krashennikov, and S. Neidle, *FEBS Lett.* **399**, 78–82 (1996).
17. M. Ibba, *Biotechnol. Genet. Eng. Rev.* **13**, 197–216 (1996).
18. J.D. Bain, C. Switzer, A. Chamberlin, and S. Benner, *Nature* **356**, 537–539 (1992).
19. D. Mendel, V.W. Cornish, and P.G. Schultz, *Annu. Rev. Biophys. Biomol. Struct.* **24**, 435–462 (1995).
20. T. Hohsaka, K. Sato, M. Sisido, K. Takai, and S. Yokoyama, *FEBS Lett.* **344**, 171–174 (1994).
21. D.R. Liu, T.J. Magliery, M. Pastrnak, and P.G. Schultz, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10092–10097 (1997).
22. G. Kreil, *Annu. Rev. Biochem.* **66**, 337–346 (1997).

23. S.D. Heck, W.S. Faraci, P.R. Kelbaugh, N.A. Saccomano, P.F. Tadeio, and R.A. Volkmann, *Proc. Natl. Acad. Sci.* **93**, 4036–4039 (1995).
24. R.A. Bradshaw, W.W. Brickley, and K.W. Walker, *Trends Biochem. Sci.* **23**, 263–267 (1998).
25. A.A. Cooper and T.H. Stevens, *Trends Biochem. Sci.* **20**, 351–356 (1995).
26. M. Martinez-Bueno, M. Maqueda, A. Gàlvez, B. Samyn, J. van den Beeumen, J. Coyette, and E. Valdivia, *J. Bacteriol.* **176**, 6334–6339 (1994).
27. J.M. Luz and W.J. Lennarz, *EXS 77*, 97–117 (1996).
28. Y.M. Li, J.C. Milene, L.L. Madison, R. Kolter, and C.T. Walsh, *Science* **274**, 1188–1193
29. R. Jack, F. Götz, and G. Jung, in H.-J. Rehm and G. Reed eds., *Biotechnology*, 2nd ed., Vol. 7, Verlag Chemie, Weinheim, 1997, pp. 323–370.
30. R.W. Jack, J.R. Tagg, and B. Ray, *Microbiol. Rev.* **59**, 171–200 (1995).
31. H.-G. Sahl, R.W. Jack, and G. Bierbaum, *Eur. J. Biochem.* **230**, 827–853 (1995).
32. K.H. Nierhaus, F. Fransceschi, A.R. Subramanian, V.A. Erdmann, and B. Wittmann-Liebold eds., *The Translational Apparatus*, Plenum, New York, 1993.
33. G.S. Beckler, D. Thompson, and T. Van Osbree, *Methods Mol. Biol.* **37**, 215–232 (1995).
34. W. Stiege and V.A. Erdmann, *J. Biotechnol.* **41**, 81–90 (1995).
35. L. Ryabova, E. Volianik, O. Kurnasov, A. Spirin, Y. Wu, and F.R. Kramer, *J. Biol. Chem.* **269**, 1501–1505 (1994).
36. S.V. Matveev, B.A. Illarionov, E.S. Vysotski, V.S. Bondar, S.V. Markova, and Y.B. Alakhov, *Anal. Biochem.* **231**, 34–39 (1995).
37. J.-M. Ghuysen and R. Hakenbeck eds., *Bacterial Cell Wall*, Elsevier, Amsterdam, 1994.
38. S.S. Eveland, D.L. Pompliano, and M.S. Anderson, *Biochemistry* **36**, 6223–6229 (1997).
39. W.E. Rauser, *Annu. Rev. Biochem.* **59**, 61–86 (1990).
40. J.C. Steffens, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **41**, 533–575 (1990).
41. P. Meuwly, P. Thibault, A.L. Schwan, and W.E. Rauser, *Plant J.* **7**, 391–400 (1995).
42. K. Ziegler, A. Diener, C. Herpin, R. Richter, R. Deutzmann, and W. Lockau, *Eur. J. Biochem.* **254**, 154–159 (1998).
43. H. Kleinkauf and H. von Döhren, in T. Anke ed., *Fungal Biotechnology*, Chapmann & Hall, Weinheim, 1997, pp. 147–158.
44. M.F. Byford, J.E. Baldwin, C.-Y. Shiau, and C.J. Schofield, *Chem. Rev.* **97**, 2631–2649 (1997).
45. C.T. Walsh, A.M. Gehring, P.H. Weinreb, L.E.N. Quadri, L.E.N., and R.S. Flugel, *Curr. Opin. Chem. Biol.* **1**, 309–315 (1997).
46. L.E.N. Quadri, P.H. Weinreb, M. Lei, M.M. Nakano, P. Zuber, and C.T. Walsh, *Biochemistry* **37**, 1585–1595 (1998).
47. D. Konz, A. Klens, K. Schörgendörfer, and M.A. Marahiel, *Chem. Biol.* **4**, 927–937 (1997).
48. F. de Ferra, F. Rodriguez, O. Tortora, C. Tosi, and G. Grandi, *J. Biol. Chem.* **272**, 25304–25309 (1997).
49. A. Schneider and M.A. Marahiel, *Arch. Microbiol.* **169**, 404–410 (1998).
50. F. Schauwecker, F. Pfennig, W. Schröder, and U. Keller, *J. Bacteriol.* **180**, 2468–2474 (1998).
51. T. Stachelhaus, A. Schneider, and M.A. Marahiel, *Science* **269**, 69–72 (1995).
52. A. Schneider, T. Stachelhaus, and M.A. Marahiel, *Mol. Gen. Genet.* **257**, 308–318 (1998).
53. H. Kleinkauf and H. von Döhren, *Acta Biochim. Polonica* **44**, 839–848 (1997).
54. H.G. Boman, J. Marsh, and J.A. Goodie eds., *Antimicrobial Peptides*, Ciba Foundation Symposium 186, Wiley, Chichester, 1994.
55. C.L. Bevins and M. Zasloff, *Annu. Rev. Biochem.* **59**, 395–414 (1990).
56. B.L. Kagan, T. Ganz, and R.I. Lehrer, *Toxicology* **87**, 131–149 (1994).
57. H.G. Boman, *Annu. Rev. Immunol.* **13**, 61–92 (1995).
58. R.I. Lehrer, A.K. Lichtenstein, and T. Ganz, *Annu. Rev. Immunol.* **11**, 105–128 (1993).
59. W.L. Maloy and U.P. Kari, *Biopolymers* **37**, 105–122 (1995).
60. R.E.W. Hancock, T. Falla, and M. Brown, *Adv. Microb. Physiol.* **37**, 135–175 (1995).
61. W.F. Broekaert, B.P.A. Commue, M.F.C. De Bolle, K. Thevissen, G.W. De Samblanx, and R.W. Osborn, *Crit. Rev. Plant Sci.* **16**, 297–323 (1997).
62. C. Binnie, J.D. Cossar, and D.I. Stewart, *Trends Biotechnol.* **15**, 315–320 (1997).
63. G. Gelissen, C.P. Hollenberg, and Z.A. Janowicz, *Bioprocess Technol.* **22**, 195–239 (1995).
64. K. Turgay and M.A. Marahiel, *Peptide Res.* **7**, 238–240 (1994).
65. J. Nielsen, *Physiological Engineering Aspects of Penicillium chrysogenum*, Polyteknisk Forlag, Lyngby, Denmark, 1995.
66. R.W. Newbert, B. Barton, P. Greaves, J. Harper, and G. Turner, *J. Ind. Microbiol. Biotechnol.* **19**, 18–27 (1997).
67. A. Brakhage, *FEMS Microbiol. Lett.* **148**, 1–10 (1997).
68. Y.K. Yang, M. Morikawa, H. Shimizu, S. Shioya, K.-I. Suga, T. Nihira, and Y. Yamada, *Biotechnol. Bioeng.* **49**, 437–444 (1996).
69. A. Ohno, T. Ano, and M. Shoda, *Process Biochem.* **31**, 801–806 (1996).
70. M. Taniguchi, K. Hoshino, H. Urasaki, and M. Fujii, *J. Ferment. Bioeng.* **78**, U122 (1996).
71. R. Kuboi, T. Maruki, H. Tanaka, I. Komasawa, *J. Ferment. Bioeng.* **78**, 431–436 (1994).
72. R. Thiericke, and R. Rohr, *Nat. Prod. Rep.* **10**, 263–289 (1993).
73. S.T. Prigge, A.S. Kolhekar, B.A. Eipper, R.E. Mains, and L.M. Amzel, *Science* **278**, 1300–1305 (1997).
74. M.R. Ladisch and K.L. Kohlmann, *Biotechnol. Prog.* **8**, 469–478 (1992).
75. J.S. Fruton, *Adv. Enzymol.* **53**, 239–305 (1982).
76. I. Gill, R. Lopez-Fandino, X. Jorba, and E.N. Vulfson, *Enzyme Microb. Technol.* **18**, 163–183 (1996).
77. J. Bongers and E.P. Heimer, *Peptides* **45**, 183–193 (1995).
78. I. Gill and E. Vulfson, *Trends Biotechnol.* **12**, 118–122.
79. M. Hansler and H.D. Jakubke, *J. Peptide Sci.* **2**, 279–289 (1996).

PHENYLALANINE

IAN G. FOTHERINGHAM
NSC Technologies
Mount Prospect, Illinois

KEY WORDS

Amino acid
Ammonia-lyase
Aromatic
Biosynthesis
Biocatalyst
Biotransformation
Fermentation
Phenylalanine
Resolution
Transaminase

OUTLINE

INTRODUCTION

The amino acid phenylalanine is currently manufactured worldwide at an annual scale exceeding 12,000 metric tons. The vast majority of this material is L-phenylalanine, synthesized specifically for incorporation into the dipeptide sweetener aspartame. Production of L-phenylalanine grew rapidly during the 1980s in parallel with the demand for aspartame and has grown steadily in recent years. L-Phenylalanine is also used in additional food and medical applications and is in increasing demand in pharmaceutical development as a chiral intermediate or as a precursor of chiral auxiliaries such as benzyloxazolidinone. In contrast, only a few metric tons of D-phenylalanine are used annually, also in pharmaceutical drug development, although this application is increasing. During the past two decades, a number of manufacturers have developed chemoenzymatic and purely biological routes for L-phenylalanine manufacture, as shown in Table 1. Chemoenzymatic syntheses have relied on resolution technology, applied to chemically prepared racemic phenylalanine derivatives or DL-5-monosubstituted hydantoins and biotransformation routes from chemically prepared achiral precursors of L-phenylalanine. Purely biological routes have concentrated on direct synthesis of L-phenylalanine through microbial fermentative means. Today, the largest and most commercially successful synthetic routes to L-phenylalanine production use large-scale fermentation of bacterial strains that overproduce L-phenylalanine. Such microorganisms have been isolated through a combination of classical mutagenesis selection procedures and molecular genetic manipulation, the latter deriving from a detailed understanding of the molecular biology of the pathways involved in bacterial L-phenylalanine biosynthesis. Conversely, the present commercial production of D-phenylalanine relies solely on chemoenzymatic resolution.

L-PHENYLALANINE-OVERPRODUCING MICROORGANISMS

Common Aromatic and L-Phenylalanine Biosynthetic Pathways

Efforts to develop L-phenylalanine-overproducing organisms have been vigorously pursued by Nutrasweet Company, Ajinomoto, Kyowa Hakko Kogyo, and others. The focus has centered on bacterial strains that have previously demonstrated the ability to overproduce other amino acids. Such organisms include principally the coryneform bacteria, *Brevibacterium flavum* (1,2) and *Corynebacterium glutamicum* (3,4) used in L-glutamic acid production. In addition, *Escherichia coli* (5) has been extensively studied in L-phenylalanine manufacture because of the detailed characterization of the molecular genetics and biochemistry of

Table 1. Commercial Processes for L-Phenylalanine Synthesis

Manufacturer	L-Phenylalanine processes developed
Ajinomoto	Microbial fermentation
Biotechnica	Microbial fermentation
Degussa/Rexim	Enzymatic resolution (aminoacylase)
DSM	Enzymatic resolution (amino acid amidase)
Genex	Biotransformation (phenylalanine ammonia lyase)
Kyowa Hakko	Microbial fermentation
Miwon	Microbial fermentation
Nutrasweet Co.	Microbial fermentation, biotransformation (aminotransferase)
PEI	Biotransformation (aminotransferase)
Tanabe Seiyaku	Enzymatic resolution, biotransformation (aminoacylase, aminotransferase)

its aromatic amino acid pathways and its amenability to recombinant DNA methodology. The biochemical pathway that results in the synthesis of L-phenylalanine from chorismate is identical in each of these organisms (Fig. 1**b**); the common aromatic pathway to chorismate is shown in Figure 1**a**. In each case, the L-phenylalanine biosynthetic pathway comprises three enzymatic steps from chorismic acid, the product of the common aromatic pathway. The precursors of the common aromatic pathway, phosphoenolpyruvate (PEP) and erythrose-4-phosphate, derive from the glycolytic and pentose phosphate pathways of sugar metabolism, respectively. Although the intermediate compounds in both pathways are identical in each of the organisms, there are differences in the organization and regulation of the genes and enzymes involved (2,6,7). Nevertheless, the principal points of pathway regulation are very similar in each of the three bacteria (4,6,8). Conversely, tyrosine biosynthesis, which is also carried out in three biosynthetic steps from chorismate, proceeds through different intermediates in *E. coli* than the coryneform organisms (6,8).

Regulation of phenylalanine biosynthesis occurs both in the common aromatic pathway and in the terminal phenylalanine pathway, and the vast majority of efforts to deregulate phenylalanine biosynthesis have focused on two specific rate-limiting enzymatic steps. These are the steps of the aromatic and the phenylalanine pathways carried out, respectively, by the enzymes 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) synthase and prephenate dehydratase. Classical mutagenesis approaches using toxic amino acid analogues and the molecular cloning of the genes encoding these enzymes have led to very significant increases in the capability of host strains to overproduce L-phenylalanine. This has resulted from the elimination of the regulatory mechanisms controlling enzyme synthesis and specific activity. The cellular mechanisms that govern the activity of these particular enzymes are complex and illustrate many of the sophisticated means by which bacteria control gene expression and enzyme activity. Chorismate mutase, the first step in the phenylalanine-specific pathway, and the shikimate kinase activity of the common aromatic pathway are subject to a lesser degree of regulation and have also been characterized in detail (9,10).

Classical Mutagenesis and Selection

Classical methods of strain improvement have been widely applied in the development of phenylalanine overproducing organisms (5,11,12,13). Tyrosine auxotrophs have frequently been used in efforts to increase phenylalanine production through mutagenesis. These strains often already overproduce phenylalanine because of the overlapping nature of tyrosine and phenylalanine biosynthetic regulation (3,14,15). Limiting tyrosine availability leads to partial genetic and allosteric deregulation of common biosynthetic steps (16). Such strains have been subjected to a variety of mutagenesis procedures to further increase the overall titer and the efficiency of phenylalanine production. In general, this has involved the identification of mutants that display resistance to toxic analogues of phenylalanine or tyrosine, such as β-2-thienyl-DL-alanine and *p*-fluoro-DL-phenylalanine (3,5,17). Such mutants can be readily identified on selective plates in which the analogue is present in the growth medium. The mutations responsible for the phenylalanine overproduction have frequently been located in the genes encoding the enzymatic activities of DAHP synthase, chorismate mutase, and prephenate dehydratase. In turn, this has prompted molecular genetic approaches to further increase phenylalanine production through the isolation and in vitro manipulation of these genes, as described next.

Deregulation of DAHP Synthase Activity

The activity of DAHP synthase commits carbon from intermediary metabolism to the common aromatic pathway converting equimolar amounts of PEP and E4P to DAHP (18). In *E. coli* there are three isoenzymes of DAHP synthase of comparable catalytic activity encoded by the genes *aroF*, *aroG*, and *aroH* (19). Enzyme activity is regulated by the aromatic amino acids tyrosine, phenylalanine, and tryptophan, respectively (16,20–23). In each case, regulation is mediated both by repression of gene transcription and by allosteric feedback inhibition of the enzyme, though to different degrees.

The *aroF* gene lies in an operon with *tyrA*, which encodes the bifunctional protein chorismate mutase–prephenate dehydrogenase (CMPO). Both genes are regulated by the TyrR repressor protein complexed with tyrosine. The *aroF* gene product accounts for 80% of the total DAHP synthase activity in wild-type *E. coli* cells. The *aroG* gene is repressed by the TyrR repressor protein complexed with phenylalanine and tryptophan. Repression of *aroH* is mediated by tryptophan and the TrpR repressor protein (20). The gene products of *aroF* and *aroG* are almost completely feedback inhibited respectively by low concentrations of tyrosine or phenylalanine (24), whereas the *aroH* gene product is subject to maximally 40% feedback inhibition by tryptophan (24). In *B. flavum*, DAHP synthase forms a bifunctional enzyme complex with chorismate mutase and is feedback inhibited by tyrosine and phenylalanine synergistically but not by tryptophan (25,26). Similarly, in *C. glutamicum*, DAHP synthase is inhibited most significantly by phenylalanine and tyrosine acting in concert (27), but unlike *B. flavum* reportedly does not show tyrosine-mediated repression of transcription (25,28).

Many examples of analogue-resistant mutants of these organisms display reduced sensitivity of DAHP synthase to feedback inhibition (3,5,7,25,29,30). In *E. coli*, the genes encoding the DAHP synthase isoenzymes have been characterized and sequenced (6,31,32). The mechanism of feedback inhibition of the *aroF*, *aroG*, and *aroH* isoenzymes has been studied in considerable detail (32), and variants of the *aroF* gene on plasmid vectors have been used to increase phenylalanine overproduction. Simple replacement of transcriptional control sequences with powerful constitutive or inducible promoter regions and the use of high copy number plasmids has readily enabled overproduction of the enzyme (33,35), and reduction of tyrosine-mediated feedback inhibition has been described using resistance to the aromatic amino acid analogues β-2-thienyl-DL-alanine and *p*-fluoro-DL-phenylalanine (33,35).

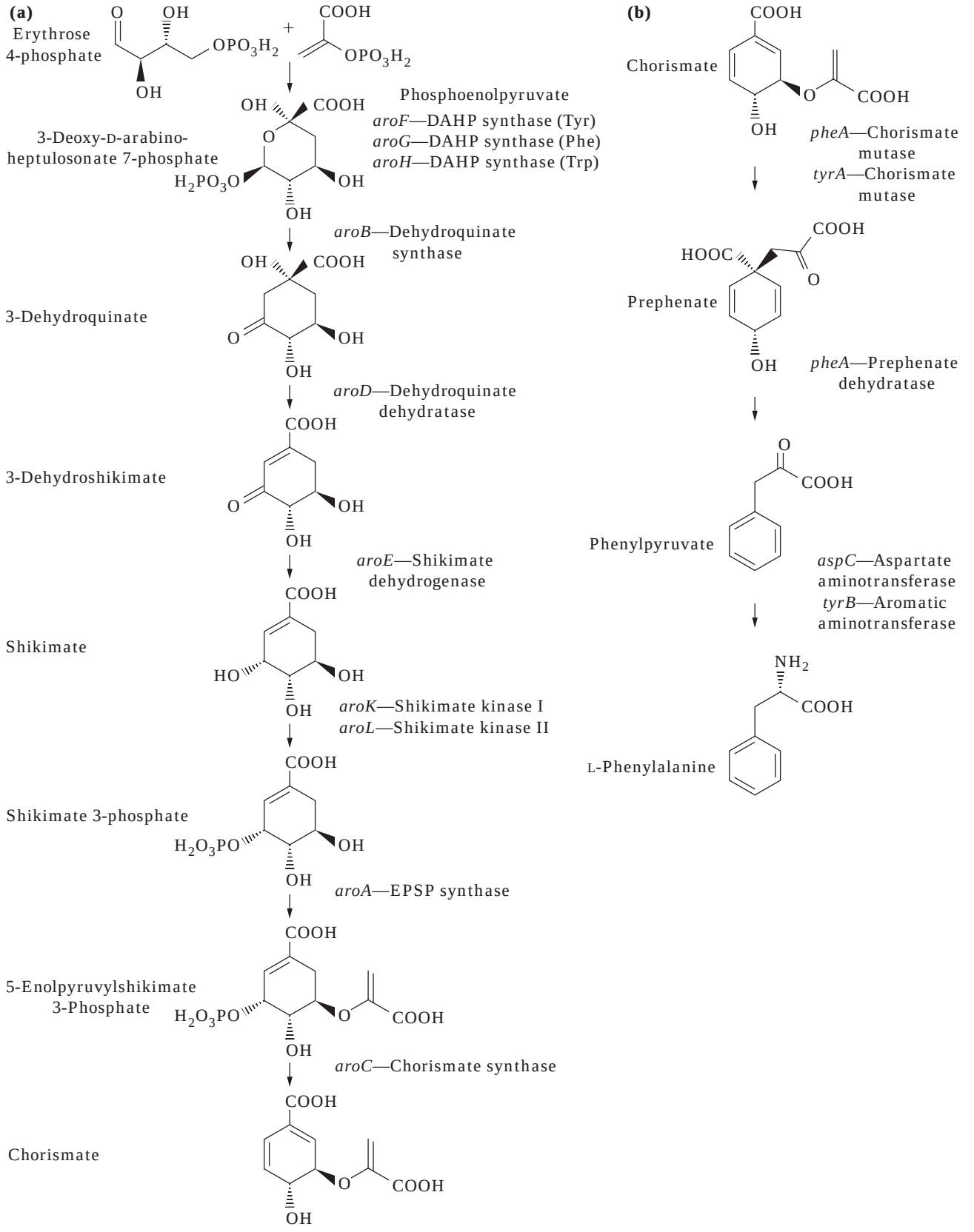


Figure 1. (**a**) The common aromatic pathway to chorismate in *Escherichia coli* K12. The enzymes responsible for each step are indicated to the right preceded by the encoding gene mnemonic. (**b**) The L-phenylalanine biosynthetic pathway from chorismate to L-phenylalanine in *Escherichia coli* K12.

Deregulation of CMPD

The three enzymatic steps by which chorismate is converted to phenylalanine appear to be identical among *C. glutamicum*, *B. flavum*, and *E. coli*, although only in *E. coli* have detailed reports appeared upon the characterization of the genes involved. In each case, the principal regulatory step is that catalyzed by prephenate dehydratase. In *E. coli*, chorismate is converted first to prephenate and then to phenyl pyruvate by the action of a bifunctional enzyme, CMPD, encoded by the pheA gene (6). The final step, in which phenyl pyruvate is converted to L-phenylalanine, is carried out predominantly by the aromatic aminotransferase encoded by *tyrB* (36). However, both the aspartate aminotransferae, encoded by *aspC* and the branched-chain aminotransferase encoded by *ilvE*, can efficiently catalyse this reaction (37). Phenylalanine biosynthesis is regulated by control of CMPD through phenylalanine-mediated attenuation of *pheA* transcription (6) and by feedback inhibition of the prephenate dehydratase and chorismate mutase activities of the enzyme. Inhibition is most pronounced upon the prephenate dehydratase activity, with almost total inhibition observed at micromolar phenylalanine concentrations (38,39). Chorismate mutase activity in contrast is maximally inhibited by only 40% (39). In *B. flavum*, prephenate dehydratase and chorismate mutase are encoded by distinct genes. Prephenate dehydratase is again the principal point of regulation, with the enzyme subject to feedback inhibition, but not transcriptional repression, by phenylalanine (40,41). Chorismate mutase, which in this organism forms a bifunctional complex with DAHP synthase, is maximally inhibited by phenylalanine and tyrosine to 65%, but this is significantly diminished by the presence of very low levels of tryptophan (42). Expression of chorismate mutase is repressed by tyrosine (29,43). The final step is carried out by at least one transaminase (44). Similarly, in *C. glutamicum*, the activities are encoded in distinct genes, with prephenate dehydratase again being more strongly feedback-inhibited by phenylalanine (45–47). The only transcriptional repression reported is that of chorismate mutase by phenylalanine. The *C. glutamicum* genes encoding prephenate dehydratase and chorismate mutase have been isolated and cloned from analogue-resistant mutants of *C. glutamicum* (8,48) and used along with the cloned DAHP synthase gene to augment L-phenylalanine biosynthesis in overproducing strains of *C. glutamicum* (48).

Many publications and patents have described successful efforts to reduce and eliminate regulation of prephenate dehydratase activity in phenylalanine overproducing organisms (5,9,34,35). As with DAHP synthase, the majority of the reported efforts have focused on the *E. coli* enzyme encoded by the *pheA* gene that is transcribed convergently with the tyrosine operon on the *E. coli* chromosome (6). The detailed characterization of the *pheA* regulatory region has facilitated expression of the gene in a variety of transcriptional configurations, leading to elevated expression of CMPD, and a number of mutations in *pheA* that affect phenylalanine-mediated feedback inhibition have been described (9,33,49). Increased expression of the gene is readily achieved by cloning *pheA* onto multicopy plasmid vectors and deletion of the nucleotide sequences comprising the transcription attenuator (Fig. 2**a**). Most mutations that affect the allosteric regulation of the enzyme by phenylalanine have been identified through resistance to phenylalanine analogues such as β-2-thienylalanine, but there are examples of feedback-resistant mutations arising through insertional mutagenesis and gene truncation (33,49). Two regions of the enzyme in particular have been shown to reduce feedback inhibition to different degrees. Mutations at position Trp338 in the peptide sequence desensitize the enzyme to levels of phenylalanine in the 2 to 5 mM range but are insufficient to confer resistance to higher concentrations of L-phenylalanine (33,49). Mutations in the region of residues 304 to 310 confer almost total resistance to feedback inhibition at L-phenylalanine concentrations of at least 200 mM (9). Feedback inhibition profiles of four such variants (JN305–JN308) are shown in Figure 2**b**, in comparison to the profile of wild-type enzyme (JN302). It is not clear if the mechanism of resistance is similar in either case, but the difference is significant to commercial application because overproducing organisms readily achieve extracellular concentrations of L-phenylalanine over 200 mM.

Precursor Supply

Rate-limiting steps in the common aromatic and phenylalanine biosynthetic pathways are obvious targets in the development of phenylalanine overproducing organisms. However, the detailed biochemical and genetic characterization of *E. coli* has enabled additional areas of its metabolic function to be specifically manipulated to determine their effect on aromatic pathway throughput. Besides efforts to eliminate additional lesser points of aromatic pathway regulation, attempts have been made to enhance phenylalanine production by increasing the supply of aromatic pathway precursors and by facilitating exodus of L-phenylalanine from the cell. The precursors of the common aromatic pathway D-erythrose 4-phosphate and PEP are the respective products of the pentose phosphate and glycolytic pathways. Precursor supply in aromatic amino acid biosynthesis has been reviewed very recently (50,51). Theoretical analyses of the pathway and the cellular roles of these metabolites suggest that the production of aromatic compounds is likely to be limited by PEP availability (52,53) because PEP is involved in a number of cellular processes, including the generation of metabolic energy through the citric acid cycle (54) and the transport of glucose into the cell by the phosphotransferase system (55). Strategies to reduce the drain of PEP by these processes have included mutation of sugar transport systems to reduce PEP-dependent glucose transport (56) and modulation of the activities of pyruvate kinase, PEP synthase, and PEP carboxylase, which regulate PEP flux to pyruvate, oxaloacetate, and the citric acid cycle (50,54,57). Similarly, the availability of E4P has been increased by altering the levels of transketolase, the enzyme responsible for E4P biosynthesis (58). In general, these efforts have been successful in directing additional flux of PEP of E4P into the aromatic pathway, although their effect has not proved to be as predictable as the deregulation of rate-limiting

(a)

−35 −10

TGTATCGCCAACGCGCCTTCGGGCGCG**TTTTTTGTTGACAGCGTGAAAACAGTACGGGTA**CTG**TACT**

AAAGTCACTTAAGGAAACAAACATGAAACACATACCGTTTTTCTTCGCATTCTTTTTTACCTTCCCC

Attenuator Region

TGAATGGGAGGCGTTTCGTCGTGTGAAACAGAATGCGAAGACGAACAATAAGGCCTCCCAAATCGGG

GGGCCTTTTTTATTGATA**ACAAAAAGGCAACACTATGACATCGGAAAACCCGTTACTGGCGCT**

−35 −10

M T S E N P L L A L

GAATTCTTTTTTGTTGACAGCGTGAAAACAGTACGGGTATAATACTAAAGTCACAAAAAGGCAACACTATGACATCGGAAAACCCGTTACTGGCGCT

Coding sequence

(b)

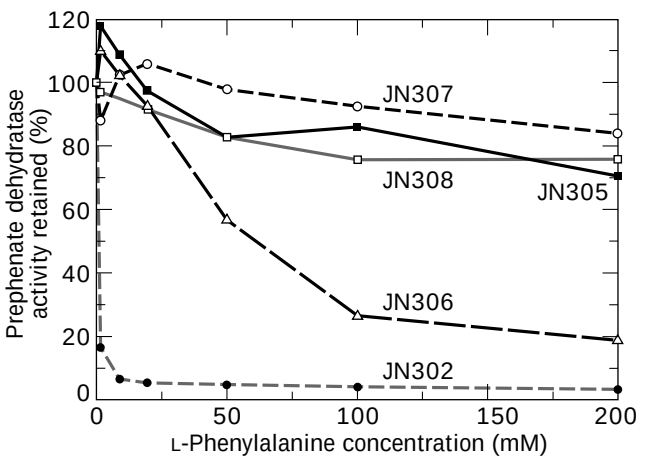


Figure 2. (**a**) The promoter region of the *E. coli* K12 *pheA* gene. The sequence of the wild-type promoter region is shown in the top part of the figure. The −35 and −10 hexamers are indicated. Bases retained in the deregulated promoter are in bold. The sequence involved in the attenuator region is shown underlined. The sequence of the deregulated *pheA* promoter region used to produce chorismate mutase/prephenate dehydratase (10) is shown in the bottom of the figure. Altered bases in the −10 region are shown underlined. (**b**) L-Phenylalanine-mediated feedback inhibition of wild-type *E. coli* K12 prephenate dehydratase (JN302) and four inhibition resistant enzyme variants (JN305-JN308). Activity is expressed as a percentage of wild-type enzyme activity.

pathway steps, and their overall impact on L-phenylalanine overproduction has not been well characterized.

Secretion of Phenylalanine from the Cell

Exodus of phenylalanine is of manifest importance in the large-scale production of phenylalanine because the amino acid is typically recovered only from the extracellular medium and washed cells. Lysis of cells to recover additional phenylalanine is not generally practical, and so L-phenylalanine remaining within the cells after washing is usually lost. Because cell biomass is extremely high in large-scale fermentations, this can represent up to 5% of total phenylalanine produced. Additionally, because L-phenylalanine is known to regulate its own biosynthesis at multiple points through feedback inhibition, attenuation, and TyrR-mediated repression, methods to reduce intracellular concentrations of phenylalanine through reduced uptake or increased exodus throughout the fermentation have been sought. Studies have addressed the means by which L-phenylalanine is both taken up and excreted by bacterial cells and whether this can be altered to increase efflux. In overproducing strains, it is likely that most phenylalanine leaves the cell by passive diffusion but specific uptake and exodus pathways also exist. In *E. coli*, L-phenylalanine is taken up by at least two permeases (59), one specific for phenylalanine encoded by *pheP* and the other encoded by *aroP* a general aromatic amino acid permease responsible for the transport of tyrosine and tryptophan in addition to phenylalanine. The genes encoding the permeases have been cloned and sequenced, and *E. coli* mutants deficient in either system have been characterized (60,61). The effect of an *aroP* mutation has been evaluated in L-phenylalanine overproduction (5). In addition, specific export systems have been identified that although not completely characterized have been successfully used to increase exodus of L-phenylalanine from strains of *E. coli* (62,63). In one such system, the Cin invertase of bacteriophage P1 has been shown to induce a metastable phenylalanine hypersecreting phenotype upon a wild-type strain of *E. coli*. The phenotype can be stabilized by transient introduction of the invertase on a temperature-sensitive plasmid replicon and is sufficient to establish L-phenylalanine overproduction in the absence of any alterations to the normal biosynthetic regulation (63).

Summary

The incremental gains made from the various levels at which phenylalanine biosynthesis has been addressed has led to the high-efficiency production strains currently in use today. Almost all large-scale L-phenylalanine manufacturing processes in present operation are fermentations using bacterial strains such as those described in this article in which classical strain development or molecular genetics have been extensively applied to bring about phenylalanine overproduction. The low substrate costs and the economics of scale associated with this approach have resulted in significant economic advantages. However, a number of alternate biotransformation and chemoenzymatic resolution strategies have been extensively developed that have also demonstrated commercially viable levels of L-phenylalanine synthesis.

BIOTRANSFORMATION ROUTES TO L-PHENYLALANINE

L-Phenylalanine Ammonia Lyase Process

The use of L-phenylalanine ammonia lyase (PAL) in various overproducing strains of the yeast *Rhodoturula* was developed by Tanabe Seiyaku Co. and by Genex Corporation in the 1980s to produce phenylalanine in the reaction shown in Figure 3. In this process, fermented cells of yeast strains *R. glutinis, R. rubra* or *R. graminis* (64–67) were recovered and washed before bioreaction under batch or immobilized conditions with *trans*-cinnamic acid and ammonia. The elevated pH and high concentration of ammonia used under the process conditions enabled the normally catabolic PAL reaction (68,69) to favor formation of L-phenylalanine. Significant development work was carried out to determine the appropriate conditions for optimal growth and maximal induction of the PAL enzyme (70) and the appropriate process conditions for retention of PAL activity (71). The principal drawbacks of the PAL process lay in substrate inhibition of the enzyme, low enzyme-specific activity, and enzyme instability. Screening regimes carried out to address these limitations resulted in the identification of an *R. graminis* strain that displayed greater enzyme specific activity and stability than the *R. rubra* and *R. glutinis* strains used previously (64). Subsequent mutagenesis and selection procedures carried out on this strain using the cinnamic acid analogue phenylpropiolic acid led to a further increase in PAL activity through a significant elevation in PAL gene expression (64). Such strain and process development strategies enabled the Genex process to achieve yields of L-phenylalanine in excess of 50 g/L, with close to 90% conversion of substrate to product (67). Nevertheless, the biotransformation could not favorably compete with the economics of fermentative production, and it is not presently used in L-phenylalanine production.

Aspartate or Aromatic Aminotransferase Process

A number of reports from PEI, Nutrasweet, and Allelix, among others, have described the use of aminotransferases (transaminases) to produce L-phenylalanine in the reaction shown in Figure 4. The reaction proceeds by a ping-pong (72,73) mechanism whereby the amino group of a donor amino acid, typically aspartic acid, is exchanged with the keto group of phenylpyruvic acid, yielding L-phenylalanine and the by-product oxaloacetate. Aminotransferases are ubiquitous in nature, and microorganisms typically possess multiple aminotransferases that are involved in the biosynthesis of a number of amino acids, often catalyzing the terminal step from the keto acid precursor (73–75). The enzymes use pyridoxal 5′-phosphate as cofactor and typically possess high catalytic rates and stereoselectivity with broad substrate specificities (76).

This is exemplified by the multiple transaminases in *E. coli* that can each catalyze the synthesis of L-tyrosine, L-aspartate, and L-valine and L-leucine in addition to L-

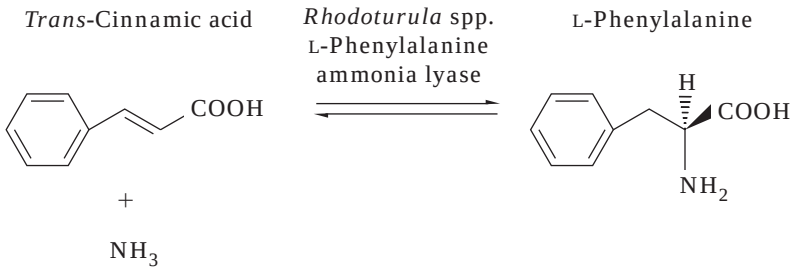


Figure 3. L-Phenylalanine-ammonia-lyase-catalyzed synthesis of L-phenylalanine from *trans*-cinnamic acid.

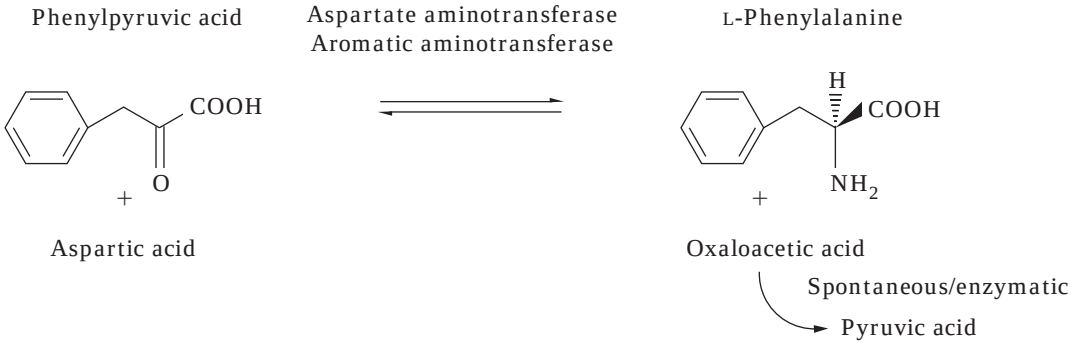


Figure 4. Aminotransferase-catalyzed synthesis of L-phenylalanine from phenylpyruvic acid using aspartic acid as the amino donor. Decarboxylation of oxaloacetic acid drives the reaction toward completion.

phenylalanine (37). Many aminotransferase encoding genes have been cloned and used to overproduce the enzymes (37,77,78). Significant sequence homology often exists among the aminotransferases of individual strains as well as between the corresponding enzymes of microbial and mammalian origin (37). In addition, the tertiary structures of a number of aminotransferases have been determined, furthering the mechanistic understanding of the enzyme and prompting mutagenesis approaches to altering enzyme properties (79). Aminotransferases from many microorganisms have been applied in L-phenylalanine biosynthesis (80–82). In the case of *E. coli*, the highly homologous aspartate and aromatic aminotransferases encoded respectively by the *aspC* and *tyrB* genes have both been developed and applied in this bioconversion (83–85). The aspartate aminotransferase has been generally preferred, having shown greater throughput and increased stability over the *tyrB* gene product (85,86). A number of L-phenylalanine biosynthetic processes involving aminotransferases have been described using either isolated enzyme or whole cells in batched or immobilized systems (82,85,87,88). In the mid 1980s, one such immobilized process was commercialized by PEI using the *E. coli* aspartate aminotransferase (85). In this case, the PPA substrate was economically derived from chemically synthesized 5-benzilidene hydantoin. The highest yields in the transaminase synthesis are generally obtained from batch processes at the expense of the greater catalyst stability of the continuous systems. The cost of catalyst regeneration in batch systems is a drawback of the aminotransferase approach, as is the need for two primary metabolites as substrates. The principal drawback, however, lies in the reaction equilibrium, which being close to 1 limits the yield of the reaction to 50% unless the keto acid by-product can be eliminated from further reaction. To overcome this limitation, commercial transamination approaches have almost invariably used aspartate as the amino donor, which upon transamination to the relatively unstable oxaloacetate will spontaneously decarboxylate to yield pyruvate. Because pyruvate is frequently not a substrate for these transaminases, this enables the reaction to proceed to a yield often exceeding 90%. The decarboxylation of oxaloacetate is more rapid at elevated temperature and has prompted the development of transaminase bioconversions carried out in excess of 50 °C (86). Aminotransferases from thermophiles have been explored in this respect as well as chemical and enzymatic steps to accelerate oxaloacetate decarboxylation and thereby optimize product yield and substrate conversion (89).

Phenylalanine Dehydrogenase Process

A similar process using the phenyl pyruvate precursor has also been developed using the enzyme phenylalanine dehydrogenase. Unlike the aminotransferases, phenylalanine dehydrogenase uses ammonia as a substrate in the reductive amination of phenyl pyruvate, thereby realizing a cost advantage. In addition, the reaction equilibrium lies heavily in favor of phenylalanine synthesis. These advantages in the reaction mechanism are diminished by the dependence of the dehydrogenase on an NADH cofactor that in turn requires enzymatic regeneration. This has been ad-

dressed through the use of a coupled enzyme system (90) in which formate dehydrogenase is used along with phenylalanine dehydrogenase in the reaction scheme shown in Figure 5. The reaction proceeds quantitatively because of the evolution of CO_2 generated in the oxidation of formate. Using partially purified enzymes in membrane reactors, this reaction has achieved phenylalanine titers and substrate conversions comparable to the aminotransferase based systems (90,91). Despite the high efficiencies (92) obtained in this bioconversion, commercial production of L-phenylalanine using this method has not been economically competitive with fermentative processes, mainly because of the costs of phenylpyruvic acid preparation. Nevertheless, in a similar process operated by Degussa using leucine dehydrogenase, this approach has been successfully applied to the synthesis of compounds of higher intrinsic value such as the unnatural amino acid tertiary leucine (93).

RESOLUTION-BASED L-PHENYLALANINE SYNTHESIS

Amino Acid Amide Resolution

In contrast to the various fermentation and bioconversion approaches that have used enzyme stereoselectivity to synthesize L-phenylalanine as a single isomer, additional methods have been developed that rely on the same stereoselectivity to resolve chemically synthesized amino acid DL-racemic mixtures. One such general approach that has been successfully commercialized in L-phenylalanine production relies on cleavage of the L-isomer of a DL-α-amino acid amide mixture by a highly enantiospecific amidase enzyme. The general procedure operated by DSM and shown in Figure 6 uses an amidase containing strain of *Pseudomonas putida* to specifically hydrolyze the L-isomer of the DL-amino acid amide mixture. The racemic amino acid mixture is prepared inexpensively from an aldehyde precursor (94,95) via the Strecker reaction, with alkaline hydrolysis of the resulting aminonitrile. Treatment with benzaldehyde after selective hydrolysis of the L-amino acid amide yields an insoluble Schiff base that forms between benzaldehyde and the residual D-amino acid amide. In this way, the residual D-amino acid amide may be readily separated (96) from the resulting L-amino acid and can then be

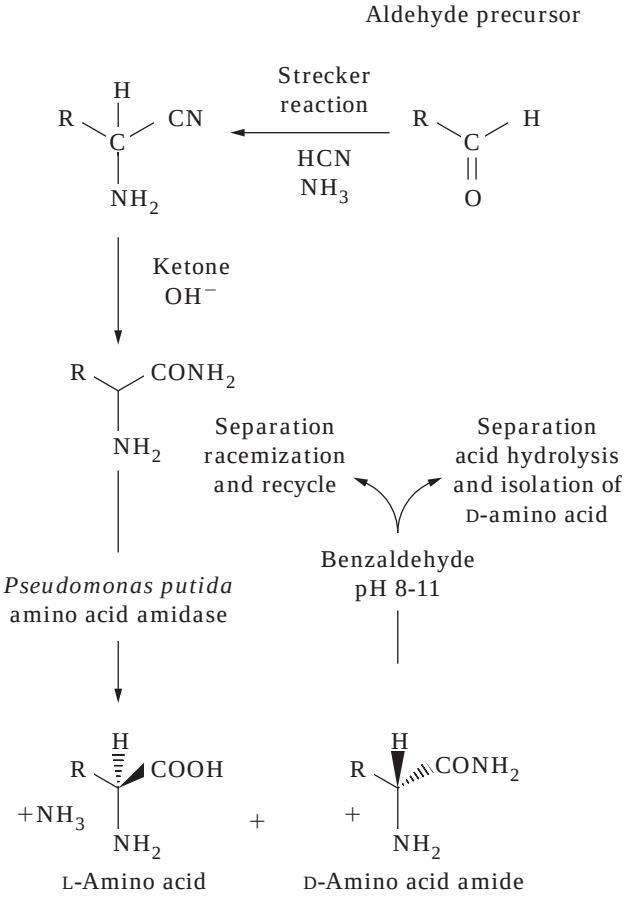


Figure 6. The DSM process for the manufacture of L- and D-α-amino acids.

chemically hydrolyzed and purified or racemized and recycled (97). The advantages of this process include general applicability to the preparation of a range of amino acids (95,98) and the opportunity to recover the D-amino acid if desired from the reaction in addition to the L-amino acid. Conversely, in the preparation of only the L-amino acid by this approach, the need to recycle the unreacted substrate is considered a drawback of the process, necessitating ad-

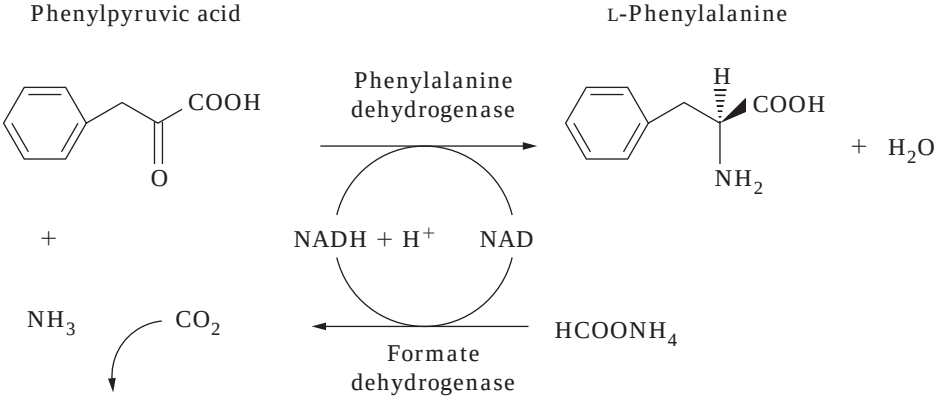


Figure 5. Reductive amination of phenylpyruvic acid to L-phenylalanine using L-phenylalanine dehydrogenase with cofactor regeneration by coupled formate dehydrogenase.

ditional complexity and cost. This aspect has been considered in a further development of this process by DSM in which *P. putida* strains possessing D- or L-amino acid amidase as well as amino acid racemase activity have been identified. The in situ coupling of either amidase with the racemase activity avoids the recycling requirement of the process and facilitates direct conversion of the DL-amino acid amide mix into the desired isomer (99).

Acylase Process

An earlier strategy involving stereoselective enzyme hydrolysis of chemically prepared racemic precursors has been operated successfully for many years by Tanabe Seiyaku. In this process, illustrated in Figure 7, a DEAE-Sephadex immobilized L-specific aminoacylase selectively hydrolyzes *N*-acetyl-L-amino acids from an *N*-acetyl-DL-amino acid mixture (100,101). After hydrolysis, the L-amino acid is isolated by selective crystallization and the remaining *N*-acetyl-D-amino acid is racemized either chemically or enzymatically (102) and recycled. The preferred aminoacylase for this reaction was obtained from the mold *Aspergillus oryzae* and is readily isolated and immobilized in the bioreactor (100). The enzyme has broad substrate specificity and has been used by Tanabe in the production of L-methionine and L-valine and addition to L-phenylalanine. More recently, this process has been further developed by Degussa/Rexim (103) to prepare bulk quantities of L-phenylalanine and a large number of L-amino acids. In addition, Tanabe has reported on the further development of the aminoacylase process through the coupling of acetamidocinnamate aminohydrolase and aminotransferase activity. In this process, L-phenylalanine is produced in two steps from acetamidocinnamic acid through a phenylpyruvic acid intermediate using immobilized cells of two independent bacterial strains (87).

DL-5-Monosubstituted Hydantoin-Based Strategies

The recycling aspect of resolution-based amino acid syntheses such as the amidase and aminoacylase processes is again considered and largely overcome in an alternative strategy using DL-5-monosubstituted hydantoins. Inexpensive chemical routes such as the Bucherer-Bergs synthesis or the condensation of aldehydes with hydantoin (104–107) are used to prepare the racemic hydantoin derivatives that are then resolved using either L- or D-enantiospecific hydantoinase and *N*-carbamoyl amino acid amidohydrolase enzymes to yield the desired amino acid isomer. The process is shown in Figure 8. In the case of phenylalanine, both L- and D-phenylalanine can be prepared from racemic DL-5-benzylhydantoin. The advantage of the hydantoinase process is that in many cases a dynamic resolution can be established whereby the unreacted isomer can be racemized in situ either chemically (108) or in the presence of a racemase (109) enzyme to enable the process to approach 100% yield. Enzymes with L- and D-specific hydantoinase activity have been described in a number of microbial species, with the D-specific activity having been identified much more frequently, often attributable to dihydropyrimidinases involved in pyrimidine catabolism (110–114). Substrate specificities vary significantly between enzyme isolates and contribute to the broad applicability of this approach (113–116). Molecular cloning has been applied in at least one case to isolate a D-hydantoinase encoding gene from a thermophilic bacillus and to subsequently overproduce the enzyme (117). The D- and L-specific *N*-carbamoyl amino acid amidohydrolases have also been shown to be widely distributed in microorganisms (113,114,116,118). The bioconversion steps are typically carried out using whole cells in immobilized systems (110,118) with racemization of the unreacted hydantoin occurring under the normal alkaline conditions of the reaction (108). In some cases, microbial racemases have been used to optimize the racemization step (109,119). The hydantoinase process offers the advantage of relatively inexpensive precursors and high conversion rates through the in situ racemization of the substrate. In addition, through the broad substrate range of hydantoinase and carbamoylase enzymes, the process is broadly applicable in L- and D-amino acid biosynthesis. However, it is in D-amino acid production that large-scale commercial application of the hydantoinase process has been more significantly applied, most notably by Recordati and Kanegafuchi in the manufacture of D-phenyl glycine and D-*p*-hydroxyphenyl glycine as components of semisynthetic penicillins and cephalosporins (120–122) but also by Degussa and Ajinomoto in the production of additional D-amino acids, including D-phenylalanine (122–124).

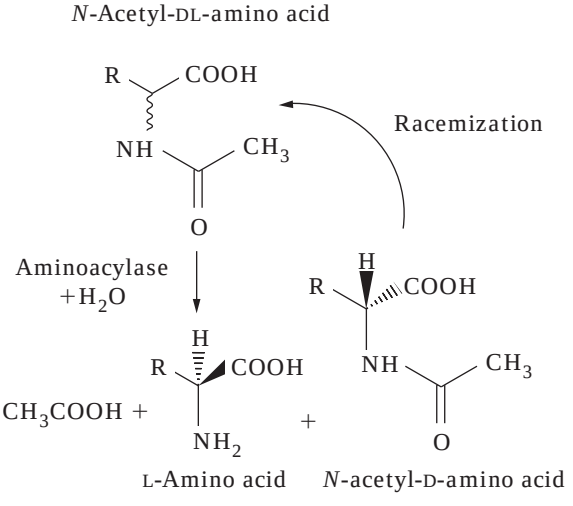


Figure 7. Resolution of *N*-acetyl-DL-amino acids using aminopeptidases yielding L-amino acid and recycling of residual *N*-acyl-D-amino acid.

CONCLUSIONS AND FUTURE PROSPECTS

Numerous commercial processes have been developed for the large-scale commercial production of L-phenylalanine since the early 1980s, fueled by the enormous increase in L-phenylalanine demand for the dipeptide sweetener aspartame. The most successful processes use microbial enzymes at some stage, either in an enantioselective hydrolysis or biotransformation step using chemically derived precursors or in whole cell fermentations using microorganisms engineered through classical and recombinant

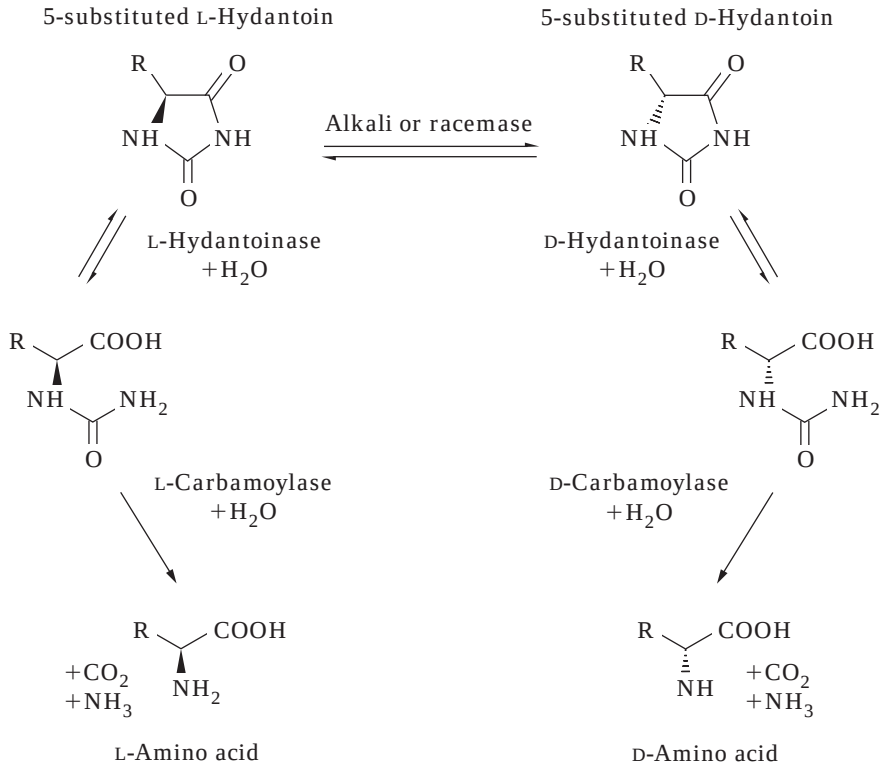


Figure 8. Hydantoinase/carbamoylase-based dynamic resolution of DL-5-monosubstituted hydantoins to either D- or L-amino acids.

strain improvement methodology. In general, the most economically competitive processes have resulted from large-scale fermentation approaches. This is mainly because of the incremental savings achieved in such processes through increased understanding of the physiology and molecular genetic regulation of L-phenylalanine biosynthesis in microbes and the low raw material costs attained through media and process development. In these processes, the concentration of L-phenylalanine can often exceed 50 g/L in crude fermentation broth with the overall incorporation of carbon into phenylalanine from a source such as glucose, equaling or exceeding the theoretical maximum calculated at around 25 to 30% (125). With this level of efficiency, it is unlikely that most major manufacturers of L-phenylalanine will continue to invest significantly in the development of fermentation processes as the return on investment is diminished.

Several elegant biotransformation approaches such as the PAL and dehydrogenase processes have been optimized through molecular genetic strategies but have been impacted by the drawbacks of greater precursor cost and in some cases limited biocatalyst stability. As a result, none of these approaches are currently used to produce L-phenylalanine. Nevertheless, a number of manufacturers, including DSM, Degussa/Rexim, and Tanabe, successfully produce L-phenylalanine through chemoenzymatic routes that are generally more versatile and are often part of more broadly applicable processes in natural and unnatural amino acid manufacture. This is apparent in the manufacture of D-phenylalanine where, in contrast to the preponderance of fermentative routes to L-phenylalanine, the synthesis is currently carried out using the hydantoinase and amidase resolution strategies mentioned earlier. Ultimately, through a greater understanding of microbial pathway engineering and the increasing availability of novel enzymatic activities, it is quite possible that the favorable economics of large-scale fermentation may also become more applicable to D-phenylalanine biosynthesis (126,127).

BIBLIOGRAPHY

1. Eur. Pat. 0264,914 (1987), R. Katsumata, and A. Ozaki.
2. I. Shiio, S. Sugimoto, and K. Kawamura, *Agric. Biol. Chem.* **52**, 2247–2253 (1988).
3. H. Hagino and K. Nakayama, *Agric. Biol. Chem.* **38**, 157–161 (1974).
4. M. Ikeda, A. Ozaki, and R. Katsumata, *Appl. Microbiol. Biotechnol.* **39**, 318–323 (1993).
5. U.S. Pat. 4,681,852, D.E. Tribe.
6. G.S. Hudson and B.E. Davidson, *J. Mol. Biol.* **180**, 1023–1051 (1984).
7. A. Ozaki, R. Katsumata, T. Oka, and A. Furuya, *Agric. Biol. Chem.* **49**, 2925–2930 (1985).
8. I. Shiio and S. Sugimoto, *Agric. Biol. Chem.* **45**, 2197–2207 (1981).
9. J. Nelms, R.M. Edwards, J. Warwick, and I. Fotheringham, *Appl. Environ. Microbiol.* **58**, 2592–2598 (1992).
10. G. Millar, A. Lewendon, M. Hunter, and J.R. Coggins, *Biochem. J.* **237**, 427–437 (1986).
11. I. Shiio, K. Ishii, K. Yokozeki, *Agric. Biol. Chem.* **37**, 1991 (1973).

12. T. Tsuchida, K. Kubota, Y. Morinaga, H. Matsui, H. Enei, and F. Yoshinaga, *Agric. Biol. Chem.* **51**, 2095–2101 (1987).
13. Y. Tokoro, K. Oshima, M. Okii, K. Yamaguchi, K. Tanaka, and S. Kinoshita, *Agric. Biol. Chem.* **34**, 1516 (1970).
14. S.O. Hwang, G.H. Gil, Y.J. Cho, K.R. Kang, J.H. Lee, and J.C. Bae, *Appl. Microbiol. Biotechnol.* **22**, 108–113 (1985).
15. U.S. Pat. 3,759,790 (1973), K. Nakayama, Sagamihara, and H. Hagino.
16. B.J. Wallace and J. Pittard, *J. Bacteriol.* **97**, 1234–1241 (1969).
17. Y.J. Choi and D.E. Tribe, *Biotechnol. Lett.* **4**, 223–228 (1982).
18. E. Haslam, in *The Shikimate Pathway*, Wiley, New York, 1974.
19. F. Gibson and J. Pittard, *Curr. Topics Cell Regul.* **2**, 29–63 (1970).
20. K.D. Brown, *Genetics* **60**, 31–48 (1968).
21. K.D. Brown and R.L. Somerville, *J. Bacteriol.* **108**, 386–399 (1971).
22. H. Camakaris and J. Pittard, *J. Bacteriol.* **115**, 1135–1144 (1973).
23. S.W.K. Im, H. Davidson, and J. Pittard, *J. Bacteriol.* **108**, 400–409 (1971).
24. K.M. Herrmann and R.L. Somerville, eds., *Amino Acids: Biosynthesis and Genetic Regulation*, Addison-Wesley, Reading, Mass., 1983.
25. I. Shiio, S. Sugimoto, and J. Miyajima, *J. Biochem.* **75**, 987 (1974).
26. S. Sugimoto and I. Shiio, *J. Biochem.* **87**, 881–890 (1980).
27. H. Hagino and K. Nakayama, *Agric. Biol. Chem.* **39**, 351 (1975).
28. H. Hagino and K. Nakayama, *Agric. Biol. Chem.* **38**, 2125–2134 (1974).
29. I. Shiio and S. Sugimoto, *J. Biochem.* **86**, 17–25 (1979).
30. S. Sugimoto, M. Nakagawa, T. Tsuchida, and I. Shiio, *Agric. Biol. Chem.* **37**, 2327–2336 (1973).
31. W.D. Davies and B.E. Davidson, *Nucleic Acids Res.* **10**, 4045–4058 (1982).
32. J.M. Ray, C. Yanofsky, and R. Bauerle, *J. Bacteriol.* **170**, 5500–5506 (1988).
33. Eur. Pat. 229,161 (1986), R.M. Edwards, P.P. Taylor, M.G. Hunter, and I.G. Fotheringham.
34. C. Förberg, T. Eliaeson, and L. Häggström, *J. Biotechnol.* **7**, 319–332 (1988).
35. S. Sugimoto, M. Yabuta, N. Kato, T. Seki, T. Yoshida, and H. Taguchi, *J. Biotechnol.* **5**, 237–253 (1987).
36. J.T. Powell and J.F. Morrison, *Eur. J. Biochem.* **87**, 391–400 (1978).
37. I.G. Fotheringham, S.A. Dacey, P.P. Taylor, T.J. Smith, M.G. Hunter, M.E. Finlay, S.B. Primrose, D.M. Parker, and R.M. Edwards, *Biochem. J.* **234**, 593–604 (1986).
38. M-J. Gething and B.E. Davidson, *Eur. J. Biochem.* **86**, 165–174 (1978).
39. T.A.A. Dopheide, P. Crewther, and B.E. Davidson, *J. Biol. Chem.* **247**, 4447–4452 (1972).
40. I. Shiio and S. Sugimoto, *J. Biochem.* **79**, 173–183 (1976).
41. S. Sugimoto and I. Shiio, *J. Biochem.* **76**, 1103 (1974).
42. I. Shiio and S. Sugimoto, *J. Biochem* **89**, 1483–1492 (1981).
43. S. Sugimoto and I. Shiio, *Agric. Biol. Chem.* **49**, 39–48 (1985).
44. I. Shiio, M. Mori, and H. Ozaki, *Agric. Biol. Chem.* **46**, 2967–2977 (1982).
45. L. de Boer and L. Dijkhuizen, *Adv. Biochem. Eng. Biotechnol.*, **41**, 1–27 (1990).
46. H. Hagino and K. Nakayama, *Agric. Biol. Chem.* **38**, 2367–2376 (1974).
47. H. Hagino and K. Nakayama, *Agric. Biol. Chem.* **39**, 331–342 (1975).
48. M. Ikeda and R. Katsumata, *Appl. Environ. Microbiol.* **58**, 781–785 (1992).
49. U.S. Pat. 4,753,883 (1988), K.C. Backman and B. Ramaswamy.
50. A. Berry, *TIBTECH* **14**, 250–256 (1996).
51. J.W. Frost and K.M. Draths, *Annu. Rev. Microbiol.* **49**, 557–579 (1995).
52. R. Patnaik and J.C. Liao, *Appl. Environ. Microbiol.* **60**, 3903–3908 (1994).
53. R. Patnaik, R.G. Spitzer, and J.C. Liao, *Biotech. Bioeng.* **46**, 361–370 (1995).
54. J.E. Miller, K.C. Backman, M.J. O, Connor, and R.T. Hatch, *J. Ind. Microbiol.* **2**, 143–149 (1987).
55. P.W. Postma, in F.C. Neidhardt ed., Escherichia coli and Salmonella typhimurium: *Cellular and Molecular Biology*. American Society for Microbiology, Washington, D.C., 1987.
56. N. Flores, J. Xiao, A. Berry, F. Bolivar, and F. Valle, *Nat. Biotechnol.* **14**, 620–623 (1996).
57. S. Bledig, Ph.D. Thesis, Univ. of Warwick, 1994.
58. K.M. Draths, D.L. Pompliano, D.L. Conley, J.W. Frost, A. Berry, G.L. Disbrow, R.J. Staversky, and J.C. Lievense, *J. Am. Chem. Soc.* **114**, 3956–3962 (1992).
59. M.J. Whipp, and A.J. Pittard, *J. Bacteriol.* **132**, 453–461 (1977).
60. J. Pi, P.J. Wookey, A.J. Pittard, *J. Bacteriol.* **173**, 3622–3629 (1991).
61. N. Honore and S.T. Cole, *Nucleic Acids Res.* **18**, 653 (1990).
62. M. Mulcahy, Ph.D. Thesis, Univ. of Dundee, 1988.
63. U.S. Pat. 5,354,672 (1994), I.G. Fotheringham, J. Ton, and C. Higgins.
64. S.A. Orndorff, N. Constantino, D. Stewart, and D.R. Durham, *Appl. Env. Microbiol.* **54**, 996–1002 (1988).
65. U.S. Pat. 4,584,269 (1986), P.J. Vollmer, J.J. Schruben, J.P. Montgomery, and Y. Huei-Hsuing.
66. S. Yamada, K. Nabe, N. Izuo, K. Nakamichi, and I. Chibata, *Appl. Environ. Microbiol.* **42**, 773–778 (1981).
67. B.K. Hamilton, H.Y. Hsiao, W.E. Swann, D.M. Anderson, and J.J. Delente, *Trends Biotechnol.* **3**, 64 (1985).
68. H.J. Gilbert and M. Tully, *J. Bacteriol.* **150**, 498–505 (1982).
69. N. Onishi, K. Yokozeki, Y. Hirose, and K. Kubot, *Agric. Biol. Chem.* **51**, 291–292 (1987).
70. C.T. Evans, K. Hanna, D. Conrad, W. Peterson, and M. Misawa, *Appl. Microbiol. Biotechnol.* **25**, 406–414 (1987).

71. C.T. Evans, D. Conrad, K. Hanna, W. Peterson, C. Choma, and M. Misawa, *Appl. Microbiol. Biotechnol.* **25**, 399–405 (1987).
72. J.F. Kirsch, G. Eichele, G.C. Ford, M.G. Vincent, and J.N. Jansonius, *J. Mol. Biol.* **174**, 497–525 (1983)
73. R.A. Jensen and D.H. Calhoun, *CRC Crit. Rev. Microbiol.* **8**, 229–266 (1981).
74. T. Tachiki and T. Tochikura, *Agric. Biol. Chem.* **37**, 1439–1448 (1973).
75. T. Tachiki, M. Moriguchi, and T. Tochikura, *Agric. Biol. Chem.* **39**, 43–50 (1973).
76. P. Christen and D.E. Metzler eds., *Transaminases*, Wiley, New York, 1984.
77. S. Kuramitsu, T. Ogawa, H. Ogawa, and H.J. Kagamiyama, *Biochemistry* **97**, 993–999 (1985).
78. M.-H. Sung, K. Tanizawa, H. Tanaka, S. Kuramitsu, H. Kagamiyama, K. Hirotsu, A. Okamoto, T. Higuchi, and K. Soda, *J. Biol. Chem.* **266**, 2567–2572 (1991).
79. E. Kohler, M. Seville, J. Jager, I. Fotheringham, M. Hunter, M. Edwards, J.N. Jansonius, and K. Kirschner, *Biochemistry* **33**, 90–97 (1994).
80. E. Bulot and C.L. Cooney, *Biotechnol. Lett.* **7**, 93 (1985).
81. C.T. Evans, W. Peterson, C. Choma, and M. Misawa, *Appl. Microbiol. Biotechnol.* **26**, 305–312 (1987).
82. K. Nakamichi, K. Nabe, Y. Nishida, and T. Tosa, *Appl. Microbiol. Biotechnol.* **30**, 243–246 (1989).
83. J. Then, A. Doherty, H. Neatherway, R. Marquardt, H.M. Deger, H. Voelskow, G. Wöhner, and P. Präve, *Biotechnol. Lett.* **9**, 680 (1987).
84. M. Robinson, R. Marquardt, and J. Then, J. McChesney, H. Neatherway, G. Wöhner, H.M. Deger, and P. Präve, *Biotechnol. Lett.* **9**, 673–679 (1987).
85. G.J. Calton, L.L. Wood, M.H. Updike, L. Lantz, J.P. Hamman, *Biotechnology* **4**, 317–320 (1986).
86. Eur. Pat. 0152,275 (1991), D. Lewis and S. Farrand.
87. K. Nakamichi, Y. Nishida, K. Nabe, and T. Tosa, *Appl. Biochem. Biotechnol.* **11**, 367–376 (1985).
88. H. Ziehr, M.R. Kula, E. Schmidt, C. Wandrey, and J. Klein, *Biotechnol. Bioeng.* **9**, 482 (1987).
89. I. Schutten, W. Harder, and L. Dijkhuizen, *Appl. Microbiol. Biotechnol.* **27**, 292 (1987).
90. W. Hummel, E. Schmidt, C. Wandrey, and M.R. Kula, *Appl. Microbiol. Biotechnol.* **25**, 175–185 (1986).
91. W. Hummel, H. Schutte, E. Schmidt, C. Wandrey, and M.R. Kula, *Appl. Microbiol. Biotechnol.* **26**, 409–416 (1987).
92. C. Wandrey, in O.M. Neijssel, R.R. van der Meer and K.Ch.A.M. Luyben eds., *Proc. of the 4th European Congress on Biotechnology, vol. 4*, Elsevier, Amsterdam, 1987, pp. 171–188.
93. U. Kragl, D. Vasic-Racki, and C. Wandrey, *Chem. Ing. Tech.* **64**, 499 (1992).
94. J. Kamphuis, W.H.J. Boesten, Q.B. Broxterman, H.F.M. Hermes, J.A.M. van Balken, E.M. Meijer, and H.E. Schoemaker, *Adv. Biochem. Eng. Biotechnol.* **42**, 134 (1990).
95. E.M. Meijer, W.H.J. Boesten, H.E. Schoemaker, and J.A.M. van Balken, in J. Tramper, H.C. van der Plas, and P. Linko eds., *Biocatalysts in Organic Synthesis*, Elsevier, Amsterdam, 1985, pp. 135–156.
96. U.S. Pat. 4,172,846 (1979), DSM/Stamicarbon.
97. Eur. Pat. 1,442,585, (1991), DSM/Stamicarbon.
98. E.M. Meijer, J. Kamphuis, J.A.M. Van Balken, H.F.M. Hermes, W.J.J. van den Tweel, M. Kloosterman, W.H.J. Boesten, H.E. Schoemaker, in V. Claassen ed., *Trends in Drug Research*, Elsevier, Amsterdam, 1990, pp. 363–382.
99. Eur. Pat. 0307023 (1989), Hermes HFM, et al.
100. T. Tosa, T. Mori, N. Fuse, and I. Chibata, *Enzymologia* **31**, 214 (1966).
101. T. Tosa, T. Mori, N. Fuse, and I. Chibata, *Agric. Biol. Chem.* **33**, 1047 (1969).
102. Eur. Pat. 304,021 (1989), Takeda Chemical Industries.
103. U.S. Pat. 4,304,858 (1981), Degussa/GBF.
104. E. Ware, *Chem. Rev.* **46**, 403 (1950).
105. Jpn. Pat. 80 136279 (1979), Ajinomoto.
106. German Pat. DE 3043259 A1 (1980), A. Kleeman and M. Samson.
107. K. Drauz, A. Kleeman, and M. Samson, *Chem.-Ztg.* **12**, 391 (1984).
108. M. Bovarnick and H.T. Clark, *J. Am. Chem. Soc.* **60**, 2426 (1938).
109. C. Gross, C. Syldatk, V. Kackowiak, and F. Wagner, *J. Biotechnol.* **14**, 363 (1990).
110. R. Olivieri, E. Fascetti, L. Angelini, and L. Degen, *Biotechnol. Bioeng.* **23**, 2173–2183 (1981).
111. K. Yokozeke, S. Nakamori, C. Eguchi, K. Yamada, and K. Mitsugi, *Agric. Biol. Chem.* **51**, 355–362 (1987).
112. A. Möller, C. Syldatk, M. Schultze, and F. Wagner, *Enzyme Microb. Technol.* **10**, 618 (1988).
113. K. Yokozeke, Y. Hirose, and K. Kubota, *Agric. Biol. Chem.* **51**, 737–746 (1987).
114. A. Yamashiro, K. Kubota, and K. Yokozeki, *Agric. Biol. Chem.* **52**, 2857–2863 (1988).
115. H. Yamada, S. Takahashi, Y. Kii, and H. Kumagai, *J. Ferment. Technol.* **56**, 484 (1978).
116. K. Yokozeke and K. Kubota, *Agric. Biol. Chem.* **51**, 721–728 (1987).
117. Y. Mukohara, T. Ishikawa, K. Watabe, and H. Nakamura, *Biosci. Biotechnol. Biochem.* **58**, 1621–1626 (1994).
118. R. Olivieri, E. Fascetti, L. Angelini, and L. Degen, *Enzyme Microbiol. Technol.* **1**, 201 (1979).
119. C. Syldatk and F. Wagner, *Food Biotechnol.* **4**, 87 (1990).
120. H. Yamada and S. Shimizu, in J. Tramper, H.C. van der Plas, P. Linko eds., *Biocatalysts in Organic Syntheses*, Elsevier, Amsterdam, 1985, pp. 19–37.
121. Eur. Pat. 0175312 (1985), Kanegafuchi.
122. K. Yokozeke, S. Nakamori, S. Yamanaka, C. Eguchi, K. Mitsugi, and F. Yoshinaga, *Agric. Biol. Chem.* **51**, 715–719 (1987).
123. German Pat. DE 3917057 (1992), Degussa.
124. Jpn. Pat. JP 86-200434 (1986), T. Ishikawa and H. Kimura.
125. C. Foerberg, T. Eliaeson, and L. Haggstrom, *J. Biotechnology*, **7** 319–332 (1988).
126. I. Fotheringham, S. Bledig, R. Senkpeil, and P. Taylor, *Abstracts of the 1995 SIM Annual Meeting*, San Jose, Calif., 1995.
127. J. Dailey, N. Grinter, R. Nelson, D. Pantaleone, R. Senkpeil, P. Taylor, J. Ton, R. Yoshida, and I. Fotheringham, *Abstracts of the IBC Enzyme Technology Symposium*, Lake Buena Vista, Fla., 1996.

PHENYLALANINE DEHYDROGENASE

YASUHISA ASANO
Toyama Prefectural University
Toyama, Japan

KEY WORDS

L-Amino acid
Bacillus badius
Bacillus sphaericus
Enantioselective synthesis
α-Keto acid
Microdetermination
NADH regeneration
Phenylalanine dehydrogenase
Phenylketonuria
Sporosarcina ureae

OUTLINE

INTRODUCTION

NAD^+-dependent phenylalanine dehydrogenase (PheDH, phenylalanine: NAD^+ oxidoreductase, deaminating (EC 1.4.1.20)) was found to occur in *Brevibacterium* sp. by Hummel et al. in 1984 (1). Asano et al. first described the enzymological properties of PheDHs with crystalline enzymes from *Soprosarcina ureae* and *Bacillus sphaericus* (2,3). Similar enzymes that catalyze a reversible oxidation-reduction of amino acids (acting on -CH-$NH_{(2)}$) with NAD^+ or $NADP^+$ as an electron acceptor are classified in EC 1.4.1., which includes glutamate dehydrogenase (GluDH, EC 1.4.1.2-4) (4), alanine dehydrogenase (AlaDH, EC 1.4.1.1), leucine dehydrogenase (LeuDH, EC 1.4.1.9), and other oxidoreductases acting on several amino acids (5).

Since its discovery (1,2,3), much attention has been paid to PheDH because it appeared to be a useful catalyst in the enantioselective synthesis of Phe and related L-amino acids from their keto analogs (6). Some microbial PheDH producers have been isolated from nature, and the enzymes have been characterized. Since 1992, PheDH has been used as a reagent in the colorimetric microdetermination of Phe to detect phenylketonuria (PKU) in the blood of neonates in Japan (7,8).

Here we deal with the occurrence, properties, and structures of bacterial PheDHs, and their applications to the synthesis and determination of amino acids.

DISCOVERY, ENZYMATIC PROPERTIES, AND FUNCTIONS OF PheDH

Hummel et al. screened for a PheDH-producing strain of bacteria from soil and isolated a *Brevibacterium* sp. by an enrichment culture technique (1). Culture conditions for PheDH formation and some reaction conditions to synthesize Phe and Tyr with the crude enzyme were optimized. Asano et al. noted PheDH production in the gram-positive spore-forming bacteria *S. ureae* SCRC-R04 (2,3), *B. sphaericus* SCRC-R79a (3), and *B. badius* IAM 11059 (7) and described its purification, crystallization, and characterization; no activity was found in yeasts. Following these studies, PheDH was given a new entry number: EC 1.4.1.20 (5). Later, *Rhodococcus maris* (9), *Nocardia* sp. (10), *Thermoactinomyces intermedius* (11), *Rhodococcus* sp. (12), *Bacillus cereus* (13), and *Microbacterium* sp. (14,15) were found to produce the enzyme. The occurrence of PheDH appears to be limited in some groups of gram-positive spore-forming bacteria and thermophilic Actinomycetes (11). The distribution of PheDH seems similar to that of LeuDH (16), although the distribution of the former is much more limited than that of the latter, which may be implicated in microbial sporulation, which shuffles carbon and nitrogen metabolism of amino acids by their reversible nature. Because PheDH is specifically distributed in *B. sphaericus* and *B. badius*, which are characterized to be poor in sugar oxidation among the genus *Bacillus* (17), PheDH seems involved in the catabolism of Phe and other amino acids in the medium.

Hummel et al. suggested that PheDH is responsible for the degradation of Phe, not its synthesis (1); however, in *S. ureae* (3), L-phenylalanine aminotransferase activity with oxaloacetic acid as an amino acceptor was detected during growth on a medium containing Phe at a range between 0.001 and 0.006 units/mg, while PheDH was highly induced (0.033 units/mg, measured in the oxidative deamination reaction). The formation of PheDH was growth associated, indicating that the enzyme is responsible for the catabolism of Phe. On the other hand, when the strain was grown in M9 medium supplemented with 0.1% yeast extract, the PheDH activity was detected only in the exponential phase at a level as high as 0.059 units/mg, which is much higher than L-phenylalanine aminotransferase activity (less than 0.001 units/mg). Because the equilibrium of PheDH favors the synthesis of the amino acid, the enzyme probably also functions in the anabolism of Phe, depending on the environment.

Enzymatic Properties

The enzymatic properties of PheDH are summarized in Table 1. The molecular weight of the PheDH subunits are in the range of 36,000 to 42,000. The higher structures vary; PheDHs from *Sporosarcina* (2,3), *Bacillus* (3), and *Microbacterium* (15) are octameric enzymes, whereas PheDH from *Rhodococcus* sp. (12) is tetrameric, that from *Rhodococcus maris* (9) is dimeric, and that from *Nocardia* (10) is monomeric. The molecular weight of PheDH ranges from 42,000 (the monomeric *Nocardia* enzyme) to 331,480 (the octameric *B. sphaericus* enzyme). The optimum pH for the reversible oxidative deamination and reductive amination varies by about 1 pH unit; the former reaction varies in the range of pH 10.4 to 11.3, whereas the latter is in the range of pH 9 to 10.3. The velocity toward the reductive amination to synthesize Phe generally is several times faster than that for the reverse reaction at optimum pH; the V_{max} values for the oxidative deamination reaction of Phe catalyzed by *Sporosarcina* and *B. sphaericus* PheDH are both 114, whereas those for the reductive amination reaction for phenylpyruvate are 598 and 416, respectively (3). The K_m values toward the preferred substrates of L-amino acids and keto acids are in the range of 0.1 to 1 mM, though much higher K_m values of around 100 mM were observed toward ammonia.

Substrate Specificity and Kinetic Properties

In the oxidative deamination reaction, Phe and L-norleucine were active as substrates for the *Sporosarcina* enzyme. The from enzymes *B. badius* and *S. ureae* show narrower substrate specificity than that from *B. sphaericus*, which has almost equal affinity toward Tyr and its keto analog (Table 2), thus *B. sphaericus* PheDH may be called *tyrosine dehydrogenase*. On the other hand, the enzymes showed wider substrate specificities in the reductive amination reaction than observed in the oxidative deamination reaction. The substrate specificity of PheDH has been extensively characterized with the *B. sphaericus* enzyme by using chemically synthesized substrates and substrate analogs (3,6). The relative rates of the reductive amination of α-keto acids and their analogs catalyzed by PheDH from *S. ureae* and *B. sphaericus* are shown in Table 3. The substrate specificity of *B. sphaericus* was examined in detail with various synthetic compounds as shown in Table 4. The enzyme was found to accommodate α-keto acids with large substituents, including compounds substituted at the β-position of pyruvic acid with a longer or bulkier group, although the relative velocity of the reductive amination reaction was low. Phenylpyruvate analogs substituted at the phenyl ring were relatively good substrates, whereas the substitution at the β-position of phenylpyruvate with a bulkier group, such as a hexyl group, greatly lowered the reaction velocity. On the other hand, pyruvate, α-ketoglutarate, α-hydroxypyruvate, benzoylformate, α-keto esters such as ethyl phenylpyruvate and ethyl α-keto-γ-phenylbutyrate, β-keto esters such as ethyl β-keto-γ-phenylbutyrate, β-keto acids such as β-keto-γ-phenylbutyrate, and α-ketoalcohols such as α-keto-β-phenylpropanol were inactive as substrates. Thus, it was shown that the enzyme acts solely on α-keto acids, and a free carboxylic acid moiety is required for a compound to be recognized as a substrate. It was also revealed that the enzyme does not differentiate the configuration of the substituent at the β-position of pyruvic acid; for example, a diastereomeric mixture of L-α-amino-β-DL-methyl-β-phenylpropionate was synthesized from α-keto-β-DL-methyl-β-phenylpropionate. Gln, Asp, methylamine, dimethylamine, trimethylamine and ethylamine (each at 200 mM) did not replace the ammonium ion in the reductive amination reaction. Thus, PheDH from *B. sphaericus* showed wide substrate specificity toward substituted pyruvic acids. The finding that the enzyme utilized α-keto-γ-phenylbutyrate and α-keto-ε-phenylvalerate, but not β-keto-γ-phenylbutyrate, shows it has a definite requirement for a distance between the carbonyl carbon and the carboxyl group of the substrates. For the *Sporosarcina* enzyme, phenylpyruvate, α-ketocaproate, α-keto-γ-methylthiobutyrate, *p*-hydroxyphenylpyruvate, and α-ketoisocaproate were active as substrates. NAD^+ was active as a cofactor for both enzymes, whereas $NADP^+$ was not. The mean value of the apparent equilibrium constant (K_{eq} = [phenylpyruvate][NADH][NH_4^+][H^+]/[Phe][NAD^+][H_2O]) was determined using the *Sporosarcina* enzyme to be 1.4×10^{-15} at the pH range of 7.0 to 11.5 and 25 °C, assuming the conventional concentration of water ([H_2O] = 1 M) (3). For the *B. sphaericus* enzyme, it was calculated to be 2.0×10^{-15} at pH 8.40 to 10.38 (6). The substrate specificities of PheDH from *B. badius* (7) and *R. maris* (9) are similar to that of *S. ureae* (3), which shows more limited activities than the *B. sphaericus* enzyme, although *R. maris* PheDH has a slightly higher affinity toward Tyr than *S. ureae* PheDH. It is notable that the PheDH of *Nocardia* sp. is 0.54 and 2.4 times more active toward indolepyruvate and α-ketoisocaproate, respectively, than toward phenylpyruvate (10), and *T. intermedius* PheDH is inactive toward Tyr and its keto analog (11).

STRUCTURE OF PheDH

The *pdh* genes have been cloned and sequenced, and overproduction of the *Bacillus* enzymes has been achieved (7,18,19). *Escherichia coli* JM 109/pBPDH1-DBL expresses about 120-fold higher activity of PheDH (7,200 units/L) than the wild type *B. sphaericus*. The nucleotide sequences for the *pdh* genes from *B. sphaericus* (19), *B. badius* (20), *S. ureae* (21), *T. intermedius* (22), and *Rhodococcus* sp. (12) have been reported.

Computer assessment of the data showed that the deduced primary structure of *B. badius* PheDH (20) is similar to PheDH of *B. sphaericus* (67.9% identical over 377 amino acids) (19) and *T. intermedius* (56.6% identical over 346 amino acids) (22), LeuDH from *B. stearothermophilus* (23) (49.9% identical over 353 amino acids), $NADP^+$-dependent valine dehydrogenase from *Streptomyces coelicolor* (24) (46.6% identical over 352 amino acids), PheDH from *Rhodococcus* sp. (34.7% identical over 334 amino acids) (12), and GluDH from *Clostridium symbiosum* (25.9% identical over 185 amino acids) (25). Similarities between these proteins are seen in the catalytic domains of amino acid dehydrogenases, G-G-(G or A)-K (26), and the glycine-rich

Table 1. Comparison of Properties of PheDHs

Property	Microorganism						
	B. sphaericus	*S. ureae*	*B. badius*	*R. maris*	*Nocardia* sp.	*T. intermedius*	*Microbacterium* sp.
Specific activity of final preparation (U/mg)	111	84	68	65	30	86	37
M_r by gene sequencing	331,480	330,608	330,800	–[a]	–	242,928	–
M_r gel filtration	340,000	310,000	335,000	70,000	42,000	270,000	330,000
M_r of subunit(s) by gene sequencing	41,435	41,326	41,350	36,000	42,000	40,488	41,000
Number of subunits	8	8	8	2	1	6	8
pI	4.3	5.3	3.5	–	–	–	5.8
pH optimum							
Oxidative deamination	11.3	10.5	10.4	10.8	10	11.0	12.0
Reductive amination	10.3	9.0	9.4	9.8	–	9.2	12.0
Apparent K_m (mM) value for							
L-Phenylalanine	0.22	0.096	0.088	3.80	0.75	0.22	0.10
L-Tyrosine	0.50	–	–	–	–	–	–
Phenylpyruvate	0.40	0.16	0.106	0.50	0.06	0.045	0.03
p-Hydroxyphenylpyruvate	0.34	–	–	1.30	–	–	–
NAD^+	0.17	0.14	0.15	0.25	0.23	0.07	0.20
NADH	0.025	0.072	0.21	0.043	–	0.025	0.07
Ammonia	78	85	127	70	96	106	85
Remaining activity after incubation for 10 min	100%	75%	50%	100%	50%	100%	100%
	55 °C	40 °C	55 °C	35 °C	53 °C, 2 h	70 °C, 1 h	60 °C
	pH9.0	pH9.0	pH8.0	pH7.4	pH9.5–10	pH7.2	
Reference	(3,9)	(3,21)	(7,20)	(9)	(10)	(11,12)	(15)

[a] –, data not available.

Table 2. Substrate Specificity of PheDH from *S. ureae* SCRC-R04 and *B. sphaericus* SCRC-R79a

Amino acid	*S. ureae* SCRC-R04			*B. sphaericus* SCRC-R79a		
	Relative activity (%)	K_m (mM)	V_{max} (units/mg)	Relative activity (%)	K_m (mM)	V_{max} (units/mg)
Phe	100	0.096[a]	114[a]	100	0.22[a]	114[a]
Tyr	5.4[b]			72[c]	0.50[a]	112[a]
Trp	5.0			1.2		
His	0.1			ND[d]		
Met	4.1			3.0		
Ethionine	7.0			3.1		
Val	3.1			1.4		
Leu	2.3			1.3		
Ile	0.7			0.45		
L-*allo*-Isoleucine	4.3			0.3		
L-α-Amino-*n*-butyric acid	1.6			NT[e]		
L-Norvaline	6.3			1.3		
L-Norleucine	15	0.25[f]	20.0[f]	3.9		

Note: The oxidative deamination reaction was carried out under the standard reaction conditions. The concentration of amino acid was 10 mM unless indicated otherwise (3). Amino acids that were inactive as substrates for the *Sporosarcina* enzyme included Gly, Ala, His, Arg, Lys, Orn, Asp, Glu, Gln, Pro, Ser, Thr, Cys, L-tyrosinamide, L-tyrosine hydroxamate, D-Phe, D-Tyr, D-phenylglycine, *tert*-L-leucine, and DL-*threo*-phenylserine.
[a]Value determined from the secondary plots of intercepts against reciprocal concentration of the substrate.
[b]Measured at 1.4 mM.
[c]Measured at 0.3 mM.
[d]ND, not detected.
[e]NT, not tested.
[f]Apparent value determined from double reciprocal plot a fixed concentration of NAD^+ (2.5 mM).

Table 3. Substrate Specificity of PheDH from *S. ureae* SCRC-R04 and *B. sphaericus* SCR-R79a in the Reductive Amination Reaction

Keto acid	*S. ureae* SCRC-R04			*B. sphaericus* SCRC-R79a		
	Relative Activity (%)	K_m (mM)	V_{max} (units/mg)	Relative activity (%)	K_m (mM)	V_{max} (units/mg)
Phenylpyruvate	100	0.16[a]	598[a]	100	0.40[a]	416[a]
p-Hydroxyphenylpyruvate (5 mM)	24			136	0.34[a]	1240[a]
Indole-3-pyruvate (2 mM)	0.73			0.39		
Imidazole pyruvate	0.04			ND[b]		
α-Keto-γ-methylthiobutyrate	27			11		
α-Ketoisovalerate	2.1			5.5		
α-Ketoisocaproate	13			7.8		
DL-α-Keto-β-methylvalerate	3.2			2.9		
α-Ketovalerate	8.8			6.2		
α-Ketocaproate	32	2.44[c]	379[c]	ND[b]		

Source: From Ref. 3.
[a]Apparent value obtained from the secondary plot of intercepts against reciprocal concentration of the substrate at a fixed concentration of NH_4Cl (400 mM).
[b]ND, not detectable.
[c]Apparent value obtained from a double reciprocal plot at a fixed concentration of NADH (0.4 mM) and NH_4Cl (400 mM).

nucleotide-binding domain G-X-G-X-X-(G or A), connecting β-strand with α-helix in the region of adenine dinucleotide phosphate (ADP)-binding $\beta\alpha\beta$ folds, which is strongly conserved among $NAD(P)^+$-dependent dehydrogenases and FAD-containing oxidoreductases (27).

Reaction Mechanism

Speculation concerning the reaction mechanism of PheDH is based on the findings with respect to the substrate specificity already described, and on other experiments with GluDH and LeuDH. The formation of 2-iminoglutarate, catalyzed by GluDH, has been proved (28). The *pro-S* hydrogen at position 4 of the reduced pyridine ring of NADH is incorporated into the product Phe in the reductive amination reaction of PheDH (B-stereospecific) (3). Studies on the steady-state kinetics of the PheDH reaction revealed that the reaction proceeds sequentially (3); that is, after all three substrates (phenylpyruvate, ammonia, and NADH) bind to the enzyme, the product Phe is released. In a homology search of the deduced primary structure of PheDH, a sequence similar to the active site of GluDH, which also resembles LeuDH, has been found (29). The reaction catalyzed by PheDH would be analogous to that of GluDH from *Clostridium symbiosum* discussed on the basis of X-ray crystallographic studies by Rice et al.

Table 4. Substrate Specificity of PheDH from *B. sphaericus* SCRC-R79a (3,6)

Substrate (10 mM)	Relative activity (%)
4-Hydroxyphenylpyruvate	136
Phenylpyruvate	100
4-Vinylphenylpyruvate	52
4-Fluorophenylpyruvate	39
α-Keto-γ-methylthiobutyrate	11
α-Keto-β-DL-phenylvalerate	8.8
α-Ketoisocaproate	7.8
α-Ketobutyrate	6.3
α-Ketovalerate	6.2
α-Ketoisovalerate	5.5
α-Keto-β-DL-(4-fluorophenyl)butyrate	5.0
α-Keto-β-DL-phenylbutyrate	4.8
α-Keto-4-phenylbutyrate	3.4
α-Keto-β-DL-methylvalerate	2.9
α-Keto-ε-phenylvalerate	2.0
α-Keto-β-DL-(3-methylphenyl)butyrate	1.0
α-Ketononanoate	0.73
α-Keto-β-(2-naphthalene)propionate	0.46
Indolepyruvate (2 mM)	0.39
α-Keto-β-DL-phenyl-γ-methylvalerate	0.10

(25,26,29–32) and to that of *B. stearothermophilus* LeuDH estimated by the kinetic studies done with wild-type and mutant enzymes by Tanizawa et al. (33). The first step would be initiated by the deprotonation of the α-amino group of Phe by an acidic residue (^{126}Asp in *B. sphaericus* PheDH, which corresponds to ^{165}Asp in GluDH) acting as a general base. Subsequently, a hydride transfer from the hydrogen attached to the α-carbon to the *Si* face of NAD^+ (a hydride transfer between the *pro-S* hydrogen of NADH and the *Re* face of the imino acid) occurs (3). Next, the iminophenylpyruvate produced is hydrated by the action of a general base (^{90}Lys in PheDH, which corresponds to ^{125}Lys in GluDH and ^{80}Lys in LeuDH) in the consensus sequence (G-G-X-K, X = G or A or S). The iminophenylpyruvate breaks into phenylpyruvate, ammonia, and a proton. The orientation of phenylpyruvate is fixed, with the phenyl group accommodated in the hydrophobic pocket and the carboxyl group as an anchor, as evidenced by the finding that the enzyme does not catalyze the reductive amination of neutral α-keto-β-phenylpropanol and the α-keto esters (6). Tanizawa et al. (33) generated LeuDH mutants of *B. stearothermophilus* in which ^{80}Lys was changed to Ala, Arg, or Gln and found that the mutants showed reduced activities, but their K_m values for substrate and coenzyme did not change significantly. They analyzed the pH dependence of the reaction and observed a solvent isotope effect, suggesting that ^{80}Lys participates in catalysis as a general base and assists a nucleophilic attack of water to the substrate α-carbon atom in the oxidative deamination reaction.

PheDH of *T. intermedius* acts preferentially on Phe and Tyr (11), whereas LeuDH of *B. stearothermophilus* acts almost exclusively on Leu and some other branched-chain L-amino acids (34). The two enzymes share a similar sequence (47%). The inherent hexapeptide segment (^{124}FVHAAR) in the substrate-binding domain of a mutant PheDH of *T. intermedius* was replaced by the corresponding part of LeuDH (MDIIYQ) (35). The catalytic efficiencies (K_{cat}/K_m) of the mutant enzyme with aliphatic amino acids and aliphatic keto acids as substrates were 0.5 to 2% of those of the wild-type enzyme. However, the efficiencies for Phe and phenylpyruvate were greatly reduced to 0.008 and 0.035% of those of the wild-type enzyme, respectively. Soda et al. (36) constructed a chimeric enzyme consisting of an N-terminal domain of PheDH containing the substrate-binding region and a C-terminal domain of LeuDH containing the NAD^+-binding region. The chimeric enzyme showed 6% of that of the parental PheDH activity on Phe and had broader substrate specificity in the oxidative deamination. The substrate specificity of the chimeric enzyme in the reductive amination was an admixture of PheDH and LeuDH. The chimeric enzyme has a Gly corresponding to position 124 of PheDH and a Val corresponding to position 307.

Engel et al. showed the same effect by changing only two sites. ^{124}Gly and ^{307}Leu of PheDH from *B. sphaericus* were altered by site-specific mutagenesis to the corresponding residues in LeuDH, Ala and Val, respectively (19,37). These two residues have previously been implicated (by molecular modeling based on the X-ray crystallographic analysis of *C. symbiosum* GluDH [25,26,28–32]), as important residues in determining the degree of substrate discrimination. Of the single mutants L307V and G124A, and the double mutant G124, L307V displayed lower activities toward Phe and enhanced activity toward almost all aliphatic amino acid substrates tested compared to the wild-type; thus an amino acid dehydrogenase that shows closer resemblance to LeuDH was made from PheDH by changing only two residues.

APPLICATION OF THE ENZYME

Application of thermostable AlaDH and LeuDH is reviewed by Ohsima and Soda (34). A review on regeneration of NADH for organic synthesis was published by Chenault and Whitesides (38). By carefully choosing the differences in the substrate specificities of PheDHs, *B. sphaericus* enzyme was applied to the synthesis of various L-amino acids (6). PheDH from *B. badius* was shown be narrower in substrate specificity (7) and, therefore, was applicable to the microdetermination of Phe and phenylpyruvate in blood (8).

Synthesis of L-Amino Acids from Their α-Keto Analogs

The production of Phe as a starting material for the artificial sweetener aspartame has been a target of industrial development (39). Several enzymatic processes of Phe synthesis have been reported: L-specific hydrolysis of benzylhydantoin (40), amination of *trans*-cinnamic acid (41), transamination from an amino donor to phenylpyruvate (42), and two-step conversion starting from acetamidocinnamic acid *via* phenylpyruvate (43,44). Because the α-keto acids are generally more expensive than the L-amino acids, it has been impractical to use the α-keto acids as starting materials for the synthesis of L-amino acids. With the development of an efficient method of phenylpyruvic acid synthesis by double carbonylation of benzylchloride in the

presence of a cobalt catalyst, a new enzymatic method for the synthesis of Phe from phenylpyruvate was sought (45). Previously, the enzymatic synthesis of Phe from phenylpyruvate was accomplished with the transamination reaction. However, this process requires an L-amino acid, such as Asp, as an amino donor, and after the reaction, a by-product such as oxaloacatate or its degradation product, pyruvic acid, are formed. With PheDH, ammonia can be used as a ammonium source.

The discovery of PheDH made it possible to devise a reaction scheme for the synthesis of Phe from phenylpyruvate with NADH as a reducing agent. The enzyme from *B. sphaericus* was chosen to study the application of the enzyme to the synthesis of various L-amino acids because it is very stable and shows broader substrate specificity than other PheDHs, acting on Tyr as well as Phe. Using PheDH from *B. sphaericus* and formate dehydrogenase (FDH) from *Candida* (Fig. 1), with a catalytic amount of NAD^+, Phe was synthesized in a good yield (120 g/L) (6,46).

In the study on substrate specificity of the enzyme in the reductive amination of α-keto acids, various optically pure L-amino acids were quantitatively synthesized using PheDH and FDH. Table 5 shows the yield of L-amino acids thus synthesized (6). The products from α-keto-β-DL-methylbutyrate and α-keto-β-DL-methyl-β-phenylpyruvate were identified as diastereomeric mixtures of Ile and L-*allo*-isoleucine, and L-α-amino-β-DL-methyl-β-phenylpropionic acid, respectively. L-α-Amino-γ-phenylbutyric acid (L-homophenylalanine) and other unnatural L-amino acids could be efficiently synthesized. L-Homophenylalanine is a building block of some of the angiotensin-converting-enzyme inhibitors (47). The solubility of these Phe homologs, such as Tyr, L-α-amino-γ-phenylbutyric acid, L-α-amino-ϵ-phenylvaleric acid, and so on, is so low that they are easily separated in crystalline forms from the reaction mixture by filtration. The filtered enzyme solution can be used for further repeated synthesis.

L-β-Chloroalanine is a useful intermediate for the synthesis of several L-amino acids. Conditions for the synthesis of optically pure L-β-chloroalanine from β-chloropyruvate using AlaDH, LeuDH, and PheDHs from *B. badius*, *S. ureae*, and *T. intermedius* with a regeneration of NADH by FDH were investigated (48). The enzymatic reaction was carried out at neutral pH because of the chemical instability of β-chloropyruvate under alkaline conditions. AlaDH from *B. stearothermophilus* IFO 12550 showed the highest activity for the production of L-β-chloroalanine at pH 7.5. L-β-Chloroalanine was produced with high chemical (>90%) and optical yields (100% enantiomeric excess) and at a high concentration (43 g/L).

Immobilized *Nocardia opaca* cells were used to synthesize Phe from phenylpyruvate and ammonia under hydrogen pressure, although no enzymological study was done (49). Thermostable L-amino acid dehydrogenases such as PheDH and formate dehydrogenase were simultaneously expressed in *E. coli*, which was utilized for the synthesis of L-amino acids from their keto analogs (50).

Continuous Synthesis of Phe

Membranes are efficiently used in the enzyme-catalyzed syntheses of optically active compounds on laboratory and industrial scales (51–53). This technique can be applied for the repeated use of the enzyme in the synthesis, especially if smaller amounts of enzyme are available. Small amounts of PheDH from *B. sphaericus* and FDH were placed in a dialysis tube to check their durability and ef-

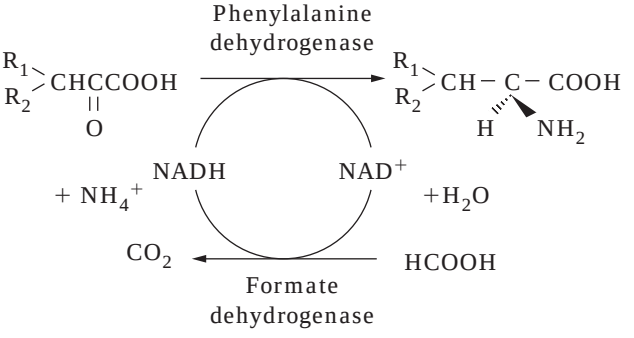


Figure 1. Synthesis of L-amino acids from their β-keto analogs by PheDH with a regeneration of NADH by formate dehydrogenase.

Table 5. Synthesis of L-Amino Acids from α-Keto Acids by Using *B. sphaericus* PheDH and FDH (6)

Substrate	Product	Yield (%)
Phenylpyruvate	L-Phe	>99
4-Hydroxyphenylpyruvate	L-Tyr	>99
4-Fluorophenylpyruvate	L-4-Fluorophenylalanine	>99
α-Keto-γ-phenylbutyrate	L-α-Amino-γ-phenylbutyric acid	99
α-Keto-ϵ-phenylvalerate	L-α-Amino-ϵ-phenylvaleric acid	98
α-Keto-β-methyl-β-phenylpropionate	L-α-Amino-β-DL-methyl-β-phenylpropionate	98
α-Ketononanoate	L-α-Aminononanoic acid	99

Note: To prevent substrate inhibition, α-keto acids have been divided into portions and added to the reaction mixture, not to exceed 50 mM.

ficiency in the continuous synthesis of Phe. The *B. sphaericus* PheDH proved to be very stable in the synthesis, with high efficiency. PheDH maintained its activity for up to 28 changes. During the 34-day operation, 10.74 g (6.51×10^{-2} mol) of Phe was synthesized, using 45 μg (5 units, 1.35×10^{-10} mol) of PheDH, ca. 7.5 mg (15 units) of FDH, and ca. 35 mg of $NAD^+ \cdot Na$ (5×10^{-5} mol). Thus, 1 mol of PheDH catalyzed the synthesis of 4.8×10^8 mol of Phe, which is 2.4×10^5 times the weight of the enzyme (6). Continuous synthesis of Phe by the use of membrane reactors has been demonstrated with PheDHs from *Brevibacterium* sp. (54) and *Rhodococcus* sp. (55,56).

To avoid the multistep purification procedure, the use of enzymes in whole cells is preferable. Acetone-dried cells of *B. sphaericus* and *C. boidinii* were also effective for Phe synthesis, providing a simple microbial method of synthesis. In a typical course of the reductive amination reaction of phenylpyruvate using the mixture of acetone-dried cells as catalysts, the concentration of Phe reached 61.5 mg/mL with a yield of more than 99% (6).

Tyr and Phe labeled with a positron-emitting radionuclide would be useful for positron emission tomography (PET) studies of tissues and neoplasms, such as measuring protein synthesis in tumor and brain tissue. ^{13}N-Labeled Phe, Tyr, and L-dopa were synthesized for PET (57). PheDH catalyzes the reductive ^{13}N amination of either phenylpyruvate or *p*-hydroxyphenylpyruvate to form [^{13}N]Phe or [^{13}N]Tyr, respectively, with NADH as a reducing agent. After short incubation of ^{13}N ammonia and either phenylpyruvate or *p*-hydroxyphenylpyruvate with the enzyme immobilized on the CNBr-activated Sepharose column, 83% of the ^{13}N ammonia was converted to [^{13}N]Phe, and 38% of the ^{13}N was converted to [^{13}N]Tyr, respectively. The labeled amino acids were >30 mCi. L-Dopa was labeled by the exchange reaction between L-dopa and ^{13}N ammonia in the presence of NAD^+ and PheDH; 9% of the label was transferred to L-dopa. In a separate experiment involving the exchange reaction between Phe and ^{13}N ammonia, transfer of label from ^{13}N ammonia to Phe was 50% (Scheme 1). Thus, utilizing *B. sphaericus* PheDH is useful for preparing [^{13}N]Tyr and [^{13}N]Phe.

Microdetermination by Cycling Assay

Quantitative determination of Phe in the plasma is important in diagnosing PKU. Several groups have been investigating the use of PheDH for the microdetermination of Phe and phenylpyruvate in blood samples (58–70). *B. badius* PheDH was found to be structurally closely related to that from *B. sphaericus*, although the former has a narrower substrate specificity suitable for the microdetermination, whereas the latter is more stable and shows a wider specificity suitable for the synthesis of various L-amino acids.

$$^{13}NH_3 + \text{L-dopa} \xrightarrow[NADH+H^+]{PheDH} NH_3 + \text{L-[}^{13}\text{N]dopa}$$

Scheme 1. Enzymatic synthesis of L-[^{13}N]-dopa by an exchange reaction.

Utilizing *B. sphaericus* PheDH, enzymatic cycling assay for the determination of Phe and phenylpyruvate in deproteinized tissue extracts was devised by Cooper et al. (61). Assay 1 couples glutamine transaminase K with PheDH (Scheme 2). Assay 2 combines PheDH, L-amino acid oxidase, and catalase (Scheme 3). In both assays, Tyr and some other amino acids (or their α-keto acid analogs) can replace Phe (or phenylpyruvate) to a small extent. Thus, when measuring Phe, a correction must be made for the nonspecificity of the reaction. By removing Phe on a cation exchange column, it was possible to measure phenylpyruvate in tissue extracts. Concentrations of phenylpyruvate (μmol/kg) in normal rat liver, kidney, and brain were 2.1 $\pm$ 1.1 (n = 8), 1.8 $\pm$ 0.4 (n = 4), and 3.3 $\pm$ 0.6 (n = 4), respectively.

A spectrophotometric recycling assay for the quantitation of Phe (and phenylpyruvate) involves the coupling of PheDH with rat kidney cytosolic glutamine transaminase K (68). The latter enzyme possesses high affinity for phenylpyruvate. Recycling results in a 50-fold increase in sensitivity over that of a conventional spectrophotometric end-point analysis procedure. The spectrophotometric recycling procedure has now been adapted to the measurement of Phe in microliter quantities of human blood. The procedure is 10 times more sensitive than the end-point assay for the spectrophotometric measurement of Phe in

$$\text{L-Phe} + NAD^+ + H_2O \xrightarrow{PheDH} PPA + NADH + H^+ + NH_3$$

$$\text{L-Gln} + PPA \xrightarrow{Gln\ Transaminase\ K} \alpha\text{-Ketoglutaramate} + \text{L-Phe}$$

Net

$$\text{L-Gln} + NAD^+ + H_2O \longrightarrow \alpha\text{-Ketoglutaramate} + NADH + H^+ + NH_3$$

Scheme 2. Enzymatic cycling assay for Phe.

$$\text{L-Phe} + H_2O + O_2 \xrightarrow{\text{L-}AAO} PPA + H_2O_2 + NH_3$$

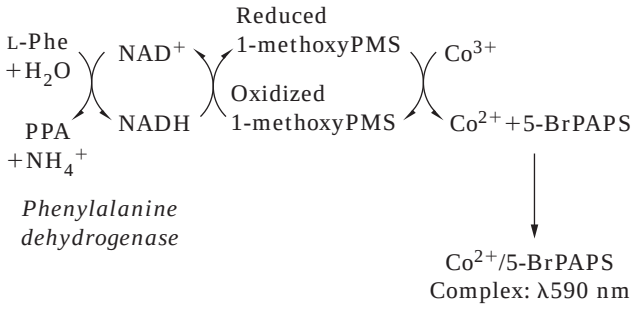
$$H_2O_2 \xrightarrow{Catalase} H_2O + 1/2\ O_2$$

$$PPA + NADH + H^+ + NH_3 \xrightarrow{PheDH} \text{L-Phe} + NAD^+ + H_2O$$

Net

$$NADH + H^+ + 1/2\ O_2 \longrightarrow NAD^+ + H_2O$$

Scheme 3. Enzymatic cycling assay for phenylpyruvate.



Scheme 4. Colorimetric determination of Phe in blood samples.

human blood. The findings suggest that the recycling procedure adapted for fluorometry will be even more sensitive.

Microdetermination by End-Point Assay

Naruse et al. have introduced a sensitive enzymatic microplate method using PheDH from *B. badius* for the mass screening of PKU among neonates (7,8) as an alternative method for the Guthrie test (71). The principle of this system involves two steps. First, Phe is oxidized by PheDH to phenylpyruvate, reducing NAD^+ to NADH, which then serves as the electron donor in a colorimetric reaction as follows. In the presence of NADH and the electron carrier l-methoxy-5-methylphenazium methylsulfate (l-MPMS), Co^{3+} is reduced to Co^{2+}, which produces a colored chelate with the metal indicator 2-(5-bromo-2-pyridyazo)-5-(*N*-propyl-*n*-sulfopropyl-amino)phenol (5-BrPAPS) (Scheme 4). This assay system detected Phe in filter paper blood in the concentration range of 0.02–1.50 mM. This method uses a simple TCA extraction method and a PheDH reaction with high substrate specificity that produces very sensitive colorimetric quantitation. The absorbance at 585 nm of this chelate can be measured by a colorimetric microplate reader. This highly specific and sensitive method has been used to screen neonates for PKU in Japan since 1992.

CONCLUSION

Since the recent discovery of PheDH in 1984, several microbial producers of the enzyme have been isolated and their enzymological properties and primary structures have been studied. Indeed, PheDH has been used to detect PKU in the blood of neonates in Japan since 1992. However, PheDH has many further applications to be studied. If the structure and function of the amino acid dehydrogenases are thoroughly understood, it might be possible to create a new amino acid dehydrogenase with a wider scope of application.

BIBLIOGRAPHY

1. W. Hummel, N. Weiss, and M.-R. Kula, *Arch. Microbiol.* **137**, 47–52 (1984).
2. Y. Asano and A. Nakazawa, *Agric. Biol. Chem.* **49**, 3631–3632 (1985).
3. Y. Asano, A. Nakazawa, and K. Endo, *J. Biol. Chem.* **262**, 10346–10354 (1987).
4. E.L. Smith, B.M. Austen, K.M. Blumenthal, and J.F. Nyc, *Enzymes* **11**, 293–367 (1975).
5. Nomenclature Committee of the International Union of Biochemistry, *Enzyme Nomenclature*, Academic Press, Orlando, Fla., 1992
6. Y. Asano, A. Yamada, Y. Kato, K. Yamaguchi, Y. Hibino, K. Hirai, and K. Kondo, *J. Org. Chem.* **55**, 5567–5571 (1990).
7. Y. Asano, A. Nakazawa, K. Endo, Y. Hibino, M. Ohmori, N. Numao, and K. Kondo, *Eur. J. Biochem.* **168**, 153–159 (1987).
8. H. Naruse, Y.Y. Ohashi, A. Tsuji, M. Maeda, K. Nakamura, T. Fujii, A. Yamaguchi, M. Matsumoto, and M. Shibata, *Screening* **1**, 63–66 (1992).
9. H. Misono, J. Yonezawa, S. Nagata, and S. Nagasaki, *J. Bacteriol.* **171**, 30–36 (1989).
10. L. de Boer, M. van Rijssel, G.J. Euverink, and L. Dijkhuizen, *Arch. Microbiol.* **153**, 12–18 (1989).
11. T. Ohshima, H. Takada, T. Yoshimura, N. Esaki, and K. Soda, *J. Bacteriol.* **173**, 3943–3948 (1991).
12. N.M.W. Brunhuber, A. Banerjee, W.R. Jacobs, Jr., and J. Blanchard, *J. Biol. Chem.* **269**, 16203–16211 (1994); DDBJ/EMBL/GenBank databases, accession number: RSU08381.
13. X.-L. Huang and Y. Asano, *Bull. Toyama Pref. Univ.* (in Japanese) **4**, 211–215 (1994).
14. Y. Asano, M. Tanetani, and N. Horita, *Bull. Toyama Pref. Univ.* (in Japanese) **6**, 101–106 (1996).
15. Y. Asano and M. Tanetani, *Arch. Microbiol.* **169**, 220–224 (1998).
16. S. Kuroda, K. Tanizawa, Y. Sakamoto, H. Tanaka, and K. Soda, *Biochemistry* **29**, 1009–1015 (1990).
17. J.R. Norris, in R.C.W. Berkeley and M. Goodfellow eds., *The Aerobic Endospore-Forming Bacteria: Classification and Identification* Academic Press, London 1981, pp. 337–357.
18. Y. Asano, K. Endo, A. Nakazawa, Y. Hibino, M. Ohmori, N. Numao, and K. Kondo, *Agric. Biol. Chem.* **51**, 2621–2623 (1987).
19. N. Okazaki, Y. Hibino, Y. Asano, M. Ohmori, N. Numao, and K. Kondo, *Gene* 63, 337–341 (1988); DDBJ/EMBL/GenBank databases, accession number: BACPDH.
20. A. Yamada, T. Dairi, Y. Ohno, X.-L. Huang, and Y. Asano, *Biosci. Biotech. Biochem.* **59**, 1994–1995 (1995); DDBJ/EMBL/GenBank databases, accession number: D50261.
21. DDBJ/EMBL/GenBank databases, accession number: AB001031
22. H. Takada, T. Yoshimura, T. Ohshima, N. Esaki, and K. Soda, *J. Biochem.* (Tokyo) **109**, 371–376 (1991); DDBJ/EMBL/GenBank databases, accession number: BTTPDH
23. S. Nagata, K. Tanizawa, N. Esaki, Y. Sakamoto, T. Ohshima, H. Tanaka, and K. Soda, *Biochemistry* **27**, 9056–9062 (1988).
24. L. Tang and R. Hutchinson, *J. Bacteriol.* **175**, 4176–4185 (1993).
25. K.L. Britton, P.J. Baker, D.W. Rice, and T.J. Stillman, *Eur. J. Biochem.* **209**, 851–859 (1992).
26. P.J. Baker, K.L. Britton, D.W. Rice, A. Rob, and T.J. Stillman, *J. Mol. Biol.* **228**, 662–671 (1992).
27. R.K. Wierenga, M.C.H. Demaeyer, and W.G.J. Hol, *Biochemistry* **24**, 1346–1357 (1985).
28. M.C. Hochreiter and K.A. Schellenberg, *J. Am. Chem. Soc.* **91**, 6530–6531 (1969).
29. K.L. Britton, P.J. Barker, P.C. Engel, D.W. Rice, and T.J. Stillman, *J. Mol. Biol.* **234**, 938–945 (1993).

30. T.J. Stillman, P.J. Baker, K.L. Britton, and D.W. Rice, *J. Mol. Biol.* **234**, 1131–1139 (1993).
31. P.J. Baker, A.P. Turnbull, S.E. Sedelnikova, T.J. Stillman, and D.W. Rice, *Structure* **3**, 693–705 (1995).
32. A.P. Turnbull, P.J. Baker, and D.W. Rice, *J. Biol. Chem.* **272**, 25105–25111 (1997).
33. T. Sekimoto, T. Matsuyama, T. Fukui, and K. Tanizawa, *J. Biol. Chem.* **268**, 27039–27045 (1993).
34. T. Ohshima and K. Soda, *Trends Biotechnol.* **7**, 210–214 (1989).
35. K. Kataoka, H. Takada, T. Yoshimura, S. Furuyoshi, N. Esaki, T. Ohshima, and K. Soda, *J. Biochem.* (Tokyo) **114**, 69–75 (1994).
36. K. Kataoka, H. Takada, K. Tanizawa, T. Yoshimura, N. Esaki, T. Ohshima, and K. Soda, *J. Biochem.* (Tokyo) **116**, 931–936 (1994).
37. S.Y.K. Seah, K.L. Britton, P.J. Baker, D.W. Rice, Y. Asano, and P.C. Engel, *FEBS Lett.* **370**, 93–96 (1995).
38. H.K. Chenault and G.M. Whitesides, *Appl. Biochem. Biotechnol.* **14**, 147–197 (1987).
39. A. Klausner, *Bio/technology* **3**, 301–307 (1985).
40. K. Yokozeki, K. Sano, C. Eguchi, H. Iwagami, and K. Mitsugi, *Agric. Biol. Chem.* **51**, 729–736 (1987).
41. C.T. Evans, K. Hanna, C. Payne, D. Conrad, and M. Misawa, *Enzyme Microb. Technol.* 417–421 (1987).
42. G.J. Calton, L.L. Wood, M.H. Updike, L. Lantz II, and J.P. Hammen, *Bio/technology* **4**, 317–320 (1986).
43. K. Nakamichi, K. Nabe, Y. Nishida, and T. Tosa, *Appl. Microbiol. Biotechnol.* **30**, 243–246 (1989).
44. C.T. Evans, W. Bellamy, M. Glesson, H. Aoki, K. Hanna, W. Peterson, D. Conrad, and M. Misawa, *Bio/technology* **5**, 818–823 (1987).
45. JP 60-61550 (April 9, 1985), K. Hirai and I. Ojima.
46. Y. Asano and A. Nakazawa, *Agric. Biol. Chem.* **51**, 2035–2036 (1987).
47. H. Urbach and R. Henning, *Tetrahedron Lett.* **25**, 1143–1146 (1984).
48. Y. Kato, K. Fukumoto, and Y. Asano, *Appl. Microbiol. Biotechnol.* **39**, 301–304 (1993).
49. T. Matsunaga, M. Higashijima, A. Sulaswatty, and S. Nishimura, *Biotechnol. Bioeng.* **31**, 834–840 (1988).
50. A. Galkin, L. Kulakova, T. Yoshimura, K. Soda, and N. Esaki, *Appl. Environ. Microbiol.* **63**, 4651–4656 (1997).
51. B. Nidetzky, D. Haltrich, and K.D. Kulbe, *Chemtech* **1996**, 31–36 (1996).
52. W. Hummel and M.-R. Kula, *Eur. J. Biochem.* **184**, 1–13 (1989).
53. M.-R. Kula and C. Wandrey, *Methods Enzymol.* **136**, 9–21 (1987).
54. W. Hummel, E. Schmidt, C. Wandrey, and M.-R. Kula, *Appl. Microbiol. Biotechnol.* **25**, 175–185 (1986).
55. W. Hummel, H. Schütte, E. Schmidt, C. Wandrey, and M.-R. Kula, *Appl. Microbiol. Biotechnol.* **26**, 409–416 (1987).
56. R. Campagna, and A.F. Bückmann, *Appl. Microbiol. Biotechnol.* **26**, 417–421 (1987).
57. A.S. Gelbard, A.J.L. Cooper, Y. Asano, E. Nieves, S. Filc-DeRico, and K.C. Rosenspire, *Appl. Rad. Isotopes* **41**, 229–233 (1990).
58. T. Ohshima, H. Sugimoto, and K. Soda, *Anal. Lett.* **21**, 2205–2215 (1988).
59. W. Hummel, H. Shütte, and M.-R. Kula, *Anal. Biochem.* **170**, 397–401 (1988).
60. U. Wendel, W. Hummel, and U. Langenbeck, *Anal. Biochem.* **180**, 91–94 (1989).
61. A.J.L. Cooper, L.K.H. Leung, and Y. Asano, *Anal. Biochem.* **183**, 210–214 (1989).
62. U. Wendel, M. Koppelkamm, W. Hummel, J. Sander, and U. Langenbeck, *Clin. Chim. Acta* **192**, 165–170 (1990).
63. U. Wendel, M. Koppelkamm, and W. Hummel, *Clin. Chim. Acta* **201**, 95–98 (1991).
64. K. Dooley, *Clin. Biochem.* **25**, 271–275 (1992).
65. M.A. Vilaseca, C. Farre, and F. Ramon, *Clin. Chem.* **39/1**, 129–131 (1993).
66. S. Keffler, R. Denmeade, and A. Green, *Ann. Clin. Biochem.* **31**, 134–139 (1994).
67. B.C. Cook, T.R. White, C.H. Smith, and M. Landt, *Clin. Chem.* **41**, 949–950 (1995).
68. K. Nakamura, T. Fujii, Y. Kato, Y. Asano, and A.J.L. Cooper, *Anal. Biochem.* **234**, 19–22 (1996).
69. E.W. Randell, and D.C. Lehotay, *Clin. Biochem.* **29**, 133–138 (1996).
70. T. Huang, A. Warsinke, T. Kuwana, F.W. Scheller, *Anal. Chem.* **70**, 991–997 (1998).
71. R. Guthrie and A. Susi, *Pediatrics* **32**, 338–343 (1963).

See also BIOCATAYLSIS DATABASES; PHENYLALANINE.

PHENYLGLYCINES, D-PHENYLGLYCINES

SATOMI TAKAHASHI
Kaneka Corporation
Hyogo, Japan

KEY WORDS

Asymmetric hydrolysis
Decarbamoylation
Dynamic resolution
Hydantoin
Hydantoinase
D-*p*-Hydroxyphenylglycine
D-Phenylglycine
Racemization

OUTLINE

INTRODUCTION

D-Phenylglycines is a general name of a representative group of industrially useful D-α-amino acids. They consist of 2-phenyl and substituted phenylglycine, which have D-form stereo configuration, that is, the opposite stereo configuration of L-amino acids.

D-Amino acids are often called "unnatural" amino acids, and they are not as commonly used as L-amino acids. However, they are widely distributed in various living metabolites and play important roles as physiologically active components.

In recent years, D-amino acids have become increasingly important in the pharmaceutical field as intermediates or chiral synthons for the preparation of β-lactam antibiotics, peptide hormone mimics, pyrethroids, and other compounds. Some D-phenylglycines (D-PGs) are important intermediates for well-known semisynthetic penicillins and cephalosporins, such as ampicillin or amoxicillin, as their side chains.

Among the various D-PGs, D-phenylglycine (D-PG) and D-*P*-hydroxyphenylglycine (D-HPG) are the most important intermediates and are produced and used in large amounts (several thousand metric tons per year) throughout the world. Although they had been produced by classical optical resolution methods until the 1980s, economical and efficient unique synthetic methods have been established that employ asymmetric inversions accomplished chemically (dynamic optical resolution) or enzymatically (hydantoinase process).

STRUCTURE AND NOMENCLATURE

D-PGs are substituted at the second-position proton in glycine (2-amino acetic acid) by phenyl or substituted phenyl radicals to form the D-stereo configuration, as shown in Figure 1. The absolute stereo configuration of the structure is (*R*)-form. D-PGs are enantiomers of L-phenylglycines (L-PGs), which have the reverse stereo configuration.

Representatives of D-PGs include D-phenylglycine (all X = H) and D-*p*-hydroxyphenylglycine (X = *p*-OH), which are also called D-2-amino-2-phenylacetic acid and D-2-amino-2-(4-hydroxyphenyl)acetic acid, respectively.

DERIVATIVES AND USAGE

The practical usage of D-PG and D-HPG is essentially limited to the creation of side chains for various semisynthetic penicillins (SSPs) and semisynthetic cephalosporins (SSCs), as shown Tables 1–3. Among of them, D-PG is used mainly for the side chain of ampicillin and cephalexin, and D-HPG is used mainly for the side chain of amoxicillin and cefatrizine, as shown in Figure 2.

When D-PGs are used for the production of these penicillins and cephalosporins, it is usually as active derivatives such as DANE-salt or acid chloride, which are reacted with 6-aminopenicillanic acid (6-APA, etc.) or 7-amino cephalospollins (7-ACA, 7-ADCA, etc.), as schematically shown in Figure 3.

The DANE-salts of D-PGs are prepared by the reaction of alkali metal salts (Na, K) of D-PGs with methyl acetoacetate (MAA). The acid chlorides of D-PG and D-HPG are prepared by the reaction of D-PGs with chlorinating reagents such as phosphorous pentachloride (PCl_5) or phosgene ($COCl_2$), followed by hydrogen chloride treatment, respectively.

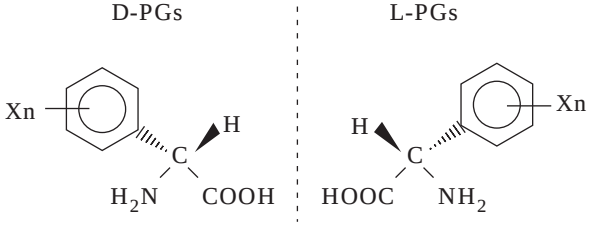


Figure 1. Stereo configurations of D-phenylglycines.

PRODUCTION METHODS

Industrially, D-PGs have been produced by resolving the racemic DL-PGs obtained by chemical syntheses. Typical optical resolution methods are resolution by fractional crystallization of the corresponding diastereomer salt, and by predominant crystallization of the corresponding aromatic sulfuric acid salt. However, these conventional methods are disadvantageous. The former method requires many reaction steps as well as an expensive resolving agent, and the optical purity of the product is not high. The yield obtained in one cycle of the latter method is very low. Furthermore, each method requires a complicated step for racemizing the residual useless L-PGs, and a step for recyclization of the optical resolution in order to increase the overall yield of D-PGs.

To overcome these difficulties, efficient synthetic methods have recently been established by employing unique

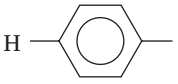
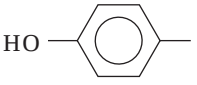
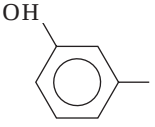
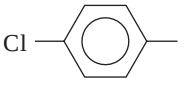
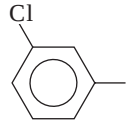
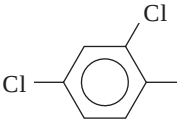
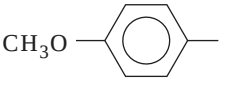
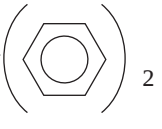
Table 1. D-PG Using Semisynthetic Penicillins (SSP) and Semisynthetic Cephalosporins (SSC)

Type	General name		
SSP	Ampicillin	Apalcillin	Azlocillin
	Bacampicillin	Helacillin	Lenampicillin
	Mezlocillin	Piperacillin	Pirbenicillin
	Pivampicillin	Sarpicillin	Sultamicillin
	Suncillin	Talampicillin	
SSC	Cefaclor	Cefpimizole	Cephalexin
	Cephaloglycin	Loracarbef	Pivcefalexin

Table 2. D-HPG Using Semisynthetic Penicillins (SSP) and Semisynthetic Cephalosporins (SSC)

Type	General name		
SSP	Amoxicillin	Aspocillin	Fumoxicillin
	Piridicillin		
SSC	Cefadroxil	Cefapadole	Cefatrizine
	Cefoperazone	Cefpiramide	Cefprozil

Table 3. Substrate Specificity of Hydantoinase from *Pseudomonas striata*

Substrate		Rerative rate
5- H-	Hydantoin	13
5- CH_3-	Hydantoin	45
5- $(CH_3)_2CH$-	Hydantoin	15
5- $(CH_3)_2CHCH_2$-	Hydantoin	48
5- $CH_3SCH_2CH_2$-	Hydantoin	48
5- 	Hydantoin	25
5- 	Hydantoin	16
5- 	Hydantoin	5
5- 	Hydantoin	19
5- 	Hydantoin	10
5- 	Hydantoin	7
5- 	Hydantoin	4
5,5- 	Hydantoin	0
5,5-$(CH_3)_2$	Hydantoin	0
Dihydrouracil		100

asymmetric transformations that are based on in situ racemization, also called deracemization, of the substrate. These new processes have none of the disadvantages of the optical resolution (conventional) method and are technically and economically advantageous. The overall result of these methods is an asymmetric transformation in which, in principle, 100% yield and 100% ee can be obtained.

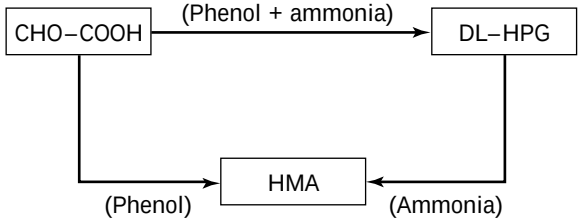
One typical method is based on the dynamic resolution process; the other is the hydantoinase process. These methods are used for the economical production of D-PGs, including D-PG and D-HPG, with good efficiency.

Dynamic Resolution Process

In the conventional optical resolution process, the maximum yield of one cycle is not higher than 50%. However, in the dynamic resolution process, the theoretical resolution yield of D-PGs from DL-PGs can be 100% because in situ racemization of L-PGs occurs through Schiff base tautomerism with the precipitation of the diastereomer salt of the D-PGs plus the resolving agent concomitantly. The mechanism of racemization through Schiff base tautomerism is shown in Figure 4. Some aldehydes or ketones are required as a racemization catalyst to form the Schiff base of the amino acid or its derivatives in an acidic solvent such as acetic acid.

D-PG Production. The dynamic resolution process for D-PG uses DL-phenylglycin amide as the substrare, mandelic acid (MA) as the resolving agent, and benzaldehyde as the racimezation calalyst, followed by acid hydrolysis (1,2). This process is shown schematically in Figure 5.

D-HPG Production. The starting material, DL-HPG, is basically produced from glyoxylic acid, phenol, and ammonia. In DL-HPG preparation, there are two processes: the "one-pot" process and the via-HMA process. In the one-pot process, glyoxylic acid, phenol, and ammonia are reacted at the same time. But the yield is not high (not more than 40%) because of the complex formation of DL-HPG produced with glyoxylic acid under the reaction conditions. So usually, DL-HPG is produced via the reaction of 4-hydroxymandelic acid (HMA), which is produced from glyoxylic acid and phenol, with ammonia or its salts, in good yield (60 to 70%).



DL-HPG can be efficiently resolved into D-HPG by two dynamic resolution processes: (*1*) a dynamic preferential resolution process using arylsulfonic acid (Ar · SO_3H) to precipitate D-HPG · arylsulfonate salt preferentially by the salt seeding (3,4), and (*2*) a dynamic diastereomer resolution process using D-bromocamphorsulfonic acid (BCSA) as the resolving agent (5), as shown in Figure 6. Because of the high reproducibility and high yield, the later process is generally used for commercial D-HPG production. The yield of D-HPG is approximately 90%, and there is almost 100% ee.

Hydantoinase Process

The hydantoinase process, which is an attractive chemicoenzymatic process from both the economical and technical standpoints, was initially developed for the production of various D-amino acids by employing microbial hydantoinase (6). In the process, racemic DL-hydantoins,

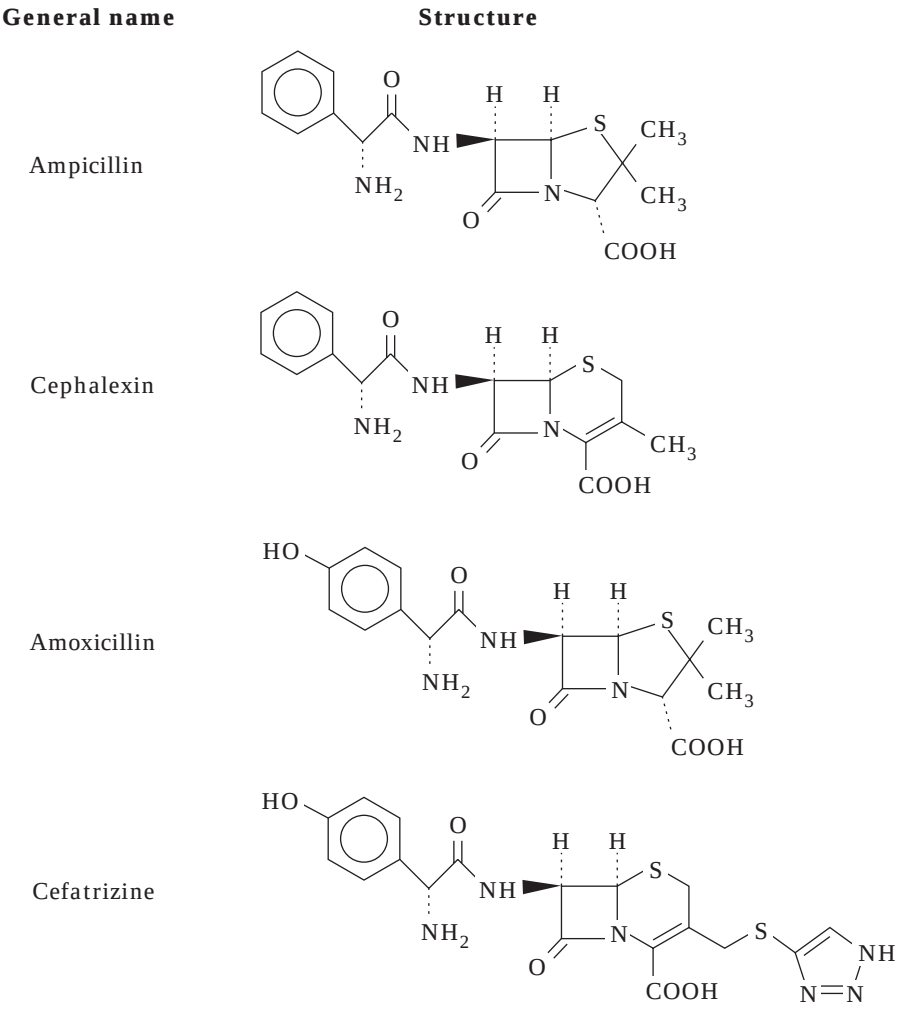


Figure 2. The representative semisynthetic penicillins and semisynthetic cephalosporins.

prepared by a Bucherer-Berg reaction from aldehyde, can be completely transformed into the corresponding optically active *N*-carbamoyl-D-amino acids, because in situ racemization of L-hydantoins occur mainly through keto–enole tautomerism of the hydantoins under the enzymatic reaction conditions. The *N*-carbamoyl-D-amino acids thus obtained can be readily transformed into D-amino acids in good yield by two methods: the chemical method using nitrous acid, or the enzymatic method using *N*-carbamoyl-D-amino acid amidehydrolase.

D-PG can be efficiently produced by the hydantoinase process from benzaldehyde, but the chemical resolution method is usually used for this purpose owing to its easy racemization of the unuseful L-PG.

After the establishment of a new synthetic method for DL-5-(4-hydroxyphenyl)hydantoin (p-HPH), the substrate for D-HPG production, the hydantoinase process was developed mainly for D-HPG production; a practical hydantoinase process was established and began to be used commercially. At present, most commercial D-HPG is produced by this process.

Hydantoinase and D-amino Acid Production. Despite many attempts at L-amino acid production from DL-hydantoins, D-amino acid production from hydantoins was not achieved until the end of the 1970s.

Dudley et al., in the course of studies on the metabolic fate of the anticonvulsants ethotoin and dilantin, found that only the D-form of 5-phenylhydantoin was hydrolyzed into *N*-carbamoyl-D-phenylglycine under the catalysis of animal dihydropyrimidinase (E.C.3,5,2,2), which catalyzed the hydrolysis of dihydropyrimidines to *N*-carbamoyl-*β*-amino acids (7). On the basis of these findings, Cecere et al., studied the specificities of the enzyme and showed that the D-form of several 5-substituted hydantoins were hydrolyzed to provide the corresponding *N*-carbamoyl-D-amino acids (8).

The distribution of hydantoin-hydrolyzing activity in microorganisms was investigated by Yamada and Takahashi (9), and high activity was found in various bacteria belonging to the genera *Pseudomonas*, *Aerobacter*, *Agrobacterium*, *Bacillus*, *Corynebacterium*, and *Actinomycetes*. A hydantoin-hydrolyzing enzyme (hydantoinase) was purified and crystallized from the cell extract of *Pseudomonas striata*, and its properties were established in some detail (10). The bacterial enzyme showed its highest activity and affinity toward dihydrouracil, suggesting its identity with dihydropyrimidinase. The enzyme also catalyzed the

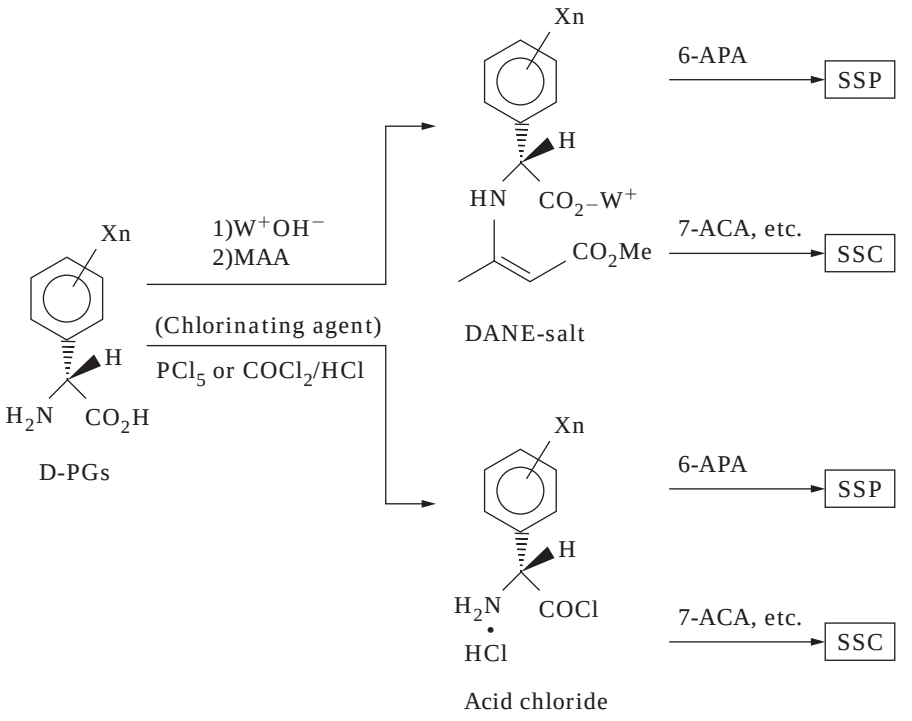


Figure 3. Derivatization of D-PGs and the synthesis of semisynthetic penicillins (SSP) and semisynthetic cephalosporins (SSC).

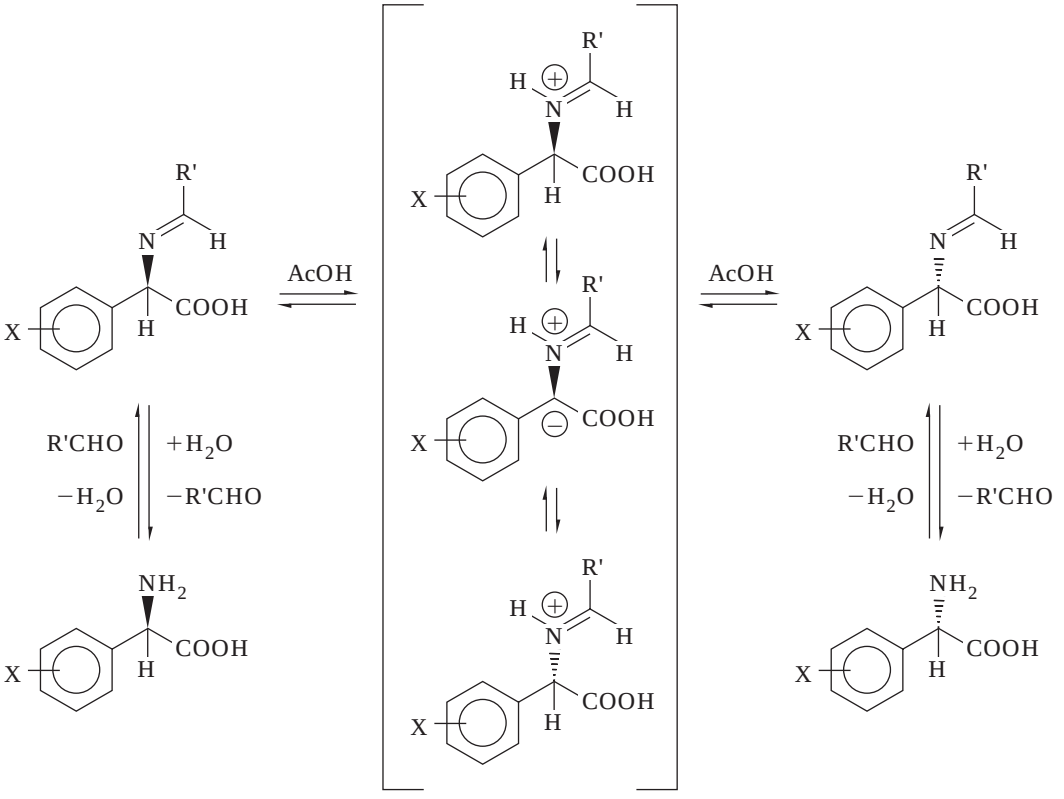


Figure 4. Racemization mechanism.

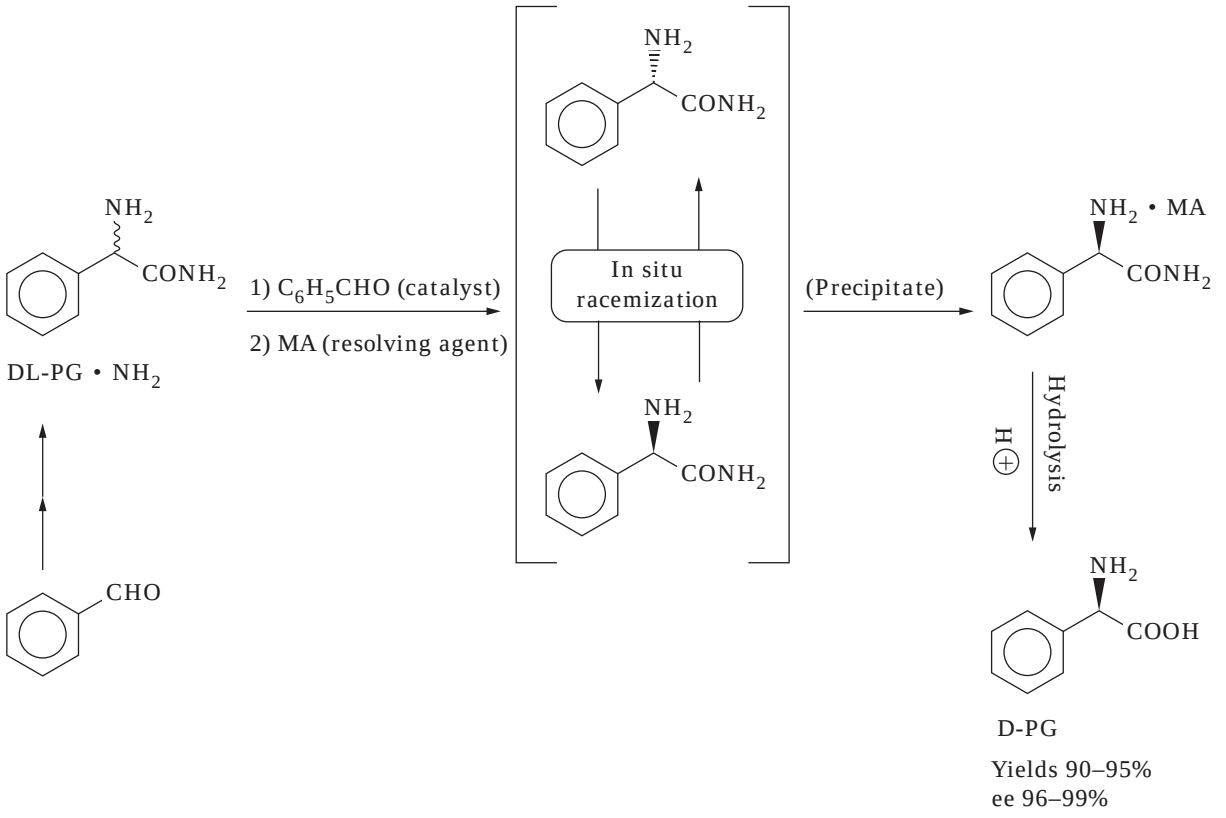


Figure 5. D-PG production.

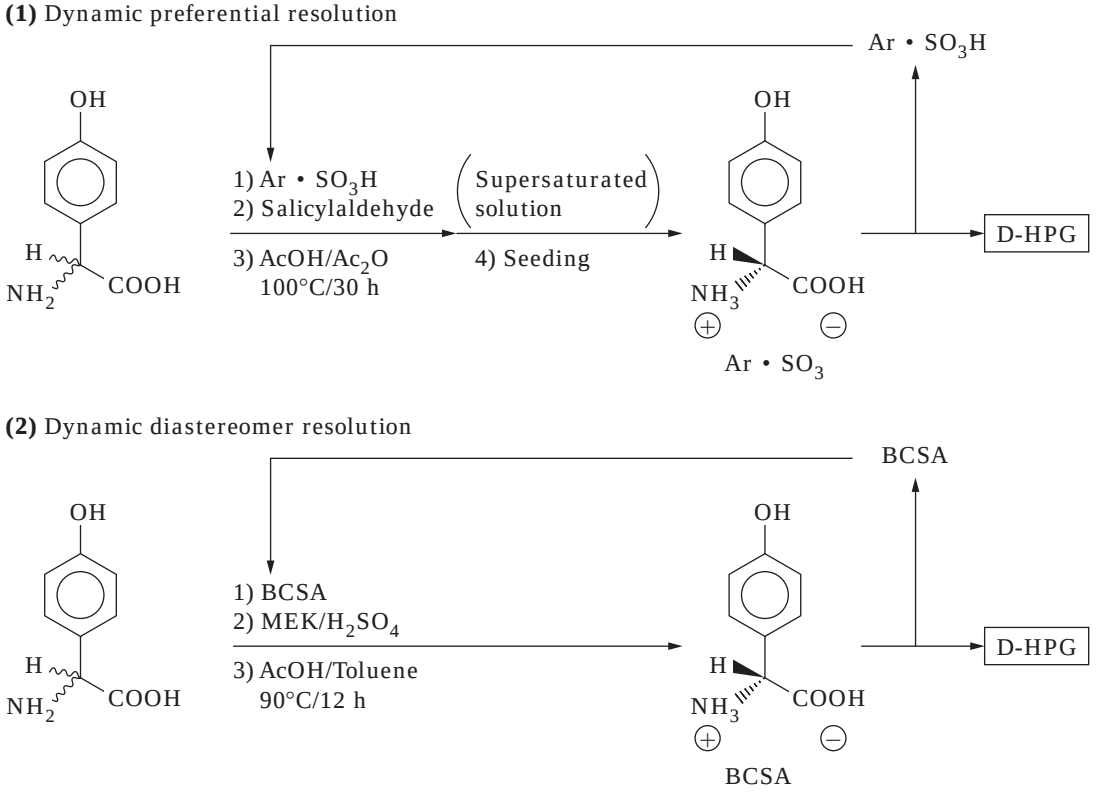


Figure 6. Resolution of DL-HPG into D-HPG.

hydrolysis of a variety of 5-substituted-D-hydantoins, as shown in Table 3. The D-form of 5-aliphatic and aromatic hydantoin such as 6-phenyl, 5-(4-hydroxyphenyl), and 5-thienylhydantoin were hydrolyzed easily to the corresponding *N*-carbamoyl-D-amino acids. However, other hydantoins having a charged group in the amino acid moieties were generally resistant to hydrolysis.

The *N*-carbamoyl-D-amino acids thus obtained were quantitatively transformed to D-amino acids by treating them with nitrous acid under acidic conditions. These results suggested that the combination of the D-stereoselective hydantoinase reaction and the decarbamoylation by nitrous acid was useful as a simple method of D-amino acid production (6).

This chemicoenzymatic process consists of two chemical steps (steps 1 and 3) and one enzymatic step (step 2), as follows:

1. Hydantoin synthesis (generally according to Bucherer–Berg synthesis)
2. Transformation of hydantoin to *N*-carbamoyl-D-amino acid (microbial hydantoinase)
3. Decarbamoylation of *N*-carbamoyl-D-amino acid (nitrous acid treatment)

D-HPG Production. The hydantoinase process for D-HPG production consists of three steps (11), as follows:

1. Hydantoin (5-(4-hydroxyphenyl)hydantoin:p-HPH) synthesis from glyoxylic acid (GOA)
2. Transformation of the hydantoin into carbamoyl-D-HPG using hydantoinase
3. Decarbamoylation of carbamoyl-D-HPG into D-HPG

However, three sub-types are differentiated in the step 3 decarbamoylation:

1. Chemical decarbamoylation using sodium nitrite
2. Enzymatic decarbamoylation using immobilized *N*-carbamoyl-D-amino acid amidehydrolase
3. Enzymatic decarbamoylation using bacterial cells that contain hydantoinase and *N*-carbamoyl-D-amino acid amidohydrolase

A summary is shown in Figure 7.

Synthesis of p-HPH. p-HPH is an important intermediate in the hydantoinase process for D-HPG production. It is known that p-HPH can be synthesized by the reaction of 4-hydroxybenzaldehyde, ammonium carbonate, and sodium cyanide, according to Bucherer–Berg's method. However, this method requires the use of dangerous sodium cyanide and expensive aldehyde. Furthermore the crude hydantoin obtained is contaminated by impurities caused by the oxidative side reaction of the phenol nucleus in an alkaline condition, or it may take on a brownish color.

A novel and unique method for preparing p-HPH involving amidoalkylation of phenol with glyoxylic acid derivatives was established (12). According to this method, p-HPH can be readily prepared in high purity and good yield from relatively inexpensive raw materials such as phenol, glyoxylic acid, and urea.

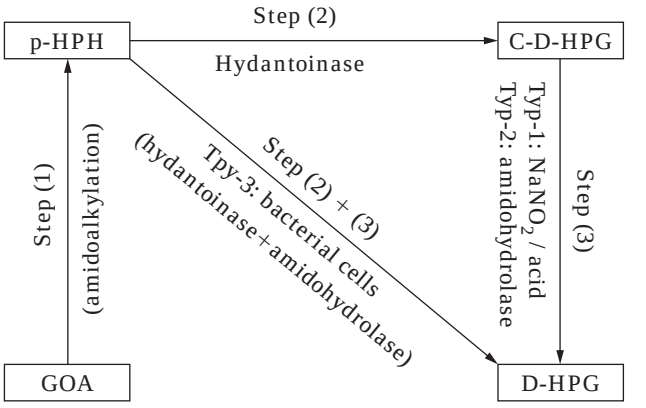
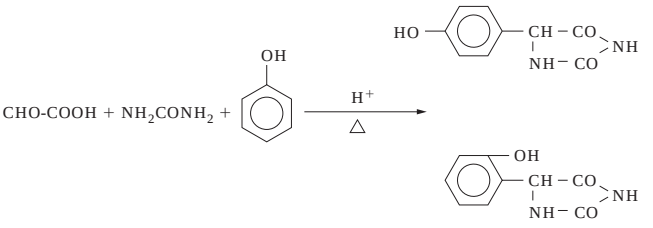


Figure 7. Hydantoinase process of HPG production.



The main product of this reaction is p-HPH, but its isomer, 5-(2-hydroxyphenyl)hydantoin (o-HPH), is produced concomitantly. For the efficient synthesis of p-HPH, the ratio of p-HPH to o-HPH and total HPH yield should be increased concurrently. However, practically speaking, because o-HPH is hard to crystallize and more soluble, p-HPH can be easily obtained in a high state of purity by merely precipitating and separating it from the reaction mixture.

More recently, various methods for p-HPH synthesis have been investigated using glyoxilic acid derivatives. Among them, the method via *p*-hydroxymandelic acid, which is prepared from glyoxylic acid and phenol, seems promising because by-product o-HPH preparation can be minimized.

Transformation of p-HPH into N-Carbamoyl-D-HPG. *Enzyme (hydantoinase) preparation.* Hydantoinase-producing microorganisms can be found in a wide range of genera; high activity is found particularly in bacteria. The stable hydantoinase from thermophilic microorganisms is useful for high productivity.

Cells with high hydantoinase activity can be obtained by culturing the microorganism in a medium supplemented by pyrimidine base or a metabolite such as uracil, thymine, or *β*-alanine as the inducer. The accumulation of hydantoinase in the cells is further increased when a metal ion such as manganese is added to the medium together with the inducer.

Recently, to drastically increase hydantoinase levels, genetic engineering methods were attempted and resulted in reports of 4- to 40-fold higher enzyme activity with an

overexpression plasmid vector. Hydantoinase from thermophilic microorganisms such as thermophilic *Bacillus* seem to be very useful because their stability is rather higher than that of mesophilic microorganisms.

Asymmetric hydrosis of p-HPH. In the asymmetric hydrosis of p-HPH, racemic p-HPH can be completely transformed into *N*-carbamoyl-D-HPH by the action of hydantoinase. Technically and economically, this is a great advantage for the industrial production of D-HPG. The mechanism of this interesting reaction seems to be as follows.

It is well known that hydantoins are readily racemized in dilute alkaline solution through base-catalyzed tautomerism. In practice, p-HPH undergoes spontaneous racemization very easily under mild conditions such as those of the enzymatic reaction. In the reaction system for the asymmetric hydrolysis of p-HPH, only the D-form of p-HPH is susceptible to enzymatic hydrolysis. Unreactive L-p-HPH undergoes rapid spontaneous racemization in the same system. However, the *N*-carbamoyl-D-HPG formed is never racemized under these conditions. Consequently, in this system the enzymatic hydrosis of the hydantoin ring and chemical racemization of the substrate proceed simultaneously, so that DL-p-HPH can be completely transformed into the D-form of *N*-carbamoyl-HPG (Figure 8).

Hydantoinase is employed via intact whole cells or immobilized enzyme resin. In the immobilized form, the stable hydantoinase from a thermophilic microorganism such as thermophilic *Bacillus* is very useful for good performance even after repetitive use.

The concentration of the substrate DL-p-HPH that is available to the reaction depends on the enzyme activity used. A large portion of p-HPH is present in suspended form because the solubility of p-HPH in water is very low (ca. 50 mM). However, this is not an obstacle to the reaction, because the substrate is successively dissolved during the progress of the reaction in alkaline pH (ca. pH9). It is preferable to maintain the pH by adding alkaline solution successively, because the pH lowers in the course of hydrolysis; a drop in pH will result in a lowered reaction rate. It is also effective to cover the reaction mixture with an inert gas such as nitrogen to avoid the oxidative side reaction of the phenol nucleus. Under these optimum conditions, the yield of *N*-carbamoyl-D-HPG formed is almost quantitative.

Decarbamoylation of N-Carbamoyl-D-HPG *N*-Carbamoyl-D-HPG produced by enzymatic hydrolysis can be readily converted into D-HPG by decarbamoylation with nitrous acid under acidic conditions. The principle of this oxidative reaction is based on the Van Slyke determination, and the reaction seems to be a consecutive reaction. With respect to the stereochemistry of the reaction, the retention of the configuration is achieved completely. Therefore, optically pure D-HPG can be readily obtained in good yield. The decarbamoylation is preferably carried out by reacting *N*-carbamoyl-D-HPG with approximately equimolar nitrous acid in an aqueous medium in the presence of a strong mineral acid such as sulfuric or hydrochloric acid. It is convenient to employ a water-soluble salt of nitrous acid such as sodium nitrite or potassium nitrite. Because this decarbamoylation is an exothermic reaction and generates large quantities of gas (N_2 and CO_2), an aqueous solution of nitrous acid is added gradually to the reaction medium with agitation to facilitate a smooth reaction. The reaction temperature is usually kept below 20

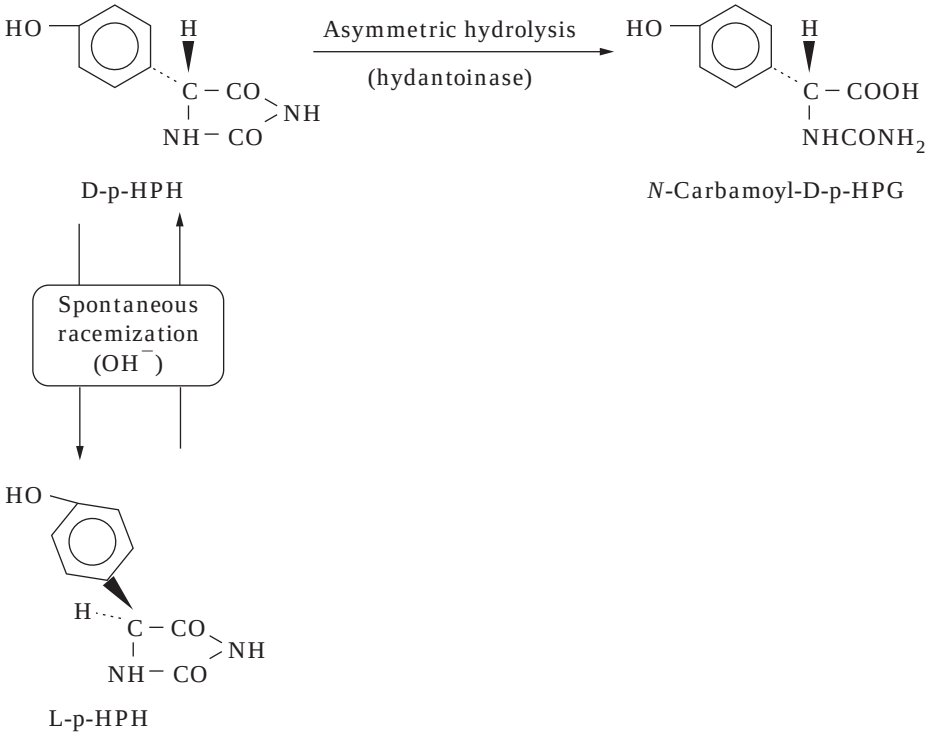
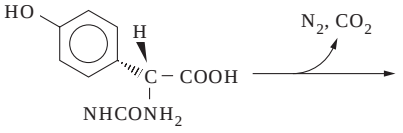
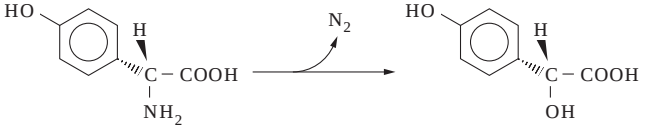


Figure 8. The mechanism of asymmetric hydrolysis of DL-p-HPH.

°C to avoid a side reaction such as the further degradation of D-HPG into 4-hydroxymandelic acid and other compounds. Under these optimum conditions, the yield of D-HPG from *N*-carbamoyl-D-HPG is almost quantitative.



N-Carbamoyl-D-HPG



D-HPG 4-Hydroxybenzaldehyde

BIBLIOGRAPHY

1. Dutch Pat. 900.0386 (September 16, 1991), W.H. Boesten and W.H.J. Boesten (to Stamicarbon BV and DSM NV).
2. Dutch Pat. 900.0387 (September 16, 1991), W.H. Boesten and W.H.J. Boesten (to Stamicarbon BV and DSM NV).
3. C. Hongo, R. Yoshioka, M. Tohyama, S. Yamada, and I. Chibata, *Bull. Chem. Soc. Jpn.* **56**, 3744–3747 (1983).
4. C. Hongo, R. Yoshioka, M. Tohoyama, S. Yamada, and I. Chibata, *Bull. Chem. Soc. Jpn.* **85**, 433–436 (1985).
5. Eur. Pat. 173921 (March 12, 1986), V.W. Jacewixz (to Beecham Group PLC).
6. S. Takahashi, T. Ohashi, Y. Kii, H. Kumagai, and H. Yamada, *J. Ferment. Technol.* **57**, 328–332 (1979).
7. K.H. Dudley, T.C. Butler, and D.L. Bius, *Drug Metab. Disp.* **2**, 103–112 (1974).
8. F. Cecere, G. Galli, and F. Morisi, *FEBS Lett.* **57**, 192–194 (1975).
9. H. Yamada, S. Takahashi, Y. Kii, and H. Kumagai, *J. Ferment. Technol.* **56**, 484–491 (1978).
10. S. Takahashi, Y. Kii, H. Kumagai, and H. Yamada, *J. Ferment. Technol.* **56**, 492–498 (1978).
11. K. Aida, I. Chibata, K. Nakayama, K. Takinami, and H. Yamada, *Progress Industrial Microbiology, Vol. 24, Biotechnology of Amino Acid Production*, Kodansha, Tokyo, 1986, pp. 269–279.
12. T. Ohashi, S. Takahashi, T. Nagamachi, K. Yoneda, and H. Yamada, *Agric. Biol. Chem.* **45**, 831–838 (1981).

PICHIA, OPTIMIZATION OF PROTEIN EXPRESSION

KOTI SREEKRISHNA
Procter and Gamble Co.
Ross, Ohio

KEY WORDS

Alcohol (methanol) oxidase promoter
Clonal variation
Expression plasmid
Fermentation
Methylotrophic yeast
mRNA secondary structure
Protease deficient strain
Secretion
Synthetic gene
Transcriptional barrier

OUTLINE

INTRODUCTION

Since the advent of the recombinant DNA technology, a wide variety of expression systems for protein production have become available. These include systems based on bacteria, lower eukaryotes (yeast and fungi), invertebrates (cells or larvae), vertebrates (cells or whole animal), and plants (cells or whole plant). Of these, only the yeast and fungal expression systems have many of the attributes of both the bacterial and higher eucaryotic expression systems. Yeasts have the robust growth characteristics and ease of manipulation of a bacterial system and at the same time perform many posttranslational modifications found in higher eukaryotes. The purpose of this article is to highlight the strategies one can apply to optimize protein expression in the *Pichia* yeast expression system.

BACKGROUND

Pichia pastoris is a methylotrophic yeast. It is able to use methanol for energy as well as for growth. It possesses a highly regulated pathway for the utilization of methanol (Fig. 1). Synthesis of the enzymes of methanol metabolism increase rapidly once the cells are placed in the methanol medium. The most dramatic effect is seen for methanol oxidase, which accounts for 35% of the cellular protein in cells adapted to grow on methanol. An extensive proliferation of the peroxisomes, known to sequester methanol oxidase and dihydroxy acetone synthase, is also observed in methanol grown cells (1).

Pichia was initially investigated at Phillips Petroleum Company as a potential source of single cell protein. Thus, a very efficient fermentation process (cell density $>$ 130 g dry cell weight per L and biomass productivity $>$ 10 g/L-h) was developed (2). Though impressive, this process could not compete with the economics of production of soy protein. After this setback, Phillips Petroleum Company directed its efforts in developing *Pichia* as an expression system for the production of recombinant proteins. This has turned out to be a worthwhile endeavor.

STRATEGIES FOR OPTIMIZATION OF PROTEIN EXPRESSION

Typical *Pichia* expression vectors (Fig. 2), for example, pHIL-A1 (3), pHIL-D2 (4), pHIL-D7 (5) and pPIC9 (6) are based on the strong methanol oxidase gene (*AOX1*) promoter (7). A wide variety of proteins have been produced in this system (Table 1). The final yield of a protein expressed in *Pichia* is largely influenced by its inherent properties. Nevertheless, the production yield of a protein can be significantly enhanced by using a strategy that takes into account the multiple factors that influence protein expression.

Cellular State of the Expression Cassette

The expression cassette can be introduced into *Pichia* cells by way of chromosomal integration or autonomous replication. However, the integrative approach is preferable because it has the following advantages: (*1*) expression cassette stability; (*2*) ability to generate clones that contain multiple copies of the expression cassette (see section on gene dosage); (*3*) control over the sites of integration (*HIS4* or *AOX1* loci); and (*4*) ability to engineer different modes of integration (with or without the eviction of the *AOX1* coding sequences) (see section on methanol utilization phenotype).

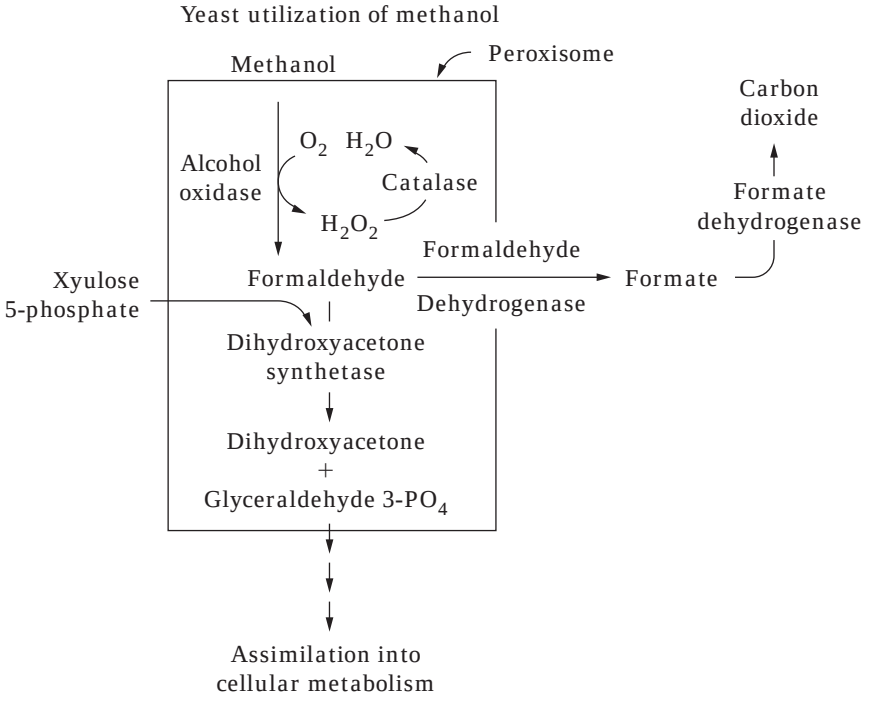


Figure 1. Methanol utilization pathway in methylotrophic yeasts.

Plasmids based on autonomous vectors such as pHIL-A1 (Fig. 2) are maintained in *Pichia* cells at low copies and are rapidly lost from the population of dividing cells. In a small number of cells, autonomous plasmids do eventually integrate. Thus, to arrive at a stable clone, a secondary selection screen should be implemented. This makes autonomous plasmids less attractive for protein expression. However, owing to their ability to transform *P. pastoris* spheroplasts at a very high frequency, the autonomous plasmids are useful for cloning genes in *Pichia* by functional complementation.

Site of Integration of the Expression Cassette

Both *AOX1* and *HIS4* sites have been used successfully for expression. Occasional loss of the expression cassette at the *HIS4* locus because of intrachromosomal crossover between mutant *his4* and the good *HIS4* has been observed. Thus, the *AOX1* locus is the preferred site for stable expression.

Methanol Utilization Phenotype, Mut^+ or Mut^- of the Host

Transformation of a *P. pastoris his4* strain using linear DNA with the ends bearing homology to the 5′ and 3′ regions of the *AOX1* chromosomal locus (e.g., *Bgl*II digested pHIL-D7, or pPIC9 or *Not*I-digested pHIL-D2 [Fig. 2]) results in the site-specific eviction of the *AOX1* structural gene, as illustrated in Figure 3. The eviction of the *AOX1* occurs at a frequency of 1 to 5 per 20 His^+ transformants. *AOX1*-deleted clones show a slower growth phenotype (Mut^-) on minimal methanol medium, as compared to *AOX1* intact clones, which have phenotypically normal growth characteristics (Mut^+) on minimal methanol medium. The Mut^+ transformants arise presumably because of circularization of the linear DNA inside the yeast cell and subsequent integration into one or more of the *AOX1* or *HIS4* chromosomal loci. Thus, in a single transformation experiment, both Mut^+ and Mut^- transformants can be obtained.

For intracellular expression, it is preferable to use Mut^- cells because they will have a lower level of alcohol oxidase protein and the expressed protein can be more readily purified. For secretion, either one of Mut^+ or Mut^- constructs can be used. There is no significant difference between these two types of host cells in the secretion of human serum albumin (74,75).

Gene Dosage: Exploiting the Clonal Variation of Expression

In many instances of *Pichia* expression of heterologous proteins, namely, expression of LACZ (8), hepatitis B surface antigen (9), invertase (33), and human serum albumin (HSA) (4), a single copy of the expression cassette was sufficient for optimal production. Deliberately increasing the copy number had no significant effect on production in these instances. Yet, in other cases, multiple copies of the expression cassette (> 10) are a must for high-level expression. A dramatic effect of copy number on protein production in *P. pastoris* is seen in the production of human tumor necrosis factor (TNF), tetanus toxin fragment C, *Bordetella pertussus* pertactin P69, and mouse epidermal growth factor (EGF) (6,13,14,74,76). In some rare instances, an increase in copy number has a negative effect on the production level (16).

Because the effect of gene copy number on expression is unpredictable, it is necessary to examine the production level as a function of gene dosage. The spheroplast method of transformation of *P. pastoris* results in transformants with a wide range of copy numbers (14,74,76). Analysis of as few as 100 individual clones for protein production is generally adequate to arrive at a good producer. If other methods of transformation (e.g., LiCl method or electroporation) that do not give rise to multicopy transformants at a high frequency are used, then more efficient screens can be used, such as colony hybridization (or dot blot analysis) with DNA probes or selection for multicopy integrants based on increased level of resistance to antibiotic G418 (77) or Zeocin (Invitrogen, San Diego, Calif.).

Alternatively, multicopy constructs of the expression cassette can be obtained by transformation with DNA concatamers or by using multicopy construction vectors such as pAO815 (3,16). This is not a method of choice because it permits production analysis over only a narrow range of gene dosage.

Translational Optimization: 5′ Untranslated Region

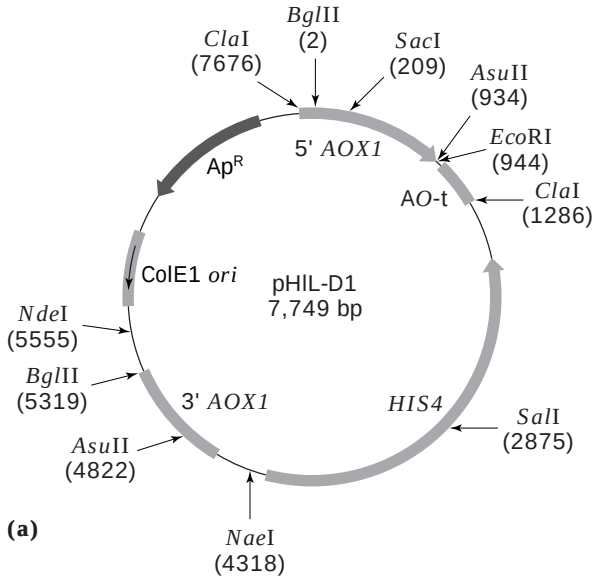
The nucleotide (nt) sequence and the length of the 5′ untranslated region (5′ UTR) can be detrimental to optimal protein translation. The leader length of the highly expressed *AOX1* mRNA is 114 nt long, and the sequence is A + U rich (7). For optimal synthesis of heterologous proteins, it is essential that the 5′ UTR of the expression cassette should closely resemble that of the *AOX1* mRNA. Ideally, it is preferable to make it identical to that of the *AOX1* mRNA. The expression level of HSA is increased more than 50-fold by adjusting the 5′ UTR to be closer to that of the *AOX1* mRNA. Expression plasmid such as pHIL-D7 can be used to readily make 5′ exact constructs.

Translational Optimization: Initiation Codon AUG Context

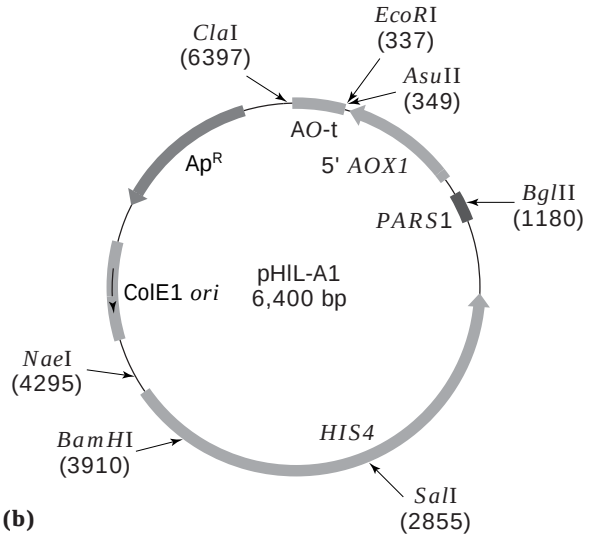
The translation initiation codon AUG should be avoided in the 5′ UTR to ensure efficient translation of mRNA from the actual translation start site. The mRNA secondary structure around the initiator AUG may be adjusted, so that AUG is relatively free of secondary structure (74). This can be accomplished by redesigning the initial portion of the coding sequences with alternate codons (3,74).

Transcriptional Optimization

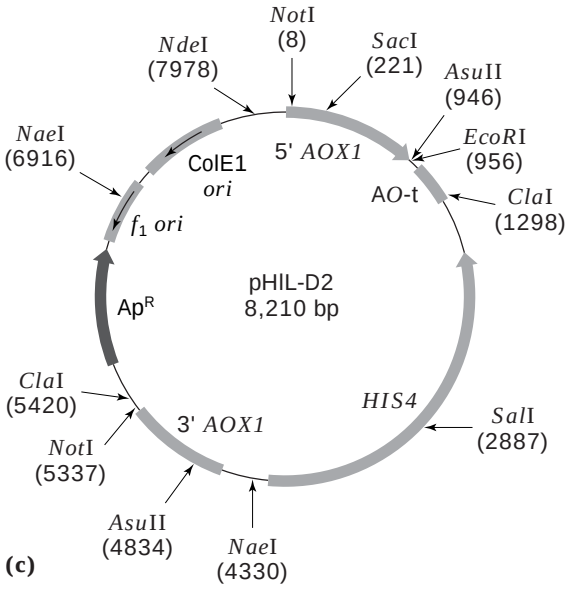
Genes with high A + T nucleotide clusters are inefficiently transcribed because of premature termination. One such sequence, ATTATTTTATAAA, present in HIV-gp120 has been identified to block transcription in *P. pastoris*, and when this stretch is altered to TTTCTTCTACAAG, the premature termination is abolished (17). Because there are many yet unidentified AT-rich stretches that act as transcriptional terminators, a general strategy would be to redesign the genes using *P. pastoris* preferred codons (3,74) so as to have an A + T content in the range of 30 to 55%. By using this approach, it has been possible to con-


ClaI (7676)
BglII (2)
SacI (209)
AsuII (934)
EcoRI (944)
ClaI (1286)
5' AOX1
AO-t
Ap^R
ColE1 ori
pHIL-D1
7,749 bp
NdeI (5555)
BglII (5319)
3' AOX1
HIS4
SalI (2875)
AsuII (4822)
NaeI (4318)

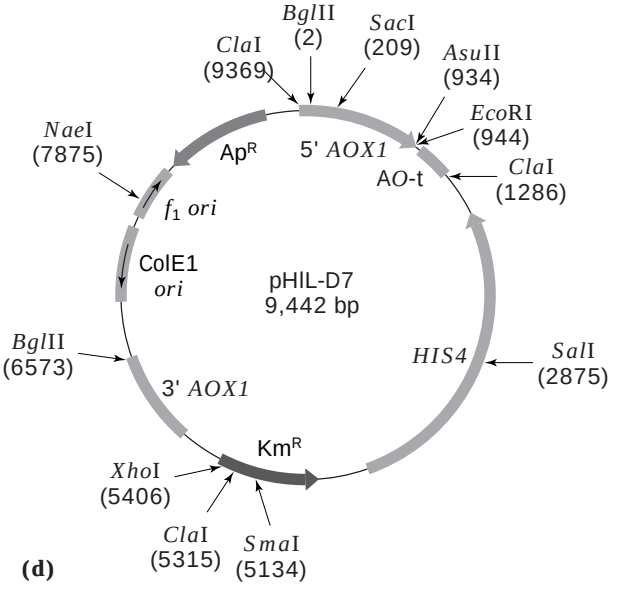
(a)


ClaI (6397)
EcoRI (337)
AsuII (349)
AO-t
5' AOX1
Ap^R
PARS1
BglII (1180)
ColE1 ori
pHIL-A1
6,400 bp
NaeI (4295)
BamHI (3910)
HIS4
SalI (2855)

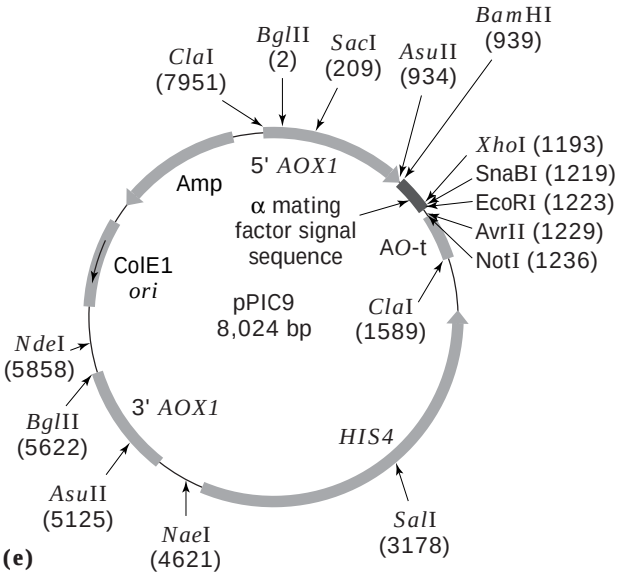
(b)


NotI (8)
NdeI (7978)
SacI (221)
AsuII (946)
EcoRI (956)
NaeI (6916)
ColE1 ori
5' AOX1
AO-t
ClaI (1298)
f1 ori
pHIL-D2
8,210 bp
Ap^R
ClaI (5420)
HIS4
3' AOX1
SalI (2887)
NotI (5337)
AsuII (4834)
NaeI (4330)

(c)


BglII (2)
SacI (209)
ClaI (9369)
AsuII (934)
EcoRI (944)
NaeI (7875)
Ap^R
5' AOX1
AO-t
ClaI (1286)
f1 ori
ColE1 ori
pHIL-D7
9,442 bp
BglII (6573)
HIS4
SalI (2875)
3' AOX1
Km^R
XhoI (5406)
ClaI (5315)
SmaI (5134)

(d)


BamHI (939)
BglII (2)
SacI (209)
AsuII (934)
ClaI (7951)
XhoI (1193)
SnaBI (1219)
EcoRI (1223)
AvrII (1229)
NotI (1236)
5' AOX1
Amp
α mating factor signal sequence
AO-t
ColE1 ori
pPIC9
8,024 bp
ClaI (1589)
NdeI (5858)
3' AOX1
BglII (5622)
HIS4
AsuII (5125)
NaeI (4621)
SalI (3178)

(e)

struct strains for efficient production of tetanus toxin fragment C (13) and *Bacillus sphaericus* mosquitocidal toxins (BSP1 and BSP2) (3,15).

Product Secretion

If a protein can be secreted, it is the desired mode of protein production because of the ease of product recovery. Furthermore, certain normally secreted proteins, such as HSA and growth hormones, remain insoluble when expressed intracellularly. For a protein that is not normally secreted, it may be difficult to cause secretion. However, there are some encouraging results with respect to yeast secretion of cytochrome P-450 (78) and rat liver elongation factor (55) that suggest that at least in some instances it may be possible to secrete a normally intracellular protein.

The choice of secretion signal is rather arbitrary (Table 1). In several cases, (HSA, invertase, bovine lysozyme, etc.), the native signal sequence is adequate. We have found that the native secretion signals present in matrix metalloproteinases 1,2,3,9 (MMP-1,2,3,9) and tissue inhibitor of matrix metalloproteinases 1 (TIMP-1) are functional in the *P. pastoris* system. If a native secretion signal sequence is not available, then a signal sequence based on the *S. cerevisiae* invertase (Suc2p), the *S. cerevisisae* α-mating factor (αMF), or the *P. pastoris* acid phosphatase (Pho1p) can be used.

The pre-pro αMF signal works very efficiently, especially for the secretion of a large variety of proteins, including the smaller-sized products such as aprotinin, EGF, thrombomodulin fragment, blood factor XII, a fragment of amyloid β-protein, antibody single-chain Fv fragment, and ghilanten (Table 1). In making protein fusions to the αMF signal, it is prudent to retain the Glu-Ala spacers adjacent to the Kex2-like protease cleavage site: (. . . V-S-S-L-E-Lys-Arg-^{vKex2p}Glu-Ala-^{vDapp}Glu-Ala-^{vDapp}fused protein). The presence of the Glu-Ala spacers help to alleviate the steric interference by the fused protein, resulting in an efficient cleavage of the pro sequence by the *P. pastoris* Kex2-like protease (16). The Glu-Ala spacer sequence is subsequently cleared by a diamino peptidase (Dapp) to yield the protein of interest free of additional N-terminal amino acid residues. In spite of taking this precaution, we have noticed that the pro sequence is not cleaved when the αMF signal is fused with human brain-derived neurotrophic factor (BDNF) (mol. wt. $\sim$ 14 kDa). In this situation, BDNF is secreted into medium as a 26 to 30 kDa product, which is the expected result if the pro portion of the αMF signal is not cleaved and is heterogeneously glycosylated at the multiple glycosylation sites present in the αMF pro peptide (5).

Acid phosphatase (Pho1p) secretion signal sequence, invertase secretion signal, and a hybrid sequence consisting of Pho1p secretion signal containing a Kex2 protease recognition sequence has also been used successfully for high level (>g/L) secretion of certain proteins (Table 1).

Choice of the Secretion Signal

In most of the examples listed in Table 1, the relative efficacies of the different secretion signals for a given product have not been examined. Only in a limited number of cases has a rough comparison been made between native signal versus pre-pro αMF. For example, in the case of invertase secretion, both the extent of glycosylation and secretion kinetics are enhanced when the invertase signal sequence

Figure 2. Typical *Pichia* expression vectors. (**a**) pHIL-D1(3) is an *E. coli–P. pastoris* shuttle vector, with sequences required for selection in each host. The left half of the plasmid is a portion of pBR322, from *Cla*I (at position 7676 in the map) to *Pvu*II site (modified to *Bgl*II [at nt 5319 in the map]). This segment of pBR322 contains the ampicillin-resistance encoding gene (Ap^R) and the origin of replication (ColE1 *ori*). The *Eco*RI site in this segment has been eliminated. The DNA elements comprising the rest of the plasmid are derived from the genome of *P. pastoris*, except for short regions of pBR322 used to link the *Pichia* sequences. The *P. pastoris* elements in the plasmid are as follows: (*1*) 5′ *AOX1*, approximately a 1,000 bp of the alcohol oxidase promoter. The *AOX1* sequences following the nt A of the ATG start codon have been removed by BAL 31 treatment and the synthetic linker 5′-GGAATTC added to generate a unique *Eco*RI cloning site. (*2*) AO-t, approximately a 300 bp of the *AOX1* terminating sequence. (*3*) *P. pastoris* histidinol dehydrogenase gene, *HIS4*, contained on a 2,800-bp fragment to complement the defective *his4* gene in the *P. pastoris* strains such as GS115, SMD1168, etc. (*4*) Region of 3′*AOX1* DNA approximately 650 bp, which together with the 5′*AOX1* region is necessary for the site-directed displacement of *AOX1* coding sequences (see Fig. 3). The unique *Sac*I (at nt 209 in the map) and *Sal*I (at nt 2875 in the map) sites present in the plasmid can be used for site-directed integration of the entire plasmid into the *AOX1* and *HIS4* loci of *P. pastoris*, respectively. (**b**) Plasmid pHIL-A1 is an *E. coli–P. pastoris* shuttle vector, with sequences required for selection and autonomous replication in each host. The left half of the plasmid is a modified portion of pBR322 containing the Ap^R gene and the ColE1 *ori*. The DNA elements comprising the rest of the plasmid are derived from the genome of *P. pastoris*, except for short segments of pBR322 used to link the yeast sequences. The various *P. pastoris* sequences are as follows: (*1*) 5′*AOX1*, approximately 750-bp portion of the *AOX1* promoter with the *Eco*RI (at position 337 in the map) engineered as previously described (Fig. 1). (*2*) AO-t and *HIS4* are same as in Figure 1. (*3*) *PARS1*, approximately 190-bp segment of a *P. pastoris* autonomous replication sequence. The unique *Sal*I (at nt 2885 in the map) site can be used to direct the integration of pHIL-A1 into the *HIS4* loci of *P. pastoris*. (*c*) pHIL-D2 (4) has a 458-bp DNA containing the f_1-*ori* inserted between the original *Dra*I sites of pBR322. The *Not*I sites at nt 8 and 5337 are used for site-directed displacement of *AOX1* coding sequences. The other details are the same as described for pHIL-D1 (**a**). (**d**) pHIL-D7 is a shuttle plasmid containing the kanamycin-resistance-encoding gene (Km^R) and is useful for selection of multicopy transformants based on increased level of resistance to G418. Unlike the commonly used vectors such as pHIL-D1 and pHIL-D2, we have designed pHIL-D7 so that it has a unique *Asu*II site at nt position 934 (the second *Asu*II site present in the 3′*AOX1* has been eliminated). Therefore, the sequence TTCGAAACG can be added immediately upstream of the start codon (ATG) of the gene of interest, and an *Eco*RI site can be added downstream of the stop codon. The modified gene can then be inserted between the *Asu*II (at nt 934 in the map) and *Eco*RI (at nt 944 in the map) region TTCGAAACGAGGAATTC of pHIL-D7. Thus, the modified gene will have identical 5′ UTR to that present in *AOX1*. The other details are same as described for pHIL-D1 (**a**). (**e**) pPIC9 (6), a secretion vector with αMF secretion signal and multicloning sites. The other details are same as for pHIL-D1 (**a**).

Table 1. Partial List of Heterologous Proteins Expressed in *P. pastoris*

Intracellular expression

E. coli β-galactosidase (LACZ) (8)
Hepatitis B surface antigen (HBsAg) (9)
Human tumor necrosis factor-α (TNF-α) (10)
Human TNF-α analogues (11)
Salmon growth hormone (SGH) (K. Sreekrishna and K. A. Parker, unpublished observations, 1987)
Streptokinase (12)
Tetanus toxin fragment C (13)
Bordetella pertussus pertactin P69 (14)
Human inteleukin 2 (W. R. McCombie, personal communication, 1988)
B. sphaericus mosquitocidal components, BSP1, BSP2, and BSP1 + BSP2 (15)
Superoxide dismutase (SOD) (16)
Human serum albumin (HSA) (4)
HIV gp120 (17)
Porcine leukocyte lipoxygenase (18)
$p110^{RB}$ retinoblastoma tumor suppressor protein 919
Hamster prion protein (20)
Spinach phosphoribulokinase (21)
Spinach cytochrome C oxidase (22)
Spinach glycolate oxidase (23)
Dogfish shark cytochrome P450c17 (hydroxylase and lyase activities) (24)
Algal phytochrome (25)
Green fluorescent protein (GFP) (26)
Arabidopsis chaperone protein ATJ2 (27)
Corn cytochrome *b* and cytochrome *c* reductases (28)
Mouse thioether methyltransferase (29)
Polyomavirus large T antigen (30)
Rat liver mitochondrial carnitine palmitoyltransferase I and II (31)
Mouse $5HT_{5A}$ serotonin receptor (32)

Secretion using native secretion signal sequence

S. cerevisiae invertase (*SUC2*) (33)
Human serum albumin (HSA) (4)
Bovine lysozyme (34)
Matrix metalloproteinases MMP-1,2,3, and 9 (5)
Tissue inhibitor of matrix metalloproteinases-1 (TIMP-1) (5)
MMP-1,2,3, or 9 coexpressed with TIMP-1 (5)
Barley α-amylases 1 and 2 (35)
Cathepsin E (36)
D-alanine carboxypeptidase (dacA) of *Bacillus stearothermophilus* (37)
Pectate lyase pel B from *Fusarium solani f.* sp. *pisi* (38)
IgE receptor (α subunit) (39)

Secretion using pre–pro α *mating factor (*α*MF) secretion signal sequence*

Aprotinin (16)
S. cerevisiae invertase (*SUC2*) (16)
Human epidermal growth factor (EGF) (16)
Mouse epidermal growth factor (EGF) (6)
Insulin-like growth factor-1 (IGF-1) (40)
Kunitz protease inhibitor (nexin-2) (41)
Human vascular endothelial factor (VEGEF) (42)
κ-Bungaratoxin (43)
Cathepsin L pro-peptide (44)
Human fibroblast collagenase, matrix metalloproteinase-1 (MMP-1) (45)
N-Lobe of human serum transferin (46)

Table 1. Partial List of Heterologous Proteins Expressed in *P. pastoris* (*continued*)

C. albican hyphal surface protein HWP1 (47)
Coffee bean α-galactosidase (48)
Major histocompatibility (MHC) class II human leukocyte antigens DRα-Fos and DRβ-Jun (49)
Human immunodeficiency virus envelope glycoprotein gp120 (17)
Ghilanten (50)
Human factor XII (51)
Cellulose binding domain (CBD)–factor X variant fusion protein (52)
Cellulose binding domain (CBD)–steel factor fusion protein (1996 Invitrogen) (52)
Human gelatinase B (matrix metalloporiteinase-9, MMP-9) fragment (53)
Measles viral glycoprotein (54)
Rat liver elongation factor-2 (eEF-2) (55)
Human growth hormone (56)
Drosophila notch protein EGF-like domains (57)
Type II fish antifreeze protein (58)
Follicle stimulating hormone, FSH α and β chains (59)

Secretion using acid phosphate secretion signal sequence

Human vascular endothelial factor (VEGEF) (60)
Pectate lyase pelC from *Fusarium solani* f. sp. *pisi* (61)
Catalytic domain of 92-kDa gelatinase (matrix metalloproteinase-9, MMP-9) (62)
Angiotensin-converting enzyme (ACE) and its mutants (63)
Rabbit plasma cholesteryl ester transfer protein (CETP) (64)
β-Cryptogein elicitor protein secreted by *Phytophthora cryptogea* (65)
Candida albicans aspartic proteinases (66)
Angiostatin (an integral fragment of plasminogen) (67)
Rabbit reticulocyte 15-lipoxygenase (68)
A single chain antibody (69)
Human monocyte chemotactic protein-3 (MCP-3) (70)

Secretion using a hybrid of acid phosphatase signal sequence with a synthetic pro sequence and a KEX2 protease cleavage site

Tick anticoagulant peptide (TAP) (71)

Secretion using invertase (SUC2) secretion signal sequence

A single-chain antibody (P. Mezes, personal communication, 1992)
Porcine tumor necrosis factor-α (TNF-α) (72)
Boophilus microplus Bm 86 antigen (73)

Note: As of November 1998, the number of proteins expressed in the *Pichia* system is 172.

is substituted with the pre-pro αMF sequence (secretion $t_{1/2}$ of 10 min for pre-pro αMF signal, compared to 90 min for the native invertase signal) (16). Thus, in this case, one would avoid using αMF signal if hyperglycosylation is not desired. However, if higher productivity (product yield per liter per hour) is desired, then it is preferable to use the pre-pro αMF signal. The secretion kinetics can also impact on the recovery of the intact protein, if the expressed protein is susceptible to proteolysis in the *Pichia* broth. For example, the yield of full length MMP-1 from the culture medium is significantly improved by using the αMF signal instead of the native signal (5,45,47).

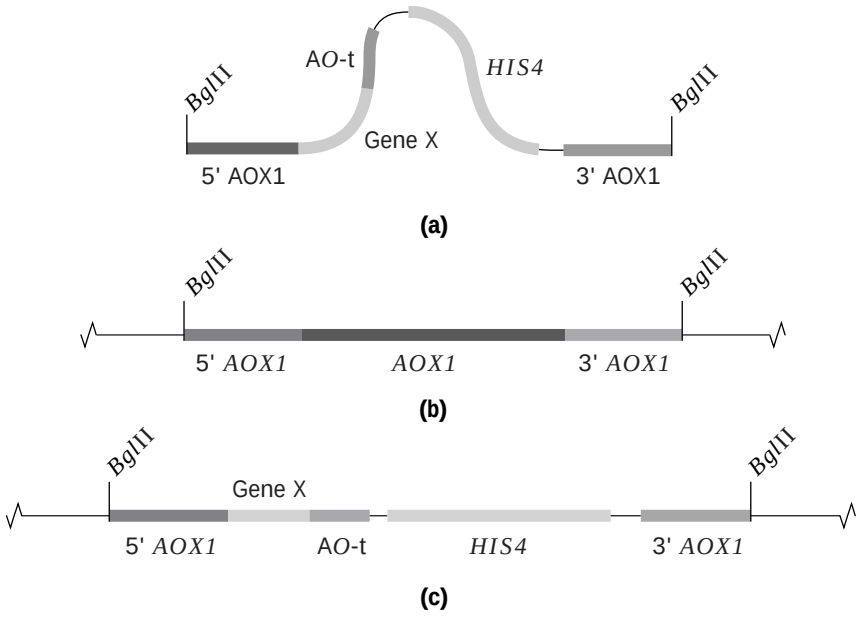


Figure 3. Site-specific eviction of *AOX1* by gene replacement. (**a**) *Bgl*II-digested DNA derived from pHIL-D1 expression plasmid used for transformation. Gene X is the gene of interest cloned at the *Eco*RI site of pHIL-D1. (**b**) *P. pastoris* chromosomal *AOX1* locus. (**c**) Chromosomal structure resulting from the replacement of the entire *AOX1* locus by the transforming DNA (shown in **a**).

Product Stabilization by Media Manipulations

In some instances, a secreted protein is rapidly degraded in the *Pichia* broth. The stability of a protein in the *Pichia* broth can be improved by manipulating the pH of the induction medium. The recommended pH range for experimentation is between 2.8 to 8.0. The pH control is best maintained in a fermenter, although it is possible to have a reasonable pH control even in shake flask cultures with the use of appropriate buffers (phosphate buffer for pH 5.7 to 8.0; alanine-HCl buffer for pH 2.5 to 3.6; citrate buffer for pH 3.0 to 6.0). The product stability is further enhanced by addition of casamino acids (1 to 4%) or yeast extract (1%) plus peptone (2%) to the medium (6,16,40,74,75). For example, secretion of HSA was significantly improved by raising the pH of the medium from 5.2 to 6.0, and the yield was further enhanced by the addition of yeast extract and peptone (4,74). Production of mouse EGF was favored at pH 3.0 in the presence of casamino acids (6). Casamino acids are preferable to yeast extract plus peptone, because the peptide components of peptone (such as bovine collagen fragments) can interfere in product analysis and recovery (5). Addition of 5 mM EDTA to the medium also improves product stability. It should be noted that media manipulation can significantly alter the profile of protein components, so that previously unnoticed proteins can become evident (5). The susceptibility of a protein to proteolysis in the *Pichia* culture medium can be determined beforehand if sufficient quantity of protein is available and proper adjustments can be made to the culture medium.

Protease-Deficient Strains

In addition to the media optimizations described in the previous section, the product yield can be further improved by using a protease-deficient *Pichia* strain. Two generally applicable techniques have been used to generate protease deficient strains. In one approach, a specific *Pichia* protease is knocked out by site-specific gene disruption, and in the other approach, a *Pichia* strain is engineered to express a specific protease inhibitor. Both types of protease-deficient strains have been found to be superior for expression of certain products. Even while using the protease deficient strains, it is essential to optimize the culture medium as discussed in the previous section.

By using the gene disruption approach, three protease-deficient strains, SMD1168 (*his4, pep4*), SMD1165 (*his4, prb1*), and SMD1163 (*his4, pep4, prb1*), that are defective in one or more of protease A (PEP4) or protease B (PRB1) have been constructed. Because a functional PEP4p is necessary for the maturation of PRB1p, the use of the single knockout strain SMD1168 alone should be adequate. The production yield of insulin-like growth factor-1 (40) and ghilanten (50) are improved twofold to threefold by using SMD1168 as compared to the strain GS115, which has intact PEP4p.

Co-expression of MMP-9 along with its inhibitor, TIMP-1, significantly improves the expression level of MMP-9 (5). Furthermore, TIMP-1 expression results in overall stabilization of several *Pichia* secreted proteins. Thus, strains engineered to express protease inhibitors such as aprotinin (16) and TIMP-1 (5) should be useful for high-level expression of proteins that are susceptible to rapid degradation.

Strategies for Coexpression of Proteins

In many instances it is required or desirable to express two or more heterologous proteins at the same time. The fact that *Pichia* can stably integrate multiple copies of an expression cassette implies that it should be possible, at least in principle, to express numerous heterologous proteins at the same time in *Pichia*. The first example of simultaneous expression of two proteins is that of the construction of *Pichia* strains expressing high levels of BSP1 (43 kDa) and BSP2 (52 kDa) components of *Bacillus sphaericus* mosquitocidal toxin (15). Since then, there have been few other

reports in literature for coexpression of two proteins in *Pichia* (5,49,51).

Two generally applicable approaches have been used for coexpression. In one approach, *Pichia* is cotransformed with two expression cassettes, the resulting His^+ transformants are screened for the presence of both the expression cassettes by PCR and subsequently examined for expression (49). In another approach, a *Pichia* expression plasmid is engineered with two expression cassettes (Fig. 4) so that the DNA used for transformation has both the expression cassettes present on the same fragment. In this situation, every His^+ transformant is expected to contain both the expression cassettes. Plasmids such as pAO815 or pAO856 (3,16) can also be used to readily construct expression plasmids for simultaneous expression of two or more products.

One can also use sequential transformation to construct *Pichia* strains that coexpress two or more different proteins, because numerous transformation markers, namely His^+, Arg^+, Suc^+, $G418^R$, and $Zeocin^R$ markers are available for *Pichia*.

Leucine Zipper Dimerization Motif Mediated Protein Assembly

Several membrane receptors, such as major histocompatibility complex (MHC) class I molecules, MHC class II molecules, T-cell receptors, are multiprotein complexes. These receptors are comprised of α- and β-chains, each of which have a large NH_2-terminal extracellular portion, a transmembrane region, and a short C-terminal cytoplasmic tail. Efforts to express soluble $\alpha\beta$-heterodimers of MHC class II molecules by coexpressing just the extracellular portions of α and β chains have proved impossible. The reason for this is that the transmembrane region of MHC class II α- and β-chains facilitate the proper assembly of $\alpha\beta$ heterodimer, and in their absence, assembly is hampered. Interestingly, this problem has been overcome (49) by substituting the transmembrane regions of α- and β-chains with the leucine zipper dimerization motifs from the transcription factors Fos and Jun, which are known to assemble as stable, soluble heterodimers. The leucine zippers are characterized by five leucines that are spaced periodically at every seventh residue (heptad repeat); each heptad repeat contributes two turns of the α-helix (3.5 residues/turn). The leucine residues play a special role in dimerization, and they form the interface between the two α-helices in the coiled coil. The Fos-Jun dimer is soluble because of the charged amino acid residues on the outer surface of the coiled coil. The Fos/Jun leucine zipper dimerization domains (Table 2) may have general applicability in the expression of extracellular domains of multiprotein membrane receptors.

Product Toxicity

In some instances, the product being expressed may be toxic to *Pichia*. In such cases, a stepwise induction is suggested. First, the cell mass is built up using glucose as sole carbon source, taking care to see that glucose is not limiting. Under these conditions, very little product is made because of strong repression of the *AOX1* promoter by glucose. Subsequent to this, the product production phase is initiated with methanol feed. Whether a particular product is toxic to *P. pastoris* can be evaluated by comparing the growth of transformants in 1% sorbitol (or 100 mM alanine) media to that on sorbitol plus methanol or alanine plus methanol media. If the expressed protein is toxic, then the growth in the presence of methanol will be drastically impaired.

Strategies for Induction of Protein Expression

Expression vectors for heterologous expression in *P. pastoris* are typically derived from *AOX1* regulatory sequences (7). One or more of the following induction strategies can be used for optimal protein expression. Refer to the "Fermentation Process" section for induction of expression in the fermenter. Refer to the "Media Composition" section for details of media composition.

Continuous Induction for Intracellular Expression. The *AOX1* promoter can be maximally activated (derepression or induction) by growing cells on methanol as the sole source of carbon and energy in shake flasks, shake tubes, or plates at 30 °C. The doubling time on methanol is $\simeq$6 h as compared to $<$2 h on glucose or glycerol medium. Even the cells in which the *AOX1* gene has been evicted (Mut^-) can grow on methanol. However, the growth rate is slow (doubling time, 18 to 24 h) because it is solely dependent on a secondary alcohol oxidase, *AOX2*, which is expressed only at low levels (79). In such cases the growth rate can be enhanced by supplementing the minimal methanol medium with 1% sorbitol or 100 mM alanine. The advantage of using sorbitol or alanine lies in the fact that these compounds do not interfere with the induction of *AOX1* promoter. Induction of expression on plates is carried out by streaking or plating cells on a methanol medium (MM) plate.

Stepwise Induction for Intracellular Expression. For liquid cultures, cells are initially grown to an OD_{600nm} of 2 to 10 on a nonactivating carbon source such as glycerol or glucose (minimal glycerol yeast extract medium [MGY] or minimal dextrose medium [MD]) in shake flasks or tubes at 30 °C. The cells are then harvested by centrifugation and shifted to MM and incubated with shaking at 30 °C for 1 to 4 days. Induction of expression on plates is carried out by streaking or plating cells on a nitrocellulose or cellulose acetate membrane placed on an MD or MGY plate. Once colonies appear, the filter is lifted and placed on an MM plate.

Efficient Secretion of Proteins. The conditions described thus far work well for the intracellular accumulation of heterologous proteins, but are rather inadequate for secreted proteins. The following protocols work well for secreted proteins for both Mut^+ and Mut^- cells (3,75).

Shake Tube Cultures. First cells are grown to saturation in 10 mL buffered minimal glycerol yeast extract medium (BMGY) placed in a 50 mL tube (2 to 3 days). The OD_{600nm} of culture will be in the range of 10 to 20. The cells are harvested by centrifugation, and the cell pellet is

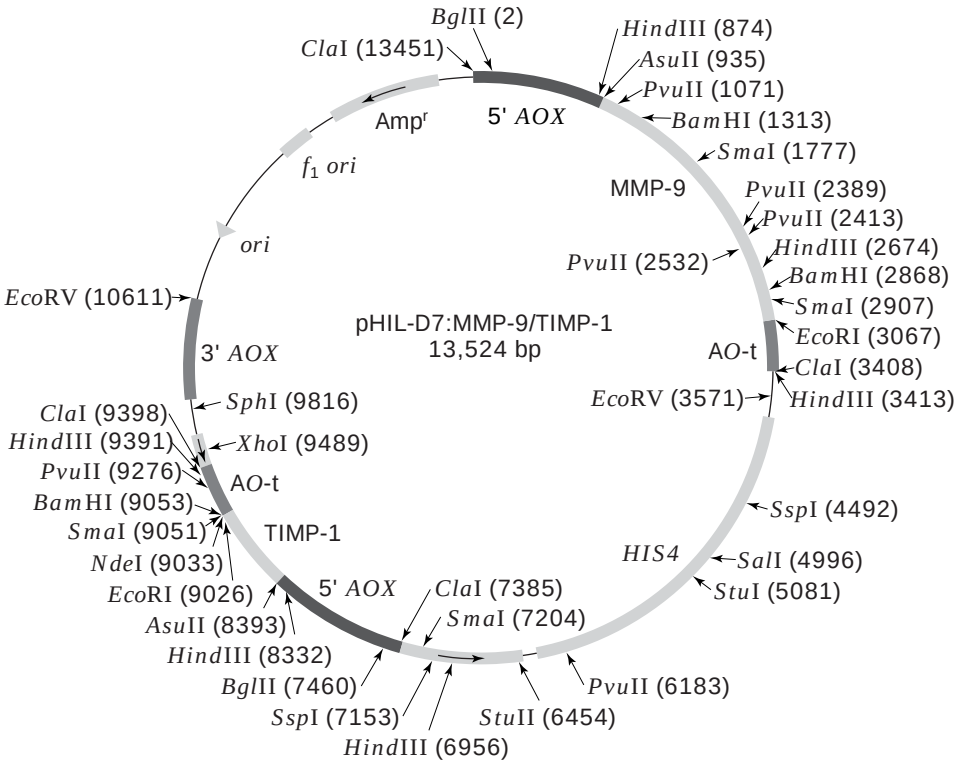


Figure 4. Expression plasmid for Coexpression of MMP-9 and TIMP-1 (5). This plasmid is derived from pHL-D7 (Fig. 2**d**) *Source*: Courtesy of Keith E. Kropp.

Table 2. Linker (7 Amino Acid) and Fos/Jun Leucine Zipper Dimerization Domains (40 Amino Acid)

VD GGGGG LTDT*L*QAETDQ*L*EDEKSA *L*QTEIAN *L*LKEKEK *L* EFILAAH
VD GGGGG RIAR*L*EEKVKT*L*KAQNSE*L*ASTANM *L*REQVAQ *L*KQKVMNH

Source: Data from Kalandadze et al. (49).

resuspended with 2 mL of buffered minimal methanol–yeast extract medium (BMMY). The tube is covered with a sterile gauze (four layers of cheesecloth; USP type VII gauze, 20 × 12 mesh from Ultimed International Inc., Glendale Heights, Ill.) and incubated in the shaker for 4 days. Aliquots of supernatant can be analyzed for product secretion at desired time intervals. Fresh media should be added to account for loss of media caused by evaporation.

Shake Flask Cultures. Cells are grown in 1 L of BMGY in a 2-L flask. Cells are harvested and resuspended in 50 mL of BMMY in a 300- to 400-mL low baffle flask, covered with four layers of cheesecloth as described in the previous section, and incubated in a 30 °C shaker for 2 to 4 days. The media supernatant is analyzed for product at different time intervals. Fresh BMMY media is added as necessary to prevent loss of liquid caused by evaporation.

Plates. Cells are patched or plated on nitrocellulose filter (sterile), placed on a square or circular BMGY plate, and incubated at 30 °C for 2 to 3 days. Once colonies have grown to ≃2 mm in size, the filter is removed and placed over a BMMY plate. The filter containing colonies is covered with a sterile nitrocellulose filter. If desired, a cellulose acetate filter can be interfaced between the two nitrocellulose filters and incubated at 30 °C for 2 to 3 days. The secreted protein on the nitrocellulose filter can be analyzed by ELISA or another suitable procedure.

FERMENTATION PROCESS

The general methods for production of biomass as well as for the production of heterologous proteins using Mut^+ and Mut^- cells in both continuous and batch modes of fermentation are described here. The process can be scaled up (>1,000 L) or scaled down (0.2 L) as desired. Fermentation can be conducted over a wide pH range (3.0 to 6.5) at 30 °C. (see "Media Composition" section for details of media.) For example, HSA secretion yield is improved more than threefold by using pH 5.85 compared to the generally used pH of 5.0 (4). In the case of secretion of the V1 domain of CD4 (amino acid residues number 1 to number 106 of mature CD4), intact product was seen only at acidic pH (2.5 to 3.5) (R.G. Buckholz, personal communication, 1990). As high as a twofold to fourfold increase in secreted yields of human EGF and human insulin-like growth factor-1 are observed at pH 3.0 (40).

Fermentation Equipment

A typical benchtop fermenter has a 2- to 20-L capacity, a 1- to 10-L operation volume, and monitors and controls for pH, dissolved oxygen (DO), agitator speed, temperature, air flow, pressure, foam, and weight.

Fermenters can be custom built or purchased from one of the numerous commercial sources such as B. Braun Biotech Inc., Allentown, Penn.; Biolaffitte, SA, France; New

Brunswick Scientific, Edison, N.J.; Porton Instruments Inc., Hayward, Calif.; Infors, UK Ltd., Crewe, Cheshire, Great Britain; Caltec Scientific Ltd., Edmonton, Alberta, Canada; Hotpack Corp., Philadelphia, Penn.; J & H Berge Inc., S. Plainfield, N.J.; Setec Inc., Livermore, Calif.; Techne, Inc., Princeton, N.J.; and VWR Scientific Products, W. Chester, Penn.

Sixfors, offered by Infors, UK Ltd., can bridge the gap between shake flasks and fermenters. They can be used for small-scale (typically 0.3 L) continuous, batch, or cascade processes where the growth from one vessel is passed into another.

Methods for Monitoring the Fermentation Process

Bio-Rad's portable fermentation monitoring analyzer can be used on the spot to provide fast HPLC analysis of the concentration of methanol, glycerol, ethanol, glucose, acetic acid, lactic acid, fructose, maltotriose, and maltose. The equipment is compatible with automatic sampling and computerized data analysis. Analysis takes only 10 min. The conditions used are as follows:

1. *Instrument*: Bio-Rad's fermentation monitoring analyzer
2. *Column*: Fermentation monitoring column (150 × 7.8 mm)
3. *Sample*: Extracellular broth, prefiltered through a 0.22-μm syringe filter; sample volume: 10 to 20 μL
4. *Injector*: Rheodyne injector valve with a 10-μL injection loop
5. *Solvent*: 0.002 N H_2SO_4 in water at a flow rate of 0.8 mL/min (isocratic, so only one pump is required)
6. *Temperature*: 65 °C
7. *Detector*: Refractive index detector

To determine the wet cell weight of the culture, 1 mL of the fermenter broth is centrifuged in a microfuge, and the cell pellet is weighed after carefully decanting the cell-free supernantant. To determine the washed dry cell weight of the culture, cells are harvested by centrifugation (3000 × *g* for 10 min at room temperature) from a known volume of the fermenter culture (typically 10 to 50 mL). The cell pellet is washed twice with water (10 to 50 mL), dried overnight at 100 °C, and weighed. Alternatively, the dry cell weight can be readily determined in less than 10 min with the use of a moisture balance such as the Mettler LP16 Infrared Dryer (Mettler Instrument Corporation, Hightstown, N.J.).

Mass transfer is determined by measuring the air flow to the fermenter and the composition of both the inlet and outlet air with a Perkin-Elmer gas analyzer. The O_2 transfer rate (OTR) is calculated as follows:

$$\text{OTR} = f(C_{\text{in}} - C_{\text{out}})/V$$

where f is flow rate L/h, C_{in} and C_{out} are the concentration in mmol of O_2 in the inlet and outlet gases, and V is the ungassed broth volume. OTR can also be established from yields using the equation

$$\text{OTR} = \mu X/Y \times O_2$$

where μ is the specific growth rate or dilution rate (h^{-1}), X is the cell density (g/L), and $Y \times O_2$ is the yield on O_2 (g cells/mmol O_2).

Heat transfer rate and heat load can be determined by measuring the temperature of the circulating water in and out of the fermenter heat exchanger.

Productivity (P) for the continuous culture is determined using $P = DX$, where D is the dilution rate (h^{-1}), and X is the cell density. An increase in D and/or X will result in higher productivity. Because D also equals the specific growth rate μ, the dilution rate is limited by the microorganism's intrinsic characteristics. Once $\mu_{\max}$ or $D_{\max}$ is reached, P can only be increased by increasing the cell density. A set of parameters determined for the high-productivity process for *P. pastoris* on methanol in a 1,500-L fermenter are given in Table 3.

Extracellular protein concentration (for secreted proteins) can be estimated by using a Lowry-type protein analysis on TCA-precipitated material and analyzed for specific product by SDS-PAGE and immunological methods. Aliquots of cells are lysed to prepare cell extracts for protein analysis.

Continuous Culture of Mut$^+$ and Mut$^-$ *Pichia* on Methanol

Fermentation is carried out in two steps, first in the batch mode on glycerol as the carbon source, followed by continuous mode on methanol-containing medium.

Inoculum for the fermenter is prepared by growing cells in 1 L of MMD, YMPD, YMPGy, MGyB, MMGY, or YPD grown to an OD_{600nm} of 2 to 10 in a 2-L shake flask. This volume of inoculum is used for inoculating a 20-L fermenter containing 9 L of basal salt FM21 medium containing 5% glycerol (higher levels may be toxic to the cells) or 5 to 10% glucose and pH adjusted to 5.0 (by using 50% NH_4OH solution or NH_3 gas). Biotin stock (4 mL) and PTM1 trace mineral mix (11 mL) are also added. Fermentation is conducted until the glycerol or glucose (carbon source) is exhausted. During the run, dissolved O_2 is maintained at >20%, with the agitator speed set between 500 to 1,500 rpm and a vessel pressure of 2 to 3 psi. Foaming is con-

Table 3. Parameters Determined for a High Cell-Density Process in a 1,500-L Fermenter

Substrate	Methanol
Substrate concentration (g/L)	263
Dilution rate (h^{-1})	0.11
pH	3.5
Temperature (°C)	30
Cell mass (g of cells/L)	105
Productivity (g/L h)	11.6
Yield (g of cells/g of methanol)	0.4
OTR (mmol O_2/L h)	880
Oxygen consumption (g O_2/g cells)	2.42
Heat release	
kcal/L h	109
kcal/mol O_2	123.9

Source: Adapted from Shay et al. (80).

trolled through the addition of a 5% Struktol J673 (Struktol Company of America, Stow, Ohio) or Mazu DF 37C (Mazer Chemicals, Inc., Gurnee, Ill.). The cell yield expected for the batch phase on 5% glycerol under these conditions is 20 to 25 g of washed dry cell weight per liter.

Continuous fermentation is established by feeding FM21-methanol (15% v/v) for Mut^+ cells or FM21-methanol (1% v/v) plus 15% glycerol, sorbitol, or alanine for Mut^- cells. The feed for continuous culture is also supplemented with PTM1 (1.1 mL of stock solution per liter) and biotin (0.4 mL of the stock solution per liter). Feed sterilization is carried out by filtration (Pall Ultipor disposable filter assembly DFA 4001 AR, 0.2 μm). Feed addition is achieved with a Milton-Roy positive-displacement metering pump (Model 2396 Duplex). Continuous culture is performed as a chemostat under steady-state conditions where the dilution rate (D) is equal to the growth rate of the population. The growth rate of cells can be controlled by adjusting the flow rate of fresh medium into the fermenter. Typical D values range from 0.056 h to 0.11 h. Dissolved oxygen is held in the 45 to 75% air saturation range by varying air flow, vessel pressure, or agitator speed. The maximum cell mass achievable under these conditions is around 80 g of washed dry cell weight per liter with a cell productivity of approximately 10 g/L h. If a higher cell mass is desired, the amount of carbon source in the feed can be raised (e.g., 25% methanol for Mut^+ and 1% methanol plus 25% glycerol, sorbitol, or alanine for Mut). This would also necessitate a proportional increase in the amount of minerals, trace elements, and biotin. As high as 110 g/L washed dry cell weight with 11.6 g/L h. productivity are achievable in this process. Highest induction of the *AOX1* promoter occurs using methanol (Mut^+ cells) or methanol plus sorbitol or alanine (Mut^-) as the carbon source in the feed. Intermediate levels of induction are seen with methanol plus glycerol feed.

Fed-Batch Fermentation of Mut^+ and Mut^- *Pichia* on Methanol

Fermentation is carried out in three steps: glycerol batch phase, glycerol fed-batch phase, and methanol-fed batch phase.

One liter of inoculum (prepared as previously described for the continuous fermentation protocol) is used for inoculating a 20-L fermenter containing 5 L of basal salt BSM medium with 5% v/v glycerol (higher levels may be toxic to the cells) or 5 to 10% glucose as carbon source; the pH of which is adjusted to 5.0 by using 50% NH_4OH solution or NH_3 gas. Biotin stock (40 mL) and PTM1 trace mineral mix (40 mL) are also added. Fermentation is conducted as previously described until all the carbon source (glucose or glycerol) is completely consumed. This phase should take 18 to 24 h. After this, a fed-batch phase on glycerol is initiated with a 50% w/v glycerol feed (500 mL of 100% glycerol + 480 mL water + 10 mL each of PTM1 and biotin stock solution) at a feed rate of 18 mL/h/L initial fermentation volume with the aid of a peristalic pump. Glycerol feeding is carried out for 4 h. The cell yield at this point will be in the range of 180 to 220 g/L of wet cells (equivalent to approximately 30 to 55 g washed dry cell weight per liter). This phase can be manipulated by varying the concentration of glycerol or the duration of fermentation to achieve optimal heterologous protein production in the methanol-fed batch phase. Some suggested cell yield ranges that should be tested for are listed in Table 4.

The fed-batch phase on methanol is initiated with a methanol feed (980 mL of 100% methanol + 10 mL of PTM1 + 10 mL of biotin stock solution) at a rate of 1 to 4 mL/L/h of the initial culture volume. The methanol-feed flow rate is adjusted such that with Mut^- strains, the methanol concentration in the fermenter is maintained in the 0.2 to 0.5% v/v level, whereas with the Mut^+ strains, methanol levels in the fermenter approach zero and the fermenter is run in a methanol-limited fashion. Dissolved oxygen in both cases is maintained at the 20 to 70% range. If dissolved oxygen falls below 20%, the methanol feed is stopped, and nothing should be done to increase oxygen rates until an upward spike in the dissolved oxygen level is seen. At this point, adjustments (rpm, aeration, vessel pressure, oxygen feed) can be made. Maintaining the dissolved oxygen above 20% may be difficult depending on the OTR of the fermenter. With stainless steel vessels, the system can be pressurized up to 15 to 30 psi to increase OTR. Also, the oxygen feed (air plus oxygen mixture) at 0.1 to 0.3 vvm can be used to maintain adequate levels of dissolved oxygen. The methanol fed-batch phase generally lasts for 70 h. However, it may prolong to more than 200 h if the feed rate is slowed down because of the fall in dissolved oxygen levels. Another factor that may also influence the fermentation time is the secretion rate of the heterologous protein. Longer times may be necessary to allow for accumulation of high levels of a slowly secreted protein in the broth. Depending on the product being produced, adjustments in the media (addition of casamino acids) and pH will have to be made to increase product yield. One technique that can be used to lower the pH is by setting the pH to the desired lower value (e.g., pH 3.0) and then letting the culture pH to decrease to the new set point of pH 3.0 as a result of cellular metabolism (40).

Strategies for Multicycle Fermentation

Considerable time is expounded on setting up a fermentation run. It would make sense to reap as many cycles of product production as possible for each start-up. Continuous fermentation on methanol, which works well for biomass production (1,2,80), is largely not applicable for recombinant protein production, especially with Mut^- strains, because they grow poorly on methanol. Use of methanol plus glycerol mixed feed restores the cell productivity, with considerable level of protein expression; optimal level is not achievable because of partial repression of the *AOX1* promoter by glycerol (11). This problem can be overcome by using a sorbitol plus methanol mixed feed, because sorbitol is a nonrepressing carbon source. However, sorbitol is a poorer carbon source than glycerol and also needs supplementation with 1% each of malt extract, peptone, and yeast extract for improved growth. Nevertheless, this strategy has been used successfully to accomplish several cycles of MMP-2 production in a Biostat-B (Braun) benchtop fermenter (5). In this approach, after recovery of 90% of the fermenter sample from the first fermentation

Table 4. Cell Yield Ranges

Mut^- intracellular expression	50–100 g/L dry cell weight or 200–400 g/L wet cell weight
Mut^+ intracellular expression	30–80 g/L dry cell weight or 140–320 g/L wet cell weight
Mut^- secretion	20–40 g/L dry cell weigth or 80–160 g/L wet cell weight
Mut^+ secretion	12–80 g/L dry cell weight or 50–300 g/L wet cell weight

Note: It is obvious from the suggested ranges for cell density that in several instances the glycerol fed-batch phase is not even necessary.

cycle, the fermenter is rebooted with an appropriate volume of fermenter growth media containing sorbitol as carbon source and supplemented with 0.1% yeast extract and 0.2% peptone. After growth for 16 to 20 h, a fed-batch fermentation with sorbitol plus methanol nutrient feed is continued for another 72 h for MMP-2 production. The whole process of product recovery and rebooting the fermenter can be repeated as many times as one desires. By this approach, using a 4-L fermenter with one start-up, we were able to generate 25 L of MMP-2 broth in less than 4 weeks. This approach is also useful for Mut^+ cells; it reduces overall methanol consumption because most of the growth is supported by sorbitol. Methanol can be added to the sorbitol feed at any desired point, to initiate induction of expression.

CONCLUSIONS AND FUTURE PERSPECTIVES

Problems generally encountered in protein expression may be overcome in most instances by taking due consideration of the factors that influence protein expression. Nearly every expression project has given new insight into the intricacies of the *Pichia* system. In spite of due consideration and careful planning, certain proteins are difficult to express in this system. Now that the expression system is readily available from Invitrogen (San Diego, Calif.), hundreds of investigators around the world are currently exploring the system, and their results will undoubtedly provide new insights as well as expand our knowledge base about the utility and limitations of the system.

Within the near future, it is anticipated that *Pichia*-produced HSA will reach the market. It will make one think of what to do with the tens of thousands of tons of *Pichia* yeast that would become available as a by-product of HSA production plants. It is tempting to speculate that the *Pichia* HSA strain would also be engineered to produce an intracellular insecticidal toxin, and thus the *Pichia* cells resulting from the HSA production can be used as a biopesticide.

Several modifications have recently been made to the *P. pastoris* expression vectors by Invitrogen (San Diego, Calif.). These newer vectors (pPICZ A,B,C series, and pPICZα A,B,C series) are smaller in size and have $Zeocin^R$ marker. Furthermore, they allow in-frame fusion to *myc* epitope and His_6 tags. It has been previously noticed that the addition of eight amino acid carboxy terminal tags to TNF rendered the tagged proteins completely insoluble under nondenaturing conditions, whereas the untagged TNF was completely soluble (11). Thus, a tag enthusiast should bear in mind that in some instances, a tag can bring about an undesirable change in the property of a protein.

Overexpression of ubiquitin (Ubi) seems to enhance the secretion of a human leukocyte protease inhibitor in *S. cerevisiae* (81). Ubi is recognized to be a normal secreted component of *P. pastoris* (5). Thus, it is presumable that the overexpression of Ubi in *P. pastoris* may also enhance heterologous protein secretion. A *P. pastoris* host strain that overexpresses Ubi can be engineered using one of the available Ubi cDNAs. Also, a Ubi expression plasmid can be introduced into any *P. pastoris* strain by using one of the dominant selection markers for transformation, such as invertase, *SUC2* (82), $G418^R$ (77,83) or $Zeocin^R$ (available from Invitrogen, San Diego, Calif.). Thus, the effect of Ubi overexpression on product secretion can be tested in any of the existing production strains. A note of concern may be warranted while using $zeocin^R$ for selecting *P. pastoris* transformants. $Zeocin^R$ is a strong mutagen, and thus cells challenged with Zeocin*R* on a complex medium may acquire some undesirable change that may interfere with growth and thus reduce the overall productivity under the growth conditions in a fermenter.

A major concern for large-scale use of *Pichia* in the production of biologicals has been the substrate methanol, which is toxic, inflammable, and volatile. With the existing technology in a typical production plant, one will have to use several tons of methanol for the production phase of the operation. It will certainly be an added advantage if the amount of methanol required can be reduced or even eliminated. Two recent developments in this direction have been made. In one case, a mutant strain of *Pichia* that is able to turn on the *AOX1* promoter in the absence of methanol has been generated (A.A. Sibirny, personal communication, 1992). This mutant still responds to repression by glucose and ethanol. In another instance, a *Pichia* strain in which both the alcohol oxidase structural genes *AOX1* and *AOX2* have been deleted has been constructed and shown to express heterologous protein LACZ in response to methanol induction (V. Chiruvolu, personal communication, 1995). This novel alcohol oxidase null strain (*aox1, aox2*) uses considerably lower amounts of methanol for production of hetrologous protein.

Another development that one can speculate is that *Pichia* will be engineered to express human post-translational modification enzymes (e.g., glycosyl transferases, amidating enzyme, vitamin K-γ carboxylase, etc.), so as to obtain *Pichia* strains that can be used to produce more authentic human proteins.

MEDIA COMPOSITIONS

Stock Solutions

Note: For filter sterilization of various solutions and liquids, filter wares (disposable or reusable types) equipped

with cellulose acetate or cellulose nitrate membranes (pore size 0.2 to 0.22 μm) from one of the several manufacturers (Nalgene Company, Rochester, N.Y.; Costar Corporation, Cambridge, Mass.; Corning Glass Works, Corning, N.Y.) can be used. For filter sterilization of methanol and methanol-containing media, only cellulose acetate membranes (0.2 to 0.22 μm) are suitable, because methanol does not filter through cellulose nitrate membranes of pore size 0.2 μm.

10× YNB. Dissolve 13.4 g of yeast nitrogen base without amino acids (YNB, Difco Labs., Detroit, Mich.) in 100 mL of water (heat if necessary) and filter sterilize. This solution can be stored for over a year at 4 °C.

500× B. Dissolve 20 mg of D-biotin (Sigma Chemicals, St. Louis, Mo.) in 100 mL of water and filter sterilize.

100× H. Dissolve 400 mg L-histidine in 100 mL of water (heat if necessary) and filter sterilize.

10× D. Dissolve 20 g of D-glucose in 100 mL water. Autoclave for 15 min or filter sterilize. Stores well for years at room temperature.

10× GY. Mix 100 mL of glycerol with 90 mL of water. Filter sterilize. Stores well for years at room temperature.

10× M. Mix 5 mL of methanol (100%) with 95 mL of water. Filter sterilize and store at 4 °C.

100% methanol. Filter sterilize pure methanol (100%). Store at room temperature in a fireproof cabinet.

Minimal Media Compositions

MD. Mix 100 mL of 10× YNB, 2 mL of 500× B, and 100 mL of 10× D with 800 mL of autoclaved water (include 15 g bacto agar for plates).

MM. Mix 100 mL of 10× YNB, 2 mL of 500× B, and 100 mL of 10× M with 800 mL of autoclaved water (include 15 g bacto agar for plates).

MGY. Mix 100 mL of 10× YNB, 2 mL of 500× B, and 100 mL of 10× GY with 800 mL autoclaved water (include 15 g bacto agar for plates).

All these liquid media and plates store well for several weeks at 4 °C.

Minimal media with other carbon sources (such as D-sorbitol, D,L-alanine) are prepared by using the desired carbon source at 10 g/L in place of glucose in MD. Minimal media containing a mixture of carbon sources can also be prepared by combining two or more desired substrates in the growth medium.

Supplemental Minimal Media Compositions

Minimal media are supplemented with necessary supplemental nutrients such as amino acids depending on the specific requirement of a given strain. For example, *P. pastoris* strains GS115 and KM71, commonly used in molecular genetic manipulations, are auxotrophic for histidine. Such strains will grow in minimal media only in the presence of supplemental histidine. However, once transformed with *HIS4* (histidinol dehydrogenase gene) they readily grow in the absence of histidine.

The composition of supplemental minimal histidine media (suitable for histidine auxotrophic strains such as GS115) is as follows. Other supplemental media can be prepared depending on the need of a particular strain in use.

MDH. Mix 100 mL of 10× YNB, 2 mL of 500× B, 100 mL of 10× D, and 10 mL of 100× H with 790 mL of autoclaved water (include 15 g agar for plates).

MMH. Mix 100 mL of 10× YNB, 2 mL of 500× B, 100 mL of 10× M, and 10 mL of 100× H with 790 mL of autoclaved water (include 15 g agar for plates).

MGyH. Mix 100 mL of 10× YNB, 2 mL of 500× B, 100 mL of 10× GY, 10 mL of 100× H with 790 mL of autoclaved water (include 15 g agar for plates).

MGyB. Dissolve 11.5 g KH_2PO_4, 2.66 g K_2HPO_4, 6.7 g YNB, pH 6.0, and 20 mL glycerol in 1 L water and autoclave.

All these liquid media and plates will store well for several weeks at 4 °C.

Supplemental minimal histidine media with other carbon sources is prepared by adding a similar amount of histidine as above to the minimal media with the desired carbon source.

Complex Media Composition

YPD. Dissolve 10 g of bacto yeast extract, 20 g of peptone, and 20 g of glucose in 1,000 mL of water (also include 15 g bacto agar for slants and plates) and autoclave for 20 min.

YMPD. Dissolve 3 g of yeast extract, 3 g of malt extract, 5 g of peptone, and 10 g of glucose in 1 L of water and autoclave.

YMPGy. Same as YMPD with the exception that 10 mL of 100% glycerol is used instead of 10 g of glucose.

Secretion Media Composition

BMGY. Mix 100 mL of 1 M potassium phosphate buffer, pH 6.0, 100 mL of 10x YNB, 2 mL of 500x biotin (refer to growth and storage section for composition of stock solutions), and 10 mL of glycerol. Filter sterilize and add to an autoclaved solution of 10 g yeast extract and 20 g peptone in 788 mL water (15 g bacto agar is included for plates).

BMMY. Same as BMGY, with the exception that 5 mL of methanol is added in the place of 10 mL glycerol.

Note: Yeast extract and peptone in the above media can be replaced by 1% casamino acids. The pH 6 suggested here may not be optimal for every secreted product. Experimentation with pH values in the range 2.5 to 8 (by using appropriate buffers) is suggested to determine the optimal pH for a particular product. Some suggested buffers are as follows:

Phosphate buffer for pH range 5.7 to 8.0
Alanine-HCl buffer for pH range 2.5 to 3.6
Aconitic acid-NaOH buffer for pH range 2.5 to 5.7
Citrate buffer for pH range 3.0 to 6.2

Avoid buffers such as succinate buffer, because succinate can serve as a carbon source and will repress activation of the *AOX1* promoter.

Fermentation Media Composition

FM21 basal salt media

Composition is for 1 L final volume in water:

Phosphoric acid, H_3PO_4 (85%)	3.5 mL
Calcium sulfate, $CaSO_4 \cdot 2H_2O$	0.15 g
Potassium sulfate, K_2SO_4	2.4 g
Magnesium sulfate, $MgSO_4 \cdot 7H_2O$	1.95 g
Potassium hydroxide, KOH	0.65 g

Biotin stock solution

Biotin	0.2 g/L

PTM1 trace salts

Composition is for 1 L final volume in water:

Cupric sulfate ($CuSO_4 \cdot 5H_2O$)	6.0 g
Manganese sulfate ($MnSO_4 \cdot H_2O$)	3.0 g
Ferrous sulfate ($FeSO_4 \cdot 7H_2O$)	65.0 g
Zinc sulfate ($ZnSO_4 \cdot 7H_2O$)	20.0 g
Silfuric acid (H_2SO_4)	5.0 mL
Cobalt chloride ($CoCl_2 \cdot 6H_2O$)	0.5 g
Boric acid (H_3BO_3)	0.02 g
Sodium molybdate ($NaMoO_4 \cdot 2H_2O$)	0.2 g
Potassium iodide (KI)	0.1 g

BSM medium composition

Composition is for 1 L final volume in water:

Phosphoric acid, H_3PO_4 (85%)	26.0 mL
Calcium sulfate, $CaSO_4 \cdot 2H_2O$	0.9 g
Potassium sulfate, K_2SO_4	18.0 g
Magnesium sulfate, $MgSO_4 \cdot 7H_2O$	14.0 g
Potassium hydroxide, KOH	4.0 g

GLOSSARY OF *P. pastoris* VECTORS

pHIL-A1	Autonomously replicating vector (3)
pHIL-D1	Integration vector with or without deletion of AOX1 structural gene (3)
pPIC3	pHILD1 type vector with multiple cloning sites (3)
pAO815	pHILD1 type vectors for making multicopy expression units (3)
pAO856	pAO815 with a unique *Bg1*II site (3)
pHIL-D2	Modified pHILD1 with *Not*I site and fl ori (3)
PHIL-D3	Derived from pHIL-D2 for making constructs with exact 5′-UTR (3)
pHIL-D4	pHILD1 with kanamycin-resistance marker (3)
pPIC3K	pPIC3 with kanamycin-resistance gene (3)
pHIL-D5	pHILD2 with kanamycin-resistance gene (3)
pHIL-D7	pHILD4 with a unique *Csp45*I (*Asu*II) site for making exact 5′-UTR constructs (75)
pHIL-S1	Secretion vector with *P. pastoris* acid phosphatase secretion signal (3)
pPIC9	Secretion vector with *S. cerevisiae* αMF pre-pro signal (6)
pPIC9K	pPIC9 with kanamycin-resistance gene (77)
pPICZ A, B and C	Zeocin^R marker and *myc* epitope–His_6 tag (Invitrogen Inc.)
pPICZα A, B, and C	Zeocin^R marker, *myc* epitope–His_6 tag, and αMF secretion sequence (Invitrogen Inc.)

GLOSSARY OF *P. pastoris* STRAINS

NRRL Y-11430-SC5 (wild type) (3)

GS115 (*his4*)—this strain is also known as GTS115 (3)

KM71 (*his4*, *aox1*::*ARG4*) (3)

PPF1 (*his4*, *arg4*) (3)

Mc100-3 (*aox1*::*ARG4*, *aox2*::*his4*, *his4*, *arg4*) (79)

Protease deficient strains (derived by protease A (*PEP4*) and/or protease B (*PRB*) gene disruption (3):

- SMD1163 (*his4*, *pep4*, *prB1*)
- SMD1165 (*his4*, *prB1*)
- SMD1168 (*his4*, *pep4*)

BIBLIOGRAPHY

1. G.H. Wegner and W. Harder, in H.W. Van Verseveld and J.A. Duine eds., *Microbial Growth on C_1 Compounds*, Martinus Nijhoff, Boston, 1986, pp. 139–149.
2. G.H. Wegner, *FEMS Microbiol. Rev.* **87**, 279–284 (1990).
3. K. Sreekrishna and K.E. Kropp, *Non Conventional Yeasts in Biotechnology*, Springer, Berlin, 1996, pp. 203–252.
4. K. Sreekrishna, K.A. Barr, S.A. Hoard, W.D. Prevatt, R.E. Torregrosa, R.E. Levingston, J.A. Cruze and G.H. Wegner, *Fifteenth International Congress of Yeast Genetics and Molecular Biology*, Hague, The Netherlands, topic number 09-37B.
5. K. Sreekrishna, R.G. Brankamp, K.E. Kropp, D.T. Blankenship, J.T. Tsay, P.L. Smith, J.D. Wierschke, A. Subramaniam, and L. Birkenberger, *Gene* **190**, 55–62 (1997).
6. J.J. Clare, M.A. Romanos, F.B. Rayment, J.E. Rowedder, M.A. Smith, M.M. Payne, K. Sreekrishna, and C.A. Henwood, *Gene*, **105**, 205–212 (1991).
7. S.B. Ellis, P.F. Brust, P.J. Koutz, A.F. Waters, M.M. Harpold, and T.R. Gingeras, *Mol. Cell. Biol.* **5**, 1111–1121 (1985).
8. J.M. Cregg and K.R. Madden, *Dev. Ind. Microbiol.* **29**, 33–41 (1988).
9. J.M. Cregg, J.F. Tschopp, C. Stilman, R. Siegel, M. Akong, W.S. Craig, R.G. Buckholz, K.R. Madden, P.A. Kellaris, G.R. Davis, B.L. Smiley, J. Cruze, R. Toregrossa, G. Velicelebi, and G.P. Thill, *Bio / Technology* **5**, 479–485 (1987).
10. K. Sreekrishna, L. Nelles, R. Potenz, J. Cruze, P. Mazzaferro, W. Fish, M. Fuke, K. Holden, D. Phelps, P. Wood, and K. Parker, *Biochemistry* **28**, 4117–4125 (1989).
11. K. Sreekrishna, R. Potenz, J.A. Cruze, W.R. McCombie, K.A. Parker, L. Nelles, P.K. Mazzaferro, K.A. Holden, R.G. Harrison, P.J. Wood, D.A. Phelps, C.E. Hubbard, and M. Fuke, *J. Basic Microbiol.* **28**, 265–278 (1988).
12. M.J. Hagenson, K.A. Holden, K.A. Parker, P.J. Wood, J.A. Cruze, M. Fuke, T.R. Hopkins, and D.W. Stroman, *Enzyme Microb. Technol.* **11**, 650–656 (1989).
13. J.J. Clare, F.B. Rayment, S.P. Ballantine, K. Sreekrishna, and M.A. Romanos, *Bio / Technology* **9**, 455–460 (1991).
14. M.A. Romanos, J.J. Clare, K.M. Beesley, F.B. Rayment, S.P. Ballantine, A.J. Makoff, G. Dougan, N.F. Fairweather, and I.G. Charles, *Vaccine* **9**, 901–906 (1991).

15. Eur. Pat. 0586892A1 (March 16, 1994), K. Sreekrishna, W.D. Prevatt, G.T. Thill, G.R. Davis, P. Koutz, K.A. Barr, and S.A. Hopkins (to Research Corporation Technologies, Inc.).
16. G.P. Thill, G.R. Davis, C. Stillman, C. Holtz, R. Brierly, M. Engel, R. Buckholz, J. Kinney, S. Provo, T. Vedvick, and R.S. Siegel, *Proceedings of the Sixth International Symposium on Genetics of Industrial Microorganisms*, Societe Francaise de Microbiologie, Paris, 1990, pp. 477–490.
17. C.A. Scorer, R.G. Buckholz, J.J. Clare, and M.A. Romanos, *Gene* **136**, 111–119 (1993).
18. R.G. Reddy, T. Yoshimoto, S. Yamamoto, and L.J. Marnett, *Biochem. Biophys. Res. Commun.* **205**, 381–388 (1994).
19. D. Antelman, D. Giroux, J. Christian, P. Shabram, and R. Gregory, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1994, p. 25.
20. S. Weiss, M. Famulok, F. Edenhofer, Y.H. Wang, I.M. Jones, M. Groschup, and E.L. Winnacker, *J. Virol.* **69**, 4776–4783 (1995).
21. H.K. Brandes, F.C. Hartman, T.Y. Lu, and F.W. Larimer, *J. Biol. Chem.* **271**, 6490–6496 (1996).
22. N. Shiraishi, *Plant Physiol.* **108**, 72–81 (1995).
23. M.S. Payne, K.L. Petrillo, J.E. Gavagan, L.W. Wagner, R. Dicosimo, and D.L. Anton, *Gene* **167**, 215–219 (1995).
24. J.M. Trant, *Arch. Biochem. Biophys.* **326**, 8–14 (1996).
25. S.H. Wu and J.C. Lagarias, *Proc. Natl. Acad. Sci. USA* **93**, 8989–8994 (1996).
26. E.Z. Monosov, T.J. Wenzel, G.H. Luers, J.A. Heyman, and S. Subramani, *J. Histochem. Cytochem.* **44**, 581–589 (1996).
27. R. Zhou, B. Kroczynska, and J.A. Miernyk, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 26.
28. N. Shiraishi and W.H. Campbell, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 38.
29. J. Hoffman, and M. Huff, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 46.
30. Y-C. Peng and N.H. Acheson, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 54.
31. Y. de Vries, D.N. Arvidson, H.R. Waterham, J.M. Cregg, and G. Woldegiorgis, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 59.
32. H.M. Weiss, W. Haase, H. Michel, and H. Reilander, *FEBS Lett.* **377**, 451–456 (1995).
33. J.F. Tschopp, G. Svelow, R. Kosson, W. Craig, and L. Grinna, *Bio/Technology* **5**, 1305–1308 (1987).
34. M.E. Digan, S.V. Lair, R.A. Brierly, R.S. Siegel, M.E. Williams, S.B. Ellis, P.A. Kellaris, S.A. Provow, W.S. Craig, G.V. Velicelebi, M.M. Harpold, and G.P. Thill, *Bio/Technology* **7**, 160–164 (1989).
35. N. Juge, J.S. Anderson, D. Tull, P. Roepstroff, and B. Svensson, *Protein Exp. Purif.* **8**, 204–214 (1996).
36. M. Yamada, T. Azuma, T. Matsuba, H. Iida, H. Suzuki, K. Yamamoto, Y. Kohli, and H. Hori, *Biochim. Biophys. Acta* **1206**, 279–285 (1994).
37. C.W. Despreaux and R.F. Manning, *Gene*, **131**, 35–41 (1993).
38. W. Guo, L.G. Candelas, and P.E. Kolattukudy, *J. Bacteriol.* **177**, 7070–7077 (1995).
39. M. Basu, M.Y. Nettleton, E. Dharm, R.F. Manning, C.W. Despreaux, J. Hakimi, and J.P. Kochan, *American Society for Biochemistry and Molecular Biology/Biophysical Society Joint meeting*, Houston, Texas, Feb. 9–13, 1992, Abstract #1507.
40. U.S. Pat. 5,324,639 (1994), R.A. Brierley, G.R. Davis, and G.C. Holtz.
41. S.L. Wagner, R.S. Siegel, T.S. Vedvick, W.C. Raschke, and W.E. van Nostrand, *Biochem. Biophys. Res. Commun.* **186**, 1138–1145 (1992).
42. D. Mohanraj, T. Olson, and S. Ramakrishnan, *Biophys. Res. Commun.* **215**, 750–756 (1995).
43. J.J. Fiordalisi, P.L. James, Y. Zhang, and G.A. Grant, *Toxicon* **34**, 213–224 (1996).
44. E. Carmona, E. Dufour, C. Plouffe, S. Takebe, P. Mason, J.S. Mort, and R. Menard, *Biochemistry* **35**, 8147–8157 (1996).
45. S.A. Rosenfeld, O.H. Ross, M.C. Hillman, J.I. Corman, and R.L. Dowling, *Protein Exp. Purif.* **7**, 423–430 (1996).
46. A.B. Mason, R.C. Woodworth, R.W.A. Oliver, B.N. Green, L-N. Lin, J.F. Brandts, B.M. Tam, A. Maxwell, and R.T.A. MacGillivary, *Protein Exp. Purif.* **8**, 119–125 (1996).
47. J.F. Staab, C.A. Ferrer, and P. Sundstrom, *J. Biol. Chem.* **271**, 6298–6305 (1996).
48. A. Zhu, C. Monahan, Z. Zhang, R. Hurst, L. Leng, and J. Goldstein, *Arch. Biochem. Biophys.* **324**, 65–70 (1995).
49. A. Kalandadze, M. Galleno, L. Foncerrada, J.L. Strominger, and K.W. Wucherpfennig, *J. Biol. Chem.* **271**, 20156–20162 (1996).
50. R.G. Brankamp, K. Sreekrishna, P.L. Smith, D.T. Blankenship, and A.D. Cardin, *Protein Exp. Purif.* **6**, 813–820 (1995).
51. C.E. White, N.M. Kempi, and E.A. Komives, *Structure* **2**, 1003–1005 (1994).
52. M.M. Guarna, A. Boraston, E.A. Amandoron, H.C.F. Cote, R.T.A. MacGillivary, R.A.J. Warren, and D.G. Kilburn, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 28.
53. N. Roy, M.M. Smith, M. Navre, and G. Das, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 56.
54. I. Duffy and B.K. Rima, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 41.
55. P. Bargis, A. Dumont-Miscopein, J.P. Lavergne, and J.P. Reboud, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 42.
56. L.L. Escamilla-Trevino, M. Guerrero-Olazaran, J.M. Viader-Salvado, and H.A. Barrera-Saldana, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 43.
57. H.A. Gonzalez and E.A. Komives, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 44.
58. M.C. Loewen, A.J. Daugulis, and P.L. Davies, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 47.
59. C.A. Seid, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 58.
60. D. Mohanraj, T. Olson, and S. Ramakrishnan, *Growth Factors* **12**, 17–27 (1995).
61. W. Guo, L. Gonzalez-Candelas, and P.E. Kolattukudy, *Arch. Biochem. Biophys.* **323**, 352–360 (1995).
62. J.M. Shipley, G.A.R. Doyle, C.J. Fliszar, Q-Z. Ye. L.L. Johnson, S.D. Shapiro, H.G. Welgus, and R.M. Senior, *J. Biol. Chem.* **271**, 4335–4341 (1996).
63. R. Sadhukan, G.C. Sen, and I. Sen, *J. Biol. Chem.* **271**, 18310–18313 (1996).
64. H. Kotake, Q. Li, T. Ohnishi, K.W.S. Ko, L.B. Agellon, and S. Yokoyama, *J. Lipid Res.* **37**, 599–605 (1996).
65. M.J. O'Donohue, G. Boissy, J-C. Huet, C. Nespoulous, S. Brunie, and J-C. Pernollet, *Protein Exp. Purif.* **8**, 254–261 (1996).
66. K. Jaton-Ogay, J. Wyniger, D. Sanglars, S. Foundling, and M. Monod, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 32.

67. B.K.L. Sim, H. Liang, R. Lapcevich, J.W. Madsen, W. He, A.H. Fortier, S. Plum, A. Ruiz, and M.S. O'Reilly, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 33.
68. K. Schwarz, S. Borngraber, and H. Kuhn, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 57.
69. J.A. Whyte and C. Cunningham, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 62.
70. S. Masure, L. Paemen, P. Proost, J. Van-Damme, and G. Opdenakker, *J. Interferon Cytokine Res.* **15**, 955–963 (1995).
71. Y. Laroche, V. Storme, J.D. Meutter, J. Messens, and M. Luwereys, *Bio / Technology* **12**, 1119–1124 (1994).
72. M.D. Moody, J. Ray, K. Murphy, and D. Tannatt, *Current Topics in Gene Expression Systems*, Invitrogen Inc., San Diego, 1996, p. 51.
73. M. Rodriguez, R. Rubiera, M. Penichet, R. Montesinos, J. Cremata, V. Falcon, G. Sanchez, R. Bringas, C. Cordoves, M. Valdes, R. Lieonart, L. Herrera, and J. de la Fuente, *J. Biotechnol* **33**, 135–146 (1994).
74. K. Sreekrishna, *Industrial Microorganisms: Basic and Applied Molecular Genetics*, American Society of Microbiology, Washington D.C., 1993, pp. 119–126.
75. K.A. Barr, S.A. Hopkins, and K. Sreekrishna, *Pharm. Eng.* **12**, 48–51 (1992).
76. K. Sreekrishna, W.R. McCombie, R. Potenz, K.A. Parker, P.K. Mazzaferro, G.M. Maine, J.L. Lopez, D.K. Divelbiss, K.A. Holden, R.D. Barr, and M. Fuke, *Fifth Annual Biotech USA Industry Conference and Exhibition*, San Francisco, 1988.
77. C.A. Scorer, J.J. Clare, W.R. McCombie, M.A. Romanos, and K. Sreekrishna, *Bio / Technology* **12**, 181–184 (1994).
78. A.C. Swart, P. Swart, S.P. Roux, K.J. van der Merwe, I.S. Pretorius, and A.J.C. Steyn, *Endocrine Res.* **21**, 289–295 (1995).
79. J.M. Cregg, K.R. Madden, K.J. Barringer, G.P. Thill, and C.A. Stilton, *Mol. Cell. Biol.* **9**, 1316–1323 (1989).
80. L.K. Shay, H.R. Hunt, and G.H. Wegner, *J. Ind. Microbiol.* **2**, 79–85 (1987).
81. Y. Chen, D. Pioli, and P.W. Piper, *Bio / Technology* **12**, 819–823 (1994).
82. K. Sreekrishna, J.F. Tschopp, and M. Fuke, *Gene* **59**, 115–125 (1987).
83. K. Sreekrishna, T.D. Webster, and R.C. Dickson, *Gene* **28**, 73–81 (1984).

See also CHINESE HAMSTER OVARY CELLS, RECOMBINANT PROTEIN PRODUCTION; EXPRESSION SYSTEMS, E. COLI; EXPRESSION SYSTEMS, MAMMALIAN CELLS; INSECT CELL CULTURE, PROTEIN EXPRESSION; INSECT CELLS AND LARVAE, GENE EXPRESSION SYSTEMS.

PILOT PLANTS, DESIGN AND OPERATION

BETH H. JUNKER
Merck Research Laboratories
Rahway, New Jersey

KEY WORDS

Biopharmaceutical
Bioprocessing
Pilot plant
Process development
Scale-up

OUTLINE

INTRODUCTION

Several diverse roles, objectives, and purposes are associated with a bioprocessing, biochemical, or biopharmaceutical pilot plant. These include, but are not limited to, manufacture of clinical supplies, process development for new products in the company pipeline, process improvement including technical support for existing products on the market, and production of biologically produced materials (which are not intended for the clinic) to support basic research efforts such as screening, assays, and disease mechanism elucidation. Unlike chemical pilot plants, pharmaceutical pilot plants are used not only to scale up laboratory processes but also to produce developmental quantities of chemicals (and biological substances) for safety, toxicological, and clinical studies (1). Thus, balancing the demands for scale-up and development research with those for clinical production is a key dilemma for pilot plant operation (2). The timing of clinical production is critical because expensive and time-consuming clinical programs can be delayed, thus postponing product application filings, approvals, and launches (2). Although traditionally pilot plants have not manufactured material for sale, accelerated launch dates, deferred manufacturing facility capital commitments, and other areas of project uncertainty have prompted consideration and use of pilot plants for manufacturing in recent years.

The types of products manufactured using biologically based methodologies include primary metabolites, secondary metabolites (ranging from pharmaceuticals to biopolymers), constitutive proteins, recombinant proteins, intracellular or extracellular enzymes and biocatalysts, antigens, viruses, polysaccharides, and DNA itself. These products are made using a wide range of cell cultivation and isolation techniques suitable for bacteria, yeast, fungi, actinomycetes, suspension animal cells, anchorage-dependent animal cells, insect cells, and plant cells. Processing equipment varies considerably among these diverse systems.

Designing for extensive processing flexibility can be expensive, potentially in excess of $100 million dollars, depending on the specific process requirements, size, and intended use of the pilot plant. Processing requirements can change dramatically as products are deemphasized and reemphasized in development. These decisions on priorities are likely not to result from the robustness of the process but rather from the state of the clinical data or potential market share. The ability to process multiple products in an acceptable manner typically is desirable to maximize facility use. Although designers may need to predict future types of processes and products, they should be cautious about overdesigning the facility in terms of containment and flexibility (3).

The general areas required in a pilot plant might include support labs for bench-scale development and quality control, production suites for scale-up, utility areas, and personnel offices, as well as ample storage areas for raw materials, supplies, product, samples, and documentation (2). There should be a clear division between controlled and noncontrolled areas, by utilizing signs at a minimum and perhaps by physically restricting access. General design criteria might focus on modularity, simplicity, future growth provisions, redundancy, construction costs, operating costs, and air and water waste containment (4). Future operating costs are an important consideration for pilot plants because ample funds are necessary for raw materials, staffing, maintenance, and additional capital modifications to permit improvements.

OPERATIONAL CONCEPTS AND DESIGN REQUIREMENTS

Several operational concepts influencing design requirements need to be resolved before initiation of design activities. Agreement should be sought among end users, operations staff, compliance auditors, environmental consultants, and safety officers as well as from appropriate levels of management directly or indirectly influencing and controlling capital expenditures. A written document summarizing consensus on these issues may then become the basis for conceptual design activities.

Major decisions influencing pilot plant capabilities include whether the facility will produce bulk drugs (thus falling under the regulation of Center for Drug Evaluation and Research (CDER) and following guidelines from 21CFR 210/211) or biologics (thus falling under the regulation of Center for Biologics Evaluation and Research (CBER) and following guidelines from 21 CFR 210/211 and 600) (5). The level of validation and the validation philosophy are directly related to the end use of the resulting pilot plant products, specifically whether they are used for process development, clinical supplies (presafety/basic research, safety, phase I, phase IIa, IIB, phase III), or material for sale. Validation strategies, as well as other operational and design issues, are also influenced by whether the pilot plant is administratively located within the manufacturing or research division. Individual company preferences and guidelines for similarity between manufacturing and pilot plant areas need to be identified and evaluated for their impact on the proposed pilot plant facility. Typical areas targeted for standardization within a company are vessel designs, computer systems and interfaces, equipment vendors, and construction specifications.

Design requirements center around the desired initial and end product stages of processing, specifically the forms of the pilot plant inputs and outputs. Example stages include master/working seed preparation, fermentation broth, initial captured product (e.g., microfiltration retentate, ultrafiltration concentrate, cell paste), crude product (e.g., rich ion exchange fraction), finished isolated or sterile bulk, and filled vials. Ambient, refrigerated, and frozen storage requirements based on these inputs and outputs then need to be addressed in the areas of seed, raw materials, intermediates, clinical supplies, and process samples.

The scale of the pilot plant relative to current or proposed manufacturing scale and relative to laboratory scale needs to be established. Generally, a 1:5 or 1:10 scale-up is common, but smaller ratios can be advantageous in terms of minimizing risk during factory start-up. The appropriate level of biosafety also must be considered, with good large scale practice (GLSP) being sufficient for most recombinant work. Some pilot plant areas may require biosafety level 1–large scale (BL1-LS) and/or biosafety level 2-large scale (BL2-LS), and careful consideration of multiple-use/decontamination issues for BL1- and BL2-containing facilities is necessary. Solvent use areas need to be identified and minimized if possible because cost implications for explosion-proof design can be substantial.

Primary motivations to design a multiuse pilot plant are the conservation of capital and improved response time for clinical manufacturing needs for new development products. The strategy for multiuse should be identified and considered throughout the design process. Common strategies (6) include using totally dedicated equipment and limiting production to one product at one particular production stage, using campaigned equipment in which multiple products are sequentially processed in the same equipment "campaign style," with documented product changeover procedures completed to minimize cross-contamination potential, and conducting concurrent manufacturing in dedicated equipment in which several products are simultaneously produced in segregated areas with segregation achieved via physical separation or closed systems. These strategies result in minimization of cross-contamination potential through engineering controls, standard operating procedures (SOPs), or temporal segregation (7). Typical pilot plants combine the second and third strategies by providing segregated and dedicated ar-

eas for certain types of processing (e.g., virus and nonvirus areas, GLSP versus BL1/BL2 areas, beta-lactam antibiotic processing, heat-resistant spore-former cultivation) (8). Example layouts for product-specific and common areas depending on the strategy selected have been described (9).

Guidelines for the use of open versus closed systems for various processing stages need to be devised and justified based on available and cost-effective technologies as well as quality requirements. In a closed system, the product is not exposed to the immediate environment. The quality of materials entering a closed system (such as water, steam, or air entering a bioreactor) is controlled by its quality as well as the manner in which these materials are entering the closed system (such as filter sterilization, autoclaving, or steam sterilization of system connections) (6). In an open system, the product is exposed directly to the surrounding environment, thus necessitating the creation of a controlled or sometimes even a closed system around the open system. One example of this strategy is the transfer of seed vial contents in a biosafety cabinet. Other examples might include the supply of higher-quality filtered air to areas for the cooldown of autoclaved materials and assembly of clean equipment. Thus, levels of containment can be constructed until all open systems are contained within controlled or closed systems to meet quality requirements for the specific processing stage.

The specific issue of potential exposure of product to spore-forming organisms has been debated extensively. The closed system concept appears to minimize the concern for airborne spores typically found in the environment. Cultivation guidelines for spore-forming organisms within a multiuse facility appear to have been redefined to evaluate the robustness of the spore type formed. Careful consideration of segregation has been given when cultivating organisms such as *Bacillus* that form difficult-to-kill spores. A separate, self-contained area with no shared equipment, air systems, or entrances minimizes exposure concerns for other products. The relatively heat-sensitive, substantially larger, fungal spores have not been highlighted as a substantial cross-contamination concern. Thus, segregation for these types of processes is not as critical, although it still might be implemented due to company preferences for separate bulk drug and biologics processing areas.

DESIGN

The initial generation of flow charts for a bioprocessing, fermentation, or isolation pilot plant is best conducted based on model processes that are developed and critiqued by several end users of the equipment and facility. These model processes should be carefully selected as to give adequate and appropriate representation to current and reasonable future needs of the pilot plant. As they form a preliminary but sound basis for equipment and utility sizing, they should include unit operations as well as expected cycle times. In some cases, simulation (10–12) or actual laboratory- or pilot-scale testing (13) of critical equipment under typical processing conditions or with expected process streams may be warranted to further define design specifications. Example model processes have been published for vaccines (14) and intracellular and extracellular products (15–18). Resulting flow charts can focus on raw materials for fermentation or isolation, seed development, fermentation, recovery, crude product isolation, final product isolation, product finishing, buffers, utilities, and waste (including chemical waste, biological waste, and any necessary offgas treatment) (15).

The requirements for these model processes, coupled with an estimate of the required facility output, can be used to construct a preliminary architectural layout that divides the facility into sections (termed modules, cubes, or suites). These sections are surrounded by controlled corridors and accessed via airlocks (3). Based on this initial layout, flow patterns are developed to permit the logical, usually unidirectional, flow of personnel, raw materials, product, waste, and clean or used equipment throughout the facility. Flows are designed primarily to minimize cross-contamination from multiple products and from clean and used equipment while maintaining access and flexibility (3). This separation often is accomplished by using clean and return corridors located on either side of the processing area as well as by temporal segregation using appropriate SOPs. Evaluation of these flow patterns should be conducted carefully and should incorporate end users as well as quality auditors. In some cases, companies may solicit outside good manufacturing practice (GMP) reviews of the project at this time from external consultants, a second design firm, or even FDA representatives. Gathering and organizing up-to-date accurate information and comments about the impact on model process needs are critical factors in forming a successful basis of design (19).

Example layouts for process flows for biotech facilities have been published (6,20,21). An extensive reference (including (piping and instrument diagrams (P&IDs), layouts, and flow diagrams) for the design of biopharmaceutical plant equipment, utilities, heating, ventilation, and air-conditioning (HVAC), and waste treatment also has been published (22). Although it was based on acceptable practices in 1991, the strategies and concepts for design presented are still quite relevant. Adequate space around the equipment needs to be reserved for operations and maintenance access as well as for related portable equipment. The movement, segregation, storage, and cleaning of portable equipment such as vessels, skids, hoses, and carts should be fully examined.

Various reports of successful and problematic aspects of project design, construction, start-up, and validation have been published. Cost control issues and the benefits of formulating a master schedule (23) need to be reviewed in terms of their potential validation impact. Construction concerns and "lessons learned" from multiple biotech installations also have been summarized (24). One case study of a laboratory facility highlights the need for teamwork among client, design firm, and construction manager (25). This trio can be extended to include the validation contractor in an effort to minimize "finger pointing" and "blame storming" when unforeseen problems are uncovered. Regardless of who may actually be perceived as being "at fault," the company owning the facility ultimately is adversely affected by substantial start-up and validation

delays. A comprehensive overview of biochemical pilot plant design and operational guidelines appeared in the mid-1980s (26).

Process Equipment

The major decision regarding process equipment is that of similarity versus diversity as it applies to specific process steps as well as multiple products. The extent of commonality required in equipment specifications is an important operational factor to be considered because equipment specifications are based on inputs from model processes. For a pilot plant to accept a manufacturing interface role, at least a portion of the equipment design and selection should match that of the factory which the pilot plant is intended to support. For example, if the factory intends to use continuously sterilized medium for large-scale fermentation, then there should be provisions in the pilot plant to perform continuous sterilization as products get closer to the transfer to production. Alternatively, depending on individual company preferences for the pilot plant role, later clinical batches (phase II and beyond) might be manufactured in the factory to facilitate eventual process transfer and ensure consistency of equipment.

Skid-mounted equipment has been often selected in recent installations because it puts the design burden entirely on one vendor, who specifies the appropriate quality, supply rate, and pressure of the required utilities. Skid-mounted equipment can be fixed or portable, which permits different types of processing in the same space as necessary. Consistency in skid designs with respect to components and spare parts may be a drawback as vendors might not be willing to build skids to individual company specifications without added costs. This situation can be particularly problematic if the skid manufacturer has arranged a low-cost deal with a vendor (e.g., diaphragm valves) or programmed software on a programmable logic controller (PLC) that may not be the vendor of choice for the rest of the plant. There also may be incompatibilities in piping, utility requirements, or software when the skid is connected (27) if miscommunications occurred between the skid vendor and the facility designer.

The traditional alternative to skid-mounted designs is to purchase the equipment components separately and request the facility designer to develop the system. One advantage to this approach is that total flexibility in equipment selection now rests with the pilot plant designers, who may be quite confident in their ability to design the equipment based on their extensive prior experience. Drawbacks of this approach include placement of the design burden on the facility as well as added design costs for one-of-a-kind systems. Specific references are available for flow charts and detailed designs for fermentation and harvest equipment (28–31) as well as for filtration units (32), chromatographic equipment (20,33,34), and centrifuges (35,36). A list of typical process, utility, and support equipment that might be required in a bioprocessing pilot plant has been compiled (Table 1).

For a bioprocessing facility, certain aspects of equipment and piping specifications should be uniform throughout the installation where possible. These specifications might include requirements for orbital welding of lines, documented weld inspections, passivation, pressure testing, self-draining (sloped) piping, minimal lengths for dead legs and pockets, restricted use of flanged connections, level of polish on product contact surfaces, stainless steel quality (typically 316L for product contact surfaces), and use of FDA-approved polymers in specific applications (gasket, o-rings, valve seats, distribution piping). Proper storage of fabricated piping awaiting installation as well as prominent identification of valves and lines also must be considered. Valve specifications can be important to determine early in the design estimation phase. Valve capital cost contribution is high because large numbers are required and any installed valves (either diaphragm or ball) need subsequent maintenance. Consideration should be given as to whether live steam, set up as a steam block, should be used to continually purge the back side of product contact valves and whether live steam should be used for tracing in between double o-rings on a port or manway (28). While such designs traditionally have increased sterility assurance for lengthy secondary metabolite fermentations, they may create detrimental localized hot spots during animal cell cultivation. Neither the equipment itself nor its disposable or replaceable parts should release any extractable substances (after the initial postinstallation rinsing and cleaning steps) into the product (28). An overview of sanitary piping design, installation, and validation concerns is available (37).

Transfer piping strategy is another major design decision. One common strategy is for several transfer lines to meet at a transfer panel or process manifold, typically located at a high point between two or more areas, and be connected using spool or “jumper” pieces to move material in the desired fashion (3). Steam entry is at the high point and condensate drainage at the low point for steam-in-place (SIP) during sterile transfers. Clean-in-place (CIP) solutions are passed along the direction of the transfer. Advantages to this approach include reduced piping distribution costs and space requirements, which can multiply rapidly when multiple tanks and multiple types of process transfers are required. Disadvantages center on operational restrictions concerning the elapsed time between successive transfers through a common manifold system as well as cross-contamination associated with transfers of nonsterile media/buffer, sterile media/buffer, inoculum, harvested broth, CIP solutions, and water through common lines. A listing of product and nonproduct transfer line/utility piping that might be required in a bioprocessing facility has been compiled in Table 2.

Both sterilization and cleaning sequences need to be defined and documented during the equipment and transfer piping specifications because retrofits after installation can be time consuming and expensive. P&IDs should be traced through these sequences before approval by both operational and validation personnel. Requirements for validation testing and in-use monitoring of equipment operation also need to be identified during the equipment specification phase. Examples include sizing condensate lines to permit adequate drainage of condensate when spore strips are present, specification of additional validation ports for thermocouples on vessels and condensate

Table 1. Types of Processing Equipment and Utilities Possibly Required for a Bioprocessing Pilot Plant

Fermentation and harvest	Isolation	Utilities	Support
Culture storage freezer	Low-pressure chromatography columns	Clean steam generator	Sterilizing autoclave
Incubator/shaker	High-pressure chromatography columns	USP purified water system	Decontamination autoclave
Biosafety cabinet	Solvent blending and delivery system	WFI system	Depyrogenation ovens
Laminar flow hood	Sterile filtration assembly	Deionized water	Laminar flow hoods for cooldown and equipment assembly
Seed fermentor(s)	Laminar flow hood	Process (city) water	Glasswasher(s) (glassware, vials, stoppers)
Production fermentor(s)	Liquid transfer pumps	Compressed clean air	Buffer preparation tanks
Nutrient feed tank(s)	Dispensing machines	Compressed instrument air	Media preparation tanks
Harvest tank(s)	Chemical fume hood	Compressed plant air	In-line mixer
Microfiltration skid	Portable tanks	Compressed gases (O_2, N_2, CO_2, NH_3)	In-process sample refrigerator/freezer
Ultrafiltration skid	Laboratory-scale centrifuge	Electricity	Raw material, cold/frozen storage
Pilot-scale centrifuge	Ultracentrifuge	Chilled water	Floor and bench scales
Homogenizer	In-process refrigerator/cold room/freezer	Cooling tower water	Process and equipment monitoring/alarm systems
Continuous sterilizer	Product refrigerator/freezer	Plant steam	Uninterrupted power source (UPS)
Filter integrity tester	Syringe/vial filling equipment	Glycol	Offgas analyzer (mass spectrometer)
Sterile tubing welder	Lyophilizer	HVAC and building automation system	Bar-code system
		CIP skid(s)	
		Biowaste inactivation system	
		Environmentally potent chemical destruction system	
		Dust collection system	

Table 2. Transfer Lines and Utility Piping Possibly Required in a Bioprocessing Pilot Plant

Product contact	Nonproduct contact
Utilities	Utilities
Clean water	Chilled/cooling water and/or glycol
Clean air	Instrument/plant air
Clean steam	Plant steam
CIP solutions	Contained sewer
Other compressed gasses (CO_2, O_2, N_2, NH_3)	Chemical (noncontained sewer)
Process	Sanitary water
Sterilized medium or buffer	Vacuum system
Sterilized nutrient/acid/base/antifoam feeding	Natural gas
Nonsterilized medium or buffer	
Inoculum	
Harvest	

lines, installation of product contact utility sampling stations (especially at worst-case locations), and installation of temperature sensors on critical lines to monitor sterilization and proper trap operation.

Utilities

Product contact utilities might include water, clean steam, compressed air and other gases, and CIP systems. Nonproduct contact utilities generally include chilled water, instrument air, plant steam, and plant air, although in some cases plant steam may be appropriate for product contact (Table 3). A listing of product and nonproduct contact utilities is compiled in Table 2. Multiple utility use points are located within and among suites for both product and nonproduct utilities; these are organized into utility stations for each equipment skid. To minimize piping, it might seem convenient to use product contact utilities in nonproduct contact applications such as compressed clean air on a vessel jacket, relying on a check valve for backflow prevention. This approach can risk disaster because the operation and quality of utilities can affect processing throughout the entire facility. Cross-contamination for product contact utilities among suites is rare because of positive pressures of utilities piping into each module.

Peak utility loads should be based on carefully investigated assumptions that are transferred from design directly to validation testing. Although it may be possible to run several large fermentors simultaneously, each at maximum airflow rate, it may be less likely that several large fermentors are cooled simultaneously after completion of sterilization. Thus, air compressors might be sized assuming maximum load while process chillers might be sized on the basis of an operationally realistic fractional load. Assumptions about processing cycle times and tank volumes are important particularly when sizing CIP systems and

Table 3. Example Plant Steam and Clean Steam Achievable Quality Comparisons

Quality attribute	Clean steam	Plant steam
Bioburden	<0.1 cfu/mL	<0.1 cfu/mL
Total organic carbon	<0.5 ppm	<2.0 ppm
Endotoxin	<0.25 EU/mL	<0.5 EU/mL
pH	5–7	5–8
Conductivity	Meets USP 23 Suppl. 5 standards	N/A

liquid waste decon systems. Shorter (20-h) *Escherichia coli* fermentations might require a smaller tank volume [1,000 L] but are run more frequently, compared with longer (400-h) fungal fermentations, which might require a larger tank volume (20,000 L) but are run less often.

Product Contact. The most critical product contact utility is often water because of the sheer volume needed during typical processing steps. It is well known that the quality of water needs to be appropriate for the intended application, but water system design decisions often are conservative. In many installations, purified water is used for upstream processes while WFI is reserved for downstream isolation and for cultures sensitive to endotoxin such as animal cells (38). A brief overview of various methods to produce pure water is available (39).

Water-for-injection (WFI) is often favored for the entire facility for uniformity. Specifications are set for parameters such as bioburden (typically 0.1 cfu/mL) and endotoxin (typically 0.25 enzyme units, EU/mL) as well as USP chemical tests that include conductivity, pH, and total organic carbon (TOC). WFI systems containing stills generally are designed as a hot 80 °C loop with point-of-use coolers that can provide water at a user-selected temperature. These coolers can form deadlegs and thus require substantial flushing before use. It is also possible to design a cold or ambient WFI loop with periodic (typically daily) heating to 80 °C or ozonization for sanitization purposes. In some applications, reverse osmosis (RO) systems are used, but care of the membranes is critical to consistent operation.

USP purified water is an attractive alternative to WFI for various upstream applications. Specifications include bioburden (typically 100 cfu/mL) and USP chemical tests. Although such systems can be operated hot or cold, recent success with ambient systems sanitized with ozone and ambient systems sanitized chemically (40) has been reported. Other noncompendial grades of water may be required. In these cases, quality attributes typically are established by the facility to address the needs of specific processing steps (Table 4). Examples might be deionized water with bioburden (typically <500 cfu/mL) and resistivity (>1 Mohm) specifications or process water (municipal water passing through a break tank).

General design elements for water systems include the desired flow rate at each use point and the desired number of simultaneous use points in operation. Use-point positioning should consider the height of the take-off valve, automation of take-off valves, and sink or sewer location for flushing, as well as cleaning, sanitization, and acceptable noncoiled storage for associated use-point hoses. These values are combined to obtain the total "take-off" flow rate; the system and piping are then sized such that velocity of the remaining flow rate is sufficient to minimize bioburden and biofilm growth. Suggested methods for obtaining this residual velocity have been outlined (41). In addition, water systems should be designed with specific sanitization procedures identified at the outset. Steam, ozone, or chemical sanitization methods are most common, although steam sanitization generally has not been recommended for plastic loop systems because of concerns about distribution pipe sagging. Plastic systems utilizing polyvinylidene fluoride (PVDF) instead of stainless steel are becoming more common and effective at achieving even WFI conditions. A direct comparison of plastic piping with stainless steel with respect to ion leaching, smoothness, and cfu counts has been made (42).

Clean steam can be generated from a purified water feed using plant steam for supplying heat of vaporization. Increased heat exchanger fouling may occur if a controlled feed water supply is not used. Clean steam distribution piping is designed with regulators to reduce the pressure from 5 atm down to about 2–2.5 atm as needed for typical SIP operations. Sanitary pressure gauges, placed near or on equipment skids, can assist in determining dynamic pressures during SIP cycles. Steam sampling stations generally take the form of a removable sanitary trap to which a sample cooling device (typically a heat exchanger) is attached. Consequently, a source of cooling water is necessary, either piped from local lines or transported by gravity via a portable holding tank about 20 L in volume. Traps should be placed in header branches and equipment supply lines such that condensate buildup does not occur when skid equipment is isolated for repair. Periodic cleaning of the clean steam system assists in cleaning fouled heat exchange surfaces and in minimization of rouging.

Compressed air in product contact generally is used for fermentation sparger air supply and pressurized transfers. Oil-free compressors are favored, with intakes positioned away from other building effluents. Heating up to and holding at 200 °F (93 °C), typically accomplished by design of the compression ratio and residence time in a receiver, is desirable to minimize airborne biological contaminants such as bacteriophage, although there is some evidence that submicrometer (0.1 μm) filtration also can be effective. Humidity can then be reduced from the compressed air using refrigeration and desiccant dryers. Moisture removal is desirable to a dew point of -40 °C. Air is then microfiltered (0.2 μm) at the point of use to remove microbes and particulates. Compressed air may be sampled for any or all of the following depending on the application: identity, hydrocarbon level, microbial content, and particulates. Other gases may be required for the facility such as oxygen enrichment for microbial cells, carbon dioxide for pH control of bicarbonate-buffered media, nitrogen and oxygen supplementation for animal or insect cells, and appropriate gas blending and control systems. Planning should consider storage of bottled gas cylinders outside the facility or in segregated maintenance areas to minimize the transport of unclean cylinders in controlled locations.

Table 4. Comparison of Example USP Purified Water, Deionized Water, and Process Water Achievable Quality Attributes

Quality attribute	USP purified water	Deionized water	Process (city) water
Bioburden	<100 cfu/mL	<500 cfu/mL	<500 cfu/mL
Coliforms	Absent	Absent	Absent
Total organic carbon	<0.5 ppm	<2.5 ppm	<5 ppm
Endotoxin	<0.25 EU/mL (for information only)	N/A	N/A
pH	5–7	5–8	5–8
Conductivity/resistivity	Meets USP 23 Suppl. 5 standards	>1 Mohm	N/A

CIP skids can be portable and moved into place in the suite at a designated utility location or one or more fixed main skids can be located permanently in the utility area. CIP systems are composed of supply and return pumps, one or more cleaning agent dilution tanks, provisions for heating cleaning agents, and a PLC for sequence control and associated data acquisition. Because multiple tankage and pumps are required (20), these skids often are located in utility areas where acid- or caustic-based cleaning agents arriving in drums can be suitably diluted. More than one CIP skid may be necessary to provide product segregation and minimize cross-contamination. Multiple skid requirements depend on design decisions concerning the level of containment and product segregation (3) as well as capacity.

Nonproduct Contact. The type of cooling/heating system desired, while not in product contact, has widespread design implications. Recirculating temperature control loops containing high surface area/unit volume heating and cooling heat exchangers, an expansion tank, and a circulation pump are common. These systems rely on indirectly heating or cooling of jacket fluid so their response can be slow, particularly if heat exchangers were not sized according to user expectations. The involvement of heat exchangers means that practical application may be limited to vessels a few thousand liters in size. Jacket fluids can be water, glycol, or novel fluids with wide temperature ranges. Operation of the circulation pump and internal (as well as external) integrity of the heat exchangers are also critical. Cooling heat exchangers can use either chilled water or glycol, whereas heating heat exchangers generally use plant steam. Dual parallel filter housings, containing 10 to 50-μm elements, are advantageous to reduce heat exchanger fouling. Sediment buildup on filters can be monitored using measured pressure differentials across the filters, and the assembly can be designed so that one filter may be replaced while maintaining flow to the second filter.

Direct application of heating or cooling fluids to process jackets is the traditional alternative. Consecutive application of chilled water and steam on the jacket is acceptable if the design evacuates the jacket using higher-pressure compressed air such that banging is minimized. Use of glycol and steam on the same jacket can have environmental implications from glycol losses. Regardless of the indirect or direct mode of application, jacket design can greatly affect heat transfer effectiveness. Dimpled or straight jackets are preferred for smaller vessels, with external half-pipe cooling coils used in larger vessels. The largest manufacturing scale vessels utilize internal cooling coils but these present additional internals for cleaning and increased sterility risks, particularly at welds, because cooling water can be at a higher pressure than the fermentor contents.

Filtered dried instrument air, typically at a pressure near 100 psi, is required for solenoids. Pressure specifications for control valve operation vary considerably among vendors. Plastic hosing, although convenient and cost effective, should be evaluated carefully for instrument air lines, especially on hot equipment where heat-induced leaks can develop and in solenoid cabinets where crimping can slice the tubing. A backup instrument air compressor can minimize operational consequences from supply pressure dips. A separate compressed air source, at a higher pressure than chilled/cooling water and jacket steam, is desirable for evacuation of vessel jackets to minimize risk of backflow into the instrument air system.

Natural gas usage, if required, might best be addressed using bottled gas for safety reasons, which discourage distribution piping throughout the facility. Favored for flaming of openings in many microbiological procedures, the use of Bunsen burners in a biosafety cabinet or laminar flow hood should be examined carefully with company safety department representatives. Care should be taken to minimize the potential for reaching the lower explosive limit near the motor of the cabinet or hood. Compatibility of plastic laboratory ware with flames also should be ensured.

HVAC. HVAC is a key and costly component of the facility. Its design and operation needs to be understood by operating, maintenance, and quality assurance personnel. Requirements for air cleanliness, as well as the ability to achieve them, are determined by HEPA filtration, airflow volume, and pressurization room design as well as the type of operation being conducted (Table 5) (3). HEPA filters are often used for inlet air but are only used for exit air when required by biosafety containment. The specific classification of processing areas depends on processing step quality concerns. Classification level, representing the maximum number of particles permitted per cubic foot of air sampled, of an area increases from fermentation (class 10,000–100,000; 20–40 room changes per hour) to purification (class 1,000–10,000; up to 100 room changes per hour) (3). Within a room there are cleaner areas surrounding open transfers (e.g., biosafety cabinets or laminar flow hoods) that are class 100 (540 changes per hour) (43). For viable particles, the range is from class 100,000, which permits 2.5 cfu/ft^3, down to class 100, which permits 0.1 cfu/ft^3 with

Table 5. Example Controlled and Noncontrolled Area Achievable Quality Comparisons Based on Room Design

Type of area	Pressurization	Airflow changes (#/h)	Inlet filtration	Nonviable (#/ft)	Viable (cfu/ft^3)
Bulk fermentation processing	None	60	Coarse	<250,000	<25
Laboratory	Positive	60	Coarse	<100,000	<2.5
Laboratory	Positive	10	HEPA	<100,000	<2.5
Laboratory, clean construction	Positive	25	HEPA	<100,000	<2.5
Laboratory, clean construction	Positive	20	HEPA	<10,000	<0.5
Laboratory, contains autoclave	Positive	100	HEPA	<100,000	<2.5
Laboratory, contains glasswasher	Negative	50	HEPA	<100,000	<2.5

intervening levels generally determined by individual application (43). Sufficient air velocities also are necessary to dilute particulate contaminants. In some cases, separate air handlers are considered for critical sterile areas, inoculum preparation areas, high particle or dust areas, and segregated product areas. In other instances, airflow is designed in a once-through pathway, which may increase heating and cooling utility costs substantially.

Inlet air enters from the top of the room, flows downward, and exits from ducts near the floor. Any recirculation, which may be desirable to minimize capital costs for heating and cooling equipment as well as operational costs, should be only from the room where the air originated, limited to areas that do not generate substantial amounts of particles, and pass through inlet HEPA filters again. This strategy is beneficial to the life span of the HEPA filters and the ease of maintaining constant temperature and humidity (3). Exit ducts should be placed near equipment within a processing area that generates the most particulates or aerosols so as not to spread them throughout the area. Specific examples include autoclaves, glass washers, and cell-disrupting homogenizers.

Room pressurizations should be higher for cleaner areas and lower for areas that generate particles (3). Higher pressurization levels should be assigned to airlocks between adjacent areas for which it is desired to minimize cross-contamination. Relative pressurizations between rooms sharing a common wall should be considered to supplement or minimize reliance on completely sealing rooms during their construction. The use of door sweeps and gasketted ceiling tiles to achieve and maintain pressurization should be evaluated relative to floor cleanability and ceiling maintenance access, respectively. Pressure differentials might be as high as 0.03 in of water between classified areas and as high as 0.05 in between classified and nonclassified areas. Lower pressure differentials might be acceptable depending on the specific product so long as the direction of airflow remains consistent. Interlocks, control, monitoring, and alarm building automation systems may be designed to reduce pressure losses when doors are opened and to alert users to unsatisfactory conditions.

Facility temperature is generally maintained on the cool side because gowning tends to make workers warm. Inaccurate assumptions concerning the impact of radiant heating may cause temperatures to be higher than desired on certain days of operation. Humidity should be comfortable to avoid drying of workers' respiratory membranes during extended hours. Care should be taken to specify tolerance limits that have quality implications relevant to the processing steps expected to be conducted in that area to avoid unnecessary facility costs. While specific example guidelines are 23 ± 3 °C for temperature and 40–60% relative humidity (44), substantially larger ranges may be acceptable for many processing applications. An extensive P&ID design of HVAC for clean spaces is available as a published reference (43). In addition, consideration should be given to Federal Standard 209E, Airborne Particulate Cleanliness Classes in Clean Rooms and Clean Zones, as well as European Economic Community (EEC) regulations that establish parallel categories of Grade D (100,000), Grade C (10,000), and Grades A + B (different designs of class 100) (43).

Containment

Similar issues exist for product, environmental, and personnel protection. Often the resulting design specifications overlap with those items necessary for sterile or sanitary operation. This concept is addressed in the literature (45) in which solutions to prevent microbial transport for the purposes of containment were compared with those to maintain sterility for fermentors. A unified approach is recommended with the goal of minimizing unnecessarily stringent specifications that can substantially augment capital and operating costs. Environmental and personnel protection issues should be evaluated for each new process undertaken by the facility, along with product protection quality concerns. Samples of broth, and other product, intermediate, and waste streams, should be submitted for aquatic and human toxicity testing in advance of processing in the pilot plant when appropriate. Preliminary laboratory data then might be available to evaluate potential hazards in conjunction with published documentation concerning the relationships of the questionable compounds, raw materials, or organism with known hazards.

Product Protection. The major concept for product protection is the closed system in which all material entering or leaving the process must be controlled in a specified and documented manner. This requirement extends to product contact utilities such as purified water, clean steam, compressed air, and controlled/monitored room air so that when equipment is opened for cleaning the quality of the surrounding environment is consistently at a known and documented level. The equipment is cleaned to an acceptable quality level. After sterilization, material enters the closed system via a sterilizing filter, steamed connection to an autoclaved bottle assembly, or sterilized transfer line.

Offgas passes through a 0.2 μm steam-sterilized vent filter, often contained in a low-pressure steam-jacketed housing preceded by a coalescing filter, condenser, or vent heater to avoid pluggage from moisture. Double mechanical seals are utilized with a barrier fluid of a controlled quality (typically condensed clean steam) at a higher pressure than the vessel. A closed system also can be used to surround an open operation such as the use of isolator technology (i.e., glove box) as an alternative to clean rooms or biosafety cabinets (46,47).

Environmental Protection. Environmental protection concerns focus on emissions in air, liquid, and solid waste. State air permits typically are required for new vessels and need to be altered if additional air treatment devices such as vent filters, incinerators, or scrubbers are added. Hydrophobic vent filters can provide emergency foam control and prevent external product release. These filters can be monitored for pluggage over cultivation time using pressure differentials with plugged filters able to be replaced during a run only if a bypass line has been installed. Liquid emissions include accidental spills/sewerings, which can be minimized by locking vessel bottom valves, blocking open sewer connections, collecting sample line flush waste, and providing clear written instructions to personnel regarding batch disposition using batch sheets and appropriate signs. Dikes, troughs, and vaults can be used to contain large spills or divert inadvertently sewered material. Sample waste can be minimized using low-volume sampling devices. Thermal or chemical (less common) batch or continuous kill tanks need to be considered for biological hazards. For chemical hazards, other appropriate waste destruction and minimization systems may include hot caustic hydrolysis, reverse osmosis, and evaporative condensers. Batch leakage into cooling water utility systems can be minimized by pressurizing utility systems higher than the batch, but this might introduce contamination if jacket cooling coils develop leaks. Alternatively, the utility system return headers can be isolated from general use until their contents are treated and/or tested. This solution might be attractive for cooling water as well as condensate, especially when it is returned to river sources. Care must be taken to avoid sequestering large quantities of potential waste for long periods of time. Solid waste can be controlled procedurally using clearly labeled containers.

Personnel Protection. Personnel protection is not as great a concern now that most recombinant organisms can be grown under typical operating conditions of GLSP. For other pilot plant processes requiring personnel protection, either the culture is rated at or above BL1-LS or a compound produced by the culture adversely affects human heath (e.g., a genotoxin, mutagen, teratogen, or carcinogen). Environmental monitoring of the operating area is necessary for the compound or culture of interest using air or swab sampling techniques. For optimal recovery, swabs should be soaked in a solvent known to dissolve the compound of interest. Screening and monitoring of personnel potentially in contact with the process (operator, mechanics, janitors, engineers, visitors) must be arranged and notification given to those personnel at unacceptable levels of risk relative to their individual health histories. A key element of personnel protection is personnel protective equipment (PPE). This equipment can include gowns, uniforms, safety glasses, lab coats, and gloves at a minimum but can be extended to include face shields and respirators as required. Spill response equipment, procedures, and training are also necessary.

Reduction to practice techniques are presented in the literature for general biocontainment issues (48) and biocontainment regulations and requirements (49). Specific examples are available for the design of waste inactivation systems (50,51), contained facility design and operation (52,53), design details for cleanliness and containment for floors, walls, ceilings, and penetrations (54), and contained sampling systems (55).

Instrumentation

The main purpose of a pilot plant is to collect data in as many ways as possible. Thus, a large fraction of the pilot plant capital allocation might be devoted to instrumentation. Traditional instruments for on-line monitoring, particularly for fermentation, include pH (56), dissolved oxygen (DO) (57–59), temperature, agitator speed, agitator power draw, backpressure, foam detection, batch level or weight, airflow rate, and vent gas analysis via mass spectrometry (60). Redundant instrumentation is desired for critical parameters such as pH and DO. Instrumentation for isolation monitoring includes monitoring of conductivity, pH, ultraviolet wavelength intensity, and refractive index detection. Useful utility instrumentation includes on-line total organic carbon (TOC) and resistivity/conductivity meters. Local readouts of transmitter values can be useful, particularly for pressure and possibly temperature, although the potential for differences between field and control room values may have to be addressed.

Many novel sensors have been developed including dissolved carbon dioxide probes (56,61), redox sensors (56), glucose sensors (62), near-infrared detectors (63,64), fiber optic biosensors (65), fluorometry of intracellular NADPH (66), and cell density probes (67–69). Another novel measurement device example incorporates fermentor side streams to measure viscosity on-line (70). A description of useful sensors may be found in the literature (71). Characteristics of on-line and in situ devices include cleanability, minimal drift, minimal interferences, versatility, sterilizability, ease of calibration, and several other factors. Utilization of any sensor should be evaluated for its effect on bioreactor sterility as well as product quality.

Automation and Control

The need for flexibility and change within the pilot plant (as compared with production) needs to be apparent in control schemes, event sequencing, data collection and presentation, and alarming (72). Increased amounts of automation restrict abilities of the end user to implement changes, but do reduce the number of manual operations and thus possibly headcount. Specifically the control system needs to readily permit set-point and alarm changes. Control systems can be distributed or centralized, with distributed control systems (DCS) favored because there are

individual controllers/computers for each unit, often connected to a central computer via a high-speed network. Thus, the loss of one centralized computer is not catastrophic (72).

While there can be "islands of automation" within a facility, connectivity in a centralized location for data monitoring, set-point/alarm changes, and data archiving might be desirable. Substantial advantages exist for localized controls on skids when many field-based inputs are required based on an operator's evaluation of field conditions. A local readout of monitoring data versus a centralized "control room" monitoring area (or even both) is dependent on the degree of operator supervision versus intervention required for the processing. This decision is related directly to the flexibility of operators to go in and out of the field in terms of gowning procedures and ease of access. The ability to remotely input set-point or alarm changes can be useful when containment or product protection levels are high. Multiple areas for setting parameter values need to be considered carefully from the aspect of operational control to avoid duplication of desired changes and to minimize batch record log omissions. Verbal communications between the field and the control room using radios or intercoms also need to be considered.

Distributed control systems are favored for process management. System architecture requirements vary depending on the processing requirements but examples for a non-PC based system have been published (73). Distributed control systems also can be PC-based using commercial PC control/data acquisition software (74). In either case, some type of process-level controller is required such as a PLC, loop controller (LC), personal computer–based controller) (PCBC) (32). The ability for those without computer expertise to easily modify and understand control sequences using newly developed commercial control software is attractive but requires effective change control and documentation procedures. The ability to readily configure the data acquisition system to download and archive data is a consideration important for later operation and process development.

There are several types of computer inputs and outputs (72), which include *(1)* digital inputs in the form of an on–off signal used for foam probe, pressure, flow, or temperature switches; *(2)* digital outputs in the form of an on–off signal used for the flow of antifoam, open–close valves, or pump/motor start/stops; *(3)* analog inputs in the form of an amplified output in the range of 4–20 mA used for temperature, pressure, pH, dissolved oxygen, flow, and power draw; and *(4)* analog outputs in the form of a proportional signal that controls the valve opening via an I/P conversion, typically used for back pressure, cooling, and airflow valves or for signals such as motor speed.

The type of control most frequently utilized is the proportional integral differential (PID) control loop. It is tuned using three parameters to minimize oscillations, which can change depending on the specific process (e.g., for dissolved oxygen cascade control of a fermentor cultivating faster-growing *E. coli* versus slower-growing yeast cells) or over the course of the same process (e.g., for glucose and ethanol consumption phases of a baker's yeast fermentation) (75). These three parameters are the size of the error relative to the proportional band with smaller bands resulting in more sensitive systems, integral of the error over time with longer error durations resulting in greater corrections, and rate of error change with time with larger increases in correction with time calculated the faster the error is growing with time (71). Typically the third parameter is at a minimal value near zero for the optimally tuned loop.

A conscious decision needs to be made concerning the desired level of automation for CIP, SIP, and vessel-to-vessel transfers. Higher levels of automation may ensure reproducibility of operation, reduce reliance on operational staff, and simplify SOPs. Raising the level of automation increases capital, validation, and maintenance costs while reducing flexibility. Operating personnel who routinely observe automated processing steps may be less familiar with the details of automated versus manual equipment operational sequences. A domino effect may exist in that automation of one valve may lead to automation of several related valves, severely increasing design complexity and cost. Limit switches often are installed that confirm to the control system that the desired automatic valve has opened or shut as requested. Automatic pulsed diaphragm valves have been successful for nutrient feeding as well as for sterilization of transfer lines in automatic plants (76). One prominent example of a highly automated but apparently reliable fermentation facility is operated for the manufacture of cephamycins as well as for pilot use (77). Typical diagrams for monitoring and control of a fermentor (20,72) and for monitoring and control of an ultrafiltration unit (32) have been published.

Data Acquisition and Archiving

The collection and management of data during pilot plant operations are critical, with the collection interval dependent on the nature of the processing step. For fermentations, intervals might be every 15–20 min during secondary metabolite or animal cell cultivations but might shrink to 5–10 min for faster-growing bacterial or yeast cultures, especially when a key process addition or change is planned to be made on the basis of offgas or dissolved oxygen data. Typically, 5- to 10-min intervals are appropriate for microfiltration or centrifugation operations. Smaller data collection intervals may create substantial data archiving demands that may be alleviated by only saving those data substantially different from previous values, based on some user-defined value. Continuous data collection in the form of chromatograms is typically required for chromatographic operations.

In-process trending of data is attractive to permit timely evaluation of process status and troubleshooting of processing problems. Automatic recording of alarms and changes to the control system, preferably sorted according to the equipment and/or batch on which they occurred, also can help trace back processing problems and actions taken. On-line data might be archived on a common server by some unique batch delimiter, according to the established facility SOP for archiving data. One formidable challenge is to incorporate off-line data in the same or a closely related reference database; this off-line data can get fragmented among the notebooks and files of individual re-

searchers. Novel customized systems for bar-coding samples taken from the field and creating a database for results can be helpful and can also create possibilities for the use of robotic systems for sample analysis. Remote access is an attractive option for monitoring processes from off-site locations. Remote alarming can be triggered to a satellite pager, thus removing the necessity for round-the-clock coverage. Remote control of operations can be administratively challenging to regulate and document as well as a potential safety issue. Security issues must always be considered for any established remote capability.

Warehousing

Warehousing needs for the facility need to be determined based on expected production capacity. Common items that might be stored include ambient, cold, or frozen solid or liquid hazardous and nonhazardous raw materials (controlled temperature and/or relative humidity storage), disposables, equipment, and spare parts as well as sample and product storage. Computerized inventory systems, organized by process or equipment, can be useful for most of these items with reminders to reorder when supplies decline to a critical level. Marker equipment also might be purchased for long-lead replacement parts for which redundancy was not desired during the original facility design. Sizing of these storage areas should consider typical lot volumes of prepared culture media and solids to minimize repetitive release testing for multiple lots. For a pilot plant, in which several diverse processes may be run, additional process-specific storage often is required. The relative amount of storage necessary in the building should be compared with other on-site and off-site options. Vendor manufacturing schedules often require purchasing larger quantities of rare raw materials to avoid future shortages. In this instance, if on-site storage is not available, vendor storage or a third-party warehouse alternative may be required. Warehouse access may need to be restricted depending upon the nature of its contents. There should be segregation for released, unreleased, and quarantined materials either by physical separation or by designated labels.

While the entire facility should have a comprehensive pest monitoring program, as well as an SOP, there may be a need to directly address pest control, particularly for flours and meals used in certain secondary metabolite fermentations. These food ingredients when purchased for pilot-scale fermentation purposes may not pass through fumigation unless specifically requested on the purchase order. Typically, larger customers in the food industry have invested capital for on-site fumigators, but commercial nutrient vendors may not have this type of equipment because of their high rate of inventory turnover.

Staging areas should be available for collecting raw materials and supplies for a specific batch. Some method of inventory assessment and control should be available; one convenient method is to use computerized bar-coding to check inventory as well as to confirm use of specified materials when charging occurs. In addition, controlled sampling areas (to provide raw material protection, environmental, and/or personnel protection) for sampling raw materials may be necessary.

Backup Systems and Redundancy

A critical design decision centers on the level of backup or redundancy required in the pilot plant. While typically none of the material created in the pilot plant is "for sale," mechanical problems can delay processing and clinical material production, which may interrupt clinical supplies and compromise the success of ongoing studies. Even if no clinical material is at risk, process development efforts are hampered when experiments are frequently interrupted for mechanical support-type equipment failures (4). Scenarios that might be protected against include loss of clean steam lubrication to agitator seals, loss of positive air pressure on a sterile tank, and loss of electrical power to a controller. Corresponding equipment needs might include redundant clean air compressors, clean steam generators, and purified or WFI water generation systems, with the peak plant capacity needs split between two systems rather than relying on a single system sized for peak usage. Such redundancy usually is achievable because utility costs are low compared to equipment costs (4). An uninterrupted power source (UPS) system sized to include the most critical equipment (typically controllers and data acquisition devices) might also be installed. Backup power for large electrical loads such as motors might not be economical; however, consideration might be given to installing connections for a gas-powered generator if future evaluations altered the economics.

Future Expansion and Modification

Because future processing needs may change, the pilot plant must adapt to maintain its usefulness to the company. Large initially underutilized spaces might be outfitted with utility headers to minimize processing disruption during future installation. Oversizing of utilities for anticipated future loads (4) as well as unforeseen procedural modifications warranted on the basis of validation testing results such as additional rinsing cycles assures subsequent flexibility. Selection of portable equipment permits exchanges for new processes, provided there are procedures developed for storage and start-up of this equipment.

One aggressive approach to increase versatility is the availability of a modular mobile validated pilot plant (78), composed of a system of interconnected and integrated modules. Modules are selected according to processing needs and can be assembled rapidly to form a processing plant capable of undergoing validation. Production modules include fermentation, recovery, and purification; the utility module includes clean water and clean steam; the personnel support module includes lockers and changing areas; and the HVAC module houses the air handling and filtration equipment. Additional modules can then be added or deleted as required.

OPERATION

Each pilot plant facility must decide for itself whether it is required to operate in a continuous mode of GMP compliance or whether it can selectively apply GMP practices to certain batches [but not to other concurrently run batches]

or during certain specific processing periods. A conservative solution might be to operate continually in GMP compliance, thus insuring facility readiness and preventing omission of a required procedure for an actual batch destined for the clinic. Dual GMP and non-GMP operation serves to create two standards and may cause confusion among operating and maintenance staff. If dual operating modes are considered, then a clear distinction between GMP and non-GMP operation must be established, documented, and maintained.

Maintenance

Preventative. Preventive maintenance (PM), performed during an annual facility shutdown, is probably the most important factor influencing smooth operation throughout the upcoming year. The scheduling of an annual facility shutdown is desirable for several reasons including the ease of lockout and tagging of hazardous energy sources, the time economies of concentrating several maintenance workers in a single area, and the ability to collect and document completion of a large fraction of maintenance work at one time. This yearly PM binder can then serve as the model for maintenance personnel for the following year's shutdown. Individual PM work orders, spread over the course of a year, result in substantial time and effort expended in scheduling, preparing, and documenting the work done. Interspersing of annual PM work among batch processing inevitably results in PM delays due to equipment and/or manpower unavailability. Guidelines for planning and executing an annual plant shutdown are available (79).

All PM work orders should be listed on a single spreadsheet to be reviewed before the facility shutdown and augmented with any modifications or one-time replacement work orders that are requested. As PM work is completed, documentation might be signed by the tradesperson as applicable and reviewed by mechanical and operational staff for completeness. Problems identified with equipment or instrumentation should be highlighted as they occur and then entered into change control work order systems so that they can be tracked for their potential effect on validation.

Before the annual shutdown, transfer line manifolds and utility supply lines, which typically cannot be shut down entirely for maintenance throughout the year, might be leak tested to identify faulty valves in need of replacement. After the annual shutdown, equipment start-up procedures should be documented, preferably through the use of equipment start-up batch sheets. These procedures might include equipment cleaning, testing of instrument calibrations such as level, temperature, and pressure (80), and vessel integrity testing for leaky valves. Instrument calibration accuracy should be evaluated for random or systematic errors, with random error within accepted tolerances being the expected outcome.

Some general PM procedures have been outlined that can apply to several areas in a pilot plant, both product and nonproduct contact. Sanitary fittings, which tend to loosen during repeated heating and cooling cycles, should be checked, with gaskets replaced and clamps tightened as required. Any flanges in product contact should be tightened to a specified torque. PM inspections of vessel internals should be conducted with lock washers on nuts that may loosen due from repeated heating and cooling. A PM should be developed for inspection of vessel cooling coil leaks that can result in inadvertent cooling water contact with product (81). For agitated systems, there should be a regular examination of foot bearings, motor/drive vibration, oil analysis, and shaft runout. Double mechanical seals, the failure of which can lead to condensate buildup or batch contamination, should be tested regularly and replaced preventatively based on a minimum life expectancy (82). An outline of specific PM procedures for laboratory fermentors can be applied to larger-scale fermentors (83). An example of documentation for the maintenance of horizontal laminar flow hoods also is available (84) that can be readily adapted to other HEPA filter units. PMs should be conducted on all utility systems, which may include rebuild of air compressor heads, descaling of heat exchange surfaces, and replacing of air dryer desiccant, among other items recommended by the equipment manufacturer. Finally, novel predictive maintenance evaluation techniques should be explored and implemented as appropriate to identify imminent equipment failures (85).

It is most appropriate to perform any annual validation work (e.g., fermenter SIP, autoclave load pattern, controlled temperature unit) at the conclusion of the annual PM shutdown. In this manner, proper equipment operation can be documented and the effectiveness of PM procedures can be evaluated. Often inadvertent mistakes or omissions during maintenance can be identified before they have an adverse, and possibly widespread, quality impact.

Ongoing. A system of tracking outstanding maintenance by equipment or by process should be designed. Each piece of equipment should have a clearly distinguished tag name. Work orders need to be initiated, evaluated for their effect on validation, marked as complete when mechanical work is done, and finalized when process personnel have checked the job. A computerized database can readily track maintenance and identify repetitive maintenance problems. Such a system permits easy review of outstanding work orders for each piece of equipment by the operation group before it is readied for service. All maintenance work orders need to be tracked within one system; care should be taken not to bypass the operations area when maintenance needs are identified by those directly controlling the mechanical staff.

Interim maintenance work, such as calibrations and subannual PM work, should be documented and filed in a manner equivalent to the yearly PM results. PM schedules should be reviewed based on the frequency and results of interim maintenance work. Unusually common equipment failures can be identified through maintenance database searches so that appropriate longer-term repair efforts or replacement can be justified and undertaken.

Calibrations. Additional concerns are associated with calibrations because important processing decisions may be based on the accuracy of the instrument readings. In

some cases, specific instruments are designated as "critical" for a processing step; in other cases, desired specification ranges for the processing step are targeted based on certain identified instruments. These two methods provide assurance that product quality is not compromised if any of the other instruments fail in the equipment being used.

Although some facilities distinguish between calibration procedures for critical and noncritical instrumentation, it may be prudent for a pilot plant to calibrate all instruments in a reliable and documented manner because their accuracy may be important for useful process development data. Calibrations should be performed using National Institute of Standards and Technology (NIST) traceable instruments and should include full loop calibration to the process monitor as appropriate. A specific calibration and maintenance SOP for the instrument is most desirable, but a general SOP referencing vendor manual might also be acceptable, particularly in the interim period while instrument calibration SOPs are being developed. The as-found and as-left conditions should be documented along with the serial number of the standard used. Establishment of an acceptable tolerance for each instrument ensures consistency. When an instrument is found to be outside the tolerance specified, the impact of this calibration error on previously executed batches may need to be evaluated. Thus, a confirmation of the error might be warranted before correcting the calibration. Records should be evaluated for the amount of instrument drift, which may indicate the need to either recalibrate on a more frequent basis or else replace or upgrade aging equipment. An example calibration sheet is available (86).

Modifications and Change Control. Any validated facility needs to devise procedures for tracking changes on equipment as well as on processes and computers. Incomplete documentation of changes can obviate extensive time and money spent on validation testing. Although the exact contents may change, such a procedure starts during the initial design phase as equipment specifications are altered, continues through the facility start-up and qualification phases, then moves directly into the facility operational phase. Adherence to change control procedures may be difficult during facility start-up as many unexpected items are discovered to be incorrect and management pressure to have the facility operational may mount. Depending on the philosophy of the individual facility, the scope of this change control procedure may include product and nonproduct contact equipment to varying degrees. The procedure identifies the individual or individuals that need to approve changes and determine before or in concert with the change implementation what additional validation testing is required. Often, processing personnel, being most familiar with the equipment being modified and the reason for its modification, can provide valuable leadership in this area and initiate contact with quality and validation personnel as required for input.

Generally replacements in kind are accepted without additional testing unless the item being replaced was specifically tested or calibrated during the initial validation effort. Example items and associated testing might include a replacement or rebuilt motor or associated drive (measure motor amperage draw at various agitator speed settings) or replacement instrumentation (calibrate transmitter). Replacement items in this category need to be tagged appropriately in the field with the same tag number as the original item and serial numbers updated in any centralized maintenance data bank as well as recorded on installation documentation.

Modifications or obvious changes in equipment or instrumentation need to be evaluated for their potential impact on the validated state of the equipment. This impact is particularly important for SIP and CIP cycles. Such modifications may be minor, such as a change in gasket or o-ring material of construction, or major, such as a change in piping. An example procedure is available that can ensure proposed changes are reviewed by operations, maintenance, and quality for their potential impact (87). Specific attention should be given to alterations needed in P&IDs, SOPs, PMs, batch sheets, or training programs. Equipment modifications, necessary to accommodate a new process, need to be identified early and their impact evaluated (2). It is critical to recognize the need to balance the request for change with an evaluation of its validation impact; changes cannot be discouraged in a multipurpose pilot plant but need to be controlled and managed in a logical manner.

Staffing

Personnel needs are best defined based on the production process (88), but in a pilot plant setting production processes change rapidly. A decision needs to be made early in the design process concerning the level of off-shift activity reasonably expected in the facility because this can severely impact items such as utility capacity and operation. Round-the-clock coverage, while demanding for personnel involved, may be necessary a large fraction of the days to ensure success and timeliness of production. The use of unionized personnel, technicians, or process development staff should be incorporated into the educational and manual workload expectations for the operating staff. A core operations group, responsible for the facility, can ensure continuity in areas such as SOP updates, maintenance work order tracking, batch sheet content, and equipment operation. Such a group might be supervised by an area manager responsible for coordination and scheduling who might interface with those involved directly in process development. Processing groups might then rotate through the pilot plant as their projects demanded, accompanied by a member of the operations group. Alternatively, a dedicated pilot production process development staff might be instituted, responsible for translation of laboratory processes into GMP batches. A shift report or log should be established and continually updated as to the mechanical and operational status of individual pieces of equipment and utilities.

Laboratory Support

To provide adequate and accurate sample analyses, laboratory support needs should be identified and addressed through either on-site or contract laboratories. Targets for testing might include final product, in-process samples,

raw materials, sterility, culture purity, environmental quality of controlled environments, and utility quality (Table 6). While the goal of these laboratories might be to work toward using validated methods and approved SOPs for all assays, this goal might not be attainable for project-specific assays until the later stages of a successful development project. Initially, in the early project stages, only notebook procedures might be available to guide laboratory testing. Information directly impacting quality should be communicated directly to operating personnel and appropriate investigations conducted both in the field and in the laboratory for failing results.

A microbiological laboratory is critical for fermentation-based operations to prepare seed cultures and to test new culture seed vials for culture purity and productivity in lab-scale processes. An examination of up-to-date sterility/culture purity results before vessel-to-vessel transfers should be undertaken as close to possible to the time of transfer to minimize adverse impacts of contaminated cultures. In a pilot plant setting, tension may exist between the time required to adequately test new cultures and the development pressure to use the latest mutant or transformant in a pilot-scale cultivation to evaluate its capabilities. Reasonable procedures need to be established in advance for the introduction of new processes or new cultures into the pilot plant. Laboratory-prepared nutrients, trace ingredients, and sterilized additions need to be evaluated for their quality impact, and the procedures used for their preparation should be documented where possible.

Standard Operating Procedures

SOPs are essential for the planning and standardization of various procedures conducted by operational staff and maintenance personnel that need to be completed before, during, and after pilot plant batches (2). Typical categories include SOPs for equipment operation, facility and equipment cleaning, sampling, sterilization/sanitization, maintenance, and documentation procedures for tasks such as product changeover, change control, and batch sheets (Table 7). When referencing SOPs as part of the batch record, working versions of SOPs can be helpful to track progress in the field. An SOP on the requirements for writing, issuing, and inactivating an SOP is also useful. SOPs focus operation, guide operations staff, document the order and scope of activities, standardize approaches, define regulatory needs and constraints, and increase the speed of new personnel training (2). They should be written by a person familiar with the facility or equipment operation that is the subject of the SOP, reviewed by another person who might be asked to follow the SOP, and approved by a quality group. Each SOP should have an effective date of issue, be available to those who need it, and be archived when a future revision is created.

The need for structure and consistency within an SOP must be balanced by the need for flexibility in a pilot plant when unexpected process requirements arise. Procedures must be defined for what to do when an SOP is altered intentionally to accommodate a new process or a surprise development in an existing process. Consideration to SOPs should be given during the equipment design phase so that tasks can be performed in a logical clockwise or counterclockwise fashion (88). SOPs should be clear so that the expected outcome is assured. Specific examples of details often omitted include statements regarding settings for pressure switches and acceptable calibration error tolerances, insertion diagrams, part numbers, and manual ref-

Table 6. Types of Testing (Sampling and Analysis) Possibly Required by On-Site or Off-Site Support Laboratories for a Bioprocessing Pilot Plant

Facility and Utility monitoring	Sterilization cycle validation	Cleaning cycle validation	Process monitoring
Total organic carbon	Indicator spore strip incubation	Total organic carbon (swab and rinse water)	Spectrophotometer
pH	Indicator water spore suspension incubation	pH	pH meter
Conductivity	Indicator media spore suspension incubation	Conductivity	HPLC for organic acids, amino acids, sugars
Endotoxin (LAL)	Sterility	Endotoxin (LAL)	HPLC for product/impurity analysis
Bioburden and coliforms	Culture purity (bacterial and fungal)	Bioburden (swab and rinse water)	LC-mass spectrometry
Airborne viable particles	Filter integrity testing	Ultraviolet light absorbance	Protein gel analysis
Airborne total particles	Microbial identification	Visible light absorbance	Osmometer
Surface, viable particles	Gram stains	Filtered solids	Microscope
Compressed gas, viable particles		Dissolved solids	Coulter counter
Compressed gas, nonviable particles			Conductivity
Compressed gas, identity			Metabolite analyzer
Noncondensable gases in steam condensate			Plate readers
			Blood gas analyzer
			Laboratory-scale control/ troubleshooting experiments

Table 7. Example Types of Procedures Possibly Required in a Bioprocessing Pilot Plant

Administrative	Process equipment	Laboratory equipment	Maintenance	Operations
Processing abnormality	Sterilization-in-place	Autoclave operation: sterilization and decontamination	Instrument calibration	Gowning
Investigation of contamination	Cleaning-in-place	Autoclave load patterns	Equipment preventative maintenance	Water system flushing/ sampling
Product changeover	Sanitization	Glasswasher, operation	Utility preventative maintenance	Steam system sampling
Writing of an SOP	Vessel setup and operation	Glasswasher, load patterns	HVAC preventative maintenance	Spill containment
Personnel training on equipment/operations	Transfer line/panel operations	Working in biosafety cabinet	Equipment change control	Facility housekeeping and cleaning
Personnel training on assays	Batch harvest (noncontained)	Working in laminar flow hood	Equipment start-up after processing hiatus	Facility start-up after processing hiatus
Personnel training on computers	Batch harvest (contained)	Incubator/shaker operation	Equipment start-up after major mechanical work	Daily utility checkout
Issuing and review of batch sheets	Equipment integrity test	Freezer	Equipment storage during processing hiatus	In-process sampling and analysis
Raw material release and storage	Utility systems operation	Refrigerator/cold room	Computer system change control	Data archiving
Investigation of monitoring excursions	HVAC systems operation	Tubing welder use and cleaning	Computer system failure recovery	Computer system operation
Forklift certification	Media/buffer preparation	Manual washing	Shutdown/interim maintenance documentation	Monitoring-controlled environments
Pest control	Specific equipment SOPs	Usage of scales and checkweights	Equipment isolation for maintenance	Tracking of process/ equipment status
Ongoing validation testing		Calibration of pH meter	Issuing and tracking of work orders	
Ongoing facility monitoring		Facility monitoring assays		
Introduction of new products/processes		In-process monitoring assays		
		Fume hood use		
		Seed preparation		

erences to aid in following the procedure, and reliance solely on valve numbers, not valve purposes, to describe the procedure. While SOPs should be reviewed periodically every few years to determine if they need clarification, care should be taken to minimize the number of similar versions of the same SOP issued to the field. Operational clarification memos, issued to all operational staff and signed off when read, might be used to provide additional SOP clarification without reissuing the SOP during the interim period between SOP reviews. All major SOP changes should be documented, evaluated for their effect on validation, and communicated to staff in a documented manner.

Safety

Because concerns regarding product exposure to personnel are similar to concerns for personnel exposure to product, GMP guidelines overlap with safety regulations in many respects such as training, SOPs, preventative maintenance including calibrations, and validation/change control (2). Safety captions can be required for certain raw materials based on review of material safety data sheet (MSDS) documentation. Broths, as well as isolated compounds, might be submitted for toxicity testing as soon as they are cultivated at a larger scale. Safety is a component of installation qualification as pressure vessel documentation and utility connections are reviewed. Hazard operability (HAZOP) reviews for pressure vessels (e.g., fermentors, autoclaves) and utilities (e.g., clean steam generators) often are required by company policy and summaries included within equipment qualification reports. HAZOP action items need to be addressed promptly and modifications to equipment or procedure made to minimize risk. Equipment design needs to be reviewed for the ease of lockout and removal of hazardous energy sources (e.g., bleed valves for the evacuation of process and utility lines, field-mounted pressure gauges to confirm depressurization) as well as the individual isolation of skids/units for maintenance. In addition, specific SOPs for preparing equipment for maintenance might be required.

Regular safety inspections with documented action item lists communicated to all affected personnel (i.e., operating staff, technicians, engineers, management) can identify and correct nonoptimal working conditions and cluttered processing areas. Exposure of personnel to excessive sound, heat, or dust should also be considered during facility design as well as monitored during facility operation. Although personal protective equipment such as ear plugs, heat-resistant gloves, and respirators is available, when

possible engineering controls such as sound dampeners, insulation, and dust collectors may be better solutions. Safety during validation testing should always be considered, particularly when groups outside of operations are involved in setting up or operating equipment and when sources of hazardous energy are involved. Whether during validation testing or operational use, fittings of any sort should be tightly secured and should not be tightened when equipment is under pressure or hot, specifically during SIP and CIP procedures. Summaries of generally accepted laboratory and pilot plant safety procedures to consider are available (89).

According to ASME code, overpressurization protection using relief devices is required for vessels more than 6 in in diameter. The background for this regulation as well as details of pressure relief valves and rupture disks is available (90). Sanitary rupture disks are most desirable for product contact; once blown, processing must stop until they are replaced. Sometimes a rupture disk is installed, followed by a pop safety valve rated at a slightly higher pressure with a pressure gauge between to indicate rupture of the disk (48). If this arrangement is used, then the cleanliness of the area between the rupture disk and pop safety valve must be demonstrated for the batch to not be adversely impacted if the disk ruptures but the pop safety valve does not release. Removal of the entire rupture disk assembly (not the removal of the disk from the assembly) for cleaning should be checked with company policy but may be acceptable in most cases.

Another often-overlooked matter is that the outlets of safety relief devices (which might be unexpectedly spewing steam, hot water, or broth after relief) should be oriented such that personnel safety is considered for those who may be standing near them. One method is to pipe outlets to outside containment troughs, but it may not be desirable to have a potential connection to the outside. A second alternative is to create a designated area within the pilot plant processing area. While outlets to the floor beside the equipment should be minimized, if they are found to be necessary they should be carefully situated outside the normal path of personnel.

Training

Initial and ongoing training programs need to be developed and documented for each individual associated with the facility, based on the facility needs. A training SOP should be developed, which might include a training manual containing useful items such as key SOPs, facility layouts, equipment diagrams, and policies. Initial training might include items from vendor-supplied training sessions to validated load pattern arrangements. Often this training falls at a peak workload period during the facility start-up and validation when attention can be diverted into seemingly more urgent areas. In other cases, more individuals require training than can reasonably be accommodated in the area surrounding the equipment. When permitted, such training might be recorded for later review by the staff when they have fully assumed responsibility for operation. Another alternative is for those just trained to provide training to others, which may be suitable for certain operations. For best results, training should occur with written documentation available to those being trained for reference.

Ongoing training should be done on a regular basis; once every 1 or 2 months seems most appropriate for an active pilot plant facility. Training should include supervisors, operators, engineers, and maintenance personnel as well as validation and GMP staff. Training might include a process review that highlights critical steps and equipment, SOP reviews, identification of personnel safety and environmental containment issues, and review of PM and ongoing maintenance needs, as well as what operational personnel should look for with respect to proper equipment function (88). Training sessions are one method to communicate recent SOP changes as well as to explain why certain procedures are necessary. These might also be used for periodic SOP reviews to obtain feedback from those using the SOP frequently. Attendance should be required and sessions recorded when possible; those new to the facility should review prior training sessions for the past few years as applicable.

Validation

The recent guideline issued by the FDA (91) makes pilot plants eligible for product licensure should company needs require a product launch from a pilot plant facility. When manufacturing clinical materials, pilot plants are evaluated under the guidelines for investigational new drug products, among other regulations which clearly state that GMPs apply for drug products approved for clinical trials in humans and animals (92). Individual company philosophies vary on the suitability or desirability of their pilot plants for this task because using a pilot facility for manufacturing reduces its availability for the production of clinical material. The level of validation effort required is best considered during the initial facility design. Validation can be done retrospectively, although the collection of the required documentation may be more challenging and modifications necessary for testing may be costly.

The major steps for validation ensure that installed equipment and utilities operate in a manner acceptable for the process to be run in that equipment. Because multiple processes, many of which are not necessarily identified at the time of validation, may be conducted in a pilot plant, operational ranges need to be set to accommodate a wide range of reasonably expected processes. For example, an incubator might be validated for temperatures between 25 and 37 °C to accommodate common bacterial, yeast, and fungal cultures, but not necessarily at 45 °C to accommodate thermophiles. The selection of worst-case scenarios can be difficult for a pilot plant if the associated worst-case conditions are unduly burdensome for normal operation. Specifically, the cleaning procedures required after a 3-week fungal production cultivation are significantly more complex than those required after a 2-day fungal seed cultivation of the same culture. Similarly, the sterilization cycle hold time for a concentrated sugar solution is substantially longer than that required for a heat-sensitive waterlike growth medium. Approaches to resolving these multiproduct validation issues can be complicated by desires to follow manufacturing validation guidelines, which most likely were established based on a single product.

Often validation, qualification, verification, commissioning, and start-up are used with specific definitions in mind for individual facilities and organizations. Generally, validation implies the highest rigor of direct reproducible testing of a specific procedure. Qualification encompasses a wider range of documentation collection and testing of capabilities. Verification might imply a single test against an identified standard. Commissioning can refer to documented troubleshooting of newly installed equipment, whereas start-up might refer to initial equipment checkouts directly after installation. The major validation phases are divided to focus on qualification followed by validation. Installation qualification (IQ) focuses of the verification of proper installation, appropriate utility connections, and adherence to manufacturer and purchase order specifications (93). Operational qualification (OQ) includes testing of equipment function and an acceptance criteria for performance (93). Performance qualification (PQ) centers on process-specific testing for SIP, CIP, and autoclave sterilizing and decontamination load patterns. For computers, a system life cycle methodology incorporating specifications, program design, coding, testing, start-up, operation, and maintenance has been adopted (93).

Process validation centers on the reproducibility of the process in the manufacturing area. Parameters influencing this reproducibility might be best studied in the pilot plant to obtain acceptable operating ranges for those processing parameters, such as cultivation temperature, pH, and dissolved oxygen, found to influence product quality. The set of criteria for proceeding to the next processing step, specifically volumetric or specific productivity, yield, and purity, must be validated in a documented fashion. The critical processing steps important to process quality should be identified and validated (93). Other noncritical process steps may be studied for their impact on manufacturing productivity. Process validation may require multiple identical pieces of equipment, such as fermentors, at the laboratory and process scale to define acceptable processing ranges in a parallel rather than sequential manner.

Batch and Facility Records

Documentation associated with a new facility can be extensive initially and can increase substantially during the time that the facility operates. While centralizing all documentation in a single area can be attractive, duplicate copies of some items often are needed in other locations as well as for archival backup purposes. For example, equipment manuals and facility drawings may be needed near the actual location of the equipment, in the maintenance area, and in the centralized documentation area. Validation documentation, consisting of IQ, OQ, PQ, and load pattern and SIP/CIP reports, is generated both during the initial validation as well as during annual retesting and ongoing environmental monitoring programs. Storage of change control documentation might be accomplished through notebooks and file folders for individual pieces of equipment or by individual suites. SOPs might be archived in a centralized location but also posted near equipment for ready reference. Access and sign-out procedures for the control of documentation need to be developed.

A robust system for batch records needs to be created, based on individual batch sheets for processing steps or for individual pieces of equipment or some appropriate combination of the two approaches. Issued batch sheets should be tracked by a unique sequential number with the product, processing step, and/or equipment number clearly identified. There is a need to track the use of equipment by product and batch, particularly for multipurpose facilities. While this might be done through individual equipment use logs, it is also necessary to include such tracking within the batch record itself.

Newly authored or completed batch sheets should be reviewed by the operations area and may be required to be approved by a quality group before execution. Batch sheets should have clear written instructions as well as a signature, date, and time, as required, for each step. Handwritten instructions, likely to be added during the processing of pilot-scale batches, may need to be reviewed by representatives from operation or quality groups. Brief justifications for alterations in batch sheet instructions also should be considered. An written explanation documenting that the material from the current step is acceptable for continued processing in a subsequent step is recommended. Care should be taken to use common templates wherever possible to simplify successive batch sheet differences and minimize the review effort. Executed batch sheets should be reviewed for completeness by the operations area, by those in charge of the batch, and by a quality group as required. A tracking system should be instituted to monitor the review of completed batch sheets as well as to ensure that all issued batch numbers are accounted for in the batch record filing cabinets. Photocopying, microfilming, or scanning of completed batch sheets serves to provide a backup in case of loss or damage. Paperless batch records, as well as electronic signatures, have also been considered to minimize documentation volume, and guidelines for implementation and validation have been published (94).

BIBLIOGRAPHY

1. P. Basu, J. Quaadgras, J. Holleman, R. Mack, and A. Noren, *Chem. Eng. Prog.* **96**, 66–75 (1997).
2. P. Basu, J. Quaadgras, R. Mack, and A. Noren, *Chem. Eng. Prog.* **94**, 67–74 (1998).
3. M. Burnett, V. Santamarina, and D. Omstead, *Ann. N.Y. Acad. Sci.* **646**, 357–66 (1991).
4. P. Barrer, *Biotechnology* **1**, 661–666 (1983).
5. F. Bader, in M. Ladisch and A. Bose eds., *Harnessing Biotechnology for the 21st Century*, American Chemical Society, Washington, D.C., 1992, pp. 228–231.
6. J. Odum, *Pharm. Eng.* **15**, 8, 10–18, 20 (1995)
7. F. Bader, A. Blum, B. Garfinkle, D. MacFarlane, T. Massa, and T. Copmann, *BioPharm* **5**, 32–40 (1992).
8. M. Hamers, *Biotechnology* **11**, 561–570 (1993).
9. S. Fitzpatrick, A. Ma'ayan, and J. Wagget, *Chem. Eng. Prog.* **86**, 26–31 (1990).
10. M. Narodoslawsky, *Chem. Biochem. Eng. Q.* **5**, 183–187 (1991).
11. N. Samsatli and N. Shah, *Trans. Inst. Chem. Eng. Part C* **74**, 221–231 (1996).

12. N. Samsatli and N. Shah, *Trans. Inst. Chem. Eng. Part C* **74**, 232–242 (1996).
13. D. Schell, *Appl. Biochem. Biotechnol.* **51/52**, 549–557 (1995).
14. M. Cannales, A. Enriquez, E. Ramos, D. Cabrera, H. Dandie, A. Soto, V. Falcon, M. Rodriguez, and J. de la Fuente, *Vaccine* **15**, 414–422 (1997).
15. E. Eliezer, *BioPharm* **6**, 24–29 (1993).
16. R. Sato and M. da Costa, *Biotechnol. Lett.* **18**, 275–280 (1996).
17. D. Petrides, E. Sapidou, and J. Calandranis, *Biotechnol. Bioeng.* **48**, 529–541 (1995).
18. N. Titchener-Hooker, D. Gritsis, K. Mannweiler, R. Olbrich, S. Gardiner, N. Fish, and M. Hoare, *BioPharm* **4**, 34–38 (1991).
19. J. Zawistowski and J. Rago, *Pharm. Eng.* **14**, 24, 26, 28–30, 32 (1994).
20. A. Shahidi, R. Torregrossa, and Y. Zelmanovich, *Pharm. Eng.* **15**, 72–83 (1995).
21. O. Doblhoff-Dier, S. Huss, R. Litos, R. Plail, F. Unterluggauer, M. Reiter, and H. Katinger, *Process Biochem.* **26**, 201–207 (1991).
22. K. Nelson, in A. Prokop, R. Bajpai, and C. Ho eds., *Recombinant DNA Technology and Applications*, McGraw-Hill, New York, 1991, pp. 509–565.
23. K. Leach, *Pharm. Eng.* **13**, 54–62 (1993).
24. J. Odum, *BioPharm* **5**, 36–38 (1992).
25. L. Scutti and S. Stark, *Pharm. Eng.* **13**, 47–52 (1993).
26. B.K. Hamilton, J.J. Schruben, and J.P. Montgomery, in A. Demain and N. Solomon eds., *Manual of Industrial Microbiology and Biotechnology*, American Society for Microbiology, Washington, D.C., 1986, pp. 321–344.
27. P. James, *Pharm. Eng.* **18**, 72–82 (1998).
28. R. Wisniewski and C. Burman, in A. Lubiniecki and S. Vargo eds., *Regulatory Practice in Biopharmaceutical Production*, Wiley-Liss, New York, 1994, pp. 407–445.
29. M. Charles and J. Wilson, in B. Lydersen, N. D'Elia, and K. Nelson eds., *Bioprocess Engineering: Systems, Equipment and Facilities*, Wiley, New York, 1994, pp. 3–67.
30. Y. Chisti, *Chem. Eng. Prog.* **88**, 55–58 (1992).
31. J. Alford, B. Allen, S. Bisch, L. Brill, D. Clapp, and L. Danaher, *Ann. N.Y. Acad. Sci.* **721**, 326–336 (1994).
32. T. Ransohoff, M. Murphy, and H. Levine, *BioPharm* **3**, 20–26 (1990).
33. A. Barry and R. Chojnacki, *BioPharm* **7**, 43–47 (1994).
34. S. Fulton, A. Shahidi, N. Gordon, and N. Afeyan, *Biotechnology* **10**, 635–639 (1992).
35. G. Aronsson and P. Zadorecki, *Aust. J. Biotechnol.* **1**, 17–20 (1987).
36. J. Mahar, *BioPharm* **6**, 42–51 (1993).
37. J. Odum, *Pharm. Eng.* **12**, 8–12 (1992).
38. K. Nelson, *BioPharm* **1**, 34 (1988).
39. J. Williams, *Biotechnology* **7**, 75–76 (1989).
40. B. Junker, M. Stanik, J. Adamca, K. LaRiviere, M. Abbatiello, and P. Salmon, *Bioprocess Eng.* **17**, 277–286 (1997).
41. G. Gray, *Pharm. Eng.*, **17**, 28–33 (1997).
42. M. Burkhart, J. Wermelinger, W. Setz, and D. Muller, *PDA J. Pharm. Sci. Technol.* **50**, 246–251 (1996).
43. M. Del Valle, *Pharm. Eng.* **15**, 14–22 (1995).
44. A. Petrossian, N. Smart, and R. Projetto, *BioPharm* **6**, 40–45 (1993).
45. C.H. Brooks and P.D. Russell, *Process Biochem.* **21**, 77–80 (1986).
46. J. Akers, *BioPharm* **7**, 43–47 (1994).
47. J. Akers, *Pharm. Technol.* **19**, 28 (1995).
48. S. Miller and D. Bergmann, *J. Ind. Microbiol.* **11**, 223–234 (1993).
49. J. Van Houten and D.O. Fleming, *J. Ind. Microbiol.* **11**, 209–215 (1993).
50. D. Janssen, P. Lovejoy, M. Simpson, and L. Kennedy, in P. Yu ed., *Fermentation Technologies: Industrial Applications*, Elsevier, New York, 1990, pp. 388–393.
51. J. Kossik *Gen. Eng. News* **18**, 21 (1998).
52. L. Kennedy, M. Boland, D. Janssen, and M. Frude, in P. Yu ed., *Fermentation Technologies: Industrial Applications*, Elsevier, New York, 1990, pp. 383–387.
53. R. Maigetter, F. Bailey, and B. Miller, *BioPharm* **3**, 22–29 (1990).
54. J. Odum, *BioPharm* **6**, 42 (1993).
55. J. Hodgson, *Biotechnology* **12**, 983–987 (1994).
56. K. Gary, *American Biotechnology Laboratory* **7**, 26–33 (1989).
57. Y. Lee and G. Tsao, in T. Ghose, A. Fiechter, and N. Blakebrough eds., *Advances in Biochemical Engineering*, vol. 13, Springer-Verlag, New York, 1979, pp. 35–86.
58. M. Johnson, J. Borkowski, and C. Engblom, *Biotechnol. Bioeng.* **6**, 457–468 (1964).
59. R. Elsworth, *Chem. Eng.* 63–71 (1972).
60. P. Salmon and B. Buckland, in M. Rhodes and P. Stanbury eds., *Appl. Microb. Physiol.* , 131–163 (1997).
61. M. Uttamlal and D. Walt, *Biotechnology* **13**, 597–601 (1995).
62. J. Bradley, A. Kidd, P. Anderson, A. Dear, R. Ashby, and A. Turner, *Analyst* **114**, 375–379 (1989).
63. G. Macaloney, J. Hall, M. Rollins, I. Draper, B. Thompson, and B. McNeil, *Biotechnol. Tech.* **8**, 281–286 (1994).
64. G. Macaloney, J. Hall, M. Rollins, I. Draper, K. Anderson, J. Preston, B. Thompson, and B. McNeil, *Bioprocess Eng.* **17**, 157–167 (1997).
65. V. Agayn and D. Walt, *Biotechnology* **11**, 726–729 (1993).
66. A. Humphrey, K. Brown, J. Horvath, and H. Semerjian, in A. Fiechter, H. Okada, and R. Tanner eds., *Bioproducts and Bioprocesses*, Springer-Verlag, Berlin, 1989, pp. 309–320.
67. D. Kell, G. Markx, C. Davey, and R. Todd, *Trends Anal. Chem.* **9**, 190–194 (1990).
68. B. Junker, J. Reddy, K. Gbewonyo, and R. Greasham, *Bioprocess Eng.* **10**, 195–207 (1994).
69. L. Olsson and J. Nielsen, *Trends Biotechnol.* **15**, 517–522 (1997).
70. W. Groot, R. Mijnhart, R. van der Lans, and K. Luyben, *Biotechnol. Tech.* **5**, 371–376 (1991).
71. S. Collins, *Appl. Microb. Physiol.*, 75–101 (1997).
72. J. Petersen, *Pharm. Technol.* **12**, 42 (1988).
73. O. Doblhoff-Dier, S. Huss, R. Litos, R. Plail, F. Unterluggauer, M. Reiter, and H. Katinger, *Process Biochem.* **26**, 201–207 (1991).
74. H. Blachere, *Gen. Eng. News* **18**, 23 (1998).
75. K. Jones, D. Williams, D. Phipps, and P. Montgomery, *IFAC Symp. Ser. (Adv. Control Chem. Processes)* **8**, 35–40 (1992).
76. F. Wilde, *Gen. Eng. News* **18**, 24 (1998).
77. H. Eiki, I. Kishi, T. Gomi, and M. Ogawa, in M. Ladisch and A. Bose eds., *Harnessing Biotechnology for the 21st Century*, American Chemical Society, Washington, D.C., 1992, pp. 223–227.
78. E. Bardone, B. Lani, and G. Cassani, *Pharm. Eng.* **14**, 34–38 (1994).
79. G. Williams, *Plant Services* **19**, 115–118 (1998).

80. B. Junker, J. Lynch, J. Leporati, J. Schmitt, J. Gieger, T. Garah, M. Stober, and P. Salmon, *Bioprocess Eng.* **17**, 279–287 (1995).
81. C. Perkowski, *Biotechnol. Bioeng.* **26**, 857–859 (1984).
82. D. Todhunter, *Bioprocess Eng. Symp.*, 97–103 (1989).
83. F. Kleppinger, *Int. Biotechnol. Lab.* **5**, 28 (1987).
84. F. Kleppinger, *Int. J. Pharm. Compd.* **1**, 344–345 (1997).
85. K. Mobley, *Plant Services* **19**, 147 (1998).
86. D. Wade, *Maint. Technol.* 44–46 (1989).
87. C. Kennedy, *BioPharm* **8**, 34–37 (1995).
88. C. Perkowski, *BioPharm*, 62–65 (1987).
89. V. Singh, *Chem. Age India* **39**, 387–388 (1988).
90. J. Kossik, *Gen. Eng. News* **18**, 25 (1998).
91. Center for Biologics Evaluation and Research, *Fed. Reg.* **60(132)**, 35750–35753 (1995).
92. Center for Drug Evaluation and Research, Food and Drug Administration, Washington, D.C., March 1991.
93. R.G. Werner, H. Langlouis-Gau, F. Walz, H. Allgaier, and H. Hoffman, *Arzneim.-Forsch.* **38**, 855–862 (1988).
94. 21 CFR Part 11, *Fed. Reg.* **62(54)**, 13429–13466 (1997).

See also CHROMATOGRAPHY, COMPUTER-AIDED DESIGN; CLEANING, CLEANING VALIDATION; METABOLITES, PRIMARY AND SECONDARY; SECONDARY METABOLITE PRODUCTION, ACTINOMYCETES, OTHER THAN STREPTOMYCES.

PLASMID DNA REPLICATION

MARCELO E. TOLMASKY
California State University, Fullerton
Fullerton, California

LUIS A. ACTIS
Miami University
Oxford, Ohio

JORGE H. CROSA
Oregon Health Sciences University
Portland, Oregon

KEY WORDS

ColE1
Covalently closed circular DNA
Extrachromosomal elements
Incompatibility
Linear plasmids
pT181
R6K

OUTLINE

INTRODUCTION

Plasmids are extrachromosomal replicons found in Gram-negative and Gram-positive bacteria as well as in some lower eukaryote organisms. They are present in bacterial cells replicating at a specific number of copies per cell. Their size varies from a few to several hundred kilobase pairs, and bacterial cells can harbor more than one plasmid species (1).

The term *plasmid* originally was used by Lederberg to describe all extrachromosomal hereditary determinants. Currently, the term is restricted to the autonomously replicating extrachromosomal DNA of bacteria. Although plasmids replicate autonomously, they generally rely on some host-encoded factors for their replication. Although not essential for the survival of bacteria, plasmids may encode a wide variety of genetic determinants that permit their bacterial hosts to survive better in an adverse environment or to compete better with other microorganisms occupying the same ecological niche. Plasmids are found in a wide variety of microorganisms, and it is as difficult to generalize about plasmids as it is to generalize about the microorganisms that harbor them. Plasmid DNA is mostly isolated as covalently closed, circular double-stranded DNA molecules, although recently linear plasmids have been described in certain bacterial species (2). Furthermore, single-stranded DNA plasmids were identified as intermediates during the process of replication in certain Gram-positive bacteria (3). Plasmids also include the replicative forms of filamentous coliphages and the prophage state of phages such as P1 (4).

The medical importance of plasmids that encode for antibiotic resistance (R plasmids) and those that contribute directly to microbial pathogenicity, such as for instance iron transport in several pathogens or the presence of adhesions, invasins, or antiphagocytic proteins, is well documented, as is the role played by plasmids in bacteria of importance in agriculture and industry (5–8). Some plasmid genes encoding antibiotic resistances and other traits are frequently located in transposable elements or integrons (9–11). This produces great variation and flexibility in the constitution of plasmids. These extrachromosomal elements are of equal importance, however, for the study of the structure and function of DNA. Plasmids have taken on paramount importance in recombinant DNA technology (12). In this section we describe the mechanisms of the control of plasmid DNA replication and examine the replication of plasmids ColE1 and R6K, found in Gram-negative organisms, and pT181, found in Gram-positive bacteria. In addition, we also discuss linear plasmids. For the replication mechanisms of other plasmids we refer the reader to other reviews (13–23).

ColE1-TYPE PLASMIDS

The ColE1-type family encompasses both a large number of naturally occurring plasmids and many of the most commonly used cloning vehicles (16,18,24–30). The mechanism controlling initiation of replication of these plasmids is among the best understood, perhaps reflecting the importance of this family of plasmids in recombinant DNA applications (11,16–18,31–36).

Replication of plasmids containing the ColE1 origin is initiated at a unique startsite *(ori)* (37,38) (Fig. 1). Unlike in other plasmid families, initiation of replication of ColE1 is not mediated by a plasmid-encoded protein; instead ColE1 requires the host DNA polymerase I (Pol I) (39). In addition, initiation of replication of ColE1 requires the presence of a host-encoded RNA polymerase and ribonuclease H (RNase H) (39–41). Replication of the plasmid is initiated by the synthesis of an RNA molecule by the host RNA polymerase 555 bases upstream of *ori* (Fig. 1, panel a). This RNA species, called RNA II, extends about 700 nucleotides from its initiation (Fig. 1, panels b and c) (40,42). The 3′ of this RNA molecule subsequently forms a duplex with the DNA of the plasmid at a region close to the origin of replication, which is located at nucleotide 555 of RNA II (Fig. 1, panel c) (40). The formation of the duplex structure (known as coupling) begins following the synthesis of a portion of the RNA II molecule. The coupling event is strongly dependent on the formation of a specific secondary structure at the 5′ end of the RNA II molecule that results in an interaction between the RNA portion on the *ori* and an upstream region of this molecule with the template DNA (Fig. 1, panel c) (43–46). The appropriate conformation of RNA II necessary to form the hybrid with DNA is shown in Figure 2 (46,47). The duplex RNA II–DNA structure acts as a substrate for the subsequent action of RNase H, which cuts the RNA at the replication origin, leaving a 3′-hydroxyl group that serves as the primer for DNA synthesis catalyzed by Pol I (Fig. 1, panels c and d) (40). It has been shown that certain point mutations in RNA II prevent RNA II from adopting the right conformation, which affects hybridization with the template DNA (44,45,47). Once Pol I begins the addition of deoxynucleotides, the remaining RNA hybridized to the DNA strand is digested by RNase H at other sites on the molecule and by the 5′-3′ exonuclease activity of Pol I. RNase H also functions to remove the RNA primer from the growing DNA strand (Fig. 1, panel d) (48). Replication of ColE1 then proceeds unidirectionally with the initiation of the lagging strand synthesis at specific ColE1 sites.

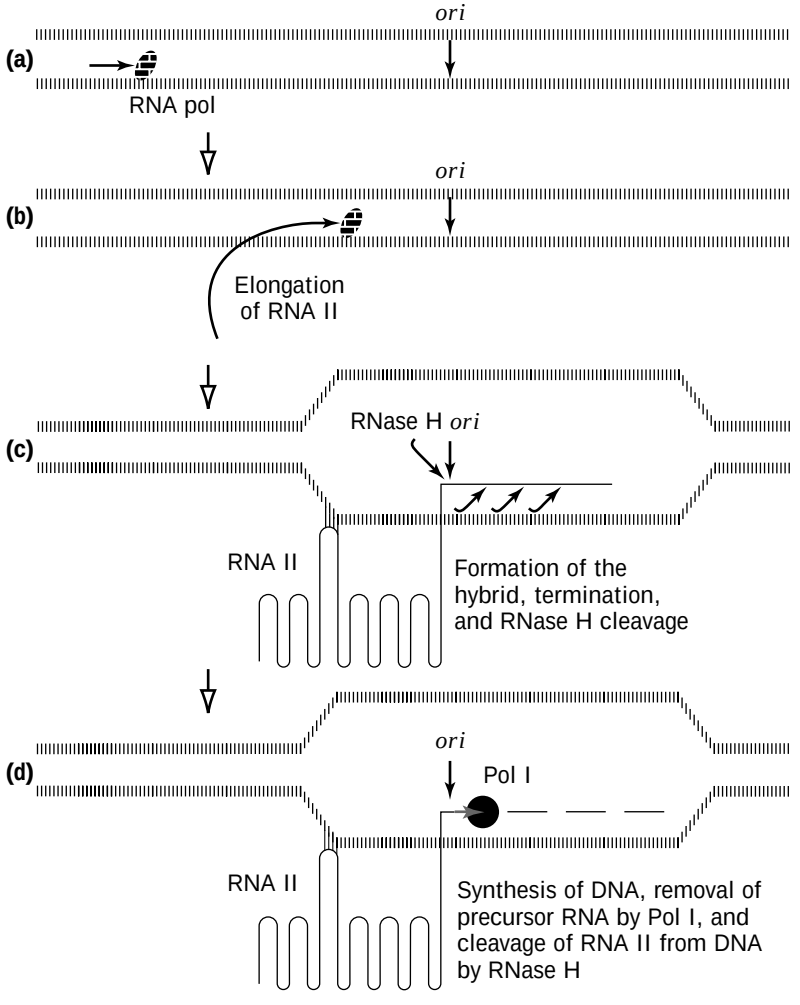


Figure 1. Mechanism of initiation of ColE1 replication in wild-type *E. coli*. The horizontal black arrow in (**a**) indicates the transcription start site (555 nucleotides upstream of *ori*) of RNA II. The origin of replication is indicated by a vertical arrow. ↖, indicates sites of action of RNase H; ||||||, DNA; ■, RNA; ⇶, RNA polymerase I. See the text for details.

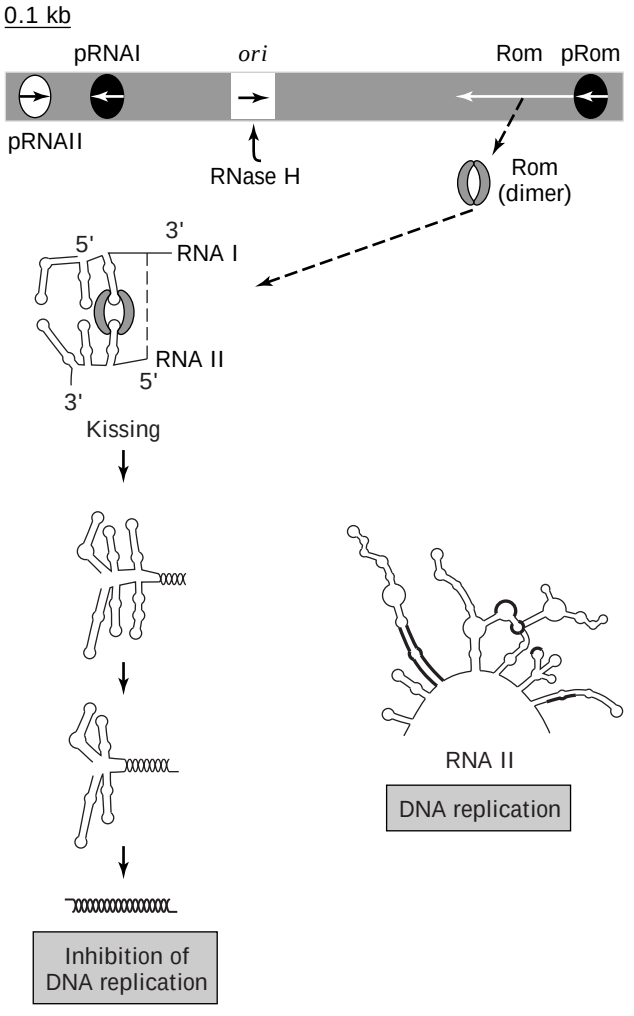


Figure 2. Proposed mechanism of the succession of events involved in the regulation of the initiation of replication of ColE1 by RNA I and Rom. The upper portion shows a genetic map of the origin of replication of ColE1. |||||| represents ColE1 DNA, with the promoters indicated by arrows inside circles. The open arrow indicates the *rom* gene. In the left panel, a schematic of the mechanism of RNA I and RNA II interaction that leads to the inhibition of DNA synthesis is shown. The first interaction between RNA I and RNA II (kissing) is reversible and stabilized by the Rom protein. The dimeric form of Rom is shown both below its gene and in the interaction with RNA I and RNA II. The right panel of the figure indicates the secondary structure that leads to DNA replication. For this to occur, an RNA II molecule that is not coupled with RNA I binds to DNA and acts as a primer after RNase H digestion. The structures reported by Fitzwater et al. as well as Masukata and Tomizawa are almost identical (46,47). A more detailed scheme of the possible interactions between the RNA and Rom species has been reported elsewhere (18). *Source*: Adapted from Ref. 18.

It has been further shown that replication of plasmids containing the ColE1 origin of replication can proceed using two alternative modes of replication that allow for replication in the absence of either RNase H alone or RNase H and Pol I (49–51) (Fig. 3). These two mechanisms appear to not be utilized in wild-type cells, however, and probably represent adaptations to specific host mutations (49,51–53). One of these mechanisms is independent of the presence of Pol I and involves the elongation of RNA II coupled to the template DNA. The lack of RNase H allows the formation of a single-stranded DNA region that can extend to a length that is adequate for assembly of a replisome and initiation of synthesis on the opposite DNA strand (lagging strand) (Fig. 3, panel a) (18,50).

Another initiation mechanism has been demonstrated in RNase H–deficient mutants that is dependent of Pol I. In this case, the extended RNA II species can be recognized by Pol I and used as a primer (Fig. 3, panel b) (49). However, this mechanism of initiation has proved to be inefficient in in vitro experiments (40). Regardless of the replication mode, however, RNA II must be produced in the correct structural configuration to act as the primer for DNA replication (49,51,54).

Control of initiation of replication of ColE1 is mediated by the interaction of RNA II with a 108-nucleotide antisense RNA transcript, called RNA I, that is encoded in a region that overlaps the coding region for RNA II (Fig. 2) (40,55). RNA I binds to nascent RNA II by complementary base pairing, triggering a conformational change in RNA II that prevents coupling with DNA, thereby inhibiting replication initiation (42,44,55–57). Within the secondary structures of both RNA species there are several stem-loops (Fig. 2). Interaction between both RNA species starts at complementary loops in a reversible process known as "kissing," which results in the generation of an unstable initial complex (Fig. 2) (58–62). After this initial contact takes place, an irreversible process leads to the formation of an RNA duplex between RNA I and RNA II (Fig. 2) (56,60,62,63). RNA I achieves its inhibitory effect only if it is present during a short, specific interval in the synthesis of RNA II (56,64). This situation occurs most frequently because RNA I transcription occurs as much as five times more often than transcription of RNA II. This results in the formation of a productive primer only in about 1 out of every 20 RNA II molecules that are started (65). The available amount of RNA I in the cell is influenced by polyadenylation (66,67) and the bacterial RNase E, which cleaves the 5′ end of RNA I, rendering it inactive (68).

A second factor involved in the control of initiation of replication of ColE1 is the plasmid-encoded 63–amino acid protein Rom (RNA one modulator), also known as Rop (Repressor of primer) (Fig. 2) (56,60,62,69,70). The gene encoding Rom is located downstream of *ori* (Fig. 2). In the early stages of plasmid replication, Rom binds to the stem-loop portions of the unstable initial complex (58,59), reducing the dissociation constant of the complex, thus creating a pathway for stable binding of RNA I and RNA II (Fig. 2) (60,71). As a consequence, ColE1 deletion derivatives that lack *rom* present a higher copy number than plasmids that have the complete ColE1 replication region (62,72). The presence of Rom, however, is not needed for viability of the plasmid, and in some RNA-regulated replicons is not present (16).

Incompatibility is the failure of two plasmids to be stably inherited in the same cell line. This phenomenon is a consequence of sharing elements of plasmid inheritance functions such as replication or partition (1,17,73). RNA I

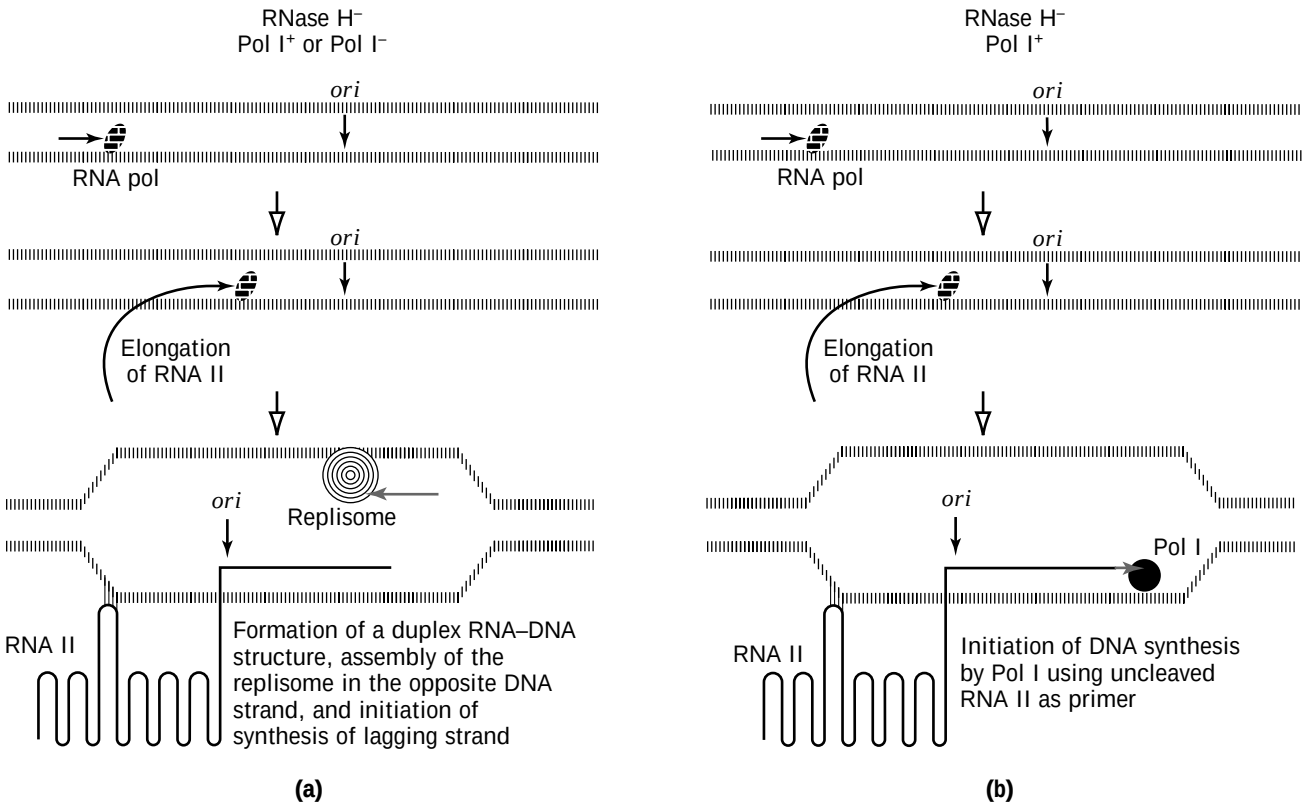


Figure 3. Alternative mechanisms of initiation of replication of ColE1 in RNase H–deficient mutants. (**a**) Proposed mechanism of initiation in RNase H–deficient *(rnh)* and Pol I positive (*polA*$^+$) or negative (*polA*) mutants (52). (**B**) Mechanism proposed to occur in *rnh, polA*$^+$ mutants. The black circle represents Pol I. The horizontal black arrow at the top indicates the transcription start site; concentric circles represent the replisome; and *ori* shows the origin of replication in wild-type cells. ||||||, DNA; ■, RNA. *Source*: Adapted from Ref. 18.

is the main incompatibility determinant in ColE1-type plasmids. Two plasmids that depend on the same RNA I species for regulation of initiation of replication can not coexist in the same cell (1,17,73). It has been shown that even single nucleotide changes can have profound effects in the incompatibility properties of ColE1-type plasmids (55,57).

THE R6K PLASMID

R6K is a naturally occurring conjugative plasmid that encodes resistance to the antibiotics ampicillin and streptomycin (74). This plasmid is an example of an extrachromosomal element of intermediate size and copy number because it is about 38 kb in size and has a copy number of 13 to 40 per cell (75). These features, together with a unique mode of replication, make R6K an attractive system to study the genetic and molecular mechanisms involved in plasmid DNA replication. In addition, this plasmid and its replication components were among the first used to generate gene fusions, transcription enhancement, protein tagging, and site-specific proteolysis (76–79); it is currently utilized for high expression and rapid isolation by affinity chromatography of a large variety of proteins. Furthermore, R6K derivatives were instrumental in designing a series of suicide vectors successfully used to generate mutants by allelic exchange or transposition mutagenesis in Gram-positive bacteria (80,81).

Plasmid R6K is a member of a growing group of replicons encoding an initiator protein that binds to nucleotide sequence repeats, called iterons, located within *ori*. This group of iteron replicons includes the *Escherichia coli oriC* (82) and the plasmids F (83), pSC101 (84), P1 (85), PMJ101 (86), Rts1 (87), the REPI replicon of pColV-K30 (88), and the RK2- and RP4-related plasmids (89). R6K has three origins of replication, named *ori*α, *ori*β, and *ori*γ (90,91), all clustered within a 4-kb DNA fragment (Fig. 4). In addition, the *pir* and *bis* genes encoding the π and Bis replication proteins, respectively, were found within this replication region. Near this region is located the replication termination site containing the τ sequence that binds the host-encoded ter protein and arrests the movement of the replication fork (90,92,93). The main active origins in vivo are α and β, while *ori*γ tends to remain inactive (90,91) due to the synthesis of a silencer RNA encoded immediately downstream from the *ori*γ sequence (94). The silencing activity of this RNA was explained by the formation of a silencer RNA-activator RNA hybrid, driven to completion of the R6K-encoded replication protein, that leads to the in-

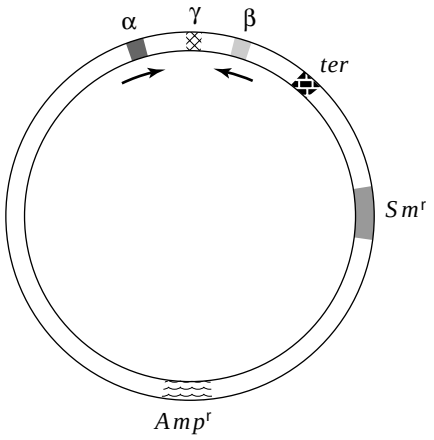


Figure 4. Diagram of R6K showing the 4-kb replication region containing the α-, β-, and γ-origins. The location of the termination site (*ter*) and the genes encoding resistance to streptomycin *(Sm)* and ampicillin *(Amp)* are also indicated. The arrows mark the in vivo direction of the initial replication from the α- and β-origins.

activation of *ori*γ (95,96). However, *ori*γ can replicate autonomously when the other two origins are deleted and the *pir* gene is provided in *cis* or in *trans* (97–101). Thus, *ori*γ has served as the simplest model system derived from R6K to study the replication of an iteron-containing DNA molecule.

The γ origin is the major initiator binding site and is required in cis for initiation of plasmid replication from the other two sites (102–104). Therefore, the γ origin behaves as a prokaryotic enhancer-type element in the sense that DNA–protein interactions at this site induce significant changes in DNA structure that facilitate initiation of DNA replication from distantly located origins (105). Figure 5a shows a diagram of the complete *ori*γ replication region. The molecular organization of this origin is similar to other plasmid origins, although it contains two functionally distinct segments (104,106). The 277-bp core segment, common to all three origins, is essential for replication, and consists of three distinct regions: (*1*) the AT-rich region, bound by π and the integration host factor (IHF) protein, (*2*) seven 22-bp repeats bound by π, and (*3*) a multiprotein-binding region interacting with DnaA, IHF, and RNA polymerase. Immediately to the left of the core lies the 106-bp origin enhancer, which includes a DnaA box and a small segment containing the *stb* locus (107). Even though the enhancer is not absolutely required for replication at low level of wild-type π-protein, it was found that the enhancer region is required for stable maintenance of γ-origin-containing plasmids. Although the *stb* locus has some similarities with the *par* locus of pSC101, the partition systems of these two plasmids differ from each other in several aspects (107). It has been proposed recently by Wu et al. (107) that a host-encoded protein may be binding to the *stb* repeats. The finding that R6K derivatives carrying the intact three origins can bind in vitro to both inner and outer membrane fractions of *E. coli* (108) appears to support this hypothesis. An alternative explanation is that *stb* mediates plasmid partition by altering the structure of γ-origin-containing plasmids (107). The host-encoded protein Fis also binds to 10 sites in the γ-origin that overlap all the previously identified binding sites for the R6K-encoded π-protein and the host-encoded DNA binding proteins DnaA, IHF, and RNA polymerase (109). However, it appears that the Fis protein is required for plasmid replication only when π-copy-up variants and the penicillin-resistant marker are simultaneously used, suggesting that plasmid sequences such as those encoding antibiotic resistance play a role in plasmid DNA replication.

The analysis of the functional α-replicon revealed that it is composed of two elements that must be present in *cis* and oriented as in the intact R6K: (*1*) a 580-bp fragment containing the α-origin, and (*2*) the 277-bp core segment located within the γ-origin (110) (Fig. 5b). These two elements are separated by approximately 3 kbp, which are not required for the α-origin activity. The 580-bp fragment contains a long 98-bp palindrome, and it was suggested that this sequence serves as the recognition signal for initiation of DNA replication in the α-origin. DNA homology analysis also revealed the presence of a 23-bp sequence that resembles the seven 22-bp iterons found in the core segment (110) and plays a role in the π-protein-mediated looping between the γ- and α-sequences (105). In addition, this fragment contains a DnaG binding site that can serve as DnaB loading site by DnaB–DnaG interaction (111).

The minimal β-replicon was defined as a 2-kbp fragment that encompasses the following elements: (*1*) the γ-core region, (*2*) the *pir* gene encoding the π-initiator protein, (*3*) the *bis* gene encoding the 17.2-kDa Bis protein, and (*4*) the β-origin (102,112–114) (Fig. 5c). The Bis protein is required only by the β-origin, and its synthesis is coupled in cis to the expression of π-protein from an unaltered *pir* gene (113). The β-origin also contains a half 22-bp iteron and a 98-bp palindrome that has 96% homology with the α-origin hairpin (110). The presence of these pallindromic sequences at the *ori*α and *ori*β regions was previously assessed by electron microscopy. The half iteron and the long palindromic sequence are required for the π-mediated looping between the γ- and β-sequences (105) and the initiation of DNA replication at the *ori*β, respectively.

The π-initiator protein is essential for replication from each R6K origin and, although it can be provided in cis or trans to regulate the activity of the γ- and α-origins, it is required in cis for activation of the β-origin (97,102, 110,114,115). This protein is a homodimer with a 36-kDa molecular weight for the monomeric form, which is lysine rich and weakly basic (98). In its native dimeric form (116), or as a hybrid with β-galactosidase (78,117) or collagen (77), it has been shown to bind to the seven iterons in the γ-origin as well as to an eighth iteron and a smaller inverted pair of repeats located in the operator-promoter region of the *pir* gene (78,116–118). In addition, this initiator protein interacts with the iterons located in the α- and β-iterons; however, these contacts are weak and require the enhancing effect of π already bound to the seven γ-iterons (105,119).

The π-protein has both positive and negative roles in the replication of R6K (120). The positive role of π is displayed by its ability to enhance replication of *ori*α and *ori*β. This activity is mediated by the promotion of conforma-

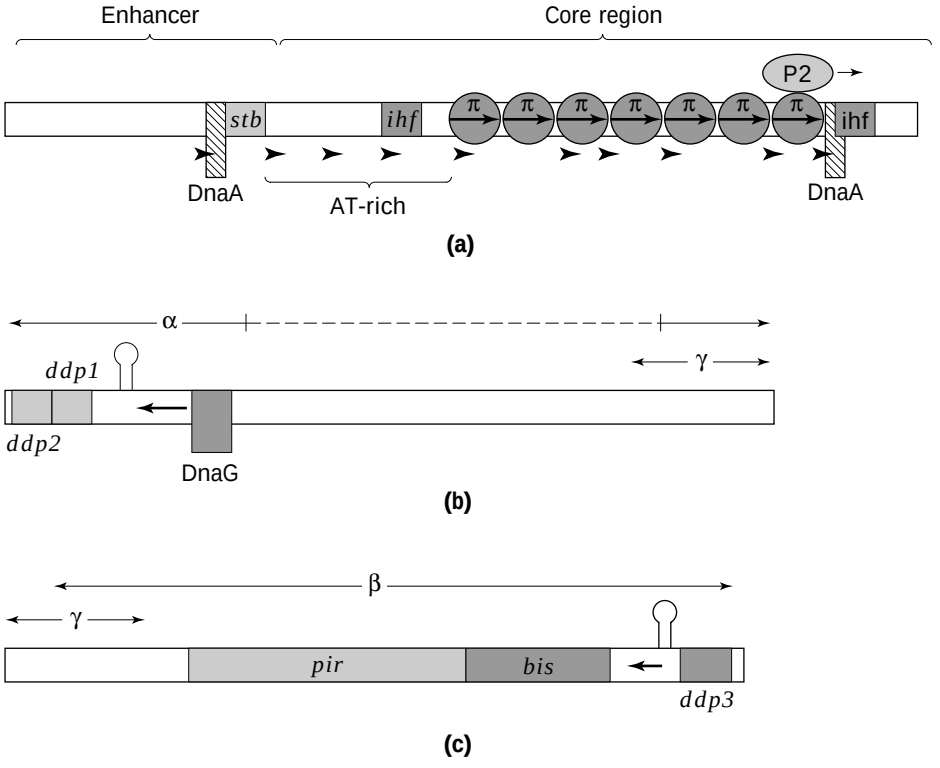


Figure 5. Diagram of the three R6K origins of DNA replication. (**A**) Components of the enhancer and core regions of the γ-origin. The long rectangle represents the DNA region containing binding sites for the host-encoded proteins DnaA (tall rectangles), IHF (short rectangles), and RNA polymerase (oval). P2 and the open arrow represent the promoter site and the direction of transcription of the *pir* gene, respectively. The thick arrows represent the seven iterons that bind the π-replication protein (circles). *stb* represents the locus involved in plasmid maintenance. The 10 arrowheads represent the location of the Fis binding sites. (**B**) Components of the active α-origin. The double-headed arrows indicate the DNA fragment containing the α-origin and the γ-core region required in cis for active replication. The dashed line represents the intervening sequences not involved in DNA replication. The hairpin marks the location of the long inverted repeat located in the α-origin, followed by the genes encoding the DDP1 and DDP2 proteins (small rectangles). The thick arrow and the large rectangle indicate the position of an iteron and a DnaG binding site, respectively. (**C**) Components of the active β-origin. The double-headed arrows indicate the DNA fragment containing the γ-core region and β-origin required in cis for active replication. The β-origin includes the *pir* and *bis* genes (long rectangles), a half iteron (half thick arrow), the *ori*β long inverted repeat (hairpin), and the gene encoding the DDP3 protein (short rectangle).

tional changes, such as DNA unwinding and bending, in the replication region that cause activation of these two origins by looping out intervening sequences located between the *ori*γ core region and the α- and β-origins (119,121). Site-directed mutagenesis showed that π-protein is necessary but not sufficient for activation of *ori*β and probably *ori*α (105). This observation suggested that the DNA looping process must also be required for the transfer of a multiprotein complex capable of initiating DNA replication. Supporting this hypothesis, it was recently reported that π specifically interacts with the host-encoded helicase DnaB replication protein (111). This result indicates that DnaB is initially recruited by π, already bound to the γ-origin, and then delivered to the two other origins by the DNA looping induced by the R6K-encoded replicator protein.

The π-protein has been shown to possess two negative activities on the replication of R6K. One of them is as an autorepressor at the level of transcription and involves its binding to iteron sequences located within the operator-promoter site of the *pir* gene. These interactions act by either preventing the binding of the RNA polymerase to the promoter or displacing the RNA polymerase from promoter–enzyme complexes (99,118,122). The other negative role of the π-protein is its negative control of R6K replication. Biochemical and genetic experiments showed that increased intracellular levels of initiator protein will either lower the plasmid copy number or completely prevent its replication (120). The copy-up phenotype of some of these mutants correlates with base changes within a 100-bp fragment of the *pir* gene and can be suppressed by the presence of the wild-type initiator protein at concentrations lower than that of the mutant protein (123). This negative regulatory activity of π was explained by molecular models involving either direct π–DNA interactions or association of π with either other π molecules or with host-

encoded proteins (120,124). One of these models, termed *handcuffing*, is based on the observation that π has the ability to associate two DNA molecules containing γ-origin sequences and to enhance the DNA ligase–catalyzed multimerization of a single DNA fragment carrying this R6K origin of replication (124). In addition, it was shown that the negative domain of π is located in the N-terminal region of this protein and that the π-mediated inhibition of R6K replication does not require direct binding to DNA (125). Consequently, the origins located within these DNA–protein complexes are unable to initiate replication, most likely because π-induced DNA structure alterations in the origins or potential π–host protein interactions are prevented. In summary, all these observations suggest that the dual activity of π and the regulation of R6K replication is the result of a competition between positive γ–α and γ–β interactions and the inhibitory aggregation of γ-containing molecules, all mediated by the π-replication protein.

Recently, three novel R6K-encoded proteins, designated DDP1, DDP2, and DDP3, were identified (126). DDP1 and DDP2 are encoded by two tandem genes located at the 5′ end of the long inverted repeat of the α-origin, while the gene encoding DDP3 was mapped at the 3′ end of the β-origin long inverted repeat. These three proteins are required for the distortion of the DNA structure of the two R6K active origins. Although the distortions caused by these proteins are potentially linked to R6K replication, they are not equivalent to those described in other replication regions previously characterized (82). It was also suggested (126) that this distortion system serves to synchronize the initiation of replication and establishes the direction of replication from the α- and β-origins by generating a "locked" preinitiation protein–DNA complex.

PT181-TYPE PLASMIDS

The rolling circle type of replication is most often utilized by a number of small, high-copy-number plasmids from Gram-positive bacteria (13,15,21,127). However, plasmids that replicate using this mechanism have recently been identified in a number of Gram-negative bacteria (128–132).

Gram-positive plasmids that replicate using rolling circle mechanisms are classified into five families based on sequence comparisons and genetic organization of the replication regions (13,16,21,133). All plasmids in these families replicate by an asymmetric rolling circle pathway that resembles that of the single-stranded filamentous bacteriophages (16,21,134). One of the best-understood replication mechanisms is that of the pT181-type family. We describe the replication of pT181 as a paradigm of rolling circle replicating plasmids. A more comprehensive description of replication of plasmids in other families can be found elsewhere (13,15,21,127).

A schematic representation of the pT181 replication region and the mechanism of initiation of replication is shown in Figure 6. This plasmid, derived from *Staphylococcus aureus*, encodes a 38-kDa initiator protein, RepC, that has sequence-specific endonuclease and topoisomerase I–like activities (137). To initiate replication, RepC nicks one of the pT181 DNA strands (leading strand) at a specific site, generating a free 3′-OH end that is used as primer for subsequent DNA synthesis (137). The functional structure of the RepC protein is a homodimer (RepC/RepC) (135,138) that recognizes and binds to a site (Rep binding site) that encompasses an inverted repeat (IR III in Fig. 6b) (136). A domain of six amino acids has been identified in RepC that is important in the recognition and interaction of the Rep binding site (Fig. 6) (139). The binding efficiency of RepC is increased by the presence of *cmp*, a 100-bp cis-acting replication enhancer located about 1 kb from the nicking site (140,141). It has been recently demonstrated that an *S. aureus* protein, CBF1, binds *cmp* and increases distortion of the already bent *cmp* locus (142). Whether this binding is associated with the enhancing activity of *cmp* is still not known. Upon binding of the RepC homodimer to IR III, DNA in this region is induced to bend (143) (Fig. 6b). This is closely followed by a change in structure of RepC, DNA melting, and formation of a cruciform structure at the IR II region (Fig. 6b) (139). The melting step is facilitated by the presence of an AT-rich inverted repeat region (IR I) located upstream of IR II (Fig. 6b) (16). The formation of a cruciform structure may result in an approximation of the nicking site of the leading strand and the active site of RepC. This process involves a tyrosine residue that appears to facilitate the generation of the nick. After the endonuclease attack, the RepC protein remains covalently bound to the 5′ end of the DNA by a phosphotyrosine bond (93), probably remaining attached throughout replication of the leading strand. This is not true for all plasmids replicating through a rolling circle mechanism, however, for it was recently shown that in a derivative of pMV158 the initiator protein does not remain covalently bound to the DNA after nicking (144). After generation of the 3′-OH terminal end by RepC, an initiation complex is formed with the probable participation of DNA polymerase III, the helicase PcrA, and single-strand binding protein (Fig. 6b) (16,145,146).

Following (or during) replication, the initiator protein RepC becomes modified (RepC*) by the addition of a 10- to 12-mer oligodeoxynucleotide identical to the sequence located immediately 3′ to the origin of the leading strand. In addition, the two enzymatic activities of RepC are lost during this time (138,147). Further analysis has demonstrated that the active initiator homodimer RepC/RepC becomes the inactive heterodimer RepC/RepC* after it has been used for replication of pT181 (135).

Replication of the lagging strand starts at a different origin than replication of the leading strand. Fig. 6a shows the location of the lagging strand origin in pT181 (148). This origin, also known as SSO (single-strand origin) or *palA*, comprises about a 160-bp palindromic DNA sequence (16,21). Synthesis of the leading strand generates a displaced single-stranded DNA that allows the origin of replication of the lagging strand to adopt the correct conformation to serve as a priming site. Replication of this strand does not require any protein encoded by the plasmid (149). In addition, some rolling circle replicating plasmids have more than one origin of replication for its lagging strand that is functional in different hosts (16). Taken to-

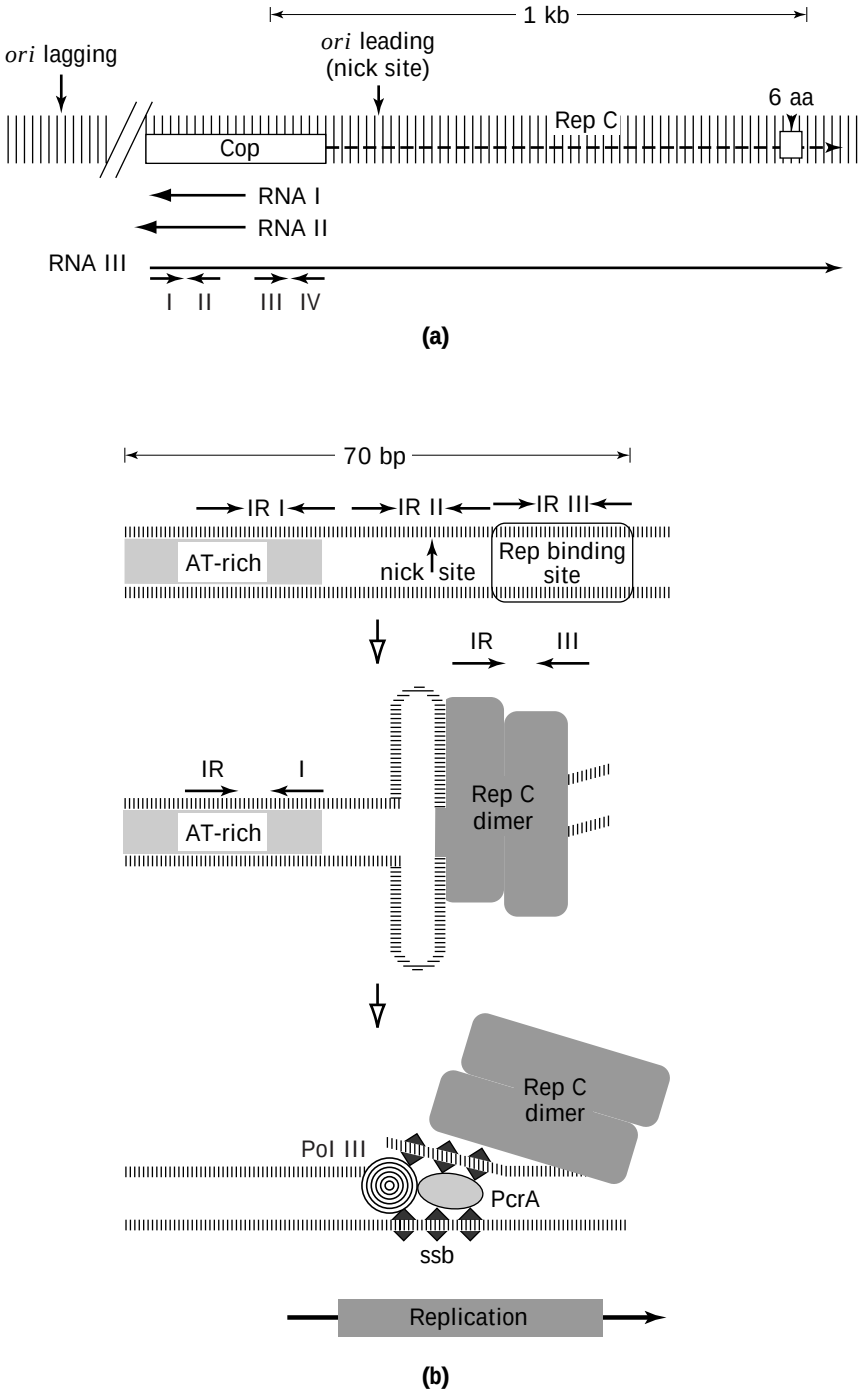


Figure 6. Replication region and model for initiation of replication of the plasmid pT181. (**a**) Diagram of the pT181 replication region. The bar represents pT181 DNA. The two origins of replication (leading and lagging strands) are indicated by different vertical arrows. The broken-line arrow represents the RepC coding region. The six–amino acid region of RepC recognized by *ori* is indicated as a white square. The two diagonal lines indicate that there is a larger distance than shown between the *ori* of the lagging strand and the *ori* of the leading strand. The black arrows represent RNA molecules. The location of the *cop* region is indicated by a box. The small arrows under RNAIII represent the inverted repeats that participate in attenuation. For the sake of clarity, these inverted repeats are not at scale. (**b**) Model for initiation of replication of pT181. The first diagram shows the region encompassing the origin of replication of the leading strand and the three inverted repeats. This region is also referred to as the DSO (double-stranded origin). The RepC binding site located at IR III is boxed. The second diagram shows the formation of the cruciform structure after binding of the RepC homodimer, the bending of the DNA at the binding region, and the change in structure of RepC. The melting process is facilitated by the presence of an AT-rich region that includes IR I. The third diagram shows that after nicking, a replisome is assembled at *ori*, and presumably, DNA polymerase III initiates replication in the presence of the helicase PcrA and single-strand binding protein (ssb). *Source*: Adapted from Refs. 16, 135, and 136.

gether, these facts suggest that lagging strand origins of replication may be important in determining the host range of the plasmid (16).

Termination of synthesis of the leading strand occurs at the nick site by a strand transfer mechanism that is mediated by the replication initiator protein RepC (150). In pT181, once the nick site has been replicated and extended a few nucleotides beyond this site, one of the subunits of the RepC dimer contacts the growing strand. This interaction initiates a strand transfer reaction, resulting in the formation of a single-stranded monomer (old strand), a double-stranded molecule where one of the strands is newly synthesized, and a dimer in which one of the monomers is attached to the oligonucleotide resulting from the extension of replication beyond the nick site (16,21). Although the presence of the nicking site is required, the initiator protein recognition site is not essential for termination (21).

Regulation of initiation of replication of pT181 is achieved through the control of the synthesis of RepC (16,21,151,152). The organization of the pT181 replication regulation region *(cop)* is shown in Fig. 6a. The mRNA

(RNA III) encoding RepC contains a leader sequence typical of genes that are regulated through attenuation (153). This leader sequence has four interacting sequences (I–IV). Sequence I, a nine-nucleotide sequence called *preemptor* can form duplexes with sequences II and III (Fig. 7). In addition, there are two antisense RNA species (RNA I and RNA II) complementary to the leader sequence of RNA III. Analysis of the secondary structure of the leader portion of RNA III in the presence or absence of antisense RNA indicated that upon interaction of the two complementary RNA species, a stem-loop structure between inverted repeats III and IV is formed followed by an AUUUUUU sequence that acts as a transcription termination signal (Fig. 7) (154). On the other hand, if antisense RNA is not present, the interaction between inverted repeats I (the preemptor) and III generates a structure that prevents the formation of the III-IV stem-loop (Fig. 7) (154). Under these conditions, transcription of *repC* takes place. Mutations in the preemptor have been shown to lead to an absolute requirement for RepC in trans (154,155). For example, Figure 7 shows the effect on the secondary structure of RNA III caused by a three-nucleotide replacement mutation in the preemptor. In absence of antisense RNA, the stem-loop structure involving inverted repeats III and IV is formed, no RepC synthesis occurs, leading to a plasmid that has an absolute requirement for RepC in trans (154,155) (Fig. 7).

The main incompatibility elements in pT181-type plasmids are the antisense RNA and the origin of replication of the leading strand (16,154,155). The origin of replication of the leading strand directly binds RepC, titrating the protein, which is present at limiting levels in the cytosol due to the controlling action of the antisense RNAs.

LINEAR PLASMIDS IN BACTERIA

A Unique Plasmid Anatomy

Bacterial genomes have long been considered to contain only circular DNA molecules. However, the development of orthogonal field-pulsed agarose gel electrophoresis in the mid-1980s allowed the discovery of the first linear bacterial chromosome in *Borrelia burgdorferi* (156,157), the causative agent of Lyme disease. Since then, linear replicons have been found in many bacterial genera including *Agrobacterium*, *Streptomyces*, *Thiobacillus*, *Nocardia*, *Rhodococcus*, and even *Escherichia* (158). The first linear plasmid was found in *Streptomyces rochei* in 1979 (159),

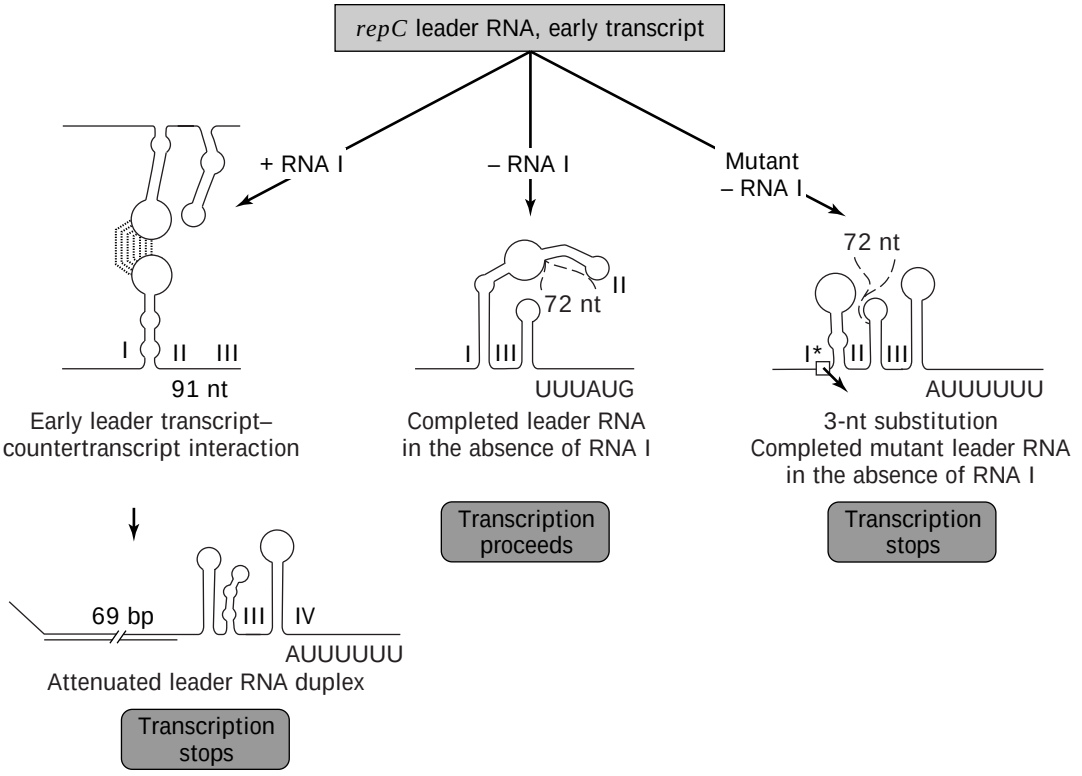


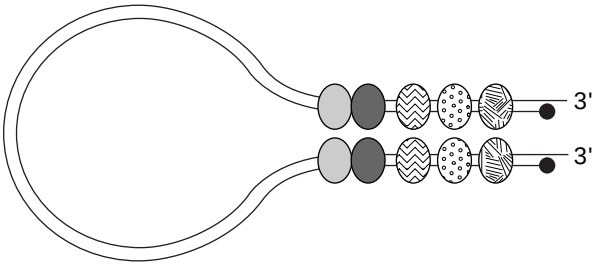
Figure 7. Mechanism of regulation of expression of RepC by antisense RNA. In this model, the early RNA III transcript can interact with antisense RNA (RNA I), allowing the formation of a stem-loop between inverted repeats III and IV that results in a transcription termination signal. In the absence of antisense RNA, the inverted repeat III is sequestered by inverted repeat I, leading to completion of the *repC* mRNA. The three-nucleotide-change mutation in inverted repeat I is also shown. The secondary structure of this mutated RNA III results in formation of the stem loop III–IV even in the absence of antisense RNA, thus leading to a complete dependence of RepC in trans for replication. *Source*: Adapted from Refs. 21 and 154.

and others have been detected in the Gram-positive bacteria *Rhodococcus fascians* and *Nocardia opaca* as well as in the Gram-negative *Thiobacillus versutus* (158). The *Streptomyces* linear plasmids range in size from nine to several hundred kilobases, and all of them contain terminal inverted repeats of different lengths according to the size of the plasmids where they are present. The linear plasmids in *Borrelia* also show a wide range of sizes, typically ranging from 15 to 200 kb and are distributed widely among the members of this species. The *Borrelia* plasmids are unique among extrachromosonal elements because some of them carry critical genes such the *guaA* and *guaB* biosynthetic genes (160), as well as genes encoding the essential major outer surface Osp or Vmp lipoproteins (161–164). In addition, it was determined that some of these linear plasmids are present stably in low copy number, about the equivalent of one per chromosome, and can be cured by exposing *Borrelia* cultures to the DNA gyrase inhibitor novobiocinin (165). These characteristics suggest that the replication and partition processes of linear plasmids are well controlled during cell division. All the facts, together with the observation that the *Borrelia* chromosome has attributes of a linear DNA molecule, has led to the idea that these linear plasmids might be regarded more properly as minichromosomes (166).

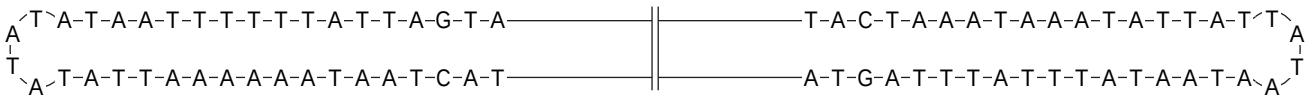
The presence of these linear DNA molecules brings an interesting biological problem associated with the protection of the DNA ends from exonuclease degradation, and the complete replication of these plasmids. These problems are solved in eukaryotes by the addition of telomere repeats by reverse transcription of a short RNA incorporated in the telomerase (167,168). On the other hand, prokaryotes can protect and replicate completely these linear plasmids using two alternative mechanisms.

The ends of linear *Streptomyces* plasmids have telomeric structures also found in adenoviruses and some prokaryotic phages as well as in almost all eukaryotic plasmids. Sequence analysis of pSL2, a linear plasmid present in *S. rochei*, revealed the presence of an inverted terminal repetition of 614 bp and 11 sites of interrupted homology (169,170). These terminal DNA structures, together with specific DNA binding proteins, are thought to be involved in the juxtaposition of the two plasmid termini containing identical or very similar nucleotide sequences. This structure, known as the racket frame–like DNA model and shown in Figure 8a, also includes a terminal protein covalently attached to the 5′ DNA termini, which is required to protect the DNA from degradation and to complete the replication of the 3′ overhanging ends, as described later.

The structure of the 16- and 49-kb linear plasmids of *Borrelia burgdorferi* consists of a double-stranded DNA chain connected at each end by a perfect palindromic AT-rich hairpin loop (Fig. 8b) (2,171). In addition, each end contains a conserved 19-bp inverted repeat sequence, a telomeric structural feature found in all linear plasmids analyzed in *Borrelia* to date (158,162,163). These features have similarities to the telomeres of other linear double-stranded replicons, including among them the genomes of vaccinia and other animal poxviruses (172), the mitochondria of the yeast *Pichia* (173), and the iridovirus, which causes African swine fever (174). The fact that the latter and at least one *Borrelia* species share a common tick vector suggests that the *Borrelia* linear plasmids may have originated by horizontal transfer between kingdoms (2). These telomeric structures indeed play a role in preserving the integrity of these plasmids, and in their replication before cell division.



(a)



(b)

Figure 8. Telomeric structure of linear plasmids. (**a**) Racket frame structure proposed for linear *Streptomyces* plasmids. The ovals represent juxtaposition proteins that recognize specific regions of palindromic symmetry, bringing together the plasmid termini. The black circles represent the terminal protein attached to the 5′ ends, which is required to complete the replication of a 280-nucleotide segment at both plasmid termini. (**b**) Structure and partial nucleotide sequence of the telomeric termini of linear plasmids isolated from *Borrelia*.

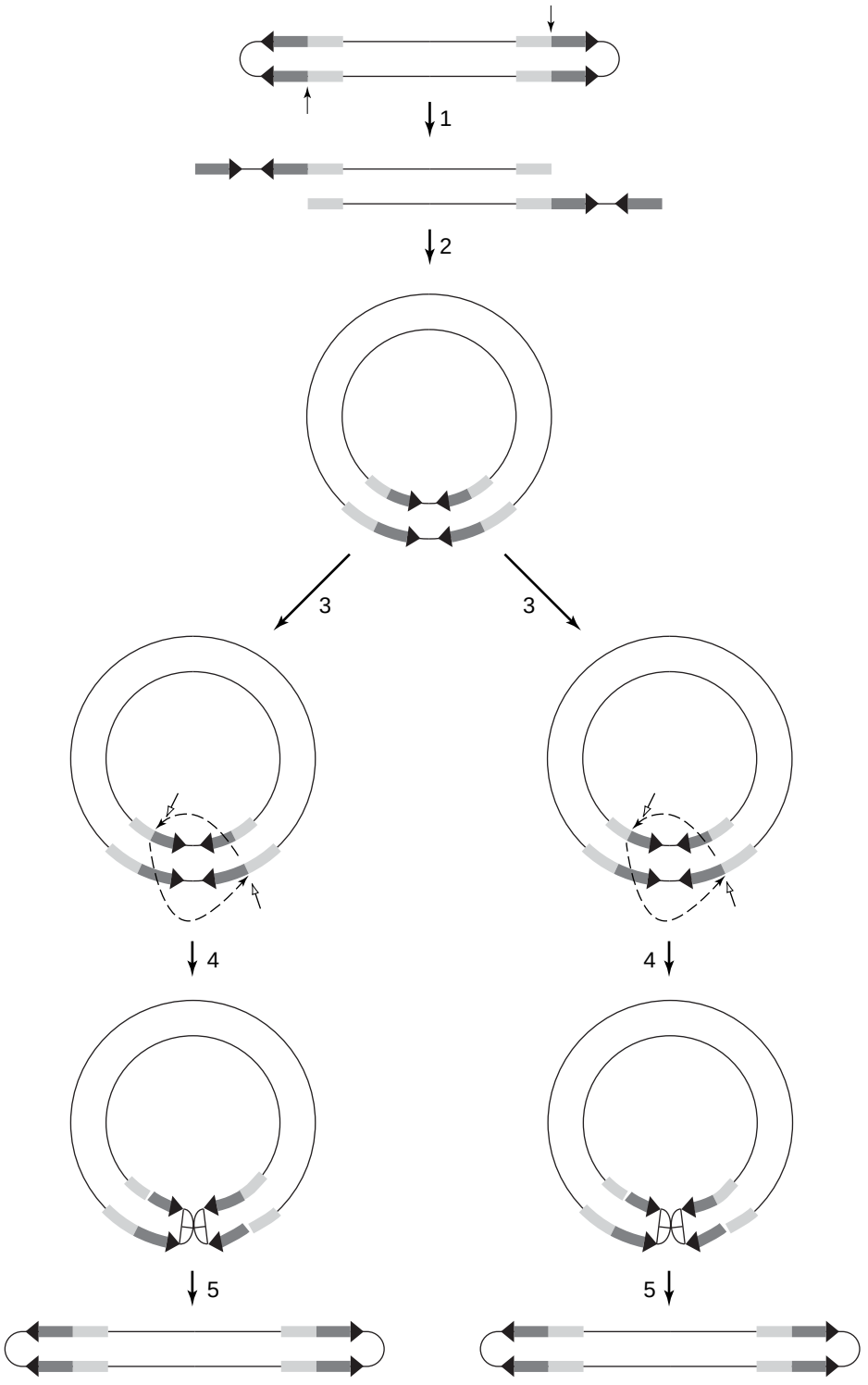


Figure 9. Model depicting the different stages leading to the formation of circular intermediates and replication of linear plasmids in *Borrelia*. According to this model, in step 1 the ends of the plasmid are open by a nick (indicated by the arrows with open arrowheads) within the inverted repeats (represented by the thick arrows) located at the plasmid termini. The open linear plasmid circularizes due to the presence of complementary sequences at its termini (step 2) and replicates as a circular replicon (step 3). A second nick is introduced within the repeated sequences, and the free single-stranded ends pair back (indicated by the dashed arrows) with their complementary copies located in the same monomer and, thus, reconstitute the hairpins at each plasmid end (step 4). After DNA ligation (step 5), two copies of the original linear plasmid are generated. *Source*: Adapted from Ref. 157.

Replication of Linear Plasmids

It was proposed initially that the linear plasmid of *Streptomyces* replicates by a protein-primed replication mechanism (158,175) that has been well characterized for adenovirus, phage ø29, and *Bacillus subtilis* (168). In this mechanism of DNA replication, the telomere is the origin of replication, and it is recognized by specific DNA binding proteins that promote the unwinding of the double helix and serve as the primer for a specific DNA polymerase. However, it was later observed that the *S. clavuligerus* linear plasmid pSCL can replicate as a circular DNA molecule when the telomeres are removed and the ends are ligated (176). This observation was further confirmed by the de-

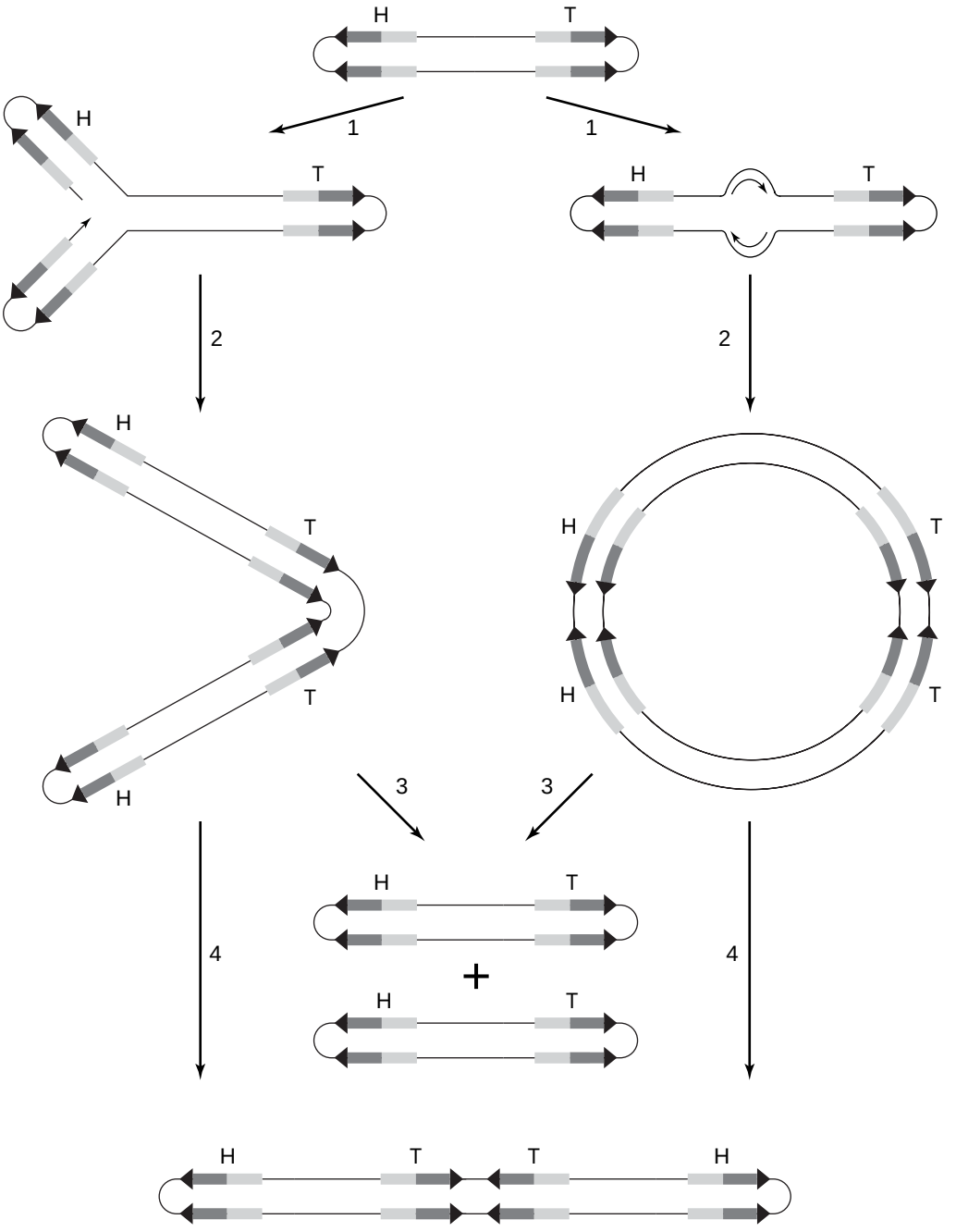


Figure 10. Model showing an alternative replication mechanism and the formation of tail (T)-to-tail dimers of linear plasmids in *Borrelia*. The model depicted in the left half starts with the cleavage of the head (H) telomere and the initiation of DNA replication (step 1). The replication of the entire linear replicon results in a tail-to-tail dimer (step 2) that can result in either two copies of the original monomers (step 3), after resolution of the tail telomeres, or remain as a tail-to-tail dimer due to a failure in the plasmid segregation mechanism (step 4). The model shown in the right half predicts the initiation of DNA replication from an internal origin (step 1) that results in the formation of a circular intermediate (step 2). Segregation of this intermediate by independent DNA cleavage events at each telomere junction leads to the generation of two copies of the original linear replicon (step 3). Conversely, a failure of the cleavage of the tail telomere results in the formation of a tail-to-tail dimer (step 4). *Source*: Adapted from Ref. 178.

tailed analysis of the mechanism of replication of pSLA2, another *Streptomyces* plasmid that contains protein covalently attached to the 5′ end (175). Ligation of pSLA2 deletion derivatives lacking the telomeric termini to a selectable marker resulted in a high-copy-number circular extrachromosomal element. This observation suggested that this linear plasmid must contain an internal site capable of promoting DNA replication. Two-dimensional agarose gel electrophoresis analysis showed that this plasmid is indeed replicated primarily bidirectionally from an internal origin located near the center of the plasmid, toward the ends, rather than by full-length strand displacement initiated at the telomeres. However, pSLA2 still requires a protein-primed strand-displacing mechanism to complete the replication of a 280-nucleotide segment at both plasmid termini. It was suggested (175) that the synthesis of the 5′ terminal segment of the lagging strand of this linear plasmid uses as a template the 3′ overhang of the leading strand and is primed by the protein covalently attached to the 5′ end of the mature plasmid.

The replication mechanisms of linear plasmids in *Borrelia* are poorly understood; however, our knowledge of these novel replicons encoding essential genes has been extended by a series of very recent publications (166,177,178). The replication of some of the *Borrelia* linear plasmids involves circular intermediates. In addition, it was recently reported that the 180-kb plasmid of *Borrelia hermsii* can exist and replicate either as a linear or a monomeric circular element (166). According to these observations, it was proposed that the monomeric circular intermediate is formed by a head-to-tail junction, after a nick introduced in the inverted repeat opens the plasmid termini (Fig. 9). Because the termini are complementary, the plasmid circularizes and can replicate as any other well-characterized circular replicon. This circular intermediate is then resolved into a linear double-stranded DNA structure by a second nick within the terminal inverted repeats. Because the terminal repeats are complementary, a linear molecule should be reconstituted.

More recently, it was found that *B. burgdorferi* sensu lato contains atypical large linear plasmids ranging from 92 to 105 kb (178). These plasmids carry *p27* and *ospAB*, genes that were also detected in other isolates on the 50-kb linear plasmid pAB50. A more detailed analysis of the larger plasmids demonstrated that they are formed by tail-to-tail dimerization of pAB50. The presence of such dimers can explain the unusual plasmid variability observed among different isolates and may provide new information regarding the replication mechanism of these linear replicons. It was postulated (178) that these linear dimers are the result of failed segregation after DNA replication by a mechanism similar to that described for vaccinia viruses. In this model proposed by Marconi et al. (178) and shown in Figure 10, initiation of plasmid replication proceeds from one of the termini, after the head hairpin loop in this particular case is nicked, allowing the formation of tail-to-tail dimer intermediates. Alternatively, these dimers can arise by DNA replication initiated from an origin located within a linear monomer that results in a circular replication intermediate. In the normal replication process, the circular intermediate is resolved into two linear monomers by independent cleavage events at each telomere junction. Cleavage failure at the tail hairpin by a specific DNA cleavage system results in two monomers linked tail to tail (Fig. 10), a possibility that was confirmed experimentally by restriction analysis and Southern blot DNA hybridization experiments (178). However, this same analysis demonstrated that not all *Borrelia* large plasmids were originated via dimer formation, and alternative DNA replication systems that are still uncharacterized must exist.

Some of the proteins involved in the replication of the linear plasmids in *Borrelia* were identified by determining the complete nucleotide sequence of the 16-kb plasmid lp 16.9 isolated from *B. burgdorferi* B31 (177). This study revealed the presence of 15 open reading frames named A to O. The predicted proteins encoded by OrfM and OrfN showed homology to proteins involved in plasmid and cell replication. OrfM is homologous to MinD, a cytoplasmic membrane protein with ATPase activity required for the correct placement of the division site (179,180). OrfM also has homologies with plasmid partition proteins such as ParA or SopA (181,182), and the RepB protein encoded by the pheromone-responsive plasmid pAD1 in *Enterococcus faecalis* and involved in the control of the copy number of this extrachromosomal element (183). The predicted features and the primary sequence of the OrfN hypothetical polypeptide are similar to those of the CopB DNA binding proteins, also known as RepB or RepA2, involved in the control of plasmid copy number in Gram-negative bacteria (184). This sequence analysis also led to the hypothesis that OrfN may interact with the short repeat sequence located within the promoter region of OrfM and, thus, control the expression of the latter (177).

In summary, sequence analysis of *Borrelia* plasmids showed that the few proteins identified are similar to prokaryotic proteins involved in plasmid replication and maintenance in Gram-positive and Gram-negative bacteria. In addition, these results do not support the previous conclusion that the linear plasmids in *Borrelia* are indeed derived from animal viruses. Furthermore, it seems that circular and linear plasmids of this bacterial genus have a common origin and share DNA replication mechanisms.

BIBLIOGRAPHY

1. M.E. Tolmasky, L.A. Actis, and J.H. Crosa, in J. Lederberg ed., *Encyclopedia of Microbiology*, vol. 3, Academic Press, San Diego, 1992, pp. 431–442.
2. J. Hinnebusch and A.G. Barbour, *J. Bacteriol.* **173**, 7233–7239 (1991).
3. H. te Riele, B. Michel, and D. Ehrlich, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2541–2545 (1985).
4. A. Abeles, L. Reaves, B. Youngren-Grimes, and S. Austin, *Mol. Microbiol.* **18**, 903–912 (1995).
5. R. Brubaker, *Clin. Microbiol. Rev.* **4**, 309–324 (1991).
6. J.H. Crosa, *Microbiol. Rev.* **53**, 517–530 (1989).
7. A. Jones, E. Lai, K. Shirasu, and C. Kado, *J. Bacteriol.* **178**, 5706–5711 (1996).
8. Y. Lee, S. Jin, W. Sim, and E. Nester, *Gene* **179**, 83–88 (1996).
9. R. Hall and C. Collis, *Mol. Microbiol.* **15**, 593–600 (1995).
10. N. Kleckner, *Annu. Rev. Genet.* **15**, 341–404 (1981).
11. M. Tolmasky and J. Crosa, *Plasmid* **29**, 31–40 (1993).

12. S. Cohen, *Gene* **135**, 67–76 (1993).
13. M. Espinosa, G. del Solar, F. Rojo, and J. Alonso, *FEMS Microbiol. Lett.* **130**, 111–120 (1995).
14. M. Filutowicz, S. Dellis, I. Levchenko, M. Urh, F. Wu, and D. York, *Prog. Nucleic Acids Res. Mol. Biol.* **48**, 239–273 (1994).
15. A. Gruss and S. Ehrlich, *Microbiol. Rev.* **53**, 231–241 (1989).
16. D. Helinski, A. Toukdarian, and R. Novick, in F. Neidhardt ed., Escherichia coli *and* Salmonella, vol. 2, ASM Press, Washington, D.C., 1996, pp. 2295–2324.
17. B. Kittel and D. Helinski, in D. Clewel ed., *Bacterial Conjugation*, Plenum Press, New York, (1993).
18. U. Kues and U. Stahl, *Microbiol. Rev.* **53**, 491–516 (1989).
19. D. Manen and L. Caro, *Mol. Microbiol.* **5**, 233–237 (1991).
20. K. Nordstrom, *Cell* **63**, 1121–1124 (1990).
21. R. Novick, *Annu. Rev. Microbiol.* **43**, 537–565 (1989).
22. J. Viret, A. Bravo, and J. Alonso, *Microbiol. Rev.* **55**, 675–683 (1991).
23. E. Wagner and R. Simons, *Annu. Rev. Microbiol.* **48**, 713–748 (1994).
24. D. Astill, P. Manning, and M. Heuzenroeder, *Plasmid* **30**, 258–267 (1993).
25. M. Bagdasarian, and M. Bagdasarian, in P. Gerhardt ed., *Methods for General and Molecular Bacteriology*, American Society for Microbiology, Washington, D.C., (1993), pp. 406–417.
26. K. Dery, R. Chavideh, V. Waters, R. Chamorro, L. Tolmasky, and M.E. Tolmasky, *Plasmid*, **38**, 97–105 (1997).
27. J. Fu, H. Chang, Y. chen, Y. Chang, and S. Liu, *Plasmid* **34**, 75–84 (1995).
28. N. Nomura and Y. Murooka, *J. Ferment. Bioeng.* **78**, 250–254 (1994).
29. R. Old and S. Primrose, Principles of Gene Manipulation, Blackwell Scientific, Oxford, 1985.
30. J. Sambrook, E. Fritsch, and T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989.
31. G. Cesareni, C. Helmer, and L. Castagnoli, *Trends Genet.* **7**, 230–235 (1991).
32. Y. Eguchi, T. Itoh, and J. Tomizawa, *Annu. Rev. Biochem.* **60**, 631–652 (1991).
33. E. Gerhardt, H. Wagner, and R. Simmons, *Annu. Rev. Microbiol.* **48**, 713–742 (1994).
34. S. Merlin and B. Polisky, *J. Mol. Biol.* **248**, 211–219 (1995).
35. K. Nordstrom and E. Wagner, *Trends Biochem. Sci.* **19**, 294–300 (1994).
36. B. Polisky *Cell* **55**, 929–932 (1988).
37. J. Inselburg, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 2256–2259 (1974).
38. J. Tomizawa, M. Ohmori, and R. Bird, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 1865–1869 (1977).
39. T. Itoh and J. Tomizawa, *Cold Spring Harbor Symp. Quant. Biol.* **43**, 409–418 (1978).
40. T. Itoh and J. Tomizawa, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2450–2454 (1980).
41. T. Itoh and J. Tomizawa, *Nucleic Acids Res.* **10**, 5949–5965 (1982).
42. J. Tomizawa, T. Itoh, G. Selzer, and T. Som, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 1421–1425 (1981).
43. L. Castagnoli, R. Lacatena, and G. Cesareni, *Nucleic Acids Res.* **13**, 5353–5367 (1985).
44. H. Masukata and J. Tomizawa, *Cell* **44**, 125–136 (1986).
45. H. Masukata and J. Tomizawa, *Cell* **36**, 513–522 (1984).
46. H. Masukata and J. Tomizawa, *Cell* **62**, 331–338 (1990).
47. T. Fitzwater, Y. Yang, X. Zhang, and B. Polisky, *J. Mol. Biol.* **226**, 997–1008 (1992).
48. G. Selzer, and J. Tomizawa, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 7082–7086 (1982).
49. S. Dasgupta, H. Masuakata, and J. Tomizawa, *Cell* **51**, 1113–1122 (1987).
50. H. Masukata, S. DasGupta, and J. Tomizawa, *Cell* **51**, 1123–1130 (1987).
51. H. Ohmori, Y. Murakami, and T. Nagata, *J. Mol. Biol.* **198**, 223–235 (1987).
52. A. Kornberg and T. Baker, *DNA Replication*, 2nd ed, W.H. Freeman, New York, 1991.
53. D. Kingsbury and D. Helinski, *Biochem. Biophys. Res. Commun.* **41**, 1538–1544 (1970).
54. S. Naito and H. Uchida, *J. Bacteriol.* **166**, 143–147 (1986).
55. R. Lacatena and G. Cesareni, *Nature* **294**, 623–626 (1981).
56. J. Tomizawa, *Cell* **47**, 89–97 (1986).
57. J. Tomizawa and T. Itoh, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6096–6100 (1981).
58. Y. Eguchi and J. Tomizawa, *Cell* **60**, 199–209 (1990).
59. Y. Eguchi, and J. Tomizawa, *J. Mol. Biol.* **220**, 831–842 (1991).
60. J. Tomizawa, *Cell* **40**, 527–535 (1985).
61. J. Tomizawa, *J. Mol. Biol.* **212**, 683–694 (1990).
62. J. Tomizawa and T. Som, *Cell* **38**, 871–878 (1984).
63. J. Tomizawa, *Cell* **38**, 861–870 (1984).
64. E. Wong, and B. Polisky, *Cell* **42**, 959–966 (1985).
65. S. Lin-Chao and H. Bremer, *Bacteriol.* **169**, 1217–1222 (1987).
66. L. He, F. Soderbom, E. Wagner, U. Binnie, N. Binns, and M. Masters, *Mol. Microbiol.* **9**, 1131–1142 (1993).
67. F. Xu, S. Lin-Chao, and S. Cohen, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6756–6760 (1993).
68. T. Tomcsanyi and D. Apirion, *J. Mol. Biol.* **185**, 713–720 (1985).
69. G. Cesareni, M. Cornelissen, M. Lacatena, and L. Castagnoli, *EMBO J.* **3**, 1365–1369 (1984).
70. M.I. Lacatena, D. Banner, L. Castagnoli, and G. Cesareni, *Cell* **37**, 1009–1014 (1984).
71. J. Tomizawa, *J. Mol. Biol.* **212**, 695–708 (1990).
72. A. Twigg and D. Sherrat, *Nature* **283**, 216–218 (1980).
73. R. Novick, *Microbiol. Rev.* **51**, 381–395 (1987).
74. N. Datta and P. Kontomichalou, *Nature (Lond.)* **208**, 239–241 (1965).
75. P. Kontomichalou, M. Mitani, and R.C. Clowes, *J. Bacteriol.* **104**, 34–44 (1970).
76. D. Bastia and J. Germino, *Cell* **32**, 131–140 (1983).
77. J. Germino and D. Bastia, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4692–4696 (1984).
78. J. Germino and D. Bastia, *Cell* **32**, 131–140 (1983).
79. J. Germino, J.G. Gray, H. Charbonneau, T. Vanaman, and D. Bastia, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 6848–6852 (1983).
80. V. De Lorenzo and K.N. Timmis, *Methods Enzymol.* **235**, 386–405 (1994).
81. V.L. Miller and J.J. Mekalanos, *J. Bacteriol.* **170**, 2575–2583 (1988).

82. D. Bramhill and A. Korngerg, *Cell* **54**, 915–918 (1988).
83. T. Morotsu, K. Matsubara, H. Sugusaki, and M. Takanami, *Gene* **15**, 257–271 (1981).
84. C. Vocke and D. Bastia, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 6557–6561 (1983).
85. A.L. Abeles, K.M. Snyder, and D.K. Chattoraj, *J. Mol. Biol.* **173**, 307–324 (1984).
86. C. Bidinost, J.H. Crosa, and L.A. Actis, *Plasmid* **31**, 242–250 (1994).
87. Y. Kamio, Y. Itoh, and Y. Terawaki, *J. Bacteriol.* **170**, 4411–4414 (1988).
88. A.E. Gammie and J.H. Crosa, *Mol. Microbiol.* **5**, 3015–3023 (1991).
89. B.L. Kittell and D.R. Helinski, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1389–1393 (1991).
90. J.H. Crosa, L. Luttropp, and S. Falkow, *J. Bacteriol.* **126**, 454–466 (1976).
91. J.H. Crosa, *J. Biol. Chem.* **255**, 11075–11077 (1980).
92. T. Horiuchi and M. Hidaka, *Cell* **54**, 515–523 (1988).
93. C. Thomas, D. Balson, and W. Shaw, *J. Biol. Chem.* **265**, 5519–5530 (1990).
94. I. Patel and D. Bastia, *Cell* **47**, 785–792 (1986).
95. I. Patel and D. Bastia, *Cell* **51**, 455–462 (1987).
96. P.R. Sista, C.A. Hutchinson III, and D. Bastia, *Genes Dev.* **5**, 74–82 (1991).
97. R. Kolter, M. Inuzuka, and D.R. Helinski, *Cell* **15**, 1199–1208 (1978).
98. J. Germino and D. Bastia, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 5475–5479 (1982).
99. A. Shafferman, R. Kolter, D.M. Stalker, and D.R. Helinski, *J. Mol. Biol.* **161**, 57–76 (1982).
100. M. Inuzuka and D.R. Helinski, *Biochemistry* **17**, 2567–2573 (1978).
101. A. Shafferman and D.R. Helinski, *J. Biol. Chem.* **258**, 4083–4090 (1983).
102. D.M. Stalker, R. Kolter, and D.R. Helinski, *J. Mol. Biol.* **161**, 33–43 (1982).
103. J.H. Crosa, L.K. Luttropp, and S. Falkow, *J. Mol. Biol.* **124**, 443–468 (1978).
104. W.L. Kelly, I. Patel, and D. Bastia, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5078–5082 (1992).
105. A. Miron, S. Mukherjee, and D. Bastia, *EMBO J.* **11**, 1205–1216 (1992).
106. F. Wu, I. Goldberg, and M. Filutowicz, *Nucleic Acids Res.* **20**, 811–817 (1992).
107. F. Wu, I. Levchenko, and M. Filutowicz, *J. Bacteriol.* **177**, 6338–6345 (1995).
108. P.F. Jemilohun, C.W. Clark, and E.R. Archibold, *Biochem. Biophys. Res. Commun.* **221**, 186–192 (1996).
109. F. Wu, J. Wu, J. Ehley, and M. Filutowics, *J. Bacteriol.* **178**, 4965–4974 (1996).
110. A. Shafferman, Y. Flashner, I. Hertman, and M. Lion, *Gen. Genet.* **208**, 263–270 (1987).
111. P.V.A.L. Ratnakar, B.K. Mohanty, M. Lobert, and D. Bastia, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5522–5526 (1996).
112. J.H. Crosa, L.K. Luttropp, and S. Falkow, *Cold Spring Harbor Symp. Quant. Biol.* **43**, 111–120 (1979).
113. P. Mukhopadhyay, M. Filutowicz, and D.R. Helinski, *J. Biol. Chem.* **261**, 9534–9539 (1986).
114. M. Shon, J. Germino, D. Bastia, *J. Biol. Chem.* **257**, 13823–13827 (1982).
115. R. Kolter and D.R. Helinski, *J. Mol. Biol.* **161**, 45–56 (1982).
116. M. Filutowicz, E. Uhlenhopp, and D.R. Helinski, *J. Mol. Biol.* **187**, 225–239 (1985).
117. J. Germino and D. Bastia, *Cell* **34**, 125–134 (1983).
118. W. Kelly and D. Bastia, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 2574–2578 (1985).
119. S. Mukherjee, H. Ericson, and D. Bastia, *Cell* **52**, 375–383 (1988).
120. M. Filutowicz, M.J. McEachern, and D.R. Helinski, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9645–9649 (1986).
121. S. Mukherjee, I. Patel, and D. Bastia, *Cell* **43**, 189–197 (1985).
122. M. Filutowicz, G. Davis, A. Greener, and D.R. Helinski, *Nucleic Acids Res.* **13**, 103–114 (1985).
123. D.M. Stalker, M. Filitowicz, and D.R. Helinski, *Proc. Natl. Acad. Sci.* **80**, 5500–5504 (1983).
124. M.J. McEachern, M.A. Bott, P.A. Tooker, and D.R. Helinski, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 7942–7946 (1989).
125. A. Greener, M.S. Filutowicz, M.J. McEachern, and D.R. Helinski, *Mol. Gen. Genet.* **224**, 24–32 (1990).
126. Y. Flashner, J. Shlomai, and A. Shafferman, *Mol. Microbiol.* **19**, 986–996 (1996).
127. G. del Solar, M. Moscoso, and M. Espinosa, *Mol. Microbiol.* **8**, 789–796 (1993).
128. D. Galli and D. Leblanc, *J. Bacteriol.* **177**, 4474–4480 (1995).
129. A. Gielow, L. Diederich, and W. Messer, *J. Bacteriol.* **173**, 73–79 (1991).
130. H. Kleanthous, C. Clayton, and S. Tabaqchali, *Mol. Microbiol.* **5**, 2377–2389 (1991).
131. H. Yasukawa, T. Hase, A. Sakai, and Y. Masamune, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 10282–10286 (1991).
132. N. Zhang and J. Brooker, *Plasmid* **29**, 125–134 (1993).
133. J. Alonso and M. Espinosa, in K. Hardy ed., *Plasmids: A Practical Approach*, 2nd ed., vol. 138, IRL Press, New York, 1993, pp. 39–63.
134. P. Baas and H. Jansz, *Curr. Topics Microbiol. Immunol.* **136**, 31–70 (1988).
135. A. Rasooly, P. Wang, and R. Novick, *EMBO J.* **13**, 5245–5251 (1994).
136. F. Wang, S. Projan, V. Herniquez, and R. Novick, *EMBO J.* **12**, 45–52 (1993).
137. R. Koepsel, R. Murray, W. Rosenblum, and S. Khan, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6845–6849 (1985).
138. A. Rasooly and R. Novick, *Science* **262**, 1048–1050 (1993).
139. F. Wang, S. Projan, V. Herniquez, and R. Novick, *J. Mol. Biol.* **223**, 145–158 (1992).
140. M. Gennaro and R. Novick, *J. Bacteriol.* **168**, 160–166 (1986).
141. V. Henriquez, V. Milisavljevic, J. Kahn, and M. Gennaro, *Gene* **134**, 93–98 (1993).
142. Q. Zhang, S. Soares de Oliveira, R. Colangeli, and M. Gennaro, *J. Bacteriol.* **179**, 684–688 (1997).
143. R. Koepsel and S. Khan, *Science* **233**, 1316–1318 (1986).
144. M. Moscoso, G. del Solar, and M. Espinosa, *J. Biol. Chem.* **270**, 3772–3779 (1995).
145. S. Iordanescu, *J. Bacteriol.* **175**, 3916–3917 (1993).
146. S. Iordanescu, *J. Mol. Biol.* **221**, 1183–1189 (1991).
147. A. Rasooly, S. Projan, and R. Novick, *J. Bacteriol.* **176**, 2450–2453 (1994).
148. A. Gruss, H. Ross, and R. Novick, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2165–2169 (1987).

149. P. Birch and S. Khan, *Proc. Nat. Acad. Sci. U.S.A.* **89**, 290–294 (1992).
150. S. Iordanescu and S. Projan, *J. Bacteriology* **170**, 3427–3434 (1988).
151. I. Carleton, S. Projan, S. Highlander, and R. Novick, *EMBO J.* **3**, 2407–2414 (1984).
152. R. Novick, S. Projan, S. Kumar, C. Carleton, S. Gruss, S. Highlander, and J. Kornblum, in D. Helinski, S. Cohen, D. Clewell, D. Jackson, and A. Hollaender ed., *Plasmids in Bacteria*, Plenum Press, New York, 1985, pp. 299–320.
153. R. Landick, C. Turnbough, and C. Yanofsky, in F. Neidhardt, R. Curtis III, J. Ingraham, C. Lin, K. Low, B. Magasanik, W. Reznikoff, M. Riley, M. Schachter, and H. Umbarger eds., Escherichia coli *and* Salmonella: *Cellular and Molecular Biology*, 2nd ed., American Society for Microbiology, Washington, D.C., 1996, pp. 1263–1286.
154. R. Novick, S. Iordanescu, S. Projan, J. Kornblum, and I. Edelman, *Cell* **59**, 395–404 (1989).
155. R. Novick, G. Adler, S. Projan, S. Carleton, S. Highlander, A. Gruss, S. Khan, and S. Iordanescu, *EMBO J.* **10**, 2399–2405 (1984).
156. C. Baril, C. Ricahud, G. Baranton, and I. Saint Girons, *Res. Micriobiol.* **140**, 507–516 (1989).
157. M.S. Ferdows and A.G. Barbour, *Proc. Nat. Acad. Sci. U.S.A.* **86**, 5969–5973 (1989).
158. J. Hinnebusch and H. Tilly, *Mol. Microbiol.* **10**, 917–922 (1993).
159. T. Hayakawa, T. Tanaka, K. Sakaguchi, N. Otake, and H. Yonehara, *J. Gen. Appl. Microbiol.* **25**, 255–260 (1979).
160. N. Margolis, D. Hogan, K. Tilly, and P.A. Rosa, *Bacteriol.* **176**, 6427–6432 (1994).
161. A.G. Barbour, *Trends Microbiol.* **1**, 236–239 (1993).
162. A.G. Barbour and C.F. Garon, *Science* **237**, 409–411 (1987).
163. T. Kitten and A.G. Barbour, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6077–6081 (1990).
164. R.H. Plasterk, M.I. Simon, and A.G. Barbour, *Nature (Lond.)* **318**, 257–263 (1985).
165. J. Hinnebusch and A.G. Barbour, *Bacteriol.* **174**, 5251–5257 (1992).
166. M.S. Ferdows, P. Serwer, G.A. Griess, S.J. Norris, and A.G. Barbour, *J. Bacteriol.* **178**, 793–800 (1996).
167. F. Meinhardt, F. Kempken, J. Kamper, and K. Esser, *Curr. Genet.* **17**, 89–95 (1990).
168. M. Salas, *Annu. Rev. Biochem.* **60**, 39–71 (1991).
169. H. Hirochika, K. Nakamura, and K. Sakaguchi, *EMBO J.* **3**, 761–766 (1984).
170. K. Sakaguchi, H. Hirochika, and N. Gunge, in S.N.C.D.R. Helinski, D.B. Clewell, D.A. Jackson, A. Hollaender eds., *Plasmids in Bacteria*, vol. 30, Plenum Press, New York, 1985, pp. 433–451.
171. J. Hinnebusch, S. Bergstrom, and A.G. Barbour, *Mol. Microbiol.* **4**, 811–820 (1990).
172. B.M. Baroudy, S. Venkatesan, and B. Moss, *Cell* **28**, 315–324 (1982).
173. N. Dinouel, R. Drissi, I. Miyakawa, F. Sor, S. Rousset, and H. Fukuhara, *Mol. Cell. Biol.* **13**, 2315–2323 (1993).
174. A. Gonzalez, A. Talavera, J.M. Almendral, and E. Vinuela, *Nucleic Acid Res.* **14**, 6835–6844 (1986).
175. P.-C. Chang, and S.N. Cohen, *Science* **265**, 952–954 (1994).
176. D. Shiffman and S.N. Cohen, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6129–6133 (1992).
177. A.G. Barbour, C.J. Crater, V. Bundoc, and J. Hinnebusch, *J. Bacteriol.* **178**, 6635–6639 (1996).
178. R.T. Marconi, S. Casjens, U.G. Munderloloh, and D.S. Samuels, *J. Bacteriol.* **178**, 3357–3361 (1996).
179. P.A.J. de Boer, R.E. Crossley, A.R. Hand, and L.I. Rothfield, *EMBO J.* **10**, 4371–4380 (1991).
180. A.W. Varley and G.C. Stewart, *J. Bacteriol.* **174**, 6729–6742 (1992).
181. A.L. Abeles, S.A. Friedman, and S.J. Austin, *J. Mol. Biol.* **185**, 261–272 (1985).
182. H. Mori, A. Kondo, A. Ohshima, T. Oruga, and S. Hiraga, *J. Mol. Biol.* **192**, 1–15 (1986).
183. K.E. Weaver, D.B. Clewell, and F. An, *J. Bacteriol.* **175**, 1900–1909 (1993).
184. J. Lopez, I. Andres, J.M. Ortiz, and J.C. Rodriguez, *Nucleic Acid Res.* **18**, 7177 (1990).

PLURONIC POLYOLS, CELL PROTECTION

DAVID W. MURHAMMER
University of Iowa
Iowa City, Iowa

KEY WORDS

Cell culture scale-up
Pluronic polyols
Sparging protection

OUTLINE

INTRODUCTION

In recent years, it has become apparent that there is a need for animal cell cultures to produce some products (e.g., complex recombinant DNA proteins) because of the inability of other expression systems (e.g., bacteria) to properly

perform post-translational modifications (e.g., glycosylation and proper folding). Unfortunately, animal cells lack a protective cell wall and thus are subject to damage from sparging and other environmental stresses. In large-scale bioreactors, however, aerating the cultures through direct sparging is the most practical method of supplying oxygen to the culture medium. Therefore, in order to realize the full potential of using animal cells to produce recombinant DNA proteins on an industrial scale, there is a critical need for large-scale sparged bioreactors capable of supporting animal cell growth. Currently, the generally accepted approach to solving this problem is to supplement the medium with a protective agent, such as the commercially available surfactant Pluronic F-68.

PROPERTIES OF PLURONIC POLYOLS

A wide range of Pluronic polyols are available that are capable of protecting animal cells in suspension cultures from the adverse effects of sparging. Of these, Pluronic F-68 is by far the most widely used. Properties of Pluronic polyols that are relevant to the selection of a protective agent are discussed in this section.

Synthesis and Structure

Pluronic polyols, commercially available from the BASF Corporation (Wyandotte, Mich.), are block copolymer nonionic surfactants consisting of a center block of poly(oxypropylene) (hydrophobe) and end blocks of poly(oxyethylene) (hydrophile). They are synthesized by the two-step process shown in Figure 1 (1). In the first step, a poly(oxypropylene) block is synthesized by the addition of propylene oxide to propylene glycol. The poly(oxyethylene) blocks are then added to both ends by reacting the resulting poly(oxypropylene) with ethylene oxide. The oxyalkylation steps are carried out in the presence of an alkaline catalyst (e.g., potassium hydroxide) that is subsequently neutralized and removed from the final product. The properties of the Pluronic polyols are altered by varying the relative sizes of the poly(oxypropylene) and poly(oxyethylene) blocks (note that the sizes of the two poly(oxyethylene) blocks are statistically identical). The properties of a variety of Pluronic polyols are given in Table 1.

Nomenclature

The letters L, P, and F in the Pluronic polyol nomenclature represent liquid, paste, and flake, respectively (Table 1) (1). The first two digits of the three-digit polyol number, or the first digit of a two-digit polyol number, are indicative of the hydrophobe [poly(oxypropylene)] molecular weight, for example, 6, 8, and 12 represent average hydrophobe molecular weights of 1,750, 2,250, and 4,000, respectively. The last digit in the polyol number represents approximately one-tenth the percentage of hydrophile [poly(oxyethylene)] in the molecule. For example, Pluronic F-68 is a flake (i.e., a white powder), has a hydrophobe molecular weight of approximately 1,750, and contains approximately 80% hydrophile.

Impurities

Commercial-grade Pluronic polyols have been demonstrated to contain low molecular weight impurities, including aldehydes, formic acid, and acetic acid (2,3). In addition, peroxide derivatives of the Pluronic polyols can be formed during steam sterilization. The formation of the peroxide derivatives can be prevented by filter sterilizing the Pluronic polyol solutions instead of using steam sterilization. The low molecular weight impurities can be removed from aqueous solutions of Pluronic polyols by silica-gel column absorption (2).

Hydrophilic–Lipophilic Balance

The hydrophilic–lipophilic balance HLB is an empirical measure of the emulsifying capabilities of surfactant mol-

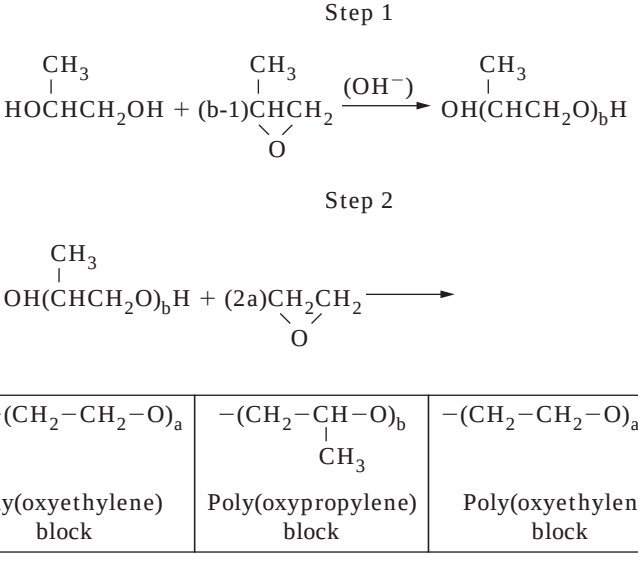


Figure 1. Pluronic polyol synthesis.

Table 1. Pluronic Polyols Tested for Their Effect on *Spodoptera frugiperda* Sf-9 Insect Cell Growth in the Absence and Presence of Sparging

Pluronic polyol	Average MW	wt % hydrophobe	HLB range	Effect on cell growth in the absence of sparging	Protects from sparging?
L-61	2,000	90	1–7	Cell lysis	N/A
L-121	4,400	90	1–7	Cell lysis	N/A
P-103	4,950	70	7–12	Cell lysis	N/A
P-123	5,750	70	7–12	Cell lysis	N/A
P-84	4,200	60	12–18	Cell lysis	N/A
P-104	5,900	60	12–18	Cell lysis	N/A
P-65	1,900	50	12–18	Inhibition	N/A
P-105	6,600	50	12–18	Inhibition	N/A
L-35	1,900	50	18–23	None	Yes
F-127	12,600	30	18–23	None	Yes
F-38	4,700	20	>24	None	Yes
F-68	8,400	20	>24	None	Yes
F-108	14,600	20	>24	None	Yes

Source: Adapted from Murhammer and Goochee (6).

ecules (4). Lower values of the HLB are representative of molecules with a higher degree of hydrophobicity. In general, molecules with high HLB values tend to be water soluble, and those with low HLB values tend to be oil soluble. The HLB values given in Table 1 are given in terms of ranges to indicate that the effective HLB of the Pluronic polyols varies with the emulsion system (5). From the results given in Table 1, it is clear that Pluronic polyols with low HLB values are unsuitable for use as protective agents because medium supplemented with 0.2% (w/v) of these Pluronic polyols either lysed *Spodoptera frugiperda* Sf-9 cells or significantly inhibited their growth in the absence of sparging, specifically in 50-mL spinner flasks oxygenation by surface diffusion (6). It was further demonstrated that at higher HLB values that Sf-9 cell growth was not inhibited in the absence of sparging and that protection from the adverse effects of sparging was provided in airlift bioreactor studies. Thus, the key to finding a Pluronic polyol that served as a protective agent for the Sf-9 cells in a sparged environment corresponded to finding one that did not inhibit cell growth in the absence of sparging. It should be noted that the results may be different with other cell lines.

Foaming Characteristics

Excessive foaming is generally undesirable in animal cell cultures. Therefore, if Pluronic polyols are used that exhibit unacceptable levels of foaming, then an antifoam should also be added to the medium (6). Of the Pluronic polyols given in Table 1, F-68, P-105, and P-65 are the best for foam generation (5). Therefore, a number of other Pluronic polyols may be preferable if foaming is problematic (e.g., L-35 or F-127).

Surface Tension

Because the Pluronic polyols are surfactants, it would be expected that they would reduce the air–water static surface tension. Indeed, according to BASF literature (5), the addition of 0.1% (w/v) Pluronic polyols to aqueous solutions reduces the static surface tension from 72 dyne/cm for pure water to a range of 33 (L-121) to 52 dyn/cm (F-38). In addition, the static surface tension of a 0.1% (w/v) aqueous solution of Pluronic F-68 is 50 dyn/cm. Michaels et al. (7), however, have given a convincing argument that the dynamic surface tension is a more suitable measure for sparged systems since interfacial properties are sensitive to the physical state of the interface. They found that the reduced dynamic surface tension afforded by Pluronic F-68 addition correlated well with reduced cell–bubble interactions and thereby provided cell protection.

Oxygen Transport

Zhang et al. (8) found that the addition of Pluronic F-68 to growth medium had minimal effect on $K_L a$ values for large bubbles (~5 mm diameter), but significantly enhanced $K_L a$ for intermediate (~1 mm diameter) and micron-sized (~100 μm diameter) bubbles. It is generally accepted that surfactants can affect oxygen transport from bubbles in two different ways: affecting the diffusional resistance and affecting the interfacial area. The first affects the K_L, and the second affects the a in $K_L a$. The addition of surfactants usually results in a decreased K_L (9) and an increased a (10). Consistent with this expected behavior, Zhang et al. (8) found that there was a significant reduction in the Sauter mean bubble size for the intermediate and micron-sized bubbles, thereby resulting in an increased a. Pluronic F-68 addition, however, had minimal effect on the size of the large bubbles. No direct measurements directly related to the K_L term were taken in this study, although the results clearly demonstrate that the presence of the Pluronic F-68 resulted in a greater increase in a than decrease in K_L.

PROTECTIVE EFFECTS OF PLURONIC POLYOLS

It has been clearly demonstrated that Pluronic polyols can protect insect and mammalian cells from the adverse effects of sparging. There is also some evidence, although not as conclusive, that Pluronic polyols can provide protection

from shear forces resulting from agitation. Evidence supporting these protective effects and the corresponding mechanisms are discussed in this section.

Sparging Damage

In 1968, Kilburn and Webb (11) were the first to report the use of Pluronic polyols to protect animal cells from sparging. Specifically, they found that supplementing the medium with 0.02% (w/v) Pluronic F-68 protected mouse LS cells from the adverse effects of sparging in short-term studies. They did not, however, show any results from long-term tests, such as an entire bioreactor run. No comprehensive studies regarding the use of Pluronic polyols to protect animal cells from the adverse effects of sparging were conducted until the late 1980s, when Handa-Corrigan et al. (12,13) and Murhammer and Goochee (6,14,15) demonstrated the ability of Pluronic F-68 to protect mammalian and insect cells, respectively. Handa-Corrigan et al. (12,13) demonstrated that the survival of a variety of mammalian cells (hybridomas, NS1 myeloma, BHK-21, and lymphoblastoid [RAJI]) in the presence of sparging depends on cell type and bubble size and frequency (i.e., small bubbles and high flow rates are more detrimental). They further demonstrated that the protective effect of Pluronic F-68 was concentration dependent, with 0.1% (w/v) being sufficient to protect cells under the most severe conditions reported. Murhammer and Goochee (14) investigated the protective effect of Pluronic F-68 in both airlift and sparged agitated bioreactors. It was found that supplementing the medium with 0.1% (w/v) Pluronic F-68 resulted in a major protective effect and that 0.2% (w/v) provided full protection to Sf-9 cells grown in both types of bioreactors. It was also demonstrated that recombinant protein synthesis (produced after infection of the cells with a recombinant baculovirus) in the airlift bioreactor was significantly reduced compared to the unsparged control even after increasing the Pluronic F-68 concentration to 0.5% (w/v). This was hypothesized to result from the increased susceptibility of virally infected cells to lysis in a sparged environment. Murhammer and Goochee (6) further demonstrated that many other Pluronic polyols could serve as protective agents in airlift bioreactors (Table 1).

After the development of conclusive evidence regarding the ability of Pluronic F-68 to protect mammalian and insect cell cultures from the adverse effects of sparging, the research focus in this field shifted to the mechanisms of this protective effect and of cell damage in the absence of protective agents. In regards to bubbles in sparged bioreactors, there are three regions in which cell damage could potentially occur: (*1*) the bubble formation region at the gas distributor, (*2*) the rising bubble region between the gas distributor and medium surface, and (*3*) the bursting bubble region at the medium surface. The results of Handa-Corrigan et al. (13) and Tramper et al. (16) clearly demonstrated that the primary source of cell damage is not in the region of rising bubbles. Specifically, they demonstrated that cell damage decreased as the height of the bioreactors increased. In addition, Michaels et al. (17) demonstrated that cell-to-bubble interactions in the bulk of the bioreactor are not responsible for cell damage in studies in which the headspace above the medium where bubbles can burst was eliminated. Murhammer and Goochee (15) demonstrated that it is possible to have cell damage in the sparger region and that this damage correlates with a high pressure drop across the sparger. In most practical applications (i.e., when the sparger is properly designed), however, significant cell damage only occurs at the medium surface in the bursting bubble region, as first proposed by Handa et al. (12). It has also been demonstrated that cells tend to adhere to air–medium interfaces and that fatal damage occurs to these attached cells as the bubbles burst at the medium surface, that is, the energy released by the bursting bubbles results in extensive cell lysis (18–21).

Michaels et al. (22) suggested that the protective effect of Pluronic polyols could result either from a biological mechanism (i.e., through changes in the cells' ability to resist shear) or from a physical mechanism (i.e., through changes in the medium properties that affect the level or frequency of forces experienced by cells). They concluded that the protective effect of Pluronic F-68 in sparged bioreactors is caused by a physical effect. Consistent with this hypothesis, it was later demonstrated that the protective effect of Pluronic F-68 resulted from its ability to cover the medium–bubble surface and thereby prevent cell adhesion to the bubbles (21,23). Thus, cells are not in the immediate vicinity of bursting bubbles at the medium surface.

It has also been demonstrated that Vero cells grown on microcarriers are damaged in sparged bioreactors (24). It was suggested that the addition of Pluronic F-68 to the medium may be able to protect cells grown on microcarriers from the adverse effects of sparging. Corresponding studies, however, were not conducted. This is certainly an issue that needs to be addressed in future research.

Agitation Damage

Although a few studies have reported that supplementing the medium with Pluronic F-68 can protect attachment-independent cells growing in suspension from agitation damage (25), this is generally not a relevant issue because insect (15) and mammalian (17) cells are commonly tolerant to relatively high agitation rates. For example, it was found that Sf-9 cell damage did not occur unless the agitation rate was high enough to induce bubble incorporation via cavitation or vortexing (15). Further, supplementing the medium with Pluronic F-68 protected the cells from damage caused by bubble incorporation by either of these mechanisms. It is assumed that the corresponding mechanisms of cell damage without Pluronic F-68 and protection by Pluronic F-68 are similar to those reported above for sparged bioreactors.

In spite of the lack of evidence demonstrating that Pluronic F-68 can protect cells from agitation damage, there is some evidence that Pluronic F-68 can make cells more resistant to shear damage. For example, Goldblum et al. (26) demonstrated that the addition of Pluronic F-68 (0.2 and 0.3% [w/v]) to the medium protected Sf-9 and *Tricho-*

plusia ni Tn-368 insect cells from shear damage in viscometer studies in a concentration-dependent manner. In contrast to these results, Michaels et al. (22) found that Pluronic F-68 did not make CRL-8018 cells more shear tolerant. These results suggest that the ability of Pluronic F-68 to increase a cell's shear tolerance may be cell dependent. Direct interaction with the cell (e.g., incorporation of the hydrophobic portion of Pluronic F-68 with the cell membrane) is a possible mechanism through which Pluronic F-68 may increase the shear tolerance of cells. Several lines of evidence are consistent with an interaction of Pluronic F-68 with the cell membrane. First, Pluronic polyols of low HLB values lysed cells (Table 1). Second, Pluronic F-68 inhibits the uptake of trypan blue dye that would normally be taken up in the absence of Pluronic F-68 (14). Third, treating sickled erythrocytes with Pluronic F-68 reduced their rigidity (27). Fourth, Pluronic F-68 increases the fluidity of hybridoma membranes, which correlated with increased shear sensitivity (28).

APPLICATION

If attachment-independent insect or mammalian cells are to be grown in sparged bioreactors, then it is recommended that the medium be supplemented with a Pluronic polyol. Many commercially available serum-free media already contain Pluronic F-68; therefore, additional supplementation may not be necessary. If supplementation is necessary, then a 10% (w/v) aqueous solution of Pluronic F-68 (obtained from BASF Corp.) can be prepared and filter sterilized. This solution can then be used to supplement the medium at 0.1 to 0.2% (w/v) Pluronic F-68. Ready-made solutions of Pluronic F-68 are also commercially available (e.g., Sigma Chemical Company, St. Louis, Mo.). The resulting medium should first be tested in the absence of sparging (e.g., in spinner or shaker flasks) to detect any potential inhibitory effects. Note that purification of the Pluronic F-68 (as mentioned above) may be necessary to remove inhibitory impurities. The next step is to test the ability of the medium to protect cells in small-scale sparged bioreactors. Success in protecting the cells at this scale is a good indicator of similar success in larger-scale sparged bioreactors. The use of Pluronic F-68 allows sparging with micron-sized bubbles that significantly enhance culture oxygenation. If Pluronic F-68 is not successful, many other Pluronic polyols are available for testing (Table 1). Another concern is that Pluronic polyols may interact with the protein of interest and interfere with downstream purification (29). Winzerling et al. (30) have demonstrated, however, that recombinant insect transferrin can be purified from Sf-9 cell supernatant containing 0.1% (w/v) Pluronic F-68 by using high metal ion capacity gel chromatography.

BIBLIOGRAPHY

1. I.R. Schmolka, *J. Am. Oil Chemists' Soc.* **54**, 110–116 (1977).
2. P.K. Bentley, K.C. Lowe, O.L. Johnson, S.K. Sharma, C. Washington, and S.S. Davis, *J. Pharm. Pharmacol.* **40** (suppl), 143P (1988).
3. P.K. Bentley, R.M.C. Gates, K.C. Lowe, D.I. de Pomerai, and J.A.L. Walker, *Biotechnol. Lett.* **11**, 111–114 (1989).
4. P. Becher and R.L. Birkmeier, *J. Am. Oil Chem. Soc.* **41**, 169–172 (1964).
5. BASF Publication, *Pluronic and Tetronic Surfactants*, BASF, Ludwigschafen, Germany, 1987.
6. D.W. Murhammer and C.F. Goochee, *Biotechnol. Prog.* **6**, 142–148 (1990).
7. J.D. Michaels, J.E. Nowak, A.K. Mallik, K. Koczo, D.T. Wasan, and E.T. Papoutsakis, *Biotechnol. Bioeng.* **47**, 420–430 (1995).
8. S. Zhang, A. Handa-Corrigan, and R.E. Spier, *Biotechnol. Bioeng.* **40**, 252–259 1992.
9. J.D. Sheppard and D.G. Cooper, *J. Chem. Tech. Biotechnol.* **48**, 325–336 (1990).
10. J.J. Heijnen and K. Van't Riet, *Chem. Eng. J.* **28**, B21–B42 (1984).
11. D.G. Kilburn and F.C. Webb, *Biotechnol. Bioeng.* **10**, 801–814 (1968).
12. A. Handa, A.N. Emery, and R.E. Spier, *Dev. Biol. Standard.* **66**, 241–253 (1987).
13. A. Handa-Corrigan, A.N. Emery, and R.E. Spier, *Enzyme Microb. Technol.* **11**, 230–235 (1989).
14. D.W. Murhammer and C.F. Goochee, *Bio / Technology* **6**, 1411–1418 (1988).
15. D.W. Murhammer and C.F. Goochee, *Biotechnol. Prog.* **6**, 391–397 (1990).
16. J. Tramper, D. Joustra, and J.M. Vlak, in C. Webb and F. Mavituna, eds., *Plant and Animal Cells*, Ellis Horwood Ltd., Chichester, England, 1987, pp. 125–136.
17. D. Michaels, A.K. Mallik, and E.T. Papoutsakis, *Biotechnol. Bioeng.* **51**, 399–409 (1996).
18. F. Bavarian, L.S. Fan, and J.J. Chalmers, *Biotechnol. Prog.* **7**, 140–150 (1991).
19. J.J. Chalmers and F. Bavarian, *Biotechnol. Prog.* **7**, 151–158 (1991).
20. M. Garcia-Briones and J.J. Chalmers, *Ann. N.Y. Acad. Sci.* **665**, 219–229 (1992).
21. D. Chattopadhyay, J.F. Rathman, J.J. Chalmers, *Biotechnol. Bioeng.* **45**, 473–480 (1995).
22. J.D. Michaels, J.F. Petersen, L.V. McIntire, and E.T. Papoutsakis, *Biotechnol. Bioeng.* **38**, 169–180 (1991).
23. M. Jordan, H. Sucker, F. Widmer, and H.M. Eppenberger, in R.E. Spier, J.B. Griffiths, and J.B. Berthold eds., *Animal Cell Technology: Products of Today, Prospects for Tomorrow*, Butterworth-Heinemann, Oxford, U.K., 1994, pp. 302–304.
24. D.E. Martens, E.A.A. Nollen, M. Hardeveld, C.A.M. van der Velden-de Groot, C.D. de Gooijer, E.C. Beuvery, and J. Tramper, *Cytotechnology* **21**, 45–59 (1996).
25. A. Mizrahi, *J. Clin. Microbiol.* **2**, 11–13 (1975).
26. S. Goldblum, Y.-K. Bae, W.F. Hink, and J. Chalmers, *Biotechnol. Prog.* **6**, 383–390 (1990).
27. C.M. Smith, R.P. Hebbel, D.P. Tukey, C.C. Clawson, J.G. White, and G.M. Vercellotti, *Blood* **69**, 1631–1636 (1987).
28. O.T. Ramirez and R. Mutharasan, *Biotechnol. Bioeng.* **36**, 911–920 (1990).
29. N. Kioukia, A.W. Nienow, M. Al-Rubeai, and A.N. Emery, *Biotechnol. Prog.* **12**, 779–785 (1996).
30. J.J. Winzerling, D.Q.-D. Pham, S. Kunz, P. Samaraweera, J.H. Law, and J. Porath, *Protein Exp. Purif.* **7**, 137–142 (1996).

POLY(3-HYDROXYALKANOATES)

BIRGIT KESSLER
BERNARD WITHOLT
Institute of Biotechnology
Zürich, Switzerland

KEY WORDS

Bacteria
Bioplastic
Biopolymer
Polyhydroxyalkanoates (PHA)
Polyhydroxybutyrate (PHB)

OUTLINE

INTRODUCTION AND HISTORY

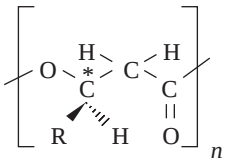
Plastics have been produced by the chemical industry since the early 1930s. They constitute a very important group of materials, because their high molecular weight and low reactivity make them especially suited for applications for which durable, inert materials are required. However, these very same properties have recently been recognized as major drawbacks of plastic materials. Whenever plastic is disposed of, it remains at the site of disposal, and its resistance to biological breakdown causes it to accumulate in the environment. The recognition of this undesired property has led to a search for new plastic materials that disappear after their disposal, the so-called biodegradable plastics. Several biological polymers that could fulfill this requirement are currently under investigation. Most of these biopolymers belong to the class of polysaccharides or polyhydroxyalkanoates (PHAs). In this review, we focus on the second class, the PHAs.

The first example of PHAs to be discovered was polyhydroxybutyrate (PHB). In 1926, Lemoigne isolated and characterized PHB from *Bacillus megaterium* (1). Since this first report, PHB accumulation has been found in many microorganisms: representatives of Gram-negative and Gram-positive species (i.e., autotrophs, heterotrophs, phototrophs, aerobes, anaerobes) and archaebacteria (as reviewed elsewhere [2–4]). Bacteria synthesize and accumulate PHAs as carbon and energy storage materials or as a sink for redundant reducing power under conditions of limiting nutrients in the presence of excess carbon source. When the supply of the limiting nutrient is restored, the PHA can be degraded by intracellular depolymerases and subsequently metabolized as a carbon and energy source.

In 1976, Imperial Chemical Industries (ICI Ltd., UK) recognized the potential applicability of PHB to replace some of the oil-derived commodity plastics. Although the bacterially produced PHB was relatively expensive compared to petrochemical plastics, it was expected that the price of crude oils would rise to high levels, making the production of PHB economically feasible. Because no rise in oil prices occurred, PHB came to be seen as a biodegradable and biocompatible plastic material. Mainly because of their high price, PHB and PHB/HV (a copolymer consisting of hydroxybutyrate and hydroxyvalerate) known under the tradename Biopol™ have only recently found their way to the market. In 1990, the German company Wella released a new shampoo packaged in bottles made of Biopol™. Early in 1996, Monsanto (USA) bought the Biopol™ business from ZENECA BioProducts, a daughter company of ICI. Monsanto is now in the process of generating transgenic plants (i.e., soybean and rapeseed) for future agricultural production of PHB and PHB/HV, which is expected to result in a significantly lower price of the final product.

The discovery of a polyester that consists mainly of hydroxyoctanoate monomers by de Smet et al. (5) in 1983 was the first example of a new group of polymers that can contain a wide variety of different monomers. Thus far, more than 90 different monomers have been found in the polymers (6). Among these are 3-hydroxy acids of 6 to 14 carbon atoms with saturated, unsaturated, straight, or branched aliphatic side chains. Furthermore, monomers with various different functional groups in the side chain, such as halogen atoms, hydroxy-, epoxy-, cyano-, carboxyl-, and esterified carboxyl groups, have been introduced into medium-chain-length (MCL) PHAs (4,6,7).

PHAs can be classified into two groups depending on the number of carbon atoms in the monomer units: short-chain-length (SCL) PHAs, which consist of 3–5 carbon atoms, and MCL PHAs, which consist of 6–14 carbon atoms (8). The general structure of PHAs is shown in Figure 1.



R = CH_3-C_2H_5 Short-chain-length PHAs
R = C_3H_7-$C_{14}H_{29}$ Medium-chain-length PHAs

Figure 1. General structure of polyhydroxyalkanoates (PHAs). PHAs are polyesters with (*R*)-3-hydroxy acids as monomer unit. The monomers of short-chain-length (SCL) PHAs consist of 3–5 carbon atoms, and those of medium-chain-length (MCL) PHAs consist of 6–14 carbon atoms. The molecular weights of these biopolymers range from 2×10^5 to 3×10^6, and the number of monomers (*n*) varies between 1,000 and 30,000.

The monomer units in these microbial polyesters are all in the *R*-configuration because of the stereospecificity of the biosynthetic enzymes. The molecular weights of PHAs range from 2×10^5 to 3×10^6, depending on the specific PHA, the microorganism from which the PHA was obtained, and the growth conditions. In this review, the term *PHA* refers to polyhydroxyalkanoates generally. PHB (or PHV) refers to specific SCL PHAs, and MCL PHA refers to all PHAs that contain monomers with six or more carbon atoms.

SCL PHA

PHB, the best-known of the SCL PHAs, not only appears in microorganisms as storage material but is also ubiquitous in nature. A fascinating development in recent years has been the discovery of the very wide distribution of PHB as a low molecular weight oligomer (120–200 monomers) in microorganisms, plants, and animals, including humans (9,10). In many cases this form of PHB is found as a PHB–calcium polyphosphate complex in membranes that seems to function as an ion channel through cell membranes (11,12). In *Escherichia coli*, the complex has been found to occur in large amounts in the membrane of cells made competent for genetic transformations (13–15). In addition, PHB of similar molecular weight has been found in eukaryotes, for example, in various plant tissues, in organs of animals (such as beef heart, chicken liver, porcine kidney), and in human blood plasma (16). In this review, we focus on the high molecular weight polymer that has considerable potential as an industrial biodegradable plastic.

Physiology and Biosynthesis of PHB

The biochemistry of PHB biosynthesis is very well studied. In most bacteria, such as *Alcaligenes eutrophus*, PHB is synthesized in a three-step reaction starting with acetyl-CoA when cultivated on carbohydrates, pyruvate, or acetate (Fig. 2) (17,18). Two acetyl-CoA molecules are coupled to form acetoacetyl-CoA in a condensation reaction catalyzed by 3-ketothiolase. The product is subsequently stereoselectively reduced to (*R*)-3-hydroxybutyryl-CoA in a re-

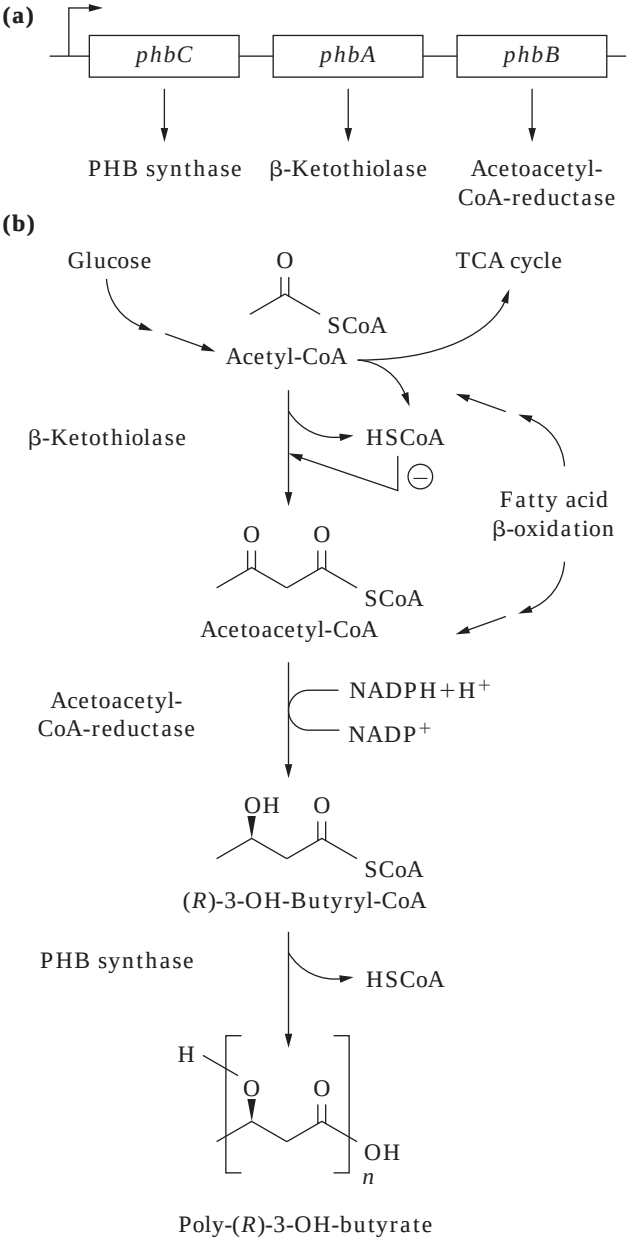


Figure 2. Polyhydroxybutyrate (PHB) synthetic genes and PHB synthesis pathway of *Alcaligenes eutrophus*. (**a**) The genes *phbA*, *phbB*, and *phbC* of *Alcaligenes eutrophus* are located in one operon and code for the *β*-ketothiolase, acetoacetyl-CoA-reductase, and PHB synthase, respectively. (**b**) In the three-step PHB synthesis pathway, two acetyl-CoA molecules are coupled to form acetoacetyl-CoA in a condensation reaction catalyzed by *β*-ketothiolase. The *β*-ketothiolase is negatively regulated by the product coenzymeA (HSCoA), which is also a product when acetyl-CoA enters the TCA cycle under nonlimited conditions. The product is subsequently stereoselectively reduced to (*R*)-3-hydroxybutyryl-CoA in a reaction catalyzed by NADPH-dependent acetoacetyl-CoA reductase. Finally, PHB is synthesized by polymerization of (*R*)-3-hydroxybutyryl-CoA molecules by the PHB synthase. *Source*: From van der Leij et al. (103).

action catalyzed by NADPH-dependent acetoacetyl-CoA reductase. Finally, PHB is synthesized by polymerization of *(R)*-3-hydroxybutyryl-CoA molecules by the PHB synthase (19). A few exceptions to this general pathway are known, such as *Rhodospirillum rubrum*, where (*S*)-3-hydroxybutyryl-CoA is generated and converted to *(R)*-3-hydroxybutyryl-CoA by two stereospecific enoyl-CoA hydratases (20).

As pointed out in the introduction, PHB biosynthesis and accumulation occur in most bacteria if a carbon source is provided in excess and if another nutrient, such as nitrogen, sulfur, phosphate, iron, magnesium, potassium, or oxygen, is limiting (21,22). PHB synthesis is regulated at the enzymatic level, and the availability of suitable substrates for PHB synthesis routes results from an imbalanced supply of nutrients to the cell via physiological regulation pathways than have been studied in *Azotobacter beijerinckii* (23) and *Alcaligenes eutrophus* (24,25). It was found that the intracellular concentration of acetyl-CoA and free coenzyme A play a central role in the regulation of polymer synthesis (26). Under balanced growth conditions acetyl-CoA is oxidized via the tricarboxylic acid (TCA) cycle. In this reaction sequence, NADH is generated and is used for biosynthetic purposes. When growth ceases the NADH concentration increases, which reduces the activity of the TCA cycle enzymes citrate synthase and isocitrate dehydrogenase. As a result acetyl-CoA cannot be oxidized via the TCA cycle and enters the PHB synthetic pathway. The 3-ketothiolase enzyme of this pathway is inhibited by free CoA, which is generated by oxidation of acetyl-CoA via the TCA cycle during normal growth.

PHB/HV Copolymer Biosynthesis

A copolymer of PHB and PHV can be synthesized by *Alcaligenes eutrophus* and many other microorganisms from either a mixed substrate of glucose and propionic acid or a direct precursor of 3-hydroxyvalerate (e.g., valeric acid) (27). If propionate is fed, essentially the same biochemical pathway as for PHB synthesis is used, but propionyl-CoA and acetyl-CoA are condensed by the ketothiolase to give 3-ketovaleryl-CoA, which leads to the incorporation of 3-hydroxyvalerate monomers in the polymer. 3-Hydroxybutyrate arises in the normal way. Alkanoic acids of odd-number carbon chain length can also serve as carbon source (28,29). In this case, the 3-hydroxyvalerate in the polymer arises directly from the β-oxidation of these fatty acids. Recently, it was shown that low dissolved oxygen concentrations enhance 3-hydroxyvalerate incorporation into the polymer produced by *A. eutrophus* (30).

Several microorganisms have been isolated that are capable of accumulating PHB/HV from substrates apparently unrelated to, and not immediate precursors of, 3-hydroxyvalerate. These organisms include *Corynebacterium*, *Nocardia*, and *Rhodococcus* strains (31) and mutants of *A. eutrophus* (32). When supplied with hexanoic acid as substrate, *Rhodococcus* sp. was able to accumulate a terpolymer containing 3-hydroxybutyrate, 3-hydroxyvalerate, and 3-hydroxyhexanoate monomers. In addition, *Rhodococcus* sp. NCIMB40126 and *Chromobacterium violaceum* are capable of producing a homopolymer consisting of 3-hydroxyvalerate when grown on valeric acid as sole carbon source (31,33).

Copolymers of 3-hydroxybutyrate and 4-hydroxybutyrate can be synthesized by *A. eutrophus* from 4-hydroxybutyric acid, 1-4-butanediol, butyrolactone, and 4-chlorobutyric acid (34). Terpolymers of 3-hydroxybutyrate, 3-hydroxyvalerate, and 5-hydroxyvalerate can be produced by *A. eutrophus* fed with either 5-chlorovalerate alone or 5-chlorovalerate or 5-hydroxyvalerate in combination with valeric acid (35). Polymers that contain 4-hydroxyvalerate and 3-hydroxypropionate have been synthesized by a variety of organisms, including *A. eutrophus* (36). The substrates used were 4-hydroxyvalerate, 4-valerolactone, and 3-hydroxypropionic acid. Most *Alcaligenes* strains tested and *Pseudomonas oxalaticus* accumulated a terpolymer of 3-hydroxybutyrate, 3-hydroxyvalerate, and 4-hydroxyvalerate from 4-valerolactone or 4-hydroxyvalerate as sole carbon source (36). *Alcaligenes eutrophus* and *Burkholderia cepacia* are able to synthesize terpolyesters that consist of 3-hydroxybutyric acid, 3-hydroxyvaleric acid, and 2-methyl-3-hydroxybutyric acid when tiglic acid is used as sole carbon source or in combination with gluconic acid (37).

Molecular Biology and Enzymology of SCL PHA Synthesis.

In *A. eutrophus*, the structural genes for PHA synthesis are organized in the *phbCAB* operon (see Fig. 2), coding for PHB synthase, β-ketothiolase, and NADPH-dependent acetoacetyl-CoA reductase, respectively (38–40). The transcription start of *phbC* was identified 307 base pairs upstream from the translation initiation codon by S1 nuclease protection assay, and a sequence that exhibits a striking homology to the *E. coli* σ^{70} promoter consensus sequence was found (39).

The PHB synthase structural genes have been cloned and sequenced from various bacteria (8). Comparison of the protein primary structures revealed that PHB synthases represent a group of highly homologous enzymes. The PHB synthase of *A. eutrophus* has been studied in detail in several laboratories. It is a large enzyme with identical subunits that exhibit a relative molecular mass of 63.9 kDa (38) and specificity for 3-hydroxyacyl-CoA of C_3 to C_5 units (41). Site-directed mutagenesis demonstrated that the cysteine$_{319}$ residue is required for PHB synthase activity in *A. eutrophus* (42). Furthermore, evidence has been obtained that the enzyme requires posttranslational modification by phosphopantetheine to provide a second thiol group per PHB synthase subunit (42), which could support the mechanistic model for the polymerization process that depends on the presence of two thiol groups (43).

A putative model for the polymerization process based on de novo fatty acid synthesis is presented in Figure 3. This model illustrates only one possible reaction mechanism, and other models are also possible. In an initiation or priming process (Fig. 3a), the 3-hydroxybutyric acid moiety from 3-hydroxybutyryl-CoA is transferred to the thiol group of the phosphopantotheine-modified serine, coupled with the release of coenzyme A. This 3-hydroxybutyryl moiety is then translocated to the thiol of the cysteine. It has been shown recently that acylation of cysteine$_{319}$ causes a shift of the monomeric form of the synthase to its dimeric form, and this shift is accompanied by an increase in its specific activity and a decrease in the lag

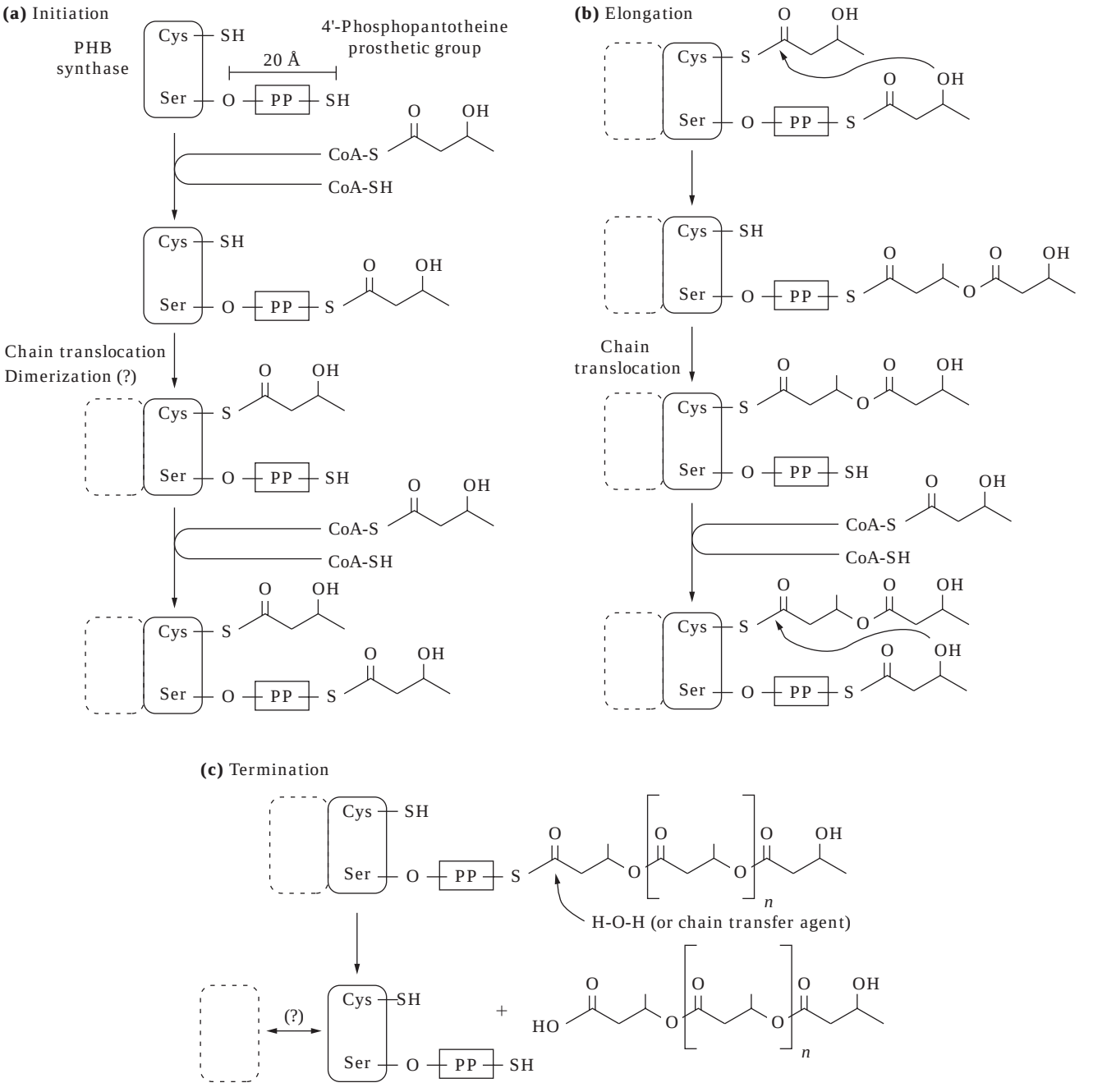


Figure 3. The hypothetical model of the catalytic PHB synthase mechanism is based on analogy to the de novo fatty acid synthesis and on information available about PHB synthases. The conserved cysteine$_{319}$ and the posttranslationally modified serine$_{259}$ of the PHB synthases that are likely to be involved in the reaction mechanism are indicated. In the initiation step (**a**) the 3-hydroxybutyric acid moiety from 3-hydroxybutyryl-CoA is transferred to the thiol group of the phosphopantotheine modified serine, coupled with the release of coenzyme A. It is supposed that this 3-hydroxybutyl moiety is translocated to the thiol group of the cysteine and that a shift of the monomeric form of the synthase to its dimeric form takes place. A second 3-hydroxyacyl moiety is added to the thiol group of the phosphopantotheine modified serine. During chain elongation (**b**) the 3′-hydroxy group of the entering monomer attacks the thiol ester of the previously entered monomer to form a covalently bound dimer/oligomer. Termination takes place when a chain transfer with water occurs, and the polymer chain is released from the enzyme (**c**). A chain transfer agent might be involved in this process. This model does not show that the polymer chain might be translocated from the posttranslationally modified serine to the thiol group of the cysteine of the other subunit as is the case in de novo fatty acid synthesis.

phase of polymer formation (44). A second monomer is then again added to the thiol group of the phosphopantotheine-modified serine, and its 3′-hydroxy group attacks the thiol ester of the first monomer to form a covalently bound dimer (Fig. 3b). After translocation of the formed dimer to the thiol group of the cysteine (either of the same subunit as shown in Fig. 3b or of the other enzyme subunit), the thiol group of the posttranslationally modified serine can accept a new 3-hydroxybutyric acid moiety, and by repeated propagation steps the covalently bound oligomer or polymer grows by one unit during each cycle. It is assumed that the polymer is released from the enzyme if a chain transfer with water occurs that terminates polyester biosynthesis (Fig. 3c). Doi and coworkers have postulated that a chain transfer agent, which might be an enzyme with a water molecule in its active site, could be involved in the termination process (45–47).

A combined chemical and enzymatic procedure has been developed to synthesize PHB granules in vitro (48). Purified PHB synthase from *A. eutrophus* was exposed to synthetically prepared *(R)*-3-hydroxybutyryl-CoA, and PHB was formed, thereby establishing the minimal requirements for PHB formation. The in vitro polymerization system yielded PHB with a molecular mass higher than 10×10^6 Da, exceeding by an order of magnitude the mass of PHB typically extracted from microorganisms. Furthermore, the molecular mass of the polymer could be controlled by the initial PHB synthase concentration.

The PHB synthases of *Chromatium vinosum* and *Thiocystis violacea* represent a different class of synthases that consists of two different types of subunits encoded by *phaE* and *phaC* (49,50). Together, these genes constitute one single operon. The C-terminus of PhaC contains high homology to the central region of the PHB synthase of *A. eutrophus*. However, the PhaC protein lacks the N-terminal part and is much smaller than the other synthases. The PhaE protein may take over essential functions of the larger PHB synthases in the polymerization process (8). The PHA synthase of *Thiocapsa pfennigii* is also composed of two different subunits, similar to those of *Chromatium vinosum* and *Thiocystis violacea*. From the substrate range, however, this synthase seems to be clearly distinguished from all other PHA synthases because it can synthesize copolyesters that consist of 3-hydroxyalkanoates ranging from C_3 to C_{14}, or even 4- and 5-hydroxyalkanoic acids (51).

Granule Structure

The PHA polymers are accumulated in the bacterial cells as defined granules (Fig. 4). Because the granule structure of SCL and MCL PHAs produced by different organisms seems to be very similar, both types of polymer granules are discussed together. Nuclear magnetic resonance (NMR) spectroscopy of various bacteria has clearly demonstrated that the polyester in the cells occurs in a metastable amorphous state and that it is stored as a mobile elastomer (52–54). The bouyant density of PHB granules is about 1.18–1.24 g cm^{-3} (55–57), in contrast to MCL PHA granules, which have a bouyant density of approximately 1.05 g cm^{-3} (58). The isolated granules consist of polyester,

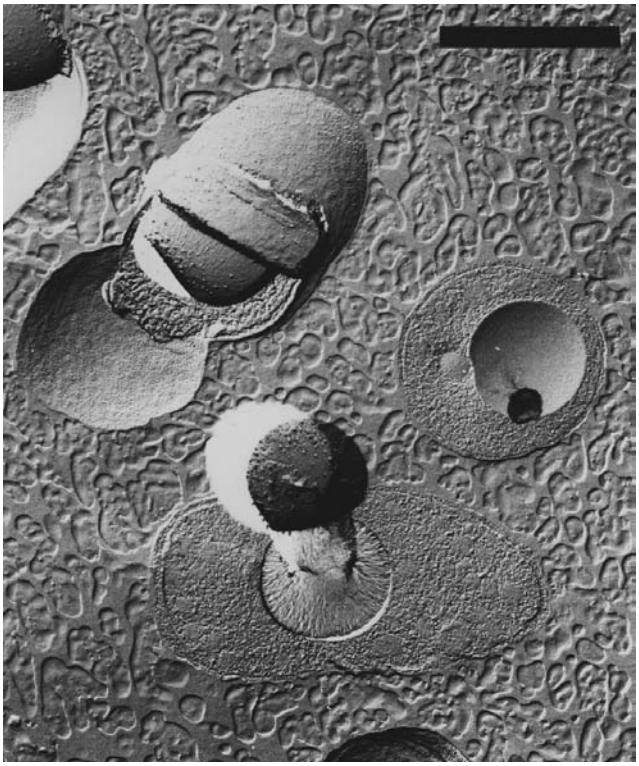


Figure 4. Polyhydroxyalkanoate (PHA) granules in *Pseudomonas oleovorans*. Freeze-fracture electron micrograph of octane-grown *P. oleovorans* cells containing PHA granules that were deformed during the fracture process and show the typical mushroom-type deformation of granules consisting of MCL PHA. The bar represents 0.5 μm.

proteins, and lipids. The composition was first determined for PHB granules of *Bacillus megaterium*, which consist of 97.7% polyester, 1.87% proteins, and 0.46% lipids or phospholipids (43).

Several different models of granule structure have been proposed. First indications that the native PHA granules in *Pseudomonas oleovorans* might be surrounded by a phospholipid monolayer in which the polymerase enzymes are embedded have been found by Witholt and coworkers (59,60). In contrast to this model, Stuart et al. (61) presented a model in which the hydrophobic PHA core is surrounded by two protein layers with a phospholipid layer in between. In this model, the polymerase is supposed to be embedded only in the inner layer. Recently, determination of the thickness of the boundary layer surrounding the polymer inclusions of different bacteria confirmed that a lipid monolayer is more likely than two protein layers with a lipid layer between them (62). In addition to the polymerases, other proteins such as depolymerases, granule-associated proteins (GAPs, also called phasins), and proteins of unknown function seem to be embedded in the phospholipid layer (63,64). Phasins are in general low molecular weight proteins and are assumed to form a close protein layer at the surface of the granules, providing the interface between the hydrophilic cytoplasm and the hydrophobic core of the PHA inclusion (63). In *A. eutrophus*,

the first evidence was found by mutant studies for the structural importance of a 24-kDa GAP (called GA24) (65). Mutants lacking the GA24 protein synthesized only one large PHB granule per cell in contrast to 10–30 granules as found in the wild type. The amino acid sequence of the GA24 protein revealed two closely related stretches consisting exclusively of nonhydrophilic amino acids at the C-terminal region, which are presumably involved in the binding of GA24 to the granules. Similar granule-forming proteins were identified and isolated from various other strains, such as *Acinetobacter* sp., *Rhodococcus ruber*, *Chromatium vinosum*, *Methylobacterium extorquens*, and *Pseudomonas oleovorans* (64,66–73).

Industrial Production of PHB and PHB/HV

Many bacteria have been screened to produce PHB or PHB/HV. However, only a few have been used for biotechnological production of these polyesters on a large scale. Here, we focus on the two examples that have found application in an industrial process.

The suitability of a bacterium for PHA production depends on many different factors such as stability and safety of the organism, growth and accumulation rates, achievable cell densities and PHA contents, extractability of the polymer, molecular weights of accumulated PHA, range of utilizable carbon sources, costs of the carbon source and the other components of the medium, and occurrence of by-products (74). Recently, the theoretical yield (the yield based on the reaction stoichiometry) of PHB has been estimated for several substrates (75). Regeneration of nicotinamide nucleotides, which are used as cofactors for PHA synthesis, has been taken into account in this analysis.

In the production plant of ZENECA BioProducts at Billingham (UK), the PHB homopolymer and the copolyester PHB/HV have been produced at scales up to 200,000 L (74). The process utilized a mutant of *A. eutrophus* (76), and the fermentation was carried out in a two-step fed-batch process. During the first step, the cells were grown in a mineral salt medium with glucose as sole carbon and energy source and with a calculated amount of phosphate based on the known requirements of the organism to allow production of a given amount of biomass. As the culture grew, phosphate was depleted from the medium, and during the second step when phosphate was limiting, the culture started to store polymer. Glucose was fed to the culture, and the fermentation was continued until the required polymer content was reached. Each phase lasted for approximately 48 h, and final biomass dry weights of 100 g/L were achieved (77). The copolymer of PHB/HV was made by substituting a mixed feed of glucose and propionic acid in the polymer accumulation phase. The 3-hydroxyvalerate content of the copolymer was controlled by adjusting the ratio of the two substrates in the feed (77).

A different industrial process has been developed at the Biotechnologische Forschungsgesellschaft in Linz (Austria) for the production of PHB homopolymer (78,79). In this process strain *Alcaligenes latus* DSM1124 was employed, which can accumulate PHB during balanced cell growth up to 80% of the cellular dry matter. Therefore, a one-step fed-batch fermentation process was carried out using a mineral salts medium with sucrose as sole carbon source. A biomass of 60 g/L was achieved (79). In 1993, however, the company stopped the production of PHB/HV.

Recovery

The cells that contain PHB/HV must be separated from the broth by conventional procedures such as centrifugation, filtration, or flocculation–centrifugation. After the biomass is harvested, cells must be disrupted to recover the polymer. Several methods have been developed for the recovery of PHAs. The method most often used involves extraction of the polymer from the biomass with solvents (e.g., chloroform, methylene chloride, propylene carbonate, dichloroethane). Large amounts of solvent are required in this process. Several other methods have been developed, such as sodium hypochloride treatments for the differential digestion of non-PHA cellular materials (80). Although this method is effective, it causes severe degradation of PHB, resulting in a reduction in the molecular weight (81).

ZENECA has developed a nonsolvent-based recovery process as an alternative to solvent extraction for the commercial production of PHB and PHB/HV by *A. eutrophus* (77,82). In this process the cells were first exposed to 80 °C and subsequently treated with a cocktail of various hydrolytic enzymes consisting of lysozyme, phospholipase, lecithinase, the proteinase alcalase, and others. Most of the cellular components were hydrolyzed by these enzymes, whereas the polymer remained intact. After washing, flocculation, and drying, the polymer was recovered as a white powder that was melted, extruded, and converted into chips, the form in which the polymer was supplied to the fabricators.

In contrast, the Biotechnologische Forschungsgesellschaft has applied a solvent-based process for the recovery of PHB from *A. latus*. The cells were harvested by centrifugation, and subsequently the PHB was extracted with methylene chloride from the suspended cells and precipitated out of the solvent by the addition of water. After drying of the polymer, a polyester with 99% purity was obtained (79). The process also included recovery of the solvent.

Properties of PHB and Its Copolymers

Comprehensive overviews on the physical properties of PHB and its copolymers, or on the material properties of the manufactured objects from those polymers, have been published previously (27,74,83–85) and are summarized next.

The PHB homopolymer is a completely stereoregular polyester, with all asymmetric carbon atoms in the (*R*)-configuration, and it is therefore highly crystalline. The high crystallinity (typically 55–80%) makes it relatively stiff and brittle. The glass transition temperature (T_g of PHB lies between 5 and 9 °C, and the melting point (T_m) lies between 173 and 180 °C (Table 1). PHB polymer decomposes at approximately 200 °C, which is close to its melting temperature. In its solid, crystalline state, the polymer is a right-handed helix, a condition that is retained in chloroform solution. The mechanical properties

Table 1. Comparison of Polymer Properties of Polyhydroxybutyrate and Its Copolymers with Polypropylene

Physical properties	P(3HB)	P(3HB/3HV) 90:10	P(3HB/3HV) 80:20	P(4HB)	P(3HB/4HB) 90:10	P(3HB/4HB) 10:90	PP
Melting point (°C)	179	150	135	53	159	50	170
Tensile strength (MPa)	40	25	20	104	24	65	34.5
Young's modulus (GPa)	3.5	1.2	0.8	149	—	100	1.7
Elongation to break (%)	3.0	20	100	1000	242	1080	400

Source: Data taken from Lee (2) and Byrom (74).
P(3HB), poly(3-hydroxybutyrate); P(3HB/3HV), poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate); P(4HB), poly(4-hydroxybutyrate); P(3HB/4HB), poly(3-hydroxybutyrate-*co*-4-hydroxybutyrate); PP, polypropylene.

Table 2. Enzymes Involved in the Pathways Leading to MCL PHA Formation and Their Respective Genes in *Escherichia coli*

Enzyme	Gene in *E. coli*
1. Acyl-CoA synthetase	*fadD*
2. Acyl-CoA dehydrogenase	*fadE*
3. Enoyl-CoA hydratase or crotonase	*fadB*
4. (*S*)-3-OH-Acyl-CoA dehydrogenase (NAD dependent)	*fadB*
5. 3-Ketoacyl thiolase	*fadA*
6. 3-ketoacyl synthase I; II; III	*fabB, F, H*
7. Ketoacyl-ACP reductase	*fabG*
8. 3-OH-Acyl-ACP dehydratase	*
8a. 3-OH-Decanoyl-ACP dehydratase	*fabA*
9. Enoyl-ACP reductase	*
10. Fatty acid thioesterase I; II	*tesA, B*
11. Ketoacyl-CoA reductase (NADPH dependent)	*
12. (*R*)-3-OH-Acyl-CoA dehydrogenase (NAD dependent)	*
13. 3-OH-Acyl-CoA epimerase	*fadB*
14. Enoyl-CoA hydratase II	*
15. Enoyl-CoA isomerase	*fadB*
16. (*R*)-3-OH-Acyl (ACP → CoA) transferase	*
17. Several acyl transferases or acylases	*
18. PHA polymerase	None

Source: From van der Leij et al. (103).
Note: Additional genes in *E. coli* have the following functions: *fadH* encodes a 2,4-dienoyl-CoA reductase; *fadL* encodes a protein that is needed for fatty acid transport across the membrane; and *fadR* codes for a regulatory protein.
*The respective gene is not known.

of PHB including Young's modulus (3.5 GPa) and the tensile strength (40 MPa) are similar to those of polypropylene (Table 1). However, the elongation to break for PHB is about 3%, which is significantly lower than that of polypropylene (400%).

The flexibility of the material is greatly improved when 3-hydroxyvalerate units are incorporated into the polymer, resulting in a decrease of Young's modulus below 0.7 and a decrease of tensile strength below 30 MPa. The elongation to break also increases as the comonomer fraction increases. The melt temperature is greatly depressed, down to 130 °C, dependent on the 3-hydroxyvalerate content (Table 1), but the degradation temperature is little affected. This property generates a relatively wide window of conditions that allow thermal processing as a melt without degradation of the material. PHB/HV copolymers with different 3-hydroxyvalerate content, ranging from 0 to 30%, are commercially available as Biopol™ resins.

Copolymers of 3-hydroxybutyrate and 4-hydroxybutyrate monomers cannot form isomorphic crystals because of the extent of the structural differences between the monomer units. The glass transition temperature decreases from 5 to −50 °C, and the melting temperature decreases from 180 to 54 °C as the mol% of 4-hydroxybutyrate increases from 0 to 100%. Young's modulus, tensile strength, and % elongation at break (for 94 mol% 4-hydroxybutyrate) are reported as 55 MPa, 39 MPa, and 500%, respectively. The T_g and T_m of copolymers consisting of 3-hydroxybutyrate and 3-hydroxypropionate decrease with an increase in the 3-hydroxypropionate (from 0 to 26 mol%) from 4 to −8.1 °C and 180 to 85 °C, respectively. The degree of crystallinity also decreases with an increase of 3-hydroxypropionate.

Recombinant Organisms

Successful biosynthesis of PHB in *E. coli* was independently achieved by two different groups (39,40). Polymer formation was only observed when all three PHB biosynthesis genes of *A. eutrophus* were expressed in *E. coli*; expression of *phbC* alone did not result in PHB production. These studies were extended by the production of the copolymer PHB/HV in recombinant *E. coli* (*fadR atoC*) mutants containing the PHB biosynthetic genes and using glucose and propionic acid as carbon source. The *E. coli* (*fadR atoC*) mutants produce the enzymes required for the catabolism of SCL fatty acid constitutively. 3-Hydroxyvalerate contributed up to 39 mol% to the accumulated polyester (86). Synthesis of PHB by recombinant *E. coli* does not require limitation of a specific nutrient, but is dependent on the amount of acetyl-CoA available and is significantly enhanced by the addition of various complex nitrogen sources, amino acids, or oleic acid (87). Application of recombinant *E. coli* in fed-batch culture resulted in cell densities of 101 g L^{-1}, containing approximately 80% PHB per cell dry weight (88) after 39 h of cultivation.

Under specific conditions, such as pH 6.0, *E. coli* produced PHB with an extremely high molecular mass of 20 MDa (47). Furthermore, it was shown that the amount of PHB synthase in recombinant *E. coli* plays a key role in controlling the molecular weight and the polydispersity of the polymer (89). With an increased amount of PHB biosynthetic enzymes, a decrease in polyester molecular weight was observed.

For the first time, PHB biosynthesis was successfully conferred to *Arabidopsis thaliana*, a small oil seed plant,

by introducing the *phbB* and *phbC* gene from *Alcaligenes eutrophus* (90). When the enzymes necessary to catalyze the synthesis of PHB from acetyl-CoA were produced in the cytoplasm, only small amounts of PHB were synthesized. Plants were viable but growth was very poor, possibly because of a reduction of acetyl-CoA needed for growth. Targeting of all three required proteins to the plastids of *Arabidopsis thaliana* resulted in normal growth, active enzymes, and PHB accumulation up to 14% of total cell dry weight (91). However, the molecular weight distribution of the plant-produced PHB was much broader than that of typical bacterial PHB (92).

The *phbC* gene of *Alcaligenes eutrophus* has also been introduced into *Saccharomyces cerevisiae*, which enabled this yeast to accumulate PHB to up to 0.5% of cell dry weight (93). A novel pathway for PHB synthesis has been engineered in insect cells (*Spodoptera frugiperda* (94,95). In these insect cells, a heterologous dehydrase-deficient fatty acid synthase from rats is produced in addition to the PHB synthase of *A. eutrophus*. The multifunctional fatty acid synthase effectively replaces the *β*-ketothiolase and acetoacetyl-CoA reductase activities of the natural prokaryotic PHB pathway and, therefore, provides the precursor 3-hydroxybutyryl-CoA for conversion to PHB (Fig. 5). This strategy suggests a new method for generating structurally diverse polyhydroxyalkanoates by metabolic engineering.

MCL PHA

Although many species of bacteria are reported to synthesize PHB and PHV, including some *Pseudomonas* strains, the ability to store polymers that contain monomers of 3-hydroxy acids of carbon chain length C_6 and longer, here referred to as MCL PHAs, is unique to the group of pseudomonads.

Physiology and Biosynthesis of PHA

Three main pathways are involved in the synthesis of PHA precursors in *Pseudomonas putida* (96,97). These were revealed by labeling studies, NMR spectroscopy, and gas chromatography–mass spectroscopy (GC-MS) analyses of PHA from cells that were grown on simple carbon substrates such as glucose and/or fatty acids. *P. putida* derives PHA from precursors made via the following routes. *(1)* Chain elongation, in which acyl-CoA is extended with acetyl-CoA, is comparable with the major route for PHB monomer synthesis, where the acyl-CoA is one of the two acetyl-CoA molecules (Fig. 6). It has been found that about 3–10% of the monomeric units of PHA are one C_2 unit longer than the fatty acid used as substrate (98–100). However, this route is less important for PHA synthesis than it is for PHB synthesis. *(2)* Fatty acid degradation by *β*-oxidation is the main pathway when fatty acids are used

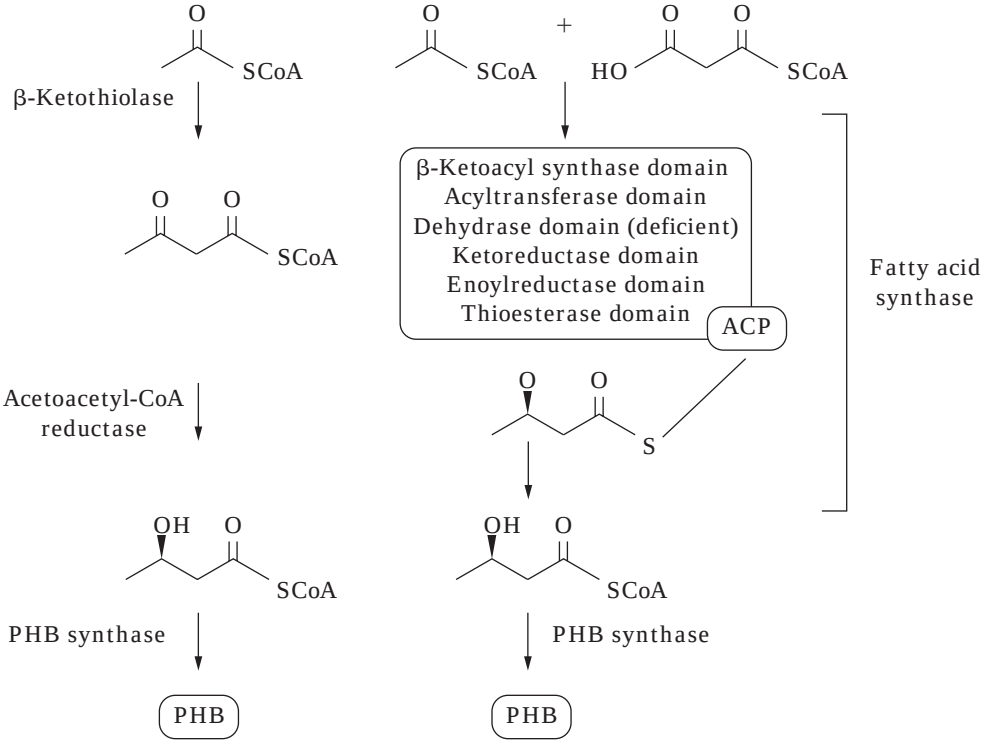


Figure 5. Engineered polyhydroxybutyrate (PHB) synthesis pathway in insect cells. The bacterial PHB synthesis pathway described in Figure 2 is shown on the left, and the engineered PHB synthesis pathway in insect cells on the right. The condensation of acetyl-CoA and malonyl-CoA followed by ketoreduction of acetoacetyl-acyl carrier protein (ACP) are performed by the fatty acid synthase multifunctional protein. The different catalytic domains of the enzyme are listed. The dehydrase domain is not functional. *Source*: From data of Williams et al. (94).

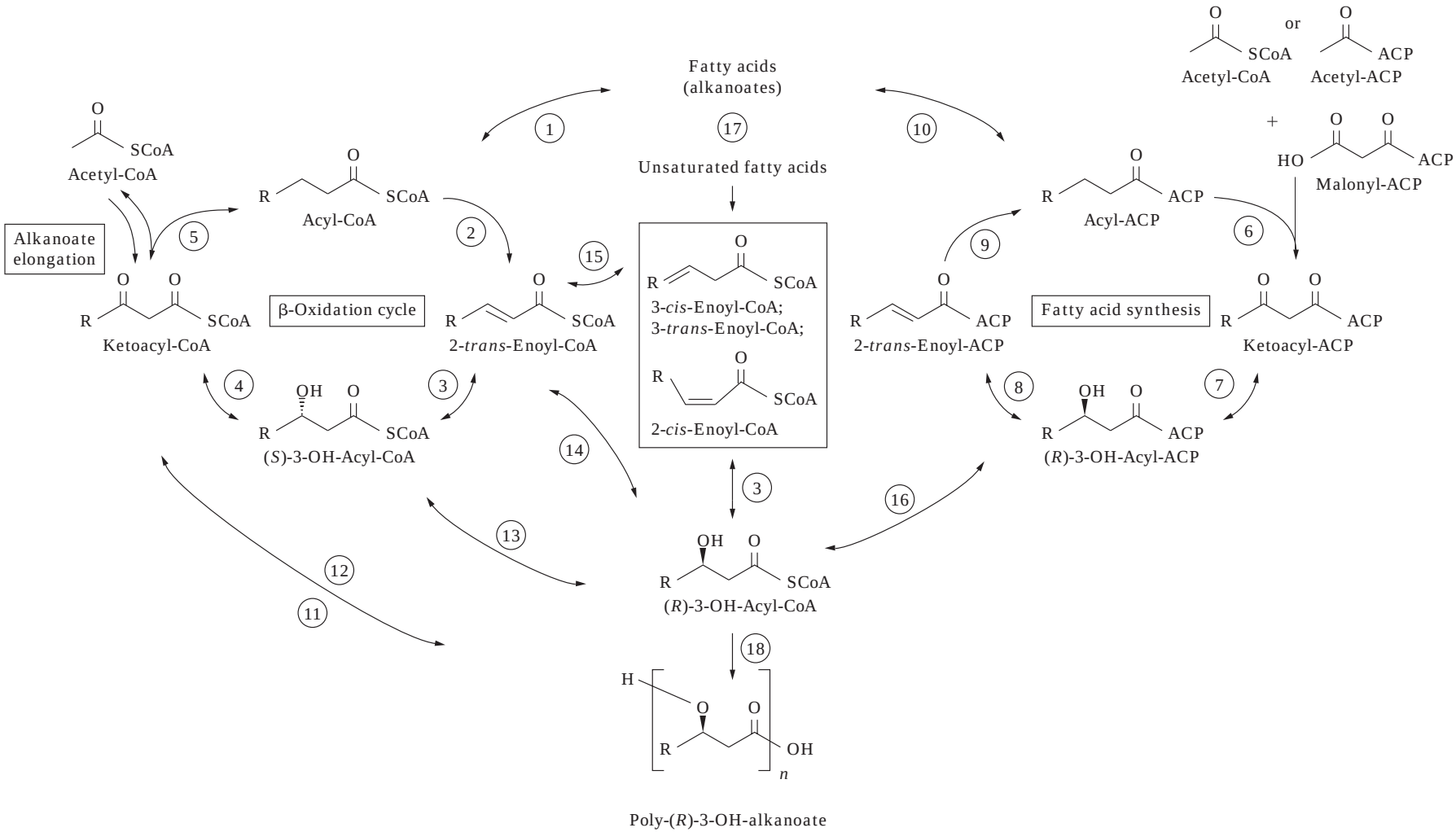


Figure 6. Major pathways involved in MCL polyhydroxyalkanoate (PHA) synthesis in *Pseudomonas*. Alkanoate elongation, *β*-oxidation, and fatty acid biosynthesis are involved in MCL PHA synthesis in *Pseudomonas*. The numbering of the enzymatic activities is as given in Table 2. *Source*: From van der Leij et al. (103).

as substrate. *(3)* De novo fatty acid biosynthesis is the main route during growth on simple carbon compounds. The enzymes involved in the pathways depicted in Figure 6 do not have the same activity for different chain length substrates. Thus, the PHA polymerase seems to have a preference for the incorporation of C_8 to C_{10} monomers. Such preferences might have substantial influences on the fluxes and the contributions of each pathway.

The intermediate precursors of PHA are either ketoacyl-CoA, (*S*)-3-hydroxyacyl-CoA, enoyl-CoA, or (*R*)-3-hydroxyacyl-acyl carrier protein (ACP). Regarding the similarity between the PHB polymerase and the PHA polymerases (see following), it seems likely that the ultimate substrate for polymerization is the (*R*)-form of the CoA-activated 3-hydroxy fatty acid intermediates. Because the monomeric units in PHA, like those in PHB, are in the (*R*)-form, and the PHB and PHA polymerases are homologous, the intermediates in fatty acid metabolism must presumably be converted to the (*R*)-form of 3-hydroxyacyl-CoA before transesterification (Fig. 6). The PHA precursors are generally considered to be available in a much lower in situ concentration than the PHB precursors because they are intermediates of ongoing fatty acid metabolic cycles.

Molecular Biology and Enzymology

Knowledge of the genetic and physiological regulation of PHA biosynthesis is relatively limited compared to what is known about PHB biosynthesis because PHA synthesis has only been studied since 1983 (5). The *pha* loci of *Pseudomonas oleovorans* (101) and *Pseudomonas aeruginosa* (102) consist of four ORFs: the ORFs *phaC1* and *phaC2* code for polymerases and are separated by an ORF called *phaZ*, which encodes a depolymerase; downstream of these genes is an ORF called *phaD*, which encodes a putative granule surface protein (103) (Fig. 7). Subsequently, more ORFs have been identified downstream and in opposite orientation to the *pha* locus of *P. putida* and *P. oleovorans*, which presumably encode granule-associated proteins (H. Valentin, unpublished results, 1996). The sequence of this region has not yet been published.

At least two promoters are involved in the transcription of the *pha* genes in *P. aeruginosa* (102). Two promoters are upstream of the *phaC1* gene and resemble the consensus sequences for σ^{54}-dependent and σ^{70}-dependent promoters. It is not clear whether a third promoter upstream of the *phaC2* gene in *P. aeruginosa* is active (102). Northern blot experiments indicated the presence of two mRNAs, one carrying *phaC1* and *phaZ*, the other *phaC2* and *phaD*. However, a transcriptional start site accompanying the putative promoter upstream of *phaC2* could not be experimentally shown, suggesting that this promoter may not be functional.

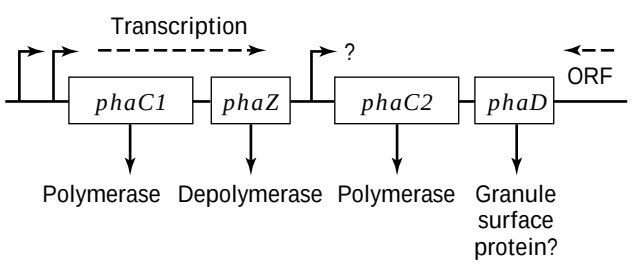


Figure 7. Organization of *pha* genes in *P. oleovorans* and *P. aeroginosa*. Two polymerases and one depolymerase are encoded by *phaC1*, *phaC2*, and *phaZ*, respectively. The *phaD* gene may encode a granule-associated protein. More open reading frames (ORF) in opposite orientation were found downstream of these genes. Putative promoters are indicated with open arrows (see text for details). *Source*: From Steinbüchel et al. (8) and van der Leij et al. (103).

The two PHA polymerases show identities of 54–59% at the protein sequence level when compared within each species. When corresponding PHA polymerases from different species are compared, they show 69–80% identity at the amino acid level. The identities between these PHA polymerases and the *A. eutrophus* PHB polymerase are nearly 40% (101,102). According to the gene sequences, the *Pseudomonas* PHA polymerases have molecular weights of 62–64 kDa. The enzymes show a small difference in substrate specificity (104).

The *phaZ* gene product, a PHA depolymerase with a predicted molecular weight of 31.5 kDa, shows homology with a range of hydrolytic enzymes (59,101). It contains the common lipase box Gly-X-Ser-X-Gly, which is also found in extracellular PHA depolymerases (105).

Production of MCL PHA

Pseudomonas strains such as *P. oleovorans, P. putida, P. aeroginosa*, and *P. resinovorans* are the major strains used for MCL PHA production (100,106,107). Batch, fed-batch, and continuous cultivation techniques have been extensively investigated at the laboratory scale (60,108,109), but MCL PHA production up to a 200-L scale has also been carried out (110). Efficient PHA production requires rapid formation of biomass and product. In batch and fed-batch fermentations, similar to the PHB production process, cells are first grown to a desired concentration without nutrient limitation, after which an essential nutrient is limited to trigger efficient PHA synthesis. During this nutrient limitation stage the residual cell concentration (defined as the cell concentration minus the PHA concentration) remains almost constant, and the cell concentration increases only because of the intracellular accumulation of PHA.

P. oleovorans has been used to produce MCL PHAs in two-liquid-phase cultivations (111,112). The organism was cultured in media consisting of an aqueous phase containing all minerals necessary for growth and a second liquid phase containing *n*-alkanes as sole carbon and energy source. Two-liquid-phase, single-stage chemostat, and fed-batch cultivations of *P. oleovorans* using *n*-octane as the carbon source resulted in highest volumetric productivities of 0.58 g L^{-1} h^{-1} (108) and 0.25 g L^{-1} h^{-1} (109), respectively. In a two-liquid-phase, two-stage continuous cultivation a volumetric productivity of 1.06 g L^{-1} h^{-1} with a final cellular PHA content of 66% per total cell dry weight has recently been reported (113). The monomer composition of the synthesized polymer is mainly dependent on the strain used and the carbon source applied. The cultivation conditions have only a minor influence on the composition of the polymer. However, it was shown that temperature

and concentration of the substrate has an influence on the molecular weight of the polymer (45). The monomer composition of PHAs produced from different alkanoates using *P. oleovorans* is shown in Table 3 (100). Recently, it was reported that *Pseudomonas putida* U grown on octanoic acid as the sole carbon source accumulated a homopolymer of poly-3-hydroxyoctanoate (114).

Different *Pseudomonas* strains have been used to produce MCL PHAs from renewable resources, such as fatty acids (29,100) and carbohydrates (96,107,115). In fed-batch cultivations, *P. putida* KT2442 was cultivated to high cell densities (92 g/L) with oleic acid as a substrate. With this process, an overall volumetric productivity of 1.6 $g\ L^{-1}\ h^{-1}$ was achieved with a final cellular PHA content of 40% per total cell dry weight (116). Furthermore, MCL PHAs could be produced from glucose, fructose, and glycerol using *P. putida* KT2442 (96). In addition to the major constituent 3-hydroxydecanoate, six other monomers were found to be present: 3-hydroxyhexanoate, 3-hydroxyoctanoate, 3-hydroxydodecanoate, 3-hydroxydodecenoate, 3-hydroxytetradecanoate, and 3-hydroxytetradecenoate (Table 4). The degree of unsaturation of PHA was slightly influenced by the cultivation temperature (96).

Various functional groups can be introduced into the polymer by applying different substrates in the growth medium. Olefin groups were introduced by adding *n*-alkenes (111,112) or alkenoates (29,117,118) to the medium. Furthermore, different halogen atoms were incorporated in the polymer when halogenated alkanoates were used as substrates (119–122). A homopolymer of 99% purity consisting of 3-hydroxy-5-phenylpentanoate is synthesized by *P. oleovorans* when grown on 5-phenylvalerate (123). Polymers containing different mol% of phenyl groups (124,125), or other aromatic compounds such as phenoxy groups (126,127) or cyanophenoxy groups (128,129), were produced by adding phenyl alkanoates, phenoxy alkanoates, or cyanophenoxy alkanoates, respectively, as a cocarbon source.

Table 3. Composition of MCL Polyhydroxyalkanoates Produced by *Pseudomonas oleovorans* Grown on Different Alkonoates as Sole Carbon Source

	Composition (mol% of total 3-hydroxy fatty acids)						
Carbon source	C6	C7	C8	C9	C10	C11	C12
Hexanoate	95		5				
Heptanoate		100					
Octanoate	8		91		1		
Nonanoate		35		65			
Decanoate	8		75		17		
Undecanoate		28		59		13	
Dodecanoate	6		57		32		5
Tridecanoate		32		48	5	14	
Tetradecanoate	7		59		30		4
Pentadecanoate		32		47	8	13	
Hexadecanoate	8		50		30		12

Source: Huisman et al. (100).
HC, 3-hydroxyhexanoate; HH, 3-hydroxyheptanoate; HO, 3-hydroxyoctanoate; HN, 3-hydroxynonanoate; HD, 3-hydroxydecanoate; HUD, 3-hydroxyundecanoate; HDD, 3-hydroxydodecanoate.

Table 4. Influence of Cultivation Temperature on MCL Polyhydroxyalkanoate Composition in *Pseudomonas putida*

	T	%	Composition (mol% of total PHA)						
Substrate	(°C)	PHA	C6	C8	C10	C12:1	C12	C14:1	C14
Glucose	15	20.5	0.4	8.4	64.8	16.7	6.0	4.0	<0.1
	30	17.0	<0.1	6.9	74.3	8.8	7.7	1.6	<0.1
Decanoate	15	48.5	5.9	50.9	43.2	—	—	—	—
	30	38.1	5.3	52.3	42.4	—	—	—	—

Source: Huijberts et al. (96).
C6, 3-hydroxyhexanoate; C8, 3-hydroxyoctanoate; C10, 3-hydroxydecanoate; C12:1, 3-hydroxy-5-*cis*-dodecenoate; C12, 3-hydroxydodecanoate; C14:1, 3-hydroxy-7-*cis*-tetradecenoate; C14, 3-hydroxytetradecanoate; —, not detectable.

Recently, it was shown that MCL PHAs can be produced by specific recombinant *E. coli* strains that are deficient in fatty acid *β*-oxidation and in addition contain the PHA polymerase-encoding gene of *P. aeroginosa* or *P. oleovorans* (130,131). The recombinant strains were cultured on LB medium or mineral salt medium supplemented with yeast extract and alkanoates as additional carbon source. Up to 21% PHA per cell dry weight could be achieved by these recombinant *E. coli* strains.

Recovery

Recovery procedures for MCL PHA resemble those originally developed for the production of PHB. A number of solvent extraction processes have been assessed to separate MCL PHAs from biomass. These usually involve the use of a chlorinated solvent such as chloroform (111) or methylene chloride. An alternative, nonsolvent-based extraction process was developed as well to make the overall production process more attractive (132,133). The biomass is separated from the medium by centrifugation, and treated with a protease cocktail and a detergent to solubilize all cell components. Removal of the solubilized cell material and concentration of the resulting PHA suspension is achieved by cross-flow microfiltration. The submicrometer MCL PHA granules display a density close to that of water (58), as a result of which a MCL PHA suspension does not settle (134); in fact, it forms a highly stable polymer latex. The overall purity of the latex amounts to 95%. MCL PHAs produced for applications that require higher purity, like those in biomedicine, must be further purified by extraction of PHA from the latex using methylene chloride and subsequent precipitation in a methanol–water mixture.

The applicability of MCL PHA as a latex product has been evaluated by Marchessault et al. (135). A typical latex product would be a coating. Unlike that reported for PHB and its derivatives (135), additional treatments (e.g., heat, solvent vapor) to enhance coalescence into a film are not necessary, because MCL PHA does not readily crystallize. Heat, UV, or electron beam sources have been applied to cross-link polymer chains and fix the film. Mcl-PHA coat-

ings are transparent with good clarity and exhibit excellent barrier properties to water.

Properties of MCL PHAs

Owing to their enzymatic synthesis, as is the case for PHB, MCL PHAs have an exceptional stereochemical regularity and therefore are fully isotactic. This allows MCL PHAs to achieve some crystallinity, which can be as high as 25% (136). The molecular weight is of the order of 2×10^5 g/mol (Table 5) (45,112). The glass transition temperature is usually well below room temperature, ranging from −43 °C to −25 °C for MCL PHAs (112). Melting temperatures depend on both the thermal history and the nature of the pendant chain. Values reported vary between 39 °C and 61 °C with MCL PHA, based on *n*-octane or *n*-octanoate (PHO) (112,136,137). The glass transition temperature of PHO is −35 °C (112,138,139). Depending on the thermal history and the test conditions, the tensile modulus varies between 2.5 and 9 MPa, the tensile strength between 6 and 10 MPa, and the maximum elongation between 300 and 450% (138,139). These data classify PHO as a thermoplastic elastomer (136).

Because the low melting temperature and crystallization rate hamper the applicability of MCL PHAs as thermoplasts, cross-linking was anticipated to overcome both limitations (138). Cross-linking was achieved by conventional techniques, namely electron-beam irradiation (138), peroxide cross-linking (140), or vulcanization (141). In all three studies, the presence of olefin groups was shown to enhance chemical cross-linking.

Sufficient cross-linking appeared to prevent polymer crystallization, so a true rubber was produced. Properties were therefore constant over a wide temperature range from the glass transition temperature (T_g, − 30 °C) up to the decomposition temperature (∼180 °C) (138). The rubber modulus could be controlled up to 10 MPa by varying the cross-link density (138). Similar to PHO, maximum elongations amounted to more than 300% for the irradiated material, but other cross-linking techniques produced lower values. Peroxide cross-linking even resulted in a loss of material integrity (140). The tensile strength of the cross-linked material never exceeded 1 MPa and tear resistance was poor (140,141).

BIODEGRADATION

PHAs can be degraded either intracellularly by intracellular depolymerases of the accumulating strain, to supply accumulated carbon and energy in the absence of a suitable exogenous sources (43), or extracellularly by extracellular depolymerases for consumption of oligomers or monomers by fungi, yeast, and bacteria. Relatively little is known about intracellular degradation of PHAs. It was shown that intracellular PHA depolymerases are unable to hydrolyze extracellular PHAs, and extracellular depolymerases cannot hydrolyze intracellular granules (142), apparently as the result of differences in the physical structures of the completely amorphous intracellular granules and the partially crystallized extracellular PHA. Under sterile or aseptic conditions, PHA is degraded by a hydrolytic mechanism, especially at high pH and high temperature (143,144). This type of degradation is important for medical applications, such as the use of PHA in drug release carriers or surgical sutures.

Because PHAs are so abundant, a wide range of bacteria and fungi have developed the capability to utilize PHAs (145). It has been shown that PHA degradation occurs in a large variety of complex ecosystems, including oxic and anoxic environments such as soil, seawater, or sludge. Microbiological parameters of the ecosystem, like population density, microbial diversity, microbial activity, and spatial distribution of microorganisms, can influence the degradation of the polymers. Furthermore, physicochemical parameters of the ecosystem, such as temperature, pH, water content, oxygen content, redox potential, and nutrient supply, can all affect biodegradation. The polymer itself influences biodegradability as well. The most important factors are (*1*) the stereoregularity, (*2*) the crystallinity, (*3*) the molecular mass, and (*4*) the monomeric composition of the polymer. Finally, material processing can also affect the extent of biodegradability.

The effect of different environments on the degradation rate of PHB and PHB/HV has been studied by different groups (146–154). It was shown that the polymer (1-mm molding) was completely degraded after 6 weeks in anaerobic sewage, after 75 weeks in soil, and after 350 weeks in seawater. Concerning functional MCL PHAs, it could be shown that even after cross-linking of unsaturated side chains the polymer remained biodegradable (138). However, it should be noted that at higher cross-link densities the rate of degradation was reduced.

Eleven groups of PHA-degrading bacteria can be distinguished by their substrate (PHA) specificities (155). In addition, 95 genera of fungi capable of degrading PHAs have been identified (155). The synthesis of PHA depolymerases is highly regulated. Most PHA-degrading bacteria repress depolymerase synthesis in the presence of a soluble carbon source that permits high growth rates. However, after exhaustion of the nutrients, the synthesis of depolymerase is derepressed (156,157).

Table 5. Physical Characteristics of Saturated and Unsaturated MCL Polyhydroxyalkanoates Isolated from *Pseudomonas* Strains

Strain	Substrate	Mw	Mn	Mw/Mn	T_g (°C)	T_m (°C)	Reference
P. oleovorans	*n*-Octane	178,000	99,000	1.8	−36.5	58.5	112
P. oleovorans	*n*-Octene	242,000	101,000	2.4	−36.6	—	112
P. putida KT2442	Oleic acid	140,000	70,000	1.9	ND	ND	96
Pseudomonas sp. NCIMB40135	Glucose	143,000	61,900	2.32	<−60	54.8	115

Note: Mw, weight-average molecular weight; Mn, absolute number of average molecular weight; ND, not determined.

The extracellular PHA depolymerases of many bacteria have been purified and characterized. Depending on the depolymerase, the hydrolysis products are monomers [e.g., *Comamonas* sp. (156)] or monomers and dimers [e.g., *Pseudomonas fluorescens* (157)] or a mixture of oligomers (monomers to trimers) as in the case of *Alcaligenes faecalis* and *Pseudomonas lemoignei* depolymerases (158). In a second step, the oligomers are hydrolyzed to monomers by oligomer hydrolases (159,160). All PHA depolymerases analyzed so far are specific for polymers that consist of monomers in the (*R*)-configuration. Chemically synthesized poly-(*S*)-hydroxybutyrate is not degraded by PHB depolymerases (161).

Nine bacterial PHA depolymerase genes that encode depolymerases with specificities for SCL PHAs have been cloned and analyzed (Ref. 155 and references cited therein). The enzymes contain a signal peptide (which is cleaved off during passage across the plasma membrane), a large catalytic domain at the N-terminus, a C-terminal substrate-binding domain, and a linking region in between. Three strictly conserved amino acids, namely serine, aspartate, and histidine, constitute the active center of the catalytic domain. The serine is part of the lipase-box motif Gly-X-Ser-X-Gly, which has been found in all known serine hydrolases such as lipases, esterases, amd serine proteases (162). The oxygen atom of the serine side chain is the nucleophile that attacks the ester bond and is supported by the imidazole ring of the histidine. The positive charge of the latter is stabilized by the carboxyl group of the aspartate.

Only one gene that encodes an extracellular MCL PHA depolymerase has been cloned and analyzed so far (157,163). In contrast to most SCL PHA depolymerases, the MCL PHA depolymerase of *P. fluorescens* GK13 consists of two identical subunits, and except for small regions adjacent to the lipase-box motif there are no significant homologies to SCL PHA depolymerases.

APPLICATIONS

PHAs display an unique combination of features. First, in contrast to synthetic polymers, the natural PHAs have the fundamental advantage of being renewable resources not dependent on the supply of petroleum. Second, as described in the previous section, they are genuinely biodegradable, that is, they can be completely metabolized into harmless, naturally occurring molecules. Although complete degradation can easily be achieved, the shelf life of the PHA products is virtually unlimited. Colonization of the hydrophobic surface is an arduous task for the degrading organisms, and thus a high biological activity is a prerequisite for biodegradation. Cheaper biodegradables, such as the starch-based plastics, are available however. Therefore, the most important feature of these polyesters may well be their hydrophobicity, which allows them to surpass carbohydrate-based biodegradable competitors in moisture resistance.

Applications of PHB

As mentioned previously, PHB and its derivates are all highly crystalline materials and therefore possess a high modulus, making them suitable for polypropylene-type applications including injection molding, extrusion blow molding, and fiber-spray gun molding. The polymers can be manufactured into many different materials with a wide range of applications (164). The first commercial application of PHB or PHB/HV, sold under the trade name Biopol™, was in packaging for personal hygiene products, for example, as biodegradable shampoo bottles. Later applications included packing containers, bottles, tubes, wrappings, bags, tubes, and disposable razors. In these cases the rather costly Biopol™ was used for its environmentally friendly image to increase product sales. In addition, applications that require biodegradability for functional reasons were developed. Among those, garbage bags, which should ideally be composted together with their contents, fishing nets, which can cause severe damage once they are lost in the marine environment, and agricultural delivery devices for nutrients, fertilizers, herbicides, and insecticides are important examples.

Applications of MCL PHA

No applications for MCL PHAs have been reported yet. Complementing the intrinsically rigid and brittle properties of PHB and its derivates, MCL PHAs are expected to support applications that require flexibility and elastomeric properties. MCL PHAs will have very limited applicability as thermoplasts. Amorphous MCL PHAs, cross-linked and non-cross-linked, are more likely to penetrate the plastics market (110). The capability of pseudomonads to incorporate many different monomer units into MCL PHAs suggests various possibilities to produce tailor-made polymers. Incorporation of unsaturated pendant groups to enhance cross-linking is one example (138). A recent study reported that nitrophenoxy groups have been introduced, aiming to create chiral polymers for nonlinear optical applications (128). Specific functionalities for surface modification could be introduced as well. In the long run, it should be feasible to custom design new polymers and then program the bacteria to produce them.

PHA Application in Blends and Composites

Polymer blending offers interesting possibilities of preparing inexpensive (biodegradable) materials with useful mechanical properties. In this context, PHAs could form an important source of biodegradable additives. Several blending studies have been performed on PHB or PHB/HV copolymers with other dissimilar compounds. For example, the miscibility and compatibility of these polyesters with chlorinated polymers, polyamides, polycarbonates, cellulose derivatives, and other functional polymers are well documented (85,165). At present, applications of PHAs as commodities are primarily limited by their high cost, and thus attention is being focused on composite products with plastics constituting only a minor part of the product, such as paper coatings like the plastic film moisture barrier in food or drink cartons and in sanitary napkins. In these cases, total product biodegradability can be attained for little extra cost.

Medical Applications

A notable feature of PHB is that it is biocompatible, producing a mild foreign-body response to any implant. Degradation of PHB in vivo is significantly faster than the in vitro hydrolysis in neutral buffer at body temperature, indicating that enzymes present in vivo catalyze the degradation process (166). The only degradation product is (*R*)-3-hydroxybutyric acid, which is a normal mammalian metabolite and is resorbed without metabolic or histological disturbance. Like PHB, MCL PHA might be biocompatible and useful in biomedical applications. The rigid PHB might be useful in bone fracture fixation (167), whereas the flexible elastomeric MCL PHAs might find uses in vascular grafts, surgical swabs, wound dressings, breast prostheses, and other implants (110).

Chiral Synthons

Hydrolysis of PHAs yields chiral (*R*)-3-hydroxy fatty acids, which could be a feedstock for stereospecific synthesis and the associated pharmaceutical industry (27,168). Concerning the enormous number of monomers that have been incorporated into various polymers, many different synthon precursors could be provided by hydrolyzing the polymers.

ECONOMICS

The market price of Biopol when produced by ZENECA BioProducts was about \$16 kg^{-1} (169). Current advances in fermentation and purification technology, as well as the application of metabolic engineering techniques for expanding the spectrum of utilizable substrates and for improving polymer concentration in the cells, are likely to lower the price of PHB to \$4 kg^{-1} (169). A preliminary economic study for the production of mcl-PHAs showed that the price for octanoate-based PHAs would be of the order of \$20 kg^{-1} for an annual production of 1,000 tons (110,132). Optimizing the polymer content of the cells, using cheaper carbon sources as substrates, reducing the chemicals used during downstream processing, and reducing energy consumption is expected to lower the MCL PHA price to a minimum of \$5 kg^{-1} (110,132).

It can be assumed that production of PHAs by using native or recombinant bacteria is in the long term only required for development purposes and for the production of specific functionalized PHAs. It is more likely that efforts in generating transgenic plants for PHA production purposes will succeed in ultimately reducing the price of the polymer to that of other large-scale agromaterials such as starch or conventional commodity plastics, or less than \$1 kg^{-1} (103,170). Therefore, an approach to insert the PHB biosynthesis genes into crop plants (e.g., soybean and rapeseed), which have the agronomic properties to provide high yields of PHB per hectare, has been undertaken by CEREGEN/Monsanto and is currently under way.

OUTLOOK

The development of PHAs as a commercial product remains limited to PHB and its copolymer, PHB/HV. As yet, no applications in the fabrication of articles made of MCL PHAs have been reported. There are probably at least two reasons for this. There has been little material available for development work, and the lower polymer concentration in the cells and the different recovery process means that the material cost of MCL PHA will remain even higher than for PHB for the foreseeable future.

Alternative strategies for PHA production are shown in Figure 8. Future large-scale production of PHB (and MCL PHA) in transgenic plants should ultimately provide a cheaper and sustainable technology for the production of these materials. This advance is essential if PHAs are to be used for commodity purposes and must compete on the plastic market. Improvements in the competitiveness of PHAs on the market could also derive from environmental legislation and from decreasing availability of oil and concomitant rising prices of synthetic plastics.

The development of polymers containing exotic monomers into commercial products will depend on the identification of niche applications that will support the high cost of production arising from both the cost of the monomers used and the small scale of the operation to make these specialty materials. Such applications are likely to include nonlinear optical devices or medical implants of the isotactic polymers, and the use of the stereospecific monomers as chiral synthons for the production of fine chemicals. Furthermore, capsules made of PHA for controlled release of bioactive agents, such as drugs or fertilizers, could find broad application in medicine and agriculture.

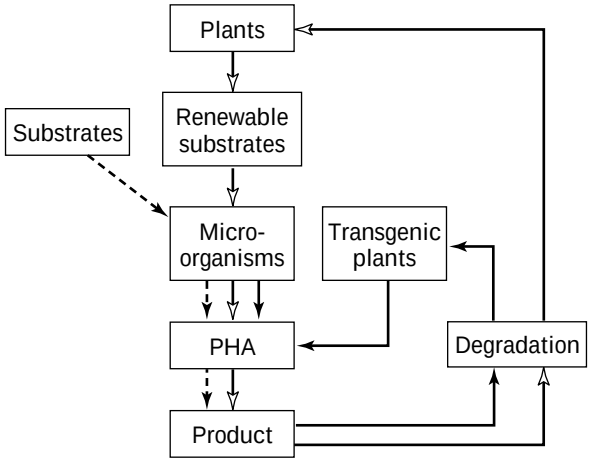


Figure 8. Overview of alternative polyhydroxyalkanoate (PHA) production strategies for different applications. Dependent on the application purpose, three different PHA production strategies can be considered. Bacterial fermentations can be used mainly for production of smaller amounts of functionalized, high-cost PHAs for specialties (- - ▶). Application of renewable sources as sole carbon source in bacterial fermentations or in addition to cosubstrates containing a specific functional group will reduce the price of the product (——▷). Low-cost products with specific properties could be produced by this strategy. Production of inexpensive and biodegradable PHAs for commodities (——▶) will be best achieved by large-scale cultivation of transgenic plants.

BIBLIOGRAPHY

1. M. Lemoigne, *Bull. Soc. Chem. Biol.* **8**, 770–782 (1926).
2. S.Y. Lee, *Biotechnol. Bioeng.* **49**, 1–14 (1996).
3. A. Steinbüchel, in D. Byrom ed., *Biomaterials. Novel Materials from Biological Sources*, Macmillan, Basingstoke, 1991, pp. 123–213.
4. C. Sasikala and C.V. Ramana, in S.L. Neidleman and A.I. Laskin eds., *Advances in Applied Microbiology*, vol. 42, Academic Press, San Diego, Calif., 1996, pp. 97–218.
5. M.J. de Smet, G. Eggink, B. Witholt, J. Kingma, and H. Wynberg, *J. Bacteriol.* **154**, 870–878 (1983).
6. A. Steinbüchel and H.E. Valentin, *FEMS Microbiol. Lett.* **128**, 219–228 (1995).
7. R.W. Lenz, Y.B. Kim, and R.C. Fuller, *FEMS Microbiol. Rev.* **103**, 207–214 (1992).
8. A. Steinbüchel, E. Hustede, M. Liebergesell, U. Pieper, A. Timm, and H. Valentin, *FEMS Microbiol. Rev.* **103**, 217–230 (1992).
9. R.N. Reusch and H.L. Sadoff, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4176–4180 (1988).
10. R.N. Reusch, *Proc. Soc. Exp. Biol. Med.* **191**, 377–381 (1989).
11. R.N. Reusch, R.P. Huang, and L.L. Bramble, *Biophys. J.* **69**, 754–766 (1995).
12. D. Seebach, A. Brunner, H.M. Burger, R.N. Reusch, and L.L. Bramble, *Helv. Chim. Acta* **79**, 507–517 (1996).
13. R.N. Reusch and H.L. Sandoff, *J. Bacteriol.* **156**, 778–788 (1983).
14. R.P. Huang, and R.N. Reusch, *J. Bacteriol.* **177**, 486–490 (1995).
15. C.E. Castuma, R.P. Huang, A. Kornberg, and R.N. Reusch, *J. Biol. Chem.* **270**, 12980–12983 (1995).
16. R.N. Reusch, A.W. Sparrow, and J. Gardiner, *Biochim. Biophys. Acta* **1123**, 33–40 (1992).
17. G. Gottschalk, *Arch. Mikrobiol.* **47**, 225–229 (1964).
18. G. Gottschalk, *Arch. Mikrobiol.* **47**, 230–235 (1964).
19. G.W. Haywood, A.J. Anderson, and E.A. Dawes, *FEMS Microbiol. Lett.* **57**, 1–6 (1989).
20. G.J. Moskowitz and J.M. Merrik, *Biochemistry* **8**, 2748–2755 (1969).
21. H.G. Schlegel, G. Gottschalk, and R. von Bartha, *Nature* **191**, 463–465 (1961).
22. A. Steinbüchel and H.G. Schlegel, *Appl. Microbiol. Biotechnol.* **31**, 168–175 (1989).
23. P.J. Senior and E.A. Dawes, *Biochem. J.* **134**, 225–238 (1971).
24. V. Oeding and H.G. Schlegel, *Biochem. J.* **134**, 239–248 (1973).
25. G.W. Haywood, A.J. Anderson, L. Chu, and E.A. Dawes, *FEMS Microbiol. Lett.* **52**, 91–96 (1988).
26. G.W. Haywood, A.J. Anderson, L. Chu, and E.A. Dawes, *FEMS Microbiol. Lett.* **52**, 259–264 (1988).
27. P.A. Holmes, in D.C. Bassett ed., *Development in Crystalline Polymers*, vol. 2, Elsevier, London, 1988, pp. 1–65.
28. M. Akiyama, Y. Taima, and Y. Doi, *Appl. Microbiol. Biotechnol.* **37**, 698–701 (1992).
29. G. Eggink, H. van der Wal, G.N.M. Huijberts, and P. de Waard, *Ind. Crops Prod.* **1**, 157–163 (1993).
30. G. Lefebvre, M. Rocher, and G. Braunegg, *Appl. Environ. Microbiol.* **63**, 827–833 (1997).
31. A.J. Anderson and E.A. Dawes, *Microbiol. Rev.* **54**, 450–472 (1990).
32. A. Steinbüchel and U. Pieper, *Appl. Microbiol. Biotechnol.* **37**, 1–6 (1992).
33. A. Steinbüchel, E.M. Debzi, R.H. Marchessault, and A. Timm, *Appl. Microbiol. Biotechnol.* **39**, 443–449 (1993).
34. Y. Doi, A. Segawa, and M. Kunioka, *Int. J. Biol. Macromol.* **12**, 106–111 (1990).
35. Y. Doi, A. Tamaki, M. Kunioka, and K. Soga, *Macromol. Chem. Rapid Commun.* **8**, 631–635 (1987).
36. H.E. Valentin, A. Schönebaum, and A. Steinbüchel, *Appl. Microbiol. Biotechnol.* **36**, 507–514 (1992).
37. B. Füchtenbusch, D. Fabritius, and A. Steinbüchel, *FEMS Microbiol. Lett.* **138**, 153–160 (1996).
38. O.P. Peoples and A.J. Sinskey, *J. Biol. Chem.* **264**, 15298–15303 (1989).
39. P. Schubert, E.F. Steinbüchel, and H.G. Schlegel, *J. Bacteriol.* **170**, 5837–5847 (1988).
40. S.C. Slater, W.H. Voige, and D.E. Dennis, *J. Bacteriol.* **170**, 4431–4436 (1988).
41. G.W. Haywood, A.J. Anderson, L. Chu, and E.A. Dawes, *Biochem. Soc. Trans.* 1046–1047 (1988).
42. T.U. Gerngross, K.D. Snell, O.P. Peoples, A.J. Sinskey, E. Csuhai, S. Masamune, and J. Stubbe, *Biochemistry* **33**, 9311–9320 (1994).
43. R. Griebel, Z. Smith, and J.M. Merrick, *Biochemistry* **7**, 3676–3681 (1968).
44. J. Wodzinska, K.D. Snell, A. Rhomberg, A.J. Sinskey, K. Biemann, and J. Stubbe, *J. Am. Chem. Soc.* **118**, 6319–6320 (1996).
45. K. Hori, K. Soga, and Y. Doi, *Biotechnol. Lett.* **16**, 709–714 (1994).
46. F. Koizumi, H. Abe, and Y. Doi, *J. Macromol. Sci. Pure Appl. Chem.* **A32**, 759–774 (1995).
47. S. Kusaka, H. Abe, S.Y. Lee, and Y. Doi, *Appl. Microbiol. Biotechnol.* **47**, 140–143 (1997).
48. T.U. Gerngross and D.P. Martin, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6279–6283 (1995).
49. M. Liebergesell and A. Steinbüchel, *Appl. Microbiol. Biotechnol.* **38**, 493–501 (1993).
50. M. Liebergesell and A. Steinbüchel, *Eur. J. Biochem.* **209**, 153–150 (1992).
51. A. Steinbüchel, in H.-J. Rehm and G. Reed eds., *Biotechnology*, vol. 6, VCH, Weinheim, 1996, pp. 405–464.
52. Y. Kawaguchi and Y. Doi, *FEMS Microbiol. Lett.* **70**, 151–156 (1990).
53. G.N. Barnard and J.K.M. Sanders, *J. Biol. Chem.* **264**, 3286–3291 (1988).
54. S.R. Amor, T. Rayment, and J.K.M. Sanders, *Macromolecules* **24**, 4583–4588 (1991).
55. K.W. Nickerson, *Appl. Environ. Microbiol.* **43**, 1208–1209 (1982).
56. H. Bauer and A.J. Owen, *Colloid Polym. Sci.* **266**, 241–247 (1988).
57. D.M. Horowitz, J. Clauss, B.K. Hunter, and J.K.M. Sanders, *Nature* **363**, 23 (1993).
58. H. Preusting, J. Kingma, G. Huisman, A. Steinbüchel, and B. Witholt, *J. Environ. Polym. Degrad.* **1**, 11–21 (1993).
59. G.W. Huisman, Ph.D. Thesis, Rijksuniversiteit Groningen, The Netherlands, 1991.

60. H. Preusting, J. Kingma, and B. Witholt, *Enzyme Microb. Technol.* **13**, 770–780 (1991).
61. E.S. Stuart, R.W. Lenz, and R.C. Fuller, *Can. J. Microbiol.* **41**, 84–93 (1995).
62. F. Mayer and M. Hoppert, *J. Basic Microbiol.* **37**, 45–52 (1997).
63. A. Steinbüchel, K. Aerts, W. Babel, C. Follner, M. Liebergesell, M.H. Madkour, F. Mayer, U. Pieper-Fürst, A. Pries, H.E. Valentin, and R. Wieczorek, *Can. J. Microbiol.* **41**, 94–105 (1995).
64. R.C. Fuller, J.P. O'Donnel, J. Saulnier, T.E. Redlinger, J. Foster, and R.W. Lenz, *FEMS Microbiol. Rev.* **103**, 279–288 (1992).
65. R. Wieczorek, A. Pries, A. Steinbüchel, and F. Mayer, *J. Bacteriol.* **177**, 2425–2435 (1995).
66. U. Pieper-Fürst, M.H. Madkour, F. Mayer, and A. Steinbüchel, *J. Bacteriol.* **176**, 4328–4337 (1994).
67. U. Pieper-Fürst, M.H. Madkour, F. Mayer, and A. Steinbuchel, *J. Bacteriol.* **177**, 2513–2523 (1995).
68. M. Liebergesell, B. Schmidt, and A. Steinbüchel, *FEMS Microbiol. Lett.* **99**, 227–232 (1992).
69. M. Liebergesell and A. Steinbüchel, *Biotechnol. Lett.* **18**, 719–724 (1996).
70. R. Wieczorek, A. Steinbüchel, and B. Schmidt, *FEMS Microbiol. Lett.* **135**, 23–30 (1996).
71. M.A. Schembri, A.A. Woods, R.C. Bayly, and J.K. Davies, *FEMS Microbiol. Lett.* **133**, 277–283 (1995).
72. C.G. Follner, W. Babel, and A. Steinbüchel, *Can. J. Microbiol.* **41**, 124–130 (1995).
73. C.G. Follner, M. Madkour, F. Mayer, W. Babel, and A. Steinbüchel, *J. Basic Microbiol.* **37**, 11–21 (1997).
74. D. Byrom, *FEMS Microbiol. Rev.* **103**, 247–250 (1992).
75. T. Yamane, *FEMS Microbiol. Rev.* **103**, 257–264 (1992).
76. H.G. Schlegel and G. Gottschalk, *Biochem. Z.* **342**, 249–259 (1965).
77. D. Byrom, in E.A. Dawes ed., *Novel Biodegradable Microbial Polymers*, Kluwer, Dordrecht, 1990, pp. 113–117.
78. U.J. Hänggi, in E.A. Dawes ed., *Novel Biodegradable Microbial Polymers*, Kluwer, Dordrecht, 1990, pp. 65–70.
79. O. Hrabak, *FEMS Microbiol. Rev.* **103**, 251–256 (1992).
80. D.R. Williamson and J.F. Wilkinson, *J. Gen. Microbiol.* **19**, 198–209 (1958).
81. E. Berger, B.A. Ramsay, J.A. Ramsay, C. Chaverie, and G. Braunegg, *Biotechnol. Tech.* **3**, 227–232 (1989).
82. D. Byrom, *Trends Biotechnol.* **5**, 246–250 (1987).
83. Y. Doi, *Microbial Polyesters*, VCH, New York, 1990.
84. P.J. Hocking and R.H. Marchessault, in G.J.L. Griffin ed., *Chemistry and Technology of Biodegradable Polymers*, Blackie A&P, Glasgow, 1994, pp. 48–96.
85. R. Sharma and A.R. Ray, *J. Macromol. Sci. Rev. Macromol.* **C35**, 327–359 (1995).
86. S. Slater, T. Gallaher, and D. Dennis, *Appl. Environ. Microbiol.* **58**, 1089–1094 (1992).
87. S.Y. Lee and H.N. Chang, *Can. J. Microbiol.* **41**, 207–215 (1995).
88. S.Y. Lee, K.S. Yim, H.N. Chang, and Y.K. Chang, *J. Biotechnol.* **32**, 203–211 (1994).
89. S.J. Sim, K.D. Snell, S.A. Hogan, J. Stubbe, C.K. Rha, and A.J. Sinskey, *Nat. Biotechnol.* **15**, 63–67 (1997).
90. Y. Poirier, D.E. Dennis, K. Klomparens, and C. Somerville, *Science* **256**, 520–523 (1992).
91. C. Nawrath, Y. Poirier, and C. Somerville, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12760–12764 (1994).
92. Y. Poirier, C. Somerville, L.A. Schechtman, M.M. Satkowski, and I. Noda, *Int. J. Biol. Macromol.* **17**, 7–12 (1995).
93. T.A. Leaf, M.S. Peterson, S.K. Stoup, D. Somers, and F. Srienc, *Microbiology UK* **142**, 1169–1180 (1996).
94. M.D. Williams, J.A. Rahn, and D.H. Sherman, *Appl. Environ. Microbiol.* **62**, 2540–2546 (1996).
95. M.D. Williams, A.M. Fieno, R.A. Grant, and D.H. Sherman, *Protein Expr. Purif.* **7**, 203–211 (1996).
96. G.N.M. Huijberts, G. Eggink, P. de Waard, G.W. Huisman, and B. Witholt, *Appl. Environ. Microbiol.* **58**, 536–544 (1992).
97. G.N.M. Huijberts, T.C. de Rijk, P. de Waard, and G. Eggink, *J. Bacteriol.* **176**, 1661–1666 (1995).
98. H. Brandl, R.A. Gross, R.W. Lenz, and R.C. Fuller, *Appl. Environ. Microbiol.* **54**, 1977–1982 (1988).
99. G.W. Haywood, A.J. Anderson, and E.A. Dawes, *Biotechnol. Lett.* **11**, 471–476 (1989).
100. G.W. Huisman, O. de Leeuw, G. Eggink, and B. Witholt, *Appl. Environ. Microbiol.* **55**, 1949–1954 (1989).
101. G.W. Huisman, E. Wonink, R. Meima, B. Kazemier, P. Terpstra, and B. Witholt, *J. Biol. Chem.* **266**, 2191–2198 (1991).
102. A. Timm and A. Steinbüchel, *Eur. J. Biochem.* **209**, 15–30 (1992).
103. F.R. van der Leij and B. Witholt, *Can. J. Microbiol.* **41**, 222–238 (1995).
104. G.W. Huisman, E. Wonink, G.J.M. de Koning, H. Preusting, and B. Witholt, *Appl. Microbiol. Biotechnol.* **38**, 1–5 (1992).
105. D. Jendrossek, B. Müller, and H.G. Schlegel, *Eur. J. Biochem.* **218**, 701–710 (1993).
106. B.A. Ramsay, I. Saracovan, J.A. Ramsay, and R.H. Marchessault, *Appl. Environ. Microbiol.* **58**, 744–746 (1992).
107. A. Timm and A. Steinbüchel, *Appl. Environ. Microbiol.* **56**, 3360–3367 (1990).
108. H. Preusting, W. Hazenberg, and B. Witholt, *Enzyme Microb. Technol.* **15**, 311–316 (1993).
109. H. Preusting, R. van Houten, A. Hoefs, E. Kool van Langenberghe, O. Favre-Bulle, and B. Witholt, *Biotechnol. Bioeng.* **41**, 550–556 (1993).
110. G.J.M. de Koning, M.B. Kellerhals, C. van Meurs, and B. Witholt, *J. Environ. Polym. Degrad.* **4**, 243–252 (1996).
111. R.G. Lageveen, G.W. Huisman, H. Preusting, P. Ketelaar, G. Eggink, and B. Witholt, *Appl. Environ. Microbiol.* **54**, 2924–2932 (1988).
112. H. Preusting, A. Nijenhuis, and B. Witholt, *Macromolecules* **23**, 4220–4224 (1990).
113. W.M. Hazenberg, Ph.D. Thesis, ETH Zürich, Switzerland, 1997.
114. D. Carnicero, M. Fernandez Valverde, L.M. Canedo, C. Schleissner, and J.M. Luengo, *FEMS Microbiol. Lett.* **149**, 51–58 (1997).
115. G.W. Haywood, A.J. Anderson, D.F. Ewing, and E.A. Dawes, *Appl. Environ. Microbiol.* **56**, 3354–3359 (1990).
116. G.N.M. Huijberts, Ph.D. Thesis, Rijksuniversiteit Groningen, The Netherlands (1996).
117. K. Fritzsche, R.W. Lenz, and R.C. Fuller, *Int. J. Biol. Macromol.* **12**, 85–91 (1990).
118. Y.B. Kim, R.W. Lenz, and R.C. Fuller, *J. Polym. Sci. Appl. Polym. Chem.* **33**, 1367–1374 (1995).
119. Y.B. Kim, R.W. Lenz, and R.C. Fuller, *Macromolecules* **25**, 1852–1857 (1992).

120. C. Abe, Y. Taima, Y. Nakamura, and Y. Doi, *Polym. Commun.* **31**, 404–406 (1990).
121. O. Kim, R.A. Gross, W.J. Hammar, and R.A. Newmark, *Macromolecules* **29**, 4572–4581 (1996).
122. Y. Doi and C. Abe, *Macromolecules* **23**, 3705–3707 (1990).
123. K. Fritzsche, R.W. Lenz, and R.C. Fuller, *Makromol. Chem.* **191**, 1957–1965 (1990).
124. J.M. Curley, B. Hazer, R.W. Lenz, and R.C. Fuller, *Macromolecules* **29**, 1762–1766 (1996).
125. Y.B. Kim, R.W. Lenz, and R.C. Fuller, *Macromolecules* **24**, 5256–5260 (1991).
126. J.J. Song and S.C. Yoon, *Appl. Environ. Microbiol.* **62**, 536–544 (1996).
127. H. Ritter and A.G. Vonspee, *Macromol. Chem. Phys.* **195**, 1665–1672 (1994).
128. O.Y. Kim, R.A. Gross, and D.R. Rutherford, *Can. J. Microbiol.* **41**, 32–43 (1995).
129. R.A. Gross, O. Kim, D.R. Rutherford, and R.A. Newmark, *Polym. Int.* **39**, 205–213 (1996).
130. S. Langenbach, B.H.A. Rehm, and A. Steinbüchel, *FEMS Microbiol. Lett.* **150**, 303–309 (1997).
131. Q. Ren, Ph.D. Thesis, ETH, Zurich, Switzerland, 1997.
132. G.J.M. de Koning, M. Kellerhals, C. van Meurs, and B. Witholt, *Bioprocess Eng.* **17**, 15–21 (1997).
133. G.J.M. de Koning and B. Witholt, *Bioprocess Eng.* **17**, 7–13 (1997).
134. R.H. Marchessault, F.G. Morin, S. Wong, and I. Saracovan, *Can. J. Microbiol.* **41**, 138–142 (1995).
135. R.H. Marchessault, C.J. Monasterios, and P. Lepoutre, in E.A. Dawes ed., *Novel Biodegradable Microbial Polymers*, Kluwer, Dordrecht, 1990, pp. 97–112.
136. R.H. Marchessault, C.J. Monasterios, F.G. Morin, and P.R. Sundararajan, *Int. J. Biol. Macromol.* **12**, 158–165 (1990).
137. R.A. Gross, C. DeMello, R.W. Lenz, H. Brandl, and R.C. Fuller, *Macromolecules* **22**, 1106–1115 (1989).
138. G.J.M. de Koning, H.H.M. van Bilsen, P.J. Lemstra, W. Hazenberg, B. Witholt, H. Preusting, J.G. van der Galiën, A. Schirmer, and D. Jendrossek, *Polymer* **35**, 2090–2097 (1994).
139. K.D. Gagnon, R.W. Lenz, R.J. Farris, and R.C. Fuller, *Macromolecules* **25**, 3723–3728 (1992).
140. K.D. Gagnon, R.W. Lenz, R.J. Farris, and R.C. Fuller, *Polymer* **35**, 4358–4367 (1994).
141. K.D. Gagnon, R.W. Lenz, R.J. Farris, and R.C. Fuller, *Polymer* **35**, 4368–4375 (1994).
142. J.M. Merrick and M. Doudoroff, *J. Bacteriol.* **88**, 60–71 (1964).
143. S.J. Holland, A.M. Jolly, M. Yasin, and B.J. Tighe, *Biomaterials* **8**, 289–295 (1987).
144. N.D. Miller and D.F. Williams, *Biomaterials* **8**, 129 (1987).
145. H. Brandl, R. Bachofen, J. Mayer, and E. Wintermantel, *Can. J. Microbiol.* **41**, 143–153 (1995).
146. J. Mergaert, G. Glorieux, L. Hauben, V. Storms, M. Mau, and J. Swings, *Syst. Appl. Microbiol.* **19**, 407–413 (1996).
147. J. Mergaert, A. Wouters, C. Anderson, and J. Swings, *Can. J. Microbiol.* **41**, 154–159 (1995).
148. J. Mergaert, A. Webb, C. Anderson, A. Wouters, and J. Swings, *Appl. Environ. Microbiol.* **59**, 3236–3238 (1993).
149. D.F. Gilmore, S. Antoun, R.W. Lenz., S. Goodwin, R. Austin, and R.C. Fuller, *J. Ind. Microbiol.* **10**, 199–206 (1992).
150. D.F. Gilmore, S. Antoun, R.W. Lenz, and R.C. Fuller, *J. Environ. Polym. Degrad.* **1**, 269–274 (1993).
151. B.H. Briese, D. Jendrossek, and H.G. Schlegel, *FEMS Microbiol. Lett.* **119**, 396 (1994).
152. L.R. Krupp and W.J. Jewell, *Environ. Sci. Technol.* **26**, 193–198 (1992).
153. C.L. Yue, R.A. Gross, and S.P. McCarthy, *Polym. Degrad. Stabil.* **51**, 205–210 (1996).
154. H. Abe and Y. Doi, *Macromolecules* **29**, 8683–8688 (1996).
155. D. Jendrossek, A. Schirmer, and H.G. Schlegel, *Appl. Microbiol. Biotechnol.* **46**, 451–463 (1996).
156. D. Jendrossek, I. Knoke, R.B. Habibian, A. Steinbüchel, and H.G. Schlegel, *J. Environ. Polym. Degrad.* **1**, 53–61 (1993).
157. A. Schirmer, D. Jendrossek, and H.G. Schlegel, *Appl. Environ. Microbiol.* **59**, 1220 (1993).
158. K. Nakayama, T. Saito, T. Fukui, Y. Shirakura, and K. Tomita, *Biochim. Biophys. Acta* **827**, 63–72 (1985).
159. F.P. Delafield, K.E. Cooksey, and M. Doudoroff, *J. Biol. Chem.* **240**, 4023–4028 (1965).
160. Y. Shirakura, T. Fukui, T. Tanio, K. Nakayama, R. Matsuno, and K. Tomita, *Biochim. Biophys. Acta* **748**, 331–339 (1983).
161. Y. Doi, Y. Kumagai, N. Tanahashi, and K. Mukai, in M. Vert ed., *Biodegradable Polymers and Plastics*, Royal Soc. of Chemistry, London, 1992, pp. 139–148.
162. K.E. Jäger, A. Steinbüchel, and D. Jendrossek, *Appl. Environ. Microbiol.* **61**, 3113–3118 (1995).
163. A. Schirmer and D. Jendrossek, *J. Bacteriol.* **176**, 7065–7073 (1994).
164. P.A. Holmes, *Phys. Technol.* **16**, 32–36 (1985).
165. H. Verhoogt, B.A. Ramsay, and B.D. Favis, *Polymer* **35**, 5155–5169 (1994).
166. E. Piskin, *J. Biomater. Sci. Polym. Ed.* **6**, 775–795 (1995).
167. C. Doyle, E.T. Tanner, and W. Bonfield, *Biomaterials* **12**, 841 (1991).
168. R.H. Marchessault, T.L. Bluhm, Y. Deslandes, G.K. Hamer, W.J. Orts, P.R. Sundararajan, and M.G. Taylor, *Macromol. Chem.* **19**, 235–254 (1988).
169. S.Y. Lee, *Trends Biotechnol.* **14**, 431–438 (1996).
170. Y. Poirier, C. Nawrath, and C. Somerville, *Biotechnology* **13**, 142–150 (1995).

POLYHYDROXYALKANOATES, SEPARATION, PURIFICATION, AND MANUFACTURING METHODS

JULIANA RAMSAY
Queen's University
Kingston, Ontario
Canada

BRUCE RAMSAY
Queen's University
Kingston, Ontario
Canada

KEY WORDS

Polyhydroxyalkanoates (PHA)
Polyhydroxybutyrate (PHB)
Processing
Production
Separation

OUTLINE

POLY-3-HYDROXYALKANOATES AS BIODEGRADABLE PLASTICS

Microbially produced biodegradable plastics can be divided into four classes: polysaccharides, polymers of amino acids, naturally occurring polyesters, and semisynthetic polyesters (Table 1). All possess linkages that are susceptible to nonspecific, extracellular, microbial enzymes that exist in nature and thus are inherently environmentally degradable. Their structures can be used as the basis for the design of synthetic biodegradable polymers. Unlike many common synthetic plastics, those microbially produced are often at least partially water soluble or tend to be easily hydrolyzed, especially at extremes of temperature and pH. These characteristics aid in their biodegradability but greatly limit their use as plastics. Although their properties as plastics may be greatly improved through chemical modification, care must be taken to maintain environmental degradability.

The ideal plastic is durable, biodegradable, recyclable, and produced from renewable resources such as carbohydrates. Poly-3-hydroxyalkanoates (PHAs) are a family of intracellular polyesters produced by a wide variety of microorganisms. They are water insoluble, relatively resistant to aqueous hydrolysis, and readily environmentally degradable. They can be produced from renewable resources, and, although temperature sensitive, they are potentially recyclable. PHAs have mechanical properties similar to those of some common synthetic plastics. Like conventional thermoplastics, PHAs can be recycled with some loss of molecular weight and mechanical properties. Unlike many conventional polymers, PHAs are completely environmentally degradable and will decompose rapidly to CO_2 and H_2O during composting. Degradation has also been demonstrated in soil burial studies and in the marine environment.

PHAs whose polymeric units are predominantly five carbons or fewer are classified as short-chain-length (SCL) PHAs, and all others are referred to as medium-chain-length (MCL) PHAs. Homopolymeric poly-3-hydroxybutyrate (PHB) has properties that are frequently compared to those of polypropylene, except that it is more brittle and its melting temperature is high enough that there may be significant thermal degradation on melt extrusion. Lower-melting, more-flexible PHB can be obtained by annealing or addition of plasticizers. However, copolymers of 3-hydroxybutyrate and 3-hydroxyvalerate (P[HB-co-HV]) and of 3-hydroxybutyrate and 4-hydroxybutyrate (P[3HB-co-4HB]) are more flexible and have lower melting temperatures than PHB, making them more suitable for the manufacture of extruded products. MCL PHAs have even lower melting points and glass transition temperatures; they are tough, rubberlike thermoelastomers, especially after cross-links between the polymer chains have been formed by heat or chemical induction.

PHAs would seem to be ideal biodegradable plastics, but their production is much different from that of conventional plastics. The following sections describe how PHAs may be produced by widely varying methodologies, with details of diverse approaches to PHA fermentation processes and the separation of the polymer from the biomass. Existing commercial processes and future possibilities for PHA production with an emphasis on cost reduction are also discussed.

Table 1. Classes of Microbially Produced Biodegradable Plastics

Class	Linkage	Example	Effects of H_2O at different pH and temperature	Fermentation characteristics
Polysaccharides	Glycosidic	Pullulan	Often soluble or hydrolyzed	High yield, simple separation but cost increases with purity; production limited by O_2 transfer
Polyaminoacids	Peptide	Poly-γ-glutamate	Often soluble or hydrolyzed	Still experimental; separation simple; yield should be high
Polyesters	Ester	Polyhydroxyalkaoates	Not soluble; resistant to hydrolysis	Yield depends on type of polymer and on the process; separation can be expensive
Semisynthetic polymers of organic acids	Ester	Polylactides	Generally not soluble but are hydrolyzed	High yield

THREE ROUTES TO PHA PRODUCTION

Chemical Synthesis

PHB can be produced synthetically from racemic β-butyrolactone, and P(HB-co-HV) is made from a similar mixture of β-butyrolactone and β-valerolactone with triethylaluminum and water as catalyst (1). Although the stereospecificity of the resulting polymer can be controlled to a certain extent by adjusting the ratio of triethylaluminum and water, synthetic PHAs are usually optically inactive and hence only partially biodegradable.

Synthesis in Transgenic Plants

SCL PHA. PHB can be synthesized in plants in which the bacterial reductase and polymerase genes have been introduced. Although virtually all organisms possess the remaining gene in the SCL PHA pathway, coding for 3-ketothiolase, its activity and specificity vary and it may not be present in all eukaryotic organelles. Production is usually increased by cloning the 3-ketothiolase gene with the other two. Although transgenic plants are easily created, productivity remains problematic. The SCL PHA pathway was first cloned into *Arabidopsis thaliana*, the *Escherichia coli* of plant biotechnology (2). The yield was low, 0.1% dry weight, and the plants were stunted. However, granules similar to those found in bacteria were produced. The initial cloning strategy targeted the cytoplasm, and granules were even found in the cell nucleus (Fig. 1). Subsequent work has targeted the plastid, which is also the site of starch and oil synthesis in plants, resulting in increased yields of up to 14% dry weight in the leaves of *A. thaliana*. These plants grew well but were chlorotic, possibly because of large amounts of PHB in the chloroplasts. Present research centers on oil- and starch-producing plants because these are already adapted to have large amounts of storage material in the plastids (3), but other strategies are also being assessed.

MCL PHA. Because of its high melting point and brittle nature, the PHB homopolymer might have limited application even if its production cost in transgenic plants was low. Although SCL PHA copolymers such as P(HV-co-HB) have been synthesized in plants (H. Valentin, personal communication), it is not obvious how the metabolic pathways of plants can be manipulated to maximize their production. It may be easier to produce MCL PHA in plants. Unfortunately progress is limited by a lack of understanding of the link between fatty acid synthesis and MCL PHA synthesis. Nevertheless, small quantities of MCL PHA have recently been produced in recombinant *A. thaliana* (Y. Poirier, personal communication).

Microbial Synthesis

SCL PHA. Many bacterial species accumulate PHB, but PHA copolymers are more commercially valuable. Most PHB-accumulating bacteria will produce P(HB-co-HV) under the appropriate conditions. When propionic acid is supplied during a fermentation process, some of the acid is converted to propionyl-CoA. A portion of this propionyl-CoA is condensed with acetyl-CoA via β-ketothiolase to produce five-carbon monomers; this is the most common pathway of P(HB-co-HV) production. Because acetate is needed for this synthetic route and some propionic acid is decarboxylated before monomer synthesis, usually less than 50% of the monomers will be HV even if a feed of pure propionic is added (i.e., no glucose or acetate). This yield can be increased in *Ralstonia eutropha* (previously *Alcaligenes eutrophus*) by decreasing the dissolved oxygen concentration in the bioreactor during the accumulation phase (4).

High yields of HV are obtained when pentanoic acid is used because pentanoic acid is metabolized directly to L-(3)-hydroxyvaleryl-CoA, which is subsequently racemized (at least in *R. eutropha*) to D-(3)-hydroxyvaleryl-CoA (5). Although some acetate is produced leading to HB monomer formation, PHAs with greater than 90% HV content can be obtained. SCL PHA producers have not been reported to accumulate monomers with side chains longer than C_2 or C_3 but can produce P(3HB-co-4HB) when fed 4-hydroxybutyric acid or 1,4-butanediol. *R. eutrophus* accumulates terpolyesters of 3HB-, 3HV-, and 5HV-monomers when fed a mixture of pentanoic and 5-chloropentanoic acids. Even among SCL-accumulating microorganisms, there may be differences in the PHA synthetic pathway. For example, in *Rhodospirillum rubrum*, L-(+)-3-hydroxybutyrl-CoA is racemized to D-(−)-3-hydroxybutyrl-CoA using enoyl-CoA hydrases that have not been found in *Azotobacter beijerinckii*.

In most SCL PHA–accumulating bacteria, high intracellular levels of reduced nicotinamide adenine dinucleotide (NADH) and acetyl-CoA greatly stimulate the rate of PHA synthesis. Thus, if the growth rate is limited by a nutrient other than the source of carbon and energy, PHA is accumulated in the cell. For this reason, PHA production is usually divided into two phases. In the first phase, balanced growth proceeds with relatively little PHA accumulation. When the concentration of the limiting nutrient in batch culture falls below a critical value, the rate of PHA synthesis increases while synthesis of other cellular components decreases during the second phase. Because cosubstrates such as propionic acid are expensive and can be used for the synthesis of biomass other than PHA, these chemicals are only added during the accumulation phase. This limitation probably results in nonhomogeneity of polymer composition in different PHA granules with some (produced during the growth phase) containing almost exclusively PHB. There is still debate about the existence and importance of monomer turnover in PHA granules. In contrast, carbon and energy limitations increase depolymerase activity, leading to a decrease in cellular PHA content. The control of these biochemical pathways varies greatly among different bacterial species. For example, *Alcaligenes latus* does not require nitrogen or phosphorous limitation; it can accumulate large amounts of PHB during balanced growth (6).

The most important economic considerations are substrate costs, reactor productivity, and separation costs. SCL PHA can be produced from hydrolyzed starch (glucose), lactose (using recombinant *E. coli* on cheese whey), cane or beet molasses (using *Azotobacter vinelandii* or *Al-*

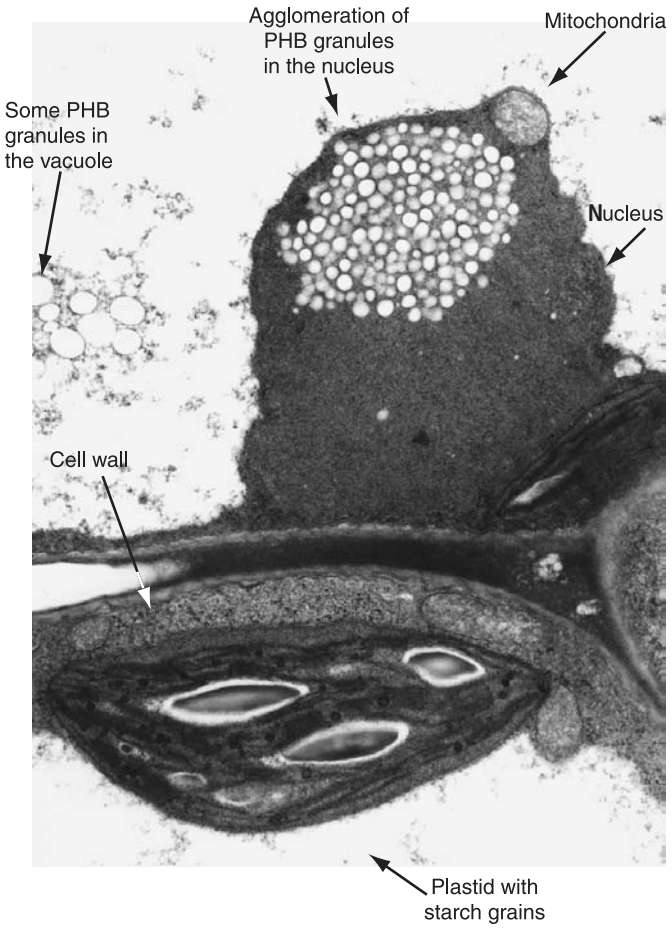


Figure 1. Polyhydroxybutyrate (PHB) granules in *Arabidopsis thaliana* (electron micrograph courtesy of Yves Poirier, Université de Lausanne; Chris Somerville, Carnegie Institute; and Karen Klomparens, Michigan State University).

caligenes latus on sucrose), waste wood (fermented, hydrolyzed hemicellulose), methanol, and even mixtures of hydrogen and carbon dioxide (using *R. eutropha*). Addition of the appropriate cosubstrate such as propionic acid can result in P(HB-co-HV) synthesis in any of these processes although the %HV can vary considerably depending on the bacterial strain. High productivity have been reported for several of these substrates. Using *R. eutropha* growing on glucose, 164 g L^{-1} of biomass containing 76% PHB has been produced in a fed-batch process in only 50 h (7). Densities well in excess of 100 g L^{-1} have been reported for fed-batch processes using *E. coli* strains with multicopy plasmids containing the key SCL PHA synthetic genes (8) as well as for methylotrophs growing on methanol (9). In recombinant *E. coli* strains, PHB content is frequently reported as being in excess of 90% of the biomass dry weight, making the separation process much simpler.

MCL PHA. The highest productivity results from using octanoic or nonanoic acids, but these are toxic at concentrations of several grams per liter and their low solubility limits the growth rate as the result of poor mass transfer. The salts of these acids are much more soluble, but the final product concentration in fed-batch culture is limited by accumulation of toxic concentrations of sodium ions in the medium. In any event, the present cost of both the salts and the acids prohibits their use in commercial production. Most microorganisms that make MCL PHA produce lipases and grow well on vegetable oils. If the lipase concentration limits the production rate, the oils can be chemically hydrolyzed to liberate their fatty acids. A *Pseudomonas putida* strain has been grown in fed-batch culture from oleic acid (a fatty acid commonly found in plant oils) to a biomass concentration of 92 g L^{-1} containing 45% PHA in 26 h (10). If of sufficient low density, these PHAs may float to the surface after breakage of the bacterial cell and can thus be easily and cheaply separated from the rest of the biomass.

SEPARATION OF PHA FROM BIOMASS

General Considerations

All separation processes to date have been developed for microbial rather than plant biomass, but many of these approaches could also be applied to recovery of PHA from plants. As a first step in the recovery from microbial biomass, the cells are separated from the fermentation broth, usually by centrifugation or flocculation. In the latter case, the pH of the fermentation broth is adjusted to 9; it is heated, then the pH is readjusted to 5 (11) and the biomass recovered by centrifugation or filtration. The wet biomass may be freeze- or spray-dried to remove water. Processes to separate the intracellular PHA from the rest of the mi-

crobial biomass may be divided into two categories: *(1)* selective solubilization of PHA in an organic solvent, leaving the majority of the other biomass components in suspension, and *(2)* destruction (i.e., enzymatic solubilization) of cellular components other than PHA, leaving the PHA granules in suspension.

Physical treatments such as drying, heating, freezing and thawing, repeated centrifugation, and exposure to acids and organic solvents change the amorphous nature of native PHA to a semicrystalline state.

Selective Solubilization of PHAs: Solvent Extraction

The majority of the patented separation processes describe the extraction of PHB from microbial biomass using organic solvents such as chlorinated hydrocarbons (e.g., chloroform or 1,2-dichloroethane), azeotropic mixtures (e.g., 1,1,2-trichloroethane with water [12]), chloroform with either methanol, ethanol, acetone, or hexane (13), and cyclic carbonates (e.g., hot [120–150 °C] ethylene carbonate or 1,2-propylene carbonate [14]) in which the polymer is soluble. With the wider variety of PHAs that can be produced today, the choice of solvents should be carefully considered. In general, solvents that are suitable for PHB should be equally good for any SCL PHA and MCL PHA. The reverse may not be true. For example, while semicrystalline PHB is insoluble in acetone, MCL PHA will dissolve in it.

Cells may be pretreated to make the polymer more accessible to the extraction solvent. This may be done by washing in a hydrophilic solvent in which the polymer is insoluble such as methanol or acetone for SCL PHA and methanol for MCL PHA or by means of physical disruption such as grinding, wet-milling, French press, or freezing and thawing cycles. Pretreatment with acetone has been shown to have the added advantage of removing certain polar lipid impurities and increasing the amount of PHB recovered (15). After solvent extraction, the polymer solution is separated from the cellular residue by filtration or centrifugation. PHB is then precipitated from the chlorinated hydrocarbons or azeotropic mixtures with a cold nonsolvent such as methanol. In the case of ethylene carbonate and 1,2-propylene carbonate, most of the PHB precipitates on cooling (14).

Difficulties are encountered at polymer solution concentrations greater than 5% by weight because these concentrations tend to be very viscous. As a result, large amounts of solvent are required to obtain sufficiently dilute solutions to enable easy separation. Even with efficient solvent recovery procedures, some solvent loss is expected. This method is not considered economical as solvent recovery and raw material costs are high. However, solvent extraction is useful when the PHA content of the biomass is low or when high polymer purity is required. A significant portion of the PHB may remain in the biomass after solvent extraction (15).

Although thermal degradation of PHB in air or a nitrogen atmosphere has been well studied (16), few examples show that PHB degrades in organic solvents. The molecular weight of PHB decreased in propylene carbonate and ethylene carbonate at 110–140 °C (14), in a chloroform–methanol mixture at room temperature (17), and in methylene chloride, chloroform, and 1,2-dichloroethane at their boiling points (15).

Although solvent extraction is commercially uneconomical, it is a commonly used laboratory technique for PHA recovery, especially for molecular weight determination. Usually, methanol-pretreated (or ethanol-pretreated), lyophilized biomass is heated in chloroform under reflux conditions for about 1 h, the polymer solution is recovered by filtration or centrifugation, and the polymer is precipitated by cold ethanol or methanol. The polymer may be further purified by redissolving in chloroform and reprecipitating with cold ethanol. It should be noted that about 25% of the original molecular weight (determined by chloroform extraction under ambient conditions) was lost when refluxed in chloroform for 1 h (15).

Solubilization of Cellular Components Other than PHA

Enzymatic Methods. Classical enzymatic treatment (involving lysozyme, proteinases, DNAses, etc.) with or without a surfactant step can be used to solubilize the cellular biomass (18). Biologists have used these methods to prepare native PHB granules to study intracellular enzymatic degradation. On a commercial level, Holmes and Lim (19) described a process in which the fermentation broth may be first heated to 80 °C to denature DNA and some proteins. A series of enzymes (such as lysozyme, phospholipase, lecithinase, or alcalase [proteinase]) are then used to obtain a product that is 90–95% pure with 6–7% proteins and some peptidoglycan as major impurities. There may be a final wash with hydrogen peroxide to make a whiter product. This is the reported method of choice for the current production of P(HB-co-HV) (20), which is sold under the tradename Biopol by Monsanto plc. When PHA granules recovered by enzymatic treatment are freeze-dried, they have a hard, crystalline shell and a noncrystalline core (21).

Chemical Methods. Strong oxidizing agents such as sodium hypochlorite have been used to dissolve the cellular biomass under conditions that result in significant loss of PHA molecular weight. However, by optimizing the operating conditions, PHB with an average molecular weight of 600,000 g/mol and 95% purity has been obtained from *R. eutropha* biomass containing 50% PHB with an initial molecular weight of 1,200,000 g/mol. When the biomass is pretreated with a surfactant, a reduced contact time with hypochlorite is needed to obtain PHB with the same purity. This method allows an even higher molecular weight (800,000 g/mol) to be obtained (22). Never-dried granules obtained by surfactant–hypochlorite treatment also have an external crystalline shell and an amorphous core (23).

Physical Methods. Although there are many methods (e.g., reciprocating pump, ball mill) to physically disrupt microbial cells, few processes that are based on these principles have been proposed for the large-scale recovery of PHAs. Holmes and Jones (24) patented a process in which a biomass suspension was heated to 220 °C under nitrogen in an autoclave, then forced through a fire-jet into a tank

of cold water. Cell breakage occurred, releasing PHB granules, which were suspended in 50% aqueous acetone to remove lipids and to cause flocculation.

COMMERCIAL PRODUCTION

History

PHB remained an academic curiosity until the 1950s and 1960s when W.R. Grace and Co. in the United States produced small quantities for commercial evaluation. They obtained patents for production processes and manufactured articles such as sutures and prosthetic devices (25) but eventually abandoned the project because of low PHB yields and difficulties in the separation process. About 10 years later, Imperial Chemical Industries (ICI) began a program in PHA production and product development. They patented the first commercial process to produce P(HB-co-HV) (26).

Current Production Methods

Tredegar Industries has used recombinant *E. coli* to produce more than 100 g L^{-1} of biomass containing at least 80% PHB by dry weight at the pilot-scale level. Mild detergent treatment and washing steps followed by spray-drying result in a white product that is 98% pure (27).

Fermentation and separation processes were developed for the pilot-scale production of PHB using an *Alcaligenes latus* strain. As much as 1 ton per week of PHB has been produced in Austria in a 15,000-L bioreactor (28). This technology is currently owned by the German company, Biomer. The Biomer fermentation process is relatively simple with multistaging from the petri dish to a shaker flask to a small fermenter, which is then used to inoculate the production reactor (Fig. 2a). If sufficient carbon source (molasses) is added during the fed-batch process, as much as 60 g L^{-1} of PHB may be obtained. The resulting biomass is washed in water. The concentrated cell suspension is then subjected to methylene chloride extraction (Fig. 2b). The resulting slurry is separated in a decanter. The solvent phase is then injected into hot water in a precipitation vessel where the methylene chloride evaporates and the PHB precipitates. After water removal (drying), the PHB can then be processed into a marketable product.

Biopol has been produced by Zeneca from glucose and propionic acid in 35,000-L airlift reactors and in stirred tank reactors up to 200,000 L (29). Monsanto uses conventional stirred-tank fermenters to produce 1,000 tons of Biopol per year. Phosphate limitation (after about 60 h of growth) is used to trigger the accumulation phase in a strain of *R. eutropha*. Phosphate limitation was found to give better results than nitrogen limitation (30). Although solvents (chloroform or methylene chloride extraction after a hot methanol wash) may be used when high purity is required, the P(HB-co-HV) is usually obtained by treatment of the biomass with a series of surfactants, enzymes, and possibly oxidizing agents to dissolve the non-PHA biomass. The product is essentially a latex that may then be washed, spray-dried, compounded with nucleating agents and plasticizers, and extruded in the form of pellets. Several grades of Biopol are available at a cost of about $8 per pound. Based on process development at IPT (Sao Paulo), Copersucar, the Brazilian ethanol cooperative is now using hydrolyzed sucrose and NH_4^+ limitation to produce SCL PHA. these are purified by solvent extraction. Pilot-scale production capacity is currently 30 to 60 tons per annum. Cost savings from the burning of bagasse to supply energy and from cheap cane sugar allow Copersucar to sell PHA for much less than $8 per pound.

PROCESSING TO FINAL PRODUCTS

Conventional equipment used to process polyethylene can be used for Biopol to form blow-molded, injection-molded, and film products. A key requirement is temperature control. It is recommended to operate at the lowest possible temperature and to keep the residence time at a minimum because the melting point of PHB is about 178 °C and its degradation temperature begins at 180 °C. The melting point of the copolymers decreases as the HV content increases, which allows HV copolymers to be processed at lower temperatures with less risk of thermal degradation. Plasticized high-HV Biopol can have a melting point as low as 136 °C. For most applications, Monsanto suggests processing at a maximum temperature of 170 °C and a maximum residence time of 3 min, and that molds and blow-pin temperatures be kept at 60 ± 5 °C (31). Mechanical properties deteriorate rapidly if the average molecular weight falls below 400,000 g mol^{-1}.

PHAs are moisture resistant and have gas barrier properties similar to the best coated films, so that they have potential use in products which require a plastic moisture barrier. Possible markets include mulch films, food and agricultural packaging, drink cartons, feminine hygiene products, disposable diapers, and disposable kitchen items such as plastic films that are difficult to separate for recycling. Coatings may be applied with PHA as an amorphous latex, which can be reconstituted from the crystalline form of the polymer or obtained directly from a bacterial suspension by dissolving the other cellular macromolecules (cell wall components, DNA, etc.). In the latter case, it is important that the granules are not dried before use as they would become crystalline. The latex may be applied as a spray to a material such as paper, dried, and cross-linked using heat or ultraviolet light. The MCL PHAs are promising as coatings because they are less crystalline than the SCL PHAs (in their crystalline state), and the microorganisms that synthesize them can introduce different functional groups on the molecule to make specific polymers (32,33).

The mechanical properties, morphology, biodegradability, and thermal and crystallization behavior of PHAs blended with nonbiodegradable polymers (such as polyvinyl acetate) and with biodegradable materials (such as wood cellulose fibers and starch) have been reviewed (34). In general, immiscible blends or blends with other biodegradable polymers are more biodegradable than miscible blends or blends with a nondegradable polymer.

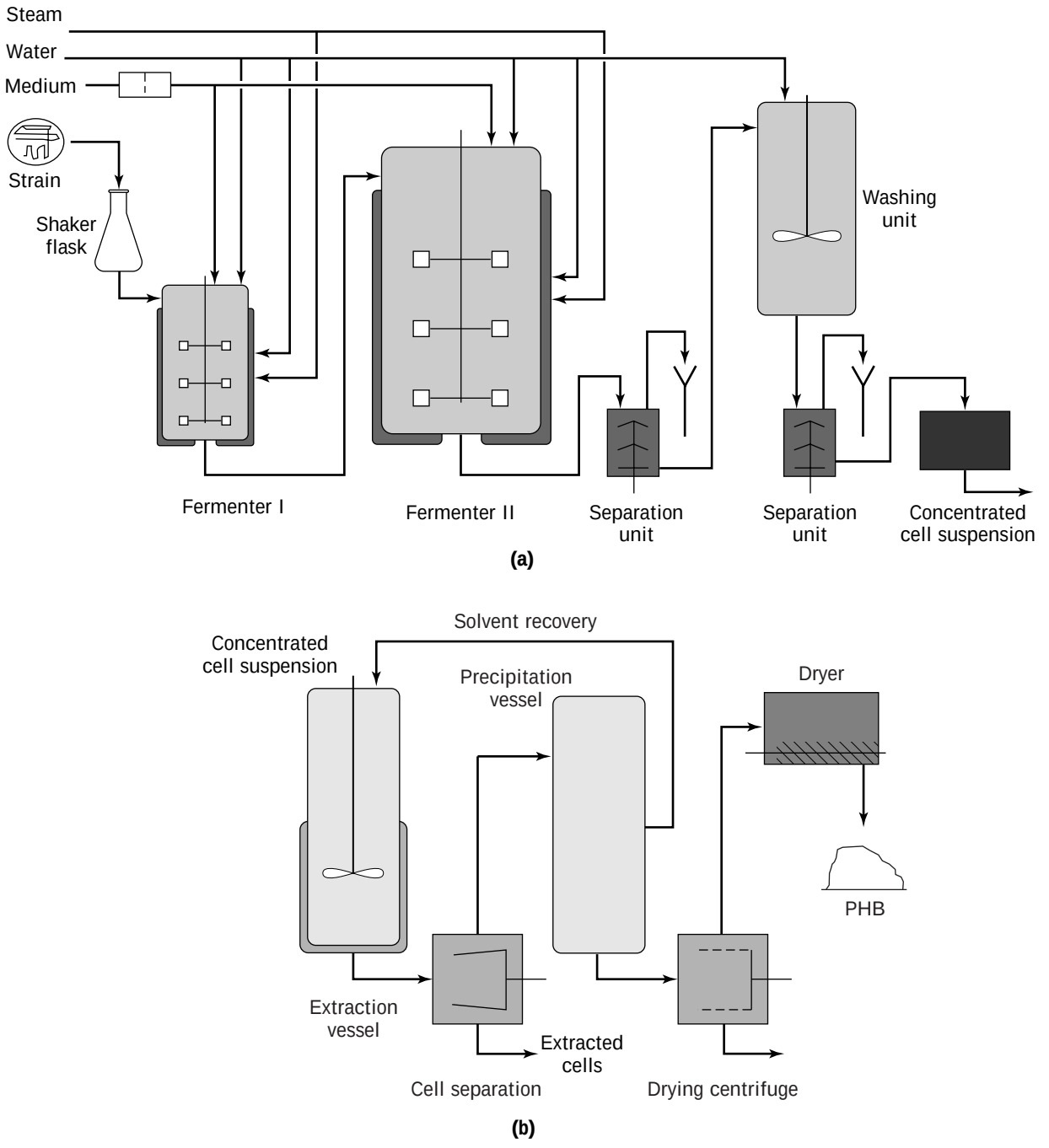


Figure 2. (**a**) Biomer polyhydroxybutyrate (PHB) fermentation process. (**b**) Biomer PHB separation process.

FUTURE PROSPECTS

Economic Considerations

Cost of production is not a major consideration if PHA is to be used exclusively for specialty applications. However, if PHA is to be economically competitive with synthetic plastics, it must be produced on a far larger scale than any other aerobically produced microbial product. An additional problem arises from the fact that PHAs are intracellular products. Thus, the amount that can be produced per cell has physical limitations.

In a typical PHA fermentation process, the cost is divided fairly evenly between substrate cost, capital and operating expenditures in the fermentation process, and the cost of separating the PHA from the microbial biomass. Typically the yield of dry biomass from a hexose sugar such as glucose is about 42% for *R. eutropha*. If the final biomass contains 75% by weight of recoverable PHA, then the yield of PHA from the total amount of substrate supplied is about 32%. Thus, if carbon substrate was available at $500 per metric ton, the cost of the carbon source alone would be $1,560 per metric ton, mak-

ing the product twice as expensive as bulk synthetic plastics.

New methods such as the incorporation of temperature-sensitive plasmids for the production of autolytic enzymes should reduce separation costs. The direct use of PHA-containing bacteria in melt processing (i.e., no separation process apart from water removal) has also been proposed. However, even assuming a negligible cost for the carbon source by using a waste material such as cheese whey permeate, and neglecting the separation costs, studies have shown that PHA produced in bioreactors could not be economically competitive with synthetic bulk plastics in the foreseeable future. The volumetric production rate cannot compete with chemical processes, and this rate is limited by the oxygen transfer rate of present-day bioreactors. New technologies must be developed for commercial PHA production.

New Technologies for the Production of PHA

In vitro synthesis of PHA granules may be commercially feasible. In *R. eutropha*, the maximum rate of PHA synthesis occurs early in the accumulation phase and deteriorates slowly thereafter. One would therefore expect that the maximum amount of PHA accumulation would depend on enzymatic activity or on NADPH or substrate supply. However, the most detailed study on this process points to a physical limitation (30) where polymer synthesis slows to a virtual stop simply because there is no more space available in the cell. If this is indeed the case, then a solution is to produce the polymer outside of the cells as has been achieved previously (35,36). Economic exploitation of this concept would require a source of the biosynthetic enzymes, substrates, NADP, and a method for NADP reduction. This would eliminate separation costs because the product could be used directly as a latex.

It is likely that, in the near future, starch, oilseed, or leguminous plants will be used to produce PHAs. Other plant materials such as cellulose and starch are already used to make plastics or plasticlike materials. Hemicelluloses have been used directly or converted to organic acids for PHA production at a laboratory scale. Thus, any waste from the separation of PHA from plants could be hydrolyzed, fermented to lactic, acetic, and propionic acids, and then fed into bioreactors for the production of specialty PHAs. Polylactides and other polymers of organic acids could be produced in the same plant as could polysaccharides. Plant materials may become a major source of plastic materials in the next century, with PHAs leading the way.

ACKNOWLEDGMENT

We acknowledge Y. Matteau for the preparation of Figure 2.

BIBLIOGRAPHY

1. D.A. Holden and R.H. Marchessault, *Polymer Preprints* **29,** 594–595 (1988).
2. Y. Poirier, D.E. Dennis, K. Klomparens, and C.R. Somerville, *Science* **256**, 520–523 (1993).
3. C. Nawrath, Y. Poirier, and C. Somerville, *Mol. Breed.* **1**, 105–122 (1995).
4. G. Lefebvre, M. Rocher, and G. Braunegg, *Appl. Environ. Microbiol.* **63**, 827–833 (1997).
5. Y. Doi, A. Tamaki, M. Kunioka, and K. Soga, *J. Chem. Soc. Chem. Commun.*, 1635–1636 (1987).
6. G. Braunegg and B. Bogensberger, *Acta Biotechnol.* **5**, 339–345 (1985).
7. B.S. Kim, S.C. Lee, H.N. Chang, Y.K. Chang, and S.I. Woo, *Biotechnol. Bioeng.* **43**, 892–898 (1994).
8. B.S. Kim, S.Y. Lee, and H.N. Chang, *Biotechnol. Lett.* **14**, 811–816 (1992).
9. D. Bourque, Y. Pomerleau, and D. Groleau, *J. Microbiol. Biotechnol.* **44**, 367–376 (1993).
10. G.N.M. Huijberts, H. van der Wal, R.A. Weusthuis, and G. Eggink, Abstract, *5th Int. Symp. on Bacterial Polyhydroxyalkanoates*, Davos, Switzerland, August 1996.
11. U.S. Pat. 4,358,583 (September 11, 1982), J. Walker and J.R. Whitton (to Imperial Chemical Industries, PLC, Great Britain).
12. U.S. Pat. 4,705,604 (November 10, 1987), N. Vanlautem and J. Gilain (to Solvay and Company, Belgium).
13. U.S. Pat. 4,562,245 (December 31, 1985), R.F. Stageman (to Imperial Chemical Industries, PLC, Great Britain).
14. U.S Pat 4,101,533 (July 18, 1978), R.M. Lafferty and E. Heinzle (to Agroferm AG, Chur, Switzerland).
15. J.A. Ramsay, E. Berger, R. Voyer, C. Chavarie, and B.A. Ramsay, *Biotechnol. Tech.* **8**, 589–594 (1994).
16. N. Grassie, E.J. Murray, and P.A. Holmes, *Polym. Degrad. Stabil.* **6**, 95–103 (1984).
17. A.H. de Mola, M. Marx-Figini, and R.V. Figini, *Makromol. Chem.* **176**, 2655–2667 (1975).
18. R. Greibel, Z. Smith, and J.M. Merrick, *Biochemistry* **7**, 3767–3681 (1968).
19. U.S. Pat. 4,910,145 (March 20, 1990), P.A. Holmes and G.B. Lim (to Imperial Chemical Industries, PLC, Great Britain).
20. D. Byrom, in E.A. Dawes ed., *Novel Biodegradable Microbial Polymers*, Kluwer, Dordrecht, Netherlands, 1990, pp. 113–117.
21. C. Lauzier, J.-F. Revol, and R.H. Marchessault, *FEMS Microbiol. Rev.* **103**, 299–310 (1992).
22. J.A. Ramsay, E. Berger, B.A. Ramsay, and C. Chavarie, *Biotechnol. Tech.* **4**, 221–226 (1990).
23. C.A. Lauzier, C.J. Monasterios, I. Saracovan, R.H. Marchessault, and B.A. Ramsay, *Tappi J.* **76**, 71–77 (1993).
24. Eur. Pat. Appl. 46,335 (February 24, 1980), P.A. Holmes and E. Jones (to Imperial Chemical Industries, PLC, Great Britain).
25. U.S. Pat. 3,225,766 (December 28, 1965), J.N. Baptist and J.B. Ziegler (to W.R. Grace and Co., New York).
26. Eur. Pat. 69,497 (April 15, 1987), P.A. Holmes, L.F. Wright, and S.H. Collins (to Imperial Chemical Industries, PLC, Great Britain).
27. F. Merchant, D.E. Dennis, R. Merchant, and C.F. Batrell, Abstract, *4th Int. Symp. on Bacterial Polyhydroxyalkanoates*, Montreal, Canada, August 1994.
28. U.J. Hänggi, in E.A. Dawes ed., *Novel Biodegradable Microbial Polymers*, Kluwer, Dordrecht, 1990, pp. 65–70.
29. D. Byrom, *Int. Biodeterior. Biodegrad.* **31**, 199–203 (1993).
30. D.G.H. Ballard, P.A. Holmes, and P.J. Senior, *Recent Adv. Mech. Synthet. Aspects Polymer.* **215**, 293–314 (1987).
31. *Biopol Nature's plastic: Properties and Processing*, Monsanto, St. Louis, Mo., 1996, pp. 13–15.

32. K. Fritzsche, R.W. Lenz, and R.C. Fuller, *Int. J. Biol. Macromol.* **12**, 85–91 (1990).
33. R.G. Lageveen, G.W. Huisman, H. Preusting, P. Ketelaar, G. Eggink, and B. Witholt, *Appl. Environ. Microbiol.* **54**, 2924–2932 (1988).
34. H. Verhoogt, B. Ramsay, and B. Favis, *Polymer* **35**, 5155–5169 (1994).
35. R.J. Griebel and J.M. Merrick, *J. Bacteriol.* **108**, 782–789 (1971).
36. T.U. Gerngross and D.P. Martin, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6279–6283 (1995).

PROCESS CONTROL, STRATEGY AND OPTIMIZATION

D. DOCHAIN
Université Catholique de Louvain
Louvain-La-Neuve, Belgium

KEY WORDS

Automatic control
Bioprocess
CSTR
Fed-batch reactor
Nonlinear control
Optimal control
Optimization
Parameter estimation
Software sensor
State estimation

OUTLINE

INTRODUCTION

Industrial-scale biotechnological processes have progressed vigorously over the past decades. Generally speaking, the problems arising from the implementation of these processes are similar to those of more classical industrial processes, and the need for monitoring systems and automatic control to optimize production efficiency, improve product quality, or detect disturbances in process operation is obvious. Nevertheless, automatic control of industrial biotechnological processes is clearly developing very slowly. There are two main reasons for this slow development:

1. The internal working and dynamics of these processes are as yet poorly grasped, and many problems of methodology in modeling remain to be solved. It is difficult to develop models that take into account the numerous factors that can influence the specific bacterial growth rate. The modeling effort is often tedious and requires many experiments to produce a reliable model. Reproducibility of experiments is often uncertain because of the difficulty in obtaining the same environmental conditions. Moreover, because these processes involve living organisms, their dynamic behavior is strongly nonlinear and nonstationary. Model parameters cannot remain constant over a long period; they vary, for example, due to metabolic variations of biomass or to random and unobservable physiological or genetic modifications. It should also be noted that the lack of accuracy of the measurements often leads to identifiability problems.
2. Another essential difficulty lies in the absence, in most cases, of cheap and reliable instrumentation suited to real-time monitoring. To date, the market offers very few sensors capable of providing reliable on-line measurements of the biological and biochemical parameters required to implement high-performance automatic control strategies. The main variables (i.e., biomass, substrate, and synthesis product concentrations) generally need to be determined through laboratory analyses. The cost and duration of the analyses obviously limit the frequency of the measurements.

Figure 1 shows a schematic view of a computer-controlled bioreactor. In the illustrated situation, the influent flow rate is modulated via a control algorithm implemented in the computer. As explained later, the control algorithm combines different informations about the process provided by on-line measurements and some knowledge about the process dynamics. The on-line measurements may undertake some mathematical treatment via software sensors to provide extra information about some key process parameters (such as specific growth rates) and unmeasured variables.

The article is organized as follows. The basic concepts of automatic control of bioprocesses, with a particular emphasis on the notions of optimization, stability, and disturbances, are discussed first. Then I introduce the basic in-

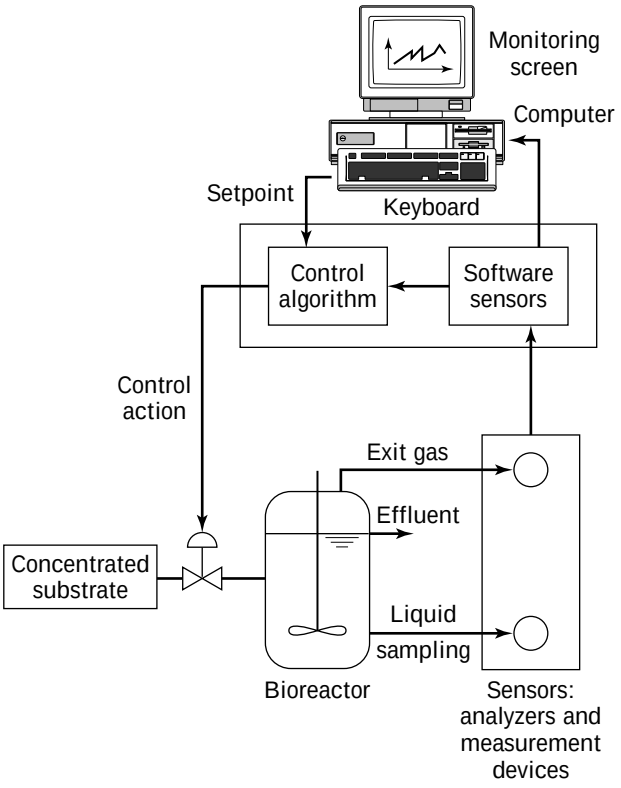


Figure 1. Schematic view of a computer-controlled bioreactor.

gredients of automatic control (dynamical model; feedback; proportional, integral, and feedforward actions), and briefly compare linear and nonlinear control and adaptive and nonadaptive control.

BIOPROCESS CONTROL: BASIC CONCEPTS

The objective of automatic control when applied to any process (including bioprocesses) is to run and maintain the process in stable, optimal operating conditions in spite of disturbances.

Let us first give some more insights about the concepts that are covered by the words *stable*, *optimal*, and *disturbances*. And let us consider the following simple example to illustrate these concepts: a simple microbial growth reaction in a stirred tank reactor. The microbial growth reaction is characterized by the following reaction scheme:

$$S \rightarrow X \tag{1}$$

with S the limiting substrate and X the biomass. Mass balances for both components S and X lead to the following set of dynamic equations:

$$\frac{dS}{dt} = -DS - \frac{1}{Y_{X/S}}\mu X + DS_{\text{in}} \tag{2}$$

$$\frac{dX}{dt} = -DX + \mu X \tag{3}$$

where D is the dilution rate (h^{-1}) (i.e., the ratio of the influent flow rate F_{in}[L/h] over the volume V[L] occupied by the reacting medium in the reactor), $Y_{X/S}$ is the yield coefficient, μ is the specific growth rate (h^{-1}), and S_{in} is the influent substrate concentration (g/L). Note the following:

1. The model is a dynamic model in order to capture the time evolution of the process: this is an essential ingredient of automatic control (discussed in more detail later).
2. The model is valid whether the operation mode is continuous (then the reactor is sometimes called a chemostat or a continuous stirred tank reactor [CSTR]), batch (then D is equal to zero), fed batch (then the volume is not constant and its time variation is given by the following equation: $dV/dt = F_{\text{in}}$).
3. We use the same notation for the components (in reaction scheme 1) and their concentration (in equations 2 and 3).

Let us now consider that the specific growth rate here is of the Haldane type:

$$\mu = \frac{\mu^* S}{K_S + S + S^2/K_I} \tag{4}$$

with K_S is the Monod constant, K_I is the inhibitor constant, and μ^* is a constant connected to the maximum specific growth rate $\mu_{\max}$ as follows:

$$\mu^* = \mu_{\max}\left(1 + 2\sqrt{\frac{K_S}{K_I}}\right) \tag{5}$$

Let us also consider that the objective is to control the process by using the dilution rate D (or, more precisely, the influent flow rate F_{in}) or the influent substrate concentration S_{in} as the control action.

Disturbances

Let us illustrate the notion of disturbances with the preceding example. If the dynamic model (equations 2–4) is assumed to be perfect, and if the dilution rate D is the only manipulated input, then the influent substrate concentration S_{in} is a disturbance: variations of S_{in} will disturb the process by inducing variations in the time evolution of the variables S and X.

Assume now that the dynamic model is not perfect, for example, that some of the parameters μ^*, K_S, and K_I of the Haldane model (equation 4) are changing with time (e.g., due to microorganism adaptation): these variations are also disturbances with regard to the considered process model. Another similar disturbance may come from the presence of a side reaction (e.g., maintenance or microorganism death) that has been neglected in the earlier dynamic model but that may not be negligible under some operating conditions. Then, in case of maintenance, for instance, the true dynamic model of the process is described by equation 3 and the following modified mass balance equation for S:

$$\frac{dS}{dt} = -DS - \frac{1}{Y_{X/S}}\mu X - k_m X + DS_{in} \qquad (6)$$

where k_m is the maintenance coefficient.

Finally, typical disturbances may be due to process failures (e.g., actuator, sensor, or even stirring system). Even if those disturbances are inherently different, it is important to consider all the disturbances that are likely to take place during the process operation when analyzing and testing the performances of a designed controller that has to be implemented on a real bioprocess.

Stability

The notion of stability has been largely considered and studied in dynamic systems and automatic control and has led to different mathematical definitions and tests (see Ref. 1). Here we shall concentrate on Lyapunov stability. Qualitatively the notion of stability can be roughly explained as follows. Let us consider that the process is at some equilibrium point (or steady state). The equilibrium point will be stable if the process does not go far from this state for small deviations. It will be unstable if the variables of the process move away from the equilibrium point and if their variations become larger and larger as time increases. Stability is obviously of vital concern in automatic control: it is essential to design controllers that keep processes in stable conditions on the one hand and that are capable of stabilizing unstable processes on the other hand.

Let us illustrate the concept of stability on the microbial growth example and how to test it. The Haldane model is represented in Figure 2. What we shall now show is that values of the substrate concentration on the left-hand side of the maximum specific growth rate

$$\mu^* \Big/ \left(1 + 2\sqrt{\frac{K_S}{K_I}}\right)$$

correspond to stable equilibrium points, whereas the ones on the right-hand side correspond to unstable equilibrium points.

Equilibrium Points. An equilibrium point is, by definition, a constant state (denoted here $[\bar{S}, \bar{X}]$) that satisfies the equation of the dynamic model, the following algebraic equations in our example:

$$\frac{dS}{dt} = 0 \Rightarrow -\bar{D}\bar{S} - \frac{1}{Y_{X/S}}\mu(\bar{S})\bar{X} + \bar{D}\bar{S}_{in} = 0 \qquad (7)$$

$$\frac{dX}{dt} = 0 \Rightarrow -\bar{D}\bar{X} + \mu(\bar{S})\bar{X} = 0 \qquad (8)$$

for given constant values of $\bar{D}$ and $\bar{S}_{in}$. If we consider the Haldane model (Fig. 2), it is straightforward to check that there are three possible equilibrium points:

1. Operational equilibrium point 1: $\bar{S}_1 \leq \sqrt{K_S K_I}$, $\bar{X}_1 = Y_{X/S}(\bar{S}_{in} - \bar{S}_1)$
2. Operational equilibrium point 2: $\bar{S}_2 > \sqrt{K_S K_I}$, $\bar{X}_2 = Y_{X/S}(\bar{S}_{in} - \bar{S}_2)$
3. Washout equilibrium point: $\bar{X}_3 = 0$, $\bar{S}_3 = \bar{S}_{in}$.

The washout may occur for any value of $\bar{D}$ and $\bar{S}_{in}$. It is also the only possible equilibrium point for $\bar{D} > \mu_{max}$. It is obviously an undesirable steady state. The values of $\bar{S}$ of the last two equilibrium points correspond indeed to the two solutions of the second-order algebraic equation obtained by introducing the Haldane equation (equation 4) into the equilibrium point equation of X (equation 8):

$$\frac{\bar{D}}{K_I}\bar{S}^2 + (\bar{D} - \mu^*)\bar{S} + \bar{D}K_S = 0 \qquad (9)$$

Stability Analysis. For the stability analysis, let us consider the linear approximation of the model (equations 2–4) around the equilibrium points (i.e., the linearized tangent model). Whatever the input u (D, S_{in}, or both), the linearized tangent model around each operational equilibrium point is written as follows:

$$\frac{dx}{dt} = Ax + Bu \qquad (10)$$

where x is the vector of the deviations of S and X from the equilibrium values

$$x = \begin{bmatrix} S - \bar{S} \\ X - \bar{X} \end{bmatrix} \qquad (11)$$

and the matrix A is given by the following expression:

$$A = \begin{bmatrix} 0 & \Omega \\ -\dfrac{1}{Y_{X/S}}\bar{D} & -\dfrac{1}{Y_{X/S}}\Omega - \bar{D} \end{bmatrix} \qquad (12)$$

with

$$\Omega = \frac{\mu^*\bar{X}\left(K_S - \dfrac{\bar{S}^2}{K_I}\right)}{\left(K_S + \bar{S} + \dfrac{S^2}{K_I}\right)^2} \qquad (13)$$

The stability of the equilibrium points is determined by the

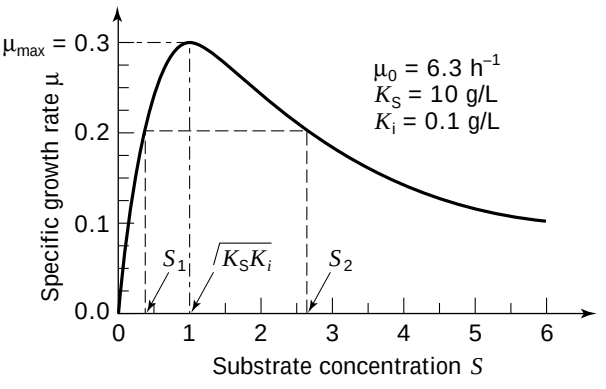


Figure 2. The Haldane model.

eigenvalues of the matrix A: the equilibrium point will be stable if the real parts of all the eigenvalues are negative, it will be unstable if one eigenvalue has a positive real part. In our example, the eigenvalues are equal to $(-\bar{D})$ and $(-1/Y_{X/S}\Omega)$. If the first eigenvalue is always negative, the second one is negative for the first equilibrium point ($\bar{S}_1$, $\bar{X}_1$) and is positive for the second one ($\bar{S}_2$, $\bar{X}_2$). This means that ($\bar{S}_1$, $\bar{X}_1$) is a stable equilibrium point and that ($\bar{S}_2$, $\bar{X}_2$) is an unstable one.

Regulation versus Tracking

Two typical situations are usually considered in automatic control: constant setpoint control (regulation) and trajectory tracking. Regulation is typical of the control of continuous reactors, where the objective is usually to maintain the process in an a priori chosen steady state despite the disturbances. But in many instances, such as in batch and fed-batch reactors but also for start-up of reactors and in case of production grade changes, regulation is not appropriate: the objective is then to drive the process from some initial state to some final desired state by following some predetermined trajectory. In the simple microbial growth example, if we consider a fed-batch reactor and if the objective is to optimize the production of biomass, then an optimal trajectory is an "exponential" profile for X: the controller must then be designed so as to maintain the process as close as possible to the optimal profile.

Optimization

The optimal conditions in which the controller is supposed to run the bioprocess are typically defined so as to optimize production efficiency, improve product quality, or improve the economic profit of the plant. Optimization is obviously a helpful tool for determining the optimal operating conditions. Static optimization techniques can be used for the determination of optimal constant setpoints, whereas dynamic optimal control techniques will be useful for determining the best operating profile.

It must be stressed, however, that the results obtained via optimization have to be handled with great care. The reason lies in the high sensitivity of the optimization techniques on the quality of the process model (including the model structure, i.e., its equations, and the precision on the values of the parameters of the chosen model). Indeed the models used for describing the dynamics of bioprocesses (often mainly more specifically the kinetics) are typically characterized by a large uncertainty (as already mentioned).

Let us illustrate these ideas with our simple microbial growth example with Haldane kinetics. Assume that only the substrate concentration S is accessible for on-line measurement (and will therefore be used for feedback control) and that the objective is to maximize the production of biomass X.

In a CSTR, the objective is to operate the process in a steady state that maximizes the productivity (i.e., the quantity of biomass per unit time). This means that the objective function is

$$J = \bar{D}\bar{X} \tag{14}$$

The solution of the optimization requires the knowledge of the process model in steady state (equations 4, 7, and 8) (equality constraints), as well as (physical) inequality constraints on the process variables

$$0 \le \bar{S} \le \bar{S}_{in},\ 0 \le \bar{X} \le Y_{X/S}\bar{S}_{in} \tag{15}$$

The value of the substrate concentration that maximizes J

$$J = \bar{D}\bar{X} = \frac{\mu^*\bar{S}}{K_S + S + S^2/K_I} Y_{X/S}(\bar{S}_{in} - \bar{S}) \tag{16}$$

is obtained by calculating the derivative of J with respect to $\bar{S}$ and by setting its value to zero

$$\frac{dJ}{dS} = 0 \rightarrow \bar{S}_{optimal} = \frac{-K_S K_I + K_I\sqrt{K_S^2 + \bar{S}_{in}(\bar{S}_{in} + K_I)}}{S_{in} + K_I} \tag{17}$$

and then by checking that the sign of the second-order derivative is negative (to have a maximum):

$$\frac{d^2J}{dS^2}(\bar{S}_{optimal}) < 0 \Rightarrow \bar{S}_{optimal} \text{ is a maximum} \tag{18}$$

and that the inequality constraints are fulfilled. This optimization problem was originally treated by D'ans et al. (2). (Indeed there is another value of $\bar{S}$ for which dJ/dS is equal to zero, but it is negative and therefore obviously rejected.)

In a fed-batch reactor, the optimization problem is somewhat different. Assume that the time of the fed-batch reactor operation is fixed. Then the objective is to obtain the highest value possible at the end of the fermentation. The objective function is then

$$J = X(t_f)V(t_f) \tag{19}$$

Intuitively, it is obvious that the optimum will be obtained when the biomass is produced at its maximum production rate, that is, if the value of substrate concentration is kept at the value that maximizes the specific growth rate

$$S_{optimal} = \sqrt{K_S K_I}$$

This result can be formally obtained as follows by considering optimal control theory.

The general problem of optimal control is to find an optimal trajectory of the control input $u(t)$ that drives the process from its initial state to its final state and at the same time minimizes a cost function

$$J = g[x(t_f),\ t_f] + \int_{t_0}^{t_f} h[x(t),\ u(t),\ t]dt \tag{20}$$

where x is the state vector, t_0 and t_f are the initial and final

time, respectively, and the process dynamics are described by the following ordinary differential equation:

$$\frac{dx}{dt} = f[x(t), u(t), y] \tag{21}$$

and the input $u(t)$ obeys the following practical constraints:

$$U_{\min} \leq u(t) \leq U_{\max} \tag{22}$$

The optimal control solution is based on the Hamiltonian H:

$$H \triangleq \lambda(t)^{\mathrm{T}}(f_1(x) + b(x)u) = \varphi + \psi u \tag{23}$$

where λ represents the adjoint variables, or costates. They are the solution of the following differential equation:

$$\frac{d\lambda}{dt} = \frac{\partial H}{\partial x} \tag{24}$$

For time-invariant systems linear in the input

$$\frac{dx}{dt} = f_1(x) + b(x)u \tag{25}$$

and the cost function J without integral term ($J = g[x(t_f), t_f]$), it can be shown that the optimal control has the following form:

$$u(t) = U_{\min} \quad \text{if } \psi < 0 \tag{26}$$

$$= U_{\max} \quad \text{if } \psi > 0 \tag{27}$$

$$= u_{\text{sing}} \quad \text{if } \psi = 0 \tag{28}$$

The optimal control solution consisting of applying only $U_{\min}$ and $U_{\max}$ is called bang-bang control. The so-called singular control input u_{sing} is obtained by computing the first- and second-order time derivatives of ψ and equating them to zero.

In our fed-batch reactor example, assume that the control input is F_{in}. The dynamics of the process are given by equations 2 and 3 plus the volume equation ($dV/dt = F_{\text{in}}$) and a Haldane specific growth rate model (equation 4). Then the Hamiltonian H is given by the following expression:

$$H = \lambda_1\left(-\frac{F_{\text{in}}}{V} S - \frac{1}{Y_{X/S}} \mu X + \frac{F_{\text{in}}}{V} S_{\text{in}}\right) + \lambda_2\left(-\frac{F_{\text{in}}}{V} X + \mu X\right) + \lambda_3 F_{\text{in}} \tag{29}$$

It is routine to check that $d\psi/dt = 0$ if $\partial\mu/\partial S = 0$ (i.e., u_{sing} corresponds to the value that maximizes μ). If we choose the initial value of the substrate concentration $S(t_0)$ corresponding to the one that maximizes μ

$$(S(t_0) = \sqrt{K_S K_I})$$

the optimal trajectory is obtained by applying the singular control input from t_0 to t_f; that is,

$$F_{\text{in,sing}} = \frac{\mu^* X V \sqrt{K_S K_I}}{Y_{X/S}(2K_S + \sqrt{K_S K_I})(S_{\text{in}} - \sqrt{K_S K_I})} \tag{30}$$

with X and V computed via the mass balance equations:

$$\frac{dX}{dt} = \frac{\mu_{\max}\sqrt{K_S K_I}}{2K_S + \sqrt{K_S K_I}} X - \frac{F_{\text{in}}}{V} X \tag{31}$$

$$\frac{dV}{dt} = F_{\text{in}} \tag{32}$$

Figure 3 illustrates the high sensitivity of optimal control to modeling errors in numerical simulation. We have considered the fed-batch example with the following model parameter and initial values:

$$\mu^* = 5\text{h}^{-1},\ K_S = 10 \text{ g/L},\ K_I = 0.1 \text{ g/L},\ Y_{X/S} = 1 \tag{33}$$

$$S(0) = 0.9 \text{ g/L},\ X(0) = 0.1 \text{ g/L},\ V(0) = 10 \text{ L},\ V(t_f) = 20 \text{ L},\ S_{\text{in}} = 10 \text{ g/L} \tag{34}$$

We have also considered the optimal control when 10% error is made on the value that maximizes μ; that is, the operator believes that the optimal substrate concentration is equal to 9 (instead of 10) g/L. This has been done by setting the saturation constant K_S to 8.1 (instead of 10) g/L in the optimal control law (equation 30). Note that even in the presence of a reasonable model error, the optimal control is moving far away from the optimal substrate concentration.

Let us mention among the large body of scientific literature on optimization and optimal control the following references: Ref. 3, which the treats the optimization of chemical processes, and Ref. 4, which covers the essential items of the optimal control of bioprocesses and treats it then in an original manner.

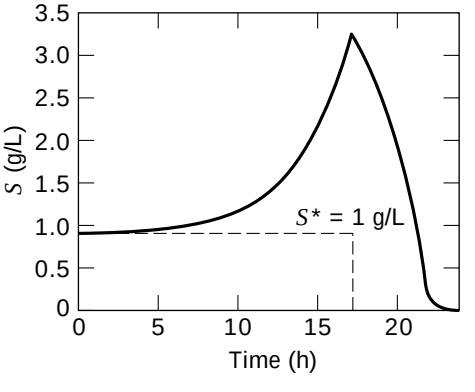


Figure 3. Sensitivity of the optimal control of a fed-batch reactor to modeling errors.

BIOPROCESS CONTROL: BASIC INGREDIENTS

Dynamic Model

The first essential basic ingredient for the design of automatic control algorithms is a dynamic model of the process. The need of dynamic models for the control design is motivated by the intrinsic need to handle transients for transferring the process from one state to another or from one setpoint to another, for following a desired trajectory, or for rejecting or reducing the effect of disturbances. In the latter, disturbances influence the dynamic behavior of the process: knowledge of the process dynamics, and in particular of the dynamic effect of the disturbance, is helpful in selecting the appropriate strategy to reduce and possibly reject their negative influence on the process operation.

The dynamic models to be considered for control design can take different forms. The preceding sections suggest that the most obvious model forms are the mass balance models. But many other forms are possible, including linear models deduced, for example, by linearization of the mass balance models or by identification from input–output data (e.g., Ref. 5) (such models are also called black-box models). The input–output models may be nonlinear; neural network models are presently a quite popular version of nonlinear black-box models (e.g., Ref. 6). A possibly attractive form of dynamic models of bioprocesses are hybrid models, which are an intermediate form between mass balance and neural network models (e.g., Ref. 7). In the following, mainly for conciseness and pedagogical reasons, we concentrate on mass balance models.

Feedback

Another essential basic ingredient of automatic control is the introduction of a feedback loop: On-line (real-time) information about the process (typically on-line measurement of the variable to be controlled) is provided to the controller, which then provides the control action to the process. The concept of feedback loop is illustrated in Figure 4.

Proportional Action

Most feedback control algorithms (except very simple controllers such as on-off controllers) modulate the amplitude of the control action proportionally to the control error (i.e., the difference between the desired value of the controlled variable and its measured value), which is known as the proportional action.

Integral Action

The proportional action is usually not sufficient to guarantee a control error equal to zero in steady state: this motivates the introduction in the control scheme of an integral action (i.e., a term proportional to the integral of the control error).

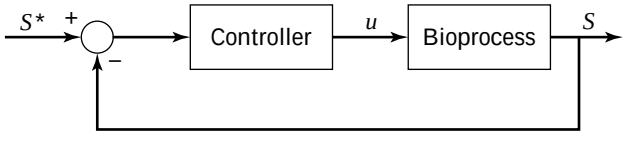


Figure 4. Feedback loop for the control of S in a bioprocess.

Feedforward Action

If the effects of a disturbance on the bioprocess dynamics are known, the controller may be designed by also incorporating a feedforward action to anticipate the possibly negative effects of the disturbance. We shall illustrate in the following how to incorporate feedforward action when the influent substrate concentration S_{in} is a measured disturbance.

Linear versus Nonlinear Control

The controller can be designed on the basis of either a linear or nonlinear model of the process. As earlier, let us illustrate the design with a simple example: the control of the substrate concentration S by acting on the dilution rate D (or more precisely the influent flow rate F_{in}) in a stirred tank reactor with a simple microbial growth reaction.

Let us start with the first option: linear control. The control design is based on a linear model of the process (e.g., the linearized tangent model [equations 10 and 11] around some steady state $[\bar{S}, \bar{X}, \bar{D}, \bar{S}_{in}]$. If we consider the control of S at a prescribed value S^* (e.g., equal to $\bar{S}$), a linear proportional integral (PI) regulator will then typically have the following form (see, e.g., Ref. 8):

$$D = \bar{D} + K_P\left[(S^* - S) + \frac{1}{\tau_i}\int_0^t (S^* - S(\tau))d\tau\right], \quad K_P > 0,\ \tau_i > 0 \qquad (35)$$

where K_P is the controller gain and τ_i is the integral time, or reset time.

Indeed a similar PI regulator can be computed on the basis of a nonlinear model of the process (e.g., the mass balance equation of the substrate concentration S) (2). The control algorithm will then be written as follows:

$$D = \frac{C_1(S^* - S) + C_2\int_0^t (S^* - S(\tau))d\tau - k_1\mu X}{S_{in} - S}, \quad C_1 > 0,\ C_2 > 0 \qquad (36)$$

The preceding controller equation could have been obtained by combining the mass balance equation, equation 2, with the following desired closed-loop second-order dynamic equation:

$$\frac{d(S^* - S)}{dt} = -C_1(S^* - S) - C_2\int_0^t (S^* - S(\tau))d\tau \qquad (37)$$

for S^* constant. (Indeed by differentiating equation 37 once with respect to time, one obtains the usual second-order linear equation $d^2x/dt^2 + C_1 dx/dt + C_2 x = 0$ with $x = S^* - S$.) In simple terms, equation 37 imposes that if at some time instant the substrate concentration S is different from its desired value S^*, the controller will force S

to converge to S^* with a rate determined by the design parameters C_1 and C_2.

Let us present at this point the advantages and drawbacks of the approaches introduced. Before doing so, it is important to note that the second algorithm (equation 36) is more sophisticated than the first one (equation 35). In particular, the second controller requires the knowledge (on-line measurement) of the biomass concentration X and of the influent substrate concentration S_{in} and the knowledge of the specific growth rate model. When comparing both approaches, it is important to keep these differences in mind.

Linear Control. The linear controller, if it is correctly tuned (see, e.g., Ref. 8 for tuning rules), will most probably do a good job if it is used for regulation around a steady state. It also has the advantage to be easily implemented via standard industrial PI regulators. However, its performance (in particular, its transient performance) might be degrading if the process is moving away from the steady state for which it has been calibrated and tuned. This typically happens for start-up and grade changes in CSTRs and might be even more critical for batch and fed-batch reactors in which the process state typically follows large variations (just think of the "exponential" growth of the biomass in this type of reactor). This is illustrated in Figure 5, where the PI regulation of S has been numerically simulated with the same model as in Figure 3 and under the following conditions:

$$S^* = 1 \text{ g/L}, K_p = 2, \tau_i = 0.5, S(0) = 0 \text{ g/L} \quad (38)$$

The values of the design parameters of the PI have been chosen so as to correspond to a closed-loop dynamics close to the mean open-loop dynamics. We note that the controller gain is too high at the beginning (and induce large variations of the controlled output S) and then too small (the controller is not able to maintain the controlled output S to its desired value during the second part of the "exponential" growth phase).

Nonlinear Control. The main advantage of the nonlinear controllers is directly related to the mentioned drawback of the linear controllers: if correctly designed and tuned, it will be able to handle a larger spectrum of the operating domain of the bioprocess. In the example, the main source of nonlinearity is the term $k_1\mu X$. The presence of this term in the controller is particularly important in the control of fed-batch reactors; as pointed out by Axelsson (9), the integration of a term that characterizes the "exponential" biomass growth is a key factor for the efficiency of the control action. The other term, $S_{in} - S$, also plays an important role; if S_{in} is a disturbance in the context of the control of the bioprocess, then this term is indeed a feedforward term that allows anticipation of the effect of variations of S_{in} on the performance of the closed-loop system (process plus controller).

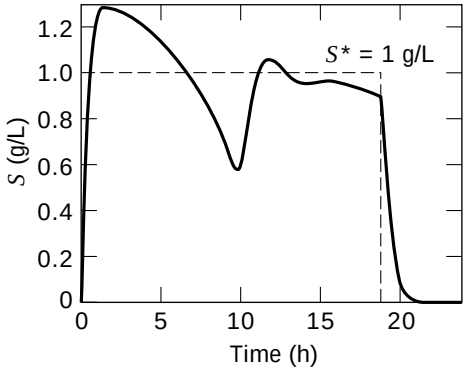


Figure 5. Numerical simulation of a linear proportional integral control of a fed-batch reactor.

However, first the implementation of the control algorithm requires a computer instead of a standard industrial PI (which today can hardly be considered a drawback). Secondly (as for any controller), the performance depends on the quality and reliability of the model on which the design has been based. In other words, in the given example, this means that we assume that the hydrodynamics in the reactor are fairly represented by completely mixed conditions and that the simple growth reaction is largely dominant with regard to possible other side reactions. Besides the implementation as such of the controller, equation 36 requires the values of X and also of S_{in}. In many cases one may expect that if S is measured on-line, it should not be difficult to also measure S_{in} (in many instances, you do not even need to measure it because the influent substrate concentration is known by user's choice). But the situation is usually different for the biomass concentration: very often X is not measured on-line, or, to say it differently, S and X are not often available at the same time for on-line measurement. Finally, the nonlinear controller (equation 36) relies on the knowledge of (the model of) the specific growth rate μ; as suggested previously, this is a major source of uncertainty in bioprocess models.

Adaptive versus Nonadaptive Control

One possible way to handle the previously mentioned difficulties (lack of on-line measurements, uncertainties on the kinetics models) is to use adaptive (linearizing) control (see Ref. 10). In the example, this means that the control law will be implemented by considering on-line estimates of X and μ. An on-line estimate of the biomass concentration can be provided by an asymptotic observer (for a systematic design, see Ref. 10). In our example, it will take the following form. We first define an auxiliary variable Z:

$$Z = S + \frac{1}{Y_{X/S}} X \quad (39)$$

The dynamics of Z are readily derived from the mass balance equations of X and S:

$$\frac{dZ}{dt} = DS_{in} - DZ \quad (40)$$

The on-line estimate $\hat{X}$ of X is then given by computing Z from equation 40 and by rewriting equation 39:

$$\hat{X} = Y_{X/S}(Z - S) \tag{41}$$

or equivalently, one can replace the term $X/Y_{X/S}$ in the controller equation with $Z - S$.

Let us now consider the on-line estimation of μ. Different methods can be used to give an on-line estimate $\hat{\mu}$ of μ. One possible estimator is the observer-based estimator (see Ref. 10), which specializes in our example as follows:

$$\frac{d\hat{S}}{dt} = -\hat{\mu}(Z - S) + DS_{in} - DS + \gamma_1(S - \hat{S}), \gamma_1 > 0 \tag{42}$$

$$\frac{d\hat{\mu}}{dt} = -\gamma_2(S - \hat{S}), \gamma_2 > 0 \tag{43}$$

$\hat{S}$ is an estimate of S whose role is to drive the estimation of the unknown value of the specific growth rate μ on-line; this is achieved by comparing $\hat{S}$ with the measured value S (this is the role of the correction terms $S - \hat{S}$ in equations 42 and 43). γ_1 and γ_2 are design parameters that have to be calibrated to obtain the best on-line estimates (10).

Because the parameter estimation introduces an integral action in the control loop (10,11), the term $C_2\int_0^t(S^* - S(\tau))d\tau$ in the control law (equation 36) is not necessary anymore. Then, in our example, the adaptive linearizing controller is given by the following equation:

$$D = \frac{C_1(S^* - S) - \hat{\mu}(Z - S)}{S_{in} - S}, C_1 > 0 \tag{44}$$

with Z given by equation 40 and $\hat{\mu}$ given by 42 and 43.

Different applications of adaptive linearizing control to bioprocesses have been described (12–15). One of the key features of adaptive linearizing can be illustrated on the basis of the example; it allows incorporation of the well-known characteristics of the process dynamics while keeping the usual features of classical controllers:

- Proportional action via the term $C_1(S^* - S)$
- Integral action via the adaptation mechanism (equations 42 and 43)
- Feedforward action via the presence of S_{in}; the controller is capable of anticipating the effect of a variation of the influent substrate concentration; for instance, an increase of S_{in} results in a decrease of the control input D inversely proportional to this increase.

As for the nonadaptive, nonlinear PI, the controller contains a state estimate via the term $k_1\mu X$ (or more precisely $\hat{\mu}[Z - S]$). If the process is in good working condition, in particular if the biomass is in a good state, then it is possible to treat large amounts of substrate, because the control action D is proportional to X.

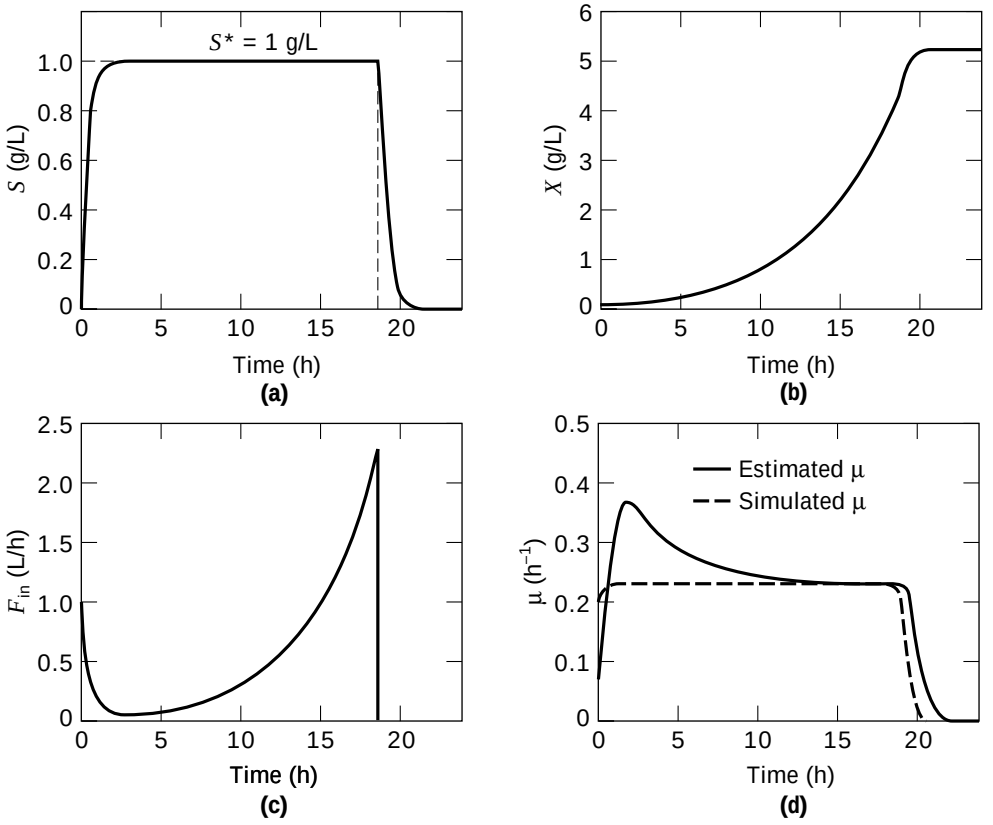


Figure 6. Adaptive linearizing controller of a fed-batch reactor. (**a**) Controlled output S; (**b**) biomass concentration X; (**c**) control input F_{in}; (**d**) specific growth rate μ.

Finally, note that besides the integral action, the estimation of physical parameters (here μ) has the further advantage of giving useful information that can be used for monitoring the process and possibly also for analyzing the internal working of the process.

Figure 6 illustrates the performance in numerical simulation of the adaptive linearizing controller. The same model parameters as in Figures 3 and 5 have been used. The following design parameters and initial values have been considered:

$$S^* = 1 \text{ g/L}, C_1 = 2, \gamma_1 = 5, \gamma_2 = \frac{\lambda_1^2}{4(Z - S)} \quad (45)$$

$$\hat{\mu}(0) = 0.1 \text{ h}^{-1}, \hat{S}(0) = S(0), Z(0) = S(0) + 0.5X(0), S(0) = 0.5 \text{ g/L} \quad (46)$$

The initial value of S has been set to 0.5 g/L, which allows the opportunity to see that the controller has no problem forcing S to converge to its desired value (see Fig. 6a). Figures 6b and 6c exhibit the exponential profile of the biomass X and the control input F_{in}. Finally, Figure 6d gives the on-line estimation of the specific growth rate μ.

Other Approaches

The (adaptive) nonlinear controller presented in the preceding section belongs to a large class of control strategies, often labeled model-based control. This includes control approaches such as model predictive control, in which constraints (such as input constraints) may be handled explicitely in the design (8). We have restricted the presentation to mass balance models, but obviously any dynamic model of the bioprocess (e.g., hybrid model), as long as it is reliable, can be considered in the control design.

CONCLUSIONS

In this article we have first concentrated on the basic concepts of automatic control of bioprocesses, with a particular emphasis on the notions of optimization, stability, and disturbances. In a second step the basic ingredients of automatic control (dynamic model, feedback, proportional, integral, and feedforward actions) have been introduced, and linear and nonlinear control and adaptive and nonadaptive control have been compared.

An appealing avenue for future applications in bioprocesses is an integrated process design approach, which is already largely used in other areas such as aeronautics and that implies the integration of the control strategies in the design of the process from the very beginning (16).

ACKNOWLEDGMENTS

This article presents research results of the Belgian Programme on Inter-University Poles of Attraction initiated by the Belgian State, Prime Minister's office for Science, Technology, and Culture. The scientific responsibility rests with its author.

BIBLIOGRAPHY

1. J.L. Willems, *Stability Theory of Dynamical Systems*, Nelson, London, 1970.
2. G. D'Ans, D. Gottlieb, and P. Kokotovic, *Automatica* **8**, 729–736 (1972).
3. T.F. Edgar and D.M. Himmelblau, *Optimization of Chemical Processes*, McGraw-Hill, New York, 1988.
4. J. Van Impe, Ph.D. Thesis, Katholieke Universiteit, Leuven, Belgium, 1993.
5. L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, Englewood Cliffs, N.J., 1987.
6. G. Montague and J. Morris, *TIBTECH* **12**, 312–323 (1994).
7. F. Feyo De Azevedo, B. Dahm, and F.R. Oliviera, *Comput. Chem. Eng.* **21**, S751–S756 (1997).
8. D.E. Seborg, T.F. Edgar, and D.A. Mellicamp, *Process Dynamics and Control*, Wiley, New York, 1989.
9. J.P. Axelsson, Ph.D. Thesis, Lund Institute of Technology, Lund, Sweden, 1989.
10. G. Bastin and D. Dochain, *On-Line Estimation and Adaptive Control of Bioreactors*, Elsevier, Amsterdam, 1990.
11. M. Perrier and D. Dochain, *International Journal of Adoptive Control and Signal Processing* **7**, 309–321 (1993).
12. L. Chen, G. Bastin, and V. Van Breusegem, *Automatica* **31**, 55–65 (1995).
13. Y. Pomerleau, M. Perrier, and D. Bourque, in A. Monack and K. Schogerl eds., *Proc. 6th Int. Conf. Computer Appl. Biotechnol.*, Pergamon Press, Oxford, 1995, pp. 107–112.
14. Y. Pomerleau and G. Viel, in N.M. Karim and G. Stephanopoulos eds., *Proc. 5th Int. Conf. Computer Appl. Biotechnol.* Pergamon Press, Oxford, 1992, pp. 315–318.
15. P. Renard, D. Dochain, G. Bastin, H. Naveau, and E.J. Nyns, *Biotechnol. Bioeng.* **31**, 287–294 (1988).
16. J. Harmand, Ph.D. Thesis, Université de Perpignan, Perpignan, France, 1997.

See also AMMONIA TOXICITY, ANIMAL CELLS; MEDIUM FORMULATION AND DESIGN; PILOT PLANTS: DESIGN AND OPERATION; PROCESS MONITORING; SECONDARY METABOLITE PRODUCTION, ACTINOMYCETES, OTHER THAN STREPTOMYCES.

PROCESS MONITORING

A.P.F. TURNER
S.F. WHITE
Cranfield Biotechnology Centre
Bedfordshire, United Kingdom

KEY WORDS

Biomass
Biosensors
Carbon dioxide
Chromatography
Oxygen
pH
Pressure
Temperature
Viscosity

OUTLINE

INTRODUCTION

For most modern industrial production methods, process monitoring plays a key role; bioprocess systems are no exception to this rule. Effective methods of monitoring are required in order to develop, optimize, and maintain biological reactors at maximum efficiently. An optimized process should lead to streamlined performance, reductions in running and material costs, and improvements in quality control.

A number of bioprocess monitoring instruments have been regular features of the industry for many years (Fig. 1), in particular, sensors capable of measuring physical parameters. The challenge is to develop cost-effective devices capable of measuring a much wider range of parameters, providing process operators with a deeper insight into the process. Eventually effective monitoring will lead to better control systems, providing cost and quality benefits.

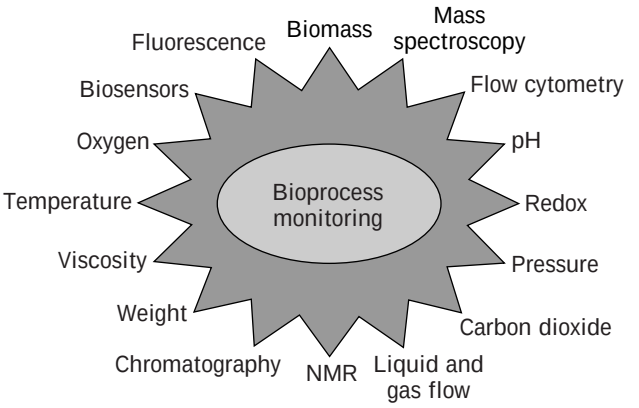


Figure 1. Major areas of bioprocess monitoring.

OPERATING FEATURES

For any measuring technique, there are a number of criteria that the device must satisfy if it is to accepted by commercial bioprocess operators (Fig. 2). The following list of desirable characteristic is not exhaustive, but does cover the main areas of concern.

Reliability

Reliability is a key issue for bioprocess monitoring equipment. Confidence in the ability of an instrument to maintain its credibility in terms of performance is a fundamental parameter. This encompasses a number of features relating to the operation of the instrument, for example, ease of use, maintenance, repair, and replacement. A successful monitoring instrument will have a low failure rate; again, this is related to reliability.

Accuracy

Accuracy can be described as the relationship between the measured value (by the instrument) and the actual value of a bioprocess variable. An accurate instrument will achieve a low percentage error between these two values. Generally this can be stated as

$$\% \text{ error} = (\text{measured value} - \text{true value})/\text{true value} \times 100\%$$

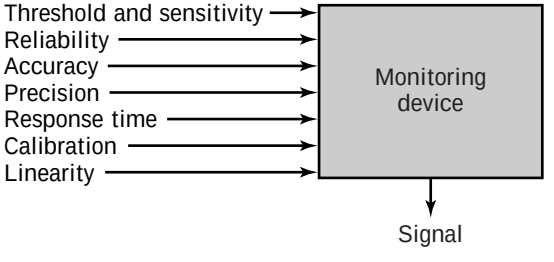


Figure 2. Main criteria for a bioprocess monitoring instrument.

Precision

Precision is a measure of instrument reproducibility, that is, the ability to obtain the same value with repeated measurements of a process variable (at a constant level). It follows that a precise instrument may not, necessarily, be accurate. Therefore it is important to distinguish between these two parameters.

Response time

The measurement of any process variable will entail a time delay between change in the parameter and display of the measured value. This response time should not be detrimental to the progress of the bioprocess, particularly if the measurement is linked to a control action. Response times will be affected not only by the type of instrument but also by the method of measurement. Off-line sampling and measuring can involve a number of time-consuming steps. In contrast, in situ devices can provide a real-time measurement.

Calibration

In order to maintain the accuracy of a sensor it is (generally) desirable to carry out a calibration step. This is carried out using set standards of comparison. The effects of calibration can impinge on the process, in other words, it may be difficult to carry out calibration of an in situ sensor, whereas an on-line device could easily incorporate a calibration step(s) during routine running.

Linearity

Under ideal conditions, the output signal from a sensor would be directly proportional to the analyte concentration. However, this is not always the case, and other models have to be used to reach a true value. Despite this drawback, linearity is not an overriding prerequisite for a successful monitoring instrument. Developments in modern software can be adapted to compensate for nonlinearity.

Threshold and Sensitivity

Sensitivity is the magnitude of the output signal per unit change in the target analyte concentration. The lowest level of detection is related to the sensitivity and the signal-to-noise ratio. A number of factors influence the sensitivity of a monitoring device, including sensor design, the operating environment, periods of maintenance, and interfering noise levels. The sensitivity will influence the dynamic range above which the device becomes saturated; the highest level of detection will be the threshold limit for the device.

METHODS OF MONITORING

In addition to the variety of monitoring devices available, a number of methods (for carrying out the measurement exist (Fig. 3). Sampling and sample handling is a vital issue, affecting both the accuracy and frequency of measurements (1). Broadly, the main approaches to sampling can be described in one of several ways.

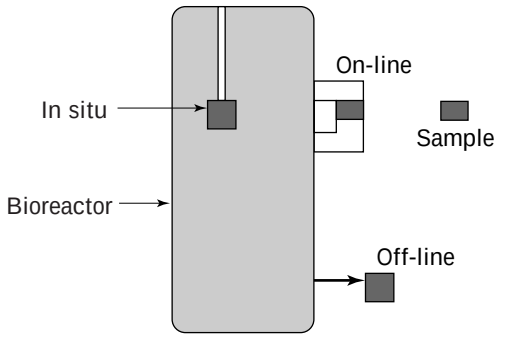


Figure 3. Methods of monitoring.

Off-Line or At-Line Monitoring

Off-line monitoring involves taking a sample from the bioreactor and carrying out the measurement at a different site, usually under laboratory conditions. Off-line sampling can be detrimental in terms of cost, efficiency, and threat to asepsis. Manually removing and measuring a sample requires technical labor. The received signal will not be "real time" because of the delays. In order to obtain the sample, the sterility of the bioreactor must be maintained. Hence, provision must be made to achieve this. In some cases, decentralized equipment has become available that allows measurements to be made simply close to the production process (at-line), reducing some of the delays associated with laboratory analysis.

In Situ or In-Line Monitoring

In situ sensors are placed directly in the bioprocess vessel. The use of in situ sensors has long been established in the bioprocess industry (2), where "dip-in" devices are used to monitor a number of parameters such as pH and dissolved gas concentrations. A number of advantages are gained by operating sensors in this fashion, including real-time monitoring (sampling-related time delays are eliminated) with the sensors operating continuously (any rapid change in the analyte concentration can be readily observed), and labor requirements and problems of contamination are both significantly reduced. However, in situ sensors can have a number of drawbacks, for example, the sensor system must be amenable to sterilization; if the lifetime is short, replacement may be difficult (during a process run); the surface of the sensor could become fouled by components of the growth medium, affecting the signal output.

On-Line Monitoring

Methods of on-line monitoring involve the automatic removal and measurement of sample or sample stream from the bioreactor. One example of this has been the development (since the early 1980s) of flow injection analysis (FIA). This is a liquid handling technique that has proved flexible in adapting to most chemical and biochemical reaction procedures (3), representing an effective compro-

mise between the desirability of in situ monitoring and the technical ease of off-line measurements. The main advantages of on-line methods include the following: sensor sterilization can be readily accomplished, sample pretreatment (e.g., gassing, dilution, removal of interferents), and sensor calibration can be built into the system. The main disadvantages are a need for an effective and reliable sampling system and the fact that the signal is discontinuous; frequency measuring rate is determined by the limitation of the overall FIA arrangement. All of these methods have their advantages and disadvantages. The choice of approach adopted depends on a number of factors, not least of which will be the availability of both sensor and system.

MEASURING DEVICES

pH Measurement

Monitoring the pH of a number of industrial biological reactions is an extremely effective way of determining their course, providing the operator with valuable information. Measuring pH, along with temperature, is one of the most common practices in bioprocess monitoring (4). Correcting action can be taken, either manually or automatically, to prevent unwanted increases or decreases in pH.

The most common form of pH sensor used for fermentation monitoring is based on the electrode design introduced by Ingold in 1947. Under operating conditions, the probe must always be overpressurized (with respect to the fermentation broth) to keep a positive flow of electrolyte out of the porous plug. A combined glass and reference electrode that is capable of withstanding repeated sterilization at temperatures of 121 °C and pressures of 138 kN/m^2 is required. In this case, the term *combined* refers to the single unit containing both electrodes. Figure 4 illustrates the main features of the pH electrode.

Typically, the reference electrodes used are silver/silver chloride (Ag/AgCl) or calomel (Hg/HgCl) and maintained in a concentrated solution of potassium chloride in a buffer solution within the body of the sensor. Under these conditions an equilibrium is established with the electrolyte:

$$Hg/HgCl_2/KCL \rightarrow 2Hg + 2Cl^- \rightarrow Hg_2Cl_2 + 2e^- \quad (1)$$

This leads to a situation whereby the reference electrode potential is dependent on the activity of the chloride ions.

Generally the sensor format is that of a single unit glass-reference design. It has been known since the early years of this century that certain forms of glass are electrically conducting. Furthermore, the electrical potential at the glass-liquid interface changes with activity of hydrogen ions. This potential is directly related to the activity of the H^+ via the Nernst equation,

$$E = E^0 + (RT/\eta F) \ln aH^+ \quad (2)$$

where E is the measured potential, R the gas constant, T the absolute temperature, F the Faraday constant, and ln aH^+ the activity. With this type of pH electrode, the potential determining process is determined by interactions between H^+ activity (in the sample solution) and sianol groups on the glass.

Autoclavable pH probes are available (for sterilization purposes), but this can lead to a reduction in useful probe life. Generally the sensor's response does degenerate over time with repeated use. Interactions with the medium constituents can lead to problems. However, proprietary cleaning agents are available to regenerate the probe. A recent advance in the field of pH monitoring has been the introduction of sensors using ion-selective field effect transistors (ISFETs). These are solid-state devices based on the use of semiconductor structures (5,6). The pH is determined by measuring the change in the voltage required to maintain a zero current between the gate and drain electrode as the transmembrane potential alters. Several commercial companies have marketed ISFET-based pH meters with potential applications in a number of areas. These sensors are robust, virtually shatterproof, and are easily cleaned. In addition, they require no filling solution and are easy to operate and maintain.

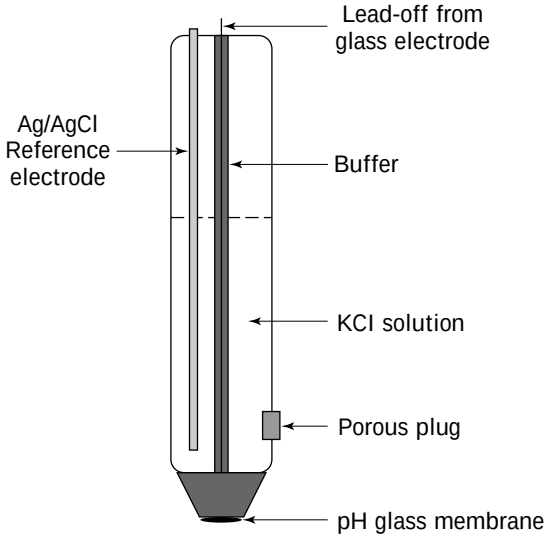


Figure 4. pH electrode.

Temperature

Maintaining optimum conditions for any bioprocess will invariably involve monitoring and controlling the temperature of the broth. Generally, bioprocesses are monitored over a temperature range of between 0 °C and 100 °C (excluding the sterilization cycle). Furthermore, this may require a control regime operating within a narrow range of temperatures. A number of devices are available for obtaining an accurate measurement of temperature conditions during a bioprocess operation, based on a range of techniques.

Thermistors. Thermistors are semiconductor devices composed of agglomerates of polycrystalline metallic oxides (e.g., magnesium oxide, cobalt oxide, ferric oxide). The operating characteristic of these devices relies on a sensitive change in resistance, resulting from temperature fluc-

tuations (7). This change is a function of the absolute temperature. The output response from thermistors is of a nonlinear temperature versus resistance curve, with the resistance decreasing as the temperature increases. Generally the response can be described by the following equation:

$$R(T) = R_0 \exp[B(1/T - 1/T_0)] \quad (3)$$

Thermistors tend to be cheap and stable, with rapid response times.

Resistance Thermometers. Resistance thermometers are based on the changes in the electrical resistance of metallic conductors, with changing temperature (8). This relationship can be described by the following equation:

$$R = R_0(1 + \alpha T) \quad (4)$$

R is resistance. T is the temperature, and α the temperature coefficient of linear expansion.

Platinum is the preferred metal for these devices, but other metals can be used (e.g., nickel, tungsten, or copper). The advantage of platinum is that it can be obtained in a highly pure form, hence the electrical properties of the metal can be well characterized. A platinum wire of 100-Ω resistance at 0 °C is typically used. Generally, calibration of the sensor is carried out at this temperature. Readings can be obtained using a Wheatstone bridge circuit. Overall, these sensors tend to be precise, accurate ($\pm 0.25\%$), reliable, and sensitive. Platinum sensors are stable in both sterilization and fermentation conditions.

Thermocouples. Thermocouples are based on the Seebeck effect, first postulated in 1821. It was noted that a circuit consisting of two dissimilar metals joined together at junctions maintained at different temperatures will generate an electrical current. The electromotive force generated will be determined by the choice of metals and the respective temperatures. For practical applications, one of the junctions (termed the *cold junction*) is maintained at a constant temperature and used as a reference point. The second junction is in contact with the media to be monitored; fluctuations in the temperature can be studied by monitoring the current response (7).

The response (E mV versus T[K]) approximates to a parabolic curve. Generally it is preferable to measure over the linear range of this response. One advantage of thermocouples is that they do not require an external electrical supply. However, they are susceptible to cold junction problems, with the need to maintain a constant temperature.

Mercury-in-Glass Thermometers. Mercury-in-glass thermometers can be used directly in small batch fermenters. However, the fragility and potential toxicity of these thermometers restricts their use. In order to overcome the problem of using mercury, alcohol can be substituted. Generally these sensors are not very accurate (especially over a wide range) and provide only an indication of the temperature. However, these thermometers have the advantage of being both portable and providing a direct reading.

Bimetallic Thermometers. Bimetallic thermometers usually consist of a bimetallic helical coil surrounded by a protective tube or wall. The coil winds or unwinds with changes in temperature and causes movement of a fixed pointer (8). By affixing a pen to the fixed point, a temperature profile of the bioprocess can be obtained. Despite being more robust than the mercury thermometer, they are less accurate (the range determined by the metals used) and, again, are limited to giving only an indication of temperature.

Pressure

Many bioprocesses operate under conditions of overpressure. Monitoring the magnitude of this pressure is an important factor, both in terms of safety and optimization of the process. Industrial and laboratory fermenters are designed to operate up to a safe working pressure. Increasing the applied pressure above the upper limit can be dangerous. Furthermore, maintaining a positive reactor head pressure can prevent contamination of the bioreactor. In order to facilitate effective sterilization, fermenters need an accurate pressure monitor. Pressure will also affect the solubility of gases (such as O_2 and CO_2). Several different approaches have been used in the development of pressure gauges.

Bourdon Tube Pressure Gauge. The Bourdon tube pressure gauge is composed of an arched or spiral-shaped tube. Part of the tube is of an elliptical shape, made of either metal or quartz. Under pressure, the elliptical section tends to become circular. The interior of the tube is influenced by the process pressure, whereas the exterior is under atmospheric pressure. The overall effect of the change in cross-sectional area is to straighten out the curved shape of the tube. One end of the tube is fixed to an inlet from the reactor. The other end is sealed and allowed to move; this is connected (via a geared section and pinion) to an indicator needle (9). Pressure can be read directly from the gauge. These devices tend to be robust and economical.

Strain Gauges. Strain gauges are based on the phenomenon that when a wire is placed under strain, the electrical resistance changes (7). This is due to both the alteration in geometry and resistivity changes. The electrical output can be related to changes in pressure. Typically the sensors are constructed using either a metallic or semiconductor element, usually in a grid structure composed of a number of conductor threads. Generally the change in resistance to pressure can be related by the following equation:

$$\Delta R/R = \Delta L/L + \Delta\rho/\rho - \Delta S/S \quad (5)$$

R is resistance, L is length of thread (a multiple of the number of threads), ρ is the resistivity of the material, and S is the cross-sectional area of the material.

Piezoelectric Manometers. Certain materials, such as quartz crystal, posses an asymmetrical electrical charge distribution. When an alternating electric field is applied

across a piezoelectric material, small mechanical deformations (to the crystal) are generated (10). Alternatively, if the crystal is subjected to pressure on one face, an electrical charge will be generated on the opposite face. This effect (the piezoelectric effect) forms the basis of operation for piezoelectric manometers. By attaching electrodes to one face of the crystal, changes in pressure can be relayed as an electrical output.

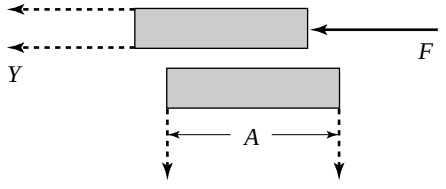
Diaphragm-Type and Pressure Bellows Sensors. Diaphragm-type sensors incorporate a membrane that deforms as a result of pressure from either a liquid or gas, forming a pressure difference across the membrane. This movement can be monitored and a corresponding readout obtained, indicating the pressure of the matter under investigation. Generally, the membranes used can be made from materials such as stainless steel, palladium, or silicon. When in operation, these sensors can suffer from the deposition of material on the surface, leading to erroneous measurements. In order to prevent this, periodic checks should be made.

Pressure bellow sensors operate, as the name implies, by monitoring the movement of a bellows when subject to a pressure. Changes in the bellows extension length can be related to pressure. The bellows can be linked directly to a variable transformer that is capable of recording the physical movement and displaying a readout.

Viscosity

Information on the rheology, or viscosity, can help in ensuring the efficiency of a biological process (11). An example of this would be the effectiveness of oxygen transfer throughout a medium and determining the degree of branching of filamentous microorganisms. In addition, viscosity can effect pumping, mass transfer, and mixing. Viscosity is the apparent shear resistance between adjacent layers of liquids or gases moving at different speeds. Typically, in fluids this is the result of molecular cohesion; rising temperatures lead to a decrease in viscosity. In contrast, viscosity increases for gases under conditions of rising temperature. This results from an increased measured molecular activity.

Newtonian Fluids. Consider two parallel surfaces with area A and distance Y apart. If a shearing force F is applied to one side, it will move with a velocity V:



It is found that for ideal or Newtonian fluids the shear stress $F/A(\tau) = dV/dY$ multiplied by a constant (μ). This constant is called the coefficient of absolute (or dynamic) viscosity. The term dV/dY is called the *shear rate* (γ). Plotting τ against γ gives the flow curve characteristics of a particular fluid. Newtonian fluids display a straight-line response, passing through the intersection of the axis. The slope of this gradient is equal to μ. Typical units of viscosity are Newtons per second per square meter ($NS\ m^{-2}$) and centipoise. Examples of Newtonian fluids include all gases, liquids with low molecular weight (notable exceptions are the colloidal suspensions and polymeric solutions with long molecules, e.g., specially developed mineral oils), water, and some oils.

Non-Newtonian Fluids. Non-Newtonian fluids are classed according to their flow curves, which may be of various shapes according to the nature of the fluid (12). For each fluid the ratio between shear stress and shear rate is not constant.

Bingham Plastic. For these fluids a certain minimum shear (τ^0) is required before viscous flow takes place, after which the relationship is linear. Examples include some fats and plastic melts, some greases, and detergent slurries. Perhaps the best everyday example is toothpaste, a certain amount of pressure must be applied to the tube before the paste begins to flow.

Pseudoplastic Fluids. Apparent viscosity decreases with shear rate, tending to a constant value at high shear rates. This behavior is typical of suspensions of long-shaped particles of high-polymer solutions such as cellulose derivatives. The process has been explained as the result of progressive alignment of particulars or molecules as shear rate increases. Examples include starch suspensions, biological fluids, and rubber solutions.

Dilatant Fluids. In contrast to the pseudoplastic fluids, dilatant fluids display an increasing apparent viscosity with increased shear rate. Examples include sugar solutions, starch, and wet cement aggregates.

Casson Plastic. These fluids display a decreasing apparent viscosity with increasing shear rate after a yield stress (τ^0) value has been exceeded. Examples include blood, printing ink, and some fruit juices.

Generally, bioprocess media display non-Newtonian behavior, that is, a nonlinear response to shear rate. These fluids do not have a constant viscosity, it varies according to the shear rate. In practice, because of the problems of maintaining sterility, most viscosity measurements are made off-line.

There are a number of commercially available instruments (viscometers), applicable to the measurement of bioprocess rheology properties (Fig. 5). Generally, these instruments can be classified into three different types:

1. *Cone and plate viscometer*. This instrument consists of a flat plate in contact with an inverted cone, the fluid to be measured fills the gap between the cone and the plate. In operation, the cone is rotated in the fluid, and the angular velocity and torque are measured. These values can be related to the shear stress and rate.
2. *Coaxial cylinder rotary viscometer*. This instrument consists of two cylinders, one located inside the other. The fluid to be measured fills the gap between the two cylinders. Again, the angular velocity and torque are measured when one of the cylinders is held stationary and the other rotated.

Figure 5. The CVO rheometer (Bohlin Instruments).

3. *Impeller viscometer*. This device is similar to the coaxial cylinder viscometer. However, the fluid measurements are made using a rotating impeller inside the stationary cylinder.

Oxygen

Generally, oxygen measurement falls into two main categories: dissolved oxygen and exit gas analysis. For most aerobic fermentations, an accurate determination of the dissolved oxygen concentration is of particular importance, for example, in maintaining the concentration above a specified minimal level. In addition, oxygen transport processes can be determined only by using an accurate method of measurement.

The exit gas analysis for oxygen, usually in conjunction with CO_2 measurement, can be used to determine the metabolic state of aerobic microorganisms and their oxygen uptake.

Dissolved Oxygen Probes. A wide range of dissolved oxygen probes based on the use of electrochemical determination are available (13). Typically these fall into two main types (14): galvanic (potentiometric) or polarographic (amperometric or Clark). In essence these electrodes measure the activity (or partial pressure) of oxygen present, rather than the concentration. The range covered by these devices is approximately 0.001 to 1000 ppm (15). Both the galvanic and polarographic electrodes use an oxygen-permeable membrane (e.g., Teflon, polyethylene, and polypropylene) to separate the electrode components from the fermentation media. Similarly, the measurement for both types is based on the electrochemical reduction of oxygen at a cathode.

Polarographic Electrodes. The principal components of the polarographic oxygen sensor (16) include a precious metal cathode (e.g., gold or platinum), a silver/silver chloride (Ag/AgCl) anode, and an electrolyte solution (e.g., sodium chloride). During operation the following reactions take place at the respective electrodes:

$$\begin{aligned} \text{Cathode:}\quad & O_2 + 2H_2O + 2e^- \rightarrow H_2O_2 \\ & + 2OH^- \rightarrow H_2O_2 + 2e^- \rightarrow 2OH^- \\ \text{Anode:}\quad & Ag + Cl \rightarrow AgCl + 2OH^- \end{aligned} \tag{6}$$

Typically, polarographic (or amperometric) sensors operate at a fixed voltage, 0.6 to 0.8V, between the two electrodes. This voltage range represents the diffusion-controlled plateau of oxygen reduction. In effect, the current recorded can be directly related to the oxygen concentration.

Galvanic Electrodes. In contrast to the polargraphic oxygen electrode, the galvanic sensor (17) employs a non–noble metal (zinc, lead, or cadmium) anode with a noble metal cathode (silver or gold). The reactions occurring at the electrodes are

$$\begin{aligned} \text{Cathode:}\quad & O_2 + 2H_2O + 4e \rightarrow 4OH^- \\ \text{Anode:}\quad & 2Pb(\text{e.g.}) \rightarrow 2Pb^{2+} + 2e^- \end{aligned} \tag{7}$$

In contrast to the polargraphic sensors, galvanic electrodes require no external voltage source. The combination of a sacrificial anode and noble metal cathode is sufficient to provide an adequate electromotive force, resulting in a spontaneous reduction of oxygen at the electrode.

Despite their differences in operation, both types of electrodes are subject to similar constraints. During steady state, oxygen flux to the cathode is dependent on a number of transport-related steps including movement of oxygen through the bulk of the fermentation media and diffusion across the membrane and through the supporting electrolyte. In relation to the fermentation media, the viscosity and local hydrodynamic characteristics (in the vicinity of the electrode face) will play an important role. An example of this would be the incorrect placing of the electrode in a highly viscous media displaying non-Newtonian characteristics. If the electrode is situated in an area of quiescent fluid, the received signal may not be an accurate representation of oxygen partial pressure throughout the bioreactor.

The design of the sensor will determine its operating characteristics. These include the membrane thickness, the membrane diffusion coefficient, and the surface area of the cathode. Temperature can also have a significant influence on the sensor response. Generally a change in temperature will affect both the solubility and diffusion of oxygen through the membrane. In order to compensate for

these effects a built-in thermistor is generally included in commercial devices.

There are several problems encountered by both types of sensor during bioprocess operations. Fouling of the sensor membrane surface by components of the bioprocess media can lead to erroneous signals. With continued use, this could include the growth of microorganisms on the membrane surface. The effect of a buildup on the membrane would be to change the diffusion characteristics of the sensor (see equation 8). Dissolved oxygen electrodes are also susceptible to signal drift during operation. In addition, gas bubbles and restricted flow toward the sensor can affect the received signal. Other important considerations are the effect of sterilization on the sensor, and calibration. Typically, calibration is carried out after sterilization, using medium gassing to obtain the full range of sensor values. In order to avoid sterilization problems, the sensor can be withdrawn into an isolating lock where it can be sterilized using milder conditions. The actual measurements can be expressed in one of several ways. One approach is to use percent of saturation on a scale of 0 to 100% oxygen saturation. Alternatively partial pressure can be used,

$$pO_2 = (p_{atm} + \Delta p - pH_2O)0.2095 \quad (8)$$

where $(p_{atm} + \Delta p)$ is the absolute pressure and pH_2O is the partial water pressure. Finally a direct concentration can be obtained using Henry's law coefficient:

$$c = pO_2/H \quad (9)$$

The coefficient value is subject to alteration through interactions with media components such as sugars, products, salts, and so forth (18). Hence, these factors must be taken into consideration when determining oxygen concentration.

O_2 Optodes. A recent development for measuring oxygen, and indeed a range of other analytes, has been the introduction of sensors based on the use of photometric transducers: optodes (19). The driving force for the development of these devices was linked to the evolution of immobilization technologies and fiber optics. The optical fibers (used to transmit information over long distances: 2–100 m) can be sited in harsh environments (e.g., high temperature and extremes of pH), allowing in situ real-time monitoring. A fiber is constructed using an optically transparent material called the *core* (this is the guide path for the light) and a cladding, surrounding the core. Light is totally reflected at the cladding-core interface, allowing it to be transported over long distances without significant loss. The light interacts with a sensitive element or layer, generating a signal.

Several spectroscopic techniques have been used in the construction of optodes; of these, fluorescence has been the most widely used in the construction of O_2 sensors. Fluorescence is the emission of light from an absorbing species that has been excited by light of a different wavelength. Optical sensors for oxygen are constructed using an immobilized flurophore that undergoes dynamic quenching. Increasing concentrations of the quencher lead to a decrease in the fluorescence. The relationship between fluorescence intensity and extent of quenching can be summarized as

$$I_0/I = 1 + K_{sv} \times [Q] \quad (10)$$

where I_O and I represent the measured fluorescence intensity in the presence and absence of the quenching agent, K_{sv} the Stern-Volmer proportionality constant, and Q the concentration of the quenching agent (20).

For a number of applications, dyes, whose fluorescent properties depend on the presence of a target analyte (e.g., oxygen), can be used. The dye can be attached directly to the surface of the fiber or onto a membrane. With the most common configuration, where sensing takes place at the end of the fiber light is returned after acting with the sensitive layer. This can occur through the same fiber or through a parallel fiber in a bifurcated device. The magnitude of the interaction is determined at the other end of the fiber using a detector system (21).

These sensors have a number of advantages compared with electrochemically based systems including that they do not require a reference signal, they can be miniaturized, multisignaling from a range of analytes measured at the same time can be achieved, and the fibers can be steam sterilized. Despite these advantages there are a number of drawbacks to these devices, including interferences from ambient light, photobleaching and washout of the sensitive indicator, and the limited range of some devices. A number of fiber-optic-based O_2 optodes have been described, using a range of indicator molecules (e.g., 22,23,24). Their use in measuring oxygen concentrations during bioprocess applications will indubitably become more widespread as the technology develops.

Tubing Methods. An alternative approach to measuring dissolved oxygen content is to use the tubing method described by Philips and Johnson (25). This method is based on the use of a polymeric tubing (e.g., permeable Teflon) immersed in the fermenter media. A stream of helium or nitrogen gas is passed through the coil. Oxygen in the fermenter medium diffuses through the tubing into the carrier gas stream. The oxygen concentration in the inert gas is then determined using a suitable detector, for example, a paramagnetic gas analyser. Despite the longer response times for this method (2 to 10 minutes for a reading) compared with the electrode approach, it is a reliable technique. In addition the tubing is robust and can withstand repeated sterilization cycles.

Exit Gas Analysis. There a number of methods available for determining the oxygen concentration of exit gas from a bioprocess reactor. Several of these are based on exploiting the strong affinity shown by oxygen to a magnetic field. The paramagnetic gases, such as oxygen, display a positive magnetic susceptibility. There are two types of detectors based on this phenomenon: the magnetopneumatic and thermomagnetic analysers.

Magnetodynamical Analyzer. The magnetodynamical analyzer is based on a dumbbell-shaped di-magnet held in

a nonuniform magnetic field (15). The dumbbell structure is free to rotate about an axis. In the presence of oxygen, the dumbbell is rotated (under the action of the increased magnetic force) in proportion to the oxygen partial pressure. By applying a known corrective electrostatic force to return the dumbbell to its original position, the current required can be directly related to the partial pressure.

Thermomagnetic Analyzers. Thermomagnetic analysers (26) are based on the observation that changes in temperature affect magnetic susceptibility (χ). Generally this is described by the Curie law,

$$\chi T = \text{constant} \tag{11}$$

whereby χ is inversely proportional to the absolute temperature. The instrument consists of a flow-through ring element divided up into two measuring chambers. The detection site is composed of a thin-walled glass tube containing resistors that heat the incoming gases. These resistors form part of a Wheatstone bridge circuit and are used to detect changes in resistance resulting from flow-rate variations. When the gas is heated, oxygen loses a high proportion of its magnetic susceptibility. This gas, in turn, is displaced by incoming cool oxygen, generating a convection current or "magnetic wind." The flow of current (or heat transport) causes a change in resistance, imbalancing the Wheatstone bridge. In effect the convection current can be related to the oxygen concentration.

Carbon Dioxide

For many bioprocesses, the measurement of CO_2 is an important feature. Increased levels of dissolved carbon dioxide can inhibit growth and reduce the production of secondary metabolites.

Electrochemical Probe. Steam-sterilizable electrochemical probes for determining dissolved CO_2 partial pressure are commercially available. These function in a similar way to oxygen probes and are based on a pH sensor immersed in a saturated bicarbonate buffer, separated from the bioreactor fluid by a hydrophobic membrane. Dissolved CO_2 gas molecules diffuse from the bioreactor media, through the hydrophobic membrane, into the carbonate buffer. The bicarbonate buffer equilibrates with the dissolved CO_2 gas molecules:

$$\begin{aligned} &CO_2 + H_2O \rightarrow HCO_3^- + H^+ \\ &aH^+ = K(aCO_2)/(aHCO_3) \end{aligned} \tag{12}$$

Here, a is the activity. Hence, increases or decreases in pH can be directly related to the concentration of dissolved CO_2. This logarithmic relationship requires a two-point calibration, which is usually carried out by measuring the pH response, following substitution of the carbonate buffer with a reference solution.

CO_2 Optrodes. Optical sensors for the detection of CO_2 have also been described. Generally, these sensors are analogous to CO_2 electrochemical systems, whereby changes in pH resulting from the formation of bicarbonate ions are detected. One example (27) described the use of a fluorescein dye as a pH indicator. All of the advantages and disadvantages already outlined for the O_2 sensor can be equally applied to these devices. Future developments in this field should lead to a more widespread use of these sensors.

Exit Gas Analysis. With a well-mixed aerobic fermentation, the concentration of dissolved CO_2 can be estimated from analysis of the exit gas. One particularly important approach used for determining the exit gas CO_2 concentration is infrared (IR) analysis. Particular molecules absorb IR radiation at characteristic wavelengths. This absorption is related to the chemical structure of the molecule, that is, the type and number of function groups present. When these molecules absorb IR radiation, the vibrational and rotational states are altered. Observing this absorbance (in either the gas, solid, or liquid phase, or combinations of these), the concentration of the target analyte can be determined (28). In the gas phase, CO_2 absorbs strongly at wavelengths <15 μm. The main components of an IR analyzer (29) consist of a suitable radiation source, a method of selecting the range of wavelengths required, a sample and reference cell, a detector, and a means of displaying the information in terms of CO_2 concentration. Generally, CO_2 is monitored using a positive filtering method. With this method, the optical system detector is filled with CO_2. The system detects the reduced radiation energy of the measuring beam reaching the detector. Comparing the reference and sample responses provides a value for the CO_2 concentration present. One of the main problems encountered with this approach is conflicting absorbance patterns from "interfering" molecules (i.e., those absorbing over the same wavelength range).

Redox Potential

Monitoring the redox potential of a bioprocess medium can provide information about the equilibrium between oxidizing and reducing species (electron acceptors and donors, respectively) present. Measurement of redox potential is achieved using a combined metal-reference electrode system. Typically the metal electrode can be gold, iridium, or platinum; with platinum being the usual choice. The reference electrode is either Ag/AgCl or calomel. The redox sensor is linked to a pH meter that is fitted to provide a readout in millivolts. During operation, the redox potential (measured between the metal and reference electrodes) varies as the logarithm of the ratio of oxidizable and reducible components in the media. The redox value varies linearly with the pH of the media and with the logarithm of dissolved oxygen tension. This can lead to complications when trying to interpret the readout from the sensor. A full description of the use and problems associated with the redox sensor was outlined by Kjaergard (28).

Chromatography

Chromatography has proved to be a valuable tool for numerous applications in diverse areas of analysis. The principle is based the separation of a mixture of compounds by

means of a suitable matrix. Target analytes are separated by their characteristic interaction with the matrix. Following separation, the compounds can be detected using a range of detectors (7) including flame ionization (used for gas chromatography applications), UV detectors, electrochemical detectors, infrared and refractive index detectors, conductivity detectors, and mass spectrometers. A number of chromatographic methods are available, several of which can be used for bioprocess monitoring. Overall chromatographic analysis is carried out off-line. However, recent developments have shown that on-line methods of analysis can be developed.

Liquid Chromatography. For the bioprocess industry, liquid chromatography is used in sample analysis as a preparative method for the large-scale purification of biomolecules and as recovery technique for high-value products such as therapeutics and pharmaceuticals. One important development in this field has been the introduction of high-pressure liquid chromatography (HPLC). This approach to liquid chromatography involves the use of well-characterized homogeneous columns that contain small (3–5 μm-diameter) particles. In operation, these columns can withstand high pressures (up to 6,000 psi). The overall effect is to achieve a fast, accurate separation of the components in a sample. Liquid chromatography permits the separation of a sample solution into consistent components through the interaction of two phases. One type incorporates a solid stationary phase in conjunction with a fluid mobile phase containing the sample of interest. During their passage over the stationary phase, sample components are separated by their degree of solubility or, alternatively, their affinity for the solid matrix. The result is a separation profile indicative of the component mixture of the sample. There are a numerous examples of HPLC systems being used on-line to monitor bioprocess production (e.g., 30,31,32). Typically, on-line monitoring is carried out using a sampling system based on membrane filtration (33,34).

Liquid–Liquid Chromatography. With the liquid–liquid chromatography method, the stationary phase is coated with a suitable fluid. The sample is carried in the liquid phase under pressure. Partitioning between the two liquids leads to a separation of the sample components. Detection of the separated components can be carried out using detectors such as UV absorption, conductivity, or refractive index. Liquid–liquid chromatography can be further divided into normal-phase type, in which the carrier liquid is less polar than the stationary fluid phase, and reverse-phase type, in which the carrier liquid is more polar. The reverse-phase method is particularly useful for separating organic compounds in an aqueous carrier solvent.

Ion Exchange Chromatography. Ion exchange chromatography is used to separate ionic compounds; typically from an aqueous carrier. This approach is particularly suited to measuring low concentrations of ions. The method is based on the ionic interaction between cations or anions in the sample solution, with the corresponding (oppositely charged) functional group on the column matrix.

Gas Chromatography. Gas chromatography (GC) employs gas as the mobile phase and is principally used for separating and measuring the composition of volatile compounds such as aldehydes and alcohols, for example, in a bioreactor headspace. Again, this is a chromatographic approach that has been widely used for bioprocess monitoring. Connected to a suitable sampling system, GC can be used in an on-line format (35,36). An important development in the field of analysis has been the introduction of Fourier transform infrared spectroscopy (FT-IR) instrumentation (37). Combined with chromatography (e.g., GC), this approach can provide a powerful tool for the accurate, sensitive analysis of gas mixtures. This technique will undoubtedly become more important over time.

Gel Filtration. Gel filtration is a method of separation based on molecular size and is carried out using an appropriate polymeric matrix. By selecting the correct pore-size exclusion limit, molecules in a sample solution can be separated. Small molecules, relative to the gel "sieve," pass rapidly through the chromatography column. However, larger molecules diffuse into the polymer matrix, where their passage is impeded. Hence, elution time from the column is determined by molecular size.

Mass Spectrometry

The use of on-line mass spectroscopy (MS) has developed since the first report of Reus et al. (38). A wide range of gasses, both free and dissolved (e.g., CO_2, O_2, CH_4), can be measured. In addition, volatile organic compounds such as methanol, ethanol, acetone, and simple organic acids can be monitored. The technique is based on the rupture of molecules by a high-energy source into corresponding ions. Magnetic separation of the charged ions may then be achieved, based on their mass: charge ratio. Each compound produces a characteristic display pattern. Typically an instrument consists of an electron bombardment source, the magnetic ion separation system, and the detector.

There are several types of mass spectrometer available. The quadrapole MS is based on the use of a radio frequency system to separate the ions. A combination of applied voltage and frequency range produces a change in the trajectory of the ions. The term *quadrapole* comes from the four rods used to produce the radio frequencies. There is an alternative MS instrument that is potentially more precise and stable. This instrument is based on the use of a magnetic field to separate ions. Two types of magnetic instruments are available. One system uses a fixed magnetic field and several detectors. The other uses a variable controlled magnetic field with one detector system.

Mass spectrometers offer a number of advantages for process monitoring including speed, specificity, precision, and stability. The major drawback to the use of mass spectrometers is cost; in addition some peaks can be difficult to resolve using the less elaborate instruments. Despite these drawbacks, MS monitoring of volatile compounds has been widely reported (39,40,41).

Weight

Monitoring fluctuations in the weight of a vessel used for bioprocesses is an effective way of measuring the contents and flow rate of additions. There are a number of load cells available for this task. The principle of a load cell (8) is based on measuring the compressive strain that is placed on the device (e.g., a solid or tubular steel cylinder) when under an axial load. Electrical resistance strain gauges can be included in the device structure. A proportional electrical resistance to the applied load can then be measured; this varies in relation to changes in load. Obviously, a temperature compensator must be included in the instrument to account for heat effects on the resistors. These devices need to be rugged and have long-term stability. For reasons of safety, it is also desirable for the device to be accurate. A load cell designed to detect tensile forces can be used for weight measurements of suspended vessels.

Liquid Level Measurement

Liquid levels in a vessel can be determined using a number of techniques (42). Determining the level can assist in formulating mass balances and measuring nutrient additions. Hydrostatic pressure can be measured using a single sensor positioned at the bottom of the vessel. Alternatively, two sensors spaced at the top and bottom of the vessel (differential design) can be employed. These sensors are electromechanical devices based on the deformation of a spring; the resulting signal is displayed as an electrical signal (capacitance or reactance).

Conductivity Sensors. Conductivity sensors are particularly suited to monitoring levels of foam (in a number of fermentations), where in excess it can cause problems. The sensor consists of a stainless steel probe, insulated except for the tip. The device can be used with noncorrosive conducting fluids. If the liquid or foam level rises to contact with the tip, an electric current is passed through the sensor; the foam acts as an electrolyte and the vessel as a ground. The sensor can be coupled to a control device that dispenses antifoam to counter the increased level.

Capacitance Sensors. Capacitance sensors operate by detecting changes in the relative dielectric constant of the media, compared with air. Measurements are made by detecting variations in electrical capacity brought about by changes in the liquid level.

Acoustic Sensors. Acoustic sensors operate via transducers that generate and detect ultrasonic waves. Two formats are available: a single device that both transmits and receives the signal, and two separate transducers with one function. By monitoring the time it takes for the sound wave to travel, the liquid level can be determined. With the single device the sound wave is directed onto the liquid surface. When the level rises, the time delay is shortened.

Temperature Probes. Liquid levels can be measured using a series of thermistors sited vertically through the vessel. During operation, an electrical current is applied to the sensors, raising their temperature above the ambient. When the liquid (or foam) level reaches a sensor positioned lower down the vessel, the sensor cools. This causes a temperature difference and, hence, a change in electrical resistance. The resistance can be displayed as an output signal indicating the liquid level.

Flow Measurement and Control

Flow measurement is an important feature for many bioprocesses (7), for both gases and liquids. These measurements are carried out on both the influent and effluent streams.

Gas Flow Measurement. The most common form of gas measurement is carried out using a variable area flow meter, or rotameter (7). This device consists of vertically mounted (usually conical) tube enclosing a free-floating body that is able to move up (floating in the gas flow of interest) and down a tapered bore running through the tube body. Typically, the tube is constructed of glass or metal and the float is a ball or hollow thimble shape. Because the position adopted by the flow rate is dependant on both the gas flow and the viscosity of the medium, the instrument requires careful calibration.

Another important instrument is the thermal mass flow meter. The measuring principle is based on the use of a heating device situated in the gas flow, which is used to heat the gas. The mass flow rate can derived from the heat balance,

$$Q = H/C_p(T_2 - T_1) \tag{13}$$

where Q is the mass flow of the gas, H is the enthalpy, C_p is the specific heat of the gas, T_1 is the temperature of the gas before heat is transferred to it, and T_2 is the gas temperature after the transfer of heat. The mass flow detected using this approach is independent of pressure and temperature.

Liquids. A number of techniques are available to measure liquid flow rates (Pons 1991). The choice of method will be determined by the nature of the liquid (e.g., viscosity, conductivity, etc.). For nonsterile liquids, instruments such as the turbine flow meter can be used. Fluid flow over a turbine blade array is proportional to the rotation speed of the turbine, hence a direct correlation can be made.

A common approach to measuring fluid flow in ancillary pipework is to use a differential pressure device, which utilizes the pressure drop caused by a constriction in the pipeline. This approach is based on the general Bernoulli equation for fluid flow. The constriction causes an increase in the velocity of the fluid, which in turn results in a corresponding pressure drop across the constriction. The differential pressure is a function of the flow velocity and density of the fluid. Measurement of the pressure difference can be used to determine fluid velocity.

Measurement of fluid flow under sterile conditions can be carried out using a magnetic-inductive flow meter. This device can operate with a fluid that includes particulate matter in suspension. The outside of the device comprises two windings that, in operation, are supplied with an al-

ternating current. This induces a magnetic field. A pair of electrodes is placed in conjunction with the device. The average flow velocity is proportional to the electrical potential induced in the field. Potential differences in the field are detected by the electrode arrangement. Hence, accurate flow measurements can be made.

Biomass

The monitoring of biomass concentration can be carried out using a number of techniques (43,44). Conventionally, biomass concentration is measured off-line using labor-intensive, time-consuming methods such as dry weight cell, plate or microscopic cell count, and measuring the optical density of diluted samples (45). However, a number of more rapid methods have been developed. Of particular interest is the development of real-time on-line methods. Generally, methods for determining biomass concentration can be divided into two classifications: direct and indirect. The former is based on determining the physical properties of the cell and its components. In contrast, indirect methods measure factors related to the cell and its activity (e.g., respiration, electrochemical behavior, and nutrient fluctuation).

Bioluminescence and Chemiluminescence. The use of bioluminescent techniques is based on determining levels of adenosine triphosphate (ATP) concentration. Generally, levels of ATP remain constant for living cells, decreasing when the cells die. ATP concentration can, therefore, be related to biomass. A particular enzyme that has proved useful for this method is luciferase, which catalyzes the following reaction:

$$\text{ATP} + \text{luciferin} + \text{O}_2 \xrightarrow{\text{luciferase}} \text{oxyluciferin} + \text{AMP} + \text{light} \quad (14)$$

By detecting the light produced by this reaction, ATP concentrations, and hence, biomass, can be determined. This method can detect cell numbers a low as 10^5 cells ml^{-1} (46); it can be automated, and the assay time is fairly rapid (45). However, the method does suffer from several drawbacks, for example, the extraction of ATP may be incomplete, there may be free ATP present from other sources, and there maybe degradation of ATP by the extraction reagents.

Chemiluminescence is based on the detection of light produced by the protein-catalyzed oxidation of luminol in the presence of hydrogen peroxide.

Acoustic Resonance Densitometry. Acoustic resonance density is based on determining the change in a resonant frequency that results from changes in cell density (47). This is a noninvasive method that does not require contact between the instrument and the culture medium. The method incorporates an oscillatory circuit, amplifier, and test cell. Theoretically, the fluid density of the sample can be calculated from the square of its oscillation. Recent reports have described using this approach to monitor cultures of hybridomas and human lymphoma cells (48). Further advantages for this method include independence from flow rate and viscosity. However, there are disadvantages including the need for a filtration system (to enable the resonance density of the medium to be measured in the absence of cells), with the inherent problems associated with such systems, including poor sensitivity and problems caused by bubbles and particulate matter.

Capacitance, Conductivity, and Electrochemical Methods. The electrical properties of cells have been exploited in the development of a number of techniques. Changes in media capacitance have been related to biomass concentration (e.g., a decreasing capacitance coupled with an increasing biomass concentration) (49). In addition, it has been suggested that cell viability is linked to capacitance measurements (50).

Several in situ commercial instruments are currently available for determining biomass concentration based on capacitance measurements. The Biomass Monitor (Aber Instruments, Aberystwyth, U.K.) measures the radio frequency conductance and capacitance of a cell suspension using a constant voltage, four-terminal, phase-sensitive detector system (Figure 6). Furthermore, the probes can be inserted directly into bioreactors using a standard 25-mm-diameter (Ingold-type) port. The probes are fully sterilizable and can be cleaned in situ during operation.

Fluorescence. Fluorescence is a characteristic possessed by a number of important biological compounds including proteins, enzymes, and coenzymes. Simply defined, it is the absorbance of light energy at a particular wavelength, followed by reemission at a longer wavelength. After the passage of light energy, the compound returns to its ground-state level. In the field of bioprocess monitoring, fluorescence has been used, in particular, to monitor reduced nicotinamide adenine dinucleotide (NADH) and reduced nicotinamide adenine dinucleotide phosphate (NADPH) concentrations. These compounds are irradiated at 340 nm and emit at 450 nm. Both compounds are present in living organisms and are vital components of metabolic processes. By monitoring the intensity of fluo-

Figure 6. The Aber Instrument Viable Cell Monitor S20, connected to an Applikon fermenter.

rescence at the characteristic emission wavelength, it should be possible to calculate the total biomass concentration.

Unfortunately, the use of fluorescence as a marker for cell concentration does suffer from a number of drawbacks including that the fluorescence signal can originate from changes in metabolic state and not from cell concentration levels, sensitivity may be affected by the presence of compounds that can quench both the excitation and emission wavelengths, and other compounds may fluoresce at the same wavelength. Nevertheless, the use of this approach has been reported in a number of papers (51,52,53).

Flow Cytometry. Flow cytometry is a measuring technique based on the irradiation of a sample solution (containing a cell population) with a suitable light source, followed by monitoring of the scattered or absorbed light. In addition, fluorescence can be used as the measuring parameter. This technique can be used to ascertain a number of cellular features, such as the accumulation of cellular components (e.g., DNA, RNA, and proteins), and cell dynamics (e.g., cell size distribution). Furthermore, flow cytometry can be used to differentiate and quantify a range of species populations present in a mixed medium.

Light Scattering and Turbidity. Optical techniques based on either the scattering of light (nephelometry) or the degree of transmitted light or optical density (turbidity) have been developed for determining biomass (44). By monitoring the degree of light scattering using nephelometry, both cell numbers and mass can be determined. This approach is particularly suited for bioprocesses that involve low cell concentrations, where the background (compared) level is near zero. Turbidity measurement can be used for both on-line and off-line determination of biomass (54). Light scattering by a turbid medium results from a number of factors including particle size, shape, and number. Quantification is based on the Lambert-Beer law (55).

Biosensors

One possible answer to the problem of monitoring metabolites, both in situ and on-line, may be solved in the near future by the use of biosensors. Since their conception in the 1960s, these devices have generated considerable interest. This has spread to a diverse range of fields including clinical diagnostics, environmental protection, bioprocess monitoring, and defence applications (56). In general the operation of a biosensor is characterized by three functional steps: recognition, physiochemical signal generation, and signal processing. The biological component (e.g., enzyme, whole cell, antibody, and cell receptor) imparts a high degree of selectivity on the biosensor. Coupled to the biological component; the transducer is designed to respond to the changing physicochemical parameters caused by the specific interaction of the biological component with the substrate. A range of transducers have been used in the development of biosensors and include the electrochemical (potentiometric and amperometric), optical, calorimetric, piezoelectric, and thermometric (57). Most of these types have been adopted in the development of biosensors intended for bioprocess monitoring. For a more detailed insight into both the fundamentals and potential for biosensors, a number of specialized publications are available (58). Enzyme-modified field effect transistors (FETs), whereby the biologically active layer is positioned on top of the ion electrode membrane, have also been used for bioprocess monitoring (59).

An off-line biosensor system (Yellow Spring, Inc.) has been available for a number of years (Fig. 7). The instrument can be used to determine a range of metabolites (e.g., glucose, lactate, and ethanol). Recently the company has introduced an on-line instrument. Indeed, on-line biosensor systems have been demonstrated using a variety of biological-transducer systems (60,61). A wide range of metabolites have been monitored using various bioprocess regimes (62). Commercially, a significant number (63) of biosensor devices are available for measuring a range of analytes. However, at present, the majority of these devices are aimed at the medical diagnostic market. The goal is to develop cheap, reliable sensors that can operate under a range of bioprocess conditions for extended periods with a minimum of maintenance.

Nuclear Magnetic Resonance

Nuclear magnetic resonance is based on the detection of the response from particular nuclei when exposed to a

Figure 7. The YSl 2700 biochemistry analyzer.

magnetic field and electromagnetic radiation. Following absorption, the resonance frequency is shifted in a characteristic response pattern, according to the environment of the sample under detection. This approach can be used to determine a range of intracellular factors such as ATP, ADP, sugar phosphate, and polyphosphate, as well as pH. In these examples, concentrations of ^{31}P are used to characterize the compounds. This is an off-line monitoring system that is currently expensive; hence, its use is primarily in the research field, not in routine production environments.

Artificial Intelligence

The data obtained using the various measuring instruments already discussed are invariably used to control and optimize the bioprocess being carried out. Recent developments in the field of artificial intelligence have led to investigations into the use of such systems for improving bioprocess control, based on the received measurement output signals. This has included the use of both knowledge-based expert systems (64,65) and neural networks (e.g., 66,67) during bioprocess operation. A recent report (68) described the successful use of a neural network as a tool for evaluating the received measurement from an enzyme (penicillin-G amidase) pH-FET sensor linked to a flow injection system.

Undoubtedly, the adaptation of such "intelligent" systems will develop over the coming years and will play an important role in the precise control of bioprocess applications.

CONCLUSION

Although a diverse range of monitoring equipment is available, only a relatively narrow range of the more reliable instruments is used routinely in practice. As culture techniques become more elaborate and high-value-added products are produced, conventional methods will either prove insufficient or require supplementation with a range of sensors able to directly monitor key process parameters. Despite a clear demand for new sensors (e.g., lactate, glutamine, and glutamate for animal cell culture), cell cultivation represents only a modest market for analytical instrument manufacturers. Hence, while progress is to be expected, it may be slower than might be wished.

BIBLIOGRAPHY

1. B. Mattiasson, H. Hakanson, *TIBTECH* 11, 136–142 (1993).
2. S.L. Brooks, I.J. Higgins, J.D. Newman, and A.P.F. Turner, *Enzyme Microb. Tech.* **13**, 946–955 (1991).
3. J. Ruzicka and E.H. Hansen, *Flow Injection Analysis*, 2nd ed., Wiley & Sons, New York, 1988.
4. C.C. Westcott, *pH Measurements*, Academic Press, New York, 1978.
5. P. Bergveld, *IEEE Trans. Biomed. Eng.* BME **19**(5), 342–353 (1972).
6. J.M. Thompson, *Med. Biol. Eng. Comput.* **28**, B29–B33 (1990).
7. M.N. Pons, in M.N. Pons ed., *Bioprocess Monitoring and Control*, Carl Hanser Verlag, Munich, 1991, pp. 4–28.
8. P.F. Stanbury and A. Whitaker, *Principles of Fermentation Technology*, Pergamon Press, Oxford, (1984).
9. B.G. Liptak, in B.G. Liptak ed., *Instrumentation in the Process Industries*, Chilton, Philadelphia, 1969, pp. 38–94.
10. H. Schafer, V. Graeger, and R. Kobs, *Sensors Actuators* **17**, 521–527 (1989).
11. A.H. Scragg, *Bioreactors in Biotechnology: A Practical Approach*, Ellis Horwood, Chichester U.K., 1991.
12. P.M. Doran, *Bioprocess Engineering Principles*, Academic Press, London, 1995.
13. Y.H. Lee and G.T. Tsao, *Adv. Biochem. Eng.* **13**, 35–86 (1979).
14. B. Atkinson and F. Mavituna, *Biochemical Engineering and Biotechnology Handbook*, 2nd ed., Macmillan, Basingstoke, U.K., 1991, Chapter 18.
15. K.G. Carr-Brion, in K.C. Carr-Brion ed., *Measurement and Control in Bioprocessing*, Elsevier Science, New York, 1991, pp. 37–66.
16. L.C. Clark and C. Lyons, *Ann. N.Y. Acad. Sci.* **102**, 29–45 (1962).
17. P. Van Hemert, D.G. Kilbrun, R.C. Righelato, and A.L. Van Wezei, *Biotechnol. Bioeng.* **11**, 549–560 (1969).
18. G. Quicker, A. Schumpe, B. Konig, and W.-D. Deckwer, *Biotechnol. Bioeng.* **23**, 635–650 (1981).
19. D.W. Lubbers and N. Opitz, in J. Grote, D. Reneau, and G. Thews eds., *Oxygen Transport to Tissue, II*, Plenum, New York, 1976, pp. 65–68.
20. M.A. Arnold, *Anal. Chem.* **64**, 1015A–1025A (1992).
21. J. Janata, in J. Janata ed., *Principles of Chemical Sensors*, Plenum Press, New York, 1989, pp. 241–283.
22. M.R. Surgi, in D.L. Wise ed., *Applied Biosensors*, Butterworths, Boston, 1989, pp. 249–290.
23. R.S.D. Wolthuis, J.C. McRae, E. Hartl, G.L. Saaski, K. Mitchell, and R. Willard, *IEEE Trans. Biomed. Eng.* **39**, 185–193 (1992).
24. K.W. Berndt and J.R. Lakowicz, *Anal. Biochem.* **201**, 319–325 (1992).
25. D.H. Philips and M.J. Johnson, *J. Biochem. Microbiol. Technol. Eng.* **3**, 261–275 (1961).
26. J.E. Brown, R.K. Kaminski, A.C. Blake, and A. Brodgesell, in B.G. Liptak ed., *Instrument Engineer's Handbook*, vol. 1, Chilton, Philadelphia, 1969, pp. 713–918.
27. Y. Kawabata, T. Kamichika, T. Imasaka, and N. Ishibashi. *Anal. Chim. Acta* **219**, 223–229 (1989).
28. L. Kjaergard, in T.K. Ghose ed., *Advanced Biochemical Engineering*, Springer-Verlag, New York, 1977, pp. 131–145.
29. M.H. Gordon and R. Macrae, in *Instrumental Analysis in the Biological Sciences*, Blackie, Glasgow, 1987, pp. 133–145.
30. K. Holzhauer-Rieger, W. Zhou, and K. Schugerl, *J. Chromatogr.* **499**, 609–615 (1990).
31. W. Smolenski and J.M. Suflita, *J. Microbiol. Methods* **6**, 71–79 (1987).
32. H. Kurokawa, Y.S. Park, S. Iijima, and T. Kobyashi. *Biotechnol. Bioeng.* **44**, 95–103 (1994).
33. K. Schugerl, in C. Christiansen, L. Munck, and J. Villadsen eds., *Proceedings of the 5th European Congress on Biotechnology*, vol. II, Munksgaad International, Copenhagen, 1990, pp. 1188–1193.
34. K. Schugerl, in K. Schurgerl ed., *Biotechnology*, vol. 4, VCH, New York, 1991, pp. 149–180.

35. D.M. Comberbach, J.M. Scharer, and M. Moo-Young, *Biotechnol. Lett.* **6**, 91–96 (1984).
36. J.K. McLaughlin, C.L. Mayer, and E.T. Papoutsakis, *Biotechnol. Bioeng.* **27**, 1246–1257 (1985).
37. C. Fujimoto and K. Jinno, *Anal. Chem.* **64**, 476A–481A (1992).
38. M. Reuss, H. Piehl, and F. Wagner, *Eur. J. Appl. Microbiol.* 1, 323–325 (1975).
39. S.J. Coppella and P. Dhurjati, *Biotechnol. Bioeng.* **29**, 679–689 (1987).
40. E. Heinzle, *Adv. Biochem. Eng. Biotechnol.* **35**, 1–45 (1987).
41. S. Chauvatcharin, K.B. Konstantinov, K. Fujiyama, T. Seki, and T. Yoshida, *J. Fermen. Bioeng.* **79**, 465–472 (1995).
42. M. Reuss, in R.K. Finn and P. Prave eds., *Biotechnology Focus I: Fundamentals.Applications.Information*, Oxford University Press, New York, 1988, pp. 153–190.
43. B. Sonnleitner, G. Locher, and A. Fiechter, *J. Biotechnol.* **25**, 5–22 (1992).
44. K. Konstantinov, S. Chuppa, E. Sajan, S. Yoon, and F. Golini, *TIBTECH* **12**, 324–333 (1994).
45. C.M. Harris and D.B. Kell, *Biosensors* **1**, 17–84 (1985).
46. E.W. Chappelle, G.L. Piccioli, and J.W. Deming, *Methods Enzymol.* **57**, 656–672 (1978).
47. B.C. Blake-Coleman, D.J. Clarke, M.R. Calder, and S.C. Moody, *Biotechnol. Bioeng.* **28**, 229–235 (1986).
48. D.G. Kilburn, P. Fitzpatrick, B.C. Blake-Coleman, D.J. Clarke, and J.B. Griffiths, *Biotechnol. Bioeng.* **33**, 1379–1484 (1989).
49. K. Mishima, A. Mimura, Y. Takahara, K. Asami, and T. Hanai, *J. Ferment. Bioeng.* **72**, 291–295 (1991).
50. R.M. Matanguihan, K.B. Konstantinov, and T. Yoshida, *Bioprocess Eng.* **11**, pp. 213–222 (1994).
51. G. McMichael, W.B. Armiger, J.F. Lee, and R. Matharasan, *Biotechnol. Techn.* **1**, 213–218 (1987).
52. J.K. Li and A. Humphrey, *J. Ferment. Bioeng.* **74**, 104–111 (1992).
53. K.B. Konstantinov, P. Dhurjati, T. Van Dyke, W. Majarian, and R. LaRossa, *Biotechnol. Bioeng.* **42**, 1190–1198 (1993).
54. B. Sonnleitner, G. Locher, and A. Fiechter, *J. Biotechnol.* **25**, 5–22 (1992).
55. K.F. Reardon and T. Scheper, in K. Schugerl ed., *Biotechnology*, vol. 4, VCH, New York, 1991, pp. 179–223.
56. E. Kress-Rogers, *Handbook of Biosensors and Electronic Noses*, CRC Press, Boca Raton, Fla., 1997.
57. R.F. Taylor and J.S. Schultz, *Chemical and Biological Sensors*, IOP Publishing, Bristol, U.K., 1996.
58. A.P.F. Turner, *Advances in Biosensors*, suppl. 1, 2, 3, JAI Press, London, 1991, 1992, 1993, 1994.
59. T. Kullick and R. Ulber, in R. Freitag ed., *Biosensors in Analytical Biotechnology*, Academic Press, San Diego, 1996, pp. 23–33.
60. B. Mattiasson and B. Danielsson, in R.F. Taylor and J.S. Schultz eds., *Chemical and Biological Sensors*, IOP Publishing, Bristol, 1996, pp. 533–549.
61. U. Bilitewski and I. Rohm, in E. Kress-Rogers ed., *Handbook of Biosensors and Electronic Noses*, CRC Press, Boca Raton, Fla., 19, 435–468 (1997).
62. U. Bilitewski and I. Rohm, in R. Freitag ed., *Biosensors in Analytical Biotechnology*, Academic Press, San Diego, 1996, pp. 163–191.
63. J.D. Newman and A.P.F. Turner, *Chemistry & Industry*, May 1994, pp. 374–378.
64. M.N. Karim and A. Halme, in N.M. Fish, R.I. Fox, and N.F. Thornhill, eds. *Computer Applications in Fermentation Technology: Modelling and Control of Biotechnological Process*, Elsevier Applied Science, Barking, U.K., 1989, pp. 37–46.
65. K.B. Konstantinov and T. Yoshida, *Biotechnol. Bioeng.* **39**, 479–486 (1992).
66. J. Thibault, V. van Breusegem, and A. Cheruy, *Biotechnol. Bioeng.* **36**, pp. 1041–1048 (1990).
67. Q. Zhang, J.F. Reid, J.B. Litchfield, J. Ren, and S.-W. Chang, *Biotechnol. Bioeng.* **43**, 483–489 (1994).
68. B. Hitzman, T. Kullick, R. Quack, and M. Venschott, in R. Freitag ed., *Biosensors in Analytical Biotechnology*, Academic Press, San Diego, 1996, pp. 129–144.

See also CONDUCTIVITY; PILOT PLANTS, DESIGN AND OPERATION; PROCESS CONTROL, STRATEGY AND OPTIMIZATION; SECONDARY METABOLITE PRODUCTION, ACTINOMYCETES, OTHER THAN STREPTOMYCES; YEAST, BAKER'S.

PROCESS VALIDATION

DANE W. ZABRISKIE
Biogen, Inc.
Cambridge, Massachusetts
THOMAS M. SMITH
ALAN R. GARDNER
SmithKline Beecham
King of Prussia, Pennsylvania

KEY WORDS

Cell culture
Clearance of contaminants
Consistency
Fermentation
Good manufacturing practices (GMP)
Operating parameters
Operating ranges
Proven acceptable ranges
Purification
Validation
Viral clearance

OUTLINE

INTRODUCTION

Process validation for pharmaceuticals has been defined by the U.S. Food and Drug Administration (FDA) as "establishing documented evidence which provides a high degree of assurance that a specific process will consistently produce a product meeting its pre-determined specifications and quality attributes" (1). In spite of the apparent clarity of this definition, pharmaceutical companies vary widely in their approaches to process validation. Operational definitions of process validation are often developed by firms to suit their own validation philosophies. Subcategories of process validation such as formal process validation, process characterization, process development, process evaluation, and the like can result, and the meaning of these phrases is highly specific to a particular company. For the purposes of this article, the term *process validation* refers to any activity consistent with the FDA definition. The discussion that follows is necessarily subjective, and it does not reflect the view of any particular company or the industry at large. Rather, it describes concepts used by those who practice this craft in industry.

This article is concerned with processes that produce the active ingredient in a biopharmaceutical product for use in humans as therapeutics or vaccines. Biopharmaceuticals, protein agents derived from recombinant gene expression or hybridoma production systems, include various biotechnology products such as monoclonal antibodies, enzymes, hormones, soluble receptors, growth factors, and immunogens. Pharmaceutical manufacturing processes are divided into two subprocesses. The primary process produces the bulk biological substance (BBS), also referred to as the active pharmaceutical ingredient (API). The secondary process transforms the BBS into the final product, which is what is actually sold. Secondary processing steps include formulation, sterile filtration, filling, lyophilization (in some cases), capping, and packaging. This article is limited to the validation of the primary processes which produce biopharmaceutical BBS.

Pharmaceutical manufacturers typically classify validation activities into the following broad categories: process validation, equipment and facility validation, cleaning validation, analytical methods validation, and computer validation. Equipment and facility validation is further subdivided into design qualification, installation qualification (IQ), operational qualification (OQ), and performance qualification (PQ). Process validation is viewed by some experts to be an element of PQ.

Process validation is a regulatory requirement to enable product licensure. Limited process validation is also critical for clinical trial materials to ensure product safety. Process validation regulations originated within the FDA, which was created in 1906, initially to prevent the adulteration of products. The sulfanilamide disaster led the Food, Drug, and Cosmetics (FD&C) Act in 1938 to ensure product safety. The thalidomide disaster in 1962 led to the Kefauver–Harris Amendment to the FD&C Act, which required demonstration of product safety and efficacy. It was in this revision that the concept of good manufacturing practices (GMP) originated. In 1978 the act was further revised to include the concepts underlying current GMPs (cGMPs) and good laboratory practices (GLPs). The origins for process validation derive from FDA draft guidelines and workshops that occurred in the late 1970s and early 1980s (2). *Guideline on General Principles of Process Validation* (1), published in 1987, represented the first (and still only) regulatory document focusing on this topic. The definition for *process validation* quoted earlier was taken from this FDA document and has gained wide acceptance by other regulatory agencies around the world. This guidance was specifically intended for the production of finished drug products. Today, the regulatory basis in the United States for process validation is contained within cGMPs for finished pharmaceuticals (drugs) (3). The *Code of Federal Regulations* (21 CFR 211.110 states that "control procedures shall be established to monitor the output and validate the performance of those manufacturing processes that may be responsible for causing variability in the characteristics of in-process material and the drug product." It (21 CFR 211.113) also states that "appropriate written procedures, designed to prevent microbiological contamination of drug products purporting to be sterile, shall be established and followed. Such procedures shall include validation of any sterilization process."

Until recently, these FDA validation requirements and guidelines only applied explicitly to drug products and did not apply to bulk biologics. Nevertheless, there was a trend for these regulatory principles to apply implicitly to areas more broadly than originally defined. There are numerous publications by industrial scientists, but currently no umbrella regulatory document exists on the validation of bulk production processes for biopharmaceuticals. Important aspects to this subject are captured within numerous other regulatory documents that focus on other aspects of manufacturing and control. Table 1 summarizes some of these references. Regulatory requirements that pertain to process validation are changing and are likely to continue to change for the foreseeable future. The most current information on process validation and many other regulatory subjects may be accessed on the Internet (Table 2). Proposals published recently by the FDA pertain to the cGMPs for finished pharmaceuticals and include content on process validation (18). How these proposals will be extended to bulk biological products has become clearer with the recent publication by the FDA of a draft guidance for the manufacturing of APIs (19) and the guidance for submission of the chemistry, manufacturing, and controls (CMC) section of a biologics license application (BLA) for most biopharmaceutical products (20).

In addition to being a regulatory requirement, process validation provides important benefits to the manufacturer. Biopharmaceuticals are among the most expensive products to manufacture, costing hundreds to thousands of dollars per gram to produce. The assurance of process consistency therefore enables a manufacturer to maximize plant productivity and obtain favorable cost efficiency by reducing the incidence of process failures, rejection of lots, and reworking of rejected but salvageable lots. In today's manufacturing environment where plant capacity is closely matched with product demand, unexpected process failures can lead to product shortages, with negative financial and company product image impacts, particularly

Table 1. Selected References on Process Validation

Validation topic	Comment	Reference
Viral and contaminant clearance, cell line stability	A good general discussion of these topics with examples from Genentech's Activase® experience	4
Purification and general topics	A special issue of BioPharm that compiles 12 papers published in this journal in 1994 and 1995	5
Viral clearance	ICH guideline common to the United States, Europe, and Japan	6
Cell line characterization	ICH guideline common to the United States, Europe, and Japan	7
General	The first and still only regulatory document in the world dedicated to this subject; some content is becoming obsolete	1
Prospective, concurrent, and retrospective studies	Focuses on sterile drug products but contains useful general guidance	8
Operating ranges, proven acceptable ranges	An early discussion of this subject	9
Cell age	Describes scientific approaches to demonstrating consistency of cell line characteristics used at Genentech for tPA	10
Chromatography	Industrial perspective on the validation of chromatography purification steps	11
General	Inclusive discussion for biotech products beginning with master cell bank and ending with the final product; includes example from the tPA process	12
Experimental design	Application of design of experiment (DOE) techniques	13
Tangential flow filtration	Industrial perspective on the validation of tangential flow filtration steps	14
Cell line characterization and viral clearance	FDA Points to Consider in the Characterization of Cell Lines Used to Produce Biologicals (1993)	15
Cell line characterization, viral clearance, and manufacturing changes	FDA Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products for Human Use (1997)	16
Viral clearance	Note for guidance published by the CPMP for use in the European Union (1996)	17

Note: ICH, International Conference on Harmonization; CPMP, Committee for Proprietary Medicine.

Table 2. Internet Sites with Current Information on Process Validation and Other Regulatory Information

For Searchable GMPs for the FDA
http://www.access.gpo.gov/cgi-bin/cfrassemble.cgi
ICH Documents
http://www.ifpma.org/ich5.html
FDA Documents
http://www.fda.gov
European Union (European Union Drug Regulatory Authorities)
http://www.eudra.org/frame/frametest1.html

for life-saving products. Process validation is a key means to minimize the occurrence of unexpected process failures. Process validation can also provide a manufacturer with operational flexibility to respond to inevitable production problems (e.g., power failures, equipment breakdown, and scheduling conflicts). To the extent that they can be anticipated in the process validation program, ranges can be established for operational parameters that, if maintained, will ensure that the products meet specifications.

These important benefits only accrue after considerable investment because process validation is costly and time consuming. It occupies the full period from producing supplies for the first clinical trial through the start of production in the commercial plant and continues throughout the commercial lifetime of the product. During product development, for example, more resources from process development, operations, quality control, and quality assurance groups can be consumed in validating the process, equipment, and facility than were used to develop the process and to produce clinical supplies. Nevertheless, careful planning and sound scientific strategies make process validation cost effective and meet regulatory requirements.

There are several different classes of process validation. *Clearance validations* demonstrate the capability of a process to remove or inactivate contaminants to and beyond the degree necessary to ensure product safety. Of greatest concern are contaminants that have toxic, pharmacological, or infectious potential in humans. Typical contaminants in biopharmaceutical processes with these potential liabilities include DNA and proteins derived from the recombinant host cell line, certain components of culture media (e.g., methotrexate and other drugs used to maintain selection pressure, insulin, bovine serum albumin, growth factors, and animal-derived components of culture media), viruses derived from the host cell line or raw materials, additives for viral inactivation (e.g., guanidine), and affinity chromatography ligands (e.g., monoclonal antibodies, protein A). *Validation of operating ranges* involves the determination of a range for critical operating parameters through which excursions can occur without affecting the quality of the product. The usage of the term *operating parameters* here is taken very broadly to include not only compositions, temperature, pH, and related factors but also holding times of in-process intermediates, variations in raw materials, cycle limits on multiuse resins, age of

raw materials, and age limit of cells used in the process. *Validation of process consistency* must be established through the characterization of multiple lots of product produced by separate production runs operated and controlled under the same conditions. Each of these classes of validation are described in greater detail in this article.

PROCESS VALIDATION STRATEGY AND DOCUMENTATION

Process validation generally starts after a process is defined in development and evolves with the progression of clinical trials. Process validation for phase I and II trials is limited to addressing potential product safety concerns. Validations for the clearance of model viruses and retroviruses and host cell DNA are sufficient. More extensive validation activities occur in parallel with production of phase II and III clinical supplies and the conducting of the phase II and III clinical studies. At this stage, the virus clearance validations are expanded to include a greater number of model viruses spanning a greater range of virus characteristics. The full range of clearance and operating range validations are undertaken in a pilot plant or development laboratories. Process consistency validation, however, is derived from the pilot and commercial plants. License applications can be filed with consistency data derived from a pilot plant. However, the awarding of a license is contingent on the filing of process and product consistency data derived from a pilot plant or commercial manufacturing facility. In the United States, three qualification lots are usually sufficient; the European Union usually requires five lots. Process validation continues after product launch for the length of time the product is manufactured. Validation maintenance programs often call for revalidation of specific aspects of the process at regular intervals. Trending or control charting of process variables can lead to refinements of operating range limits. Revalidation is also required when the process is modified because of improvement, equipment alterations, raw material changes, process scale-up, substitution of a processing step with an alternate, resequencing process step, and so forth.

Validation is subdivided into prospective, concurrent, and retrospective according to the timing relationship between the validation work and production (1). Prospective validation studies "are conducted prior to the [commercial] distribution of a new product, or product made under a revised manufacturing process, where the revisions may affect the product's characteristics" (1). Concurrent validation studies employ data derived from production runs for product in distribution. How prospective and concurrent validations are conducted is often defined in preapproved validation protocols. Retrospective validation studies demonstrate, "in some measure, the adequacy of the process by examination of accumulated test data on the product and records of the manufacturing procedures used. Retrospective validation can also be useful to augment initial premarket prospective validation for new products or changed processes" (1). Retrospective validation is generally not conducted according to a preapproved validation protocol, but this practice may change in the future.

In reality, most comprehensive process validation programs involve all three types of validation throughout the lifetime of a product. Consistency validations are most often run prospectively in full-scale production equipment for clinical supplies or commercial purposes. Concurrent and retrospective validations can be used to set new operating ranges not covered by previous validation work or to modify (usually narrow) operating ranges based on manufacturing experience using statistical process control, control charting, or related techniques. A final precaution should be taken regarding recent regulatory concerns on retrospective validation. Retrospective validations are acceptable as an adjunct to a well-defined and comprehensive prospective validation program, but they cannot substitute for it.

Another technical consideration to be defined is the scale at which process validation studies are to be conducted. One could assume that the most authoritative results would be obtained by conducting validation experiments at the scale intended for commercial manufacturing. This assumption normally leads manufacturers to validate process consistency using the equipment to be licensed. It is not unusual, however, to identify 50 operating parameters for which low and high operating range limits are sought. This could involve 100 experiments (if interactions among parameters were disregarded) and up to 1×10^{15} experiments (if interactions are included), which is clearly not feasible in full-scale equipment. The infectivity and limited quantity of viral reagents used in clearance validations prevent these studies from being executed at full scale or in a GMP facility. GMP prohibits the intentional introduction of infectious agents into processing areas (21). In other clearance validations, limitations on the quantity of reagents used for spiking (e.g., host cell proteins, host cell DNA, and affinity ligands) prevent these studies from being conducted in full-scale equipment.

One approach to these problems is to scale down processing steps. Scale-down factors from 10- to 10,000-fold have been employed in industry. In some cases, validation experiments can be run in parallel using multiple identical scaled-down units (e.g., fermentors), which enable statistical design of experimental techniques and a significant shortening in the overall time expended for validation experiments. In other cases a hybrid approach can be taken in which samples of in-process streams taken from the full-scale process are used as feed streams to scaled-down unit operations to assess performance. Virus clearance studies that employ live viruses and retroviruses are conducted in this manner. The use of scaled-down systems for virus clearance validations also facilitates adequate biological containment.

If scaled-down models are to be employed, then it is necessary to establish that the models are predictive of full-scale performance. A useful scaled-down model may be developed by attending to three considerations: performance, design, and quality (PDQ). The performance of the scaled-down model should be identical to or at least predictive of the full-scale process. Typical performance characteristics are titer, cell density, viability, processing rates, and step yields. In cases where there is a performance difference between the scales, it should be quantified and the basis for the difference should be understood from technical and

scientific principles whenever possible. The degree of variation in the model should not be greater than variation at full scale to ensure that process changes are distinguishable from natural variation in the model. During the design of the scaled-down model, it is important to observe the same technical principles used to scale up a unit operation. Typical examples are maintaining constant culture split ratios in cell culture and to preserve chromatography bed heights and linear flow velocities in purification (22,23) during scale-up or scale-down. The quality of the intermediate produced by each scaled-down unit operation should be identical or similar to its full-scale counterpart in its analytical attributes. Finally, it is important to document the qualification of scaled-down models used for process validation.

As the definition of *process validation* states, documentation of validation work is essential. The validation master plan defines the overall validation philosophy, the systems to be validated, the general testing scheme to be employed, the responsibilities of various groups, and the approval mechanisms of the program. It may also define the evolution of validation throughout the life cycle of the process, from its origins in process development through validation maintenance during commercial manufacturing. Another class of document is the validation protocol, which prospectively defines the details for testing of each validation stage. The validation protocol is defined by the FDA as "a written plan stating how validation will be conducted, including test parameters, product characteristics, product equipment, and decision points on what constitutes acceptable test results" (1). The final document is the process validation report, which includes the validation protocol, deviations from the protocol, the results from the validation experiment, and the conclusion regarding the validation status for the processing stage. An important phrase in the definition of process validation is establishing documented evidence. Validation experiments must be documented in reports to be acceptable to regulatory agencies. Copies of the most critical validation reports and summaries of the others are included as part of the license application (20). All other process validation reports must be available to the regulators for inspections as part of the licensing process or during routine postlicense inspections. Failure to submit essential process validation reports can lead to a refusal to file determination by the FDA in the United States. A failure to produce appropriate validation reports during regulatory inspections can lead to a rejected license application or other regulatory actions such as the issuance of an FDA 483 form.

CLEARANCE VALIDATIONS

Clearance validations are generally concerned with the capacity of the recovery and purification process to remove process contaminants. Of particular concern are contaminants with toxic, pharmacological, or infectious potential in humans; studies of these risks are thus among the first to be undertaken in a validation program. The overall approach to this form of validation can be illustrated using the clearance of DNA.

DNA contaminants could contain genes that encode harmful proteins (e.g., toxins and oncogene products), viruses, or retroviruses endogenous to the host. If these gene contaminants could transfect patient cells (resulting in retention and duplication of the gene), then expression of the gene product is a theoretical possibility. The World Health Organization (WHO) has established a 100 pg/dose limit to DNA contaminants in biopharmaceutical products. Proposals are under consideration to increase this limit to 10 ng/dose. One approach to validating adequate clearance of DNA by the process is to measure the DNA level in the BBS over the full range of process operating conditions. If these measurements fall below the 100 pg/dose limit, then the process is validated for this clearance capability. This approach is taken for contaminants whose assays are sufficiently sensitive to quantify levels in the BBS.

Often the assays for DNA and many other process contaminants are not sufficiently sensitive (24) to quantify the contaminant in the bulk product, and an alternate validation approach must be used. Clearance factors are established for upstream unit operations for which DNA levels can be measured in the in-process streams. For downstream steps for which DNA levels are immeasurable, host cell DNA is added (spiked) to the feed stream for each unit operation at a level sufficient to detect it in the output stream. Clearance factors for the downstream steps can then be calculated. The clearance factors are then summed together to give a total clearance factor for the process. This total may be adjusted downward to eliminate double counting steps that have similar mechanisms of clearance. If, for example, a process includes two anion exchange steps, only one would be included in the analysis of clearance. This is based on the notion that virus particles capable of being removed by this mechanism (i.e., anion exchange) would only be removed by the first step. No further clearance could be expected in the second or additional anion exchange steps. The total clearance factor is then applied to the range of DNA levels present in unprocessed bulk harvests. If the calculated DNA level in the BBS is below the 100 pg/dose level, then the process is validated with respect to DNA clearance.

Spiking studies frequently use a hybrid of the full-scale process and scaled-down process models during validation. During full-scale operation, samples of in-process streams are collected and stored. Portions of these samples are spiked with the appropriate contaminant(s) and are then processed over a scaled-down model. In some cases, the spiked species may be labeled with a radioactive label to facilitate quantitative detection in in-process streams. This approach is often used for measuring DNA clearance.

When animal cells are employed during production, clearance of putative viruses constitutes the largest group of clearance validations. The general approach is similar to that described for DNA clearance (25). The product safety concern is to minimize the risk of viruses or retroviruses contaminating the product. The viruses can originate as active or latent viruses associated with the host cells (i.e., endogenous viruses) or can be introduced with raw materials (i.e., adventitious viruses), particularly those derived from animals. In most cases, if viruses exist at all, they are present at levels below detection in in-

process streams, making spiking–clearance studies necessary for all steps in purification. For infectious contaminants such as viruses, there are two modes of clearance. The first involves removal by each purification step, similar to the manner in which DNA is removed. The second involves inactivation, where additives or conditions are used to disable the virus without removing it.

The determination of the viral load to the process depends on the specifics of the host cell system. For Chinese hamster ovary (CHO) cells, the load can be estimated by measuring the level of retrovirus-like particles (RVLPs) present in the unprocessed bulk harvest. CHO cells are known to produce RVLPs, which resemble intracytoplasmic type A retroviruses and cell membrane budding type C retroviruses (26). Even though these particles are defective in their ability to infect other cells, they nevertheless provide a basis for estimating a retroviral load in the process. Levels of 10^5 to 10^7 RVLPs per milliliter of unprocessed bulk are typical. The level of viral clearance to target depends on the perceived risk to the patient. One target for retroviral clearance in processes involving CHO cells is to allow no more than 1 RVLP per million doses of final product (4). The total process retroviral clearance and the level of RVLPs in the unprocessed bulk are used to calculate an estimate of the RVLPs in the final product. The process is considered to be validated for retroviral clearance when the estimated level of RVLPs in the final product is less than or equal to the target. Process clearance factors of 15–20 logs are typical for murine retroviruses. The FDA has suggested that a minimum of 3 logs of clearance is necessary based on the initial level of contamination entering the purification process (16).

For most viruses other than endogenous retroviruses, the validation exercise becomes more academic; there is no reasonable basis for establishing the virus load in the process because there should not be any present. In these cases, measured clearance factors are reported to provide some assurance that if these types of viruses were inadvertently introduced into the process, there would be a defined capacity to clear them.

The spiking studies use model viruses that are representative of the most likely contaminants and when grouped together cover a broad spectrum of virus characteristics. Clearance studies are often performed for a number of viruses to test the ability of the process to clear viruses in general. The choices of model viruses depend on the cell line employed and the viruses could contain. For CHO cells, typical model viruses include a small, nonenveloped virus (e.g., simian virus [SV40], human polio virus 1, animal parvovirus), a medium-to-large enveloped RNA virus (parainfluenza virus, influenza virus, Sinbis virus), a medium-to-large DNA virus (e.g., herpes simplex virus 1 and pseudorabies virus), and a murine retrovirus (e.g., murine leukemia virus) (6). The first three viruses are chosen because they cover a range of virus characteristics, and the murine retrovirus is chosen to model one that could infect the CHO host cell line.

OPERATING RANGES

The purpose of validating process operating ranges is to demonstrate the robustness of the process to excursions in operating parameters that are likely to be encountered in a practical manufacturing environment. The process can be subdivided into a series of validation stages or modules. In some cases, a stage corresponds to a specific processing step (e.g., production fermentation or a chromatography step). In other cases, a stage may be a series of steps (e.g., media preparation or seed expansion). The choice of defining validation stages depends on the specifics of the process and the manner in which the process is controlled. One validation staging strategy for a typical biopharmaceutical process is shown in Figure 1.

Once the stages are identified, inputs and outputs to each stage are defined. Typical inputs include the operating parameters, properties and storage duration of the feedstock derived from the previous upstream step, and quality attributes of critical raw materials used in the step. Typical outputs include the performance attributes for the stage (e.g., titer and yield) and the properties of the product stream (product homogeneity, purity, contaminant levels, and chromatogram shape). For chromatographic steps, the shape of the chromatogram has been used to indicate step performance. Suggested quantitative measures of chromatogram shape have been proposed, such as SPEW, for the retention times (or chromatogram distance) when

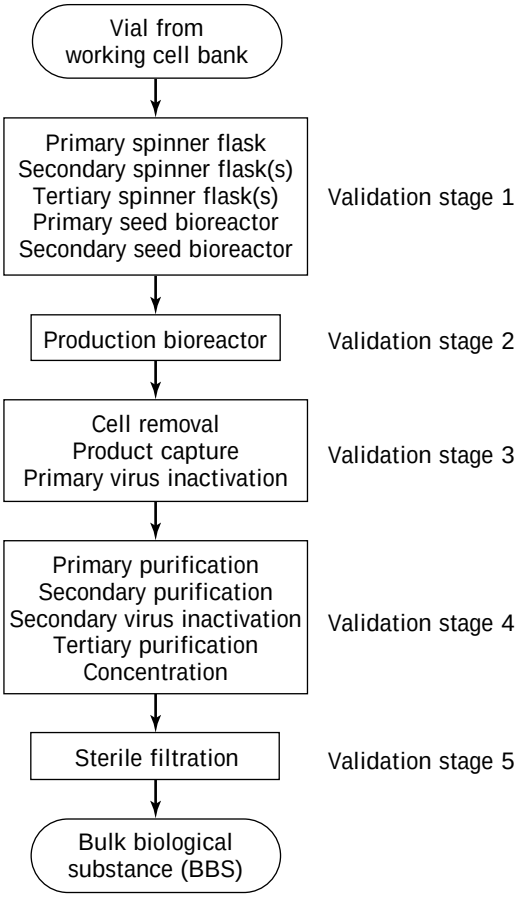


Figure 1. Validation of a process can be simplified by dividing the process into a sequence of validation stages, some of which include multiple processing steps.

the product collection starts, peaks, and ends and the width of the elution peak (27). In general, all in-process controls must be included within the list of process step outputs. Representative inputs and outputs for a fermentation stage and a purification stage are illustrated in Figure 2.

It is usual to prioritize the inputs to be validated based on the *(1)* sensitivity of process outputs to variations in each input, *(2)* proximity of the validation stage to the final product, and *(3)* technical difficulty in controlling an input to within a specified range. Assessment may be based on an analysis of all data that are available and documented, including process development, monitoring of clinical manufacturing, and any experiments specifically designed to gauge the sensitivity of the process to variations. The most important stage inputs must be included in the operating range validation program.

Having identified the parameters to be validated, the next step is to define the criteria on the outputs that will be used to determine whether a validation experiment is successful. These criteria are typically a mixture of process performance results (e.g., titer, yield, and concentration) and quality of the product intermediate (e.g., protein purity, product homogeneity, and contaminant levels). The need for both types of criteria is based on the assumption that no amount of testing of the end product can ensure consistency without also showing the process to be controlled. These criteria are based on a variety of considerations and experience with the process. In some cases, these criteria are based on the output range results from operational history of the process during development or clinical supply production. In other cases, they may be based on equipment limitations. For example, a minimum titer from the production fermentation tank may be established by the resin bed size and lower capacity limit for the first chromatography step in a process. In still other cases, output criteria may be derived from product specifications. If, for example, product purity is primarily determined by the performance of the first purification step of a five-step process producing a monoclonal antibody, then the product purity specification can be used to define a purity output criterion for the first step.

It is then necessary to define an overall experimental paradigm. Usually, the experimental program begins with a set of proposed operating ranges for the inputs to be evaluated for each validation stage. These ranges may be based on equipment capability, theoretical considerations, process development data, operating history, the need for flexibility in manufacturing, or other considerations. Each validation stage is then run using inputs at the limits of the proposed operating ranges. These experiments can be run to investigate one input at a time or in parallel using statistical design of experiment (DOE) techniques. The outputs from these experiments are obtained, and the proposed operating ranges become proven acceptable ranges (PARs) (9) if the output criteria are met. If the output cri-

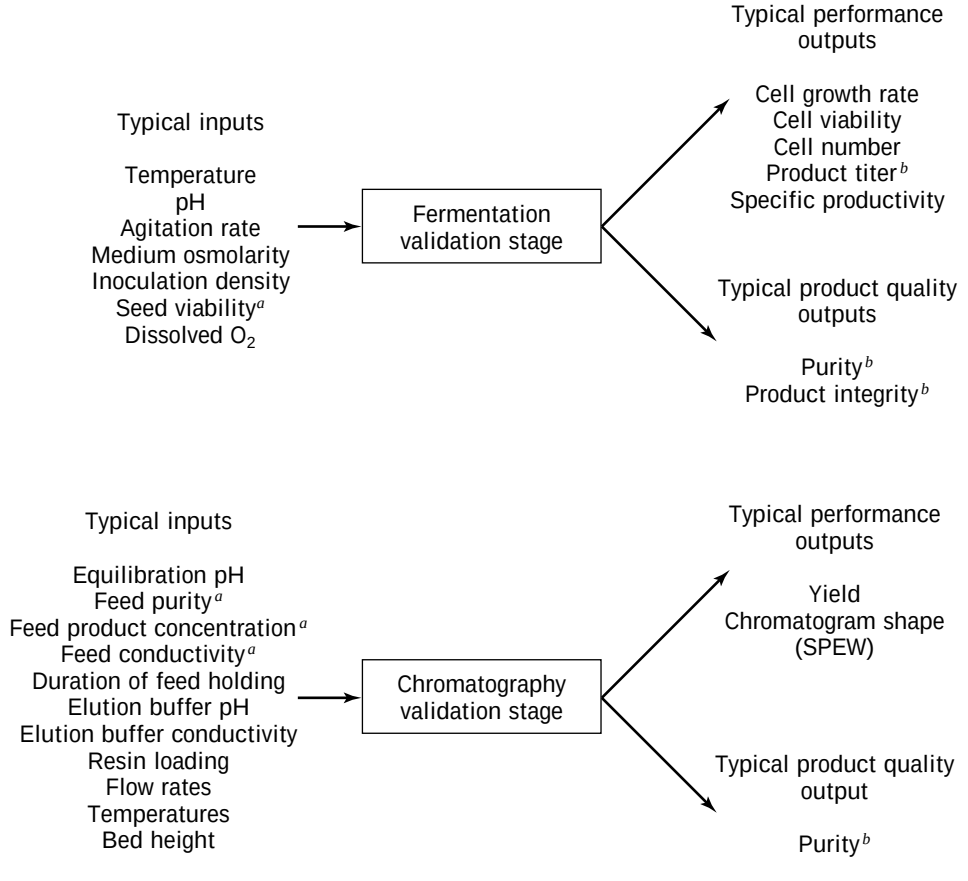


Figure 2. Each validation stage is driven by a set of inputs and produces a set of outputs. The outputs may be related to process stage performance or product quality. The general goal is to validate that the outputs are maintained within satisfactory ranges when the inputs are permitted to vary within defined ranges. [a]Forward linkage input–output from previous stage; [b]forward linkage output–input to next stage.

teria are not met, then the input ranges are narrowed and the experiments are repeated until the criteria are met.

This approach is taken for each validation stage. However, it may be necessary to explore interactions among the validation stages. For example, the performance of a purification step or the quality of the product it produces often depends on the quality of the input feed stream. Because the input feed stream is an output from the previous purification step, one way to handle these interactions is to include the input stream properties as other inputs to the validation stage. Input streams of high and low quality are then required. One approach to obtain these streams is to operate the upstream purification step under best- and worst-case conditions. Alternatively, by knowing the function of the purification step to be validated (e.g., endotoxin removal, host cell protein removal, DNA removal, or removal of product isoforms), one can prepare input streams of high and low quality by spiking in the appropriate species to be removed. The pair-wise validation of processing steps has been referred to as forward linkage validation. *Worst case* and *most appropriate challenge* are terms that describe the validation of a process run with its critical operating parameters set at the limits of their operating ranges (11).

PARs are generally used to establish maximum operating ranges (MORs) bounded by action limits for the important process variables. When a process variable exceeds the outer limits of an MOR, then the impact on the process and the product may be unknown. This event often requires corrective action to be taken by operations staff and various QA systems. Depending on the importance of the variable, the duration and extent of the excursion beyond the MOR, and other operating data from process operations, a variety of follow-up actions may be appropriate. These actions can include relatively minor responses such as requiring a more detailed batch record review or more analytical testing than is normally required for product release when appropriate analytical systems and databases are available. Other QA responses include launching an investigation to look for the cause of the problem, requiring the demonstration of lot stability before release, requiring reprocessing of the lot, or rejecting the lot.

VALIDATION OF PROCESS CONSISTENCY

From the standpoint of process control, it is necessary to determine the extent to which the process and product vary naturally under normal operating conditions. This knowledge is important in defining the normal operating range (NOR) bounded by alert limits for each important process variable. When a NOR is exceeded, operations staff may have to take corrective action. NORs must fall within the MORs established in most cases by the PAR studies. When NORs exceed MORs, the process is not sufficiently controlled to provide "a high degree of assurance that a specific process will consistently produce a product meeting its pre-determined specifications and quality attributes" (1). NORs are typically calculated from replicate runs of the process that produces product that meets specifications. The calculation of a 99% or 95% confidence interval is a typical means of establishing NORs.

The relationship among ranges for any process variable is shown in Figure 3. Process development or optimization often establishes a desired value known as a target or setpoint (e.g., 37 °C) for a process variable. Under normal operating conditions, however, a process variable varies within an NOR (e.g., 37 ± 0.5 °C). Validation studies can determine a PAR (e.g., 37 ± 2.0 °C), which can be used to set the MOR (e.g., 37 ± 2.0 °C). The MOR falls within a broader range bounded by the edge of failure (e.g., 32–40 °C). A process variable that exceeds the edges of failure will cause the process or product to fail, but the edges of failure are not typically determined during process development or validation except by accident. The ranges over which the variable can be controlled (e.g., 20–80 °C) and for which it can exist (e.g., ≥ −273 °C) are still broader (9).

POSTLAUNCH PROCESS VALIDATION

Process validation is a requirement for product licensing. However, it continues throughout the lifetime of the product. Experience with a process in the commercial manufacturing facility is often very limited when the facility files license applications. Experience may mean production of as few as three or five lots to meet the qualification lot requirements for filing in the United States or Europe, respectively. These numbers of lots are insufficient to generate meaningful process or product specifications. As a consequence, initial process validation data are generally derived from pilot plant experiences obtained during the production of supplies for clinical trials or the development of the process. Regulatory authorities accept this information as sufficient for licensing, but many of the MORs, NORs, and product specifications may be refined during actual commercial production experience. In most cases, changes to these ranges are reported to regulatory authorities in annual updates, particularly when the change re-

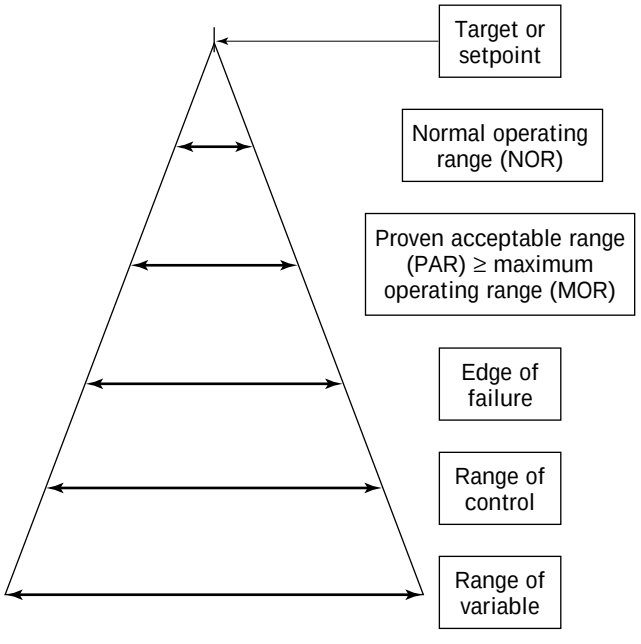


Figure 3. Relationships among ranges for process variables.

sults in a narrowing of a control range or an improvement in product quality. In other cases, however, the change may need to be proposed to the FDA and regulatory approval gained before implementation.

FUTURE PROSPECTS

The significance of process validation has grown from its origins in the mid 1970s to the present time. Of major importance to the biopharmaceutical industry has been the total absence of medical incidents traceable to process control problems, largely because of the application of validation principles. More recently, regulatory authorities have reconsidered their former positions that most changes to the manufacturing process, including relatively minor ones, led to a product that might be different. In the past, clinical studies were required to establish the safety and efficacy of the new product. Today, many changes can be justified using analytical data and process validation studies without the need to undertake a clinical trial. More recently, the FDA has dropped the requirement for filing an establishment license application (ELA) for "therapeutic recombinant DNA-derived products and/or monoclonal antibodies" (20). The FDA has also dropped the requirement for the FDA Center for Biologics Evaluation and Review (CBER) to release test each lot of product with this classification. These changes are due largely to the increase in powerful analytical methods that are available to test these types of products and the greater sophistication of validation to characterize the process.

Because of the growing importance of process validation and the absence of all but the most general guidance from regulatory agencies, it is likely that process validation will become the subject of future regulatory focus by the FDA and the European Union. Most cGMP regulatory requirements are specified for final product, and efforts in the United States to define cGMPs specifically for the production of bulk pharmaceutical chemicals (also known as APIs) are in progress (19). It is in these guidelines that we can expect more practical directives on process validation to be detailed. Given the expense of process validation work, guidance that leads to improved definition of requirements and elimination of process validation studies of marginal significance will be welcome. A reexamination of the balance among prospective, concurrent, and retrospective validation approaches will also be useful. Industry will need to contribute to these debates to ensure that validation costs and their impact on the overall health care system cost are directed toward areas with the greatest impact on the safety and efficacy of biopharmaceutical products.

ACKNOWLEDGMENTS

We thank Tony Lubiniecki, Peter Calcott, and John Bennan for their thoughtful reviews of this manuscript.

BIBLIOGRAPHY

1. U.S. Food and Drug Administration Center for Drugs and Biologics and Center for Devices and Radiological Health, *Guideline on General Principles of Process Validation*, May 1987.
2. K.G. Chapman, *Pharmaceutical Technology* **15**, 82–96 (1991).
3. Current Good Manufacturing Practice in Manufacturing, Processing, Packing, or Holding of Drugs, *Code of Federal Regulations*, Part 210, Title 21, 1998; Current Good Manufacturing Practice for Finished Pharmaceuticals, *Code of Federal Regulations*, Part 211, Title 21, 1998.
4. A.L. Lubiniecki, M.E. Wiebe, and S.E. Builder, in A.S. Lubiniecki ed., *Large-Scale Mammalian Cell Culture Technology*, Marcel Dekker, New York, 1990, pp. 515–541.
5. Biotech Product Validation, *BioPharm Validation Supplement* (1995).
6. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), *Viral Safety Evaluation of Biotechnology Products Derived from Cell Lines of Human and Animal Origin*, March 5, 1997.
7. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), *Quality of Biotechnological Products: Analysis of the Expression Construct in Cells Used for Production of r-DNA Derived Protein Products*, November 1995.
8. PMA's Validation Advisory Subcommittee, *Pharmaceutical Technology* **9**, 78–82 (1995).
9. K.G. Chapman, *Pharmaceutical Technology* **8**, 22–36 (1984).
10. M.E. Wiebe and S.E. Builder, in F. Brown and A.L. Lubiniecki eds., *Genetic Stability and Recombinant Product Consistency*, Karger, Basel, 1994, pp. 45–54.
11. Parenteral Drug Association, *Journal of Parenteral Science and Technology* **46**, 87 (1992).
12. R.G. Werner, H. Langlouis-Gau, F. Walz, H. Allgaier, and H. Hoffmann, *Arzneim.-Forsch.* **38**, 855–862 (1988).
13. L.D. Torbeck and R.C. Branning, *Pharmaceutical Technology* **20**, 108 (1996).
14. Parenteral Drug Association Biotechnology Task Force on Purification and Scale-Up, *Journal of Parenteral Science and Technology* **46**, s3–s13 (1992).
15. U.S. Food and Drug Administration, Center for Biologics Evaluation and Research (CBER), *Points to Consider in the Characterization of Cell Lines Used to Produce Biologics*, 1993.
16. U.S. Food and Drug Administration, Center for Biologics Evaluation and Research (CBER), *Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products for Human Use*, April 28, 1997.
17. CPMP Biotechnology Working Party, *EU—Note for Guidance on Virus Validation Studies: The Design, Contribution, and Interpretation of Studies Validating the Inactivation and Removal of Viruses*, 1996.
18. Current Good Manufacturing Practice: Proposed Amendment of Certain Requirements for Finished Pharmaceuticals, *Fed. Regist.*, 61(87), 20104 (1996).
19. U.S. Food and Drug Administration, *Guidance for Industry: Manufacturing, Processing, or Holding Active Pharmaceutical Ingredients*, Discussion Draft, March 1998.
20. U.S. Food and Drug Administration, Center for Biologics Evaluation and Research (CBER) and Center for Drug Evaluation and Research (CDER), *Guidance for Industry for the Submission of Chemistry, Manufacturing, and Controls Information for a Therapeutic Recombinant DNA-Derived Product or a Monoclonal Antibody Product for In-Vivo Use*, August 1996.
21. Biological Products: General; Physical Establishment, Equipment, Animals, and Care, *Code of Federal Regulations*, Part 600.11a, Title 21, 1998.

22. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), *Guideline on Quality of Biotechnology Products: Viral Safety Evaluation of Biotechnology Products Derived from Cell Lines of Human or Animal Origin*, December 1, 1995.
23. G. Sofer, *BioPharm* **9**, 51–54 (1996).
24. A. Riggen, G.C. Davis, and T.L. Copmann, *BioPharm* **9**, 36–41 (1996).
25. A.J. Darling and J.J. Spaltro, *BioPharm* **9**, 42–50 (1996).
26. K.P. Anderson, M.-A. L. Low, Y.S. Lie, G.-A. Keller, and M. Dinowitz, *Virology* **181**, 305–311 (1991).
27. R.J. Seely, H.D. Wight, H.F. Fry, S.R. Rudge, and G.F. Slaff, *BioPharm* **7**, 41–48 (1994).

See also ASTM STANDARDS FOR BIOTECHNOLOGY; CLEANING, CLEANING VALIDATION; GOOD MANUFACTURING PRACTICE (GMP) AND GOOD INDUSTRIAL LARGE SCALE PRACTICE (GLSP); STERILIZATION-IN-PLACE.

PRODUCTION OF L-AMINO ACIDS BY AMINOACYLASE

TADASHI SATO
TETSUYA TOSA
Tanabe Seiyaku Company, Ltd.
Osaka, Japan

KEY WORDS

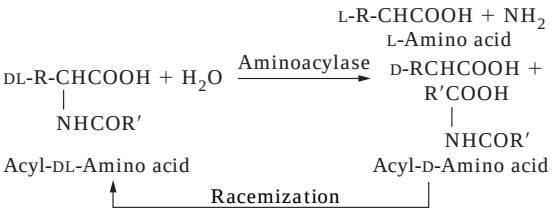
Aminoacylase
L-Amino acid
Optical resolution
Racemization

OUTLINE

Bibliography

At present, fermentative and chemical synthetic methods are used for the industrial production of L-amino acids instead of conventional isolation from protein hydrolysates. However, chemically synthesized amino acids are optically inactive racemic mixtures of L- and D-isomers. The L form is the physiologically active natural form and is therefore the required form for medicine and food. To obtain L-amino acid from the chemically synthesized DL form, optical resolution is necessary.

Generally, optical resolution of racemic amino acids can be carried out by physicochemical, chemical, enzymatic, and biological methods. Among these methods, an enzymatic method developed by Chibata et al. (1) using mold aminoacylase is one of the most advantageous procedures, yielding optically pure L-amino acids. The reaction catalyzed by the enzyme is as follows.



Chemically synthesized acyl-DL-amino acid is asymmetrically hydrolyzed by aminoacylase to give L-amino acid and the unhydrolyzed acyl D-amino acid. After concentration, both materials are easily separated on the basis of the difference in their solubilities. Acyl-D-amino acid is racemized and reused for the resolution procedure.

Chibata et al. (1) studied enzymes catalyzing the reaction and found that aminoacylase produced by *Aspergillus oryzae* has high activity and broad substrate specificity, that is, it can asymmetrically hydrolyze many kinds of acyl-DL-amino acids.

From 1953 to 1969, this mold aminoacylase was used for the industrial production of several L-amino acids. This method was one of the most advantageous procedures for the industrial production of L-amino acids. However, because the enzyme reaction was carried out batchwise by incubating a mixture containing the substrate and soluble enzyme, the procedure had some disadvantages for industrial purposes. For instance, to isolate L-amino acid from the enzyme reaction mixture, it was necessary to remove enzyme protein by pH and/or heat treatment. Thus, even if enzyme activity remained in the reaction mixture, the enzyme had to be discarded because there was no suitable procedure for isolating the active enzyme from the mixture. In addition, a complicated purification procedure was necessary for the removal of proteins and coloring materials contaminating the crude enzyme preparations usually used for industrial purposes.

To overcome these disadvantages and to improve this enzymatic method, we extensively studied the continuous optical resolution of DL-amino acids using a column packed with immobilized aminoacylase (2). An efficient and automatically controlled enzyme reactor system was achieved, and since 1969, we have been industrially operating a series of these reactors in our plants.

Various immobilization methods were investigated, and relatively active and stable immobilized aminoacylases were obtained by ionic binding to DEAE-Sephadex (3), covalent binding to iodoacetylcellulose (4), and entrapping into polyacrylamide gel (5). However, for the industrial operation of a continuous enzyme reaction using an immobilized enzyme, it was necessary to satisfy many conditions. Thus, the characteristics of these three immobilized aminoacylases were compared, and aminoacylase bound to DEAE-Sephadex was chosen as the most advantageous enzyme preparation for the industrial production of L-amino acids, because (*1*) preparation is easy, (*2*) the activity is high, (*3*) the stability is high, and (*4*) regeneration of deteriorated immobilized enzyme is possible.

To design the most efficient enzyme column, the following are important factors: (*1*) the effect of flow rate of substrate on the reaction rate, (*2*) the flow of the substrate solution, (*3*) the effect of column dimension on the reaction rate, and (*4*) the pressure drop in the column.

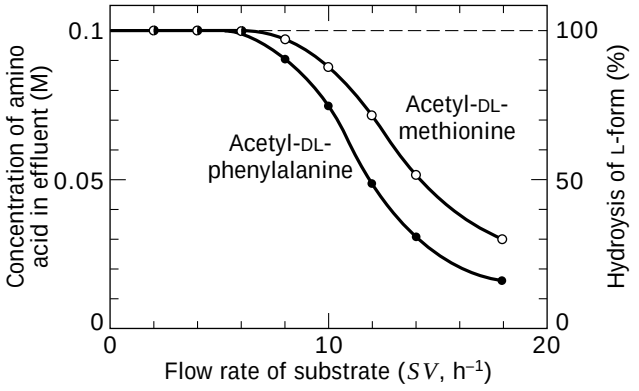


Figure 1. Continuous hydrolysis of acetyl-DL-amino acids by an immobilized aminoacylase column. Relationship between flow rate of substrate and extent of hydrolysis. Aminoacylase was immobilized by ionic binding to DEAE-Sephadex. Concentration of substrate, 0.2 M (containing 5×10^{-4} M Co^{2+}); reaction temperature, 50 °C.

Because it is one of the factors considered in the design of an enzyme reactor, the relationship between the flow rate of a substrate solution passing through the enzyme column and the extent of the reaction must be investigated in advance. We investigated the relationship between the flow rate of a solution of acetyl-DL-methionine or acetyl-DL-phenylalanine and the extent of the reaction, and the results are shown in Figure 1 (2). An aqueous solution of acetyl-DL-amino acid (0.2M, pH 7.0 to 7.5, containing 5×10^{-4} M Co^{2+}) was passed through a DEAE-Sephadex-aminoacylase column at various flow rates at 50 °C. In Figure 1, flow rate is expressed in terms of space velocity. The space velocity is the volume of liquid passing through a given volume of immobilized enzyme in 1 h divided by the latter volume. If the space velocity is reduced, that is, the flow rate is slowed, the reaction proceeds more effectively and the relationship between the flow rate and the extent of the reaction is sigmoidal. As shown in Figure 1, when the concentration of L-amino acid in the effluent reached 0.1 M, the reaction stopped. This result indicates that the L-form of 0.2 M acetyl-DL-amino acid was completely hydrolyzed and its D-form was not hydrolyzed, that is, asymmetric hydrolysis occurred.

The pressure drop of the enzyme column is also important factor in the design of an enzyme reactor. Thus, the relationship between the flow rate of substrate solution and the pressure drop of the column was investigated using columns of various lengths. The pressure drop of the DEAE-Sephadex-aminoacylase column was found to be proportional to the flow rate of substrate solution and the column length.

Table 1. Production of L-Amino Acids on 1,000-L DEAE-Sephadex Aminoacylase Column

		Yield (theory) of L-amino acids per	
L-Amino acid	Space velocity	24 h (kg)	20 days (kg)
L-Alanine	1.0	214	6,420
L-Methionine	2.0	715	21,450
L-Phenylalanine	1.5	594	17,820
L-Tryptophan	0.9	441	13,230
L-Valine	1.8	505	15,150

In addition, by considering some other properties of the immobilized enzyme, such as optimum pH, the effect of temperatures on the reaction rate and stability, and so on, we designed an enzyme reactor system for continuous production of L-amino acids, as shown in Figure 2. In this system, the flow rate and pH of substrate solution and the operating temperature can be automatically controlled and recorded.

Since 1969, continuous optical resolution of acetyl-DL-amino acids has been carried out using this system, and several kinds of optically active amino acids such as methionine, phenylalanine, valine, and others have been industrially produced at Tanabe Seiyaku Company, Japan. Several examples of the production of L-amino acids are summarized in Table 1, which shows the space velocity and the theoretical yield for each amino acid produced in a 1,000-L aminoacylase column. That was the first industrial application of immobilized enzymes.

The immobilized aminoacylase was very stable, as shown in Figure 3, and maintained 60 to 70% of its initial activity after continuous operation for 30 days at 50 °C. The enzyme column can be completely regenerated after prolonged operation by the addition of an amount of aminoacylase corresponding to the lost activity. The carrier, DEAE-Sephadex, is also very stable and has been used for more than 5 years in our column process without any change in adsorption capacity for the enzyme, shape, or pressure drop. A continuous enzyme reaction using such a stable immobilized enzyme gives not only high productivity per enzyme unit, but also high product yield, because contamination of the effluent by such impurities as proteins and coloring substances does not occur, so the product can be obtained by simple purification and the amount of substrate required is reduced. Further, automatic operation reduces labor costs, and a significant reduction in production cost can be expected in comparison with the conventional batch process using soluble enzyme. A comparison of the production costs of L-amino acids by the conventional batch process using soluble enzyme and by the continuous process using immobilized enzyme in our plant is shown in Figure 4. In the immobilized enzyme process, the overall production cost is more than 40% lower than that of the conventional batch process using soluble enzyme. Savings of enzyme and labor costs are the main contributors, as are the increase in product yield from the easy isolation of L-amino acids from the reaction mixture. Although DEAE-Sephadex is a relatively expensive carrier, production costs are not greatly affected by the cost of the carrier, because it can be used for a very long period, as mentioned.

Besides immobilization by ionic binding, mold aminoacylase from *Aspergillus* sp. was also immobilized by covalent binding to alkylaminosilanized porous glass with glutaraldehyde or to the diazonium derivative of acylaminosilanized porous glass, and these immobilized aminoacylases were used for continuous preparation of L-amino acids from acetyl-DL-amino acids (6).

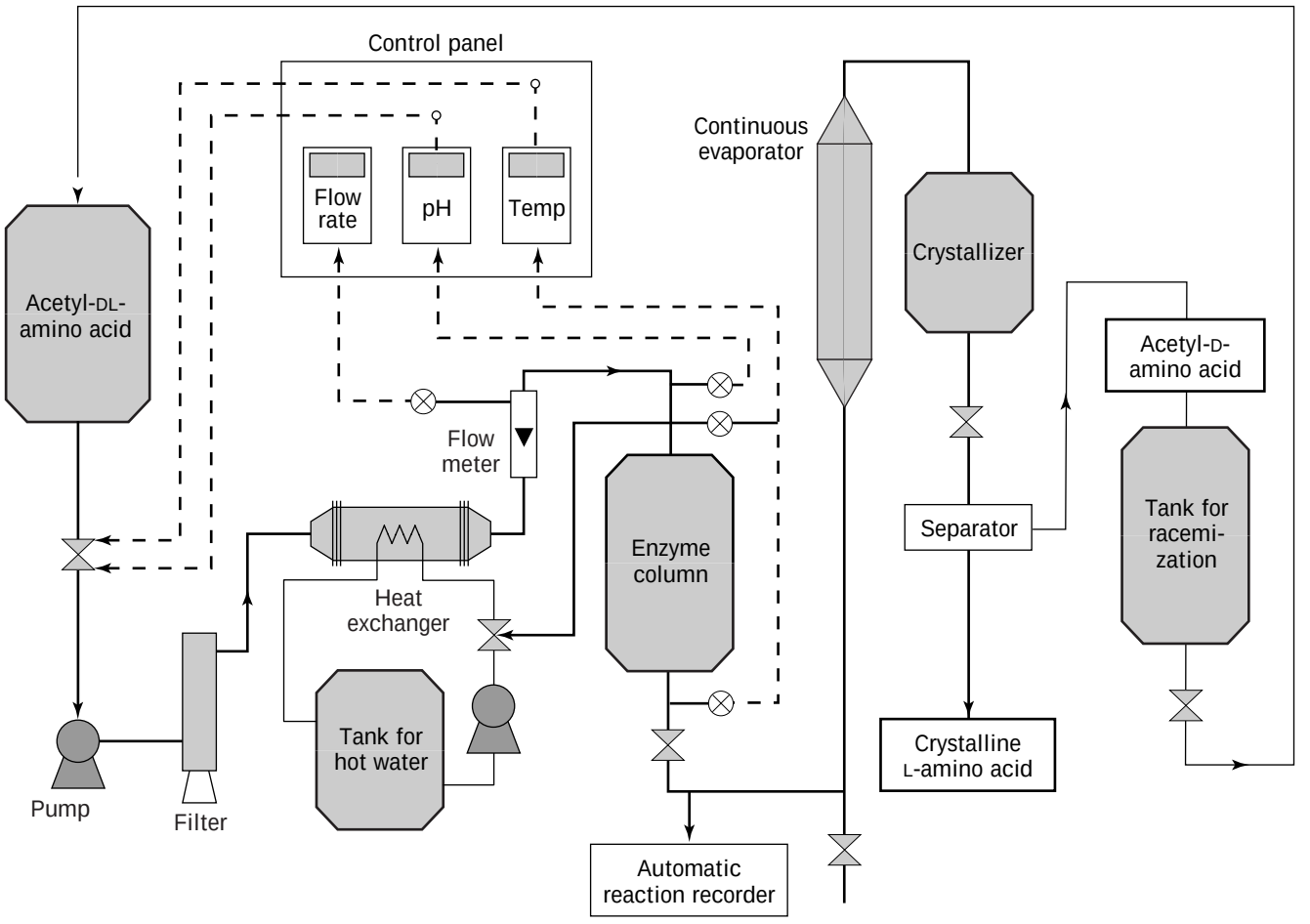


Figure 2. Flow diagram for the continuous production of L-amino acids by immobilized aminoacylase.

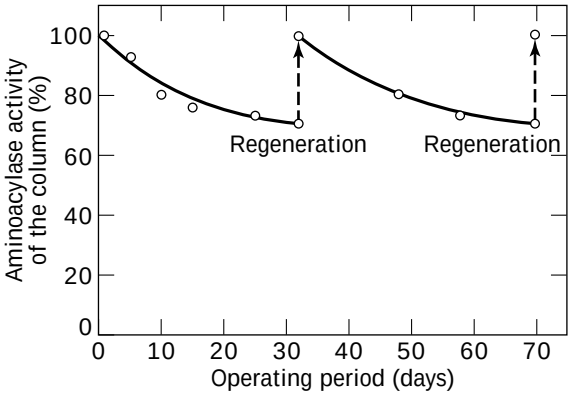


Figure 3. Stability and regeneration of an immobilized aminoacylase column. Aminoacylase was immobilized by ionic binding to DEAE-Sephadex, and 0.2 M acetyl-DL-methionine (containing 5×10^{-4} M Co^{2+}) was continuously passed through a column packed with the immobilized aminoacylase at a flow rate of SV = 2 at 50 °C.

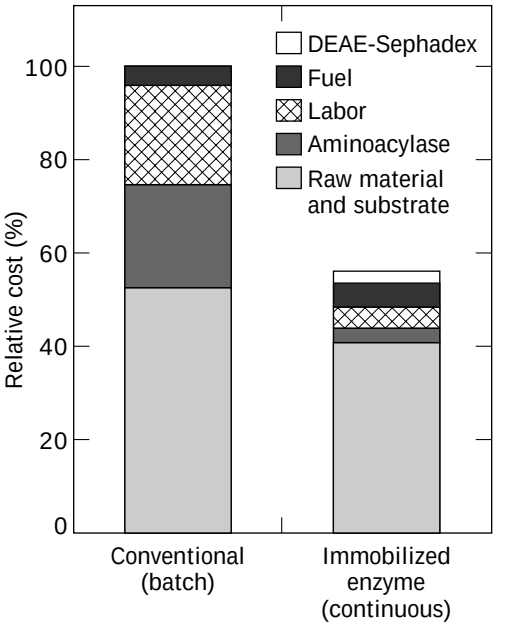


Figure 4. Comparison of production costs of L-amino acids by batch and continuous processes.

Bacterial and animal aminoacylases have been also immobilized. Bacterial aminoacylase has been covalently bound to diazotized polyaminostyrene (7,8). The aminoacylase from pig kidney was immobilized by ionic binding to DEAE-cellulose (9) and by covalent binding to the azide derivative of Enzacryl AH (10) or diazotized Enzacryl AA (11). However, these preparations were used batchwise and on only a laboratory scale for the preparation of L-amino acids from acyl-DL-amino acids.

BIBLIOGRAPHY

1. I. Chibata, T. Ishikawa, and S. Yamada, *Bull. Agric. Chem. Soc. Jpn.* **21**, 300–303 (1957).
2. I. Chibata, T. Tosa, T. Sato, T. Mori, and Y. Matuo, *Proceedings of the 4th International Fermentation Symposium: Fermentation Technology Today*, 1972, p. 383–389.
3. T. Tosa, T. Mori, N. Fuse, and I. Chibata, *Biotechnol. Bioeng.* **9**, 603–615 (1967).
4. T. Sato, T. Mori, T. Tosa, and I. Chibata, *Arch. Biochem. Biophys.* **147**, 788–796 (1971).
5. T. Mori, T. Sato, T. Tosa, and I. Chibata, *Enzymologia* **43**, 213–226 (1972).
6. H.H. Weetall and C.S. Detar, *Biotechnol. Bioeng.* **16**, 1537–1544 (1974).
7. M.A. Mitz and L.J. Summaria, *Nature* **189**, 576–577 (1961).
8. M.D. Lilly, C. Money, W.E. Hornby, and E.M. Crook, *Biochem. J.* **95**, 45p (1965).
9. T. Barth and H. Maskova, *Collect. Czech. Chem. Commun.* **36**, 2398–2402 (1971).
10. Y. Ohno and M. Stahmann, *Macromolecules* **4**, 350–352 (1971).
11. H. Maskova, T. Barth, B. Jirovsky, and I. Rychlik, *Collect. Czech. Chem. Commun.* **38**, 943–947 (1973).

See also AMINO ACIDS, PRODUCTION PROCESSES; BIOCATALYSIS DATABASES; FOOD PROCESS ENGINEERING.

PRODUCTION OF L-GLUTAMIC ACID

SHUKUO KINUSHITA
Kyowa Hakko Kogyo Co., Ltd.
Tokyo, Japan

KEY WORDS

Amino acid fermentation
Biotin
Corynebacterium
Fermentation process
Food taste enhancer
Glutamic acid
Penicillin

OUTLINE

INTRODUCTION

L-Glutamic acid monosodium salt was first discovered by Dr. Kikunae Ikeda as a taste component of *konbu*, a sea tangle, which is a food taste-enhancing material traditionally used in Japan. Based on this discovery, L-glutamic acid monosodium salt was industrially produced from wheat gluten by hydrolysis with hydrochloric acid (1). Later, the raw material was replaced by soybean protein, and a synthetic method was studied and finally industrialized. In addition attempts were made to produce α-ketoglutaric acid, which is a precursor of L-glutamic acid, by organic acid fermentation. These attempts were aimed at producing L-glutamic acid by transamination from the resulting α-ketoglutaric acid.

In 1957, Kinoshita et al. succeeded in establishing a process to produce L-glutamic acid by direct fermentation from a carbohydrate and ammonia using a microorganism, *Corynebacterium glutamicum*, which they had isolated from nature (2). Compared to the then conventional hydrolytic process comprising the hydrolysis of natural protein, this direct fermentation process made it easy to secure the production raw material, and at the same time facilitated a drastic reduction of cost.

With establishment of the fermentative process, the monosodium glutamate production industry showed rapid growth. In recent years, the production worldwide has been estimated to reach to about 600,000 tons per year as L-glutamic acid (Table 1). As the production cost was cut

Table 1. Production of Monosodium Glutamate

Years	Production of MSG (t/year)
1954	7,652
1960	22,177
1970	92,935
1988	340,000
1994	592,000

Note: MSG, monosodium glutamate.

owing to improvement in production technologies, the production plants, initially operated only in Japan, have been extended worldwide, particularly to South East Asia where there are abundant supplies of raw materials and also consumption of the product (Table 2).

SYNTHETIC PATHWAY OF L-GLUTAMIC ACID

In the biosynthesis of L-glutamic acid (Fig. 1), a key step is a reductive ammonia-addition reaction of α-ketoglutaric acid, which constitutes the TCA cycle:

$$\alpha\text{-ketoglutaric acid} + \text{NADPH} + \text{H}^+ + \text{NH}_4 \xrightarrow[\text{glutamate dehydrogenase}]{} \text{L-glutamic acid} + \text{NADPH}^+ + \text{H}_2\text{O}$$

NADPH, the coenzyme of this reaction, is supplied through coupling in the reaction to form α-ketoglutaric acid:

$$\text{isocitric acid} + \text{NADPH}^+ \xrightarrow[\text{isocitrate dehydrogenase}]{} \alpha\text{-ketoglutaric acid} + \text{CO}_2 + \text{NADPH} + \text{H}^+$$

L-GLUTAMIC ACID FERMENTATION

Initially, a glutamic acid–producing bacterium was isolated as a biotin-requiring mutant of *C. glutamicum*, and a fermentative production process in which the strain is cultured in a medium containing a limited concentration of biotin was industrialized. Afterward, use of cheap cane molasses attracted attention as a raw material carbohydrate; however, cane molasses contains a large amount of biotin. To enable use of cane molasses, the penicillin shot method was developed, which made it possible to accumulate L-glutamic acid in the presence of excess biotin, and the industrial technology had been further advanced. Subsequently, techniques to use oleic acid–requiring strains or to add saturated higher fatty acid esters (or saturated higher fatty acid) such as polyoxyethyleneglycol parmityl ester were developed. Also, use of various carbon source materials other than carbohydrate was examined. L-Glutamic acid fermentation from *n*-paraffin (3,4) has been much studied, and the production process has become almost as well established as the use of carbohydrate. However, the production from *n*-paraffin has never been commercialized because of the high cost of the main raw material, as well as other reasons. Currently, the penicillin shot method using cane molasses, beet molasses, glucose, and so forth as raw materials is most frequently employed in industry because of cost, stable production, and other reasons.

Table 2. Worldwide Capacity for Monosodium Glutamate, 1994

Area	Capacity (estimated) (t/year)
U.S.	60,000
Japan	279,000
Far East (except Japan)	392,700
Europe	72,500
South and Central America	53,000
Total	857,200

METABOLIC MECHANISMS OF L-GLUTAMIC ACID PRODUCTION

Change of the Membrane-Permeation System

As already mentioned, the L-glutamic acid production process was initially established using a biotin-requiring bacterial strain and strict control of biotin concentration. It was later found that production could be accomplished even under excess biotin conditions by adding penicillin, using oleic acid-requiring microorganisms, and adding saturated higher fatty acid esters such as polyoxyethyleneglycol parmityl ester or fatty acid itself. It was also found that, under high oleic acid concentration, although L-glutamic acid was not produced by the addition of higher fatty acids or their esters, the penicillin shot method was still effective. Today, these phenomena are explained as attributable to dependency of L-glutamic acid accumulation on the changes of the membrane-permeation system (Fig. 2). More particularly, biotin plays a role in the biosynthesis of oleic acid by glutamic acid–producing bacteria, and under the biotin-controlled condition, oleic acid formation is repressed. This causes a change in fatty acid composition and phospholipids that constitute the cell membrane and, consequently, a change in permeability of L-glutamic acid through the membrane. Oleic acid–requiring mutants and the direct addition of saturated higher fatty acid bring about a similar change in composition whether or not biotin is present, and produce a similar effect. Hence these conditions do not give rise to L-glutamic acid accumulation at a high concentration of oleic acid. On the other hand, it is believed that the addition of penicillin inhibits synthesis of cell walls rather than changes in the composition of fatty acid that constitutes the membrane, thereby causing a difference in the osmotic pressure between the outside and the inside of the membrane. It is this pressure differential that causes the membrane to undergo a change that leads to L-glutamic acid accumulation.

When *n*-paraffin is used as a raw material, the mechanism of L-glutamic acid production is explained by a similar change of membrane permeability, although a threonine-requiring strain is used in place of the biotin-requiring strain used in carbohydrate fermentation; and the presence of a copper ion additionally affects L-glutamic acid formation (4).

Other Factors for Excess Formation of L-Glutamic Acid

There has been a report on the change in the activity of the L-glutamic acid–producing enzyme system in its L-glutamic acid formation phase (6). Furthermore, in some cases, L-glutamic acid–producing strains with an enhanced biosynthetic enzyme system have been obtained by recombinant technology. Thus, L-glutamic acid production may not be totally explained by "leakage" of the amino acid. This may be a problem to be studied in the future.

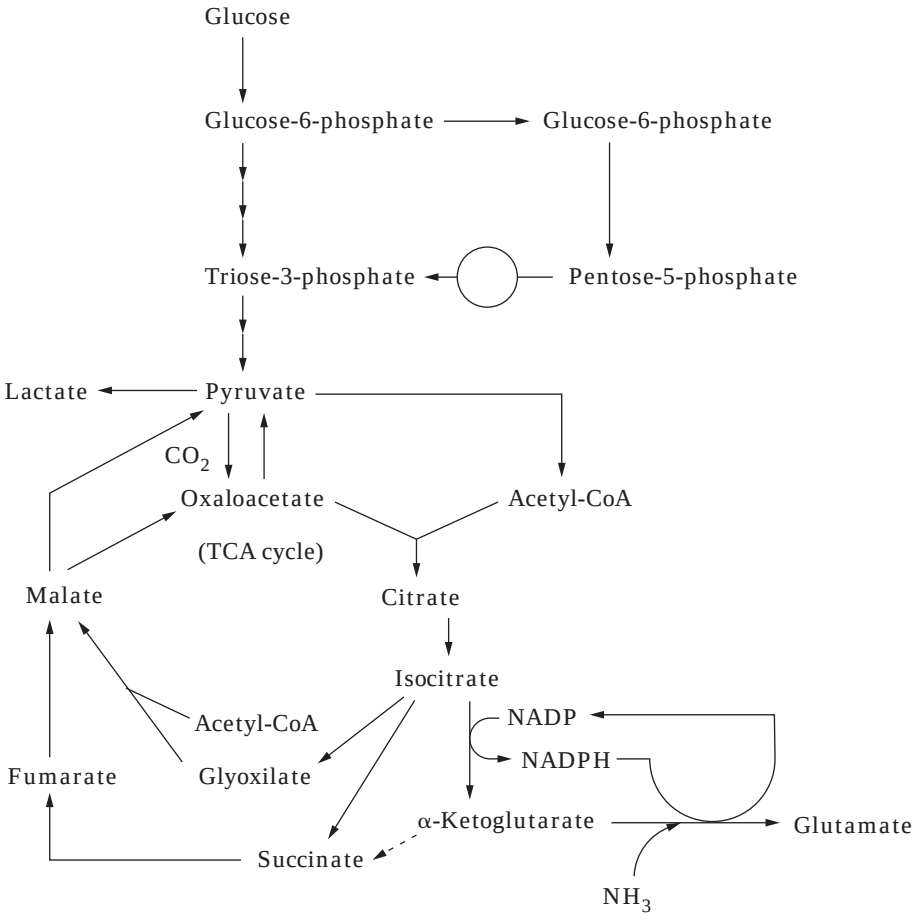


Figure 1. The biosynthetic pathway of glutamate.

In a recent report (7) a gene for detergent-sensitivity (*dtsR*) carried by a wild-type strain was identified using a polyoxyethylene sorbitan fatty acid ester–sensitive strain of *Brevibacterium lactofermentum*, and a mutant lacking this gene was found to accumulate L-glutamic acid in a medium containing excess biotin and in the presence of polyoxyethylene sorbitan oleate ester. Under these conditions the wild-type strain did not accumulate the amino acid. This strain had its specific activity of 2-oxoglutarate dehydrogenase complex (an enzyme involved in the TCA cycle that converts α-ketoglutaric acid to succinyl-CoA) reduced to one-third of that possessed by the wild-type strain. From these findings, Kimura et al. (8) proposed a mechanism of L-glutamic acid fermentation: "The target of controlling biotin concentration and addition of surfactants is biotin enzyme complex containing DtsR, and the reduction of the activity of this biotin enzyme complex may possibly cause glutamic acid formation." According to this report, the sensitive strain propagates in response to the concentrations of oleic acid and polyoxyethylene sorbitan oleate ester in the medium. Protein, a product of the DtsR gene, had homology with a biotin enzyme derived from other organisms and was involved in the biosynthesis of higher fatty acid. Furthermore, accumulation of L-glutamic acid has been reported by a mutant strain having membrane permeability to L-glutamic acid. The mutation was induced from a wild-type methanol-assimilating strain that does not accumulate L-glutamic acid in the medium (9).

These findings suggest some ideas for the improvement of techniques to construct L-glutamic acid–producing strains by the recombinant technology. In the future, more efficient fermentation production of L-glutamic acid may be possible through diverse analyses of production processes and mechanisms, and construction of mutant strains.

INDUSTRIAL PRODUCTION OF MONOSODIUM GLUTAMATE

Fermentation Processes

The typical L-glutamic acid–producing wild strain of *C. glutamicum* isolated from nature requires biotin for growth and accumulates L-glutamic acid in the medium when cultured under conditions that include sufficient ammonium ions and controlled biotin concentration. This glutamic acid–producing strain shows different forms of fermentation, depending upon fermentation conditions other than biotin concentration. Basic fermentation conditions and the forms of fermentation corresponding to them are shown in Table 3. Glutamic acid production requires sufficient oxygen, appropriate amounts of phosphoric acid, and a nitrogen source, as well as controlled biotin concentration. This strain grows vigorously in a medium containing excess biotin and does not accumulate L-glutamic acid. However, fermentation production under

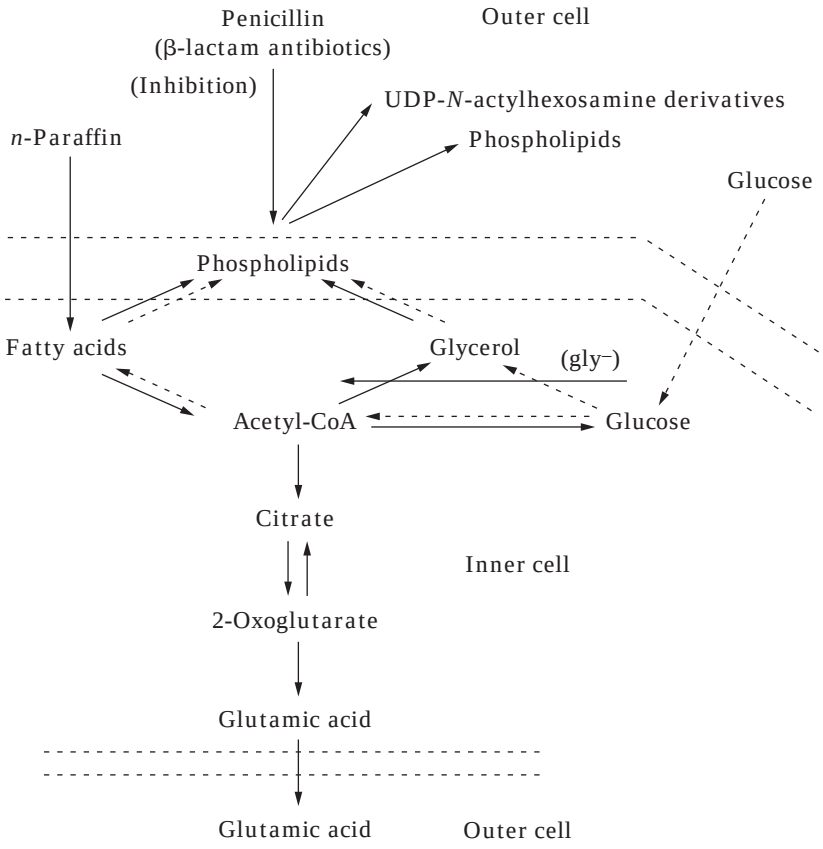


Figure 2. The excretion model of L-glutamic acid in glutamate-producing bacteria. *Source*: From Ref. 5.

Table 3. Coversion of Fermentation by Glutamic Acid–Producing Bacteria with Nutrient and Atmospheric Regulation

Regulatory factors	Coversion of fermentation
Oxygen	Lactate or succinate ↔ Glutamate (anaerobic) (aerobic)
NH_4^+	α-ketoglutarate ↔ Glutamate ↔ Glutamate (unsatisfy) (optimal) (oversatisfy)
pH	*N*-Acetylglutamine ↔ Glutaminate (or Glutamine) (acidic) (basic)
Phosphate	Valine ↔ Glutaminate (high)
Biotin	Lactate or succinate ↔ Glutamate (satisfy) (deficiency)

Source: From Ref. 10.

controlled biotin concentration involved two difficulties: *(1)* available raw materials were restricted (particularly, it was difficult to use cane molasses due to its high biotin content) and *(2)* fermentation production lacked stability because cell growth was not stable. Hence in order to solve this problem, various processes were attempted to accumulate L-glutamic acid under excess biotin conditions. Among these, there are two methods that have become employed as typical methods for industrial production. One is the penicillin shot method in which penicillin is added to cells cultured with sufficient biotin, and the other is the addition of saturated fatty acids or their esters to accumulate L-glutamic acid. The latter method was developed with the elucidation of the physiological effect of the penicillin shot. At present, the penicillin shot method is considered to be advantageous because it is stable in industrial operation. This method is discussed in detail later.

Raw Materials

From the viewpoints of stable supply and low cost, cane molasses (hereinafter referred to simply as "molasses"), beet molasses, and glucose are popular as the main raw material carbon sources. When molasses is used, it is necessary to employ a technique for excess biotin conditions described in the article GLUTAMIC ACID PRODUCING MICROORGANISMS because molasses contains excess biotin. On the other hand, when beet molasses and glucose are used, it is necessary to add biotin to the medium, whether using conditions of excess biotin or limited biotin. The following discusses the penicillin shot method using molasses.

When molasses is used as a main raw material, necessary components such as carbon, nitrogen, phosphoric acid, potassium, sodium, sulfuric acid radical, various metal ions, vitamins, and so on can be derived from this raw material. Any components revealed to be lacking by preliminary tests are supplemented to the medium.

In the following description of the penicillin shot method of L-glutamic acid fermentation, *C. glutamicum* and molasses are used, respectively, as the strain and the raw material, and culturing is carried out by a fed-batch system.

The fermentation tank is a basic aeration/agitation tank, shown in Figure 3. Generally, jacket-type (small fermenter) and coil-type (large fermenter) cooling devices are used for controlling the temperature, but cooling may also be effected by a system that cools the fermentation liquor with an external heat exchanger. In the L-glutamic acid fermentation from carbohydrate, optimization of oxygen supply is an essential condition, and equipment needs be selected to satisfy this condition. A seed medium prepared in a seed tank is transferred to a main fermenter containing a sterilized (using the equipment mentioned later) initial medium, and culturing is initiated after the culturing conditions are adjusted as predetermined by earlier tests. Penicillin G potassium salt is added (8–10 units/L) during the logarithmic growth phase. By adding penicillin, the strain undergoes changes, mainly in membrane permeability, to produce L-glutamic acid, and L-glutamic acid accumulates in the medium. Culturing is then carried out by continuously adding sterilized liquid molasses and controlling the pH around neutrality with ammonia. Liquid molasses is added until the strain no longer produces L-glutamic acid. As shown in Figure 4, which gives one example of a time course of fermentation, generally, the L-glutamic acid production phase continues only 30 or so hours after the addition of penicillin. L-Glutamic acid is finally produced in a concentration of 90–105 g/L. Unlike the culturing process with controlled biotin concentration, it is a characteristic feature of the penicillin shot method that culturing can be carried out under optimum conditions for microbial growth during the early phase of culturing. The advantages of this process are that culturing can be controlled easily, cell concentration is maintained almost at a constant level, and process control (i.e., the rate of molasses addition) is relatively easy. As shown in Table 3, glutamic acid fermentation undergoes different patterns as a result of certain factors. For example, an excessively high rate of carbohydrate addition leads to increased con-

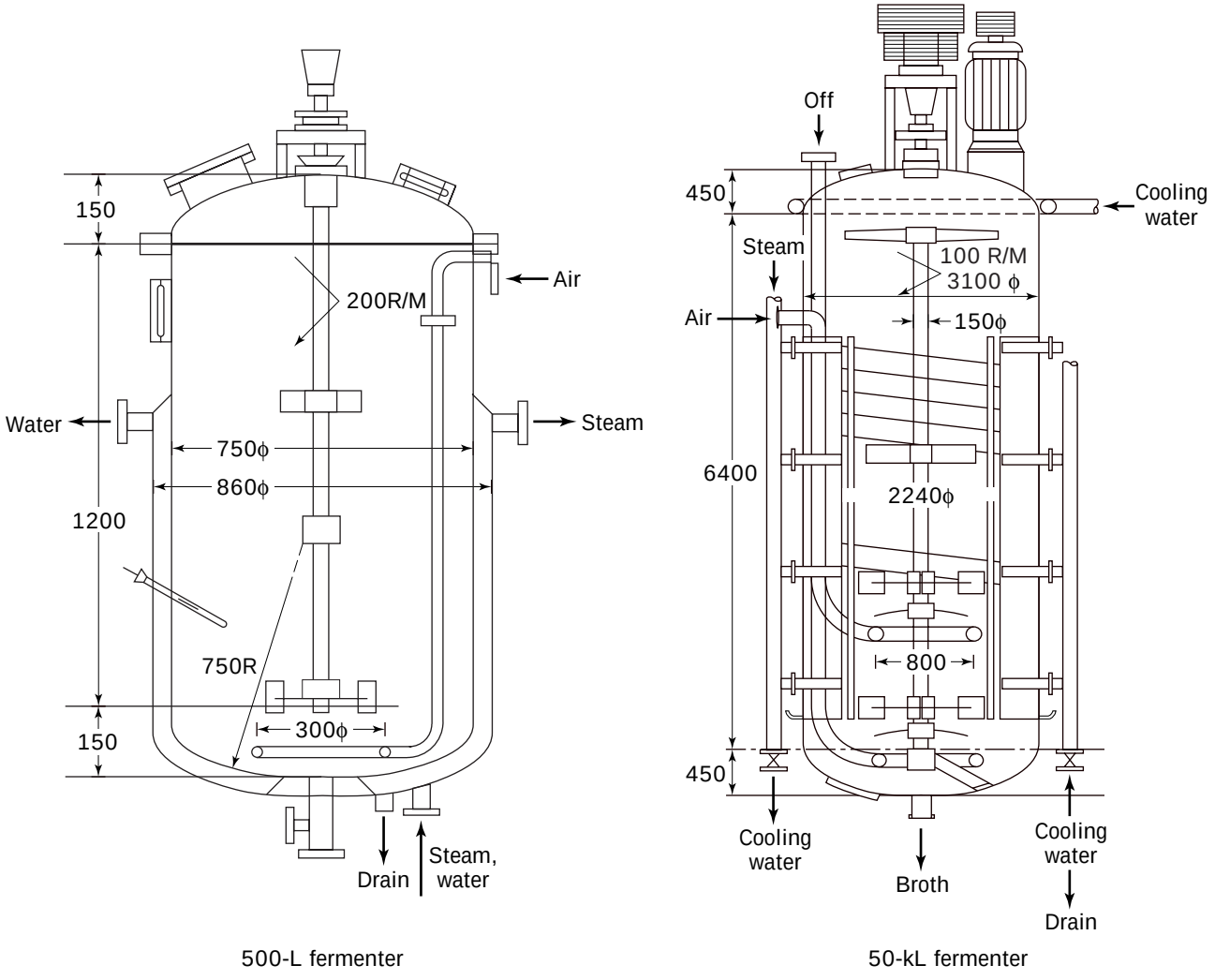


Figure 3. Example of a fermenter. *Source*: From Ref. 11.

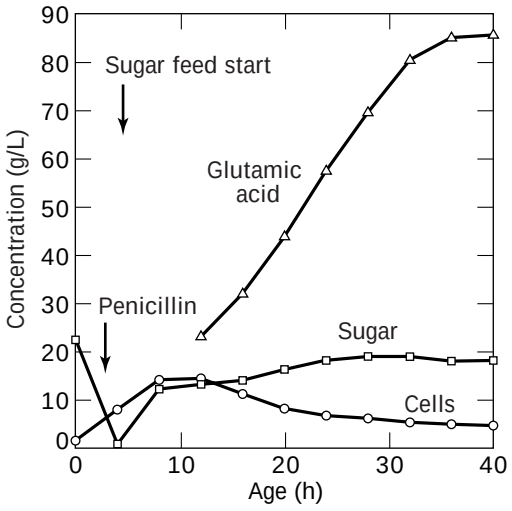


Figure 4. A fed-batch culture producing L-glutamate (*C. glutamicum*).

sumption of oxygen, resulting in oxygen shortage and, consequently, lactic acid and other organic acid formation. The control of cell growth is delicate when accomplished by the control of required substances, and the penicillin shot method is superior in ensuring stable productivity. Furthermore, the penicillin treatment concomitantly reduces the fraction of insoluble cells to a remarkable extent, which makes the recovery of L-glutamic acid from the fermentation broth much easier. This makes it possible to omit the mechanical separation of cells, resulting in simplification of the whole process.

It is, however, a problem that under the conditions currently used, the L-glutamic acid productivity lasts only about 30 h, as already mentioned. Generally, continuous processes are used because they improve productivity, but there is no continuous process that can practically operated due to the limited period of L-glutamic acid formation.

Purification: Direct Concentration and Crystallization

The outline of the purification step is shown in Figure 5. This step consists of two sub steps: recovering the slurry of L-glutamic acid from the fermentation broth (GA step), and producing the final product, monosodium L-glutamate, from L-glutamic acid (MSG step).

The GA step is directed to efficient recovery of high-purity GA crystals from the fermentation broth and the MSG-crystallized mother liquor resulting from the MSG step, and discharges the depleted GA-crystallized mother liquor out of the system. According to the penicillin shot method, in which most of the cells in the fermentation broth are lysed and solubilized, physical separation of the cells, such as with centrifugation, can be excluded, and the cell components are discharged into the GA-crystallized mother liquor. Precipitation is effected at the isoelectric point. GA crystals are known to take either the α- or the β-form. The α-crystal, which can be readily isolated, needs to be selectively precipitated because removal of impurities is accomplished by washing the crystals with water.

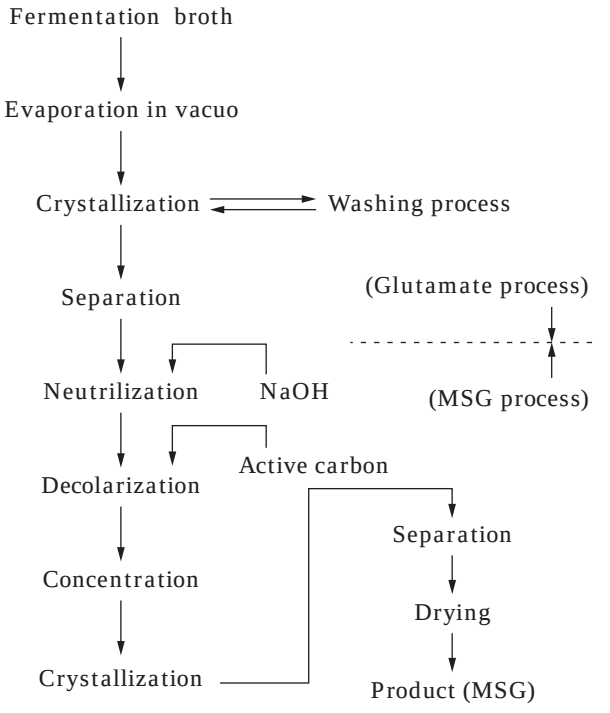


Figure 5. Process flow of purification of monosodium glutamate (MSG).

In the MSG step, ensuring high quality of the final product is essential, and improvement of the yield is required. To this end, both high purity of the GA crystals and efficient use of activated carbon must be attained. The resulting MSG-crystallized mother liquor, which still contains a high concentration of dissolved MSG, is sent back to the GA step to enhance the overall purification yield. In the whole process, from fermentation to final product, the GA-crystallized mother liquor and a decolorized cake of activated carbon are the main wastes discharged from the system. If the high quality of the product is to be ensured, it is apparent that the properties of the fermentation broth strongly affect the yield as a whole. Also, efficient use of the GA-crystallized mother liquor in view of utility and cost is an important factor in evaluating the process. Thus, in MSG production by the direct concentration and crystallization method, the process as a whole, from the selection of raw materials to the final product production, consists of linked steps that cannot be simply compared with other methods such as the resin methods. The purification yield of MSG from L-glutamic acid contained in the fermentation broth is estimated to be nearly 100% (78% in terms of ratio by weight and L-glutamic acid recovery). If the GA-crystallized mother liquor, which is rich in nutrient sources of plant origin and is used as an organic fertilizer, is efficiently used, this process as a whole constitutes an efficient, recycleable production system.

As is already mentioned, yield of L-glutamic acid is strongly affected by the ratio of this substance to the total solids in the fermentation broth. This ratio varies depending upon the sugar purity of the raw material used (ratio

of sugar to total solids) and the conversion to L-glutamic acid based on sugar. Because sugar purity varies greatly, depending on the different raw materials used, selection of appropriate raw materials is important.

FUTURE PROSPECTS OF COMMERCIAL MONOSODIUM L-GLUTAMATE PRODUCTION

Genetic Recombination Technology

There have been many reports concerning the application of current genetic recombination technology to the construction of amino acid–producing bacterial strains. Among them are studies on *Escherichia coli*, for which there has been advanced analysis on the genetic level, and those strains of coryneform bacteria that are most widely used for amino acid production. As shown in Figure 1, the biosynthetic pathway of L-glutamic acid is relatively simple compared with that of other amino acids and requires strong assimilability of ammonia, which indicates that coryneform bacteria are preferred producers of this substance. Figure 6 shows one example of improving L-glutamic acid productivity using recombinants of coryneform bacteria. In this example, productivity was improved by introducing a variant-type gene of L-glutamic acid biosynthetase into *C. glutamicum*. Although the productivity did not reach a practical level, this suggests the possibility of constructing an industrially feasible strain by making further improvements to both the donor gene and receptor microorganism. Currently, success in the application of recombinant technology to construction of amino acid–producing strains has been reported in connection with amino acids that have a relatively complicated mechanism of metabolic regulation, such as the aromatic amino acids histidine and threonine. With respect to amino acids that are produced in good yields as a result of conventional construction method, such as L-glutamic acid, so far there have been no reports published on remarkable productivity improvements achieved by recombinant technology. However, there is still high technical potential, and drastic improvement in production may be expected in the future using this technology.

Strain		L-Glutamic acid produced (mg/dL)
No. 5707	(parent)	600
No. 12	(donor, glu^-)	0
No. 26	(donor, glu^-)	0
AJ 11566	(No. 12/pAM 286)	1,010
AJ 11567	(No. 26/pAM 286)	1,000

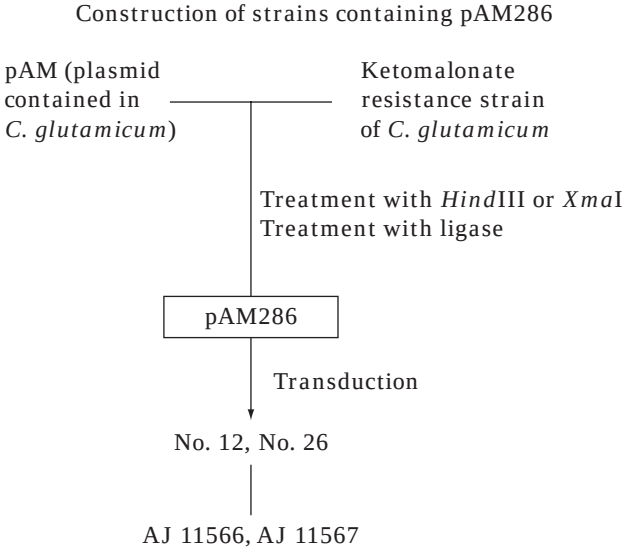


Figure 6. Glutamic acid production by recombinant *C. glutamicum*. *Source*: From Ref. 12.

Continuous Fermentation

The penicillin method batch-type L-glutamic acid fermentation gives only a short period of high L-glutamic acid productivity. Though continuous fermentation is desired to improve productivity, it is difficult to establish a continuous process with the currently available technologies. Studies on measures to stabilize the productivity, and establishment of the process with techniques other than the penicillin method are desired. Further improvement in productivity may be expected in combination with genetic recombination technology.

Replacement of Raw Materials: L-Glutamic Acid Fermentation with Petrochemicals

In the fermentation processes, the main raw materials are carbohydrate derived from molasses, corn, beet molasses, and so forth, with ammonia as a nitrogen source. Although industrial production has been carried out by fermentation with carbohydrate, the possibility of using petrochemicals including normal paraffin, ethanol, methanol, and acetic acid as fermentation raw material has been extensively studied (13–15). However, fermentation production using petrochemicals as a raw material has not been industrialized, except for partial use of acetic acid, because of the cost (16).

Acetic acid, derived from petrochemicals, may partially replace molasses. Glutamic acid–producing microorganisms from carbohydrate can easily utilize acetic acid if it is in low concentration, and acetic acid may be used in industrial processes if its concentration in the culture medium is kept at a low level. Acetic acid is supplied during the fermentation. With this method, although the cost of raw materials for culturing increases, the purification yield can be enhanced by increasing the ratio of L-glutamic acid to the total solid of fermentation broth. The feasibility must be evaluated with respect to the process as a whole.

Due to the rising cost of petroleum, L-glutamic acid production by fermentation from petrochemicals, particularly normal paraffin, is presently not practical, and the prospects for the future are not optimistic. However, this technology is discussed in this section because L-glutamic acid

fermentation (and, similarly, fermentation to produce other amino acids) from hydrocarbons may provide one option for replacing raw materials. Naturally isolated, hydrocarbon-assimilating L-glutamic acid–producing microorganisms did not show a requirement for biotin, unlike those producing L-glutamic acid from a carbohydrate such as molasses; but they did require other vitamins for growth. Typical L-glutamic acid producers of the genera *Corynebacterium* and *Arthrobacter* had a requirement for thiamin. As in the case with fermentation using carbohydrate, L-glutamic acid was accumulated in the medium under the controlled concentration of the required vitamin, and the membrane permeability underwent a change when the microorganism was in the L-glutamic acid production phase. Also, the addition of penicillin or saturated fatty acids or their esters to the cells growing in the presence of a sufficient concentration of the required substance converted the cell from the growing phase to the L-glutamic acid production phase. Whereas carbohydrates are metabolized through glycosylation, hydrocarbons are metabolized through β-oxidization. However, subsequently, L-glutamic acid formation undergoes similar metabolic regulation in both processes.

With respect to the reaction processes, carbohydrates are used basically as an aqueous solution, whereas hydrophobic hydrocarbons separate into a layer of oil on water. Furthermore, oxidization of hydrocarbons requires much more oxygen compared with the oxidation of carbohydrates, although an advantage of the hydrocarbon process is that continuous feeding of the raw material, as required with carbohydrates, is not necessary, and, therefore, high-concentration culturing is possible in a batch culturing manner. The equipment for glutamic acid production must provide efficient dispersion of the oil to enlarge the surface area of the oil layer and also must provide efficient aeration and agitation to supply the large volume of oxygen required. Under optimum conditions, glutamic acid production from hydrocarbons is very similar to production from carbohydrate. There is no major difference between the two processes in the purification step, except that fermentation from hydrocarbons does involve steps for removing trace residual oil and cells.

Environmental Problems

The main raw materials, molasses, beet molasses, and glucose, are agricultural products. It is necessary to establish a production process that takes into consideration the possibility of recycling the wastes produced. The current industrial process uses the remaining fractions to produce organic fertilizers, and the process as a whole is designed to be a closed system.

BIBLIOGRAPHY

1. S. Kinoshita ed., *Hakko to Kogyo*, Dainihon Tosho, Tokyo, 1975, p. 125.
2. S. Kinoshita, S. Udaka, and M. Shimono, *J. Gen. Appl. Microbiol.* **3**, 193–205 (1957).
3. K. Yamada, J. Takahashi, and K. Kobayashi, *Agric. Biol. Chem.* **27**, 773–783 (1963).
4. K. Tanaka, K. Yamaguchi, and S. Kinoshita, *Amino Acid Nucleic Acid* **17**, 136–142 (1968).
5. H. Soda, K. Takinami, I. Chibata, K. Nakayama, and S. Yamada eds., *Amino Acid Fermentation*, Gakkai Shuppan Center, Tokyo, 1986, p. 201.
6. K. Sonntag, J. Schwinde, A.A. de Graaf, A. Marx, B.J. Eikmanns, W. Wiechert, and H. Sahm, *Appl. Microbiol. Biotechnol.* **44**, 489–495 (1995).
7. E. Kimura, C. Abe, Y. Kawahara, T. Nakamatsu, and H. Takuda, *Biochem. Biophysic. Research Communication* **234**, 157–161 (1997).
8. E. Kimura, C. Abe, Y. Kawahara, and N. Nakamatsu, *Protein Nucleic Acid Enzyme* **42**, 2633–2640 (1997).
9. H. Motoyama and H. Anazawa, *Kagaku to Seibutsu* **35**, 123 (1997).
10. T. Uemura and H. Soda eds., *Fermentation and Microorganisms*, Asakura Shoten, Tokyo, 1971, p. 53.
11. Y.C. Su and K. Yamada, *Bull. Agric. Chem. Soc. Jpn.* **24**, 69–74 (1960).
12. Tsuchida, K. Miwa, S. Nakamori, and H. Momose, Japanese Unexamined Patent Publication No. 148296/81, November 17, 1981.
13. K. Miyai, I. Tsuruo, R. Ogawa, S. Hayakawa, Y. Akimoto, and K. Goto, *Nippon Nogieikagaku Kaishi* **37**, 32–36 (1963).
14. S. Okumura, R. Tokawa, T. Tsunoda, K. Kohno, T. Matsui, and N. Miyaji, *Nippon Nogeikagaku Kaishi* **36**, 141–159 (1962).
15. Association of Amino Acid and Nucleic Acid ed., *Amino Acid Fermentation*, Kyoritsu Shuppan, Tokyo, 1972, p. 15.
16. T. Tsunoda, I. Shiio, and K. Mitsugi, *J. Gen. Appl. Microbiol.* **7**, 18 (1961).

PROFESSIONAL SOCIETIES, ASSOCIATION OF BIOMOLECULAR RESOURCE FACILITIES

RONALD L. NIECE
University of California–Irvine
Irvine, California

KEN WILLIAMS
Yale University
New Haven, Connecticut

CLAYTON NAEVE
St. Jude Children's Research Hospital
Memphis, Tennessee

OUTLINE

INTRODUCTION

The primary purpose of this article is to benefit the tens of thousands of individuals who either utilize biomolecular analyses and syntheses provided by resource laboratories or help to oversee, administer, and regulate the hundreds of biopolymer core laboratories that are dedicated to support research, quality control, and quality assurance programs. The core laboratories that provide these biopolymer analyses and syntheses are in turn supported by the Association of Biomolecular Resource Facilities (ABRF), an international, nonprofit organization of more than 300 core laboratories that is truly remarkable in terms of the ever-widening expanse of its research and outreach activities; the dedication, volunteerism, and extent of participation of its members; and the laudable goals to which it aspires. These are well summarized in its mission statement:

> A. To promote and support resource facilities, research laboratories, and individual researchers regarding operation, research, and development in the areas of methods, techniques, and instrumentation relevant to the analysis and synthesis of biomolecules.
> B. To provide mechanisms for the self-evaluation and improvement of procedural and operational accuracy, precision, and efficiency in resource facilities and research laboratories.
> C. To provide a mechanism for the education of resource facility and research laboratory staff, users, administrators, and interested members of the scientific community.

Following a brief history, which describes the genesis of bioanalytical core laboratories and the organization they engendered, the ABRF, are nine sections that chronicle the work of ABRF's research committees. These sections are followed by a section on the effective use of biomolecular resource facilities, and the article concludes with a prognostic look into the expanding role that core laboratories are likely to play in the postgenome era.

HISTORY AND ACTIVITIES OF THE ABRF

RONALD L. NIECE
University Of California–Irvine
Irvine, California

The formation of shared biotechnological instrumentation laboratories in the 1980s provided a major and highly efficient mechanism to bring state-of-the-art biomolecular syntheses, analyses, and advice to tens of thousands of investigators at minimal cost. Growth in the number of biotechnological core laboratories and the parallel growth in the diversity of services they offer reflect the increasing sophistication of the protocols and instrumentation needed to conduct research in the life sciences. The ever-increasing complexity of the equipment increases the cost of conducting research in multiple ways. In addition to higher capital costs, operating expenses are higher due to high-purity custom reagents and increased service contract and maintenance costs. Specialized parts and highly trained instrument service representatives contribute to the cost of high-performance instrumentation. Laboratory staff costs are higher because of the need for experts in operation of the instrumentation and interpretation of the resulting data. The establishment of biomolecular resource laboratories has provided expensive and sophisticated technology at minimal cost.

Core laboratories began to evolve in the early 1980s from individual research laboratories and programmatic centers that were often equipped with an automated amino acid analyzer and/or spinning cup protein/peptide sequencer and that provided analyses on an informal or collaborative basis. This evolution was spurred on by the introduction and commercialization of the gas phase protein/peptide sequencer (1) and automated oligonucleotide synthesizers (2) and by the inauguration of the Shared Instrumentation Grant Program of the National Institutes of Health in 1981.

An awareness that many laboratories were providing services for other researchers (i.e., they were carrying out noncollaborative analyses on samples not prepared in their own laboratories) emerged at the Symposium of American Protein Chemists in San Diego, California, in 1985. This led to the informal association of Research Resource Facilities (RRF) (3), which met annually to exchange views on sample handling in the core facility setting, improvements in protocols, and advances in technology. A significant innovation of RRF was the inauguration of research committees devoted to analyzing instrumentation capabilities and to providing a realistic definition of state of the art in terms of what could be accomplished by the majority of the practitioners of the art. The first "unknown" sample presented to affiliates was a peptide designed to assess the performance of automated protein sequencers and amino acid analyzers (4). In 1989 the informal RRF group incorporated as the Association of Biomolecular Resource Facilities (ABRF). ABRF research committees based on other themes including peptide synthesis, mass spectrometry, nucleic acids chemistry, DNA sequencing, laboratory quality and compliance, and carbohydrate analysis were established subsequently (see following sections).

As it has been from the beginning, a major focus of ABRF is to establish realistic expectations of what can be accomplished with currently available technology in resource and research laboratories. The laboratories using the technology provide the best definition of requirements for sample preparation or design, instrument performance, and personnel proficiency. Since 1988 ABRF research committees have designed experiments and distributed test samples and syntheses for membership participation (see later sections) following the model described (4) for the 1988 sequence and amino acid analysis sample. An "unknown" sample or synthetic project that reflects a typical task is distributed to members with the recommendation that it receive no more nor less care than would any other request made to the laboratory.

The research committees analyze the resulting data and synthesized products as described in the sections that follow. The results help establish a useful assessment of the range of performance and capabilities on instruments in member laboratories. Because the data and products are submitted anonymously, individual core facilities cannot be identified. The anonymity helps ensure that samples and data returned are in fact representative of what happens on a day-to-day basis and allows member laboratories to compare their performance relative to that of their peers and relative to the level appropriate to their constituency. Unless a laboratory knows that it is performing below the average level defined by peers using similar instruments, reagents, and protocols, there is little incentive to improve. The best analyses and highest-quality synthetic products highlight the level of performance possible under actual operating conditions and identify the conditions needed to attain this high level of performance. Surveys of the composition, capabilities, and performance of core laboratories provide a complementary means of defining realistic expectations (5–7). In addition, survey data on operations and financial support and expenses of core facilities is collected anonymously as with the "unknown" samples.

Comparison of performance in terms of sensitivity, throughput, turnaround time, and so forth with the average provides a performance baseline that can serve as an incentive for performing as effectively as one's peers. The detailed financial data provided by the surveys have documented the economics of core facilities. Although the centralized laboratory with shared instrumentation provides considerable economies of scale and pooled expertise, the survey data demonstrate that only a small percentage are able to recover 100% of operating expenses from user fee income.

Results of the research studies are distributed directly to the members who participated and are also published (3–50). Originally, the research results were presented at an annual satellite meeting held in conjunction with the Symposium of the Protein Society. Today symposia, workshops, and tutorials are presented by the ABRF. Symposia on leading-edge technologies presented at national and international scientific meetings introduce the scientific community to the range of capabilities available in core facilities while highlighting the constraints on sample preparation or design required for success. Because ABRF consists of the users of the technology, the workshops emphasize the practical tools needed for improved and high performance. The tutorials are popular and effective tools for helping core facility staff expand their repertoire of capabilities and increase their level of expertise.

The research undertaken by the ABRF has been funded by federal grants, corporate sponsors, institutional contributions, and membership dues. These funds are leveraged by voluntary contributions of individual research committee members, their laboratories, and their institutions. Research committees consist of volunteer ABRF members with a designated liaison from the executive board. ABRF is an incorporated, nonprofit organization led by an elected, volunteer executive board. Business activities are handled by the Federation of American Societies for Experimental Biology (FASEB) business office (ABRF, 9650 Rockville Pike, Bethesda, MD 20814-3998; 301-571-8300; FAX: 301-530-7049; E-mail: *abrf@faseb.org*). Official ABRF business is conducted at an annual business meeting.

The ABRF Corporate Sponsorship program provides a unique opportunity to enhance communication between the manufacturers and suppliers of biotechnological instrumentation and the laboratories that depend on this equipment. ABRF research activities help define how effective the instruments and protocols are in the field, the level of training required for their optimum use, and what happens to this performance as the instruments age. Identifying performance levels, new applications, and unforeseen problems that accompany the running of "real" samples that were not anticipated by manufacturers promotes improvements and corrective actions. ABRF surveys of satisfaction with instrument performance, service support, reagent and supply costs, and application support offer unbiased group feedback from the day-to-day users who best understand the instrumentation and protocols.

Since 1994 the ABRF has presented an annual award for outstanding contributions to the development of technologies used in core facilities. An independent, anonymous panel of respected scientists presents the ABRF executive board with candidates for this award. This ABRF award recognizes contributions to science through technological advances and is accompanied by $10,000 and an opportunity to address the ABRF members at their annual meeting. In 1994 and 1995 Beckman Instruments funded the awards; in 1996 Hewlett-Packard began sponsoring of the award. The first award was presented to Dr. Frederic Sanger, the winner of two Nobel prizes, for his development of methods for sequencing proteins (chemistry, 1958) and for sequencing DNA (chemistry, 1980). In 1995 Dr. Klaus Biemann was honored for his contributions in the development of mass spectrometry for analysis of biomolecular polymers and for his training of mass spectrometrists. Dr. David Lipman received the 1996 award for developments of computational methods for analysis of protein and DNA sequences. The 1997 award was presented to Dr. Lloyd Smith for his contributions to the advancement of automation in DNA sequencing. The 1998 award went to Nobel laureate (chemistry, 1984) Dr. Bruce Merrifield for developing the concept and strategies for solid-phase synthesis of peptides, which opened the way for peptide and other organic syntheses to become routine.

ABRF educational initiatives target member laboratories, the scientific community, which represents its user base, the manufacturers of instrumentation and reagents used in core facilities, and the administrators and decision makers who deal with science, scientists, and the delivery of research resources. To fulfill the educational needs, ABRF promotes several activities beyond conducting and publishing research reports. Since 1986 symposia have been organized in conjunction with the Protein Society and the International Congress of Biochemistry and Molecular Biology and joint symposia have been presented at the FASEB and American Society of Biochemistry and Molecular Biology (ASBMB) annual meeting. Workshops and symposia have been presented at meetings of the American Peptide Society, American Society of Cell Biology, American Society of Mass Spectrometry, European Protein Society, International Genome Sequencing and Analysis Conference, and the Methods in Protein Structure Analysis. From 1986 to 1998 the ABRF organized workshops and symposia at 32 different meetings and held three independent meetings, each of which were attended by 500 to 1,000 individuals.

Another significant medium for information exchange and education of members has been the newsletter, *ABRFnews*. The newsletter contains peer-reviewed articles including summaries of meetings, methods and review papers, and laboratory tips. A new dimension in publication of methods in biomolecular research began in 1997 with the launch of the electronic *Journal of Biomolecular Techniques: The Official Methods Journal of the ABRF (http://www.abrf.org/JBT/JBT.html)* edited by John Shively. Beginning in 1998, *ABRFnews* and the electronic journal were merged into one under the title *Journal of Biomolecular Techniques*, with Clayton Naeve as the editor in chief. To maintain the advantages of the electronic journal, accepted rapid communications are immediately posted on the Web site followed by reprinting in the next issue of the journal. Electronic communications using an E-mail bulletin board (*abrf@aecom.yu.edu*) connect more than 900 members around the world; the novice and expert can almost instantaneously assess instrument and reagent problems, discover if other laboratories have similar concerns, and help one another design experimental protocols. The ABRF Web site (*www.abrf.org*) contains written protocols distributed at workshops and tutorials, the results of many research studies, and comprehensive archives of the discussions on the E-mail bulletin board. The Web site averages more than 100 inquiries per hour to one or more documents.

The concept that developed into ABRF began with six scientists in 1986. In 1998 there were more than 900 members listed in the ABRF directory. Members reside in 26 different countries on six continents. Three-quarters (76%) are in the United States, from 40 different states, half (55%) are in academic institutions, 13% in research institutions, 28% in industry, and the balance in government laboratories. Although these numbers are illustrative, they fail to adequately convey the underlying reason for the success of ABRF: The ABRF is uncommon in terms of the unrelenting volunteer spirit and commitment of its members and its single-minded purpose of helping to bring state-of-the-art biotechnology to bear on the myriad array of biochemical and biomedical challenges that are its reason for being. It is unique in its multipronged, targeted research by its research committees, which study technical capabilities, current protocols, and emerging technologies.

AMINO ACID ANALYSIS

DANIEL J. STRYDOM
Bio Nebraska, Inc.
Lincoln, Nebraska
JOHN W. CRABB
Cleveland Clinic Foundation
Cleveland, Ohio

Amino acid analysis (AAA) is one of the major analytical techniques used in the biochemical and biotechnical environments. It is the method of choice for reliable quantification of peptides and proteins and also provides useful qualitative information in the evaluation of synthetic peptides and recombinant proteins, the identification of proteins based on composition, and the detection of unusual amino acids. AAA also has applications in protein sequence studies, such as in the design of fragmentation strategies, the identification of amino- and carboxy-peptidase reaction products, and the quantification of samples subjected to Edman degradation to detect those with blocked N-termini. In addition, AAA is used for the characterization of feed and food samples and physiological fluids.

Two basic approaches to quantitative AAA receive widespread use, namely (*1*) classical postcolumn derivatization followed by ion exchange chromatography and (*2*) precolumn derivatization followed by reverse-phase HPLC. Excellent instrumentation and more than one method is available for these basic approaches. Because AAA applications are of such varied character, no single methodology can be shown to be superior in all respects. The technology is deceptively simple but actually is one of the more complex and difficult in protein chemistry. The greatest benefit of AAA is its quantitative capability; however, measuring many different amino acids with quantitative accuracy is precisely what makes the analysis so difficult. AAA is a statistical type of measurement; amino acids vary widely in their stability/lability to acid hydrolysis, oxidation, and modification. Trp and Cys are destroyed by HCl hydrolysis, Ser and Thr are partially destroyed, Ile and Val are slow to cleave, Met is subject to oxidation, and Asn and Gln are completely deamidated. Variable, but ubiquitous, background (most frequently Gly, Ala, and Ser) plus variable instrument and analyst performance can confound the analysis, particularly at and below the low picomole level. As a consequence, realistic expectations concerning the accuracy and precision of AAA must be tempered with an appreciation for the difficulties of the technology.

The annual ABRF AAA studies between 1988 and 1997 provide a broad performance comparison of AAA methods and instruments (Table 1). In addition to the reports in Techniques in Protein Chemistry (Table 1), the ABRF AAA studies have been highlighted in the *ABRFnews* and reported elsewhere (12,13,49). The number of participating

Table 1. Summary of ABRF Amino Acid Analysis Studies, 1988–1996

Year	Sample description	Number of participants	Sample amount[a] (μg)	Average % error[b]	Study focus	Reference
1988	Synthetic peptide (STD-2)	45	5	19.3	Accuracy and sensitivity	4
1989	β-Lactoglobulin (89AAA)	17	0.2	25.3	Accuracy, precision, sensitivity	10
		26	5.0	13.9	Hydrolysis methods; and Cys quantification	
1990	Myoglobin control (90AAA1)	41	1.3	13.5	AAA in the presence of buffer salts and on PVDF membrane	15
	Myoglobin + salt (90AAA2)	35	1.3	19.7		
	Myoglobin on PVDF (90AAA3)	30	0.9	26.9		
1991	Bovine serum albumin (91AAA)	51	150	9.5	Accuracy with unlimited sample; Trp and Cys quantification	18
1992	Chymotrypsin (92AAA)	59	56	10.5	Trp and Cys quantification	21
1993	Synthetic peptide (93AAA)	50	23	12.5	Phosphoamino acids and hydroxyproline analyses	26
1994	RNase + angiotensin (94AAA1)	58	30	10.9	Hydrolysis and chromatography methods; cys and glucosamine analyses	30
	Prehydrolyzed 94AAA1 + glucosamine (94AAA2)	60	30	6.5		
1995	Myoglobin on PVDF (95AAA)	77	30	21.4	AAA on PVDF membrane	35
1996	Triose phosphate isomerase (96AAA)	71	10	11.9	Protein identification from AAA databases	48
1997	Bovine serum albumin, lysozyme on PVDF (97AAA)	40	1	30	AAA on electroblotted PVDF	49
			5	19	Protein identification from AAA	
			10	19		

[a]Sample amounts provided since 1991 were meant to be adequate for all methods and not to restrict performance.
[b]Average percentage compositional error for each study is shown for all methods combined without results for Cys, Trp, or unusual residues.

laboratories per year has varied from about 40 to 77 (Table 1); approximately 38% of the participants since 1989 have used postcolumn ninhydrin methodology, approximately 48% have used precolumn phenylthiocarbamyl (PTC) methods, and 14% have used other AAA methods. The ABRF AAA studies are crafted to examine some aspect of the AAA technology to help participating laboratories evaluate and improve their capabilities; participation is organized in an anonymous and noncompetitive manner. In general, these studies demonstrate that more than one AAA method works well and that the skill of the analyst is probably the most important determinant of success. More specifically, these collaborative trials have evaluated AAA performance with high and low amounts of sample, in the presence of salt, on PVDF membrane, with emphasis on difficult residues (e.g., cysteine/cystine, tryptophan, and phosphoamino acids), with special focus on hydrolysis and chromatography conditions, or with attention to protein identification methods based on amino acid composition data. Each year, these studies have published the average levels of AAA performance and also typically spotlighted the highest-quality analyses attained with different AAA methods.

The overall AAA accuracy in the past 10 years of ABRF studies has ranged from 6.5% to 30% average compositional error per study (Table 1). Accuracy is strongly dependent on the type and amount of sample provided, hydrolysis conditions, and level of background contamination during the analysis. The ninhydrin-based postcolumn analyses require more sample (typically 5–10 μg per analysis) than precolumn derivatization–based techniques (0.5–1 μg per analysis). Each year a broad range of accuracy is apparent, including some analyses that are grossly wrong and others that are exceptionally good (e.g., 3–5% average error). The studies indicate that knowledgeable and skillful analysts can expect high-quality results with approximately 10% average compositional error from more than one AAA method.

Instrument calibration and peptide/protein hydrolysis to individual amino acids are critically important for high-quality AAA. General parameters for sample hydrolysis have been considered frequently in the ABRF studies (e.g., vacuum, temperature, acid additives, liquid acid versus acid vapor), and two trials have specifically evaluated the hydrolysis process by providing both a hydrolyzed and an unhydrolyzed sample for analysis (10,30). Study results clearly demonstrate the variability of the hydrolysis process from laboratory to laboratory (30) and emphasize the importance of controlling this step. Vapor-phase acid hydrolysis has become the most popular methodology over the course of these trials. Instrument calibration has been examined also, and results demonstrate that replicate analysis of amino acid standards such as Pierce standard H provided superior results, whereas

hydrolyzing the standard appears to be disadvantageous (18).

Quantification of cysteine is difficult because hydrolysis in constant-boiling HCl destroys cysteine and partially destroys the disulfide-linked, dimeric form of the residue, cystine. Tryptophan is destroyed by routine HCl hydrolysis and is even more challenging to quantify than cysteine. Cys is best analyzed by various pretreatments of the sample, such as oxidation, reduction and alkylation, or perhaps optimally by adding disulfide exchange reagents during acid hydrolysis (21). Trp appears best analyzed by adding dodecanethiol to the hydrochloric acid used for hydrolysis (21).

Rare and modified residues are of continual interest to many investigators. Phosphoamino acids and hydroxylated and glycosylated residues have been examined in the ABRF collaborative trials; such analyses are not routine for most laboratories. No current AAA methodology allows quantitation of phosphoamino acids in a single analysis in conjunction with all the common amino acids (26). Optimal hydrolysis conditions must be determined for each sample; however, precolumn analyses such as the phenylthiocarbamyl (PTC) method were more effective in the chromatographic resolution of phosphoamino acids than the postcolumn ninhydrin method (26). In analyses of the 93AAA sample, 41% of the facilities detected an unknown amino acid and 75% of these correctly identified it as hydroxyproline. Glucosamine in the 94AAA2 sample was detected by only 13% of the participants, although all had been alerted to its possible presence.

Comparison of ninhydrin and PTC methodologies in the ABRF AAA studies reveals statistically significant differences. For example, the interlaboratory variability of Arg, His, Leu, Lys, Met, Pro, and Val is dramatic between these two techniques, perhaps due to chromatography issues and reagent-related peaks. Pro is the only of these amino acids that appears to be analyzed more consistently and accurately by PTC than by ninhydrin methods, primarily due to its low color yield with ninhydrin. Met is underestimated by ninhydrin sites, probably due to oxidation and to chromatographic disturbances, and it is overestimated by PTC sites. The latter can be due to overlap of the PTC Met peak with reagent-related peaks. The oxidation of Met leads to high Arg values in PTC analyses due to coelution of the oxidation product with Arg, whereas ninhydrin methods frequently underestimate Arg, most likely due to the very broad chromatography peak and consequent integration problems. Other residue-specific problems were addressed in the 1991 study (18), including low values for PTC Asp and PTC Glu caused by metal contamination from glass hydrolysis tubes that often can be remedied by drying the samples with EDTA prior to hydrolysis. Although more than one ABRF AAA study has shown precolumn derivatization techniques more sensitive than ninhydrin methods, the 1990 study demonstrated that postcolumn derivatization methods generally provided better AAA results in the presence of buffer salts.

Two ABRF studies have evaluated AAA on PVDF membranes (15,35), and both resulted in compromised accuracy (21–27% average error) relative to nonimmobilized samples. The 1995 study further demonstrated significantly high backgrounds on blank PVDF membranes (35), suggesting that perhaps the higher error from PVDF immobilized samples is due in part to background contamination that is difficult to extract from the membranes. The 1997 ABRF AAA study readdressed AAA on PVDF by probing whether PVDF electroblotted samples exhibit less background contamination. Results in 1997 indicated that the average error for samples blotted to PVDF was still about twofold higher than for solution samples (49).

The 1996 ABRF AAA study (48) demonstrated that 90% of participating laboratories could correctly identify an unknown sample protein (triose phosphate isomerase) from their AAA data (11.9% average compositional error) using the World Wide Web and either ExPASy *(http://expasy.hcuge.ch/ch2d/aacompi.html)* or PROPSEARCH *(http://www.EMBL-heidelberg.de/aaa.html)* database computer query programs. Results from the 1997 AAA study on PVDF membrane suggest that correct protein identifications can be made with AAA data containing up to 15% average error. In the years to come, AAA will likely become a more widely recognized and valuable tool for the identification of proteins.

PROTEIN SEQUENCING

JOSEPH FERNANDEZ
Rockefeller University
New York, New York

KATHRYN STONE
Yale University
New Haven, Connecticut

The ABRF Protein Sequence Research Committee focuses primarily on providing member laboratories with "unknown" protein/peptide samples that are representative of samples routinely submitted for protein sequence analysis. Participating laboratories characterize the primary structure of the protein/peptide, or the sequence, typically from the amino-terminal end of the sample utilizing automated Edman degradation chemistry.

Chemistry and Sample Preparation

The chemistry used during Edman degradation (51) consists of consecutive cycles; one cycle contains the following steps: (*1*) coupling of phenylisothiocyanate (PITC) with the α-amino group or amino-terminal amino acid of the protein/peptide at pH 9–10 to form a phenylthiocarbamyl (PTC) group, (*2*) removal of chemistry by-products using organic washes, (*3*) cleavage with anhydrous trifluoroacetic acid (TFA) to generate a protein/peptide minus the first amino acid (or N-1 protein/peptide, which contains a free α-amino group for a subsequent cycle of Edman sequencing), along with an anilinothiazolinone (ATZ) derivative of the released amino-terminal amino acid, (*4*) extraction of the ATZ derivative from the N-1 protein/peptide, and (*5*) conversion of the ATZ derivative to the more stable phenylthiohydantoin (PTH) form. The PTH amino acid derivative can then be identified by its retention time on reverse phase HPLC, during which amino ac-

ids, chemistry artifacts, and derivatized amino acids are resolved from each other. Figure 1 shows chromatograms distributed to ABRF members as a sequence assignment exercise (46) and represent an example of a PTH amino acid standard HPLC profile, along with Edman cycles showing residues 1–7 of the peptide. Cysteine is destroyed during the Edman chemistry and thus is usually reduced and alkylated before automated sequence analysis (52).

The performance and capabilities of automated Edman degradation chemistry are generally characterized by initial yields, repetitive yields, and lag (carryover). Initial yield refers to the recovery of the first amino acid, typically in picomoles, and is generally expressed as a percentage of the total picomole amount of protein/peptide analyzed (typically 50–80%). Repetitive yield provides a measure of the average amino acid recovery after each cycle. Although it usually ranges from 90 to 99%, repetitive yield also varies with sample and instrumentation. Lag or carryover means the yield, expressed as a percentage, of the amino acid that appears in the subsequent cycle ($n + 1$) relative to the actual cycle (n). Repetitive yield and lag indicate the degree of completion of the Edman chemistry and generally determine how far a protein can be sequenced before the signal becomes indistinguishable from the noise.

Edman degradation is only possible when the protein/peptide has a free amino (N) terminus. If a naturally occurring or other chemical group is present at the N-terminus, no data can be obtained and the sample is considered blocked. Some of the more common naturally occurring blocking groups are acetylated amino acids, formylated methionine, and pyroglutamic acid. Generally, the percentage of naturally blocked proteins isolated from an organism increases along with the complexity of the organism. Hence, while only a few bacterial proteins are blocked, approximately 80% of soluble proteins isolated from a mammalian tumor cell line appear to be blocked (53). Artificial blockages can occur during electrophoresis, but certain precautions, such as using high-purity reagents, can be taken to help decrease this type of blocking (54). A lack of sequence or PTH amino acid signal can also be due to an insufficient amount of protein/peptide used during sequencing. The reason for a lack of sequence (a blocked amino terminus or an insufficient amount of protein/peptide) can often be inferred by performing hydrolysis and an AAA (see previous section) of the sample and its sequencing support, after sequencing.

Samples for sequence analysis by Edman degradation are generally prepared in one of three forms: (*1*) in solutions that often contain nonvolatile buffers (phosphate, Tris, SDS, etc.), (*2*) in volatile buffers, such as the trifluoroacetic acid/acetonitrile solutions used to elute reverse-phase HPLC columns, and (*3*) noncovalently bound (via electrophoretic transfer) to PVDF membranes (55). All samples in these forms can be directly sequenced, although certain instruments are better designed for one sample form over another.

ABRF Protein Sequence Studies

Over the past 10 years, the ABRF Protein Sequence Research Committee has distributed annually an unknown amino acid sequencing sample, along with a survey, to member laboratories. The purpose of these studies is to provide participating laboratories with a mechanism for evaluating their own sequencing capabilities, particularly as new instruments and techniques are introduced. Because many laboratories are involved in these studies (the average number of participating laboratories is 69; see Table 2), they provide a reasonable approach to determine realistic expectations regarding the average amount of protein/peptide that can be routinely sequenced. In addition, the quality or accuracy of the resulting data can be determined because the identity of the sample was unknown to the participants. Each study has been designed with a particular emphasis: sensitivity of sequencing (STD-1, ABRF-93SEQ, and 97SEQ), posttranslational modification (ABRF-92SEQ), heterogeneity (ABRF-89SEQ, ABRF-95SEQ, ABRF-96SEQ, and ABRF-97SEQ), protein bound peptides (which mimic large proteins) on PVDF or in solution (ABRF-90SEQ and ABRF-91SEQ), identification of problematic residues such as cysteine and tryptophan (ABRF-94SEQ, ABRF-95SEQ, and ABRF-96SEQ), and distance sequenced (ABRF-95SEQ). In addition, emphasis is placed on positive correct calls or positive accuracy, which is defined as the number of correct positive calls per total number of positive calls. A summary of these studies is found in Table 2.

The amount of sample supplied in these studies has ranged from a high of 500 pmol (ABRF-92SEQ) to a low of 10 pmol (ABRF-97SEQ). In the first study (STD-1, 1988 [4]), 100 pmol of a unique 40-residue synthetic peptide was supplied. The positive accuracy for this study was quite good, with ~95% correct. Because real samples often contain some heterogeneity, the following study (ABRF-89SEQ [8]) contained a mixture of two unique synthetic peptides at a 5:1 molar ratio (248:48 pmol). The lengths of these peptides were similar to the STD-1 sample, and the positive sequence calling accuracy for the major sequence was again ~95%. This indicated that the participating laboratories were quite good at differentiating a major sequence from a minor sequence when the abundance of the minor sequence was ~20% of the major. Furthermore, it provided a guideline of about 80% for the minimal requirement of purity for samples destined for N-terminal sequencing. A sequence mixture was also provided as a sequence assignment exercise in the ABRF-96SEQB (46) study, where participating laboratories received the HPLC chromatograms of a sequence mixture in place of an actual sample. The positive accuracy for the major, 10-pmol, 22-mer sequence was still 95%. In this study, the accuracy for the minor 14-mer sequence, which was at the 2-pmol level, was also determined and found to be 85%. This lower sequence calling positive accuracy most likely resulted from either the very low amount of sample sequenced or the fact that it was a minor component.

The ABRF-93SEQ (25) sample had the same composition and length as the STD-1 sample. However, only 50 pmol of ABRF-93SEQ was supplied as compared with the 100 pmol of the STD-1 sample that was distributed for the 1988 study (Table 2). A comparison of several sequencing performance parameters (positive sequence calling accuracy, number of positive correct assignments, and average

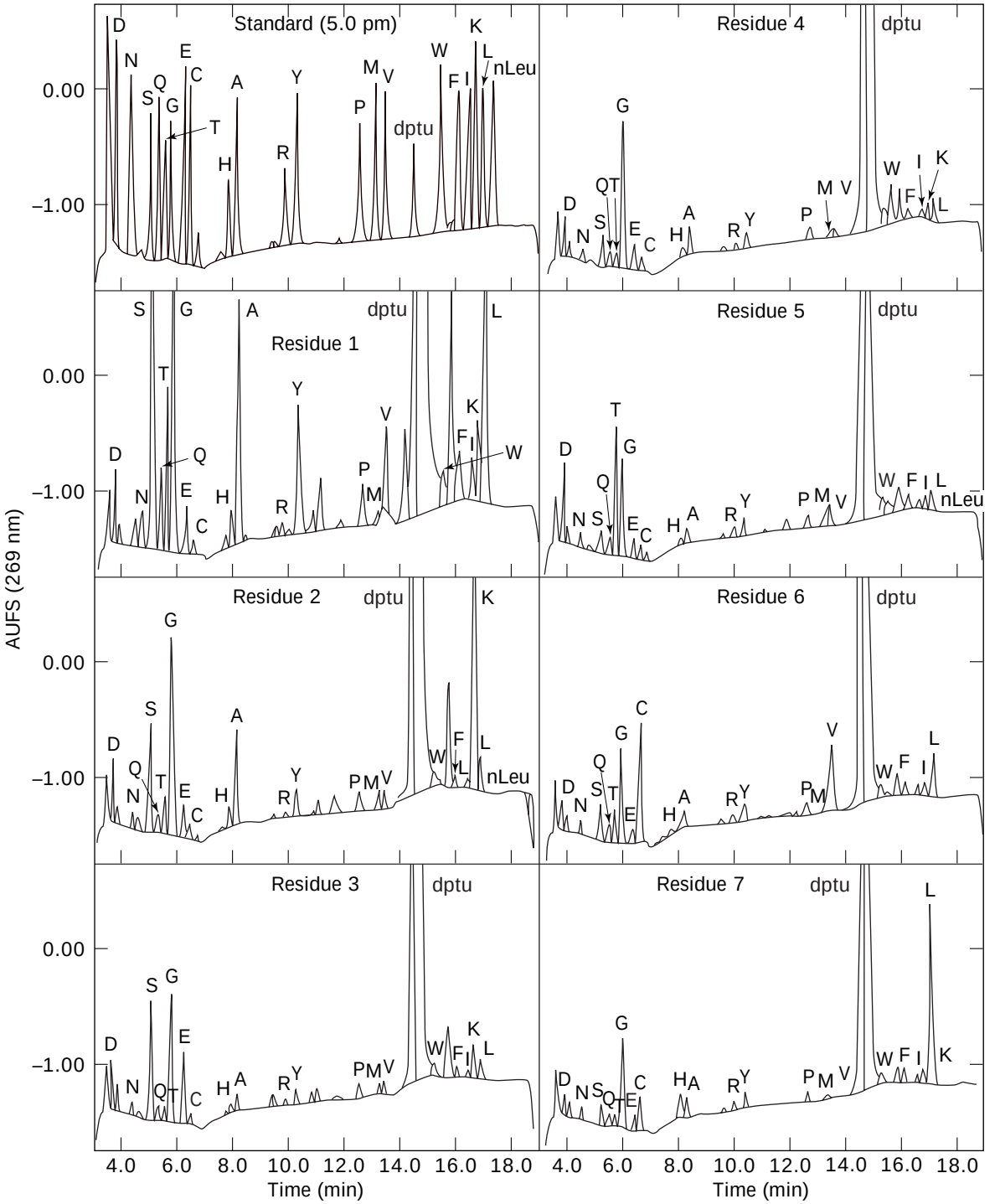


Figure 1. PTH chromatograms from ABRF-96SEQ Dataset B that was distributed as a sequence assignment exercise. Cysteine was present in both the standard and sample as the carboxamidomethylated derivative. The derivative nLeu represents PTH norleucine and is present only in the standard, and diphenylthourea (dptu) is an Edman chemistry by-product. The sequence of the major peptide indicated was leucine-lysine-serine-tryptophan-threonine-cysteine-leucine (LKSWTCL) and the minor was tyrosine-alanine-glutamic acid-glycine-aspartic acid-valine-histidine (YAEGDVH).

Table 2. Summary of ABRF Protein Sequencing Studies, 1988–1997

Sample	Sample description	Number mailed	Number responses[a]	Amount sent (pmol)	Positive accuracy (%)	Reference
STD-1	40-residue synthetic peptide	103	47 (47%)	100	~95	4
ABRF-89SEQ	40-residue synthetic peptide (major)	123	49 (40%)	240	~95	8
	43-residue synthetic peptide (minor)			48		
ABRF-90SEQ	29-residue synthetic peptide covalently attached to acetylated transferrin (90%) bound to PVDF	132	56 (42%)	30	83	14
ABRF-91SEQ	29-residue synthetic peptide dried with acetylated transferrin (90%)	198	90 (45%)	80	83	16
ABRF-92SEQ	37-residue synthetic peptide with two unusual amino acids	225	74 (32%)	500	94	22
ABRF-93SEQ	40-residue synthetic peptide with same composition as STD-1	274	80 (29%)	50	91	25
ABRF-94SEQ	Lactoferrin to assist sequencing labs in identification of cysteine and tryptophan	258	78 (30%)	50	95	31
ABRF-95SEQ	Recombinant protein to determine maximum length of sequences that can be obtained; sample also contained four heterogeneous positions	252[b]	71 (28%)	45	78	37
ABRF-96SEQA	Sequence calling exercise (single sequence)	210[b]	95 (45%)	40	99.8	46
ABRF-96SEQB	Sequence calling exercise (major)	210[b]	95 (45%)	10	96	46
	Sequence calling exercise (minor)			2	85	46
ABRF-97SEQ	21-residue synthetic peptide (major)	215	50 (23%)	10	92	
	14-residue synthetic peptide (minor)			2	72	

[a]Percentage based on number of responses divided by number mailed.
[b]Samples for ABRF-95SEQ, ABRF-96SEQ, and ABRF-97SEQ were mailed only to facilities that perform protein sequencing.

repetitive yield) between the STD-1 and ABRF-93SEQ studies indicated there was no significant improvement in sequencing performance over this 5-year period. Again, however, the amount of peptide provided in the ABRF-93SEQ study was half that provided in the STD-1 study.

To better study the factors that affect sequence calling accuracy, the ABRF-96SEQA and ABRF-96SEQB (46) studies were sequence calling exercises. In each case, the PTH chromatograms were supplied to participants rather than an actual sample to be sequenced. The ABRF-96SEQA sample was at a relatively high level (40 pmol) and had the highest degree of positive correct calls for any ABRF sequence study, with an accuracy of 99.8%. This study indicated that PTH chromatograms generated from straightforward, high-level samples with no heterogeneity can be used to assign a positively called sequence with a very high level of confidence. ABRF-96SEQB was a sequence mixture as described previously, and here the positive sequence calling accuracy was 96%. ABRF-97SEQ also contained a mixture of two peptides and was similar to ABRF-96SEQB in quantity (~10:2 pmol ratio) and peptide composition. The ABRF-97SEQ study, which involved member laboratories running the sample and interpreting the resulting data, was designed to be comparable with the ABRF-96SEQB sequence calling exercise. The positive sequence calling accuracy of ABRF-97SEQ major was 92%, which was slightly lower than the ABRF-96SEQB major (96%). However, a comparison of the positive sequence calling accuracy obtained from ABRF-97SEQ samples run by the committee using the same instrumentation as the ABRF-96SEQB analysis shows the same positive sequence calling accuracy (96%). Thus, the positive sequence calling accuracy appears to be more dependent on instrumentation (model and upkeep) and sample handling and loading than on data interpretation.

ABRF-90SEQ (14) differed from all of the other studies in that this sample was adsorbed to a PVDF membrane. A major advancement in Edman sequencing was realized when Matsudaira (55) discovered that proteins could be sequenced from sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) after electroblotting to a PVDF membrane. Sequencing from PVDF membranes propelled SDS PAGE from an analytical tool to a preparative procedure; it is still the method of choice for the final purification of proteins destined for N-terminal and internal (see next section) sequencing. The goal of the ABRF-90SEQ sample was to mimic a large electroblotted protein. Hence, a 29-residue synthetic peptide was covalently attached to an acetylated transferrin, mixed with yeast glucose-6-phosphate isomerase, and adsorbed onto a PVDF membrane. The sequence identified for this sample achieved only an 83% accuracy with respect to positive calls. This decrease from ~95% positive accuracy for previous studies may have been due to the difficulty of the sequence, the PVDF support, or background from the acetylated protein.

Therefore, the following study, ABRF-91SEQ (16), used a sample similar to ABRF-90SEQ. In this study, the 29-residue synthetic peptide was not covalently attached to the transferrin, and the sample (also containing the glucose-6-phosphate isomerase) was supplied in a dried form, which was dissolved before loading onto the sequencer. The positive sequence calling accuracy was also 83% for this study, and thus the PVDF appeared to not

have influenced the ability to call this sequence. The lower sequence calling accuracy in these two studies was probably due to the sample itself, which demonstrates nicely the dependence of the quality of the sequence data on the sample. In this instance, the lower sequence calling accuracy for the ABRF-91SEQ (83%) as compared, for instance, with the STD-1 sample (95%) may well be attributed to the increasing background that would be expected to accompany sequencing of the relatively high molecular weight (~80,000) transferrin protein that was in the ABRF-91SEQ sample. That is, the small fraction of nonspecific cleavage that occurs during each cycle of Edman degradation becomes more significant as the size of the protein (and thus the probability that it will contain an increased number of acid labile peptide bonds) is increased.

Two other challenges faced in amino acid sequencing are identification of posttranslational modifications and identification of difficult residues. ABRF-92SEQ (22) contained two posttranslationally modified amino acids—hydroxyproline and phosphoserine. Hydroxyproline was identified by 45% of the study participants using Edman sequencing alone, whereas only 13.9% were able to identify phosphoserine. Participants were encouraged to use amino acid composition, mass spectrometry, capillary zone electrophoresis (CZE), or other techniques to properly identify these residues. However, very few participants used techniques other than Edman degradation for sequence identification.

ABRF-94SEQ (31) focused on tryptophan and cysteine identification, which tend to be problematic because PTH cysteine is destroyed during Edman chemistry, and PTH tryptophan, recovered in low yields, often elutes on or near an Edman chemistry by-product, diphenylurea. Tryptophan identification in this study was somewhat improved (86%) over several previous studies (~70%), possibly because of the addition of isopropanol to solvent B. The isopropanol helps to separate the diphenylurea from tryptophan. Because cysteine degrades during Edman sequencing to dehydroalanine, it is usually necessary to reduce and alkylate the cysteines in the sample before sequencing. The ABRF-94SEQ study provided participants with seven different techniques that could be used to modify cysteines, with the hope of improving cysteine (Cys) identification. These techniques included cysteine modification to carboxymethyl-Cys, carboxamidomethyl-Cys, pyridylethyl-Cys, Cys-S-propionamide, Cys-S-dimethylpropionamide, *N*-isopropylcarboxyamidomethyl-Cys, and aminopropyl-Cys. Cysteine identification improved from ~53% in the ABRF-93SEQ sample to 82% in the ABRF-94SEQ sample.

In the ABRF-96SEQB (46) study, the cysteines were carboxamidomethylated with iodoacetamide by the committee, the sample was sequenced by the committee, and the PTH HPLC chromatograms were sent out to participating laboratories. In this study, 99% of the cysteines were identified correctly. This suggested that some difficulty was being encountered by facilities in alkylating the sample, willingness to alkylate the sample, or separating the PTH cysteine derivative from the Edman chemistry by-products or other PTH amino acids.

The ABRF-97SEQ study attempted to address this issue by alkylating the cysteines to Cys-S-propionamide using acrylamide before distributing the sample. This test sample contained a 10:2 pmol mix of two synthetic peptides and was the lowest amount of sample provided thus far in these studies. Two controls were also included with the sample, a PTH-Cys-S-propionamide standard and an internal sequencing standard. The PTH-Cys-S-propionamide standard was included so that separation of this derivative could be obtained before sequencing the sample. The cysteine calling accuracy for the ABRF-97SEQ sample was 88% for the cysteine in cycle 4 and 97% for the cysteine in cycle 13. Thus, although the positive sequence calling accuracy for cysteine was not quite as high in the ABRF-97SEQ study (Cys-13 = 97%) as in the ABRF-96SEQ study (Cys-13 = 99%), there was definite improvement over the ABRF-95SEQ study (Cys-15 = 63%). The 17-residue internal sequencing standard peptide (56) was composed of norleucine in cycles 1, 6, 11, and 16 and succinylated lysine in the other positions. Because this peptide is composed of nonnaturally occurring amino acids, it can be mixed with "unknown" sequencing samples. This allows independent and continuous monitoring of the sequencer's performance during all sequencing runs and rapid detection of instrument problems. The low amount of the ABRF-97SEQ sample was expected to be a challenge for the participating laboratories. However, based on the positive sequence calling accuracy for the 50 participating laboratories, the accuracy (92%) was not significantly decreased even for this low amount of sample (see Table 2).

In the ABRF-95SEQ (37) study, the sample was a recombinant 20K protein that contained microheterogeneity (2:1 molar ratio) in positions 6–9. The purpose of this sample was to assess participating laboratories' abilities to recognize microheterogeneity and to determine the length of sequence assignment that is possible. The longest correct sequence observed was 62 residues, including 1 cysteine and 2 tryptophans. The best responses (based on number of residues sequenced and sequence calling accuracy) were obtained on new-generation sequencers, specifically the Hewlett Packard G1005A and Perkin Elmer/Applied Biosystems Division 494 sequencers. Hence, this study suggests that these new instruments may result in improved performance. The overall average positive sequence calling accuracy on this sample was only 78%, which increased to 80% if the four cycles of microheterogeneity were excluded. The majority of the errors occurred in the later cycles and could have been minimized by terminating the sequence earlier, which illustrates another common problem, overcalling of sequences. Basically, the confidence of sequence calling decreases as the sequencing run progresses due to both loss of signal and (especially with relatively high molecular weight proteins) increased background.

Instrument Technological Improvement/Design over the Course of ABRF Studies

The first automated protein sequencer, based on Edman degradation chemistry, was designed by Edman and Begg in 1967 (57) and was manufactured by Beckman. This instrument, a Series 890 Beckman Automated Protein/

Peptide Sequenator (a liquid-phase sequencer), used a spinning cup to perform the coupling and cleavage reactions and was a major step forward in protein sequencing. A solid-phase automated sequencer was developed by Laursen in 1971 (manufactured by Sequemat) in which a polypeptide was covalently attached to derivatized controlled-pore glass beads to minimize sample washout. The PTH conversion and PTH amino acid analysis for both the liquid-phase sequencer and the solid-phase sequencer were done off-line and required several nanomoles of starting material.

A major advance in instrumentation occurred in 1981 with the introduction of the gas-phase sequencer (1), which was made commercially available by Applied Biosystems, Inc. The commercial instrument performed the entire Edman chemistry (including conversion), delivered the coupling base and TFA as gases (greatly reducing background), miniaturized the chemistry plumbing, and, most importantly, increased sensitivity to ~500 pmol of starting material. Samples were analyzed by application to a glass fiber disk coated with a polymer support (polybrene), which helps retain the sample. Shortly after the introduction of the gas-phase sequencer, Applied Biosystems introduced an on-line HPLC to serve as a dedicated PTH analyzer. This change automated the sequencing process, increased the quality of the data, and decreased the amount of starting material required to ~100 pmol. Minor improvements were made by Applied Biosystems such as incorporation of computer control of the instrument, data collection, and the introduction of pulsed-liquid delivery of TFA, which increases repetitive yields. Another sequencer, manufactured by Porton, was one of the first instruments to run multiple consecutive samples in series on a single instrument.

The next major change in Edman chemistry involved the introduction of biphasic columns (half reverse-phase-type material and half hydrophilic-silica-type material) by Hewlett-Packard to analyze samples. This allows samples to remain in the sequencer without use of a polymer support and also allows loading of samples in various salt-containing solutions. In the past few years, Perkin Elmer/Applied Biosystems Division has further miniaturized the instrument design and incorporated in-line sensors to regulate exactly how much reagent is delivered, thus optimizing the chemistry. Perkin Elmer/Applied Biosystems Division recently introduced a sequencer that uses a capillary HPLC, which increases sensitivity to below 1 pmol of starting material, so that PTH amino acid signals of ~50 fmol can be identified. Hence, the overall amount of protein/peptide required for Edman sequencing has decreased over the years from nanomoles (Beckman spinning cup) to subpicomole amounts (Procise cLC) and represents a greater than 10,000-fold increase in sensitivity in 17 years.

Due to these developments, it is currently relatively routine for a laboratory that specializes in high-sensitivity sequencing to obtain sequence from ~5 pmol of sample and, if the sequencer and PTH analyzer are carefully optimized, from 1–1.5 pmol of sample. In this regard, it is important to note that although reverse-phase HPLC-purified peptides typically have initial sequencing yields of ~50%, initial yields for SDS-PAGE purified samples are often less than 25% if electroblotting is not optimized or artificial N-terminal blockage occurs. Assuming 50% electroblotting efficiency from an SDS polyacrylamide gel to a PVDF membrane and an initial sequencing yield of ~15% for a sample that has been subjected to SDS-PAGE, 700 femtomoles of protein would need to be subjected to SDS PAGE to provide the minimum PTH amino acid signal of ~50 fmol required for the cLC instrument. It is important to note that although such a sample might yield 5–10 residues of sequence, a 10-fold greater amount would almost certainly result in higher-quality data and more positively called residues. Thus, it is difficult to respond to the often asked question of how little material is required because the answer depends on the quality and quantity of the sequencing data, the extent of N-terminal blocking that may have occurred during sample preparation, the molecular weight of the protein, the sequence itself, the instrumentation available, and the level of experience of the instrument operator. Current trends in amino acid sequencing include the use of mass spectrometry for obtaining a primary sequence from low and in some instances subpicomole amounts of peptides generated from gel-separated samples (58). In addition, much greater emphasis is being placed on using amino acid composition and mass spectrometric data to identify proteins via protein database searches without resorting to Edman degradation.

INTERNAL PROTEIN SEQUENCING

SHEENAH MISCHE
Rockefeller University
New York, New York

DAVE SPEICHER
Wistar Institute
Philadelphia, Pennsylvania

ULF HELLMAN
Ludwig Institute for Cancer Research
Uppsala, Sweden

KEN WILLIAMS
Yale University
New Haven, Connecticut

One- and two-dimensional PAGE offer powerful methods for the isolation of microgram quantities of individual proteins (59,60). Two-dimensional gels (in which the first dimension is usually isoelectric focusing and the second is SDS PAGE) are generally considered the method of choice for highest-resolution separation of complex protein mixtures at the microgram level and are often used to study the phenotypic-dependent alterations of protein expression in total cellular extracts or enriched cell fractions.

Either one- or two-dimensional PAGE is also often used for final purification of proteins destined for N-terminal sequence analysis. In either case the separated proteins are then electroblotted onto a PVDF membrane, first described as a support for protein microsequencing 10 years ago (55), stained with a dye such as Amido Black or Coomassie Blue, and the band of interest excised and placed in an automated peptide/protein sequencer. However, if the

protein has a blocked (e.g., acetylated) amino terminus (estimated occurrence of this modification is approximately 80% in soluble proteins from Ehrlich Ascites cells [53]), no useful sequence data will result. In addition, there are other reasons that may eliminate the direct N-terminal sequencing approach, including the need for noncontiguous stretches of primary sequence data for DNA cloning and/or the need to generate peptides that could be used for rapid protein identification via an increasing array of mass spectrometric approaches. All these considerations inspired the development of protein digestion methods that can be carried out in gel slices and on PVDF membranes. Although proteolytic digestion and peptide purification have been essential methods for the protein chemist in the study of protein structure and function, novel approaches were needed that conserved the limited quantities of available protein.

In 1987 Aebersold et al. (61) described a method for generating internal sequence from tryptic peptides recovered from proteins electroblotted to nitrocellulose. In rapid succession, methods were reported for in situ digestion of protein bands from Coomassie Blue–stained gels (62,63), proteins bound to PVDF (64), and cationic derivatized PVDF (65). It soon became apparent that microscale generation and isolation of internal peptides were extremely useful in obtaining primary sequence information from N-terminally blocked proteins, and internal sequence analysis radically changed the approach taken in protein sequencing laboratories. Additional refinements that improved each procedure (66–71) resulted in the rapid emergence of internal sequence analysis as a technique for obtaining sequence data from low abundance proteins and has created demands for these analyses that often are associated with unrealistic expectations.

In keeping with its role of education and implementation of new techniques and methodologies, the ABRF created the Internal Protein Sequence Research Committee in 1995 with the purpose of helping member laboratories evaluate or implement internal protein sequence analysis in their laboratories. The study goals that the Internal Protein Sequence Research Committee seeks to achieve are fivefold: (*1*) provide a mechanism for ABRF laboratories to anonymously compare their internal sequencing capabilities with those of other core laboratories; (*2*) provide a reasonable sample and well-proven protocols to facilitate introduction of this technology into laboratories that do not offer these services; (*3*) obtain data that can help determine the relative efficacy of internal sequencing from PVDF blots versus gel samples; (*4*) determine if there are any significant commonalities among the best in-gel and PVDF digestion protocols to help optimize these protocols, and (*5*) compile data obtained by multiple laboratories on the same "unknown" sample, to define more realistic expectations for internal sequencing.

The results of the two ABRF studies completed provide a basis for estimating the amounts of protein that are likely to be required for internal sequence analysis as well as the protein- and methodological-dependent limitations inherent to this approach and their probable impact on the results. Refer to the work of Williams and coworkers (47) for the details of the 1996 study. In brief, each ABRF study involved a microscale enzymatic digestion of protein bound to a membrane support or within an SDS PAGE, Coomassie Blue–stained band, followed by elution of the resultant peptides into a volume of buffer that often ranges from 50 to 200 μL. This peptide mixture can then be subjected directly to mass spectrometry and/or the peptides can be fractionated and a limited number individually sequenced by traditional chemical (Edman) degradation (see "Protein Sequencing" for a brief description of Edman degradation).

Although it is possible to use mass spectrometry to directly sequence individual peptides from complex protein digests, the expense (often more than $400,000) of the necessary equipment and the considerable expertise needed to successfully carry out this approach and to accurately interpret the data has so far limited its availability to a very small fraction of core laboratories. Increasingly, however, a much larger fraction of core laboratories are being equipped with the instrumentation needed to determine the masses (but not the sequences) of many of the peptides present in proteolytic digests. The resulting peptide masses are then searched against peptide mass databases that contain theoretical masses for all peptides expected from digesting known proteins with trypsin and other proteases. Although peptide mass searching provides an extremely rapid approach to identifying known proteins (72–75), which usually requires the use of only a small fraction of the digest, we believe most laboratories recommend that such an identification be confirmed by mass spectrometry or Edman sequencing of at least one peptide.

Microanalytical preparation of peptides from in situ proteolysis for Edman sequence analysis usually employs reverse-phase HPLC using a C18 [$CH_3(CH_2)_{17}$,-] stationary phase that is covalently linked to a silica-based resin packed into a 2.1-mm or smaller internal diameter column that is most often equilibrated with 0.05–0.1% trifluoroacetic acid in water. Peptides are eluted by linear acetonitrile gradients that extend from 0–5% (v:v) in the beginning of the gradient to 80–100% (v:v) at the end of the gradient. One or more peptides are selected for Edman degradation, with or without prescreening by mass spectrometric analysis (76), and the primary sequence is used to search protein and DNA sequence databases (GenBank, Protein Identification Resource [PIR], European Molecular Biology Laboratory [EMBL], Swiss Protein) to determine whether the protein has been previously identified.

As mentioned, two studies of these analytical strategies have been carried out by the ABRF Internal Protein Sequence Research Committee. The 1996 study was a comparative HPLC mapping study in which the participants received 70 pmol each of two recombinant proteins, 28 kDa and 30 kDa, which were identical except for an insert in the 30-kDa protein that generated a 15-residue tryptic peptide unique to the 30-kDa protein. A mixture of these two proteins was separated by one-dimensional SDS PAGE and either provided as Coomassie Blue–stained in-gel samples or electrophoretically transferred to PVDF membranes and stained with Amido Black. A synthetic peptide in which the sequence of the 15-residue unique peptide had been randomized served as an external standard calibrant. This standard was used to normalize data from various laboratories for use of different flow cell path lengths, flow

rates, absorbance wavelengths, and so forth. In-gel and PVDF blanks and representative protocols for each type of sample were provided. Participants who requested samples were asked to (*1*) perform a tryptic digestion of the 28-kDa, 30-kDa, and blank samples and subject all three digests plus the external standard to comparative HPLC; (*2*) identify the target peptide by its presence in the 30-kDa chromatogram, absence in the 28-kDa chromatogram, and elution near that of the external standard peptide; and (*3*) determine the amino acid sequence of the target peptide by traditional Edman degradation and/or mass spectrometric analysis.

For the 1997 Internal Sequencing Study, participating laboratories were asked to perform a tryptic digest on 50 pmol of a known 41-kDa recombinant protein, either in-gel or on PVDF, and to identify the target peptide contained within. A predigested sample of a recombinant protein containing a 23.5-kDa portion of the 41-kDa protein was supplied to assess the ability of the participating laboratory's instrumentation to perform the chromatography required to separate tryptic digestion mixtures. An external peptide standard was supplied that had similar, but not identical, chromatographic behavior to the target peptide. This peptide standard was used by the participants to help identify the target peptide and by the Internal Protein Sequence Research Committee to calculate the relative recoveries of tryptic peptides from each participating laboratory. In-gel and PVDF blanks and representative protocols for each type of sample were provided. In addition to providing a predigested sample, the 1997 study differed further from the 1996 study in that whereas the 1996 target peptide sequence was not in any commonly used database and was thus unique, the 1997 target peptide corresponded to an expected tryptic peptide from a protein whose sequence is included in all commonly used databases. Thus, the 1997 target peptide sequence could be deduced by utilizing observed peptide masses and/or partial Edman or mass spectrometric fragmentation data to search sequence databases.

Figure 2 shows one of the data sets that were returned in the 1997 study. In this instance, the participating laboratory carried out both in situ gel and on-PVDF membrane digests; as evident from this figure, both approaches provided excellent results. Although almost all significant absorbance peaks are unique to the samples, the digests carried out on "blank" gel and PVDF pieces reveal a few absorbance peaks that can be readily eliminated from further study based on their appearance in these control digests. These peaks may result from trypsin autolysis products, buffer-related contaminants, reagents used in SDS PAGE, or residual Coomassie Blue or Amido Black stain.

Several important conclusions follow from the data that resulted from the 1996 and 1997 studies, as summarized in Table 3. First, because hydrolysis and amino acid analysis of representative gel and PVDF bands from the 1997 study indicated that almost threefold more protein was present in the gel bands, the 41-kDa protein used in this study probably has a blotting efficiency of only 19/50 pmol (Table 3), or 38%, which is much lower than the >80% efficiency reported previously (77). Second, although the apparent recovery of the 41-kDa gel band following SDS

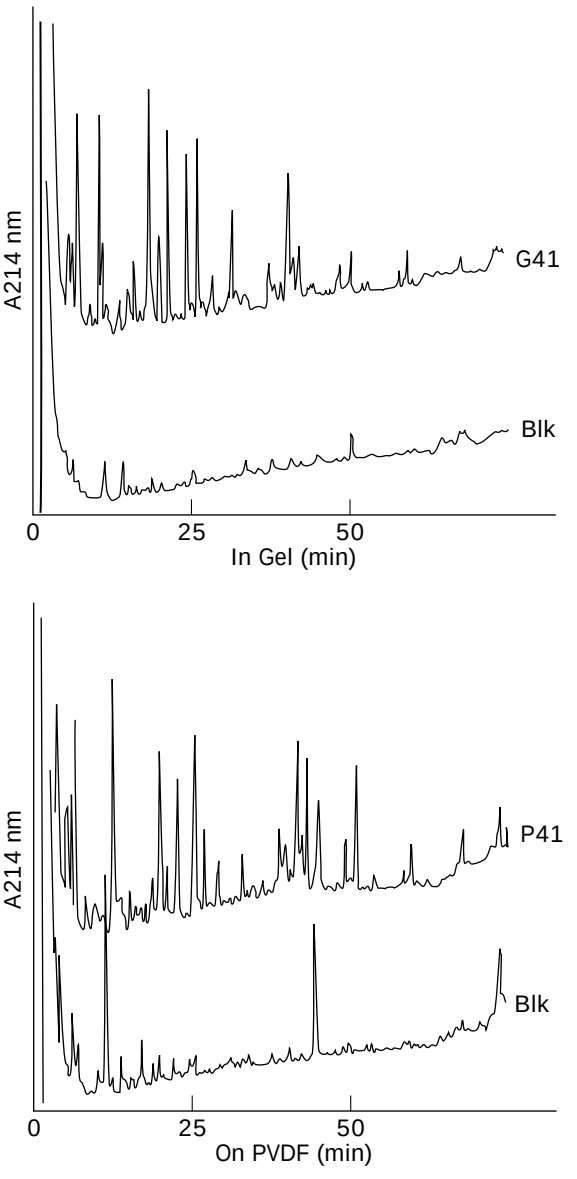


Figure 2. HPLC chromatograms submitted by participant 5 in the 1997 ABRF Internal Sequencing Study. The upper two chromatograms were obtained from tryptic digests of 50 pmol of the 41-kDa protein that were carried out either in the Coomassie Blue–stained gel band (upper panels) or on the Amido Black–stained PVDF membrane (lower panels). In both cases the corresponding digests carried out on blank sections of gels or PVDF membranes are shown in the lower chromatograms. HPLC was carried out at 50 μL/min on a 1- $\times$ 100-mm C18 column that was equilibrated with buffer A (2% acetonitrile, 0.057% trifluoroacetic acid in water) with detection at 214 nm. Peptides were eluted with linear gradients to 40% B (60 min), 60% B (65 min), and finally to 100% B (100 min). Although hydrolysis and amino acid analysis of several gel and PVDF protein bands indicated almost three-fold more protein in the gel samples (Table 3), the similar peak heights suggest a similar final yield of peptides from both the PVDF and gel samples. This suggests the facility that submitted the data may have obtained higher relative peptide yields from its PVDF as opposed to their in-gel digests.

Table 3. Summary of ABRF Internal Sequencing Studies

Parameter	In-Gel		On PVDF	
	1996[a]	1997	1996[a]	1997
Amount of protein subjected to SDS PAGE	70 pmol	50 pmol	70 pmol	50 mol
Amount of protein in stained gel or PVDF band[b]	ND	52 pmol	ND	19 pmol
Median sequencing yield from target peptide[c]	2.3 pmol	8.9 pmol	2.5 pmol	2.9 pmol
Overall target peptide yield[d]	6.4%	24%	7.6%	8.0%
Target peptide recovery from stained gel or PVDF band[e]	ND	23%	ND	21%

ND, no data.
[a]See Ref. 47 for further information regarding assumptions that were made to calculate overall peptide yield and other data from the 1996 study.
[b]As estimated by acid hydrolysis of the gel or PVDF band followed by amino acid analysis.
[c]Yield of valine (4) and of phenylalanine (3) in the target peptides used in the 1996 and 1997 studies, respectively.
[d]Calculated from the estimated initial peptide sequencing yield in cycle 1 (which was based on the median sequencing yield from the target peptide corrected to cycle 1 based on an estimated repetitive yield of 90%) divided by the estimated amount of protein loaded onto the SDS polyacrylamide gel.
[e]Calculated from the estimated initial peptide sequencing yield divided by the amount of protein in the stained gel or PVDF band.

PAGE and staining with Coomassie Blue was high, a previous study (70) demonstrated recoveries as low as 44% following SDS PAGE, Coomassie Blue staining, and washing of gel bands that contained relatively low molecular weight proteins. Hence, in the case of the six proteins (which had a median size of 15 kDa) used in this study (70), the median recovery of protein after washing of the Coomassie Blue–stained gel bands before in-gel digestion was 68% of the amount loaded onto the gel. Taken together, these data demonstrate that some proteins exhibit unusually high losses during electrotransfer to PVDF membranes, whereas other proteins exhibit unusually high losses using in-gel digestion methods.

Comparison of the overall target peptide yields for the two ABRF studies (Table 3) suggests that even when similar in-gel approaches are used, there can be substantial differences in the recovery of an individual peptide from two different proteins. As calculated from the estimated initial amount of peptide that actually sequenced versus the amount of protein subjected to SDS PAGE, the overall target peptide recovery from the 1996 in-gel samples was 6.4% as compared with 24% in 1997. This four-fold difference in peptide recovery makes it difficult to provide a precise answer to the often asked question of how much protein is needed for internal sequencing. Clearly, the answer can vary by at least fourfold, depending on the protein of interest, assuming, of course, that the amount of the protein of interest that was subjected to SDS PAGE is accurately known. Because we believe the most common means of estimating the amount of protein submitted for internal sequencing (comparative staining intensity) is also subject to at least two- to fourfold variability (depending again on the protein of interest), it is quite possible that a state-of-the-art core laboratory might correctly provide a fourfold or greater range in response to the question of how much protein is needed for internal sequencing. Alternatively, they may suggest an amount of protein that is sufficiently large that the internal sequencing project will likely succeed even if the protein in question stains more intensely with Coomassie Blue or Amido Black than an average protein and/or if peptide yields from this protein are below average. In view of the apparent fourfold or greater range in peptide yields following in situ enzymatic cleavage of different proteins, it is also apparent that studies and publications based on results from only a few nonrepresentative proteins may underestimate the amount of an average protein that may be needed to succeed with the same procedure. In this regard, it appears that neither of the two proteins used in the studies summarized in Table 3 are average. That is, the average recovery of 409 peptides that were sequenced from in-gel digests of 191 "unknown" proteins was 12.2% (70). In contrast, the 6.4% target peptide recovery for the 30-kDa protein used in the 1996 ABRF study (Table 3) suggests that this protein may be more difficult than an average protein to digest, and the 24% average in-gel target peptide recovery in the 1997 ABRF study (Table 3) suggests the 41-kDa protein used in this study may digest in a gel matrix much better than does an average protein.

As one means of approaching the question of how much protein is generally needed for internal sequencing, this question was asked of the participants in the 1997 study, which determined that the median response was 50 pmol protein. The results from the 1997 study indeed suggest that a core laboratory that routinely offers internal sequencing will indeed usually succeed if 50 pmol or more protein is subjected to SDS PAGE. Hence, 83% of the core laboratories that attempted to sequence a single peptide derived from the 41-kDa protein distributed in the 1997 ABRF study succeeded in identifying the protein based on database searches of the resulting sequence. Obviously, the percentage of laboratories that could have succeeded in this study is higher than 83%, because many laboratories isolated numerous other peptides from this protein in addition to the target peptide and many of these were probably suitable for sequencing. It appears from the 1997 ABRF Internal Sequencing Study that the 50-pmol estimate is sufficiently robust that even when one step in the procedure does not proceed optimally (such as the much lower than average blotting efficiency of the 41-kDa protein used in the 1997 study), usable internal sequences are nonetheless likely to be derived from the sample.

Another important response in the survey portion of the 1997 ABRF study was the answer to the question of what percentage of received samples are already in the databases. Because the median response was 75%, it is clear that the overall goals of internal protein sequencing will increasingly shift to protein identification (i.e., proteome

research) and to the elucidation of posttranslational modifications that usually cannot be predicted a priori from the ever-increasing wealth of cDNA sequences in nonredundant databases and that mass spectrometry will play an ever-increasing role in this endeavor.

Finally, the answer to the question of which approach (in-gel or on-PVDF digestion) is best for internal protein sequencing is that the two methods appear to be quite comparable. That is, when peptide recoveries are calculated from the Edman sequencing yields divided by the amount of protein in the stained gel or PVDF band (as opposed to the amount that was loaded onto the gel), the target peptide recoveries for both PVDF and gel samples were just above 20% in the 1997 study (Table 3). The optimal method for a specific project depends on the protein of interest and the personal preferences and level of expertise that the core laboratory workers to whom the sample will be submitted has with each of these approaches. Based on two ABRF studies, each of which involved about 40 core laboratories, we conclude that if a 50-pmol protein sample is subjected to SDS PAGE and then prepared according to the protocol recommended by the core laboratory to which the sample will be submitted, the probability of obtaining internal peptide sequences is likely to be much higher than 83%.

PEPTIDE SYNTHESIS

Lynda F. Bonewald
University of Texas Health Science Center
San Antonio, Texas

Ruth Hogue Angeletti
Albert Einstein College of Medicine
Bronx, New York

The goal of peptide resource laboratories is not only to supply the biomedical researcher with the high-quality peptides critical for scientific research at an affordable cost but also to actively research, test, and design new investigative peptides. To meet this challenge, the peptide chemist must be aware of, be able to learn, and be capable of instituting the latest techniques of peptide chemistry and evaluation. The ABRF Peptide Synthesis Research Committee was established to assist personnel of its member facilities in attaining this goal.

Use of peptides is diverse, ranging from use as antigens for the production of antibodies and as enzyme substrates or inhibitors to use for structure–function studies, for screening purposes, or as agonists or antagonists in biological assays and systems. For these applications, either large amounts of a small number of peptides or small amounts of a large number of peptides are usually necessary. More peptides with modifications are required to address important scientific questions such as the roles of phosphorylation and dephosphorylation in cell signaling. Peptide chemists are in the process of improving and developing new strategies for synthesis and evaluation of modified peptides. Examples of other modifications include biotinylated and fluoresceinated peptides for detection and quantitative assays, modifications that allow cross-linking on exposure to ultraviolet light, cyclized peptides for enhanced stability leading to enhanced biological activity, and a variety of other modifications of both the amino- and carboxy-termini and modifications of amino acid side chains.

To achieve the goal of supplying necessary, critical peptides for research, the following two components are essential: skilled, knowledgeable personnel and state-of-the-art instrumentation. Peptide synthesis laboratory personnel must be flexible and learn and institute new technologies quickly. Automated instrumentation belies the actual skill and experience required for peptide laboratory personnel. Peptide synthesis is not just "button pushing." Each peptide can vary considerably in the amount of design and planning required for synthesis. Some peptides require only routine synthesis, cleavage, desalting and/or purification, and analysis, whereas others may generate unique challenges at any or all phases of production. Many of these challenges lead to research projects involving collaboration between the peptide chemist and the biomedical researcher.

The laboratory must contain or have access to state-of-the-art instrumentation. The standard peptide synthesis instrument that produces a single peptide at one time is becoming inefficient except in cases of producing longer peptides. New instruments are being designed for multiple arrays and for more stringent chemistries and synthesis conditions. New instrumentation under development that can produce multiple peptides at one time with on-line cleavage is becoming state of the art. Though not perfected, these instruments can meet the demand for the production of large numbers of peptides. With the ability to produce large numbers of peptides comes the need to desalt and purify peptides efficiently and in a timely manner. HPLC systems are being developed to provide autoinjection and autocollection of either large numbers or large amounts of sample. As discussed later, the optimal means to ensure that the correct peptide has been made is by mass spectrometric analysis.

The Peptide Synthesis Research Committee's raison d'être has been to evaluate the quality of peptides produced by its member laboratories and to introduce and analyze new chemistries and technologies in solid-phase peptide synthesis. Since its inception in 1990 the Peptide Synthesis Research Committee has addressed several aspects of peptide production. The first two studies (17,20) were designed to determine and address potential problems in synthesis, cleavage, and side-chain deprotection. For peptide analysis, mass spectrometry with an accuracy of one atomic mass unit was found to be essential. HPLC and capillary electrophoresis are useful for quantitation but can be misleading. A single peak on a chromatogram can represent an incorrect peptide. HPLC is most useful for purification and collection. AAA is useful for quantitation but cannot identify specific problems, only that a problem may or does exist. Amino acid sequence analysis is costly, time-consuming, and cannot aid in identifying protected amino acids. Mass spectrometry proved to be most useful in identifying problems in synthesis, cleavage and deprotection, and identification of correct peptides.

Because cleavage from the resin was shown to be more problematic than synthesis or side-chain deprotection in

the first and second study, this step was specifically addressed in the third study (27). However, no correlations could be found between common side reactions occurring during cleavage and specific cleavage methods. The human element may play a critical role in the quality of peptide production.

The fourth study (32) was undertaken to determine if core laboratories had improved from the first study; therefore, the exact same peptide was requested for analysis as was requested in the first study. The submitted peptides were found to have increased considerably in quality from the first to the fourth study (in the first study, 78% of submitted samples had less than 25% of the correct product in contrast to the fourth study, in which 93% contained greater than 25% of the correct product). The sixth study (45) showed even greater improvement (almost 95% of the submitted samples contained greater than 70% correct product).

The fifth study (36) addressed problems in disulfide bond formation and introduced cyclization to the member laboratories. It was found that peptide cyclization can be performed on a routine basis with high yield. The sixth study (45) was undertaken to examine potential problems in racemization. It was found that histidine racemization may not be a major problem, but researchers should evaluate their peptide application beforehand to determine whether racemization is important. If even low levels of racemization could pose a potential problem, the racemized peptide can be separated from the correct peptide by HPLC techniques.

In the seventh study, the Peptide Synthesis Research Committee introduced and analyzed phosphorylated peptides requested from its member laboratories. This study was initiated because peptide modifications are becoming necessary for a number of applications. One very important modification is the role of phosphorylation in cell signaling and transduction. This study should determine whether it is feasible to synthesize posttranslational peptide modifications, such as tyrosine phosphorylations.

Other important observations and conclusions drawn from the studies conducted by the Peptide Synthesis Research Committee include the following: A shift in the type of chemistry used in solid-phase peptide synthesis has occurred from Boc to Fmoc chemistry (from 50% in 1991 [17] to 98% in 1996 [45]). This is probably due to Fmoc chemistry being safer, more flexible, and less problematic. One observation that has been repeated with each year's study is the unexpected occurrence of human errors in peptide synthesis. Each year's study contains a few surprises. For example, one year a submitted test peptide was synthesized in the wrong direction. It is not known whether these incorrect peptides are mainly due to inexperience or to human error, but they should convince all concerned that postsynthesis peptide analysis is critical. The best combination of techniques to ensure the correct product is HPLC (see Fig. 3) and mass spectrometric analysis (see Refs. 17, 20, 27, 32, 36, 38, 44, and 45).

The peptide laboratory must be flexible and capable of constantly changing to meet the needs of biomedical research. Academic peptide laboratories cannot compete with commercial peptide facilities without support from

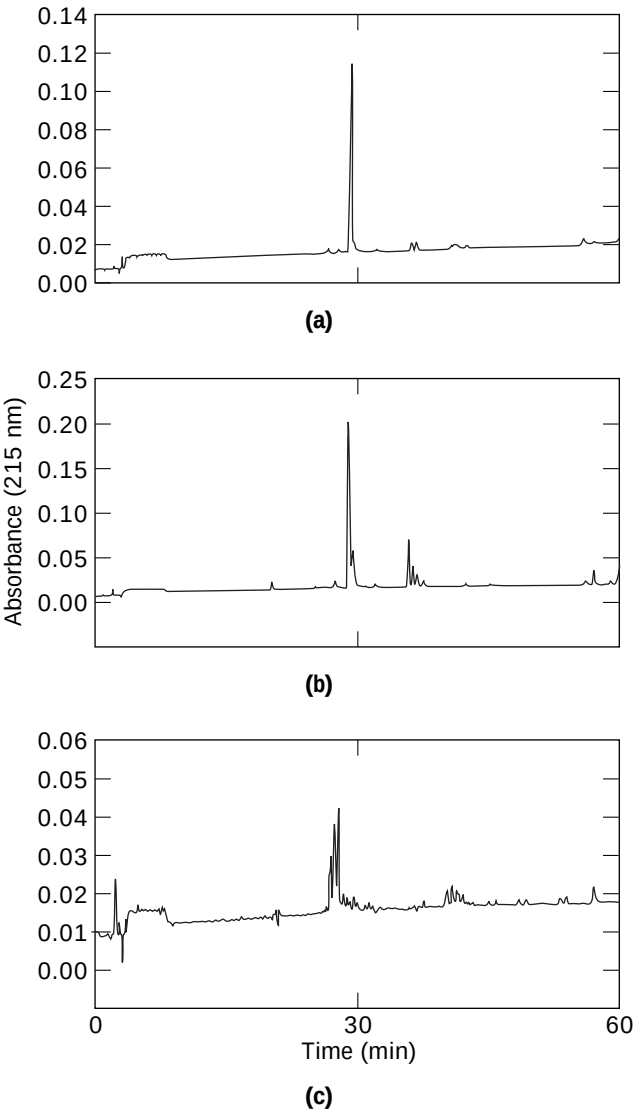


Figure 3. Analytical reverse-phase HPLC chromatograms from three peptide samples submitted in the ABRF 1995 Peptide Synthesis Study. Biomedical researchers should determine whether their studies require purified (greater than 70% correct peptide) or highly purified (greater than 95% correct) peptides. If the study does not require highly purified material, then considerable time and cost can be saved. Usually 70% correct peptide can be obtained with a good synthesis followed by a desalting step. To obtain highly purified peptides, HPLC is usually necessary. Examples of the linear peptide used from the ABRF 1995 study are shown (36). **(a)** Greater than 96% pure peptide; **(b)** good peptide, greater than 76% pure; **(c)** poor synthesis. The major peak by HPLC from any synthesis should always be analyzed by mass spectrometry to ensure that the peak represents the correct peptide.

their institutions. The rationale for institutional support is the availability of advice, consultation, potential collaboration, potential for modifications and chemistries not readily available through commercial ventures, and reduced cost to the investigator. The common goal of the peptide chemist and biomedical researcher should be scientific endeavor and not profit or monetary gain. The future of

peptide laboratories will definitely include new chemistries, new peptide and amino acid modifications, the synthesis of longer, larger peptides, and new instrumentation. Perhaps in the future, combinatorial libraries may be generated for basic research or such exotic chemistries as peptide nucleic acids will be utilized. The ABRF Peptide Synthesis Research Committee will continue to assist its members with these challenging and changing goals.

CARBOHYDRATE ANALYSIS

R. REID TOWNSEND
University of California–San Francisco
San Francisco, California

Determining the monosaccharide content of glycoproteins (*1*) provides the basis for further structural studies, (*2*) suggests classes of oligosaccharide chains, (*3*) provides unequivocal evidence that the polypeptide chain is glycosylated, and (*4*) may serve as a measure of production consistency for therapeutic glycoproteins. The presence of GalNAc (as $GalNH_2$ after de-*N*-acetylation from acid hydrolysis) suggests mucin-type *o*-linked structures. Only Man and GlcNAc suggest oligomannosidic-type structures; GlcNAc, Man, Gal, and Fuc suggest lactosamine-type oligosaccharides; and Glc and Gal suggest glycosylation of hydroxylysine and hydroxyproline residues. Historically, monosaccharide analysis has been a three step process of (*1*) release of residues from oligosaccharide structures and the polypeptide backbone (*2*) derivatization for gas chromatography (GC) or HPLC, and (*3*) chromatography with quantitative detection. It is well recognized that breakage of glycosidic and protein–sugar bonds by acid hydrolysis or methanolysis is often incomplete and results in some destruction of monosaccharide residues, although many studies have been performed to optimize the yield of released monosaccharides. For GC, volatile monosaccharide derivatives must be prepared (78,79). For HPLC, derivatization is used to increase hydrophobicity for reverse-phase chromatography and to increase the sensitivity of detection (for review, see Ref. 80). The chemical diversity and lability among neutral, amino, and anionic classes of monosaccharides necessitates different release and derivatization methods to analyze all of the sugars from a glycoprotein. The detection of carbohydrates on gold electrodes in a low-volume flow cell after separation using high-pH anion-exchange chromatography (HPAEC) obviates the need for derivatization of neutral, amino, or anionic monosaccharides in samples that contain ~100 pmol of individual monosaccharides (for review, see Ref. 81). The commercialization of HPAEC with pulsed amperometric detection (PAD) has resulted in a wider availability of monosaccharide analysis using a single method. Most recently, fluorophore-based methods, which are available in a kit format from vendors, should further increase the general availability of monosaccharide analysis. Finally, the oligosaccharides on a single protein occur as an array of structures whose proportions are difficult to quantify, and the type and proportions of structures can vary from lot to lot, even after isolation from the same source.

In 1995 the ABRF Carbohydrate Analysis Research Committee (CARC) initiated a study to determine the ability of research laboratories and core facilities to identify and quantify carbohydrates (43), which are commonly found in N- and O-linked oligosaccharides of glycoproteins from serum, cell membranes, and recombinant expression systems. Test sample A was ultrapure grade bovine serum albumin from Calbiochem (San Diego, California), and test sample B was bovine fetuin, purified by the Spiro method from Gibco (Grand Island, New York). The two test samples were lyophilized from water and quantified by hydrolysis followed by amino acid analysis. Based on the average of the values for alanine and valine, test sample B contained 25 nmol fetuin, which was chosen because it is one of the most well-characterized glycoproteins (82–88). All participating laboratories ($n = 29$) correctly identified which sample was glycosylated. In 14 cases the glycosylated sample was identified by AAA or colorimetric assays. The amino acid compositions were analyzed further using PROPSEARCH (89), which can be accessed via the Web site *(www.embl-heidelberg.de/aaa.html)* to identify the test samples as bovine serum albumin and bovine serum fetuin. Monosaccharides were quantified using HPAEC PAD (14 laboratories), HPLC after precolumn derivatization with 1-phenyl-3-methyl-5-pyrazolone (90) (one laboratory), gas chromatography and mass spectrometry (four laboratories), or gel electrophoresis after derivatization with 2-aminoacridone (one laboratory) (91).

Analysis of the glycosylated test sample by 4 CARC laboratories and 10 participating laboratories using HPAEC PAD gave the following results (in mol/mol of protein): GalN, 2.3 ± 39%; GlcN, 16 ± 23%; Gal, 12 ± 44%; and Man, 9 ± 55%. (The error is expressed as percent standard

Table 4. Monosaccharide Composition of Bovine Fetuin as Determined by HPLC (Mole Sugar per Mole Protein)

Monosaccharides	Literature[a] (1984–1991)	CARC[b] (1995–1996) (%)	External laboratories (%)
GalNAc (3)[c]	1.2–3.5	2.7 ± 20[d]	2.2 ± 41
Man (9)	6.4–8.3	9.6 ± 15	8.0 ± 50
GlcNAc (15)	11.5–16.4	17.2 ± 9	15.2 ± 27
Gal (12)	9.3–13.0	12.4 ± 24	11.6 ± 43
Neu5Ac (12)	7.2	14.8 ± 9.5	15.2 ± 27

[a]Reviewed in Ref. 78.
[b]Values after iterative analyses by CARC.
[c]Values in parentheses are theoretical and estimated.
[d]Based on the amino acid analysis of the bovine fetuin test sample (25 nmol from an average of alanine and valine content). The percent standard deviation is given in terms of σ_{n-1}.

deviation.) The results from GC analysis (alditol acetate and trimethylsilyl derivatives) from four laboratories also showed large errors: GalNAc, 1.7 ± 43%; GlcNAc, 9.2 ± 36%; Gal, 13.6 ± 35%; and Man, 9.6 ± 26%. The analysis of *N*-acetylneuraminic acid by both the CARC and participating laboratories gave a lower percent standard deviation by both HPLC (15.2 ± 17%) and GC (13.2 ± 23%). Finding these errors too large, CARC undertook a systematic internal study to identify sources of error in monosaccharide analysis by HPAEC PAD. First, it was established that interlaboratory variability was significantly greater than the analysis of the same bovine fetuin sample when performed on a single instrument over a 6-month period: GalN, 2.4 ± 8.8%; GlcN, 16.4 ± 8.6%; Gal, 12.8 ± 15% and Man, 7.8 ± 4.6%. Because all studies were performed with external standards, monosaccharide standards from four CARC laboratories were analyzed in the same laboratory. The percent standard deviation was no greater than 15% and for most monosaccharides was less than 5%, indicating that standard preparation was not a major contributor to the error among CARC laboratories. Based on literature reports and experience within CARC laboratories, standard hydrolysis conditions were adopted for neutral and amino sugars (2 M trifluoroacetic acid for 3 h at 100 °C) and sialic acids (2 M acetic acid for 3 h at 80 °C). Because monosaccharides can be destroyed under acid hydrolysis conditions used for their release, CARC used external standards that had been hydrolyzed under the mentioned conditions. Untreated external standards and standards that had been subjected to the same hydrolysis conditions as the glycoprotein were then compared. Unexpectedly, the amino sugar electrochemical responses in the hydrolyzed standards were increased as much as twofold, suggesting that the higher values observed for amino sugar determination, both in some CARC and participating laboratories, were due to this phenomenon. The lower peak areas were attributed to either selective loss of amino sugars through the injection path or reaction of free amino groups on the standard nonacetylated sugars with the aldehyde functions of the monosaccharides in both the standard mixture and hydrolyzates.

To control further for variability associated with hydrolysis of samples, a single hydrolyzate of the bovine fetuin sample was prepared in one laboratory (in a single vessel) and then distributed to the other CARC laboratories for analysis by HPAEC PAD. Hydrolyzed, external monosaccharide standards were also used. The following reduction in percent standard deviation was observed from four CARC laboratories: GalN, 2.7 ± 20%; GlcN, 17.2 ± 9.1%; Gal, 12.4 ± 22%; and Man, 9.6 ± 15%. In summary, by adopting standard hydrolysis conditions, ensuring that the external standards prepared in different laboratories gave similar responses and using hydrolyzed external standards, the interlaboratory error of HPAEC PAD analysis was significantly improved; however, interlaboratory error remained greater than those observed using a single instrument.

Table 4 compares the monosaccharide compositions of bovine fetuin from literature values with those obtained by CARC after the iterative analyses discussed earlier and with those from participating laboratories, all using HPLC. Overall, the mean values obtained by both the CARC laboratories and the participating laboratories vary by no more than two residues; however, the error range associated with the external study was significantly greater. The theoretical values (indicated in parentheses in Table 4) were estimated from the oligosaccharide structures and the number of modified sites on the protein. Bovine fetuin has three N-linked sites that contain biantennary (84,86) and triantennary structures (82,87,88; Fig. 4). In addition, at least three O-linked sites, occupied with sialylated di- and tetrasaccharides (83,85; Fig. 4), are present (83). The value of three GalNAc residues assumes complete site occupancy of three O-linked sites; however, it is well recognized that glycosylation at serine and threonine residues is frequently not stoichiometric. An approximately threefold range of values has been reported in the literature and by the external laboratories participating in the 1995–1996 study for the GalNAc composition of fetuin (Table 4). The CARC laboratories found an average value of 2.7, with a range of 2–3 residues.

If all three N-linked sites are fully occupied, nine Man residues would be present in fetuin. The value obtained by CARC approached this value, with an average of eight reported in the external study and a larger error. The theoretical value of 15 for GlcNAc accounts only for the composition of three triantennary structures; however, the O-linked tetrasaccharide also contains GlcNAc (Fig. 4). Whether the value of 17 residues from the CARC analyses reflects a significant amount of the tetrasaccharide in fetuin will require further studies. The 12 theoretical Gal residues are estimated from three triantennary structures and three O-linked disaccharides, and the 1995–1996 results approach this value (Table 4). The determination does not indicate the presence of biantennary structures which have one less Gal residue, nor the tetrasaccharide, which has one more Gal residue. Complete sialylation of 12 Gal residues would result in 12 Neu5Ac residues. The greater value for Neu5Ac indicates complete sialylation and is also consistent with the presence of the known tetrasialylated triantennary structures in bovine fetuin (87,88).

Overall, these results show that quantitative monosaccharide analysis in combination with amino acid analysis, to determine the amount of protein, gives molar ratios of sugars to protein that are in agreement (within one to two residues) with the oligosaccharide content. However, these studies have also shown that there is a significantly larger error associated with analyzing the same hydrolyzate on different instruments (43). Future studies will focus on the reanalysis of the same test sample to identify factors that are related to differences in instrument configurations and interlaboratory analyses.

NUCLEIC ACIDS

C.W. NAEVE
St. Jude Children's Research Hospital
Memphis, Tennessee

P.S. ADAMS
Trudeau Institute
Saranac Lake, New York

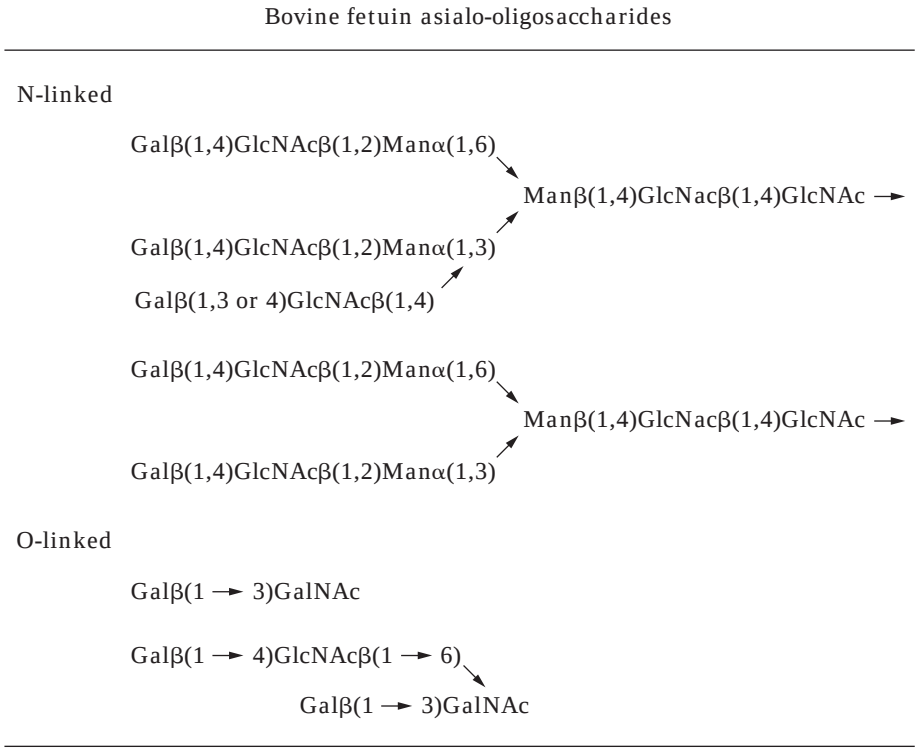


Figure 4. Structures of bovine fetuin asialo-oligosaccharides.

The dramatic growth in the use of molecular biology methods in biomedical research over the past 15 years has seen a corresponding increase in the number of core laboratories engaged in providing essential DNA chemistry services, especially automated DNA synthesis and DNA sequencing. To effectively utilize these resources, it is important for the investigator to have realistic expectations regarding the current capabilities of this technology. To determine the level of technology that can be routinely reached, provide a means of self-evaluation, and gain valuable comparative data beneficial to the field, the ABRF regularly surveys member laboratories and tests facility capabilities through distribution of "unknown" samples for analysis. Two ABRF committees have been established to evaluate nucleic acid chemistries. The Nucleic Acids Committee has published three studies since its inception in 1991, the first designed to assess the type of nucleic acid services provided by core facilities (24), the second designed to assess the accuracy of automated DNA sequencing services (29), and the most recent designed to determine the quality of synthetic oligonucleotides produced by member laboratories (33). The tremendous increase in use of automated DNA sequencing led to the establishment of a second committee in 1995 whose focus is specifically DNA sequencing. The DNA Sequence Research Committee recently completed a general survey and a study designed to assess member laboratories' capabilities in sequencing a moderately difficult GC-rich template (39; *http://www.dfci.harvard.edu/abrfdnaseq*). The results of the most recent of these studies are summarized in the following and provide an overview of the current capabilities of core laboratories that provide DNA synthesis and DNA sequencing services.

ABRF DNA Synthesis Studies

The ABRF has conducted two studies regarding DNA synthesis in core laboratories; the first in late 1992 was a survey designed to obtain information on instrumentation used, favored chemistries, modifications, charge-back fees, and so forth (24). The most recent study was conducted in April 1995 when member laboratories were again asked to respond to a brief survey but were also asked to provide the crude trityl-off, deprotected, and lyophilized products from a 25-base and 50-base 0.2-mmol-scale DNA synthesis for analysis by the Nucleic Acids Committee (33). Seventy-one laboratories participated in the survey and submitted 208 oligonucleotides prepared on a distribution of instruments from various vendors that was essentially unchanged from the 1992 survey. New developments in instrumentation (higher capacity or throughput, lower reagent consumption), lower reagent prices, and increasing competition from commercial providers has led to substantial reductions in charge-back fees; the median rates for oligonucleotide synthesis decreased 20–30% from the previous study to $1.90/base for the 40-nmol-scale synthesis (the actual total cost of synthesis is estimated to be $3.80/base, assuming an institutional subsidy of 50%). Median annual output from these facilities had increased more than 20% to 1,440 oligonucleotides per year. Although new developments in DNA synthesis chemistry have occurred (e.g., fast deblocking reagents), most respondents (80%) continued to use phosphoramidites with the conventional *N*-benzoyl and *N*-isobutyryl protection groups and controlled pore glass (CPG) supports, which suggests that these traditional reagents continue to provide the highest overall satisfaction.

The synthetic oligonucleotides submitted to the committee were analyzed by capillary electrophoresis (CE) to estimate overall yield and purity and were used as primers in DNA sequencing reactions as a functional measure of quality. Overall yield by CE of most of the oligonucleotides was very good. Purity ranged from 14 to 95%, with an average of 74% for 25-mers (Fig. 5) and from 1 to 92% with an average of 50% for 50-mer (data not shown). Average median coupling efficiencies of 98.9% for 25-mers and 98.7% for 50-mers were calculated. Coupling efficiencies of >98% are usually guaranteed by reagent and instrument manufacturers, and this number has been used previously as a minimum standard. Thus, the vast majority (approximately 85%) of crude oligonucleotides submitted for analysis exceeded the minimum standard. To put coupling efficiencies into perspective, it should be noted that with a coupling efficiency of 98%, the maximum yield of desired product for a 25-mer, 50-mer, and 100-mer synthesis would be 61%, 37%, and 14%, respectively, with the remaining material in each preparation resulting from $n - 1$, $n - 2$, $n - 3$, . . . species. Because most facilities perform quality assurance measures and offer purification services, many of the substandard oligonucleotides would have been detected or eliminated before distribution.

As an additional measure of quality, the crude 25-mers were tested for their ability to function as primers in automated DNA sequencing reactions. A total of 103 of 104 25 mer submitted generated sequence data, and most of the data were remarkably good. Four primers produced perfect sequences over a 400-base window downstream from the primer target; half generated fewer than four errors (including insertions, deletions, substitutions, and ambiguities), and only those in the bottom quartile generated more than seven errors. The number of sequencing errors generated by each 25-base primer was compared with the CE-measured purity of that primer (see Fig. 5). Surprisingly, primer purity does not strictly correlate with sequencing fidelity. Primers that were >70% pure generally produced sequence data with an accuracy >97.5% over the 400-base window. Primers with a purity <70% yielded mixed sequencing results, although 18 of 25 primers that were only 54–70% pure still generated good (<10 errors) sequence data. Overall, more than 85% of the crude 25-mer oligonucleotides yielded unedited sequence data that were >97.5% error free over the 400-base test window. Because unprocessed postsynthesis samples were requested, these results also suggest that purification or desalting provides only marginal benefit for the synthesis in which the average coupling efficiency exceeds 98.5%.

ABRF DNA Sequencing Studies

The ABRF has conducted two formal studies that focus exclusively on DNA sequencing services. A 1995 survey, conducted by the Nucleic Acids Committee (29), was designed to determine the accuracy of sequences obtained using dye-primer or dye-terminator chemistry and *Taq* polymerase. A standard double-strand DNA template was mailed to 80 member laboratories; 44 responded by submitting 83 DNA sequence data sets for analysis. These were grouped by chemistry utilized (dye terminator or dye primer) and whether postrun editing was employed (edited or unedited). The results showed that, using *Taq* polymerase, dye-primer chemistry provided the longest and most

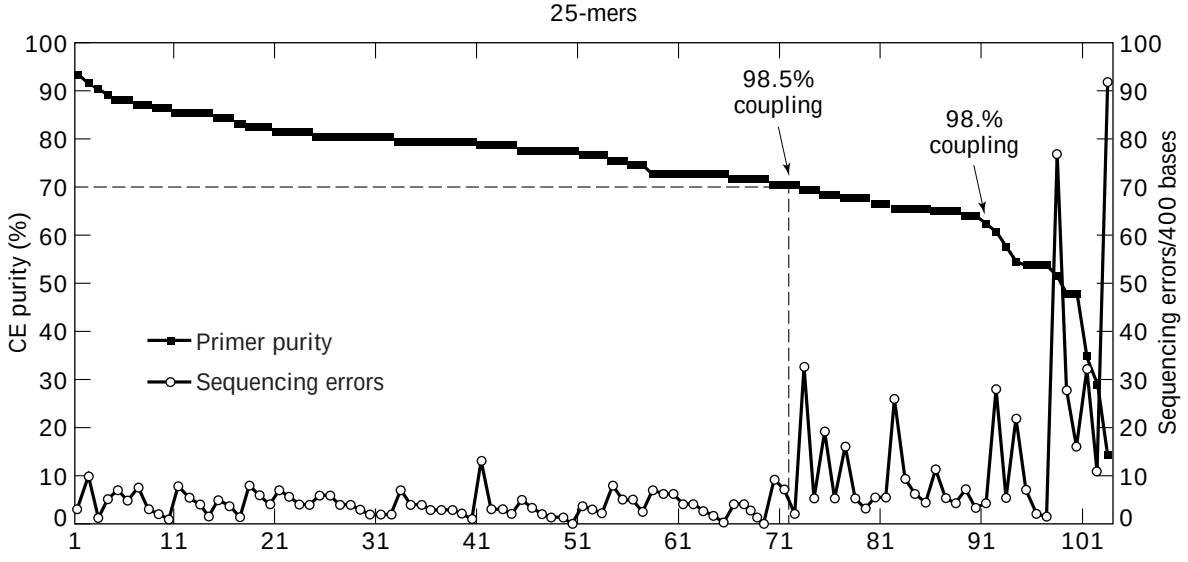


Figure 5. Oligonucleotide purity and performance as sequencing primers. Purity and sequence performance of 25 mer. The capillary electrophoresis (CE) determined purity (■) of the 25-mer oligonucleotides is illustrated in descending order from highest (95%) to lowest (14%) purity. Also plotted is the number of sequencing errors (○) (deletions, insertions, miscalls, and ambiguities) found in the first 400 bases of unedited, automatically called sequence when the 25 mer were used as sequencing primers. CE purity of 70%, representing approximately 98.5% average coupling efficiency, is indicated. The industry standard 98% average coupling efficiency is indicated.

accurate sequence. Edited dye-primer data were >95% accurate out to 400–450 base pairs, whereas edited dye-terminator data were accurate at this level only to 300–350 base pairs. However, most laboratories in this sampling (75%) prefer the dye-terminator protocol, presumably because of its versatility and convenience. Laboratories that manually edit dye-primer data were able to obtain an additional 100 base pairs of data. Surprisingly, editing of dye-terminator data did not increase the length of read for this template.

The DNA Sequence Research Committee's 1996 survey was designed to assess member laboratories' ability to sequence a moderately difficult double-stranded plasmid DNA sample containing a GC-rich insert (69% GC over 926 base pairs) while studying the effectiveness of protocols developed for GC-rich samples, changes in sequencing chemistry, performance according to sequencing hardware, the type of products used, and the effect of editing. In addition, a questionnaire was included to allow comparisons with previous studies (6,7,24). Responses to the questionnaire revealed that the number of facilities performing automated DNA sequencing and the number of samples sequenced per year continues to grow. In 1989, 12 laboratories participated in the survey and reported sequencing an average of 240 templates per year. In comparison, in 1993, 37 laboratories reported sequencing an average of 2,100 templates per year, and the 1996 survey showed 53 laboratories sequencing an average of 4,560 templates per year, of which 76% were double-stranded plasmids and 21% were polymerase chain reaction (PCR) products. Fluorescent techniques were used by all contributing facilities in 1996, dye-terminator chemistry was used by 82%, and the new TaqFS (Perkin Elmer/Applied Biosystems Division [PE/ABD]) enzyme was used by two-thirds. The 53 laboratories that participated in the 1996 study generated 78 sequencing data sets for analysis. The majority of respondents submitted data obtained using dye-terminator sequencing chemistry and the enzyme TaqFS. The sequences were ranked according to length of entirely correct sequence. Keeping in mind that this was a difficult sequence and a routine run was requested, the average length of correct unedited sequence for all machines was 284 bases and the longest was 615 bases. Gel length, manual data review, and sequencing chemistry all influenced sequence accuracy and length. The average correct unedited length of read for PE/ABD Stretch instruments equipped with 36- or 48-cm plates was 362 bases compared with an average of 232 bases for all other instruments. In several direct comparisons, it was shown that an average of 157 more bases of correct sequence were obtained by manual editing for this difficult template. Some facilities submitted data plus or minus dimethylsulfoxide (DMSO) in the sequencing reactions. In most cases DMSO increased the length of correct sequence; however, many of the top-ranked responses did not use DMSO. The 1997 study focused on elucidating the benefits of DMSO, altered thermocycling procedures, and editing when sequencing difficult GC rich templates. A summary of the average error-free length of read depicted by machine, well to read (length of the gel), and a comparison of edited versus unedited sequence can be seen in Figure 6.

In summary, studies conducted by the ABRF reveal that the vast majority of DNA core facilities that offer DNA synthesis services produce synthetic oligonucleotides of high quality. The typical facility in 1996 utilized cyanoethyl phosphoramidite chemistry to produce more than 1,400 synthetic oligonucleotides per year. The average charge-back fee was $1.90/base for 40-nmol scale (an estimated total cost of $3.80/base assuming an institutional subsidy of 50%). Surprisingly, even crude primers of relatively low purity were found to function well when used in automated DNA sequencing. Primers with a purity of >70% were found to function well as DNA sequencing primers without any further purification or even desalting, whereas primers with purity <70% were less likely to function well in DNA sequencing reactions. The average DNA core laboratory that offered automated DNA sequencing services in 1996 typically used dye-terminator chemistry and TaqFS polymerase. A median of 4,560 templates per year were sequenced for an average fee of $29.50/template (estimated total cost $59/template assuming an institutional subsidy of 50%). Minimum template requirements were 0.5–1.0 μg DNA, from which the user could expect to obtain 400–800 base pairs depending on the length of the gel. A recent study (39) has shown that postrun editing of the raw data can substantially increase the sequence accuracy from difficult GC-rich templates when using the most popular chemistry and enzyme.

MASS SPECTROMETRY

GARY HATHAWAY
California Institute of Technology
Pasadena, California
KRISTINE M. SWIDEREK
ZymoGenetics
Seattle, Washington

Often regarded as a supportive or adjunct technique by biomolecular resource laboratories, and initially employed by them largely as a procedure for the quality control of synthetic peptides (17), mass spectrometry is rapidly becoming a major analytical tool for biopolymer analysis (in several instances it has become the only available tool for their structural analysis). With the advent of electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) methods, mass spectrometry has become a popular and powerful method for analyzing large biopolymers such as peptides, proteins, and both native and modified oligonucleotides (92–97). No longer an exercise reserved for specialty laboratories, the use of mass spectrometry for sequencing picomole quantities of peptides and proteins is becoming widespread (see test sample results later in this article), with a few laboratories able to achieve results in the femtomole range (98,99). Used in conjunction with automation and databases, such sequence and fragmentation mass data can provide identification of biologically interesting proteins isolated by two-dimensional gel electrophoresis (100–102). This technique, when applied to the study of cellular response to a given stimulus as reflected by its entire complement of proteins

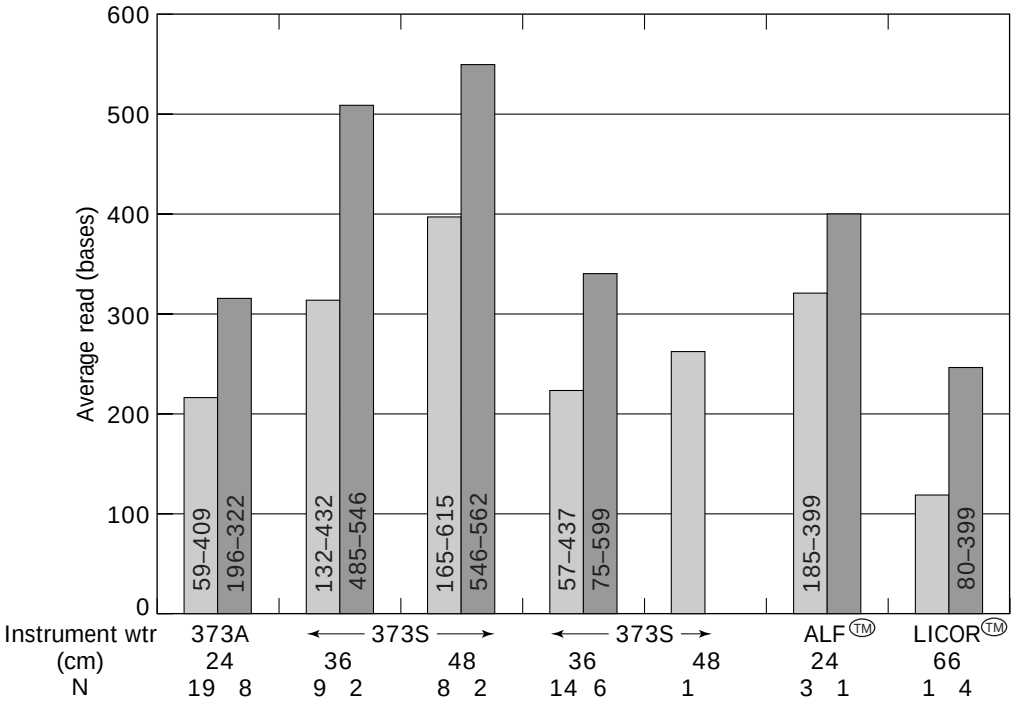


Figure 6. A summary of the average error-free length of read depicted by machine (instrument) and well to read (WTR; length of the gel in centimeters). Values for unedited sequences are open bars, and values for edited sequences are shaded bars. Ranges and the number of samples per group *(N)* are shown.

at a particular point in time, has come to be known as proteome science (103) and is an area of intense effort. If compared with the conventional N-terminal sequence analysis by Edman degradation, the direct generation of peptide fragments and their sequence analysis within coupled mass analyzers (MS/MS) is faster and more sensitive (although data interpretation can sometimes take longer) (104). It has the advantage that modified amino acids can be identified more easily and that a blocked N-terminus does not present a problem for a successful analysis. On the other hand, amino acids of equal mass (such as isoleucine and leucine) generally cannot be distinguished, and glutamine and lysine, which differ by only 0.036 mass units, are determined only with difficulty or after chemical modification of lysine. Another disadvantage is that sequence analysis utilizing mass spectrometric techniques is limited to relatively short peptides. Peptides longer than about 25 amino acids can be sequenced with mass spectrometry, but only if special techniques (which may not be generally available in core laboratories) such as ladder formation through cyclic chemical degradation are employed or by the use of exoproteases such as carboxypeptidase and aminopeptidase (105,106). This allows readout of the sequence by a single mass determination. Automated sequencing by Edman degradation is often less ambiguous and easier to interpret than de novo sequence analysis of peptides by mass spectrometry, which remains a somewhat laborious technique that requires great expertise. Nevertheless, for the identification of peptides and proteins of known sequence, mass spectrometry has become the method of choice.

The goals of the ABRF Mass Spectrometry Committee include determining current state-of-the-art mass spectrometry capabilities in resource laboratories, providing test samples to permit anonymous self-evaluation, and fostering mass spectrometric techniques relevant for the biochemist. As part of an initial research study, the committee conducted a survey designed to obtain information about current mass spectrometric capabilities in 284 ABRF laboratories. The survey revealed that close to 90% of the respondents were engaged in mass spectrometry, and the respondents expressed a desire to receive a test sample designed to address basic issues of mass accuracy and resolution. In addition, the test sample would measure how well the laboratories identified peptides in mixtures that utilize mass spectrometric data with database analysis. To assist them in the identification of the unknown sample, participants were allowed to use adjunct methods such as HPLC separation of components, amino acid analysis, and enzymatic digestion by endo- and exoproteases. The more advanced laboratories were invited to perform collisional activated dissociation (CAD) and/or postsource decay (PSD) fragmentation to achieve direct mass information on fragments and to use this information for the process of identification.

Initial Results of the Survey

As mentioned, 95 out of 107 (88.8%) laboratories that responded reported using some form of mass spectrometry. The majority of the laboratories had only recently added the technique; 19 had less than 1 year experience, 60 re-

ported between 1 and 10 years experience, and only 8 reported more than 10 years experience in mass spectrometry. More than 98% of the mass spectrometers were reported to have been obtained commercially, and although most laboratories had only one, surprisingly, six laboratories reported having more than three. The choice of hardware showed about half (52.5%) that used MALDI or time of flight (TOF) and close to 37% used ESI/quadrupole instrumentation. A few used the newer technology of ESI quadrupole ion trap analysis. Only two laboratories reported using fourier transform ion cyclotron resonance mass analyzers (FTICR). According to the results, the most frequently analyzed samples were peptides and proteins, but 36 laboratories reported doing carbohydrate or glycoprotein sample analysis and 24 reported analyzing oligonucleotides, and other samples were lipids, peptide nucleic acids (oligonucleotide analogs that have a peptidelike backbone), and small organic molecules. Most sites ask for 1 to 10 pmol of sample for a mass determination. Surprisingly, and despite recently reported and dramatic results using mass spectrometry as the sole means for the identification of proteins (98,99), a majority surveyed said they perceived mass spectrometry as a supportive technique.

Finally, to gauge current interest, respondents were asked what they felt was most important in terms of addressing the future needs of the core laboratory in the field of mass spectrometry. The highest responses were increased testing, more education in mass spectrometry, mass spectrometric techniques, evaluation of instrumentation, and methods development.

Preliminary Results of the Analysis of ABRF MS1 and MS2 Test

Participants were asked to analyze two samples. MS1 was *N*-acetylated porcine angiotensinogen with average mass 1,801.085. MS2 was the same peptide mixed in equal portions with the *N*-acetylated peptide from human, mass 1,802.075. The two peptides differed only in four residues at their C-termini. Laboratories were asked to determine these values with highest accuracy and to report both mass values and resolution (calculated by the full width, half mass [FWHM] method) as a test of ability to separate closely spaced ions. Participants were told that the second sample contained two peptides and that HPLC might be required to separate them. Also, it was suggested that certain other procedures such as AAA, exoprotease or endoprotease digestion, and CAD or PSD fragmentation might be helpful in identifying the peptides.

Preliminary results based on 26 sites reporting show an average error of 387 ppm (this is equivalent to 99.96% accuracy and gives an uncertainty of ± 0.8 Da for a 1.8-kDa peptide). This amount of error is close to the difference in mass of the two peptides in MS2. Routine introduction of error at this level would prevent the distinction of an amidated from a free carboxylic acid at the C-terminus of a 2-kDa peptide. Error ranged from a high of 1,873 ppm (99.81% accuracy) to a low of 3 ppm (99.9997% accuracy). Interestingly, the error spread for two samples reported by a single site ranged from a low of 5 ppm to a high of 1,327 ppm, with an average spread of 227 ppm. Because it is assumed that the same technique and hardware were used to determine both samples, the explanation for this difference in accuracy may lie with the particular methodology employed or operator technique. The overall dependence of accuracy on hardware was similar for the two most prominent types, with values ranging from 8 to 541 ppm for ESI or quadrupole equipment and 3 to 1,435 ppm for MALDI or TOF. Resolution varied widely, with a high of 10,606 measured with a quadrupole ion trap. Values ranged from 514 to 3,001 FWHM for ESI or quadrupole and 120 to 5,457 for MALDI or TOF instruments. Five laboratories correctly identified both peptides, and 13 identified at least one of the peptides. Thus nearly 70% (18 of 26) of the respondents were able to identify one or both peptides. This result demonstrates that a significant number of resource laboratories possess a high degree of development both in the ability to use mass spectrometry and in the sophisticated adjunct methods such as database analysis for the identification of peptides and proteins.

Conclusions

The results of the survey and test sample analyses showed that although the average resource laboratory is relatively young in the field of mass spectrometry it already possesses a high degree of expertise not only for the accurate measurement of peptides but also in the ability to analyze a diverse set of sample types, to employ advanced techniques, and in general bring mass spectrometry to bear in a concerted way with other established methodologies in solving difficult analytical problems in biology. This was exemplified by the many laboratories that correctly identified test samples not only from mass measurements but also by combining data from AAA enzymatic digestion, and advanced mass spectrometric techniques. Results clearly demonstrated the high general expertise and awareness of many laboratories of a wide variety of database search algorithms and the ability to access them via the Internet.

This initial research study better defines the work of the mass spectrometry research committee. Based on the results of the survey and test samples, future studies may now be devised to focus exactly on the interests of laboratories engaged in the application of mass spectrometry to problems of biological significance.

SURVEYS OF CORE LABORATORIES

RONALD L. NIECE
University of California–Irvine
Irvine, California

KATHRYN M. IVANETICH
University of California–San Francisco
San Francisco, California

The typical, or average, biotechnology core laboratory can be characterized by a description of the instruments, space, staff, budget, and performance parameters for the resources provided. This information has been periodically obtained from general surveys of ABRF member laboratories. The first survey to provide a formal description of biomolecular core laboratories was conducted in 1987 (5);

it identified several services including protein/peptide sequencing, AAA, HPLC purification of proteins and peptides, peptide synthesis, and DNA synthesis, which are now considered traditional services. The most recent published data from a survey of facility operations (7) have shown that core facilities have been quick to incorporate new instrumentation into their repertoire and now offer technologies as diverse as capillary electrophoresis, DNA sequencing, mass spectrometry, and carbohydrate analysis (see Fig. 7). The survey data also show that many core facilities are resources not only for their home institutions but for outside investigators as well. These core facilities offer a broad range of services (Fig. 7), with outside charges usually being higher than in-house charges (7).

Combined data from several surveys of ABRF member core laboratories (7,34) have shown that academic institutions and research institutes house more than three-quarters of member laboratories, and industry has most of the remaining 25%, with only a small percentage of such core facilities in governmental settings (universities excluded). Core facilities require instrumentation, suitable laboratory and office space, appropriately trained personnel, and financial resources. A typical core facility is staffed with four full-time staff members, including one member with a doctoral degree, one with a master's degree, and two with bachelor's degrees. Directors have a median of 13 years experience in the analysis and synthesis of biological molecules, and staff have a median of 6.9 years experience (34). A typical facility houses six major instrument systems (50) in 1,000 sq. ft. of laboratory space. An average, noncapital, annual operating budget is $250,000, half of which comes from charge-back fees for services and reagents. Nearly all core facilities require substantial subsidization, averaging 50% of total operating budget, to cover the difference between user fee income and total expenses.

Performance with respect to the quantity of sample needed or scale of synthesis, turnaround time, and charge for selected services is listed in Table 5. The level of subsidy an institution provides its core facility is largely determined by the portion of operating budget that the facility obtains from its users. For highly subsidized facilities the user charges are relatively low, reflecting the high subsidy; for the 1–2% of facilities that operate without subsidy, the charge is the actual cost of performing the service, which very likely is much higher than the corresponding charge from a heavily subsidized facility. The spread in charges as shown by the large standard deviations may reflect both differences in subsidy and specifics of what is included in the standard charge. Both lower subsidy from institutional sources and higher levels of performance by the facility (including sensitivity, scale, extent of service, and turnaround time) are reflected in higher charges for users. Services for outside users generally incur an additional charge.

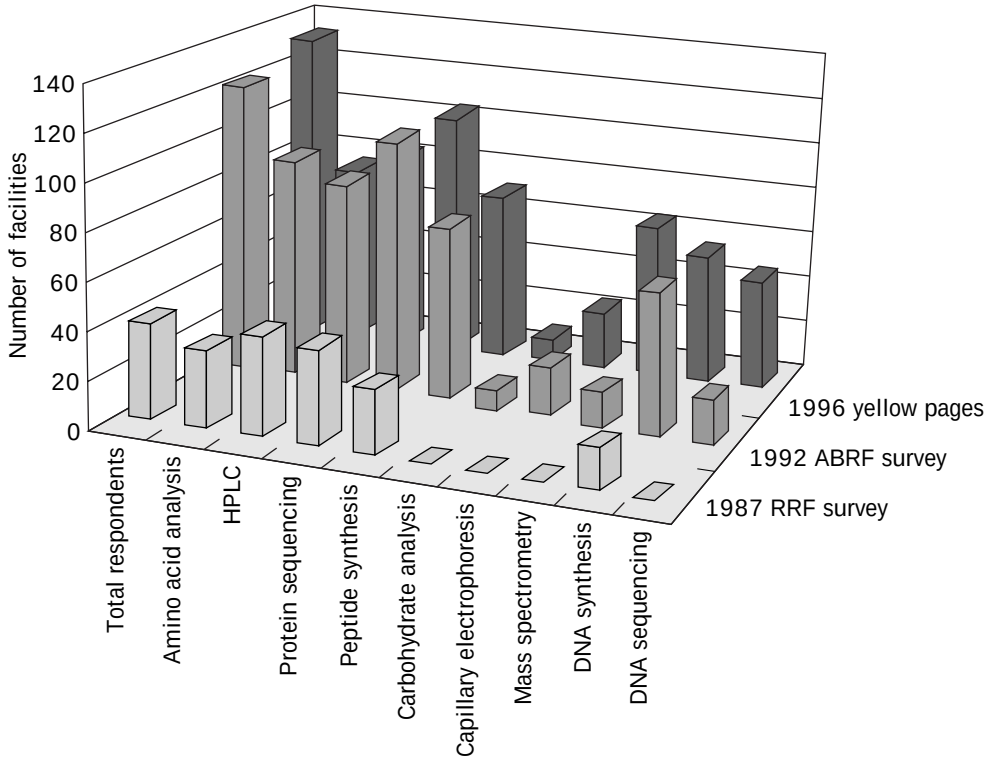


Figure 7. The number of core facilities offering selected services. Data from survey of laboratories interested in the informal organization, Research Resource Facilities (RRF) (5), respondents to survey conducted by the ABRF Survey Committee in 1992 (7), and the 1996 ABRF Membership Directory, Section II (Yellow Pages): Facilities Offering Services to Outside Investigators (available from ABRF business office or ABRF Web site).

Table 5. Services, Turnaround, Charges, and Costs Determined from ABRF Surveys

	Scale of service		Turnaround (days)	In-house charge ($)		Subsidy (% of operating costs)	Calculated cost without subsidy ($)
	Mean ± SD	Units	mean ± SD	Setup mean ± SD	Per cycle mean ± SD	mean ± SD	
Amino acid analysis	32 ± 220	μg	6 ± 4	35 ± 19		45 ± 45	64
HPLC	0.51 ± 1.4	nmol	7 ± 6	106 ± 74		47 ± 39	200
Protein sequencing	73 ± 92	pmol	9 ± 7	117 ± 79	17 ± 10	42 ± 41	788[a]
Peptide synthesis	0.56 ± 1.3	mmol	20 ± 12	194 ± 156	35 ± 12	35 ± 36	1,214[a]
Carbohydrate analysis	81 ± 74	μg	6 ± 4	20 ± 14			
Capillary electrophoresis	17 ± 18	pmol	3 ± 2				
Mass spectrometry	0.15 ± 0.2	nmol	6 ± 4	74 ± 48		30 ± 33	106
DNA synthesis	0.2 ± 0.1	μmol	4 ± 2	22 ± 13	3 ± 1	21 ± 33	123[a]
DNA sequencing	2 ± 2	μg	6 ± 4	41 ± 32		30 ± 37	59

Note: Mean ± standard deviation (SD) for scale of service with units noted, anticipated turnaround time for completion of the synthesis or analysis, in-house charge with setup and per-cycle charges (where appropriate), subsidy, and calculated charge for full cost recovery.
[a]Protein sequence (25 residues), synthetic peptide (15 residues), oligodeoxyribonucleotide (25 bases).
Source: Data from Ivanetich et al. (7).

The self-reported data on performance in the member laboratories collected by the surveys describe analytical capabilities in operating laboratories. This assessment of performance level is corroborated by the analysis of results from the "unknown" samples distributed by the ABRF research committees (see other sections of this article). In addition, members of ABRF have critiqued instrumentation by rating performance, service, technical support, and manufacturer-provided reagents. These evaluations, provided by many instrument users, have helped guide manufacturers in meeting the needs of the research and quality control communities by identifying deficiencies and useful features (11,50). A continuing concern of core facilities and their users has been the cost of consumables (7,11) because these expenses most frequently are included in the portion of operating costs that is not subsidized.

A number of trends can be discerned from the responses to these surveys conducted over a period of nearly 10 years. First, the number of core facilities continues to increase, from 146 ABRF member facilities in 1988 to well over 300 today. The average operating budgets of these facilities have increased from $158,000 to $279,000 per year; the level of institutional subsidization has remained fairly constant at approximately 50%. Although the majority of ABRF member facilities are rooted in protein chemistry, the number of core facilities that provide DNA chemistry services has increased substantially. For example, the number of ABRF core laboratories that offer DNA sequencing has doubled in the most recent (1996) survey. Although costs for protein chemistry services have either remained steady (e.g., amino acid analysis) or increased (e.g., peptide synthesis and protein sequencing), the cost of DNA chemistry services has declined dramatically. DNA synthesis costs decreased ~74% between 1988 (5) and 1993 (7). Core facilities are often the providers of advanced technology, as evidenced by the rapid implementation of new instrumentation and methodologies. Mass spectrometry was not listed as a service by any of the respondents in 1987, but in 1992 more than 10% offered the service and in 1998 nearly half of the respondents offered mass spectrometry (K. Ivanetich, unpublished data).

In summary, the average core facility characterized by survey data includes six instrument systems operated by four staff in about 1,000 sq. ft. Services offered include the traditional protein/peptide sequencing, AAA HPLC purification of proteins and peptides, peptide synthesis, and DNA synthesis and newer technologies such as capillary electrophoresis, DNA sequencing, mass spectrometry, and carbohydrate analysis. Most core facilities receive about 50% of their operating budgets from institutional funds to supplement their cost-recovery income. Core facilities evaluate instrument performance and service and technical proficiency through anonymous surveys. These data identify instrument performance levels and define the state of the art in real-world settings.

LABORATORY QUALITY AND COMPLIANCE

NADINE RITTER
Abbott Laboratories
Abbott Park, Illinois

BETH FOWLER
Autoimmune, Inc.
Lexington, Massachusetts

JOHN DOUGHERTY
Lilly Research Laboratories
Indianapolis, Indiana

AUGUSTINE SMITH
Abbott Laboratories
Abbott Park, Illinois

ELEANOR CANOVA-DAVIS
Genentech
San Francisco, California

BARBARA GHRIST
Genentech
San Francisco, California

The roles of analytical resource facilities in biotechnology range from supporting early research and discovery efforts to conducting validated testing for product development. Facilities that comply with the specific regulations governing product manufacturing and testing can also test drug products manufactured for commercial purposes. There are significant differences in the documentation and compliance requirements placed on analytical laboratories depending on how a client intends to use the data generated. Discussions between the client and the facility should identify these considerations, preferably before the analyses are conducted. If the facility can meet the client's quality and compliance needs, using a contract analytical facility can provide numerous advantages in biotechnology product development.

Most resource facilities have the instrumentation and expertise needed to investigate and physicochemically characterize many biomolecules of interest. In the research and discovery phase, most investigations are aimed at identifying active moieties. A company's research and development team could contract with a resource facility to perform selected analytical techniques that may yield critical clues about the nature of the biomolecules being studied. This strategy leverages the research and development resources across a greater breadth of technologies than may be available in-house.

In later developmental phases, an analytical profile or physicochemical "fingerprint" of the product in its final purified form may be needed. The objectives include determination of the composition of the biological material, its impurities or degradation products, and its stability under various conditions. These data may be used to establish the product's identity, purity, potency, and strength. As in discovery, the analytical methods employed may go beyond a company's in-house capabilities.

Some analytical methods, such as protein or DNA sequence analysis, are needed for only short periods in the product's development cycle. Final product characterization studies may utilize numerous methods but only on a few lots of material. Methods used in stability studies may be required only at periodic intervals over the study's lifetime. Rather than establishing and maintaining a complex analytical laboratory, clients can choose a facility with the exact instrumentation needed to perform a given analysis. A client can stay abreast of advances in technology each time a new study is contracted by choosing laboratories with the desired analytical capabilities. In addition, highly skilled resource facility scientists can assist in planning the appropriate experiments needed to obtain the best results. Finally, outsourcing selected analytical services allows a company to focus more of its internal resources on controlling the critical quality parameters of the product itself.

The level of quality assurance in a resource facility can be matched to the needs of the client base. For any laboratory, clients (and staff) are principally concerned with data integrity. Properly functioning instrumentation, adequate documentation of experiments, and traceability of results to the samples tested are minimal expectations in all laboratories. If the objective of the analysis is to provide information used in product discovery, additional quality measures may be employed. Calibration of equipment, controlled temperature storage of materials, expiration dating of reagents and materials, and documentation of these and other routine laboratory activities demonstrate the quality control of facility operations. In addition, when sample storage, handling and preparation, general and specific experimental protocols, instrument operation parameters, and data analysis methods are well documented according to established facility procedures, the client may consider such a laboratory well suited to provide reliable, reproducible analyses on which to base product design decisions.

A sometimes misused designation for general laboratory quality measures is the term good laboratory practices (GLPs). GLPs are a specific set of U.S. Food and Drug Administration (FDA) requirements for nonclinical (i.e., animal-based) laboratory studies conducted to support product development in the preclinical safety phase (107). The GLPs apply to the procedures used to characterize the test article (the substance used for dosing in the animal and toxicology experiments) to determine its stability, homogeneity, and concentration (108).

The GLP regulations include requirements that must be met by contract facilities that perform specified parts of the total study. At this phase, product's dosing formulation—usually a mixture—is being analyzed rather than the purified active biomolecule. Although a biomolecular resource facility may support the study with selected biochemical analyses, it typically does not perform the entire GLP study. The extent to which the GLP regulations apply to a laboratory depends on the role it will have in designing, conducting, documenting, archiving, and auditing the quality assurance of the study.

One element of setting up a GLP study is the designation of two key parties: a study director and a quality reviewer. The responsibilities of these entities are outlined in the regulations. Also included in the regulations are laboratory operational measures aimed at ensuring the quality of the testing results, such as those listed earlier. These measures are sometimes called good laboratory practices in casual use. When selecting a facility that claims to be a GLP laboratory, a client should clarify the meaning of this designation, discuss the role of the laboratory in the overall study plan, and define the directing and reviewing activities to prevent any misunderstandings.

Analytical procedures to support the chemical, manufacturing, and controls (CMC) sections of pharmaceutical product applications are more likely performed by contract resource facilities. Typically, this function is termed *characterization* of the product. Analytical tests such as protein sequence analysis, mass spectrometry, and AAA may be among these procedures. For these types of analyses, GLP conditions are not formally required, particularly the establishment of the study director and quality reviewer. However, facilities that implement some of the operational elements of quality described in the GLP regulations are better able to provide traceable, auditable proof of the analyses performed. As such, clients can have more confidence in reporting these results in the product application for regulatory approval.

Analytical resource facilities may be contracted to perform lot release testing on commercial products. The U.S.

regulations governing such testing are called current good manufacturing practices (cGMPs) and are published in the *Code of Federal Regulations* section 610 for biological products (109) and sections 210 and 211 for pharmaceuticals (110). There are significant regulatory requirements for virtually every aspect of laboratory operations, including the physical environment in which the laboratory is housed. An analytical laboratory that is designated cGMP must meet stringent quality measures including instrument and equipment standard operating procedures, established method testing procedures, personnel training procedures, data review and approval, and archiving of all files associated with laboratory activities and results. In addition, most process and testing methods are expected to be validated. Modifications to any of these operating procedures or validations must be controlled via an established change control system. Before contracting with (or designing) a laboratory for cGMP testing, the facility should be audited by individuals knowledgeable in both the federal regulations and analytical laboratory operations to ensure the facility meets or exceeds the level of quality expected.

Another type of quality designation that some analytical laboratories may obtain is the ISO 9000 series of measures (American National Standards Institute, 11 West 42nd St., 13th Floor, New York, NY [111]). These measures were established to provide defined levels of quality metrics in businesses ranging from law firms to lamp manufacturing and can be applied to resource facilities. The basis of these measures is to define and document the quality system of an organization: namely, to designate the responsible individual and the method of data management. ISO 9000 registration is obtained when the organization seeks and passes inspection from an accredited third party (112,113). Certification is conferred for a specified period of time, after which reinspection and recertification are required. ISO 9000 certification adds to the quality of an analytical facility in direct proportion to the level of operations already present. Usually, such laboratories are relatively small organizations with clearly defined roles and responsibilities. Interviews with facility directors and audits of laboratories allow the client to assess the quality system in place and to evaluate its integrity.

An additional global regulatory initiative is taking shape in the European Union (European Standard EN 46001 [see Ref. 114]). Similar to the ISO 9000 requirements, EN 46000 is currently expected to apply to a specific classification of medical products, although the nature of the regulations may expand in the future. Likewise, quality principles are being codified in many other countries around the world. This trend reflects a movement toward global harmonization of quality practices and product regulation, as seen in activities by the International Committee on Harmonization (see Ref. 115).

It is important to note that these quality and compliance designations are not mutually inclusive or exclusive. A laboratory may be ISO 9000 registered without complying with the GLP or cGMP regulations. Samples designated for a GLP study may be handled in the same facility that performs analyses for other, nonregulated studies, as long as the GLP samples are handled in accordance with applicable FDA regulations. Both the client and the resource facility must understand the differences among the various quality measures to have confidence in the compliance of laboratory operations and the appropriate use of the data generated.

In general, analytical resource facilities are ideally equipped to provide the various types of biomolecular analyses required for biotechnology products. As analytical methods and techniques evolve, these facilities are positioned to incorporate technical advances as ongoing improvements to their services. In addition, many of these laboratories may be able to provide the level of quality needed to support specified regulatory uses of the data obtained. As partners in product development, these facilities provide highly skilled, technically advanced teams of scientists with the instruments and expertise needed to perform complex, ever-changing methods. Additionally, as regulatory agencies begin to accept specified biotechnology products for licensing (116), adopting a strategy of outsourcing special bioanalytical services may prove to be an effective way to manage product development costs while ensuring the quality of biotechnology products.

EFFECTIVE USE OF BIOMOLECULAR RESOURCE FACILITIES

CLAYTON NAEVE
St. Jude Children's Research Hospital
Memphis, Tennessee

Progress in biomedical research is becoming increasingly dependent on the utilization of shared resources or core facilities. These facilities typically provide biotechnology reagents and services that require skilled staff and expensive instruments that are most effectively and economically provided by an institution in the form of shared resources. Such core facilities span a broad range in the size and spectrum of services and reagents offered and may include protein or DNA sequencing; DNA, RNA, or peptide synthesis; mass spectroscopy; carbohydrate analysis; and many other possible services. Because the use of sophisticated instrumentation is high in these core facilities, the investigator often assumes that these are turnkey operations (i.e., submit a request, get a reagent or result). In reality, these resources play a critical partnership role in the research endeavor and should be viewed as such to succeed. If you have a core facility available that provides the services you require, the following suggestions may help optimize your use of their expertise.

Before Submission of a Request

Familiarize Yourself with the Core Facilities Available to You and the Range of Expertise, Services, and Reagents They Provide. Knowing that the protein core facility at your institution can or cannot sequence proteins at the femtomole level or knowing that the DNA core facility at your institution will only accept DNA sequencing templates cloned into specific vectors may influence your experimental design. If cost is a deterrent, communicate with the facility to understand the basis for its service charges; some fa-

cilities and services are highly subsidized by their host institutions and some are not (see the "Surveys of Core Laboratories" section of this article). One of the reasons for the seemingly high charges at your facility may be that its degree of subsidization via nonuser fee income is much less than the ~50% average level (7); hence, its charges may cover much more than just 50% of the cost of providing the requested service.

Talk to Your Core Facility Director and Staff before You Write the Grant, Design the Protocol, Plan the Experiment, or Run the Sample That Will Utilize Their Facility. Communicating with facility staff will ensure that funds can be budgeted for the needed analyses and syntheses, and the core laboratory can verify that the biotechnological aspects of the proposal are indeed reasonable. Proposing to load an SDS PAGE gel slice directly onto a protein/peptide sequencer or proposing a postsource decay (sequencing) study on a laser desorption or TOF mass spectrometer that is not equipped with a reflectron would not enhance the probability of your grant application being funded. Take advantage of the resource of expertise and services available.

Understand the Limits of the Technology. For instance, although protein hydrolysis and amino acid analysis provide the amount of individual amino acids in a protein or peptide sample, they do not provide the number of amino acid residues unless an accurate mass is available, the sample is pure, and corrections are made for peptide bonds that are difficult to hydrolyze and for loss of less-stable amino acids.

Follow Your Core Facility's Sample Preparation Guidelines. Proper sample preparation is probably the most important step in ensuring success with most biotechnology chemistries. Preparation of clean DNA templates for automated sequencing is a classic example; follow the guidelines given to you explicitly and you will greatly increase your probability of success. Consult with facility staff in planning the appropriate sample preparation strategy before sample submission.

After Receipt of Your Data or Reagents

Understand the Significance of the Data Produced. If your DNA sequencing facility provides you with four-color electropherograms and unedited sequence data, you must be aware that miscalls occur and that it is your responsibility to evaluate the quality of the data and correct errors. Suppose you have submitted a sample for MALDI or TOF mass analysis and the observed mass differs from the predicted by 0.01% using reflectron mode analysis. Is this within the accuracy range of the instrument (thus helping to confirm the identity of the peptide) or does this suggest the presence of a modification? To assess significance, you must familiarize yourself with the technology in use to judge the results obtained. Obviously, the core facility staff will help you with this assessment, but the ultimate responsibility is yours.

Work with the Facility to Correct Any Real or Perceived Deficiencies. If you are unhappy with the results obtained, talk to the facility's staff or director. If problems do occur, your input and experience can often be critical in diagnosing and correcting difficulties. On the other hand, the problems that you believe are present (e.g., absence of cysteine during the amino acid sequencing of a peptide known to contain this amino acid) may reflect an oversight or lack of knowledge on your part (in this case the failure to appropriately modify all cysteine residues before submitting the sample for amino acid sequencing).

What Should You Do If Your Core Laboratory Does Not Provide the Service You Need? First, ask why this particular service is not available. There are many reasons why a given service may not be available, including insufficient demand, staff, expertise, equipment, or space to justify offering the service. For instance, many core laboratories do not offer automated C-terminal protein sequencing because of the relatively meager demand, the large amounts of protein required, and the relatively few amino acid residues that can be sequenced. It is thus important to first determine why the service is not available. Second, if there is sufficient demand to justify the addition of this service, help the core laboratory bring this new service to fruition. Third, ask for a referral, because your core laboratory may well know an outside core laboratory that is particularly well suited to providing the requested service. In this instance, it is important to tell staff at your core laboratory what happened at the outside core laboratory so they may better advise others who follow.

Support Your Core Laboratory. A close partnership between the core laboratory and its users is critical for success. Just as the user needs access to state-of-the-art biotechnologies, the core laboratory can only succeed with the support of its users. Support includes acknowledging all services provided by the core laboratory and forwarding reprints of these publications to the core laboratory, responding to requests made by the core laboratory for research proposals and other support of their shared instrumentation grant applications, and helping the core laboratory gain institutional support, which may range from subsidization of operating costs (to help lower user fees) to equipment acquisitions to gaining additional space. Finally, because much of the work carried out in core laboratories goes without reward in terms of publications, be appreciative of the efforts of core laboratory staff and do not hesitate to coauthor publications with them when their efforts and input go beyond that which justifies a simple acknowledgment.

FUTURE OF CORE LABORATORIES

JOHN STULTS
Genentech
San Francisco, California

What does the future hold for core laboratories? The needs of the biological community are changing rapidly, as are

the economics and technology of science. As a result, the twenty-first century will present challenges and exciting new possibilities for core laboratories.

Paradigm Shift in the Postgenome Era

The single most important transforming factor in the immediate future is the availability of enormous amounts of genomic sequence. Partial nucleotide sequences from human expressed sequence tags (ESTs) (117,118) already number in excess of 1,000,000 (119). Complete genomes have been sequenced for a number of organisms, including *Saccharomyces cerevisiae* (120). The imminent completion of a number of other genomes, including human, and the advent of the postgenome era will provide an information resource previously unimaginable. This information glut will trigger a paradigm shift in the ways that core laboratories operate and in the services they provide. The emphasis on sequence determination for new proteins, particularly by DNA sequencing and cloning, an overwhelming and perhaps inordinate goal of biology for the past 20 years, will give way to renewed emphasis on biological principles and protein function and its relationship to structure (121). For core laboratories, this information will likely mean that much of traditional protein sequence determination will give way to protein identification via an array of rapidly evolving approaches and technologies, many of which utilize mass spectrometry, and that traditional DNA sequencing will give way to powerful genetic analysis techniques based on hybridization to high-density DNA oligonucleotide arrays (122).

Automation

Automation of many tasks will continue to be a key factor in further reducing operating costs and increasing productivity and reproducibility. Proper, efficient sample preparation for analysis has been and will continue to be a bottleneck for rapid and reliable analyses, making it a primary target for automation. DNA synthesis chemistries and instrumentation have already reached a state of extremely high throughput and reliable operation that is perhaps manifested best by the 10-fold reduction in the in-house charge for a 25-mer DNA oligonucleotide that has occurred over the past 10 years (5,33). Instrumentation is now available that can synthesize and, in some cases, automatically purify 48 or more DNA oligonucleotides. DNA sequencing is rapidly approaching high-throughput operation such that some instruments can analyze as many as 96 samples in parallel. To keep pace with this instrumentation, increased attention is being given to the use of robotics to automate DNA template preparation. Peptide synthesis is moving more and more toward automation and high-throughput operation with the availability of instruments capable of synthesizing 12 or more peptides followed by the cleavage of the newly synthesized peptides from their solid supports. In the case of both DNA and peptide synthesis, increased automation accompanied by decreased reagent costs due to the smaller synthesis scales adopted has resulted in further cost reductions. Amino acid analysis has been a reasonably high-throughput method for years, with a typical instrument able to complete an analysis in 1 h. Protein sequencing remains slow by comparison, yet its operation continues to be simplified and its sensitivity of detection increased, and automated operation for analyzing multiple samples in tandem is now available. In contrast, mass spectrometry is inherently a rapid technique. Automation of mass spectrometry, although done for years with GC and MS, has only recently become available for protein and peptide analysis. The capabilities of the instrumentation are sure to keep advancing; in general, the higher the throughput of the instrumentation, the more suited it is for use in core laboratories.

Future Opportunities

Undoubtedly, the availability of the sequences of the approximately 100,000 human proteins will pose many more questions than it answers and will spawn a new area of investigation called proteome research. This availability will also result in a renaissance in studies directed toward understanding how protein structure relates to function. In their simplest form, the questions are as follows: What are the functions of these proteins? How do they accomplish their tasks? How can we take advantage of this knowledge to intervene in a bewildering array of diseases? It is not difficult to illustrate our ignorance in this regard. Currently, the only reliable means for predicting three-dimensional structure and function from a newly discovered protein sequence is to carry out a database search and hopefully identify homologous proteins in the same family or in other organisms. If the sequence is truly unique, no laboratory currently is capable of accurately predicting the function of this protein, its three-dimensional structure, or how the intricate folding of its one-dimensional amino acid sequence gives genesis to function. Compared with the 100,000 human proteins, the current availability of only a few dozen structures for protein–ligand complexes surely pales in comparison and bodes extremely well for increasing the need for identifying sites of protein–protein and protein–ligand cross-linking and for the use of synthetic peptide analogs and other approaches to identify active sites in proteins. Similarly, opportunities for core laboratories abound in terms of the biophysical instrumentation (e.g., spectrofluorometers, scanning and isothermal calorimeters, analytical ultracentrifuges, surface plasmon resonance, and circular dichroism instruments) needed to quantitatively characterize these interactions.

All these factors suggest that the future of biotechnology core laboratories is indeed bright and far more interesting than in the past, with the tedium of determining primary sequences being replaced rapidly by the more interesting and pertinent task of determining the reason for that order. Core laboratories are a tremendous intellectual and biotechnological resource whose importance can only be increased by the increasing financial constraints that helped promote their rapid emergence and by the ever-increasing complexity that seems to lie beneath each biochemical and biomedical discovery. Although core laboratories are synonymous with sophisticated instrumentation, in the end it is the efforts of people that lead to advancements in science. Ultimately, it is the interdisciplinary knowledge, research experience, dedication, and

ingenuity of core laboratory personnel that will lead to their continued importance and success.

ACKNOWLEDGMENTS

We especially thank all past and present members of all of the ABRF Research Committees that compiled the data on which much of this article is based and all of the member laboratories who participated in this research. Special thanks are given to all the individuals who reviewed sections of this article for their critical insights and to the National Science Foundation (DIR9003100), the Department of Energy (DE-FG02-95ER61839), and all past and present ABRF members and corporate sponsors who have provided support for the ABRF.

BIBLIOGRAPHY

1. R.M. Hewick, M.W. Hunkapiller, L.E. Hood, and W.J. Dreyer, *J. Biol. Chem*. **256**, 7990–7997 (1981).
2. M.J. Gait, in M.J. Gait ed., *Oligonucleotide Synthesis: A Practical Approach*, IRL Press, Oxford, 1984, pp. 1–22.
3. R.L. Niece, D. Atherton, A.V. Fowler, R. Kutny, and A.J. Smith, in K.A. Walsh ed., *Methods in Protein Sequence Analysis*, Humana Press, Clifton, New Jersey, 1987, pp. 317–319.
4. R.L. Niece, K.R. Williams, C.L. Wadsworth, J. Elliott, K.L. Stone, W.J. McMurray, A.V. Fowler, D. Atherton, R. Kutny, and A.J. Smith, in T. Hugli ed., *Techniques in Protein Chemistry*, Academic Press, San Diego, Calif., 1989, pp. 89–110.
5. K.R. Williams, R.L. Niece, D. Atherton, A.V. Fowler, R. Kutny, and A.J. Smith, *FASEB J*. **2**, 3124–3130 (1988).
6. R.L. Niece, C.M. Beach, R.F. Cook, G.M. Hathaway, and K.R. Williams, *FASEB J*. **5**, 2756–2760 (1991); **6**, 792 (1992).
7. K.M. Ivanetich, R.L. Niece, M.F. Rohde, E. Fowler, and T.K. Hayes, *FASEB J*. **7**, 1109–1114 (1993).
8. D.W. Speicher, G.A. Grant, R.L. Niece, R.W. Blacher, A.V. Fowler, and K.R. Williams, in J.J. Villafranca ed., *Current Research in Protein Chemistry*, Academic Press, San Diego, Cal., 1990, pp. 159–166.
9. R.L. Niece, D. Atherton, C.M. Beach, R.F. Cook, A.V. Fowler, G.M. Hathaway, R. Kutny, A.J. Smith, and K.R. Williams, *J. Protein Chem*. **9**, 262–263 (1990).
10. J.W. Crabb, L.H. Ericsson, D. Atherton, A.J. Smith, and R. Kutny, in J.J. Villafranca ed., *Current Research in Protein Chemistry*, Academic Press, San Diego, Cal., 1990, pp. 49–61.
11. C.M. Beach, G.M. Hathaway, T.K. Hayes, A.J. Smith, and R.L. Niece, *ABRF News* **2**, Supplement 1–7 (1991).
12. R.L. Niece, L.H. Ericsson, A.V. Fowler, A.J. Smith, D.W. Speicher, J.W. Crabb, and K.R. Williams, in H. Jörnvall, J.-O. Höög, and A.-M. Gustavsson eds., *Methods in Protein Sequence Analysis*, Birkhäuser Verlag, Basel, 1991, pp. 133–141.
13. L.H. Ericsson, D. Atherton, R. Kutny, A.J. Smith, and J.W. Crabb, in H. Jörnvall, J.-O. Höög, and A.-M. Gustavsson eds., *Methods of Protein Sequence Analysis*, Birkhauser Verlag, Basel, 1991, pp. 143–150.
14. K.Ü. Yüksel, G.A. Grant, L.M. Mende-Mueller, R.L. Niece, K.R. Williams, and D.W. Speicher, in J.J. Villafranca ed., *Techniques in Protein Chemistry II*, Academic Press, San Diego, Calif., 1991, pp. 151–162.
15. G.E. Tarr, R.J. Paxton, Y.-C.E. Pan, L.H. Ericsson, and J.W. Crabb, in J.J. Villafranca ed., *Techniques in Protein Chemistry II*, Academic Press, San Diego, Calif., 1991, pp. 139–150.
16. D.L. Crimmins, G.A. Grant, L.M. Mende-Mueller, R.L. Niece, C.A. Slaughter, D.W. Speicher, and K.Ü. Yüksel, in R. Hogue-Angeletti ed., *Techniques in Protein Chemistry III*, Academic Press, San Diego, Calif., 1992, pp. 35–43.
17. A.J. Smith, J.D. Young, S.A. Carr, D.R. Marshak, L.C. Williams, and K.R. Williams, in R. Hogue-Angeletti ed., *Techniques in Protein Chemistry III*, Academic Press, San Diego, Calif., 1992, pp. 219–229.
18. D.J. Strydom, G.E. Tarr, Y.-C.E. Pan, and R.J. Paxton, in R. Hogue-Angeletti ed., *Techniques in Protein Chemistry III*, Academic Press, San Diego, Calif., 1992, pp. 261–274.
19. R.L. Niece and R.H. Angeletti, *FASEB J*. **7**, A1056 (1993).
20. G.B. Fields, S.A. Carr, D.R. Marshak, A.J. Smith, J.T. Stults, L.C. Williams, K.R. Williams, and J.D. Young, in R. Hogue-Angeletti ed., *Techniques in Protein Chemistry IV*, Academic Press, San Diego, Calif., 1993, pp. 229–238.
21. D.J. Strydom, T.T. Andersen, I. Apostol, J.W. Fox, R.J. Paxton, and J.W. Crabb, in R. Hogue-Angeletti ed., *Techniques in Protein Chemistry IV*, Academic Press, San Diego, Calif., 1993, pp. 279–288.
22. S.M. Mische, K.Ü. Yüksel, L.M. Mende-Mueller, P. Matsudaira, D.L. Crimmins, and P.C. Andrews, in R. Hogue-Angeletti ed., *Techniques in Protein Chemistry IV*, Academic Press, San Diego, Calif., 1993, pp. 453–461.
23. R.L. Niece and L.H. Ericsson, *FASEB J*. **8**, A1263 (1994).
24. R.T. Pon, G.A. Buck, R.L. Niece, M. Robertson, A.J. Smith, and E. Spicer, *BioTechniques* **17**, 526–534 (1994).
25. J. Rush, P.C. Andrews, D.L. Crimmins, J.E. Gambee, G.A. Grant, S.M. Mische, and D.W. Speicher, in J.W. Crabb ed., *Techniques in Protein Chemistry V*, Academic Press, San Diego, Calif., 1994, pp. 133–141.
26. K.Ü. Yüksel, T.T. Andersen, I. Apostol, J.W. Fox, J.W. Crabb, R.J. Paxton, and D.J. Strydom, in J.W. Crabb ed., *Techniques in Protein Chemistry V*, Academic Press, San Diego, Calif., 1994, pp. 231–240.
27. G.B. Fields, R.H. Angeletti, S.A. Carr, A.J. Smith, J.T. Stults, L.C. Williams, and J.D. Young, in J.W. Crabb ed., *Techniques in Protein Chemistry V*, Academic Press, San Diego, Calif., 1994, pp. 501–507.
28. C.W. Naeve and R.L. Niece, *FASEB J*. **9**, A1260 (1995).
29. C.W. Naeve, G.A. Buck, R.L. Niece, R.T. Pon, M. Robertson, and A.J. Smith, *BioTechniques* **19**, 448–453 (1995).
30. K.Ü. Yüksel, T.T. Andersen, I. Apostol, J.W. Fox, R.J. Paxton, and D.J. Strydom, in J.W. Crabb ed., *Techniques in Protein Chemistry VI*, Academic Press, San Diego, Calif., 1995, pp. 185–192.
31. J.E. Gambee, P.C. Andrews, K. DeJongh, G. Grant, B. Merrill, S. Mische, and J. Rush, in J.W. Crabb ed., *Techniques in Protein Chemistry VI*, Academic Press, San Diego, Calif., 1995, pp. 209–217.
32. G.B. Fields, R.H. Angeletti, L.F. Bonewald, W.T. Moore, A.J. Smith, J.T. Stults, and L.D. Williams, in J.W. Crabb ed., *Techniques in Protein Chemistry VI*, Academic Press, San Diego, Calif., 1995, pp. 539–546.
33. R.T. Pon, G.A. Buck, K.M. Hager, C.W. Naeve, R.L. Niece, M. Robertson, and A.J. Smith, *BioTechniques* **21**, 680–685 (1996).
34. L. Bibbs, N.D. Denslow, C. Glabe, K. Ivanetich, C.W. Naeve, and R.L. Niece, *ABRF News* **7**, 18–21 (1996).
35. A.M. Mahrenholz, N.D. Denslow, T.T. Andersen, K.M. Schegg, K. Mann, S.A. Cohen, J.W. Fox, and K.Ü. Yüksel, in D.R. Marshak ed., *Techniques in Protein Chemistry VII*, Academic Press, San Diego, Calif., 1996, pp. 323–330.

36. R.H. Angeletti, L. Bibbs, L.F. Bonewald, G.B. Fields, J.S. McMurray, W.T. Moore, and J.T. Stults, in D.R. Marshak ed., *Techniques in Protein Chemistry VII*, Academic Press, San Diego, Calif., 1996, pp. 261–274.
37. K. De Jongh, J. Fernandez, J.E. Gambee, G.A. Grant, B. Merrill, K.L. Stone, and J. Rush, in D.R. Marshak ed., *Techniques in Protein Chemistry VII*, Academic Press, San Diego, Calif., 1996, pp. 347–358.
38. G.B. Fields, L. Bibbs, L.F. Bonewald, J.S. McMurray, W.T. Moore, A.J. Smith, J.T. Stults, L.C. Williams, and R.H. Angeletti, in P.T.P. Kaumaya and R.S. Hodges eds., *Peptides: Chemistry, Structure, and Biology*, Mayflower Scientific, 1996, pp. 52–54.
39. P.S. Adams, M.K. Dolejsi, S. Hardin, S. Mische, B. Nanthakamur, H. Riethman, J. Rush, and P. Morrison, *BioTechniques* **21**, 678 (1996).
40. P.S. Adams, M.K. Dolejsi, S. Hardin, S.M. Mische, B. Nanthakamur, H. Riethman, J. Rush, and P. Morrison, *Eighth Annual Genome Sequencing and Analysis Conference*, 1996, B-26.
41. P.S. Adams, M.K. Dolejsi, S. Hardin, S.M. Mische, B. Nanthakamur, H. Riethman, J. Rush, and P. Morrison, *http://mbcf.dfci.harvard.edu/abrfdnaseq* (last updated February 8, 1999).
42. G.A. Grant and M.W. Crankshaw, in B.J. Smith ed., *Methods in Molecular Biology*, vol. 64: *Protein Sequencing Protocols*. Humana Press, Clifton, 1997, New Jersey, pp. 197–215.
43. R.R. Townsend, A. Manzi, R.K. Merkle, M.F. Rohde, M. Spellman, A.J. Smith, and S. Carr, *ABRF News* **8**, 14–21 (1997).
44. R.H. Angeletti, L.F. Bonewald, and G.B. Fields, in G.B. Fields ed., *Methods in Enzymology*, Academic Press, San Diego, Calif., 1997, pp. 697–717.
45. R.H. Angeletti, L. Bibbs, L.F. Bonewald, G.B. Fields, J.W. Kelly, J.S. McMurray, W.T. Moore, and S.T. Weintraub, in D.R. Marshak ed., *Techniques in Protein Chemistry VIII*, Academic Press, San Diego, Calif., 1997, pp. 875–890.
46. J. Fernandez, A. Admon, K. DeJongh, G.A. Grant, W. Henzel, W.S. Lane, K.L. Stone, and B. Merrill, in D.R. Marshak ed., *Techniques in Protein Chemistry VIII*, Academic Press, San Diego, Calif., 1997, pp. 69–78.
47. K.R. Williams, U. Hellman, R. Kobayashi, W.W. Lane, S.M. Mische, and D.W. Speicher, in D.R. Marshak ed., *Techniques in Protein Chemistry VIII*, Academic Press, San Diego, Calif., 1997, pp. 99–109.
48. K.M. Schegg, N.D. Denslow, T.T. Andersen, Y.A. Bao, S.A. Cohen, A.M. Mahrenholz, and K. Mann, in D.R. Marshak ed., *Techniques in Protein Chemistry VIII*, Academic Press, San Diego, Calif., 1997, pp. 207–216.
49. A. Mahrenholz, T.T. Andersen, Y. Bao, S.A. Cohen, N.D. Denslow, J. Hulmes, P.E. Hunzijer, K. Mann, K.M. Schegg, and K. West, ABRF Amino Acid Analysis Survey: Identification of Proteins Electroblotted to PVDF. *http://www.abrf.org/ABRF/ResearchCommittees/aaarescomm.html* (last updated February 9, 1999).
50. K.M. Ivanetich, L. Bibbs, R.L. Niece, N.D. Denslow, C.W. Naeve, M. Rohde, and L.H. Ericsson, *Genetic Engineering News* **17**, 17, 19, 47 (1997).
51. P. Edman, *Acta Chem. Scand.* **4**, 283 (1950).
52. M.W. Hunkapiller, in A.S. Bhown ed., *Protein/Peptide Sequence Analysis: Current Methodologies*, CRC Press, Boca Raton, Fla., 1988, pp. 87–117.
53. J.L. Brown and W.K. Roberts, *J. Biol. Chem.* **251**, 1009–1014 (1976).
54. M.W. Hunkapiller, E. Lujan, F. Ostrander, and L.E. Hood, *Methods Enzymol.* **91**, 227–236 (1983).
55. P. Matsudaira, *J. Biol. Chem.* **262**, 10035–10038 (1987).
56. J.I. Elliott, K.L. Stone, and K.R. Williams, *Anal. Biochem.* **211**, 94–101 (1993).
57. P. Edman and G. Begg, *Eur. J. Biochem.* **1**, 80 (1967).
58. A. Shevchenko, M. Wilm, O. Vorm, and M. Mann, *Anal. Chem.* **68**, 850–858 (1996).
59. U.K. Laemmli, *Nature* **227**, 680–685 (1970).
60. P.H. O'Farrell, *J. Biol. Chem.* **250**, 4007–4021 (1975).
61. R.H. Aebersold, J. Leavitt, R.A. Saavedra, L.E. Hood, and S.B.H. Kent, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6970–6974 (1987).
62. H. Kawasaki, Y. Emori, and K. Suzuki, *Anal. Biochem.* **191**, 332–336 (1990).
63. J. Rosenfeld, J. Capdevielle, J.C. Guillemot, and P. Ferrara, *Anal. Biochem.* **203**, 173–179 (1992).
64. J. Fernandez, M. DeMott, D. Atherton, and S.M. Mische, *Anal. Biochem.* **201**, 255–264 (1992).
65. S.D. Patterson, D. Hess, T. Yungwirth, and R. Aebersold, *Anal. Biochem.* **202**, 193–203 (1992).
66. U. Hellman, C. Wernstedt, J. Góñez, and C.-H. Heldin, *Anal. Biochem.* **224**, 451–455 (1995).
67. S. Best, D.F. Reim, J. Mozdzanowski, and D.W. Speicher, in J.W. Crabb ed., *Techniques in Protein Chemistry V*, Academic Press, New York, 1994, pp. 205–213.
68. J. Fernandez, L. Andrews, and S.M. Mische, *Anal. Biochem.* **218**, 112–117 (1994).
69. H. Erjument-Bromage, M. Lui, D. Sabatini, S.H. Snyder, and P. Tempst, *Protein Sci.* **3**, 2425–2446 (1994).
70. K. Williams, M. LoPresti, and K. Stone, in D. Marshak ed., *Techniques in Protein Chemistry VIII*, Academic Press, San Diego, Calif., 1997, pp. 79–90.
71. U. Hellman, in R.M. Kamp, T. Choli-Papadoupoulou, and B. Whittmann-Liepold eds., *Protein Structure Analysis: Preparation, Characterization, and Microsequencing*, Springer-Verlag, Berlin, 1997, pp. 97–104.
72. D.J.C. Pappin, P. Hojrup, and A.J. Bleasby, *Curr. Biol.* **3**, 327–332 (1993).
73. J.R. Yates, III, S. Speicher, P.R. Griffin, and T. Hunkapiller, *Anal. Biochem.* **214**, 397–408 (1993).
74. W.J. Henzel, T.M. Billeci, J.T. Stults, S.C. Wong, C. Grimley, and C. Watanabe, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5011–5015 (1993).
75. M. Mann, P. Hajrup, and P. Roepstorff, *Biol. Mass Spectrom.* **22**, 338–345 (1993).
76. K.R. Williams, S.M. Samandar, K.L. Stone, M. Saylor, and J. Rush, in J.M. Walker ed., *The Protein Protocols Handbook*, Humana Press, Totowa, New Jersey, 1996, pp. 541–555.
77. J. Mozdanowski and D.W. Speicher, in J.J. Villafranca ed., *Current Research in Protein Chemistry: Techniques, Structure, and Function*, Academic Press, San Diego, Calif., 1990, pp. 87–93.
78. B. Lindberg and J. Lonngren, *Methods Enzymol.* **50**, 3–33 (1978).
79. R.K. Merkle and I. Poppe, *Methods Enzymol.* **230**, 1–15 (1994).
80. R.R. Townsend, in C. Horvath and L.S. Ettre eds., *Chromatography in Biotechnology*, American Chemical Society, Washington, D.C., 1993, pp. 86–101.
81. R.R. Townsend, in Z. El Rassi ed., *Carbohydrate Analysis: High Performance Liquid Chromatography and Capillary Electrophoresis*, Elsevier, New York, 1995, pp. 181–209.

82. R.G. Spiro, *J. Biol. Chem.* **237**, 382–388 (1962).
83. R.G. Spiro and V.D. Bhoyroo, *J. Biol. Chem.* **249**, 5704–5717 (1974).
84. R.R. Townsend, M.R. Hardy, T.C. Wong, and Y.C. Lee, *Biochemistry* **25**, 5725–5731 (1986).
85. A.S.B. Edge and R.G. Spiro, *J. Biol. Chem.* **262**, 16135–16141 (1987).
86. E.D. Green, G. Adelt, J.U. Baenziger, S. Wilson, and H. van Halbeek, *J. Biol. Chem.* **263**, 18253–18268 (1988).
87. B. Bendiak, M. Harris-Brandts, S.W. Michnick, J.P. Carver, and D. Cumming, *Biochemistry* **28**, 6491–6499 (1989).
88. D.A. Cumming, C.G. Hellerqvist, M. Harris-Brandts, S.W. Michnick, J.P. Carver, and B. Bendiak, *Biochemistry* **28**, 6500–6512 (1989).
89. U. Hobohm, T. Houthhaeve, and C. Sander, *Anal. Biochem.* **222**, 202–209 (1994).
90. D.J. Strydom, *J. Chromatogr.* **678**, 17–23 (1994).
91. C.M. Starr, R.I. Masada, C. Hague, E. Skop, and J.C. Klock, *J. Chromatogr.* **720**, 295–321 (1996).
92. M. Dole, L.L. Mack, R.L. Hines, R.C. Mobley, L.D. Ferguson, and M.B. Alice, *J. Chem. Phys.* **49**, 2240–2249 (1968).
93. J.B. Fenn, M. Mann, C.K. Meng, S.F., Wong, and C.M. Whitehouse, *Science* **246**, 64–71 (1989).
94. M. Karas and F. Hillenkamp, *Anal. Chem.* **60**, 2299–2301 (1988).
95. R. Beavis and B.T. Chait, *Rapid Commun. Mass Spectrom.* **3**, 432–435 (1989).
96. G.J. Currie and J.R. Yates, III, *J. Am. Soc. Mass Spectrom.* **4**, 955–963 (1993).
97. U. Pieles, W. Zurcher, M. Schar, and H.E. Moser, *Nucleic Acids Res.* **21**, 3191–3196 (1993).
98. D.C. Stahl, K.M. Swiderek, M.T. Davis, and T.D. Lee, *J. Am. Soc. Mass Spectrom.* **7**, 532–540 (1996).
99. M. Wilm, A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fotsis, and M. Mann, *Nature* **379**, 466–469 (1996).
100. P.R. Griffin, M.J. MacCoss, J.K. Eng, R.A. Blevins, J.S. Aaronson, and J.R. Yates, III, *Rapid Commun. Mass Spectrom.* **9**, 1546–1551 (1995).
101. S.D. Patterson and R. Aebersold, *Electrophoresis* **16**, 1791–1814 (1995).
102. C.H. Wheeler, S.L. Berry, M.R. Wilkins, J.M. Corbett, K. Ou, A.A. Gooley, I. Humphery-Smith, K.L. Williams, and M.J. Dunn, *Electrophoresis* **17**, 580–587 (1996).
103. P. Kahn, *Science* **270**, 369–370 (1995).
104. K. Bieman, *Protein Sci.* **4**, 1920–1927 (1995).
105. M. Bartlett-Jones, W.A. Jeffery, H.F. Hansen, and D.J.C. Pappin, in J.W. Crabb ed., *Techniques in Protein Chemistry VI*, Academic Press, San Diego, Calif., 1995, pp. 3–11.
106. D.H. Patterson, G.E. Tarr, F.E. Regnier, and S.A. Martin, *Anal. Chem.* **67**, 3971–3978 (1995).
107. *Fed. Regist.* Part VI, 21 *Code of Federal Regulations*, 58.
108. U.S. FDA, *Digest of Questions and Answers*, June 20, 1981.
109. U.S. FDA, 21 *Code of Federal Regulations* 610.
110. U.S. FDA, 21 *Code of Federal Regulations* 210 and 211.
111. A. Thayer, *Chem. Eng. News* 12–17 (1993).
112. M. Breitenberg, NISTIR No. 4721, U.S. Department of Commerce, National Institute of Standards and Technology, Standards Code and Information Program, Office of Standard Service, November 1991.
113. J.H. Stratton, *Quality Progress* 67–69 (1992).
114. European Standard EN 46001, CEN/CENELEC, Brussels, Belgium, August 1996.
115. Discussions of the Activities of the International Committee on Harmonization of Technical Requirements for the Registration of Pharmaceuticals for Human Use, *Fed. Regist.* 59 (1994).
116. Elimination of the Establishment Licensing Application for Specified Biotechnology and Specified Synthetic Biological Products, *Fed. Regist.* **61**, 94 (1996).
117. M.D. Adams, J.M. Kelley, J.D. Gocayne, M. Dubnick, M.H. Polymeropoulos, H. Xiao, C.R. Merril, A. Wu, B. Olde, R.F. Moreno, A.R. Kerlavage, W.R. McCombie, and J.C. Venter, *Science* **252**, 1651–1656 (1991).
118. L. Hillier, G. Lennon, M. Becker, M.F. Bonldo, B. Chiapelli, S. Chissoe, N. Dietrich, T. DuBuque, A. Favello, W. Gish, M. Hawkins, M. Hultman, T. Kucaba, M. Lacy, M. Le, N. Le, E. Mardis, B. Moore, M. Morris, J. Parsons, C. Prange, L. Rifkin, T. Rohlfing, K. Schellenberg, M.B. Soares, F. Tan, J. Thierry-Meg, E. Trevaskis, K. Underwood, P. Wohldman, R. Waterston, R. Wilson, and M. Marra, *Genome Res.* **6**, 807–828 (1996).
119. National Center for Biotechnology Information dbEST, *http://www.ncbi.nlm.nih.gov/dbEST/* (released February 9, 1999).
120. A. Goffeau, B.G. Barrell, H. Bussey, R.W. Davis, B. Dujon, H. Feldmann, F. Galibert, J.D. Hoheisel, C. Jacq, M. Johnston, E.J. Louis, H.W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, and S.G. Oliver, *Science* **274**, 546–549 (1996).
121. S.M. Tilghman, *Genome Res.* **6**, 773–780 (1996).
122. M. Chee, R. Yang, E. Hubbell, A. Berno, X.C. Huang, D. Stern, J. Winkler, D.J. Lockhart, M.S. Morris, and S.P. Fodor, *Science* **274**, 610–614 (1996).

PROFESSIONAL SOCIETIES, SOCIETY FOR INDUSTRIAL MICROBIOLOGY (SIM)

CHRISTINE L. CASE
Skyline College
San Bruno, California

ROBERT SCHWARTZ
Abbott Laboratories
North Chicago, Illinois

OUTLINE

BACKGROUND OF THE ORGANIZATION

The Society for Industrial Microbiology (SIM) is a nonprofit professional association dedicated to the advancement of applied microbiological sciences, especially as they apply to industrial products, biotechnology, materials, and processes. A primary objective of SIM is to serve as a liaison among the various specialized fields of applied microbiology. It promotes the exchange of scientific information through meetings and publications.

Industrial microbiology is the application of scientific and engineering principles to the processing of materials by microorganisms or other biological agents to create useful products for the benefit of humankind. These microorganisms may be natural isolates, laboratory-selected mutants, or genetically engineered microbes. Members of SIM include bacteriologists, molecular biologists, mycologists, chemists, engineers, geologists, marine biologists, and plant pathologists. These scientists, engineers, technicians, administrators, and educators represent many organizations in industrial, educational, federal, and state laboratories, and are involved in research, process development, production, and quality control for a wide variety of products and services including antibiotics and other pharmaceuticals, vaccines, diagnostic tests, foods, and beverages; biocides; microbial products such as enzymes, amino acids, and carbohydrates; agricultural biotechnology; oil recovery and mining; and bioremediation. SIM sponsors an annual scientific meeting as well as special topic conferences.

In 1995, then-president-elect of SIM, Steve Orndorff, commissioned the Strategic Planning Committee with responsibility for charting a course for SIM for the next 5 years. The committee developed a statement of the mission, visions, and values of SIM as well as a strategic plan. These concepts were published for comment by the membership, and after due consideration were formally adopted by the board to become part of the society's formal charter.

Mission: To enhance and strengthen the professional activities of all those who practice, teach, or study microbiological and related sciences to solve practical or industrial problems for the benefit of humanity. We do this by providing high-quality publications, informative and congenial scientific meetings, synergistic career development interactions, practical education of our successors, and effective advocacy of relevant issues.

Vision: To be the most valued representative of professionals in our field. To be the most productive catalyst of career development interactions in our field. To be the most widely acclaimed organizer of scientific conferences in our field. To be the most respected publisher of scientific work in our field.

Values: We perform to the highest standards of professional integrity and scientific ethics. We direct our intellectual resources to benefit and protect humankind. We respect and treat fairly our employees, colleagues, employers, and suppliers. We inform and educate the public within the framework of our nonprofit status. We rely for our effectiveness on the productive volunteerism of our members. We value the benefit of interactions between scientists with common interests. We hold ourselves accountable for being responsible public citizens and scientists. We take responsibility for educating and nurturing our successors.

In 1996, the board of directors recommended to the membership to change the society's name to the Society for Industrial Microbiology and Biotechnology to reflect the careers and technologies that have developed since the 1980s. In 1998, the society membership rejected the name change (390 to 369 votes) in favor of remaining the Society for Industrial Microbiology.

The common denominator of applied or industrial microbiology is microbial science in the service of humans. In fact, the science of microbiology developed from problems associated with some microbial processes that had immediate application to humans. Microorganisms have been used for the production of food, clothing, and drink for centuries but leavening bread, fermenting wine, producing vinegar, and retting flax had been practiced in ignorance. It was not until 1864 that Louis Pasteur defined the role of microorganisms in fermentations. Industrial microbiology was slow to develop after Pasteur's initial experiments, probably because of the lack of good culture methods and selective media. In the late 1800s, Jokichi Takamine brought the Koji process for producing amylase to the United States. He received the first U.S. patent on an enzyme production process and established the first fermentation company in the United States (1). Except for a few outstanding contributions mainly involving lactic acid, acetone–butanol fermentations, and production of yeast and yeast products, industrial microbiology did not progress very rapidly until after 1900. Samuel C. Prescott initiated the first course in industrial biology in 1896 at the Massachusetts Institute of Technology. The first textbook in industrial microbiology was not published until 1940, however (2).

Industrial microbiologists in the 1940s worked with deterioration of military material, solvent fermentations, and penicillin production. As these scientists began to accumulate experimental results, there was a natural desire to share their findings with scientific colleagues at a national society meeting and in scientific journals. However, there was no central organization in which to share their findings because these scientists came from a variety of scientific specialties including plant pathology, medical laboratories, materials engineering, and academic mycology and bacteriology. Moreover, it was difficult for a purely practical paper in microbiology to secure a place on the program of existing societies or to find an opportunity for publication in the established journals. These scientists became unhappy, and a conflict developed between basic and applied microbiology practitioners. In 1951, the Society of American Bacteriologists (SAB) authorized a new journal called *Applied Microbiology*. However, the first issue was not published until January 1953.

The decision to publish *Applied Microbiology* took too long in the minds of the industrial microbiologists of the

late 1940s. By that time, industrial microbiologists felt that applied and industrial aspects of microbiology were being neglected by the SAB. Walter Ezekiel, Bureau of Ordnance, Department of the Navy (Washington, D.C.) wrote to and spoke with a number of individuals during the summer and autumn of 1949 and suggested formation of a new Society for Industrial Microbiology. Ezekiel arranged with Raymond L. Taylor, associate administrative secretary of the American Association for the Advancement of Science (AAAS), to schedule and advertise a meeting of interested persons during the AAAS meeting in New York (3). Ezekiel published a short announcement of the meeting in *Science*, in which he stated that "a new society of industrial microbiologists is to be organized to cover the borderline work in application of microbiologic knowledge or processes to work with industrial materials. The society plans to handle . . . microbiologic manufacturing processes and microbiologic assay. The organization meeting will be held December 29 at 4:00 P.M. at the Hotel McAlpine Ballroom" (4).

On that date, more than 250 interested persons met in the ballroom of the Hotel McAlpine in New York City. There, it was decided to form a new society to affiliate with the newly formed American Institute of Biological Sciences (AIBS). Charles Thom, Northern Regional Research Laboratory (Peoria, Ill.) was elected acting chair and Charles L. Porter, Purdue University (West Lafayette, Ind.), acting secretary. Ezekiel's motion to establish a scientific organization known as the Society for Industrial Microbiology, with its general scope the field of microbiology as applied to industrial materials and processes, was passed. It was decided at that meeting, over the objections of a few, that the group would meet the following year, again with AIBS. Thom appointed the following members of the Organizing Committee: Maynard M. Baldwin (Battelle Memorial Institute, Columbus, Ohio), F. G. Walter Smith (University of Miami Marine Laboratory, Coral Gables, Fla.), Earnest A. Walker (U.S. Department of Agriculture, Washington, D.C.), William L. White (Farlow Herbarium, Harvard University), Ezekiel, and Porter (5).

The committee met several times during the first year to prepare a program and write a constitution for the new society in preparation for the first annual meeting to be held with AIBS at Ohio State University, September 11–13, 1950. In Ohio, Thom was elected the first president and Porter the secretary-treasurer. In 1951, the new SIM met with AIBS at the University of Minnesota and Thom was elected president for a second term. In 1952, the meeting was held at Cornell University.

The need for an interdisciplinary society first conceived by Ezekiel became apparent, and membership grew at a rapid pace. In 10 years, membership in the new society had grown from 75 to 550. SIM was incorporated as a nonprofit organization in 1960 in Washington, D.C. Membership reached 2,000 in 1989. In 1989, Paula Myers-Keith was the first woman elected president of SIM. For many years, the annual meetings were held along with AIBS on university campuses. In 1967, SIM held its first independent annual meeting, at the University of Western Ontario. Since 1985, all annual meetings have been held at commercial convention centers rather than universities. The annual meeting is a weeklong event featuring plenary sessions, roundtables, poster sessions, workshops, and extensive technical exhibits.

Between 1960 and 1970, SIM sponsored summer institutes on microbiological problems of current interest with topics including industrial microbial genetics and federal regulations for disinfectants, drugs, and cosmetics. In 1960, SIM sponsored the initial conference on antimicrobial agents. This was a mammoth undertaking for the small organization, and sponsorship was subsequently released to the American Society for Microbiology (Interscience Conference on Antimicrobial Agents and Chemotherapy, ICAAC) (6). In 1978, the Alma Dietz actinomycete roundtable discussions were included and have since become a part of all subsequent annual meetings.

In 1987, reaffirming its commitment of service to the microbial biotechnologist, SIM approved the organization of the Special-Topic Conference Series. These conferences run 2 to 3 days and usually result in the production of a monograph. Initiated by George Somkuti (president, 1985–1986), the first event in the series was the International Conference on the Biotechnology of Microbial Products: Novel Pharmacological and Agrobiological Activities (BMP), held March 13–16, 1988. BMP is held every 3 years. Since then, SIM has sponsored or cosponsored special conferences on a regular basis. Other conference series include Recent Advances in Fermentation Technology (RAFT), initiated in 1995. RAFT is held biannually and is cosponsored with the Biochemical Technology Division of the American Chemical Society. The aim is the free exchange of newly developed fermentation technologies related to current highly productive, consistent, scaleable, and economic processes. Fermentation microbiology, biochemistry, biocatalysis, and biochemical engineering aspects are explored. The Sixth Genetics and Molecular Biology of Industrial Microorganisms (GMBIM) Conference, cosponsored with Indiana University Institute of Molecular and Cellular Biology, was held in 1996 at Indiana University. This continuing series provides a forum for academic and industrial scientists to exchange information on the latest developments in the genetics and biochemistry of microbes of industrial interest.

The business of SIM is conducted by the board of directors (which is elected by the membership), 12 standing committees, and the executive secretary. The board consists of the president, president-elect, secretary, treasurer, past president, and four directors. In 1959, as president of SIM, Porter said, "There are a number of organizations including the Society of American Bacteriologists and the American Chemical Society which have grown so large that they have become unwieldy. The administration of these groups has lost personal contact with the membership . . . They must enforce rules and regulations which hamper individual initiative, incentive, and recognition. At their meetings the sessions devoted to volunteer papers are so clogged that each presentation is limited to 10 or 12 minutes. . . . Membership beyond 1200–1500 should be discouraged, for then, we like other groups mentioned, would lose that personal contact which is one of the principal attractions" (7).

At present, most of the membership is in the United States, but interest in industrial microbiology is international and SIM has more than 200 members in 38 countries outside the United States. Employment opportunities in industrial microbiology and biotechnology have increased dramatically since 1959; consequently, SIM now has a membership of more than 1,900, in addition to 75 corporate members. However, SIM continues to foster an environment in which individuals are recognized and have opportunities to participate. Modern communications including the Internet are available so that members have fast and easy access to society resources via electronic mail and a World Wide Web home page. At the annual meetings, speakers give 30-minute presentations of their work. Additionally, there are numerous opportunities for members to serve on committees and become officers. Local sections provide more opportunities for scientists, engineers, technicians, and students to meet and form personal contacts.

AWARDS

SIM has established the Committee on Awards and Honors to provide monetary and honorary awards to its members and others eminent in the field. Awards are made at the annual meeting. In 1967, the first Charles Thom Award was presented to Kenneth B. Raper (president, 1952–1953). This award is bestowed annually to persons making outstanding contributions in the field of industrial microbiology. The second Thom Award was presented to Arthur M. Kaplan (president, 1960–1961) at the 1970 annual meeting held at the University of Rhode Island. The Charles Porter Award of Merit is given in recognition of outstanding contributions to the growth and success of SIM. The first Porter Award was given to Earnest A. Walker in 1960.

A fellowship status was approved by the SIM membership at the 1984 Annual Business Meeting. Fellowship is a membership grade of distinction in the Society acknowledging significant research and/or service contributions to the professions of applied microbiology. A sustained record of such contributions while a member of SIM is the main criterion for consideration to fellowship. No more than 10% of the members are eligible for fellowship status at any given time. The first group elected as SIM fellows in August 1985 were Arnold L. Demain (president, 1990–1991), Arthur E. Humphrey, Richard P. Elander (president, 1973–1974), Warren P. Iverson, Clifford W. Hesseltine (president, 1957–1958), H. Boyd Woodruff (president, 1954–1956), Leland A. Underkofler (president, 1967–1968), Kaplan, and Raper.

The Selman A. Waksman Teaching Award was established in 1989, and the first recipient was Douglas E. Eveleigh. The award is presented to a recipient every other year at the annual meeting. The nominee must be an active, full-time professor at a recognized institution of higher education for a minimum of 10 years or have attained emeritus status. The nominee must be actively involved in research in an area of industrial or applied microbiology or biotechnology. Annually, SIM present the Student Abstract Award for the best abstract submitted for the annual meeting.

The Panlabs Award Lecture is sponsored by Panlabs and is presented by individuals of world renown in the field of microbiology. Each Panlabs Lecturer presents an address at the annual meeting. The Schering–Plough Research Institute Young Investigator Award is made to a society member who is under 36 years of age at the time of nomination. The nominee must have made a significant contribution in industrial microbiology or biotechnology that indicates promise of a professional career of merit. This award is made on alternate years with the Waksman Teaching Award.

LOCAL SECTIONS

Regional groups called local sections are governed by their own constitutions and elected officers. Local sections provide an opportunity for individuals involved in biotechnology to meet, discuss topics of interest, attend talks given by noted industrial and academic speakers, and attend local symposia on important contemporary subjects. Membership includes all interested microbiologists, chemists, and engineers from students to advanced industrial, academic, and government professionals.

PUBLICATIONS

SIM News is the official publication of the society. It was first published in 1951 and entitled *SIM News Letter*; the name was changed in 1972. *SIM News* is distributed bimonthly to all members. It contains technical review articles, news of special significance in the field of applied microbiology, and reports of many activities of the society including local sections reports, placement opportunities, and meeting notices.

The first volume of *Developments in Industrial Microbiology (DIM)* was published in 1960. This peer-reviewed annual publication contained papers from the annual meeting. *DIM* was published annually for 30 years (31 volumes).

The *Journal of Industrial Microbiology and Biotechnology (JIMB)* is an international, peer-reviewed journal that seeks to further scientific knowledge and to disseminate information in biotechnology, fermentation and industrial microbiology, biodegradation, biodeterioration, quality control, and other areas of applied microbiology. The concept for a new technical publication for the society was presented by SIM President C. Herb Ward (1983–1984), and the first issue of *JIMB* was published in March 1986 as the *Journal of Industrial Microbiology (JIM)*. The first editor in chief was George E. Pierce. The name *JIM* was used continuously through 1996, when the board of directors adopted the new name, *JIMB*. *JIMB* will consider publication of original research papers, critical reviews, short communications, and letters to the editor. All papers submitted are subject to peer review.

The *Developments in Industrial Microbiology Series* complements *JIMB* and *SIM News* by providing in-depth reviews of specific subject areas. The monographs are focused on topics of interest to scientists in industrial and environmental microbiology. Authors for the individual

series are identified from presentations at special conferences sponsored by SIM and selected symposia at the annual meeting. The address for SIM is 3929 Old Lee Highway, Suite 92A, Fairfax, VA 22030-2421. Internet: ⟨*info@simhq.org*⟩⟨*www.simhq.org*⟩.

BIBLIOGRAPHY

1. J.W. Bennett, *Takamine: Documents from the Dawn of Industrial Biotechnology*. Miles, Berkeley, Calif., 1988.
2. G. Reed, *Prescott and Dunn's Industrial Microbiology*, 4th ed., AVI Publishing, Westport, Conn., 1983.
3. W.N. Ezekiel, in *Developments in Industrial Microbiology*, vol. 8, Society for Industrial Microbiology, Washington, D.C., 1967, pp. 431–434.
4. Announcements, *Science* **110**, 674 (1949).
5. R.M. Rogers, *ASM News* **41**, 642–647 (1975).
6. L.A. Underkofler and R.D. Schwartz, *SIM News* **42**, 260–262 (1992).
7. C.L. Porter, in *Developments in Industrial Microbiology*, vol. 1, Society for Industrial Microbiology, Washington, D.C., 1960, pp. 261–267.

PROFESSIONAL SOCIETIES, THE PROTEIN SOCIETY

ROBERT NEWBURGH
The Protein Society
Bethesda, Maryland

Protein science is the study of proteins. The development of new techniques, including those from molecular biology, analytical chemistry, immunology, spectroscopy, diffraction, and computations, has provided opportunities to probe structural and functional relationships of proteins in detail undreamed of until recent times. These opportunities underscore the importance of broad interactions among investigators of diverse backgrounds to resolve fundamental questions. The Protein Society was established in 1986 to promote these interactions by providing regular opportunities for protein scientists to exchange ideas, findings, and techniques via both scientific meetings and the society's journal, *Protein Science*.

The goal of the society is to promote the discipline of protein science by these means:

1. Promoting research and scientific endeavor in the field of protein science
2. Promoting education in protein science to a wide variety of constituencies including established investigators, young scientists, students, others interested in science, and physicians interested in diseases pertinent to the field
3. Promoting and extending current understanding of the importance of protein science, particularly in regard to the methods, concepts, and problems of structural biology, cell and molecular biology, and biotechnology, as they relate to the behavior and properties of proteins.

The society attempts to foster effective interactions between investigators who share these interests. To further this objective, the society seeks to expand its membership and activities in ways that will enhance interpersonal communication among scientists.

To achieve these goals, the society invests its resources in

1. Management of the Society
2. Conduct of scientific meetings
3. Publication of scientific research
4. Educational activities
5. Fostering relations with scientists in other countries interested in protein science

Further information regarding the society can be obtained from the executive officer, Dr. Robert Newburgh, The Protein Society, 9650 Rockville Pike, Bethesda, MD 20814-3998 (phone 301-571-0662; Fax 301-571-0666; e-mail newburgh@protein.faseb.org).

PROTEIN ADSORPTION, EXPANDED BED

SIDDARTHA GHOSE
Aston University
Birmingham, United Kingdom

KEY WORDS

Adsorbent
Chromatography
Expanded bed chromatography
Expanded beds
Fluidized beds
Whole-broth adsorption

OUTLINE

INTRODUCTION

Traditional biochemical processes can be characterized into three major operational divisions: *(1)* the upstream process, which deals primarily with preparation and pre-

treatment of the raw material from which the specific target product is to be directly or indirectly obtained, *(2)* the actual production phase, in which the chemical transformation of the raw material takes place to produce the end product in which the desired target product is obtained along with associated by-products, and *(3)* the downstream processing, which is specifically involved with the recovery and purification of the target product. Progress in the area of recombinant DNA technology has made it commonplace to clone and express specific products for diagnostic and therapeutic applications within a variety of host microorganisms. This advancement has resulted in a pace of development of production phase technology that has far surpassed that of the recovery and purification areas. Recent years have seen significant advances in this latter area, and, with the incorporation of affinity-based interactions, it is now possible to include more powerful and specific operations for improving the overall efficiency of the recovery process (1).

Downstream processing refers to the collective set of operations involved in obtaining the target product in its required end form. The name derives from its relative position to the actual production stage in the process flow sheet. Factors such as the starting concentration of the target product, its final required purity, and the type of impurities to be removed dictate the exact choice of operations, and the sequence is necessarily adapted for each particular process. The use of microorganisms in biochemical processes makes cell harvesting, disruption (for the product residing within the cytoplasmic spaces as an intracellular product), and solid–liquid separation essential initial stages in the sequence of operations. Processes in which the product is expressed as an extracellular component require no cell disruption stage, although clarification of the feedstock from whole cells or cellular debris is always essential before subsequent treatment. The common solid–liquid separation methods traditionally used are centrifugation (2,3) and cross-flow microfiltration (4). The use of aqueous two-phase systems for primary treatment of feedstock (5) and cell debris removal has also been reported (6–9).

Biochemical product recovery nearly always involves one or more packed bed chromatographic processes as an often essential step of the separation protocol, usually in the secondary and final purification stages (10). It has now developed into a powerful tool for product separation at the analytical and preparative (11) scale of operation. The wide variety of modes of interaction between the solubilized product and ligands on the solid adsorbent (11–13), ease of scale-up and cleaning-in-place procedures, commercial availability of equipment and consumables, and the well-documented operation have given this technology wide acceptance in academic and commercial research and development. There is now a significant amount of information on the optimization of chromatographic separations (14). However, full utilization of its potential implies clarification of the biological feedstock before application onto a packed bed, failing which bed blockage is inevitable.

This article describes in some detail the use of expanded bed chromatography, which has been described as the only new unit operation that the bioprocess industry has seen for decades (15). This technique offers the possibility of directly applying the unclarified feedstock onto the bed, bypassing the conventional solid–liquid separation techniques while simultaneously capturing the desired product from the liquid phase. A typical sequence of separation operations is shown schematically in Figure 1, in which the relative position of expanded bed chromatography is also indicated. As will become clear from the description of its operation, the incorporation of this unit will enable facilities to replace cost-intensive and often problematic solid–liquid separation techniques such as centrifugation and cross-flow microfiltration. In addition, there is the added benefit of a chromatographic step early in the separation protocol. One must bear in mind that this technology is still in its development phase, and it is too early yet to conclude that it will become an integral part of a downstream processing protocol. However, as witnessed from the Second International Conference on Expanded Bed Adsorption recently held in the United States, there has been a growing awareness of the potential of the technology and it may indeed prove to have a promising future in bioseparations.

Before a detailed description of various aspects of expanded bed systems, a point of explanation should be made

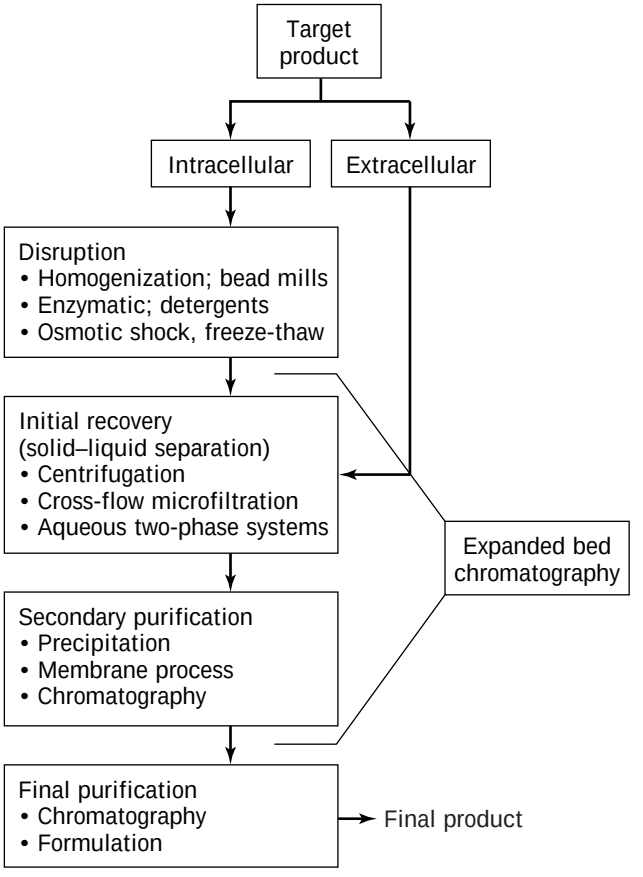


Figure 1. Typical sequence of operation in the recovery and purification of proteins from biological feedstock. The relative position of expanded bed as a viable alternative operation is also indicated.

regarding the use of the terms *fluidized bed* and *expanded bed*. During the initial development of this technology, the terms were often used exclusively of one another. In classical chemical engineering operations, fluidized beds were associated with catalytic reactors in which a high degree of solid–fluid mixing was maintained to ensure good heat and mass transfer. Expanded beds, on the other hand, would benefit product capture if good plug-flow-type hydrodynamics could be obtained. Clearly mixing needed to be minimized and bed stability ensured. The two terms were clearly in contradiction with each other; to maintain the existence of the critical difference, the use of the term *fluidized beds* was avoided in the context of biological separations by *expanded beds*. Over the years, however, this distinction has eroded, and now both are interchangeably used to mean essentially the same thing. The same applies in this article.

THEORY

To understand the physical process by which an expanded bed is obtained and some of the basic expressions describing its various states, we begin with a settled bed of adsorbent packed in a column with no top flow adapter. When an upward flow of liquid is introduced at the lower end, there is a tendency for the adsorbent particles to move in the direction of flow. This movement occurs only when the liquid velocity exceeds the minimum fluidization velocity of the particles. The bed then undergoes a transient expansion phase, at the end of which the adsorbent particles attain a steady position when the upward force due to the liquid flow exactly counterbalances the downward gravitational force on the particles. The bed is then said to be fluidized and will remain in this state as long as the two forces counterbalance each other.

When the fluid velocity equals the minimum fluidization velocity of the particles, the bed is said to be in a state of incipient fluidization. At this point the bed can be considered packed for the purpose of estimating the pressure drop across it and fluidized for calculating the buoyant weight of the particles. By using the Ergun equation for pressure drop and equating the two forces, we obtain an expression for the minimum fluidization velocity:

$$150\,\frac{(1-\varepsilon_{mf})^2}{\varepsilon_{mf}^2}\cdot\frac{\mu\cdot u_{mf}}{d_p^2}+1.75\,\frac{1-\varepsilon_{mf}}{\varepsilon_{mf}^2}\cdot\frac{\rho u_{mf}^2}{d_p}=(1-\varepsilon_{mf})(\rho_p-\rho)g \quad (1)$$

The pressure drop term on the left-hand side of equation 1 has two parts; the first term denotes the contribution due to laminar flow, and the second is due to turbulent flow. Depending on the flow regime, one could simplify the expression by neglecting the appropriate term. In the laminar flow region, the minimum fluidization velocity can be derived from equation 1 as

$$u_{mf}=\frac{\varepsilon_{mf}^2}{150(1-e_{mf})}\cdot\frac{1}{\mu}\,d_p^2(\rho_p-\rho)g \quad (2)$$

The physical parameters that affect the velocity at which a settled bed of particles begins to fluidize can be seen from equation 2 and are the viscosity of the fluid, the particle size of the adsorbent beads, and the density difference between the adsorbent and the liquid. An accurate estimate of the bed voidage at the point of minimum fluidization is also important for estimating u_{mf}. For spherical particles, 0.4 is always a good first approximate valid when the packing is not overly compact. Wen and Yu (16) have found a value of 0.42 to be satisfactory.

Once the bed is in a fluidized state, a simple relation exists between the bed voidage and the fluid velocity. This relation is valid over a given range of the latter, and the most commonly used relationship is that described by Richardson and Zaki (17), which relates the superficial fluid velocity, the terminal settling velocity of the particle, and the bed voidage. The relationship is expressed as

$$\frac{u}{u_t}=\varepsilon^n \quad (3)$$

where n is the Richardson–Zaki coefficient. The determination of n and u_t is usually obtained from a logarithmic plot of experimentally determined values of the bed voidage and the corresponding flow velocity and forms a fundamental part of the study of these systems.

However, bed voidage is not an easily measured parameter and would in turn require further experimentation. A more convenient approach is to use the relationship between the bed voidage at a given bed expansion, ϵ, in terms of the expanded bed height, H_{exp}, the initial (or settled) bed height, H_o, and the voidage of the settled bed ϵ_o. Thus we have

$$\varepsilon=1-\frac{H_o}{H_{exp}}\cdot(1-\varepsilon_o) \quad (4)$$

The ratio H_{exp}/H_o is referred to as the degree of bed expansion. Experimentally, one would measure the stable expanded bed height, H_{exp}, at a given flow velocity through the column (18–20). The data can then be plotted to obtain n and u_t for the bed. Analysis of equations 3 and 4 shows us that a high u_t for a particle results in a lower bed expansion and vice versa.

The terminal velocity is the upper limit of validity of the Richardson–Zaki expression, indeed of fluidization in general. At a fluid velocity equal to the terminal velocity of the particles, the particles are no longer suspended in the fluid but travel along with it. In the context of expanded bed chromatography, this leads to what is known as elutriation, where the adsorbent particles are physically removed from the column. Clearly, the latter phenomenon is highly undesirable, and it is therefore essential that for effective operation of expanded beds, the velocity range lies well within the limits of the minimum fluidization velocity at the lower end and the terminal velocity at the higher. The value of u_t can also be determined theoretically from Stokes law; thus

$$u_t=\frac{d_p^2(\rho_p-\rho)g}{18\mu} \quad (5)$$

The coefficient n has been estimated by Richardson and Zaki (17) to be a function of both the particle to column diameter and the Reynolds number. The correlations are valid for uniform spherical particles $>100\ \mu$m in size. Garside and Al Dibouni showed that a different set of correlations are required when the fluidized bed voidage is greater than 0.85. Two expressions were proposed by them—one general type and a second similar to the Richardson–Zaki type of formulation. Both are valid over the entire range of Reynolds number. Some common expressions developed by different workers for estimating the values of n and u_t are shown in Table 1. Values determined experimentally or through the correlations are used for comparative evaluation of different adsorbents for their potential use as fluidized bed adsorbents. The equations are a simple and effective starting point for evaluating some of the fundamental characteristics of an expanded bed.

PRINCIPLES OF OPERATION

The operational procedure for protein recovery using an expanded bed is similar in principle to that of a conventional packed bed column and is schematically depicted in Figure 2. As mentioned in the previous section, the initial step involves formation of the stable expanded bed from its settled or packed state. This is done by first moving the top flow adapter to a sufficient height to permit free movement of the top of the bed in the upward direction of liquid flow. When flow of the equilibrating buffer is initiated, bed expansion occurs and the top of the bed rises. The bed height increase is usually characterized by a linear phase followed by a progressively reducing nonlinear rate of change until a final stable height is achieved. This expanded height depends on a number of physical factors of the system, which are described in the next section. Once a stable bed has been obtained, the top flow adapter is usually repositioned to lie anywhere between 1 and 15 cm above the top of the bed. The actual position depends on the mode of operation employed (see later section). With the top adapter in position, the bed is essentially ready for initiation of feedstock load.

The flow is then switched from the buffer to the unclarified feedstock, usually by an appropriate valve mechanism. During the feedstock load, the solid matter, consisting of cell debris or sometimes whole cells, travels directly through the column while some of the solubilized products are adsorbed onto the column. Ideally there will be no retention of the particulate debris within the column, but in practice one could well observe a degree of nonspecific binding to the adsorbent particles. This issue becomes of particular relevance during cleaning-in-place and column regeneration. Feed load is usually terminated when the outlet concentration of the product being recovered is approximately 5–10% of its concentration in the feed. At this point the inlet flow is then switched back to the equilibrating buffer to wash out the residual particulate matter from the expanded bed.

When the outlet liquid is completely clear of insoluble material, the product can be eluted from the column, most commonly in the packed mode and after the bed is first allowed to settle by terminating the flow of liquid. The top flow adapter is then lowered onto the settled bed, bringing it to a packed mode. The bound products are then eluted by the appropriate elution buffers, usually in a single step because specific separation is not the primary objective at this stage of the recovery process. When elution is complete, the column is regenerated and where required cleaned. The bed is then expanded with the equilibration buffer and made ready for the next cycle of feedstock load.

Bed Stability

The sequence of steps in the operation of an expanded bed is relatively straightforward and simple to incorporate. As

Table 1. Fluidization Parameters for Uniform-Sized Spherical Particles

Reference	Validity range	Equations
Richardson and Zaki (17)	100–6,350 d_p (μm) 1.06–11.25 ρ_p (g/mL) 0.81–2.89 ρ_1 (g/mL) 10^{-3}–113.10^{-3} μ (Ns/m^2)	$\frac{u}{u_t} = \varepsilon^n$ where $n = 4.65 + 20\frac{d_p}{d_c}$ (Re < 0.2) $n = \left(4.4 + 18\frac{d_p}{d_c}\right)\mathrm{Re}^{-0.03}$ ($0.2 <$ Re < 1) $n = \left(4.4 + 18\frac{d_p}{d_c}\right)\mathrm{Re}^{-0.01}$ ($1 <$ Re < 200) $n = 4.4\mathrm{Re}^{-0.1}$ ($200 <$ Re < 500) $n = 2.4$ (Re > 500)
Garside and Al Dibouni (xx)	167–3,070 d_p (μm) 2.8 ρ_p (g/mL) 1.0 ρ_1 (g/mL) 10^{-3} μ (Ns/m^2)	$\frac{(u/\varepsilon \cdot u_t) - A}{B - (u/\varepsilon u_t)} = 0.06\mathrm{Re}^{(\varepsilon+0.2)}$ where $A = \varepsilon^{4.14}$ $B = 0.8\varepsilon^{1.28}$ for $\varepsilon \leq 0.85$ $B = \varepsilon^{2.65}$ for $\varepsilon > 0.85$
Wen and Yu (16)	200–350 d_p (μm) 2.37–7.84 ρ_p (g/mL) 1.0 ρ_1 (g/mL) 10^{-3} μ (Ns/m^2)	$\varepsilon^{4.7}\left(\frac{d_p \cdot g \cdot (\rho_p - \rho_1)\rho_1}{\mu^2}\right) = 18\mathrm{Re} + 2.7\mathrm{Re}^{1.687}$ $0.01 <$ Re $< 10^4$

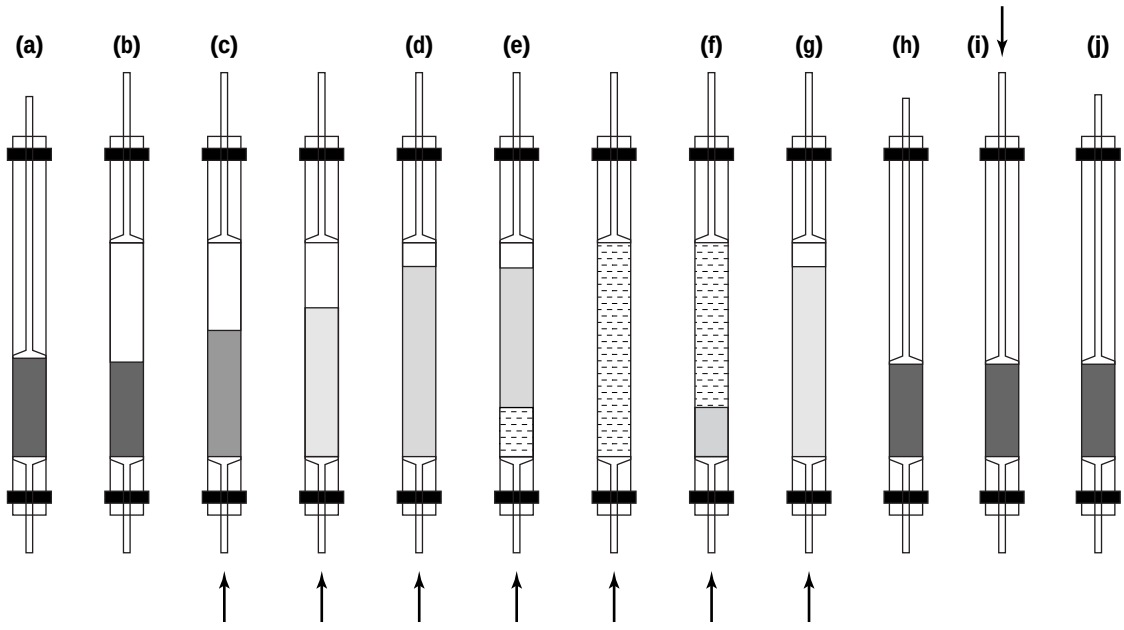


Figure 2. The various stages of operation of expanded bed chromatography from start-up to column regeneration. **(a)** Starting with a settled bed of adsorbent, **(b)** the adapter is moved up to remove any limitations on the movement of the bed. **(c)** Liquid flow through the column is then initiated from the bottom in the upward direction, during which time the bed undergoes a transient phase until **(d)** it finally reaches a stable steady-state height. The bed is equilibrated with buffer and **(e)** feedstock load initiated until the column is saturated, as seen from the breakthrough of bound product from the column. **(f)** The load is terminated and buffer wash started until **(g)** all the homogenate is washed out from the column. **(h)** The flow is then terminated, the bed allowed to settle, and the flow adapter lowered to attain a packed bed configuration, after which **(i)** the products are eluted, usually in the reverse direction. **(j)** The column is then cleaned and regenerated and made ready for the next run.

described earlier, the first step is obtaining a stable expanded bed; although not immediately obvious, this is a critical requirement with implications for overall process efficiency. A stable bed can be defined as one in which the liquid–liquid and solid–liquid mixing is low. Bed stability becomes a necessary requirement for good plug-flow-type hydrodynamics, which in turn affects the product binding characteristics. This is best quantified through residence time distribution (RTD) studies by analyzing the outlet profile of a tracer pulse in exactly the same manner in which it would be done for a packed bed system. Tracers such as acetone are effectively used in the absence of any particulate material. However, RTD experiments conducted in the absence of such particulates only indicate the state of the bed before feedstock load is initiated. Unlike packed bed chromatography, in which the hydrodynamics of the bed can be assumed to be more or less the same throughout the load phase, expanded beds are more than likely affected to a certain degree by the presence of the particulate matter. Although the influence need not necessarily be significant, it is nevertheless an important point to remember when conducting stability tests in these systems.

The tracer outlet profile can be evaluated by the moments rule to obtain an estimate of the Peclet number or the vessel dispersion number (21), defined as

$$\mathrm{Pe} = \frac{u}{D_{\mathrm{ax}} \cdot L} \tag{6}$$

where D_{ax} is the liquid axial dispersion coefficient. Values of Pe greater than 40 are considered acceptable in terms of the flow hydrodynamics. The Peclet number is used interchangeably with the Bodenstein number, defined in the same manner and used extensively by Karau et al. (20) and Thommes et al. (22).

For the experienced user of expanded bed chromatography columns, the power of visual observation, although highly qualitative and impossible to validate, can often prove to be sufficient in ascertaining the extent of bed stability. Liquid–liquid (and thus solid–liquid) mixing is often a result of radial flow inequalities. Some of the common causes are poor distribution at the inlet to the column (which often transmits through the full length), inaccurate vertical alignment of the column (which is particularly important in scale-down systems), and aggregation of the adsorbent particles (which leads to significant variation in the bed voidage along the length of the column and thus affects the flow pattern). The consequential gross flow inadequacies can be more than sufficiently detected by the naked eye, and although an accurate determination of the Peclet number is not possible, visual observation often gives sufficient indication when a bed has yet to reach a

stable configuration. Experience clearly plays a significant role in the use of visual observation as a tool for determining bed stability, and it has only limited applicability. Nevertheless, it may often prove sufficient when a rigorous residence time distribution experiment is not essential.

Stability of expanded beds is achieved through the inherent particle size distribution of the adsorbent. From the theory of fluidization outlined in the previous section, smaller particles fluidize more readily than larger ones in a given liquid stream, assuming that the particles are uniformly dense. For fluid velocity ranging between the minimum fluidization velocity of the smaller particle to the terminal velocity of the larger one, different situations arise, as depicted in Figure 3. As the flow velocity is increased such that $u_{mf,large}$, $u_{t,large}$, $u_{t,small} > u > u_{mf,small}$ the smaller particles fluidize and attain a steady-state position somewhere along the length of the column. Continual increase in the flow until $u_{t,large}$, $u_{t,small} > u > u_{mf,small}$, $u_{mf,large}$ results in both the small and large particles being fluidized. However, the increased flow velocity requires a longer column length for the inertial forces to drop to the level to be sufficiently counterbalanced by the buoyant force of the smaller particles. It therefore attains its steady state higher up along the column relative to both its initial position and that of the larger particle. A further increase such that $u > u_{mf,large}$, $u_{t,small}$ would wash the small particles out from the column and move the steady-state position of the larger particles higher up. Finally, when $u > u_{t,large}$, the larger particles are also elutriated.

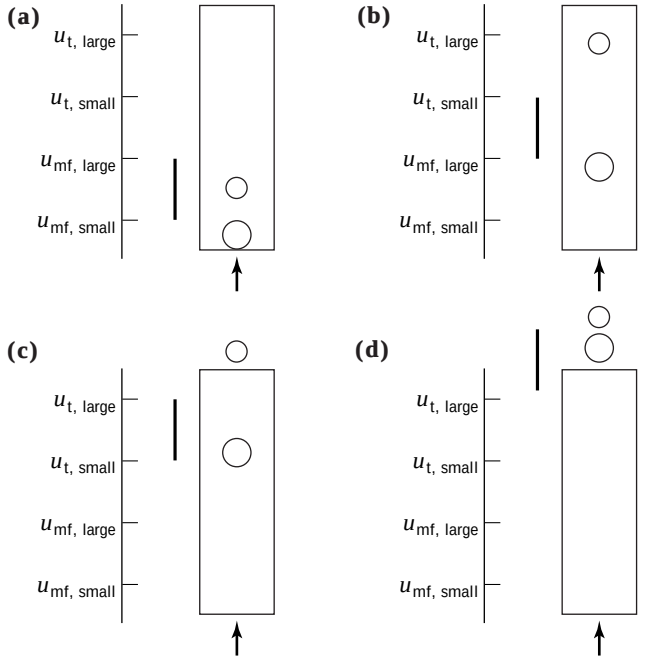


Figure 3. Schematic diagram showing the effect of flow velocity on the fluidization of two particles of different sizes. **(a)** As the flow velocity is increased such that $u_{t,large}$, $u_{t,small}$, $u_{mf,large} > u > u_{mf,small}$ the smaller particles fluidize and attain a steady-state position somewhere along the length of the column. **(b)** Continual increase in the flow until $u_{t,large}$, $u_{t,small} > u > u_{mf,large}$, $u_{mf,small}$ results in both the small and large particles being fluidized. The increased flow velocity, however, requires a longer column length for the inertial forces to drop to the level to be sufficiently counterbalanced by the buoyant force of the smaller particle. It therefore attains its steady state higher up along the column relative to both its initial position and that of the larger particle. **(c)** A further increase such that $u_{t,large} > u > u_{t,small}$, $> u_{mf,large}$, $u_{mf,small}$ would wash the small particles out from the column and move the steady-state position of the larger particles higher up. Finally, when $u > u_{t,large}$ the larger particles are also elutriated. The vertical line shows the flow velocity range on a relative scale between the limits.

Extending the preceding discussion to an expanded bed of adsorbent with a defined particle size distribution means that the particles will be stratified along the length of the column. This is schematically depicted in Figure 4 for a bed of adsorbent in the settled and stable expanded states. In the absence of gross flow inequalities, each particle remains in its steady-state position, with the larger ones at the bottom or inlet to the column and the smaller ones near the top (23). This configuration can be maintained indefinitely. In reality there is localized movement of the particles, but it contributes little to liquid–liquid mixing within the system.

The other most notable method for stabilizing fluidized beds is through the use of magnetic fields (24–27). A magnetic material is incorporated within the adsorbent beads, and the superimposition of the magnetic field acts to limit the movement of the adsorbent particles. The flow of particulate-laden feedstock can often be conducted at relatively higher flow velocities, enhancing the efficiency of its wash through the column.

Modes of Operation

Although stability has been discussed in terms of liquid–liquid and solid–liquid mixing, bed height movement must be considered. Because the top of the bed is not restricted in any way, changes in the flow or physical properties of

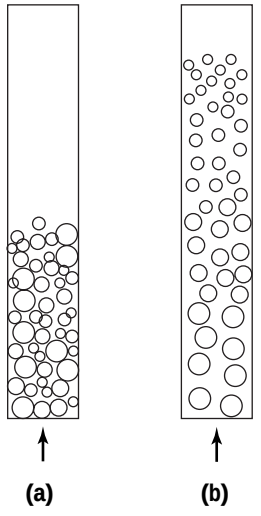


Figure 4. Schematic diagram showing the stratification of particles based on size along the length of the column going from **(a)** the randomly distributed settled bed state to **(b)** the stable expanded one.

the inlet liquid directly, although not always instantly, influence the bed height. Thus, when switching between buffers of different viscosities, densities, or both, the bed exhibits a concomitant movement to achieve a stable height corresponding to the new physical conditions to which it has been subjected. This movement is also seen when feedstock load onto a stable expanded bed is initiated. The significantly different physical properties of the particulate containing feedstock from the equilibrating buffer implies that on switching to feedstock, the bed height will change to a new stable position, assuming that the flow rate through the entire operation is kept constant. Consequently, there are two distinct possibilities for modes of operation of the bed:

1. In a *constant flow rate* mode, the bed is operated with the flow rate maintained at its initial value, allowing the bed height to vary if necessary. Because biological feedstocks are in general more viscous and dense than the standard buffers commonly used in biochemical processes, switching from buffer to feedstock load is manifested in terms of the expanded bed height increasing through an initial transient phase to a new position. This increase is further influenced by the presence of particulate matter.Two operational implications arise as a result. Firstly, the position of the top flow adapter needs to be managed appropriately, either through continuous repositioning during the course of the run or by selecting an initial position significantly higher than the top of the bed before the feedstock load. In this mode the gradually increasing bed voidage during the course of feedstock load would be beneficial in the context of solids removal. Protein adsorption could be adversely affected, however, because of increasing mass transport distances, likely making it a critical rate limiting step in the process. The latter effect would be more pronounced for high molecular weight proteins.
2. In a *constant bed height* mode the bed is maintained at its initial expansion, and any tendency for the bed height to change is corrected by appropriate adjustment in the flow rate. An obvious advantage is that there is no need to adjust the position of the top flow adapter, simplifying equipment design. It is imperative, however, to be able to monitor the top of the expanded bed and any changes that it undergoes to effectively control flow rate.Protein adsorption in this case would perhaps be more efficient but would be offset by slower solids removal because of the progressive reduction in the flow rate to maintain the bed height at its initial level. The extent to which flow rate adjustment is required depends on the percentage of solids in the feedstock, with their removal being hindered at the higher solid content.

EQUIPMENT

Because of the nature of the feedstock typically used in expanded bed chromatography applications, the equipment varies from that used in conventional packed bed chromatography, although in principle the equipment undergoes only minor modifications. Certain design considerations with respect to the column and flow distributor would need to be satisfied, as discussed in the following sections.

Column

Expanded bed chromatography is a low-pressure operation that requires minimal modification of conventional packed bed columns. A principle design consideration is the inlet flow distributor, which must be able to handle particulate material in the feedstock being loaded onto the column. This aspect is discussed in detail later. Another point of practical relevance is that the column must be of sufficient length to handle the desired degree of bed expansion. Two- to threefold bed expansions are typical; for a settled bed height of 20 cm, this translates to a minimum of 60-cm column length. In practice, columns are usually 1 m long so that the top flow distributor can be suitably positioned to allow for an increase in bed expansion when unclarified feedstock is being processed in the column.

The commercially available columns now range from 2.0 to 1,200 cm in diameter. The first column to be marketed was the STREAMLINE 50™, which is extensively used by researchers today. Given the potential for commercial application, the subsequent generation of columns was scaled up to 20, 60, and 120 cm in diameter. Although useful for process scale operations, the size clearly limited their use for methods development and process validation work. Scale-down of the expanded bed columns followed, and column sizes of 2.0 and 2.5 cm are now available. I conducted the first scaled-down validation, and data from columns sizes of 0.5, 1.0, 5.0, and 25.0 cm in diameter showed excellent correlation in expansion properties and protein adsorption breakthrough.

Table 2 shows the major suppliers of the expanded bed chromatography columns and available specifications. Included in the table is the small glassware company that tailor-made the small-diameter columns I used in the scaled-down studies described earlier. These columns are fitted with a sintered glass disk as the inlet and outlet flow distributor. Columns up to 20 cm in diameter are constructed of glass, and larger ones are made from stainless steel. Glass columns have the advantage of allowing visualization of the process. This observation is particularly useful when column blockage results in excessive back pressure, reducing the flow through the column and causing the bed to rapidly settle.

Flow Distributor

The importance of the flow distributor at the inlet has been a neglected issue in column design. There are several parallel requirements, beginning with the need for a uniform distribution of the liquid at the column inlet. A degree of variation from this ideal situation is always encountered, and this particular parameter is not considered to be critical in determining the overall performance of the bed. One of the reasons is the higher ratio of length to diameter, which offers a dampening effect for poor flow profiles that

Table 2. Commercial Suppliers of Columns for Expanded Bed Chromatography

Supplier	Column diameters (cm)	Materials	Distributor type
Amersham Pharmacia Biotech AB.	2.5, 5.0, and 20.0	Glass	Stainless steel disk and mesh arrangement
Amersham Pharmacia Biotech AB.	60 and 120	Stainless steel	Stainless steel disk and mesh arrangement
UpFront Chromatography A/S.			Distribution through stirred baffles in the column base

result from maldistribution at the inlet. A more relevant requirement of flow distributors in the context of expanded bed chromatography is that it should allow free passage of the particulate material present in the feedstock. The high voidage of the bed itself permits almost unhindered flow, and bed blockage is therefore more likely to be a problem at the flow distributor. Although larger pore sizes are clearly beneficial in overcoming this limitation, it is also important to remember that the adsorbent particles themselves must be retained within the column.

The most common distributor type is the perforated stainless steel plate fitted with a stainless steel wire mesh on top. This type is found in all of the Amersham Pharmacia Biotech (APB) columns available commercially. The number of holes in the perforated plate depends on the column diameter it fits. The wire mesh serves to prevent the liquid from jetting into the bed at the column inlet by offering a distribution surface. It also has a dual purpose of preventing the adsorbent particles from leaking out from the bottom of the column. De Luca et al. (19) have investigated bed expansion characteristics using distributors with various numbers of orifices and arrangements from the center. The settled bed height and bed expansion were found to be important parameters that influence mixing within the bed.

Although all expanded bed work using APB columns relies on these distributors, little has been published about the problems of blockage when using unclarified feedstock. This is not to say that the problem is nonexistent. Column back pressure buildup occurs because of gradual accumulation of the particulate material in either the pores of the perforated disk or the space between the disk and the mesh. In new columns and those seeing their first feedstock challenge subsequent to cleaning and regeneration, the buildup is relatively slow, and a much longer load time is possible before the column needs to be back flushed to clear the appropriate blockage. Back flushing is usually done by reversing flow direction for a short period, often at an enhanced rate. When the loading is resumed, the back pressure will be at its initial value; however, the buildup to the next back flush condition often occurs more quickly. The procedure can in principle be repeated until the end of the load cycle. The exact length of the initial and subsequent load periods depends on the parameters of the systems in question, including the solids concentration in the feedstock, the type of flow distributor, and the flow rate. Thus no reports on investigations of the influence of different parameters on the load efficiency of flow distributors have been published.

An interesting approach to flow distributor design has been adopted by UpFront Chromatography (Denmark). Its novel method essentially bypasses the conventional type of distributor and mesh from the flow path of the feedstock. In contrast to the commonly used method for load described earlier, the UpFront system introduces the feedstock from a side inlet and above the bottom support mesh. Distribution of the material is ensured through the use of a stirring device located immediately above the bottom support mesh. Although distributor blockage is clearly avoided in this system, the stirring speed needs to be carefully selected, balancing the requirement for uniform distribution over the inlet area with the need to avoid vortex formation and increased liquid–liquid mixing in the column itself. The scale-up potential of this distributor type needs to be addressed. The company manufactures 2-, 5-, and 20-cm-diameter columns using this technology.

Bioprocessing (England) has another approach to obtaining uniform flow distribution. Standard chromatography columns in which beds of glass ballotini are placed directly above the bottom support mesh are used. The expanded bed adsorbent is then added on top. When flow of buffer is initiated through the column, the bed of adsorbent fluidizes but the glass ballotini do not because of their significantly higher density. Uniform distribution is obtained, and the mixing of adsorbent and ballotini beads is avoided. Proper selection of ballotini size is important. The smaller ones have been shown to offer better distribution, equivalent to the use of a perforated disk with many orifices. However, one must remember that smaller particles have a lower minimum fluidization velocity, and the selection needs to balance the need for small-sized ballotini with maintaining a larger difference in the minimum fluidization velocity between it and the adsorbent beads.

Adsorbent

As with packed bed chromatography, the adsorbent must offer a basis for interaction with the solubilized components of the feedstock to enable product capture. In addition, expanded bed adsorbents must also possess good fluidization characteristics. During the flow of particulate bearing feedstock through an expanded bed, the adsorbent should, under the flow conditions employed, remain in a fluidized state and not tend to elutriate from the column. This can be achieved if the terminal settling velocity of the adsorbent is significantly higher than any of the other solid matter present in the system. The three physical properties of the system that most significantly influence the terminal velocity of the adsorbent—particle diameter, density difference between the particle and the liquid, and liquid

viscosity—can be seen from the Stokes equation, described earlier.

Liquid viscosity is essentially dependent on the feedstock and is for the most part not an adjustable parameter, although some feed pretreatment is certainly possible. However, regarding the adsorbent, one can essentially experiment with the particle diameter and the density difference between the adsorbent and the liquid. The problem of suitably dense adsorbents can be offset to some extent by using size as a parameter for increasing the terminal velocity of an adsorbent particle. Thus a material of lower density and higher particle size can be used in expanded bed chromatography. However, a significant disadvantage lies in the resulting increase in mass transport resistance. The diffusion time is related to the square of the particle diameter; pore diffusion thus becomes the rate-limiting step in larger particles. Another disadvantage of expanded beds is increased interparticle distances, which may lead to slower transport of solute from the bulk to the surface of the adsorbent. The presence of particulate debris may cause further mass transport limitations and pose problems of adsorbent fouling and pore blockage.

The optimal adsorbent for use in expanded bed chromatography should therefore be small and very dense, and, as with any chromatography adsorbent, the dynamic bed capacity (DBC) should be acceptably high. There is an intrinsic link between these different parameters for effective expanded bed operation. DBC can be improved by using smaller particle sizes; however, efficient removal of the particulate debris present in the feedstock is an essential process requirement, and the two conditions can be simultaneously satisfied through the use of high-density adsorbents. Although good adsorbent design is critical in expanded bed operation, only limited work has been undertaken on its improvement, and most applications of this technology thus far have centered on the use of commercially available material. Only a few such materials exist currently. Table 3 summarizes some of the more common adsorbents currently used in expanded bed applications.

APPLICATIONS

The growing realization of the potential of expanded bed chromatography is borne out by the rapidly increasing number of reported applications of this technology in biological product recovery. The growth of this subject area over the last decade can be judged by following the literature for reported work during that period. The pioneering work in this field by Chase and coworkers relied on the use of conventional packed bed chromatographic adsorbents (28,29). Yeast cell homogenates were also commonly used in model systems. They are a suitable starting material given the many proteins present, any one of which can be targeted for recovery by expanded bed operation and offering the scope for use with ion exchange and affinity adsorbents (30,31).

Applications of expanded bed chromatography for the recovery of product directly from feedstock has moved on from the yeast homogenate systems that were studied a few years ago. There are now examples of applications for the recovery of both native and recombinant proteins from yeast (32) and bacterial cells (33–35), milk (36), and monoclonal antibodies from hybridoma cells.

The examples mentioned are but a few of the growing list of new applications that are being reported. The recently held Second International Conference on Expanded Bed Chromatography gives further evidence of the growing awareness of the potential of this technology in biological product recovery (see the book of abstracts for further information). In following the work in this area, it is obvious that the applications are for the most part based on conventional operational procedures and differ only in the system for which the recovery is being studied. Although the theory-oriented investigations have strived to tackle interesting issues of this technology, the applications mostly been bound by conventional operating protocols, similar to what has been described in the previous sections.

Two applications warrant specific mention because of their novel and interesting approach and are briefly highlighted. The work by Chase and coworkers on the use of a continuous countercurrent extractor (37–39) extends the

Table 3. Commercial Suppliers of Tailor-Made Adsorbents for Expanded Bed Chromatography

Supplier	Adsorbent trade name	Ligands	Density (g/mL)	Size	Matrix type
Amersham Pharmacia Biotech AB.	STREAMLINE™	DEAE SP Phenyl Recombinant protein A	1.2	100–300 μm, mean size ~200 μm	Cross-linked agarose with crystalline quartz core
Bioprocessing Ltd.	PROSEP™	Protein A Thiosorb	1.3	100 μm; pore size ~70–1,000 Å	Porous, dense glass beads
UpFront Chromatography A/S.	Mimo™	Mixed mode	1.0–1.6		Cross-linked agarose with core of one glass particle
	FastMabs™	Antibody	1.0–1.5		
	UFC	Ion exchange Hydrophobic Metal chelate	1.0–1.6 1.0–1.6 1.0–1.6	40–300 μm	
Biosepra	HyperD				Mineral oxide based

principles of expanded bed chromatography to continuous protein separation. A perfluorocarbon-based adsorbent (40) was pumped between four different contactors, each representing in sequence one of the four stages (i.e., load, wash, elution, and regeneration). Crude unclarified feedstock was pumped into the first contactor and product obtained from the third. The use of a specific affinity ligand permitted highly purified product to be collected at the outlet of the elution stage.

Another interesting methodology has been demonstrated by Lyddiatt and coworkers in which a fluidized bed was coupled with batch fermentation for direct product removal through either continuous or intermittent recycle of the reactor contents through the fluidized bed. Inherent problems of holding the fermentation material outside the controlled reactor environment were addressed through the use of a multibed manifold comprising smaller beds. Each bed was operated until the adsorbent was saturated and then removed from the flow loop for product recovery, cleaning, and regeneration.

Expanded bed protein adsorption has now reached a stage in its development where a good understanding of the potential and drawbacks can be effectively assessed. Issues related to the bed design, performance, and application have been addressed in the literature. Although the potential of this technology is well accepted, it is dangerous to assume that it will offer the solution to many recovery and separation problems. Many aspects are still either not considered important or not understood. The worth of the technology will be seen only when they become an integral part of bioprocesses. Toward this goal, the first step may already have been taken: There is now at least one report of a biotechnology company having patented a process for plasmid DNA production in which expanded bed adsorption is incorporated in the purification protocol.

BIBLIOGRAPHY

1. J.E. Ramirez-Vick and A.A. Garcia, *Separation Purif. Methods* **25**, (1996).
2. A.I. Clarkson, M. Bulmer and N.J. Titchener-Hooker, *Bioprocess Eng.* **14**, 81–89 (1996).
3. J.P. Maybury, K. Mannweiler, N.J. Titchener-Hooker, M. Hoare, and P. Dunnill, *Bioprocess Eng.* **18**, 191–199 (1998).
4. A. Pessoa and M. Vitolo, *Appl. Biochem. Biotechnol.* **70**, 505–511 (1998).
5. D.P. Harris, A.T. Andrews, G. Wright, D.L. Pyle, and J.A. Asenjo, *Bioseparations* **7**, 31–37 (1997).
6. S.G. Walker and A. Lyddiatt, *J. Chromatogr., B* **711**, 185–194 (1998).
7. G.M. Zijlstra, M.J.F. Michielsen, C.D. de Gooijer, and L.A. van der Pol, *Bioseparations* **6**, 201–210 (1996).
8. M. Ritopalomares and A. Lyddiatt, *J. Chromatogr., B* **680**, 81–89 (1996).
9. G.M. Zijlstra, M.J.F. Michielsen, C.D. de Gooijer, and L.A. van der Pol, and J. Tramper, *Biotechnol. Bioeng.* **12**, 363–370 (1996).
10. I.J. Isaacs, *Australasion Biotechnol.* **6**, 88–92 (1996).
11. S.R. Narayanan, *J. Chromatogr., A* **658**, 237–258 (1994).
12. M. Leonard, *J. Chromatogr.* **699**, 3–27 (1997).
13. E. Boschetti, *J. Chromatogr., A*, **658**, 207–236 (1994).
14. Q.M. Mao and M.T.W. Hearn, *Biotechnol. Bioeng.* **52**, 204–222 (1996).
15. D.K. McCormick, *Biotechnology* **11**, 1059 (1993).
16. Y.U. Wen and Y.H. Yu, *Chem. Eng. Prog.* **62**, 100–111 (1966).
17. J.F. Richardson and W.N. Zaki, *Trans. Inst. Chem. Eng.* **32**, 35–53 (1954).
18. G.S. Finette, Q.M. Mao, and M.T.W. Hearn, *J. Chromatogr., A*, **743**, 57–73 (1996).
19. L. De Luca, D. Hellenbroich, N.J. Titchener-Hooker, and H.A. Chase, *Bioseparations* **4**, 311–318 (1994).
20. A. Karau, C. Benken, J. Thömmes, and M.-R. Kula, *Biotechnol. Bioeng.* **55**, 54–64 (1997).
21. O. Levenspiel, *Chemical Reaction Engineering*, 2nd ed., Wiley, New York, 1972.
22. J. Thommes, M. Weiher, A. Karau, and M.-R. Kula, *Biotechnol. Bioeng.* **48**, 367–374 (1995).
23. H.A. Chase, *TIBTECH* **12**, 296–303 (1994).
24. M. Goto, T. Imamura, and T. Hirose, *J. Chromatogr.* **690**, 1–8 (1995).
25. L. Nixon, C.A. Koval, R.D. Noble, and G.S. Slaff, *Chem. Mater.* **4**, 117–121 (1992).
26. B.E. Terranova and M.A. Burns, *Biotechnol. Prog.* **5**, 98–104 (1989).
27. A.S. Chetty and M.A. Burns, *Biotechnol. Bioeng.* **38**, 963–971 (1991).
28. H.A. Chase and N.M. Draeger, *J. Chromatogr.* **597**, 129–145 (1992).
29. Y.K. Chang, G.M. McCreath, N.M. Draeger, and H.A. Chase, *Trans. Inst. Chem. Eng.* **71**, 299–303 (1993).
30. Y.K. Chang, G.E. McCreath, and H.A. Chase, *Biotechnol. Bioeng.* **48**, 355–366 (1995).
31. Y.K. Chang and H.A. Chase, *Biotechnol. Bioeng.* **49**, 204–216 (1996).
32. F. Raymond, D. Rolland, M. Gauthier, and M. Jolivet, *J. Chromatogr.* **706**, 113–121 (1998).
33. M. Hansson, S. Ståhl, R. Hjorth, M. Uhlén, and T. Moks, *Biotechnology* **12**, 285–288 (1994).
34. G. Maurizi, V. DiCioccio, G. Macchai, P. Bossu, C. Bizzarri, U. Visconti, D. Boraschi, A. Tagliabue and P. Ruggiero, *Prot. Expression Purif.* **9**, (1997).
35. H.J. Johansson, C. Jagersten, and J. Shiloach, *J. Biotechnol.* **48**, 9–14 (1996).
36. A. Degener, M. Belew, and W.H. Velander, *J. Chromatogr.* **799**, 125–137 (1998).
37. R.O. Owen, G.E. McCreath, and H.A. Chase, *J. Mol. Recognit.* **9**, 575–584 (1996).
38. R.O. Owen, G.E. McCreath, and H.A. Chase, *Biotechnol. Bioeng.* **53**, 427–441 (1997).
39. R.O. Owen and H.A. Chase, *J. Chromatogr., A* **757**, 41–49 (1997).
40. G.E. McCreath and H.A. Chase, *J. Mol. Recognit.* **9**, 607–616 (1996).

PROTEIN AGGREGATION, DENATURATION

CATHERINE H. SCHEIN
University of Texas Medical Branch
Galveston, Texas

KEY WORDS

Aggregation
Biophysical techniques
Co-solvent effects
Homology modeling
Kinetics
Measuring
Protein folding
Surface force
Three-dimensional structure

OUTLINE

INTRODUCTION

Because some protein components, such as the tyrosine side chain, have negligible solubility in water, proteins in solution exist in a precarious equilibrium. Most proteins form oligomers or aggregate at higher densities, a phenomenon that can to some extent be prevented by the addition of solvent additives (1) and mutations targeted to surface areas involved in aggregate formation (2). Solvent additives have proved invaluable in areas ranging from protein formulation for drug purposes to sample preparation for high-resolution structure determination techniques. Using cosolvents is, however, a matter of trial and error, with little framework for predicting their effects or even optimizing concentration. This is partly because of the inherent problems in following the early stages of protein aggregation.

This article summarizes biophysical techniques useful for measuring solubility and following precipitation. Maintaining solubility depends on subtle interactions of side chains and backbone atoms to balance the cavitation energy needed to maintain a solvent pocket for the bulk of the protein by the energy of interaction of hydrophilic groups with the solvent. Traditionally, proteins are presented as unique but ill-characterized globules remarkable only with respect to size and net surface charge. Direct methods to determine the characteristics of the protein surface may allow a better accounting for protein–solvent interactions in the presence of cosolvents. Energy functions that account for the various forces on the protein have been derived; their application to different solvents is now being tested. Interpreting empirical data to derive atomic solvation parameters or deriving bulk solvent parameters may lead to a thermodynamic or kinetic accounting for the effect of cosolutes on protein stability and solubility.

Designing mutations that prevent aggregation can be greatly helped by a high resolution three-dimensional structure of the protein–protein complex. A relatively unsophisticated home computer or work station now allows easy assess to the steadily increasing data banks of detailed 3-D protein structures, which were previously accessible only with expensive viewing systems. Faster and more efficient computational methods have greatly expanded our abilities to model protein structures by homology to proteins of known structure or ab initio using protein threading and other algorithms from the primary sequence. Thus, a section is also dedicated to methods for obtaining and using structural information to predict areas of the protein that may be involved in oligomerization and aggregation.

METHODS FOR MEASURING SOLUBILITY AND PROTEIN ASSOCIATION

For details about the molecular interactions involved in complex formation, many methods measure the absorption or scattering of some form of electromagnetic radiation by the sample (UV/vis and IR/Ramen spectroscopy, light scattering, laser light scattering, neutron diffraction and X-ray crystallography) (3). A beam of radiation sent through the sample can:

1. Pass through the sample unchanged
2. Be absorbed by the sample and emitted at the same wavelength
3. Be absorbed by the sample and emitted at a different wavelength
4. Be scattered by the sample

Various methods can be used to detect the effects of the sample on the beam (4). Figure 1 is a summary of methods aligned according to the electromagnetic radiation spectrum. At opposite ends of this spectrum lie the two highest resolution methods, X-ray crystallography and NMR spectroscopy, which yield complete structural information at the atomic level. However, combining results from lower resolution methods can be used to develop a working model of complexes and to obtain information about the aggregation process. In the simplest and most commonly used method to measure protein concentration and aggregation, a coherent beam of visible or UV light is passed through the sample. The light emerging on the other side of a cuvette of known path length is measured and related to the concentration of the sample. Some information about the nature of the sample can be obtained by observing how much light is absorbed when the wavelength is varied over the whole spectrum (a so-called UV/vis scan). In large proteins, the tryptophan side chains within the protein may be shielded, and as the protein unfolds, this tryptophan absorption at 280 nm becomes more pronounced. Absorption in the visible range is usually due to the presence of a chromophore (e.g., Haem absorption) or particulates. Fluorescence spectroscopy measures the energy of the emitted light, and laser light scattering measures the angle of the reflected light to determine particulate size (5). Many other, more advanced methods, summarized in Table 1, have been used to measure the quaternary state of proteins in solution.

Combining Methods to Follow Aggregation and Precipitation and Determine the Structure of Complexes

The size of small complexes in solution can be accurately determined by equilibrium centrifugation, native gel electrophoresis, or gel sizing chromatography (using for example Sephadex beads or HPLC). However, these methods are not particularly useful for studying aggregated protein. In vitro aggregation can have many forms, extending from a drastic, sudden precipitation of large flocculent masses of protein, often visible to the naked eye, to the gradual formation of tiny, invisible particles. The most valuable experiments look at the initial phases. The simplest direct quantitation of aggregation is thus to measure the absorption of visible light by the sample, if the particles are small enough to remain in solution, or to follow the precipitation by the drop in UV absorption of a protein solution where the particles are allowed to settle or are removed by centrifugation. If aggregation is related to unfolding, it may be possible to follow the increase in OD_{280} (as the tryptophan residues become exposed upon unfolding) and then its decrease as the protein aggregates.

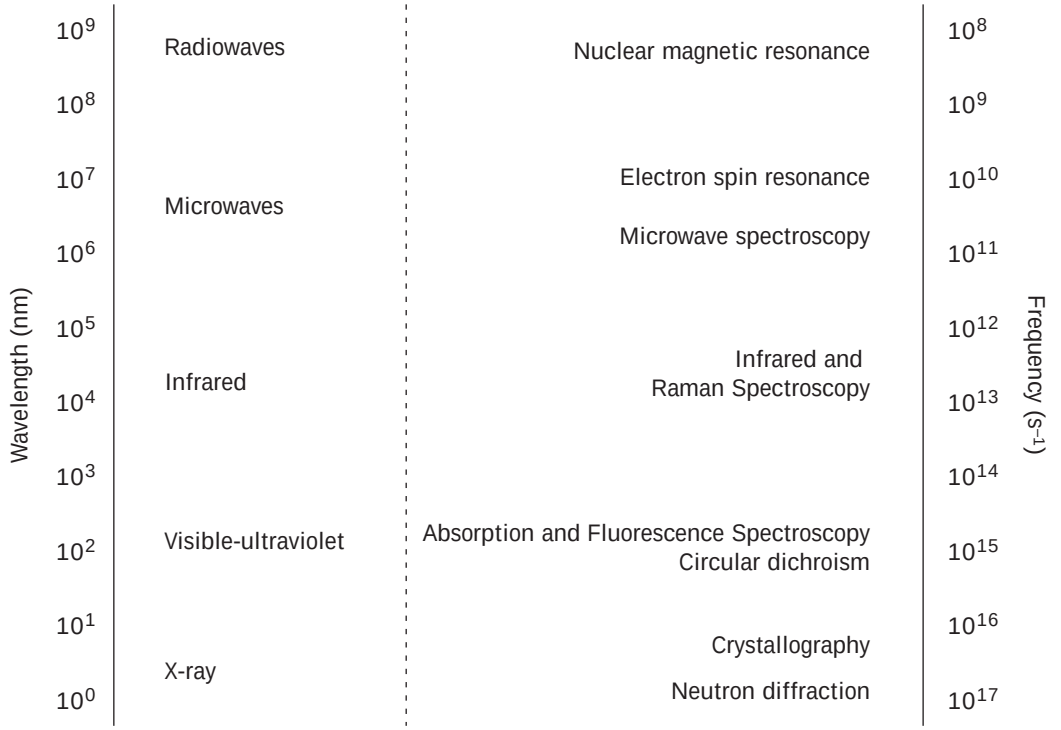


Figure 1. Methods used for analyzing the structure of biological molecules and complexes thereof as a function of position in the electromagnetic spectrum.

Table 1. Physical Methods for Monitoring Changes in Protein Structure and Interprotein Interactions Involved in Aggregation

Method (abbreviation)	Brief description	Protein needed[a]	Sample preparation/ buffer interference[b]	Best used for
Equilibrium (ultra) centrifugation (EC)	Protein complex is centrifuged at high speed in a density gradient, and its equilibrium position used to determine the size of the molecule or complex.	High	Gradient components may affect protein association.	Determining the oligomeric or aggregation state of soluble complexes. Insoluble aggregates settle at the bottom of the tube or float on the surface of the gradient.
Gel sizing chromatography (GSC)	Soluble proteins are passed over a column of porous polymer. Molecules small enough to enter the pores are held back and elute according to size, larger ones pass through unimpeded with the void volume. Can be done with capillary HPLC columns on the microscale.	Moderate to high	Sample should be concentrated for application, a step that may induce aggregation or oligomerization.	Sizing small, soluble complexes of proteins; large aggregates either elute with the front or precipitated at the top of the column (latter may slowly redissolve, giving unreliable data).
Polyacrylamide gel electrophoresis (PAGE)	Proteins are separated according to size by electrophoresis through acrylamide gels of varying density. Proteins may be detected in a variety of ways: by staining, autofluorescence, or after transfer to a membrane, by reaction with a specific antibody or nucleic acid fragment.	Concentrated sample required, but small amounts of protein may be visualized by silver or gold staining.	If the sample is not to be eluted for further analysis, apply less protein and use more sensitive dyes for accurate size determination.	Sizing small, soluble complexes of proteins; large aggregates stick in the slot.
UV/Vis absorption (UV-abs)	A beam of coherent light is passed through the sample, and the absorption by the protein constituents is measured. Proteins absorb between 210 and 220 nm because of orbital transitions of the peptide backbone, which are overwhelmed by that of the side chains of D, N, E, Q and R, and at 280 nm because of aromatic side chains of W, Y, F; chromophores in the protein may absorb throughout the visible range as well.	Low	Only for protein in solution; samples should be diluted to be in the linear range of the spectrophotometer	The denaturing effect of cosolvents on proteins; following aggregation by decrease in the fluorescence absorption and/or increase in absorption in the visible range.
Fluorescence transfer (FT)	A fluorescence acceptor and donor group are coupled to a protein. When excited by light, the quenching of donor fluorescence by the acceptor serves as a measure of the distance in the complex between them.	High	Only for protein in solution, buffer interference; requires coupling of donor and acceptor molecules usually to free cysteine groups in the proteins.	Distance between labeled side chains in a single molecule or oligomer; most useful with protein complexes with a well-established 3-D structure.

Table 1. Physical Methods for Monitoring Changes in Protein Structure and Interprotein Interactions Involved in Aggregation (*continued*)

Method (abbreviation)	Brief description	Protein needed[a]	Sample preparation/ buffer interference[b]	Best used for
Optical rotatory dispersion (ORD)/ Circular dichroism (CD)	In ORD, the rotation of a beam of polarized light by the sample is measured by a coupled polarizer/photomultiplier on the exit side of the cuvette. The angle of rotation of the second polarizer to obtain the highest intensity of emitted light is plotted as a function of the wavelength. In CD, molar ellipticity is measured as the difference in absorbence by the sample of an incident beam rotated to the left or right (I_L and I_R) by a timed electric field change. A photomultiplier on the exit side measures absorbence as a function of time, and a microprocessor interprets the data to determine the mean residue ellipticity (which is proportional to $I_L - I_R$) as a function of wavelength.	Moderate	Only for protein in solution, buffer interference possible.	Secondary structure of proteins in various media; also to follow changes in the secondary structure as a function of denaturation or aggregation. Computer-aided spectral interpretation can be used to obtain more complicated data about the protein configuration.
Electron microscopy (EM)	In vacuo, a beam of electrons is passed through a target with the (usually stained or coated) object affixed to it or imbedded in a matrix to obtain a 2-D image. In essence, one measures the electron density of the sample after coating to obtain a 2-D image. Cryo-EM techniques use unstained samples embedded in vitreous ice. High-speed calculations can be used to phase many images of particles (6,7).	Low	Fixation/staining may be needed. Consistent sample preparation, crucial to the cryo-EM technique, can be quite difficult.	Determining the ultrastructure of large, electron-dense, complexes; particularly useful to determine the shape of particles and their size distribution. Current resolution is in the range of 10 Å.
Electron diffraction (ED)	By tilting the target during EM and measuring reflection of the electron beam at various angles from the incident, a 3-D image can be reconstructed.	Low to high	2-D crystals or membrane patches; viral particles.	Obtaining ultrastructural data about regularly structured monolayers such as the membranes of purple bacteria.
Scanning tunneling microscopy (STM)	The electron stream from the object is detected across a vacuum barrier by a sensitive probe. The microscope itself is in reality the probe, and the name comes from the tunnel current, which flows when there is a small electric current between the tip of the probe and the conducting surface to be scanned (which should be less than 1 nm away).	Low	Resolution is higher for conductive samples; fixation/ staining may be needed for proteins and nucleic acids.	Determining the shape and size of fixed and coated large complexes or protein-DNA complexes; resolution is theoretically at the atomic level; suited to the study of larger protein aggregates because it gives a 3-D image directly without the image reconstruction methods used for 3-D transmission EM images.
Atomic force microscopy (AFM)	Technique (also called scanning probe microscopy) related to STM and using similar apparatus, more suited to biological samples because it is done under a solvent phase rather than in vacuo (8). Very little different from hydration STM, based on the electrical conductivity of molecularly thin water layers that adsorb to the sample surfaces in a humid atmosphere.	Low	Attachment to a surface necessary, can be done in a variety of solvents.	Direct viewing of the ultrastructure of large, not necessarily regular biomolecular complexes without staining. Methods are being developed for viewing protein-RNA complexes.

Table 1. Physical Methods for Monitoring Changes in Protein Structure and Interprotein Interactions Involved in Aggregation (*continued*)

Method (abbreviation)	Brief description	Protein needed[a]	Sample preparation/ buffer interference[b]	Best used for
Fourier transform-infrared absorption (FTIR)	Absorption of light in the infrared region (2.5–250 μm) by the sample is measured and the output beam resolved by Fourier transformation. *IR difference spectra*, deconvoluted in the amide I band region (between 1620 and 1700 cm^{-1}) can show secondary structure in terms of % α-helix, β-sheet, etc. (9,10).	High	Can also be used for precipitates.	Determining secondary structure elements and alterations thereof under different conditions.
Infrared attenuated total reflection spectroscopy (IR-ATR)	Infrared beam is reflected through an optically transparent germanium plate coated with sample; internal reflections along the plate increases sensitivity.	Low to moderate	Membrane patches attached to a special plate.	Membrane protein interactions with each other and external molecules.
Scanning fluctuation correlation spectroscopy (FCS)	Measures particle number concentrations by monitoring spontaneous equilibrium fluctuations in the local concentration of fluorescent species in a small (femtoliter) subvolume of a sample.	Low	Little problem with buffer interference	Detect molecular aggregation for dilute, submicromolar samples by directly counting particles.
Elastic (classical, static) laser light scattering (SLLS)	For elastic (classical, static) light scattering, the scattering of a beam of polarized laser light as it passes through the sample is measured and the obtained scatter factor can be used to calculate both the weight average molecular weight of the particles in solution as well as the radius of gyration.	Moderate	Scattering methods are usually not greatly affected by the buffer.	Determining the size of aggregates and precipitates. To obtain the scattering weight of a particle, one measures the scattering factor at several different scattering angles (q) and then plots the scattering factor as a function of q. Extrapolating back to $q = 0$ will yield the scattering weight.
Quasielastic (dynamic) laser light scattering (DLLS)	Similar to SLLS; but changes in the frequency from the translational (Brownian) movement of the scattering particles are also measured. The broadness of the intensity distribution of the emitted light for frequencies around the primary monochromatic beam frequency is directly related to the diffusion coefficient of the particles, which can then be related to the hydrodynamic radius if a model for the particle shape is available.	Moderate	The beam should not be strong enough to damage the sensitive biological specimen.	Kinetics of particle coagulation by following the decrease in diffusion coefficient as the particle size increases (11); determine particle molecular weight as a function of time; to follow course of aggregation in any solvent; compare samples in different buffers if viscosity is accurately measured.

Table 1. Physical Methods for Monitoring Changes in Protein Structure and Interprotein Interactions Involved in Aggregation (*continued*)

Method (abbreviation)	Brief description	Protein needed[a]	Sample preparation/ buffer interference[b]	Best used for
Raman, Raman resonance (RRS)	Light scattered at the incoming frequency, so-called Raleigh scattering, is measured in classical (elastic) and quasielastic (dynamic) light scattering. A small fraction of the laser light is scattered at frequencies higher (anti-Stokes) and lower (Stokes) than the incident beam. These scattered beams are measured in Raman scattering, which has many variations.	High	Sample fluorescence may interfere with Raman resonance measurement of precipitates.	Follow quaternary interactions of soluble complexes and to compare the secondary structures of proteins under various conditions; to compare the effects on secondary structure of precipitating agents.
Neutron scattering (NS)	Scattering of a neutron beam by the sample is measured.	High	Beam time is limited.	Determining the size of larger proteins and complexes.
Mass spectrometry (MS)	The magnetic field induced deflection of a vaporized, ionized molecule is used to determine its mass. Types are characterized by the method used to ionize and detect the sample. Advanced computational methods can be used to interpret structural data from the cracking pattern of larger molecules.	Varies	Usually protein is bound to a solid (e.g., nitrocellulose).	Very accurate molecular weight of monomers; use in combination with chemical cleavage methods may aid in determining tertiary and quaternary structure. Electrospray ionization (ESI-MS) has been used to characterize complexes of proteins and protein/nucleic acids (12).
Electron spin resonance (ESR)	Microwaves are absorbed by paramagnetic substances, and the change in the energy level of the electron spins measured.	Moderate		Structure and dynamics of lipid membranes; free radical detection in proteins; state of transition metals in proteins.
Nuclear magnetic resonance (NMR)	A concentrated solution of sample, possibly isotopically enriched for nuclei with a magnetic moment different from 0 (^{1}H, ^{13}C, ^{15}N, ^{31}P), is subjected to a high magnetic field. In the most common method, after the sample is pulsed, the decay in amplitude of the emitted radiowaves with time is measured and converted to a frequency spectrum by Fourier transformation. The magnitude and direction of the pulse can be altered to detect certain types of interactions between nuclei. The strength of the local magnetic field around an individual atom is decreased by the presence of covalently bound or spatially close neighboring groups, causing changes in the position of the resonance lines (chemical shifts) in the output frequency spectrum. Peak intensities in NOESY spectra reflect the distances between individual spins. Distance geometry methods and restrained molecular dynamics calculations are used to calculate the 3-D structure of proteins or protein–protein complexes.	High	Most solvents interfere with measurements. Labeling of the sample with stable isotopes may be necessary. Oligomers may be artificially generated by the high protein concentration required for measurement.	Determining the exact tertiary structure of small proteins, nucleic acids, and complexes. Combining results from different NMR experiments can be used to monitor changes in the resonance frequencies upon complex formation, and thus give specific information on binding sites. Various techniques (e.g., resonance transfer) can be used to measure interactions between two molecules even without establishing a structure. See Ref. 13 for a discussion of NMR techniques to study protein complexes.

Table 1. Physical Methods for Monitoring Changes in Protein Structure and Interprotein Interactions Involved in Aggregation (*continued*)

Method (abbreviation)	Brief description	Protein needed[a]	Sample preparation/ buffer interference[b]	Best used for
Protease mapping (PM)	Fragments left after cleaving with a site-specific protease are analyzed (by gel chromatography and sequencing or MS/MS sequencing) to determine surface accessible areas	High	Need to have specific sites for proteases in the protein; buffers should not interfere with techniques used to identify protein fragments.	Tertiary structure of an isolated protein or protected areas in complexes.
Chemical-based cleavage (CC)	Like PM, but an activated metal ion cleavage agent is incorporated at a free sulfhydryl group, and the cleavage is triggered by chemical or physical stimuli.	High	Buffers used cannot stimulate cleavage or interfere with subsequent methods to determine the size and constitution of fragments; need free sulfhydryl groups in the protein	Tertiary structure of an isolated protein or areas of contact in complexes.
Surface plasmon resonance biosensors (SPR)	A surface plasmon is an oscillation of free electrons that propagates across the surface of a conductor, typically a thin film of gold or silver. The sample to be measured is immobilized on one side of the gold film, e.g., in a hydrogel, and plasmon excitation is started with a beam of light. The intensity of plasmon generation (measured as the reflected light reaching the detector) is related to the refractive index of the sample. This can be altered by flowing a binding substance over the surface where it is immobilized.	Low to moderate	No labeling of the biomolecules is required, but the two media must have different refractive indices. To avoid rebinding artifacts, the bound protein should not be overloaded.	Real time analysis of (unlabeled) protein binding reactions (14).
Total internal reflection fluorescence (TIRF)	Fluorescently labeled proteins in aqueous solution interact with a solid phase (e.g., fused silica) coated with bilayers of synthetic phospholipids, unlabeled proteins, or even whole cells.	Moderate	Two media must have different refractive indices.	Determining interactions between membrane-bound and free proteins.
Small-angle X-ray scattering (SAXS)	The scattering of X-rays at small angles can be used to obtain greater resolution of X-ray crystal structures.	Moderate	Crystals required and special detection/computation systems.	Resolving the structure of large complexes to higher resolution than possible with diffraction data alone.
X-ray diffraction (XRD)	A protein crystal is bombarded with a high intensity X-ray beam of known wavelength and polarity. The diffraction pattern of the emitted energy, measured as spots on X-ray film or using an area detector, can be used to determine the position of individual atoms in the 3-D structure of the protein or complex.	High	Crystals required and extensive experimentation must be done to solve the phase problem.	The only method now in use that can give the 3-D structure of large, regular protein complexes at the atomic level. Adaptations are now being developed to use it to determine dynamics as well.

Note: Data presented with respect to the protein requirements, impedance by common buffers, and the type of information obtainable about the protein in solution or precipitate. For more details, see related references in Table 2; for general information about the methods, see Galla (4), cited specific references, or other books on protein characterization techniques.

[a]Protein concentration required for measurement. Low is in the nM–μM range, high means 0.1 mM or higher. Note that highly concentrated protein solutions are usually required to obtain crystals and that a lot of protein is wasted during the process of crystallization. Thus, one needs to start with about 100 mg of purified protein for X-ray analysis, although only a few μgs may find their way to the beam.

[b]Buffers used for sample preparation for all methods should be sterile and ultrafiltered (0.45 μ or smaller pore diameter filter) to remove particulate matter. Buffer requirements for various methods are discussed in the references given.

Although rapid onset suggests aggregation is a noncontrolled process, there are reports of sequence specificity based on the pattern of higher molecular weight aggregates seen on gels (15). An approximately similar pattern of bands was seen, whether a protein formed inclusion bodies alone or in the presence of another aggregating protein. Although this would indicate that each protein was only aggregating with itself, the authors could only study aggregates small enough to enter a nondenaturing gel, and a substantial portion of the protein in all the experiments remained at the top of the polyacrylamide gel (PAG). More recent data using computer simulations and light scattering suggest a cluster–cluster mechanism rather than a stepwise monomer addition polymerization (16). The aggregation of polypeptide chains during refolding occurred by multimeric polymerization, in which two multimers of any size could associate to form a larger aggregate and did not require a sequential addition of monomeric subunits.

Generally, combinations of methods are applied to proteins, such as the HIV-Rev, where it is important to determine the 3-D structure, but problems arise in obtaining this by high resolution methods. The Rev protein, whose function is to move viral unspliced RNA from the nucleus to the cytoplasm, is essential for the establishment of infection. Although establishing its 3-D structure would be very useful for designing inhibitors that could be used therapeutically, severe aggregation problems have hindered its characterization by X-ray or NMR spectroscopy. Data from STEM, cryo-EM of lysine cross-linked filaments, and Raman spectroscopy have been used to develop a low resolution (21 Å) model of Rev filament structure that may be useful in designing inhibitors (17). Structural modeling tools possibly useful for this protein are discussed later. To further illustrate how the methods of Table 1 can be applied to study the interaction of proteins, Table 2 is a list of some of the protein complexes that have been studied by combining biophysical methods.

CHARACTERIZING PROTEIN–SOLVENT INTERACTIONS

Protein Stability in a Practical Sense

As Privalov has pointed out (85), if one assumes that all deviations from the X-ray crystal structure represent a new and separate state of a protein, the net thermodynamic stability of a protein at any temperature above a few degrees Kelvin is essentially zero. The thermodynamic framework works well for small molecules that can assume a limited number of stable conformations. However, proteins change conformation continuously; phenylalanine side chains are often free to rotate around the Cβ-phenyl bond and as many as 40% of the amide protons of a protein exchange with the solvent in a few seconds (86). Solvent effects make these dynamic considerations even more complicated; the rate of hydrogen exchange of a protein changes when even the isotopic nature of the solvent is changed (87).

To simplify the problem, thermodynamic stability has been defined as the energy needed to switch the protein structure from a folded, active native state (in reality a collection of substates around a mean native structure) to a completely unfolded, denatured state. Formation of the denatured state in a dilute protein solution can be induced by slowly increasing temperature until the transition is indicated by extreme changes in the protein's activity or protease susceptibility, or in the protein's characteristic UV absorption, circular dichroism, or more recently, NMR spectrum. The temperature at which this change occurs is referred to as the thermal denaturation temperature, or T_d. Thermodynamic stability data can be collected in a few hours, but the measurements cannot be done at high-protein density because aggregation usually precedes total denaturation. The T_d varies considerably from one protein to another and is a function of the pH and other properties of the solution in which the measurement is made. The free energy of unfolding ($\Delta G^\circ_{\text{unfolding}}$), which can be calculated from the T_d, is approximately 12 $\pm$ 5 kcal/mol at 30 °C for all proteins. This number is equated with the free energy of stabilization of proteins in aqueous solution according to the two-state model (88). As native, properly folded structures aggregate less than unfolded, denatured ones, there is an intimate relationship between solubility and thermodynamic stability. However, although extreme differences in the T_d can be used as a direct indication of the stability of the protein structure (peptides may lose their secondary structure at T $<$ 12 °C, whereas thermophilic proteins retain it in a boiling water bath), the T_d of typical proteins is in too narrow a range to serve as an indicator of structure or purity. We know that proteins vary greatly in their stability in solution or the dry state. It has been shown that the $\Delta G_{\text{unfolding}}$ for RNase T1 can be almost doubled by adding 0.2 M Na_2HPO_4 or certain other salts, but the increase amounts to only 5 kcal/mol (86). Many additives that stabilize proteins against aggregation do not affect the T_d. Aggregation occurs at temperatures well below the T_d for proteins, indicating that a protein need not unfold completely to aggregate.

Thus, other methods must be used to determine protein stability in solution under nonextreme conditions. A two-state model ignores the less abrupt (and therefore less easy to measure) changes in the protein structure that precede complete unfolding. Such changes, coupled with an increased rate of aggregate-producing encounters, probably lead to temperature-dependent protein precipitation. Stability for practical purposes is the empirically determined resistance of a protein to oligomerization or to inactivation by other mechanisms. Stability can then be equated with the persistence of solubility or activity and determined kinetically as the protein half-life (measured by protein concentration or activity after centrifugation) in a solution of specified pH, solute concentration, and temperature (1).

The kinetic method of measuring stability requires several days and moderate amounts of protein. To shorten the time required for kinetic measurements of solution stability, an additional stress can be put on the system by agitating it or adding a protease. For example, resistance to mechanical shaking was found to be a useful indicator of solution half-life for 10 enzymes present in liver extracts (89). Use of more sensitive methods to detect protein aggregation should reduce the protein required for stability measurements. Osmolytes useful for formulating recombinant human growth hormone and interferon-γ were de-

Table 2. Model Systems for the Study of Protein Aggregation

Model system	Studied with	Examples and key results
		Multiprotein complexes
Ribosomes	XRCrys, SANS, EM, neutron diffraction	Early crystals of ribosomes from thermophilic organisms diffracted to >20 Å were used for comparing the exterior structure with EM photos. Crystals diffracting to <5 Å were used, combined with image reconstruction, to determine structural details (18). Neutron diffraction of ribosome crystals was used to localize the RNA within the core of the particles using difference maps prepared at different D_2O/H_2O (altering the relative concentrations of D_2O and H_2O in the solvent changes the relative diffraction caused by RNA or protein in the crystals) (19). Freezing crystals stabilized them in the X-ray beam. A 37-Å resolution 3-D structure of ribosomes from rabbit reticulocytes was calculated from electron micrographs of uranyl-acetate stained single-particle specimens (20); whereas cryo-EM and neutron scattering data can be used to model the protein/nucleic acid interactions (21).
Nucleosomes	EM, XRCrys	The crystal structure of the nucleosome core particle was developed to high resolution using a variety of isolation and protein techniques and modern image analysis (22).
Photosynthetic reaction center and other membranes	EM, XRCrys, NMR, electron diffraction	Crystals diffracting to ~3 Å resolution required the development of small amphiphilic molecules (23). Other techniques include NMR of ^{19}F-labeled bacteriorhodopsin fragments in a membrane-like milieu (24) and IR difference spectra to monitor structural changes in ^{15}N-labeled or deuterated bacteriorhodopsin (25). EM of purple membranes from bacteria are a natural 2-D crystalline array of lipids and bacteriorhodopsin. As the array is only one unit cell thick, it is not suitable for X-ray analysis, but yields electron diffraction patterns with spots to beyond 4 Å resolution. By tilting the sample, a series of patterns are obtained that can be converted to a 3-D model of the structure of bacteriorhodopsin in the membranes (26, 27). EM images the hexagonally packed regular protein monolayer from the cell wall of the bacterium *Deinococcus radiodurans* (28).
Virus assembly	CD, EM, XRCrys, SAXS, STM	CD showed the membrane-associated coat protein of bacteriophage M13 (5.2 kD) changed from an α-helical conformation when it was in a membrane environment (i.e., in solution with cholate or in the capsid) to a β-sheet form when aggregated (aggregation followed with PAGE) (29). NMR studies of detergent-solubilized M13 coat protein, in combination with the results of sedimentation equilibrium, Raman and CD studies, indicated two conformers that represent the nonequivalent monomers of an asymmetric dimer (30) STM was used to view bacteriophage T4 polyheads (31). Small-angle X-ray scattering was used to follow the rapid assembly of tobacco mosaic virus protein induced by a temperature jump. The minimum counting time was 7.5 seconds (32). Assembly of the icosahedral shells of the dsDNA bacteriophages, herpesviruses, and acid adenoviruses requires proteins not found in the mature virion, termed scaffolding proteins. The bacteriophage P22 precursor procapsid contains approximately 300 scaffolding molecules within a shell composed of 420 coat protein subunits (33). Cryo-EM has been used to characterize the scaffolding/assembly of herpes simplex virus particles (34).
Antibody–antigen interactions	XRCrys, NMR, SLLS, DLLS, STM	Crystal structures of antibodies or bacterially produced antibody fragments with peptide antigen (35) or proteins can diffract to <2.5Å (36) and can be used to identify the atomic features of the binding site. NMR studies of binding of an antibody fragment (Fab) to a peptide (37). As the on/off rate of the peptide is short compared to the spin-lattice relaxation time of the Fab and peptide protons, a NOESY spectrum in the presence of a vast excess of peptide has extra cross-peaks (transfer NOEs) that are due to magnetic exchange between the bound and free peptide fractions and are not present when the peptide and Fab are present in a 1:1 ratio (38). Dynamic light-scattering measurements of antibody–antibody complex aggregate formation indicated fractal dimensions rather than a defined geometric shape (39). STM ultrastructural images of unstained immunoglobulin G, air dried on the surface of highly oriented pyrolytic graphite (HOPG), required only 1 μl of a 0.5 mg/mL solution of protein in 50 mM of tris buffer (40).
Casein micelles	SANS, light scattering, EM	Neutron scattering from suspensions of casein micelles measurements were combined with EM data to characterize the structure as packed micelle monomer subunits (41).

Table 2. Model Systems for the Study of Protein Aggregation (*continued*)

Model system	Studied with	Examples and key results
Inclusion body formation	EM, SLLS	The traditional method to determine the presence of aggregation after differential centrifugation is gel electrophoresis (also coupled with immunoblotting or isoelectric focusing [42]). Inclusion bodies contain protein polymers (43,44) resistant to SDS and β-mercaptoethanol solubilization. Inclusion bodies (IBs) are seen as electron dense particles in EM thin sections of whole *E. coli* cells; the morphology of isolated IBs was studied by scanning and freeze-etch EM (45). IBs (at least the polymeric forms that will enter a PAG) appear to consist of only one protein, not nonspecific coaggregates with other proteins (15). More recent data using computer simulations and light scattering suggest a cluster–cluster mechanism rather than a stepwise monomer addition polymerization (16).
β-amyloid fibril formation	CD, static and dynamic light scattering	β-amyloid peptide (A-β) is the primary protein component of senile plaques in Alzheimer's disease patients. Assembly of synthetic A-β into amyloid fibrils in phosphate-buffered saline was followed by static and dynamic light scattering. CD was used to show A-β's structure varied when dissolved in organic solvents commonly used in neurotoxicity assays. In pure DMSO, A-β had no detectable β-sheet content; in 0.1% trifluoroacetate, the peptide contained one third β-sheet; and in 35% acetonitrile/0.1% trifluoroacetate, A-β was two-thirds β-sheet, equivalent to the fibrillar peptide in physiological buffer (46).
Bacterial flagella assembly, microtubules	EM and cryo-EM	EM demonstrated in vitro reconstruction of organelles such as *Salmonella* flagellar filaments (phosphotungstic acid stained preparations) (47) and microtubules (frozen hydrated or glutaraldehyde/tannic acid fixed and stained) (48).
Lipoxygenase substrate binding	EMR, X-ray absorption spectroscopy, infrared CD magnetic CD	Spectroscopic studies, including electron magnetic resonance, X-ray absorption spectroscopy, infrared circular dichroism, and magnetic circular dichroism, have been applied to compare lipoxygenases from varied sources and with different substrate positional specificity. Soybean lipoxygenase-1 contains a novel three-turn π-helix near the iron center (49).
Nucleotide binding	NMR	The conformation of MgATP bound at the active site of *Salmonella typhimurium* 5-phospho-α-D-ribose 1-diphosphate synthetase (PRibPP synthetase) (50) and conformations of the adenosine moiety of MgADP and MgATP bound to rabbit muscle creatine kinase (51) have been investigated by 2-D transferred-NOE spectroscopy (TRNOESY).
Fibrinolase/Zn binding	NMR and CD	The Zn binding enzyme fibrolase (pI is 6.7) shows a minimum in its solubility profile (which can be followed with CD or NMR spectroscopy as a structural change) below pH 5. Zn loss, induced by rising temperature or lowered pH or adding chelators (EDTA, DTT), leads to a rapid loss of α-helical structure, exposure of hydrophobic residues, and aggregation (52).
Calmodulin/Ca	XRCrys, SANS	Small angle X-ray and neutron scattering of calmodulin in complex with two peptides (53).
		Specific protein interactions
Profilin/actin	XRCrys	Crystals of profilin/actin diffracting to 1.8-Å resolution required ATP hydrolysis to form and a 3.2 M ammonium sulfate/ATP bathing solution for stability (54).
RecA protein/DNA	EM, SANS, STM	RecA protein complexes with itself and with DNA were characterized under various conditions by combining EM and small-angle neutron scattering (55) STM images to a resolution of a few nM has been obtained for coated specimens of RecA bound to DNA (56).
Other protein DNA complexes	EM, STM	A combination of electron microscopy, image analysis, 3-D reconstruction (57) and STM has been used to show that the active form of the DnaB protein, the primary replicative helicase in *E. coli*, is a hexamer. Depending on the presence of different nucleotide cofactors (ATP, ATP γ-S, AMP-PNP or ADP), two different hexameric forms of the DnaB ring were found, one with a three-fold symmetry and one with six-fold symmetry. EM and image analysis was used to establish the relative orientation *E. coli* RuvA and RuvB bound to model Holliday structures (58,59).
Clotting factors with inhibitors	XRCrys	Crystals structures of the complex of thermitase (279 amino acids) with the inhibitor Eglin C (70 amino acids), resolved to <2 Å resolution (60) and human α-thrombin with its inhibitor hirudin to 2.3 Å resolution (61).
Cyclosporin/cyclophilin	XRCrys, NMR	NMR of the binding of cyclophilin to cyclosporin indicate that cyclosporin A and FK506 (but not rapamycin) undergo drastic conformational changes when bound to their respective binding proteins (62).

Table 2. Model Systems for the Study of Protein Aggregation (*continued*)

Model system	Studied with	Examples and key results
α-Crystallins	EM, SLLS, DLLS	Dynamic light scattering data (determine particle molecular weight as a function of time) combined with EM to determine shape of aggregates of bovine α-crystallin. Static light-scattering data indicated that the sample preparation methods used for (negative staining) EM had induced some particle breakdown (63).
Making lac repressor tetramer into a dimer or monomer; mechanism of isopropyl 1 thio-β-D-galactoside (IPTG) binding	SAXS, SANS, XRCrys, SDM, equilibrium sedimentation, fluorescence emission quenching	Tetrameric (4×360 amino acid) wild-type lac repressor was modeled based on scatter data (size and shape of tetramer) initially until a crystal structure was available. Targeted SDM-generated monomers formed preferentially dimers or did not associate, while retaining DNA binding function. A point mutation (Y282D) was sufficient to disrupt tetramer formation and generate monomeric protein (64). IPTG (inducer) binding reduced the fluorescence emission intensity of the repressor in solution. Low concentrations of IPTG reduced the fluorescence quenching effect of KI on the emission intensity (65).
Making a dimer from monomeric staphylococcal nuclease	XRCrys, sedimentation equilibrium	A serendipitous mutant of staphylococcal nuclease (6 amino acids in a surface loop, Δ114–119) forms a stable dimer ($Kd < 10^{-8}$ M). The carboxyterminal helices of each monomer swap places in the dimer (66).
		In vitro aggregation
	Light scattering, laser light transmission fluctuation	For monoclonal antibody aggregation in the presence of antigen, the estimates of the radius of gyration in solution, as measured by both classical and dynamic light scattering, agreed very well with those estimated from electron micrographs of the same aggregates (67). The aggregation of low-density lipoprotein (LDL) induced by acetylation, carbamylation, maleylation, or oxidation was evaluated by laser light transmission fluctuation. Aggregated LDL, but not unmodified monomer, stimulates the uptake of cholesteryl ester in arteriosclerotic cells (68).
Protein denaturation	CD, FTIR, IR-ATR, RS, SLLS, NMR	Raman spectroscopy has been used extensively to study the quaternary interactions of hemoglobin and to compare the secondary structures of proteins under various conditions (69,70). For example, 10–20% solutions of α-chymotrypsin and chymotrypsinogen were used for a Raman spectroscopic analysis of secondary changes induced by pH and pressure (71). A relationship between the salt type and concentration used for precipitation of α-chymotrypsin and the disordering of the secondary structure was determined from Raman spectroscopy (72). When the method was extended to another 11 proteins, the general result was that the precipitated protein contained on the average a higher percentage of β-sheet structure as determined by the amide I band intensity and location; the effect was more pronounced when the chaotrophic salt KSCN was used compared to the structure stabilizing salt, Na_2SO_4 (73).
Salt precipitation of protein	RS, SANS, fluorescence scatter peak, SLLS, DLLS	SANS in combination with ultracentrifugation has been used to determine the particle weight of malate dehydrogenase from a halophilic bacteria in different salt solutions. In 1–2 M NaCl solution, the protein is a stable dimer and binds much more water and NaCl than in low salt solution, where the salt-loving protein unfolds (74). Stopped-flow turbidimetry has been used to derive a kinetic model for the precipitation of α-chymotrypsin by different salts (75).
Crystallization pathways	Fluorescent light scattering, EM, SLLS, DLLS	Quasielastic light scattering study of the effect of salts on lysozyme and concanavalin A showed precipitating salts decrease translational diffusion coefficient, whereas those furthering crystallization do not change it and solutions remain monodisperse with increasing salt (76). Light scattering studies of solutions of canavalin at a fixed salt concentration (where the protein was known to form rhombohedral crystals) (77). Stereo-electron microscopy of frozen specimens used to characterize crystallization methods for lysozyme and tomato bushy stunt virus particles (330 Å). Crystallization intermediates had a more compact structure than aggregates that lead to amorphous precipitates (78). A symmetric dimer of a polypeptide in an organic solvent suggested as a model for the effects of crystal packing and solution dynamics on structure determination by X-ray crystallography and solution NMR methods (79).

Table 2. Model Systems for the Study of Protein Aggregation (*continued*)

Model system	Studied with	Examples and key results
Complexes with phospholipids and surfactants	Solid state NMR, NMR, ATR/IR, ESR	The complete structure determination of a polypeptide in a lipid bilayer derived from solid-state NMR observations (80). IR has also been used to study the binding of CO_2 to carbonic anhydrase and mutants thereof (81), and combinations of ^{31}P-NMR and ATR-IR have been used to study pulmonary surfactants (82,83). The interaction of phospholipid with (*R*)-3-hydroxybutyrate dehydrogenase, a phosphatidylcholine-requiring membrane enzyme, has been studied using ESR spectroscopy of spin-labeled lipids, both as ordered multibilayers and in lipid vesicle suspensions (liposomes) (84).

Note: The systems described in this chapter are amenable to characterization by several methods. See the section under the appropriate method for details and references. CD = circular dichroism; EM = electron microscopy; FTIR = Fourier-transformed infrared spectroscopy; IR-ATR = infrared attenuated total reflection spectroscopy; RS = Raman spectroscopy; SE = sedimentation equilibrium; SANS = small-angle neutron scattering; SAXS, small-angle X-ray scattering; SLLS/DLLS = static or dynamic laser light scattering; STM = scanning tunneling microscopy; XRCrys = X-ray crystallography.

tected by their ability to protect the native state of the protein against the stress of emulsification in organic solvent and re-extraction into buffer. Recovery of native protein was determined by native size exclusion HPLC followed by circular dichroism (CD), both rapid techniques that require relatively little protein (90).

Determining the Surface Charge and Hydrophobicity of a Protein

Again, practicality dictates the use of chromatographic methods to determine the approximate surface charge and hydrophobicity of a protein in solution. Because aggregation of integral membrane proteins after transfer to an aqueous environment is a well-described phenomenon (91), there is a tendency to regard precipitation in aqueous solvents as an indication that a protein is hydrophobic. In reality, all proteins are to some extent hydrophobic, with tightly packed cores that exclude water (92,93), and even proteins with many polar residues at their surfaces, such as interferon-γ and RNase A, will aggregate at high concentrations.

As with stability, the surface charge of a protein is not easily defined. An isoelectric point (pI) measured by isoelectric focusing or chromatofocusing indicates the pH at which the protein has no net charge at very low salt concentrations (high ionic strength prevents proper electrophoresis and interferes with the establishment of the ampholyte-pH gradient). These methods cannot be used for proteins (e.g., T7RP) that precipitate in low salt buffers. Further, the effective pI of a protein is greatly affected by buffer conditions. Due to the binding of salt, one cannot assume that a protein in solution will be negatively charged at a pH above its pI in a solution containing salts or metal ions (e.g., acidic caseins bind Ca^{2+} and appear positively charged at pH 7 (16). Oligomerization can alter the apparent isoelectric point of a protein by more than a pH unit, as has been shown for insulin (94).

At pH 7.5 and 50 mM salt, most proteins will bind to DEAE-coupled (positively charged) resins if they are negatively charged and to phosphate and other negatively charged resins if they are positively charged. The degree of surface charge can be estimated from the salt concentration required for elution from DEAE or CM-conjugated resin columns; this can also give indirect evidence about the degree of surface exposure of altered residues. For example, when positively charged residues (Arg10 or Lys7) were changed to uncharged glutamine by site-directed mutagenesis from the N-terminus of RNase A, the mutants eluted at lower salt from CM-Sepaharose at pH 8 than did the wild type; the double mutant did not bind to the column at this pH (95). Similarly, C-terminal deletions of human interferon-γ eluted at progressively lower salt concentrations from CM-Sephadex as more positively charged residues (Lys [K] and Arg [R]) were deleted (Table 3).

Later in the purification, after the solubility of the protein has been better defined, chromatofocusing in an appropriately stabilizing buffer can be used to more accurately estimate the protein's pI. Once one has enough purified protein, the pI can be determined by measuring the amount of KOH necessary to increase the pH of a solution of protein in buffer from 2.5 to 11 compared to that needed to titrate the buffer alone. Gel methods for following the changes in surface charge during protein folding and aggregation have also been developed (42).

Similarly, binding to resins coupled with hydrophobic groups, such as Phenyl- or Octyl-Sepharose® (Pharmacia) indicates the presence of hydrophobic residues at the protein surface. Proteins are applied in high salt (0.7 to 1 M ammonium sulfate), which increases hydrophobic interactions, and then eluted with a decreasing salt gradient. Most proteins elute between 0.5 and 0.1 M salt; very hydrophobic proteins will not elute into low salt buffer unless the polarity is decreased by adding ethylene glycol. If a protein does not bind to Phenylsepharose, it either has a very hydrophilic surface (e.g., RNase A) or is aggregated.

One can determine the hydrophobicity of a purified protein or follow changes in exposure of hydrophobic groups during folding by measuring interaction with a hydrophobic dye or radioactive tracer [e.g., 1-anilino-8-napthalenesulfonate (98), or ^{125}I-TID, 3-(trifluoromethyl)-3-(*m*-^{125}iodophenyl)diazirine (99)].

Empirical Models for Salting-in and Precipitation with Various Agents

Empirical Models to Account for Salt/pH Effects. Generally, charged proteins can be salted-in by counterions. This

Table 3. Proteolytically Cleaved Human IFN-γ Elutes from CM-Sephadex at Progressively Lower Salt Concentrations Depending on the Number of Positively Charged Residues Removed from the C-Terminal

Human IFN-γ	C-Terminal	[NACl] eluted
Full length	– PAAKTGKRKRSQMLFRGRRASQ	300–500 mM
IFN-$\gamma\Delta$12	– PAAKTGKRKR	200–300 mM
IFN-$\gamma\Delta$14	– PAAKTGKR	80–100 mM

Note: After overnight cleavage at room temperature with one unit of either the protease Factor Xa (Biolabs) (for the Δ12 protein) or Endopeptidase Arg-C (Sigma) for the Δ14 protein, samples were diluted fivefold with 10 mM $NaHPO_4$ buffer, pH 5.7, and applied to small columns of CM-Sephadex C25 (Pharmacia) (96,97). The indicated eluates were then concentrated by centrifuging in Centricon-3 microconcentrators (Amicon). After automated Edman degradation, which indicated that the N-terminus of the proteins was unaltered, the number of amino acids deleted was determined by electrospray ionization mass spectroscopy.

is caused by the ion binding to the protein and neutralizing the net charge at its surface (see Table 4 for dielectric constants of common ions). The solubility of lysozyme, a positively charged protein, was shown to vary more with the nature of the anion added than the cation; the anion dependence followed the Hofmeister series (100). On the other hand, the solubility of caseins with pI between pH 3 and 5 varies with the nature of the cation: sodium, potassium, and ammonium caseinates are soluble; calcium caseinates form micelles (101). Casein solubility varies inversely with the atomic radius of the divalent cations added; Zn^{2+} and Cu^{2+} are more effective precipitants than Co^{2+} (102).

The interaction of salts as a function of pH with very soluble proteins can be accounted for mathematically, on a case-by-case basis. Detailed studies have been reported for a few proteins, including porcine pancreatic amylase (103) and ovalbumin (104). The latter study found that at 30 °C, the solubility of ovalbumin (C_o) in ammonium sulfate solutions of varying concentration C_A could be described by the simple quadratic equation:

$$\log_{10}(C_o) = 5.06 - 0.006t\text{-}0.205\ C_A + 0.5(\text{pH-}4.58) + 1.1(\text{pH-}4.58)^2$$

where concentrations of both ovalbumin and ammonium sulfate are expressed in g/100g H_2O, and 4.58 is the isoelectric point of ovalbumin. The data agreed quite well with earlier measurements made at 18 °C if the temperature correction was added.

Precipitation with Polyethylene Glycol and Other High MW Polymers. Precipitation as a means of purifying or storing a protein in a dry state must be done in such a way that after removal of the precipitant, a soluble and active product is obtained. Organic solvents that perturb the aqueous shell around the protein (ethanol, acetone, and even trichloroacetic acid [105]) are used industrially for large-scale protein precipitation, where maintaining a low temperature during the procedure is usually required, but the most useful general precipitants are those that do not denature the protein, such as polyethylene glycol (PEG). Models for precipitation with such compounds are usually based on excluded volume, meaning that the protein is forced into a progressively smaller area as the polymer occupies more of the solvent until the critical solubility limit of the protein is reached.

One recent study analyzed precipitation of common proteins (bovine serum albumin, human serum albumin, and ovalbumin) with various weights of PEG, and assumed:

1. the protein is globular and has homogenous surface properties
2. the precipitating polymer is linear, nonionic, and homogenous

With regard to the first assumption, the model is only useful for charged proteins if the ionic strength of the solvent is adjusted to minimize electrostatic protein–protein interactions (106). To predict the phase diagram, the Gibbs free energies (evaluated using a second-order perturbation

Table 4. Electronegativity Values for Protein Atoms and Common Ions Associated with Proteins

Electronegativity of protein atom constituents		Ionic character of bond and (percent ionic)		Electronegativity of ions	
H	2.1	C-H, S-H	0.4 (4%)	Li	1.0
C	2.5	C-O	1.0 (22%)	Na	0.9
N	3.0	N-C	0.5 (7%)	K, Rb	0.8
O	3.5	OH	1.4 (39%)	Ca	1.0
S	2.5	NH	0.9 (19%)	Cu	1.9
P	2.1	P-O	1.4 (39%)	Zn	1.6
				Fe	1.8
				Cl	3.0
				F	4.0

theory around a hard sphere as the reference) of protein in the solid and liquid phases were equated. Protein solubility was lowest (concentration of protein in the solid phase increased) for higher polymer concentrations, larger protein sizes, higher molecular weight of polymer, weaker protein–polymer interactions, poorer solvent, higher ionic strengths, and pH values close to the isoelectric point (107).

Because precipitation models are typically based on data for large, quite soluble proteins with the polymer used for precipitation of the same size and at about the same density as the protein being precipitated, it is difficult to apply the results directly to most laboratory-scale processes. Here, typically, the polymer greatly outweighs the protein in solution, and the intrinsic solubility of the protein is much less than the standard proteins.

Models for Determining the Effect of Co-solvents on Protein Folding

Accounting for the interaction between parts of the protein and the surrounding solvent is one of the most rapidly developing areas in protein folding. There are several theories to explain the effects of denaturants and osmolytes on protein solubility. The eventual outcome of this work is to understand how a protein folds, with an eye to predicting tertiary from primary structure. Determining the effects of osmolytes and denaturants on protein stability in solution should enable us to predict a priori what cosolutes are best suited to a protein of a given composition.

Osmolytes have strikingly different effects on different proteins. One cannot simply add any osmolyte to a protein solution and assume it will increase solubility; indeed, addition of carbohydrates or other osmolyte to a concentrated protein solution will often induce precipitation. The osmolytic amino acids betaine and proline, which protect *Escherichia coli* against urea and salt stress (108), can destabilize DNA helices (109). Adding side chains to betaine generated forms that are protective against high salt but not urea and vice versa (108). Curiously, although osmolytes have the general property of increasing the net solubility of proteins in solution, that is, one can achieve a higher concentration of the protein before aggregation occurs, none of the individual amino acids are more soluble in the osmolytic solutions for which there are data (0.5 M sucrose [110]; 1 M trimethyl amine *N*-oxide [TMAO] [111]). There is also no really significant increase, regardless of side chain, in amino acid solubility in the presence of 2 M urea, although a cyclic glycine dipeptide designed to mimic the backbone unit is about 20% more soluble (111).

The Thermodynamic Approach. The thermodynamic approach is based on the relative ability of molecules to stabilize the native state or the denatured state, thus accounting for their effects in a thermodynamic framework. As one cannot explain the specificity of osmolytes by difference in solubilities of the amino acid side chains, the preferential hydration theory suggests that osmolytes are repelled from the surface of the protein, whereas denaturants bind. The basic theory resembles the excluded volume theory of protein precipitation and deals with the protein as a hydrated ion. The binding of water molecules to the surface of the protein serves to shield charges and protect the protein from interaction with other proteins. According to the model, the concentration of the cosolvent in the immediate vicinity of the protein is either less than (for osmolytes) or greater than (for denaturants) that in the bulk solution. Thus, the water shell around a protein in osmolytic solution will be relatively richer in water molecules than the bulk solution. The theory assumes first that such a concentration gradient surrounding the protein exists and that this can serve as a driving force for stabilization.

The concentration of cosolvent near the protein surface is not easily measured. Experimental results to demonstrate that this is different from that in the bulk solution are based on measurements of the partial molar volume of the protein in various solutions or more recently, infrared spectroscopy to measure O-H stretching rates in the vicinity of the protein (112). Barrier effects caused by the presence of high concentrations of cosolvents in the binding of CO to myoglobin have been analyzed by frequency resolved calorimetry (113). To develop more general ways to account for the effects of solvent at the protein surface, computer models have been developed. Explicit modeling of solvent, by inserting the protein into a shell of solvent of varying thickness, is computationally inefficient. Drastic perturbations have to be imposed in the unfolding simulation, such as high temperature or pressure or by applying special algorithmic means to unfold a protein (114–116), which are likely to induce pathways not found in nature. Alternatively, one can implicitly model solvent molecules, in an approach based on the continuum approximation and the calculation of the solvent accessible surface area (SASA).

This approach was experimentally validated by a recent statistical data analysis that showed that for 45 proteins, the measured free energy and heat capacity of unfolding correlates directly with the change in SASA during the process (117). A recent method was introduced to efficiently calculate the SASA (118). By changing the protein–solvent interaction via the SASA and extrapolating to zero cosolvent concentration, a folding pathway that should have more physiological meaning can be induced and monitored.

In this case, the protein–solvent interaction term E_{hyd} is computed in the continuum approximation by

$$E_{\mathrm{hyd}} = \sum_{i=1,\mathrm{atoms}} \sigma_i A_i$$

where A_i is the SASA of atom i and σ_i is a solvation-parameter (119–121) depending on the atom type and solvent composition. A variety of data are available for the determination of σ_i values for individual protein atoms. Several such atomic solvation parameter (ASP; σ_i) sets have been derived (119,120,122–126) based on deriving a transfer free energy (ΔG_{tr}) for moving organic molecules (usually amino acids or their analogues) from one medium to another. Determination of the ΔG_{tr} can be calculated from the relative solubility of the groups in the two media. In the latter case, solubilities in the two solutions (C_w and C_s for maximal solubility in water or alternative solvent, respec-

tively) yield ΔG_{tr}, by assuming that the ratio of the activity coefficients is ~1, according to the equation (127):

$$\Delta G_{tr} = RT \ln \frac{C_w}{C_s}$$

These ΔG_{tr} values can be directly related to the accessible surface area of each amino acid (A_i) by the equation: $\Delta G_{tr} = \Delta\sigma_i A_i$, and it is assumed for parameter extraction that the ΔG_{tr} of a residue R is the sum of the free energies of each atom:

$$\Delta G_{tr}^{R} = \sum_{\text{atoms}j} \Delta\sigma_j A_j$$

Eisenberg and McLachlan (119), in one of the earliest systematic studies, used data from Fauchère and Pliska on the energy of transfer of amino acid side chains from *n*-octanol to water (128) to derive a set of five ASP values ($\Delta\sigma$(C), $\Delta\sigma$(N/O), $\Delta\sigma$(O−), $\Delta\sigma$(N+), $\Delta\sigma$(S)). Their calculated ΔG_{tr} for the side chains using this set was linearly related to the observed values, and others have shown that the parameter sets, with slight modification, can drive disturbed structures back to their native state (118,129–131). Although one can also expand the parameter list by distinguishing aliphatic, methylene or carbonyl-C, and backbone or guanidyl nitrogen atoms, too many parameters may be detrimental in determining accurately the effect of the interaction. Other data well suited to the derivation of ASPs are those for the transfer potential of side chain analogues to a vapor phase (132) or based on measurement of the hydration enthalpy of *N*-acetyl amino acid amides (133), or the isoentropic pressure coefficients of tripeptides (134). Data are also available to derive ASP sets for solvents containing denaturants (135) or osmolytes (110).

There are data available for the solubility of protein components in ethanol, octane, and other solvents that could be used to derive ASPs for calculation purposes and direct data for purifying proteins with aqueous ethanol extraction, or following the unfolding of proteins or peptides in various solvent conditions (136,137). As the ability of water-miscible organic solvents to solubilize naphthalene is logarithmically related to the concentration at which they denature well-known proteins (trypsin, cytochrome *c*, chymotrypsinogen, chymotrypsin, laccase, and myoglobin) (138), ASP sets derived for one solvent may be related to other solvents according to simple correlation constants.

In support of using such ASPs, the hydrations seen around atom types in protein crystals (Table 5) show specific chemical interactions of the individual atoms with water that are unaffected by the crystal lattice. Neutral and negatively charged oxygen atoms are the most hydrophilic, followed by positively charged nitrogen atoms, neutral nitrogens, carbons, and sulfurs ($O^- \sim O > N + > N \gg C \sim S$, where ~ means approximately the same).

Interactions between protein and bound water are fundamentally similar to those between protein and bulk fluid water, measured in partition experiments (139). The linear correlation coefficient of the data from Table 5 with side-chain free energy of transfer between cyclohexane and water (132,140) was $r = -0.95$ and with residue retention on reversed-phase HPLC using a pH 7.5 average scaler $= -0.93$ and $r = -0.86$ for the data of Fauchère (128).

The Kinetic Approach. A kinetic-based accounting for the effects of cosolvents on aggregation is based on their alteration of bulk solution properties, primarily the viscosity, macroscopic dielectric constant, and the chemical potential, which affect also the unfolding of the protein and its thermodynamic equilibrium energy. These changes can also have pronounced effects on the kinetics of unfolding, refolding, and transitions between substrates, which must be dealt with to account for the effects of osmolytes on aggregation and solubility. Aggregation, the major barrier to obtaining clear thermodynamic data, accelerates with increasing protein concentration. This strongly suggests that protein precipitation is directly related to the rate of protein–protein interaction. Kinetically, the role of osmolytes in maintaining proteins in equilibrium solution can be described by equations initially based on Newton's law of viscosity and Archimedes' principle. A protein in solution is in buoyant equilibrium with the water structure and viscosity-based forces accelerating its precipitation. To remain in aqueous solution, the protein must inhabit a cavity in the solvent structure somewhat larger than its surface area. The energy to create such a cavity has been estimated, based on the surface tension of water being 72 dyn/cm, to be 104 cal/Å/mol (141). This energy must be offset by the energy of binding of water molecules to the charged groups on the protein molecule. This may explain why more highly charged proteins, such as RNase A, are quite stable in solution in the presence of moderate amounts of salt, in contrast to proteins with large hydrophobic surface areas. A compact globular shape has the lowest surface area; unfolding the protein increases its surface area (and area for interaction with other molecules) and alters its specific density. For example, for a protein in a temperature gradient, the change in velocity of the molecule can be summarized as a function of the (temperature-dependent) viscous forces [$\nabla \cdot \tau$] and buoyant forces [$\rho\beta g\Delta T$]:

$$\rho \frac{Dv}{Dt} = -[\nabla \cdot \tau] - \rho\beta g(T - T)$$

where ρ is the specific density (mass/unit volume), g is the gravitational constant, and β is the coefficient of volume expansion.

The second term in kinetic effects in protein unfolding states simply that the viscosity of the solution will affect the time required for folding, an effect summarized (142) as:

$$\bar{t}(n) = t_0 e^{(\Delta E(n)/k_b T)^2}$$

where $t(n)$ is the lifetime of a typical microstate, and t_0 is the time scale for a typical motion of a large segment of the chain, which is a factor of the local barriers and the solvent viscosity. For the case of viscosity (friction) effects on geminate binding reactions, please see Ansari et al. (143) for a restatement of Kramer's law in order to apply it to proteins.

In addition, one must account for changes in the dielectric constant of the medium and for the polarizability (α) of the protein and its cosolvents on the interaction energy between protein particles as a function of distance (R). Classic thermodynamics says that this energy, for an oscillating system with no limits on the sphere of motion, should be:

$$\Delta E(R)_{\text{dis}} = -3hv/4(\alpha^2/R^6)$$

where the oscillator energy hv can be equated with the ionization potential (IP). Thus, the energy of interaction decreases as the distance between particles decreases to the sixth power and the polarizability increases. As the distance between particles decreases directly with the concentration of the protein, small increases in concentration will greatly accelerate the aggregation rate.

With respect to the ionization potential, the constituent atoms of proteins, with the exception of hydrogen, have very similar electronegativity values (Table 4), and most of the bonds within a protein are basically covalent. Note that hydration is seen in crystal structures at precisely the atoms in proteins where the ionic character of the bonds is most pronounced (Table 4). Obviously, substituting a hydrogen atom for any of the alkali metals will increase the local dipole moment within the protein. The polarizing effect of substituting a larger ion is intermediate. These considerations may be used as the basis for analysis of salt interactions with proteins, if the surface charge is known. Although such models will still require experimentation, they may be used as an optimization framework that may reduce the data needed.

COMPUTER DESIGN OF MORE SOLUBLE PROTEINS

Combinations of the methods discussed in the first section have been used to generate mutants of insulin with altered oligomerization and solubility and to change the lac repressor from a tetramer to a dimer or monomer, to name but a few examples (see Table 2). Residues involved in aggregation were identified with the help of a detailed 3-D structure, generated by high-resolution methods, and site-directed mutagenesis was used to test which residues could be altered without affecting function or the global fold of the protein (144). All these studies required years of work. Fortunately, computational advances now allow us to have a reasonable view of the protein structure for many proteins based on their sequence homology to other proteins. This sort of modeling can be implemented to solve problems even during the initial isolation of a protein of known sequence. As discussed elsewhere (2), there are a few empirical rules for identifying areas of proteins that cause aggregation. The few good examples of controlling aggregation of proteins through alteration of primary structure started from high resolution X-ray crystallographic or NMR structures of protein complexes or oligomers. It is now possible to view the 3-D structure of proteins whose coordinates have been deposited in the protein data base (PDB) (for downloading from the Internet, see the notes to the Internet site "Pedro's Macromolecular Tools") on a regular Macintosh or PC terminal; stereo viewing software can also be had for a reasonable price. Good viewing systems such as the Silicon Graphics Indigo series or the Evans and Sutherland Molecular Graphics systems allow more detail and sophisticated, real-time rotation of the molecule. Even when the 3-D fold of the protein is known, it is still not easy to identify aggregation sites; one can, however, identify surface-exposed amino acid stretches that might be involved in aggregation or intermolecular contacts rapidly and test the effects of site-directed mutagenesis at these positions on solubility. New methods for mimicking protein folding by computer allow hope that we may soon be able to identify problem areas in the primary sequence. Two basic assumptions may greatly simplify calculations of the protein aggregation problem:

1. Protein folding proceeds first by folding into a series of domains, which then interact with each other specifically to form the final stable global structure.
2. During this process, certain domains may form inter- rather than intramolecular contacts. Once two molecules are so intertwined, their residual domains may be forced to associate with other molecules, leading rapidly to a mass of partially folded, interlocked protein monomers. This idea, based on the domain swapping hypothesis of Bennett et al. (145), is from structural analysis of oligomeric proteins, including *Diphtheria* toxin, staphylococcal nuclease, α-spectrin, and antisialidase, with swapped domains ranging from 14 to 252 amino acids.

The problem would then become identifying probable domains within the protein that are sufficiently separated during folding so as to allow another molecule to intercalate during the folding process. One can view surface groups on proteins, searching for areas of exposed hydrophobic surface, for example, that may be involved in aggregation. There are also specific programs that allow one to view the dynamics of interaction between molecules (DOCK, from Kuntz's group at UCSF, can be obtained over the Internet). DOCK is most suitable for viewing the interaction between small sections of the molecule and small

Table 5. Hydrations Seen in Crystal Structures

Protein atom group	Atom types included in group	Water ions/occurrence
Neutral O	Q, N, D, T, Y and all carbonyl O atoms	0.53
Negative O	D and E side chain O	0.51
Positive N	H, K, R side chain N except Arg Nε, which is fairly nerutral	0.44
Neutral N	W, N and Q side chain N and Arg Nε, plus all main chain N	0.35
Carbon/sulfur	All carbon and sulfur atoms	0.08

ligands. For example, the 3-D structure of neurotrophin (also called p75NGFR or LANR) was modeled by its homology to regions in the receptor of tumor necrosis factor. Hypothetical complexes of this neurotrophin model structure with nerve growth factor were then modeled with DOCK 3.5. Viewing of hundreds of predicted orientations with MIDAS revealed similarities between motifs in ligand-binding domains of the neurotrophin and tumor necrosis factor receptor and pinpointed differences that could account for specificity (146).

However, experimental structural data are not available for most proteins, and structures of complexes of proteins are even rarer. For a significant percentage of newly identified protein sequences, which have homology to a protein of known sequence homology, one can predict with some degree of confidence their overall topology and stretches of amino acids likely to be surface exposed. This does not mean that such models can replace experimental data. The protein-folding problem is mathematically described as a system with frustration, because the number of computing steps required to solve the problem increases faster than any power of the size of the system (147). This section covers some of the possibilities for using sequence data and molecular graphics tools to generate model 3-D structures. Some programs from the Internet are presented in addition to the self-correcting distance geometry (SECODG) approach to the modeling being developed in the computational biology group of the UTMB. Several different pathways, starting from an initial sequence alignment, can be used to obtain such constraints (Fig. 2).

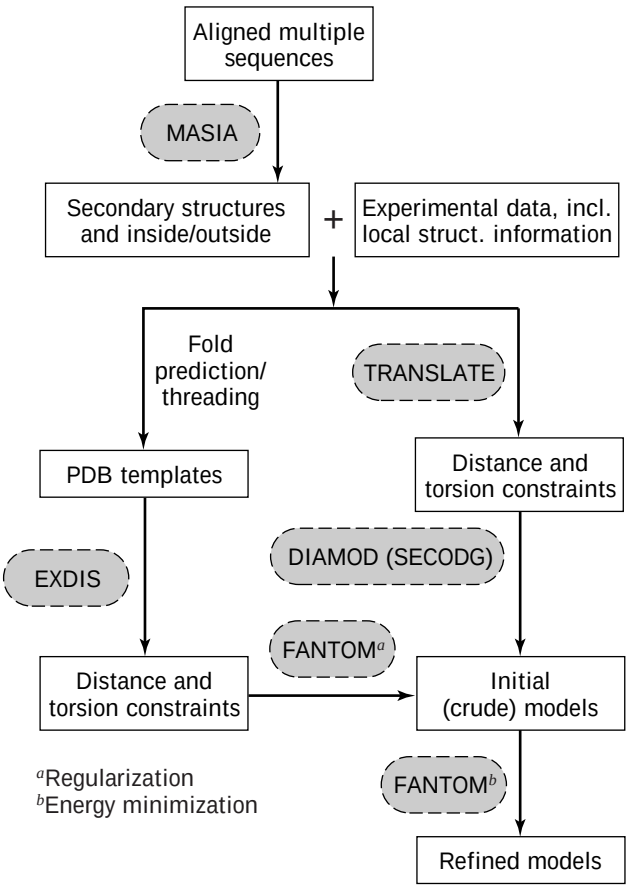


Figure 2. Flow chart describing protein modeling strategies depending on the degree of homology of the protein sequence to those of known structure. The *shaded boxes* enclose names of computational biology software that is being developed in the computational biology group at UTMB. The left and right sides of the flow chart represent procedures for generating models by independent routes. See text for details. *Source*: Figure kindly supplied by Dr. Kizhake Soman of UTMB, who retains copyright.

Automatic Homology Modeling

If a structure for a protein with significant (>30 to 40%) sequence identity exists, a 3-D-structure model can be obtained over the Internet. For example, we tested SWISSMODEL using a 62-amino-acid section that aligns with other homeobox proteins from the Pem protein (involved in control of hormone generation in the testes). The murine Pem homeodomain homologous area

RQMPLQGSRFAQHRLRELESILQRTNSFDVPREDL
DRLMDACVSRVQNWFKIRRAAARRNRR

sequence was cut out of a text file and pasted into the appropriate area of the form for submission to SWISSMODEL, a program on the Internet that is underwritten by Glaxo Research and constructed by Peitsche. An e-mail response with coordinates for a 3-D structure based on the structure of the Fujitaratsu homeodomain (PDB files 21FJL and 11 FJL) was returned quickly. Although for proteins with sufficiently high homology to one in the PDB the template selection is automatic, it is always a good idea to choose an appropriate model oneself by extensive homology (BLAST or equivalent) searches before requesting the modeling.

SWISSMODEL uses the PDB to obtain structural information and automatically generates a model for sequences that share significant similarities with at least one protein of known 3-D structure. Given an unknown sequence, SWISSMODEL uses a FastA and BLAST search to determine what protein in the PDB could be a possible structural homologue based on the degree of primary sequence similarity. A sequence with a FastA score 10 SD above the mean of the random scores and a Poisson unlikelihood probability P(N) of 10–5 will be considered. The program as of 1998 requires at least 35% residue identity with 40% of the target sequence (or overall 25% identity). The automatic alignment tool used within the program suite is SIM, but the user on a second level can choose to submit other alignments or structures for use in the modeling (148). Hyperlinks between SWISSMODEL and the 2-D gel electrophoresis data bank of SWISS-2DPAGE (see *http://expasy.hcuge.ch*) can be used to map *E. coli*, *Haemophilus influenzae*, *Mycoplasma genitalium*, *M. tuberculosis*, and *Bacillus subtilis* sequences on the basis of charge and size to proteins seen on the gel grid. A similar gel database is being accumulated for yeast and human proteins (149).

The ease, speed, and impressive level of prediction of SWISSMODEL is certainly a tribute to the Glaxo team. A significant amount can be learned from visualizing the structure using Midas software on a Silicon graphics ter-

minal, and residues involved in DNA binding, on the basis of the homologous protein used to develop the model, can be predicted. However, the program only works for proteins where a high degree of sequence homology exists to protein of known structure. By some estimates, only 10 to 15% of *E. coli* proteins have a suitable modeling template in the PDB using the above cutoff levels of residue identity; one would anticipate an even lower probability for an unknown mammalian sequence. For our example, when the whole PEM protein sequence, or even the 63-amino-acid sequence mentioned earlier that aligns with many other homeodomain proteins, was sent in the same fashion to SWISSMODEL and submitted without naming a target protein for modeling, the program was unable to generate even a model of the area that contained the homeobox. A second problem with using such simple models to design site-directed mutagenesis experiments is that although the program will be able to predict a global fold for the protein, local conformations that can account for specific properties are less easy to predict. For this reason, developing a working model (as of the date of this article) means using manually interactive programs that require more work on the part of the investigator.

Other methods are available from the Web and should also be investigated. For example, one can use the TOPITS program, which threads protein sequences onto model structures (150). Here, sequence residue identity is less important than the stability of the final structure. An example of using this program is shown in Figure 3.

Modeling Using Self-Correcting Distance Geometry with the Programs CLUSTAL, MASIA, NOAH, DIAMOD, and FANTOM to Develop a 3-D Model of a Protein

Figure 2 summarizes methods used by our group for structure prediction based on a unique algorithm based on self-correcting distance geometry (SECODG) from Dr. Werner Braun and coworkers. The SECODG-based program suite has the advantage that experimental results can be directly incorporated into the modeling process by manual insertion of distance constraints (DCs), limits on the distance between two atoms in a 3-D structure. Experimental DCs can be derived experimentally from NMR or crystal data or from other types of biophysical data, for example, from knowing where secondary structure elements lie, which two cysteine residues form a disulfide bond, or which residues lie within the active site (which means they must lie within a certain distance from each other). The method starts with BLAST searches and multiple sequence alignment to determine whether a structure exists for a protein with reasonable sequence homology. If a homologous structure is available, the PDB coordinates can be regularized and DCs extracted automatically by related software programs.

Step 1: Proteins with a significant identity pattern to the test protein are searched for in a protein sequence database (for example, SWISSPROT).

Step 2: CLUSTALW (downloaded from the Internet) is used to align the best-matched proteins in a block to identify areas of high homology with other proteins. These sequences are then used to search for known 3-D structures in the PDB, a composite of tertiary structure details.

Step 3: The program MASIA (formerly MULTAN [130]) analyzes the aligned sequences and identifies areas of probable secondary structure.

Step 4: The results from steps 2 and 3 are combined to generate a list of distance and dihedral angle constraints. These constraints can be combined with results from other biophysical methods, such as knowledge of where disulfide links and active site residues are in proteins.

Step 5: The SECODG-based package of NOAH/DIAMOD then uses these constraints to generate a family of structures that best fit these constraints. The NOAH program is a structure-based filter, which means in simple terms that it uses a rough 3-D structure to determine whether a list of distance constraints is internally consistent.

Step 6: The local structure is refined using the energy minimization program FANTOM, to determine structures with lowest energy.

For examples of how this method can be applied to real situations, see the discussion by Mumenthaler and Braun (130). Figure 4 is an example of the structure generated by this method, before energy minimization, for the rat and murine Pem homeodomains by the author and Hong Yao Zhu (Univ. of Texas Medical Branch).

Modeling Complexes

Modeling the interaction of several molecules is important for our understanding of aggregation. Extensive modeling studies for protein complexes has been done, for example, to analyze the allosteric mechanism of the tetrameric lac repressor. Here, years of site-directed mutagenesis studies were done before a crystal structure of the repressor was available for study (152). If crystal structures of the proteins involved are available, and there is some biophysical data about the binding site, one can make relatively exact models using graphics programs such as MIDAS.

Of course, a model must be validated experimentally and theoretically. Tests for structure validity can be based on a binary code where all residues are considered polar except for seven nonpolars (CILMFWV). The arrangement of hydrophobic and polar residues alone as evaluated by a scoring scheme (hydrophobic fitness score) (153) recognized the native fold out of 2,000 near-native structures generated for each of five small monomeric proteins. Other tests for structural validity based on identifying structures with minimum thermodynamic energy are being developed.

Protein Design

The field of protein design (designing proteins that fold in a certain manner) is proving to have nearly as much impact on our understanding of protein folding as the direct measurement of natural proteins. Proteins can be designed to have a specific overall structure with defined secondary elements and selected for solubility (154,155). There have been several recent protein design projects that have illus-

RANK	EALI	LALI	IDEL	NDEL	ZALI	PIDE	LEN2	PDB-ID	NAME
1	76.87	124	22	4	1.76	10	625	1nfk	Nuclear factor κB
2	76.25	117	47	6	1.72	10	378	7api	Modified α1 antitrypsin
3	75.98	120	16	6	1.71	14	576	1vnc	Chloroperoxidase
4	75.40	124	48	8	1.68	13	272	1dih	Dihydrodipicolinate reductase
5	75.20	123	33	4	1.67	11	373	2ach	α1 Antichymotrypsin
6	75.15	124	19	3	1.66	15	1189	1prc	Photosynthetic reaction center
7	74.43	122	20	4	1.62	15	285	2ebn	Endo-β-*N*-acetylglucosaminidase
8	74.38	120	42	6	1.62	15	1078	1nir	Nitrite reductase
9	74.15	110	22	5	1.61	14	280	1pyp	Inorganic pyrophosphatase
10	73.68	123	33	6	1.58	20	403	1atl	Atrolysin C
11	73.62	117	25	4	1.58	14	228	1udh	Uracil DNA-glycosylase
12	73.57	112	15	6	1.58	19	314	1tag	Transducin α + GDP + Mg
13	73.22	117	26	5	1.56	12	449	1rtm	Mannose-binding protein a
14	73.18	120	25	5	1.56	17	421	1ses	Seryl-tRna Synthetase
15	73.15	123	41	4	1.55	10	279	1thm	Thermitase
16	72.38	124	38	3	1.51	6	338	1cse	Subtilisin carlsberg

Figure 3. Output files from the TOPITS program (151) for template proteins when an alignment of HIV-Rev protein sequences was submitted as input. Listed are the best 16 hits of protein structures whose folds Rev is predicted to adopt. EALI = alignment score; LALI = length of the alignment (between rev and each of these proteins); IDEL = number of inserted residues; NDEL = number of insertion points; ZALI = alignment score based on which the list has been ranked; PIDE = % pairwise sequence identity; and PDB-ID = the 4-letter PDB code for each protein. For reliable predictions, ZALI > 3.0. Thus, the fold prediction only reflects remote homology with Rev. *Source*: Unpublished data used with the kind permission of Kizhake Soman and Werner Braun, Structural Biology, UTMB Galveston, who retain copyright.

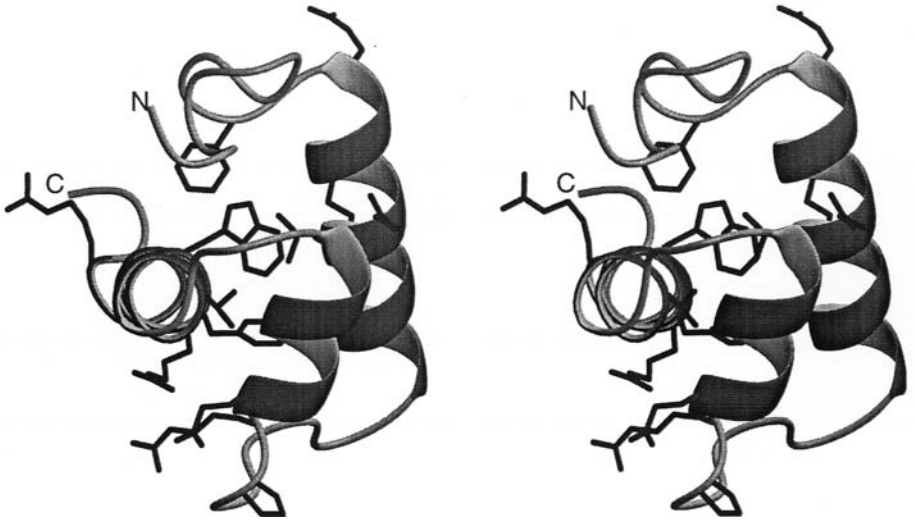


Figure 4. Stereo view of a model structure for the murine PEM homeodomain region, derived by homology modeling based on the 1TTF structure in the PDB using the DIAMOD program. The side chains shown are identical in both the modeled sequence and the template structure. *Source*: Unpublished data from the author and Dr. Hong Yao Zhu, who retain the copyright.

trated how well we can plan a primary sequence that will form a desired structure. Although the design of α-helical proteins is fairly advanced, β-sheet design is still hindered by solubility problems.

For this reason, the Paracelsus prize (small in cash value but honorable in the extreme) was claimed many years after it was offered to any group who could change the structure of a protein from helical to a sheet (or v.v.) without altering more than 50% of the residues. The winning protein was a composite of an area of protein G that formed a β-sheet, which was altered to a four-helix bundle by substituting residues at discrete positions from the ROP protein. Helix-forming residues were inserted, and residues known to favor sheet formation were removed (156). Despite the apparent difficulty humans have in changing the structure of a peptide from α-helical to β-sheet, such a mechanism has been proposed for natural puzzles. For example, aggregation of the proteins associated with scrapies is suggested to be a conversion from a largely helical conformation to a β-sheet-based structure in the aggregate (157,158).

Computational methods have also been introduced into the area of protein design. For example, a protein that contained no histidine, cysteine, or chelated metal ion was designed that exactly fit the X-ray crystal structure of the zinc finger protein. The structure of the synthesized model protein, which required no metal ion for maintenance, was determined by NMR to be identical to the original Zn finger region (159). With time, similar modeling algorithms will presumably allow the design of proteins that match in both structure and their functional abilities, for example, DNA binding (160,161).

CONCLUSION

The study of protein oligomerization with various methods is leading to a better understanding of the atomic basis for aggregation. Many different forms of biophysical data can be used to develop better models of the processes. Understanding how the interactions between solvents and the structure of proteins affects interprotein contact will help design better methods for controlling aggregation and processes dependent on resolubilizing proteins after precipitation.

ACKNOWLEDGMENTS

I wish to thank all my coworkers at UTMB, particularly Kizhake Soman, Hong Yao Zhu, Yuan Xu, Robert Frackiewicz, and Werner Braun, Wayne Bolen for sharing unpublished work and critical thoughts, and David Volk for critically reading the manuscript. The work related to Figures 2 to 4 was supported by grants DE-FG03-96ER62267 from the Department of Energy and #454940 from the John Sealy Memorial Endowment Fund. I also thank those who sent reprints during the preparation of this work and apologize in advance to any authors whose work, as a result of my oversight, is not adequately cited here.

BIBLIOGRAPHY

1. C.H. Schein, *Bio / Technology* **8**, 308–317 (1990).
2. C.H. Schein, *Pharm. Acta Helv.* **69**, 119–126 (1994).
3. C.H. Schein, in G. Georgiou and E. de Bernardez-Clark eds., *Protein Refolding* ACS Books, Washington, D.C., 1991, pp. 21–34.
4. H.-J. Galla, *Spektroskopische Methoden in der Biochemie*, Georg Thieme Verlag, Stuttgart, 1988.
5. B.J. Berne and R. Pecora, *Dynamic Light Scattering*, Wiley, New York, 1976.
6. E. Mancini, F. de Haas, and S. Fuller, *Structure* **5**, 741–750 (1997).
7. Z. Zhou, W. Chiu, K. Haskell, H. Spears, J. Jakana, F. Rixon, and L. Scott, *Biophys. J.* **74**, 576–588 (1998).
8. L. Haggerty and A.M. Lenhoff, *Biotechnol. Prog.* **9**, 1–11 (1993).
9. H.L. Casal, U. Köhler, and H.H. Mantsch, *Biochim. Biophys. Acta* **957**, 11–20 (1988).
10. R.C. Hester and J.C. Austin, in E.D. Schmid, F.W. Schneider, and F. Siebert, eds., *Spectroscopy of Biological Molecules: New Advances*, Wiley, Chichester, England, 1988, pp. 3–10.
11. H. Versmold and W. Härtl, *J. Chem. Phys.* **79**, 4006–4009 (1983).
12. T.D. Veenstra, L.M. Benson, T.A. Craig, A.J. Tomlinson, R. Kumar, and S. Naylor, *Nat. Biotechnol.* **16**, 262–266 (1998).
13. A.J. Wand and S.W. Englander, *Curr. Opin. Biotechnol.* **7**, 403–408 (1996).
14. D.G. Myszka, *Curr. Opin. Biotechnol.* **8**, 50–57 (1997).
15. M. Speed, D. Wang, and J. King, *Nature Biotechnol.* **14**, 1283–1287 (1996).
16. M. Speed, J. King, and D. Wang, *Biotechnol. Bioeng.* **54**, 333–343 (1997).
17. N.R. Watts, M. Misra, P.T. Wingfield, S.J. Stahl, N. Cheng, B.L. Trus, A.C. Steven, and R.W. Williams, *J. Struct. Biol.* **121**, 41–52 (1998).
18. A. Yonath and H.G. Wittmann, *TIBS* **14**, 329–335 (1989).
19. M. Eisenstein, R. Sharon, Z. Berkovitch-Yellin, H.S. Gewitz, S. Weinstein, E. Pebay-Peyroula, M. Roth, and A. Yonath, *Biochemie* **73**, 879–886 (1991).
20. A. Vershoor and J. Frank, *J. Mol. Biol.* **214**, 737–749 (1990).
21. F. Mueller and R. Brimacombe, *J. Mol. Biol.* **271**, 545–565 (1997).
22. K. Luger, A. Mader, R. Richmond, D. Sargent, and T. Richmond, *Nature* **389**, 251–260 (1997).
23. J. Deisenhofer and H. Michel, *EMBO J.* **8**, 2149–2170 (1989).
24. V.F. Bystrov, A.S. Arseniev, A.L. Barsukov, A.L. Lomize, G.V. Abdulaeva, A.G. Sobol, I.V. Maslennikov, and A.P. Golovanov, in O. Jardetsky ed., *Protein Structure and Engineering*, Plenum Press, New York, 1989, pp. 111–138.
25. K.J. Rothchild, Y.W. He, D. Gray, P.D. Roepe, S. Pelletier, R.S. Brown, and J. Herzfeld, *Proc. Natl. Acad. U.S.A.* **86**, 9832–9835 (1989).
26. T.A. Ceska and R. Henderson, *J. Mol. Biol.* **213**, 539–560 (1990).
27. R. Henderson, J.M. Baldwin, T.A. Ceska, F. Zemlin, E. Beckmann, and K.H. Downing, *J. Mol. Biol.* **213**, 899–929 (1990).
28. M. Amrein, Z. Wang, and R. Guckenberger, *J. Vac. Sci. Technol.* **9**, 1276–1281 (1991).
29. R.B. Spruijt, C.J.A.M. Wolfs, and M.A. Hemminga, *Biochemistry* **28**, 9158–9165 (1989).
30. G.D. Henry and B.D. Sykes, *J. Mol. Biol.* **212**, 11–14 (1990).
31. M. Amrein, R. Dürr, H. Winkler, G. Travaglini, R. Wepf, and H. Gross, *J. Ultrastruct. Mol. Struct. Res.* **102**, 170–177 (1989).
32. Y. Hiragi, H. Inoue, Y. Sano, K. Kajiwara, T. Ueki, and H. Nakatani, *J. Mol. Biol.* **213**, 495–502 (1990).

33. B. Greene and J. King, *Virology* **225**, 82–96 (1996).
34. Z. Zhou, S. MacNab, J. Jakana, L. Scott, W. Chiu, and F. Rixon, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2778–2783 (1998).
35. R.L. Stanfield, T.M. Fieser, R.A. Lerner, and I.A. Wilson, *Science* **248**, 712–719 (1990).
36. G. Boulot, J.-L. Eisele, G.A. Bentley, T.N. Bhat, E.S. Ward, G. Winter, and R.J. Poljak, *J. Mol. Biol.* **213**, 617–619 (1990).
37. R. Levy, T. Assulin, T. Scherf, M. Levitt, and J. Anglister, *Biochemistry* **28**, 7168–7175 (1989).
38. J. Anglister, *Q. Rev. Biophys.* **23**, 175–203 (1990).
39. J.G. Rarity, R.N. Seabrook, and R.J.G. Carr, *Proc. R. Soc. London, Sec. A.* **423**, 89–102 (1989).
40. R.J. Leatherbarrow, M. Stedman, and T.N.C. Wells, *J. Mol. Biol.* **221**, 361–365 (1991).
41. P. Stothart, *J. Mol. Biol.* **208**, 635–638 (1989).
42. P.J.M. van den Oetelaar, B.M. de Man, and H.J. Hoenders, *Biochim. Biophys. Acta* **995**, 82–90 (1989).
43. C.H. Schein and M.H.M. Noteborn, *Bio/Technology* **6**, 291–294 (1988).
44. C. Albiges-Rizo and J. Chroboczek, *J. Mol. Biol.* **212**, 247–252 (1990).
45. G.A. Bowden, A.M. Paredes, and G. Georgiou, *Bio/Technology* **9**, 725–730 (1991).
46. C. Shen and R. Murphy, *Biophys. J.* **69**, 640–651 (1995).
47. T. Ikeda, S. Asakura, and R. Kamiya, *J. Mol. Biol.* **209**, 109–114 (1989).
48. R.H. Wade, D. Chrétien, and D. Job, *J. Mol. Biol.* **212**, 775–786 (1990).
49. B. Gaffney, *Ann. Rev. Biophys. Biomol. Struct.* **25**, 431–459 (1996).
50. G.M. Jarori, N. Murali, R.I. Switzer, and B.D. Rao, *Eur. J. Biochem.* **230**, 517–524 (1995).
51. N.J. Murali, G.K. Jaori, S.B. Landy, and B.D. Rao, *Biochemistry* **32**, 12941–12948 (1993).
52. D. Pretzer, B. Schulteis, C.D. Smith, D.G. Vander Velde, J.W. Mitchell, and M.C. Manning, *Pharm. Biotechnol.* **5**, 287–314 (1993).
53. J. Trewhalla, D.K. Blumenthal, S.E. Rokop, and P.A. Seeger, *Biochemistry* **29**, 9316–9324 (1990).
54. C.E. Schutt, U. Lindberg, J. Myslik, and N. Strauss, *J. Mol. Biol.* **209**, 735–746 (1989).
55. E. DiCapua, M. Schnarr, R.W.H. Ruigrok, P. Lindner, and P.A. Timmins, *J. Mol. Biol.* **219**, 557–570 (1990).
56. M. Amrein, A. Stasiak, H. Gross, E. Stoll, and G. Travaglini, *Science* **240**, 514–516 (1988).
57. A. Stasiak, I.R. Tsaneva, S.C. West, C.J. Benson, X. Ya, and E.H. Egelman, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7618–7622 (1994).
58. X. Yu, S.C. West, and E.H. Egelman, *J. Mol. Biol.* **266**, 217–222 (1997).
59. X. Yu, M.J. Jezewska, W. Bujalowski, and E.H. Egelman, *J. Mol. Biol.* **259**, 7–14 (1996).
60. P. Gros, C. Betzel, Z. Dauter, K.S. Wilson, and W.G.J. Hol, *J. Mol. Biol.* **210**, 347–367 (1989).
61. T.J. Rydel, K.G. Ravichandran, A. Tulinsky, W. Bode, R. Huber, C. Roitsch, and J.W. Fenton, *Science* **249**, 277–280 (1990).
62. W. Braun, W. Kallen, V. Mikol, M.D. Walkinshaw, and K. Wüthrich, *FASEB J.* **9**, 63–72 (1995).
63. P. Schurtenberger and R.C. Augusteyn, *Biopolymers* **31**, 1229–1240 (1991).
64. J.M. Chen and K.S. Matthews, *Biochemistry* **33**, 8728–8735 (1994).
65. W.I. Chang and K.S. Matthews, *Biochemistry* **34**, 9227–9234 (1995).
66. S.M. Green, A.G. Gittis, A.K. Meeker, and E.E. Lattman, *Nat. Struct. Biol.* **2**, 746–751 (1995).
67. R.M. Murphy, H. Slayter, P. Schurtenberger, R.A. Chamberlin, C. Colton, and M.L. Yarmush, *Biophys. J.* **54**, 45–56 (1988).
68. V.V. Tertov, I.A. Sobenin, Z.A. Gabbasov, E.G. Popov, and A.N. Orekhov, *Biochem. Biophys. Res. Commun.* **163**, 489–494 (1989).
69. D.L. Rousseau and M.R. Ondrias, in D.L. Rousseau ed., *Optical Techniques in Biological Research*, Academic Press, Orlando, Fla., 1984, pp. 65–132.
70. R.W. Williams, *Methods Enzymol.* **130**, 311–331 (1986).
71. L. Heremans and K. Heremans, *Biochim. Biophys. Acta* **999**, 192–197 (1989).
72. T.M. Przybycien and J.E. Bailey, *Biochim. Biophys. Acta* **995**, 231–245 (1989).
73. T.M. Przybycien and J.E. Bailey, *Biochim. Biophys. Acta* **1076**, 103–111 (1991).
74. G. Zaccai and H. Eisenberg, *Trends Biol. Sci.* **15**, 333–337 (1990).
75. T.M. Przybycien and J.E. Bailey, *AIChE J.* **35**, 1779–1790 (1989).
76. V. Mikol, E. Hirsch, and R. Giege, *J. Mol. Biol.* **213**, 187–195 (1990).
77. W. Kadima, A. McPherson, M.F. Dunn, and F.A. Jurnak, *Biophys. J.* **57**, 125–132 (1990).
78. S.D. Durbin and G. Feher, *J. Mol. Biol.* **212**, 763–774 (1990).
79. S. Pascal and T. Cross, *J. Mol. Biol.* **241**, 431–439 (1994).
80. R. Ketchem, K. Lee, S. Huo, and T. Cross, *J. Biomol. NMR* **8**, 1–14 (1996).
81. J.F. Krebs, F. Rana, R.A. Dluhy, and C.A. Fierke, *Biochemistry* **32**, 4496–4505 (1993).
82. F.R. Rana, A.J. Mautone, and R.A. Dluhy, *Biochemistry* **32**, 3169–3177 (1993).
83. F.H. Rana, J.S. Mautone, A.J. Dluhy, *Biochemistry* **32**, 27–31 (1993).
84. K.R. Klein, B. McIntyre, J.O. Fleischer, S. Trommer, *Biochemistry* **35**, 3044–3049 (1996).
85. P.L. Privalov, *Adv. Protein Chem.* **33**, 167–241 (1979).
86. G. Wagner, *Q. Rev. Biophys.* **16**, 1–57 (1983).
87. M.J. Parker and A.R. Clarke, *Biochemistry* **36**, 5786–5794 (1997).
88. R. Jaenicke, *Naturwissenschaften* **75**, 604–610 (1988).
89. D.F. Mann, K. Shah, D. Stein, and G.A. Snead, *Biochim. Biophys. Acta* **788**, 17–22 (1984).
90. J.L. Cleland and A.J. Jones, *Pharm. Res.* **13**, 1464–1475 (1996).
91. M. McCloskey and M. Poo, *Int. Rev. Cytol.* **87**, 19–81 (1984).
92. G.D. Rose, A.R. Geselowitz, G.J. Lesser, R.H. Lee, and M.H. Zehfus, *Science* **229**, 834–838 (1985).
93. P.E. Wright, H.J. Dyson, and R.A. Lerner, *Biochemistry* **27**, 7167–7175 (1988).
94. N.C. Kaarsholm, S. Havelund, and P. Hougaard, *Arch. Biochem. Biophys.* **263**, 496–502 (1990).
95. E. Boix, M.V. Nogués, C.H. Schein, S.A. Benner, and C.M. Cuchillo, *J. Biol. Chem.* **269**, 2529–2534 (1994).
96. C.H. Schein and M. Haugg, *Biochem. J.* **307**, 123–127 (1995).

97. C.H. Schein, *In Vitro Toxicology* **10**, 275–285 (1997).
98. Y. Goto and A.L. Fink, *Biochemistry* **28**, 945–952 (1989).
99. R.D. Mitchell, H.K.B. Simmerman, and L.R. Jones, *J. Biol. Chem.* **263**, 1376–1381 (1988).
100. M.M. Ries-Kautt and A.F. Ducruix, *J. Biol. Chem.* **264**, 745–748 (1989).
101. J.E. Kinsella, *CRC Crit. Rev. Food Sci. Nutr.* **21**, 197–262 (1984).
102. H.M. Farrell, T.F. Kumosinski, P. Pulaski and M.P. Thompson, *Arch. Biochem. Biophys.* **265**, 146–158 (1988).
103. R. Boistelle, *J. Cryst. Growth* **123**, 109–120 (1992).
104. R. Judge, M. Johns, and E. White, *J. Chem. Eng. Data* **41**, 422–424 (1996).
105. Eur. Pat. Appl. 61,250 (1982), C. Weissmann and C.H. Schein (to Biogen N.V.).
106. M. Guo and G. Narsimhan, *Sep. Sci. Technol.* **31**, 1777–1804 (1995).
107. R. Guo, M. Guo, and G. Narsimhan, *Ind. Eng. Chem. Res.* **35**, 3015–3026 (1996).
108. K. Randall, M. Lever, B.A. Peddie, and S.T. Chambers, *Biochim. Biophys. Acta* **1291**, 189–194 (1996).
109. C.S.V. Rajendrakumar, T. Suryanarayana, and A.R. Reddy, *FEBS Lett.* **410**, 201–205 (1997).
110. Y. Liu and D.W. Bolen, *Biochemistry* **34**, 12884–12891 (1995).
111. A.J. Wang and D.W. Bolen, *Biophys. J.* **71**, 2117–2122 (1996).
112. F. Demmel, W. Doster, W. Petry, and A. Schulte, *Eur. Biophys. J* **26**, 327–335 (1997).
113. T. Kleinert, W. Doster, H. Leyser, W. Petry, V. Schwartz, and M. Settles, *Biochemistry* **37**, 717–733 (1998).
114. V. Daggett and M. Levitt, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5142–5146 (1992).
115. M. Hao, M. Pincus, S. Rackovsky, and H. Scheraga, *Biochemistry* **32**, 9614–9631 (1993).
116. P. Hunenberger, A. Mark, and W. van Gunsteren, *Proteins: Struct., Funct., Genet.* **21**, 169–213 (1995).
117. J.K. Myers, C.N. Pace, and J.M. Scholtz, *Protein Sci.* **4**, 2138–2148 (1995).
118. R. Fraczkiewicz and W. Braun, *J. Comp. Chem.* **19**, 319 (1998).
119. D. Eisenberg and A.D. McLachlan, *Nature* **316**, 199–203 (1986).
120. T. Ooi, M. Oobatake, G. Némethy, and H.A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 3084–3090 (1987).
121. J. Vila, R.L. Williams, M. Vasquez, and H.A. Scheraga, *Proteins* **10**, 199–218 (1991).
122. L. Wesson and D. Eisenberg, *Protein Sci.* **1**, 227–235 (1992).
123. G.I. Makhatadze and P.L. Privalov, *J. Mol. Biol.* **232**, 639–659 (1993).
124. M. Delarue and P. Koehl, *J. Mol. Biol.* **249**, 675–690 (1995).
125. J.S. Evans, S.I. Chan, and W.A. Goddard III, *Protein Sci.* **4**, 2019–2031 (1995).
126. M.D. Cummings, T.N. Hart, and R.J. Read, *Protein Sci.* **4**, 2087–2099 (1995).
127. A.J. Wang and D.W. Bolen, *Biochemistry* **36**, 9101–9108 (1997).
128. J.L. Fauchère and V. Pliska, *Eur. J. Med. Chem. Chim. Ther.* **18**, 369–375 (1983).
129. B. von Freyberg, T.J. Richmond, and W. Braun, *J. Mol. Biol.* **233**, 275–292 (1993).
130. C. Mumenthaler and W. Braun, *J. Mol. Model.* **1**, 1–10 (1995).
131. B. von Freyberg and W. Braun, *J. Comput. Chem.* **12**, 1065–1076 (1991).
132. A. Radzicka and R. Wolfenden, *Biochemistry* **27**, 1664–1670 (1988).
133. G. Barone, G.G. Della, V.P. Del, C. Giancola, and G. Graziano, *Biophys. Chem.* **51**, 193–199 (1994).
134. G.R. Hedwig and H. Hoiland, *Biophys. Chem.* **49**, 175–181 (1994).
135. J.K. Myers, C.N. Pace and J.M. Scholtz, *Protein Sci.* **4**, 2138–2148 (1995).
136. I. Karle, J. Flippen-Anderson, K. Uma, and P. Balaram, *Proteins: Struct., Funct., Genet.* **7**, 62–73 (1990).
137. I. Karle, J. Flippen-Anderson, K. Uma, and P. Balaram, *Biopolymers* **33**, 827–837 (1993).
138. C.M. Rosell, A.M. Vaidya, and P.J. Halling, *Biochim. Biophys. Acta* **1252**, 158–164 (1995).
139. N. Thanki, J.M. Thornton, and JM. Goodfellow, *J. Mol. Biol.* **202**, 637–657 (1988).
140. P.R. Gibbs, A. Radzicka, and R. Wolfenden, *J. Am. Chem. Soc.* **113**, 4714–4715 (1991).
141. A. Ferscht, *Enzyme Structure and Mechanism*, W.H. Freeman, New York, 1984.
142. J. Bryngelson, J. Onuchic, N. Socci, and P. Wolynes, *Proteins: Struct., Func., Genet.* **21**, 167–195 (1995).
143. A. Ansari, C.M. Jones, E.R. Henry, J. Hofrichter, and W.A. Eaton, *Biochemistry* **33**, 5128–5145 (1994).
144. C.H. Schein, *Curr. Opin. Biotechnol.* **4**, 456–461 (1993).
145. M. Bennett, M. Schlunegger, and D. Eisenberg, *Protein Sci.* **4**, 2455–2468 (1995).
146. B.S. Chapman and I.O. Kuntz, *Protein Sci.* **4**, 1696–1707 (1995).
147. U.H.E. Hansmann and Y. Okamoto, *Preprint SC-94-20*, Konrad Zuse Zentrum, Berlin, available via anonymous-ftp: *ftpserv01.zib-berlin.deSC94-20cd / pub get README* (1994).
148. M.C. Peitsch, *Biochem. Soc. Trans.* **24**, 274–279 (1996).
149. M.C. Peitsch, M.R. Wilkins, L. Tonella, J.C. Sanchez, R.D. Appel, and D.F. Hochstrasser, *Electrophoresis* **18**, 498–501 (1997).
150. B. Rost, *Methods Enzymol.* **266**, 525–539 (1996).
151. B. Rost, in C. Rawlings ed., *Third International Conference on Intelligent Systems for Molecular Biology* AAAI Press, Menlo Park, Calif., 1995, pp. 314–321.
152. L. Li and K.S. Matthews, *J. Biol. Chem.* **270**, 10640–10649 (1995).
153. E. Huang, S. Subbiah, J. Tsai, and M. Levitt, *J. Mol. Biol.* **257**, 716–725 (1996).
154. M.W. West and M.H. Hecht, *Protein Sci.* **4**, 2032–2039 (1995).
155. S. Roy, K. Helmer, and M.H. Hecht, *Fold. Design* **2**, 89–92 (1997).
156. S. Dalal, S. Balasubramanian, and L. Regan, *Nature Struct. Biol.* **4**, 548 (1997).
157. S.B. Prusiner, G. Telling, F.E. Cohen, and S.J. Dearmond, *Semin. Virol.* **7**, 159–173 (1996).
158. J. Safar, P.P. Roller, D.C. Gajdusek, and C.J. Gibbs Jr., *Protein Sci.* **2**, 2206–2216 (1993).
159. B. Dahiyat, C. Sarisky, and S. Mayo, *J. Mol. Biol.* **273**, 789–796 (1997).
160. J.L. Pomerantz, S.A. Wolfe, and C.O. Pabo, *Biochemistry* **37**, 965–970 (1998).
161. J.S. Kim and C.O. Pabo, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2812–2817 (1998).

PROTEIN EXPRESSION, SOLUBLE

CATHERINE H. SCHEIN
University of Texas Medical Branch
Galveston, Texas

KEY WORDS

Bacillus subtilis
Escherichia coli
High-density cultivation
Proteolysis
Secretion
Temperature effect

OUTLINE

INTRODUCTION

My first publications on techniques to avoid the formation of inclusion bodies in *Escherichia coli* (1,2) were met with many positive responses. Many others had experienced the problems summarized in the second column of Table 1 and were tired of dealing with "solutions of renatured proteins" that were in reality particulate microdispersions of inactive protein. They also shared my concern that proteins treated with strong denaturants were not the best starting point for biophysical studies, especially in cases where the specific activity had not been determined previously. Techniques such as protein NMR revealed structural heterogeneity in proteins purified from inclusion bodies (IBs), and few proteins purified from IBs have been useful therapeutically.

However, a recent review of the literature would indicate that renaturation from IBs, probably because it was the earliest-described method for producing large quantities of protein in *E. coli*, is still the predominant method. Generally, if isolation from IBs does not work, usually for the reasons in the second column of Table 1, the next step is to express the protein in yeast, insect cells, or transfected mammalian cells. Although expression in yeast is a fine alternative (see Refs. 3 and 4 for a summary of methods), both of the latter methods are significantly more expensive than bacterial cell culture. Insect cell cultivation is an excellent method for production of larger proteins, but stably transformed lines are given to variability, tests for expression are time-consuming, and the scale-up is more difficult than with bacteria or yeast.

Isolation from inclusion bodies is frequently used to purify proteins for characterizing the results of site-directed mutagenesis, precisely the area where accurate determination of specific activity and physical properties is most needed. University groups, who often appear to be racing with competitors to study the greatest number of mutants in the shortest time, frequently resort to IB purification schemes. Having little access to bioreactors and fermentors, they rely on shaker flask cultivation. The medium during growth in shaker flasks rapidly becomes anaerobic and acidic, meaning that the bacteria are struggling to grow and not able to produce optimal levels of product. Under these conditions, significant production requires completely subverting the bacterial metabolism into recombinant protein production, which usually means the use of induction methods such as temperature shock that favor IB production.

Obtaining an apparently pure protein in solution after renaturing from IBs is no guarantee of activity (5). For example, human tyrosine kinase Lck was completely inactive if expressed insolubly in *E. coli* and renatured from inclusion bodies. The protein was active, however, as measured by autophosphorylation, if solubly expressed by coproduction with either gro ESL or TRX (6).

Further, a mutant of a protein that renatures well from IBs may differ considerably from the wild type in solubility as well as specific activity. Ribonuclease A is probably the most common model for protein refolding experiments. However, a graduate student with experience in expressing RNase A mutants in *E. coli* and isolating them from IBs was unable to obtain any of a mutant that represented the ancestral protein of the hooved animals (the gene contains 16 amino acid changes chosen by evolutionary considerations). Undergraduate students in a lab course recloned the gene into an *E. coli* secretion system (7), expressed the mutant protein in a soluble form, and were able to determine its specific activity with respect to that of the modern protein (author's unpublished results).

As illustrated by the this example, expressing recombinant proteins solubly does not necessarily require great expertise. However, unlike expression in IBs, which can be done under relatively uncontrolled conditions, high-level expression of proteins in a soluble form requires control over many variables during growth and induction. These include temperature, medium pH, oxygenation, and concentrations of essential nutrients such as glucose (or glyc-

Table 1. Perceived Advantages and Actual Disadvantages of Expressing Recombinant Proteins as Inclusion Bodies in *Escherichia coli*

Advantages of IBs	Disadvantages
Many mutants can be tested because a single flask (or multiplate well) per mutant suffices for comparative studies.	Mutants of proteins can differ greatly in their specific activity and ability to refold.
High expression level means reduced fermentation costs.	Refolding shifts problems and costs downstream.
Production can be monitored with PAGE or immunoblotting.	Production cannot be monitored directly by activity.
Cytoplasmic proteins are washed away, simplifying purification. 95% purity can be quickly achieved.	The other 5% contaminants are hydrophobic, insoluble membrane proteins and cell wall fragments
Removal of cytoplasmic proteins early means less proteolysis.	Significant proteolysis accompanies refolding in dilute solution
The major contamininants are oligomers and misfolded or proteolyzed forms of the desired protein.	Separation of multiple forms of the same protein is the most difficult purification step.
The pL promotor induced by raising temperature often yields where other systems fail.	If the protein does not refold well, another expression system will still be needed.

erol) and toxic by-products of the bacteria, including acetate and carbon dioxide. Several industrial biotechnology groups have developed elegant solutions for all the problems encountered in soluble expression (8,9). Presumably because of stability problems with respect to both the host strain and plasmid, continuous culture methods are generally not used. The highest-yielding processes to date are based on fed-batch technology, which is also the standard for antibiotic production.

Most researchers wish to obtain their protein in a soluble form with a minimum amount of work. I thus begin by suggesting alterations in growth conditions that can be tested with most expression systems. The second section then discusses alternative host strains and plasmids that are better for directing soluble protein expression. The third section is devoted to discussing ways of altering the protein itself to avoid IB formation.

Today, bacterial culture and cloning techniques are used by many whose specialties are far removed from the area of microbiology. For this reason, the fourth section contains detailed suggestions for preparing and maintaining stock cultures and inocula, as well as a phage-free environment. One must stress the importance of good sterile technique, a topic that is never presented in other areas of the physical sciences. The last section includes tips on maintaining solubility during harvesting and the preliminary washing/lysis steps. Cosolvents for retaining solubility of proteins in solution are discussed elsewhere (10).

ALTERING GROWTH CONDITIONS TO INCREASE SOLUBLE PROTEIN EXPRESSION

Lowering Growth Temperature

For intracellular production, simply lowering the growth temperature (*T*) (1,11) can greatly increase the soluble fraction of protein. Table 2 lists many recombinant proteins from many different sources that are more soluble when expressed in *E. coli* at a lower growth temperature. The list includes several fusion proteins, for example, mouse DNA polymerase delta expressed as a glutathione-*S*-transferase conjugate (19). Although most of the examples in Table 2 used 25–30 °C for production, there is evidence that one can lower the temperature much further to increase the yield of soluble protein. The vaccinia virus large subunit of ribonucleotide reductase (vvR1, 87 kDa) was completely insoluble in *E. coli* when expressed at 37 °C, but mostly soluble if induced at 15 °C (29). The quantity of soluble and insoluble human androgen receptor (AR) fusion protein produced at different times following induction at 12, 22, and 37 °C indicated that while much more insoluble protein was expressed at higher temperature, soluble receptor expression was much higher at the lower temperatures. The amount of soluble protein did not correlate well with the ligand-binding capability of the receptor, indicating there was some bacterial factor needed for folding to a binding-competent state whose capacity was exhausted at higher expression levels (there is little evidence of proteolytic degradation of the cytosolic fraction from their gels) (15). Six different areas of the protein were converted to insoluble protein at the higher temperatures.

There is a lower limit of about 8–10 °C below which *E. coli* will not grow at all. In general, K12 strains grow very slowly below 23 °C; at 10 °C only a few dozen proteins are produced (42).

Point mutations may make a protein more or less prone to temperature-induced aggregation. For example, human triose phosphate isomerase is soluble at 37 °C during production in *E. coli*, as are two mutants, Met14Gln and Arg98Gln, which are partially and completely inactive, respectively. The (inactive) protein containing both mutations, however, was produced in a soluble form only at a growth temperature of 28 °C (43).

Although in some cases the production of soluble protein may appear to be the same at different growth temperatures (44), lowering the growth temperature has two additional great advantages. The first is that at lower temperature the O_2 solubility in the growth medium is much higher. Thus the cell density can be increased before the culture grows anaerobically. Anaerobic growth, and particularly the production of acetate, has been shown to limit cell growth and the production of protein (45).

The second advantage is that several plasmid expression systems have been shown to be more stable when bacteria are grown at lower temperatures. This may be due to the fact that the half-life of the antibiotics used to select plasmid-bearing bacteria during the cultivation also have a longer half-life at the lower temperature.

Table 2. A Few Proteins That Are More Soluble When Produced in *Escherichia coli* Grown at 30 °C or Less

Protein	Ref.
Animal / mammalian	
Bovine heart mitochondrial ubiquinol–cytochrome *c* reductase (fusion protein)	12
Chick brain β-tubulins, 33 kDa	
Chicken riboflavin carrier protein	13
Cystic fibrosis transmembrane conductance regulator (fused to glutathione-*S*-transferase)	14
Human androgen receptor (AR) fusion protein	15
Human initiation factor-4E	16
Human Interferon-α2	
Human Interferon-β	17
Human Interferon-γ	
Human transforming growth factor-β	18
Human triose phosphate isomerase double mutant (M14Q/R98Q)	
Mouse DNA polymerase delta	19
Murine cAMP-dependent protein kinase (catalytic subunit)	20
Murine Mx	
Rabbit muscle glycogen phosphorylase	21
Rat phosphatidylinositol transfer protein	22
Regulator of chromosome condensation (RCC1)	23
Xenopus Cu, Zn superoxide dismutase	24
Plant	
Radish 3-*OH*-3-methyl glutarylcoenzyme A reductase	25
Rice lipoxygenase L-2	26
Yeast α-glucosidase P1	27
Viral	
Herpes simplex thymidine kinase fusion protein	28
Vaccinia virus large subunit of ribonucleotide reductase (vvR1, 87-kDa)	29
Yeast / fungal	
Cephalosporin-C acylase from Pseudomonas N176	30
Isopenicillin-*N*-synthases	31
Bacterial	
Bacteroides fragilis metallo-β-lactamase (truncated protein)	32
Diphtheria toxin and fusion proteins thereof	33
Heparinase II and III from *Flavobacterium heparium*	34
Lycopene cyclase from *Erwinia uredovora*	35
P22 tailspike	36
Ricin A-chain	37
Salmonella typhimurium CheY mutant	38
Salmonella typhimurium tryptophan synthetase C-terminal deletions	39
Shigella flexneri invasion plasmid antigens his-fusion	40
T4 DNA polymerase	41

Note: Where no reference is cited, the data are from Ref. 11. Proteins are not listed that are produced in a soluble form at 30 °C but for which no comparative data is provided in the reference concerning solubility at a higher growth temperature.

Higher growth temperature can also inhibit processing. When chicken riboflavin carrier protein, a phosphoglycoprotein, was expressed in the T7 system, at 37 °C, two products were seen in whole cells lysates for protein with and without the pelB leader sequence encoded in the pET20b vector. Protein solubilized primarily from IBs did not contain the signal sequence, which was also absent in soluble protein expressed at 22 °C (13).

Possible Reasons for the Effect of Temperature. The reasons for the effect of lowering growth temperature are hard to pinpoint. The temperature effect is not simply due to slowing the production rate of the protein or to a decreased growth rate. The rate of synthesis of a particular protein does not necessarily increase with temperature. The protein Mx was sequestered into IBs at 37 °C, regardless of the absolute production rate (which was controlled by varying the distance between the Shine–Dalgarno sequence and the initial ATG in otherwise similar expression plasmids) (11). Further, decreasing the growth rate in other ways does not increase soluble expression; to the contrary, substrate limitations lower recombinant protein expression (9).

It is possible that compartmentalization of protein production is more efficient at the lower growth temperature. For example, RNase A was insoluble at 37 °C when produced cytoplasmically, but soluble at this temperature when directed to the periplasm/medium in a secretion system (7). In mammalian cells, proteins expressed in the nucleus were more susceptible to thermal denaturation in situ compared with those expressed in the cytoplasm (46).

Another possibility is the presence or lack of some proteins important for folding or aggregation at the different temperatures. Phosphorylation may occur at the higher temperature, thus altering the characteristics of the protein, as has been shown for Gro EL (8). Proteolysis increases at higher temperatures, as does the catabolism of amino acids involved in central metabolic pathways, such as serine (47). Studies of the expression of various heat shock proteins, which are induced by growth at higher temperatures, have indicated that their presence alone cannot mimic the effects on cell growth of increasing temperature. Indeed, as we will see in the section "Altering the Host Strain and Vector to Direct Soluble Expression", coexpression of some of these proteins can selectively increase the amounts of soluble recombinant protein. A possible explanation for the effect of primary sequence on IB formation is discussed in the section "Altering the Protein for Solubility."

Medium Composition

Use of a defined medium (i.e., one in which poorly defined nitrogen sources such as yeast extract or casein hydrolysate are not used) is preferable. Several defined media have been developed to allow high-density bacterial growth (48). Glycerol may be a better carbon source than glucose because less acetate is produced (45). In general, growth limitations should be avoided, with the exception of expression with negatively regulated induction systems, such as those based on the tryptophan promotor. Batch

feeding during induction has been shown to increase protein yield considerably. Mathematical models have also been developed to aid in increasing cell density (49).

Before eliminating casein hydrolysates from the growth medium, as one does for example in producing labeled proteins for NMR experiments, one should insure that their presence is not essential for controlling protease activity. Addition of protease inhibitors to the growth medium, in the form of an ill-defined extract of potato, was essential for production of interferon in *Bacillus subtilis* (50). Specific inhibitors of toxic enzymes may also be added to the growth medium. Yields of vvR1, for example, were improved by adding the ribonucleotide reductase inhibitor hydroxyurea (29).

Adding compatible cosolvents during the culture may also help maintain solubility, whereas in some cases their positive effect may actually be during harvesting (see section "Controlling Other Variables That Affect Solubility"). Increasing medium osmolarity by the addition of nonmetabolizable sugars may also increase protein solubility at higher temperatures: 0.4 M sucrose in the growth medium increased the production of soluble β-lactamase protein (51). There seem to be relatively few examples, however, where compatible cosolvents have been tried, perhaps because the levels of the compounds must be optimized and their use can greatly increase medium costs. The solubility of bovine heart mitochondrial ubiquinol–cytochrome *c* reductase, expressed as a fusion protein with glutathione, was improved by adding 2.5 mM betaine and 440 mM sorbitol to the growth medium. Presumably, the increased osmolarity of the growth medium forced the uptake of betaine and a corresponding increase of glycyl-betaine within the cells (12).

Good aeration, even beyond the O_2 uptake requirements of the culture, can increase production of proteins in *E. coli*. Rapid aeration reduces the concentration of CO_2 and acetate, both of which can be toxic. High acetate production has been implicated in reducing yields, especially in shaker flasks, even when the pH is controlled by using highly buffered media. Strains that produce less acetate have been created by metabolic pathway engineering; however, these may grow too slowly to be useful. Controlling the growth conditions of a prototrophic strain with few mutational markers so that it did not experience anerobiosis or stress has been shown to be a more useful way to reduce acetate production in large-scale, high-density culture (9).

Proteins induced under stress conditions may serve to control the production of acetate. The universal stress protein (USPA), a 13.2-kDa protein, is induced by a variety of stress conditions. This protein's role during stress may be to control the pathways of glycolysis, because mutants in the USPA gene secrete large amounts of acetate (52).

Other methods to increase the oxygen availability may be preferable to heavy aeration, which can lower the yields of secreted proteins by increasing the shear force (increased agitation) and denaturation at the air–liquid interface of bubbles. For example, surface aeration with supplemental O_2 was used to produce interferon-α in *Bacillus subtilis* because large activity losses were found when air or even plain nitrogen was bubbled through a solution of medium and the purified protein (53). The antifoam agent needed to control foaming at high volumetric airflow rates (VVM) can also seriously lower yields, especially for secreted proteins, and should be used sparingly.

Induction Conditions

The amount and time of inducer addition should be strictly controlled. Many have found it advantageous to induce at low cell density in shaker flask culture to insure that the cells are growing actively. However, in a bioreactor with efficient batch feeding and good aeration, much higher densities can be used. A higher cell density may be preferable with induction at very low temperatures, as for example when using a cold shock promotor, because the cells will grow quite slowly after that point. Most industrial processes use very high cell density; for example, production of the fusion protein of fibroblast growth factor–Saponin (a mitotoxin designed to treat cancer and restenosis) relied on fed-batch cultivation with a rich medium feed to reach A_{600} = 85 before inducing with isopropyl-D-thiogalactopyranoside (IPTG) at 30 °C (pET11a vector/BL21 [DE3] host). The yield of soluble mitotoxin, which could be isolated in good yield using heparin affinity chromatography, was 2.2 mg/L/A_{600} (54). The concentration of inducer may also be important; several authors have reported better yields of soluble protein if the IPTG inducer concentration was reduced from the usual 1 mM (22,23,55). Yields of vvR1, for example, were improved by inducing at a low cell density and using a low concentration (0.05 mM) of IPTG (29).

ALTERING THE HOST STRAIN AND VECTOR TO DIRECT SOLUBLE EXPRESSION

Choice of Host Strain

Expression versus Transformation. Correct host strain selection is also essential. Table 3 lists the markers carried by a few commonly used *E. coli* strains. Although some of their markers may determine how well a strain produces protein, several of them are potentially dangerous leftovers selected for the use for which the strain was originally designed. For example, *supE* is an amber suppressor that inserts a glutamine at UAG codons; it was added to strains of *E. coli* to allow the replication of early M13 vectors. These vectors were required to contain amber mutations that would control their propagation on accidental release, according to NIH guidelines that no longer apply (62). For this reason, older bacterial strains with fewer markers may be preferable for production purposes (9).

Careful cultivation and preservation of host strains is essential for optimal production. Because host strains for protein production often do not transform well, strains such as JM101 or HB101 are used for transformations after ligation or for vector propagation. Choose a host strain that is suitable for the problem at hand.

Low Protease Strains and Those with Better Redox Characteristics. Use of protease minus host strains (63) is often an advantage. *E. coli* BL21 (DE3), a protease-deficient strain, was used to express bacterial glutamyl tRNA re-

Table 3. Several Common *Escherichia coli* Cloning or Host Strains and Their Genotypes

Strain	Genotype	Ref.
C600	*supE44, thi-1, hsdR, thr-1, leuB6, lacY1, tonA21, mcrA*	56
HB101	*supE44, hsdS20*(r-Bm-B)*recA13ara-14proA2lacY1gal/K2rpsL20xyl-5mtl-1*	57
JM101	F′, *traD36, proAB, lacZ, Δ*M15/*Δ*(*lac,pro*), *supE44, thi-1*	58
JM109(DE3)	[F′, *traD36, proAB, lacI*^{q}, *lacZ, Δ*M15] *λ*(DE3), *Δ*(*lac,pro*), *endA1,hsdR17*(Rk^-Mk^+)*supE44, thi-1, recA1, gyrA96,relA1*	45
RV308	*lac74-galISII*::OP308-*strA*	59
XL1-Blue	*endA1, hsdR17*(Rk^-Mk^+)*supE44, thi-1, λ-, recA1, gyrA96, relA1, lac*[F′, *proAB,lacI*^{q}*Δ*M15,TN10(*tet*^{R})]	60

Note: See tables by Sambrook et al. (61) for more strains and an explanation of the abbreviations.

ductase. The synthesis of 5-aminolevulinic acid (ALA) and porphyrins, which were secreted into the medium, could be used to follow the production of the protein, which was soluble when produced at 37 °C (64).

Protease-minus strains of *E. coli* are particularly valuable for production of soluble proteins and those expressed from secretion vectors. A *lon*$^-$ *htpr*$^-$ strain (lacking the lon protease of *E. coli* [65]) significantly increased the production of RNases in a secretion system (7), and the strain SG21173 (*lon*$^-$, *htpR*$^-$, *clp*) was used to express the human T-cell receptor V beta 5.3 under the control of the hybrid trc promoter containing the *E. coli* heat-stable enterotoxin II (STII) signal sequence for protein secretion (66). There are also strains that facilitate the formation of disulfide bonds in the reducing environment of the cytoplasm (67). These should be considered for expression of proteins with multiple cysteine disulfides.

Coexpressing Other Proteins. Cocloning of molecular chaperones, which can include such diverse molecules as thioredoxin reductase (68) and, for membrane proteins, phospholipids (69) can be of help, but they must be selectively chosen (Table 4). Coexpression of the *E. coli* heat shock protein GroESL has solubilizing effects only on certain mammalian proteins. For example, coexpression of GroESL increased the solubility of four of eight mammalian proteins (c-myb, CRE-BP1, p53, and the *Xenopus* protooncogene product Mos), while coexpression of thioredoxin increased the solubility of these and the other four (SnoN, c-Myc, adenovirus E1A, and the human tyrosine kinase Lck) (6). Thioredoxin also increased the production of antibodies (68).

There is also a report that association with GroEL actually decreased the solubility of the cystic fibrosis transmembrane conductance regulator (expressed as a glutathione-*S*-transferase fusion protein). A 60-kDa protein coeluted from affinity columns with the protein and was shown by sequencing to be GroEL. The authors were able to break up the complex and purify the protein (which was induced at 25–28 °C) by adding 5 mM MgATP to the bacterial extract and incubating at 37 °C (14).

Table 4. Useful Expression Partners for Recombinant Proteins

DNA-k and DNA-J
E. coli GroESL
Phospholipids
Thioredoxin
Antibodies to the protein
Vitreoscilla hemoglobin

Other reports indicate that GroESL may have a solubilizing effect only at certain temperatures. The aggregation of preS2-*S′*-*β*-galactosidase, a three-domain fusion protein in the cytoplasm of *E. coli*, was reduced three- to sixfold by cooverexpression of the DnaK and DnaJ molecular chaperones over a wide range of growth temperatures (30–42 °C), while an increase in the concentration of the GroEL and GroES chaperonins had a significant effect only at 30 °C. Coimmunoprecipitation experiments confirmed that preS2-*S′*-*β*-galactosidase formed a stable complex with DnaK, but not with GroEL, at 42 °C. When the intracellular concentration of chromosomal heat shock proteins was increased by overproduction of the heat shock transcription factor *σ* (33), or by addition of 3% ethanol (v/v) to the growth medium, a two- to threefold higher recovery of active enzyme was observed at 30 and 42 °C, but not at 37 °C. The overexpression of all heat shock proteins or specific chaperone operons did not significantly affect the synthesis rates or stability of preS2-*S′*-*β*-galactosidase and did not lead to the disaggregation of preformed inclusion bodies. Rather, the improvements in the recovery of soluble and active fusion protein resulted primarily from facilitated folding and assembly. The authors suggested that titration of the DnaK–DnaJ early folding factors may trigger the formation of preS2-*S′*-*β*-galactosidase inclusion bodies (70).

Coexpression of proteins that improve solubility or the overall growth of the bacteria can also be useful. For example, coexpression of *Vitreoscilla* hemoglobin can increase the total cellular protein of various common *E. coli* cloning strains (JM101, pRed2 and W3110) and *B. subtilis* (71). Their production characteristics should also be improved because onset of oxygen limitation is delayed (72). Cocloning demonstrated that the function of at least one of the six homologous insulin-like growth factor binding proteins is to assist in the formation and maintenance of the native disulfides of IGF-I (73).

Cocloning of specific antibodies to the protein may increase yields. For example, the heat-induced aggregation of carboxypeptidase A could be greatly reduced in vitro in the presence of monoclonal antibodies that did not inhibit its activity (74). In addition, coexpression of tRNA genes for the preferred mammalian codons may increase the expression of mammalian genes in *E. coli*. For example, the

expression of argU tRNA enhances expression of interferon (coded from the mammalian sequence) in *E. coli* (75).

Alternative Hosts to *E. coli*. Despite the millions of different bacteria in nature, very few are used currently as hosts for protein production. *E. coli* itself, isolated originally from infant diarrhea, was chosen as an experimental strain by Jacques Monod soon after World War II because it was relatively nontoxic and grew quickly. Other bacteria that have been used as hosts for recombinant protein production were selected because they were known for medical reasons (*Straphylococci*, *Bacillus calmette-guerin*), were approved for use in some food-related applications (*Lactobacillus* [76]), or were used industrially for enzyme or antibiotic production (*Bacillus subtilis* and *Bacillus amyloliquefaciens*).

B. subtilis systems, based on previous industrial processes for enzyme production, are mostly suitable for the production of stable proteins that are resistant to proteases (77). If the basic problems of protease degradation and processing problems that may lead to insoluble product (53,78) can be avoided, there are several good expression vectors, including one based on the α-amylase signal sequence and promotor (50). *Bacillus licheniformis*, useful for producing bacterial enzymes on the industrial scale (79), has been used to produce active human IL-3 of therapeutic quality (80). *B. calmette-guerin* has been used to produce a *Schistosoma mansoni* antigen (81). *Lactobacillus lactis*, found in cheese cultures, is also a useful expression host (82) and may have promise for vaccines (83).

Recently, *Staphylococcus carnosus* was used for surface display of a malaria antigen (84), and a secretion/expression system in *Salmonella* has been suggested for vaccine production (85). There are "shuttle vectors" available that can function in both *E. coli* and various alternative host bacteria. For example, expression of *Trympanosoma cruzi* polyantigens using the vector pRIT21, which contains the origins of replication for both *E. coli* and *Staphylococcus aureus*, allowed expression in both strains. Seven different antigens showed less proteolysis when secreted from *S. aureus*; with purification yields of 80%, between 4 and 16 mg/L of the different antigens for vaccine production were obtained (86).

Various *Vibrio* species were used to express the nontoxic B-subunit of *E. coli* heat-labile enterotoxin linked to a 27-residue C-terminal fragment of Pol, the catalytic subunit of DNA polymerase of herpes simplex virus type 1. The highest yields were reported to be 3.5 mg/L; 0.3 mM EDTA added to the growth medium blocked proteolysis of the secreted protein (87).

Altering the Vector

Choice of Promotor. Using a prokaryotic promoter that is tightly regulated and can be induced to express large amounts of protein at low temperatures (23–30 °C) is essential for soluble expression. For example, expression levels in the T7 system can be increased by tighter transcriptional control before induction (88). In view of the effect of temperature on solubility, it is interesting that a whole group of *E. coli* proteins is induced when the culture temperature is lowered to below 20 °C. The promotor for one of these genes, cold shock promotor CspA, is the basis of a vector developed specifically for low-temperature induction of recombinant proteins (89). The expression levels possible with this promotor at the optimal induction temperature compare favorably with those obtained with the tac promotor (90). The Csp promotor has the advantage that temperature shift alone induces production. It should certainly be considered for expression of highly aggregation-prone or unstable gene products.

Older promotors, such as those based on the tryptophan operon, are useful for low-temperature expression, as is a modified tryptophanase promotor (91). Secretion of specific proteins from *E. coli* is already possible using a variety of vectors (7,92). Secretion expression systems are particularly suited to mutant analysis (93) because the culture supernatant can be directly analyzed. Further study of the factors affecting transport of proteins is needed.

Selecting Signal Sequence or Fusion Partners. The sequence of the protein itself can be altered by adding linkers to improve protein folding, solubility, protease resistance, yield, and secretion into the culture medium (Table 5). Expression of small proteins or fragments of a larger protein may be possible only if the desired sequence is expressed as a fusion protein (98). The *E. coli* maltose binding protein (MalE) is particularly useful because fusion molecules with it, besides being easily purified, are often secreted. For example, CD4, the human T-lymphocyte receptor for the AIDS virus, HIV, as well as many different functional deletions that retained their ability to bind monoclonal antibodies and HIV gp160 and inhibit HIV binding to $CD4^+$ cells, were fused as multiple repeats to both ends of a single MalE molecule. The hybrid molecules (expressed in L-broth at 30 °C) were exported into the periplasm (95). Pollen major allergen Bet v1 from *Betula verrucosa* (White Birch) has been cloned and expressed in *E. coli* as a fusion with maltose-binding protein and a factor Xa proteolytic cleavage site. The cloning strategy, based on a polymerase chain reaction, positioned the factor Xa proteolytic site so that the authentic amino terminus of Bet v1 was generated after cleavage. Fusion protein (expressed solubly at 37 °C) was isolated by amylose affinity chromatography and enzymatically cleaved by incubation with factor Xa (99). The kringle V domain of the lipoprotein apolipoprotein(a) from human liver was solubly incubation with factor Xa (99). The kringle V domain of the lipoprotein apolipoprotein(a) from human liver was solubly expressed in *E. coli* as a MalE fusion protein. As with the other MalE fusions, this protein was purified by amylose–agarose affinity chromatography by eluting with 10 mM maltose (100).

Several linkers, such as hexahistidine, can aid in the detection and purification of recombinant proteins. Most vectors also include a protease site between the linker and the protein. The 54-kDa subunit of the canine signal recognition particle (SRP) was expressed *E. coli* using a T7 expression system with an amino-terminal extension encoding an initiating methionine and 10 histidine residues followed by an enterokinase cleavage site. His-SRP (54 kDa) was purified to give a single band on SDS-PAGE in 20% yield (0.3 mg from 500 mL culture) (101). The hexo-

Table 5. Linkers That Improve the Solubility of Fusion Proteins or Allow More Specific Purification

Linker	Comments
Calmodulin (94)	Calmodulin binds to phenol sepharose tightly.
E. coli maltose binding protein (MBP or MalE)	Purify fusion protein by binding to amylose beads and eluting with maltose. Secretion into the periplasm of soluble fusion protein is often possible even with 37 °C growth. Proteins and deletions thereof may be cloned at either or both ends (95).
Glutathione-*S*-transferase (GST)	Purify fusion protein by binding to glutathione beads. Fusions may be more soluble.
Hexa- (or more) polyhistidine linker	Allows purification with metal chelate columns.
Thioredoxin (TRX)	Very commonly used vector system. Antibodies to thioredoxin may be useful in purifying the fusion proteins (96).
Ubiquitin (97)	Easily cleavable fusion protein.

kinase of *Schistosoma mansoni* has been expressed in *E. coli* as a fusion protein including an N-terminal polyhistidine tag and enterokinase cleavage site. The enzyme was purified by metal chelate chromatography to > 95% homogeneity, based on analysis by SDS gel electrophoresis and isoelectric focusing (102). This sort of expression system has the advantage that the protein can often be expressed without an N-terminal methionine, but the disadvantage that yields are reduced due to the additional processing step. Even specific proteases may cleave the protein itself. Thus initially high expression levels may turn into disappointing bottom-line figures.

Thioredoxin fusions of many proteins are soluble, especially if produced at a lower growth temperature, active in some cases before removal of the thioredoxin portion, and may be purified by heating lysates to 80 °C. A-DDDDK-linker, the recognition site for mammalian intestinal protease enteropeptidase, allows removal of the thioredoxin moiety (103). IGF-1 was more than 100 times more soluble when fused to one or two Z-domains (from Staph protein A), and the solubilizing effect could also be seen during in vitro refolding (104).

There is a fine interplay between the temperature effect on solubility and the linker. For example, the solubility of the his fusions with the IpaC invasion plasmid antigens of *Shigella flexneri* (antibodies to which may be useful in treating Shigellosis dysentery) was improved by production at 30 °C. However, no protein was found when the IpaB antigen was similarly cloned. Both histidine-linked proteins were partially soluble at 37 °C if expressed as fusions with Trx (40).

Human osteocalcin (hOC), a 49–amino acid peptide produced mainly by bone osteoblasts, was fused to the glutathione S-transferase (GST) gene, and the fusion protein was produced mainly in a soluble form. The affinity-purified fusion protein was cleaved with activated protease factor X, releasing the rhOC portion. The structure of rhOC was confirmed by mass spectrometry and amino acid sequencing (105).

Changing Codon Use without Altering Sequence. Bacterial and mammalian codon usage differ significantly. Genes for several proteins have been synthesized to reflect the bacterially preferred codons for degenerate amino acids and to eliminate elements in the noncoding region of genes that lower productivity. For example, RNase A was expressed at a much higher level when synthesized from a synthetic gene that reflected codon usage than was its murine counterpart expressed from the natural mammalian gene (7). Alternatively, one can look for specific mammalian codon usage in the cDNA that may inhibit translation. For example, 4.9% of the adult isoform of human cardiac troponin (288 amino acids) is arginine coded by the (rarely used in *E. coli*) codons AGG and AGA. Changing two areas where these codons occurred consecutively (AGG [165] and AGG [166], and AGG [215] and AGG [216]) to the bacterially preferred codon CGT increased expression 10-fold (one pair); there was a 40-fold increase when both pairs of the rare arginine codons were replaced (106). In a similar example, electrospray ionization mass spectrometry of recombinant link module from human tumor necrosis factor–stimulated gene 6 protein revealed that much of the protein was a mixture of forms where between one and four of the arginine residues had a lysine substitution. When the AGA codons for the arginine residues were substituted to the bacterially preferred codon CGT, the substitution problem was eliminated, and the level of expression was greatly increased (107). One can also try using strains that express tRNAs for the mammalian preferred codons, as already discussed.

A combination of methods may be necessary. For example, the aggregation of the vaccinia virus large subunit of ribonucleotide reductase (vvR1, 87 kDa) (29) was prevented by lowering the induction temperature to 15 °C, the cell density during induction, and the concentration of the inducer (to 0.05 mM IPTG). Addition of hydroxyurea, an inhibitor of ribonucleotide reductase, further increased production of soluble vvR1 in a dose-dependent manner.

ALTERING THE PROTEIN FOR SOLUBILITY

Why Do IBs Form?

The reason for the formation of IBs has been the subject of much debate since the phenomenon was first observed. An answer to this question is vital to achieving maximum soluble expression in *E. coli*. Lowering growth temperature to achieve more soluble protein does not work for all proteins; further, we do not understand the physical characteristics that account for why some proteins are perfectly soluble even at 42 °C, but point mutations thereof can render the proteins insoluble at all temperatures. Size is not

a variable because even very large proteins (e.g., heparinases II and III, see Table 2, are over 700 amino acids long) can be expressed in soluble form.

Although even point mutations can render proteins more or less soluble (for reviews see Refs. 2 and 108), specific effects cannot be directly related to hydrophobicity (Gravy scores [31]) or content of a specific amino acid such as cysteine or proline. A multifactorial analysis (109) also did not yield a satisfactory explanation for the effect of sequence on solubility.

The only general observation is that there is a good deal of misfolded protein and proteolytic fragments in IBs. IB formation in *E. coli* was probably first observed when cells were treated with amino acid analogues (110); this relationship between abnormal protein and aggregation has been born out by other studies. In general, proteolysis accompanies IB formation (11); indeed several authors attributed the effects of a lower growth temperature on solubility to reduced proteolytic cleavage of their proteins (17,55). Certain intermediate states in protein aggregation are detectable by some physical techniques but difficult to define at the sequence level (111).

Two recent observations may shed some light on the process of aggregation. The first is that protein aggregates, as viewed by PAGE, appeared identical in the presence or absence of another protein. Thus the protein appeared to aggregate only with itself and not to coaggregate with a potential partner (112). Of course, the authors could not analyze the more insoluble, larger molecular weight aggregates that did not penetrate into the gel, but the results strongly suggest that aggregation follows a very specific pathway.

The second observation is that the oligomers seen in many crystal structures are not due to the monomers aligning themselves at surface-exposed hydrophobic sites. Rather, the monomers actually exchange a whole section of the protein to form a compact structure. Requirements for the swapped domain are that it have both a hydrophilic and hydrophobic face and that it be joined to the rest of the protein by a flexible linker. "Domain swapping" is one possible mechanism for aggregation that accounts for many of the observations about IB formation (113):

- *Specificity*. The swapped domain is connected to the oligomeric partner in much the same way as to the original monomer.
- *Speed*. Thermodynamically speaking, the major barrier to overcome during the initial formation of the dimer molecule is the barrier toward simultaneously opening two monomers that lie within a permissible distance of one another to permit establishing a bond. Once a dimer has formed, however, addition of another monomer requires only the opening up of one more molecule at a time. Thus aggregation will proceed quite quickly once the initial template exists.
- *Proteolysis*. IB formation is usually accompanied by proteolysis, and proteolytic fragments are often seen in solubilized preparations from IBs. If one assumes that a long, flexible linker joins the swapped domains, this essentially unstructured area would present a good site for the attack of proteases. The domains liberated by hydrolysis of the linker might serve as fragments to further participate in the aggregate formation.
- *Association with synthesis*. Pulse labeling studies with *E. coli* overexpressing the *Salmonella typhimurium* gene *CheY* have shown that newly formed protein aggregates more quickly than existing protein (38). However, preexisting protein can also become involved in IBs if incubation is continued for a sufficiently long time (11). This makes sense if one assumes that protein immediately after synthesis is not yet completely folded. Possibly the stable domains fold first, followed by the folding of the whole protein. A protein in solution exhibits a certain amount of folding-out (so-called breathing) that can be observed by techniques such as hydrogen exchange. Presumably, new protein whose secondary structure is formed but is still unfolded at the tertiary level can exchange domains easily, whereas existing protein can join the aggregate only at the appropriate stage in its breathing state. The movement increases with temperature, which can also account for the observed temperature dependence of aggregation. Aggregation would thus be one aspect of the protein solution half-life, which can be related to temperature by an Arrhenius plot ($\ln k$ vs. $1/T$, where k is an intrinsic instability constant and T is the temperature in °K) (114).
- *Irreversibility*. Because the area of the domain that is swapped is substantial, the energy required to form (and hence to separate) oligomers is very large. The domain-swapping model can also explain to some extent the influence of heat shock proteins and chaperonins on solubility.
- *Other Miscellaneous Observations*. It has also been noted that there is active protein in IBs (115), which suggests strongly that the protein is not completely misfolded. Some IBs have the ability to be separated by ion exchange. Also there is a lack of a general relationship between protein sequence and aggregation. Domain swapping accounts quite neatly for the effects of specific point mutations (116) and their corresponding supressors on the solubility of proteins (117).

If one accepts that domain swapping is involved in oligomer formation, designing peptides to prevent aggregation may be a greatly simplified process, especially if a crystal structure is available or one has knowledge of interacting faces withing the protein. One can direct changes in the protein itself to inhibit its aggregation, or design peptides that will bind reversibly to these sites and prevent the initiation of aggregation (106).

These peptides need not be very large; for example, a simple tetramer, Arg-Gly-Asp-Ser, from fibrinogen was sufficient to inhibit the aggregation of platelets (118). A few discrete amino acid changes at the end of the B-chain of insulin considerably reduced the oligomerization state (119).

Two examples show that this mechanism can be exploited if one can detect amino acid regions in turns. In the first, a monomeric λ-cro protein was made from the homodimer by duplicating the dimer interface region (six amino acids) at the C-terminus of the molecule. The monomeric mutant was more thermostable and appeared from NMR studies to have a solution structure similar to that of the dimer (120). *E. coli* maltose-binding protein is soluble, but a double point mutant (Gly32Asp/Ile33Pro) is insoluble whether it is targeted to the cytoplasm or periplasm by expression without or with its signal sequence, respectively. These mutations are located in a turn/loop region between two helical regions in the crystal structure (121).

Mutation to Increase Solubility

Single point mutations may greatly improve solubility, whereby the observed increase in soluble protein may also be a secondary effect of increased stability to proteolysis or improved interaction with a chaperone (122). Single point mutations may alter the effect of temperature on solubility. The solubility and stability to proteolysis during production in *E. coli* of a 101–amino acid fragment from the human respiratory syncytial virus (RSV) major glycoprotein (G-protein) was significantly improved by altering several phenylalanine residues. Two cysteine-to-serine changes, did not affect the solubility or stability of the gene product. The mutant G-proteins appeared antigenically similar to the parent and had similar secondary structures (123).

Destruction due to proteases can be minimized by identifying degradation "hot spots" in proteins (124). Some of these sites have been discussed elsewhere (2,108). It should be noted that such sites in mammalian proteins may be one mechanism for controlling their activity. For example, proteolytic cleavage at the C-terminal of human interferon-γ, which also occurs in blood, can increase activity several fold. Efficient production of this cytokine in *E. coli* was possible only if the N-terminal residues Cys-Tyr-Cys, which are also not found in serum isolates, were removed from the coding sequence (125).

CONTROLLING OTHER VARIABLES THAT AFFECT SOLUBILITY

Inoculum Preparation and Storage

One of the greatest causes of variation in the production of proteins in bacteria is plasmid and host strain instability. The plasmid should be checked periodically during large-scale cultivation. Loss of plasmid, or alterations, including changes in the sequence of the desired protein or even deletion of the whole segment and insertion of unrelated DNA into the plasmid, can occur. Several different insertion sequences are known in *E. coli*. Correct preparation of inocula to minimize variation is essential and is thus dealt with here in some detail.

Most strain stability problems are not discussed openly in papers dedicated to studying the proteins themselves (with notable exceptions [126]). Because many expression problems can be ameliorated by care in preparing the inoculum, a detailed protocol is included here for several stages.

The following general recommendations are made for the start of culture:

1. Larger inocula are preferred.
2. Glycerol cultures should not be reused, and they should never be stored at temperatures above -70 °C.
3. Glycerol cultures should be made from the first culture after a transformation.

Glycerol Culture Preparation and Storage. The following protocol proved reliable for culture of various *E. coli* strains producing interferon and other proteins in differing vector backgrounds:

1. A transformation was done by standard protocols (62), and the bacteria were grown up on selective solid medium.
2. Five to six colonies from the transformation were picked individually into numbered flasks containing 50 mL of selective media (L-broth plus antibiotic) and grown up on a shaker to midexponential phase. Then between 50 and 100 glycerol cultures were made from each flask by combining 0.4–0.5 mL of the culture with 0.5 mL 50% glycerol in L-broth. (In practice, sterile Eppendorf tubes—with a freezer-safe lock closure—or cryotubes were filled with the glycerol solution, closed, and autoclaved. The culture was added sterilly and the tube was reclosed.) The cultures were stored on ice for 2 h, frozen overnight at -20 °C, and then moved to -70 °C or liquid nitrogen storage. In addition, $2\times$ 1 mL of each culture was centrifuged (5 min centrifugation in a cold room in an Eppendorf centrifuge) and the pellet frozen for plasmid analysis.
3. The remaining culture was induced in an appropriate fashion, the cells were lysed, and the amount of protein or its activity was determined. A plasmid prep was made of the selected colony, and the restriction map compared to that expected. The culture with the highest yield was chosen for further analysis (assuming its restriction map was in order), and the other cultures and the glycerol cultures made from them were discarded.*Note*: It is unwise to ignore oddities in the restriction map. Because 1 μL of a commercial *Eco*RI or *Bam*HI preparation will cut an entire miniprep of plasmid completely in a matter of minutes, unexplainable high molecular weight bands are probably not residual host DNA. Unexpected bands may indicate either latent bacteriophage contamination or insertion sequences in the plasmid. When in doubt, throw it out (the culture that is) and do another transformation.
4. To each of the chosen glycerol cultures, afix a freezer-stable label with the plasmid name, host strain, resistance marker, date, and initials of the experimenter. Cultures should be stored in numbered boxes, and a file kept of their location in the -70 °C

or liquid nitrogen storage. Excess glycerol cultures of strains no longer in use should be weeded at regular intervals (it is best not to wait until the compressor needs replacing or the contents of the freezer must be moved).

Cultures prepared and selected in this fashion are relatively stable; the author recently prepared Mx protein from a culture that had been frozen in a −70 °C freezer (with all the temperature variations that can represent) for more than 10 years. Nonetheless, one should maintain a good plasmid library of minipreps of vectors as the ultimate backup.

Trouble-Shooting Protein-Producing Recombinant Bacteria

Most host *E. coli* strains carry a portfolio of mutations that may be essential to their success in producing protein. Several cloning and expression strains and their genotypes are listed in Table 5. The requirements for a cloning strain (basically, high efficiency of transformation) are different from those for a good expression host. The latter should grow vigorously on minimal media and be low in proteases, whereas the former, such as HB101, should have few host restriction systems and high sensitivity to the selection marker. Do not store host strains on agar plates in the cold room; the best method is to prepare them as glycerol cultures and store at −70 °C or liquid nitrogen temperature.

It is vital to continuously monitor the quality of these host strains, especially when several researchers share stock cultures. For best results, one should prepare several hundred glycerol cultures of common host strains when they are obtained in the lab. One of these should be used per experiment, preferably by adding the whole 1 mL to 50 mL of growth media used to induce transformation phenotype. Although not suitable for long-term storage or in cases where a high yield is necessary from the transformation, competent cells can be stored frozen and used for several weeks after. When performing a transformation, a control plate of competent cells with no DNA added should be used to test for development of antibiotic resistance. Look for differences in the characteristic color, size, and odor of host strain colonies streaked on a nonselective agar medium.

If necessary, one can test the host strain for the presence of markers. Reversion to a protease-producing phenotype can be detected by plating on agar containing 1% casein. A complete listing of the meaning of various markers for *E. coli*, with alternate notation possibilities, can be found on the internet (61); the more common mutations are described by Sambrook et al. (62).

Protecting Your Cultures from Contamination. When bacteriophages were first introduced, in the 1960s, several labs were sufficiently afraid of their invasiveness as to limit work with them to discrete areas. This procedure should be adhered to today. A wide-spread M13 phage contamination can be caused from one shaker accident that is not adequately cleaned up, and months and even years of subsequent work can be wasted if the problem is ignored. Because switching to an F^- *E. coli* strain will limit possible hosts, the best method is to prevent the original contamination.

Frequent hand washing with antiseptic hand soap and even wearing disposable gloves is essential to avoid cross-contamination of cultures when handling plasmids and other replication-competent nucleic acids. Surfaces should be covered with absorbant paper that is changed regularly, preferably at the end of every experiment. The counter underneath and all cold room surfaces should be washed down with 2% chlorine bleach or Javell water, or 0.5% hydrogen peroxide, frequently. Isolation precautions such as those used throughout the medical field and the food preparation industry should be automatic during work with recombinant organisms. One should avoid touching hair or face during work, and long hair should be tied back. Protective coats should be removed at the door to the lab and never worn in public areas, especially lunchrooms.

Discourage other groups from using your shakers unless they are willing to conform to the following rules, which seem to be obvious:

1. All cultures must be stoppered properly (preferably with a cotton wad wrapped in cheese cloth or some other thick filtering substance) and the stopper fixed in position with tape so that it will not fly off during culture. Aluminum foil is not adequate. Flasks should fit the holder, and one should avoid wrapping paper towels and such around smaller flasks to accommodate them to a larger holder. Multiple test tube cultures must be in a properly secured rack, not taped to the side of a flask holder.
2. Flasks should be removed from the shaker area or room if additions are made after inoculation.
3. The shaker should be cleaned often with an antiseptic solution.
4. UV sterilization using a fluorescent light (one with a time switch can be installed for little expense in any area) is an excellent method to halt the growth of microorganisms, if the light has not been used for more than 300 h of operation and is not left on for more than 15 min at a time. Improperly used, UV irradiation is worse than no sterilization at all because a weak light can select for mutant bacteria that will be resistant to most killing methods. An easy way to determine if the lamp is still in order is to inoculate an agar plate with 1 mL of a dense culture of bacteria (preferably an aerobic spore former such as *Bacillus*, but *E. coli* will do) and leave it open under the lamp for 15 s (time accurately!). If any colonies grow on the plate after 24 h at 30–37 °C, install a new fluorescent bulb.

Fermentor Culture Starting from Glycerol Cultures. Even a 30,000-L fermentor can be inoculated from 1 mL of culture if the plasmid is stable. Do not reuse a glycerol culture; use it once and dispose of it.

Several methods can be used to determine plasmid stability as a function of culture time. One can compare the specific plasmid content of the bacteria at the end of a plasmid miniprep and determine the colony-forming units of a transformation into HB101 or a high frequency of transformation strain.

Stability of Selection Stress. Most antibiotics (in particular, penicillin derivatives) should not be stored in frozen aqueous solution, nor should antibiotic-containing agar plates be stored before use. The half-life of the closed β-lactam ring is extremely short in aqueous solution, particularly at neutral pH. Ampicillin should be dissolved in sterile water, filter sterilized, and added to the growth medium as close to time of inoculation as humanly or mechanically possible.

Before ignoring this caveat, note that those frozen ampicillin solutions inherited from the previous lab generation, while most certainly lacking their ability to lyse bacteria, may inhibit growth of nontransformed bacteria. Week-old refrigerated antibiotic-containing plates used for transformation selection will show no colonies for the control Ca^{2+} cells. As the transformed bacteria grow, they secrete β-lactamase that, coupled with the instability of the antibiotic itself at 37 °C, will rapidly eliminate any selective pressure in that area of the plate (or in the vessel during a large-scale cultivation). The invisible nontransformed bacteria will start to grow. Colonies picked from the transformation plates will perforce contain nontransformed but not necessarily nonviable bacteria. The initially inhibited nonproducers will be able to grow, diluting the plasmid bearers and lowering yield. If the plasmid confers any negative growth characteristics, the plasmid-containing bacteria will rapidly become the minority in the culture vessel. Additional proteins from all those nonproducing bacteria will render the purification even more difficult.

Maintaining Solubility during Harvesting and Cell Lysis

If one is fortunate, the protein to be isolated has distinct characteristics that separate it from the pack. Unfortunately, the majority of naturally occurring proteins are moderately hydrophobic, with a pK_a of between 6 and 7. This means they will elute from hydrophobic interaction columns at salt concentrations between 0.1 and 0.5 N, and from ion exchange resins with the vast majority of bacterial proteins. Thus a typical protein, if expressed at low concentration in the bacterial cytoplasm, will be difficult to separate by general methods of purification and may be destroyed by proteases before ever being seen on a gel.

Here, IB formation would appear to confer an enormous advantage. The protein is clearly visible from the earliest stages of expression on a polyacrylamide gel, and an apparently purified preparation can be quickly obtained. The denaturants used at the early stages of preparation inhibit most proteases, and the overall yield, estimated visually, would seem to be very high.

The apparent ability to wash away contaminants from IBs is often a Trojan horse. Membrane fragment and other insoluble proteins persist through many detergent washes (indeed, there is a report that some proteins are rendered insoluble by these agents, particularly Triton X-100). These high molecular weight, often highly modified proteins are not always visible on gel electrophoresis because they "stick in the slot" or do not stain with silver stain reagents or Coomassie blue. An even greater problem arises: the protein is contaminated with proteolytic fragments, alternate folded forms, and oligomers of itself. If solubly expressed protein requires isolation from bacterial proteins with similar characteristics, the problem in IB schemes comes down to isolating the protein from what is essentially the same protein. This is the hardest of all purification steps.

Even proteins that are initially soluble in the bacteria may be induced to aggregate, sometimes irreversibly, during cell breakage and washing. Aggregation and proteolysis can be minimized by adding a proteolysis inhibitor to the cell culture at the end of growth and maintaining a low temperature during centrifugation. Especially on the large scale, this may require special cooling equipment for centrifuges, and small batch processing for producing cell lysates using high-pressure equipment such as a Manton/Gaulin or French press.

As noted earlier, detergents such as Triton X-100 and Nonidet P-40 can render proteins insoluble when added during cell washing/breakage steps. However, others have noted that extraction with 1% Triton X-100 was necessary for obtaining high yields of bovine heart mitochondrial ubiquinol–cytochrome *c* reductase (fusion protein with glutathione). They solved later aggregation problems by resolubilizing the protein in the detergent dodecyl maltoside in 300 mM NaCl; the resultant protein appeared active because it gave a blue shift in the visible spectrum when incubated with ubiquinone (12). The preliminary breakage steps to maintain the solubility of T7 RNA polymerase had to be done in buffer containing large amounts of glycerol and detergent to prevent the protein from precipitating. Once its extremely narrow solubility range was determined (10), the purification was straightforward.

Even the pH of the extraction buffer can have a profound effect on solubility. For example, UMP-kinase from *E. coli* was insoluble if extracted from cells with a pH 7.4 Tris buffer, but was soluble if 1 mM UTP was added to this buffer or if 100 mM borate buffer pH 9 was used for the extraction. Reversible aggregation could be induced by adding Mg to the UTP-containing solutions or lowering the pH. A point mutant, D159N, was soluble at up to 5 mg/mL at pH 7.4 (127).

As the last example demonstrates, the effects of point mutations on solubility do not fit easily into a clear set of empirical or theoretical rules. The reader is referred to Refs. 92 and 108 for further examples and discussion.

CONCLUSIONS

Proteins produced in a soluble form in bacteria are easier to purify and work with than those refolded from inclusion bodies. Expression of soluble proteins can be improved by judicious choice of the vector and host strain, and manipulation of the inoculation, growth, and induction conditions. Very high per liter yields of purified product have been achieved, and the overall yield for some secretion systems is higher than those possible with inclusion bodies.

BIBLIOGRAPHY

1. C.H. Schein, *Bio / Technology* **7**, 1141–1149 (1989).
2. C.H. Schein, *Curr. Opin. Biotechnol.* **2**, 746–750 (1991).

3. M.C. Hensing, R.J. Rouwenhorst, J.J. Heijnen, J.P. van Dijken, and J.T. Pronk, *Antonie van Leeuwenhoek* **67**, 261–279 (1995).
4. A.R. Mistry, L. Falciola, L. Monaco, R. Tagliabue, G. Acerbis, A. Knight, R.P. Harbottle, M. Soria, M.E. Bianchi, C. Coutelle, and S.L. Hart, *BioTechniques* **22**, 718–729 (1997).
5. J.W. Harper and N.J. Logsdon, *Biochemistry* **30**, 8060–8066 (1991).
6. T. Yasukawa, C. Kanei-Ishii, T. Maekawa, J. Fujimoto, T. Yamamoto, and S. Ishii, *J. Biol. Chem.* **270**, 25328–25331 (1995).
7. C.H. Schein, E. Boix, M. Haugg, K.P. Holliger, S. Hemmi, G. Frank, and H. Schwalbe, *Biochem. J.* **283**, 137–144 (1992).
8. P. Carter, R.F. Kelley, M.L. Rodrigues, B. Snedecor, M. Cavarrubias, M.D. Velligan, W.L.T. Wong, A.M. Rowland, C.E. Kotts, M.E. Carver, M. Yang, J.H. Bourell, H.M. Shepard, and D. Henner, *Bio/technology* **10**, 163–167 (1992).
9. U. Horn, W. Strittmayer, A. Krebber, U. Knupfer, M. Kujau, R. Wenderoth, K. Muller, S. Matzku, A. Plückthun, D. Riesenberg, *Appl. Microb. Biotechnol.* **46**, 524–532 (1996).
10. C.H. Scein, *Bio/Technology* **8**, 308–317 (1990).
11. C.H. Schein and M.H.M. Noteborn, *Bio/Technology* **6**, 291–294 (1988).
12. L. Yu, K.P. Deng, and C.A. Yu, *J. Biol. Chem.* **270**, 25634–25638 (1995).
13. Sooryanarayana, P.R. Adiga, and S.S. Visweswariah, *Protein Express. Purif.* **7**, 147–154 (1996).
14. C. Randak, A.A. Roscher, H.B. Hadorn, I. Assfalgmachleidt, E.A. Auerswald, and W. Machleidt, *FEBS Lett.* **363**, 189–194 (1995).
15. B. Cooper, J.A. Gruber, and M.J. McPhaul, *J. Steroid Biochem. Mol. Biol.* **57**, 251–257 (1996).
16. S. Morino, H. Hazama, M. Ozaki, Y. Teraoka, S. Shibaata, M. Doi, H. Ueda, T. Ishida, and S. Uesugi, *Eur. J. Biochem.* **239**, 597–601 (1996).
17. T. Mizukami, Y. Komatsu, N. Hosoi, S. Itoh, and T. Oka, *Biotechnol. Lett.* **8**, 605–610 (1986).
18. H.F. Seow, C.R. Goh, A.G. Porter, and L. Krischnan, *Gene* **83**, 117–129 (1989).
19. R. Hindges and U. Hubscher, Gene **158**, 241–246 (1995).
20. L.W. Slice, and S.S. Taylor, *J. Biol. Chem.* **264**, 20940–20946 (1989).
21. M.F. Browner, P. Rasor, S. Tugendreich, and R.J. Fletterick, *Protein Eng.* **4**, 351–357 1991.
22. J.M. Tremblay, G.M. Helmkamp, and L.R. Yarbrough, *J. Biol. Chem.* **271**, 16139–16143 (1996).
23. A.M. Tsai, M.J. Betenbaugh, and J. Shiloach, *Biotechnol. Bioeng.* **48**, 715–718 (1995).
24. A. Battistoni, M.T. Carri, A.P. Mazzetti, and G. Rotilo, *Biochem. Biophys. Res. Commun.* **186**, 1339–1342 (1992).
25. A. Ferrer, C. Aparicia, N. Nogués, A. Wettstein, T.J. Bach, and A. Boronat, *FEBS Lett.* **266**, 67–71 (1990).
26. Y. Shirano, and D. Shibata, *FEBS Lett.* **271**, 128–130 (1990).
27. E. Kopetzki, G. Schumacher, and P. Buckel, *Mol. Gen. Genet.* **216**, 149–155 (1989).
28. J. Fetzer, and G. Folkers, *Pharm. Pharmacol. Lett.* **2**, 112–114 (1992).
29. M.B. Slabaugh, R.E. Davis, N.A. Roseman, and C.K. Mathews, *J. Biol. Chem.* **268**, 17803–17810 (1993).
30. Y. Saito, Y. Ishii, T. Fijimura, H. Sasaki, Y. Noguchi, H. Yamada, M. Niwa, and K. Shiomura, *Ann. N.Y. Acad. Sci.* **782**, 79–86 (1996).
31. S. Tiow-suan, and D.S.H. Tan, *Biochem. Mol. Biol. Int.* **35**, 1069–1078 (1995).
32. J.H. Toney, J.K. Wu, K.M. Overbye, C.M. Thompson, and D.L. Pompliano, *Protein Express. Purif.* **9**, 355–362 (1997).
33. W.R. Bishai, R. Rappuoli, and J.R. Murphy, *J. Bacteriol.* **169**, 5140–5151 (1987).
34. H.S. Su, F. Blain, R.A. Musil, J.J.F. Zimmermann, K.F. Gu, and D.C. Bennett, *Appl. Environ. Microbiol.* **62**, 2723–2734 (1996).
35. G. Schnurr, N. Misawa, and G. Sandmann, *Biochem. J.* **315**, 869–874 (1996).
36. A. Mitraki, and J. King, *FEBS Lett.* **307**, 20–25 (1992).
37. M. Piatak, J.A. Lane, W. Laird, M.J. Bjorn, A. Wang, and M. Williams, *J. Biol. Chem.* **263**, 4837–4843 1988.
38. J. Klein and P. Dhurjati, *Appl. Environ. Microbiol.* **61**, 1220–1225 (1995).
39. L.H. Yang, S. Ahmed, and E.W. Miles, *Protein Express. Purif.* **8**, 126–136 (1996).
40. W.L. Picking, J.A. Mertz, M.E. Marquart, and W.D. Picking, *Protein Express. Purif.* **8**, 401–408 (1996).
41. T.C. Lin, J. Rush, E.K. Spicer, and W.H. Konigsberg, *Proc. Natl. Acad. Sci.* **84**, 7000–7004 (1987).
42. P.G. Jones, R.A. VanBogelen, and F.C. Neidhardt, *J. Bacteriol.* **169**, 2092–2095 (1987).
43. V. Mainfroid, P. Terpstra, M. Beauregard, J.M. Frere, S.C. Mande, W.G.J. Hol, J.A. Martial, and K. Goraj, *J. Mol. Biol.* **257**, 441–456 (1996).
44. M.-C. McKenna, C. Muchardt, R. Gaynor, and D. Eisenberg, *Protein Express. Purification* **5**, 105–111 (1994).
45. S. Sakamoto, I. Terada, Y.C. Lee, K. Uehara, H. Matsuzara, and M. Iijima, *Appl. Microbiol. Biotechnol.* **45**, 94–101 (1996).
46. A.A. Michels, V.-T. Nguyen, A.W.T. Konings, H.H. Kampinga, and O. Bensaude, *Eur. J. Biochem.* **234**, 382–389 (1995).
47. R.G. Matthews, and F.C. Neidhardt, *J. Bacteriol.* **171**(5), 2619–2625 (1989).
48. D. Riesenberg, *Curr. Opin. Biotechnol.* **2**, 380–384 (1991).
49. L. Andersson, L. Strandberg, L. Haggstrom, and S.O. Enfors, *FEMS Microbiol. Rev.* **14**(1), 39–44 (1994).
50. I. Palva, P. Lehtovaara, L. Kèèrièinen, M. Sibakov, K. Cantell, C.H. Schein, K. Kashiwagi, and C. Weissmann, *Gene* **22**, 229–235 (1983).
51. G.A. Bowden, A.M. Paredes, and G. Georgiou, *Bio/Technology* **9**, 725–730 (1991).
52. T. Nystrom and F.C. Neidhardt, *J. Bacteriol.* **175**, 3949–3956 (1993).
53. C.H. Schein, K. Kashiwagi, A. Fujisawa, and C. Weissmann, *Bio/Technology* **4**, 719–725 (1986).
54. J.R. McDonald, M. Org, C. Shen, Z. Parandoosh, B. Sosnowski, S. Bussell, and L.L. Houston, *Protein Purif. Express.* **8**, 97–108 (1996).
55. G.R. Smith, R. Ford, J.D. Bryant, R.L. Gambley, T.K. Meghia, R.M. Harding, and J.L. Dale, *Arch. Virol.* **140**, 1817–1831 (1995).
56. T.V. Huynh, R.A. Young, and R.A. Davis, D. Glover ed., *DNA Cloning*, Oxford University Press, Oxford, 1985, pp., 56–110.
57. F. Bolivar and K. Backman, *Methods Enzymol.* **68**, 245 (1979).
58. C. Yanish Perron, J. Viera, and J. Messing, *Gene* **33**, 103–199 (1985).
59. R. Maurer, B.J. Meyer, and M. Ptashne, *J. Mol. Biol.* **139**, 147–161 (1980).
60. W.O. Bullock, J.M. Fernandez, and J.M. Short, *Biotechniques* **5**, 376–379 (1987).

61. M. Kroger, and R. Wahl, *Nucleic Acids Res.* **25**, 39–42 (1997).
62. J. Sambrook, E.F. Fritsch, and T. Maniatis, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989.
63. G.T. Yarranton and A. Mountain, in R.R. Rees, M.J.E. Sternberg, R. Wetzel, eds., *Protein Engineering: A Practical Approach*, IRL Press, New York, 1992, pp. 303–325.
64. W. Chen, L. Wright, S. Lee, S.D. Cosloy, and C.S. Russell, *Biochim. Biophys. Acta* **1309**, 109–121 (1996).
65. S.A. Goff, L.P. Casson, and A.L. Goldberg, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6647–6651 (1984).
66. B. Andrews, H. Adari, G. Hannig, E. Lahue, M. Gosselin, S. Martin, A. Ahmed, P.J. Ford, E.G. Hayman, and S.C. Makrides, *Gene* **182**, 101–109 (1996).
67. A.J. Derman, W.A. Prinz, D. Belin, and J. Beckwith, *Science* **262**, 1744–1747 (1993).
68. K. Proba, L. Ge, and A. Plückthun, *Gene* **159**, 203–207 (1995).
69. M. Bogdanov, J. Sun, H.R. Kaback, and W. Dowhan, *J. Biol. Chem.* **271**, 11615–11618 (1996).
70. J.G. Thomas and F. Baneyx, *J. Biol. Chem.* **271**, 11141–11147 (1996).
71. P.T. Lallio and J.E. Bailey, *Biotechnol. Prog.* **12**, 31–39 (1996).
72. J.E. Bailey, A. Sburlati, V. Hatzimanikatis, K. Lee, W.A. Renner, and P.S. Tsai, *Biotechnol. Bioeng.* **52**, 109–121 (1996).
73. S. Hober, A. Hansson, M. Uhlen, and B. Nilsson, *Biochemistry* **33**, 6758–6761 (1994).
74. T. Katzav-Gozansky, E. Hanan, and B. Solomon, *Biotechnol. Appl. Biochem.* **23**, 227–230 (1996).
75. O.L. Garcia, B. Gonzalez, A. Menendez, A.E. Sosa, J.R. Fernandez, H. Santana, and N. Meneses, *Ann. N.Y. Acad. Sci.* **782**, 79–86 (1996).
76. H. Hashiba, R. Takiguchi, K. Jyoho, and K. Aoyama, *Biosci. Biotechnol. Biochem.* **56**, 190–194 (1992).
77. M. Takagi, S.P. Lee, and T. Imanaka, *J. Ferment. Bioeng.* **81**, 557–559 (1996).
78. D. Dahan, R. Srikumar, R. Laprade, and J.W. Coulton, *FEBS Lett.* **392**, 304–308 (1996).
79. D.M. Rothstein, P.E. Devlin, and R.L. Cate, *J. Bacteriol.* **168**, 839–842 (1986).
80. H. Burger, R.W. van Leen, L.C. Dorssers, N.L. Persoon, P.J. Lemson, and G. Wagemaker, *Blood* **76**, 2229–2234 (1990).
81. L. Kremer, G. Riveau, A. Baulard, A. Capron, and C. Locht, *J. Immunol.* **156**, 4309–4317 (1996).
82. C. Rush, L. Hafner, and P. Timms, *Appl. Microbiol. Biotechnol.* **47**, 537–542 (1997).
83. P.M. Norton, R.W.F. Le Page, and J.M. Wells, *Folia Microbiol.* **40**, 225–230 (1995).
84. P. Samuelson, M. Hansson, N. Ahlborg, C. Andreoni, F. Gotz, T. Bachi, T.N. Nguyen, H. Binz, M. Uhlen, and S. Stahl, *J. Bacteriol.* **177**, 1470–1476 (1995).
85. S. Liljeqvist, D. Haddad, K. Berzins, M. Uhlen, and S. Stahl, *Biochem. Biophys. Res. Commun.* **218**, 356–359 (1996).
86. J.I. Moreno, *Protein Express. Purif.* **8**, 332–340 (1996).
87. A. Loregian, T.R. Hirst, H.S. Marsden, and G. Palu, *Protein Express. Purif.* **8**, 381–389 (1996).
88. N. Mertens, E. Remaut, and W. Fiers, *Biotechnology* **13**, 175–179 (1995).
89. J.A. Vasina and F. Baneyx, *Appl. Environ. Microbiol.* **62**, 1444–1447 (1996).
90. J.A. Vasina and F. Baneyx, *Protein Express. Purific.* **9**, 211–218 (1997).
91. K.C. Sitney, M.B. Mann, G.W. Stearns, A.D. Menjares, J.L. Stevenson, M.D. Snavely, J.C. Fieschko, C. Curless, and L.B. Tsai, *Ann. N.Y. Acad. Sci.* **782**, 297–310 (1996).
92. C.H. Schein, *Curr. Opin. Biotechnol.* **4**, 456–461 (1993).
93. E. Boix, M.V. Nogués, C.H. Schein, S.A. Benner, and C.M. Cuchillo, *J. Biol. Chem.* **269**, 2529–2534 (1994).
94. Y. Ishii, T. Nakano, N. Honma, N. Yuyama, Y. Yamada, H. Watarai, T. Tomura, M. Sato, H. Tsumura, and T. Ozawa, *J. Immunol. Methods* **186**, 27–36 (1995).
95. J.M. Clement, M. Jehanno, O. Popescu, W. Saurin, and M. Hofnung, *Protein Express. Purif.* **8**, 319–331 (1996).
96. R.R. Dickason, R.A. Edwards, J. Bryan, and D.P. Huston, *J. Immunol. Methods* **185**, 237–244 (1995).
97. M.H.M. Koken, H.H. Odijk, M. van Duin, M. Fornerod, and J.H. Hoeijmakers, *Biochem. Biophys. Res. Comm.* **195**, 643–653 (1993).
98. M. Larsson, E. Brundell, L. Nordfors, C. Hoog, M. Uhlen, and S. Stahl, *Protein Express. Purif.* **7**, 447–457 (1996).
99. M.D. Spangfort, H. Ipsen, S.H. Sparholt, S. Aasmul-Olsen, M.R. Larsen, E. Mortz, P. Roepstorff, and J.N. Larsen, *Protein Express. Purif.* **8**, 365–373 (1996).
100. X. Chenivesse, T. Huby, J. Chapman, D. Franco, and J. Thillet, *Protein Express. Purif.* **8**, 145–150 (1996).
101. S. Patel and B.M. Austen, *Protein Express. Purif.* **8**, 283–294 (1996).
102. R.L. Armstrong, J.E. Wilson, and C.B. Shoemaker, *Protein Express. Purif.* **8**, 374–380 (1996).
103. E.R. LaVallie, E.A. DiBlasio, S. Kovacic, K.L. Grant, P.F. Schendal, and J.M. McCoy, *Bio/Technology* **11**, 187–193 (1993).
104. E. Samuelsson and M. Uhlen, *Ann. N.Y. Acad. Sci.* **782**, 486–494 (1996).
105. S.M. Kakonen, J. Hellman, K. Pettersson, T. Lovgren, and M. Karp, *Protein Express. Purif.* **8**, 137–144 (1996).
106. X. Hu, Q. Shi, T. Yang, and G. Jackowski, *Protein Express. Purif.* **7**, 289–293 (1996).
107. A.J. Day, R.T. Aplin, and A.C. Willis, *Protein Express. Purif.* **8**, 1–16 (1996).
108. C.H. Schein, *Pharm. Acta Helv.* **69**, 119–126 (1994).
109. D.L. Wilkinson and R.G. Harrison, *Bio/Technology* **9**, 443–448 (1991).
110. W.F. Prouty, M.J. Karnovsky, and A.L. Goldberg, *J. Biol. Chem.* **250**, 1112–1122 (1975).
111. J. King, C. Haase-Pettingell, A.G. Robinson, M. Speed, and A. Mitraki, *FASEB J.* **10**, 57–66 (1996).
112. M.A. Speed, D.I.C. Wang, and J. King, *Nat. Biotechnol.* **14**, 1283–1287 (1996).
113. M.J. Bennett, M.P. Schlunegger, and D. Eisenberg, *Protein Sci.* **4**, 2455–2468 (1995).
114. C.Ó. Fágáin, *Biochim. Biophys. Acta* **1252**, 1–14 (1995).
115. K. Tokatlidis, P. Dhurjati, J. Millet, P. Béguin, and J.-P. Aubert, *FEBS Lett.* **282**, 205–208 (1991).
116. R. Wetzel, J.L. Perry, and C. Veilleux, *Bio/Technology* **9**, 731–737 (1991).
117. A. Mitraki, B. Fane, C. Haase-Pettingell, J. Sturtevant, and J. King, *Science* **253**, 54–58 (1991).
118. S. Krishnamurthi, T.A. Dickens, Y. Patel, C.P. Wheeler-Jones, and V.V. Kakkar, *Biochem. Biophys. Res. Comm.* **163**, 1256–1264 (1989).
119. D.N. Brems, L.A. Alter, M.J. Beckage, R.E. Chance, R.D. DiMarchi, L.K. Green, H.B. Long, A.H. Pekar, J.E. Shields, and B.H. Frank, *Protein Eng.* **5**, 527–533 (1992).

120. M.C. Mossing and R.T. Sauer, *Science* **250**, 1712–1715 (1990).
121. J.M. Betton and M. Hofnung, *J. Biol. Chem.* **271**, 8046–8052 (1996).
122. P.A. Nygren, S. Stahl, and M. Uhlen, *Trends Biotechnol.* **12**, 184–188 (1994).
123. M. Murby, E. Samuelsson, T.N. Nguyen, L. Mignard, U. Power, H. Binz, M. Uhlen, and S. Stahl, *Eur. J. Biochem.* **230**, 38–44 (1995).
124. M. Murby, M. Uhlen, and S. Stahl, *Protein Express. Purif.* **7**, 129–136 (1996).
125. C.H. Schein and M. Haugg, *Biochem. J.* **307**, 123–127 (1995).
126. M.-C. McKenna, C. Muchardt, R. Gaynor, and D. Eisenberg, *Protein Express. Purif.* **5**, 105–111 (1994).
127. L. Serina, N. Bucurenci, A.M. Gilles, W.K. Surewicz, H. Fabian, H.H. Mantsch, M. Takahashi, I. Petrescu, G. Batelier, and O. Bârzu, *Biochemistry* **35**, 7003–7011 (1996).

See also β-GALACTOSIDASE, ENZYMOLOGY; ENZYMES, PULP AND PAPER PROCESSING; GENE TRANSFER, GRAM-POSITIVE BACTERIA; INSECT CELLS AND LARVAE, GENE EXPRESSION; PROTEIN SECRETION, *SACCHAROMYCES CEREVISIAE*; *PICHIA*, OPTIMIZATION OF PROTEIN EXPRESSION.

PROTEIN GLYCOSYLATION

NIGEL JENKINS
Eli Lilly, Inc.
Indianapolis, Indiana
PIR M. SHAH
De Montfort University
Leicester, U.K.

KEY WORDS

Carbohydrates
Glycosylation
Recombinant proteins
Review

OUTLINE

INTRODUCTION

Protein glycosylation, the addition of sugar residues to the polypeptide backbone, is the most extensive of all the cellular posttranslational modifications made to proteins in the eukaryotic cell. It entails attachment and remodeling of oligosaccharides in the endoplasmic reticulum (ER) and Golgi apparatus prior to secretion of the protein or its appearance at the cell surface (1,2). Almost all proteins on the mammalian cell surface (including receptors) are glycoproteins (3), and the majority of secreted proteins are also glycosylated (with notable exceptions such as insulin, growth hormone, and some interleukins). Glycosylation has been shown to influence the secretion, solubility, stability, antigenicity, and blood clearance of glycoproteins. In some cases, carbohydrates form part of the ligand binding site. In cases where this does not occur, the influence of carbohydrates may be overlooked until the protein enters animal and human clinical trials. There have been four major areas of advance in recent years:

1. Development of methods for the accurate structural analysis of carbohydrates attached to glycoproteins (glycans)
2. Cloning of key glycosyltransferase enzymes in the glycan biosynthetic pathways.
3. Understanding the cellular influences on protein glycosylation.
4. Understanding the varied biological functions of carbohydrates in glycoproteins.

In this article the methods and findings that underpin the science of glycobiology are reviewed, with particular focus on their application in the biotechnology industry. Differences in carbohydrate structures that may arise from choosing alternative gene expression systems and cell culture conditions are discussed, together with methods of carbohydrate analysis that are appropriate at each stage of product development. For reviews of other aspects of glycobiology the reader should consult reviews on biosyn-

thetic pathways (4,5), the biological properties of carbohydrates (6), and the cellular influences on the glycosylation process (2,7).

Oligosaccharide Structures

Three distinct types of glycosylation are known to occur in higher plants and animals, two of which are shown in Figure 1:

1. N-linked glycosylation of proteins; oligosaccharides are attached to the amino group of asparagine residues, usually within the sequence Asn-X-Ser/Thr.
2. O-linked glycosylation; smaller oligosaccharides are linked to serine or threonine residues.
3. Glucosamine and mannose residues are incorporated into the glycosyl–phosphatidylinositol (GPI) anchor of some membrane proteins.

In addition, single fucose residues (8) or small glycans linked via Asn-X-Cys have been found on some blood proteins (9).

N-Linked Glycosylation

N-linked glycosylation is characterized by a glycosidic bond to the R-group of an asparagine residue of the peptide backbone. A tripeptide core or the consensus peptide sequence Asn-X-Ser/Thr (where X is any amino acid apart from proline) must be present for this particular form of glycosylation. The presence of this amino acid sequence by no means guarantees its glycosylation, because other factors such as protein secondary structure and its folding status within the ER are important influences (10,11). For example, the immunomodulator interferon-γ (IFN-γ) has two potential N-glycosylation sites (Asn_{25} and Asn_{97}), but exists in three glycoforms: doubly glycosylated (at both Asn_{25} and Asn_{97}), singly glycosylated (at Asn_{25} only), and nonglycosylated (12). Similarly, in the blood protease tissue plasminogen activator (tPA), only two N-glycosylation sites (Asn_{117} and Asn_{448}) are fully occupied: Asn_{184} is only 50% glycosylated due to steric hindrance (11).

N-linked oligosaccharides fall into three structural categories: complex, high-mannose (or oligomannose), and hybrid, dependent on the extent of glycan remodeling in the later stages of the glycosylation process (Fig. 1). All three categories contain the same chitobiose core structure of two *N*-acetylglucosamine (GlcNAc) units attached to the asparagine containing sequon, which is extended by three mannose (Man) units, with a single fucose (Fuc) residue sometimes attached to the GlcNAc of the reducing end (next to its peptide attachment site).

1. Complex oligosaccharides contain two or more branches attached to the core, and the terms *biantennary*, *triantennary*, and *tetraantennary* reflect the number of branches. Sugar residues such as GlcNAc, galactose (Gal), *N*-acetylgalactosamine (GalNAc), Fuc, and sialic acids (*N*-acetyl [NeuAc] and *N*-glycolyl [NeuGc] neuraminic acids) extend the structures.
2. Oligo (5–9) and high-mannose (>9 Man) structures contain only mannose residues attached to the core structure.

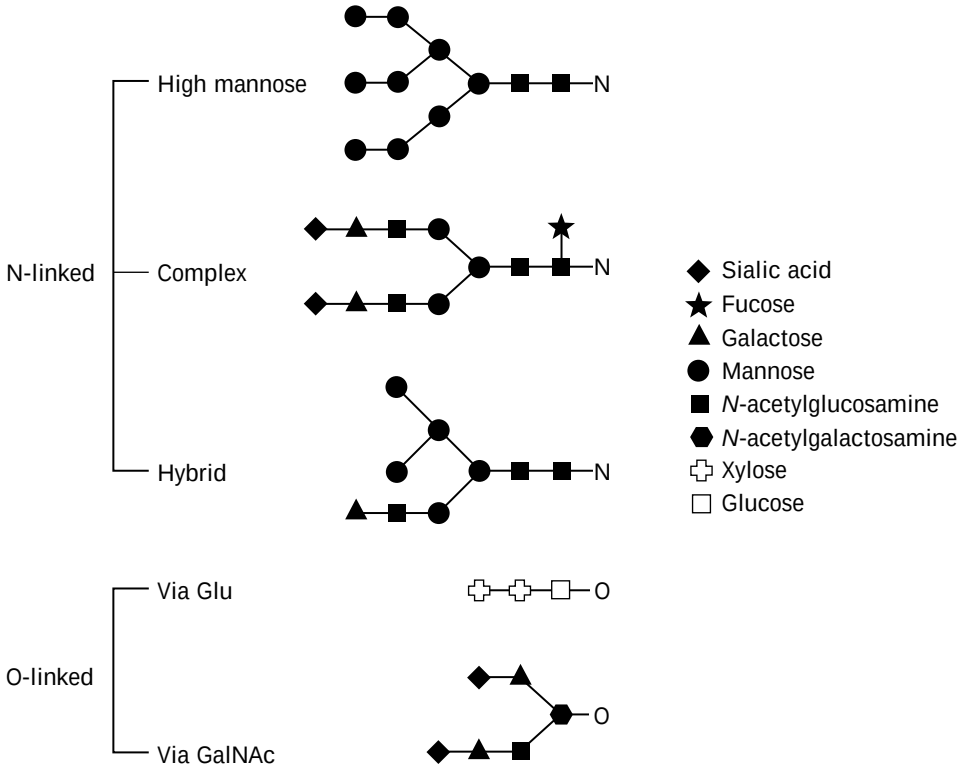


Figure 1. Common oligosaccharide structures found on glycoproteins.

3. Hybrid structures contain at least complex sugar-containing antennae, with other α1-3 branches containing only mannose residues.

N-glycosylation occurs in both the ER and Golgi apparatus and is controlled by a large number of glycosyltransferases and glycosidases. Glycosyltransferases are grouped into families, based on the type of sugar they transfer to the protein (e.g., galactosyltransferases and sialyltransferases) (5). Theoretically, if all enzymes in the pathway are fully active, the required donor sugar nucleotides are present, and no structural hindrances are present, all glycans will be fully processed into complex structures terminating in sialic acids. But in practice, this is rarely observed and a mixture of partially and fully processed glycans is more typical of N-linked glycoproteins.

Figure 2 shows a simplified pathway in the biosynthesis of *N*-glycans. A common core glycan ($Glc_3Man_9GlcNAc_2$) is built on a dolichol pyrophosphate lipid carrier, first in the cytoplasm and then in the ER. This oligosaccharide core structure is attached to the target asparagine of the nascent polypeptide chain as it enters the ER. The enzyme responsible for catalyzing this reaction, oligosaccharyl transferase, is closely associated with the ribosome (step 1). Within the ER lumen the three terminal glucose units are removed by α-glucosidases to give $Man_9GlcNAc_2$ (steps 2 and 3) that is further modified to $Man_8GlcNAc_2$ by α2-mannosidase (step 4). This trimmed structure is no longer recognized by the ER chaperone calnexin and is transported to the *cis*-Golgi via an intermediate compartment. Here it is further trimmed by additional mannosidases until a $Man_5GlcNAc_2$ structure remains (step 5), which is a substrate for GlcNAc transferase I to add GlcNAc to the 1-3 antenna (step 6). The presence of the GlcNAc on this arm of the glycan acts as a signal for possible further modifications. These include the removal of further Man residues (mannosidase II, step 7); addition of antennary GlcNAc (via GlcNAc transferase II, step 8), and bisecting GlcNAc residues (via GlcNAc-transferase III); creation of further branches (via GlcNAc-transferase IV and V); and core α6-fucosylation (step 9).

Therefore, the absence GlcNAc transferase I activity, or the subsequent removal of GlcNAc, as has been suggested for some insect cell lines (13), may result in simple oligomannose or hybrid structures being present on the mature glycoprotein. There appears to be no requirement for more complex glycan processing to precede secretion of glycoproteins from the cell or their appearance on the plasma membrane. Certain pathways are mutually exclusive, for example, the addition of a bisecting GlcNAc residue via GlcNAc-transferase III can prevent further branching of the antennae by GlcNAc-transferase V. The final stages of *N*-glycan processing take place in the *trans*-Golgi and involve the addition of galactose and sialic acids (steps 10 and 11), and α1-6- or α1-3-linked fucose residues. It is the presence and activity of these Golgi enzymes that gives rise to differences in glycan structures (microheterogeneity) produced by alternative cell hosts; these factors vary with culture conditions. Certain cell types posses unusual glycosyltransferases that facilitate the addition of cell-specific glycan structures. For example, up to 20% of IgG molecules produced by human B-lymphocytes possess a bisecting GlcNAc residue β1-4-linked to the central β-linked mannose of the core glycan, which can play a role in antibody-dependent cell-mediated cytotoxicity (ADCC). The GlcNAc transferase III enzyme required for this modification is not active in all cells, indeed only certain rodent cell lines such as the rat Y0 myeloma are able to produce recombinant antibodies containing this bisecting residue (14). Another example is sulfation of antennary GalNAc residues in certain pituitary glycoprotein hormones such as luteinizing hormone (LH), thyrotrophin, and proopiomelanocortin (15). The specific GalNAc transferase and a terminal GalNAc sulfotransferase recognize protein motifs in the nascent peptide (e.g., Pro-Leu-Arg) and are mainly restricted to the anterior pituitary gland. Therefore, only cell lines derived from the pituitary gland or endothelium (such as At20 and 293 cells) are able to perform this sulfation (16).

O-Linked Glycosylation

O-linked glycosylation is a simpler process involving a smaller number of sugar residues (typically 1–6) and occurs exclusively in the Golgi apparatus, where a core GalNAc residue attaches via an O-glycosidic bond to the R-group of serine or threonine (Fig. 1). There is no obvious and invariant peptide sequon surrounding the serine or threonine attachment site, although a computer algorithm has been devised to predict the likely occupation of potential sites, based on sequences in other O-linked glycoproteins (17). Large structural proteins such as mucins have multiple O-linked oligosaccharides, which may dominate the mass of the mature glycoprotein. Also, regions that contain high proportions of serine, threonine, and proline have been shown to be extensively glycosylated. For example, immunoglobulin class A_1 contains a proline-rich hinge region between the Fab and Fc regions. Of the 21 amino acids within this hinge, there are 10 Pro, 5 Ser, and 4 Thr; and the 5 Ser are all O-glycosylated (18). Several core O-linked oligosaccharides have been described attached to serine or threonine, which can be elongated by addition of further Gal, GlcNAc, Fuc, and sialic acid residues.

Glycophosphatidylinositol Residues

Carbohydrates are also components of the glycophosphatidylinositol (GPI) anchor used to secure certain proteins to cell membranes (19). All have the common core structure ethanolamine-PO_4-Man-Man-Man-$GlcNH_2$-myoinositol-PO_4-lipid. Ethanolamine forms a bridge to the membrane protein, and the inositol phospholipid is inserted onto the lipid bilayer of the cell's plasma membrane. There are certain protein-specific and species-specific modifications to this basic structure (for further information see Ref. 19).

INFLUENCES ON THE GLYCOSYLATION PROCESS

Ultimately, the process of glycosylation is controlled by the activity of glycosyltransferases and glycosidases, the ac-

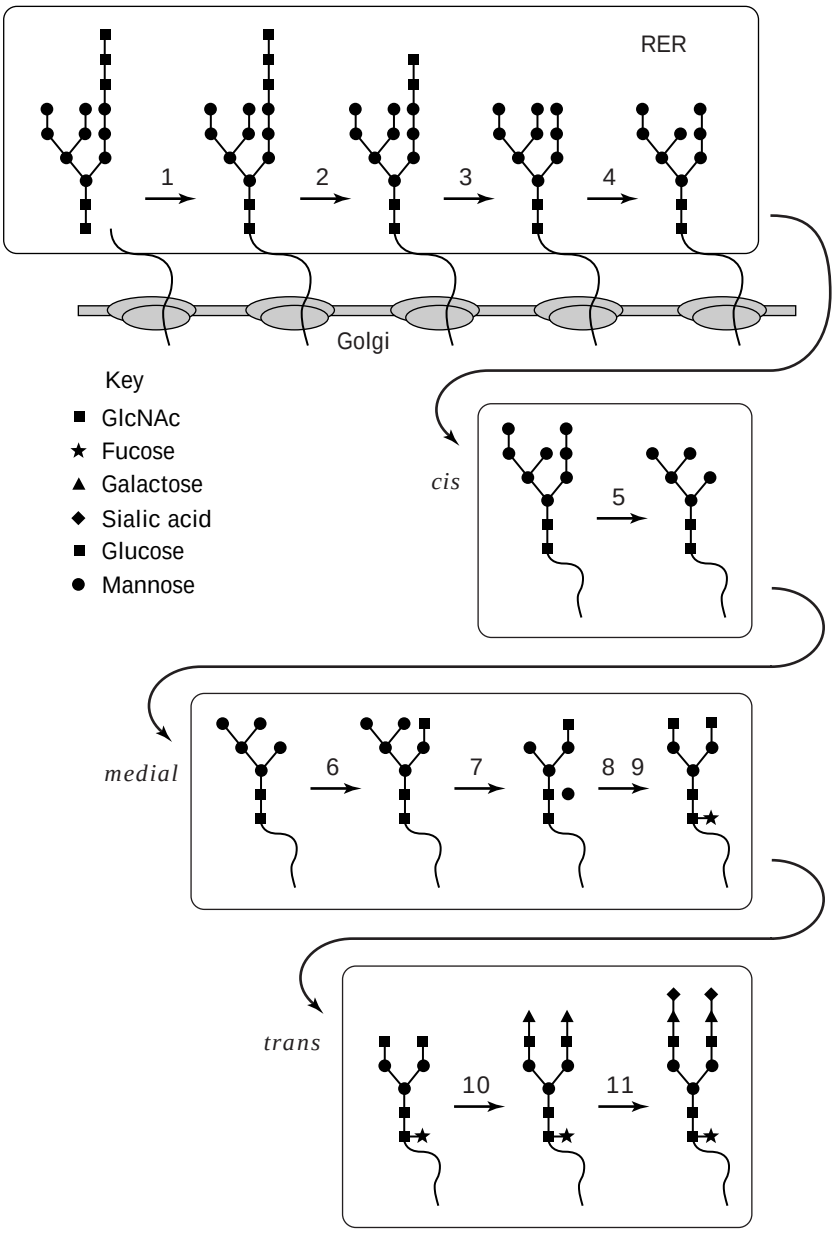


Figure 2. Steps in the N-glycosylation pathway. RER, rough endoplasmic reticulum.

cessibility of their glycan substrates on the glycoprotein, and the availability of cofactors and nucleotide sugar donors (e.g., CMP-NeuAc, GDP-Man, and UDP-GlcNAc). At the cellular level, these control points are influenced by the secondary and tertiary structure of the protein in the ER and Golgi, and the host cell type chosen for the recombinant protein expression and its physiological status.

As more recombinant glycoproteins are used therapeutically and for vaccines, the importance of correct glycosylation has become evident. The availability of alternative expression systems for producing recombinant glycoproteins, together with advances in analytical methods for glycans, has enabled comparisons to be made between glycoproteins from different sources. Crucially, the authenticity of recombinant glycoproteins compared to their natural human counterparts can be assessed, and choices made, concerning the most appropriate expression system.

Host Cell Line

Microbial cells are often the first choice for recombinant protein expression because of their high yields and simple media components. However, common bacterial expression systems such as *Escherichia coli* are incapable of expressing large proteins with authentic mammalian-type glycosylation. Hypermannosylation (the addition of a large number of mannose residues to the core oligosaccharide) is a common property of most yeast strains and can compromise the efficacy of recombinant protein vaccines (20). There is also some evidence to suggest that different O-glycosylation sites are used by yeast and mammalian cells (21).

The few studies reporting the production of therapeutic proteins in plants have suggested smaller *N*-glycan structures are present compared with mammals and lack outer residues such as sialic acids, which may compromise activity. Another obstacle may be the presence of potentially allergenic residues such as core α1-3-linked fucose and xylose (22) and substantial activity of endoglycosidases causing glycan cleavage.

The baculovirus-infected insect cell expression system has been used to express complex proteins and its short production time gives it an advantage over stable mammalian expression systems. Several cell lines have been developed using this technique, including *Spodoptera frugiperda*–derived cell lines Sf9 and Sf21 and *Estigmena acrea*–derived cell line Ea4. Some studies have shown that this system can attach complex oligosaccharides (23) using Sf9 cells, although this result is controversial because other researchers have shown that only simple oligomannose structures can be observed with Sf9. Other insect cell lines differ in their glycosylation capacity; for example, the Ea4 and *Trichoplusia ni* (TN-368) cell lines have been shown to add some complex antennary (GlcNAc and Gal) residues to glycoproteins (24,25).

Although most mammalian cell lines studied possess the intracellular machinery to perform complex N- and O-linked glycosylation to produce proteins that are not dissimilar to those of humans, there are some notable exceptions. Mouse and pig cells have a tendency to add NeuGc instead of NeuAc as the terminal sialic acid of glycoproteins. In contrast, NeuGc is absent from adult humans (due to inactivation of the CMP-NeuAc hydroxylase gene), and the NeuGc residue can be immunogenic (26). Rodents and pigs, like most mammals, express the enzyme α1-3-galactosyltransferase, which generates Galα1,3-Galβ1,4-GlcNAc residues on membranes and secreted glycoproteins. In humans, apes, and Old World monkeys the gene has become inactivated through frameshift mutations. For further details of specific glycosylation machinery in different cell types the reader is referred to another review (2).

Relatively few studies have been reported on the glycosylation of recombinant proteins expressed in the milk of transgenic animals. Those published from experiments in transgenic goats (27) and mice suggest that a lower level of sialylation and higher fucosylation may be achieved. It is now feasible to remodel glycoproteins in situ by the transgenic expression of extra glycosyltransferases in the mouse mammary gland (28). The use of human host cell lines poses a significant risk of retroviral transmission and is not generally favored by the regulatory authorities. Additionally, the transformation event required in most cases to produce a stable cell line may itself result in altered glycosylation profiles (29).

Cell Culture Conditions

It is well documented that the cell culture environment can influence both the macroheterogeneity and microheterogeneity of oligosaccharides in recombinant glycoproteins by affecting both the glycan biosynthetic process and the secretion or release of degradative enzymes, such as exoglycosidases (30). The process of adaptation from serum-containing to serum-free medium may lead to more heterogeneous glycans production in recombinant BHK-21 cells (31) and may improve galactosylation in recombinant CHO cells (14).

The ambient glucose concentration affects the amount of glycosylation of monoclonal antibody produced by human hybridomas in batch culture (32), and of IFN-γ produced by CHO cells in continuous culture (33). However, both lipids and nucleotides are also required for glycan biosynthesis. Lipid supplements alone or in combination with lipoprotein carriers can improve the N-glycosylation site occupancy of IFN-γ (34), and provision of cytidine and uridine can also alter protein glycosylation capacity of hepatocytes by increasing the availability of nucleotide sugars (35). Substances that are commonly used to increase productivity, such as sodium butyrate, should be checked for their influence on glycosylation (36).

Over time, an increase in the percentage of oligomannose (predominantly $Man_5GlcNAc_2$) and truncated structures has been observed in batch and fed-batch cultures of NSO myelomas producing recombinant IgG (37) and in interferon-γ made by recombinant CHO cells (38). Perfusion or chemostat culture of BHK-21 and CHO cells leads to a more consistent glycosylation pattern over time, and perturbations of nutrient availability (glucose or amino acids) can lead to changes in macroheterogeneity (33). Mild hypoxia has minimal effects on the glycosylation of tPA produced by recombinant CHO cells (39) but influences the level of sialylation of recombinant FSH (40). Similarly, pH changes within the range 6.9 to 8.2 in the cell culture medium do not have a dramatic effect on the glycosylation profile of recombinant placental lactogen expressed in CHO cells, however there is some evidence for underglycosylation outside this range (41). Increases in the concentration of ammonium ion in the culture medium above 2 mM may compromise sialylation (42). Degradative glycosidase enzymes have been detected in CHO cell lysates and culture supernatants, and sialidase is most active at neutral pH (43). This sialidase has been purified from culture supernatant of CHO cells (44) and can degrade glycans from proteins such as recombinant gp120. Although CHO cells also produce an α-L-fucosidase, this enzyme is incapable of releasing core fucose from intact recombinant glycoproteins (45). Sialidase, β-galactosidase, β-hexosaminidase, and fucosidase can also be detected at low levels in supernatants from mouse 293, NSO, and hybridoma cells, but the sialidase activity is much lower than that found in CHO cells (46). For more information, see GLYCOSYLATION OF RECOMBINANT PROTEINS.

EFFECTS OF GLYCOSYLATION

The presence, extent, and type of glycosylation influence the functional and structural properties of glycoproteins. Therefore, data acquired on the structure–function relationships of glycans present on one recombinant glycoprotein cannot always be applied to another. A complete listing of the structure–function relationships of glycans found on glycoproteins is beyond the scope of this article

(see Refs. 6 and 47); however, the main processes on which protein glycosylation has been shown to exert influence are the following:

1. Pharmacokinetics and clearance in vivo (particularly influenced by the extent of sialylation)
2. Immunogenicity, which is carbohydrate residue specific
3. Protein solubility and protease resistance (6,7)

REGULATORY CONSIDERATIONS

Because of the efficacy and safety considerations, both the Food and Drug Administration in the United States and the European Medicines Agency now demand more comprehensive carbohydrate analysis before licensing recombinant glycoproteins. It is important to understand that some degree of heterogeneity exists in most natural as well as recombinant glycoproteins. Therefore, current guidelines accept glycan heterogeneity, but it must be shown to be consistent within prescribed boundaries between lots of the finished product. The inherent glycan heterogeneity and its manipulation by genetic or enzymatic engineering poses interesting challenges to the patenting of glycoprotein drugs, and several patents have already been filed that claim to alter the glycosylation process for biotechnological purposes.

STRATEGIES FOR ANALYSING GLYCOPROTEINS

See also GLYCOSYLATION OF RECOMBINANT PROTEINS. Much of the increased knowledge concerning the influences on the glycosylation pathways and the biological significance of glycan variations has derived from recent advances in complex carbohydrate analysis. Using the appropriate equipment and techniques, it is now possible to unequivocally recognize most glycan structures and deduce their relative proportions in a complex glycoprotein. Current efforts are focused on improving the sensitivity of such techniques and making them more user friendly.

Sample Preparation

The glycoprotein must first be isolated and purified, and it is essential to know if it is secreted or membrane bound. The two classes of glycoproteins have different physicochemical properties; hence, the isolation protocol has to be adjusted for each type, with detergents usually required for membrane-bound proteins. Common precipitation and chromatographic techniques used for protein purification can generally be applied to glycoproteins. Chromatographic techniques tend to be based on charge, size, or affinity, and it is worth noting that of all the common sugar residues found in glycoproteins only the sialic acids are charged (negative). Also, protein A from *Staphylococcus aureus* Cowan I strain or proteins G and C from *Streptococcus*, which are commonly used to purify IgG, are both glycosylation independent. The purity of the glycoprotein can be assessed using SDS-polyacrylamide gel electrophoresis (expect multiple bands for complex glycoproteins and/or oligomers). Its yield can be quantified using protein stains (Bradford or Lowry method), A_{280}, HPLC, or immunoassays. We recommend performing at least two different assays to estimate protein yield because errors at this stage may have disastrous consequences for further analysis.

Preliminary Analysis

The simple SDS-PAGE gel analysis can be extended to yield data on the extent of protein glycosylation. For example, a portion of the sample can be digested with PNGase F (Oxford Glycosciences, U.K.) (to remove any *N*-glycans) and a combination of sialidase and *O*-glycanase (to remove *O*-glycans), and its mass shift compared on gels to that of undigested glycoprotein. Alternatively, kits are available, such as the FACE system from Dextra Laboratories (Reading, U.K.), which use a combination of isoelectric focusing gels and carbohydrate stains. Bear in mind that there are many other causes of protein heterogeneity than glycosylation: modifications such as proteolysis are common and will result in multiple bands. These preliminary tests are useful in estimating the degree and type of glycosylation and will help determine the final glycosylation analysis strategy.

Lectin blots may also be used to detect certain sugar residues or linkages (48). Lectins are carbohydrate-binding proteins that have been used successfully to identify and even isolate some glycoproteins. Lectin blots are a convenient way of identifying glycoproteins run on SDS-PAGE gels and then transferred to nitrocellulose sheets by techniques analogous to Western blotting. The blots are washed, blocked, and then incubated with the desired lectin that is already conjugated to a convenient enzyme (e.g., alkaline phosphatase or peroxidase) or light-generating compound. Control experiments using an appropriate competing saccharide should be run to verify that specific binding is being observed. Even using these controls, there are incidents of false identification, and we are reluctant to rely on these methods entirely. Similarly, colorometric assays are available to approximate the percentage of total sugars in each sample, but these assays are not very sensitive (5–10 nmol).

Glycopeptide Preparation

Glycopeptide preparation is often a prerequisite to establishing the location and structures of individual glycans within a complex glycoprotein. Controlled proteolytic cleavage may be carried out with enzymes such as trypsin and chymotrypsin (49). The exact protease to be chosen for depends on the size of the glycopeptides and its peptide substrate specificity; chemical cleavage, such as with cyanogen bromide, may also be employed. Care should be taken to prevent incomplete or nonspecific digestion because this can compromise the subsequent resolution of individual glycopeptides by reverse-phase HPLC (see later). Before enzyme or chemical cleavage it is desirable to modify any disulfide bonds present in the glycoprotein (50). This can be accomplished by reduction and carboxymethylation in the presence of a denaturant (8 M urea).

Releasing Glycans from Glycoproteins

There are a plethora of chemical and enzymatic methods that have been used to release carbohydrates from glycoproteins, with varying degrees of success (Table 1). The enzymatic methods (such as peptide-*N*-glycosidase F [PNGase F]) (see Table 2) used to cleave the glycosidic linkage between GlcNAc and the asparagine of N-linked glycans) generally preserve intact the peptides and glycans, but sometimes there are concerns over their efficiency (51). Most chemical cleaving methods such as hydrazinolysis (a nonselective method for releasing glycans using anhydrous hydrazine to cleave both N- and O-glycosidic bonds [Fig. 3]), tend to destroy the peptide at the glycosylation site(s). There is also a risk of further breakdown to monosaccharides and their degradation products; therefore the reaction must be carefully controlled (52). A clean-up step is often required before progression to further analysis. Following these preliminary steps (not all may be required), the analysis should progress to a combination of one or more of the techniques listed next. The strategy routinely used by us is described in Figure 4.

Gas–Liquid Chromatography

Gas–liquid chromatography GLC methods are highly efficient for the analysis of carbohydrates released from glycoproteins, especially volatile monosaccharide derivatives. A mass spectrometer can be connected to the GLC and used to characterize these resolved peaks, in comparison with known standards. Identification of each peak is achieved by the combination of its mass and the relative retention times (53). However, these GLC methods have largely been superseded for more complex oligosaccharides by HPLC alternatives, due to increased sensitivity, which enables smaller sample sizes to be analyzed.

High-Performance Liquid Chromatography

Analysis of glycopeptides or carbohydrate chains released from glycoproteins often requires high-resolution methods such as HPLC. The most common mode of HPLC for glycopeptide purification is reverse-phase HPLC (RP-HPLC), which separates on the basis of peptide, glycopeptide, or glycan hydrophobicity. Other modes employed are normal-phase partition and ligand exchange (49). RP-HPLC requires the presence of hydrophobic groups such as N-Ac-, deoxy-, -OMe, or OAc groups for longer retention time and better separation. Oligomannose- or high-mannose-type oligosaccharides (which have only two N-Ac groups) are retained on the RP-HPLC column and are not well resolved compared with complex-type oligosaccharides. Glycopeptides, released sugars (chemically or enzymically), or derivatized glycans can be resolved by HPLC and can be analyzed on a mass spectrometer. A common use for this technique is the separation of glycopeptides corresponding to individual glycosylation sites within a glycoprotein, following controlled proteolysis (50). The resultant glycopeptides are then analyzed by mass spectrometry and exoglycosidase arrays.

Table 1. Glycosidase Enzymes Used for Selective Cleavage of Oligosaccharides for Structural Analysis

Methods	Comments
Mild acid hydrolysis	Quantitative release of terminal nonreducing fucose
Strong acid hydrolysis	Quantitative release of hexose, deoxyribose, pentoses, and hexosamines
Methanolysis	Quantitative release of neutral monosaccharides and hexosamines
Hydrazinolysis	Nonselective release of glycans by cleaving both N-linked and O-linked bonds; can be optimized for specific N-linked or O-linked cleaving
Smith degradation	Quantitative release of monosaccharides
Acetolysis	Cleaves 1-6-linkages in oligosaccharides

High-Performance Anion Exchange Chromatography with Pulsed Amperometric Detection

Another procedure for the direct analysis of carbohydrates uses the high-performance anion-exchange mode (utilizing pellicular resin columns) of HPLC with pulsed amperometric detection (HPAE-PAD, pioneered by Dionex [54]). This is suitable for the analysis of released N- and O-linked glycans in quantities >0.2 nmol (55). Oligomannose-type glycans are eluted at low alkali concentrations, and an increase in the content of neutral sugar residues results in higher retention times. The negatively charged glycans that contain sialic acids are retained most. The method can also be used for analyzing released monosaccharides, such as resolving different forms of sialic acids, with a detection limit of around 20 pmol.

Capillary Electrophoresis

Capillary electrophoresis (CE) is an emerging technology that may be used for direct analysis of oligosaccharide derivatives (56). The main advantages of CE are its simplicity and high resolving power. There are a number of modes of operation that have been developed for CE, for example, free-zone capillary electrophoresis (CZE), micellar electrokinetic capillary chromatography (MECC) (57), affinity electrophoresis, gel electrophoresis, isoelectric focusing (IEF), and isotacophoresis. The method can be used for resolving released glycans, glycopeptides, or glycolipids (58), and the sensitivity of CE can be improved by employing a laser-induced fluorescence detector (59).

Mass Spectrometry

Mass spectrometry has proved to be an invaluable tool for glycosylation analysis. There are a number formats that can be used, such as electrospray ionisation mass spectrometry (ESI-MS) (60), liquid chromatography mass spectrometry (LC-MS) (61), and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) (49,50) (Fig. 4). In MALDI-TOF, which is more tolerant to common salts, a variety of matrices have been developed to maximize ionization of glycopeptides or released glycans (62).

Table 2. Glycosidase Enzymes Used for Selective Cleavage of Oligosaccharides for Structural Analysis

Enzyme	Source	Specificity	pH
PNGase F	*Flavobacterium Meningosepticum*	N-Glycosidic linkage between GlcNAc and asparagine	7.0
Endoglycosidase H	*Streptomyces plicatus*	Oligomannose and hybrid-type *N*-glycans within the chitobiose core	5.0
α-L-Fucosidase	*Aspergillus niger*	Fucose α1-2-galactose	3.8
	Bacillus fulminans	Fucose α1-2-galactose	6.6
	Clostridium perfringens	Fucose α1-2-galactose	6.0
	Corynebacterium	Fucose α1-2-galactose	8.5
	Almond emulsin	Fucose α1-3 and 4-GlcNAc	5.5
	Bovine epididimis	Nonspecific	6.5
	Charonia lampas	Nonspecific	3.3
	Turbo cornutus	Nonspecific	4.0
α-D-Galactosidase	*Aspergillus niger*	Galα1-3Gal	4.5
	Coffee bean	Nonspecific	6.0
	Fig	Nonspecific	6.0
β-D-Galactosidase	Jack bean	Galβ1-6GlcNAc > Galβ1-4GlcNac > Galβ1-3GlcNAc	3.5
	Streptococcal 6646K	Galβ1-4GlcNAc > Galβ1-3GlcNAc	5.5
	Aspergillus niger	Galβ1-4GlcNAc > Galβ1-6GlcNac > Galβ1-3GlcNAc	4.0
α-D-Mannosidase	Jack bean	Manα1-2 > Manα1-6 > Manα1-3	4.0
	Aspergillus niger	Manα1-2 and Manα1-4 > Manα1-6 > Manα1-3	4.5
	Aspergillus saoti	Manα1-2	4.5 5.0
β-D-Mannosidase	*Helix pomitia*	$ManGlcNAc_2Asn$	4.5
β-*N*-Acetylgalactosaminidase	*Aspergillus niger*	GalNAc (terminal)-Ser	3.5
β-*N*-Acetylhexosaminidase	Jack bean	GlcNAcβ1-2,3,4,6	6.0
	Aspergillus niger	GlcNAβ1-3,4	5.0

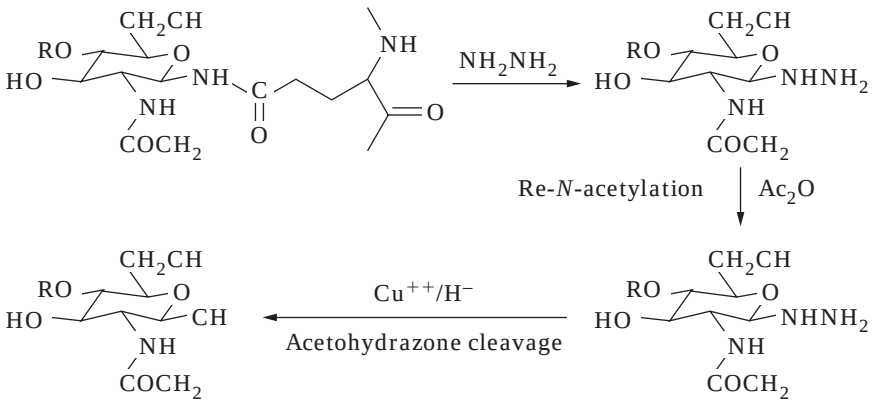


Figure 3. Hydrazinolysis reaction used to release glycans from glycoproteins.

Exoglycosidase Arrays

One of the most powerful methods for the elucidation of glycan structure uses an array of exoglycosidases (Table 2). These enzymes recognize predominantly the terminal nonreducing monosaccharide residues of glycans, glycopeptides, or whole glycoproteins and can be used to determine the anomeric configuration and glycosidic linkage of this terminal. Time can be saved by constructing a parallel array of these enzymes in combination (63), and the resultant mixtures can be analyzed by mass spectrometry (30,49) or by HPAE (64). Here, differences in mass or resolution time corresponding to the loss of specific sugar residues are determined.

Sialic Acid Analysis

Because of the importance of sialic acid in influencing half-life and protein stability, sialic acid analysis acquires a special significance, particularly for biopharmaceuticals destined for human therapy. Sialic acid is a collective term used to describe several derivatives of neuraminic acid. The neuraminic acid occurring in glycoproteins is usually substituted with an *N*-acetyl or *N*-glycolyl group in a species-specific fashion (2). Sialic acids can also be derivatized with *O*-acetyl, *O*-lactyl, or *O*-methyl groups, but these are rarer in natural glycoproteins. Each type of sialic acid has a characteristic mass/charge ratio that is recognizable by mass spectrometry; however, for detailed quan-

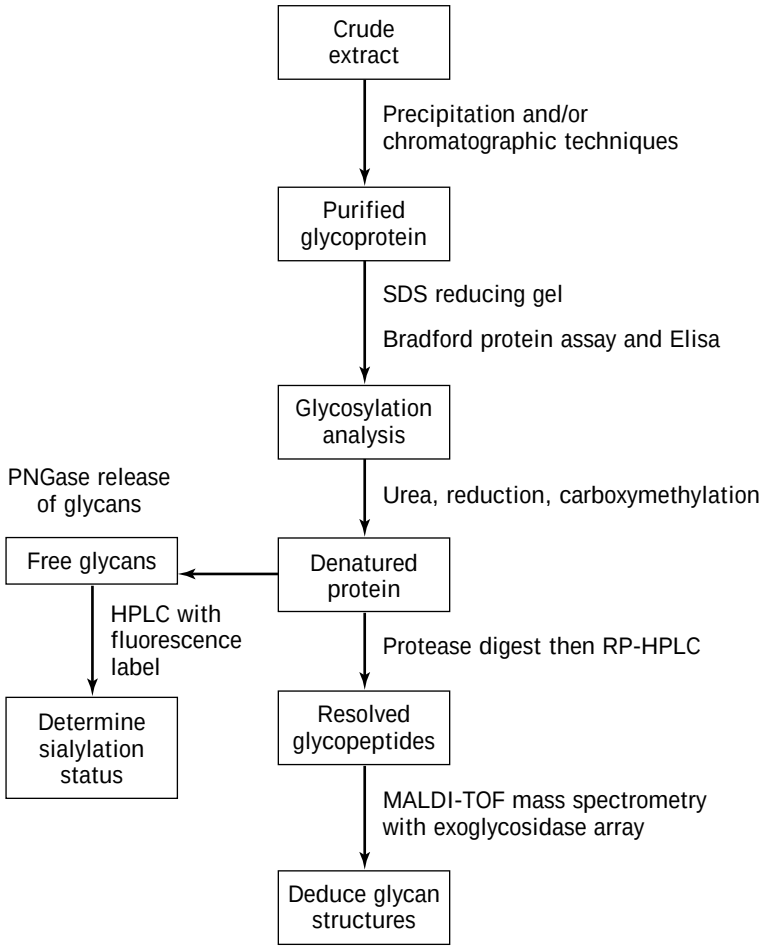


Figure 4. A strategy for glycan analysis at each glycosylation site.

tification it is often necessary to release the sialic acids from the glycoprotein by acid hydrolysis or by linkage-specific neuraminidases (65). Released sialic acids can then be analyzed by HPLC or HPAE (see earlier).

CONCLUSIONS

Glycobiology has advanced dramatically as people become aware of the biological implications of carbohydrate heterogeneity. Over the past five years, there has been a move away from simple monosaccharide composition analysis or gel estimates of glycan heterogeneity toward more sophisticated forms of analyses using techniques such as CE, mass spectroscopy, fluorescence HPLC, and HPAE. Regulatory authorities such as the U.S. Food and Drug Administration are demanding increasingly sophisticated carbohydrate analyses as part of the product or process validation (66). In addition to these analytical trends, there is increasing awareness of the relevance of glycosylation in producing a product with a good therapeutic profile and a high degree of batch consistency.

ACKNOWLEDGMENTS

The authors have been supported by the Biotechnology and Biological Research Council of the United Kingdom, and the European Union Biotechnology Framework Programme (DGXII).

BIBLIOGRAPHY

1. C. Abeijon and C.B. Hirschberg, *Trends Biochem. Sci.* **17**, 32–36 (1992).
2. N. Jenkins, R.B. Parekh, and D.C. James, *Nat. Biotechnol.* **14**, 975–981 (1996).
3. C.G. Gahmberg and M. Tolvanen, *Trends Biochem. Sci.* **21**, 308–311 (1996).
4. B.K. Hayes and G.W. Hart, *Curr. Opin. Struct. Biol.* **4**, 692–696 (1994).
5. M.C. Field and L.J. Wainwright, *Glycobiology* **5**, 463–472 (1995).
6. A. Varki, *Glycobiology* **3**, 97–130 (1993).
7. N. Jenkins and E.M. Curling, *Enzyme Microb. Technol.* **16**, 354–364 (1994).
8. R.J. Harris, V.T. Ling, and M.W. Spellman, *J. Biol. Chem.* **267**, 5102–5107 (1992).
9. S.B. Yan, Y.B. Chao, and H. Vanhalbeek, *Glycobiology* **3**, 597–608 (1993).
10. S. Allen and N.J. Bulleid, *Biochem. J.* **328**, 113–119 (1997).
11. R. Wilson, A.J. Allen, J. Oliver, J.L. Brookman, S. High, and N.J. Bulleid, *Biochem. J.* **307**, 679–687 (1995).
12. E.M. Curling, P.M. Hayter, A.J. Baines, A.T. Bull, K. Gull, P.G. Strange, and N. Jenkins, *Biochem. J.* **272**, 333–337 (1990).

13. T.A. Hsu, N. Takahashi, Y. Tsukamoto, K. Kato, I. Shimada, K. Masuda, E.M. Whiteley, J.Q. Fan, Y.C. Lee, and M.J. Betenbaugh, *J. Biol. Chem.* **272**, 9062–9070 (1997).
14. M.R. Lifely, C. Hale, S. Boyce, M.J. Keen, and J. Phillips, *Glycobiology* **5**, 813–822 (1995).
15. A.A. Bergwerff, J. Vanoostrum, J.P. Kamerling, and J.F.G. Vliegenthart, *Eur. J. Biochem.* **228**, 1009–1019 (1995).
16. P.L. Smith and J.U. Baenziger, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 329–333 (1992).
17. E.F. Hounsell, M.J. Davies, and D.V. Renouf, *Glycoconjugate J.* **13**, 19–26 (1996).
18. A. Allen and J. Feehally, *Adv. Exp. Med. Biol.* **435**, 175–183 (1998).
19. J. Takeda and T. Kinoshita, *Trends Biochem. Sci.* **20**, 367–371 (1995).
20. L. Lehle, A. Eiden, K. Lehnert, A. Haselbeck, and E. Kopetzki, *FEBS Lett.* **370**, 41–45 (1995).
21. A.O. Herscovics and P. Orlean, *FASEB J.* **7**, 540–550 (1993).
22. G. Garcia-Casado, R. Sanchezmonge, M.J. Chrispeels, A. Armentia, G. Salcedo, and L. Gomez, *Glycobiology* **6**, 471–477 (1996).
23. F.J. Castellino and D.J.C. Davidson, *Am. Chem. Soc. Symp. Ser.* **202**, 20–20 (1991).
24. O.W. Ogonah, R.B. Freedman, N. Jenkins, K. Patel, and B.C. Rooney, *Biotechnology* **14**, 197–202 (1996).
25. T.R. Davis and H.A. Wood, *In Vitro Cell. Dev. Biol.* **31**, 659–663 (1995).
26. A. Noguchi, C.J. Mukuria, E. Suzuki, and M. Naiki, *J. Biochem.* **117**, 59–62 (1995).
27. E.S. Cole, E. Higgins, R. Bernasconi, L. Garone, and T. Edmunds, *J. Cell Biochem.* **265**, S18D (1994).
28. P.A. Prieto, P. Mukerji, B. Kelder, R. Erney, D. Gonzalez, J.S. Yun, D.F. Smith, K.W. Moremen, C. Nardelli, M. Pierce, Y.S. Li, X. Chen, T.E. Wagner, R.D. Cummings, and J.J. Kopchick, *J. Biol. Chem.* **270**, 29515–29519 (1995).
29. K. Yamashita, N. Koide, T. Endo, Y. Iwaki, and A. Kobata, *J. Biol. Chem.* **264**, 2415–2423 (1989).
30. D.C. James, R.B. Freedman, M. Hoare, O.W. Ogonah, B.C. Rooney, O.A. Larionov, V.N. Dobrovolsky, O.V. Lagutin, and N. Jenkins, *Biotechnology* **13**, 592–596 (1995).
31. M. Gawlitzek, U. Valley, M. Nimtz, R. Wagner, and H.S. Conradt, *J. Biotechnol.* **42**, 117–131 (1995).
32. H. Tachibana, K. Taniguchi, Y. Ushio, K. Teruya, K. Osada, and H. Murakami, *Cytotechnology* **16**, 151–157 (1994).
33. P.M. Hayter, E.M. Curling, A.J. Baines, N. Jenkins, I. Salmon, P.G. Strange, J.M. Tong, and A.T. Bull, *Biotechnol. Bioeng.* **39**, 327–335 (1992).
34. N. Jenkins, P.M.L. Castro, S. Menon, A.P. Ison, and A.T. Bull, *Cytotechnology* **15**, 209–215 (1994).
35. W.R.P. Rijcken, B. Overdijk, D.H. Vandeneijnden, and W. Ferwerda, *Biochem. J.* **305**, 865–870 (1995).
36. C.A. Gebert and P.P. Gray, *Cytotechnology* **17**, 13–19 (1995).
37. C.C. Yu-Ip, W.J. Miller, M. Silberklang, G.E. Mark, R.W. Ellis, L.H. Huang, J. Glushka, H. van Halbeek, J. Zhu, and J.A. Alhadeff, *Arch. Biochem. Biophys.* **308**, 387–399 (1994).
38. A.D. Hooker, M.H. Goldman, N.H. Markham, D.C. James, A.P. Ison, A.T. Bull, P.G. Strange, I. Salmon, A.J. Baines, and N. Jenkins, *Biotechnol. Bioeng.* **48**, 639–648 (1995).
39. A.A. Lin, R. Kimura, and W.M. Miller, *Biotechnol. Bioeng.* **42**, 339–350 (1993).
40. W. Chotigeat, Y. Watanapokasin, S. Mahler, and P.P. Gray, *Cytotechnology* **15**, 217–221 (1994).
41. M.C. Borys, D.J.H. Linzer, and E.T. Papoutsakis, *Bio/Technology* **11**, 720–724 (1993).
42. D.C. Andersen, C.F. Goochee, G. Cooper, and M. Weitzhandler, *Glycobiology* **4**, 459–467 (1994).
43. M.J. Gramer, C.F. Goochee, V. Chock, D.T. Brousseau, and M.B. Sliwkowski, *Bio/Technology* **13**, 692–698 (1995).
44. T.G. Warner, J. Chang, J. Ferrari, R. Harris, T. Mcnerney, G. Bennett, J. Burnier, and M.B. Sliwkowski, *Glycobiology* **3**, 455–463 (1993).
45. M.J. Gramer, D.V. Schaffer, M.B. Sliwkowski, and C.F. Goochee, *Glycobiology* **4**, 611–616 (1994).
46. M.J. Gramer and C.F. Goochee, *Biotechnol. Bioeng.* **43**, 423–428 (1994).
47. A. Varki, *FASEB J.* **11**, 248–255 (1997).
48. R.D. Cummings, *Methods Enzymol.* **230**, 66–86 (1994).
49. C.W. Sutton, J.A. O'Neill, and J.S. Cottrell, *Anal. Biochem.* **218**, 34–46 (1994).
50. D.C. James, M.H. Goldman, M. Hoare, N. Jenkins, R.W.A. Oliver, B.N. Green, and R.B. Freedman, *Protein Sci.* **5**, 331–340 (1996).
51. K. Yamamoto, *J. Biochem.* **116**, 229–235 (1994).
52. G.D. Roberts, W.P. Johnson, S. Burman, K.R. Anumula, and S.A. Carr, *Anal. Chem.* **67**, 3613–3625 (1995).
53. J.H.G. Mutsaers, J.P. Kamerling, R. Devos, Y. Guisez, W. Friers, and J.F.G. Vliegenthart, *Eur. J. Biochem.* **156**, 651–654 (1986).
54. M. Weitzhandler, C. Pohl, J. Rohrer, L. Narayanan, R. Slingsby, and N. Avdalovic, *Anal. Biochem.* **241**, 128–134 (1996).
55. P. Hermentin, R. Witzel, R. Doenges, R. Bauer, H. Haupt, T. Patel, R.B. Parekh, and D. Brazel, *Anal. Biochem.* **206**, 419–429 (1992).
56. P.L. Weber, T. Kornfelt, N.K. Klausen, and S.M. Lunte, *Anal. Biochem.* **225**, 135–142 (1997).
57. D.C. James, R.B. Freedman, M. Hoare, and N. Jenkins, *Anal. Biochem.* **222**, 315–322 (1994).
58. R.S. Rush, H.J. Boss, V. Katta, and M.F. Rohde, *Electrophoresis* **18**, 751–756 (1997).
59. X.C. Le, Y. Zhang, N.J. Dovichi, C.A. Compston, M.M. Palcic, R.J. Beever, and O. Hindsgaul, *J. Chromatogr. A* **781**, 515–522 (1997).
60. G.J. Rademaker, S.A. Pergantis, L. BlokTip, J.I. Langridge, A. Kleen, and J.E. Thomasoates, *Anal. Biochem.* **257**, 149–160 (1998).
61. P. Camilleri, D. Tolson, and H. Birrell, *Rapid Commun. Mass Spectrom.* **12**, 144–148 (1998).
62. D.I. Papac, A. Wong, and A.J.S. Jones, *Anal. Chem.* **68**, 3215–3223 (1996).
63. C.J. Edge, R.B. Parekh, T.W. Rademacher, M. Wormald, and R.A. Dwek, *Nature* **358**, 693–694 (1992).
64. F.H. Routier, M.J. Davies, K. Bergemann, and E.F. Hounsell, *Glycoconjugate J.* **14**, 201–207 (1997).
65. L. Monaco, A. Marc, A. EonDuval, G. Acerbis, G. Distefano, D. Lamotte, J.M. Engasser, M. Soria, and N. Jenkins, *Cytotechnology* **22**, 197–203 (1996).
66. D.T.Y. Liu, *Trends Biotechnol.* **10**, 114–120 (1992).

See also GLYCOSYLATION OF RECOMBINANT PROTEINS.

PROTEIN PURIFICATION, AQUEOUS LIQUID EXTRACTION

MARIA-REGINA KULA
KLAUS SELBER
Heinrich Heine University Düsseldorf
Jülich, Germany

KEY WORDS

Affinity extraction
Aqueous two-phase systems
Detergent-based systems
Large-scale
Protein extraction
Protein purification
Recycling
Reverse micelles
Thermodynamics
Thermoseparating polymers

OUTLINE

INTRODUCTION

For the large-scale production of proteins, complex separation procedures are necessary after protein synthesis by living cells is completed. In comparison to downstream processing in the chemical industry, the separation of proteins must be very gentle to conserve the native structure, which means that in general near-neutral pH values and moderate temperatures ($\leq$30 °C) must be maintained. Important objectives in separation are high selectivity and yield because multistep procedures are usually required. Because of the polar character and high molecular mass of proteins, several unit operations common in the chemical industries cannot be applied, for example, distillation or extraction into organic solvents. Aqueous two-phase systems, however, can be used with advantage for protein isolation by liquid–liquid extraction. The technique is based on the incompatibility of two different hydrophilic polymers in the common solvent water or the reversible hydration of oxygen in polyethers.

In an aqueous liquid–liquid extraction process, cells and cell debris and other unwanted substances should remain in one phase while the desired protein is partitioned into the other phase, preferentially the top phase. The phases are separated and further processed until the desired level of product purity is reached. The extraction of proteins is performed preferentially at early stages of an isolation process using cell homogenates or whole fermentation broth because this allows the concomitant removal of solids, a product enrichment and concentration in a single step.

Partitioning of proteins by aqueous liquid–liquid extraction has been well described (1–6) and finds increasing application for large-scale processes in industry. For further reading the monographs by Albertsson (4) and especially by Walter et al. (5) are recommended. The latter gives detailed information about handling principles and applications. The Aqueous Two-Phase (a.k.a. Biphasic) Systems Community home page has a bibliography compiled by Harry Walter that contains more than 1,000 files of journal publications on aqueous two-phase systems published between 1956 and 1996: *http://radar.ch.ua.edu/~aq2phase/*

BIOCHEMICAL FUNDAMENTALS

Miscibility gaps exist in different aqueous systems; these are based on polymer–polymer interactions, polymer–salt interactions, or thermoseparation. A special case is affinity extraction, which can be viewed as a reactive extraction process. For further descriptions see “Alternative Two-Phase Systems”.

Principles of Aqueous Two-Phase Extraction

Most of the hydrophilic, natural, and synthetic polymers are miscible with water and show phase separation in a mixture with another polymer or salt. The phase diagram of polyethylene glycol (PEG) and potassium phosphate in Figure 1 demonstrates the miscibility of polymer and salt phases. Above the curve (binodal) the system separates into two liquid phases. If PEG and salt are mixed in a percentage represented by a point *M* above the binodal, the mixture separates along the tie-line. The composition of the two resulting phases are represented by the cross-points of the tie-line and the binodal. The volume ratio of the PEG-rich phase to the salt-rich phase is given approximately by the ratio of the tie-line sections *MB* and *TM*. The critical point is reached when the tie-line is represented by only one point on the binodal. At the critical point, the addition of an infinitesimally small amount of water results in two phases of equal volume and composition; therefore, any added component should have a partition coefficient *K* of 1, where *K* is defined as the ratio of concentration in the top ($c_{i,\mathrm{T}}$) and bottom phase ($c_{i,\mathrm{B}}$)

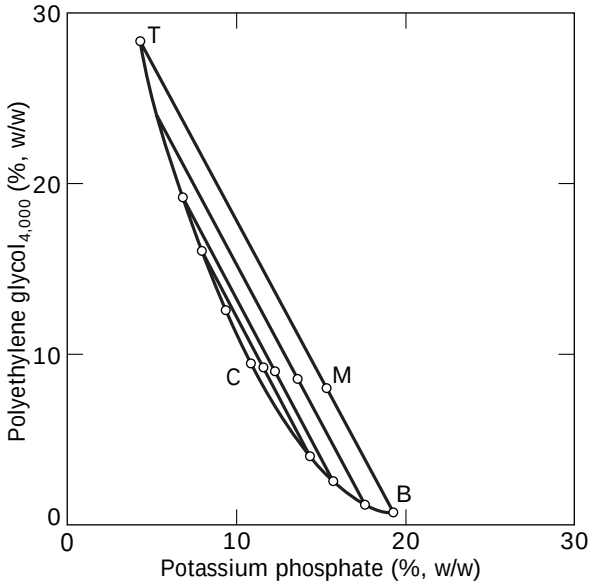


Figure 1. Phase diagram of PEG 4,000–potassium phosphate at 20 °C. *Source*: Ref. 4.

$$K = \frac{c_{i,T}}{c_{i,B}} \tag{1}$$

With increasing length of the tie-line the composition of the phases at equilibrium diverge and consequently K changes.

It must be noted that phase diagrams are altered by addition of other components in sufficient concentration such as cell homogenates and different salts (7,8). Nevertheless, the phase diagram is useful for the selection of phase systems. Many phase diagrams of polymer–polymer and polymer–salt systems can be found in Refs. 4 and 6. An extraction system is usually designed such that the wanted protein partitions predominantly to the top phase (PEG, in this case) and the undesired material including cells and cell debris partitions to the bottom phase (salt). The separation power of a system is characterized by the yield Y, the separation factor G, depending on the volume ratio R, and K. The yield in the top phase is defined as

$$Y_T = \frac{n_T}{n_T + n_B} \tag{2}$$

or consequently as

$$Y_T = \frac{1}{1 + \left[\frac{V_B}{V_T} \cdot \frac{1}{K}\right]} \tag{3}$$

where n is the molar quantity of the desired protein in the volume of the top (V_T) and bottom phase (V_B), respectively. G is defined as the product of the partition coefficient K and the volume ratio R. The definition of the volume ratio is given in equation 5.

$$G = K \cdot R \tag{4}$$

$$R = \frac{V_T}{V_B} \tag{5}$$

Y, K, and R affect each other. Usually Y_T increases if K is increased at a constant R or if R is increased at a constant K (Figure 2).

K does not depend on the concentration of the protein i, so long as the concentration c_i is much smaller than the concentration of the phase-forming polymers, and no complex formation takes place. According to Bronstedt (9), there exists an exponential relation between K and the surface area (molecular mass) of i:

$$\ln K_i = \frac{\alpha \cdot A_i}{k \cdot T} \tag{6}$$

α is a factor related to the chemical potential, A_i is the surface for interaction, k is the Boltzmann constant, and T is the absolute temperature. Thus, cells and cell debris should exhibit one-sided partition approaching zero or infinity even at low values of α because of their large value of A, while metabolites should have K values approximately 1 and proteins usually between 0.1 and 10. The relation between $\ln K$ and A has been capitalized in a process to isolate hepatitis B surface antigen particles (molecular weight, $\simeq$ 1 to 2 million Da) as a hepatitis B vaccine from recombinant yeast cultures (10). Equation 1 is valid for the equilibrium between two liquid phases only; however, sometimes a solid phase seems to be involved and rather large values of K may result as a consequence of preferential solubility of i under extraction conditions in the top phase of PEG–salt systems (11). In such a case, K will increase with c_i.

Factors Determining Partitioning

The forces that cause partitioning can be explained in a first approximation by splitting the partition coefficient into increments:

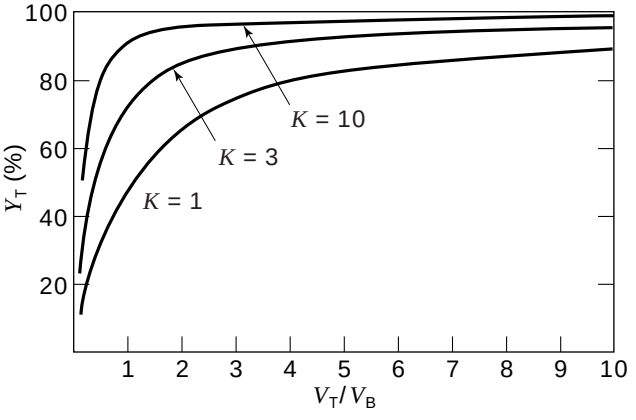


Figure 2. Interdependence of yield (Y), partition coefficient (K), and volume ratio (R). *Source*: Ref. 8.

$$\ln K = \ln K_{ion} + \ln K_{hphob} + \ln K_{hphil} + \ln K_{conf} + \ln K_{lig} \quad (7)$$

where K_{ij} denotes the partition coefficient from charges or hydrophobic or hydrophilic forces or its dependence on conformation or ligand interaction, respectively (12). Only $\ln K_{lig}$ is to some extent predictable. For this reason, affinity extraction is much studied, because the affinity interaction can be made to dominate ln *K*. Even in the absence of affinity interactions many parameters will influence the partitioning of proteins in aqueous two-phase systems, affecting the partition coefficient and the selectivity.

These factors are as follows:

- Selection of component(s) to form a two-phase system
- Concentration of polymers used
- Molecular weight of the polymers
- Type and concentration of ions
- pH
- Temperature
- Addition of other agents (e.g., chaotropic agents) (13,14)
- Concentration of cells and cell debris added

The strong influence on the partition of only small changes in the system composition is shown in Figure 3, where the partition coefficient for L-2-hydroxyisocaproate dehydrogenase is changed significantly by addition of potassium phosphate and the resulting change of the tie-line.

The molecular mass of a given polymer can strongly influence the partition of a protein. The partition coefficient can be increased by increasing the molecular mass of the bottom phase forming polymer (for example, dextran 1) or by lowering the average molecular mass of the top-phase-forming polymer, most often PEG (1,4). The latter represents a fast and convenient way to increase protein yield in the top phase and is aided by the fact that the polymerization of ethyleneoxid can be easily adjusted to yield different molecular weight fractions of PEG. Several fractions are commercially available and are included in the pharmacopoeia of many countries. Changes in pH affect the ionization of side chains of the protein and therefore the charge and polar/hydrophobic balance of the surface; it will also alter the ionization of some weak and multivalent acids present in the extraction mixture leading to changes in *K*. pH values are easily manipulated in the range of protein stability. The addition of ions has a remarkable effect on *K* (1,4) by altering the length of the tie-line and the solvating power in PEG–salt systems as well as changes in the electrochemical potential across the interface in polymer–polymer systems.

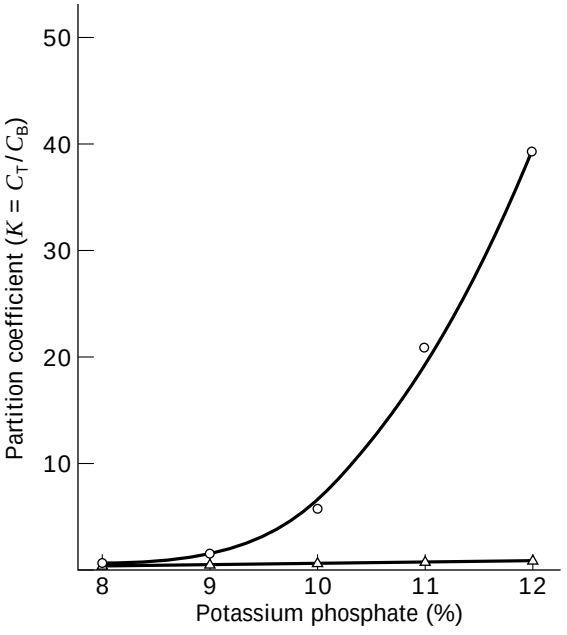


Figure 3. Partition of L-2-hydroxyisocaproate dehydrogenase (○) and D-lactate dehydrogenase (△) as a function of potassium phosphate concentration. *Source*: Ref. 15.

The search for a separating system of a new protein is usually carried out by trial and error guided by the heuristic rules discussed previously. A series of phase systems differing in one parameter are produced in the scale of 5.0–10.0 g using a statistical design. A nice example of a statistical approach is shown by Hart et al. for the isolation of nonnative IGF-I (16). *Y*, *K*, and *R* are determined as well as protein concentration. From this information, a useful extraction system can be derived that can be further optimized with regard to specific goals, such as product yield and quality. Other criteria for optimization are biodegradability, ease of recycling, toxicity of phase-forming chemicals, physical properties of extraction systems, and last although not least the total costs (see "Economic Aspects").

For research purposes, the most investigated system is the PEG–dextran (1) combination. Both polymers are commercially available in different molecular weight fractions. For technical applications, fractionated dextran is too expensive; crude dextran may be used in low concentrations (17,18), but results in high-viscosity bottom phases. A less expensive possibility is PEG–salt systems, which are commonly used for large-scale applications (8). Other alternatives are reported (19,20). When substituting other polymers for dextran, for example, the limiting concentration for phase separation needs to be considered as well as product yield and quality. Combinations of the different systems are possible, like polymers and surfactants (21,22) or affinity partitioning combined with temperature-induced phase separation (23).

For large-scale operations, questions of waste disposal and the possibility of recycling auxiliary chemicals are of increasing interest for environmental as well as economic reasons. These topics are discussed separately under "Applications".

Thermodynamics

Phase separation in aqueous systems and partitioning of proteins are governed by thermodynamic forces. Because thermodynamics including mathematical models of aqueous two-phase systems cannot be explained in a few sentences, the reader is referred to the literature. For a

broader introduction, the review by Cabezas (24) should be consulted.

Phase separation is based on the Gibbs free energy of mixing and will occur if ΔG^{mix} of two phases becomes energetically more favorable than ΔG^{mix} of a single phase:

$$\Delta G^{mix} = \Delta H^{mix} - T\Delta S^{mix} \quad (8)$$

where ΔH^{mix} and ΔS^{mix} are the enthalpy and entropy of mixing and T is the temperature. ΔG^{mix} is related to the excess enthalpy G^e by

$$\Delta G^{mix} = G^e + RT \sum_i x_i \ln x_i \quad (9)$$

where x_i is the mole fraction of the different phase components i.

There is a large diversity of polymer–polymer models available, such as Edmond and Ogston-based models using McMillan–Mayer or Hill virial expansions (25,26), Flory–Huggins–based lattice models using the classical polymer solution theory (27,28), the UNIQUAC lattice model (29), integral equation theory-based models (30), the VERS model based on group contributions of salt and polymer (31), and the excluded volume theory based on statistical geometric arguments (32). A comprehensive discussion can be found in a recent paper by Johansson et al. (33). The problem for most of these models is that they include theories derived for gases or crystalline solids. They have advantages and disadvantages, such as too many parameters, time-consuming experiments to obtain parameters, and derived parameters that are valid only for a certain range of conditions.

An interesting approach to solve binodales for PEG–dextran systems is shown by Baughman and Liu (34). They used about 100 different binodales and protein partitioning data from other investigators as raw data for an expert system based on a neural network, resulting in quantitative predictions for unknown phase diagrams. To predict protein partition away from the isoelectric point the results were not satisfactory, however. In polymer–salt systems most of the electrostatic interactions are screened by the high salt concentration of the phases, which makes modeling more complicated. Only a few models have been developed based on virial expansions (31), integral equation theory (30), and a potential of mean force model including fluctuation solution theory (35).

Until now a reliable prediction of protein partitioning in aqueous two-phase systems for a random protein has not been possible. The systems are too complex, depending on pH, salt, polymer, protein structure, temperature, specific interactions, and charge distribution on the protein surface. The aim to minimize experiments for certain systems and predict trends is advancing. The problem with the model experiments is that they rely upon commercially available cheap proteins, which have strong specific interactions, however, such as lysozyme or α-chymotrypsin in which dimerization or self-digestion occurs. These effects in aqueous two-phase systems have not been studied extensively. Work has been done to predict peptide partitioning (27,36–38). Predictions have improved over recent years for certain peptides and amino acids, but depending on the sequence, predictions do not work well yet in general.

The diverse models describe phase separation reasonably well but a state-of-the-art model is not available (24). The interaction between proteins and solutions needs more attention to understand protein partition and describe it in quantitative terms based on fundamental principles. This goal will be elusive for quite some time to come. Fortunately it is very easy to determine K values experimentally so long as an appropriate specific assay is available.

ALTERNATIVE TWO-PHASE SYSTEMS

In addition to the commonly used polymer–polymer and polymer–salt systems, there are some alternative, also well-investigated, aqueous two-phase systems.

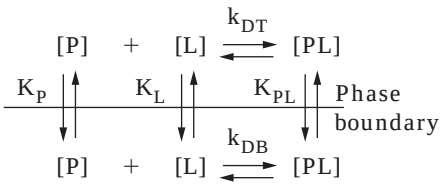
Affinity Extraction

Ligands covalently bound to one of the phase-forming components are useful to direct the extraction of certain proteins because the ligand will predominantly partition with the polymer in one phase. There exists a linked equilibrium of partition and complex formation, as shown in Figure 4. The extraction efficiency depends on the number of binding sites of the protein and the dissociation constant of the complex. The latter may be different in the top and bottom phase (39,40).

Affinity ligands can be cofactors or inhibitors of the target protein. General ligands based on textile dyes are more often used than natural ligands. Besides textile dyes, metal chelators are attractive as general ligands. These are robust, cheap chemicals applicable in PEG–salt systems and offering useful selectivity. The syntheses of PEG derivatives has been described by Harris (41). Johansson et al. (42) investigated the partition of phosphofructokinase from yeast. The partition coefficient increased by a factor of 10^4 compared to physical extractions using the pseudo-affinity ligand cibacron-blue in an affinity process (Fig. 5).

The effectivity of affinity extraction was also demonstrated by Kroner et al. (43). With increase of the amount of PEG blue, the partitioning coefficient K increases. A yield of 98% could be obtained thereby with a single extraction step. Other parameters than the ligand concentration (43) are the polymer concentration (44), the kind of salt and its concentration (42), the pH value (45), the temperature (45), and the addition of free ligands (44,45). Useful models describing affinity extraction have been derived by Cordes et al. (40) and Suh and Arnold (46).

An example of screening for a suitable dye can be found in a patent of Guliano and Szlag (47) describing experiments with the triazin dyes Cibacron-Blue FGF, Procion Turquoise H-A, Procion Green HE-4BDA, and Procion Red HE-3B to separate alcohol dehydrogenase. Large-scale applications of affinity extraction are listed later. Reviews of polymer–ligand systems have been published by Johansson (48) and Johansson and Tjerneld (49). It can be expected that specific ligands originally selected from libraries of suitable synthetic chemicals or peptides for affinity



[P] Concentration of protein
[L] Concentration of ligand
[PL] Concentration of protein–ligand complex
K_P Partition coefficient of the protein
K_{PL} Partition coefficient of the ligand
K_L Partition coefficient of the complex
k_{DT} Equilibrium constant of the complex in the top phase
k_{DB} Equilibrium constant of the complex of the bottom phase

Figure 4. Equilibrium of partition and complex formation in affinity extraction processes. [P], concentration of protein; [L], concentration of ligand; [PL], concentration of protein–ligand complex; K_P, partition coefficient of the protein; K_{PL}, partition coefficient of the ligand; K_L, partition coefficient of the complex; k_{DT}, equilibrium constant of the complex in the top phase; and k_{DB}, equilibrium constant of the complex in the bottom phase.

chromatography will find applications also in affinity extraction.

Detergent-Based System

A recent development is the two-phase extraction in aqueous two-phase systems based on nonionic detergents carrying ethylenoxide groups as the hydrophilic moiety. The protein is partitioned using only one auxiliary chemical. The principle is shown in Figure 6. At low temperature, a homogeneous phase containing micelles of the detergent and fermentation broth exists. The detergent interacts preferably with hydrophobic proteins. As soon as the temperature is increased above a certain system-dependent temperature, turbidity develops as an indicator of phase separation. The so-called cloud point depends mainly on the average number of ethylene oxide groups; a narrow range is desirable. A C12EO5 detergent, for example, has a cloud point between 23 and 25 °C. The phase separation depends on the reversible, temperature-induced dehydration of the ether oxygens in the detergent and leads to a detergent-rich and a detergent-depleted phase in equilibrium. The micelles are believed to aggregate into laminar structures carrying hydrophobic proteins into the so-called coazervate phase, which may still contain 70–80% water (50).

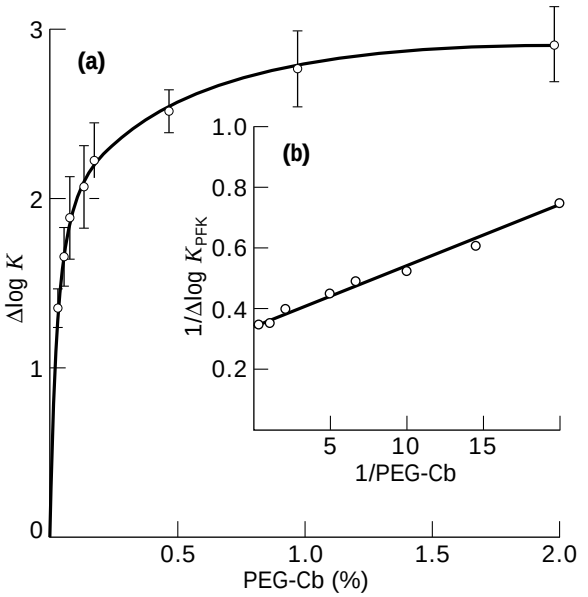


Figure 5. Changes of the partition coefficient of phosphofructokinase depending on the amount of added ligand. *Source*: Ref. 42.

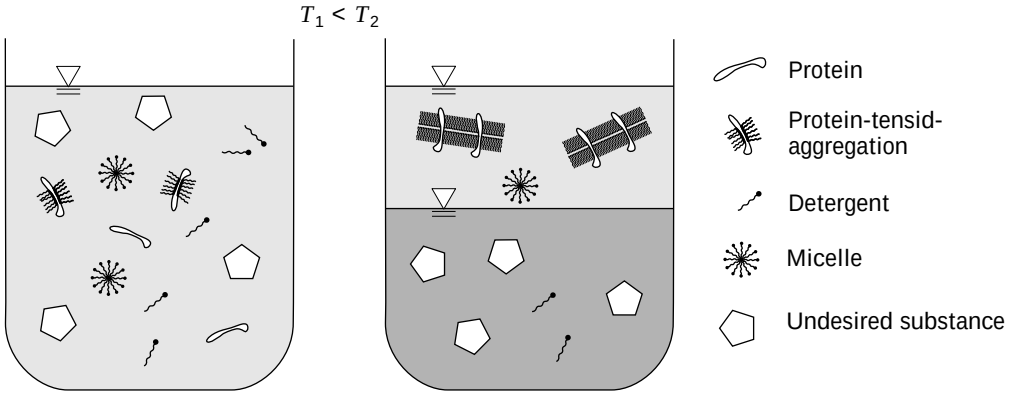


Figure 6. Schematic illustration of aqueous detergent solutions below and above the cloud point.

Bordier (51), Minuth et al. (52), and others investigated cloud-point extraction of proteins. Nonionic or zwitterionic detergents are most useful for the separation of membrane proteins (53,54). A complete process design for the purification of cholesterol oxidase is presented in Figure 7. A scale-up study indicated that the rather low density difference in detergent-based systems limits phase separation by centrifugation in the initial step (56).

It has been shown that the detergent for extraction of hydrophilic proteins and membrane-bound proteins should be chosen with respect to its hydrophile–lipophile balance (HLB) (57). Table 1 shows partitioning results using Tritons with different HLBs for the separation of an endopeptidase in the presence of sodium chloride. Yield, concentration factor, and partitioning coefficient increased significantly with an increasing HLB.

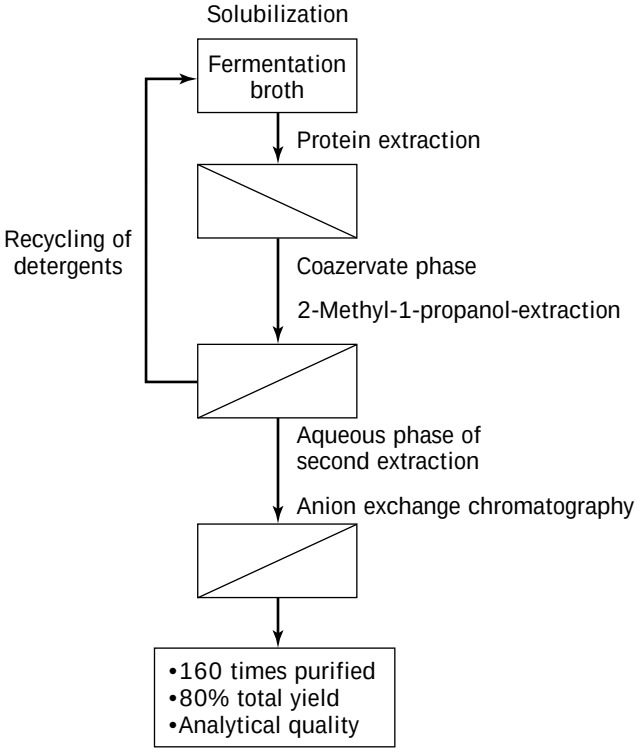


Figure 7. Optimal process design for the extraction of cholesterol oxidase from *Nocardia rhodochrous* using detergent-based systems in the initial step. (*Source*: Ref. 55.)

Table 1. Influence of the Surfactant Hydrophile–Lipophile Balance (HLB) on the Partition Coefficient, Concentration Factor, and Yield of the Purification Process for the Example of Endopeptidase

Surfactant	HLB	Partition coefficient	Concentration factor	Yield (%)
Triton X-100	13.5	0.81	0.7	15
Triton X-165	15.8	4.17	3.65	49
Triton X-305	17.3	8.93	3.1	69
Triton X-405	17.9	53	6.63	100

Source: Ref. 57.

Reversed Micelles

Because it is similar to aqueous liquid–liquid extraction, reversed micellar protein extraction is mentioned here briefly. Reversed micelles extract proteins into the water pool contained in the micelles present in an organic solvent depending on the isoelectric point of the protein and therefore pH and ionic strength in the process solution. The principle is shown in Figure 8. The organic solvent allows the organization of the reversed micelles and the separation of the bulk water from the water pool in the micelles.

Extraction into reversed micelles has been investigated for several industrially relevant proteins, but these are apparently not yet used in industry (58,59). The major remaining problem is the difficult reextraction of the bioactive protein into an aqueous phase and the reuse of reversed micelles.

Thermoseparating Polymers

Certain polyethylene–polypropylenoxide copolymers also show reversible dehydration of the oxygen groups in an interesting temperature range and have been investigated by Alred et al. (60), Johansson et al. (61), and recently by Persson et al. (62) for application in aqueous two-phase systems. Usually a polymer–polymer system is used in the first step selected to partition the protein in the copolymer-rich phase. After primary phase separation, the solution is heated above the cloud point, which results in phase separation under conditions in which the protein is almost exclusively found in the polymer-depleted phase.

APPLICATIONS

Technical Aspects

Aqueous liquid–liquid extraction uses the same principles as the well-known chemical process of liquid–liquid ex-

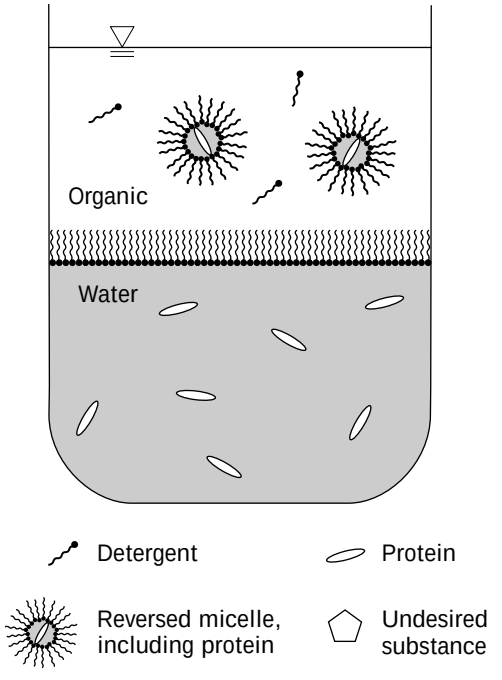


Figure 8. Operating principle of extraction into reversed micelles.

traction. Therefore similar, commercially available equipment may be used for mixing and phase separation. The physical properties of aqueous two-phase systems, however, are very different from those of solvent–water systems. The most notable difference is the dramatically lower interfacial tension (Table 2).

In general, for the purification of an intracellular protein the cells are disrupted, the phase-forming components added, and the pH adjusted; the suspension is then mixed to reach equilibrium. In large-scale processes this is done in a stirred vessel with baffles or using static mixers (63). An extremely low energy input is needed because of the low interfacial tension of the whole system. The approach to equilibrium is surprisingly fast (≤30 s), as demonstrated by Fauqueux et al. (64), provided turbulent mixing is ensured. After equilibration is reached the phases should separate into a top and bottom phase. This separation can be done under unit gravity or can be speeded up by centrifugation. The latter is recommended if the viscosity is high or the density difference between the phases is low (8). Centrifugation gives higher yields as entrainment is avoided at higher *g* forces; optimal feed rates of the suspension up to 3 L/min corresponding to residence times of about 16 s have been reported using small disk stack separators (8).

After the initial separation step, the top phase is often mixed with salts to generate a secondary phase system. Conditions are selected such that the protein of interest is partitioned to the bottom phase and can be further processed after phase separation (Fig. 9). Other possible procedures for the isolation of proteins from polymer phases are centrifugation, adsorption, ultrafiltration, or electrophoreses (66). In large-scale processes, ultrafiltration or diafiltration is used to concentrate and condition the product from the secondary bottom phase. The top phase may be recycled (52,67). A secondary bottom phase may be applied under certain conditions to a chromatographic support, such as hydrophobic interaction matrices.

Extraction usually is performed in the batch mode, but often it may be more cost-effective to use a continuous process. In general there are three possibilities of continuous processing in production lines: the co-, cross-, and countercurrent flow. The three modes are illustrated in Figure 10, and a comparison of specific properties is shown in Table 3. In large-scale processes the counter-current method should be the most efficient because of a higher yield and lower consumption of chemicals (63). Countercurrent extraction has been studied for the multistep procedures discussed next. The continuous cross-current extraction has been investigated experimentally in pilot scale. Enzyme extraction has been successfully carried out using more than 500 kg yeast per day in a computer-controlled process (68). The procedure could also be applied to extract enzymes continuously from bacterial cell homogenates (69).

To achieve a higher purification, multiple partition steps would be required. Not every multistage apparatus used in the chemical industry for liquid–liquid extraction is suitable for operation with aqueous phases as the physical properties of the two phases are too similar and the interfacial tension too low. The performance of the Graesser contactor has been best investigated. For the separation of α-lactalbumin from whey, a space time yield of 29 g/L reactor volume and day was achieved (70). The speed of rotation in the "raining bucket contactor" had to be reduced from the usual 20 U/min to 2–5 U/min, working with PEG–salt systems (71). In addition, the rotational direction was reversed to reduce the intensity of mixing and to avoid flooding. Other apparatus investigated for multistage extractions are mixer-settler, Kühni columns (72), the Podbielniak extractor (1), and spray columns (73). All can be operated successfully with specific precautions against flooding. Changing operating conditions such as pH or the length of the tie-line, however, may be more effective for protein purification than multiple partition steps.

At Miles Inc., a conventional downstream process was compared to an operation using aqueous two-phase systems and *Bacillus* alpha-amylase as an example (74). The latter was carried out in a 50,000-L large-scale process (Fig. 11). In this example the extraction process in aqueous two-phase-systems required fewer steps, resulting in higher yield and lower costs.

Recycling of Process Chemicals

To lower high chemical and waste disposal costs, recycling or recirculation of the auxiliary chemicals is recommended. Recycling of phosphate by crystallization at 6 °C has been described from solid-free process streams (8). The successful recirculation of most of the PEG has been demonstrated by Papamichael et al. (75) and of detergents by Minuth et al. (55) (see Fig. 7). The recovery of salt from the primary bottom phase was accomplished by Greve and Kula (76) making use of miscibility gaps in the ternary-phase systems—water–salt–lower aliphatic alcohols (Fig. 9). The authors also presented an economic analysis (76). In principle, the material balance of the auxiliary chemicals introduced into the process stream to accomplish phase separation can be closed more than 90% by recycling, yielding cost-effective and environmentally acceptable procedures.

Large-Scale Applications

Examples of Large-Scale Applications. Direct information about industrial applications of aqueous two-phase extrac-

Table 2. Physical Properties of Aqueous Two-Phase Systems

System composition	$\Delta\rho$ (g/cm^3)	η_T (mPa s)	η_B (mPa s)	$\eta_{dispersion}$ (mPa s)	δ (mN/m)	Reference
9% PEG 4000, 2% dextran T-500	1,010	3	94	3.4	—	16
7% PEG 4000, 1.25% crude dextran	1,046	2.7	2,000	4	—	16
16% PEG 4000, 13% phosphate	1,085	12	2.1	—	1.25	58

Source: PEG, polyethylene glycol. Ref. 8.

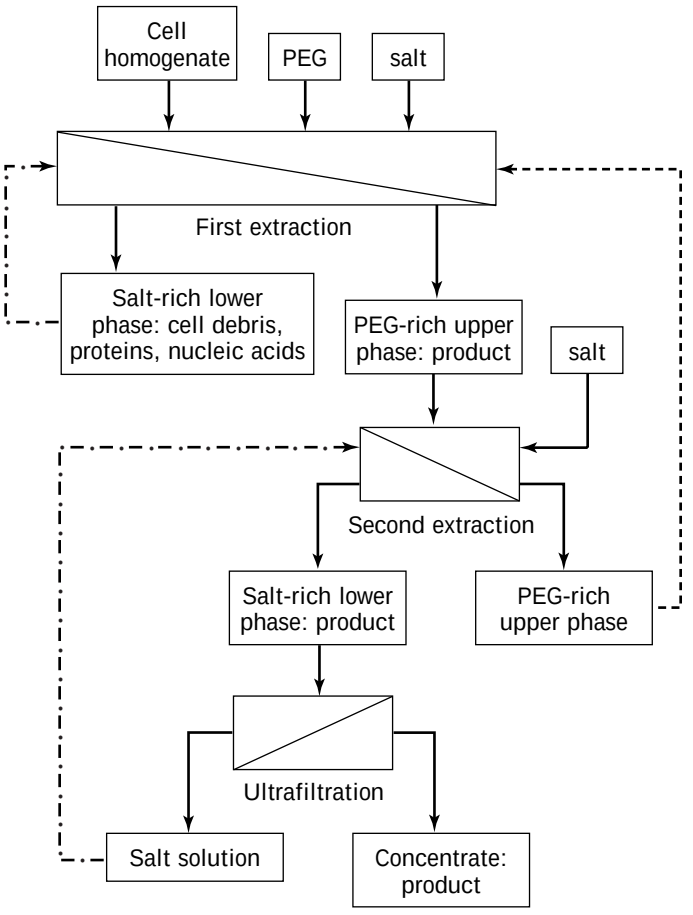


Figure 9. Flow sheet for product, chemicals, and waste flows in an extraction process using a PEG–salt aqueous two-phase system. ·–·–·–·, recycling of salt; ––––, recycling of PEG. *Source*: Ref. 65.

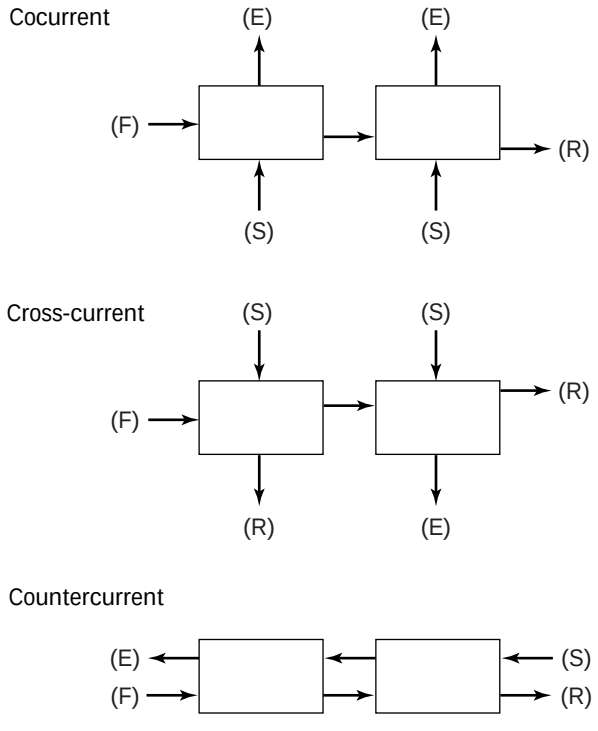


Figure 10. Principles of co-, cross-, and counter-current flow. F, feed; S, solvent; E, extract; R, raffinade.

tion is scarce. In Table 4, a list of patents and large-scale applications (beginning in 1982; minimum, 24-kg cells) described in the literature is compiled. The size of the operation reached 50,000 L or more (74). Usually centrifugation is preferred in the initial phase separation. (An example is illustrated in Fig. 11.)

Scale-Up. One big advantage of aqueous liquid–liquid extraction is the ease of scaling up. Hart et al. (89), Schütte et al. (93), and Kroner et al. (94) showed that scale-up is feasible by proportional linear increase of the amount of ingredients. The necessary centrifugation equipment is commercially available and can be selected using the Sigma concept. In Figure 12, scale-up data of Kroner et al. (94) are presented for the purification of formate dehydrogenase from *Candida boidinii* in PEG–potassium phosphate. A scale-up of nearly 40,000 was carried out for different separation steps without significant changes in yield.

Economic Aspects. The costs of the protein recovery and enrichment by aqueous liquid–liquid extraction is of course highly dependent on the actual system, scale, etc. As an example Kroner et al. (95) analyzed the cost structure for extraction and compared the economics with other downstream separation methods employed as alternatives for the purification of intracellular fumarase from *Brevibacterium ammoniagenes* in a 100-kg batch process. With a

Table 3. Advantages and Disadvantages of Different Flow Regimes in Continuous Processing

	Suitable for	Consumption requirements	Yield	Purification factor
Cocurrent	Low K	As cross-current	Medium	Medium
Cross-current	High K	As co-current	Lowest	Highest
Countercurrent	Low K	Lowest	Highest	Lowest

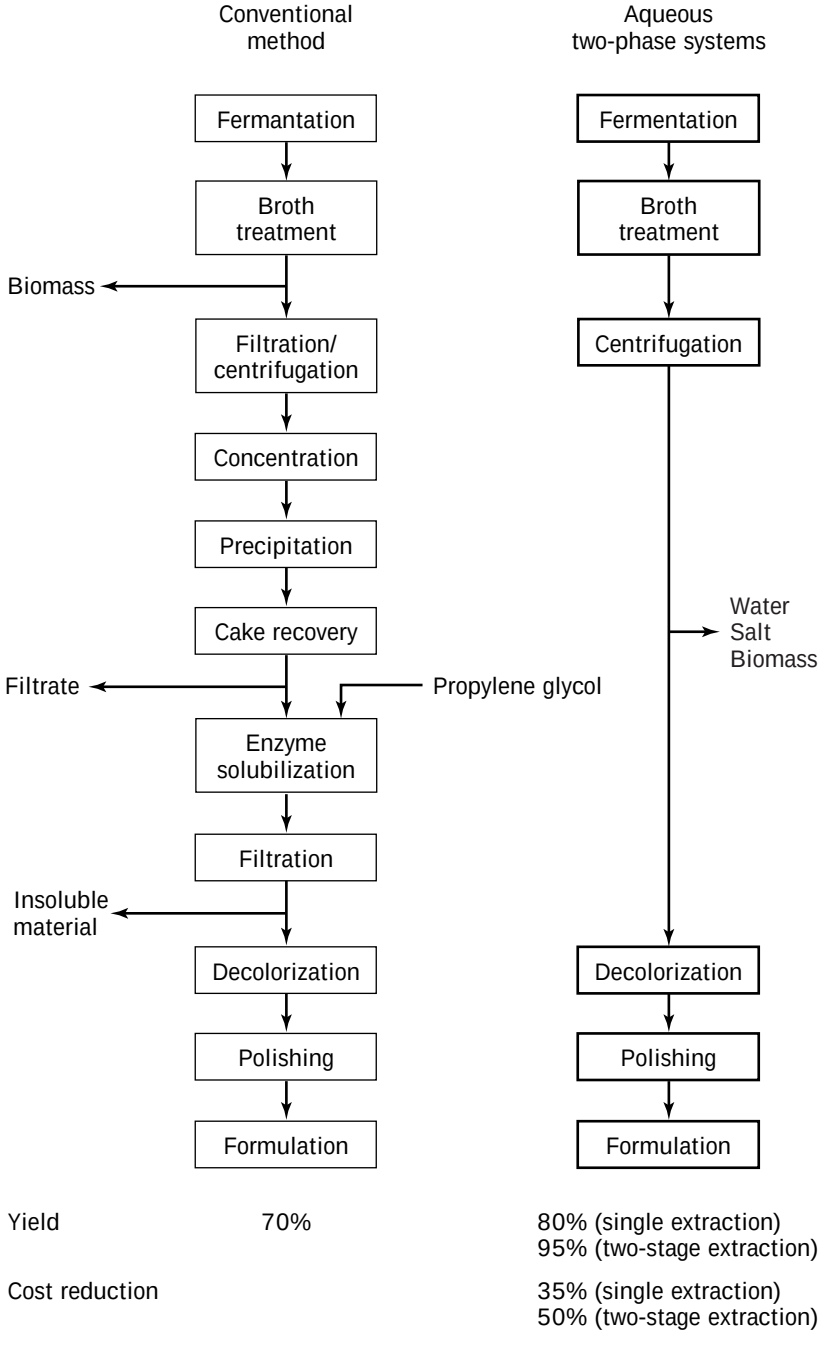


Figure 11. Comparison of the downstream process of alpha-amylase, using a conventional method (left column) and using aqueous two-phase systems (right column). *Source*: Ref. 74.

Table 4. Patents and Large-Scale Applications of Protein Extraction Using Aqueous Two-Phase Systems

Literature		Author	Date	Enzyme or protein	System used	Ligand	Origin
77	Patent	Kula	1982	Interferon	PEG–dextran		Raw material
15	Journal article	Hummel et al.	1983	D-Lactate dehydrogenase	PEG–phosphate		*Lactobacillus cellubiosus*
78	Patent	Kim et al.	1985	Protease, amylase	PEG–dextran		*Aspergillus orizae*
79	Journal article	Schütte et al.	1985	L-Leucin dehydrogenase	PEG–phosphate		*Bacillus cereus*
80	Patent	Gustafsson et al.	1986	ADH, hexokinase	PEG–K_2HPO_4, KH_2PO_4	Yes[c]	Yeast
80	Patent	Gustafsson et al.	1986	Transferrin	PEG–K_2HPO_4, KH_2PO_4	Yes	Blood plasma
81	Patent	Paul et al.	1986	Dextran-sucrase	PEG–dextran		*Leuconostoc mesenteroides*
82	Patent	van Wijnendaele et al.	1986	Hepatitis B antigen (HBsAg)	PEG–$(NHY_4)_2SO_4$		Yeast
82	Patent	van Wijnendaele et al.	1986	Alpha-1-antitrypsin	PEG–$(NH_4)_2SO_4$		Yeast
83	Patent	Dove et al.	1987	Therapeutically active proteins[a]	PEG–KH_2PO_4		Blood plasma
44	Journal article	Tjerneld et al.	1987	Lactate dehydrogenase	PEG–aquaphase PPT	Yes	Pig muscle
84	Patent	Ananthapadmanabhar	1988	Alcaline protease	PEG–Na_2SO_4	Yes	
85	Patent	Brewer et al.	1988	Protease	PEG–Na_2SO_4		Whole fermentation beer
14	Patent	Sieron et al.	1991	Recombinant proteins	PEG–polyvinylalcohol	[b]	
86	Patent	Enfors et al.	1992	Human immunoglobulin G	PEG–phosphate		Staphylococcus
47	Patent	Giuliano et al.	1992	Alcohol dehydrogenase (ADH)	PVP–maltodextrin	Yes	Baker's yeast
87	Patent	Heinsohn et al.	1992	Chymosin	PEG–Na_2SO_4		*Aspergillus niger* var. *awamori*
88	Patent	Kirchberger et al.	1992	Alkalische phosphatase	PEG–dextran		Calf intestine
71	Journal article	Coimbra et al.	1994	Beta-lactoglobulin	PEG–phosphate		Cheese whey
71	Journal article	Coimbra et al.	1995	Alpha-lactalbumin	PEG–phosphate		Cheese whey
18	Journal article	Cordes et al.	1994	Formate dehydrogenase	PEG–phosphate	Yes	*Candida boidinii*
89	Journal article	Hart et al.	1994	IGF[a]	PEG–Na_2SO_4		*E. coli*
90	Patent	Lorch et al.	1994	EG[c]	PEG–$(NH_4)_2SO_4$		Cellulase mixture
13	Patent	Builder et al.	1995	IGF-I or mammalian polypeptide	PEG–citrate	[b]	*E. coli*
91	Patent	Heinssohn et al.	1995	Chymosin	PEG–Na_2SO_4		Bovine stomach
92	Patent	Lee and Kahn	1995	Hemoglobin	PEG–phosphate buffer		Bovine blood
57	Patent	Braunstein et al.	1995	Different lipases and proteases	Detergents		Different organisms

Patents often cover wide ranges; therefore, in some cases only examples from the patent are given.
[a]Albumin, IgM, IgG, alpha-1-antitrypsin.
[b]Use of chaotropic substances.
[c]Adsorption at Sepharose particles.

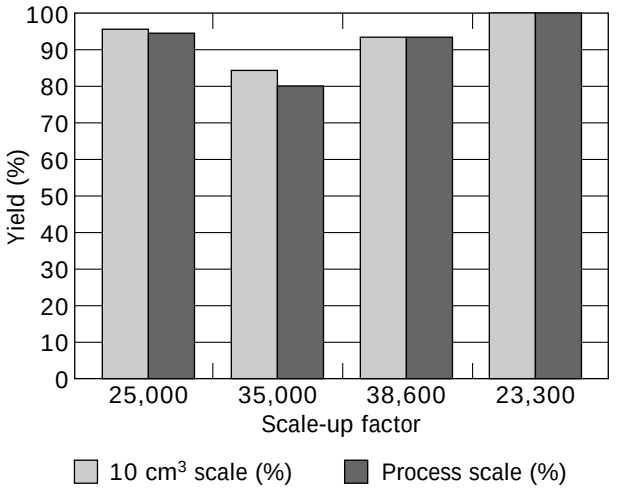


Figure 12. Scale-up of various extraction steps during the isolation and purification of formate dehydrogenase from *Candida boidinii*. *Source*: Ref. 94.

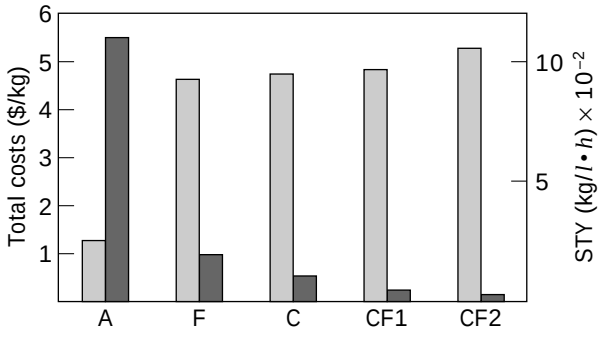


Figure 13. Cost factors and total costs comparing different initial isolation procedures. A, aqueous-phase system operation; F, filtration; C, centrifugation; CF, cross-flow filtration (1, retention coefficient $R = 0$; 2, $R = 0.7$). *Source*: Ref. 95.

system of PEG and potassium phosphate, an overall yield of 70% was achieved by a two-step extraction. The cost structure of the extraction is shown in Table 5. Even if prices have changed meanwhile, the table still gives some idea about the relative cost distribution of material, labor, energy, and investment for the different process steps.

In Figure 13, a comparison of the cost factors and the total costs of the initial recovery operations is shown (95). Data in Table 5 and in Figure 13 are calculated from equation 10, which is derived from coefficients used for the cost calculation in the chemical industry and normalized for the yield:

$$T = \frac{1.13M + 2.6L + 1.13E + 0.13I}{Y} \qquad (10)$$

where the letters represent costs for T, total production; M, material; L, labor; E, energy; I, investment (capital); and Y, yield.

Aqueous two-phase extraction is a low-energy, labor-requiring technique with comparatively low investment costs and high material costs. These high costs for auxiliary chemicals can be lowered by a good choice of volume ratio and by new phase systems and especially by the development of recycling processes (65). The process time can be decreased dramatically by centrifugal separation, resulting not only in high space time yields but also in improved product quality because of the absence of significant proteolytic degradation (96,97). Economic aspects of affinity extraction have been discussed by Johansson et al. (45) and Cordes et al. (18). In such cases, economics rely heavily on the initial costs and the reusability of the affinity ligand.

For a valid evaluation, it is important to realize that extraction constitutes an integrative technology; it achieves removal of solids and an increase in the specific activity of the product in a single step. In addition, in most cases the product can be separated from high molecular weight nucleic acids and its concentration can be increased, improving the performance of subsequent separation steps. An industrial total cost comparison between a downstream process using aqueous two-phase systems and a conventional one is shown in Figure 11. A cost reduction of 35% to 50% was achieved for the separation of *Bacillus* alpha-amylase.

CONCLUSION

Liquid–liquid extraction of proteins using aqueous two-phase systems is a comparatively new but nevertheless

Table 5. Cost Structure of an Extraction Process

Step	Feedstock ($/kg)	Material ($/kg)	Labor ($/kg)	Energy ($/kg)	Instrumentation ($/kg)	Total ($/kg)	Percentage of total
Fermentation	5.80					6.55	40.1
Disruption		0.03	0.92	0.03	2.76	2.82	17.2
1. Extraction		1.89	0.48	<0.01	0.94	3.52	21.5
2. Extraction		0.17	0.29	<0.01	0.94	1.08	6.6
Ultrafiltration		0.20	0.48	0.01	0.90	1.60	9.8
Waste water						0.78	4.8
Sum ($/kg)	5.80	2.29	2.17	<0.06	5.54	16.35	
Percentage of total	35.5	14.0	13.3	0.4	33.9		

Source: Ref. 95.

well-investigated technique for the isolation and separation of proteins. It is apparently not yet widely used in industry, but the number of industrial applications seems to have increased in recent years due to the effectiveness and ease of operation on a large scale. Intensive research is going on worldwide with respect to practical applications and theoretical understanding. Together with increases in production size for pharmaceutical proteins, this will make liquid–liquid extraction in aqueous two-phase systems a more common unit operation in the biotechnological industry.

BIBLIOGRAPHY

1. M.-R. Kula, K.H. Kroner, and H. Hustedt, *Adv. Biochem. Eng.* **24**, 73–118 (1982).
2. M.-R. Kula, *Bioseparation* **1**, 181–189 (1990).
3. H. Walter and G. Johansson, *Methods in Enzymology*, vol. 228, Academic Press, San Diego, Calif., 1994.
4. P.A. Albertsson, *Partition of Cell Particles and Macromolecules*, Wiley, New York, 1986.
5. H. Walter, D.E. Brooks, and D. Fisher, *Partitioning in Aqueous Two-Phase Systems*, Academic Press, Orlando, Fla., 1985.
6. B.Y. Zaslavsky, *Aqueous Two-Phase Partitioning*, Dekker, New York, 1994.
7. M.-R. Kula, in A. Humphrey, C.L. Cooney eds., *Comprehensive Biotechnology*, vol. 2, Pergamon Press, New York, 1985, pp. 451–471.
8. H. Hustedt, K.H. Kroner, and M.-R. Kula, in H. Walter, D.E. Brooks, and D. Fisher eds., *Partitioning in Aqueous Two-Phase Systems*, Academic Press, Orlando, Fla., 1985, pp. 529–587.
9. J.N. Bronsted, *Z. Phys. Chem. Abt. A* **157**, 257 (1931).
10. D.J. Kubek, in R.G. Harrison ed., *Protein Purification Process Engineering*, Dekker, New York, 1994, pp. 87–114.
11. U. Menge, M. Morr, U. Mayr, and M.-R. Kula, *J. Appl. Biochem.* **5**, 75–90 (1983).
12. P.A. Albertsson, *Endeavour* **1**, 69 (1977).
13. U.S. Pat. 5,407,810 (April 18, 1995), S. Builder, R. Hart, P. Lester, J. Ogez, and D. Reifsnyder (to Genentech, Inc.).
14. GDR Pat. DD 288 837 (April 11, 1994), R. Sieron, R. Wondraczek, K. Binder, H. Dittmar, K. Lambrecht, and H. Thessa (to Zentralinstitut für Mikrobiologie und experimentelle Therapie).
15. H. Schütte, W. Hummel, and M.-R. Kula, *Appl. Microbiol. Biotechnol.* **19**, 167–176 (1984).
16. R.A. Hart, J.R. Ogez, and S.E. Builder, *Bioseparation* **5**, 113–121 (1995).
17. K.H. Kroner, H. Hustedt, and M.-R. Kula, *Biotechnol. Bioeng.* **24**, 1015–1045 (1982).
18. A. Cordes and M.-R. Kula, in H. Walter and G. Johansson eds., *Methods of Enzymology*, vol. 228, Academic Press, San Diego, Calif., 1994, pp. 600–617.
19. D.C. Szlag, K.A. Giuliano, and S.M. Snyder, J.F.P. Hamel, J. Hunter, and S.K. Sikdar eds., *Downstream Processing and Bioseparation*, ACS Symposium Series, vol. 419, American Chemical Soc., Washington, D.C., 1990, pp. 71–86.
20. M.-R. Kula, in H. Brauer ed., *Biotechnology*, vol. 2, VCH-Verlagsgesellschaft, Weinheim, 1985, pp. 725–760.
21. U. Sivars, K. Bergfeld, L. Piculell, and F. Tjerneld, *J. Chromatogr. B* **680**, 43–53 (1996).
22. H.-O. Johansson, G. Karlström, and F. Tjerneld, *Macromolecules* **26**, 4478 (1993).
23. P.A. Alred, A. Koslowski, and J.M. Harris, *Bioseparation* **2**, 363–373 (1992).
24. H. Cabezas, Jr., *J. Chromatogr. B* **680**, 3–30 (1996).
25. D. Forciniti and C.K. Hall, *ACS Symp. Ser.* **419**, 53–70 (1990).
26. F. Döbert, A. Pfennig, and M. Stumpf, *Macromolecules* **28**, 7860–7868 (1995).
27. A.D. Diamond, K. Yu, and J.T. Hsu, in *Protein Purification*, ACS Symposium Series, American Chemical Soc., Washington, D.C., 1990, pp. 52–65.
28. A. Gustafsson, H. Wernerström, and F. Tjerneld, *Polymer* **27**, 1768–1770 (1986).
29. H. Hartounian and S.I. Sandler, *Biotechnol. Prog.* **7**, 279 (1991).
30. C.A. Haynes, F.J. Benitez, H.W. Blanch, and J.M. Prausnitz, *AIChE J.* **39**, 1539 (1993).
31. C. Großmann and G. Maurer, *Fluid Phase Equilib.* **106**, 17–25 (1995).
32. Y. Guan, T.H. Lilley, and T.E. Treffry, *Macromolecules* **26**, 3971 (1993).
33. H.-O. Johansson, G. Karlsström, F. Tjerneld, and C.A. Haynes, *J. Chromatogr. B* **711**, 3–17 (1998).
34. D.R. Baughman and Y.A. Liu, *Ind. Eng. Chem. Res.* **33**, 2668–2687 (1994).
35. M. Kabiri-Badr and H. Cabezas, Jr., *Fluid Phase Equilib.* **115**, 39–58 (1996).
36. J. Bringmann, B. Keil, and A. Pfennig, *Fluid Phase Equilib.* **101**, 211–225 (1994).
37. W.-Y. Chen, C.-G. Shu, J.Y. Chen, and J.-F. Lee, *J. Chem. Eng. Jpn.* **27**, 688–690 (1994).
38. M.A. Eiteman, C. Hassinen, and A. Veide, *Biotechnol. Prog.* **10**, 513–519 (1994).
39. M.-R. Kula, A. Walsdorf, and A. Cordes, *Ber. Bunsen-Ges. Phys. Chem.* **93**, 968–970 (1989).
40. A. Cordes, J. Flossdorf, and M.-R. Kula, *Biotechnol. Bioeng.* **30**, 514–520 (1987).
41. J.M. Harris, *J. Macromol. Sci.* **C-25**, 325 (1985).
42. G. Johansson, G. Kopperschläger, and P.A. Albertsson, *Eur. J. Biochem.* **131**, 589–594 (1983).
43. K.H. Kroner, A. Cordes, A. Schelper, M. Morr, A.F. Bückmann, and M.-R. Kula, in T.C.J. Gribnau, J. Visser, and R.J.F. Nivard eds., *Affinity Chromatography and Related Techniques*, Elsevier, Amsterdam, 1982, pp. 451–501.
44. F. Tjerneld, G. Johansson, and M. Joelsson, *Biotechnol. Bioeng.* **30**, 809–816 (1987).
45. G. Johansson and M. Andersson, *J. Chromatogr.* **303**, 39 (1984).
46. S.-S. Suh and F.H. Arnold, *Biotechnol. Bioeng.* **35**, 682–690 (1990).
47. U.S. Pat. 5,093,254 (March 3, 1992), K.A. Giuliano and D.C. Szlag (to the United States).
48. G. Johansson, in S.K. Sikdar and M. Bier eds., *Frontiers in Bioprocessing*, CRC Press, Boca Raton, Fla., 1989, pp. 271–284.
49. G. Johansson and H. Tjerneld, *J. Biotechnol.* **11**, 135–142 (1989).
50. G.C. Terstappen, R.A. Ramelmeier, and M.-R. Kula, *J. Biotechnol.* **28**, 263–275 (1993).
51. C. Bordier, *J. Biol. Chem.* **250**, 1604–1606 (1981).

52. T. Minuth, J. Thömmes, and M.-R. Kula, *J. Biotechnol.* **38**, 151–164 (1995).
53. R.P. Frankewich and W.L. Hinze, *Anal. Chem.* **66**, 944–954 (1994).
54. W.L. Hinze and E. Pramauro, *Crit. Rev. Anal. Chem.* **24**, 133–177 (1993).
55. T. Minuth, J. Thömmes, and M.-R. Kula, *Biotechnol. Appl. Biochem.* **23**, 107–116 (1996).
56. T. Minuth, H. Gieren, H. Pape, C. Raths, J. Thömmes, and M.-R. Kula, *Biotechnol. Bioeng.* **55**, 339–347 (1997).
57. Int. Pat. WO 96/23061 (January 27, 1995) E.L. Braunstein, N.T. Becker, G. Ganshaw, and T.P. Graycar (to Genencor International, Inc.).
58. B.D. Kelley, D.I.C. Wang, and T.A. Hatton, *Biotechnol. Bioeng.* **42**, 1199–1208 (1993).
59. M.J. Pires and J.M.S. Cabral, *J. Chem. Technol. Biotechnol.* **61**, 219–224 (1994).
60. P.A. Alred, A. Kozlowski, J.M. Harris, and F. Tjerneld, *J. Chromatogr. A* **659**, 289–298 (1994).
61. H.-O. Johansson, G. Karlsström, B. Mattiasson, and F. Tjerneld, *Bioseparation* **5**, 269–279 (1995).
62. J. Persson, L. Nyström, H. Ageland, and F. Tjerneld, *J. Chromatogr. B* **711**, 97–109 (1998).
63. N. Papamichael and H. Hustedt, *Methods Enzymol.* **228**, 573–584 (1994).
64. P.-F. Fauqueux, H. Hustedt, and M.-R. Kula, *J. Chem. Technol. Biotechnol.* **35B**, 51–59 (1985).
65. A. Greve and M.-R. Kula, *J. Chem. Technol. Biotechnol.* **50**, 27–42 (1991).
66. G. Johansson, in H. Walter and G. Johansson eds., *Methods of Enzymology*, vol. 228, Academic Press, San Diego, Calif., 1994, pp. 569–573.
67. H. Hustedt, K.-H. Kroner, U. Menge, and M.-R. Kula, *Trends Biotechnol.* **3**, 139–144 (1985).
68. H. Hustedt, B. Börner, K.H. Kroner, and N. Papamichael, *Biotechnol. Tech.* **1**, 49–54 (1987).
69. H. Hustedt and N. Papamichael, *Enzyme Eng.* **9**, 135–139 (1988).
70. J. Dos Reis Coimbra, J. Thömmes, A. Meirelles, and M.-R. Kula, *Bioseparation* **5**, 259–268, (1995).
71. J. Dos Reis Coimbra, J. Thömmes, and M.-R. Kula, *J. Chromatogr. A* **668**, 85–94 (1994).
72. H. Hustedt, K.-H. Kroner, U. Menge, and M.-R. Kula, *Enzyme Eng.* **5**, 45–47 (1988).
73. K.R. Jafarabad, S.B. Sawant, J.B. Joshi, and S.K. Sikdar, *Chem. Eng. Science* **47**, 57–68 (1992).
74. C.Y. Kim, J.W. Brewer, C.E. Brothers, T.F. Farver, and E.K. Lee, *3rd Chemical Cong. of North America*, Toronto, Canada, June 5–10, 1988.
75. N. Papamichael, B. Börner, and H. Hustedt, *J. Chem. Technol. Biotechnol.* **54**, 47–55 (1992).
76. A. Greve and M.-R. Kula, *Bioprocess Eng.* **6**, 173–177 (1990).
77. FRG Pat. DE 2943016 C2 (September 6, 1984), U. Menge, M. Morr, and M.-R. Kula (to the Gesellschaft für Biotechnologische Forschung mbH).
78. U.S. Pat. 4,508,825 (April 2, 1985), C.Y. Kim, T.F. Farver, and J.W. Brewer (to Miles Laboratories, Inc.).
79. H. Schütte, W. Hummel, H. Tsai, and M.-R. Kula, *Appl. Microbiol. Biotechnol.* **22**, 306–317 (1985).
80. U.S. Pat. 4,579,661 (April 1, 1986), S.J. Gustafsson, P.O. Hedman, T.G.I. Ling, and B.G. Mattiasson (to Pharmacia AB, Sweden).
81. U.S. Pat. 4,591563 (May 27, 1986), F. Paul, P. Monsan and D. Auriol (to Societe Nationale Elf Aquitaine, France).
82. E.U. Pat. 0 199 698 B1 (August 28, 1991), F. van Wijnendaele, D. Gilles, and G. Simonet (to Smithkline Biologicals S.A.).
83. U.S. Pat. 4,684,723 (August 4, 1987), G.B. Dove and G. Mitra (to Miles Laboratories, Inc.).
84. U.S. Pat. 4,743,550 (May 10, 1988), K.P. Ananthapadmanabhan and E.D. Goddard (to Union Carbide Corporation).
85. U.S. Pat. 4,728,613 (March 1, 1988), J.W. Brewer, C.E. Brothers, T.F. Farver, C.Y. Kim, and E. Lee (to Miles Laboratories, Inc.).
86. Int. Pat. WO 92/07868 (May 14, 1992), S.-O. Enfors, K. Köhler, Ch. Ljundquist, B. Nilsson, and A. Veide (to Pharmacia AB, Sweden).
87. U.S. Pat. 5,139,943 (August 18, 1992), H.G. Heinsohn, J.D. Lorch, and R.E. Arnold (to Genencor International, Inc.).
88. FRG Pat. DD 298 424 (February 20, 1992), J. Kirchberger, G. Kopperschläger, M. Rockstroh, and K. Eisenbrandt (to Universität Leipzig, Germany).
89. R.A. Hart, P.M. Lester, H. Riefsnyder, J.R. Ogez, and S.E. Builder, *BioTechnology* **12**, 1113–1117 (1994).
90. U.S. Pat. 5,328,841 (July 12, 1994), J.D. Lorch K.A. Clarkson, E. Larenas, B.S. Bower, and G.L. Weiss (to Genencor International, Inc.).
91. E.U. Pat. 0 477 284 B1 (August 16, 1995), H.G. Heinssohn and K.J. Hayenga (to Smithkline Biologicals S.A.).
92. U.S. Pat. 5,407,579 (April 18, 1995) C.-J. Lee and P. Khan (to National Science Council, Taipei).
93. H. Schütte, K.H. Kroner, W. Hummel, and M.-R. Kula, *Ann. N.Y. Acad. Sci.* **413**, 270–282 (1983).
94. K.H. Kroner, H. Schütte, W. Stach, and M.-R. Kula, *J. Chem. Technol. Biotechnol.* **32**, 130–137 (1982).
95. K.H. Kroner, H. Hustedt, and M.-R. Kula, *Process Biochem.* **19**, 170–179 (1984).
96. T.R. Pulliam, S. Winston, and W.E. Bentley, *Enzyme Microb. Technol.* **20**, 46–51 (1997).
97. H. Hellebust, A. Veide, and S. Enfors, *J. Biotechnol.* **7**, 185–198 (1988).

See also CELL DISRUPTION AND LYSIS; CRYSTALLIZATION, PROTEIN, KINETICS.

PROTEIN SECRETION, *SACCHAROMYCES CEREVISIAE*

K.D. WITTRUP
Kyowa Hakko Kogyo Co., Ltd.
Tokyo, Japan

KEY WORDS

Cell engineering
Chaperones
Expression vector
Foldases
Heterologous secretion
Protein expression

Saccharomyces cerevisiae
Yeast

OUTLINE

INTRODUCTION

The yeast *Saccharomyces cerevisiae* is an excellent host for production of nonglycosylated secreted proteins possessing disulfide bonds. As a microbial eukaryote, yeast possesses the protein folding and conformational proofreading apparatus of the endoplasmic reticulum and Golgi apparatus, yet it is amenable to high cell density fermentation in large, process-scale fermenters. Effective exploitation of the yeast cell's secretory capacity can produce yields in the gram per liter range of secreted, active protein.

YEASTS

The most widely studied yeast is *S. cerevisiae*, which is used as a model system for many studies of eukaryotic cell biology (1). As a result, a rich knowledge base is available concerning the physiology and genetics of *S. cerevisiae*, for example, the entire genome sequence has been obtained. Alternative yeasts have also been exploited for protein expression, such as *Pichia pastoris*, *Kluyveromyces lactis*, and *Hansenula polymorpha*. In particular, use of *P. pastoris* is widespread, and a kit is commercially available. In most cases, these nontraditional yeasts have been found to secrete many-fold higher levels of heterologous protein than *S. cerevisiae* (2,3). More recent studies have indicated, however, that the intrinsic secretory capacity of *S. cerevisiae* is inefficiently exploited by the most commonly used multicopy plasmid vectors; when chromosomal integrating vectors are used, the secretory capacity of *S. cerevisiae* is comparable to or greater than that of *P. pastoris* (4,5). Reviews of *P. pastoris* are available elsewhere (2); this article focuses on *S. cerevisiae*. Rather than exhaustively reviewing all facets of *S. cerevisiae* expression, the approaches expected to be of most practical value are highlighted. An excellent, more comprehensive review of yeast expression is available elsewhere (6).

PROTEINS SUITABLE FOR SECRETION IN YEAST

Yeast expression is used to best advantage for proteins whose native route of synthesis is the eukaryotic secretory pathway, a classification that includes most pharmaceutical proteins. Yeast is not well suited for producing glycosylated proteins destined for therapeutic administration, because of the introduction of immunogenic sugar linkages and the absence of sialic acid in yeast glycosylation. Nonmammalian glycosylation does not preclude yeast expression of glycoproteins for analytical, diagnostic, or other in vitro applications, however. A partial list of proteins industrially produced in *S. cerevisiae* for clinical trials is presented in Table 1 (7). Noteworthy yeast-produced commercial products are human insulin, granulocyte-macrophage colony-stimulating factor, and hepatitis B vaccine. A potentially important product class for yeast expression is single-chain antibody fragments and immunotoxins; many such therapeutics are currently in development (8).

VECTORS

Although the most commonly used *S. cerevisiae* vectors are based on the 2μ multicopy plasmid, amplifiable integrating vectors can offer strong advantages and are generally preferable for high-level secretion.

2μ Plasmids

Most laboratory strains of *S. cerevisiae* possess 50 to 100 copies of a circular plasmid called the *2*μ *plasmid* because of its dimensions observed by electron microscopy. The native 2μ plasmid does not confer a measurable growth phenotype, but represents parasitic selfish DNA and is apparently maintained simply as a result of effective amplification, replication, and partitioning mechanisms. Numerous expression vectors containing all or part of the 2μ plasmid have been constructed (9). 2μ-Based vectors exhibit broad clonal variation in copy number and are mitotically unstable. Controlled amplification of 2μ-based plasmids has been accomplished by overexpression of the FLP recombinase that converts θ replication intermediates to a rolling circle mode of amplification (9). 2μ vector copy numbers can also be amplified by use of faulty selection genes such as *leu2-d* or *ura3-d*, which are required at several hundred copies per cell to support growth (10,11).

Integrating Vectors

Chromosomal integration of heterologous expression cassettes takes advantage of the extremely high fidelity of the

Table 1. Proteins Expressed in *Saccharomyces cerevisiae* by Pharmaceutical and Biotechnology Companies for Clinical Trials

Immunomodulators
Interferon-α
Interferon-γ
Interleukin-1
Interleukin-2
Interleukin-3
Tumor necrosis factor
Growth factors
Granulocyte colony-stimutating factor
Macrophage colony-stimulating factor
Granulocyte-macrophage colony stimutating factor
Fibroblast growth factor
Transforming growth factor-α
Transforming growth factor-β
Epidermal growth factor
Platelet-derived growth factor
Fibronectin
Hormones
Insulin-like growth factor-1
Insulin-like growth factor-2
Human growth hormone
Human growth hormone releasing factor
Calcitonin
Human chorionic gonadotropin
Insulin
Proinsulin
Blood proteins
Human serum albumin
Hemoglobin
Factor VII
Factor XIII
Tissue-type plasminogen activator
Hirudin
α-1-Antitrypsin
Enzymes
Lysozyme
Superoxide dismutase
Gastric lipase protein
Vaccines
Hepatitis B
Malaria

Source: Data from Hodgson (7).

replication and partitioning apparatus used by the yeast cell to maintain its own genome. Vectors can be integrated either by homologous recombination or retrotransposition.

Homologous Recombination. The predominant majority of recombination events in *S. cerevisiae* are homologous, and vectors can be precisely targeted to particular chromosomal locations by restriction digestion in a segment of the vector identical to the chromosomal target sequence (12). Several integrating vectors have been developed that are targeted to multiple chromosomal targets. These vectors are typically integrated at one to three locations in iterated repeats, a configuration that can lead to instability if gene amplification reduces growth rate due to constitutive heterologous gene expression. The ribosomal DNA repeats have been used as integration targets (13), and Ty retrotransposon sequences have also been targeted. Several groups have developed δ vectors that are targeted to the 150 to 200 copies of the Ty long terminal repeat that are dispersed throughout the yeast chromosomes (14–19). The δ vectors have shown particular promise, increasing heterologous secretion 5 to 10 fold over levels achieved using 2μ plasmids (15,16). The recombination event that integrates the δ vector is reversed at a frequency of approximately 10^{-4}, looping out the vector sequences from the chromosome. If constitutive expression is used, clones with reduced copy number can quickly outgrow the amplified clones, resulting in substantial instability (18,20). If a strongly repressible promoter such as the GAL promoter is used instead, amplified expression cassettes are completely stable for more than 50 generations in the absence of antibiotic selection, because there is no growth selection against the amplified clones (16,20).

Retrotransposition. Expression cassettes can be inserted within inducible yeast retrotransposons and chromosomally integrated at dispersed sites via the processes of transcription, reverse transcription, and integration used by native yeast retrotransposons. Vectors of this type have been constructed from Ty1 and Ty3 retrotransposons (21,22). A key advantage of this approach is the dispersed nature of gene amplification, which reduces instability caused by recombination between iterated repeats. One disadvantage of this approach is that relatively low numbers of integration events occur at each round of amplification. A potential regulatory difficulty with this approach is the error-prone nature of the reverse transcription process, which could potentially introduce novel mutations in each integrated expression cassette.

SELECTION MARKERS

Nutritional Markers

Almost all yeast genetic studies not related to heterologous protein production use recessive chromosomal mutations with nutritional deficiency phenotypes. These mutations are complemented by plasmid-borne wild-type genes that allow transformants to propagate on medium lacking key amino acids or nucleotides. The genes and corresponding nutrients most commonly used are *URA3* (uracil), *TRP1* (tryptophan), *HIS3* (histidine), *LEU2* (leucine), and *ADE2* (adenine.) These genetic selection markers are often inserted into 2μ plasmids to select for plasmid maintenance. A refinement of this approach is to use a promoter-deleted version of the selection gene that is required at high amplification to support growth. The *leu2-d* and *ura3-d* alleles have been used in this way to select for very high 2μ copy numbers (10,11).

Dominant Markers

Dominant markers have two key advantages: the capability to transform polyploid industrial yeast strains and the capability to select for increasing gene dosage by increasing levels of the toxic selection compound. Several antibiotic resistance genes have been used as dominant markers for yeast transformation, but the most commonly used is the *G418* resistance gene from the *E. coli* transposon Tn903 (23). When used without a yeast promoter, expression of $G418^R$ is weak enough to provide a selectable advantage by increasing gene dosage. *G418* resistance is approximately proportional to the number of integrated copies of the Tn903 $G418^R$ gene (16). The *CUP1* copper metallothionein gene has also been used as a dominant selection marker, with the phenotype of increased copper resistance (24).

Autoselection

An essential gene can be used as the selectable marker in a strain with the chromosomal copy of that gene deleted. This approach allows the use of complex media without antibiotics, and has been used with *ura3fur1* (25) and Δ*tpi* (26) strains. Cells that lose the plasmid and its associated essential gene by segregational instability simply stop growing.

TRANSCRIPTIONAL PROMOTERS

Constitutive

Given that individual glycolytic enzymes can constitute several percent of the cell's total protein, the strong promoters from these genes are often used for constitutive heterologous expression. The most commonly used constitutive promoters are from the *PGK* and *GAP* genes (27,28). However, the use of a constitutive promoter for high level heterologous secretion is inadvisable because of inevitable negative growth phenotypes associated with high-level secretion, which in numerous studies have been shown to select for outgrowth of nonexpressing cells or cells with reduced expression (6,18,29,30).

Inducible

The GAL1 promoter is particularly well suited for heterologous expression because transcription is very tightly repressed in the presence of glucose, yet it is induced more than 1,000-fold by switching the carbon source to galactose. The wild-type levels of the Gal4p transcriptional activation protein can become limiting when multiple GAL1 promoters are induced, but this limitation can be overcome by single-copy inducible overexpression of Gal4p from a GAL promoter (31). Glucose-repressible promoters such as ADH2 and hybrid GAP/ADH2 promoters have also been used (32). However, difficulties may arise with such promoters during high cell density fermentations in balancing the central carbon metabolism against the requirement to repress heterologous expression during the growth phase. Glucose concentration is typically kept low during high cell density fermentation to prevent overflow metabolic production of ethanol resulting from the Crabtree effect, but this may lead to early induction of gene expression and growth defects. The CUP1 copper inducible promoter can be tuned across a 25-fold induction range by varying copper concentration in the growth medium (33). However, maximal induction occurs at copper concentrations that can have toxic effects on cell physiology, and basal expression in the absence of added copper could produce counterselection against gene amplification.

SIGNAL PEPTIDES

The most widely used leader peptide for heterologous secretion in *S. cerevisiae* is that of pre-pro α factor (34). This leader peptide is often found to be the most effective when compared with other signal peptides. The signal peptide is cleaved by signal peptidase after endoplasmic reticulum membrane translocation, and the pro region is cleaved C-terminal to a Lys-Arg site by the Kex2 protease in the Golgi apparatus. Another signal peptide that has been used frequently is the invertase signal peptide, which does not have an associated pro region.

POST-TRANSLATIONAL PROCESSING

Secreted proteins undergo a number of processing steps after translocation into the lumen of the endoplasmic reticulum: disulfide bond formation, folding, subunit assembly, glycosylation, and site-specific endoproteolytic processing. The endoplasmic reticulum and Golgi apparatus contain high concentrations of specialized chaperones, foldases, and enzymes dedicated to the task of conformational maturation of secreted proteins before release to the extracellular space (35). Many of these processing steps are performed inefficiently or not at all by bacteria, but yeast's secretory processes closely mirror those of mammalian cells (with the exception of glycosylation).

Protein Folding and Conformational Proofreading

Proteins do not exit the endoplasmic reticulum until they have folded correctly (36). The mechanistic details of this conformational proofreading process have not been completely elucidated, but they involve addition and trimming of glucose residues for glycoproteins and binding to the lectinlike chaperone calnexin. Two practical consequences of this proofreading step are of critical significance for heterologous protein production by secretion: (*1*) only correctly folded protein appears in the growth medium, obviating the need for refolding procedures; and (*2*) protein folding in the endoplasmic reticulum constitutes the key rate- and yield-limiting step in high-level heterologous protein secretion. Consequence *1* provides one of the major advantages of using yeast for protein production. Consequence *2* has motivated researchers to intensively study and manipulate the endoplasmic reticulum protein folding machinery in order to increase secretory capacity.

Most secreted eukaryotic proteins possess disulfide bonds. These covalent cross-links between cysteine side chains stabilize the folded conformation of a protein. As

the number of disulfides in a protein increases, the combinatorial possibilities for incorrect cysteine pairings during folding grows geometrically. The endoplasmic-reticulum-resident foldase enzyme protein disulfide isomerase (PDI) catalyzes thiol–disulfide exchange to shuffle among possible disulfide pairings until the most stable folded conformation is attained. Overexpression of PDI in yeast can significantly improve secretion of some heavily disulfide bonded proteins (37,38); it is not a panacea, however. Secretion of some proteins is unaffected by PDI overexpression despite possession of disulfide bonds (37).

Another endoplasmic-reticulum-resident chaperone whose function in heterologous secretion has been closely examined is BiP, an hsp70 homologue. Although fulfilling an essential role in membrane translocation and subunit assembly, BiP levels are clearly not limiting for heterologous secretion in yeast or mammalian cells. Less clear is whether reduction of BiP levels is beneficial for secretion, because secretion of three heterologous proteins was decreased in yeast by reduction of BiP levels (39), whereas secretion of some proteins has been increased in mammalian cells by reducing BiP (40). The likely source of this discrepancy lies in the proteins studied; those proteins whose secretion is increased at low BiP levels are either mutant proteins with folding defects or subunits lacking stabilizing partners for a complex. It is possible that overexpression of other endoplasmic-reticulum-resident soluble or membrane bound chaperones and foldases could prove fruitful for heterologous secretion; such candidates include GRP94 chaperones, peptidyl prolyl isomerases, and redox foldases related to PDI such as EUG1 and ERP72.

Kex2p Endoproteolytic Processing

The Golgi-resident Kex2p endoprotease cleaves polypeptides C-terminal to Lys-Arg in unstructured regions of the protein. Kex2p removes the α-factor pro region from the α-factor peptide and heterologous fusion proteins. The sequence requirements for the amino acids adjacent to the Lys-Arg cleavage site have not been completely determined. In many cases, cleavage is efficient and the protein's native N-terminus is produced. For some proteins, however, Kex2 cleavage is inefficient, and a dipeptide spacer must be inserted at the junction to achieve efficient cleavage (41). This dipeptide may be tolerable as part of the protein product, or alternatively it may be removed by cathepsin-C processing (42). Undesirable Kex2 cleavage may also occur at internal Lys-Arg sites of a protein (43). This problem may be circumvented either by deletion of the *KEX2* gene or by alteration of the Lys-Arg sites by site-directed mutagenesis.

Glycosylation

S. cerevisiae covalently attaches polysaccharide structures at the same Asn-X-Ser/Thr sites recognized by mammalian cells. Although the core glycosylation structures initially added in the ER are identical for yeast and mammals, yeast subsequently adds extended mannose chains that are not capped with sialic acid and can be immunogenic. Although hypermannosylation can be reduced or eliminated by use of a *mnnl mnn9* strain (44), the absence of mammalian-type outer-chain glycosylation makes all yeasts undesirable hosts for expression of therapeutic glycoproteins.

SUPERSECRETION MUTATIONS

A classical genetic approach to strain improvement is to screen for spontaneous mutations that increase secretion of the protein of interest. Although such mutants may be difficult to analyze because of mutations at multiple loci or dominant alleles, the practical value of this strategy can be significant (45–48). One of the best characterized supersecretion mutants is the *pmr1* mutation, which involves loss of function of a calcium pump localized in the Golgi (45). How this causes substantial increases in heterologous secretion is not clear.

HIGH CELL DENSITY FERMENTATION

S. cerevisiae can be grown to cell densities of more than 100 g cell dry weight per liter in industrial-scale fed-batch fermentation (49,50). In such fermentations, glucose feed can be controlled by measurement of the respiratory quotient to prevent overfeeding and generation of ethanol. In the absence of instrumentation to measure the respiratory

Table 2. Genetic Components of *S. cerevisiae* Expression/Secretion Systems

Vector	Selection marker	Promoter	Leader peptide
Integrating	Dominant	Inducible	Pre-pro α
Recombination	***G418*^{R}**	**GAL1**	Invertase
Retrotransposition	*CUP1*	ADH2	Acid phosphatase
2μ plasmid	Recessive	CUP1	
	URA3	Constitutive	
	TRP1	GAP	
	HIS3	PGK	
	LEU2	TPI	
	ADE2		
	Autoselection		
	ura3fur1		
	Δ*tpi*		

quotient, an exponentially increasing feed rate can also achieve high cell densities.

CONCLUSIONS AND RECOMMENDATIONS

Some of the most commonly used genetic components of heterologous secretion systems for *S. cerevisiae* are summarized in Table 2, and the recommended approach is highlighted in boldface. Amplifiable chromosomal integration vectors provide excellent mitotic stability and can improve secretion yields by an order of magnitude relative to 2μ plasmids. The *G418*^{R} dominant selectable marker enables tuning of gene amplification via increasing antibiotic resistance and allows use and optimization of complex media without auxotrophic nutritional constraints. Constitutive transcriptional promoters are not recommended because of counterselection against high-level expression and consequent strain instability. The GAL promoter provides a combination of near complete repression during a glucose fed growth phase, followed by strong induction upon feeding of galactose. The pre-pro α-factor leader peptide generally provides maximal targeting efficiency, and there is generally little benefit in examining several different leader peptides for comparison. The recommended combination of a δ integrating vector with *G418*^{R} selection, a GAL promoter, and pre-pro α-type leader peptide has been applied to secrete 400 mg/L of active bovine pancreatic trypsin inhibitor (BPTI) in shake flask cultures of *S. cerevisiae*, by comparison with 10 to 40 mg/L secreted with a constitutive promoter and 2μ plasmid (16). This specific productivity exceeds that reported for BPTI secretion in *P. pastoris* (4), demonstrating that efficient exploitation of *S. cerevisiae* can eliminate apparent differences among yeast species in their capacity for heterologous protein secretion.

BIBLIOGRAPHY

1. J.R. Broach, J.R. Pringle, and E.W. Jones eds., *The Molecular and Cellular Biology of the Yeast* Saccharomyces. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1991.
2. M. Romanos, *Curr. Opin. Biotechnol.* **6**, 527–533 (1995).
3. K.N. Faber, W. Harder, G. Ab, and M. Veenhuis, *Yeast* **11**, 1331–1344 (1995).
4. T. Vedvick, R.G. Buckholz, M. Engel, M. Urcan, J. Kinney, S. Provow, R.S. Siegel, and G.P. Thill, *J. Ind. Microb.* **7**, 197–202 (1991).
5. R.N. Parekh, M.R. Shaw, and K.D. Wittrup, *Biotechnol. Prog.* **12**, 16–21 (1996).
6. M.A. Romanos, C.A. Scorer, and J.J. Clare, *Yeast* **8**, 423–488 (1992).
7. J. Hodgson, *Bio / Technology* **11**, 887–893 (1993).
8. D.M. Goldenberg, *Am. J. Med.* **94**, 297–312 (1993).
9. A.B. Rose and J.R. Broach, *Methods Enzymol.* **185**, 234–279 (1990).
10. E. Erhart and C.P. Hollenberg, *J. Bacteriol.* **156**, 625–635 (1983).
11. G. Loison, A. Vidal, A. Findeli, C. Roitsch, J.M. Balloul, and Y. Lemoine, *Yeast* **5**, 497–507 (1989).
12. R. Rothstein, *Methods Enzymol.* **194**, 281–301 (1991).
13. T.S. Lopes, J. Klootwijk, A.E. Veenstra, P.C. van der Aar, H. van Heerikhuizen, H.A. Raue, and R.J. Planta, *Gene* **79**, 199–206 (1989).
14. J.R. Shuster, H. Lee, and D.L. Moyer, *Yeast* **6**, 579 (1990).
15. A. Sakai, F. Ozawa, T. Higashizaki, Y. Shimizu, and F. Hishinuma, *Bio / Technology* **9**, 1382–1385 (1991).
16. R.N. Parekh, M.R. Shaw, and K.D. Wittrup, *Biotechnol. Prog.* **12**, 16–21 (1996).
17. N. Shiomi, H. Fukuda, F. Murata, and A. Kimura, *Appl. Microbiol. Biotechnol.* **42**, 730–733 (1995).
18. X. Wang, Z. Wang, and N.A. DaSilva, *Biotechnol. Bioeng.* **49**, 45–51 (1996).
19. D. Mochizuki, K. Miyahara, D. Hirata, H. Matsuzaki, T. Hatano, S. Fukui, and T. Miyakawa, *J. Ferment. Bioeng.* **77**, 468–473 (1994).
20. T. Fujii, K. Kondo, F. Shimizu, H. Sone, J.-I. Tanaka, and T. Inoue, *Appl. Env. Microbiol.* **56**, 997–1003 (1990).
21. J.D. Boeke, H. Xu, and G.R. Fink, *Science* **239**, 280–282 (1988).
22. X. Wang and N.A. DaSilva, *Biotechnol. Bioeng.* **51**, 703–713 (1996).
23. T.D. Webster and R.C. Dickson, *Gene* **26**, 243–252 (1983).
24. R.C.A. Henderson, B.S. Cox, and R. Tubb, *Curr. Genet.* **9**, 133–138 (1985).
25. G. Loison, M. Nguyen-Juilleret, F. Alouani, and M. Marquet, *Bio / Technology* **4**, 433–437 (1986).
26. Eur. Pat. 017,114 (1986), G. Kawasaki.
27. S.M. Kingsman, D. Cousens, C.A. Stanway, A. Chambers, M. Wilson, and A.J. Kingsman, *Methods Enzymol.* **185**, 329–341 (1990).
28. S. Rosenberg, D. Coit, and P. Tekamp-Olson, *Methods Enzymol.* **185**, 341–351 (1990).
29. S. Elliott, J. Giffin, S. Suggs, E.P. Lau, and A.R. Banks, *Gene* **79**, 167–180 (1989).
30. A.S. Robinson and K.D. Wittrup, *Biotechnol. Prog.* **11**, 171–177 (1995).
31. L.M. Mylin, K.J. Hofmann, L.D. Schultz, and J.E. Hopper, *Methods Enzymol.* **185**, 297–308 (1990).
32. V.L. Price, W.E. Taylor, W. Clevenger, M. Worthington, and E.T. Young, *Methods Enzymol.* **185**, 308–318 (1990).
33. T. Etcheverry, *Methods Enzymol.* **185**, 319–329 (1990).
34. A.J. Brake, *Methods Enzymol.* **185**, 408–420 (1990).
35. A. Helenius, T. Marquardt, and I. Braakman, *Trends Cell Biol.* **2**, 227–231 (1992).
36. C. Hammond and A. Helenius, *Curr. Opin. Cell Biol.* **7**, 523–529 (1995).
37. A. Robinson, V. Hines, and K.D. Wittrup, *Bio / Technology* **12**, 381–384 (1994).
38. L. Schultz, H. Markus, K. Hofmann, D. Montgomery, C. Dunwiddie, P. Kniskern, R. Freedman, R. Ellis, and M. Tuite, *Ann. N.Y. Acad. Sci.* **721**, 148–157 (1994).
39. A.S. Robinson, J.A. Bockhaus, A.C. Voegler, and K.D. Wittrup, *J. Biol. Chem.* **271**, 10017–10022 (1996).
40. A.J. Dorner and R.J. Kaufman, *Biologicals*, **22**, 103–112 (1994).
41. R.N. Parekh, K.J. Forrester, and K.D. Wittrup, *Protein Exp. Purif.* **6**, 537–545 (1995).
42. C. Lauritzen, E. Tuchsen, P.E. Hansen, and O. Skovgaard, *Protein Exp. Purif.* **2**, 372–378 (1991).
43. O.S. Gabrielsen, R. Reppe, O. Saether, O.R. Blingsmo, K. Sletten, J.O. Gordeladze, A. Hogset, V.T. Gautvik, P. Alestrom, T.B. Oyen, and K.M. Gautvik, *Gene* **90**, 255–262 (1990).

44. C.E. Ballou, *Methods Enzymol.* **185**, 440–470 (1990).
45. R.A. Smith, M.J. Duncan, and D.T. Moir, *Science* **230**, 1219–1224 (1985).
46. J.R. Shuster, D.L. Moyer, H. Lee, A. Dennis, B. Smith, and J.P. Merryweather, *Gene* **83**, 47–55 (1989).
47. G.S. Gill, P.G. Zaworski, K.R. Marotti, and E.F. Renberg, *Bio/Technology* **8**, 956–958 (1990).
48. V. Chisholm, C.Y. Chen, N.J. Simpson, and R.A. Hitzeman, *Methods Enzymol.* **185**, 471–482 (1990).
49. G.M. O'Connor, F. Sanchez-Riera, and C.L. Cooney, *Biotechnol. Bioeng.* **39**, 293–304 (1992).
50. O. Mendoza-Vega, J. Sabatie, and S.W. Brown, *FEMS Microb. Rev.* **15**, 369–410 (1994).

See also ENZYMES, DIRECTED EVOLUTION; ENZYMES, EXTREMELY THERMOSTABLE; SECRETION FROM ANIMAL CELLS.

PROTEIN ULTRAFILTRATION

ROBERT VAN REIS
Genentech, Inc.
South San Francisco, California
ANDREW L. ZYDNEY
University of Delaware
Newark, Delaware

KEY WORDS

Buffer Exchange
Concentration
Cross Flow Filtration
Desalting
Diafiltration
Filtration
Membrane
Protein Purification
Tangential Flow Filtration
Ultrafiltration

OUTLINE

INTRODUCTION

Protein ultrafiltration (UF) is a pressure-driven membrane process used for the concentration or purification of protein solutions. UF membranes typically have mean pore size between 10 and 500 Å, which is intermediate between reverse osmosis (RO) and microfiltration (MF). Although UF has often been viewed as a purely size-based separation, with species larger than the membrane pores being fully retained while smaller species pass freely, separation in UF actually occurs because of differences in the rate of filtration of different components across the membrane in response to a given pressure driving force. Solute filtration rates, and thus the membrane selectivity, are determined by both thermodynamic (e.g., partitioning) and hydrodynamic (e.g., transport) interactions. UF is currently used throughout downstream processing for protein concentration, buffer exchange and desalting, protein purification, virus clearance, and clarification. Buffer exchange and desalting are typically accomplished using a diafiltration (DF) mode in which the low molecular weight components are washed away from the protein by simultaneously adding fresh buffer (or solvent) to the feed during UF. Protein purification is accomplished using high-performance tangential flow filtration (HPTFF), with the product collected in either the retentate or filtrate depending on the relative filtration rates. Virus filtration (VF) uses larger pore UF membranes that are permeable to protein but provide significant virus removal.

PRINCIPLES OF ULTRAFILTRATION AND HIGH-PERFORMANCE TANGENTIAL FLOW FILTRATION

Background and Theory

General Operating Principles. Although protein UF has been used in the laboratory since the turn of the century, large-scale industrial applications of UF only date back to about 1970. The key breakthrough was the development of the asymmetric membrane by Loeb and Sourirajan and the subsequent extension of this technique to the production of UF membranes by Michaels and colleagues at Amicon. These asymmetric membranes have a very thin skin (approximately 0.5 μm thick) that provides the membrane its selectivity and a more macroporous substructure that provides the required mechanical and structural integrity. The thin skin results in much higher filtration rates than obtainable with homogeneous membranes, significantly reducing the required membrane area or process time. Typical process fluxes in UF range from 25 to 250 L m^{-2} h^{-1}. Operating pressures are generally in the range of 0.2 to 4 bar, which are much lower than the 25 to 80 bar required in RO to overcome the large osmotic pressure of the retained salts. UF membranes are available with a variety of pore sizes and can be cast from a range of base polymers, each having unique physical and chemical characteristics.

Most large-scale UF devices use tangential flow filtration (TFF), also referred to as a cross-flow configuration, in which the feed flow is parallel to the membrane and thus perpendicular to the filtrate flow (1). This allows retained species to be swept along the membrane surface and out

the device exit, significantly increasing the process flux compared to that obtained with dead-end operation. A concentration boundary layer consisting of a high concentration of retained solutes typically develops at the upstream surface of the membrane. This highly concentrated region reduces the effective pressure driving force and can cause extensive membrane fouling from protein adsorption, denaturation, precipitation, or aggregation. The performance of many TFF systems is determined almost entirely by the rate at which these retained solutes are transported away from the membrane and back into the bulk solution, a phenomenon referred to as concentration polarization. A variety of membrane modules have been developed to achieve the desired mass transport rates while maintaining high membrane packing densities, and these are discussed in some detail in "Modules and Devices". Most units are produced in modular form to allow easy replacement and scale-up. Membrane devices are available with surface areas as small as 50 cm^2 for bench-top operation to more than 200 m^2 for large-scale processing. Dead-end (or normal flow) filtration is used primarily for laboratory-scale separations and for systems in which the retained species are present at very low concentration. For example, dead-end MF cartridges have been used extensively for sterile filtration, in which the retained species are present at very low concentration. Similar modules can be used for virus removal.

Conventional UF has traditionally been used to separate solutes that differ by more than 10-fold in size, making it ideal for protein concentration and buffer exchange (2). UF processes are thus used throughout the downstream purification, often to condition the product before or immediately after a chromatographic column. The fairly limited selectivity of UF has generally been attributed to the wide pore size distributions in commercial membranes, the presence of significant bulk mass transfer limitations in existing modules, and membrane fouling phenomena. Recent work (3) has demonstrated that membrane systems can be used for protein–protein separations, even for proteins with very similar molecular weight, using a process known as HPTFF. High selectivity values are obtained in HPTFF by careful choice of both membrane and buffer conditions, with the latter chosen to maximize the difference in hydrodynamic volume for the two proteins. Concentration polarization is exploited to enhance, rather than limit, solute separation. Membrane fouling is controlled using appropriate fluid dynamic and mass transfer conditions. High yields and purification factors are obtained using a diafiltration mode. The net result is that HPTFF provides a competitive separation process that complements chromatographic techniques for protein purification.

Membrane Transport. Theoretical analyses of solute and solvent transport through UF membranes are generally developed using either the Kedem-Katchalsky or Stefan-Maxwell descriptions of irreversible thermodynamics (4). The filtrate flux J (equal to the volumetric filtration rate normalized by the membrane surface area) is proportional to the effective pressure driving force (5):

$$J = L_p[\Delta P - \Sigma(\sigma_i \Delta \Pi_i)] \quad (1)$$

where L_p is the overall permeability (equal to the reciprocal of the overall resistance), and ΔP is the transmembrane pressure. L_p is only equal to the clean membrane permeability (evaluated from water flux data obtained with the clean membrane) in the absence of irreversible fouling. The osmotic reflection coefficient (σ_i) provides a measure of the membrane selectivity: $\sigma_i = 1$ for a solute that is completely retained and $\sigma_i = 0$ for a solute that is completely permeable. The osmotic pressure difference across the membrane ($\Delta \Pi_i = \Pi_w - \Pi_f$) is a function of the solute concentration at the upstream surface (C_w) and in the filtrate solution (C_f). Although protein osmotic pressures are typically small for $C < 10$ g/L, $\Delta \Pi_i$ can be comparable to ΔP during UF because of the buildup of retained protein at the membrane surface (see "Concentration Polarization").

The solute flux N_s (in units of grams protein collected in the filtrate per membrane surface area per time) is conveniently expressed as (6)

$$N_s = J S_\infty C_w \quad (2)$$

where S_∞ is the intrinsic (or asymptotic) sieving coefficient, which is approximately equal to $1 - \sigma_i$ based on Onsager reciprocity. Equation 2 is valid only under conditions where the solute flux is dominated by convection; a more detailed analysis accounting for the coupling between convection and diffusion is given by Opong and Zydney (6). The asymptotic sieving coefficient is determined by thermodynamic partitioning of the solute into the membrane and the relative rate of solute transport through the pore (7):

$$S_\infty = \phi K_c = C_f / C_w \quad (3)$$

where K_c is the hydrodynamic hindrance factor, C_f is the filtrate concentration, and C_w is the retentate concentration at the membrane wall. The hydrodynamic hindrance factor accounts for the additional drag on the solute because of the presence of the pore wall. Theoretical analyses for a hard sphere in a cylindrical pore indicate that K_c only varies between 1.00 and 1.47 (8); thus S_∞ is determined primarily by the equilibrium partition coefficient (ϕ), which is a function of both steric (size) and long-range (e.g., electrostatic) interactions. The electrostatic energy of interaction between a protein and a pore has three distinct contributions (9). In UF systems, the dominant factor is typically the charge on the solute (σ_s). The ϕ decreases with increasing σ_s because of the increase in free energy associated with the distortion of the electrical double layer surrounding the protein caused by the presence of the pore wall. This can be interpreted as an increase in effective protein size (or hydrodynamic volume) associated with the diffuse ion cloud (10). This effect is the basis for the 100-fold decrease in S_∞ with decreasing salt concentration seen for bovine serum albumin (11). There are also contributions from the double layer surrounding the pore wall (which reduces the effective pore size) and from direct charge–charge interactions, with the latter causing an increase in ϕ when the solute and pore have opposite charge.

In the absence of long-range interactions, the sieving coefficient for a spherical solute in a cylindrical pore is (7)

$$S_\infty = (1 - \lambda)^2 \quad (4)$$

where λ is the ratio of the solute (r_s) to pore radius (r_p). Thus, even a membrane with a single pore size has a gradual decrease in S_∞ because of the reduction in accessible pore volume with increasing r_s. The retention characteristics of commercial UF membranes are more complex because of the broad distribution of irregularly shaped pores. A simple approximate analysis yields (12)

$$S_\infty = \exp(-r_s/s) \quad (5)$$

where s provides a measure of the mean pore size. In this case, S_∞ approaches zero asymptotically because of the presence of a small number of very large pores. The effects of different pore size distributions on S_∞ are discussed elsewhere (12).

Concentration Polarization. One of the critical factors determining the overall performance of any UF device is the rate of protein transport in the bulk solution adjacent to the membrane. The filtrate flow causes an accumulation of partially (or completely) retained solute at the upstream surface of the membrane, a phenomenon referred to as concentration polarization. The solute concentration thus varies from its maximum value at the membrane surface (C_w) to its bulk value (C_b) over the thickness of the concentration boundary layer (δ). Most analyses of concentration polarization have used the simple stagnant film model originally presented by Michaels (13):

$$J = k \ln([C_w - C_f]/[C_b - C_f]) \quad (6)$$

where the solute mass transfer coefficient (k) is equal to the ratio of the solute diffusivity (D) to the boundary layer thickness (δ). More rigorous analyses of solute transport indicate that equation 6 is valid at low degrees of concentration polarization, with the actual concentration dependence becoming more complex at high C_w/C_b (14). The wall concentration increases with increasing filtrate flux, causing an increase in $\Delta\Pi$ and a reduction in the effective pressure driving force (equation 1). This negative feedback balances any increase in applied pressure, with the net result that the flux approaches a nearly constant, pressure-independent value at high ΔP (Fig. 1). High wall concentrations can also reduce L_p through irreversible fouling or the formation of a protein gel on the membrane surface. In the latter case, the pressure-independent flux is determined by the maximum solubility or gel concentration for the protein. Although this gel polarization model has been used quite extensively, the calculated values of C_w are device-dependent and are often in poor agreement with independent estimates of protein solubility (8).

Theoretical analyses of solute mass transfer coefficients (k) in laminar flow are generally based on the Leveque solution (1):

$$k = 0.807(D^2\gamma/L)^{1/3} \quad (7)$$

which predicts a two-thirds power dependence on the protein diffusivity (D) and a one-third power dependence on shear rate (γ). Empirical correlations for k have been developed from pressure drop data using the Chilton-Colburn analogy with a power law dependence on feed flow rate around 0.5 for spacer-filled channels and near 1.0 for turbulent flow (8). The most accurate method for evaluating the mass transfer coefficient, particularly for modules with complex fluid flow characteristics (see "Modules and Devices") is to fit equation 6 to actual filtrate flux data. The simplest approach is to evaluate k from the slope of a plot of the pressure-independent flux as a function of $\ln(C_b)$ (assuming a fully retentive membrane). This requires minimal experimental data, but the results can be in considerable error because of uncertainties in the limiting flux. In addition, the resulting value for k is only valid under conditions of very high polarization; the behavior at lower degrees of polarization can be considerably different because of the concentration dependence of the solution viscosity and protein diffusivity (8). The mass transfer coefficient can also be evaluated by fitting filtrate flux versus ΔP data to equations 1 and 6 simultaneously, which was the approach used to generate the solid curves in Figure 1. This requires an estimate of the process membrane permeability, which can be evaluated from buffer flux data obtained with the fouled membrane, and an expression for the protein osmotic pressure as a function of C_w. The osmotic pressure can be evaluated independently, or it can be described using a simple polynomial expansion, in which case k and the osmotic coefficients are fit simultaneously. van Reis et al. (15) developed a novel approach that eliminates the need to assume a specific functional relationship for $\Delta\Pi$ by simultaneously analyzing J versus ΔP data obtained at different C_b.

The accumulation of retained solute at the membrane also affects protein retention. Equation 6 can be rearranged to evaluate the observed protein sieving coefficient using $N_s = JC_f$, yielding (6)

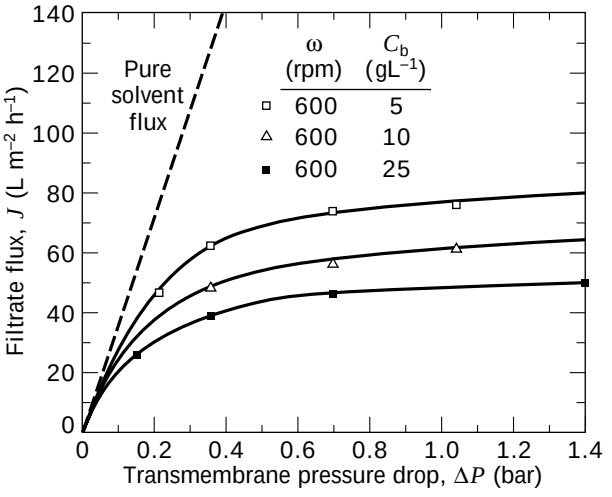


Figure 1. Filtrate flux as a function of transmembrane pressure for stirred-cell UF of BSA solutions of different bulk concentrations (C_b) using a constant stir speed (ω). Solid curves are model calculations based on equations 1 and 6. *Source*: Figure adapted from Opong and Zydney (6).

$$S_o = C_f/C_b = S_\infty/((1 - S_\infty) \exp(-J/k) + S_\infty) \quad (8)$$

At low filtrate flux, $S_o = S_\infty$ because C_w is approximately equal to C_b. The observed sieving coefficient increases with increasing flux, approaching a value of 1 (i.e., $C_f = C_b$) at very high flux as long as the membrane is at least partially permeable to the solute and the required wall concentration ($C_w = C_b/S_\infty$) is physically attainable. Typical experimental results for bovine serum albumin are shown in Figure 2. The solid curve is the model calculation using equation 8. The slight increase in S_o at very low filtrate flux is because of the diffusive contribution to protein transport through the membrane (6).

Membrane Properties

Membrane Chemistry and Morphology. Almost all UF membranes have an anisotropic or asymmetric structure, with the membrane selectivity determined by a thin skin layer and the mechanical integrity provided by a much thicker macroporous support. Most asymmetric UF membranes are prepared by immersion casting (8). A thin film of a polymer solution is spread on a suitable porous or woven support (e.g., polyethylene). The cast film is then immersed in a nonsolvent (often water). Diffusional exchange occurs between the solvent and nonsolvent, causing phase separation of the polymer solution and subsequent gelation of the polymer-rich phase. This gives the membrane a nodular morphology with the pore size determined by the size and spacing of the polymer-rich globules. The asymmetric structure develops because of mass transfer (kinetic) effects, such as growth and coalescence of polymer-rich globules before gelation (8). Large tear-shaped macrovoids often form beneath the membrane skin because of interfacial instabilities or the growth of polymer-lean regions in the film (16). Membranes with macrovoids have high permeability but lower maximum pressure limits. More importantly, defects can form if the macrovoids extend up to (or through) the skin. Hollow fiber membranes are also made by phase inversion with the polymer solution extruded through an orifice into a gelation bath while a second gelation medium is pumped into the fiber lumen. Skin layer(s) develop on the inner or outer surface, depending on the properties of the gelation media. Composite membranes are produced using multiple casting steps. For example, very uniform UF membranes that are free of macrovoids can be produced by casting a thin (5 to 10 μm) polymer film on a MF substrate (17). These membranes are particularly attractive for virus filtration where the presence of even trace defects leads to poor virus removal.

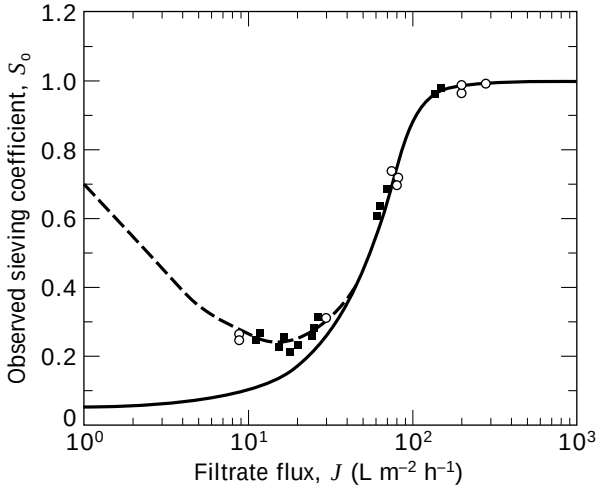


Figure 2. Observed sieving coefficient as a function of filtrate flux for the UF of BSA through a 100-kD poly(ether sulfone) membrane. Solid curve is model calculation based on equation 8. Dashed curve accounts for diffusive contribution. *Source*: Figure adapted from Opong and Zydney (6).

UF membranes can be cast from a wide range of polymers, including polysulfone, poly(ethersulfone), polyamides, polyimides, poly(vinyl chloride), polyacrylonitrile, poly(vinylidene fluoride), regenerated cellulose, and cellulose acetate. Cellulosic membranes bind the least amount of protein but can be damaged by harsh cleaning protocols. The synthetic polymers have much greater chemical and thermal stability but are more susceptible to protein fouling. Ceramic membranes, produced by slip-casting of Al_2O_3 or acid leaching of SiO_2 (18), have very high thermal and chemical stability but tend to be considerably more expensive than polymeric membranes. Many membranes are modified during or after casting by inclusion of polyvinylpyrrolidone in the initial polymer solution, sulfonation of cast polysulfone, radiation-grafting of methacrylate monomers to polyamides, or coating with hydroxypropyl cellulose (19). These modifications increase membrane hydrophilicity and alter membrane surface charge, but they can reduce long-term chemical and thermal stability.

Membrane Characterization. Information on UF membrane characterization most commonly provided to end users consists of water flux measurements, solute retention tests, selected protein binding data, chemical compatibility, and thermal limits. Water flux data provide a measure of the membrane permeability as determined by evaluating the filtrate flux (filtration rate normalized for surface area) as a function of the transmembrane pressure (ΔP). Normal flow filtration devices yield a linear relationship between flux (J) and transmembrane pressure as long as membrane compaction is avoided. The situation is more complex in TFF because the conversion of feed flow to filtrate flow creates a nonlinear pressure drop through the feed channel. A correct value for the permeability of a membrane in a TFF device can be obtained by measuring flux versus transmembrane pressure and extrapolating $J/\Delta P$ versus ΔP to $\Delta P = 0$. Membrane permeability is primarily determined by pore size distribution, porosity (pore density), membrane thickness, and solvent viscosity. Several issues should be considered when judging membrane permeability. The type and quality of water used for the tests can have a significant influence on the results. The pH and ionic strength of the water can influence the flux because of charge interactions, and contaminants can foul the membrane during the test. It is also important to carefully control and report the temperature of the test because permeability is linear with viscosity. As previously men-

tioned, membrane compaction and device configuration can also play a significant role in these tests. It should be noted that membrane water permeability data are of minimal value unless used in conjunction with other membrane characterization information. Water permeability data are not predictive of process fluxes because of significant differences in fouling behavior of various membrane chemistries. Polysulfone membranes typically have much higher water permeability values than regenerated cellulose membranes, whereas the latter membranes typically have higher process permeability values as a result of their inherent low fouling characteristics with most biotechnology feedstreams.

It is important to consider the intended use of membranes when selecting a retention test. The process objectives, and hence the criteria for membrane selection, are very different for protein concentration, buffer exchange, virus–protein separations, clarification, or protein purification. Retention testing of UF membranes involves measurements of solute transmission. Sieving coefficients (S) or retention coefficients (R) are determined by analyzing the solute in the feed and filtrate streams under specific process conditions.

$$S = C_f/C_F \tag{9}$$

$$R = 1 - S \tag{10}$$

Tests are carried out using either single solutes or multiple solutes in separate experiments. The most commonly used single solutes are polymers [such as poly(vinylpyrrolidone)], proteins and viruses. The purpose of the retention tests are to provide nominal molecular weight cutoff values for each membrane. Single-solute retention tests are the easiest to perform and provide either the least or most useful data, depending on the chosen solute and test conditions. Single-solute tests performed by membrane manufacturers typically involve the use of one or several experiments with solutes that are partially retained by the membrane. This information is of limited value for protein concentration and buffer exchange applications because the coefficients of partially retained solutes are not predictive of the retention values one would obtain for a highly retained solute. The retention of solutes with $R = 0.3$ and $R = 0.9$, for example, do not provide guidance on the molecular weight of a solute with a retention coefficient of $R = 0.999$, which is what one would strive for in a protein concentration–diafiltration operation. In all applications (UF, VF, and HPTFF), retention coefficients will also be highly dependent on fouling, buffer chemistry, and fluid dynamics. Furthermore, there is a lack of standardization of retention testing protocols between manufacturers. Different solutes, buffers, and fluid dynamic conditions are used. In addition, the protocols for assigning nominal molecular weight designations (for example, the molecular weight of a solute that has a retention coefficient of $R = 0.9$) are not even standardized within a given company. It is therefore imperative that the end users perform their own retention testing on the solutes of interest under conditions that are a linear scale-down of the intended process.

The mixed dextran test (20) provides an alternative methodology for membrane characterization. A mixture of several polydisperse dextrans are used with the feed and filtrate streams analyzed using size exclusion chromatography (SEC) with an on-line refractive index (RI) detector. Column calibration is performed using several narrow molecular weight distribution standards. The SEC chromatograms represent the solute concentrations (RI absorbance) as a function of dextran molecular weight (elution time). A continuous retention curve (retention versus molecular weight) can hence be generated by constructing a curve using equations 9 and 10 with the feed and filtrate concentrations represented by the respective chromatograms. Because the SEC chromatograms provide elution times for every molecular weight fraction, it is not necessary to use dextrans with tight molecular weight distributions. Mixed dextran tests performed under conditions of constant mass transfer coefficient and flux are a useful tool in comparing the relative retention characteristics of various UF membranes. The method is a good quality-control test for both membrane manufacturers and end users. The method can also be used to derive information on pore size distributions (21). Some predictive power can also be derived from this methodology by converting both dextran and protein molecular weights to equivalent hydrodynamic volumes (22).

A more sensitive method of characterizing UF membranes involves the principle of liquid–liquid pore intrusion. Two highly immiscible liquids, such as solutions of a sulfate salt and a poly(ethylene glycol) are contacted through mixing to reach equilibrium partitioning. The membrane to be tested is primed with one of the liquids so that all the pores are filled. After draining the feed channels, the second fluid is introduced into the system. The first fluid is then displaced out of the pores by the second fluid, and the flow rate is measured as a function of transmembrane pressure. The resulting data provide information on pore size distribution and can be correlated with the nominal molecular weight cutoff. It has also been demonstrated that this methodology can be used as an excellent quality-control test for membranes used in HPTFF applications because performance can be correlated with protein sieving data (23).

Protein binding studies are often performed to contrast the relative behavior of various membrane chemistries. The data generated in these studies can be useful as a general guide. It is important, however, to recognize that protein adsorption will be highly dependent on pH, ionic strength, and the specific isoelectric point and hydrophobicity of a particular protein. It is also useful to note that most membranes will not cause significant protein losses from adsorption because of the relatively low binding capacities (1 to 10 mg/m^2), the small surface areas required (0.01 m^2/g), and the high protein concentrations (0.05 to 1 g/L) obtained in modern biotechnology processes. This is even true for processes such as MF, VF, and HPTFF, in which the product may be exposed to the internal surface area, which is 10 to 500 times the frontal membrane area. Chemical and thermal limits for membranes are determined by characterizing the membranes (using the techniques described herein) both before and after controlled

exposure. Membrane manufacturers often perform additional tests to determine pore size distribution, porosity, morphology, thickness, charge, process flux, and mechanical strength (compaction, tensile strength, and delamination). Pore size distribution and porosity can be determined by either mixed dextran, liquid–liquid intrusion, mercury intrusion, air–liquid intrusion (for large pore sizes), thermoporometry, scanning electron microscopy, or atomic force microscopy. Membrane morphology and thickness are determined by either scanning electron microscopy or transmission electron microscopy. Membrane charge can be determined by streaming potential measurements and is of particular use in HPTFF. Process fluxes provide some complementary information to the water permeability values provided by manufacturers but are best carried out by the end user on the specific application of interest. Determination of mechanical strength should include testing of membrane compaction, tensile strength, and ability to withstand reverse pressure. Soft membranes are susceptible to membrane compaction under process conditions. Membrane compaction can easily be determined by measuring water permeability as a function of transmembrane pressure throughout the desirable ΔP range. Tensile strength is most relevant to the membrane manufacturers and is analyzed using specialized equipment that measures the elongation of the membrane as a function of applied force. The reverse pressure limits of membranes are important to the end user, because unintended pressure spikes may occur during processing with resultant delamination of the membrane and subsequent product loss. Testing is done by subjecting the membrane to reverse pressure cycles at various pressures, followed by retention testing as outlined earlier.

Integrity testing of UF membranes can be performed by bubble-point measurement, air-diffusion flow measurement, or the liquid–liquid intrusion method just described. Bubble-point determinations involve measuring a gas flow rate through a membrane at increasing levels of pressure. The bubble point is determined by an inflection in the flow rate versus pressure curve indicative of liquid displacement out of the pores. Bubble-point measurements are primarily used for MF membranes that have bubble point values in the 1 to 6 bar range (for 0.65 to 0.1 μm pore sizes) using water as solvent. Bubble-point measurements of UF membranes are only practical when using much lower surface tension solvents, because the bubble point values using water as a solvent often exceed the pressure ratings of either the membrane or device. UF membrane integrity is therefore commonly performed using air (or nitrogen) diffusion methods. Air diffusion measurements are done at a single pressure and strive to detect abnormally high flow rates because of convective flow through membrane defects that are large enough to permit the bubble point to be exceeded.

Membrane Fouling and Regeneration. Membrane fouling is a critical factor in almost all UF processes. Fouling refers to the irreversible alteration in membrane properties caused by specific physical or chemical interactions between the membrane and the various components present in the feed. Fouling can occur because of protein adsorption, deposition, precipitation, denaturation, and aggregation, all of which affect membrane performance as well as product yield and quality. The rate and extent of fouling are determined by the morphological and chemical characteristics of the membrane, the fluid mechanics in the module, and the buffer chemistry (24). In addition, protein aggregation or denaturation can occur during repeated pump passes, with these species then deposited on the membrane during UF. Static protein adsorption (in the absence of filtration) generally attains monolayer values on the accessible membrane surface (25), with less adsorption on hydrophilic membranes. As discussed previously, adsorptive losses are small even when the protein has access to the internal pore surface. The convective filtrate flow during UF causes significant additional fouling that can be well beyond that associated with simple adsorption. Thus, considerable care must be taken in predicting fouling levels from static protein binding data. The initial fouling during MF is often caused by physical deposition of large protein aggregates on the membrane surface (26). Protein fouling in UF is strongly influenced by the magnitude of the local protein concentration at the membrane surface, with high values of C_w leading to increased adsorption, precipitation, aggregation, or denaturation. Protein fouling is generally maximum near the protein isoelectric point (27) and is enhanced under conditions where the protein and membrane have opposite charge. The extent of fouling is very sensitive to start-up conditions, with less fouling seen when the system is slowly ramped to the desired filtrate flux or transmembrane pressure (3).

Membrane cleaning or regeneration is performed whenever the flux or selectivity drop below some minimally acceptable level. In addition, membrane disinfection and sterilization must be performed between process runs even in the absence of significant fouling. Cleaning can be accomplished by physically removing the foulants from the membrane, for example, by fluid recirculation at zero transmembrane pressure, backflushing with negative ΔP, or mechanical scrubbing (generally limited to large diameter tubular modules). Chemical cleaning is done using appropriate detergents, acids, alkalis, enzymes, or chelating compounds. These cleaners must effectively displace, solubilize, or chemically modify the foulants while not exceeding the mechanical, thermal, or chemical limits of the membrane polymer or module. For example, cellulosic membranes are highly susceptible to damage by extremes of pH and temperature (28). Polysulfone and poly(ether sulfone) membranes are much more stable, with some manufacturers reporting acceptable performance even after limited exposure to temperatures as high as 125 °C (29). These membranes can also be operated over a wide range of pH (1 to 13) and can tolerate low levels of chlorine (up to 200 ppm). Alkali cleaners (e.g., NaOH) are particularly effective for removing biological foulants (30). Many alkali cleaners are used in combination with chelating agents (e.g., citrate or EDTA) to remove free Ca^{2+} and Mg^{2+} to prevent precipitation of saponified fats and oils (8). Detergents such as SDS are also effective in many systems. Enzymatic cleaners can be used for membranes that are unable to withstand elevated temperatures, strong chemicals, or pH extremes, but these solutions tend to be

very expensive and care must be taken to ensure complete removal of residual proteases. Membrane cleaning remains very much an art, with the development of an optimal cleaning cycle for a given application determined largely by trial and error.

Sodium hypochlorite (NaOCl) is used extensively for chemical disinfection of membrane systems. Hypochlorite is most effective at low pH, but it is very corrosive under these conditions (31). Disinfection is typically performed at pH≈10 using fairly concentrated solutions (0.25 to 4% NaOCl). Other oxidizing agents used for membrane disinfection include hydrogen peroxide (H_2O_2) and peracetic acid ($CH_3COCOOH$) (32). Peracetic acid is attractive because it is fast-acting, and it decomposes into the nontoxic acetic acid. Sodium bisulfite ($NaHSO_3$), a strong reducing agent, can be used for disinfection of membrane systems that would be degraded by strong oxidizing agents such as hypochlorite (33). Steam sterilization can be performed on disassembled membrane systems in an autoclave, but the membrane unit must then be aseptically attached to the remainder of the presterilized system (34). Steam-in-place (SIP) is the preferred method of sterilization, with the entire unit exposed to flowing steam as part of the completely assembled filtration system. Minimum requirements for an effective steam sterilization are 15 min exposure to steam at 121 °C and 1 bar pressure. Effective validation of SIP procedures is essential (35). Membrane systems can also be sterilized using ethylene oxide (EtO) or chlorofluorocarbons (CFCs), but these gas sterilization procedures require fairly long incubation periods to ensure complete deactivation of all microorganisms and spores (35). Gamma irradiation can also be used, although some membrane materials (e.g., PTFE and cellulosics) are damaged on exposure to high radiation doses (35).

Modules and Devices

The membrane module is the physical unit that houses the UF membrane in an appropriately designed configuration. The module must provide physical separation of the retentate and filtrate streams, mechanical support for the membrane, high membrane packing densities, easy access for cleaning or replacement, and scalability. In addition, the feed channel must provide adequate bulk mass transfer rates without excessive feed flow rates. Many of these criteria are to a large extent contradictory, for example, devices with large membrane packing densities are generally more susceptible to particulate fouling and more difficult to clean. A number of commercial UF modules have been developed, differing primarily in the size and shape of the feed (and filtrate) flow channels. These are conveniently classified into five distinct types: flat-sheet cassettes, spiral-wound modules, hollow-fiber cartridges, tubular modules, and enhanced mass transfer modules. The latter category includes a wide range of system geometries and operating strategies. Additional information on the economics, fluid mechanics, and mass transfer characteristics of these modules are available in Zeman and Zydney (8).

Flat-Sheet Cassettes. Flat-sheet (also known as plate-and-frame) cassettes were among the earliest configurations developed for large-scale commercial applications. The module uses a support plate (which also defines the filtrate flow path), membrane, and channel spacer (which defines the retentate flow path) in a sandwich configuration (Fig. 3). The membranes can be sealed to the plates using gaskets, in which case the module (stack) is hydraulically clamped to form a tight fit. Alternatively, the membranes can be directly bonded or glued to the plates using heat sealing or an appropriate adhesive to form an integral membrane element. Several of these elements or plates are then stacked together and clamped to form a complete module. The feed flow is distributed among the different channels at one end of the device, and the retentate is collected at the opposite end. The filtrate is manifolded separately from the feedstream and can be combined within or outside the module depending on the particular device design. The feed channel is typically 0.03 to 0.1 cm in height. The channel can either be fully open or it can use an appropriate screen (typically a polypropylene mesh) to promote local mixing and improve mass transfer. The small channel height leads to relatively high membrane packing densities (generally 300 to 500 m^{-1}) and low hold-up volumes, which is particularly attractive for recovery of high-value products. Open channel cassettes are less susceptible to plugging and operate with smaller pressure drops (feed minus retentate pressure). However, backflushing may not be possible because the membranes are effectively supported only on one side.

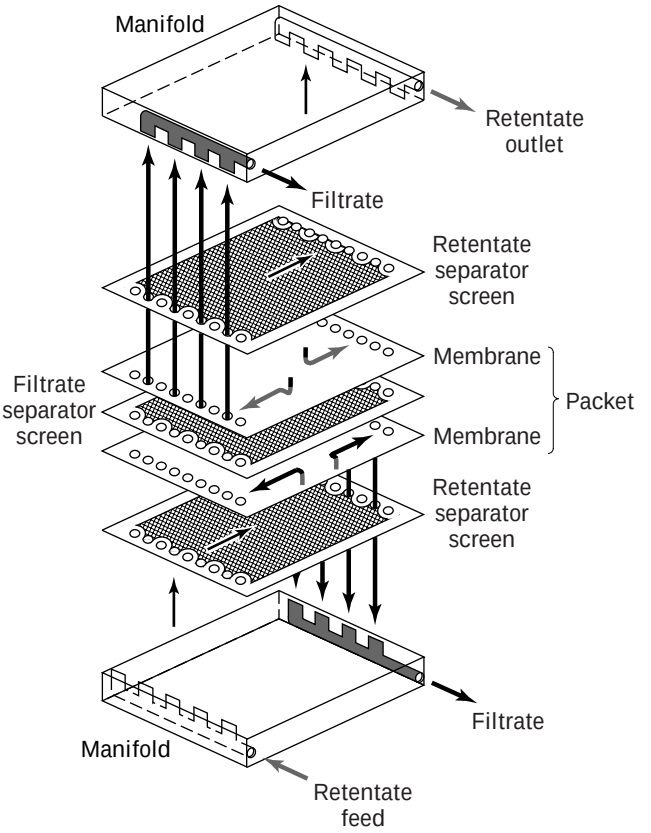


Figure 3. Schematic drawing of flat-sheet cassette. *Source*: Adapted from Ref. 8. Courtesy of Marcel Dekker, Inc.

Spiral Cartridges. Spiral-wound cartridges are constructed using flat-sheet membranes in the form of a pocket consisting of two membrane sheets separated by a highly porous permeable mesh that provides physical support for the membrane and defines the filtrate flow path (Fig. 4). The membranes are sealed along three edges to form a pocket using an appropriate epoxy or polyurethane adhesive. The open side of the pocket is attached (glued) to a central perforated tube that collects the filtrate flow. Several pockets are wound in a spiral configuration around a single collecting tube using a feed-side (polypropylene) mesh as a spacer between the pockets to establish the desired feed channel thickness. The use of multiple pockets reduces the total path length and pressure drop for the filtrate flow. The entire spiral is generally wrapped with fiberglass tape or shrink wrap for added mechanical strength and to prevent unwrapping of the spiral under pressure. The spiral-wound element is fitted into a tubular steel or plastic pressure vessel with an antitelescoping device placed at both ends to prevent distortion of the spiral during operation. This also allows several spiral elements to be inserted into a single tube. Pressurized feed solution flows through the plastic screens along the surface of the membranes with the retentate collected out the other end of the device. Filtrate flows into the closed membrane pockets and flows in a spiral path inward to the central tube. The region between the outer edge of the spiral element and the inside of the cylindrical housing is sealed to avoid fluid bypass. Spiral cartridges that direct the feed flow in the spiral direction and the filtrate flow in the axial direction have also been produced. Spiral-wound cartridges have relatively high membrane packing densities and very effective mass transfer characteristics because of the feed-side spacers. The cartridges also have fairly low energy costs because they provide effective mass transfer at relatively low flow rates. The primary disadvantage of the spiral wound modules is that they are highly susceptible to particulate fouling because of the narrow and irregular flow path through the spacers. In addition, the dead space between the outer edge of the element and the cylindrical housing is difficult to clean and highly susceptible to bacterial contamination. This is sometimes overcome by allowing a very limited feed bypass to continuously flush the annular space. There can also be problems with the integrity and chemical resistance of the seals used to form the pockets, particularly for repeated applications using strong cleaning solutions.

Hollow-Fiber Cartridges. Hollow-fiber cartridges use an array of narrow bore (self-supporting) fibers, typically with an anisotropic structure. Fiber diameters range from about 200 to 2,500 μm, with the fiber wall around 200 μm in thickness. The dense skin layer is typically at the lumen side of the fiber, although it can also be placed on the outside. A parallel array of 50 to 10,000 fibers is potted at the ends in an epoxy or polyurethane resin within a cylindrical cartridge of plastic or steel to form a tube that is machined to expose the open bores (lumens) of the fibers (Fig. 5). These modules are generally operated with the feed flow into the fiber lumens, with the filtrate moving in a radial direction outward through the fiber walls. The cartridge is typically equipped with a single feed inlet, a retentate outlet, and two filtrate ports (one near each end of the unit). The feed and filtrate flows can be cocurrent or countercurrent. Hollow fiber devices have high packing density because of the narrow diameter of the individual fibers and the ability to pack the fiber bundles tightly in the cylindrical cartridge (shell). The use of small diameter fibers

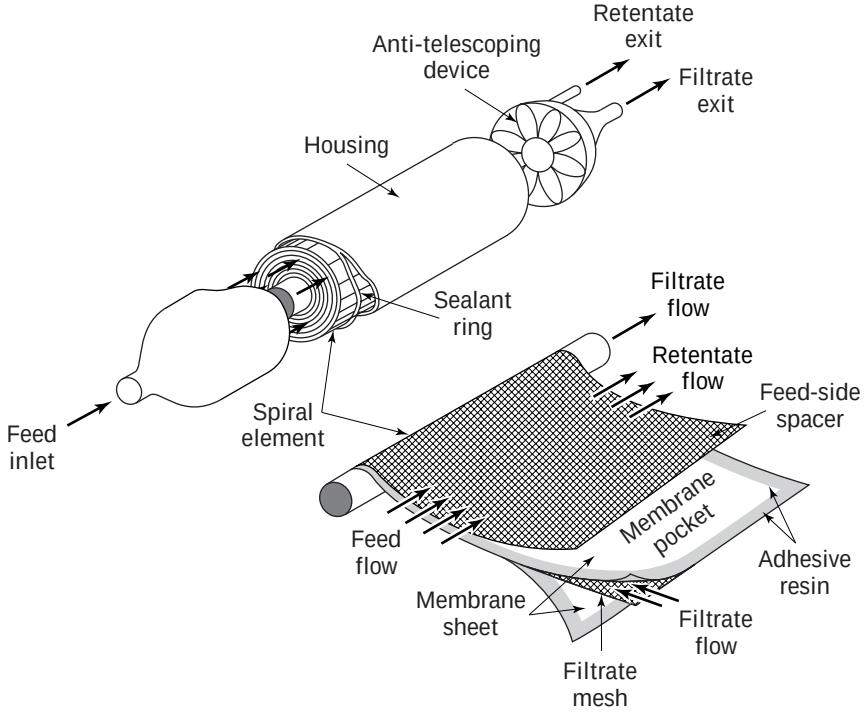


Figure 4. Schematic drawing of spiral-wound module showing membrane pocket. *Source*: Adapted from Ref. 8. Courtesy of Marcel Dekker, Inc.

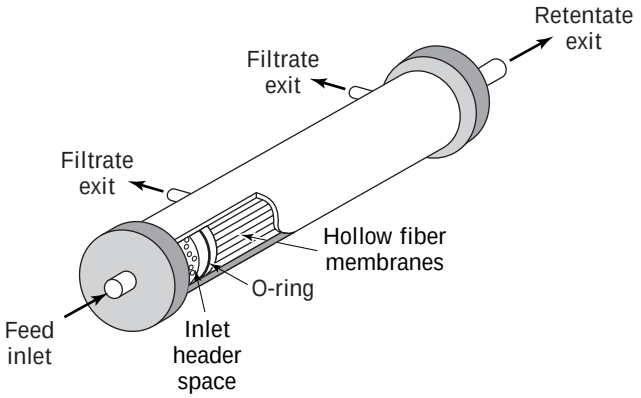


Figure 5. Schematic drawing of hollow-fiber module. *Source*: Adapted from Ref. 8. Courtesy of Marcel Dekker, Inc.

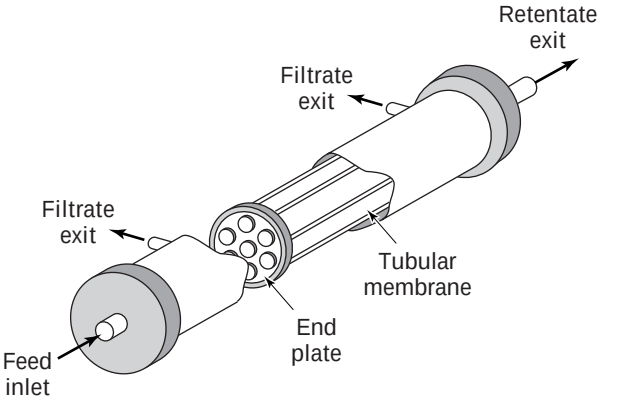


Figure 6. Schematic drawing of tubular module. *Source*: Adapted from Ref. 8. Courtesy of Marcel Dekker, Inc.

can provide effective mass transfer at relatively low feed flow rates. The absence of any spacers or multiple sealing procedures reduces labor costs involved in assembly. The hollow fiber membranes are self-supporting, which often allows them to be cleaned by backflushing of the device either by reversing the direction of the filtrate flow or by reversing the feed flow with the filtrate ports closed. One disadvantage of the hollow fiber modules is that they are susceptible to particulate plugging. Fiber breakage can also be a problem, so it is critical to follow the manufacturer's recommendations with respect to operating pressures. Fiber breakage can often be difficult to detect. Several manufacturers recommend using a modified bubble-point test to detect ruptured fibers.

Tubular Modules. Tubular modules are very similar in design to the hollow-fiber cartridges described in the previous section but use much larger diameter tubes (typically 0.3 to 2.5 cm). The tubular membranes are not self-supporting (with the exception of some inorganic membranes). The membranes are usually cast in place within a porous support tube made of fiberglass, ceramic, plastic, or stainless steel. The individual tubes can be placed inside a plastic or stainless steel sleeve to form a single tube cartridge, or they can be packed together in small bundles that are kept in place using an appropriate end plate (Fig. 6). The feed flows through the bore of the tubes, while the filtrate flows in a radial direction outward across the membrane and support tube where it is collected from the filtrate outlet ports. The two streams may be cocurrent or countercurrent. Inorganic membranes are typically made in a honeycomb monolith (18) in which the membranes are arranged in a parallel array. The unit is housed in an appropriate plastic or stainless steel cartridge. Tubular modules are typically operated in the turbulent flow regime to obtain effective mass transfer in the large diameter tubes. The primary advantages of the tubular modules are their resistance to particulate plugging and ease of chemical or mechanical cleaning. Physical cleaning of the membranes can be affected by passage of a rubber sponge, a scouring ball, or a narrow rod through the bore of the tubes. It is also possible to replace individual tubes in the multielement units in the field, resulting in a considerable reduction in membrane replacement costs and an increase in the overall unit lifetime. The primary disadvantages of the tubular modules are the low membrane packing densities ($<100\ m^{-1}$), high hold-up volumes, large feed flow rates, and high capital costs (including high floor space requirements).

Enhanced Mass Transfer Devices. Enhanced mass transfer modules exploit flow instabilities or turbulence to increase the transport of retained solutes. A variety of devices have been designed using pulsatile flow or membranes with pronounced surface corrugations (36). Alternatively, systems have been developed in which the feed flow generates secondary flow patterns or vortices. The most successful of these to date has been the rotating cylinder device originally proposed by Hallstrom and Lopez-Leiva (37). The feed flows into a thin (0.05 to 0.10 cm) annular region between two concentric cylinders, one (or both) of which is porous and has a membrane bound to the surface. The filtrate is collected in the central chamber (for a device with membrane on the inner cylinder) or in a separate annular filtrate region adjacent to the outer porous cylinder. The inner cylinder in these devices is rotated at high speed (typically >3,000 rpm) to induce Taylor vortices, which are caused by the centrifugal forces acting on the rotating fluid as it moves around the cylindrical flow path. The vortices effectively mix the fluid near the surface of the membrane, increasing the feed-side mass transfer coefficients. Mass transfer in the rotating system is determined almost entirely by the rotation rate of the inner cylinder, which effectively decouples the mass transfer characteristics from the feed flow rate. Consequently, rotating devices can be operated at very low feed flow rates. The use of low flow rates can be particularly attractive for many high viscosity applications. However, the rotating systems have very large capital costs, they are difficult to scale-up, and there is concern about long-time performance of a device with so many moving parts. In addition, membrane packing densities are low, and it is difficult to develop and maintain effective seals in such a rotating system.

Belfort et al. (38) and Winzeler (39) independently proposed using a vortex-flow system based on Dean, rather than Taylor, vortices. Dean vortices are induced by centrifugal forces exerted on a fluid as it flows inside a highly curved channel. This can be accomplished using spiral wound flat sheet membranes, with the channel curvature adjusted to maintain the Dean vortices throughout the system, or with hollow-fiber membranes wound around a central core at an appropriate pitch in a screwlike geometry (40). Both systems generate counter-rotating vortices that effectively depolarize the concentration boundary layer, but without the need for any moving parts or seals. Available data on the Dean vortex systems are very encouraging (38). These units have not yet been commercialized, so it is impossible to evaluate their economics and overall performance characteristics.

Another approach to achieving high mass transfer is to move the membrane, with the bulk of the fluid remaining largely stationary. For example, Vigo et al. (41) developed a module in which a cylindrical membrane is vibrated (longitudinally) inside a concentric cylindrical housing. Culkin and Armando (42) developed vibratory shear enhanced processing (VSEP) in which a stack of parallel circular membrane disks is spun at high speed inside a concentric cylindrical housing using a torsional spring. Shear rates at the membrane-feed interface can be as high as 150,000 s^{-1} for a 60-Hz oscillation. The high shear rates obtained in these systems can result in dramatic improvements in filtrate flux; however, important questions remain about the scalability and robustness of these modules as well as the overall economics.

High-frequency transmembrane pressure pulsing can also be used to enhance mass transfer. In contrast to classical backflushing, in which the reverse filtration occurs after several minutes or hours of normal filtration, transmembrane pressure pulsing is performed with a frequency around 0.1 to 1 Hz (43). The back pressure pulse is applied using nitrogen or air pressurization, with the frequency and duration of the pulse set using computer-controlled solenoid valves (43). Very short forward filtration times are used to prevent significant concentration polarization or fouling, thereby maintaining high average filtrate flux. Data for filtration of albumin–immunoglobulin mixtures showed more than an order of magnitude increase in flux (44), but a corresponding reduction in protein retention. Although transmembrane pressure pulsing systems appear to have excellent performance characteristics, they have not yet been used in large-scale commercial applications.

Process Configurations

Batch. UF and HPTFF processes consist of various combinations of concentration and diafiltration steps. A concentration step is either the main objective of the process or may be used as a means to optimize a subsequent diafiltration step (see "Design Equations" and "Process Optimization"). Concentration steps are carried out in either batch or fed-batch mode. In a batch mode, the entire volume of feed is contained within the recycle tank, as illustrated in Figure 7. Batch mode operations use a minimum

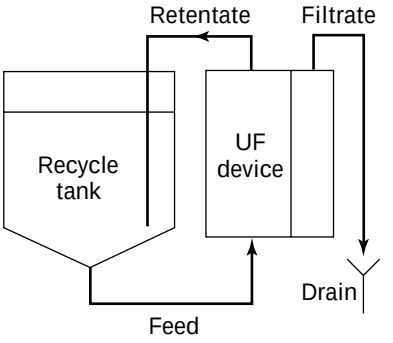


Figure 7. Batch-mode UF for a retained product.

of hardware, provide simple manual or automatic control, and provide the highest flux rates. Although the initial and final retentate concentrations are the same regardless of process configuration, the batch operation gives higher flux rates than a fed-batch operation because the bulk concentration follows a more dilute path from initial to final concentration. The disadvantages of a batch configuration can include less flexibility in using the same system for multiple processes, increased difficulty in designing a well-mixed system, and difficulties in obtaining high concentration factors. Mixing is important both to keep the bulk concentration at a minimum and for efficient buffer exchange during diafiltration. High concentration factors can in some cases be obtained in a batch mode using tanks with two or more sections of progressively decreasing diameters.

Fed Batch. A fed-batch configuration uses an additional tank to feed into the recycle tank, as shown in Figure 8. Fed-batch configurations are commonly used in industry to obtain high concentration factors, provide flexibility for use in multiple processes, and provide well-mixed low hold-up retentate reservoirs for efficient diafiltration. In choosing a fed-batch configuration, several trade-offs must be carefully considered. For a given membrane area, process time will increase with the use of a fed-batch configuration because of the increase in bulk concentration compared to a batch operation. In addition, the number of times the retained solute passes through pumps and valves in the recycle line will increase because of the longer

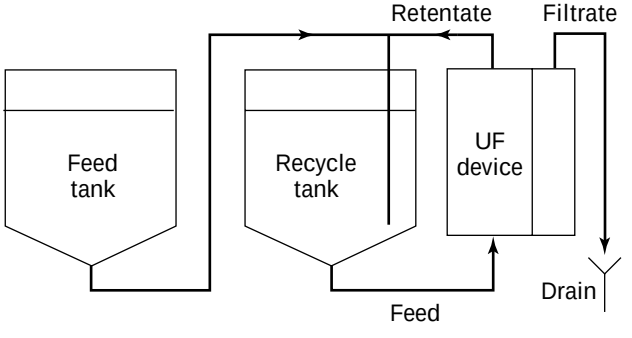


Figure 8. Fed-batch mode UF for a retained product.

processing time and the smaller retentate volume. An increased number of pump and valve passes can lead to shear-induced cell lysis and protein denaturation and aggregation. Protein denaturation in pumps and valves is believed to be the result of protein–gas interfaces (45) formed as a result of microcavitation.

Diafiltration. Diafiltration is an operation in which a solvent or solution (typically a buffer) is added to the recycle tank simultaneously with filtration (Fig. 9). The most common form is constant retentate volume diafiltration in which the buffer is added at the same rate as the filtrate rate. In UF processes, diafiltration is used to perform buffer exchange if the product is in the retentate, or to enhance yields if the product is in the filtrate. In HPTFF, diafiltration is used to achieve purification. In certain processes, it can be advantageous to perform simultaneous dilution or concentration with diafiltration (46). Optimization of diafiltration processes will be described in "Fluid Dynamics and Process Control" for UF and in "Process Optimization" for HPTFF.

Cascade Systems. In several plant designs, it may be advantageous to use cascaded process configurations. These configurations include both open-loop and closed-loop designs. In both configurations, the filtrate from the first stage is used as a feed to the second stage. In the open-loop design, the filtrate from the second stage is directed to drain while a replacement buffer is added to the first stage. In the closed-loop design (Fig. 10), the filtrate from the second stage is used as the diafiltration buffer for the first stage, thereby significantly reducing the amount of buffer required for the process. Cascade systems can be used to couple multiple process steps into a single operation. For example, a MF step used to harvest protein from a culture can be cascaded with a second-stage UF step to achieve simultaneous harvest and protein concentration. In HPTFF, two-stage cascade systems can be used in a similar manner when the product of interest is in the first-stage filtrate. Once a normal diafiltration has been performed on the first stage to achieve optimum buffer conditions (see "Buffer Chemistry"), a closed-loop configuration can be used to obtain a very large number of diavolumes by reusing a volume of buffer that is equal to the volume of retentate. It should be noted that the retentate volumes in the first- and second-stage tanks do not have to be the same. In cases where the buffer cost is very high, one might also consider using such a configuration even when the product is in the first-stage retentate if economically justified. Cascade configurations with more than two stages can also be used in processes where it is desirable to separate a feedstream into more than two fractions.

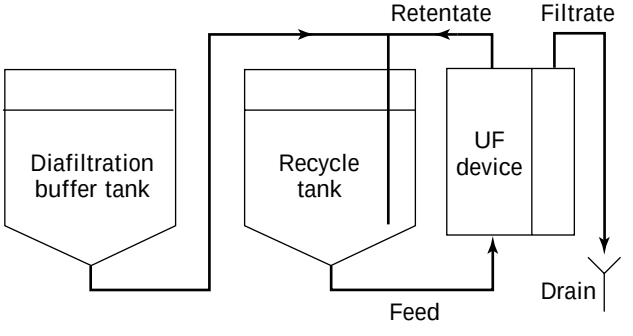


Figure 9. Diafiltration mode UF for a retained product.

PROCESS DESIGN

Ultrafiltration

Buffer Chemistry. Buffer exchange is often used to condition a product for subsequent loading onto a chromatographic column or for final formulation. The buffer chemistries used in ultrafiltration are hence often dictated by the preceding and subsequent steps in the process. Because of the very large amounts of protein processed per surface area of membrane, ultrafiltration can be used with a wide range of ionic strengths and pH values without significant adsorptive product loss. As discussed previously, adsorptive yield losses are usually insignificant ($<0.01\%$) for products in the retentate and can be limited to less than 5% for products in the filtrate. Although UF is routinely used to provide buffer exchange in preparation for subsequent processing, it can be advantageous to consider selecting buffer conditions for the sake of optimizing the UF step. Changes in pH and ionic strength can be made to affect protein solubility, stability, adsorption, and hydrodynamic volume. These parameters are directly linked to yield losses incurred at the membrane surface (see "Fluid Dynamics and Process Control"). In addition, buffer chemistry plays an essential role in membrane fouling, which can have a significant impact on process fluxes even when adsorptive product losses are minimized. The mass transfer coefficient can also be significantly enhanced by changes in buffer composition resulting in enhanced fluxes, reduced membrane area, and minimal volumetric hold-up losses. In general, higher fluxes are obtained at pH values either above or below the protein isoelectric point and at low to moderate ionic strength.

The theoretical removal of low molecular weight components during diafiltration is discussed in "Design Equations." It is important to note that one cannot assume that the retention coefficient for these components is zero despite the fact that the membrane may have a molecular weight cutoff an order of magnitude larger than the molecular weights of the components. Deviations in theoretical clearance of low molecular weight components can result from poor mixing in the recycle tank but can also result from buffer conditions. The use of detergents in either the feed material or in the diafiltration buffer can cause incomplete diafiltration if the critical micelle concentration is exceeded. Buffer components are entrained in the micelles, which may be retained by the membrane depending on micelle size and membrane molecular weight cutoff. It is generally advisable to avoid the use of detergents in UF for these reasons. Charged membranes may cause Donnan exclusion effects that result in increased retention of ionic solutes of all sizes. It has also been observed that low molecular weight ionic solutes can have

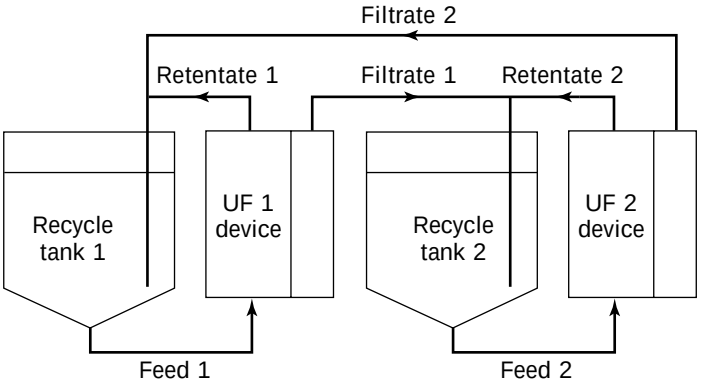


Figure 10. Two-stage closed-loop cascade configuration.

increased clearance, possibly as a result of Donnan exclusion by retained solutes.

Membrane Selection. Membrane selection should start with the choice of a high-quality vendor because robustness, reliability, and reproducibility of biotechnology manufacturing operations are of paramount importance. Consistent membrane and device characteristics can be as important to product quality, yield, and economics as are the inherent differences between various membranes and devices. The most important engineering parameters to consider during membrane selection are product retention, process permeability, mechanical strength, and chemical compatibility. Adsorptive properties play an important role in determining process permeability but are often less important with regard to product loss. Membrane chemistry should therefore be chosen in conjunction with the choice of buffer chemistry. Membrane charge should also be considered with regard to potential Donnan exclusion effects, as discussed previously. Water permeability values should not be used as a selection criteria since they are not predictive of process performance. Pore size distributions, porosity, morphology, and membrane thickness are only of indirect importance, because these parameters affect the ultimate retention, permeability, and mechanical characteristics of the membrane. As previously discussed ("Membrane Characterization"), retention properties should be determined on the actual solute under process conditions because marker studies can be completely misleading. Rules of thumb for selection of a membrane molecular weight cutoff relative to a protein molecular weight should also be avoided because protein hydrodynamic volumes vary drastically with pH and ionic strength (10) and because membrane manufacturers have not standardized their membrane characterization and nomenclature for UF membranes. The mechanical strength of a membrane can be important with regard to membrane compaction and reverse pressure tolerance. Membrane compaction can influence process permeability if retentate pressure limits are exceeded. The reverse pressure limits of a membrane are also important, because reverse pressure spikes are very difficult to avoid in industrial UF systems. The resulting membrane delamination can have a catastrophic impact on process yields. Chemical compatibility needs to be considered for the intended feedstream and the regeneration chemicals. Polysulfone and modified polysulfone membranes tend to have much broader chemical compatibility than regenerated cellulose membranes, but they also require much harsher chemical conditions for regeneration because of their inherent fouling characteristics. Long-term stability of membranes in bacteriostatic storage solution must also be carefully considered for economic reasons. The long-term stability of UF membranes should be evaluated by studying product retention as a function of membrane storage time in the chosen storage solution.

Fluid Dynamics and Process Control. Ultrafiltration processes can been performed using constant transmembrane pressure (ΔP), constant flux (J), or constant retentate wall concentration (C_w). The primary benefit of constant transmembrane pressure UF processes is the inherent simplicity of control, especially in manual systems. The feed rate is ramped up to the set point value, and a retentate valve is partially closed to achieve a specified retentate or transmembrane pressure:

$$\Delta P = (P_F + P_R)/2 - P_f \quad (11)$$

where P_F is the feed pressure, P_R is the retentate pressure, and P_f is the filtrate pressure. Note that this equation is an approximation because the pressure drop along the feed channel is not truly linear because of conversion of feed to filtrate. If the feed stock viscosity changes during the process, one can adjust the retentate valve closure to maintain either constant retentate or transmembrane pressure. This is easily accomplished in both manual and automatic systems and often requires minimal adjustments after the initial set point has been achieved. The retentate pressure can also be controlled by using a gas overlay on the recycle tank. The feed rate and transmembrane pressure are empirically optimized by measuring the flux as a function of bulk solute concentration at various feed rates and transmembrane pressures. It is also important to measure solute sieving as a function of bulk concentration and flux. Even very small sieving coefficients may result in substantial product loss when a large number of diavolumes are required in DF processes. For example, a sieving coefficient of $S_o = 0.01$ will result in a 10% product loss during a 10-diavolume buffer exchange process (see equation 17).

In certain applications, it is not possible to maintain constant transmembrane pressure without severe reductions in filtrate rate. It has been demonstrated that one can obtain higher overall throughput in these applications (such as clarification) by operating at constant flux (47). This can be done by regulating the retentate pressure control valve or by using a pump on the filtrate line. The latter alternative provides more precise control and is easier to use in manual systems. Optimization is typically done through empirical studies varying feed rate and flux and determining process throughput, filtrate clarity, and product sieving.

In a third method of process control, the flux is varied with both retentate concentration and mass transfer coefficient to maintain a constant wall concentration of retained species (evaluated from equation 6). Control is performed using a control loop that measures flux and controls the transmembrane pressure with a retentate control valve to maintain a constant wall concentration (Fig. 11) throughout the concentration (15):

$$J = k \ln(C_w V / C_0 V_0) \tag{12}$$

where J is the flux, k is the mass transfer coefficient, C_w is the wall concentration of the retained solute, V is the retentate volume, V_0 is the initial feed volume, and C_0 is the initial retentate concentration. The benefits of constant C_w control are that product yield is maximized, product quality is ensured, membrane area is minimized, and process time is consistent and independent of membrane permeability. Process control for constant C_w during a concentration phase requires an automated system. Constant retentate volume DF using constant C_w, however, can easily be performed using either manual or automatic control because the C_w set point is maintained at a constant flux. Adjustments for changing mass transfer coefficients can be made during the DF process. The optimum C_w is determined by minimizing yield losses attributable to sieving, adsorption, solubility, and volumetric hold-up. (Fig. 12). The feed rate is optimized by evaluating a tradeoff between

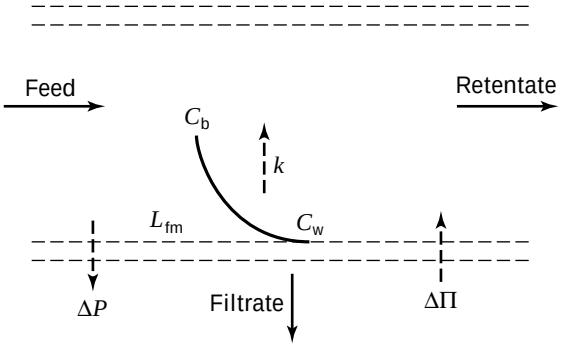


Figure 11. The concentration of a fully retained solute at the membrane wall (C_w) depends on the bulk concentration (C_b), the transmembrane pressure (ΔP), the fouled membrane resistance (L_{fm}), the osmotic pressure ($\Delta\Pi$), and the mass transfer coefficient (k). *Source*: Reprinted from R. van Reis et al. (15). With kind permission of Elsevier Science–NL, Amsterdam, The Netherlands.

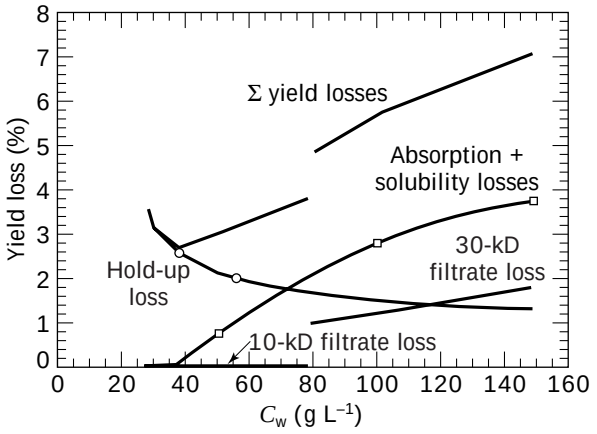


Figure 12. Yield losses in UF as a function of the wall concentration (C_w). Filtrate, adsorption, and solubility losses were determined with 0.1 to 0.4 m^2 linear scale-down devices. Hold-up losses were derived from data obtained with 5-, 21-, 63-, and 85-m^2 systems. *Source*: Reprinted from R. van Reis et al. (15). With kind permission of Elsevier Science–NL, Amsterdam, The Netherlands.

increased mass transfer coefficient (which increases flux and reduces membrane area for a given C_w) and an increased number of pump/valve passes (which may increase protein denaturation). The mass transfer coefficient can be calculated as:

$$k = aQ^b D^{2/3} \tag{13}$$

where a and b are coefficients determined for the specific UF module, Q is the feed flow rate normalized for membrane area, and D is the solute diffusivity. A typical constant C_w process is shown in Figure 13.

Design Equations The yield and purification obtained in UF can be evaluated from a mass balance, typically assuming a constant sieving coefficient. The fractional retentate product yield in a batch process is

$$Y_r = (V_0/V_F)^{-S_0} \tag{14}$$

where V_F is the final retentate volume, V_0 is the initial retentate volume, and S is the sieving coefficient. The analogous yield for a product in the filtrate can simply be calculated as

$$Y_f = 1 - Y_r \tag{15}$$

For a fed-batch process, the retentate yield is given by

$$Y_r = [1 - (1 - S_o) \exp(-S_o(V_0/V_F - 1))]/S_o(V_0/V_F) \tag{16}$$

and the corresponding filtrate yield can be calculated by combining equations 15 and 16. The retentate yield during constant retentate volume diafiltration is derived as

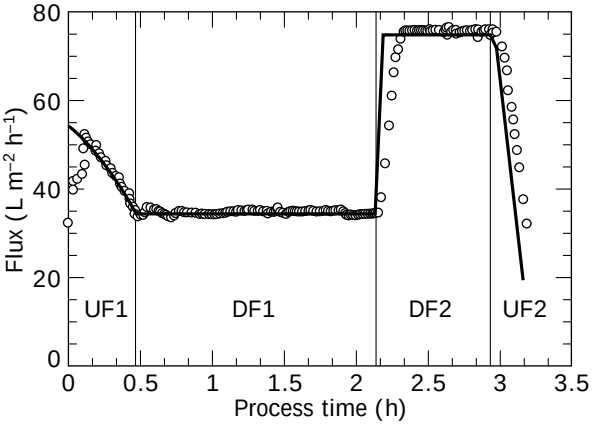


Figure 13. Predicted filtrate flux (solid line) and actual filtrate flux (open circles) for a constant C_w UF/DF process as a function of process time. Concentration of a recombinant antibody from 4 to 10 g L^{-1}, followed by a 10-diavolume buffer exchange and final concentration to 45 g L^{-1}. The increased fluxes during DF2 and UF2 process steps are the result of an increase in mass transfer coefficient and optimum C_w after the initial buffer exchange (DF1).

$$Y_r = \exp(-S_o N) \tag{17}$$

where the number of diavolumes is given by

$$N = V_D/V \tag{18}$$

where V_D is the DF buffer volume, and V is the constant retentate volume. It is often useful to combine and rearrange equations 14 and 17 to determine the retention coefficient required to obtain a desired retentate yield for a batch UF and constant volume DF process as

$$R = 1 + \ln Y/(\ln(V/V_0) + N) \tag{19}$$

For a filtrate product, the required retention coefficient is

$$R = 1 + \ln(1 - Y)/((\ln(V/V_0) + N) \tag{20}$$

DF processes are often used to remove undesirable low molecular weight components. The residual concentration can be calculated as

$$C = C_0 \exp(-S_o N) \tag{21}$$

Buffer exchange may also involve the addition of a new component:

$$C = C_i(1 - \exp(-N)) \tag{22}$$

where C_i is the bulk concentration of the component that is being added during diafiltration. Simultaneous addition and removal of a component may also occur in cases where the initial buffer and the DF buffer have a common component. The concentration of such a component can be calculated by adding equations 21 and 22. Process time and membrane area can be minimized by operating DF processes at an optimum bulk concentration (48):

$$C_b^* = C_w/e \tag{23}$$

By combining equations 23 and 6, with $C_f = 0$ for a fully retentive membrane, it has been shown (15) that the optimum diafiltration flux at C_b^* can be derived as

$$J^* = k \tag{24}$$

High-Performance Tangential-Flow Filtration

UF has been limited to separating solutes that are at least 10-fold different in size. In contrast, HPTFF enables the separation of solutes without limit to their relative size. HPTFF can even be used to perform reverse separations in which the higher molecular weight solute passes through the membrane while the lower molecular weight solute is highly retained. High-resolution separations are accomplished by exploiting concentration polarization, optimizing buffer chemistry, and using highly selective UF membranes. The performance of HPTFF is quantified by two dimensionless numbers, Ψ and $N\Delta S$:

$$\Psi = S_2/S_1 \tag{25}$$

$$N\Delta S = (JAt/V)(S_2 - S_1) \tag{26}$$

where Ψ is the selectivity, S_1 and S_2 are the respective sieving coefficients for the less and more highly retained species, N is number of diavolumes, A is membrane area, t is process time, and V is retentate volume. Membrane area and process time can be chosen to satisfy economic and manufacturing criteria independent of other variables. The retentate volume can be optimized separately. It is therefore useful to initially evaluate the selectivity (equation 25) and $J\Delta S$ values, which only depend on three simple experimental parameters:

$$J\Delta S = J(S_2 - S_1) \tag{27}$$

The next three sections will describe how buffer chemistry, membrane selection, and fluid dynamics are optimized. The final section will describe the mathematical basis and use of optimization equations and diagrams.

Buffer Chemistry. Selecting buffer pH and ionic strength are of paramount importance to achieving optimum separation performance in HPTFF. By operating close to the pI of the lower molecular weight solute and far away from the pI of the larger molecular weight solute, one can take advantage of differences in hydrodynamic volume and charge (3). The electric double layer that surrounds charged solutes increases the effective hydrodynamic volume as determined by SEC (10). Direct charge effects can also be exploited by using a membrane that has the same charge as the retained solute. In addition, the conformation of a molecule, and hence the hydrodynamic volume, may be changed under highly charged conditions. Use of charge effects are optimized by operating at low ionic strength (Fig. 14). The optimum ionic strength is generally achieved

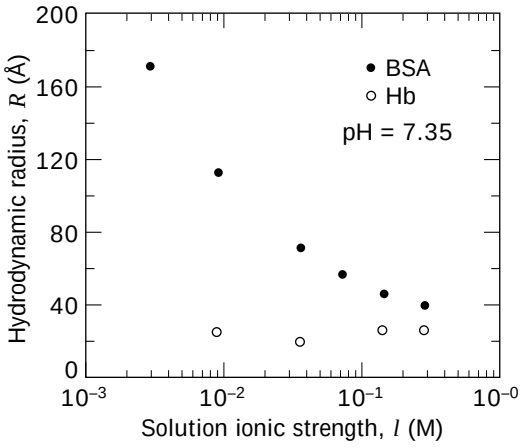


Figure 14. Effective hydrodynamic radius of BSA and hemoglobin (Hb) as a function of solution ionic strength. Calculated values of R were determined from retention time measurements in SEC using narrow molecular weight distribution dextran standards for calibration (10).

by using the lowest possible value while still maintaining an adequate pH buffering capacity. The buffering capacity of proteins at high concentrations near the membrane wall must be taken into account. Even in cases where there is no difference in pI between the solutes, it is advantageous to optimize both pH and ionic strength because the hydrodynamic volumes can be adjusted to increase selectivity and $J\Delta S$ for a given membrane molecular weight cutoff (3). This also reduces the need to have a large selection of different molecular weight cutoff membranes for HPTFF. Hydrodynamic volume and charge effects can also be altered by using buffering ions that have different protein binding characteristics (49). Rapid process development can, however, often be accomplished by testing just a few pH values around the pI of the lower molecular weight solute with a single buffering species at a concentration of 10 mM. As in all purification process development, it is important to ensure that the buffer composition does not cause undesirable biochemical alterations to the product and that product solubility is maintained at the concentrations incurred at the membrane wall (see "Fluid Dynamics and Process Control").

Membrane Selection. The general considerations for UF membrane selection outlined in "Membrane Selection" also apply to HPTFF. The selectivity and $J\Delta S$ that can be achieved with a membrane is determined by the pore-size distribution and membrane chemistry. Pore-size distribution affects both selectivity and $J\Delta S$ by altering the solute sieving coefficients and the filtrate flow distribution (8). Membrane chemistry has an indirect impact on the pore-size distribution because membrane formation conditions are linked to the casting solution chemistry. Membrane chemistry also plays a significant role in fouling, which in turn influences the selectivity and $J\Delta S$ values. Finally, the charge versus pH profile of the membrane can be selected to influence both charge repulsion and fouling. It should be noted that selective fouling of a membrane can actually be used to enhance separations because the pore-size distribution can be reduced and the surface charge can be altered. Using these guidelines, it is often possible to narrow down the choice to one or two membranes before proceeding with fluid dynamic optimization.

Fluid Dynamics and Process Control. The intrinsic sieving coefficient for a solute is defined by equation 3. The intrinsic sieving coefficient depends on the size and charge characteristics of both the solute and the membrane pores. Buffer optimization and membrane selection have hence focused on establishing conditions that provide the best differential intrinsic sieving coefficients for the solutes to be separated. The observed sieving coefficient of a solute, however, depends not only on the intrinsic sieving coefficient but also on the concentration of solute at the membrane wall. The wall concentration is determined by the mass transfer coefficient and the filtrate flux. The observed sieving coefficient is related to the intrinsic sieving coefficient (S_∞), mass transfer coefficient (k) and flux (J) by equation 8. It is therefore clear that the optimization of HPTFF is strongly influenced by concentration polarization effects. Concentration polarization has often been cited in the literature as one of the reasons why UF has not been successfully used for high-resolution separations. It has been demonstrated (3), however, that concentration polarization can be used to enhance, rather than limit, the resolving power of UF membranes. Key differences used in HPTFF methodology (50,51) include operating below the transition point (Fig. 15) and minimizing the transmembrane pressure gradient along the length of the filtration module. Although conventional UF has often sought to maximize flux by operating above the transition point, it has been shown that the optimum selectivity and $J\Delta S$ (a differential mass flux) occur below the transition point (3). It has also been demonstrated that membrane fouling is more consistent and can be controlled below the transition point (15). Optimization of HPTFF fluid dynamics can therefore be reduced to a set of simple experiments in which module geometry, feed flow rate, and filtrate flux are investigated. The results are evaluated by measuring selectivity (Ψ) and differential mass flux ($J\Delta S$) as a function of these variables.

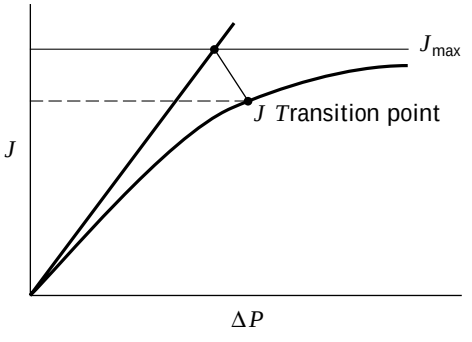


Figure 15. Typical flux versus transmembrane pressure curve showing the transition point flux at the inflection point (50). The optimum selectivity (Ψ) and differential mass flux ($J\Delta S$) occur below the transition point flux.

Process Optimization. The goal of process optimization is to achieve product quality specifications with the best combination of yield and purification factor using an economic amount of membrane area and providing an acceptable process time. Yield, purification factor, membrane area, and process time can be evaluated by using an optimization diagram (52) for either retentate product (Fig. 16) or filtrate product (Fig. 17). The optimization diagrams consist of a family of curves with the dimensionless parameters selectivity and $N\Delta S$. The family of curves in the optimization diagrams are mathematical expressions that link the yield and the purification factor to selectivity and $N\Delta S$ values. The equations are derived (52) as

$$P_R = Y_R^{1-\Psi} = \exp(N\Delta S) \quad (28)$$

$$P_F = Y_F/(1 - (1 - Y_F)^{1/\Psi}) = Y_F/(1 + (Y_F - 1)\exp(N\Delta S)) \quad (29)$$

where P is the purification factor, and Y is the fractional yield in the retentate (R) and filtrate (F). Both retentate and filtrate product processes operate along lines of constant selectivity. In the case of a retentate product, the process starts at a purification factor of 1 and a fractional yield of 1 and proceeds along a constant selectivity line up to the limit of $N\Delta S$ with an increasing purification factor and a decreasing yield. Conversely, the filtrate process starts at zero yield and a purification factor that numerically equals the selectivity (52) and proceeds along a constant selectivity curve with increasing yield and decreasing purification factor up to a limiting $N\Delta S$ value. Different experimental conditions (buffer chemistry, membrane chemistry and pore size, and fluid dynamics) can now be compared by evaluating the impact of selectivity and $J\Delta S$ on yield and purification factor. Selectivity and $J\Delta S$ often reach maxima under different experimental conditions. The best combination of selectivity and $J\Delta S$ depends on the overall process objectives and can be evaluated by comparing yield and purification factors for different combinations of these parameters using the optimization diagrams. As part of this analysis, $J\Delta S$ is converted to $N\Delta S$ by including membrane area, process time, and retentate volume according to equation 26. The membrane area and process times are chosen to provide acceptable economics, system hold-up volume, and process time. The retentate volume must be optimized. Decreasing the retentate volume (while maintaining a constant mass of solute) decreases the amount of buffer required for the process and can also decrease the amount of membrane area required for a given process time. This must be balanced against the increase in the number of pump and valve passes that occurs at increased concentrations, which can adversely impact product quality (see "Fluid Dynamics and Process Control").

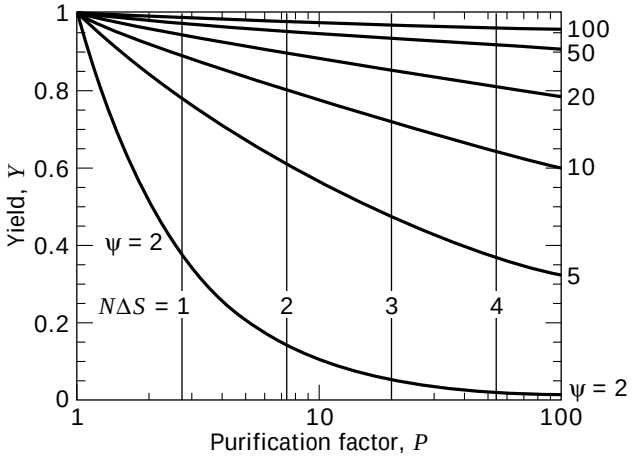


Figure 16. Optimization diagram for product in the retentate. The process follows a constant selectivity (Ψ) curve starting at a fractional yield $Y = 1$ and can be operated up to a maximum $N\Delta S$, which depends on the experimental $J\Delta S$ value and the ratio of membrane area (A), process time (t), and retentate volume (V). *Source*: Adapted from R. van Reis and S. Saksena (52). With kind permission of Elsevier Science–NL, Amsterdam, The Netherlands.

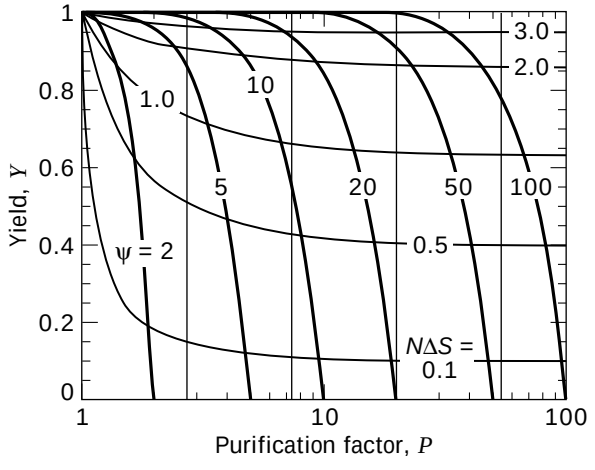


Figure 17. Optimization diagram for product in the filtrate. The process follows a constant selectivity (Y) curve starting at a fractional yield $Y = 0$ and can be operated up to a maximum $N\Delta S$, which depends on the experimental $J\Delta S$ value and the ratio of membrane area (A), process time (t), and retentate volume (V). *Source*: Adapted from R. van Reis and S. Saksena (52). With kind permission of Elsevier Science–NL, Amsterdam, The Netherlands.

Scale-Up

Effective commercial implementation of UF and HPTFF requires accurate and dependable scale-up of these membrane processes. It is also important to be able to scale-down existing manufacturing processes for further optimization, validation, and troubleshooting. Scale-up in UF and HPTFF obviously requires use of the same membrane material and pore size as well as the same module configuration and channel height at all scales of operation, with the membrane area scaled proportional to the filtrate volume. The most effective approach to scale-up is linear scaling, in which the pressure, fluid flow, and concentration profiles along the length of the filtration module are all kept constant when changing scale of operation (53). Be-

cause the local retentate velocity, solute concentration, and fluid pressure vary with position in the channel because of fluid removal and frictional pressure losses, linear scaling can only be achieved by keeping the channel length constant. Thus, the increase in membrane surface area required for large filtrate volume operation must be accomplished by increasing the number of parallel channels or fibers. It is also possible to increase the channel width in flat-sheet cassettes, although some care must be taken to ensure that this does not alter the fluid flow profiles in different size modules. Linear scaling of hollow fiber systems is relatively straightforward using modules with equal length but different numbers of fibers. Equal flow distribution between fibers and between fiber cartridges is easily accomplished with appropriate cartridge housings and piping manifolds. Linear scaling of flat-sheet designs has also been achieved using careful system design to ensure equivalent flow distribution, channel height compression, and entrance and exit effects at the small and large scale (53). Linear scale-up of hollow-fiber and flat-sheet devices over at least a 400-fold range of membrane area has been done successfully without intermediate pilot scale tests (53). Spiral-wound modules have a more limited range of linear scalability because pressure losses on the filtrate side will vary as the membrane area is changed, particularly in the design of very low area ($<0.1\ m^2$) modules. Linear scaling can also be developed for spiral devices in which the feed flow is in the spiral direction (with the filtrate flow in the axial direction), and these modules should provide enhanced mass transfer through the formation of Dean vortices. Linear scaling of the other enhanced mass transfer modules discussed in "Modules and Devices" can often be difficult to achieve over a wide range of membrane surface area because of practical limitations on system geometry and design.

BIBLIOGRAPHY

1. W.F. Blatt, A. Dravid, A.S. Michaels, and L. Nelsen, in J.E. Flinn ed., *Membrane Science and Technology*, Plenum Press, New York, 1970, pp. 47–97.
2. R.T. Kurnik, A.W. Yu, G.S. Blank, A.R. Burton, D. Smith, A.M. Athalye, and R. van Reis, *Biotech. Bioeng.* **45**, 149–157 (1995).
3. R. van Reis, S. Gadam, L.N. Frautschy, S. Orlando, E.M. Goodrich, S. Saksena, R. Kuriyel, C.M. Simpson, S. Pearl, and A.L. Zydney, *Biotech. Bioeng.* **56**, 71–82 (1997).
4. E.N. Lightfoot, *Transport Phenomena in Living Systems*, Wiley, New York, 1974.
5. O. Kedem and A. Katchalsky, *Biochim. Biophys. Acta.* **27**, 229–246 (1958).
6. W.S. Opong and A.L. Zydney, *AIChE J.* **37**, 1497–1510 (1991).
7. W.M. Deen, *AIChE J.* **33**, 1409–1425 (1987).
8. L.J. Zeman and A.L. Zydney, *Microfiltration and Ultrafiltration: Principles and Applications*, Dekker, New York, 1996.
9. N.S. Pujar and A.L. Zydney, *J. Colloid Interface Sci.* **192**, 338–349 (1997).
10. N.S. Pujar and A.L. Zydney, *J. Chromatography A*, **796**, 229–238 (1998).
11. N.S. Pujar and A.L. Zydney, *Ind. Eng. Chem. Res.* **33**, 2473–2482 (1994).
12. S. Mochizuki and A.L. Zydney, *J. Membrane Sci.* **82**, 211–228 (1993).
13. A.S. Michaels, *Chem. Eng. Prog.* **64**, 31–44 (1968).
14. A.L. Zydney, *J. Membrane Sci.* **23**, 275–282 (1997).
15. R. van Reis, E.M. Goodrich, C.L. Yson, L.N. Frautschy, R. Whiteley, and A.L. Zydney, *J. Membrane Sci.* **130**, 123–140 (1997).
16. F.G. Paulsen, S.S. Shojaie, and W.B. Krantz, *J. Membrane Sci.* **91**, 265–282 (1994).
17. A.J. DiLeo and A.E. Allegrezza, *Nature* **351**, 420–421 (1991).
18. H.P. Hsieh, *AIChE Symp. Ser.* **84**, 1–18 (1988).
19. G. Belfort and A.L. Zydney, in M. Malmsten ed., *Biopolymers at Interfaces*, Dekker, New York, 1998, pp. 513–559.
20. G. Tkacik and S. Michaels, *Bio/Technol.* **9**, 941–946 (1991).
21. M. Meireles, P. Aimar, and V. Sanchez, *J. Membrane Sci.* **56**, 13–28 (1991).
22. M. Meireles, A. Bessieres, I. Rogissart, P. Aimar, and V. Sanchez, *J. Membrane Sci.* **103**, 105–115 (1995).
23. S. Gadam, M. Phillips, S. Orlando, R. Kuriyel, S. Pearl, and A.L. Zydney, *J. Membrane Sci.* **133**, 111–125 (1997).
24. G. Belfort, R.H. Davis, and A.L. Zydney, *J. Membrane Sci.* **96**, 1–58 (1994).
25. B.C. Robertson and A.L. Zydney, *J. Colloid Interface Sci.* **134**, 563–575 (1990).
26. S.T. Kelly, W.S. Opong, and A.L. Zydney, *J. Membrane Sci.* **80**, 175–187 (1993).
27. J.H. Hanemaajer, T. Robbertsen, T. van den Boomgard, and J.W. Gunnik, *J. Membrane Sci.* **40**, 199–212 (1989).
28. S.B. McCray and J. Glater, in S. Sourirajan and S. Matsuura eds., *Reverse Osmosis and Ultrafiltration*, ACS Symposium Series, Washington, D.C., 1985.
29. M. Cheryan, *Ultrafiltration Handbook*, Technomic, Lancaster, Penn., 1986.
30. G. Tragardh, *Desalination* **71**, 325–335 (1989).
31. C.A. Lawrence and S.S. Block, *Disinfection, Sterilization, and Preservation*, Lea and Febiger, Philadelphia, 1968.
32. V. Gekas, G. Tragardh, and B. Hallstrom, *Ultrafiltration Membrane Performance Fundamentals: 10 Years of Characterization and Fouling Studies*, The Swedish Foundation for Membrane Technology, Lund University, Sweden, 1993.
33. K.E. Smith and R.L. Bradley, in D. Lund, E. Plett, and C. Sandu eds., *Fouling and Cleaning in Food Processing*, Extension Duplicating, University of Wisconsin, Madison, Wis., 1985, pp. 410–423.
34. D. Berman, T. Meyers, and S. Chrai, *J. Parenteral Sci. Tech.* **40**, 119–121 (1986).
35. V. Goel, M.A. Accomazzo, A.J. DiLeo, P. Meier, A. Pitt, M. Pluskal, and R. Kaiser, in W.S.W. Ho and K.K. Sirkar eds., *Membrane Handbook*, Chapman and Hall, New York, 1992, pp. 506–570.
36. G. Belfort, *J. Membrane Sci.* **35**, 245–270 (1988).
37. B. Hallstrom and M. Lopez-Leiva, *Desalination* **24**, 273–279 (1978).
38. K.-Y. Chung, R. Bates, and G. Belfort, *J. Membrane Sci.* **81**, 139–150 (1993).
39. H.B. Winzeler and G. Belfort, *J. Membrane Sci.* **80**, 35–47 (1993).
40. C.H. Gooding, *Membrane Technology and Separations Planning Conference*, Newton, Mass., October 1997.
41. F. Vigo, C. Uliana, and E. Ravina, *Sep. Sci. Technol.* **25**, 63–82 (1990).

42. B. Culkin and A.D. Armando, *Filtr. Sep.* **29**, 376–378 (1992).
43. S.G. Redkar and R.H. Davis, *AIChE J.* **41**, 501–508 (1995).
44. V.G.J. Rodgers and R.E. Sparks, *J. Membrane Sci.* **68**, 149–168 (1992).
45. T.J. Narendranathan and P. Dunnill, *Biotech. Bioeng.* **24**, 2103–2107 (1982).
46. U.S. Pat. 5,597,486 (Jan. 28, 1997), H. Lutz (to Millipore Investment Holdings Limited).
47. J.J. Sheehan, B.K. Hamilton, and P.F. Levy, *ACS Symp. Ser.* **419**, 130–155 (1990).
48. P. Ng, J. Lundblad, and G. Mitra, *Sep. Sci.* **2**, 499–502 (1976).
49. M. Menon and A.L. Zydney, *American Chemical Society Annual Meeting*, San Francisco, 1997.
50. U.S. Pat. 5,256,294 (Oct. 26, 1993), R. van Reis (to Genentech Inc.).
51. U.S. Pat. 5,490,937 (Feb. 13, 1996), R. van Reis (to Genentech Inc.).
52. R. van Reis and S. Saksena, *J. Membrane Sci.* **129**, 19–29 (1997).
53. R. van Reis, E. Goodrich, C.L. Yson, L.N. Frautschy, S. Dzengeleski, and H. Lutz, *Biotech. Bioeng.* **55**, 737–746 (1997).

See also MEMBRANE CHROMATOGRAPHY; MEMBRANE SEPARATIONS.

PROTEOLYTIC CLEAVAGE, REACTION MECHANISMS

THOMAS P. GRAYCAR
Genencor International Inc.
Palo Alto, California

KEY WORDS

Aspartic proteinase
Cysteine protease
Endopeptidase
Exopeptidase
Metallopeptidase
Proteolytic cleavage
Reaction mechansim
Serine protease
Site-specific mutagenesis
X-ray crystallography

OUTLINE

INTRODUCTION

Proteolytic cleavage, the hydrolysis of peptide bonds in a protein, is a ubiquitous reaction in nature catalyzed by enzymes called proteases. Whether the purpose served is digestive scavenging of nutrients or regulation of a critical physiological process such as blood coagulation, proteases play an essential role in the growth and survival of all living organisms. Because of their abundance and importance, the reaction mechanisms of proteases have been studied more extensively than any other class of enzyme. Many of the first high-resolution X-ray crystal structures of proteins were obtained for proteases. Recently, our understanding of the mechanism of rate acceleration used by proteases and the structural basis for their specificity has improved dramatically through the use of a powerful recombinant DNA technique called site-specific mutagenesis. Developed in the early 1980s, site-specific mutagenesis is a method whereby one or more specific residues in the amino acid sequence of a protein can be removed or, more commonly, replaced with another amino acid via mutagenesis of its corresponding cloned gene (1). The contribution to catalysis of an amino acid functional group in the active site of an enzyme can then be determined directly (2). Because of the valuable insight into the reaction mechanism of proteases afforded by site-specific mutagenesis studies and high-resolution X-ray crystallography, there has been significant progress made in the design of new protease-inhibitor drugs as well as the development of proteases with improved properties for medicinal and industrial applications.

REACTION MECHANISMS

Mechanistic Classification of Proteases

Enzymes that hydrolyze peptide bonds can be grouped into two subclasses, exopeptidases and endopeptidases, depending on where the reaction takes place in the polypeptide substrate. Exopeptidases cleave peptide bonds at the amino or carboxyl ends of the polypeptide chain, whereas endopeptidases cleave internal peptide bonds. Based on their reaction mechanism, proteases may be further classified into four groups: serine, cysteine, aspartic, and metallo (3). The name of each class is derived from a distinct catalytic group involved in the reaction. The active sites of serine and cysteine proteases use serine hydroxyl and cysteine thiol side groups, respectively, as the attacking nucleophiles during catalysis. Their reaction pathway entails formation of a covalent intermediate between the nucleophile and the carbonyl carbon atom of the scissile peptide bond. By contrast, the mechanism of hydrolysis for aspartic and metallo proteases does not involve a covalent intermediate because the nucleophile for both of these enzymes is a water molecule. The different functional groups for each protease reaction mechanism affect their active pH range as well as the types of inhibitors suitable for

inactivation. Indeed, measurements of the pH dependence for activity and reactivity with a set of class-specific inhibitors are typical laboratory methods for determining the mechanistic class of an unknown protease (4). Before describing each mechanism of catalysis in detail, it should be noted that the reaction pathways in all four protease classes share multiple common elements:

- Binding of the substrate polypeptide into a channel or pocket on the enzyme surface
- Nucleophilic attack on the carbonyl carbon of the scissile peptide bond by either oxygen or sulfur
- General base-assisted catalysis for removal of a proton from the nucleophile
- Stabilization of the tetrahedral transition state intermediate formed at the carbonyl carbon of the scissile peptide bond
- General acid-assisted catalysis for transfer of a proton to the amine leaving group

A description of substrate binding and the contribution of binding forces to catalysis and specificity will be discussed in a later section. The focus of this section will be the catalytic groups responsible for carrying out the chemical steps in the reaction.

Serine Proteases

The name *serine protease* refers to the nucleophilic serine residue located in the enzyme active site. In addition to the catalytic serine residue, this protease class is distinguished by having essential aspartate and histidine residues that together form a catalytic triad (Fig. 1). Serine proteases are prevalent in both prokaryotic and eukaryotic organisms and can be further classified into three families according to their structural homology to chymotrypsin, subtilisin, and wheat serine carboxypeptidase II (5,6). A fourth family of serine proteases has recently been identified based on amino acid sequence homology to prolyl oligopetidase (7). The catalytic serine and histidine residues in this enzyme have been identified by covalent modification with active-site-specific chemical reagents (8,9). Although it is expected that a catalytic triad is present in the prolyl oligopetidase family of serine proteases, identification of the catalytic aspartate awaits having a three-dimensional structure of the active site.

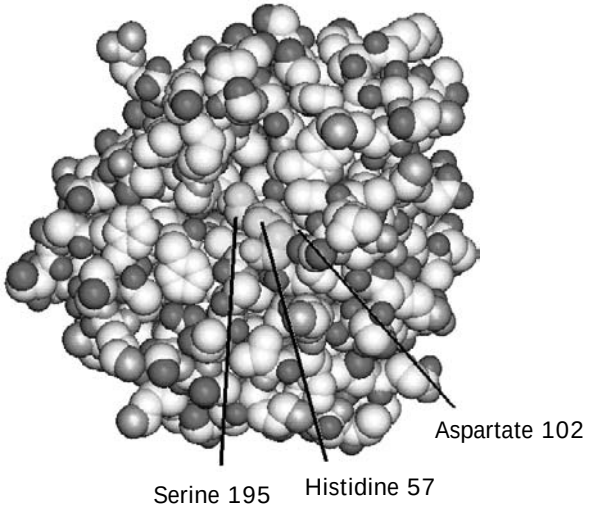


Figure 1. Three-dimensional structure of α-chymotrypsin. Residues forming the catalytic triad—serine 195, histidine 57, and aspartic acid 102—are located at the base of the substrate binding cleft. *Source*: Courtesy of Dr. Richard Bott, Genencor International Inc.

Chymotrypsin, a mammalian digestive enzyme, was the first protease to have its three-dimensional structure determined by X-ray crystallography (10). Other mammalian proteases structurally homologous to chymotrypsin but with different specificity are the digestive enzymes trypsin and elastase, and thrombin, which has specificity similar to trypsin and plays a critical role in blood clot formation. Subtilisin is a bacterial serine protease secreted extracellularly for the purpose of scavenging nutrients. Not all subtilisin-like serine proteases are bacterial in origin. Precursors of bioactive proteins such as hormones and neuropeptides are cleaved by serine proteases with greater structural homology to subtilisin than to chymotrypsin (11). The structure of wheat serine carboxypeptidase II has only recently been determined (12). Not only are the tertiary protein folds completely different for each structural family, the order of the catalytic triad residues in the amino acid sequence are also different. The arrangements are histidine-aspartate-serine, aspartate-histidine-serine, and serine-aspartate-histidine for chymotrypsin-like, subtilisin-like, and serine-carboxypeptidase-II-like serine proteases, respectively. Despite these differences, however, their reaction mechanism and geometry of catalytic residues are nearly identical and are believed to have arisen from different ancestral proteins through a process of convergent evolution.

Hydrolysis of amide and ester substrates by serine proteases follows an acyl transfer reaction pathway:

$$E + S \rightleftharpoons ES \xrightarrow{P_1} EP_2 \xrightarrow{H_2O} E + P_2$$

where ES represents the noncovalent enzyme-substrate Michaelis complex, EP_2 is the acylenzyme intermediate, and P_1 and P_2 are the reaction products generated during acylation and deacylation, respectively. Evidence for the acyl transfer mechanism was obtained from numerous kinetic studies using activated ester substrates where deacylation is the rate-determining step in the reaction such that accumulation and partitioning of the intermediate could be monitored (13). Further support came from direct crystallographic observation of a stable elastase acylenzyme intermediate formed at low temperature (14). X-ray crystal structures also revealed key structural features related to the reaction mechanism of serine proteases: the catalytic triad comprising serine 195, histidine 57, and aspartate 102 (residue numbers are from the chymotrypsin

amino acid sequence), and the oxyanion hole formed by the main chain NH groups of serine 195 and glycine 193 (15). A schematic of the reaction mechanism for serine proteases is illustrated in Figure 2. Formation of the covalent acylenzyme intermediate involves the nucleophilic attack of serine 195 on the substrate carbonyl carbon atom and release of the product P_1. This step is believed to proceed via general base-catalyzed transfer of a proton from the side chain hydroxyl group of serine 195 to the imidazole nitrogen of histidine 57, stabilization of the tetrahedral transition state by hydrogen bonds formed between the negatively charged carbonyl oxygen atom and the main chain NH groups of serine 195 and glycine 193, and general acid-catalyzed transfer of the proton of histidine 57 to the amine leaving group. Deacylation is essentially the same reaction except the attacking nucleophile is a water molecule bound at the site vacated by the P_1-leaving group. The pH dependence for catalysis by serine proteases follows the protonation state of the catalytic histidine residue. Chymotrypsin and subtilisin are active at neutral pH and above where the histidine side group is deprotonated. In contrast, wheat serine carboxypeptidase II is most active in the pH range of 4.5 to 5.5, suggesting an unusually low pK_a for its catalytic histidine residue. Although the role of the catalytic triad and oxyanion hole in the acyl transfer reaction mechanism of serine proteases was de-

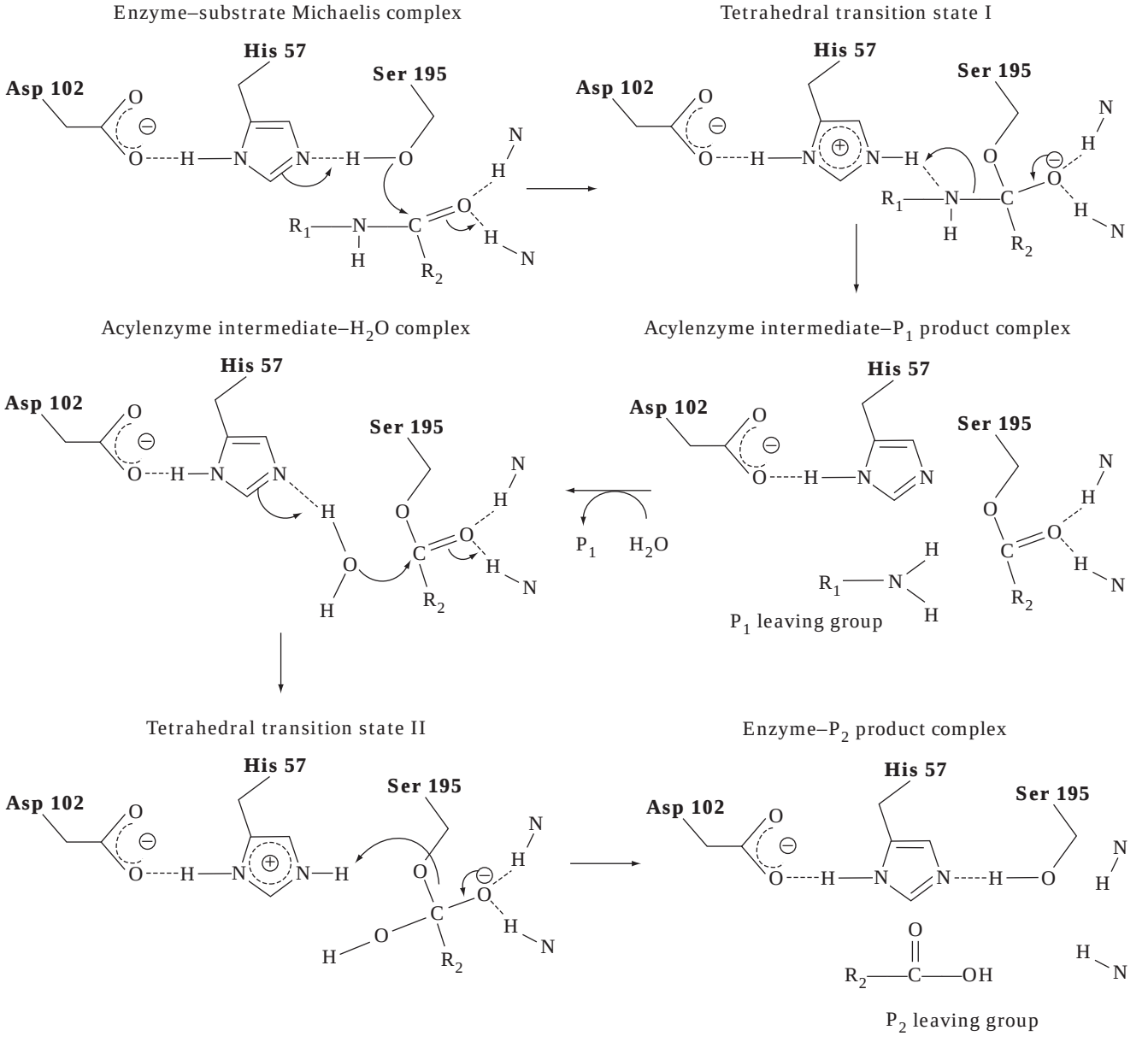


Figure 2. Schematic of the reaction mechanism in serine proteases. Nucleophilic attack of serine 195 on the peptide substrate bound in the substrate binding cleft forms an acylenzyme intermediate that is then hydrolyzed by a solvent water.

scribed more than 20 years ago, there continues to be considerable interest in understanding the fundamental source of their catalytic power.

Site-specific mutagenesis has been used extensively as a tool to better understand the catalytic mechanism of serine proteases. To assess the importance of transition-state stabilization, this technique was first applied to the oxyanion hole of subtilisin BPN' (16,17). Unlike chymotrypsin, where both hydrogen bonds stabilizing the substrate oxyanion are formed with main chain NH groups, the oxyanion hole in subtilisin BPN' consists of the main chain amide group of the catalytic serine 221 and one of the two protons on the side chain nitrogen atom of asparagine 155. Removal of the latter hydrogen bond via substitution of asparagine 155 with an amino acid lacking the necessary functional group or side chain geometry to form a good hydrogen bond to the oxyanion results in a reduction in the rate of catalysis by two to three orders of magnitude. In a similar fashion, the catalytic significance of aspartate 102 in trypsin was investigated by substitution with asparagine and replacement of each residue in the catalytic triad of subtilisin BPN' (serine 221, histidine 64, and aspartate 32) with alanine (18,19). The substitution of alanine for serine and histidine separately in subtilisin BPN' reduced the rate of catalysis by a factor of 10^6, whereas replacement of aspartate 32 with alanine led to a 10^4-fold loss in activity. Simultaneous substitution of all three amino acids in the catalytic triad of subtilisin with alanine also reduces catalysis by a factor of 10^6, illustrating the cooperative nature of these residues in catalysis. Assigning a functional role for aspartate 102 in the mechanism of serine proteases has been less straightforward than for either serine 195 or histidine 57. Initially, it was believed the high nucleophilicity of serine 195 was generated via a charge relay network in which histidine 57 accepts a proton from serine 195 and transfers another proton to aspartate 102 (20). This double-proton transfer reaction, however, has not been supported experimentally. It is now believed that aspartate 102, through a hydrogen bond to the side chain of histidine 57, plays the crucial roles of orienting histidine 57 in the correct tautomeric form to accept a proton from serine 195 and stabilizing the positive charge formed on histidine 57 in the transition state. Recently, it has been debated whether formation of a low-barrier hydrogen bond between the catalytic aspartate and histidine residues in the reaction transition state is an important feature in the catalytic mechanism of serine proteases (21–25). Although this issue remains to be resolved, it is evident that interactions between multiple residues work synergistically to stabilize the oxyanion transition state formed in the reaction mechanism of serine proteases (26).

Cysteine Proteases

The functional roles of cysteine proteases are quite broad, and members of this enzyme class are found in animals, plants, bacteria, and viruses. Like serine proteases, hydrolysis of amide and ester substrates by cysteine proteases follows an acyl transfer reaction pathway (equation 1), but the attacking nucleophile is the thiol group of an active-site cysteine residue instead of a serine hydroxyl group. The plant enzyme papain is the most extensively studied cysteine protease and the first to have its three-dimensional structure determined (27). It is a globular protein with the active site located in a cleft at the interface of two domains. The pH activity profile of papain is a bell-shaped curve centered about pH 6, with activity dependent on formation of a thiolate-imidazolium ion pair between the catalytic residues cysteine 25 and histidine 159. The loss in activity in the acidic pH range is attributed to protonation of cysteine 25 with an apparent pKa of 4, whereas the drop in activity in the basic pH range is attributed to deprotonation of histidine 159 with an apparent pKa of 8.5 (28). The side chain of asparagine 175 is hydrogen bonded to the imidazole of histidine 159, forming a cysteine-histidine-asparagine triad similar to the serine-histidine-aspartate catalytic triad found in serine proteases. The functional role attributed to asparagine 175 is that of orienting the imidazole ring of histidine 159. The importance of this residue for catalysis has been confirmed by screening of site-specific random amino acid replacements at position 175 in papain for catalytic activity (29). Glutamine at position 175 was found to have weak activity whereas no activity was observed for 11 other amino acids incorporated at position 175. Although a complete description of the acylation and deacylation steps in the reaction pathway of papain remains to be defined, it is clear there are significant differences in the mechanisms of cysteine and serine proteases.

Cysteine proteases have been referred to as activated enzymes because their thiolate-imidazolium ion pair present in the ground state obviates the need for general base-catalyzed removal of a proton to generate the nucleophilic thiolate anion in acylation and for general acid catalyzed proton transfer from histidine 159 during deacylation (30). A schematic of the reaction pathway for papain is shown in Figure 3. After formation of the noncovalent enzyme-substrate Michaelis complex, acylation proceeds via nucleophilic attack of the cysteine 25 thiolate anion on the carbonyl carbon of the substrate scissile bond. Breakdown of the tetrahedral transition state into a covalent acyl–enzyme intermediate is facilitated by rotation of the imidazole ring of histidine 159 and general acid-catalyzed transfer of its proton to the amide nitrogen of the leaving group. Hydrolysis of the acyl–enzyme intermediate during deacylation is base catalyzed with transfer of a proton from a water molecule to histidine 159 to generate the OH-nucleophile. Until recently, the existence of an oxyanion transition state in both acylation and deacylation has been the subject of considerable debate. A putative oxyanion hole in papain has been described comprising the main chain amide proton of cysteine 25 and the side chain NH_2 group of glutamine 19. Substitution of alanine and serine for glutamine at position 19 in papain via site-directed mutagenesis resulted in mutant enzymes having 60- and 600-fold decreases in catalytic efficiency, respectively, compared to wild-type enzyme (31). The effect of these mutations on the catalytic activity of papain are somewhat smaller than that observed for analogous mutations in the oxyanion hole of subtilisin. However, these results clearly demonstrate the importance of oxyanion stabilization in the reaction mechanism of cysteine proteases.

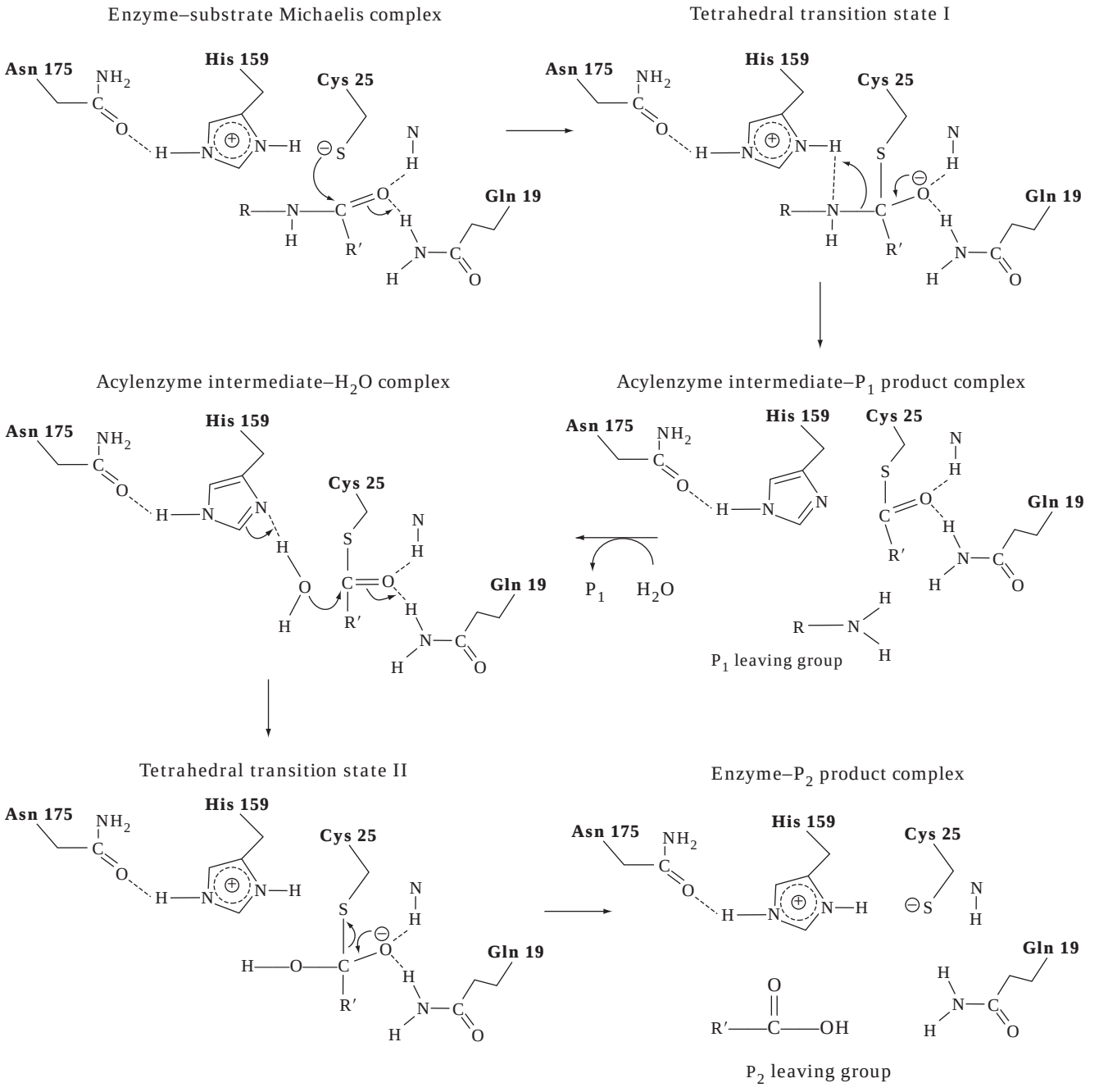


Figure 3. Schematic of the reaction mechanism in cysteine proteases. Nucleophilic attack of the thiolate anion of cysteine 25 on the peptide substrate bound in the substrate binding cleft forms an acylenzyme intermediate that is then hydrolyzed by a solvent water.

Aspartic Proteinases

The aspartic proteinases are named for the essential catalytic role of two aspartic acid residues. The three-dimensional structures for aspartic proteases of mammalian, fungal, and viral origin are quite comparable to each other. These enzymes are composed of two structurally similar domains with the active site formed within the cleft between each domain, much like the active site in cysteine proteases. Further, each domain contributes one of the two aspartate residues, aspartate 32 and aspartate 215, critical for catalysis (residue numbers are from the pepsin amino acid sequence) (32). The pH activity profile of aspartic proteinases is a bell-shaped curve, with activity dependent on protonation and ionization of the carboxylate group of aspartates 32 and 215, respectively. Most aspartic proteases are active in the range of pH 2 to 4, and in the past, this class of enzymes was referred to as acid proteases. Renin is one exception where activity is in the pH

range of 5.5 to 7.5. Inhibition by pepstatin is an important feature common to all aspartic proteases. Pepstatin, a natural microbial hexapeptide containing a statine residue, is believed to be a mimic of the reaction-transition state. Crystal structures of aspartic proteases with various analogues of pepstatin have provided valuable insight in elucidating the reaction mechanism and for the purpose of designing inhibitor drug molecules (33).

Unlike serine and thiol proteases, proteolytic cleavage of peptide bonds by aspartic proteases does not proceed via a covalent acyl intermediate. A water molecule bound at the active site between the carboxylates of aspartates 32 and 215 is believed to be the attacking nucleophile. Attack of the water on the scissile carbonyl is facilitated by the general base and general acid functions of aspartates 215 and 32, respectively. As shown in Figure 4, the ionized aspartate 215 abstracts a proton from the water nucleophile, and aspartate 32 transfers its proton to the carbonyl oxygen. Stabilization of the tetrahedral intermediate is achieved through hydrogen bonds formed between the gem-diol and the negatively charged aspartate 32 and between the aspartate 215 and one of the hydroxyl oxygens. Breakdown of the intermediate and release of products is facilitated by transfer of the proton from aspartate 215 to the amine nitrogen of the leaving group and deprotonation of the statinelike hydroxyl by aspartate 32. Additional stabilization of the tetrahedral intermediate via hydrogen bonds from the side chain of tyrosine 75 has also been proposed. Substitution of tyrosine 75 with other amino acid residues in calf stomach chymosin and *Rhizomucor pusillus* pepsin by site-directed mutagenesis results in modified enzymes with dramatically altered catalytic activity, suggesting a critical role of this residue for catalysis, presumably via formation of hydrogen bonds to one of the hydroxyls formed in the transition state (34).

Metallopeptidases

The catalytic activity of metallopeptidases is dependent on the presence of one or two metal ions, most commonly zinc, bound at the active site. The reaction mechanism for the two-metal ion enzymes remains to be fully understood, although they clearly comprise a class of metallopeptidases distinct from that of the single-metal ion enzymes. Significant differences in amino acid sequence and tertiary structure are observed both for single- and two-metal ion metallopeptidases isolated from bacteria, fungi, and higher organisms. Crystal structures have recently been determined for three different aminopeptidases having two-metal ions at their active site (35–37). The identification of catalytic groups and elucidation of a reaction mechanism for single-metal ion metallopeptidases is based largely on crystallographic structural analyses of the bacterial endopeptidase thermolysin and the pancreatic exopeptidase carboxypeptidase A (38). Although the three-dimensional structures of these two enzymes are very different, they share a similar orientation of reactive groups, suggesting convergent evolution of a common reaction mechanism. Like cysteine and aspartic proteases, single-metal ion metallopeptidases are typically bilobal proteins with the active site located at a cleft formed at the junction of the two domains. Another feature in common with aspartic proteases is formation of a noncovalent intermediate during catalysis, because the attacking nucleophile in the single-metal ion metallopeptidases is a zinc-bound water molecule.

Figure 5 shows a schematic of the reaction mechanism for thermolysin. Binding of substrate at the active site displaces the water nucleophile toward glutamate 143, with the carbonyl oxygen of the scissile bond polarized by the positively charged zinc ion. Glutamate 143 functions as a proton shuttle during catalysis, similar to the role of histidine in the catalytic triad of serine proteases. General base-catalyzed transfer of a proton from the nucleophilic water to glutamate 143 facilitates attack of the hydroxyl anion on the substrate carbonyl carbon atom. This proton is then passed to the amide nitrogen of the leaving group. Not surprisingly, substitution via site-specific mutagenesis of glutamate 143 with alanine in the metallopeptidase from *Bacillus subtilis* results in complete loss of enzyme activity (39). The negatively charged oxyanion formed in the transition state is believed to be stabilized via hydrogen bonds to the side groups of histidine 231 and tyrosine 157. Replacement of histidine 231 with alanine or phenylalanine by site-specific mutagenesis of the thermolysin-like endopeptidase isolated from *B. stearothermophilus* leads to a large reduction in k_{cat} and little change in K_M

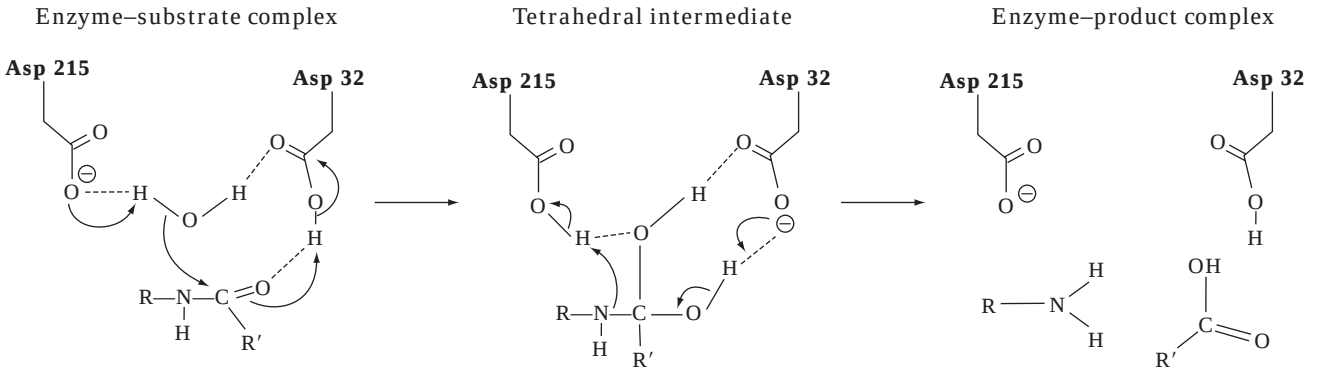


Figure 4. Schematic of the reaction mechanism in aspartic proteinases.

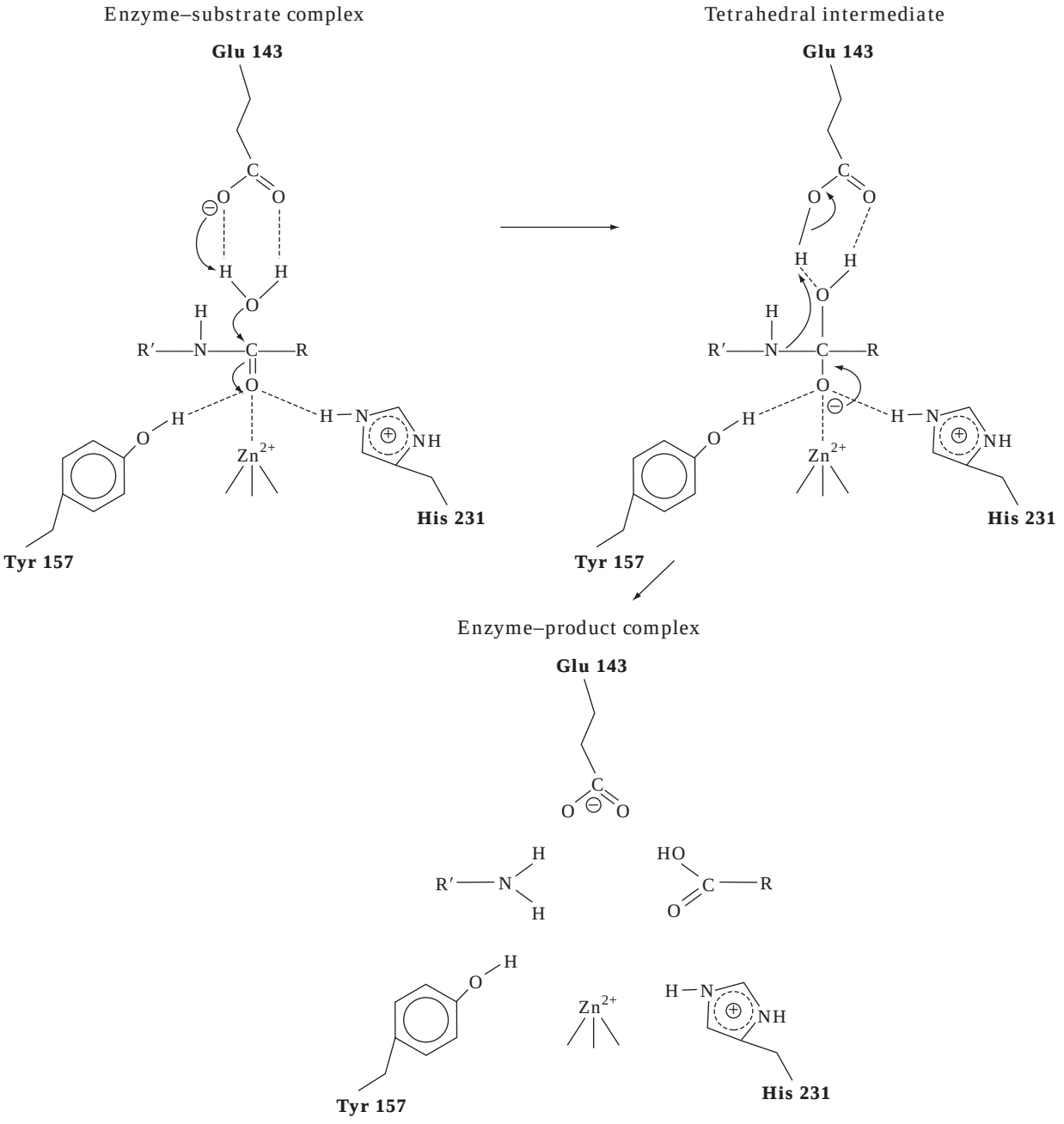


Figure 5. Schematic of the reaction mechanism of metallopeptidases.

(40). This result is consistent with assigning the role of transition state stabilization for histidine 231. A similar effect is obtained where arginine 127 in carboxypeptidase A, a mammalian exopeptidase, is substituted by site-specific mutagenesis with either alanine or methionine (41). The X-ray crystal structure of carboxypeptidase A reveals that arginine 127 is located at an active-site position equivalent to histidine 231 in thermolysin. Stabilization of the transition state in carboxypeptidase A is achieved via electrostatic interaction between the oxyanion and the positively charged guanidinium side group of arginine 127.

BINDING ENERGY IN CATALYSIS

A key feature used for rate acceleration by all four protease families is lowering of the reaction activation barrier through stabilization of the tetrahedral intermediate. Removal of the catalytic triad in the serine protease subtilisin via site-specific mutagenesis leads to an enzyme still capable of a 1,000-fold enhancement in the rate of peptide bond hydrolysis compared to the nonenzymatic reaction (16). Although much of the rate acceleration observed in the catalytically impaired subtilisin enzyme is likely attributable to the presence of a functional oxyanion hole,

the contribution of additional binding energy to catalysis is also possible. The γ-hydroxyl group of threonine 220 in subtilisin has been shown to stabilize the oxyanion transition state independent of the oxyanion hole (42). In addition, the role of long-range electrostatic interactions in subtilisin catalysis has been explored via measurement of the contribution of charged residues on the enzyme surface to binding of a transition state analogue inhibitor (43). Interaction of the substrate peptide chain with residues along the protease binding cleft is another source of stabilization energy.

The substrate binding cleft of proteases is often quite large, with numerous regions of contact, called subsites, for binding amino acid residues of the polypeptide substrate (44). Figure 6 shows the structure of a peptide substrate bound into the binding cleft of the serine protease subtilisin. Hydrogen bonding, salt bridge formation, and hydrophobic interactions between substrate and subsite residues create significant levels of binding energy important in determining substrate specificity and in facilitating catalysis. The rate of catalysis for the aspartic proteinase pepsin is observed to increase progressively as the length of the substrate polypeptide chain increases (45). Dramatic alterations in the substrate specificity profiles and catalytic activity of numerous proteases have been achieved via substitution of a limited number of subsite residues by site-directed mutagenesis (46). Protein structure outside of the binding cleft has also been found to affect the rate of peptide bond hydrolysis. For example, a recent site-specific mutagenesis experiment that altered the substrate specificity of trypsin to match that of chymotrypsin revealed that in addition to changes in the specificity pocket of the binding cleft, recruitment was required of two surface loop sequences that do not contact bound substrate (47). Although a description of the structural mechanism for how these loops influence catalysis remains to be determined, the importance of long-range enzyme–substrate interactions for transition-state stabilization is clearly apparent.

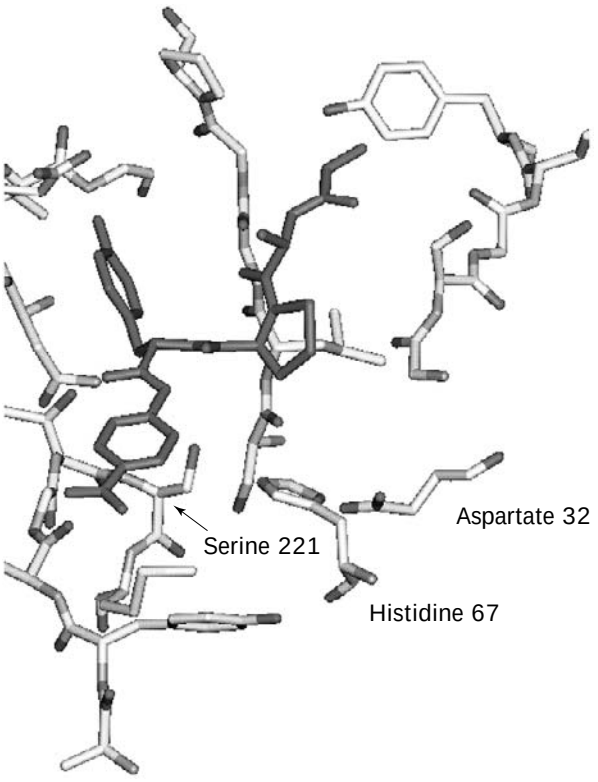


Figure 6. Close-up X-ray crystal structure view of the tetrapeptide succinyl-alanine-alanine-proline-phenylalanine-*p*-nitroanilide (solid dark lines) bound into the substrate binding cleft of subtilisin BPN′. The catalytic triad residues serine 221, histidine 64, and aspartic acid 32 are also shown. *Source*: Courtesy of Dr. Richard Bott, Genencor International Inc.

FUTURE DIRECTIONS

More than 50 years ago, Linus Pauling proposed that enzymes accelerate chemical reactions by providing a flexible template for binding the transition-state structure of a substrate molecule rather than its ground state (48). The structure-activity relationships elucidated from investigations on the reaction mechanisms of proteases provide an experimental framework that supports Pauling's theory on how enzymes catalyze chemical reactions. Although protease reaction mechanisms are still not completely understood, it is evident that essential catalytic groups can be readily identified from X-ray crystallographic and site-specific mutagenesis studies. The importance of having detailed information about the active-site geometry and chemical bonds stabilizing the reaction transition state has been established for the development of inhibitor drugs to treat protease-related diseases (49). Our knowledge of protease structure and function will also enable the facile discovery of new proteases from genetic sequence databases. The amount of genomic information available to researchers is growing at a fast rate. There are now 45 in-progress and 20 complete prokaryotic genome sequences, and 17 in-progress and 1 complete eukaryotic genome sequences available for analysis on the Internet (*www.mcs.anl.gov/home/gaasterl/genomes.html*). Computer searches for homologous sequence alignments and functional motifs such as metal–ion binding sites can rapidly identify protein sequences having similarity to a particular protease family. Monitoring of expression or genetic variability of the newly discovered gene via high-density DNA array technology can further our understanding of the physiological role of the new protease and its possible clinical significance (60).

The wealth of information known about protease reaction mechanisms has also led to exciting new applications of protein engineering in the creation of novel biocatalysts. For example, the activity of serine proteases for hydrolysis of amide and ester bonds can be modified by altering their active site structure via chemical modification or site-specific mutagenesis (51–55). Such a change in reaction specificity minimizes product degradation in protease-catalyzed peptide synthesis reactions. Incorporation of selenium at the active center of subtilisin results in an enzyme having significant peroxidase activity (56). More

recently, a protease having novel specificity was created via incorporating the catalytic triad of serine proteases into the substrate binding site of a nonprotease enzyme (57). The success of these and other experiments with proteases demonstrate the feasibility of redesigning the structure of a protein to produce a catalytic activity or specificity unobtainable from a naturally occurring enzyme.

BIBLIOGRAPHY

1. P. Carter, *Methods Enzymol.* **154**, 382–403, (1987).
2. R.J. Leatherbarrow and A.R. Fersht, *Protein Eng.* **1**, 7–16 (1986).
3. A.J. Barrett, *Methods Enzymol.* **244**, 1–15, (1994).
4. B.M. Dunn, in R.J. Benyon and J.S. Bond eds., *Proteolytic Enzymes: A Practical Approach*, IRL Press at Oxford University Press, Oxford, 1989, pp. 57–81.
5. B.W. Mathews, in H. Neurath and R.L. Hill eds., *The Proteins*, vol. 3, Academic Press, New York, 1977, pp. 404–590.
6. D.-I. Liao and S.J. Remington, *J. Biol. Chem.* **265**, 6528–6531 (1990).
7. N.D. Rawlings, L. Polgar, and A.J. Barrett, *Biochem. J.* **279**, 907–908 (1991).
8. D. Rennex, B.A. Hemmings, J. Hoffsteenge, and S.R. Stone, *Biochemistry* **30**, 2195–2203 (1991).
9. S.R. Stone, D. Rennex, P. Wikstrom, E. Shaw, and J. Hoffsteenge, *Biochem. J.* **276**, 837–840 (1991).
10. B.W. Mathews, P.B. Sigler, R. Henderson, D.M. Blow, *Nature* **214**, 652–656 (1967).
11. W.J.M. Van de Ven, A.J.M. Roebroek, and H.L.P. van Duijnhoven, *Crit. Rev. Oncol.* **4**, 115–136 (1993).
12. D.-I. Liao, K. Breddam, R.M. Sweet, T. Bullock, and S.J. Remington, *Biochemistry* **31**, 9796–9812 (1992).
13. M.L. Bender and J.V. Killheffer, *CRC Crit. Rev. Biochem.* **1**, 149–199 (1973).
14. T. Alber, G.A. Petsko, and D. Tsemoglou, *Nature* **263**, 297–300 (1976).
15. J. Kraut, *Annu. Rev. Biochem.* **46**, 331–358 (1977).
16. J.A. Wells, B.C. Cunningham, T.P. Graycar, and D.A. Estell, *Philos. Trans. R. Soc. Lond., Ser. A* **317**, 415–423 (1986).
17. P. Bryan, M.W. Pantoliano, S.G. Quill, H.-Y. Hsiao, and T. Poulos, *Proc. Natl. Acad. Sci. USA* **83**, 3743–3745 (1986).
18. C.S. Craik, S. Roczniak, C. Largman, and W.J. Rutter, *Science* **237**, 909–913 (1987).
19. P. Carter and J.A. Wells, *Nature* **332**, 564–568 (1988).
20. D.M. Blow, J.J. Birktoft, and B.S. Hartley, *Nature* **221**, 337–340 (1969).
21. P.A. Frey, S.A. Whitt, and J.B. Tobin, *Science* **264**, 1927–1930 (1994).
22. A. Warshel, A. Papazyan, and P.A. Kollman, *Science* **269**, 102–104 (1995).
23. W.W. Cleland and M.M. Kreevoy, *Science* **269**, 104 (1995).
24. P.A. Frey, *Science* **269**, 104–106 (1995).
25. S.-O. Shan, S. Loh, and D. Herschlag, *Science* **272**, 97–101 (1996).
26. A. Warshel, G. Naray-Szabo, F. Sussman, and J.-K. Hwang, *Biochemistry* **28**, 3629–3637 (1989).
27. I.G. Kamphuls, K.H. Kalk, M.B.A. Swarte, and J. Drenth, *J. Mol. Biol.* **179**, 233–256 (1984).
28. A.C. Storer and R. Menard, *Methods Enzymol.* **244**, 486–500 (1994).
29. T. Vernet, J. Chatellier, D.C. Tessler, and D.Y. Thomas, *Protein Eng.* **6**, 213–219 (1993).
30. L. Polgar and B. Asboth, *J. Theor. Biol.* **121**, 323–326 (1986).
31. R. Menard, J. Carriere, P. Laflamme, C. Plouffe, H.E. Khouri, T. Vernet, D.C. Tessler, D.Y. Thomas, and A.C. Storer, *Biochemistry* **30**, 8924–8928 (1991).
32. D.R. Davies, *Annu. Rev. Biophys. Biophys. Chem.* **19**, 189–215 (1990).
33. B. Veerapandian, J.B. Cooper, A. Sall, T.L. Blundell, R.L. Rosati, B.W. Dominy, D.B. Damon, and D.J. Hoover, *Protein Sci.* **1**, 322–328 (1992).
34. Y.-N. Park, J.-I. Aikawa, M. Nishiyama, S. Horinouchi, and T. Beppu, *Protein Eng.* **9**, 869–875 (1996).
35. N. Strater and W.N. Lipscomb, *Biochemistry* **34**, 14792–14800 (1995).
36. S.L. Roderick and B.W. Mathews, *Biochemistry* **32**, 3907–3912 (1993).
37. B. Chevrier, C. Schalk, H. D'Orchymont, J.-M. Rondeau, D. Moras, and C. Tamus, *Structure* **2**, 283–291 (1994).
38. B.W. Mathews, *Acc. Chem. Res.* **21**, 333–340 (1988).
39. S. Toma, S. Campagnoli, E. De Gregorils, R. Gianna, I. Margarit, M. Zamal, and G. Grandi, *Protein Eng.* **2**, 359–364 (1989).
40. A. Beaumont, M.J. O'Donohue, N. Paredes, N. Rousselet, M. Assicot, C. Bohuon, M.-C. Fournie-Zaluski, and B.P. Roques, *J. Biol. Chem.* **270**, 16803–16808 (1995).
41. M.A. Phillips, R. Fletterick, and W.J. Rutter, *J. Biol. Chem.* **265**, 20692–20698 (1990).
42. S. Braxton and J.A. Wells, *J. Biol. Chem.* **266**, 11797–11800 (1991).
43. S.E. Jackson and A.R. Fersht, *Biochemistry* **32**, 13909–13916 (1993).
44. I. Schechter and A. Berger, *Biochem. Biophys. Res. Commun.* **27**, 157–162 (1967).
45. K. Medzihradzky, I.M. Voynick, H.M. Schweiger, and J.S. Fruton, *Biochemistry* **9**, 1154–1162 (1970).
46. J.J. Perona and C.S. Craik, *Protein Sci.* **4**, 337–360 (1995).
47. L. Hedstrom, L. Szllagyl, and W.J. Rutter, *Science* **255**, 1249–1253 (1992).
48. L. Pauling, *Chem. Eng. News* **24**, 1375–1377 (1946); see also *Nature* **161**, 707–709 (1948) and *Am. Sci.* **38**, 51–58 (1948).
49. E.W. Petrillo and M.A. Ondetti, *Med. Res. Rev.* **2**, 1–41 (1982).
50. L. Wodicka, H. Dong, M. Mittmann, M.-H. Ho, and D.J. Lockhart, *Nature Biotechnol.* **15**, 1359–1367 (1997).
51. J.B. West, W.J. Hennen, J.A. Bibbs, J.L. Lalonde, Z. Zhong, E.F. Meyer, and C.-H. Wong, *J. Am. Chem. Soc.* **112**, 5313–5320 (1990).
52. T. Nakatsuka, T. Sasaki, and E.T. Kaiser, *J. Am. Chem. Soc.* **109**, 3808–3810 (1987).
53. Z.-P. Wu and D. Hilvert, *J. Am. Chem. Soc.* **111**, 4513–4514 (1989).
54. P.R. Bonneau, T.P. Graycar, D.A. Estell, and J.B. Jones, *J. Am. Chem. Soc.* **113**, 1026–1030, (1991).
55. T. Graycar, R. Bott, R. Caldwell, J. Dauberman, P. Lad, S. Power, H. Sagar, R. Silva, G. Weiss, L. Woodhouse, and D. Estell, *Ann. N.Y. Acad. Sci.* **672**, 71–79 (1992).
56. Z.-P. Wu and D. Hilvert, *J. Am. Chem. Soc.* **112**, 5647–5648 (1989).
57. E. Quememeur, M. Moutiez, J.-B. Charbonnier, and A. Menez, *Nature* **391**, 301–304 (1998).

PSEUDOMONAS, PROCESS APPLICATIONS

ANDREW SCHMID
MARCEL WUBBOLTS
MARTIN HELD
BERNARD WITHOLT
Swiss Federal Institute of Technology
Zurich, Switzerland

KEY WORDS

Biocatalysis
Bioconversion
Bioprocesses
Biotransformation
Industrial processes
Pseudomonas

OUTLINE

INTRODUCTION

Pseudomonads are known to be widely encountered in nature, occupying niches ranging from contaminated soil to infected mammalian tissues. The extraordinary metabolic versatility of these species has been noticed since the early days of their study (1,2), probably because of the capability of many pseudomonads to share and acquire genetic information required to degrade a large number of compounds (3,4).

Because of the wide variety of degradation reactions and metabolic pathways existing in nature, it is not surprising that screening programs have led to the isolation of numerous bacterial strains, capable of using countless substrates. Interestingly, many substrates that can be degraded by *Pseudomonas* species include not only water-soluble chemicals but also highly hydrophobic and toxic compounds. The observation that (partially) purified enzymes and even intact cells can be efficiently used in mixed-solvent systems or in the presence of anhydrous solvents has led to widespread study and application of biocatalysts in organic synthesis during the past two decades.

The growing demand for optically pure drugs and agrochemicals, devoid of enantiomeric ballast, and the improvement of chemical catalysts have initiated development toward enantiospecific synthesis routes for chiral substances. However, the production of enantiopure products using chemical catalysts is not always feasible, and alternative chemical routes usually tend to be tedious, involving costly and dirty process steps. Therefore, biocatalysis exploiting the enzymatic activity of cells evolved in nature presents a valuable tool, with a large economical potential, for organic synthesis (Fig. 1).

BIOSYNTHESIS

A well-known biotechnological production process for a natural compound on an industrial scale with the help of *Pseudomonas* strains is the production of vitamin B-12 by *P. denitrificans* (7), which is especially efficient in media (such as beet molasses) that are rich in betaine (8,9). The production of vitamin B-12 by methylotrophic *Pseudomonas* sp. AM-1 on the cheap carbon source methanol (10) has not been able to replace the successful original industrial process (7).

Microbial polysaccharides are important biotechnological products that have found applications in the food, agricultural, and chemical industries (11–13). Exopolysaccharide production is common to *P. fluorescens*, *P. putida*, *P. aeruginosa*, *P. cepacia*, and other *Pseudomonas* (14–16). These compounds are believed to serve as a defense mechanism against the host immune system and as protection against bacteriophages or dehydration. Plant pathogens may use it as an anchoring material (17). Surfactants excreted by hydrocarbon degrading species improve the mobilization and bioavailability of mineral oils in soil (18,19).

The production of *Pseudomonas* exopolysaccharides such as alginate from *P. aeruginosa* and gellan gum from *P. elodea* have been commercialized, but these polysaccharides have not reached the level of application of xanthan gum, the principally used microbial polysaccharide from *Xanthomonas campestris* (17). The role of alginates in cystic fibrosis has triggered numerous scientific efforts to elucidate the molecular biology of their synthesis (20–22), which may well result in more efficient ways to produce this polysaccharide that is frequently applied to immobilize cells and enzymes in biotransformation procedures.

Other biopolymers that are of industrial importance are the bacterial polyesters: poly-3-hydroxyalkanoates (23,24). These biopolymers are biodegradable and provide an environmentally friendly alternative to common plastics. The commercialization of Biopol, a poly-3-hydroxybutyrate-poly-3-hydroxyvalerate copolymer from *Alcaligenes eutrophus* by Zeneca (ICI) has been realized (25), and the development of other biodegradable plastics is underway.

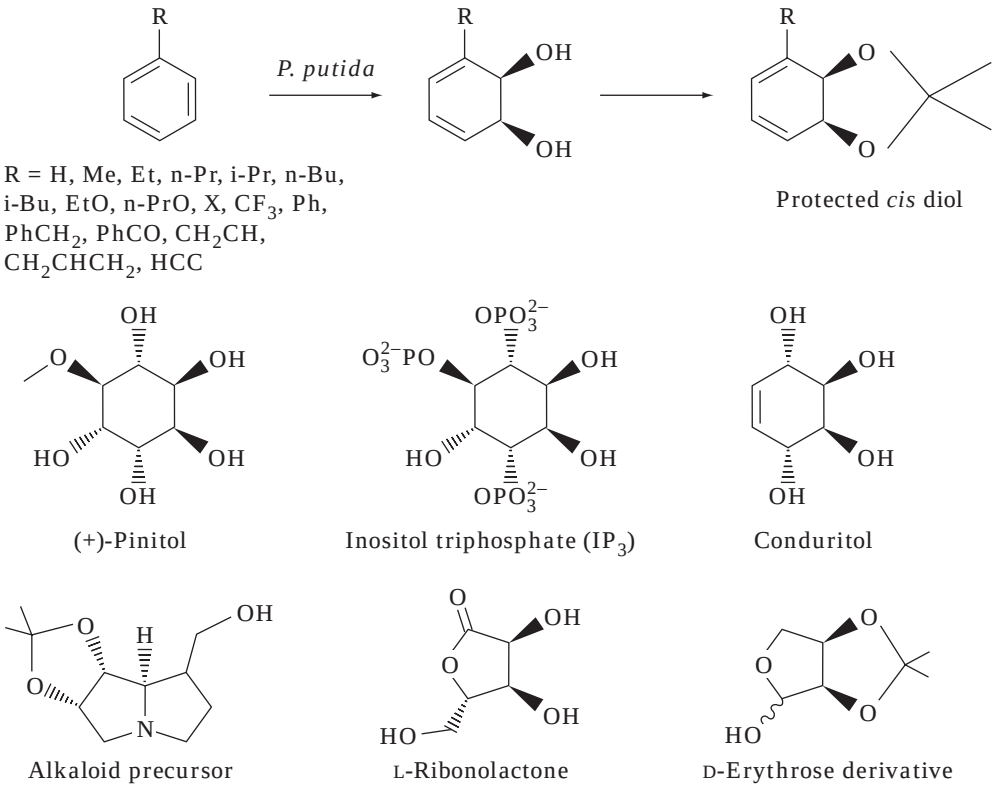


Figure 1. Synthesis of variously substituted *cis*-dihydrodiols by *P. putida* and several examples of complex molecules made using this synthon (5,6).

Poly-3-hydroxyalkanoates from *P. oleovorans* (26), an alkane-degrading microorganism (27,28), are of particular interest, because these bacterial polyesters can be produced with variable side groups, depending on the (co)substrate added to the culture (24,29,30). Biopolymers with variable-length side chains, unsaturated side chains, and aromatic pendant groups have been produced by this microorganism (24), but large-scale commercial application of these medium-chain-length poly-3-hydroxyalkanoates has yet to be realized.

BIOTRANSFORMATIONS

The use of enzymes in biocatalytic processes, either in (partially) purified form or in whole cell systems, is not evenly distributed over the different enzyme classes, defined by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (formerly Enzyme Committee [EC]) (31).

Biocatalysts based on the hydrolases (EC class 3) are most used in synthetic reactions, because these enzymes are cofactor independent; are often secreted from the cells into the extracellular medium, facilitating their isolation, and in some cases are commercially available. Furthermore, many hydrolases have been shown to maintain their activity in media containing bulk amounts of hydrophobic, organic solvents. Oxidoreductases (EC class 1) are relatively complex enzymes that require cofactors and frequently consist of more than one protein component and are therefore mostly used as whole cell biocatalysts. Despite their complexity and difficulty of handling, these biocatalysts are important tools by virtue of their regio- and stereoselectivity. The enzymes that catalyze addition–elimination reactions, lyases (EC class 4), are underrepresented in biotransformations, although some biocatalysts of this type are used on a large scale in industrial applications, such as for the production of acrylamide (32). The development of biocatalysts that catalyze group transfer reactions (transferases, EC class 2) or isomerizations (isomerases, EC class 5) as well as ligase-based biotransformation reactions (EC class 6) is in a relatively early stage, with few realized industrial applications. Enzymes of these classes generally are highly substrate specific, which makes them less suited for wide-range biocatalytic applications.

Redox Reactions

Oxidation and reduction reactions catalyzed by oxidoreductases require cofactors, such as $NAD(P)^+$ or NAD(P)H, as electron acceptors or donors. Because the addition of cofactor to an enzyme reaction mixture in stoichiometric quantities is economically unfeasible, oxidative and reductive bioconversions are usually performed using whole cell cultures, where metabolically active cells efficiently recycle the cofactor. However, recent developments of efficient in situ cofactor regeneration systems (5,33) have stimulated the use of dehydrogenases in enzymatic biocatalyses.

Useful oxidoreductases encompass predominantly dehydrogenases, which catalyse the addition or elimination of hydrogen, and oxygenases, catalyzing the direct incorporation of molecular oxygen into organic molecules. Known hydrogenation reactions catalyzed by pseudomonads include the reduction of various aliphatic, cyclic, and aromatic ketones and aldehydes (34–36). *Pseudomonas*-catalyzed oxidations of hydroxy groups have been reported for substrates such as morphine, codeine, and (*S*)-malic acid (37,38).

Monooxygenases are a class of oxygenases that catalyze the addition of single oxygen atoms introducing hydroxy groups or epoxy moieties to aliphatic and olefinic compounds as well as aromatic rings. Additional reactivities include Baeyer-Villiger reactions, sulfoxidation reactions, and oxidative dehalogenations (5,39). The former two reaction types appear as promising alternatives to chemical synthesis pathways because of their (stereo-) specificity. Monooxygenase-catalyzed dehalogenation reactions performed by pseudomonads have not proved useful for synthetic biotransformation reactions. On the other hand, these reactions are of importance for biodegradation purposes.

Dioxygenases, a second class of oxygenases, drive the dihydroxylation of aromatic compounds to optically active *cis*-diols (Fig. 1). Often, these monooxygenase and dioxygenase-catalyzed hydroxylation reactions represent the initial steps in microbial degradation pathways of organic compounds. With the necessary genes in hand, recombinant strains have been developed that catalyze these initial oxidation reactions only, without further degradation, yielding specifically oxidized organic bioconversion products (40,41).

Generally, these enzyme systems possess a broad substrate range as well as a high regio- and stereospecificity (Fig. 2). In addition, different enzyme systems have been shown to ideally complement each other, such as the alkane hydroxylase system (EC 1.14.15.3) from *P. oleovorans* (43) and xylene oxygenase (EC 1.14.15.-) from *P. putida* mt-2 (44). For example, xylene oxidizes pendant methyl groups on aromatic rings (45), whereas alkane hydroxylase does not accept toluene derivatives as substrates but will introduce a terminal hydroxy group on ethyl side chains (46,47).

Because of their versatility and specificity, these enzyme systems represent valuable biocatalysts for the synthesis of chiral organic compounds, and they are being exploited in industrial applications for the production of fine chemicals (48,49) and pharmaceuticals (50).

Group Transfer Reactions

Transferases catalyze group transfer reactions. Although some acyltransferases, alkyltransferases, amino transferases, glycosyl transferases and phosphoryl transferases have found their way into biocatalytic processes, the application of transferases from pseudomonads on an industrial scale is rare.

Transferase-catalyzed reactions reported in the academic and patent literature include amino- (51,52), glycosyl- (53,54), and phospho-transfer reactions (55). The alkyl transfer reaction for the formation of L-serine from the achiral amino acid glycine and methanol was realized using serine hydroxymethyl transferase (EC 2.1.2.1) from *Pseudomonas* MS31, depicted in Figure 3 (56).

Hydrolyzations

Most industrial and laboratory enzyme processes are based on hydrolases. These enzymes do not use cofactors, usually consist of one catalytic component, and are readily available from commercial suppliers. Furthermore, the ability of hydrolases to catalyze synthesis reactions in nonaqueous environments has been thoroughly investigated and has provided paths to synthesize complex esters and amides.

Processes with pseudomonad hydrolases include reactions for the cleavage of ester, amide, and glycosidic bonds, the hydrolysis of epoxides to diols and nitriles to carboxylic acids, and the dehalogenation of chlorinated organic compounds (5,57). Their principle value lies in their contribution to the enantiopure synthesis of chemicals, often achieved by resolving racemic starting compounds or intermediates readily obtained from inexpensive precursors by chemical synthesis (Fig. 4). Also, many enzymatic reactions proceed under milder conditions and produce less-hazardous waste products compared to the chemical reaction alternative (5), which explains why especially nitrilases are chosen even when chemo-, regio- or enantioselectivity is not required for the synthesis of carboxylic acids from nitriles.

Hydrolytic reactions catalyzed by *Pseudomonas* enzymes are used as intermediate steps in various industrial processes for the synthesis of compounds such as natural and unnatural amino acids and peptides, organic acids, alcohols, and esters. Applications of these products range from food additives and agrochemicals to pharmaceuticals, including antibiotics, anticancer drugs, anti-inflammatory drugs, and β-blockers.

Lipases are the most applied hydrolases in biotransformations because these enzymes are highly enantioselective and are readily available from commercial suppliers. The pseudomonads are particularly rich in catalytically useful lipases. Lipases (EC 3.1.1.3) from *P. fluorescens*, *Pseudomonas* sp., *P. cepacia*, and *P. aeruginosa* have been and continue to be major players in the field of enantioselective ester hydrolysis and esterification reactions with technical or commercial applications (59–61). (Fig. 5)

Interestingly, proteases (EC 3.4.-.-) from pseudomonads have not received much attention, and they are hardly commercially available, although these hydrolytic enzymes obtained from other sources (e.g. subtilisin, chymotrypsin, papain) are very frequently used in many industrial and laboratory applications.

The discovery that many hydrolytic enzymes function well in organic solvents has lead to an even broader application spectrum of these biocatalysts, because it is possible to use substrates and to produce products unstable in water systems, and synthesis reactions have become feasible in low-water systems by eliminating the hydrolytic reaction dominant in aqueous media (62). For example lipases can be directed to catalyze esterifications or transesterifi-

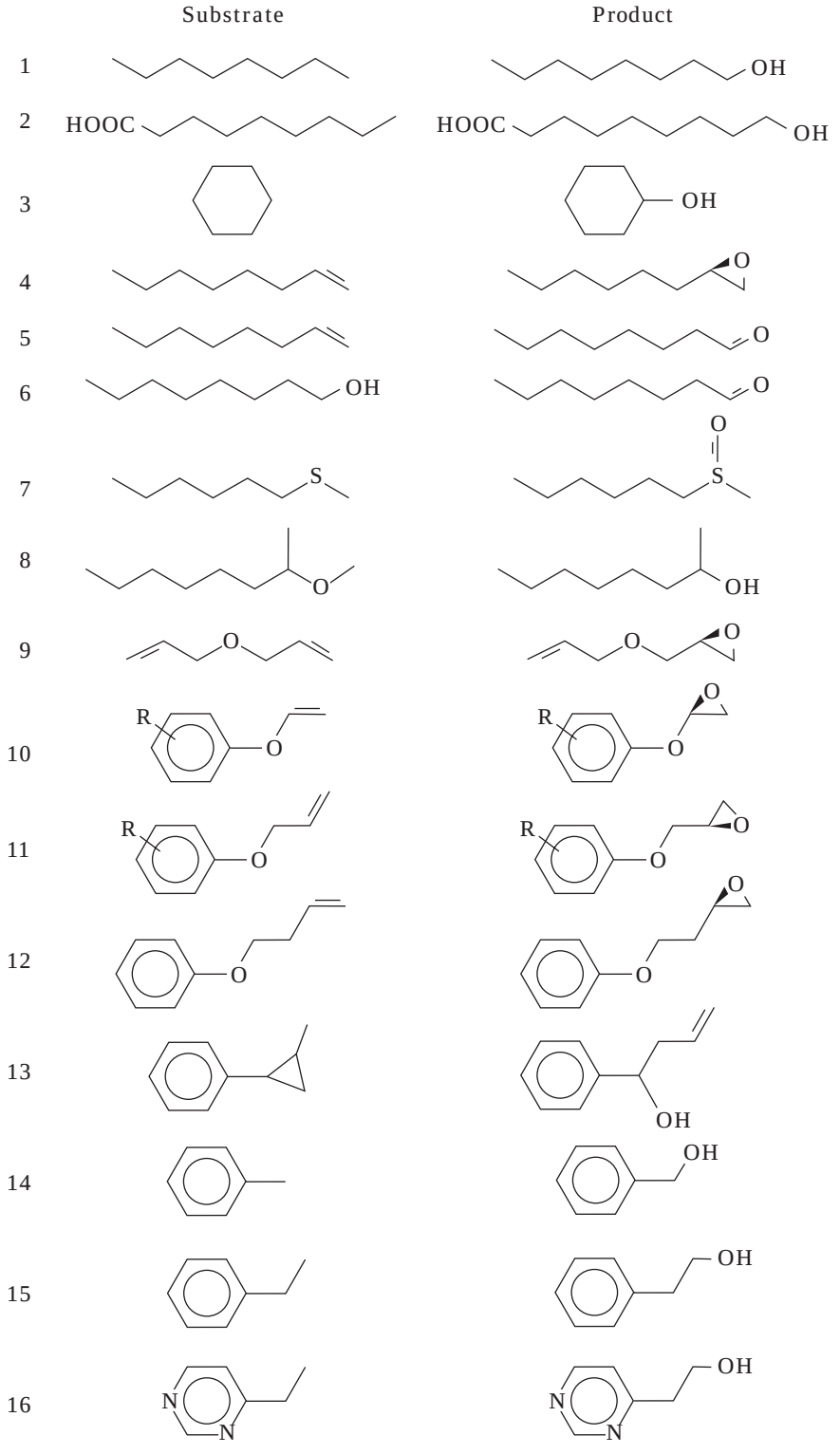


Figure 2. Substrate range of the alkane hydroxylase system (EC, 1.14.15.3) (42). Examples of the classes of substrates are shown. 1, C6–C14 alkanes to corresponding 1-alkanols; 2, C6–C12 fatty acids to ω-hydroxy fatty acids; 3, cyclohexane to cyclohexanol; 4, C3–C12 1-alkenes to 1,2-epoxyalkanes; 5, 1-octene to 1-octanal; 6, 1-octanol to 1-octanal; 7, methyl hexyl thioether to corresponding sulfoxide (and heptylthiol); 8, branched ethers to alcohol (demethylation); 9, diallyl ether to corresponding epoxide; 10, ring-substituted allylbenzene to 3-phenyl-1,2-epoxypropane; 11, ring-substituted allyl phenyl ethers to corresponding epoxides; 12, substituted allyl benzyl ethers to corresponding epoxides; 13, (*R*,*S*)-*trans*-2-phenyl-1-methylcyclopropane to 1-phenyl-3-butene-1-ol; 14, toluene to benzylalcohol; 15, ethylbenzene to 2-phenylethanol; 16, ethyl-substituted heterocyclic 5- and 6-membered aromatic rings to corresponding alcohols.

cation reactions such as acylation or alcoholysis. Increased conformational and thermal stability of water-soluble enzymes in hydrophobic solvents allows features such as molecular imprinting, by lyophilizing the enzyme from an aqueous solution at an ideal pH or in the presence of the preferred substrate, enzyme activity and specificity can be influenced (63).

Addition and Elimination Reactions

Lyases catalyze additions to double bonds or eliminations that generate double bonds and are specific for carbon–carbon, carbon–oxygen, carbon–nitrogen, carbon–sulfur, carbon–halide, or phosphorus–oxygen bonds. A strong enantioface discrimination during addition reactions is typ-

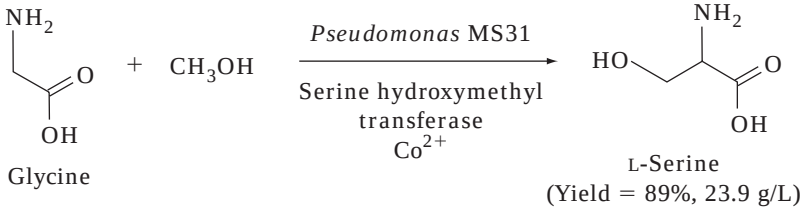


Figure 3. Biotransformation of glycine to L-serine using *Pseudomonas* (56).

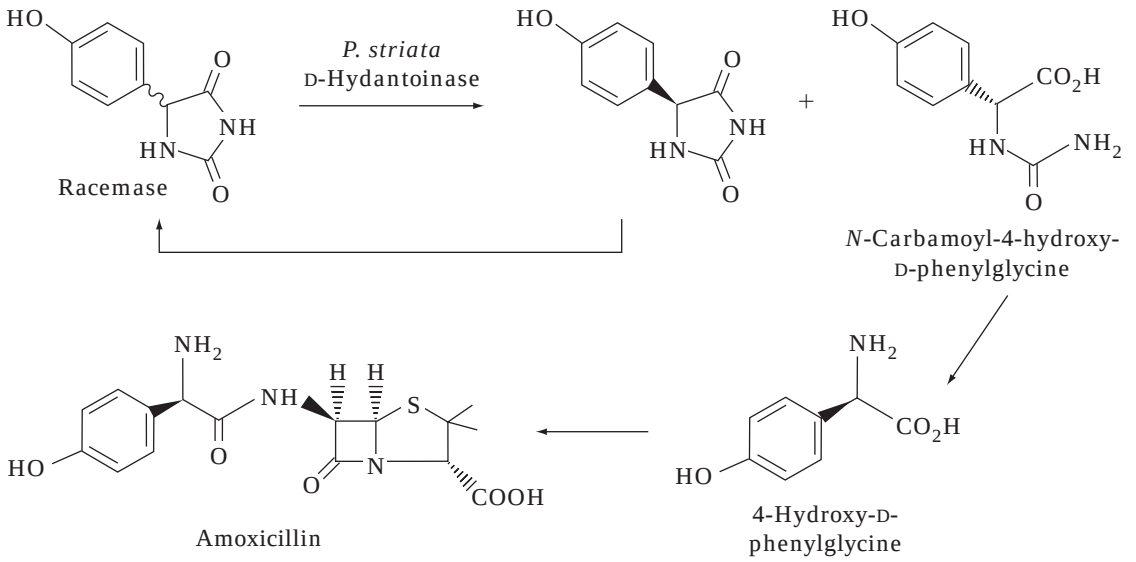


Figure 4. Amoxicillin production with a *P. striata* hydantoinase, a hydrolase that cleaves hydantoin rings to linear hydantoic acids (58).

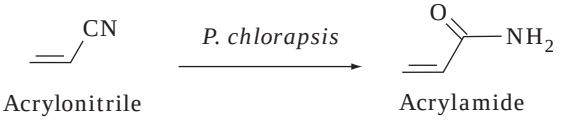


Figure 5. The production of acrylamide from acrylonitrile by Nitto using *P. chlorapsis* or *R. rhodochrous* J1 acrylonitrile hydrolyase (nitrile hydratase) (32).

ical of lyases, and lyase-catalyzed eliminations are as a rule also highly enantioselective, which makes them useful tools for chiral synthesis (64). Unfortunately, most lyases have a narrow substrate spectrum, which limits the applicability of these enzymes in biocatalytic processes (5).

Nevertheless, a number of industrial applications exploiting *Pseudomonas* lyases do exist (57). These include aminolyases for the synthesis of the artificial sweetener aspartame, L-aspartyl-L-phenylalanine methyl ester (51,65,66); carnitine hydrolyase (EC 4.2.1.89), containing whole cells of *Pseudomonas* sp. used by Lonza to produce the food additive L-carnitine (67,68), and haloyases, used for the synthesis of D-cysteine derivatives, important building blocks for the semisynthetic production of cephalosporin, (*S*)-carboxymethxy-D-cysteine. Originally, the hydrolyase acrylonitrile hydrolase from *P. chlororaphis* B23 (EC 4.2.1.84) was used as a catalyst for the production of acrylamide from acrylonitrile by the Nitto company on a scale of 30,000 tons per year (59), but since 1991 *R. rhodochrous* J1 has been used for production because of its superior reaction characteristics (69).

Isomerizations

Amino acid racemases (EC 5.1.1.10) catalyze the interconversion of D- and L-amino acids and play an important role in the industrial production of some amino acids, because racemization of the undesired or nonconverted enantiomer makes a 100% yield of the desired enantiomer attainable (59). In a number of processes, multienzyme reaction mixtures that include racemases have been used for the stereoselective synthesis of various amino acids (70,71) (Fig. 6).

The use of carbohydrate isomerases from pseudomonads is limited, but some processes have been described for the synthesis of fructose, mannose, and psicose (53,72,73).

Ligations

Ligases catalyze the joining of two molecules at the expense of the hydrolysis of a high-energy phosphate bond and are of particular use in molecular biology. In the bio-

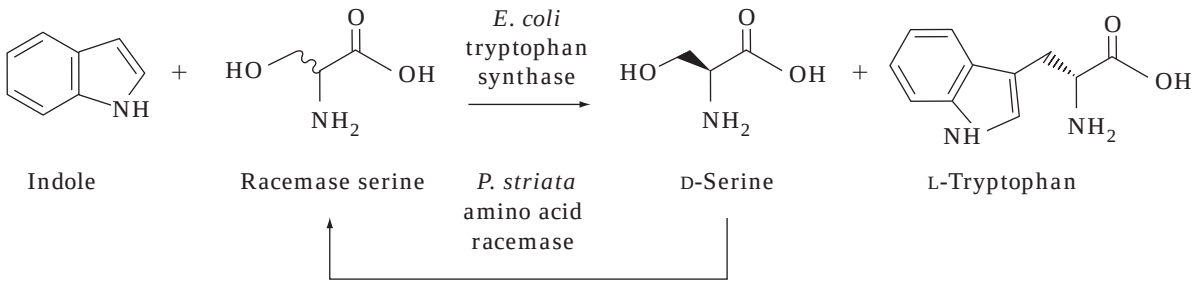


Figure 6. The Mitsui Toatsu process for production of L-tryptophan using serine racemase from *P. striata* (69).

transformation field, however, these enzymes are applied to a very limited extent (31).

Phenyl acetyl CoA ligase (EC 6.2.1.21) from *P. putida* M has been used in conjunction with *Penicillum chrysogenum* acyl-CoA:6-APA acyltransferase to produce a variety of penicillin analogues (74). Because of the wide substrate range of the enzymes involved (75), this system allows the in vitro synthesis of a wide variety of β-lactam antibiotics such as 3-furylmethylpenicillin and several ketoalkylpenicillins (76,77).

BIOPROCESS TECHNOLOGY

Toxicity of Apolar Organic Solvents

Many substrates and products involved in biosynthesis and biotransformation reactions inhibit the catalytic activity of cells or enzymes above critical concentrations. Which organic solvents cause inhibitory effects is most reliably predicted by the logarithm of the octanol-water partitioning coefficient (log*P*) (78). In general, complete biocatalyst inhibition is observed when apolar compounds with a log*P* below 2 are added to liquid cell suspension cultures at or above water-saturating concentrations. Above values of 4 to 5, inhibitory effects are markedly reduced. Gram-positive bacteria, which remain active in the presence of solvents with log*P* values above 4.5 to 6, show a higher sensitivity to solvents than Gram-negative bacteria, of which *Pseudomonas* species belong to the most resistant, where growth is observed in the presence of solvents with log*P* values as low as 3 (79–81).

Unless specific interactions with enzymes exist that limit growth, apolar chemicals generally harm cells as a result of their partitioning into the cell membrane bilayers. As a consequence, higher membrane fluidity and permeability results, leading to the loss of cell functionality and eventually cell lysis (82–84).

To withstand the exposure to organic solvents, cells react by increasing their membrane hydrophobicity and by decreasing membrane fluidity. The latter compensates the fluidizing effects of solvent molecules intercalated in the lipid bilayers, whereas changes in hydrophobicity have been observed to reduce the partitioning of solvents to the membranes (85,86). These effects are achieved by the incorporation of longer C_{18} fatty acids and proteins into the cell membrane, by modulation of the ratio between saturated and unsaturated fatty acids, by changes in the phospholipid composition, and by conversion of *cis*-unsaturated fatty acids to their *trans*-isomers (87–90). For a comprehensive review on the subject of cellular adaptation to organic solvents, see the work by Weber and de Bont (91).

Strain selection (92,93) and genetic engineering (94–96) are increasingly providing strains with improved resistance against elevated concentrations of organic chemicals. Alternative strategies to minimize cell damage and growth inhibition include cell immobilization (97–99). A range of carrier materials has been used with cells either adsorbed to the surface or entrapped within beads and gels. Examples of such materials are Celite, calcium alginate, κ-carrageenan, polyurethane, polyacrylamide, silicon, and chitosan (100,101). However, it must be considered that cell immobilization introduces an additional diffusion barrier, which can lead to mass transfer limitations (102).

Despite these advances, special procedures during bioprocess operations are nevertheless required to avoid toxic effects by harmful substrates and products when high productivities and high final product titers are to be attained in commercial applications. Strategies to attain this goal are described in the next section.

Controlled Substrate Feed

To avoid substrate inhibition, it is necessary to feed toxic starting compounds to a cell culture or an enzyme-containing reaction mixture in a controlled way, at a rate that avoids substrate accumulation to inhibitory levels. This is achieved either by supplying volatile compounds via the gas phase (103,104) or via controlled liquid feeds (105,106) in laboratory as well as production-scale reactors.

An alternative way—limited to laboratory studies—to eliminate substrate inhibition is to add toxic starting material dissolved in apolar carrier solvents, such as hexadecane and other high log*P* solvents. Because hydrophobic substrates mainly partition into the apolar solvent, they are less exposed but still well available to the biocatalyst residing in the aqueous phase of the two-liquid phase mixture (Fig. 7). Thus, higher quantities of toxic material can be added to the reaction mixture without encountering inhibitory effects. An example where such a two-liquid phase operation may in the future be scaled up for production purposes is the biosynthesis of functionalized poly-3-hydroxyalkanoates (PHA) from toxic precursors (24,107).

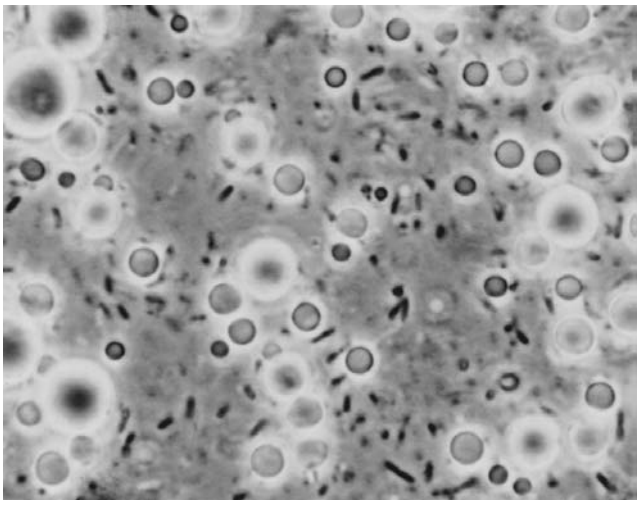


Figure 7. Micrograph of two-liquid phase medium. *P. oleovorans* cells (dark rods) are suspended throughout the continuous aqueous phase. The organic solvent *n*-octane is dispersed as fine droplets, with an average diameter in the order of 10 μm.

These precursors must be present in concentrations below inhibiting levels, but above limiting concentrations, because cells accumulate PHA as a carbon storage material only in abundance of the carbon source (108). Because the second liquid phase will buffer fluctuations of substrate uptake rate, as a result of continuously varying substrate uptake rates in batch and fed-batch processes, concentrations should be controlled more easily within the required narrow range.

In Situ Product Extraction

To obtain high titers of toxic or inhibitory products, it is necessary to continuously remove product from the reaction vessel. Depending on the nature of the product and the bioprocess involved, various methods are feasible.

Two-Liquid Phase Operation. By operating a two-liquid, aqueous-organic phase system, it is possible to continuously extract hydrophobic products formed by the biocatalyst residing in the aqueous phase (Fig. 7). As a result, cells and enzymes see less of the toxic products accumulating mainly in the organic phase during the course of the bioprocess, thus enabling higher final product titers to be attained. Compounds successfully produced in such in situ extraction processes include aliphatic and aromatic epoxides (41,109,110), alcohols (111,112) optically pure derivatives of citronellol (113) and organic acids (114). These biotransformations have been limited to laboratory scales; however, various industries are developing related processes.

Model studies have shown that the extractant can be separated from the aqueous phase by thermal and physical means (115) or, as recently developed, by continuously removing the organic solvent from the reaction mixture during the biotransformation step by in situ centrifugation (116). Product isolation and purification from the organic phase proceeds via conventional, chemical downstream processing techniques (117,118).

Solid Phase Extraction. Solid adsorbents have been successfully used to recover small molecules such as acetone, butanol, terpenes, aromatic amino acids, and steroids from fermentation broth. In these procedures, cells are first removed from the medium (119,120) or they are immobilized on solid carrier material (121). Solid adsorbent in situ recovery procedures have also been used in the presence of living cells, for example, to achieve enhanced production of antibiotics by streptomycetes (122), or mycobacteria (123). The addition of solid adsorption material such as XAD-4 to cultures was especially useful for adsorption of alkaloids from plant cell cultures (124–126). Solid adsorbents can be added into the bioreactor either directly or separated from the cells by a membrane. Alternatively, culture medium can be cycled from the bioreactor via an external loop through a packed or fluidized adsorption bed (127).

Membrane Separation. When using (partly) purified enzymes in biotransformation processes, it is often important to recycle the biocatalyst to optimize productivity and to minimize operation costs. This may be achieved by immobilizing the enzymes to a solid support and filtering the reaction mixture either continuously during the biotransformation reaction or batchwise at the end of the bioconversion step. Alternatively, soluble enzymes may be used in a homogeneous reaction mixture with ultrafiltration membranes used to separate the high molecular weight enzymes from the low molecular weight metabolites (128).

In oxidoreductase catalyzed-reactions, cofactors such as $NAD(P)^+$ and NAD(P)H are required as electron acceptors or donors. Because the addition of these metabolites in stoichiometric amounts is uneconomical, they must be regenerated. This can be achieved by supplementing the reaction mixture with a second enzyme, which catalyzes a redox reaction running in the opposite direction (129). For example, formate dehydrogenase (EC 1.2.1.2) added together with formate catalyzes the oxidation to carbon dioxide, thereby recycling the electron acceptor NAD^+ to NADH, which can again donate electrons to the desired enzymatic reaction (35). Because the native cofactor permeates through ultrafiltration membranes, NAD has been covalently bound to a water-soluble polymer, such as poly(ethylene glycol) (PEG).

OUTLOOK

Ongoing studies to elucidate anabolic and catabolic pathways and to characterize the enzymes that catalyze the individual reactions will, together with a multitude of screening programs, undoubtedly lead to further discoveries and applications of new and useful biotransformation reactions. In this process, genetic engineering will continue to play an important role for the characterization of enzymes as well as their production by cultures of cloned enzyme producing organisms. Genetically engineered whole cell biocatalysts can be beneficial when undesired

side reactions of the host organism hamper a process and when space and time yields are insufficient. Modern techniques promise to allow an even more effective construction of biocatalytically active cell strains than is achieved today. By exerting a more sophisticated control over enzyme expression and regulation as well as by designing whole cell biocatalysts containing genetic information from various sources, whole cell biocatalysts tailor-made for specific processes will evolve.

New process technologies, useful for biotransformation reactions, that have been developed during recent years include in situ product removal (127), membrane reactor systems (128), and two-liquid phase bioprocesses (130). These techniques significantly widen the potential of biocatalysis, opening applications involving cofactor-dependent enzyme processes, bioconversions of toxic compounds, and syntheses of organic molecules in organic media. Thus, reactions catalyzed by biological systems, which formerly were restricted to the field of chemistry, have become feasible. Chemistry and biology are moving together to provide interchangeable tools for the efficient production of important chemicals useful in diverse applications in medicine, agriculture, and nutrition.

BIBLIOGRAPHY

1. L.E. den Dooren de Jong, Ph.D. Thesis, Leiden University, The Netherlands, 1926.
2. R.Y. Stanier, N.J. Palleroni, and M. Doudoroff, *J. Gen. Microbiol.* **43**, 159–271 (1966).
3. B. Holloway, M. Escuadra, A. Morgan, R. Saffery, and V. Krishnapillai, *FEMS Microbiol. Lett.* **100**, 101–105 (1992).
4. J.R. van der Meer, *FEMS Microbiol. Rev* **15**, 239–249 (1994).
5. K. Faber, *Biotransformations in Organic Chemistry*, Springer, Berlin, 1997.
6. T. Hudlicky, H.F. Olivo, and B. McKibben, *J. Am. Chem. Soc.* **116**, 5108–5115 (1994).
7. J. Florent, in H.-J. Rehm, and G. Reed, eds., Biotechnology, vol. 4, VCH Verlaggesellschaft, Weinheim, Germany, 1986, pp. 117–158.
8. Y.H. Fa, J.P. Kusel, and A.L. Demain, *Appl. Environ. Microbiol.* **47**, 1067–1069 (1984).
9. J.P. Kusel, Y.H. Fa, and A.L. Demain, *J. Gen. Microbiol.* **130**, 835–841 (1984).
10. Y. Tsuchiya, and N. Nishio, *J. Ferment. Technol.* **58**, 485–487 (1980).
11. J.D. Desai and I.M. Banat, *Microbiol. Mol. Biol. Rev.* **61**, 47–64 (1997).
12. V.J. Morris, *Agro Food Industry Hi-Tech* **3**, 3–8 (1992).
13. I.W. Sutherland, *Biotechnology of Microbial Exopolysaccharides*, Cambridge Univ. Press, Cambridge, U.K. 1991.
14. D.G. Allison and M.J. Goldsbrough, *J. Basic. Microbiol.* **34**, 3–10 (1994).
15. A.M. Marques, I. Estanol, J.M. Alsina, C. Fuste, and D. Simon-Pujol, *Appl. Environ. Microbiol.* **52**, 1221–1223 (1986).
16. I.W. Sutherland, E. Conti, A. Flaibani, and M. O'Regan, *Microbiology* **140**, 1125–1132 (1994).
17. A.N. Glazer, and H. Nikaido, *Microbial Biotechnology: Fundamentals of Applied Microbiology*, Freeman, New York, 1995.
18. D.S. Francy, J.M. Thomas, R.L. Raymond, and C.H. Ward, *J. Ind. Microbiol.* **8**, 237–245 (1991).
19. Y. Zhang, and R.M. Miller, *Appl. Environ. Microbiol.* **61**, 2247–2251 (1995).
20. V. Deretic, D.W. Martin, M.J. Schurr, M.H. Mudd, N.S. Hibler, et al., *Biotechnology* **11**, 1133–1136 (1993).
21. L.O. Martins and I. Sa-Correia, *Enzyme Microbial Technol.* **13**, 385–389 (1991).
22. M.J. Schurr, D.W. Martin, M.H. Mudd, and V. Deretic, *J. Bacteriol.* **176**, 3375–3382 (1994).
23. A.J. Anderson and E.A. Dawes, *Microbiol. Rev* **54**, 450–472 (1990).
24. A. Steinbüchel, and H.E. Valentin, *FEMS Microbiol. Lett.* **128**, 219–228 (1995).
25. *Eur. Plastic News* **12**, 38 (1985).
26. M.J. de Smet, G. Eggink, B. Witholt, J. Kingma and H. Wynberg, *J. Bacteriol.* **154**, 870–878 (1983).
27. R.D. Schwartz, *Appl. Microbiol.* **25**, 574–577 (1973).
28. R.D. Schwartz, and C.J. McCoy, *Appl. Microbiol.* **26**, 217–218 (1973).
29. R.G. Lageveen, G.W. Huisman, H. Preusting, P. Ketelaar, G. Eggink, and B. Witholt, *Appl. Environ. Microbiol.* **54**, 2924–2932 (1988).
30. R.W. Lenz, Y.B. Kim, and R.C. Fuller, *FEMS Microbiol. Rev.* **103**, 207–214 (1992).
31. WWW site *http://alpha.gmw.ac.uk/~ugca000/iupac/enzyme*.
32. Eur. Pat. 93782 (1983), H. Yamada, and Y. Tani.
33. C. Wandrey, and B. Bossow, *Biotechnol. Bioind.* **3**, 8–13 (1986).
34. C.W. Bradshaw, H. Fu, G.J. Shen, and C.H. Wong, *J. Org. Chem.* **57**, 1526–1532 (1992).
35. W.O. Pat. 9318138 (Sept. 16, 1993) M.R. Kula and J. Peters, (to Forschungszentrum Juelich GmbH).
36. S. Shimizu, and H. Yamada, *J. Syn. Org. Chem. Japan* **49**, 52–70 (1991).
37. M. Deamici, C. Demicheli, G. Molteni, D. Pitre, G. Carrea, S. Riva, S. Spezia, and L. Zetta, *J. Org. Chemistry* **56**, 67–72 (1991).
38. S.I. Suye, S. Yokoyama, and A. Obayashi, *J. Ferment. Bioeng.* **68**, 301–304 (1989).
39. D.B. Janssen, F. Pries, and J.R. Vanderploeg, *Annu. Rev. Microbiol* **48**, 163–191 (1994).
40. N. Mermod, J.L. Ramos, and K.N. Timmis, *J. Bacteriol.* **167**, 447–454 (1986).
41. M.G. Wubbolts, O. Favre-Bulle, and B. Witholt, *Biotechnol. Bioeng.* **52**, 301–308 (1996).
42. J.B. van Beilen, Ph.D. Thesis, Rijksuniversiteit Groningen, 1994.
43. M. Lee and A.C. Chandler, *J. Bacteriol.* **41**, 373–386 (1941).
44. D.A. Kunz and P.J. Chapman, *J. Bacteriol.* **146**, 179–191 (1981).
45. M.G. Wubbolts, P. Reuvekamp, and B. Witholt, *Enzyme Microb. Technol.* **16**, 608–615 (1994).
46. J.B. van Beilen, J. Kingma, and B. Witholt, *Enzyme Microb. Technol.* **16**, 904–911 (1994).
47. M.G. Wubbolts, R. Noordman, J.B. van Beilen, and B. Witholt, *Recl. Trav. Chim. Pays-Bas* **114**, 139–144 (1995).
48. A. Kiener, *Angew. Chem., Int. Ed. Engl.* **31**, 774–775 (1992).
49. A. Kiener *Chemtech* 31–35 (Sept. 1995).

50. J.B. Johnston, and V. Renganathan, *Enzyme Microb. Technol.* **9**, 706–708 (1987).
51. Eur. Pat. 289,846 (Nov. 9, 1988), R. Marquardt, J. Then, B. Braeu, P. Praeve, and G. Woehner, (to Hoecst).
52. H.I. Rhee, K. Murata, and A. Mimura, *Agric. Biol. Chem.* **51**, 1701–1702 (1987).
53. U.S. Pat. 5,049,494 (Sept. 17, 1991), P. Allenza, (to Allied-Signal Inc.)
54. S. Maruo, H. Yamamoto, M. Toda, N. Tachikake, M. Kojima, and Y. Ezure, *Biosci. Biotechnol. Biochem.* **57**, 499–501 (1993).
55. J. Alvarez-Jacobs, D. Court, and G. Guarneros, *Biotechnol. Lett.* **12**, 425–430 (1990).
56. M. Watanabe, Y. Morinaga, and H. Enei, *J. Ferment. Technol.* **65**, 617–620 (1987).
57. M.G. Wubbolts, submitted for publication.
58. S. Takahashi, *Hakko Kogaku Kaishi* **61**, 139–151 (1983).
59. K. Drauz and H. Waldmann, *Enzyme Catalysis in Organic Synthesis: A Comprehensive Handbook*. VCH Verlagsgesellschaft, Weinheim, Germany, 1995.
60. K.E. Jaeger and S. Wohlfarth, *Bioengineering* **9**, 39–46 (1993).
61. F. Theil, *Chem. Rev.* **95**, 2203–2227 (1995).
62. A.L. Gutman and M. Shapira, *Adv. Biochem. Eng. Biotechnol.* **52**, 87–128 (1995).
63. J.S. Dordick, *Curr. Opin. Biotechnol.* **2**, 401–407 (1991).
64. J. van der Werf, W.J.J. van den Tweel, J. Kamphuis, S. Hartmans, and J.A.M. de Bont, *Trends Biotechnol.* **12**, 95–103 (1994).
65. J.S. Takagi, R. Fukunaga, M. Tokushige, and H. Katsuki, *J. Biochem.* **96**, 545–552 (1984).
66. Eur. Pat. B 1 0102529 (May 10, 1989), H. Tuneo, T. Hisao, and I. Tatsuo, (to Toso Corp.).
67. G.H. Kulla, *Chimia* **86**, 295–323 (1991).
68. Eur. Pat. 410,430 (Jan. 30, 1991), F. Hocks (to the Lonza AG, Gampel/Wallis, Switzerland).
69. T. Nagasawa, H. Shimizu, and H. Yamada, *Appl. Microbiol. Biotechnol.* **40**, 189–195 (1993).
70. K.I. Ishiwata, M. Fukuhara, M. Shimada, N. Makiguchi, and K. Soda, *Biotechnol. Appl. Biochem.* **12**, 141–149 (1990).
71. U.S. Pat. 4,605,625 (Aug. 12, 1986), H. Yamada, S. Shimizu, and S. Shiozaki, (to Nippon Zeon).
72. H. Itoh, T. Sato, and K. Izumori, *J. Ferment. Bioeng.* **80**, 101–103 (1995).
73. Y. Takasaki, K. Hinoki, Y. Kataoka, S. Fukuyama, and N. Nishimura, *J. Ferment. Bioeng.* **76**, 237–239 (1993).
74. H. Martinez-Blanco, A. Reglero, J. Martin-Villacorta, and J.M. Luengo, *FEMS Microbiol. Rev.* **72**, 113–116 (1990).
75. M. Fernandez-Valverde, A. Reglero, M.-B. H., and J.M. Luengo, *Appl. Environ. Microbiol.* **59**, 1149–1154 (1993).
76. M.A. Ferrero, A. Reglero, H. Martinez-Blanco, M. Fernandez-Valverde, and J.M. Luengo, *Antimicrob. Agents Chemother.* **35**, 1931–1932 (1991).
77. H. Martinez-Blanco, A. Reglero, and J.M. Luengo, *J. Antibiotics* **44**, 1252–1258 (1991).
78. C. Laane, S. Boeren, K. Vos, and C. Vergeer, *Biotechnol. Bioeng.* **30**, 81–87 (1987).
79. A. Inoue and K. Horikoshi, *J. Ferment. Bioeng.* **71**, 194–196 (1991).
80. A.N. Rajagopal, *Enzyme Microb. Technol.* **19**, 606–613 (1996).
81. M. Vermue, J. Sikkema, A. Verheul, R. Bakker, and J. Tramper, *Biotechnol. Bioeng.* **42**, 747–758 (1993).
82. Q. Chen, A. Nijenhuis, H. Preusting, J. Dolfing, D.B. Janssen, and B. Witholt, *Enzyme Microb. Technol.* **17**, 647–652 (1995).
83. J. Sikkema, J.A.M. de Bont, and B. Poolman, *J. Biol. Chem.* **269**, 8022–8028 (1994).
84. J. Sikkema, J.A.M. de Bont, and B. Poolman, *Microbiol. Rev.* **59**, 201–222 (1995).
85. M.C. Antunes-Madeira, and V.M.C. Madeira, *Biochim. Biophys. Acta* **901**, 61–66 (1987).
86. M.A. Ascon-Cabrera and J.M. Lebeault, *J. Ferm. Bioeng.* **80**, 270–275 (1995).
87. H. Alexandre, B. Mathieu, and C. Charpentier, *Microbiology* **142**, 469–475 (1996).
88. Q. Chen, D.B. Janssen, and B. Witholt, *J. Bacteriol.* **177**, 6894–6901 (1995).
89. H.C. Pinkart, J.W. Wolfram, R. Rogers, and D.C. White, *Appl. Environ. Microbiol.* **62** 1129–1132 (1996).
90. F.J. Weber and J.A.M. de Bont, *Meded. Fac. Landbouwwet. Rijksuniv. Gent* **59**, 2295–2392 (1994).
91. F.J. Weber and J.A.M. de Bont, *Biochim. Biophys. Acta* **1286**, 225–245 (1996).
92. K. Moriya, S. Yanagitani, R. Usami, and K. Horikoshi, *J. Mar. Biotechnol.* **2**, 131–133 (1995).
93. J.L. Ramos, E. Duque, M.J. Huertas, and A. Haidour, *J. Bacteriol.* **177**, 3911–3916 (1995).
94. R. Aono, T. Negishi and H. Nakajima, *Appl. Environ. Microbiol.* **60**, 4624–4626 (1994).
95. H. Asako, H. Nakajima, K. Kobayashi, M. Kobayashi, and R. Aono, *Appl. Environ. Microbiol.* **63**, 1428–1433 (1997).
96. H. Nakajima, H. Kobayashi, T. Negishi, and R. Aono, *Biosci. Biotech. Biochem.* **59**, 1323–1325 (1995).
97. R. Bar and J.L. Gainer, *Biotechnol. Prog.* **3**, 109–114 (1987).
98. L.E.S. Brink, J. Tramper, K.C.A.M. Luyben, and K. Van't Riet, *Enzyme Microb. Technol.* **10**, 736–743 (1988).
99. M.D. Hocknull and M.D. Lilly, *Appl. Microbiol. Biotechnol.* **33**, 148–153 (1990).
100. A.C.P. Dias, J.M.S. Cabral, and H.M. Pinheiro, *Enzyme Microb. Technol.* **16**, 708–714 (1994).
101. J.Y. Houng, W.P. Chiang, K.C. Chen, and C. Tiu, *Enzyme Microb. Technol.* **16**, 485–491 (1994).
102. E.G. Ceen, J.P.R. Herrmann, and P. Dunnill, *Appl. Microbiol. Biotechnol.* **25**, 491–494 (1987).
103. U.S. Pat. 4,833,078 (May 23, 1989), J.H. Hsieh (to the Celgene Corp.).
104. U.S. Pat. 5,236,832 (Sept. 7, 1993), A. Kiener (to Lonza Ltd.).
105. D.G.H. Ballard, A.J. Blacker, J.M. Woodley, and S.C. Taylor in D.P. Mobley ed., *Plastics from Microbes: Microbial Synthesis of Polymers and Polymer Precursors*, Hanser Publishers, Munich, 1994, pp. 139–168.
106. K. Gbewonyo, B.C. Buckland, and M.D. Lilly, *Biotechnol. Bioeng.* **37**, 1101–1107 (1991).
107. H. Preusting, R. van Houten, A. Hoefs, E.K. van Langenberghe, O. Favre Bulle, and B. Witholt, *Biotechnol. Bioeng.* **41**, 550–556 (1993).
108. H.G. Schlegel, G. Gottschalk, and R. von Bartha, *Nature* **191**, 463–465 (1961).

109. M.J. de Smet, J. Kingma, H. Wynberg, and B. Witholt, *Enzyme Microb. Technol.* **5**, 352–360 (1983).
110. K. Kawakami, in H., Okada, A. Tanaka, and H.W. Blanch, eds., *Enzyme Engineering X*, vol. 613, New York Academy of Sciences, New York, 1990, pp. 707–711.
111. A. Bosetti, J.B. van Beilen, H. Preusting, R.G. Lageveen, and B. Witholt, *Enzyme Microb. Technol.* **14**, 702–708 (1992).
112. S.R. Roffler, T.W. Randolph, D.A. Miller, H.W. Blanch, and J.M. Prausnitz, in B. Mattiason and O. Holst, eds., *Extractive Bioconversions*, Dekker, New York, 1990, pp. 133–172.
113. S. Oda, Y. Inada, A. Kato, N. Matsudomi, and H. Ohta, *J. Ferment. Bioeng.* **80**, 559–564 (1995).
114. V.M. Yabannavar and D.I.C. Wang, *Biotechnol. Bioeng.* **37**, 1095–1100 (1991).
115. R.G. Mathys, Ph.D. Thesis, Eidgenössische Technische Hochschule Zürich, Switzerland, 1997.
116. A. Kollmer, Ph.D. Thesis, Eidgenössische Technische Hochschule Zürich, Switzerland, 1997.
117. R.G. Mathys, O.M. Kut, and B. Witholt, *J. Chem. Technol. Biotechnol.*, in press (1998).
118. R.G. Mathys, A. Schmid, O.M. Kut, and B. Witholt, *J. Chem. Technol. Biotechnol.*, in press (1998).
119. J.P. Corry, W.L. Reed, and W.R. Curtis, *Biotechnol. Bioeng.* **42**, 503–508 (1993).
120. L. Nielsen, M. Larsson, O. Holst, and B. Mattiasson, *Appl. Microbiol. Biotechnol.* **28**, 335–339 (1988).
121. M.H.L. Ribeiro, D.M.F. Prazeres, J.M.S. Cabral, and M.M.R. da Fonseca, *Bioprocess Eng.* **12**, 95–102 (1995).
122. H.P. Fielder, M. Nega, C. Pfefferle, I. Groth, C. Kempter, H. Stephan, and J.W. Metzger, *J. Antibiot.* **49**, 758–764 (1996).
123. V. Hecht, J. Vorlop, H. Kalbitz, K. Gerth, and J. Lehmann, *Biotechnol. Bioeng.* **29**, 222–227 (1987).
124. K.D. Green, N.H. Thomas, and J.A. Callow, *Biotechnol. Bioeng.* **39**, 195–202 (1992).
125. R.J. Robins, A.J. Parr, and M.J.C. Rhodes, *Biochem. Soc. Trans.* **16**, 67–71 (1988).
126. R.D. Williams, N. Chauret, C. Bedard, and J. Archambault, *Biotechnol. Bioeng.* **40**, 971–977 (1992).
127. A. Freeman, J.M. Woodley, and M.D. Lilly, *Bio / Technology* **11**, 1007–1012 (1993).
128. A.S. Bommarius, K. Drauz, U. Groeger, and S. Wandrey, in A.N. Collins, G.N. Sheldrake, and J. Crosby eds, *Chirality in Industry*, Wiley, Chichester, U.K., 1992, pp. 371–397.
129. P. Adlercreutz, *Biocat. Biotransf.* **14**, 1–30 (1996).
130. B. Witholt, M.J. de Smet, J. Kingma, J.B. van Beilen, M. Kok, R.G. Lageveen, and G. Eggink, *Tibtech* **8**, 46–52 (1990).

PULLULAN, MICROBIAL PRODUCTION METHODS

JULES THIBAULT
ANH LEDUY
Laval University
Sainte-Foy, Canada

KEY WORDS

Aureobasidium pullulans
Fermentation
Hyperbaric
Microbial polysaccharide
Mixing
Oxygen transfer
Pigmentation
Pullulan
Pullulanolysis
Reciprocating plate bioreactor
Rheology

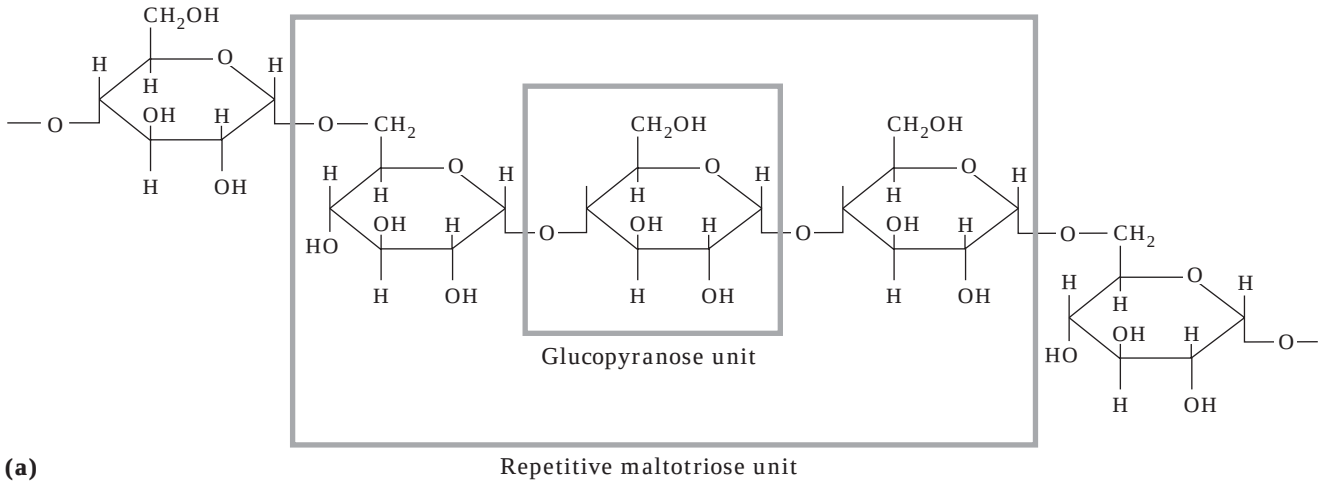
OUTLINE

WHAT IS PULLULAN?

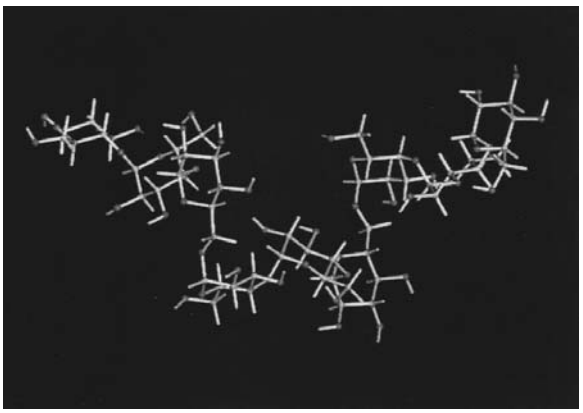
Pullulan is a water-soluble extracellular polysaccharide synthesized by the yeastlike fungus *Aureobasidium pullulans* (formerly named *Pullularia pullulans*) from a variety of mono- and disaccharides (1). It is generally accepted that pullulan is a neutral glucan that consists of either a linear chain of D-glucopyranosyl units that alternate regularly between one (1→6)-α-D and two (1→4)-α-D linkages, or a linear polymer of maltotriosyl units connected by (1→6)-α-D linkages (Fig. 1). However, pullulan containing 5–7% maltotetraosyl units has also been reported for several strains of *A. pullulans* (2,3).

Freeze-dried pullulan is white in color, tasteless, odorless, nonhygroscopic, nontoxic, biodegradable, and edible. It is insoluble in many solvents including methanol, ethanol, and acetone, but soluble in water to form a transparent, colorless, viscous, and adhesive solution. Shaped bodies as well as very thin pullulan films 0.01 mm thick can be formed. Thin films of pullulan are colorless, transparent, tasteless, odorless, tenacious, resistant to oil and grease, unaffected by small thermal variations, impermeable to oxygen, nontoxic, biodegradable, and edible (4,5).

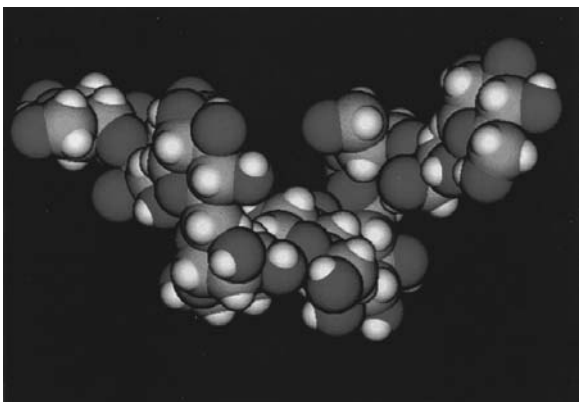
Numerous applications of pullulan are claimed in a large number of patents. A comprehensive review of the applications of pullulan in food and drugs, as well as on the potential uses of pullulan in industrial and medical applications, is reported elsewhere (6).



(a)



(b)



(c)

Figure 1. Generally accepted chemical structure of pullulan: (**a**) chemical formula; (**b**) ball-and-stick model for one hypothetical conformation of three maltotriose units; (**c**) space-filled model for one hypothetical conformation of three maltotriose units. *Source*: The ball-and-stick and space-filled models were kindly provided by Professor Josée Brisson from the Chemistry Department, Laval University.

PROCESS TECHNOLOGY FOR THE PRODUCTION OF PULLULAN

The general process flowchart in Figure 2 illustrates the five main steps in the production of pullulan by fermentation: culture medium preparation, inoculum preparation, fermentation, product separation and purification, and product alteration. This five-step general process flowchart for the production of pullulan by fermentation applies for small-scale laboratory experiments as well as large-scale commercial production of pullulan.

Step A. During the culture medium preparation step, appropriate quantities of raw materials to be used as culture medium ingredients are to be formulated, weighed, and dissolved in water. The ingredients are chosen and formulated in such a way to provide the exact amounts of carbon, nitrogen, and mineral salts required for cell growth and pullulan synthesis during the fermentation step. The pH of the solution is adjusted to the desired value; the medium is then sterilized, either in a fermenter or in a separate feed tank.

Step B. The inoculum preparation step consists of producing a large quantity of microorganism to seed or to inoculate the sterilized culture medium obtained at the end of step A in order to start the fermentation (step C). The appropriate strain of microorganism is selected, then propagated by successive fermentations from a small volume of a pure culture to obtain a large volume of microbial cell suspension required for inoculum purposes. In the industrial production of pullulan, the amount of inoculum required to seed the production fermenter may vary from hundreds to thousands of liters of microbial cell suspension. The inoculum culture-to-medium ratio used may also vary from 0.1% v/v to 10% v/v. Usually, larger inoculum

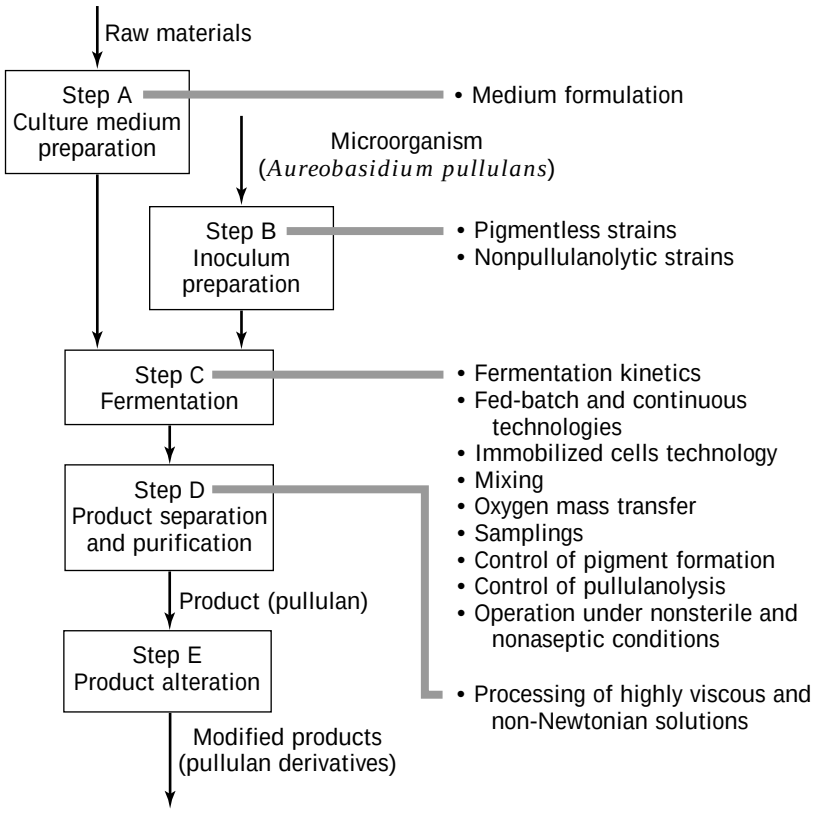


Figure 2. Five-step general process flowchart for the production of pullulan by fermentation. Research and development requirements for technology improvement are indicated for each step.

ratios are suitable for laboratory-scale fermentations, and smaller ratios are more practical for industrial-scale productions.

Step C. The fermentation step is the biosynthesis of pullulan from a carbon source by living and growing microbial cells in the fermenter under sterile and aseptic conditions. The fermentation technology for pullulan production may be batch mode, fed-batch mode, or continuous mode, using either free or immobilized cells. Batch fermentation using free cells of *A. pullulans* is still the most popular technology used in pullulan production in small as well as in large scales. The duration and the performance of the fermentation step are determined by the kinetic behavior of the combined system of microorganism–medium–fermentation conditions. It means that each specific strain of microorganism, in a specific culture medium (ingredients and composition), and under a specific set of fermentation conditions will respond with a specific fermentation kinetic behavior. Therefore, knowledge of fermentation kinetics is important for choosing the best fermentation technology and conditions for each specific strain of microorganism in a specific culture medium. During the fermentation step, operating conditions such as pH, temperature, oxygenation (by aeration), and mixing (by agitation) are either artificially controlled by maintaining each one at a certain value for a specific time period, or simply uncontrolled.

Step D. When the fermentation step is completed, the resulting fermented broth is a mixture of solids in suspension (which consists of microbial cells and cellular debries) together with dissolved solids in the aqueous solution (which comprise water-soluble pullulan synthesized during the fermentation, residual ingredients from the culture medium, and extracellular metabolites produced and excreted during the fermentation). In the product separation and purification step, solids in suspension are first separated from the culture medium by centrifugation or filtration. The obtained liquid phase containing soluble pullulan is then decolorized using activated carbon to remove the black pigmentation that is produced by *A. pullulans* during fermentation. The decolorized liquid phase is further purified to remove small molecular compounds (residual ingredients, extracellular metabolites), then dewatered and dried to obtain pullulan as a final product in solid form. The dewatering operation usually consists of the precipitation of pullulan from the aqueous solution using acetone, methanol, or ethanol as a solvent (Fig. 3).

Step E. The product alteration step is necessary only when modified pullulan or pullulan derivatives are desired as final products. Technologies are available to chemically modify pullulan in order to improve its existing properties and/or to impart newly desired properties to pullulan. Examples of such pullulan derivatives are presented in Table 1 (5,7–15).

COMMERCIAL PRODUCTION OF PULLULAN

At the present time, the only known commercial producer of pullulan by the fermentation process is the Hayashibara Co., Ltd. in Japan. Figure 4 illustrates the detailed flowchart for the commercial production of pullulan by this particular company (16). Note that it matches very well the

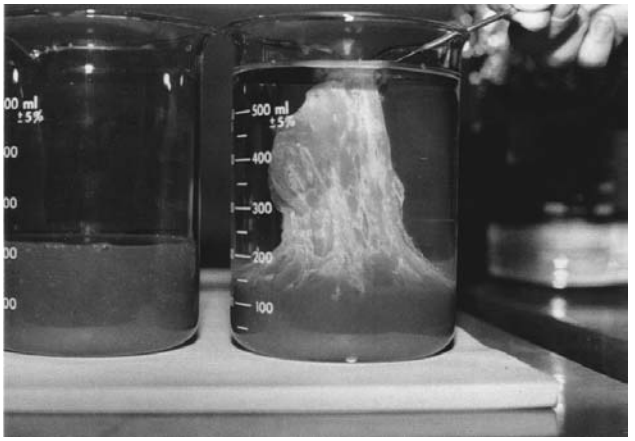


Figure 3. Solvent precipitation of pullulan: Pullulan, which is soluble in the fermentation broth (yellow solution), becomes insoluble in methanol (clear liquid). The fibrous texture appearance of the precipitated pullulan indicates that this polysaccharide has very high molecular weight.

five-step general process flowchart for the production of pullulan by fermentation shown in Figure 2.

The culture medium used consists of a carbon source (partial starch hydrolyzate with dextrose equivalent [DE] around 50% at a concentration from 10 to 20%), a nitrogen source (formulated from corn gluten, soybean protein, peptone, nitrate salts, and ammonium salts), and other mineral salts. The initial pH is within 5–6.8.

The inoculum is prepared from selected mutant strains of *A. pullulans* AHU 9553, *Dematium pullulans* IFO 4464, and *Pullularia fermentans* var. *fermentans* IFO 6401. The mutant strains used as production strains must have the least black pigment, require the shortest duration of fermentation, and give the highest yield of pullulan.

The fermentation is carried on in batch mode for approximately four days in a conventional fermenter, with special consideration to ensure adequate aeration and agitation of the highly viscous fermentation broth. The temperature is maintained at 30 °C and aeration at 0.5 VVM (volume of air per volume of culture medium per minute). The kinetic behavior of commercial pullulan fermentation corresponds very well with those reported in the scientific literature (17–20). The pH of the culture medium decreases sharply during the first day and remains constant at around pH 3.5 to the end of the fermentation. Cell growth and pullulan synthesis occur during this low and constant pH period. The concentrations of cell mass and pullulan continue to increase in the fermentation broth until the end of the fermentation. However, the molecular weight of pullulan in the fermentation broth increases to a maximum value of around 2,000,000–3,000,000 in the early stage of fermentation, then drastically decreases to very low molecular weight toward the end of the fermentation. This decrease is presumably caused by the action of a pullulan-decomposing enzyme, which is also synthesized by the microorganism during the fermentation period. Consequently, the viscosity of the fermentation broth, which results from both pullulan concentration and pullulan molecular weight, increases rapidly in the early stage of fermentation to over 300 cP, then reduces to less than 100 cP at the end of the fermentation. By controlling the fermentation conditions (phosphate content and initial pH), Hayashibara Co., Ltd. obtains a pullulan product that ranges from 50,000 to 500,000 MW.

The pullulan recovery and purification operations depend on the desired quality of pullulan product for specific applications (Fig. 4). Hayashibara Co., Ltd. defines three different pullulan grades: (*1*) Product I (crude pullulan culture and its concentrate) for use as an industrial adhesive, dispersant, and coagulant; (*2*) Product II (cell-free, decolorized pullulan, containing some minerals, mostly sodium chloride) for use as a raw material for various industries, including the food processing industries, as well as a base material for the production of pullulan derivatives; and (*3*) Product III (cell-free, decolorized, and desalted pullulan), which is highly purified pullulan in white powder for making films and shaped articles.

The annual production of pullulan by Hayashibara Co., Ltd. is 200 metric tons and the ex-factory sale price of pullulan PF-20 (Product III) is 2300 JP_1 (approximately 17 US$) per kilogram (personal communication with Dr. K. Okada, Overseas Business Development, Hayashibara Co., Ltd., April 6, 1998).

RESEARCH AND DEVELOPMENT REQUIREMENTS FOR TECHNOLOGY IMPROVEMENT

Pullulan fermentation, like numerous other extracellular polysaccharide fermentations, is one of the most challeng-

Table 1. Pullulan Derivatives

Modified pullulan	Imported properties	References
Esterified pullulan	Various degrees of water insolubility	5
Etherified pullulan	Various degrees of water insolubility	5
Hydrogenated pullulan	High heat resistance	7
Pullulan sulfate and its salts	Anti–peptidic ulcer; low blood anticoagulation	8
Carboxylated pullulan	High solubility in cold water	9
Cross-linked pullulan	Hydrophilic gel; various degrees of water absorbability	10
Dialdehyde pullulan	Water resistance	11
Pullulan aminoalkyl ether	Cationic polymer	12
Ionic pullulan gel	Charged gel	13
Cyanoethylated pullulan	High heat resistance; insolubility in water; high dielectric constant	14
Cyanoethyldihydroxylakylated pullulan	High dielectric constant	15

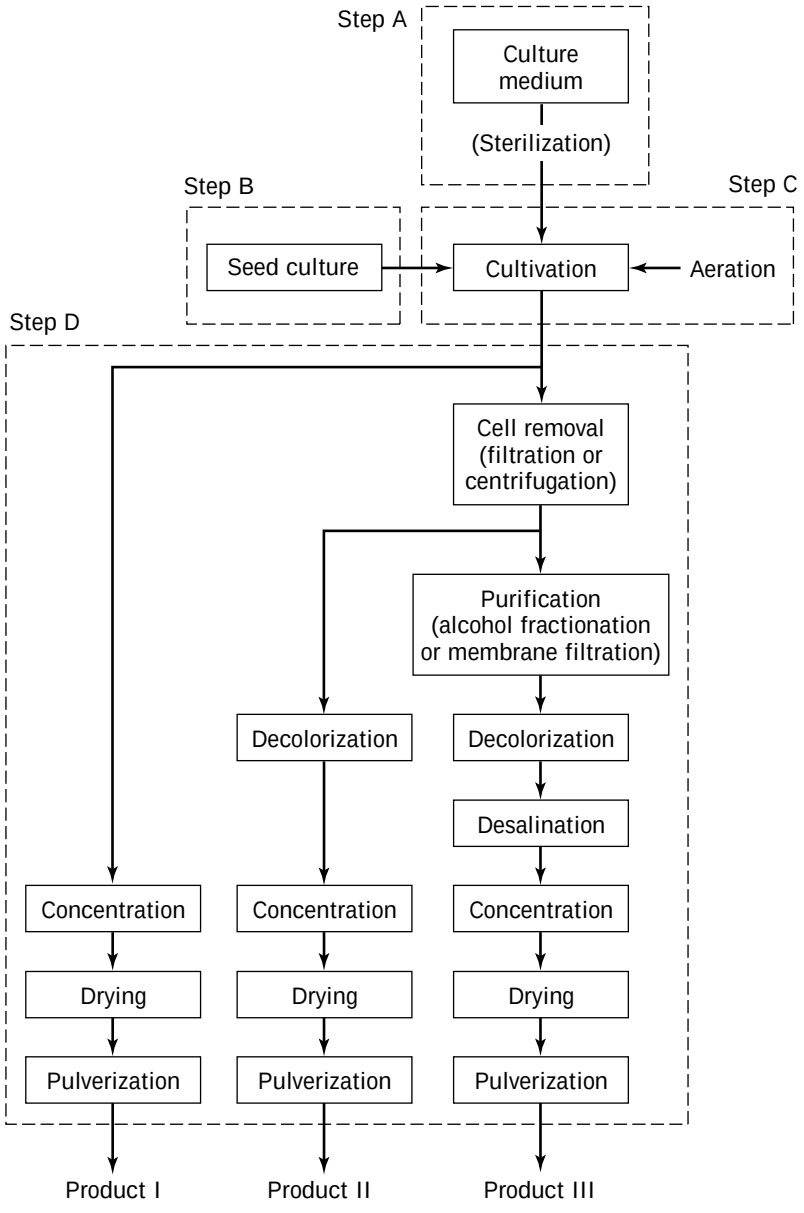


Figure 4. Process flowchart for the commercial production of pullulan by Hayashibara Co., Ltd. *Source*: Adapted from Ref. 16, reproduced with permission.

ing bioprocesses for engineers and scientists seeking ways to improve fermentation technology. Four groups of problems associated with pullulan fermentation still require better solutions in order to increase the performance of the fermentation process and lower product recovery costs.

Problems Due to Viscosity of the Fermentation Broth

The evolution of rheological properties of the fermentation broth during batch fermentation of *A. pullulans* has been studied in detail (17). In batch fermentation, the culture medium initially exhibits Newtonian behavior, and then the rheology of the broth progressively changes to a highly viscous and non-Newtonian behavior. The rheological properties of the fermentation broth may be described by two different parameters: the field apparent viscosity and the non-Newtonian flow behavior index, which can be expressed with the following rheological model:

$$\eta = K\dot{\gamma}^{n-1} \tag{1}$$

where η is the apparent viscosity, K is the consistency index, n is the flow behavior index, and $\dot{\gamma}$ is the shear rate. This equation shows that, in non-Newtonian fluids, the value of viscosity obtained depends on the rate of shear at which the viscosity is measured. The viscosity of a non-Newtonian fermentation broth is commonly standardized by its field value, which is the value measured at the unit rate of shear. Such values of the field apparent viscosity are independent of either rotational speed or geometrical configuration of the viscometer used.

It would be tempting to believe that the viscosity of the fermentation broth increases steadily throughout the fermentation to follow the increase in pullulan concentration. In fact, in the case of pullulan, the field apparent viscosity of the fermentation broth increases slowly during the ini-

tial phase of fermentation, then rises rapidly to its maximum value, and finally decreases toward its initial value. The maximum values of the apparent field viscosity, as well as the time when these maximum values occur, depend on fermentation conditions such as initial pH of the culture medium and age of inoculum (Fig. 5). The results of Figure 5 are very informative in the sense that the large differences observed for the maximum values of the field apparent viscosity can be directly attributed to the length of the pullulan chain or to the average molecular weight of pullulan, because the concentrations of pullulan in the fermentation broths are approximately the same around the fifth day of fermentation. The decrease in the field apparent viscosity during later stages of each fermentation, when the polysaccharide concentration has leveled off, is called pullulanolysis, which relates to the depolymerase secreted by *A. pullulans*. The evolution of the field apparent viscosity of the fermentation broth depends also on other factors such as the aeration rate (VVM) and the level of dissolved oxygen concentration (mg/L) during pullulan fermentation (21).

The flow behavior index characterizes very well the average length of the pullulan chain (22). It is therefore interesting to monitor this rheological parameter during the course of the fermentation to observe changes in pullulan average molecular weight. At the beginning of a typical pullulan fermentation, the culture medium is Newtonian (non-Newtonian index $n = 1$). The non-Newtonian index of the fermentation broth decreases slowly during the initial phase of fermentation, then drops rapidly to its minimum value, and finally rises toward its initial value. Therefore, throughout the fermentation the broth becomes more and more pseudoplastic (non-Newtonian index <1), then returns toward its initial Newtonian behavior. The minimum values of the non-Newtonian index as well as the time when these minimum values occur depend on fermentation conditions such as initial pH of the culture medium and the age of inoculum (Fig. 6). The rheological behavior of such non-Newtonian solutions has a profound impact on the flow characteristics within the fermenter. The apparent viscosity decreases when the fluid is in motion. This reduction in apparent viscosity with the velocity of the fluid is more pronounced as the flow behavior index is reduced.

The observed rheological properties of the fermentation broth in Figures 5 and 6 result from the combined contribution of the pullulan concentration and the average length of the pullulan chain. Both pullulan concentration and pullulan average chain length in the fermentation broth are continuously changing during batch fermentation. Therefore, from the results in Figures 5 and 6, it is not possible to say how and how much these two parameters contribute to the observed rheological properties. LeDuy et al. (17) have recovered pullulan from fermentation broths at 1, 3, 5, and 8 days of a batch fermentation

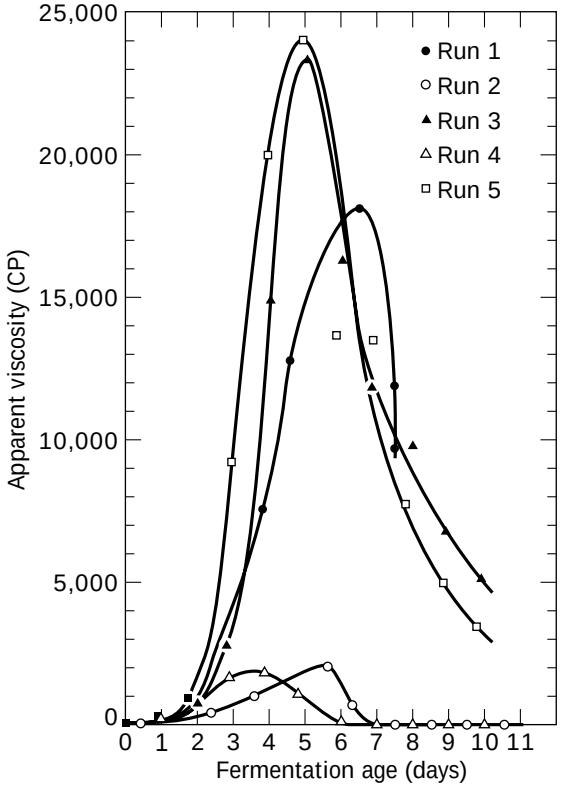


Figure 5. Evolution of the field apparent viscosity of the fermentation broths under various conditions of initial pH and inoculum age. *Source*: From Ref. 17, reproduced with permission.

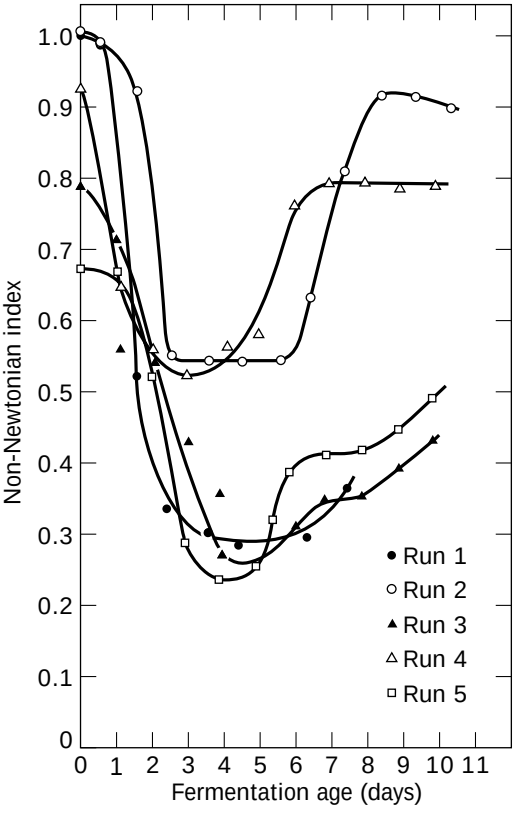


Figure 6. Evolution of the non-Newtonian index of the fermentation broths under various conditions of initial pH and inoculum age. *Source*: From Ref. 17, reproduced with permission.

(run 3 from Figs. 5 and 6), freeze-dried it, and then dissolved the recovered pullulan in water to form solutions having a pullulan concentration varying from 0.125 to 2.0 g/L. The field apparent viscosities and the non-Newtonian indices of these solutions were measured and are presented in Figure 7. At the same concentration of pullulan, the field apparent viscosity of pullulan solution recovered from the early stage of fermentation is higher than that recovered from the later stage, indicating that the pullulan recovered in the first days of fermentation has higher molecular weight than in subsequent days. The opposite phenomenon is observed for the non-Newtonian index. At the same concentration of pullulan, the non-Newtonian index of pullulan solution recovered from the early stage of fermentation is lower than that recovered from the later stage, indicating that the pseudoplasticity of the pullulan solution is directly related to the high molecular weight of pullulan. Each set of curves prepared with pullulan recovered after the same fermentation time in Figure 7 shows the very important contribution of pullulan concentration on the field apparent viscosity as well as on the non-Newtonian index of the pullulan solution. Higher pullulan concentration leads to higher field apparent viscosity and larger pseudoplasticity.

Thus, any fermentation process aiming to achieve high-yield production of high molecular weight pullulan still provides tremendous engineering challenges in mixing, oxygenation, and sampling during the fermentation, as well as the major challenge of removal of microbial cells from the highly viscous fermentation broth in the product separation and purification step. To clearly illustrate the difficulties involved in dealing with a highly viscous fermentation broth, Figure 8 shows a typical fermentation broth flowing out of a fully opened large-size sampling tube under the compression of the aeration gas inside the fermenter. During pullulan fermentation, the nonhomogeneous mixing problem is always observed, even in a laboratory-scale fermenter. The fermentation broth inside the fermenter tank segregates into two regions having different characteristics, with an active core around the impeller where the culture medium is well mixed and properly aerated. This core is separated from the innermost zone by a fictitious boundary, a stagnant fluid region characterized by a total lack of agitation and aeration (Fig. 9). Some authors even discuss the formation of three distinct mixing regions: a small central region (often referred to as a cavern) in the vicinity of the impellers where the turbulence is important, an annular section characterized by slow fluid motion, and finally, a stagnant outer region where the

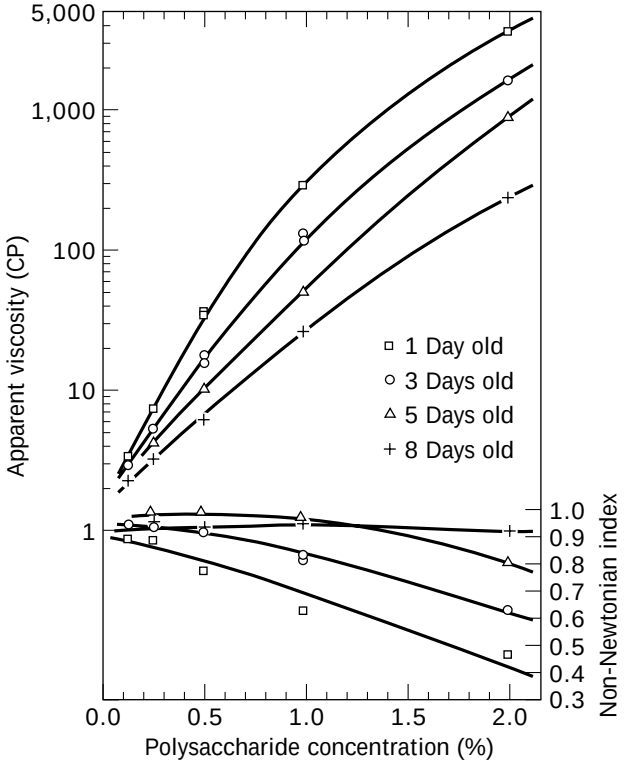


Figure 7. Variations of field apparent viscosity and non-Newtonian index as a function of the concentration of pullulan in a solution and the pullulan chain length. Pullulan molecules have longest chain lengths after one day and shortest chain lengths after eight days of batch fermentation. *Source*: From Ref. 17, reproduced with permission.

Figure 8. Highly viscous fermentation broth of *A. pullulans* in a sucrose medium: the broth slowly flowing out of a fully opened large-size sampling tube under the compressed air inside the fermenter.

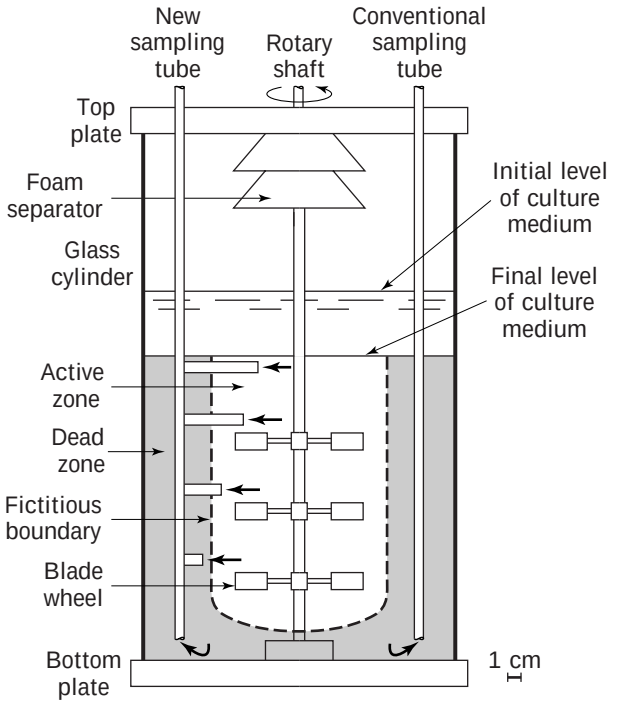


Figure 9. Active zone and stagnant zone of the highly viscous fermentation broth in the fermenter during the pullulan fermentation. Configurations of both conventional and newly improved sampling devices are shown. *Source*: From Ref. 23, reproduced with permission.

fluid is motionless and the apparent viscosity very high (24). The energy imparted to the fluid by the rotating impellers is dissipated in a restricted region of the total volume of the fermentation broth. The high viscosity combined with the non-Newtonian behavior creates major difficulties in sampling, oxygen mass transfer, mixing, and downstream processing. These are briefly analyzed in turn.

Process Monitoring and Sampling. With the absence of homogeneous conditions throughout the fermentation vessel, it is only possible to monitor locally the dissolved oxygen concentration and the pH. This also applies to any novel sensor for monitoring other fermentation process variables such as the *A. pullulans* population (25). These values will therefore depend greatly on the position of the various probes. The same phenomenon applies also to the fermentation broth sampling port. The geometry and the location of the conventional sampling tube, as provided in fermenters by manufacturers (Fig. 9), shows that the samples are taken very frequently from the stagnant zone, where the data are not representative for the entire culture medium content. A newly improved sampling device was designed and tested in a 7-L fermenter (23) and enables one to obtain a more representative sample because it draws a sample from different locations within the culture vessel (Fig. 9). Nevertheless, because of the evolving and shear-dependent viscosity, it is not guaranteed that a uniform sampling in all branches of the sampling tube will be achieved. Thus, the sampling of representative samples and process monitoring of highly viscous fermentation broth in larger-scale fermenters having nonhomogeneous mixing remains a challenge that one should not overlook.

Oxygen Transfer. Pullulan fermentation is an aerobic fermentation. It was demonstrated that oxygen is essential for the biosynthesis of pullulan by *A. pullulans*. In a growth medium, pullulan yield and synthesis rates are proportional to the oxygen availability. However, under a controlled oxygen environment in a nongrowth (nitrogen-free) medium, pullulan yield and synthesis rates are inversely proportional to the oxygen concentration in the culture medium (26). During the fermentation, air is continuously provided to the culture medium to supply oxygen to the microbial cells. However, oxygen molecules in the air bubbles have to move into the fermentation broth and then diffuse through this liquid phase to reach the microbial cells in order to be used for the growth and biosynthesis of pullulan. With the highly viscous fermentation broth in pullulan fermentation, the oxygen mass transfer process is very difficult and remains a major engineering challenge for the future. Indeed, as the fermentation broth becomes increasingly viscous, the gas bubbles in the peripheral stagnant zone of the fermenter remain trapped in the motionless fluid and are progressively depleted from their oxygen content. Figure 10 shows a photograph taken during a typical pullulan production where the fermentation broth and gas bubbles in the outer region remained stagnant for long periods.

The evaluation and comparison of the performance of oxygen mass transfer within a mixing vessel is usually quantified by the volumetric mass transfer coefficient $K_L a$. The overall oxygen transfer rate from air bubbles to the fermentation broth (dC_L/dt) is the product of the volumetric oxygen transfer coefficient ($K_L a$) and the mean concen-

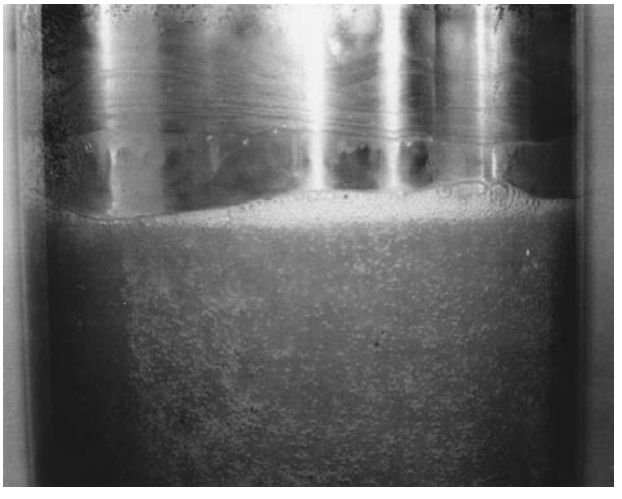


Figure 10. Mixing and mass transfer challenges in pullulan fermentation. The picture is taken when the agitator equipped with Ruston turbines is rotating at 500 rpm. The layer of fermentation broth adjacent to the fermenter wall remains stagnant with trapped air bubbles.

tration driving force of oxygen in the fermentation broth ($C^* - C_L$):

$$\frac{dC_L}{dt} = K_L a(C^* - C_L) \quad (2)$$

where C_L is the dissolved oxygen concentration in the fermentation broth and C^* is the saturated dissolved oxygen concentration in the fermentation broth that is the solubility of oxygen under the fermentation conditions. According to equation 2, one possible way to improve the oxygen mass transfer rate is to improve the mean concentration driving force of oxygen in the fermentation broth ($C^* - C_L$) by increasing the solubility of oxygen in the liquid phase. The solubility of oxygen in the fermentation broth, C^*, during the fermentation can be enhanced either by using oxygen enriched gas, up to pure oxygen, in aeration or by pressurizing the fermenter headspace (commonly referred to as hyperbaric fermentation). For practical and economic reasons, the use of pure oxygen is not feasible in large-scale fermenters. On the other hand, enriched gas would be used only when air would not suffice to maintain the desired dissolved oxygen concentration within the fermenter, because high oxygen concentration in the fermentation broth may enhance the biosynthesis of the black melanin pigment, which is undesirable (26).

The effects of pressure on the growth of *A. pullulans* and the synthesis of pullulan have been studied in a specially designed fermenter and fermentation system (Figs. 11 and 12) (27). It is reported that for all volumetric airflow rates in the range of 0.2 to 2.0 VVM, the production of biomass increases with pressure up to a critical value, which ranges from 0.50 MPa to 0.75 MPa, at which point a drastic decrease in biomass production and a change in cellular morphology occur. Above the critical pressure, *A. pullulans* forms aggregates of irregular shape that are fairly uniform in size under each set of fermentation conditions. In general, larger aggregates of approximately 8 mm are associated with a higher pressure level (1.48 MPa) and aeration rate (2 VVM). The details of *A. pullulans* aggregates are shown in the scanning photomicrographs in Figure 13. The same pattern is also observed for pullulan synthesis: it increases with pressure up to approximately the same critical pressure, then decreases drastically beyond this value. Much more investigation is required to better understand the behavior of hyperbaric pullulan fermentation before reaching a practical and economic solution for this technology.

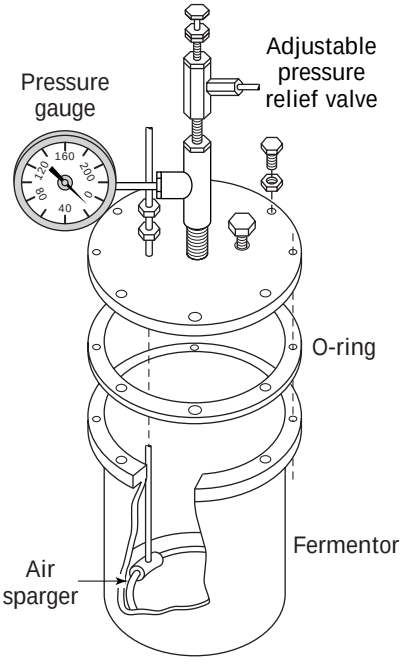


Figure 11. The specially designed fermenter for studying pullulan fermentation under hyperbaric conditions. *Source*: From Ref. 27, reproduced with permission.

A second way to improve the oxygen mass transfer rate is to increase the volumetric oxygen transfer coefficient ($K_L a$) by using better mixing devices. It has been clearly shown that $K_L a$ is not uniform throughout most mixing vessels. It is usually large in the vicinity of the mixing element and lower further away. It is evident from Figures 9 and 10 that any mixing device that can reduce or even eliminate the dead zones in the fermenter will have a chance to provide better volumetric oxygen transfer to the pullulan fermentation broth and in addition create more homogeneous operating conditions throughout the fermentation vessel. An adequate uniform mixing would also contribute to the maintenance of gas bubbles with high oxygen content and therefore keep the oxygen gradient at a high level. Several unconventional mixing devices have been designed and studied. These are the reciprocating plate mixer, the helical ribbon mixer, and the helical ribbon screw mixer.

Although the physical properties of the fluid and operating conditions, used by many authors, are hardly comparable over a wide range, their results have a common trend of increasing oxygen mass transfer coefficients with increasing agitation and gasflow rate. The efficiency of mixing devices, evaluated in terms of $K_L a$, is usually correlated in terms of power input per unit volume (P_G/V_L) and superficial gas velocity (U_G), using the following expression:

$$K_L a = C\left(\frac{P_G}{V_L}\right)^{\alpha} (U_G)^{\beta} \quad (3)$$

The correlation coefficients C, α, and β depend on the nature of the fluid, the geometrical configuration of the mixing element, and the hydrodynamic regimes of the fluid (28,29).

The biochemical engineering group at Laval University has been interested for the last decade in developing a relatively new type of bioreactor, the reciprocating plate bioreactor (RPB). The RPB (Fig. 14a) consists of a stack of perforated plates that undergoes reciprocating vertical motions inside the fermenter. The RPB, with its particular geometry, possesses the desired attributes to produce efficient mixing, with the result that microorganisms, nutrients, and air bubbles are properly distributed (31). In the RPB, the perforated plates occupy the total cross section

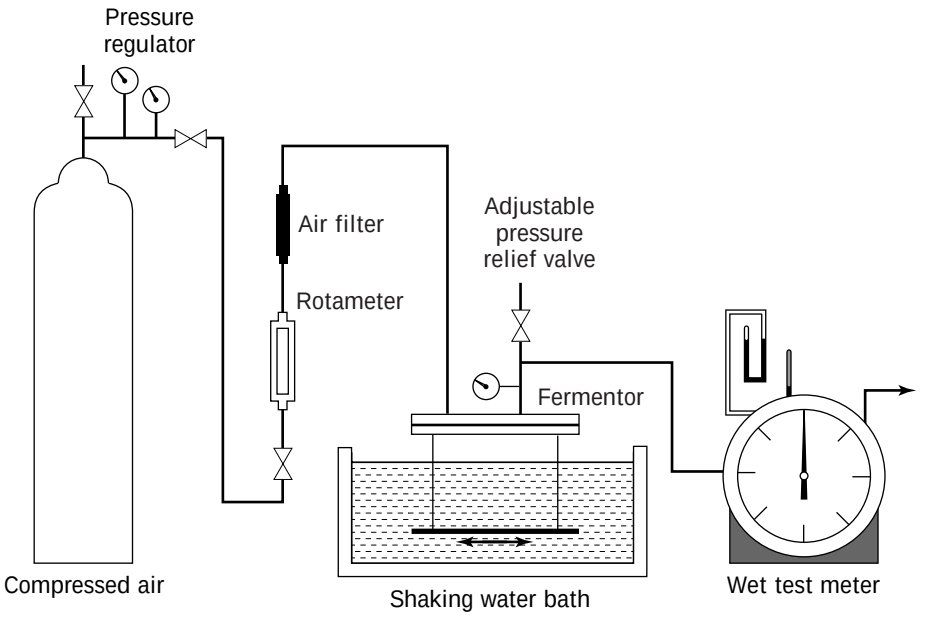


Figure 12. The fermentation setup for studying pullulan fermentation under hyperbaric conditions. *Source*: From Ref. 27, reproduced with permission.

of the fermentation vessel so that the upward and downward movements of the plate stack lead to uniform mixing. Three different RPBs with aspect ratios ranging from 0.08 to 0.5 have been characterized with fluid models and pullulan fermentation (21,28,30). Audet et al. (29) have performed experiments to evaluate the influence of polysaccharide concentration and molecular weight on oxygen mass transfer in the RPB. They used model fluids consisting of different concentrations of dextran (another water-soluble microbial polysaccharide) of different molecular weights to simulate the pullulan fermentation broth at different stages of fermentation. They correlated their results using equation 3 and reported that the correlation coefficients C, α, and β in this equation also depend on the polysaccharide concentration, the polysaccharide molecular weight, and the hydrodynamic regime of the solution.

The most recently constructed RPB was designed to be identical, except for the mixing mechanism, to a stirred-tank (radial-type mixing device) bioreactor that was built simultaneously. The schematic diagram of the experimental setup of the RPB is shown in Figure 15. Each bioreactor has a total volume of 22.5 L and a working volume of 18 L. The two bioreactors, using a series of mixing elements, were initially characterized in water for their ability to produce high K_La values. A total of five plate stacks were used in the RPB. Each plate stack consists of 6 (plate stacks PA, PB, and PC) or 12 (plate stacks PD and PE) perforated stainless steel plates, 221 mm in diameter and 1.25 mm thick, mounted on a central shaft and uniformly distributed with a distance of, respectively, 50 and 25 mm between adjacent plates. The diameters of perforations were 6.35 mm for plate stacks PA and PD, 12.7 mm for plate stacks PB and PE, and 19.05 mm for plate stack PE. A total of five mixing elements were used with the radial-type bioreactor: assemblies containing one, two, or three identical Rushton turbines (RT1, RT2, and RT3, Fig. 14b), and a helical ribbon impeller with or without surface baffles (HR and HRB, Fig. 14c). A more complete description is presented by Gagnon et al. (32).

These bioreactors were first characterized in terms of oxygen mass transfer (32). Typical plots of the oxygen mass transfer coefficient in water as a function of the power input per unit volume are presented in Figure 16 for all mixing devices at one superficial gas velocity. All mixing devices show identical patterns. For lower power inputs per unit volume, the oxygen mass transfer, K_La, is constant and varies only with the gas superficial velocity. At a certain power input per unit volume, K_La starts to increase rapidly with the power input per unit volume. Perez and Sandall (33) and Nishikawa et al. (34) previously observed these two distinct regimes of operation with Rushton turbines.

The two bioreactors were then used in an investigation undertaken with the objective of comparing experimentally the performance of four different mixing devices for the production of pullulan (30). The four mixing devices include two different reciprocating plate stacks (PA and PB), the assembly of three Rushton turbines (RTB), and the helical ribbon impeller (HR). These four fermentations were performed with an equal level of power input per unit volume (1,000 W/m^3) and gasflow rate (0.5 VVM or 9 L/min) during the entire fermentation. The evolution of the concentrations of biomass, pullulan, and dissolved oxygen during each fermentation are presented in Figures 17–19. Under identical power inputs per unit volume, better cell growth and pullulan synthesis were obtained with the reciprocating plate mixer that had larger perforations. The helical ribbon mixer gave the poorest cell growth and pullulan synthesis. However, pullulan of the highest molecular weight, as indicated by its highest fermentation broth viscosity, was obtained with the helical ribbon mixer.

The evolution of dissolved oxygen (Fig. 19) clearly shows that the RPBs are more efficient at maintaining a high and homogeneous level of dissolved oxygen within the fermen-

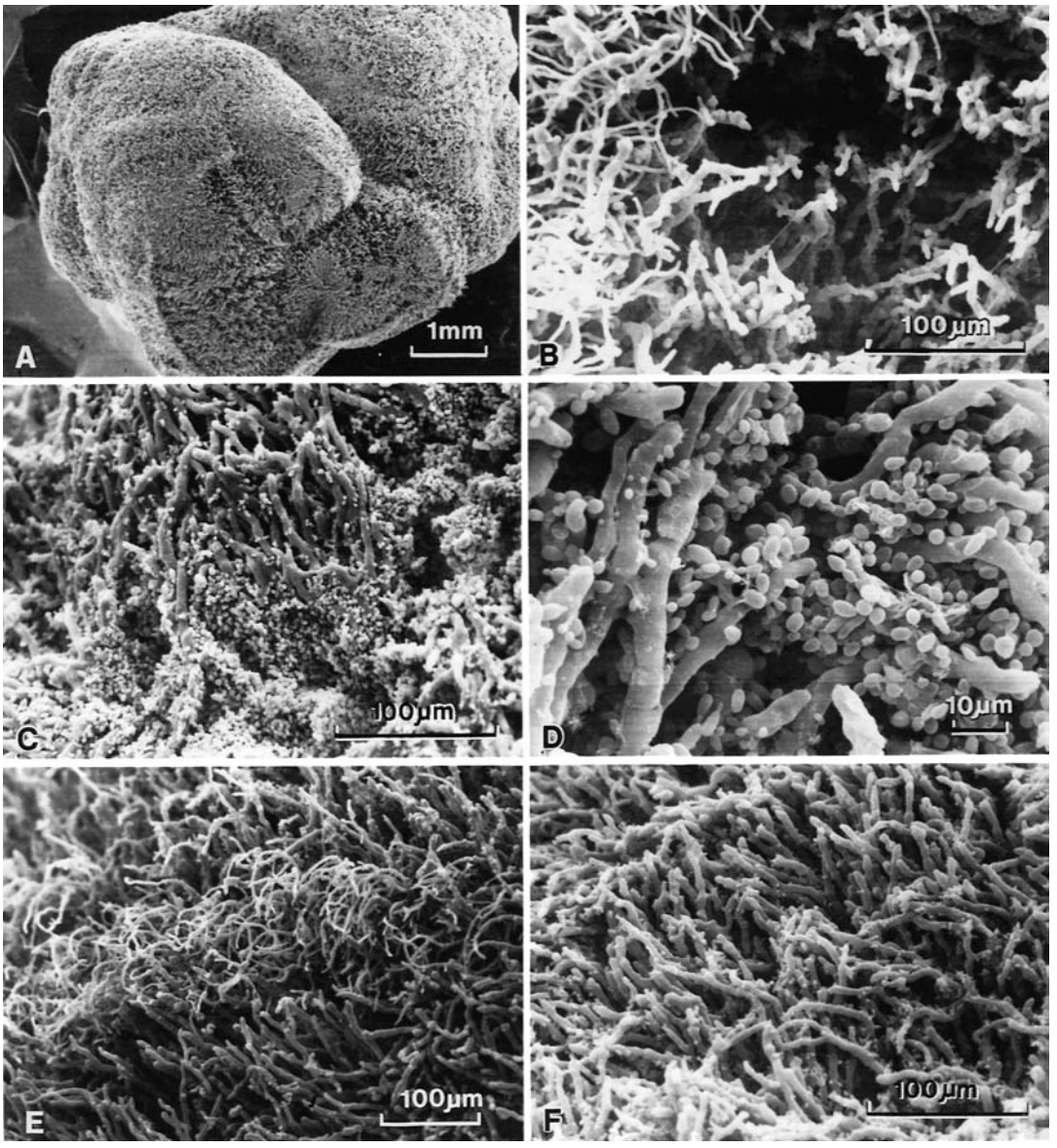


Figure 13. Agglomerates of *A. pullulans* in pullulan fermentation under 1.13 MPa and 0.2 VVM: (**a**) Global view of agglomerate; (**b**) cavity region; (**c**) and (**d**) regions where yeast cells dominate; (**e**) and (**f**) regions where mycelial cells dominate. *Source*: From Ref. 27, reproduced with permission.

tation broth. For the other two mixing devices, it can be stated that the local dissolved oxygen varied greatly and was always very low in the vicinity of the dissolved oxygen probe.

Guerinik (35) has modified the helical ribbon impeller by adding a central screw installed on the rotating shaft to provide the pumping action on the fermentation broth in order to enhance its circulation inside the fermenter. He also found that pullulan fermentations using the helical ribbon screw mixer resulted in a higher concentration of pullulan of higher molecular weight as compared with that obtained using standard Ruston turbines.

In a recent study, Tanguy et al. (36) proposed a dual coaxial impeller composed of a Rushton turbine and a helical ribbon impeller, both rotating clockwise but at different speeds. They have shown that the hybrid mixer system outperforms the standard helical ribbon impeller when the fluid rheology changes with time. These hybrid systems could be used advantageously in pullulan fermentation.

Product Separation and Recovery. The first operation in the product separation and purification step (Step D in Figures 2 and 4) is the removal of *A. pullulans* cells from the highly viscous fermentation broth. This is done by either centrifugation or filtration. The fermentation broth has to be diluted many fold to lower its viscosity prior to the centrifugation or filtration operation. The resulting large volume of diluted processing broth leads to higher operating

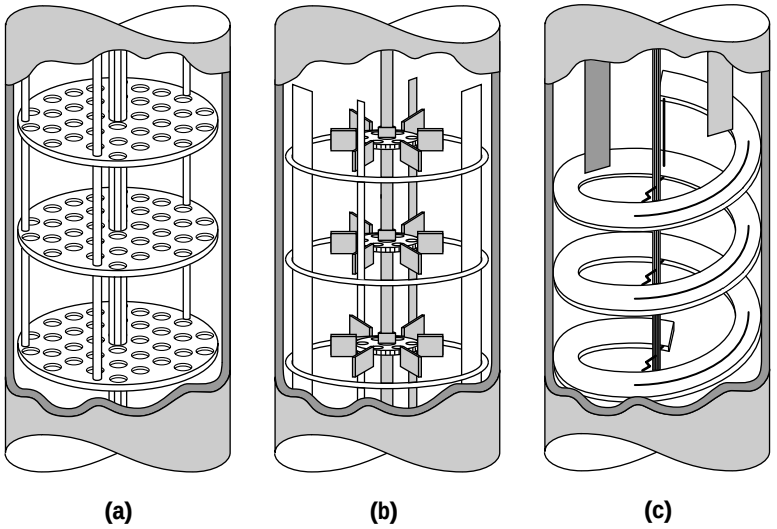


Figure 14. Mixers used in pullulan fermentations: (**a**) the reciprocating plate mixer; (**b**) the Rushton turbines; (**c**) the helical ribbon mixer. *Source*: From Ref. 30, reproduced with permission.

costs in the subsequent steps of pullulan purification, especially in the solvent precipitation step. In order to eliminate the presence of cells in the highly viscous fermentation broth, attempts to immobilize *A. pullulans* cells on solid supports and to use immobilized cells in pullulan fermentation has been reported. The utilization of immobilized cells for this purpose is not feasible because considerable numbers of cells leak into the fermentation broth (37). In addition, because pullulan synthesis is partially growth associated, it is inevitable that some cells will be present in the fermentation broth.

Problems Due to the Sterile and Aseptic Conditions Requirement

At the present time, pullulan production by fermentation is carried under sterile and aseptic conditions, which require very high capital and operating costs as in the case of pharmaceutical production. If pullulan is to be produced for use as bulk chemicals and/or bulk commodities, its production costs have to be greatly lowered. The technological challenge is to develop a nonsterile and nonaseptic fermentation process for the production of pullulan. A bistaged pH process has been developed for this purpose (18). In batch fermentation, the pH of the culture medium decreases rapidly from its initial value of 5.0 to the final value of 2.5 within the first 12–36 h and then self-stabilizes at this value until the end of the fermentation. It is observed that most of the polysaccharide from the batch fermentation process was synthesized during the self-stabilized acidic phases. Experiments on the effect of the initial pH on the fermentation reveal that at very low initial pH values, such as pH 2, the polysaccharide production is insignificant. However, the biomass concentration obtained under these conditions is very high. This interesting phenomenon served as the basic principle for the development of the bistaged pH fermentation process. In this process, the first stage of fermentation is conducted at a very acidic pH for the best production of biomass. When the biomass concentration reaches its maximum value, the second stage of fermentation is initiated by adjusting the medium pH to a higher value for promoting the synthesis of pullulan. The bistaged pH process enhances the pullulan concentration in the medium, influences the rheological properties of the fermentation broth, and has a potential of operation under nonsterile and nonaseptic conditions. During the first stage of fermentation, the microbial contamination could be eliminated due to the very low pH of the medium. During the second stage of fermentation, the contamination could be minimized to an insignificant level due to the high concentration of *A. pullulans* already present in the culture medium. This bistaged pH process could be applied to batch or continuous fermentation processes.

Problems Due to Melanine Pigmentation

Aureobasidium pullulans is also known as a black yeast because of its characteristic black pigmentation associated with mature cells (Fig. 20). The pigment, melanin, is synthesized from the pentaketide pathway in which 1,8-dihydroxynaphthalene is the direct precursor (38). Both intracellular and extracellular melanin are synthesized during the later stage of fermentation where the cell morphology changes from swollen cells to true chlamydospores (39). Thus, melanin production occurs simultaneously with pullulan synthesis, resulting in a dark green to black fermentation broth at the end of the fermentation. This is an undesirable feature for industrial production of pullulan because decolorization (melanin removal) is required during the product separation and purification operations (Step D of Figures 2 and 4). The synthesis of melanin by *A. pullulans* depends on the culture conditions, such as carbon source and concentration, nitrogen source and concentration, oxygen concentration, and trace minerals (26,40,41) (Fig. 20). Therefore, such conditions can serve as control variables to minimize melanin synthesis during the pullulan fermentation process. However, the most promising solution to the problem of melanin production during fermentation may be the use of nonpigmented strains of *A. pullulans*, which may be obtained by mutation (42).

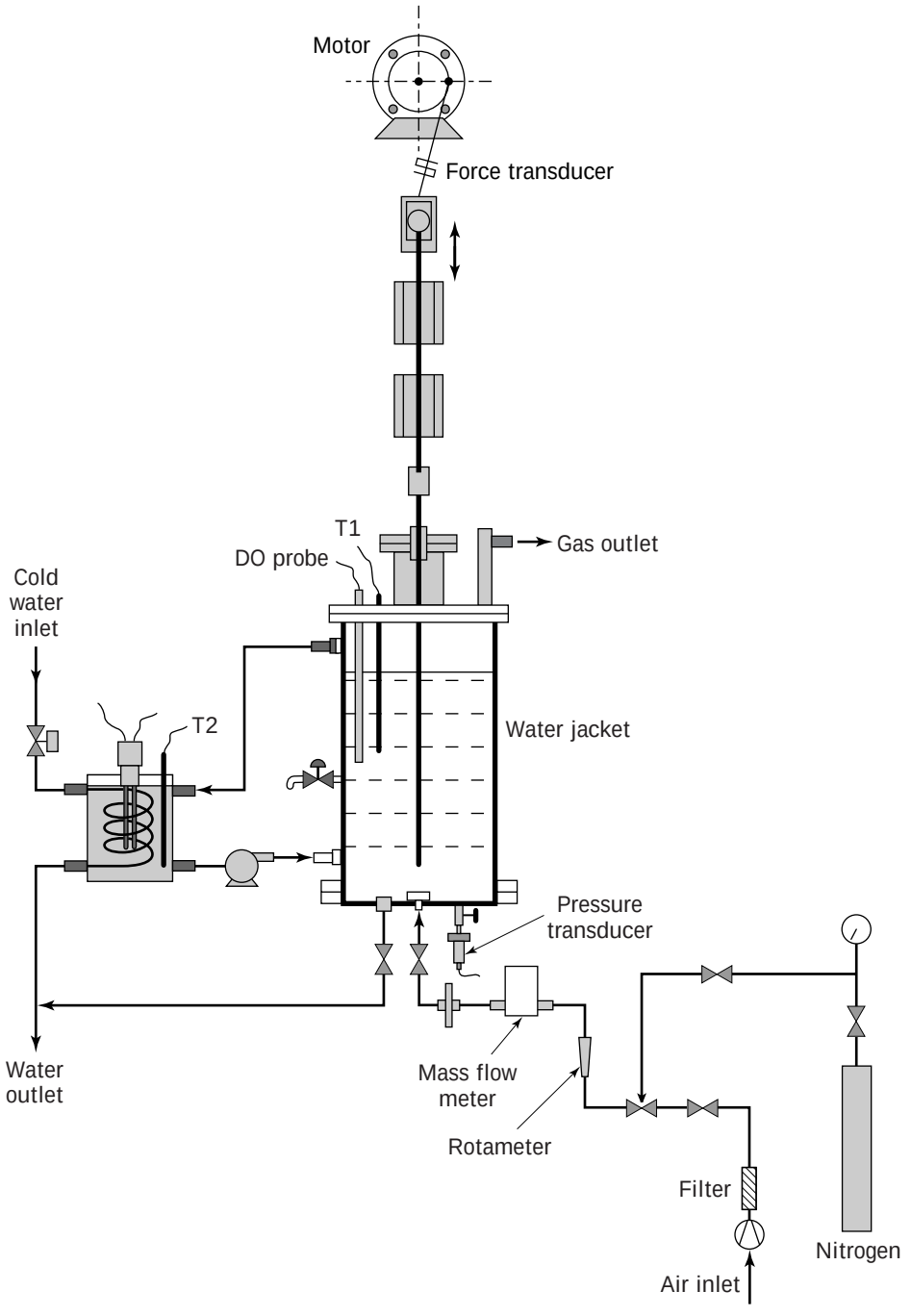


Figure 15. Schematic diagram of the experimental system showing the reciprocating plate bioreactor. *Source*: From Ref. 32, reproduced with permission.

Problems Due to Pullulanolysis during Fermentation

Pullulanolysis is the breakdown of high molecular weight pullulan in the fermentation broth into lower molecular weight pullulan by the depolymerase activity of *A. pullulans*. The pullulanolysis always occurs during the progress of fermentation and is responsible for the decrease of viscosity of the fermentation broth during the later stage of fermentation. The presence of an extracellular endoamylase is suggested as the cause of the observed phenomenon (43). It is not clearly understood why *A. pullulans* produces the pullulan-destructive endoamylase during the synthesis of pullulan. The resulting consequence of the pullulanolysis is that only low molecular weight pullulan is obtained at the end of the fermentation. The pullulanolytic breakdown of the pullulan product is worsen when a high-shearing mixing device such as the Rushton turbine is employed in the fermenter (30,35). Thus, if the high molecular weight pullulan is desired as a final product, the fermentation has to be stopped after approximately 24 h, when the molecular weight of pullulan is at its highest value (44,45). This is not, however, an economically feasible solution for industrial practice because the fermentation

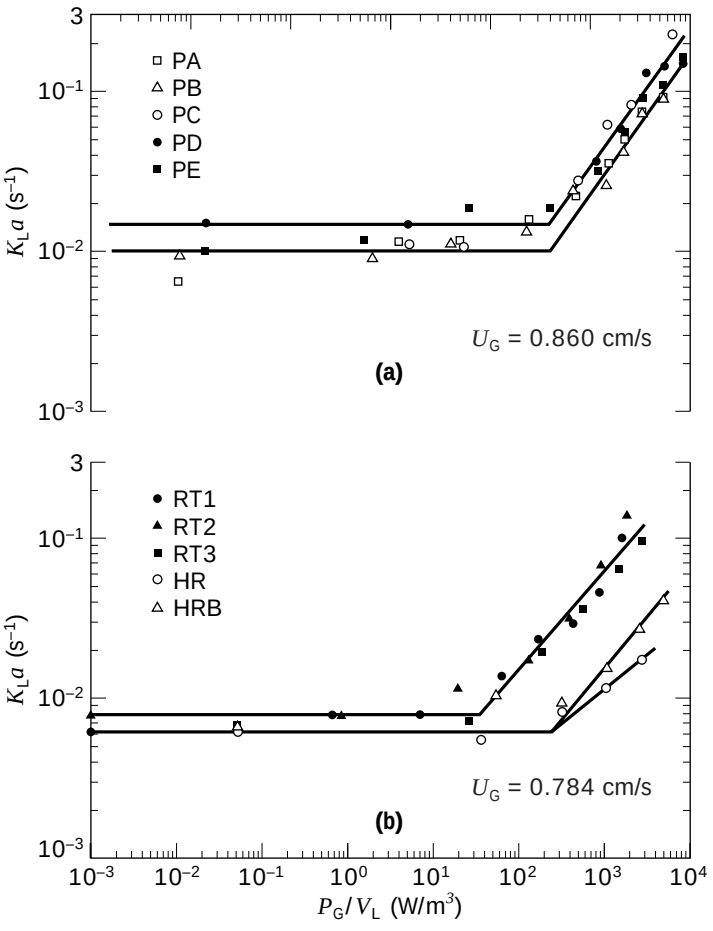


Figure 16. K_La as a function of the power input per unit volume for four mixing devices. *Source*: From Ref. 32, reproduced with permission.

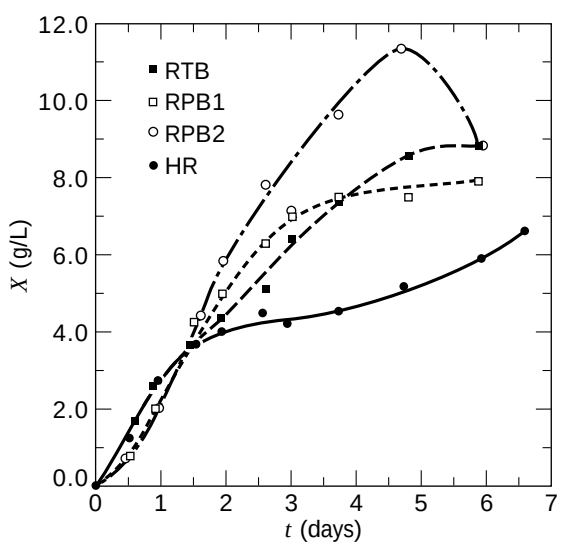


Figure 17. Evolution of biomass dry weight concentration as a function of fermentation time for four mising devices. *Source*: From Ref. 30, reproduced with permission.

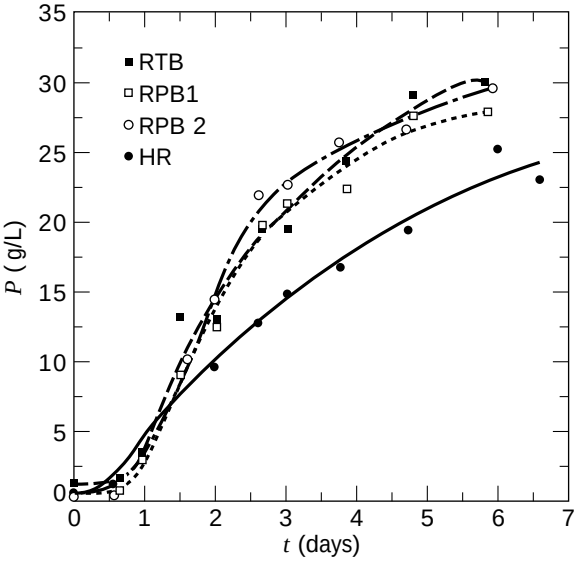


Figure 18. Evolution of pullulan concentration as a function of fermentation time for four mising devices. *Source*: From Ref. 30, reproduced with permission.

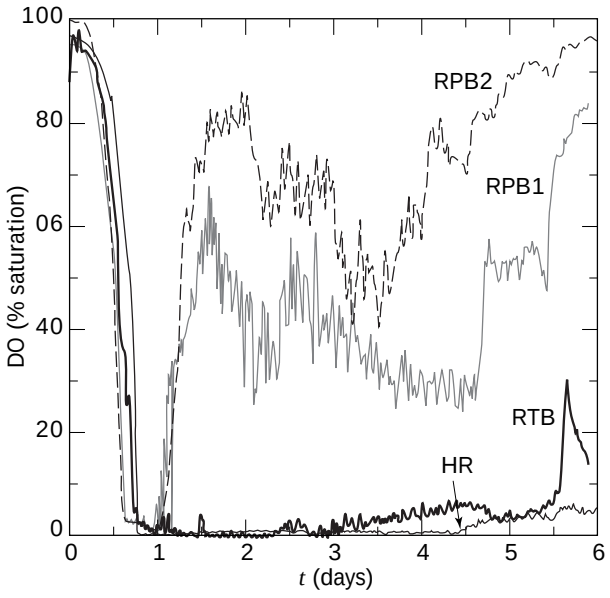


Figure 19. Evolution of dissolved oxygen concentration as a function of fermentation time for four mising devices. *Source*: From Ref. 30, reproduced with permission.

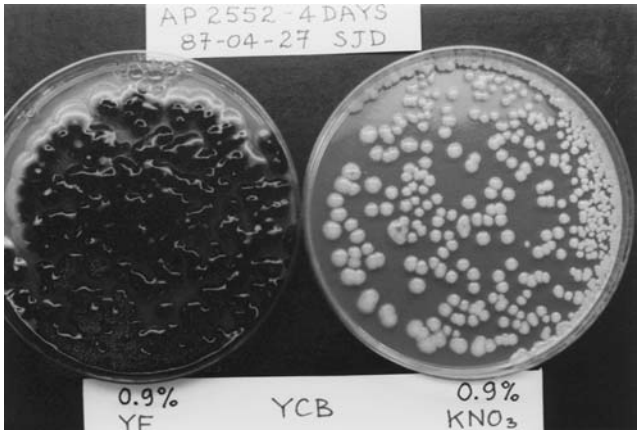


Figure 20. Melanin pigment of *A. pullulans* can be controlled by the nitrogen source. Potassium nitrate as nitrogen source instead of yeast extract in an yeast-carbon-base medium greatly reduces the melanin synthesis (*A. pullulans* 2552 after 4 days of incubation).

efficiency (expressed as grams of pullulan produced per 100 g of supplied substrate) is very low after this short period of fermentation. A more effective but challenging alternative solution is to acquire more knowledge on the pullulanolysis phenomenon by research, then to develop the technology for the suppression or the inhibition of endoamylase synthesis by *A. pullulans* during the pullulan fermentation.

BIBLIOGRAPHY

1. H. Bender, J. Lehmann, and K. Wallenfels, *Biochim. Biophys. Acta* **36**, 309–316 (1959).
2. R. Taguchi, Y. Kikuchi, Y. Sakano, and T. Kobayashi, *Agric. Biol. Chem.* **37**, 1583–1588 (1973).
3. B.J. Catley and W.J. Whelan, *Arch. Biochem. Biophys.* **143**, 138–142 (1971).
4. Ger. Pat. 1,096,850 (January 12, 1961), K. Wallenfels and H. Bender (to Farbwerke Hoechst AG).
5. S. Yuen, *Process Biochem.* **9**(9), 7–9, 22 (1974).
6. A. LeDuy, L. Choplin, J.E. Zajic, and J.H.T. Luong. in H.F. Mark, N.M. Bikales, C.G. Overberger, and G. Menges eds., *Encyclopedia of Polymer Science and Engineering*, 2nd ed., vol. 13, Wiley, New York, 1988, pp. 650–660.
7. U.S. Pat. 3,931,146 (January 6, 1976), K. Kato and M. Shiosaka (to Hayashibara Biochemical Laboratories, Inc.).
8. French Pat. 2,287,911 (May 14, 1976), Anonymous (to Sumitomo Chemical Co., Ltd. and Hayashibara Biochemical Laboratories, Inc.).
9. U.S. Pat. 4,090,016 (May 16, 1978), K. Tsuji, M. Fujimoto, F. Masuko, and T. Nagase (to Sumitomo Chemical Co., Ltd. and Hayashibara Biochemical Laboratories, Inc.).
10. U.S. Pat. 4,152,170 (May 1, 1979), T. Nagase, K. Tsuji, M. Fujimoto, and F. Masuko (to Sumitomo Chemical Co., Ltd. and Hayashibara Biochemical Laboratories, Inc.).
11. U.S. Pat. 4,186,024 (January 29, 1980), M. Fujimoto, K. Fukami, K. Tsuji, and T. Nagase (to Sumitomo Chemical Co., Ltd. and Hayashibara Biochemical Laboratories, Inc.).
12. Can. Pat. 1,078,830 (June 3, 1980), F. Fujita, K. Fukami, and M. Fujimoto (to Sumitomo Chemical Co., Ltd. and Hayashibara Biochemical Laboratories, Inc.).
13. Can. Pat. 1,094,550 (January 27, 1981), F. Fujita, K. Fukami, M. Fujimoto, and T. Nagase (to Sumitomo Chemical Co., Ltd.).
14. French Pat. 2,461,718 (February 6, 1981), H. Muto and H. Suzuki (to Shin-Etsu Chemical Co., Ltd.).
15. European Pat. 76,698 (April 3, 1983), I. Murase, F. Fujita, T. Ohnishi, and T. Tamura (to Sumitomo Chemical Co., Ltd.).
16. K. Sugimoto, *Ferment. Ind.* **36**(2), 98–108 (1978).
17. A. LeDuy, A.A. Marsan, and B. Coupal, *Biotechnol. Bioeng.* **16**, 61–76, (1974).
18. C. Lacroix, A. LeDuy, G. Noël, and L. Choplin, *Biotechnol. Bioeng.* **27**, 202–207 (1985).
19. A. Mulchandani, J.H.T. Luong, and A. LeDuy, *Biotechnol. Bioeng.* **32**, 639–646 (1988).
20. J.M. Boa and A. LeDuy, *Biotechnol. Bioeng.* **30**, 463–470 (1987).
21. J. Audet, M. Lounes, and J. Thibault, *Bioprocess Eng.* **15**, 209–214 (1996).
22. S. Onogi, T. Masuda, and K. Kitagawa, *Macromolecules* **3**, 109–116 (1970).
23. A. LeDuy and J.E. Zajic, *Biotechnol. Bioeng.* **15**, 579–581 (1973).
24. G.B. Tatterson, *Fluid Mixing and Gas Dispersion in Agitated Tanks*, McGraw-Hill, New York, 1991, pp. 417–528.
25. H. Guterman and Y. Shabtai, *Biotechnol. Bioeng.* **51**, 501–510 (1996).
26. D. Rho, A. Mulchandani, J.H.T. Luong, and A. LeDuy, *Appl. Microbiol. Biotechnol.* **28**, 361–366 (1988).
27. R. Dufresne, J. Thibault, A. LeDuy, and R. Lencki, *Appl. Microbiol. Biotechnol.* **32**, 526–532 (1990).
28. M. Lounes, J. Audet, J. Thibault, and A. LeDuy, *Bioprocess Eng.* **13**, 1–11 (1995).
29. J. Audet, J. Thibault, and A. LeDuy, *Biotechnol. Bioeng.* **52**, 507–517 (1996).

30. J. Audet, H. Gagnon, M. Lounes, and J. Thibault, *Bioprocess Eng.* **19**, 45–52 (1998).
31. H. Brauer, *Bioprocess Eng.* **6**, 939–944 (1991).
32. H. Gagnon, M. Lounes, and J. Thibault, *Can. J. Chem. Eng.*, **76**, 379–389 (1998).
33. J.F. Perez and O.C. Sandall, *AIChEJ.* **20**, 770–775 (1974).
34. M. Nishikawa, M. Nakamura, H. Yagi, and K. Hashimoto, *J. Chem. Eng. Jpn.* **14**, 219–226 (1981).
35. K. Guerinik, Ph.D. Thesis, Université Laval, Sainte-Foy, Canada, 1989.
36. P.A. Tanguy, F. Thibault, E. Brito de la Fuente, T. Espinosa-Solares, and A. Tecante, *Chem. Eng. Sci.* **52**, 1733–1741 (1997).
37. A. Mulchandani, J.H.T. Luong, and A. LeDuy, *Biotechnol. Bioeng.* **33**, 306–312 (1989).
38. D.J. Siehr, *J. Coatings Technol.* **53**, 23–25 (1981).
39. G.M. Gadd, *FEMS Microbiol. Lett.* **9**, 237–240 (1980).
40. Y. Shabtai and I. Mukmenev, *Appl. Microbiol. Biotechnol.* **43**, 595–603 (1995).
41. T.P. West and B. Strohfus, *World J. Microbiol. Biotechnol.* **13**, 233–235 (1997).
42. T.J. Pollock, L. Thorne, and R.W. Armentrout, *Appl. Environ. Microbiol.* **58**, 877–883 (1992).
43. B.J. Catley, *FEBS Lett.* **3**, 190–193 (1970).
44. B. McNeil and B. Kristiansen, *Biotechnol. Lett.* **9**, 101–104 (1987).
45. K.Y. Lee and Y.J. Yoo, *Biotechnol. Lett.* **15**, 1021–1024 (1993).

PULP AND PAPER PROCESSING. See ENZYMES, PULP AND PAPER PROCESSING; HEMICELLULASES; HEMICELLULOSE CONVERSION.

PUMPS, INDUSTRIAL

BOB STOVER
ED DOMANICO
Tri-Clover
Valencia, California

KEY WORDS

Cavitation
Centrifugal
Head
Horsepower
Mechanical seals
Motors
Net positive suction head (NPSH)
Positive displacement
Specific gravity
Viscosity

OUTLINE

INTRODUCTION

Liquid transfer is a vital component of most biotechnology processes. There is possibly more information available on pumps than any other single piece of equipment (pumps being a very common piece of rotating equipment, second only to the electric motor) existing in a typical plant, so this section must be a rough introduction to hydraulic theory. This article provides a basic understanding of pump mechanics, nomenclature, and principles and sufficient information to begin the selection process of a new pump.

THEORY

Before a specific selection can be initiated, the operating conditions the pump will be expected to meet must be understood. Breaking down the system into clearly defined and agreed-upon components will not only simplify calculations, and help guarantee nothing is overlooked, but also make it easier to evaluate the consequences of changes to any part of the system. There are universally understood and applied ways of breaking down a system, recognized by all pump manufacturers and users.

Flow

Capacity is expressed in flow/unit of time, commonly gallons per minute (gpm) or per hour (gph), or as cubic meters per hour (m^3/h).

To convert from volume to gallons:

$$1\ ft^3 = 7.48\ gallons$$
$$1\ m^3 = 264.2\ gallons$$

It is important to be realistic with estimates, because accurate pressure loss calculations depend on an accurate flow rate. Overestimating will result in equipment that will need to be valved back or slowed down to produce the demands actually required or in oversized/overpriced pumps.

Head

Head is the accepted hydraulic term for pressure. It is most commonly expressed in pounds/inch^2 (psi), or feet of head. The concept of a linear unit being used to express pressure can be most easily visualized as taking a resistance to flow and mentally converting it to a vertical column of water, such that the weight of this imaginary column is equivalent to the resistance of the considered section. Any pump that can push fluid against a column of water of the height thus calculated will be able to deliver flow against the system used to make the calculation.

To convert from psi to feet:

$$1\ psi = 2.31\ ft\ of\ water$$
$$1\ ft\ of\ water = .433\ psi$$

Most pipe friction loss charts express their losses as feet of head per 100 feet run of pipe, and all losses given as psi can be converted to feet by the above conversion factors (see Specific Gravity for adjustments given liquids heavier or lighter than water).

This linear unit has become the most common choice for hydraulic pressure (in the sense of resistance to flow) because:

- It is easily visualized.
- A linear unit fits into most hydraulic formulas (velocity, horsepower, specific speed, etc.).
- It compensates for various specific gravities before inclusion.
- It provides a common unit for all components of the total head that can then be added together to arrive at total resistance.

Types of Head. The total head referred to above is made up of four components. Any resistance to flow will fall into one of the following categories.

Static Head (ΔZ). Static head refers to the elevation differential between the source level of the liquid and the elevation at the destination. In other words, the elevation differential the pump must overcome. Caveats to this are:

- If a line enters at or near the bottom of a destination tank, the liquid level *in* the tank is considered the appropriate one, not where the pipe makes its entry (the pump must overcome the weight of the column *inside* the tank as part of its pressure capability).
- The vertical ups and downs are irrelevant (they will be accounted for later under Friction Head); the elevations at the beginning and end of the line are the ones used for ΔZ.
- The pump's position in relationship vertically to either the source or destination point is also irrelevant.

Static head is easily expressed in feet.

Pressure Head (ΔP). Pressure head is the differential in pressure at the ends of the system, the inlet and outlet, expressed in feet. An example would be a vacuum tank at the source; it provides resistance to flow (the vacuum holds back the fluid) and must be overcome by the pump. Resistance to flow from vacuum can be converted to feet of head by the following:

$$1\ inch\ mercury\ (Hg) = 1.133\ ft\ of\ water$$

Pressure at the discharge end (a pressurized tank or injection into a pressurized line) also needs to be added into the total and is included in ΔP.

Velocity Head (H_V). Velocity head is that energy necessary to overcome inertia. It is calculated by

$$H_V = V^2/2g$$

where V = velocity through the piping lines (in feet/second), and g = 32.2 ft/s^2, the acceleration caused by gravity at sea level.

Velocity head is reasonably estimated by taking the average velocity in a system and applying the formula. However, given a maximum average velocity of less than 8 ft/s in most systems, H_V is usually a relatively small number (less than 1 ft in the average 8 ft/s example above) in a typical hydraulic calculation and almost always ignored.

Friction Head (H_F). Friction loss is usually the largest individual component of total head for most calculations. There are formulas available for its calculation given various types of pipe, fittings, valves, and filters in texts, but fortunately there are charts in existence giving the calculated heads. Pipe losses themselves are usually expressed in feet of head/unit of length (usually 100 ft). Fitting losses are expressed either in equivalent feet of straight pipe to be added to the run of actual straight pipe for what is called an *equivalent feet of pipe* to be used in the calculations. Or, charts may give the actual loss per each fitting. Losses through valves, filters, and such are often expressed in psi or feet of loss per valve. Each component in a line must be considered, and a grand total for the system is arrived at as a total of the individual pieces.

The friction loss numbers will always vary considering the flow through the line; higher flows mean more drag and therefore more resistance. Filters will vary with flow and also condition of the filter element. It is common practice to use worst-case conditions.

Total Head (H_M). The four individual elements listed previously are added together for the *total mechanical*

head (H_M, also known as mechanical head, total dynamic head [TDH], or simply total head). They are the components for Bernoulli's formula, modified for the pump industry, given as:

$$H_M = \Delta Z + \Delta P + H_V + H_F$$

Adjustments to Head. There are qualities a liquid can possess that will affect the H_M.

Specific Gravity. Specific gravity (SG) is the direct numeric ratio of the weight of a particular liquid to the weight of water. Given that at certain standard conditions water weighs approximately 8.34 lb/gallon, anything heavier than that would have a specific gravity greater than 1.0 (for example, a liquid weighing 16.68 lb/gallon would have an SG of 2.0), anything lighter less than 1.0. Specific gravity will affect a conversion from psi to feet as follows:

$$\text{psi} = (\text{feet of water} \times \text{SG})/2.31$$

or

$$\text{feet of water} = (\text{psi} \times 2.31)/\text{SG}$$

Specific gravity will also directly affect horsepower because of the extra torque needed to overcome the additional weight of the liquid (as will be shown under Horsepower).

The second quality a liquid possesses that directly affects total head calculations is its *viscosity*. Most pump selection charts assume waterlike viscosity (1 centipoise [cps] or 31 Saybolt second universal [SSU]). The viscosity of liquids thicker than water affects pumps in several ways:

- It affects friction loss calculations; losses are higher with more viscous liquids. There are separate charts used for liquids more viscous than water.
- It affects horsepower. Manufacturers often provide specific charts to give the added torque needed to overcome viscosity.
- It affects the very capability of a pump to even *deal* with a product. Centrifugal pumps especially have limited abilities to handle viscous products.

Viscosity is usually expressed in the referenced units (cps or SSU), but there are many others. Viscosity can vary under dynamic conditions. Specifically:

- If a liquid becomes thinner under shear, it is *thixotropic* (example: paint).
- If it becomes thicker under shear, it is *dilatant* (example: taffy).
- If it becomes thinner only after exceeding a certain minimum threshold of shear, it is *plastic* (example: ketchup).
- If a liquid is unaffected by shear, it is *Newtonian* (example: water).

As mentioned above, viscosity affects centrifugal pumps in their very ability to pump a liquid; with viscous fluids their pressure capability falls off, the flow capability falls off even more, and their efficiency may fall off to a point where their use is totally unacceptable. Positive displacement pumps are affected in the speed they can turn and still move product without starving at the inlet. Manufacturers provide charts giving maximum speed for varying viscosities, for adjustments needed to horsepower to drive the unit with a thicker product, and perhaps for minimum port sizes necessary to allow product to get in quickly enough.

If the viscosity of a product is unknown, many pump manufacturers provide services wherein a liquid's viscosity will be tested, often at no charge, in exchange for the opportunity to quote their pumps.

Horsepower

Given the results of the total head calculations above, it is possible to make accurate calculations of the horsepower necessary to drive a particular pump.

The basic horsepower formula is

$$\text{HP} = (\text{gpm} \times H_M \times \text{SG}) \div (3{,}960 \times \text{eff.})$$

where gpm = flow (in gallons/min); H_M = total mechanical head (in feet); SG = specific gravity; eff. = efficiency (expressed as a decimal; e.g., 50% would be .50). *Note*: 3,960 is a constant, derived from the horsepower units and the weight of water.

A variation of this formula, commonly used for positive displacement pumps wherein specific gravity has already been calculated into the total head before being expressed as psi, is

$$\text{HP} = (\text{gpm} \times \text{psi}) \div (1{,}715 \times \text{eff.})$$

The 3,960 constant has been changed to 1,715 to provide for psi as a pressure unit by dividing by 2.31.

Often, viscosity will affect the horsepower as calculated above; in those cases manufacturers will provide correction factors for differing viscosity. The Hydraulic Institute has generic charts for centrifugal pumps that seem to work fairly well for almost all manufacturers, although some especially high-efficiency pumps claim to outperform the charts and therefore provide their own correction factors.

Net Positive Suction Head

Net positive suction head (NPSH) is the hydraulic term referring to the pressure under which liquid enters the pump, not only in terms of what the existing system is able to provide but also in reference to what any particular pump needs in order to function properly. It is an absolute unit, always expressed relative to absolute zero (as opposed to gauge units, whose values are always in reference to normal atmospheric pressure, called zero gauge). If a pump does not receive the liquid under sufficient pressure, it will cavitate.

Cavitation. Liquids will stay in liquid form only if sufficient atmospheric or system pressure is available to prevent it from going off to vapor. Even cold water, for exam-

ple, will flash off into vapor in a strong enough vacuum, outer space as an obvious example. All liquids have a predictable vapor pressure necessary to keep them in their liquid state, varying up or down respectively as the liquid is heated or cooled.

This is of particular interest in hydraulics because of the often high vacuum present at the inlet of a pump. If the pressure provided by the system is less than the liquid's vapor pressure, the pump will undergo a physical process called cavitation. This action varies in specifics from centrifugal to positive displacement type pumps, but the general process is always approximately the same:

- The liquid finds itself in an area of low pressure, the pump's inlet port.
- If the pressure is *too* low (less than the liquid's vapor pressure at the operating temperature), it undergoes a state change from liquid to vapor. Because the vapor is not free to escape, it stays entrained in the liquid and assumes a bubble (spherical) shape.
- The vapor bubbles are carried further into the pump with the flow of surrounding liquid.
- Eventually the low-pressure bubbles and surrounding liquid find themselves in a higher pressure area as they near the pump discharge.
- The higher pressure causes the vapor bubbles to implode (collapse). This implosion is called cavitation.

The implosion of the low pressure vapor bubbles makes a very identifiable sound; in centrifugal pumps it produces a rattle, sounding like gravel is being carried through the housing; in positive displacement pumps the sound is more percussive, often *very* loud. In both cases the implosion of the vapor bubbles is damaging to the pump itself and often to the product. Centrifugal impellers and positive displacement pump rotors will actually have material removed by the very powerful implosions, shafts can be deflected to the point of fatigue, higher horsepower is required, and performance will be negatively affected.

Net Positive Suction Head Required. Manufacturers can test their pumps to see exactly how much positive suction pressure is necessary at various flows to keep their pumps from cavitating. This is called the net positive suction head required ($NPSH_R$). Typically, the greater the flow, the more pressure necessary because of the higher velocities going through the housing. A possible exception: at extremely low flows, the slippage caused by recirculation within the housing in some centrifugal pumps may create low pressure areas, sometimes low enough to fall under a liquid's vapor pressure and allow cavitation. This is called either suction or discharge recirculation cavitation, depending on where it physically occurs in the pump housing. Manufacturers determine the $NPSH_R$ from empirical data. They set up a test skid wherein they can gradually restrict flow into a pump, constantly monitoring the performance with pressure gauges and flow meters. Eventually they find a point at which the restrictions affect performance an agreed-upon minimum amount (3% below normal curve is the current standard). The pump is considered to be cavitating, and this point is registered on a chart. The same process is performed for a number of flows, and an $NPSH_R$ curve is generated. $NPSH_R$ curves are assumed to be drawn using 60 °F water; adjustments can be made later for other liquids.

Net Positive Suction Head Available. Net positive suction head available ($NPSH_A$) is the pressure any particular system makes available at a pump's inlet under operating conditions. It is a total of several factors. After calculation (or determination by an actual physical test), it will be compared to the pump's $NPSH_R$. Also by previous agreed-upon standard, $NPSH_A$ must exceed $NPSH_R$ by a minimum of 2 ft for a pump to be considered acceptable. This is a controversial issue; some people believe that this differential is insufficient, that the $NPSH_R$ chart represents the point at which the pump is *already* cavitating, a 2-ft differential is often too low; and that a sliding scale of minimal percentages by which $NPSH_A$ should exceed $NPSH_R$ would serve the user better. However, at this time the 2-ft differential is still accepted by most hydraulic engineers.

$NPSH_A$ is calculated by the formula:

$$NPSH_A = H_{ATM} \pm \Delta Z - H_F - H_{VAP}$$

where H_{ATM} = atmospheric pressure in absolute units, expressed in feet (atmospheric pressure at sea level = 14.7 psi = 33.9 ft); ΔZ = elevation differential from liquid level to pump's inlet (in feet). If the liquid level is above the pump inlet, the number will be positive (+), if below it will be considered a negative (−); H_F = friction loss in suction line only (in feet); and H_{VAP} = vapor pressure of liquid at operating temperature (in feet).

H_{ATM} and H_{VAP} are available from various reference books; H_F and ΔZ are calculated as in the section on Friction Head.

$NPSH_A$ can be determined on an operating pump by installing either a compound or vacuum gauge at the pump's inlet and by taking this reading and subtracting the liquid's vapor pressure (which will not be indicated on the gauge: H_{ATM}, H_F, and ΔZ will all be indicated by the gauge reading).

As stated earlier, the results of the calculation (or gauge reading $-H_{VAP}$) should be compared to the manufacturer's $NPSH_R$, a minimum 2-ft differential necessary to guarantee a pump will not cavitate. The $NPSH_A$ calculation is valuable for several reasons:

- It can show us beforehand whether a pump will cavitate after installation.
- It indicates the area most easily altered to eliminate existing cavitation and how much it needs to be altered. For example, if H_F is high, it may be lowered with larger pipe or fewer elbows. If ΔZ is low (or even negative), we can pick up positive head by raising the source tank.

If a system is still cavitating, and the $NPSH_A$ calculation does not indicate an easy correction, it may be that a pump with a lower $NPSH_R$ is needed. There are whole

classes of pumps considered low NPSH that typically have requirements below 1 ft or 2 ft.

CENTRIFUGAL PUMPS

Pumps generally are divided into two main classifications, centrifugal and positive displacement (PD). There are many subclassifications. Several will be listed here later according to their suitability for use in bioprocessing, including a few that seem to cross the boundaries of centrifugal and PD. The vast majority of pumps in use are centrifugal because of the many advantages of centrifugal design (to be listed later), but bioprocessing has unusual demands, and the percentages of positive displacement pumps in use is significantly higher than in other industries.

Principals of Operation of Centrifugal Pumps

Centrifugal pumps (Fig. 1) consist of

- A rotating impeller
- An impeller housing (volute) to channel the flow from the suction port to the discharge port
- A power source, usually an electric motor
- A connection from the motor to the impeller; either the motor shaft itself, or sometimes a separate shaft extension
- A mechanical seal assembly to seal liquid inside the liquid-end assembly at the point where the shaft goes through the backplate.

There are many designs of impeller, having to do with whether the vanes are enclosed by shrouds or not (open or enclosed; open is cleaner, but less efficient; enclosed has less slip and is therefore more efficient, but is much harder to clean and to inspect); number and angle of the vanes themselves (influencing the shape of the pump curve); and diameter and thickness of the vanes themselves (deeper vanes give more flow, larger diameters mean more pressure capability). Manufacturers can create impeller designs for very specific sets of conditions. A centrifugal pump's simple design means that

- Centrifugal pumps are usually less expensive (often *considerably* less expensive) than PD pumps.
- Centrifugal pumps can be close-coupled, meaning they can be mounted directly on the face of specially designed close-coupled motors, a rare feature on PD pumps, helping keep them inexpensive (and also compact).
- Their flow is relatively pulsation free.
- Because of the inherent slip caused by the generally open nature of the impeller/volute package, there is a maximum pressure (shut-off head) a centrifugal pump can develop. This also means they can be easily controlled by regulating valves and often do not need pressure relief valves or controls in the system because HP loads actually decrease when flow is valved back.
- They are easy to maintain and repair.
- They are easily adapted to sanitary standards.

There are also disadvantages possessed by centrifugal pumps that are often especially significant in bioprocessing.

- Their inherent slip means they do not handle viscous materials well. The rule of thumb is anything over 2,000 SSU (about 500 cps) is usually considered too viscous for a centrifugal pump.
- Because pressure is generated by vane tip speeds, centrifugal pumps are not by nature easily adapted to high-pressure applications. This can be overcome to a degree by running them especially fast (designs exist that run at over 10,000 rpm) or by incorporating multistage designs wherein a chain of impellers is mounted on a common shaft, each impeller boosting the pressure of the impeller preceding it.
- Centrifugal pumps must be run at relatively fast speeds to develop sufficient impeller vane tip velocity to generate reasonable head (usually a minimum of 1,200 rpm). PD pumps, on the other hand, are often run efficiently at well under 100 rpm. This minimum speed means that centrifugal pumps are not usually good in low shear applications.
- The minimum speed also means that centrifugal pumps are not efficient with low-flow applications; turning slowly to limit flow means all pressure capability is gone. Even very small centrifugal pumps are often rated at 10 gpm or more; they will run at lesser flows by regulating them back in any of various fashions (regulating valves, inverters, bypass lines), but they are not efficient at very low flows.
- The presence of a mechanical seal means that abrasive products are a potential problem, not only because of the limited life of the components, but also for the danger of materials wearing off the faces and

Figure 1. Centrifugal pump. *Source*: Courtesy of Tri-Clover Inc., Kenosha, Wis.

entering the product stream. There are designs that remove seals from the product flow (double mechanical seals) or eliminate seals altogether (magnetically driven pumps, or pumps with gas injection seals), but they are often expensive and sometimes not available in sanitary designs.

Types of Centrifugal Pumps

End Suction. By far the most popular general classification of centrifugal pump is the end-suction design. Liquid enters through the suction port in the center of the housing, is picked up by the leading edge of the impeller, has energy imparted by the impeller's rotation, and exits tangentially at the discharge port on the volute's periphery. End-suction centrifugal pumps are available as close coupled (mounted directly onto a motor's face) or pedestal mounted (the pump is a separate entity with its own shaft/bearing housing, connected to the motor by a flexible coupling, the assembled package bolted to a common base). Close-coupled design is generally preferred because of lower cost, its more compact design, and its inherent alignment between components, but pedestal-mounted pumps often have more space for special seal options and often will adapt more easily to any special motor requirements.

Magnetically Coupled. One way to arrive at a seal-less design and remove the difficulties of seal compatibility (see Mechanical Seals) is to adapt the standard end-suction design. Rather than the motor shaft extending directly through the mechanical seal and into the liquid end of the pump, it drives a large, hollow rotating magnet. The liquid end of the pump is in a sealed housing and fits up against the motor with a magnetized impeller extended into the hollowed-out area of the drive magnet. As the motor shaft rotates the hollow drive magnet, the impeller is pulled along. Because of the lack of a direct connection between the motor and impeller, magnetically driven pumps can handle only limited viscosity or specific gravity before the impeller decouples (uncouples). They can be very inexpensive in smaller designs, but the price escalates quickly as they get larger. Often, soft-start motors are necessary to prevent decoupling. At present, there are no sanitary-design magnetically driven pumps, but they will undoubtedly be available soon.

Multistage. Centrifugal design is normally best suited for medium-to-high flow, low-head applications such as liquid transfer. When the many advantages of centrifugal pumps makes them the design of preference, but the total head necessary exceeds that available from standard models, multistage pumps may provide an alternate to PD pumps. Impellers are staged, one after another in a line, and are driven by a long common shaft. Internal channels see that the first impeller's discharge is fed to the inlet side of the second, which discharge feeds the third, and so forth. Pressure is cumulative, but the flow is limited to that which can be carried by any individual impeller. The only restriction to the number of stages, and therefore the total pressure possible, is the housing-pressure limitations. This design carries all the advantages of standard centrifugal pumps, except they are not as easy to repair (impellers must be removed in series, and clearances can be difficult to adjust), and they are not necessarily inexpensive. There are sanitary designs available.

POSITIVE DISPLACEMENT PUMPS

Design and Principles of Operation

Positive displacement pumps make-up the second major classification of pumps used in the bioprocess industry today. In contrast to centrifugal pumps, discussed earlier, PD pumps function on a very different principle: that compressing a fluid volume will increase the ability of the fluid to do work. This change in volume takes place within the pump cavity as the pump drives or pushes that volume to the outlet, thereby increasing the pressure on that volume of fluid to be equal to the pressure at the pump outlet or discharge. An example of this principle is a simple syringe. As the syringe plunger is pushed in, the fluid within the syringe will begin to be exposed to a higher pressure as the volume holding the fluid decreases. The higher pressure then drives the fluid out the end of the syringe. In both the syringe and a pump, the pressure that the pump will generate is controlled by the resistance to flow that the fluid is exposed to as it leaves the pump discharge. Unlike dynamic or centrifugal pumps, this creates a situation whereby the PD pump can become overpressurized. The PD pump will continue to drive fluid forward regardless of the level of resistance. This overpressurization will only be relieved by either a pressure relief device on the discharge side of the pump or through the slip that is experienced as the pump operates.

Slip, in a positive displacement pump, is defined as an efficiency calculated by taking the amount of fluid that escapes the encapsulated volume as the pressure of the fluid volume increases and dividing by the amount of fluid delivered. The escaping fluid is driven back to the inlet or low-pressure side of the pump through clearances within the pump. By knowing the amount of slip that the pump is experiencing, one can calculate the pump efficiency. Typically, PD pump efficiencies are relatively high, 80 to 90%, and will increase as the fluid viscosity, discussed earlier, increases.

Advantages. Advantages that can be found in using PD pumps in the bioprocess industry include:

- The ability to deliver a constant flow rate; the discharge rate for a PD pump is relatively constant, against a varying discharge pressure.
- They function well at a low-flow rate; the PD pump can deliver a low-flow rate at a high discharge pressure and still maintain a high efficiency as compared to centrifugal pumps.
- They function well at high viscosity; the PD pump can easily be adjusted to handle high viscosity liquids and, in fact, will run at a higher efficiency as the higher viscosity acts to reduce the amount of slip experienced by the pump.

- They have the ability to self-prime; most PD pumps will act to create a vacuum at the pump inlet. This vacuum will then pull liquid into the pump cavity.
- PD pumps can easily be made to sanitary design specifications.

Disadvantages. Disadvantages that can be found in using PD pumps in the bioprocess industry include:

- Overpressurization can occur; a flow restriction on a PD pump will build pressure until that pressure is relieved. An internal or external pressure relief valve may be required for safe operation of the pump.
- Pulsing flow will occur; the design of PD pumps is such that usually they will cause pressure pulsation as they maintain flow rate. This pressure pulsation can act to damage discharge piping caused by excessive vibration. A pulsation dampener may be required.
- High flow rates are hard to maintain; high flow rates may be hard to achieve even with large pumps because the maximum pump speed may be limited by the viscosity and suction port size.
- Power limitation; because the discharge pressure is directly proportional to the horsepower required, very high discharge pressures may require a very high horsepower, which may exceed the physical strength of the pump.

Types of Positive Displacement Pumps

Gear Pumps. Gear pumps (Fig. 2) comprise, perhaps, the most common PD pump in industry. They have wide acceptance because of low cost and simple construction. In general, gear pumps are made in two very distinct designs. The first is the external gear pump. This design uses two meshed gears that rotate in opposite directions. The first gear is driven by the pump motor, referred to as the driven gear, and the second gear is driven by the first and is called the pump gear. Fluid enters and exits the gear housing perpendicular to the gear mesh, with the incoming fluid being trapped within the gear teeth as the opening mesh creates a vacuum pulling liquid in. This trapped liquid rotates on the outside of the gears and is expelled at the pump discharge port as the gear teeth again start to mesh. The fluid is forced out of the gear housing at the higher pressure of the discharge side of the pump.

The other type of gear pump used in industry is the internal gear pump. The internal gear pump works under the same principle as the external gear pump; however, the gears are configured such that a smaller, off-center, internal gear is placed inside a planetary gear. The internal gear meshes with the inner teeth of the planetary gear to create the pumping action. In this design, the larger planetary gear is the motor driven gear, with the inner gear acting as an idler. The gap created by mounting the internal gear off center is filled with a crescent-shaped scraper. This scraper acts to seal the inlet port from the discharge port and reduces the amount of slip in the pump. This pump is slightly more expensive, but is quieter and often longer lived.

Air-Driven Double-Diaphragm Pumps. Double-diaphragm pumps (Fig. 3) are a unique type of PD pump. The design of a double-diaphragm pump is made up through the porting of four check valves and two dia-

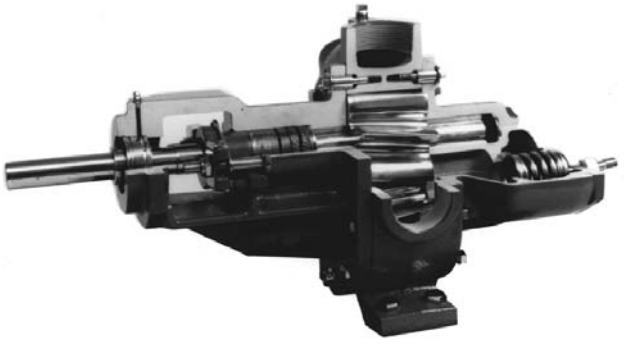


Figure 2. Gear pump. *Source*: Photo used by permission of Roper Pump Company.

Figure 3. Double-diaphragm pump. *Source*: Courtesy of Wilden Pump Company.

phragms mounted in parallel. The two diaphragms are connected to an air-driven piston. The operation of the pump has liquid entering a common inlet port and the liquid separating and alternately flowing through one of two check valves, the check valve that is exposed to the opening diaphragm. At the same time, on the opposite side of the closed check valve, the closing diaphragm is expelling fluid at a discharge pressure (limited by the air supply pressure that powers the pump). The fluid being expelled pushes open the check valve exiting the closing diaphragm section and pushes closed the check valve on the opening diaphragm. This action cycles back and forth as the air piston that drives the two diaphragms moves in and out. Because both diaphragms are attached to the same shaft, when one diaphragm is opening, the other is closing.

Progressing Cavity Pumps. Progressing cavity pumps (Fig. 4) have a special design that allows for very gentle pumping of highly viscous liquids. Most progressing cavity pumps are made up from a single helical (metal) rotor rotating inside a (elastomeric) double helical stator. Cavities formed by this combination of rotor and stator progress down the length of the rotating shaft as the shaft turns. The function of the rotor is very similar to that of a screw pump slowly pushing the pumped liquid forward toward the discharge of the pump. Pressure is developed because close clearances between the rotor and stator minimize the amount of slip that is found in the pump. Further, because the pump is well suited to pumping high-viscosity liquids, this action is even better utilized.

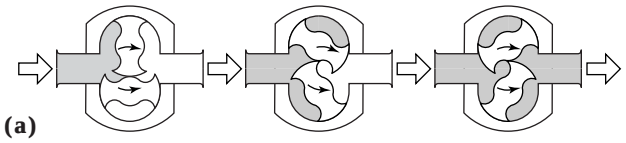
Flexible Impeller Pumps. Flexible impeller pumps (Fig. 5) make up a class of very inexpensive PD pumps that can be easily adapted to many applications. Flexible impeller pumps work on the principle of a multivaned rubber impeller rotating in a housing that is cammed such that the vanes are pushed close as the impeller rotates. This closed impeller expels fluid out the discharge of the pump and at the same time pulling liquid into the suction port. This action creates a vacuum at the suction port. Vacuum values as high as 22 in Hg are easily achievable. The major drawback of this style of PD pump is the inability of the pump to run dry. Dry running of this type of pump will quickly erode the rotor.

Figure 4. Progressing cavity pumps. *Source*: Photo used by permission of Roper Pump Company.

Figure 5. Flexible impeller pump. *Source*: Courtesy of ITT Jabsco.

Circumferential Piston Pumps. Circumferential piston pumps (Fig. 6a) are used in the bioprocess industry to transfer a wide range of both raw material and finished products. The circumferential piston pump works with two opposing hemispherical rotors spaced 180° apart. These two rotors turn in opposite directions much like that of the external gear pump; however, the action of the circular piston moving past the suction port creates a vacuum that draws fluid into the pump cavity. This fluid, which is trapped between the sides of the rotating piston and the pump housing, travels around the pump casing and is discharged at the pump outlet as the piston of the opposing rotor pushes the trapped liquid into the outlet of the pump. This action alternates from pump rotor to driven rotor as trapped fluid is pushed into the pump outlet. This pump design requires that the rotors of the pump not come in contact with one another. To accomplish this, the rotors are driven with a gear housing that contains supporting shafts, bearings, and timing gears. It is therefore very clean and ideally adds no metal to the product.

Lobe Pumps. Very similar in design to circumferential piston pumps, lobe pumps (Fig. 6b) also use a gear housing containing supporting shafts, bearings, and timing gears. The rotors of the lobe pump rotate in opposite directions, creating a low pressure or vacuum at the pump inlet as the lobes separate. The fluid drawn into the pump cavity



(a)

(b)

Figure 6. (**a**) Circumferential piston pump; (**b**) Lobe pump. *Source*: (**a**) courtesy of Waukesha Fluid Handling; (**b**) courtesy of ITT Sherotec.

is trapped in the valley of the lobe and rotated to the outlet of the pump where the trapped fluid is forced into the pump outlet as the opposing rotor lobe rotates into the lobe valley containing the trapped liquid. This action alternates from rotor to rotor similar to that found on the circumferential piston pump. The lobe pump is somewhat less shear intensive as compared to the circumferential piston pump and is more easily cleaned when cleaning a pump in place.

Peristaltic Pumps. Peristaltic pumps (Fig. 7) are very common throughout the bioprocess industry. They can be found in all size ranges, handling flows on the laboratory bench of 3 to 4 mL/min to flow rates on the production floor of 30 to 40 gpm. The basic design of the pump functions through the progressive squeezing of tubing by rollers mounted to a rotating plate. The rollers trap fluid within the tubing and rotate it to the outlet. They are self-priming and have the major advantage that the fluid only contacts the pump tubing so sealing the pump is not required. The action also minimizes the amount of shear that the pump fluid is exposed to.

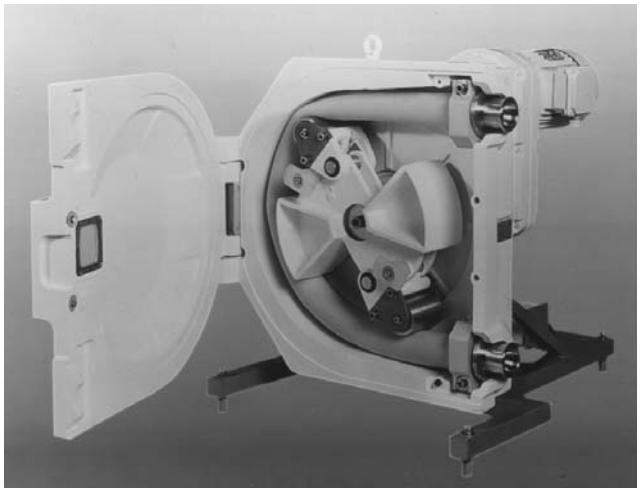


Figure 7. Peristaltic pump. *Source*: Courtesy of Watson-Marlow Inc.

Nutating Disk-Type Pumps. Nutating disk-type pumps (Fig. 8) are new to the bioprocess pump market and provide for advantages typically found in several different types of older pump designs. The principle that drives this class of PD pumps is the action of a wobbling disk within the pump cavity. This disk can be used to trap a volume of fluid between it and the stationary pump housing. The trapped fluid is pushed toward the pump outlet through this wobbling action and the use of a scraper gate that rides on the width of the rotating disk.

DRIVERS

The most popular choice of drivers for pumps, by far, is the AC motor. They have become the standard because they

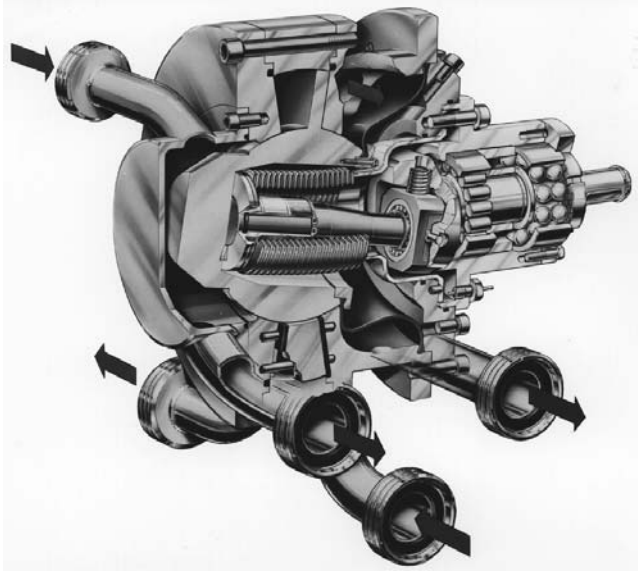


Figure 8. Nutating disk-type pump. *Source*: Courtesy of KSI.

are inexpensive and long lived, have relatively few moving parts, and are available from a large variety of sources, in a large variety of styles and enclosures. They are relatively quiet (typically between 70 and 85 decibels), efficient (often over 90%), and institutions exist (National Electrical Manufacturer Association [NEMA]) that guarantee a certain degree of interchangeability from manufacturer to manufacturer. Typically, only if an AC motor is unacceptable are other choices considered. Considerations include style, electrical specifications, and enclosure.

Style

Close coupled, foot-mounted motors are common for centrifugal pumps that will accept this. Standard horizontal foot-mounted motors are needed when pedestal-mounted centrifugal pumps have been chosen. PD pumps usually need motors with gear boxes integrally mounted to achieve the low speed necessary. Because most pump suppliers sell their pumps as motor and pump packages, the user is not usually concerned with physically matching the components, but should understand the principles.

Electrical Specifications

Unlike the general motor style, it is the user's responsibility to know the electrical systems the drive will be expected to operate under. Concerns the supplier must address will be:

- *Phase*: Single phase is available for small horsepower drivers (below 5 hp), and common for fractional horsepower, but three phase is available for fractional and large integral motors.
- *Voltage*: in the United States, standards are 115 or 230 V for single phase and 230 or 460 V for three phase, but others voltages exist. Because motors are good for $\pm 10\%$, there is some flexibility.
- *Frequency*: 60 Hz is the standard in the United States, but 50 Hz exists all over the world. Motors tolerate only $\pm 5\%$ frequency variation.

Enclosure

The desired enclosure is an important consideration. Standard types available are open drip-proof for environments where the motor will be protected from atmospheric dangers; totally enclosed, fan cooled (TEFC) for areas where the motor will be exposed to light moisture; explosion proof where a UL rating is necessary and the motor's ability to keep sparks from escaping is crucial for safety considerations; and common in bioprocessing, the wash-down enclosure where the area the unit is to be mounted in will be periodically hosed down.

NEMA is an organization that has written a standardized list of motor dimensions such that a user may substitute one brand for another in confidence that the new motor will physically interchange. This is done by a series of frame designations, such that a brand A motor in, for example, frame 145TC will exactly fit in place of a brand B motor of the same frame.

An accessory often considered today is the frequency inverter, which varies frequency and voltage in such a fashion as to allow an AC motor to act as a variable speed device. This has many advantages in addition to allowing a user to dial in more precisely the desired flow, such as the ability to control speed remotely, and to be able to start up a pump slowly to minimize start-up damage to the pump and the motor.

Other options to an AC motor are gasoline or diesel engines for areas where electrical power is unavailable, and more and more commonly, compressed air drives. Air motors have the advantages of being inherently explosion proof and being able to stall without incurring damage, but have the disadvantages of needing very dry compressed air and of being relatively inefficient. DC power is used in some very small pump packages, but is uncommon.

SPECIAL CONSIDERATIONS FOR BIOPROCESSING PUMPS

Sanitary Standards

The main standard for the sanitary design and construction of pumps to be used in the bioprocess industry is the 3A Standards (Code of Federal Register Title 21, part 177). These standards govern certain design modifications that are key elements in maintaining sanitary conditions within a pump. These guidelines cover surface finish, materials of construction, fabrication, and other areas. The sanitary design of a pump must also comply with all other applicable standards, including FDA, cGMP, and the ANSI/ASME standards for food, drug, and beverage equipment.

Materials of Construction and Surface Finishes

The most commonly found material of construction for pumps in the bioprocess industry is austenitic stainless steel alloys. Austenitic stainless steel alloys used in the construction of sanitary stainless steel pumps include AISI types 304, 304L, 316, and 316L. Analysis of the makeup of these alloys reveals that a major element in all stainless steel alloys is iron. Iron alone is quickly attacked by most fluids, but in the presence of chromium in correct combinations of Ni, C, Mn, P, Mo, Si, and S, the ability of stainless steel to withstand corrosion is excellent. This, plus its low carbon content and its ability to be easily welded, machined, and polished, makes 316L the most popular material of construction for pumps in bioprocessing. Product-contact surface finish is critical in most bioprocess systems because of the need to preserve sterile or sanitary conditions. All product contact surfaces must be free of cracks, crevices, and occlusions that can harbor and promote bacteria growth. Surface finish in most pumps is obtained through the use of mechanical polishing and electropolishing. This surface treatment will result in a surface finish that would measure at approximately 15 Ra surface roughness.

Mechanical Seals

Theory. Mechanical seals (Fig. 9) are a standard product used in the manufacture of most pumps to seal between

Figure 9. Mechanical seals. *Source*: Courtesy of the Durametallic Company.

the rotating pump shaft and the pump cavity. The type of mechanical seal used to seal the pump cavity is a key component of most pumps used in the bioprocess industry. The mechanical seal that is placed into pumps must serve two functions:

- Provide a barrier to the atmosphere
- Contain the fluid in the pump cavity without affecting fluid quality

Bioprocess pump seals are designed to be easily cleansed and not harbor bacteria. However, for them to be used in pumps for sterile fluids, particular modifications are required. These modifications include changes in the seal design, placement of the seal within the pump housing and selection of seal materials.

A discussion on mechanical seals is best started with a brief background on how mechanical seals work. A single mechanical seal works with two very flat surfaces, one a rotating face element and the other a stationary face element. These faces ride on one another with an ultrathin layer of fluid between the seal faces. When the fluid enters the zone between the two faces, the fluid temperature rises owing to the frictional resistance of one face on the other. If the seal is functioning correctly, the fluid will vaporize as it exits this zone and enters the atmosphere. This principle applies to most single mechanical seals in bioprocess pumps. The key to this functioning correctly is to have proper force pushing the rotating element on to the stationary seat. The force can be obtained either mechanically or hydraulically. The mechanical force is provided by a compressed spring. The hydraulic force, found in a hydraulically balanced seal, uses the pressure inside the pump head to push the rotating element to the stationary seat. Both of these methods are used in bioprocess pumps seals to produce a force that will generate enough friction between the faces to vaporize the fluid just before it leaves the seal faces without causing the faces to wear too quickly or generate too much heat.

If the mechanical seal is to function as described and provide a sterile barrier between the atmosphere and the fluid within the pump cavity, certain modifications are required to the seal design. First, a secondary mechanical seal is needed behind the primary inboard seal. This combination is referred to as a double mechanical seal. Second, both of these seals are placed into a stuffing box or gland that is attached to the pump where the rotating shaft enters the pump head. This is done so that a flushing fluid can be run between the two seals. The fluid barrier will guard against any contaminants from the atmosphere penetrating into the fluid within the pump cavity. The flushing fluid is pushed through the stuffing box under pressure so that a positive pressure on the outboard seal stops contaminants from entering the stuffing box. A secondary function of the seal flush liquid is to provide a certain amount of cooling to the seal faces.

The requirement that the seal contain the fluid within the pump cavity without affecting the fluid quality necessitates changes in the way the inboard seal functions. Fluid in the pump cavity is kept pure by ensuring that a positive pressure is maintained on the pump side of the inboard seal. This pressure functions in the same manner as that on the outboard seal; however, it should be greater than the stuffing box pressure to ensure that any leakage through the seal passes from inside the pump into the stuffing box. This will prevent the passage of contaminants from the stuffing box into the pump cavity. The seal designed for this configuration creates some problems.

First, if the inboard seal fails, product can be lost though the flush outlet of the stuffing box. Second, if the fluid in the pump contains active organisms that must be contained, these organisms could pass into the seal flush. The common flushing procedure for double mechanical seals in this situation is to maintain the stuffing box at a higher pressure than the pump side inboard seal pressure. In the case of seal failure, the flush fluid would move into the pump head, and no product would be lost. This arrangement can present problems if particulate, worn from the seal faces, is forced into the pump cavity by the high stuffing box pressure, thereby contaminating the fluid in the pump. Applications should be reviewed to determine which flushing configuration is most applicable.

The mechanical seal chosen for an application should be constructed of the highest quality FDA-approved materials of construction to provide a reliable, long-lasting seal. The face and rotating element materials should be selected for fluid compatibility, hardness, and flatness. Commonly used materials include tungsten carbide, silicone carbide, ceramics, carbon graphite, and stainless steel. The supporting items of the seal, including the springs, cups, O-rings, drive collar and set screws, should also be selected for fluid compatibility and performance characteristics. These pieces will function as a system to provide a reliable seal.

Static Seals

Static seals within pumps used in the bioprocess industry fall into two categories. These categories are elastomer O-rings and flat gaskets. O-rings are more commonly used in modern pump designs because they can be sized to properly seal flat or round pump components within the pump

design. Further, O-rings and their corresponding O-ring grooves can be easily configured to minimize any fluid entrapment and provide a cleanable crevice-free seal. O-rings are commonly used to seal pump shafts to mechanical seals and pump coverplates to pump housings. The O-ring material is selected based on the required durometer (hardness) and material compatibility with the sealed fluid. Commonly found materials include Viton®, EPDM, Buna-N, Silicone, and Kalrez®.

Flat gaskets are less common in modern pump designs; however, they are used when maintaining clearances within the pump design is critical. As with O-rings, they can also be constructed of an elastomeric compound, but it is also common to see them made from such materials as paper or PTFE (Teflon®).

Aseptic Modifications

Sterile fluids require systems that are capable of being steamed-in-place (SIP). To minimize downtime and preserve the integrity of the sterile atmosphere within the pumping system, pumps can be modified so they can be fully steamed-in-place. Most pumps can be steamed-in-place if they are designed in accordance with 3A standards. These standards, as mentioned earlier, will specify proper surface finish, impeller, and seal design. Further, the pump must utilize a seal that will allow for full steam penetration under pressure. Because proper steam flow is required to maintain saturated steam within the pump being sterilized, steam flow is enhanced in pumps by providing a means of condensate removal. As steam sterilizing the pump cools, it will form condensate. For proper sterilization to occur, this condensate must be removed from the pump quickly, preventing cold spots from forming. Hence, there is often the need for the pump to be self-draining by adding a tangential bottom drain.

Water-for-Injection Modifications (Centrifugal Pumps)

The design of centrifugal pumps make them very well suited to the application of pumping sterile water for injection (WFI). Because WFI must be kept moving at high temperature and high velocity within a piping system, a standard sanitary centrifugal pump can be modified to handle this application. The first modification is to the surface finish and materials of construction. The pump should be constructed of AISI Type 316L stainless steel that has been finished so that the surface roughness measures approximately 15 Ra. Second, all elastomeric components of the pump should be selected to be compatible with the WFI. This requires matching compatibility for temperature. Typical materials include steam resistant Viton® or EPDM. The third measure is the correct configuration of the mechanical seal to be used. This should be selected to reduce the chances of contamination. The last typical modification includes the addition of a casing drain to promote the removal of condensate during the steam sterilization of the WFI pump.

TROUBLESHOOTING

There is a general theme in successful troubleshooting. Specifically:

1. Make sure the system is fully understood (to eliminate oversights).
2. Narrow down or isolate potential sources of the problem.
3. Determine all possible causes within these areas.
4. In order of probability, go through them one by one.

The difficult step is often number 3, but all four hold potential hazards and solutions.

1. System analysis can be tedious. Fear of asking what appear to be obvious questions can tempt one to give this step short shrift. However, if problems began after a change in the system, system analysis will usually hold the answer. Walking through a typical flow pattern and questioning each step is often helpful. Keep in mind that flow at any division seeks equilibrium; flow will divide such that resistance to flow on one side given its gallonage will exactly equal the resistance to flow on the other side with the remainder of flow. Double checking that the pump in question will, in theory, meet the system conditions may be also be appropriate.

2. In many cases, narrowing down whether a problem is in the system or the pump can be affected by switching pumps in an installation containing one that is working well. If the problem follows the pump, it is in the pump; if it stays with the line, it is in the system. In other cases, noises can be isolated by separating various components of a system to determine if they continue or cease. The motor sometimes can be separated from the pump, for example, or suspect components temporarily removed. Valving can sometimes isolate pumps or piping sections to check the results.

3. There are usually more possible explanations for a problem than a typical manufacturers' lists provides. Following is a sampling of potential causes for a few specific complaints. Creativity is essential in this step, with nothing *ever* taken for granted. Often, a marginal explanation tentatively or casually offered up will be immediately recognized by an engineer and the problem solved.

4. There will always be possibilities (listed in the following section) that are more likely to occur than the others, depending on the symptoms. Of these, the most easily checked should be the first inspected, and then the more difficult. At times there are two simultaneous flaws in a system, making checking of individual items fruitless. In these cases, logic will be of more value than experiment.

Suggestions for a few common problems follow.

Low Flow

Centrifugal Pumps Only

1. Reverse rotation (*Note*: a centrifugal impeller flings the liquid as opposed to scoops it, if there is ever a question as to rotational arrows. A centrifugal pump will pump if turning in reverse, but with low flow and pressure and higher horsepower because of lowered efficiency).

2. More head than originally calculated.
3. Clogged impeller.
4. Wrong impeller diameter for conditions.
5. Cavitation (see earlier section).
6. Impeller clearances worn too large.
7. Motor speed slower than nameplate indicates.
8. Impeller spins loosely on shaft.
9. Broken coupling.
10. Viscosity has increased.
11. Vortexing in source tank.
12. Air in tank introduced from splashing or churning close enough to tank outlet to allow air to enter with stream.
13. Wrong pump for conditions.
14. If pump is above source tank, air leak in suction pipe.

Positive Displacement Pumps Only

1. Excess wear on components.
2. Stuck internal check valves, if any.
3. Torn diaphragms, if present.
4. Stuck internal (or external) pressure relief valve.
5. Speed too slow.
6. Cavitation (see preceding section).
7. Wrong pump for conditions.
8. If belt driven, a slipping belt.
9. Broken coupling.
10. Speed not slow enough for viscosity, leading to cavitation.
11. If pump is above source, air leak in inlet pipe.

Cavitation

Note: Cavitation is indicated by crackling noise for centrifugal pumps, by banging for most PD pumps, or by finding that TDH has fallen at least 3% below the standard curve for the particular centrifugal pump.

1. Blocked inlet.
2. Too much friction loss in suction line.
3. Elbow too close to pump suction port (*Note*: The rule-of-thumb is that the elbow should be located no closer than 6 × the suction port diameter).
4. For centrifugal pumps only, running at extreme left end of the curve causing low pressure areas to develop.
5. Suction line with high spot (as opposed to continually rising), therefore, air trap to disturb or restrict flow.
6. Liquid too hot and therefore vapor pressure too high.
7. Elevation differential too great (too much lift).
8. If drawing from closed tank, vacuum level too high, or vacuum level that builds to an unacceptable point as liquid is pumped out.
9. Excess prerotation at inlet of centrifugal, causing flow to enter in wrong pattern for design (*Note*: Rotation against impeller rotation is a positive factor and will move the curve upward, but the more common prerotation in the same direction as the impeller is harmful and may drop the curve).

Motor Overheats

Centrifugal Pumps Only.

1. Too much flow; therefore, further out on curve than calculated.
2. Wrong rotation.
3. Oversized impeller.
4. Unbalanced impeller.

Positive Displacement Pumps Only.

1. Too much backpressure.
2. Binding or rubbing components.
3. In case of progressive cavity, flexible impeller, or other pump with rubbing components, insufficient motor starting torque to get up to full speed

Any Pumps

1. Voltage above ±10% allowed by manufacturers.
2. Frequency beyond ±5% allowed by manufacturers.
3. Bad motor bearing(s).
4. Three-phase motor not getting power to all three lines (single phasing).
5. Cavitation.
6. Misaligned motor/pump shaft, or misaligned pulleys.
7. Undersized cable to motor.
8. Loose connection at motor outlet box or control box.
9. Insulation has broken down through time and use.
10. Cable too long to handle amperage.
11. Ambient temperature too high.
12. Overgreased bearings, causing binding when bearings heat up.
13. On TEFC design, broken or slipping fan.
14. Speed too high for pump.
15. Not enough air movement to keep motor cool because of confined or poorly vented area.
16. If single phase motor, starting windings do not disengage.
17. Viscosity higher than calculated.
18. Pipe torque, causing misalignment and binding.
19. Motor stops and starts too often (consult manufacturer for maximum stops and starts for applicable horsepower, it is probably lower than you think!).
20. Dirty surfaces not allowing efficient cooling.
21. On drip-proof design, internal diffuser or fans not channeling air flow properly.

Noise

Note: Various designs have various degrees of noise considered standard and acceptable. The manufacturer will be able to supply the numbers if given operating conditions.

1. Motor or pump bearings worn.
2. Cavitation.
3. If a PD pump, components hitting or rubbing.
4. If rigidly piped, amplification of what might be considered normal noise through lines into abnormal amounts.
5. If mounted on channel steel base, amplification of noise from pocket of air beneath motor/pump package.
6. Coupling hitting coupling guard.
7. Motor fan hitting fan cover guard.
8. Check valves chattering as they stick.
9. Water hammer (valves must close slowly enough for momentum in line to be absorbed). A rough rule of thumb is $t = 2L/a$, where t = time in seconds for valve to close (minimum); L = length of pipe via longest route, in feet; a = speed of pressure wave in feet/second (an average for water is 4,000 ft/s). An example would be a 200-ft length of pipe handling water, which would demand a valve closing in no less than $2 \times 200/4{,}000 = 0.1$ s to avoid water hammer.

Abnormal Seal Wear

Note: A mechanical seal in a continuous-duty, clean application can be expected to last several years, as a minimum, to as many as 10 years or longer if under absolutely ideal conditions.

Seat Fracture

1. Thermal shock from dry running followed by sudden introduction of cooler liquid, or any sudden drastic change in the temperature or the liquid. Usually appears as a hairline fracture.
2. Seal head installed incorrectly, causing the head to strike the seat.
3. Physical shock from a blow such as being dropped.
4. Water hammer.
5. Pipe misalignment, twisting the bracket holding the seat.

Carbon Washers Wear Prematurely

1. Abrasives (clearly visible by scored faces).
2. Chemical incompatibility between the chemical and either the washer itself, or the binder holding the carbon together.
3. Improper installation, usually with the seal head cocked.

Elastomer Problems

1. Chemical incompatibility. Elastomer may become brittle, softened, swollen, or even totally eaten away.
2. Dry running. Elastomer may look burned or cracked or become brittle and hard.

ADDITIONAL READING

Hydraulic Institute Standards, Hydraulic Institute.

I. Karassik, W. Krutzsch, W. Fraser, and J. Messina eds. *The Pump Handbook*, 2nd ed., McGraw Hill, New York, 1985.

R. Stover, in B. Lydersen, N. D'Elia, and K. Nelson eds. *Bioprocess Engineering: Systems, Equipment and Facilities*, Wiley, New York, 1994, pp. 253–315.

V. Streeter, *Handbook of Fluid Dynamics*, 1961.

C.R. Westaway and A.W. Loomis, *Cameron Hydraulic Data*, 16th ed., Ingersol Rand, 1984.

PYRUVATE, PRODUCTION USING DEFECTIVE ATPASE ACTIVITY

ATSUSHI YOKOTA
SEIGO AMACHI
FUSAO TOMITA
Hokkaido University
Sapporo, Japan

KEY WORDS

Energy metabolism
Escherichia coli
F_1-ATPase
Futile cycle
Glycolytic pathway
Hemoglobin
Phosphotransferase system
Phosphoglycerate kinase
Pyruvate kinase
Pyruvic acid

OUTLINE

INTRODUCTION

Much work has been done on the development of producers of useful metabolites. Metabolic engineering has enabled more scientific and strategic breeding than the conventional method based on mutagenesis followed by selection. In general, the targets for breeding are biosynthetic pathways of the metabolites, and the best example is the metabolic regulatory fermentation of various kinds of amino acids. In contrast, few studies consider energy metabolism as the target of breeding. However, it seems that manipulation of energy metabolism opens up a new dimension in breeding, because the energy status of the cell deeply concerns both regulation of energy production and biosynthesis and assembly of cell materials, although it does not necessarily directly affect the biosynthetic pathways of the desired metabolites. In this article, the development of producers of useful metabolites by the manipulation of energy metabolism is described with special emphasis on our work, which is improvement of the pyruvic acid productivity of a pyruvic acid–producing mutant of *Escherichia coli* K-12 by a defect in F_1-ATPase activity.

ENHANCEMENT OF CENTRAL METABOLISM BY A DEFECT IN THE ACTIVITY OF F_1F_0-ATP SYNTHASE

This section mainly describes our work, which has demonstrated both the enhancement of glucose metabolism and the improvement of pyruvic acid productivity by the introduction of a defect in an energy-producing reaction in an *E. coli* strain.

Improvement of Pyruvic Acid Production by Strain TBLA-1, an F_1-ATPase-Defective Mutant of *E. coli* W1485*lip2*

For the development of producers of useful metabolites, it seems important to consider points both upstream and downstream of the metabolic flow as well as the biosynthetic pathway itself, such as increases in the supply of precursor metabolites that flow into the biosynthetic pathways and the efficient transport of substrates and metabolites across the membranes. Under these concepts, we have focused on the improvement of fermentative production by enhancement of the glycolytic pathway. That is, it was expected that the enhancement of the glycolytic pathway would lead to the improvement of producers through smooth assimilation of substrate carbohydrates.

As a model system to assess this idea, we have established pyruvic acid production by *E. coli* W1485*lip2* (ATCC 25645) (1), a lipoic acid auxotroph of *E. coli* K-12 (2). This system was employed for several reasons. First of all, it was expected that the effect of enhancement of the glycolytic pathway would be obvious because pyruvate is the end product of glycolysis. Second, *E. coli* strain K-12 is suitable for characterization because much information about its genetics and physiology is available. Strain W1485*lip2* requires lipoic acid for its growth and produces a large amount of pyruvic acid in cultures in which lipoic acid is limiting, where the activity of the pyruvate dehydrogenase complex becomes almost negligible (1).

Then, how can we enhance the glycolytic pathway? Much work has revealed that flow in the glycolytic pathway is enhanced by a decrease in the energy level of the cell. That is, under those conditions, phosphofructokinase (PFK) I, one of the isozymes of *E. coli* PFK, and pyruvate kinase (PYK) II, one of the isozymes of *E. coli* PYK, have been shown to be activated by ADP and AMP, respectively (3,4). Therefore, decrease in the energy level of the cell seemed to be effective for enhancement of the glycolytic pathway. To decrease the energy level, it seemed appropriate to impair the generation of ATP through oxidative phosphorylation. For these reasons, we have introduced a defective gene of the α-subunit of F_1-ATPase, *atpA401* (5), into the strain W1485*lip2* by transduction, and obtained an F_1-ATPase-defective mutant, TBLA-1 (6). Both strains

were cultured aerobically in the pyruvic acid production medium, Medium 7 (6), containing 50 g/L of glucose, using jar fermentors with pH controlled at 6.0. As shown in Figure 1, strain TBLA-1 was found to produce pyruvic acid much more effectively than the parental strain in spite of some decreases both in growth rate and in growth level (6). That is, the parental strain, W1485*lip2*, consumed 50 g/L of glucose and produced 25.5 g/L of pyruvic acid after culture for 32 h; under the same conditions, strain TBLA-1 consumed glucose and produced more than 30 g/L of pyruvic acid after culture for 24 h.

Table 1 shows F_1-ATPase activities, fermentation data, and fermentation parameters of various strains. The fermentation data were obtained after culture for 16 h, at which time maximum growth was attained in the strains listed. As seen in Table 1, the level of growth in strain TBLA-1 decreased to 67% of the parental strain because of impairment of oxidative phosphorylation. However, in strain TBLA-1, both the amount of glucose consumed per cell ([B]/[A] in Table 1) and the amount of pyruvic acid produced per cell ([C]/[A] in Table 1) increased by factors of 1.9 and 2.8, respectively, compared with strain W1485*lip2* (6). In strain TBLA-1, the conversion yield of pyruvic acid from consumed glucose ([C]/[B] in Table 1) also increased by a factor of about 1.5 (6). Furthermore, as all of these changes have not been observed in a revertant from strain TBLA-1 (no. 63-1) having a normal level of F_1-ATPase activity (Table 1), it was concluded that the enhancement of pyruvic acid productivity was brought about by the defect in F_1-ATPase activity (6).

Table 1. F_1-ATPase Activities and Fermentation Profiles of Various Strains

Strain	W1485*lip2*	TBLA-1	No. 63-1
F_1-ATPase activity[a]	6.3	0.09	5.7
Growth (A_{590}) (A)	8.5	5.7	7.2
Glucose consumed (g/L) (B)	32.1	40.4	35.6
Pyruvic acid produced (g/L) (C)	13.5	25.9	14.8
(B)/(A)	3.8	7.1	4.9
(C)/(A)	1.6	4.5	2.1
(C)/(B)	0.42	0.64	0.42

[a] μmol min^{-1} mg protein^{-1}.

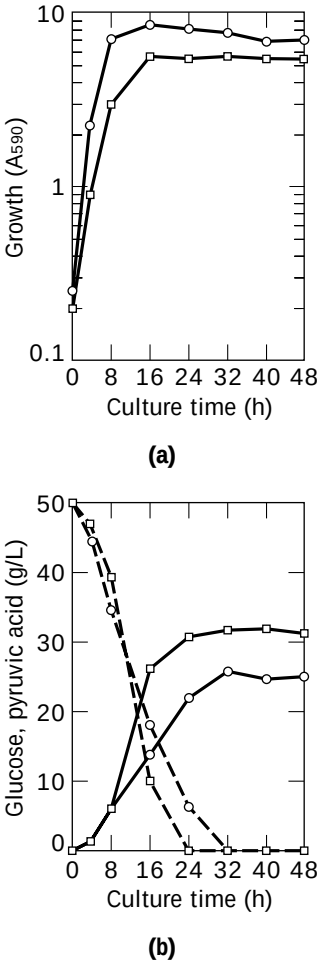


Figure 1. Time courses of pyruvic acid production by strains W1485*lip2* and TBLA-1 (6) showing data for growth (**a**) and pyruvic acid production (**b**). In **b**, *dashed lines* indicate glucose consumption and *solid lines* indicate pyruvic acid production. ○, strain W1485*lip2*; □, strain TBLA-1.

These results seem to be the first example of the application of a mutant defective in energy metabolism in the field of biotechnology. Although we have not optimized culture conditions for pyruvic acid production, the productivity of pyruvic acid in strain TBLA-1 in terms of grams of pyruvic acid produced per liter per hour is quite efficient among the pyruvic acid producers so far reported (7,8). The highest accumulation of pyruvic acid has been reported for the wild-type *Torulopsis glabrata* IFO 0005 (8), which requires nicotinic acid, thiamine, pyridoxine, and biotin for its growth. This yeast strain is capable of producing 57 g/L of pyruvic acid from 100 g/L of glucose in 59-h culture. The commercial importance of pyruvic acid includes its use as a starting material for the synthesis of many drugs and agrochemicals. Especially, pyruvic acid has been used as one of the substrates for the enzymatic production of L-dihydroxyphenylalanine (L-DOPA).

Enhancement of Glucose Metabolism in Strain TBLA-1

In the course of these studies, Jensen and Michelsen (9) reported the physiological characteristics of *E. coli* mutants having deletions in the *atp* operon coding for H^+-ATPase. In the mutants, they demonstrated increases in the oxygen consumption rate and in the level of *b*-type cytochromes by factors of 1.4 and 1.8, respectively, compared with the wild-type strain. It was also found in the *atp* mutant that the flow through the glycolytic pathway increased, and twice as much acetate was produced as a main by-product from consumed glucose. The increased oxygen consumption rate in the *atp* mutants is considered to contribute to the enhanced operation of the glycolytic pathway and to enhanced ATP production by substrate-level phosphorylation through the smooth reoxidation by the electron transport chain of NADH generated from the glycolytic pathway. Also, increase in the content of *b*-type cytochromes seems to contribute to enhanced respiration.

In the cells of strain TBLA-1 growing in pyruvic acid–producing medium, oxygen consumption rates (Fig. 2a), and the content of *b*-type cytochromes were also found to increase by factors of 1.7 and 1.4, respectively, compared with strain W1485*lip2* (10). These changes suggested increases in the amount of some components of the respiratory chain.

In the *E. coli* respiratory chain, two types of NADH dehydrogenases, designated NDH-1 and NDH-2, have been found (11). The expression of the *nuo* operon coding for NDH-1 has been shown to be repressed anaerobically by ArcA, a gene product from *arcA* (*arc* for aerobic respiration control) (12). NDH-2 is specified by the *ndh* gene. The expression of this gene is subject to *fnr* (fumarate nitrate reduction) gene product (Fnr) -mediated anaerobic repression (13). Thus, these two dehydrogenase genes are expressed aerobically and repressed anaerobically. Two types of terminal oxidases, Cyt *o* and Cyt *d*, are known in *E. coli* (14); the former is coded by the *cyo* operon and the latter by the *cyd* operon. The *cyo* operon is expressed aerobically and is repressed anaerobically by both ArcA and Fnr (15,16). On the other hand, the expression of the *cyd* operon has been shown to be regulated by both ArcA and Fnr in such a way that the expression is maximum under microaerobic conditions (15,17,18). To clarify the mechanism(s) for the enhanced respiratory activities observed in strain TBLA-1, it seemed necessary to measure and compare the amounts of the NADH dehydrogenases and terminal oxidases of strains TBLA-1 and W1485*lip2*. This kind of approach might allow us to consider how the energy status of the cells can affect the activities of the oxygen-responding regulators ArcA and Fnr.

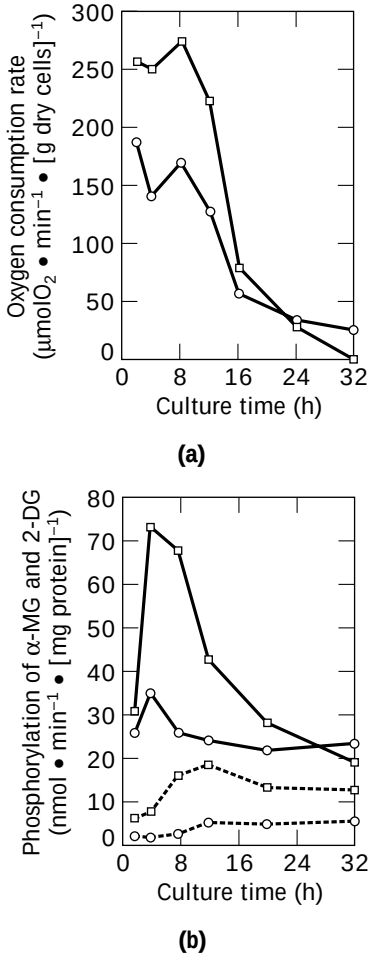


Figure 2. Oxygen consumption rates (**a**) and phosphotransferase (PTS) activities (**b**) of strains W1485*lip2* and TBLA-1 during pyruvic acid production (10) were measured using the cells harvested from the cultures for pyruvic acid production presented in Figure 1. In **b**, *solid lines* and *dashed lines* indicate PEP-dependent phosphorylation of methyl α-D-glucopyranoside (α-MG) and 2-deoxy-D-glucose (2-DG), respectively. ○, strain W1485*lip2*; □, strain TBLA-1.

To elucidate the enhancement of glucose consumption in strain TBLA-1, activities of the phosphotransferase system (PTS) for glucose uptake were measured by phosphoenolpyruvate (PEP) dependent phosphorylation of methyl α-D-glucopyranoside and 2-deoxy-D-glucose using the toluenized cells. We found that these activities were higher in strain TBLA-1 than in strain W1485*lip2* (Fig. 2b) (10). These changes were considered to be important explanations for the higher activities of glucose consumption per cell in strain TBLA-1. PTS not only catalyzes the formation of glucose-6-phosphate on uptake of glucose but also catalyzes the conversion of PEP to pyruvate, seemingly the same reaction as that of pyruvate kinase. Therefore, increase in the activity of PTS found in strain TBLA-1 is considered to be equivalent to the enhancement of glycolysis. Enhancement of glucose consumption has been reported in some *atp* mutants (9,19). However, PTS activities of these mutants under glycolytic conditions have not yet been reported. Therefore, our results are the first implication of the importance of PTS activities in the regulation of glycolysis in *atp* mutants. Furthermore, the increase in PTS activity is considered to lead to increased yield of pyruvate from consumed glucose in strain TBLA-1 because PEP is converted to pyruvate much more effectively with the increased operation of PTS. Rohwer et al. (20) studied PTS activities of a strain of *E. coli* K-12 growing on succinate by measuring initial uptake rates of methyl α-D-glucopyranoside. They found that the uptake rate of the strain decreased when the expression level of the *atp* operon decreased and suggested that the PTS activity decreases under energy-deficient conditions. Thus, it seems quite interesting to examine PTS activities of *atp* mutants in more detail under various culture conditions.

Measurements of the activities of all the glycolytic enzymes revealed that those of phosphoglycerate kinase and PYK-I were higher in strain TBLA-1 than in strain W1485*lip2* during the logarithmic growth phase by factors of about 1.5 at their maxima (10). These differences were considered to be not so distinct as those found in PTS activities. Moreover, so far as the activities of PYK-I were concerned, differences were not observed after the culture had continued for 12 h. However, as these two enzymes are involved in substrate-level phosphorylation, increased activities of these enzymes seem to contribute to the supply of ATP required for the synthesis of cell materials during the logarithmic growth phase. Also, the increases seemed to contribute to increase in the rate of glucose consumption

per cell during the logarithmic growth phase in strain TBLA-1. Because activities of all the glycolytic enzymes were measured with crude enzymes that had been gel-filtered to eliminate low molecular weight materials, the increases in these enzyme activities were not related to activation by low molecular weight ligands such as ADP and AMP.

These interesting changes found so far in strain TBLA-1 are summarized in Figure 3 (21). It became apparent that various kinds of changes that seem to compensate for energy shortage take place in strain TBLA-1, which is defective in F_1-ATPase activity. At present, it is not clear if these changes are indigenous to *atp* mutants. These points need to be investigated using continuous cultures in which the physiological state of each cell is homogeneous and well defined. Investigation of the mechanisms of these changes seems to be quite important and will shed light on the energy metabolism of *E. coli*.

Applicability of a Defect in F_1F_o-ATP Synthase Activity to Other Industrially Important Organisms

It seems interesting to consider whether the enhancement of glucose metabolism and the concomitant improvement of the productivity of a metabolite by the introduction of a gene defective for F_1-ATPase observed in *E. coli* can be applicable to other industrially important microorganisms. To date, other than *E. coli*, physiological changes similar to that observed in *E. coli* mutants (6,9), the enhancement of glucose metabolism and increase in respiratory activity, have also been reported in *atp* mutants of *Bacillus subtilis* (19). It was also demonstrated that the expression of genes coding for some of the respiratory enzymes has also been increased by disruption of the *atp* operon (19). From these observations, it seems that increases in the activities of both rate-limiting enzymes of the glycolytic pathway and respiration for the reoxidation of NADH generated via glycolysis are prerequisite for the enhancement of glucose metabolism. Here, for simplicity, let us consider the possibility of the enhancement of glucose metabolism only from the former aspect, increased activity of rate-limiting enzymes of the glycolytic pathway. Generally in bacteria, PFK and PYK, both of which catalyze the irreversible reactions in glycolysis, have been shown as the control points for the glycolytic flux. In many cases these enzymes are reported to be activated or inhibited allosterically by metabolites of the glycolytic pathway and by adenine nucleotides, and activation is normally observed in response to decrease in the energy level of the cell. Thus, it is expected that a defect in the activity of F_1F_o-ATP synthase lowers the energy level of the cell, thereby leading to the enhancement of glycolysis from the activation of PFK or PYK under those conditions.

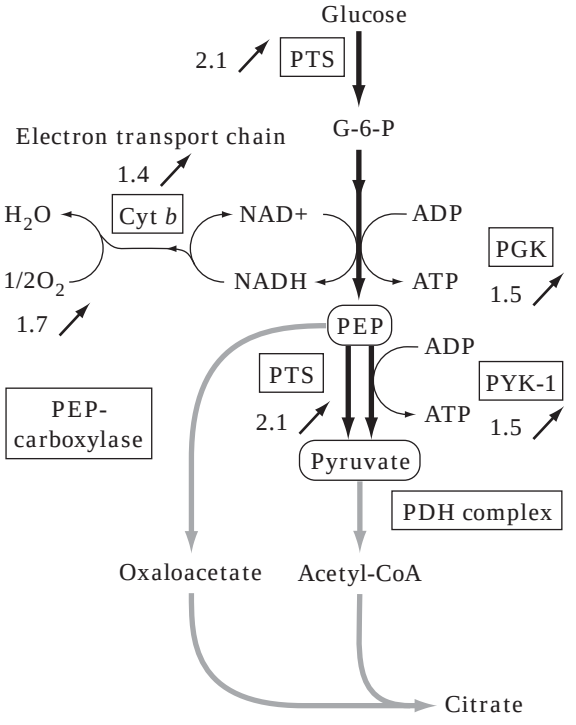


Figure 3. Changes in activities of glucose metabolism in strain TBLA-1 (from 21, with permission of Japan Society for Bioscience, Biotechnology, and Agrochemistry). Each number with an *arrow* shows an increase of the indicated reaction compared with strain W1485*lip2*. The increase for PTS is the average value for PEP-dependent phosphorylation of α-methylglucoside after culture for 4–12 h (Fig. 2b). The increase for oxygen consumption rate is also the average value for 4 to 12 h in Figure 2a. G-6-P, glucose-6-phosphate; Cyt b, b-type cytochromes; PDH, pyruvate dehydrogenase; PGK, phosphoglycerate kinase. Hatched lines indicate weak reactions.

Table 2 summarizes effectors for PFKs and PYKs from some of the industrially important bacteria. As reported (6,9,19), defects in F_1F_o-ATP synthase activity resulted in the enhancement of glucose metabolism in *E. coli* and in *B. subtilis*. Consistently, PFK-1 and PYK-2 of *E. coli* are activated by ADP and AMP, respectively (3,4). In strain TBLA-1 defective in F_1F_o-ATP synthase activity, an increase in the activity of PYK-1 that is different from allosteric activation by fructose-1,6-diphosphate (FDP) (22) was observed (10), suggesting synergistic effect for the increase of PYK activity under glycolysing conditions. On the other hand, the control point of glycolysis has not been established in *B. subtilis* (23). A PFK from *B. stearothermophilus* (24) and PYKs from *B. licheniformis* (25) and *B. stearothermophilus* (26) have been shown to be activated by AMP or ADP, suggesting the possibility of enhancement of glycolytic flux by a defect in F_1F_o-ATP synthase activity. In the case of glutamic acid–producing bacteria, *Brevibacterium flavum* no. 2247 (28,29) and *Corynebacterium glutamicum* (30), the control point for the regulation of glycolysis appeared to be the PYK reaction. A mutant of *B. flavum* no. 2247 defective in ATPase activity has already been isolated in our laboratory. In this mutant enhancement of glucose metabolism has also been observed as in the case of *E. coli* mutants (Yokota et al., unpublished data). Little information is available concerning the regulation of glycolysis in industrially important actinomycetes that produce antibiotics. Purification and characterization of PFK from *Streptomyces coelicolor* A3 (2) have been reported to indicate the enzyme as a regulatory step in the glycolytic pathway (31).

Table 2. Effectors for Phosphofructokinase (PFKs) and Pyruvate Kinase (PYKs) from Industrially Important Bacteria

Organisms	PFK	PYK	References
Escherichia coli K-12	PFK-I Cooperativity (Yes) ADP (+) PEP (−)	PYK-I Cooperativity (Yes) FDP (+) GTP (−) ATP + succinyl-CoA (−)	3,4,22
	PFK-II Cooperativity (No)	PYK-II Cooperativity (Yes) AMP (+)	
Bacillus subtilis 60015	ND	Cooperativity (Yes)	23
Bacillus licheniformis A-5	Cooperativity (No) PEP (−)	Cooperativity (Yes) ADP (+), AMP (+) ATP (−), Pi (−)	25,27
Bacillus stearothermophilus NCA1503 (ATCC7954)	Cooperativity (No) ADP (+) (+), GDP (+) PEP (−)	Cooperativity (Yes) AMP (+), GMP (+) R5P (+) Pi (−)	24,26
Brevibacterium flavum No. 2247 (ATCC14067)	Cooperativity (No)	Cooperativity (Yes) AMP (+) ATP (−), GTP (−)	28,29
Corynebacterium glutamicum M2	ND	Cooperativity (Yes) AMP (+) ATP (−)	30
Streptomyces coelicolor A3(2)	Cooperativity (No) PEP (−)	ND	31
Streptococcus lactis ML3	ND	Cooperativity (Yes) FDP (+) Pi (−)	32

Presence (Yes) or absence (No) of homotropic cooperativity (cooperativity) in the substrate saturation curve for fructose-6-phosphate in the case of PFK and for PEP in the case of PYK. (+), positive effector; (−), negative effector; ND, not determined; FDP, fructose-1,6-diphosphate; Pi, inorganic phosphate; R5P, ribose-5-phosphate; PEP, phosphoenolpyruvate.

The control point for the regulation of glycolytic flux in lactic acid bacteria has extensively been investigated using *Lactococcus* (formerly *Streptococcus*) *lactis* (32). In *L. lactis*, glycolytic flux appeared to be controlled in PYK that has FDP as an activator and inorganic phosphate as an inhibitor. In a glycolysing cell of *L. lactis*, a high intracellular concentration of FDP and a low intracellular concentration of inorganic phosphate have been reported (33), suggesting the presence of an already fully activated form of PYK in the cell. Moreover, in the case of lactic acid bacteria, F_1F_o-ATP synthase has been shown to work in the direction of pumping out protons on hydrolysis of ATP for keeping intracellular pH at neutral under acidic culture conditions. In fact, our work (34) showed that a mutant of *L. lactis* ssp. *lactis* C2 with reduced membrane-bound ATPase activity had almost the same productivity of lactic acid from glucose in a pH-controlled fermentor as that of the parental strain. Thus, glucose consumption per cell did not increase as a result of the mutation. On the other hand, the mutant showed higher acid sensitivity than that of the parental strain, indicating the importance of F_1F_o-ATP synthase in the pH homeostasis of the cell. In acidic culture conditions, higher transcriptional levels of the gene encoding the β-subunit of F_1F_o-ATP synthase and higher amounts of the β-subunit protein than those observed under neutral culture conditions were detected in both strains (35). In contrast, no increase in the transcriptional level of the *atp* operon encoding F_1F_o-ATP synthase was found in *E. coli* in response to acidic culture conditions (36), suggesting the difference of the function of F_1F_o-ATP synthase in different microorganisms. The industrial importance of the mutant defective in the activity of F_1F_o-ATP synthase in the prevention of postacidification of yogurt and fresh cheese has been proposed (34).

From all these examples, mutations in energy metabolism caused by defects in the activity of F_1F_o-ATP synthase are quite interesting for the breeding of industrially important microorganisms. Because of the pleiotropic nature of the mutation, the development of a novel type strain can be expected.

OVEREXPRESSION OF GENES OF ENZYMES INVOLVED IN THE FORMATION OF FUTILE CYCLES

Futile cycles are cyclic metabolic pathways that involve biosynthesis and degradation of certain metabolic intermediates with energy spilling. For example, the coexistence of PFK and fructose bisphosphatase and of PYK and PEP synthase forms futile cycles. The activities of these enzymes are under kinetic control by adenine nucleotides. That is, when the energy level of the cell decreases, glycolytic flow is enhanced, while under energy-sufficient conditions, gluconeogenetic flow is enhanced and thus both enzymes are not activated simultaneously.

Patnaik et al. (37) cloned the PEP synthase gene on a multicopy plasmid, and overexpressed it in *E. coli* K-12 to

form a futile cycle as shown in Figure 4. When the PEP synthase activity was overexpressed 15-fold, an increase in the oxygen consumption rate per cell was observed. Furthermore, an increase in the glucose consumption rate per cell and the excretion of pyruvate and acetate were observed when the activity was overexpressed 30-fold. Patnaik et al. inferred that the increase in PEP concentration due to the overexpression of PEP synthase gene caused the increase in PTS activity, resulting in the enhancement of glucose consumption rate. Similarly, Chao and Liao (38) overexpressed genes of PEP carboxylase and PEP carboxykinase simultaneously in *E. coli* K-12 to form a futile cycle (Fig. 4). Also in this case, with the incremental induction of both enzymes, an increase in the oxygen consumption rate per cell was first observed, and an increase in the glucose consumption rate per cell and excretion of pyruvate and acetate then followed. Chao and Liao concluded that the *arc* system had no relationship to the increase in the oxygen consumption rate.

These physiological changes caused by the overexpression of the genes of the enzymes involved in futile cycles resemble those observed in *atp* mutants (9,19). The operation of futile cycles causes waste of ATP. Therefore, these cells might be suffering from shortage of energy as in *atp* mutants. Although no information is available as to the changes in the activities of PTS and glycolytic enzymes in these cells, it seems interesting to use these futile cycles to improve the fermentative production of useful compounds through enhancement of glucose metabolism.

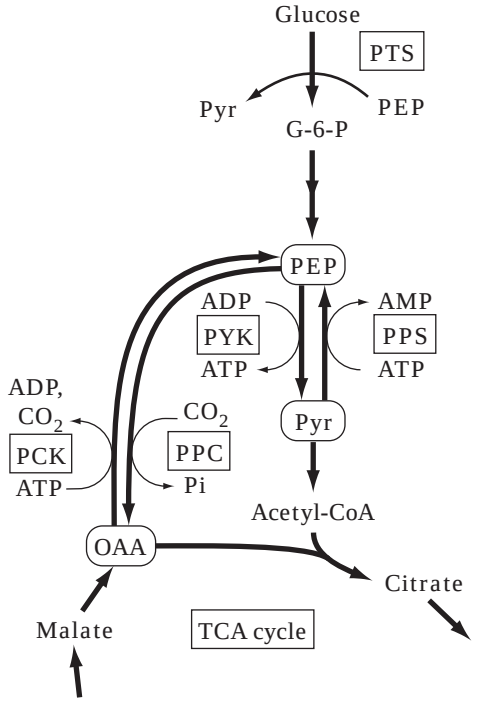


Figure 4. PEP pyruvate and PEP oxaloacetate futile cycles formed in *E. coli* K-12 (from 21, with permission of Japan Society for Bioscience, Biotechnology, and Agrochemistry). Pyr, pyruvate; OAA, oxaloacetate; PPS, PEP synthase; PPC, PEP carboxylase; PCK, PEP carboxykinase, G-6-P, glucose-6-phosphate.

IMPROVEMENT OF PRODUCTION OF USEFUL COMPOUNDS BY THE EXPRESSION OF THE HEMOGLOBIN GENE FROM *VITREOSCILLA*

Finally, in contrast to our concept, we discuss examples in which production of various kinds of useful compounds was improved by increased efficiency of energy production. Bacteria of the genus *Vitreoscilla* are Gram-negative chemoorganotrophs that move by gliding. They are aerobes and use molecular oxygen as the terminal electron acceptor when they live in hypoxic conditions. Under these conditions, a strain of *Vitreoscilla* was found to synthesize a kind of hemoprotein (*Vitreoscilla* hemoglobin, VHb) similar to that of the eukaryotic hemoglobin (39,40). VHb is considered to contribute to the growth of the organism under hypoxic conditions through enhanced diffusion of oxygen into the cytoplasm (41).

When the VHb gene containing its promoter region was introduced into *E. coli* K-12, active VHb protein was produced in response to the decrease in oxygen concentration in the culture medium, as in *Vitreoscilla* cells (42). The regulation of the gene expression was found to be at the transcription level (42). By expression of the VHb gene, various interesting physiological changes were observed in *E. coli* K-12 cells, including increases in growth rate, final growth level, and oxygen consumption rate per cell (43). Moreover, expression of the VHb gene improves various kinds of fermentative production under hypoxic conditions. For example, production of enzymes by *E. coli* (44,45) and by *B. subtilis* (46), production of actinorhodin by *Streptomyces coelicolor* (47), and production of cephalosporin C by *Acremonium chrysogenum* (48) were found to be improved. It should be noted that in contrast to *atp* mutants, enhancement of glucose consumption has not been observed in the *E. coli* K-12 strain expressing the VHb gene (44). Efficient production of these metabolites under hypoxic conditions has much merit in industrial-scale production using large fermentors in which the supply of oxygen frequently becomes a limiting factor.

Figure 5a shows an *E. coli* K-12 cell expressing the VHb gene. It has been assumed that VHb enhances diffusion of oxygen from periplasm to cytoplasm across the cell membrane by reversible association with free oxygen and increases the concentration of effective oxygen ([Oc] + [HbOc]) in the cytoplasm (41,49). In cells of *E. coli* K-12 expressing the VHb gene, a twofold increase in ATP content was found compared with cells without the VHb gene (49). In general, the main terminal oxidase of *E. coli* cells is believed to be Cyt *o* under aerobic conditions; in contrast, under oxygen-limited conditions Cyt *d* becomes the main oxidase. The cells exhibiting increased ATP production with expression of VHb gene were found to possess both types of cytochromes. It has been found that Cyt *d* dose not have proton-pumping activity, although it has a higher affinity to oxygen than Cyt *o*. On the other hand, Cyt *o* has been reported to have proton-pumping activity (50). Based on this information, it has been assumed that in cells without the VHb gene, Cyt *d* works as a main terminal oxidase owing to the low oxygen concentration in the cells (Fig. 5b), whereas in the cells expressing the VHb gene, Cyt *o* as well as Cyt *d* work under the conditions of increased effective

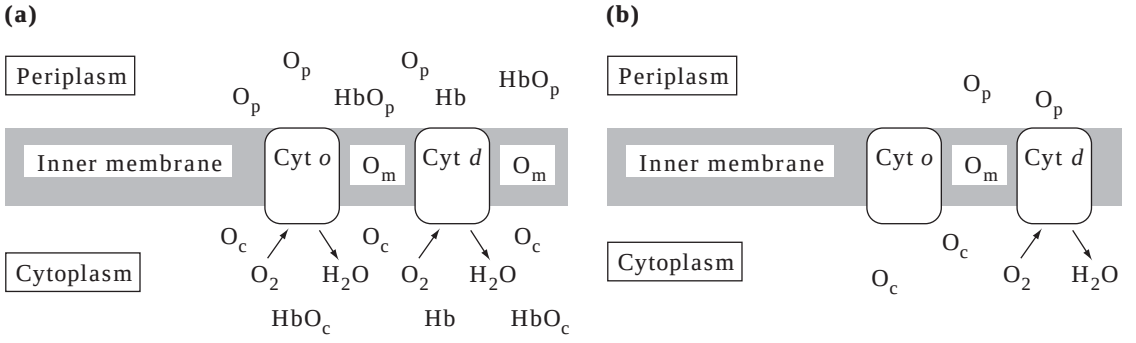


Figure 5. Effect of VHb gene expression on energy metabolism of *E. coli* K-12 (from 21, with permission of Japan Society for Bioscience, Biotechnology, and Agrochemistry). (**a**) Cells expressing VHb gene. (**b**) Cells without VHb gene. O, free oxygen (subscripts p, m, and c denote periplasm, inner membrane, and cytoplasm, respectively); Hb, deoxyhemoglobin; HbO, oxyhemoglobin.

oxygen in the cells (Fig. 5a). Therefore, in the latter case, ATP production by F_1F_o-ATP synthase would be enhanced under the hypoxic conditions caused by the increase in proton motive force. Under these conditions, various kinds of fermentative production would be improved through the enhancement of biosynthetic reactions (49).

CONCLUSIONS

In this article, we have described recent work on the energy metabolism and strain improvement with special emphasis on our work on pyruvic acid production by an *atp* mutant. Improvement of producers of useful metabolites by manipulation in energy metabolism provides not only a new direction in strain improvement but also provides useful information in the basic research of energy metabolism. From the work described, it can be concluded that defects of energy production from the *atp* mutation cause enhancement of glycolysis and thereby are suitable for the improvement of the production of primary metabolites such as pyruvate or amino acids, while the enhancement of energy production attained by the expression of the VHb gene seems to be favorable for the improvement of production of such products as protein and secondary metabolites, which require more energy for biosynthesis.

BIBLIOGRAPHY

1. A. Yokota, H. Shimizu, Y. Terasawa, N. Takaoka, and F. Tomita, *Appl. Microbiol. Biotechnol.* **41**, 638–643 (1994).
2. A.A. Herbert and J.R. Guest, *J. Gen. Microbiol.* **53**, 363–381 (1968).
3. J. Babul, *J. Biol. Chem.* **253**, 4350–4355 (1978).
4. D. Kotlarz, H. Garreau, and H. Buc. *Biochim. Biophys. Acta* **381**, 257–268 (1975).
5. S.D. Dunn, *Biochem. Biophys. Res. Commun.* **82**, 596–602 (1978).
6. A. Yokota, Y. Terasawa, N. Takaoka, H. Shimizu, and F. Tomita, *Biosci. Biotechnol. Biochem.* **58**, 2164–2167 (1994).
7. A. Yokota and S. Takao, *Agric. Biol. Chem.* **48**, 2663–2668 (1984).
8. T. Yonehara and R. Miyata, *J. Ferment. Bioeng.* **78**, 155–159 (1994).
9. P.R. Jensen and O. Michelsen, *J. Bacteriol.* **174**, 7635–7641 (1992).
10. A. Yokota, M. Henmi, N. Takaoka, C. Hayashi, Y. Takezawa, Y. Fukumori, and F. Tomita, *J. Ferment. Bioeng.* **83**, 132–138 (1997).
11. K. Matsushita, T. Ohnishi, and H.R. Kaback, *Biochemistry* **26**, 7732–7737 (1987).
12. J. Bongaerts, S. Zoske, U. Weidner, and G. Unden, *Mol. Microbiol.* **16**, 521–534 (1995).
13. J. Green and J.R. Guest, *Mol. Microbiol.* **12**, 433–444 (1994).
14. L. Thöny-Meyer, *Microbiol. Mol. Biol. Rev.* **61**, 337–376 (1997).
15. S. Iuchi, V. Chepuri, H.-A. Fu, R.B. Gennis, and E.C.C. Lin, *J. Bacteriol.* **172**, 6020–6025 (1990).
16. P.A. Cotter and R.P. Gunsalus, *FEMS Microbiol. Lett.* **91**, 31–36 (1992).
17. H.-A. Fu, S. Iuchi, and E.C.C. Lin, *Mol. Gen. Genet.* **226**, 209–213 (1991).
18. P.A. Cotter, S.B. Melville, J.A. Albrecht, and R.P. Gunsalus, *Mol. Microbiol.* **25**, 605–615 (1997).
19. M. Santana, M.S. Ionescu, A. Vertes, R. Longin, F. Kunst, A. Danchin, and P. Glaser, *J. Bacteriol.* **176**, 6802–6811 (1994).
20. J.M. Rohwer, P.R. Jensen, Y. Shinohara, P.W. Postma, and H.V. Westerhoff, *Eur. J. Biochem.* **235**, 225–230 (1996).
21. A. Yokota, *Nippon Nógeikagaku Kaishi* **71**, 11–16 (1997) (in Japanese).
22. E.B. Waygood and B.D. Sanwal, *J. Biol. Chem.* **249**, 265–274 (1974).
23. M. Diesterhaft and E. Freese, *Biochim. Biophys. Acta* **268**, 373–380 (1972).
24. B.C. Valdez, S.H. Chang, and E.S. Younathan, *Biochem. Biophys. Res. Commun.* **156**, 537–542 (1988).
25. F.W. Tuominen and R.W. Bernlohr, *J. Biol. Chem.* **246**, 1746–1755 (1971).
26. H. Sakai, K. Suzuki, and K. Imahori, *J. Biochem.* (*Tokyo*) **99**, 1157–1167 (1986).
27. C.K. Marschke and R.W. Bernlohr, *Arch. Biochem. Biophys.* **156**, 1–16 (1973).
28. S. Sugimoto and I. Shiio, *Agric. Biol. Chem.* **53**, 1261–1268 (1989).

29. H. Ozaki and I. Shiio, *J. Biochem.* (*Tokyo*) **66**, 297–311 (1969).
30. M.S.M. Jetten, M.E. Gubler, S.H. Lee, and A.J. Sinskey, *Appl. Environ. Microbiol.* **60**, 2501–2507 (1994).
31. A.M.C.R. Alves, G.J.W. Euverink, M.J. Bibb, and L. Dijkhuizen, *Appl. Environ. Microbiol.* **63**, 956–961 (1997).
32. L.B. Collins and T.D. Thomas, *J. Bacteriol.* **120**, 52–58 (1974).
33. J. Thompson and D.A. Torchia, *J. Bacteriol.* **158**, 791–800 (1984).
34. A. Yokota, S. Amachi, S. Ishii, and F. Tomita, *Biosci. Biotechnol. Biochem.* **59**, 2004–2007 (1995).
35. S. Amachi, K. Ishikawa, S. Toyoda, Y. Kagawa, A. Yokota, and F. Tomita, *Biosci. Biotechnol. Biochem.* **62**, 1574–1580 (1998).
36. E. Kasimoglu, S.-J. Park, J. Malek, C.P. Tseng, and R.P. Gunsalus, *J. Bacteriol.* **178**, 5563–5567 (1996).
37. R. Patnaik, W.D. Roof, R.F. Young, and J.C. Liao, *J. Bacteriol.* **174**, 7527–7532 (1992).
38. Y.P. Chao and J.C. Liao, *J. Biol. Chem.* **269**, 5122–5126 (1994).
39. S.J. Boerman and D.A. Webster, *J. Gen. Appl. Microbiol.* **28**, 35–43 (1982).
40. S. Wakabayashi, H. Matsubara, and D.A. Webster, *Nature* **322**, 481–483 (1986).
41. C. Khosla and J.E. Bailey, *J. Mol. Biol.* **210**, 79–89 (1989).
42. C. Khosla and J.E. Bailey, *J. Bacteriol.* **171**, 5995–6004 (1989).
43. C. Khosla and J.E. Bailey, *Nature* **331**, 633–635 (1988).
44. C. Khosla, J.E. Curtis, J. DeModena, U. Rinas, and J.E. Bailey, *Biotechnology* **8**, 849–853 (1990).
45. M. Khosravi, D.A. Webster, and B.C. Stark, *Plasmid* **24**, 190–194 (1990).
46. P.T. Kallio and J.E. Bailey, *Biotechnol. Prog.* **12**, 31–39 (1996).
47. S.K. Magnolo, D.L. Leenutaphong, J.A. DeModena, J.E. Curtis, J.E. Bailey, J.L. Galazzo, and D.E. Hughes, *Biotechnology.* **9**, 473–476 (1991).
48. J.A. DeModena, S. Gutiérrez, J. Velasco, F.J. Fernández, R.A. Fachini, J.L. Galazzo, D.E. Hughes, and J.F. Martin, *Biotechnology* **11**, 926–929 (1993).
49. P.T. Kallio, D.J. Kim, P.S. Tsai, and J.E. Bailey, *Eur. J. Biochem.* **219**, 201–208 (1994).
50. A. Puustinen, M. Finel, T. Haltia, R.B. Gennis, and M. Wikström, *Biochemistry* **30**, 3936–3942 (1991).

REVERSE MICELLES, ENZYMES

FRAGISKOS N. KOLISIS
National Technical University of Athens
Athens, Greece

KEY WORDS

Biocatalysis
Enzyme extraction
Enzymes
Lipases
Reverse micelles

OUTLINE

REVERSE MICELLES

Reverse micelles are fine dispersions of water in organic solvent stabilized by a surfactant molecule. The main property of this system is the coexistence of two completely different microphases in an apparently macrohomogeneous solution. The nature of the surfactant molecules, known as amphiphiles, induces molecular organizates, among which reverse micelles are thermodynamically stable and are formed spontaneously. The solutions are optically isotropic because one solvent is finely dispersed into the other, forming microdroplets that have a diameter on the order of 100 Å.

Formation of Reverse Micelles

The formation of the reverse micelles depends on the relative concentrations of water, surfactant, and organic solvent and can be explained by the energy change caused by dipole–dipole interactions between the polar head groups of the surfactant molecule. This mechanism can be understood by taking into consideration the values of the interfacial tension between water and oil (the term *oil* refers to any hydrophobic solvent). This interfacial tension (with values varying from 30 to 60 dyn/cm) causes a reduction of the surface between the two solvents, which results in their separation. The addition of surfactants induces the formation of micelles because of the reduction of the interfacial tension down to very low values (10^{-5} dyn/cm). The amphiphilic nature of the surfactant molecule causes its self-association in solution. In a reverse micellar solution, the polar head of the surfactant is directed toward the aqueous core of the micelle, and the hydrophobic tail extends into the continuous organic phase (Fig. 1). Depending on the hydrophilic–lipophilic balance (HLB) of the surfactant (i.e, the rate of the weight percentage of hydrophilic groups to the weight percentage of hydrophobic groups in the surfactant molecule), the interface will curve spontaneously toward water or the oil phase, favoring the formation of water-in-oil or oil-in-water structures, respectively. Various types of surfactants have been used as amphiphiles to formulate reverse micellar systems, including natural membrane lipids and artificial surfactants. The most commonly used surfactants are the anionic AOT [bis(2-ethylhexyl) sodium sulfosuccinate]; the cationic CTAB (cetyltrimethyl ammonium bromide); various nonionic surfactants of the Brij, Triton and Tween series; and the zwitterionic biological membrane-forming phospholipids. They can be grouped in three main categories: anionic, cationic, and nonionic, depending on the nature of their polar moiety (Fig. 2).

Cationic micelles are very small ($[H_2O]/[\text{surfactant}] \leq 3$). To increase the micelle size, and thus the amount of water solubilized, a cosurfactant is included in the system, usually an aliphatic alcohol, the cosurfactant being partitionated between the interface and the apolar phase. Anionic surfactants form micellar aggregates of large size ($[H_2O]/[\text{surfactant}] \leq 60$). There is a surfactant concentration under which the reversed micelle formation does not occur the critical micellar concentration (CMC). The CMC depends on the chemical structure of the surfactant, solvent, temperature, and pressure.

The solubilization properties of surfactants are often expressed by a three- or four-component phase diagram, af-

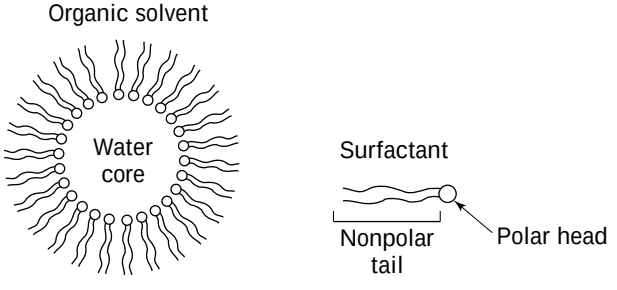
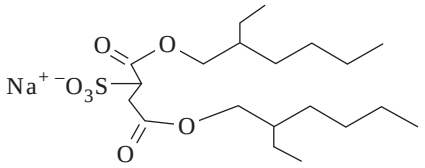


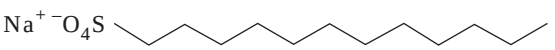
Figure 1. Reverse micelle.

Anionic surfactants

AOT (bis(2-ethylhexyl) sodium sulfosuccinate)

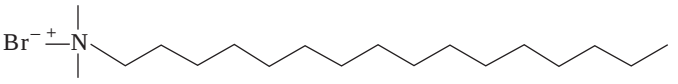


SDS (sodium dodecyl sulfate)

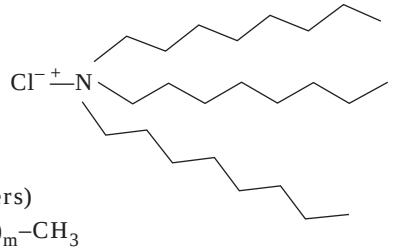


Cationic surfactants

CTAB (cetyltrimethyl ammonium bromide)



TOMAC (trioctylmethyl ammonium chloride)



Nonionic surfactants

Brij (polyoxyethylene glycol ethers)

$$H–(O–CH_2CH_2)_n–(CH_2)_m–CH_3$$

Triton X-100

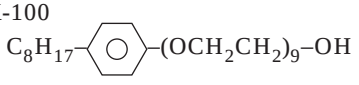


Figure 2. Some surfactants used for reverse micelles.

ter determination of regions of optical transparency. From these diagrams, it is easy to calculate the volume of each component and mix them by microinjection of the aqueous buffer solution into a stirred solution of surfactant dissolved in organic solvent in order to obtain a transparent solution. Figure 3 depicts the phase diagram of AOT/isooctane/water, showing the reverse micelles region.

Micelle Size and Structure

The size, shape, and solubilizing ability of the reverse micelles depend strongly on the nature of the amphiphile molecule. The structure of microemulsions has been studied using various techniques such as NMR, electron paramagnetic resonance (EPR), IR, fluorescence spectroscopy, chemical relaxation, and various scattering techniques such as dynamic light scattering, small-angle X-ray scattering, and small-angle neutron scattering.

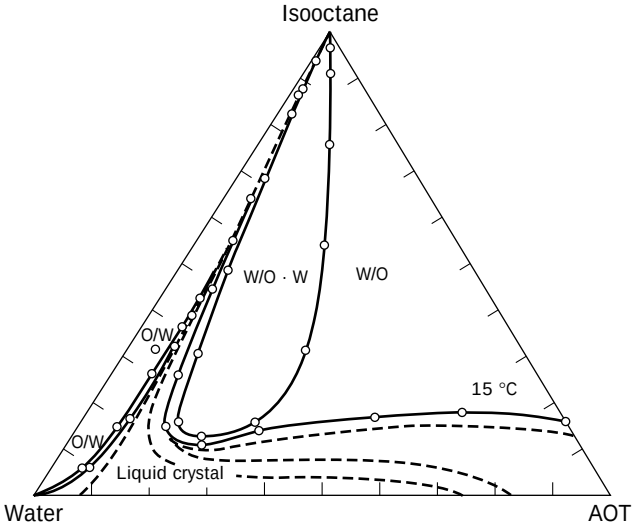


Figure 3. Phase diagram of AOT/isooctane/water, showing the reverse micelles region. O = oil; W = water.

In the majority of the works performed with reverse micelles in biological systems, sodium(bis-2-ethylhexyl)-sulfosuccinate (Aerosol-OT or AOT) has been used as surfactant. This anionic surfactant has the ability to form reverse micelles in a great number of hydrophobic solvents (oils), incorporating considerable amounts of water. The water pools that are formed can have radii from 15 to >100 æ, allowing the solubilization of hydrophilic biomolecules of various sizes.

The structure of the AOT reverse micelles strongly depends on the amount of incorporated water. The water content can be expressed by the molar ratio w_0, with $w_0 = [H_2O]/[AOT]$, reflecting the sizes of the reverse micelles. Many experimental studies using scattering techniques have shown the existence of empirical relations between w_0 and the radius of the reverse micelles (whole microdroplet), for example, $r(nm) = 1.5 + 0.175\ w_0$ (1). Also, for a given w_0 value, the increase of the respective quantities of water and AOT lead to an increase of the number of micelles.

It has been well established from various structural studies that the AOT reverse micelles are spheres formed by a monolayer of AOT molecules surrounding a water core. The water molecules that are located in the close vicinity of the AOT membrane are quite immobilized and possess different properties as compared to bulk water. This is because these water molecules serve to hydrate the polar heads of the surfactant. On the contrary, the water molecules that are situated in the center of the sphere have almost the same properties as bulk water. By increasing

w_0, that is, increasing the amount of water, the overall behavior of the water phase tends to that of normal water.

In general, the physical characteristics of the water in the reverse micelles strongly depends on the value of w_0 and on the nature of the surfactant, and the presence of these differently organized water molecules can be categorized in three different domains in every type of reverse micellar system: *free water*, whose properties and structure become closer to those of bulk water as w_0 increases; *bound water*, whose properties are qualitatively different than those of bulk water because of hydrophilic interactions with the polar head groups of the surfactant; and *surfactant tails*, which penetrate the apolar solvent. When the amount of water is not sufficient to hydrate the surfactant molecules, there is not free water and reverse micelles have only two domains: bound water and surfactant.

Reverse Micelles Dynamics

Reverse micelles are not rigid, impenetrable spheres; they are dynamic entities. In addition to the normal Brownian motions, a number of relaxation processes characterize the monomer, that is, the individual surfactant molecules participating in the micellar aggregate and the water molecules. In contrast to micelles formed in the aqueous solution, the reverse micelles are electrically neutral in their exterior shell, so that no Coulombic repulsions occur, and collisions are frequent. The collision of reverse micelles can be predicted from simple diffusion theory to be on the order of nanoseconds. Not all collisions result in mixing of the aqueous pool. When micelles collide with each other, they may build a transient dimer with a communication channel, which permits the exchange of materials.

Exchange of the contents of reverse micelles is believed to occur by means of a collision–fusion–fission process. The characteristic time for exchange of solubilizates between water pools in AOT reverse micelles has been found to be 10^{-4} to 10^{-6} seconds. The exchange is a function of droplet size and temperature, decreasing with increasing droplet size at a given temperature, and increasing with increasing temperature at a given w_0. The rate of exchange for reverse micelles stabilized by the surfactant alkylarylsulfonate is much slower, of the order of hours, whereas that of CTAB reverse micelles is on the order of microseconds.

pH and Effect of the Ionic Strength in the Reverse Micelles

The pH of the solubilized water in reverse micelles can differ form that of the bulk water, as a consequence of the high concentration of its electric charges caused by the ionization of the polar head groups of the surfactant molecule. For example, in an anionic surfactant reverse micellar system, a decrease the apparent pH of the water within the reverse micelles is observed, and this difference is larger for low w_0.

In studies concerning the role of salts, when salts are present in the aqueous pool of the micelles, it has been found that the ionic strength, the type of salt, and the type of buffer used can affect the size as well as the shape of the reversed micelles.

These two factors, pH and ionic strength, are of particular importance for the encapsulation of enzymes inside reverse micelles in their application in biocatalysis and liquid–liquid extraction of proteins.

ENZYMES, IN REVERSE MICELLES

Enzymes represent novel routes to a number of important products, ranging from β-blockers to synthetic cocoa butter (2). Although they have outstanding catalytic properties, including regioselectivity, stereoselectivity, and enantioselectivity, which can be applied for the preparation of compounds not available by chemical means, their use in organic synthesis is quite limited. The main reason is that most organic compounds of commercial interest are very sparingly soluble in aqueous media used in conventional enzymology. Water is a poor solvent for nearly all applications in industrial chemistry.

Biocatalysts—enzymes, microbial cells, and others—are traditionally applied to aqueous reaction systems, including water-soluble reactants. Since the recent development of biotechnology, much wider applications of biocatalysts have been demanded, for example, bioconversions of biological and xenobiotic compounds having lipophilic or hydrophobic characters. In such bioconversion reaction systems, it is indispensable to introduce organic solvents into the system to improve the poor solubility of the reactants in water. Furthermore, in an attempt to apply biocatalysts for the conversion of petrochemicals, the substrates and products are themselves organic solvents in such biochemical reactions.

The recent realization that enzymes can function also in organic solvents has given the synthetic potential of enzymes a powerful boost for many processes. Other uncommon biocatalysis media are supercritical gases such as CO_2(CO_2SF), liquid membranes, and reverse micelles. The activity and specificity of the enzymes in these uncommon environments are changed. The enantiomeric excesses are normally increased. Several substrates (e.g, lipophilic ones) are better soluble in these media and downstream processing is often easier. Various enzymes (such as alcohol dehydrogenase and esterases, to mention some of them), catalyzed asymmetric redox reactions, stereospecific transesterification, regioselective esterification of sugars, and others have been studied in nonaqueous systems.

Enzymes Inside Reverse Micelles

The use of reverse micelles to solubilize enzymes in organic solvents has attracted considerable interest in the past decade (3–5). A large variety of enzymes, including cofactor-dependent enzyme, have been successfully encapsulated in reverse micellar systems, retaining their catalytic activity. The enzyme molecules can be entrapped in the reverse micelles, avoiding direct contact with the organic medium, potentially limiting their denaturation. The interior of the reverse micelles acts like a microreactor that provides a favorable aqueous microenvironment for enzyme activity as well as an enormous interfacial area through which the conversion of hydrophobic substrates can be catalyzed. Increasing the interfacial area is of great technological interest since this results in the increase of the number of substrate molecules available to react.

When biomolecules such as enzymes are solubilized in reverse micelles, they can be localized in the different microcompartments, depending on their chemical structure or relative solubility in these microphases. The degree of hydrophobicity of the biomolecule plays an important role to the preferential localization among the various microenvironments (water core, bound water, surfactant membrane, and organic solvent). Hydrophilic enzymes are entrapped in the water core, whereas others with an amphiphilic character can be anchored to the surfactant layer, penetrating even in the continuous organic phase (Fig. 4). Recent studies relate the enzyme localization in the different domains of the reverse micelles with their selectivity for various substrates with a different degree of polarity. This hypothesis at least stands for lipase-catalyzed esterification reactions. In addition, this interaction between the enzyme molecule and the interface can explain the different enzyme behavior found in several cases with respect to bulk water.

Solubilization of Enzymes in Reverse Micelles. The most common technique used for the encapsulation of enzymes in the polar core of reverse micelles is the injection method. The protein already solubilized in a concentrated stock aqueous (or buffered) solution is added to the surfactant-containing organic solution upon stirring. With this method, the solubilization is almost instantaneous, and it is possible to control the amount of water present in the system. Another method is to contact the protein solution with the micellar solution upon stirring. Phase transfer will occur, depending on the charge of distribution of the enzyme, surfactant charge, and pH and ionic strength of the aqueous phase. With this method, the enzyme solubilization is slow, but a large amount can be solubilized with minimal values of w_0. A third way of introducing the enzyme is to add the dry enzyme into the empty reverse micelles while aggitating. This technique causes denaturation of the enzyme to a degree. The injection and dry addition methods are commonly used in biocatalytic applications, the latter being suited to hydrophobic proteins. The phase-transfer technique is the basis for extraction of proteins from aqueous solutions. In addition to the method of protein inclusion to the reverse-micelles, other variables that can influence protein solubilization are the size and isoelectric point of the protein, the size of the reverse micelles, the nature of the surfactant used, and the pH and ionic strength of the aqueous solution.

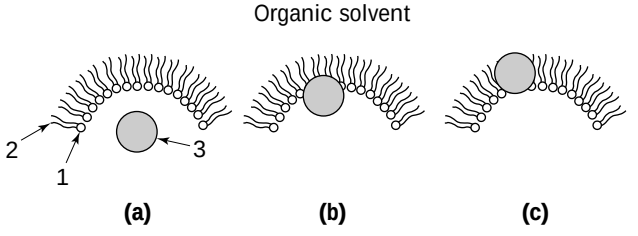


Figure 4. Enzyme localization in reverse micelles depending on the degree of hydrophobicity of the molecule. (**a**) Hydrophilic; (**b**) Amphiphilic, and (**c**) Hydrophobic enzyme. *1*, hydrophilic head of the surfactant; *2*, hydrophobic tail of the surfactant; and *3*, enzyme molecule.

Factors Affecting Enzyme Solubilization. For both the injection and phase contact methods, proteins with a net charge opposite to the surfactant charge and a low ionic strength will favor solubilization. As far as the size of protein molecule is concerned, their solubilization may occur even when the radius of the molecule is larger than the micellar polar core radius. The protein itself probably stabilizes the enlarged micelle, constructing a protective surfactant shell with a certain amount of water. Data support the possibility that there is not a definite maximal molecular size upon which encapsulation in surfactant aggregates does not occur.

The hydrophilic character of large proteins may have a significant role on the formation of reversed micelles. It seems that the hydrophilic proteins will favor the formation of surrounding surfactant shell. This fact could account for the solubilization of large entities, even when the total charge of the macromolecules is identical to the charge of the surfactant heads, thus excluding the formation of micelles via direct electrostatic interactions between the protein and the surfactant layer. This type of interaction within a micellar system can be particularly important for the phase transfer of proteins, applied to the liquid–liquid extraction of proteins and other molecules via reverse micelles. The various parameters affecting this process are discussed in more detail in the section on liquid–liquid extraction.

Enzyme Conformation in Reverse Micelles. Enzyme-containing reverse micellar solutions are optically clear. Thus, spectroscopic techniques, such as ultraviolet (UV), fluorescence, circular dichroism (CD), EPR, and NMR, are able to be used for monitoring the conformation changes that may occur in enzyme molecules upon their solubilization in reverse micelles. Moreover, in this system it possible to precisely control the range of water microenvironments (interfacial/bulk), the size of the micelles by varying the water/surfactant molar ratios, and the role of the chemical structure of the surfactant.

Generally, a more ordered structure has been identified by CD for myelin basic proteins (6) and poly(α-amino acids) (7) in AOT reverse micelles and trypsin and α-chymotrypsin in lecithin reverse micelles (8). UV spectra of enzymes in reverse micelles show little change from aqueous solutions. Fluorescence spectra in reverse micelles are red shifted for some enzymes and blue shifted for others (9,10). An increase in quantum yield has been observed with α-chymothrypsin (10), trypsin (11), pepsin (11), lysozyme (10), and a decrease in quantum yield has been observed for lactate dehydrogenase (12). ^{1}H-NMR studies of some peptides and small proteins show that met-encephalin became more folded, and pancreatic secretory trypsin inhibitor became highly flexible (13). Furthermore, EPR studies have been performed for investigation of structural changes to α-chymotrypsin (14), serum albumin (15), and cutinase (16) upon solubilization in reverse micelles.

The information derived by several studies shows that proteins can be categorized according to the perturbation

of their structure induced by reverse micelles: conformations that are not changed or only slightly changed after incorporation into reverse micelles; conformations that are altered in reverse micelles but unaffected by the amount of water present in the system; and conformations that are altered and that also change as a function of the water/surfactant molar ratio.

Biocatalysis in Reverse Micelles

The first report on enzymes retaining their catalytic activity in reverse micellar systems was in 1978 (2,17,18). Today the number of active enzymes described in reverse micelles has increased to more than 50, including hydrolases, dehydrogenases, and oxidases used in single or multienzyme catalyzed reactions. Enzymatic catalysis in microemulsions has been used for a variety of reactions, such as synthesis of esters, peptides and sugar acetals, transesterifications, various hydrolysis reactions, and steroid transformations. Several applications with potential interest for biotechnological conversions have been developed; some of them are summarized in Table 1.

Enzymatic reactions in reverse micelles are found to obey normal Michaelis-Menten kinetics, but some attention should be given to the way of expressing the kinetic constant K_m. K_m is the substrate concentration where the reaction rate is half the maximum rate. The substrate concentration in a reversed micellar system can be based on the total volume of the system or the water-pool volume of the micelles, but the latter is very difficult to measure. Other factors that have to be taken under consideration is the compartmentalization of the enzyme and substrates between the various microdomains of the reverse micelles and the degree of hydrophobicity of substrate, product, and enzyme molecule, which is related to its localization in the micelles. Several theoretical models have been developed to explain enzyme kinetics in reverse micelles. They are based on different assumptions, such as all droplets in the medium are equal, and there are not any changes in the enzymes kinetic parameters (K_m and V_{max}) during their entrapment in the reverse micelles (the diffusional models). The nondiffusional models are based on the fact that the differences in the kinetic constants with respect to bulk water are caused by conformational changes produced in the reverse micellar microenvironment. Another important parameter obtained from Michaelis-Menten kinetics is the K_{cat}, the turnover number, which is related to maximum reaction velocity. Because K_{cat} is concentration independent, its determination is more straightforward. Turnover numbers for enzymes in reverse micelles are usually dependent on the amount of water present in the system, which is expressed by the molar water-to-surfactant ratio, w_0. This dependence is often a bell-shaped curve, explained by the existence of an optimal inner cavity for catalytic activity, which corresponds to the size of the enzyme. Maximum K_{cat} is sometimes higher than that in aqueous solutions, giving rise to the concept of superactivity (31). This effect can be explained by a greater reactivity of the structured water in the micelle, increased rigidity of the enzyme (32), and enhanced substrate concentration at the enzyme surface (33). Thus, the first parameter that should be considered when studying the enzyme activity in reverse micelles is the amount of the water in the system (w_0). The electric properties of the substrate and surfactant can often have a marked effect on the K_m and K_{cat} constants. When the charges of substrate and surfactant are the same, the differences of K_{cat} and K_m values with those found in water are not significant, but when the charges are different, partition effects are usually observed because of the increase of the substrate concentration near the vicinity of the surfactant layer, hampering their access to the enzyme, which is surrounded by water. α-Chymotrypsin and trypsin encapsulated in AOT and CTAB show a decrease in the second-order rate constant (K_{cat}/K_m). The specificity of the enzymes can also be altered because of the distribution of the substrate between the system phases or interactions with the surfactant layer. An example is the oxidation of aliphatic primary alcohols, catalyzed by alcohol dehydrogenase encapsulated in an AOT/octane system (24). In the reverse micellar system, octanol ($n = 8$) is the optimum substrate, whereas in water the preferred substrate is butanol ($n = 4$) (17). It is profound that these parameters are more accurately determined using the active enzyme concentration, which can be obtained typically by active site titration, for example, based on fluorescence techniques for the active site titration of α-chymotrypsin and lactate dehydrogenase (1). Another parameter to take into account is the pH. When the enzyme is entrapped in reverse micelles, significant shifts of the pH profile usually occur, because of the heterogeneous dis-

Table 1. Examples of Bioconversions in Reversed Micellar Systems

Type of reaction	Enzyme	Reverse micelles
Reduction of decanal	Alcohol dehydrogenase	AOT/isooctane (19)
Polymer synthesis	Horseradish peroxidase	AOT/isooctane (20)
Reduction of ketosteroids	Alcohol dehydrogenase	AOT/isooctane (21)
Oxidation of polyunsaturated fatty acids, linoleic acid	Lipoxidase	AOT/isooctane (22)
Cholesterol oxidation	Cholesterol oxidase	CTAB/octanol/octa ne (23)
Oxidation of aliphatic alcohols	Alcohol dehydrogenase	AOT/octane (24)
Peptide synthesis	α-Chymotrypsin	AOT/isooctane (25)
Hydrolysis of olive oil	Lipase	AOT/isooctane (26)
Inter- and *trans*-esterification	Lipase	AOT/isooctane (27)
Bioluminescent assay	Luciferase	Brij-96/octane (28)
Fatty acid esterification with alcohols	Lipase	AOT/isooctane (29)
Polyol–fatty acid esters	Lipase	AOT/isooctane (30)

tribution of protons inside the polar core of the micelle. Larger shifts of optimal pH occur for small values of w_0. The shifts reported in literature for systems with anionic surfactants are of one to two unities (3).

Lipases in Reverse Micelles

A particular case of enzymatic studies in microemulsions is that of lipases (34). Among the enzymes so far investigated, lipase is one of the most advantageous because it is stable, inexpensive, and widely used in the development of various applications in the detergents, oils and fats, dairy, and pharmaceutical industries. In addition to hydrolysis of triglycerides to glycerol and free fatty acids, lipases can be used in esterification and transesterification reactions in low-water content media. This catalytic process is heterogeneous and can be favored by the use of microemulsion. Lipases are active almost exclusively near interfaces in a classical heterogeneous procedure. There are various reports in the literature concerning studies on the hydrolytic and synthetic action of lipases in different types of microemulsions. Table 2 lists examples of the studies performed so far with lipases from various sources in water-in-oil microemulsions.

Hydrolytic Reactions. Lipase catalyzes the hydrolysis of long-chain aliphatic esters to glycerol and free fatty acids. Several hydrolytic reactions performed by lipases in microemulsion systems have been reported, such as the hydrolysis of triolein and tributyrin by pancreatic lipase in an AOT/octane microemulsion system (35). The same system has been used for the hydrolysis of palm kernel olein by *Rhizopus arrhizus* lipase (36). Hydrolysis of triolein and tributyrin has been reported with the use of *Rhizopus delemar* lipase in AOT/isooctane, CTAB/isooctane, and $C_{12}E_4$/decane systems (37–41). Olive oil hydrolysis by *Candida rugosa* lipase and glycerolysis of triolein by *Chromobacterium viscosum* lipase in AOT microemulsion systems have been performed (42,43). A bioreactor system has been designed that is suitable for continuous glycerolysis of fats and oils, using liposomes as a matrix for adsorbing enzyme and AOT reverse micellar systems, and as a supplier of interfacial area for glycerolysis (45). Palm oil and trimyristine hydrolysis by *Rhizopus* sp. lipases for the production of monoglycerides and diglycerides with a good yield has been also reported in AOT and $C_{12}E_3$ based systems (40,41).

Table 2. Lipase-Catalyzed Reactions in Microemulsions

Lipase source	Microemulsion system
Hydrolytic reactions	
Pancreas	AOT/octane (35)
R. arrhizus	AOT/octane (36)
R. delemar	AOT/isooctane (37)
	CTAB/isoctane (38)
	$C_{12}E_4$/decane (38)
	Lecithin/isooctane (39)
Rhizopus sp.	AOT/isooctane (40)
	$C_{12}E_3$/isooctane (41)
Candida rugosa	AOT/isooctane (42)
Chromobacterium viscosum	AOT/isooctane (45)
Candida cylinracea	Lecithin/cyclooctane (44)
Synthetic reactions	
Rhizopus delemar	Phosphatidylcholine/hexane (27)
	$C_{12}E_3$/nonane (47)
	$C_{12}E_4$/decane (49)
	AOT/isooctane (30)
	AOT/hexane (47)
Rhizopus arrhizus	AOT/isooctane (50)
Chromobacterium viscosum	AOT/heptane (46)
Candida cylindracea	Brij-35/olive oil (48)
	AOT/isooctane (30)
Penicillium simplicissimum	AOT/isooctane (50)

Several parameters play an important role in lipase-catalyzed hydrolytic reactions in microemulsion systems. The reverse micellar structural characteristics, such as the size, shape, and solubilizing ability, which depend strongly on the nature of the polar moiety of the amphiphile used, strongly affect the enzyme activity. Anionic, cationic, and nonionic types of surfactants have been used in various hydrolytic reactions. Among the different systems studied, the anionic surfactants seem to be more suitable for the hydrolytic activity of lipase, as compared to the cationic ones, whereas nonionic systems do not seem adequate for this type of reaction.

Other important parameters affecting lipase activity that have been studied in detail by several groups are pH, temperature, water content (w_0), fatty acid chain length of the triglycerides used as substrates, and enzyme stability. In studies concerning the hydrolysis of triolein by *Rhizopus delemar* lipase, the pH optima were reported to be 6.5 in the anionic systems and 5.8 in the cationic systems, considering as pH values the values of the buffered enzyme solutions used for the preparation of the various microemulsion systems. The pH optimum for the same reaction performed in classic heterogeneous biphasic systems was found to be 5.6. This difference is obviously related to the fact that the pH values within the water-dispersed phase do not necessarily correspond to that in the bulk water. The difference observed between AOT and CTAB systems is possibly the result of the different amounts of bound water and the nature of the ions present in the water pools of each system. In the case of AOT systems, it has been reported that 8 to 10 water molecules are indispensable for hydrating each Na^+ and sulphonate group, whereas for CTAB, only 1 to 3 water molecules per Br^- and ammonium head are enough. In other studies on AOT-formulated microemulsion systems, a pH optimum of 7.1 was reported for lipase from *Candida rugosa* 7.0 for lipase from *Rhizopus* sp.; and a maximum in the range 6.0–7.5 for lipase from *Chromobacterium viscosum*. Temperature also plays an important role in lipase activity in microemulsions; in the case of anionic systems, the optimum temperature is reported to be 30 °C; in cationic systems, 22.5 °C. However, at temperatures lower than 15 °C and higher than 40 °C, the structure of these systems may be considerably altered.

Water content in microemulsions appears to be the most important factor that determines the reaction rates. The molar ratio w_0 ($w_0 = [H_2O]/[\text{surfactant}]$) values are used

to express the water content of the system. Several reports concerning the activity of lipase from various sources as a function of w_0 shows a maximum in the range of 9 to 11. The same pattern is also observed for other enzymes studied in the same microemulsion systems, such as chymotrypsin and pyrophosphatase. However, when reverse micelles based on lecithin were used, the maximum reaction rate was observed at an extremely low water content ($w_0 = 2.2$).

In studies concerning the effect of the fatty acid chain length as well as the type of microemulsion system on the substrate conversion rate, it was found that for microbial origin lipases, triolein was a slightly better substrate as compared to tributyrin, but pancreatic lipase cleaves triolein 15 times more efficiently than tributyrin. Also, the enzyme in anionic microemulsion systems shows a higher affinity as compared to the cationic systems, whereas nonionic systems do not seem adequate for hydrolytic reactions.

One serious problem for the application of lipase to triglyceride hydrolysis is the stability of the enzyme in the microemulsion systems. The size of the microemulsion microdroplet seems to play an important role on the enzyme stability. It has been reported that lipase stability is improved as w_0 decreases. Also, the presence of various additives, such as $CaCl_2$ and casein, considerably increases the stability of lipase in AOT microemulsion systems.

Synthetic Reactions. The hydrolytic reaction catalyzed by lipase is reversible because the enzyme can also catalyze the synthesis and transesterification of triglycerides, under certain conditions. This catalytic process is of great technological interest and may be favored by the use of microemulsions. The lipophilic substrates and products can be solubilized in the continuous organic phase, whereas lipase remains active in the water core of the microemulsion droplets. In this case, specific synthetic reactions can be catalyzed, under mild conditions. The synthesis of triacylglycerol by *Rhizopus delemar* lipase has been reported in a phosphatidylcholine reverse micellar system (27). The esterification of glycerol with oleic acid, catalyzed by *Chromobacterium viscosum* lipase in glycerol-in-oil–type microemulsions of AOT in heptane, and the synthesis of various fatty acid esters in water-in-oil microemulsions, stabilized by CTAB, have also been reported (46). Similar systems have been used for the transesterification of palm oil with stearic acid (47). Synthesis of heptyloleate and the esterification of various triglycerides by *Candida cylindracea* lipase in a microemulsion-like system has been reported using Brij-35 as surfactant (48). The esterification of various alcohols with fatty acids by *Rhizopus delemar* lipase has been performed using nonionic microemulsion systems stabilized by $C_{12}E_4$, with significantly fast reaction rates also reported (49). In this work, it was shown that pH ranging from 3.45 to 8.55 does not affect the reaction rates in this type of microemulsion system. Also, studying the effect of the system's water content showed that when $w_0 < 5$ the conversion is considerably slow, but for $w_0 > 7$ a 90% conversion is attained in 4 h. Esterification reactions of alcohols with fatty acids and synthesis of polyols have been reported (30) using *Rhizopus delemar* and *Candida cylindracea* lipase in AOT microemulsion systems. Recent studies have been reported on the activity of lipases from *Rhizopus delemar, Rhizopus arrhizus*, and *Penicillium simplicissimum* in esterification reactions of various aliphatic alcohols with fatty acids in microemulsions formulated by AOT in isooctane (50). The results showed a remarkable selectivity of the lipases regarding the chain length and the structure of the substrates. *P. simplicissimum* lipase presented higher reaction rates in the esterification of long-chain alcohols as well as secondary alcohols. Primary alcohols had a low reaction rate, and tertiary alcohols had a very slow rate of esterification. Long-chain fatty acids were better catalyzed as compared to the shorter ones. *R. delemar* and *R. arrhizus* lipases showed a preference for the esterification of short-chain primary alcohols, but the secondary alcohols had a low rate of esterification and the tertiary could not be catalyzed. Medium-chain fatty acids were also better catalyzed than the long ones. Fluorescence quenching measurements suggest that there are not any significant changes of the shape and size of the microemulsions caused by the substrates, which could have been related to the enzyme selectivity. In extensive studies using a large number of substrates with different degrees of polarity as well as spectroscopic techniques, this selectivity appears to be related to the localization of the enzyme molecule within the micellar microstructure, because of the hydrophobic/hydrophilic character of the protein and not of a specificity of the enzyme itself. This hypothesis can further be supported if one takes into consideration the structural differences of these lipases.

Kinetic Studies. Lipase-catalyzed hydrolysis of triglycerides in reverse micelles follows Michaelis-Menten kinetics. The release of fatty acids as a function of reaction time at different enzyme concentrations and fixed substrate concentrations (20–40%, v/v, where $[S] \gg K_m$) is proportional to the enzyme concentration. The K_m and V_{max} parameters of the Michaelis-Menten kinetic equation can be determined from the Lineweaver-Burk double-reciprocal plot of the lipase activity only as apparent ones. To find a true K_m, the substrate concentration at the enzyme surface must be used. This requires knowledge of the substrate partitioning between various microphases, as the water pool, the interface, and the bulk organic solvent. The partition coefficient is difficult to be determined experimentally. Another constant obtained from Michaelis-Menten kinetics is the turnover number, K_{cat}, which is related to the maximum reaction velocity. Determination of K_{cat} depends on accurate measurement of active-enzyme concentration in the microemulsion system. Usually, the total enzyme concentration and not the active-enzyme concentration is used for the calculation of K_{cat}. The latter can be achieved by active site titration studies.

Kinetic studies of the esterification reaction on the reaction of ethylene glycol and fatty acids (30) and of lauric acid and (-)menthol (29), catalyzed by lipase in AOT/isooctane/water microemulsion systems, have been described. In the second case, the authors have shown that the kinetics of this synthesis follow a ping-pong Bi-Bi mechanism. According to this mechanism, lipase reacts

with lauric acid to form the lipase–lauric acid complex. The complex is then transformed to a carboxylic–lipase intermediate and water is released. This is followed by an attack of (-)menthol to this intermediate, forming the (-)menthyl laureate. The values of all apparent kinetic parameters were determined, and a model concerning this type of reaction involving two hydrophobic substrates was proposed. According to this, the partitioning of the two hydrophobic substrates between the continuous organic phase and the reverse micelles must be taken into account to determine the real kinetic constants.

Lipase Applications in Reverse Micelles

The potential applications of lipases in biotechnological processes have been the subject of increased scientific, economic, and industrial interest and efforts in recent years. Regioselectivity, stereospecificity, substrate specificity, and low energy consumption are some of the features that make lipase-catalyzed processes an alternative to conventional nonenzymatic reactions. Special attention is given to the transformation of fats and oils as well as the preparation of synthetically useful enantiomers. Among the systems under investigation, the encapsulation of lipases in reverse micelles seems to present several advantages such as greater interfacial area, lower mass transfer limitations, and enhancement of catalytic activity.

The development of a bioreactor that enables reactions in a semibatch or continuous mode allowing the separation of substrates, products, and enzymes is a critical demand in reversed micellar technology today. Recently, two types of reactors have been developed for the glycerolysis of olive oil by lipases. In the first case, the bioreactor, abbreviated as LRM system, uses liposomes for adsorbing the enzyme and reversed micelles as a good supplier of interfacial area for the glycerolysis reaction (45). According to this system, the lipase was adsorbed on the liposomes prepared in a simple manner, and reverse micelles containing the immobilized enzyme were formed by injecting appropriate amounts of glycerol and predetermined amounts of water in AOT/isooctane system. Olive oil in isooctane is fed at controlled flow rates into the bioreactor by pressurizing the reservoir with nitrogen gas. The reaction takes place at 37 °C. The products are obtained from the outlet in a single phase, but the microemulsions are not broken. A microporous membrane is used to separate products formed (oleic acid, monoolein, and diolein) and unreacted substrate (triolein) from the microdroplets containing immobilized lipase. The membrane, made of polysulfone, is about 12.5 μm thick, and the micropores have a maximum diameter of 25 nm. This system presents a long operational stability. The second bioreactor system is a membrane bioreactor for simultaneous lipolysis and bioproducts separation in a reversed micellar system (51). The main purpose of this type of bioreactor is to retain lipase behind the membrane where reaction occurs and to recover the products on the permeate side. The membrane is a ceramic one, resistant to organic solvents, made of a zirconium oxide layer over a porous carbon support with a 10,000 molecular weight cutoff. The inside diameter is 6 mm; the wall thickness, 2 mm; and the effective length, 20 cm. The internal permeation area is 38 cm^2. The reaction mixture is recirculated from the thermostated vessel into the reactor by means of a gear pump flowing tangentially to the membrane. Part of the solution is permeated through the wall of the semipermeable membrane (by connective flow) as a result of the hydrostatic pressure difference, and it is recycled to the reactor. All connecting piping is solvent resistant.

Enzymatic catalysis in reverse micelles has also been performed mostly in batch-type reactors in various other processes. Some examples are the synthesis of peptides by α-chymotrypsin in reverse micelles in a hollow fiber reactor (25), or the production of tryptophan by tryptophanase in a continuous reversed micellar reactor (52).

Liquid–Liquid Extraction of Enzymes Using Reverse Micelles

The application of reverse micelles in the extraction and isolation of enzymes is a promising technique at the downstream biotechnological processes, having interesting advantages over existing processes for large-scale recovery of extracellular enzymes (53). The traditional process of transfer proteins into organic solvents often causes irreversible denaturation of proteins and loss of biological function. A reverse micellar extraction combines the potential for concentration and purification of the enzymes through transfer between aqueous and organic micellar phases. Enzymes can be extracted without loss of activity. Examples of enzyme extraction using reverse micellar systems are summarized in Table 3.

The recovery of enzymes from the aqueous phase involves two steps: forward and back extraction. In the first step, the desired enzyme is transferred selectively from the aqueous phase to the reversed micellar phase, and in the second step, the enzyme is subsequently recovered from the reversed micellar phase by extraction with a second aqueous phase, the micellar phase being recovered for further extractions. The rates of forward transfer from the aqueous to the micellar phase are extremely fast. Upon steering, for example, only a few seconds are necessary for the transfer of cytochrome *c* into a system of AOT/isooctane. The process of back extraction into the new aqueous phase is much slower and requires constant agitation. In ideal conditions, the back extraction of cytochrome *c* can be achieved in several minutes. The distribution of an enzyme between the reversed micellar and the aqueous phase is influenced by several factors such as pH, ionic strength, and surfactant type (anionic, cationic, or nonionic). The pH determines the ionization state of the surface-charged groups of the protein molecule. The solubilization is favored when attractive electrostatic interactions occur between the protein molecule and the surfactant head groups; in this case, the overall charge of the surfactant is opposite to the overall charge of the protein. This implies that, for cationic surfactants, solubilization of the protein in reverse micelles is favored at pH values above the isoelectric point (pl) of the protein, whereas the opposite is true for anionic surfactants. Together with the pH, the ionic strength plays a very important role in protein recovery. The presence of salts in the aqueous phase can alter the equilibrium size of the surfactant aggregates.

Table 3. Examples of Enzyme Extraction Using Reversed Micelles

Enzyme	Reversed micellar system
Trypsin (54)	AOT/isooctane, octane, heptane
Extracellular alkaline protease (55)	AOT/isooctane
α-Amylase (56)	TOMAC/isooctane
Separation of a mixture of cytochrome *c*, ribonuclease, lysozyme (57)	AOT/isooctane
Lysozyme, α-chymotrypsin, pepsin (58)	AOT/isooctane
Isocitrate dehydrogenase, β-hydroxybutyrate dehydrogenase, glucose-6-phosphate dehydrogenase (59)	CTAB/octane/exanol
α-Chymotrypsin (60,61)	NR_4^+/cycloexane
	AOT/octane/water/glycerol
Ribonuclease A (62)	AOT/isooctane

A higher ionic strength will decrease the size of the micelle, and this effect can be stronger for larger ions, which causes shorter electrostatic interactions. Thus, it is possible to achieve a selective transfer based on size exclusion effects. The attraction between the surfactant and the protein is weaker, and this contributes to exclude the protein from the micellar core. When the ionic strength is very low, phase transfer does not occur. The transfer of proteins from an aqueous to a micellar phase requires a minimum of ionic strength in the aqueous phase. In addition to the surfactant charge, other surfactant-dependent parameters have to be taken under consideration, such as the size of the reversed micelles that are formed, the energy required to enlarge the reverse micelles, and the charge density on the inner surface of the reverse micelles.

The liquid–liquid extraction technique in general is well known, and apparatus and scale-up rules have been established for numerous applications, including the use in recovery processes for antibiotics and organic acids from fermentation broths. The forward and back extractions can be performed in a continuous mode, with the reverse micellar phase circulating between the two extraction units. No specialized equipment is required. The process developed for conventional liquid–liquid extraction is, in principle, also suitable for this application. The most important types of equipment are mixer/settler, agitated column, centrifugal, and membrane extractors. At the present time, process development has centered on the use of mixer/settlers and membrane extractors.

CONCLUDING REMARKS

So far, the results derived from studies with reversed micelles show that the use of these systems can reach a considerable biotechnological relevance and be extended to a wide range of applications. In the field of bioconversion, the enzyme solubilization in reversed micelles clearly offers advantages over other techniques, including the very fast reaction rates or the small quantity of enzyme to be used in comparison to other approaches of enzymology in extreme conditions. However, there are a few problems that must be solved in order to take full advantage of this system, such as the need to remove the excess surfactant from the products and to develop appropriate reactors, which will allow the continuous production and in situ recovery of the bioproducts applying new separation techniques, for example, membrane polymers. Another promising field for reverse micellar systems application is in the downstream processing for the separation of other bioproducts besides proteins and peptides, such as amino acids, nucleotides, nucleic acids, solvents, organic acids, antibiotics and steroids.

BIBLIOGRAPHY

1. J.D. Nicholson and J.H.R. Clarke, in K.L. Mittal and B. Lidman eds., *Surfactants in Solution*, vol. 3, Plenum, New York, 1984, p. 1663.
2. H.W. Blanch, in F.N. Kolisis, B. Macris, and D. Kekos, eds., *Biocatalysis in Organic Solvents*, EEC Advanced Workshops on Biotechnology, 1992.
3. M.J.M. Castro and J.M.S. Cabral, *Biotech. Adv.* **6**, 151–167 (1988).
4. M. Waks, *Proteins: Struct. Funct. Genet.* **1**, 4–15 (1986).
5. P.L. Luisi and R. Wolf, in Mittal and Fedler eds., *Solution Behavior of Surfactants*, vol. 2, Plenum, New York, 1982, pp. 887–905.
6. C. Nicot, M. Vachet, M. Vincent, J. Gallay, and M. Waks, *Biochemistry* **24**, 7024–7032 (1985).
7. M. Seno, H. Noritomi, Y. Kuroyanagi, K. Iwamamoto, and G. Ebert, *Colloid Polym. Sci.* **262**, 727–733 (1984).
8. O. Peng and P.L. Luisi, *Eur. J. Biochem.* **188**, 471–480 (1990).
9. S. Barbaric and P.L. Luisi, *J. Am. Chem. Soc.* **103**, 4239–4244 (1981).
10. C. Grandi, R.E. Smith, and P.L. Luisi, *J. Biol. Chem.* **256**, 837–843 (1981).
11. P.L. Luisi, F. Henninger, M. Joppich, A. Dossena, and G. Casnati, *Biochem. Biophys. Res. Comm.* **74**, 1384–1389 (1977).
12. G.B. Strambini and M. Gonnelli, *J. Phys. Chem.* **92**, 2850–2853 (1988).
13. A. de Marco, L. Zetta, E. Menegatti, and P.L. Luisi, *J. Biochem. Biophys. Methods* **12**, 335–347 (1986).
14. P. Marzola, C. Forte, C. Pincino, and C.A. Veracini, *FEBS Lett.* **289** (1991).
15. L.V. Belonova, R.M. Davydov, and V.P. Timofeev, *Russ. J. Phys. Chem.* **57**, 1670–1673 (1983).
16. V. Papadimitriou, A. Xenakis, C.T. Cazianis, H. Stamatis, M. Egmont, and F.N. Kolisis, *Ann. N.Y. Acad. Sci.* **799**, 275–280 (1996).
17. A. Sanchez-Ferrer and F. Garcia-Ramona, *Enzyme Microb. Technol.* **16**, 409–415 (1994).

18. K. Martinek, I.V. Berezin, Y.L. Khmelnitski, N.L. Klyachko, and A.V. Levashov, *Biocatalysis* **1**, 9–15 (1987).
19. K. Martinek and A.N. Semenov, *Biochim. Biophys. Acta* **659**, 90–101 (1981).
20. A.M. Rao, V.T. John, R.D. Gonzalez, J.A. Akkara, D.L. Kaplan, *Biotechnol. Bioeng.* **41**, 531–540 (1993).
21. R. Hillhorst, C. Laane, and C. Veeger, *FEBS Lett.* **159**, 225–228 (1983).
22. J.F.G. Vliegenhart, G.A. Veldink, and J. Boldingh, *Agric. Food Chem.* **27**, 623–626 (1979).
23. C. Laane and J. Spruijt, in J.M.C. Duarte ed., *Recent Developments in Biotechnology*, Plenum, pp. 142–150.
24. K. Martinek, A.V. Levashov, Y.L. Khmelnissky, N.L. Klyachko, and I.V. Berezin, *Science* **218**, 889–891 (1982).
25. P. Luthi and P.L. Luisi, *J. Chem. Soc.* **106**, 7285–7286 (1984).
26. D. Han and J.S. Rhee, *Biotechnol. Lett.* **7**, 615–621 (1985).
27. S. Morita, H. Narita, T. Matoba, and M. Kito, *J. Am. Oil Chem. Soc.* **61**, 1571–1574 (1984).
28. E.I. Belyaeva, L. Brovko, N.N. Ugarova, N.L. Klyachko, A.V. Levashov, K. Martinek, and I.V. Berezin, *Doklady Akad. Nauk. SSSR* **273**, 494–497 (1983).
29. H. Stamatis, A. Xenakis, U. Menge, and F.N. Kolisis, *Biotechnol. Bioeng.* **42**, 103–110 (1993).
30. D.G. Hayes and E. Gulari, *Biotechnol. Bioeng.* **40**, 110–118 (1992).
31. N.L. Klyachko, A.V. Levashov, and K. Martinek, *Molekulyarnaya Biologiya (Russ)* **18**, 1019–1032 (1984).
32. K. Martinek, I.V. Berezin, Y.L. Khmelnissky, N.L. Klyachko, and A.V. Levashov, *Coll. Czech. Chem. Commun.* **52**, 2589–2602 (1987).
33. E. Ruckenstein and P. Karpe, *J. Colloid Inter. Sci.* **139**, 408–436 (1990).
34. A. Ballesteros, U. Bornscheuer, A. Capewell, D. Combes, J.S. Condoret, K. Koening, F.N. Kolisis, A. Marty, U. Menge, T. Scheper, H. Stamatis, A. Xenakis, *Biocatalysis* **13**, 1–42 (1995).
35. E.A. Malakhova, B.I. Kurganov, A.V. Levashov, I.V. Berezin, and K. Martinek, *Doklady Akad. Nauk. SSSR* **270**, 474–477 (1983).
36. T. Kim and K. Chung, *Enzyme Microb. Technol.* **11**, 528–532 (1989).
37. A. Xenakis, T.P. Valis, and F.N. Kolisis, *Progr. Colloid Polym. Sci.* **79**, 88–93 (1989).
38. T.P. Valis, A. Xenakis, and F.N. Kolisis, *Biocatalysis* **6**, 267–279 (1992).
39. R.K. Schmidli and P.L. Luisi, *Biocatalysis* **3**, 367–376 (1990).
40. K. Holmberg and E. Osterberg, *Progr. Colloid Polym. Sci.* **82**, 181–189 (1990).
41. K. Holmberg and E. Osterberg, *J. Am. Oil Chem. Soc.* **65**, 1544–1548 (1988).
42. S.W. Tsai, C.L. Chiang,*Biotechnol. Bioeng.* **38**, 206–211 (1991).
43. K. Holmberg, B. Lassen, and M.B. Stark, *J. Am. Oil Chem. Soc.* **66**, 1796–1800 (1989).
44. R. Scartazzini and P.L. Luisi, *Biocatalysis* **3**, 377–380 (1990).
45. P.S. Chang, J.S. Rhee, and J.J. Kim, *Biotechnol. Bioeng.* **38**, 1159–1165 (1991).
46. P.D.I. Fletcher, R.B. Freedman, B.H. Robinson, G.D. Rees, and R. Schomacker, *Biochim. Biophys. Acta* **912**, 278–282 (1987).
47. K. Holmberg and E. Osterberg, *Progr. Colloid Polym. Sci.* **74**, 98–102 (1987).
48. M. Bello, D. Thomas, and M.D. Legoy, *Biochem. Biophys. Res. Commun.* **146**, 361–367 (1987).
49. F.N. Kolisis, T.P. Valis, and A. Xenakis, *Ann. N. Y. Acad. Sci.* **613**, 674–680 (1990).
50. H. Stamatis, A. Xenakis, M. Provelegiou, and F.N. Kolisis, *Biotechnol. Bioeng.* **42**, 931–937 (1993).
51. D.M.F. Prazeres, F.A.P. Garcia, and J.M.S. Cabral, *Biotechnol. Bioeng.* **41**, 761–770 (1993).
52. D.K. Eggers and H.W. Bloch, *Bioproc. Eng.* **3**, 83–91 (1988).
53. M. Dekker, R. Hilhorst, and C. Laane, *Anal. Biochem.* **178**, 217–226 (1989).
54. Q. Chang and J. Chen, *Biotechnol. Bioeng.* **46**, 172–174 (1995).
55. R.S. Rahaman, J.Y. Chee, J.M.S. Cabral, and T.A. Hatton, *Biotechnol. Prog.* **4**, 218–225 (1988).

RHEOLOGY OF FILAMENTOUS MICROORGANISMS, SUBMERGED CULTURE

S.M. HEYDARIAN
A.P. ISON
P. AYAZI SHAMLOU
University College London
London, United Kingdom

KEY WORDS

Culture viscosity
Filamentous microorganisms
Mixer-viscometers
Morphology-viscosity
Non-Newtonian flow
Rheology
Rheometers

OUTLINE

INTRODUCTION

Many biotechnology products are manufactured using microorganisms and mammalian cells that behave in a Newtonian manner. In other cases, the product of the culture or the cellular material itself may cause the culture to be

highly non-Newtonian. It is the processing of non-Newtonian materials that can cause major design and operational problems. For example, in most filamentous fermentations in submerged cultures, at the start of fermentation the broth rheology is Newtonian, not too different from water. However, often the microbiological requirements of the fermentation are such that the prevailing operating conditions lead to the development of diverse and complex non-Newtonian broth rheology, and extreme changes in the apparent viscosity of the broth can occur during fermentation (1–3). These changes can cause practical difficulties in maintaining an adequate power input per unit volume in the bioreactor. Subsequently, significant heterogeneity in dissolved oxygen tension and shear rate distribution can occur in the vessel contents (4–6). Consequently, transport processes such as heat and mass transfer and metabolic activity of the cells can suffer, cell damage can occur, and a fall in productivity can result (7–9). The extent of non-Newtonian behavior of the broth is closely related to the growth of the biomass and the development of a three-dimensional network of highly branched hyphae in the case of mycelia microorganisms (10–12). Considerable work has been done in recent years to understand these complex structures by measuring the mycelia morphology in dispersed (filamentous) and pelleted (aggregated) forms (13–17). However, there are as yet no general predictive constitutive equations that can be used confidently for design and scale-up, and even for the same strain there is no consensus of opinion among researchers on the precise effects of mycelia morphology on broth rheology or how the rheology of such systems should be described. The uncertainties associated with the use of predictive rheological equations are to a large extent understandable considering the wide range of parameters that are thought to influence the rheology of the cultures. These include the mode of fermentation (e.g., batch, fed-batch, or continuous), growth rate, level of the dissolved oxygen concentration and pH levels of the culture, level of mixing and shear intensity in the fermentor, and presence of solid debris, impurities, surfactants, and additives. These parameters critically affect the properties of the individual hyphae, for example, their length and length distribution, diameter, and mechanical integrity (strength, rigidity, and flexibility), as well as the physicochemical interactions between the hyphae (compactness, roughness) and between the hyphae and the continuous liquid phase (3,15,18–22). The analysis of rheological data is further complicated by the degree of hyphae aggregation (i.e., formation of clumps) and by the flocculation of clumps (16). The superimposed effects of all these parameters on broth rheology can be difficult to simulate.

Few publications are available on the rheology of non-filamentous cultures. In cases where data are published (e.g., for yeast suspensions), there is considerable disagreement between the results of different researchers (11,23–25). Fermentation cultures and biological solutions that do not exhibit mycelia growth but can show non-Newtonian behavior include high concentrations of single-cell suspensions (e.g., yeasts), high molecular weight fermentation products (e.g., polysaccharides), solutions containing globular proteins, supercoiled plasmid DNA and macromolecules (e.g., membrane protein complexes), chromosomal DNA, and animal cells in solutions in high concentrations (11,26).

This review article focuses largely on the rheology of the microbial suspensions in order to examine some of the key factors that are perceived to affect their flow properties. The intention is to highlight the results of some of the recent research on the interactions between mycelia morphology and culture rheology without emphasizing the effect of rheology on, for example, transport processes such as heat and mass transfer. For these aspects, the reader is referred to other relevant reviews on the subject (4–6, 27–29).

VISCOMETRY

In addition to biomass concentration, the extent of non-Newtonian behavior of fermentation broths is determined largely by the primary properties of the hyphae, which include hyphal morphology and dimensions, hyphal flexibility, and ability to form clumps. The primary properties of the hyphae in turn are critically dependent on the detailed fermentation conditions, making it difficult to develop a satisfactory theoretical simulation of the non-Newtonian behavior of broths.

An alternative engineering approach to the study of flow properties of a fermentation culture is to experimentally develop the flow curve, or shear diagram, which is a plot of shear stress, τ, against shear rate, γ, in the laminar region (Fig. 1). This information is usually obtained in a conventional laboratory viscometer such as the cup-and-

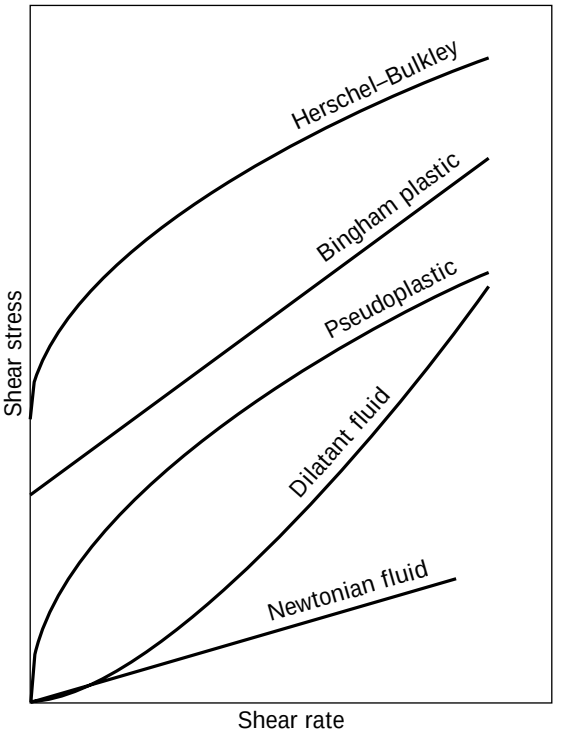


Figure 1. Rheological models.

bob (coaxial cylinder), capillary tube, or a cone-and-plate viscometer. However, most cultures are highly heterogeneous and contain large amounts of insoluble debris and particulate material, which can be fragile and can separate and sediment during rheological measurements. This renders many of the conventional rheometers unsuitable for characterization of biological suspensions. In such cases useful rheological data can be generated indirectly in laminar flow by using a calibrated, scaled-down, "engineering" viscometer, for example, an instrumented pipe flow viscometer or impeller viscometer (16–34). The former is widely used in the biochemical industry for flow characterization and is essentially a cup-and-bob viscometer except that the rotating inner cylinder is replaced by a mixer, (e.g., a turbine, paddle, disc, or helical ribbon) (Fig. 2). From the rheological viewpoint, the basic difficulty in using engineering viscometers is that these instruments do not provide a well-defined and uniform shear field, and therefore, individual elements of fluid are likely to undergo periods of exposure to different levels of shear during measurements. In the case of the impeller viscometer, the evaluation of an average shear rate in the cup is based on an empirical technique originally developed for mechanically agitated vessels by Metzner and Otto (35) and later modified for relatively simple rotating units (e.g., a disc spindle) by Mitschka (36). These techniques have been fully described elsewhere (31,37) and are therefore only briefly reviewed here. The estimation of the average shear rate, γ_{av}, in the cup of an impeller viscometer is based on the assumption that the average shear rate, γ_{av}, is directly proportional to the rotational speed of the impeller, N; the constant of proportionality, k_s, which is generally known as the impeller shear rate constant, is uniquely dependent on the geometry of the impeller viscometer but is independent of the fluid rheology. Thus, $\gamma_{av} = k_s N$. Figure 3 summarizes the experimental procedure used in evaluating the shear rate constant, k_s, for an impeller (viscometer) based on the Metzner and Otto technique. The key steps in the determination of k_s are as follows.

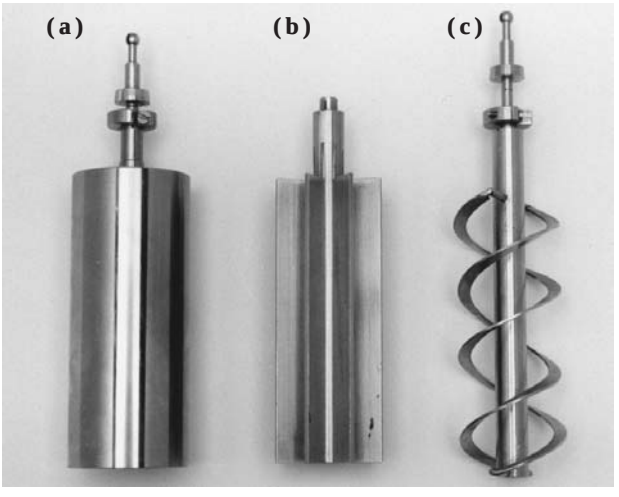


Figure 2. Mixer for rheometers. (**a**) solid cylinder; (**b**) turbine; (**c**) helical ribbon.

1. The laminar Newtonian power curve is obtained for the impeller-cup-and-bob viscometer from a series of measurements of torque and speed using one or more Newtonian liquids.
2. The flow curve is developed for a (model) non-Newtonian liquid, for example, carboxymethycellulose (CMC) solution using a conventional cup-and-bob viscometer.
3. The cup of the impeller viscometer is filled with the (model) non-Newtonian liquid.
4. The impeller speed is set at a predetermined value, N_1, and the corresponding torque, T_1, is measured. These values are used to calculate the corresponding dimensionless power number, $(P_o)_1$, $[P_o = P/(\rho N^3 D^5)]$.
5. From the Newtonian power curve, the Reynolds number, $(Re)_1$ corresponding to $(P_o)_1$ is obtained.
6. Using the definition of the general Reynolds number, $(Re = \rho N D^2/\mu_a)$ and the value of the Re_1 from step 5, the apparent viscosity, $(\mu_a)_1$, is calculated.
7. From the non-Newtonian flow curve the value of the shear rate, $(\gamma_a)_1$, is obtained corresponding to $(\mu_a)_1$.
8. Steps 4 through 7 are repeated for a range of speeds (N_2, N_3, etc.), and these values are plotted against the corresponding values of shear rates ($(\gamma_a)_2$, $(\gamma_a)_3$, etc.).

The plot obtained in step 8 is often a straight line through the origin and has a slope equal to the shear rate constant, k_s. It should be noted that the above procedure is valid only if the measurements of speed and torque (power) fall in the laminar flow region of the power curve, this being the region where the power number is a function of the Reynolds number. For most turbine impellers, the Metzner and Otto technique gives a value of the shear rate constant, k_s, in the range between 10 and 13 (35), whereas for helical ribbon and large paddle impellers its value falls in the range of 26 to 30 (38).

An important practical consideration is the selection of an appropriate rheometer (e.g., a rotating viscometer or a capillary flow rheometer). Different methods of measuring the rheological parameters of a non-Newtonian liquid do not necessarily yield the same result. This is not to say that one method is better than another but merely that it reflects the fact that the liquid flow field is different in the different devices, and the response of a non-Newtonian liquid depends on the flow field to which it is exposed. Therefore, in specifying a value for the apparent viscosity, it is important to quote the means by which the parameter has been measured. Moreover, the method of measurement should be closely related to the application of the final product. For example, for the design of a mechanically agitated bioreactor, an appropriate viscometer is the cup-and-bob or the turbine impeller viscometer, whereas the capillary tube or the instrumented pipeflow viscometer can be used effectively to gather rheological data for pipeline transportation of non-Newtonian biological suspensions. The following discussion is based on rheological data obtained in the former group of viscometers.

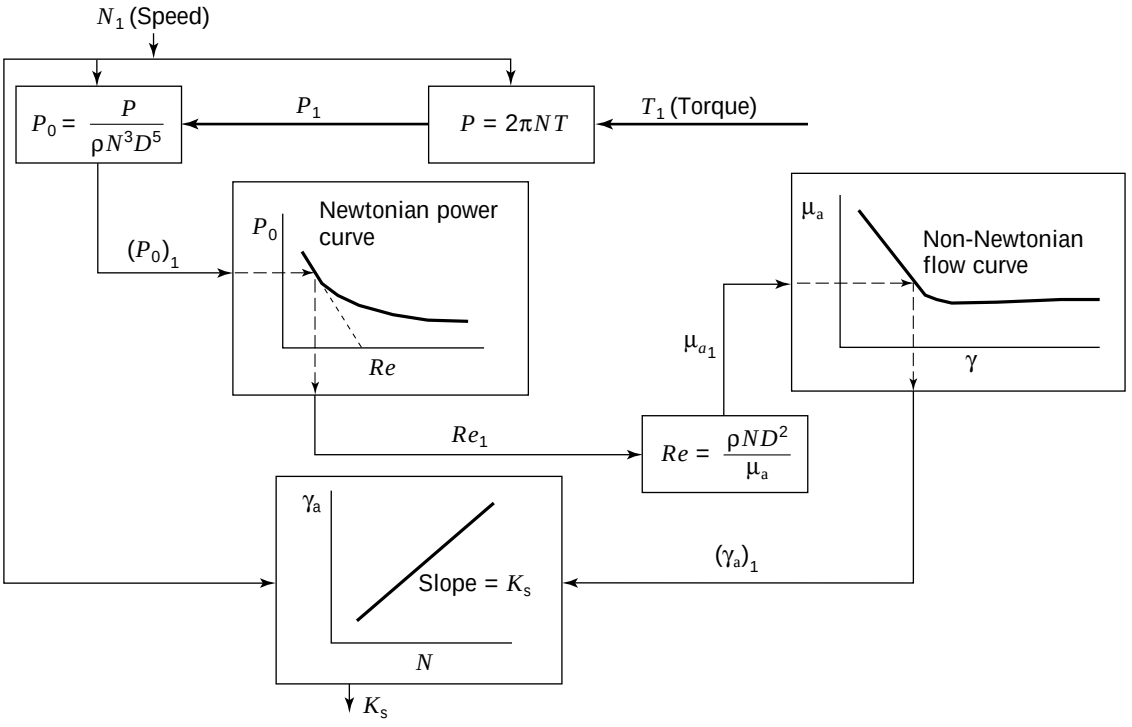


Figure 3. Experimental procedures for calculation of K_s from Metzner and Otto equation.

Commercial Viscometers

Bohlin, Haake, Brookfield, Rheomat, Carrimed, and Rheomatrics are just a few of the more commonly used devices used for the rheological characterization of liquids (39). The first four viscometers are available in both laboratory and portable (battery-operated) form, and some, such as the Brookfield, can also be used on-line. The state of technology has vastly improved over the past two decades, and current instruments such as the Bohlin can be fitted with different heads (e.g., coaxial cylinders, parallel plates, or cone-and-plate geometry) and can perform a variety of measurements including controlled stress and controlled strain measurements. Additionally, the Bohlin rheometer can accurately determine the response of the liquid to more complex types of motion imposed on the sample, including oscillatory motion from which rheological information can be obtained on the viscoelastic properties of the liquid.

These rheometers can be used to develop the non-Newtonian flow curve directly over a wide range of shear rates (10^{-6}–10^{7} s^{-1}) of interest to most process engineering operations (39). The low shear rates represent processes involving, for example, sedimentation of fine particles (10^{-6}–10^{-4} s^{-1}), leveling due to surface tension (10^{-2}–10^{-1}), and gravity drainage (10^{-1}–10^{1}). The medium range of shear rates covers processes such as extrusion, chewing and swallowing (10^{0}–10^{2} s^{-1}), mixing in mechanically agitated vessels, and pipe flow (10^{1}–10^{3} s^{-1}). Shear rates of 10^{3} s^{-1} to 10^{7} s^{-1} occur in operations such as spraying, centrifugation, and high-pressure homogenization.

RHEOLOGICAL MODELS

For design purposes it is often desirable to describe mathematically the information contained in a flow curve by fitting empirical equations to experimental measurements. Numerous flow equations have been proposed to describe the shape of the flow curves, including the Bingham plastic, Herschel–Bulkley, Casson, and power law equations. There is, however, no theoretical justification in recommending any of these equations a priori, and the selection of an appropriate flow equation should be based on practical considerations; the model equation should give an accurate fit of the experimental flow curve, it should have a minimum number of independent constants that are readily evaluated, and preferably the constants should have some physical basis.

A few researchers have reported time-dependent rheology (6,40,41) and viscoelastic behavior (42) for filamentous cultures. However, in the majority of cases the rheology of the cultures appears to exhibit time-independent shear thinning behavior, which can be described adequately by a two-parameter power law equation (31,43,44). This is characterized by a flow curve that, over a reasonable range of shear rate, is a straight line on a logarithmic plot of shear stress against shear rate (Fig. 4). Thus the relation between shear stress, τ, and shear rate, γ, can be written as $\tau = K\gamma n$, in which K is the consistency index and n is known as the flow behavior index and has a value less than unity for most cultures. The value of the flow behavior index gives a measure of the extent of the non-Newtonian behavior of the broth. For $n = 1.0$, the culture has Newtonian flow properties, with $K = \mu$. As n deviates from 1.0, the broth becomes increasingly non-Newtonian.

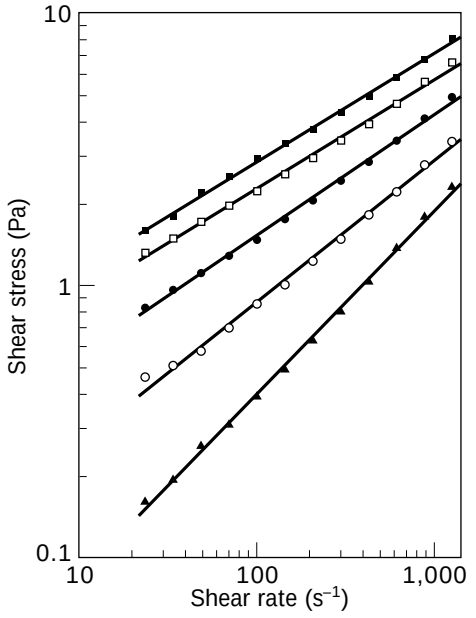


Figure 4. Shear stress against shear rate in log-log scale. Samples from *S. erythraea* (CA 340), fermentation in 7-L stirred bioreactor. Biomass concentration of the samples was between 3 and 12 g L^{-1}. Details of the fermentation areas in Ref. 45.

Flow equations have also been recommended for fermentation cultures based on the Bingham plastic, Casson, and Herschel–Bulkley models, all of which contain a yield stress parameter that needs to be evaluated from experimental measurements (16,42,46). The flow curves of many biological suspensions rise rapidly and then flatten out at relatively low shear rate. Under these conditions it is difficult to differentiate between the power law model and a model containing a yield stress, and very accurate data at quite low shear rates would be required to prove or disprove the existence of "true" yield stress (42,46). In the absence of such information, it seems prudent to use the power law model to describe the rheological data of biological materials over a specified range of shear rate.

RHEOLOGY OF MICROBIAL CULTURES

Filamentous cultures exhibit two extreme types of morphology in submerged culture, highly dense pelleted and free mycelia, with loosely packed mycelium forming a transition between the two types (Fig. 5). The freely dispersed mycelia and its entanglement to form clumps can cause the fermentation to become highly viscous and non-Newtonian. Past studies on morphology have attempted to distinguish between the different types by using various morphological parameters including compactness, roughness, and fractal dimension. These studies show that the compactness of *Penicillium chrysogenum* varies between about 0.3 for highly filamentous structure to close to unity for pellets (16). The fractal of *P. chrysogenum, Streptomyces griseus,* and *Streptomyces tendae* have been measured by Patankar et al. (47) and in our laboratory (45) and seem to vary between 1.9 and 2.7 in filamentous form and between 1.45 and 2.00 in pelleted form (47).

The pelleted type of growth consists of compact discrete masses of hyphae that generally lead to Newtonian behavior of the broth (48). However, in many cases as the size of the pellets increases, the diffusion of nutrients, particularly O_2, into the center of the pellets becomes limited, resulting in low productivity (6,49). Moreover, there are reported cases of cultures consisting of pelleted mycelium that exhibit clear non-Newtonian behavior (50). Great caution must therefore be exercised in generalizing the flow behavior of microbial cultures.

The flow curves presented later are intended to serve as typical rheological examples of real biological microorganisms extensively studied in the past. The discussion examines the effects of some of the important parameters that are perceived to influence broth rheology. For these reasons, the details of fermentation conditions are not included, but where necessary, the reader is referred to other publications and references where further information on the relevant aspects of the work can be found. Additionally, with the exception of the information given in Table 1, a conscious decision has been made to avoid recommending specific constitutive equations because the absolute values of the constitutive parameters K and n are highly system specific and ideally should always be based on experimental flow curves obtained under the actual fermentation conditions for which they are required.

Filamentous Mold Cultures

Numerous publications are available on the rheology of mold fermentation cultures; mostly the mycelium is in the filamentous form because these cultures show strong non-Newtonian behavior. The organisms most widely studied are *P. chrysogenum* and *Aspergillus niger*. Table 1 gives details on the range of parameters used in the rheological characterization of these materials. The non-Newtonian behavior of most of the cultures investigated has been modeled either in terms of a flow behavior index, n, and consistency coefficient, K, of a power law shear thinning equation, or in terms of a yield stress, τ_0, and a plastic viscosity, μ_p, of the Bingham plastic model. In a few cases the Casson model is used, in which case the yield stress, τ_0, and the shear rate coefficient, K_c, are given as the two model parameters. In these studies, plastic viscosity, μ_p, and yield stress, τ_0, in the Bingham plastic model and the consistency coefficient, K, in the power law equation were found to increase and the flow behaviour index, n, decrease with increasing biomass concentration. The increase was found to be nonlinear, and its extent was highly strain dependent (Table 1).

An examination of Table 1 shows that equations are also available in which the rheological parameters have been related to the morphological properties of the individual hyphae. However, recent measurements have shown that in many cases as much as 90% of the mycelium exists in flocculated (clump) form and not in a freely dispersed form, suggesting that rheological parameters are more likely to be related to clump properties and not to the de-

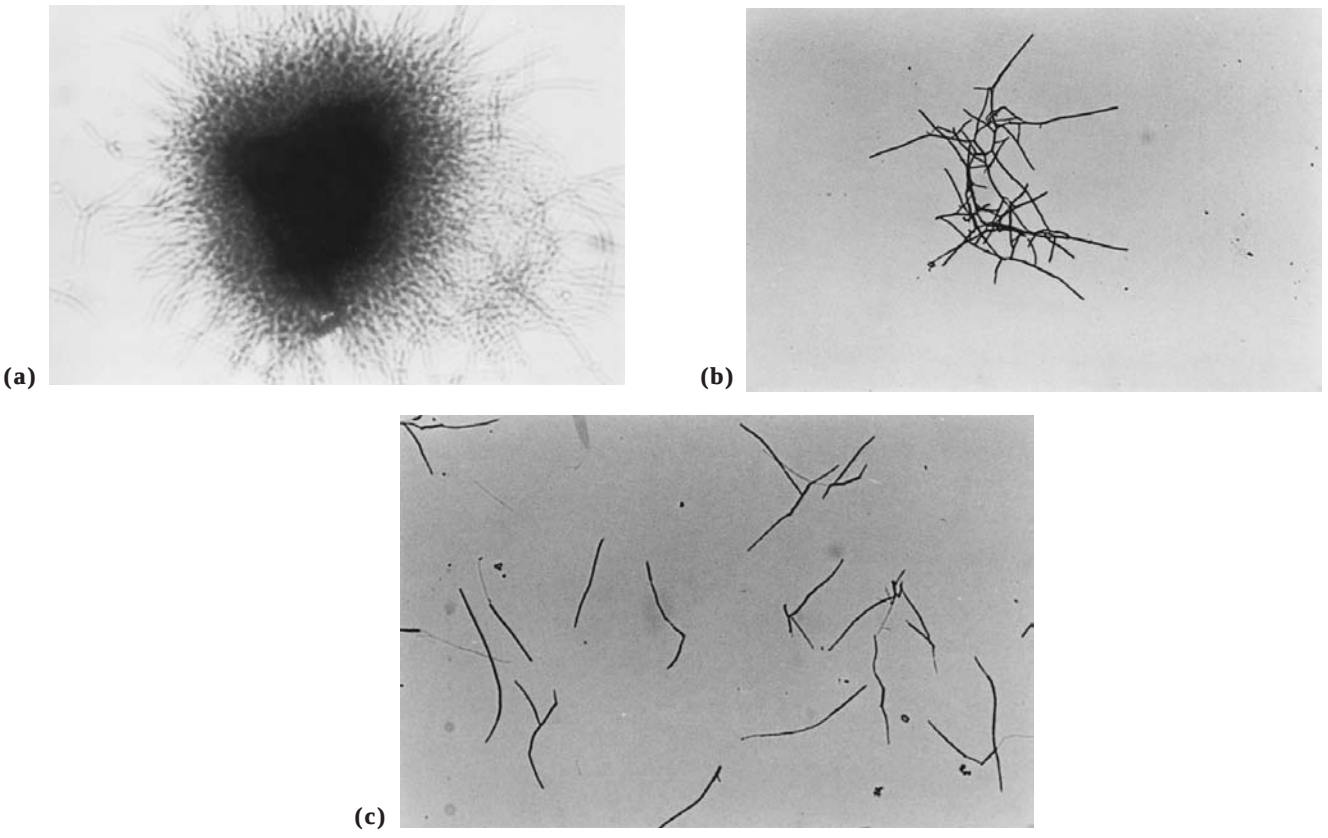


Figure 5. Example of *S. erythraea* (CA 340) morphology. (**a**) Pellet; (**b**) clump; (**c**) dispersed mycelium.

tailed characteristics of the individual hyphae. Recent measurements carried out by Tucker and Thomas (56) and Olsvik et al. (16) appear to support this conclusion and suggest that, for example, the compactness and the roughness of hyphae flocs can be successfully related to the rheological parameter (Table 1).

Filamentous Actinomycetes Cultures

In filamentous form, the sizes of the individual hyphae in the mycelium are generally less than 100 μm, and the broths exhibit considerable non-Newtonian flow behavior (31,46). However, recent experimental evidence obtained in our laboratory using three industrial strains of Actinomyces ranging in size from 15 μm to 25 μm indicated that the properties of the individual hyphae have little effect on broth rheology (3,44). The results of these experiments, discussed in more detail later, also revealed that there was little change in the size of the mean main length of the hyphae, and furthermore very few clumps were observed during hyphal characterization.

Figure 6 shows the variations of n and K of *Saccharopolyspora erythraea* (strain CA 340) mycelia cultures as a function of time of fermentation (45). The data were obtained in a standard coaxial cylinder viscometer using fresh samples of the culture that were removed during fermentation in a 7-L stirred bioreactor (45). The plots shown in Figure 6 are effectively snap shots of a developing rheology, showing the dramatic changes in the non-Newtonian behavior of the fermentation broth during the growth period.

Similar rheological data are shown in Figures 7, 8, and 9, respectively, for the fermentation of industrial strains of *S. erythraea*, *Aspergillus roseorufa*, and *Streptomyces rimosus* (3,44). The data for *S. rimosus* were obtained from an industrial-scale mechanically agitated bioreactor; in the case of *S. erythraea* and *A. roseorufa*, samples were taken during fermentation in a small, laboratory-scale mechanically stirred unit. In all three cases, the broths contained a significant amount of particulate matter, which made the use of a standard cup-and-bob viscometer inappropriate. The rheological parameters plotted in Figures 7–9 were therefore obtained by using a Brookfield 2HA-TDVII Synchro-Lectric rheometer fitted with a disc-spindle impeller, for which the average shear rate at each rotational speed was obtained by using Mitschka's technique, which is described fully elsewhere (36).

Figure 10 shows experimental data on the mean main hyphal length as a function of fermentation time. The width of the plotted lines indicates the 95% confidence limits of the data. Variations of the power law coefficient, K, with time are also shown in Figure 10 for comparison. The mean main hyphal length fluctuates throughout the fermentation time, but no observable trend in the data can be detected for any of the microorganisms. Additionally these plots indicate no correlation between the mean main hyphal length and broth rheology. The mean main hyphal

Table 1. Rheological Correlations for Biological Cultures

Microorganism	Rheological model	Variation of theological parameters[a]	Shear range (s^{-1})	Viscometer	Comments	References
A. niger (dispersed mycelia)	Power law	$K = 0.308X^{3.15}d_p^{0.229}$ $n = 0.559X^{-0.058}d_p^{-0.082}$	30–200	Pipe flow	$2.3 < X < 11.2$	51
A. niger (dispersed mycelia)	Power law Bingham Casson	$K = 4.3 \times 10^{-4}X^{3.3}$ $\tau_0 = 0.02X^{2.3}$ $K_c = 0.048\ X^{0.26}$ $\tau_0 = 2.5 \times 10^{-3}X^3$	50–650 for all three models	Coaxial cylinder	Power law was preferred; morphology changes were not considered.	31
A. niger (dispersed mycelia)	Power law	$K = -0.56 + 1.8 \times 10^{-3}RX^{1.7}$ for continuous culture and $K = 0.38 + 4.8 \times 10^{-5}RX^{2.9}$ for fed-batch culture	10–50	Turbine impeller	–	16
A. niger (dispersed mycelia)	Power law for early stage ($x < 4$ g/L), casson for later stage	–	50–150	Turbine impeller and pipe flow	Results for two different viscometers were not consistent.	32
A. niger (dispersed mycelia)	Herschel–Bulkley	$\tau_0 = 4.2 \times 10^{-6} \times X^{2.6}L_{hgu}^{0.65}$ $\tau_0 = 4.8 \times 10^{-7}X^{2.5}R^{3.2}$	–	Haake Rotovisco RV12	–	46
A. awamori	Power law	$K = 6.9 \times 10^{-5}X^{3.96}$	–	Coaxial cylinder	Morphological changes were not considered.	31
P chrysogenum (dispersed mycelia)	Casson	$\tau_0 \alpha X^{2.5}$ $K_c = 1.4 \times 10^{-3}X^1L_{hgu}^{0.6}$	–	Turbine impeller	Flexibility can affect rheology.	20
P. chrysogenum (dispersed mycelia)	Power law Bingham Casson	$K = 3.6 \times 10^{-3}X^{2.5}$ $\tau_0 = 0.043X^{2.1}$ $\mu_p = 2.2 \times 10^{-3}X^{0.78}$ $\tau_0 = 8.3 \times 10^{-3}X^{2.5}$ $K_c = 0.047X^{0.19}$	50–650 for all three models	Coaxial cylinder	Morphology changes not considered; power law was preferred.	31
P. chrysogenum (dispersed mycelia)	Herschel–Bulkley Casson	$K = aX^{2.8}R^{0.7}\mathrm{Com}^{1.2}$ $1 - n = bX^{0.7}R^{0.6}\mathrm{Com}^{0.9}$ $\tau_0 = cX^{3.2}R^{0.6}\mathrm{Com}^{1.5}$ $\tau_0 = eX^{2.9}R^{0.6}\mathrm{Com}^{1.0}$ $K_c = fX^{1.2}R^{0.1}\mathrm{Com}^{1.0}$	3.5–75.8 for both models	Turbine impeller	Morphological measurement was for clump, which is more than 90% of total.	16

Table 1. Rheological Correlations for Biological Cultures (continued)

Microorganism	Rheological model	Variation of theological parameters[a]	Shear range (s^{-1})	Viscometer	Comments	References
P. chrysogenum (dispersed mycelia)	Casson	$\tau_0 = aX^2$	–	Turbine impeller	–	41
P. Chrysogenum (dispersed mycelia)	Bingham	$\tau_0 = aX$	–	MacMichael viscometer	–	52
P. chrysogenum (pellet)	Casson	$K_c = aE_x + b$ $\tau_0 = c(E_x/d)^3$	<20	–	E_x = pellet volume fraction	50
S. levoris	Power law Bingham Casson	$K = 0.27X^{0.7}$ $\tau_0 = 1.1X^{0.75}$ $\mu_p = 5.9 \times 10-3X^{0.32}$ $\tau_0 = 0.64X^{0.79}$ $K_c = 0.042X^{0.2}$	50–650 for all three models	Coaxial cylinder	Morphological changes not considered; power law was preferred.	31
S. levoris	Herschel–Bulkley	$\tau_0 = 4.2 \times 10^{-6} \times X^{2.6}L_{hgu}^{0.65}$	–	Haake Rotovisco RV12	–	46
S. fradiae	Herschel–Bulkley and Casson	Apparent viscosity, consistency index, and yield stress change with morphology.	<1300	Coaxial cylinder	Mathematical relation was not shown; Casson was preferred.	53
S. grisens	Bingham	τ_0 and plastic viscosity varied sinousidally with fermentation age.	–	–	–	54
S. erythraea *S. rimoscus* *A. roseourufa*	Power law	Rheology is related to biomass, not morphology.	<250	Disk spindle impeller	Mean hyphal length was between 15 and 25 μm.	44
S. erythraea	Power law	$K\alpha X^{2.14}$	24–880	Coaxial cylinder	–	45
Baker's yeast	Power law	$K\alpha e^{3.43Ec}$; no relation for n	10^4–10^5	Coaxial cylinder	Acidic pH lowered the viscosity.	23
Baker's yeast, *S. cerevisiae* and *C. utilis*	Newtonian	$\mu/\mu_0 = 1/1 - (h_sE_c)^a$; $a = 0.5$ for *S. cerevisiae*, $X < 150$ g/L	–	Falling sphere viscometer	*C. utilis* is considerably smaler than *S. cerevisiae* and had a higher value of viscosity.	25
Baker's yeast	Newtonian	$\mu = \mu_0(1 + 2.5E_c + 7.25 \times E_c^2)$	–	–	–	55
X. campestris	Power law and Herschel–Bulkley	The non-Newtonian behavior is controlled by molecular properties of dissolved polysaccharide.	<200	–	Thixotrophy was shown.	5

[a] a, b, c, d, e and f are constant parameters.

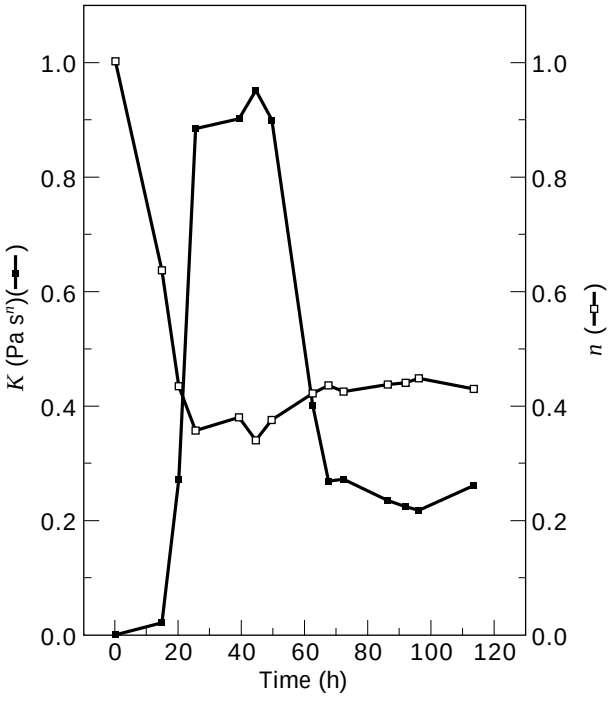


Figure 6. Variation of consistency index, *K*, and flow behavior index, *n*, of *S. erythraea* (strain CA 340) culture as a function of time of the fermentation.

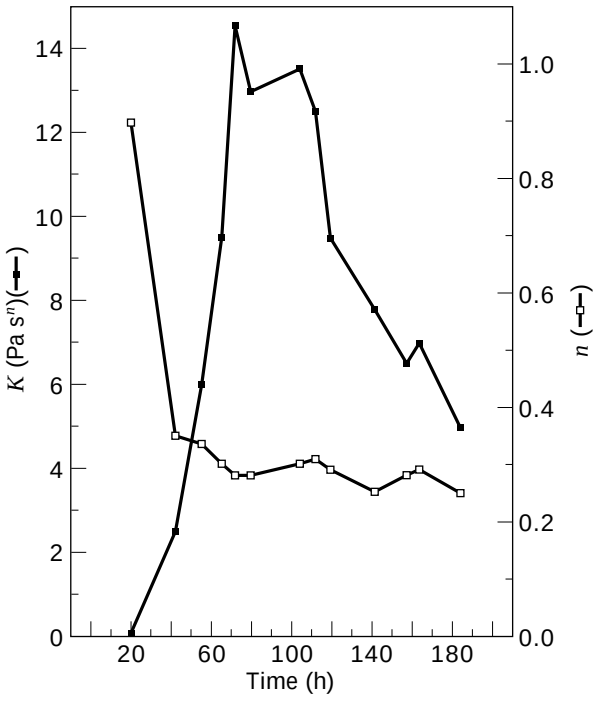


Figure 8. Variation of consistency index, *K*, and flow behavior index, *n*, of *S. roseorufa* industrial strain culture as a function of time of the fermentation.

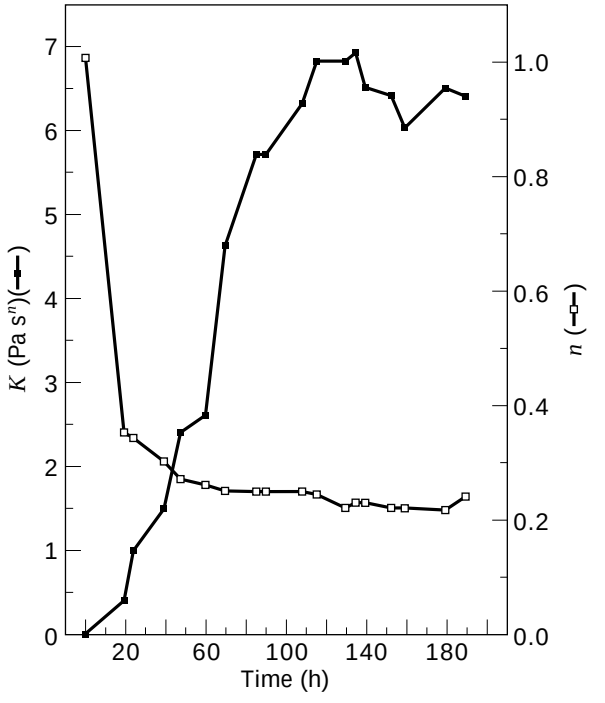


Figure 7. Variation of consistency index, *K*, and flow behavior index, *n*, of *S. erythraea* industrial strain culture as a function of time of the fermentation.

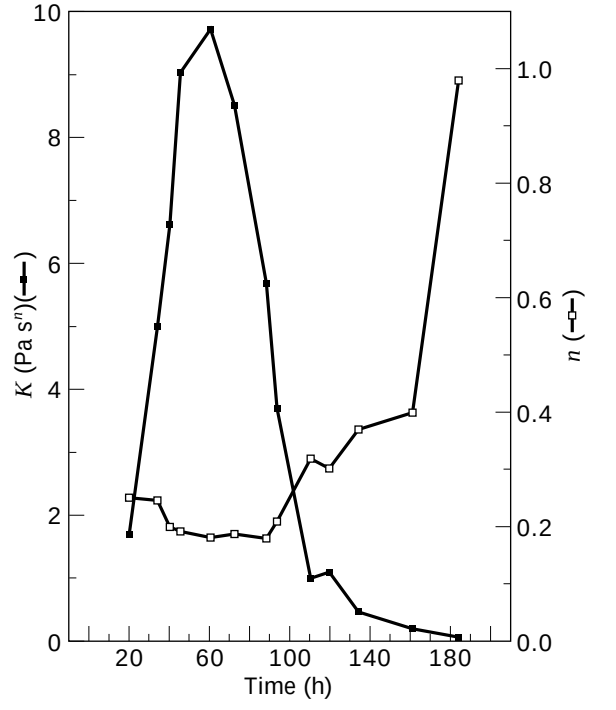


Figure 9. Variation of consistency index, *K*, and flow behavior index, *n*, of *S. rimosus* industrial strain culture as a function of time of the fermentation.

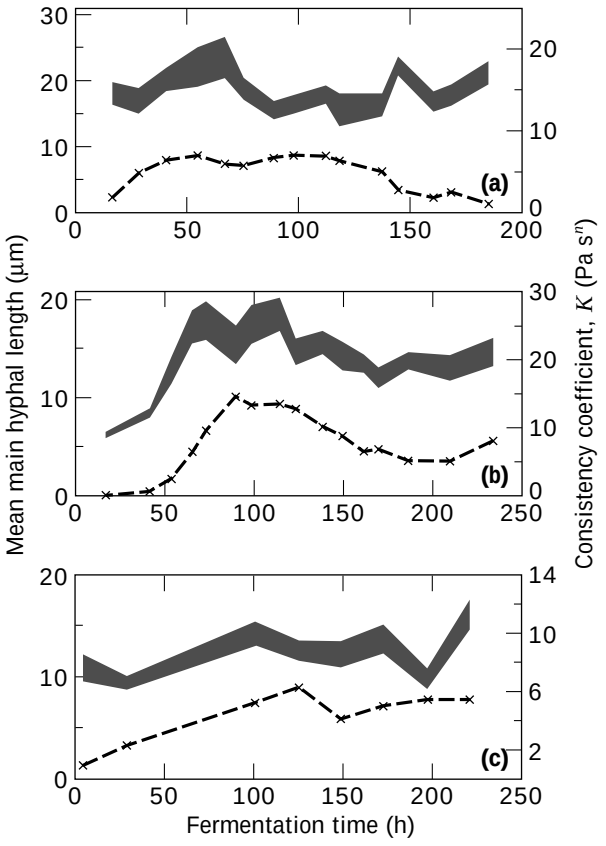


Figure 10. Variation of mean main length and consistency index, K, of fermentation cultures of (**a**) *S. erythraea*, (**b**) *A. roseorufa*, and (**c**) *S. rimoscus*.

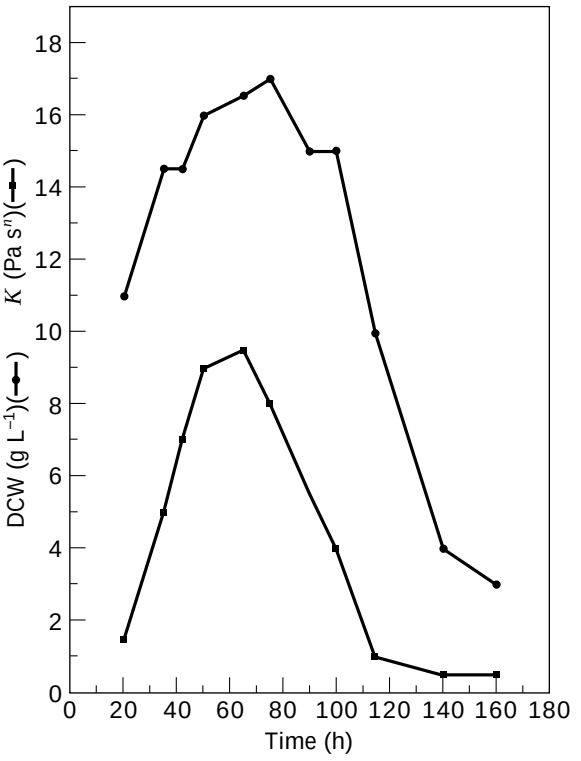


Figure 11. Variation of dry cell weight and consistency index for *S. erythraea* culture.

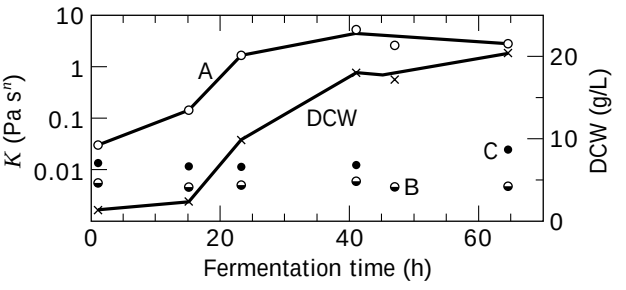


Figure 12. Effect of biomass, insoluble particulate matter, and soluble compounds on rheology of *S. erythraea* culture.

length of the three microorganisms shown in Figure 10 are significantly smaller than those reported previously for *Streptomyces clavuligerus* (45–120 mm), and *P. chrysogenum* (e.g., 130–420 μm and 60–120 μm).

An examination of the information given in Table 1 shows that, compared with the fungal cultures, the rheological parameters, K, τ_o, μ_p, and K_c, for example, are less sensitive to changes in the biomass concentration. The effect of dry cell weight (DCW) on the consistency coefficient, K, of *S. erythraea* culture is shown in Figure 11. The similarity of the two profiles in Figure 11 is evident from these plots, and a series of experiments were conducted to examine the factors responsible for the observed dependency of the broth consistency coefficient, K, on DCW. Two fermentation batches were prepared and run in the same way except that only one batch was inoculated with *S. erythraea*. Several samples were taken periodically from each batch for rheological measurements and DCW evaluation. Some of the samples were also centrifuged prior to analysis to remove *S. erythraea* cells and/or medium particulate and insoluble debris. Figure 12 shows the results of these experiments and indicates that for the inoculated batch (A) the profiles of the consistency coefficient, K, and the DCW are similar and increase with fermentation time. However, once the *S. erythraea* biomass was removed by centrifugation (curve B) the broth rheology changed dramatically, the variation in K disappeared, and the flow behavior index increased from its value of about 0.2 to a value close to unity, indicating a change from a highly shear thinning to an almost Newtonian behavior. Samples from the uninoculated batch showed very little change in their rheology, which remained Newtonian practically throughout (curve C). Similar results shown in Figure 13 were observed for the fermentation of *A. roseorufa*, confirming that the high apparent viscosity of the filamentous cultures depends on the presence and concentration of the organisms, and that other constitutive components (e.g., insoluble particulate matter) appear to have little effect on broth rheology, at least for the cases presented here.

The effect of pH of the culture on broth rheology was also investigated, and it was found to have a strong

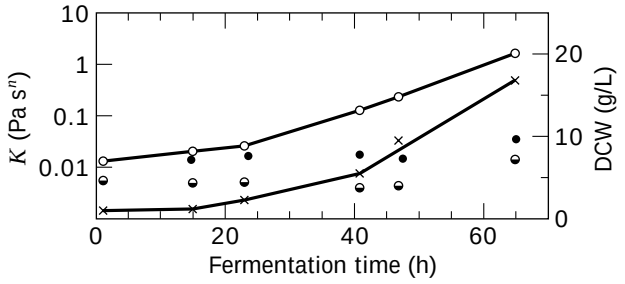


Figure 13. Effect of biomass, insoluble particulate matter, and soluble compounds on rheology of *S. roseorufa* culture.

effect on the extent of non-Newtonian behavior of the culture. This effect was somewhat unexpected and was further investigated for one of the microorganisms, *S. rimosus*. A large sample of the broth was removed from the bioreactor and divided into several specimens. Using concentrated hydrochloric acid and concentrated sodium hydroxide (60% by weight) the pH of each sample was altered to cover the range of pH between 3 and 10, in increments of 0.5 pH. The rheological properties of each sample were then measured using the Brookfield viscometer. Figure 14 shows the effect of pH on both the consistency coefficient, K, and the flow behavior index, n, of the broth. Measurements of the hyphal morphology (not shown) revealed no effect of pH on the mean main hyphal length of the individual hyphae. An important implication of the data in Figure 14, however, is that alteration in charges on and/or close to the surface of the microorganisms, arising from vari ations in pH, could play a role in determining the observed changes in broth rheology. The zeta potential characterizes the nature and the extent of surface charges, and the zeta

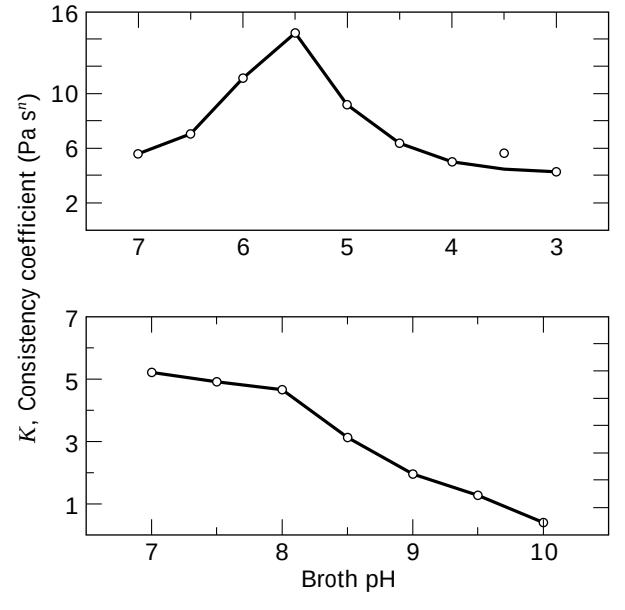


Figure 14. Effect of pH on consistency index of *S. rimosus* culture.

potential–pH profiles of the microorganisms should reflect these effects. No electrophoretic measurements are available in the literature concerning the zeta potential–pH profiles of microorganisms. These profiles should be developed experimentally and studied to examine how they are affected by broth conditions. The results could provide valuable information on the interrelationship between broth rheology, hyphal morphology, and biomass concentration.

CONCLUSION

Considerable experimental work has been published on the rheology of biological suspensions since the 1960s. Advances made in image analysis techniques in the early 1990s resulted in further work on relating the rheological properties to the morphological characteristics of filamentous microorganisms. No consistent opinion has yet emerged on the interaction between the two. Recent developments on the measurement of fractal properties of mycelium may provide a useful basis for further elucidation of these complex interactions.

NOMENCLATURE

Com	Compactness of mycelial aggregate
D	Tank diameter (m)
DCW	Dry cell weight (g L^{-1})
d_p	Diameter of mycelial aggregate (m)
E_c	Volume fraction of wet cells
h_s	Packing factor
K	Consistency index (N sn m^{-2})
K_c	Casson viscosity (N s m^{-2})$^{1/2}$
K_s	Impeller shear rate constant
L_e	Length of the hyphae (μm)
L_{hgu}	Length of hyphal growth unit (μm)
n	Flow behavior index
N	Rotational speed of impeller (s^{-1})
P	Power (W)
P_0	Impeller power number
R	Roughness of mycelial aggregates
T	Torque (N m)
X	Concentration of biomass (g L^{-1})
γ	Shear rate (s^{-1})
γ_{ave}	Average shear rate (s^{-1})
μ_0	Viscosity of cell-free medium (N m^{-2} s)
μ	Viscosity of culture (N m^{-2} s)
μ_a	Apparent viscosity (N m^{-2} s)
μ_p	Plastic viscosity (Pa s)
ρ	Density (kg m^{-3})
τ	Shear stress (N m^{-2})
τ_0	Yield stress (N m^{-2})

BIBLIOGRAPHY

1. M. Borovic, A. Cimerman, W. Steiner, and T. Koloini, *Appl. Microbiol. Biotechnol.* **34**, 579–581 (1991).
2. P. Ayazi Shamlou, D.J. Pollard, A.P. Ison, and M.D. Lilly, *Chem. Eng. Sci.* **49**, 303–312 (1994).
3. S.J. Warren, E. Keshavarz-More, P.A. Shamlou, M.D. Lilly, C.R. Thomas, and K. Dixon, *Bioprocess Eng.* **13**, 45–48 (1995).
4. C.M. Tuffile and F. Pinho, *Biotechnol. Bioeng.* **12**, 849–871 (1970).
5. M. Charles, *Adv. Biochem. Eng.* **8**, 1–62 (1978).
6. H.W. Blanch and S.M. Bhavaraju, *Biotechnol. Bioeng.* **18**, 745–790 (1976).
7. F. Vardar and M.D. Lilly, *Appl. Microbiol. Biotechnol.* **14**, 203–211 (1982).
8. G. Larsson and S.O. Enfors, *Bioprocess Eng.* **3**, 123–127 (1988).
9. H. Funahashi, H. Harada, H. Taguchi, and H. Yoshida, *J. Ferment. Technol.* **20**, 277–282 (1987).
10. W. Ruohang and C. Webb, *Biotechnol. Tech.* **9**, 55–58 (1995).
11. M. Reuss, D. Debus, and G. Zoll, *German Biochem. Eng.* 233–236 (June 1982).
12. J.C. Van Suijdam and J.C. Metz, *Biotechnol. Bioeng.* **23**, 111–148 (1981).
13. H.L. Packer and C.R. Thomas, *Biotechnol. Bioeng.* **35**, 870–881 (1990).
14. K.G. Tucker, T. Kelly, P. Delgrazia, and C.R. Thomas, *Biotechnol. Prog.* **4**, 353–359 (1992).
15. E. Olsvik, and B. Kristiansen, *Biotechnol. Adv.* **12**, 1–39 (1994).
16. E. Olsvik, K.G. Tucker, C.R. Thomas, and B. Kristiansen, *Biotechnol. Bioeng.* **42**, 1046–1052 (1993).
17. S.K. Treskatis, V. Oregeldinger, H. Wolf, and E.D. Gilles, *Biotechnol. Bioeng.* **53**, 191–201 (1997).
18. N. Ohta, Y.S. Park, K. Yahiro, and M. Okabe, *J. Ferment. Bioeng.* **79**, 443–448 (1995).
19. L. Ju, C.S. Ho, and J.F. Shanaham, *Biotechnol. Bioeng.* **38**, 1223–1232 (1991).
20. B. Metz, N.W.F. Kossen, and J.C. Van Suijdam, *Adv. Biochem. Eng.* **11**, 103–156 (1979).
21. J.C. van Suijdam and B. Metz, *Biotechnol. Bioeng.* **23**, 111–148 (1981).
22. H. Eiki, H. Ishida, H. Saito, Y. Oka, and T. Osono, *J. Ferment. Bioeng.* **67**, 345–349 (1989).
23. S. El-Temtamy, L. farahat, A. Nour el-din, and A. Gaber, *Eur. J. Appl. Microbiol. Biotechnol.* **5**, 156–160 (1982).
24. B.W. Shimmons, W.Y. Svrcek and J.E. Zajic, *Biotechnol. Bioeng.* **18**, 1793–1805 (1976).
25. M. Reuss, M. Popovic and W.K. Bronn, *Appl. Microbiol. Biotechnol.* **8**, 167–175 (1979).
26. R.P. Tomkiewicz, A. Biviji, and M. King, *Biorheology* **31**, 511–520 (1994).
27. B. McNeil and L.M. Harvey, *Crit. Rev. Biotechnol.* **13**, 275–304 (1993).
28. K. Gbewonyo, G. Hunt, and B. Buckland, *Bioprocess Eng.* **8**, 1–7 (1992).
29. G. Li, H. Qiu, Z. Zheng, Z. Cai, and S. Yang, *J. Chem. Tech. Biotechnol.* **42**, 385–391 (1995).
30. U. Bjorkman, *Biotechnol. Bioeng.* **29**, 114–129 (1987).
31. D.G. Allen and C.W. Robinson, *Chem. Eng. Sci.* **45**, 37–48 (1990).
32. N. Blakebrough, W.J. McManamey, and K.R. Tart, *J. Appl. Chem. Biotechnol.* **78**, 453–461 (1978).
33. E.S. Olsvik and B. Kristiansen, *Biotechnol. Bioeng.* **40**, 375–387 (1992).
34. N. Blakebrough, W.J. McManamay, and W.J. Tart, *J. Appl. Chem. Biotechnol.* **28**, 453–461 (1978).
35. A.B. Metzner and R.E. Otto, *AIChEJ.* **3**, 3–10 (1957).
36. P. Mitschka, *Rheological Acta* **21**, 207–209 (1982).
37. Z. Kemblowski and B. Kristiansen, *Biotechnol. Bioeng.* **28**, 1474–1483 (1986).
38. P. Ayazi Shamlou and M.F. Edwards, *Chem. Eng.* **40**, 1773–1781 (1985).
39. H. Barnes, *Chem. Eng.* **24**, 17–24 (1993).
40. J.J.T.M. Bongenaar, N.W.F. Metz, and F.W. Meijboom, *Biotechnol. Bioeng.* **15**, 201–206 (1973).
41. J.A. Roels, J. Van Den Berg, and R.M. Voncken, *Biotechnol. Bioeng.* **16**, 181–208 (1974).
42. M. Mohseni, H. Kautola, and D.G. Allen, *J. Ferment. Bioeng.* **83**, 281–286 (1977).
43. T. Oalman and H.W. Blanch, *Crit. Rev. Biotechnol.* **4**, 133–184 (1986).
44. S.J. Warren, E. Keshavarz-Moore, P. Ayazi Shamlou, M.D. Lilly, C.R. Thomas, and K. Dixon, *Biotechnol. Bioeng.* **45**, 80–85 (1995).
45. S.M. Heydarian, Ph.D. Thesis, University of London, London, U.K. (1998).
46. M. Mohseni and D.G. Allen, *Biotechnol. Bioeng.* **48**, 257–265 (1995).
47. D.B. Patankar, T. Liu, and T. Oolman, *Biotechnol. Bioeng.* **42**, 571–578 (1993).
48. J.H. Kim, J.M. Lebeault, and M. Reuss, *Eur. J. Appl. Microbiol. Biotechnol.* **18**, 11–16 (1983).
49. E. Ujokova, Z. Fencl, M. Muzilkova, and L. Seichert, *Biotechnol. Bioeng.* **22**, 227–241 (1980).
50. R. Witter, R. Matthes, and K. Schugerl, *Appl. Microbiol. Biotechnol.* **18**, 17–23 (1983).
51. I.A. Fatile, *Appl. Microbiol. Biotechnol.* **21**, 60–64 (1985).
52. F.H. Deindoerfer and E.L. Gaden, *Appl. Microbiol.* **3**, 253–257 (1970).
53. N.P. Ghildyal, M.S. Thakur, S. Srikanta, S.A. Jaleel, S.G. Prapulla, M.S. Prasad, P.N. Devi, and B.K. Lonsane, *J. Chem. Tech. Biotechnol.* **38**, 221–234 (1987).
54. J.W. Richards, *Prog. Ind. Microbiol.* **3**, 253–260 (1961).
55. B.W. Shimmons, W.Y. Svrcek, and J.E. Zajic, *Biotechnol. Bioeng.* **18**, 1793–1805 (1976).
56. K.G. Tucker and C.R. Thomas, *Food Bioprod. Proc.* **71**, 111–117 (1993).

See also *ASPERGILLUS*; SECONDARY METABOLITE PRODUCTION, ACTINOMYCETES, OTHER THAN STREPTOMYCES; XANTHAN GUM.

ROLLER BOTTLE CULTURE, MIXING

D.R. UNGER
M. LIU
P.D. SWANSON
F.J. MUZZIO
Rutgers University
Piscataway, New Jersey
JOYE BRAMBLE
JAMES SEARLES
JOHN AUNINS
Merck and Company
West Point, Pennsylvania

KEY WORDS

Cell propagation
Mixing
Particle settling
Roller bottle
Virus infection

OUTLINE

INTRODUCTION

The settling behavior of biological cells with densities different than the fluid is a key factor affecting the performance of roller bottle bioreactors. The fully three-dimensional velocity field in a roller bottle bioreactor was simulated using a finite-element method, and the settling behavior of cells was simulated for both steady and time-dependent bottle rotation rates using particle dynamics algorithms. These simulations were validated experimentally using fluorescent particle tracers. Under steady flow conditions the flow is divided into two regions, one in which the particles settle to the wall and one in which the particles remain suspended indefinitely, but unsteady flow conditions (characterized by repeated reversal of the flow direction) lead to complete cell settling. These observations showed qualitative agreement with culture infection kinetic data in which changes in the roller bottle rotation rate resulted in enhanced infection. The main mixing limitation observed experimentally in roller bottles is very slow mixing in the axial direction, caused by flow segregation at the central plane of symmetry of the bottle. This limitation is readily overcome by imposing a small-amplitude vertical rocking motion to the bottle, forcing fluid to flow through the segregation plane and leading to a completely homogeneous mixture in just a few revolutions of the bottle.

BACKGROUND

Manual and automated roller bottle systems have been used for more than 40 years in the pharmaceutical, biochemical, and medical fields for processes such as cell growth and infection, heterologous glycoprotein production, vaccine preparation, and high-density plant cell cultivation (1–10). Despite efforts by numerous investigators to develop unit operations such as microcarrier culture for the production of anchorage-dependent cells or cell products (11–19), roller bottle systems still persist in research and industry. Additionally, for industrial-scale cell culture production (e.g., vaccines), cells are frequently expanded in roller bottles before transfer to microcarrier cultures for the final growth phase even when unit operation-based systems are utilized (17).

Widespread use of the roller bottle has several reasons. Most notably, the process relies on very simple technology: a horizontal cylindrical vessel is filled approximately one-fourth full and rotated axially; increased production is accomplished simply by increasing the number of bottles. Thus, scale-up development is straightforward, resulting in reduced developmental time lines for industry and faster introduction to market for new products. The system allows constant fluid–gas contact and easy addition of nutrients without interruption of the process. In addition, the process is capable of maintaining sterile conditions for prolonged times, contamination of one or more roller bottles does not necessarily result in the failure of an entire lot, precise control of nutrient and waste product levels is possible, and the direct monitoring of cell growth and metabolism is relatively simple (6).

On the other hand, roller bottles are limited in surface area available for growth and in the volume of harvest fluid obtained. Manpower and facility space requirements are higher than for unit operation systems such as microcarriers because hundreds or thousands of roller bottles can be operated for a single production run, although new automation systems are addressing the manpower issue (10,20). In addition, cell growth and infection are believed to be suboptimal in certain vaccine products because of flow and mixing dynamics, perhaps by preventing infected cells from attaching to host cells attached to the bottle walls (4). Although these issues point toward an obvious need for flow analysis and process design criteria, there have been no published results to date on either of these topics.

The following sections detail computational and experimental characterizations of the fluid flow profiles within rotating roller bottles, including particle trajectories, fluid mixing patterns, and unsteady-state flow strategies. The results point to ways in which cell proliferation and infection can be optimized by simple modifications to roller bottle rotation. This article is organized as follows. First, the numerical method for obtaining the three-dimensional flow solution is described. The flow solution is then validated by comparison with experimental measurements obtained via particle imaging velocimetry. Computationally predicted mixing patterns are compared with mixing patterns obtained by visualization of fluorescent dye streaklines. The model is then used to examine cell motion in the flow field. Computationally predicted cell motion patterns are compared with experimentally observed particle motion patterns obtained via long-time exposure photography of fluorescent particles. Then, particle settling results for the roller bottle under steady flow conditions are discussed and related to cell growth results obtained for various bottle rotation speeds. Finally, attention turns to improving the performance of roller bottles. The use of unsteady motions on the cylinder to enhance cell settling is explored, and the use of rocking flow perturbations to enhance axial mixing is demonstrated. The final portion of this chapter is devoted to conclusions and recommendations for future research.

DEVELOPMENT AND EXPERIMENT VALIDATION OF A COMPUTATIONAL MODEL OF FLUID AND CELL MOTION IN A ROLLER BOTTLE

Flow Field Solution

The three-dimensional flow field for a standard 850-cm^2 roller bottle geometry was obtained using a commercial CFD software package, FLUENT. This and similar software packages have been used by several researchers in the past to solve fluid dynamics problems (21–26). The equations of motion for the flow are the well-known Navier–Stokes equations, which, for a two-dimensional flow, include the momentum equation in two components and the continuity equation:

$$\boldsymbol{u}/t + \boldsymbol{u}\nabla\boldsymbol{u} = -\nabla p + (1/\mathrm{Re})\nabla^2\boldsymbol{u} + (1/\mathrm{Fr}) \quad (1)$$

$$\nabla\boldsymbol{u} = 0 \quad (2)$$

where $\boldsymbol{u}$ is the dimensionless velocity and p is the dimensionless pressure. Re in Equation 1 is the Reynolds number, $\mathrm{Re} = \rho vL/\mu$, where ρ and $\boldsymbol{\mu}$ are the fluid density and viscosity, and v and L are the bottle's linear rotational velocity and the liquid height, respectively; Fr is the Froude number, $\mathrm{Fr} = gL/v^2$, where g is gravity. The rationale for selecting liquid height as the natural length scale for the flow is that the liquid height is the diffusional length scale in the system. Viscous forces can be viewed as diffusive transport of momentum, with the kinematic viscosity (μ/ρ) playing the role of a momentum diffusivity. Hence, the liquid depth is the natural length scale for the action of viscous forces.

The experimental geometry is shown in Figure 1a. To obtain the velocity field, it was assumed that (*1*) the flow is steady, (*2*) the flow is in the Stokes regime (creeping flow conditions), (*3*) the free boundary is both planar and horizontal (effects of surface tension and viscous drag on the boundary are neglected). For the conditions used in the biotechnology industry, Re = 10, confirming the validity of the creeping flow assumption.

The physical domain was discretized using a curvilinear mesh; although the bottle has a total length to diameter ratio of 2:1, because of its symmetry only one-half of the system needed to be simulated, giving a length to diameter ratio of 1:1. The liquid height in the bottle was one-fourth the total diameter, which is typical of industrial practice. A 80 × 20 × 94 structured computation grid was used; Figure 1b represents a schematic of the z-plane grid (for the computation a higher node density was used). An acceptable node density for the computational mesh was found by creating several meshes with different node densities. The number of nodes in the mesh was successively increased by 25% until an average velocity difference of less than 3% was achieved for two successive meshes. The lower node density was then used to generate the computational mesh.

Velocity and pressure fields were obtained via iteration using a Sun SPARC 20 workstation. Approximately 140 Mb of RAM and 60 h of CPU time were required for solution convergence. To determine convergence, a criterion of 10^{-6} was used for each normalized residual velocity and pressure component. Residuals for each iteration were normalized versus the residual values obtained after the second iteration of the solver. To test the sufficiency of the convergence criterion, the simulation was run using convergence criteria ranging from 10^{-5} to 10^{-6} with no significant change in the velocity field when the convergence criteria were varied over this range (<1.0% change in velocity magnitude, on average). Sensitivity analysis (i.e., trajectory closing) showed that such a mesh was sufficiently accurate. These flow data were then used for all subsequent particle tracking simulations.

The condition adopted in the base case simulation was $\Omega = 0.25$ rpm, $\mu = 0.125$ Pa s at room temperature, giving a Reynolds number of Re 1.32. At this Re, the flow can be considered to be in the creeping flow regime. The flow was assumed to be sufficiently slow so that it could be assumed that the free surface was always horizontal and no surface tension effects were present at the bottle walls. The free surface was therefore modeled as a frictionless wall (i.e., a surface that exerts no stresses on the fluid). Several assumptions were also made regarding the particles simulated in the postprocessing program. The particles in the flow were sufficiently dilute so that they did not interact with each other or affect the fluid flow in any way. The particles were assumed large enough so that Brownian forces could be ignored. All the assumptions used here are valid for the majority of industrial uses of roller bottles.

The Fluid Velocity Field in a Roller Bottle

Two flow visualization experiments were performed to validate the simulations. The experiments were conducted in

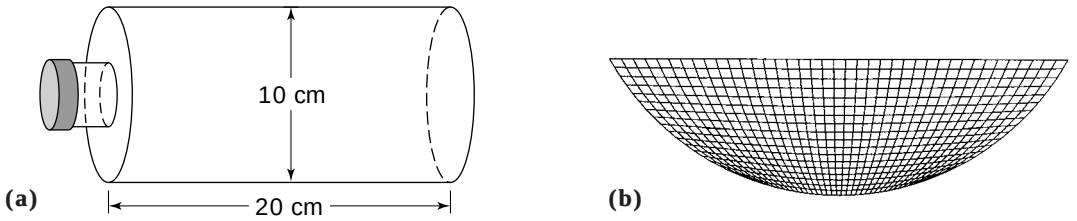


Figure 1. (**a**) Schematic diagram of the roller bottle. (**b**) Schematic diagram of the computational grid used for the simulations.

a glass roller bottle 10 cm in diameter and 20 cm in length. The working fluid was glycerin, which has a density of 1.26 g/cm^3 and a viscosity of 1.25 Pa s. The bottle was rotated using an apparatus that consisted of a set of rollers whose rotation speed and direction were accurately controlled by a computer. The rotation rate for the two experiments performed was 0.25 rpm, giving a Re of 0.13. This flow condition can be assumed to be representative of all conditions in the creeping flow regime, because in this regime the velocity field becomes independent of Re. The velocity field was first validated via comparison with particle imaging velocimetry (PIV) results. The use of particle imaging techniques to study fluid flow problems is well documented in the literature and has been reviewed by Adrian (27). Information that is pertinent to our specific system is provided here. The flow was seeded with 12-μm-diameter silver-coated hollow glass spheres (Potter's Industries) that have a specific gravity of 1.17. These particles are sufficiently small as to provide accurate measurements of the instantaneous velocity field. A 10-mJ New Wave Research pulsed mini-YAG laser was used as the illumination source. The laser beam passed through a series of optical components to produce a sheet of light 1 mm thick (at a focal distance of 1 m). A 32 $\times$ 32 grid was analyzed using cross-correlation data acquisition, with 10 ms between laser pulses. Resultant data from PIV was in the form of velocity direction and magnitude at the center of each interrogation area.

The experimental velocity field was measured in the vertical plane normal to the bottle axis in the middle of the bottle. Figure 2 shows the experimental (Fig. 2a) and computational (Fig. 2b) velocity vector fields. The length and color of the arrows represent the velocity magnitude in the plane, ranging from low (dark blue) to high (red). As seen in this figure, the velocity fields are qualitatively identical, with the exception of some small experimental scattering near the free surface of the liquid.

The computational velocity field was further validated by comparing computational and experimental mixing patterns obtained for passive tracers at the center plane of the bottle. This first step for simulating mixing patterns is to determine the trajectories followed by point particles as a result of the equation of motion:

$$d\boldsymbol{x}/dt = \boldsymbol{v} \tag{3}$$

where $\boldsymbol{x}$ is the particle position and $\boldsymbol{v}$ is the fluid velocity. Figure 3 shows the flow pathlines predicted using 4th-order Runge–Kutta integration of equation 3, which are also the flow streamlines because the flow is steady. A stagnation point is located on the symmetry line at 0.6 unit lengths below the free surface (the unit length is the height of the liquid). A similar approach was used to simulate the evolution of a vertical filament initially placed at the center of the bottle (Fig. 4a). The "line" consisted of 10,000 particles moving at the flow velocity. Figures 4b–4e show the configuration adopted by the line, as revealed by the particle positions, at 4, 8, 16, and 32 bottle revolutions, respectively. As the particles are moved by the flow, the line is twisted around the stagnation point. The line stretch in this steady two-dimensional flow is linear with time. The evolution of a fluid line was also observed ex-

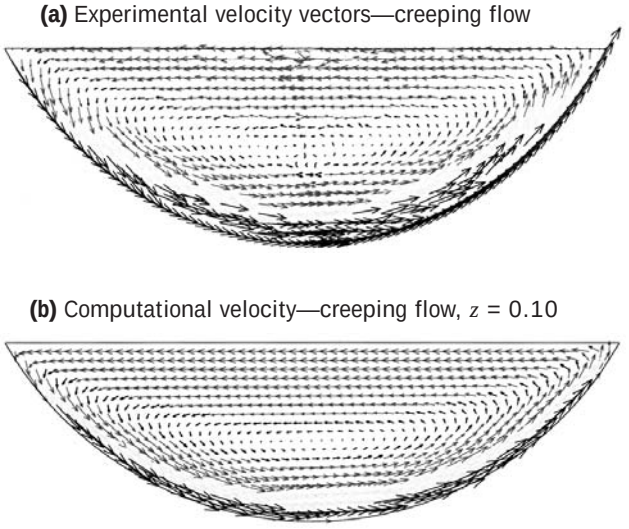


Figure 2. (**a**) Particle imaging velocimetry determination of the velocity field at the center plane of the bottle. (**b**) Finite volume results of velocity field at the center plane of the bottle.

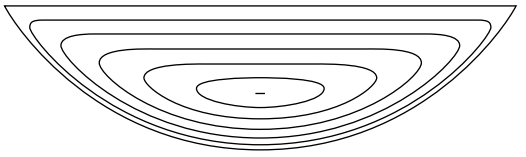


Figure 3. Computational stream lines of the flow at the center plane of the bottle under steady flow conditions.

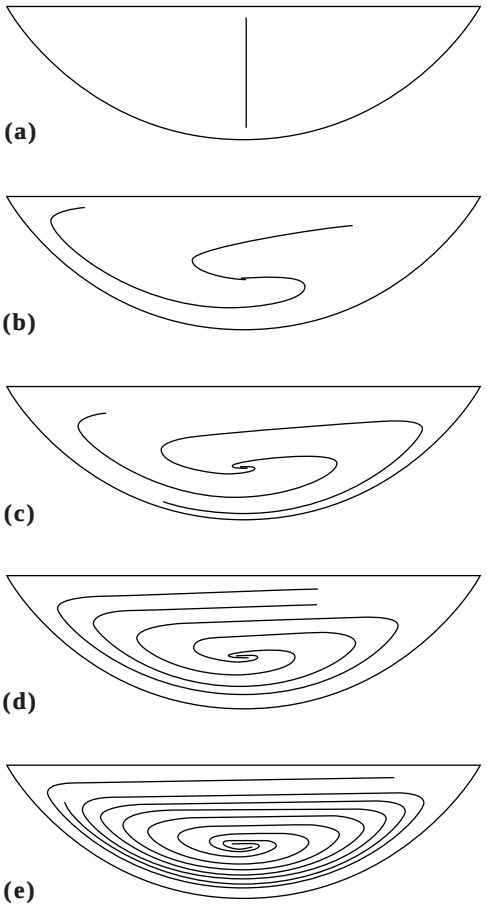


Figure 4. Evolution of a vertical line placed in the flow field. (**a**) Initial condition. (**b**–**e**) The line-after 4, 8, 16, and 32 bottle revolutions, respectively.

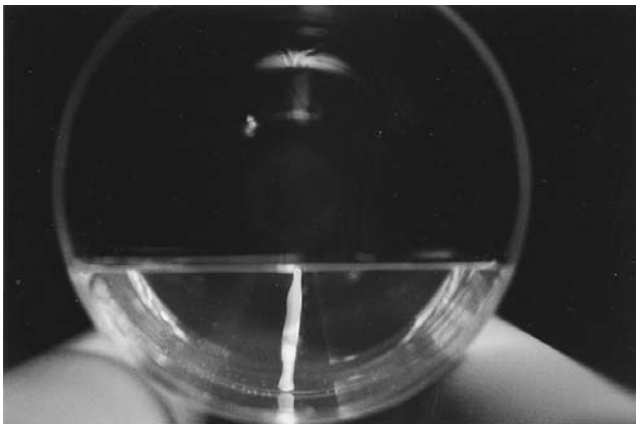


Figure 5. Initial condition for the flow visualization experiment. The fluorescent dye was placed in the center of the bottle axially so that end effects would cancel.

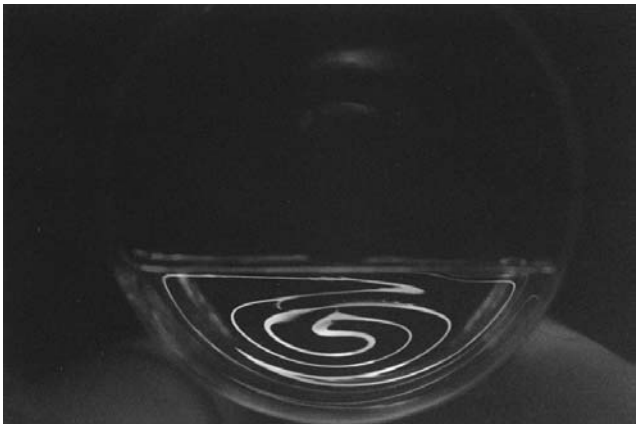


Figure 6. Experimental validation of the evolution of a vertical line placed in the flow field.

perimentally to verify the simulations. Figure 5 represents the initial condition, which consisted of a vertical streak of dyed fluid placed in the center plane of the roller bottle, and Figure 6 shows the evolution of the streak of dye after 20 revolutions. Once again, excellent qualitative agreement between experiments and simulations was obtained, and it was concluded that the simulations indeed provide an accurate representation of the velocity field in the roller bottle.

Once the computational velocity field was validated, it was used to obtain a detailed characterization of fluid motion in the bottle. To provide a clear illustration of a three-dimensional velocity vector field is not a trivial task. This is accomplished by using two-dimensional cuts of the velocity field, where the in-plane components are shown using a vector and the total magnitude of the velocity field is shown using color contours, with red signifying high velocity and blue indicating low velocity.

Figure 7 shows vertical cross sections at 2-cm intervals through the bottle (i.e., $z = 0.0$ is 0.0 m from the end wall, and $z = 0.10$ cm is at the symmetry plane in the center of the roller bottle). Figure 7f corresponds to the center of the bottle and is therefore identical to Figure 2b. Moreover, Figures 7d and 7e are also identical to Figure 7f, indicating that end effects play little or no role in the fluid flow past 5 cm from the end of the bottle. The velocity field shows both reflection symmetry relative to the center plane (expected from the symmetry of the boundary) and antisymmetry between the left and right sides of the bottle (expected from the creeping flow condition and the rotational symmetry of the bottle). As discussed next, this observation can have important implications because the flow induced by the end walls of the bottle is the only means of axial mixing.

Figure 7a represents the velocity field at the end of the roller bottle. Due to the no-slip condition at the bottle wall, all fluid follows the bottle wall in a counterclockwise direction. When the fluid reaches the free surface (at the right side), it detaches from the wall and flows outward in the z-direction. At the left side, the fluid immediately adjacent to the wall is drawn downward; this motion is "fed" by an inward flow from the z-direction. Figures 7b and 7c show the presence of a high-velocity component near the surface of the liquid, caused by the end-wall effects. However, detailed examination of the velocity vectors displayed

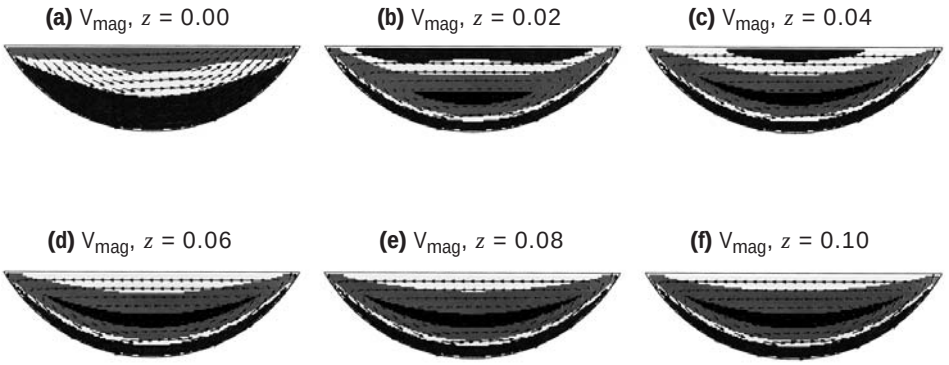


Figure 7. Finite-volume results. Velocity field at several vertical planes spaced along the axis of the bottle.

in Figures 7b and 7c indicates that the *x*- and *y*-components of the velocity field are largely unaffected by end-wall effects; such effects appear primarily in the axial (*z*) component of the velocity field.

The end-wall effects are more clearly shown by a top view of the bottle. Figure 8 shows a velocity contour plot just below the liquid free surface. The end of the bottle is at the bottom of this figure, and the symmetry plane is at the top. Flow in the axial (*z*) direction is clearly visible in this figure. The fluid flows upward from the end wall in the lower right-hand corner, travels in a counterclockwise loop, then flows downward at the lower right-hand corner as a result of the end-wall flow. There is a stagnation point 0.46 cm from the end of the bottle, as shown by the dark blue color. Moving toward the center of the bottle, end effects have less influence on the flow; there is less than 2% difference in the velocity field between the two-dimensional symmetry plane and the plane at 5 cm from the end wall. Similar flow patterns were observed for other values of Re of less than 10.

An alternative illustration of the three-dimensional flow pattern in the roller bottle is obtained by computing the pathlines of the flow. Figure 9 shows six pathlines for fluid particles initially placed 0.2 cm from the bottom of the bottle. The center of the bottle is at the left-hand side of the figure, and the end wall is at the right. Examination of the pathlines from left to right reveals the effect of the end wall on the flow field. The pathline at the center plane is a two-dimensional vertical loop identical to the pathlines observed in Figure 3. Toward the end of the bottle, the pathlines begin to bend near the free surface toward the center of the bottle. All the pathlines shown are closed loops. The pathlines at the center plane of the bottle are two-dimensional loops, and the pathlines away from the symmetry plane are three-dimensional loops. In either situation, however, the same result ensues: fluid particles in the flow remain trapped in the closed periodic orbits and do not contact other portions of the flow. Such regular motion patterns can pose significant barriers to the mixing process. Experimental evidence of this obser-

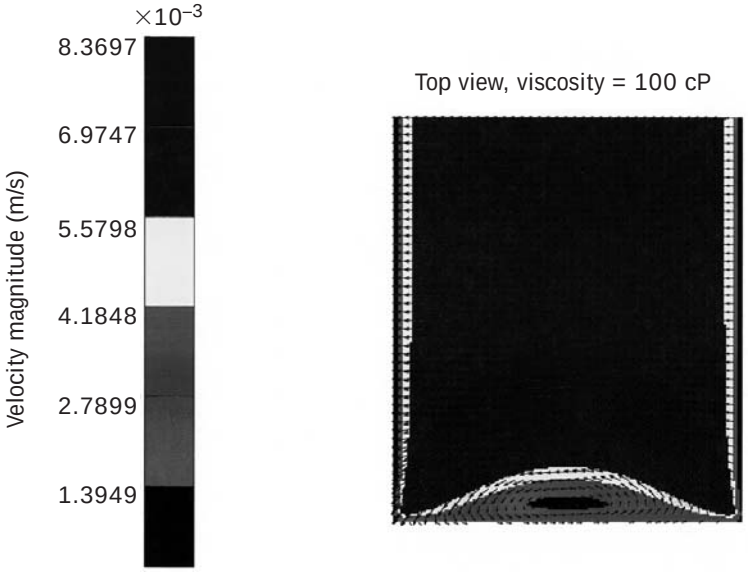


Figure 8. Finite-volume results. Velocity field (*left*) at a horizontal plane near the upper surface of the fluid. *Right*, top view.

vation and approaches for overcoming mixing limitations follow.

The Motion of Cells in a Roller Bottle

The initial point for the modeling of small particles in the flow field is given by the equations of motion proposed by Maxey and Riley (28):

$$m_p \frac{d\mathbf{v}}{dt} = (m_p - m_F)\boldsymbol{g} + m_F \frac{D\mathbf{u}}{Dt} - 6\pi a^2 \mu \int [\pi\nu(t-\tau)]^{-1/2}\left(\frac{d\mathbf{X}}{dt}\right)d\tau - \left(\frac{1}{2}\right) m_F \frac{d[\mathbf{u} - v - (1/10)a^2\nabla^2\mathbf{u}]}{dt} - 6\pi a\mu \boldsymbol{X} \quad (4)$$

where $\boldsymbol{X} = \boldsymbol{u} - \boldsymbol{v} - (1/6)a^2\nabla^2\boldsymbol{u}$.

In these equations, $\boldsymbol{u}$ is the (dimensional) fluid velocity vector, $\boldsymbol{v}$ is the (dimensional) cell velocity vector, a is the cell radius, ν is the fluid kinematic viscosity; and m_p and m_F correspond to the weight of the cell and that of the fluid displaced by the cell, respectively. The five terms on the right-hand side of the equation correspond to buoyancy forces, pressure forces, flow history effects, added mass effect, and Stokes drag forces. This equation can be significantly simplified for the roller bottle problem. Because $\boldsymbol{u}$ is known from the numerical solution of the flow, the relative magnitude of each term can be evaluated. Analysis has shown that the leading order terms are inertia, buoyancy, and drag forces, thus the equation becomes:

$$\frac{d\boldsymbol{v}}{dt} = G\boldsymbol{g} + \frac{1}{S_{tk}}(\boldsymbol{u} - \boldsymbol{v}) \quad (5)$$

where $G = 2(\gamma - 1)gL/[(2\gamma + 1)U^2]$, and $S_{tk} = (2\gamma+1)a^2U/(9\nu L)$; γ is the particle density relative to the fluid density, L is the depth of the fluid, and U is the bottle rotation velocity. If we assume that the particles are cells and the rotation speed and liquid are typical of those used in the biotechnology industry, this equation can be further simplified. Under such conditions, $G \sim O(10^{+3})$ and $S_{tk} \sim O(10^{-6})$, so that the inertial term can be neglected, greatly simplifying the problem. The particle velocity is obtained immediately as

$$\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{V_s}, \quad \boldsymbol{V_s} = \frac{2a^2(\gamma - 1)}{9\nu U}\boldsymbol{g} \quad (6)$$

where $\boldsymbol{v}$, the total particle velocity, is simply the sum of $\boldsymbol{u}$, the fluid velocity, and $\boldsymbol{V}_S$, the particle terminal settling velocity. Once $\boldsymbol{v}$ is known, particle positions are found by numerically integrating the equations

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{v} \quad (7)$$

using the standard 4th-order Runge–Kutta method. To understand small particle behavior in a steady rotating roller bottle, equations 6 and 7 are used to simulate the motion of particles on the center plane of the bottle. Figure 10 shows the particle trajectories for a settling velocity $V_s = 0.05\ U$, where U is the nondimensional tangential velocity of the bottle wall. Six particles were initially placed on the vertical center line. The particles showed different behaviors: while the two particles nearest to the free surface reached the rotating wall in a fraction of a bottle rotation, the remaining four particles moved in closed orbits. Such particles will never reach the bottom wall. This simulation indicates that whether a particle settles or not depends not only on its settling velocity, but also on its initial location within the flow field. Moreover, as the settling velocity used in this simulation is much larger than those corresponding to real systems, Figure 10 actually indicates that a substantial fraction of the cells in a roller bottle could actually remain trapped indefinitely in recirculating trajectories within the bottle flow, never reaching the bottle wall.

This simulation result was investigated experimentally by using long-exposure-time photography to reveal the pathline followed by a fluorescent particle with a settling velocity similar to the one corresponding to Figure 9. A 200-μm fluorescent polystyrene particle was placed in the center plane of the roller bottle. This particle had a terminal settling velocity of approximately 0.5 cm/min, or 0.06 U, where U is the linear rotation velocity of the roller bottle. Using long-exposure photography, the particle pathline was photographed for one bottle revolution under fluorescent light (Fig. 11). The original position of the particle was approximately 0.6 length units from the free surface (where the unit length is the height of the fluid). The pathline revealed in Figure 11 closely resembles the pattern shown in Figure 10, which is a simulation for particles of roughly similar settling velocity. When the particle was originally placed in a different location, closer to the bottle center, the particle followed a different closed pathline. When placed close enough to the upper surface, the particle eventually settled and remain attached to the bottle wall.

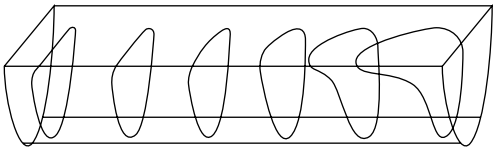


Figure 9. Pathlines of passive tracer particles starting from different positions inside the bottle.

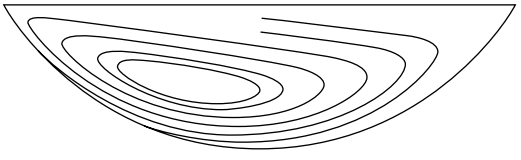


Figure 10. Pathlines of six particles with settling velocity of $V_s = 0.05\ U$ placed in the flow.

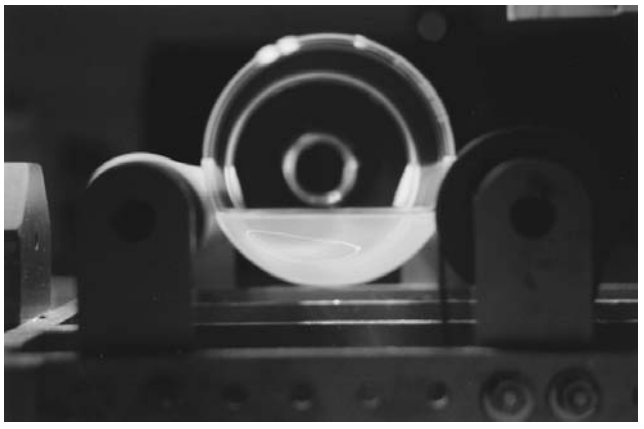


Figure 11. Experimental validation of a particle pathline in the flow. The particle had a settling velocity of approximately 0.06 *U*.

DETAILED ANALYSIS OF CELL SETTLING

Particle Settling Results

A more complete picture of cell behavior can be obtained by following a large number of particles over a long period of time. Such simulations were conducted by placing 20,000 particles uniformly distributed across the entire central cross section of the bottle and following them for 20 rotations of the bottle. Whenever a particle touched the bottle wall, it was (arbitrarily) assumed that the particle had "settled" and was removed from the simulation. Figure 12 shows Poincaré diagrams obtained from such a simulations for settling velocities $V_s = 0$, 0.02 *U*, 0.05 *U*, 0.1 *U*, and 0.2 *U*. For these simulations, the bottle was rotating counterclockwise at a nondimensional linear velocity of $U = 1.0$. For $V_s = 0$ (Fig. 12a), the structure of the particle positions is identical to those of the flow streamlines, because such particles are in fact fluid particles. As V_s increases, the particle paths are perturbed away from streamlines, but as it is shown in Figure 12b, for $V_s = 0.02$ *U* most particles remain trapped in closed periodic orbits and never reach the bottle walls; only those particles initially very close to the free surface and the bottle wall are able to reach the bottle boundary. As settling velocity increases, more particles settle, but only for large values of V_s (i.e., $V_s = 0.2$ *U*; Fig. 12e) are most particles able to reach the bottle walls.

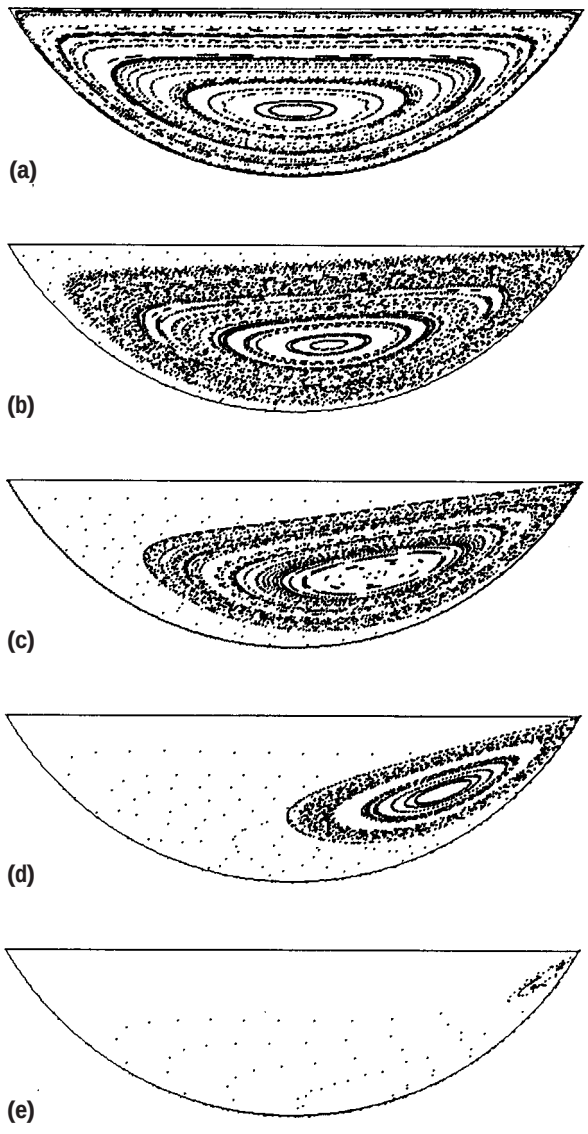


Figure 12. Poincaré diagram of particles, initially distributed uniformly in the central vertical cross section of the bottle, for settling velocities of $V_s = 0$, 0.02 *U*, 0.05 *U*, 0.1 *U*, and 0.2 *U*, respectively, from top to bottom of figure (**a**–**e**).

Although such simulations give information about the dynamic behavior of particle motion, the settling rate cannot be inferred from them. To obtain the spatial distribution of particle settling rate, the settling time of each particle was recorded and plotted as a function of initial particle location. The settling time distribution in the bottle is shown in Figure 13. The bottle is rotated clockwise at a dimensionless linear velocity of $U = 1.0$ at the wall. Results correspond to $V_s = 0.02$ (Fig. 13a), 0.05 (Fig. 13b), 0.1 (Fig. 13c), and 0.2 (Fig. 13d). The colors from dark gray to white represent the particle residence time from 0 to the maximum computing time. The smooth core region in dark gray represents the particles that do not settle at all. For small $V_s = 0.02$ (Fig. 13a), only a narrow strip of particles under the top free surface and near the bottom on the right side of the bottle settle. The width of this strip increases as V_s increases and the size of the red core decreases. The tip of this core region is always at the upper-left corner of the flow and leans to the left side of the bottom of the bottle. At $V_s = 0.1$ (Fig. 13b), most of the particles in the bottle will settle. When V_s is increased to 0.2 (Fig. 13c), almost all the particles settle to the wall. The spatial distribution of particle settling time in the flow is shown clearly by the shades of gray. Short-lived particles are always on the right side of the bottle for this flow direction. If the flow direction were reversed, Figure 13 would appear as a mirror image because of the symmetry of the system. These results clearly reveal that the particle residence time

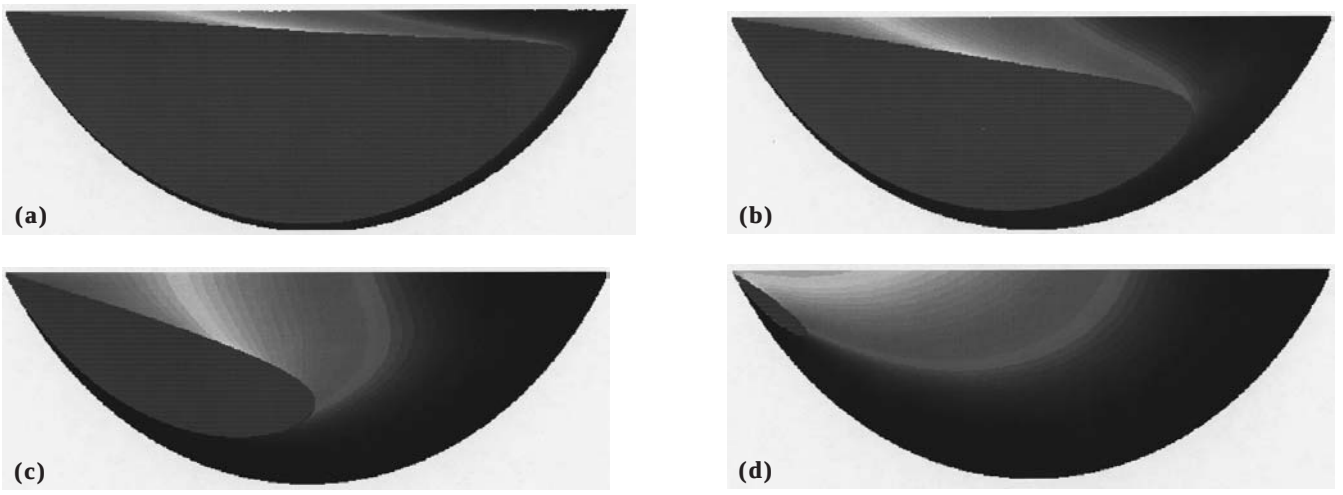


Figure 13. Contour plots of residence time of particles against particle initial position for V_s = 0.02 U, (**a**) 0.05 U, (**b**) 0.1 U, (**c**), and 0.2 U (**d**) after 20 bottle revolutions. The shades in this figure range from dark gray to white, corresponding to zero residence time to infinite residence time.

distribution in suspension is strongly affected by the recirculation of the flow.

The rate of particle settling is shown in Figure 14; the fraction of particles suspended in the flow at a given time is plotted against the dimensionless time, where one bottle revolution corresponds to 12.56 time units. All the settling processes are nonlinear (except for V_s = 0.2). Particles settle to the wall relatively fast at the initial stage, and then the rate of settling (the slope of the curves) levels off in a relatively short time, about 8 to 12 time units. With large settling velocity (V_s = 0.2, the heaviest particles examined), the particles settle at almost a constant rate.

Effects of Bottle Rotation Speed on Cell Culture Experiments

Previous experiments demonstrated that the growth of anchorage-dependent cells and cell culture products (e.g., varicella virus) in roller bottles was much less efficient than that observed in stationary cultures such as T-flasks. It was hypothesized that these differences were caused in part by decreased cell (uninfected and infected cells) sedimentation rates present in the roller bottle cultures as the result of circulating liquid flows that trapped the particles in circular orbits. The decreased sedimentation times resulted in lower cell plating efficiencies for uninfected cells and increased degradation times for virus-infected cells, that were unable to reach the cell culture surface for attachment or infection. To further investigate this hypothesis, the kinetics of virus-infected cell sedimentation and attachment were studied at different rotation rates. The attachment to and growth of cells on the roller bottle surface were also studied to determine the primary mechanism controlling the differences observed in cell growth rates as a function of rotation rate.

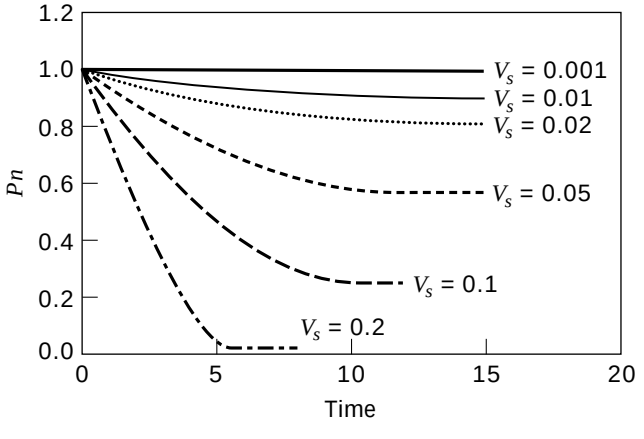


Figure 14. Fraction of particles (Pn) suspended in the flow as a function of time for varied settling velocities (V_s).

Materials and Procedures. Oka strain of varicella was obtained from the Biken Institute, Japan. MRC-5 cells were obtained from ATCC or NIBSC and used to generate a manufacturer's working cell bank (MWCB) at Merck. MEM was manufactured by Merck and supplemented at the time of use with fetal bovine serum (FBS), neomycin, and glutamine. Eagle's basic medium (BME) was purchased from Gibco. Horseradish peroxidase–conjugated donkey antigoat antibody (affinity-purified F[ab′]2 fragment) was obtained from Jackson Immunoresearch Labs (West Grove, Pa. [Cat# 705-036-147]). Goat antivaricella antiserum was obtained from Lampire (Lot 2200). 3,3-Diaminobenzidine (DAB) was obtained from Sigma, prepared as a 25 mg/mL solution in water, and stored at −20 °C. Immediately before use, DAB stock solution was thawed and diluted with a peroxide/nickel enhancing solution (Amersham) in phosphate buffer.

A freshly trypsinized, concentrated cell suspension was placed in roller bottles with MEM growth medium supplemented with 10% FBS, 50 μg/mL neomycin, and 2 mM glutamine. The bottles were incubated at 37 °C and rotated at the desired rate. After 4–6 days (depending on the process used), the roller bottles were refed with fresh nutrient

medium. Roller bottle cultures with an established cell monolayer on the roller bottle surface were infected by the addition of freshly trypsinized, varicella virus-infected MRC-5 cells. The infected cells were added to the liquid nutrient medium in the roller bottle 6–8 days after the initial roller bottle cell plant. Particle counting and size analysis utilized an Elzone particle size analyzer model with a 48-μm orifice.

An in situ immunostaining procedure was used. The medium was aspirated from the culture vessel (T-flask or roller bottle) using vacuum. The culture vessel was then rinsed three times with PBS with 1% BSA. After the final rinse with PBS, a fixing solution of 5% acetic acid, 5% water, and 90% methanol was added to the culture vessel and exposed to all surfaces for 15–30 min. The methanol is then aspirated from the culture vessel with vacuum. The culture vessels are then rinsed three more times with PBS, and stored with a PBS overlay at 4–8 °C until stained. Immunostaining of the fixed bottles was carried out by a dual antibody method. First, 2–3 mL of a 1:500 dilution of goat antivaricella antiserum in PBS with 1% BSA was added to the culture vessels, covering the entire growth surface of the T-flask and a 3- to 4-cm-wide strip along one side of the roller bottle. The vessels were then placed on a rocker platform and rocked gently at room temperature for 1–2 h to allow binding of the primary antibody. The culture vessels were then rinsed three times with PBS to remove unbound antibody, and 3 ml of a 1:500 dilution of the anti-goat antibody was added. The vessels were again incubated at room temperature with gentle rocking for 30–45 min. Unbound secondary antibody was then removed by three PBS rinses. Then, 10 mL of peroxidase substrate solution was added to each culture vessel and the vessels were incubated at 37 °C until foci of infected cells appeared. The peroxidase solution was then aspirated from the vessels with vacuum and the vessels again rinsed three times with PBS. The stained areas were left under a PBS overlay, and the vessels were stored at 4–8 °C. Infectious foci were counted using a microscope at 10× magnification.

Varicella plaque assay was conducted in the following manner: MRC-5 cells were grown in 60-mm plates with 5 mL/plate of BME with 10% FBS, 0.2% glutamine, and 0.05% neomycin; 10 plates were prepared per sample with 5 plates per dilution. The plates were then incubated at 35.5 °C with a 3% CO_2 gas environment for 2 days. Before addition of the samples, the cell growth medium was removed and 5 mL/plate of MEM with 2% FBS, 0.2% glutamine, and 0.05% neomycin was added to each plate. Diluted virus samples of 0.1 mL were added to each plate. Samples were diluted to ensure a plaque counting range of 15–80 plaques per plate. The plates were mixed well and incubated for 7 days at 35.5 °C with a 3% CO_2 gas environment. Following incubation, the medium was removed by aspiration and 2 mL of 0.2% Coomassie blue stain in 100% ethanol containing 1% v/v glacial acetic acid was added to each plate. After 15 min at room temperature, the stain was removed and the plate was rinsed with cold water. The plates were inverted to dry. Plaques were then counted in each plate using a light box to assist in visualizing the plaques.

Results. In the first study, the kinetics of virus-infected cell sedimentation and attachment were studied at three rotation rates: 0 (T-flask stationary culture), 1/4 rpm (roller bottle culture), and 1/2 rpm (roller bottle). A monolayer of MRC-5 cells was established on the surface of the T-flasks and roller bottles. Following growth of these cells to confluence, varicella-infected MRC-5 cells in nutrient medium were added to the T-flasks or roller bottles and rotated at 0, 1/4, and 1/2 rpm. Samples of the supernatant were collected at various times from 0 to 36 h and were measured for infected cell number remaining. At the same time intervals, flasks and roller bottles were sacrificed and stained with a varicella-specific monoclonal antibody marker to determine the number of infectious foci on the cell monolayer. Infectious foci appear when an infected cell has attached to, fused with, and infected a cell attached to the roller bottle surface followed by infection of neighboring cells. Thus, the disappearance of infected cells from the supernatant could be correlated to the appearance of infectious foci on the cell monolayer (where it is assumed that one infected cell equals one infectious focus).

The infectivity of varicella-infected cells degrades exponentially at culture temperatures. To optimize the use of this material and to provide for as synchronous an infection as possible, the time that the infected cells remain in solution is to be minimized. Figure 15 demonstrates that as the rotation rate is reduced the fraction of infected cells remaining in the supernatant decreases at a given time. This finding is further supported by the increased appearance of infectious foci on the cell monolayer (Fig. 16) as the rotation rate is decreased, thus establishing the correlation between number of infected cells disappearing from the supernatant with the increase in infectious foci. Additionally, a mass balance was performed to demonstrate that as a function of time the appearance of infectious foci equaled the initial number of cells in the supernatant minus the intrinsic degradation rate in solution (degradation rates were determined from independent experiments).

In the second study, lower rotation rates were used to further optimize the roller bottle rotation rate for the infection of the cell monolayer. Rotation rates lower than 1/8 rpm were not selected because the cell monolayer remains without nutrient medium coverage for too long, resulting in poor cell growth and viability. Thus, the following three rotation rates were selected in the second study:

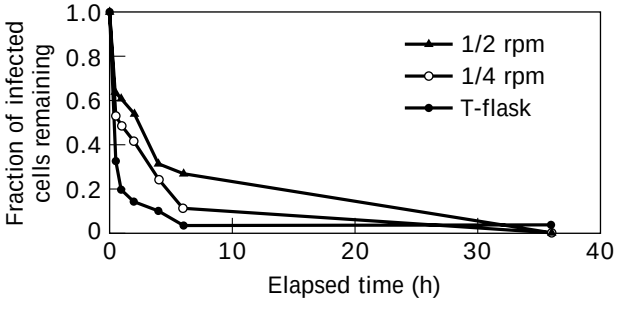


Figure 15. Fraction of cells remaining in the supernatant as a function of roller bottle rotation rate (rpm).

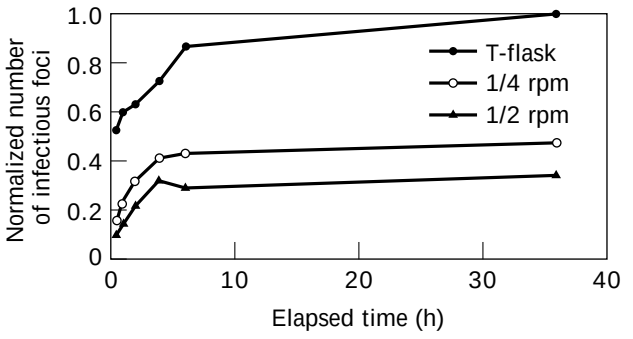


Figure 16. Quantity of infectious foci on the cell monolayer as a function of rotation rate (rpm).

0 (T-flask stationary culture), 1/8 rpm (roller bottle culture), and 1/4 rpm (roller bottle). In Figure 17, the disappearance of infected cells correlates with a decrease in rotation rate, further supporting the hypothesis.

In the final study performed, cell growth rates and final cell densities were determined as a function of roller bottle rotation rate to evaluate if rotation rate was a critical parameter in achieving high cell densities once the cells attached to the culture vessel surface. Two rotation rates were evaluated in this study: 1/4 and 1/2 rpm. Higher cell numbers were achieved when higher rotation rates were utilized (Fig. 18). The higher cell densities as rotation rate is increased may be attributed to improved transport of nutrients, gases, and waste products to and from the bulk

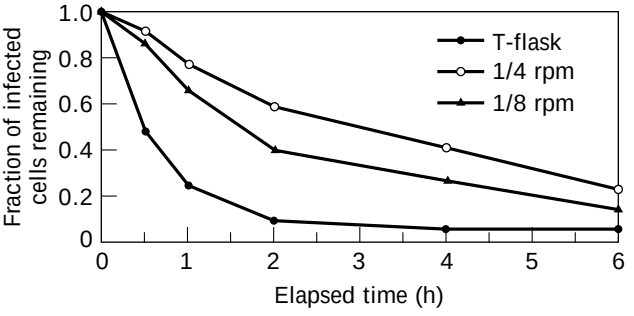


Figure 17. Fraction of infectious foci on the cell monolayer as a function of rotation rate, for lower rotation rates.

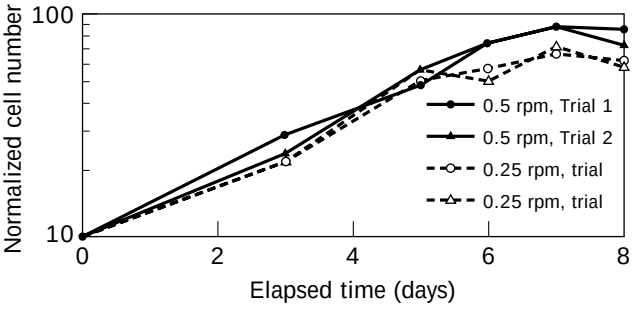


Figure 18. Cell growth rates and final cell densities as a function of rotation rate.

solution to and from the cell monolayer. The results of this experiment pointed to the need to further optimize the cell proliferation steps of the process. These studies were carried out, and the resulting process has shown much greater productivity (data not shown).

The liquid flow modeling shows good qualitative agreement with the experimental data. In both cases, the settling dynamics are nonlinear, in which the rate at which particles are removed from the suspension starts out high and decreases thereafter. In addition, the model correctly predicted that decreases in rotation rate resulted in improved sedimentation of cells onto the cell monolayer.

IMPROVING THE PERFORMANCE OF ROLLER BOTTLES

Using Unsteady Rotation to Enhance Settling

Comparison of Figure 3 and Figure 12 indicates that incomplete settling is caused by the strong recirculation present in the flow, which in two-dimensional cuts of the flow field causes the flow to behave as sets of nested closed loops. When a particle is in the lower part of the flow (below the stagnation point), it falls through loops toward an outer loop. However, if the particle does not reach the wall, it then moves to the upper part of the flow and then falls toward an inner loop. The combination of these two effects causes the particle to move up and down in an closed orbit. In fact, without Brownian motion a particle can exhibit only one of two phenomena: settle to the bottom of the bottle or move in a closed orbit.

Intuitively, such closed orbits can be disrupted by perturbing the flow in a time-dependent manner, perhaps resulting in enhanced settling. This hypothesis was tested by examining settling in time-periodic flows, which are the simplest and most easily controllable time-dependent flows. There are different ways of introducing time periodicity into the flow. A simple approach is to reverse the flow direction periodically. For the creeping flow used in this study, transient effects that occur when the flow reverses direction are negligible because inertial terms in the Navier–Stokes equations are very small; therefore, the periodic flow can be simulated using two identical steady flows that are mirror images of each other, each one acting for a predetermined amount of time in an alternating fashion.

Although mixing of passive tracers will not improve at all in such a periodic flow because the tracers would simply move back and forth along the same streamlines, settling of heavy particles is greatly enhanced by periodic reversal of the direction of rotation. The settling behavior of particles depended strongly on the value of the flow period T, defined as the total displacement of the bottle wall (measured in dimensionless units, where one unit is equal to half the depth of the liquid) between a pair of reversals (i.e., a flow period). For example, if the bottle completes a rotation between reversals, the flow period is two total rotations and one obtains a value of $T = 25.13$. For some values of T, complete settling can be achieved in short times even for a much smaller settling velocity than those studied for steady flow.

The particle settling rate in the periodic flows was studied by placing 6,540 particles uniformly in the flow domain.

As before, particles that hit the wall were considered to be settled at the wall. Two small values of settling velocity were studied, $V_s = 0.002$ and 0.005. The particles were followed for 30 periods for each case, and the fraction of particles suspended in the flow was plotted against time. The results for $V_s = 0.005$ are shown in Figure 19, which demonstrates that the settling processes for all values of T are linear. At small values of T, $T < 25$, almost all particles settle very fast, and the smaller the value of T, the faster the settling. However, at $T = 25$, almost no settling occurs, and the same behavior is observed for $T = 50$. The particle settling rate increases from small T to $T = 25$ and then decreases until $T = 35$. The settling rate for $T = 35$ and 40 are about the same during most of the flow process, except at the later stage when settling in the $T = 40$ case continues at a linear rate while settling in the $T = 35$ case slows down. Therefore, the fastest settling is achieved at small T or between $T = 35$ and $T = 40$. Of those two conditions, the latter is preferable because for $T < 25$, the bottle does not complete a rotation between reversals, and therefore some regions of the bottle wall never touch the liquid. Such a condition would lead to drying and lysis of the cell monolayer in such regions of the bottle wall and therefore to reduced process efficiency.

Using Rocking Motion to Enhance Axial Mixing

As mentioned, the flow patterns revealed by simulations suggest slow fluid mechanical mixing, especially in the axial direction. This expectation was confirmed experimentally by using a pH indicator (bromothymol blue) and an acid–base neutralization to study mixing behavior (29). The experiments were carried out on an apparatus that consisted of a set of rollers whose rotation speed and direction were accurately controlled by a computer (Fig. 20). To achieve flows within the creeping regime, glycerin was used as the working fluid. The glycerin initially was mixed with the indicator and 0.5 mL 1 M NaOH, causing the glycerin to turn a deep blue color. With the roller bottle on the apparatus, approximately 20 mL of glycerin was then removed from the bottle with a syringe and mixed with 1.0 mL of 1 M HCl, turning the fluid yellow. This fluid was then injected back into the roller bottle near the back wall (the bottom of the bottle when upright) with a syringe and carefully mixed with the surrounding fluid at the base of the bottle using a specially modified spatula.

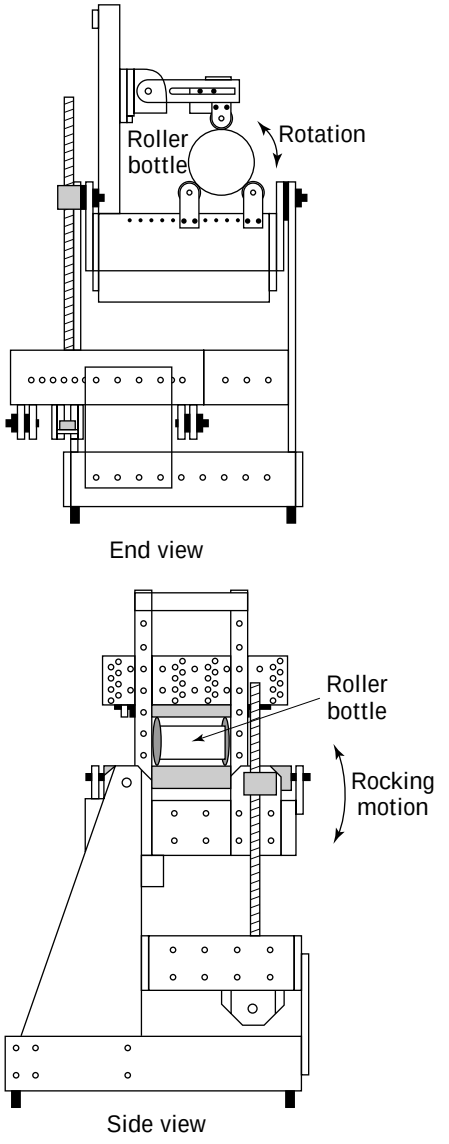


Figure 20. Experimental apparatus used to investigate mixing enhancement by introduction of rocking motion.

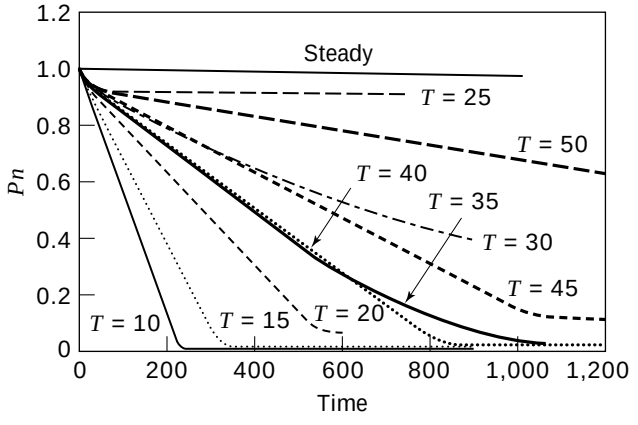


Figure 19. Fraction of particles (Pn) suspended in the flow as a function of time for varied period lengths. The settling velocity for all figures was the same, $V_s = 0.003$ U.

After the acidic glycerin was injected back into the roller bottle, the mixer was turned on and the bottle was rotated at constant speed. Figure 21a–c shows the results of this experiment. The initial condition is shown in Figure 21a, mixing after 32 revolutions in Figure 21b, and 64 revolutions in Figure 21c. These figures clearly reveal the effect of the symmetry plane located in the center of the bottle. After 32 revolutions, mixing within each half of the bottle is complete. This mixing results from the axial motion of fluid caused by the end-wall effects. However, even after 64 bottle revolutions, mixing is minimal between the two halves of the bottle; neutralization of the base on the right-hand side will occur only through the action of diffusion.

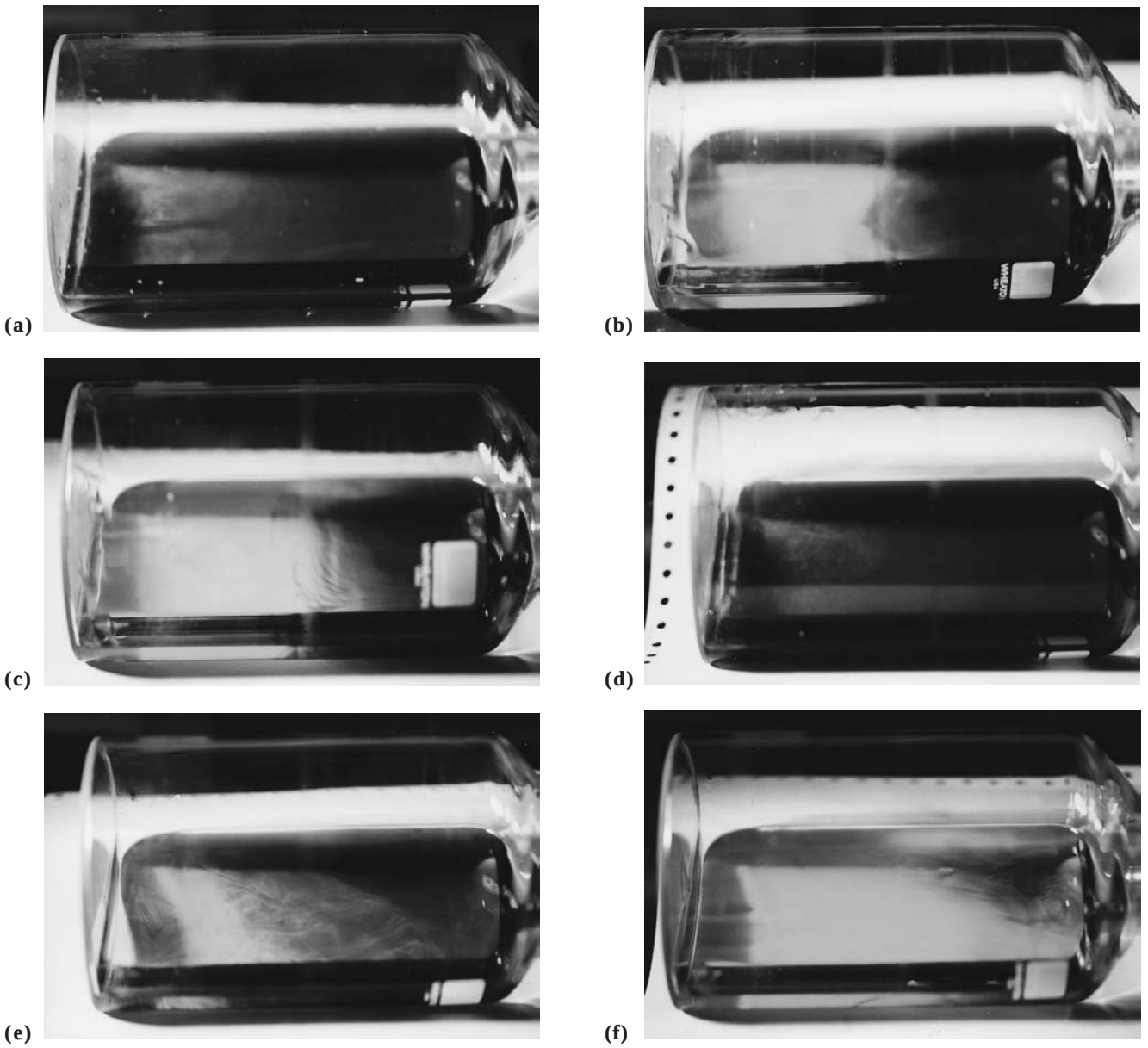
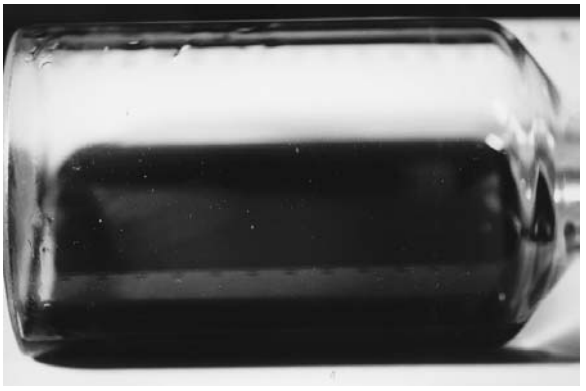
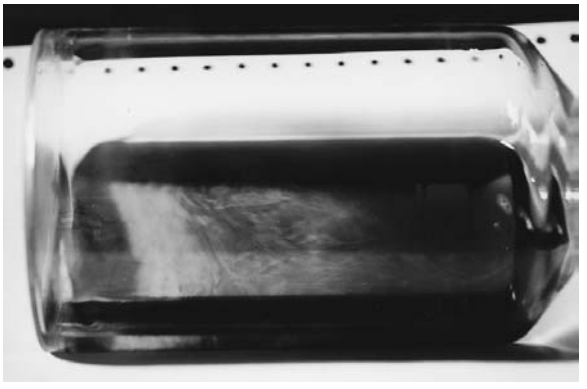


Figure 21. Results from mixing experiments performed (**a–c**) without rocking motion, corresponding to (**a**) the initial condition, (**b**) 32 revolutions, and (**c**) 64 revolutions; (**d–f**) with rocking motion at amplitude of 12 degrees, frequency equal to 1.6 revolutions per rock, corresponding to (**d**) the initial condition, (**e**) 32 revolutions, and (**f**) 64 revolutions; and (**g–i**) with rocking motion at amplitude of 12 degrees, frequency equal to 3.2 revolutions per rock, corresponding to (**g**) the initial condition, (**h**) 32 revolutions, and (**i**) 64 revolutions.

Recent experiments dealing with mixing of dry powders in a rotating and rocking cylinder have shown that a small-amplitude rocking motion in the vertical direction can greatly enhance mixing in the axial direction (30). This approach was implemented here to attempt to enhance mixing in the roller bottle. Two conditions were examined: (*1*) mixing at a rocking frequency of 1.6 bottle revolutions per rocking cycle and (*2*) mixing at a rocking frequency of 3.2 bottle revolutions per rocking cycle.

Figure 21d–f represents the effect of rocking on mixing in the system. The roller bottle was rocked approximately 12 degrees at a rocking rate of 1.6 bottle revolutions per rock, and mixing improved substantially in the system as a result of the rocking. Similarly, Figure 21g–i represents mixing for a rocking rate of 3.2 bottle revolutions per rock; rocking indeed enhanced mixing, which is essentially complete after 64 revolutions. Moreover, because rocking disrupts the recirculating flow patterns described earlier, it is also likely to disrupt cyclical cell motions and enhance settling. It is important, however, to realize that if rocking were used to enhance mixing in a cell growth process, the choice of rocking frequency and amplitude should take into


(g)


(h)

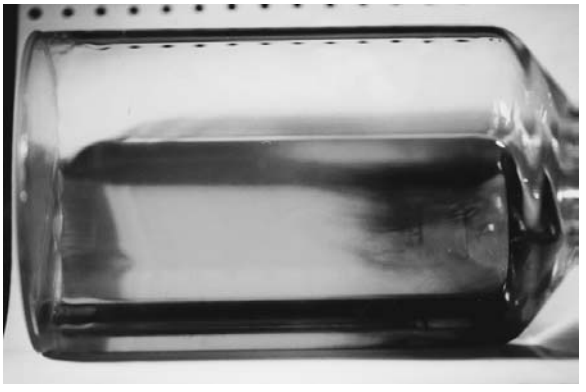

(i)

Figure 21. Continued.

account the need to keep the cells submerged in the nutrient medium as uniformly as possible. If an integer number of revolutions per rocking cycle were chosen, then the cells at the top of the bottle when the rocking reached its maximum angle of inclination would always be the same and would be exposed to nutrients for a substantially shorter fraction of time than those at the bottom. The rocking frequencies of 1.6 and 3.2 revolutions per rocking cycle were chosen because these frequencies should keep the cells fairly uniformly exposed to the nutrients, although the cells closer to the neck of the bottle will still be exposed less to the nutrients.

CONCLUSIONS

The three-dimensional flow in a roller bottle was simulated using a finite-volume method and used to study the particle settling behavior in the bioreactor. From the steady flow simulations of cell motion in two-dimensional cuts of the velocity field, it was shown that a large number of particles in the flow remain trapped in circulating loops and never reach the walls of the roller bottle. Only for large settling velocities (i.e., $V_s = 0.2\,U$) do most particles reach the bottle walls. Thus, in many situations, poor mixing and settling of cells could occur.

It was shown that settling could be enhanced by destroying the closed-orbit motion of the cells. A simple time-dependent flow that periodically reversed the rotation direction was an effective approach for this purpose. Particles were found to settle faster for small values of the flow period T, but such short periods are undesirable because such conditions fail to wet the entire bottle wall, which is essential to keep the cell monolayer intact and viable. Longer periods were also shown to be effective for enhancing settling, and nearly optimal values were identified.

Major limitations in mixing performance in the axial direction were caused by the presence of a vertical plane of symmetry at the center of the bottle. Such limitations, however, can be readily overcome by introducing a small-amplitude vertical rocking motion that disrupts the plane of symmetry, leading to much faster and complete axial mixing. The frequency of such motion was shown to have a significant effect on mixing rate. Additional work is needed to identify optimum values of frequency and amplitude.

BIBLIOGRAPHY

1. H. Tanaka, F. Nishijima, M. Suwa, and T. Iwamoto, *Biotechnol. Bioeng.* **25**, 2359–2365 (1983).
2. H. Tanaka, *Process Biochem.* **22**, 106–115 (1987).
3. C.Y. Hong, T.P. Labuza, and S.K. Harlander, *Biotechnol. Prog.* **5**, 137–143 (1989).
4. Y.A. Elliot, *Bioprocess Technol.* **10**, 207–216 (1990).
5. V.G. Kalthod, *Novel Carrier and Reactor for Culture of Attachment Dependent Mammalian Cells*, D.Sc. Thesis, Washington University, St. Louis, Mo., 1991.

6. E.I. Tsao, M.A. Bohn, D.R. Omstead, and M.J. Munster, *Ann. N.Y. Acad. Sci.* **665**, 127–136 (1992).
7. R. Pennell and C. Milstein, *J. Immunol. Methods* **146**, 43–48 (1992).
8. E. Olivas, B.B.D.-M. Chen, and W.S. Walker, *J. Immunol. Methods* **182**, 73–81 (1995).
9. R. Singhvi, J.F. Markusen, B. Ky, B.J. Horvath, and J.G. Aunins, *Cytotechnology* **22**, 79–87 (1996).
10. R. Kunitake, A. Suzuki, H. Ichihashi, S. Matsuda, O. Hirai, and K. Morimoto, *J. Biotechnol.* **52**, 289–294 (1997).
11. E. Van Hemert, D.G. Kilburn, and A.L. Van Wezel, *Biotechnol. Bioeng.* **11**, 875–885 (1969).
12. C. Horng and W. McLimans, *Biotechnol. Bioeng.* **17**, 713–732 (1975).
13. R.E. Spier and J.P. Whiteside, *Biotechnol. Bioeng.* **18**, 649–659 (1976).
14. D.W. Levine, D. Wang, and W.G. Thilly, *Biotechnol. Bioeng.* **2**, 821–845 (1979).
15. J.J. Clark and M.D. Hirtenstein, *J. Interferon Res.* **1**, 391–400 (1981).
16. B.J. Montagnon, B. Fanget, and A.J. Nicolas, *Dev. Biol. Stand.* **47**, 55–56 (1981).
17. V.G. Edy, *Adv. Exp. Med. Biol.* **172**, 169–176 (1984).
18. E. Rivera, C.G. Sjosten, R. Bergman, and K.A. Karlsson, *Res. Vet. Sci.* **41**, 391–400 (1986).
19. R.M. Gallegos Gallegos, E.L. Espinosa Larios, L.R. Ramirez, R.K. Schmid, and A.G. Setien, *Arch. Med. Res.* **26**, 59–63 (1995).
20. R. Archer and L. Wood, *Proceedings of the 11th Annual Meeting of the European Society for Animal Cell Technology*, Brighton, U.K., September 2–6, 1991.
21. A.K. Majumdar, R.H. Whitesides, S.L. Jenkins, and D.L. Bacchus, *J. Propulsion and Power* **6**, 5–10 (1990).
22. V. Dupont, M. Pourkashanian, A. Williams, and R. Wooley, *Fuel* **72**, 497–504 (1993).
23. J. Seth and W.R. Wilcox, *J. Cryst. Growth* **114**, 357–363 (1991).
24. D.Y. Tang, J.J. Ou, R.H. Heist, S.H. Chen, and A.J. Dukat, *Ind. Eng. Chem. Res.* **32**, 1727–1733 (1993).
25. P.D. Swanson, F.J. Muzzio, A. Annapragada, and A. Adjei, *Int. J. Pharmaceutics* **142**, 33–52 (1996).
26. D.M. Hobbs and F.J. Muzzio, *Chem. Eng. Sci.* **53**, 3199–3214 (1998).
27. R.J. Adrian, *Annu. Rev. Fluid Mech.* **23**, 261–304 (1991).
28. M.R. Maxey and J.J. Riley, *Fluids* **26**, 883–889 (1983).
29. D.J. Lamberto, F.J. Muzzio, and A.L. Tondovich, *Chem. Eng. Sci.* **51**, 733 (1996).
30. C. Wightman, R.M. Paul, F.J. Muzzio, R.E. Riman, and E.K.Gleason, *Powder Technol.* **84**, 231–240 (1995).

See also AMMONIA TOXICITY, ANIMAL CELLS; CULTURE MEDIA, ANIMAL CELL, LARGE SCALE PRODUCTION.

SAMPLING METHODS (REACTORS, CONTAMINATION)

M.P. NANDAKUMAR
H. HÅKANSON
B. MATTIASSON
Lund University
Lund, Sweden

KEY WORDS

Biosensors
Coaxial catheter
Expanded bed
Flow injection analysis
On-line analysis
On-line cell disintegration
Sample handling
Sampling
Sampling probes

OUTLINE

INTRODUCTION

To optimize fermentation processes, that is, to ensure maximum growth and/or productivity, the fermentation conditions should be controlled and maintained at their optimum values. However, the media for fermentations and other bioprocesses are very complex in nature. These are often multiphase systems. The liquid phase contains soluble nutrients (e.g., sugar, proteins, vitamins, salts, and products of metabolism such as amino acids, alcohols, and organic acids). Furthermore, there may be a solid phase of suspended material (insoluble substrate/product, carrier materials, etc.). The gas phase is in equilibrium with dissolved gases in the medium. Analysis of any of these compounds by traditional techniques is tedious and time-consuming and cannot provide on-line information of the process. Recently, the increasing demand of high-value products has focused interest on well-controlled, optimized processes. This has initiated the development of new on-line measurement techniques (1–3). The ideal situation would be to have specific probes that can operate in situ during the fermentation. However, at present there is only a limited number of such probes available for measuring pH, dissolved oxygen, and carbon dioxide. There is still no suitable technology for operating biosensors containing a biological macromolecule. In the last few years research on biosensors for bioprocess control has been growing. These sensors offer the ability to measure concentrations of various biomolecules, as well as biological parameters, very specifically using immobilized enzymes or cells or other biochemical/biological entities. Unfortunately, these sensors are inactivated by steam sterilization (4). Because sensor technology is not yet sufficiently advanced to permit this type of monitoring, the majority of analysis is still carried out outside of the bioreactor, a process requiring sampling and sample-handling strategies. This review highlights the importance of sampling and sample handling and the techniques used.

STRATEGIES FOR SAMPLING

For analysis of any bioprocess, off-line or on-line sampling is a prerequisite and is the basis for the final analytical result. The following strategies/criteria are important when designing a sampling device (5).

1. The sterility of the process must be maintained during sampling.
2. The delay between sampling and analysis may result in changes in the composition of the sample. Precautions must be taken to prevent or compensate for that.
3. The analysis must be accurate and reproducible.
4. When constructing the sampling and sample treatment unit, the selection of material must be governed by compatibility criteria such that the process being monitored is not negatively influenced.
5. Protein adsorption and cell colonization on the sampling and sample-handling device may lead to an imparted mass transport and formation of unwanted immobilized catalytic activities that may severely influence the composition of the sample before analysis can be carried out.

Considering these parameters, a sampling system for on-line monitoring of fermentation processes must meet the following criteria.

1. The system can take representative samples from fermenter but should not disturb the fermentation process.
2. The system should be sterilizable and provide a sterile barrier to avoid contamination.
3. Separation of cells, solids, or macromolecules should be carried out with as short a delay as possible within the sampling devices (6).

There are many different types of sampling units available to achieve these goals (7). This review highlights a detailed description of some of the sampling devices used for on-line monitoring of bioprocesses, especially when biosensors are used in the analytical step. It should, however, be stressed that the sampling and sample handling as described here will be applicable when setting up any kind of analysis to monitor a bioprocess. Figure 1 presents a schematic description of different sampling modes.

NATURE OF THE SAMPLE

The character of the reaction mixture to be sampled determines the choice of sampling method. Different kinds of analytes require different sampling protocols.

Volatile Compounds and Compounds Present in Gaseous Phase

Volatile compounds may be present in the head space of a fermentation or be dissolved in the liquid medium. At equilibrium, the concentration in the head space reflects that in the liquid. The composition of the gas phase is usually less complex than that of the aqueous phase; therefore, when the metabolites are volatile, it is probably simpler to monitor the concentration in the gas phase. The risks of infection of the fermentation are lower when sampling from the gas phase. This has made it attractive to use analysis of the gas phase when monitoring industrial fermentations.

There are two main modes described in the literature for the sampling of gases and volatile compounds:

1. Analysis of the outlet gases or of the headspace gases.
2. Analysis of the volatile gaseous compounds dissolved in the culture medium.

A sampling method is applied that is based on a technology for supplying oxygen to bioprocesses that are disturbed by bubbling oxygen. Oxygen is supplied in a tubing made of hydrophobic material. The oxygen penetrates the walls of the tubing and is released to the cultivation broth. Similar phenomena can take place for gaseous compounds with an opposite concentration gradient (8). Tubings made from materials that are permeable to organic solvent molecules are immersed in the fermentation broth, and an inert gas that flows through the lumen of the tubings transports the solvent molecules to an analyzer (e.g., a gas-sensitive semiconductor sensor placed downstream of the tubing) (9,10).

Recently a new approach for monitoring carbon dioxide during fermentation was reported based on a piezoelectric crystal (11). A change in the oscillation frequency of the crystal occurs due to selective sorption of a gaseous compound to the surface, and the change in frequency is directly proportional to the mass deposited. Fast response, simple use, and low cost may lead to a more extensive use of this technique. More details on sampling of gases during fermentations are described in the literature (12,13).

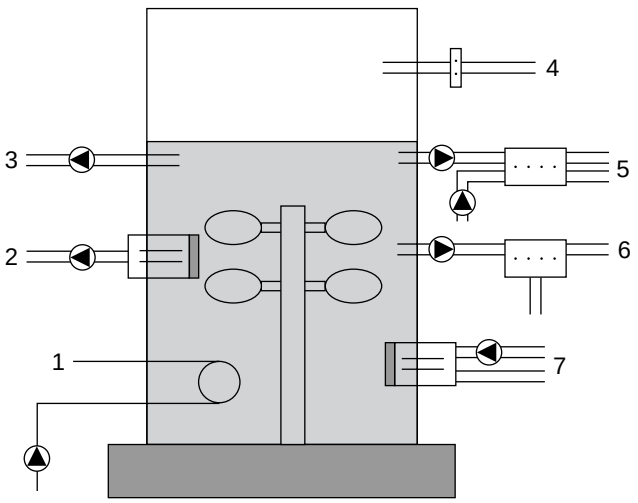


Figure 1. Schematic presentation of different modes of sampling from a bioreactor. The different concepts are (1) dialysis, (2) a membrane unit for removing a particulate-free sample from the turbulent liquid around the impeller, (3) direct sampling, (4) sampling of headspace, (5) sampling and filtering the sample outside the fermenter using a receiving buffer to catch the permeate, (6) sampling with filtering where the permeate is collected, (7) a membrane dialysis probe with the receiving buffer being pumped through the unit.

Analysis of Nonvolatile Compounds

Analytes Dissolved in the Liquid Medium. Automatic sampling, either continuously or by a frequently repeated procedure, is a prerequisite for on-line or "at-line" monitoring. The sampling system normally serves also as a sterile barrier between the bioreactor and the analytical system. Different analytical systems can be used. Flow injection analysis (FIA) is a common and convenient way to carry out the analysis (14). The samples are heterogeneous, with compounds dissolved in the medium, some forming complexes with other constituents, some not fully soluble, some particulate, and so forth. Manual sampling and off-line analysis are still in practice; however, they result in delays in analysis, risks of infection, and less reproducible sample handling (15). Sampling can be done using sampling units mounted either inside or outside the bioreactor. Several systems of both kinds have been reported recently (16,17). The advantages and disadvantages of internal and external sampling modules are presented in Table 1.

Table 1. Advantages and Disadvantages of Internal and External Sampling Modes

Mode of sampling	Advantages	Disadvantages
In situ modules	Short answering time, simple design, less contamination, suited for coupling on on-line analyzers	Problem of long-term stability, not exchangeable, adhesion of microorganisms and fouling, positioning is crucial, size limitation
External modes	Exchangeable, suited for coupling of on-line analysers	Complex design, longer answering time, fouling, shear stress

Direct Sampling. Direct sampling from a bioreactor is possible through a catheter with a small inner diameter. It is possible to connect this sample stream directly to a FIA system or any other analytical device. However, for precise measurements, the metabolic activity in the sample must be stopped immediately after sampling, otherwise unwanted reactions may take place during the period after sampling but before analysis. If a deactivation is carried out with addition of an inhibitory agent (18), one has to examine the possibility of interference from the reagent on the analysis (19).

One device that addresses these challenges is the coaxial catheter unit (20,21), which is connected to the bioreactor for direct sample withdrawal. It consists of two concentrically placed tubings, an inner and an outer (Fig. 2); the outer protrudes beyond the inner a few millimeters. The sampling is done via the inner tubing. An inhibitor is fed into the space between the tubings, and the broth plus the inhibitor are mixed in the small mixing zone at the tip of the catheter. Inhibition thus takes place at the moment of sampling. A major advantage of the direct sampling system is that a complete sample is obtained, and it is therefore possible to measure the biomass concentration on-line (22) and even perform analysis on intracellular compounds (23,24). The coaxial catheter has not yet been used for sample withdrawal from cultures of filamentous fungi and highly viscous fluids. This method is so far limited to studies of unicellular organisms.

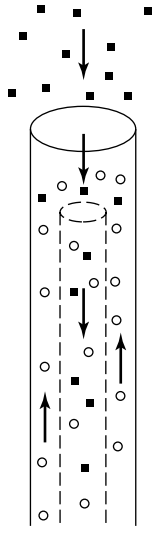


Figure 2. The coaxial catheter. The unit consists of two concentrically placed catheters, the inner being a few millimeters shorter that the outer. Sample is sucked into the inner and is transported away for analysis. The space between the two catheters is used to transport an inhibitor solution toward the tip of the coaxial catheter, where it is mixed with the sample and then withdrawn. If the inner catheter sucks at, for example, double the rate of the inhibitor feed, one volume inhibitor is mixed with one volume of sample and then transported to analysis.

Membrane-Based Sampling Techniques. The sampling system acts as an aseptic barrier between the bioreactor and the analytical equipment. The basic idea behind this type of sampling device is to utilize a membrane for separation during the sampling process. In most cases separations based on differences in size are carried out. The membrane-based sampling techniques can also be applied to cultures of filamentous fungi and in other viscous media. The major disadvantage of the membrane module is clogging of the membrane, mainly due to deposition of proteins on the membrane surface. Reviews on membrane modules are present in the literature (3,25,26).

When using a membrane sampling module, there are three types of separation forces that can be utilized: pressure-driven difference, concentration difference, and electrical potential difference. Applying any of these forces, transport over the membrane will occur and may serve several purposes:

1. Cells are removed from the sample, thus stopping the metabolic activity.
2. Compounds that may interfere with the analysis are retained.
3. Membrane sampling acts as a barrier between the analytical system and the bioreactor.

The details of four major types of membrane separation processes, used in combination with FIA, are presented in Table 2.

Membrane modules can be categorized into two groups: those placed in a recycle loop connected to the bioreactor (filtration outside the reactor), and those placed in situ (as with dialysis filtration). When the membrane unit is placed outside the reactor, fermentation broth is usually pumped through the membrane unit and then recycled to the fermentor. Some characteristics both pro and con of such a device are listed in Table 3.

What is special with the external loop membrane unit is that upon passage of the fermentation broth through the loop, special conditions may prevail. Thus, oxygen limitations may be severe, and, damaging shear forces may appear. This type of device is thus less suited to dealing with high-cell-density fermentations of aerobic cells or with

Table 2. Membrane Separation Procedures Used for Sampling

Mode of separation	Particles separated	Driving force
Filtration	Water and dissolved species	Gravity
Microfiltration	Water and dissolved species (possible to separate unicellular organisms)	Pressure difference
Ultrafiltration	Water and salts (separates biologicals, colloids, and macromolecules)	Pressure difference
Electrodialysis	Charged molecules and ions	Electrical potential
Dialysis	Low molecular weight species	Concentration difference

Table 3. Advantages and Disadvantages in Using a Membrane Unit Placed Outside the Bioreactor

Pro	Con
Replacement of membrane possible	Oxygen depletion may appear in the external loop
Quick response	Shear forces
	Risk of infection

shear-stress-sensitive cells. The great advantage to using an external membrane is the greater flexibility. Thus, a membrane may be easily replaced when needed.

There are two basic modes of operation: the permeate may be collected and used for analysis, or alternatively, a receiving buffer is pumped through the unit on the collection side of the membrane to collect substances passing the membrane. It is important that analyte collection is carried out under sterile conditions. This is usually secured by intermittent sterilization (27–29).

The membranes used for external filtration are made of a variety of materials (e.g., nylon, polypropylene, acrylnitril copolymer, cellulose acetate, cellulose nitrate, polyvinylpyrroliden fluoride) and are available in different pore sizes. Normally, membranes with pore sizes of 0.2 μm are preferred because they guarantee the separation of medium from the biomass and other suspended material (30). Sampling units with external filters are available commercially. The Bio-Pem (B. Braun, Melsungen, Germany) consists of a chamber (vol. 175 mL) containing a magnetically stirred rod. The stirrer ensures a tangential flow across the membrane, which is placed in the bottom of the chamber (31). This unit was successfully used for sample withdrawal during a cultivation process involving unicellular microorganisms but was found unsuitable for sampling cultures of fungi. A major draw back of the Bio-Pem is the large dead volume in the membrane module, which may cause substrate depletion in the loop as well as high response times. An alternative method involves a cross-flow filtration module in which the dead volume is less, resulting in a much shorter response time ($<$10 s). This module is applied efficiently in fed-batch penicillin fermentations on complex media (32) and for wastewater treatment (33).

There is still another unit based on a planar membrane (Fig. 3). The membrane is clamped between two metal blocks, which are then screwed together. During filtration, the incoming broth (feed) is pumped into grooves in the upper block, and the filtrate is collected in the lower chamber. The broth is separated into two streams: the particular-free filtrate (permeate), which passes the membrane, and the particulate-enriched retentate, which is returned to the fermentor. The filtrate is then led to an injection loop for later injection into the analytical system. After passing the loop, two possibilities arise: the filtrate is either returned to the bioreactor or sent to the waste collector. The latter would quickly empty a small fermentor. To avoid this, the filtrate can be recycled to the fermentor, but this increases the risk of contamination. Placing a sterile barrier (e.g., a microfilter) between the injection loop and the bioreactor reduces the risk (34). This may be acceptable in microbial fermentations, but in certain types of processes, such as human cell line production, such a strategy cannot be applied because such cells are very slowly growing and must be cultivated during long time periods and at the same time have a high value. An alternative approach is called discontinuous filtration. The filtrate line is closed between injections and opened shortly before the injection to flush the tubings and the loop (35).

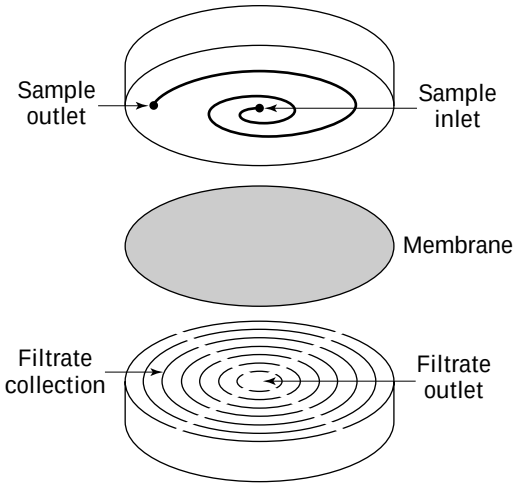


Figure 3. Schematic presentation of the flux membrane dialysis module. The membrane is mounted between the two holders, and sample is introduced on one side. The filtrate is collected and transported for future analysis.

There are several units with planar membranes available on the market. These units have been used for different applications. The A-SEP TM module (Applicon, Schiedam, The Netherlands) has been used for monitoring acetate and phosphate during *Escherichia coli* fermentations (36), utilizing a standard 47-mm-wide 0.45-μm MF membrane. The Minitan TM system (Millipore, Bedford, Mass.), using two membranes, has been employed for the determination of glucose in *Candida rugosa* cultivations

(37), followed by several others for the determination of glucose, diols, and ethanol (38,39).

Membrane filters may also be available in the form of hollow fibers of different diameter and porosity (Fig. 4). Depending on the intended application, the cartridge contains different numbers of hollow fibers. In accordance with the cross-flow filtration principle, the medium is pumped at high velocity (1 m s^{-1}) through the fiber. A pressure is applied, and the filtrate is collected on the outer side of the fibers. It is also possible to pump the broth on the outer side of the fiber(s) and collect the filtrate inside (40). However, depending on the membrane geometry, this latter may be less attractive. When dealing with anisotropic membranes it is advisable to apply the complex medium to the smooth side of the membrane.

A range of different applications are described in the literature. A stainless steel module with only one fiber inside has been reported (41). The fiber had an inner diameter of 5.5 mm and a length of about 200 mm. The dead volume of the module was 2.5 cm^3; and it was successfully used for sampling in conjunction with determination of ammonium, glucose, and phosphate. Reports on the use of several in-house built hollow-fiber modules can be found in the literature, covering vastly different applications in mammalian cell culture (42), lignocellulose hydrolysate fermentations (43), and biodegradation of naphthalene sulphonic acid (44). Systems are also available commercially from Bio-FIO (Glasgow, U.K.) and used for monitoring different types of fermentation components (45,46).

Despite the potential complications associated with operating an external loop filtration system, it remains a popular approach. The great advantage is that when membrane clogging occurs, it is possible to intervene without ruining the fermentation.

In dialysis, a membrane is used to separate a complex solution from a more well defined solution. Pores in the membrane allow exchange of small molecules between the two liquid volumes. The driving force is concentration difference. The different modes of dialysis originate from whether the donor and acceptor phases are stagnant or moving. In equilibrium dialysis both are stagnant and, at equilibrium, the concentration in both phases will be the same of those molecules able to pass the membrane. In continuous dialysis, compounds passing the membrane are continuously removed. This results in a larger concentration gradient, and a higher flux is maintained.

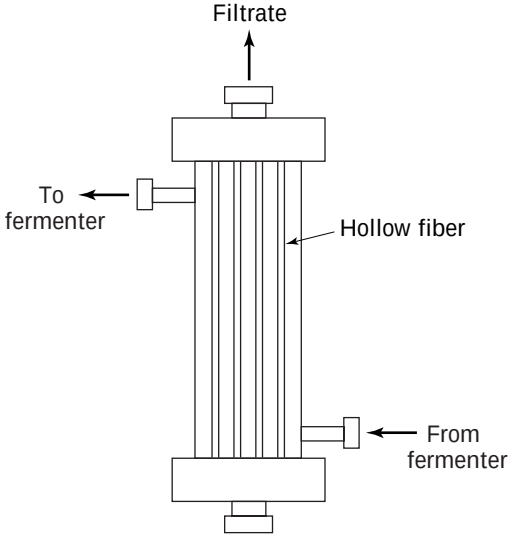


Figure 4. Schematic presentation of the hollow-fiber modules for sampling. A perfusion buffer is passed through the lumen of the fibers, and the fermentation broth is introduced on the shell side of the unit.

The membranes used for dialysis are classified according to their molecular weight cut-off. Most dialysis membranes are made from cellulose materials, but because many organisms are able to hydrolyze cellulosic materials, this limits their applicability. There are several sampling units that incorporate dialysis membranes. They are usually placed in the fermentor, and a receiving dialysis carrier buffer is circulated on the permeate side of the membrane. The carrier buffer is thus pumped into the bioreactor and, after having received the dialysate molecules, removed again. The major problem associated with this approach is the deposition of cells and fouling of the membrane on the fermentor side. The accumulated layer of cells may reduce the mass transfer from the fermentor to the accepting solution and also cause decreases in concentration of metabolites passing through the membrane because they run the risk of being metabolized by the cell layer before entering the dialysis membrane. To help reduce fouling it is possible to place the dialysis probe into the zone of the fermentor where the impeller creates the greatest turbulence (47).

Another means of obtaining similar results, but with fewer design restrictions, is to construct a dialysis probe that besides containing the membrane unit also has a stirring function, so that a high tangential liquid flow is generated over the membrane. This approach was used by Mandenius et al. (48) in constructing a system for use in dense particulate suspensions. The membrane is placed in the front of the probe, and membrane fouling is avoided by means of the rotating magnet driven by an electrical motor inside the probe (Fig. 5).

Zabriskie and Humprey (49) developed a dialysis probe situated within one baffle of a stirred fermentor. They were able to take samples for glucose analysis during cultivation of yeast cells on a synthetic medium. The disadvantage of this dialyzer is that it can be used only in this type of bioreactor. Another sampling device was presented by Glaxo Laboratories, Ltd. (50). This unit consist of an elongated support with a milled helical groove. A tubular membrane made of cellulosic materials or polytetrafluoroethylene was fixed to the support using ring seals.

An in situ dialysis membrane module was developed at the University of Hannover and is marketed by Advanced Biotechnology Corporation (ABC). The module consists of a unit supporting the membrane, and one or more extension units. The membrane module is usually inserted through a standard 19- or 25-mm port in the top of the bioreactor or, alternatively, through a 25-mm port below the liquid surface. When inserted from the top of the bioreactor, one or more extension tubes are needed to place the membrane below the liquid surface. Liquid passing the

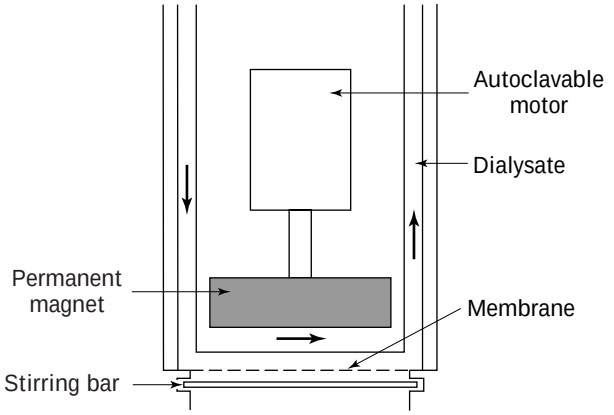


Figure 5. Cross section of the dialysis probe. The stirring bar is rotated by the rotating magnetic field when the motor is on. Receiving solution is pumped over the inner side of the membrane where dialysate is received and removed for analysis *Source*: From Ref. 48.

membrane is collected in 16 grooves, and these 16 permeate streams are joined in four openings placed evenly along the membrane holding unit. Through these four openings the permeate enters the central channel (10 mm i.d.), and a continuous sample stream to the analyzer leaves the unit from the top of the probe (51).

Gibson and Woodward (52) have also described an internal dialysis module, based on the same principle but without a stirrer. Ethanol and glucose could be sampled with this unit, and the results from the subsequent analyses agreed well with the on-line data.

Another modified technique is the microdialysis used for continuous in vivo sampling, which has been described in reviews (3,53,54). The microdialysis sampling unit is a probe, and it is a miniaturized form of a dialysis unit. It is manufactured in several different configurations, one is seen in Figure 6.

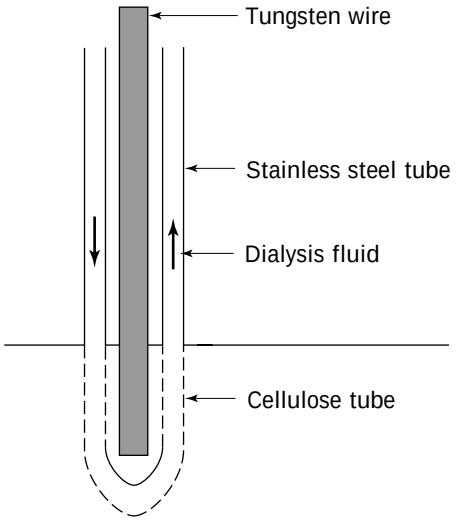


Figure 6. The microdialysis unit.

Typical flow rates of the perfusion liquid are 0.5–25 μL/min, and to avoid pulses in the flow, a peristaltic pump or preferably a syringe pump should be used. Due to its ease of operation and the general approach, the number of applications increases rapidly. It can easily be coupled on-line to a number of separation and detection techniques. This technique was used in fermentation processes for the production of ethanol, penicillin, and so on (55). However, the slow flux of liquid and thus the long response time from an analysis based on this sampling technique, compared with the conventional sensing devices, should be considered.

The major problem in dialysis units is the fouling and clogging by macromolecules on the membrane surface. This can be overcome by adding molecules that compete for the binding site of the protein, adding a displacer (56), changing the pH (57), or pumping the solution tangentially over the membrane surface. An excellent review on dialysis for flow injection analysis has been published (26).

MODES THAT AVOID OR AT LEAST REDUCE THE NEED FOR SAMPLE TREATMENT

As already noted a major concern in sample handling is clogging. Quite often packed-bed systems are used, and then particulate matter constitutes a major obstacle. In these cases, particulate matter has to be removed prior to analysis. However, with the introduction of expanded-bed chromatography, it is possible to handle large-scale chromatographic processes without separating out the cells (58). This made it tempting to investigate the possibility of using the same approach in an analytical system.

Enzyme-based analysis was studied using glucose oxidase immobilized to Streamline® particles. These particles are characterized by being of heterogeneous size and density distribution, such that when used in an expanded bed, a stable expanded bed can be obtained. Glucose was monitored using this technique (59). A more challenging issue is whether the same system would allow binding reactions to take place. If so, one should be able to run immunoassays in whole broth and maybe even in cell homogenates. Assays of human serum albumin in the presence of increasing amounts of heat-killed yeast cells was used as a model system. Antibodies against albumin were present on the Streamline® (60). The assays seemed successful, and therefore the model was also studied to see if the intracellular enzyme β-galactosidase could be quantified in a cell suspension or a cell homogenate. The results were so encouraging that it is safe to say that this technique allows binding assays without the preceding separation of particulate matter from the sample.

PRETREATMENT OF SAMPLES

Another vital part of the analysis is the pretreatment of the samples. This includes dilution, centrifugation, extraction, and filtration, depending on the nature of the samples. Certain samples can be analyzed without pretreatment; however, for complex samples extensive clean-up is required before introduction into the flow system. This topic has been discussed in several reviews (61,62). The

reproducibility and rapidity of flow injection analysis promotes its applicability for on-line pretreatment of samples with less manual intervention (63–65) (Fig. 8). Additionally pretreatment steps by FIA are rapid. The pretreatment and analysis, including chemical conversion, usually require less than a minute. Furthermore, when samples are toxic or pathogenic, FIA offers the advantage of removing, or at least minimizing, human exposure. An excellent review on sample pretreatment by FIA was published by Clark et al. (66) and includes a description of the theoretical background of sample pretreatment.

Disintegration of Cells

Monitoring levels of intracellular compounds has so far been possible only via off-line analyses. However, after the recent development of an on-line disintegration system that can be used together with expanded-bed adsorption in a flow system, a new possibility seems to have emerged. The important issues are that the whole broth is taken and treated and after that the whole homogenate can be processed in, for example, a binding assay.

The whole broth is introduced into an FIA system where it first passes an expanded bed containing Streamline® with immobilized lysozyme. On passage of *Micrococcus lysodeiticus* cells it has been proven that a high degree of disintegration is achieved, as judged from the amount of protein released (23). When adding a subsequent step involving sonication using a flow cell, the disintegration is even more efficient. The experimental set-up is shown in Figure 7. A culture of *E. coli* was studied and its level of β-galactosidase production was monitored. The detection step consisted of an expanded bed containing Streamline® with immobilized antibodies to β-galactosidase.

Upon introduction of the whole broth, homogenization takes place by treatment in the two units; and afterward liberated enzyme is trapped by the immunosorbent. The enzyme is quantified by reading the amount of *o*-nitrophenol liberated after passing a pulse of *o*-nitrophenyl-β-galactopyranoside over the sorbent. After the assay the immunocomplex is broken and the system is ready for another assay (24).

Centrifugation

Centrifugation is commonly used for separating cells and particulate materials from the fermentation broth. It is possible to include a small-scale centrifuge in an FIA system, with recycling of the sample to the reaction vessel (67).

Liquid–Liquid Extraction

Liquid–liquid extraction is based on the partitioning of the analytes between an aqueous and an organic phase and is widely used for gas chromatographic analysis. Integrating this technique into FIA is difficult because it is difficult to automate, laborious, and expensive. Solid-phase extraction (SPE) is based on sorption of analytes in solution to a sorbent packed in a cartridge or reactor, typically the size of a precolumn (10–20 $\times$ 2–4.6 mm in length and i.d., respectively) (68). While the analyte is trapped into the packing material, compounds with little or no affinity for the sorbent pass through the column. After washing the column to further remove the interfering substances, the analytes are then desorbed and eluted by a small volume of an appropriate solvent. Recently, Vreuls et al. (69) demonstrated the coupling of on-line SPE to gas chromatography. A precolumn packed with mixed anion/cation exchange materials was used on-line in a flow system for removal of ionic substances in aliquots of a filtrate from fermentation.

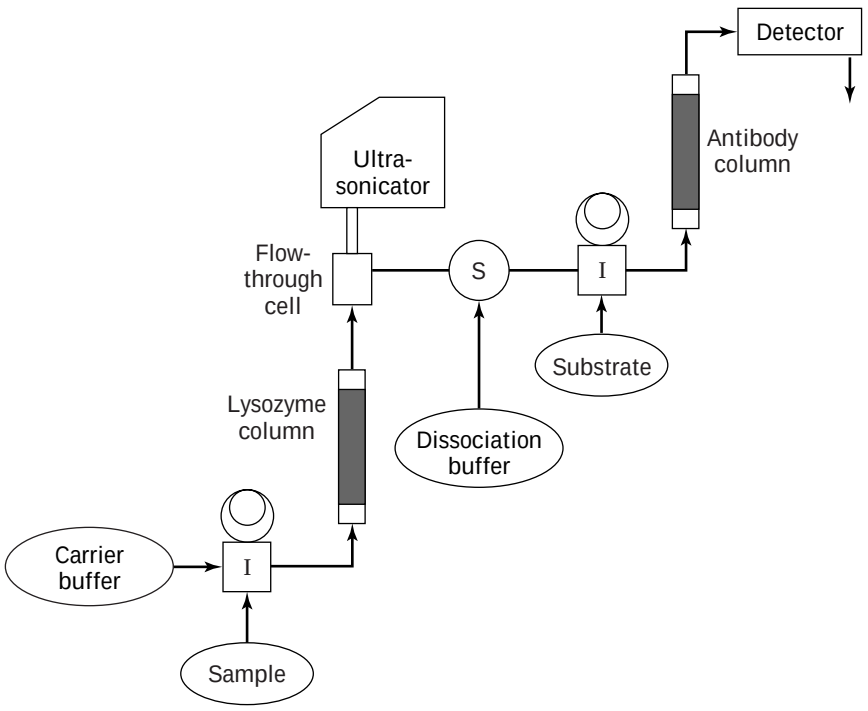


Figure 7. Schematic presentation of a complete flow system, including an expanded flow column with immobilized lysozyme and ultrasonicator flow cell. An expanded column with immobilized antibodies is also included for the subsequent ELISA.

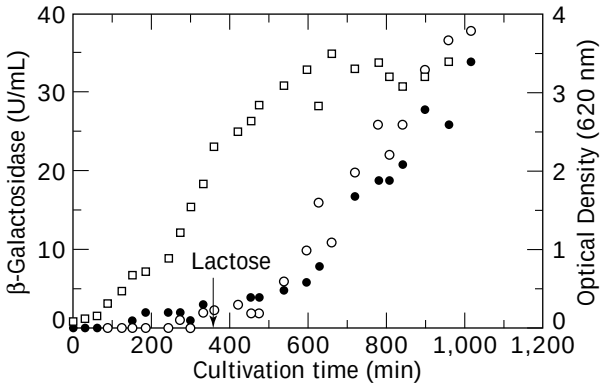


Figure 8. Monitoring of a shake-flask experiment. On-line (●) and off-line (○) *β*-galactosidase activity and o.d. (□) measurements. At the arrow the inducer lactose was added.

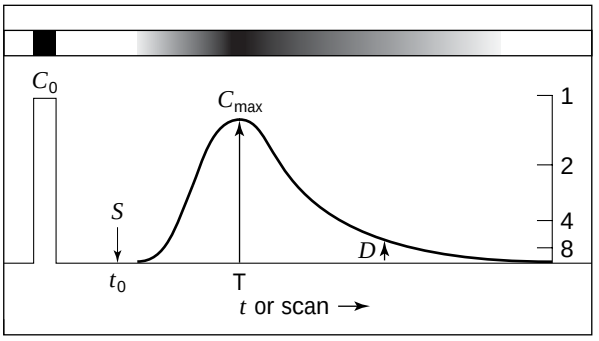


Figure 9. In addition to peak height a dispersed sample zone yields an infinite number of vertical readouts. The initial concentration, C_0, is injected at time t_0, and the pulse appears as a continuum of concentrations as exemplified by the peak. At time T, C_{max} is the highest value read, but a reading could be done anywhere along the peak (e.g., a diluted value at point D).

Preconcentration Techniques

If the analyte concentration is less than the concentration range of the detection method, the sample must be concentrated. Most of the manual pretreatment methods are time-consuming and expensive. Microcolumns with ion exchange materials can be used to significantly reduce the volume of the analyte by adsorbing the sample into a solid support in the column and by a further elution (70).

Dilution of Samples

If the analyte concentration is more than that of the detectable range of the analytical system, the sample must be diluted. The same goes for a sample in which inhibiting or otherwise disturbing substances are present in high concentrations. Manual dilution is used for off-line analysis. On-line dilution is now widely applied for fermentation processes (66) and includes the use of gradient dilution, zone sampling, and cascade dilution.

In the gradient dilution, an originally homogeneous sample zone disperses during its movement through the FIA conduit. The zone undergoes a change from an initial plug shape with a concentration C_0, to an exponentially modified gaussian shape with a maximum concentration in the zone of C_{max}. If the peak maximum does not lie in the calibration range of the measurement, the zone can be made at a time on the profile that yields a decreased response. By selection of measurement times longer than the time of peak maximum (T in Fig. 9), an element with a decreased concentration is measured, yielding a proportionally decreased absorbance. This measurement is reproducible from sample to sample and can be calibrated.

In zone sampling a small portion of the sample zone injected into a second flowing stream results in dilution (71,72). In cascade dilution the sample volume is reduced through stream splitting using differential pumping coupled with a true dilution by merging with a subsequent carrier stream.

Sample Injection

Sample handling and pretreatment can be achieved by using many techniques in FIA. The method of injection is equally important. The purpose of sample injection in FIA is to deliver a well-defined sample bolus reproducibly into a continuously flowing carrier stream. Ruzicka and Hansen (73) described that changing the injected sample volume is a powerful way to change the dispersion. An increase in peak height and sensitivity of measurement is best achieved by increasing the volume of the injected sample solution. Conversely, dilution of overly concentrated sample material is achieved by reducing the sample volume.

There are two methods of injection: time-based injection (a sample loaded for a precise time at a given flow rate) and volume-based injection (the sample fills a geometrically defined volumetric cavity). This well-defined sample bolus is then inserted into the carrier stream. Volume-based injections are performed using rotary valves, commutators, and hydrodynamic injections. A detailed methodology used for injection has been published (74).

Calibration of Sampling System

For any kind of analysis, whether it is on-line or off-line, the system needs to be calibrated to ensure that the compound of interest is being measured reliably. It is easy to calibrate an off-line system. However, the continuous systems described so far commonly have no possibility for calibration and recalibration on-line without disrupting the monitoring procedure. Thus, any changes occurring in the sample during the recalibration, are not registered (5). Another important limitation is the fact that the calibration solution and the sample solution are usually not treated in the same manner.

In-line calibrating systems are used in medical monitoring. Attempts were made for in-line calibration during fermentation with *E. coli* using a double-lumen catheter. The calibration was done by closing the sample inlet from the fermentor and substituting it with a sample of known analyte concentration (20) (Fig. 10). A further development of this technology was the use of the coaxial catheter for sampling. In on-line monitoring of glucose and lactate in blood, the calibration solution was introduced as a pulse

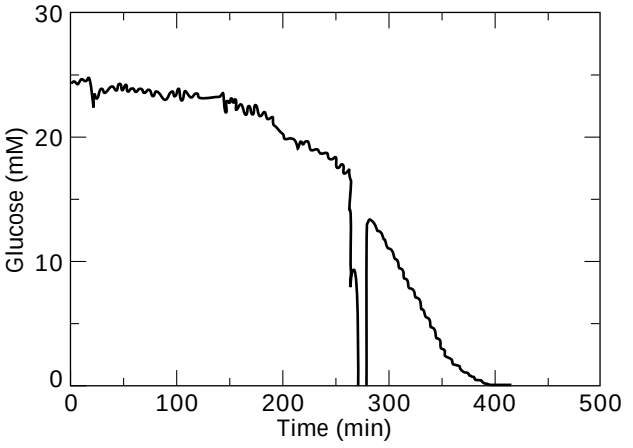


Figure 10. Monitoring of glucose in batch cultivation of *Escherichia coli*. After approx. 270 min a calibration solution is introduced, and the reading is back approx. 10 min later. *Source*: From Ref. 20.

in the heparin stream fed via the outer space in the coaxial catheter. This led to an on-line monitoring of the process without having to interrupt the reading (Fig. 11). Furthermore, by introducing the calibration solution at the tip of the sampling catheter, the whole system is calibrated. Thus a more true calibration value is obtained (75,76).

CONCLUSION

Sampling and sample treatment are extremely important when applying good analytical technique to the monitoring and control of biotechnological processes. Too often much effort is spent on developing a sensitive assay, but very little or no concern has been given to sampling and sample handling. The examples given in this article clearly demonstrate that many good techniques can be applied. Which method to use will be governed by the system studied, the equipment available, and also the earlier experience of the researcher.

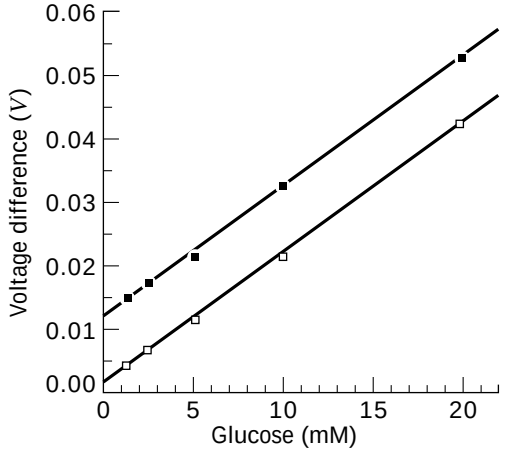


Figure 11. On-line outlet signal (V) as a function of glucose concentrations (mM); (■) glucose standard solution with added internal standard (5 mM), $y = 1.14 \times 10^{-2} + 2.05 \times 10^{-3}\,x$, $r^2 = 1.000$; (□) glucose standard solution, $y = 1.09 \times 10^{-3} + 2.04 \times 10^{-3}\,x$, $r^2 = 1.000$.

BIBLIOGRAPHY

1. D.J. Clarke, B.C. Blake-Coleman, M.R. Calder, R.J.G. Carr, and S.C. Moody, *J. Biotechnol.* **1**, 135–158 (1984).
2. R.T. Hatch, *Annu. Rep. Ferment. Proc.* **5**, 231–235 (1982).
3. N.C. van de Merbel, H. Lingeman, and U.A. Th. Brinkman, *J. Chromatogr. A* **725**, 13–27 (1996).
4. I. Karube, T. Matsunaga, S. Suzuki, T. Asano, and S. Hob, *J. Biotechnol.* **1**, 286–291 (1984).
5. B. Mattiasson and H. Håkanson, *Trends Biotechnol.* **11**, 136–142 (1993).
6. Hun.-C. Shu, Ph.D. Thesis Lund University, Lund, Sweden, 1994, ISRN LUTKDH/TKBT–94/1020-SE.
7. T. Buttler, Ph.D. Thesis Lund University, Lund, Sweden, 1996, ISRN LUNDKL/(NDKAK-1034).
8. K.D. Vorolop, J.W. Becke, and J. Klein, *Biotechnol. Lett.* **5**(8), 509–512 (1983).
9. J.H. Lee, J.C. Woodard, R.J. Pagan, and P.L. Rojers, *Biotechnol. Lett.* **3**, 351–356 (1981).
10. C.F. Mandenius and B. Mattiasson, *Eur. J. Appl. Microbiol. Biotechnol.* **18**, 197–200 (1983).
11. O.F. Fiho, J.F. de Andrade, A.A. Suleiman, and G.G. Guilbault, *Anal. Chem.* **61**, 746–748 (1989).
12. T. Buttler, C. Nilsson, L. Gorton, G. Marko-Varga, and T. Laurell, *J. Chromatogr. A* **725**, 41–56 (1996).
13. T. Lorenz, W. Schimidt, and K. Schugerl, *Chem. Eng. J.* **35**, B15–B22 (1987).
14. B. Mattiasson and B. Danielsson, (1996) in R.F. Taylor and J.S. Schultz eds., *Handbook of Chemical and Biological Sensors*, IOP Publishing, Philadelphia, pp. 533–552.
15. J. Nielson, *Process Qual. Control.* **2**, 371–384 (1982).
16. N. Vriezen and J.P. Van Dijken, *Biotechnol. Tech.* **11**(6), 427–430 (1997).
17. T. Kullick, U. Bock, J. Schubert, T. Scheper, and K. Schügerl, *Anal. Chim. Acta* **300**, 25–31 (1995).
18. K. Nikolajsen, J. Nielsen and J. Villadsen, *Anal. Chim. Acta* **214**, 137–145 (1988).
19. J. Nielsen, K. Nikolajsen, and J. Villadsen, *Biotechnol. Bioeng.* **33**, 1127–1134 (1989).
20. O. Holst, H. Håkanson, A. Miyabayashi, and B. Mattiasson, *Appl. Microbiol. Biotechnol.* **28**, 32–36 (1988).
21. H. Håkanson, M. Nilson, and B. Mattiasson, *Anal. Chim. Acta* **249**, 61–65 (1991).
22. S. Benthin, J. Nielsen, and J. Villadsen, *Anal. Chim. Acta* **247**, 45–50 (1991).
23. M.P. Nandakumar, A. Tokaj, and B. Mattiason, *Bioseparation*, in press.
24. A. Tokaj, M.P. Nandakumar, O. Holst, and B. Mattiasson, *Bioseparation*, in press.
25. K. Schügerl, in H.J. Rehm, G. Reed, A. Puhler, and P. Stadler eds., *Biotechnology*, 2nd ed., vol. 4, VCH, Weirheim, 1991, pp. 152–156.
26. J.F. van Staden, *Fresenius J. Anal. Chem.* **352**, 271–302 (1995).
27. K.H. Kroner and N. Papamichel, *Process Biochem.* **23**(10), 3–4 (1991).

28. A. Reckenwald, K.H. Kroner, and M.R. Kula, *Enzyme Microb. Technol.* **7**, 607–612 (1985).
29. M. Garn, M. Gasin, C. Thommen, and P. Cevey, *Biotechnol. Bioeng.* **34**, 423–428 (1989).
30. T. Lorenz, W. Schimidt, and K. Schügerl, *Chem. Eng. J.* **35**, B15–B22 (1987).
31. K.H. Kroner and M.R. Kula, *Anal. Chim. Acta* **163**, 3–15 (1984).
32. L.H. Christensen, J. Nielsen, and J. Villadsen, *Chem. Eng. Sci.* **46**(12), 3304–3307 (1991).
33. K.M. Pedersen, M. Kummel, and H. Soeberg, *Anal. Chim. Acta* **238**, 191–199 (1990).
34. N.C. van de Merbel, I.M. Kool, H. Lingeman, U.A.Th. Brinkman, A. Kolhorn, and L.C. de Rijke, *Chromatographia* **33**, 525–532 (1992).
35. T.A. Butler, A.J. Kristina Johanson, L. Gorton, and G.A. Markovarga, *Anal. Chem.* **65**(19), 2628–2636 (1993).
36. L.F. Forman, B.D. Thomas, and F.S. Jacobson, *Anal. Chim. Acta* **249**, 101–111 (1991).
37. F. Valero, J. Lafuente, M. Poch, C. Sola, A.N. Aranjo, and J.L.F.C. Lima, *Biotechnol. Bioeng.* **36**, 647–651 (1990).
38. M.J. Hayward, T. Kotiano, A.K. Lister, R.G. Cooks, G.D. Austin, R. Narayan, and G.T. Tsao, *Anal. Chem.* **62**, 1978–1804 (1990).
39. U. Brand, B. Reinhardt, F. Ruther, T. Scheper, and K. Schügerl, *Sensors Actuators* **4**, 315–318 (1991).
40. H. Hoffman, Ph.D. Thesis, University of Hannover, Hannover, Germany, 1985.
41. K. Schügerl, *Anal. Chim. Acta* **213**, 1–9 (1988).
42. S.S. Ozturk, J.C. Thrift, J.D. Blackie, and D. Nareh, *Biotechnol. Bioeng.* **48**, 201–208 (1995).
43. C.F. Mandenius, *Biotechnol. Bioeng.* **32**, 123–129 (1988).
44. J. Meschke, H. Bennemann, M. Herbst, S. Dormeier, and D.C. Hempel, *Bioprocess Eng.* **3**, 151–157 (1988).
45. N.C. van de Merbel, H. Lingman, U.A.Th. Brinkman, A. Kolhorn, and L.C. de Rijke, *Anal. Chim. Acta* **279**, 39–50 (1993).
46. N.C. van de Merbel, P. Zuur, Frijlink, J.J.M. Holthuis, H. Lingeman, and U.A.Th. Brinkman, *Anal. Chim. Acta* **303**, 175–186 (1995).
47. T. Lorenz, W. Schimidt, and K. Schügerl, *Chem. Eng. J.* **35**, B15–B22 (1987).
48. C.F. Mandenius, B. Danielsson, and B. Mattiasson, *Anal. Chim. Acta* **163**, 135–141 (1984).
49. D.W. Zabriskie and A.E. Humphery, *Biotechnol. Bioeng.* **20**, 1295–1301 (1978).
50. U.S. Pat. 3,830,106 (1974), Glaxo Laboratories, Ltd.
51. L.H. Christensen, J. Nielsen, and J. Villadsen, *Anal. Chim. Acta* **249**, 123–136 (1991).
52. T.D. Gibson and J.R. Woodward, *Anal. Chim. Acta* **213**, 61–68 (1988).
53. H. Benreniste and P.C. Huttemeier, *Prog. Neurobiol.* **35**, 365–372 (1991).
54. T.E. Robinson and J.B. Justice, Jr., Eds., *Microdialysis in the Neurosciences*, Elsvier Science, Amsterdam, 1991.
55. G. Marko-Varga, T. Buttler, L. Gorton, and C. Gronstervall, *Chromatographia* **35**, 285–289 (1993).
56. N.C. van de Marbel, J.M. Teule, H. Lingeman, and U.A.T. Britman, *J. Pharm. Biomed. Anal.* **10**, 225–231 (1992).
57. J.D.H. Cooper, D.C. Turnell, B. Green, and F. Verillon, *J. Chromatogr.* **456**, 53–69 (1988).
58. H. Chase, *Trends Biotechnol.* **12**, 296–303 (1994).
59. M.P. Nandakumar, A. Lali, and B. Mattiason, (1998) Submitted.
60. B. Mattiason and M.P. Nandakumar, (1998) Submitted.
61. M.C. Hennion, *Trends. Anal. Chem.* **10**, 317–322 (1991).
62. U.A.Th. Brinkman, *J. Chromatogr. A* **665**, 217–231 (1994).
63. S. Olsen, J. Ruzicka, and E.H. Hansson, *Anal. Chim. Acta* **136**, 101–112 (1982).
64. E. Martins, M. Bengtsson, and G. Johansson, *Anal. Chim. Acta* **169**, 31–42 (1985).
65. B. Karlsberg and S. Thelander, *Anal. Chim. Acta* **98**, 1–7 (1978).
66. G.D. Clark, D.A. Whitman, G.D. Christian, and J. Ruzicka, *Crit. Rev. Anal. Chem.* **21**(5), 357–375 (1990).
67. N.F. Thornhill and N.B. Fish, *Biotechnol. Tech.* **7**, 19–24 (1993).
68. G.R. van der Hoff, R.A. Banmann, U.A.Th. Brinkman, and P. van Zoonen, *J. Chromatogr.* **644**, 367–373 (1993).
69. J.J. Vreulls, G.J. De Jong, R.T. Ghijsen, and U.A.Th. Brinkman, *J. AOAC Int.* **77**, 306–327 (1994).
70. J. Ruzicka and E.H. Hansen *Anal. Chim. Acta* **214**, 1–27 (1983).
71. B.F. Reis, A.O. Jacintho, J. Moratti, F.J. Krug, E.A.G. Zagatto, F.H. Bergamin, and L.C.R. Pessenda, *Anal. Chim. Acta* **123**, 221–228 (1981).
72. J. Ruzicka and E.H. Hansen, *Anal. Chim. Acta* **145**, 1–15 (1983).
73. J. Ruzicka and E.H. Hansen, *Flow Injection Analysis*, 2nd ed., Wiley, New York, 1988.
74. F.J. Krag, F.F. Bergamin, and E.A.G. Zagatto, *Anal. Chim. Acta* **179**, 103–107 (1986).
75. M. Kyrölainen, H. Håkanson, and B. Mattiasson, *Biotechnol. Bioeng.* **45**, 122–128 (1995).
76. R. Kok and P. Hoasi, *Biosensors* **3**, 89–100 (1987/88).

SCALE-UP. See DIMENSIONAL ANALYSIS, SCALE-UP; MAMMALIAN CELL CULTURE REACTORS, SCALE-UP; SCALE-UP, STIRRED TANK REACTORS.

SCALE-UP, STIRRED-TANK REACTORS

D.C. HEMPEL
H. DZIALLAS
Technische Universität Braunschweig
Braunschweig, Germany

KEY WORDS

Dispersing capacity
Mass transport
Mixing time
Newtonian fluids
Non-Newtonian fluids
Power input
Pumping capacity
Scale-up

Shearing stress
Stirred-tank reactor

OUTLINE

INTRODUCTION

The scale-up of a biotechnological process developed in the laboratory often presents problems that, owing to the many parameters involved and their complexity, do not permit any generalized solution. As a result, criteria for scale-up have always been developed on the basis of specific processes. The main problem in all considerations of scale-up is that no strictly defined criteria exist for conducting biotechnological processes and, hence, their scaled translation. Of the biological and physical factors important in scale-up, generally valid relationships can be established only for the latter. Thus in the following, the experimentally proven physical fundamentals of scale-up for conventional stirred-tank reactors are brought together in a way that reflects their importance in practical operation. Such considerations have been published in the past but were based on the scientific knowledge available at that time (1–5). This article is an updated short version. For further information about scale-ups of fermentation processes and extended literature see Ref. 5.

SCALE-UP METHODS

Balance Equation

A scale-up using a solution of the complete equations for momentum, energy, and mass balances is exceedingly difficult to determine for a complex microbiological process. Thus for a scale-up on the basis of mathematical models, there are still no practicable equations available at present that take into account the kinetics and transport processes in the stirred-tank reactor. Nonetheless, the application of simplified balance equations is often helpful in order to understand the processes involved (6,7).

Dimensional Analysis

Dimensional analysis, introduced successfully into chemical technology and already a classical method of scale-up, is also helpful when dealing with biotechnological processes. Using it, a number of dimensionless quantities (π-quantities) are derived from the relevant process-influencing quantities, which using the mathematical interrelationships of the similarity principle, provide criteria for scale-up (see also Refs. 8 and 9). The invariance of all dimensionless π-quantities relevant to the problem provides the simplest rule for scale-up, but the rule is dependent upon complete similarity. Whereas physical processes often do behave similarly upon change in scale, similarity in the presence of a chemical or microbial process is often unattainable. Even for purely physical processes, a complete similarity can often not be attained. As shown in the section "Performance Characteristics of the Stirred-Tank Reactor," the translation rule Re = idem (idem means that the numerical value of Re is held constant in scale translation) is valid for the energy input in scaling-up stirred-tank reactors, without formation of vortices. But if, nonetheless, a vortex should form, as in unbaffled vessels, then both the Reynolds number and the Froude number must retain their numerical value. From the definition of Re and Fr, it immediately becomes obvious that the requirements Re = idem and Fr = idem cannot be fulfilled because $n \cdot d^2$ = idem cannot equal $n^2 \cdot d$ = idem.

Because it is rarely possible to maintain all dimensionless characteristic quantities at the same numerical value, it becomes necessary to restrict oneself to partial similarities in which only some of the π-quantities remain fixed. To achieve this, the rate-determining step of the reaction for the complex microbial system must be known. That means that from the many process parameters, such as nutrient requirements, physicochemical conditions of growth, or transport processes, the limiting (or inhibiting) quantities have to be determined by careful analysis. Problems then arise when there is a shift in the effect of the individual limiting factors when the scale is altered. Such an approach to scale-up, in which each rate-determining step, both of the microbial reaction and of the transport processes, is mathematically correlated, has been put forward from time to time (for example, see Ref. 10) and has also been introduced highly successfully (11–14).

Maintenance of Identical Operating Conditions

A further simplification of the scale-up criteria, which follows from accepting the notion of partial similarity, is the maintenance of identical operating conditions (usually assuming that there is geometrical similarity). The most common examples of scale-up rules that have been published in the specialized biotechnological literature are based on this concept; extrapolation is often performed ac-

Table 1. Effects of Different Criteria in Linear Scaling-Up by a Factor of 5

		Scale-up criterion			
	Laboratory scale	P/V (kW/m^3)	n (s^{-1})	$n \cdot d$ (m/s)	Re
Diameter D (m)	1	5	5	5	5
Specific power input P/V (kW/m^3)	1	1	25	0.2	1.6×10^{-3}
Power input P (kW)	1	125	3125	25	0.21
Rotational speed of stirrer n (s^{-1})	1	0.34	1	0.2	0.04
Tip speed of stirrer $n \cdot d$ (m/s)	1	1.71	5	1	0.2
Reynolds number	1	8.55	25	5	1

Source: From Ref. 20.

cording to empirical rules of thumb that are still used today in the fermentation industry. Typical examples of this are the following:

- Identical volumetric power input P/V
- Identical mixing time or adaption of the mixing time to different stirred-tank reactor sizes according to purely empirical translation rules
- Identical tip speed of the agitator
- Identical volumetric mass transfer coefficient k_La
- Identical Reynolds number, Re

In fact, according to Ref. 15, about one third of plants employed the translation rules P / V = idem and k_La = idem, and about 20% used the tip speed of the stirrer (i.e., assuming shear stress). Another 20% of industrial plants employ a scale-up on the basis of mixing time and Reynolds number. The remainder scaled-up on the basis of a limiting or inhibiting substrate or product component—most commonly on the basis of the concentration of dissolved oxygen.

A scale-up founded on a single operating condition has diverse effects on the other operating quantities, as shown in Table 1 (see also Refs. 16–19). Using a linear scaling factor of 5 (corresponding to a volume scaling factor of 125) each column shows the change in the operating quantity that results from the respective scale-up criterion. Table 1 shows that it is not possible to simultaneously fulfill several of the scale-up criteria mentioned (20).

PHYSICAL FUNDAMENTALS OF SCALE-UP CONDITIONS

Performance Characteristics of the Stirred-Tank Reactor

The stirring power, P, which produces the three-dimensional flow field can be calculated from the dimensionless power number or Newton number:

$$Ne \equiv \frac{P}{\rho \cdot n^3 \cdot d^5} \tag{1}$$

From considerations of the similarity principle, the formal correlation of the performance characteristic for a given geometry is described by

$$Ne = f(Re, Fr, Q) \tag{2}$$

where the Froude number is $Fr \equiv n^2 \cdot d/g$, the Reynolds number is $Re \equiv n \cdot d^2/\nu$, and the dimensionless gasflow number is $Q \equiv q_G/n \cdot d^3$.

Power Input in Nongassed Systems of Newtonian Fluids. In systems that have no gas input ($Q = 0$) and no vortex formation (as is effectively the case in a vessel with baffles), the effect of the Froude number is negligible. The performance characteristics of the nongassed system of Newtonian fluids, which is then solely a function of the Reynolds number (21–23) is depicted in Figure 1 for a selection of stirrers. The geometrical arrangement and the working range of the stirrers have been standardized in some countries (e.g., DIN 28131 in Germany). Except for the helical and anchor stirrers which extend right to the wall, stirred vessels are equipped with baffles to avoid vortex formation. The reduction in the number of baffles from 4 to 3 or 2 generally lowers the energy input by 10 or 20% (23). In a completely turbulent system (a condition fulfilled with most stirrers at $Re > 10^4$), the performance characteristic is independent of the Reynolds number, that is, Ne = constant. If the scale-up considerations are based on a constant volumetric power input, $P/V \sim P/d^3$ = idem, then the following is valid for the scaled translation:

$$n^3 \cdot d^2 = \text{idem} \tag{3}$$

Power Input in Nongassed Systems of Non-Newtonian Fluids. In many fermentation broths, for example, those containing polysaccharides or mycelium-forming microor-

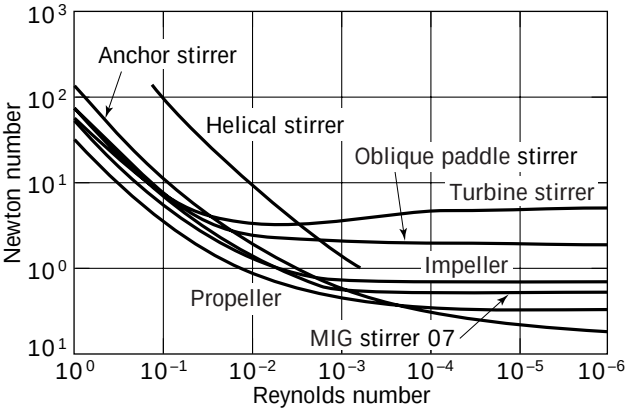


Figure 1. Performance characteristics of some types of stirrer with Newtonian fluids.

ganisms, non-Newtonian flow behavior is encountered: there is a complicated nonlinear relationship between the shearing stress, τ, and shear rate, $\dot{\gamma}$. For non-Newtonian fluids, the viscosity property can be defined only in the form of an apparent viscosity

$$\eta_a \equiv \frac{\tau}{\dot{\gamma}} \tag{4}$$

A number of models have been put forward to describe the non-Newtonian flow behavior. The most efficacious originate from the Ostwald-De Waele law for pseudoplastic fluids

$$\tau = K' \cdot \dot{\gamma}^{m'} \tag{5}$$

and from Herschel–Bulkley law for plastic fluids

$$\tau = \tau_0 + K'' \cdot \dot{\gamma}^{m''} \tag{6}$$

In order to determine the energy input in such fluids, the approach adopted by Metzner and Otto (24) and Metzner et al. (25) has proved itself. According to this concept, there is a direct proportionality between the rate of shear and the rotational speed of the stirrer

$$\dot{\gamma} = K \cdot n \tag{7}$$

whereby the apparent viscosity, η_a, is obtained from the rheological flow curve of the fluid involved. The coefficient K depends only slightly on the rheological behavior of the fluid but mainly on the stirring system, for example, $K \approx 11$ to 13 for turbine agitators and $K \approx 15$ to 25 for anchor agitators (15,26). In Figure 2 the performance characteristic of the turbine stirrer for pseudoplastic fluids is plotted versus the Reynolds number which has been modified by η_a (24,27).

A similar performance characteristic was measured by Sanchez et al. (28) and Velasco et al. (29) for a single and a dual Rushton turbine when mixing a mycelial broth under nonaerated conditions. As with Newtonian fluids, a constant Newton number is attained here at sufficiently high Reynolds numbers (i.e., the performance characteristic is independent of the viscosity). Whereas this independence of the viscosity is unequivocally valid in a nongassed system, considerable dependence on the viscosity occurs in a gassed system of non-Newtonian fluids, as is illustrated in a following section.

The apparently simple approach advocated by Metzner and Otto (24) is made more complicated by the fact that the rheological behavior of fermentation broths often changes drastically during the fermentation, as is shown in Figure 3 for a decrease of the apparent viscosity during the cultivation of *Cellulomonas uda* grown on printed newspaper (26).

Power Input in Gassed Systems at Low Viscosities. In gassed systems of aqueous-like fluids, a decrease in the stirring power input can be observed with increasing gas flow. This decrease is caused by gas cushions that form behind the stirrer blades and diminish the rotational resistance of the stirrer. The effect was first studied by Oyama and Endoh (30); the results are shown in Figure 4. Since then a series of measured values have been published; particularly extensive are the measurements made by Zlokarnik (31) and Judat (32). Judat (32) gives a relationship valid for a turbine stirrer in a turbulent region of flow ($Re > 10^4$) in water. This relationship is also applicable, in a somewhat modified form, to large vessels of up to $V = 900$ m^3:

$$Ne_G = Z \frac{Ne_0 + 187QFr^{-0.32}(d/D)^{1.53} - 4.6Q^{1.25}}{1 + 136Q(d/D)^{1.14}} \tag{8}$$

Here $Ne_0 \approx 4.9$ is the Newton number of the turbine stir-

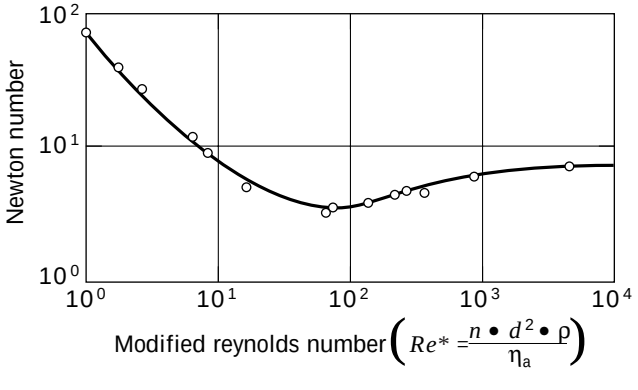


Figure 2. Performance characteristic of turbine stirrer with pseudoplastic fluid.

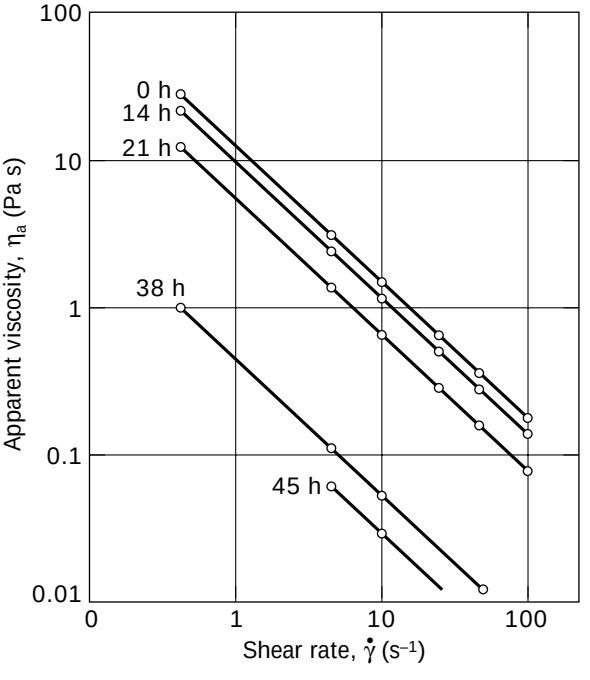


Figure 3. Time course of changes in apparent viscosity in the fermentation of cellulose.

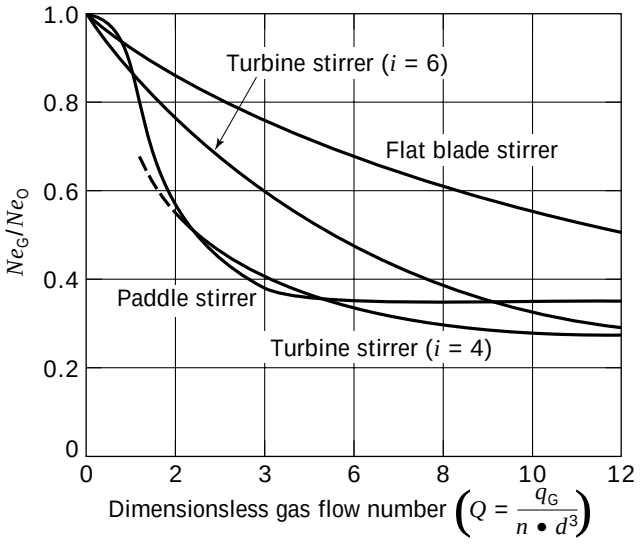


Figure 4. Power input of several types of stirrer in Newtonian fluids, dependent on the gas throughput.

ring in a turbulent region of flow without gassing, and Z is the number of stirrer blades. Equation 8 is valid for $Re > 10^4$, $Fr \leq 0.07(d/D)^3$, $0.2 \leq d/D \leq 0.42$, and $h/d > 0.75$.

Detailed investigations with other types of stirrers and multiple impellers were carried out by Schlueter and Deckwer (33). Their experimental results were well described by equation 8 with a maximum deviation of only 30%. Further investigations are necessary to describe properly the power input in stirred vessels if the impeller position is varied and multiple stirrer arrangements are used.

The influence of the Reynolds number on the power input in a gassed system has been studied by Zlokarnik (31) and is given, for the turbine stirrer, by

$$\frac{Ne_G}{Ne_0} = 1 + [(3.9Re^{0.12} + 6 \cdot 10^{-12}Re^{3.45}) \cdot (0.22Q^{0.1} + 6.25Q^3)]^{-1} \quad (9)$$

which is valid for $Q < 0.05$ and $800 \leq Re \leq 10^4$. Ne_G is the Newton number of the turbine stirrer gassed with Q at $Re > 10^4$.

Power Input in Gassed Systems of Newtonian and Non-Newtonian Fluids at Higher Viscosities. Whereas the influence of the gas on the energy input in aqueous-like fluids is about the same for the different types of stirrer, each stirrer type displays a very different effect when the viscosity increases. Detailed measurements on this have been published by Hoecker and Langer (34) and Hoecker (35). To illustrate the results, the form completely equivalent to equation 2 was chosen,

$$Ne = f(Q, Fr, Ga) \quad (10)$$

with the Galilei number

$$Ga \equiv \frac{Re^2}{Fr} = \frac{d^3 \cdot g}{\nu^2}$$

With turbine stirrers, a decrease in the power input is observed with increasing viscosity (Fig. 5). As in the correlation of the power input for nongassed systems of non-Newtonian fluids, the results with sparging can also be expressed by introducing an apparent viscosity, η_a, as results from the approach taken by Metzner and Otto (24) (i.e., equations 4 and 7). As Figure 5 shows, even the results for Newtonian glycerine–water solutions and for pseudoplastic carboxymethylcellulose–water solutions can be expressed by the same form of curve (34). With viscoelastic polyacrylamide–water solutions (Figure 5b), the decrease in the power input with viscosity is even more pronounced than with Newtonian and pseudoplastic fluids; this can probably be traced back to the development of orthogonal tension in viscoelastic fluids, which tends to damp the turbulence. The illustration of the performance characteristic for turbine stirrers in aqueous glycerine solutions in terms of the gas flow in Figure 6 clearly shows that at high viscosities a decrease in power input with the gas throughput hardly makes itself felt, but rather the Froude number exerts a more marked effect (34,35). The slight influence of the dimensionless gasflow number is due to the formation of gas cushions behind the stirrer blades, which is independent of the gas throughput.

The relationship given in equation 9 to calculate the influence of the Reynolds number on the power input with gassed turbine stirrers is roughly applicable to non-Newtonian fluids, too. In Figure 7 the results obtained by

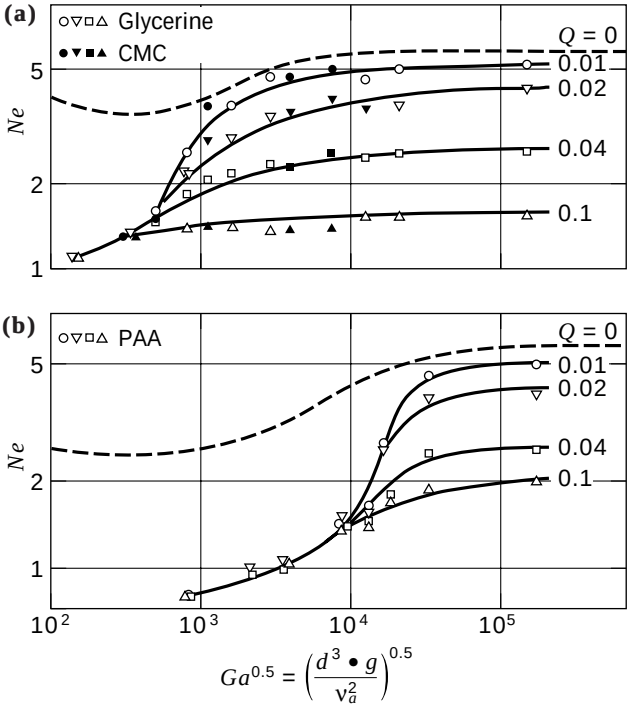


Figure 5. Power input with turbine stirrer ($d/D = 0.33$) in (**a**) Newtonian (glycerine) and pseudoplastic (CMC) and (**b**) viscoelastic (PAA) fluids.

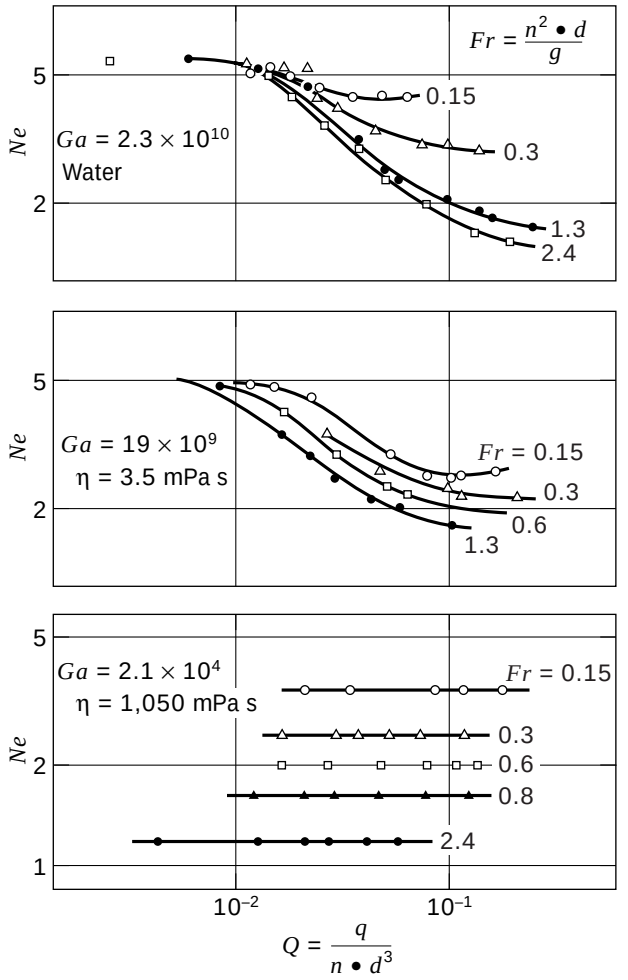


Figure 6. Power input with turbine stirrer ($d/D = 0.33$) in glycerine–water solutions.

Hoecker and Langer (34) are depicted for water, glycerine, carboxymethylcellulose (CMC), and polyacrylamide (PAA) solutions, plotted as $Ne \cdot f(Re, Q)$ corresponding to equation 9 as a function of $Q(1 + 38(d/D)^5)$. The dotted curve shows the results obtained by Zlokarnik (31). From the values taken from the line of the curve, the performance characteristics can be calculated exactly enough from Figure 7 with $Ne = Ne_G/f(Re, Q)$ with equation 9; this is the case for Newtonian and pseudoplastic fluids in the region $Ga > 10^6$ ($Re > 10^3$) and for viscoelastic fermentation broths with $Ga > 10^9$ (34,36).

Completely different behavior from that seen with a turbine stirrer is observed with a multistage impulse countercurrent (MIG) agitator: In Figure 8 the results obtained by Hoecker and Langer (34) and Hoecker (35) are plotted and are analogous to those in Figure 5. Here a decrease in the power input comparable in magnitude to that with the turbine stirrer cannot be established; this is due to the less marked formation of gas cushions with a MIG agitator. Because the interference multistage impulse countercurrent (INTERMIG) and the oblique arm paddle stirrer both exhibit hydrodynamic behaviour similar to that with a MIG stirrer, this result is to be expected in these cases, too.

Dispersing Capacity of the Stirrer

The dispersing capacity of a stirrer is limited by the so-called flooding point, which gives the maximum gas throughput that can be distributed by the stirrer, $q_{G,max}$. If this maximum gas throughput is exceeded, the stirrer is completely "flooded" by the gas, and the circulatory flow provided by the agitator breaks down. This flooding point exists in the transitional zone leading to turbulent flow and is particularly marked in the turbulent flow region. It manifests itself in a sudden decrease in the power input and mass transfer capacity. In the turbulent area of flow ($Re > 10^4$) the measured data can be correlated using the dimensionless form

$$Q = f(Fr, d/D) \tag{11}$$

Judat (32) reports that the flooding point for turbine stirrers for $Re > 10^4$ and $0.2 \leq d/D \leq 0.42$ is

$$Q_{max} = \frac{0.21Fr^{2.1d/D}}{[(d/D)^{-1} - 2.04]^{1.3}} + \frac{0.14Fr^{7.54d/D}}{[(d/D)^{-1} - 2.25]^{1.5}} \tag{12}$$

The flooding characteristic for different stirrers is depicted in Figure 9 (32). Because these results have all been acquired in laboratory experiments, comparison measurements must be carried out in order to achieve a dependable scale-up. A model to describe the flooding characteristic of stirrers by means of dimensional analysis was suggested by Zehner (37).

For fermentation processes that require low rotational speeds of the stirrer at high sparging rates or at very high gas throughput, Breucker et al. (38) made an interesting observation. For different types of turbine, propeller, or INTERMIG stirring systems they showed that sparge rings near the wall prevent the introduced gas flow from flooding the impeller. In the rotational speed range in which the stirrer is normally flooded, flooding did not occur. This extension of the operation range results in greater safety in vessel design. From the observations available for media of higher viscosity and for non-Newtonian fluids, conservative design considers the dispersing capacity of the agitator as lower than with Newtonian fluids.

Mixing Time Characteristics of Stirred-Tank Reactors

The mixing time is commonly defined as the time necessary to reduce concentration differences in the volume in question to a minimum. The mixing time should thus always be related to the method of measurement used and the degree of homogenity desired. From theoretical considerations of similarity, the mixing time characteristic in the case of Newtonian fluids (with negligible density and viscosity differences) is given by

$$n \cdot \theta = f(Re) \tag{13}$$

with the dimensionless characteristic mixing number $n \cdot \theta$.

In Figure 10 the mixing time characteristics of some types of agitator are illustrated. Those conditions of geometry (which deviate from the German norm DIN 28131) are given in the diagram. In the range of turbulent flow

Symbol				Ga	Fr
Water	Glycerine	CMC	PAA		
●	▲	△	○	$1.3 \times 10^9 - 2.3 \times 10^{10}$	
	▼			$1.37 \times 10^8 - 4.6 \times 10^8$	
	■	□		$10^7 - 6.8 \times 10^7$	> 0.6
	×	+		$1.3 \times 10^6 - 6.4 \times 10^6$	
	◄			$7.1 \times 10^5 - 9.06 \times 10^5$	
– – – – – – – – – – – – – – – – –				$2.2 \times 10^6 - 7.15 \times 10^7$	> 0.9

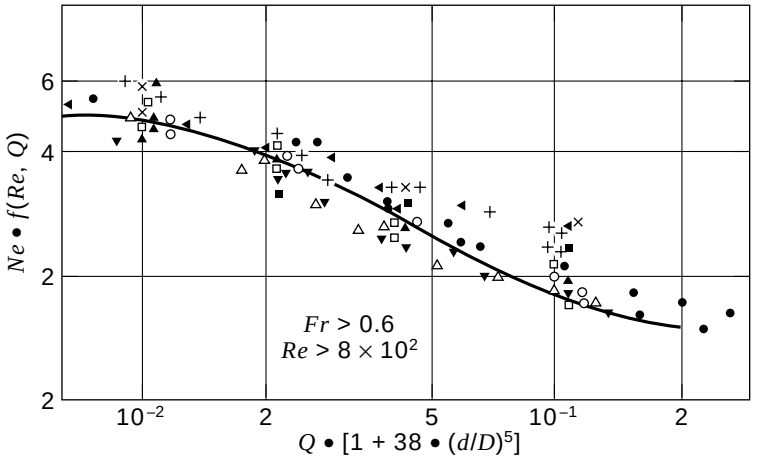


Figure 7. Power input with turbine stirrer (d/D = 0.33): Influence of Reynolds number and gas throughput [$f(Re, Q)$] according to equation 9.

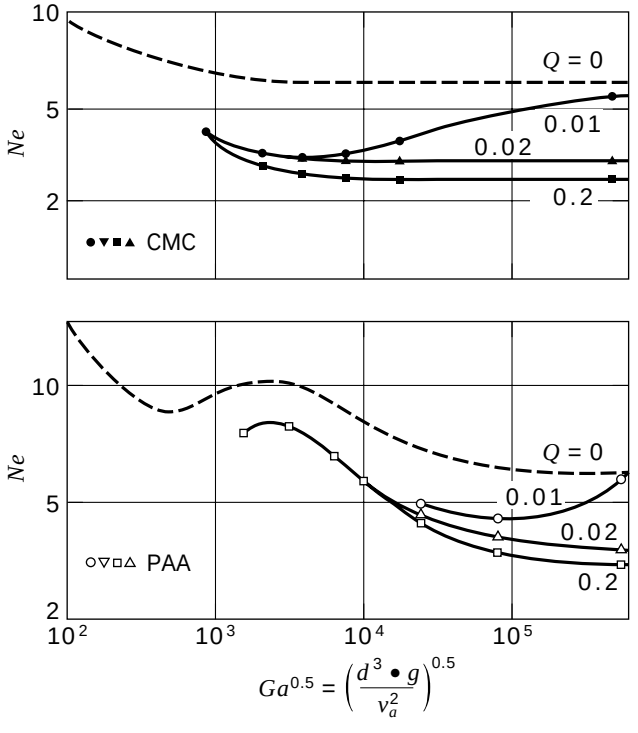


Figure 8. Power input with MIG stirrer (d/D = 0.6, 3 stirrers, H/D = 2), substance systems as in Figure 5.

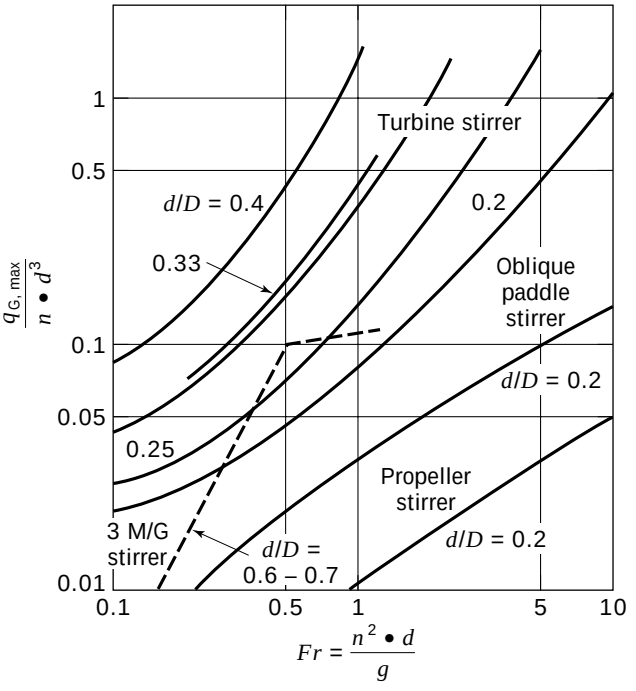


Figure 9. Dispersing capacity of several types of stirrer, $Re > 10^4$.

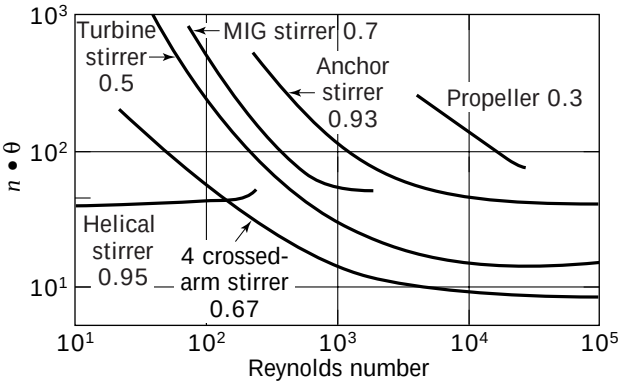


Figure 10. Mixing time characteristics of several types of stirrer in Newtonian fluids.

($Re > 10^4$) the characteristic mixing number is practically independent of the Reynolds number (39–41). Figure 11 shows a plot of the power input that is required to achieve those mixing times in the turbulent region given in Figure 10. The characteristic dimensionless numbers used in this diagram, which were formulated by Zlokarnik (42) on a theoretical consideration of similarity, permit the stirring power input to be determined for a given mixing time, or vice versa (43). For the most favorable conditions of geometry and operation, the stirring power input required in the region of turbulent flow (Ne = constant, $n \cdot \theta$ = constant) is obtained from the relationship

$$\frac{P \cdot \theta^3}{\rho \cdot D^5} = 300 \qquad (14)$$

For a scale-up with P/V = idem there is, from equation 14, a corresponding increase in the mixing time in the region of turbulent flow by

$$\frac{\theta}{D^{2/3}} = \text{idem} \qquad (15)$$

for example,

$$\frac{\theta}{V^{2/9}} = \text{idem} \qquad (16)$$

A scale-up on the basis of constant mixing times (θ = idem) requires, according to equation 14, an increase in the power input by

$$\frac{P/V}{D^2} = \text{idem} \qquad (17)$$

which leads to required energies that cannot be defended either on a technical or an economic basis. Bearing in mind the need for an economical stirring power input, it is indeed meaningful to allow an even longer mixing time when scaling-up than that given by equation 16.

The influence of non-Newtonian flow behavior on the mixing time characteristics can be derived from Figure 12

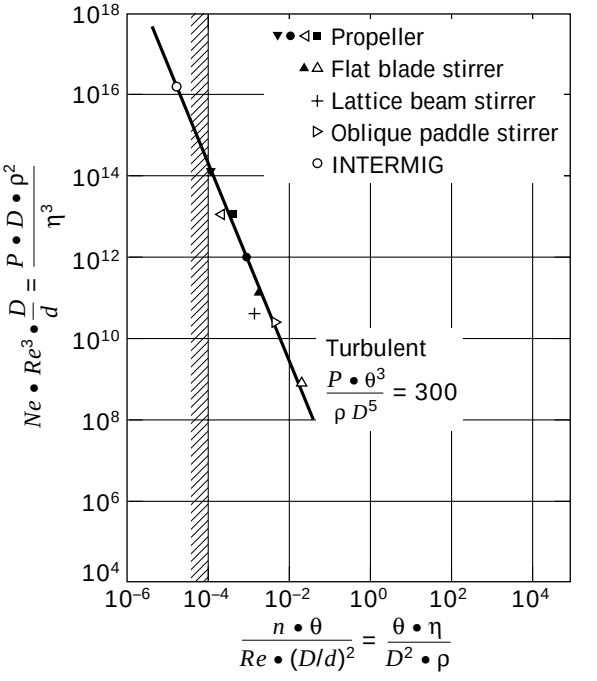


Figure 11. Power input to obtain given mixing times with favorable geometrical and operating conditions.

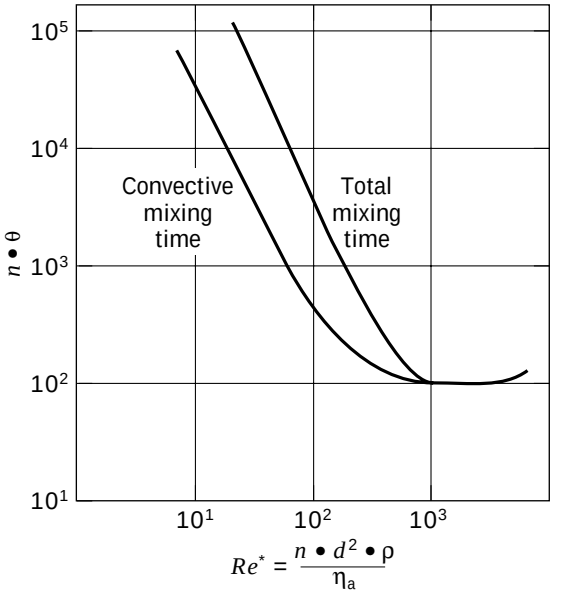


Figure 12. Mixing time characteristics in non-Newtonian fluids (crossed-arm stirrer in CMC solution).

(44). The Reynolds number here has again been modified by the apparent viscosity, as determined according to the method of approach used by Metzner and Otto (24) (equations 4 and 7).

For pseudoplastic fluids Opara (44) observed that in the laminar flow region, about 95% of the liquid is distributed very rapidly by convection, whereas the remainder is

mixed very slowly by diffusion. This convective mixing time corresponds roughly to the mixing time in Newtonian fluids (Fig. 10). The total mixing time in non-Newtonian fluids can be as much as a factor of 10 longer than the convective mixing time.

Mass Transport between Gas and Liquid Phase

Apart from the stirring power, mixing time, and dispersing capacity of the stirrer, mass transfer between gas and liquid is often the prime influence during an aerobic fermentation. The oxygen demand of the aerobic culture has to be the same as the oxygen transfer rate (OTR = $\dot{n}_{O_2}$) in the steady-state condition

$$\dot{n}_{O_2} = \frac{\dot{N}_{O_2}}{V} = k_L a (c^*_{O_2} - c_{O_2})_{av} \tag{18}$$

In this, it has been assumed that because oxygen is such a poorly soluble gas, the overall rate of mass transport is determined essentially by the mass transfer resistance in the liquid phase. In fact the transport coefficient in the gas phase is larger than that in the liquid phase by a factor of 100. The liquid-phase mass transfer coefficient, k_L, multiplied by the interfacial exchange area per unit volume, a, is termed the *volumetric mass transfer coefficient*.

The averaged concentration difference at the phase boundary $(c^*_{O_2} - c_{O_2})_{av}$ in the whole reaction vessel is, in the case of an ideally mixed liquid phase and complete segregation of the gas phase, equal to the logarithmic mean concentration

$$(c^*_{O_2} - c_{O_2})_{av} = \frac{c^*_{O_2,in} - c^*_{O_2,out}}{\ln\left(\dfrac{c^*_{O_2,in} - c_{O_2}}{c^*_{O_2,out} - c_{O_2}}\right)} \tag{19}$$

In stirred tank reactors the average concentration difference can be calculated from equation 19 with sufficient accuracy.

Mass Transfer in Aqueous Fluids. Many equations have been published for the calculation of the volumetric mass transfer coefficient, $k_L a$, in low-viscosity aqueous fluids (see Ref. 5). Regardless of the measuring system used, these equations give differences in the $k_L a$ value of up to a factor of 10. In particular, the influence of the coalescence behavior of the fluid system on $k_L a$ has not been elucidated. Consequently the $k_L a$ values found with the model system of sulfite oxidation in water must be viewed critically when applying such values to fermentation broths.

Theoretical considerations of the similarity of mass transfer between gas and liquid phase made by Zlokarnik (45) led to the formal relationship

$$k_L a = f\left(\frac{P}{V}, \frac{q_G}{V}, \ \ldots\right) \tag{20}$$

with a gas throughput per unit volume q_G/V. However, Judat (46) showed, by evaluating a series of published $k_L a$ values, that in order to fit the measured data better, the gas throughput related to the cross-sectional area of the reactor (i.e., the superficial gas velocity)

$$u_G = \frac{q_G}{A} \tag{21}$$

should be taken as the characteristic gassing quantity. This finding is in agreement with the empirical equation of the type

$$k_L a = K_1 \left(\frac{P}{V}\right)^{\alpha} u_G^{\beta} \tag{22}$$

which was compiled by Moo-Young and Blanch (3) for coalescent and noncoalescent aqueous solutions; their findings are listed in Table 2 (SI units, i.e., $k_L a$ in s^{-1}, P/V in W/m^3, and u_G in m/s). According to Van't Riet (47), differences of up to 35% arise in the $k_L a$ values in coalescent and noncoalescent systems. Van't Riet (47) gives the following correlations for low-viscosity fermentation broths from a series of published data (SI units as in equation 22):

$$\text{Coalescing medium:} \quad k_L a = 0.026\left(\frac{P}{V}\right)^{0.4} u_G^{0.5} \tag{23}$$

$$\text{Noncoalescing medium:} \quad k_L a = 0.002\left(\frac{P}{V}\right)^{0.7} u_G^{0.3} \tag{24}$$

In modifying the dimensionless quantities proposed by Zlokarnik there follows, for the calculation of the volumetric mass transfer coefficient, $k_L a$, the functional relationship

$$k_L a \left(\frac{\nu}{g^2}\right)^{1/3} = \left(\frac{P}{V \cdot \rho \cdot (g^4 \cdot \nu)^{1/3}}, \frac{u_G}{(g \cdot \nu)^{1/3}}, Sc, \sigma^*, S_i\right) \tag{25}$$

In this, $Sc \equiv \nu/\mathcal{D}$ is the Schmidt number, $\sigma^* \equiv \sigma/\rho(g \cdot \nu^4)^{1/3}$ is the dimensionless surface tension, and S_i represents characteristic quantities that are intended to describe the coalescence behavior of the fluid system but whose composition, as already mentioned, still remains largely unclear, though Breucker et al. (38) found that qualitatively the mass transfer coefficient for the noncoalescing system is up to a factor of 4 higher than for coalescing systems at comparable conditions. In the absorption of oxygen in fermentation broths, σ^* and Sc are practically determined only by the viscosity. The surface tension itself changes only very slightly and is not significant. The diffusion coefficient of aqueous solutions exhibiting Newtonian flow behavior can be correlated with the viscosity. On the other hand, the diffusion coefficient in non-Newtonian aqueous solutions is independent of ν and has the same value as water (48,49).

In the mass transfer of gases other than O_2, the Schmidt number and surface tension must also be considered in some cases. However the transport of CO_2, which is also of interest in fermentations, can be calculated without much error by using the same $k_L a$, owing to the very slight differences in diffusion coefficients between O_2 and CO_2. Indeed the overall transport rate is greater for CO_2 because it is 40 times more soluble in water than O_2.

Table 2. Sorption Characteristics of Aqueous Systems for Turbine Stirrers

Liquid	K_1	α	β	Range of validity P/V (kW/m^3)	Range of validity u_G (cm/s)
Water	0.024	0.4	0.5	0.26–5.3	0.3–1.8
Electrolyte solution	0.018	0.74	0.26	0.26–5.3	0.3–1.8
Water	–	0.4	0.35	0.03–18	0.1–0.5
KCl solution					
0.22 N	–	0.71	0.36	0.03–18	0.1–0.5
0.1N	–	0.63	0.62	0.03–18	0.1–0.5
Water	0.0275	0.42	0.43	0.44–10	0.37–1.11
Electrolyte solution (Na_2SO_4 + KOH)	0.017	0.52	0.43	0.44–10	0.37–1.11

Note: $Z = 6$, $d/D = 0.33$; according to Equation 22. For more information see Ref. 5.

By evaluating published data, Judat (46) proposed a correlation that for aqueous systems covers the entire region of technical interest:

$$k_L a\left(\frac{\nu}{g^2}\right)^{1/3} = \frac{9.8 \cdot 10^{-5}\left(\frac{P}{V \cdot \rho(g^4 \cdot \nu)^{1/3}}\right)^{\alpha}}{(B^{-\beta} + 0.81 \cdot 10^{-0.65/B})^{1.05}} \qquad (26)$$

The dimensionless gassing number is here given by

$$B \equiv \frac{q_G/D^2}{(\nu \cdot g)^{1/3}} \qquad (27)$$

The exponents are given as $\alpha = 0.42$ and $\beta = 0.6$. Although the exponents are dependent on the fluid system, the addition of the exponents, $\alpha + \beta \approx 1$, is approximately valid for all substances. Henzler (4) used this circumstance to evaluate the experimental results available with a relationship of the form

$$k_L a\left(\frac{\nu}{g^2}\right)^{1/3} = K_2\left(\frac{P}{V \cdot \rho(g^4 \cdot \nu)^{1/3}}\right)^{\alpha}\left(\frac{u_G}{(g \cdot \nu)^{1/3}}\right)^{\beta} \qquad (28)$$

Taking into account $\beta \approx 1 - \alpha$, the simplified form was derived:

$$\frac{k_L a}{u_G}\left(\frac{\nu^2}{g}\right)^{1/3} = K_2\left(\frac{P}{V \cdot u_G \cdot \rho \cdot g}\right)^{\alpha} \qquad (29)$$

which reflects the measured values with almost the same accuracy as equation 26.

In Figure 13 the sorption characteristics according to equation 29 are given for water and aqueous salt solutions. The results were obtained in reactor volumes of between 2.51 and 900 m^3. They lie close together, independent of the vessel volume. Equation 29 can thus be taken as a translation rule in scale-up on the basis of the oxygen transfer. In Table 3 the constant K_2 and exponent α are given for coalescent and noncoalescent fluid systems for turbine stirrers having $d/D = 0.14 - 0.5$.

Mass Transfer in Fluids of Higher Viscosity. For fluids of higher viscosity, in particular those with non-Newtonian flow behavior, there have been few systematic investigations published that deal with the determination of the volumetric mass transfer coefficient, $k_L a$. Figure 14 shows a plot of several sorption characteristics for millet pulp, aqueous glucose solutions, and aqueous CMC solutions using the same correlation (equation 29) as for low-viscosity fluids (4,35,49). Thereby the translation rule given by equation 29 is confirmed for all fluid systems, even though the viscosity and the diffusion coefficient change by a factor of 1,000 and 30, respectively. To illustrate the experimental results for non-Newtonian fluids it is necessary, however, as with the determination of the power input, to calculate the apparent viscosity with equations 4 and 7 according to the concept of Metzner and Otto (24).

In Table 3 all sorption characteristics compiled by Henzler (4) are shown. Because these sorption characteristics are valid only for the particular rheological behavior and coalescence properties of each system, such properties have to be measured in the fermentation medium before an exact calculation of the mass transport is possible. Perez and Sandall (50) proposed the following correlation to calculate the volumetric mass transfer coefficient for pseudoplastic non-Newtonian fluids according to an approach of Sideman et al. (51):

$$k_L a\left(\frac{d^2}{\mathcal{D}}\right) = 21.24 Re^{111} Sc^{0.5}\left(\frac{\eta_a \cdot u_G}{\sigma}\right)^{0.447}\left(\frac{\eta_G}{\eta_a}\right)^{0.694} \qquad (30)$$

The range of the flow behavior indices according to equation 5 is $m' = 0.916 - 1.00$ and $K' = (0.009 - 0.04)$ Pa $\cdot$ s$^{m'}$ with a vessel diameter of $D = 0.1524$ m. Kawase and Moo-Young (52) showed that the proposed model correlates reasonably well with the relatively wide variety of data measured by different authors for a range of vessel diameters of $D = 0.15 - 0.6$ m and for the flow behavior indices $m' = 0.59 - 0.95$ and $K' = (0.00355 - 10.8)$ Pa $\cdot$ s$^{m'}$.

A completely different correlation for the calculation of the volumetric mass transfer coefficient, $k_L a$, was suggested by Mooyman (53):

$$k_L a = C(270\varphi_G + 135\varphi_G \log(P_a/V)) \qquad (31)$$

Equation 31 expresses the overall transfer coefficient as an enhancement formula dependent on two variables: the gas holdup fraction, φ_G, that is given by

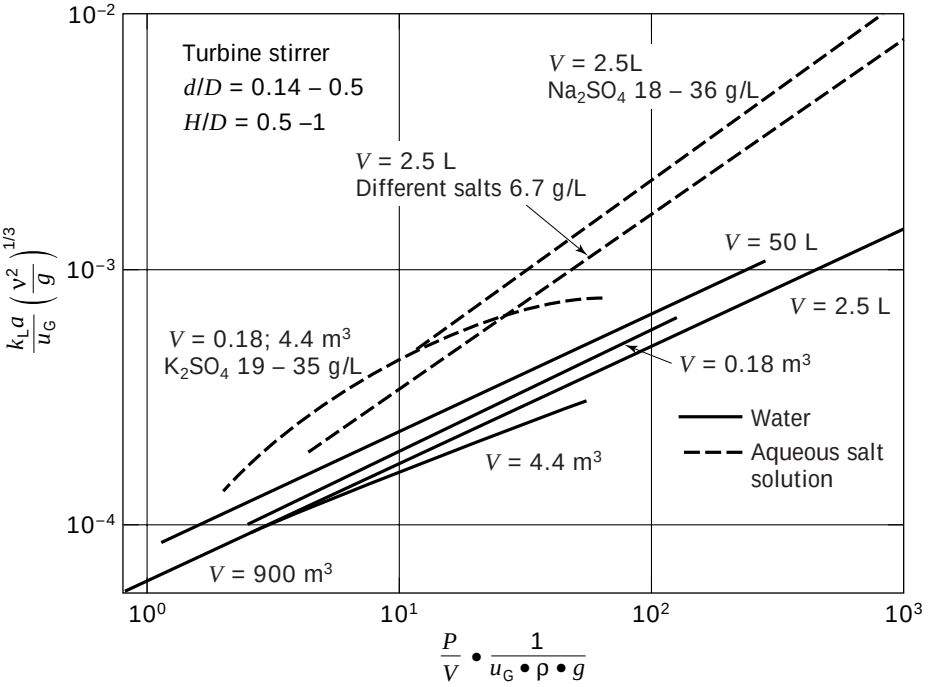


Figure 13. Sorption characteristics of aqueous systems for turbine stirrers.

Table 3. Sorption Characteristics of Different Fluid Systems for Turbine Stirrers

Liquid	Liquid	α	Range of validity		
			P/V (kW/m^3)	u_G (cm/s)	η (mPa s)
Water	7.5×10^{-2}	0.43	0.01–16.7	0.07–1.8	1
Aqueous salt solution					
6.7 g/L	7.2×10^{-5}	0.68	0.15–16	0.12–0.47	1
18–36 g/L	8.3×10^{-5}	0.71	0.23–18	0.12–1.8	1
Aqueous glucose solution	2.65×10^{-5}	0.7	–	–	12–267
Millet pulp	9.0×10^{-5}	0.7	0.23–2.2	0.19–0.78	1.3–70.2
Aqueous CMC solution, $m = 0.4$–0.82, equation 5	3.6×10^{-4}	0.55	0.06–6.2	0.22–1.9	16–1500

Note: $Z = 6$, $d/D = 0.14$–0.5; according to equation 29. For more information see Ref. 5.

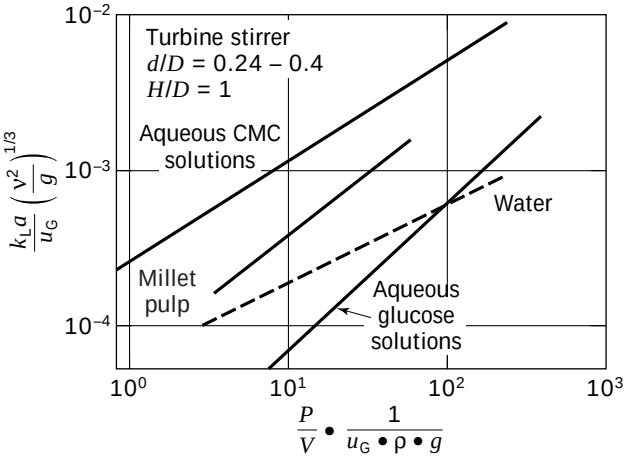


Figure 14. Sorption characteristics for millet pulp, glucose solutions, and CMC solutions.

$$\varphi_G = 0.1\left(\log\left(10.915 u_G \frac{p}{p_{av}}\right) - 0.9\right) \tag{32}$$

and the mechanical plus the adiabatic gas expansion power input per unit of volume, P_a/V; C is a multiplication factor required to correct published or calculated k_La data to actual values observed in industrial fermentations with operating volumes above 250 L (53).

The effects of the suspension of solids on the mass transfer in sparged tank reactors have been investigated by various authors (e.g., Mills et al. [54] and Oguz et al. [55]). Of particular note is the dimensionless correlation developed by Oguz et al. (55) for the k_La in air-sparged agitated tanks containing slurries of different materials:

$$k_L a \frac{V}{q_G} = 2.173 \cdot 10^{-3}\left[\frac{P/q_G}{\rho^{1/3}(\eta_a \cdot g)^{2/3}}\right]^{0.607} \tag{33}$$

The influence of the solid concentration becomes evident

in the calculation of the apparent viscosity of the slurries, which all showed non-Newtonian behavior. Consequently the viscosity was calculated with equations 4 and 7 according to the concept of Metzner and Otto (24). Oguz et al. (55) showed that an increase in the solid concentration leads to an increase of the slurry viscosity that corresponds to a decrease in k_La. The deviation of the measured data from the values calculated with equation 33 was in the range of 12.5%.

For a scale-up on the basis of constant volumetric mass transfer coefficients, there follows from the generally valid equation 29

$$\left(\frac{P}{V}\right)u_G^{(1-\alpha)/\alpha} = \text{idem} \tag{34}$$

for the power input required. The exponent α is approximately 0.67 for noncoalescing aqueous systems as well as for more highly viscous Newtonian fermentation fluids. From equation 34 it follows that

$$\left(\frac{P}{V}\right)u_G^{0.5} = \text{idem} \tag{35}$$

For example, with $u_G \sim q_G/D^2$

$$\left(\frac{P}{V}\right)\frac{q_G^{0.5}}{D} = \text{idem} \tag{36}$$

If surface-active substances, such as chemical antifoam agents, are added to the fermentation medium, the volumetric mass transfer coefficient at the interfacial surface usually drops, and the mass transfer behavior tends toward a coalescing fluid system. In a coalescing fluid system and for non-Newtonian fluids, the exponent α is approximately 0.5. Consequently it follows from equation 29 that

$$\left(\frac{P}{V}\right)u_G = \text{idem} \tag{37}$$

and

$$\left(\frac{P}{V}\right)\frac{q_G}{D^2} = \text{idem} \tag{38}$$

In applying this translation rule (equations 35–38), the dispersing capacity of the stirrer, as given by equation 12, also has to be taken into consideration.

Shearing Stress and Pumping Capacity of the Stirrer

The mixing of the fermentation broth, the distribution and dispersion of the gases, and the suspending of solid constituents of the culture medium and microorganism pellets are effected in the stirred tank by a complex three-dimensional flow field. In only a few isolated cases (those that have a defined flow field) can the shearing stress and the pump capacity of the stirrer be calculated. With turbulent mixing it is possible, however, only to estimate both the longitudinal mixing by the pumping around of the vessel contents, and the shear stress on the fluid caused by the fluctuations in turbulence.

In a theoretical consideration of turbulent flow, Reynolds split up the measured local liquid velocity, u_L, into the sum of a mean (deterministic) velocity, $\bar{u}$, and a superposed stochastic fluctuation velocity, u'. According to this hypothesis, the turbulent Reynolds shear stress produced by the stirrer is proportional to the square of the local fluctuation velocity of the turbulence and is approximately given by the tip speed of the stirrer

$$\tau_R \sim (n \cdot d)^2 \tag{39}$$

With the Reynolds shear stress, τ_R, the effects of the stochastic liquid fluctuation velocity, u', are mathematically seized. τ_R can be interpreted as a virtual shear stress, that is, a consequence of the impulse transport caused by the turbulent motions.

Hoffmann et al. (56) determined the shear stress in stirred-tank reactors by means of a model system consisting of flocculated clay. Their investigations showed that the proportionality between shear stress of the clay floc system and the tip speed of the stirrer is of little influence; this is caused by the distinctness of the length scale of the vortices resulting from the Reynolds shear stress and the size of the relatively small flocs. The clay floc system is suitable for describing the influence of shear stress on tissue cell cultures (57).

Therefore, the model of Cherry and Kwon (58) is proposed for the calculation of shear stress. It is based on Kolmogoroff's theory of local homogenous isotropic turbulence: the length scale of the microeddies in the dissipative area (with laminar flow of the eddies) yields

$$\lambda = \left(\frac{\nu^3}{\varepsilon}\right)^{0.25} \tag{40}$$

where ε is the energy dissipation rate (i.e., the energy input to the system that is completely dissipated by the formation of vortices)

$$\varepsilon = \frac{P}{\rho \cdot V} \tag{41}$$

To employ this model by Cherry and Kwon (58), the particles exposed to shear have to be smaller than 12 times the length scale, λ, a condition that is usually fulfilled in fermentation processes. Under these circumstances the microeddies induced by the energy input or the energy dissipation rate, respectively, cause the shear stress. The particles are put into rotation by the velocity gradients in the microeddies, which are significantly larger than the gradients caused by the liquid fluctuation velocities in Reynolds hypothesis. The pulsating shear stress that acts upon the flocs during the contact with the microeddies leads to

$$\tau = 5.33\rho(\varepsilon \cdot \nu)^{0.5} \tag{42}$$

that is,

$$\tau = 5.33\left(\frac{P}{V}\nu \cdot \rho\right)^{0.5} = K_3\left(\frac{P}{V}\right)^{0.5} \quad (43)$$

The influence of the liquid fluctuation velocities (i.e., the influence of the shear stress caused by the microeddies) has its maximum value if the dimension of the particles is close to the length scale of the microeddies. In Figure 15 the relationship between the experimentally determined shear stress of the clay floc system and the volumetric energy input for different types of stirrers is shown, using a relationship in analogy to equation 43

$$\tau = K_3\left(\frac{P}{V}\right)^{\alpha} \quad (44)$$

The experimentally determined maximum shear stress deviates from the data calculated with the model of Cherry and Kwon (58) by two to six times. This can be explained by the fact that, according to the type of impeller system, there exist areas with increased energy dissipation rates in the reactor in the area close to the stirrer. This effect was not taken into consideration in the model, which uses a mean energy dissipation rate over the entire reactor volume that can be up to 30 times smaller than the maximum value of the energy dissipation rate close to the stirrer (59). Because the main part of the energy input is dissipated near the impeller, the model has to be extended by a supplementary corrective term containing information about the type of stirrer and its geometry (i.e., about the area with increased dissipation rates). By analogy with the work of Liepe (60), who studied the ratio between the maximum and the mean dissipation rate, $\varepsilon_{max}/\varepsilon$, of different impeller systems for dispersion tasks, Hoffmann et al. (56) suggested a corrective function

$$f = f\left(\frac{V}{V_S \cdot Ne}\right)^{\beta} \quad (45)$$

with the effective energy dissipation volume, V_S, that is proportional to the volume of the stirrer ($V_S \sim d^2 \cdot h_B$). The values of V/V_S and Ne are given in Figure 15. By fitting the experimental data, Hoffmann et al. (56) finally determined the following correlation for the calculation of shear stress in stirred-tank reactors:

$$\tau = 0.367\left(\frac{P}{V}\left(\frac{V}{V_S \cdot Ne}\right)^{0.42}\right)^{0.55} \quad (46)$$

Figure 16 shows the calculated shear stress according to equation 46 compared with the experimentally determined data (56). The deviation of the model from the measured data is in a range of up to 30%, which means the model is suitable for a scale-up of the reactor performance on the basis of constant shear stress. In a sparged stirred-tank reactor the shear stress is also dependent on the gas input. Experimental investigations of Henzler and Biedermann (61) showed that there is a qualitative dependency of the shear stress on the pneumatic power input and the type of sparging system. They observed that with constant volumetric power input the shear stress increases with the increasing portion of pneumatic power input and with increasing superficial gas velocity. They described some of

Stirrer		d/D	V/V_S	Ne
•	Propeller	0.33	200.7	0.31
▪	INTERMIG	0.65	40.7	0.65
▲	Turbine	0.33	310.1	5.0
△	Turbine	0.65	42.5	5.5
□	4-blade	0.33	135.7	9.6
——	Model according to Ref. 58			

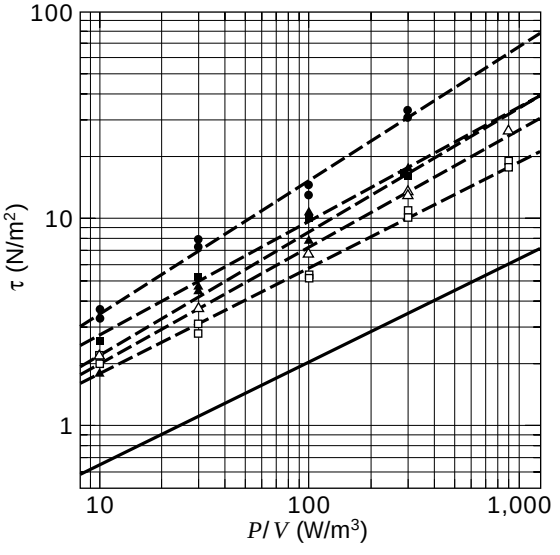


Figure 15. Experimentally determined maximum shear stress versus the volumetric power input, according to equation 44.

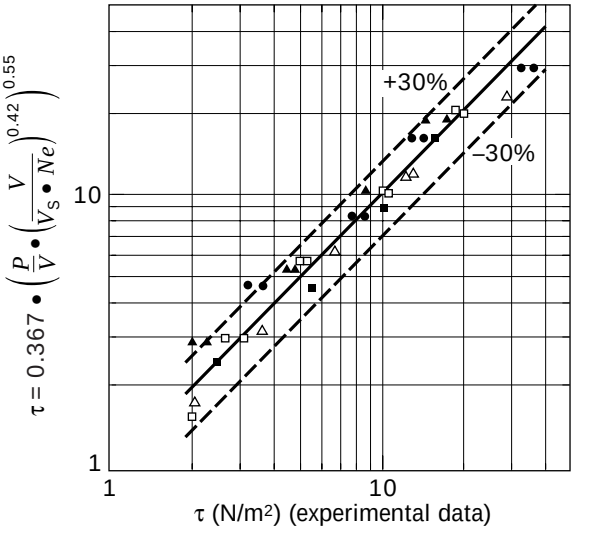


Figure 16. Experimentally determined shear stress compared with calculated data, according to equation 46; for symbols, see Figure 15.

the observed effects causing the increased shear stress, such as the turbulence in the lag zone of rising bubbles, coalescence or dispersion of bubbles close to the sparger, and the bursting of bubbles at the liquid surface (as is observed in bubble columns) (62). Because a model or correlations to describe these effects quantitatively are not yet known, further investigations are necessary to describe the effects of sparging and bubbles on the mechanical shear stress in stirred vessels.

A relationship for the circulating volumetric flow, which is produced by the pumping capacity of the stirrer, can be derived in a manner analogous to the principle that applies to rotary pumps. In this way Van de Vusse (39) calculated the circulating volumetric flow of propellers with the assumption that counterpressure is "zero" and obtained

$$\dot{V} = 0.5 \frac{\pi}{4} n \cdot d^2 \cdot h_S \tag{47}$$

where h_S is the pressure head produced by the propeller. With $h_S \sim d$, the generalized relationship

$$\dot{V} = K_4 \cdot n \cdot d^3 \tag{48}$$

follows from equation 47. K_4 is a constant that is essentially dependent on the type of stirrer (63) (see also Refs. 16, 19, and 64).

Varying details have been published regarding the suspension of solids in stirring apparatuses. Most authors say that in achieving the same suspension result in a larger vessel, the power input can be diminished (43). The circulating volumetric flow given by equation 48 is sufficiently large under the conditions of stirring that generally prevail in stirred-tank reactors so that a uniform distribution of cell pellets and solid constituents in the culture medium is guaranteed. This uniform distribution is supported by the very slight density differences between solid and liquid in the fermentation medium, and the relatively small solid particles and the small fraction of solid material ($d_P < 2$ mm, $\varphi_S < 20\%$) in the fermentation medium. Consequently the suspension process in general takes on no particular significance in scaling up stirred-tank reactors.

Mass Transport Processes in Mycelium Agglomerates

Mycelium-forming microorganisms can build up many morphological forms while submerged fermentation, depending on the conditions of fermentation. Forms include loose hyphae that are distributed homogenously throughout the medium or collect in agglomerates that appear as compact pellets. Whereas the filamentous forms result in highly viscous, non-Newtonian fermentation fluids, the mycelial pellets are usually present in Newtonian fluid systems at a considerably lower viscosity (65). An adequate shearing of the mycelium-forming microorganisms thereby leads to the preferred formation of pellets rather than a filamentous mycelial structure.

Regardless of whether the formation of pellets and compact cell agglomerates leads to fluids of lower viscosity (and thus to more turbulent conditions), problems regarding the oxygen supply to the cells can arise, in particular with large, compact pellets. When the inner and outer mass transport resistances are too large, the danger of autolysis exists, and only the external pellet layer is involved in the biotechnological process.

The effective diffusion coefficient of oxygen in mycelial pellets of *Aspergillus niger* was found to be about $\mathcal{D}_{Pe} = 10^{-4}$ cm^2/s (2). This value is larger by a factor of 10 than the molecular diffusion coefficient, $\mathcal{D}$, in the culture medium. Because the diffusion coefficient determined by Miura (2) also varied with the volumetric power input, this result must be interpreted as indicating that the mycelium forms a loose pellet structure that is penetrated by the turbulent microflow. Thus in such cultures an adequate but not excessive turbulence (i.e., shear stress) is required in order for pellets with a loose structure to form.

Whereas the mass transport inside mycelial aggregates has still not been widely researched, reliable equations exist for determining the mass transfer from fluid to pellet. The mass transfer coefficient k_{Pe} on the surface of the pellet of diameter d_{Pe} can be calculated from the formal relationship

$$Sh = f(Re_{Pe}, Sc) \tag{49}$$

where $Sh \equiv k_{Pe} \cdot d_{Pe}/\mathcal{D}$ is the Sherwood number, $Re_{Pe} \equiv u_{Pe} \cdot d_{Pe}/\nu$ is the Reynolds number formed with the relative velocity between pellet and fluid, and $Sc \equiv \nu/\mathcal{D}$ is the Schmidt number. The relative velocity at the pellet is determined solely by local turbulent fluctuation velocities, u', which are effected by the turbulent eddies of the scale of the pellets. According to the theory of isotropic turbulence put forward by Kolmogoroff, it follows for this velocity that

$$\sqrt{\overline{u'^2}} \sim (\varepsilon \cdot d_{Pe})^{1/3} \tag{50}$$

where ε is the mass-related energy of dissipation calculated with equation 41. Then the Reynolds number, formed with the relative velocity

$$u_{Pe} \sim \sqrt{\overline{u'^2}}$$

is

$$Re_{Pe} = \left(\frac{P}{V} \frac{d_{Pe}^4}{\rho \cdot \nu^3}\right)^{1/3} \tag{51}$$

In Figure 17 the experimental results according to equation 49, from different authors, are shown; the experimental data, following the theory, have been plotted as

$$\frac{Sh - 2}{Sc^{1/3}} = f(Re_{Pe}) \tag{52}$$

The only published equations that are mentioned here are those relationships reported by Liepe and Moeckel (66); with the relative velocity defined here they are given by

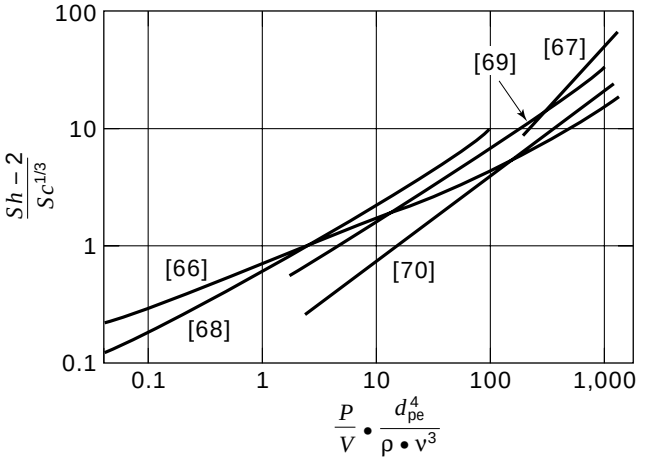


Figure 17. Mass transfer at the pellet surface as dependent on the particle Reynolds number. Numbers in brackets indicate reference.

$$\frac{Sh - 2}{Sc^{1/3}} = 0.35\left(\frac{\Delta\rho}{\rho}\right)^{1/3}\left(\frac{P}{V}\frac{d_{Pe}^4}{\rho \cdot \nu^3}\right) \qquad (53)$$

The curve given by equations 52 and 53 is illustrated in Figure 17, marked according to Liepe and Moeckel (66). Apart from the measurements made by Sicardi et al. (67), all measurements were performed in a nongassed system. Nevertheless the comparison of results shows that the presence of gas has virtually no influence on the mass transfer from fluid to pellet.

SCALE-UP EXAMPLES OF BIOTECHNOLOGICAL PROCESSES BY MAINTENANCE OF IDENTICAL OPERATING CONDITIONS

Scale-Up on the Basis of the Oxygen Transfer Rate

A scale-up of aerobic biotechnological processes is often performed successfully using the criterion of identical oxygen transfer rate (OTR = idem), or the same volumetric mass transfer coefficient (k_La = idem). The reason is that in the aerobic culture of yeasts and bacteria, which have high rates of respiration, the rate-limiting reaction step is often located in the oxygen transfer resistance at the gas–liquid interface. In the fermentation of mycelium-forming microorganisms there arises, under certain conditions, a further limiting mass transport resistance at the surface of or inside the mycelial pellet. That is the reason why a scale-up on the basis of both mass transport resistances is more successful.

Karow et al. (71) took the oxygen transfer rate as a scale-up criterion for the production of penicillin and streptomycin. In Figure 18 the relative penicillin concentrations are shown as a function of the oxygen transfer rate. The results for the production of streptomycin are very similar. The scale-up covers 4 orders of magnitude. At oxygen transfer rates of more than 0.5 mol/(L h) the product yield is maximal and nearly constant. The successful translation from the laboratory to a production scale, performed by

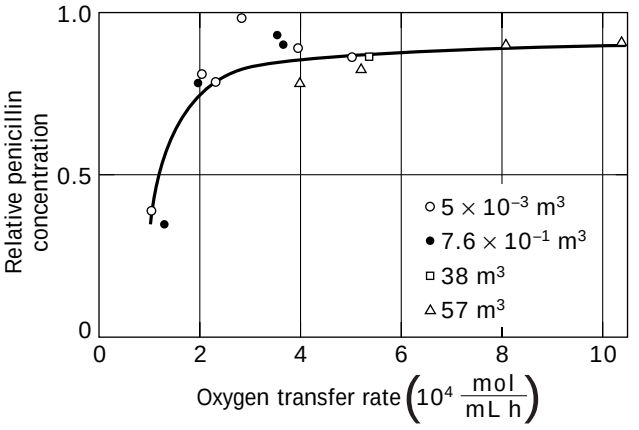


Figure 18. Production of penicillin as a function of the oxygen transfer rate.

Karow et al. (71), was also confirmed by Bylinkina et al. (72) for the fermentation of penicillin and streptomycin. They carried out a scale-up from the 15-L fermentor via a 100-L and 3-m^3 fermentor up to a 63-m^3 production fermentor on the basis of the mass transfer coefficient at the gas–liquid interface. Wegrich and Shurter (73) also described a successful scale-up of penicillin production. The same yields of penicillin were attained in an 8-m^3 fermentor as in a 100-m^3 fermentor. They did not only keep the superficial gas velocity constant but also the volumetric power input of the stirrer. According to equation 37, holding these parameters constant in non-Newtonian fluids corresponds to the scale-up rule k_La = idem.

A further, now "historical," example of scale-up on the basis of the oxygen transfer is shown in Figure 19, in which the production of baker's yeast is depicted (74). Although the oxygen transfer rate in the model system (sulfite oxidation), used as the scale-up criterion, was measured, no significant difference in the production of yeast could be

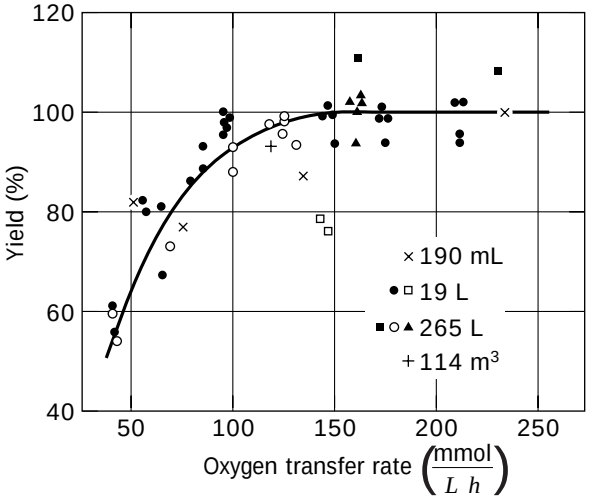


Figure 19. Scale-up of fermentation of baker's yeast based on oxygen transfer rate (sulfite oxidation).

established, which was dependent only on the size of the reactor. Those points plotted in Figure 19 were obtained using shaker flasks with a 190-mL liquid content and in 19-L, 265-L, and 114-m^3 fermentors. Jari (75) was able to show for the production of the primary metabolites alkaline and acid proteases and the secondary metabolites nystatin and fumagillin that a scale-up can be carried out using the same oxygen transfer rate (k_La = idem).

In the bacterial production of vitamin B_{12}, Bartholomew (76) established that the yield at high oxygen transfer rates is negatively influenced by a not very clearly described effect (Fig. 20). Since, moreover, this effect is dependent on the fermentor size, the oxygen transfer rate in this region cannot be used as a scale-up criterion. The dotted curve in Figure 20 applies to the yield only when small-scale equipment is used, and the continuous line represents the yield achieved in a production fermentor. In other words, the production fermentor is oversized. Takei et al. (77) studied protease production by *Streptomyces* sp. in a 0.03-m^3- and a 0.2-m^3-sized fermentor and found the same effect. The productivity of the cells in synthesizing protease is indeed determined by the oxygen transfer rate but is nonetheless diminished at higher oxygen transfer rates by a fermentor size–dependent effect.

In the production of glucoamylase by *Endomyces* sp., the enzyme yield is also not solely determined by the oxygen transport. A scale-up on the basis of identical k_La values also leads to oversizing of the production fermentor. The yields depicted in Figure 21, expressed in relative enzyme activities as a function of the volumetric mass transfer coefficients, were measured in fermentors of 0.06 m^3, 3 m^3, and 30 m^3 (78). In the latter three examples of fermentation described, no adequate scale-up among the fermentor sizes was achieved by using solely the oxygen transfer rate (i.e., the k_La value). The reasons for such deviations can be diverse. For example, in applying the same k_La values as a scale-up criterion, the influence of the surface aeration has to be taken into consideration. Fuchs et al. (79) showed that this effect makes itself felt at fermentor volumes smaller than 200 L.

The translation of the oxygen transfer rate from a laboratory fermentor to a production scale can still lead to

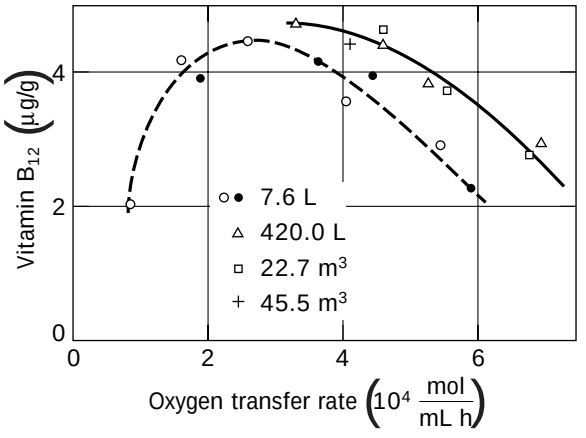


Figure 20. Production of vitamin B_{12} as function of the oxygen transfer rate.

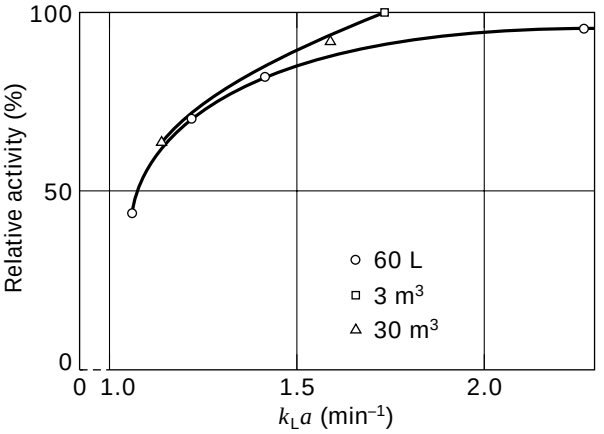


Figure 21. Relative glucoamylase activity as function of the k_La value.

considerable errors for other reasons. Thus, with the same k_La value, the oxygen transfer rate in the production fermentor can be larger than in the laboratory fermentor because, owing to the higher hydrostatic pressure, the driving force (concentration difference) will be larger. But with the same bubble size (identical rising velocity of the bubbles) the average residence time of the air bubble in the larger apparatus is longer; it is thus deficient in oxygen, and consequently, the positive effect of hydrostatic pressure, just cited, is diminished.

Scale-Up on the Basis of the Volumetric Power Input

One-third of the scale-up translations performed in the fermentation industry use the scale-up criterion "identical volumetric power input," as Margarites and Zajic (15) reported. Apart from the simplicity of the measurement in applying this scale-up rule, the reason for its industrial success probably lies in the fact that almost all hydrodynamic and mass transport phenomena can be correlated using the dissipated energy (and thus using P/V). Nevertheless, or maybe simply because of it, this criterion provides an extremely rough rule of thumb, which permits a scale-up only on an empirical basis.

Maxon (80) studied the influence of the volumetric power input on the production of novobiocin using *Streptomyces niveus*. In Figure 22 the novobiocin concentration produced is plotted against the volume-related power input for three different sizes of stirrer. Above P/V = 1 kW/m^3 the yields of the antibiotic are maximal and largely independent of P/V. Scale-up based on this criterion is not dependable because it has been established that there is a dependance on the diameter of the stirrer. According to Wang and Fewkes (81) a scale-up of this process is possible only owing to the partial similarity of the turbulent shearing stress and the pumping capacity of the stirrer.

In the fermentation of penicillin, the volumetric power input, P/V, was often used in the past as a criterion for scale-up. In Figure 23 the penicillin concentration is plotted as a function of the volume-related stirring power, as described by Gaden (82), who gave P/V values between 1.5

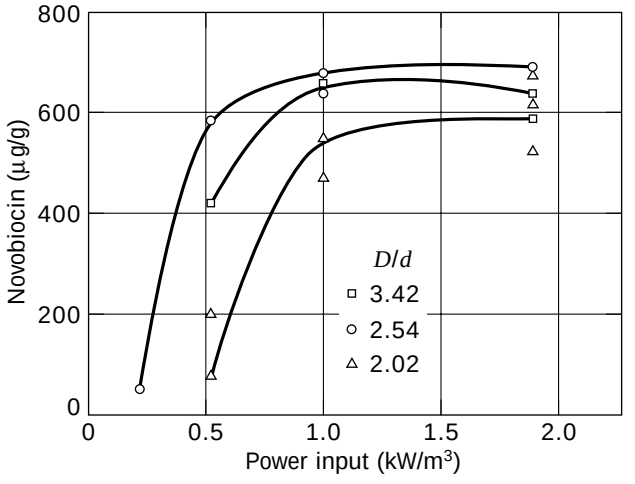


Figure 22. Novobiocin concentration after 115 h of fermentation duration, dependent upon the volumetric power input.

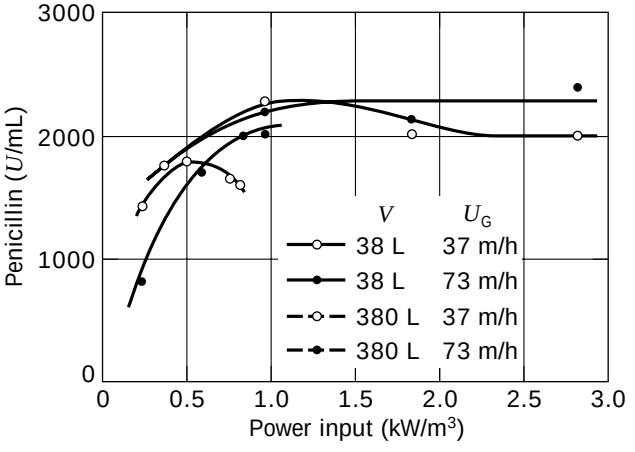


Figure 23. Influence of the volumetric power input on the synthesis of penicillin.

and 3 kW/m^3 as a scale-up rule. Similar results were achieved by Humphrey (83,84), who established a correlation between penicillin production and the volumetric power input and carried out a scale-up at a volume-related stirring power between 1 and 2.5 kW/m^3. Humphrey (84) reports a volumetric power input of 1.5 kW/m^3 for a scale-up of streptomycin production.

NOMENCLATURE

A	Reactor cross-sectional area (m^2)
a	Interfacial exchange area per unit volume (m^2/m^3)
$B \equiv (q_G/D^2)/(v \cdot g)^{1/3}$	Dimensionless gassing number
C	Factor
$c^*_{O_2}$	Saturation concentration of O_2 in fermentation medium (mg/L)
c_{O_2}	Concentration of O_2 in fermentation medium (mg/L)
$c^*_{O_2,in}$, $c^*_{O_2,out}$	Saturation concentration of O_2 at entrance and exit of reactor (mg/L)
D	Diameter of stirred-tank reactor (m)
$\mathscr{D}$	Diffusion coefficient of the liquid phase (m^2/s)
$\mathscr{D}_{Pe}$	Effective diffusion coefficient in mycelial pellets (m^2/s)
d	Diameter of the stirrer (m)
d_P	Diameter of solid particles (m)
d_{Pe}	Diameter of mycelial pellets (m)
$Fr \equiv n^2 \cdot d/g$	Froude number
$Ga \equiv d^3 \cdot g/v$	Galilei number
g	Acceleration due to gravity (m/s^2)
H	Liquid filling height in the reactor (m)
h	Distance between stirrers (m)
h_B	Height of stirrer blade (m)
h_S	Pressure head of the stirrer (m)
i	Number of stirrer blades
K	Proportionality constant
K'	Consistency index [(Pa s)m']
K''	Consistency index [(Pa s)m'']
K_1, K_2, K_3, K_4	Constants
k_L	Liquid-side mass transfer coefficient at gas–liquid interface (m/s)
k_{Pe}	Mass transfer coefficient at mycelium pellet surface (m/s)
k_La	Volumetric mass transfer coefficient at gas–liquid interface (s^{-1})
m', m''	Flow behavior parameters
$Ne \equiv P/\rho \cdot n^3 \cdot d^5$	Newton number
Ne_G	Newton number in gassed reactor at $Re > 10^4$
Ne_0	Newton number without gassing
$\dot{N}_{O_2}$	Oxygen mass flow (kg/s)
n	Rotational speed (s^{-1})
$\dot{n}_{O_2}$	Oxygen transfer rate (mass flow density) [kg/(s · m^3)]
P	Stirring power (kW)
P_a	Mechanical plus adiabatic gas expansion power input (kW)
p	Atmospheric pressure (N/m^2)
p_{av}	Average pressure at half the liquid height (N/m^2)
$Q \equiv q_G/n \cdot d^3$	Dimensionless gassing number
Q_{max}	Dimensionless gas flow number at flooding point of stirrer
q_G	Gas throughput (m^3/s)
$q_{G,max}$	Maximum gas throughput (m^3/s)

$Re \equiv n \cdot d^2/\nu$	Reynolds number
$Re^* \equiv n \cdot d^2 \cdot \rho/\eta_a$	Modified Reynolds number
$Re_{Pe} \equiv u_{Pe} \cdot d_{Pe}/\nu$	Pellet Reynolds number
S_i	Characteristic quantities for coalescence of fluids
$Sc \equiv \nu/\mathcal{D}$	Schmidt number
$Sh \equiv k_{Pe} \cdot d_{Pe}/\mathcal{D}$	Pellet Sherwood number
u'	Turbulent fluctuation velocity of the fluid (m/s)
$\bar{u}$	Mean velocity of the fluid (m/s)
u_G	Gas velocity based on empty cross section area of the reactor (m/s)
u_L	Local liquid velocity (m/s)
u_{Pe}	Relative velocity between pellet and fluid (m/s)
V	Reactor volume (m^3)
$\dot{V}$	Circulating volumetric flow by stirrer pumping (m^3/s)
V_S	Volume stirred up by the stirrer blades (m^3)
Z	Number of stirrer blades
α	Exponent
β	Exponent
$\dot{\gamma}$	Shear rate (s^{-1})
ε	Mass-related energy dissipation rate (kW/kg)
ε_{max}	Maximum energy dissipation rate (kW/kg)
η	Dynamic viscosity (Pa s)
η_a	Apparent dynamic viscosity (Pa s)
η_G	Dynamic viscosity of the gas phase (Pa s)
θ	Mixing time (s)
λ	Length scale of microeddies (m)
ν	Kinematic viscosity (m^2/s)
ρ	Density of fermentation fluid (kg/m^3)
σ	Surface tension (N/m)
$\sigma^* \equiv \sigma/[\rho(g \cdot \nu^4)^{1/3}]$	Dimensionless surface tension
τ	Shear stress (Pa)
τ_R	Reynolds shear stress (Pa)
τ_0	Yield stress (Pa)
φ_G	Gas holdup
φ_S	Solid holdup

BIBLIOGRAPHY

1. D.I.C. Wang and A.E. Humphrey, in D.J.D. Huckenhull ed., *Progress in Industrial Microbiology*, vol. 8, Temple Press Books, London, 1968, pp. 1–34.
2. Y. Miura, in T.K. Ghose, A. Fiechter, and N. Blakebrough eds., *Advances in Biochemical Engineering*, vol. 4, Springer-Verlag, Berlin, 1976, pp. 3–40.
3. M. Moo-Young and H.W. Blanch, in A. Fiechter ed., *Advances in Biochemical Engineering*, vol. 19, Springer-Verlag, Berlin, 1981, pp. 1–69.
4. H.-J. Henzler, *Chem.-Ing.-Tech.* **54**, 461–476 (1982).
5. D.C. Hempel, in R.K. Finn, P. Praeve, M. Schlingmann, W. Crueger, K. Esser, R. Thauer, and F. Wagner eds., *Biotechnology Focus 1*, Hanser Publishers, Munich, 1988, pp. 51–94.
6. J. Bryant, in T.K. Ghose, A. Fiechter, and N. Blakebrough eds., *Advances in Biochemical Engineering*, vol. 5, Springer-Verlag, Berlin, 1977, pp. 101–124.
7. R. Mann, P.P. Mavors, and J.C. Middleton, *Trans. Inst. Chem. Eng.* **59**, 271–278 (1981).
8. J. Pawlowski, *Die Ähnlichkeitstheorie in der physikalisch-technischen Forschung: Grundlagen und Anwendung*, Springer-Verlag, Berlin, 1971.
9. J. Stichlmaier, *Chem.-Ing.-Tech.* **63**, 38–51 (1991).
10. F. Kudrewizki, *Acta Biotechnol.* **4**, 89–104, 247–254 (1984).
11. M. Reuss, R.K. Bajpal, D. Debus, R. Lenz, H. Niebelschuetz, and A. Papalexiou, *6th Int. Ferment. Symp.*, London, Ontario, Canada, 1980.
12. R.K. Bajpal and M. Reuss, *Can. J. Chem. Eng.* **60**, 384–392 (1982).
13. N.M.G. Oosterhuis, N.M. Groesbeek, A.P.C. Olivier, and N.W.F. Kossen, *Biotechnol. Lett.* **5**, 141–146 (1983).
14. N.M.G. Oosterhuis, *Scale-Up of Bioreactors*, Proefschrift, Techn. Hogeschool, Delft, Netherlands, 1984.
15. A. Margarites and J.E. Zajic, *Biotechnol. Bioeng.* **20**, 939–1001 (1978).
16. J.Y. Oldshue, *Fluid Mixing Technology*, McGraw-Hill, New York, 1983.
17. J.Y. Oldshue, *Chem. Eng. Prog.* **5**, 33–42 (1989).
18. J.Y. Oldshue, *AlChE Symp. Ser.* **89**, 158–163 (1993).
19. J.Y. Oldshue, *Chem. Eng. Prog.* **93**, 70–73 (1997).
20. J.Y. Oldshue, *Biotechnol. Bioeng.* **8**, 3–24 (1966).
21. J.H. Rushton, E.W. Costich, and H.J. Everett, *Chem. Eng. Prog.* **46**, 395–404, 467–476 (1950).
22. M. Zlokarnik, in E. Bartholome, E. Biekert, H. Hellmann, H. Ley eds., *Ullmanns Enzyklopaedie der Technischen Chemie*, vol. 2, Verlag Chemie GmbH, Weinheim/Bergstrasse, 1972, pp. 259–281.
23. K. Kipke, *GVC-Vortragstagung: Verfahrenstechnische Fortschritte beim Mischen, Dispergieren und bei der Wärmeuebertragung in Fluessigkeiten*, VDI-Verlag, Duesseldorf, 1978, 21–36.
24. A.B. Metzner and R.E. Otto, *AlChE J.* **3**, 3–11 (1957).
25. A.B. Metzner, R.H. Feehs, H.L. Ramos, R.E. Otto, and J.D. Toothill, *AlCHE J.* **7**, 3–9 (1961).
26. R. Rapp, H. Reng, D.C. Hempel, and F. Wagner, *Biotechnol. Bioeng.* **26**, 1167–1175 (1984).
27. P.H. Calderbank and M.B. Moo-Young, *Trans. Inst. Chem. Eng.* **37**, 26–33 (1959).
28. A. Sanchez, A. Martinez, L.G. Torres, and E. Galindo, *Process Biochem.* **27**, 351–365 (1992).
29. D. Velasco, A. Martinez, L.G. Torres, and E. Galindo, in A.W. Nienow ed., *3rd International Conference on Bioreactor and Bioprocess Fluid Dynamics, BHR Group Conference Series Publication No. 5*, Mechanical Engineering Publications Limited, London, 1993, pp. 101–116.
30. Y. Oyama and K. Endoh, *J. Chem. Eng. Jpn.* **19**, 2–11 (1955).

31. M. Zlokarnik, *Chem.-Ing.-Tech.* **45**, 689–692 (1973).
32. H. Judat, Dissertation, Universitaet Dortmund, Germany, 1976.
33. V. Schlueter and W.D. Deckwer, *Chem.-Ing.-Tech.* **64**, 644–645 (1992).
34. H. Hoecker and G. Langer, *Rheol. Acta* **16**, 400–412 (1977).
35. H. Hoecker, Dissertation, Universitaet Dortmund, Germany, 1979.
36. H.-J. Henzler, *GVC-Vortragstagung: Verfahrenstechnische Fortschritte beim Mischen, Dispergieren und bei der Wärmeuebertragung in Fluessigkeiten*, VDI-Verlag, Duesseldorf, 1978, pp. 91–110.
37. P. Zehner, *Chem.-Ing.-Tech.* **60**, 531–539 (1988).
38. C. Breucker, A. Steiff, P.-M. Weinspach, *Proceedings of the European Conference on Mixing and Centrifugal Separation / Fluid Engineering*, vol. 6, British Hydromechanics Research Association, Cranfield, Bedford, 1988, pp. 399–406.
39. J.G. Van de Vusse, *Chem.-Ing.-Tech.* **31**, 583–587 (1959).
40. S. Nagata, *Mixing: Principles and Applications*, Kodansha, Tokyo, and Wiley, New York, 1975.
41. H.-J. Henzler, *Untersuchungen zum Homogenisieren von Fluessigkeiten oder Gasen, VDI-Forschungheft*, vol. 587, VDI-Verlag, Duesseldorf, 1978.
42. M. Zlokarnik, *Chem.-Ing.-Tech.* **39**, 539–548 (1967).
43. A. Mersmann, W.D. Einenkel, and M. Kaeppel, *Chem.-Ing.-Tech.* **47**, 953–996 (1975).
44. M. Opara, *Verfahrenstechnik* **9**, 446–449 (1975).
45. M. Zlokarnik, in T.K. Ghose, A. Fiechter, and N. Blakebrough eds., *Advances in Biochemical Engineering*, vol. 8, Springer-Verlag, Berlin, 1978, pp. 133–151.
46. H. Judat, *Chem.-Ing.-Tech.* **54**, 520–521 (1982).
47. K. Van't Riet, *Ind. Eng. Chem. Process Des. Dev.* **18**, 357–375 (1979).
48. K. Akita and F. Yoshida, *Ind. Eng. Chem. Process Des. Dev.* **13**, 84–91 (1974).
49. H. Yagi and F. Yoshida, *Ind. Eng. Chem. Process Des. Dev.* **14**, 488–493, (1975).
50. J.F. Perez and O.C. Sandall, *AlChE J.* **20**, 770–775 (1974).
51. S. Sideman, O. Hortacsu, and J.W. Fulton, *Ind. Eng. Chem. (Int. Ed.)* **58**, 32–47 (1966).
52. Y. Kawase and M. Moo-Young, *Chem. Eng. Res. Des.* **66**, 285–288 (1988).
53. J.G. Mooyman, *Biotechnol. Bioeng.* **29**, 180–186 (1987).
54. D.B. Mills, R. Bar, and D.J. Kirwan, *AlChE J.* **33**, 1542–1549 (1987).
55. H. Oguz, A. Brehm, and W.-D. Deckwer, *Chem. Eng. Sci.* **42**, 1815–1822 (1987).
56. J. Hoffmann, K. Buescher, and D.C. Hempel, *Chem.-Ing.-Tech.* **67**, 210–214 (1995).
57. J. Hoffmann, J. Tralles, and D.C. Hempel, *Chem.-Ing.-Tech.* **64**, 953–956 (1992).
58. R.S. Cherry and K.-Y. Kwon, *Biotechnol. Bioeng.* **36**, 563–571 (1990).
59. R.K. Geisler, Dissertation, Technische Universitaet Munich, Germany, 1991.
60. F. Liepe, *Verfahrenstechnische Berechnungsmethoden, vol. 4: Stoffvereinigen in fluiden Phasen*, VCH, Weinheim, 1988.
61. H.-J. Henzler and A. Biedermann, *Chem.-Ing.-Tech.* **68**, 1546–1561 (1996).
62. K. Buescher and D.C. Hempel, *Chem.-Ing.-Tech.* **68**, 1452–1455 (1996).
63. B.K. Revill, *4th Eur. Conf. on Mixing*, Paper B1, Noordwijkerhout, Netherlands, 1982.
64. J.Y. Oldshue, in C.S. Ho and J.Y. Oldshue eds., *Biotechnology Processes: Scale-Up and Mixing*, AlChE, New York, 1987, pp. 3–5.
65. B. Metz, N.W.F. Kossen, J.C. Van Suijdam, in T.K. Ghose, A. Fiechter, and N. Blakebrough eds., *Advances in Biochemical Engineering*, vol. 11, Springer-Verlag, Berlin, 1979, pp. 103–156.
66. F. Liepe, H.O. Moeckel, *Chem. Tech. (Leipzig)* **28**, 205–209 (1976).
67. S. Sicardi, R. Conti, G. Baldi, L. Franzino, *Chem.-Ing.-Tech.* **53**, MS 870/81 (1981).
68. P.L.T. Brian, H.B. Hales, T.K. Sherwood, *AlChE J.* **15**, 727–733 (1969).
69. P. Harriot, *AlChE J.* **8**, 93–102 (1962).
70. P.H. Calderbank, M.B. Moo-Young, *Chem. Eng. Sci.* **16**, 39–54 (1961).
71. E.O. Karow, W.H. Bartholomew, M.R. Sfat, *J. Agric. Food. Chem.* **1**, 302–312 (1953).
72. E.S. Bylinkina, V.V. Birukov, S.A. Voroshilova, in H. Dellweg ed., *Abstracts 5th International Fermentation Symposium*, Berlin, 1976.
73. O.G. Wegrich, R.A. Shurter, *Ind. Eng. Chem.* **45**, 1153–1160 (1953).
74. J. Strohm, R.F. Dale, H.J. Peppler, *Appl. Microbiol.* **7**, 235–238 (1959).
75. M. Jarai, in G. Terni ed., *4th International Fermentation Symposium Proceedings*, Fermentation Technology Today, Osaka, 1972, pp. 97–105.
76. W.H. Bartholomew, in Umbreit ed., *Advances in Applied Microbiology 2*, Academic Press, New York, 1960, pp. 289–300.
77. H. Takei, K. Mizusawa, and F. Yoshida, *J. Ferment. Technol.* **53**, 151–158 (1975).
78. H. Taguchi, T. Imomanaka, S. Teramoto, M. Takatsu, and M. Sato, *J. Ferment. Technol.* **46**, 823–828 (1968).
79. R. Fuchs, D.D.Y. Ryu, A.E. Humphrey, *Ind. Eng. Chem. Process Des. Dev.* **10**, 190–196 (1971).
80. W.D. Maxon, *J. Biochem. Microbiol. Tech. Eng.* **1**, 311–324 (1959).
81. D.I. Wang, R.C.J. Fewkes, in L.A. Underkofler ed., *Developments in Industrial Microbiology*, American Institute of Biological Science, Washington, D.C., 1977, pp. 39–56.
82. E.J. Gaden, Jr., *Sci. Rep. Inst. Superore di Sanita* **1**, 161–176 (1961).
83. A.E. Humphrey, *J. Ferment. Technol.* **42**, 265–278 (1964).
84. A.E. Humphrey, *J. Ferment. Technol.* **42**, 334–345 (1964).

See also BIOREACTORS, AIR-LIFT REACTORS; BIOREACTORS, FLUIDIZED-BED; BIOREACTORS, CONTINUOUS STIRRED-TANK REACTORS; DIMENSIONAL ANALYSIS, SCALE-UP; FERMENTATION MONITORING, DESIGN AND OPTIMIZATION; MAMMALIAN CELL CULTURE REACTORS, SCALE-UP; MASS TRANSFER.

SECONDARY METABOLITE PRODUCTION, ACTINOMYCETES, OTHER THAN STREPTOMYCES

ANN C. HORAN
Schering Plough Research Institute
Kenilworth, New Jersey

KEY WORDS

Actinomycetes
Cell wall composition
Chemotaxonomy
Fatty acids
Fermentation
Menaquinones
Phospholipids
Secondary metabolites

OUTLINE

INTRODUCTION

Actinomycetes are an industrially important group of microorganisms, producing, in fermentation, numerous therapeutically relevant natural products. In the pharmaceutical and biotechnology industries, they continue to be the focus of many natural product screening programs, especially in the search for novel antibiotics active against

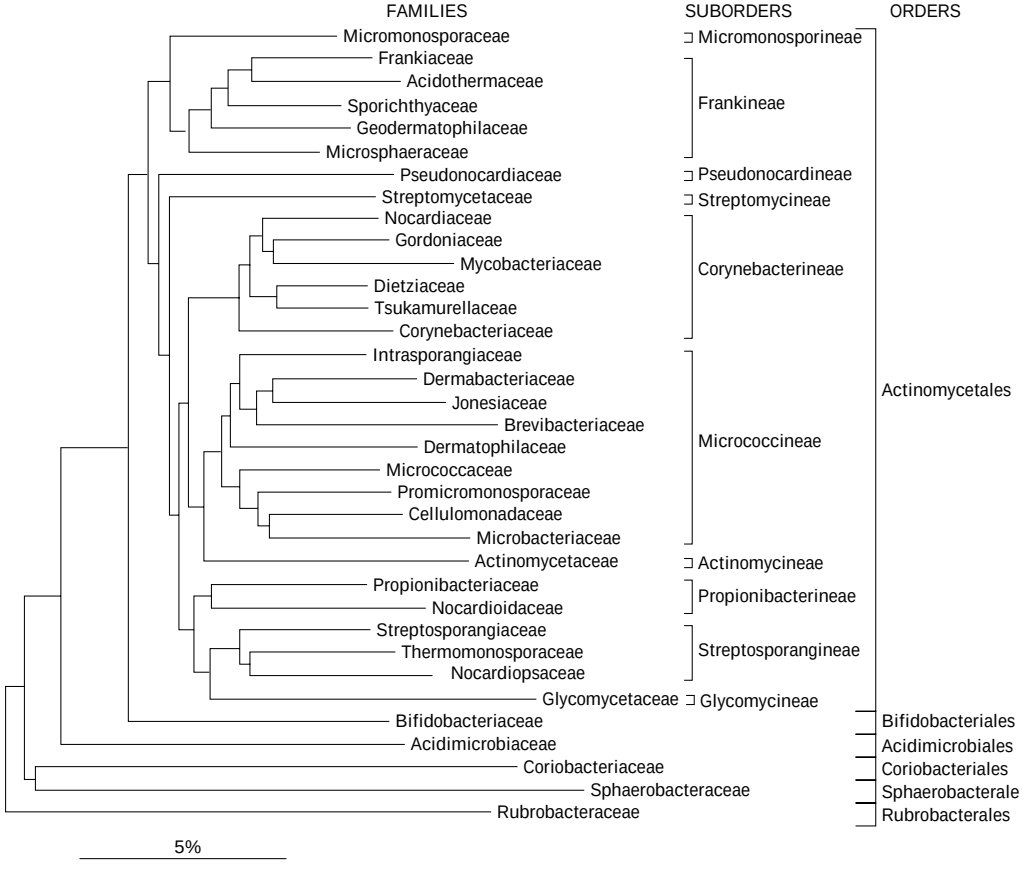


Figure 1. Relatedness of actinobacteria showing the presence of 6 orders and 10 suborders of the Actinomycetales based on 16S rDNA/rRNA sequence comparison. The scale bar represents 5 nucleotide substitutions per 100 nucleotides. *Source*: From Stackebrandt et al. (2).

Table 1. Morphological Characteristics of Members of the Families Streptosporangiaceae, Nocardiopsaceae, Thermomonosporaceae, Pseudonocardiaceae, Micromonosporaceae, and Glycomycetaceae

Family/genus	Substrate mycella	Aerial mycelia	Spore chains	Spores
		Streptosporangiaceae		
Streptosporangium	Branching, nonfragmenting; yellow, brown, orange, red	Present; white, pink to greenish-gray; spore vesicles borne on long or short sporangiophores, singly or clustered	Spore vesicles contain a coiled chain of arthrospores formed by septation of an unbranched, spiral hyphae within an expanded sporangiophore sheath.	Nonmotile; smooth, spherical, oval or rod-shaped
Herbidospora	Branching, nonfragmenting; colorless, tan to yellow-brown	Absent; spore chains, in mass, white	Straight, short chains, 10 to 30 spores, borne on tips of branching sporophores, in clusters, directly from the vegetative mycelium	Nonmotile; smooth, oval
Microbispora[a]	Branching, nonfragmenting; yellow, brown, orange, violet	Present; pink to white; bearing longitudinal pairs of spores often closely arranged along the hyphae, either sessile or on short sporophores	Spores in pairs	Nonmotile; smooth, spherical to oval
Microtetraspora	Branching, nonfragmenting; yellow, brown, orange, violet	Present; white, pink, blue-gray; short, sparsely branched	Spores in chains of 4	Nonmotile; smooth, warty, spherical to oval
Planobispora	Branching, nonfragmenting; colorless or rose	Present; white to light rose; bearing cylindrical to clavate sporangia, singly or in bundles, on short sporangiophores	Longitudinal pairs of spores in sporangia	Motile; peritrichous flagella; smooth, straight to slightly curved
Planomonospora	Branching, nonfragmenting; grayish-yellow, pink, reddish, orange	Present; white to rose; bearing cylindrical to clavate sporangia, sessile or on sporangiophores	Single spore in each sporangium	Motile; peritrichous flagella; straight to curved.
		Nocardiopsaceae		
Nocardiopsis	Branching, fragmenting into rods and coccal elements; yellow, yellowish-brown, olive	Present; white, yellowish-gray; long, moderately branched; straight to flexous, zigzag fragmentation	Fragmenting into long (greater than 50) chains of spores	Nonmotile; smooth, irregular
		Thermomonosporaceae		
Thermomonospora[b]	Branching, nonfragmenting; colorless to pale yellow	Present; white; differentiates into a single, heat sensitive spore	Single, sessile, or on the end of short branched or unbranched sporophores	Nonmotile; smooth to ridged
Actinomadura	Branching, nonfragmenting; colorless to yellow, orange, brown, red-brown, red, violet, gray-green	Present; white, yellow, pink, blue, green, gray; branching, carrying short to long chains of arthrospores	Straight, flexous, hooked, open loops, irregular spirals of 1 to 4 turns	Nonmotile; folded, irregular, smooth, spiny, or warty
Spirillospora[c]	Branching, nonfragmenting; white, pale yellow, pale pink, red to reddish brown	Present; white; bearing spherical sporangia on sporangiophores	Each spore vesicle contains coiled, branched aerial hyphae that differentiate into spores.	Motile, 1 to 7 subpolar flagella; smooth, rod-shaped, curved

Table 1. Morphological Characteristics of Members of the Families Streptosporangiaceae, Nocardiopsaceae, Thermomonosporaceae, Pseudonocardiaceae, Micromonosporaceae, and Glycomycetaceae (continued)

Family/genus	Substrate mycella	Aerial mycelia	Spore chains	Spores
		Micromonosporaceae		
Micromonospora	Branching, well developed; tan to orange, brown, blue-green, purple	Absent or a gray to white bloom	Borne singly, sessile or on short sporophores, in branched clusters	Nonmotile; smooth to warty, oval
Actinoplanes	Branching, nonfragmenting; tan to red, brown, violet	Absent or scanty, sporangia, spherical, subspherical, clavate, irregular, arising from the substrate mycelia, sessile or on short sporangiophores	Within the sporangia	Motile, polar flagella; spherical or short rods
Catellatospora	Branching, nonfragmenting; tan, yellow, mustard gold, orange, reddish brown.	Absent	Short, straight to flexous, branched chains of spores directly from the substrate mycelia	Nonmotile; cylindrical, ovoid, smooth to slightly rough
Couchioplanes	Branching, nonfragmenting; dark blue	Present; branching	Short spore chains in irregular spirals, directly from the substrate mycelia or on short aerial mycelia	Motile, polar flagellation; oval, short rods, smooth
Catenuloplanes	Branching, nonfragmenting; orange, red, brown	Present; sparse, branching	Short spore chains in spirals of 1 to 2 turns, from the substrate or aerial mycelia	Motile, peritrichous flagella; rod-shaped, straight to curved, smooth
Dactylosporangium	Branching, nonfragmenting; orange, rose, or wine.	Absent; sporangia present, borne on short sporangiospores, finger-shaped to claviform	Within the sporangia a single row of 3 to 4 spores	Motile, polar flagella; oblong, ovoid
Pilimelia	Branching, nonfragmenting; pale lemon-yellow, golden yellow, orange	Absent; sporangia present, spherical ovoid, pyriform, cylindrate, borne on sporangiospores	Within the sporangia, in parallel or swirl rows, and as conidia, in chains	Motile, sporangiospores, laterally inserted tuft of flagella; nonmotile, rod-shaped conidiospores
Spirilliplanes	Present, nonfragmenting, branching, yellow to orange	Absent	Short chains of spores arranged in spirals and clusters directly from the substrate mycelia	Motile, oval to short rods, smooth, flagella type not given
		Pseudonocardiaceae		
Pseudonocardia	Zigzag fragmentation and long chains of spore-like structures; mycelial wall with a characteristic electron dense outer layer	Present; white; fragmenting	Spore chains long, 50 or more per chain	Nonmotile; smooth, spiny, irregular
Actinopolyspora	Branching, occasionally fragments; buff	Present; white, with an outer sheath	Long chains of spores	Nonmotile; smooth
Actinosynnema	Fragmenting	Present; form synnemata (dome-like structures)	Chains of arthrospores at the tip of the synnemata	Motile arthrospores (flagellation not given)
Amycolatopsis	Fragments into long chains of squarish to ellipsoid structures	Rarely formed; when present, white; may fragment into chains of spores	Chains of squarish to oval fragments	Nonmotile; oval, smooth, warty
Kibdelosporangium	Zigzag fragmentation	Present; white; sporangia-like structures that do not contain spores but germinate when placed on agar	Long chains	Nonmotile; smooth, irregular

Table 1. Morphological Characteristics of Members of the Families Streptosporangiaceae, Nocardiopsaceae, Thermomonosporaceae, Pseudonocardiaceae, Micromonosporaceae, and Glycomycetaceae (*continued*)

Family/genus	Substrate mycella	Aerial mycelia	Spore chains	Spores
Kutzneria	Branching, nonfragmenting	Present; white, olive gray; branching with globose sporangia	Long chains formed by separation of coiled, unbranched hyphae	Nonmotile; spherical, rod-shaped or oval
Lentzea	Branching, nonfragmenting; yellow to yellow brown	Present; white to whitish-yellow; fragmenting into rod-shaped elements	Chains of rod-like structures	Nonmotile; smooth, rod-like
Saccharomonospora	Branching, rarely fragments; beige, lilac, green	Present; white becoming green; with an outer sheath	Single densely packed, on the aerial mycelia	Nonmotile; oval, smooth, warty
Saccharopolyspora	Fragmenting; spores often present	Present; white, pink to brownish gray, often sparse; with an outer sheath	Long and short chains	Nonmotile; spore sheath ornamentation hairy, spiny, smooth
Saccharothrix	Zigzag fragmentation; yellow to yellow brown	Present; white to yellowish-gray, zigzag fragmentation	Chains of ovoid to irregular elements	Nonmotile; smooth, ovoid
Streptoalloteichus	Branching, bearing spore-like vesicles containing one to several spores	Present; spore chains in clusters and sclerotia	Spore chains on the aerial mycelia, sporangiospores on the substrate mycelia	Motile sporangiospores by a single long polar flagellum
Thermocrispum	Branching, nonfragmenting; yellow to light brown	Present; white; aggregating into clusters that fragment into rod-like structures	Chains of rod-like structures	Nonmotile; smooth, rod-like
		Glycomycetaceae		
Glycomyces	Branching, nonfragmenting; pale yellow to tan	Present; white; forming spores	Short chains	Nonmotile; square-ended
		Genera of Uncertain Affiliation		
Actinobispora	Branching, nonfragmenting; bearing paired spores; light orange	Present; white; sparse; bearing paired spores	Longitudinal pairs	Nonmotile; smooth, spherical
Actinocorallia	Branching, nonfragmenting; ivory to pale yellow	Absent	Coralloid sporophores arising from the substrate mycelia bearing long chains (30+) of spores	Nonmotile; smooth, cylindrical
Actinokineospora	Branching, nonfragmenting; (on agar); colorless to yellow brown	Present; white; bearing chains of spores in a sheath	Irregularly curved; less than 50 per chain	Motile, peritrichous flagella; smooth, ovoid to squarish
Planotetraspora	Branching, nonfragmenting; ivory	Present; whitish gray; cylindrical sporangia on short sporangiophores	Single row of 4 spores per sporangium	Motile, a single polar flagellum
Thermobispora	Branching, nonfragmenting; yellow to yellow brown	Present; white; monopodially branched; bearing pairs of spores	In longitudinal pairs on the aerial hyphae	Nonmotile; smooth, spherical to oval

[a] *Microbispora bispora* has been removed from the genus to form a new, thermophilic single species genus, *Thermobispora* (7).
[b] The genus *Thermomonospora*, as defined by Kroppenstedt and Goodfellow (8), contains two species, *T. curvata* and *T. formosensis*.
[c] The genus *Spirillospora* was last placed in the family Streptosporangiaceae (9) based on 16S rDNA but has now been moved to the family Thermonomosporaceae (2).

Table 2. Cell Wall Types and Corresponding Whole-Cell Sugar Patterns of Aerobic Actinomycetes Containing *meso*-Diaminopimelic Acid

Cell Wall		Whole-cell sugar pattern		
Type	Distinguishing major constituents	Type	Diagnostic sugars	Representative genera
II	Glycine	D	Xylose, arabinose	*Micromonospora*
III	None	B	Madurose[a]	*Actinomadura*
		C	None	*Nocardiopsis*
IV	Arabinose, galactose	A	Arabinose, galactose	*Amycolatopsis*

Source: According to Lechevalier and Lechavalier (13).
[a]Madurose is 3-*O*-methyl-D-galactose.

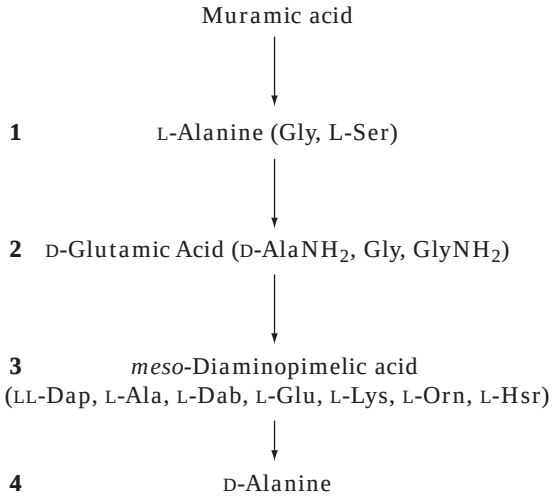


Figure 2. The peptide stem of the actinomycete (and other Gram-positive bacteria) peptidoglycan and the possible variations in amino acid content (in parentheses). Ala, alanine; Dab, diaminobutyric acid; Glu, glutamic acid; Gly, glycine; Hsr, homoserine; Lys, lysine; Orn, ornithine; Ser, serine. *Source*: From Suzuki et al. (15).

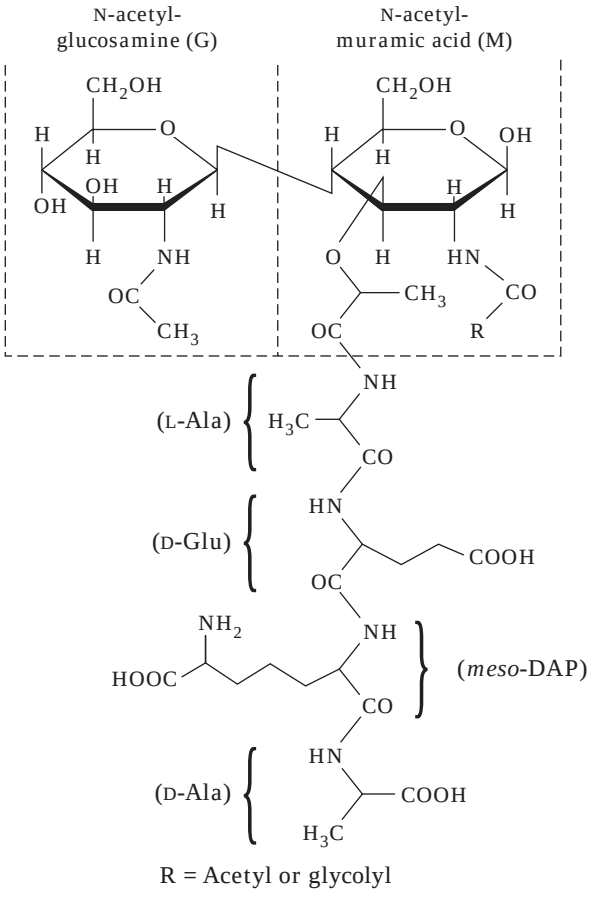


Figure 3. Complete structure of a single subunit of the peptidoglycan showing the linkage between the two amino sugars that make up the glycan strand and between muramic acid and the four amino acids in the peptide. Also indicated is the position of acetyl/glycolyl on muramic acid. *Source*: Modified from Suzuki et al. (15).

highly resistant, infectious Gram-positive pathogens. In an actinomycete-based natural product screening program, the organisms must be isolated from the environment, grown in pure state, preserved, and fermented, and the fermentation products must be tested for biological activity. Essential to this process is a knowledge of actinomycete ecology, taxonomy, and physiology, all of which contribute to the isolation of new strains with the potential of producing novel compounds.

Actinomycetes are Gram-positive, filamentous bacteria with a mol % G + C content of DNA greater than 50%. They can be divided into two broad groups: the fermentative organisms that are found in the natural cavities of man and animals and the larger group of oxidative organisms found in soil (1). The majority of oxidative soil actinomycetes can be further subdivided into two large groups based on the isomer of diaminopimelic acid in the peptidoglycan: the streptomycetes, containing LL-diaminopimelic acid (DAP) and the remainder containing *meso*-DAP. Further subdivisions to family, genus, and species relies on analysis of actinomycete morphology, determination of chemical constituents of cell walls and whole cells, biochemical characteristics, and resistance patterns. Sequence data generated from the analysis of small subunit RNA (16S) have been applied to the understanding of actinomycete phylogeny as well as molecular identification. A new hierarchical classification based on 16S rDNA/rRNA sequences (Fig. 1), proposed by Stackebrandt et al. (2), will be followed as phenotypic data are presented. Aerobic, saprophytic, filamentous actinomycetes containing *meso*-DAP in their cell

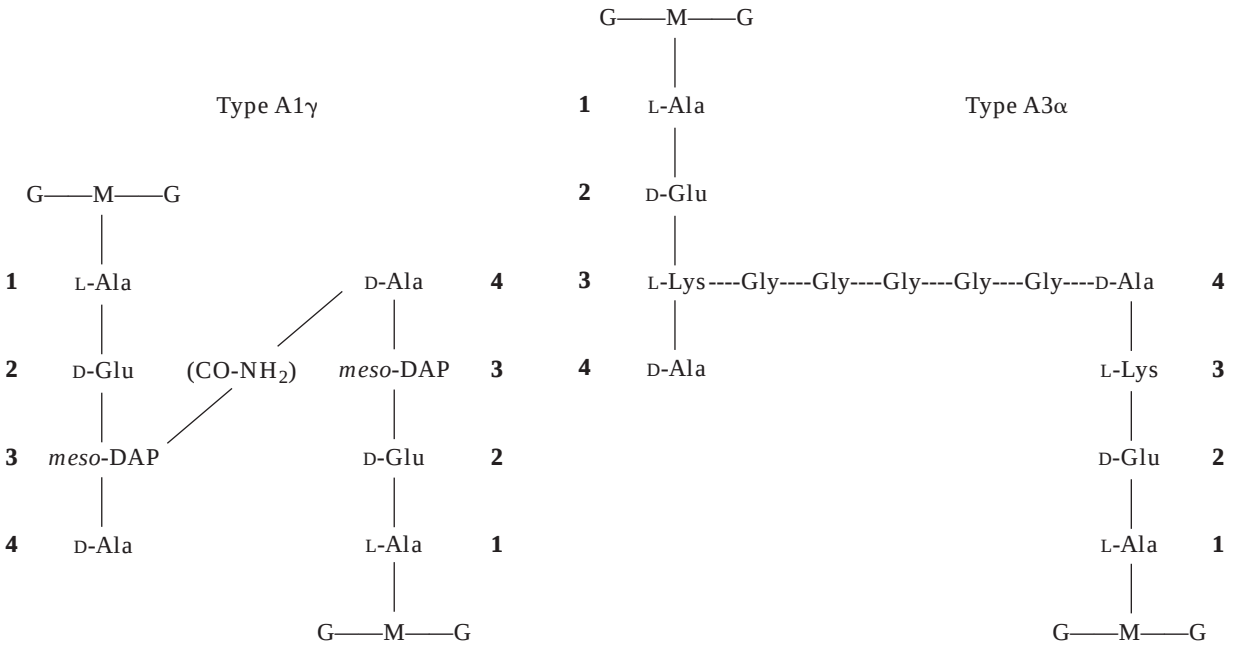


Figure 4. Example of a directly cross-linked *meso*-DAP–containing peptidoglycan of type A (cross-linkage between positions 3 and 4 of two peptide subunits), 1 (no peptide bridge), and γ (*meso*-DAP in position 3), and an example of cross-linking through a peptide bridge, A (as above), 3 (the bridge is an oligopeptide), and α (L-lysine in position 3. Ala, alanine; Dap, diaminopimelic acid; G, *N*-acetylglucosamine; Glu, glutamic acid; Gly, glycine; Lys, lysine; M, *N*-acetyl- or *N*-glycolylmuramic acid. *Source*: Modified from Suzuki et al. (15).

walls, which have not been placed in any of the actinomycete families defined by 16S rDNA sequence data, including Actinobispora (3), Actinocorallia (4), Actinokineospora (5), Planotetraspora (6), and Thermobispora (7) will also be described. The suborders Actinomycineae, Micrococcineae, Corynebacterineae, Propionibacterineae and Frankineae (2) will not be discussed in this article.

IDENTIFICATION OF NONSTREPTOMYCETE ACTINOMYCETES

Morphological Characterization

Meso-DAP–containing soil actinomycetes have a complex morphology. When grown on agar, a substrate mycelia forms that may penetrate the agar and that can bear spores, sporelike structures, or sporangia. Aerial mycelia may arise from the substrate mycelia and can be branching. The aerial mycelia may be in a sheath or in a sporangium and may fragment into motile or nonmotile spores. Morphological characteristics can be observed by growing the organisms on minimal media, for example, water agar (crude agar, 15 g; tap water, 1,000 mL), dilute Bennett's agar (agar, 15 g; glucose, 1 g; pancreatic digest of casein, 0.2 g; yeast extract, 0.1 g; beef extract, 0.1 g; distilled water, 1,000 mL), or dilute yeast–dextrose agar (agar, 15 g; yeast extract, 1 g; dextrose, 1 g; tap water, 1000 mL). After 7–28 days of incubation at 28 °C, plates can be examined directly under a light microscope using a long working distance, 40× objective. Detailed morphological analysis, such as spore and sporangiophore ornamentation, location and attachment of flagella on motile spores, and sporangial structures and spore formation within the sporangia will require either scanning or electron microscopic observations. In some cases, actinomycetes can be identified to genus based on their characteristic morphology (Table 1).

Chemical Characterization

Although the majority of actinomycetes readily form morphological structures, many do not, making preliminary identification based on this characteristic impossible. The laboratories of Huber and Mary Lechevalier at the Waksman Institute, Rutgers University, New Jersey, pioneered the extensive use of chemical markers to delineate genera of actinomycetes. The first large-scale, systematic application of chemical criteria to taxonomy focused on the analysis of cell-wall and whole-cell sugars and amino acids. Representatives of all the described actinomycete genera were analyzed, and distinct patterns were discerned (11,12) (Table 2). In conjunction with the data generated by Schleifer and Kandler (14) delineating the mode of cross-linking between subunits of the Gram-positive peptidoglycan (Figs. 2 to 4), a clearer understanding of the relationship of wall structure to wall chemotype was obtained (15) (Tables 3 and 4). The presence of either glycolyated or acetylated muramic acid in the cell wall (17,18) and lipids such as mycolic acids (19), phospholipids (20), menaquinones (21), and fatty acids (22,23) resulted in additional taxonomic discrimination. Phospholipid patterns are presented in Table 5, and the structure of typical ac-

Table 3. Peptidoglycan Classification

Position of cross-link	Peptide bridge	Amino acid at position 3
	Peptidoglycan A	
Cross-linkage between positions 3 and 4 of two peptide subunits	1. None	α-L-Lysine β-L-Ornithine γ-*meso*-DAP
	2. Polymerized subunits	α-L-Lysine
	3. Monocarboxylic L-amino acid or glycine or oligopeptides	α-L-Lysine β-L-Ornithine γ-LL-DAP
	4. Contains a dicarboxylic amino acid	α-L-Lysine β-L-Ornithine γ-*meso*-DAP δ-L-DAB
	Peptidoglycan B	
Cross-linkage between positions 2 and 4 of two peptide subunits	1. Contains an L-amino acid	α-L-Lysine β-L-Homoserine γ-L-Glutamic acid δ-L-Alanine
	2. Contains a D-amino acid	α-L-Ornithine β-L-Homoserine γ-L-DAB

Source: From Suzuki et al. (15) based on Schleifer and Kandler (14).
Note: DAB, diaminobutyric acid.

Table 4. Relationship Between Wall Chemotype and Peptidoglycan Type in Actinomycetes

Wall type[a]	Major wall amino acids[a]	Other distinguishing characteristics[b]	Peptidoglycan type[c]	Representative genera
I	Glycine, LL-DAP	None	A3γ	*Streptomyces*
II	Glycine, *meso*- or OH-DAP	None	A1γ	*Actinoplanes* *Micromonosporaa*
III	*meso*-DAP	Madurose in whole cells	A1γ	*Actinomadura* *Streptosporangium*
IV	*meso*-DAP	Arabinose and galactose in cell wall	A1γ	*Amycolatopsis* *Nocardia*
V	Lysine, ornithine	None	A5β	*Actinomyces*
VI	Lysine (aspartic acid, var.)	Galactose (var.)	A3α, A4α	*Couchioplanes, Oerskovia*
VII	Diaminobutyric acid, glycine (lysine, var.)	None	B2α	*Agromyces*
VIII	Ornithine	None	B2β	*Bifidobacterium*

Source: From Minnikin and Goodfellow (16).
Note: var. = variable
[a]Major constituents of actinomycete cell walls according to Lechevalier and Lechevalier (13).
[b]All cell wall preparations contain major amounts of alanine, glutamic acid, glucosamine, and muramic acid.
[c]From Schleifer and Kandler (14).

tinomycete fatty acids and the patterns obtained from their analysis are presented in Tables 6 and 7. Chemotaxonomic markers aid in the identification of actinomycetes to the genus level. Further discrimination to the species level can be achieved using physiological tests, including utilization of various carbon and nitrogen sources; hydrolysis of complex carbohydrates and macromolecules; and resistance to chemicals, including antibiotics (24,25). Sequence analysis, particularly of the 16S rDNA, may play a significant role in future definitions of actinomycete species. However, because of the highly conserved nature of small subunit RNA, very high similarity levels (greater than 97%) among related strains will be observed. In order to delineate species within these groups, DNA–DNA reassociation may be required (26).

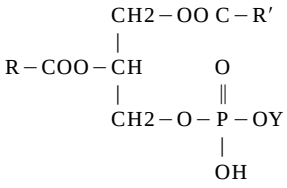
The chemotaxonomic characteristics for each genus of the families described in the next section are presented in Table 8.

CHARACTERISTICS OF ACTINOMYCETE FAMILIES

Streptosporangiaceae

Streptosporangiaceae (9,27) are aerobic, Gram-positive, non-acid-fast, chemoorganotrophic actinomycetes that form stable substrate and aerial mycelia. Aerial mycelia

Table 5. Phospholipid Patterns of Aerobic Actinomycetes



R, R′ = Long chain alkyl

Y = Polar head groups, PG, DPG,[a] PI,[a] PIM, PE, PC, PME

Phospholipid Pattern	Phosphatidyl-ethanolamine (PE)	Phosphatidylmethyl-ethanolamine (PME)	Phosphatidyl-choline (PC)	Phosphatidyl-glycerol (PG)	GluNU[a]
PI	Absent	Absent	Absent	Variably present	Absent
PII	Present	Absent	Absent	Absent	Absent
PIII	Variably present	Variably present	Present	Variably present	Absent
PIV	Variably present	Variably present	Absent	Absent	Present
PV	Absent	Absent	Absent	Present	Present

Source: According to Lechevalier et al. (20).
[a]DPG, diphosphatidy glycerol; PI, phosphatidylinositol; GluNU, phospholipids of unknown structure containing glucosamine.

Table 6. Structure of Representative Actinomycete Fatty Acids

Fatty acid type	Common example and structure
Straight-chain	*Hexadecanoic* CH_3-$(CH_2)_{14}$-COOH
Cis-unsaturated	*Oleic* CH_3-$(CH_2)_7$-CH = CH-$(CH_2)_7$-COOH *cis*
Tuberculostearic	*10-Methyloctadecanoic* CH_3 CH_3-$(CH_2)_7$-CH-$(CH_2)_8$-COOH
Cis-cyclopropane	*11,12-Methyleneoctadecanoic* CH_2 CH_3-$(CH_2)_5$-CH-CH-$(CH_2)_9$-COOH *cis*
Iso	*13-Methyltetradecanoic* CH_3 CH_3-CH-$(CH_2)_{11}$-COOH
Anteiso	*12-Methyltetradecanoic* CH_3 CH_3-CH_2-CH-$(CH_2)_{10}$-COOH

Source: From Minnikin and O'Donnell (16).

may carry chains of two or more spores or be differentiated into sporangia (spore vesicles) containing one, two, or more spores. Spores are either motile or nonmotile. Cell walls contain *meso*-DAP and N-acetylated muramic acid, but lack characteristic sugars; whole cells contain madurose. The peptidoglycan is of the A1γ type, having a cross-linkage between positions 3 and 4 of adjacent peptide subunits, with *meso*-DAP at position 3 of the peptide stem (14). Lipid profiles contain saturated, *iso*- and *anteiso*-fatty acids; major proportions of tetrahydrogenated menaquinones, with nine isoprene units saturated at sites III and VIII; and glucosamine-containing polar lipids. Mycolic acids are absent. The mol % G + C content of the DNA ranges from 66 to 74.

Nocardiopsaceae

Nocardiopsaceae (28,29) are aerobic, Gram-positive, non-acid-fast, chemoorganotropic actinomycetes that form well-developed, branching substrate mycelia that may fragment into coccoid to bacillary forms. Aerial mycelia are long and moderately branching, with zigzag fragmentation into irregular spores. The family is defined on the basis of chemotaxonomic markers because the morphological characteristics are the same for the genus *Saccharothrix* (family Pseudonocardiaceae). Cell walls contain *meso*-DAP, N-acetylated muramic acid, and no characteristic sugars. The peptidoglycan is A1γ. Lipid profiles contain phosphatidylcholine and phosphatidylmethylethanolamine; major portions of variably saturated menaquinones with 10 isoprene units; *iso*-branched, *anteiso*-branched, and 10-methyl-branched fatty acids; and high levels of octadecanoic acid. The combination of 15 to 20% *anteiso*-C17:0 (14-methylhexadecanoic acid) with 20 to 25% 10-methyl-C18:0 (tuberculostearic acid) or its precursor, oleic acid, is diagnostic and unique among bacteria. Mycolic acids are absent. The G + C content of the DNA is 64 to 69 mol %. The family contains only one genus, *Nocardiopsis*.

Thermomonosporacea

Thermomonosporacea (8) are aerobic, Gram-positive, non-acid-fast, chemoorganotrophic actinomycetes producing branched substrate mycelia bearing aerial mycelia that differentiate into single or short chains of arthrospores. Cell walls contain *meso*-DAP, N-acetylated muramic acid, and no characteristic sugars. The peptidoglycan type is

Table 7. Fatty Acid Types

Type	*Iso*-15:0	*Anteiso* 15:0	*Iso*-16:0	16:0	10-Methyl 16:0	*Iso*-17:0	*Anteiso*-17:0	17:0	17:1	10 Methyl 17:0	*Iso*-18:0	18:0	18.1	10-Methyl 18:0	OH_2	Cyclo-19
1a	−	−	−	+ + +	−	−	−	−	−	−	−	+	+ + +	−	−	−
1b	−	−	−	+ + +	+ +	−	−	−	−	+	−	+	+ +	+ + +	−	−
1c	−	−	−	+ + +	−	−	−	−	−	−	−	+	+ +	−	−	+ +
2a	+	−	+ + +	+ +	−	+	+	+	+	−	+	−	−	−	+ +	−
2b	+ + + +	+ +	+ +	+	−	+ + +	+	−	−	−	−	−	−	−	−	−
2c	+ +	+ + +	+ + +	+	−	+	+ +	−	−	−	−	−	−	−	(V)	−
2d	+ + +	+ +	+ + + +	+ +	−	+ +	+ +	+	+	−	+	+	+ +	−	−	−
3a	−	−	+ +	+ + +	+	−	−	+	+	+	+	+	+ +	+ +	−	−
3b	+ + + +	+	+ +	−	−	+ +	+ +	+	+ +	+	−	−	+	−	−	−
3c	+ +	+	+ + +	+	+	+	+	+ +	+	+ +	+	+	−	−	+ +	−
3d	+	+	+ + +	+	−	+	+ + +	+	+	+ +	+	+	+ +	+ +	−	−
3e	+ +	−	+ + + +	+	−	+	+	−	+	+	−	−	−	−	−	−
3f	+ +	+	+ + + +	+	−	+	+	+ +	+ +	+	−	−	−	−	+	−
3g	+ +	+	+ + + +	+	−	+	+ +	+	+	+	−	−	−	−	+	−

Source: From Kroppenstedt (23).
Note: + = 1.5%; + + = 5–15%, + + + = 15–25%, + + + + = 25%; (V) = variable, usually less than 2% for one component.

A1γ. Whole cells contain madurose. Lipid profiles contain mixtures of straight- and branched-chain fatty acids; hydrogenated menaquinones with nine isoprene units; and major amounts of phosphatidylglycerol, phosphatidylinositol, and phosphatidylmannosides, some strains contain minor amounts of phosphatidylglycerol. The G + C content of the DNA is mol % 66 to 72.

Micromonosporaceae

Micromonosporaceae (10,30,31) are aerobic, Gram-positive, and non-acid-fast, forming nonfragmenting, branching substrate mycelia that, in the genus *Micromonospora*, form nonmotile spores, singly either sessile or on short sporophores. Aerial mycelia are rarely formed, but when present are scanty. Sporangia, which may be present and can contain two to five or numerous spores, are formed on sporangiophores arising directly from the substrate mycelia. Globose, subglobose, or rod-shaped spores are formed in the sporangia or in chains arising from the substrate mycelia. In water the spores become motile by means of polar or lateral tufts of flagella. Cell walls contain *meso*- or 3-hydroxy-DAP, or both. The peptidoglycan is of the A1γ type; glycine is the first amino acid (acetate in *Pilimelia*) in the stem peptide attached to a glycolated muramic acid. Whole-cell hydrolysates contain arabinose and xylose. Lipid profiles have saturated, *iso*-, and *anteiso*-fatty acids, and diphosphatidylglycerol, phosphatidylethanolamine, phosphatidylglycerol, and phosphatidylinositol as characteristic phopholipids; menaquinones are of varying lengths. Mycolic acids are absent. The mol % G + C content of the DNA ranges from 71 to 73.

Pseudonocardiaceae

Pseudonocardiaceae (32,33) are a heterogeneous grouping of organisms based on 16S rDNA/rRNA sequence data. Aerobic, Gram-positive, mesophilic or thermophilic, and chemoorganotrophic, one genus is halophytic. Heterogeneous morphologies include stable or fragmenting substrate and aerial mycelia; single or short chains of nonmotile spores on both the aerial and substrate mycelia; sporangial-like bodies not containing spores; and long chains of interwoven spores with motile spores at the tip. The peptidoglycan contains *meso*-DAP and is acetylated; many genera contain arabinose and galactose in their cell walls and whole cells, others do not. Rhamnose in combination with galactose in whole-cell hydrolysates appears to be diagnostic for some genera. Lipid profiles include mono-methyl-branched, straight-chain saturated and unsaturated, and hydroxy fatty acids and are genera specific; tetrahydrogenated menaquinones with 8, 9, or 10 isoprene units; phospholipid patterns II, III and IV are represented; mycolic acids are not present. The mol % G + C content of the DNA ranges from 64 to 79.

Glycomycetaceae

Glycomycetaceae (34) are aerobic, Gram-positive, non-acid-fast, and chemoorganotrophic, forming stable substrate mycelia and aerial mycelia that fragment into short chains of square-ended spores. The cell wall contains *meso*-DAP and glycine and N-glycolated muramic acid; the whole-cell sugar pattern is xylose and arabinose. The peptidoglycan is A1γ. Lipid profiles contain diphosphatidylglycerol and phosphatidylinositol; nitrogenous phospholipids are not found, nor are mycolic acids; tetra- and hexahydrogenated menaquinones with 10 isoprene units predominate. The mol % G + C content of the DNA is 71. The family contains only one genus, *Glycomyces*.

SCREENING ACTINOMYCETES FOR NATURAL PRODUCTS

Isolation from Natural Habitats

Studies have been performed (35–38) to discover appropriate conditions for isolating various groups of actinomycetes from natural habitats, thereby enhancing diversity and increasing the chances of discovering novel natural products. A testing paradigm was described by Horan (39) that evaluated the effect of soil type, media

Table 8. Chemotaxomic Characteristics of Members of the Families Streptosporangiaceae, Nocardiopsaceae, Thermomonosporaceae, Micromonosporaceae, Pseudonocardiaceae, and Glycomycetaceae

Family/genus	Diagnostic whole-cell sugars	Diagnostic cell-wall components[a]	Acyl type of muramic acid	Menaquinones	Phospholipid type	Fatty acid type[b]	mol % G+C
			Streptosporangiaceae				
Streptosporangium	Madurose	None	Acetyl	MK-9($H_{0,2,4}$)	IV	3c	69–71
Herbidospora	Madurose	None	Acetyl	MK-10($H_{2,4,6}$)	IV	3c	69–71
Microbispora	Madurose	None	Acetyl	MK-9($H_{0,2,4}$)	IV	3c	67–74
Microtetraspora	Madurose	None	Acetyl	MK-9($H_{0,2,4}$)	IV	3c	64–69
Planobispora	Madurose	None	Acetyl	MK-9($H_{0,2,4}$)	IV	3c	70–71
Planomonospora	Madurose	None	Acetyl	MK-9($H_{0,2,4}$)	IV	3c/a	72
			Nocardiopsaceae				
Nocardiopsis	None	None	Acetyl	MK-10($H_{4,6}$)	III	3d	64–69
			Thermomonosporaceae				
Thermomonospora	+/−Madurose	None	Acetyl	MK-9($H_{2,4,6,8}$)	I	3a/c	72
Actinomadura	Madurose	None	Acetyl	MK-9($H_{4,6,8}$)	I	3a	66–70
Spirillospora	Madurose	None	Acetyl	MK-9($H_{4,6}$)	I/II	3a	71–73
			Micromonosporaceae				
Micromonospora	Arabinose Xylose	OH-DAP Glycine	Glycolyl	MK-9(H_4) MK-10($H_{4,6}$) MK-12	II	3b	71–73
Actinoplanes	Arabinose Xylose +/−Galactose	OH-DAP Glycine	Glycolyl	MK-9(H_4) MK-10(H_4)	II	2d/2c	72–73
Catellatospora	Arabinose Xylose	OH-DAP Glycine 3-*O*-Methyl-rhamnose	Glycolyl	MK-10($H_{2,6}$) MK-9($H_{4,6}$)	III	2c	70.6 71.5
Couchioplanes	Arabinose Xylose Galactose	L-Lysine glycine Serine	Glycolyl	MK-9($H_{4,6}$)	II	2c	69.9 72.1
Catenuloplanes	Xylose +/−Galactose	L-Lysine Serine Glycine	Glycolyl	MK-9(H_8) MK-10(H_8)	III	2c	71–72
Dactylosporangium	Arabinose Xylose	Glycine	Glycolyl	MK-9($H_{2,6,8}$)	II	3b/2d	72–73
Pilimelia	Arabinose Xylose	Glycine	Glycolyl	MK-9($H_{2,4}$)	II	2d/2b	NR
Spirilliplanes	3-*O*-Methylmannose Xylose Galactose	Glycine	Glycolyl	MK-10(H_4)	II	2d	69
			Pseudonocardiaceae				
Pseudonocardia	Arabinose Galactose	Arabinose	Acetyl	MK-8(H_4)	III	3f/3e	68–79
Actinopolyspora	Arabinose Galactose Ribose	Arabinose Galactose	Acetyl	MK-9($H_{4,6}$)	III	2e/2c	64
Actinosynnema	None	None	Acetyl	MK-9($H_{4,6}$)	II	3d	71–73
Amycolatopsis	Arabinose Galactose	Arabinose Galactose	Acetyl	MK-9($H_{2,4}$)	II	3f	66–69
Kibdelosporangium	Arabinose Galactose tr-Madurose Rhamnose	Trace galactose	Acetyl	MK-9(H_4)	II	3f	66
Kutzneria	Galactose Rhamnose	None	Acetyl	MK-9(H_4)	II	3c/3e	70.3 70.7
Lentzea	None	None	Acetyl	MK-9($H_{0,2}$)	II	3d	68.6
Saccharomonospora	Arabinose Galactose	Arabinose Galactose	Acetyl	MK-9(H_4) MK-8(H_4)	II	2a	66–70

Table 8. Chemotaxomic Characteristics of Members of the Families Streptosporangiaceae, Nocardiopsaceae, Thermomonosporaceae, Micromonosporaceae, Pseudonocardiaceae, and Glycomycetaceae (*continued*)

Family/genus	Diagnostic whole-cell sugars	Diagnostic cell-wall components[a]	Acyl type of muramic acid	Menaquinones	Phospholipid type	Fatty acid type[b]	mol % G+C
Saccharopolyspora	Arabinose Galactose	Arabinose Galactose	Acetyl	MK-9($H_{2,4}$) MK-10(H_4)	III	3c	70–71.5
Saccharothrix	Galactose Rhamnose	None	Acetyl	MK-9(H_4)	II	3f	70–73
Streptoalloteichus	Galactose Rhamnose	Galactose	Acetyl	MK-9(H_6) MK-10(H_6)	II	NR	NR
Thermocrispum	Arabinose Trace galactose	None	Acetyl	MK-9(H_4)	II	3f	69–73
		Glycomycetaceae					
Glycomyces	Xylose Arabinose	Glycine	Glycolyl	MK-10($H_{2,4}$)	I	2c	71–73
		Genera of Uncertain Affiliation					
Actinobispora	Arabinose Galactose Xylose	None	Acetyl	MK-7(H_2) MK-9(H_2)	IV	NR	71
Actinocorallia	None	None	Acetyl	MK-9($H_{4,6}$)	II	1a	73
Actinokineosporia	Arabinose Galactose Rhamnose	Arabinose Galactose	Acetyl	MK-9(H_4)	II	3c	69–72
Planotetraspora	Arabinose Galactose Xylose Ribose	Glutamic acid	NR	NR	NR	NR	NR
Thermobispora	Madurose Galactose	None	Acetyl	MK-9($H_{0,2,4}$)	IV	3c	70

Note: NR, not reported.
[a]All organisms contain *meso*-DAP.
[b]Growth conditions may affect the profile (15).

constituents, and selective pressure on the isolation of actinomycetes. Resistance to antibiotics and physiological data, derived from taxonomic analysis of the various species of actinomycetes, formed the basis for the choice of selective pressures (antibiotics) and soil isolation media constituents. Soil samples were suspended in distilled water and serially diluted. The suspensions plated onto the surface of agar media containing various carbon and nitrogen sources and antibiotics, including the antifungal agent nystatin at 50 μg/mL. Resulting colonies were isolated to purity and identified microscopically to either genus or broad morphological group. Antibiotics in the agar medium had the most profound effect on the types of actinomycetes isolated, followed by soil type and media constituents. Novobiocin and spectinomycin at 20 μg/mL selectively enhanced the isolation of micromonosporae: up to 90% of isolates belonged to this group. Rifamycin (15 μg/mL), rosaramicin (10 μg/mL), everninomicin (10 μg/mL), and gentamicin (5 μg/mL) significantly reduced the streptomycete population, allowing diverse, *meso*-DAP containing actinomycetes to proliferate. Pretreatment of the soil sample, particularly air drying for 48 hours before plating, and rifamycin in the agar positively affected the isolation of nocardioform (fragmenting aerial and substrate mycelia) actinomycetes. Table 9 summarizes selective isolation techniques for non-streptomycete actinomycetes.

Growth and Fermentation of Isolates

Organisms are isolated from soil plates using sterile toothpicks, streaked onto the surface of a rich medium, usually American Type Culture Collection (ATCC) medium 72 (53); and grown at 28 °C for 5 to 7 days. Isolated colonies are transferred to broth (ATCC medium 172) in 25-mm tubes containing 10 mL of media. After 3 to 5 days' incubation at 28 °C on a rotary shaker at 200 to 250 rpm, the tubes are removed. Sterile glycerol is added to a final concentration of 10%, and 3 mL aliquots are distributed into sterile vials. The contents of a frozen vial are used as the inoculum that initiates the fermentation process. One milliliter of a thawed suspension is added to 10 mL of germination medium in a 25-mm tube stoppered with a Morton closure, and the tube is incubated for 3 to 5 days at 28 °C and shaken (200 to 250 rpm) on rotary shaker. After incubation, 1 mL of the germination tube growth (G1) is transferred into 10 mL of fermentation medium in 25-mm tubes and incubated as above. For primary screening, large numbers of diverse actinomycetes should be fermented in me-

Table 9. Methods for the Isolation of Actinomycetes from Natural Habitats

Genus	Habitat	Method
Streptosporangium *Kurtzneria*	Soil, possibly slightly acidic	Air dry soil, plate on AV agar containing penicillin G (0.8 μg/mL) and/or polymyxin (4 μg/mL) or gentamicin (2–5 μg/mL) (38,40)
Microbispora and *Microtetraspora*	Soil	Air dry soil, dry heat 120 °C/h, treat with 1.5% phenol, plate on AV, MGA-SE, and HV agars alone or on minimal media with rubromycin or streptomycin without phenol treatment (9,41)
Dactylosporangium, *Planobispora*, *Planomonospora*, *Spirillospora*	Soil	Baiting using organic substrates including grass, centrifugation for motile spored organisms, and plating on egg white agar containing nystatin (50 μg/mL) (42–44)
Herbidospora	Soil and plant material	Plant samples desiccated at 28 °C for 1 week, ground in blender with water, suspension incorporated into yeast extract agar containing nystatin and cyclohexamide; pour plates incubated 28 °C >2 weeks (45)
Thermonomonospora	Soil, mesophiles; overheated substances, bagasse, compost, fodders and manure, thermophiles	Sample heated to 100 °C, dilution plate onto media containing rifamycin (12 μg/mL) and/or kanamycin (25 μg/mL) (8,35)
Actinomadura	Soil	Air dry soil, dilution plate on media containing rifamycin (15 μg/mL), mefoxitin (10 μg/mL), or gentamicin (2–5 μg/mL) (38,39)
Micromonospora	Soil, lake muds, sediments	Air dry soil, dilution plate on soluble starch-yeast extract agar plus novobiocin (20 μg/mL), spectinomycin (20 μg/mL), or gentamicin (5–50 μg/mL) (39)
Actinoplanes *Couchioplanes* *Dactylosporangium*	Soil, sediments	Air dry soil, dilution plate on water agar, colloidal chitin agar or HV agar containing 0.1% potassium tellurite (30), or novobiocin (25 μg/mL) (27), or baiting and centrifugation (42)
Pilimelia	Soil, rarely encountered	Baiting using hair (30)
Catenuloplanes	Soil	Dilution plating on starch-casein agar plus nalidixic acid (25 μg/mL), kanamycin (12.5 μg/mL), cefsulodin (5 μg/mL), and kabicidin (6.25 μg/mL) (46)
Spirilliplanes	Soil	Dry heat 120 °C/h, dilution plating onto HV agar (10)
Amycolatopsis	Soil	Dilution plate on AV agar plus vancomycin (1–10 μg/mL) and polymyxin B (5 U/mL) (38); some species isolated from human specimens (47)
Saccharomonospora	Soil, thermophilic strains isolated from leaf litter, manure, compost	Isolate using Anderson Sampler, sedimentation chamber with penicillin G and polymyxin (5 μg/mL) added to half-strength nutrient agar (35)
Saccharopolyspora	Soil and decaying plant material	Dilution plate on yeast extract (1 g/L), glucose (1 g/L), agar (15 g/L) plus novobiocin (5 μg/mL) (A. Horan, unpublished data)
Kibdelosporangium	Soil	Ampicillin (4 μg/mL) and nalidixic acid (10 μg/mL) in complex medium incubated at 43 °C (48)
Saccharothrix	Soil	AV or starch-casein agars plus penicillin G (5–10 μg/mL) and nalidixic acid (15 μg/mL) or starch-casein-nitrate agar plus rifamycin (10 μg/mL) (38,39)
Streptoalloteichus	Soil, arid	Dilution plate on ISP-2 or Bennett's agars plus gentamicin or kanamycin (10 μg/mL), incubate 43 °C (49)
Thermocrispum	Municipal waste and mushroom composts	Dilution plate and Anderson Sampler, TSA plus rifamycin (10 μg/mL), erythromycin and oleandomycin (100 μ/mL), novobiocin (15 μg/mL), 50 °C (50)
Glycomyces	Soil	Dilution plate on Czapek-sucrose agar plus novobiocin (25 μg/mL) and streptomycin (15 μg/mL) (30)
Actinocorallia	Soil	Dilution plate on colloidal chitin-vitamin agar plus kabicidin (10 μg/mL) (4)
Actinokineospora	Soil, fallen leaves	Dilution plates plus kanamycin and/or nalidixic acid (50 μg/mL) or baiting and centrifugation (5,51)

Note: Detailed selective procedures are not reported for the following general: *Nocardiopsis*, *Actinoploy spora* (not isolated from a natural habitat, appeared as a contaminant), *Catellatospora*, *Pseudonocardia*, *Actinobispora*, and *Planotetraspora*. *Lentzea* has been isolated from human specimens (52).
[a]All media contain nystatin at 50–100 μg/mL.

dia that support the production of a wide variety of secondary metabolites. More than one medium is often used, three to four are optimal. The fermentation is harvested after 4 to 6 days and can either be extracted with solvent (butanol, methyl isobutyl ketone, or ethyl acetate) or filtered (paper or 10,000 mw cutoff filters). The extract or filtrate is tested for secondary metabolite production using biological activity or chromatography. Primary fermentations of wild-type strains, under the conditions outlined, can yield from 0.01 μg/mL to as much as 400 μg/mL of

Table 10. Methods for the Production of Actinomycete Secondary Metabolites

Parameter	Example 1 (54): saccharocarcins by *Saccharothrix aerocolonigenes*	Example 2 (55): megalomicin by *Micromonospora megalomicea*
Germination medium (per liter)	Glucose, 10 g; trehalose, 10 g; hydrolyzed casein, 5 g; soy flour, 5 g; yeast extract, 5 g; calcium carbonate, 2 g; tap water	Dextrose, 1 g; potato starch, 24 g; beef extract, 3 g; tryptone, 5 g; yeast extract, 5 g; calcium carbonate, 2 g; tap water
Germination conditions	G1 48 hr at 28 °C, 250 rpm; G2 48 h at 28 °C, 250 rpm	G1 72 h at 35 °C, 300 rpm
% inoculum	10	5
Fermentation medium (per liter)	Dextrin, 30 g; molasses, 10 g; soluble starch, 5 g; tap water	Dextrose, 10 g; starch, 20 g; yeast extract, 5 g; casein hydrolysate, 5 g; calcium carbonate, 4 g; tap water
Fermentation conditions	5 days at 30 °C, 250 rpm	60–70 hr at 31 °C, 300 rpm
Extraction procedure	Whole broth with ethyl acetate	Adjust whole broth to pH 9.5 and extract with ethyl acetate

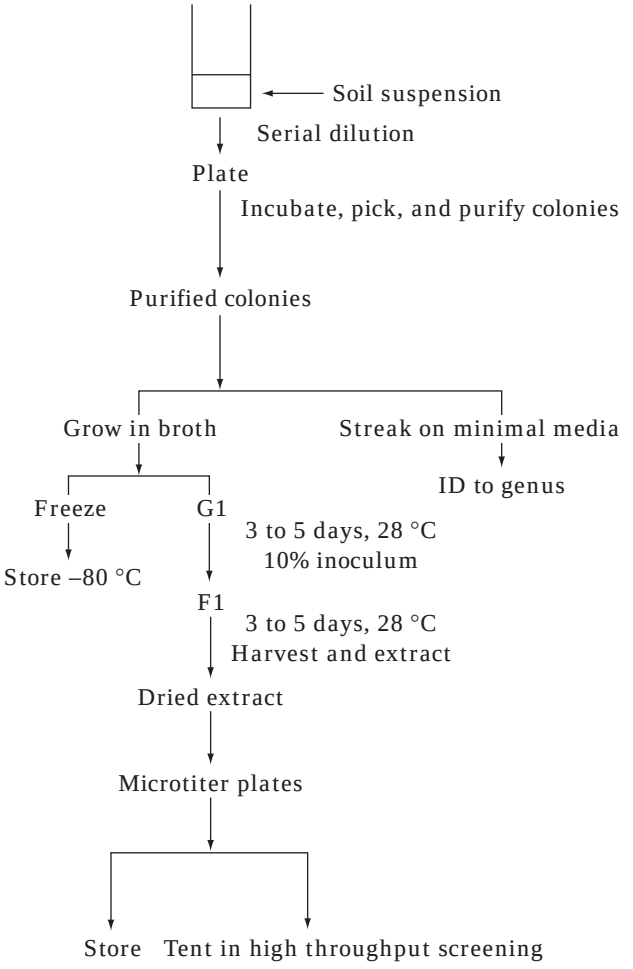


Figure 5. Outline of the primary screen process.

product. Examples of the production of secondary metabolites by two different genera of actinomycetes are presented in Table 10.

The advent of high throughput screening has resulted in a dramatic increase in the number of samples assayed, including natural products, with a concurrent decrease in the time required to test samples. In addition, an assay's lifetime has also decreased because all available samples can be tested in a relatively short time span (one month). For natural products to effectively compete with synthetic compounds and combinatorial libraries for available assay slots, a readily available source of samples is required. High throughput screening requires small amounts of test sample; a 10-mL fermentation results in enough extract for as many as 100 assays. Many natural product screening programs have addressed the issue of sample availability by developing libraries of extracts stored in 96-well microtiter plates. For each extract that is stored, the corresponding producing culture must also be preserved and its location catalogued. The nonstreptomycete actinomycetes described in this section can be stored at −80 °C in 10% glycerol for many years. However, data on the stability of natural product extracts stored over a period of years as well as the ability of preserved nonstreptomycete actinomycetes to maintain productivity have yet to appear in the literature. Figure 5 outlines the steps in a typical primary screen. Once activity has been detected in an extract or filtrate, the producing culture is fermented in a large enough volume to isolate enough material for characterization and structure elucidation. Many of the secondary metabolites produced by the *meso*-DAP containing actinomycetes are listed in Table 11.

SUMMARY

Nonstreptomycete actinomycetes continue to be a source of novel, therapeutically relevant natural products. Procedures described to effectively isolate a wide variety of genera rely on (*1*) collecting ecologically and geographically distinct soil samples, (*2*) using selective pressures (incorporating antibiotics and other chemicals into isolation media) that suppress streptomycetes while allowing diverse genera to survive, and (*3*) incorporating media constituents that favor the growth of rare actinomycetes. Evaluating the effectiveness of soil isolation procedures in obtaining diverse nonstreptomycete actinomycetes can quickly be achieved by identifying isolates macro- and microscopically. More definitive identification currently re-

Table 11. Secondary Metabolites Produced by *meso*-DAP-Containing Actinomycetes

Genus	Secondary metabolites[a]
Streptosporangium	Anthracyclines (figaroic acids, fragilomycins, carminomycin, carborubicin, sibiromycin; ansamycins (ansamitocin); bleomycins (platomycins, victomycin); glycopeptides (A-84575); macrolides (aculeximycin); peptides; phenazines; sugars (sporacuracins, 21906A); selenomycin; WS79089
Microbispora	Peptides (cochinmicins); phenazines; EV22 (SCH31828)
Microtetraspora	Amino acid containing (azaserine); talisomycin B; kistamicins; pyralomicins; macrolactam (fluvirucins); Bu-2313
Planobispora	Peptides; GE 2270
Planomonospora	Peptides (sporangiomycin); polyene
Thermomonospora	Naphthaquinone (SCH38519)
Actinomadura	Anthracyclines (akrobomycins, barminomycins, benanomicins, carminomycins, pradimicin, rubeomycins, rubstomycins); ansamycin (rifamycins); amino acid containing (madumycin); anthraquinones (oxanthromicin, maduramycin); indolocarabozole (AT2433); nucleosides (chloropentastatin, adenosine, coformycin, pentostatin); macrolactam (SCH38516, SCH38518, fluvirucin); macrocyclic lactone (kijanimicin); prodigiosins; polyethers (cationomycin, hidamicin, kijimycin); peptide related (formestin, luzopeptin, parvodicin), sugar containing (veractamycins, esperamicins); tetracyclin related (ES-119, SCH33256, SCH36969)
Spirillospora	Peptide macromolecular (Spirillomycin); polyene.
Micromonospora	Aminoglycosides (gentamicin, fortimicin, kanamycin, neomycin, sagamicin, sisomicin, verdamicin, G-52, G-418, Jl20); anthracyclines (daunorubicin, sibanomicin); anthraquinone (dynemicin); ansamycins (halomycins, rifamycins); benzodiazepine (sibanomicin); macrocyclic lactones (antlermicin, tetrocarcins); macrolides (clostomicin, erythromycin, juvenimicin, megalomicin, mycinamicin, izenamicin, rustmicin, primycin, rosaramicin); nucleosides (astromycin, dapiramicin); orthosomicins (everninomicins); peptides and amino acid containing (actinomycin, bottromycin, chalcidin, micromonosporin, 68-1147, glycine, negamicin); pigments (genistein, rubradirin, nocardiorubin); sugar containing (calicheamicin); xanthone (citreamicins); LL-E-19085
Actinoplanes	Glycopeptides (actaplanin, teicoplanin, ramoplanin); macrolides (lipiarmycin); peptides and amino acid containing (gardimycin, mycoplanecin, planothiocin, taitomycin, azetidine, azaserine, plauracins, virginiamycin, thristreptone); orthosomicin (SE737/74D); polyenes (octamycin, SE293); quinone (purpuromycin); chuangxinmycin; A/15104Y
Couchioplanes	Polyene (67-121)
Catenuloplanes	CP-54715
Dactylosporangium	Aminoglycosides (dactimicin, G-367); macrolides (clostomicin, tiacumin); orthosomicins (SF-2107); peptide containing (capreomycin, SF-2329); polyether (AC-7230, CP44161); quinone (DK-7814); tetracycline (SCH34164)
Pseudonocardia	Glycopeptides (azureomycins, boxazomycins, helvedardins); peptide (41043); polyene; quinones
Actinosynnema	Ansamycins (macbecins, C-33196); β-lactams (nocardicins)
Amycolatopsis	Ansamycins (rifamycins, kanglemycin); glycopeptides (avoparcin, vancomycin, orienticins, decaplanin, ristocetin, balhimycin); macrocyclic lactones (quartromycins); macrolides (A-59770); quinones (benzathrins, epoxyquinomicins, azicemicins); sugar (efrotomycin type); tetracycline (cetocycline); octacosamicins
Kibdelosporanglum	Glycopeptides (aridicins, kibdelins, decaplanin, A-80407); cycloviracins
Kutzneria	Chloramphenicol; glycoside macrolides (sporaviridins)
Saccharomonospora	Glycopeptide (AB-65)
Saccharopolyspora	Aminoglycosides (nebramycin, destomycin, saccharocin, sporaricin); alkaloids (hatomamicin); cinodines (coumamidines); lactones (belactins, nodusmicin); macrolides (erythromycin, sporeamicin); peptides (vanoxonin); CL307-24
Saccharothrix	Alkaloids (tetrazomine); glycopeptides (galacardins); indolocarbazoles (staurosporin, K-252a, rebeccamycins, SCH43228); macrolides (kayamycin, sekothrixide, N-814-103 factors); macrocyclic lactones (saccharocarcins); peptides (dopsisamine, madumycin, polynitroxin, nocamycin, sandramycin); polyether (portmicin); sugar (LL-C-19004); phosphonothrixin; karnamicins; tetrazomine
Streptoalloteichus	Aminoglycosides (kanamycins); bleomycins (talisomycins); siderochelin
Thermocrispum	Undescribed antifungal antibiotic (50)
Glycomyces	Amino acid containing (azaserine); LL-D05139
Actinocorallia	Peptide (azaserine related)

Note: Secondary metabolites not reported from *Herbidospora, Catellatospora, Pilimelia, Spirilliplanes, Actinopolyspora, Lentzea, Actinobispora*, and *Planotetraspora*.

[a]Data from Actinomycetes Antibiotic Data Base, Technomics Inc., C. P. O. Box 882, Tokyo, 00-91, Japan.

quires extensive chemical analysis of cell-wall and whole-cell components. Future identification of actinomycetes from natural habitats will rely on molecular approaches, particularly sequencing the gene encoding the small subunit of RNA (16S). This gene is a likely target because it is relatively small (1.5 kb), many actinomycete-type strains have been sequenced, and the data are readily available on-line for comparison to soil isolated, unknown strains. Numerous isolates can be fermented in small volumes. The fermentations are extracted or filtered, and the resulting samples are stored for testing against novel targets using high throughput screening methodologies. And so the quest for new metabolites with therapeutic potential from new organisms continues.

BIBLIOGRAPHY

1. M.P. Lechevalier and H.A. Lechevalier, in A.L. Demain and N.A. Solomon eds. *Biology of Industrial Microorganisms*, Benjamin/Cummins, Menlo Park, California, 1985, pp. 315–358.
2. E. Stackebrandt, F.A. Rainey, and N.L. Ward-Rainey, *Int. J. Syst. Bacteriol.* **47**, 479–491 (1997).
3. C. Jiang, L. Xu, Y. Yang, G. Guo, J. Ma, and Y. Liu, *Int. J. Syst. Bacteriol.* **41**, 526–528 (1991).
4. S. Iinuma, A. Yokota, T. Hasegawa, and T. Kanamaru, *Int. J. Syst. Bacteriol.* **44**, 230–234 (1994).
5. T. Hasegawa, *Actinomycetologica* **2**, 31–45 (1988).
6. H. Rumano, W. Guizhen, and L. Junying, *Int. J. Syst. Bacteriol.* **43**, 468–470 (1993).
7. Y. Wang, Z. Zhang, and J. Ruan, *Int. J. Syst. Bacteriol.* **46**, 933–938 (1996).
8. R.M. Kroppenstedt and M. Goodfellow, in A. Ballows, H.G. Trüper, M. Dworkin, W. Harder, K.H. Schleifer eds. *The Prokaryotes, Volume 2: A Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Application*, 2nd ed., Springer-Verlag, New York, 1992, pp. 1085–1114.
9. M. Goodfellow, in A. Ballows, H.G. Trüper, M. Dworkin, W. Harder, K.H. Schleifer eds. *The Prokaryotes, Volume 2: A Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Application*, 2nd ed., Springer-Verlag, New York, 1992, pp. 1115–1138.
10. T. Tamura, M. Hayakawa, and K. Hatano, *Int. J. Syst. Bacteriol.* **47**, 97–102 (1997).
11. M.P. Lechevalier, *J. Lab. Clin. Med.* **71**, 934–944 (1968).
12. M.P. Lechevalier and H.A. Lechevalier, *Int. J. Syst. Bacteriol.* **20**, 435–443 (1970).
13. H.A. Lechevalier and M.P. Lechevalier, in M.P. Starr, H. Stolp, H.G. Trüper, A. Balows, and H.G. Schlegel eds. *The Procaryotes: A Handbook on Habitats, Isolation and Identification of Bacteria*, Springer-Verlag, New York, 1981, pp. 1915–1922.
14. K.H. Schleifer and O. Kandler, *Bacteriol. Rev.* **36**, 407–477 (1972).
15. K. Suzuki, M. Goodfellow, and A.G. O'Donnell, in M. Goodfellow and A.G. O'Donnell eds. *Handbook of New Bacterial Systematics*, Academic Press, New York, 1993, pp. 195–238.
16. D.E. Minnikin and A.G. O'Donnell, in M. Goodfellow, M. Mordarski, and S.T. Williams eds. *The Biology of the Actinomycetes*, Academic Press, New York, 1984, pp. 337–388.
17. K. Uchida and K. Aida, *J. Gen. Appl. Microbiol.* **23**, 249–260 (1977).
18. K. Uchida and A. Seino, *Int. J. Syst. Bacteriol.* **47**, 182–190 (1997).
19. M.P. Lechevalier, A.C. Horan, and H.A. Lechevalier, *J. Bacteriol.* **105**, 313–318 (1971).
20. M.P. Lechevalier, C. DeBievre, and H.A. Lechevalier, *Biochem. Syst. Ecol.* **5**, 249–260 (1977).
21. M.D. Collins, in M. Goodfellow and D.F. Minnikin eds. *Chemical Methods in Bacterial Systematics*, Soc. for Applied Biotech., Tech. Series No. 20, Academic Press, London, 1985, pp. 267–287.
22. R.M. Kroppenstedt, in M. Goodfellow and D.F. Minnikin eds. *Chemical Methods in Bacterial Systematics*, Soc. for Applied Biotech., Tech. Series No. 20, Academic Press, London, 1985, pp. 173–199.
23. R.M. Kroppenstedt, in A. Ballows, H.G. Trüper, M. Dworkin, W. Harder, K.H. Schleifer eds. *The Prokaryotes, Volume 2: A Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Application*, 2nd ed., Springer-Verlag, New York, 1992, pp. 1139–1156.
24. E.B. Shirling and D. Gottlieb, *Int. J. Syst. Bacteriol.* **16**, 313–340 (1966).
25. A.C. Horan and B.C. Brodsky, *Int. J. Syst. Bacteriol.* **32**, 195–200 (1982).
26. E. Stackebrandt and B.M. Goebel, *Int. J. Syst. Bacteriol.* **44**, 846–849 (1994).
27. M. Goodfellow, L.J. Stanton, K.E. Simpson, and D.E. Minnikin, *J. Gen. Microbiol.* **136**, 19–34 (1990).
28. J. Meyer, in S.T. Williams, M.E. Sharpe, and J.G. Holt eds. *Bergey's Manual of Systematic Bacteriology, Volume 4*, Williams & Wilkins, Baltimore, 1989, pp. 2562–2568.
29. F.A. Rainey, N. Ward-Rainey, R.M. Kroppenstedt, and E. Stackebrandt, *Int. J. Syst. Bacteriol.* **46**, 1088–1092 (1996).
30. G. Vobis, in A. Ballows, H.G. Trüper, M. Dworkin, W. Harder, K.H. Schleifer eds. *The Prokaryotes, Volume 2: A Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Application*, Springer-Verlag, 2nd ed., New York, 1992, pp. 1029–1060.
31. C. Koch, R.M. Kroppenstedt, F.A. Rainey, and E. Stackebrandt, *Int. J. Syst. Bacteriol.* **46**, 765–768 (1996).
32. M.T. Embley, J. Smida, and E. Stackebrandt, *Syst. Appl. Microbiol.* **11**, 44–52 (1988).
33. M.T. Embley, in A. Ballows, H.G. Trüper, M. Dworkin, W. Harder, K.H. Schleifer eds. *The Prokaryotes, Volume 2: A Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Application*, 2nd ed., Springer-Verlag, New York, 1992, pp. 996–1027.
34. D.P. Labeda, R.T. Testa, M.P. Lechevalier, and H.A. Lechevalier, *Int. J. Syst. Bacteriol.* **35**, 417–421 (1985).
35. T. Cross, in M.P. Starr, H. Stolp, H.G. Trüper, A. Balows, and H.G. Schlegel eds. *The Procaryotes: A Handbook on Habitats, Isolation and Identification of Bacteria*, Springer-Verlag, New York, 1981, pp. 2091–2102.
36. M. Goodfellow and S.T. Williams, *Ann. Rev. Microbiol.* **37**, 189–216 (1983).
37. S.T. Williams and E.M.H. Wellington, in J.D. Bu'lock, L.J. Nisbet, and D.J. Winstanley eds. *Bioactive Microbial Products: Search and Discovery*, Academic Press, London, 1982, pp. 9–26.
38. M.C. Shearer, *Dev. Ind. Microbiol.* **28**, 91–97 (1987).
39. A.C. Horan, in V.P. Gullo ed. *The Discovery of Natural Products with Therapeutic Potential*. Butterworth-Heinemann, Boston, 1994, pp. 3–30.
40. T.S. Whitham, M. Athalye, D.E. Minnikin, and M. Goodfellow, *Antonie Van Leeuwenhoek* **64**, 387–429 (1993).
41. M. Hayakawa and H. Nonomura, *J. Ferment. Technol.* **65**, 501–509 (1987).
42. C.E. Bland and J.N. Couch, in M.P. Starr, H. Stolp, H.G. Trüper, A. Balows, and H.G. Schlegel eds. *The Procaryotes: A Handbook on Habitats, Isolation and Identification of Bacteria*, Springer-Verlag, New York, 1981, pp. 2004–2010.
43. N.S. Makkar and T. Cross, *J. Appl. Bacteriol.* **52**, 209–218 (1982).
44. F.P. Mertz, *Int. J. Syst. Bacteriol.* **44**, 274–281 (1994).
45. T. Kudo, T. Ioth, S. Miyadoh, T. Shomura, and A. Seino, *Int. J. Syst. Bacteriol.* **43**, 319–328 (1993).
46. A. Yokota, T. Tamura, T. Hasegawa, and L.H. Huang, *Int. J. Syst. Bacteriol.* **43**, 805–812 (1993).
47. A.F. Yassin, B. Haggene, H. Budzikiewicz, and K.P. Schaal, *Int. J. Syst. Bacteriol.* **43**, 414–420 (1993).
48. K. Tomita, Y. Hoshino, and T. Miyaki, *Int. J. Syst. Bacteriol.* **43**, 297–301 (1993).
49. K. Tomita, in S.T. Williams, M.E. Sharpe, and J.G. Holt eds. *Bergey's Manual of Systematic Bacteriology, Volume 4*, Williams & Wilkins, Baltimore, 1989, pp. 2569–2572.

50. F. Korn-Wendisch, F. Rainey, R.M. Kroppenstedt, A. Kempf, A. Majazza, H.J. Kutzner, and E. Stackebrandt, *Int. J. Syst. Bacteriol.* **45**, 67–77 (1995).
51. T. Tamura, M. Hayakawa, H. Nonomura, A. Yokota, and K. Hatano, *Int. J. Syst. Bacteriol.* **45**, 371–378 (1995).
52. A.F. Yassin, F.A. Rainey, H. Brzezinka, K.D.-Jahnke, H. Weissbrodt, H. Budzikiewica, E. Stackebrandt, and K.P. Schaal, *Int. J. Syst. Bacteriol.* **45**, 357–363 (1995).
53. R. Gherna, P. Pienta, and R. Cote eds. *ATCC Catalogue of Bacteria and Bacteriophages*, 18th ed., American Type Culture Collection, Rockville, Md., 1992.
54. A.C. Horan, M.C. Shearer, V. Hegde, M.L. Beyazova, B.C. Brodsky, A. King, R. Berrie, K. Cardaci, and M. Nimeck, *J. Antibiot.* **50**, 119–125 (1997).
55. M.J. Weinstein, G.H. Wagman, J.A. Marguez, R.T. Testa, E.M. Oden, and J.A. Waitz, *J. Antiobiot.* **22**, 253–258 (1969).

See also CEPHALOSPORINS; FERMENTATION MONITORING, DESIGN AND OPTIMIZATION; MEDIUM FORMULATION AND DESIGN, *E. COLI* AND *BACILLUS* SPP.; METABOLITES, PRIMARY AND SECONDARY; MUTAGENESIS.

SECONDARY METABOLITES, ANTIBIOTICS

WILLIAM R. STROHL
Merck Research Laboratories
Rahway, New Jersey

KEY WORDS

Antibacterial drugs
Antibiotic drug discovery
Antibiotic markets
Antibiotic resistance
Anticancer drugs
Antifungal drugs
Beta-lactams
Macrolides
Natural products
Polyketides

OUTLINE

INTRODUCTION

Definition of Natural Product Antibiotics

The field of natural products owes its origins to the discovery of penicillin and its ability to kill bacteria. Since Fleming's original discovery in 1928 of the antibacterial effects of *Penicillium chrysogenum* and Flory and Chaim's efforts more than a decade later to bring that discovery to production, the field of natural products has undergone enormous changes.

Vuillemin first introduced the concept of antibiotic activity in 1889, using the term *influences antibiotiques* (antibiotic influences) to describe negative interactions between plants and animals (1). Waksman later coined the term *antibiotic* to describe "a chemical substance derived from microorganisms which has the capacity of inhibiting growth, and even destroying, other microorganisms in dilute solutions" (2). Natural product antibiotics are almost universally secondary metabolites, compounds generally produced at low growth rates or after growth has ceased, nonessential for growth of the producing organisms in pure culture, and typically possessing unusual structures not found in primary metabolites of central metabolism. Although natural product antibiotic biosynthesis is not essential to the producing organism in pure culture, these compounds are critical to the producing organisms in their natural environment, both for survival and competitive advantage (3).

In more recent years, the term *antibiotics* sometimes has been used more broadly to define anti-infective agents, essentially in line with Waksman's definition, as well as *bioactive natural products*. This latter group includes immunomodulatory agents, statins, antitumor drugs, antiparasitic compounds, insecticides, and other bioactive natural products, many of which also possess weak antibacterial or antifungal activities. Indeed, the immunosuppressive drug cyclosporin was originally discovered in an antifungal screening program. Additionally, the biosynthesis of bioactive antibiotics is often very similar to that of other bioactive natural products. As a result, cyclosporin, a weak antifungal antibiotic but an excellent immunosuppressive agent, is usually included in discussions of other peptide antibiotics. Similarly, when reviews are written about macrolide antibiotics, other bioactive natural products such as tacrolimus (formerly called FK-506) and rapamycin, both of which are immunosuppressive agents (4), are typically included in discussions with erythromycin, a bacteriostatic antibiotic; avermectin, a potent antiparasitic compound; and tylosin, an agricultural feed additive.

This certainly makes sense because all these bioactive natural products possess similar modes of biosynthesis.

An extensive table including all marketed antibiotics (both natural product-derived and chemosynthetic) as well as all other marketed bioactive natural products, as of December 1996, was recently compiled for the book entitled *Biotechnology of Antibiotics* (5). Additionally, a recent review by Shu (6) highlighted most of the recent important discoveries in the field of natural products, ranging from new antifungal activities to natural products active as neuromodulators. Thus, this review will focus on natural product anti-infective and antitumor agents.

Overview of the Natural Product Antibiotics Field

Cragg et al. (7) published in 1997 a well-cited review of the natural products field in which they analyzed the percentage of natural products in various fields. Of 520 new approved drugs between 1983 and 1994, 55.6% were chemically synthesized drugs, 24.4% were semisynthetic compounds, 5.8% were unmodified natural products, 8.8% were chemically synthesized drugs patterned after natural products, and 5.4% were "biologics" (e.g., recombinant proteins) (7). Thus, approximately 40% of all newly approved drugs over that period were derived from natural product leads. In breaking down the categories, however, natural products dominated the antibacterial markets (63% were derived from natural product leads) and the anticancer market, in which 61% of the new drugs were derived from natural product leads. Natural product leads also strongly contributed to the antihyperintensive market (48% were derived from natural product leads) and the anti-inflammatory market, in which one-third of the new drugs approved between 1983 and 1994 were derived from natural product leads. According to Cragg et al., however, natural products did not contribute in the period of 1983–1994 to newly approved drugs in the antifungal, analgesic, anxiolytic, antihistamine, antidepressant, cardiotonic, or hypnotic fields (7).

Strohl (5) conducted a similar analysis of bioactive natural products with antibiotic activities, identifying a list of 252 natural and synthetic compounds used clinically and agriculturally in the United States with selected compounds included from other countries. Thus, although this analysis was not all-inclusive, it represented the majority of commercially important antibiotics and antimetabolic drugs. Of those compounds identified, 75 (30%) were natural products, another 75 (30%) were semisynthetic compounds derived from natural products, and the remaining 40% were chemically synthesized compounds (5).

The discovery of new, novel natural products continues to rise dramatically (5). Based on an analysis of papers published in *Journal of Antibiotics*, Bérdy (8) indicated that about 11,900 antibiotics and bioactive natural products had been discovered through 1994, with increasing rates of discovery every year through 1990. He projected that the total of new, novel natural product structures discovered by the year 2000 would top 16,000 (8). The annual rate of new natural product descriptions published is now more than 500 per year, up from 200 to 300 per year 20 years ago (5,8). Nevertheless, there is the widespread perception in both the scientific literature and the popular press that the discovery of new natural products and natural product antibiotics is on the wane (5). This anomaly, however, can easily be explained. Because most new natural product drugs still are concentrated in the fields of anti-infectives and anticancer drugs, and natural product leads have been around long enough in both the anti-infectives and anticancer areas, the law of diminishing returns is most prevalent in these fields. For example, from 1962, when nalidixic acid was developed as the first quinolone antibiotic, to 1996, no new classes of clinical natural-product-derived antibacterials have been introduced to the market. In 1993, only one new natural product antibiotic was approved by the Food and Drug Administration for use in the United States, none in 1994, and only a few in 1995 and 1996. Overall, from 1980 to 1990, the worldwide discovery of more than 4,600 new bioactive natural products was published, yet only 20 new natural-product-derived drugs (3 natural products and 17 semisynthetic products) were commercialized in that period. Finally, in the period spanning from World War II to 1975, approximately 1.6% of all natural product antibiotics reported (either by publication or patent) were eventually commercialized. Since 1975, this number has dropped to about 0.11%. Thus, fewer antibiotics are being commercialized from an increasingly larger pool of isolated candidate compounds. These data suggest that the Golden Era of antibiotics, generally recognized as having begun shortly after World War II ended, lasted nearly 30 years before the law of diminishing returns caught up to it. Another major contributing factor for the dearth of commercialized natural product antibiotics from 1975 to the present is that the U.S. government, academic researchers, and pharmaceutical companies alike turned their efforts toward the defeat of cancer and heart disease, believing that infectious diseases were under control (5,9).

On the other hand, in pharmacological areas in which natural products are making new inroads, many discoveries of bioactive natural products have resulted in compounds that are still in preclinical or clinical trials. Since the late 1960s, Umezawa and colleagues in Japan showed interest in discovering non-anti-infective bioactive natural products, particularly in the area of protease inhibitors (10). Although researchers in U.S. pharmaceutical companies showed interest in developing natural products as bioactive molecules in areas other than the anti-infective and anticancer markets (i.e., whole cell assays), many viewed the Japanese mechanism-based assay approach more as a curiosity than a serious discovery effort. It was not until the publication of Merck's patent and paper in *Proceedings of the National Academy of Science* (11) in 1980 on lovastatin as a remarkable inhibitor of hydroxymethylglutaryl-CoA reductase, and ultimately an outstanding cholesterol-lowering drug, that the inclusion of natural products in mechanism-based screens was considered seriously.

Since that time much has changed, and although some pharmaceutical companies have minimized natural product screening efforts because of the perceived cost-to-benefit ratio, others have sought to expand the scope of natural products in their screening programs. As evi-

denced by publications in journals such as *Journal of Antibiotics*, patent applications, and clinical trial candidates (6), it has now become apparent that natural products can significantly augment synthetic chemical approaches for the discovery of new drug entities in a wide range of pharmacologically important areas. The two most significant recognized drawbacks of natural products in a broad-based screening program are the inclusion of crude extracts (which usually include salts, proteins, fatty acids and other potentially interfering materials) in in vitro assays (e.g., receptor-based, enzyme inhibition, or protein–protein interaction assays) and the time it takes to turn a crude-extract hit into a chemical structure. Despite these drawbacks, natural products now have been found to contribute to significant leads in most therapeutic areas (6).

The sources of natural products are also likely to change. Bérdy (8) showed that of the 11,900 antibiotics that had been discovered through 1994, approximately 6,600 (55%) were produced by *Streptomyces* spp. The remainder of producing organisms included the filamentous fungi, which produced 2,600 (22%); nonactinomycete bacteria, which produced 1,400 (12%); and non-*Streptomyces* strains of actinomycetes, which produced 1,300 (11%) (8). In the broader pharmacological field, it is widely recognized that a great percentage of pharmacologically derived products in use today are derived from or patterned after plant products (6). It can be expected that in the future a wider range of microorganisms, including well-known but underused groups such as the filamentous fungi, myxobacteria, bacilli, and pseudomonads, will contribute to the pool of natural products in larger ways (5).

Market for Natural Product Antibiotics

The total world sales for prescription pharmaceuticals was estimated in a September 1998 news release to be $308 billion. In a recent report of the 200 most-prescribed medicines in the United States (12), natural products accounted for 10 of the top 35 prescribed drugs (Table 1). All the top-prescribed natural products or semisynthetic drugs derived from natural products are anti-infective agents except for three statins: simvastatin, lovastatin and pravastatin. Although it is difficult to identify the entire contribution of natural products to the world pharmaceutical market economy, it should easily total more than $30 billion, including the natural-product-derived anti-infective agents (>$20 billion), the natural-product-derived anticancer agents (in which the anthracyclines and taxanes alone represent more than $3 billion in market size), the statins (with a market size estimated at more than $6 billion and rising at a rate of more than 20% annually), and the large pool of plant-derived natural products or compounds patterned after plant products.

The total world market for anti-infective agents in 1996 was reportedly $23 billion (13). The 1995 U.S. anti-infectives market alone was greater than $8 billion, with cephalosporins (45%), penicillins (15%), fluoroquinolones (11%), tetracyclines (6%), and macrolides (5%) comprising the majority of sales. Of these top-selling anti-infective agents, only the fluoroquinolones lack the distinction of being natural products, and even they are patterned after the original quinolone, nalidixic acid. Thus, in deference to their origins, antibacterial drugs are still natural products, semisynthetic compounds derived from fermentation products, or patterned after natural products.

The world antifungal market in 1995 was approximately $3 billion. Although this represents a much smaller market than the antibacterial market, the antifungal market is projected to grow at 20% annually. This rapid growth in antifungals unfortunately is driven by the dramatic increase in life-threatening fungal diseases, mostly as a result of opportunistic fungal pathogens infecting immunocompromised patients.

The antiviral market was $1.8 billion in 1995, and the projected 1998 market is $3 billion. Although the top product in 1995 was acyclovir (50% of the market [14]), more than 17 million people were infected with HIV, a number that is expected to reach 30 to 40 million by year 2000 (14). Because of the development of successful new anti-HIV drugs such as the HIV-proteinase inhibitors, it is expected that the antiviral market should increase substantially within the next few years.

Biogenesis of Natural Product Antibiotics

Since the discovery and elucidation of the structure of penicillin, there has been an enormous effort placed into finding new natural product antibiotics, particularly those of the antibacterial, antifungal, and anticancer therapeutic classes. Natural product antibiotics, for the most part, fall into several major classes of structures, as shown in Table 2, based on the precursors from which they are assembled and the mechanisms of assembly. Precursors common for the synthesis of natural products include short chain acyl-CoAs such as acetyl-CoA, malonyl-CoA, and propionyl-CoA, which are used to synthesize polyketides; activated amino acids, which are used to build peptides and β-lactams; various pyrimidines, which are the building blocks for nucleoside natural products; isoprene units, which form the basis of terpene, alkaloid, and substituted steroid natural products; and sugars and amino sugars, which are the building blocks for aminoglycosides as well as for the sugar moieties that decorate a wide variety of natural products. Many natural products result from the fusion of moieties arising from two or more different biogenic classes. For example, 6-deoxyerythronolide B (DEB), the macrolide backbone of erythromycin, is synthesized by a type I modular polyketide synthase, but the molecule is not biologically active until the sugars desosamine, attached at C-5, and cladinose, attached at C-3, are added. Several other macrolides, including spiromycin and tylosin, have this same characteristic. Similarly, anthracyclines are comprised of the aglycone portions, derived from the activity of a type II iterative polyketide synthase system, and daunosamine, a 2,3,6-trideoxy-3-aminohexose synthesized from TDP-glucose (15). Glycopeptides such as vancomycin also contain core units derived from the activity of a nonribosomal peptide synthetase that are decorated by sugar units to make bioactive molecules. An exaggerated example of this is the antifungal compound pradimicin, which is synthesized by a type II iterative polyketide synthase and then decorated both with sugar moi-

Table 1. Ranking of Natural Product Pharmaceuticals as Prescription Medicines in the United States (1997)

Natural product[a]	Brand name	Company	Rank[b]
Amoxicillin	Trimox	Apothecon	1
Azithromycin	Zithromax	Pfizer	9
Amoxicillin-clavulanate	Augmentin	SmithKline Beecham	14
Simvastatin	Zocor	Merck	15
Clarithromycin	Biaxin	Abbott	24
Amoxicillin	Amoxil	SmithKline Beecham	25
Cephalexin	Cephalexin	Teva	26
Penicillin V (K^+)	Veetids	Apothecon	32
Amoxicillin	Amoxicillin	Teva	33
Pravastatin	Pravachol	Bristol-Meyers Squibb	34
Cephalexin	Cephalexin	Apothecon	54
Lovastatin	Mevacor	Merck	57
Cefprozil	Cefzil	Bristol-Meyers Squibb	58
Cefuroxime	Ceftin	Glaxo Wellcome	72
Erythromycin	Ery-Tab	Abbott	82
Neomycin/polymyxin B/HC	Neomycin/polymyxin B/HC	Schein	122
Mupirocin	Bactroban	SmithKline Beecham	130
Cefaclor	Cefaclor	Mylan	139
Doxycycline	Doxycycline	Zenith	168
Amoxicillin	Amoxicillin	Warner-Chilcott	184
Clarithromycin	Biaxin suspension	Abbott	192
Cefixime	Suprax	Wyeth-Ayerst	196

[a]Includes both natural products and semisynthetic compounds derived from natural products.
[b]Ranking based on number of prescriptions sold.
Source: Data from Ref. 12.

eties and D-amino acid moieties (16). Other molecules, such as the rifamycins, contain a unique building block such as 3-amino-5-hydroxybenzoic acid (AHBA), which is used as the primer with which to initiate the synthesis of the macrocycle (17). Thus, many natural products are derived from mixed biogenic origins. For more detailed information on specific antibiotics, recent books edited by Vining and Stuttard (18) and by Strohl (19) cover the biogenesis and biosynthesis of virtually all the major classes of antibiotics and bioactive natural products.

ANTIBACTERIAL AGENTS

Introduction

The most widely prescribed, best known, best characterized, and perhaps also the most widely disparaged natural products are the antibacterials. After the discovery of penicillin by Fleming in 1928 and its subsequent development as a drug during World War II, the pharmaceutical industry and well-known academics such as Waksman of Rutgers University worked feverishly to discover new antibacterials that would eliminate bacterial infections of all kinds. The natural product antibiotics gramicidin (1939), streptomycin (1944), chlortetracycline (1948), chloramphenicol (1948), cephalosporin C (1948), neomycin (1949), oxytetracycline (1950), erythromycin (1952), tetracycline (1953), vancomycin (1955), rifamycins (1959), lincomycin (1962), pristinamycin (1962), and nalidixic acid (1962), the predecessor to fluoroquinolones, were all discovered by 1962 (22). No new, truly novel classes of natural product antibacterial drugs have been marketed since then. Even in the cases of modern synthetic antibacterials, all successful synthetic antibacterials have been biomimetically derived from natural product structures (the best example being the fluoroquinolones, which are based on nalidixic acid). There are new classes of candidate antibacterials, such as the oxazolidinones (23), which are not derived from natural products, but these have yet to break into the market, in part because of associated toxicities exhibited by this class of compounds. The targets for these drugs, including cell wall biosynthesis, protein biosynthesis, and nucleic acid biosynthesis, all have been targeted many times over by analogues of them, as well as by other compounds, both synthetic and naturally derived. As we enter the 21st century, there is a critical need for new, potent antibacterials to which resistance is not easily gained.

Needs and Markets

With the emergence of new infectious diseases as well as the emergence of bacterial strains resistant to existing antibiotics, there is an enormous challenge to pharmaceutical companies, researchers, and governments to develop new methods for treating existing infectious diseases and those emerging as new health threats. As recently reviewed by Strohl (5), there are several reasons for the need to develop new antibiotics, preferably with novel structures and activities.

First and foremost, pathogenic bacteria are acquiring or developing resistance to existing antibiotics and classes of antibiotics in direct correlation with the use of those antibiotics to treat them (24). As Julian Davies stated, "it is frightening to realize that one single base change in a gene encoding a bacterial β-lactamase can render useless $100 million worth of pharmaceutical research effort" (25). Although this is an oversimplification of the problem, it

Table 2. Biogenic Classes of Natural Product Antibiotics

Biogenic class and subclass[a]	Example compounds or groups	Example activities
	Nonribosomal peptide synthetases	
Diketopiperazines	Gliotoxin	Antibiotic
Cyclic peptides	Cyclosporin	Immunosuppressive
	Gramicidin	Antibacterial
Depsipeptide	Enniatin, valinomycin	Ionophores
Lipopeptides	Echinocandins	Antifungal
Glycopeptides	Vancomycin	Antibacterial
	Bleomycins	Antitumor
β-Lactams	Penicillins, cephalosporins	Antibacterial
Lantibiotic	Nisin	Antibacterial
Peptidolactones	Actinomycin D	Antitumor
Phosphinopeptide	Bialaphos	Phytotoxic
	Polyketides	
Type I modular		
Macrolides	Erythromycin, tylosin	Antibacterial
Pentacyclic lactones	Avermectin, milbemycin	Insecticidal acaricidal
Ascomycins	Tacrolimus	Immunosuppressive
Ansamycins	Rifamycin, ansatrienins	Tuberculostatic, antibacterial
Polyenes	Nystatin, amphotericin B	Antifungal
Polyethers	Monensin	Growth promotants
Type I iterative		
Statins	Lovastatin, pravastatin	Cholesterol-lowering
Type II iterative		
Tetracyclines	Oxytetracycline	Antibacterial
Anthracyclines	Doxorubicin, aclarubicin	Cytotoxic antitumor
	Aminoglycosides	
Aminoglycosides	Streptomycin, kanamycin	Antibacterial
Aminocyclitols	Spectinomycin	Antibacterial
Cyclitol	Kasugamycin	Antibacterial
	Isoprenoids	
Steroidal	Squalamine	Antitumor, antibacterial
Sesquiterpene	Artemisinin (Qinghao), epolone	Antimalarial
Diterpene	Paclitaxel (Taxol)	Antitumor
Triterpene	Fusidic acid	Antibacterial
Indole alkaloid	Vinblastine	Antitumor
	Others	
Lincosamides	Lincomycin, clindamycin	Antibacterial
Chorismic acid-derived	Chloramphenicol	Antibacterial
Amino acids	Cycloserine	Antibacterial
Quinolones	Nalidixic acid, fluoroquinolones	Antibacterial
Nucleosides	Nikkomycins, polyoxins	Antifungal
Mitosanes	Mitomycin C	Antitumor, cytotoxic
Oligosaccharides	Everninomycin, avilamycin	Antibacterials

Source: Data from Refs. 5, 18–21.

succinctly sums up the problem with the development of antibiotic resistance by pathogenic bacteria. The most critical reason for the development of new antibiotics is the race to remain a step ahead of antibiotic-resistance development, which has proceeded at an alarming rate (26). Even since the earliest use of antibacterial agents, the emergence of resistance to the drugs used has been observed. For example, within 4 years after the introduction of penicillin during World War II, several strains of penicillin-resistant bacteria were isolated from infected patients (27). After the war, outbreaks of dysentery in Japan, caused by sulfonamide-resistant *Shigella* spp., presented a serious health hazard to the recovering population. Within just a few years after introduction in 1950 of the newly developed antibiotics, streptomycin, tetracycline, and chloramphenicol to help quell the problem, resistance to each of these antibiotics was observed (28). By 1969, 69% of shigellae were resistant to sulfonamides, streptomycin, tetracycline, and chloramphenicol (28). Thus, in fewer than 20 years, nearly 70% of all *Shigella* spp. in Ja-

pan became resistant to four antibiotics to which there was minimal resistance before their use (28). Widespread antibiotic use has resulted in the rapid spread of multidrug resistant pathogens, and it has become both a local and global health hazard of epidemic proportions. The incidence of methicillin-resistant *Staphylococcus aureus* (MRSA), vancomycin-resistant enterococci (VRE; particularly by *Enterococcus faecium* and *Enterococcus faecalis*), and β-lactam-resistant *Streptococcus pneumoniae* are among the most serious problems, but are by no means the only ones (29). The incidence of nosocomial infections (i.e., hospital-acquired) is on the rise, and of those, an increasingly greater percentage of the infectious agents (typically MRSA and VRE) are drug-resistant (30). The U.S. Centers for Disease Control and Prevention (CDC) estimates that of the 40 million hospitalizations in the United States each year, 2 million patients (5%) contract nosocomial infections, and that as many as 58,000 of them (0.15%) die of them (31). Unfortunately, VRE are nearly untreatable with current antibiotics, with the possible exceptions of the experimental drugs Synercid or teichoplanin, which have been used successfully to treat some cases of VRE (32). The mortality rate of patients infected with multiple-drug resistant enterococci is 70%, making this a serious health care problem (32).

Second, new pathogens are evolving at an alarming rate. For example, today there are more than 30 new infectious diseases (e.g., AIDS, Ebola, Legionnaires disease, hantavirus, Lyme disease, and food-borne *E. coli* O157:H7 infections) that were unknown 20 years ago (5,33). Third, even with the development of remarkable anti-infective drugs with outstanding biological activities, there are still certain pathogens that naturally defy even the best efforts to defeat them. Perhaps the best example of this is our current inability to treat successfully *Pseudomonas aeruginosa* infections resulting from cystic fibrosis (5).

Finally, many potentially important antibiotics have associated toxicities that limit their use. The best example of this is gentamycin and other aminoglycosides, which are limited in effectiveness because of their associated nephrotoxicity and ototoxicity (5). Thus, critical challenges lay ahead for the entire anti-infectives field, many of which may be met by the discovery and development of new natural product leads.

There are three critical needs in the antibacterial market. First, new parenteral drugs that can successfully treat MRSAs and VREs are absolutely required. This need will grow dramatically in the coming years as the once-rare incidence of vancomycin-resistant *S. aureus* strains, the single greatest worry of infectious disease experts (34), expands. Second, even with a seemingly crowded pediatric market, there is still a critical need (and market place) for an orally absorbed, pediatric-safe (which eliminates fluoroquinolones, which are not considered safe for children) antibacterial agent that can successfully treat upper respiratory tract infections (primarily otitis media caused by β-lactam resistant streptococci, moraxellae, and *Haemophilus influenzae*) and that is not compromised by β-lactam, aminoglycoside, or MLS (macrolide-lincosamine-streptogramin) resistance or degraded by extended spectrum β-lactamases (35).

The pediatric antibacterial market, which has grown sharply since the early 1980s, has become one of the most important anti-infective markets today. Pediatric antibacterials typically are prescribed for middle-ear infections (otitis media) caused by *H. influenzae*, *Moraxella* spp., and *Streptococcus pneumoniae*. By the year 2000, it is expected that the U.S. pediatric market alone will be $2.2 billion, with annual increases of approximately 15%. The current leading pediatric antibacterials are amoxicillin, Augmentin (amoxicillin plus clavulanic acid), azithromycin, and clarithromycin. It is expected that the ketolides may make a significant impact on this market. The final critical need is for a new tuberculocidal drug. Antibiotic resistance in *Mycobacterium tuberculosis* is considered to be a major growing problem. Currently, a combination treatment with isoniazid and rifamycin (both tuberculostats) is the method of choice, but many strains of *M. tuberculosis* have recently been found that are resistant to one or both of these antibiotics, and approximately 14% of all tuberculosis cases reported today involve *M. tuberculosis* strains that are resistant to one or more antibiotics (36). The problem facing doctors today is that there is no proved, effective alternative yet on the market to treat isoniazid- and rifamycin-resistant *M. tuberculosis* (36).

Antibiotic-Resistant Bacteria

Antibiotic resistance is a major clinical problem that will only get worse with time. In the case of bacteria, resistance can be mediated by several different types of mechanisms, including modification of the target to make it unsusceptable to the antibiotic, enzymatic modification of the antibiotic to render it inactive, exclusion of the antibiotic by modification of membrane permeability, active export of the antibiotic, and bypass of the critical physiological mechanisms or structures inhibited by the antibiotic. Resistance to some these mechanisms, encoded by chromosomal mechanisms, is attained at a rate of 1×10^{-8}, and others are obtained at much higher frequencies through horizontal plasmid transfer.

Plasmid-encoded TEM- and SHV-type β-lactamases are the most common mechanisms responsible for β-lactam resistance in Gram-negative bacteria (37). These β-lactamases have generally broad substrate specificities for penicillins as well as some first- and second-generation cephalosporins and are responsible for much of the resistance to β-lactams observed in the clinics. It is well documented that as new β-lactams have been developed to resist plasmid-borne β-lactamases, the β-lactamases themselves have evolved (25,37). In an extraordinary example of natural directed evolution, TEM- and SHV-type β-lactamases recently have been found containing specific point mutations, all falling in close proximity to the active site cavity, which confer resistance to third-generation cephalosporins and monobactams (37).

Vancomycin has long been the last line of defense against antibiotic-resistant bacteria, particularly MRSAs and β-lactam- and aminoglycoside-resistant enterococci. Just a little over a decade ago, it was thought that microorganisms would never become vancomycin resistant because of its unusual mechanism of action (38). This illusion

was shattered, however, when VREs were reported in England and France in 1987 (39). VREs now are being discovered at an alarming rate, with no new last line of defense substitute firmly entrenched on the market (40).

Not all clinically significant antibiotic resistance is plasmid- or transposon-mediated. *S. pneumoniae*, the primary causative agent in pediatric otitis media and pneumonia as well as bacterial meningitis and bacteremia, remained sensitive to penicillin for many years. Recently, however, an alarming number of penicillin-resistant *S. pneumoniae* strains have recently been isolated from infected patients, correlating with the use of penicillins to treat them (41). These penicillin-resistant *S. pneumoniae* were not found to produce *β*-lactamases, the common plasmid-borne mechanism for resistance. Instead, they had incurred chromosomal mutations giving rise to altered penicillin-binding proteins (PBPs), the target for penicillin activity (41). This discovery has a touch of irony, because *S. pneumoniae* is the bacterium with which plasmid-mediated transformation was first observed. Similarly, the Oxford strain of *S. aureus*, isolated 50 years ago, is sensitive to nearly every antibiotic tested against it (42). The staphylococci, however, acquired resistance to aminoglycosides very quickly, and by 1992, 32% of *S. aureus* strains tested were found to be resistant to methicillin, the drug of choice to treat staphylococcal infections, up from 2% in 1975 (42). Like pneumococci, MRSAs have altered PBPs, giving them a resistance mechanism that cannot be countered easily. Thus, vancomycin has become the drug of last resort to treat MRSAs. Ever since *Enterococcus faecalis* was shown to transfer vancomycin resistance to *S. aureus* in the laboratory in 1992 (43), one of the greatest fears of health professionals universally has been the emergence of vancomycin-resistant, methicillin-resistant staphylococci (44). Recently, such doubly resistant, nearly untreatable strains have appeared in sporadic cases, confirming the greatest fears of infectious disease professionals (45).

A new mechanism of antibiotic resistance recently was discovered in meropenem-resistant *Pseudomonas aeruginosa* strains (46). *P. aeruginosa* is an opportunistic pathogen that contributes to nosocomial infections, eye infections, and opportunistic infections of burn patients and those with cystic fibrosis. The mechanism was found to be based on the alteration of outer membrane proteins, causing a reduction in permeability of several antibiotics, including meropenem, through the outer membrane (46). This is a significant finding, because carbapenems (e.g., imipenem) have been considered the antibiotics of choice for the treatment of opportunistic pseudomonads.

β-Lactams

As is well known, the discovery of penicillin was the genesis of modern pharmaceutical efforts to produce antibiotics, the magic bullets that at one time were thought to cure the world of disease. *β*-Lactams function by binding to PBPs, especially transpeptidases, that normally function in peptidoglycan biosynthesis. Because peptidoglycan is novel to bacteria, these drugs specifically act on a bacterial target. Since bacteria cannot grow in normal environments without their cell wall, inhibition of cell wall biosynthesis results in killing of the bacteria; thus, all *β*-lactams by definition are bacteriocidal drugs.

The total antibacterial market tops $20 billion annually, with nearly all the top-selling antibacterial drugs being natural products, semisynthetics derived from natural products, or compounds patterned after natural products. Even more than a half century after the first widespread use of penicillin during World War II to treat bacterial infections, *β*-lactams as a whole still make up 60%, or more than $12 billion worth, of the total antibacterial market. The commercial *β*-lactam antibiotics are comprised of many types of several different classes, ranging from penicillin G to the monobactam biomimetic drug, aztreonam (Table 3).

The natural product *β*-lactams can be separated into five separate biogenic subclasses (47): penams (e.g., penicillins), cephems (cephamycin C, cephalosporins), clavams (clavulanic acid, clavamycins), carbapenems (thienamycin), and monobactams (aztreonam) (Table 4). The ring structures of the various *β*-lactams differ from the classical penicillin core as follows: cephems have an expanded 6-membered ring containing a double-bond but retaining the sulfur atom; clavams have an oxygen atom replacing the sulfur of penicillins, and they tend to differ at C-2 and C-6; in carbapenems, the sulfur is replaced by a carbon and a double bond is present; and the monobactams are monocyclic (48).

β-Lactams of the penicillin, cephamycin, and cephalosporin subclasses share a common biogenic origin, i.e., the condensation of three activated amino acids, L-*α*-aminoadipate, L-cysteine, and L-valine by aminoadipyl-cysteinyl-valine (ACV) synthase (*pcbAB* gene products) to form the tripeptide *δ*-L-(*α*-aminoadipyl)-L-cysteinyl-D-valine (ACV), which is then cyclized by the enzyme isopencillin N (IPN) synthase (*pcbC* gene product) to form isopenicillin N (49). ACV synthases comprise a subset of a larger class of nonribosomal peptide synthases, which have been characterized in detail, both biochemically and molecularly (49,50). Both ACV synthases (49,50) and IPN synthases (49) of all *β*-lactam producers are highly conserved, indicating common ancestral origins. IPN is either directly converted to penicillin G via IPN acyltransferase (*penDE* gene product), or via IPN epimerase (*cefD* gene product) to penicillin N, which is converted via deacetoxycephalosporin C (DAOC) synthase (*cefE* gene product; also called expandase because it expands the 5-membered ring of penicillins to the 6-membered ring of cephalosporins) to deacetoxycephalosporin C (DAOC) (49).

The clavam *β*-lactamase inhibitor, clavulanic acid, is produced by *Streptomyces clavuligerus*, which also produces cephamycins (49). Interestingly, the biogenesis of the bicyclic nucleus of clavulanic acid occurs via a completely different mechanism than does the biosynthesis of other *β*-lactams (49,51). In *S. clavuligerus*, the clavulanic acid biosynthesis genes are just downstream of the cephamycin C biosynthesis genes (51).

Not all *β*-lactam-producing taxonomic groups of organisms synthesize all subclasses of *β*-lactams, as depicted in Table 4. Filamentous fungi, for example, produce *β*-lactams only of the penam and cephem classes, whereas various actinomycetes produce *β*-lactams of all subclasses

Table 3. Important Clinical and Experimental β-Lactams

Subclass	Drugs on market[a]
Penicillins	Ampicillin, *amoxicillin*, bacampicillin, cloxacillin, floxacillin, mezlocillin, nafcillin, oxacillin, penicillin G, penicillin V
Penicillinase-resistant penicillins	*Methicillin*, dicloxacillin
Antipseudomonal penicillins	Carbenicillin indanyl, piperacillin, *ticarcillin*
First-generation cephalosporins	Cefadroxil, cefazolin, *cephalexin*, cephalothin, cephapirin, cephradine
Second-generation cephalosporins	*Cefaclor*, cefmetazole, cefonicid, cefotetan, *cefuroxime*
Third-generation cephalosporins	Cefamandole, *cefixime*, cefoperazone, cefotaxime, cefpodoxime proxetil, *cefprozil*, ceftazidime, *ceftibuten*, ceftizoxime, ceftriaxone
Oxycephams (third generation)	Flomoxef, latamoxef
Fourth-generation cephalosporins	*Cefepime*
Cefam	*Cefoxitin*
Carbacefem	*Loracarbef*
Carbapenems	*Imipemen/cilistatin*, meropenem, MK-826, panipenem (Japan only)
Monobactams	Aztreonam
Clavams (β-lactamase inhibitors)	*Clavulanate*, sulbactam, tazobactam
Penicillins/β-lactamase inhibitors	*amoxicillin/clavulanate (Augmentin*), ampicillin/sulbactam (Unasyn), pipericillin/tazobactam (Zosyn), ticarcillin/clavulante (Timentin)

[a]Compounds in italics are most important commercial antibiotics.

Table 4. Biogenic Classes of β-Lactams

		β-Lactams produced by[a]		
Subclass	Example compounds	Fungi	G^+ bacteria	G^- bacteria
Penams	Penicillin G	*Penicillium, Aspergillus*	—	—
Cephems	Cephalosporin C	*Cephalosporium*	—	—
	DAOC[b]	—	—	Various strains
	Cephabacins	—	—	*Lysobacter*
	Cephamycin C	—	*Streptomyces, Nocardia*	—
Carbapenems	Thienamycin	—	*Streptomyces cattleya*	*Serratia, Erwinia*
Monobactams	Aztreonam[c]	—	*Nocardia*	*Pseudomonas*
Clavams	Clavulanic acid	—	*Streptomyces*	—
	Clavamycins	—	*Streptomyces*	—

Note: See Cohen and Aharonowitz (47) for more information. Dash means not produced by this group.
[a]Example taxonomic groups are given by genus names rather than an exhaustive listing of producing organisms.
[b]DAOC = deacetoxycephalosporin C.
[c]Example given is the biomimetic synthetic compound of the class.

except the penams and cephalosporins of the cephem subclass. The cephems in general are produced by the taxonomically widest range of microorganisms, including filamentous fungi, Gram-negative bacteria (e.g., lysobacters), and actinomycetes. The cephem products accumulated by each group differs, however, likely indicating evolutionary branches in terms of enzymes present. The Gram-negative bacteria, *Flavobacterium* sp. SC 12.154 and *Lysobacter lactamgenes* accumulate a 7-formylamino analogue of cephalosporin called cephabacin (49,52), an unusual natural β-lactam structure that is highly resistant to β-lactamases (49). The actinomycetes *Streptomyces clavuligerus* and *Nocardia lactamdurans* produce cephamycin C, which is formed from DAOC by the enzymes (gene names in parentheses) DAOC hydroxylase (*cefF*), deacetylcephalosporin C (DAC) *O*-carbamoyltransferase (*cmcH*), cephalosporin 7-α-hydroxylase (*cmcI*), and 7-α-hydroxy-*O*-carbamoyldeacetoxycephalosporin C methyltransferase (*cmcJ*) (49). The filamentous fungus, *Cephalosporium acremonium*, on the other hand, converts DAOC to cephalosporin C in two steps: DAOC hydroxylase (*cefF*), and DAC acetyltransferase (*cefG*) (49).

As is well known, penicillin was the first widely used clinical antibiotic. The first oral penicillin, used primarily to treat Gram-positive bacteria, was the fermentation product phenoxymethylpenicillin (penicillin V), still sold today under the name Veetids (5). Second-generation penicillins include the aminopenicillin group of ampicillin, amoxicillin, and bacampicillin, which exhibited much better gastrointestinal tract absorption and possess longer half lives. These aminopenicillins also exhibit greater activity against Gram-negative bacteria at the expense of activity against Gram-positives. Unfortunately, these penicillins are highly susceptible to β-lactamases. Development of β-lactamase inhibitors such as clavulanic acid, sulbactam, and tazobactam, however, have greatly extended the use of these broad-spectrum penicillins (47,49). In fact, Augmentin (oral amoxicillin/clavulanic acid), Unasyn (injectable ampicillin/sulbactam), Timentin (injectable ticarcillin/clavulanic acid), and Zosyn (injectable pipericillin/ta-

zobactam) are widely used antibiotics. Carbenicillin indanyl, the first oral antipseudomonal β-lactam, and ticarcillin, a potent parenterally administered drug, are third-generation penicillins (53). The development of antipseudomonal fluoroquinolones, however, has rendered carbenicillin indanyl obsolete. Pipericillin, azlocillin, and mezlocillin are parenteral fourth-generation penicillins that have not found wide usage clinically (53).

First-generation cephalosporins typically inhibited Gram-positive bacteria other than the enterococci. Second- and third-generation cephalosporins typically inhibit Gram-negative bacteria at the expense of activity against Gram-positives. A new fourth-generation parenteral cephalosporin, cefepime (Maxipime), possesses activity against both Gram-positives and Gram-negatives (54). Additionally, cefipime penetrates bacteria well, is not susceptible to β-lactamases, and has a lower propensity to induce β-lactamases (55). Unfortunately, this broad-spectrum cephalosporin does not effectively treat either VREs or MRSAs (54). Ceftriaxone was very recently approved as a one-time-only injectable for treatment of pediatric otitis media. The concept behind the development of this drug is that the greatest problem, and reason for failure, with treatment of otitis media is patient compliance; the option of a one-time-only injectable removes this as an obstacle.

Carbapenems are very broad spectrum, β-lactamase-resistant β-lactams that contain a carbon atom instead of a sulfur atom in the 5-membered ring (47). Imipenem/cilistatin (Primaxin) was the first clinical carbapenem antibiotic, developed in the mid-1970s and made available commercially since 1986, as a synthetic biomimetic analogue of the natural product thienemycin (56). Imipenem is administered with cilistatin, an inhibitor of human renal dehydropeptidase, to protect the antibiotic from enzymatic degradation. Imipenem has broad-spectrum activity against Gram-negatives and anaerobic bacteria and moderate activity against Gram-positive bacteria (53). Unfortunately, imipenem also is limited by potential renal toxicity (56). Meropenem (Merrem) (57) is a new carbapenem antibiotic approved in 1996 for use as a broad-spectrum injectable antibiotic for treatment of intra-abdominal infections and for pediatric bacterial meningitis. The advantage of meropenem over imipenem is that meropenem is stable in the presence of renal dehydropeptidase, thereby bypassing the need for cotherapy with cilistatin (53). A third carbapenem called panipenem is available in Japan but not in the United States (58). Merck currently has a new, very potent, broad-spectrum, long-acting carbapenem, MK-826, in phase-II clinical trials (59). This new carbapenem drug candidate has been demonstrated to be highly resistant to β-lactamases, including plasmid-encoded extended spectrum β-lactamases (ESBLs) (59). Nevertheless, certain rarely pathogenic bacteria (particularly *Acinetobacter baumannii*) contain specific carbapenem-degrading enzymes known as carbapenemases (60), which may, through horizontal gene transfer, enter the broader pathogen field in the future.

An interesting new parenteral β-lactam drug candidate has recently been described. The new Glaxo-Wellcome experimental compound, GV129606, called a trinem, is a highly potent antibacterial with efficacy against Gram-positives, Gram-negatives, aerobes and anaerobes alike, plus marked resistance to extended spectrum, clinically relevant β-lactamases (61). This compound may mark a new path to be followed in the β-lactam field.

Macrolides

The macrolide antibiotics are a group of compounds produced by actinomycetes via type I modular polyketide synthases (62). The core aglycone of the prototype macrolide, erythromycin, is synthesized by the huge multifunctional enzyme, 6-deoxyerythronolide B synthase (DEBS), which consists of three, multidomain subunits, each subunit with a molecular weight around 300,000 Da. The macrolide chain is synthesized by a set of enzymatic functional domains within the proteins in a linear, stepwise fashion. Interestingly, the enzymatic active sites are arranged in the polypeptides in the same order as their apparent function (62).

Although macrolides theoretically can have ring sizes ranging from as few as 6-membered to over 30-membered, most of the clinically relevant macrolides possess 14-membered or 16-membered rings (Table 5). Current clinical antibacterial macrolides are mostly derived from erythromycin, although a wide variety of biologically active macrolides are used for several purposes across several therapeutic areas (Tables 5 and 6).

The macrolide antibiotic, erythromycin, was first discovered in 1952, the year the double-helical structure of DNA was first announced. Macrolides, which are bacteriostatic agents, inhibit bacterial protein synthesis by structure-based mechanisms: 14-membered macrolides such as erythromycin block translocation of peptidyl-tRNA, whereas 16-membered macrolides inhibit peptidyl-transfer reactions (64). There are apparently additional interactions of macrolides of both groups with ribosomes, but even after 30 years of research, these are not fully understood (64). Resistance to macrolides is most often via the MLS resistance mechanism, which is based on mechanism of action. MLS resistance, conferred by several different genes, results in methylation of adenosine-2058 of 23S rRNA. MLS-I resistance, exemplified by the *tlrD* gene, is result of monomethylation, whereas MLS-II (*ermE* -type) is the result of dimethylation (65).

Erythromycin itself is a poorly soluble, acid-labile molecule that nevertheless has enjoyed a long history of successful use. Erythromycin has been an important orally active antibiotic for many years. The efforts for years in the macrolide field have been aimed at developing new drugs that are acid-stable, possess a longer half-life, have broader ranges of activity, and are not subject to MLS resistance mechanisms. Recently, several semisynthetic macrolides arising from modifications of erythromycin, including clarithromycin (Biaxin) the 15-membered azalide, azithromycin (Zithromax), and most recently, dirithromycin (Dynabac), have made significant contributions to the antibacterials market. Among them, these macrolides represent more than $2 billion in market share.

The most exciting research area in antibacterials today may be the discovery of ketolides, erythromycin-analogue macrolides in which the L-cladinose at the 3-position of

Table 5. Clinical and Experimental Macrolides of Interest as Sorted by Ring Size

Compounds	Ring size	Comments
Spinosad (spinosyns)	12	New insecticidal (esp. against lepidopterans) drugs from Dow AgroSciences
Methymycin	12	Antibacterial; produced by same PKS and gene cluster as picromycin (63)
Clarithromycin	14	Biaxin; 6-*O*-methylerythromycin; acid stable form available in 1-day dosing
Dirithromycin	14	Dynabac; approved in 1995 as orally available, single dose/day formulation
Erythromycin A	14	Original 14-membered macrolide antibacterial agent
Flurithramycin	14	8-Fluoro-erythromycin
Galbonolides	14	Group of 14-membered antifungal macrolides
Ketolides	14	HMR3004 and HMR3647 are new, experimental, 3-oxo-semisynthetic macrolides
Oleandomycin	14	Antibacterial compound used as agricultural feed additive
Picromycin	14	Synthesized from same PKS and gene cluster as methymycin (63)
Roxithramycin	14	9-[*O*]-(2-Methoxyethoxy)methyloxime] derivative of erythromycin
Rustmicin	14	Experimental 14-membered ring antifungal agent of the galbonolide group
Troleandomycin	14	Acid-stable acetylated ester of oleandomycin; antibacterial
Azithramycin	15	Zithromax; Orally active azalide; semisynthetic analogue of erythromycin
Avermectin	16	Veterinary use; also use to treat human onchocerciasis (river blindness)
Doramectin	16	Mutational biosynthetic analogue of avermectin; antiparasitic
Epothilone	16	Experimental myxobacterial antitumor drug; targets microtubule formation
Ivermectin	16	Semisynthetic analogue of natural product avermectin; antiparasitic
Josamycin	16	Used in both human and veterinary treatments
Midecamycin A_1	16	Antibacterial agent
Milbemycin	16	Used in Europe as antiparasitic agent
Miocamycin	16	Acetylated ester of midecamycin gives oral bioavailability
Rokitamycin	16	Propionyl ester of leucomycin A5; active against macrolide-resistant Gram-positives
Spiramycin	16	Used in France as antibacterial agent
Tylosin	16	Veterinary use as antibacterial feed additive
Soraphen	18	Myxobacterial macrolide patented as an agricultural antifungal agent
Ascomycin	23	Also called FK-520 and immunomycin; immunosuppressive compound
Tacrolimus (FK-506)	23	Immunosuppressive agent
Rapamycin	31	Immunosuppressive agent

erythromycin has been replaced by a keto moiety. The ketolides are particularly interesting because they do not induce MLS-resistance mechanisms in pathogenic organisms. This is critical because MLS resistance is typically not constitutive. The two leading ketolide drug candidates, RU64004 (also known as HMR 3004) and RU66647 (also known as HMR 3647), have been shown to be potent inhibitors of Gram-positive bacteria, particularly the pneumococci, including those that show marked resistance to other macrolides such as erythromycin and azithromycin (66). The ketolides, however, show poor activity against Gram-negative bacteria.

Recently, a series of new, 14-membered macrolide antifungal compounds known collectively as galbonolides was discovered (67), of which one, rustmicin, has been investigated in more detail (68). The data indicate that rustmicin and, by extension, possibly the other galbonolides appear to interact with the α-mannan of the cell wall by a potentially novel mode of action (68). Several years ago, interest was sparked when a macrolide known as FK506 (Tacrolimus) was found to possess significant immunosuppressive action (69). Additional compounds of interest in this area included rapamycin, which is expected to be approved soon, and immunomycin, which was dropped because of a narrow efficacy-to-toxicity window (69). Both the milbemycins and avermectins have been used for several years as antiparasitic drugs (70). Ivermectin has been used and approved to treat human diseases such as river blindness and intestinal threadworm. Additionally, macrolides such as tylosin have long been important as animal-feed additive growth promotants in the agricultural area (71).

During the past few years, interest in macrolides as bioactive natural products has grown significantly in areas other than the anti-infective area. As shown in Table 6, there currently is active experimental interest in macrolides in multiple therapeutic areas, including antibacterial, antifungal, antiparasitic, anticancer, immunosuppressive, neuroimmunophilin, gastrointestinal, respiratory, and agricultural areas.

There are five human therapeutic areas in which macrolides have not been used clinically, but in which there is currently considerable interest. An acid-catalyzed breakdown product of erythromycin, called motilide, has been shown to be a potent agonist of the motilin receptor, but it is unstable. Abbott has the lead drug candidate in this area, ABT-229, a 12-deoxyerythromycin homologue possessing modification in the sugars, which is stable and has reasonably high potency (74). It has recently been demonstrated that tacrolimus (formerly FK506) and rapamycin possess additional activities other than those related to immunosuppression. Both tacrolimus and rapamycin have been demonstrated to promote neuron cell growth under conditions in which they can still bind the FK-binding protein in the absence of the second ligand (calcineurin or RAFT, respectively) (73). The macrolide, roxithramycin, currently an investigational drug for treatment of asthma (75), potentially has multiple modes of action for this indication, including inhibition of glycoconjugate secretion to

Table 6. Therapeutic Areas in which Macrolides Are Being Studied as Potential Bioactive Molecules

Therapeutic area	Example macrolides	Comments
Antibacterial	Erythromycin analogues	Traditional area of strength based on erythromycin (64)
Antifungal	Rustmicin, galbonolides	New activities of interest; interaction with mannans (67,68)
Antiparasitic	Avermectin, doramectin	Avermectin recently approved for human use in U.S. (70)
Anticancer	Epothilone	Tubulin stabilizer activity similar to Taxol (72)
Immunosuppressive	Tacrolimus, rapamycin	Substitutes for cyclosporin (69)
Neuroimmunophilin	Rapamycin and analogues	Stimulation of neuronal growth (73)
Motilin receptor	Motilide, ABT-229	Treatment of gastroesophageal reflux disorder (74)
Respiratory	Roxithramycin	Antiasthmatic by multiple potential mechanisms (75)
Agricultural	Tylosin	Antibacterial and growth promoting feed additive (71)

increase mucous clearance; clearance of *Chlamydia pneumoniae*, a possible causative or implicated agent in asthma; suppression of T-cell cytokines interleukins 2, 3, and 4; and inhibition of cholinergic neuroeffector transmission in airway smooth muscle cells. Finally, a macrolide produced by the myxobacterium, *Sorangium cellulosum*, was found to possess tubulin-stabilizing activity similar to Taxol (72), a $1 billion per year anticancer drug. Several companies and research laboratories have shown strong interest in trying to develop epothilone analogues as potential antitumor drugs.

Streptogramins

For more than 25 years, pristinamycin, a combination of two antibiotics, pristinamycin I_A (a peptidic macrolactone belonging to the streptogramin B group) and pristinamycin II_A (a polyunsaturated macrolactone belonging to the streptogramin group A) (76), has been used in Europe as a drug to combat severe infections caused by both Gram-positive and Gram-negative bacteria, without development of resistance against it (76). Each component of pristinamycin is bacteriostatic with poor efficacy, but the combination is bacteriocidal with good potency against a broad range of bacteria (76). The usefulness of pristinamycin, however, has been limited because of its poor solubility (77). In February 1996, however, Rhône-Poulenc Rhorer introduced Synercid, a new variation on pristinamycin. A mixture of the two water-soluble, semisynthetic streptogramins, quinupristin (RP57669) and dalfopristin (RP54476) (78), Synercid (RP 59500) is currently in phase III clinical trials worldwide and is currently available in the United States and Europe in an emergency-use program. More than 700 patients have received Synercid in this program. In one study, 70% of 115 patients with VRE infections were successfully treated (i.e., infection cleared) with Synercid (5). Synercid has been found to be bactericidal for most Gram-positive bacteria, but bacteriostatic for *Listeria monocytogenes* (78). In one study, 96% of all staphylococci tested were sensitive to Synercid in vitro (76), suggesting that this drug may become the "vancomycin of the future," a drug of last resort for use against severe bacterial infections. One hopeful aspect concerning Synercid is that it is unaffected by classical rRNA-methylase-conferred MLS resistance mechanisms, the most common resistance mechanisms found in staphylococci (79). Other MLS resistance mechanisms such as acetylation or hydrolysis of streptogramins still affect these compounds, however, so they are not completely impervious to bacterial resistance mechanisms (79). A recent study showed that Synercid at 2 μg/mL inhibited nearly 99% of vancomycin-resistant *E. faecium* clinical isolates tested in the United States (80), giving rise to hopes that Synercid will help to contain the growing VRE problem.

Tetracyclines

The tetracyclines are among the earliest of antibacterial antibiotics to be discovered, with chlortetracycline being discovered in 1948. The natural tetracyclines are produced by various species of actinomycetes; *Streptomyces aureofaciens* produces both chlortetracycline and tetracycline, the production of each which can be optimized by specific fermentation conditions (81), *Streptomyces rimosus* produces oxytetracyclines, and dactylocyclines are produced by *Dactylosporangium* spp. and *Actinomadura brunnea* (81).

Tetracyclines function as antibiotics by inhibiting protein translation, specifically by disrupting codon-anticodon interactions of tRNA and mRNA at the acceptor site of the ribosome. Because they inhibit protein synthesis, tetracyclines, like macrolides, are bacteriostatic. Tetracyclines are widely used not only in human health, but also in animal health programs; perhaps the widest current use of chlortetracycline and oxytetracycline is as prophylactic drugs in the fish husbandry industry (81).

Bacteria typically develop resistance to tetracyclines by acquiring a plasmid- and transposon-borne resistance gene (*tetA*) encoding an efflux pump (81). Thus, tetracycline resistance is one of the primary resistances observed as being transmitted horizontally among different groups of bacteria. Additionally, a second mechanism, mediated by TetM, gives resistance by inhibition of the interaction of the tetracyclines with the ribosome. This, unfortunately, has led to the diminution of use of the classical tetracyclines for most human health indications. Some of the more modern tetracyclines, such as minocycline, are important for treating MRSAs (81).

The most exciting advance in the tetracycline area has been the development of the glycylglycine tetracyclines (81). Two new semisynthetic glycylglycines, *N*,*N*-dimethylglycylamido-9-aminominocycline (DMG-MINNO) and 9-amino-6-demethyl-6-deoxytetracycline (DMG-DMDOT), have been developed by Lederle for their very broad spec-

trum of activity against Gram-positive, Gram-negative, aerobic, and anaerobic bacteria (54). Their spectrum includes, perhaps surprisingly, bacteria as widely disparate as mycoplasmas, MRSAs, VREs, *Neisseria* spp., *Haemophilus* spp., and *Moraxella* spp. Most importantly, glycylglycine tetracyclines bind the ribosome more tightly than traditional tetracyclines so that TetM, which mediates ribosomal resistance mechanism, cannot interfere with their activity, and they are not substrates for the TetA-mediated efflux mechanisms (81). Thus, there currently is no effective resistance mechanism yet known for these modified tetracyclines. For more information, Hunter and Hill (81) have recently described the tetracyclines in detail.

Glycopeptides

Vancomycin was first isolated in the early 1950s, making it another of the group of long-standing, well-known antibacterial drugs. For a long time, vancomycin was not used to a significant degree, but as the incidence of infections with drug-resistant organisms increased, the use for vancomycin became significantly greater. Because vancomycin is absorbed poorly by the gastrointestinal tract, it is rarely used as an oral drug. Instead, its primary use today comes as a parenteral drug as the last line of defense against many potentially lethal pathogenic bacteria, including MRSAs and antibiotic-resistant *Enterococcus faecium*. In 1987, however, after 30 years of vancomycin being on the market, resistant strains of *E. faecium* (VREs) were finally observed (39). The increased incidence of vancomycin resistance, especially in *S. aureus*, is of enormous concern (44).

Little more than a decade ago, resistance to vancomycin was unheard of and totally unsuspected, because its mechanism of action was so different from that of any other antibiotic class. Nevertheless, resistance to vancomycin and other glycopeptides was discovered and is on the rise. Vancomycin and other glycopeptides exert their activity by inhibiting peptidoglycan biosynthesis, making them potent bacteriocidal drugs. Peptidoglycan consists of a sugar backbone to which pentapeptide chains are attached. The terminal residues of the pentapeptide chain are D-alanine residues, which are critical to processing the peptidoglycan matrix as it is built. Vancomycin and other glycopeptides bind, through multiple hydrogen bonding, to the terminal D-alanine residue of the pentapeptide chain and prevent both transglycosylation and transpeptidation (82). Resistance occurs when a bacterium contains special genes (*vanAHXY*) that allow for incorporation of a D-lactic acid residue in place of the terminal D-alanine residue of the peptidoglycan pentapeptide chains, rendering the peptidoglycan unrecognizable to vancomycin binding. Two genetically distinct forms of resistance have emerged, although both operate by the same mechanism. VanA-mediated resistance gives high-level resistance to both vancomycin and teichoplanan, whereas VanB-mediated resistance is restricted to vancomycin (82). A recent study of clinical VRE isolates in the United States showed that 72% possessed the VanA phenotype, whereas only 28% possessed the VanB phenotype (80).

There are several relatively new analogues of vancomycin on the market to combat vancomycin resistance. Teichoplanan has been around for several years, but its use is restricted to Europe so far. Ramoplanan, another glycopeptide that has been around for a few years, is currently in phase II clinical trials at BioSearch Italia. Eli Lilly has developed a new chlorobiphenyl analogue of vancomycin, LY333328, that is considerably more potent than vancomycin (83). Moreover, LY333328 appears not only to inhibit bacterial growth, but to kill existing bacteria, which may give it a distinct clinical advantage over other existing antibiotics. LY333328 is very potent against VRE in vitro (84), giving hope for treatment of vancomycin-resistant pathogens.

Aminoglycosides

Aminoglycosides are not widely used in clinical therapy in the United States because of their relative toxicity profile as compared to other available antibacterials and to the rapid development of plasmid-borne resistance mechanisms through pathogen populations. Streptomycin and neomycin were first described in 1944 and 1949, respectively, so this group of antibacterials has a rich history. An incredible amount of knowledge about the biosynthesis, the genes, and the molecular mechanisms for regulation of aminoglycoside production is now known, and Piepersberg, the architect of much of this effort, has coalesced this enormous wealth of information in a recent review (85). One new aminoglycoside worth mentioning is the development in Japan of arbekacin, a new aminoglycoside that has potent activity against methicillin-, gentamycin- and tobramycin-resistant *S. aureus* strains (86). The incidence of methicillin-resistance in Japan has risen to nearly 60%, and only vancomycin and the newly developed arbekacin are approved for treatment of these strains. The new aminoglycoside is apparently effective because the widely distributed resistance mechanism, a bifunctional aminoglycoside phosphotransferase (APH[2″])-6′-acetyltransferase (AAC[6′]) which confers resistance to other aminoglycosides such as gentamycin and tobramycin, is relatively ineffective at modifying the new drug (86). It is probable that within a few years, APH(2″)-AAC(6′) will evolve so that it will be able to modify arbekacin.

Other Antibacterial Compounds and New Activities of Interest

Several new antibacterial drugs recently have been introduced to the market or in clinical trials that hold hope for containing the problems of newly emergent pathogens and increased resistance to existing drugs. Several of these are listed in Table 7. A comparison of the minimal inhibitory concentrations (MICs) to kill 90% of *S. pneumoniae* treated in vivo (MIC_{90}s, in μg/mL) was made of several of the new antibacterials with the following results (in order of efficacy) (87): ketolides (new macrolides containing keto groups at C-3), 0.016 to 0.25; everninomicins (novel oligoglycoside from marine organism), 0.06 to 0.25; glycylglycines (modified tetracyclines), 0.12 to 0.5; new fluoroquinolones (e.g., sparfloxacin and trovafloxacin), 0.12 to 0.5; oxazolidinones (new class of chemosynthesized antibacterials), 0.5 to 1.0; and new streptogramins (e.g., Synercid), 0.25 to 2.0. These can be compared to MIC_{90}s for the ex-

Table 7. Newly Approved and Experimental Antibacterial Agents

Compound or drug	Class[a]	Comments
Azithromycin	Azalide	New 15-membered macrolide containing a nitrogen in the ring; approved for pediatric use against otitis media
Trovafloxacin	Quinolone	New very broad-spectrum, high potency quinolone (Pfizer)
Mupirocin	Pseudomonic acid	The pseudomonad product, which inhibits bacterial isoleucyl tRNA synthetase, was recently approved in the U.S. as a topical drug
Cefepime	Cephalosporin	Fourth generation very broad spectrum cephalosporin; has potency of cefotaxime against Gram-positive bacteria and of ceftazidime against Gram-negative bacteria
Oxazolidinones	Chemically made	Two new oxazolidinones are being tested as potential broad-spectrum antibiotics for use against VRE and staphylococci (Upjohn)
Synercid	Streptogramins	Now in phase III clinical trials; combination of quinupristin (RP57669) and dalfopristin (RP54476), both semisynthetic streptogramins
Ziracin (SCH27899)	Everninomycin	Although discovered in 1979, everninomycins were not pursued because of their narrow spectrum; now Ziracin is a leading clinical candidate for treatment of VRE and staphylococci
PA 1648	Ansamycin	Rifampin-derivative being developed that has excellent activity against *Mycobacterium tuberculosis* and *M. avium–M. intracellulare* complex, the most common opportunistic bacteria infection of AIDS patients (PathoGenesis)
Meropenem	Carbapenem	New, safer penem with better activity against Gram-negative bacteria
Rifapentine	Ansamycin	FDA approved in 1998 for treatment of pulmonary tuberculosis
MK-826	Carbapenem	Long-acting, stable, broad-spectrum carbapenem now in phase II clinical trials (Merck)
Trinem	Novel β-lactam	New, broad-spectrum β-lactam being tested that is resistant to β-lactamases (Glaxo-Wellcome)
LY333328	Glycopeptide	New chlorobiphenyl vancomycin analogue with improved potency over vancomycin; currently in phase II clinical trials (Eli Lilly)

[a]Natural product class, except where noted.

isting drugs on various (and variably resistant) strains of *S. pneumoniae*: amoxicillin-clavulanate (Augmentin), 0.06 to 1.0; cefuroxime (semisynthetic cephalosporin), 0.12 to 2.0; cefpodoxime (semisynthetic cephalosporin), 0.25 to 2.0; and cefixime (semisynthetic cephalosporin), 0.5 to 16 (87). As important as MICs are a low MIC does not alone make a good drug. The other critical factor is the development of antibacterials that can be taken orally in once-per-day doses to increase patient compliance. Thus, the time for a particular drug dose to maintain in vivo efficacy above MIC ($T > MIC$) value is also critical to dial into the structure of newly developed drugs (87).

Most of the new antibacterials on the market are modified analogues of existing, well-known classes of antibacterials: azithromycin and clarithromycin, as well as the new experimental ketolides, are based on modifications to the erythromycin core structure; trovafloxacin and sparfloxacin are new fluoroquinolones; cefepime is a new cephalosporin; meropenem and MK-826 are new carbapenems; trinem is a novel β-lactam; LY33328 is a new glycopeptide; and Synercid is comprised of two new streptogramins. Only the oxazolidinones, a chemical synthesis class, represents a new, novel class of antibacterials.

Pharmacia-Upjohn is testing two new oxazolidinones for their ability to treat multiple-drug-resistant strains of *Mycobacterium tuberculosis* (23). These new chemically synthesized antibiotics inhibit protein synthesis in a manner different from natural product protein synthesis-inhibitory antibiotics, have novel structures unrelated to structures of other antibiotics, and have a wide spectrum of activity (23). Experiments have shown that, at least in in vitro tests, pathogens do not form resistance against these antibiotics at significant frequencies.

Mupirocin (Bactroban) was approved in 1997 by the U.S. FDA for use as a topical agent, especially against nasal MRSA. Mupirocin has been used in the U.K. since 1985 and is currently approved in 90 countries. Produced as pseudomonic acid by *Pseudomonas fluorescens*, mupirocin is interesting because of its novel mode of action; it inhibits bacterial protein synthesis specifically through its binding to isoleucyl tRNA synthetase (88). Unfortunately, mupirocin-resistant MRSA have emerged, which is not surprising since the target is a single enzyme.

Ziracin is an interesting, if still-evolving story. The everninomycins have been known since the early 1960s as bacterial protein synthesis inhibitors, but they have attracted little attention for many years (89). Ziracin, an aromatic moiety-containing oligosaccharide, was first isolated in 1979 from a marine actinomycete, *Micromonospora cabonacea* (89). At that time, the everninomycin was not pursued because of its relatively narrow spectrum of being potent against only Gram-positive bacteria. With the increased resistance of Gram-positive bacteria to existing drugs, interest in the everninomycin, Sch27899, was revived. Now, Sch27899 is in clinical development as a novel antibacterial. The problem is that a similar everninomycin, avilamycin (MaxusG), has been used for several years in the European Union as a feed additive in the agricultural market. Recently, avilamycin-resistant isolates of *E. faecium* have been found in broiler chickens that are cross-resistant to Ziracin (90). This discovery is of great concern, potentially cutting short the effectiveness, and certainly the life span, of Ziracin if it is ever marketed (90).

ANTIFUNGAL AGENTS

The need for new antifungal drugs has never been greater for a variety of reasons. Fungal infections fall into two ma-

jor categories, those that are localized, dermal, and mostly annoyances, such as ringworm, athlete's foot, and nail infections (onychomycoses), and those that are disseminated and life threatening. Life-threatening, disseminated fungal diseases have risen dramatically during the past few decades, making this an important area for new drug development. Typically, life-threatening fungal infections occur in patients with impaired defenses, such as patients with AIDS; with other debilitating diseases; who are under treatment with immunosuppressive drugs (usually after organ transplants, to prevent tissue rejection); who are receiving chemotherapy; or with severe burns. Additionally, with the potentially broader use of immunosuppressive drugs for other indications such as arthritis, the potential for fungal infections may be expanded.

The most important pathogenic fungi are *Candida albicans* (disseminated candidiasis), *Candida* spp., *Aspergillus fumigatus* (systemic aspergillosis), *Cryptococcus neoformans* (disseminated cryptococcosis leading to meningoencephalitis), *Pneumocystis carinii* (*P. carinii* pneumonia [PCP]), *Histoplasma capsulatum* (histoplasmosis), and *Coccidioides immitis*, the final example being a fungal pathogen very rich in chitin (91). The most important fungal pathogens associated with immunocompromised patients, such as those with AIDS or receiving immunosuppressive therapy, are *P. carinii*, *C. neoformans*, and *Candida* spp.

The current antifungal drugs and leading clinical candidates are listed in Table 8. Current potent antifungal drugs have problems that make them less than optimal for clinical use. For example, amphotericin B, the antifungal drug of choice in life-threatening cases, causes acute toxicity, resulting in it not being well tolerated in a large percentage of patients, although it is efficacious against PCP (92). Azoles (e.g., clotrimazole, fluconazole, ketoconazole), which target membrane ergosterol synthesis, are static in vitro; they stop fungi from growing but do not kill them. Such agents, which depend on the immune system to clear the existing fungal infection, are used in immunocompromised patients but, because they do not clear the infections, require lifelong therapy to prevent recurrence. Additionally, *Aspergillus* spp. are naturally resistant to azoles, preventing their use in typically aggressive systemic aspergillosis infections that result in high mortality rates. Finally, the incidence of azole-resistant fungi is increasing, particularly among *Candida* spp. such as *C. glabrata* and *C. krusei*.

Because most of the current antifungal drugs suffer from either toxicity (amphotericin B) or the limitations of being static in vitro and increasing resisted by target populations (azoles), there is a critical need for new antifungal agents that act against new targets. Azoles target lanosterol demethylase, a P450 enzyme involved in the synthesis of ergosterol (93), a sterol found in fungal cell membranes but not mammalian cell membranes. Antifungal polyenes, including amphotericin B, nystatin, candicidin, and faerifungin, also target the cell membrane of fungi, but in this case by the formation of transmembrane pores and other membrane-associated effects (94). Although polyene antifungal agents are typically fungicidal, they also are not generally well tolerated by a large majority of patients, making them drugs used as last resorts.

Drug resistance in fungi, which more closely resembles drug resistance in tumor cells, is inherently different from drug resistance in bacteria. First, there is no evidence for plasmid-borne drug resistance mechanisms in fungi, so resistance is limited to chromosomal and nontransferable resistance genotypes. Most fungal drug resistance is effected

Table 8. Clinical and Lead Experimental Antifungal Drugs

Antifungal drug	Type	Comments
Amphotericin B	Polyene (NP)	Clinical; broad-spectrum antifungal agent
Atovaquone	Substituted naphthoquinone (CS)	Clinical; used for *P. carinii* infections
Clotrimazole	Oral fluconazole (CS)	Clinical; for vaginal candidiasis and dermal use
Fluconazole	Bis-triazole (CS)	Clinical; Orally active, relatively broad spectrum
Flucytosine	Flurocytosine	Clinical; for cryptococcal meningitis and candidiasis
Itraconazole	Triazole (CS)	Clinical; used for histoplasmosis and blastomycosis
Ketoconazole	Triazole (CS)	Clinical; broad spectrum
LY303366	Lipopeptide echinocandin (NP)	Phase II clinical trials (Eli Lilly)
Miconazole	Azole (CS)	Clinical; oral, parenteral, and topical; broad spectrum
MK-0991	Lipopeptide pneumocandin (NP)	Phase III clinical trials (Merck); broad spectrum
Natamycin	Polyene (NP)	Clinical; for aphthalmic treatment
Nikkomycin Z	Nucleoside dipeptide (NP)	Phase I clinical trials; chitin synthesis inhibitor
Nystatin	Polyene (NP)	Clinical; orally active and broad spectrum
Pentamidine isethionate	Aromatic diamidine (NP)	Clinical; for *P. carinii* infections
Pradamycins	Benzonaphthacene quinone (NP)	Phase I clinical trials
Rustmicin	Macrolide (NP)	Unusual target; plasma protein interactions
Sordarin analogue	Unusual natural product (NP)	Experimental compound
Sulconazole nitrate	Imidazole (CS)	Clinical; for topical treatment
Sulfanilamide	Sulfonamide (CS)	Clinical; for topical vulvovaginitis
Terbinafine	Allylamine (CS)	Clinical (Lamisil); orally active; for nail fungal infections
Thiabendazole	Benzimidazole (CS)	Clinical; for treatment of worms and fungi
Trimetrexate	Diaminoquinazoline (CS)	Clinical; for treatment of *P. carinii*

Notes: This list may not represent all clinical or experimental antifungal drugs. Generic names are used. NP = natural product; CS = chemically synthesized.

by changes in membrane permeases, leading to exclusion of the drugs, induction of membrane-bound efflux mechanisms, modifications of activases required for the activation of some antifungal drugs into their active forms, and site-specific mutations of target enzymes. Degradation of antifungal drugs by target organisms (e.g., similar to the action of β-lactamases) or modification of antifungal drugs to inert forms (e.g., aminoglycoside phosphotransferases) has not been observed thus far.

With the sequencing of the entire *Saccharomyces cerevisiae* genome, a study was conducted to search for novel targets, unique to fungi or different enough in fungi, which might be used to develop new antifungal agents with novel mechanisms of action (95). Of 6,000 genes analyzed, only a dozen or so potential targets were identified, including chitin synthetase, β-1,4-glucan synthetase, tubulin, elongation factor 2,*N*-myristoyl transferase, acetyl-CoA carboxylase, inositol phosphoryl ceramide synthase, membrane ATPase, mannosyl transferase, tRNA synthetases, lanosterol dehydrogenase, lanosterol synthase, and squalene epoxidase.

The most widely considered new targets for fungicidal activity are all involved with cell wall biosynthesis or structure. Yeast cell walls, which make up 15 to 25% of the cell dry weight, are typically comprised of β-1,3- and β-1,6 glucan polymers (about 50%), α-mannan (highly glycosylated mannoproteins; about 50%), and chitin (about 2%). Although different species of fungi contain different percentages of these components in their cell walls, the major constituents are typically present (93). Thus, the biosynthesis or structural integrity of β-1,3-glucans, β-1,6-glucans, and α-mannans represent potential excellent targets. The pneumocandins and echinocandins, novel lipopeptides produced by *Glarea lozoyensis* and similar fungi, inhibit the synthesis of β-1,3-glucans (96). Two new drug candidates, Merck's MK-0991 and Eli Lilly's LY303366, are semisynthetic lipopeptide echinocandin derivatives that specifically target the Rholp dissociable subunit of the heterodimeric transmembrane 1,3-β-(D)-glucan synthase (92,93,96). Because there is apparently only a single gene that is responsible for the synthesis of this subunit, this is considered an excellent target. The pneumocandin experimental drug, MK-0991, has been shown to possess efficacy against *Aspergillus fumigatus, C. albicans, H. capsulatum*, and *P. carinii* (but not against *Cryptococcus neoformans*), suggesting that this would make an outstanding potential broad-spectrum candidate for disseminated, life-threatening fungal diseases, particularly in immunocompromised patients (92).

Other targets of interest that have been pursued recently include chitin and sphingolipid biosynthesis. Fungi contain three chitin synthetases (CSI, CSII, CSIII) that catalyze the synthesis of the polymer from individual units of UDP-*N*-acetylglucosamine. The leading antifungal chitin biosynthesis-inhibitor drug candidate is nikkomycin Z, a nucleoside dipeptide substrate analog of UDP-*N*-acetylglucosamine which competitively inhibits chitin synthases (91). Although nikkomycin Z is effective in vitro against *C. albicans, C. immitis*, and other fungi with relatively high chitin contents, it does not kill *Saccharomyces cerevisiae* (92,94). The polyoxins, also peptidyl-nucleoside natural products, also inhibit chitin synthase, but none of these has been developed yet as a clinical candidate. Rustmicin (also called galbonolide A), a 14-membered macrolide produced by *Micromonospora chalcea*, inhibits inositol phosphoceramide synthase at pM levels, resulting in the loss of all complex sphingolipids and the accumulation of ceramide (68). Although rustmicin was reasonably efficacious in a murine cryptococcosis model, it suffers from severe in vivo instability, brought about by serum protein binding, and by being an excellent substrate for *S. cerevisiae* efflux pump Pdr5 (68). The pradamycins are D-amino-acid-substituted benzonaphthacene quinone glycosides that appear to act by binding to the saccharide moieties of the mannoproteins in fungal cell walls, leading to a loss of intracellular potassium that results in gross morphological changes (6). Another unusual antifungal compound is sordarin, which exerts antifungal activity via inhibition of the elongation step in protein synthesis (6,97). Glaxo-Wellcome has developed a semisynthetic sordarin, GM237354, that possesses a broad spectrum of activity and high in vivo potency (6).

ANTICANCER AGENTS

The search for natural product anticancer agents began in the late 1950s and became a reality with the discovery in 1964 of daunorubicin, a cytotoxic, antitumor drug of the anthracycline class with efficacy against a wide variety of solid tumors and leukemias (15). Five years later, doxorubicin, the 14-hydroxy analogue of daunorubicin, was discovered at Farmitalia (15). Doxorubicin, which showed a wider range of efficacy for treatment of various cancers, has been a mainstay in chemotherapy for more than 25 years. It was shown recently that doxorubicin as well as many other cytotoxic antitumor drugs (e.g., etoposide) act by trapping topoisomerase II with DNA in what has been referred to as the cleavable complex, resulting in multiple DNA strand breaks (15). Another class of anthracyclines, exemplified by aclarubicin (aclacinomycin A), were recently shown to inhibit by a different mechanism of action, by inhibition of topoisomerae II–DNA interaction (15).

Table 9 shows a list of the major antitumor drugs derived from natural products. Other natural product antitumor chemotherapy agents used for several years, but more restrictively than doxorubicin, include mithramycin, discovered in the early 1950s; the mitomycins, first described in 1956; bleomycin, described by Umezawa in 1965; and the vinca alkaloids, vinblastine and vincristine, which were first described in the period of 1958–1962. In all these cases, the mode of action is cytotoxicity through drug-DNA interactions. With the ultimate desire to get away from cytotoxic drugs that are detrimental to all cells and not just cancerous cells, these cytotoxic natural product antitumor drugs are considered to be drugs of the past and not of the future.

Interest in natural product antitumor drugs was renewed, however, with the discovery of taxanes from the pacific yew tree. Paclitaxel (Taxol), a complex diterpenoid alkaloid, was discovered as a potential antitumor agent in 1971 (98), and it has taken nearly three decades to bring

Table 9. Clinical and Lead Experimental Natural Product-Derived Anticancer Drugs

Anticancer drug	Type	Comments
Aclacinomycin A	Anthracycline	Produced by *Streptomyces galilaeus*; used in Japan and France
Bleomycin	Glycopeptide	Produced by *Streptomyces verticillus*
Carminomycin	Anthracycline	Produced by *Actinomadura carminata*; used in Russia
Dactinomycin	Acylpeptidolactone	Produced by *Streptomyces parvulus*
Daunorubicin	Anthracycline	Produced by *Streptomyces peucetius*
Docetaxel	Taxane	Taxotere; recently approved for treatment of ovarian cancer
Doxorubicin	Anthracycline	Adriamycin; Produced by *S. peucetius* subsp. *caesius*
Epothilone	Substituted macrolide	Produced by *Sorangium cellulosum*; experimental only
Idarubicin	Anthracycline	Synthetic (mimetic), orally available analogue of doxorubicin
Irinotecan	Camptothecin	Approved in 1996 for metastatic colon and rectal cancers
Mitomycin C	Cytotoxic drug	Produced by *Streptomyces caespitosus*
Mitoxantrone	Anthracenedione	Biomimetic compound
Paclitaxel	Taxane	Produced by *Taxus* spp. (yew trees); approved for use in 1994
Pentostatin	Nucleoside analogue	Produced by *Streptomyces antibioticus*
Plicamycin	Polyketide	Mithramycin; Produced by *Streptomyces argillaceus*
Tenipocide	Podophyllotoxin	Cytotoxic antitumor drug
Topotecan	Camptothecin	Semisynthetic; from *Camptotheca acuminata* (Chinese tree)
Vinblastine sulfate	Vinca alkaloid	Produced by *Catharanthus roseus* (periwinkle)
Vincristine sulfate	Vinca alkaloid	Produced by *Catharanthus roseus* (periwinkle)
Vinorelbine tartrate	Vinca alkaloid	Semi-synthetic vinblastine analogue

it to the market. One of the unique features that helped move the taxanes from experimental drugs to clinical candidates was the finding that they acted as tubulin stabilizers, a novel mode of action (99). Docetaxel (Taxotere) and paclitaxel (Taxol) have excellent activity against ovarian and breast cancers. Additionally, Taxol and semisynthetic taxanes may be efficacious against a wide range of tumors. Thus far, the greatest impediment to the development of these drugs on a large scale was that they are only found at levels of about 0.01% (w/w) in the bark of several related yew trees, *Taxus brevifolia* (Pacific yew, an endangered species), *Taxus baccata*, *Taxus cuspidata* (Japanese yew), and others (100). Moreover, the chemical synthesis of paclitaxel is very difficult, requiring more than 30 steps. Thus, paclitaxel is currently made semisynthetically from 10-deacetylbaccatin III, produced by the needles of the Himalayan yew (100). Other novel antitumor drugs developed in recent years include irinotecan (Camptosa) and topotecan (Hycamptin), both which are inhibitors of topoisomerase I.

FUTURE DIRECTIONS IN NATURAL PRODUCTS DRUG DISCOVERY

Several approaches are being taken to address the issues of MRSA, VRE, other antibiotic-resistant bacteria, newly emerging infectious diseases and related problems. One approach is the use of genomics to help to identify new targets amongst bacteria and fungi. As mentioned previously, sequencing of the yeast genome helped to identify about a dozen unique targets for new antifungal drugs. It appears that the incidence of fungal resistance to new drugs targeted against single, critical enzymes, such as glucan synthase, is very low. Similarly, it is hoped that genomic analysis of bacteria is as useful in identifying new targets for novel antibacterials. Unfortunately, there is a significant difference between bacteria and fungi and between bacterial infections and fungal infections that must be considered. First, any single-site bacterial target for which single mutations can render resistance cannot be seriously considered because of the high frequency of mutations (10^{-8}) that occur in bacteria (77). Second, many resistance mechanisms are passed horizontally among bacterial populations by plasmid- and transposon-mediated transfer, which renders many antibiotics useless. Even the fluoroquinolones, which for years were thought to target only DNA gyrase, are now known to target at least two separate enzymes (101). This is critical, too, because single mutations in bacterial gyrases have been shown recently to yield resistance to several fluoroquinolones (102). What is the best answer? Although there is probably no single best answer for all resistance problems, the clear-cut best approach is still to inhibit cell wall biosynthesis, preferably with a novel mechanism of action or at several sites. Alternatively, inhibition of systems or processes, such as protein or nucleic acid synthesis, again at multiple sites, appears to be back in vogue. The conundrum, however, is that protein synthesis inhibitors are bacteriostatic agents, which in recent years have been thought not to be desired. Nevertheless, the newer macrolides, dirithromycin, clarithromycin, and azithromycin, all protein synthesis-inhibiting bacteriostats, are effective drugs.

The Pharmaceutical Research and Manufacturers of America organization announced that as of September 1998, 136 new anti-infective drugs were currently in advanced stages of development by 78 different companies. Of these, 27 compounds are antibacterials, 31 are antivirals, 12 are antifungals, 42 are vaccines, and 24 are in the "other" category, but still relating to anti-infection indications. As described throughout this review, there are unmet needs in all fields of antibiotics. The wealth of new

drugs on the horizon bodes well for the future of antibiotics of all classes.

BIBLIOGRAPHY

1. S.B. Levy, *The Antibiotic Paradox: How Miracle Drugs Are Destroying the Miracle*, Plenum, New York, 1992.
2. E.J. Vandamme, in E.J. Vandamme ed., *Biotechnology of Industrial Antibiotics*, Dekker, New York, 1984, pp. 3–31.
3. A.L. Demain, in C.L. Hershberger, S.W. Queener, and G. Hegeman eds., *Genetics and Molecular Biology of Industrial Microorganisms*, American Society for Microbiology, Washington, D.C., 1989, pp. 1–11.
4. K.A. Reynolds and A.L. Demain, in W.R. Strohl ed., *Biotechnology of Antibiotics*, Dekker, New York, 1997, pp. 497–520.
5. W.R. Strohl, in W.R. Strohl ed., *Biotechnology of Antibiotics*, Dekker, New York, 1997, pp. 1–47.
6. Y.-Z. Shu, *J. Natl. Prod.* **61**, 1053–1071 (1998).
7. G.M. Cragg, D.J. Newman, and K.M. Snader, *J. Natl. Prod.* **60**, 52–60 (1997).
8. J. Bérdy, in V.G. Debabov, Y.V. Dudnik, V.N. Danilenko eds., *Proceedings of the Ninth International Symposium on the Biology of the Actinomycetes*, All-Russia Scientific Research Institute for Genetics and Selection of Industrial Microorganisms, Moscow, 1995, pp. 13–34.
9. E. Culotta, *Science* **264**, 362–363 (1994).
10. H. Umezawa, *Annu. Rev. Microbiol.* **36**, 75–99 (1982).
11. A.W. Alberts, J. Chen, G. Kuron, V. Hunt, J. Huff, C. Hoffman, J. Rothrock, M. Lopez, H. Joshua, E. Harris, A. Patchett, R. Monaghan, S. Currie, E. Stapley, G. Albers-Schonberg, O. Hensens, J. Hirschfield, K. Hoogsteen, J. Liesch, and J. Springer, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3957–3961 (1980).
12. American Druggist, *http://www.rxlist.com/top200a.htm*, Feb. 1998.
13. N. Pfeiffer, *Gen. Eng. News* **16**, 18–19 (1996).
14. SCRIP's 1995 Yearbook, PharmaBooks Ltd., New York.
15. W.R. Strohl, M.L. Dickens, V.B. Rajgarhia, A.J. Woo, and N.D. Priestley, in W.R. Strohl ed., *Biotechnology of Antibiotics*, Dekker, New York, 1997, pp. 577–657.
16. T.J. Walsh and N. Giri, *Eur. J. Clin. Microbiol. Infect. Dis.* **16**, 93–97 (1997).
17. P.R. August, L. Tang, Y.J. Yoon, S. Ning, R. Muller, T.W. Yu, M. Taylor, D. Hoffman, C.G. Kim, X. Zhang, C.R. Hutchinson, and H.G. Floss, *Chem. Biol.* **5**, 69–79 (1998).
18. L.C. Vining and C. Stuttard eds., *Genetics and Biochemistry of Antibiotic Production*, Butterworth-Heinemann, Boston, 1995.
19. W.R. Strohl ed., *Biotechnology of Antibiotics*, Dekker, New York, 1997.
20. L.C. Vining, in L.C. Vining and C. Stuttard eds., *Genetics and Biochemistry of Antibiotic Production*, Butterworth-Heinemann, Boston, 1995, pp. 499–504.
21. J. Mann, R.S. Davidson, J.B. Hobbs, D.V. Banthorpe, and J.B. Harborne, *Natural Products. Their Chemistry and Biological Significance*, Addison-Wesley Longman, Essex, U.K., 1996.
22. S. Budavari ed., *The Merck Index: An Encyclopedia of Chemicals, Drugs, and Biologicals*, 12th ed., Merck Research Laboratories, Merck and Co., Whitehouse Station, 1996.
23. C.W. Ford, J.C. Hamel, D. Stapert, J.K. Moerman, D.K. Hutchinson, M.R. Barbachyn, and G.E. Zurenko, *Trends Microbiol.* **5**, 196–200 (1997).
24. J.E. McGowan, *Rev. Infect. Dis.* **5**, 1033–1048 (1983).
25. J. Davies, *Science* **264**, 375–382 (1994).
26. M.L. Cohen, *Science* **257**, 1050–1055 (1992).
27. J. Travis, *Science* **264**, 360–362 (1994).
28. S. Falkow, *Infectious Multiple Drug Resistance*, Pion Ltd., London, 1975.
29. J. Lederberg, R.E. Shope, and S.C. Oaks eds., *Emerging Infections: Microbial Threats to Health in the United States*, Institute of Medicine, National Academy of Sciences Press, Washington, D.C., 1992.
30. R.A. Weinstein, *Emerg. Infect. Dis.* **4**, 416–420 (1998).
31. Centers for Disease Control and Prevention, *Morb. Mortal. Wkly. Rep.* **42**, 597–599 (1993).
32. E. Rubinstein and F. Bompart, *J. Antimicrob. Chemother.* **39**, 139–143 (1997).
33. *World Health Report 1996*, Office of World Health Reporting, World Health Organization, Geneva, 1996.
34. J. May, K. Shannon, A. King, and G. French, *J. Antimicrob. Chemother.* **42**, 189–197 (1998).
35. K. Sunakawa, H. Akita, S. Iwata, Y. Sato, and R. Fujii, *Infection* **23**, S74–S78 (1995).
36. B.R. Bloom and C.J.L. Murray, *Science* **257**, 1055–1064 (1992).
37. E. Collatz, R. Labia, and L. Gutman, *Mol. Microbiol.* **4**, 1615–1620 (1990).
38. P.M. Rowe, *Lancet* **347**, 252 (1996).
39. A.H.C. Uttley, C.H. Collins, J. Naidoo, and R.C. George, *Lancet* **1**, 57–58 (1988).
40. G.L. French, *Clin. Infect. Dis.* **27**, S75–S83 (1998).
41. P. Moreillon and A. Wenger, *Schweiz. Mediz. Wochensch.* **126**, 255–263 (1996).
42. W. Brumfitt and J.M.T. Hamilton-Miller, *Drugs Exptl. Clin. Res.* **20**, 215–224 (1994).
43. W.C. Noble, Z. Virani, and R.G.A. Cree, *FEMS Microbiol. Lett.* **93**, 195–198 (1992).
44. H.C. Neu, *Scand. J. Infect. Dis.* **91** 7–13 (1993).
45. K.J. Christensen and P.O. Gubbins, *Ann. Pharmacother.* **30**, 288–290 (1996).
46. Y. Sumita and M. Fukusawa, *Chemotherapy* **42**, 47–56 (1996).
47. G. Cohen and Y. Aharonowitz, in P.A. Hunter, G.K. Darby, and N.J. Russell eds., *Fifty Years of Antimicrobials: Past Perspectives and Future Trends*, Soc. Gen. Microbiol. Ltd., Cambridge, 1995, pp. 139–163.
48. H. Vanderhaeghe, *Verh. K. Acad. Geneeskd. Belg.* **53**, 39–58 (1991).
49. A.S. Paradkar, S.E. Jensen, and R.H. Mosher, in W.R. Strohl ed., *Biotechnology of Antibiotics*, Dekker, New York, 1997, pp. 241–277.
50. P. Zuber and M. Marahiel, in W.R. Strohl ed., *Biotechnology of Antibiotics*, Dekker, New York, 1997, pp. 187–216.
51. B.O. Bachmann, R. Li, and C.A. Townsend, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9082–9086 (1998).

52. H. Palissa, H. von Dohren, H. Kleinkauf, H.H. Ting, and J.E. Baldwin, *J. Bacteriol.* **171**, 5720–5728 (1989).
53. D.M. Allen and M.L. Ling, *Singapore Med. J.* **35**, 626–630 (1994).
54. P.M. Southern, *Compr. Ther.* **23**, 575–582 (1997).
55. M.A. Wynd and J.A. Paladino, *Ann. Pharmacother.* **30**, 1414–1424 (1996).
56. R.C. Moellering Jr., G.M. Eliopoulos, and D.E. Sentochnik, *J. Antimicrob. Chemother.* **24**, 1–7 (1989).
57. D.N. Fish and T.J. Singletary, *Pharmacotherapy* **17**, 644–669 (1997).
58. J. Shimada and Y. Kawahara, *Drugs Exptl. Clin. Res.* **20**, 241–245 (1994).
59. C.J. Gill, J.J. Jackson, L.S. Gerckens, B.A. Pelak, R.K. Thompson, J.G. Sundelof, H. Kropp, and H. Rosen, *Antimicrob. Agents Chemother.* **42**, 1996–2001 (1998).
60. D.M. Livermore, *J. Antimicrob. Chemother.* **39**, 673–676 (1997).
61. E. Di Modugno, R. Broggio, I. Erbetti, and L. Lowther, *Antimicrob. Agents Chemother.* **41**, 2742–2748 (1997).
62. L. Katz and S. Donadio, in L.C. Vining and C. Stuttard eds., *Genetics and Biochemistry of Antibiotic Production*, Butterworth-Heinemann, Boston, 1995, pp. 385–420.
63. Y. Xue, D. Wilson, L. Zhao, H.-W. Liu, and D.H. Sherman, *Abstr. 8th Intl. Symp. Genet. Industr. Microorg.*, Jerusalem, 1998, p. 39.
64. T. Mazzei, E. Mini, A. Novelli, and P. Periti, *J. Antimicrob. Chemother.* **31**, 1–9 (1993).
65. J.L. Pernodet, S. Fish, M.H. Rouault-Blondelet, and E. Cundliffe, *Antimicrob. Agents Chemother.* **40**, 581–585 (1996).
66. R.R. Reinert, A. Bryskier, and R. Luetticken, *Antimicrob. Agents Chemother.* **42**, 1509–1511 (1998).
67. H. Achenbach, A. Muhlenfeld, U. Fauth, and H. Zahner, *Ann. N.Y. Acad. Sci.* **544**, 128–140 (1988).
68. S.M. Mandala, R.A. Thornton, J. Milligan, M. Rosenbach, M. Garcia-Calvo, H.G. Bull, G. Harris, G.K. Abruzzo, A.M. Flattery, C.J. Gill, K. Bartizal, S. Dreikorn, and M.B. Kurtz, *J. Biol. Chem.* **273**, 14942–14949 (1998).
69. K.W. Mollison, T.A. Fey, R.A. Krause, V.A. Thomas, A.P. Mehta, and J.R. Luly, *Ann. N.Y. Acad. Sci.* **685**, 55–57 (1993).
70. D.J. MacNeil, in L.C. Vining and C. Stuttard eds., *Genetics and Biochemistry of Antibiotic Production*, Butterworth-Heinemann, Boston, 1995, pp. 421–442.
71. F.G. Day, *J. Am. Vet. Med. Assoc.* **208**, 655–656 (1996).
72. D.M. Bollag, P.A. McQueney, J. Zhu, O. Hensens, L. Koupal, J. Liesch, M. Goetz, E. Lazarides, and C.M. Woods, *Cancer Res.* **55**, 2325–2333 (1995).
73. J.P. Steiner, M.A. Connolly, H.L. Valentine, G.S. Hamilton, T.M. Dawson, L. Hester, and S.H. Snyder, *Nature Med.* **3**, 421–428 (1997).
74. P.A. Lartey, H.N. Nellans, R. Faghih, A. Petersen, C.M. Edwards, L. Freiberg, S. Quigley, K. Marsh, L.L. Klein, and J.J. Plattner, *J. Med. Chem.* **38**, 1793–1798 (1995).
75. S. Konno, K. Asano, M. Kurokawa, K. Ikeda, K. Okamoto, and M. Adachi, *Int. Arch. Allergy Immunol.* **105**, 308–316 (1994).
76. J.C. Barrière, D.H. Bouanchaud, J.M. Paris, O. Rolin, N.V. Harris, and C. Smith, *J. Antimicrob. Chemother.* **30** 1–8 (1992).
77. L.L. Silver and K.A. Bostian, *Antimicrob. Agents Chemother.* **37**, 377–383 (1993).
78. T. Nichterlein, M. Kretschmar, and H. Hof, *J. Chemother.* **8**, 107–112 (1996).
79. R. Leclercq, L. Nantas, S.-J. Soussy, and J. Duval, *J. Antimicrob. Chemother.* **30**, 67–75 (1992).
80. G.M. Eliopoulos, C.B. Wennersten, H.S. Gold, T. Schuelin, M. Souli, M.G. Farris, S. Cerwinka, H.L. Nadler, M. Dowzicky, G.H. Talbot, and R.C. Moellering Jr., *Antimicrob. Agents Chemother.* **42**, 1088–1092 (1998).
81. I.S. Hunter and R.A. Hill, in W.R. Strohl ed., *Biotechnology of Antibiotics*, Dekker, New York, 1997, pp. 659–682.
82. T.I. Nicas and R.D.G. Cooper, in W.R. Strohl ed., *Biotechnology of Antibiotics*, Dekker, New York, 1997, pp. 363–392.
83. T.I. Nicas, D.L. Mullem, J.E. Flokowitsch, D.A. Preston, N.J. Snyder, M.J. Zweifel, S.C. Wilkie, M.J. Rodriguez, R.C. Thompson, and R.D.G. Cooper, *Antimicrob. Agents Chemother.* **40**, 2194–2199 (1996).
84. R.S. Schwalbe, A.C. McIntosh, S. Qaiyum, J.A. Johnson, R.J. Johnson, K.M. Furness, W.J. Holloway, and L. Steele-Moore, *Antimicrob. Agents Chemother.* **40**, 2416–2419 (1996).
85. W. Piepersberg, in W.R. Strohl ed., *Biotechnology of Antibiotics*, Dekker, New York, 1997, pp. 81–163.
86. M. Inoue, M. Nonoyama, R. Okamoto, and T. Ida, *Drugs Exptl. Clin. Res.* **30**, 233–240 (1994).
87. W.A. Craig, *Diagn. Microbiol. Infect. Dis.* **27**, 49–53 (1997).
88. J.S. Bertino Jr., *Am. J. Health Syst. Pharm.* **54**, 2185–2191 (1997).
89. A.L. Demain, *Nature Biotechnol.* **16**, 3–4 (1998).
90. F.M. Aarestrup, *Microb. Drug Resist.* **4**, 137–141 (1998).
91. J.R. Graybill, L.K. Najvar, R. Bocanegra, R.F. Hector, and M.F. Luther, *Antimicrob. Agents Chemother.* **42**, 2371–2374 (1998).
92. M.A. Powles, P. Liberator, J. Anderson, Y. Karkhanis, J.F. Dropinski, F.A. Bouffard, J.M. Balkovec, H. Fujioka, M. Aikawa, D. McFadden, and D. Schmatz, *Antimicrob. Agents Chemother.* **42**, 1985–1989 (1998).
93. M.B. Kurtz, *ASM News* **64**, 31–39 (1998).
94. J.A. Gil and J.F. Martin, in W.R. Strohl ed., *Biotechnology of Antibiotics*, Dekker, New York, 1997, pp. 551–575.
95. Y. Koltin, *Abstr. 8th Intl. Symp. Genet. Industr. Microorg.*, Jerusalem, 1998, p. 5.
96. W.W. Turner and W.L. Current, in W.R. Strohl ed., *Biotechnology of Antibiotics*, Dekker, New York, 1997, pp. 315–334.
97. M.C. Justice, M.-J. Hsu, B. Tse, T. Ku, J. Balkovec, D. Schmatz, and J. Nielsen, *J. Biol. Chem.* **273**, 3148–3151 (1998).
98. M.C. Wani, H.L. Taylor, M.E. Wall, P. Coggen, and A.T. McPhail, *J. Am. Chem. Soc.* **93**, 2325–2327 (1971).
99. P.B. Schiff, J. Fant, and S.B. Horowitz, *Nature* **277**, 665–667 (1979).
100. Y. Yukimune, H. Tabata, Y. Higashi, and Y. Hara, *Nature Biotechnol.* **14**, 1129–1132 (1996).
101. X.-S. Pan and L. Fisher, *Antimicrob. Agents Chemother.* **41**, 471–474 (1997).
102. B. Waters and J. Davies, *Antimicrob. Agents Chemother.* **41**, 2766–2769 (1997).

SECRETION FROM ANIMAL CELLS

RANDAL J. KAUFMAN
University of Michigan Medical School
Ann Arbor, Michigan

KEY WORDS

Disulfide bond formation
Endoplasmic reticulum
Glycosylation
Golgi apparatus
Posttranslational modification
Protein chaperones
Protein folding
Protein trafficking
Proteolytic processing
Tyrosine sulfation

OUTLINE

INTRODUCTION

The specialized process of secretion in mammalian cells is directed through membrane-enclosed compartments composed of vesicles. Each compartment is unique to ensure specificity and directionality for the secretion process. The compartments are maintained by mechanisms of selective targeting of newly made proteins to their correct vesicle. As proteins destined for the cell surface transit the different compartments, they attain their final conformation through a maturation process that involves protein folding facilitated by molecular chaperones, covalent modification of the polypeptide backbone, and assembly into higher-order structures.

PROTEIN TRANSLOCATION INTO THE ENDOPLASMIC RETICULUM

Targeting to the Endoplasmic Reticulum

All proteins destined for the cell surface, the extracellular space, or intracellular organelles such as lysosomes or the Golgi apparatus are first translocated across the membrane of the endoplasmic reticulum (ER). In mammalian cells the process of protein transport across the ER membrane occurs in the majority of cases while the polypeptide is being translated, which is called cotranslational translocation. In contrast, in yeast, translocation can occur either cotranslationally or after translation of the polypeptide is complete, that is, posttranslational translocation. Proteins that are targeted for transport across the ER must contain a signal that directs the process. Most secretory proteins contain an amino-terminal set of amino acids that compose the signal that directs cotranslational translocation of the protein across the ER membrane. The signal peptide is exposed as the polypeptide emerges from the 60S ribosomal subunit and mediates association of the nascent polypeptide with the cytosolic face of the ER. Although the exact amino acid sequence of signal peptides is not conserved for a given protein across species, its general character is very conserved. The signal sequence may be approximately 30–60 amino acids in length and is composed of three regions: (*1*) an amino terminal segment of variable length that has a net positive charge, (*2*) a central hydrophobic core of 6–15 residues, and (*3*) a C-terminal region that often has a helix-breaking amino acid residue such as glycine proline or serine (1). In yeast, the more hydrophobic the character of the signal peptide, the more likely the transport process occurs by a cotranslational mechanism.

The signal recognition particle (SRP) is a complex of six polypeptides and one 7S RNA molecule that targets substrates for cotranslational translocation (for review see Ref. 2). The 300-nucleotide RNA component is very conserved across all species and is homologous with the highly repetitive Alu DNA family. The mechanism by which cotranslational translocation occurs was elucidated using in vitro studies of protein import into canine pancreatic microsomes, because they are efficient at directing SRP-dependent translocation. Analysis of the SRP has identified subcomplexes that perform different functions. SRP54 (the number refers to the molecular weight of the protein) binds the signal sequence and targets it to the ER membrane. SRP54 contains a stretch of methionine residues that provide a flexible surface, much like bristles of a brush, for interaction with the variable hydrophobic stretches of amino acids in different signal peptides. After interaction with SRP54, translation elongation pauses by a subcomplex of SRP9 and SRP14. This pause helps ensure proper targeting to the ER membrane before significant portions of the polypeptide are synthesized and begin to fold.

After SRP binds the signal sequence of the nascent chain–ribosome complex, it targets the complex to the ER translocation machinery (Fig. 1). The SRP receptor (SRPR) on the ER membrane is composed of two subunits SRα and SRβ and receives the nascent polypeptide after release from SRP. SRP then recycles back to the cytosol. Both SRP54 of the SRP and each of the SRP receptor subunits are GTPases. SRP and SRPR reciprocally stimulate each others' GTPase activity. SRP in the GTP-bound form promotes high-affinity association with SRPR (4). Subsequent transfer of the signal to the translocation apparatus depends on GTP binding by SRα (5). For recycling, SRP must be released from SRPR, a process that requires GTP hydrolysis by SRP54 (6). The nascent chain–associated complex (NAC), an ER-localized protein complex, ensures fidelity of cotranslational targeting to the translocon in the ER membrane by binding nascent polypeptides that lack signal sequences, preventing their targeting to the ER membrane, and preventing free ribosomes from interacting with the ER membrane (7–9). In addition, some polypeptides also require the translocating chain–associated membrane protein (TRAM) for translocation. TRAM can be cross-linked to the signal peptide of some secretory proteins (10). The signal-containing polypeptide inserts into the membrane, and the polypeptide threads through the protein channel, composed of Sec61α, into the lumen of the ER as an extending loop.

Translocation into the ER

The proteins that are required for translocation in vitro and in vivo have been identified through genetic studies in yeast. The most important protein is Sec61p (40 kDa), a protein predicted to span the ER membrane a number of times; it is the primary subunit of the protein translocation channel (11). Sec61p is part of a heterotrimeric complex with Sbh1p and Sss1p, which are the yeast homologues of mammalian Sec62β and Sec61γ (12). In addition to the Sec61p complex, another complex, Sec62p, Sec63p, Sec71p, and Sec72p, is required for posttranslocational translocation in yeast. Sec63p has a luminal region homologous to the J-domain that is highly conserved in the DnaJ protein family, a protein family that stimulates the ATPase activity of Hsp70 family members. The yeast homologue of BiP, Kar2p, is required for posttranslational translocation, probably through its ATPase activity, which is regulated by the J-domain of Sec63p. In mammalian cells, purified SRP, SRPR, and Sec61p are the only proteins required for cotranslocational translocation (12). TRAM is required for only a subset of precursors and stimulates translocation of others.

The role of resident proteins in the lumen of the ER was studied by selectively solubilizing and reconstituting microsomal translocation activity into proteoliposomes (13). Alkaline extraction at pH 9.5 releases a class of luminal proteins that is essential for completion of polypeptide translocation (14). Among the proteins released are the stress proteins BiP/GRP78, GRP94, protein disulfide isomerase, calreticulin, and peptidyl prolyl *cis-trans*-isomerase enzymes (cyclophilin and FKBP). These studies demonstrated that BiP/GRP78 was required for translocation in these yeast-derived proteoliposomes. However, experiments in mammalian systems have not provided compelling evidence that BiP/GRP78 is required for translocation. In addition, expression of an ATPase-defective BiP only prevented the secretion of selective proteins, supporting the view that the BiP ATPase activity is not absolutely required for the secretion of all proteins (15).

Proteins that become integral membrane proteins pose a unique problem for the translocation machinery. Type I transmembrane proteins contain a cleavable N-terminal signal peptide and an internal transmembrane domain. They are oriented such that their amino terminus is in the extracellular space. For these proteins, a span of hydro-

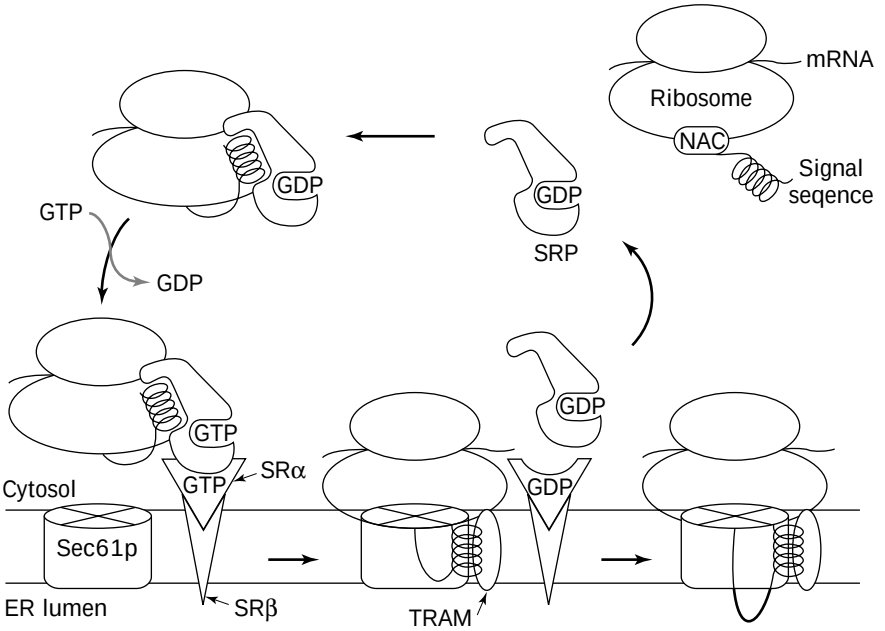


Figure 1. SRP-dependent cotranslocational translocation into the ER. SRP binds to the signal sequence of a nascent polypeptide emerging from the 60S ribosome to form a targeting complex in which elongation is arrested. SRP in the targeting complex then binds SRP receptor (composed of SRα and Srβ subunits) in the ER membrane. After GTP binding, SRP interacts with SRP receptor to result in formation of the ribosome-translocon junction, and translation elongation resumes. GTP hydrolysis stimulates the dissociation of SRP from SRP receptor so they can be recycled. As the nascent chain grows, the translocon pore allows passage of the nascent protein across the membrane. In most cases, the signal sequence is removed on the luminal side of the ER membrane by signal peptidase (not shown). *Source*: Reprinted from Ref. 3, with permission.

phobic amino acids (called the stop-transfer sequence) anchors the polypeptide in the lumen of the ER. In type II transmembrane proteins the carboxy terminus of the protein is in the extracellular space. The topology of type I and type II membrane proteins is determined by the charge distribution of polar amino acids flanking the signal-anchor domain (16). The side of the signal anchor that has more positively charged flanking residues remains intracellular, likely reflecting interactions with the negatively charged phospholipid bilayer. For multiple-spanning transmembrane proteins, each transmembrane domain displaces the prior one from the protein conducting channel where it moves laterally into the lipid bilayer. During the process of translocation, the channel is always gated, either by the ribosome on the cytoplasmic side or by BiP/GRP78 on the luminal side (17).

The cleavage site for signal peptidase is marked by small amino acids (alanine or glycine) in the 3- and 1-positions, relative to the cleavage site. Insertion into the ER membrane involves the formation of a loop structure with the amino terminus remaining in the cytoplasm and the growing carboxy terminus being continuously translocated across the membrane through Sec61p, a protein-conducting channel (18). Cleavage of the signal sequence by the signal peptidase (19,20) then releases the mature amino terminus into the lumen of the ER and is required for translocation of proteins into the secretory pathway.

The transport of a polypeptide across the ER membrane requires energy. GTP is used in the cycling of SRP and SRPR. The energy in cotranslational translocation is derived from the process of translational elongation. For posttranslational translocation, ATP is required for both function of Hsp70 in the cytosol and BiP function in the lumen of the ER. Depletion of intracellular ATP can block the secretion of some proteins, presumably by preventing dissociation of proteins from BiP (21). An ER ATP transporter has been characterized (22), and studies have directly detected ATP in the lumen of the ER by directing firefly luciferase expression to be localized in the ER (23). It is possible that increases in the intraluminal concentration of ATP will increase the secretion capacity in mammalian cells.

POSTTRANSLATIONAL MODIFICATIONS WITHIN THE SECRETORY PATHWAY

Protein translocation into and transport through the secretory pathway in eukaryotic cells is accompanied by a multitude of covalent modifications that occur on the polypeptide backbone. The sequence of posttranslational modifications that occur on a polypeptide is carefully regulated both temporally and spatially; specific modifications occur only within vesicles of the ER or Golgi compartment (Fig. 2). In many cases these modifications are required for proper polypeptide folding and secretion. In addition, posttranslational modifications can affect the half-life of the protein in the plasma and may be required for functional activity of the polypeptide. The mechanisms that direct posttranslational modifications recognize specific structural determinants within the polypeptide backbone. The efficiency of these modifications is determined by the specific host cell enzymatic machinery, the availability of substrates and cofactors, as well as structural properties of the polypeptide. To evaluate the role of posttranslational modification in protein function, investigators have studied proteins treated with chemicals or enzymes to remove modifications, proteins synthesized in the presence of inhibitors of specific modification reactions, proteins expressed in either different cell types or in cell mutants with defects in specific enzymatic machinery required to perform modifications, or proteins engineered through recombinant DNA technology that contain mutations that prevent modifications. The interpretation of results utilizing different strategies needs to be qualified to consider secondary effects due to the approach utilized. As a consequence, it is desirable to utilize several independent approaches to confirm the importance of any specific modification.

Protein Folding and Catalysts of Protein Folding in the ER

Analysis of protein folding in vitro has identified several key steps that occur during spontaneous refolding of a protein upon dilution from a denaturant. First, the polypeptide chain collapses into a compact shape to bury hydrophobic structures; this coincides with formation of secondary structures. Then native tertiary structures form as the protein passes through kinetically defined intermediates; this process may take several minutes. The partially folded intermediate is called the molten globule. Finally, the rate-limiting step lies close to the native folded state and involves formation of hydrogen bonds and disulfide bonds. Although proteins can fold into correct tertiary conformations in vitro (25), additional factors such as protein chaperones assist protein folding in vivo. Molecular chaperones compose a group of unrelated proteins that mediate the correct assembly of other proteins but are not themselves components of the final folded structure. Some molecular chaperones actually catalyze protein folding, whereas others maintain proteins in a folding-competent state and prevent protein aggregation.

A number of protein chaperones bind and stabilize partially folded polypeptides. One of the most well studied members of this family is the immunoglobulin binding protein (BiP) that is the same as the glucose-regulated protein of 78 kDa (GRP78) within the lumen of the ER (26). BiP is a member of the heat shock protein family, which exhibits a peptide-dependent ATPase activity (27), and for which expression is induced by the presence of abberently folded protein or unassembled protein subunits within the ER (28,29). Polypeptide release from BiP and transport out of the ER requires high levels of intracellular ATP (21); however, not all proteins require functional BiP for secretion (15). Intensive investigation into the role of BiP in protein folding has led to two different, but not mutually exclusive, functions for BiP. One hypothesis is that BiP assists protein folding by maintaining proteins in a conformation where they are folding competent, (30,31). This model is supported by the transient association with polypeptides destined for secretion (32–35) and an ATP dependence for proper folding and disulfide bond formation (36). In

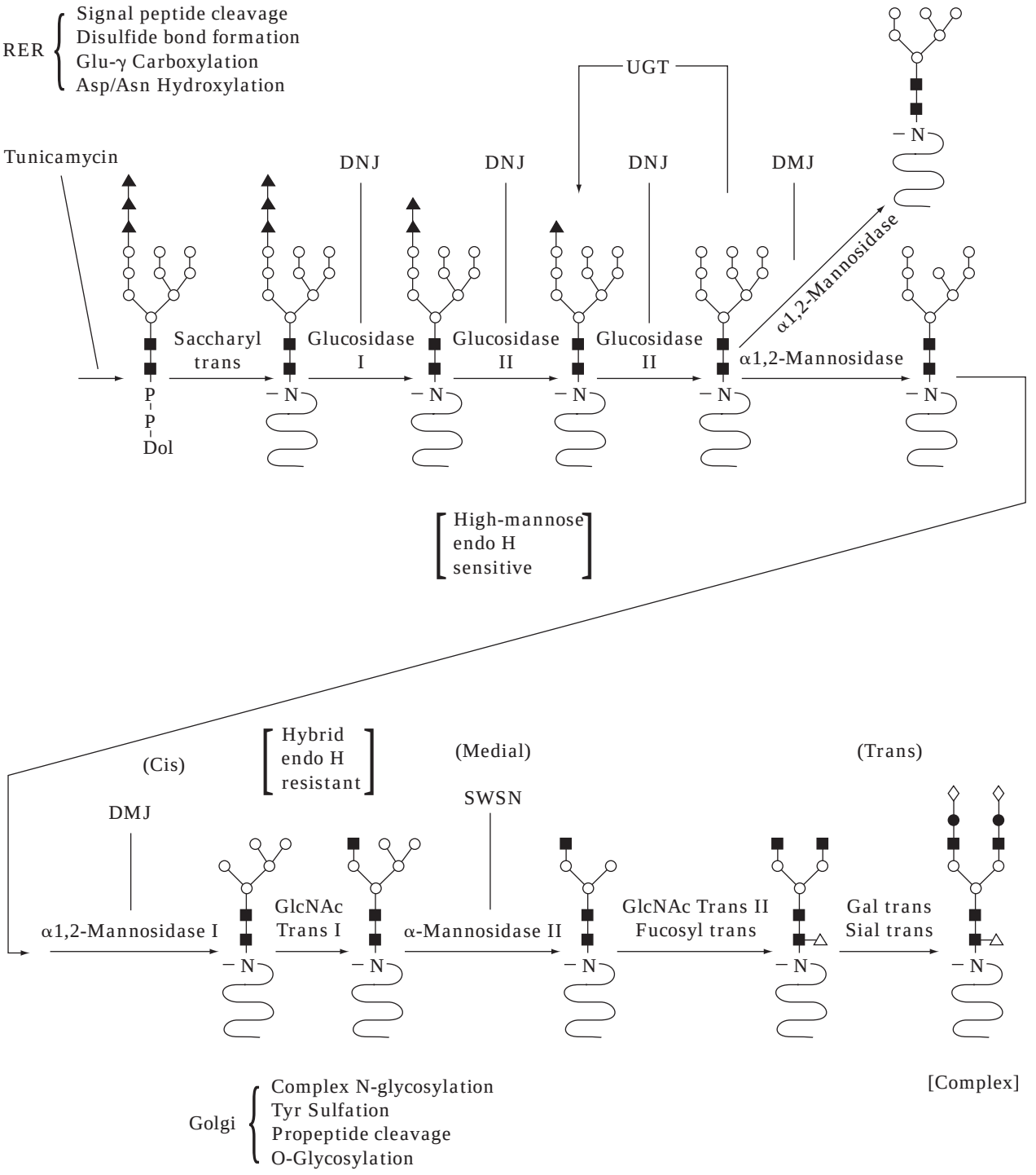


Figure 2. Posttranslational modifications within the secretory pathway. The posttranslational modifications that occur within the rough endoplasmic reticulum (RER, top) and the Golgi compartment (below) are shown. In addition, the sequence of oligosaccharide processing of N-linked oligosaccharides is shown, including the enzymes responsible and the known inhibitors (DNJ, deoxynojirimycin; DMJ, deoxymannojirimycin; SWSN, swainsonine). The point at which N-linked oligosaccharides become resistant to endoglycosidase H is shown. Sugars: ▲, glucose; ○, mannose; ■, *N*-acetylglucosamine (GlcNAc); △, fucose; ●, galactose (Gal); ◇, sialic acid (Sial) trans, transferase. UGT, UDP-glucose:glycoprotein glucosyltransferase. *Source*: Adapted from Ref. 24, with permission.

contrast, other studies (37–40) suggest that these proteins act as a retention mechanism for quality control to prevent aberrantly folded proteins from exiting the secretory pathway. However, despite many efforts to date, there is no direct demonstration that BiP binding to protein substrates actually catalyzes protein folding (41).

Disulfide bond formation occurs in the oxidizing environment of the ER. The tripeptide glutathione (γ-glutamyl, cyteinyl, glycine) is the major thiol-containing molecule in eukaryotic cells and prevents disulfide bond formation in the cytosol and catalyzes their formation in the ER. Glutathione is in equilibrium between a reduced form (GSH) and an oxidized form (GSSG). In the ER the ratio of GSH to GSSG is 5:1, an optimum for formation of disulfide bonds. In contrast, the ratio is 50:1 in the cytosol (42). Oxidized glutathione is reduced by glutathione reductase in a reaction that uses NADPH as a cofactor and occurs in the cytosol. One identified catalyst of protein folding is protein disulfide isomerase (PDI), which exchanges disulfide bonds on substrates and ensures proper disulfide bond formation, and exchange occurs prior to exit from the ER (43). PDI is also the β-subunit of prolyl hydroxylase, which catalyzes posttranslational hydroxylation of proline residue in pro-α-collagen. PDI is a soluble protein of the ER and is the most effective catalyst of protein folding known. PDI does not change the intermediates in the folding reaction but increases the kinetics of the overall process. Proper pairing of disulfide bonds is essential for transport through the ER, and pairing does not always occur in a sequential order.

Another rate-limiting step in protein folding is proline isomerization. Proteins that catalyze isomerization around peptidyl-prolyl bonds are peptidyl-prolyl isomerases (PPIases). There are two classes of these enzymes, based on their different substrate specificities and sensitivities to cyclophilin or FK506. Cyclophilin binds to cyclosporin A, a PPIase that has broad specificity. FK506 binds to FKBP, a PPIase that has narrower specificity with preference for hydrophobic residues in the P1 position. There are some very specific PPIases, one of the best studied is NinaA, a PPIase of the ER that specifically folds rhodopsins 1 and 2. *Drosophila* that have mutations in NinaA lack proper visual responses.

Asparagines and Serine/Threonine-Linked Glycosylation

High-mannose-containing oligosaccharide structures are added to asparagine residues of glycoproteins as an obligatory event for the folding and assembly of many newly synthesized polypeptides (44). This is an essential process in eukaryotic cells (45). The presence of oligosaccharides may be required for the efficient transport of individual glycoproteins through the secretory pathway (33,46), to increase the plasma half-life, or for the biological activity of glycoproteins. The luminal enzyme oligosaccharyltransferase catalyzes the transfer of a preassembled high-mannose-containing oligosaccharide core structure [glucose $(Glc)_3$mannose $(Man)_9$*N*-acetylglucosamine $(GlcNAc)_2$] from a dolichol pyrophosphate precursor onto asparagine acceptor sites within the consensus sequence Asn-X-Ser/Thr, where X can be any amino acid except proline. The mammalian oligosaccharyltransferase is a heterotrimeric transmembrane complex of ribophorin I, ribophorin II, and the oligosaccharyltransferase, which has been purified and molecularly cloned (47). The utilization of a particular consensus site for N-linked oligosaccharide attachment is determined by the structure of the growing polypeptide. As a consequence, proteins expressed in heterologous cells most frequently exhibit occupancy of N-linked sites very similar to that of the native polypeptide (48). To date there is no evidence that the N-linked oligosaccharide addition machinery can be saturated at high expression levels of glycoproteins.

After addition of the high-mannose-containing oligosaccharide core structure, trimming begins with the removal of the three terminal glucose residues, which is mediated by the action of glucosidases I and II (Fig. 2). Glucosidase I removes the terminal α1-3-glucose, and glucosidase II subsequently removes the two α1-2-glucose residues. Glucose trimming is required for binding to the protein chaperones calnexin and calreticulin within the lumen of the ER (49) (Fig. 3). Calnexin (IP90, p88, or CNX; an integral transmembrane protein) (50) and calreticulin (CRT; an ER luminal protein) (51) are homologous lectin-binding protein chaperones that transiently and selectively bind to overlapping sets of newly synthesized glycoprotein folding intermediates, thereby preventing their transit through the secretory pathway. Prolonged association with CNX and/or CRT is observed when proteins are unfolded, misfolded, or unable to oligomerize. CNX and CRT bind most avidly to monoglucosylated forms of the N-linked core structure and promote folding, delay oligomerization, and prevent degradation of some glycoproteins (52). In contrast, interaction with CNX correlates with intracellular degradation of some glycoproteins (53). Removal of the third glucose from the oligosaccharide core structure correlates with release from CNX and CRT and transport to the Golgi apparatus. It is proposed that the selective binding of unfolded glycoproteins to CRT and CNX is mediated by reglucosylation of the deglucosylated N-linked oligosaccharide. This reglucosylation activity is performed by a UDP-glucose:glycoprotein glucosyltransferase (UGT). The activity of the UGT is activated by unfolded protein (54). Thereby, only unfolded, mutant, or unassembled proteins are subject to reglucosylation. Reglucosylated proteins can rebind CNX and/or CRT and in this manner; unfolded proteins are retained in the ER through a cycle of CNX/CRT interaction, glucosidase II activity, and UGT activity. Inhibition of glucose trimming by inhibitors of glucosidase I and II, such as deoxynojirimycin or castanospermine, can inhibit this cycle and the secretion of some proteins. Because glucosidase inhibitors are fairly nontoxic and can inhibit the secretion of HIV envelope glycoprotein gp160, they have been used in the treatment of AIDS (55). Subsequent to glucose trimming in the ER, at least one α1-2-linked mannose is removed by an ER α1-2-mannosidase prior to transport out of the ER. Transit out of the ER is the rate-limiting step in secretion for the majority of proteins and may vary from 15 min to days, depending upon the rate at which a polypeptide attains a properly folded conformation.

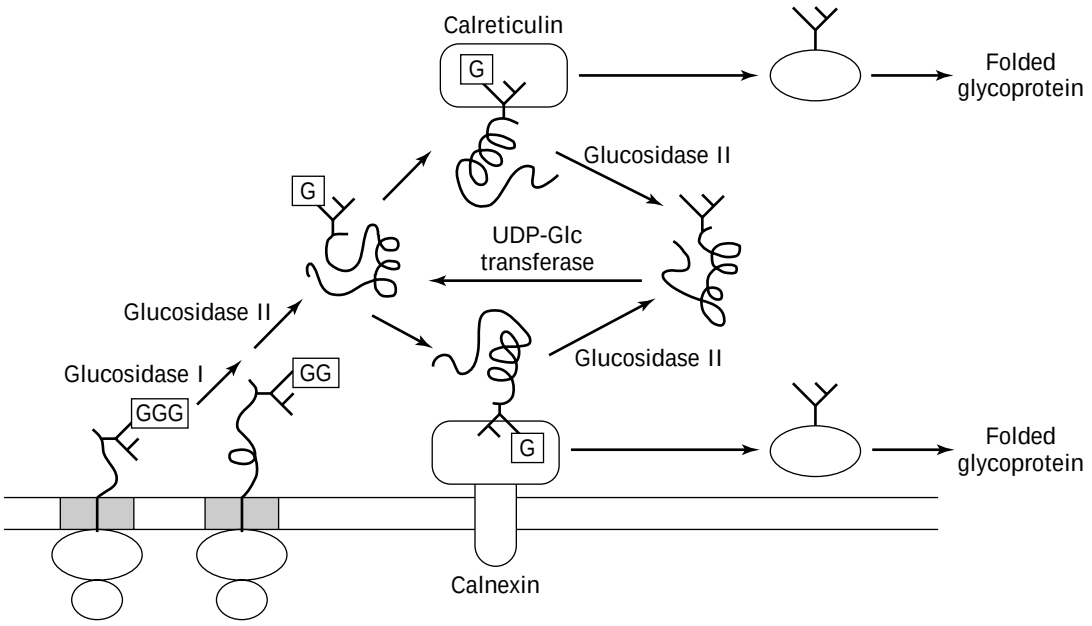


Figure 3. Role of calnexin/calreticulin and UDP-glucose:glycoprotein glucosyltransferase in the recognition and retention of unfolded glycoproteins in the ER. As a protein is cotranslationally translocated into the lumen of the ER, high-mannose core oligosaccharides are attached to asparargine residues. Subsequently, the glucose residues on the core-N-linked oligosaccharides are trimmed by the action of glucosidases I and II to expose a monoglucosylated core oligosaccharide that is a ligand for calnexin (a transmembrane protein) and/or calreticulin (a luminal ER protein) binding. After further action by glucosidase II, the glycoprotein is released from calnexin and calreticulin interaction. Unfolded glycoproteins are recognized as a substrate for the UDP-glucose:glycoprotein glucosyltransferase to form a monoglucosylated ligand to promote another round of calnexin and/or calreticulin binding. Folded glycoproteins bypass reglucosylation and are subsequently transported to the Golgi compartment. *Source*: Adapted from Ref. 44, with permission.

A sequential series of additional carbohydrate modifications occur that are separated spatially and temporarily as the protein transits through the *cis*, *medial*, and *trans* Golgi compartments. Mannose residues are removed by Golgi mannosidases I and II, and then *N*-acetylglucosamine, fucose, galactose, and sialic acid residues are added. These reactions are catalyzed by specific glycosyltransferases that modify the high-mannose carbohydrate to complex forms. Also within the Golgi apparatus, O-linked oligosaccharides are attached to the hydroxyl of serine or threonine residues through an O-glycosidic bond to *N*-acetylgalactosamine. Serine and threonine residues subject to glycosylation are frequently clustered together and contain an increased frequency of proline residues in the region, especially at positions -1 and $+3$, relative to the glycosylated residue (56). Galactose, fucose, and sialic acid are frequently attached to the serine/threonine-linked *N*-acetylgalactosamine. O-glycosylation occurs in the Golgi complex concomitant with complex processing of N-linked oligosaccharides.

The role of N- and O-linked oligosaccharide addition in glycoprotein function can be studied by using specific enzymatic inhibitors, glycosidases, and mutant cell lines (24). N-linked glycosylation is most easily evaluated by using specific glycosidases. The acquisition of resistance to endo-β-*N*-acetylglucosaminidase (Endo-H), which cleaves high-mannose and some hybrid-type oligosaccharides at the GlcNAcβ1-4GlcNAc linkage to leave a single GlcNAc residue attached to the asparagine (57), is frequently used to monitor movement of the protein from the ER to the medial Golgi apparatus. Resistance to Endo-H occurs in the medial Golgi following action of GlcNAc transferase I and mannosidase II. Peptide-*N*4-(*N*-acetyl-β-glucosaminyl)-asparagine amidase (N-gly) removes all N-linked oligosaccharides, regardless of the complexity of their structure, to leave a free aspartic acid residue (58). Also of use are a number of inhibitors that are specific for selective steps in the N-linked oligosaccharide processing pathway (see Fig. 2). In addition, a number of cell lines have been isolated that are defective in specific steps in oligosaccharide processing (59).

There are no specific inhibitors that can be used to study the requirement for O-linked glycosylation. It is possible to inhibit O-glycosylation, as well as complex modification of N-linked oligosaccharides, by depletion of the divalent metal ion manganese from the secretory pathway, and this may be used as an indication of whether O-linked sugars are present on the polypeptide (60). Of greater utility is a mutant cell line, ldlD, that is deficient in the UDP-galactose and UDP-*N*-acetylgalactosamine-4-epimerase and cannot synthesize Gal or GalNAc under normal growth conditions in the presence of glucose (61,62).

Finally, it is possible to enzymatically remove O-linked glycans with endo-α-*N*-acetyl-D-galactosaminidase (*O*-glycanase) that cleaves the Galβ1-3GalNAc disaccharide unit linked to serine or threonine residues (63). Sialic acid residues on the Gal or GalNAc will inhibit *O*-glycanase cleavage and therefore need to be removed by prior digestion with neuraminidase.

Glycosylphosphatidylinositol Anchor Addition

The addition of glycosylphosphatidylinositol (GPI) to many cell-surface-associated proteins occurs in order to anchor the protein in the lipid bilayer. The addition of GPI anchors occurs in the ER. A signal in the carboxy terminus of the polypeptide substrate contains a hydrophobic sequence of 17–30 amino acids followed by a spacer of 5–10 amino acids and a cleavage/attachment site. The attachment site must be an amino acid with a small side chain. The reaction is catalyzed by GPI:protein transamidase. This activity is defective in individuals with paraxysmal nocturnal hemoglobinuria. This disease results from a defect in the cell surface expression of inhibitors of complement-mediated cell lysis (such as decay accelerating factor). As a consequence, hemolytic anemia occurs because blood cells are unprotected from complement-mediated lysis.

Gamma-Carboxylation of Glutamic Acid Residues

The precursor of the vitamin K–dependent coagulation factors and a number of bone proteins and growth factors contain a propeptide that directs γ-carboxylation of up to 12 glutamic acid residues at the amino terminus of the mature protein (64). The propeptides of these proteins exhibit amino acid conservation within the γ-carboxylase recognition site and the site for cleavage of the propeptide. NMR structural analysis identified that the propeptide is an amphipathic α-helix with the carboxylase recognition site N-terminal to the helix (65).

The vitamin K–dependent γ-glutamyl carboxylase enzyme converts glutamate residues to γ-carboxyglutamate Gla residues. In the presence of CO_2, O_2, and vitamin K hydroquinone (KH_2) the enzyme is able to carboxylate a peptide containing glutamic acid residues, yielding a Gla-containing peptide, vitamin K epoxide, and H_2O (66). The vitamin K epoxide formed is subsequently reduced to regenerate KH_2 by either a thiol or the enzyme vitamin K epoxide reductase. The bovine (67) and human (68) cDNAs encoding the vitamin K–dependent γ-glutamyl carboxylase were isolated and demonstrate that the protein is a single-chain polypeptide of 94,000 kDa that spans the membrane 3 to 5 times. The overexpressed protein product directed increased carboxylation activity in vitro using isolated microsomes from transfected mammalian or insect cells and a synthetic peptide substrate (67,69). The ability to express γ-carboxylase activity from the cloned cDNA has permitted identification of the propeptide binding site between residues 50 and 225 (70) and the γ-carboxylase active site for glutamate binding within 218 residues of the amino terminal (71), whereas the vitamin K–reactive site is in the carboxy terminus of the protein (72). Vitamin K epoxidase activity is also catalyzed by the enzyme, and the carboxy terminus of the enzyme is required for this activity.

High-level expression of the vitamin K–dependent plasma proteins in transfected mammalian cells is limited by the ability of the mammalian host cell to efficiently perform γ-carboxylation of amino-terminal glutamic acid residues and also to efficiently cleave the propeptide (73,74). Analysis of factor IX expressed in CHO cells revealed that the protein had a much lower specific activity compared with the natural human plasma–derived protein. The reduced specific activity was attributed to both the limited ability of CHO cells to cleave the propeptide of factor IX and its ability to efficiently perform γ-carboxylation (67,73). Overexpression of the γ-carboxylase cDNA yielded increased activity in microsomal fractions when measured in vitro using small substrates (67). However, overexpression of the γ-carboxylase did not improve γ-carboxylation of factor IX when coexpressed in transfected mammalian cells (67). These results suggest that the amount of carboxylase protein is not a limiting factor to direct vitamin K–dependent γ-carboxylation in vivo. Several possibilities exist for the inability of the overexpressed γ-carboxylase to improve γ-carboxylation in vivo. First, the overexpressed γ-carboxylase may be mislocalized within the secretory pathway. It is possible that another protein, such as a protein chaperone, may be required to utilize a more complex protein substrate as opposed to a small peptide substrate. It is possible that another cofactor, possibly reduced vitamin K, is limiting for factor IX carboxylation in vivo. Further information on the mechanism of γ-carboxylation reaction in vivo is required in order to elucidate the rate-limiting step for γ-carboxylation in vivo.

β-Hydroxylation of Amino Acids: Aspartic acid, Asparagine, Lysine, and Proline

Proline and lysine residues are hydroxylated in procollagen by prolyl-3-hydroxylase and prolyl-4-hydroxylase. Prolyl-4-hydroxylase acts on prolines if the amino acid sequence is G-X-P, where X is any amino acid. This reaction requires O_2, Fe^{2+}, ascorbic acid, and α-ketoglutarate. This modification is important for stability of the collagen triple helix. Vitamin C deficiency results in scurvey because of reduced proline hydroxylation to form a stable collagen triplex.

A number of proteins that contain epidermal growth factor (EGF) domains, such as blood coagulation factor IX, have an aspartic acid or asparagine that is β-hydroxylated (75). This modification occurs by posttranslational hydroxylation of aspartic acid and/or asparagine within the ER. The β-hydroxylase has been molecularly cloned and characterized to some extent in recent years (76,77). β-Hydroxylation does not require the propeptide, vitamin K, or concomitant γ-carboxylation. A consensus β-hydroxylation site within EGF domains (Cys-X-Asp/Asn-X-X-X-X-Phe/Tyr-X-Cys-X-Cys) was proposed (75). Hydroxylation of both aspartic acid and asparagine is catalyzed by aspartyl β-hydroxylase, requires 2-ketoglutarate and Fe^{2+} (78,79), and is inhibited by agents that inhibit 2-ketoglutarate-dependent dioxygenases (80). It is interesting that only 0.3 mol/mol of plasma factor IX is modified by β-hydroxylation at Asp64, and this same amount of β-hydroxylation occurs

in recombinant factor IX expressed at high levels in CHO cells (80). Thus, the low efficiency of β-hydroxylation appears to be a consequence of the factor IX polypeptide backbone and independent of the cell type used for expression.

Sulfation of Tyrosine Residues

Sulfate addition to tyrosine as an *O*4-sulfate ester is a common postranslational modification of secretory proteins that occurs in the trans Golgi apparatus (81) and is mediated by tyrosylprotein sulfotransferase that utilizes the activated sulfate donor 3′-phosphoadenosine 5′-phosphosulfate (PAPS) (82,83). Although the enzyme has been biochemically characterized (83), to date the enzyme has not been molecularly cloned. This modification occurs on many secretory proteins including a number of proteins that interact with thrombin, such as hirudin (84), fibrinogen (85), heparin cofactor II (86), α2-antiplasmin (87), vitronectin (88), bovine factor X (89), factor V, and factor VIII (90–92). Comparison of all known tyrosine sulfation sites yielded a consensus sequence that is primarily characterized by a large number of acidic amino acid residues. The consensus tyrosine sulfation site has three aspartic acid and/or glutamic acid residues near the sulfated tyrosine (± 5 amino acids), a turn-inducing residue within $+7/-7$, and the absence of cysteine or N-linked glycosylation sites within 7 residues of the sulfated tyrosine (81). Although tyrosine sulfation modification is found on many proteins that transit the secretory apparatus, there are few examples where this modification is required for secretion or functional activity of the molecule; for example, tyrosine sulfation in the hormone cholecystokinin is required for its biological activity (93). However, tyrosine sulfation can modulate the biological activity, binding affinities, and secretion of specific proteins (81). For example, tyrosine sulfation at the carboxy terminus of hirudin increases its binding affinity to the anion binding exosite of thrombin (94,95).

The most direct method to measure sulfation of a particular protein is to measure incorporation of [^{35}S]-sulfuric acid into protein as it is synthesized in cultured cells as described in Ref. 24. [^{35}S]-Sulfuric acid will not be incorporated in cysteine. Sodium chlorate is an inhibitor of ATP sulfurylase, the first of two enzymes involved in the synthesis of 3′-phosphoadenosine 5′-phosphosulfate (PAPS). PAPS is the donor for sulfation of both tyrosine and carbohydrate residues in intact cells (96). Analysis of proteins synthesized in the presence of sodium chlorate can provide a useful means to study the role of tyrosine sulfation.

Proteolytic Processing of Precursor Polypeptides

Early studies demonstrated that the budding yeast *Saccharomyces cerevisiae* gene product Kex2 could cleave mammalian precursor polypeptides after paired basic residues (97), suggesting that a human homologue of Kex2 may be the protease required for propeptide processing. Kex2 is a membrane-bound Ca^{2+}-dependent subtilisin-like serine protease that cleaves substrates, such as alpha mating factor, within the *trans*-Golgi compartment (98,99). A computer search identified a human homologue of Kex2 upstream of the c-*fes*/*fps* protooncogene and was subsequently named furin (fes/feps upstream coding region) or PACE, an acronym for "paired basic amino acid cleaving enzyme." The cDNA was subsequently cloned and shown to encode a protein that could cleave pro-vWF (100,101), pro–nerve growth factor (102), proalbumin (103), complement pro-C3 (104), and pro–factor IX (73) after pairs of basic amino acids. PCR-based cloning strategies subsequently identified over a half dozen members of this subtilisin-like serine protease family (105). Expression of some of these enzymes (PC1/PC3, PC2, and PC4) is restricted to neuroendocrine tissues, and the activities are likely responsible for processing of neuropeptides and endocrine hormones (106,107). In contrast, furin/PACE and PACE4 are ubiquitously expressed, but to a greater extent in the hepatocyte (108,109), and are likely candidates for processing of many other proteins including coagulation factors.

Propeptide cleavage occurs in the trans Golgi compartment just prior to secretion from the cell. The activity of the propeptide processing enzymes is uniquely regulated where a calcium-dependent induced autocatalytic activation occurs within the trans Golgi compartment to activate these proteases to their active form (110–113). In the process of activation, a propeptide is cleaved away from the zymogen. The specific localization of propeptide processing to the trans Golgi compartment ensures that the propeptide is associated with the mature polypeptide as proteins transit the secretory compartment.

Characterization of the amino acid requirements around the propeptide cleavage site has identified that both the P1 and P4 arginine are important for efficient processing mediated by furin/PACE and PACE4 (114,115). Cotransfection experiments were performed to test whether these enzymes could process pro–factor IX and or pro-vWF. Whereas both furin/PACE and PACE4 were effective at enhancing pro-vWF processing, only furin/PACE was capable of improving pro–factor IX processing. Thus, it appears that different members of this class of enzymes recognize and process overlapping sets of substrates. Overexpression of furin/PACE in transfected cells (103) as well as in transgenic animals (116) improves the ability to yield fully processed proteins. Recombinant factor IX is produced by coexpression with furin/PACE to ensure complete processing of the propeptide.

VESICULAR PROTEIN TRANSPORT

Transport vesicles are the intermediates for the trafficking of proteins between compartments of the secretory pathway (117). The principal mechanism of transport vesicle formation is through cytosolic coat proteins that bind to specific sites on the membrane (118). Interactions between these coat proteins and the cytosolic tails of membrane proteins play a fundamental role in sorting cargo into vesicles (119,120). Binding of coat proteins to the membrane induces the configuration of the membrane into a coated bud. Membrane fission at the neck of this bud releases the vesicle. The most well-characterized coat protein is the clathrin coat that is responsible for endocytic vesicle formation

at the plasma membrane. Two additional types of coat components mediate vesicle formation and fusion within the secretory pathway. COP II (*co*at *p*rotein) components are required for the formation of vesicles that export cargo from the ER. COP I components were identified as coat proteins that form Golgi-derived transport vesicles (121,122) (Fig. 4).

Protein transport through the secretory pathway begins at the ER. Secretory proteins leave the ER in vesicles that bud from a specialized region of the ER. Protein transport between the ER and the intermediate compartment occurs in both directions. Retrograde transport from the intermediate compartment to the ER was first demonstrated for ER-resident proteins. If such proteins are erroneously incorporated into transport vesicles and leave the ER, they are efficiently recycled by retrograde transport. COP I–coated vesicles mediate the recycling of ER-resident proteins from intermediate compartments back to the ER, and COP II–coated vesicles mediate the export from the ER. The interaction of COP I and COP II components with the cytosolic tail domains of secretory transmembrane and ER-resident transmembrane proteins is likely to play an important role in sorting cargo into transport vesicles (119,120). Recently studies have used green fluorescent protein fused to vesicular stomatitis virus glycoprotein to visualize protein trafficking between the ER and Golgi compartments in living cells. The results demonstrate that the COP II–coated vesicles that are derived from the ER membrane have their coat proteins replaced by COP I and then move in a manner that depends on microtubules to the Golgi complex (Fig. 4). In addition, the vesicles are more elongated (appearing more like tubules) and are juxtaposed to microtubules that probably guide their movement (122,123).

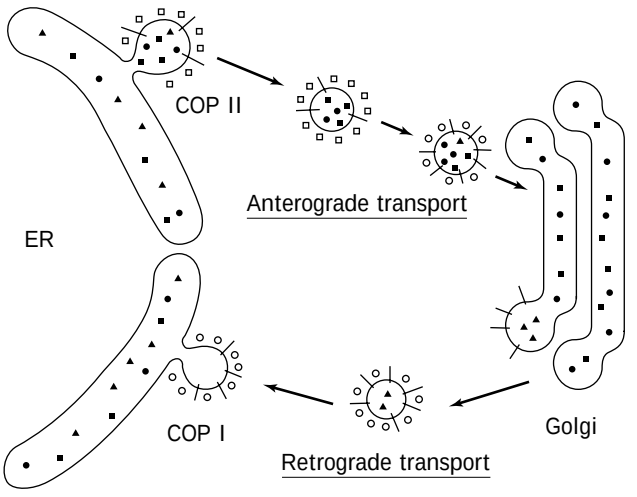


Figure 4. Proposed roles of COP I and COP II in vesicle traffic between the ER and cis Golgi. COP II vesicles transport normal cargo, some escaped ER resident proteins, and some anterograde targeting membrane proteins from the ER to the intermediate compartment where the COP II coat is exchanged for a COP I coat. The COP I–coated vesicle then fuses with the cis Golgi compartment. COP I vesicles are also shown mediating retrograde transport to the ER for escaped ER proteins.

Secretory proteins pass through the Golgi in the cis to trans direction. As proteins pass through the Golgi compartments they are modified by sequence of glycosylation and deglycosylation steps (124). Best understood is the modification of N-linked glycans in the Golgi. Initial trimming of mannose residues is followed by attachment of *N*-acetylglucosamine (GlcNAc), which is followed by more trimming of mannose residues and attachment of more GlcNAc. Mannose trimming and GlcNAc addition occur in cis and medial Golgi compartments. Although each Golgi subcompartment differs from its neighboring compartments in the specific set of enzymes contained within it, this separation is not complete. The same enzymes may even have different localization in different cell types. Golgi enzymes are not restricted to just one cistern. When the distributions of two different Golgi enzymes were determined relative to each other in the same samples, both were found spread out over several cisternae. The maximum concentration of each enzyme was in different in different compartments, but their distribution showed considerable overlap. Furthermore, Golgi enzymes seem to be in flux. Experiments with yeast demonstrated that a Golgi enzyme that is predominantly localized in the early Golgi was modified by late-acting Golgi enzymes (125). In a similar study using mammalian cells it was shown that an enzyme that was localized in the medial Golgi could be modified by an enzyme in the cis Golgi, even after the medial Golgi enzyme had reached its localization (126). These results indicate that Golgi enzymes must be efficiently recycled from late to early Golgi compartments and move again forward with the secretory flow. Golgi enzymes are not immobile in Golgi cisternae, but rather are freely diffusable.

Two models are proposed to explain protein transport through the Golgi: the vesicular transport model and the cisternal maturation model (127–129). The vesicular transport model assumes several distinct subcompartments that exchange cargo by vesicular transport in the anterograde and retrograde direction. The cisternal maturation model proposes that cisternae form de novo by fusion of ER-derived transport vesicles. These ER-derived transport intermediates would acquire Golgi enzymes by fusion with Golgi-derived retrograde transport vesicles from the cis-most compartment of the Golgi stack. At the same time, resident Golgi enzymes would move by retrograde transport in the trans to cis direction, from cisternae that formed earlier to more recently formed cisternae. Ultimately, cisternae would emerge on the trans face of the Golgi stack and disperse into vesicles of the constitutive and regulated secretory pathways and into transport vesicles to endosomes and lysosomes. Both models were first formulated shortly after the Golgi apparatus was characterized by electron microscopy, and which one best describes the Golgi has been controversial ever since.

Fusion of the vesicle membrane with its target membrane requires the cytosolic protein NSF (NEM-sensitive factor for fusion). NSF interacts with another cytosolic protein, SNAP (soluble NSF attachment protein), and SNARE (SNAP receptor) membrane proteins (130). It is the interaction of vesicle SNARE (v-SNARE) with target SNARE (t-SNARE) that is thought to play an important role in

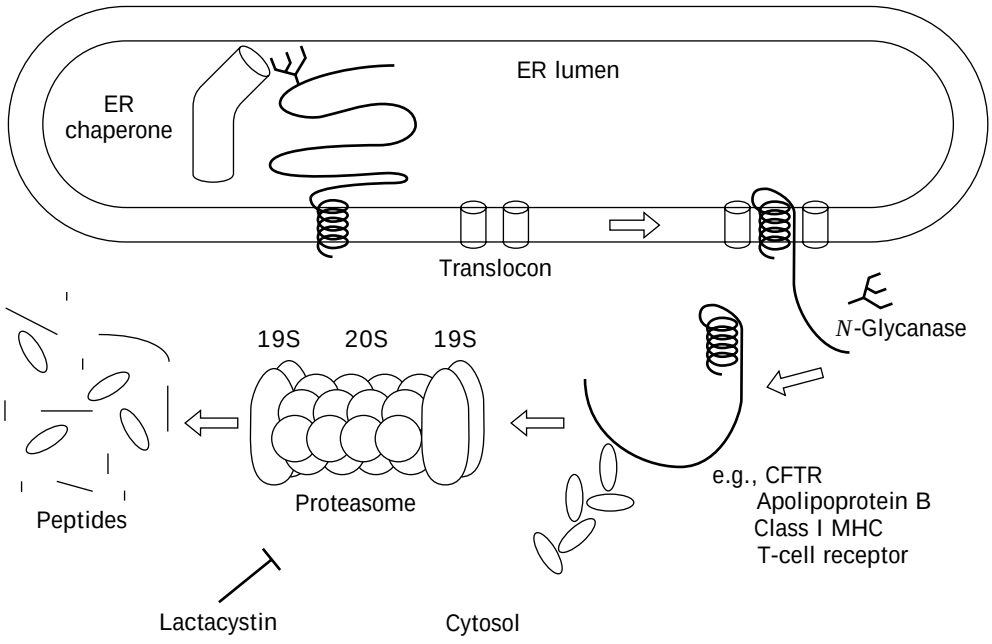


Figure 5. Degradation of ER-associated proteins. A putative ER chaperone facilitates association of a protein, with the Sec61 translocon. Reverse translocation from the lumen to the cytosol occurs with subsequent deglycosylation by *N*-glycanase. The protein may be subject to modification by ubiquitination and eventual targeting to the proteasome. Soluble proteins of the ER may also be degraded by reverse translocation through the Sec61 translocon, or alternatively they may associate with transmembrane proteins, such as CNX, that mediate their degradation. *Source*: Adapted from Ref. 141, with permission.

controlling the specificity of the vesicle fusion reaction. The precise role of NSF in this reaction is unclear. NSF action disrupts a v-SNARE/t-SNARE complex (131). This disruption could be what mediates the rearrangement of SNAREs that is required during fusion (132). These results therefore support the hypothesis that NSF directly mediates vesicle fusion. However, in yeast NSF mutants, vesicles accumulate in the cytosol, and this observation lends support to a model in which NSF is required for vesicle docking (133). Consistent with this hypothesis is that although yeast vacuole fusion is NSF dependent, the NSF action in this process precedes the docking of the membranes (134). These results indicate that NSF acts by activating the SNAREs for vesicle docking, rather than being involved in fusion itself. Small GTP-binding proteins of the rab family, as well as other proteins such as p115/Uso1p, are likely to be involved in further control of vesicle docking and fusion (135–138).

Mammalian cells constitutively transport proteins within vesicles from the ER to the plasma membrane, where they fuse through exocytosis. There are also specialized secretory cells, for example neuroendocrine cells, that store granules and release them upon appropriate stimulation through a regulated secretory pathway. Different proteins are sorted to the regulated secretory granules in different cell types through a common mechanism. The sorting process involves selective protein aggregation into clathrin-coated vesicles at the trans Golgi network. The aggregation occurs at pH 6.5 with 1 mM Ca^{2+}.

PROTEIN RETENTION AND DEGRADATION IN THE ER

A protein within the ER has several possible destinations that include transport to the Golgi compartment, degradation, and retention in the ER. What determines the fate of any particular polypeptide? For resident proteins of the ER, there are two primary mechanisms that ensure that proteins do not exit the ER compartment (139). Soluble luminal proteins of the ER, such as BiP and PDI, contain a carboxy-terminal tetrapeptide, with the consensus sequence KDEL, that is necessary and sufficient for preventing secretion of the polypeptide. A KDEL receptor, that is, a transmembrane-spanning protein, binds KDEL-containing proteins that leave the ER and retrieves them to the ER from the intermediate compartment. For transmembrane-resident proteins of the ER, there is a signal within the cytosolic domain that mediates retrieval from the intermediate compartment by COP proteins. Type I transmembrane-resident ER proteins contain a dilysine motif in the carboxy terminus (i.e., cytosolic tail). For type II transmembrane proteins, a diarginine motif in the amino terminus mediates the interaction with COP proteins. COP I vesicles then retrieve these proteins to the ER from the intermediate compartment.

As mentioned earlier, for many years investigators searched for signals that direct protein transport out of the ER. Although it is now known that specific short sequences exist that facilitate export from the ER, it appears now that the majority of polypeptides have retention signals

that prevent their export. As a polypeptide folds, these retention signals become buried within the polypeptide, and thus, the polypeptide can transit to the Golgi compartment. A variety of polypeptides that recognize and bind unfolded proteins have been identified, including BiP/GRP78, CNX, and CRT, which were discussed earlier. These protein chaperones play a fundamental role in preventing unfolded proteins from exiting the ER compartment.

Proteins that are unable to attain their native conformation are targeted for degradation. A nonlysosomal pathway has been identified in the degradation of these proteins (140,141). The discovery that lactacystin, a fungal metabolite that specifically binds and inhibits the cytosolic 26S proteasome, inhibits the degradation of many ER-associated proteins pointed to the involvement of the cytosolic proteasome in the degradation of ER-localized proteins. There are now a number of examples of ER-associated proteins, such as the cystic fibrosis transmembrane conductance regulator, the T-cell receptor, factor VIII, class I MHC molecules, and so forth, that are degraded by this pathway (53,142–146). Insight into the mechanism of degradation was obtained by studying how human cytomegalovirus (HCMV) evades the immune system (Fig. 5). HCMV encodes two gene products, US2 and US11, that are able to dislocate ER MHC class I proteins to the cytosol (146,147). This dislocation occurs through the Sec61 protein conducting channel (148). Once in the cytosol, the class I molecules are deglycosylated by an *N*-glycanase and are then polyubiquitinated. The ubiquitinated protein is then degraded by the cytosolic 26S proteasome. How are unfolded proteins identified to be degraded? The answer to this question is unknown, but present data suggest that CNX may be involved in this quality control decision. Proteins that are destined to be degraded are bound to CNX but not to CRT (53,143,149,150). CNX is a transmembrane protein that is in close proximity to the Sec61p translocon. It is speculated that CNX may mediate the transfer of unfolded proteins in the lumen of the ER to the proteasome.

ACKNOWLEDGMENTS

I thank Joe Nowak and Renee Strauss for assistance in preparation of this article.

BIBLIOGRAPHY

1. H. Nielsen, J. Engelbrecht, S. Brunak, and G. von Heijne, *Protein Eng.* **10**, 1–6 (1997).
2. P. Walter and A.E. Johnson, *Annu. Rev. Cell Biol.* **10**, 87–119 (1994).
3. A.K. Corsi and R. Schekman, *J. Biol. Chem.* **271**, 30299–39302 (1996).
4. P.J. Rapiejko and R. Gilmore, *J. Cell Biol.* **117**, 493–503 (1992).
5. P.J. Rapiejko and R. Gilmore, *Mol. Biol. Cell* **5**, 887–897 (1994).
6. J.D. Miller, H. Wilhelm, L. Gierasch, R. Gilmore, and P. Walter, *Nature* **366**, 351–354 (1993).
7. B. Wiedmann, H. Sakai, T.A. Davis, and M. Wiedmann, *Nature* **370**, 434–440 (1994).
8. B. Lauring, G. Kreibich, and M. Weidmann, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9435–9439 (1995).
9. B. Lauring, H. Sakai, G. Kreibich, and M. Wiedmann, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5411–5415 (1995).
10. S. Voigt, B. Jungnickel, E. Hartmann, and T.A. Rapoport, *J. Cell Biol.* **134**, 25–35 (1996).
11. C.J. Stirling, J. Rothblatt, M. Hosobuchi, R. Deshaies, and R. Schekman, *Mol. Biol. Cell* **3**, 129–142 (1992).
12. D. Gorlich and T.A. Rapoport, *Cell* **75**, 615–630 (1993).
13. T.A. Rapoport, *Science* **258**, 931–936 (1992).
14. C.V. Nicchitta and G. Blobel, *Cell* **73**, 989–998 (1993).
15. J.A. Morris, A.J. Dorner, C.A. Edwards, L.M. Hendershot, and R.J. Kaufman, *J. Biol. Chem.* **272**, 4327–4334 (1997).
16. S. High and B. Dobberstein, *Curr. Opin. Cell Biol.* **4**, 581–586 (1992).
17. B.D. Hamman, L.M. Hendershot, and A.E. Johnson, *Cell* **92**, 747–758 (1998).
18. B. Martoglio, M.W. Hofmann, J. Brunner, and B. Dobberstein, *Cell* **81**, 207–214 (1995).
19. R.E. Dalbey, M.O. Lively, S. Bron, and J.M. van Dijl, *Protein Sci.* **6**, 1129–1138 (1997).
20. K.U. Kalies, T.A. Rapoport, and E. Hartmann, *J. Cell Biol.* **141**, 887–894 (1998).
21. A.J. Dorner, L.C. Wasley, and R.J. Kaufman, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 7429–7432 (1990).
22. P. Mayinger, V.A. Bankaitis, and D.I. Meyer, *J. Cell Biol.* **131**, 1377–1386 (1995).
23. A.J. Dorner and R.J. Kaufman, *Biologicals* **22**, 103–112 (1994).
24. A.J. Dorner and R.J. Kaufman, *Methods Enzymol.* **185**, 577–596 (1990).
25. C.B. Anfinsen, E. Haber, M. Sela, and F.H. White, *Proc. Natl. Acad. Sci. U.S.A.* **47**, 1309–1314 (1961).
26. S. Munro and H.R.B. Pelham, *Cell* **46**, 291–300 (1986).
27. G.C. Flynn, T.G. Chappell, and J.E. Rothman, *Science* **245**, 385–390 (1989).
28. Y. Kozutsumi, M. Segal, K. Normington, M.J. Gething, and J. Sambrook, *Nature* **332**, 462–464 (1988).
29. A.S. Lee, *Trends Biochem. Sci.* **12**, 20–24 (1987).
30. M.-J. Gething and J. Sambrook, *Nature* **355**, 33–45 (1992).
31. J.E. Rothman, *Cell* **59**, 591–601 (1989).
32. D.T.W. Ng, R.E. Randall, and R.A. Lamb, *J. Cell Biol.* **109**, 3273–3289 (1989).
33. A.J. Dorner, D.G. Bole, and R.J. Kaufman, *J. Cell Biol.* **105**, 2665–2674 (1987).
34. T. Hai, F. Liu, E.A. Allegretto, M. Karin, and M.R. Green, *Genes Dev.* **2**, 1216–1226 (1988).
35. P. Blount and J.P. Merlie, *J. Cell Biol.* **113**, 1125–1132 (1991).
36. I. Braakman, J. Helenius, and A. Helenius, *Nature* **356**, 260–262 (1992).
37. L. Hendershot, D. Bole, G. Kohler, and J.F. Kearney, *J. Cell Biol.* **104**, 761–767 (1987).
38. A.J. Dorner, M.G. Krane, and R.J. Kaufman, *Mol. Cell. Biol.* **8**, 4063–4070 (1988).
39. A.J. Dorner, L.C. Wasley, and R.J. Kaufman, *EMBO J.* **11**, 1563–1571 (1992).

40. B.A. Pollok, R. Anker, P. Eldridge, L. Hendershot, and D. Levitt, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 9199–9203 (1987).
41. A. Zapun, T.E. Creighton, P.J.E. Rowling, and R.B. Freedman, *Proteins* **14**, 10–15 (1992).
42. C. Hwang, A.J. Sinskey, and H.F. Lodish, *Science* **257**, 1496–1502 (1992).
43. R. Noiva and W.J. Lennarz, *J. Biol. Chem.* **267**, 3553–3556 (1992).
44. A. Helenius, *Mol. Biol. Cell* **5**, 253–265 (1994).
45. S. Silberstein, P.G. Collins, D.J. Kelleher, P.J. Rapiejko, and R. Gilmore, *J. Cell Biol.* **128**, 525–536 (1995).
46. J.L. Guan, C.E. Machamer, and J.K. Rose, *Cell* **42**, 489–496 (1985).
47. D.J. Kelleher, G. Kreibich, and R. Gilmore, *Cell* **69**, 55–65 (1992).
48. Y. Gavel and G. von Heijne, *Protein Eng.* **3**, 433–442 (1990).
49. D.N. Hebert, B. Foelimer, and A. Helenius, *Cell* **81**, 425–433 (1995).
50. J.J. Bergeron, M.B. Brenner, D.Y. Thomas, and D.B. Williams, *Trends Biochem. Sci.* **19**, 124–128 (1994).
51. K.H. Krause and M. Michalak, *Cell* **88**, 439–443 (1997).
52. D.N. Hebert, B. Foellmer, and A. Helenius, *EMBO J.* **15**(12), 2961–2968 (1996).
53. S.W. Pipe, J.A. Morris, J. Shah, and R.J. Kaufman, *J. Biol. Chem.* **273**, 8537–8544 (1998).
54. S.E. Trombetta and A.J. Parodi, *J. Biol. Chem.* **267**, 9236–9240 (1992).
55. A.D. Elbein, *Semin. Cell Biol.* **2**, 309–317 (1991).
56. I.B. Wilson, Y. Gavel, and G. von Heijne, *Biochem. J.* **275**, 529–534 (1991).
57. A.L. Tarentino and F. Maley, *J. Biol. Chem.* **249**, 811–817 (1974).
58. A.L. Tarentino, C.M. Gomez, and T.H. Plummer, Jr., *Biochemistry* **24**, 4665–4671 (1985).
59. P. Stanley, in M.M. Gottesman ed., *Molecular Cell Genetics*, Wiley, New York, 1985, pp. 745–772.
60. R.J. Kaufman, M. Swaroop, and P. Murtha-Riel, *Biochemistry* **33**, 9813–9819 (1994).
61. C.L. Wasley, P. Horgan, G. Timony, J. Stoudemier, M. Krieger, and R.J. Kaufman, *Blood* **77**, 2624–2632 (1991).
62. D.M. Kingsley, K.F. Kozarsky, L. Hobbie, and M. Krieger, *Cell* **44**, 749–759 (1986).
63. J. Umemoto, V.P. Bhavanandan, and E.A. Davidson, *J. Biol. Chem.* **252**, 8609–8614 (1977).
64. M.J. Jorgensen, A.B. Cantor, B.C. Furie, C.L. Brown, C.B. Shoemaker, and B. Furie, *Cell* **48**, 185–191 (1987).
65. D.G. Sanford, C. Kanagy, J.L. Sudmeier, B.C. Furie, B. Furie, and W.W. Bachovchin, *Biochemistry* **30**, 9835–9841 (1991).
66. J.W. Suttie, *FASEB J.* **7**, 445–452 (1993).
67. A. Rehemtulla, D.A. Roth, L.C. Wasley, A. Kuliopulos, C.T. Walsh, B. Furie, B.C. Furie, and R.J. Kaufman, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4611–4615 (1993).
68. S.M. Wu, W.F. Cheung, D. Frazier, and D.W. Stafford, *Science* **254**, 1634–1636 (1991).
69. D.A. Roth, A. Rehemtulla, R.J. Kaufman, C.T. Walsh, B. Furie, and B.C. Furie, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8372–8376 (1993).
70. M. Yamada, A. Kuliopulos, N.P. Nelson, D.A. Roth, B. Furie, B.C. Furie, and C.T. Walsh, *Biochemistry* **34**, 481–489 (1995).
71. A. Kuliopulos, N.P. Nelson, M. Yamada, C.T. Walsh, B. Furie, B.C. Furie, and D.A. Roth, *J. Biol. Chem.* **269**, 21364–21370 (1994).
72. D.A. Roth, M.L. Whirl, L.J. Velazquez-Estades, C.T. Walsh, B. Furie, and B.C. Furie, *J. Biol. Chem.* **270**, 5305–5311 (1995).
73. L.C. Wasley, A. Rehemtulla, J.A. Bristol, and R.J. Kaufman, *J. Biol. Chem.* **268**, 8458–8465 (1993).
74. R.J. Kaufman, L.C. Wasley, B.C. Furie, B. Furie, and C.B. Shoemaker, *J. Biol. Chem.* **261**, 9622–9628 (1986).
75. J. Stenflo, A. Lundwall, and B. Dahlback, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 368–372 (1987).
76. K. McGinnis, G.M. Ku, W.J. VanDusen, J. Fu, V. Garsky, A.M. Stern, and P.A. Friedman, *Biochemistry* **35**, 3957–3962 (1996).
77. S. Jia, W.J. VanDusen, R.E. Diehl, N.E. Kohl, R.A. Dixon, K.O. Elliston, A.M. Stern, and P.A. Friedman, *J. Biol. Chem.* **267**, 14322–14327 (1992).
78. R.S. Gronke, W.J. VanDusen, V.M. Garsky, J.W. Jacobs, M.K. Sardana, A.M. Stern, and P.A. Friedman, *Proc. Natl. Acad. Sci.* **86**, 3609–3613 (1989).
79. J. Stenflo, E. Holme, S. Lindstedt, N. Chandramouli, L.H. Huang, J.P. Tam, and R.B. Merrifield, *Proc. Natl. Acad. Sci.* **86**, 4440–4447 (1989).
80. C.K. Derian, W. VanDusen, C.T. Przysiecki, P.N. Walsh, K.L. Berkner, R.J. Kaufman, and P.A. Friedman, *J. Biol. Chem.* **264**, 6615–6618 (1989).
81. W.B. Huttner, *Annu. Rev. Physiol.* **50**, 363–376 (1988).
82. P.A. Baeuerle and W.B. Huttner, *J. Cell Biol.* **105**, 2655–2664 (1987).
83. C. Niehrs and W.B. Huttner, *EMBO J.* **9**, 35–42 (1990).
84. P.J. Braun, S. Dennis, J. Hofsteenge, and S.R. Stone, *Biochemistry* **27**, 6517–6522 (1988).
85. D.H. Farrell, E.R. Mulvihill, S.M. Huang, D.W. Chung, and E.W. Davie, *Biochemistry* **30**, 9414–9420 (1991).
86. G. Hortin, D.M. Tollefsen, and A.W. Strauss, *J. Biol. Chem.* **261**, 15827–15830 (1986).
87. G. Hortin, K.F. Fok, P.C. Toren, and A.W. Strauss, *J. Biol. Chem.* **262**, 3082–3085 (1987).
88. D. Jenne, A. Hille, K.K. Stanley, and W.B. Huttner, *Eur. J. Biochem.* **185**, 391–395 (1989).
89. T. Morita and C.M. Jackson, *J. Biol. Chem.* **261**, 4008–4014 (1986).
90. D.D. Pittman, J.H. Wang, and R.J. Kaufman, *Biochemistry* **31**, 3315–3323 (1992).
91. D.D. Pittman, K.N. Tomkinson, D. Michnick, U. Selighsohn, and R.J. Kaufman, *Biochemistry* **33**, 6952–6959 (1994).
92. D.A. Michnick, D.D. Pittman, R.J. Wise, and R.J. Kaufman, *J. Biol. Chem.* **269**, 20095–20102 (1994).
93. V. Mutt, in G.B.J. Glass ed., *Gastrointestinal Hormones* Raven Press, New York, 1980, pp. 169–221.
94. C. Niehrs, W.B. Huttner, D. Carvallo, and E. Degryse, *J. Biol. Chem.* **265**, 9314–9318 (1990).
95. T.J. Rydel, K.G. Ravichandran, A. Tulinsky, W. Bode, R. Huber, C. Roitsch, and J.W. Fenton, *Science* **249**, 277–280 (1990).
96. E. Friederich, H.J. Gritz, and W.B. Huttner, *J. Cell Biol.* **107**, 1655–1667 (1988).

97. D.C. Foster, R.D. Holly, C.A. Sprecher, K.M. Walker, and A.A. Kumar, *Biochemistry* **30**, 367–372 (1991).

98. I.C. Bathurst, S.O. Brennan, R.W. Carrell, L.S. Cousens, A.J. Brake, and P.J. Barr, *Science* **235**, 348–350 (1987).

99. D. Julius, A. Brake, L. Blair, R. Kunisawa, and J. Thorner, *Cell* **37**, 1075–1089 (1984).

100. R.J. Wise, P.J. Barr, P.A. Wong, M.C. Kiefer, A.J. Brake, and R.J. Kaufman, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9378–9382 (1990).

101. W.J. van de Ven, J. Voorberg, R. Fontijn, H. Pannekoek, A.M. van den Ouweland, H.L. van Duijnhoven, A.J. Roebroek, and R.J. Siezen, *Mol. Biol. Rep.* **14**, 265–275 (1990).

102. P.A. Bresnahan, R. Leduc, L. Thomas, J. Thorner, H.L. Gibson, A.J. Brake, P.J. Barr, and G. Thomas, *J. Cell Biol.* **111**, 2851–2859 (1990).

103. S.O. Brennan and R.J. Peach, *J. Biol. Chem.* **266**, 21504–21508 (1991).

104. Y. Misumi, K. Ohkubo, M. Sohda, N. Takami, K. Oda, and Y. Ikehara, *Biochem. Biophys. Res. Commun.* **171**, 236–242 (1990).

105. P.A. Halban and J.C. Irminger, *Biochem. J.* **299**, 1–18 (1994).

106. S. Benjannet, N. Rondeau, R. Day, M. Chretien, and N.G. Seidah, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 3564–3568 (1991).

107. L. Thomas, R. Leduc, B.A. Thorne, S.P. Smeekens, D.F. Steiner, and G. Thomas, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5297–5301 (1991).

108. A.J. Roebroek, J.A. Schalken, J.A. Leunissen, C. Onnekink, H.P. Bloemers, and W.J. van de Ven, *EMBO J.* **5**, 2197–2202 (1986).

109. M.C. Kiefer, J.E. Tucker, R. Joh, K.E. Landsberg, D. Saltman, and P.J. Barr, *DNA Cell Biol.* **10**, 757–769 (1991).

110. A. Rehemtulla, A.J. Dorner, and R.J. Kaufman, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8235–8239 (1992).

111. R. Leduc, S.S. Molloy, B.A. Thorne, and G. Thomas, *J. Biol. Chem.* **267**, 14304–14308 (1992).

112. E.D. Anderson, J.K. VanSlyke, C.D. Thulin, F. Jean, and G. Thomas, *EMBO J.* **16**, 1508–1518 (1997).

113. S.S. Molloy, L. Thomas, J.K. VanSlyke, P.E. Stenberg, and G. Thomas, *EMBO J.* **13**, 18–33 (1994).

114. A. Rehemtulla and R.J. Kaufman, *Blood* **9**, 2349–2355 (1992).

115. A. Rehemtulla, P.J. Barr, C.J. Rhodes, and R.J. Kaufman, *Biochemistry* **32**, 11586–11590 (1993).

116. R. Drews, R.K. Paleyanda, T.K. Lee, R.R. Chang, W.N. Drohan, A. Rehemtulla, and R.J. Kaufman, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10462–10466 (1995).

117. J.E. Rothman, *Protein Sci.* **5**, 185–194 (1996).

118. R. Schekman and L. Orci, *Science* **271**, 1526–1533 (1996).

119. M.J. Kuehn, J.M. Herrmann, and R. Schekman, *Nature* **391**, 187–190 (1998).

120. P. Cosson and F. Letoumeur, *Science* **263**, 1629–1631 (1994).

121. R. Schekman and I. Mellman, *Cell* **90**, 197–200 (1997).

122. S.J. Scales, R. Pepperkok, and T.E. Kreis, *Cell* **90**, 1137–1148 (1997).

123. J.F. Presley, N.B. Cole, T.A. Schroer, K.J. Hirschberg, K.J. Zaal, and J. Lippencott-Schwartz, *Nature* **389**, 81–85 (1997).

124. R. Kornfeld and S. Kornfeld, *Annu. Rev. Biochem.* **54**, 631–664 (1985).

125. J. Ting, S.K. Wooden, K.S. Kelleher, R. Kritz, R.J. Kaufman, and A.S. Lee, *Gene* **55**, 147–152 (1987).

126. R.J. Jenny, D.D. Pittman, J.J. Toole, R.W. Kriz, R.A. Aldape, R.M. Hewick, R.J. Kaufman, and K.G. Mann, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4846–4850 (1987).

127. S.I. Bannykh and W.E. Balch, *J. Cell. Biol.* **138**, 1–4 (1997).

128. B.S. Glick, T. Elston, and G. Oster, *FEBS* **414**, 177–181 (1998).

129. H.D. Love, C.C. Lin, C.S. Short, and J. Ostermann, *J. Cell Biol.* **140**, 541–551 (1998).

130. T. Sollner, S.W. Whiteheart, M. Brunner, H. Erdjument-Bromage, S. Geromanos, P. Tempst, and J.E. Rothman, *Nature* **362**, 318–324 (1993).

131. J.E. Rothman, and G. Warren, *Curr. Biol.* **4**, 220–233 (1994).

132. T. Sollner, M.K. Bennett, S.W. Whiteheart, R.H. Scheller, and J.E. Rothman, *Cell* **75**, 409–418 (1993).

133. C.A. Kaiser, and R. Schekman, *Cell* **61**, 723–733 (1990).

134. A. Mayer, W. Wickner, and A. Haas, *Cell* **85**, 83–94 (1996).

135. V. Rybin, O. Ullrich, M. Rubino, K. Alexandrov, I. Simon, C. Seabra, R. Goody, and M. Zerial, *Nature* **383**, 266–269 (1996).

136. V.V. Lupashin, and M.G. Waters, *Science* **276**, 1255–1258 (1997).

137. S.K. Sapperstein, V.V. Lupashin, H.D. Schmitt, and M.G. Waters, *J. Cell Biol.* **132**, 755–767 (1996).

138. N. Nakamura, M. Lowe, T.P. Levine, C. Rabouille, and G. Warren, *Cell* **89**, 445–455 (1997).

139. H.R. Pelham, *Curr. Opin. Cell Biol.* **7**, 530–535 (1995).

140. R.D. Klausner and R. Sitia, *Cell* **62**, 611–614 (1990).

141. R.R. Kopito, *Cell* **88**, 427–430 (1997).

142. C.L. Ward, S. Omura, and R.R. Kopito, *Cell* **83**, 121–127 (1995).

143. R. Halaban, E. Cheng, Y. Zhang, G. Moellmann, D. Hanlon, M. Michalak, V. Setaluri, and D.N. Hebert, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6210–6215 (1997).

144. D. Qu, J.H. Teckman, S. Omura, and D.H. Perlmutter, *J. Biol. Chem.* **271**, 22791–22795 (1996).

145. M.M. Hiller, A. Finger, M. Schweiger, and D.H. Wolf, *Science* **273**, 1725–1728 (1996).

146. E.J. Wiertz, T.R. Jones, L. Sun, M. Bogyo, H.J. Geuze, and H.L. Ploegh, *Cell* **84**, 769–779 (1996).

147. E.J. Wiertz, D. Tortorella, M. Bogyo, J. Yu, W. Mothes, T.R. Jones, T.A. Rapoport, and H.L. Ploegh, *Nature* **384**, 432–438 (1996).

148. M. Pilon, R. Schekman, and K. Romisch, *EMBO J.* **16**, 4540–4548 (1997).

149. K.S. Cannon, D.N. Hebert, and A. Helenius, *J. Biol. Chem.* **271**, 14280–14284 (1996).

150. W. Chen, J. Helenius, I. Braakman, and A. Helenius, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6229–6233 (1995).

See also PROTEIN SECRETION, *SACCHAROMYCES CEREVISIAE*.

SHEAR SENSITIVITY

YUSUF CHISTI
University of Almería
Almería, Spain

KEY WORDS

Air-lift bioreactors
Animal cell culture
Bioreactor hydrodynamics
Bubble columns
Cell damage
Shear damage
Shear protection
Shear rate
Shear stress

OUTLINE

INTRODUCTION

Susceptibility to hydrodynamic and mechanical shear forces affects performance of cultured cells of animals, plants, microalgae, and cyanobacteria. In addition, hydromechanical forces affect or otherwise damage some commercially relevant mycelial fungi, filamentous bacteria, microbial flocs, and biofilms. In specific cases, intense shear fields may also damage the larger bioactive molecules such as enzymes. This article details the shear sensitivity of the major types of biocatalysts and the approaches available for mitigating the damaging effects. Methods for estimating shear rate and shear stress are discussed for various configurations of bioreactors operated in process-relevant conditions.

Depending on its intensity, a shear field may be stimulatory, inhibitory, or outright destructive. Intense shear fields will disintegrate even the most robust of microorganisms (1). Lesser levels of turbulence impact upon cellular morphology (2,3), biochemistry and physiology (4), floc size (5), and attachment and detachment from surfaces (6,7). These influences of shear stresses are reflected in the industrially important aspects of production rate, productivity, product spectrum, and characteristics. Among the most susceptible to shear effects are the animal cells, plant cells, and cells of some microalgae and cyanobacteria. Shear-associated problems have the greatest economic impact in the animal cell culture industry because it is among the largest ones based on shear-sensitive cells and because of the high-value nature of most cell-culture-derived products (8). Damaging consequences of intense shear fields are also well documented for other microorganisms. For example, shear effects may alter characteristics of extracellular microbial polysaccharides by affecting the secretery process in some bacteria (9,10). Considerations of shear stress levels are also important in design and operation of biomedical devices such as dialysis machines, heart-lung machines, and transfusion filters. Shear stresses associated with implanted devices such as artificial heart valves can have medically significant implications. Increasingly, shear-sensitive human cells are being used to regenerate damaged tissue in vitro for later implanting or grafting. In other instances, susceptibility to shear may be beneficial, for example, during recovery of intracellular material by mechanical disruption (1,11,12). Some of the factors that determine shear sensitivity are summarized in Table 1.

SHEAR FORCES IN BIOREACTORS

Substantial information exists on the effects of hydrodynamic forces on cells in defined flow geometries such as viscometers and capillaries (5,13–18), but little is known about shear fields in bioreactors (19–21). Whereas selection of more shear-tolerant cell lines can be helpful, successful culture of shear-sensitive biocatalysts requires attention to bioreactor design and operation.

Table 1. Factors Determining Shear Sensitivity

1. Type of cell and species
2. Composition and thickness of cell wall when present
3. Size and morphology of cell
4. The intensity and nature of shear stress, whether turbulent or laminar, or associated with interfaces (e.g., during bubble rise and rupture)
5. Growth history, both short-term (e.g., starvation) and long-term adaptation
6. Growth environment (pH, temperature, agitation intensity, light in photosynthetic cultures)
7. Growth medium (trace elements, vitamins, carbon and nitrogen sources)
8. Growth rate
9. Growth stage
10. Type and concentration of shear-protective agents if present

The magnitude of the fluid mechanical forces is often expressed as shear stress, τ, or shear rate, γ. These quantities are related; thus, in laminar newtonian flow,

$$\tau = \gamma \mu_L \tag{1}$$

where μ_L is the viscosity of the fluid. Shear rate is a measure of spatial variation in local velocities in a fluid. Cell damage in a moving fluid is sometimes associated with the magnitude of the prevailing shear rate or the associated shear stress, but these quantities are neither easily defined nor easily measured in the relatively turbulent environment of most bioreactors. Moreover, shear rate varies with location within a vessel. Attempts have been made to characterize an average shear rate or a maximum shear rate in various types of bioreactors and process flow devices, as discussed later for the more common cases.

Bubble Columns and Airlift Bioreactors

Expressions for mean shear rate in bubble columns have been summarized by Chisti (21), Chisti and Moo-Young (19), and Shi et al. (22). These expression generally correlate the average shear rate with the superficial gas velocity; thus

$$\gamma = k U_G^a \tag{2}$$

where the parameter a equals 1.0 in most cases, but the k value has been reported variously as 1,000, 2,800, 5,000 m^{-1}, and so forth (20). The enormous disagreement among various authorities is obvious in Figure 1, where the shear rate calculated according to several available equations is shown as a function of the superficial aeration velocity in air–water system in a bubble column.

Equation 2 has also been applied to airlift bioreactors, using the superficial gas velocity in the riser zone as a correlating parameter; however, that usage is incorrect (19). A more suitable form of equation 2 for airlift reactors is

$$\gamma = \frac{k U_{Gr}}{1 + \dfrac{A_d}{A_r}} \tag{3}$$

where U_{Gr} is the superficial gas velocity in the riser, A_r is the cross-sectional area of the riser, and A_d is the cross-sectional area of the downcomer. In addition to the already-noted discrepancies (Figure 1), equations 2 and 3 have other significant flaws. The shear rate should depend also on the momentum transfer capability of a fluid, that is, on the density and the viscosity of the fluid, but equations 2 and 3 show no such dependence. Indeed, it is well known that the bubble size in a turbulent field depends on the viscosity and the density of the fluid as well as on the specific energy input rate. It is therefore reasonable to assume that correlations that express the shear rate as a function of U_G (or U_{Gr}) alone are incomplete (19). Furthermore, correlations such as equation 2 have generally been based on observations of phenomena at solid–liquid interfaces (e.g., heat transfer from coils or jackets), and their extension to phenomena at the gas–liquid interface or the bulk fluid is absurd at best. Shear stress and, hence, shear rate at walls of riser and downcomer zones of an airlift device are readily calculated using methods developed for pipes and channels, as discussed in a later section.

Another equation for estimation of an "effective" shear rate in airlift reactors is

$$\gamma = 3.26 - 3.51 \times 10^2 U_{Gr} + 1.48 \times 10^4 U_{Gr}^2 \tag{4}$$

which was developed for $0.004 < U_{Gr}\ (m\ s^{-1}) < 0.06$ (22); the shear rate range covered was 2–35 s^{-1}. Equation 4 was developed in an external-loop airlift reactor. First, the effect of viscosity on the induced liquid circulation velocity in the downcomer was established using Newtonian glycerol solutions at various gas flow rates. The maximum Reynolds number in the downcomer was about 3,200, or barely in the turbulent regime. In a second step, pseudoplastic media were used in the reactor, and the effective viscosity (μ_{ap}) of these fluids in the circulation loop was determined as being equal to the viscosity of the Newtonian glycerol solutions, when the two systems were at identical aeration rates and liquid circulation rates. The effective viscosity and the known values of K and n were used in the power law equation to calculate the prevailing shear rate:

$$\gamma = \left(\frac{\mu_{ap}}{K}\right)^{1/n-1} \tag{5}$$

The calculated shear rates were correlated with the superficial gas velocity in the riser, as noted in equation 4 (22). Although written in terms of the superficial gas velocity in the riser, equation 4 may usefully be expressed in terms of the specific power input in the reactor, as recommended elsewhere (19).

Because the procedure employed by Shi et al. (22) equated the viscosity-associated reduction in the liquid cir-

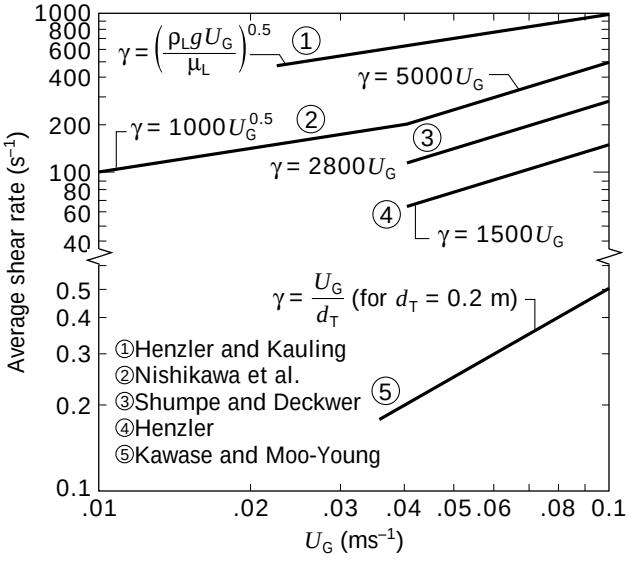


Figure 1. Average shear rate in air–water system in a bubble column according to various sources. *Source*: Based on Ref. 21.

culation velocity in different fluids, it gave in some sense a shear rate in the vicinity of the interface between the fluid and the walls of the reactor; shear rate in the bulk flow, which is the quantity of interest in most cases, was not quantified. Furthermore, the method of analysis used (22) applies strictly to a laminar flow regime, quite unlike the flow situations encountered in most practical operations.

Compared with bubble columns, equation 4 yields quite low values for shear rates in airlift reactors, as noted by Shi et al. (22). In such comparisons, care needs to be taken to ensure that the devices are being compared at identical values of specific power inputs (19,21). Although Shi et al. (22) did not adhere to this criterion, the specific geometry of the reactor they used was such that the error was small. Unlike what Shi et al. (22) concluded, equation 4 is not suitable for correlating mass transfer from gas bubbles or suspended solids because it does not give shear rates at gas–liquid or particle–liquid interfaces. Similarly, the usefulness of the shear rate calculated using equation 4, for correlating survival of fragile biocatalysts, remains questionable. An alternative, mechanistic approach to quantifying the bulk shear rate in various zones of airlift bioreactors has been advanced by Molina Grima et al. (14). Because the hydrodynamic environment in various zones of airlift reactors tends to be quite different, characterization of shear rate by a single global value is not sensible. The overall shear rates can be deceptively low even though damaging levels may be experienced in the high-shear zones (14); hence, the approach of Molina Grima et al. (14) is preferred.

Yet other attempts at characterizing the hydrodynamic forces relied on measurement of noises associated with events such as bubble formation at the sparger, bubble disengagement at the surface, and liquid flow over the baffle into the downcomer (23). Spectra of acoustic signals were measured in water, a salt solution, and aqueous glycerol media in a bubble column and split-cylinder airlift reactor (23). Because several events occur simultaneously in bubbling reactors, there were substantial ambiguities in assignment of the noise signals to specific events. The practical significance of the results was unclear even though the intended aim was to somehow relate the data to damage of fragile cells that has been reported by others in such reactors. Yet other work (24) examined the structure of turbulence in water and power law solutions ($K = 0.0194$ Pa $s^{0.973}$ and 0.0596 Pa $s^{0.958}$) in an external-loop airlift reactor. The reactor achieved complete gas–liquid separation and there was no gas in the downcomer. Measurements of local root mean square velocity fluctuations as an indicator of turbulence intensity showed a slight decrease from the center of the riser to the wall. These measurements were at a constant gas velocity of 2.05×10^{-2} m s^{-1}. The magnitudes of the velocity fluctuations were similar for all media; however, the velocity fluctuations were much lower in the gas-free downcomer than in the riser despite similar values of Reynolds numbers in the two zones. Based on measurements of one-dimensional energy spectra in the center of the riser (24), turbulence could not be considered isotropic, particularly in power law fluids. Other substantial evidence for the absence of isotropic turbulence in airlift and bubble column reactors has been documented (25,26).

Stirred-Tank Fermenters

The local velocity at a fixed position in any vessel fluctuates around a mean value; hence, the shear rate fluctuates. The extent of fluctuations is position dependent. In stirred tanks with radial-flow impellers such as Rushton turbines, the velocity fluctuations are greatest near the impeller tip and decline rapidly as one moves radially outward from the tip. Elsewhere in the vessel, the velocity fluctuations are reduced yet further. The time-averaged mean shear rate generated by a six-bladed Rushton turbine agitating a Newtonian fluid in a baffled vessels of diameter d_T is given as (27)

$$\gamma_{av} = 4.2N\left(\frac{d_i}{d_T}\right)^{0.3}\frac{d_i}{W} \tag{6}$$

where W is the width of the impeller blade, d_i is the impeller diameter, and N is it's rotational speed. The time averaged maximum shear rate value is 2.3-times γ_{av} (27). Another expression for the maximum shear rate at Rushton turbine blades is (28)

$$\gamma_{max} = 3.3N^{1.5}d_i\left(\frac{\rho_L}{\mu_L}\right)^{1/2} \tag{7}$$

which applies to Newtonian and non-Newtonian liquids when $100 \leq (Nd_i^2\,\rho_L/\mu_L) \leq 29{,}000$. The uncertainty in the coefficient (equation 7) is said to be $\pm 20\%$. For non-Newtonian fluids, μ_L in equation 7 is the zero shear viscosity. A different equation has been proposed by Wichterle et al. (28):

$$\gamma_{max} = N(1 + 5.3n)^{1/n}\left(\frac{N^{2-n}d_i^2\rho_L}{K}\right)^{1/1+n} \tag{8}$$

Equation 8 provides the maximum shear rate on a Rushton turbine blade. An average shear rate expression applicable to a broader range of impellers and media is (29):

$$\gamma_{av} = k_i\left(\frac{4n}{3n+1}\right)^{n/n-1}N \tag{9}$$

where n, the flow index of a fluid, equals 1.0 for a Newtonian liquid. In equation 9, k_i is an impeller-dependent constant. Some typical k_i values are noted in Table 2. The various shear rate correlations discussed here are compared in Figure 2 for water in a standard stirred tank (30) agitated

Table 2. The k_i Values for Use in Equation 9

Impeller	k_i
Six-bladed disc turbines	11–13
Paddle impellers	10–13
Propellers	~10
Helical ribbon impellers	~30

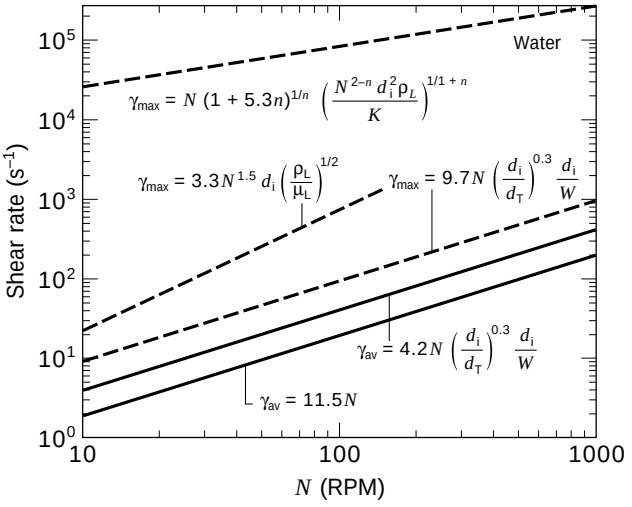


Figure 2. Comparison of the various shear rate correlations for water in a standard stirred-tank agitated with a 0.1-m-diameter six-bladed Rushton turbine. The dashed lines are for the maximum shear rate; the solid ones are for the average shear rate. See text for additional details.

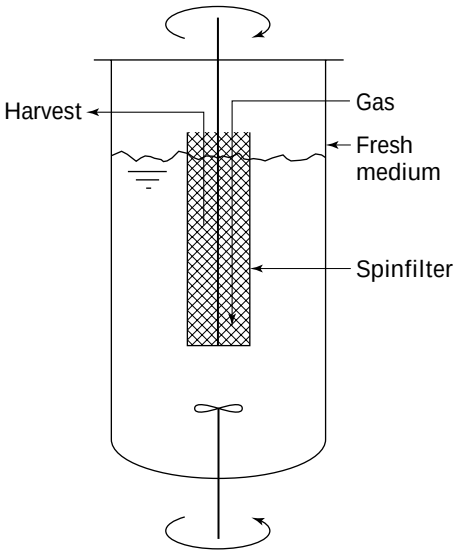


Figure 3. Perfusion culture with spinfilter for hydrodynamics-based cell retention in bioreactors.

with a 0.1-m diameter six-bladed Rushton turbine. For some contexts, the shear rate around a rising bubble may be approximated as the ratio of the terminal rise velocity to the bubble diameter (or radius); hence,

$$\gamma_{av} = \frac{2U_B}{d_B} \tag{10}$$

In air–water, under conditions typical of bubble columns and airlift reactors, the bubble rise velocity is about 0.2 m s^{-1}, and the bubble diameter is about 0.006 m. Thus, the interfacial shear rate approximates to 67 s^{-1} if the interface is nonmobile. Lower shear rates are expected at circulating interfaces.

Spinfilter

Spinfilters—rotating cylinders made of wire screen with openings that are significantly bigger (e.g., 25 μm) than the cells—are used to retain suspended animal cells in perfusion culture continuous bioreactors. Most of the cells are kept on the upstream side of the wire mesh by a hydrodynamic mechanism that depends on rapid rotation (e.g., 500 rpm) of the cylindrical screen (Fig. 3). The spent, largely cell-free medium is withdrawn from the zone within the rotating screen. Shear stresses associated with rotation of spinfilters do not generally damage animal cells (31). Sometimes the cell-free zone within the confines of the screen is also used for submerged aeration.

Other Flow Devices

Bioreactors and process vessels are where much of the shear-sensitive material is processed; however, at least part of the processing usually requires transfer of material between vessels. Transfer is accomplished through pipework, and centrifugal pumps are sometimes used. Methods of estimating shear stress levels in these devices are briefly summarized next.

Pipes, Tubings, and Flow Channels. ***Newtonian Fluids.*** In developed laminar flow of a Newtonian fluid through a straight tube of diameter d, the wall shear rate is given as

$$\gamma_w = \frac{8U_L}{d} \tag{11}$$

where U_L is the mean flow velocity. For rectangular channels of height h, the maximum or wall shear rate in developed laminar flow is

$$\gamma_w = \frac{6U_L}{h} \tag{12}$$

where U_L is again the mean flow velocity.

The wall shear stress (i.e., the maximum value) in a flow channel such as the riser of an airlift reactor is related to the pressure drop (ΔP), the length L of channel, and the hydraulic diameter (20); thus,

$$\tau_w = \frac{d}{4L}\,\Delta P \tag{13}$$

Consequently, in turbulent flow, the wall shear stress is

$$\tau_w = \frac{1}{2}\,C_f\rho_L U_L^2 \tag{14}$$

where ρ_L is the liquid density and C_f is the Fanning friction factor. The latter is related with the Reynolds number as follows:

$$C_f = 0.0792\left(\frac{\rho_L U_L d}{\mu_L}\right)^{-0.25} \quad (15)$$

where d is the hydraulic diameter of the flow channel or pipe. From equations 14 and 15, the wall shear stress can be shown to depend on Newtonian viscosity of the fluid: $\tau_w \propto \mu_L^{-0.75}$ (14). Use of equations 11 and 12 and others given later that are similar presupposes that the liquid velocity is known. This is normally the case in pipes and channels; however, with airlift vessels the induced liquid circulation rate will often need to be estimated using available methods (20,21,32).

Non-Newtonian Power Law Fluids. Methods for calculation of wall shear stress in laminar and turbulent flow of power law fluids with and without dispersed gas, have been discussed by Chisti (21) and Chisti and Moo-Young (33). For a non-Newtonian pseudoplastic medium flowing through a tube such as the riser or the downcomer of an airlift bioreactor, the flow is laminar when (34)

$$Re \leq 2300\,\frac{3n+1}{4n} \quad (16)$$

where the Reynolds number is calculated using the equation

$$Re = \frac{8U_L^{2-n}d^n\rho_L}{K\left(\frac{6n+2}{n}\right)^n} \quad (17)$$

In the laminar regime, the wall shear stress for flow of a gas-free liquid has been expressed as (35)

$$\tau_w = \frac{8\rho_L U_L^2}{Re} \quad (18)$$

When the flow is turbulent (i.e., the Reynolds number exceeds the value calculated with equation 16), the wall shear stress in a gas-free liquid is (34,35)

$$\tau_w = \frac{0.316n^{0.121}\rho_L U_L^2}{8Re^{2/n(x-1)+2}} \quad (19)$$

In presence of a gas, the shear stress is a complex function of τ_w, the gas holdup, the properties of the liquid (ρ_L, K, n), and the superficial velocities of the gas and the liquid phases, as discussed elsewhere (21,33). In equations 16–19, d is the channel diameter, U_L is the superficial liquid velocity, K is the consistency index, and n is the flow behavior index. The parameter x in equation 19 is given as

$$x = 0.70n^{-0.591} \quad (20)$$

as reported by Chisti and Moo-Young (33) based on data of Sokolov and Metkin (34).

Submerged Jets. Submerged jets are encountered wherever a pipe or nozzle discharges a fluid beneath the surface of the same fluid in a larger reservoir. For example, culture broth recirculating from a tank to an ultrafiltration module and back to the tank could form a submerged jet. If the cross-sectional area of the discharge nozzle is less than about 25% of that of the reservoir, the wall effects can be neglected and the jet is said to be "free." In a stable submerged turbulent jet, the maximum shear stress occurs in the direction of discharge, six to seven nozzle diameters downstream from the orifice. This shear stress is given as

$$\tau_{max} = 0.018\rho_L u_o^2 \quad (21)$$

where u_o is the velocity at the orifice.

Centrifugal Pumps. Centrifugal pumps are commonly employed in bioprocessing; hence, a knowledge of the shear rate in these devices is essential to assessing their suitability for a given use. The shear rate in a Newtonian fluid adjacent to the rotating impeller of a centrifugal pump depends on the radial position. The local shear rate at any position has been expressed as

$$\gamma = 6.30NRe_L^{0.5} \quad (22)$$

where the local Reynolds number at the local diameter d_L is

$$Re_L = \frac{Nd_L^2\rho_L}{\mu_L} \quad (23)$$

Equation 22 is based on theoretical considerations, but it has closely correlated measured shear rate data over the range $10^2 \leq Re_L \leq 2 \times 10^6$ in a centrifugal pump with the impeller rotating in a standard volute casing (36). Over the approximate impeller rotational speed range of 0.01–100 s^{-1}, the shear rate values ranged over 1–10^5 s^{-1} in aqueous electrolyte solutions (36).

As a general guideline, centrifugal pumps may be used to transfer most low-viscosity microbial broths and enzyme solutions so long as gas–liquid interfaces are rigorously excluded. Centrifugal pumps should not be used for animal cells, biofilm supporting particles, soft immobilized enzyme supports, suspensions of nematodes and protozoa, protein precipitates, and plant cells. These pumps will be satisfactory for some microalgae but not for others.

Shear Rate in Isotropic Turbulence

Irrespective of the bioreactor configuration, a turbulence field is deemed isotropic when the size of the primary eddies generated by the turbulence-producing mechanism is a thousandfold or more compared with the size of the energy-dissipating microeddies. The length scale of the primary eddies is often approximated as the width of the impeller blade or the diameter of the impeller in a stirred tank. In bubble columns and airlift bioreactors, the length scale of primary eddies is approximated as the diameter of the column (or the riser tube), or the diameter of the bubble issuing from the gas sparger.

Shear stress, shear rate, the dimensions of microeddies, and other characteristics of flow are linked ultimately to the energy input and dissipation rates in the fluid. The local shear rate in the vicinity of an eddy in an isotropically

turbulent field may be estimated as the ratio of the velocity and the length of the eddy; hence,

$$\gamma_i = \frac{u}{l} = \frac{\mu_L}{\rho_L l^2} \tag{24}$$

where ρ_L and μ_L are, respectively, the density and the viscosity of the fluid. The mean length, l, and the velocity, u, of the microeddies are related with the energy dissipation in the turbulence field; thus,

$$l = \left(\frac{\mu_L}{\rho_L}\right)^{3/4} E^{-1/4} \tag{25}$$

and

$$u = \left(\frac{\mu_L E}{\rho_L}\right)^{1/4} \tag{26}$$

where E is the rate of energy dissipation per unit mass of fluid. In most cases, all the energy input to the fluid is dissipated in fluid eddies, and E equals the rate of energy input. Equations 24–26 apply when local isotropic turbulence prevails.

Energy dissipation rate in bubble columns and airlift vessels is a function of the superficial aeration velocity (20,21); thus

$$E = gU_G \text{ (bubble columns)} \tag{27}$$

and

$$E = g\frac{U_{Gr}}{1 + \frac{A_d}{A_r}} \text{ (airlift bioreactors)} \tag{28}$$

Methods for calculating the specific energy dissipation rate in stirred vessels have been discussed elsewhere (30,37). The energy dissipation varies greatly from the mean value in a stirred tank. The maximum energy dissipation occurs in the vicinity of the impellers. This maximum value can be calculated using the equation

$$E = \frac{\mathrm{Po}N^3 d_i^2}{\rho_L} \tag{29}$$

where Po is the power number, d_i is the diameter of the impeller, and N is the rotation speed (s^{-1}). In developed turbulent flow in stirred vessels, the power number is generally constant for a given type of impeller and tank geometry. For single-phase pipe flow, the specific energy dissipation rate is

$$E = \frac{U_L \Delta P}{\rho_L} \tag{30}$$

where ΔP is the pressure drop per unit length, and U_L is the mean flow velocity.

Generally, if the dimensions of the biocatalyst particle are much smaller than the calculated length, l, of the microeddies, the particle is simply carried around by the fluid eddy; the particle does not experience any disruptive force. On the other hand, a particle that is larger than the length scale of the eddy will experience pressure differentials on its surface, and if the particle is not strong enough it could be broken by the resulting forces.

In addition to turbulence within the fluid, other damage-causing phenomena in a bioreactor include interparticle collisions; collisions with walls, other stationary surfaces, and the impeller; shear forces associated with bubble rupture at the surface of the fluid; phenomena linked with bubble coalescence and breakup; and bubble formation at the gas sparger. Some of these aspects are discussed elsewhere in this article. Summarizing, several possible shear rate values may be calculated for a given situation in a bioprocess device. Not every calculated value is appropriate or relevant to the problem at hand. For pneumatically agitated bioreactors, when the turbulence characteristics in the bulk fluid are the relevant ones, the preferred approach is to use equation 24 for shear rate in the vicinity of eddies. The same applies to mechanically stirred tanks. In some cases, the relevant shear rate may be that at the interface of a rising bubble unless turbulence is so intense that bubbles do not rise freely. Other situations would be controlled by the bubble rupture events, as discussed later. In yet other cases, the fluid eddy shear rate and the maximum shear rate at the impeller will need to be taken into account. Factors such as the frequency of passage of a sensitive biocatalyst through high-shear regions may need consideration.

RESPONSE TO SHEAR

Response to shear has been characterized using various indices, including growth rate (38–40); cell viability (41); regrowth potential (42); release of intracellular material (38,43,44); changes in oxygen uptake rate, ATP, productivity of metabolites, and biochemical composition of the cells; and alterations in morphology of cells and cell aggregates (2,42,44).

Most studies of shear response of cells have utilized batch culture with or without aeration. Such cultures contain cells at different stages of growth; hence, the size and metabolic state of cells are nonuniform. In contract, continuous culture is time-consuming (38), differs substantially from the culture methodologies that are typically employed commercially, and, through natural selection, the cells in continuous culture adapt to shear fields, hence, masking the effect of shear. For example, Wu et al. (45) noted for insect cells that the beneficial effect of shear-protective agents (serum, methyl cellulose, Pluronic F68) were not as important as the effect of passage number. Similar results have been alluded to for plant cells (38). To overcome problems associated with cell growth and adaptation to shear, Zhang et al. (43) recommended the use of erythrocytes for comparative evaluation of the hydrodynamic environment in bioreactors. A nongrowing, well-defined, and shear-susceptible cell population such as that provided by erythrocytes can help identify suitable regimes of operation of a given reactor. In addition, shear-

damaging potential of different bioreactor configurations can be assessed on a common basis using cells such as erythrocytes. Specific information on shear susceptibility of animal, plant, and various microbial cells is discussed in the following sections.

Bacteria, Yeasts, and Mycelial Microfungi

Yeasts, bacteria, and mycelial microfungi are generally quite robust and not easily damaged under typical processing conditions; however, even the most robust cells are broken when subjected to suitable combinations of high stress and exposure time (1). Often, genetically modified strains of an otherwise robust microorganism are less shear tolerant (1,10,46–48). With microorganisms, the composition of the growth medium and the specific growth rate used for the culture are known to influence the robustness of the cell (1). Cells grown in fully defined media that may lack one or more components essential to synthesis of a normal cell wall are often less robust (1). Growth at a higher specific rate will often produce weaker cells (1). When the shear field is insufficiently intense to damage cells, it may still affect product formation and characteristics (9,10). In one case, properties of a shear-sensitive polymer were affected, but not the producing bacterial cells (10). Cells experiencing physiological stress (e.g., starvation) are more susceptible to shear-stress-related damage (44).

Reports cited by Märkl et al. (44) document increasing release of intracellular material from mycelium of *Rhizopus javanicus* and *Mucor javanicus* as the stirrer speed increased. Increasing release of nucleotides was noted with increasing speed of agitation, or with increasing agitation time at a constant agitation speed. Effect of Rushton turbine tip speed on performance of a process for converting cellulose to protein-enriched microbial biomass were clearly demonstrated by Moo-Young et al. (39). The process utilized the filamentous microfungus *Neurospora sitophila* for the conversion. Increasing impeller tip speed from 2.35 m s^{-1} to 3.29 m s^{-1} reduced the specific protein production rate and the final yield (Fig. 4). In addition, a distinct lag phase was observed at the higher agitation rate (Fig. 4) when the inoculum grown in a more quiescent environment was transferred to the fermenter. At the highest tip speed shown in Figure 4, the specific protein production rate was only 55% of that at the lowest tip speed. Similarly, the yield of crude microbial protein was reduced to ~67% of the value obtained at the lowest agitation speed used (39). Further work revealed that agitation-associated damage could be reduced substantially by replacing the Rushton turbines with axial flow hydrofoil impellers (40); thus confirming that different types of impellers generate very different shear fields at identical tip speeds. For equal or lower power draw, hydrofoil impellers are typically better bulk mixers than Rushton turbines (30,37).

In *Penicillium chrysogenum* fermentations performed in turbine-impeller-agitated vessels with working volumes of 0.005–1 m^3, operated at impeller tip speeds of 2.5–6.3 m s^{-1}, the specific penicillin production rate and the mean length of the main hyphae were lower at high agitation intensities (2). High agitation speeds affected the fungal

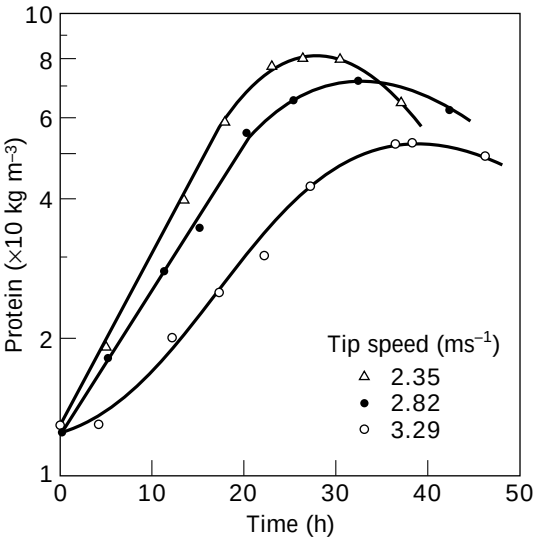


Figure 4. Effect of Rushton turbine tip speed on protein production by the mold *N. sitophila*. *Source*: From Refs. 39 and 40.

morphology, promoting mycelial fragmentation and a greater frequency of branching (2). The behavior of the fermentation at various scales correlated poorly with the impeller tip speed, but correlation was superior in terms of a parameter that took into account the frequency of passage of the culture through the intensely agitated zones around the impellers (2). This behavior suggests that the degree of hyphal damage in stirred bioreactors depends on a combination of the shear intensity in the impeller zone and the cumulative time of exposure of the microbial solids to the shear field.

Sometimes, especially with certain mycelial microfungi, the biomass yield may not be affected by agitation rate, but the cell morphology and production of metabolites may be influenced greatly. Using different strains of citric acid–producing *Aspergillus niger*, Ujcova et al. (3) showed that the yield of citric acid first increased with increasing agitation rate, presumably because of improved oxygen supply, but the yield declined because of cell damage if the agitation speed was increased further. The yield maxima occurred at different agitation rates for different strains, and the sensitivity to agitation differed among strains. Other morphological effects of nondamaging shear levels include changes in growth form, whether pelletlike or pulplike; changes in the degree of branching; the extent of entanglement; and whether growth occurs in filamentous or spherical cell form in certain dimorphic species.

Evidence suggests that even a supposedly robust cell such as *Saccharomyces cerevisiae* (1,11) is susceptible to mechanical shear damage in fermenters (49,50). In one study, *S. cerevisiae* suspended in isotonic saline lost viability at increasing rate over a 25-h period as the agitation speed of the paddle impeller was increased (50). At 1.43 m s^{-1} impeller tip speed, >90% of the cells were dead over the course of 25 h. In contrast, at 0.66 m s^{-1} tip speed, only about 15% of the cells died over the same period (50). Excessive shear has been alleged to also influence the pro-

ductivity of ethanol in fermentations employing *S. cerevisiae* (49).

Among methane bacteria, *Methanobacterium thermoautotrophicum* with rodlike cells that exceeded 2 μm in length was highly resistant to shear stress (44). In contrast, and surprisingly, the relatively smaller (<2 μm) spheroidal cells of *Methanococcus vanielii* were quite sensitive to shear stresses (44). Supplementing the medium with carboxymethyl cellulose viscosity enhancer increased the extent of the cell damage (44). Failure to provide an essential nutrient, selenium is this case, also enhanced shear susceptibility of the culture (44). Although many authors have successfully correlated cell damage with the tip speed of the impeller, Märkl et al. (44) believe that power input per unit culture volume is a superior correlating parameter that is also suitable for comparing the cell damage effects of mechanically stirred and other nonmechanically agitated bioreactors. A constant power input per unit culture volume has also been suggested as a bioreactor scale-up criterion for fragile biocatalysts (44). This recommendation notwithstanding, average power input alone cannot be a universally satisfactory determinant of cell damage; as noted in an earlier section, the maximum shear rate and the frequency of passage (or cumulative exposure) through the high-shear zone need to be taken into account, especially when the shear rate in a given device is strongly position dependent. Several studies confirm that frequency of passes through a high-shear environment have a clear impact on biocatalyst survival (1,2,11,12,16,20,51). Effects of variously intense shear fields on microbial cells are further discussed by Thomas (5,52), Prokop and Bajpai (53), and Chisti and Moo-Young (1). Shear effects on microbial flocs and pellets are treated in a separate section of this article.

Microalgae and Cyanobacteria

Cell fragility has been referred to as the key problem of culture of microalgae in photobioreactors (54); however, the real situation is not quite that bad. Among microalgae, the sensitivity to shear varies greatly with the species. Algae such as *Dunaliella* are indeed extremely fragile, but many other species will tolerate high levels of turbulence and sparging with gas. Thus, gas-sparged airlift devices are commonly used to circulate algal cultures through the solar receivers of photobioreactors (20,55–57). Similarly, relatively low-shear positive displacement rotary pumps have been successfully employed. High-shear centrifugal pumps have also been used; however, at least in some cases, the centrifugal pumps are known to be more damaging than the other fluid-moving systems. In one case, for otherwise equivalent conditions, replacement of the centrifugal pump with a positive displacement device doubled the productivity of *Porphyridium cruentum* (54), a producer of phycoerythrin and polyunsaturated fatty acids. Irrespective of the method used to circulate the culture, the culture flow velocities through tubular solar receivers typically range over 0.3 to 0.5 m s^{-1} in tubes up to 0.06 m in diameter. These conditions have been used with a variety of algae.

In cultures of *Haematococcus pluvialis*, a producer of astaxanthin, early onset of the stationary phase has been associated with turbulence-induced deflagelation of the cells (54). Fragility of species such as *H. pluvialis* depends also on the phase of growth; thus, tailoring the bioreactor operational regime to the growth phase can improve overall productivity (54). The stationary phase red cysts of *H. pluvialis* are significantly more robust than the green vegetative cells (54). The cyanobacterium *Spirulina*—a healthfood and source of the food colorant phycocyanin—is another species that is susceptible to damage by excessive shear forces. The marine alga *Dunaliella* is particularly fragile: the cell lacks a ridged wall; only a thin cytoplasmic membrane encloses it. *Dunaliella* is a good source of natural β-carotene; hence, it is commercially useful. The cell also produces glycerol. Silva et al. (58) showed that *Dunaliella* was susceptible to damage by gas bubbles as well as by turbulence in the liquid. Supplementing the culture with carboxymethyl cellulose or agar improved cell survival.

In one study with *Dunaliella tertiolecta*, culture in a bubble column was quite successful, but when the bubble column was converted to an airlift device by inserting a vertical baffle, the productivity declined (59). Under conditions that were earlier identified as optimal, no growth was observed in the airlift reactor, whereas good growth occurred in the bubble column. Microscopic examination showed significant disruption of the cells in the airlift device (59). This was associated with the hydrodynamic stresses generated as the culture flowed over the upper edge of the baffle into the downcomer (59). This effect could have been avoided, or at least minimized, by hydrodynamic smoothing of the upper and lower parts of the baffle to prevent flow separation (20). In the bubble column, the growth was sensitive to aeration rate: growth rate increased with increasing superficial gas velocity until a velocity of about 0.6 m min^{-1}. Further increase in aeration rate reduced growth, apparently because of hydrodynamic stresses in the fluid (59). Under nongrowth conditions (no light), the specific death rate in the bubble column was shown to increase with superficial gas velocity for velocities exceeding 0.6 m min^{-1} (59). At a fixed aeration velocity ($U_G = 1$ m min^{-1}), the specific death rate decreased with increasing height of the culture fluid in the column (59), probably because the specific power input and, hence, the turbulence intensity declined (20). Similar behavior has been reported with animal cells in bubble columns (60–64).

For otherwise fixed conditions, increasing the number of aeration nozzles (i.e., decreasing the gas jet velocity) caused a marginal reduction in the specific death rate of *D. tertiolecta* (59). This led to the conclusion that the events at the gas sparger did contribute to cell death (59), although the effect was minimal. Unlike Suzuki et al. (59), others (56,57) have quite successfully cultured various microalgae in pilot-scale airlift devices. Continuous culture of the green alga *Chlorella pyrenoidosa* in a rectangular airlift device with multiple light-emitting diodes was reported by Matthijs et al. (65). With a relatively robust photosynthetic organism such as *Chlorella*, increasing turbulence within limits improves productivity in photobioreactors because of enhanced fluid interchange between the dimly lit interior and the better-illuminated exterior parts of a photobioreactor.

In addition to an organism's morphology, the biochemical composition of the cell wall has been associated at least qualitatively with the cell's susceptibility to hydrodynamic stresses. The microalga *Chlorella vulgaris*, which is shear tolerant, has a thick (100–200 nm) cell wall composed of cellulose, pectin, and sporopollenin (44). The somewhat more sensitive alga *Chlorella reinhardii* has only glycoproteins in its relatively thin (40 nm) cell wall (44), whereas a highly sensitive cell wall defective mutant of *C. reinhardii* is devoid of cell wall (44). Similarly, the cyanobacterium *Synechococcus*, with a murein-containing cell wall, and the methane bacterium *M. thermoautotrophicum*, with a pseudomurine-containing wall, are noticeably more resistant than methane bacteria such as *Methanococcus vanielli*, which has a protein-based wall.

Although earlier work with the mold *M. javanicus* showed a reduced extent of damage to mycelia as the viscosity was increased at constant agitation rate, and similar reports exist with animal cells (66,67) and the halotolerant alga *Dunaliella* (58), Märkl et al. (44) observed an opposite effect: with a highly shear-sensitive strain of *C. reinhardii*, even slight increase in viscosity enhanced cell damage (44). The viscosity was increased by adding carboxymethyl cellulose to the culture broth. With a relatively less shear sensitive strain of *C. reinhardii*, increasing the viscosity reduced the impeller agitation speed at which the damaging effects were first discerned (44). Apparently, at least with the less sensitive strain, supplementing the medium with the viscosity enhancer caused the cells to grow in clumps, and aggregation supposedly was an effective shield against cell damage (44). Increasing shear damage with increasing culture viscosity has been seen before with a protozoan (68). These apparently contradictory results are easily explained: whether changes in viscosity increase or reduce shear damage depends on the peculiarities of the mechanism causing the damage. For a given impeller tip speed, increasing viscosity will not affect the tip speed–associated shear rate, but the turbulence will be reduced; hence, the shear rate in the fluid should be lower. However, the shear stress around a particle may increase if the fluid eddy length scale remains comparable to the dimensions of the particle. In other cases, increased viscosity may reduce both shear rate and shear stress: only eddies that are larger than the particle may survive and there may be little or no relative motion between the fluid and the particle. In yet other cases, the frequency of passage through high-shear zones may decline with increasing viscosity.

Plant Cells

Cultured plant cells are a potential source of numerous high-value compounds, although few production processes have been commercialized (13,69). Low culture productivities continue to be the primary limitation. Investigational culture in large bioreactors has been carried out since at least 1959. A variety of reactors have been used, including bubble columns, airlift devices, stirred vessels, rotary-drum reactors, and immobilized cell systems (13). Airlift bioreactors have proven especially effective for large-scale culture of heterotrophic, photomixotrophic, and photosynthetic plant cell suspensions (20,70). Relatively low values of pneumatic power inputs have given satisfactory results during cultivation of plant cells. For example, Fischer and Alfermann (70) obtained good results with about 13 W m^{-3} specific power input during culture of *Chenopodium rubrum* cells in a concentric draft-tube airlift device. Similarly, Ballica and Ryu (71) obtained maximum biomass and product yield at an approximate power input value of 35 W m^{-3}, also in a draft-tube airlift reactor, while culturing *Datura stramonium* cells. In the latter case, power inputs above ~60 W m^{-3} caused large declines in biomass and product yields. The size of the cell aggregates declined with increasing specific power input (71). Reinhard et al. (72) described culture of *Digitalis lanata* in a 300-L airlift bioreactor for producing β-methyldigoxin.

Plant cells in suspension tend to be large: 20–50 μm in diameter and 100–500 μm in length. The cells have a cellulose-based wall. Most of the cell's volume is taken up by a vacuole that may occupy >95% of the available space. The fluid in the vacuole is much less viscous than the cytoplasm; hence, a cell with a large vacuole is likely to be less resistant to certain types of deformations than one with a smaller vacuole. Cells frequently occur as aggregates of up to a hundred cells. The aggregates settle easily. Plant cells grow slowly, with doubling times of the order of 20–100 h. Cells may secrete extracellular polysaccharides, producing non-Newtonian broths. Also, the biomass concentration may be high, up to 70 kg m^{-3}. Heterotrophically growing plant cells require oxygen, but demands are low, of the order of 10^{-6} kg O_2 kg^{-1} cell s^{-1}. Critical dissolved oxygen concentration generally tends to be ≤20% of air saturation. The polysaccharide-containing cultures tend to foam a lot when aerated. Also, excessive aeration may reduce productivity by stripping carbon dioxide and other essential volatiles produced by the cells themselves (5,69).

As has been demonstrated with some microorganisms (1), the composition of the culture medium, the culture conditions, and the specific growth rate of a cell may affect the wall strength of plant cells also. For otherwise identical conditions, agitation intensity is reported to influence the composition of plant cell walls. Thus, the relative amounts of pectin, cellulose, and hemicellulose in walls of *Catharanthus roseus* were affected by the intensity of agitation in one case (73). In another case cited by Doran (13), the mass ratio of cell wall relative to the whole cell increased with increasing intensity of shear, suggesting adaptation of the cell to the prevailing environment even over short periods. Even though the turbulence level may be insufficient to affect individual cells, aggregates may be affected; hence, a shear-related metabolic response may occur due to altered cell–cell interactions (69).

Based on laminar flow viscometric measurements, a critical shear stress level of 80–200 N m^{-2} has been suggested for *Morinda citrifolia* cells. Whereas for *Daucus carota*, a shear stress level of 50 N m^{-2} has been associated with cell damage. In another study, carrot cells in a laminar flow Couette viscometer lost the ability to grow and divide in the shear stress range of 0.5–100 N m^{-2} (53). The intracellular enzyme activity was impaired at shear stress levels above 3,000 N m^{-2}, but significant lysis did not occur until a shear stress level of 10,000 N m^{-2} was applied over a prolonged period (≥1 h) (53). In contrast to behavior in

laminar flow, the cells were quite sensitive to turbulent impeller agitation. Impeller tip speeds of ~1.1 m s^{-1} lysed a significant proportion of the cells within 40 min (53).

In studies with *Perilla frutescens* cell suspensions, Zhong et al. (74) noted that the cell viability declined with increasing shear rate (constant exposure time) or increasing shear time at constant rate of shear in a concentric cylinder shearing device. A constant shear rate of about 140 s^{-1} produced a noticeable loss of viability within 20 min. At higher shear rate of ~860 s^{-1}, less than 75% of the cells were viable after 20 min (74). In marine propeller–agitated fermenters, the impeller tip speed affected the specific growth rate and the yield of the cells (74). Tip speeds >0.6 m s^{-1} reduced the specific growth rate and the cell biomass yield (74). In other cases, impeller tip speeds as low as 0.35 m s^{-1} have extensively lysed certain plant cells.

With *M. citrifolia* cells subjected to 25–350 N m^{-2} turbulent shear stress in a capillary, the viability decline followed first-order kinetics in which the rate constant increased with the applied shear stress (75). The culture was circulated through the capillary by a peristaltic pump, which also contributed to cell damage (75). The specific death rate constant was linearly related with the cumulative energy dissipated (75). Cell aggregates subjected to shear were reduced in size. For the same species, damage was reported when cells grown in a quiescent environment were subjected to turbulent jet flow through 1- and 2-mm nozzles (51). Cells lost viability with increasing number of passes through the nozzles. The viability loss was first order in the number of passes (51). For a fixed number of nozzle passes, increased nozzle flow velocity enhanced the death rate constant. Similarly, the maximum and the mean aggregate size were reduced when the suspension was subjected to jet flows (51). The average aggregate size was lower for higher jet velocities. The average energy dissipation rates in the jet ranged over 10^3–10^5 W kg^{-1} (51).

Meijer et al. (38) assessed the hydrodynamic stress sensitivity of *C. roseus, Nicotiana tabacum, Tabernaemontana divaricata,* and *Cinchona robusta* in a 3-L Rushton turbine-stirred vessel. Effects of stress were quantified in terms of culture growth rate, respiration rate, and release of organic compounds from the cells. Cultures of *C. roseus* carried out at 0.59 m s^{-1} and 2.36 m s^{-1} impeller tip speeds showed identical behavior in terms of biomass growth rate and glucose consumption rate; hence, hydrodynamics-associated stress was concluded to not affect the cells over the ranges examined (38). Similarly, *N. tabacum* was little affected by differences in agitation rate, although the growth rate and the final yield were noticeably greater at the lower agitation speed. Cells of *C. robusta* responded differently: at 0.59 m s^{-1} impeller tip speed the cells grew and there was little release of intracellular protein; however, at 2.36 m s^{-1}, there was no growth and the specific protein concentration increased with culture time, indicating leakage from damaged cells (38). The *T. divaricata* cells also were damaged at 2.36 m s^{-1}, but less than the cells of *C. robusta*. The damage to *T. divaricata* was reduced at a lower impeller tip speed of 0.59 m s^{-1}. In separate experiments, cells of *N. tabacum* that had earlier withstood higher shear stress levels were shown to be more susceptible to damage after a prolonged period (10 days) of starvation (38). This effect is similar to that noted with other cells (44).

Table 3 is a summary of the damage thresholds levels—either impeller tip speed or the specific power input due to aeration—noted for various plant cell suspensions in stirred fermenters and airlift bioreactors. The studies just referenced clearly affirm that the susceptibility to shearing forces varies a great deal with cell line. Moreover, for the same cell line, cultivation history (e.g., prolonged starvation) and age of cell affect the cell's response to shear fields (38,44,69), which is consistent with similar findings for microorganisms (1). That plant cells adapt to shear fields is also well known: cells cultured in high-shear environments better withstand hydrodynamic forces after many generations.

Protozoa

Midler and Finn (68) provided a clear demonstration of shear-related damage to cells of the large (~80-μm cell diameter) ciliated protozoan *Tetrahymena pyriformis*. Lysis correlated with impeller tip speed. Agitation at impeller tip speeds approximately ≥0.5 m s^{-1} destroyed most of the cells. In a Couette viscometer, the critical shear stress level when lysis first occurred was ~24 N m^{-2} (68). Eighty percent of the cells lysed within the first minute when exposed to shear stress levels of 10^3 N m^{-2}. The extent of damage increased with shear stress level and exposure time. At a given shear rate, increasing the broth viscosity made the protozoan more susceptible to shear damage, implying that the damage was associated with shear stress rather than shear rate (68). The cells were more sensitive to turbulent shear stress: an average turbulent shear stress of ~10 N m^{-2} lysed over half the population within 10 min. In contrast, in laminar flow, similar levels of damage occurred at shear stress levels of almost 300 N m^{-2}. Note that *T. pyriformis* cells are larger than most animal cells, but smaller than most cultured plant cells.

Animal Cells

Animal cells are the most shear-susceptible of biocatalysts. Animal cells are relatively large—typically 10–20 μm in diameter—and they lack a protective cell wall. Osmotic

Table 3. Damaging Threshold Values of Impeller Tip Speed or Specific Power Input for Some Plant Cells

Stirred fermenters	Impeller tip speed (m s^{-1})
Catharanthus roseus	≥2.36
Nicotiana tabacum	~2.36
Cinchona robusta	<2.36
Daucus carota	<1.10
Tabernaemontana divaricata	~0.59
Perilla frutescens	0.60
Others	≤0.35
Airlift bioreactors	**Power input (W m^{-3})**
Chenopodium rubrum	≥13
Datura stramonium	~60

shock is often sufficient to lyse animal cells. Other than blood cells, most normal animal cells are anchorage dependent. Cancerous cells, hybridomas, and established lines will normally grow in suspension. Because of these distinct growth characteristics, two types of culture methods are employed in industrial practice: free suspension culture, in which cells are freely suspended in a nutrient broth; and microcarrier culture, in which anchorage-dependent cells grow on surfaces of solid carriers that are suspended in a culture fluid. Irrespective of the culture method, the suspending fluid is invariably directly sparged with air or other gas mixture to supply the cells with oxygen. Although other aeration alternatives are available (see MASS TRANSFER article), direct sparging with a gas is likely to remain the dominant method of providing oxygen to animal cells (76). Both stirred tanks and airlift devices have been successfully employed for commercial processing of animal cells (77–79). Anchorage-dependent cells may also be cultured on surfaces of large static particles held in packed beds (80,81), but this method is uncommon. Devices used in generation of implantable tissue typically grow normal cells supported on stationary surfaces or within a static matrix of gel or woven fabric (82). The supporting surface may be flat or variously shaped.

Cells in bioreactors are subject to two main kinds of fluid mechanical forces: (*1*) those due to turbulence within the liquid, and (*2*) forces associated with the gas–liquid interface. The latter category includes effects of bubble formation, rise, and rupture. In addition in microcarrier culture, collisions among carriers and carrier–impeller (or carrier–wall) interactions may play a role. Furthermore, in tall bioreactors cells may be subjected to substantial changes in hydrostatic pressure as they circulate with the fluid. Cells in tissue regeneration bioreactors experience the same kind of forces as occur in flow past a flat plate, wall, or other submerged object. The specific considerations relevant to the various culture methodologies are treated in the following sections.

Resistance to shear stress is determined at least partly by the cytoskeletal structure of the cell. The cytoskeleton determines a cell's mechanical properties, shape, and movements such as occur during phagocytosis. The cytoskeleton is made of protein filaments, mainly of actin. The filaments interact closely with the cell membrane by anchoring to specific membrane proteins. A cell exposed to mechanical forces undergoes cytoskeletal reorganization. Cytoskeleton-mediated stress signal transmission to intracellular organelles (Fig. 5) has been associated with various biochemical, physiological, and gene-level responses in vascular cells (4). Some cells are known to have surface stress receptors (e.g., Ca^{2+} ion channels) that have been implicated in biochemical responses (4,83,84) independent of cytoskeletal mediation (Fig. 5). Cytoskeletal properties are known to affect the shear tolerance of hybridomas. Effects of turbulence on the cytoskeleton and other biological responses have been treated by several authors (4,53,85). Work cited by Al-Rubeai et al. (86) suggests that inhibitors of actin polymerization decrease a cell's resistance to shear forces, whereas inducers of polymerization make the cells more robust. Thus, supplementing the culture medium with actin polymerization inducers may be one way of improving performance of bioreactor-cultured animal cells, so long as the requisite production biochemistry is not adversely affected.

Plasma membrane fluidity (PMF), which is a measure of the degree of packing of the membrane and its constituent's freedom to move about laterally through the bilayer membrane, to rotate about their major axes, and to oscillate in the plane of the membrane, is another factor that appears to influence shear susceptibility of cells (41). Increasing membrane fluidity correlates with increasing shear sensitivity (41). PMF increases with increasing temperature and can be manipulated in both directions using various additives. In studies with a freely suspended mouse–mouse hybridoma in a laminar shear stress Couette flow device, conditions that yielded higher membrane fluidity produced more fragile cells (41). For example, cells sheared at elevated temperatures showed greater loss of viability than unsheard controls at the same temperature (41).

Shear Effects in Bubble-Free Environments ***Suspended Cells.*** Freely suspended cells in bubble-free bioreactors are not damaged by mechanical agitation even at intensities much greater than the ones used in typical processing. Exceptions occur in extensional or elongational flow in certain high-shear devices even when the flow is laminar. Extensional or elongational flow occurs whenever the cross-sectional area of the flow channel reduces (e.g., at an orifice on the wall of a tank or at the entrance of a capillary connected to a larger reservoir). The fluid elements undergoing extensional flow stretch and thin. Suspended particles also experience elongational forces in the direction of flow, and compression perpendicular to the flow streamlines. Drops subjected to extension flow can rupture. Similarly, extensional flow through orifices of high-pressure homogenizers contributes to rupture of even the very robust microbial cells (1). Rupture of erythrocytes at entrances to capillaries is well known (87). Shear effects on suspended erythrocytes are discussed in detail in a later section.

For a hybridoma line, Born et al. (88) reported that exposure to laminar shear stress (208 N m^{-2}) in unaerated flow in a cone-and-plate viscometer led to substantial loss in cell count and viability within 20 min. At a constant 180-s exposure, increasing shear stress over 100–350 N m^{-2} linearly enhanced cell disruption, with >90% of the cells being destroyed at 350 N m^{-2} stress level (88). Shear stress levels associated with bubble rupture at the surface of a bioreactor may range over 100–300 N m^{-2} (13). These values are remarkably consistent with shear rates that damaged hybridomas in *unaerated* laminar flow experiments (88).

For a suspended mouse myeloma line in turbulent capillary flow, McQueen et al. (18) noted a threshold average wall shear stress value of 180 N m^{-2} when lysis first commenced. Although the flow caused lysis, it had no effect on viability (18), suggesting that cells at various growth stages were equally affected. The sudden flow contraction at the entrance to the capillary may have contributed to cell lysis, but the residence time in the capillary also had an effect at otherwise constant average wall shear stress level. The rate of lysis was first order in cell number. Above

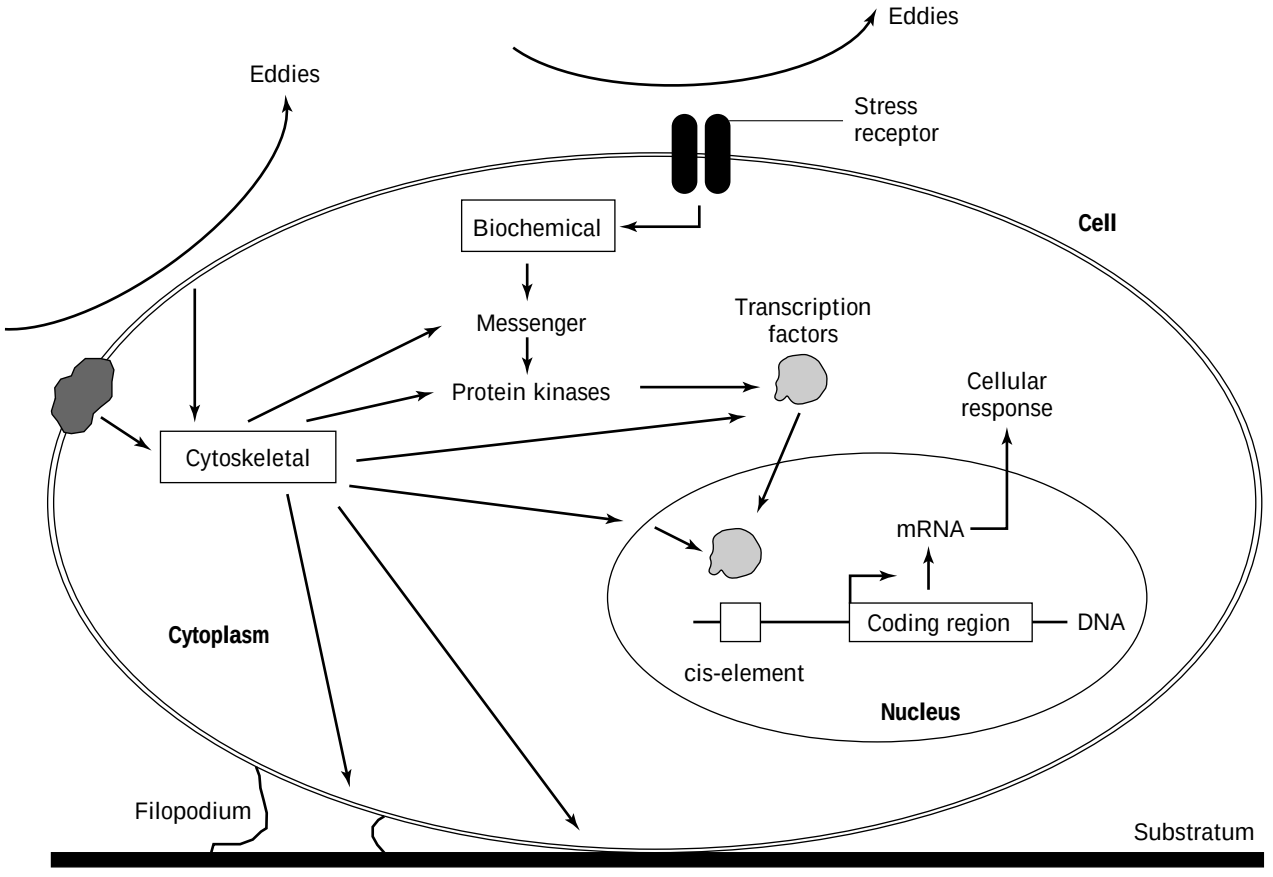


Figure 5. Stress signal transmission and transduction via cytoskeletal and biochemical routs in an anchored cell. *Source*: Adapted from Ref. 4.

the threshold shear stress value, the specific lysis rate increased with increasing level of shear stress (18). The growth rate and the DNA synthesis rate of the cells exposed to the shearing environment were unaffected when the surviving cells were returned to a normal quiescent growth environment (18). In other studies cited by McQueen et al. (18), the shear stress threshold for damage has been reported as 0.87 N m^{-2} for a mouse hybridoma and 1.5 N m^{-2} for insect cells. Higher shear sensitivity of another mouse cell line relative to a human carcinoma has been reported (16).

For human cervical carcinoma HeLa S3 and mouse abdominal fibroblast L929, Augenstein et al. (16) observed lysis of suspended cells in turbulent flow through stainless steel capillaries. Cell death could be correlated with the average wall shear stress level or the power dissipation within the capillaries. The L929 line was more sensitive than the human cell. Control experiments showed that the positive displacement pumps used to circulate the cells through capillaries contributed little to cell lysis (16). Average wall shear stress levels of 0.1–2.0 $\times$ 10^3 N m^{-2} were sufficient to induce cell inactivation for the two lines. According to Shiragami (17), the mean shear stress acting on cells suspended in capillary flow is 4/3 the shear stress at the capillary walls, so long as the ratio of the cell's diameter to that of the capillary is <0.08.

For a hybridoma examined by Shiragami (89), the specific rate of monoclonal antibody production in a surface-aerated spinner flask depended on the agitation speed. In a 250-mL spinner vessel an agitation rate of $\sim$180 rpm gave the highest specific antibody production rate. The specific productivities were reduced at higher or lower values of agitation speed. The increased antibody production with increasing agitation was associated supposedly with enhanced secretion in a more turbulent environment (89). Oxygen transfer effects may have better explained the observations (see MASS TRANSFER), but no data were reported on this aspect.

Damage to murine hybridomas was observed by Jan et al. (31) in stirred tanks equipped with marine impellers agitated at sufficiently high speeds that vortexing occurred and gas entrained into the medium. Even at these high speeds, damage could be prevented by baffling the tank, which suppressed vortex formation. Usually though, vortexing is not a problem in large-scale cell culture. Unbaffled, marine impeller–stirred tanks were successfully used by Chisti (78) in industrial culture of several hybridoma lines. Effects of agitation on hybridoma culture in the absence of sparging or surface entrainment were further examined by Smith and Greenfield (90). Culture growth was unaffected by agitation intensity (100 or 600 rpm) in the RPMI medium supplemented with fetal bovine serum

(10% vol/vol). However, in PFHM II medium supplemented with either Pluronic F68, fetal bovine serum, or bovine serum albumin, and agitated at 600 rpm (impeller tip speed = 1.6 m s^{-1}, power input = 1 kW m^{-3}) the results were different: the agitation intensity did not affect the exponential growth rate, but once growth had ceased, the decline phase was substantially faster than in control experiments (90).

A certain level of hydrodynamic shear is generally beneficial to culture processes, especially in intraparticle immobilized culture where mass transfer limitations can be severe. In producing engineered tissue of smooth muscle cells/polymer matrices (biodegradable 2-mm-diameter fibers of polyglycolic acid assembled into a nonwoven matrix), Kim et al. (82) noted that seeding of the matrix under agitation led to significantly higher intramatrix cell densities than when cells were seeded under static conditions. Moreover, the higher cell densities were attained more rapidly than the lower densities of static culture. In addition, the relative rates of synthesis of elastin and collagen were significantly greater in seeded matrices cultured with agitation than in ones grown statically (82). A lower possible supply of oxygen or other nutrient may have reduced the performance of static seeding methodology. In view of the above referenced studies and similar others, sufficiently intense fluid mechanical forces other than those associated with aeration do affect cells.

Adherent Cells. Quite low shear stress levels, for example, between 0.25 and 0.6 N m^{-2} in laminar flow (91), can interfere with the process of cell attachment to surfaces; however, once the cells are attached and spread out, they may tolerate higher stresses. Laminar shear stress of the order of 0.5–10 N m^{-2} may remove adherent cells from surfaces (92), but even lower values (e.g., 0.1–1.0 N m^{-2}) are known to affect cellular morphology, permeability, and gene expression (92). Sublethal shear stress levels cause no obvious physical damage but are known to produce various biochemical and physiological responses (4,83,84). In studies with rat aortic endothelial cells anchored on the internal walls of glass capillaries, laminar shear stress was shown to affect the cytosolic pH because of preferential leakage of certain ions out of the cells into the buffer saline (84). This reversible permeability enhancement occurred even at stress levels as low as 0.05 N m^{-2} applied over short durations ($\sim$2 min). Similar effects were noted with human aortic endothelial cells but not with human skin fibroblasts or rat intestinal epithelial cells exposed to shear stress levels of up to 1.34 N m^{-2} (84). With rat aortic cells, the flow-induced cytosolic acidification could be maintained for at least 30 min at 1.34 N m^{-2} shear stress level, but the cytosolic pH returned to unstressed levels within about 15 min when the sustained shear stress was $\leq$0.027 N m^{-2} (84).

In studies with anchorage-dependent cells attached to the flat glass walls of a rectangular flow channel, Shiragami and Unno (93) observed increased activity of lactate dehydrogenase (LDH) in cells that had been exposed to a steady-state shear stress of 0.5 N m^{-2} for 12 h: the activity was fourfold greater relative to controls. The LDH activity correlated with the transmission of energy from the fluid to the attached cells (93).

In vivo hemodynamic forces have been implicated in various physiological and pathophysiological processes (4). Atherosclerotic lesions in humans tend to develop in zones of flow separation (4) such as regions of arterial branching and sharp curvature. Arteries adapt to chronic changes in blood flow, increasing in circumference under high flow and declining under reduced flow (4). Shear stress signals transmitted throughout the vascular cell via cytoskeletal and biochemical elements result in changes to structure, metabolism, and gene expression (4).

Shear effects need to be considered also in perfusion devices that retain cells over long periods in continuous culture (94). Shear effects on adhering erythrocytes and leukocytes are treated separately.

Erythrocytes and Leukocytes. Mammalian erythrocytes, or red blood cells, are among the best studied of animal cells. Properties of erythrocytes and other blood cells are relevant to various pathologies and biomedical processes. Erythrocytes and leukocytes, being suspended cells in vivo, likely experience the kind of stresses encountered in bioreactors and various other industrial processing devices; hence, these cells may provide a broad general insight into mechanical behavior of other cells of mammalian origin. The normal lifetime of a human erythrocyte in circulation is 121 days. The normal daily destruction rate is about 2 $\times$ 10^{11} cells (87). Some cells lyse during circulation; most are destroyed in the spleen. Aging cells resist deformation more than the younger ones. All cells unable to deform in the 3-μm openings of the splenic circulation are delayed and ultimately phagocytized (87).

Erythrocytes are highly deformable cells that orient in laminar flow so that least possible surface area of the disc-shaped cell is perpendicular to the flow. The cytoplasm of erythrocytes is a viscous Newtonian fluid (95). The cell membrane behaves as an elastic solid: the cell deforms but almost instantly recovers its shape when the deforming force is removed. The membrane has little resistance to bending but substantially resists increase in area (95). Erythrocytes suspended in turbulent isotonic saline (viscosity $\sim 1 \times 10^{-3}$ Pa s) have been observed to undergo elongation and deformation; however, the cell appears to be less vulnerable to turbulent shear stress than a cell at the same stress level in a viscous suspending medium (87). Based on measurements in turbulent jets, a critical lytic shear stress level of 4,000 N m^{-2} has been reported for very brief exposures ($\sim 10^{-5}$ s) (87). Measurements on erythrocytes of different mammals reveal that the critical shear stress increases dramatically as the cell volume declines.

According to Blackshear and Blackshear (87), "Areal change occurs when the cell membrane is subjected to stress; hemolysis occurs when the area increases by approximately 6.4 percent." Hemolysis is associated at least in part with physical factors and flow, which produce the hemolysis threshold strain in membranes of erythrocytes (87). Once the threshold strain is exceeded, membrane pores open and the membrane eventually tears (87). A briefly (e.g., 10 ms) imposed uniaxial tension of 0.058 N m^{-1} is a sufficient criterion for lysis (87). When the cell is subjected to biaxial stress, a tension of about 0.029 N m^{-1}

may produce lysis. In viscometric stress for prescribed periods, time to lysis declines as the imposed stress is increased. However, it has been conclusively shown that shear stress alone is not a sufficient predictor of hemolysis rate or thresholds (87); cell shape and tumbling also play a role. Membrane shear appears to be far lower when the cells are suspended in higher-viscosity media in viscometric studies.

Erythrocytes allowed to adhere to a glass surface and then subjected to a fluid shear commence movement when the fluid shear force exceeds about 10^{-11} N (95). During this process the cell gradually moves downstream, but the membrane may remain attached to the surface (95). The membrane can be deformed permanently when the deforming force persists for more than a few minutes (95). Shear elasticities of nonnucleated mammalian red cells are generally similar, but elasticities are about an order of magnitude greater for cells of nucleated species. Small amounts of thiol reagents are known to decrease the elongation of human red cells suspended in a shear field (95), presumably by producing some sort of cross-linking. Potentially, this methodology may improve survival of any fragile cell with a significant amount of cross-linkable protein in the membrane. Other additives may render the cell more susceptible to shear damage. Certain chemical lysins and some antigen–antibody reactions cause perforation of the cell membrane and leakage of intracellular material (87). Cholesterol enrichment or depletion of the human erythrocyte membrane does not affect its viscosity.

Hemolysis of red cells is known to occur intravascularly in vivo as well as in various in vitro flow systems (87). A number of studies have correlated hemolysis to flow in pumps, valves, heart–lung machines, blood dialyzers, and transfusion filters. In flow through tubes, wall roughness of the scale of erythrocyte correlates with hemolysis (87). Bubbles trapped in surface imperfections appear to aid lysis. In flow through tubes, hemolysis correlates with the shear rate and the surface-to-volume ratio (87). This type of lysis occurs at shear stress thresholds lower than the ones required to produce lysis in a fluid shear field. Wall contact–associated lysis has been observed to depend on the chemical nature of the wall material. Lysis may decline with time as surfaces become passivated by prolonged contact with plasma proteins (87).

In capillaries of ~1 mm in diameter, an upper limit on the mean tube velocity of 6 m s^{-1} has been suggested for capillaries with sharp-edged entrances, and blood with a viscosity of 4×10^{-3} Pa s (87). This corresponds to a Reynolds number of 1,500 inside the capillary and an average wall shear rate of about 4,800 s^{-1}. Velocities as high as 17 m s^{-1} may be employed inside capillaries with carefully flared entrances (87). As with other animal cells, erythrocytes subjected to bubbling are susceptible to bubble rupture–associated damage, as shown in Figure 6 for porcine erythrocytes. The normalized cell number (Fig. 6) declined more rapidly in a sparged airlift bioreactor in comparison with cells in a surface-aerated shake flask (43). In surface-aerated shake flasks, increasing shaker platform speed over 100–400 rpm increased the specific cell lysis rate (Fig. 7). The slight decline in lysis rate at 100 rpm (Fig. 7) was associated with improved surface aeration

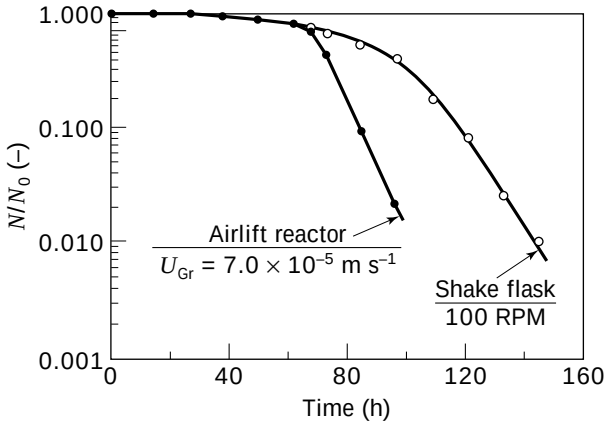


Figure 6. Survival of suspended porcine erythrocytes in a sparged airlift bioreactor and a surface aerated shake flask. Normalized cell count (N/N_0) is plotted as a function of time. *Source*: From Ref. 43.

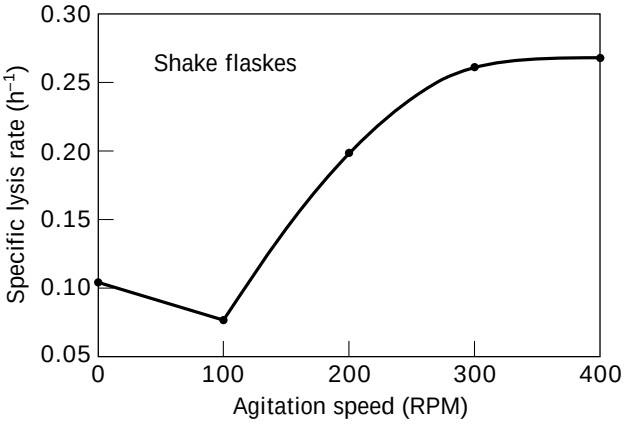


Figure 7. Effect of shaker platform agitation speed on specific lysis rate of suspended porcine erythrocytes in surface aerated shake flasks. *Source*: From Ref. 43.

relative to static flask. Because erythrocytes do not multiply in vitro, the effects of cell-damaging forces are not masked by growth-associated adaptation, changes in cell size, and stages of growth; hence, suitably chosen erythrocytes provide a well-defined cell population for comparative assessment of the damaging potential of various hydrodynamic environments (43).

In comparison with erythrocytes, the cytoplasm of leukocytes has markedly different rheological properties (96), but the properties of the leukocyte membrane are similar to those of the red cell membrane. Leukocytes adhering to vascular endothelium detach when the shear stress is between 26.5 and 106.0 N m^{-2} (96). According to work cited by Prokop and Bajpai (53), a shear stress level of 60 N m^{-2} applied over 10 min should lyse about one-fourth of a leukocyte population. In another study, sublethal shear stresses of 10 and 20 N m^{-2} applied over 10 min in a Couette viscometer affected the biochemical response of human T cells relative to unsheared controls (83). It seems,

therefore, that cells in vivo in circulation are apparently more shear resistant than ones studied in vitro.

Microcarrier Culture. Both stirred-tank and airlift bioreactors (97) may be used for suspended microcarrier culture. Microcarrier-supported cells likely experience hydrodynamic forces of greater magnitude than do freely suspended cells. This is because in highly agitated or aerated systems, the length scale of fluid eddies can be of the same order as the dimensions of microcarriers, resulting in high local relative velocities between the solid and the liquid phases (15,78,98). Furthermore, the carriers have greater inertia than free cells; hence, collisions among microcarriers and between the impeller and microcarriers likely damage attached cells. Similarly, fluid eddy impact and shear stress forces on a high-inertia particle are greater than on free cells.

Resistance of a cell to rupture by impact (i.e., burst resistance) is a possible measure of the cell's survivability in the culture environment. Measurements of burst force are potentially useful also for comparative assessment of cell strength and in establishing culture conditions that give rise to more robust cells. It may eventually be possible to correlate the sensitivity to shear of a cell to its resistance to mechanical rupture (99). Burst strength may be directly relevant in microcarrier culture where impact-associated damage is likely; however, in view of the cytoskeleton-mediated pressure signal transmission to internal parts of a cell, a more likely scenario is that the force required to cause impact-associated damage is far lower than the rupture threshold. This issue notwithstanding, resistance to rupture of a mouse hybridoma grown in serum-containing continuous culture has been measured by squeezing single cells between flat surfaces of micromanipulator arms (100). The bursting strength increased with cell size (101). The bursting force correlated with the initial cell diameter as follows:

$$S = 0.2d_c - 0.1 \times 10^{-6} \tag{31}$$

where the intercept was not significantly different from zero (100). The cell diameter ranged over $\sim(10–17) \times 10^{-6}$ m. Other similar data (101) have been correlated (99) as follows:

$$S = 0.27d_c - 0.86 \times 10^{-6} \tag{32}$$

where the burst strength S is in newtons, and the hybridoma cell diameter, d_c, is in micrometers. The calculated bursting tension had a mean value of $(1.8 \pm 0.5) \times 10^{-3}$ N m^{-1}, which was essentially independent of cell size (100). Because the bursting tension was size independent, the bursting pressure fell with increasing cell diameter. At $(0.8 \pm 0.3) \times 10^{-3}$ N m^{-1}, the calculated mean compressibility modulus of the cells was roughly independent of cell diameter (100).

Sinskey et al. (102) advanced the concept of an integrated shear factor (ISF)—a measure of the strength of the shear field between the impeller and the spinner vessel walls—to correlate shear damage to mammalian cells. The ISF was defined as

$$\text{ISF} = \frac{2\pi N d_i}{d_T - d_i} \tag{33}$$

or, effectively, the laminar shear rate between the impeller tip and the walls. For a range of stirred vessels (0.25–2.0 L, $0.032 \leq d_i$ (m) ≤ 0.085), cell damage occurred once the ISF value exceeded about 18 s^{-1} during culture of microcarrier-supported human fibroblasts (103). Damage could be correlated also with the impeller tip speed, but unlike the ISF, the damaging value of the tip speed depended on the size of the culture vessel (103).

Relying on the earlier work of Nagata, Croughan et al. (103) arrived at the following expression for a time-averaged shear rate:

$$\gamma_{av\cdot T} = \frac{112.8 r_i^{1.8}(r_T^{0.2} - r_i^{0.2})(r_c/r_i)^{1.8}}{r_T^2 - r_i^2} \tag{34}$$

where r_i and r_T are the impeller and the tank radii, respectively. The radius r_c of the formed vortex zone would need to be estimated using Nagata's expression:

$$\frac{r_c}{r_i} = \frac{Re_i}{1{,}000 + 1.6Re_i} \tag{35}$$

The impeller Reynolds number in equation 35 is calculated as follows:

$$Re_i = \frac{N d_i^2 \rho_L}{\mu_L} \tag{36}$$

Equation 34 was developed for unbaffled stirred tanks operated in the turbulent regime (103). For human fibroblasts on microcarriers in various stirred vessels, cell damage occurred when the time-averaged shear rate exceeded about 2.5 s^{-1} (103). For chicken embryo fibroblasts, also on microcarriers, the damage threshold was 6 s^{-1} time-averaged shear rate (103). Cell damage correlated also with the Kolmogoroff eddy length scale (l): for human fibroblasts the cell damage occurred when the length scale declined to approximately below 125 μm, whereas for the chicken embryo cells damage was observed when the l-value declined to below 100 μm (103).

Analyzing data from several sources, Croughan et al. (103) further showed that the specific death rate correlated with the average energy dissipation rate per unit mass; thus,

$$k_d \propto E^m \tag{37}$$

where m was 0.72, 0.76, and 0.82, respectively, for microcarrier-supported Vero cells, similarly supported human fibroblasts, and a freely suspended protozoan. For attaining a more homogeneous shear field in stirred culture vessels (i.e., for the minimum value of the maximum-to-average shear rate ratio) Croughan et al. (103) noted an optimal vessel geometry corresponding to $r_i/r_T = 0.74$.

In microcarrier culture, collisions between microcarriers and interactions between carriers and internals of a reactor are other possible causes of cell damage, particu-

larly in stirred bioreactors (15,104,105). Cherry and Papoutsakis (98,104,105) advanced the concept of "severity of collision" to account for at least some of the damage to cells in suspended microcarrier culture in stirred vessels. Severity of collision combined collision frequency and energy of the impact. Two collision severities were defined: a turbulent collision severity (TCS) for turbulence-associated particle-to-particle impacts, and an impeller collision severity (ICS) for particle-to-impeller collisions; thus,

$$\mathrm{TCS} = \left(\frac{E\rho_L}{\mu_L}\right)^{3/2}\left(\frac{\pi^2\rho_S d_p^5 \varepsilon_S}{72}\right) \tag{38}$$

and

$$\mathrm{ICS} = \frac{9\pi^4\rho_S n_B N^3 d_t^4 d_p^4}{512 V_L} \tag{39}$$

where E is the energy dissipation rate per unit liquid mass, ρ_S and d_p are the density and diameter of the microcarriers, ε_S is the volume fraction of the carriers, n_B is the number of impeller blades, N and d_i are the impeller rotational speed and diameter, and V_L is the volume of the liquid in the vessel. Improved cell growth was observed with smaller microcarriers, as predicted by equations 38 and 39. The specific cell death rate increased with increasing values of TCS and ICS (104,105); however, the influence of hydrodynamic forces on culture performance correlated also in terms of the ratio of Kolmogoroff eddy scale to bead diameter: the specific death rate declined as l/d_p increased. As culture viscosity was raised, damage to cells declined in conformance with equation 38 and in agreement with interpretations based on the l/d_p ratio (105).

Unlike in stirred tanks, animal cell microcarriers suspended in airlift bioreactors under typical operating conditions do not significantly interact with each other or with the walls of the vessel (106); instead, the microcarrier particles follow the laminar streamlines of the fluid (106). Consequently, in airlift reactors at least, the effects of particle–particle or particle–wall collisions on monolayers of cells may be disregarded. Observations of Ganzeveld et al. (106) spanned microcarrier loadings of up to 30 kg m^{-3}, with carriers of 150–300 μm diameter, and 1,030–1,050 kg m^{-3} density. The observations covered a power input value of up to 33 W m^{-3}, which is about the upper limit for cell culture in pneumatically agitated bioreactors. These results applied to a split-cylinder airlift bioreactor with an aspect ratio of 7.6, which would not normally be exceeded in large-scale cell culture systems. As an additional design constraint, the Reynolds number in the riser and the downcomer should not exceed about 3,000, or the flow will be more chaotic (106).

Numerical analysis of forces exerted by laminar flow on anchorage-dependent cells attached to flat surfaces suggests that shear stress between 0.25 and 0.6 N m^{-2} is sufficient to detach round cells, but much higher values are needed to dislodge spread out cells (91). These observations are relevant to microcarrier culture where, during inoculation, the round cells must first attach to microcarriers before spreading on the solid surface. During proliferation, bead-to-bead transfer of cells may also require a level of turbulence that is not so high as to hinder the reattachment process, yet not so low that bead-to-cell encounters are few.

Typically, the shear rate values in airlift bioreactors range over 250–4,000 s^{-1} for operational conditions that are relevant to animal cell culture (14). These shear rates are substantially lower than the $\sim 10^5$ s^{-1} that would be needed to damage cells if a 100-N m^{-2} shear stress value (13,88) is taken as the threshold of mechanical damage. However, based on the shear stress data of Olivier and Truskey (91), during the process of attachment of cells to microcarriers, shear rate levels of 250–600 s^{-1} may well be detrimental. Thus, during initial attachment of cells, the reactor would need to be operated at reduced aeration rates—a practice that is well established through empirical experience, but previously explained only intuitively (14).

Under conditions typical of microcarrier culture in airlift bioreactors, the specific energy dissipation rates are different in different zones of the vessel. The specific energy dissipation rates increase in the following order: downcomer < riser < bottom, for any fixed value of the aeration rate (14). The dissipation rates in all zones decline with increasing loading of microcarriers; however, the prevailing shear rates are not particularly sensitive to the density or the diameter of microcarriers within the ranges that are relevant to anchorage-dependent cell culture. In one case, typical shear rates ranged over 250–4,000 s^{-1} in microcarrier-containing systems, but much higher values, up to 12,000 s^{-1} could occur in solids-free media (14). These values compare favorably with shear rates of $\sim 10^5$ s^{-1} that have been reported as the threshold of damage to cells. Cells in airlift bioreactors would experience a substantial increase in the riser shear rate only when the fluid eddy length–to–microcarrier diameter ratio declines to ~ 1 (14).

Hydrostatic Pressure As the scale of operation increases, suspended animal cells will inevitably experience substantial hydrostatic pressure changes as they move with fluid elements in a large bioreactor. At least for hybridomas, artificially imposed pressure cycling of up to 1.2 atm (18 psig) has shown no harmful effect (79). This level of pressure fluctuation will be encountered in bioreactors that are about 12 m tall; hence, for a 2:1 aspect ratio, an animal cell culture vessel could be scaled up to about 340 m^3. If, as recommended for airlift and bubble column bioreactors (20,21), an aspect ratio of 5:1 is employed, a maximum vessel volume of about 50 m^3 should be feasible. Suspension cell culture vessels in current commercial use do not generally exceed 10 m^3, or about 20% of the feasible volume based on hydrostatic pressure considerations.

Bubble-Associated Damage That sparging damaged cells has been known since the 1960s; however, the specific events responsible for the damage—whether bubble formation at the sparger, bubble detachment, rise through the fluid, or events at the disengagement surface—and the mechanism(s) of damage remained unclear until recently.

The growth rate, viability, and antibody productivity of a mouse hybridoma in a surface-aerated stirred vessel were not affected by the agitation rate, but damage occurred when gas was sparged through the culture (107). In surface-aerated operation, the average and the maximum shear stress values varied over the respective ranges of 0.001–0.4 N m^{-2} and 0.06–6.4 N m^{-2}. Subsequent work showed that bubbles larger than 5 mm were significantly less damaging to cells than smaller bubbles (108). Furthermore, placing the sparger below the impeller enhanced cell damage, presumably because the impeller-associated bubble breakup gave rise to the more damaging smaller bubbles (108).

Cell lines differ tremendously in sensitivity to aeration as was demonstrated by Handa et al. (109). At constant aeration rates in geometrically identical bubble columns, smaller diameter bubbles (0.2 mm in diameter) were shown to be significantly more damaging to a hybridoma than larger bubbles (1.62 mm in diameter). In industrial-scale culture of several hybridomas, Chisti (78) preferred yet larger bubbles (10–20 mm diameter) for aeration. Similarly, Tramper et al. (63) observed that smaller bubbles were more damaging to insect cells than larger ones. However, the effect of bubble size may not be prevalent across cell types. Thus, in another study, baby hamster kidney (BHK-21) cells were not sensitive to the size of gas bubbles (110). Over a superficial gas velocity range of (0.42–8.5) $\times$ 10^{-4} m s^{-1}, higher velocities led to faster decline in cell viability of a hybridoma (109), which is consistent with other similar observations (61).

In commercial stirred-tank culture of several hybridoma lines at up to 0.3 m^3, Chisti (78) employed submerged aeration with bubbles of 0.01–0.02 m in diameter. Because of high rise velocities and mobile interfaces, such bubbles did not carry many attached cells to the surface. In contrast, during aeration with sintered metal spargers that produced slower rising bubbles of ~0.002–0.003 m in diameter, the cells were rapidly carried into a persistent foam layer at the top by a froth flotation mechanism (78). The foam produced by larger bubbles collapsed soon after the bubble rose to the surface. The reactors employed were agitated by magnetically coupled agitators with support bearings submerged in the culture. Reactors with submerged mechanical seals were also used. No loss in culture performance occurred relative to roller bottles. The impeller tip speeds >1 m s^{-1} produced no ill effects. Under typical operating conditions, the length scale of the microeddies was about 130 μm, or about 10-fold greater than the dimensions of the cells (78).

Based on energetics of bubble-associated damage to freely suspended cells, Wang et al. (111) concluded that the principal determinants of cell damage were the cell–bubble encounter rate, the rate of bubble breakup within the fluid, and the bursting rate at the surface. Cell death correlated linearly with specific gas–liquid interfacial area, with the proportionality constant being 0.0125 m h^{-1} for a murine hybridoma (111).

Using an elegant mechanistic analysis, Tramper et al. (62,63) related the first-order rate constant for cell death due to all causes to aeration rate in a bubble column; thus,

$$k_d = \frac{24QV_k}{\pi^2 d_B^3 d_T^2 h_L} \tag{40}$$

where k_d is the first-order death rate constant, Q is the volumetric aeration rate, h_L is the height of fluid, d_T is the column diameter, and V_k is a hypothetical killing volume associated with the bubbles. The latter is actually the total volume of fluid associated mainly with the bubble wakes; the cells entrapped in the circulating wake are carried with the bubble to the surface where most of the damaging events take place. According to equation 40, the specific cell death rate is proportional to the frequency of bubble generation, or $6Q/\pi d_B^3$. Alternatively, because at hydrodynamic steady state the bubble generation frequency equals the frequency of rupture at the surface, the death rate may be interpreted as being dependent on the frequency of bubble rupture and the killing volume associated with the bubbles.

Based on equation 40, the death rate constant should decline with increasing height of fluid, as experimentally confirmed: for otherwise fixed conditions, increasing liquid height in bubble columns improved survival of insect cells (62,63), myelomas, and hybridomas (60,64,110), supporting the view that cell death occurred predominantly in the bubble disengagement zone at the surface (109,110). For a myeloma grown in serum-supplemented medium, the minimum culture height that provided a performance equivalent to that of surface aerated flasks, was ~0.7 m, corresponding to a bubble column aspect ratio of ~14 (64). The 0.05-m diameter column was aerated through a sintered disc sparger (150- to 250-μm pore size). The aeration rate was 5 mL min^{-1} (64), corresponding to a specific power input of about 0.42 W m^{-3}.

If bubble rise were to predominantly contribute to cell death, then the specific cell death rate should be independent of the height of fluid (112). Because this is contrary to experimental observations, the rise events may be disregarded as principal contributors to cell damage (112). Whereas, if gas entry at the sparger contributed predominantly to cell damage, then at a given gas flow rate in the reactor, decreasing the number of sparger holes should enhance the specific cell death rate constant (112). In bubble columns, Jöbses et al. (112) noted that increasing the sparger hole gas velocity over 0.6–2.5 m s^{-1} at a constant volumetric gas flow in the reactors had no effect on the first-order rate constant for cell death; hence, events at the sparger could be disregarded as contributing to cell damage. This process of elimination identified bubble breakup at the culture surface as the principal cause of damage to suspended cells in sparged bioreactors such as bubble columns (112). Equation 40 applies also to airlift devices (61), especially ones in which gas is not carried into the downcomer. Because most agitated tanks used in animal cell culture operate at quite low mechanical power inputs and the impeller does not usually serve as a bubble breaking device, equation 40 should provide a lower limit on k_d also in such tanks.

Using a hybridoma culture van der Pol et al. (60) showed that the first-order death rate constant increased with increasing $1/d_T^2$ in bubble columns, as expected from equation 40. Equation 40 suggests that cell damage in-

creases in direct proportion to the superficial gas velocity in the column, or in direct proportion to the specific power input. Note that increasing liquid level at constant aeration rate in a column of fixed diameter decreases the specific power input (14,20,21). Equation 40 further suggests that for a given V_k, larger bubbles should be less damaging than smaller ones. This agrees with substantial empirical evidence (62,63,78,108,110,113).

That the death rate increases linearly with increasing aeration rate has been proven (62,63,113). Also, the killing volume, V_k, has been shown to be independent of the aeration rate and the height of fluid, so long as the bubble diameter is not affected (62,63). The killing volume apparently depends on bubble diameter, increasing with diameter but in such a way that the ratio V_k/d_B^3 declines as the bubble diameter increases. This has been confirmed by Wu and Goosen (113) using *Spodoptera frugiperda* insect cells in bubble columns that were sparged with bubbles of different mean diameters while maintaining a constant volumetric gas flow rate. Increasing bubble diameters reduced the specific death rate, as shown in Figure 8. In view of the numerous contrary observations, analyses suggesting lesser cell damage with smaller bubbles are clearly flawed.

The dimensionless specific killing volume, or $6V_k/\pi d_B^3$, in equation 40 depends on the concentration of serum in the culture medium. Van der Pol et al. (60) showed that the dimensionless killing volume declined as the serum concentration was raised over the range 0–2.5% (vol/vol); the dimensionless killing volume in the absence of serum was ~2.7-fold that at 2.5% serum in a hybridoma culture (60). This and other similar observations (61) at least partly explain the well-known protective effect of serum. Cell-protective effects of other proteins may be similarly explained. Serum-free media with 1 g protein L^{-1} (equivalent to supplementing the medium with 2% vol/vol serum) are commercially used in sparged airlift culture vessels (79).

Why rupture of larger bubbles is less damaging to cells than breakup of smaller ones has become clear through analyses of mechanics of rupture (114,115). Numerical simulations and experimental observations of bubble rupture at an air–water interface confirm that smaller bubbles rupture much more violently than larger ones (114,115). A bubble reaching the surface rises to different heights above the flat liquid surface before rupturing. For a bubble that is stationary at the surface, a part is always submerged below the flat surface of the liquid. The degree of submergence is generally greater for smaller bubbles. During rupture, the liquid drains from the film between the raised liquid surface and the bubble (116,117). As the film thins, a hole develops in the center of the dome, and the film rapidly retreats into the bulk liquid. The rupture of the film and consequent elongational and accelerational flow have been postulated to damage cells in the film and those adhering to the bubble cavity (116). A cell on the surface of a rupturing film would apparently not experience any motion until struck by the torroidal ring of fluid at the receding edge of the film (116). The force experienced by the cell will depend on the film thickness at the instance of rupture (116). Rupture of thinner films is likely to be less damaging than the breakage of a thicker film. Once the bubble ruptures, fluid rapidly accelerates and moves down the interior walls of the bubble cavity. The flow from around the walls impacts at a stagnation point located immediately below where the bubble had been. The high pressure produced by the impacting streams forces a jet of fluid upward into the gas phase above the liquid surface. A second opposing jet is forced into the fluid below the stagnation point. Intense accelerational flows around the bubble cavity and formation of liquid jets occur mostly during rupture of small bubbles. Bubbles larger than about 6 mm rupture much less violently.

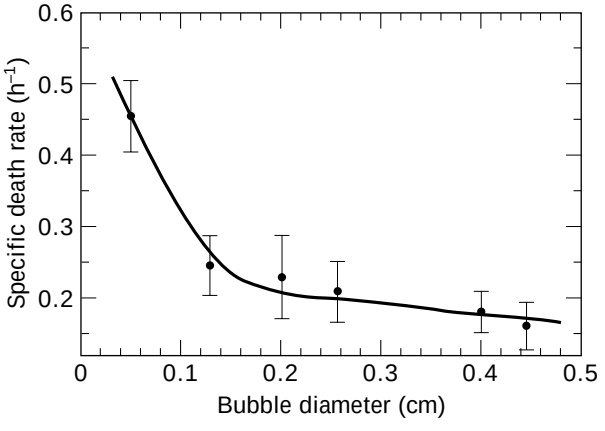


Figure 8. Effect of bubble diameter on the specific death rate of *S. frugiperda* insect cells in bubble columns aerated at a constant volumetric gas flow rate of 10 mL min^{-1}. *Source*: Adapted from Ref. 113.

According to Boulton-Stone and Blake (114), "one of the most important factors determining the motion following film rupture, in terms of the energy released, is the height of the top of the bubble above the equilibrium free-surface" of the liquid. Bubbles rupturing from a lower position beneath the equilibrium surface of the liquid release more energy because of their higher internal pressure. The energy is released as high-speed liquid jets. Bubbles larger than about 6 mm bursting at the surface do not form significant jets (114) which is consistent with numerous observations dating from 1950s and earlier. When a jet is produced, the speed of the jet declines as the size of the rupturing bubble increases: for example, for air bubbles rupturing on water, a maximum jet speed of 6.4 m s^{-1} has been calculated for rupture of a 1-mm bubble; for rupture of a 6-mm bubble the maximum jet speed declines to 0.94 m s^{-1} (114). In addition to a liquid jet rising into the atmosphere above the liquid, a second liquid jet moves down into the fluid from the region that had been below the base of the bubble. The maximum pressure produced during bursting of bubbles of various sizes declines with increasing bubble diameter, as illustrated in Figure 9. The maximum energy dissipation rate occurs just beneath the bubble immediately before the jet forms (114). The calculated maximum energy dissipation rates for the smallest bubbles are equivalent to a stress of about 10.4 N m^{-2} (114). The maximum energy dissipation rates decline with increasing bubble diameter, as depicted in Figure 10.

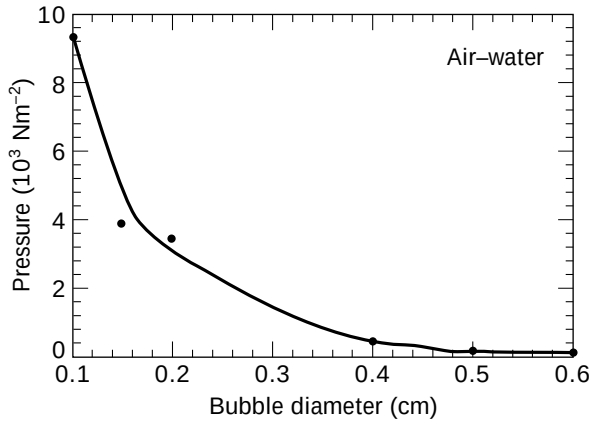


Figure 9. The maximum pressure produced during bursting of air bubbles of various sizes on the surface of water. *Source*: Adapted from Ref. 114.

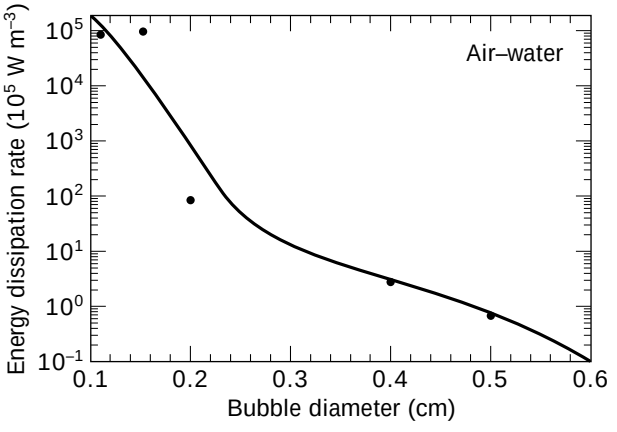


Figure 10. The maximum energy dissipation rates produced during bursting of air bubbles of various sizes on the surface of water. *Source*: Adapted from Ref. 114.

Even a violent bubble rupture event would be inconsequential if there were no cells in the vicinity. Unfortunately, cells attach to rising bubbles, and the fluid behind bubbles is especially rich in cells. That some animal cells will attach to rising bubbles has been shown with *Trichoplusia ni* and *S. frugiperda* insect cells (116,118). The cells adhered to rising microbubbles (diameter $\leq$3 mm) of the type used commonly in froth flotation. Hydrophobic interactions are apparently responsible for cell attachment—a mechanism observed also with ore particles attaching to bubbles during foam flotation. Although, very small bubbles are generally not suited to aeration of cell cultures (78), larger bubbles also transport cells to the surface, but not predominantly as adhering cells. Instead, the cells are trapped in the circulating wake carried behind large bubbles. The fluid associated with the wakes is known to be richer in particles than the bulk fluid.

More than 95% of the cells in the liquid ejected as jet during rupture of bubbles have been shown to be nonviable even when the cells in the bulk culture contained fewer than 10% dead cells (115). Clearly, the cells carried in the bubble wake are destroyed by the rupture event. Based on simulations of rupture of bubbles with diameters of 0.77 mm, 1.70 mm, and 6.32 mm, Garcia-Briones et al. (115) calculated maximum energy dissipation rate values (W m^{-3}) of 9.52×10^7, 1.66×10^7, and 9.40×10^4, respectively. Comparing these values with energy dissipation rates that have damaged cells in various unaerated flow devices (capillaries and viscometers), Garcia-Briones et al. (115) noted that values of the order of 5.81×10^2 W m^{-3} did not damage the cells significantly, whereas appreciable damage occurred when the energy dissipation rate exceeded about 2.25×10^4 W m^{-3}. Clearly, therefore, a 6.32-mm bubble still has a significant potential to damage cells, and a smaller bubble would produce instantaneous damage. For example, in gas-free capillary flow in experiments reported by Augenstein et al. (16), more than 80% of the cells died within 0.3 s of exposure to energy dissipation rates of about 4.80×10^7 W m^{-3} (115). Interpreting data from several sources, Garcia-Briones et al. (115) further note that average energy dissipation rates of the order of 2,990–29,000 W m^{-3} have generally produced a significant decline in growth rate in impeller-agitated bioreactors. These dissipation rates are average values; the cells probably experienced much higher local energy dissipation rates in the vicinity of the impellers.

If stable foam is not allowed to form, and relatively large bubbles (diameter $\approx$ 10–20 mm) are used for aeration so that the energy dissipated during rupture remains below the cell damaging threshold, then the cells carried in the wake will not be damaged or removed into a foam layer. For bubbles of a given size, the volume of the circulating wake can be reduced by increasing the viscosity of the culture broth. Small increases will reduce the rise velocity of the bubbles, and the circulation in the wakes will be slowed down. Consequently, slight enhancements in viscosity could reduce the transport of particles in the wake. This may indeed be one explanation of why viscosifying additives have been frequently observed to reduce cell damage. Also, for a given volumetric aeration rate, use of larger bubbles reduces bubbling frequency; hence, the frequency of bubble rupture is reduced.

Based on the studies reviewed, several important observations can be made regarding freely suspended animal cells in sparged and agitated bioreactors: (*1*) the aeration rate should be kept as low as possible; (*2*) the mean bubble size should larger than 7 mm, preferably about 10–20 mm; (*3*) the location of the sparger should be such that the rising bubbles do not interact with any impellers; (*4*) with small airlift and bubble column reactors the aspect ratio should be about 14; however, in larger columns where the sparged portion of the cross section can be a small fraction of the total cross-sectional area, more realistic aspect ratios of about 6 or 7 are satisfactory; and (*5*) when an impeller is used, the average energy dissipation (input) rate should remain below about 1.0×10^3 W m^{-3}. Observations (*1–5*) are consistent with industrial practice (78). The sole purpose of the impeller should be to suspend the cells and mix the fluid gently so that the oxygen transferred from the bubbles is distributed throughout the vessel. Consistent with this purpose, the impeller should be of a kind that

does not produce excessively high local rates of energy dissipation. For example, relatively large axial flow hydrofoils are the preferred impellers (30,78). In addition, a suitably selected additive such as Pluronic F68 should be used whenever feasible.

Mitigation of Shear Damage Various approaches have been used to reduce bubble rupture–associated damage to cells. Some common approaches are the use of cell protective additives as noted here; compartmentalization of the culture into cell-free and cell-containing zones, with sparging confined to cell-free regions (80,81); use of bubble-free oxygenation through microporous or diffusion tubing made of silicone rubber or other polymer; use of liquid oxygen vectors such as perfluorocarbons; and surface aeration (see MASS TRANSFER). Alternatives to bubbling are suited typically only to specific cases; direct sparging remains the predominant method of aeration of cell cultures, and it is not likely to be displaced in the foreseeable future (76).

One adaptation of submerged aeration is the "bubble bed" bioreactor, which can substantially reduce surface rupture of bubbles in short-term batch culture (119). The bubble bed device (Fig. 11) is similar to a concentric draft-tube airlift bioreactor, but a downward pumping axial flow impeller located in the draft tube circulates the fluid. The downward liquid flow in the draft tube is sufficiently fast that gas bubbles injected into the draft tube cannot escape; hence, the bubbles do not rupture at the surface of the culture. Pure oxygen or oxygen-enriched gas is used to reduce the needed gas injection rate to very low values. The gas holdup in the downcomer increases during batch operation, and a noncirculating bed of bubbles develops. The cross-sectional area of the draft tube increases downward (Fig. 11) so that the superficial velocity of the liquid declines as it moves down; consequently, the downflowing liquid cannot drag the bubbles out of the draft tube and into the bubble-free annular zone. According to Sucker et al. (119), the cell growth and antibody production in bubble bed devices is comparable to values obtained in surface-aerated spinner flasks even when the sparged culture in the bubble bed units is not supplemented with serum, Pluronic F68, or other protective additives.

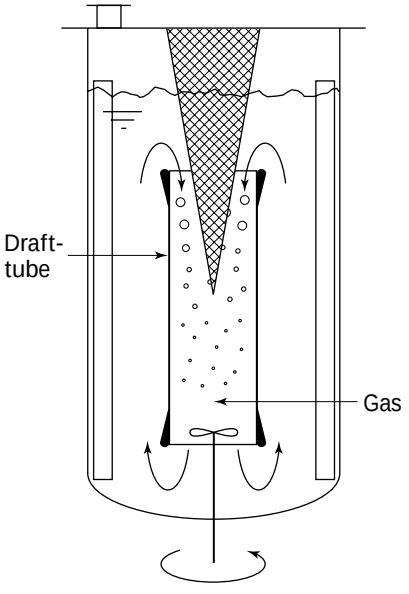


Figure 11. A bubble bed reactor.

Shear Protectants. Bubble-associated damage can be significantly reduced by adding shear-protective substances to the culture medium. Many additives have been examined for shear-protective function (120). Some of the protective additives are noted in Table 4. The protective effect depends on the properties of the specific additives, the concentration used, the type of cell, and the mechanism(s) of protection. Some common protective agents and their mechanisms of protection are treated next.

Serum. A concentration-dependent protective effect of fetal calf serum, horse serum, and others, has been widely observed in aerated and bubble-free suspension cultures of many types of animal cells (41,60,61,66,110,120–122). Typically, protection increases with increasing serum concentration until about 10% (vol/vol) serum. In one study reduction in fetal calf serum level from 2.5% to 0% substantially increased the death rate constant of a hybridoma in aerated culture (60). Investigations revealed that serum had no long-term biological effect, but only a direct nonspecific physical protective effect (60). Thus, cells growing in low-serum medium were immediately protected upon transfer to a medium with 2.5% serum. A physical protective effect of serum has been claimed to originate in its ability to reduce plasma membrane fluidity (41,121). According to Martens et al. (61), serum has an immediate physical protective effect, but also a physiological protective effect that takes longer to become effective. The latter finding concurs with the results of Michaels et al. (122). In bubble-free experiments in a Couette viscometer (shear rate = 5,000 s^{-1}, shear stress = 5 N m^{-2}), prolonged exposure of cells to fetal bovine serum (FBS) reduced their shear sensitivity in laminar flow studies (122); however, shorter (30–120 min) exposure to FBS did not affect shear tolerance of cells. Based on these results, the protective effect of FBS was claimed to have both physical and biological (biochemical) origins (122).

The precise nature of the physical protective effect of serum is not clear. Supplementation with serum enhances the viscosity of the culture medium, but the effect is small (92,123). Nevertheless, Croughan et al. (66) and others (67,104) have associated the observed protection to the turbulence-dampening effect of serum. In contrast, other evidence suggests a lack of relationship between the amount of protection afforded by the serum and enhance-

Table 4. Media Additives Associated with a Cell-Protective Effect

Serum	Pluronics
Proteins (e.g., serum albumin)	Polyvinyl alcohols
Derivatized celluloses (e.g., methyl cellulose)	Polyvinyl pyrrolidone
Derivatized starches (e.g., hydroxyethyl starch)	Polyethylene glycols
Dextrans	

ment of viscosity (124,125). In addition to dampening turbulence, the serum proteins may coat the cells so that the fluid eddies no longer penetrate to the vicinity of the cell membrane (66). In one study, increasing the concentration of fetal bovine serum over 0–10% (vol/vol) reduced plasma membrane fluidity of a hybridoma line, making the cells more resistant to shear damage (41,121).

Pluronic® F68. Another well-known shear protectant is the nonionic surfactant polyol Pluronic® F68, a block copolymer of poly(oxyethylene) and poly(oxypropylene). Highly pure Pluronic F68 is added to cell culture media, typically at 0.5–3 kg m^{-3} (92,109,110,112,120,122). The protective effect of the surfactant is concentration dependent, increasing with concentration but leveling off at ~0.5 g L^{-1} (Fig. 12).

The protective effect of Pluronic F68 is commonly associated with its ability to suppress attachment of cells to bubbles (110): in the presence of the surfactant, fewer cells are carried to the surface where much of the cell damage occurs in aerated culture. However, evidence exists for other protective mechanisms. For example, a cell-protective effect of Pluronic F68 has been observed even in unaerated culture (41,112). In one case this effect was observed in a strongly agitated culture sample (112). Although no bubbles were present, a deep vortex was drawn into the fluid by the marine impeller used. Extensive cell damage occurred, but the damage was significantly reduced when the culture was supplemented with Pluronic F68. Based on these observations, Jöbses et al. (112) concluded that the protective effect of Pluronic F68 was due to its direct influence on cells and to some effect of the surfactant on the gas–liquid interface. One possible explanation for the protection observed by Jöbses et al. (112) in the absence of the bubbles follows: the damage may have been occurring at the rapidly moving surface of the vortex (126), and supplementation of the culture with Pluronic F68 may have reduced the surface adherent tendency of the cells; hence, fewer cells may have attached to the surface during its formation at the periphery of the vessel.

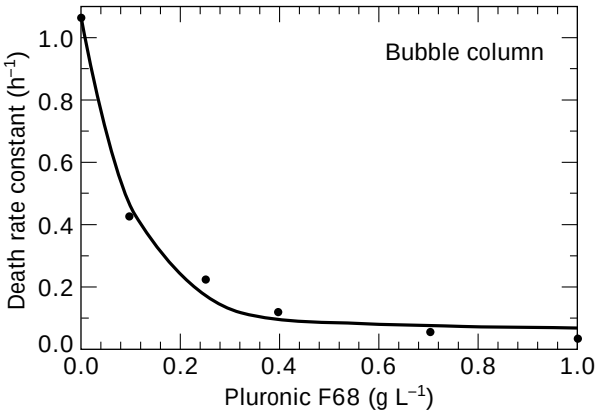


Figure 12. Cell protective effect of Pluronic® F68 in a bubble column. The specific first-order cell death rate constant is plotted as a function of the concentration of Pluronic F68 in the medium. All data were obtained at a constant specific power input of ~40 W m^{-3}. *Source*: Adapted from Ref. 112.

Using flow cytometry, Al-Rubeai et al. (86) showed that leakage of a positively charged dye, fluorescein diacetate, from murine hybridomas increased with rising intensity of agitation, suggesting a possible link between the permeability of the cell membrane and the shear rate. However, because intense agitation was accompanied by entrainment of gas bubbles into the culture (86), the observed leakage could not be conclusively associated with fluid turbulence. The authors noted that for otherwise fixed agitation conditions, the leakage of the dye reduced when the culture medium was formulated with 1% Pluronic F68. The protective effect of the surfactant was apparently linked at least partly with its biochemical association with the cell membrane. The agitation- and vortex-associated hydrodynamic forces were reported to also affect the antigen receptors embedded in the membrane of the hybridoma (86). The number of active surface receptors declined with increasing intensity of agitation and the accompanying increase in the size of the vortex; however, the viability of the cells was unaffected. Pluronic F68 had a protective effect on the receptors.

Pluronic F68 is known to reduce the plasma membrane fluidity (PMF) of cells, and this has been suggested as a possible mechanism of protection (41,121). Other additives that affect PMF also affect a cell's shear tolerance. For example, in one case cholesterol enrichment of the culture medium reduced membrane fluidity and enhanced the shear resistance of hybridomas, whereas supplementing the culture with benzyl alcohol increased PMF and made cells more shear sensitive relative to controls (41,121).

Other protective mechanisms speculated for Pluronic F68 include its stabilizing action on the gas–liquid interface and consequent slower film drainage during bubble rupture (110), improved nutritional transport because of a reduced cell–fluid interfacial tension, and incorporation of the surfactant into the cell membrane accompanied with membrane stabilization (127,128). Although the protective effect of Pluronic F68 is said to be solely due to physical factors (122), the surfactant is known to produce physiological and biochemical effects in at least some cases. Growth suppressing and stimulatory effects of Pluronic F68 have been seen even in some static cultures. Such effects are sometimes associated with impurities in Pluronic F68 samples. Pluronic F68 affects the permeability of some cells. Exposure of *Saccharomyces cerevisiae* to 0.2% (wt/vol) Pluronic F68 for 30 min at 20 °C made the cells more permeable, and they became susceptible to otherwise nonlethal doses of certain antibiotics (129). The mean activity of alcohol dehydrogenase in permeablized cells was more than 3.5-fold of that in normal cells. Although Pluronic F68 could permeablize the yeast cells, it did not inhibit growth or affect cell viability relative to controls when added to culture media at up to 1% (wt/vol), which was the highest concentration tested (129).

Other data suggest that the cell-protective effect of Pluronic F68 may not be quite general (43). For porcine erythrocytes suspended in isotonic buffer, Zhang et al. (43) noted that addition of Pluronic F68 actually increased cell lysis in agitated environments relative to results in the surfactant-free medium. Whether Pluronic F68 incorporates into an erythrocyte's membrane is not known; how-

ever, cholesterol enrichment or depletion of erythrocyte membrane is reported to not affect membrane viscosity and presumably, therefore, fluidity. In addition to Pluronic F68, several other pluronics are known to have a shear-protective capability.

Poly(ethylene glycol). Poly(ethylene glycol) (PEG) supplementation of the medium has a protective effect on some cells but not on others. In one study (129) that employed different PEGs with molecular weights of 400–35,000 Da in the concentration range of 0–2 kg m^{-3}, the polymer had a protective effect in sparged cultures when the molecular weight and the concentration exceeded 1,000 Da and 0.25 kg m^{-3}, respectively. The extent of protection afforded depended on the molecular weight and the concentration used. Under the best conditions, the PEG supplementation reduced the first-order cell death rate constant to 0.5% of the value in the PEG-free medium (130). The PEG additives affected the surface tension of the culture medium; however, in extreme cases the surface tension varied only between $\sim 60 \times 10^{-3}$ and 70×10^{-3} N m^{-1} (130). The pattern (but not the magnitude) of variation in surface tension with addition of PEG was similar to the pattern of change in the death rate constant. This led the authors to suggest that the presence of PEG at the gas–liquid interface was somehow related to the protective mechanism.

Michaels et al. (122) showed that, in sparged bioreactors, supplementing the culture medium with Pluronic F68, fetal bovine serum (FBS), or PEG (8,000 Da, 1 kg m^{-3}) protected cells irrespective of whether the CRL-8018 hybridomas were grown in the presence of the additive, or the additive was added after the damaging levels of agitation were imposed.

Composite Particles and Cell Aggregates

Microbial Biofilms on Suspended Particles. Gas–liquid–solid three-phase suspension reactors are increasingly being used in wastewater treatment (20,131). The suspended solid phase (usually sand, glass beads, or carbon particles) supports a biofilm of waste-degrading microorganisms (132). The hydrodynamic environment of such systems must be tailored to control the biofilm thickness without dislodging it altogether (20). Biofilm slurry reactors are conceptually similar to microcarrier-supported suspension culture of animal cells, but there the similarity ends. Unlike animal cell microcarrier suspensions, the biofilm reactors are much more turbulent, the solids loading is generally higher, the density difference between the solid and the liquid phase is much greater, and the biofilm is typically four or five times thicker than the monolayer of cells on microcarriers.

In studies of degradation of naphthalene-2-sulfonate by *Pseudomonads* immobilized as biofilms on sand particles suspended in an external-loop airlift bioreactor, Wagner and Hempel (133) noted that at solids loadings of 20–30 kg m^{-3}, only about 15–20% of the particles were covered with biofilm, but high solids loadings were necessary to control the film thickness at 75 ± 10 μm by providing a high-attrition environment (133). These observations were for particles about 200 μm in diameter, covered with a biofilm ~ 75 μm thick, suspended in the reactor under fairly turbulent conditions (230–440 W m^{-3} specific power inputs). Even when oxygen limitation is not an issue, diffusion of other substrates limits the useful thickness of microbial biofilm to less than 400 μm in most wastewater treatment operations (131).

The development of microbial biofilms on static surfaces and carriers suspended in airlift bioreactors has been investigated quite extensively (6,7,132,134–140). In one case, 21 carriers were tested at a constant carrier loading of 125 kg m^{-3} (132). The superficial gas velocity was constant at 0.038 m s^{-1}. The hydraulic retention time of wastewater was one or two hours. Based on the tests, carrier diameter and surface roughness were the major factors that influenced biofilm formation (132). Smaller particles, 0.1–0.3 mm in diameter, tended to show better biofilm development. This was said to be due to more frequent collisions among larger particles, causing biofilm detachment (132); however, with larger particles, a greater *severity* of collisions would have been a better explanation (20). For equal mass loadings of solids in two systems, the system with larger particles should have fewer particles per unit volume and, therefore, a lower frequency of particle–particle collisions. Cell disruption experiments in agitated slurries of glass beads confirm this (11). The explanation for the impact of surface roughness on biofilm formation lay in the fact that a rough surface provided many relatively protected, low-shear sites where microbial colonies could take hold in the process of gradually spreading over the whole surface. Based on the criteria of surface roughness, particle size, cost, and the requirement of a high-density carrier for ease of separation, basalt (ρ_S = 2,900 kg m^{-3}) was the best carrier (132).

Irrespective of carrier type, biofilm formation is favored at relatively high dilution rates when any detached cells are rapidly washed out (132,136). Tijhuis et al. (136) have shown that during start-up at high dilution rates, the concentration of bare carriers declines as an increasing number are covered with the biofilm. The total amount of the biomass in the reactor increases, but the amount per carrier remains constant until most of the carriers are covered. With few bare carriers remaining, the extent of abrasive loss of film declines, and the film thickness begins to increase until a stable thickness is reached (136). These observations suggest that varying the reactor operation to suit different stages of film development may be worthwhile (20).

In other similar work, detachment of biomass from nongrowing biofilms supported on carriers suspended in four configurations of draft-tube sparged internal-loop airlift reactors was examined (139). The superficial aeration velocity in the riser varied over the approximate range 0.03–0.09 m s^{-1}. The detachment of biomass from biofilm-covered carriers in the presence of bare solids was dominated by collisions between bare carriers and the covered particles (139,140). The concentration of bare solids and the particle roughness strongly influenced the detachment rate (139). For otherwise identical conditions, biofilm detachment due to bare carriers increased in the following order: glass, sand, basalt. The glass particles were the most round and smooth; the basalt carriers were the

roughest. Whereas biofilms develop more easily on rough surfaces (132), once developed, the film is readily damaged by biofilm-free rough carriers. Gjaltema et al. (139) observed that a "change in flow regime from bubbling to slug flow considerably increased the detachment rate." This change apparently occurred at a U_{Gr} value of $\sim$0.06 m s^{-1}. The "slug flow" regime noted by Gjaltema et al. (139) was more likely the highly chaotic churn turbulent regime in which particle-to-particle collisions are more frequent and severe than in the bubble flow regime. The bottom clearance of the draft tube had no effect on detachment rate (139), but only a small range of the A_d/A_b ratios was investigated. The volume fraction of carriers was 3.3–3.8%, and other properties were: basalt, d_p = 0.32 mm, ρ_S = 3,010 kg m^{-3}; sand, d_p = 0.31 mm, ρ_S = 2,606 kg m^{-3}; and glass, d_p = 0.28 mm, ρ_S = 2,887 kg m^{-3} (139). Another similar study focused on the effects of particle shape, structure, and concentration on biofilm detachment (140).

Strategies for development of biofilms with nitrification activity on carriers suspended in draft-tube sparged internal-loop airlift reactors were investigated by Tijhuis et al. (141) and van Benthum et al. (138). The development of the nitrifying films was similar to what had been reported for heterotrophic biofilms (134–137); however, the nitrifying films were more robust (141). This suggests that biofilms formed by different microorganisms can have markedly different physical properties and strength of attachment to substrate (20); hence, the operating regime of the reactor and, in extreme cases, the design peculiarities are likely to be affected.

In further studies of the dynamics of biofilm detachment from carriers suspended in airlift reactors, Tijhuis et al. (135,137) noted that the common assumptions of uniform growth and detachment of biomass over the surface of the film under fixed hydrodynamic conditions were inconsistent with observations. Instead, nonuniform local detachment of biofilm with formation of cracks and fissures on the surface and rapid filling of the spaces formed with new growing biomass was proposed as a more realistic picture of film dynamics (135).

During development of mixed bacterial biofilm on a static glass surface in a rectangular turbulent flow cell, Stoodley et al. (7) noted that parts of the biofilm sloughed off at a wall shear stress level of 1.32 N m^{-2}. In the region of detachment, pieces of the biofilm were still attached to the surface, suggesting that the glass-to-biofilm bond was stronger than the biofilm structure (7). Relative to the clean flow cell, the pressure drop through the channel was higher in presence of the biofilm (7), presumably because a patchy development of the film increased the effective surface roughness.

Flocs and Agglomerates. Microbial flocs are encountered in several commercially relevant processes, especially in wastewater treatment and beer fermentation with autoflocculating yeasts. Similarly, agglomerates of protein precipitates form during certain recovery operations. The size of flocs and agglomerates often needs to be preserved or increased during processing. Stable size of aggregates and flocs, and whether the flocculated mass further increases in size or breaks up, depends on the level of turbulence in the suspending fluid. Turbulent stress levels below about 0.75 N m^{-2} are generally too mild to break aggregates (142). Physical disruption of aggregates commences at stress levels of $\sim$2.6 N m^{-2}, and especially when the stress level exceeds about 5 N m^{-2} (142). In a turbulent field, aggregates may deform or breakup because of the influence of instantaneous velocity differences acting on different parts of the floc (143). Also, turbulent drag forces originating from the local motion of the fluid relative to the aggregate may contribute to fragmentation (142). The shear force across the aggregate due to instantaneous velocity difference is related to the energy dissipation rate in the fluid and the size of the aggregate (142); thus,

$$F_s = \frac{\pi}{4}(\rho_L \mu_L E)^{1/2} d_F^2 \tag{41}$$

The dynamic pressure fluctuations acting on opposite sides of the floc create a normal stress. Assuming that the eddy size is in the Kolmogoroff microscale range and that the particle dimensions are similar to those of the eddies, the normal force can be expressed as follows (142):

$$F_n \propto \left(\frac{E}{\mu_L}\right)\rho_L^2 d_F^2 \tag{42}$$

The turbulent drag force acting on an aggregate may be calculated using the expression (142)

$$F_D \propto \frac{C_D}{2}\left(\frac{d_F}{2}\right)^3 (\Delta u d_F)^2 \rho_L \tag{43}$$

where Δu is the difference in velocities of the particle and the fluid, and C_D is the drag coefficient. For low-density particles in the Stokes regime ($Re < 1$), the drag coefficient equals $24/Re$, where Re is the Reynolds number, or $(\rho_L \Delta u d_F)/\mu_L$. Reynolds shear stresses (i.e., turbulent stresses associated with the fluctuating component of the mean velocity) may be calculated using the equation (144)

$$\tau = 0.37\rho_L\left(\frac{E\rho_L}{\mu_L}\right)d_p^2 \tag{44}$$

which applies to stresses around a spherical particle. Reynolds stresses are independent of viscosity.

The maximum stable size of the aggregate in a turbulent field can be established by a force balance on a particle: at stable size the instantaneous turbulence force accelerating the particle just equals its mechanical strength (142). In the inertial subrange of turbulence when the eddy size is greater than the Kolmogoroff scale (l),

$$d_{F\ max} \propto \left(\frac{\rho_L \sigma_M}{\mu_L}\right)^{1/3} E^{-1} \tag{45}$$

whereas in the viscous subrange of turbulence when the eddy size is $\leq l$, we have

$$d_{F\ max} \propto \left(\frac{\sigma_M \mu_L}{\rho_L^2 E}\right)^{1/2} \quad (46)$$

where σ_M is the mechanical strength of the aggregate. In practice, the maximum stable floc size has correlated with the energy dissipation per unit mass of fluid,

$$d_{F\ max} \propto \Gamma E^{-m} \quad (47)$$

where the coefficient Γ depends on the nature of the particle, and m is usually between 0.2 and 0.4 (145).

In stirred tanks, the maximum floc size for particles of uniform composition has been theoretically and experimentally correlated with the rotational speed of the impeller (143); thus,

$$d_{F\ max} \propto N^{-1.5} \quad (48)$$

which applies to particles smaller than the microscale of turbulence. The dependence shown in equation 48 can be expressed in terms of the energy dissipation per unit mass; hence,

$$d_{F\ max} \propto E^{-0.5} \quad (49)$$

Expression 49 is consistent with that reported by Thomas (143), but the exponent on E is somewhat stronger than the values noted by Gregory (145) (see equation 47). In precipitation vessels, shear rate values of $\sim 80\ s^{-1}$ appear to be a damaging threshold for soya protein precipitate particles of 20–30 μm initial size (5). Particles exposed to approximate shear rates of $\geq 80\ s^{-1}$ over 10–12 min would be reduced to 5–8 μm. The shear rate numbers were calculated from the specific energy dissipation rates in the vessels, as follows:

$$\gamma = \left(\frac{\rho_L E}{\mu_L}\right)^{0.5} \quad (50)$$

Protein precipitates are easily damaged by pumping. Moyno screw pumps, centrifugal pumps, and gear pumps all damage precipitates to similar levels (5). Peristaltic pumps are apparently the only suitable ones.

Fragmentation of mycelial pellets of the molds *Lentinus edodes* and *Aspergillus niger* in a standard Rushton turbine stirred tank was examined by Taguchi et al. (145). Two types of pellet fragmentation patterns were noted: (*1*) decrease in diameter by erosion of surface, and (*2*) direct rupture of the spherical pellet. The rates of these processes were expressed as

$$-\frac{dd_p}{dt} = k_j (N d_i)^{5.5} d_p^{5.7} \quad (51)$$

for pellet surface erosion, and

$$-\frac{dN_p}{dt} = k' N_p = \alpha (d_p^{3.2} N^{6.65} d_i^{8.72})^{5.5} N_p \quad (52)$$

for rupture of the particle. In equations 51 and 52, d_p is the pellet diameter at time t, N is the rotational speed of the impeller, d_i is the impeller diameter, N_p is the concentration of unbroken pellets at time t, and k_j, k', and α are constants. Clearly, the rates of pellet erosion and rupture depend strongly the agitation intensity–associated factors (i.e., on N and d_i) and the dimensions of the pellet.

In submerged stirred-tank culture of the filamentous fungus *Aspergillus awamori*, Cui et al. (147) correlated the pellet size with the specific energy dissipation rate; thus,

$$d_p \propto E^{-0.16} \quad (53)$$

The E values varied over 0.2–5.7 W kg^{-1} (146), whereas the pellet diameter ranged roughly over 1–2 mm. The length of the "hairy" hyphal extensions from the well-defined surface of the denser pellet also correlated with the specific energy dissipation rate (147); thus,

$$L_h = 1.17 \times 10^{-4} E^{-0.25} \quad (54)$$

where L_h (m) is the length of the hairy extensions.

Bioactive Proteins

Enzymes are not generally susceptible to the shear rates that are commonly encountered during processing including high-shear operations such as cell disruption (1). Exceptions occur when shear stresses are applied in the presence of gas–liquid interfaces, at which proteins tend to unfold (1,8,148). Also, membrane-associated proteins are more prone to shear damage (1). In addition to processes such as high-pressure homogenization, high-shear processing of proteins occurs during liquid–liquid emulsification, which is used to produce microencapsulated proteins for drug delivery systems (149). Effects of extremely high shear rate levels ($>10^5\ s^{-1}$) on recombinant human growth hormone (rhGH) and recombinant human deoxyribonuclease (rhDNase) were examined by Maa and Hsu (149). The rhGH and rhDNase had respective molecular masses of 22.13 kDa and 32.74 kDa. The shearing was done at a controlled temperature of 20 °C. The aggregation and the bioactivity of the two proteins were not significantly affected at shear rates greater than $10^5\ s^{-1}$ applied over 60 min in the absence of gas–liquid interfaces (148,149). Thermodynamic properties of the proteins were determined before and after shearing using scanning microcalorimetery (SM). Alterations in secondary and tertiary structure of a protein are manifested in its scanning calorimetery data. The SM thermograms of rhDNase were not significantly affected by shearing; however, for rhGH, the SM indicated possible conformational changes (149). The silver-stained SDS-PAGE gels of rhGH showed no significant differences in the high molecular weight region, but low molecular weight fragments were observed in the sheared sample, which suggested some clipping of rhGH in the intense shear field. No significant changes were noted in the SDS-PAGE gels of sheared and unsheared rhDNase (148). If a globular protein's molecular weight is assumed to indicate it's size, the rhDNase molecule should be larger by a factor of approximately $(32.74/22.13)^{1/3}$, or 1.14, relative to the rhGH. It seems therefore that the molecular size is

not the sole determinant of a protein's susceptibility to damage; the specific structure or morphology also appear to play a role.

The damaging effect of shear in presence of gas–liquid and solid–liquid interfaces may be reduced by supplementing the protein solution with a surfactant. Added surfactants tend to reduce gas–liquid interface–induced aggregation and precipitation of rhGH (148). Proteins also adsorb on hydrophobic solid surfaces. Adsorption of insulin at hydrophobic surfaces, accompanied by denaturation, is well known. Similarly, tPA is known to adhere to solid surfaces (150), but the surface adherent tendency is reduced in presence of surfactants such as Tween-80 (polyoxyethylene sorbitane monooleate) added at 0.01% typical concentration (150). Shear susceptibility of proteins is further discussed elsewhere (5,151).

CONCLUDING REMARKS

Biocatalysts, especially molds, filamentous bacteria, certain microalgae, and plant and animal cells are susceptible to damage by hydrodynamic and mechanical forces encountered in bioprocessing. Typically, recombinant strains are more prone to shear damage than equivalent wild types. Extent of damage depends on the cell and the specific damaging mechanism operating. Often, the damage may be reduced by one or more of several approaches: adaptation of cells, supplementing culture media with shear-protectants, modification of processing operations and equipment, and other methods. In many cases, the mechanisms of protection and damage are not entirely clear, but multiple mechanisms seem to operate. As noted throughout this monograph, some sound guidance is available for engineering and operation of fragile culture processes; nevertheless, a priori quantitative prediction of culture behavior is not reliably possible. Empirical approaches must continue to be relied upon in engineering practice.

NOMENCLATURE

A_b	Cross-sectional area under the baffle or draft tube (m^2)
A_d	Cross-sectional area of the downcomer (m^2)
A_r	Cross-sectional area of the riser (m^2)
a	Parameter in equation 2
C_D	Drag coefficient
C_f	Fanning friction factor
d	Diameter or hydraulic diameter (m)
d_B	Bubble diameter (m)
d_c	Cell diameter (μm)
d_F	Floc diameter (m)
d_{Fmax}	Maximum stable diameter (m)
d_i	Impeller diameter (m)
d_L	Local diameter (m)
d_p	Microcarrier or particle diameter (m)
d_T	Tank or column diameter (m)
E	Energy dissipation rate per unit mass (W kg^{-1})
F_D	Turbulent drag force on a particle (N)
F_n	Normal force on a particle (N)
F_s	Shear force across the aggregate due to instantaneous velocity difference (N)
g	Gravitational acceleration (m s^{-2})
h	Channel height (m)
h_L	Height of liquid (m)
ICS	Impeller collision severity defined by equation 39 (kg m^{-2} s^{-3})
ISF	Integrated shear factor defined by equation 33 (s^{-1})
K	Consistency index (Pa s^n)
k	Parameter in equation 2 (m^{-1})
k'	Constant in equation 52 (s^{-1})
k_d	Death rate constant (s^{-1})
k_i	Constant in equation 9
k_j	Constant in equation 51 ($s^{4.5}$ $m^{-10.2}$)
L	Channel length (m)
L_h	Length of hairy extensions from pellet surface (m)
l	Mean microeddy length (m)
m	General exponent
N	Rotational speed (s^{-1}) or cell concentration (m^{-3})
N_o	Initial cell concentration (m^{-3})
N_p	Concentration of unbroken pellets at time t (m^{-3})
n	Flow behavior index
n_B	Number of impeller blades
Po	Power number
ΔP	Pressure drop (Pa)
Q	Volume flow rate of gas (m^3 s^{-1})
Re	Reynolds number
Re_i	Impeller Reynolds number defined by equation 36
Re_L	Local Reynolds number defined by equation 23
r_c	Radius of the vortex zone defined by equation 35 (m)
r_i	Impeller radius (m)
r_T	Tank radius (m)
S	Burst strength (N)
TCS	Turbulent collision severity for particle-to-particle collisions (kg m^{-2} s^{-3})
t	Time (s)
U_B	Bubble rise velocity (m s^{-1})
U_G	Superficial gas velocity (m s^{-1})
U_{Gr}	Superficial gas velocity in riser (m s^{-1})
U_L	Average liquid velocity (m s^{-1})
u	Mean velocity of the microeddies (m s^{-1})
Δu	Difference in velocities of the particle and the fluid (m s^{-1})
u_o	Jet velocity at orifice (m s^{-1})
V_k	Dimensionless hypothetical specific killing volume

V_L Volume of liquid (m^3)
W Width of the impeller blade (m)
x Parameter in equation 19

Greek Symbols

α Constant in equation 52 ($s^{35.575}\ m^{-65.56}$)
Γ Coefficient in equation 47
γ Shear rate (s^{-1})
γ_{av} Average shear rate (s^{-1})
$\gamma_{av\cdot T}$ Time averaged shear rate (s^{-1})
γ_i Isotropic turbulence shear rate defined by equation 24 (s^{-1})
γ_{max} Time averaged maximum shear rate (s^{-1})
γ_w Wall shear rate (s^{-1})
ϵ_S Volume fraction of microcarriers or solids (−)
μ_{ap} Effective or apparent viscosity (Pa s)
μ_L Viscosity of liquid (Pa s)
π Pi
ρ_L Density of liquid (kg m^{-3})
ρ_S Density of microcarriers or solid (kg m^{-3})
σ_M Mechanical strength of aggregate (N m^{-2})
τ Shear stress (N m^{-2})
τ_w Wall shear stress (N m^{-2})

BIBLIOGRAPHY

1. Y. Chisti and M. Moo-Young, *Enzyme Microb. Technol.* **8**, 194–204 (1986).
2. H.Y. Makagiansar, P.A. Shamlou, C.R. Thomas, and M.D. Lilly, *Bioprocess Eng.* **9**, 83–90 (1993).
3. E. Ujcová, Z. Fencl, M. Musi'lková, and L. Seichert, *Biotechnol. Bioeng.* **22**, 237–241 (1980).
4. M. Papadaki and S.G. Eskin, *Biotechnol. Prog.* **13**, 209–221 (1997).
5. C.R. Thomas, in M.R. Winkler ed., *Chemical Engineering Problems in Biotechnology*, Elsevier, London, 1990, pp. 23–93.
6. W.K. Kwok, C. Picioreanu, S.L. Ong, M.C.M. van Loosdrecht, W.J. Ng, and J.J. Heijnen, *Biotechnol. Bioeng.* **58**, 400–407 (1998).
7. P. Stoodley, Z. Lewandowski, J.D. Boyle, and H.M. Lappin-Scott, *Biotechnol. Bioeng.* **57**, 536–544 (1998).
8. Y. Chisti, in G. Subramanian ed., *Bioseparations and Bioprocessing: A Handbook*, vol. 2, Wiley-VCH, New York, 1998, pp. 3–30.
9. E. Dreveton, F. Monot, J. Lecourtier, D. Ballerini, and L. Choplin, *J. Ferment. Bioeng.* **82**, 272–276 (1996).
10. M. Moo-Young and Y. Chisti, *Biotechnology*, **6**(11), 1291–1296 (1988).
11. F. Garrido, U.C. Banerjee, Y. Chisti, and M. Moo-Young, *Bioseparation*, **4**, 319–328 (1994).
12. I.M. Tamer, M. Moo-Young, and Y. Chisti, *Ind. Eng. Chem. Res.* **37**, 1807–1814 (1998).
13. P.M. Doran, *Adv. Biochem. Eng. Biotechnol.* **48**, 115–168 (1993).
14. E. Molina Grima, Y. Chisti, and M. Moo-Young, *J. Biotechnol.* **54**, 195–210 (1997).
15. E.T. Papoutsakis, *Trends Biotechnol.* **9**, 427–437 (1991).
16. D.C. Augenstein, A.J. Sinskey, and D.I.C. Wang, *Biotechnol. Bioeng.* **13**, 409–418 (1971).
17. N. Shiragami, *Bioprocess Eng.* **10**, 47–51 (1994).
18. A. McQueen, E. Meilhoc, and J.E. Bailey, *Biotechnol. Lett.* **9**, 831–836 (1987).
19. Y. Chisti and M. Moo-Young, *Biotechnol. Bioeng.* **34**, 1391–1392 (1989).
20. Y. Chisti, *Appl. Mech. Rev.* **51**, 33–112 (1998).
21. Y. Chisti, *Airlift Bioreactors*, Elsevier, London, 1989.
22. L.K. Shi, J.P. Riba, and H. Angelino, *Chem. Eng. Commun.* **89**, 25–35 (1990).
23. L.A. Glasgow, J. Hua, T.-Y. Yiin, and L.E. Erickson, *Chem. Eng. Commun.* **113**, 155–181 (1992).
24. K. Okada, S. Shibano, and Y. Akagi, *J. Chem. Eng. Jpn.* **26**, 637–643 (1993).
25. A. Lübbert and B. Larson, *Chem. Eng. Sci.* **45**, 3047–3053 (1990).
26. A. Lübbert, B. Larson, L.W. Wan, and S. Bröring, *Inst. Chem. Eng. Symp. Ser.* **121**, 203–213 (1990).
27. R. Bowen, *Chem. Eng.* June 9, 55–63 (1986).
28. B. Robertson and J.J. Ulbrecht, in C.S. Ho and J.Y. Oldshue eds., *Biotechnology Processes: Scale-Up and Mixing*, American Institute of Chemical Engineers, New York, 1987, pp. 31–35.
29. P.H. Calderbank and M.B. Moo-Young, *Trans. Inst. Chem. Eng.* **37**, 26 (1959).
30. Y. Chisti and M. Moo-Young, in V. Moses and R.E. Cape eds., *Biotechnology: The Science and the Business*, Harwood Academic, New York, 1991, pp. 167–209.
31. D.C.-H. Jan, A.N. Emery, and M. Al-Rubeai, *Biotechnol. Techn.* **7**, 351–356 (1993).
32. Y. Chisti, B. Halard, and M. Moo-Young, *Chem. Eng. Sci.* **43**, 451–457 (1988).
33. Y. Chisti and M. Moo-Young, *J. Chem. Technol. Biotechnol.* **42**, 211–219 (1988).
34. V.N. Sokolov and V.P. Metkin, *Zh. Prikl. Khim.* **54**, 1321–1326 (1981): *J. Appl. Chem. USSR* **54**, 1103–1107 (1981).
35. V.P. Metkin and V.N. Sokolov, *J. Appl. Chem. USSR* **55**, 558–563 (1982).
36. K. Wichterle, V. Sobolík, M. Lutz, and V. Denk, *Chem. Eng. Sci.* **51**, 5227–5228 (1996).
37. A.W. Nienow, *Appl. Mech. Rev.* **51**, 3–32 (1998).
38. J.J. Meijer, H.J.G. ten Hoopen, Y.M. van Gameren, K.Ch.A.M. Luyben, and K.R. Libbenga, *Enzyme Microb. Technol.* **16**, 467–477 (1994).
39. M. Moo-Young, Y. Chisti, and D. Vlach, *Biotechnol. Lett.* **14**, 863–868 (1992).
40. M. Moo-Young, Y. Chisti, and D. Vlach, *Biotechnol. Adv.* **11**, 469–479 (1993).
41. O.T. Ramírez and R. Mutharasan, *Biotechnol. Bioeng.* **36**, 911–920 (1990).
42. E.H. Dunlop, P.K. Namdev, and M.Z. Rosenberg, *Chem. Eng. Sci.* **49**, 2263–2276 (1994).

43. Z. Zhang, Y. Chisti, and M. Moo-Young, *J. Biotechnol.* **43**, 33–40 (1995).
44. H. Märkl, R. Bronnenmeier, and B. Wittek, *Int. Chem. Eng.* **31**, 185–197 (1991).
45. J. Wu, G. King, A.J. Daugulis, P. Faulkner, D.H. Bone, and M.F.A. Goosen, *J. Ferment. Bioeng.* **70**, 90–93 (1990).
46. Y. Chisti and M. Moo-Young, *Trans. Inst. Chem. Eng.* **74**A, 575–583 (1996).
47. M. Moo-Young and Y. Chisti, in G. Durand, L. Bobichon, and J. Florent eds., *Proceedings, 8th International Biotechnology Symposium*, vol. 1, Société Française de Microbiologie, Paris, 1988, pp. 454–466.
48. M. Moo-Young, Y. Chisti, Z. Zhang, F. Garrido, U. Banerjee, and D. Vlach, *Ann. N. Y. Acad. Sci.* **782**, 391–401 (1996).
49. A. Converti, C. Sommariva, M.D. Borghi, and G. Ferraiolo, *Bioprocess Eng.* **9**, 183–189 (1993).
50. A. Converti, M.D. Borghi, G. Ferraiolo, and C. Sommariva, *Chem. Eng. J.* **62**, 155–167 (1996).
51. P.F. MacLoughlin, D.M. Malone, J.T. Murtagh, and P.M. Kieran, *Biotechnol. Bioeng.* **58**, 595–604 (1998).
52. C.R. Thomas, in P.A. Shamlou ed., *Processing of Solid–Liquid Suspensions*, Butterworth-Heinemann, Oxford, 1993, pp. 158–191.
53. A. Prokop and R.K. Bajpai, *Adv. Appl. Microbiol.* **37**, 165–232 (1992).
54. C. Gudin and D. Chaumont, *Bioresource Technol.* **38**, 145–151 (1991).
55. Y. Watanabe, J. de la Noüe, and D.O. Hall, *Biotechnol. Bioeng.* **47**, 261–269 (1995).
56. E. Molina Grima, J.A. Sánchez Pérez, F. García Camacho, J.L. García Sánchez, F.G. Acién Fernández, and D. López Alonso, *J. Biotechnol.* **37**, 159–166 (1994).
57. E. Molina Grima, J.A. Sánchez Pérez, F. García Camacho, J.M. Fernández Sevilla, F.G. Acién Fernández, and J. Urda Cardona, *Appl. Microbiol. Biotechnol.* **42**, 658–663 (1995).
58. H.J. Silva, T. Cortiñas, and R.J. Ertola, *J. Chem. Technol. Biotechnol.* **40**, 41–49 (1987).
59. T. Suzuki, T. Matsuo, K. Ohtaguchi, and K. Koide, *J. Chem. Technol. Biotechnol.* **62**, 351–358 (1995).
60. L. van der Pol, W.A.M. Bakker, and J. Tramper, *Biotechnol. Bioeng.* **40**, 179–182 (1992).
61. D.E. Martens, C.D. de Gooijer, E.C. Beuvery, and J. Tramper, *Biotechnol. Bioeng.* **39**, 891–897 (1992).
62. J. Tramper, D. Joustra, and J.M. Vlak, in C. Webb and F. Mavituna, eds., *Plant and Animal Cells: Process Possibilities*, Ellis Horwood, Chichester, 1987, pp. 125–136.
63. J. Tramper, D. Smit, J. Straatman, and J.M. Vlak, *Bioprocess Eng.* **2**, 37–41 (1987).
64. A.N. Emery, M. Lavery, B. Williams, and A. Handa, in C. Webb and F. Mavituna, eds., *Plant and Animal Cells: Process Possibilities*, Ellis Horwood, Chichester, 1987, pp. 137–146.
65. H.C.P. Matthijs, H. Balke, U.M. Van Hes, B.M.A. Kroon, L.R. Mur, and R.A. Binot, *Biotechnol. Bioeng.* **50**, 98–107 (1996).
66. M.S. Croughan, E.S. Sayre, and D.I.C. Wang, *Biotechnol. Bioeng.* **33**, 862–872 (1989).
67. A. McQueen and J.E. Bailey, *Biotechnol. Lett.* **11**, 531–536 (1989).
68. M. Midler and R.K. Finn, *Biotechnol. Bioeng.* **8**, 71–84 (1966).
69. P.M. Kieran, P.F. MacLoughlin, and D.M. Malone, *J. Biotechnol.* **59**, 39–52 (1997).
70. U. Fischer and A.W. Alfermann, *J. Biotechnol.* **41**, 19–28 (1995).
71. R. Ballica and D.D.Y. Ryu, *Biotechnol. Bioeng.* **42**, 1181–1189 (1993).
72. E. Reinhard, W. Kreis, U. Barthlen, and U. Helmbold, *Biotechnol. Bioeng.* **34**, 502–508 (1989).
73. H. Tanaka, *Biotechnol. Bioeng.* **23**, 1203–1218 (1981).
74. J.-J. Zhong, K. Fujiyama, T. Seki, and T. Yoshida, *Biotechnol. Bioeng.* **44**, 649–654 (1994).
75. P.M. Kieran, H.J. O'Donnell, D.M. Malone, and P.F. MacLoughlin, *Biotechnol. Bioeng.* **45**, 415–425 (1995).
76. J.J. Chalmers, *Appl. Mech. Rev.* **51**, 113–120 (1998).
77. W.R. Arathoon and J.R. Birch, *Science*, **232**, 1390–1395 (1986).
78. Y. Chisti, *Bioprocess Eng.* **9**, 191–196 (1993).
79. J.R. Birch, K. Lambert, P.W. Thompson, A.C. Kenney, and L.A. Wood, in B.K. Lydersen ed., *Large Scale Cell Culture Technology*, Hanser, New York, 1987, pp. 1–20.
80. Y. Chisti and M. Moo-Young, *Trans. Inst. Chem. Eng.* **71C**, 209–214 (1993).
81. Y. Chisti and M. Moo-Young, *Trans. Inst. Chem. Eng.* **72C**, 92–94 (1994).
82. B.-Y. Kim, A.J. Putnam, T.J. Kulik, and D.J. Mooney, *Biotechnol. Bioeng.* **57**, 46–54 (1998).
83. K.K. Chittur, L.V. McIntire, and R.R. Rich, *Biotechnol. Prog.* **4**(2), 89–96 (1988).
84. R.C. Ziegelstein, L. Cheng, and M.C. Capogrossi, *Science*, **258**, 656–659 (1992).
85. R.S. Cherry, *Biotechnol. Adv.* **11**, 279–299 (1993).
86. M. Al-Rubeai, A.N. Emery, S. Chalder, and M.H. Goldman, *J. Biotechnol.* **31**, 161–177 (1993).
87. P.L. Blackshear and G.L. Blackshear, in R. Skalak and S. Chien eds., *Handbook of Bioengineering*, McGraw-Hill, New York, 1987, pp. 15.1–15.19.
88. C. Born, Z. Zhang, M. Al-Rubeai, and C.R. Thomas, *Biotechnol. Bioeng.* **40**, 1004–1010 (1992).
89. N. Shiragami, *Bioprocess Eng.* **16**, 345–347 (1997).
90. C.G. Smith and P.F. Greenfield, *Biotechnol. Bioeng.* **40**, 1045–1055 (1992).
91. L.A. Olivier and G.A. Truskey, *Biotechnol. Bioeng.* **42**, 963–973 (1993).
92. J.G. Aunins and H.-J. Henzler, in H.-J. Rehm and G. Reed eds., *Biotechnology*, vol. 3, 2nd ed., VCH Weinheim, 1993, pp. 219–281.
93. N. Shiragami and H. Unno, *Bioprocess Eng.* **10**, 43–45 (1994).
94. T.L. LaPorte, J. Shevitz, Y. Kim, and S.S. Wang, *Bioprocess Eng.* **15**, 1–7 (1996).
95. R.M. Hochmuth, in R. Skalak and S. Chien eds., *Handbook of Bioengineering*, McGraw-Hill, New York, 1987, pp. 12.1–12.17.
96. G.W. Schmid-Schönbein, in R. Skalak and S. Chien eds., *Handbook of Bioengineering*, McGraw-Hill, New York, 1987, pp. 13.1–13.25.

97. X. Dai, F. Ouyang, and Y. Qi, in W.K. Teo, M.G.S. Yap, and S.K.W. Oh, eds., *Better Living through Biochemical Engineering*, University of Singapore, Singapore, 1994, pp. 408–410.
98. R.S. Cherry and E.T. Papoutsakis, *Bioprocess Eng.* **1**, 29–41 (1986).
99. Y. Chisti and M. Moo-Young, *Chimica Oggi*, **11**(3–4), 25–27 (1993).
100. Z. Zhang, M.A. Ferenczi, and C.R. Thomas, *Chem. Eng. Sci.* **47**, 1347–1354 (1992).
101. Z. Zhang, M.F. Ferenczi, A.C. Lush, and C.R. Thomas, *Appl. Microbiol. Biotechnol.* **36**, 208–210 (1991).
102. A.J. Sinskey, R.J. Fleishaker, M.A. Tyo, D.-J. Giard, and D.I.C. Wang, *Ann. N. Y. Acad. Sci.* **369**, 47–59 (1981).
103. M.S. Croughan, J.-F. Hamel, and D.I.C. Wang, *Biotechnol. Bioeng.* **29**, 130–141 (1987).
104. R.S. Cherry and E.T. Papoutsakis, *Biotechnol. Bioeng.* **32**, 1001–1014 (1988).
105. R.S. Cherry and E.T. Papoutsakis, *Bioprocess Eng.* **4**, 81–89 (1989).
106. K.J. Ganzeveld, Y. Chisti, and M. Moo-Young, *Bioprocess Eng.* **12**, 239–247 (1995).
107. S.K.W. Oh, A.W. Nienow, M. Al-Rubeai, and A.N. Emery, *J. Biotechnol.* **12**, 45–62 (1989).
108. S.K.W. Oh, A.W. Nienow, M. Al-Rubeai, and A.N. Emery, *J. Biotechnol.* **22**, 245–270 (1992).
109. A. Handa, A.N. Emery, and R.E. Spier, *Dev. Biol. Stand.* **66**, 241–253 (1987).
110. A. Handa-Corrigan, A.N. Emery, and R.E. Spier, *Enzyme Microb. Technol.* **11**, 230–235 (1989).
111. N.S. Wang, J.-D. Yang, R.V. Calabrese, and K.-C. Chang, *J. Biotechnol.* **33**, 107–122 (1994).
112. I. Jöbses, D. Martens, and J. Tramper, *Biotechnol. Bioeng.* **37**, 484–490 (1991).
113. J. Wu and F.A. Goosen, *Enzyme Microb. Technol.* **17**, 1036–1042 (1995).
114. J.M. Boulton-Stone and J.R. Blake, *J. Fluid Mech.* **254**, 473–466 (1993).
115. M.A. Garcia-Briones, R.S. Brodkey, and J.J. Chalmers, *Chem. Eng. Sci.* **49**, 2301–2320 (1994).
116. J.J. Chalmers and F. Bavarian, *Biotechnol. Prog.* **7**, 151–158 (1991).
117. R.S. Cherry and C.T. Hulle, *Biotechnol. Prog.* **8**, 11–18 (1992).
118. F. Bavarian, L. S. Fan, and J.J. Chalmers, *Biotechnol. Prog.* **7**, 140–150 (1991).
119. H.G. Sucker, M. Jordan, H.M. Eppenberger, and F. Widmer, *Biotechnol. Bioeng.* **44**, 1246–1254 (1994).
120. E.T. Papoutsakis, *Trends Biotechnol.* **9**, 316–324 (1991).
121. O.T. Ramirez and R. Mutharasan, *Biotechnol. Prog.* **8**, 40–50 (1992).
122. J.D. Michaels, J.F. Petersen, L.V. McIntire, and E.T. Papoutsakis, *Biotechnol. Bioeng.* **38**, 169–180 (1991).
123. M. Lavery and A.W. Nienow, *Biotechnol. Bioeng.* **30**, 368–373 (1987).
124. K.T. Kunas and E.T. Papoutsakis, *J. Biotechnol.* **15**, 57–70 (1990).
125. S. Goldblum, Y.-K. Bae, W.F. Hink, and J.J. Chalmers, *Biotechnol. Prog.* **6**, 383–390 (1990).
126. Y. Chisti and M. Moo-Young, *J. Chem. Technol. Biotechnol.* **58**, 331–336 (1993).
127. D.W. Murhammer and C.F. Goochee, *Biotechnology*, **6**, 1411–1418 (1988).
128. D.W. Murhammer and C.F. Goochee, *Biotechnol. Prog.* **6**, 142–148 (1990).
129. L. Laouar, K.C. Lowe, and B.J. Mulligan, *Enzyme Microb. Technol.* **18**, 433–438 (1996).
130. L.A. van der Pol, I. Beeksma, and J. Tramper, *Enzyme Microb. Technol.* **17**, 401–407 (1995).
131. M.C.M. van Loosdrecht and S.J. Heijnen, *Trends Biotechnol.* **11**, 117–121 (1993).
132. J.J. Heijnen, M.C.M. van Loosdrecht, A. Mulder, and L. Tijhuis, *Water Sci. Technol.* **26**, 647–654 (1992).
133. K. Wagner and D.C. Hempel, *Biotechnol. Bioeng.* **31**, 559–566 (1988).
134. L. Tijhuis, H. Zwols, K.-H. van Meekeren, M.C.M. van Loosdrecht, and J.J. Heijnen, L. Alberghina, L. Frontali, and P. Sensi eds., *Progress in Biotechnology*, vol. 9, *Proceedings of the 6th European Congress on Biotechnology*, Elsevier, Amsterdam, 1994, pp. 953–956.
135. L. Tijhuis, W.A.J. van Benthum, M.C.M. van Loosdrecht, and J.J. Heijnen, *Biotechnol. Bioeng.* **44**, 867–879 (1994).
136. L. Tijhuis, M.C.M. van Loosdrecht, and J.J. Heijnen, *Biotechnol. Bioeng.* **44**, 595–608 (1994).
137. L. Tijhuis, M.C.M. van Loosdrecht, and J.J. Heijnen, *Biotechnol. Bioeng.* **45**, 481–487 (1995).
138. W.A.J. van Benthum, J.M. Garrido-Fernández, L. Tijhuis, M.C.M. van Loosdrecht, and J.J. Heijnen, *Biotechnol. Prog.* **12**, 764–772 (1996).
139. A. Gjaltema, L. Tijhuis, M.C.M. van Loosdrecht, and J.J. Heijnen, *Biotechnol. Bioeng.* **46**, 258–269 (1995).
140. A. Gjaltema, J.L. Vinke, M.C.M. van Loosdrecht, and J.J. Heijnen, *Biotechnol. Bioeng.* **53**, 88–99 (1997).
141. L. Tijhuis, J.L. Huisman, H.D. Hekkelman, M.C.M. van Loosdrecht, and J.J. Heijnen, *Biotechnol. Bioeng.* **47**, 585–595 (1995).
142. P.A. Shamlou and N. Titchener-Hooker, in P.A. Shamlou ed., *Processing of Solid–Liquid Suspensions*, Butterworth-Heinemann, Oxford, 1993, pp. 1–25.
143. D.F. Bagster, in P.A. Shamlou ed., *Processing of Solid–Liquid Suspensions*, Butterworth-Heinemann, Oxford, 1993, pp. 26–58.
144. T. Matsuo and H. Unno, *J. Environ. Eng. Div. Am. Soc. Civ. Eng.* **107**, 527–545 (1981).
145. J. Gregory, in P.A. Shamlou ed., *Processing of Solid–Liquid Suspensions*, Butterworth-Heinemann, Oxford, 1993, pp. 59–92.
146. H. Taguchi, T. Yoshida, Y. Tomita, and S. Teramoto, *J. Ferment. Technol.* **46**, 814–822 (1968).
147. Y.Q. Cui, R.G.J.M. van der Lans, and K.Ch.A.M. Luyben, *Biotechnol. Bioeng.* **57**, 409–419 (1998).
148. Y.-F. Maa and C.C. Hsu, *Biotechnol. Bioeng.* **54**, 503–512 (1997).
149. Y.-F. Maa and C.C. Hsu, *Biotechnol. Bioeng.* **51**, 458–465 (1996).
150. S.A. Rouf, M. Moo-Young, and Y. Chisti, *Biotechnol. Adv.* **14**, 239–266 (1996).
151. D.J. Bell, M. Hoare, and P. Dunnill, *Adv. Biochem. Eng.* **26**, 1–72 (1983).

SOLID STATE FERMENTATION, MICROBIAL GROWTH KINETICS

DAVID A. MITCHELL
Universidade Federal do Parana
Curitiba, Brazil

DEIDRE M. STUART
Queensland University of Technology
Brisbane, Australia

ROBERT D. TANNER
Vanderbilt University
Nashville, Tennessee

KEY WORDS

Growth kinetics
Mathematical modeling
Solid-state fermentation
Transport phenomena

OUTLINE

INTRODUCTION

Solid-state fermentation (SSF) involves the growth of microorganisms on moist solid substrates in the absence of free water. The absence of free water makes the system quite different from submerged liquid fermentations and makes SSF superior for the production of some products. However, compared with submerged liquid fermentation, relatively little is known about how to design and operate bioreactors for large-scale SSF processes.

Despite this, there are commercially successful large-scale SSF processes for the production of soy sauce koji and other traditional fermented foods, citric and gluconic acids, and fungal enzymes such as cellulases, amylases, lipases, and pectinases (1,2). In addition, fungal spores have been produced by SSF for use in steroid transformations and as inocula for production of blue-vein cheeses (2). Composting is another SSF application that is widely practiced, although composting is different from other SSF processes, because self-heating of the fermenting substrate is desired in composting processes, whereas avoidance of overheating is a key problem in most SSF processes.

In addition to these commercial processes, over the last 20 years there has been a growing interest in other SSF applications (2), including the following:

- Upgrading of solid wastes from agriculture and food processing for use as fermented animal feeds. In some cases the microorganism is used to enrich the protein content; in other cases it is used to degrade toxic substances in the waste.
- Production of fungal spores for use as mycopesticides. SSF is superior for the production of fungal spores because most fungi do not sporulate well in liquid culture, and even when they do, spores produced in SSF tend to be more robust.
- Production of other enzymes such as dehairing enzymes, fungal rennets, beta-glucanases and xylanases.
- Biopulping of woodchips during the production of paper.
- Production of various other products such as gibberellic acid, aroma and flavor compounds, antibiotics, and ethanol.

Unlike the situation with submerged liquid fermentations, there is no systematic framework guiding the design and operation of large-scale bioreactors for SSF. As a result, there has been relatively little success in developing successful large-scale processes for these newer SSF products, although large-scale bioreactors have been used to

make soy sauce koji. Intermediate-scale processes can be carried out using tray fermentations, but such operations are quite labor intensive. Mathematical models of bioreactor operation will be key tools in the development of strategies for the design and operation of large-scale SSF bioreactors. This article summarizes the current state of development of modeling of SSF processes.

AIMS OF THIS ARTICLE

This article addresses the importance, to the development of large-scale SSF technology, of understanding and modeling microbial growth kinetics and transport phenomena, which combine to determine the overall performance of a solid-state bioreactor. Transport phenomena are of special importance in SSF due to the unavoidable heterogeneity of SSF systems. Good reactor performance will be achieved only through an understanding of these transport phenomena and how they interact with the microbial growth kinetics. Mathematical models represent a convenient, concise, and powerful way of describing these phenomena and their interactions and provide a sound foundation for process development, control, and optimization. They can also guide us in learning where the problems are and how to attack them.

The main message of this article is that modeling of phenomena in SSF and of the performance of SSF bioreactors has not yet progressed to a stage where it can be usefully applied to industrial applications. In fact, it was only in 1990 that the first model was proposed for SSF bioreactor operation (3). Although advances have been made since then, currently we still do not have mechanistic models of SSF that are able to make accurate predictions based fully on independently determined parameter values. To date the focus has been on appropriate model structures. Relatively little attention has been given to parameter estimation, with many parameters being borrowed from other systems, systems that often involve different substrates, organisms, and cultivation systems. As a result, none of the models has been properly validated against experimental data, and although the models are now being used to predict improved operating conditions, they have not yet been used to improve operation of actual bioreactors. In most cases some of the model parameters have been estimated by adjusting them to enable the model to fit the experimental data. Furthermore, although the ultimate aim of SSF processes is to produce useful products industrially, to date models have focused on growth. In fact, in some cases the main aim has been to predict temperatures rather than growth itself.

The aim of this article is to draw together the basic principles from the work that has been done so far, to show the general concepts that underpin all models, to identify the features that will make models most useful for application in industrial processes, and thereby to provide a foundation for continued progress in the modeling of SSF.

ORGANIZATION AND SCOPE

Filamentous fungi are the most important group of microorganisms used in SSF, and most processes are aerobic. Most of the work discussed in this article addresses this type of system, although other systems are mentioned as appropriate.

The article begins by describing the biological and transport phenomena occurring in an SSF bioreactor and identifying the various phenomena that can potentially limit the kinetic performance of the bioreactor:

- Microbial growth phenomena
- Local transport phenomena (occurring at the level of a single particle)
- Bulk transport phenomena (occurring at the level of the whole bioreactor)

Current models of SSF address the influence of either local or bulk transport phenomena on growth and are therefore labeled as microscopic or macroscopic models. The general aims and usefulness of these two types of models are discussed. The mathematical expressions appropriate for describing the three types of phenomena that may be incorporated into these models are discussed in turn.

Readers interested in modeling of SSF processes are urged to consult the original references because no attempt has been made to present the complete models of individual workers. The original articles do that better than we could hope to do. Rather we have organized this article around the ways in which the various kinetic and transport phenomena have been handled. Therefore, where we present equations, these have been separated from the author's discussion of their underlying assumptions and the notation and units they have used.

DESCRIPTION OF KEY FEATURES OF SOLID-STATE FERMENTATION

Growth in an SSF bioreactor results from and is influenced by processes operating at different scales (Fig. 1): microbial processes and local and bulk transport processes. Figure 1 also highlights some of the complex interactions between these phenomena during SSF. This article addresses the approaches that have been made to modeling the various phenomena shown in Figure 1, and the interactions between the various phenomena. Thermodynamic phenomena, such as the ability of solid substrates to absorb heat and the ability of gases to carry water vapor, also affect the overall bioreactor performance; however, these phenomena are not addressed in this article.

In developing a mathematical model, it is important to identify the processes occurring before attempting to describe them mathematically. This focuses attention on the processes that are potentially rate limiting. An understanding of the rate-limiting processes facilitates development of strategies to overcome the limitation, although SSF processes can be difficult to optimize because the limiting step might be different under different conditions, and at different stages of an SSF process the growth rate may be limited by different phenomena. Therefore this section provides a qualitative description of our current understanding of the phenomena occurring in SSF.

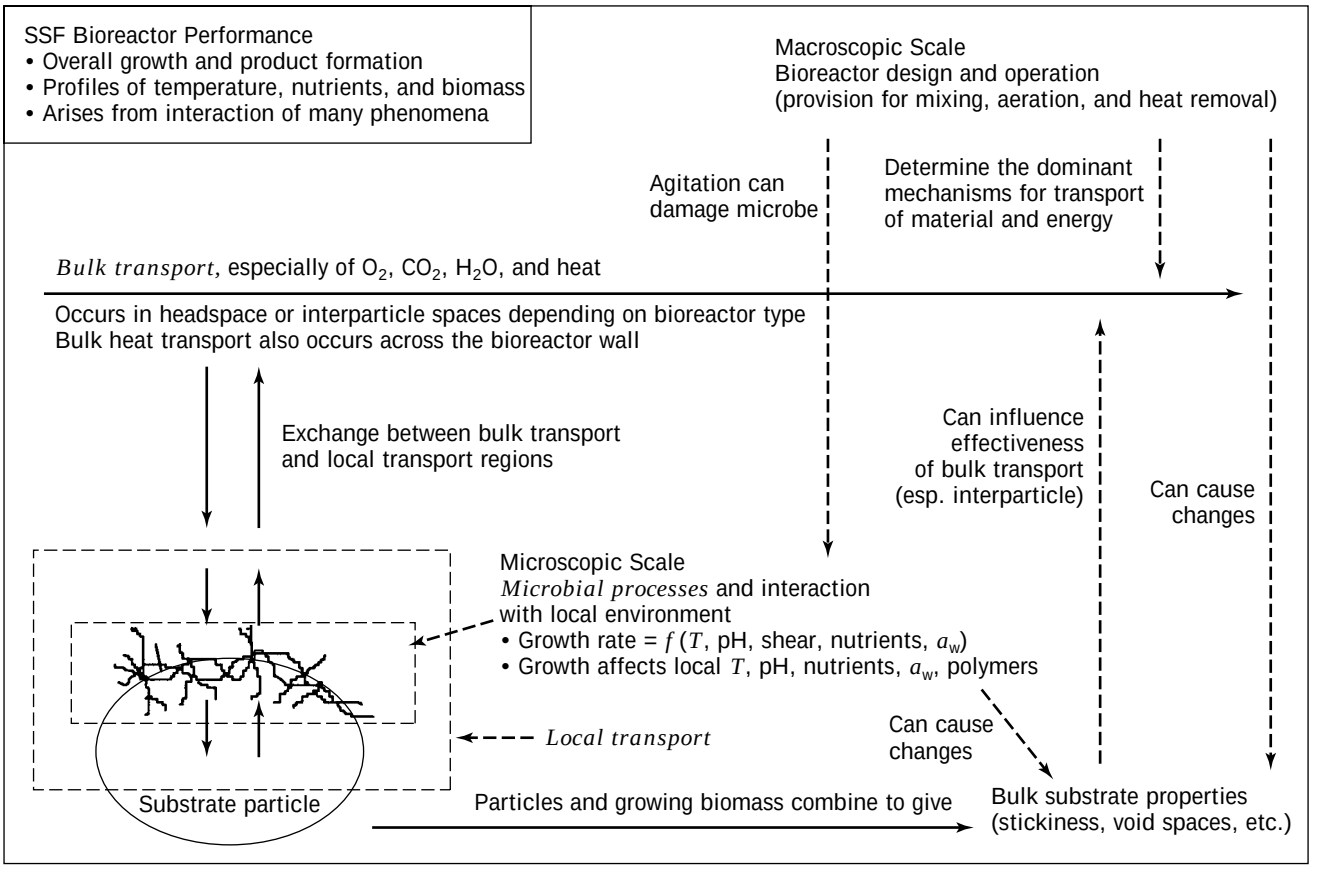


Figure 1. An overview of the microscopic and macroscopic processes that combine to determine the overall performance of an SSF bioreactor. This diagram describes transport processes in general terms and does not identify all the individual mass and heat transfer processes occurring. Although the diagram represents an SSF bioreactor, it is not drawn to scale. The nature and extent of the local and bulk transport regions depend on the particular bioreactor and how it is operated.

The Physical Nature Of Solid-State Fermentation Systems

From a macroscopic viewpoint, up to four subsystems are apparent (Fig. 2), depending on the bioreactor type: the material from which the bioreactor is constructed, a headspace containing gases, the solid particles themselves, and the interparticle gas phase entrapped between the solid particles. It is often appropriate to think of the solid particles and interparticle gas as a pseudohomogeneous "substrate bed."

From a microscopic perspective the system is considered in greater detail. The description focuses on the region near an individual particle and can include (Fig. 3) the following:

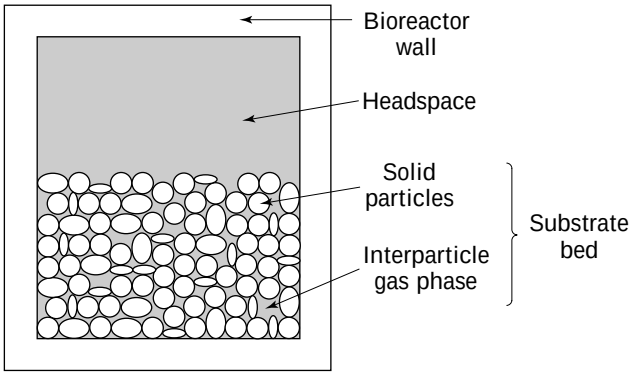


Figure 2. Subsystems within a solid-state fermentation bioreactor.

- The biomass, the majority of which is concentrated as a layer at the particle surface. Part of the fungal mycelium is immersed in the liquid film at the surface, but many hyphae extend into the gaseous region. These aerial hyphae may support films or droplets of water. Other hyphae may penetrate into the substrate particle. In contrast, yeasts and bacteria form a compact moist layer on the particle surface.
- The solid matrix, which provides the physical support for the liquid and biomass phases, and which may or may not be degraded during the fermentation.
- A liquid phase consisting of a stagnant liquid film at the substrate surface and the moisture inside the particle. This phase may be continuous, depending on the internal structure of the substrate particle.

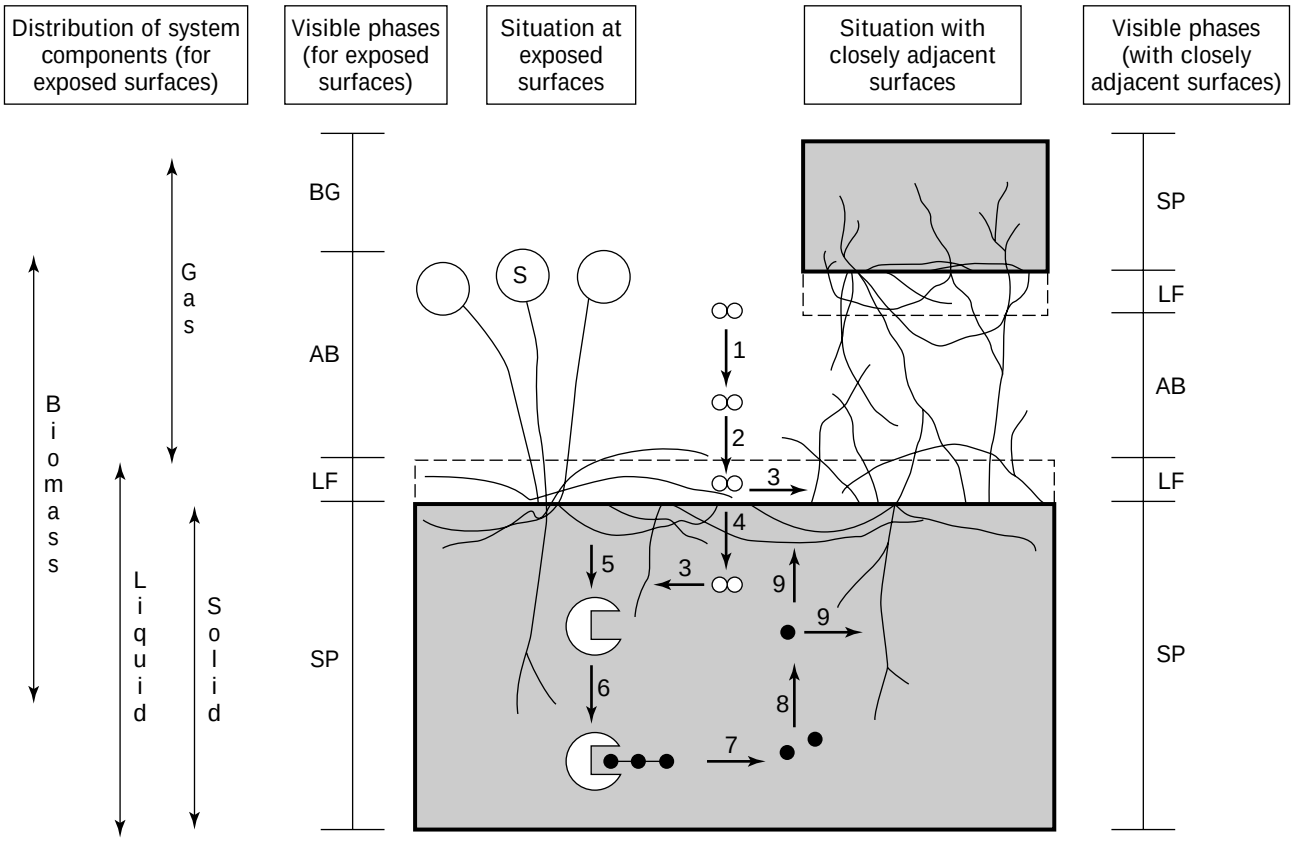


Figure 3. Structure of SSF systems at the microscopic scale and local mass transport processes. The example is for growth of a filamentous fungus, such as *Rhizopus* sp., on a starchy substrate. The diagram represents a cross-sectional view near the surfaces of the substrate particles and indicates the difference between substrate particles exposed to the bulk gas phase and substrate particles that are packed closely together. Simplifications, such as the absence of penetrative hyphae, and the liquid film as the only site of O_2 uptake, have been used in modeling studies. ∞, oxygen; ●, glucose; ●—●—●, starch; S, sporangium; SP, substrate particle; LF, liquid film; AB, aerial biomass; BG, bulk gas. The processes occurring are 1, bulk transport of oxygen; 2, diffusion of O_2 through stagnant gas regions and diffusion across the gas–liquid interface; 3, uptake of oxygen by hyphae; 4, diffusion of oxygen through the liquid phase of the substrate; 5, release of glucoamylase from the hyphae; 6, diffusion of glucoamylase through the solid substrate particle; 7, hydrolysis of starch by the glucoamylase; 8, diffusion of the released glucose through the solid substrate particle; 9, uptake of glucose by the hyphae.

- A gaseous phase. Although this gas phase is continuous, it may be conceived as comprising up to three regions. First, the aerial hyphae of the biomass are surrounded by regions of stagnant gas, and there will be a stagnant gas film adjacent to the liquid film. Second, there may be a stagnant gas region above the biomass region (i.e., part of the interparticle space). Third, there may be a mobile region in which there is bulk flow. The locations and extents of the stagnant and mobile gas regions will depend on whether the substrate bed is agitated and how it is aerated. For example, in static packed beds the interparticle spaces are filled with a network of hyphae, but there is still bulk gas flow through parts of this region. On the other hand, in tray bioreactors, the gas phase within the substrate bed is relatively stagnant, although there may be some free convection due to thermal gradients.

The Processes Occurring During Solid-State Fermentation

Aufeuvre and Raimbault (4) described the changes in morphology during growth of the filamentous fungus *Aspergillus niger* in a packed bed bioreactor, on a granular substrate prepared from cassava flour. Initially, the spores used as inoculum are spread across the substrate surface. During the germination period of about 8 h, the spores swell and then extend germ tubes across the substrate surface. These germ tubes extend for up to 50 μm before branching begins. Branching of these germ tubes then rapidly leads to a loose covering of hyphae across the whole surface. As branching continues, the hyphal density increases, and it is soon impossible to distinguish individual mycelia. The network of hyphae increases in thickness, so that many hyphae are not in contact with the substrate surface. Some hyphae form together into multihyphal filaments, which can extend up to 80 μm above the substrate

surface. These filaments form bridges between the substrate granules and bind the whole bed of substrate particles together into a single compact mass.

This description addresses only the features of the growth process that can be seen with a microscope. Many less obvious processes are occurring and influencing growth (Fig. 3). The growth activities of the microbe are determined by the local environment experienced by the microbe, especially the local temperature, pH, water activity (a_w), oxygen concentration, and nutrient concentrations. In turn, growth processes also affect the local environment, for example, hydrolytic enzymes are secreted, carbon and nitrogen substrates are absorbed into the mycelium, oxygen is consumed, and various products are excreted. In addition, waste metabolic heat is released by the microorganism. Therefore the temperature, pH, water activity, and oxygen and nutrient concentrations in the local environment are affected by the microorganism. This induces the transport processes, which tend to equilibrate chemical potential and temperature throughout a material. Within solid particles and within stagnant gas regions, mass transfer is limited to diffusion and the exchange of gases (O_2, CO_2, and water vapor) at gas–liquid interfaces, and heat transfer is limited to conduction, free-convection, and evaporation of water at the gas–liquid interface. These dissipative processes can be relatively slow compared with the rate of the microbial processes, leading to the establishment of local concentration and temperature gradients around the microbe.

The purpose of a bioreactor is to maintain an optimal local environment for the desired microbial activity, within the limits of physical and economical constraints. In SSF the ability to control the local environment is largely limited to manipulating either the bulk gas flow or agitation of the substrate, or both, providing oxygen for aerobic processes, removing carbon dioxide and evaporated water, and removing heat. Almost nothing can be done during the fermentation to prevent nutrient concentration gradients within individual substrate particles, because it is not possible to have bulk mixing below the scale of the substrate particles. This is an intrinsic consequence of the physical form of SSF systems.

Given that control of the environment is achieved by manipulating the bulk phases, the rate of removal of heat and the rate of supply of oxygen can be maximized by minimizing gradients within the bulk gas phase and minimizing the thickness of the stagnant gas film that separates the bulk gas phase from the local environment. Different reactor types enable this to be achieved to different extents. In bioreactors with bulk flow of air past the solids, due either to forced airflow or to mixing of the substrate bed, gases diffuse from the stagnant gas regions into the bulk flow and are carried away. Heat can be removed by convection and by evaporation of water into this air. Evaporative heat removal can be maximized by aerating with dry air, but this will tend to dry the substrate. Depending on the bioreactor operation there might be significant gas concentration gradients and temperature gradients in the bulk air phase across the bioreactor. Plug flow of gas (and maybe substrate) will tend to favor macroscopic gradients, whereas mixing of the gas phase and of the substrate bed will tend to minimize such gradients. In bioreactors without bulk flow of air past the solids (i.e., trays), mass transfer and heat transfer occur mainly by diffusion and conduction, respectively. In addition to heat and mass transfer within the bioreactor, there are also heat transfers between the substrate bed and the bioreactor wall, the headspace gases and the bioreactor wall, and the bioreactor walls and the surroundings (Fig. 1).

During microbial growth in SSF, growth activities and the aeration and agitation strategies can affect the bulk substrate properties, which in turn can influence the bulk transport processes, as indicated in Figure 1. For example, in an unstirred packed bed, fungal mycelia grow in the interparticle spaces and can significantly increase the pressure drop through the bed, affecting airflow patterns within the bed (5). Microbial activities on the substrate may decrease the strength or increase the stickiness of the particles, leading to compaction or cohesion of particles when the substrate is agitated, thereby eliminating interparticle spaces and preventing bulk transport to or from the microbes within the bed.

In conclusion, there can be significant profiles of temperature, nutrients, and biomass at both the microscopic scale (looking at a single particle with a microbe growing on it, as shown in Fig. 3) and the macroscopic scale (looking across the bioreactor as a whole, as shown in Fig. 1). The success of transport processes at both these scales in dissipating thermal and concentration gradients determines how readily the local conditions can be maintained near the optimal state for the desired microbial activity. Therefore, the overall performance of an SSF bioreactor arises from the interaction of biological growth kinetics, local transport processes, and bulk transport processes. The next section discusses the nature of models of SSF in general terms, whereas the remaining sections address how the underlying phenomena have been described in mathematical models of SSF processes.

Approaches to Modeling Solid-State Fermentation

A fully mechanistic description of the bulk and local transport processes occurring within SSF (Figs. 1 and 3), when combined with the growth kinetics, would lead to a very complex model of the system. Models of SSF to date have focused on either local or bulk transport phenomena, and therefore it is convenient to classify the models as being either microscopic models or macroscopic models. Microscopic models concern themselves with growth at the level of individual particles, whereas macroscopic models describe overall bioreactor performance. Both types of models have a role to play in the development of SSF technology. However, in the future, models should be developed that integrate mechanistic descriptions of both local and bulk transport phenomena, although this will depend on further advances in computing power and methods for constructing and solving complex dynamic models. Such models will describe performance under a wide range of operating conditions. This trade-off between the power of the model and the ease of model solution should be kept in mind. As well, the purpose of the model and the ability to estimate parameters should be considered when develop-

ing models that incorporate both microscopic and macroscopic phenomena. Because microscopic and macroscopic models are currently developed separately, the modeling work usually has related but different aims.

Microscopic models combine local transport phenomena (Fig. 3) and the effect of the local environment on growth, in an attempt to identify the mechanisms by which these may control the growth rate and the total amount of growth that occurs during SSF. For example, due to the absence of mixing at the microscopic scale, diffusive phenomena within the substrate particle can limit the growth rate (6). Microscopic models tend to consider the biomass in more detail than do macroscopic models, and it may be appropriate to describe variations in hyphal density in space, or the presence of differentiated forms of biomass. However, for the modeling of bioreactor performance, it is not appropriate to describe the extension and branching of individual fungal hyphae. Models incorporating these growth phenomena are powerful tools for fundamental studies aimed at understanding the mechanisms of fungal growth but are not discussed in this article.

Macroscopic models attempt to describe the performance of the whole bioreactor (Fig. 1). Such models may be developed to contribute to various aspects of bioreactor development such as design, operation, scale-up, control, and optimization. The main aim is to describe growth, although this may require prediction of temperature or nutrient or oxygen concentrations, or properties of the substrate bed. Their level of sophistication is in keeping with the ability or lack of ability to measure fermentation parameters such as biomass, nutrient concentrations, gas concentrations, temperature, and pH. Macroscopic models should describe the phenomena controlling the bioreactor performance, but current macroscopic models simply combine bulk transport phenomena with empirical descriptions of microbial growth. They do not address local transport phenomena, because of the complexity that these introduce. Simple empirical descriptions of growth may subsume biological and local transport limitations on growth but conceal the mechanisms of the limitation. Therefore, current macroscopic models are most useful for determining whether bioreactor performance is limited by gas or heat exchange, and guiding bioreactor design and operation to overcome any such limitations.

Several other characteristics of SSF lead to common features of models:

- The vast majority of SSF processes are batch processes, and the few continuous processes involve plug flow systems rather than well-stirred vessels. Therefore dynamic models are most useful for describing SSF. Steady-state models are usually not appropriate.
- The presence of mass and thermal gradients means that in most cases it is desirable to include spatial coordinates in addition to the temporal coordinate. Unfortunately this makes the equation sets more complicated to set up and solve.
- Balances are commonly set up over more than one of the subsystems shown in Figure 2.
- Depending on what is expected to limit growth, either heat or mass transfer may be described. Models incorporating both of these transport phenomena will be most flexible in describing what controls the rate of growth under a wide range of operating conditions.
- Due to the complexity that transport phenomena introduce into problems, plus the difficulties of measuring biomass and metabolic states, the majority of models for SSF are unstructured and nonsegregated, although some do have very simple structure or segregation.

The remaining sections discuss the various expressions that have been used within models of SSF to describe microbial, local transport, and bulk transport phenomena in SSF.

MODELING OF MICROBIAL PHENOMENA

This section addresses microbial phenomena and how models of SSF have treated these phenomena. It focuses on studies of either SSF itself or similar systems involving overculture on solid nutrient media. Overculture occurs when a fungus is inoculated across the whole surface, and therefore the mycelium develops simultaneously from many points. This contrasts with colony growth, where the organism is inoculated at one point, and a colony spreads outward across the surface. The kinetics of colony growth on agar plates are not considered in this article. Note also that the word biomass is used here in the restrictive sense of microbial biomass, and not in the broad sense of agricultural biomass. This distinction is important because many substrates are of agricultural origin.

A word of warning is appropriate about modeling growth kinetics. The basis for expression of the microbial biomass must be clear (e.g., total biomass in the bioreactor, or biomass per kilogram of initial dry substrate, or biomass per kilogram of wet substrate present, etc.), and conservation equations must originally be formulated on a conserved quantity (i.e., absolute amounts of biomass rather than biomass concentrations). This is crucial in SSF because there is a significant decrease in substrate mass due to conversion of dry matter to CO_2, and therefore increases in biomass concentration result from two effects: the increase in the overall amount of biomass, plus the decrease in the overall amount of dry matter. Water transfer between the substrate and gas phases causes further complications if concentrations are expressed on a wet substrate basis. Due to the wide range of ways that different workers have expressed biomass, we have used general equations in which the variables X and S represent absolute amounts of biomass and substrate respectively, whereas the variables C_X and C_S represent concentrations. Rates are expressed as mass or energy per unit time and are represented by the symbol r with a subscript. Readers should refer to original articles to find the units and nomenclature used by the authors.

Substrate-Independent Growth Kinetic Equations

Sangsurasak et al. (7) note that four general shapes of growth kinetic profiles in SSF have been reported. In SSF

models the simplest kinetic expressions describing these profiles relate the growth rate (r_g) directly to the biomass:

- Linear growth kinetics, described by

$$r_g = \text{constant} \tag{1}$$

  where r_g is the rate of growth.
- Exponential growth kinetics, described by

$$r_g = \mu_m X \tag{2}$$

  where μ_m is the observed maximal specific growth rate.
- Early acceleration followed by deceleration, with the rates of acceleration and deceleration being nearly equal, such that the biomass profile is symmetrical by rotation around the transition point. These kinetics are described by the logistic equation

$$r_g = \mu_m X\left(1 - \frac{X}{X_m}\right) \tag{3}$$

  where X_m is the maximum possible amount of biomass.
- Early rapid acceleration, followed by an extended period during which growth decelerates slowly. This profile can be approximated by raising the ratio X/X_m to an exponent n to give the power-law logistic model (8,9)

$$r_g = \mu_m X\left(1 - \left(\frac{X}{X_m}\right)^n\right) \tag{4}$$

This allows the logistic curve to be skewed (i.e., for the rates of acceleration and deceleration to be different). The value of n is a measure of the relative sensitivity of the culture to self-inhibition at high biomass densities (8). For values of n less than 1, the organism is relatively sensitive to self-inhibition, and self-inhibition occurs for quite low values of X. For $n = 1$ the logistic equation is obtained. Finally, for values of n greater than 1 the organism is relatively resistant to self-inhibition, and significant self-inhibition occurs only at values of X close to X_m. The exponent n may have biological significance. The mycelium arising from a spore undergoes a maximum number (k_m) of branching events before growth ceases, and mycelia arising from a large number of different spores have a distribution of k_m values. If n is less than 1 then many of the mycelia have low k_m values, whereas if n is greater than 1 then many of the mycelia have high k_m values (8).

The expressions for both the linear and exponential growth kinetics do not describe any limitation on growth and are of little use for incorporating into models of growth in SSF. On the other hand, the logistic equation (and by extension, its power-law modification) describes a limitation on growth, and for this reason has quite commonly been incorporated into models of growth in SSF (10–12), especially macroscopic models. However, the significance of X_m is not clear. Growth might potentially be limited by the amount of substrate or diffusional limitations, but it can also be limited sterically.

Assumption of symmetric branching enables a simplified description of the mycelial mode of growth and enables the macroscopic parameter μ_m in the logistic equation to be related to the specific growth rate of a segment of biomass just before it is about to divide (μ'_m). The value of μ'_m can be determined from microscopic observation of extending hyphae. After a few branching events the following equation holds (8,9):

$$\mu_m = \mu'_m \frac{(X + X_c)}{2X} \approx \frac{\mu'_m}{2} \tag{5}$$

where X_c is the biomass in one hyphal segment just prior to a branching event. The approximation arises because X_c is small compared with the total biomass. This relationship between the two specific growth rates occurs because, after a few symmetric branchings, the biomass in the actively extending segments of a mycelium comprises approximately half the total biomass. The remaining biomass is in the older section of the mycelium where there are no actively extending hyphal tips. The microscopic growth parameters (X_c and μ'_m) can be determined by image analysis, which might even be used to monitor growth during SSF itself (9). If the effect of environmental parameters on growth and branching is known, some prediction can be made about how the environment will influence macroscopic growth kinetics.

Substrate-Dependent Growth Kinetic Equations

Kinetic expressions that rely only on biomass are simple but lack flexibility in describing different systems because new values for the parameters must be determined for each new system. Another approach is to assume that growth is related to the concentration of a growth-limiting substrate (such as a soluble sugar or oxygen) by the Monod equation (6,13,14),

$$r_g = \mu_m X\left(\frac{C_S}{K_S + C_S}\right) \tag{6}$$

where K_S is the saturation constant for the substrate. The models of Muck et al. (13) and Kaiser (14), which are based on the Monod equation, are relatively complex because they describe composting processes involving several organisms and several substrates. Muck et al. have four microbial groups (fungi and yeasts, each of which is divided into thermophilic and mesophilic subgroups), each with different Monod growth parameters, and there are several substrates that are assumed to be used sequentially.

The substrate concentration in the Monod equation must be the actual concentration that the microorganism experiences, which may not be simple to define due to the presence of both solids and water. Due to the concentration gradients that arise during growth, substrate concentrations at the surface of a substrate where the microorganism is located can be quite low compared with the average concentration (6). Therefore the effect of substrate concentration should be incorporated into kinetic models only if

the growth equations are coupled with local transport equations. Neither Muck et al. (13) nor Kaiser (14) modeled local transport, which limits the usefulness of their models.

Various modifications of simple Monod kinetics have been used by different workers:

- Simultaneous limitation of growth by oxygen according to Monod kinetics and self-inhibition by the biomass according to the logistic equation (15)

$$r_g = \mu_m\left(\frac{C_{O_2}}{K_{O_2} + C_{O_2}}\right)X\left(1 - \frac{X}{X_m}\right) \quad (7)$$

- Simultaneous limitation by oxygen and glucose (16)

$$r_g = \mu_m X\left(\frac{C_{O_2}}{K_{O_2} + C_{O_2}}\right)\left(\frac{C_G}{K_G + C_G}\right) \quad (8)$$

- Self-inhibition of growth by biomass late in growth (according to logistic kinetics) combined with substrate inhibition kinetics (17)

$$r_g = \mu_m\left(\frac{C_S}{K_S + C_S + C_S^2/K_i}\right)X\left(1 - \frac{X}{X_m}\right) \quad (9)$$

where K_i is the inhibition constant for the substrate.

Modeling Effects of Microbial Activities on the Environment

Microbial activities affect the environment in SSF, and the changes in the local environment in turn affect growth. The effect of microbial activities on the local pH and water activity in SSF have received almost no attention, with the exception of Muck et al. (13) who related pH changes to acetic and lactic acid production rates and the buffer index of the solids. This section therefore is restricted to describing how workers have modeled consumption of nutrients (either the carbon source or oxygen), the generation of metabolic heat, and effects of growth on porosity in the substrate bed.

In models that describe the dependence of the growth rate on substrate concentration, it is also important to describe the effect of the growth activities on the substrate (i.e., substrate consumption). Sometimes a constant growth yield coefficient has been assumed,

$$r_s = -\frac{1}{Y_{XS}}\, r_g \quad (10)$$

where Y_{XS} is the observed yield coefficient. The negative sign indicates that substrate is consumed during the growth reaction. However, many workers have assumed that substrate is used for both growth and maintenance and have assumed the relationship

$$r_s = -\frac{1}{Y'_{XS}}\, r_g - m_s X \quad (11)$$

where Y'_{XS} is the true growth yield on the substrate, and m_s is the maintenance coefficient. This equation has been applied both to nutrients within the substrate and to oxygen.

If the particle structure is provided by a macromolecule that is degraded by the microorganism, such as during the utilization of predominantly starchy materials by amylolytic microbes, or the degradation of cellulosic substrates by cellulolytic organisms, then the particle will shrink during the fermentation. This can affect bioreactor performance through effects on bulk transport. For example, a decrease in particle size caused the substrate bed in a packed-bed bioreactor to pull away from the walls, allowing the air to bypass the bed and reducing the availability of oxygen within the bed (18). A simple model for particle size reduction assumes that oxygen is the limiting substrate and that it is consumed at the substrate–biomass interface (19,20). As the fermentation proceeds, the oxygen must diffuse through a biomass layer that increases in thickness, lowering the rate of diffusion and therefore the rate of reaction. Further developments of shrinking-substrate models are necessary because most substrates have complex compositions, and microbes can produce a range of enzymes, often in sequential order according to preferences for different macromolecules.

The generation of large amounts of metabolic waste heat can cause overheating of the substrate, and the kinetics of waste heat generation must be described in models of heat transfer in SSF. The simplest approach is to assume a constant relationship between the rate of production of new biomass and the rate of generation of metabolic heat, r_Q (11,15):

$$r_Q = Y_{QX} r_g \quad (12)$$

where Y_{QX} is the differential yield coefficient for waste heat production during growth. However, if maintenance metabolism is significant, then a growth and maintenance model is appropriate (21),

$$r_Q = Y'_{QX} r_g + m_Q X \quad (13)$$

where Y'_{QX} is the differential yield of heat from the metabolism associated with growth, and m_Q is the maintenance coefficient for heat production.

The rate of metabolic heat production does not have to be directly related to the growth rate. Saucedo-Castaneda et al. (3) related heat generation to the carbon dioxide evolution rate, which in turn was related to the substrate utilization rate by a constant yield,

$$r_Q = Y_{QCO_2} r_{CO_2} = Y_{QCO_2}[Y_{CO_2S}(-r_s)] \quad (14)$$

where Y_{QCO_2} is the differential relationship between heat released and CO_2 evolved (J/g), and Y_{CO_2S} is the yield of carbon dioxide from the substrate (g/g). This equation was indirectly related to the growth equation because logistic growth kinetics were assumed, and substrate consumption was modeled by a growth and maintenance model.

Instead of simply using experimentally determined heat yield coefficients (e.g., Y_{QX}) several workers have applied a bioenergetic approach. Gutierrez-Rojas et al. (17)

proposed a stoichiometric equation for microbial growth on sucrose and estimated the enthalpy change associated with the growth reaction from the enthalpies of formation of the reactants and products. Larroche and Gros (22) considered the energetics of growth in SSF in even greater detail, estimating ATP generation and expenditure during growth. Muck et al. (13) related the rate of energy release to the rate of consumption of each of four substrates (water-soluble carbohydrate, ethanol, lactic acid, and acetic acid), estimating the fraction of the energy of combustion of the substrates that is actually released as waste metabolic heat. Finally, Rodriguez Leon et al. (23) estimated the heat yield coefficient from considerations of the degree of reductance of the substrate and biomass.

The growth of the microorganism into the interparticle voids affects the porosity of the substrate bed according to the relationship (24)

$$E = E_0 - X/(\rho_x W_s V_r) \tag{15}$$

where ρ_x is the wet biomass density (g/cm^3); W_s is the biomass dry matter content (g/g); V_r is the reactor volume, and E_0 is the initial porosity. This reduction in void fraction affects the pressure drop through packed beds (5) and reduces the effective diffusion coefficient within the substrate bed in tray bioreactors (24).

Death Kinetics

Overheating is a common problem in SSF, and several models have attempted to describe both growth and death of microorganisms. To date, death kinetics have largely been assumed to be first order (11,12),

$$r_d = k_d X_v \tag{16}$$

where the specific death rate, k_d, is a function of temperature. The effect of temperature on the parameters in growth and death kinetic expressions will be discussed later. The incorporation of death kinetics into a model divides the population into viable (X_v) and dead (X_d) subpopulations, representing a very simple level of segregation. On the other hand, Sargantanis et al. (21) assumed that their organism was killed instantly when the temperature reached 43 °C. Not all models considering the effect of temperature actually have death kinetics in them; some describe only a decrease in net growth rate with temperature (3).

Effects of Environmental Variables on Growth and Death Kinetics

The effect of the environment on kinetic parameters is quite important because it is difficult to control environmental variables; some may change quite significantly during growth, and some might limit growth. The effect of nutrient concentrations (including O_2) have already been considered. Of most interest here are temperature, water activity, and pH. Of these, temperature has received most attention due to the overheating problems that are commonly encountered in SSF. For bioreactors with large temperature gradients, heat transport equations must be built into the model if growth is to be adequately described.

Regarding the effect of temperature, most attention has been given to the effect of the temperature on the rate constants μ_m and k_d. Simple empirical equations are normally fitted by least squares regression to growth rate data collected between the minimum and maximum growth temperatures. Polynomial expressions (11,14,15) or double exponential expressions (3,17) have been used to describe the variation in specific growth rate with temperature. Arrhenius-type relationships have also been used to describe the effect of temperature on both the specific growth rate and the specific death rate (12),

$$\mu_m = \mu_{mo} \exp\left(\frac{-E_g}{RT}\right) - k_{do} \exp\left(\frac{-E_d}{RT}\right) \tag{17}$$

where R is the universal gas constant, T is the absolute temperature, E_g and E_d are the activation energies for growth and death, and μ_{mo} and k_{do} are rate constants. Least squares regression usually allows an adequate fit of such equations, although the biological significance of the parameters is not clear because the equations are not formulated on any theoretical basis.

Other growth parameters may be expressed as functions of the environmental conditions. Sargantanis et al. (21) expressed not only μ_m but also X_m in the logistic equation as empirically fitted functions of temperature and moisture content. Saucedo-Castaneda et al. (3) expressed the maximum biomass, X_m, in the logistic equation as an empirically fitted fourth-order polynomial function of temperature.

The model of Muck et al. (13) has described the effect of the environment on growth in most detail. In addition to using Arrhenius-type equations to describe the specific growth and death rates, with different parameter values for the different microbial groups, they also described the effect of pH and water activity on growth. However, for both these environmental variables the effects on the specific growth rates of the various microbial groups were incorporated simply by using empirical equations, fitted to experimental kinetic data, to modify the maximum specific growth rate. Gervais and Simatos (25) have addressed the effect of water activity on growth on solid surfaces in greater depth, although they consider the radial extension rates of fungal colonies. The effect of water activity on the overculture situation in SSF has not been addressed.

Modeling Branching, Penetration, and Differentiation

Fundamental phenomena in the growth of filamentous fungi on solid surfaces have received some attention. Some models address extension of the hyphal tip at a very fundamental level, but these models have not been extended to describe morphological development. Others can describe branching and certain aspects of hyphal morphogenesis but have not been extended to describe overall growth kinetics (8). In any case, models describing the extension and branching of individual hyphae describe the system in far more detail than is required in models of bioreactor performance, where only overall mycelial con-

centrations are of interest. However, models describing hyphal phenomena might be useful if they concern themselves with the overall population of hyphae or hyphal tips, and this section addresses such models briefly. The equations are not presented but can be found in the original articles.

The symmetric branching model of Viniegra-Gonzalez et al. (8,9), described earlier in conjunction with the logistic equation, makes many simplifying assumptions about the branching process, essentially that all hyphae branch after reaching a certain length, and the two branches arising are identical. These do not happen in practice but may be approximated on average in the large population of hyphae within a mycelium. As shown in equation 5, on this basis the relationship between some microscopic and macroscopic growth parameters can be derived.

Edelstein and Segel (26) used a modeling approach based on average tip and hyphal concentrations per unit area of substrate surface and describing the generation of tips by branching and the movement of hyphal tips across the surface due to hyphal extension. Growth of tips depended on the local nutrient concentration. This approach avoids the great complexity that would be involved in constructing a model on the basis of intracellular events, tip extension, and branching events, and some of the parameters such as tip densities can be estimated experimentally (27). Laszlo and Silman (28) used a cellular automata approach to model mycelial densities during fungal colony expansion over solid surfaces. In this approach a matrix of small squares is superimposed on a flat substrate surface. Initially, one or more squares are designated as living, to represent an inoculum. The fungus can then grow into adjacent squares, with simple rules based on the number of occupied squares adjacent to a vacant square being used to determine whether that vacant square will be colonized. This approach can give insights into density patterns within fungal colonies (28) but otherwise is of limited usefulness for modeling growth kinetics in bioreactors. Edelstein (29) developed a model describing colony expansion, assuming no nutrient limitations, but based on processes of branching, anastomosis, tip death, and hyphal death. Only the average frequencies of these processes need to be known. Different frequencies for the various processes can lead to very different colony morphologies. These approaches could be adapted to describe growth during overculture in SSF.

Penetration into the substrate is an interesting aspect that has not received much attention, although Jurus and Sundberg (30) gave a qualitative account of the penetration of *Rhizopus oligosporus* into soybeans. More recently Ito et al. (31) observed an exponential decrease in the concentration of hyphae below the surface of the rice in koji. Experimentally they observed that the quality of the koji was strongly correlated with the degree of mycelial penetration: at lower degrees of penetration, the activities of undesirable enzymes was high, leading to poorer quality koji. Although they proposed a mathematical equation to describe the decrease in hyphal concentration with depth, it was simply fitted to the data collected at each time point. A dynamic model describing the increase in concentration of penetrative hyphae as a function of time and distance was not proposed. Such dynamic models might be important because penetration into the substrate has the potential to increase substrate degradation.

Differentiation can be quite important in SSF processes as an essential step in the production of spores or secondary metabolites (32). Georgiou and Shuler (33) modeled growth and differentiation of *Aspergillus nidulans* on a solid medium containing glucose. The model could predict the increases in density of the various types of biomass (vegetative, competent, differentiated, and conidial) during overculture.

MODELING OF LOCAL TRANSPORT PHENOMENA

Local mass transfer processes have the potential to limit the rate of growth in solid-state fermentation. Most of the local transport phenomena occurring in SSF, such as diffusion through porous media and stagnant films, have been well studied in other systems. However, the translocation of nutrients within the fungal mycelium itself is poorly understood. The work done on nutrient diffusion during radial expansion of colonies on agar surfaces will not be described here because SSF usually involves overculture of the organism across the whole surface, for which nutrient diffusion horizontally toward the organism is not important. In overculture the key spatial dimension is the one passing vertically through the substrate (Fig. 3).

This section indicates how modeling of these local transport phenomena has improved our understanding of what controls growth in SSF. Emphasis is placed on the mathematical descriptions of the phenomena. Model predictions about system behavior are mentioned but not discussed in detail. This section concentrates on mass transport phenomena because local heat transfer has not been considered in any mathematical model of SSF to date. Indeed, local heat transport is less likely than local mass transfer to limit SSF performance, although it might have unknown significance, for example, through indirect effects on mass transfer coefficients.

Diffusion and Reaction

Early modeling studies relevant to SSF considered the overculture of microorganisms on solid agar media, with glucose as the substrate (33,34). Two growth periods are predicted: an early period during which the glucose concentration at the substrate surface is relatively high and consequently the microbe grows at its maximum specific growth rate (i.e., exponential growth), followed by diffusion-limited growth during which the glucose concentration at the substrate surface is very low and the growth rate depends on the flux of glucose to the surface. The duration of the exponential growth phase depends strongly on the initial concentration of glucose because this determines how long it takes for the surface glucose concentration to fall to limiting levels. Analytical solutions to the diffusion equation can be obtained for short time periods but cannot describe the whole growth period. In any case, the usefulness of these models is limited because relatively few SSF processes contain a soluble sugar as the major carbon source.

Mitchell et al. (6) proposed and solved numerically a model for growth of *R. oligosporus* in overculture with starch as the carbon source. This allows some insight as to how mass transfer within the substrate itself can limit growth. In the experimental system a membrane filter prevented penetration of hyphae into the substrate. Five steps in the growth process were modeled (these correspond to processes 5 to 9 in Fig. 3):

- Glucoamylase release into the substrate by the fungus located in a plane at the substrate surface; an empirical equation was used
- Diffusion of glucoamylase through the substrate according to Fick's law
- Hydrolysis of starch by the glucoamylase according to Michaelis–Menten kinetics
- Generation of glucose by the glucoamylase according to Michaelis–Menten kinetics and the diffusion of the glucose through the substrate; this is a process of simultaneous reaction and diffusion
- The uptake of glucose by the fungus at the substrate surface according to Monod kinetics based on the surface glucose concentration.

The expression of interest here is that for simultaneous reaction and diffusion. For diffusion in an overculture system in which net diffusion occurs only in the vertical direction, as shown in Figure 3, the balance equation for compound A has the general form

$$\frac{\partial A}{\partial t} = r_a + D_a \frac{\partial^2 A}{\partial z^2} \tag{18}$$

where r_a is the rate of generation of compound A, D_a is the diffusivity of A, and z is the vertical spatial coordinate. This equation, with the appropriate sign on the reaction term, can refer to the generation and diffusion of hydrolysis products within the substrate, or to the diffusion and consumption of nutrients within a biomass layer.

Microscopic models incorporating local transport equations are useful tools for understanding the role of local transport phenomena in limiting growth in SSF. The model of Mitchell et al. (6) predicts a short phase of exponential growth during the early period when the glucoamylase is still near the surface and the surface glucose concentration is much higher than the saturation constant for growth. As the biomass at the surface increases, the increased glucose uptake rate causes the surface glucose concentration to fall below the saturation constant. This is followed by a period of slow deceleration of growth. This kinetic profile has been observed in SSF by several workers. The model is able to explain how such kinetics might arise from interactions among the processes of glucoamylase diffusion, starch hydrolysis, and glucose diffusion.

Recently, Rajagopalan and coworkers have made two extensions to this model. Rajagopalan and Modak (16) incorporated an expanding biomass film of constant density, with simultaneous reaction and diffusion of oxygen from the bulk air above the biomass layer and of glucose from the substrate below the biomass layer. This model suggests that at different times during growth either glucose or oxygen may be limiting, but that overall oxygen limitation might be more important than glucose limitation. They ran simulations under different bulk oxygen concentrations, such as might occur in different positions in a bioreactor, foreshadowing the linking of both local and bulk transport equations in a single model. Rajagopalan et al. (35) then further modified the model to describe a decrease in particle size as glucose is used up. Their predictions agreed reasonably well with the experimental data of Nandakumar et al. (20), who measured decreases in particle size during growth of *Bacillus coagulans* on wheat bran particles. The predictions also confirm the observations of Mitchell et al. (6) that glucoamylase diffusion plays a crucial role. At low diffusivities most of the glucoamylase remains near the surface. This quickly depletes the starch at the surface, leaving much of the enzyme without substrate to act upon. As a consequence, the rate of glucose production within the substrate particle falls, and growth becomes limited by the rate of glucose production in the deeper regions of the particles.

These modeling studies have used homogeneous gel substrates with very high water contents in which diffusion is the major mass transport phenomenon. However, SSF substrates may have high solute concentrations and contain cellular structures. Sorption phenomena may restrict diffusion (25), capillary action may occur, and cell walls may present an impermeable barrier. The mechanisms of mass transport within real solid substrates deserve greater attention.

Film Transfer of Oxygen

Another important local transport process is that of diffusion of oxygen across the stagnant gas film into the thin liquid film at the substrate surface (Fig. 3). If the biomass is assumed to be surrounded by a thin liquid film that offers the major resistance to mass transfer, then the oxygen transfer rate, N, can be expressed as (36)

$$N = K_L a(C^* - C) \tag{19}$$

where C^* is the concentration of oxygen in the bulk gas phase, C is the concentration of oxygen at the biomass–liquid interface, a is the area over which mass transfer occurs, and K_L is the mass transfer coefficient. This equation was used to estimate the overall mass transfer coefficient ($K_L a$) for a packed-bed bioreactor. It can be applied also at the microscopic scale, although it can be difficult to measure local oxygen consumption rates and oxygen concentrations (36).

Alternatively, the biomass film can be assumed to be directly in contact with the gas phase, with the transport resistance residing in a stagnant gas film. In this case the variable C in the above equation is replaced by $HC_{O_2,F}$, where H is Henry's constant, and $C_{O_2,F}$ is the concentration of oxygen in the biomass film. This expression has been used to describe oxygen transfer to the biomass in a tray bioreactor (15). Film transfer of oxygen in SSF needs more attention because significant amounts of fungal mycelium are in direct contact with the air phase and are not immersed in the liquid film at the substrate surface (37).

MODELING OF BULK TRANSPORT PHENOMENA

Recognizing Subsystems within SSF Bioreactors

In developing a model of bulk transport phenomena occurring during SSF, we must consider the level of detail to be modeled and identify simplifications of the model structure. As pointed out earlier, at the macroscopic scale, up to four subsystems can be identified in SSF bioreactors: the bioreactor wall, the headspace gases, the solid particles themselves, and the interparticle gases (Fig. 2). The presence and arrangement of these various subsystems and whether the subsystems themselves are homogeneous or have concentration or temperature gradients depend on the particular bioreactor and how it is operated. Macroscopic models contain transport equations to describe the heat and mass transfer processes within and between the subsystems recognized by the model, and between the subsystems and the surroundings.

It is often convenient to ignore the bioreactor wall in order to simplify the model. This may be appropriate if the wall is maintained at a constant temperature by a cooling system, making the bioreactor wall part of the surroundings. Alternatively, if the resistance to heat transfer between the bioreactor wall and the outside air is much greater than the resistance to heat transfer from the substrate material to the bioreactor, the overall heat transfer coefficient can be assumed to be similar to the outside heat transfer coefficient. This also has the implicit assumption that the energy stored in the bioreactor wall is negligible.

It is usually convenient and appropriate in macroscopic models to lump two or more of the phases together as a pseudohomogeneous subsystem having the mass or volume-averaged properties of the individual phases (Table 1). Most commonly, a substrate bed consisting of the moist solid particles and the interparticle gases is defined as a single subsystem. In the pseudohomogeneous matrix that comprises this system, the temperatures and water potentials of the gas and solid at the same location are assumed to be the same. The validity of this assumption depends on whether the bioreactor operation allows the gas and solid to approach equilibrium. In addition, note that the temperature and water potential of this pseudohomogeneous matrix can vary from position to position.

Depending on how well subsystems are mixed, macroscopic models may only need to describe transfer between subsystems, or they may also need to describe transfer within subsystems. If a subsystem is well mixed, then there will be no thermal and concentration gradients across the subsystem, and there will only be transfer equations to describe the transport between subsystems. Therefore the spatial coordinate does not appear in the transfer equations for such subsystems. For subsystems that are static or poorly mixed, there will also be transfer equations to describe the transfer processes within subsystems. For example, equations incorporating the spatial coordinate may be required to describe solid or gas flow patterns in the bioreactor, and to describe temperature and concentration gradients within the substrate bed. Such models can be used to explore bioreactor design and operation strategies to minimize these gradients.

The Bulk Transport Phenomena Occurring in SSF Bioreactors

Through their interactions with local transport processes, bulk transport processes determine the local environment experienced by the microorganism and therefore influence the performance of SSF bioreactors. Macroscopic models predict bioreactor performance by combining empirical descriptions of growth kinetics with expressions describing bulk transport between and within the subsystems (solid, gas, and bioreactor wall), and between these subsystems and the surroundings. Local transfer mechanisms are currently ignored in these models. Momentum transfer, heat transfer, and mass transfer may be involved, and any of these may potentially limit the performance of an SSF bioreactor. Table 2 indicates the mechanisms for these phenomena in various SSF bioreactor types. Of these phenomena, heat transfer has received the most attention.

Conversely, there is very little information regarding momentum transfer processes, that is, those processes involving material flow and that affect mixing and shear phenomena. Within bioreactors, both mixing of gas molecules within the gas phases in the bioreactor and mixing of solid substrate particles may be important. To date, the mixing of solids in agitated SSF systems has not been studied, and there are only two investigations of gas mixing and flow behavior (40,42). Residence time distributions for gas molecules indicate that operational parameters such as the gas velocity and solids loading affect the presence and extent of stagnant gas regions within the substrate bed and the homogeneity of the headspace gas phase (40,42). Further, such investigations are crucial because these flows underlie the other transport phenomena, such as heat and mass transfer. Due to the paucity of information, current macroscopic models assume ideal flow behavior such as plug flow of gases in packed beds or perfect mixing of the solids and gases in rotating drums.

The flow of solids leads to shear forces due to interactions of the solid substrate particles with each other and with the bioreactor wall. These shear forces are complex and therefore difficult to characterize. Thus, despite their potential to affect substrate characteristics and adversely affect microbial growth, and significant speculation regarding their importance in SSF systems, shear forces are not described in any macroscopic model of SSF.

Heat transfer processes have received attention because overheating due to release of metabolic waste heat is a very significant problem in SSF even at small scales. Rates of metabolic heat production reported in SSF range from about 3 to 330 kJ (kg dry matter)$^{-1}$ h^{-1}. Removal of metabolic heat is more difficult in SSF than in liquid fermentation processes for a number of reasons. First, compared to liquid media, solid substrates have low heat capacities and low thermal conductivities, which means that although less heat is required to raise the temperature of the solid substrate by the same increment, the solid substrate loses heat less readily than liquid media. As well, due to the concentrated nature of SSF systems, biomass concentrations are higher and heat production on a volumetric basis is therefore also correspondingly higher. Heat removal in SSF system occurs by convection, conduction,

Table 1. Subsystems in SSF Bioreactors and Their Simplification

Subsystem and references	Detailed view of the system (macroscopic viewpoint)	Simplification of the system appropriate for modeling	Mixing
Trays: may be stirred (the whole room is the bioreactor) (10,15,24)	Headspace gas (flowing around tray)	Headspace gas	Assume headspace infinite and well mixed, such that one tray represents all trays
	Solid particles (static) Interparticle gas (no forced flow)	Assume the bed consists of a pseudohomogeneous matrix	May be intermittently mixed, but usually not appropriate to assume that the bed is well mixed; gradients will exist
Packed beds and similar stirred reactors (3, 38, 39)	Solid particles (mostly static) Interparticle gas (flowing due to movement within bed)	Assume the bed consists of a pseudohomogeneous matrix	If mixed, usually intermittently; therefore will still have end-to-end gradients within the bed
Rotating drums and similar stirred reactors (21, 40)	Headspace gas (flowing) Solid particles (flowing) Interparticle gas (flow depends on movement within bed)	Headspace gas Assume the bed consists of a pseudohomogeneous matrix	Depending on operation may assume these two subsystems are each well mixed, or may allow for axial variation; if well mixed, may assume that headspace and bed are in equilibrium (i.e., 1 subsystem)
Air–solid fluidized beds (growth in this bioreactor has not yet been modeled)	Solid particles (flowing) Interparticle gas (flowing)	Assume a pseudohomogeneous matrix, with mixing keeping the whole bed homogeneous, and that aeration air is moist so that system is close to moisture equilibrium	Typically well mixed

Note: Subsystems are in addition to the bioreactor wall itself.

and evaporative cooling, with the contributions of each mechanism depending on the bioreactor design and operation.

The important bulk mass transfers in SSF bioreactors involve the gas phase: oxygen supply and carbon dioxide removal for aerobic processes, and the removal of water vapor during evaporative cooling. The transfer processes described by macroscopic models may include transfer across interfaces, diffusion within stagnant gas regions, and convection by bulk gas flow. As noted in Table 1, it may be appropriate to divide the gas phase into a headspace gas phase and an interparticle gas phase. The degree of continuity between these phases depends on the type of bioreactor and how it is operated.

Solid-state bioreactors differ in the types of mixing that they provide and in the relative importance of flow and diffusion processes in the bulk transport phenomena that occur. This section discusses mathematical modeling of bulk transport in tray, packed-bed, and rotating drum bioreactors. Models have not been developed to describe transport phenomena within other bioreactor types. The following sections discuss, for each bioreactor, the approaches to modeling the transport phenomena identified in Table 2. Although these transport phenomena have also been studied in non-SSF systems, the discussion is limited to SSF itself.

Macroscopic Transport Phenomena in Tray Bioreactors

Tray bioreactors often consist of a large number of individual trays incubated within a room in which the temperature and the humidity of the air may be controlled. However, mathematical models of trays usually describe only a single tray, assuming homogeneous conditions within the headspace surrounding the multiple trays in the bioreactor. This single tray comprises either an intact or a perforated tray on which rests a thin layer of static solid substrate. Two distinct gas phases exist in tray SSF systems: the first is the interparticle voids within the substrate itself; the second is the bulk gas phase, which consists of the headspaces above and below the tray.

Heat Transfer in Tray Bioreactors. Only Rajagopalan and Modak (15) have described heat transfer during tray SSFs. In their model, sensible heat transfer occurs in the vertical direction by conduction through the substrate,

$$\frac{\partial T}{\partial t} = \frac{k}{\rho_s C_p}\frac{\partial^2 T}{\partial y^2} + \frac{\Delta H r_g}{C_p} \tag{20}$$

and by convection from the substrate surfaces to the headspaces above and below the substrate bed (the base of the tray was assumed to be open),

$$-k\frac{\partial T}{\partial y} = h(T_0 - T) \tag{21}$$

where T is the temperature of the bed; t is time; k, C_p and ρ_s are the thermal conductivity, heat capacity, and density of the bed, respectively; ΔH is the enthalpy of the growth reaction; y is the spatial coordinate; T_0 is the temperature

Table 2. Important Transport Phenomena within SSF Bioreactors

Bioreactor and subsystem	Heat transport phenomena	Mass transport phenomena	Momentum transport phenomena
All bioreactor systems: the microscopic phenomena within the substrate particle are usually greatly simplified in macroscopic models describing bioreactor performance. Also, as noted in Table 1, the solid and gas subsystems are often assumed to form a single pseudohomogeneous matrix.			
Within the solid particle	This is the site of heat generation by the microbe growing at the surface of the solid particles.	Diffusion of enzymes and nutrients within the solid particle (6, 11, 33, 34)	
Between the solid and gas subsystems within the substrate bed	Conduction, evaporation of water, and free or forced convection (depending on bioreactor operation) from solids to interparticle air	Evaporation of water, and transfer of O_2 and CO_2 between the solid particles and the interparticle air (15)	
After assumption of a pseudohomogeneous matrix	Solid and gas in thermal and moisture equilibrium	Transfer of O_2 and CO_2 between solid particles and interparticle air Transfer of water to maintain saturated air (38,39)	
Trays: the solid particles and interparticle air may or may not be considered as a single pseudohomogeneous substrate bed. The whole room is taken as the bioreactor. Individual trays are assumed to be open at the top and to have perforated bases.			
Transfer across the substrate bed	Conduction (15) and natural convection within the bed, leading to spatial gradients	Diffusion (10, 15) and free convection of gases and water vapor within the bed, leading to spatial gradients	
Involving the tray wall or the headspace gases	Conduction to the tray walls, conduction within the tray wall, and then convection between the tray wall and the headspaces Evaporation and free or forced convection (depending on airflow), between the exposed bed surfaces and the headspace gases	Convection of gases (15) from the exposed bed surfaces into the headspace Evaporation of water from the substrate into the headspace at the exposed bed surfaces	Flow of headspace gases around the trays (15)
At inner surface of the room walls	Convection to the tray room walls Condensation on room walls		
	Packed beds and similar stirred reactors, assumed to be vertical cylinders		
Transport through the pseudohomogeneous matrix	Conduction, convection and evaporation axially along the substrate bed (38,39) Conduction radially within the substrate bed (3,38) Both axial and radial gradients may exist (38)	Convection of air in the axial direction due to the forced aeration of the bed. Although axial O_2 gradients occur, they are usually not as important as the temperature gradients Radial mass transfer should be negligible	Gas flow patterns through the bed (plug flow without dispersion is often assumed) (3,38,39) Pressure drop between the bottom and top of the substrate bed (40)
At the inner surface of the bioreactor wall	Conduction radially between the substrate and the inside surface of the bioreactor wall (3,38)		
	Rotating drums and similar stirred reactors, assumed to be horizontal cylindrical drums		
Transport through the pseudohomogeneous matrix	The role of conduction depends on the degree of mixing; may assume that bed is well mixed (21,41) Convective flow within the bed due to mixing Usually well mixed radially, but axial gradients may exist	The role of diffusion depends on the degree of mixing Convection of gases within substrate voids due to mixing Axial O_2 gradients may occur but are usually not important	Mixing and flow of the solid substrate Shear forces; impact and abrasion due to solids interacting with each other

Table 2. Important Transport Phenomena within SSF Bioreactors (*continued*)

Bioreactor and subsystem	Heat transport phenomena	Mass transport phenomena	Momentum transport phenomena
Between the bed and the headspace	Convection between the substrate bed surface and the headspace (41) Evaporative cooling at the surface of the substrate bed (41) Bulk flows of gases between the interparticle voids and the headspace carry sensible energy	Gas and water vapor exchange between the substrate bed and headspace gases (41). This may involve bulk flows exchanging gas between the headspace and interparticle voids.	Bulk gas flows between the interparticle voids and the headspace
Within the headspace	Convective flow of the headspace gases, carrying sensible energy (41)	Convective flow of the headspace gases (41)	Flow patterns of the headspace gases (41)
At the inner surface of the drum wall	Conduction from the substrate to the drum wall and convection between the headspace gases and the inside surface of the drum wall (41)		Shear forces, impact, and abrasion due to solids interacting with the drum wall
Air–solid fluidized beds, assumed to be vertical cylinders. The aeration air is assumed to be moist so that the bed is close to moisture equilibrium. Growth in these bioreactors has not yet been modeled.			
Within the well-mixed substrate bed	Whole bed well mixed such that there are no thermal or moisture gradients	Whole bed well mixed, such that there are no water vapor, CO_2, or O_2 gradients	Pressure drop across the bed itself, fluidization phenomena Shear, impact, and abrasion due to interactions between particles in the bed
At the inner surface of the bioreactor wall	Convection from the bed to the bioreactor wall		Shear, impact, and abrasion due to interactions between the solid particles and the wall
All bioreactors			
Involving the entry and exit of the airstream	Energy enters and leaves the system in the airstream blown through the bioreactor (3,21,38,41).	Gases and water vapor enter and leave the system in the airstream blown through the bioreactor (21,41)	Pressure drop between air pump and air exhaust
Within the bioreactor wall and between the wall and the surroundings	Conduction through the wall (the room wall for the tray system) and loss to the surroundings, often by free convection to air (21,41) or forced convection to cooling water (3,38)		

Note: These are phenomena that should be considered in the construction of mathematical models of bioreactor performance. As discussed in the text, most models greatly simplify the system by including only the most important transport phenomena or lumping phenomena together. In this table not all phenomena have been mentioned. In some places where transport processes occur in parallel, minor contributors have been ignored. In addition, the importance of the transport phenomena mentioned will depend on how the particular bioreactor is designed and operated. After identifying the phenomena to be modeled, the next task in modeling is to find appropriate expressions to include in the mass and energy balances that describe the system. Expressions that have been used to date in macroscopic models of SSF bioreactors are discussed in the text. Numbers in parentheses indicate references.

of the surrounding air, and h is the heat transfer coefficient between the bed and the surroundings. The model predicted that substrate temperatures would be highest in the center of the tray, and that the incubation temperature and substrate depth were critical to the success of tray SSFs. For a tray bed 6.35 cm high, an incubation temperature of 30 °C produced the greatest total biomass, although the maximum specific growth rate occurred at 40 °C. Likewise, substrate depths of 3 cm and under gave significantly better growth when incubated at 38 °C than when incubated at 40 °C. Simulations involving simultaneous changes of incubation temperature and substrate depth indicate that, as the bed height decreases, the incubation temperature supporting optimal growth increases and overall growth is optimal in trays less than 1 cm high. Clearly, such shallow trays would not be practical at large scales. Gas flow velocity around the trays had little effect on heat transfer between the gas phase and the bed. This model did not

describe evaporative cooling, convective heat transfer through the substrate bed due to free convection, or heat removal through the tray material.

Mass Transfer in Tray Bioreactors. A number of workers (10,15,24) have modeled mass transfer in trays. All of these models describe mass transfer within the interparticle voids; however, only the model of Rajagopalan and Modak (15) describes mass exchange between the interparticle voids and the bulk headspace.

In describing mass transfer within the interparticle voids, all three models consider transfer by diffusion in the vertical direction only. The models consider transfer of either oxygen only (10) or both oxygen and carbon dioxide (15,24). None of the models describes the mass transfer of water. All three models neglect the possibility of convective mass transfer due to thermal gradients within the substrate. The equations used in these models for mass transfer within the substrate bed take the general form of a balance of compound A (either O_2 or CO_2) in the interparticle gas phase,

$$\frac{\partial \varepsilon C_a}{\partial t} = D_a \frac{\partial^2 C_a}{\partial y^2} - R_a \tag{22}$$

where t is time, y is vertical distance within the bed, ε is the porosity, and D_a and C_a are the diffusivity and the concentration, respectively, of A in the bed. In the case where A is oxygen, R_a is the rate of oxygen consumption (10,24) or the rate of oxygen dissolution into a liquid film layer on the substrate particles (15). For film transfer, $R_a = K_a(C_{O_2} - HC_{O_2,F})$, where K_a is an overall gas side mass transfer coefficient, H is Henry's constant for oxygen and $C_{O_2,F}$ is the oxygen concentration in the film liquid. In the case where A is carbon dioxide, the minus sign in front of R_a in the previous equation is changed to a plus sign, and R_a is the rate of carbon dioxide production (15,24).

The porosity in the previous equation is an important term and received further attention within the models of Rajagopalan and Modak (15) and Auria et al. (24). Rajagopalan and Modak (15) calculated D for the oxygen and carbon dioxide gases as

$$D_i^b = \frac{\varepsilon}{\tau} D_i' \tag{23}$$

where ε is the porosity of the substrate bed, τ is the tortuosity factor (taken to equal 1.7), i is the gas (either carbon dioxide or oxygen), and D^b and D' are the diffusivities in the bed and in air. However, the porosity was held constant for any one simulation; therefore porosity changes caused by growth were neglected. Auria et al. (24) used their model to estimate the diffusivity within static beds consisting of *A. niger* growing on beads of ion exchange resin impregnated with a nutrient medium. The diffusivity decreased with increasing biomass concentration, and at high biomass concentrations of 27 mg g^{-1} dry matter, the diffusivities of oxygen and carbon dioxide were less than 5% of their bulk diffusivities in air. The effect of growth on porosity and gas diffusivity deserves further attention.

The tray models of Auria and Rajagopalan are dynamic models. The model developed by Ragheva Rao et al. (10) is different in that it consists of only a single pseudo-steady-state equation. From a balance for oxygen within the bed, they developed an equation for the maximum thickness, H_c, that a tray could be without becoming depleted of oxygen at any depth within the substrate layer during the fermentation,

$$H_c = \sqrt{\frac{2D_e C_{O_2} Y}{\mu_m X_m}} \tag{24}$$

where μ_m and X_m are the maximum specific growth rate (h^{-1}) and maximum biomass concentration (mg cm^{-3}), respectively. Use of μ_m and X_m represents the worst case scenario for which the rate of oxygen demand would be greatest. This model has a number of implications for the design and operation of tray SSFs. First, with the use of perforated rather than unperforated trays, the critical bed thickness, H_c, could be increased by a factor of up to two, depending on the fraction of the area of the underside of the bed exposed to air. Second, greater substrate depths can be used if the air above the bed is oxygen enriched or if the headspace is continually replenished during operation (thereby keeping C_{O_2} as high as possible). This model is one that could be readily applied in practice, requiring only values of the D and kinetic parameters μ_m and X_m, although it has not yet been tested experimentally.

The model of Rajagopalan and Modak (15) goes further than the other models in not only describing diffusion within the bed, but also the convection of gases from the upper and lower bed surfaces into the headspaces. Headspace air is assumed to move horizontally across the substrate surface, being well mixed normal to the flow direction but with no axial dispersion. At the plane of contact with the substrate, gases diffuse into or out of the bed. The balance of oxygen or carbon dioxide at the bed surface is given by

$$\frac{\partial C_a^h}{\partial t} = -v_x \frac{\partial C_a^h}{\partial x} - \frac{D_a}{z} \left.\frac{\partial C}{\partial y}\right|_{y=L} \tag{25}$$

where A is either carbon dioxide or oxygen, x is the horizontal coordinate, y is the vertical coordinate, L is the thickness of the substrate, z is the thickness of the headspace above the bed, C_a^h is the headspace concentration, and v_x is the headspace gas flow velocity. Model simulations showed that substrate depth and porosity greatly affected mass transfer within the bed; however, the gas flow velocity around the trays had little effect on mass transfer between the headspaces and the bed. These predictions can guide design and operation of tray bioreactors.

Combined Heat and Mass Transfer in Tray Bioreactors. Finally, both heat and mass transfer processes may be important at different times during SSFs in trays. Therefore, mathematical models describing these systems will need to describe both transport mechanisms. Currently, only the model by Rajagopalan and Modak (15) does this. The phenomena described by their model are shown in Figure 4.

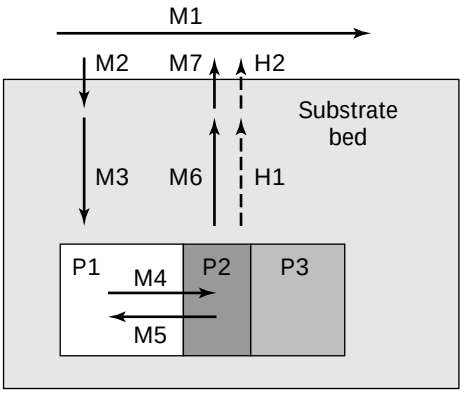


Figure 4. Diagrammatic representation of the tray bioreactor as modeled by Rajagopalan and Modak (15), indicating the three phases within the substrate bed. P1, Interparticle gas phase; P2, thin water film on the surface of substrate particles, which is the location of oxygen consumption, carbon dioxide production, and generation of waste metabolic heat; P3, solid substrate particle. The heat and mass transfer processes occurring are M1, bulk flow of air above the tray surface; M2, transfer of oxygen into the bed; M3, diffusion of oxygen within the bed; M4, transfer of oxygen from the interparticle air phase to the liquid film; M5, transfer of carbon dioxide from the liquid film to the interparticle air phase; M6, diffusion of carbon dioxide within the bed; M7, transfer of carbon dioxide from the bed to the surrounding air; H1, conduction of heat through the bed; H2, convective removal of heat at the bed surface.

The predictions of this model for substrate temperature, biomass, and gas phase concentrations of oxygen and carbon dioxide showed trends similar to those noted experimentally by Rathbun and Shuler (43) during the tempe fermentation of soybeans by *R. oligosporus*. The model predicts that the substrate temperature is the dominant influence on growth in the initial and final stages of the fermentation. During intermediate stages of the fermentation, oxygen has a large effect at the bottom of the bed (in an unperforated tray), whereas temperature effects are more important in the center of the bed.

Macroscopic Transport Phenomena in Packed-Bed Bioreactors

Heat Transfer in Packed-Bed Bioreactors. Heat transfer in packed beds has received most modeling attention because the temperature gradients obtained are more important than the gas concentration gradients, and if aeration satisfies the heat removal requirements it will also satisfy aeration requirements (44,45). In bioreactors of different designs and operated under different conditions, both axial and radial temperature gradients have been observed (45–47). Under some operating conditions, temperatures in the middle of the bed can be over 20 °C higher than the temperature of the inlet air, leading to large variations in growth and product formation across the bioreactor (46).

Saucedo-Castaneda et al. in 1990 (3) modeled radial heat transfer in an experimental packed bed in which *A. niger* was grown on cassava meal. This model was later extended by Sangsurasak to describe both axial and radial heat transfer (38). A general expression for heat transfer in a pseudohomogeneous substrate bed within a cylindrical packed-bed bioreactor, which takes into account the heat transfer processes shown in Figure 5, can be expressed as (38)

$$\rho_b C_{pb} \frac{\partial T}{\partial t} + (\rho_{ma} C_{pma} + \rho_a f \Delta H_v) V_z \frac{\partial T}{\partial z} = \frac{k_b}{r} \frac{\partial T}{\partial r} + k_b \frac{\partial^2 T}{\partial r^2} + k_b \frac{\partial^2 T}{\partial z^2} + r_Q \quad (26)$$

where ρ_a, ρ_{ma}, and ρ_b are the densities of air, moist air, and the bed, respectively; C_{pma} is the heat capacity of moist air; C_{pb} the heat capacity of the bed; f describes the relationship between saturation humidity and temperature; ΔH_v is the enthalpy of vaporization of water; k_b is the thermal conductivity of the bed; V_z is the superficial air velocity; z and r are the axial and radial spatial coordinates; and r_Q is the rate of heat generation from the growth reaction. The factor $\rho_a f \lambda$ arises because the equilibrium assumed between the substrate and air means that the evaporation of water to keep the air saturated as the air temperature increases gives the air a higher apparent heat capacity. This is a simplified approach to handling evaporative cooling,

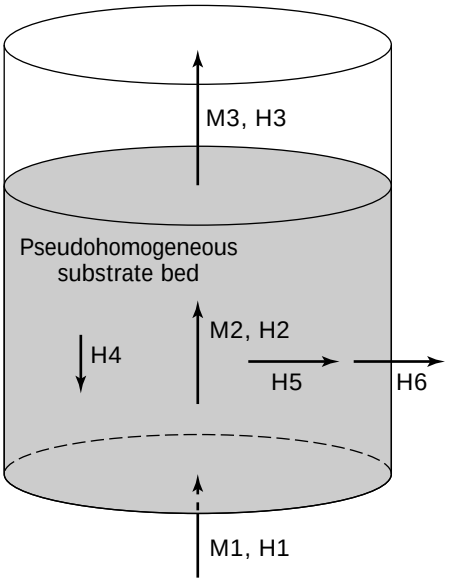


Figure 5. Diagrammatic representation of the packed-bed bioreactor as modeled by Sangsurasak (38). The substrate bed is treated as a single pseudohomogeneous phase. The microbe grows in the bed and releases waste metabolic heat. The heat and mass transfer processes occurring are M1, flow of moist air into the pseudohomogeneous bed; M2, flow of moist air within the bed, with the air being saturated at the bed temperature; M3, flow of moist air out of the pseudohomogeneous bed; H1, bulk transfer of energy into bioreactor with airflow; H2, convective heat transfer within the bed; H3, bulk transfer of energy out of bioreactor with airflow; H4, axial conduction within the bed (opposes the temperature gradient caused by convection); H5, radial conduction within the bed; H6, conduction across the bioreactor wall and convective removal by cooling water.

which can contribute as much as 65% of the heat removal from a packed bed (47).

For thin packed-bed bioreactors of about 6 cm diameter, operated with low superficial air velocities (e.g., 0.01 m s^{-1}), radial conduction is the major heat removal mechanism (3). In this particular case the terms for convection and evaporation and axial conduction can be omitted.

In addition to transfer phenomena within the bed, the other important transfer process is from the outer edge of the substrate bed to the surroundings (the bioreactor wall itself has been assumed to offer negligible resistance to heat transfer). Convective cooling at this boundary has been assumed (3,38):

$$\left.\frac{\partial T}{\partial r}\right|_{\text{wall}} = \frac{ha}{k_b}(T_{\text{surr}} - T_{\text{wall}}) \qquad (27)$$

Overheating problems occur at the top of the column when it is aerated from the bottom because the air reaching higher regions has been prewarmed in the lower regions and therefore does not cool the top as effectively. Even for quite small columns, when they are operated with the cross-sectional airflow rates of 0.01 to 0.025 m s^{-1} that are commonly used in packed-bed bioreactors, temperatures that can cause significant microbial death are predicted. The model of Sangsurasak (38) can be used as a tool in bioreactor design, operation, and scale-up. Sangsurasak used the model to explore strategies for minimizing the overheating problem at large scale. Higher aeration rates and lower height-to-diameter ratios appear to be the best strategies. However, the model has not yet been used in the actual construction of a larger-scale packed bed.

Gutierrez-Rojas et al. (17) used a different approach to modeling heat transfer during growth of *A. niger* in a packed bed on beads of ion exchange resin impregnated with nutrient medium. They defined an "elementary representative volume" of 20 mL, over which scale variations in the process variables were assumed to be negligible. However, they modeled only an individual elementary representative volume, and although their model was able to describe the experimental results obtained in a 20-mL packed bed, it is not able to describe the gas concentration and temperature profiles that occur across larger-scale packed-bed bioreactors (46).

Heat and Mass Transfer in Packed-Bed Bioreactors. The reliance on evaporative cooling means that water balances over the bed are important (45). A combined heat and mass transfer model is required to take these into account. Such a model, considering only gradients in the vertical direction, was proposed by Van Lier et al. (39), who modeled the composting of a mixture of wheat straw and horse manure. The energy balance is similar to that used by Sangsurasak (38), but because a material balance was also done for water, a simplified description of evaporation was not required. The equation was of the form

$$\rho_b C_{pb}\frac{\partial T}{\partial t} + \rho_{ma} C_{pma} V_z \frac{\partial T}{\partial z} = k_b \frac{\partial^2 T}{\partial z^2} - h(T - T_0) - \Delta H_v r_w + r_Q \qquad (28)$$

where h is the heat transfer coefficient for conduction through the side wall to the surroundings at temperature T_0, and r_W is the rate of evaporation of water from the solid phase into the gas phase.

A general mass balance equation for compound A in the gas phase, for example, oxygen or water, can be written in the following form with terms for axial convection, axial diffusion, and exchange with the metabolizing organism attached to the solid phase:

$$\frac{\partial(\varepsilon C_A)}{\partial t} = -V_Z \frac{\partial C_A}{\partial z} + \frac{\varepsilon}{\beta^2} D_A \frac{\partial^2 C_A}{\partial z^2} + r_A \qquad (29)$$

where C_A is the concentration of A at a location, D_A is the diffusivity of A in air, ϵ is the bed void fraction, β is the tortuosity and r_A is the rate of exchange of A with the solid phase. In the water balance Van Lier et al. assumed that the vapor phase remained saturated with water at the local temperature in the bed, and used this equation to calculate the evaporation rate (r_W). Film transfer of oxygen was not incorporated into the equations.

Solution of the equation set was complicated because the porosity and height of the bed changed during the composting process. However, good agreement was achieved between the model predictions and measured temperatures and bed heights. In composting, unlike other SSF processes, some self-heating is desired, and this reduces problems with overheating. Despite this, the aeration rate required to remove excess heat is still about 10-fold greater than that required to supply oxygen (39).

Other Phenomena in Packed-Bed Bioreactors. As air flows through the porous bed in a packed-bed bioreactor, there are viscous and kinetic phenomena (5). Substantial pressure drops can occur in packed-bed bioreactors (5,18,41). To date characterization of this phenomenon has been limited to demonstrating the relationship between pressure drop and the bed void fraction, which decreases as the microbe fills in the interparticle spaces. At high biomass densities, the microorganism can occupy up to 34% of the free interparticular space, reducing the relative permeability of the bed to around 1% of the original value (40). In the future, such considerations could be built into momentum balances to describe how the flow through the bed changes during the fermentation.

Macroscopic Transport Phenomena in Rotating-Drum Bioreactors

Only heat transfer phenomena (and the related mass transfer of water) have been modeled in rotating-drum bioreactors and the related rocking-drum bioreactor (RDBs). As in packed-bed systems, it is likely that in order to achieve adequate heat removal in RDBs, aeration requirements will already be met. Neither of the models developed so far (21,40) describes all of the heat transfer processes mentioned in Table 2.

Sargantanis et al. (21) modeled heat transfer during growth of *R. oligosporus* on corn grit in a rocking-drum bioreactor with a 1.2-L working volume. They described heat removal very simplistically by assuming equilibrium

between the headspace and the substrate bed; in essence, this was a single homogeneous subsystem. The model was useful for determining better operation of their rocking-drum bioreactor system; however, it is not very versatile and can not readily be adapted to describe larger-scale SSFs in RDBs, where the assumption of thermal and moisture equilibrium between the solid substrate and the headspace could be unreasonable, particularly if high aeration rates are used. For such cases a model predicting heat removal in terms of heat transfer rather than equilibrium relationships is likely to be more useful.

Stuart (41) modeled heat transfer in a 20-L rotating-drum bioreactor in which *A. oryzae* was grown on both an artificial gel substrate and on wheat bran. The model is more mechanistic than that of Sargantanis et al. (21) and considers several of the heat transfer processes listed in Table 2. As shown in Figure 6, it simplifies the bioreactor system into three subsystems: the two internal subsystems, the fermenting substrate and the headspace gas, which are contained by the third subsystem, the fermenter wall. Each subsystem is assumed to be homogeneous with respect to composition and temperature in both the radial and axial directions. The model, therefore, does not describe heat transfer processes within any of the subsystems, and the temperature of each subsystem is represented by a single time-dependent value. Stuart (41) neglected water mass transfer within the substrate bed by assuming that water at 25 °C is added to maintain the water activity within this pseudohomogeneous subsystem close to unity. Mass transfer from this bed to the headspace gas is assumed to be limited by the rate at which water vapor in the gas sublayer of the substrate bed diffuses into the bulk headspace gas.

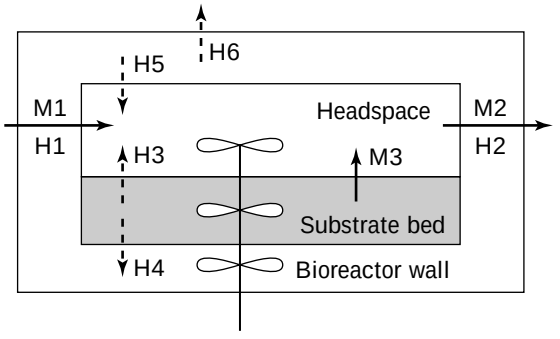


Figure 6. Diagrammatic representation of the rotating-drum bioreactor as modeled by Stuart (41), showing the three subsystems that were assumed. The pseudohomogeneous substrate bed is the site of generation of water and waste metabolic heat during growth. The mixer indicates that each subsystem is assumed to be at the same temperature at all locations within the subsystem. The heat and mass transfer processes occurring are M1, flow of moist air into the headspace; M2, flow of moist air out of the headspace; M3, transfer of water from the substrate bed to the bioreactor headspace; H1, bulk transfer of energy into bioreactor with airflow; H2, bulk transfer of energy out of bioreactor with airflow; H3, convection from the substrate bed to the headspace; H4, conduction from the substrate bed to the bioreactor wall; H5, convection from the bioreactor wall to the headspace; H6, convection from the bioreactor wall to the surroundings.

Three energy balance equations, one for each subsystem, describe heat transfer phenomena within the RDB system, and heat exchange between the RDB system and the surroundings (41). Metabolic heat is generated within the substrate and then dissipated by conduction to the bioreactor wall and by convection and evaporative cooling to the headspace gases. The overall energy balance for the substrate is

$$\frac{d[T_s M(C_{pm} + C_{pw}W)]}{dt} = r_Q - h_{sf}A_{sf}(T_s - T_f) - h_{sa}A_{sa}(T_s - T_a) - (1.2 \times 25)C_{pw}(W_i - W)M - kA_{sa}(C_I - C_B)[T_s C_{pw} + \Delta H_v - (T_s - T_a)C_{ph}] \quad (30)$$

Energy enters the headspace gas with the evaporated water from the substrate and with incoming aeration gases, and through convection from the substrate and from the bioreactor wall:

$$\frac{d[T_a G(C_{pg} + C_{ph}H)]}{dt} = T_i F_i(C_{pg} + C_{ph}H_i) + kA_{sa}(C_I - C_B)T_a C_{ph} - T_a F_o(C_{pg} + C_{ph}H) + h_{sa}A_{sa}(T_s - T_a) + h_{fa}A_{fa}(T_f - T_a) \quad (31)$$

The exiting air flow carries energy out of the headspace and thereby away from the RDB system. During fermentation, the bioreactor wall heats up due to its contact with the substrate. Energy, when it leaves the bioreactor wall, passes either to the surroundings or to the headspace gas inside the RDB:

$$\frac{d(T_f V_f \rho_f C_{pf})}{dt} = h_{sf}A_{sf}(T_s - T_f) - h_{fa}A_{fa}(T_f - T_a) - h_{fe}A_{fe}(T_f - T_e) \quad (32)$$

Thus there are two overall routes by which metabolic heat is removed from the substrate subsystem to the surroundings: with the exiting air stream or as heat lost through the bioreactor wall.

In these equations, T is temperature, M is the mass of dry solid, W is the mass of water per mass of dry solid, G is the mass of dry gas, H is the mass of water per mass of dry gas, V_f and ρ_f are the volume and density of the bioreactor material, respectively, C_P is a heat capacity, r_Q is the rate of production of metabolic heat, h is a heat transfer coefficient, A is an area of heat or mass transfer, k is a mass transfer coefficient, C is a concentration of water at either the interface (subscript I) or within the headspace bulk (subscript B), ΔH_v is the heat of vaporization of water, F is a flow of dry gas either entering (i) or leaving (o) the bioreactor system during aeration. The subscripts S, A, and F denote the three subsystems: the substrate, headspace gas, and fermenter wall, respectively. Combinations of these subsystem subscripts indicate heat or mass transfer between these materials, for example, the subscript combination "SF" represents heat transfer from the substrate

to the bioreactor wall. The other subscript terms denote dry solid matter (m), dry gases (g), liquid water (w), water vapor (h), and the external surroundings (e).

Estimation of the coefficients and areas involved in the heat and mass transfer processes during SSF in RDBs is difficult, Stuart (41) used empirical equations to estimate transfer coefficients and assumed that the substrate and headspace gas subsystems occupy constant volumes throughout the fermentation. Likewise, transfer areas were assumed to be constant and were calculated based on the interfacial areas between the substrate, headspace gas, and the bioreactor wall, when the surface of the substrate bed forms a flat plane. Heat and mass transfer coefficients and interfacial areas of contact urgently require experimental attention in SSF systems.

In order to incorporate evaporative cooling into a model of heat transfer during SSF, water balances must be determined over the substrate and headspace gas subsystems. Stuart (41) assumed that the substrate and headspace gas subsystems each consist of a water component and a dry component. Therefore four material balance equations were written, two equations to describe each subsystem.

- Balance of dry matter in the substrate:

$$\frac{dM}{dt} = r_m \tag{33}$$

- Balance of liquid water in the substrate:

$$\frac{d(MW)}{dt} = -kA_{sa}(C_I - C_B) + r_w + 1.2(W_i - W)M \tag{34}$$

- Balance of dry gas in the headspace gas subsystem:

$$\frac{dG}{dt} = F_i - F_o + r_{dg} \tag{35}$$

- Balance of water vapor in the headspace gas subsystem:

$$\frac{d(GH)}{dt} = F_iH_i - F_oH_o + kA_{sa}(C_I - C_B) \tag{36}$$

In these equations, r_m, r_w, and r_{dg} are the net rates due to microbial metabolism of substrate dry matter consumption, water production, and dry gas production, respectively.

Due to consumption of substrate dry matter during microbial growth in SSF and loss of water from the substrate to the headspace gas, the flow of gases out of the RDB ($F_o[1 + H]$) will often be greater than the air flow into the RDB ($F_i[1 + H_i]$). Because RDBs are open systems and large pressure variations in the headspace are unlikely, ideal gas behavior and a constant pressure within the headspace gas are assumed. This gives an equation (not shown) for the flow of gas out of the bioreactor (F_o) and enables the calculation of C_B (the headspace concentration of water vapor).

This model was used to explore SSF performance in RDBs over a range of scales, operations, and designs. The model predictions of temperature profiles for the substrate, bioreactor wall, and headspace gas agreed reasonably well with experimental results for SSFs in real small-scale RDB systems (41). Model simulations indicate that, whereas at small scales heat removal from the substrate bed through the bioreactor wall is most important, the required contribution of evaporative cooling increases with scale. This requires very high aeration rates with relatively dry air: simulations showed that the aeration rate in air volumes per bioreactor volume per minute (vvm) required to achieve adequate heat removal from the substrate might increase by a factor of 25 over a scale change of 1,000. Such high aeration rates would be costly. Therefore heat transfer through the bioreactor wall should be maximized to minimize reliance on evaporative cooling and to limit the amount of water that needs to be added to the system during the fermentation. At all scales, heat removal from the substrate to the headspace gases by sensible heat transfer was insignificant compared with heat transport from the substrate subsystem by conduction and evaporative cooling. However, sensible heat transfer to the headspace gases was always indirectly important in the model predictions because it affected the temperature of the headspace gases and thereby determined the amount of moisture that the headspace could hold and the driving force for evaporative cooling. This investigation of RDB performance demonstrates how models can be used as tools to explore interactions between important processes occurring during SSFs as well as operation and control options for the design of larger-scale bioreactors. This model, with its mechanistic basis, represents a first step toward developing semifundamental scale-up criteria for SSF in RDBs.

PARAMETER ESTIMATION

As highlighted in the introduction to this article, much more attention needs to be given to the estimation of the parameters involved in models of SSF systems. Models involve four main types of parameters: operational and design parameters, microbial parameters, transport parameters, and thermodynamic parameters. In the modeling work to date, many of these parameters have not been independently determined. In some cases parameters have been allowed to vary to enable the model to fit the experimental data. In other cases parameters have been borrowed from other systems involving different substrates and microorganisms, and even from work done in submerged liquid culture. Even physical parameters may depend on the substrate and its preparation and the bioreactor and how it is operated, making it important for parameters to be measured in the actual system studied. This section describes the work that has been done to determine model parameters in SSF systems.

Operational and Design Parameters

Operational and design parameters come directly from the experimental system being modeled. They include bioreac-

tor geometry, air temperatures and flow rates, and rotation or mixing rates. These parameters are important in optimizing bioreactor performance, but these considerations are beyond the scope of this article. For modeling purposes, these parameters are easy to obtain because they are determined by the operator.

Microbial Parameters

Microbial parameters include specific growth and death rates, saturation constants, yields, activation energies, lag periods, and biomass composition. Ideally these parameters should be measured in SSF systems, although this is difficult for those parameters requiring direct measurement of biomass because it is usually impossible to remove the biomass from the substrate particles. Parameters such as specific growth rates and yields are often based on indirect indicators of growth, such as oxygen uptake or protein content, but the relationship of these indicators to growth can often vary during the growth phase. In other cases, parameter values determined in liquid culture have been used, but due to the different physical environment in SSF, these values may not be appropriate for SSF.

Specific growth rates depend on the particular system used. A wide range of values have been reported, from very low values of 0.05 h^{-1} to values as high as 0.5 h^{-1} (45). In some modeling work the specific growth rate has been used as one of the fitting parameters (3,6). The stoichiometry of growth will also be highly variable, and most models consider overall yields and do not balance elements. Only Larroche and Gros (22) have done detailed stoichiometric analyses (for growth and sporulation of *Penicillium roqueforti* on buckwheat seeds). For the anabolic reactions during biomass production from starch they obtained

$$1.104CH_{1.667}O_{0.833} + 0.087NH_3 \rightarrow CH_{1.882}O_{0.603}N_{0.087} + 0.014CO_2 + 0.109H_2O \quad (37)$$

Because a range of different empirical equations have been used to describe the effect of temperature on growth, the values of the parameters do not have general significance. In any case, these parameters are typically determined by incubating cultures at different temperatures, but with any one culture being held isothermally throughout the whole growth phase. The models of Stuart (41) and Saucedo-Castaneda et al. (3) suggest that the relationships obtained from such experiments may not describe the situation in SSF where the organism is subjected to high temperatures after an initial period at the optimum temperature for growth, because their models could describe their data only if the specific growth rate was maintained unchanged at the value for the optimum temperature as the temperature increased.

Transport Parameters

Transport parameters include heat and mass transfer coefficients, diffusivities, and thermal conductivities. Relatively few efforts have been made to determine these in SSF systems. With respect to mass transfer, Auria et al. (24) estimated the diffusivities of oxygen and carbon dioxide within substrate beds as 2 to 25% of their diffusivities in air, with the values decreasing with increasing biomass content. For transfer of oxygen into the moisture films on substrate surfaces, values of $K_L a$ have been estimated by a sulfite oxidation method to range from 1,140 to 3,050 h^{-1}, with the value increasing with increasing airflow rate (48). For nutrient diffusion within particles, Mitchell et al. (6) estimated the diffusivity of glucose at 37 °C as 8.2×10^{-3} cm^2 h^{-1} compared with a bulk diffusivity in water of 2.5×10^{-2} h^{-1}. However, because they used an artificial gel substrate, this value is not of general applicability. They were not able to measure glucoamylase diffusivity using their experimental system. Diffusion has not been investigated for real SSF substrates containing structures such as fibers and cell walls.

With respect to heat transfer, coefficients for heat transfer between subsystems in SSF bioreactors have typically been estimated from literature values for non-SSF systems. However, because these coefficients can depend strongly on bioreactor design and operation, they need to be determined experimentally for the particular system used. Thermal conductivities of composts varied from 0.25 W m^{-1} K^{-1} at 20% moisture content to 1.0 W m^{-1} K^{-1} at 70% moisture content (39).

Thermodynamic and Physical Parameters

Thermodynamic parameters include enthalpies of vaporization and saturation humidities that are usually straightforward to obtain from thermodynamic databases. However, it may be difficult to estimate other parameters such as the heat capacities of substrates with complex compositions, and therefore experimental measurements should be made. Physical parameters such as substrate densities and bulk packing densities are readily measured experimentally. They vary widely between different systems.

SUMMARY AND NEEDS FOR THE FUTURE

Growth in SSF results from the interaction of three types of phenomena: microbial growth, local transport, and bulk transport. This article has surveyed the ways in which these phenomena have been described in mathematical models of SSF processes. In these models, many simplifications are made, reducing both the complexity of the system and the complexity of the processes occurring within the system. Despite these simplifications and assumptions, the usefulness of the models has been identified. Microscopic models give insights into how diffusion of enzymes, their hydrolysis products, and oxygen can limit growth. Macroscopic models give insights into how interparticle oxygen transport, and more importantly, how heat transport can limit growth within bioreactors.

SSF and traditional liquid culture systems are quite different: compared with liquid culture systems, in SSF systems water availability is more restricted, and diffusional limitations play a greater role. It is these differences that can give SSF advantages over liquid culture in certain cases; the unique SSF environment can stimulate certain products, and the low water content can lead to higher concentrations and therefore higher volumetric productivities and reduced recovery costs. These differences also lead to

some differences in approaches to modelling SSF and liquid culture. Due to the heterogeneity of SSF systems, differential equations must often be formulated and solved for each of several subsystems, and due to the importance of transport processes, many of the equations involve partial differential terms. Also, due to diffusional limitations within substrate particles, it is not simple to measure the substrate concentration experienced by the microbe. For this reason, growth kinetic expressions for the majority of models in SSF will continue to be substrate independent. This is in contrast to liquid culture, where even simple models usually take substrate concentration into account, often assuming a Monod-type relationship. However, in the future, some models will be developed that integrate the microscopic- and macroscopic-scale processes occurring within SSF. These models will give interesting insights into how the interactions among the various phenomena control growth at various stages during SSF.

More work needs to be done in establishing the validity and flexibility of current models. Due to the difficulties of measuring key process parameters in SSF, many of the parameters used in these models have been estimated, or borrowed from other systems. Therefore, more effort is required to measure model parameters in SSF systems. More attention must be paid to microbial physiology. Current expressions describing the effect of elevated temperatures on microbial growth are based on experiments where a number of cultures have been incubated isothermally at a range of temperatures. Better expressions will be developed from experiments that mimic the temporal temperature profiles characteristic of SSF systems. In addition, product formation is usually the main objective of an SSF process. Accurate description of product formation kinetics will require an understanding of how the SSF environment influences microbial product formation.

Our current ability to apply SSF technology at large scale is limited by our relatively poor knowledge of the basic engineering principles underlying SSF bioreactor performance. Rational rules for the design and operation of large-scale SSF bioreactors are crucial to the future commercial success of SSF technology, and mathematical models of bioreactor performance will be essential tools in the development of such rules. The power of the modeling approach is obvious. Modeling is a more cost-effective way of searching for optimal bioreactor designs than building large bioreactors and encountering problems experimentally. It does not replace the need to experiment, but can reduce the experimental program by avoiding poor designs and can guide the program by raising questions about the dominant mechanisms and the key process variables. However, because the models to date have been based on laboratory-scale systems, work needs to be done to confirm that the mechanisms incorporated into the models remain dominant at large scale. Several of the bioreactor designs, such as stirred bioreactors and air solid fluidized beds, have not received any modeling attention. Air solid fluidized beds have the potential to overcome many of the problems faced in other SSF bioreactors, and therefore urgently need modeling attention.

BIBLIOGRAPHY

1. R.E. Mudgett, in A.L. Demain and N.A. Solomon eds., *Manual of Industrial Microbiology and Biotechnology*, American Society for Microbiology, Washington, D.C., 1986, pp. 66–83.
2. A. Pandey, *Process Biochem.* **27**, 109–117 (1992).
3. G. Saucedo-Casteneda, M. Gutierrez-Rojas, G. Bacquet, M. Raimbault, and G. Viniegra-Gonzalez, *Biotechnol. Bioeng.* **35**, 802–808 (1990).
4. M.A. Aufeuvre and M. Raimbault, *Comptes Rendus des Seances de L'Academie des Sciences, Serie III Sciences de la Vie* **294**, 949–956 (1982).
5. R. Auria, M. Morales, M., E. Villegas, and S. Revah, *Biotechnol. Bioeng.* **41**, 1007–1013 (1993).
6. D.A. Mitchell, D.D. Do, P.F. Greenfield, and H.W. Doelle, *Biotechnol. Bioeng.* **38**, 353–362 (1991).
7. Sangsurasak, M. Nopharatana, and D.A. Mitchell, *J. Sci. Ind. Res.* **55**, 333–342 (1996).
8. G. Viniegra-Gonzalez, G. Saucedo-Castaneda, F. Lopez-Isunza, and E. Favela-Torres, *Biotechnol. Bioeng.* **42**, 1–10 (1993).
9. G. Viniegra-Gonzalez, C.P. Larralde-Corona, and F. Lopez-Isunza, *Adv. Bioprocess Eng.* **11**, 183–189 (1994).
10. K.S.M.S. Ragheva Rao, M.K. Gowthaman, N.P. Ghildyal, and N.G. Karanth, *Bioprocess Eng.* **8**, 255–262 (1993).
11. P. Sangsurasak and D.A. Mitchell, *J. Chem. Technol. Biotechnol.* **64**, 253–260 (1995).
12. K.W. Szewczyk and L. Myszka, *Bioprocess Eng.* **10**, 123–126 (1994).
13. R.E. Muck, R.E. Pitt, and R.Y. Leibensperger, *Grass Forage Sci.* **46**, 283–299 (1991).
14. J. Kaiser, *Ecol. Model.* **91**, 25–37 (1996).
15. S. Rajagopalan and J.M. Modak, *Chem. Eng. Sci.* **49**, 2187–2193 (1994).
16. S. Rajagopalan and J.M. Modak, *Chem. Eng. Sci.* **50**, 803–811 (1995).
17. M. Gutierrez-Rojas, R. Auria, J.C. Benet, and S. Revah, *Chem. Eng. J.* **60**, 189–198 (1995).
18. E. Gumbira-Sa'id, P.F. Greenfield, D.A. Mitchell, and H.W. Doelle, *Biotechnol. Adv.* **11**, 599–610 (1993).
19. M.P. Nandakumar, M.S. Thakur, K.S.M.S. Raghavarao, and N.P. Ghildyal, *Process Biochem.* **29**, 545–551 (1994).
20. M.P. Nandakumar, M.S. Thakur, K.S.M.S. Raghavarao, and N.P. Ghildyal, *Enzyme Microb. Technol.* **18**, 121–125 (1996).
21. J. Sargantanis, M.N. Karim, V.G. Murphy, D. Ryoo, and R.P. Tengerdy, *Biotechnol. Bioeng.* **42**, 149–158 (1993).
22. C. Larroche and J.B. Gros, *Biotechnol. Bioeng.* **39**, 815–827 (1992).
23. J.A. Rodriguez Leon, A. Torres, J. Echevarria, and G. Saura, *Acta Biotechnol.* **11**, 9–14 (1991).
24. R. Auria, J. Palacios, and S. Revah, *Biotechnol. Bioeng.* **39**, 898–902 (1992).
25. P. Gervais and D. Simatos, in S. Thorne ed., *Mathematical Modelling of Food Processing Operations*, Elsevier, London, 1992, pp. 137–183.
26. L. Edelstein and L.A. Segel, *J. Theor. Biol.* **104**, 187–210 (1983).
27. L. Edelstein, Y. Hadar, I. Chet, Y. Henis, and L.A. Segel, *J. Gen. Microbiol.* **129**, 1873–1881 (1983).
28. J.A. Laszlo and R.W. Silman, *Biotechnol. Adv.* **11**, 621–633 (1993).

29. L. Edelstein, *J. Theor. Biol.* **98**, 679–701 (1982).
30. A.M. Jurus and W.J. Sundberg, *Appl. Environ. Microbiol.* **32**, 284–287 (1976).
31. K. Ito, A. Kimizuka, N. Okazaki, and S. Kobayashi, *J. Ferment. Bioeng.* **68**, 7–13 (1989).
32. C. Larroche, *J. Sci. Ind. Res.* **55**, 408–423 (1996).
33. G. Georgiou and M.L. Shuler, *Biotechnol. Bioeng.* **28**, 405–416 (1986).
34. G. Lazarova, M. Kancheva, and I. Panayotov, *Acta Microbiol. Bulgarica* **15**, 63–71 (1984).
35. S. Rajagopalan, D.A. Rockstraw, and S.H. Munson-McGee, *Bioresource Technol.* **61**, 175–183 (1997).
36. M.K. Gowthaman, K.S.M.S. Ragheva Rao, N.P. Ghildyal, and N.G. Karanth, *Process Biochem.* **30**, 9–15 (1995).
37. D.A. Mitchell, P.F. Greenfield, and H.W. Doelle, *World J. Microbiol. Biotechnol.* **6**, 201–208 (1990).
38. P. Sangsurasak, Ph.D. Thesis, The University of Queensland, Brisbane, Australia, 1996.
39. J.J.C. Van Lier, J.T. Van Ginkel, G. Straatsma, J.P.G. Gerrits, and L.J.L.D. Van Griensven, *Netherlands J. Agric. Sci.* **42**, 271–292 (1994).
40. R. Auria, I. Ortiz, E. Villegas, and S. Revah, *Process Biochem.* **30**, 751–756 (1995).
41. D.M. Stuart, Ph.D. Thesis, The University of Queensland, Brisbane, Australia, 1996.
42. P. Gervais, C. Bazelin, and A. Voilley, *Biotechnol. Bioeng.* **28**, 1540–1543 (1986).
43. B.L. Rathbun and M.L. Shuler, *Biotechnol. Bioeng.* **25**, 929–937 (1983).
44. M.K. Gowthaman, N.P. Ghildyal, K.S.M.S. Raghava Rao, and N.G. Karanth, *J. Chem. Technol. Biotechnol.* **56**, 233–239 (1993).
45. G. Saucedo-Castaneda, B.K. Lonsane, M.M. Krishnaiah, J.M. Navarro, S. Roussos, and M. Raimbault, *Process Biochem.* **27**, 97–107 (1992).
46. N.P. Ghildyal, M.K. Gowthaman, K.S.M.S. Raghava Rao, and N.G. Karanth, *Enzyme Microb. Technol.* **16**, 253–257 (1994).
47. M. Gutierrez-Rojas, S. Amar Aboul Hosn, R. Auria, S. Revah, and E. Favela-Torres, *Process Biochem.* **31**, 363–369 (1996).
48. A. Durand, P. Pichon, and C. Desgranges, *Biotechnol. Tech.* **2**, 11–16 (1988).

SOLID SUBSTRATE FERMENTATION, AUTOMATION

J. Ricardo Pérez-Correa
Eduardo Agosin
Pontificia Universidad Católica de Chile
Santiago, Chile

KEY WORDS

Bacterial α-amylase
Differential pressure measurement
Evaporative cooling
Fermented foods
Fixed-bed bioreactor
Fluidized-bed bioreactor
Gibberellic acid
Packed-bed bioreactor
Stirred bioreactor
Thermocouples

OUTLINE

INTRODUCTION

Solid substrate fermentation (SSF) processes may be defined as four-phase systems. The first, an air phase, is a continuous system that usually flows through a solid bed. The solid second phase consists of a water-insoluble support containing the third phase, a nutrient-rich aqueous solution. This solution is tightly absorbed within the matrix of the insoluble support; some water, however, may exist in a free state in capillaries, although this can never exceed the water-holding capacity of the matrix (1). The fourth phase is the microorganism itself, which grows in or around the support and may grow in the interparticular free space.

SSF has been used by humans for centuries. Traditional uses of SSF systems include the production of fermented foods, pigments, and koji in the Far East. Within the past decade, the production of other, higher-value microbial metabolites such as antibiotics (2,3), biopesticides (4), aromas (5), gibberellic acid (6–9), and bacterial α-amylase (10), to name a few, have been evaluated with highly promising results. Polymeric and insoluble humid agricultural products or by-products are usually employed in SSF processes. These serve a dual functions (*1*) as a source of carbon, energy, nitrogen, and other nutrients and (*2*) as the support for microorganism growth (11). A novel system has been recently developed (12) whereby a solution of nutrients is adsorbed on an inert support.

This type of process has several features that make SSF particularly attractive to industry. One of the main advantages is the use of agroindustrial by-products as raw materials, because these are not only low in cost but are usually very difficult to dispose of properly, given the organic nature of the by-products and the large volumes generated. Differential expression of certain microbial metabolites, which are not expressed in liquid medium (13), is another significant attribute of this culture system, as, for example, in the production of active *Trichoderma* propagules for bio-

logical control programs as well as in bioremediation applications.

Despite certain drawbacks of SSF, higher yields, measured in terms of product per volume, make the process economically viable. Reports of SSF metabolite concentrations 100-fold higher than those obtained from submerged fermentation (14) and biomass concentrations up to 200 g/L have been published. SSF does not need large bioreactors, and downstream processing costs might also be lower. The low energy consumption of SSF bioreactors, which require only occasional agitation of the solid bed, is a further asset from the standpoint of operating costs. Last, but not least, SSF generates less effluent, thereby reducing the environmental impact of the process as well as treatment costs.

In spite of these promising features, few applications of SSF processes are currently utilized in industry. Most SSF research has been carried out only at the laboratory level, where its interest as a fermentation system has been highlighted. Unfortunately, only a few of these processes have been evaluated at pilot or preindustrial levels because of problems inherent to solid substrate cultivation technology:

- Relatively slow rates of production, due to the additional barrier to growth imposed by the solid substrate
- Limited capacity of dissipating metabolic heat generated during cultivation
- Mass transfer problems, in particular, slow diffusion of nutrients from substrate to microorganism
- Lack of readily available, appropriate sensors for monitoring
- Limited research on SSF reactor models, control strategies, and scaling-up procedures

The latter topics are considered in this article in an effort to highlight areas that call for substantial development to bring about widespread industrial acceptance of this unique process.

The production of gibberellic acid by the filamentous fungus *Gibberella fujikuroi* is used throughout to illustrate the technical difficulties in the automation and scaling-up of SSF technology. Gibberellic acid is a biologically active compound, naturally present in plants, which acts as an hormone and regulates many plant growth and development processes. Gibberellic acid is a typical secondary metabolite, which in *G. fujikuroi* is triggered on exhaustion of assimilative nitrogen sources.

This production process presents a real challenge to those involved with automation and scaling. *G. fujikuroi* is a slow-growing filamentous fungus, generates significant amounts of metabolic heat during its exponential growth phase, and is highly susceptible to microbial contamination. Hence, successful advances in this process will boost the industrialization of SSF technology.

AUTOMATION OF SSF BIOREACTORS

Design

This section covers the design and operation of current SSF bioreactors, many of which are still the subject of research and development.

Bioreactor Types. SSF bioreactors may be of static (fixed-bed) or agitated (stirred) design. Tray and packed bioreactors are most commonly used among the former, whereas rotating drums, periodically agitated tanks, and rocking and fluidized-bed bioreactors are examples of the latter. Most SSF bioreactors operate in batch form, although water and nutrients are added periodically in many systems. Very few SSF bioreactors have been designed to operate continuously (15–17), and no industrial continuous processes have been reported.

Tray Bioreactors. Trays, one of the simplest types of SSF bioreactor, have been employed in the Far East for many years. They are readily available commercially (18) from laboratory to industrial scale. The main characteristics of this type of bioreactor follow (11):

1. A set of trays are covered with a thin layer of solid substrate.
2. Trays are kept inside a chamber, normally under controlled environment.
3. Forced air is not commonplace, although trays can be perforated allowing air to circulate smoothly through their bases.
4. In some cases the solid bed is agitated gently, either manually or automatically, using simple mechanical devices.

The main drawback of this type of bioreactor is that it is highly labor intensive, so that automation of large-scale processes is very difficult. Also, because the solid layer in the tray must be thin and the trays cannot be too close to one another, very large chambers are required to achieve industrial levels of production. Industrial applications of this type of reactor include production of mushrooms (*Agaricus bisporus*) and tempeh as well as aseptic production of soy sauce koji mold spore inoculum.

Packed-Bed Bioreactors. Packed-bed bioreactors are normally cylindrical tanks, filled with a solid substrate suspended on a perforated surface through which air is forced. The inlet air conditions (temperature, humidity, and flow rate) are generally regulated. At laboratory scale, the most common types of packed bioreactors are known as Raimbault columns. The latter are submerged in a thermoregulated bath and supplied with saturated air. In larger systems, cooling fluid is forced through double walls, coils, or plates (19) immersed in the solid bed to overcome the difficulty of removing metabolic heat. The main disadvantage of this type of bioreactors is its propensity for heterogeneous growth that hinders scaling up to industrial size.

Plastic bags filled with solid substrate provide a typical example of commercial application of packed bioreactors. The latter includes mushroom cultivation (*Agaricus bisporus, Pleurotus ostreatus*, etc.), production of biopesticides (*Trichoderma harzianum*), tempeh, and koji mold spore inoculum. In addition, biofilters (microbial flora adsorbed to a porous solid support) have recently been developed and commercialized for treating toxic gaseous and liquid streams.

Rotating Drums. Features of rotating drums used as bioreactors (11) include the following:

1. A partially inclined cylinder is filled with substrate up to a predefined level.
2. Smooth cylinder rotation mixes the solid bed; this process is usually enhanced with baffles.
3. Aeration is occasionally provided by means of tubes placed in the free space of the cylinder or in its shaft.

Rotation must be controlled to avoid damaging the microorganisms or causing substrate agglomeration or particle breakup. The control of temperature in this type of bioreactor is very difficult and becomes more difficult when increasing the scale of the reactor. Rotating drums are particularly appropriate for bioremediation of contaminated soils. Klein (16) has described several industrial applications of this technology to kerosene-contaminated soils and to degradation of mineral oil hydrocarbons, chlororganics, and polycyclic compounds. Bioreactor sizes reach 40 m in length and 3 m in diameter, with a treatment capacity of 0.5 to 1 ton/h.

Rocking Bioreactor. Although similar to the rotating drum, the rocking bioreactor cylinder rocks gently instead of performing a complete rotation. It achieves acceptable agitation and mixing without damaging either the microorganism or the substrate. This bioreactor has only been built and tested at a laboratory scale (20).

Fluidized-Bed Bioreactors. This bioreactor is a vertical cylinder inside which the solid substrate rests on a perforated plate. Air forced through the plate at high flow rates suspends the solid in the air, causing the solid–air mixture to behave like a fluid. The main advantage of this type of bioreactors is (11) the high rate of heat transfer and aeration resulting from the enormous solid–gas interface area. Additionally, gases and volatile metabolites are easily removed from the bioreactor, a very good degree of mixture of the solid is achieved, and the absence of mechanical components simplifies the automatic control of the process. These factors permit high yields, reducing the needs for large bioreactors.

Most investigations with this technology have included yeasts, either for cell biomass or ethanol production. Other microorganisms such as filamentous fungi have also been assayed. Even though these cells could be damaged by the vigorous agitation, this system causes less harm than conventional, stirred bioreactors. To our knowledge, this technology has not yet been applied at an industrial scale.

Stirred Bioreactors. Stirred bioreactors are probably the type of SSF bioreactor most widely used at the industrial scale. Here the solid bed is agitated by screws, palettes, or other mechanical devices. Agitation is intermittent and is varied during the process. These bioreactors may be either vertical or horizontal.

- Among vertical bioreactors, the most popular are the Japanese Fujiwara bioreactors (21), which are expensive. These are specifically designed and mostly used for koji production in the Far East and in the United States.
- Among the horizontal type, the best known is the INRA-Dijon bioreactor (22), which is used by some French companies in the production of enzymes (Lyven) and biopesticides (Calliope). This open, rectangular reactor is based on a design for a barley malting process. The solid bed rests on a perforated plate through which air is forced. Agitation is performed by vertical rotating screws mounted on a conveyor that moves them back and forth across the main axis of the tank. The shape and size of the screws must be designed specifically for each substrate and microorganism to minimize abrasion and breakup damage.

In both types, water, nutrients, and preservatives can be added by aspersion over the solid mass. Air can be conditioned at the inlet, and the temperature and pH of the solid may be measured at several places within the tank.

Neither the Fujiwara bioreactors nor the INRA-Dijon bioreactor is designed for a completely aseptic operation. However, results obtained using pilot stirred bioreactors with aseptic operation have been published (23–26). These required a high degree of automation and are described extensively next.

Continuous Countercurrent Bioreactor. The continuous countercurrent bioreactor consists of a ribbon-flighted screw inside a horizontal cylindrical vessel. The screw transports the solid matter backward and forward. Roussos and Pyle (17) evaluated this design for the continuous cultivation of *Aspergillus niger* at a substrate feed rate of 0.5 kg/h in the production of glucoamylase. Regulation of aeration, cooling, and water content was achieved through natural convection. It was shown that the mycelium did not suffer any damage. The general performance of this bioreactor was comparable to that obtained in laboratory-scale batch cultivation. Currently, this kind of bioreactor is used for large-scale soil bioremediation in a manner similar to rotating drums (16).

Critical Operating Conditions. The most critical operating variables in controlling an SSF bioreactor are the bed temperature, pH, and water content of the solid support, because microorganisms can only grow and produce within a relatively small range of these operating conditions. Agitation and aeration of the solid bed are further significant factors. Finally, the concentration of nutrients and metabolic gases affect biomass and metabolite production rates, although these variables are very difficult to control.

Temperature. Significant temperature gradients, to 3 °C/cm, can be observed with almost every type of static SSF bioreactor. Any microorganism generates energy (metabolic heat) during its growth. This heat must be dissipated to avert undesirable solid bed temperature conditions. In certain types of reactor, this requirement may become critical. Keeping laboratory-scale reactors, such as Raimbault columns, in a thermostatic bath overcomes the problem. Controlling the chamber temperature of tray reactors suffices in regulating the solid bed temperature; but only if the solid bed is thin enough. Larger pilot- or industrial-scale reactors require additional cooling techniques. Because of simplicity, the most commonly used cooling method is to force air through the solid bed (27). In extreme cases, very cold air or very high flow rates are required to ensure proper regulation of the solid bed temperature. In these situations, good temperature regulation

may be achieved using refrigerated jackets, cooled coils, or plates (19), although these may diminish the useful volume of the reactor. Evaporative cooling is likely one of the most effective methods; the relative humidity of the inlet air is regulated to control the rate of evaporation from the solid bed, thereby governing the amount of heat removed by the air stream. At some point in the process, however, very dry inlet air may be required, which points out the main drawback of this method; excessive drying of the solid bed requires efficient water content regulation in the control system. Ryoo et al. (20) used the technique successfully in controlling the solid temperature of a rocking laboratory-scale reactor. Fernández et al. (25) reported that evaporative cooling was capable of regulating the solid temperature in a 50-kg pilot-scale reactor. Recent simulation studies (28,29) have established the convenience of regulating the temperature of the solid bed through simultaneously manipulating the relative humidity, temperature, and flow rate of the inlet air.

Water Content of the Solid Bed. Regulation of water activity in the solid bed is crucial because water content directly defines microbial growth and metabolite production rates. Laboratory studies should thus determine the optimal water activity for the rates of both growth and production. The manner in which water activity of the solid bed varies with water content as the result of substrate consumption and subsequent biomass accumulation also needs to be assessed.

Control of water content is difficult because on-line sensors are expensive and unreliable. However, this limitation can be overcome using soft sensors based on secondary on-line measurements such as solid bed weight (22). Periodic samples could also be taken manually and analyzed for water content in the laboratory (25). There are several ways to regulate bed humidity, although the easiest is the direct addition of freshwater to the solid bed. Water may also be supplied through the air stream as fog or vapor. Sato et al. (30) were able to control water content by incorporating particles with high water retention capacity in the solid bed. The method ultimately adopted depends on the type of reactor selected and the way in which control of the solid temperature is managed.

pH. The high buffering capacity of the fermenting bed and the lack of appropriate sensors for solid substrate conditions makes pH control difficult to implement. However, pH regulation could be achieved by selecting supplementary nutrients that can be fed periodically during the cultivation; this is particularly true for the nitrogen source (22).

Agitation. Agitation in SSF reactors accomplishes several functions:

- Maintains homogeneous operating conditions
- Permits uniform distribution of water, nutrients, and preservatives supplied during cultivation
- Renews the air occluded between the particles
- Avoids particle agglomeration, crucial when a filamentous fungus is cultivated

The choice of agitation policy is determined by its main purpose. Normally, several objectives must be considered simultaneously, and it is very difficult to establish a priori how agitation could influence each objective independently. The process engineer also must ensure that the agitation neither harms the active biomass (e.g., mushroom culture) nor the substrate. Consequently, the agitation policy, at this point, is fundamentally empirical. Parameters associated with agitation are speed, sensing, and, for reactors employing intermittent agitation, frequency and duration. In fixed-bed reactors, for example, intermittent agitation could assist in the removal of heat upon the saturation of the other temperature loop control variables. Additionally, stirring could be used when air pressure loss exceeds a certain value. When using these types of reactors, however, it is essential to agitate every time water or nutrients are added.

Aeration. Because most SSF processes involve aerobic microorganisms, oxygen transfer toward the biomass is fundamental (11). Aeration also plays a key role in the removal of metabolic heat, CO_2, and volatile metabolites occluded within interparticular spaces. The following mechanisms are potentially important as relates to mass transfer efficiency between the gas phase and the microorganisms in SSF bioreactors:

- Convection from the bulk of the gas phase toward the interparticular spaces.
- Diffusion through the stagnant gas layer that surrounds the particles.
- Diffusion through the liquid layer that covers the surface of the particles. The high solute concentration of the liquid layer characteristic of SSF processes may significantly affect gaseous diffusion through this layer. To our knowledge, no research has directly targeted this important factor in SSF reactor technology.
- Diffusion inside the porous particles.
- Microbial oxygen consumption and CO_2 production.

Only the first two of these mechanisms could be influenced by either reactor design or operating conditions. Convection in particular is important in packed reactors or in those in which agitation is very slow or infrequent or aeration is very weak. Interparticular oxygen transfer is also retarded in substrates formed by a mixture of fine and thick particles, above all when filamentous microorganisms are cultivated. This situation creates an additional barrier to air circulation. The other mechanisms are fundamentally determined by the characteristics of the substrate (particle size, form, porosity, texture, etc.) and of the microorganism.

Case Study: An Aseptic Pilot Bioreactor for SSF Processes. An SSF general-purpose aseptic pilot reactor of 50-kg nominal capacity that has been built and evaluated for the production of gibberellic acid (GA_3) is now described.

General Features. The structure of the stainless steel reactor encompasses the following parts:

1. A hermetic lid with perforations (for air sampling and nutrient feeding) and an observation window;

the agitation system (motor, double arm, and blades) and the feeding system are mounted on the lid.

2. The principal body has a double-jacket system for sterilization.
3. An air chamber, located in the lower part of the reactor, is used for inlet air homogenization. Air temperature and relative humidity are measured here.
4. A rotating basket 1.15 m in diameter and 0.28 m high has a nominal capacity of 50 kg of wet solid. The basket base contains perforations of 2 mm in diameter with a separation of 3.5 mm.

During reactor operation, the following main steps can be distinguished: (*1*) sterilization, (*2*) substrate loading, (*3*) inoculation, (*4*) cultivation, (*5*) substrate unloading, and (*6*) cleaning. All these steps except cultivation are carried out manually.

Air Conditioning. As shown in Figure 1, the air is forced into the reactor by a fan (a) with a prefilter (b). The speed of the air can be regulated manually between 0.25 and 7 m/s through a purge system (c). An absolute filter (d) ensures the purity of the air, it is able to retain particles larger than 0.3 μm with a minimal efficiency of 99.99%. A heating system (e), with 6-kW electric resistance, can heat the air for reactor sterilization and process needs. In addition, the air can be cooled to 0 °C with fins (f) and a refrigeration system (g). The relative humidity is controlled by vapor addition through a solenoid valve (h) connected to a steam boiler (i) with a working pressure of 5–10 psi.

Agitation. The bed agitation system includes two arms: one contains blades that operate like a plough, and the other has curved paddles that scrape the basket base. Both arms can be manually regulated with cranks, according to process needs. The basket movement is provided by a reducing motor of 1.5 kW that is connected to the basket shaft. An inverter driver controls the rotation speed, achieving a minimum of 5 rpm with a torque of 5,500 Nm. An aseptic water seal ensures that all the inlet air passes through the solid bed. The basket is supported on roller bearings for a uniform and soft movement.

Water and Nutrient Feeding. A set of three sprinklers, mounted on the lid, are used to add water and dissolved nutrients periodically. The solution is based forced into the reactor with a peristaltic pump with a maximum flow of 10 L/min. To distribute the solution over the entire solid bed, the agitation system is turned on each time the solution is fed. The monitoring and control system of this reactor is described later.

Instrumentation and Control

Automation of SSF bioreactors is a complex problem that has yet to be resolved satisfactorily before scaling up the process. In fact, this is why much laboratory-scale research has not yet been scaled up to industrial level (1,6,8). Many of the difficulties complicating the operation and control of SSF bioreactors concern the regulation of temperature and water content in the solid bed. If the high rate of heat generation during the growth period is poorly controlled, overheating of the bed can result. The control of water content is particularly cumbersome. Part of the water is strongly bound to the solid, while the rest may remain free in the capillary areas of the material. Care must be taken to

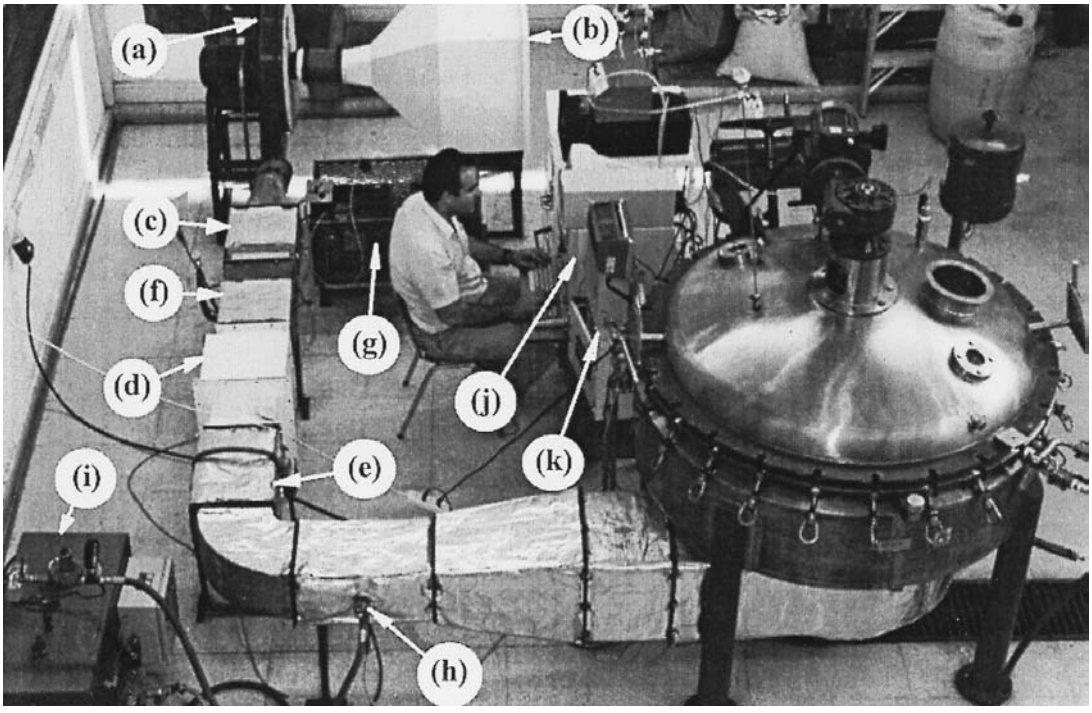


Figure 1. Aseptic pilot solid substrate fermentation (SSF) reactor with its air conditioning system: (**a**) fan, (**b**) prefilter, (**c**) purge system, (**d**) absolute filter, (**e**) heating system, (**f**) fins, (**g**) refrigeration system, (**h**) solenoid valve, (**i**) steam boiler, (**j**) control system, and (**k**) programmable logic controller (PLC).

avoid saturation of the solid matrix, which will lead to anaerobic conditions. The lack of appropriate and inexpensive sensors for monitoring water content on-line results in difficult, slow, and imprecise regulation of water content. The development of new instruments and control strategies specific to SSF processes would unquestionably simplify the development of scalable and economically competitive SSF bioreactor technology.

Instrumentation One of the main limitations of SSF as opposed to submerged fermentation (SMF) bioreactor development is the lack of specific and inexpensive instrumentation enabling proper monitoring and control of the processes. Much of the instrumentation used in SSF bioreactors is associated with measurements of the gaseous phase only. The following section discusses instrumentation commonly used in the monitoring and control of solid substrate bioreactors.

Temperature. The solid bed temperature is a primary control objective because overheating can seriously impair process performance. The temperature of the air stream is an important manipulated variable in many SSF bioreactor control strategies. Several relatively inexpensive and efficient temperature sensors are available, with thermocouples being the most widely used. When two wires of different metals are welded at two junctions and kept at different temperatures, an electric current will flow in the circuit, proportional to the temperature difference. This is the principle behind thermocouples, of which there are of several types, differing with the kind of metals used: J, iron-constantan; T, copper-constantan; K, chromel-alumel; etc. These types are appropriate for different temperature ranges and ambient conditions (reducing or oxidizing). Error limits range from ± 0.5 °C to ± 2.2 °C, depending on thermocouple type and quality. Their main advantages are fast response, shock resistance, small size, probe flexibility, heavy use capability, low cost, and suitability for both high and low temperatures. They must be calibrated periodically.

The electrical resistance of metals changes with temperature. This principle is applied in resistance temperature detectors (RTDs), which are generally more expensive than thermocouples and are suitable over a narrower temperature range; they are however more accurate and do not need periodic calibration. Platinum RTDs are most commonly used.

Air Flow Rate. Differential pressure measurement, with accuracy ranging from $\pm 0.8\%$ to $\pm 5\%$, relies on the pressure drop created as a fluid posses an obstruction (such as a concentric orifice, venturi, or pitot tube), and is therefore proportional to flow rate. The main limitation of these devices is that the energy cost from the loss of pressure can be considerable. A differential pressure transducer is needed to include this measurement in a digital control system.

Another air velocity measurement instrument commonly used is the turbine anemometer. A rotating device is placed in the path of the fluid, where its rotational speed is proportional to the fluid velocity and provides accurate flow measurement over wide ranges. Thermal anemometers, in which a thin wire immersed in the fluid is heated by an electric current, can also be used to measure air velocity. The fluid velocity is related to the heat dissipated from the wire. Full calibration is required for anemometers used for flow rate measurements because velocity depends on the anemometer's transversal position inside the duct.

Relative Air Humidity. Electrochemical sensors are most commonly used for determining relative humidity. Two widely used electronic relative humidity (RH) sensors are the Dunmore element and the Pope cell. The former employs a solid substrate coated with a LiCl solution. The hygroscopic nature of this salt causes it to take up water vapor from the surrounding atmosphere. Dunmore cells, while excellent RH sensors, are normally designed to cover only a narrow range (from 40% to 60% RH). The Pope cell is similar to the Dunmore cell with the substrate being made from sulfuric acid-treated polystyrene. The main advantage of Pope cell is its wide range of response covering from 15% to 99% RH.

Pressure and Differential Pressure. When certain asymmetric crystals are deformed along specific axes, an electrical potential created within the crystals produces an electric current in external circuits. Such devices are called piezoelectric sensors. Cultured quartz is both readily available and reasonably priced for this purpose. In addition, it presents near-perfect elasticity and stability and is temperature insensitive. Instrument response times are of the order of submicroseconds. Differential pressure transducers consist of two chambers connected internally to a rod and externally to high and low pressure measuring diaphragms. When a difference of pressure is encountered the rod is displaced in the oil-filled chambers, producing a signal.

Air Composition. Analysis of exhaust gases is very useful in the monitoring and control of SSF bioreactors (31). Information on the physiological state and respiration rate of the culture can be obtained through continuous measurement of CO_2 and O_2 concentrations in the gas outlet stream. The specific growth rate (μ), biomass concentration, and respiratory quotient (RQ) can also be resolved on-line. Other volatile compounds associated with the metabolite of interest can be measured on-line, facilitating production control and optimization. On-line measurement of CO_2 and O_2 can be performed by gas chromatography or special-purpose gas analyzers. The main advantages of the first method are that many compounds can be monitored with the same instrument and over a wide range of values. Other gas analyzers may be more precise and provide faster response. The several kinds of instruments include paramagnetic analyzers for CO_2, O_2, and ammonia, infrared instruments for CO_2, and electrochemical probes for O_2. Care must be taken to dry the sampling air beforehand and to provide a regulated flow rate entering the instrument to obtain meaningful results.

Solid Bed Water Content. Water content of the reactor bed is very difficult to measure on line. There are no instruments commercially available that can measure water content in solid substrates at levels greater than 80% with good performance and a small error. Conduction and capacitance principles are applied, to good effect, in measuring water content up to 50%. The main problem in measuring high water content in solid substrates, with

instruments based on conductance/capacitance, is that the apparent density varies unpredictably during cultivation, either because of agitation or growth of the biomass. Devices based on infrared (IR) measurements of solid samples yield very good results, but, they are expensive and their application for on-line measurement in aseptic SSF bioreactors is not well developed. A probe that can measure the water activity of a solid medium consists of a chamber which is separated from the medium by a hydrophobic membrane and contains a capacitive device to measure the relative air humidity within it (32).

Control Strategies. Proper regulation of the temperature and water content of the solid bed over several days has long been recognized as crucial to optimum performance (33). The major challenge in temperature control is to overcome the limited dissipation capacity of the metabolic heat generated during the cultivation. Temperature gradients across a bed can be as much as 20 °C, especially during the period of high heat generation. Although agitation of the bed can help in reducing the temperature gradient, in practice this heat cannot be removed by vigorous or frequent agitation, as this often has a negative impact on microbial growth and metabolism. Several methods, linked to reactor type, have been proposed and experimentally tested in attempting to control the temperature of the solid bed. Forced air and the addition of freshwater have been successfully used (27) in controlling the temperature of a packed column (35 cm height × 6 cm diameter). Temperature control has also been achieved with intermittent agitation in a laboratory rotating drum reactor (34) that homogenized the bed and also decreased its compactness. However, agitation can damage the hyphae of filamentous fungi and may alter structure morphology.

The most effective method of controlling bed temperature may be evaporative cooling. In this strategy, the rate of water evaporation from the bed can be modified by manipulating the relative humidity of the inlet air, considerably increasing the rate of heat removal. Applying this principle, good temperature control has been attained at laboratory scale in a rotary drum bioreactor (31) and a rocking bioreactor (20). A further consideration, however, is that freshwater must be supplied to avoid drying of the bed. In a packed bed, the rotation or rocking motion has to be replaced with discontinuous agitation. Our experience with 50-kg and 200-kg aseptic pilot reactors indicates that none of these strategies is sufficient to maintain proper control of the solid temperature (25); periodic agitation of the bed and manipulation of the inlet air temperature are also required. Simulation studies with SSF pilot-scale reactor models have shown that a multivariable algorithm can control solid temperature efficiently by simultaneous manipulation of relative humidity, flow rate, and inlet air temperature (28,29).

Few publications have considered the regulation of solid water content. For example, wet wood pulp particles have been placed in the solid substrate during cultivation of *Aspergillus oryzae* to provide a reservoir of free water (35). However, this practice reduces the effective bioreactor volume. In other works, the water content of the bed has been controlled by manipulating the ambient air temperature (36); the algorithm was based on energy and water balances. While this method was successfully applied in a laboratory bioreactor (20 × 16 cm), it is difficult to scale up to pilot level. Fernández et al. (25) controlled the water content to keep the water activity constant in the solid medium in a 50-kg pilot reactor. As no on-line measurement of water content was available, samples were taken manually to define the addition of freshwater based on a water balance and heuristics.

Carbon dioxide and oxygen in the outlet gas can be controlled by manipulating the inlet air flow rate (31) and pH can be controlled by spraying an acid or base solution over the solid bed (22). Nutrient concentration, can be maintained by spraying a solution directly over the solid bed (37) or into the inlet air (38).

Control Algorithms. Discrete PID and on–off control algorithms are currently used in both pilot and industrial SSF bioreactors. In light of the established process difficulties, these controllers prove difficult to tune and demand manual supervision to ensure effective operation (25).

There are remarkably few reports concerning applications of advanced techniques for controlling SSF bioreactors. The work of Ryoo et al. (20) can be mentioned as an isolated example in which an adaptive algorithm was used to control the temperature of the solid in a rocking reactor by evaporative cooling. These authors also estimated biomass concentration and total dry mass on-line using an extended Kalman filter. Optimizing the addition of nutrients to maximize the production of secondary metabolites in fed-batch SSF reactors remains a challenge, although some attempts have been made to solve the problem empirically (37,38). A short description of a number of advanced control methodologies that have been applied successfully in SMF bioreactors follows. We believe that these techniques can also be useful in the automation of SSF reactors.

Adaptive Linear Control. In essence, an adaptive controller is an on-line algorithm that estimates parameters that are, directly or indirectly, associated with the dynamics of the process. These parameters are then used to "tune" a control law on-line. Several adaptive control algorithms have been described (39,40). These controllers are particularly appropriate for processes that have variable time delays and that exhibit nonlinear dynamic behavior. Many successful applications of adaptive control in biotechnology have been published (41–44). As mentioned, application of this technique in SSF automation has been experimentally assessed by Ryoo et al. (20).

Nonlinear Control. Because real processes present nonlinear dynamics, nonlinear control techniques have attracted the attention of many researchers, both as to the theory and in the development of real-time applications. The employment of adaptive techniques in estimating uncertain or varying model parameters greatly improves the likelihood of using nonlinear control successfully in industrial processes. Several applications of this technique in biotechnology have been reported. Bastin and Dochain (45) have even developed a general methodology for designing nonlinear bioreactor controllers. Successful industrial ap-

plications have been reported by both Pomerleau and Viel (46) and Chen et al. (47).

Predictive Controllers. Paján et al. (28) compared the performance of a dynamic matrix control (DMC) predictive algorithm (48) with that of PID in the control of a simulated fed-batch SSF process and established that employing DMC greatly simplified the design and start-up of the control system. DMC possesses a flexible structure that can be adapted to the process facilities and limitations. In this particular application, the effort devoted to tuning the temperature control loop was significantly reduced and a better closed-loop response was achieved with a DMC algorithm than with a PID one. Real-time application has not yet been described, but the authors believe that this could be on optimal control strategy.

Case Study. To illustrate how the foregoing ideas can be applied to a real SSF process, a control system is now fully described for pilot SSF bioreactors (24,25).

The Process. Two aseptic packed-bed bioreactors (50 kg and 200 kg of total solid nominal capacity) using periodic agitation and forced air were built to scale-up the production of gibberellic acid by the filamentous fungus *Gibberella fujikuroi*. This process is demanding because *G. fujikuroi* is sensitive to high temperatures (above 36 °C) and to contamination by other microorganisms. Because it is filamentous, this fungus is sensitive to mechanical stresses when the solid bed is agitated, and its growth increases bed compactness, which creates a reduction in heat and mass transfer rates. In addition, bed overheating can easily occur as the result of the high rate of metabolic heat generated during the growth period. Finally, the processing time is long, demanding a robust control system.

Monitoring. Variables measured on-line are as follows (Fig. 2):

- Temperatures: in the solid bed and in the inlet and outlet air (K-type thermocouples)
- Relative humidity of the inlet air
- CO_2 concentration in the outlet air
- Air pressure loss throughout the bed
- Inlet air speed

Bed water content was measured off-line every hour.

The following remote commands are available for manipulation:

- Inlet air electric heaters (on–off; 6 kW maximum for sterilization and 3 kW maximum for operation)
- Inlet air cooler (on–off; 6.6 kW capacity)
- Bed agitation speed (regulated between 3 and 15 rpm) for both reactors
- Local screw stirrers (inverted drive) for 200-kg reactor only
- Water addition flow rate (on–off; peristaltic pump, 600 rpm)
- Steam addition (on–off; solenoid valve)

Finally, the inlet air flow rate can be set manually between 50 and 500 m^3/h.

A programmable logic controller (PLC) is used for data acquisition and control of primary actuators (Fig. 2). Operator and control calculation interaction are carried out via an IBM-compatible PC computer linked to the PLC. Project-specific software was developed for the PC with a graphic interface to handle the control systems in either automatic or manual mode (Fig. 3).

Control System. The primary objectives of this control system are to regulate solid bed temperature and to control bed water content according to a varying reference. The secondary objectives are the minimization of temperature gradients within the bed, as well as bed compactness.

Control of the bed temperature is based on evaporative cooling, by manipulating the relative humidity of the inlet air and maintaining its temperature at a low value. During the period of high heat generation, manual operation of inlet air flow rate, inlet air temperature, and agitation is also necessary to avoid bed overheating. The average bed temperature is fed into a digital PID control algorithm to drive the setpoint of the inlet air relative humidity. This, in turn, is controlled with an on–off algorithm with a dead band and hysteresis that commands a solenoid valve adding steam. The inlet air temperature is further controlled with another on–off algorithm with a dead band and hysteresis that manipulates the heaters or coolers, according to process needs.

The water content of the bed is controlled through the periodic addition of freshwater. The reference trajectory of the water content was computed in the laboratory based on water activity studies. A solid sample was taken each hour from the reactor for the laboratory measurement of the water content. The amount of water to be added is determined by the operator using an approximated water balance and their own experience. The bed is agitated at each addition of water.

To keep the bed as homogeneous as possible and to avoid excessive interparticular aerial growth, which would reduce porosity, a periodic agitation policy was established. The degree of homogeneity is defined by the temperature gradient inside the bed, while interparticular aerial growth is estimated by the loss in air pressure through the bed, as suggested by Auria et al. (49); this is a semiautomatic loop that employs heuristic logic. The operator may establish the agitation speed, its duration, and, as for the 200-kg reactor, also its path (left or right, up or down for individual stirrers).

Control Performance. The control strategy adopted here provides efficient operation of an aseptic pilot SSF bioreactor. When the reactor was run manually, it required the constant attention of at least two operators. The current control system, however, requires only one operator, who has relatively little interaction with the process; the system even allows bioreactor operation without direct supervision at certain times. Moreover, the control of bed temperature and water content is more precise than was possible without automation. Although the inlet air conditioning control system exhibits limitations, such as oscillations and offsets (25), these do not influence bed temperature control. The latter loop is critical to the achievement of high biomass and product concentrations.

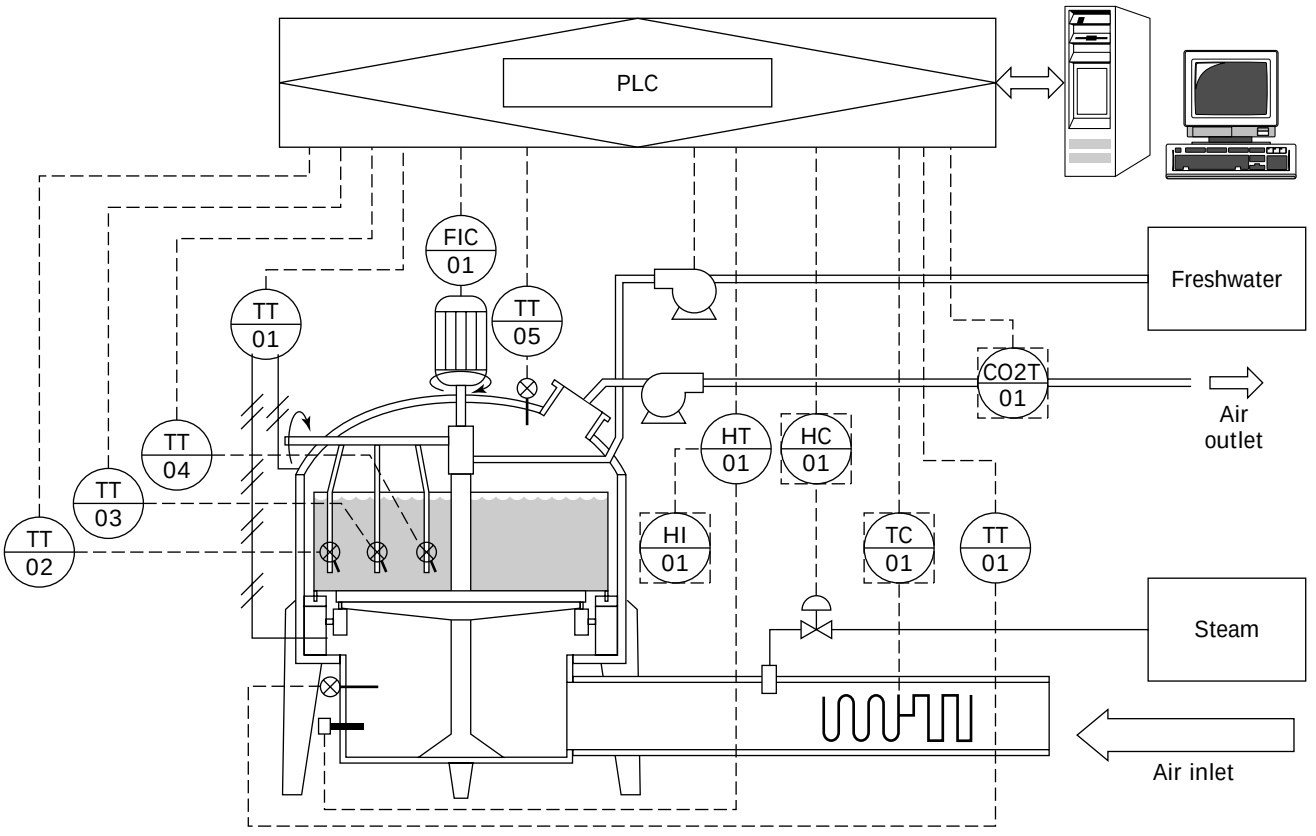


Figure 2. Process and instrumentation diagram (P&ID) of a pilot SSF reactor.

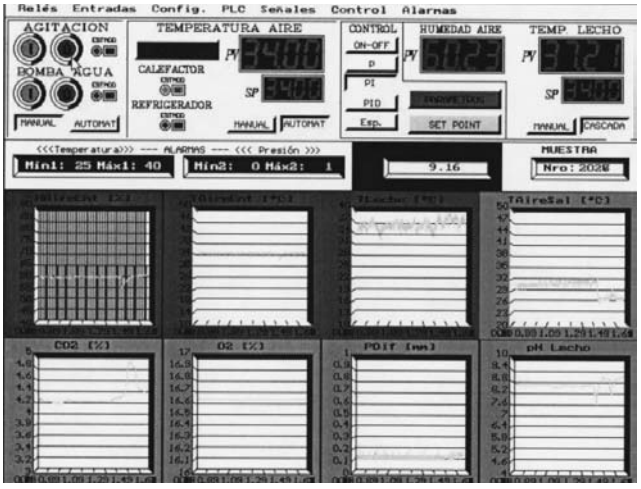


Figure 3. Example of graphics interface for SSF reactor operation.

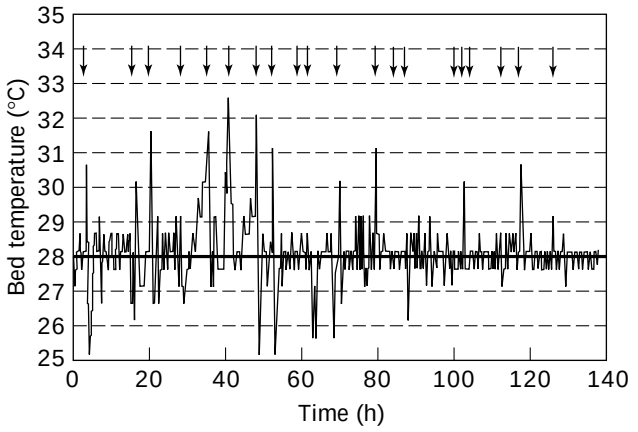


Figure 4. Bed temperature control of a SSF reactor. Agitation time is shown by arrows.

Adequate average bed temperature (Fig. 4) and water content control (Fig. 5) are achieved, although they are not fully automatic. For temperature control, evaporative cooling is insufficient during high heat generation and must be complemented with agitation and direct manual operation of the inlet air temperature and flow rate. Lacking appropriate sensors, control of bed water content is performed manually. Another significant limitation of the current system is that it is difficult to control the temperature gradient of the bed during high heat generation (not shown). However, this is undoubtedly more a limitation of the current agitation system than of the controller. It is also noteworthy that the control strategy was successfully scaled up from a 50-kg reactor to a 200-kg design with only minor adjustments.

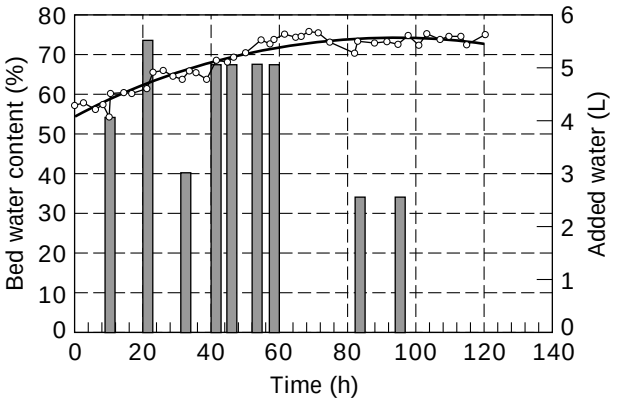


Figure 5. Water content control of a SSF reactor. *Bars* indicate addition of freshwater.

Future Developments. Before scaling the bioreactor up to industrial capacities (2 or 3 tons), a multivariable model-based control algorithm that has already been tested using simulations (29) will be installed in the 200-kg bioreactor to achieve control of the bed temperature. In this strategy, the inlet air temperature, flow rate, and relative humidity will be manipulated automatically considering prespecified priorities.

Research is under way to develop on-line estimators of water content, biomass, and product. Thus, through the addition of freshwater and nutrient solutions, these variables will be controlled automatically in the near future.

Modeling. The first step in modeling is to clearly establish the purpose of the model (50,51). This section mainly covers physical models that may be used in scaling-up, optimization, and control of solid substrate cultivation of filamentous microorganisms; black-box (fuzzy, neural networks, etc.) and microscopic models are purposely excluded. Because industrial SSF processes are carried out in batch or semibatch mode, only dynamic models are discussed.

To simplify matters, it is convenient to divide the modelling of solid state bioreactors into three stages. First, it is necessary to describe *biomass growth* and its relationship with respiratory activity (CO_2 and O_2 evolution). The effect of culture conditions such as temperature, water content, pH, CO_2, O_2, and limiting nutrients on biomass growth must be defined. Then, the *production* and *degradation* rates of the metabolite of interest are studied. Finally, *mass* and *heat transfer* phenomena must be included using appropriate energy and mass balances. This step is critical when it comes to scaling-up from laboratory to pilot and on to industrial reactors.

Biomass Growth. Almost all published work on solid substrate fermentation models considers logistic growth kinetics (27,52,53), which assumes that steric effect is the only limitation for microbial growth:

$$\frac{dx}{dt} = \mu_{\max} \cdot \left(1 - \frac{x}{x_{\max}}\right) \cdot x$$

This model may be useful in describing the evolution of the biomass during the growth phase, although it cannot be applied directly during stationary or lytic phases. If the aim of the fermentation is to produce a secondary metabolite, substrate limitation must be included in the kinetic growth model. The most widely used expression for this is the Monod equation, in which the growth rate decreases steadily after the limiting substrate begins to be depleted. This equation may be derived theoretically assuming that the growth rate is controlled by a single enzymatic reaction. Although real systems are much more complex, the Monod equation fits experimental results well. Many other kinetic equations also can be used to fit experimental results, for example, the Contois, Tessier, and Moser (54). Each of these equations considers a single limiting substrate; multiple substrate models are discussed by Blanch (55). Few SSF models consider growth kinetics other than the logistic equation. For example, Mitchell et al. (56) modeled the growth of *Rhizopus oligosporus* on a solid substrate using the Monod equation with glucose as the limiting substrate, and Gutiérrez-Rojas et al. (57) presented a growing kinetic expression that simultaneously included a logistic model and a substrate-dependent model with substrate inhibition.

Temperature and water content strongly influence biomass growth. Few models consider the effect of water content on the growth kinetic in SSF processes. Sargantanis et al. (52) modeled the dependence of the logistic kinetic parameters on temperature and water content using polynomials. Gervais et al. (58) considered the effect of water activity on the biomass growth rate; this variable better represents the physicochemical state of the water in the fermenting solid substrate. The effect of temperature on biomass growth is modeled in several works using different mathematical expressions, such as polynomials (53,59), the Ratkowsky equation (60,61), the Esener equation (27), and the Arrhenius equation (62).

Other important aspects that have received little research attention include dependence on other environmental variables (pH, CO_2, O_2) and the description of the reduction of biomass activity during the lytic phase. Rajagopalan and Modak (59,63,64) attempted modeling the effect of O_2 concentration on biomass growth using the Monod equation. Sangsurasak and Mitchell (53) and Smits (61) included a death rate in their growth models to describe the reduction of biomass activity. Finally, almost all authors use the Pirt linear growth model (65), which relates biomass with respiration activity. This model assumes that respiration activity is proportional to the biomass growth rate, and that a little respiration activity is still present (maintenance) even after the growth has ceased.

Metabolite Production. Although many authors have presented experimental results on metabolite production in SSF processes (66–68), modeling is in its infancy (54,69).

Heat and Mass Transfer. Many of the experimental systems used in the study of fungal growth on solid substrates are small enough (shake flasks, Raimbault columns, thin trays) to ensure homogeneity of the medium; that is, biomass concentration, temperature, water content, O_2 concentration, pH, and nutrient concentration are approximately equal throughout the solid bed. Because these

systems are small, it is relatively easy to keep environmental variables (temperature, water content, O_2) under control. In larger bioreactors (pilot and industrial scale), the situation is very different; in these, it is extremely difficult to regulate the environmental variables and almost impossible to ensure homogeneity of the solid medium.

Few researchers have developed models aimed at predicting or explaining temporal and spatial dependence of temperature, water content, and O_2. These models can be classified as lumped parameter or distributed parameter and as pseudo-homogeneous or heterogeneous. Lumped parameter models are those that do not consider space distribution. In this scheme, process variables depend only on time and are mathematically described by ordinary differential equations. When space distribution is considered, process variables are described by partial differential equations. Lumped parameter models are used to describe very small bioreactors or stirred bioreactors. Larger static bioreactors should be modeled as a distributed parameter system. Models not explicitly considering distribution inside the solid particles are termed pseudo-homogeneous (70), while models that describe gradients inside the solid particles are called heterogeneous.

An early complete model describing a SSF process that was developed by Raimbault (71) purports that the growth of the filamentous fungus *Aspergillus niger* on amylaceous substrates is limited by the lack of surface and free water. These limitations are specific to the fermentation on solid substrates and hence are not applicable to liquid fermentation. Thus, a physiological base is established to explain the logistic model of growth. Raimbault obtained a global model, based on these principles, to predict the evolution of relevant process variables, such as sugars, biomass, CO_2, O_2, water, and metabolic heat. The determination of the maintenance coefficients for sugars, O_2, and CO_2, and the establishment of a global stoichiometric equation of growth, are significant contributions of this work. This author used small laboratory columns (known as Raimbault columns) specifically designed to minimize temperature and humidity gradients.

Another significant contribution to SSF modeling has been presented by Ryoo (72). The main feature of this lumped parameter model is the incorporation of the latent heat of evaporation in the energy balance. This model also considers the degradation kinetics of the substrate (extruded corn) by the filamentous fungus *Rhizopus oligosporus*. Like Raimbault, Ryoo used a logistic model to describe the fungal growth. The model also includes the evolution of CO_2, water content, and temperature. The growth kinetic parameters are calculated in terms of the water content and temperature of the solid bed. This work was carried out in a 20- $\times$ 25-cm rocking reactor.

More recently, Gutiérrez-Rojas et al. (57) proposed a lumped parameter model for the growth of *Aspergillus niger* on inert support in a Raimbault column. The model describes the evolution of biomass, sugar, water, CO_2, O_2, and temperature. It incorporates several new features:

- Growth kinetic with simultaneous steric and substrate limitations
- A temperature-dependent μ_{max}
- An energy balance that includes conduction, convection, and metabolic heat

Using this model, which did not consider evaporative cooling, the authors concluded that the main growth rate-limiting factor was sugar depletion; water and void spaces remained largely available.

Saucedo-Castañeda et al. (27) applied a distributed pseudo-homogeneous model in describing the evolution of the radial temperature gradient of a cylindrical static packed-bed SSF bioreactor with forced air in laminar regime. The energy balance included metabolic heat generation and heat transfer by two mechanisms, conduction inside the solid bed and convection through the reactor wall. The model was calibrated and validated with experimental data obtained in a reactor of 6 cm radius $\times$ 35 cm height during the growth of *Aspergillus niger* on cassava meal. In this example, conduction presented the largest heat transfer resistance.

A similar model was developed by Rajagopalan and Modak (59) to describe a 6-cm-deep tray reactor without forced air. The model was able to reproduce the trend of experimental observations, such as maximum temperatures at the center of the bed, decrease of oxygen concentration, and increase of carbon dioxide from the top to the bottom of the tray. However, it failed to quantitatively fit the experimental data. The authors argued that this is because the model neglects some observed phenomena, for example, evaporative cooling, loss of moisture, and solid shrinkage. Smits (61) included evaporative cooling in a tray bioreactor model, concluding that evaporation was not significant because the differences in water vapor concentrations between the tray and the environment were small.

A heterogeneous form of their original model was developed by Rajagopalan and Modak (64); that is, oxygen diffusion inside a solid particle is also considered. Assuming a zero-order oxygen consumption rate, they found an upper limit for cell density. In addition, the model predicted that, despite aerobic conditions in the gas phase, anaerobic conditions may exist within the biofilm. The same authors (63) also developed a model for fungal growth over an spherical particle of wheat bran, where the growth kinetics included simultaneous O_2 and glucose limitations, using the Monod equation in both cases. Simulations with this model showed that for a typical tray fermentation, oxygen depletion inside the biofilm is the primary cause of growth rate decline. However, as fermentation proceeds, the depletion of glucose causes growth to halt completely. In contrast, for a highly aerated solid bed of spherical particles, J. Thibault, K. Pouliot, E. Agosin, and R. Pérez-Correa (unpublished), using an oxygen transfer model calibrated with published data, concluded that in this case diffusion of oxygen through the biofilm is the main limiting step for oxygen transfer. Therefore, a direct consequence of this result is that for large bioreactors, optimizing operating conditions (air flow rate, oxygen concentration in the gas phase, bed agitation) is not enough to maximize oxygen availability to the microorganism because this also significantly depends on the thickness and the diffusion coefficient of the biofilm. In our opinion, data resulting from this exciting research subject can

significantly improve the overall performance of SSF technology.

Case Study: Modeling of Fungal Growth and Metabolite Production by *Gibberella fujikuroi*. A homogeneous, lumped parameter model that described the evolution of biomass of *G. fujikuroi* growing on an inert support in a Raimbault column was developed. The lag phase was not incorporated either in the model or with the data used for calibration. The main assumptions are as follows:

- Oxygen transfer resistance is negligible.
- Available nitrogen is the limiting substrate.
- Carbon source is not limiting.
- Temperature and water activity are constant throughout cultivation.

The major novelties of this model are the inclusion of the lytic phase of growth and the production of a secondary metabolite in SSF. A more detailed description of the model is under preparation.

Fungal Growth. The Contois equation is used to model the growth kinetics. Total biomass is determined by the glucosamine method (73). To be consistent with the respiration activity, the model also considers active and inactive biomass. Moreover, it was assumed that biomass lysis occurs.

The active biomass concentration is described by

$$\frac{dX}{dt} = \mu \cdot X - k_d \cdot X$$

where μ is the specific growth rate: and k_d is the death rate. The measured biomass, X_m, includes both active and inactive biomass:

$$\frac{dX_m}{dt} = \mu \cdot X - k_m$$

k_m represents a constant lysis rate. The consumption rate of the nitrogen source, urea, is given by

$$\frac{dU}{dt} = -k \cdot U$$

where k is the conversion rate from urea to available nitrogen for the microorganism, N_l, which can be directly metabolized into active biomass. The following expression represents its evolution during cultivation:

$$\frac{dN_I}{dt} = 0.47 \cdot k \cdot U - \mu \cdot \frac{X}{Y_{X/N_l}}$$

where 0.47 corresponds to the nitrogen content of the urea.

The specific growth rate is then

$$\mu = \mu_m \cdot \left(\frac{N_I}{k_N \cdot X + N_I}\right)$$

The evolution of the carbon source, soluble starch, is given by

$$\frac{dG}{dt} = -\mu \cdot \frac{X}{Y_{X/G}} - m_G \cdot X$$

where m_G is the maintenance coefficient. Finally, the respiration activity is described by the accumulated CO_2 using a standard linear growth model:

$$\frac{dCO_2}{dt} = \mu \cdot \frac{X}{Y_{X/CO_2}} + m_{CO_2} \cdot X$$

The following figures compare experimental data, obtained at 31 °C, with simulations using this model (the parameters are given in Table 1).

Figure 6 shows the expected three growth phases: exponential, deceleration, and lysis. The model reproduces accurately the first and second phases, as well as the trend of the lytic phase. The data of the latter show high dispersion, which makes it difficult to fit it adequately. Urea concentration disappeared before 20 h of cultivation, much earlier than the maximum biomass concentration was reached (Fig. 7). This result supports the assumption that urea is not directly used by the microorganism for growth.

Table 1. Model Parameters

Parameter	31 °C
μ_m (h^{-1})	0.20
k_N (g N_l/g X)	0.55
k_d (h^{-1})	0.022
k_m (h^{-1})	5.0×10^{-5}
$Y_{X/G}$ (g X/g G)	0.35
m_G (g G/g X h)	0.028
$Y_{X/N\,l}$ (g X/g N_l)	4.4
m_{CO_2} (g CO_2/g X h)	0.12
Y_{X/CO_2} (g X/g CO_2)	0.19
β_m (h^{-1})	1.22×10^{-3}
γ (g/g i.s.)	0
K_l (g i.s./g)	996
k_p (h^{-1})	1.0×10^{-6}
k (h^{-1})	0.1330

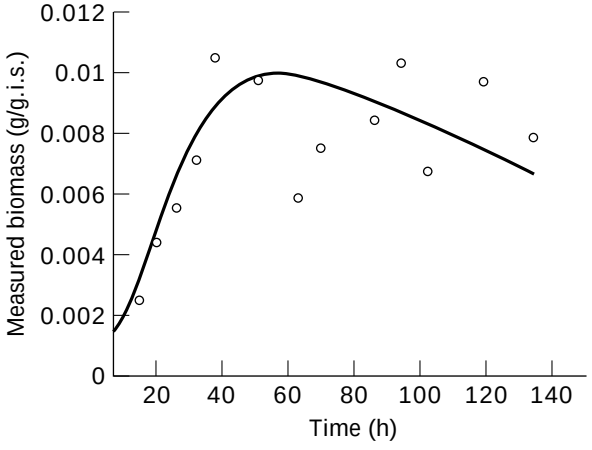


Figure 6. Biomass: measured (*circles*) and simulated (*line*).

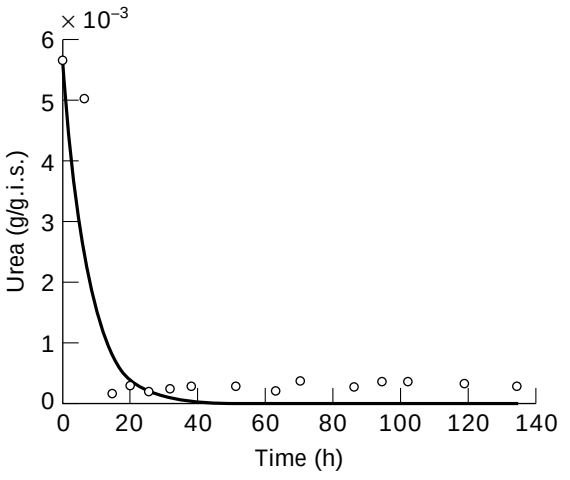


Figure 7. Urea concentration: measured (*circles*) and simulated (*line*).

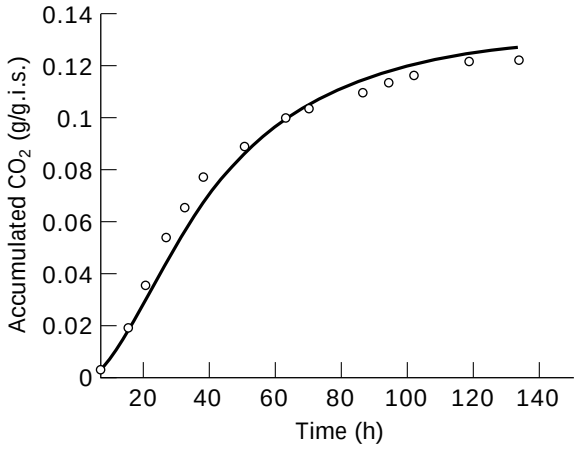


Figure 8. Accumulated CO_2: measured (*circles*) and simulated (*line*).

The model is able to reproduce well the observed decline in biomass respiration activity (Fig. 8). The carbon source is not limiting for growth as seen in Figure 9, because enough starch remains even after growth has stopped.

These figures show that the model is able to describe the complex interaction between active biomass concentration, nutrients, and respiration activity. This model can serve several purposes, such as establishing a fed-batch policy to maximize biomass concentration or developing on-line biomass estimators (74).

Production. The product of interest is a growth hormone called gibberellic acid (GA_3), which is a secondary metabolite; its production starts after nitrogen depletion. GA_3 production in SSF is difficult because it decomposes in water (75), so that if cultivation continues too long, most of the product will be lost. The following model can reproduce this observed behavior:

$$\frac{dP}{dt} = \beta \cdot X - k_P \cdot P$$

where k_p is the degradation rate and $\boldsymbol{\beta}$ is the specific production rates, which depends on the available nitrogen:

$$\beta = \frac{N_I \cdot \beta_m}{\gamma + N_I + K_l \cdot N_I^2}$$

Figure 10 shows the evolution of product concentration. The model fits the experimental data well, so it can be used to developed an on-line GA_3 estimator or a fed-batch policy to maximize product concentration at a given time.

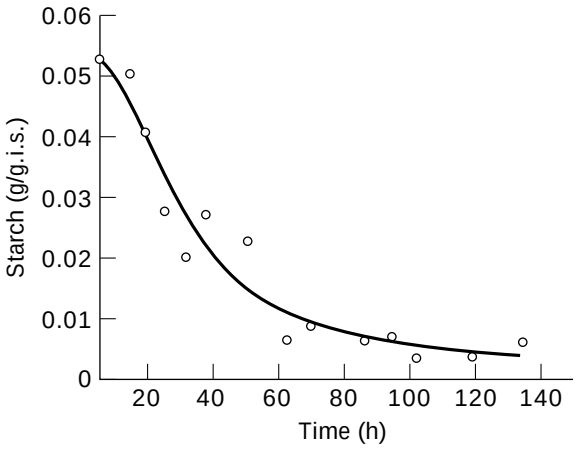


Figure 9. Carbon source evolution: measured (*circles*) and simulated (*line*).

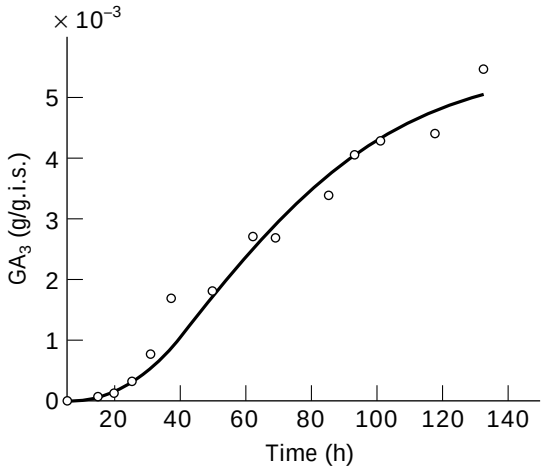


Figure 10. GA_3 concentration: measured (*circles*) and simulated (*line*).

INDUSTRIAL IMPLEMENTATION

To develop a complete SSF process that can be scaled up to industrial level, production must be integrated with the following two critical stages.

Upstream Processing

In SSF processes, supports are normally made from either agricultural products (soybean, rice, wheat, corn) or by-products (wheat bran, sugarcane bagasse, cassava bagasse). These materials are generally supplemented with easily metabolized carbon and energy sources (starch, maltodextrin, molasses, etc.) and other nutrients (urea, calcium, phosphates, etc.). The preprocessing of these substrates, mainly thermal treatment, has a double objective:

- Sterilize the support
- Change substrate physicochemical properties (starch gelatinization, increase porosity, etc.) to favor the support's structural strength and also the degradability and accessibility of its polymeric components (polysaccharides and proteins)

Minimizing the endogenous microbial load present in these substrates to levels compatible with the growth of the microorganism of interest is fundamental to attaining proper cultivation. It is easy to understand that this operation shares a significant part of the production costs when considering the high volumes of solids that are handled in industrial reactors. Lonsane et al. (76) have proposed the use of batch sterilization for 3–6 h at 121 °C, emphasizing that each solid particle must be subjected to 121 °C for at least 60 min. This, however, does not seem feasible for industrial-scale processes.

Traditionally, this stage is performed through prolonged cooking of the support, that is, rice and soybean in the particular case of koji production. For the latter, toasting has been recently been introduced as an alternative to cooking, with encouraging results in terms of productivity (Tomomori, personal communication, 1992). In 1995 Chou and Rwan (77) evaluated the extrusion cooking of rice for koji production. Here, *Aspergillus oryzae* was cultivated on rice extruded at different temperatures and water content levels, and achieved enzymatic yields superior to those obtained with traditional cooking techniques. Extrusion is a technology widely applied in the formulation of foods for human and animal consumption. However, its use in the preprocessing of substrates for SSF has been limited (78,79). Extrusion is essentially a fast process at high temperature in which proteic nutrients, rich in starch, are subjected to considerable heat and mechanical stresses, which modify the structure of the original material (80). Extrusion provides cooked substrates with the following characteristics:

- Elimination or at least significant reduction of the initial microbial load
- Liquefaction and gelatinization of the starch
- Denaturation of proteins and inactivation of antinutritional factors
- Solid form and porosity, which improve microbial growth and resistance to mechanical stresses
- Homogeneous substrate granulometry, increasing interparticular gas mass transfer

Special substrates made from products rich in fiber, such as sugar beet (81) or wheat bran (82), and amylaceous products, have also been developed. Recent results obtained in our laboratory in the production of gibberellic acid by *G. fujikuroi* have demonstrated the benefits of using this treatment. A mixture of wheat bran and starch (80%: 20%) extruded at 160 °C and at 22% humidity achieved a volumetric productivity 25% higher than that obtained when using nonextruded material.

In conclusion, this technique appears particularly appropriate for industrial-level SSF preprocessing. An aseptic support is produced at shorter processing times, with higher productivities and reduced costs. However, the process requires considerable investment. Other promising techniques that are currently being evaluated are sterilization by microwave (83), high voltages, and magnetic fields.

Downstream Processing

Downstream processing is one of the most neglected aspects of SSF processes in spite of its importance from both technical and economical standpoints. Traditional SSF-derived products generally encompass the whole fermented mass (koji, ang-kak, tempeh, etc.), which is, at most, dried (tane koji, for example) to improve keeping quality. However, there is a growing number of products that now require purification and formulation before being marketed. It is worth noting that the usual operations tied to these stages, although well developed for liquid cultivation, cannot be directly applied to SSF. Indeed, it is necessary first to leach the product from the fermented solids.

Following solid substrate cultivation, the fermented mass consists of unused substrate, biomass, spores (when formed), the product of interest, and a series of other cometabolites (76,84). The product in question must be extracted from this mass with an appropriate solvent. Efficiency of leaching depends on a series of factors:

- Effectiveness of the selected solvent
- Solid-solvent proportion
- Product and solvent diffusion
- Solvent retention by the solids
- Time, degree of mixing, temperature, and pH
- Techniques used to separate the leached solution

Among the available separation techniques, the most commonly used are percolation, pulsed extraction in plug flow, multistage countercurrent extraction, hydraulic press, and supercritical extraction. It is imperative to the economic viability of the SSF process that the selected extraction technique maintains the high concentration of the product to the greatest extent possible. Indeed, if the leaching process resulted in the dilution of the product, even at concentrations similar to those obtained by submerged fermentation, one of the main advantages of SSF would be lost. As a result, while common, percolation is not an appropriate technique because the high dilution of the resulting product (1:10 on the average) demands an additional concentration stage (84).

Plug flow pulsed extraction is a variation of percolation (85) in which the solvent is added by pulses through as-

persion of the substrate. With such a flow, the solvent completely wets the solid, achieving optimum solid to solvent proportions (1:2 to 1:3). A further advantage is that this is a simple, low-cost technique that requires minimum supervision. On the negative side, the extraction period is long, which may result in product modification or degradation.

Countercurrent extraction (86) allows adequate product concentration (1:1 solid:solvent proportion). In this system, the solids and the solvent move in opposite direction. Initially, the fresh solvent makes contact with the leached solids then with partially leached solids, contacting increasing concentrations of product. Finally, the enriched solvent contacts fresh solid. The required number of contact stages varies between 2 and 5 (86,87). In spite of the high concentrations achieved, the method presents some limitations in that the operation needs constant attention from trained operators. In addition, the extract obtained can be highly viscous, resulting from the extraction of exopolysaccharides. Extraction by hydraulic press is an efficient operation, although it requires high pressures and a second press to improve efficiency. Moreover, the operation is slow and tedious.

Product extraction with supercritical fluids has been studied for solid-state fermentation; in particular with CO_2 (88). Solvents have greater extraction ability at near-critical temperatures and pressures. This property is being used increasingly in the chemical industry as an alternative to conventional distillation and to other extraction processes. Here, the separation stages are simplified because the products are volatilized or precipitated from the supercritical fluid. However, the capital and operational costs of this process are high, restricting its applicability to high-value products. Once the product is leached from the solids, the extract is similar to a cultivation broth obtained following submerged fermentation. Conventional bioseparation (concentration, purification) and product formulation methods normally used for liquid cultivation processes can then be applied.

Case Study: Production of Gibberellins by Solid Substrate Cultivation of *Gibberella fujikuroi*

All commercially available gibberellins, which are typical secondary metabolites, are produced by liquid cultivation (LC) of the ascomycetous fungus *G. fujikuroi* (88). On exhaustion of the nitrogen source, exponential fungal growth ceases. At this point, secondary metabolism is triggered, with the concomitant biosynthesis of gibberellins. Commercially, the most important resultant product is gibberellic acid (GA_3). Several authors have reported on gibberellic acid production in solid substrate fermentation with higher yields than those obtained from liquid fermentation (6,9,37,89). A flowsheet of a pilot plant (Fig. 11) used for the production of gibberellic acid is presented and discussed.

This process, which has been scaled up from laboratory (12 g) to prepilot level (50 kg) and, recently, on up to a 200-kg reactor (Fig. 12), is carried out in three stages. The first, *upstream*, includes fungal propagation and substrate

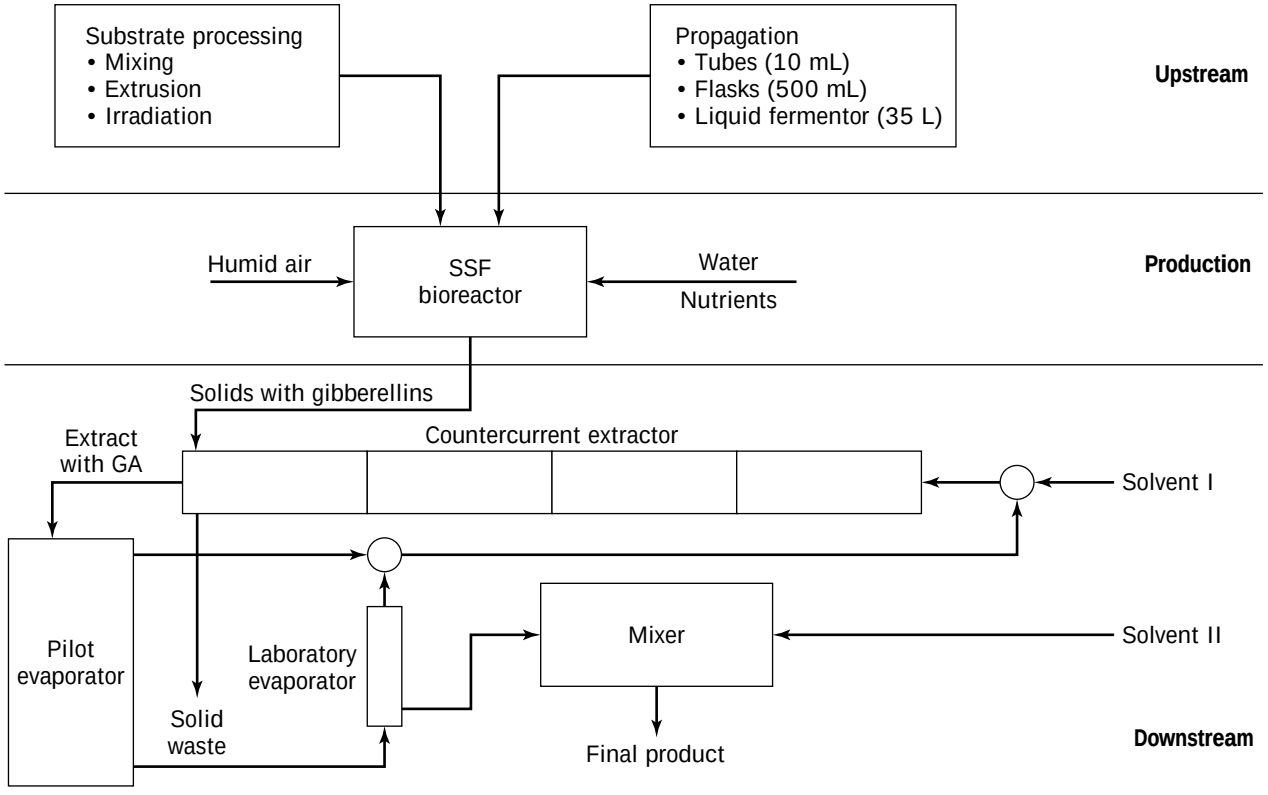


Figure 11. Flowsheet of a SSF pilot plant for gibberellin production.

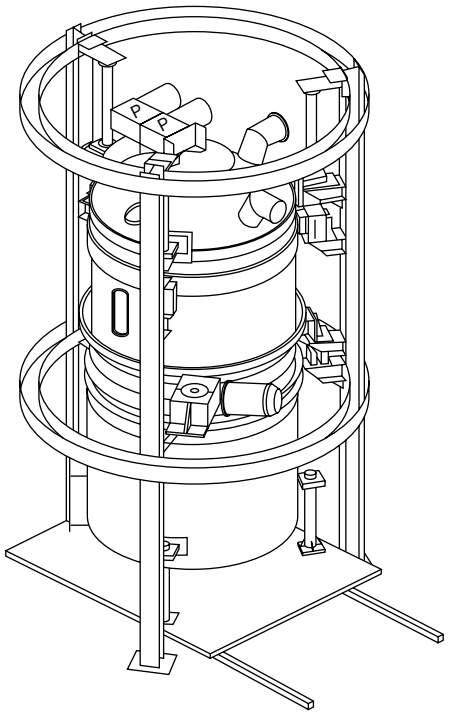


Figure 12. A 200-kg SSF pilot bioreactor.

preparation; second, the *production phase*; and finally, *downstream*, which consists of extraction and formulation. Substrate preparation involves mixing wheat bran with starch, extrusion of the mixture at high temperatures, and gamma ray irradiation. Mycelial propagation of the fungus was achieved in 72 h, from 10 mL up to 35 L. Production is carried out in batch form following a defined sequence: reactor sterilization (3 h), substrate loading (1 h), and inoculation (0.5 h), water content adjustment (0.5 h), cultivation (5 days), unloading (2 h), and bioreactor cleaning (4 h).

Extraction is done in batch form in four countercurrent stages. A primary concentration is then performed in a 15-L batch evaporator followed by final concentration in a laboratory evaporator. The last step is the formulation of the product according to market needs. The results obtained are easily reproducible (Fig. 13), and the economic viability and technical feasibility of the whole process have been demonstrated.

On field evaluation in production-scale vineyards, the product obtained (FITOPLUS) via this process exhibited considerable benefits over commercially available gibberellic acid products. Thompson seedless grapes of superior quality (berry size, reduction in berry loss, berry stalk color, etc.) were consistently obtained when treated with the SSF-derived product. This product has been placed under patent protection for agricultural applications (90).

NOMENCLATURE

G	Starch concentration (g/g i.s.)
k	Urea transformation constant (L/h)

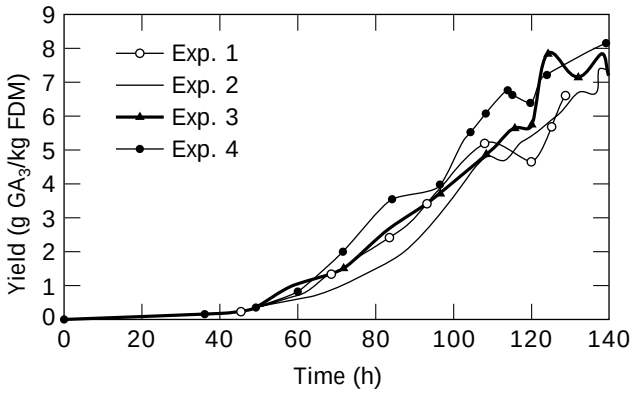


Figure 13. Gibberellic acid production in a SSF pilot bioreactor.

k_d	Biomass death constant (L/h)
K_i	Inhibition constant (g i.s./g)
k_m	Lysis constant (L/h)
k_N	Contois model parameter (g N_I/g X)
k_p	GA_3 degradation constant (L/h)
m_G	Starch maintenance constant (g G/g X h)
m_{CO_2}	CO_2 maintenance constant (g CO_2/g X h)
N_1	Available nitrogen (g/g i.s.)
P	Gibberellic acid (GA_3) (g/g i.s.)
U	Urea (g/g i.s.)
X	Active biomass (g/g. i.s.)
X_m	Measured biomass (g/g. i.s.)
$Y_{X/G}$	Biomass/starch yield coefficient (g X/g G)
$Y_{X/NI}$	Biomass/nitrogen yield coefficient (g X/g N_I)
Y_{X/CO_2}	Biomass/CO_2 yield coefficient (g X/g CO_2)
β	Specific production rate (L/h)
β_m	Maximum specific production rate (L/h)
γ	Parameter in the production rate (g/g i.s.)
μ	Biomass specific growth rate (L/h)
μ_m	Maximum specific growth rate (L/h)

ACKNOWLEDGMENTS

Financial support through the projects Fondecyt 1960360 and FONDEF 2-50, and from DICTUC S.A. and BASF-Chile S.A., is greatly appreciated. The work accomplished would not have been possible without the fruitful and continual discussions with our colleagues Iván Solar, Luciano Chiang, and Héctor Jorquera (Pontificia Universidad Católica de Chile), Gonzalo Acuña (Universidad de Santiago de Chile), Georges Corrieu and Erik Latrille (INRA, France), Paul Molin (ENSBANA, France), Maurice Raimbault (ORSTOM, France), and Jules Thibault (Université de Laval, Canada). Our dedicated postgraduate students undertook most of the experimental and computational work and also contributed several original ideas; our thanks go to Andrés Corona, Armin Ebner, Mario Fernández, Claudio Gelmi, Mauricio González, Harold Paján, Marcial Peña y Lillo, Ariel Rossenblitt, and Daniel Volpe. The skillful technical assistance of Andrea Belancic, Lenka Torres, and Cristián Valenzuela is also acknowledged. The English translation was carefully revised by Alex Crawford.

BIBLIOGRAPHY

1. R.E. Mudgett, in A. Demain and N. Solomon eds., *Manual of Industrial Microbiology and Biotechnology*, American Soc. for Microbiology, Washington, D.C., 1986, pp. 66–83.
2. A. Barrios-González, A.G. Tomasini, G. Viniegra-González, and J. Lopez, *Biotechnol. Lett.* **10**, 793–798 (1988).
3. M. Jermini and A.L. Demain, *Experientia (Basel)* **45**, 1061–1065 (1989).
4. S. Roussos, A. Olmos, M. Raimbault, G. Saucedo-Castañeda, and B.K. Lonsane, *Biotechnol. Tech.* **5**, 415–420 (1991).
5. P. Christen, E. Villegas, and S. Revah, *Biotechnol. Lett.* **16**, 1183–1188 (1994).
6. P.K. Kumar and B.K. Lonsane, *Biotechnol. Bioeng.* **30**, 267–271 (1987).
7. P.K. Kumar and B.K. Lonsane, *Appl. Microbiol. Biotechnol.* **34**, 145–148 (1990).
8. X. Qian, J. du Preez, and S. Kilian, *World J. Microbiol. Biotechnol.* **10**, 93–99 (1994).
9. E. Agosin, M. Maureira, V. Biffani, and F. Pérez, in S. Roussos, B.K. Lonsane, M. Raimbault, and G. Viniegra-González eds., *Advances in Solid State Fermentation*, Kluwer, Paris, 1997, pp. 355–366.
10. M.V. Ramesh and B.K. Lonsane, *Appl. Microbiol. Biotechnol.* **33**, 501–505 (1990).
11. D.A. Mitchell and B.K. Lonsane, in H.W. Doelle, D.A. Mitchell, and C.E. Rolz eds., *Solid Substrate Cultivation*, Elsevier, London, 1992, pp. 1–13.
12. R. Auria, S. Hernández, M. Raimbault, and S. Revah, *Biotechnol. Tech.* **4**, 391–396 (1990).
13. G. Muñoz, E. Agosin, M. Cotoras, R. San Martin, and D. Volpe, *FEMS Microbiol. Lett.* **125**, 63–70 (1995).
14. E. Agosin and J.M. Aguilera, in G.E. Harman and C.P. Kubicek eds., *Trichoderma and Gliocladium Vol. 2*, Taylor and Francis, London, 1998, pp. 205–228.
15. W.R. Gibbons, C.A. Westy, and T.L. Dobbs, *Biotechnol. Bioeng.* **26**, 1098–1107 (1984).
16. J. Klein, in D.L. Wise ed., *Global Environmental Biotechnology*, Kluwer, Dordrecht, 1997, pp. 489–504.
17. S. Roussos and D.L. Pyle, in M. Raimbault, R. Soccol, and G. Chuzel eds., *International Training Course on Solid State Fermentation*, ORSTOM, Montpellier, 1998, pp. 21–35.
18. M.G. Byndoor, N.G. Karanth, and G.V. Rao, in S. Roussos, B.K. Lonsane, M. Raimbault, and G. Viniegra-González eds., *Advances in Solid State Fermentation*, Kluwer, Paris, 1997, pp. 113–119.
19. A.A. De Araujo, C. Lepilleur, S. Delcourt, P. Colavitti, and S. Roussos, in S. Roussos, B.K. Lonsane, M. Raimbault, and G. Viniegra-González eds., *Advances in Solid State Fermentation*, Kluwer, Paris, 1997, pp. 93–111.
20. D. Ryoo, V.G. Murphy, M.N. Karim, and R.P. Tengerdy, *Biotechnol. Tech.* **5**, 19–24 (1991).
21. E. Cannel and M. Moo-Young, *Process Biochem.* **15**, 24–28 (1980).
22. A. Durand and D. Chereau, *Biotechnol. Bioeng.* **31**, 476–486 (1988).
23. Y. Chameliec, R. Renaud, J. Maratray, S. Almanza, M. Diez, and A. Durand, *Biotechnol. Tech.* **8**, 245–248 (1994).
24. M. Fernández, J. Ananias, I. Solar, R. Pérez, L. Chang, and E. Agosin, in S. Roussos, B.K. Lonsane, M. Raimbault, and G. Viniegra-González eds., *Advances in Solid State Fermentation*, Kluwer, Paris, 1997, pp. 155–168.
25. M. Fernández, R. Pérez-Correa, I. Solar, and E. Agosin, *Bioprocess Eng.* **16**, 1–4 (1996).
26. E. Agosin, R. Pérez-Correa, M. Fernández, I. Solar, and L. Chiang, in D.L. Wise ed., *Global Environmental Biotechnology*, Kluwer, Dordrecht, 1997, pp. 233–243.
27. G. Saucedo-Castañeda, M. Gutierrez-Rojas, G. Bacquet, M. Raimbault, and G. Viniegra-González, *Biotechnol. Bioeng.* **35**, 802–808 (1990).
28. H. Paján, R. Pérez-Correa, I. Solar, and E. Agosin, in D.L. Wise ed., *Global Environmental Biotechnology*, Kluwer, Dordrecht, 1997, pp. 221–232.
29. M. Fernández, R. Pérez-Correa, and M. Peña y Lillo, in G. Lefranc, M. Duarte, and R. Pérez-Correa eds., *Anales del VIII Congreso Latinoaméricano de Control Automático*, ACCA, Santiago, 1998.
30. K. Sato, M. Nagatani, and S. Sato, *J. Ferment. Technol.* **60**, 607–610 (1982).
31. G. Saucedo-Castañeda, M. Trejo-Hernández, B.K. Lonsane, J.M. Navarro, S. Roussos, D. Dufour, and M. Raimbault, *Process Biochem.* **29**, 13–24 (1994).
32. A. Durand, R. Renaud, J. Maratry, and S. Almanza, in S. Roussos, B.K. Lonsane, M. Raimbault, and G. Viniegra-González eds., *Advances in Solid State Fermentation*, Kluwer, Paris, 1997, pp. 71–111.
33. M.V. Ramana, N.G. Karanth, and K.S.M.S. Raghava Rao, *Adv. Appl. Microbiol.* **39**, 214–265 (1993).
34. J. de Reu, M.H. Zwietering, F.M. Rombouts, and M.J.R. Nout, *Appl. Microbiol. Biotechnol.* **40**, 261–265 (1993).
35. L.M. Barstow, B.E. Dale, and R.P. Tengerdy, *Biotechnol. Tech.* **2**, 237–242 (1988).
36. H. Narahara, Y. Koyama, T. Yoshida, P. Atthasampunna, and H. Taguchi, *J. Ferment. Technol.* **62**, 453–459 (1984).
37. S. Bandelier, R. Renaud, and A. Durand, *Process Biochem.* **32**, 141–145 (1997).
38. S. Tao, L. Beihui, and L. Zuohu, *J. Chem. Technol. Biotechnol.* **69**, 429–432 (1997).
39. C.J. Harris and S.A. Billings, *Self-Tuning and Adaptive Control: Theory and Applications*, IEEE Control Engineering Series, Peregrinus, London, 1981.
40. G.C. Goodwin and K.S. Sin, *Adaptive Filtering Prediction and Control*, Prentice Hall, Englewood Cliffs, N.J., 1984.
41. K. Schügerl, J. Müller, J. Niehoff, R. Hiddessen, K. Früh, and A. Lübbert, in N.M. Fish, R.I. Fox, and N.F. Thornhill eds., *Computer Applications in Fermentation Technology*, Elsevier, London, 1989.
42. S. Freyer, R. Eberhardt, D. Schurbuscher, and C. Wandrey, in N.M. Fish, R.I. Fox, and N.F. Thornhill eds., *Computer Applications in Fermentation Technology*, Elsevier, London, 1989.
43. J.M. Flaus, A. Cheruy, and J.M. Engasser, *J. Process Control* **1**, 271–281, 1991.
44. G. Montague, A. Morris, A. Wright, M. Aynsley, and A. Ward, *Can. J. Chem. Eng.* **64**, 567–580 (1986).
45. G. Bastin and D. Dochain, *On-Line Estimation and Adaptive Control of Bioreactors*, Elsevier, Amsterdam, 1990.
46. Y. Pomerleau and G. Viel, in *Proc. IFAC Symp. on Modelling and Control of Biotechnical Processes*, Keystone, Colo., 1992, pp. 315–318.

47. L. Chen, G. Bastin, and V. Van Breusegem, *Automatica* **31**, 55–65 (1995).
48. C.R. Cutler and B.L. Ramaker, *Proc. Joint Auto. Control Conf.*, paper WP5-B, San Francisco, 1980.
49. R. Auria, M. Morales, E. Villegas, and S. Revah, *Biotechnol. Bioeng.* **41**, 1007–1013, (1993).
50. J.M. Brass, F.W.J.M.M. Hoeks, and M. Rohner, *J. Biotechnol.* **59**, 63–72 (1997).
51. R. Simutis, R. Oliveira, M. Manikowski, S. Feyo de Azevedo, and A. Lübbert, *J. Biotechnol.* **59**, 73–89 (1997).
52. J. Sargantanis, M. Karim, V. Murphy, D. Ryoo, and R.P. Tengerdy, *Biotechnol. Bioeng.* **42**, 149–158 (1993).
53. P. Sangsurasak and D.A. Mitchell, *Chem. Eng. J.* **60**, 199–204 (1995).
54. J. Bailey and D. Ollis, *Biochemical Engineering Fundamentals*, McGraw-Hill, New York, 1986.
55. H.W. Blanch, *Chem. Eng. Commun.* **8**, 181–211 (1981).
56. D. Mitchell, D. Duong, P. Greenfield, and H. Doelle, *Biotechnol. Bioeng.* **38**, 353–362 (1991).
57. M. Gutiérrez-Rojas, R. Auria, J.C. Benet, and S. Revah, *Chem. Eng. J.* **60**, 189–198 (1995).
58. P. Gervais, P. Molin, W. Grajek, and M. Bensoussan, *Biotechnol. Bioeng.* **31**, 457–463 (1988).
59. S. Rajagopalan and J.M. Modak, *Chem. Eng. Sci.* **49**, 2187–2193 (1994).
60. J.C. De Reu, Ph.D. Thesis, Agricultural University of Wageningen, 1995.
61. J.P. Smits, *Solid-State Fermentation: Modelling Fungal Growth and Activity*, Ph.D. Thesis, Agricultural University of Wageningen, 1998.
62. K.W. Szewczyk and L. Myszka, *Bioprocess Eng.* **10**, 123–126 (1994).
63. S. Rajagopalan and J.M. Modak, *Chem. Eng. Sci.* **50**, 803–811 (1995).
64. S. Rajagopalan and J.M. Modak, *Bioprocess Eng.* **13**, 161–169 (1995).
65. S.J. Pirt, *Proc. R. Soc. London* **163**, 224–231 (1965).
66. P.J. Blanc, M.O. Loret, and G. Goma, in S. Roussos, B.K. Lonsane, M. Raimbault, and G. Viniegra-González eds., *Advances in Solid State Fermentation*, Kluwer, Paris, 1997, pp. 379–391.
67. S. Kumaran, C.A. Sastry, and S. Vikineswary, in D.L. Wise ed., *Global Environmental Biotechnology*, Elsevier, Amsterdam, 1997, pp. 239–248.
68. S. Vikineswary, K.S. Kumaran, S.K. Ling, N. Dinesh, and Y.L. Shim, in D.L. Wise ed., *Global Environmental Biotechnology*, Kluwer, Dordrecht, 1997, pp. 301–305.
69. E.L. Gaden, Jr., *J. Biochem. Microbiol. Technol. Eng.* **1**, 413 (1959).
70. G.F. Froment and K.B. Bischoff, *Chemical Reactor Design and Analysis*, Wiley, New York, 1979.
71. M. Raimbault, *Fermentation en Milieu Solide: Croissance de Champignons Filamenteux sur Substrat Amylacé*, ORSTOM, Paris, 1981.
72. D. Ryoo, Ph.D. Thesis, Colorado State University, 1990.
73. P.D. Sharma, P.J. Fisher, and J. Webster, *Trans. Br. Mycol. Soc.* **69**, 479–483 (1977).
74. A. Ebner, I. Solar, G. Acuña, R. Pérez-Correa, and E. Agosin, in D.L. Wise ed., *Global Environmental Biotechnology*, Kluwer, Dordrecth, 1997, pp. 211–220.
75. F. Pérez, A. Vecchiola, M. Pinto, and E. Agosin, *Phytochemistry* **41**, 675–679 (1996).
76. B.K. Lonsane, G. Saucedo-Castañeda, M. Raimbault, S. Roussos, G. Viniegra-Gonzalez, N. Ghildyal, M. Ramakrishna, and M. Krishnaiah, *Process Biochem.* **27**, 259–273 (1992).
77. C.C. Chou and J.H. Rwan, *J. Ferment. Bioeng.* **79**, 509–512 (1995).
78. A. Fujita, Y. Watanabe, K. Kishi, G. Ogawa, and T. Yoshizaki, *Nippon Shokuhin Kogyo Gakkaishi* **27**, 483–488 (1980).
79. Japanese Pat. Showa 61-205479 (1986).
80. J.M. Harper, *CRC Crit. Rev. Food Sci. Nutr.* **11**, 155–215 (1981).
81. S. Lue, F. Hsieh, and H.E. Huff, *Cereal Chem.* **68**, 227–234 (1991).
82. W.M. Wang, F. Klopfenstein, and G. Ponte Jr., *Cereal Chem.* **70**, 707–711 (1993).
83. M. Torrejón, Tesis de Ingeniero Civil de Industrias, Pontificia Universidad Católica de Chile, 1996.
84. B.K. Lonsane and M.M. Kriahnaiah, in H.W. Doelle, D.A. Mitchell, and C.E. Rolz eds., *Solid Substrate Cultivation*, Elsevier, London, 1992, pp. 147–171.
85. N.P. Ghildyal, M. Ramakrishna, B.K. Lonsane, and N.G. Karanth, *Process Biochem. Int.* **25** (1990).
86. R.W. Treybal, *Mass Transfer Operations*, Kin Keong, Singapore, 1981, pp. 737–765.
87. P.K. Kumar and B.K. Lonsane, *Process Biochem.* **22**, 139–143 (1987).
88. A. Borrow, E.G. Jefferys, R.H.J. Kessels, E.C. Lloyd, P.B. Lloyd, and I.S. Nixon, *Can. J. Microbiol.* **7**, 227–255 (1964).
89. P.K. Kumar and B.K. Lonsane, *Biotechnol. Lett.* **9**, 179–182 (1987).
90. U.S. Pat. Pending (1998), M. Pinto, E. Agosin, J.R. Pérez, F. Pérez, and A. Vecchiola.

SOLID SUBSTRATE FERMENTATIONS, ENZYME PRODUCTION, FOOD ENRICHMENT

YUSUF CHISTI
University of Almería
Almería, Spain

KEY WORDS

Bioreactors
Enzymes
Fermentors
Fermented foods
Indigenous fermented foods
Leaching
pH control
Product recovery
Scale-up
Temperature control
Water activity

OUTLINE

INTRODUCTION

Solid substrate fermentations involve microbial modification of a solid, undissolved substrate. This definition includes solid-state fermentations in which microbial cultures grow on a moist solid with little or no free water, although capillary water may be present (1). Examples of this type are seen in mushroom cultivation (2–4), bread making (5,6), and production of cheeses (7). Examples in which the solid substrate is submerged in large amounts of free water are pickling of vegetables (8) and brewing of soy sauce and sake (9,10), as well as bioconversion of cellulosic materials (11,12).

Solid substrate fermentations have been practiced worldwide since ancient times (9,13,14). Bread, sausages, and soy sauce are a few familiar products of solid substrate fermentation. Such fermentations are used in processing coffee and cocoa, as well as in environmentally significant operations: composting, silage production, soil bioremediation, biodesulfurization of coal, and metal recovery from low-grade ores (15). Solid substrate fermentations are employed also in producing enzymes and chemicals and in biotransformations.

Despite long history and widespread use, most solid substrate fermentations are not well understood. Processing is often an art, commonly employing poorly defined, naturally occurring, mixed microbial populations. Mixed cultures are sometimes essential to development of the desired product characteristics, including appearance, aroma, texture, and taste. With few exceptions, fermentations are labor-intensive batch operations. The level of technical sophistication and capital outlay vary a great deal. Solids concentration is generally high, whereas the water content is usually low (16). The product may be highly concentrated in comparison with submerged fermentations. Similarly, water consumption is comparatively lower, but space requirements can be substantial. Solid substrate culture generally requires less power because agitation is not excessive. A low-pressure blower is sufficient for air supply, unlike the higher-pressure compressors required in submerged fermentations (17). Unlike submerged culture, solid-state processes require extensive solids handling, which is difficult; nevertheless, highly mechanized, automated, and continuous processing of solids is being practiced in large *koji* (or molded grain) factories (9,10,18), as well as for other processes. Processing machinery tends to be complex. Hygienic processing practices are followed, but sterility standards that are common in submerged culture production of pharmaceuticals are not attained. Because of fungal spores and product- or microbe-contaminated dust from substrate solids, solid substrate fermentations often pose a greater health risk to operators than do submerged fermentations.

FERMENTATION EQUIPMENT

Solid substrate fermentation devices vary in technical sophistication from the very primitive banana leaf wrappings to highly automated machines used mainly in Japan. Simple fermentation systems (e.g., fermentation of cocoa beans in heaps) are quite effective in some large-scale processes. Commonly used fermentation devices are detailed in the following sections.

Tray Fermentors

One of the simplest and widely used fermentors is a wooden, metal, or plastic tray, often with a perforated or wire mesh bottom to improve air circulation (1,19,20). A shallow layer, usually less than 0.15 m deep, of pretreated (e.g., steamed) substrate is placed on the tray for fermentation. Individual trays or stacks may be located in temperature- and humidity-controlled chambers (Fig. 1) or simply in ventilated areas. A spacing of at least one tray height is usually allowed between stacked trays. Trays may be covered with cheesecloth to reduce contamination (10), but strict monosepticity is not attempted. Inoculation and occasional mixing are done manually, often by hand

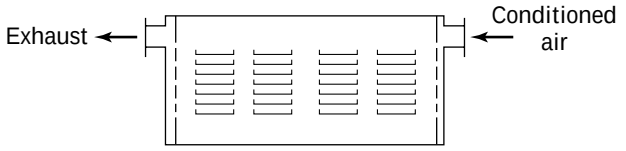


Figure 1. Tray fermentation chamber.

(10). This type of technology is quite common in small- and medium-scale koji operations in Asia. In some cases, filling, emptying, and other aspects of handling (4) of trays may be automated. For example, hinged trays with open ends may be dropped for emptying (19). Despite some automation, tray fermentors are labor intensive, require a large area, and difficulties with processing hundreds of trays limit their scalability (19). In parts of Asia, the bamboo baskets historically used instead of trays (16) continue to be employed.

Static Bed and Tunnel Fermentors

A commercially used modification of the tray fermentor employs a single, larger and deeper, static bed of substrate with forced aeration through the bed (Figs. 2 and 3). The substrate is located in an insulated chamber. In one version of this device, the tunnel fermentor (Fig. 3), the bed of solids may be quite long (1,20), but is usually no deeper than 0.5 m. Tunnel fermentors may be highly automated with mechanisms for mixing (Fig. 3), inoculation, continuous feeding, and harvest of substrate.

Rotary Disk Fermentors

Some large-scale commercial koji-making operations in Japan use a rotary disk fermenter configuration shown schematically in Figure 4 (18). The fermentor consists of upper and lower chambers, each with a circular perforated disk

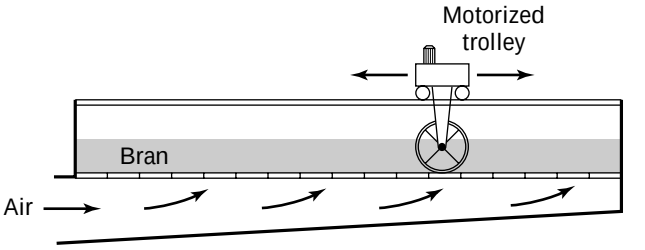


Figure 3. Mechanical mixing of fermenting substrate in a tunnel fermentor.

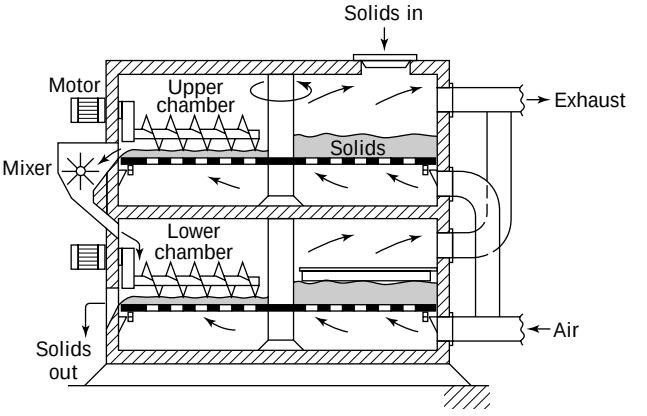


Figure 4. Automatic rotary koji fermentor.

to support the substrate. A common central shaft rotates the disks. Inoculated substrate is introduced in the upper chamber and slowly moved to the transfer screw. The upper screw transfers the partly fermented solids through a mixer to the lower chamber where further fermentation occurs. The mixer breaks up the partly fermented

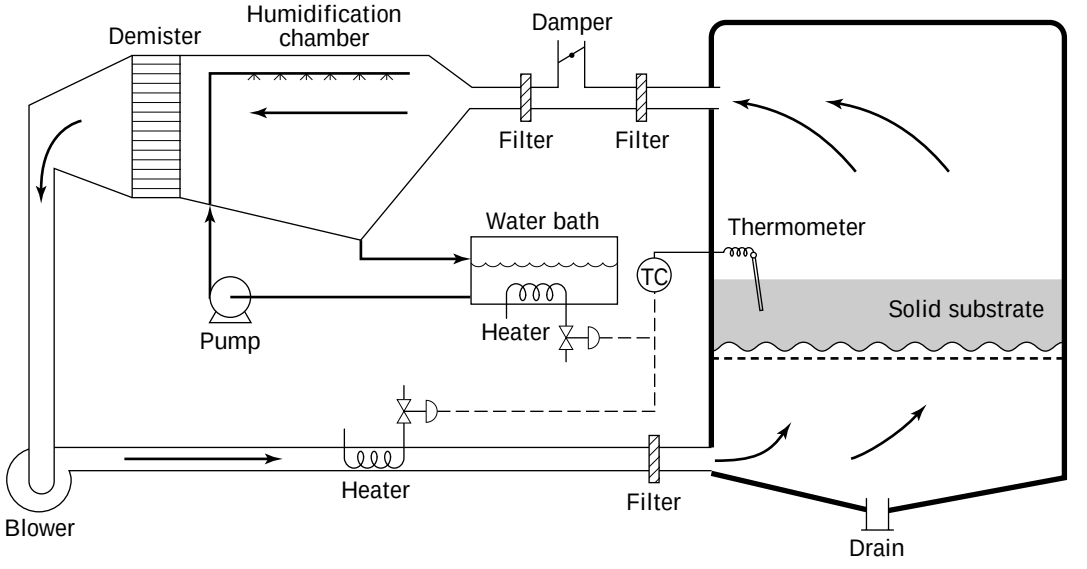


Figure 2. Static bed-type koji fermentor.

substrate–mycelium aggregates halfway through the fermentation process. Fermented substrate is eventually harvested using the lower transfer screw (Fig. 4). Both chambers are aerated with humidified, temperature-controlled air.

A tower fermentor similar in concept to the rotary koji fermentor is used in certain composting operations. The solids are supported on plates or trays that may be perforated. A stack of several tray chambers forms the tower. Mechanical devices mix and move solids down the tower (20–22) in a manner similar to that described for the rotary koji device. The gradually moving bed of solids is aerated either by forced aeration through the solid mass or solely from the surface exposed to the air above the bed.

Rotary Drum Fermentors

Rotary drum fermentors consist of a cylindrical vessel mounted horizontally on rollers and rotated around the long axis (Fig. 5). The rotational speed is commonly 1–5 rpm, but speeds up to 15 rpm are used occasionally. Rotation may be intermittent, and the speed may vary with the fermentation stage. The vessel may have straight or curved internal baffles that aid aeration and temperature control by imparting a tumbling motion to the substrate (23). Sometimes the drum may be inclined, causing the substrate to move from the higher inlet end to the lower outlet during rotation. The amount of substrate that can be processed in a batch operation depends directly on the length of the vessel and the square of its diameter. Aeration is through coaxial inlet and exhaust nozzles (Fig. 5). The air inlet pipe may extend into the vessel and may branch into several arms (19). Air is either supplied by a compressor or may be sucked in by an exhaust fan located at the air outlet (19).

Fluidized Beds

Fluidized-bed fermentors (Fig. 6) are relatively uncommon, but sufficiently promising that large-scale devices are being tested (9). A fluidized-bed fermentor consists of a relatively shallow bed of substrate supported on a perforated plate. Compressed conditioned air is normally used for fluidization, but carbon dioxide or nitrogen may be employed for anaerobic processes. Superficial gas velocity for fluidization is typically 0.24–1 m s^{-1}; the specific value

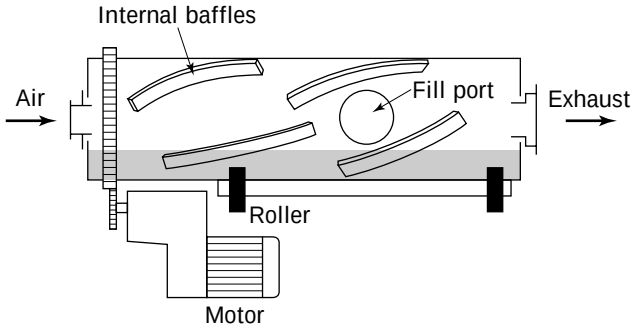


Figure 5. Rotary drum fermentor.

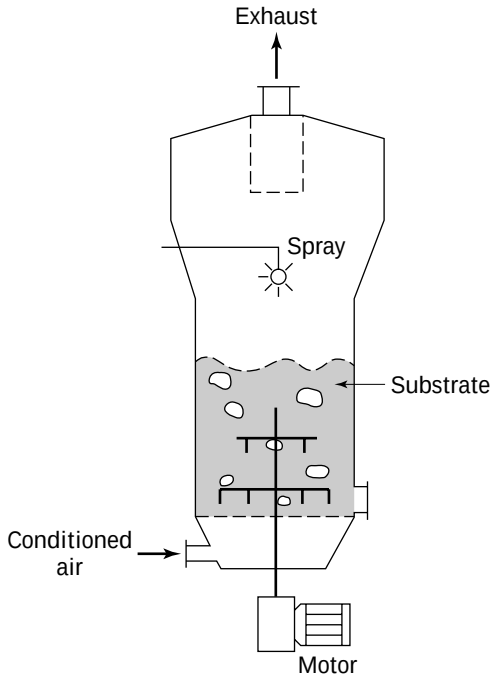


Figure 6. Fluidized-bed fermentor.

depends on the density and the particle size of the substrate. In the upper regions of the fermentor, the cross section is expanded (Fig. 6) to reduce the gas velocity to values that are insufficient to suspend the larger particles. Loss of fines in the exhaust gas is prevented by screens, filters, or cyclonic separators located at the exhaust nozzle. Fluidized beds achieve good aeration, mixing, and temperature control. Intermittent water spray may be necessary to prevent dehydration. A mechanical agitator as in Figure 6 is useful for periodic loosening of clumped solids.

Agitated Tank Fermentors

Helical ribbon-stirred tank fermentors have been employed for solid-state culture of fungi such as *Chaetomium cellulolyticum* on wheat straw (24). Other similar designs have utilized multiple helical screws for agitation of large rectangular tanks. The screws extend into tanks from mobile trolleys that ride horizontal rails located above the tanks (25). Yet another stirred tank configuration is the paddle fermentor. This design is similar to the rotary drum device (Fig. 5), except that the drum is stationary and motor-driven paddles on a concentric shaft provide periodic mixing of the substrate (20). As with rotary drums, the quantity of the substrate in paddle fermentors is generally restricted to less than 50% of the volume of the vessel. Although simpler than the rotary drum, the paddle fermentor is unlikely to be as effective in moving the entire mass of solids.

Continuous Screw Fermentors

A screw fermentor suitable for continuous fermentations is shown in Figure 7. Sterilized, cooled, and inoculated sub-

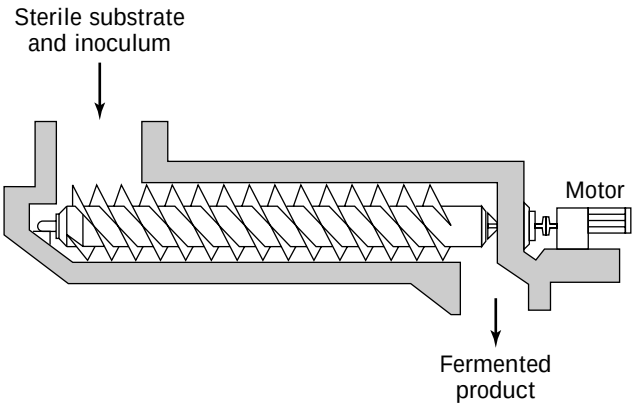


Figure 7. Screw fermentor for continuous fermentation.

strate is fed at the inlet. The screw moves the fermenting solids toward the harvest port. The fermentation time depends on the rotational speed and the length of the screw. Because the device is not aerated, only anaerobic or microaerophilic fermentations may be carried out.

Other Systems

Other miscellaneous types of fermentation devices that are used mostly in small-scale studies have been described by Lonsane et al. (19). Equipment for fermenting undissolved solid substrate slurries is identical to that used in submerged culture processes. Exceptions may be in devices used to feed slurried solids to the fermentor and remove the fermented product. Commonly used slurry fermentors include stirred tanks, bubble columns, and air-lift vessels (26–28). Simple vats are frequently employed in anaerobic slurry fermentations. Fermentors used in various types of composting have been described (14,21,22).

Construction and Cleaning

Modern commercial solid-state fermentors are constructed predominantly of stainless steel, but pressure vessel construction that is the norm for submerged culture processes (29) is not employed. Large concrete or brick fermentation chambers, or koji rooms (Fig. 1), may be lined with steel. Type 304 steel is used typically, but more corrosion-resistant grades such as Typed 304L and 316L are also utilized. Materials of construction have been discussed further by Chisti (29).

Because pressure vessel construction is not used, the fermentors cannot be steam sterilized; however, the design usually allows for atmospheric steaming. Typically, the fermentation chamber and process vessels are washed with alkaline detergent, sanitized with sodium hypochlorite, formaldehyde, or quaternary ammonium salts, and steamed between successive batches (17). Automated clean-in-place methods similar to those noted for submerged culture processes (30) are increasingly employed.

Design and Scale-Up

Design and scale-up of solid-state fermentation equipment relies heavily on empirical experience because the process kinetics and transport phenomena, such as heat transfer, oxygen transfer, and mixing, in these systems are insufficiently understood for more formal approaches to be usefully applied (20). Potentially useful strategies for scale-up have been discussed by Lonsane et al. (31), and Moo-Young et al. (22) have considered aspects of transport phenomena.

ENVIRONMENTAL FACTORS IN SOLID SUBSTRATE FERMENTATIONS

Moisture

Moisture content of the substrate is an important consideration in solid substrate processes. Free water indicates a saturated substrate. The moisture content at which free water is apparent depends on the nature of the solid (Table 1). Many solid substrate fermentations are carried out with little or no free water; typically, the substrate moisture content is 30–80% w/w (1,32). High moisture levels lead to aggregation of substrate particles, poor aeration, and possible anaerobic conditions (24). Steamed rice, a common substrate, becomes sticky when the moisture level exceeds 30–35% w/w. Optimal moisture content depends on the microorganism and the substrate. For the commonly used fungi, the optimal moisture requirement varies between 40% and 80% w/w. For the same organism growing on different substrates, the optimal moisture levels may differ widely; hence, percent moisture by itself is unreliable for predicting growth (32), and water activity is preferred instead. The water activity or relative humidity of the substrate is the ratio of the vapor pressure of water in the substrate to the vapor pressure of pure water at the temperature of the substrate. Water activities below 0.9 do not support most bacterial growth, but yeasts and fungi can grow at water activities of 0.7 and greater. Thus, the low moisture environment of many solid substrate fermentations favors yeasts and fungi. The water activity in a substrate may be determined by hygrometric measurement of relative humidity of a small air space in equilibrium with the substrate.

In addition to affecting growth, water activity affects product formation. The spectrum of products produced and characteristics such as aroma are affected. In some cases, the optimal water activities for growth and product formation differ (32). The optimal water activity depends also on factors such as agitation rate and cultivation temperature (32). Because the water activity depends on the concentration of dissolved solutes, sometimes salts, sugar, or other solutes are added to alter the activity. Different additives may influence the fermentation differently even though the water activity produced may be the same. Furthermore, the fermentation process itself leads to changes

Table 1. Saturation Moisture Content of Some Substrates

Substrate	Percent moisture (w/w)
Maple bark	>40
Rice and cassava	>50–55
Straw	75

in water activity as products are formed and the substrate is hydrolyzed. Oxidation of carbohydrates produces water.

During fermentation the water activity is controlled by aeration with humidified air and, sometimes, by intermittent water spray. Aeration with water-saturated air has commonly been found to increase the moisture content of the substrate. Relative humidity of the aeration gas is typically 60–80%. Ideally, prevention of water loss or gain by the uninoculated substrate would require that the water activity of the aeration gas and the substrate be identical. In practice, water generated during fermentation, and the additional requirement of evaporative cooling, demand aeration gas that is somewhat drier than the substrate. The temperature and humidity of the aeration gas are controlled in the same way as in air conditioning systems: temperature and humidity sensors (hygrometers) in the fermentation chamber provide the necessary data, and cooler air is humidified by water spray or injection of clean steam, and warmed to reduce the humidity to the required value (Fig. 2). Humidity–temperature relationships available in standard psychrometric charts are the basis of control.

Temperature

At 10–30 kg m^{-3}, the maximum microbial biomass concentration in solid-state fermentations is lower than 40–50 kg m^{-3} typically seen in submerged culture (24); but, because there is little water, the heat generation per unit fermenting mass tends to be much greater in solid-state fermentations. Temperature can rise rapidly, again, because there is little water to absorb the heat (i.e., the mean specific heat capacity of the fermenting mass is much lower than that of water). Consequently, temperature control in large-scale fermentations can be particularly difficult. Removal of the metabolic heat sometimes becomes the major limitation, especially because of the poor thermal conductivity of the porous fermenting mass. Cumulative metabolic heat generation in koji fermentations for a variety of products has been noted at 419–2,387 kJ kg^{-1} solids (14). Higher values, up to 13,398 kJ kg^{-1}, have been observed during composting (32). Peak heat generation rates in koji processes range over 71–159 kJ kg^{-1} h^{-1}, but the average rates are more moderate, at 25–67 kJ kg^{-1} h^{-1} (32). Peak metabolic heat production rate during fermentation of readily oxidized substrates such as starch can be much greater than in typical koji processes. For example, in one case, a generation rate of 330 kJ kg^{-1} h^{-1} was observed for *Aspergillus niger* growing on cassava starch. Heat generation during composting has been treated in depth by Stentiford and Dodds (33).

Temperature control during fermentation is obtained mostly through evaporative cooling (34); hence, drier air provides a better cooling effect. Air temperature is controlled. Intermittent spray of cool water is sometimes necessary to prevent dehydration of the substrate. Occasionally, the substrate-containing metal trays are also cooled by a circulating coolant, even though most relatively dry and porous substrates are poor conductors. Intermittent agitation further aids heat removal, particularly in large heaps and piles. Despite much effort, temperature gradients in the substrate do occur, particularly during peak growth. Gradients as steep as 3 °C cm^{-1} have been recorded during rapid growth of *Rhizopus oligosporus* on soybeans in *tempe* fermentations, even though the substrate layer was less than 7 cm deep (35). Similar values were noted in solid-state citric acid production by *Aspergillus niger* (34).

pH

Solid-state fermentations are practiced without pH control (19) other than any adjustments made during substrate preparation. In some nonsterile processes the initial pH may be adjusted to pH 4 or less to suppress bacterial contamination. Unlike most bacteria, yeasts and fungi are generally tolerant of more acidic conditions. Either mineral acids (e.g., 0.2 M hydrochloric acid) or organic acids (e.g., 0.5% w/w acetic acid) may be used for pH adjustments. Other than initial adjustments, the buffering capacity of substrates is relied on to check large changes in pH during fermentation (19). Many substrates are effective buffers. This is particularly true of protein-rich substrates, especially if deamination of protein is minimal. Some pH stability can be obtained by using a combination of urea and ammonium sulfate as the nitrogen source in the substrate. Decomposition of urea produces ammonia that takes up a proton to become NH^{+}_{4}, causing an increase in pH. This effect is countered by ejection of a proton from cells that take up the ammonium ion but incorporate it into proteins as —NH^{+}_{3}. Uptake of ammonium causes rapid acidification. In the absence of other contributing nitrogen sources, an equimolar combination of ammonium sulfate and urea is expected to yield the greatest pH stability. In media that are rich in protein or amino acids, deamination may contribute to pH rise. Attention to stabilizing the pH during fermentation can be particularly important for certain processes. Stability of enzymes and some secondary metabolites is pH dependent, and even though the rate of production may not be affected by changes in pH, the overall process productivity may decline because of destruction of the product. This behavior has been observed in producing pectinases by *Aspergillus niger* cultured on glucose-enriched sugar cane bagasse (36).

PRODUCTION OF ENZYMES

Process Technology

Only extracellular enzymes, those secreted by the microbial cell into the extracellular medium, can be produced by solid-state fermentation technology (37). Predominantly, these are hydrolytic enzymes (38,39), as noted in Table 2. Commercial production of many of those enzymes in the West utilizes submerged culture, but solid-state fermentation persists particularly in Asia where koji-type processing is commonly used. Koji is molded grain that has been used in oriental food preparations for thousands of years (40,41). Koji is a source of fungal enzymes that digest proteins, carbohydrates, and lipids into nutrients used by other microorganisms in subsequent fermentation. Koji

Table 2. Enzymes That May Be Produced by Solid-State Fermentation

Enzyme	Producing organisms
Amylase	*Aspergillus oryzae, Aspergillus niger, Bacillus licheniformis, Bacillus subtilis, Bacillus megaterium*
Catalase	*A. oryzae, Rhizopus niveus*
Cellulase	*Trichoderma reesei, Trichoderma viride, Trichoderma koningi, Trichoderma harzianum, Aspergillus ustus, A. oryzae, Sporotrichum pulverulentum, Penicillium spinulosum, Penicillium capsulatum, Pestalotiopsis versicolor*
Chitinase	*A. niger*
Glucoamylase	*A. niger, A. oryzae, Rhizopus* sp.
Invertase	*Aspergillus awamori, A. niger, A. oryzae*
Lactase	*A. oryzae*
Linamarase	*Aspergillus sydowi, Penicillium steckii*
Lipase	*A. niger, Aspergillus luchuensis, Rhizopus delemar, Penicillium candidum*
Pectinase	*A. niger, Aspergillus sojae, Aspergillus fumigatus, Bassochlumys fulva, Penicillium expansum*
Phytase	*Aspergillus ficuum, Rhizopus oligosporus*
Protease	*A. oryzae, A. niger, Aspergillus flavus, Mucor dispersus*
Rennet	*Mucor pusillus, Mucor meihei, R. oligosporus*
Ribonuclease	*Aspergillus candidus*
Xylanase	*A. niger, A. fumigatus, Aspergillus terreus, Gibberella fujikuroi*

comes in many varieties depending on the mold, substrate, method of preparation, and stage of harvest. Traditional and mechanized koji production has been described by several sources (9,10,18), and concise outlines of food koji processes are given later in this article. In typical processing for enzymes, a suitable solid substrate is mixed with water and mineral salts, heat sterilized/cooked, cooled, and inoculated. Incubation at controlled temperature (20–45 °C) and humidity follows. Subsequently, the fermented substrate is extracted with buffer to yield an enzyme-containing liquor. Further purification and recovery methods are identical to those that have been described for submerged culture processes (17,27,42,43). The crude liquor is filtered or centrifuged to remove suspended solids and concentrated by ultrafiltration. The concentrate may be stabilized and packaged. Extensive purification is uncommon for many bulk enzymes. In some cases, the fermented substrate may be dried at 35–40 °C, and ground before extraction (44).

The commonly used solid substrates include wheat bran, wheat, soybean, rice, barley, oats, and other cereals, but many more substrates have been used occasionally (Table 3). Studies with nutritionally inert substrates (e.g., polyurethane foam) wetted with dissolved nutrients have been reported (45). Two or more substrates are sometimes used in combination. Cereal grain may be used whole or cracked into pieces. Substrate particle size affects the extent and the rate of microbial colonization, air penetration, and carbon dioxide removal as well as the downstream extraction and handling characteristics (19,46). Relatively small particles are preferred because they present a larger surface for microbial action. However, particles that are too small pack together to reduce the interparticle voids that are essential for aeration (46). Similarly, too many fines in a batch of larger particles will fill up the voids. Particle shapes that pack together tightly (e.g., flat flakes, cubes) are undesirable. Lignocellulosic substrates such as straw, sugar cane bagasse, and wood chips are sometimes extensively pretreated to ease microbial colonization and substrate utilization. Pretreatments include size reduction, steaming, steam explosion, solubilization of lignin with organic solvents, acid and alkaline hydrolysis, and irradiation (12,47,48). Most of those pretreatments have proven too expensive for commercial utilization. For cereal substrates, common pretreatments include dehulling, pearling, size reduction, soaking, and cooking or steaming (9,10,18).

Table 3. Solid-State Fermentation Substrates

Animal	Milk solids
Apple pulp	Millet
Bagasse	Newspaper
Banana meal	Nutrient wetted inerts
Barley	Oats
Beet pulp	Orange peel
Buckwheat	Paper pulp
Canola meal	Peanut press cake
Cardboard	Potato starch
Cassava starch	Rice
Coffee pulp	Rice bran
Corn cobs	Rye meal
Corn grits	Sawdust
Cotton seed meal	Sewage sludge
Fish	Soybean
Fruit pulp	Straw
Jerusalem artichoke meal	Sweet potato meal
Lignocellulosic waste	Wheat bran
Logs	Wood chips
Meat	Vegetables

The selection of substrate depends to some extent on the enzyme of interest. Thus, pectinases are often produced using fungi grown on fruit pulp (e.g., apple, avocado), or on a combination of pulp and bran, or bran and fruit juice. Similarly, cellulose-rich substrates such as rice and wheat straw are commonly employed in producing cellulases.

Enzyme production processes rely overwhelmingly on filamentous microfungi (43), the main producer genera being *Aspergillus*, *Mucor*, *Penicillium*, *Trichoderma*, and *Rhizopus*. *Bacillus* species predominate among bacteria used in enzyme production (38,43,44). Examples of pro-

duction from yeasts are few. Some specific enzyme producers are listed in Table 2. Production of β-galactosidase by the yeast *Kluveromyces lactis* grown on corn grits and wheat bran moistened with deproteinated milk whey has been reported (49). Although β-galactosidase is an intracellular enzyme in yeasts and cultivation on solid substrate does not promote secretion, solid-state fermentation has been claimed to be superior to liquid culture for this enzyme (49); however, recovery questions have not been addressed.

Microbial inocula for commercial enzyme production are maintained as pure cultures (17). Routine checks on culture purity, freedom from variant forms, physiological characteristics, and enzyme productivity are used to ensure process consistency (17). The sterile solid substrate is inoculated in one of several ways: by mixing with a liquid suspension culture inoculum; by mixing with solid substrate-grown inoculant; by using a suspension of solid substrate-produced spores in a sterile liquid; or by blowing the solid substrate-produced spores with sterile air going into the fermentation chamber. Aseptic processing practice is observed (17), but fermentations are never entirely contamination free. Fermentations last 1–7 days, and a significant amount of the substrate is oxidized to carbon dioxide and water (17). Thin uncompacted and porous layers of nonadhering solid substrate particles generally ensure good oxygen penetration from the exposed surface. Excessive agitation is not wanted. Occasional turning and mixing improve oxygen transfer and reduce compaction and mycelial binding of substrate particles. Because agitation continually damages the surface hyphae, mixing suppresses often-unwanted sporulation (46). The frequency of agitation may be purely experience based, as in occasional turning of a fermenting cocoa heap, or it may be determined by a temperature controller.

Production of spore inocula requires static conditions. Deep substrate layers and heaps may require forced aeration and agitation. Oxygen supply affects product formation. Depending on the microorganism and the product, both enhanced and reduced productivities have been observed to accompany improved aeration or agitation. Aeration plays an important role in removing carbon dioxide and in controlling temperature and moisture. In some cases, increased concentrations of carbon dioxide have been severely inhibitory, whereas enhanced oxygen partial pressures have improved productivity. If used, forced aeration rates may vary widely; a typical range being 0.05–0.2×10^{-3} m^3 kg^{-1} min^{-1} (1).

Regulatory Considerations

Production of food, medicinal, and medical–diagnostic enzymes must comply with standards of "good manufacturing practices" as established by the U.S. Food and Drug Administration (FDA) or other regulatory agency appropriate to the particular jurisdiction. Safety of enzyme preparations, especially those derived from non-GRAS species (GRAS, generally recognized as safe; i.e., microorganisms such as baker's yeast that have proven to be safe through extended use and have been recognized as such by placement on the GRAS list of the FDA), must be demonstrated to gain marketing approval. Specific requirements vary with the intended use of the product. Freedom from toxins, antibiotics, other contaminants, and unwanted effects may have to be ascertained. Multigeneration feeding trials and studies of teratogenicity may be necessary. Requisite purity, potency, and stability may have to be demonstrated. The quality assurance requirements are similar to those for therapeutic proteins produced by other methods (42,50–52). Enzymes for industrial and household use (e.g., detergent enzymes) also must comply with product safety criteria, although the requirements are less stringent than in food, feed, veterinary, and medical applications.

OTHER PRODUCTS AND APPLICATIONS

Secondary Metabolites

Although solid-state surface culture technology was employed in the earliest attempts at large-scale production of penicillin, submerged culture is now the only method used in commercial production of antibiotics. Nevertheless, promising research on solid-state production methods continues, albeit at a slow pace. Solid-state production of oxytetracycline by *Streptomyces rimosus* grown on corncob has been described (53). Cephalosporins have been produced using *Cephalosporium acremonium*, *Streptomyces clavulgerus*, and *Aspergillus chrysogenum*. *Streptomyces cinnamonensis* has been used to produce monensin, an anticoccoidal agent for poultry. Other potentially antibiotic-producing *Streptomyces* that have been cultured in solid state include *S. badius*, *S. aureofaciens*, and *S. flavovirens* (1). Studies of penicillin production by several strains of *Penicillium chrysogenum* have shown that the best producers in submerged culture are not necessarily the best in solid-state culture (54). Improving productivity demands selection of strains suited specifically to solid-state fermentation.

Solid substrate fermentation for producing gallic acid used in tanning, printing, and other applications has been described (13,14). Koji technology is employed in citric acid production in Japan (16). In producing citric acid by *Aspergillus niger*, Pallares et al. (45) noted that solid-state culture had doubled the productivity (g L^{-1} d^{-1}) of submerged culture, reducing the fermentation time from 14 to 6 days. Medium-impregnated polyurethane foam cubes were used as substrate (45). Productivity was sensitive to substrate particle size; 0.3-cm cubes gave the best results (45). Other products that have been produced by solid-state culture include several mycotoxins (14,46,55) and ergot alkaloids from various fungi, gibberellic acid from *Gibberella fujikuroi* (56,57), and kojic acid (14). In one case involving production of gibberellic acid and proteases, the yield of products was improved by fed-batch strategies (57), but nutrient feeding of solid-state batches other than at start-up remains a rare practice.

Ethanol has been commercially produced by aerobic–anaerobic solid-state fermentation of sweet sorghum grains in Taiwan (24). Production of γ-linolenic acid, a polyunsaturated fatty acid, by *Cunninghamella japonica* grown on rice and millet has been reported (58). The fungus *Monascus purpureus* or *Monascus anka* is used to com-

mercially produce a bright red polyketide pigment for food coloring and flavoring in Asia (41). An additional and developing application of solid-state fermentations is in producing fungal spores for use in a variety of biotransformations (1,14). Use of phytopathogenic and entomopathogenic fungi as herbicidal and insecticidal biocontrol agents promises further applications of solid-state fermentation technology for spore production (14,59–61).

Environmental Remediation

Because of their ability to degrade otherwise persistent toxic organics, white rot fungi—for example, *Phanerochaete chrysosporium*, *Phanerochaete sordida*, and *Trametes versicolor*—are being tested for bioremediation of soils contaminated with polychlorinated biphenyls, pentachlorophenol, and other polycyclic aromatics (62). In the preferred mode of implementation, solid-state cultured fungal carriers such as corncobs, wood chips, and straw are mixed with the soil in situ. Commercial implementation requires, among other factors, the ability to economically and consistently produce large inocula. Solid substrates enriched with nutrients and concentrates of fugal spores, and formulated into pellets, are being developed to improve consistency of performance (62). Other pollution remediation applications of solid-state fermentations are composting (22,33,63), anaerobic digestion for producing methane in slurry fermentations and landfills (64–66), and treatment of gaseous pollutants and vapors in biofilters (65).

Animal Feeds

Solid substrate fermentation methods are used to enrich many wastes and low-grade substrates for animal feed (1,11,12,22,25,46,65,67). One such widely used process is ensiling. Ensiling is used to stabilize and preserve forage crops; any nutrient enrichment is incidental. Typically, crops are ensiled immediately on harvest. An uncontrolled anaerobic primary fermentation by naturally occurring or inoculated (68,69) lactic acid bacteria takes place. Sugars present in the fodder are converted to lactic and acetic acids, causing a drop in pH. A low pH suppresses secondary fermentation by *Clostridia* that leads to butyric acid and spoilage. A good silage should generally have a pH of 4.0 to 4.2 and a butyric acid content of less than 5% of dry weight. A high dry matter content, for example, 20–30% w/w, of biomass at ensiling is preferred because secondary fermentation is difficult to suppress when dry solids are less than 15% w/w. Unlike *Clostridia*, the lactic acid bacteria remain active up to 70% w/w dry matter. Microbiology of silage has been detailed by Whittenbury (70) and Fenlon et al. (68). Silage inoculants have been described by Seale (69). Certain silages may contain animal and food processing waste (e.g., cattle and poultry manure, fish waste, or offal) and grain (e.g., corn), in addition to grass and other fodder crops. Fish silage has been described by Wignall and Tatterson (71); in-depth treatments of silage fermentations have been reported by McDonald (72) and Woolford (73).

PRODUCT RECOVERY FROM FERMENTED SOLIDS

Unless the fermented solids are the desired product, the solids must be further processed to recover the enzyme or other metabolite. Volatile products such as ethanol may be recovered by direct distillation of the fermented mash (24). Most other products require extraction from fermented solids. The solids may be extracted moist or dry (44); moisture content can affect the extraction behavior. Because drying generally improves keeping quality, dried material from several batches can be accumulated for later extraction. The solid cake should be crumbled or otherwise size reduced for improved extraction. Either batch or continuous extraction may be employed. When feasible, continuous extraction with countercurrent flow of solids and solvent allows greater product recovery in a more concentrated solution. A suitable continuous extraction device is the Chisti contactor shown in Figure 8. Leaching machinery developed for the chemical industry can also be potentially adapted for use with fermented solids. A series of mixer-settlers has been used in small-scale extractions (74), but that arrangement is too cumbersome for commercial practice. Between two and five contacting stages are generally sufficient, but a larger number may be practicable with high-value products.

Cumulative contact periods typically do not exceed 50 min. Water, aqueous buffers, or dilute solutions of salts (e.g., 1% sodium chloride) or glycerol (e.g., 1% glycerol) are used to leach enzymes (74). Aqueous solutions of ethyl acetate, acetone, ethanol, or chloroform may be needed for nonprotein products such as most antibiotics and mycotoxins. In one case, aqueous ethanol (10%) was the best solvent for extracting gibberellic acid from fermented wheat bran (74). Pure methylene chloride has been used to extract mycotoxins (1). A suitable solvent must be selected empirically for each application. Factors that affect solvent choice are cost, toxicity, flammability, corrosion effects, disposal issues, solubility and stability of product, solubilities of other solutes, and compatibility with end use. The

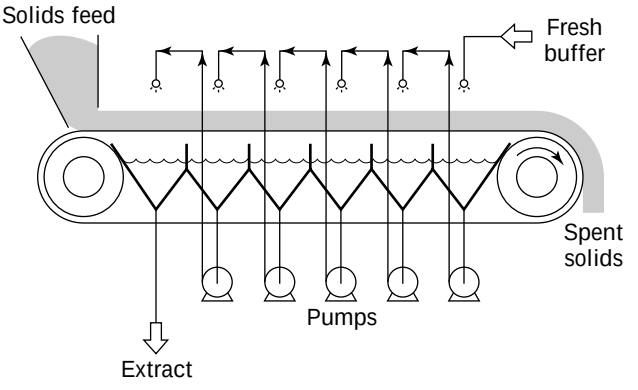


Figure 8. Continuous countercurrent leaching of solids in an extractor. Solids are placed continuously on the stainless steel conveyer mesh (0.6 m wide, 6 m long) that moves at 0.1–0.2 m min^{-1}. Fresh buffer is sprayed at the downstream end of the conveyer; the percolated extract is collected for extracting solids further upstream.

solvent-to-solid ratio generally ranges over 1 to 8; higher values increase the fraction of the product leached at the expense of producing a more dilute solution. Extraction temperature does not normally exceed 30 °C, but enzymes are commonly extracted at lower temperatures (e.g., 4–10 °C) to reduce denaturation, proteolysis, and microbial growth. Extraction performance and stability of the product may depend critically on pH. The pH of greatest stability may not be the same as that required for optimal leaching (75). Enzymes are generally extracted at pH values close to 7; however, to prevent precipitation, the extraction pH must not be the same as the isoelectric pH. Once the product has been extracted into solution, further purification employs the same well-known methods that are used in recovering extracellular products from submerged fermentations (27,42,43).

FERMENTATION-ENRICHED AND FERMENTATION-MODIFIED FOODS

Solid-state fermented foods occur worldwide. In some regions, such foods are the primary source of calories, proteins, and other nutrients for the consumers. Historically, fermented foods arose more by accident than by guided effort; nonetheless, recognition of the numerous advantages of the fermented products led to their established use. In comparison with the unfermented substrate, the fermented material may have one or more of the following characteristics: enhanced flavor, texture, and digestibility; reduced or eliminated undesirable compounds or flavors; protein enrichment or otherwise improved nutritional value; changed physical character; improved keeping quality; fast food quality of easier or quicker preparation; and other economic advantages such as reduced cooking time and hence lower energy demand. Table 4 notes some of the world's significant fermented products. Further details appear in the *Handbook of Indigenous Fermented Foods* (76), as well as other sources (6,8,40,41,77). The contribution of fermented products to the human food supply is likely to remain strong (78), and increasing cross-cultural spread is expected considering the popularity of tempe, tofu, and soy sauce in nonindigenous regions. The following sections discuss a small selection of the important foods as representative illustrations of the solid-state and solid substrate fermentation technologies.

Miso

Miso is a semisolid fermented soybean product that is widely consumed in Japan (18,40,41,79). Similar foods are produced in other parts of Asia (Table 4). In Japan miso is made into soup with added ingredients such as vegetables, meat, and seafood (18). Modern miso is generally produced from protein-rich yellow soybeans, but other varieties can be used. Cleaned soybeans are continuously washed, soaked overnight, drained, and cooked (Fig. 9). Cooking conditions vary. Large plants employ continuous pressure cookers with cooking done at 1.3–2.0 kg cm^{-2} for 2–7 min (18). Cooked beans are ground using a meat grinder with 5-mm-diameter holes (18) before being mixed with separately produced koji for fermentation. Barley koji is used in making "rural miso"; commercial processes employ rice koji (18). Miso koji is made from milled, cleaned, and usually cracked rice that is washed and soaked overnight in water at approximately 15 °C (18). Soaked rice is drained (1 h) and steamed. A horizontal belt steamer (Fig. 10) is commonly used. Steamed rice is unloaded onto a cooling conveyer, where a stream of air is blown through the rice layer to lower the temperature to 35 °C (18). Rice is then inoculated with the starter culture (called *tane-koji* in Japanese) of *Aspergillus oryzae*. Tane-koji is an olive-green, sporulated culture that has been grown on cooked rice mixed with a small amount of wood ash (18). Tane-koji is produced under aseptic and strictly controlled conditions. One kilogram of tane-koji is sufficient to inoculate 1,000 kg of rice (18). Many varieties of tane-koji are available with different capabilities for hydrolyzing proteins, carbohydrates, and lipids (18). An appropriate variety should be selected with regard to the amounts of proteins, oils, and carbohydrates in the mix of miso substrates (18).

Inoculated rice is transferred to a koji fermentor. Either a tunnel fermentor (Fig. 3) or, in larger plants, a rotary fermentor (Fig. 4) is used. The tunnel fermentor is mixed twice during the fermentation, but the rotary machine is mixed only once. Fermentation occurs at 30 °C for 40 h (18). Rice koji is then harvested and mixed with ground beans, salt, and a miso starter. The starter consists of halophilic yeasts, *Zygosaccharomyces rouxii* and *Candida versatilis*, and a halophilic lactic acid bacterium, *Pediococcus halophilus* (18). Fermentation is carried out in wooden vats, or tanks made of stainless steel or fiberglass-reinforced plastic. During fermentation miso is occasionally transferred from one vat to another to maintain homogeneity (18). The fermentation temperature is 25–30 °C (18). Ripened miso is taken out of vats, blended, pasteurized, and packaged. Process variations occur depending on the type of product desired (41).

The miso koji mold, *A. oryzae*, grows well around 30 °C (18), but optimal conditions for spore germination have been found to be 36 °C, ≥97% relative humidity, 0.1% carbon dioxide, and 20% oxygen (18). Optimal conditions for mycelial growth are 35–37 °C, ≥75% relative humidity, <1% carbon dioxide, and 20% oxygen (18). In practice, a lower than optimum temperature (28–30 °C) is used for germination, and 30–35 °C for cultivation in the koji chamber (18), apparently to avoid overheating. Desirable temperatures for production of enzymes are as follows: glucoamylase, 30 °C; α-amylase, 35 °C; protease, 25–30 °C; and carboxypeptidase, 35 °C (18). The optimal moisture level of cooked rice is 36–38%, and the optimal pH is 6.0–6.4 (18). Fermentation pH can be effectively maintained at 6–7 by supplementing the substrate with sodium glutamate or sodium succinate. Controlled pH favors higher-activity proteases (18). The microbiology and biochemistry of miso fermentation are further discussed by Abiose et al. (82).

Sake

Sake is a Japanese alcoholic beverage made from rice (10,40). Rice koji preparations for sake brewing are obtained by methods similar to those described for miso. The amount of tane-koji inoculum is 0.6–1 kg per kilogram of

Table 4. Solid Substrate Fermented Foods of the World

Sufu. A Chinese cheese made from soybean curd (*tofu*) by fermentation with molds such as *Actinomucor elegans* and *Mucor hiemalis* (41).

Idli. Small, white, steamed rice cake made by lactic acid fermentation of a thick batter of coarse-ground rice and fine-ground black gram in India. Idli cakes are moist, soft, and spongy, with a slightly sour flavor. Several varieties are produced. Other similar products of India are *dhokla* and *khaman*.

Dosa. Pancake made with thin pourable fermented batter of fine-ground rice and gram. Important part of diet in South India and Sri Lanka.

Ambali. A cooked product of fermented millet flour batter and rice. India.

Puto. Steamed cake of milled, acid-fermented rice. Similar to idli, but without legume. Important food in Philippines. Several varieties.

Enjera. Very flexible, soft, sour leavened pancake containing many small uniformly distributed gas bubbles. Made from thin, pourable fermented batter of *tef* (*Eragrostis abyssinica*) flour, or other ground grain (corn, sorghum, millet, barley, wheat). Consumed as staple in Ethiopia.

Hopper or ***appa***. Steam-baked product made from acid-fermented dough (rice or wheat flour) formulated into liquidy batter with coconut water. Sri Lanka.

Kisra. Flat bread made from fermented sorghum flour batter in Sudan.

Ogi (also *furah* and other names). Fermented cereal flour (corn, sorghum, or millet) made into a slightly sour gruel. Important source of calories in Nigeria.

Uji (and other names). Suspension of cereal (corn, millet, sorghum) or cassava flour (relatively large particles) in water. Sour-acid fermented either before or after cooking. East Africa.

Mahewu (also *magou* or *mageu*). Suspension of acid-fermented corn flour used as high-calorie beverage/food. More dilute than uji or ogi. Cooked cornmeal in water is fermented after adding a small amount of wheat flour. Product is very slightly alcoholic. Produced in large commercial operations in South Africa. Similar cereal-derived alcoholic beverages/foods occur throughout sub-Saharan Africa.

Sake. Rice wine of Japan, and other similar products: *brem bali* (Indonesia); *mie-chiu* (China); *tapoi* (Philippines); *sato* (Thailand); *yakju* (Korea); and *sonti* (India). See text.

Gari. A dry granular starchy product of West Africa (80,81). Made into staple food by adding hot or cold water. Made from peeled, grated cassava by fermenting in cloth bags while squeezing out the water over 12–96 h. The anaerobic acid fermentation produces a sour taste. Fermented solids are sun dried and lightly toasted in iron pans. A small amount of palm oil may be added during toasting to impart a pale-yellow color. Solids are sieved and sold in markets. Very poisonous varieties of cassava may be used in making gari. Crushing of cassava releases naturally present linamarase, which breaks down the cyanogenic linamarin to reduce the cyanide content. Cyanide produced boils off as hydrogen cyanide gas, indicating the need for good ventilation during processing. Acid pH produced during fermentation enhances the action of linamarase (pH optimum of 5.5). In addition, microbial action may produce enzymes that enhance the breakdown of cyanoglycosides.

Miso. Fermented soybean paste of Japan. Other similar products: *jang* (China); *doenjang* and *kochujang* (Korea); *taucho* or *tauco* (Indonesia); *tao-tsi* (Philippines) (18,40,41,79,82). See text.

Soy sauce. Japanese *shoyu*, and similar products throughout Asia: *chiang-yu* (China); *kan jang* (Korea); *kecap* (Indonesia) (9,40,41,76,79). See text.

Natto. Soaked and cooked soybeans fermented by bacteria, principally *Bacillus natto*, consumed in Japan. Other similar products: *thua-nao* of Thailand; *kinema* of Nepal, Bhutan, Sikkim, and parts of northeastern hills of India; *dagé*, an Indonesian produce made of bacterial fermented peanuts and other oilseeds.

Tempe (or *tempeh*) and similar products: *tempe kedele*, *tempe gembus*, *tempe benguk*, *tempe bongkrek*, *oncom*, or *ontjom*. See text. A Japanese solid food, *hamanatto*, is also made from soy beans fermented by fungi (41).

Dawadawa. Fermented African locust bean (*Parkia filicoidea*) product used as flavoring agent in West Africa (14).

Kenkey (and other similar products). Fermented corn dough used as staple in Ghana (8).

Mawé. Acid-fermented corn dough used in Benin and Togo (8).

Pozol. Fermented corn dough used as staple in parts of Mexico.

Kocho. Unleavened flat bread made from fermented pulp of ensete (*Ensete ventricosum*) or 'false banana' plant (a banana-like plant that does not bear bananas). Used as staple in certain regions of Ethiopia.

Tapai. Malaysian fermented glutinous rice or cassava gratings. Sweet and mildly alcoholic. Used as dessert. Similar products occur elsewhere: *lao chao* in China (41); *tapé ketan* and *tapé ketella* of Indonesia (40); and *peujeum* in Sudan.

Nham. Fermented minced raw pork similar to sausages. Thailand.

Fermented seafood. Usually fish or shrimp pastes, or sauces: *bagoong*, *patis* (Philippines); *ngapi* (Burma); *trassi*, *kccap ikan* (Indonesia); *belachan*, *budu* (Malaysia); *mam*, *nuoc-mam* (Vietnam); *jeotkal* (Korea); *prahoc* (Cambodia); *kapi* (Thailand); *padec* (Laos); *shottsuru*, *shiokara*, *katsuobushi* (Japan); *faseich* (Sudan) (14); *kobi* (Ghana) (14); and fermented fish of Northern Europe.

Balao balao. Fermented rice–shrimp mixture eaten in Philippines. Cooked, somewhat pasty rice is blended with fresh, unshelled, salted shrimps, and fermented for 7–10 days. Shrimp shell becomes quite soft. Product cooked before consumption.

Burong dalag. A fermented fish and rice product of the Philippines.

Fermented milk products. Yogurt; sour cream; cheeses; *ergo* and others (Ethiopia); *kefir*, *koumiss* (Russia); *dahi* (India); *tairu* (Malaysia); *laban rayeb*, *laban zeer*, *laban zabadi* (Egypt); *liban argeel*, *liban khather*, *mast*, *mass taw*, *shenina*, *dabbo* (Iraq).

Kishk. A fermented milk–wheat mixture of Egypt. Stored as dry balls.

Trahanas (and other names). Made from crushed wheat and fermented sheep milk in Greece, Turkey, and Cyprus. Stored as dry biscuits or dry crumbled solids.

Fermented vegetables (8). Sauerkraut (Europe, North America); Chinese sauerkraut *hum choy*; acid-pickled olives and cucumbers; various Indian and Malaysian pickles; various varieties of *kimchi* (Korea); kimchilike products of Thailand and elsewhere.

Source: Refs. 6, 76, and 77.

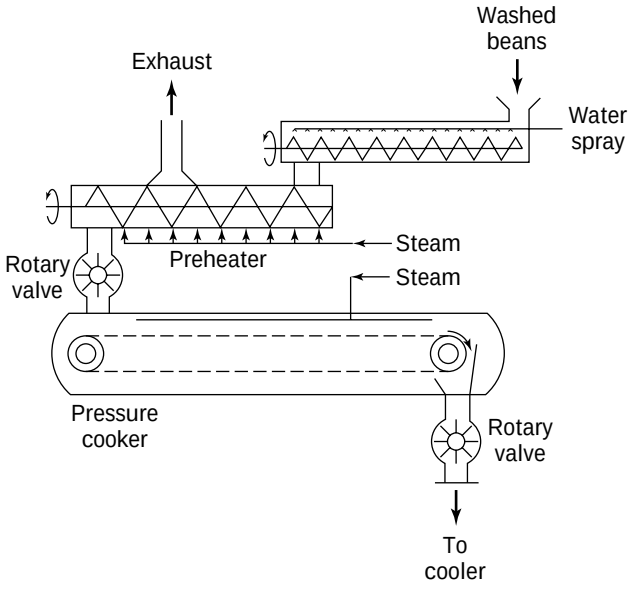


Figure 9. Continuous washing, soaking, preheating, and pressure cooking of soybeans.

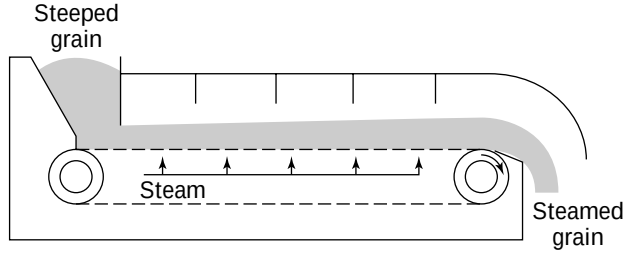


Figure 10. Continuous grain steamer.

the original rice, corresponding to about 2,000 fungal spores for each rice grain (10). The koji used is white because mold growth is arrested before sporulation (10). The unsporulated koji is relatively dry; the rice kernels are relatively hard and unclumped (10). The koji is then used in preparing the sake yeast starter mash, or *moto*. The traditional method of moto making required about a month-long fermentation of koji (300 kg) and water (1 m^3) (10). A succession of microbial populations developed, including lactic acid bacteria that reduced the pH and created an environment in which only the sake yeast, *Saccharomyces cerevisiae*, thrived (10). In modern moto production, fermentation time is reduced by adding lactic acid to water (pH 3.5–3.8) before mixing with koji. A large inoculum of sake yeast and steamed rice is also added. In a few days the starch is hydrolyzed, and fermentation begins. The next step is a *moromi* fermentation; it involves further addition of steamed rice, koji, and water to moto. Three additions are made at 1-day intervals (10). Temperature is lowered at each addition to control contamination because the added materials dilute the acids and alcohol. The temperature of moromi is controlled carefully, and after the final addition the temperature is 7–8 °C (10). Saccharification of starch and fermentation occur simultaneously in moromi (unlike in moto) until the alcohol content reaches nearly 20% v/v (10). Some steamed rice may be added during the final stages of fermentation to produce a sweeter product (10). Once the fermentation has almost ceased, alcohol is added to the mash to bring the concentration to 20–22% by volume (10). Solids are filtered out; the liquid product is further clarified, pasteurized, aged, blended, diluted to 15.0–16.5% v/v alcohol, and packaged (10). Unlike Western beer fermentations, sake mash is very high in solids that help to keep a large amount of yeast in suspension (10).

Beverages similar to sake are also found in China and many other Asian countries (Table 4). Other quite different indigenous alcoholic beverages, some made through solid substrate fermentation, are found in Africa and elsewhere (76). Some of these products are now produced in large industrial operations (77).

Soy Sauce

Soy sauce originated in China, but it is widely produced and consumed throughout Southeast Asia, Japan, and, increasingly, Europe and North America. The history, variety, and production technology of soy sauce have been detailed by Fukushima (9) and others (40,41,76,79). Here, a short description of only the latest, large-scale Japanese production methods is given.

Manufacture of soy sauce (*shoyu* in Japanese) involves a solid substrate koji fermentation, a brine or *moromi* fermentation, and refining steps (40,41,79). In koji making, soybeans or, more usually, defatted soybean flakes or grits are moistened and cooked in continuous pressure cookers (Fig. 9). The latest cooking methods employ a cooking time of 0.25 min or less, at approximately 7 kg cm^{-2} gauge pressure, equivalent to 170 °C (9). Rapid high-temperature cooking reduces loss of proteins and improves the yield of hydrolyzed amino acids in the final product. Cooked beans are mixed with roasted wheat that has been cracked into four or five pieces (9). Wheat roasters operate in a continuous mode at 170–180 °C (9). The ratio of wheat to beans is commonly 1:1, but it varies with the variety of shoyu (9). The mixed substrate is inoculated with a pure culture of *Aspergillus oryzae* (or *A. sojae*). The fungal spore density at inoculation is about 2.5×10^8/kg of wet solids (1). After a 3-day fermentation the substrate mass becomes green-yellow because of sporulation (9). Too long an incubation causes excessive, unwanted sporulation, production of ammonia from deamination of amino acids, and off-flavors in the final product (41). Koji is now harvested for use in the moromi step. Koji production is highly automated and continuous (9,18). Among the largest of koji-making processes is that of Kikkoman Corporation; it produces up to 4,150 kg h^{-1} koji, which is equivalent to processing 1,650 kg h^{-1} raw soybeans and 1,775 kg h^{-1} wheat (9).

Harvested koji is mixed with brine containing at least 20% salt for the moromi fermentation step (9,41). Large outdoor tanks are used for moromi fermentation. A specially selected culture of lactic acid bacteria is added to moromi mash, and the temperature is kept relatively low (15–20 °C) for the first month or so (9). Low temperature

prevents a too rapid growth and attendant rapid decline in pH. A pH drop that is too rapid causes undesirable early arrest of the proteolytic activity of koji enzymes; action of koji glutaminase is also obstructed (9). When the pH of moromi has declined to pH 5, a separately grown culture of the yeast *Zygosaccharomyces rouxii* is added (9). The temperature is then gradually raised to nearly 30 °C until a vigorous alcoholic fermentation occurs (9). On cessation of the alcoholic fermentation, the temperature is again reduced to about 25 °C and maintained there for the last 2 months (9). Aged moromi is filtered using automatic filter presses; the liquor is pasteurized in continuous flow plate heat exchangers and packaged in plastic bottles (9). The production schemes vary a great deal within Japan, and many quite primitive methods are encountered elsewhere in Asia. Chinese practices have been outlined by Wood and Min (41).

Tempe

Tempe (or tempeh) kedele is a fermented soybean product consumed in Indonesia, Malaysia, Singapore, the Netherlands, the West Indies, and, increasingly, the United States and Canada (8,40,41,76). Tempe kedele is made from partly cooked soybeans that are fermented with the mold *Rhizopus oligosporus*. The mold mycelium completely covers, penetrates, and binds the beans into a cake. In addition to the beans, lower-quality tempe varieties may contain solids such as cassava grits, soybean coats, or tofu that are added before fermentation. Freshly fermented tempe is creamy white; it has a mushroomlike flavor that becomes nutty and peppery on deep frying (8,76). Raw tempe does not keep well. Unless consumed within 24–48 h, it should be deep fried, dried, or blanched and refrigerated (76).

Production processes vary a great deal. In typical cottage-level processing in Indonesia, cleaned dry soybeans are washed and partially cooked by boiling for 2.5–3 h. An overnight soaking in clean water follows (8). A bacterial acid fermentation occurs and the pH drops to 4.5–5.3. The beans are now dehulled by rubbing with hands or feet. Hulls float on top and are washed away. Well-drained beans are inoculated using a previous batch of sporulated tempe that may have been sun dried. Alternatively, mold grown and dried on large leaves of *Hibiscus tiliaceus* may be used (8,76). Inoculum densities of 10^7 spores (or 10^7 cfu) per kilogram of wet solids have been reported (1,8). Overinoculation may produce excessive heat that inactivates fungal mycelium; *Bacillus* species proliferate, causing spoilage (8). Similarly, suboptimal inoculum density may cause outgrowth of yeasts and bacteria and again spoilage (8). Inoculated beans are lightly packed in relatively flat packets of *Hibiscus*, banana, or other broadleaves for 30–40 h of fermentation at ambient conditions (76). Similar general schemes are followed in large-scale production processes operated in the United States, Asia, and other countries; exceptions are mechanized processing, better-defined inoculum, and fermentation in perforated stainless steel trays (2- to 3-cm-deep bean layer) or perforated plastic bags (76). Fermentation is nonsterile; the bacterium *Klebsiella pneumoniae* that is invariably present in commercial preparations leads to vitamin B_{12} enrichment of the product (76).

Studies on production of *R. oligosporus* spores (pure culture) for commercial tempe inoculum indicate that viable spore counts of 10^9–10^{10}/kg dry substrate can be obtained on a variety of grains (1). Spores may be freeze-dried, but typically about one-log reduction in viability accompanies drying (1). Freeze-dried spores generally keep for 6 months (1).

Other tempelike products are consumed in Indonesia, including *tempe gembus*, made from waste produced by soybean curd factories; *tempe benguk*, produced using the legume *Mucuna pruriens* instead of soybeans; *tempe mata kedele*, made from soybean hulls; and *tempe bongkrek*, made from coconut press cake. The latter type can be potentially contaminated with the bacterium *Pseudomonas cocovenenans*, which produces lethal toxins (76). *P. cocovenenans* suppresses the growth of the tempe mold; hence, tempe bongkrek must never be consumed if a luxuriant growth of the mold has not developed (76). Formation of bacterial toxins is suppressed below pH 6, and an incubation temperature of 37 °C or higher prevents development of *P. cocovenenans* (76). The mold that is primarily responsible for most tempe fermentations is *Rhizopus oligosporus*, but other *Rhizopus* species, such as *R. oryzae, R. achlamydosporus*, and *R. arrhizus*, as well as non-*Rhizopus* fungi, may be involved (8).

Oncom

Oncom (or ontjom) is another tempelike product made from peanut press cake in Indonesia (41,76). *Rhizopus* species are used to produce the black oncom (*oncom hitam*), whereas the preferred red oncom (*oncom merah*) is produced using *Neurospora* species (76). *Neurospora sitophila*, *N. crassa*, and *N. intermedia* are typical oncom molds (76). In quite different processes, *N. sitophila* has been investigated also for protein enrichment of low-grade substrates such as sugar cane bagasse for animal feed (11).

Sausages

Fermented meat products such as bologna, summer sausage, thuringer, teewurst, dry sausage, pepperoni, and salami are produced by solid-state anaerobic fermentations utilizing acid-forming bacteria, particularly *Lactobacillus*, *Pediococcus*, and *Micrococcus* species. Typically, ground meat is mixed with curing agents and a starter culture at low temperature (-1 °C or less), and placed in cloth or other casings. Fermentation occurs at 20–40 °C. In dry sausage production in the United States, for example, a fast acid fermentation is employed to prevent growth of spoilage microorganisms. The curing mixture is predominantly salt and nitrite. Nitrite reacts with myoglobin in meat to produce the characteristic dark-red color of the product. Either *Pediococcus acidilactici* or *P. pentosaceus* is used for starter, depending on the selected fermentation temperature. Slower fermentation at 24 °C is commonly used in Europe. The curing mixture consists of salt, nitrate, and other spices. Naturally present micrococci and staphylococci convert the nitrate to nitrite for the desired coloring effect. After the lactic acid fermentation, a further incu-

bation period of up to 7 days at low temperature is practiced, with the meat held in 0.2-m-deep layers in metal pans.

Coffee

Coffee, after oil, is the second largest traded commodity, and more than 1.2 billion cups are consumed daily (83). Coffee processing employs a solid substrate fermentation for removing adhering mucilage (mesocarp) from ripe, mechanically pulped coffee cherry. This method is used predominantly only for *Coffea arabica* (84), which supplies the high-quality end of the market (83). Apparently, fermentation is not essential to flavor development, serving merely to remove the mesocarp from the parchment that encloses the bean. Nevertheless, the fermentation needs to be carefully conducted, or off-flavors develop that affect the final cup quality (84). In the preferred mode of fermentation, pulped cherry is placed in concrete or wooden tanks arranged for continuous drainage of the liquors produced (84). An alternative fermentation scheme utilizes under water fermentation with or without continuous water flow (84). The fermentation lasts 20–100 h, depending on factors such as pH, the coffee variety and stage of ripeness, aeration, and concentration of ions (for example, Ca^{2+} enhances the activity of certain enzymes) (84). Once the mucilage has degraded, the solids are washed and graded by water in concrete channels, and dried to produce parchment coffee (84). Milling (dehulling) removes the parchment to yield green beans used in commerce. The well-known coffee grounds are obtained by crushing blended and roasted green beans. Further processing is required to produce the instant varieties of coffee (83).

Cocoa

Cocoa beverages and chocolate are made from fermented cocoa (*Theobroma cacao*) seed or beans (85). Mature cocoa fruits or pods are harvested and split open to release the pulp-enveloped seeds. The pulp ferments as the seeds are piled in 2- to 4-ft-diameter heaps placed on perforated banana leaves that allow drainage of the liquids released (14,85). Alternative fermentation schemes employ perforated wooden boxes and other methods (85,86). The fermentation usually lasts 2–8 days, depending on the cocoa variety. The cocoa mass is turned several times during fermentation to ensure aeration and homogeneity and to avoid clumping (85,86). Turning also discourages mold growth on the better-aerated, exposed parts (86). A succession of microorganisms develop, including yeasts that are followed in turn by lactic acid producers, acetic acid bacteria, and, finally, aerobic spore-forming bacteria (86). The heat developed during the two-stage anaerobic–aerobic fermentation kills the cocoa embryo. Fermentation degrades the pulp, making removal easier, and products of fermentation penetrate the beans, stimulating complex biochemical reactions that are essential to flavor development (85,86). Fermented beans are dried, cleaned, and sold to chocolate processors or marketing organizations.

Mushrooms

The variety of edible mushrooms is tremendous, but only a few types are commercially cultivated. Production of mushrooms, mushroom compost and other substrates, and mushroom spawn utilize solid substrate fermentation methods that are comprehensively described elsewhere (2–4).

SAFETY CONSIDERATIONS

Microorganisms used in solid substrate fermentations have the potential to produce harmful effects in humans and animals. The history of occupational disease in solid-state cultivation is reflected in names such as cheesemaker's lung and mushroom worker's lung (87). Allergenic reactions and immunological hypersensitivity have been associated with exposure to fungal spores, including those of *Aspergillus niger, Aspergillus oryzae, Aspergillus awamori,* and *Aspergillus batatae* in industrial settings (88,89). Conditions such as dermatitis, asthma, vasomotor rhinitis, and itchy skin have been observed among exposed workers (87,88). Exposure to *Bacillus thuringiensis* has produced similar effects in humans (88). *Aspergillus* and *Pseudomonas* species have caused fatal infection in immunocompromised individuals and serious lung infection in cystic fibrosis patients (89). Fatal infections with *Aspergillus fumigatus* in previously healthy adults have been observed (89).

Inhalation of spores of edible mushrooms has long been associated with disease among mushroom growers (87). Allergenic reactions to spores of *Pleurotus* species, *Lentinus edodes* (3), and others have been documented. *Pleurotus* spores survive for several days after inhalation in the human respiratory tract, suggesting the potential for germination in the body (3). Allergenic reactions encountered in cultivation of *Agricus* species are linked to other molds that inhabit the substrate (3). In best-practice cultivation, attention to handling and processing of mushroom substrate and product should accompany ventilation of work areas and the use of personnel protective clothing.

In addition to microbes, fermentation products such as antibiotics and enzymes have been implicated in occupational asthma (87,89). Severe allergenic reactions to *Bacillus subtilis* proteases are well known (89). Under certain environmental conditions, organisms such as *Aspergillus flavus* and *Aspergillus oryzae* are known to produce lethal toxins (89). Specific strains of the blue-veined cheese mold, *Penicillium roqueforti*, also produce mycotoxins under narrowly defined environmental conditions (7). Species of the genus *Claviceps* and some members of other genera produce toxic ergot alkaloids. Some other toxin-producing fungi are noted in Table 5. Potentially, most physiologically active fermentation products can be disruptive to health, and certain products are highly toxic. Aflatoxins are potent carcinogens. For a given fungus, the product spectrum often depends on the fermentation conditions (temperature; pH; moisture content; type of substrate; agitation; metabolic energy source; amounts of oxygen and carbon dioxide; and nature and concentration of micronutrients, metal ions, and other chemicals); hence, evaluations under relevant conditions are essential. In particular, information from submerged culture may be an unsatisfactory indicator of behavior in solid-state fermentation.

Table 5. Toxin-Producing Fungi

Aspergillus alliaceus
Aspergillus clavatus
Aspergillus melleus
Aspergillus niger
Aspergillus ochraceus
Aspergillus parasiticus
Aspergillus sclerotiorum
Aspergillus sulphureus
Chaetomium cellulolyticum
Fusarium moniliforme
Fusarium graminearum
Penicillium islandicum
Penicillium roqueforti
Penicillium viridicatum

Handling of the substrate used in solid substrate fermentation requires care. A seemingly innocuous substrate may be contaminated with microbial spores or may generate allergenic dust. Raw material for composting may be heavily contaminated with human and animal pathogens, including eggs of parasites. Hence, certain high temperatures and holding time combinations are essential in composting to ensure destruction of pathogens and parasites. Depending on the composting process and the end use of the product, a uniform maximum temperature of 55 °C maintained for 3–15 days may be considered acceptable (33). Special care in process design and operation may be necessary to prevent recontamination of the final product. Cases of such contamination do occur perennially (e.g., infection from vegetables fertilized with manure; *Listeria* contamination in cheese). Sometimes inadequacies in processing may lead to a dangerous product: cases of cyanide poisoning from consumption of insufficiently processed *gari* occur from time to time in West Africa. Normally, fermentation and other process steps in gari making should substantially eliminate the cyanogenic glycosides (mostly linamarin and lotaustralin) that occur naturally in cassava tubers. Cases of fatal poisonings from consumption of poorly prepared tempe bongkerk have occurred in Indonesia.

Poor operational practice and failings in process and plant design can exacerbate the hazard. Experience suggests the following guidelines for minimizing process risks: (*1*) the number of operators should be kept to a minimum through mechanization and automation; (*2*) personnel–product contact should be minimal to reduce contamination of the product and ensure safety of operators; (*3*) operations should be enclosed and should be designed to minimize generation of aerosols and particulates; (*4*) the applicable GMP standards for food and pharmaceutical products should be followed; (*5*) the environment in and around the process plant must meet acceptable workplace air quality standards, especially for particulates, microorganisms, spores, humidity, and temperature; (*6*) the process equipment and the facility (including the air conditioning system) must be frequently sanitized; (*7*) facility design and operation must eliminate the risk of recontaminating the processed product; (*8*) there should be checks on the quality of the in-process material and the final product; and (*9*) satisfactory disposal of solid, liquid, and gaseous wastes must be ensured.

Operators should be fully informed of risks, and they should be given the training necessary for safe processing. Where required, personnel protective equipment should be provided. There should be adequate supervision, including routine health surveillance. Personnel who are ill, and those with open sores and cuts, should be sent home or moved to other, noncritical work. Attention to protection of peripheral support staff, for example, maintenance and cleaning personnel, is especially important because they may not have the knowledge or training for the potential risks. Thus, for example, trained, regular process operators should clean and sterilize a bioreactor before maintenance work.

The solid substrate fermentation facility should be designed and operated to prevent cross-contamination between dirty and clean areas. This goal requires attention to plant layout; movement of personnel, in-process material, process machinery; and flow of air in the facility. A minimum 12-Pa (0.05-in water gauge) pressure differential is recommended between adjacent clean and dirty areas. Negative pressure with respect to surroundings should be maintained in most solid-state fermentation chambers. Alternatively, the fermentation chamber may be under positive pressure with respect to the adjacent area, but that area should be under negative pressure with respect to other surrounding areas. Most work areas should comply with a relative humidity standard of 40–50%. Relative humidities lower than 40% cause problems with static electricity, whereas values greater than 50% promote corrosion. Between 10 and 15 air changes per hour are recommended for most plant areas. Workspaces for substrate preparation, downstream processing, and fermentation should attain 15–20 air changes per hour. Generally, conditioned air should be supplied at the ceiling level and return at lower wall locations. Preferably, particle counts should not exceed 3.5×10^6 m^{-3} (10^5 per cubic foot) for particles 0.5 μm or larger, and 2.5×10^4 m^{-3} (700 per cubic foot) for particles 5.0 μm or larger. There should be no more than 25 cfu/10 ft^3 of air. Air entering and exiting the fermentation chamber should be filter sterilized with 0.45-μm (or better) absolute filters that are also capable of in situ steam sterilization. A filter maintenance and replacement program should be in place. The practice of sterilizing solid substrate aeration gas by bubbling it through caustic or acid baths is misguided. Admittedly, the standards recommended here are not intended for processes such as composting and mushroom cultivation, but for plants producing foods, food additives, bulk pharmaceuticals, and secondary metabolites. The recommended practices are based on the author's experience. Note that the standards for packaging, sterile processing, and final dosage pharmaceuticals processing areas are significantly higher than the ones given here (90,91).

Much of the indigenous fermented food production in Asia, Africa, and South America is carried out at the cottage- or village-scale level with little consideration for hygienic processing. With few exceptions, even the larger plants do not come close to meeting the minimum standards expected in the developed world. The practices rec-

ommended in this review are for best-case production processes in the developed world where attention to hygiene and safety have transformed the traditional processing methods. Finally, as examples of toxic gari and tempe bongkerk prove, a thorough understanding of the fermentation microbiology and biochemistry is essential if disasters are to be prevented (78). Such knowledge is indispensable to developing large-scale production processes and in gaining increased acceptance of products beyond the regions of origin.

BIBLIOGRAPHY

1. R.E. Mudgett, in A.L. Demain and N.A. Solomon eds., *Manual of Industrial Microbiology and Biotechnology*, American Soc. for Microbiology, Washington, D.C., 1986, pp. 66–83.
2. T.E. Tautorus, *Adv. Biotechnol. Processes* **5**, 227–273 (1985).
3. F. Zadrazil, D. Ostermann, and G.D. Compare, in H.W. Doelle, D.A. Mitchell, and C.E. Rolz eds., *Solid Substrate Cultivation*, Elsevier, London, 1992, pp. 283–319.
4. W.A. Hayes and N.G. Nair, in J.E. Smith and D.R. Berry eds., *The Filamentous Fungi*, vol. 1, E. Arnold, London, 1975, pp. 212–248.
5. W. Röcken and P.A. Voysey, *J. Appl. Bacteriol. Symp. Suppl.* **79**, 38S–48S (1995).
6. A.H. Rose, ed., *Fermented Foods*, Academic Press, London, 1982.
7. B.A. Law, in A.H. Rose ed., *Fermented Foods*, Academic Press, London, 1982, pp. 147–198.
8. F.M. Rombouts and M.J.R. Nout, *J. Appl. Bacteriol. Symp. Suppl.* **79**, 108S–117S (1995).
9. D. Fukushima, in K.H. Steinkraus ed., *Industrialization of Indigenous Fermented Foods*, Dekker, New York, 1989, pp. 1–88.
10. K. Yoshizawa and T. Ishikawa, in K.H. Steinkraus ed., *Industrialization of Indigenous Fermented Foods*, Dekker, New York, 1989, pp. 127–168.
11. M. Moo-Young, Y. Chisti, and D. Vlach, *Biotechnol. Adv.* **11**, 469–479 (1993).
12. U.C. Banerjee, Y. Chisti, and M. Moo-Young, *Resour. Conserv. Recycl.* **13**, 139–146 (1995).
13. L.M. Miall, in J.E. Smith and D.R. Berry eds., *The Filamentous Fungi*, vol. 1, E. Arnold, London, 1975, pp. 104–121.
14. K.E. Aidoo, R. Hendry, and B.J.B. Wood, *Adv. Appl. Microbiol.* **28**, 201–237 (1982).
15. K. Bosecker and M. Kürsten, *Process Biochem.* **13**, 2–4 (1978).
16. E. Cannel and M. Moo-Young, *Process Biochem.* **15**, 2–7 (1980).
17. L.A. Underkofler, *Chem. Eng. Prog. Symp. Ser. No. 69* **62**, 11–20 (1966).
18. H. Ebine, in K.H. Steinkraus ed., *Industrialization of Indigenous Fermented Foods*, Dekker, New York, 1989, pp. 89–126.
19. B.K. Lonsane, N.P. Ghildyal, S. Budiatman, and S.V. Ramakrishna, *Enzyme Microb. Technol.* **7**, 258–265 (1985).
20. P. Weiland, in F. Zadrazil and P. Reiniger eds., *Treatment of Lignocellulosics with White Rot Fungi*, Elsevier, London, 1988, pp. 64–76.
21. E. Cannel and M. Moo-Young, *Process Biochem.* **15**, 24–28 (1980).
22. M. Moo-Young, A.R. Moreira, and R.P. Tengerdy, in J.E. Smith, D.R. Berry, and B. Kristiansen eds., *The Filamentous Fungi*, vol. 4, E. Arnold, London, 1983, pp. 117–144.
23. C.W. Hesseltine, *Process Biochem.* **12**, 29–32 (1977).
24. R.P. Tengerdy, *Trends Biotechnol.* **3**, 96–99 (1985).
25. A. Durand and D. Chereau, *Biotechnol. Bioeng.* **31**, 476–486 (1988).
26. Y. Chisti, *Airlift Bioreactors*, Elsevier, London, 1989.
27. Y. Chisti and M. Moo-Young, in V. Moses and R.E. Cape eds., *Biotechnology: The Science and the Business*, Harwood, New York, 1991, pp. 167–209.
28. Y. Chisti and M. Moo-Young, *Trans. I. Chem. E.* **74A**, 575–583 (1996).
29. Y. Chisti, *Chem. Eng. Prog.* **88**, 55–58 (1992).
30. Y. Chisti and M. Moo-Young, *J. Ind. Microbiol.* **13**, 201–207 (1994).
31. B.K. Lonsane, G. Saucedo-Castaneda, M. Raimbault, S. Roussos, G. Viniegra-Gonzalez, N.P. Ghildyal, M. Ramakrishna, and M.M. Krishnaiah, *Process Biochem.* **27**, 259–273 (1992).
32. B.A. Prior, J.C. Du Preez, and P.W. Rein, in H.W. Doelle, D.A. Mitchell, and C.E. Rolz eds., *Solid Substrate Cultivation*, Elsevier, London, 1992, pp. 65–85.
33. E.I. Stentiford and C.M. Dodds, in H.W. Doelle, D.A. Mitchell, and C.E. Rolz eds., *Solid Substrate Cultivation*, Elsevier, London, 1992, pp. 211–246.
34. M. Gutiérrez-Rojas, S.A.H. Hosn, R. Auria, S. Revah, and E. Favela-Torres, *Process Biochem.* **31**, 363–369 (1996).
35. B.L. Rathbun and M.L. Shuler, *Biotechnol. Bioeng.* **25**, 929–938 (1983).
36. S. Solis-Pereyra, E. Favela-Torres, M. Gutiérrez-Rojas, S. Roussos, G. Saucedo-Castañeda, P. Gunasekaran, and G. Viniegra-González, *World J. Microbiol. Biotechnol.* **12**, 257–260 (1996).
37. Y. Chisti, *Biotechnol. Adv.* **11**, 385 (1993).
38. A. Pandey, *Process Biochem.* **27**, 109–117 (1992).
39. P.W. Lambert, in J.E. Smith, D.R. Berry, and B. Kristiansen eds., *The Filamentous Fungi*, vol. 4, E. Arnold, London, 1983, pp. 210–237.
40. K.H. Steinkraus, in J.E. Smith, D.R. Berry, and B. Kristiansen eds., *The Filamentous Fungi*, vol. 4, E. Arnold, London, 1983, pp. 171–189.
41. B.J.B. Wood and Y.F. Min, in J.E. Smith, and D.R. Berry eds., *The Filamentous Fungi*, vol. 1, E. Arnold, London, 1975, pp. 265–280.
42. Y. Chisti and M. Moo-Young, *I. Chem. E. Symp. Ser.* **137**, 135–146 (1994).
43. K. Aunstrup, in L.B. Wingard, Jr., E. Katchalski-Katzir, and L. Goldstein eds., *Applied Biochemistry and Bioengineering*, vol. 2, Academic Press, New York, 1979, pp. 27–69.
44. B.K. Lonsane and N.P. Ghildyal, in H.W. Doelle, D.A. Mitchell, and C.E. Rolz eds., *Solid Substrate Cultivation*, Elsevier, London, 1992, pp. 191–209.
45. J. Pallares, S. Rodríguez, and A. Sanromán, *Bioprocess Eng.* **15**, 31–33 (1996).
46. C.W. Hesseltine, *Biotechnol. Bioeng.* **14**, 517–532 (1972).
47. A. Singh and P. Mishra, *Microbial Pentose Utilization: Current Applications in Biotechnology*, Elsevier, Amsterdam, 1995, pp. 71–98.
48. M.A. Millett, A.J. Baker, and L.D. Satter, *Biotechnol. Bioeng. Symp.* **5**, 193–219 (1975).
49. M. Becerra and M.I. González Siso, *Enzyme Microb. Technol.* **19**, 39–44 (1996).
50. V.K. Garg, M.A.C. Costello, and B.A. Czuba, in S. Seetharam and S.K. Sharma eds., *Purification and Analysis of Recombinant Proteins*, Dekker, New York, 1991, pp. 29–54.

51. V.R. Anicetti, B.A. Keyt, and W.S. Hancock, *Trends Biotechnol.* **7**, 342–349 (1989).
52. S.A. Rouf, M. Moo-Young, and Y. Chisti, *Biotechnol. Adv.* **14**, 239–266 (1996).
53. S.S. Yang and W.J. Swei, *World J. Microbiol. Biotechnol.* **12**, 43–46 (1996).
54. J. Barrios-González, T.E. Castillo, and A. Mejía, *Biotechnol. Adv.* **11**, 525–537 (1993).
55. C.W. Hesseltine, *Process Biochem.* **12**, 24–27 (1977).
56. P.K.R. Kumar and B.K. Lonsane, *Process Biochem.* **22**, 139–143 (1987).
57. P.K.R. Kumar and B.K. Lonsane, *Process Biochem.* **23**, 43–47 (1988).
58. E.V. Emelyanova, *Process Biochem.* **31**, 431–434 (1996).
59. R.W. Silman, *Biotechnol. Adv.* **11**, 559–560 (1993).
60. R.W. Silman, R.J. Bothast, and D.A. Schisler, *Biotechnol. Adv.* **11**, 561–575 (1993).
61. C. Desgranges, C. Vergoignan, A. Léréec, G. Riba, and A. Durand, *Biotechnol. Adv.* **11**, 577–587 (1993).
62. D. Leštan, M. Leštan, J.A. Chapelle, and R.T. Lamar, *J. Ind. Microbiol.* **16**, 286–294 (1996).
63. J.H. Crawford, *Process Biochem.* **18**, 14–18 (1983).
64. D. Augenstein, D.L. Wise, N.X. Dat, and N.D. Khien, *Resour. Conserv. Recycl.* **16**, 265–279 (1996).
65. M. Moo-Young and Y. Chisti, *Resour. Conserv. Recycl.* **11**, 13–24 (1994).
66. J.F. Rees and J.M. Grainger, *Process Biochem.* **17**, 41–44 (1982).
67. H.W. Doelle, D.A. Mitchell, and C.E. Rolz eds., *Solid Substrate Cultivation*, Elsevier, London, 1992.
68. D.R. Fenlon, A.R. Henderson, and J.A. Rooke, *J. Appl. Bacteriol. Symp. Suppl.* **79**, 118S–131S (1995).
69. D.R. Seale, *J. Appl. Bacteriol. Symp. Suppl.* **61**, 9S–26S (1986).
70. R. Whittenbury, *Process Biochem.*, 27–31 (February 1968).
71. J. Wignall and J. Tatterson, *Process Biochem.* **11**, 17–19 (1976).
72. P. McDonald, *The Biochemistry of Silage*, Wiley, Chichester, 1981.
73. M.K. Woolford, *The Silage Fermentation*, Dekker, New York, 1984.
74. B.K. Lonsane and M.M. Kriahnaiah, in H.W. Doelle, D.A. Mitchell, and C.E. Rolz eds., *Solid Substrate Cultivation*, Elsevier, London, 1992, pp. 147–171.
75. L. Ikasari and D.A. Mitchell, *Enzyme Microb. Technol.* **19**, 171–175 (1996).
76. K.H. Steinkraus ed., *Handbook of Indigenous Fermented Foods*, Dekker, New York, 1983.
77. K.H. Steinkraus ed., *Industrialization of Indigenous Fermented Foods*, Dekker, New York, 1989.
78. C.W. Hesseltine, *Process Biochem.* **16**, 2–6, 13 (1981).
79. B.J.B. Wood, in A.H. Rose ed., *Fermented Foods*, Academic Press, London, 1982, pp. 39–86.
80. N. Okafor and A.O. Ejiofor, *Process Biochem.* **25**, 82–86 (1990).
81. C.O. Ikediobi and E. Onyike, *Process Biochem.* **17**, 2–5 (1982).
82. S.H. Abiose, M.C. Allan, and B.J.B. Wood, *Adv. Appl. Microbiol.* **28**, 239–265 (1982).
83. M. Lockie, *Chem. Eng. (Lond.)* **615/616**, 22–25 (1996).
84. R.O. Arunga, in A.H. Rose ed., *Fermented Foods*, Academic Press, London, 1982, pp. 259–274.
85. J.G. Carr, in A.H. Rose ed., *Fermented Foods*, Academic Press, London, 1982, pp. 275–292.
86. R.F. Schwan, A.H. Rose, and R.G. Board, *J. Appl. Bacteriol. Symp. Suppl.* **79**, 96S–107S (1995).
87. B.K. Lonsane, N.P. Ghildyal, M. Ramakrishna, and F. Stutzenberger, in H.W. Doelle, D.A. Mitchell, and C.E. Rolz eds., *Solid Substrate Cultivation*, Elsevier, London, 1992, pp. 385–401.
88. A. Rimmington, in P. Hambleton, J. Melling, and T.T. Salusbury eds., *Biosafety in Industrial Biotechnology*, Chapman & Hall, London, 1994, pp. 67–89.
89. A.M. Bennett, in P. Hambleton, J. Melling, and T.T. Salusbury eds., *Biosafety in Industrial Biotechnology*, Chapman & Hall, London, 1994, pp. 109–128.
90. Y. Chisti, in G. Subramanian ed., *Bioseparations and Bioprocessing: A Handbook*, vol. 2, VCH, New York, 1998, pp. 3–300.
91. Y. Chisti, in G. Subramanian ed., *Bioseparations and Bioprocessing: A Handbook*, vol. 2, VCH, New York, 1998, pp. 379–415.

See also FOOD PROCESS ENGINEERING; MEMBRANE SURFACE LIQUID CULTURE, MICROORGANISMS, FUNGI.

SOLVENT FERMENTATION.

See ANAEROBES; ANAEROBES, INDUSTRIAL USES; CLOSTRIDIA, SOLVENT FERMENTATION.

SOYBEAN (FERMENTATION, MEAL OIL)

ZIVKO L. NIKOLOV
ROGER FUENTES-GRANADOS
Iowa State University
Ames, Iowa

KEY WORDS

Concentrates
Extraction
Fermented products
Flour
Isolates
Meal
Non-fermented products
Processing
Refining
Soybean oil

OUTLINE

INTRODUCTION

Soybean, *Glycine max*, is a crop native to Asia. It is not known when soybeans were introduced to the United States. Morse (1) reports that the first written mention of soybeans in the United States occurred in 1804, although there are references dating back to 1770, when Benjamin Franklin sent soybeans home from England (2). It is certain, however, that in the United States soybeans as a crop did not begin to gain popularity until the early 1920s. The primary reason for soybean's increased popularity was a newly developed ability to extract oil from it. Technological advances allowing utilization of soybeans for oil production led to the development of a large soybean industry. Today, the annual soybean crop value is about $14.5 billion. The advent of crushing facilities (continuous expellers) in the 1920s encouraged domestic demand for soybeans, and so for several decades since the 1920s, steady increases in the production and consumption of soybeans have occurred (Table 1) (3).

Economically, meal production today is more important than oil production. One metric ton (MT) of soybeans yields about 180 kg of oil and 800 kg of defatted meal. The unit price of oil is more than twice that of meal; however, a metric ton of soybeans produces meal and oil worth $160 and $80, respectively. The meal usually is the driving force behind soybean price because, on the average, it accounts for approximately two-thirds of the value of the beans (4). Soybean oil, the other direct product of soybean processing, is also important especially in the food industry. Soybean oil has 77% of the edible fats and oils market in the United States (4).

Most soy protein in the form of defatted meal is used as animal feed. Because of an ever-increasing population, however, demand for food protein has increased tremendously in recent years. Many types of soy food products have been developed, mostly from defatted soybean meal. These include soy flour and grits, soy protein concentrate, soy protein isolates, texturized soy proteins, defatted soy flakes, soy meal, full-fat soy flour, and enzyme-active soy flour. Advances in processing technology have resulted in soy protein products that can perform many functions in food while retaining their nutritive value. The development of technologies and processes has pioneered the way for soybean production and utilization. Improving existing and developing new ways to process and utilize soybeans will continue to be necessary if increasing demand for soybean products is to be met. This article reviews major soybean products and commercial processes currently in use.

Table 1. U.S. Production of Soybeans, Soybean Oil, and Total Crop Value from 1925 to 1995

Year	Total soybean production (million Bu.)	Total oil production (million lb)	Total crop value ($ million)
1925	4.9	2.5	11.4
1930	14	14	19
1935	49	105	36
1940	78	533	70
1945	193	1,392	402
1950	299	2,075	738
1955	374	2,827	831
1960	555	4,420	1,185
1965	846	5,800	2,151
1970	1,127	8,265	3,205
1975	1,548	9,630	7,618
1980	1,798	11,270	13,560
1985	2,099	11,617	10,571
1990	1,926	13,408	11,042
1995	2,152	15,155	14,564

Source: Soya Bluebook Plus, 1997 (7).

VARIETIES AND COMPOSITION

A great number of soybean varieties are grown in the United States. Most differ in agronomic characteristics, so maximum yield under diverse growing conditions can be achieved. Because chemical composition is not considered in the grading standards for trading soybeans, plant breeders have focused plant improvement programs on increasing yield. But in the last few years, new market opportunities for specialty soybeans have led to an increased release of soybean varieties with improved quality traits. Some quality characteristics recently developed in soybeans include lower amounts of linolenic and palmitic acids and higher levels of oleic acid in the oil and reduced or complete elimination of lipoxygenase and lower concentration of trypsin inhibitor in the meal.

Processing and utilization of soybeans have led to the creation of soybean varieties of two types: those processed for production of oil and meal and those processed for direct food use (5). Soybeans used directly for food are generally lighter in color, have clear hilum, and contain more protein and less oil than soybeans used for oil and meal do (2). The most desired characteristics of soybeans for food

include a large seed, high protein content, and high quality at a reasonable price (6). Typical dry weight composition of soybeans is 12 to 25% oil, 35 to 50% protein, 33 to 35% carbohydrate, and 5% ash (6).

PROCESSING SOYBEANS FOR OIL AND MEAL

Soybeans are the dominant oilseed crop among the seven major oilseeds grown in the United States. More than half of the U.S. soybean crop is used by the domestic crushing industry to produce meal and oil. More than 15 billion pounds of soybean oil were produced in 1996 (7). Improved understanding of lipid chemistry and enhanced oil processing technologies were the most vital factors in the acceptance of soybean oil as a food-grade product. Refined, bleached, and deodorized soybean oil is used by a variety of industries, including food processors and industrial concerns, by food service organizations, and by consumers (8). Modern soybean crushing for oil and meal production involves these unit operations: preparation and handling, cleaning, drying, cracking and dehulling, conditioning, flaking, extraction, and product desolventizing.

Preparation and Handling

Conventional preparation and handling of soybeans before extraction consists of cleaning, drying, cracking, dehulling, conditioning, and flaking (9).

Cleaning. Sand, stones, dust, and other foreign material usually are removed by conventional seed conditioners, which consist of a two-deck vibrating screen (5). The upper screen retains the stones and coarser materials but allows the passage of soybeans, and the lower screen retains the soybeans and filters out finer particles. Light trash and hull particles are removed by aspiration, and magnetic separators are used to remove tramp iron (5).

Drying. Drying is conducted to efficiently remove the hulls from cotyledons. Beans are dried to a moisture content of 10% or less. Soybeans are usually dried in countercurrent open-flame grain dryers heated with natural gas or fuel oil (10). Soybean temperature must be raised sufficiently to achieve the desired final moisture of 10%, but it must not exceed 76 °C because discoloration and protein denaturation will occur (10). Dried soybeans are stored in bins for 2 to 5 days of tempering to allow for moisture equilibration by diffusion. The tempering bins also serve as working bins for uninterrupted feeding of processing plants (5). Improvements have been made in reducing fuel and energy costs. Some relatively new drying concepts include microwave vacuum drying and solar drying (10).

Cracking and Dehulling. To break the soybeans into small pieces for dehulling and flaking, cracking is achieved by counter-rotating corrugated rollers (9). To save power and to obtain more uniform granulation with fewer fines in the cracked beans, the corrugation of the cracking rolls was changed some years ago by widening the pitch between the teeth of the corrugated rolls (11). This alteration resulted in cleaner cuts or shearing with less power consumption (11). Cracking rollers usually are 25 cm in diameter and at least 107 cm long, having a processing capacity as great as 500 to 600 MT/day (9). The cracking step is very important in traditional dehulling. The cotyledons of properly dried and cracked soybeans should separate easily from the hulls (12). If the soybeans are too dry, they will pulverize in the cracking rolls and produce excessive fines, which could cause problems in the extractor. If beans are too moist, they tend to mash and not to release the hulls (12).

Conditioning. Cotyledon fragments (meats) are conditioned by heat and moisture to obtain the plasticity required for flaking. The temperature of the meats at the end of conditioning is between 65 and 70 °C (9). Conditioners may be vertical stack cookers or rotary horizontal cookers in which steam or water spray is added to adjust moisture to approximately 11% (12). Effective preheating before flaking decreases power consumption of the flaker (11). Flaking cold beans (25–30 °C) requires a power of about 5 to 6 kWh/ton of soybeans; however, flaking at 55 to 60 °C requires only 3 to 4 kW h/ton (11).

On-line dryers have been developed that combine the drying, dehulling, and conditioning steps into a continuous operation. Soybeans are cracked hot, which decreases the amount of fines produced. Different systems have been designed for this combined operation. The Escher-Wyss system uses a fluid bed dryer-heater to heat and dry soybeans so that the hulls detach from the meat. This is followed by cracking rolls and impactors to break up the soybeans and loosen the hulls. Hulls then are separated by air aspiration and meats are sent to a fluidized-bed conditioner-cooler for final moisture and temperature adjustment (13).

Flaking. Flaking is achieved by passing the conditioned meats between two horizontal smooth rolls. Flaking mills consist of a pair of smooth surfaced rolls with a minimum diameter of 50 cm (12). Rolls usually are 120 cm long and 70 cm in diameter. Pressure between the rolls is generated by mechanical or hydraulic systems, and flake thickness is controlled by the pressure (12). Flaking before solvent extraction is important because (*1*) solvent can flow easier through a bed of flaked particles than through a bed of cotyledon fragments or fine particles and (*2*) pressure generated between the rolls ruptures oil cells and facilitates subsequent solvent extraction (9).

Thickness of the flakes is crucial because solvent and oil diffusion are the rate-limiting steps during hexane extraction. Optimal flake thickness is 0.025 to 0.037 cm; thickness outside this range does not yield satisfactory results. Thicker flakes offer greater diffusion resistance; thinner ones easily crumble, producing fines that hinder hexane penetration into the soy-flakes bed (9).

A common problem in flaking is greater wear at the center than at the ends of the rolls. As the wear progresses, soy flakes at the center of the rolls become thicker than those at the ends. Continued pressure adjustment to correct flake thickness near the center may cause excessive pressure at the roll ends and chipping or pitting of the metal (12).

Improvements in Pre-Extraction Technology. The use of expanders in preparing oilseeds for extraction has been one of the most significant developments in the oilseed industry (13). Expanders produce collets that improve drainage and reduce solvent carryover to the desolventizer toaster. The three main benefits of the expander in oil extraction are (*1*) as great as a 40% increase in bulk density, potentially as great as 40% increase in contact time; (*2*) increase of bed porosity resulting in more rapid drainage rates; and (*3*) additional rupture of oil cells, which substantially increases the oil diffusion rate (14).

Solvent Extraction

The most prevalent solvent extraction method in modern soybean processing facilities is hexane extraction (15). The oil extraction process is basically a process of solid–liquid extraction (5). The transfer of oil from solids to the surrounding oil–solvent solution may be described in three steps: (*1*) diffusion of solvent into the solids, (*2*) oil solubilization, and (*3*) diffusion of oil from the solid particles into the surrounding liquid phase.

The driving force of diffusion is the oil concentration gradient. Although hexane has been the solvent of choice in most commercial extraction plants, other solvents also have been considered (16,17). The properties of an ideal solvent have been discussed in detail by Johnson (17); the most important of these include high-solvent power, nontoxicity, triglyceride selectivity, nonflamability, low specific heat and heat of vaporization, and low cost. The two streams exiting from the solvent extractor are the miscella, or oil-hexane mixture, and the marc, or defatted flakes containing hexane (Fig. 1).

Throughout the world, mechanical extraction often is preferred by small extraction plants, particularly those processing a variety of oilseeds. Screw presses consist of a shaft with an interrupted worm gear rotating in a cage of metal bars separated by small spaces. High pressure forces the oil out between the cage bars as the press cake moves along the shaft. Mechanical extraction can be used for the cold pressing of oils and the production of meal with a high oil content. Advantages are low initial capital cost and no solvent requirement; the efficiency of oil extraction, however, is less than with other methods.

Extraction Equipment. The basic types of commercial oil extractors in use today fall into two main categories: percolation and immersion (15,18). Percolation types are used worldwide and are considered more efficient. Immersion extractors are not commercially used for soybean oil extraction (15).

Four types of percolation extractors are actively marketed at present: the rotary basket extractor (Rotocel), the stationary basket extractor, the horizontal belt extractor, and the continuous loop extractor (12). A description of each extractor is beyond the scope of this article, and an excellent review of commercial extractors has been given by Milligan (19). In percolation extractors, hexane or miscella is pumped over the bed of soy flakes, percolates through the bed, and leaves at the bottom through a perforated plate or mesh screen. Extractors can be classified as deep-bed ($>$ 5 m deep) and shallow-bed ($<$ 5 m deep) extractors. Deep-bed extractors (rotary or stationary basket extractors) have a material bed whose overall length is approximately five times greater than its depth (14). The bed surface area per miscella stage is relatively small, allowing the required miscella flow to permeate the bed completely. The solvent-flakes contact time is maximized, and these extractors can operate with relatively thick flakes. Shallow-bed extractors (horizontal belt and continuous loop extractors) are designed so that overall length of the bed material is 50 times greater than its depth. Bed surface area per miscella stage is large, and contact time between solvent and flakes is not maximized. Shallow-bed extractors operate effectively by using thin flakes, thereby increasing the diffusion rate (14).

Miscella Stripping. Full miscella contains about 30% oil; for every ton of crude oil, therefore, about 2.5 tons of solvent must be removed by distillation (5). The hexane must be efficiently removed and recycled to ensure the oil safety and to conform to environmental regulations (9).

A solvent recovery system consists of a two-stage stripping evaporator and an oil stripping column. Solvent is removed from the miscella by double-effect evaporators and steam stripping. The first stage usually consists of a long-tube vertical evaporator, where the miscella is heated and solvent evaporated using vapors from the desolventizer-toaster (DT) unit at below atmospheric pressure, concentrating the miscella to about 70% oil (12,15). The second-stage evaporator is a rising-film evaporator heated with steam at 10 to 20 psi in which the miscella is further concentrated to 90 to 95%. Vapors generated go to the condenser, and oil goes to the oil stripper, a disk-donut type operating under vacuum provided by steam injection (12,15). Oil flows downward in contact with stripping steam introduced at the bottom. At the end of this operation, the amount of residual solvent is sufficiently low to achieve a flash point of 150 °C. A flash point of 121 °C corresponds to 1,000 ppm of hexane in the crude oil, which is the required level for transfer to the refinery (9).

Desolventizing and Toasting Soybean Meal for Feed. Desolventizing is a key operation in solvent extraction (13). Defatted flakes from the extractor contain about 30% (w/w) hexane, and before they can be processed further for feed they need to be processed through a desolventizer. The meal, which is used usually as livestock feed, must be heated sufficiently to inactivate antinutritional factors such as trypsin inhibitors (9). This process is known as *toasting*. It is carried out commercially in conjunction with desolventizing in a DT unit (9). Hexane-wet flakes leave the extractor at temperatures of about 57 °C. Live steam is injected at the bottom of the DT unit, and the steam that condenses provides the latent heat required for hexane evaporation. The condensed steam raises the moisture level of the flakes to 16 to 24%, facilitating the toasting operation at 100 to 105 °C (20). The residual hexane content of the flakes is less than 500 ppm. The flakes then are hot-air dried to about 10% moisture, cooled to a temperature below 32 °C, and ground and screened to a desired grit size.

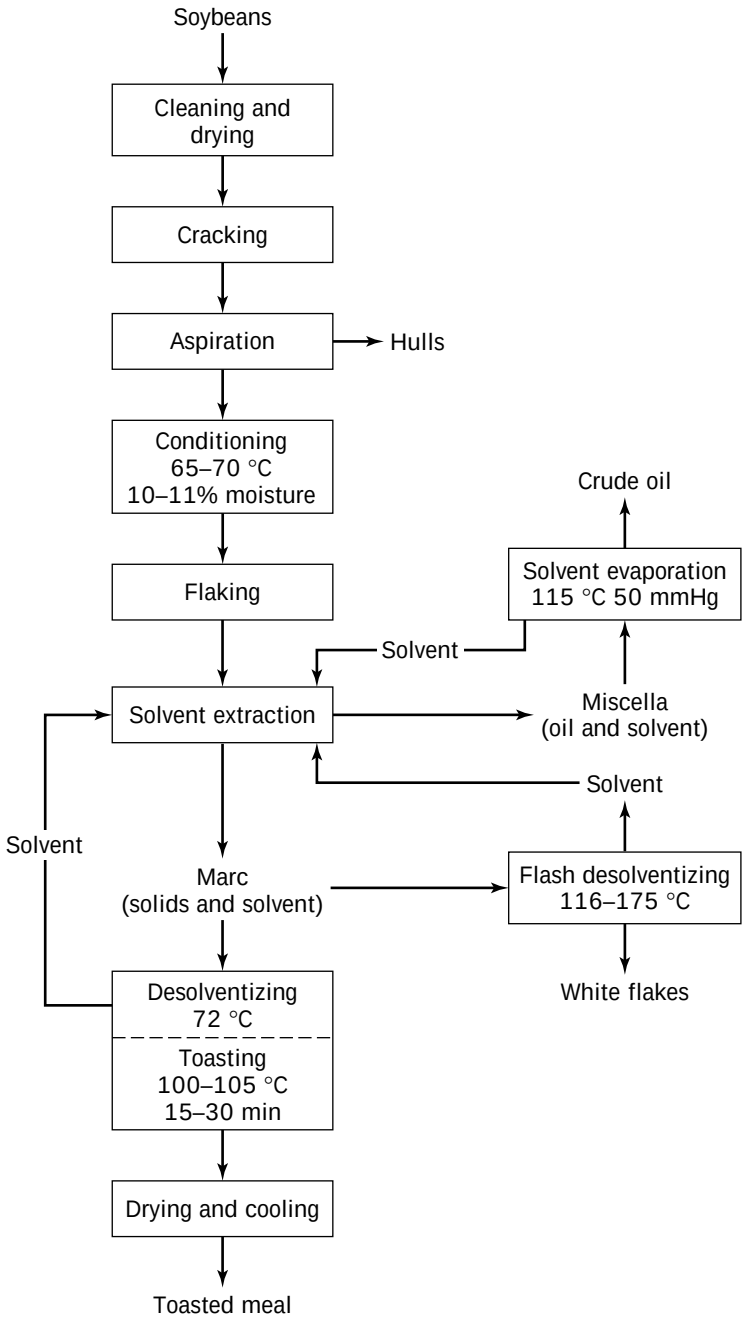


Figure 1. Principal unit operations in soybean oil extraction facilities.

Soybean meal with greater than 50% (w/w) protein, which is produced by grinding toasted flakes, provides greater than 60% of the value of processed products from soybeans. In the United States, meal is used primarily in swine and poultry rations (8). The other components of meal are carbohydrates (30 to 35%), lipids (1%), ash, and moisture.

A predesolventizing step and the use of expanders in flake preparation have marked effects on the desolventizing operation (20). Predesolventizing reduces the amount of hexane going into the DT and the amount of water condensed on the flakes and, consequently, decreases the steam requirements (20). Moisture reduction of the flakes, however, increases the possibility of overheating and decreases the nutritive value of the meal (20). Expanded flakes have higher bulk density and lower residual solvent content than unexpanded flakes do (9).

Desolventizing and Deodorizing Soy Flakes for Food. For food and industrial applications of meal, it is necessary to minimize the heat denaturation of soy proteins and to maintain their water solubility (9). Desolventizing to produce white flakes can be performed (*1*) in steam-jacketed conveyor tubes arranged in a vertical stack; (*2*) by using superheated circulating solvent vapor in a horizontal drum equipped with an agitator/conveyor pipe to facilitate con-

tact between the flakes and the superheated vapor; and (*3*) by using superheated solvent vapor to contact the flakes in a conveyor pipe, where the conveying action results from the velocity of the superheated vapor (20). The last process, also called *flash desolventizing* because of the relatively short residence time of flakes in the apparatus, is the best method for desolventizing white flakes (5). The superheated vapor provides energy for the evaporation of solvent from the flakes. The turbulent nature of the flake-vapor flow permits extremely rapid heat and mass transfer. In seconds, the superheated hexane will flash off the residual liquid hexane, leaving desolventized flakes to be collected in the cyclone (21). Protein denaturation is minimized primarily because of the short heating time. Hexane losses in white flake desolventizing systems are generally higher than those in DT systems (20). After desolventizing, white flakes are deodorized with live steam to remove the remaining hexane. Depending on end use, deodorization conditions can be adjusted to produce flakes with varying protein solubilities, such as with protein dispersibility indices (PDIs) of 10 to 90%.

One negative impact on the environment from the oilseed processing industry is the unrecovered solvent from the oil extraction (14). By November 15, 2000, maximum achievable control technology standards for hexane emissions will be established (23). What the standards will be is unknown; it will not, however, be easy to meet them, even by the most efficient plants (24). The most significant hexane loss points in the extraction plants include desolventizing, meal drying and cooling, vent systems, water and oil carrying solvents, fugitive emissions, and plant shutdowns (24). In most plants, hexane loss is less than half that of 10 years ago. This decrease was accomplished by emphasizing the basic engineering and processing principles upon which most plants were designed and built (24).

Storing and Handling Crude Soybean Oil. Crude soybean oil contains a significant amount of nontriglyceride components that should be removed before bulk storage, handling, and transport (25). Deterioration of oil in storage depends not only on external factors but also on the initial conditions of oil, including moisture content, free fatty acids content, organic impurities, and contamination with molds (25). Crude oil must be stored under specific conditions to avoid oxidation caused by moisture, temperature, air, and metals promoting oxidation (9). To avoid the settling of fines and gums, oil with a moisture content below 0.2% is cooled to 40 °C and stored in stainless steel tanks with agitation. Oil stability can be improved if impurities are removed before storage (9).

REFINING SOYBEAN OIL

Refining freshly extracted oil means less product loss and improved product quality (13). Crude oil contains both insoluble and oil-soluble impurities. Insoluble impurities include seed fragments, excess moisture, and sometimes a waxy fraction appearing in refrigerated oils (9). Oil-soluble impurities include phosphatides, free fatty acids, gummy or mucilaginous material, color bodies, tocopherols, sterols, hydrocarbons, ketones, and aldehydes. The refining process removes those impurities, converting the crude oil into food grade soybean oil (9). Oil refining consists of several major unit operations: degumming, neutralization, bleaching, hydrogenation, and deodorization, each of which is designed to remove one or more specific compounds (Fig. 2).

Crude Oil Clarification

The removal of fines from crude oil is done traditionally by filtration. In integrated facilities, filters are being replaced by centrifugal decanters (13). In facilities with acid and water degumming and when lecithin is not produced, the clarification step is combined with the degumming step by removing both solids and gums in a high-speed decanter (13).

Degumming. The degumming step removes the phospholipids, also referred to as *phosphatides*, *lecithin*, or *gums*, which are the major source of oil phosphorus. Removal of phospholipids is essential in the production of oil with optimal color, flavor, and oxidative stability. Direct mixing of warm oil (70 °C) and soft water (1 to 3%) is the most commonly used method for degumming. Hydrated and insolubilized phospholipids partition to the water phase, and the oil temperature disrupts the possible formation of emulsion (9). Degumming of the oil from undamaged soybeans decreases phosphorus content in crude oil from 500 to 900 ppm to 12 to 170 ppm (9). Soy phospholipids also are extracted, because they are used in the food industry as soy lecithin (a mixture of phosphatidylcholine, phosphatidylinositol, and phosphatidylethanolamine). Because of the limited market for lecithin, a common practice in the United States is to alkali refine crude soybean oil and to dispose of lecithin in soapstock (26). The next important development in the degumming process probably will be the implementation of membrane filtration (13).

Neutralization

Neutralization is the removal of free fatty acids from the oil. Free fatty acids must be removed because they reduce the smoke point of oil, increase foaming, and may impart soapy flavor in foods (9). There are several methods of neutralizing crude soybean oil, alkali refining, Zenith refining, and physical refining are commercially used.

Alkali Refining. Alkali refining, also called *caustic refining*, is the most common method of neutralization and can be performed on crude or degummed oil. It involves adding alkali solution, in-line mixing with oil, centrifugation, in-line washing with water, and a second centrifugation (27). The sodium salt of the fatty acid (soaps), which are produced during neutralization, are removed during centrifugation. Some loss of triglycerides may occur with soaps. A typical level of free fatty acids before refining is 0.3 to 0.7%, and refining lowers this to $<0.05\%$ (9).

Zenith Refining. The Zenith refining process was developed in Sweden in 1960 and consists of degumming with

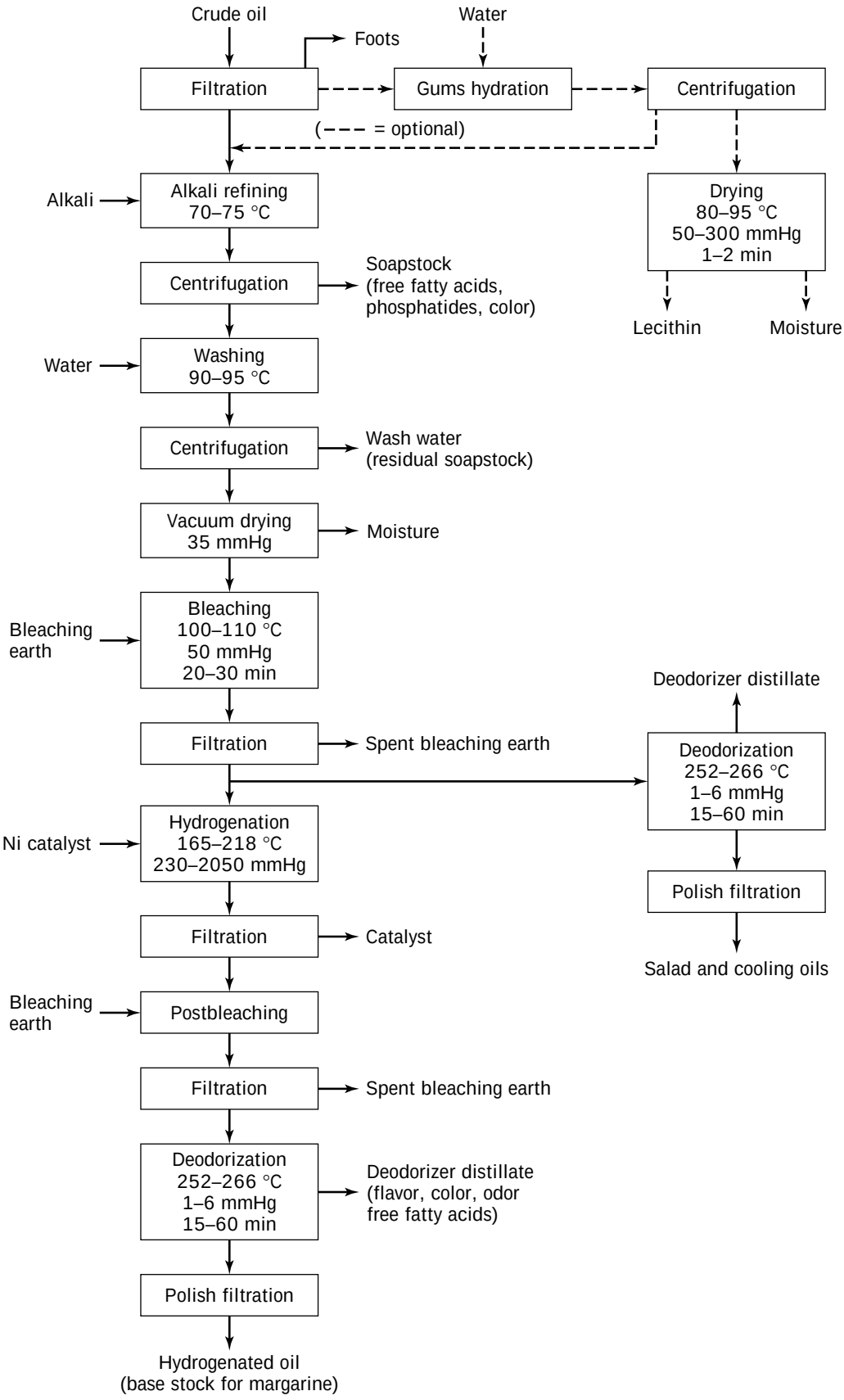


Figure 2. Unit operations involved in soybean oil refining for the production of hydrogenated oil, salad and cooking oils, and optional lecithin production.

0.1 to 0.4% (w/w) of phosphoric acid and alkali refining with a dilute sodium hydroxide solution (0.35 M). The use of dilute alkali obviates washing the oil with water. The neutralized oil contains about 1% water and 50 ppm soap and requires bleaching with citric acid and bleaching earth (27).

Physical/Steam Refining. This process is essentially a deodorization of lecithin-free oil under high vacuum at high temperatures. The process conditions are the same as those used in the deodorization. Physical refining reduces water usage and generates less wastewater than alkali refining does. Because of flavor instability, physically refined oils are not acceptable in the United States (27).

Bleaching

Treatment of refined oil with a bleaching adsorbent removes pigments, metal compounds, residual soap, and trace amounts of prooxidants that were not removed by alkali refining (15). Neutral earth, acid-activated earth, activated carbons, and silicates are used as adsorbents for soybean oil bleaching (28). This process has evolved basically from an open batch system to a continuous process under vacuum with alternating filters for clay removal (13). Typically, neutralized oil is mixed at about 70 °C with an appropriate dosage of earth (0.3 to 0.6%), heated to bleaching temperature (100 to 110 °C), mixed for 20 to 30 min, cooled to 70 to 80 °C, and filtered (28). Bleaching is an important step in enhancing oil appearance, flavor, and stability and in obtaining a light color appealing to the consumer (9).

Hydrogenation

The goal of hydrogenation is to decrease the amount of unsaturated fatty acids and to convert liquid oils to fluid, semisolids, or plastic fats (9,29). Hydrogenation is performed with hydrogen gas in the presence of a nickel catalyst. Reaction conditions differ with the desired product, and the ranges of process temperatures, pressures, and catalyst concentrations are 165 to 218 °C, 0.3 to 2.7 atm, and 0.02 to 0.05%, respectively (29). Hydrogenation increases the oxidative stability by reducing the number of double bonds in fatty acids, which are the sites of peroxide formation and subsequent off-flavor development (9).

Deodorization

The main purpose of deodorization is to remove free fatty acids, odors, and flavor compounds from the oil (13). Deodorization is a steam-stripping process in which good quality water is injected into soybean oil under a strong vacuum (1 to 6 mm Hg) and a temperature sufficiently high (252 to 266 °C) to vaporize undesirable volatile compounds (30). Deodorization is the last step in refining edible oils to be used as ingredients in salad dressings, cooking oil, margarines, shortenings, hard butters for the confectionery industry, and other food products (9).

Winterization

Winterization is the removal of solids that settle out from the oil at temperatures between 4 and 10 °C. The solid fraction (stearine) is separated so that it does not precipitate during refrigeration of salad oils (9). Winterization begins with dry, hydrogenated oil being pumped into a large insulated tank fitted with cooling coils. The temperature is lowered slowly by controlling the amount of propylene glycol circulating through the cooling coils. The rate of cooling must be controlled to promote the formation of larger crystals for easy separation. Crystal formation and growth is determined by nucleation temperature, triglyceride composition, and cooling rate. Commercial winterization is done by cooling the oil from 30 to 35 °C to about 6 °C. The temperature difference between oil and coolant is maintained at about 13 °C to ensure slow cooling. As the final temperature is approached, the cooling rate is decreased further. The cooling process lasts between 12 and 24 h (9). At the end of the cooling period, the external cooling is stopped and the winterized oil is allowed to stand for 6 to 8 h at 6 °C before filtration (15). Typical oil recovery is 75 to 80%.

Industrial Uses of Soybeans

Industrial (nonfood, nonfeed) uses of the two major soybean products, protein and oil, comprise less than 0.5% of total soy protein and about 2.6% of the total soy oil produced in the United States (31). Economically viable industrial uses today of soy protein and oil include paper coatings (protein), wood adhesives (protein), alkyd resins (oil), printing inks (oil), and oleochemicals. Potential uses include plastics and textile fibers from soy protein and soybean-based fuel. The most important current and potential applications of major soybean products are summarized in Table 2. For in-depth review of the industrial uses of soybeans, interested readers are referred to the work of Johnson and Myers (31).

SOYBEAN PROTEIN PRODUCTS

The term *soy protein* is given to processed edible dry soybean products with protein contents greater than 50% (w/w). Typically, food-grade soy proteins are produced from white flakes, which are processed on different lines than those used for producing oil and feed meal (32).

Defatted Soybean Grits and Flour

Grits are obtained by grinding defatted flakes coarsely followed by screening. Grits are sold as coarse (10 to 20 mesh), medium (20 to 40 mesh), or fine (40 to 80 mesh) grade. Soy flour is produced by grinding soy flakes to very fine particles to achieve 97% product pass through a 100-mesh screen. Soybean flour and grits are used in a variety of food products, including soups, sweets, beverages, desserts, bakery goods, breakfast cereals, and meat products.

Soy Protein Concentrates

Soy protein concentrates are prepared by removing soluble carbohydrates and volatile compounds from defatted meal or flour to obtain a product with a protein content of 65 to 75% (w/w). Nondenatured soybean protein is most soluble at pH values between 1.5 and 2.5 and between 7 and 12

Table 2. Current and Potential Industrial Uses of Soybean Products

Product	Main processes	Uses
Adhesives	Alkali dispersion of soy flour, mixing and treatment with other ingredients such as sodium silicate, antifoam, and preservatives	Wood adhesives
Plastics	Mixing of soybean meal with formaldehyde and phenol followed by heating	Potential use for foam cups, food packaging, and other fast-food industry items
Textile fibers	Treatment of soybean meal with alkaline solution and with acid to precipitate protein; dissolving, wet-spinning, recoagulating in reel, and stretching	Blending with other fibers such as cotton, rayon, nylon, wool
Paper coating	Alkali dispersion, mixing with additives	Decorative papers
Paints, varnishes, and coatings	Heating of soybean oil, and mixing with maleic anhydride; esterification of maleinized oil to a polyol	Alkyd resins
Plasticizers	Oxidation of soybean oil with performic acid, peracetic acid, or with hydrogen peroxide in the presence of a strong ion-exchange resin	To improve heat and light stability of plastics
Diesel fuel	Dissolving of 0.1–0.5% (based on oil) sodium methoxide in methanol, and mixing with heated soybean oil (80 °C); vacuum distilling, drying, and blending with diesel fuel, of the ester layer	Interest in biodiesel in cities with air pollution problems
Printing ink	Polymerization of oil by heat, blending the polymerized oil with 1% butilated hydroxitoluene and 6–20% pigments	Lithographic and letterpress newsprint applications

Source: Adapted from Johnson and Myers, 1995 (31).

and least soluble at its isoelectric region of pH 4.2–4.6 (32). Three basic processes have been used to remove carbohydrates (Fig. 3): acid extraction at pH 4.5, 60 to 80% aqueous ethanol extraction, and hot water extraction (33). In all three processes, protein remains insoluble while solubilized carbohydrates are removed by centrifugation. Solids from the acid extraction process are resuspended in water and neutralized to pH 6.8, whereas those from alcohol extraction are flash desolventized before spray drying. The most obvious difference between soybean protein concentrates is the lower ash content of concentrates prepared by acid or hot water extraction, indicating more thorough removal of minerals (32). Figure 3 summarizes basic unit operations and process conditions required to make soy protein concentrate by each of the three methods.

Soy Protein Isolates

Many processing options exist for making soy protein isolates. The preferred approach is to tailor the extraction process so that the isolate is made optimally compatible with the intended consumer product (32). Soy protein isolates are prepared traditionally from defatted soybean meal using aqueous or mild alkali extraction (pH 7 to 10) of proteins and soluble carbohydrates (33). The insoluble residue, mostly carbohydrates, is removed by centrifugation, and soluble protein is then precipitated at the isoelectric point (4.2 to 4.5) of soy proteins (Fig. 4). The precipitated protein is separated by mechanical decanting, washed, neutralized to about pH 6.8, and spray dried to obtain a product with protein content greater than 90% (w/w).

Textured Soy Protein

Textured soy protein products are obtained by thermoplastic extrusion of soy flour or concentrates to obtain a meatlike texture. Flours and concentrates are mixed with water and additives to form a dough and extruded under high pressure and temperature to obtain fibrous texture (33). Processes for textured proteins include thermoplastic extrusion, fiber spinning, direct steam texturization, shaping and heating, enzymatic texturization, and coagulation (34). Texturizing processes differ based on the type of raw protein material used, which can include defatted soy flour, soy protein concentrate, soy protein isolate, or a blend of several products (34). The most popular raw material has been slightly toasted, defatted soy flour.

Textured soy products have been produced with raw materials ranging in PDI from 20 to 70 (35). The PDI is the percentage of total protein dispersible in water under controlled conditions of extraction. During extrusion, the raw material is placed in the feeder and moved to a conditioning section where the food is mixed with 80 °C steam. In the conditioning section, the moisture content is raised to between 10 and 25%. The mixture then is sent to the extrusion cooker, where it is processed under conditions depending on type of raw material, equipment, and desired final product. Processing conditions include pressures of 400 to 1,500 psi and temperatures of 105 to 113 °C (35).

PROCESSING OF WHOLE SOYBEANS FOR FOOD

In the United States, only 3 to 5% of the soybean crop is used for food. This proportion is much larger in other countries, especially those in Asia, where soybeans are a staple of the human diet (8). In Japan alone, about 1 million tons

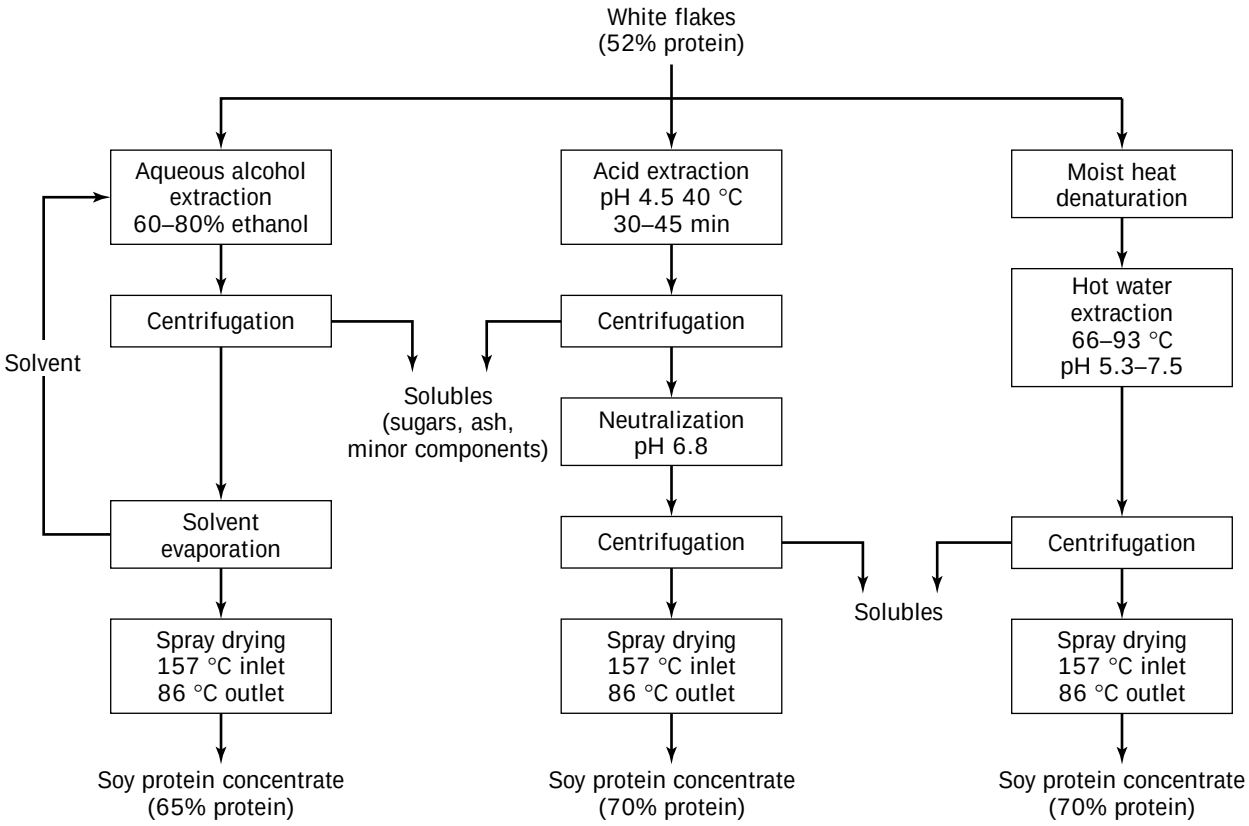


Figure 3. Alternative processes for the production of soy protein concentrates.

of soybeans are used to produce traditional Japanese foods such as tofu, shoyu, and miso (36). Both fermented and nonfermented foods are made from soybeans, and a brief discussion of most important products will be presented next.

Nonfermented Foods

A variety of nonfermented food products are made from whole soybeans. Raw soybeans have an undesirable bitterness and beany off-flavor and are rarely consumed without some sort of processing. In addition, raw soybeans contain several enzymes and antinutritional factors that need to be eliminated or inactivated if soybeans are to be palatable and digestible for humans (5,37). Growing interest in soybean foods in western countries has led to modernization of the manufacturing of traditional products and to the development of totally new products such as tofu-based ice cream and soybean yogurt (5). Soy milk, tofu, toasted full-fat soy flour, and enzyme-active soy flour are among the most important and best known commercial nonfermented food products.

Soy Milk. Soy milk is basically a water extract of soybeans obtained by wet grinding of soaked soybeans and subsequent filtering of solids (Fig. 5). Soy milk quality is related directly to raw soybean quality. Differences among soybean varieties affect not only yield but also important properties such as aroma and taste of soy milk. Soybean standards for soy milk production have been developed in Asian countries, and their use in the selection the raw soybeans have contributed to the consistently good quality of soy milk.

There are many different methods for producing soy milk, and the majority are variations of those described by Nelson et al. (38) (the Illinois method) and by Wilkens et al. (39) (the Cornell method). Both methods were developed primarily to improve soy milk flavor and stability. Regardless of process variations, the basic steps and principles of soy milk preparation are very similar and consist of soaking and grinding the beans and heating the extracted soy milk (Fig. 5). Modern large-scale soy milk plants use several additional steps and more sophisticated equipment. Several new processes have been described by Chen (40), Wilson (41), Kwok and Niranjan (42), and Liu (43).

Cleaning. Cleaning and washing, which remove dirt, dust, and microorganisms from the surface of soybeans, is crucial if contamination of the final product is to be avoided. The precleaning steps also include the removal of foreign material such as straws, stones, metal, weeds and their seeds, and damaged soybeans.

Dehulling. There are arguments in favor of and against dehulling. Most soy milk manufacturers use whole beans because dehulling disrupts the cotyledon tissue and releases the lipoxygenase enzymes, which diminishes the quality of soy milk. In addition, the presence of hulls helps filtration. Dehulling supporters believe that dehulling im-

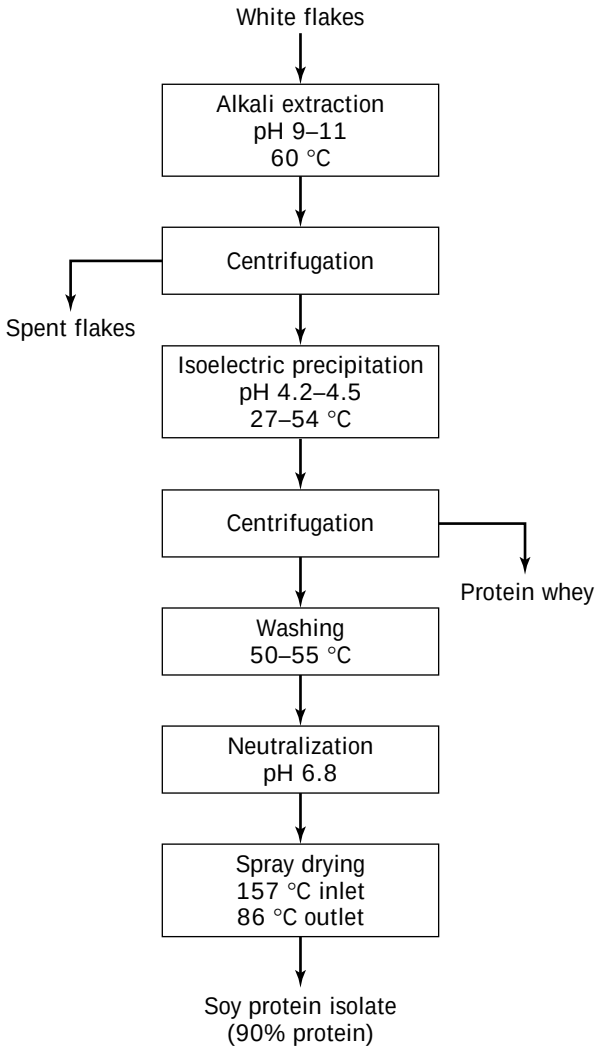


Figure 4. Production of soy protein isolate from white flakes.

proves the flavor and color of soy milk. Dehulling of soybeans is a process that consists of three steps, a heat treatment (93 °C for 15 min) to break the bonds between the hulls and the cotyledon, soybean cracking, and hull aspiration. A final dehulling efficiency of 88% is required for good quality soy milk (40). Most dehulling is done dry, but wet dehulling is also possible.

Soaking. Soybeans are soaked in water at a 1:3 solids-to-liquid ratio. The length of soaking depends on water temperature; the cooler the water, the longer the soaking. Typically, soaking time is 8 to 10 h at 20 °C in the summer and 14 to 20 hours at 10 °C in winter (40). Soybeans can be soaked in circulating water to be rinsed continuously throughout the soaking period. Most commercial soy milk production calls for soaking of soybeans to decrease the energy required for grinding, to achieve better dispersion and suspension of the solids, to increase yield, and to decrease cooking time (40,43).

Increased soaking time increases loss of water-soluble solids. Soaking of dehulled soybeans for 24 h results in a 5% solids loss as compared with a 1.5 % loss for whole (not dehulled) soybeans. Dehulled soybeans reach full hydration much faster than whole soybeans do and require only 2 to 3 h of soaking at 30 °C, or 1 h at 50 °C; whole soybeans require 7 h at 30 °C or 3 h at 50 °C (40). The addition of alkaline salts such as sodium bicarbonate, sodium citrate, or sodium hydroxide to the soaking water has been shown (*1*) to increase the protein and solids recovery, (*2*) to decrease beany off-flavor, (*3*) to inactivate trypsin inhibitors, (*4*) to tenderize soybeans, and (*5*) to affect soy milk quality by increasing soy protein solubility (43).

Fine Grinding. The purpose of fine grinding is to disrupt soybean tissue at both the cellular and the subcellular level and to suspend or to solubilize proteins, lipids, and other solids (43). Soaked soybeans may be ground with hot or cold water by a stainless steel disintegrator (Rietz), hammermill (Fiztpatrick), pin mill, or large blender (Waring) to produce coloid solution that will pass through a 150-mesh screen (40). Grinding in hot water (above 80 °C) inactivates lipoxygenases and prevents the formation of the beany off-flavor. Both very fine and rather coarse grinding are undesirable. Finely ground slurry contains maximum amount of soy protein but makes filtration difficult, thereby lowering soy milk yield. Coarse grinding may not result in the highest solubilization of the soy protein and other soluble solids and will be retained in the okara after the filtration (43). To improve flavor and mouthfeel, insoluble soybean residue is removed from the soy slurry by filtration or centrifugation.

Thermal Treatments. One or more heat treatments are used in soy milk processing. For example, in addition to cooking the slurry, the Illinois method calls for blanching (steaming), whereas the Cornell method calls for grinding the beans in hot water. Commercial processes also require pasteurization or sterilization of milk before packaging. Thermal treatment is one of the most important variables in the processing of soy milk (42). The positive effects of heat on soy milk include inactivation of antinutritional factors and lipoxygenases, increased extraction efficiency, denaturation of soy protein, and increased product shelf life. It is recommended that soy slurry or soy milk be heated for about 30 min at 100 °C to obtain optimal nutritional value and flavor. Heating also is applied to lower viscosity of the slurry to increase recovery of proteins and solids (40–43). Different soy milk processes use different temperature-time regimes during thermal processing. Effects of thermal processing on soy milk have been extensively reviewed by Kwok and Niranjan (42).

Homogenization. Homogenization is an optional step that is used to stabilize the soy milk emulsion and to disperse solids that otherwise may settle on the bottom of the container. Homogenization makes soy milk creamier and more uniform in consistency. For most soy milk products, a single pass at 2000 to 3500 psi and 90 °C through a dairy-type homogenizer is sufficient to obtain a good product. Some soy milk products exhibit a chalky taste caused by relatively large particles (>150 mesh) suspended in the milk (40). Chalkiness may be reduced by increasing the alkalinity of the blanching solution, by adjusting the final pH of soy milk to about 7.5, by decreasing solids content, or by homogenizing at high temperatures and pressures.

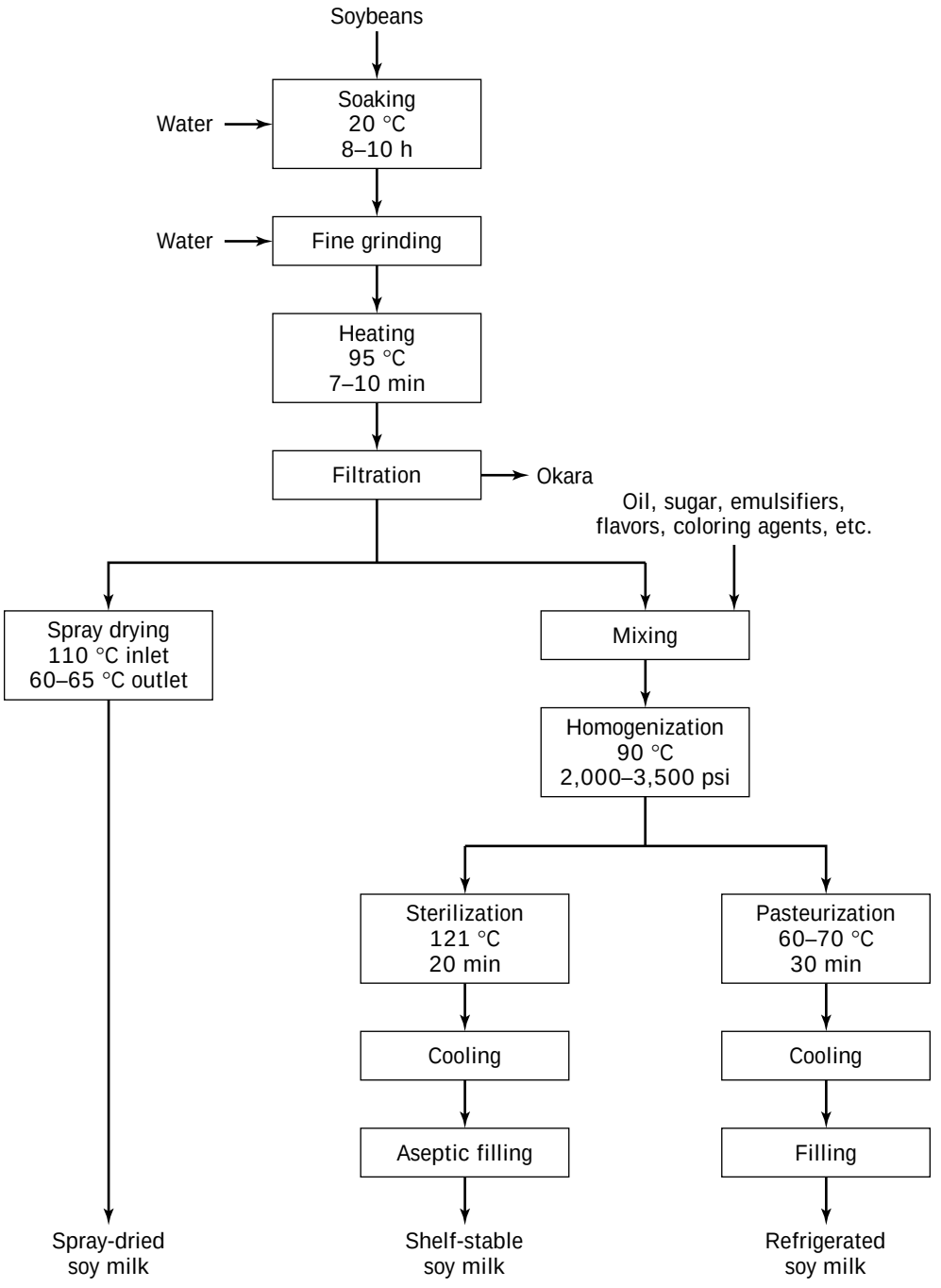


Figure 5. Alternative processing methods for soy milk production.

Pasteurization/Sterilization. After homogenization, soy milk can be pasteurized (75 °C for 15 min) or sterilized (12 °C for 15 to 20 min) and packaged for distribution. Pasteurized soy milk requires refrigeration at 4 °C to maintain a typical shelf life of about 1 week, whereas sterilized milk is shelf stable (Fig. 5).

Tofu. Tofu is a curd that is made directly from soybeans and resembles a soft white cheese. Basically, tofu is water-extracted and salt- or acid-coagulated soy protein gel with water, lipids, and other constituents trapped in its network (43). Typical pressed tofu has a moisture content of about 85% and contains about 7.8% protein, 4.2% lipids, and 2 mg/g calcium. On a dry basis, it contains about 50% protein and 27% oil.

In the United States, the tofu industry has shown steady growth during the past 10 years, partly because of the demand for healthy and low-fat ingredients (44). The quality and yield of produced tofu depend on several factors including soybean variety. For example, the ratio of

the 7S and 11S proteins in soybeans contribute to the textural characteristics of tofu (45). Tofu made from crude 11S protein is significantly harder than that made from 7S protein, and the crude 11S fraction adds to the springiness, chewiness, and gumminess of tofu (43,45). During the preparation of soy extract (milk), water-to-bean ratio, temperature and hardness of water, degree of maceration, extraction time, and grinding equipment affect amount of extracted protein (46).

There are three main steps involved in the making of tofu: preparation of soy milk, coagulation of protein, and pressing and molding (Fig. 6) (43,45,46). The processing steps for preparation of soybean extract are similar to those discussed in the previous section. A greater yield and better quality of tofu was obtained when the bean slurry was filtered before heating (45). The bean slurry is filtered through a cloth sack to separate the undispersible fiber (okara) from the soy milk. The okara normally is washed with hot water, stirred, and pressed again to increase dispersible solids recovery. Previous studies have shown that tofu yield and texture depend on the heating temperature. The texture of tofu produced from soy milk heated below 70 °C is soft and watery, between 70 and 80 °C is good and above 80 °C is hard and uneven (48).

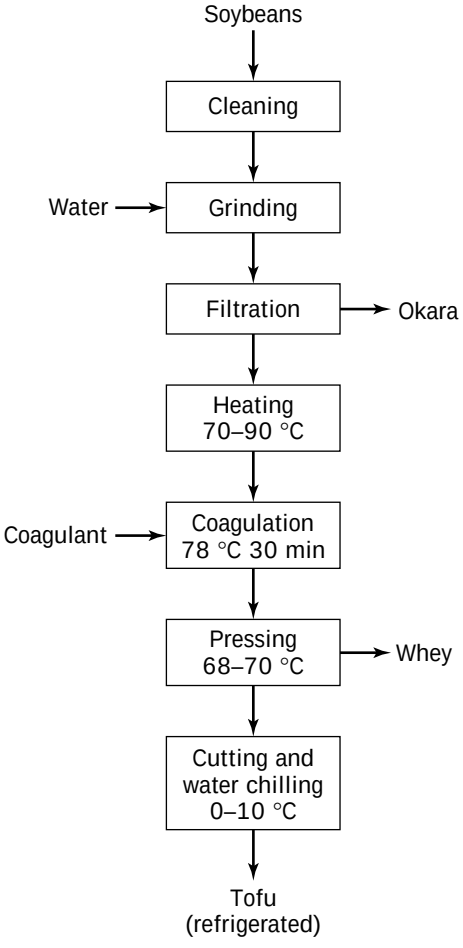


Figure 6. Basic unit operations involved in processing of soy milk into tofu.

Coagulation. Coagulation of soy milk is the most important step in tofu making and the most difficult to control because it depends on the complex interaction of many variables (49,50). Temperature of the soy milk, coagulant amount and concentration, and mode and rate of stirring all affect yield and texture of tofu. The ideal pH for soy milk coagulation is at or near 6. Soy proteins start coagulating near pH 6, and the addition of coagulant should be stopped at that pH (45). The amount and type of coagulant are important in producing reproducible tofu. Insufficient amount of coagulant would fail to coagulate the soy milk, but too much would result in low yield of small, hard aggregates with an uneven texture and a bitter (magnesium chloride) or chalky (calcium sulfate) taste (41). Data indicate that a coagulant concentration of 0.02 to 0.04 M is likely to produce a firm product (45). The characteristics of the five most common coagulants (magnesium sulfate, magnesium chloride, calcium sulfate, calcium chloride, and glucono-*d*-lactone) have been described by Karta (45). Stirring is necessary to keep the coagulant suspended, and the speed of stirring must be sufficient to maintain uniform distribution of the coagulant. A short mixing time is preferred because prolonged stirring may result in breakage of protein aggregates (50). In general, higher speed and shorter mixing time increase tofu yield.

Pressing. The soy protein coagulate (curd) is stirred and transferred to a shallow box for pressing. The purpose of breaking the gel is to disrupt the protein network to release the whey. Part of the whey is separated from the broken curd before transferal to the press box, and the rest is separated during pressing. Degree of curd breakage affects tofu yield and texture. Greater curd breakage generally leads to firmer texture and lower bulk yield. Variations in the pressing step include the amount of pressure and time applied (43). A study by Gandhi and Bourne (51) indicated that increasing the pressure from 0.19 Pa to 0.74 Pa reduced decrease moisture content from 82 to 60% and yield from 2 to 1.2 kg fresh tofu/kg soybeans. It also increased the hardness, chewiness, and gumminess of the final product. The ideal pressing temperature is between 68 and 70 °C; and, from a microbiological quality standpoint, tofu temperature should not fall below 60 °C (41).

Toasted Full-Fat Soy Flour. Full-fat soy flour, with different degrees of heat treatment, may be considered as modern industrial version of oriental soybean powder (5). It is made from dehulled soybeans and used mainly in bakery and dietetic foods. In the preparation of full-fat soy flours, soybeans first are steamed under light pressure for 20 to 30 min to inactivate lipoxygenases (52). Beans then are dried, cracked, dehulled, and ground finely to obtain full-fat soy flour.

Enzyme-Active Full-Fat Soy Flour. Soybeans used for producing enzyme-active full-fat soy flour are dried at low temperatures to minimize protein denaturation (52). Enzyme-active flour is produced by a process similar to that described for full-fat soy flour except that the initial steaming step is omitted. The resulting product is used widely in the food industry for bleaching wheat flour and for conditioning doughs in western-type breads. Enzyme-

active full-fat flour contains several lipoxygenases that bleach the carotenoid content of wheat flour to produce a whiter bread (5). The enzyme-active flour also possesses β-amylase activity, which improves the textural properties of dough. The U. S. Food and Drug Administration permits a maximum of 0.5% (w/w) of enzyme-active soy flour in wheat flour.

Fermented Foods

Fermented soybean food products are consumed commonly in Asian countries and are becoming increasingly popular in human diets in the United States and other western countries. Different procedures are used to produce somewhat similar and yet unique fermented soybean products, whose names differ between countries. This section will focus on manufacturing soy sauce, soy paste, and tempeh because they are representative and the commercially most important fermented products and because their process operations are similar to those of other fermented foods.

Soy Sauce. Soy sauce is a dark-brown liquid extracted from a fermented mixture of soybeans and wheat. Among the fermented soy foods, soy sauce is the most widely accepted, not only in the Far East but also in western countries (47). Depending on the geographic region, there are great variations in processing method and variety of soy sauce. Household production of soy sauce is still popular in some regions, but most soy sauce is made in commercial plants. Regardless of production level and method used, the basic steps and principles involved in soy sauce manufacture are very similar. Based on the method of preparation, soy sauce is divided into three types: fermented, chemical, and semichemical (47,53,54). In fermented soy sauce, proteins and carbohydrates in the raw material are enzyme hydrolyzed very slowly under mild conditions, below 30 °C for longer than 6 months. In chemical soy sauce, protein and carbohydrates are hydrolyzed quickly with hydrochloric acid at greater than 80 °C for 8 to 10 h. Semichemical soy sauce usually involves chemical hydrolysis of the raw material followed by lactic acid or yeast fermentation (47,53,54).

Fermented Soy Sauce. A typical process for the manufacture of a representative Japanese soy sauce (shoyu) generally involves treatment of raw material, koji preparation, brine fermentation, pressing, and refining.

Cleaned soybeans are soaked overnight in water at ambient temperature and then cooked for several hours under pressure. Soybeans are cooled quickly to 40 °C to prevent growth of undesirable microorganisms before fermentation (Fig. 7). Wheat kernels are roasted at 170 to 180 °C for a few minutes and then cracked into several pieces to be able to adsorb moisture from soybeans upon mixing (47,54).

Two fermentations are used in making soy sauce (Fig. 7). The first fermentation is a solid state and occurs during koji preparation, during which various enzymes are produced under aerobic conditions. The second fermentation, known as *brine fermentation*, is mainly anaerobic and begins after the addition of brine to the koji (47,54). The koji is the source of hydrolytic enzymes for converting polysaccharides and proteins into simple sugars, peptides, and amino acids, all of which are utilized by lactic acid bacteria and yeast in the subsequent brine fermentation (53). The soybean-wheat mixture is inoculated with a 0.1 to 0.2% (w/w) culture of *Aspergillus oryzae* or *A. sojae* and allowed to ferment in shallow, perforated vats for 3 to 4 days at 30 °C. Before brine fermentation, mature koji is mixed with 17 to 23% NaCl solution to achieve a final salt concentration of 17 to 19%. Lower salt concentrations allow growth of undesirable bacteria, whereas higher concentrations (greater than 23%) may retard growth of desirable halophilic bacteria and osmophilic yeast. The mash (*moromi*) is kept in large wooden containers or concrete vats fitted with aeration devices.

Temperature is an important factor during brine fermentation. In general, the higher the temperature, the shorter the fermentation time. Fermentation at lower temperatures results in a better product because enzyme inactivation rate is lower. A good quality soy sauce can be produced in 6 months by maintaining the mash temperature at 15 °C for the first month, at 28 °C the next 4 months, and at 15 °C the last month. Occasional stirring is necessary to provide aeration for yeast growth, to prevent growth of undesirable microorganisms, to maintain uniform temperature, and to facilitate CO_2 removal (47,54).

At the end of the fermentation and aging period, matured mash should have a bright reddish brown color, a pleasant aroma, and a salty but palatable taste. Raw soy sauce is commonly removed from the mash by pressing. Filtrate is stored in tanks to separate sediment (on the bottom) and floating oil (47,54). Refining of raw soy sauce consists of pasteurization (60 to 70 °C) and clarification by either sedimentation or filtration. Kaolin, diatomite, or alum may be added to enhance clarification. Heating during pasteurization inactivates enzymes and microorganisms, darkens the sauce, enhances the aroma, and induces formation of flocculants, which also facilitate clarification (47,54).

Chemical Soy Sauce. The chemical process, much simpler, faster, and cheaper than the fermentation process, results in a product lacking some of the fermented soy sauce flavor. Chemically processed soy sauce has a very limited market in Asia. In the western countries, where consumers are unaccustomed to the fermented product, acceptance of chemical soy sauce has been more rapid (54). In the chemical process, raw materials (soybeans and wheat) are hydrolyzed by boiling with 18 to 20% hydrochloric acid for 8 to 16 h (47,54). Hydrolysate is neutralized with sodium hydroxide or sodium carbonate to pH 4.5, pasteurized, and filtered. The resulting product is a clear dark-brown liquid with an approximate concentration of 18% (w/v) NaCl (47,54).

Semichemical Soy Sauce. Chemical soy sauce does not have the taste and flavor of fermented soy sauce. To improve its quality, chemical soy sauce is often blended with fermented soy sauce. A semichemical process has been developed whereby chemical hydrolysis is followed by lactic acid or yeast fermentation. Dilute hydrochloric acid (7 to 8%) is used to partly hydrolyze defatted soybeans, koji is added to the neutralized hydrolysate, and the mixture is inoculated with yeast. Fermentation can be shortened by

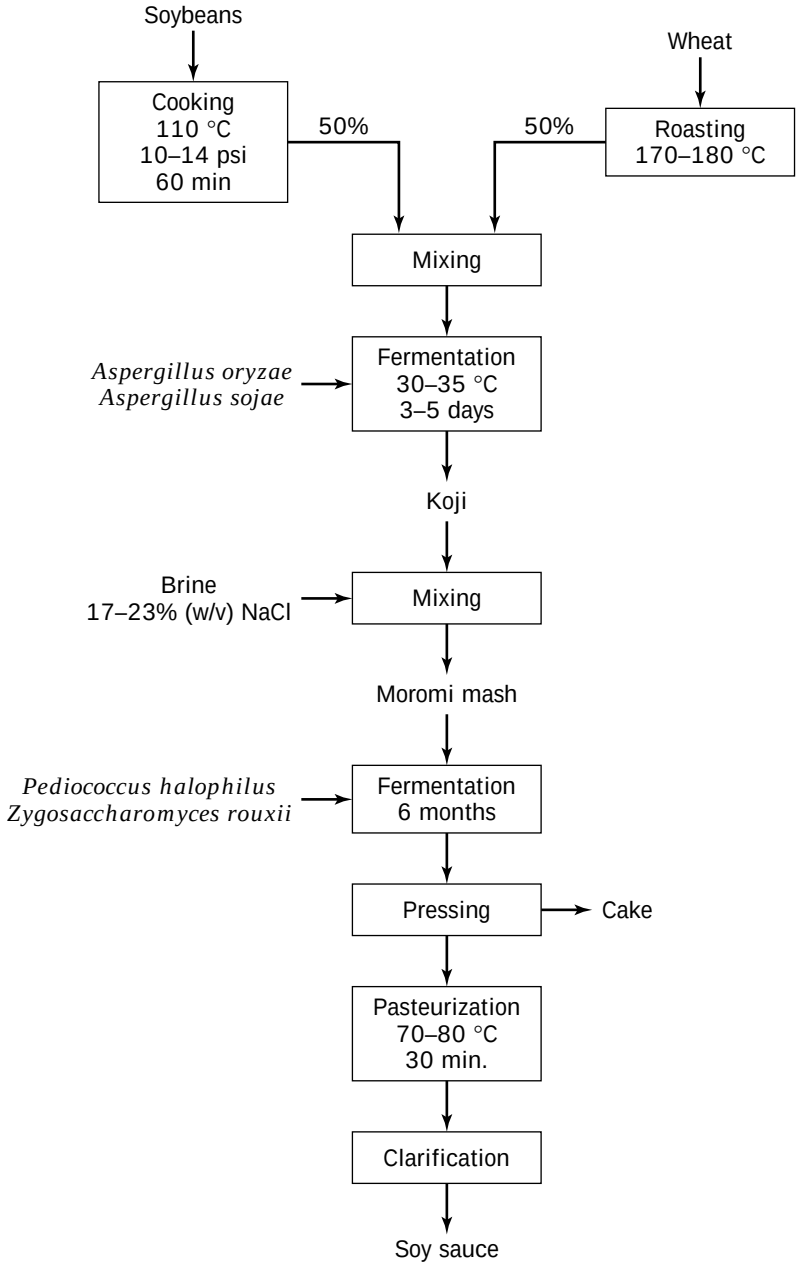


Figure 7. Flow diagram of the unit operations involved in the production of fermented soy sauce.

2 to 3 months, but the product retains some odors characteristic of chemically hydrolyzed beans (47,54).

Soy Paste. Soy paste is one of the most important fermented oriental soy foods. It is known commonly as *jiang* in China, *miso* in Japan, *jang* in Korea, *taucho* in Indonesia, and *taotsi* in Philippines. The process of making soy paste depends on the region, but the basic processes are essentially similar. The production of Japanese rice miso involves five steps (rice koji preparation, soybean treatment, ingredient mixing, fermentation, pasteurization, and packaging) that are very similar to those used in producing soy sauce (47,54). The major differences are that koji for miso is grown only on rice and that, because the whole fermented mixture is pasteurized and packaged for distribution, no pressing and filtering is required, as it is in the soy production. In this section, pertinent processing steps and conditions for making Japanese rice miso will be briefly discussed (Fig. 8).

Rice koji is prepared by soaking polished rice in water at 15 °C overnight or until the grain has a moisture content of 35%. After draining, the rice is cooked in an open cooker for 40 min, cooled to 35 °C, inoculated with a 0.1% (w/w) koji starter with *A. oryzae*, and incubated at 30 °C and 96% relative humidity (RH) for 15 h. After 15 h incubation, rice is spread on trays and maintained at 35 °C for an additional 25 to 30 h (47,54).

Soaked soybeans are steam cooked under pressure (10 to 15 psi) for 20 to 30 min. For white or light colored miso, soybeans are boiled with water at a 1:4 solids-to-water ratio to prevent browning. The advantage of boiling over steam cooking is the retention of soybean color, but the disadvantage is a 10 to 20% solids loss. Batch-type cookers are used widely, but continuous cookers with rotary bulbs and belt conveyers have been used recently in larger factories (47,54).

Pretreated soybeans are mixed with the rice koji and salt and blended in a closed blending container or ground with a sausage-type grinder. The mixture is packed tightly into wood, concrete or steel vats, covered with resin sheets or plates, and allowed to ferment at 30 to 38 °C. Fermentation for sweet miso production lasts between 10 and 15 days and for salty miso—2 and 12 months.

After ripening, miso is blended and, if needed, mashed again through a chopper with a plate cutter with perforations of 1 to 2 mm. The mashed miso is packaged in resin bags or cubic containers. For preservation, miso can be pasteurized or mixed with preservatives such as 2% ethyl alcohol or less than 0.1% sorbic acid (47,54).

Tempeh. Tempeh is a compact, cakelike product made by fermentation of dehulled and cooked soybeans with *Rhizopus oligosporous*. Tempeh is one of the most popular traditional foods in Indonesia and is becoming a popular vegetarian dish in the United States because it serves as a major source of protein, calories, and vitamins (55). A broad range of tempeh preparation methods have been reported; principles are the same, however, regardless of the method. Steinkraus et al. (54) have described a large-scale tempeh production method consisting of size grading, dry dehulling, hydration with optional addition of lactic acid, boiling, and fermentation (Fig. 9).

First, soybeans are cleaned to remove dirt, stones, weed seed, damaged and decomposed beans, and any other matter. Cleaned and graded soybeans are shrivelled slightly by dry heating for 10 min at 93 °C, cooled, and cracked to facilitate hull removal (47). Dehulled soybeans are soaked

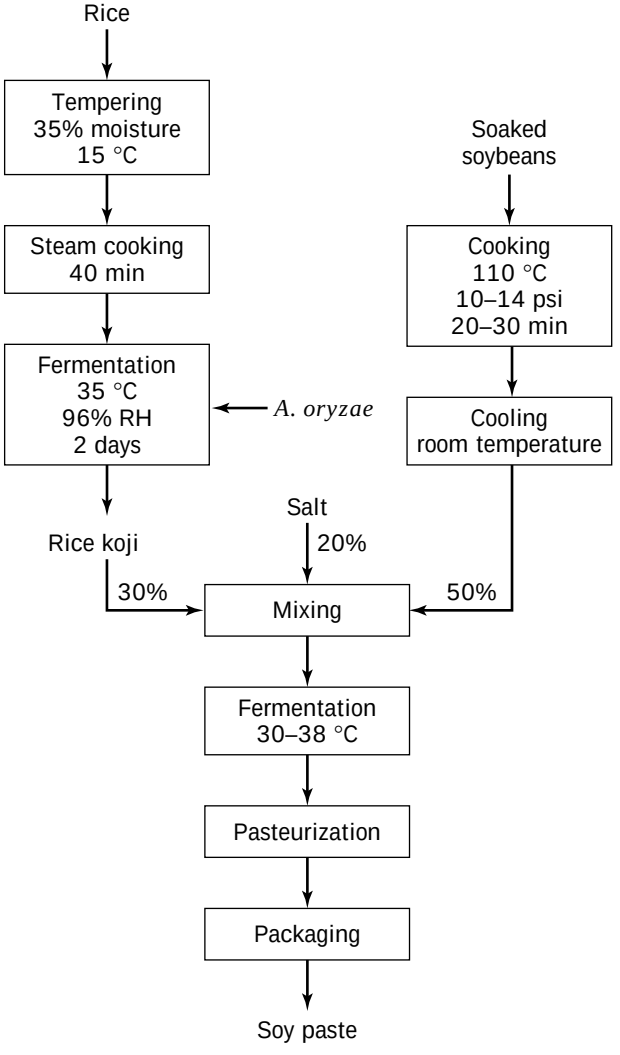


Figure 8. Processing steps in the production of Japanese rice miso.

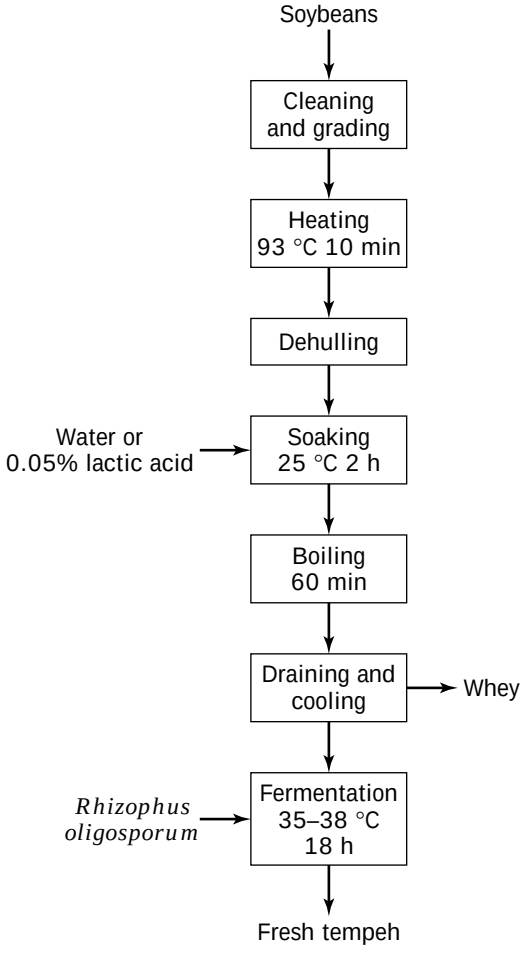


Figure 9. Basic unit operations in the processing of soybeans into tempeh.

Table 3. Principal Processing Steps and Uses of Fermented Soy Food Products

Food product (product description)	Main processing steps	Uses
Natto (cakelike product)	Soaking and cooking, fermenting with *Bacillus natto*	Side dish
Tao tio (soy paste type)	Cooking, fermenting with *Aspergillus* sp., sun-drying	Side dish, flavoring
Kochu chang (soy paste)	Boiling, mashing, fermenting, and drying	Flavoring
Ketiap (soy sauce)	Soaking, cooking, fermenting with *Aspergillus oryzae*, filtrating and washing, concentrating, and bottling	Flavoring
Sufu (soft cheese type)	Cutting, sterilizing, fermenting with *Mucor* and *Antinomucor* sp., brine aging	Side dish

Source: Data adapted from Johnson (1989) and Lui (1997) (47,52).

in water or 0.85% lactic acid solution for 2 h at 25 °C or for 30 min at 100 °C. Addition of lactic acid results in a pH drop to 4.3 to 5.3 and controls contaminating bacterial growth.

Hydrated soybeans are boiled for 60 min, drained, and cooled before inoculation with a pulverized pure *R. oligosporus* starter containing both mycelia and spores. Inoculated beans are spread out on perforated trays and covered with waxed paper to prevent dehydration and excessive aeration. Trays are incubated at 35 to 37 °C and 75 to 78% RH for about 18 h. At the end of the fermentation period, fresh tempeh is harvested, dried, and packaged before storage and shipping (47).

There are other traditional fermented products made from soybeans. Table 3 summarizes some of the best known, their principal processing steps, and their uses. Until now, most have been prepared on a small scale for family or regional consumption. Globalization of the world trade and economy is expected to increase demand for some of those products, whose production, in turn, will require further scale-up and modernization.

CONCLUDING REMARKS

Processing of soybeans into food, feed, and industrial products has experienced significant growth in the past, but new developments and challenges are foreseen. Continuous efforts must be made to further reduce energy consumption and to address safety issues and various environmental concerns such as air pollution, odor control, and waste treatment. New products and associated processing issues will emerge as a result of innovations in agricultural biotechnology. Plant biotechnology progress will not bypass the soybean industry; on the contrary, soy growers and processors will play an important role in capturing added value from new plant traits. For example, the precise alteration of metabolic processes in soybean plant will allow modification of quality, functional properties, and quantity of the protein, oil, and phytochemicals to fit end application. The new traits imparted to soybeans will require better understanding and control of soybean handling and processing to achieve the full impact and capture the added value of improved food and feed products. The ability to transfer new genes into soybeans (transgenic soybeans) will enable their use not only as a source for modified oil and protein but also as bioreactors to produce pharmaceutical proteins and drugs, industrial enzymes, biodegradable polymers, and nutritional supplements. Transgenic soybeans producing novel biomolecules will require significant processing adjustments and, in many instances, dedicated processing facilities. The success of newly derived products from genetically modified soybeans will probably depend on the close interaction or integration of the production, processing, and end application. The developments in processing of genetically engineered soybeans will ultimately affect processing and utilization of commodity soybeans.

ACKNOWLEDGMENTS

The authors thank Drs. L. A. Johnson and L. A. Wilson for reviewing the manuscript and for providing useful suggestions throughout its preparation. This work has been supported in part by the Center for Crops Utilization Research of Iowa State University. This is Journal Paper no. J-17843 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project no. 3331.

BIBLIOGRAPHY

1. M.J. Morse, in K.S. Markley ed., *Soybeans and Soybean Products*, Interscience Publishers, New York, 1950, pp. 3–60.
2. K.S. Liu, in K. Liu ed., *Soybeans: Chemistry, Technology, and Utilization*, Chapman & Hall, New York, 1997, pp. 1–24.
3. R.E. Fiedler, *JAOCS* **48**, 43–46 (1971).
4. D.D. Asbridge, in D.R. Erickson ed., *Practical Handbook of Soybean Processing and Utilization*, AOCS Press, Champaign, Ill., 1995, pp. 1–8.
5. Z. Berk, *Technology of Production of Edible Flours and Protein Products from Soybeans*, FAO Agricultural Services, Bulletin 97, Rome, Italy, 1992, p. 178.
6. K.J. Smith, in E.W. Lusas, D.R. Erickson, and W.K. Nip eds., *Food Uses of Whole Oil and Protein Seed*, AOCS Press, Champaign, Ill., 1989, pp. 1–11.
7. *Soya Bluebook Plus*, Soyatech, Inc., Bar Harbor, Maine, 1997.
8. T.H. Applewhite, in Council for Agricultural Science and Technology ed., *Quality of U.S. Agricultural Products*, Report no. 126, Council for Agricultural Science and Technology, Ames, Iowa, 1996, pp. 86–110.

9. A. Proctor, in K. Liu ed., *Soybeans: Chemistry, Technology, and Utilization*, Chapman & Hall, New York, 1997, pp. 297–346.
10. J.B. Woerfel, in D.R. Erickson ed., *Practical Handbook of Soybean Processing and Utilization*, AOCS Press, Champaign, Ill., 1995, pp. 39–55.
11. H. Schumacher, in T.H. Applewhite ed., *Vegetable Protein Utilization in Human Foods and Animal Feedstuffs*, AOCS Press, Champaign, Ill., 1989, pp. 37–40.
12. J.B. Woerfel, in D.R. Erickson ed., *Practical Handbook of Soybean Processing and Utilization*, AOCS Press, Champaign, Ill., 1995, pp. 65–91.
13. K.F. Carlson and J.D. Scott, *Inform* **2**, 1034–1060 (1991).
14. T.G. Kemper, *Inform* **6**, 1231–1236 (1995).
15. E.F. Sipos and B.F. Szuhaj, in Y.H. Hui ed., *Bailey's Industrial Oil and Fat Products*, 5th ed., Wiley, New York, 1996, pp. 497–601.
16. L.A. Johnson and E.W. Lusas, *JAOCS* **60**, 229–242 (1983).
17. L.A. Johnson, in P.J. Wan and P.J. Wakelyn eds., *Technology and Solvents for Extracting Oilseeds and Nonpetroleum Oils*, AOCS Press, Champaign, Ill., 1997, pp. 4–47.
18. G.C. Mustakas, in D.R. Erickson, E.H. Pryde, O.L. Brekke, T.L. Mounts, and R.A. Falb eds., *Handbook of Soy Oil Processing and Utilization*, AOCS, Champaign, Ill., 1980, pp. 49–66.
19. E.D. Milligan, *JAOCS* **53**, 286–290 (1976).
20. N.H. Witte, in D.R. Erickson ed., *Practical Handbook of Soybean Processing and Utilization*, AOCS Press, Champaign, Ill., 1995, pp. 93–116.
21. R.W. Fulmer, in T.H. Applewhite ed., *Vegetable Protein Utilization in Human Foods and Animal Feedstuffs*, AOCS Press, Champaign, Ill., 1989, pp. 55–61.
22. T.G. Kemper, *Inform* **5**, 898–901 (1994).
23. P.J. Wakelyn, in P.J. Wan and P.J. Wakelyn eds., *Technology and Solvents for Extracting Oilseeds and Nonpetroleum Oils*, AOCS Press, Champaign, Ill., 1997, pp. 48–74.
24. T. Gum, *Inform* **5**, 902–905 (1994).
25. G.R. List and D.R. Erickson, in D.R. Erickson, E.H. Pryde, O.L. Brekke, T.L. Mounts, and R.A. Falb eds., *Handbook of Soy Oil Processing and Utilization*, AOCS Press, Champaign, Ill., 1980, pp. 267–354.
26. D.R. Erickson, in D.R. Erickson ed., *Practical Handbook of Soybean Processing and Utilization*, AOCS Press, Champaign, Ill., 1995, pp. 174–183.
27. D.R. Erickson, in D.R. Erickson ed., *Practical Handbook of Soybean Processing and Utilization*, AOCS Press, Champaign, Ill., 1995, pp. 184–202.
28. D.R. Erickson, in D.R. Erickson ed., *Practical Handbook of Soybean Processing and Utilization*, AOCS Press, Champaign, Ill., 1995, pp. 203–217.
29. D.R. Erickson and M.D. Erickson, in D.R. Erickson ed., *Practical Handbook of Soybean Processing and Utilization*, AOCS Press, Champaign, Ill., 1995, pp. 218–238.
30. C.T. Zehnder, in D.R. Erickson ed., *Practical Handbook of Soybean Processing and Utilization*, AOCS Press, Champaign, Ill., 1995, pp. 239–257.
31. L.A. Johnson and D.J. Myers, in D.R. Erickson ed., *Practical Handbook of Soybean Processing and Utilization*, AOCS Press, Champaign, Ill., 1995, pp. 380–427.
32. E.W. Lusas and K.C. Rhee, in D.R. Erickson ed., *Practical Handbook of Soybean Processing and Utilization*, AOCS Press, Champaign, Ill., 1995, pp. 117–160.
33. N. Hettiarachchy and U. Kalapathy, in K. Liu ed., *Soybeans: Chemistry, Technology, and Utilization*, Chapman & Hall, New York, 1997, pp. 379–411.
34. K.S. Liu, in K.S. Liu ed., *Soybeans: Chemistry, Technology, and Utilization*, Chapman & Hall, New York, 1997, pp. 412–441.
35. J.P. Kearns, G.J. Rokey, and G.R. Huber, in T.H. Applewhite ed., *Vegetable Protein Utilization in Human Foods and Animal Feedstuffs*, AOCS Press, Champaign, Ill., 1989, pp. 353–362.
36. K. Seguro and M. Motoki, *Inform* **5**, 308–313 (1994).
37. L.P. Hanson, *Vegetable Protein Processing*, Noyes Data Corporation, New Jersey, 1974.
38. A.I. Nelson, M.P. Steinberg, and L.S. Wei, *J. Food Sci.* **41**, 57–61 (1976).
39. W.F. Wilkens, L.R. Mattick, and D.B. Hand, *Food Technol.* **21**, 86–89 (1967).
40. S. Chen, in T.H. Applewhite ed., *Vegetable Protein Utilization in Human Foods and Animal Feedstuffs*, AOCS Press, Champaign, Ill., 1989, pp. 341–352.
41. L.A. Wilson, in D.R. Erickson ed., *Practical Handbook of Soybean Processing and Utilization*, AOCS Press, Champaign, Ill., 1995, pp. 428–459.
42. K.C. Kwok and K. Niranjan, *Int. J. Food Sci. Technol.* **30**, 263–295 (1995).
43. K.S. Liu, in K.S. Liu ed., *Soybeans: Chemistry, Technology, and Utilization*, Chapman & Hall, New York, 1997, pp. 137–217.
44. K. Kevin, *Food Process. U.S.A.* **55**, 81–82 (1994).
45. S.S. Karta, in T.H. Applewhite ed., *Vegetable Protein Utilization in Human Foods and Animal Feedstuffs*, AOCS Press, Champaign, Ill., 1989, pp. 382–387.
46. C.G. Beddows and J. Wong, *Int. J. Food Sci. Technol.* **22**, 15–21 (1987).
47. K.S. Liu, in K.S. Liu ed., *Soybeans: Chemistry, Technology, and Utilization*, Chapman & Hall, New York, 1997, pp. 218–296.
48. C.G. Beddows and J. Wong, *Int. J. Food Sci. Technol.* **22**, 23–27 (1987).
49. C.G. Beddows and J. Wong, *Int. J. Food Sci. Technol.* **22**, 15–21 (1987).
50. H.J. Hou, K.C. Chang, and M.C. Shih, *J. Food Sci.* **62**, 824–827 (1997).
51. A.P. Gandhi and M.C. Bourne, *J. Textural Studies* **19**, 137–142 (1988).
52. D.W. Johnson, in E.W. Lusas, D.R. Erickson, and W.K. Nip eds., *Food Uses of Whole Oil and Protein Seeds*, AOCS Press, Champaign, Ill., 1989.
53. D. Fukushima, in C.W. Hesseltine and H.L. Wang eds., *Indigenous Fermented Foods of Non-Western Origin*, J. Cramer, Berlin, 1986, pp. 121–150.
54. K.H. Steinkraus, R.E. Cullen, C.S. Pederson, L.F. Nellis, and B.K. Gavitt, *Handbook of Indigenous Fermented Foods*, Dekker, New York, 1983.
55. F.G. Winarno, in T.H. Applewhite ed., *Vegetable Protein Utilization in Human Foods and Animal Feedstuffs*, AOCS Press, Champaign, Ill., 1989, pp. 363–368.

STAINLESS STEELS

C.P. DILLON
C.P. Dillon & Associates
St. Albans, West Virginia

KEY WORDS

Cleaning
Corrosion
Design
Materials selection
Passivation
Passivity
Pickling
Stainless steel
Superaustenitic
Welding

OUTLINE

Stainless steels offer an unusual combination of strength, ductility, formability, weldability, corrosion resistance, and amenability to cleaning that makes them uniquely suitable for pharmaceutical and bioprocessing operations. This article discusses the origin and nature of stainless steels, the several groups and many grades of stainless steels, together with the fundamentals of the metallurgical differences. Specific corrosion phenomena that pose potential problems, together with preventive and remedial measures, are presented. I explain the basic nominal workhorse compositions, together with their potential weaknesses, and discuss the possible utilization of other more highly alloyed substitute alloys. Particular emphasis is laid upon measures for quality assurance, postfabrication cleanup procedures, and passivation treatments to ensure improved resistance toward specific types of localized corrosion.

THE NATURE OF STAINLESS STEELS

As a whole, stainless steels and similar chromium-rich alloys are characterized by their *passivity*. Passivity is a condition in which a base metal, such as iron, exhibits the corrosion behavior of a more noble metal or alloy. As described further here, passivity cannot be maintained under some specific conditions.

Steel is an alloy of iron and carbon plus other minor elements and is notorious for susceptibility to rusting. With even minimal corrosion, it is a source of iron contamination and is therefore objectionable in many services. However, when a minimum of 12% chromium (Cr) is alloyed with steel, the resulting metal is generally not susceptible to rusting under ordinary circumstances of exposure to the atmosphere or in substantially neutral aqueous solutions. (Certain exceptions, however, are described further here.) Actually, several types of stainless steels, which are characterized by specific properties, have been developed.

TYPES OF STAINLESS STEELS

The 12% Cr grades of stainless steels are predominantly "martensitic." Martensite is a needlelike structure resulting from heating and quenching iron-base alloys, which gives increased hardness and strength with a concurrent diminution of ductility. (A certain amount of ductility is restored by tempering, i.e., heating to some intermediate temperature.) The martensitic grades are exemplified by Type 410 (UNS S41000) (1) and are characteristically magnetic.

When the chromium is increased to about 17%, the material is magnetic but not amenable to hardening by heat treatment. Such alloys, exemplified by Type 430 (S43000), are classified as "ferritic" and are more corrosion resistant than the martensitic grades by virtue of their higher chromium content, although they have lower strength than heat-treated martensitic grades.

The "workhorse" materials for the process industries, however, are the "austenitic" grades, often loosely called "18-8" stainless steels. When about 8% nickel (Ni) is added to an 18% Cr steel, a face-centered crystal structure called austenite is formed that is characteristically nonmagnetic. (This structure exists only at elevated temperatures with steels and the straight chromium grades.) Austenitic stainless steels, exemplified by Type 304 (S30400) and Type 316 (S31600), have in general a remarkable combination of properties that occasion widespread use in the process industries as compared with the martensitic and ferritic grades, which are used much less frequently.

Stainless steels that contain 18% Cr or more are also alloyed with a lesser amount of nickel (about 5%) to form "duplex" grades, which are about 50:50 austenite and ferrite. This method results in alloys of higher strength than Types 304 (S30400), 316 (S31600), and so forth, with an improved resistance in aqueous environments of high chloride content. The newer duplex grades are exemplified by Alloys 2205 (S31803) and 255 (S32550).

Finally, there are "superferritic" grades such as Alloy 26-1 (S44626) and several "superaustenitic" alloys, exemplified by Alloys 254SMO (S31254) and 6XN (N08367). The superaustenitic grades have superior resistance to localized corrosion such as pitting, crevice corrosion, and stress corrosion cracking (SCC) in chloride-bearing environments. Also, they are as strong as duplex grades by virtue of nitrogen addition.

The compositions of some typical specific grades in these basic categories are given in Table 1, wherein the midrange of major elements is given with the maximum allowable concentration for minor elements. More detailed listings of stainless steel types and compositions are available (2).

CORROSION MECHANISMS

Corrosion of metals is electrochemical in nature, metal ions being formed at anodic sites by oxidation, accompanied by a discharge of electrons through the external (metallic) circuit. At cathodic sites, electrons provide reduction of specific species (e.g., dissolved oxygen, cations) in the electrolyte to complete the reaction equilibrium. This change is most apparent in galvanic (i.e., bimetallic) corrosion in which the anodes and cathodes are different materials (e.g., the galvanic corrosion of steel by contact with copper, or of zinc by contact with steel). However, it also occurs in monometallic corrosion, such as steel in water, wherein discrete local anodes and cathodes of a transient nature provide the reactions.

$$Fe^{\circ} \rightarrow Fe^{2+} + 2e \text{ (anodic)}$$

$$2H^{+} + 1/2O_2 \rightarrow H_2O \text{ (cathodic)}$$

The formation and dissipation of discrete anodes and cathodes cause the general, uniform corrosion typical of base metals.

When a metal or alloy is protected by a superficial layer of a composition different from the substrate (mill-scale on steel; oxide films on copper-, nickel-, or chromium-bearing alloys), anodic sites may be relatively stable at defects in the otherwise protective film, causing a more localized type of attack. This reaction is somewhat analogous to galvanic corrosion, the film acting as a cathode to the more anodic substrate.

CORROSION SUSCEPTIBILITY OF STAINLESS STEELS

The stainless steels in general resist corrosion because of the formation of an invisible oxide film that renders them

Table 1. Composition of Typical Stainless Steels

Grade	UNS #	C	Cr	Ni	Mo	Other
			Martensitic grades			
410	S41000	0.15	12.5	—	—	1 Si
420	S42000	>0.15	13.0	—	—	1 Si
			Ferritic grades			
409	S40900	0.08	11.0	—	—	1 Si, Ti 6 × C
430	S43000	0.12	17	—	—	1 Si
444	S44000	0.025	18.5	1.0	2.0	Ti, Nb, N
			Austenitic grades			
304	S30400	0.08	19	9	—	1 Si
304L	S30403	0.03	19	10	—	1 Si
316	S31600	0.08	17	12	2.5	1 Si
316L	S31603	0.03	17	12	2.5	1 Si
317L	S31703	0.03	19	13	3.5	1 Si
Duplex grades						
329	*S32900*	*0.08*	*25*	*3.5*	*1.5*	*0.75 Si*
2205	*S31803*	*0.03*	*22*	*5.5*	*3.0*	*0.08–0.20 N*
255	*S32550*	*0.04*	*25*	*5.5*	*3.0*	*1.5–2.5 Cu*
			Superaustenitic grades			
254SMO	S31254	0.02	20	18	6.0	0.8 Cu, 0.20 N
367XN	N08367	0.03	21	24.5	6.5	1 Si
28	N08028	0.03	27	31	3.5	1.0 Cu
20Mo6	N08026	0.03	26	35	5.5	0.5 Si

passive in oxidizing environments. The semipermanent surface passive film is very thin (about 0.1 nm), analogous to ice on a pond. This passive film of chromic and silicon oxides plus adsorbed oxygen can be increased in thickness by exposing it to oxidizing solutions or by anodic passivation using an electrochemical technique. (Note: Electropolishing is followed by another electrochemical treatment to reform and enhance the passive film following anodic dissolution of the alloy. If the film is solubilized, as by reducing acids such as hydrochloric or dilute sulfuric, stainless steels change from a passive to an active state and are actually less resistant than ordinary iron and steel.) Most applications of stainless steels are in air, aerated water, and other oxidizing media wherein the passive film is retained. It should be noted that stainless steels are *autopassivating*; the film reforms immediately if removed by mechanical means (e.g., abrasion) when the alloy is subsequently exposed to air, water, or other oxidizing environments. They do not demonstrate the active state except under continuous abrasion or conditions of immersion in reducing environments.

Stainless steels are resistant to general corrosion in most applications, with corrosion rates <4–5 mpy. When data are reported showing general corrosion rates of 10, 20, or 50 mpy, this indicates that the alloy is fluctuating between an active and a passive state and should not be used in that specific environment. Unfortunately, despite their passive nature, stainless steels are subject to several forms of localized attack. These are described further for the austenitic grades, which are the primary alloys for bioprocessing applications (3).

FORMS OF LOCALIZED CORROSION

Pitting

Pitting can occur on boldly exposed surfaces in high-chloride environments, such as brackish cooling waters or dilute hypochlorite solutions (e.g., during sanitizing operations). The molybdenum-free alloys such as Type 304L (S30403) are more susceptible to pitting than molybdenum-bearing alloys in chloride-bearing environments, and the problem is combated by specifying successively higher grades, that is, Type 316L, Type 317L (S31703), and the 6% Mo grades as required by the severity of the environment.

Crevice Corrosion

Crevice corrosion is a special form of pitting occurring in physical or mechanical crevices (e.g., flanges, threaded fasteners) or under surface films. When it occurs under calcareous deposits, biomasses, or other foreign material, it is known as “under-deposit corrosion” or UDC. Typically, crevice corrosion is driven by a dissolved oxygen (DO) cell, the oxygen-starved crevice functioning as an anode to the boldly exposed cathodic area where the DO constitutes a cathodic depolarizer. However, the cell can also be driven by dissolved sulfur (DS) or by differences in metal ion concentrations. Crevice corrosion is aggravated by chloride contamination, the chlorides being attracted to the anodic sites and producing a locally low pH (e.g., pH 2) from acid chlorides within the crevice.

Sensitization

When the regular carbon grades such as Types 304 (S30400; 18 Cr–8 Ni–0.08 C maximum) or 316 (S31600) are heated in the temperature range from about 800° to 1,500 °F (as by slow cooling from the annealing process or from thermal stress-relieving treatment, or in the heat-affected zones of welds), carbon migrates to the grain boundaries and precipitates as chromium carbides. This leaves a chromium-depleted zone around each grain and causes a susceptibility to intergranular corrosion (IGA) in many environments. When IGA is observed in the heat-affected zone of welds, it is commonly called by the misnomer “weld decay,” although it is not the weld itself that suffers corrosion. Sensitized areas are also more prone to chloride pitting than are unaffected areas such as those at some distance from welds.

Conventional sensitization by chromium carbide precipitation was originally combatted by addition of stabilizing elements, that is, columbium (niobium) or titanium, to an 18-8 alloy such as Type 304 (S30400). When stabilizing elements are present in concentrations several times the carbon content, it is their carbides that precipitate, thereby avoiding chromium depletion. The stabilized grades such as Types 321 (S32100) and 347 (S34700) are therefore derivatives of Type 304 containing Ti and Cb, respectively, and are still used in some special applications. It should be noted that titanium-stabilized grades such as S32100 require a stabilizing anneal at about 1,800 °F to tie up the carbon in the form of titanium carbides before welding. IGA may occur even in stabilized or low-carbon grades if weldments are carbon-contaminated by oil or grease.

The stabilized grades can themselves become sensitized by multipass welding or cross-welding, causing knife-line attack (KLA); this is caused by a resolubilizing of the columbium or titanium carbides by the heat of welding, followed by slow cooling. This step requires about 1,250 °C (2,300 °F) reheating in 18-8 grades, moving the susceptible area to the fusion zone, and permitting a very narrow band of corrosion immediately adjacent to the weld (KLA). With higher nickel content, the resolubilization temperature drops; Alloy 20Cb3 (N08020) and Alloy 825 (N08825) with nickel contents in the 35% range are more easily sensitized to KLA (about 980 °C or 1,800 °F) than is type 347.

In modern practice, the problems of sensitization have been largely alleviated by use of low-carbon grades (Type 316L, S31603; 17 Cr–12 Ni–2.5 Mo–0.030 C maximum), although sensitization will occur with very prolonged exposure in the sensitizing range (e.g., in long-time service >800 °F or when stainless steel–clad vessels are thermally stress relieved). Modern-day steel-making practice using argon–oxygen decarbonation (AOD) makes very low carbon grades of stainless steels routinely available.

IGA may also occur due to mechanisms other than carbide precipitation. A ferrite phase, if present for any reason, may be selectively attacked by reducing acids such as hydrochloric or sulfuric; its thermal conversion product, sigma phase, is selectively attacked by oxidizing acids (e.g.,

nitric). Chi phase can form in molybdenum-bearing grades (S31603) as the result of prolonged heating in the sensitizing range (e.g., thermal stress relief of clad-steel vessels), producing IGA in alloys of <0.030% carbon in certain aggressive environments.

Selective Weld Corrosion

Selective or preferential attack is sometimes encountered in weldments, particularly in the molybdenum-bearing grades. This can occur either in conventional welds or autogenous welds (i.e., welds made without the addition of filler metal) in some environments. This type of attack is ascribed to microsegregation, which causes a maldistribution of the molybdenum component of the alloy.

Selective weld corrosion is the reason for the cold-work requirement for welded stainless steel condenser tubes (see ASTM A249). Cold-working to minimum requirements before the final anneal allows the weld to recrystallize and homogenize, alleviating the microsegregation.

For butt welds in piping and fabrication welds in tanks and vessels, no such cold-work is feasible. Instead, an overmatching rod is used to confer higher resistance to the weld itself, for example, alloy 254SMO (S31254) to weld S31603 or S31703 or alloy 625 (N06625) to weld S31254.

Stress Corrosion Cracking

The predominant form of environmental stress cracking (ESC) in the 18-8 varieties of stainless steels is anodic stress corrosion cracking (SCC) under either applied or residual tensile stress. Compressive stress will not cause SCC, and indeed application of *compressive* surface stress, as by shot-peening, is sometimes used as a palliative surface treatment against SCC. Only the martensitic grades are subject to hydrogen-assisted cracking (HAC), a form of ESC found, for example, in "sour" service due to hydrogen sulfide effects.

The most common form of SCC of 18-8 austenitic grades is caused by chloride contamination. The ubiquitous chloride is sodium chloride (NaCl, table salt), as found in water and atmospheric exposures. NaCl usually causes SCC only between about 50 °C (120 °F) and 200 °C (390 °F). However, SCC can occur at room temperature and even at cryogenic temperatures in the presence of other chlorides, such as ferric chloride, sulfuric–NaCl mixtures, aqueous solutions of hydrogen sulfide–NaCl, and hydrogen chloride (HCl). For boldly exposed surfaces, there are rough correlations between pH and chloride concentration at which SCC may be anticipated. There is no minimum chloride concentration below which SCC will not occur if there is any possibility of concentration, either by evaporation or by occlusion or adsorption in films adhering to the stainless surface (e.g., mill scale, heat tints, welding slag, rust deposits, calcareous deposits, biomasses).

External stress corrosion cracking (ESCC) occurs when chlorides from the atmosphere, water leaks, or insulation concentrate on the surface of an 18-8 type stainless steel pipe or vessel. This is best combatted by coating vessels and piping that are more than 4 in in diameter and that operate between 60 and 200 °C, using a zinc-free and chloride-free paint system (4,5).

Hot caustic can also cause SCC of 18-8 stainless steels, as frequently happens when high-pressure steam is contaminated by carryover of alkaline boiler-treating chemicals. This is a potential problem in steam lines and during sterilization procedures. Use of nickel-base alloys such as N06625 or better (but not N08800 or N08825) is effective against caustic SCC.

Rouging in High-Purity Water

Austenitic stainless steels are subject to superficial rusting under some circumstances, either because of tramp iron embedded in the surface or from chloride contamination by corrosive atmospheres or water environments. The former problem is addressed through standard cleaning and passivation techniques described further next.

An apparently separate problem, "rouging," is sometimes observed with clean stainless steels in high-purity water at about 80 °C (175 °F) (6). Strictly speaking, rouge consists of a thin red deposit of α-Fe_2O_3 which is usually loosely adherent, wiping off easily. Generally, very little corrosion of the surface is evident. Typically, rouging is experienced in high-purity deionized water such as WFI ("water for injection") in the pharmaceutical industry, and UHPW ("ultrahigh-purity water") in the electronics industry and in nuclear reactor systems. When the U.S. Food and Drug Administration and solutions for human injection are involved, the presence of red-colored oxide on the surfaces of stainless steel equipment is not acceptable. Also, the standards for ultrahigh-purity water in the electronics industry do not permit this sort of condition.

One explanation that has been advanced to explain the fairly recent development of this phenomenon relates to modern steel-making processes. The AOD process itself gives a precise control that permits the manufacturer to work at the lowest ranges permissible for alloying elements, for example, chromium, within the nominal range specified. During long-time hot-working and annealing, the stainless steels form a chromium oxide scale that depletes the metal substrate of Cr. Modern short-time "spray pickling" may not be adequate to dissolve the depleted layer, leaving a surface of somewhat lesser corrosion resistance than the nominal analysis. However, even minor mechanical abrasion restores the proper surface chemistry, and electropolishing, although not simple nitric acid "passivation," which does not attack alloys with as little as 11% Cr, also does so.

A rouge of loose ferric oxide has been developed on Type 304L in laboratory tests in deionized water at 250 °C, simulating a reactor loop. When the water was cooled for sampling and filtered through a 0.1-μm PTFE filter supported on a 0.2-μm silver filter, no deposits were found on the PTFE but large crystals of spinel ferric oxide (with traces of Cr and Ni) had developed on the silver filter and grown back into the PTFE from the backside. This reaction suggests either soluble or colloidal iron in the nearly supercritical (288 °C or higher) water. Loose rouge has also been observed in hot aggressive waters (e.g., WFI) as low as 80 °C (175 °F) from sterilizing procedures with stainless steels in the pharmaceutical industry. Some operators have used solutions of sodium hydrosulfite ($NaHSO_3$) to reduce and remove rouge from their systems.

Some authorities state that the term *rouging* is a misnomer because the films may be either the powdery red iron oxide or adherent films of various colors including gray, black, purple, blue, and even yellowish green. In practice, it is often difficult to distinguish between superficial corrosion (involving a black substrate under a reddish-brown crust, associated with craters and white debris) and extraneous deposits (7). Heat tints and similar films exhibit colors related to their thickness and, although they can occlude and concentrated corrosive species such as chlorides, should not be considered rouge.

In one case, an electropolished Type 316L filter plate developed a blue-black tint, with no apparent corrosion of the parent metal, in ultrahigh-purity water of reportedly ambient temperature or at least not more than about 55 °C (130 °F). Analysis determined that the film contained more than 85% iron as a black oxide, probably derived from equipment upstream and entering the system, perhaps during regeneration of the mixed ion exchange beds.

FACTORS CONTRIBUTING TO LOCALIZED CORROSION

Some factors that contribute to the occurrence of localized corrosion in the stainless steels are surface defects or foreign matter. Surface defects are undercut welds, scratches, and grinding marks, which can initiate localized attack.

Foreign matter constitutes embedded iron or iron oxides (usually from mill or shop contamination of the stainless steel surface), the bluish heat tints adjacent to welds, weld splatter or arc strikes, and paint or crayon markings. Iron oxides and heat tints can occlude and concentrate chlorides and are themselves susceptible to corrosive attack. Weld splatter and arc strikes weaken the surface film and may constitute crevices for the initiation of attack. Paints and crayon markings may contain organic chlorides, which can hydrolyze or thermally degrade to release traces of hydrochloric acid and promote localized attack.

Bacterial effects in unsterile water, for example, residual hydrotest water, can also cause localized attack. The production of traces of hydrogen sulfide by sulfate-reducing bacteria (SRBs) and the excretion of iron or manganese oxides by aerobic strains can cause serious pitting in a very short time. Sulfides compete with oxygen and weaken the passive film, contributing to pitting and SCC. Microbiologically induced corrosion (MIC) has attracted increasing interest in the process industries.

PREVENTION OF LOCALIZED ATTACK

To prevent or minimize the possibility of localized corrosion, a specific sequence of events must be established. The methodology for this starts with proper materials and design and the establishment of certain procurement details in the materials selection, specification, and inspection phases, plus postfabrication cleanup and preservice sanitizing or other preparatory steps.

Materials

Materials must be properly chosen for the expected conditions of fabrication and service. The alloys chosen should be indicated by the proper alloy designation or corresponding UNS number, and these should be allied with an appropriate ASTM or ASME specification and any necessary supplements.

For best results, cast components, such as pumps and valves (which are predominantly CF8M and CF3M, the cast equivalents of types 316 and 316L), should be purchased to ASTM A743 with a special annealing temperature requirement. Annealing at about 1,120 °C reduces the tendency toward molybdenum gradients within the casting, ensuring a more uniform corrosion resistance.

To avoid materials mixups, such as may occur either at the manufacturer's plant or in warehousing, the composition of the alloy should be verified not only by checking the mill analysis but also by spot-test or nondestructive metal identification procedures; this is particularly necessary when a molybdenum-bearing alloy has been chosen. Inadvertent substitution of Type 304L for 316L has caused many problems in process applications. The fact that a piece of stainless steel is clearly marked as Type 316L, Heat No. *X*, is not a guarantee.

Finally, special finishes may be required. The several commercial finishes (in decreasing order of surface roughness) are numbered 1 (hot-rolled, annealed, and pickled) through 8 (a mechanically polished and buffed specular finish, with no visible abrasive scratches). For bioprocessing applications, a number 2B finish (cold-rolled, annealed, pickled, and temper-rolled) or a number 4 (which adds a 150-grit abrasive belt grinding to the 2 B finish) is usually specified.

For extra smoothness in special nonstick applications or to facilitate removal of bacterial contaminants, stainless steel may be electropolished to a mirror finish in lieu of a number 8 finish. Electropolishing aids surface homogenization and autopassivation, and an actual anodizing can be added as the final step to enhance the passive film.

Design

Because localized corrosion causes about 90% of the failures experienced with stainless steels in process applications, vessels, piping, and ancillary equipment must be designed, purchased, and maintained to minimize the risk of such attack. This requires that crevices, pockets, and traps should be avoided, thereby minimizing the probability of localized corrosion.

All fabrication steps, such as bending, forming, and welding, should conform to standard procedures. The welds must be full penetration to prevent crevice corrosion in incompletely fused joints. Fillet welds should be smooth and continuous, with no undercutting. Butt welds should be smooth and corner welds should be avoided if possible; side-wall welds in tanks are preferable. Lines and vessels should be sloped and self-draining, and top nozzles, side ports, and bottom nozzles must be accessible for cleaning.

A postweld chemical cleaning with an ammoniated citric acid solution, as described next, will remove the heat tints from the welded heat-affected zones (HAZ).

REMEDIAL MEASURES

Pickling

Because stainless steels are chosen for appearance, corrosion resistance, and freedom from product contamination

in storage and shipment, pickling (to remove contaminants) or passivation (to augment the naturally occurring surface film) is often required. A surface of assured cleanliness, uniformity, and corrosion resistance is desirable and, in some services, absolutely required.

During fabrication, stainless steels are subject to contamination by embedded steel particles or iron oxides, as well as heat tints caused by welding. The alloy surface may be cleaned with acid solutions or pastes, mechanically cleaned with a stainless steel wire brush or a clean abrasive disk, or hydroblasted with high-pressure water of potable water quality (i.e., <250 ppm Cl^-). ASTM A 380, "Cleaning and Descaling Stainless Steel Parts, Equipment, and Systems," provides detailed procedures. Demineralized water is usually specified for critical services, but a 25-ppm maximum chloride limit should be adequate.

Embedded iron is subject to air oxidation to hydrated ferric oxide ($Fe_2O_3 \cdot H_2O$). Foreign rust particles or mill scale (Fe_3O_4) may also be present. "Pickling" (see following) is intended to effect the desired cleanliness while simultaneously homogenizing the surface. A solution of 10–15% nitric acid and 1–3% hydrofluoric (HF) acid has been used as a "bright pickle" to remove visible oxides, while pastes of similar composition are available for localized application. Such solutions and pastes constitute somewhat of a safety hazard to personnel, as well as potential disposal problems, but are very effective. (Note: There is increasing resistance to using HF mixtures today because of the safety concerns and disposal problems related to environmental considerations.)

A HNO_3–HF mixture, unlike nitric acid itself, is a reducing acid intended to remove iron oxides and other metallic oxides by chemical reduction. This mixture is not a passivating medium, and corrosion rates are high during the actual exposure. However, a uniform passive oxide film will reform over the freshly cleaned surface after removal of the pickling environment by autopassivation.

Heat tints from welding tend to adsorb chlorides, with attendant pitting or stress corrosion cracking. They may be removed by abrasive disks or with a nitric–HF paste. Citric acid solutions, such as a 5% solution adjusted to pH 3.5 with ammonia and used at about 80 °C, are very effective and are to be preferred for removing heat tint from bioprocessing equipment. This solution is a chelating agent as well as an acid cleaning medium. Unlike the nitric–HF materials, citric acid solutions are benign and quite readily disposable from the environmental standpoint. However, chelants incorporate the metal ions within a clawlike ringed structure that may pass through conventional waste treatment systems. When heavy metals (i.e., chromium, copper, nickel, and zinc) are involved, special waste treatment systems, rather than simple sanitary drain systems, may be employed.

Antisplatter pastes are sometimes used to minimize weld-splatter sticking. If such splatter has adhered to the adjacent areas, it should be removed by a flapper wheel, wire-brushing, or hydroblasting. A cleaning operation with organic reducing acids may be followed with an alkaline sodium nitrite neutralizing–oxidizing rinse. (Note: Adequate provision must be made for proper disposal of the chemical solutions in accordance with environmental regulations.)

Electropolishing, in which the work serves as the anode during electrochemical dissolution to produce a highly polished nonstick and uncontaminated surface, is a specialized form of chemical cleaning in which the surface is actively corroded during the operation. Although the stainless steel surface will passivate on rinsing and exposure to air, many users specify a final anodizing step, which is tantamount to chemical passivation.

Passivation

Cleaning and pickling, as described, and passivation of stainless steels, particularly in the widely used austenitic 300 series, for example, Types 304 (S30400), 304L (S30403), 321 (S32100), 347 (S34700), 316 (S31600), 316L (S31603), and 317L (S31703), are widely misunderstood.

Cleaning and pickling are intended to remove foreign materials and permit the surface to equilibrate or homogenize, effecting an electrochemically uniform surface for resistance to localized corrosion. Passivation has been defined as a reduction in the rate of the anodic reaction of an electrode involved in corrosion, as by formation of a protective oxide film.

As previously noted, stainless steels are self-passivating on exposure to air and moisture as the result of their high chromium content (i.e., $>11\%$). Passivation treatment consists of augmenting the naturally occurring protective oxide film by chemical or electrochemical methods. The enhanced passive film is somewhat thicker and more persistent than that formed by autopassivation. For example, in some iron-sensitive organic chemicals (e.g., acetoacetanilide, glyoxal), stainless steel vessels are specially passivated and repassivated at intervals thereafter whenever chromophoric corrosion products are detected in the chemical product. This practice suggests a gradual diminution in thickness (or localized weakening of) the passive film with time in some services. Passivating is accomplished by treatment with oxidizing solutions, such as warm dilute nitric acid (not nitric-HF), ammonium persulfate (which is very safe to handle), dilute hydrogen peroxide, and citric acid–sodium nitrate mixtures.

Table A2.1 of ASTM A 380, Part II, recommends passivation with 20–50% nitric acid at 49–71 °C (120–160 °F) for 10–30 min or 21–38 °C (70–100 °F) for 30–60 min for dull or nonreflective surfaces. In the latter case, a 10% caustic–4% potassium permanganate solution is suggested for smut removal after the final rinse. A 20–40% nitric acid with 2–6% sodium dichromate is suggested for bright-machined or polished surfaces, instead of straight nitric acid. Such solutions are also used for precipitation-hardening (PH) grades, 12–14% chromium grades, and high-carbon martensitic grades, such as Type 440 (S44000). However, chromates are not compatible with pharmaceutical products. (Note: In chemical plants, large tanks or vessels are usually passivated with 10–15% nitric acid at 66 °C (150 °F) for about 4 h.) A 1% citric acid–1% sodium nitrate solution at 21 °C (70 °F) is discussed in ASTM A 380, Part III. The dairy industry has used a 0.5–1.0% nitric acid for 10 min at 65–82 °C (150–180 °F).

A special sequence of alkaline–acid–alkaline (A-A-A) treatment is recommended for free-machining grades (8). Because oil or grease interferes with passivation and because attack on sulfur or selenium inclusions in Type 303 (S30300) and its variants can leave minute pits with residual corrodents on the metal surface, the machined parts are (*1*) degreased for 30 min in 5% sodium hydroxide at about 70–80 °C (160–180 °F) and rinsed, (*2*) immersed in 20% nitric plus 3 oz. per gallon sodium dichromate at 50–60 °C (120–140 °F) for 30 min and rinsed, and finally (*3*) neutralized in the original caustic solution.

Because oil or grease can interfere with self-passivation by preventing ready access of water and oxygen to the metal surface, degreasing may be required. Such contamination occurs primarily from machining operations and should be removed by degreasing with nonchlorinated solvents before passivation is attempted. (Note: Chlorinated solvents may introduce inorganic chloride ions, Cl^-, by hydrolysis or by thermal degradation, and are proscribed.)

Note that nonoxidizing solutions, such as ammonium citrate, hydroxyacetic–formic acid mixtures, and EDTA (ethylenediaminetetraacetic acid) instead of a nitric acid passivation are described in ASTM A 380, Part III. The dairy industry uses a 0.5–1.0% phosphoric acid solution at 65–82 °C (150–180 °F) for 10 min. These nonoxidizing solutions used to remove iron or iron oxides only equilibrate the surface, facilitating autopassivation. They do not actually enhance the film as do the true passivation procedures.

Sanitizing

As a final step (or between batch operations), it is frequently desirable or even necessary to sterilize stainless steel equipment. Sodium or calcium hypochlorite is most frequently utilized, but the oxidizing capacity and chloride content pose serious potential dangers for pitting or crevice corrosion. The conditions and time of exposure and the proper steps to flush the hypochlorites from the system, including low spots, is of paramount importance. Serious failures have occurred in numerous stainless steel installations when adequate measures to preclude attack during treatment or by residual hypochlorites were not taken. Treatment with ozone or dilute hydrogen peroxide, as well as by ultraviolet light, seems to provide useful alternatives.

STANDARD OPERATING PROCEDURES

To summarize, the recommended sequence of treatment of fabricated stainless steel equipment for pharmaceutical applications should embrace the following steps:

1. A pretreatment with solvent or detergent to remove oil, grease, and the like
2. A rinse with water of suitable quality (deionized water being preferred)
3. Necessary pickling or other cleaning operations
4. Passivation
5. Water rinsing
6. Sanitizing
7. Final water rinse; drain and dry

A proper understanding of service conditions and requirements is essential if the several possible cleaning, pickling, and passivation treatments are to be successfully employed for their respective end purposes.

The most valuable resources for practical information on corrosion and corrosion prevention of stainless steels are as follows:

- NACE International (standards and publications), Houston, Tex.
- The Nickel Development Institute (AISI, INCO, and NiDI publications), Toronto, Ontario, Canada
- ASTM International (specifications), Barr Harbor, Pa.
- The Materials Technology Institute of the Chemical Process Industries, Inc. (publications), St. Louis, Mo.

BIBLIOGRAPHY

1. *Metals & Alloys in the Unified Numbering System*, SAE HS-1086, ASTM DS-56 E, 6th ed., American Soc. for Testing and Materials, Philadelphia, 1993.
2. C.P. Dillon, *Corrosion Resistance of Stainless Steels*, Dekker, New York, 1995.
3. C.P. Dillon, D.W. Rahoi, and A.H. Tuthill, *BioPharm* **5**, 3–6 (1992).
4. C.P. Dillon, *Corrosion Control in the Chemical Process Industries*, 2nd ed., Publ. 45, MTI, St. Louis, Mo., 1994.
5. *Guidelines for Preventing Stress Corrosion Cracking in the Chemical Process Industries*, Materials Technology Institute of the Chemical Process Industries, St. Louis, Mo., 1985.
6. P.H. Banes, *Bioprocess Eng.* (1993).
7. D.C. Coleman and R. Evans, *Pharm. Eng.* **11** (1991).
8. T. DeBold, *Machine and Tool Blue Book*, November 1986.

STATIC MIXING, IN FERMENTATION PROCESS

RADU Z. TUDOSE
Technical University Gh.Asachi Iaşi
Iaşi, Romania

MARIA GAVRILESCU
Research Centre for Antibiotics
Iaşi, Romania

KEY WORDS

Fermentation
Heat transfer
Hydrodynamics
Mass transfer
Nonmechanical bioreactors
Non-Newtonian systems
Static mixer

OUTLINE

INTRODUCTION

The intensification of transfer processes in bioreactors using techniques that do not require additional energy costs is of great interest in biotechnology. Traditionally, most bioprocesses, especially aerobic fermentations, are conducted in stirred-tank bioreactors, which require considerable power input to attain adequate oxygen transfer rates. Besides the conventional mechanically agitated fermenters, other types of gas–liquid contactors are used, as pneumatic or jet bioreactors with or without a loop. Advantages of these nonmechanically agitated bioreactors are known to include efficient use of energy for aeration, simple design, and construction. To improve transfer processes in these bioreactors, considerable efforts have been made to study all the essentials of the major relationship for flow, mass and heat transfer over plates, cylinders, baffles, and spheres as well as through sieves and trays. These are the simplest and the oldest contacting bodies, and also are the basic shapes for other contacting devices of practical interest. Both theoretical and experimental studies on flow around these internals together with heat and mass transfer approaches are available (1–4).

Recently, new constructive types of internals known as static mixers or motionless mixers have been developed that supply spectacular effects concerning transfer rate increase. The first of these unconventional mixers was developed for viscous fluids about 30 years ago (5). After a slow start, the development of static mixers has been proceeding since 1970. Static mixers are applied in many operations in heterogeneous systems that require larger contact areas, thermal homogenization, and mass transfer. Today, these devices placed in column bioreactors are increasingly considered for use in fermentation industries, sometimes replacing the traditional mechanically agitated bioreactor design. Static mixers are suitable particularly for use in shear-sensitive cell cultures because of their lower shear rates, compared to those in mechanically agitated bioreactors. They proved to consider some specific criteria in the design of a bioreactor system for a definite process: population damage, energy distribution, and oxygen demand of a culture. The relatively numerous investigations and works that refer to static mixers are concerned with their influence on different process transfer parameters not only in fermentation systems, most of them being performed in model systems, in monophasic and polydispersed laminar or turbulent flow (usually gas–liquid dispersions), using low and viscous liquids, with Newtonian or non-Newtonian rheological behavior.

Static mixers frequently require no additional space (in-line assembly) and can be used as compact inserts or with a distance between the elements. They act on hydrodynamic conditions for turbulence development, the increase of contact areas, or the intensification of driving forces, separately or simultaneously. In all cases, the energy for process intensification is extracted from the fluid flowing through the mixers.

STRUCTURAL TYPES AND CONSTRUCTION

Static mixers can be of numerous constructive types, made of tubes, sheets, helixes, or wire matrix. They can be manufactured by steel, ceramic, or plastics. Their effect depends on the construction of the mixing element and the capacity to produce continuous separation, distribution, and reuniting the fluid stream. Besides their multiple advantages, static mixers often have the difficulty of a complicated structure. At the present time, most of the problems in the construction of static mixers have already been solved, and a number of different types of static mixing systems are being manufactured on an industrial scale, so that about 30 types of static mixers are known today. One of the most important moments in static mixing presentation was the ACHEMA '79 Exhibition in Frankfurt/Main (5). This section discusses some of the most interesting, familiar and, at the same time, most structurally dissimilar static mixers.

Static Mixers from Tubes or Short Pipes

Static mixers from tubes or short pipes generally affect the spatial distribution of the fluid by twisting, distorsion, and expansion of fluid elements in specially designed flow arrangements. For instance, the Ross-ISG static mixer (Aachener Misch- und Knetfabrik) is made by elements turned through 90°, placed one behind other. Over the cross-section of an element there are four obliquely sloped channels twisted with an angle of 120° (Fig. 1**a**), forming a tetrahedral space between elements. The PRM mixer (Prematechnik, Frankfurt/Main) consists of a tubular packet containing obliquely arranged partial tube sections (Fig. 1**b**), which change the direction of the fluid stream (5). Another static mixer used in laminar flow conditions, especially to equalize the distribution of the flow residence time, consists of six guiding holes, directed in the direction of flow (Fig. 1**c**). A portion of the fluid stream, which flows through the central section in front of the device, has a

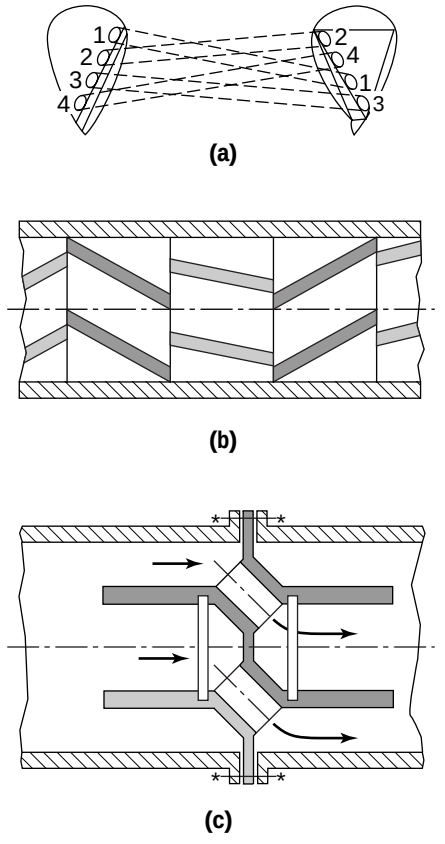


Figure 1. Static mixers made from tubes or short pipes. *Source*: Reprinted from Ref. 5, with permission from VCH Verlagsgesellschaft mbH, and from Ref. 7, with permission from Wiley-VCH.

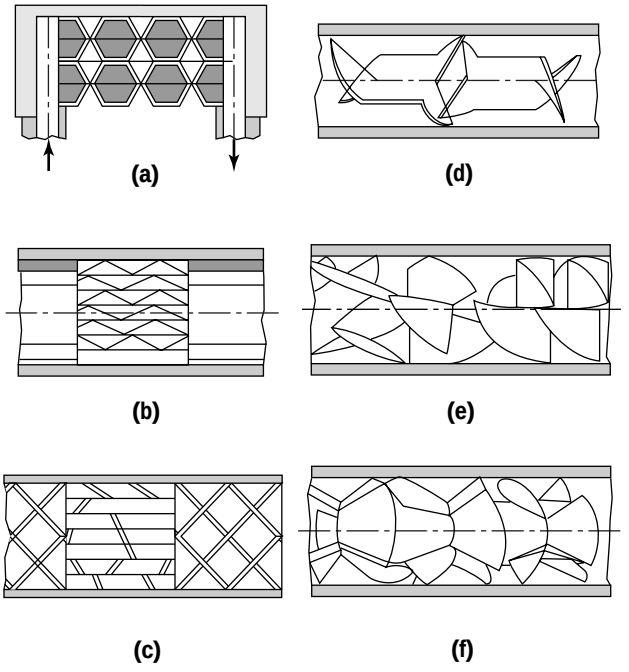


Figure 2. Static mixers made from plates and sheets. *Source*: Reprinted from Ref. 5, with permission from VCH Verlagsgesellschaft mbH and from Ref. 7, with permission from Wiley-VCH.

higher mean velocity and crosses the channels that direct it to a zone with a low mean velocity. Other parts of the fluid stream that flow near the pipe wall cross over the channels that direct them to the central zone of the pipe behind the static mixer. Therefore, the positions of the fluid streams that flow with different velocities are inverted (6–8).

Static Mixers of Plates or Sheets

The effects of these static mixers depend on the continuous cutting, subdividing, distribution, and reuniting the stream of fluid. The PSM mixer (Petzhold, Frankfurt/Main) (Fig. 2**a**) contains rectangular and conical channels combined with conical holes resulted from a special arrangement of individual plates. The fluid flows from the entry to the exit of the static mixing element, being subdivided through a central hole, a conical space created by tangentially arranged slots and furthermore through other channels and openings (5). Sulzer mixers (Sulzer Brothers Ltd., Switzerland) are adequate for both laminar and turbulent regimes. They are constructed in many different ways (SMV, Fig. 2**b**; SMX, Fig. 2**c**) from stacks of corrugate sheets with open crossing channels. The SMX mixer consists of strips inclined to one another that form frameworks with the axis perpendicular to the column axis. The mixing elements are arranged successively in the column turned through 90° one to another. The SMLX mixer type is constructed from corrugate sheets with a distance between the strips increasing in the flow direction. The crossing angle of the sheets is 60°. These static mixers made of corrugate sheets can be mounted in packed form in compact arrangement or with distance between packets in different nonmechanical bioreactors (9). The Komax mixer (Long Beach, Calif.) contains triangular-shaped channels, made by the bent ends of the slotted sheets that form the mixing element (Fig. 2**d**). The Koch mixer (10,11) contains a numer of in-line elements mounted alternately at right angles to each other. The mixing elements are made from layers of corrugate metal sheets joined one to another, forming opened intersecting channels. C_1 and C_2 mixer types (12) (Figs. 2**e** and **f**, respectively) developed in the Central Institute of Physical Chemistry of the German Academy of Sciences and the Pyrapak 480 type developed at the CLG Leipzig-Grimma were also manufactured out of metal sheets.

Helical Static Mixers

The Kenics mixer (Ott Vertriebsgeselschaft, Leonberg) is made of counterclockwise and clockwise rotational elements twisted at 180°, being mounted in tubes alternating left and right with perpendicular leading edges (Fig. 3**a**). A particular form of this helical static mixer consists of an unbroken continuous helical surface that may have the diameter of a screw identical or smaller than the inner di-

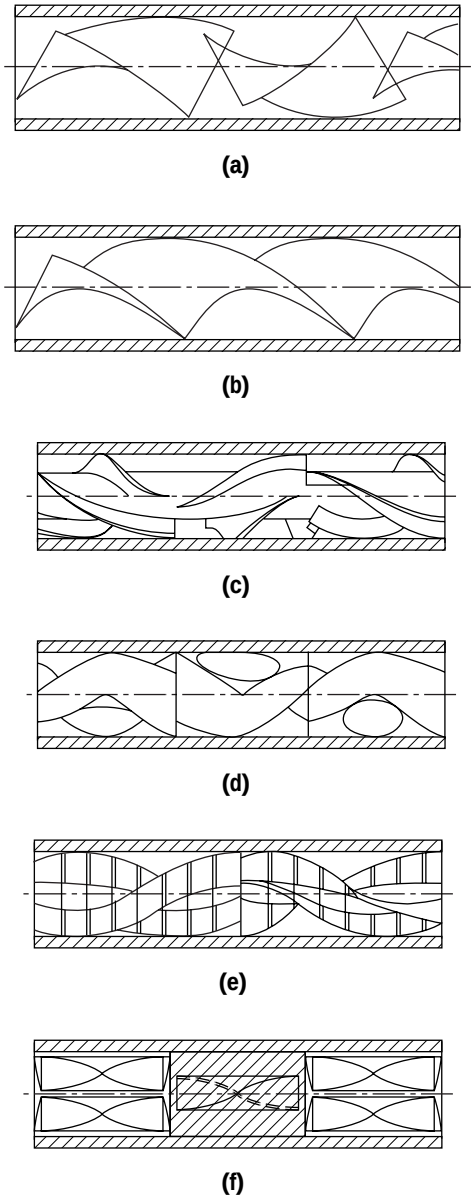


Figure 3. Helical static mixers. *Source*: Reprinted from Ref. 5, with permission from VCH Verlagsgesellschaft mbH, and from Ref. 7, with permission from Wiley-VCH.

ameter of the tube (13) (Fig. 3**b**). Helical static mixers can be disposed in columns as in-line elements or in packed form (14). The Erestat mixer (Yxzet, Dr. Karg GmbH, Affalterbach) consists of four baffles twisted in an alternate manner arranged in a left- or right-handed arrangement on an inner support (Fig. 3**c**). This assembly creates various flowing channels. The N-shaped mixer (Bran and Lübbe, Norderstadt) consists of twisted mixing elements (Fig. 3**d**) that determine variable cross-sectional areas, owing to the formation of variable channels. The Lightin type (Lightin, Rochester, NY) combined in-line mixers obtained by assembling three-bladed elements to form helixes with different senses of rotation (Fig. 3**e**). The HI mixer (Toray, Tokyo) is constructed by helixes twisted at 180°, fitted into holes performed parallel to the axis in a solid cylinder (Fig. 3**f**).

Wire Matrix Turbulators

The Heatex inserts (15) consist of a central wire core onto which are wound a series of wire loops. Each loop is inclined at a certain angle to the core. These inserts are pulled into a tube, the loops being so inclined as to be oversized for the tube and come into a close contact with the tube wall in the form of an arc (Fig. 4). A degree of resistance is offered as the insert is pulled into the tube, indicating that some wall contact is achieved and that the insert is self-supporting. Heatex inserts may have different loop densities, characterized by the number of tube wall contacts achieved per unit insert length.

EFFECTS OF STATIC MIXERS ON MOMENTUM TRANSFER

Flow Structure

Static mixers produce variations in the flowing cross-sectional area that result in modifications of velocity profiles. The fluid is split into individual streams at the open intersecting channels of the mixing element. At each intersection, a partial quantity is sheared off into a crossing channel and stagnant zones are usually eliminated. Also, most of the static mixers divide and reunite continually the main fluid stream and give rise to the formation of elementary layers of fluid. The number of n elementary layers or partial streams produced in laminar flow over a static mixer resulted through assembly of N mixing elements can be calculated from the following expression (16):

$$n = ca^{N} \tag{1}$$

where c and a are parameters characteristic for the static mixer in discussion that are usually given in the literature (16). In the laminar regime, the flow profiles are modified by cutting and twisting, displacement and distortion, or separation and expansion. In the turbulent flow, additional effects occur (5). For instance, the fluid streams are alternating directed from the center to the wall of the column

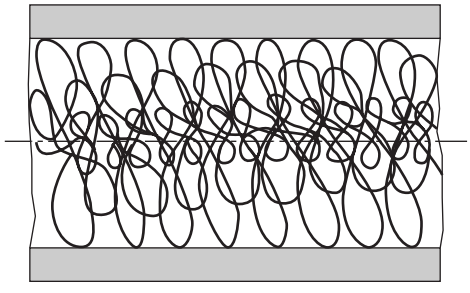


Figure 4. Static mixers of wire matrix. *Source*: Reprinted from Ref. 7, with permission from Wiley-VCH.

and from the wall to the center in the Ross-ISG mixer (Fig. 1**a**) or in the static presented in Fig. 1**c**. The presence of Heatex inserts (Fig. 4) causes a tube wall–fluid disturbance, and radial mixing is expected to be produced as the fluid flows through the mesh of loops nearer the pipe axis (15). In a Kenics mixer (Fig. 3**a**) a tracer injected below the first element in a certain point can perform a helical rotation strictly geometrically and reach the element edge in an opposite point situated on the same diameter, or it flows along a shortened path after a distortion and arrives at the end of the helix in a intermediate position on the diameter (5). The flow pattern in a Sulzer SMV mixer was presented by Pahl and Muschelknautz (5) using a model that consists in two triangular opened ducts crossing at 90° (Fig. 5) forming two intersected channels. Tracer introduced at the entrance of the channel I flows up to the intersection of the channels, and then about two-thirds of the tracer stream takes the shortest path flowing into duct II owing to the displacement that is produced at the crossing point. The trajectory of the tracer is also distorted; the fluid that exists along the left cathetus of the triangle representing the cross-section of the duct I shifts to the right side of duct I at the exit, performing a rotation. Also, the stream that starts from the opposite position at the entrance of duct I is greatly distorted at the crossing point and flows into both ducts I and II. In gas–liquid dispersions, a reasonably uniform distribution of the gas velocity over the cross-section of a column with static mixers is one of the favorable results. An equalizing effect is achieved by the resistance to flow, which forces the gas to spread out gradually in the direction of flow. This fact is very important in bioreactors with viscous non-Newtonian fermentation fluids, where the bubbles coalesce in the sparged vertical columns, leading to the formation of large bubbles, even at very low gas flow rates. In absence of static mixers, these bubbles rose with large velocities and their size distribution is controlled by bubble coalescence. The velocity profile of a fluid flowing through a tube containing motionless mixers is also considerably altered. There are some works in the literature that give information about radial and axial velocity profiles in the presence of static mixers (13,17,18). The axial and radial velocity profiles in a Kenics mixer change periodically owing to the use of alternating elements of right- and left-hand pitch that results in more intensive mixing of fluids. The screw having the form of a continuous helical surface divides the inside of the tube into two parallel channels, which causes a secondary motion of the fluid in the tube (13). In these motionless mixers, circulating flow occurs in the r-θ plane (r and θ being radial and tangential coordinates, respectively), which is caused by a slow twist in the partitioning chord so that it forms a helix. The actual velocity profiles can be calculated using a complex procedure (17), but an approximation can be made, leading to the following expressions:

$$v_r = \frac{dr}{dt} = \beta r(1 - r^{\upsilon}) \sin 2\theta \tag{2}$$

$$v_\theta = \frac{d\theta}{dt} = \beta[2 - (2 + \upsilon)r^{\upsilon}] \sin^2\theta \tag{3}$$

where

$$\upsilon = \left(\frac{11}{3}\right)^{0.5} - 1 \tag{4}$$

and β is a parameter that governs the intensity of circulation (17).

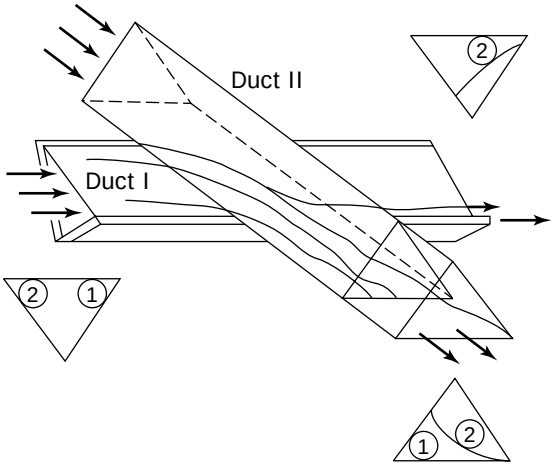


Figure 5. Flow path in a Sulzer-SMV static mixer (two triangular channels crossing at 90°). *Source*: Reprinted from Ref. 5, with permission from VCH Verlagsgesellschaft mbH, and from Ref. 7, with permission from Wiley-VCH.

Pressure Drop

The changes in the velocity field caused by the static mixers cause modifications in the pressure loss as well. The pressure drop in static mixers depends very strongly on the geometrical arrangement of the packed elements, the nature of the flowing fluids, and the flow regime (19,20). A simple form representing the pressure loss equation can be referred to the empty tube, together with a geometrical correction factor (5,7,13). For Newtonian media, the equation for calculating the pressure drop is in the following form:

$$\frac{\Delta p_M}{\rho v_{SL}^2}\frac{2D}{L} = f_D(Re_D, \text{geometry}) \tag{5}$$

The mean liquid superficial velocity $\mathbf{v}_{SL}$ can be calculated on the basis of the empty tube cross-section as follows:

$$v_{SL} = \frac{4\dot{V}_L}{\pi D^2} \tag{6}$$

In laminar flow, $f_D\ Re_D$ = constant. Rearranging equation 5, a relationship is obtained whose denominator becomes the pressure drop for laminar flow through an empty pipe:

$$\frac{\Delta p_M}{\Delta p_L} = \frac{\Delta p_M}{32\eta v/D^2} \tag{7}$$

Defining a pressure drop factor as

$$z = f_D \frac{Re_D}{64} \tag{8}$$

it follows that

$$\Delta p_M = z\Delta p_L \tag{9}$$

In Figures 6 and 7, the pressure drop and friction factors for different static mixers are presented (5,7). The plot

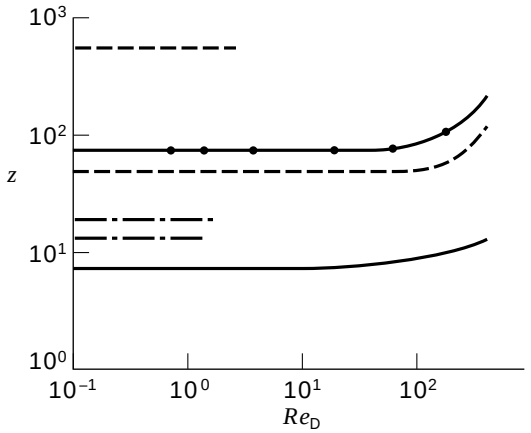


Figure 6. Relative pressure drop for different static mixers versus Reynolds number. — = Kenics; —·— = Erestat; —··— = N-shaped; — — — = HI; — □ — □ — = Sulzer-SMX; - - - - - - = Ross-ISG. *Source*: Reprinted from Ref. 5, with permission from VCH Verlagsgesellschaft mbH, and from Ref. 7, with permission from Wiley-VCH.

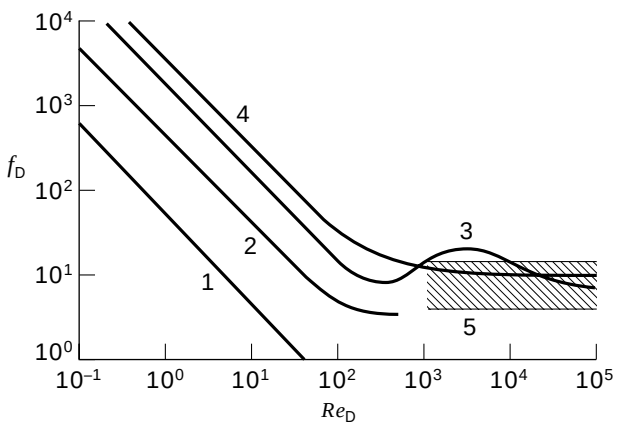


Figure 7. Friction factor as a function of Reynolds number for some static mixers. 1 = empty pipe; 2 = Kenics; 3 = HI; 4 = Sulzer-SMX; 5 = Sulzer-SMV. *Source*: Reprinted from Ref. 5, with permission from VCH Verlagsgesellschaft mbH, and from Ref. 7, with permission from Wiley-VCH.

$f_D - Re_D$ could be also used for comparing static mixers in the turbulent flow regime, since f_D is then approximately independent of the Reynolds number (Fig. 7). Compared with an empty tube, the pressure drop in static mixers is 7 to 200 times greater for laminar flow and 100 to 600 times greater for turbulent flow. Kenics mixers show the smallest pressure drop, and the Ross-ISG affords the largest pressure drop (5,16). Also, the pressure losses in a tube containing continuous screws calculated using equation 5 are lower than those corresponding to a tube containing Kenics mixers for the same value of the parameter h/d. The following analytical expressions were developed for the different types of helical static mixers (13). For Kenics type (for $15 \le Re_D \le 1{,}200$):

$$f_D = \frac{213.5 + 224.0(h/d)^{-1}}{Re_D} + [-0.549 + 4.775(h/d)^{-1}] \tag{10}$$

For static mixer with a continuous screw surface for $15 \le Re_D \le 90$,

$$f_D = \frac{191.3 + 229.0(h/d)^{-1}}{Re_D} \tag{11}$$

for $90 \le Re_D \le 5{,}000$,

$$f_D = \frac{79.8}{Re_D^{0.839 - 0.454(h/d)^{-1}}} \tag{12}$$

With increasing of h/d, the pressure loss decreases for any given type of fittings (Fig. 8). The specification of notation is given in Table 1.

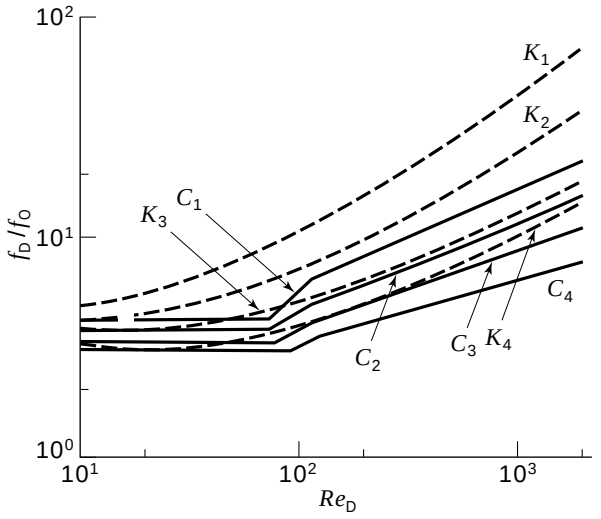


Figure 8. Dependence of the ratio between the friction coefficient in a tube with a screw f_D and that in the same tube without a screw f_0, on the Reynolds number, Re_D, and h/d (specification of C and K in Table 1). *Source*: Reprinted from Ref. 13, with permission from Elsevier Science, and from Ref. 7, with permission from Wiley-VCH.

Table 1. Geometrical Characteristics and Notations for Helical Static Mixers

Type of screw	Notation	h/d	s/d	l/d
Continuous screw	C_1	2.28	0.095	46.85
	C_2	2.95	0.095	47.57
	C_3	3.85	0.095	47.57
	C_4	4.71	0.095	47.69
Kenics	K_1	2.02	0.095	48.26
	K_2	3.03	0.095	49.14
	K_3	4.05	0.095	47.86
	K_4	5.02	0.095	47.86

Source: Reprinted from Ref. 7, with permission from Wiley-VCH.

In Newtonian liquids, the presence of wire matrix turbulators caused an increase in pressure drop by factors up to 20 compared with the empty tube case (15). In non-Newtonian fluids, the pressure drop in a static mixer assembly can be correlated as follows:

$$\Delta p_M = \frac{A}{Re_D} + \frac{B}{Re_D^m} \tag{13}$$

where the values of A, B, and m are dependent on rheological behavior of the liquid phase (7,21). Olivier and Oosterhuis (22) studied the pressure drop in the laminar flow regime for a Sulzer-SMV-type in the xanthan fermentation and found that the pressure drop is approximately 15 times greater than that in an empty tube. Fan et al. (23) studied the pressure drop in a gas–liquid system in a bubble column packed with Sulzer-Koch mixing elements and observed that the ratio of the pressure drop through the packed column to that of the empty one was slightly greater than 1. Data referring to the pressure drop in a column equipped with a static mixer of the Kenics type were obtained for a single (gaseous) phase and two-phase countercurrent flow, respectively (14,24). For one mixing element, the pressure drop, which is a function of the Reynolds number for the gaseous phase, Re_G, is assessed as follows:

$$\Delta p' = \frac{\Delta p_M}{N} \tag{14}$$

The pressure drop in the system was calculated using the gas dynamic pressure for the single phase as

$$\Delta p' = f_{Do}\rho \frac{v_{SG}^2}{2} \tag{15}$$

and for the diphasic flow:

$$\Delta p'_{TP} = f_{DTP}\rho \frac{v_{SG}^2}{2} \tag{16}$$

The plots show three domains with different slopes (Fig. 9). The friction factor also differs in the domains. For a single phase flow, it was correlated as follows:

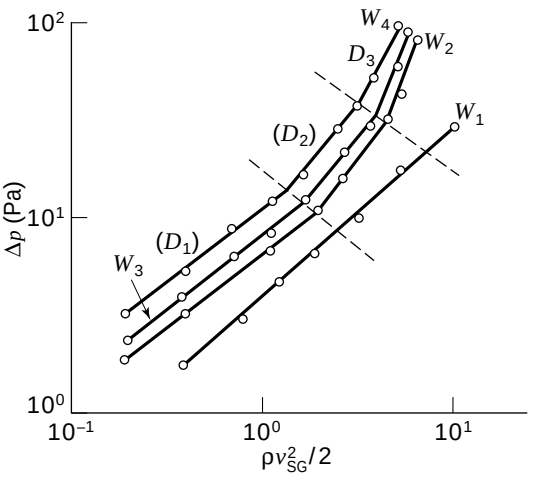


Figure 9. Pressure drop through a single static mixing element as a function of the dynamic pressure of the gaseous phase for different spraying intensities. $W_1 = 0$; $W_2 = 60.93$ m^3/m^{2}h; $W_3 = 81.24$ m^3/m^{2}h; $W_4 = 101.56$ m^3/m^{2}h. *Source*: Reprinted from Ref. 7, with permission from Wiley-VCH.

$$\text{Domain D}_1: \quad f_{Do} = 12.86 Re_G^{-0.1708} \tag{17}$$

$$\text{Domain D}_2: \quad f_{Do} = 1.039 Re_G^{0.1062} \tag{18}$$

$$\text{Domain D}_3: \quad f_{Do} = 0.9742 Re_G^{0.1137} \tag{19}$$

For a diphasic flow, the following relation was established (24):

$$\frac{\Delta p'_{TP}}{\Delta p'} = B'^{W(1+b'Re_G)} \tag{20}$$

The values of the spraying intensity, B', and of the parameter b', are different in the three domains of the curves (Table 2) (24).

For a two-phase flow, Shah and Khale (21) applied a correction factor for the manometer reading, assuming the gas holdup to be proportional with $\dot{V}_G/(\dot{V}_G + \dot{V}_L)$. For an empty tube, the correction factor was found to be 0.83. They presented their data in the form of a Lockhart-Martinelli correlation for three types of static mixers, Kenics, Sulzer, and Komax, for two-phase non-Newtonian liquid–air systems in the following form:

$$\Delta p_{TP} = \varphi_L^2 \Delta p_L = \varphi_G^2 \Delta p_G \tag{21}$$

$$\varphi_L^2 = 1 + (c'/X) + (1/X^2) \tag{22}$$

$$X = (\Delta p_L/\Delta p_G)^{0.5} \tag{23}$$

with $c' = 2.5 \pm 0.3$ for glycerol solution–air and $c' = 3.27 \pm 0.1$ for CMC solution–air diphasic systems, respectively. Generally speaking, the pressure drop in the gas–liquid systems decreases as the gas flow rate increases (10).

Gas-Holdup

In tower bioreactors with or without a loop (bubble columns, airlifts, jet contactors), liquid flow and gas holdup

Table 2. The Values of Parameters B' and b' from Equation 20

Domain	B'	b'
1	1.0150	$-5.596 \cdot 10^{-5}$
2	1.00925	$2.2348 \cdot 10^{-5}$
3	$0.947 + 2.6217 \cdot 10^{-4}$ W	$-8.5243 \cdot 10^{-9}\,W^2 + 8.262 \cdot 10^{-7}$ W $- 1.4499 \cdot 10^{-4}$

Source: Reprinted from Ref. 7, with permission from Wiley-VCH.

are of basic importance in the determination of their performance (25). Because microbial aerobic cultures often behave as non-Newtonian fluids, much attention was directed toward the performance of these gas–liquid contactors equipped with static mixers, especially when they handle non-Newtonian fermentation liquids. The rheological behavior of biofluids plays an important role in bubble coalescence and breakup, because the large velocities of coalesced bubbles lead to gas holdup decreasing (26). The presence of static mixers in diphasic flow diminishes the bubble coalescence rates owing to their shear effect. Therefore, the average bubble size should decrease and gas holdup is enhanced (7,10) (Fig. 10). The results of MacLean et al. (27) confirm these findings, suggesting that the motionless mixers help to retain more gas in the column. Similar effects can be obtained in triphasic sparged reactors (28). By using Ross-LPD and Sulzer SMV static mixers, larger gas holdup values resulted, compared with the empty contactor (Fig. 11). Kenics mixers mounted in an airlift bioreactor contribute to increase gas holdup values during the batch fermentation of a yeast culture (29). Gas holdup in downcomer (annular section) decreased because of a lower liquid recirculation rate, which implied a more intense gas disengagement at the upper part of the bioreactor, reducing the quantity of the recirculated gas. In the region of greater airflow rates (the riser), gas holdup values are, however, improved. This behavior had a favorable influence on oxygen transfer rates (29,30). The presence of static mixers does not change the correlation between gas holdup and gas superficial velocity, which was established in the following form:

$$\varepsilon_G = a v_{SG}^b \tag{24}$$

For example, static mixers manufactured through assembly of corrugated stainless steel sheets and mounted in the riser section in an external-loop airlift bioreactor can contribute to gas holdup intensification at any given gas flow rate and liquid phase properties, relative to bioreactor configuration without static mixers (Fig. 12), also proved by the modified values of the parameters a and b in equation 24. This effect is found in low-viscous Newtonian and viscous non-Newtonian liquids as well (Table 3) (25). It is evident from Table 3 that static mixers show an increased efficiency in more viscous liquids. Also, they affect the values of the coefficient a in equation 24 in a greater measure than those of b. Therefore, larger values of gas holdup in columns with static mixers resulted because of smaller bubble sizes and hindering action of the packing, which contribute to increasing the gas phase residence time (21,31).

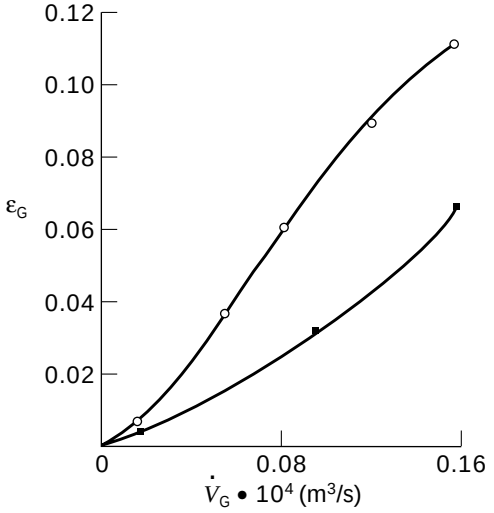


Figure 10. Gas holdup in a Koch mixer column in comparison to the witness one. ● = without static mixers; ○ = with static mixers. *Source*: Reprinted from Ref. 7, with permission from Wiley-VCH.

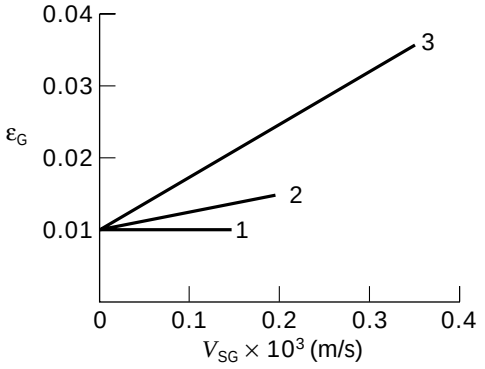


Figure 11. Gas holdup in a triphasic sparged reactor. 1 = witness reactor, 2 = Ross-LPD static mixers; 3 = Sulzer-SMV static mixer. *Source*: Reprinted from Ref. 7, with permission from Wiley-VCH.

Liquid Velocity

The presence of static mixers changes the liquid circulation rates, because they act as supplementary resistances in the flow path. Besides changes in gas flow pattern, the presence of static mixers in gas–liquid contactors having a net liquid flow (continuous bubble columns, airlifts, con-

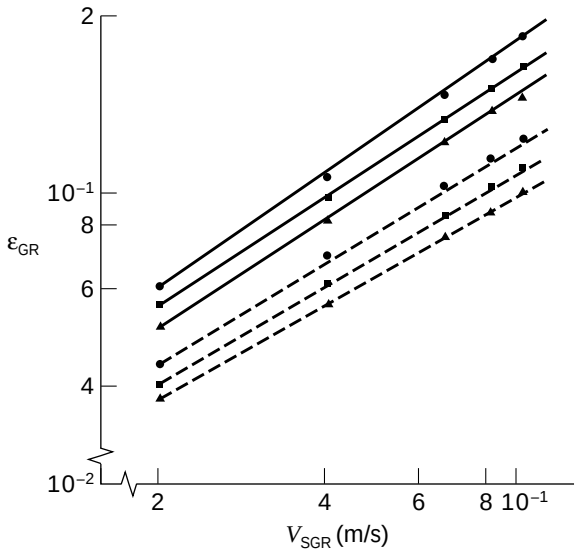


Figure 12. Gas holdup in an external-loop airlift bioreactor without (--------) and with (▬▬) static mixers, in non-Newtonian starch solutions. ● = 0.5% w/v × 2.5 w/v; ▲ = 4.0 w/v. *Source*: Reprinted from Ref. 25, with permission from Springer-Verlag, Germany.

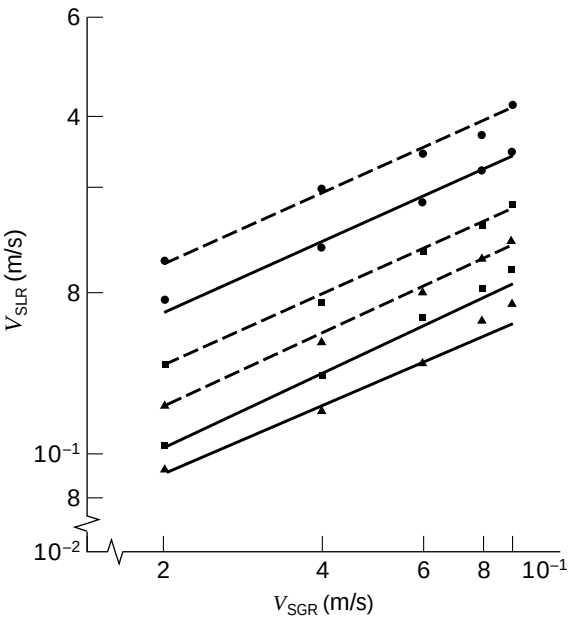


Figure 13. Liquid superficial velocity versus gas superficial velocity in an external-loop airlift bioreactor without (--------) and with (▬▬) static mixers. Symbols as in Figure 12. *Source*: Reprinted from Ref. 25, with permission from Springer-Verlag, Germany.

Table 3. Ratios Between Parameters *a* and *b* in Equation 24 in the Configuration with and without Static Mixers for an External-Loop Airlift Bioreactor

Ratio between parameters in equation 24	K (Pa s^n)		
	0.061	0.438	3.518
a_M/a	1.600	1.845	1.911
b_M/b	1.085	1.159	1.134

Source: Reprinted from Ref. 25, with permission from Springer Verlag.

Table 4. Ratios between the Parameters *f* and *m* in Equation 25 in an External-Loop Airlift Bioreactor with and without Static Mixers, in Non-Newtonian Liquids

Ratio between parameters in equation 25	K (Pa s^n)		
	0.061	0.438	3.518
f_M/f	0.829	0.841	0.856
m/m	0.997	0.994	0.994

Source: Reprinted from Ref. 25, with permission from Springer Verlag.

tactors with a liquid jet) changes the liquid circulation velocities, their effects being complementary to those of several other parameters, including gas superficial velocity, gas holdup, liquid phase viscosity, pressure drop along the flow path, reactor geometry. For instance, the static mixers mounted in an external-loop airlift bioreactor diminish the liquid superficial velocity in the riser section both in Newtonian and non-Newtonian liquids. The parallel lines in the diagram presented in Fig. 13 show that in all cases the liquid velocity increases with gas superficial velocity. This behavior can be expressed by the following mathematical expression as:

$$v_{SL} = f v_{SG}^{m} \tag{25}$$

In non-Newtonian liquids, it was found that the presence of static mixers affects f and m values, dependent on consistency index values (25) (Table 4).

The relationship between the ratio of f values with and without static mixers, respectively, and the consistency index in an airlift contactor can be represented by the following relationship (25):

$$\frac{f_M}{f} = 0.85K^{0.008} \tag{26}$$

The effects of the consistency index on the liquid velocity could be considered as being less significant proving that static mixers are effective in viscous liquids. Popovic and Robinson (32) found that the circulating liquid velocity in the riser section of an airlift reactor is dependent mainly on the downcomer-to-riser cross-sectional areas ratio, A_D/A_R, the apparent viscosity, η_{ap} and gas superficial velocity, as described by the following correlation:

$$v_{SLR} = c' v_{SGR}^{0.32} (A_D/A_R)^{0.97} \eta_{ap}^{-0.39} \tag{27}$$

Considering that equation 27 correlates well a large number of experimental data, it was found that the differences in the values of v_{SLR} should be treated as particular cases of diminuation up to 40% of the geometrical parameter A_D/A_R, as a result of the presence of static mixers (25).

Mixing

As shown, homogenization in static mixers is attained by means of motionless elements using the flow energy of the fluids. The mixing effect depends on the continuous separation, distribution, and reuniting streams of material (5,7). The energy requirements of static mixers when used for mixing are lower than those of conventional dynamic mixers and, moreover, they are of simple construction. In the case of two components, the two quantities are related to each other by the volumetric flow rate ratio, $\dot{V}_2/\dot{V}_1$. The degree of mixing may be described using the variation coefficient $s/\bar{C}_V$ or the relative standard deviation s/τ_0 (5). For component 1, the following expression held:

$$s_1/\bar{C}_V = \sqrt{\dot{V}_2/\dot{V}_1}\, s_1/\tau_0 \tag{28}$$

where

$$s_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(C_{V_i} - \bar{C}_{V_1})^2 \tag{29}$$

$$\tau_0 = [\bar{C}_{V_1}(1 - \bar{C}_{V_1})]^{0.5} \tag{30}$$

$$\bar{C}_{V_1} = \frac{1}{n}\sum_{i=1}^{n} C_{V_{1,i}} \tag{31}$$

$$C_{V_{1,i}} = \dot{V}_{1,i}/(\dot{V}_{1,i} + \dot{V}_{2,i}) \tag{32}$$

The degree of mixing is strongly dependent on the static mixer type and normalized mixing length, L/D (Fig. 14). It can be seen that the Sulzer-SMX mixer and HI-mixer require shorter homogenization lengths. Also, the degree of mixing is dependent on the entry point of fluids to be mixed as well as by the operation scale (5). Static mixers achieve rapid mixing of fluids with low energy requirements, and the fluids may have either similar or different viscosity. Operation in fermentation processes involves very viscous fluids, and the flow will be predominantly in the laminar region for both liquid and gaseous phases. Suitability of static mixers to mix fluids of different viscosities was investigated in the literature (12,22,33,34).

Mixing behavior is usually described by mixing time, t_M, which is defined as the time required to achieve a definite degree of homogeneity after a trace pulse has been injected into the reactor (35). The mixing time is highly dependent on the presence of static mixers (34). In gas–liquid dispersions, the mixing time is dependent to a great extent on the gas superficial velocity and the presence of static mixers and to a low extent on the liquid phase properties. In Figure 15 the mixing time in water and starch solutions is plotted against superficial velocity. With static mixers, mixing times at a given superficial velocity are essentially

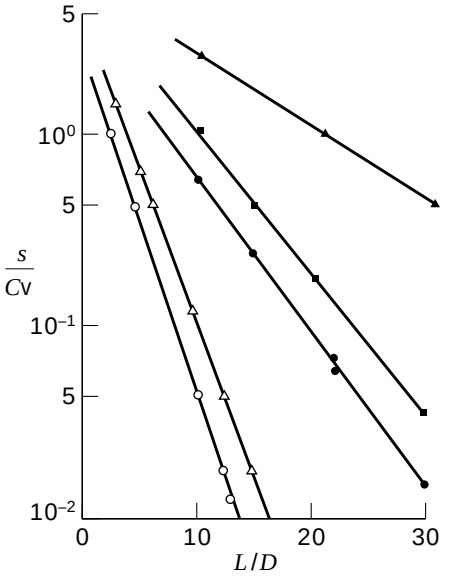


Figure 14. Variation coefficient as a function of the normalized mixing length for various static mixers. ● = Kenics; ○ = Sulzer; ■ = Komax; ▲ = Lightin; △ = HI. $C_{V_1} = 0.1$; $Re_D < 1$. *Source*: Reprinted from Ref. 5, with permission from VCH Verlagsgesellschaft mbH, and from Ref. 7, with permission from Wiley-VCH.

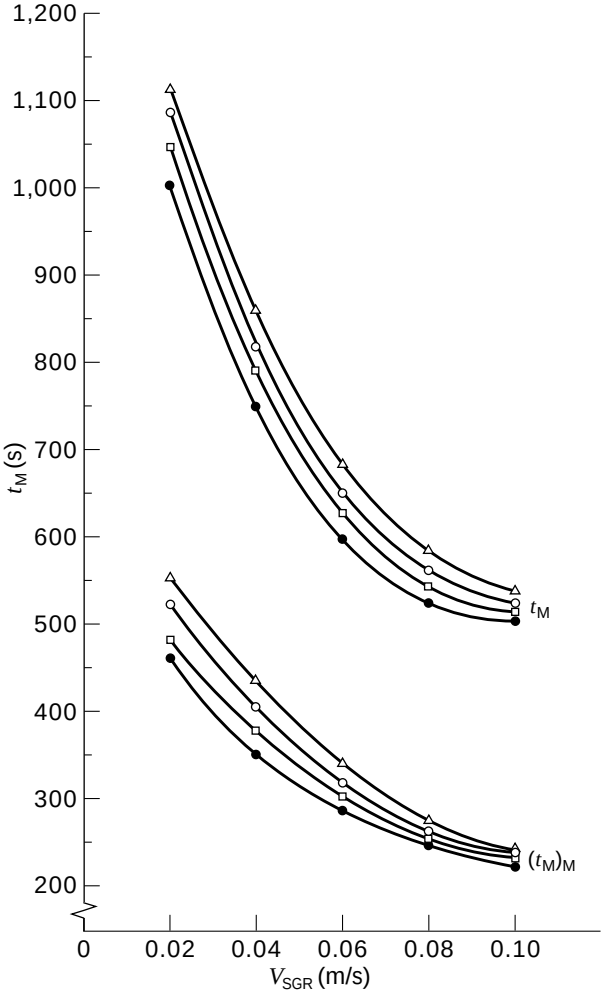


Figure 15. Mixing time versus gas superficial velocity in an external-loop airlift bioreactor without and with static mixers, operating with starch non-Newtonian solutions as liquid phase. ● = water; Starch solutions: □ = 0.5% w/v; ○ = 2.5% w/v; △ = 4% w/v. *Source*: Reprinted from Ref. 34, with permission from Wiley-VCH.

up to 50% shorter than in the reactor without static mixers. The dependence of mixing time on gas superficial velocity can be expressed by the following equation:

$$t_M = k v_{SG}^q \tag{33}$$

The presence of static mixers does not affect q very much, but significantly influences k. Mixing times increase with increasing pseudoplasticity, but not very greatly (Fig. 15). In non-Newtonian systems, the consistency index had no significant influence on these parameters in the presence of static mixers (Fig. 16).

Other properties of a suitable static mixer are good radial mixing on as short as possible flow path and as low as possible axial mixing. Some static mixers contribute to diminish axial dispersion and influence the residence time distribution in laminar monophasic liquid flow, dependent on the Reynolds values (6). In the presence of static mixer seen in Figure 1**c**, the pulse response at the tube exit indicates a smaller axial dispersion (Fig. 17) (6). Also, the dead zones are diminished, and the distribution function of the elements of fluids with ages less than the residence time indicates an evident equalization of the residence times in laminar flow (Fig. 18) (6). In turbulent flow regimes, this effect is not very important. A good radial mixing of a static mixer does not imply in all cases also low axial mixing. The individual static mixers display a different dependence of the axial dispersion on the flow rate of the phases both for single and two-phase flow system. A good static mixer from the standpoint of axial dispersion in single-phase flow does not necessary display a low axial dispersion in two-phase flow (12). Data on axial mixing for the individual types of static mixers are usually missing in the literature. Fialova et al. (12) presented the value of the Péclet number given by Sulzer for its SMX product (*Pé* > 110, d = 0.1 m; L/D = 22). These properties can be ascertained by experimental means.

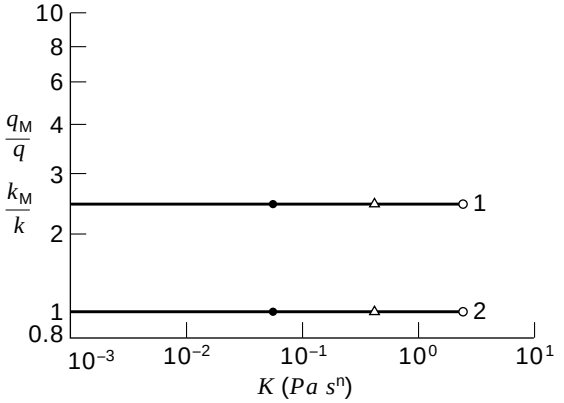


Figure 16. The influence of the consistency index K on the parameters in equation 33. (1 = k_M/k; 2 = q_M/q. Symbols as in Figure 15. *Source*: Reprinted from Ref. 34, with permission from Wiley-VCH).

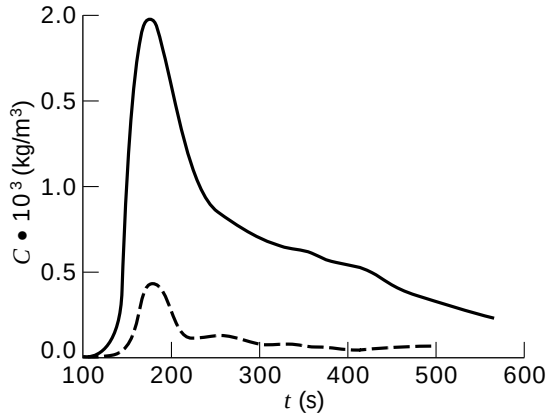


Figure 17. Pulse response at the exit of a tube equipped with the static mixer in Figure 1**c**, - - - - - - = without static mixer; — = with static mixer.

MASS TRANSFER IN PRESENCE OF STATIC MIXERS

The application of the static mixers built in nonmechanically bioreactors, especially in aerobic bioprocesses, should lead to attain the most extensive transport rate of oxygen into the liquid phase at the low possible energetic requirement. The influence of static mixers on hydrodynamic behavior has an important impact on mass and heat transfer characteristics and transfer rate intensification. The improvement of the oxygen mass transfer in bioreactors depends on a number of factors, such as bubble formation, bubble size distribution, bubble shear, gas holdup, rate of coalescence, and breakup. Also, bubble or drop coalescence behavior in gas–liquid dispersions, especially in viscous fluids, demonstrates the necessity for redispersing the gaseous or liquid phase to prevent the reduction of interfacial

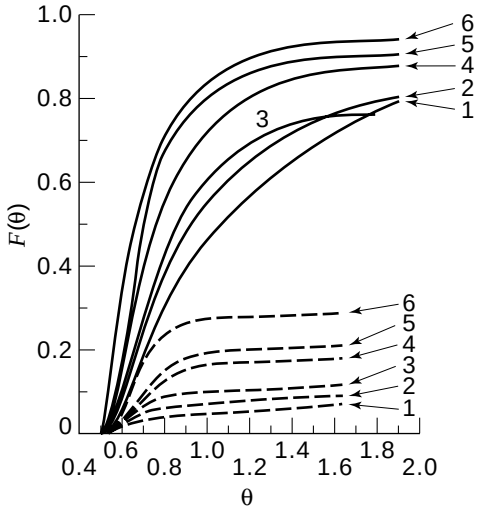


Figure 18. The distribution function of the fluid elements with the age less than the residence time versus Reynolds number in the tube without and with the static mixer from Figure 1**c**. 1 = Re_D = 630; 2 = Re_D = 950; 3 = Re_D = 1,260; 4 = Re_D = 1,580; 5 = Re_D = 1,890; 6 = Re_D = 2,2000. - - - - - - = without static mixer; — = with static mixer.

areas and mass transfer rates. A general solution to the mass transfer problems that appear in gas–liquid systems can be the use of static mixers installed at different intervals for increasing the degree of the gas dispersion and mass transfer rates (7,33,36). Usually, the intensity of oxygen transfer is characterized by its volumetric mass transfer coefficient, *Kla*. Static mixers improve *Kla* at any given gas flow rate relative to the bioreactor configuration without static mixer (7). The mass transfer rates are very high during bubble formation, and the values of mass transfer coefficient vary significantly being associated with large surface renewal rates. Coalescence and break-up, bubble formation, bubble shear, and turbulent liquid surface are examples of conditions where mass transfer coefficient is large. The introduction of static mixers has two important roles in *Kla* enhancement:

1. Bubble redispersion produces increased values of interfacial areas, *a*.
2. The interfacial turbulence during bubble coalescence and redispersion increases *Kl*.

At low gas flow rates, the latter mechanism becomes important (37). Small bubble size can be obtained with a homogeneous bubble distribution over the column cross-section and with a narrow size distribution. However, small rigid bubbles less than 2 mm in diameter have frequently been found to have much smaller mass transfer coefficients (38). Several studies were presented concerning the influence of the Sulzer type (22,33,36,39), Koch (10,11,20,24,40), Kenics (29,30) on *Kla* in bubble columns and external-loop airlift bioreactors in waterlike fluids as well as in non-Newtonian pseudoplastic media. Static mixers significantly enhance the mass transfer rates in waterlike fluids (Fig. 19). In non-Newtonian systems, the improvement in *Kla* due to static mixers depends on the rheological parameters. Stejskal and Potucek (30) and Chisti et al. (39) mounted static mixing elements in the risers of internal- and external-loop airlift reactors that

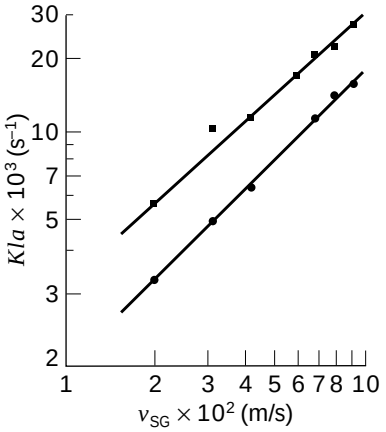


Figure 19. Effects of static mixers on the mass transfer coefficient in aqueous salt solutions. ● = without static mixers, ■ = with static mixers. *Source*: Reprinted from Ref. 7, with permission from Wiley-VCH.

enhanced mass transfer in viscous pseudoplastic CMC aqueous solutions that simulated the rheological behavior of filamentous biofluids. In pseudoplastic liquids, the thicker the fluid was, that is, the higher its value of the consistency index, *K*, the greater the effect of the static mixer was on *Kla* (36,39). Usually, the accumulation of biomass or product in the culture medium leads to increased viscosity and non-Newtonian flow characteristics that contribute to diminish *Kla* (26). This leads to a drastic decrease in oxygen transfer capability, and the aerobic culture eventually becomes oxygen limited (33,41). The significant decrease in *Kla* is mainly attributed to the formation of large bubbles and a decrease in specific interfacial areas that can be counterbalanced by the static mixers presence. The values of *Kla* can be correlated with the gas superficial velocity, v_{SG}, in tower bioreactors as

$$Kla = u v_{SG}^{t'} \tag{34}$$

The presence of static mixers affects u and t' values. They do not affect t' very much, but significantly enhance u dependent on the consistency index (Fig. 20) (36). The degree of *Kla* enhancement depends on the fluid viscosity and the pattern of fluid circulation through the reactor. The results of Hsu et al. (20) also show that mass transfer coefficient in a bubble column in the presence of Koch type static mixers is improved. A comparison of the oxygen transfer efficiency in terms of kilograms of oxygen transferred per kilowatt hour consumed at low flow rates indicated that the Koch mixed column was almost twofold as efficient as the empty column. However, at high gas flow rates, the efficiency was about the same for the two columns. Also, the ratio Kla/P_G can be used to characterize the efficiency of energy utilization for mass transfer per unit volume of the liquid phase or dispersion. Potucek (29) found that an airlift bioreactor with static mixers showed higher Kla/P_G values for absorption of oxygen in water or polyacrylamid solutions in the region of gas velocities of $v_{SGR} = 0.01 - 0.06$ m/s. Also, in the batch yeast cultivation, Kla/P_G was larger than in the witness bioreactor; gas superficial velocities ranged from $v_{SGR} = 0.03 - 0.06$ m/s, whereas with $v_{SGR} = 0.01 - 0.03$ m/s, lower values of Kla/P_G were obtained. As energy costs increase, more capital should be allocated to motionless mixers that increase the mass transfer efficiency. Olivier and Oosterhuis

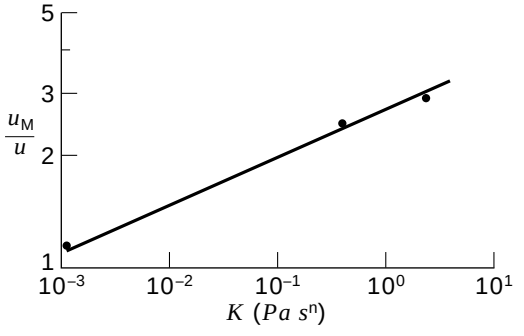


Figure 20. Variation of the ratio u_M/u with the consistency index.

(22) considered that by keeping liquid and gas velocities as well as the volume ratio of empty tube parts and internals constant, the oxygen transfer would be higher and nearly independent on the scale. They also considered that oxygen transfer takes place not only in the internals, but also in the empty tube parts, and correlated the overall oxygen transfer coefficient as (22):

$$Kla = \frac{(Kla\ V_M)_{\text{internals}} + (Kla\ V_E)_{\text{empty tube}}}{(V_M)_{\text{internals}} + (V_E)_{\text{empty tube}}} \quad (35)$$

The results obtained in different gas–liquid contactors show larger interfacial areas in the systems with static mixers (42) (Fig. 21). Data also indicate that the static mixer contactors are superior to other nonmechanically contacting devices, regarding mass transfer coefficient. Hsu et al. (20) investigated mass transfer in a bubble column with Koch mixers comparative to that with sieve trays. They found a different dependence of Kla on gas superficial velocity:

$$\text{for KOCH mixed column:} \quad Kla \approx v_{SG}^{0.95} \quad (36)$$

$$\text{For sieve-tray bubble column:} \quad Kla \approx v_{SG}^{0.89} \quad (37)$$

Also, data presented by Atkinson and Mavituna (42) show that in static mixed gas–liquid contactors a larger range of high Kla/V values was obtained, compared to those resulted in other gas–liquid contacting devices (Fig. 22).

The optimization of the number of motionless mixing elements and their positioning can enhance the mass transfer rates.

HEAT TRANSFER USING STATIC MIXERS

The role of the static mixers in heat transfer to flowing homogeneous or heterogeneous or heterogeneous fluids consists in fetching fresh fluid to the surface of the heat exchanger by promoting secondary flow, mixing, and turbulence (15). Most heat exchanger performances depend on hydrodynamics, and there are evident advantages that heat transfer may be enhanced by using built-in elements that change the velocity fields inside the tubes. Several studies investigated the heat transfer enhancement by using twisted strip inserts (43), helical static mixers (44,45), and wire matrix (15). The static mixers mostly used for heat transfer enhancement were helical static mixers, which promote a secondary motion in the fluid, increasing the heat transfer between the fluid and the tube wall by convection.

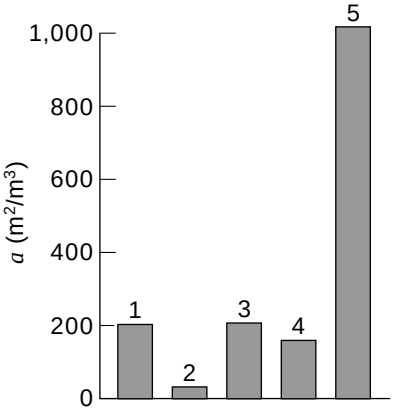


Figure 21. Comparison of specific interfacial areas in different gas–liquid contacting devices. 1 = Baffled agitated tank; 2 = bubble column; 3 = packed tower; 4 = plate tower; 5 = static mixers. *Source*: Data from Ref. 42.

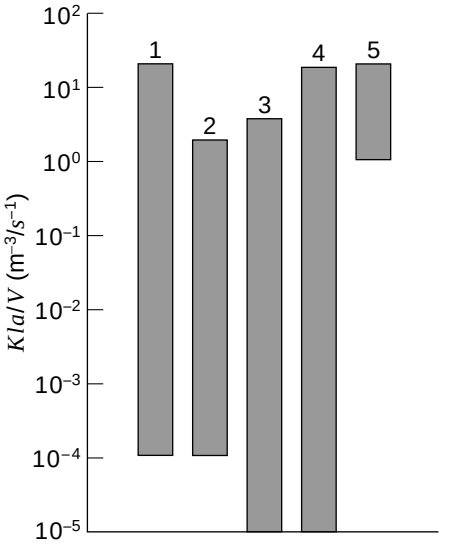


Figure 22. Comparison of different gas–liquid contacting devices regarding Kla/V values (significance of numbers as in Figure 21. *Source*: Data from Ref. 42.

The heat transfer performances of a given piece of equipment is usually estimated in terms of a Nusselt number, Nu, which is then related to the fluid properties and the Reynolds number, Re, characterizing the fluid environment. For instance, the individual heat transfer coefficient α_1 for the inside surface of a tube equipped with a helical static mixer is assessed from the following expression (44):

$$Nu_1 = \frac{\alpha_1 D}{\lambda_1} = \Phi(Re_1, Pr_1, h/d, \eta W_1/\eta_1) \quad (38)$$

with

$$Re_1 = \frac{D v_1 \rho_1}{\eta_1} \quad (39)$$

$$Pr_1 = \frac{c_{p1}\eta_1}{\lambda_1} \quad (40)$$

The overall heat transfer coefficient, K_T, of a tubular heat exchanger is related to the individual coefficients as follows:

$$K_T = \frac{1}{\frac{1}{\alpha_1}\frac{D_e}{D} + \frac{D_e}{2\lambda_s}\ln\frac{D_e}{D} + \frac{1}{\alpha_2}} \quad (41)$$

The values of K_T in a double-tube heat exchanger with hot water flowing through the annular space and cold water–glycerol solutions circulating through the inner tube containing helical static mixers were determined using the individual heat transfer coefficients calculated with the following equations (44): For the tube ($90 \leq Re_1 \leq 5800$; $6 \leq Pr_1 \leq 5800$);

$$\mathrm{Nu_1} = 1.86\left(\mathrm{Re_1 Pr_1}\frac{\mathrm{d}}{\mathrm{L_T}}\right)^{0.33}\left(\frac{\eta_1}{\eta_{W_1}}\right)^{0.14} \quad (42)$$

For the annular space ($Re_2 = 1520$):

$$Nu_2 = 1.02 Re_2^{0.45} Pr_2^{0.50}\left(\frac{D_e - D}{L_T}\right)^{0.40} \cdot \left(\frac{D_e}{D}\right)^{0.80} Gr^{0.05}\left(\frac{\eta_2}{\eta_{W_2}}\right)^{0.14} \quad (43)$$

Lecjacks et al. (44) reported the following relationship for helical elements of continuous and Kenics type:

$$Nu_1 Pr_1^{-1/3}\left(\frac{\eta_{W_1}}{\eta_1}\right)^{0.14} = a_1 Re_1^{n_1} \quad (44)$$

with

$$a_1 = \Phi_a\left(Re_1, \frac{h}{D}\right)$$

and

$$n_1 = \Phi_n\left(Re_1, \frac{h}{D}\right)$$

Also, they extended their investigations in the turbulent flow regime, which allowed them to assess the favorable effect of the built-in elements on heat transfer, especially in the laminar regime, for low values of h/d in Kenics static mixer. In the continuous helical static mixer, an increase in friction factor and heat transfer coefficient of 4.4% and 5.3% was obtained, whereas in the Kenics insert this increase was of 19.3% and 10.2%, respectively, comparative with the empty tube (45). Therefore, the heat transfer coefficient was found to be dependent on the flow conditions and geometrical properties of the internals. Olivier and Aldington (15) found a noticeable heat transfer enhancement by factors up to five in round tubes containing wire matrix turbulators, comparative to the empty tube, in laminar regime. Also, the wires caused the falling of the effective viscosity in non-Newtonian solutions at high liquid flow rates. Static mixers are used for thermal homogenization when a breakdown can occur on the heat transfer surface, especially in viscous media. Shear-thinning fluids should give good heat transfer/pressure drop behavior in the presence of tube inserts because the viscosity is reduced in the higher shear rates that are obtained. Static mixers are insensitive and nonresponsive to temperature. They can be well sealed against the surroundings. Maintenance and wear are small, and they are very economical, because they do not require additional space (in-line disposed).

SCALE-UP CONSIDERATIONS

Static mixers are used for many monophasic or gas–liquid, two-phase operations, which involve momentum, mass, and heat transfer. They are largely used in biotechnology and other applications such as processing of natural gases, wastewater treatment, dissolution of gases, continuous mixing of mutually soluble fluids of different viscosities, and so on. It is widely agreed that static mixers are suitable to mix fluids of very widely different viscosities, being especially effective for use with highly viscous fluids. They also could remove any mistakes made by the equipment designed for heat and mass transfer.

As shown in the fermentation field, factors taken into consideration on the choice and the design of an equipment for a definite process, besides the general factors, should include the population safety and damage, the uniform distribution of the phases and energy in phase transfers, the oxygen demand of the aerobic cultures, and so on (4,7). Static mixers can fulfill these objectives. Their action gives rise to elementary fluid layer formation. The maximum thickness of an elementary layer is given by the following expression (16):

$$\delta = \frac{D}{n} = \frac{D}{ca^N} \quad (47)$$

As in the case of any apparatus, the static mixer behavior is tested first in small-scale equipment and then designed for large-scale operation, on the basis of the experimental data acquired for the model. For geometric similarity between the model and the actual mixer, the same degree of mixing means that (16):

$$\delta_m = \delta_a \quad (48)$$

or

$$\frac{D_m}{ca^{N_m}} = \frac{D_a}{ca^{N_a}} \quad (49)$$

Boss and Czastkievicz (16) introduced a linear scale-up coefficient, expressed as follows:

$$S = \frac{D_m}{D_a} \quad (50)$$

when equation 49 becomes

$$N_a = N_m + \frac{lgS}{lga} \quad (51)$$

This reasoning is valid in the laminar flow regime. For

constant physicochemical properties of the fluids that flow through the model and the actual mixer, the changes in Reynolds number for the actual mixer relative to the model can be attributed to modifications occurring in the flow rate of the fluids or in the diameter of the static mixer. The relationships between the volumetric flow rates and Reynolds numbers in the model and the actual mixers, respectively, are as follows (16):

$$\dot{V}_a = \frac{z_m}{z_a} \frac{S^3}{1 + \lg S/N_m lga} \dot{V}_m \quad (52)$$

$$Re_a = \frac{z_m}{z_a} \frac{N_m}{N_a} S^2 Re_m \quad (53)$$

These general equations, applicable to scale-up of mixing processes in static mixers, are inferred starting from the pressure drops associated with laminar flow of a Newtonian fluid through circular pipes, which are taken into account by the pressure drop coefficient z, defined in equation 9. Boss and Czastkievicz (16) gave several examples of scale-up calculations for some types of static mixers, namely Kenics, Ross-ISG, and Sulzer.

CONCLUDING REMARKS

An overview of the published results on transfer processes in fluid flow through static mixers has been presented in the preceding pages. The velocity fields inside a tube containing motionless mixers can be modified using static mixers. As a result, a significant enhancement in mixing, mass, and heat transfer are achieved. Static mixers can be used in chemical and biochemical processes to perform a rapid mixing of monophasic, diphasic, and poliphasic systems with low energy requirements, and the fluids may have similar or different viscosities. The energy for mixing is extracted from the fluid flowing through the mixers. Static mixer effects result from specially designed flow arrangements, cutting and twisting, displacement and distortion, or separation and expansion. Significant efforts are devoted by specialists for developing new static mixers and optimal strategies for using minimum number of static mixers located at different positions in tubes, to ensure greater efficiency on transfer processes intensification, especially mixing and oxygen transfer in aerobic fermentation processes.

NOMENCLATURE

A	Cross-sectional area [m^2] (1 in^2 = 6.4516.10^{-4} m^2; 1 ft^2 = 9.290304.10^{-2} m^2)
a	Specific interfacial area [m^2/m^3]
c_p	Specific heat at constant pressure [J/kg °K] (1 ft pdl $°R^{-1}$ lb^{-1} = 0.1672255 J/kg °K)
D	ID of the tube [m] (1 in = 2.54.10^{-2} m; 1 ft = 3.084.10^{-1} m)
D_e	Outside tube diameter [m]
d	Diameter of static mixing element [m]
f_D	Friction factor
Gr	Grashof number
h	Height (length) of a static mixing element [m]
K	Consistency index in power-law rheological model [Pa s^n] (1 lb s ft^{-2} = 1.488163944.10 Pa s)
K_T	Overall heat transfer coefficient [W/m^2 °K] (1 lb s^{-3} $°R^{-1}$ = 0.816466266 W/m^2 °K)
L	Length of the tube [m]
N	Number of static mixing elements
Nu	Nusselt number
n	Number of elementary layers produced in laminar flow over a static mixer
Pr	Prandtl number
Δp	Pressure drop [Pa] (1 lb in^{-2} = 6.89476.10^{-3} Pa)
Re	Reynolds number
Re_D	Reynolds number in a tube with inner diameter D
t_M	Mixing time [s]
V	Volume [m^3] (1 ft^3 = 2.83168 10^{-2} m^3)
$\dot{V}$	Volumetric flow rate [m^3/s] (1 ft^3 s^{-1} = 2.83168.10^{-2} m^3/s)
$\mathbf{v}_S$	Superficial velocity ($\mathbf{v}$ = $4\dot{V}/\varphi D$) [m/s] (1 ft s^{-1} = 3.048 m/s)
α	Individual coefficient of heat transfer [W/m^2 °K]
ϵ_G	Gas holdup
δ	Maximum thickness of an elementary layer [m]
λ	Thermal conductivity [W/m °K] (1 ft lb s^{-3} $°R^{-1}$ = 0.248858918 W/m °K)
η	Dynamic viscosity [Pa s]
ρ	Density [kg/m^3] (1 lb ft^{-3} = 1.60185 10 kg/m^3)

Indices

a	Referring to an actual static mixer
D	Downcomer section of an airlift reactor
E	Referring to the empty tube
G	Gaseous phase
L	Liquid phase
M	Referring to static mixer presence
m	Referring to model static mixer
o	Referring to a single phase
R	Riser section of an airlift reactor
s	Referring to solid surface
T	Referring to heat transfer
TP	Referring to diphasic flow
W	Referring to wall temperature
1	Fluid 1
2	Fluid 2

BIBLIOGRAPHY

1. B.H. Chen, *Ind. Eng. Chem. Process Des. Dev.* **9**, 20–24; 121–126 (1970).
2. C.J. Slattery, *AlChE J.* **16**, 345–352 (1970).
3. H. Brauer and D. Sucker, *Int. Chem. Eng.* **18**, 367–380 (1978).

4. V.D. Prokopenko and U.E. Viesturs, *Acta Biotechnol.* **6**, 325–338 (1986).
5. M.H. Pahl and E. Muschelknautz, *Chem.-Ing.-Tech.* **52**, 285–291 (1980).
6. R.Z. Tudose and A. Moise, in *Proc. of the First Natl. Symposium on Inventing* Sept. 7–9, Iaşi, Romania, 1984, pp. 209–212.
7. M. Gavrilescu and R.Z. Tudose, *Acta Biotechnol.* **15**, 3–26 (1995).
8. Pat. RO 73630 (1979), R.Z. Tudose, F. Vitan, and L. Bujor (to Technical University of Iaşi).
9. Pat. RO 98240 (1989), M. Ionescu, R.V. Roman, M. Gavrilescu, A. Sauciuc, A. Pintilie, A. Pascal, S. Matache and T. Rez (to Chemical-Pharmaceutical Research Institute of Bucharest).
10. K.H. Hsu, L.E. Erickson, and L.T. Fan, *Biotechnol. Bioeng.* **19**, 247–265 (1977).
11. J.R. Gutierrez and L.E. Erickson, *Biotechnol. Bioeng.* **20**, 487–501 (1978).
12. M. Fialova, K.H. Redlich, and K. Winkler, *Collect. Czech. Chem. Commun.* **51**, 1925–1932 (1986).
13. Z. Lecjaks, I. Machak, and J. Sir, *Chem. Eng. Process.* **18**, 67–72 (1984).
14. Pat. RO 83269 (1983), R.Z. Tudose, A. Bacaoanu, and L. Bujor (to Technical University of Iaşi).
15. D.R. Olivier and R.W.J. Aldington, *Chem. Eng. Res. Des.* **66**, 555–565 (1988).
16. J. Boss and W. Czaskievicz, *Int. Chem. Eng.* **22**, 362–367 (1982).
17. E.B. Naumann, *Chem. Eng. J. (London)* **47**, 141–148 (1991).
18. R. Daraktschiev, *Int. Chem. Eng.* **30**, 222–227 (1990).
19. F. Gross-Roll, *Int. Chem. Eng.* **20**, 542–553 (1980).
20. K.H. Hsu, L.E. Robinson, and L.-T. Fan, *Biotechnol. Bioeng.* **27**, 499–514 (1975).
21. N.F. Shah and D.D. Khale, *AlChE J.* **38**, 308–310 (1992).
22. A.A.P.C. Olivier and N.M.G. Oosterhuis, in D.G. Bobichon and J. Florent eds., *8th International Biotechnology Symposium*, Société Française de Microbiologie, Paris, 1998, pp. 388–409.
23. L.T. Fan, H.H. Hsu, and K.B. Wang, *J. Chem. Eng. Data* **20**, 26–32 (1975).
24. R.Z. Tudose and F. Vitan, *Rev. Chim. (Romania)* **32**, 1003–1006 (1981).
25. M. Gavrilescu and R.Z. Tudose, *Bioprocess Eng.* **16**, 93–99 (1997).
26. M. Gavrilescu and R.V. Roman, *Acta Biotechnol.* **14**, 27–36 (1994).
27. G.T. MacLean, L.E. Robinson, K.H. Hsu, and L.T. Fan, *Biotechnol. Bioeng.* **19**, 493–505 (1977).
28. L. Rusnak and R. Vlađea, *Rev. Chim. (Romania)* **42**, 597–601 (1990).
29. F. Potucek, *Collect. Czech. Chem. Commun.* **54**, 3213–3219 (1989).
30. J. Stejskal and F. Potucek, *Biotechnol. Bioeng.* **18**, 1552–1572 (1976).
31. K.B. Wang and L.T. Fan, *Chem. Eng. Sci.* **33**, 945–952 (1978).
32. M. Popovic and C.W. Robinson, *Biotechnol. Bioeng.* **32**, 301–312 (1988).
33. M. Gavrilescu and R.V. Roman, *Acta Biotechnol.* **15**, 323–335 (1995).
34. M. Gavrilescu and R.V. Roman, *Acta Biotechnol.* **16**, 145–153 (1996).
35. O. Levenspiel, *Chemical Reaction Engineering*, John Wiley, New York, 1972.
36. M. Gavrilescu, R.V. Roman, and A. Sauciuc, *Biotechnol. Bioequip. (Bulgaria)* **6**, 60–64 (1992).
37. Y. Kawase and M. Tsujimura, *Biotechnol. Bioeng.* **44**, 1115–1121 (1994).
38. P.H. Calderbank, in N. Blakebrough ed., *Biochemical and Biological Engineering Science*, Academic Press, London, 1967, pp. 101–180.
39. Y. Chisti, M. Kasper, and M. Moo-Young, *Can. J. Chem. Eng.* **68**, 45–50 (1990).
40. C.H. Lin, B.S. Fang, C.S. Wu, H.Y. Fang, T.F. Kuo, and C.Y. Hsu, *Biotechnol. Bioeng.* **18**, 1552–1572 (1976).
41. M. Gavrilescu, R.V. Roman, and V. Efimov, *Acta Biotechnol.* **13**, 59–70 (1993).
42. B. Atkinson and F. Mavituna, *Biotechnical Engineering and Biotechnology Handbook*, Stockton Press, New York, 1991, p. 711.
43. A. Date, *Int. J. Heat Mass Transfer* **17**, 845–859 (1974).
44. Z. Lecjaks, I. Machac, and J. Sir, *Int. Chem. Eng.* **27**, 210–217 (1987).
45. D. Burfoot and P. Rice, *Chem. Eng. Res. Des.* **62**, 128–132 (1984).

See also MASS TRANSFER.

STERILIZATION, AIR.

See FILTRATION, AIR; STERILIZATION-IN-PLACE.

STERILIZATION-IN-PLACE

P.T. NOBLE
Fluor Daniel GmbH
Wiesbaden, Germany

KEY WORDS

Clean steam
Dead legs
Microbiological control
Overkill
Ozone
Steam-in-place
Steam traps
Validation
Vaporized hydrogen peroxide
Vent filters

OUTLINE

INTRODUCTION

Sterilization-in-place (SIP) refers to the practice of sterilizing process equipment or materials in process at their place of use or installed location. Sterilization can be simply defined as a procedure for rendering a system free of living organisms. SIP is a very important technology, needed for microbiological control during production. Although the process of microbial sterilization is a well-studied area of science, SIP varies widely in its interpretation and industrial application. Steam is usually employed as the sterilizing agent, although many sterilizing methods are conceivable. The term SIP in some places is used loosely for a sanitization process whereby the goal is not necessarily achievement of a high assurance level of sterility (SAL).

APPLICATIONS

SIP supplements or replaces the use of process equipment whose purpose is to sterilize materials such as autoclaves, sterilization tunnels, and ovens. Introduction of the equipment into a sterilization chamber often is not practical. SIP is also preferred when several components can be sterilized together via SIP rather than separately treated and subsequently connected under aseptic conditions.

Clean piping systems, such as a water for injection (WFI) supply ring, will usually be sterilized in place before start-up and after service interruptions. Here, the goal is to begin or resume microbiological control but not achieve a specified sterility assurance level. Fermentors, especially those used to culture genetically altered cells, need to be sterilized before innoculation, which can be achieved to a high level of assurance by SIP. Fermentation media are sometimes first introduced into the fermentor and then sterilized with the fermentor. Aseptic manufacture of parenteral drugs and devices requires sterile process equipment, which is usually achieved by SIP. For example, modern filling equipment often undergoes an SIP of the filling lines and nozzles. It is also a current Good Manufacturing Practice (cGMP) standard that lyophilizers for parenteral products be outfitted with an SIP cycle to sterilize the drying chamber and process connections that are open to the product. Finally, as part of biocontainment practice (especially when using recombinant or hazardous organisms), processing equipment may receive an SIP after use but before opening to inactivate any residual microorganisms and prevent their release into the environment.

STEAM-IN-PLACE TECHNOLOGY

The use of saturated steam for sterilization has been relied on for many years. It is natural that SIP would evolve from this experience. One of the first publications on the use of steam for SIP was provided by Myers and Chrai (1) and later reviewed by Agalloco (2). The preference for steam as the sterilizing agent can also be attributed to its advantages in heat transfer, which are elaborated in the following.

Sterilization in general is a kinetic process and is a subject of sufficient breadth for textbooks (3). It follows first-order kinetics for individual microbial populations, whereby the rate constant is species specific as well as being dependent on the sterilizing conditions. When only water is present at the saturated steam condition, the temperature alone is sufficient for completely defining these conditions and hence the rate constant for a particular organism. The rate constants are exponentially dependent on temperature, and the dependencies are significant for temperature variations as small as 1 °C.

For scientific studies, sterilization processes are usually quantified in terms of the D and F values. The D value provides the kinetic information and is simply related to the rate constant. It is species and temperature dependent, and its value is the time in minutes for a 10-fold reduction in the population. The F value gives the time in minutes for a desired reduction in the population, as calculated by $F = (\log N_0 - \log N)\, D$, where N_0 is the starting count and N is the desired count.

The F_0 value has become a standard for defining and comparing sterilization processes. It is based on the sterilization kinetics that have been found for heat-resistant organisms, such as *Bacillus stearothermophilus*. The F_0 value can be calculated from thermal data using the following integral equation, $F_0 = \int 10^{(T-121.1)/10}\, dt$, where T is the sterilization temperature in degrees Celsius and t is the sterilization time in minutes. SIP operations typically should demonstrate an F_0 value in the range of 8 to 20 min, depending on the application.

Steaming conditions other than with pure saturated steam have been found by experience to be less effective (3). Apparently, the presence of both phases, liquid and gas, is important because superheated steam at the same temperature is not as lethal. Also, the presence of air reduces sterilization effectiveness. If air is present, and the sterilizing pressure is kept constant, then the contribution of air to the total pressure lowers the partial pressure of the steam and hence its saturation temperature. The lower temperature has a direct effect on the kinetics. If an outside heating source raises the temperature, the system departs from saturated steam conditions; that is, water is present then only as superheated steam. If the total pressure is raised, so that the steam saturation temperature is maintained, such as with an overpressure autoclave cycle, then the kinetics predicted with saturated steam are followed.

Validation Issues

Important validation issues arise with SIP. The general problem is the estimation of sterility, but the additional

process complexities associated with an SIP operation often add problems of control for the process parameters. Validation is needed to provide assurance that the probability of surviving organisms is small enough to be of insignificant risk to the operation. Because the initial bioburden (number and type of microorganisms) can only be partially known from past monitoring experience, kinetic models cannot predict with accuracy the outcome of SIP. Measurements of sterility are in general destructive tests; that is, they violate the integrity of the enclosed sterile system.

Accurate temperature and pressure measurement and control is of special concern for steam-in-place processes. It is especially important to know the location of the coldest points in the system to estimate conservatively the sterilization effect. An accurate measurement of pressure is needed to verify saturated steam conditions. Validation testing include extensive thermal measurements, the use of biological indicators (e.g., spore strips) and often sampling methods via swabbing or media contact. Such testing should yield consistent, favorable, and reproducible results to give such assurance.

Reflecting the uncertainties and risks associated with a sterilization process, regulatory bodies, such as the U.S. Food and Drug Administration (FDA) (4), have specified relatively severe process parameters to meet for general use. These conditions are often termed overkill. In the cited reference, 121.5 °C is to be maintained with saturated steam for 20 min, or equivalency to these conditions must be demonstrated as by achieving the same F_0 value. Using the formula provided, this value would be $F_0 = 21.9$ min.

Process Description

A technical description and a theoretical modeling of the steam-in-place process can be found in Ref. 5 and is relied on here. A steam-in-place process follows the general operations of an autoclave, which have the following cycle steps:

1. Air purge
2. Heat-up
3. Sterilization holding period
4. Displacement and collapse of the steam with sterile air

SIP of a vessel or piping system usually accomplishes the air purge step by displacing the air with steam. The air exits the system through drain and vent openings, which are usually equipped with steam traps.

A residual presence of air is to be avoided because it introduces a poorly controllable heat transfer resistance, which is described next. Its contribution to the total gas pressure also lowers the effective steam pressure and hence saturation temperature, which affects directly the sterilization kinetics. Therefore, this first step is critical to the reproducibility and success of the operation. Unfortunately, displacement of air with steam is often not very effective. Some autoclaves utilize a series of vacuum pump-outs followed by steam repressurization to remove the air from the chamber. Vacuum pump-outs are seldom part of an SIP operation unless an efficient vacuum system is already present for process needs, as with freeze dryers.

Convective displacement of air with steam always results in some mixing of air with the steam. In branched piping systems with poor flow distribution or in large vessels, this mixing can be extensive. In theory (5), turbulent steam flow is desirable because the additional kinetic energy helps to break up any air pockets and transport the air (with steam) out. The air purge step can be successful when the piping design meets appropriate criteria (to be described later) and restrictions to flow are minimized such that maximal steam flow can be reached.

Using steam to heat an enclosure combines both heat and mass transport phenomena. In the gas phase, a comparable situation (condensation of vapor in the presence of a noncondensable gas) has been analyzed and modeled for psychrometric calculations (6). Geometric complications prevent an exact analysis of a typical SIP application. Additionally, steam-in-place must consider transport in the condensed phase as well as heat transport in the walls of the system.

The combined heat and mass transport created by steaming yields important advantages for heat transfer. When saturated steam contacts a colder surface, it condenses, and the resulting void created by the collapse of water vapor molecules into a condensed phase drives more steam to the region of the colder surface. The pumping effect of the condensation helps bring steam heat to the colder regions of the enclosure and thereby improves the heat transfer. Reliance on latent heat also reduces temperature drops, which are otherwise needed for sensible heat transfer.

During the heat-up step, the system pressure is allowed to build up to the steam supply pressure by restricting any vents or drains or via the functioning of the installed steam traps. Any trapped air left will be largely swept to the cold boundaries by the dragging effect of the moving steam. Thus, the residual air will not be uniformly distributed but will be concentrated near the colder boundaries. As a stagnant layer, through which the steam must pass to condense on the walls, residual air can bring a significant resistance to heat transfer. Enclosures with a large volume-to-surface ratio, such as tanks, are particularly prone to interference from trapped air. In Ref. 5, temperature drops of the order of 1 °C are predicted from residual amounts of air equivalent in some cases to less than 0.1% of the original volume. Such exquisite sensitivity is possible in piping dead legs (to be described later) and is greatly amplified when the equipment is not thermally insulated. Insulation reduces heat transport requirements and thereby mitigates the sweeping of air into stagnant boundary layers.

Condensate removal during steam-in-place is a serious concern. Condensate forms on all cold surfaces and will subcool if allowed to accumulate. Because of its low thermal diffusivity, liquid water can maintain a substantial temperature difference between the steam phase and the equipment walls. Condensate layers will collect at all low points. Flow in these layers to the next steam trap is usually laminar. A one-dimensional analysis (5) is useful to predict order-of-magnitude effects. In an insulated pipe, a

condensate layer 1 mm in depth will have a temperature drop of the order of 0.5 °C. Without insulation, only a 0.2-mm depth would show such a temperature drop.

The boundary layer phenomena discussed here are not usually observable without extensive temperature monitoring, which may only be performed during validation as a thermal mapping activity. From this study the coldest locations are identified. Thermal data from these locations are then used to define a sterilization holding period to achieve equivalence to the overkill process, or a predefined F_0. In normal operations only one or a few sensors, which for practical reasons sometimes cannot be positioned at the coldest locations, are used to monitor temperature. Temperature nonuniformities bring an uncertainty into the estimation of the sterilization effect; their existence demonstrates the presence of heat transfer resistances, usually resulting from boundary layer phenomena, which are not under direct control. Statistical analysis of temperature data during process validation will be used to add additional time to the holding period to provide the desired degree of sterility assurance.

On completion of the sterilization process, the steam supply is shut, and the equipment can be cooled either actively or passively before use. The collapse of residual steam will ultimately create a vacuum in the enclosure, which is to be avoided because a vacuum increases the risk of recontamination of the system. Therefore, the steam is usually displaced with sterile-filtered process air while there is still overpressure present.

Equipment Design

Process requirements for steam-in-place operations have determined to a large degree the special equipment design specifications associated with bioprocesses. The sterilization conditions usually chosen are 121 °C or greater, resulting in a saturated steam pressures above 1 bar gauge. Materials of construction and construction methods must be suitable for these conditions.

Materials of Construction. 316L stainless steel is usually chosen because of its strength and chemical resistance to the high-purity steam used. Vessels need to be pressure rated for these conditions and protected with safety relief valves or rupture disks.

Specifications for surface finishes and construction methods have the purpose of reducing the potential for the buildup of soil and deposits on the equipment surfaces. Surface deposits provide an additional barrier to the transport of heat and steam. They may also introduce an uncontrolled bioburden level. A high degree of polish will be required; electropolished is preferred. Corners and joints should meet with a generous radius of curvature. Ideally, sealing methods would be seamless. For this reason, welding is preferred, but the welds must be polished and passivated. Seiberling (7) provides general specifications of equipment to receive SIP.

All low points inside the enclosures must have drains; these are needed to drain the condensate and are also required to effectively clean the equipment so as to prevent soil buildup. Thermal insulation of the equipment is needed to reduce the temperature variations in the system. If the equipment is located in a clean room, the insulation must be clad with a housekeeping skin, so that particles from the insulation do not enter the clean rooms and the surfaces can be maintained clean. Stainless steel sheathing is usually used on vessels, and plastic is easier to install on piping. The enclosed insulation must not chemically release agents that would attack the equipment, such as chlorides, which are known to attack stainless steel.

Piping Design. Piping particularly should be thermally insulated because of its large surface-to-volume ratios. All piping must be sloped to drain condensate. Because the mass transport of steam and condensate are driven by heat transport and gravity, respectively, and not by large pressure gradients, it is important in the piping design to remove any uncertain flow paths. A clear example of this concern can be made with parallel flow arrangements. If two or more identical items such as heat exchangers are steamed in parallel, it is likely that one item will not receive the proper treatment because of poor steam distribution. In an extreme case, one item may stay filled with air or become blocked with accumulated condensate, effectively stopping the flow of steam.

Dead legs are to be avoided, but this is not usually possible. Dead legs are commonly understood to be sections of piping with no outlet. Connections that are closed with a valve during SIP are typical dead legs. Instrument ports can be another. Dead legs also need to be sloped to allow for draining. Depending on their length, flow of steam in and air out can be largely a diffusive process. Dependence on diffusion means some air is always trapped in dead legs during SIP, contributing to the difficulty of their sterilization (8). General design guidelines (7,9) limit dead legs in length by a scale factor expressed in terms of pipe diameters, but this rule cannot be defended in theory (5). The best guideline is to minimize their length because they will usually be the most difficult locations to sterilize. Practically, dead legs can easily be limited to four pipe diameters with commonly available pipe fittings.

Dead leg–free valves are commercially available (Fig. 1); they function by using the pipe wall on one side as the valve seating surface. The seating design employs a flexible diaphragm or membrane, which is pushed into the flow path to press against the seating surface. Such valves belong to a type, commonly referred to as diaphragm or membrane valves, which are preferred for equipment that receives SIP or clean-in-place (CIP). Such valves have no dynamic seals for moving parts such as valve stems that may enter or leave the sterile side of the valve. Dead leg–free valves are only such on one side of the valve. On the other side, the valve body itself is a short dead leg.

Instrumentation and Monitoring. The critical parameters to monitor are temperature and pressure. By measuring both, one can confirm that saturated steam conditions are present. Instrumentation must be of a sanitary design, which follows the general considerations given earlier. Temperature is usually measured with resistance temperature detectors (RTD) because their signal can be recorded, they are robust, and they can be accurately cali-

Figure 1. Dead leg–free valve. *Source*: Reproduced with permission of SÜDMO Schleicher AG, Riesbürg, Germany.

brated. A calibration accuracy of no more than ± 0.5 °C is recommended because such variations significantly affect the sterilization effectiveness.

Clean Steam Supplies. If the SIP application is for parenteral pharmaceutical products, high-purity steam, commonly known as clean steam, must be specified. Product contact surfaces in such applications must meet only strictly controlled media. Clean steam is usually defined by the purity of its condensate, and this must usually have WFI quality. Normal plant steam will contain more dissolved solids and potentially also some suspended debris. From these discussions, the steam supply should in any case be free of air, which is normally the case even with plant steam.

The steam supply system (boiler and distribution headers) normally is at higher pressure than the desired steam sterilization pressure. The steam connection to the equipment will then include a pressure-regulating valve to reduce the pressure. This design is preferred because the supply system then has reserve to supply large amounts of steam, without a serious loss of steam pressure. Too large a pressure reduction should not be expected from the pressure regulation valve, however. When too large a reduction is made just before entering the equipment, the steam acquires significant superheat. Superheated steam is known to be less effective as a sterilizing agent and as a heat transfer medium. Its presence affects the sterilization kinetics but is not easily observable. An industrially based guideline is to design the pressure reduction for 2:1 (10).

Sizing the steam supply for a steam-in-place is not straightforward, and there is no consensus in the industry. The greatest demand for steam occurs during the air purge and subsequent heat-up phases. The demand during the air purge is really determined by the restrictions to flow, often created largely by the vents and drains, as is discussed in the next section. During the heat-up step, the capacity of the steam supply could limit the heat-up rate, but the actual rate is not normally considered a critical variable.

Steam Traps and Vents. From the previous process description, it is clear that these components are critical to a successful steam-in-place operation. Steam traps are devices that are designed to not pass steam but rather air and condensate. There are many types available (11), and some models have been created especially for the biotech industry (Fig. 2). These devices operate as valves that are self-actuating, can sense by physical property differences whether they are passing steam, air, or water, and thus open or close appropriately. For steam-in-place applications, it is critical that they remove air and condensate. However, the steam trap designs that have evolved from other industries have the major goal of conserving steam. The types commonly selected for biotechnological applications allow some accumulation of condensate to occur upstream and may bring a significant restriction to the flow of air out of the system.

As a consequence, pharmaceutical companies have in the past dispensed with the use of steam traps and employed their own simpler solutions. The simplest solution

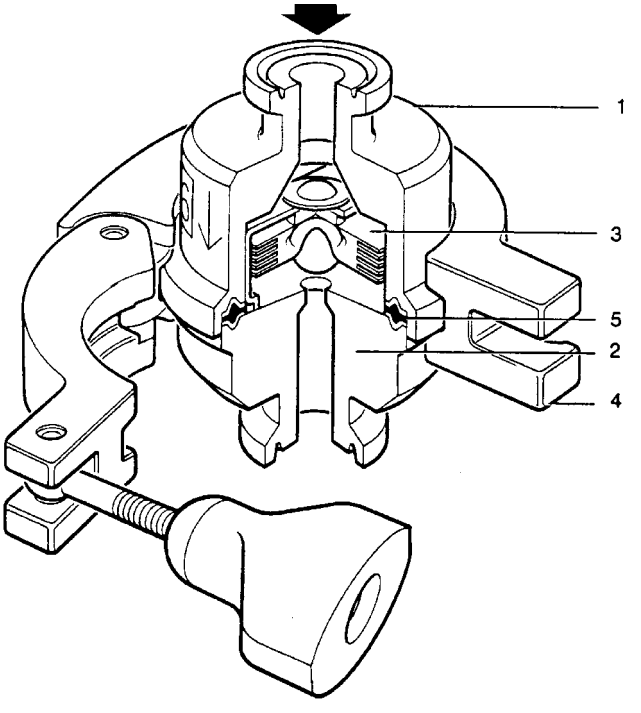


Figure 2. Sanitary blanced pressure steam trap. *Source*: Reproduced with permission of Spirax Sarco GmbH, Konstanz, Germany.

is to install manual valves instead of the steam traps, which must then be manually adjusted during the course of the procedure. During the air purge step, the valves may be fully opened. They are subsequently throttled, but only so far that steam continues to emerge, that is, condensate does not back up. A more automated approach is to employ restriction orifices instead (Fig. 3). During the air purge step, the orifice is bypassed with an open isolation valve to allow a relatively unrestricted flow of air and condensate. During the subsequent steps in the process, the isolation valve is closed, forcing remaining flow through the orifice. The orifice must be custom sized to ensure is large enough to allow all condensate to freely drain. Correctly sized, the orifice will be continuously passing (and wasting) steam and will prevent any buildup of condensate.

It is usual now to encounter sanitary steam traps of the type illustrated in Figure 2. Provided one realizes their limitations, successful steam-in-place operations can be performed with them. Typically, the steam traps are placed a significant piping distance from the limits of the system to be sterilized. The limit is usually an isolation valve, which will close immediately on completion of the operation. The piping downstream of this valve should then be noninsulated to allow for a sufficient condensation of steam and subcooling of condensate to keep the steam trap open. Steam traps must also be selected for the ability to pass air, and a sufficient number must be planned to allow for a successful air purging operation and condensate removal.

Filters. Sterilizable enclosures are almost always equipped with sterile filters, acting as barriers for sterile exchange of air or other media with the outside. Sterile vent filters are needed for pressure equalization, and process liquids are introduced though sterilizing liquid filters. Sterile filter installations are often sterilized together with the enclosure, although there is also no consensus on this practice. The filter element and enclosure may also be sterilized separately, but this raises concerns about ensuring that the separately sterilized components can subsequently be connected together aseptically.

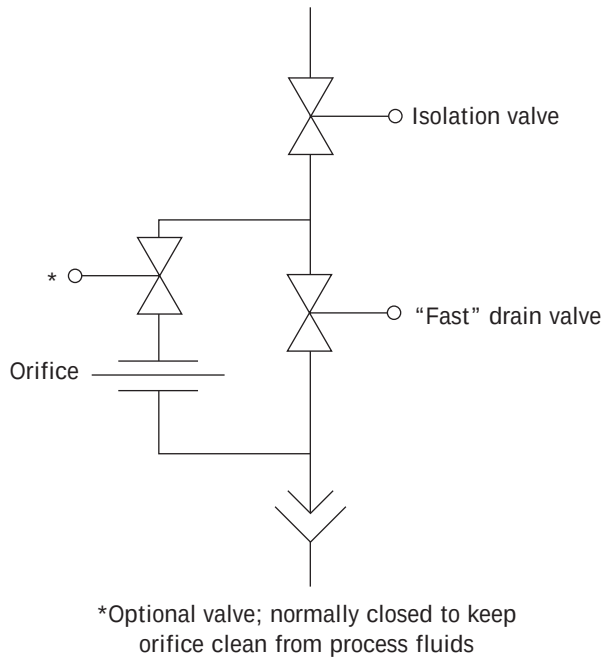


Figure 3. Use of orifices in place of steam traps.

The steaming-in-place of filters is not straightforward (12). Of special concern is the passage of steam, air, and condensate through the filter element. Also, the pressure drop across the filter element is limited by design. These concerns result in special piping designs and controls to ensure that both sides of the filter receive sterilizing conditions. It is not practical to attempt only a sterilization of the future sterile side of the filter, because without flow through of steam the filter will not be air purged successfully.

When steaming-in-place an enclosure, the vent filter will not be employed as the major vent for the air purge operation nor will it be used as the connection to the steam supply because it is typically too limited in flow capacity. It usually receives a secondary source of steam from the vessel during the sterilization process. Vent filter elements are typically nonwetting, which allows them to be quickly functional after steaming. A sterilizing liquid filter is sometimes steamed through to the enclosure (i.e., becomes the source of steam for the enclosure). Because of its limited ability to pass steam, the enclosure downstream is limited in size, often only piping. Liquid filters do not easily pass air after steaming because their hydrophilic pores become blocked with condensate.

ALTERNATIVE SIP TECHNOLOGIES

Equipment design requirements for steaming-in-place constrain opportunities for applying SIP, which raises interest in alternative technologies that are less demanding on mechanical designs. By employing chemical agents, sterilizing conditions can be reached without elevated temperatures or pressures. Sterilizing methods using radiation are not usually considered for SIP applications because the distribution of radiation throughout an extended enclosure is not practical.

When a new alternative SIP technology is chosen, a large amount of development work follows to make it a reliable industrial process. Much of this work involves scientific studies of the sterilization effect on the diverse microflora potentially present. A large amount of work is needed to develop monitoring and testing procedures for the validation and subsequent monitoring studies. Chalumeau has outlined the concerns raised with alternative sterilization methods in general (13). Some representative SIP methods receiving industrial attention are now used as examples for the technical issues encountered with alternative methods.

SIP with Superheated Water

The application of superheated water is an obvious extension to the use of saturated steam. This method can rely largely on the scientific knowledge base of the thermal

sterilization effect with saturated steam. Häggström has reported its effectiveness (14), but it has not been widely practiced. The advantage of using superheated water over steam is that it can be pumped and reheated to distribute the thermal conditions needed. Häggström claims that no steam traps or vents are needed and that a more uniform temperature distribution is attained.

A theoretical analysis (15) points to a fundamental weakness of the technology when compared with steaming-in-place. Heat transport by conduction through water is less effective than the combined heat and mass transport occurring in steam. The analysis considered heat transport in dead legs of hot WFI systems, where significant temperature drops of the order of 10 °C were predicted for typical dead legs. Although acceptable for WFI operation, such temperature variations are very unattractive for SIP. The advantages of superheated water have no benefit for the dead legs in the system, which are typically the most difficult to SIP. Also, to properly heat an enclosure with superheated water, it must be completely filled, or the headspace must be purged of air as in a classical steam-in-place process; this means that steam traps and vents are still required, although fewer would be needed. When the technical features are considered that would be needed to prepare and then discard large amounts of superheated water, it is not surprising that this method has not gained popularity.

SIP with Chemical Agents in a Liquid Phase

Chemical agents used for desinfection or sanitization are sometimes considered for SIP. Sanitizing solutions are not generally useful as sterilizing agents for SIP. To be useful they would have to have biocidal activity against the full spectrum of microorganisms yet be relatively noncorrosive and safe to handle. Commercially available agents that approach this ideal are often termed cold sterilants but still require many hours of exposure to achieve sterilizations.

In general, one must design a system to uniformly distribute the agent, monitor the operation for the proper exposure, and finally remove the agent, so that the equipment can be put back in operation. Monitoring of the operation must consider all parameters that can significantly affect the chemical reactions that determine the sterilization kinetics. Temperature may be important, but the concentration of the agent certainly is. Other material components or chemicals present can potentiate or interfere with the desired reactions. Unless a sensor is available, routine sampling and subsequent chemical analysis are required to verify sterilizing conditions.

High-purity water systems, such as those claiming to be sterile, have been reported using ozone as a circulating agent (16). Ozone has many advantages as a sterilant. It is extremely effective even at low concentrations because of its high reactivity, and it is easily generated electrically from ambient oxygen. In water systems, its decomposition product is also oxygen, which is usually not considered a contaminant. Before use of the media, residual ozone is catalytically decomposed with the aid of ultraviolet radiation. Widespread acceptance of this agent, even for simple microbiological control, has not been reached because of some underlying problems. First, it is an extremely aggressive agent that will degrade the piping system, especially gaskets. Second, it is difficult to remove it to levels sometimes desirable for end use. (See analyses in Ref. 15. Finally, penetration of chemical agents such as ozone into dead legs can be a difficult problem. Where diffusion dominates the transport, the relatively low liquid-phase diffusivities coupled with the low bulk concentrations result in very slow penetration of agent into dead legs (15).

At another extreme, caustic (sodium hydroxide) at relatively high concentrations can be considered as a sterilizing agent. Caustic is often used as a cleaning agent, such as in a CIP operation. If the concentration is high enough, it can be used as a sterilizing agent. Validation of the sterilization effect can be difficult because of the lack of scientific knowledge regarding such use of this agent with microorganisms. Also, the agent must be washed from the installation with large amounts of sterile water. However, it can be appropriate if the goal is to inactivate a biological agent left in equipment or in a waste stream before opening to the environment. Here, the removal of caustic after the operation is not critical; it will be subsequently removed in a washing step.

SIP with Chemical Agents in a Gaseous Phase

The use of gaseous chemical agents is probably the most active area of interest for alternative SIP technologies. Unlike liquid-phase agents, gaseous agents are easier to distribute and remove. There is also a long history of use, and consequently a significant knowledge base, with some of the biocidal gases. Formaldehyde and ethylene oxide have long been used for room and equipment disinfections in hospitals. As with all chemical agents, their use requires measurements of all important reaction conditions for the sterilization. One must also determine their reactivity with any system components and the ultimate fate of their decomposition products.

Both formaldehyde and ethylene oxide have become restricted in their use because of their toxicity and their activity as carcinogens. Ethylene oxide also has physical hazards; not only is it flammable, its chemical instability can lead to explosive decompositions. Recently, vaporized hydrogen peroxide as a sterilizing agent has been developed (17) to the point of a commercially available generator for SIP applications (Fig. 4). Its commercial success has been spurred by the active interest in the use of isolator technology for aseptic processing. Haas (18) has reviewed isolator design issues and listed gaseous agents that have been attempted for their SIP. At present, there does not appear to be a more promising chemical sterilant than hydrogen peroxide, which is likely to become the method of choice for pharmaceutical applications.

The AMSCO VHP™ systems have a basic three-step cycle: dehumidification, sterilization, and sterilant removal. Their commercial systems perform these functions on an attached isolator or enclosure. Often the isolator has operating components that assist the process. The reader is directed to the vendor for specific details on the operation of the commercial AMSCO VHP™ systems; the basic physical and chemical phenomena are considered here.

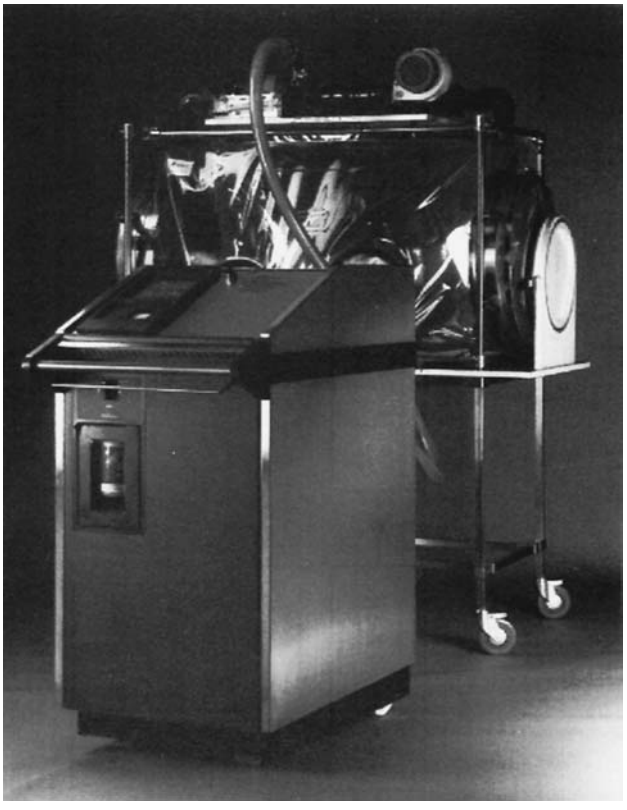


Figure 4. Vaporized hydrogen peroxide generator. *Source*: Reproduced with permission of AMSCO Scientific, Apex, North Carolina.

Hydrogen peroxide solutions have long been recognized for their disinfecting value. The greater lethality of vaporized hydrogen peroxide is a more recent discovery (17). H_2O_2 is an unstable material, available commercially in aqueous solutions up to 70%. High concentrations of H_2O_2 require proper handling because it can explosively decompose or cause fires (19). Its decomposition products are water and oxygen, and it will react with any oxidizable substance, particularly organic materials. Aqueous solutions of H_2O_2 are corrosive, but they have been found to be relatively benign to materials when only vapors are present (17).

A relatively safe solution concentration of 30% is usually the raw material consumed by a H_2O_2 generator. H_2O_2 has a lower vapor pressure than water, and considerable water vapor will also be generated when a 30% solution is vaporized. The water vapor released with the H_2O_2 limits the gas concentration of H_2O_2 achievable because water can easily reach saturation conditions in air. When water condenses, H_2O_2 is also lost from the gas phase because it will dissolve readily in the condensate. SIP with vaporized hydrogen peroxide must therefore control not only the concentration of H_2O_2 but also water in the air when the solution is vaporized. The first step in such an SIP is usually to dehumidify the enclosure so that reproducible initial conditions are established. The temperature of the isolator and the air within is also important because this has a strong effect on the relative humidity. The kinetics of the sterilization reaction do not appear to be strongly a function of temperature.

During the sterilization step, the H_2O_2 reagent is vaporized, and the vapors are introduced into the chamber. The chamber must be leak-tight to prevent emissions and exposures to personnel. Forced convection is needed to distribute the vapors throughout the enclosure and ensure a uniform exposure to, and concentration of, H_2O_2. To maintain the gas concentration, additional H_2O_2 must usually be generated for the duration of the sterilization to replace H_2O_2 that has decomposed. The sterilization step is timed to achieve the desired level of assurance of sterility. This activity on a closed system results in a steady rise in both water vapor content and air temperature. The commercial generators are successful in part because they continuously remove water vapor from the returning gas stream, as they are generating more H_2O_2 in the supply stream. They are thus able to achieve and maintain the necessary H_2O_2 concentrations needed for a rapid sterilization (0.5–2 mg/L) while avoiding reaching water saturation conditions.

As the air temperature rises during the sterilization, a temperature difference develops between the air and the enclosure and its fixtures. Saturation conditions are likely to occur at these surfaces, and process development efforts must focus on avoiding this condition by parameter adjustment. SIP with gas requires good gas circulation to achieve a uniform exposure. Modern isolator designs (18) usually have the necessary vigorous air circulation needed because they employ it to also remove airborne particles via HEPA filtration. Fixtures inside an isolator may have restricted spaces where air circulation is ineffective. Transport of biocidal gas into these spaces then is dependent on diffusion, and dead leg effects for transport can be expected.

Graham et al. (17) reported sterilization conditions could be reached up to 60 pipe diameters into a dead leg with a pulsed cycle, which included a vacuum pulldown between cycles. The vacuum pulldown ensures that transport into the dead leg is not entirely dependent on diffusion because breaking the vacuum with sterilant creates a pressure difference for convection up the tube. Most isolators are not designed for vacuum conditions, and reliance on diffusion reduces this impressive penetration result.

The analysis of diffusional transport of H_2O_2 into dead legs can be modeled analogously to the transport of ozone into dead legs (15). In the gas phase, diffusional transport is much more rapid, but the decomposition kinetics of H_2O_2 are also larger than for ozone. The diffusivity of H_2O_2 in air is reported to be 0.188 cm^2/s at 60 °C. For such gas systems, the temperature dependence of the diffusivity should be of the order of $T^{1.5}$. This corresponds to a value of 0.167 cm^2/s at 35 °C. In comparison to ozone in water, such gas-phase diffusivities are four orders of magnitude larger. The defining parameter for penetration is the diffusivity divided by the rate constant for decomposition (15). Because H_2O_2 decomposition is observable in terms of minutes and ozone decomposition in hours, one would expect that the penetration of H_2O_2 into dead legs would be dramatically better than predicted for ozone in water (15). An industrial guideline to use for limiting dead leg

lengths with H_2O_2 has not yet evolved but would be strongly affected by the specific conditions created by the enclosure, the sterilization cycle, and the material of the dead leg, and hence not generally useful.

Finally, it must be realized that the SIP of an enclosure or isolator has the main goal of sterilizing the exposed surfaces. The air in the enclosure is usually recirculated through a sterilizing HEPA filter to keep microbes out of the air, and it therefore does not have the major bioburden. Surfaces must be clean. Microorganisms encased in dirt or a biofilm will not be exposed to the gas-phase H_2O_2 concentration because of transport resistance through the film and decomposition reactions occurring with the film. Reactive surface materials will in general reduce the local concentration of agent at the surfaces.

After the sterilization step, a H_2O_2 removal step is needed before normal operations can resume. The AMSCO VHP™ systems employ a catalytic reactor in the recirculating air stream to the enclosure to accelerate the decomposition and removal of H_2O_2. Plastic components are generally porous to H_2O_2, and thus a significant length of the cycle time often is devoted to removing the outgassing amounts of H_2O_2. This phase is again a diffusional process; but now in a solid phase and hence relatively slow.

The permeability of plastics to biocidal gases, notably ethylene oxide, has been relied on to sterilize materials sealed in plastic bags via exposure from the outside. This method has also been tested successfully with H_2O_2, but a vacuum pulsed cycle is needed (17); this step is not normally considered an SIP operation. A sterilization chamber, similar in concept to an autoclave, would be needed to alternatively evacuate the air and introduce H_2O_2.

BIBLIOGRAPHY

1. T. Myers and S. Chrai, *J. Parenter. Sci. Technol.* **35**, 8–12 (1981).
2. J. Agalloco, *J. Parenter. Sci. Technol.* **44**, 253–256 (1990).
3. K.H. Wallhäuser, *Praxis der Sterilisation Desinfektion-Konservierung*, 4th ed., G. Thieme Verlag, Stuttgart, Germany, 1988.
4. FDA, *Code of Federal Regulations Title 21*, Ch. 1, part 606.60 (Nov. 18, 1975).
5. P.T. Noble, *Biotechnol. Prog.* **8**, 275–284 (1992).
6. R.B. Bird, W.E. Stewart, and E.N. Lightfoot, *Transport Phenomena*, Wiley, New York, 1960.
7. D. Seiberling, in W.P. Olson and M.J. Groves eds., *Aseptic Pharmaceutical Manufacturing*, Interpharm Press, Prairie View, Ill., 1987, pp. 247–314.
8. J.H. Young and B.L. Ferko, *J. Parenter. Sci. Technol.* **46**, 117–123 (1992).
9. C.E. Meyrick, *Pharm. Eng.* **9**, 20–27 (1989).
10. B.D. Reeks, *British Parenteral Society Tutorial Booklet No. 2: The Validation of Steam Sterilisers*, P.S. Swindon, Wiltshire, 1990, p. 11.
11. D. Coleman and P. Smith, *Pharm. Eng.* **12**, 8–14 (1992).
12. T. Myers and S. Chrai, *J. Parenter. Sci. Technol.* **36**, 108–112 (1982).
13. H. Chalumeau, *J. Parenter. Sci. Technol.* **47**, 9–15 (1993).
14. M. Häggström, *Biotechnol. Forum Eur.* **3**, 164–167 (1992).
15. P.T. Noble, *J. Parenter. Sci. Technol.* **48**, 108–112 (1994).
16. C. Nebel and W.W. Nezgod, *Solid State Technol.* **10**, 185–193 (1984).
17. G. Graham, J. Rickloff, and J. Dalmasso, in PDA, Inc., *Proceedings of the International Congress*, Basel, Switzerland 1992, pp. 32–51.
18. P. Haas, *Pharm. Technol.* **2,** 26–38 (1995).
19. J. MacKenzie, *Chem. Eng.* **6**, 85–90 (1990).

See also BIOREACTORS, AIR-LIFT REACTORS; FILTRATION, AIR; PROCESS VALIDATION; SCALE-UP, STIRRED TANK REACTORS.

SUSPENSION CULTURE, ANIMAL CELLS

JOHN R. BIRCH
Lonza Biologics PLC
Berkshire, United Kingdom

KEY WORDS

BHK cells
Bioreactor
Cell damage
Chinese hamster ovary cells
Continuous (perfusion) suspension culture
Mammalian cell
Monoclonal antibodies

OUTLINE

INTRODUCTION

Mammalian cells can be distinguished by their requirement to grow attached to a surface (anchorage dependence) or in free suspension. The ability to grow in suspension is frequently associated with cell lines that demonstrate an "immortal" or infinite lifespan phenotype. Suspension culture systems are preferred for most large-scale manufacturing processes because scale-up is more straightforward. Relatively homogeneous conditions can be achieved in a suspension bioreactor, allowing efficient monitoring and control of key process parameters.

Suspension culture technology for animal cell culture started in the 1950s with the demonstration that several types of cell could be grown in simple agitated systems such as tumbling tubes and shaken flasks (1,2). By the end of the decade methods had been developed for growing cells in magnetically stirred spinner vessels (3,4) and in fermenter vessels similar to those used for microorganisms (5,6). By the 1960s pilot plant reactors at scales of hundreds of liters were in operation (7).

The initial drive to develop an industrial process based on mammalian cell suspension culture came from the need to produce very large volumes of vaccines against foot-and-mouth disease (FMD). Processes were developed using baby hamster kidney (BHK) cells growing in stirred-tank reactors up to 3,000 liters (8–10). Subsequently stirred-tank reactors of up to 8,000-L were used for the production of interferon α from human Namalwa cells (9). The industrial application of animal cell culture has increased dramatically over the last 20 years, driven by the need to produce monoclonal antibodies and recombinant proteins in addition to vaccines. It is the demonstration that these products can be made safely in immortal cell lines that has made possible their large-scale manufacture. Very large scale suspension processes (up to 10,000-L) have been described for the production of proteins such as tissue plasminogen activator (tPA) (11).

CELL TYPES USED FOR LARGE-SCALE PRODUCTION IN SUSPENSION CULTURE

A very wide range of cell types can be grown in suspension culture. The most important cells for industrial processes are Chinese hampster ovary (CHO), BHK, hybridomas, and mouse myelomas; these are discussed in more detail later. Other cell types grown on a large scale include insect cells (for example, see Ref. 12), especially for the production of research materials. Human cell lines such as the human kidney 293 cell line are becoming increasingly important for the production of virus vectors for gene therapy (13).

CHO Cells

The cell type most commonly used for recombinant protein production is the CHO line. CHO cells have been used to produce a wide range of therapeutic proteins (hormones, growth factors, thrombolytics, blood clotting factors, immunoglobulins). The choice of the CHO cell is based on several factors: compatibility with efficient gene expression systems leading to good productivity, ability to carry out important posttranslational modifications of proteins, and freedom from detectable pathogenic agents. In addition the cell type can be grown in large-scale suspension bioreactors. CHO cells can also grow as attached cultures, and in fact, growth in suspension generally requires a period of adaptation after the production cell line has been created. This requirement for adaptation, which can take several weeks, or months, can be circumvented by using host cells for gene transfection that have been preadapted to grow in suspension (14). Kurano et al. (15) isolated several anchorage-independent sublines of CHO, one of which grew in suspension even in static flasks. A different approach was taken by Renner et al. (16) who demonstrated that expression of recombinant cyclin E (a cell cycle regulator) in CHO cells prevented surface attachment and additionally permitted growth in protein-free medium.

BHK Cells

BHK cells have been used for large-scale production of foot-and-mouth disease vaccine because of their ability to propagate the virus and their capacity to grow in large-scale suspension culture (10). Rabies vaccine for veterinary use is also manufactured in BHK cells (17). BHK cells are also used (albeit less frequently than CHO) for production of recombinant proteins such as factor VIII (18).

Hybridomas and Myeloma Cells

Rodent monoclonal antibodies are typically produced in hybridoma cells, which can readily be grown in suspension culture (see, for example, Ref. 19). Increasing therapeutic use is being made of genetically engineered antibodies that have been "humanized" to reduce their immunogenicity in humans. Engineered antibodies can be made in CHO cells (20), but rodent myeloma cells (e.g., the mouse NSO line) are also frequently used (21), and like hybridomas, these can be grown in large-scale reactors (22).

Cells such as CHO and BHK, which can grow as attached cultures, have a tendency to form aggregates when grown in suspension (23). Aggregation can be reduced by rigorous adaptation to suspension culture and by design of culture media. Boraston et al. (24) showed that aggregation of CHO cells was reduced by adjusting the inorganic salt and amino acid composition of the culture medium. When developing cell lines for suspension culture, careful screening in an appropriate system (e.g., shake-flask culture) is essential. Brand et al. (25) compared the specific production rate in attached and in suspension culture of six closely related CHO cell lines, all making the same monoclonal antibody. Production rates were higher, lower, or the same depending on the particular cell line. The same authors found similar results with NSO myeloma cells; measurement of specific production rate in static culture for clones from the same transfection was of limited value in predicting results in suspension culture.

SUSPENSION CULTURE REACTORS

Stirred Reactors

The most commonly used systems for suspension culture are based on stirred reactors used in both batch and per-

fused configurations (Table 1). Typically reactors are stainless steel with height-to-diameter ratios in the range 1:1 to 3:1. In some cases mammalian cell bioreactors have been developed by retrofitting microbial vessels, usually by changing the agitators and aeration system. Backer et al. (19) described the retrofitting of 150- and 1,300-L reactors. Rushton impellers were replaced with marine propellers, and the agitator drives were altered to allow operation between 25 and 250 rpm. Air was sparged through a sintered stainless steel sparger with 10-μm pore size. Garnier et al. (12) describe the retrofitting of a 150-L microbial bioreactor (height-to-diameter ratio 3). They used two large pitched-blade impellers (45°), three surface baffles, and a polypropylene porous sparger (80-μm pore size). Agitation was controlled at 60–120 rpm. Nienow et al. (29) found that Intermig impellers improved mixing times in a mammalian reactor compared with the Rushton turbines commonly used in microbial reactors. Impellers may be directly driven via a shaft or indirectly driven through a magnetically coupled drive (30,31). Magnetic drives remove the potential for microbial ingress through mechanical shaft seals. The operation of mammalian reactors differs from that of microbial reactors in one very important respect: power inputs are very low to minimize the risk of damage to cells. For an 8,000-L reactor it has been reported (29) that a maximum agitator speed of 1.5 rev/s is used, giving a maximum power input of 30 W/m^3. This is around two orders of magnitude less than the value typically achieved in microbial fermentations. Similarly very low aeration rates are used in animal cell reactors compared with microbial cultures. This combination of low power input and low aeration rate means that particular attention has to be paid to the design of vessels to ensure good bulk mixing and adequate mass transfer of gases (especially oxygen and carbon dioxide). Nienow et al. (29) describe a study of mixing characteristics and oxygen uptake rates in an 8,000-L reactor and discuss ways of improving reactor performance.

Airlift Reactors

The airlift reactor is used less commonly than the stirred reactor for cell culture and was originally developed for microbial fermentation. Sparging of air or other gas mixtures is used to aerate the culture and to provide mixing. Mixing is achieved by separating the vessel into riser and downcomer sections using a draught tube, divider plate, or external recirculation loop. The presence of gas bubbles in the riser section creates a density difference between riser and downcomer that causes the fluid to circulate, providing mixing. Airlift reactors tend to have larger height-to-diameter ratios than stirred reactors (typically as high as 10:1). Their advantage is simplicity (no mechanical agitators) combined with good mass transfer and low shear characteristics (32,33). The basic design features of airlift reactors have been described in reviews (34,35). Airlift reactors have been used for the culture of BHK and human lymphoblastoid cells (36), insect cells (37), and at scales up to 2,000 L for a variety of cell types including CHO, hybridomas, and myelomas (38,39) (Fig. 1). Growth and productivity profiles for a GS-NSO cell line making a recombinant antibody in an airlift reactor are shown in Figure 2 (from Ref. 40).

Operational Considerations

While the type of vessel used for suspension culture is usually a stirred reactor (occasionally airlift), the mode of reactor operations can vary significantly. Production methods have been developed based on batch, fed-batch, and continuous (perfusion) systems. The choice of process will be dictated by several factors. Batch and fed-batch culture have the advantage of simplicity, reliability, flexibility, and reduced timescales for development and validation (compared with continuous systems). On the other hand continuous systems may have significant economic advantages, particularly for high-volume products, and thorough cost modeling is necessary to determine the most appropriate route for a given product. Griffiths (41) has reviewed the benefits of continuous processes.

Batch and Fed-Batch Culture

Batch culture is very commonly used in industry. Cells and medium are added to the reactor, and the culture is harvested after an appropriate period of culture. The time of harvest will depend on the kinetics of product accumulation and, in some cases, on the stability of product within the reactor. A commonly used variation on this type of culture is repeated batch or "draw and fill" culture. This is a batch culture in which a proportion of the culture is retained in the reactor as an inoculum, and fresh medium is added. This allows repeated batch growth cycles within the reactor, obviating the need to repeatedly clean, sterilize,

Table 1. Examples of Industrial-Scale Mammalian Cell Suspension Culture Processes

Product	Cell Line	Reactor	Reference
Foot-and-mouth disease vaccine	BHK 21	3,000 L	Pullen et al. (9)
Interferon	Namalwa	8,000 L	Pullen et al. (9)
Tissue plasminogen activator (tPA)	CHO	10,000 L	Lubiniecki et al. (11)
Factor VIII	BHK 21	Perfused 500-L reactor with cell recycle system	Bödeker et al. (18)
Factor VIII	CHO	2,500-L batch refeed	Adamson (26)
Monoclonal antibodies	Hybridoma/myelomas	Perfused 500-L spin filter reactor	Deo et al. (27)
Monoclonal antibody	NSO myeloma	2,000-L fed-batch reactor	Ray et al. (22)
Monoclonal antibody	Hybridoma	1,300-L batch, draw and fill	Backer et al. (19)
Monoclonal antibody	Myeloma	2,000-L airlift reactor	Brown et al. (28)

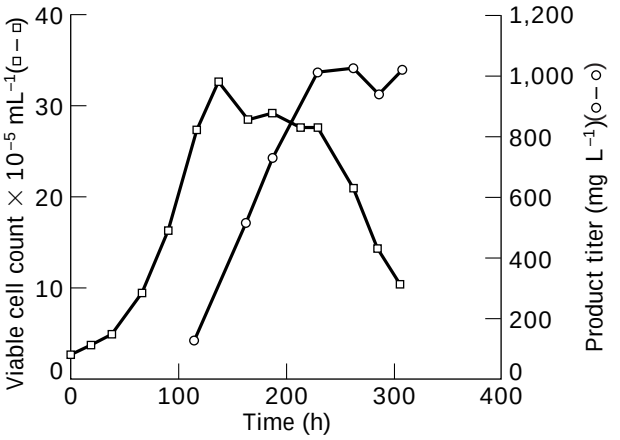


Figure 1. Growth and productivity of a GS-NSO cell line making a recombinant antibody in airlift culture. *Source*: From Ref. 40.

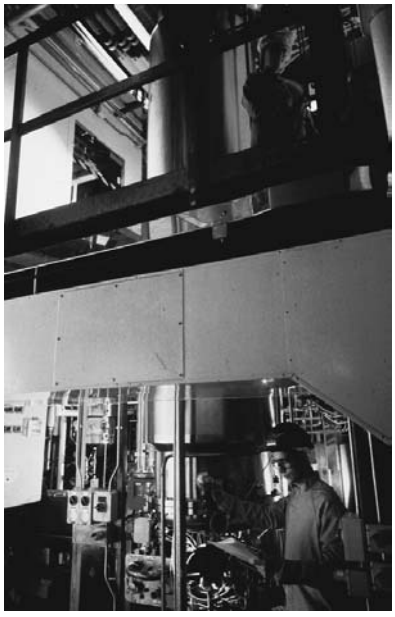


Figure 2. A 2,000-liter airlift bioreactor.

and reinoculate the vessel. Backer et al. (19) described such a system, which was operated for several months at a time, for monoclonal antibody production in a 1,300-L reactor. Every 84 h, the reactor was harvested leaving, 10% of the culture in the reactor as inoculum for the next culture.

A further commonly used variant on batch culture is fed-batch culture in which small volumes of selected nutrients are fed to the culture during the growth cycle to improve productivity. Bibila and Robinson (42) and Bibila et al. (43) describe the optimization of a feeding strategy for production of monoclonal antibodies. The use of a multinutrient feed to increase cell culture longevity resulted in antibody titers in the range of 1 to 2 g/L. A similar feeding strategy led to a 10-fold increase in productivity for large-scale myeloma cultures producing recombinant mouse and rat growth hormones (44). Ray et al. (22) described the fed-batch culture of GS-NSO cells making a humanized monoclonal antibody in a 2,000-L stirred-tank bioreactor, using a supplemental feed containing glucose and amino acids. Xie and Wang (45–48) have developed a feeding strategy based on a stoichiometric model of the demands of nutrients for growth and product formation. The strategy is aimed at preventing nutrient depletion while reducing the formation of toxic products of metabolism such as ammonia and lactate (by avoiding excess concentration of the nutrient precursors of these products). Application of this strategy in 2-L fed-batch cultures of hybridoma cells (45) resulted in very high cell densities (1.7×10^7 cells/mL), extended culture span (550 h), and a final monoclonal antibody concentration of 2.4 g/L. This approach clearly demonstrates the potential for improving the productivity of suspension culture processes. It may ultimately be possible to use on-line measurement to estimate nutrient utilization rate and feed stoichiometric amounts of nutrient to support growth and reduce accumulation of toxic metabolites. Zhou et al. (49) demonstrated such a system using on-line measurement of oxygen uptake rate to automatically control the feed to a 500-mL hybridoma culture. Maximum cell concentrations were increased from 2×10^6/mL in batch culture to 1.36×10^7/mL in the fed culture. Kurokawa et al. (50) used on-line measurement of glucose and glutamine, using HPLC, to control the concentration of these nutrients in fed batch culture at low levels (0.2 and 0.1 g/L, respectively). This strategy reduced the accumulation of lactate and ammonia and gave an approximately twofold increase in cell growth and antibody production.

Continuous Culture Systems

Continuous (perfusion) suspension culture systems are now being used for large-scale manufacturing operations. For a discussion of perfusion culture technology and the engineering issues involved in designing and operating such systems, the reader is referred to the review by Ozturk (51). Perfusion culture uses a continuous supply of culture medium to the reactor and a continuous offtake of spent culture fluid containing product. By retaining cells in the reactor (by preventing them leaving the vessel or by an external recycle device), very high cell densities and consequently very high product throughput can be achieved. The earliest device of this kind was the spin filter culture developed by Himmelfarb et al. (52). A spinning filter within the reactor allowed cell retention, and with continuous perfusion of culture medium, cell densities approaching 10^8/ml were achieved.

Several groups have described production systems based on the spin filter principle. Avgerinos et al. (53) described a 20-L perfused bioreactor for production of a plasminogen activator from CHO cells. Deo et al. (27) described the design and operation of a 500-L perfused spin filter reactor for monoclonal antibody production. They note that continuous perfusion generally gives an approximately 10-fold greater volumetric productivity in comparison with batch and fed-batch processes. Hence less reactor capacity is required. At an antibody titer of 400 to 500 mg/L the 500-L reactor is capable of producing greater than 1 kg of

antibody per week, which they estimate would require a 5,000- to 10,000-L bioreactor operating in batch or fed-batch mode. Successful operation of spin filter systems depends on careful optimization of the filter system to prevent fouling (27,54).

Other cell retention devices have also been described in the literature. Roth et al. (55) describe an external vortex flow filtration device used with perfused 3-L cultures of CHO, HEK 293, HeLa, and Sp2 cell lines. Hülscher et al. (56) describe the use of a settling tank for cell retention in a perfused 5.4-L airlift loop reactor. Pui et al. (57) have developed an acoustic resonator to cause cell aggregation and sedimentation back into the culture. At industrial scale Bödeker et al. (18) (Table 1) have used a perfused 500-L bioreactor with an external cell retention/recycle system to manufacture recombinant factor VIII from BHK cells. This is apparently the first approved biopharmaceutical from recombinant mammalian cells made using perfusion culture. The authors discuss the advantages of this process for factor VIII production, which requires the generation of greater than one million liters of harvest. These advantages are a high degree of culture control compared with batch operation (operates in steady-state conditions) and reduction in reactor size (hence easier to operate, clean, and sterilize). The main disadvantage reported was the extended process validation required. This validation included process stability, stability of product characteristics, and genetic stability of the cell line over the duration of the culture, which was up to 185 days. Several full-length campaigns were run to establish process consistency.

An alternative approach to continuous culture is to use a chemostat in which growth rate is controlled by the supply of a single limiting nutrient. The advantage of this system is that true physiological steady-state conditions can be maintained indefinitely, allowing very precise control of culture conditions, and this can be extremely valuable in a variety of research applications for studying cell physiology (see for example Refs. 58–60). To date, however, the technique has not been applied in industrial processes.

Process Monitoring and Control

The key parameters that are usually monitored and controlled in industrial-scale suspension cultures are pH, temperature, and dissolved oxygen. There is also increasing interest in the use of devices, which are now available, that allow real-time monitoring of biomass concentration (61). Such devices will allow improved monitoring and control of fermentations, particularly where fed-batch and perfusion strategies are adopted.

pH Control. Mammalian cell growth metabolism and productivity can be profoundly influenced by even small changes in pH of the culture. Ozturk and Palsson (62) found that the growth rate for a hybridoma was optimum at pH 7.2 but that the specific antibody production rate increased twofold at pH values below 7.2. They also showed that glucose and glutamine uptake rates and production rates for lactate and ammonia increased at higher pH values. Wayte et al. (63) reported that differences in pH of as little as 0.1 unit could profoundly affect cell growth and/or productivity of hybridoma and myeloma cell lines. The effects observed were cell line specific. pH may also influence product characteristics, especially glycosylation. Borys et al. (64) showed that pH affected the glycosylation pattern and specific expression rate of a recombinant lactogen protein made in CHO cell culture. Hence it is essential to establish the optimum pH for a given cell line and to measure and control pH precisely. pH is typically controlled at a value in the range 7.0 to 7.5. Most cell culture media rely on a bicarbonate-CO_2 buffering system. Hence when a bioreactor is sparged with air the pH can be controlled by injecting a controlled flow of CO_2 on demand to maintain a constant pH value. It is frequently found that lactic acid accumulates to levels that cannot be controlled by sparging with air and CO_2; it is then common practice to control the pH by addition of a sterile base solution such as sodium bicarbonate (19) or sodium carbonate (29). Care is needed in designing bioreactors to avoid pH gradients during addition of base solution. Regions of high pH may arise, especially if base is added to a poorly mixed zone (often the upper region) in both stirred reactors (29) and airlift reactors (63). The problem can be minimized by adding base to a well-mixed region of the reactor.

Dissolved Oxygen Concentration. Animal cells have an absolute requirement for oxygen. Their are several ways in which it can be provided to cultures, but the most common method is to bubble or "sparge" air or oxygen into the reactor. A less commonly used method involves gas exchange across gas permeable membranes inside or outside the reactor (51,65). Dissolved oxygen concentrations can be measured in situ with a dissolved oxygen electrode and controlled automatically by adjusting the flow rate of air or oxygen to the reactor. The optimum dissolved concentration varies from one cell line to another, although most cell types will grow over quite a wide range of dissolved oxygen concentrations. Boraston et al. (58) found that growth of a hybridoma cell was relatively unaffected over the range 8–100% air saturation. Ozturk and Palsson (62) showed that growth rates of hybridoma cells were not influenced at dissolved oxygen concentrations in the range 20–80% air saturation. Specific death rates were lowest in the range 20–50% air saturation. Nienow et al. (29) found that a CHO cell line producing recombinant interferon γ grew well and had similar productivity over the range 20–80% saturation, but growth and productivity were reduced at 5 and 100% saturation. The oxygen uptake rate of animal cells is typically in the range 0.05 to 0.5 mmol $O_2/10^9$ cells/h (66). It is of course necessary to ensure that oxygen transfer into the culture is sufficient to meet the uptake rate of the cells, and this becomes an increasingly important issue as higher cell densities are achieved through improvements in process design. Factors influencing oxygen transfer include height-to-diameter ratio for the vessel, sparger and impeller designs, and gas flow rates. Tramper (67) has pointed out that gradients of dissolved oxygen concentration may occur in large bioreactors. It is not clear what impact such gradients have in practice, but they should be borne in mind and minimized when design-

ing reactors. Aeration in cell culture bioreactors has been reviewed in detail by Aunins and Henzler (65).

Dissolved Carbon Dioxide Concentration. In addition to CO_2 added to the culture as part of the pH buffering system, account must be taken of CO_2 generated by cell metabolism. Although systems for in situ monitoring and control of CO_2 are not yet commonly used, it is useful to measure CO_2 levels off-line, for example, using a clinical blood gas analyser. In recent years it has become apparent that in some large-scale processes CO_2 can accumulate to levels that inhibit cell growth and recombinant product formation (12,65–69). It has been suggested (65) that CO_2 accumulation may be the factor that will ultimately limit the reactor scale for some processes. Kimura and Miller (70) point out that inhibition may be caused by an effect of CO_2 on intracellular pH (even if the culture pH is controlled) and may also be caused by an increase in osmolarity caused by the additional base required to maintain pH.

Gray et al. (69), working with CHO cells in a high density perfused process, showed that productivity of a recombinant product was maximized when CO_2 was maintained in the range 30–76 mm Hg. Levels greater than 105 mm Hg resulted in inhibition of growth and productivity. CO_2 concentration may also influence glycosylation of recombinant proteins (70). Aeration systems designed for efficient oxygen transfer will not necessarily be efficient at stripping out CO_2. Gray et al. (69) found that sparging with microbubbles of pure oxygen gave very efficient oxygen transfer but was inefficient for CO_2 removal because the bubbles dissolved before reaching the surface of the culture. Sparging with large bubbles (2–3 mm diameter) improved CO_2 removal while retaining adequate oxygen transfer. Zhou et al. (44), working with GS-NSO cell lines making recombinant mouse and rat growth hormones in a 250-L fed batch reactor, also found that use of a sintered steel frit delivering pure oxygen gave good oxygen transfer, but CO_2 accumulated to 200 mm Hg, depressing cell growth. By using a less efficient sparger (12-mm $\times$ 1.5-mm orifice ring sparger) CO_2 accumulation was reduced to a noninhibitory level.

Temperature. Temperature is typically controlled at around 37 °C. Bloemkolk et al. (71) studied the effect of temperature on growth and antibody production of a hybridoma in suspension culture; 37 °C was the optimum temperature for cell growth and antibody production. At 39 °C growth was completely inhibited. In some instances the optimum temperature for product accumulation may be below 37 °C (72).

Culture Media for Suspension Culture

In general media for suspension culture are similar to those used for other processes, with the exception that an agent to protect against mechanical damage is frequently added (see later section on mechanical damage). For reactors that are aerated by sparging it may also be necessary to use an antifoam agent, which is usually a silicone emulsion (65,73). In some instances changes in nutritional requirements have been observed in suspension as opposed to static culture. For example, Murakami (74) showed that a myeloma cell line in suspension culture had a requirement for a phospholipid component that was not required in static culture.

Significant progress is being made in the optimization of media composition as well as of nutrient feeds for fed-batch culture (see earlier section). Brown et al. (28) described how an iterative program of medium development gave a six- to sevenfold improvement in productivity of antibody from a recombinant NSO myeloma cell line. Addition of a feed gave a further twofold increase in productivity. Jo et al. (75) increased maximum cell concentrations significantly for several cell lines by increasing the nutrient concentrations in RPMI 1640 medium.

There is an increasing drive to remove animal-derived raw materials from cell culture processes to reduce the risk of inadvertently introducing adventitious agents. Encouraging progress has been made, and there are now several examples in the literature of industrially important cell lines being grown in chemically defined protein-free medium. Keen (76) described the growth of myeloma and hybridoma cell lines in a protein-free medium in shake flask culture. Zang et al. (77) showed that CHO cell lines making recombinant proteins could be grown in protein-free culture medium.

PREVENTION OF CELL DAMAGE IN BIOREACTORS

Two mechanisms have been identified that can potentially cause cell damage in bioreactors; the first is the interaction of cells with bubbles in sparged reactors, and the second is the stresses created by impellers at high agitation rates in stirred reactors. In practice, given the low impeller speeds used in animal cell reactors, interaction with bubbles is the only likely source of damage in aerated cultures (65,78). The problem is readily overcome by incorporation of a protective polymer such as the nonionic surfactant Pluronic F68 into the culture medium. For a review of protective additives see Ref. 78. The role of these polymers is not totally clear. They may act by preventing cells from attaching to bubbles, by altering the way in which bubbles burst when they disengage at the surface of the culture, or by adsorbing to the surface of cells and strengthening them (65,79,80).

CONCLUSIONS

Large-scale animal suspension cell cultures have become an extremely important source of biopharmaceutical products, and with the steadily increasing number of biotechnology products in clinical trial (81), this trend seems set to continue. New areas of medical development, for example in gene therapy, give rise to new process development challenges that will lead to new uses for large-scale

suspension culture. The requirement for large volumes of some animal-cell-derived pharmaceuticals drives the need to develop processes that are highly productive and cost efficient. Currently two approaches are being adopted: the use of very large bioreactors (to benefit from economies of scale), and the use of continuous (perfusion) culture using high-density cell culture at smaller scale. Advances in reactor design and operations are being complemented by the use of efficient gene expression technology and by continuing improvement in the design of culture media, especially chemically defined media.

BIBLIOGRAPHY

1. O. Owens, M.K. Gey, and G.O. Gey, *N.Y. Ann. Acad. Sci.* **58**, 1039–1055 (1954).
2. W.R. Earle, E.H. Schilling, J.C. Bryant, and V.J. Evans, *J. Natl. Cancer Inst.* **14**, 1159–1171 (1954).
3. W.R. Cherry and R.N. Hull, *Anat. Rec.* **124**, 483 (1956).
4. W.F. McLimans, E.V. Davis, F.L. Glover, and G.W. Rake, *J. Immunol.* **79**, 428–433 (1957).
5. W.F. McLimans, and F.R. Giardinello, *J. Bacteriol.* **74**, 768–771 (1957).
6. D.W. Ziegler, E.V. Davis, W.J. Thomas, and W.F. McLimans, *Appl. Microbiol.* **6**, 305–310 (1958).
7. G.E. Moore, P. Hasenpusch, R.E. Gerner, and A.A. Burns, *Biotechnol. Bioeng.* **10**, 625–640 (1968).
8. P.B. Capstick, R.C. Telling, W.G. Chapman, and D.L. Stewart, *Nature* **195**, 1163–1164 (1962).
9. K.F. Pullen, M.D. Johnson, A.W. Phillips, G.D. Ball, and N.B. Finter, *Dev. Biol. Stand.* **60**, 175–177 (1985).
10. P.J. Radlett, T.W.F. Pay, A.J.M. Garland, *Dev. Biol. Stand.* **60**, 163–170 (1985).
11. A. Lubiniecki, R. Arathoon, G. Polastri, J. Thomas, M. Wiebe, R. Garnick, A. Jones, R. van Reis, and S. Builder, in R.E. Spier, J.B. Griffiths, J. Stephenne, P.J. Crooy, eds., *Advances in Animal Cell Biology and Technology for Bioprocesses*, Butterworth, London, 1989, pp. 442–449.
12. A. Garnier, R. Voyer, R. Tom, S. Perret, B. Jardin, and A. Kamen, *Cytotechnology* **22**, 53–63 (1996).
13. I. Nadeau, A. Garnier, J. Côté, B. Massie, C. Chavarie, and A. Kamen, *Biotechnol. Bioeng.* **51**, 613–623 (1996).
14. M.S. Sinacore, T.S. Charlebois, S. Harrison, S. Brennan, T. Richards, M. Hamilton, S. Scott, S. Brodeur, P. Oakes, M. Leonard, M. Switzer, A. Anagnostopoulos, B. Foster, A. Harris, M. Jankowski, M. Bond, S. Martin, and S.R. Adamson, *Biotechnol. Bioeng.* **52**, 518–528 (1996).
15. N. Kurano, C. Leist, F. Messi, C. Gandor, S. Kurano, and A. Fiechter, *J. Biotech.* **16**, 245–258 (1990).
16. W.A. Renner, K.H. Lee, V. Hatzimanikatis, J.E. Bailey, and H.M. Eppenberger, *Biotechnol. Bioeng.* **47**, 476–482 (1995).
17. T.W.F. Pay, A. Boge, F.J.R.R. Menard, and P.J. Radlett, *Dev. Biol. Stand.* **60**, 171–174 (1985).
18. B.G.D. Bödeker, R. Newcomb, P. Yuan, A. Braufman, W. Kelsey, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology, Products for Today, Prospects for Tomorrow*, Butterworth-Heinemann, Oxford, 1994, pp. 580–583.
19. M.P. Backer, L.S. Metzger, P.L. Slaber, K.L. Nevitt, and G.B. Boder. *Biotechnol. Bioeng.* **32**, 993–1000 (1988).
20. C.R. Wood, A.J. Dorner, G.E. Morris, E.M. Alderman, D. Wilson, R.M. O'Hara, R.J. Kaufman, *J. Immunol.* **145**, 3011–3016 (1990).
21. C.R. Bebbington, G. Renner, S. Thomson, D. King, D. Abrams, and G.T. Yarranton, *BioTechnology* **10**, 169–175 (1992).
22. N.G. Ray, R. Rivera, R. Gupta, and D. Mueller, in M.J.T. Carrondo et al. eds., *Animal Cell Technology*, Kluwer Academic, Dordrecht, 1997, pp. 235–241.
23. J.L. Moreira, P.M. Alves, J.G. Aunins, and M.J.T. Carrondo, *Appl. Microbiol. Biotechnol.* **41**, 203–209 (1994).
24. R. Boraston, C. Marshall, P. Norman, G. Renner, and J. Warner, in R.E. Spier, J.B. Griffiths, and C. Macdonald eds., *Animal Cell Technology: Developments, Processes and Products*, 1992, pp. 424–426.
25. H.N. Brand, S.J. Froud, H.K. Metcalfe, A.O. Onadipe, A. Shaw, and A.J. Westlake, in R.E. Spier, J.B. Griffiths, and W. Berthold eds., *Animal Cell Technology Products of Today, Prospects for Tomorrow*, Butterworth-Heinemann, Oxford, 1994, pp. 55–60.
26. R. Adamson, *Ann. Hematol*, **68**, S9–S14 (1994).
27. Y.M. Deo, M.D. Mahadevan, and R. Fuchs, *Biotechnol. Prog.* **12**, 57–64 (1996).
28. M.E. Brown, G. Renner, R.P. Field, and T. Hassell, *Cytotechnology* **9**, 231–236 (1992).
29. A.W. Nienow, C. Langheinrich, N.C. Stevenson, A.N. Emery, T.M. Clayton, and N.K.H. Slater, *Cytotechnology* **22**, 87–94 (1996).
30. J. Cameron and E.I. Godfrey, *Biotechnol. Bioeng. Symp.* **4**, 821–835 (1974).
31. N.B. Finter, in A.S. Lubiniecki ed., *Large Scale Mammalian Cell Culture Technology*, Marcel Dekker, New York and Base, 1990, pp. 1–14.
32. H. Katinger and W. Scheirer in R.E. Spier, and J.B. Griffiths eds., *Animal Cell Biotechnology*, vol. 1, Academic, London, Orlando, 1985, pp. 167–193.
33. J.R. Birch, K. Lambert, R. Boraston, P.W. Thompson, S. Garland, and A.C. Kenney, in J.M.C. Duarte, L.J. Archer, A.T. Bull, and G. Holt eds., *Perspectives in Biotechnology*, Plenum, New York and London, 1987, pp. 97–109.
34. K.C. Merchuk and M.H. Siegel, *J. Chem. Tech. Biotechnol.* **41**, 105–120 (1988).
35. U. Onken and P. Weiland, in *Adv. Biotechnol. Processes* **1**, 67–95 (1983).
36. H.W.D. Katinger, W. Scheirer, and E. Krömer, *Ger. Chem. Eng.* **2**, 31–38 (1979).
37. B. Maiorella, D. Inlow, A. Shauger, and D. Harano, *BioTechnology* **6**, 1406–1410 (1988).
38. W.R. Arathoon, and J.R. Birch, *Science* **232**, 1390–1395 (1986).
39. M. Rhodes and J. Birch, *BioTechnology* **6**, 518–523 (1988).
40. J.R. Birch, J. Bonnerjea, S. Flatman, and S. Vranch, in J.R. Birch and E.S. Lennox eds., *Monoclonal Antibodies: Principles and Applications*, Wiley-Liss, New York, 1995 pp. 231–265.
41. J.B. Griffiths, in R. Sasaki and K. Ikura eds., *Animal Cell Culture and Production of Biologicals*, Kluwer Academic, Dordrecht, 1991, pp. 401–410.
42. T.A. Bibila and D.K. Robinson, *Biotechnol. Prog.* **11**, 1–13 (1995).
43. T.A. Bibila, C.S. Ranucci, K. Glazomitsky, B.C. Buckland, and J.G. Aunins, *Biotechnol. Prog.* **10**, 87–96 (1994).
44. W. Zhou, T. Bibila, K. Glazomitsky, J. Montalvo, C. Chan, D. Di Stefano, S. Munshi, D. Robinson, B. Buckland, and J. Aunins, *Cytotechnology* **22**, 239–250 (1996).
45. L. Xie and D.I.C. Wang, *Cytotechnology* **15**, 17–29 (1994).

46. L. Xie and D.I.C. Wang, *Biotechnol. Bioeng.* **43**, 1164–1174 (1994).
47. L. Xie and D.I.C. Wang, *Biotechnol. Bioeng.* **43**, 1175–1189 (1994).
48. L. Xie and D.I.C. Wang, *Biotechnol. Bioeng.* **51**, 725–729 (1996).
49. W. Zhou, J. Rehm, and W.-S. Hu, *Biotechnol. Bioeng.* **46**, 579–587 (1995).
50. H. Kurokawa, Y.S. Park, S. Iijima, and T. Kobayashi, *Biotechnol. Bioeng.* **44**, 95–103 (1994).
51. S. Ozturk, *Cytotechnology* **22**, 3–16 (1996).
52. P. Himmelfarb, P.S. Thayer, and H.E. Martin, *Science* **164**, 555–557 (1969).
53. G.C. Avgerinos, D. Drapeau, J.S. Socolow, J.-I. Mao, K. Hsiao, and R.J. Broeze, *BioTechnology* **8**, 54–58 (1990).
54. V.M. Yabannavar, V. Singh, and N.V. Connelly, *Biotechnol. Bioeng.* **40**, 925–933 (1992).
55. G. Roth, C.E. Smith, G.M. Schoofs, T.J. Montgomery, J.L. Ayala, and J.I. Horwitz, *Pharmaceutical Technology*, October 1997, and *BioPharm* **10**, 30–35 (1997).
56. M. Hülscher, U. Scheibler, and U. Onken, *Biotechnol. Bioeng*, **39**, 442–446 (1992).
57. P.W.S. Pui, F. Trampler, S.A. Sonderhoff, M. Groeschl, D.G. Kilburn, and J.M. Piret, *Biotechnol. Prog.* **11**, 146–152 (1995).
58. D.K. Robinson, and K.W. Memmert, *Biotechnol. Bioeng.* **38**, 972–976 (1991).
59. P.M. Hayter, E.M.A. Curling, M.L. Gould, A.J. Baines, N. Jenkins, I. Salmon, P.G. Strange, and A.T. Bull, *Biotechnol. Bioeng.* **42**, 1077–1085 (1993).
60. R. Boraston, P.W. Thompson, S. Garland, and J.R. Birch, *Dev. Biol. Stand.* **55**, 103–111 (1984).
61. K. Konstantinov, S. Chuppa, E. Sajan, Y. Tsai, S. Yoon, and F. Golini, *Tibtech* **12**, 324–333 (1994).
62. S.S. Ozturk, and B.O. Palsson, *Biotechnol. Prog.* **7**, 481–494 (1991).
63. J. Wayte, R. Boraston H. Bland, J. Varley, and M. Brown, *Genet. Eng. Biotechnol.* **17**, 125–132 (1997).
64. M.C. Borys, D.I.H. Linzer, and E.T. Papoutsakis, *BioTechnology* **11**, 720–724 (1993).
65. J.G. Aunins, and H.-J. Henzler, in G. Stephanopoulos ed., *Biotechnology*, vol. 3, *Bioprocessing*, VCH, Berlin, 1993, pp. 219–281.
66. R.J. Fleischaker, and A.J. Sinskey, *Eur. J. Appl. Microbiol. Biotechnol.* **12**, 193–197 (1981).
67. J. Tramper, *Cytotechnology* **18**, 27–34 (1995).
68. R. Kimura and W.M. Miller, *Biotechnol. Bioeng.* **52**, 152–160 (1996).
69. D.R. Gray, S. Chen, W. Howarth, D. Inlow, and B.L. Maiorella, *Cytotechnology* **22**, 65–78 (1996).
70. R. Kimura, and W.M. Miller, *Biotechnol. Prog.* **13**, 311–317 (1997).
71. J.W. Bloemkolk, M.R. Gray, F. Merchant, and T.R. Mosmann, *Biotechnol. Bioeng.* **40** 427–431 (1992).
72. J.P. Mather, and M. Tsao, in A.S. Lubiniecki ed., *Large-Scale Mammalian Cell Culture Technology*, Marcel Dekker, New York and Basel, 1990, pp. 161–177.
73. L.A. van der Pol, D. Bonarius, G. van de Wouw, and J. Tramper, *Biotechnol. Prog.* **9**, 504–509 (1994).
74. H. Murakami, T. Edamoto, K. Shinohara, and H. Omura, *Agric. Biol. Chem.* **47**, 1835–1840 (1983).
75. E.-C. Jo, D.-I. Kim, and H.M. Moon, *Biotechnol. Bioeng.* **42**, 1218–1228 (1993).
76. M.J. Keen, *Cytotechnology* **17**, 193–202 (1995).
77. M. Zang, H. Trautmann, C. Gandor, F. Messi, F. Asselbergs, C. Leist, A. Fiechter, and J. Reiser, *BioTechnology* **13**, 389–392 (1995).
78. T. Papoutsakis, *Tibtech* **9**, 316–324 (1991).
79. A. Handa-Corrigan, A.N. Emery, and R.E. Spier, *Enzyme Microb. Technol.* **11**, 230–235 (1989).
80. Z. Zhang, M. Al-Rubeai, and C.R. Thomas, *Enzyme Microb. Technol.* **14**, 980–983 (1992).
81. J.R. Birch, *Eur. J. Parent Sci.*, Biotechnology Special Issue, pp. 3–10 (1997).

See also MAMMALIAN CELL BIOREACTORS.

SWEETENERS. See ASPARTAME.

THERMAL UNFOLDING, PROTEINS

LI SHI
DAVID B. VOLKIN
Merck Research Laboratories
West Point, Pennsylvania
GAUTAM SANYAL
Astra Research Center Boston
Cambridge, Massachusetts

KEY WORDS

Protein folding
Protein pharmaceuticals
Protein stability
Secondary structure
Tertiary structure
Thermal denaturation
Unfolding kinetics

OUTLINE

INTRODUCTION

Recombinant therapeutic proteins have emerged as important pharmaceuticals for a wide variety of indications in a number of clinical disciplines. These include immunoregulatory proteins, peptide and protein hormones, growth factors, and immunoglobulins. Developing stable formulations of proteins offers challenges related to the inherent instability of their native three-dimensional structures that must usually be maintained for biological function (1,2). The pharmaceutical aspects of developing stable protein formulations include designing ways to preserve their chemical and conformational integrity as measured by sensitive, stability-indicating assays.

Protein inactivation can occur by either or both of two distinct pathways: chemical degradation, including hydrolytic and proteolytic clipping, and physical inactivation, including global unfolding, a more subtle local conformational change involving a functionally critical site. Protein stability studies must include assays that can quantitatively determine the extents of physical and chemical changes as well as changes in biological activity that may result from perturbations introduced by storage, handling, and formulation conditions. A well-designed study using various biophysical and biochemical techniques is essential to monitoring protein structure and activity as a function of formulation processing. Protein conformational stability studies under various stress conditions have been widely applied in protein formulation research. The techniques include both thermal, pH, and denaturant induced unfolding. In this article, we discuss specifically the application of thermal unfolding studies to formulation development of protein-based pharmaceuticals.

MECHANISMS OF PROTEIN UNFOLDING

The native or functional structure of a protein is a minimum energy state that is held together by an intricate balance of a multitude of covalent and noncovalent interactions. Perturbations of these interactions can lead to unfolding of the native conformation and will usually lead to the loss of biological activity. In general, there are two major types of noncovalent interaction contributing to the stability of a folded protein structure: hydrophobic interactions between apolar side chains that prefer to exclude water and electrostatic or polar interactions mediated by water. The latter includes hydrogen bonding and salt bridges. Although from a macroscopic point of view a protein molecule may appear to be in a completely folded state, the so-called native protein conformation is in fact in a fast exchange between folded or native (N) and unfolded (U) states in a dynamic equilibrium, with an overwhelming majority of the population being in the folded state. The equilibrium constant, K, describing the population distribution in these two states, is strongly dependent on the environment of the protein and is correlated to the Gibbs energy, ΔG, of the protein molecule through relation $\Delta G = -RT\ln K$.

Protein unfolding occurs when the balance of forces between the intramolecular interactions in the protein and the protein's interaction with its environment is disrupted. This may be a result of perturbation of normal water struc-

ture around the protein. The conformational stability of most proteins is sensitive to temperature. Increased temperatures lead to increased thermal motions in proteins, which, in turn, perturb intermolecular interactions as well as solvent–protein interactions. When the environment of a protein is altered because of changes of temperature, pressure, pH, or solvent polarity, the equilibrium balance between unfolded and native states is altered, and a new equilibrium is reached in which the unfolded state may be favored. However, this population redistribution is not significantly noticeable until an apparent conformational transition is approached in the unfolding process.

Disruption of Secondary and Tertiary Structure

It is important to understand the mechanism of the disruption of protein secondary and tertiary structures when the protein is exposed to stress conditions. The secondary structure of a protein refers to the spatial arrangement of amino acid residues that are near one another in the linear sequence (3–6). Some of these steric relationships manifest a regular repeat pattern, giving rise to periodic structures. The α-helices, β-sheets, turns, and random coils are elements of secondary structure. Tertiary structure refers to the overall three-dimensional architecture of the spatial arrangement of amino acid residues, including those that may be far apart in the linear sequence (3–6). This spatial arrangement is also referred to as the packing assembly of secondary structure units.

The net stabilization free energy that maintains the native folded structure of a globular protein is in the order of only 2 to 20 kcal/mol. Therefore, a few kcal of additional energy gained from the interaction changes of the protein with the environment can result in the destabilization of the protein molecule resulting in either local or global conformational changes (7). Information about obtaining the ΔG value of a protein is available (8).

It is necessary to point out that thermodynamic analysis of the free energy of protein conformational change is limited to reversible conformational change or unfolding, because the relation $\Delta G = -RT\ln K$ is only valid when the mass balance is retained for all species involved in the conformational change. For a thermally induced reversible unfolding transition, the dependence of ΔG on T (and hence the dependence of K on T) is given by $\Delta G = \Delta H - T\,\Delta S$. ΔH and ΔS are, respectively, the enthalpy and entropy differences between U and N states. The values of ΔH and ΔS are dependent on temperature through a positive heat-capacity change, ΔCp. The positive ΔCp is thought to be related to the hydrophobic exposure of apolar side chains as a result of unfolding of globular proteins.

Two-Step versus Multistep Unfolding

The simplest model for the unfolding of a protein is the two-state model in which transition from N state to U state occurs in a cooperative process, that is, in a single unfolding transition step N ↔ U (4,9,10). A typical two-state unfolding of a protein follows a sigmoidal shaped unfolding curve, as shown in Figure 1 (for a review see Ref. 11). The N state is generally considered to have a narrow distribution of conformations, which are relatively ordered (low in entropy), whereas the U state involves a broader distribution of microscopic conformational states, which are relatively disordered (high in entropy). In recent years, protein biochemists have focused on the multistep unfolding model, in which protein unfolding proceeds from N to U by passing through one or more intermediate (I) states:

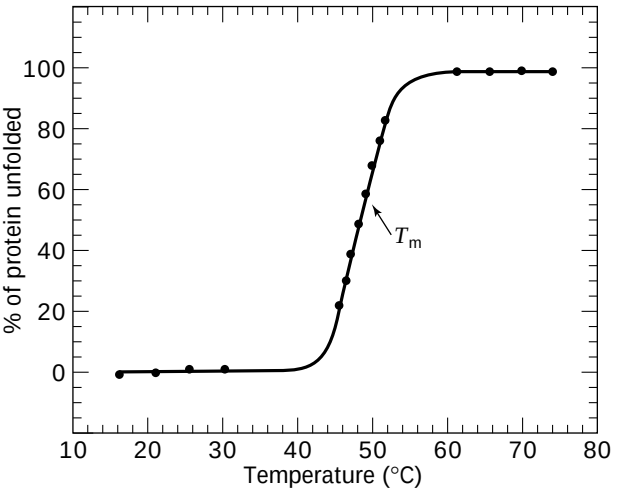


Figure 1. Two-state protein unfolding curve: Ribonuclease T1 thermal unfolding as measured by specific optical rotation at 295 nm. Data were adapted from reference 10 with permission.

$$N \leftrightarrow I \leftrightarrow \ldots \leftrightarrow U \quad (1)$$

If an I state is stable in the unfolding process, both N ↔ I and I ↔ U transitions can be characterized biophysically from equilibrium and kinetic unfolding studies (12–14). Several features of a partially unfolded state as a folding intermediate have been identified:

1. Relatively well-defined secondary structure (15,16) folded in a compact globular shape (17)
2. A non-cooperative folding/unfolding pathway (18)
3. Weak tertiary contacts and substantial solvent exposure of hydrophobic groups (18,19)
4. Relatively high affinity for hydrophobic ligands and tendency to aggregate (17,19)

For a multidomain protein, each individual domain may unfold sequentially or simultaneously (13).

Protein unfolding can lead to either reversible or irreversible loss of the folded structure. By definition, proteins denatured by a reversible pathway can regain their native structures and biological functions. An example of this process is the refolding of a protein after solubilization from inclusion bodies using denaturing conditions. Irreversible denaturation often leads to protein aggregation and results in loss of biological function (20,21).

APPLICATION OF THERMAL UNFOLDING STUDIES TO DEVELOPMENT OF STABLE PROTEIN FORMULATIONS

The unfolding of a protein during storage could be induced by perturbations such as elevated temperatures, surface

adsorption, and physical agitation. A successful formulation of a protein pharmaceutical would require the development of a dosage form in which the physicochemical properties and biological activity of the protein are maintained during storage. The design of a pharmaceutical formulation should be biocompatible, possess pharmaceutical elegance, and be suitable for commercial scale-up. The development of a stable formulation requires a thorough understanding of the causes and mechanism of inactivation, including both physical and chemical pathways. In approaching this goal, thermal unfolding studies are frequently used in the following ways. Intrinsic conformational stability data such as unfolding transition temperature can be obtained. If the transition is reversible, thermodynamic information such as ΔCp can be derived, and the results obtained from these accelerated stability measurements can often be extrapolated to make reasonable predictions about stability at different temperatures. Thermal unfolding studies can be used to investigate the mechanism of inactivation. The inactivation process is generally accelerated at an elevated temperature. Potential stabilizers can be screened by observing their effects on the thermal unfolding transition. An excipient that provides protection against temperature-induced denaturation or inactivation of a protein is often effective in prolonging its shelf life at refrigeration temperature (22). For example, an unstabilized respiratory syncytial virus (RSV) preparation lost most of its infectivity at −86 °C within 2 to 3 weeks, at 37 °C within 24 h, at 56 °C within 3 min. This study aided the development of a stabilized RSV formulation that was stable at −86 °C for 3 years, at 37 °C for 3 days, and at 56 °C for 6 min (23).

Compared to pH or denaturant-induced unfolding studies that also address the conformational stability of a protein, thermal unfolding studies are more direct and avoid the use of nonnative solution conditions (denaturants). In most cases, proteins and vaccines are likely to be exposed during handling to temperatures higher than the temperature usually recommended for their stable storage (e.g., 2 to 8 °C). Monitoring biological activity and structure of proteins as a function of increasing temperatures is the most convenient method to perform an accelerated stability study, although cold denaturation and freezing-induced aggregation of proteins have been studied as well (24–28). A major benefit of the accelerated stability studies is to expedite the process of screening suitable excipients for development of a stable formulation.

There have been many recent developments in using the approaches of stabilizing proteins by stabilizing the folded state, destabilizing the unfolded state, and altering the kinetics of unfolding. The factors that affect stability of a folded state versus the unfolded state as well as factors affecting rates of folding and unfolding play a role in maintaining stability (29). More detailed information on approaches to minimize inactivation and develop stable formulations of proteins are available (1,4,20,22,28–33).

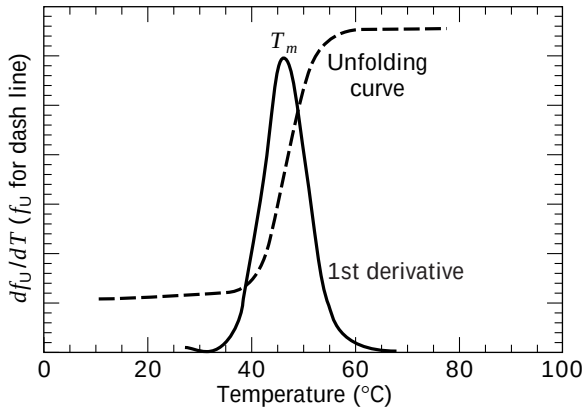
PROTEIN STABILITY EVALUATED BY MEASUREMENT OF THERMAL UNFOLDING TRANSITION

The denaturation transition temperature (T_m) has been widely used as a reference to evaluate the thermal stability of proteins and polynucleotides (34–43). T_m is defined as the temperature at the midpoint of protein unfolding transition (Figs. 1 and 2). At T_m, 50% of the total protein is unfolded (1,10). Therefore, for a reversible two-state unfolding process, at T_m, the unfolding equilibrium constant K is 1 and the Gibbs free energy ΔG is 0. Based on the complete unfolding curve, a two-state or multistate unfolding process can be suggested. For well-resolved, reversible, multiple transitions, the transition midpoint between states and the apparent unfolding ΔG can be determined for each transition under various conditions (37). The fractions of different conformational states at equilibrium can also, in such ideal cases, be determined from the unfolding transition curves (10,37,44).

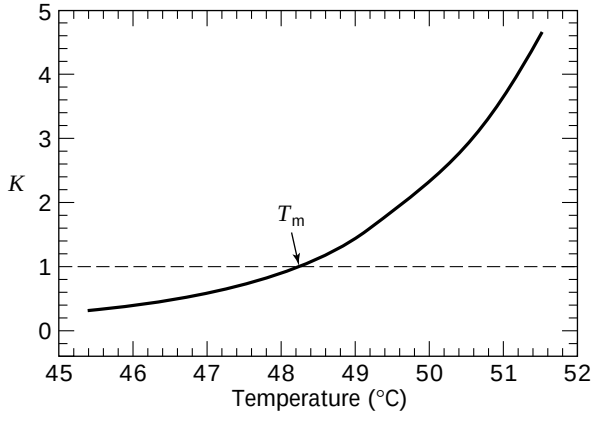
There are several methods to determine T_m of protein unfolding with reasonable accuracies. (1) The first derivative of the unfolding curve yields T_m as the sharp peak in the middle of the unfolding transition (Fig. 2**a**). (2) A plot of K values against temperature yields T_m at the temperature at which $K = 1$ (Fig. 2**b**). (3) A plot of ΔG value against temperature yields T_m at the temperature at which $\Delta G = 0$ (Fig. 2**c**).

An increase in T_m is indicative of structural stabilization of a protein in solution. The stabilization or destabilization effect of an excipient can be studied by measuring the change in T_m upon the addition of the excipient. For example, addition of heparin, a sulfated polysaccharide, to a solution of acidic fibroblast growth factor (aFGF) causes a dramatic increase in the T_m for aFGF unfolding, as measured by circular dichroism (CD) and fluorescence spectroscopy (31,45). A related study shows that the binding of heparin results in about 2.5 kcal/mol of stabilization for the native form of aFGF (46). Another example is the differential scanning calorimetric studies of the addition of osmolytes such as glycine, sarcosine, *N*,*N*-dimethylglycine, and betaine to ribonuclease A and lysozyme. The data indicate a 22 to 23 °C increase in T_m of the proteins upon the addition of the osmolytes (43). For ribonuclease, the 22 °C increase in T_m corresponds to an increase in stabilization free energy of 7.2 kcal/mol. Changes in T_m as a function of pH, salt, and osmolyte concentrations have been used to monitor thermal stability of proteins under different conditions (43,47,48). The T_m of tissue-type plasminogen activator (TPA) in phosphate buffer is approximately 66 °C. The addition of arginine (the stabilizer selected for the formulation) shifts the T_m up to 71 °C (28). Binding of nucleotides and substrates plays an important role in the thermal stabilization of a 70-kDa heat shock protein. Addition of nucleotide (ATP and/or ADP) or substrate protein (unfolded protein or peptide) to the solution increased the T_m of DnaK by 10 to 20 °C (38,49).

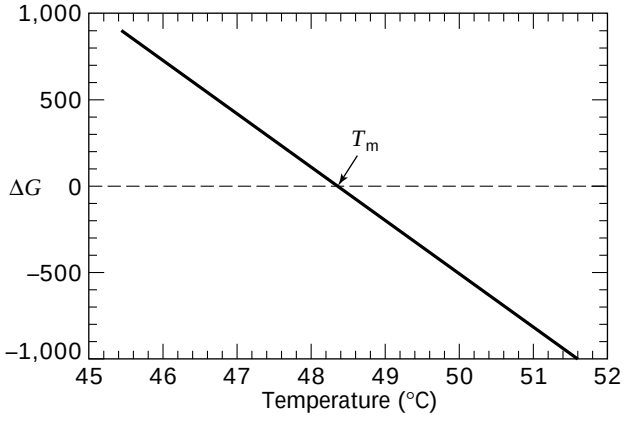
Protein unfolding can facilitate aggregation or chemical degradation (Scheme 1). Where U is fully or partially unfolded protein, I_d is the protein inactivated by degradation and I_a is the protein inactivated by aggregation. Unfolding could be a reversible process, but inactivation is normally an irreversible process. Study of the correlation of protein unfolding T_m with the rate of protein inactivation could provide useful information in understanding the mechanism of the inactivation process (7,33). In the case of reversibly denaturing of protein, the shift of T_m upon the addition of excipient is often related to the difference be-



(a)

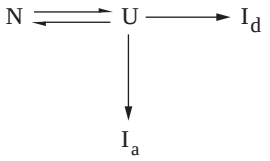


(b)



(c)

Figure 2. Methods for determining T_m value from protein unfolding curves (the plots were based on data shown in Figure 1) A) f_U (% protein unfolded) and df_U/dT versus T; B) K versus T; C) ΔG versus T. **(a)** First derivative of the fraction of unfolded protein as a function of temperature. **(b)** Equilibrium constant versus temperature. **(c)** Gibbs free energy versus temperature.



Scheme 1. Protein unfolding related inactivation pathways.

tween the binding affinities of the excipient for native and denatured states (50). For example, heparin increases the T_m of antithrombin III by 7 °C through a preferential binding to the native state of the protein (51).

Protein preparations are well known to be more stable in the dried state. Conformational flexibility is believed to be restricted in the dry state because interactions with water molecules are minimized. However, unfolding of proteins during the process of freeze-drying has been observed with FTIR spectroscopy. Many sugars and even some polymers, such as poly(vinylpyrrolidone) have been found to inhibit lyophilization-induced unfolding (24,52).

Many proteins irreversibly aggregate as they start to unfold with increasing temperature, making it difficult to obtain a complete thermal unfolding curve, and therefore the unfolding data can not be used to calculate thermodynamic parameters such as ΔG (31). As an alternative means for obtaining thermodynamic parameters, unfolding by urea or guanidine hydrochloride is frequently used, provided that these processes are reversible. These denaturants help to maintain the solubility of unfolded forms of these proteins. For the same reason, when performing the thermal unfolding studies of such proteins, a relatively low concentration (0.2 to 1 M) of chemical denaturants such as urea or guanidine hydrochloride could be added to the protein solution to prevent fully or partially unfolded protein from falling out of the solution because of aggregation (53). When this is the case, the contribution of the added denaturant to the apparent unfolding T_m should be taken into consideration. Normally, unfolding should be carried out at a series of denaturant concentrations, and T_m and other thermodynamic parameters can be obtained by extrapolation to zero denaturant concentration. A two-dimensional protein unfolding (thermal unfolding profile under various guanidine hydrochloride concentration conditions) study aimed at elucidation of biomolecular energetics has been carried out by Straume (54).

PROTEIN STABILITY EVALUATED BY KINETICS OF TEMPERATURE-INDUCED PROTEIN INACTIVATION OR UNFOLDING

Another approach to quantitative characterization of the stability of a protein is to study the protein unfolding kinetics and inactivation rate (29,55,56). Rapid unfolding, aggregation, or inactivation at high temperature indicates a less-stable protein product. A formulation that can slow down these denaturation and inactivation processes at el-

evated temperatures may be able to increase the stability of the protein product at storage temperature.

Arrhenius Relationship

For most proteins, unfolding can be simplified as a first-order kinetic process displaying single-exponential unfolding kinetics. The equilibrium unfolding curve can be predicted from the rate constant assuming a two-state kinetic model. The unfolding rate can be expressed by the decrease in population of the native structure with time as

$$-d[\mathrm{N}]]/dt = k[\mathrm{N}] \quad (2)$$

where [N] is the population of the structure remaining native at time t during unfolding, and the proportionality constant, k, is the apparent rate constant of the unfolding. The population of the native structure remaining can be expressed as

$$[\mathrm{N}] = [\mathrm{N}]_0 \exp(-kt) \quad (3)$$

where $[\mathrm{N}]_0$ is the population of native structure at time zero. Arrhenius in 1889 suggested that a reasonable equation for relating the rate constant, k, with an activation energy term (Ea) for the forward part of a reversible reaction is

$$d\ln k/dT = Ea/RT^2 \quad (4)$$

where R is the gas constant, and T is temperature. At temperature T, the above equation can be expressed as

$$k = A\exp(-Ea/RT) \quad (5)$$

where A is a frequency factor or preexperiential factor that can be determined by performing a multimeasurement of k at various temperatures (assuming A and Ea are independent of temperature). Equation 5 shows that for a reaction with a positive activation energy, the reaction rate increases with temperature.

The goal of a stable formulation is to minimize unfolding and other degradative pathways. According to equation 5, this can be achieved by either decreasing storage temperature or increasing the activation energy (Ea) for unfolding. The apparent activation energy, Ea, of unfolding determines the rate of the transition between the folded and unfolded states. This can be affected by altering solution conditions. Stabilizing excipients will increase the activation energy of unfolding and will, therefore, reduce the rate of unfolding under the formulation condition. The calculation of Ea requires measurement of a structural or functional property of the protein that directly reflects unfolding at several different temperatures. The Arrhenius plot of ln k versus $1/T$ should be linear (the slope of which yields the activation energy), and this may be true only over a limited temperature range. For example, the amount of aggregated protein might increase beyond a certain temperature to the point of causing interference with rate measurements. A biphasic Arrhenius plot in the absence of aggregation would indicate either a change in the pathway by which temperature affects the protein structure or the breakpoint that occurs at the temperature at which pathway changes. Measurement of unfolding activation energy has been applied to protein stability studies (22,28).

Experimental Approaches

Careful considerations must be given to the selection of methods for measuring the unfolding rate. Far-UV CD detects changes in secondary structure, but subtle and functionally important changes might remain undetected. Near-UV CD and tryptophan fluorescence measurements are well suited to monitor tertiary structural changes. The active sites of proteins, including receptor-binding sites, catalytic sites, and antigen-binding sites, would often be perturbed (locally) before a global unfolding occurs. In such cases, the unfolding rate will be slower than inactivation rate. A good strategy is to use a number of different techniques, for example, CD and fluorescence spectroscopy to monitor the unfolding rate and hydrodynamic methods to monitor rate of aggregation.

APPLICATION OF ACCELERATED PROTEIN STABILITY STUDIES TO FORMULATION DEVELOPMENT

Although the expiration dating of a biological product is based on real-time and real-temperature storage stability data, accelerated stability studies of biological products under stress conditions usually constitute an important part of the stability data package (57). It is common to determine the degradation rate of a chemical compound under a stress condition, for example, an elevated temperature (37 °C or higher), and then extrapolate the results to a lower (storage) temperature (22). When exposed to sufficiently elevated temperatures, all proteins will lose biological activity within limited a time.

Thermal stability studies include two main approaches: temperature-induced conformational change and kinetics of unfolding and degradation at elevated temperatures. The first approach is to follow the unfolding transition temperature to test, for example, effects of various excipients on T_m. The second approach is to accelerate the process that may result in the destabilization of the protein samples. At elevated temperatures, protein unfolding rate increases and most chemical interactions become faster compared to lower temperatures (37). Under these conditions, various formulations could be screened for their ability to inhibit the inactivation process. Under accelerated stability study conditions, the intrinsic stability of protein pharmaceuticals is often decreased dramatically. The addition of various excipients or stabilizers may enhance the stability of the protein pharmaceuticals under the same condition. For example, a properly buffered solution of acidic FGF (aFGF) complexed with heparin is stable at 5 °C for 1 year, but the protein is inactivated when stored at 30 °C for 2 to 4 months. Addition of 0.15 mM EDTA to the formulation enhanced storage stability at 30 °C to be roughly equivalent to that at 5 °C, and more than 70% of the protein mass remained after 1 year of storage at 30 °C (58).

Questions have been raised about the validity of accelerated protein stability studies in defining stable formulations. There are two simple but critical questions that deserve attention. One is the question of reliability of extrapolating results obtained at elevated temperatures to a lower storage temperature, such as 5 °C (22). Another question is how high the temperature for an accelerated condition could be pushed to. So far, there is no generally correct answer to the first question. The answer to the second question is that for quantitative extrapolation of kinetic data to low temperatures, the accelerated studies should be kept well below the lowest protein unfolding transition temperature (T_m). Pearlman and Nyguyen (59) were able to successfully extrapolate data obtained in an accelerated stability study at 40 °C down to 5 °C.

If protein unfolding is a first-order kinetic process, the unfolding activation energy could be obtained by carrying out the unfolding measurement at two or more different temperatures, and then using the temperature coefficient of kinetics to calculate the unfolding rate of the protein at storage temperature. Recently, Eberlein et al. (60) used a turbidity test for 15 h at 30 to 40 °C by UV spectrophotometry to predict storage stability of recombinant human bovine fibroblast growth factor (rhbFGF) with 95% confidence. Figure 3 shows a theoretical calculation for different activation energies of a first-order inactivation process using the Arrhenius equation. The results indicate that the storage stability increases as a function of activation energy of the unfolding and inactivation process. The higher the activation energy, the longer the shelf life is expected to be.

TECHNIQUES USED IN THERMAL UNFOLDING STUDIES

The spectral property of a protein molecule is sensitive to the molecular environment and the mobility of its chromophores. Far-UV CD (180–250 nm) and amide I Fourier transform infrared (FTIR) (1600–1700 cm^{-1}) are convenient spectroscopic tools for monitoring changes in the secondary structures of proteins (61,62). Fluorescence and near-UV CD (250–320 nm) are used to monitor environmental changes around aromatic chromophores of proteins and can often detect subtle changes in tertiary structure even in the absence of a global structural change in a protein (1,44,63,64). Differential scanning calorimetry (DSC), which measures changes in heat capacity, provides a direct method for monitoring thermal stability of proteins (43,65–67). Changes in hydrodynamic properties of proteins, such as aggregation, often induced at elevated temperatures, are monitored by techniques such as dynamic light scattering (DLS), small-angle X-ray scattering, analytical ultracentrifugation, and size exclusion chromatography (SEC). Specific applications of some of the more commonly used techniques that readily provide useful information are described in this section.

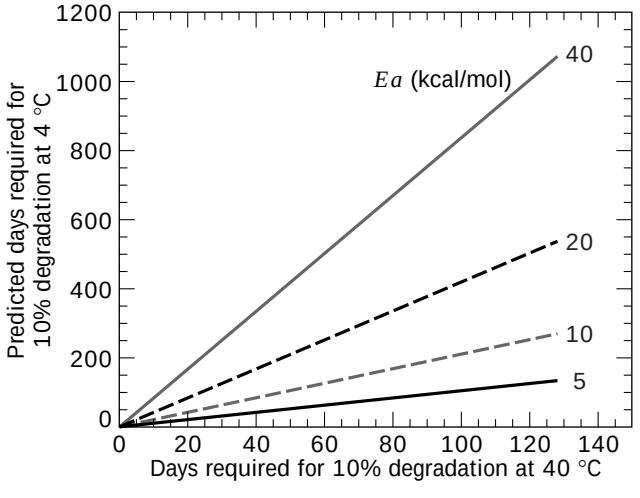


Figure 3. Theoretical prediction of 4 °C shelf-life time based on stability data at 40 °C.

Spectroscopic Methods

Tryptophan and tyrosine residues are the most commonly used intrinsic aromatic chromophores and fluorescence probes of proteins. Phenylalanine fluorescence exhibits a low quantum yield and has not found much use in protein unfolding studies. Tryptophan fluorescence spectral maxima of proteins typically range from 315 to 345 nm, and a fully water-exposed tryptophan manifests maximal fluorescence around 355 nm. The intrinsic fluorescence of tryptophan residues of most proteins is very sensitive to changes in the polarity of their immediate environments. Protein unfolding is usually accompanied by a change in the fluorescence (emission) spectrum of tryptophan, caused by increased exposure of the tryptophan residue(s) to solvent. The spectral maxima of tyrosine fluorescence are limited to a relatively narrow range (~300 to 310 nm), and they often do not manifest significant shifts upon unfolding. However, changes fluorescence intensity can provide valuable information. In proteins containing tryptophan and tyrosine residues, energy transfer from tyrosines to tryptophans can occur, and this process may be affected because of unfolding-induced changes in the distances between these residues. For most proteins containing tyrosine and tryptophan residues, excitation at or near 280 nm (close to the absorption maxima for both residues) results in tryptophan fluorescence. Fluorescence from the phenolic side chain of tyrosines in proteins is often quenched because of different mechanisms, including energy transfer to tryptophan. However, in a number of proteins, tyrosines contribute significantly to the fluorescence spectrum in the 300 to 350 nm wavelength range. To selectively observe tryptophan fluorescence in such proteins, excitation at the red edge of the tryptophan absorption spectrum (typically 295 to 300 nm) is used (68). Although this reduces the sensitivity of excitation (and consequently of emission), the differential absorption between tryptophan and tyrosine is high at these wavelengths where tyrosine absorption is minimal. For aFGF in its properly folded form, the fluorescence intensity of its single tryptophan residue is dramatically quenched, presumably because of electrostatic interactions with charged residues, and can only be observed, with a very low quantum yield, upon red edge excitation (69). When excited at 280 nm, folded aFGF man-

ifests tyrosine fluorescence and no detectable tryptophan fluorescence. As the temperature is raised and the protein starts to unfold, tryptophan fluorescence begins to emerge with a maximum wavelength characteristic of solvent-exposed tryptophan. Changes in tyrosine and tryptophan fluorescence intensities were simultaneously monitored as a function of temperature to study thermal unfolding of aFGF (55,69). These measurements have helped identify a number of polyanionic stabilizers of aFGF.

The relatively simple measurements of steady-state intrinsic fluorescence spectra of proteins, as described above, can provide a wealth of information about protein unfolding and aid in developing stabilizing formulations of proteins. Extrinsic probes such as anilinonaphthalene sulfonic acid (ANS) and (p-toluidiny)naphthalene sulfonic acid (TNS) that noncovalently bind to hydrophobic surfaces have also been widely used in protein unfolding studies, especially in the context of studying partially unfolded intermediates (70). These and other applications of fluorescence spectroscopy to unfolding and thermal stability studies, such as nonradiative energy transfer measurements, have been recently reviewed (32). It should be noted that light-scattering artifacts make it difficult for fluorescence spectroscopy to be used for protein formulations that are not clear solutions or that contain large aggregates or particles. Front face fluorometry or alternative detection geometries, coupled with careful light scattering corrections, can still yield valuable information in such cases (31).

Near-UV CD signals in the 250 to 300 nm range result from asymmetric environment experienced by aromatic amino acid side chains and in the tertiary structures of folded proteins. Unfolding creates increased mobility and decrease or loss of induced asymmetry. This technique is, therefore, complementary to fluorescence in monitoring subtle changes in tertiary structure of proteins. Phenylalanine, tyrosine, and tryptophan are usually well resolved. For guanidine-hydrochloride-induced unfolding of a chimeric protein, TP40, with potentially chemotherapeutic activity, near-UV CD, and fluorescence of tryptophan residues produced comparable unfolding transitions (63). The sensitivity of this technique, in terms of sample concentration requirement, is lower than that of fluorescence. Furthermore, the extraction of quantitative information from near-UV CD spectra is often cumbersome. Theoretically, a quantitative description of a protein's tertiary structure based on its near-UV CD spectrum is attainable, but it is rarely done because detailed information of transition dipole moment orientations, interactions, and environments are not available.

Far-UV CD spectroscopy is frequently used as a powerful technique to monitor changes in the protein's secondary structure that might arise from thermal and other kinds of perturbation. Each element of the secondary structure of a protein's peptide backbone, that is, α-helices, β-sheets, turns, and random coils, manifests different but overlapping spectral features. Protein unfolding results in a decrease in the amount of ordered secondary structure, for example, α-helices and β-sheets, which, for thermal unfolding, can be conveniently monitored by measuring the CD signal at one or more wavelengths as a function of temperature. For an accurate description of the thermal unfolding transition, careful baseline correction may be necessary. This has been achieved by simultaneously monitoring, as a function of temperature, the background signal at a wavelength where proteins do not manifest CD (71,72).

Amide I FTIR spectroscopy is complementary to CD in monitoring the secondary structure of the peptide backbone. The biggest advantage of FTIR, especially in the context of protein formulation, is that the protein sample does not have to be in a clear solution state. Samples in different physical states can be studied, including solids, gels, suspensions, and solutions. Different sample geometries may be used, including transmission, attenuated total reflection, and diffuse reflectance. The choice of the experimental design often depends on the physical state of the sample being used (62). For example, lyophilized samples can be directly studied in the dry form using the diffuse reflectance technique. Although KBr pellets of dry protein samples have often been used to make transmission measurements, it should be recognized that KBr is a chaotropic agent and may cause perturbation of the protein structure, although chances for protein unfolding are reduced in the dry state. The effect of freeze-drying on structures of proteins has been studied by FTIR spectroscopy and potential stabilizers against freezing-induced structural damage have been identified (52). Infrared spectroscopy has also been used to evaluate the effectiveness of formulations in protecting the secondary structural integrity of proteins both in solution and in dried states. This necessitates making quantitative comparisons of the overall similarity of infrared spectra in the conformational sensitive amide I region. The study of temperature-induced protein unfolding using FTIR is generally performed by following spectral change with temperature at a fixed absorption band (e.g., amide I or I′) or wavelength (e.g., 1,640 cm^{-1}). The plot of FTIR intensity at a fixed wavelength versus temperature should yield the protein unfolding transition temperature (73). Detailed discussions on the application of FTIR to study protein conformational stability and thermal unfolding are available elsewhere (62,74–81). Protein aggregation as a result of thermal unfolding can also be detected by appearance of bands around 1,610 and 1,685 cm^{-1}, as demonstrated for chymotrypsinogen (82). The major disadvantages of FTIR are (*1*) In solution, water signal interferes with the amide I signal and must be carefully subtracted or moved away from the spectral range of interest by dialyzing the sample into D_2O. (*2*) Deconvolution of the rather broad amide I band can, to some extent, be subjective. (*3*) There is significant spectral overlap between peptide bonds in random coil and α-helical conformation. (*4*) The sensitivity for aqueous protein samples is approximately two orders of magnitude lower compared to far-UV CD. As for all the biophysical techniques described here, FTIR should be used for monitoring changes in conformation as a result of a perturbation, such as temperature, and should not be viewed as a technique to determine structure contents in absolute terms. To this end, second-derivative FTIR spectroscopy has been shown to be a valuable comparative tool, and second derivatives of amide I spectra have been particularly useful in monitoring structural changes.

It should be noted that temperature-induced protein unfolding can often be monitored by simply following UV absorbance changes in the far-UV (peptide bond absorption) and near-UV (aromatic side chains) regions. This was demonstrated for ribonuclease A and equinatoxin II by Fink and Painter (83) and Poklar et al. (84). However, this is not as commonly used a method as fluorescence and CD because (*1*) the magnitude of UV absorbance change between folded and unfolded protein states is usually much smaller than the change observed by fluorescence and CD and (*2*) the potential light scattering contribution resulting from protein aggregation that often accompanies unfolding can introduce significant inaccuracy. The resolution of near-UV absorbance spectroscopy, which is inherently a low-resolution technique, can be significantly improved by using second-derivative analysis (85). In addition, UV spectrophotometry offers a simple method to monitor temperature-induced protein aggregation. This method, which is usually called the turbidity assay and frequently used in accelerated thermal stability studies, takes advantage of the increase in total optical density resulting from light scattering as large protein aggregates are formed. The intensity of light scattering is proportional to the inverse fourth power of the wavelength of light so that greater sensitivity in measuring light scattering is achieved at lower wavelengths. Typically, a satisfactory balance between measurable light-scattering intensity and minimal protein absorbance is accomplished in the 300 to 350 nm wavelength range (4). The turbidity assay was one of the methods used to study thermal stabilization of basic and acidic FGF (45,56).

The need to use multiple techniques to study thermal unfolding of proteins cannot be overemphasized. When different techniques yield different T_m values, the pathway of unfolding most likely involves one or more partially folded intermediates. For aFGF in the presence of the stabilizer, heparin, fluorescence, CD, and DSC measurements yielded apparent T_m values of 62 °C, 72 °C and 71 °C, respectively (31). This suggests that local unfolding in tertiary structures precede a more global unfolding in which the ordered secondary structure of the peptide backbone is disrupted. For another potentially therapeutic protein, TP4O, produced as chimera of a truncated form of *Pseudomonas exotoxin* and transforming growth factor-α, tryptophan fluorescence, far-UV CD, and DSC gave apparent T_m values of 42, 55, and 48 °C. In this case also, a local unfolding in one or more parts of the tertiary structure precedes global unfolding. The in vitro cell killing activity of the protein was found to be irreversibly lost at 50 °C, and the hydrodynamic radius was increased to 12 nm, compared to 4 nm at 23 °C as measured by dynamic light scattering, suggesting aggregation (71,86).

Calorimetry

Differential scanning calorimetry (DSC) is another frequently used method in protein structural stability studies. Changes in heat capacity as a function of temperature, obtained from DSC measurements, provide a direct description of the thermal unfolding transitions of proteins. A good example (Fig. 4) is the study of the effects of various

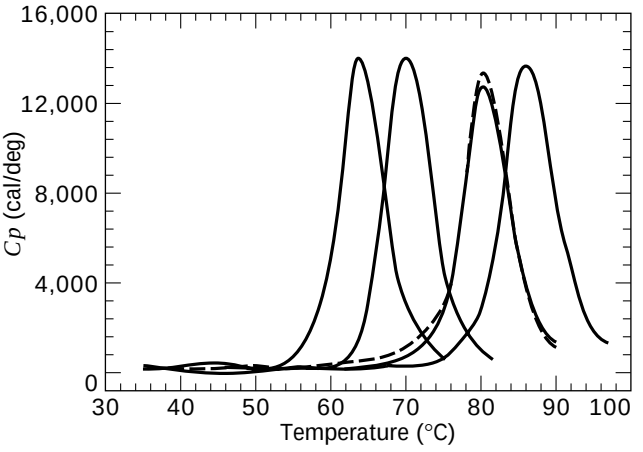


Figure 4. DSC determination of sarcosine effect on the thermal unfolding transition (T_m) of ribonuclease A. Endotherms from lower to higher T_m represent sarcosine concentrations of 0.0, 2.0, 5.5, and 8.2 M. Both of first scan and the rescan are shown for sample with 5.5 M sarcosine (solide and dashline). Data were adapted from reference 47 with permission.

excipients on thermal unfolding transitions of ribonuclease A (43,47,48). The rescanned heat flow profile indicates that thermal unfolding of ribonuclease A is a reversible process (43). In theory, calculation of thermodynamic parameters such as Gibbs free energy can only be made for reversible transitions (66,87). Thermal unfolding of proteins is often irreversible because of aggregation of the unfolded or partially unfolded states. Nevertheless, in protein formulation studies, potential stabilizers can be screened by monitoring their effects on the heat flow profile of a protein scanned from low to high temperature. For example, DSC studies provided critical data in support of stabilizing effects of heparin and osmolytes against thermal unfolding and (irreversible) aggregation of aFGF and lysozyme (43,45,48).

Hydrodynamic Measurements

Unfolding of a protein is usually associated with an apparent increase in its hydrodynamic size and changes in sedimentation and diffusion coefficients. In addition, if the partial or fully unfolded protein states are unstable and form aggregates, the overall hydrodynamic properties of the protein in the solution will change. The techniques for monitoring changes in hydrodynamic properties of proteins include capillary electrophoresis (CE), SEC-HPLC, analytical ultracentrifugation, dynamic light scattering (DLS), and small-angle X-ray scattering (SAXS).

With the recent introduction of a compact and easy-to-use analytical ultracentrifuge that offers both optical absorption and interferometric optics for real-time monitoring of sedimentation, interest has grown exponentially in using equilibrium and velocity sedimentation techniques to measure self-association of proteins (88–90). Equilibrium sedimentation offers a thermodynamically rigorous way of determining molecular weights. The sedimentation

properties (sedimentation coefficient) of a protein sample are controlled by several physical factors of the protein and protein complex with other solution components. These factors include protein mass, geometry, and density and are affected by protein oligomerization, aggregation, unfolding, and degradation. Both velocity and equilibrium sedimentation experiments can provide information about protein hydrodynamic properties (33,71). Velocity sedimentation experiments can be performed in only a few hours, although diffusion coefficient is required for obtaining a semiquantitative measure of mass change caused by aggregation.

Based on the difference between the electrophoretic mobility of the native protein and that of the unfolded one, the temperature-induced unfolding transitions of proteins can be analyzed by CE. This technique offers the advantage of very high sensitivity, although it has not yet found wide usage in protein unfolding studies. Ishihama et al. (91) used an in-column incubation method to monitor thermal unfolding of nanogram quantities protein. The thermodynamic parameters of the unfolding transition process were obtained by analyzing the temperature dependence of the electrophoretic mobilities. More information on the application of CE in protein analysis can been found in literature (91–93).

SEC allows separation of proteins based on their molecular shape and size. An unfolded or aggregated protein generally possesses a larger hydrodynamic size than the native protein and is eluted out of the column earlier (i.e., with a smaller elution volume) than the native protein. This method has been used to estimate molecular sizes of proteins under different denaturing conditions (13,63,94). The application of various signal detection techniques in protein gel–filtration analysis has been discussed (32).

Static light scattering (SLS), DLS, and SAXS from protein solutions can all be applied to the measurement of changes in the molecular sizes of proteins induced, for example, by thermal unfolding. SLS measurements normally permit reliable data of molecular weights and radii of gyration for relatively small particles since light scattering of large particles is a very complex function of particle shape and sensitive to detection angle. DLS measures the time-dependent intensity fluctuations of scattered light. The correlation of these fluctuations is related to the diffusion coefficient of the scattering particle, a quantity linked to the hydrodynamic radius through the Stokes-Einstein equation (95). SAXS is a very useful, but underused, technique for characterizing size, compactness, and shape of protein molecules in solution. In addition to determining the overall radius of gyration and the maximum dimension, SAXS can quantitatively describe the globular conformation of a protein molecule, as well as denaturant- and ligand-induced conformational changes in solution (96–98). Detailed discussions on the applications of SLS, DLS, and SAXS on protein conformational change and unfolding are available in the literature (27,32,99–102).

Kinetic Techniques

For a complete description of the thermal unfolding kinetics, a fast time resolution may be necessary. Stopped-flow instrumentation with absorption and fluorescence optics allows studies of unfolding kinetics with a dead time of about 2 ms. Stopped-flow CD instrumentation is also available. The instrumentation of time-resolved CD (TRCD) has been reviewed by Lewis (103). Time-resolved FTIR spectroscopy (104) can complement kinetic CD studies in identifying the faster phases of folding or unfolding of a protein's secondary structure. Techniques for studying fast protein structural change at an atomic level include hydrogen–deuterium exchange in combination with multidimensional NMR, electron spray mass spectrometer, and FTIR spectroscopy (105,106). These techniques monitor the spectral and mass changes induced by exchange of amide protons with deuterium during the folding or unfolding process.

CONCLUSION

Thermal unfolding studies provide formulation scientists a way to evaluate the intrinsic stability of pharmaceutical proteins and to screen excipients that might lead to stabilizing formulations during storage, handling, and administration. The data obtained from thermal unfolding studies can support the design and execution of the formulation processing steps leading to pharmaceutical products. Because temperature is one of the most practically important parameters in testing stability of a protein from small- and large-scale production, detailed information of a protein's conformational stability at various temperatures is critically needed. A rational approach to stable formulation development requires a thorough understanding of the mechanism of protein inactivation. Normally, thermal unfolding experiments provide an understanding of both the physical and chemical events leading to protein inactivation, thereby opening up a faster path to development of stabilizing formulations for protein-based pharmaceuticals and vaccines.

BIBLIOGRAPHY

1. M.C. Manning, K. Patel, R.T. Borchardt, *Pharmacol. Res.* **6**, 903–918 (1989).
2. R.D. Schmid, *Adv. Biochem. Eng.* **12**, 41–118 (1979).
3. C.R. Cantor and P.R. Schimmel, *Biophysical Chemistry, Part I: The Conformation of Biological Macromolecules*, Freeman, New York, 1980, pp. 7–12.
4. R.A. Copeland, *Methods for Protein Analysis: A Practical Guide to Laboratory Protocols*, Chapman & Hall, New York, 1994.
5. T.E. Creighton, *Protein: Structure and Molecular Properties*, Freeman, New York, 1984, pp. 159 and 220.
6. L. Stryer, *Biochemistry*, 3rd ed., Freeman, New York, p. 31, 1988.
7. D.B. Volkin and C.R. Middaugh, in T.J. Ahern and M.C. Manning eds., *Stability of Protein Pharmaceuticals, part A, Chemical and Physical Pathways of Protein Degradation*, Plenum, New York, 1992, pp. 109–134 and 215–247.
8. D. Shortle, *J. Biol. Chem.* **264**, 5315–5318 (1989).
9. C.N. Pace, *CRC Crit. Rev. Phys. Chem.* **38**, 463–488 (1975).
10. B.A. Shirley, in T.J. Ahern, and M.C. Manning eds., *Stability of Protein Pharmaceuticals, part A, Chemical and Physical Pathways of Protein Degradation*, Plenum, New York, 1992, pp. 167–194.

11. C.G. Anfinsen and H.A. Scheraga, *Adv. Protein Chem.* **29**, 205–301 (1975).
12. P.A. Jennings and P.E. Wright, *Science*, **262**, 892–896 (1993).
13. R.D. Palleros, L. Shi, K.L. Reid, and A.L. Fink, *Biochemistry* **32**, 4314–4321 (1993).
14. F.M. Hughson, P.E. Wright, R.L. Baldwin, *Science* **249**, 1544–1548 (1991).
15. D. Atkinson and D.M. Small, *Annu. Rev. Biophys. Chem.* **15**, 403–456 (1986).
16. R.T. Nolte and D. Atkinson, *Biophys. J.* **63**, 1221–1239 (1993).
17. D.L. Barbeau, A. Jonas, T. Teng and A.M. Scanu, *Biochemistry* **18**, 362–369 (1979).
18. J.A. Reynolds, *J. Biol. Chem.* **251**, 6013–6015 (1976).
19. J.M. Donovan, G.B. Benedek, and M.C. Carey, *Biochemistry* **26**, 8116–8125 (1987).
20. J.L. Cleland, M.F. Powell, and S.J. Shire, *Crit. Rev. Ther. Drug Carrier Systems* **10**, 307–377 (1993).
21. L.R. De Young, A.L. Fink, and K.A. Dill, *Acc. Chem. Res.* **26**, 614–620 (1993).
22. Y.J. Wang and M.A. Hanson, *J. Parenter. Sci. Technol.* **42**, S3–S26 (1988).
23. C.K. Gupta, J. Leszcynski, R.K. Gupta, and G.R. Siber, *Vaccine* **14**, 1417–1420 (1996).
24. J.F. Capenter, S.J. Prestrelski, T.J. Anchordoguy, and T. Arakawa, in J.L. Cleland and R. Langer eds., *Formulation and Delivery of Proteins and Peptides*, American Chemical Society, 1994, pp. 137–147.
25. A.T. Da Poian, A.C. Oliveira, and J.L. Silva, *Biochemistry* **34**, 2672–2677 (1995).
26. B.M. Eckhardt, J.Q. Qeswein, and T.A. Bewley, *Pharm. Res.* **8**, 1360–1364 (1991).
27. I. Nishii, M. Kataoka, F. Tokunaga, and Y. Goto, *Biochemistry* **33**, 4903–4909 (1994).
28. R. Pearlman and T.J. Nguyen, *Pharm. Pharmacol.* **44**, 178–185 (1992).
29. A. Shaw and R. Bott, *Curr. Opin. Struct. Biol.* **6**, 546–550 (1996).
30. S.N. Timasheff, in B.A. Shirley ed., *Protein Stability and Folding: Theory and Practice*, Humana, Totowa, N.J., 1995, pp. 219–270.
31. D.B. Volkin and C.R. Middaugh, in R. Pearlman and Y.J. Wang ed., *Formulation, Characterization, and Stability of Protein Drugs*, Plenum, New York, 1996, pp. 181–217.
32. D.B. Volkin, G. Sanyal, C.J. Burke, and C.R. Middaugh, in M. Akers and S. Nail ed., *Development and Manufacture of Protein Pharmaceuticals*, Plenum, in press.
33. D.B. Volkin and A.M. Klibanov, in T.E. Creighton ed., *Protein Function: A Practical Approach*, Oxford Press, Oxford, U.K., 1989, pp. 1–24.
34. D.R. Bachvarov, G.G. Markov, and I.G. Ivanov, *Int. J. Biochem.* **19**, 963–971 (1935).
35. P. Gunning, T. Mohun, S.Y. Ng, P. Ponte, and L. Kedes, *J. Mol. Evol.* **20**, 202–214 (1984).
36. E. Odermatt, J. Risteli, V. van-Delden, and R. Timpl, *Biochem. J.* **211**, 295–302 (1983).
37. C.N. Pace, B.A. Shirley, J.A. Thomson, in T.E. Creighton ed., *Protein Structure: A Practical Approach* IRL, Oxford, U.K., 1989, p. 311.
38. R.D. Palleros, L. Shi, K.L. Reid, and A.L. Fink, *J. Biol. Chem.* **269**, 13107–13114 (1994).
39. T. Porumb, P. Yau, T.S. Harvey, and M. Ikura, *Protein. Eng.* **7**, 109–115 (1994).
40. L.R. Scolnick, Z.F. Kanyo, R.C. Cavalli, D.E. Ash, and D.W. Christianson, *Biochemistry* **36**, 10558–10565 (1997).
41. O. Slodowski, J. Bohm, B. Schone, and B. Otto, *Eur. J. Biochem.* **202**, 1133–1140 (1991).
42. H. Takagi, T. Takahashi, H. Momose, M. Inouye, Y. Maeda, H. Matsuzawa, and T. Ohta, *J. Biol. Chem.* **265**, 6874–6888 (1990).
43. M.M. Santoro, Y. Liu, S.M.A. Khan, L. Hou, and D.W. Bolen, *Biochemistry* **31**, 5278–5283 (1992).
44. M. Eftink, *Biophys. J.* **66**, 482–501 (1994).
45. P.K. Tsai, D.B. Volkin, J.M. Dabora, K.C. Thompson, M.W. Bruner, J.O. Gress, B. Matuszewska, M. Keogan, J.V. Bondi, and C.R. Middaugh, *Pharm. Res.* **10**, 649–659 (1993).
46. C.J. Burke, D.B. Volkin, H. Mach, and C.R. Middaugh, *Biochemistry* **32**, 6419–6426 (1993).
47. C.Q. Hu, J.M. Sturtevant, J.A. Thamson, R.E. Erickson, and C.N. Pace, *Biochemistry* **31**, 4876–4882 (1992).
48. Y. Liu and J.M. Sturtevant, *Biochemistry* **35**, 3059–3062 (1996).
49. L. Shi, D.R. Palleros, and A.L. Fink, *Biochemistry* **33**, 7536–7546 (1994).
50. J.A. Schellman, *Biopolymers* **15**, 999–1000 (1976).
51. T.F. Busby, D.H. Atha, and K.C. Ingham, *J. Biol. Chem.* **256**, 12140–12147 (1981).
52. S.J. Prestrelski, N.T. Tedeschi, T. Arakawa, and J.F. Carpenter, *Biophys. J.* **65**, 661–671 (1993).
53. C.N. Pace, *CRC Crit. Rev. Biochem.* **3**, 1–43 (1975).
54. M. Straume, *Methods Enzymol.* **240**, 530–568 (1994).
55. M.K. Eidsness, K.A. Richie, A.E. Burden, D.M. Kurtz Jr., R.A. Scott, *Biochemistry*, **36**, 10406–10413 (1997).
56. Y.J. Wang, Z. Shahrokh, S. Vemuri, G. Eberlein, I. Beylin, M. Busch, *Pharm-Biotechnol.* **9**, 141–180 (1996).
57. Food and Drug Administration (FDA), *Fed. Regist.* **61**, 36466–36469 (1996).
58. D.B. Volkin, A.M. Verticelli, K.E. Marfia, C.J. Burke, H. Mach, and C.R. Middaugh, *Biochim. Biophys. Acta* **1203**, 18–26 (1993).
59. R. Pearlman and T.J. Nguyen, in D. Marshak and D. Liu eds., *Therapeutic Peptides and Proteins: Formulation, Delivery and Targeting*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989, pp. 23–31.
60. G.A. Eberlein, P.R. Stratton, and Y.J. Wang, *PDA J. Pharm. Sci. Technol.* **48**, 224–230 (1994).
61. M. Bloemendal and W.C. Johnson, Jr., in J.N. Herron, W. Jiskoot, and D.J.A. Crommelin ed., *Physical Methods to Characterize Pharmaceutical Proteins*, Plenum, New York, 1995.
62. C.R. Middaugh, H. Mach, J.A. Ryan, G. Sanyal, and D.B. Volkin, in B.A. Shirley ed., *Protein Stability and Folding: Theory and Practice*, Humana, Totowa, N.J., 1995, pp. 137–156.
63. J.O. Gress, D. Marquis-Omer, C.R. Middaugh, and G. Sanayl, *Biochemistry* **33**, 2620–2627 (1994).
64. W. Jiskoot, V. Hlady, J.J. Naleway, and J.N. Herron in J.N. Herron, W. Jiskoot, and D.J.A. Crommelin, eds., *Physical Methods to Characterize Pharmaceutical Proteins*, Plenum, New York, 1995, pp. 1–53.
65. J.I. Boye, I. Alli, and A.A. Ismail, *J. Agric. Food. Chem.* **45**, 1116–1125 (1997).

66. K. Gekko, *J. Biochem.* **91**, 1197–1204 (1982).
67. P.L. Privalov and N.N. Khechinashvili, *J. Mol. Biol.* **86**, 665–684 (1974).
68. J.R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Plenum, New York, 1983.
69. R.A. Copeland, H. Ji, A.J. Halfpenny, R.W. Williams, K.C. Thompson, W.K. Herber, K.A. Thomas, M.W. Bruner, J.A. Ryan, D. Marquis-Omer, G. Sanyal, R.D. Sitrin, S. Yamazaki, and C.R. Middaugh, *Arch. Biochem. Biophys.* **289**, 53–61 (1991).
70. K. Kuwajima, *Proteins: Struct., Funct., Genet.* **6**, 87–103 (1989).
71. G. Sanyal, D. Marquis-Omer, J.O. Gress, and C.R. Middaugh, *Biochemistry* **32**, 3488–3497 (1993).
72. G. Sanyal, D. Marquis-Omer, J.O.L. Waxman, H. Mach, J.A. Ryan, J.O. Gress, and C.R. Middaugh, *Biochim. Biophys. Acta.* **1249**, 100–108 (1995).
73. M. Menedez, M. Gasset, J. Laynez, C. Lopez-Zumel, P. Usobiaga, and E. Toper-Petersen, *Eur. J. Biochem.* **234**, 887–896 (1995).
74. J. Bandeker, *Biochim. Biophys. Acta.* **1120**, 123–143 (1992).
75. A. Dong, S.J. Prestrelski, S.D. Allison, and J.F. Carpenter, *J. Pharm. Sci.* **84**, 415–424 (1995).
76. K. Kato, T. Matsui, and S. Tanaka, *Appl. Spec.* **41**, 861–864 (1987).
77. H. Fabian, C. Schultz, D. Naumann, O. Landt, O. Hahn, and W. Saenger, *J. Mol. Biol.* **232**, 967–981 (1993).
78. I.E. Holzbaur, A.M. English, and A.A. Ismail, *Biochemistry* **35**, 5488–5494 (1996).
79. A. Nabet, J.M. Boggs, and M. Pezolet, *Biochemistry* **33**, 14792–14799 (1994).
80. S. Prestrelski, G.M. Fox, and T. Arakawa, *Arch. Biochem. Biophys.* **293**, 314–319 (1992).
81. I.H.M. Van Stokkum, H. Linsdell, J.M. Hadden, P.I. Haris, D. Chapman, and M. Bloemendal, *Biochemistry* **34**, 10508–10518 (1995).
82. A.A. Ismail, H. Mantsch, and P.T.T. Wong, *Biochim. Biophys. Acta* **1121**, 183–188 (1992).
83. A.L. Fink and B. Painter, *Biochemistry* **26**, 1665–1671 (1987).
84. N. Poklar, J. Lah, M. Salobir, P. Macek, and G. Vesnaver, *Biochemistry* **36**, 14345–14352 (1997).
85. H. Mach, G. Sanyal, D.B. Volkin, and C.R. Middaugh, in Z. Sharokh, V. Shuzky, J.L. Cleland, S.J. Shire, and T.W. Randolph eds., *Therapeutic Protein and Peptide Delivery*, ACS Symposium Series 675, 1997. pp. 186–205.
86. G. Sanyal, D. Marquis-Omer, and C.R. Middaugh, in R. Pearlman and Y.J. Wang eds., *Formulation, Characterization and Stability of Protein Drugs: Case Histories*, Plenum, New York, 1996, pp. 365–392.
87. A. Tanaka, J. Flanagan, and J. Sturtevant, *Protein Sci.* **2**, 567–576 (1992).
88. J. Behlke, *Eur. Biophys. J. Biophys. Lett.* **25**, 319–323 (1997).
89. H. Durchschlag, P. Zipper, G. Purr, and R. Jaenicke, *Colloid Polymer Sci.* **274**, 117–137 (1996).
90. P.G. Varley, A.J. Brown, H.C. Dawkes, and N.R. Burns, *Eur. Biophys. J.* **25**, 437–443 (1997).
91. Y. Ishihama, Y. Oda, N. Aksakawa, and M. Iwakura, *Anal. Sci.* **13**, 913–938 (1997).
92. Z. Deyl and R. Struzinsky, *J. Chromatogr.* **569**, 63–122 (1991).
93. T.A. Van De Goor, *Pharm-Biotechnol.* **7**, 301–327 (1995).
94. Y. Endo, H. Nagai, Y. Watanabe, K. Ochi, and T. Takagi, *J. Biochem. (Tokyo)* **112**, 700–706 (1992).
95. D.E. Koppel, *J. Chem. Phys.* **57**, 4814–4820 (1972).
96. M. Kataoka, J.F. Head, B.A. Seaton, and D.M. Engelman, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6944–6948 (1989).
97. M. Kataoka, Y. Hagihara, K. Mihara, and Y. Goto, *J. Mol. Biol.* **229**, 591–596 (1993).
98. L. Shi, M. Kataoka, and A.L. Fink, *Biochemistry* **35**, 3297–3308 (1996).
99. S. Beretta, L. Lunelli, G. Chirico, and G. Baldini, *Appl. Optics* **35**, 3763–3770 (1996).
100. J.M. Flanagan, M. Kataoka, D. Shortle, and D.M. Engelman, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 748–752 (1992).
101. M. Kataoka, I. Nishii, T. Fujisawa, T. Ueki, F. Tokunaga, and Y. Goto, *J. Mol. Biol.* **249**, 215–228 (1995).
102. E.E. Lattman, K.M. Fiebig, and K.A. Dill, *Biochemistry* **33**, 6158–6166 (1994).
103. J.W. Lewis, R.A. Goldback, D.S. Kliger, X. Xie, R.C. Dunn, and J.D. Simon, *J. Phys. Chem.* **96**, 5243–5245 (1992).
104. D. Reinstadler, H. Fabian, J. Backmann, and D. Naumann, *Biochemistry* **35**, 15822–15830 (1996).
105. J. Backmann, H. Fabian, and D. Naumann, *FEBS Lett.* **364**, 175–178 (1995).
106. J. Backmann, C. Schultz, H. Fabian, U. Hahn, W. Saenger, and D. Naumann, *Proteins* **24**, 379–387 (1996).

See also DENATURATION, PROTEINS, SOLVENT MEDIATED; PROTEIN AGGREGATION, DENATURATION.

THERMOLYSIN

SATOSHI HANZAWA
Tokyo Research Laboratory
Hayakawa, Japan

SHUN-ICHI KIDOKORO
Sagami Chemical Research Center
Kanagawa, Japan

KEY WORDS

Aspartame
Site-directed mutagenesis
Thermolysin

OUTLINE

INTRODUCTION

Thermolysin (EC 3.4.24.4) is a thermostable metalloprotease produced by *Bacillus thermoproteolyticus* Rokko (1). It is the most thermostable protease currently produced for industrial use. Thermolysin and similar microbial metalloproteases require a metal ion, usually Zn^{2+}, for their catalytic activity. Formerly, they were called neutral proteases because their optimum pH is around pH 7. They also have a common property of substrate specificity in hydrolyzing peptide bonds involving an amino group of hydrophobic and bulky amino acid residues such as Leu or Phe (2). Among these proteases, thermolysin has been extensively studied regarding its structure and catalytic mechanism (3–6). Currently, we can easily obtain structural data of thermolysin from the Protein Data Bank. Although metalloproteases have been produced not only for laboratory use but also for industrial use, the market size has been small, because their substrate specificity is narrower compared to that of other proteases such as serine proteases or cysteine proteases. Their use has been restricted to hydrolysis of proteins that cannot be hydrolyzed by other proteases, for example, that of barley protein, which contains a specific inhibitor toward serine protease (7). Recently, however, thermolysin has been used extensively in aspartame (APM) production at Holland Sweetener Company (HSC), a joint venture of TOSOH (Japan) and DSM (the Netherlands). APM, L-α-aspartyl-L-phenylalanine methyl ester, is an artificial sweetener that is 200 times sweeter than sucrose. In APM production, thermolysin plays an important role because it catalyzes the synthesis of an APM precursor, *N*-benzyloxycarbonyl-APM (Z-APM), from Z-L-Asp and L-Phe methyl ester (L-PheOMe).

$$\text{Z-L-Asp} + \text{L-PheOMe} \xrightarrow[\text{thermolysin}]{} \text{Z-L-Asp-L-PheOMe}$$

Because the coupling reaction is the key step in APM production, improvement of this reaction is highly desired. Immobilization of thermolysin has been attempted as one of the improvements to establish efficient recycling of thermolysin (8,9). We sought to improve thermolysin activity by site-directed mutagenesis, that is, by protein engineering, because such activity enhancement not only results in the reduction of enzyme consumption but also circumvents the enzyme recovery step, if enzyme cost is reduced to a negligible amount. It thus also can affect the fixed cost of APM production. In this article, we explain the results of our study to improve thermolysin.

ENHANCEMENT OF THERMOLYSIN ACTIVITY

Before introduction of our improvement study, several other reports regarding activity enhancement of thermolysin are described herein (Table 1). It is well known that enzyme activity is affected by additives in the reaction mixture, such as salts and organic solvents. Thermolysin is also affected by additives, which results in either enhancement or reduction of the activity. Thermolysin activity is enhanced by almost all salts consisting of monovalent anions and cations (10,11). In the case of alcohol, *n*-pentanol at concentrations below saturation enhances activity while *n*-butanol is inhibiting (12). The effect of alcohol is highly dependent on physicochemical properties such as log P or ϵ (13). Pressure is known to be another factor that enhances activity (14). The mechanism of the activation by salts, organic solvents, or high pressure is not clear, although a slight structural change of the enzyme by salts has been observed spectrophotometrically (10). We thought that activity enhancement by these methods on an industrial scale is not feasible from the economical point of view, because expensive equipment would be required to avoid corrosion by salts or explosion of solvents or to resist high pressure. Additionally, such additives result in the increase of industrial wastes.

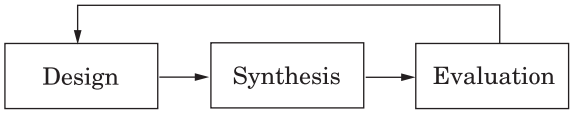
Kubó et al. (15) demonstrated activity enhancement of thermolysin for casein hydrolysis (Table 1). Mutant enzymes can be expected to be more effective for industrial use compared to salts, organic solvents, or high pressure. However, activity for casein hydrolysis does not correlate with that for hydrolysis or synthesis of dipeptide derivatives as described later. Two thermolysin mutants, Y110W and Y211W, have not been evaluated for their activity for Z-APM hydrolysis or synthesis.

IMPROVEMENT BY PROTEIN ENGINEERING

As described, thermolysin activity can be enhanced by additives. However, we preferred improvement by protein engineering because the use of thermolysin mutants is more cost effective. The strategy, method, and results of our study are described next.

Strategy

In recent protein engineering studies, thermal stability of proteins can be enhanced by introducing or reinforcing one or more of several physicochemical intramolecular interactions, such as hydrogen bonding, S–S bridges, and so forth. Most attempts to modify the activity, however, have resulted in suppression of biological activity. To find a novel strategy for producing more active enzymes, we conducted a "protein engineering cycle" consisting of three elements—design, synthesis, and evaluation—as shown in the following scheme:



In this scheme, feedback from evaluation to design is highly important. Feedback can revise our working hypothesis for the structure–function relationships. Working with this cycle will not only give us novel useful proteins

Table 1. Activity Enhancement of Thermolysin by Additives, Pressure, and Site-Directed Mutagenesis

Additive	Condition	Substrates	Reaction	Relative activity[a]	Reference
NaCl	4 M	Z-APM	Hydrolysis	9.0	10
NaCl	2.7M	Z-APM	Synthesis	4.2	10
NaCl	4 M	FaGLa	Hydrolysis	13	11
N-Pentanol	60% saturation	Z-FFM	Synthesis	4.7	12
Pressure	1,000 psi	FaGLA	Hydrolysis	4.5	14
Mutation	Y110W	Casein	Hydrolysis	1.6	15
Mutation	Y211W	Casein	Hydrolysis	1.5	15

Z-APM, *N*-benzyloxycarbonyl-aspartame; Z-FFM, Z-Phe-PheOMe; FaGLa, Fa-Gly-Leu-Ala.
[a]Relative activity compared with activity in absence of the additives.

as products but also gives us worthwhile knowledge about the mechanisms underlying the relationship between protein structure, properties, and functions.

Although enzymes do not change the thermodynamic equilibrium between starting materials and products, they accelerate reaction rates by stabilizing the transition state of the reaction. This stabilization is realized enthalpically or entropically, and the Gibbs energy of the transition state is decreased. It is well known that one of the stabilization forces for the transition state is electrostatic interaction (including hydrogen bonding) between an enzyme and its ligand. Several ionizable amino acids such as Asp, Glu, and His play an important role in this stabilization. The equilibrium between the protonation and deprotonation states of these residues, which is mainly determined by electrostatic interaction, controls enzyme activity. Usually the change of protonation state from its active form completely eliminates activity. The pH profile of enzyme activity is generally thought to reflect the protonation state. Therefore, changes in electrostatic potential at the active site may affect both the stability of the transition state and the protonation equilibrium. The optimum pH for the activity of thermolysin is known to be around 7 (1). The synthesis of Z-APM is, however, carried out industrially under a more acidic pH because the methyl ester of the substrate, phenylalanine methyl ester, and the product are hydrolyzed rapidly above pH 7 (16). If the optimum pH can be shifted to the acidic pH region, enzymatic activity in this region may be enhanced. To accomplish this we decided to change the electrostatic potential around the active site.

Design for Higher Activity

There are two residues, Glu 143 and His 231, and zinc ion, Zn^{2+}, in the active site of this enzyme (5,6). We decided to change first the electric charge of Gln 119 to affect the electrostatic potential at Glu 143 (17,18). Gln 119 is known to form a hydrogen bond to the side chain of Ser 103. When Gln 119 is replaced by other amino acids, that hydrogen bond is broken, thereby affecting dynamic properties and stability of the enzyme.

Synthesis of Thermolysin Mutants

The site-directed mutagenesis at the 119th site was performed by the overlap extension method (18,19) using the polymerase chain reaction. The mismatched primer and its complementary sequence were synthesized and used as inside primers, which anneal to the same segment of *npr*M gene coding for thermolysin. The mismatches lead to random sequence alteration in the 119th amino acid residue. The amplified DNA fragment containing random mutation was digested with *Bam*HI and *Sph*I, and substituted for the same region of the native *npr*M gene in pUBTZ2. The *npr*M gene was cloned from *Bacillus stearothermophilus* MK232, which was isolated by TOSOH from soil to provide a protease for APM production (20).

In the early stages of investigation of this metalloprotease, it was believed that its primary structure was different from that of thermolysin at the 37th and the 119th sites (21), because the amino acid sequence of thermolysin was reported to be Asp at the 37th site and Glu at the 119th site by Titani et al. (4). More recent studies, however, revised the sequence of thermolysin as Asn at the 37th and Gln at the 119th (22,23), exactly the same as the metalloprotease from *npr*M gene. The *npr*M gene was first cloned by Kubo et al. (20) in *Bacillus subtilis* MT-2 whose genotype is *trp*C2 *leu*C7 *hsd*R *hsd*M Npr^- (24). Expression plasmid pUBTZ2, which is a shuttle vector between *Escherichia coli* and *B. subtilis*, was constructed from pUC9, pUB110 (25), and a fragment of pMK4 (26) involving *npr*M. The thermolysin mutants were expressed in *B. subtilis* MT-2 by cultivation in 2 L broth (10 g of yeast extract, 20 g of Bacto peptone, and 5 g of NaCl in 1 L of H_2O) containing 20 μg/mL kanamycin at 37 °C for 20 h in a culture flask. The mutants were recovered by ammonium sulfate precipitation of culture supernatant and further purified by two-step column chromatography with Butyl Toyopearl 650 M (Tosoh, Japan) and TSK gel G2000SW (Tosoh, Japan).

Evaluation of Thermolysin Mutants

The amino acids at the 119th site were confirmed by sequencing the plasmid DNA. The purity and the isoelectric point were determined by gel isoelectric focusing using Phast System (Pharmacia, Sweden). The *pI* value of three mutants, namely Lys mutant (Q119K), Arg mutant (Q119R), and Met mutant (Q119M), is 5.3 while that of the other mutants and the wild type (WT) is 5.0. The difference in pI was thought to result from a change in one electric charge of the mutant, judging from the *pI* calculation of the amino acid composition.

The hydrolytic activities were evaluated by measuring hydrolysis of *N*-(3-{2-furyl}acryloyl)-glycyl-L-leucine amide (FaGLa) and *N*-(3-{2-furyl}acryloyl)-L-aspartyl-L-phenyl-

alanine methyl ester (Fa-APM). The theoretical curve of an integral Michaelis–Menten equation (27) fitted well with all experimental data within experimental error (10^{-3} optical density). The decrease in absorbance at 345 nm was monitored by a spectrophotometer (UB-35; Jasco, Japan) connected to a thermostat circulator (EL-15; Taitech, Japan) and a personal computer (PC9801; NEC, Japan) for data acquisition. The stored data were fitted with an integral Michaelis–Menten equation by a nonlinear least-squares method to determine k_{cat} and K_m. All measurements were done under a solution condition of 200 mM Tris-maleate buffer with 10 mM $CaCl_2$ at 37 °C. The concentrations of substrate and enzyme were about 1–2 mM and 1 nM, respectively.

Protein concentration was determined by photometry with an extinction coefficient of 1.67×10^4 OD g/mL. The computer program was described in BASIC on PC9801 and FORTRAN with a program package SALS on a workstation (Titan 750; Kubota Computer, Japan). Fa-APM was synthesized using *N*-(3-{2-furyl}acryloyl)-L-Asp and L-PheOMe following a modified method of Oyama et al. (28). To check the electrostatic effect of the 119th site on activity, we selected the potentially positively charged residues, His, Lys, and Arg, and the potentially negatively charged residues, Glu and Asp. In contrast to the expectations based on the electrostatic effect, hydrolytic activity of all mutants was higher than that of WT, indicating that the main reason for the activity enhancement is not the electrostatic effect. We therefore decided to generate all mutants at this site.

The hydrolytic activity toward FaGLa of the mutants depending on pH is presented in Figure 1. The mutant Q119E exhibited the highest activity, five times higher than WT. It should be noted that the hydrolytic activities of almost all these mutants exceeded that of the WT, indicating that the amino acid at this position is selected by nature to show the lowest activity.

We then substituted Ala and Gly for Ser 103, expecting to cause the disruption of the hydrogen bond between this site and the 119th site. These two mutants, S103A and S103G, showed activity 3.0 and 2.6 times higher, respectively, than that of the WT. The similar increase in activity observed as a result of amino acid substitution at the 103rd site indicates that disruption of hydrogen bond is the main cause of activity enhancement.

We found low K_m values for another substrate, Fa-APM, and this enabled us to determine k_{cat} and K_m values separately. The $1/K_m$ values of the mutants showed a strong correlation to k_{cat}/K_m value, while k_{cat} values showed little change from the WT (18). This result indicates that the mutation increases the affinity of the substrate so that it enhances k_{cat}/K_m for FaGLa and Fa-APM (18). The hydrogen bond of Gln 119 seems to fix the part of the β sheet from the 112th to 128th residues to the other strand. Judging from the three-dimensional structure, the strand from the 112th to 118th residue is considered to contact the pep-

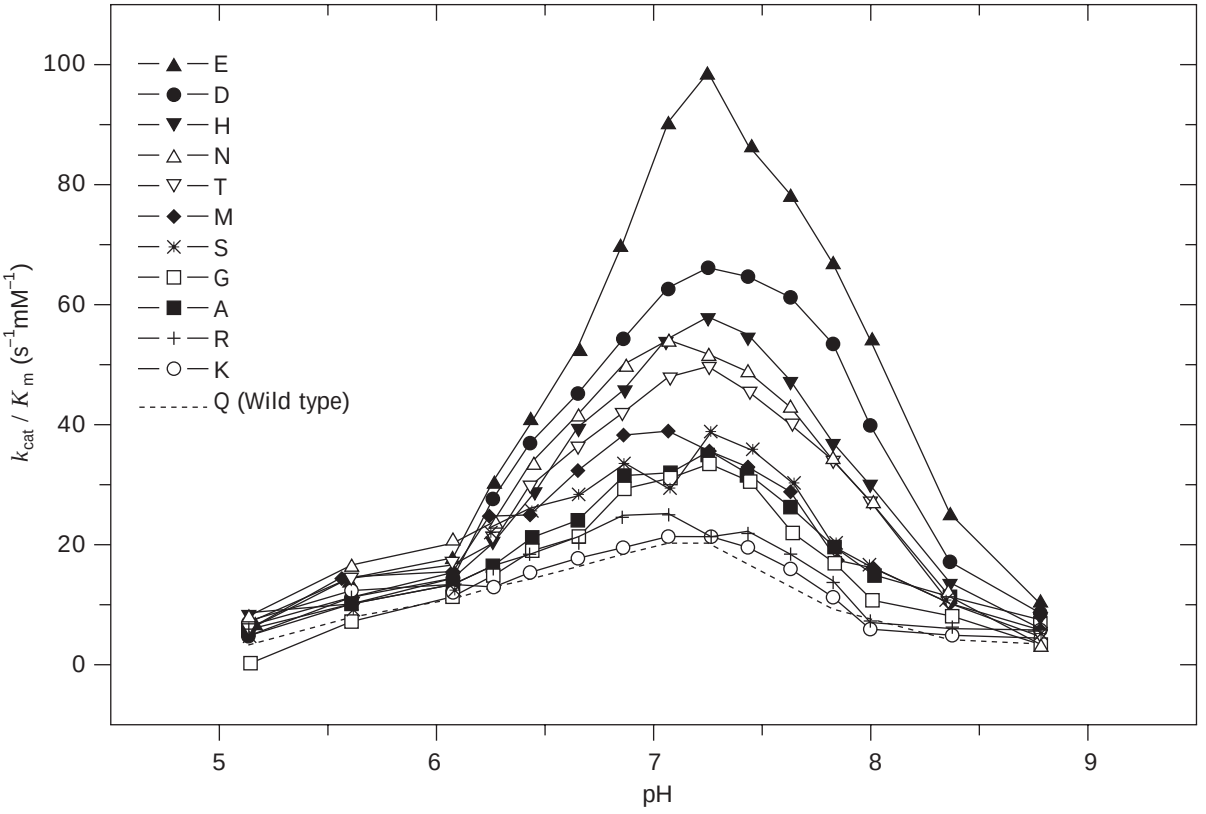


Figure 1. pH dependence of the hydrolytic activity of thermolysin and its mutants at the 119th position in 200 mM Tris-maleate buffer with 10 mM $CaCl_2$ at 37.0 °C.

tide chain of the substrate directly. The flexibility of the β strand may be increased by breakage of the hydrogen bond, thereby causing an increase in affinity of the binding site to the substrate. When flexibility is increased, it is usually observed that the thermal stability of a protein is decreased (18). In this case, a negative correlation between activity and stability is seen. If we assume that the activity is enhanced by flexibility, then there is negative correlation between flexibility and stability. Two mutants, Q119K and Q119R, however, markedly deviated from the negative correlation. From the shift of the *pI* values of these mutants, the Lys and Arg at the 119th site were supposed to be positively charged. The positive charge at this site may suppress the activity. It is reasonable that this change in electric charge affects the activity because this site is selected as a site that affects the electrostatic potential of Glu 143. Measurement of *pI* values of Q119D and Q119E showed the Asp and Glu at this site to be neutral (protonated) around pH 5.0.

From these results, we concluded that flexibility around the substrate-binding site plays an important role in enhancing the activity. The electrostatic potential also influences the change of activity. Measurement of the synthetic activity of these mutants, however, disclosed that synthetic activity of the 119th series mutants did not exceed that of WT (see Table 2). Under the conditions for the coupling reaction, the concentration of the substrates is very high so that the K_m value, which represents the affinity to substrate, does not affect the reaction rate. The important factor for enzymatic activity under the synthetic conditions, therefore, is not the K_m value but the k_{cat} value.

Reiterating the Protein Engineering Cycle

The 213th and 227th sites were then selected as two of the sites possibly most effective in changing the electrostatic potential of His 231 at the active site. Mutations of *npr*M at the 213th and 227th sites were performed by the M13 phage mutagenesis method (29).

In the case of Asp 213 mutants, four mutations, D213N, D213H, D213K, and D213R, were introduced in thermolysin. The hydrolytic activity of these mutants against FaGLa became lower than that of the WT, which led to a negative correlation between the activity and *pI* value of the mutants (30). Considering the three-dimensional structure and electrostatic calculations, it was concluded from these experimental data that the contribution of the pK shift of residues other than the His 231 and the 213th residue must be taken into account in explaining the change in activity.

The mutant N227H, in which a positively charged residue His is introduced instead of Asn 227, showed higher hydrolytic and synthetic activities (Table 2). In contrast, in the mutant N227D, where a negatively charged residue, Asp, is introduced at this site, lower hydrolytic and synthetic activities were seen. The positive charge enhanced enzymatic activity at this site, while the positive charge at the 213th site suppressed the activity. The mechanism of enhancement and suppression of activity by electrostatic interaction seems to be complex and cannot be explained by a simple model. Further theoretical and experimental research is necessary to explore this process.

The Asp 150 was replaced by neutral Asn or by positively charged His so as to influence the electrostatic potential of the Zn ion at the active site. The mutations were performed by PCR (31). In this case the k_{cat}/K_m for hydrolysis of Fa-APM decreased (both k_{cat} and K_m are increased) below that of WT. However, these mutants, D150N and D150H, respectively, exhibited increased activity for Z-APM synthesis (Table 2).

As shown in the Q119 series mutants, flexibility around the active site may have a great effect on the activity. To modify the flexibility, we searched for cavities in the protein molecule using its three-dimensional structure. Figure 2 shows the density map of a cross section of thermolysin; the darker and lighter portions represent higher and lower densities, respectively. As can be seen, the central part of the N-terminal domain has lower density than the surrounding part, clearly indicating the presence of a cavity. Computer graphics suggested that the size of the cavity can be easily modified by replacing several amino acids such as Leu 144. This cavity was also remarked by Vriend and Eijsink in stabilization of thermolysin-like metalloprotease from *B. stearothermophilus* CU21 (32), although the Phe mutant at Leu 144 (L144F), designed to fill the cavity, did not show an increase in thermal stability. We introduced mutation at the 144th site by PCR. It was found that several mutants, such as the Ser mutant (L144S), showed higher hydrolytic and synthetic activities than WT (Table 2).

Based on the three-dimensional structure, several single-point mutants having higher synthetic activity for Z-APM were prepared. We combined these single mutations to obtain more active enzymes, and their synthetic activities are shown in Table 2. We obtained fivefold higher synthetic activity for a triple mutant, L114S-D150H-N227H (33).

Table 2. Activity of Thermolysin Mutants with Z-APM and Casein as Substrates

	Z-APM[a]		
Thermolysin mutant	Synthesis (kat/mol): pH 6.0	Hydrolysis (kat/mol): pH 7.0	Casein[b] (U/mg): pH 7.2
Wild type	0.06 (1.0)	3.5 (1.0)	9,200 (1.0)
Q119E	0.07 (1.1)	5.6 (1.6)	4,000 (0.4)
L144S	0.09 (1.5)	6.3 (1.8)	4,600 (0.5)
D150H	0.12 (2.0)	6.3 (1.8)	9,200 (1.0)
D150W	0.26 (4.3)	11.9 (3.4)	7,400 (0.8)
N227H	0.09 (1.5)	7.4 (2.1)	7,400 (0.8)
L144S-D150H	0.14 (2.4)	6.4 (1.8)	2,700 (0.3)
L144S-N227H	0.11 (1.8)	10.5 (3.0)	3,800 (0.4)
D150H-N227H	0.21 (3.5)	11.3 (3.2)	9,800 (1.1)
D150W-N227H	0.45 (7.5)	22.6 (6.5)	6,700 (0.7)
L144S-D150H-N227H	0.30 (5.0)	21.5 (6.1)	3,200 (0.4)
L144S-D150W-N227H	0.63 (10.5)	37.4 (10.7)	1,900 (0.2)

Note: Numbers in parentheses are activity relative to wild type.

[a]The activity for Z-APM, that is, the synthesis or hydrolysis of 1 mol Z-APM in 1 s, was defined as 1 katal (kat).

[b]The activity for casein, that is, the release of acid-soluble peptides corresponding to the absorbance of 1 μg of tyrosine in 1 min, was defined as 1 unit (U).

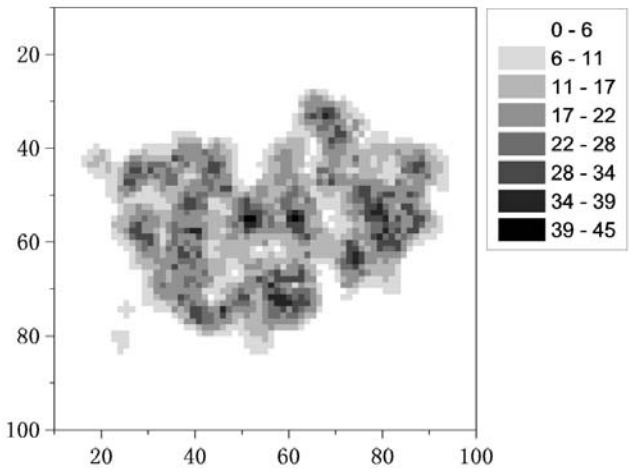


Figure 2. Cross section of average number–density of thermolysin. The density is represented by shading, the highest density being represented by black. *N*-terminal domain is plotted on the left.

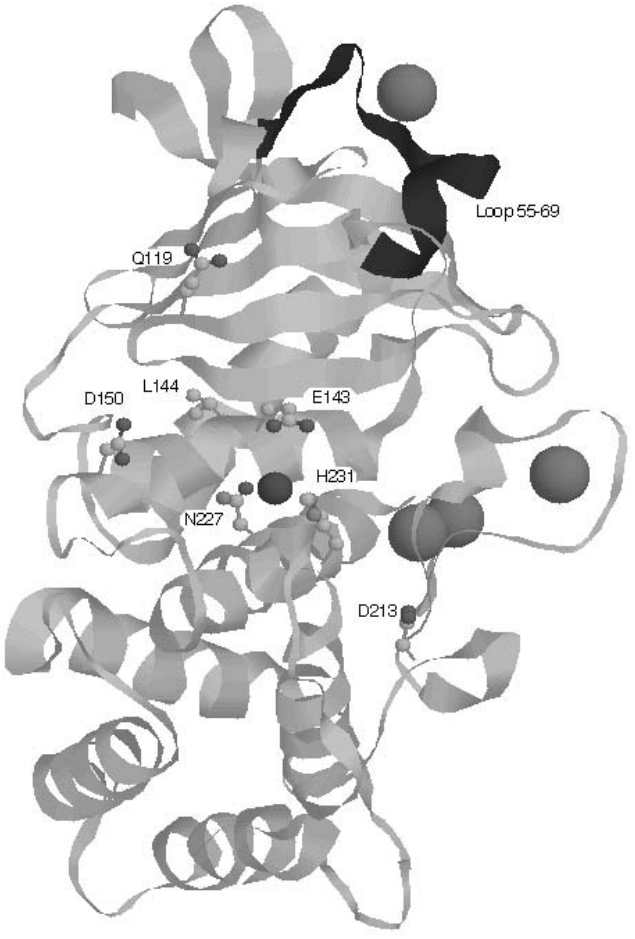


Figure 3. Three-dimensional structure of thermolysin. The peptide chain is represented by a ribbon model. The small ball in the center of this model represents the zinc ion, which is essential for catalytic activity. The four larger balls represent the calcium ion required for stabilization of the three-dimensional structure. E143 and H231 are essential for the catalytic activity. Mutated residues are located as shown here. Loop 55–69 is the dominant region in the stability of thermolysin (38).

Activity for the Aspartame Precursor Z-APM

The activity of the thermolysin mutants for Z-APM synthesis is described in detail in this section. Because mutants at Gln 119, Leu 144, Asp 150, and His 227 sites, specifically the Q119E, L144S, D150H, and N227H mutants, exhibited higher activity for FaGLa or Fa-APM hydrolysis, we examined their activity for Z-APM synthesis. Positions of these sites are indicated in Figure 3. Z-APM synthetic activity was evaluated with 0.1 M Z-L-Asp and 0.1 M L-PheOme as substrate in 0.1 M Tris-maleate buffer at pH 0.6 at 35 °C. After 30 min incubation, the reaction was terminated by addition of EDTA. Z-APM in the mixture was measured by a HPLC system (Tosoh, Japan) equipped with an ODS column (Nacali Tesque, Japan). Z-APM hydrolytic activity, the reversal of Z-APM synthesis, was also analyzed according to the modified method of Inouye (10) by monitoring absorbance at 224 nm at 35 °C with 1 mM of Z-APM as substrate in 0.1 M Tris-HCl buffer (pH 7.0) containing 10 mM $CaCl_2$.

L144S, D150H, and N227H exhibited enhancement of activity both for synthesis and hydrolysis of Z-APM while Q119E only exhibited enhancement in the hydrolysis (see Table 2). The greatest effect on the activation was obtained by D150H wherein synthetic activity was 2-fold higher than for WT; activities of both L144S and N227H were only 1.5-fold higher. The combination of L144S and another mutation enhanced the activity by simple addition of the effects of each mutation. The activity of the double mutants L144S-D150H and L144S-N227H were 2.5-fold and 2-fold higher than WT, respectively. However, when D150H and N227H were combined, the resulting activity was higher than the simple addition of the effect of each mutation. For example, if the enhancing effect were caused by simple addition, the activity of the double mutant D150H-N227H would be 2.5-fold higher than WT, instead of the observed 3.5-fold higher activity (Table 2). The triple mutant, L144S-D150H-N227H, exhibited the highest activity, which was 5-fold higher than WT. We called this thermolysin mutant TZ-1, and its efficiency was further evaluated by small-scale production under the production conditions of the HSC plant; see also ASPARTAME.

These mutants did not always exhibit enhanced activity for other substrates. For example, the activity for casein hydrolysis, analyzed following the method of Endo (1), was decreased in two mutants, Q119E and L114S, while the activity of D150H and N227H was the same as that of the WT (Table 2).

The kinetic parameters of Z-APM synthesis were examined at various concentrations of Z-Asp and 800 mM of L-PheOMe. The reaction rate increased with increasing concentration of Z-Asp up to 100 mM and then declined. This inhibition of the synthetic reaction at high concentration of Z-Asp was also reported by Nagayasu et al. (34) when the reaction was performed with immobilized ther-

molysin in MES buffer saturated with ethyl acetate. Nagayasu reported that the K_m value of the reaction was 2.8 mM and 250 mM for Z-Asp and PheOMe, respectively, and the K_i value for Z-Asp was 100 mM. Although we have not calculated the K_i value of the condensation reaction, k_{cat} and K_m values of the reactions were calculated from Lineweaver–Burk plots at the range of Z-Asp concentration in which no inhibition is observed (Fig. 4a). The kinetic parameters for Z-Asp are summarized in Table 3.

In our results, the K_m of WT for Z-Asp was 13 mM, which is four to five times higher than the results obtained by Nagayasu et al. (34), apparently because of differences in reaction conditions such as solvent, concentration of substrates, or immobilization of enzyme. L144S decreased the K_m value significantly with slight increase of k_{cat} value, D150H increased both k_{cat} and K_m, and N227H increased k_{cat} without affecting K_m. These results, except for the k_{cat}/K_m value of D150H, were similar to the kinetic parameters obtained when Fa-APM is the substrate. In Z-APM synthesis, the k_{cat}/K_m value of D150H increased while there was a decrease for Fa-APM hydrolysis. This change was caused by the greater effect of this mutation on K_m for Fa-APM than that for Z-APM and the difference of the Fa group and Z group at the N-terminal of APM. The Fa group is less hydrophobic than the Z group.

When k_{cat}/K_m values for Z-Asp of TZ-1 and WT were compared, the efficiency of the former was seven times higher. We were not able to estimate the K_m value for PheOMe in Z-APM synthesis because the Lineweaver–Burk plot at 100 mM Z-Asp and variable PheOMe passed through the graph origin (Fig. 4b), suggesting that the K_m values for PheOMe were greater than 900 mM under this reaction condition.

Optimization of Amino Acids at the 150th, 144th, and 227th Sites

Because D150H exhibited the greatest activity in the complete set of mutants tested, we further examined the substitution of other amino acids for Asp 150. The mutations were introduced by PCR using a random primer. In the complete set of 19 mutants at this site, the Trp mutant (D150W) exhibited the highest activity both for synthesis and for hydrolysis of Z-APM (Table 2) (35), about four times higher than the activity of the WT. The activity of the series of the mutants apparently exhibits correlation with hydrophobic parameters (33), such as the π value (36) or ΔH value (37) of substituted amino acids. From the analysis of three-dimensional structure, we assumed that the enhancement of activity might result from reinforcement of the hydrophobic interaction between the Z group of the substrate and the amino acids at the 150th site.

We also evaluated the substitution of other amino acids for Leu 144 and Asn 227 by random mutation. In the 12 variants of the mutants obtained at the 227th site, that is, Asp, Glu, His, Lys, Arg, Thr, Ala, Met, Trp, and Leu mutants, only N227H exhibited apparent activity enhancement. We could not find any relationship between activity and the physicochemical properties of amino acids at the 227th site. The activation mechanism of N227H is not yet clear, but it is believed that the electrostatic effect caused by the positive charge of His might be the main cause of the activity enhancement. In case of the 144th site, 10 variants of mutants, that is, the Glu, Asn, Gln, Ser, Thr, His, Cys, Ala, Tyr, Phe, and Trp mutants, were obtained. Although this series of mutants exhibited almost the same activity for Z-APM synthesis, activity for the hydrolysis of Z-APM exhibited a negative correlation with the hydrophobic properties of the substituting amino acids. This result seems to support our hypothesis described in the previous section; that is, the activity is affected by the flexibility of the enzyme molecule, which thereby influences the cavity in the center of thermolysin. Hydrophilic amino acids at the 144th site appear to contribute not only to cavity expansion but also to weakening hydrophobic interactions within the thermolysin molecule.

Finally, the triple mutant obtained, L144S-D150W-N227H (named TZ-5), exhibited about 10-fold higher activity for Z-APM synthesis and hydrolysis (Table 3) (35). D150W apparently enhanced the k_{cat} value of the reaction without an effect on K_m value. The k_{cat}/K_m value of TZ-5 was 11 times higher than that of WT.

STABILIZATION OF THERMOLYSIN

Recently, Van den Burg et al. (38) reported remarkable stabilization of thermolysin-like metalloprotease in boiling conditions (100 °C) by site-directed mutagenesis at eight sites (Table 4). This protease (TLP-ste) from *B. stearothermophilus* CU21 differs from thermolysin at 44 sites in its sequence and is less stable than thermolysin. Five mutations, A4T, T56A, G58A, T63F, and A69P, are substitutions of thermolysin-type residues for TLP-ste. In these mutations, G58A was originally suggested by Imanaka et al. (39). In a previous study, Van den Burg et al. (40) replaced almost all residues of TLP-ste by thermolysin-type amino acids and found that substitutions in the N-terminal domain are more effective than those in the C-terminal domain. They presented a hypothesis that local denaturation at the 55–69 loop in the N-terminal domain determines thermal denaturation of TLP-ste (41,42). The region corresponding to the 55–69 loop in thermolysin is shown in Figure 3. The three mutations, S65P, G8C and N60C, were rationally designed to fix the flexible region by substitution of Pro (42) and by introduction of an S–S bridge between the 8th and the 60th sites (38). Thus, the resultant 8-point mutant exhibited a half-life 170 times longer than thermolysin at 100 °C. These mutations did not affect the specific activity of TLP-ste (38). This result suggests that a highly active and extremely stable thermolysin mutant is possibly produced when their mutations are combined with our mutations. Three of these mutations, S65P, G8C, and N60C, can be introduced into thermolysin. Other mutations that stabilize thermolysin (15) are also summarized in Table 4. Thermolysin will be further stabilized when this mutation is combined with others.

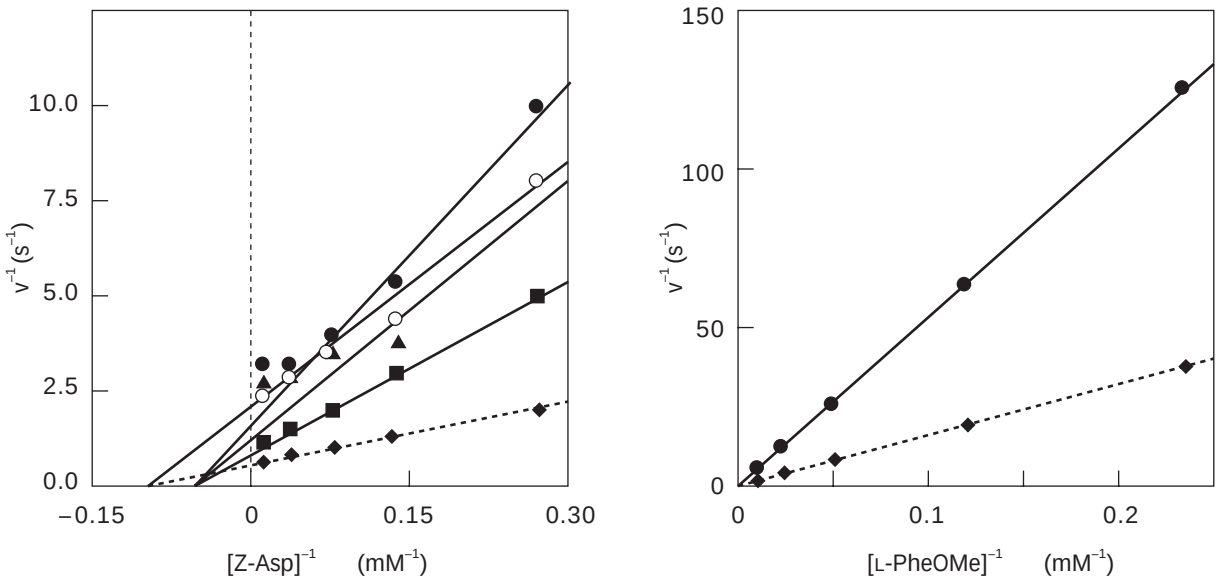


Figure 4. Lineweaver–Burk plots of the condensation reaction of Z-L-Asp (**a**) and L-PheOMe (**b**) catalyzed by thermolysin mutants. The reactions were performed at a constant concentration of L-PheOMe and various concentrations of Z-L-Asp (**a**) or at constant Z-L-Asp and various L-PheOMe (**b**) at pH 6.0 and 35 °C. Closed circles, open circles, squares, triangles and rhombuses represent wild type, L144S, D150H, N227H, and L144S-D150H-N227H, respectively.

CONCLUSION

In designing thermolysin mutants, we started by influencing the electrostatic environment around the active site. It was found during the protein engineering cycle that two additional factors influence thermolysin activity, namely, molecular flexibility and the hydrophobic interaction between the N-protective group of the substrate and the side chain of the 150th site of the enzyme. We were able to locate sites that effectively influence the activity of thermolysin, although some of the mutants at the sites do not enhance the activity, the reasons being only partly inter-

Table 3. Kinetic Parameters for Z-APM synthesis

Thermolysin variant	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} mM^{-1})
Wild type	0.4	13	0.03
L144S	0.4	9	0.04
D150H	0.9	16	0.06
D150W	1.5	12	0.12
N227H	0.6	13	0.05
L144S-D150H-N227H	1.6	8	0.20
L144S-D150W-N227H	3.0	9	0.33

The kinetic parameters were determined by Lineweaver–Burk plots at 3–100 mM Z-Asp and constant L-PheOMe concentration (800 mM).

Table 4. Stabilization of Thermolysin and Thermolysin-Like Metalloprotease from *Bacillus stearothermophilus* CU21 by Site-directed Mutagenesis

Enzyme	Mutation	Half-life[a] (min)	$T_{1/2}$[b] (°C)	Incubation time/temp.	Reference
TLN	Y93S/G	+36/24	—	80 °C	15
TLP-ste[c]	A4T	—	+1.8	30 min	32
TLP-ste	T56A	—	+1.5	30 min	32
TLP-ste	G58A	—	n.d.	30 min	32,39
TLP-ste	T63F	—	+6.2	30 min	32
TLP-ste	A69P	—	+5.6	30 min	32
TLP-ste	S65P	—	+4.2	30 min	40
TLP-ste	6 points[d]	—	+23.5	30 min	41
TLP-ste	8 points[e]	+169	—	100 °C	38

TLN, Thermolysin.
[a]Difference in half-life between wild-type TLN and mutants of TLN or P-TLP-ste during incubation at indicated temperature.
[b]Difference in temperature required for 50% denaturation between mutants and wild type.
[c]TLP-ste indicates the metalloprotease from *B. stearothermophilus* CU21.
[d]6 points is the combination of A4T, T56A, G58A, T63F, A69P, and S65P.
[e]8 points is the combination of G8C and N60C in the 6-points mutant. The mutated sites are indicated according to the sequence alignment by Vriend and Eijsink (32).

preted (30). Further studies, integrating up mutations and down mutations, must be carried out to establish a more reliable design principle. Apart from the theoretical aspect, we found it quite successful to enhance the synthetic activity of thermolysin by 10-fold. The production cost of APM will be lowered when the thermolysin mutant is applied in the actual industrial process.

ACKNOWLEDGMENTS

The improvement of thermolysin activity was supported by Holland Sweetener Company. We are grateful to Drs. Hiromasa Nagao, Toshio Miyake, Atsuo Aoyama, Takashi Yoneya, Ken-ichi Kai, Seigo Ooe, Yoshikazu Tanaka, You-ichiro Miki, and Kimiko Endo for preparation and evaluation of thermolysin mutants. We thank Dr. Kakidani and Dr. Dionisio-Sese for their critical reading of this manuscript. We would like to express special thanks to Dr. Kiyotaka Oyama, Mr. Ichiro Hiraki, Mr. Akira Tokuda, and Prof. Akiyoshi Wada for the organization of this work.

BIBLIOGRAPHY

1. S. Endo, *J. Ferment. Technol.* **40**, 346–353 (1962).
2. K. Morihara, *Biochem. Biophys. Res. Commun.* **26**, 656–661 (1967).
3. K. Titani, M.A. Hermodson, L.H. Ericsson, K.A. Walsh, and H. Neurath, *Biochemistry* **11**, 2427–2435 (1972).
4. K. Titani, M.A. Hermodson, L.H. Ericsson, K.A. Walsh, and H. Neurath, *Nat. New Biol.* **238**, 35–37 (1972).
5. P.M. Colman, J.N. Jansonius, and B.W. Matthews, *J. Mol. Biol.* **70**, 701–724 (1972).
6. M.A. Holmes, and B.W. Matthews, *J. Mol. Biol.* **160**, 623–639 (1982).
7. K. Aunstrup, in A.H. Rose ed., *Microbial Enzymes and Bioconversions*, Academic Press, New York, 1980, pp. 50–114.
8. K. Oyama, S. Irino, and N. Hagi, in K. Mosbach ed., *Methods in Enzymology*, vol. 136, *Immobilized Enzyme and Cells, Part C*, Academic Press, New York, 1987, pp. 503–516.
9. K. Nakanishi, A. Takeuchi, and R. Matsuno, *Appl. Microb. Biotechnol.* **32**, 633–636 (1990).
10. K. Inouye, *J. Biochem. (Tokyo)* **112**, 335–340 (1992).
11. K. Inouye, S. Lee, and B. Tonomura, *J. Biochem. (Tokyo)* **315**, 133–138 (1996).
12. M.N. Alam, K. Tadasa, and H. Kayahara, *Biotechnol. Lett.* **18**, 45–50 (1996).
13. M.N. Alan, K. Tadasa, T. Maeda, and H. Kayahara, *Biotechnol. Lett.* **19**, 1129–1133 (1997).
14. H. Kitano and N. Ise, *Biotechnol. Bioeng.* **31**, 507–510 (1987).
15. M. Kubo, Y. Mitsuda, M. Takagi, and T. Imanaka, *Appl. Environ. Microbiol.* **58**, 3779–3783 (1992).
16. B.E. Homler, in L.D. Stegink and L.J. Filer, Jr. eds., *Aspartame Physiology and Biochemistry*, Dekker, New York, 1984, pp. 247–262.
17. S. Kidokoro, Y. Miki, K. Endo, and A. Wada, in A. Pullman and B. Pullman eds., *Modeling of Biomolecular Structures and Mechanisms*, Kluwer, Dordrecht, 1995, pp. 399–407.
18. S. Kidokoro, Y. Miki, K. Endo, A. Wada, H. Nagao, T. Miyake, A. Aoyama, T. Yoneya, K. Kai, and S. Ooe, *FEBS Lett.* **367**, 73–76 (1995).
19. S.N. Ho, H.D. Hunt, R.M. Horton, J.K. Pullen, and L.R. Pease, *Gene* **77**, 51–59 (1989).
20. M. Kubo, K. Murayama, K. Seto, and T. Imanaka, *J. Ferment. Technol.* **66**, 13–17 (1988).
21. M. Kubo and T. Imanaka, *J. Gen. Microbiol.* **134**, 1883–1892 (1988).
22. EP Pat. 0418625A1 (September 6, 1989), R. Marquardt, R. Hilgenfeld, and R. Keller (to Hoechst).
23. Y. Miki, *J. Ferment. Bioeng.* **77**, 457–458 (1994).
24. M. Fujii, M. Takagi, T. Imanaka, and S. Aiba, *J. Bacteriol.* **154**, 831–837 (1983).
25. T.J. Gryczan, S. Contente, and D. Dubnau, *J. Bacteriol.* **134**, 318–329 (1978).
26. M. Yamada, M. Kubo, T. Miyake, R. Sakaguchi, Y. Higo, and T. Imanaka, *Gene* **99**, 109–114 (1991).
27. I.A. Nimmo and G.L. Atkins, *Biochem. J.* **141**, 913–914 (1974).
28. K. Oyama, K. Kihara, and Y. Nonaka, *J. Chem. Soc. Perkin Trans. II*, 356–360 (1981).
29. M.A. Vandeyer, M.P. Weiner, C.J. Hutton, and C.A. Batt, *Gene* **65**, 129–133 (1988).
30. Y. Miki, S. Kidokoro, K. Endo, A. Wada, T. Yoneya, A. Aoyama, K. Kai, T. Miyake, and H. Nagao, *J. Mol. Catal. B Enzym.* **1**, 191–199 (1996).
31. R. Higuchi, B. Krummel, and R. K. Saiki, *Nucleic Acids Res.* **16**, 7351–7367 (1988).
32. G. Vriend and V. Eijsink, *J. Comput.-Aided Mol. Des.* **7**, 367–396 (1993).
33. EP Pat. 0616033A1 (March 17, 1993), H. Nagao, T. Yoneya, T. Miyake, K. Kai, S. Kidokoro, Y. Miki, K. Endo, and A. Wada (to Holland Sweetener Company and Sagami Chemical Research Center).
34. T. Nagayasu, M. Miyanaga, T. Tanaka, T. Sakiyama, and K. Nakanishi, *Biotechnol. Bioeng.* **43**, 1108–1117 (1994).
35. WO Pat. 9516029 (June 15, 1995), Y. Tanaka, Y. Miyake, S. Hanzawa, S. Oe, S. Kidokoro, Y. Miki, K. Endo, and A. Wada (to Sagami Chemical Research Center and Holland Sweetener Company).
36. J. Fauchere and V. Pliska, *Eur. J. Med. Chem. Chim. Ther.* **18**, 369–375 (1983).
37. R. Wolfenden, L. Andersson, P.M. Cullis, and C.C.B. Soughgate, *Biochemistry* **20**, 849–844 (1981).
38. B. Van den Burg, G. Vriend, O.R. Veltman, G. Venema, and V.G.H. Eijsink, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2056–2060 (1998).
39. T. Imanaka, M. Shibazaki, and M. Takagi, *Nature* **324**, 695–697 (1988).
40. O.R. Veltman, G. Vriend, P.J. Middlehoven, B. Van den Burg, G. Venema, and V.G.H. Eijsink, *Protein Eng.* **9**, 1181–1189 (1996).
41. V.G.H. Eijsink, O.R. Veltman, W. Aukema, G. Vriend, and G. Venema, *Struct. Biol.* **2**, 374–379 (1995).
42. H. Hardy, G. Vriend, O.R. Veltman, B. Van der Vinne, G. Venema, and V.G.H. Eijsink, *FEBS Lett.* **317**, 89–92 (1993).

THERMOPHILIC MICROORGANISMS

PAULA M. HICKS
ROBERT M. KELLY
North Carolina State University
Raleigh, North Carolina

KEY WORDS

Archea
Continuous culture
Extreme thermophiles
Hyperthermophiles
Phylogenetic diversity
Pyrococcus
Thermoacidophiles
Thermophiles
Thermotoga

OUTLINE

INTRODUCTION

It has been 25 years since it was first reported by Brock et al. (1) that microbial life was indeed possible above 80 °C and in that time the number and diversity of known thermophilic, or heat loving, microorganisms has continued to expand. The so-called extreme thermophiles, defined here as organisms having an optimal growth temperature of at least 75 °C, can be found in a variety of geothermal sites ranging from terrestrial hot springs to seeps associated with volcanic activity to deep sea hydrothermal vents (2). The most thermophilic of these will be referred to as *hyperthermophiles* (growth at temperatures of 90 °C or above), some of which prefer to grow at temperatures of 100 °C or higher (3). Much has been accomplished in developing protocols for culturing these novel organisms in laboratory settings (4,5). As such, their physiological and metabolic characteristics have been examined (6,7), and numerous highly thermostable proteins have been purified and studied (8,9). Efforts have also begun to address the genetic manipulation of certain extreme thermophiles, not only to probe the intracacies of their replication apparatus, but also to develop cloning and expression systems that are functional at elevated temperatures (10,11).

The driving force to study extremely thermophilic microorganisms is based on both scientific and technological possibilities. The origin of life on earth has been proposed to be traced to this group of microorganisms (12), and prospects for extraterrestrial life have been fortified by the increasing variety of inhabitable environments on this planet, which were long thought to be barren of life. If one accepts the premise that among this group are the most primitive of all known life forms (13), the templates for all existing metabolic strategies and bioenergetic processes can be identified and examined for clues to not-yet-understood cellular phenomena. Fundamental insights into the basis for biomolecular stability, particularly proteins, have been sought through the study of extreme thermophiles (8,14). Encouraged by the incredible utility of DNA polymerases from extreme thermophiles in the polymerase chain reaction (15), commercial interest in enzymes from these organisms is high. Biocatalyst instability often has been a major drawback in many existing applications, and it also diminishes enthusiasm for untried enzymatic routes to new products. Extreme thermophiles represent a fertile source of robust biocatalysts that may change this perspective.

Bioprocessing has proved to be an enabling expertise in the study of extreme thermophiles and will be central to the eventual use of these organisms or their products for commercial purposes. The earliest efforts in this field of inquiry were facilitated by the development of unusual bioreactors and cultivation protocols to provide biomass for characterization (16,17). Imaginative approaches for using extreme thermophiles, ranging from mixed communities of viable cells to thermostable enzymes catalyzing multiple step reactions, will be required in future years. As such, bioprocess engineering and technology will have to accommodate the anticipated application of extreme thermophiles and their constituent biomolecules.

In this review, the physiological and metabolic characteristics of extremely thermophilic microorganisms are discussed with regard to bioprocessing issues that arise in their cultivation. The physiological and phylogenetic diversity of currently known members of this group is covered in addition to efforts to define the untapped microbial content of geothermal locales. In a separate review in this volume, the characteristics of enzymes from extreme thermophiles are examined.

OBTAINING EXTREMELY THERMOPHILIC MICROORGANISMS FROM GEOTHERMAL SITES

A variety of natural biotopes associated with geothermal activity have yielded extremely thermophilic microorgan-

isms (18–20). Samples from these sites (in the form of hot water, sediments, or solid rock or even from intestines of animals inhabitating these sites [21]), inoculated into various types of liquid growth media at temperatures below approximately 120 °C and pH between 0 and approaching 11, can contain one or more culturable species of extreme thermophiles. In fact, it has been demonstrated that samples from geothermal sites can contain any number of microorganisms whose cultivation in a laboratory setting may be difficult at best (22,23). Obtaining samples containing extreme thermophiles from geothermal biotopes can be as simple as drawing liquid into a syringe at the edge of a hot spring, as has been done at Yellowstone National Park (19), or as difficult as having to dive in a submersible several kilometers into the ocean and then expertly manipulate external sampling equipment to procure material that cannot be contaminated by the cold surrounding seawater (3).

Although no connection has been found between the degree of difficulty in sampling or remoteness of location to the isolation of the most unusual organisms, sites that have unique characteristics are continually sought. Table 1 lists some examples of geothermal locations from which extreme thermophiles have been isolated and some characteristics of these sites. As more geothermal locales are investigated, it has become clear that similar organisms may be isolated from locations that have very different physicochemical characteristics and are geographically distant. For example, members of the hyperthermophilic genus *Thermococcus* have been found in shallow marine settings (36), deep sea vents (37), and terrestrial surface waters (32) and petroleum reservoirs of various depths (38). The ubiquity of certain extreme thermophiles may reflect evolution under similar environmental conditions or a common source manifested through global volcanic activity.

Enrichment cultivation of samples taken from geothermal sites to isolate extreme thermophiles follows methodology used for less thermophilic microorganisms, with several key exceptions (18,19,39). For aerobic thermoacidophiles, typically growing optimally at temperatures of approximately 75 °C and pH 2.0, iron pyrite or elemental sulfur may be added to the medium to support chemolithotrophy (1). Many of these organisms grow better in the presence of added complex media such as yeast extract, although the nutritional requirements met by particular complex media are often unclear. Most hyperthermophiles, which grow above 90 °C, are anaerobic and thus require careful attention to ensure that oxygen is eliminated from the media. Fortunately, the reduced solubility of oxygen at such high temperatures somewhat simplifies anaerobic technique to the extent that glove boxes are usually unnecessary. Enrichment media for hyperthermophiles typically contain an array of trace elements, basic salts, oligosaccharides, complex substrates (e.g., yeast extract, tryptone, etc.), and elemental sulfur (18). Adams et al. demonstrated the importance of tungsten as a trace element for several hyperthermophiles that incorporate this metal into specific oxidoreductases (40). For example, a threefold stimulation of cell density levels of one hyperthermophile, *Pyrococcus furiosus*, was noted in the presence of trace amounts of tungsten (41). Although there have been reports of chemically defined growth media for certain hyperthermophiles, such media often result in sig-

Table 1. Examples of Geothermal Sites Yielding Extreme Thermophiles

Site	Location	Geothermal characteristics	Examples of extreme thermophiles found
Yellowstone National Park	Wyoming, Montana, Idaho	Varied; includes alkaline and acid hot springs	*Sulfolobus acidocaldarius* (1)
Vulcano Island	Mediterranean Sea, off western coast of Italy	Geothermally heated sediments and shallow marine hot springs	*Pyrococcus furiosus* (24), *Staphylothermus* sp. (11), *Thermococcus* sp. (25)
Juan de Fuca Ridge	Off coast of Washington, Oregon	Deep sea hydrothermal vents	*Pyrococcus endeavori* (ES4) (26)
Macdonald Sea Mount	Polynesia	Erupting submarine volcano	*Pyrodictium occultum* (27), *Archaeglobus fulgidus* (28), *Pyrococcus furiosus* (24), *Thermococcus celer* (25)
Icelandic hot springs	Various locations in interior Iceland	Varied; includes alkaline and acidic hot springs	*Desulfurococcus mucosus* (29), *D. mobilis* (29)
East Pacific Rise	21 °N, depth 2,610 m	Deep-sea hydrothermal vent	*Methanococcus jannaschii* (30)
Oil reservoirs	3,000 m below the bed of the North Sea	Deep oil-bearing strata	*Archaeoglobus* and *Thermococcus* species (31)
Shallow hot springs	New Zealand (central North Island)	Thermal pools, samples taken from 1-m depth	*Thermococcus* ANI (32)
Solfatara fields	Various terrestrial locations; many in Iceland and Italy	Acidic, sulfur-rich	*Acidianus brierleyi*, *A. infernus* (33), *Thermofilum pendens* (34)
Guaymas Basin	Off the coast of Guaymas, Mexico in Gulf of California	Deep-sea hydrothermal vent	*Methanopyrus kandleri* (35), *Methanococcus* strains, *Pyrococcus* sp., *Archaeoglobus profundus* (18)

nificantly lower growth yields (42). Thermal lability of many medium components is always a concern (43), especially for amino acids, such as glutamine and asparagine (42), certain vitamins, and some monosaccharides. Maltose and cellobiose, however, serve as excellent carbon and energy sources for *P. furiosus* growing at 98 °C, although glucose does not.

Extreme thermophiles can be very difficult to establish in pure, laboratory culture, whether the inoculum is directly from isolation efforts or even from storage. This is especially true for some of the hyperthermophilic, sulfur-reducing chemolithotrophs, such as *Pyrodictium occultum* (44), which grow to very low cell densities and are highly sensitive both to shearing forces and the presence of various metals (Schicho and Kelly, unpublished observation). Problems with culture stability can be encountered; cultures that have transferred successfully for long periods of time will abruptly stop growing and may require reinoculation from stock cultures.

Cultures can be stored in various ways. In some cases, extreme thermophiles will remain viable for years in liquid media. Attempts to freeze-dry cultures for long-term storage have met with mixed success (Rinker and Kelly, unpublished data); this may have something to do with the unusual cell membrane characteristics that members of the archaea possess. Lyophilization has been used successfully for some hyperthermophilic archaea, although there are also reports of limited storage times for organisms stored in this way (45). *P. furiosus* (24) can be stored in dimethyl sulphoxide (DMSO) for extended periods in glass capillary tubes over liquid nitrogen, although plastic cryotubes were found to be too permeable to oxygen and thus ineffective for extended storage (45,46).

The inability to use standard microbiological techniques on solid media has been a limitation in establishing genetic systems for extreme thermophiles. However, plating techniques for *Thermotoga* species (47,48), *Aquifex pyrophilus* (49), *P. furiosus* (50), and *Sulfolobus* strains (51) have been reported. Methodology to carry out mutagenesis experiments for the hyperthermophilic pyrococci has been developed and used to produce a uracil auxotrophic mutant of *Pyrococcus abyssi* (52).

BIODIVERSITY OF EXTREMELY THERMOPHILIC MICROORGANISMS

There are now more than 50 different species of extreme thermophiles that are classified into 25 genera and 11 orders (2); these are summarized in Table 2. The pioneering work of Carl Woese and co-workers, using a 16S rRNA-based phylogenetic tree, helped to distinguish the domain Archaea (which contains most extreme thermophiles) from the domains Bacteria and Eucarya (13). On this basis, new extremely thermophilic isolates can be compared to those already discovered on the basis of 16S phylogeny, in addition to distinguishing them by physiological differences. By using both phylogenetic and physiological yardsticks, a framework for differentiating among the many extreme thermophiles that are known and are newly isolated can be established. This is important not only for scientific reasons, but because it also helps directing attention to a certain species that may contain particular enzymes of interest or has desirable physiological characteristics for application as a whole cell.

Phylogenetic Diversity

Using the sequence of 16S rRNA as a chronometer, evolutionary trees showing the relationships among extreme thermophiles can be developed, and such sequenced-based maps of biodiversity have proved to be valuable tools in studying these microorganisms (23). Figure 1 shows one such representation. The extreme thermophiles include two genera that group with the Bacteria: *Aquifex* and *Thermotoga*. The rest are of the extreme thermophiles are found within one of two archaeal kingdoms: the Crenarchaeota and the Euryarchaeota. It should be pointed out that the Archaea include members that are not thermophilic, such as certain methanogens and the halophiles.

In some cases, it can be difficult to sort out members within a given extremely thermophilic genus based strictly on 16S RNA sequence. Within the genus *Thermococcus*, for example, differences of only a few bases may exist between species that have otherwise distinguishing physiological characteristics (83). Whether these differences are in phenotype or genotype are not completely clear, and this issue will remain unresolved until the genomes of the species being compared are examined. In view of the fact that 16S sequence alone is not sufficient in some cases for phylogenetic differentiation of microorganisms, more detailed bases for such differentiation have been proposed, including using DNA polymerase restriction fragment length polymorphisms (RFLP) (102) and using 16S/23S rRNA spacer regions (103). Of course, the best approach is to combine elements of phylogeny with classical approaches to taxonomy and growth physiology.

Physiological Diversity

The extreme thermophiles are a physiologically diverse group of microorganisms that share a common dependence on high temperatures for growth. Within this group are fermentative anaerobes, methanogens, extreme acidophiles, and to a lesser extent thus far, nitrate reducers and sulfate reducers. Some details on these classifications are provided in this section.

Fermentative Anaerobes. Geothermal sites, both terrestrial and marine, have yielded an array of fermentative microorganisms that use a variety of organic carbon and energy sources and reduce sulfur as a consequence of growth. Most of these fermentative heterotrophs require sulfur for growth, with the exception of some species in the genera *Pyrococcus*, *Thermococcus* and *Thermotoga*. Of these fermentative anaerobes, species of the archaeal genus *Pyrococcus* and of the bacterial genus *Thermotoga* have been the most carefully studied. As such, species from each of these will be discussed in more detail to illustrate the current level of understanding the physiology of extremely thermophilic fermenters.

Pyrococcus furiosus, A Hyperthermophilic Archaeon. Members of the genus *Pyrococcus* typically ferment pep-

tides and saccharides to CO_2, H_2, and fatty acids (24). Alanine is also produced in significant amounts in the absence of sulfur (104). Hydrogen sulfide will be produced in copious amounts, if a polysulfide source (e.g., elemental sulfur) is present in the growth medium (105). *Pyrococcus furiosus*, the type strain in this genus, was first isolated from geothermally heated sediments near Vulcano Island, Italy (24), and has been the subject of a number of physiological, enzymatic, and bioenergetic studies (6,7). *P. furiosus* seems to grow best on α-linked polysaccharides, such as starch and glycogen (106), and the disaccharides maltose (41) and cellobiose (104). Growth on β-linked polysaccharides larger than cellobiose has not been reported. Although *P. furiosus* will grow on pyruvate (107), it will not grow on glucose even though it has been shown to take up and metabolize this substrate (108). Defined media for this organism have been difficult to develop (105), but one such successful attempt has been reported (43). The significance of sulfur in the bioenergetics of this organism is not completely clear. Schicho et al. (109) showed that sulfur addition led to a twofold improvement in maximum growth yield of *P. furiosus* grown in continuous culture on maltose. Ma et al. (110) showed that the hydrogenase from *P. furiosus* also catalyzes sulfide formation from polysulfide, implicating this enzyme in both sulfur and nonsulfur growth. It was later shown that *P. furiosus* produces significant amounts of alanine in the absence of sulfur (104), which is seemingly an energetically inefficient alternative to acetate formation. This organism may use sulfur as a more productive electron acceptor in that alanine formation decreases in its presence. It remains to be seen why *P. furiosus* is capable of growth in the absence of sulfur, which is an unusual feature among hyperthermophilic heterotrophs.

There has been progress in elucidating the central metabolic pathways of *P. furiosus*, and this has served as a basis for examining other heterotrophic hyperthermophiles (6,7). Studies on glycolysis have shown that this microorganism possesses a version of the Embden-Meyerhof-Parnas (EMP) pathway, which has some interesting alterations compared to the one found in bacteria (Fig. 2). For example, this pathway involves ADP-dependent kinases, glucokinase, and phosphofructokinase (7). Although ADP and ATP have similar thermodynamic properties in regard to the free energy released upon hydrolysis, it is not clear why AMP-forming enzymes are involved (7). Another interesting feature in this modified pathway is the presence of a tungsten-based glyceraldehyde-3-phosphate–ferredoxin oxidoreductase (111), which appears to oxidize glyceraldehyde-3-phosphate to 3-phosphoglycerate directly. Yet another apparent difference that appears to be present in all archaea is the direct conversion of acetyl-CoA to acetate by means of an acetyl-CoA synthetase (ADP forming) (112). The transfer of reducing equivalents generated by glycolysis have been traced to several potential sinks. From ferredoxin (113), electrons may go to the formation of H_2, when growth is in the absence of sulfur, or to polysulfides that can be formed upon nucleophilic attack of elemental or colloidal sulfur (105). Ferredoxin seems to supplant the need for large pools of NAD^+/NADH in *P. furiosus*, which may be a strategy to offset the thermal lability of this coenzyme.

P. furiosus is an attractive candidate for enzymological studies because of its rapid growth rate (doubling time of ~40 minutes, in some cases), its ability to grow in the absence of sulfur, and the cell densities that can be achieved (more than 10^9 cells/mL) (43). Thus, biomass yields of over a kilogram have been reported from 600-L fermentations (8). A number of enzymes spanning most known classifications have been purified from *P. furiosus* and characterized; these are shown in Table 3. An effort is underway to sequence the entire genome of this organism, which will a valuable tool in further examining its physiological and metabolic characteristics (141).

***Thermotoga maritima* and *Thermotoga neapolitana*, Hyperthermophilic Bacteria.** The genus *Thermotoga* contains obligately anaerobic bacteria with optimum growth temperatures of up to about 80 °C, with some species capable of growth up to 90 °C (98). Species of *Thermotoga* have been found in high-salt, marine environments and low-salt, terrestrial settings. A sheathlike outer structure envelopes these rodlike organisms and balloons over the ends, hence the designation *toga* (66,98). During the stationary phase, cells may become spherical but remain within the *toga*. Nutritionally, *Thermotoga* species are very diverse and can use peptides, simple sugars, and polysaccharides, such as starch (98), xylan (142), and galactomannan (143), as carbon and energy sources.

Thermotoga maritima (98), the type species, has been studied to the greatest extent in terms of its physiological and metabolic characteristics, and *Thermotoga neapolitana* (99) has also been studied in some detail. *T. maritima* ferments sugars to CO_2, acetate, lactate, and H_2 by a standard EMP pathway (144,145); hydrogen sulfide appears to replace H_2 as a product in the presence of polysulfide. Alanine is also produced from sugar fermentation (146). For *T. neapolitana*, sugar transport was found to involve substrate level phosphorylation (147), which is presumably also the case for *T. maritima* species. *T. maritima* and *T. neapolitana* appear to be inhibited by H_2 produced during fermentation; this inhibition is minimized or avoided in the presence of elemental sulfur or if paired with a complementary extremely thermophilic methanogen (148). Both *T. maritima* and *T. neapolitana* are able to reduce thiosulfate to sulfide, which appears to improve growth rates and yields (146). *T. maritima* has been shown to be motile and migrates at a speed that is proportional to temperature (149).

Because *T. maritima* and *T. neapolitana* can be cultured in the absence of sulfur and are nutritionally diverse, these species have been examined for the presence of an array of enzymes (Table 4). In particular, several glycosyl hydrolases have been characterized from these and other Thermotoga species. These include amylases (176), xylanases (142), mannanases (143), galactosidases (143,152), and glucanases (150). The localization of these activities is interesting because some have been found to be associated with the toga (176).

Methanogens. Biotic methane generation is characteristic of many anaerobic geothermal environments, even at

Table 2. Extremely Thermophilic Microorganisms ($T_{opt} \geq 75$ °C)

Order	Genus	Species	Habitat	DNA G + C	T_{opt} [T_{max}]	Growth mode	Carbon source	Metabolic notes	Ref.
				Archaea					
Pyrodictiales	*Pyrodictium*	*abyssi*	Marine	60	97 [110]	Hetero	Cbh, pep	Fermentative	53
		occultum	Marine	62	105 [110]	Auto/hetero	Cbh, pep, CO_2	Forms pyrite in fermenter	27,54
		brockii	Marine	51	105 [110]	Auto	CO_2	Stimulated by YE	54
	Hyperthermus	*butylicus*	Marine	55.6	98 [108]	Hetero	Pep	Fermentative	55
	Thermodiscus	*maritimus*	Marine	49	88 [98]	Hetero/auto	Complex, CO_2	Fermentative?	56
Sulfolobales	*Acidianus*	*infernus*	Terrestrial	31	88 [95]	Auto	CO_2	Facultative aerobe	33
		ambivalens	Terrestrial	32.7	80 [87]	Auto	CO_2	Facultative aerobe	57
	Sulfolobus	*acidocaldarius*	Terrestrial	39	75 [80]	Auto/hetero	Cbh, Pep, CO_2	S^0, Fe^{2+D}; O_2, Fe^{3+A}	1,58
		solfataricus[a]	Terrestrial	38	87 [87]	Hetero	Cbh	Aerobe	58,59
		islandicus	Terrestrial	—	—	Hetero	Pep, Cbh	Inhibited by virus SIRV	59
		shibatae	Terrestrial	34.6	81 [86]	Auto/hetero	Cbh, pep, CO_2	10-mM phosphate inhibits growth	60
	Metallosphaera	*sedula*	Terrestrial	45	75 [80]	Auto/hetero	CO_2, rich organics	Oxidizes metal sulfides	61
		prunae	Terrestrial	46	75 [80]	Auto/hetero	CO2, Pep, rich organics	H_2S, pyrite, sphalerite, chalcopyrite, H_2^D	62
	Stygiolobus	*azoricus*	Terrestrial	38	80 [89]	Auto	CO_2	$S^0 \rightarrow H_2S$	63
	Desulfurolobus	*ambivalens*	Terrestrial	31.7	81 [87]	Auto	CO_2	Facultative aerobe	64
Desulfurococcales	*Desulfurococcus*	*mucosus*	Terrestrial	51.3	85 [97]	Hetero	Pep	Fermentative	29
		mobilis	Terrestrial	50.8	85 [95]	Hetero	Pep	Fermentative	29
		saccharovorans		50	[97]	Hetero		S^0 respiring	56
		amylolyticus			[97]				65
		S	Marine	52.0	85 [90]	Hetero	Pep	Growth strongly S^0 dependent	66
		SY	Marine	52.4	90 [95]	Hetero	Pep	Growth strongly S^0 dependent	66
	Staphylothermus	*marinus*	Marine	34.5	92 [98]	Hetero	Pep	Requires S^0	67
Thermoproteales	*Pyrobaculum*	*aerophilum*[a]	Marine	52	100 [104]	Auto/hetero	CO_2, Pep, acetate	Dissimilatory nitrate reduction	68
		islandicum	Terrestrial	46	100 [102]	Auto/hetero	Pep, CO_2	S^0, thiosulfate, sulfite, cystine[A]	69
		organotrophum	Terrestrial	46	102 [102]	Hetero	Pep	Obligate organotroph	69
	Thermofilum	*librum*	Terrestrial	57	80 [95]	Hetero		Sulfur respiring	56
		pendens	Terrestrial	57	88 [95]	Hetero	Pep	Sulfur respiring	34
	Thermoproteus	*tenax*	Terrestrial	56	88 [97]	Auto/hetero	Cbh, Pep, CO_2	Requires H_2S and S^0	70
		uzoniensis	Terrestrial	56.5	90 [97]	Hetero	Pep	Fermentative	71
		neutrophilus	Terrestrial		85 [97]	Auto/hetero	CO_2, acetate	Reductive citric acid cycle	56,72
Unclassified		FC89	Marine	31.2	85 [88]	Auto	CO_2	Hydrogenotrophic methanogen	73
		TY	Marine	42.0	88 [90]	Hetero	Pep, Cbh	Fermentative	74
		TYS	Marine	46.0	88 [90]	Hetero	Pep, Cbh	Fermentative	75
		GB-4	Marine	48.1	90 [92]	Hetero	Pep, Cbh	Not S^0 dependent	74
		GB-6	Marine	35.2	90 [92]	Hetero	Pep, Cbh	Not S^0 dependent	74
		NA10	Terrestrial	37.7	75 [85]	Hetero	Cbh	Ferments cellulose	75

Table 2. Extremely Thermophilic Microorganisms ($T_{opt} \geq 75$ °C) (continued)

Order	Genus	Species	Habitat	DNA G + C	T_{opt} [T_{max}]	Growth mode	Carbon source	Metabolic notes	Ref.
	Caldococcus	*litoralis*	Marine	41	88 [100]	Hetero	Pep	Fermentative	76
	Aeropyrum	*Pernix*	Marine	67	90-5 [100]	Hetero	Pep	Strict aerobe	77
Thermococcales	*Pyrococcus*	*furiosus*[a]	Marine	38	100 [103]	Hetero	Cbh, pep	S^0 not required	24
		GB-D	Marine	39.5	95 [103]	Hetero	Pep, aa	S^0, cystine[A]	74
		AL-2	Marine	41	100 [108]	Hetero	Pep	Barophilic	78
		woesei	Marine	38	100 [105]	Hetero	Cbh, Pep	S^0 respiration	79
		abyssi GE5	Marine	44.7	96 [105]	Hetero	Pep, aa	Fermentative	80
		ES4	Marine	55.0	90-9 [110]	Hetero	Cbh, Pep, aa	S^{0A}	26
	Thermococcus	*celer*	Marine	56.6	87 [93]	Hetero	Pep	Fermentative	25
		litoralis A3	Marine	40	85 [98]	Hetero	Cbh, pep	Fermentative	36
		stetteri	Marine	50.2	76 [98]	Hetero	Cbh, pep	Grows on pectin	81
		litoralis NSC	Marine	39.1	88 [98]	Hetero	Cbh, Pep	Produces exopolysaccharide, forms biofilm	82,42
		barossii	Marine	60	82.5	Hetero	Cbh, Pep	—	83
		AN1	Terrestrial	46.2	75–80 [85]	Hetero	Cbh, Pep	S^0, L-cystine^	32
		chitonophagus	Marine	46.5	85 [93]	Hetero	Cbh, Pep	Grows on chitin	84
		peptonophilus OG-1	Marine	52	85 [100]	Hetero	Pep	Rifampicin inhibits growth	85
		profundus	Marine	52.5	80 [90]	Hetero	Cbh, Pep	Requires sulfur	86
		fumicolans ST557	Marine	54.5	85 [103]	Hetero	Pep, aa	Cystine, polysulfide^	87
		ES1	Marine	58.6	82 [91]	Hetero	Rich organics	Fermentative	37
		alcaliphilus AEDII2	Marine	42.4	85 [90]	Hetero	Pep	Fermentative	88
		AL-1	Marine	51	90 [94]	Hetero	Pep	Barophilic	89
		DT1331	Marine	52.3	80 [93]	Hetero	Pep	S^0 required	90
Archaeoglobales	*Archaeoglobus*	*fulgidus* VC16[a]	Marine	46.0	83 [95]	Auto/hetero	Cbh, Pep, CO_2	Oxidizes lactate completely to CO_2	31
		fulgidus Z	Marine	45.0	75–80 [90]	Auto/hetero	Cbh, Pep, CO_2	Sulfate[A], thiosulfate[A], sulfite[A]	28
		profundus	Marine	41.0	82 [90]	Hetero	Cbh, Pep	Mixotrophic	91
Methanopyrales	*Methanopyrus*	*kandleri*	Marine	60	98 [110]	Auto	CO_2	S^0 inhibitory	35
Methanococcales	*Methanococcus*	*jannaschii*[a]	Marine	31.4	85 [91]	Auto	CO_2	Requires selenium	30
		CS-1	Marine	31.4	85 [94]	Auto	CO_2	Formate[D]	92
		AG86	Marine	33	85 [92]	Auto	CO_2	Requires selenite or tungstate	93
		igneus	Marine	31	88 [91]	Auto	CO_2	Inhibited by H_2S	94
Methanobacteriales	*Methanothermus*	*sociabilis*	Terrestrial	33	88 [97]	Auto	CO_2	Ammonia as nitrogen source	95
		fervidus	Terrestrial	33	83 [97]	—	—	—	96
	Methanobacterium	*thermoautotrophicum*[a]	Terrestrial	—	75	Auto	CO_2	Does not use formate	97
					Bacteria				
Aquificiales	*Aquifex*[a]	*pyrophilus*	Marine	40	85 [95]	Auto	CO_2	Tolerates 6% O_2	49
Thermotogales	*Thermotoga*	*maritima*[a]	Marine	46	80 [90]	Hetero	Cbh	Fermentative	98
		neapolitana NSE	Marine	41.3	77 [95]	Hetero	Cbh, Pep	Inhibited by sulfide	99
		sp ZB	Marine	46	[85]	Hetero	Cbh	Fermentative	100
	Thermosipho	*africanus*	Marine		75 [77–80]	Hetero/auto	Pep, CO_2	Requires cysteine	101

Notes: [A] = electron acceptor; [D] = electron donor; Cbh = carbohydrates; Pep = peptides; hetero = heterotroph; auto = autotroph; aa = amino acid.
[a]Genomes corresponding to sequencing projects finished or in progress.

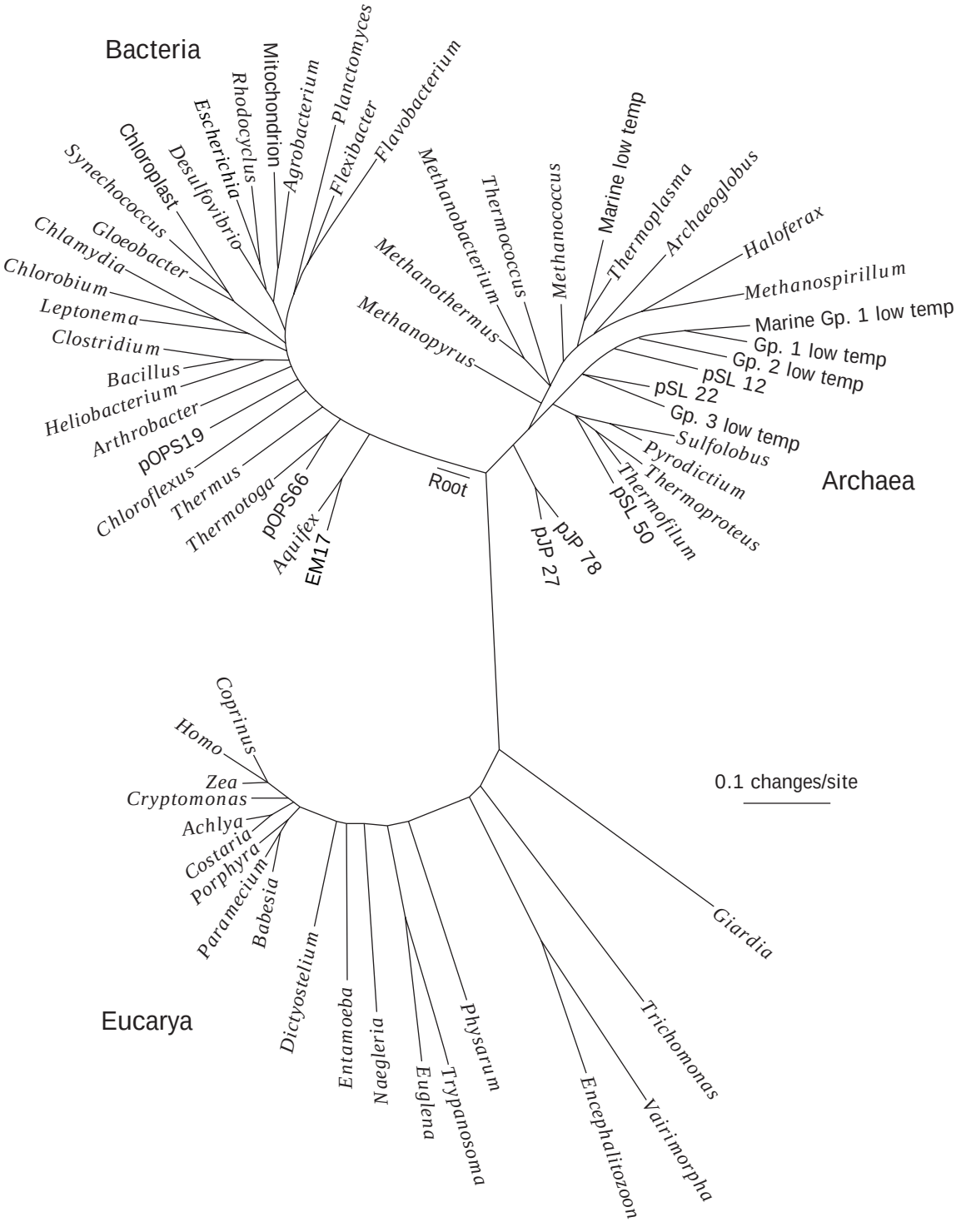


Figure 1. 16S rRNA phylogenetic tree showing placement of three domains of life. *Source*: Reprinted with permission from N.R. Pace, *Science* **276**, 734–740 (1997). Copyright 1997 American Association for the Advancement of Science.

the most extreme temperatures. Methanogenic microorganisms have been isolated from such environments. For example, *Methanococcus jannaschii* was isolated from the base of a deep-sea white-smoker chimney on the East Pacific Rise and found to grow optimally at 85 °C with a doubling time of 26 min (30). Subsequent work with this organism found it to be barophilic with elevated pressures, even higher than those associated with its isolation habitat, promoting both increased growth rate and methane production (177). *M. jannaschii* grows well when

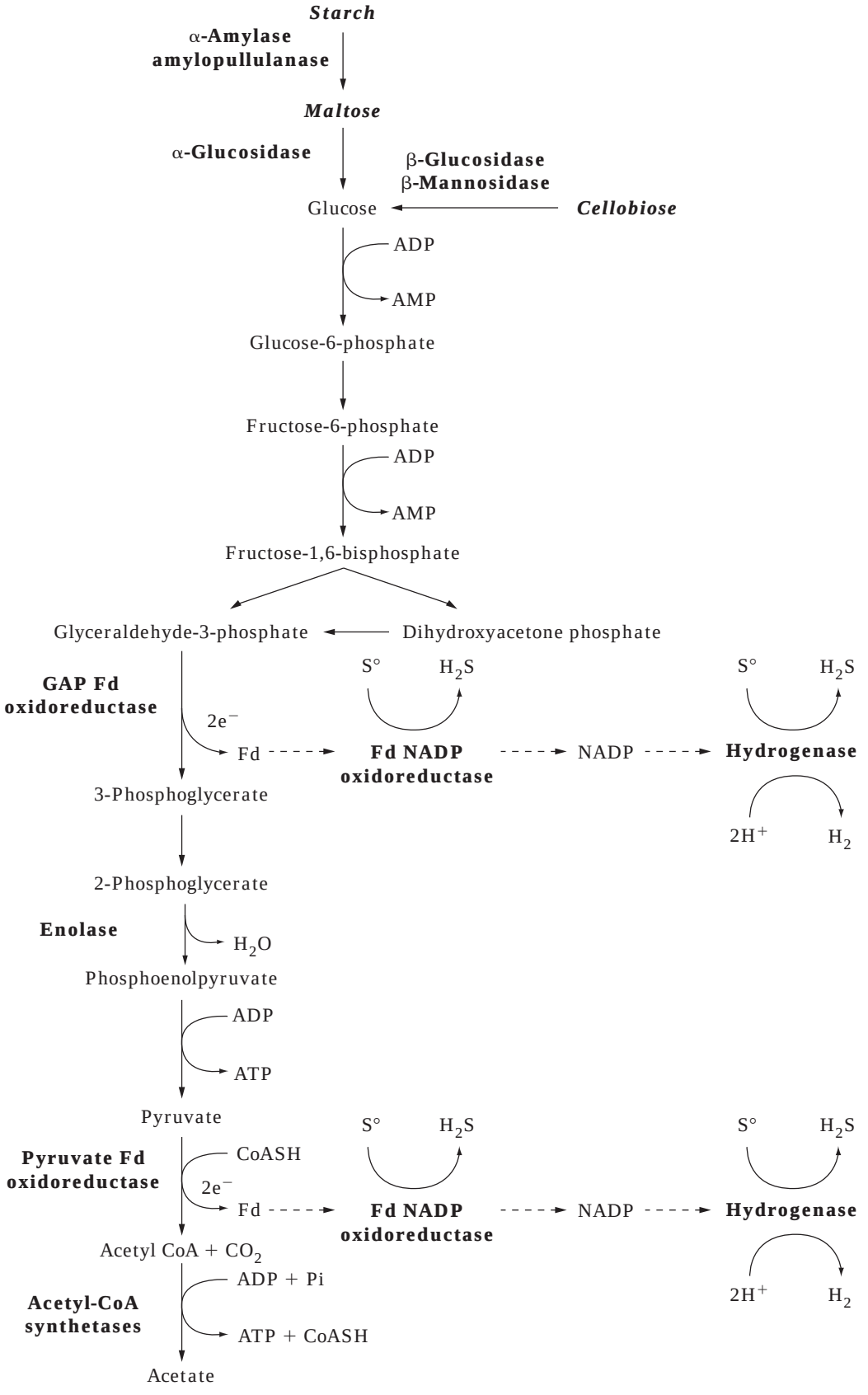


Figure 2. Proposed pathway for glucose oxidation and production of acetate, H_2, and H_2S in *Pyrococcus furiosus*. Growth substrates are shown in bold and italics Fd = ferredoxin. *Source:* Modified from Ref. 111.

Table 3. Enzymes Purified from *Pyrococcus furiosus*

Enzyme	References
Hydrolases	
α-Glucosidase	114
Amylopullulanase	106
Protease PfpI-H (S66)	115
Protease PfpI-C1, PfpI-C2	116,117
α-Amylase	118
β-Glucosidase	119,120
Prolyl oligopeptidase (putative)	121
β-Mannosidase	120
Proteasome	122
Protease pyrolysin	123
Hydrogenases / Redox Proteins	
Hydrogenase/sulfhydrogenase	124
Ferredoxin	113
Rubredoxin	125
Glutaredoxin (putative)	126
Thioredoxin	127
Oxidoreductases	
Aldehyde Fd oxidoreductase	128
Glyceraldehyde-3-Pi Fd oxidoreductase	111
Indolepyruvate Fd oxidoreductase	129
Pyruvate Fd oxidoreductase	130
2-Ketoisovalerate Fd oxidoreductase	131
Ferredoxin NADP oxidoreductase	132
Dehydrogenase	
Glutamate dehydrogenase	133
Isomerases / invertases	
Sucrose α-glucohydrolase	134
Nucleic acid modifying enzymes	
DNA polymerase	135
Synthetases	
Acetyl CoA synthetase (I + II)	136
Citrate synthase (putative)	137
Kinases	
Glucokinase	7
Other	
P1 protein	138
Enolase	139
Aromatic transaminase	140

Fd = ferredoxin.

Table 4. Enzymes purified from Thermotoga species

Enzyme	Species	References
Hydrolases		
endo-β-1,4-Glucanase	*maritina*	150
endo-Glucanase	*neapolitana*	151
exo-β-1,4-Glucanase	*maritima*	150
β-Mannanase	*neapolitana*	143
β-Mannosidase	*neapolitana*	143
α-Galactosidase	*neapolitana*	143
β-Galactosidase	*maritima*	152
β-Glycosidase	FjSS3-B.1	153
β-Xylosidase	FjSS3-B.1	153
α-Arabinofuranosidase	FjSS3-B.1	154
exo-glucanase	FjSS3-B.1	155
endo-1,4-β-D-Xylanases (2)	*thermarum*	156
β-Glucosidase (cloned)	*maritima*	152
endo-Xylanase A (cloned)	*neapolitana*	157
endo-Xylanase	*neapolitana*	151
endo-Xylanase A and B	*maritima*	142
endo-Xylanase A	FjSS3-B.1	158
Enolase	*maritima*	159
Hydrogenases / dehydrogenases		
Hydrogenase	*maritima*	160
D-Glyceraldehyde-3-phosphate	*maritima*	161
Dehydrogenase (cloned)		
L-Lactate dehydrogenase (cloned)	*maritima*	162
Oxidoreductases / redox proteins		
Pyruvate ferredoxin oxidoreductase	*maritima*	163
Ferredoxin	*maritima*	164
Nucleic acid modifying proteins		
RNA polymerase	*maritima*	165
DNA gyrase	*maritima*	165
DNA polymerase	sp.	166
RNAse P (putative from sequence)	*maritima, neapolitana*	167
RuvB (cloned)	*maritima*	168
RecA protein (cloned)	*maritima*	169
Isomerases		
Triosephosphate isomerase	*maritima*	170
Xylose (glucose) isomerase	*maritima, neapolitana*	171,172
Other		
Phosphoglycerate kinase	*maritima*	170
Omp alpha	*maritima*	173
Chemotaxis proteins	*maritima*	174
4-α-Glucanotransferase (cloned)	*maritima*	175

cocultured with hydrogen-producing extreme thermophiles (148,178), which may reflect its ecological role. A novel group of rodlike methanogens have also been isolated from marine geothermal systems that do not appear to belong to the three main methanogenic linkages defined by 16S rRNA phylogeny. *Methanopyrus kandleri*, the type species of the novel methanogenic genus, grows optimally at 98 °C and up to 110 °C and has some unusual cell wall features, including a new type of pseudomurein containing ornithin (35). Both *M. jannaschii* and *M. kandleri* exhibit rapid growth rates under optimal conditions (less than 1 h doubling times) relative to many mesophilic methanogens,

which suggests dynamic interspecies hydrogen transfer in geothermal systems (148). To date, only hydrogenotrophic methanogens have been found among the extreme thermophiles. The isolation of a rapidly growing, acetate-using species would be interesting in view of attempts to accelerate mesophilic and moderately thermophilic anaerobic digestion processes.

Extreme Thermoacidophiles. Brock's pioneering work in investigating microbial life in hot springs in Yellowstone National Park revealed the existence of thermoacidophiles that grew in temperatures in excess of 70 °C and at pH under 2.0 (19). Subsequently, a number of thermoacidophilic genera have been described (Table 2), and organisms that grow at a pH below 1.0 have been reported (179). The metabolism of these organisms is complex and may involve the mixotrophic utilization of both organic carbon and CO_2. Extreme thermoacidophiles typically require some complex organic carbon source, but it is not always clear as to what components of this source are being used to fulfill carbon and energy requirements. Many thermoacidophiles oxidize reduced metal and inorganic sulfur species and have, thus, been evaluated for use in the desulfurization of coal and the recovery of metals from ores (180).

Sulfolobus acidocaldarius was first isolated from Locomotive Spring (pH 2.4, 83 °C) in Yellowstone National Park and subsequently found to be distributed widely in hot, acidic geothermal environments (1). Although the pH optimum is very acidic, *S. acidocaldarius* was also found to grow at a pH up to 5.8. Another member of the order Sulfolobales, *Metallosphaera sedula* grows optimally at pH 2.0 and 75 °C (61). Unlike *S. acidocaldarius*, this organism does not use saccharides as a carbon or energy source, but will grow on peptide-based media supplemented with ferrous iron or elemental sulfur. *M. sedula* will solubilize metals from ores very rapidly, even though its optimal doubling time is more than 5 hours (61).

The identification of an HSP60-like heat shock protein (designated TF 55) from *Sulfolobus shibatae* (181) has led to studies focusing on heat shock in the extreme thermoacidophiles. Although survival rates for *S. shibatae* at supraoptimal temperatures could be increased by thermal acclimation (181), similar experiments with *M. sedula* revealed that thermotolerance was fleeting, and no residual thermotolerance was observed once cultures were returned to normal growth temperature ranges (182). Thermal stress in *M. sedula* appears to result in damage to the proton pumping system or leakage of protons caused by membrane damage. This caused the cytoplasmic pH to drop from around 6 to lethal levels of 4 or below upon extended exposure to supraoptimal temperatures (183).

Other Extremely Thermophilic Physiologies. ***Sulfate Reduction.*** Although the reduction of sulfate to sulfide appears to be a physiological feature of less thermophilic microorganisms, the hyperthermophilic archaeon *Archaeglobus fulgidus*, isolated from Italian geothermal systems, will use sulfate, sulfite, and thiosulfate, but not sulfur, as electron acceptors (31). This organism is a facultative autotroph, growing heterotrophically on a variety of complex media or only on H_2, CO_2, and thiosulfate. *A. fulgidus* grows optimally at 83 °C.

Nitrate Reduction. Nitrate reduction, although rare to this point among the extreme thermophiles, has been observed to occur in both a hyperthermophilic archaeon and a hyperthermophilic bacterium. *Pyrobaculum aerophilum*, isolated from a marine water hole in Ischia, Italy, grows optimally at 100 °C and pH 7.0 by dissimilatory nitrate reduction forming dinitrogen as a product (68). This same organism also grows microaerophically on organic and inorganic compounds by aerobic respiration. In either growth mode, elemental sulfur was inhibitory. Similarly, a hyperthermophilic bacterium, *Aquifex pyrophilus*, isolated from Icelandic hot springs, a strict autotroph, also was found to grow by nitrate reduction or under low concentrations of oxygen (49). Molecular hydrogen, thiosulfate, and elemental sulfur were used as electron donors, forming either sulfuric acid or H_2S, depending on whether growth was aerobic or anaerobic. For *A. pyrophilus*, growth was optimal at 85 °C, with a doubling time of 75 min (49). Of further interest, *A. pyrophilus* represents the most deeply rooted bacterial genus on the 16S rRNA phylogenetic tree (49).

CULTIVATION METHODOLOGY FOR EXTREMELY THERMOPHILIC MICROORGANISMS

The cultivation of extremely thermophilic microorganisms for physiological studies or for generation of biomass has relied on methodology previously developed for mesophiles (138). Modifications have been made to equipment and protocols to accommodate the requirement of high temperature and other extreme conditions, such as pressure and pH. In some cases, specialized bioreactors and approaches have been developed to deal with the unusual bioprocessing conditions. For example, because most of these organisms either reduce (anaerobes) or oxidize (aerobes) sulfur, the highly corrosive conditions created by high levels of hydrogen sulfide or sulfuric acid make materials of construction an important issue. Gaseous products and substrates associated with these organisms also include H_2 and CH_4, so explosive conditions can exist. In any case, many extremely thermophilic microorganisms have been successfully cultivated, in some cases producing in excess of a kilogram of biomass (wet) from a single fermentation (184). Table 5 illustrates some of the reactor types used in the bioprocessing of extreme thermophiles.

Batch Cultivation

Batch cultivation of extreme thermophiles has been done in a variety of ways and on different scales. Hungate tubes or serum bottles are often used for small scale culture, which are made anaerobic by addition of reducing agents and degassing procedures when required. Aerobic thermoacidophiles are often grown at pH of approximately 2.0, making contamination problems minimal, although the cultures must be oxygenated in some way. Details on the types of media used and protocols for establishing cultures have been described elsewhere (4,5,18,89).

Table 5. Bioreactors for Growth of Extreme Thermophiles

Reactor characteristics	Operating mode(s)	Examples of organisms grown	Volume (L)	Operation
Serum bottle	Batch	Most extreme thermophiles	~0.1	Sparge with inert gas before inoculation for anaerobes
Glass resin flask; with 4-5 ports; condenser	Batch, fed-batch, continuous	*Pyrocccus furiosus* *Thermococcus litoralis* *Thermococcus barossii* *Thermotoga neapolitana* *Methanococcus jannaschii* *Metallosphaera sedula*	1–10	Sparge with inert gas H_2S scrubbed with alkali for S^0 growth Can be agitated with magnetic mixing bar or impeller
Glass fermenter; with stainless steel heating mantle; condenser; baffles	Batch, fed-batch	*Pyrocccus furiosus* *Thermococcus litoralis* *Thermococcus barossii* *Thermotoga neapolitana* *Methanococcus jannaschii*	4–18	Sparge with inert gas H_2S scrubbed with alkali for S^0 growth
Ceramic-lined fermenter	Batch	Most extreme thermophiles	50–500	Sparge with inert gas
Stainless steel fermenter	Batch	*Sulfolobus* species *Pyrococcus furiosus* *Thermococcus litoralis* *Thermotoga maritima* *Pyrococcus endeavori* (ES4)	100–600	Sparge with inert gas Mostly non-S^0 growth Carefully clean after S growth to prevent corrosion Can be used for themoacidophile cultivation
High pressure reactor (hyperbaric)	Batch	*Methanococcus jannaschii* *Pyrococcus endeavori* *Sulfolobus acidocaldarius*	~1	Use He for gas phase + substrate gases
High pressure reactor (hydrostatic)	Batch	*Pyrococcus endeavori* (ES4)	~0.5	Flexible gold bag used to contain sample Can be subsampled
Gas-lift reactor	Fed-batch, continuous	*Pyrococcus furiosus*	2	Low hydrodynamic stress
Dialysis membrane reactor	Batch, fed-batch	*Pyrococcus furiosus*	1.2	Both outer and inner chambers are flushed Very high cell yields

Fermenters with modifications for high-temperature operation have be used to culture anaerobic extreme thermophiles. Glass vessels with working volumes up to approximately 20 L (0.7 ft^3) must also have provision for condensing large amounts of water that vaporize because of the high operating temperatures. Although mixing is important to suspend many medium components that are sparingly soluble (e.g., polysaccharides, elemental sulfur), many cultures are sensitive to even modest shearing forces. Also, glass fermenters that have heating mantels or baffle cages that are metallic can be susceptible to corrosion. In some cases, enough metal is leached from stainless steels to inhibit the growth of certain hyperthermophiles.

Growth at High Pressure

Many extremely thermophilic archaea have been isolated from ocean depths of up to 3,000 m (pressures of about 200 atm) and from within temperature gradients that go from 350 °C to 4 °C over distances much less than 1 m. Nonetheless, most isolates are enriched at atmospheric pressure or low hyperbaric pressure (less than 3 atm) and temperatures below 115 °C (78) for experimental convenience. It has been shown that there is a trend toward barotolerance and barophily at pressures greater than those encountered in the native environment (185), and this has spurred efforts to understand the interaction of pressure and temperature as parameters influential to microbial growth (186). It should be noted that special precautions must be used at high pressures to ensure that changes in growth or gas production are actually coupled to increased pressure and not a side effect related to other factors, such as protonation and deprotonation or gas solubility differences (185).

For hydrostatic pressure, a hydraulic fluid must be used to provide overpressures; often distilled water or growth media is sufficient for these purposes (78). Pressure is controlled manually during heating by adjusting hydraulic fluid volume. Hydrostatic pressure bioreactors may be as simple as a syringe or a metal cannister (89). To ensure proper sampling and mixing within the hydrostatic chamber, more sophisticated systems have been developed. A compressible gold bag, fitted with sampling ports and placed in an hydrostatic chamber within a heating mantel, has been used to culture extreme thermophiles (16). This system can be operated semicontinuously by adding and removing material through syringe pumps and rocked to promote mixing. Using such a system, Reysenbach and Deming (78) found that two hyperthermophilic strains showed tolerance to pressures up to 440 atm, and in one case, an improvement in growth rate was observed. In addition, cells grown under pressure appeared to have more

uniform size and were less irregular in shape compared to the same cells grown at low pressure (78).

In a hyperbaric pressure reactor, pressure is transmitted to the culture by the gas phase. Miller et al. (187) developed a hyperbaric pressure reactor that was constructed from a transparent sapphire cell, allowing visual observation of the culture medium during growth experiments. The reactor and a recirculating pump were both contained in an oven to facilitate experiments with extreme thermophiles. One version of this system could be operated at pressures of over 100 bar, with a 10-mL working volume. A magnetically operated pump was used to sparge inert gases through the medium, providing some degree of agitation and promoting thermodynamic equilibrium between vapor and liquid phases. The culture medium was supplied directly from a glove box into the cell, and a gas chromatograph, used to measure volatile products of growth, was directly interfaced to the reactor. The sapphire cell could be replaced with a 167-mL stainless steel vessel, which was then suitable for pressures up to 1,000 atm (177). With the larger volume system, a more powerful pneumatic pump was necessary to achieve desired flow rates to the reactor. Growth of the extremely thermophilic methanogen *M. jannaschii* in the 10-mL hyperbaric bioreactor demonstrated that increases of pressure from 7.8 to 100 bar resulted in accelerated production of biomass and methane and extended the maximum growth temperature from 90 to 92 °C (187). In the stainless steel reactor, the high-temperature limit for methanogenesis was increased to 98 °C by hyperbaric helium pressures up to 750 atm (177).

Continuous Culture

In fed-batch or continuous culture systems used for the growth of extreme thermophiles, the same considerations must be taken as those described for batch systems. Internal parts are typically glass or Teflon-coated to avoid corrosion. Media feedstreams are heat sterilized before use, or in some cases, in situ, given the normally high growth temperatures of these microorganisms. Inorganic feedstreams are filtered through a 0.2-μm porous filter to avoid contamination. Continuous culture systems have been described in some detail for *M. sedula* (182) and *P. furiosus* (41,108). Figure 3 illustrates both a schematic of and an actual continuous culture system.

The gas-lift bioreactor is an alternative type of continuous culture and has been demonstrated for the extreme thermophile *P. furiosus* (50,189). A protocol for setup and operation of an anaerobic gas-lift reactor is described in the work by Sharp and Raven (89). As mentioned previously, some hyperthermophiles are sensitive to high shear rates during growth. Gas-lift systems characteristically have low shearing forces in conjunction with high rates of mass transfer and are well suited for cultivation of extreme thermophiles (189). Such reactors for extreme thermophile growth have been constructed of glass components, with a gas distribution tube for inert gas (filtered with 0.2 μm) supply. The glass vessel can be heated through a heat transfer fluid-containing jacket; an oversized condensor is used to recover water vapor from the exhaust. Both pH and temperature can be controlled continuously. Experiments by Raven et al. (189) have shown that nitrogen is the preferred inert gas to use for sparging, allowing higher cell densities under otherwise equivalent conditions with other gases. The optimal dilution rate for cell mass production of *P. furiosus* was found to be 0.4 h^{-1}, at which point the bioreactor generated more than 1.5 g (wet weight) cells l^{-1} (189). A defined medium was optimized for *P. furiosus*, resulting in reported cell densities of 10^{10} mL^{-1} (50).

Dialysis Membrane Reactor

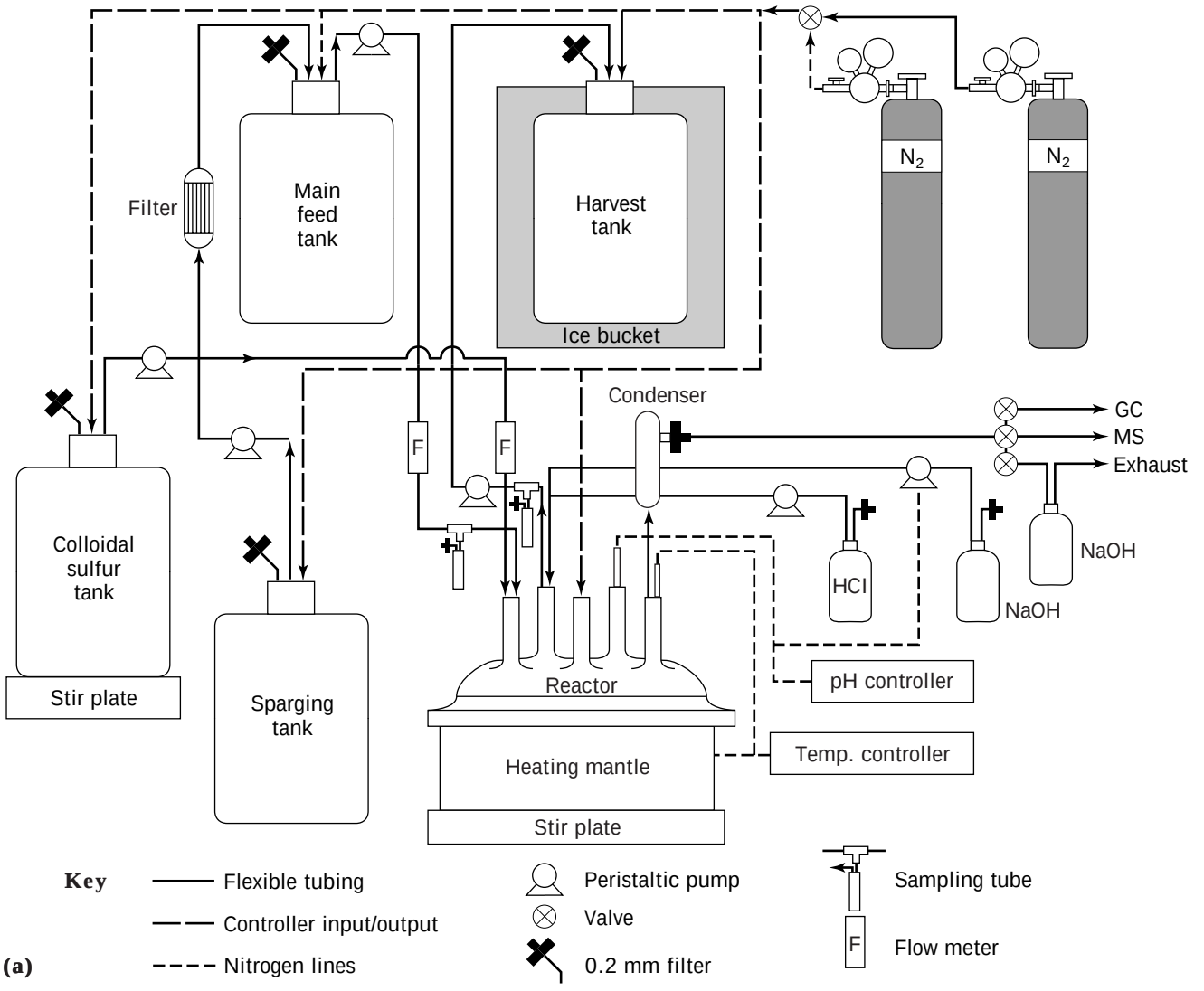
The impetus for using a membrane bioreactor is twofold: carbon and energy sources can be supplied directly to the cells through the membrane, and inhibitory or toxic end products of fermentation can be removed through the membrane, improving cellular growth rate and cell yield. The dialysis reactor described by Krahe et al. (43) is composed of two reactor chambers of unequal volumes, separated by a membrane. Both of these compartments can be sparged and agitated. The chamber that houses the cells is typically several times smaller than the chamber that contains the dialysing medium. Using complex media for both chambers, a 100-fold increase (as compared to serum bottles) in cellular concentration of *P. furiosus* has been reported, along with a 10-fold higher cell density at stationary phase where cell lysis can be significant and lead to loss of whole cells for subsequent recovery steps.

NEW DEVELOPMENTS IN THE ISOLATION AND STUDY OF EXTREMELY THERMOPHILIC MICROORGANISMS

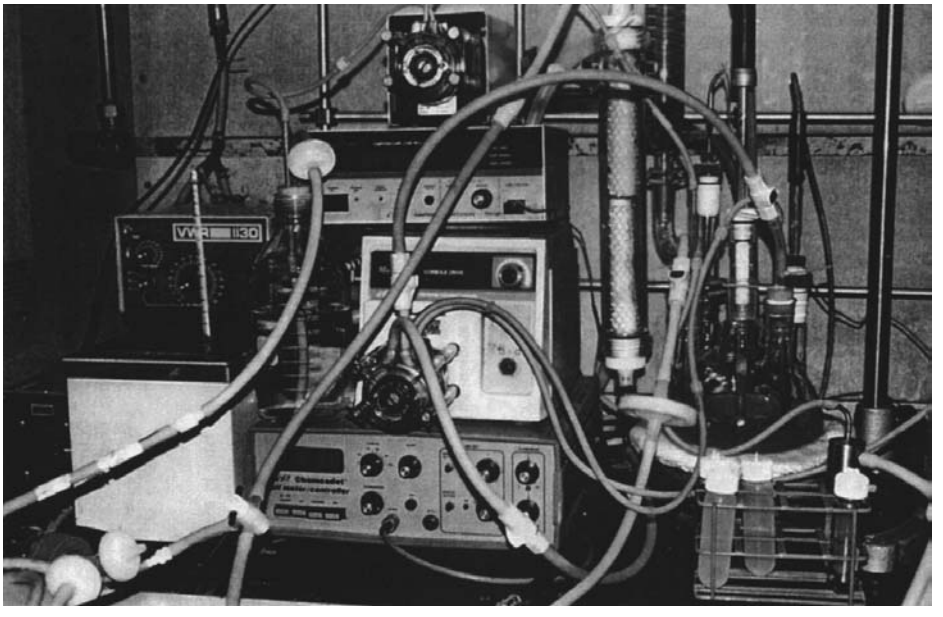
The recognition that the number of extreme thermophiles isolated to date is but a small fraction of those inhabiting geothermal environments has become apparent from analysis of 16S rRNA signatures from geothermal samples (23). Although phylogenetic diversity may not correlate well with physiological diversity, it nonetheless illustrates the untapped potential that exists for new types of microorganisms that have proved difficult to isolate into pure culture.

Molecular phylogeny has been the basis for recent methods to isolate extreme thermophiles. Single cultures of archaea containing 16S rRNA signatures, previously identified from in situ analysis of samples from hot springs, were obtained by the use of optical tweezers (190,191). This method makes use of an inverse microscope and an infrared laser to pull a single cell with a particular 16S rRNA signature into a separate section of a capillary and then put it into cultivation media, where it can be cloned and characterized (192). If effective, this short-circuits the often laborious methods of serial dilution, from which the result might be an organism that has already been isolated and characterized.

Whole genome sequencing efforts have thus far focused to a significant extent on extreme thermophiles. Table 2 indicates which of the genomes of extreme thermophiles are currently being sequenced or are finished. The Institute for Genomic Research (TIGR) maintains a list available on-line of current sequencing projects. In general, extreme thermophiles are an attractive target for sequencing



(a)



(b)

Figure 3. (**a**) Schematic and (**b**) Actual continuous culture setup for growth of hyperthermophiles. *Source*: Schematic modified from Rinker et al. (188). Photo and schematic courtesy of K.D. Rinker, North Carolina State University.

efforts given their evolutionary significance, their commercial potential, and the fact that their genomes are relatively small. The first genome sequence for an extremophile to be completed was that of *M. jannaschii* (193), undoubtedly many more will soon follow. Not only does this present an important new approach to assessing metabolic diversity, but it provides a tool to study the regulation of key pathways within a given organism and a means to study genetic, biochemical, and physiological differences and similarities among this group. The explosive increase in sequence information provides yet another challenge in that mechanisms to use these data must be developed.

SUMMARY

This review illustrates the progress that has been made during the past two decades in exploring the basis for life at extremely high temperatures. Bioprocessing has made important contributions to this effort in developing ways in which extreme thermophiles can be grown for purposes of biomass generation and physiological study. The remaining challenges include developing methods for culturing extreme thermophiles that have, thus far, not been cultured under laboratory conditions; this includes establishing mixed cultures that might be used for particular biotransformations, such as bioremediation. In the future, bioprocessing may even find a role in determining the prospects for life on the subsurface of other planets and moons that may harbor extreme thermophiles not unlike those inhabiting volcanic subsurface environments on earth. The development of bioreactors mimicking such environments will be an important contribution in this regard.

BIBLIOGRAPHY

1. T.D. Brock, K.M. Brock, R.T. Belly, and R.L. Weiss, *Arch. Microbiol.* **84**, 54–68 (1972).
2. K.O. Stetter, *FEMS Microbiol. Rev.* **18**, 149–158 (1996).
3. K.O. Stetter, *ASM News* **61**, 285–290 (1995).
4. F.T. Robb, A.R. Place, K.R. Sowers, H.J. Schreier, S. DasSarma, and E.M. Fleischmann eds., *Archaea: A Laboratory Manual (Thermophiles section)*, Cold Spring Harbor Laboratory Press, New York, 1995, pp. 1–51.
5. K.D. Rinker, C.J. Han, M.W.W. Adams, and R.M. Kelly, in *ASM Manual of Industrial Microbiology and Biotechnology*, 1997, in press.
6. R.M. Kelly and M.W.W. Adams, *Antonie van Leeuwenhoek* **66**, 247–270 (1994).
7. S.W.M. Kengen, A.J.M. Stams, and W.M. de Vos, *FEMS Microbiol. Rev.* **18**, 119–137, (1994).
8. M.W.W. Adams, F.B. Perler, and R.M. Kelly, *Bio / Technol.* **13**, 662–668 (1995).
9. C. Leuschner and G. Antranikian, *World J. Microbiol. Biotechnol.* **11**, 95–114 (1995).
10. C. Aagaard, I. Leviev, R.N. Aravalli, P. Forterre, D. Prieur, and R.A. Garrett, *FEMS Microbiol. Rev.* **18**, 93–104 (1996).
11. W. Zillig, D. Prangishvilli, C. Schleper, M. Elferink, I. Holz, S. Albers, D. Janekovic, and D. Gotz, *FEMS Microbiol. Rev.* **18**, 225–236, (1996).
12. C.R. Woese, *Microbiol. Rev.* **51**, 221–271 (1987).
13. C.R. Woese, O. Kandler, and M.L. Wheelis, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4576–4579 (1990).
14. M.W.W. Adams and R.M. Kelly, *Chem. Eng. News* **73**, 32–42 (1995).
15. F.B. Perler, S. Kumar, and H. Kong, *Adv. Protein Chem.* **48**, 377–435 (1996).
16. R.M. Kelly and J.W. Deming, *Biotechnol. Prog.* **4**, 47–62 (1988).
17. D.S. Clark and R.M. Kelly, *CHEMTECH* **20**, 641–648 (1990).
18. J.A. Baross and J.W. Deming, in D.M. Karl ed., *Deep-Sea Hydrothermal Vents*, CRC Press, Boca Raton, 1995, pp. 169–217.
19. T.D. Brock, *Thermophilic Microorganisms and Life at High Temperatures*, Springer-Verlag, New York, 1978.
20. E. Blöchl, S. Burggraf, G. Fiala, G. Lauerer, G. Huber, R. Huber, R. Rachel, A. Segerer, K.O. Stetter, and P. Völkl, *World J. Microbiol. Biotechnol.* **11**, 9–16.
21. S.C. Cary, M.T. Cottrell, J.L. Stein, F. Comacho, and F. Desbruyeres, *Appl. Environ. Microbiol.* **63**, 1124–1130 (1997).
22. S.M. Barnes, R.E. Fundyga, M.W. Jeffries, and N.R. Pace, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1609–1613.
23. N.R. Pace, *Science* **276**, 734–740 (1997).
24. G. Fiala and K.O. Stetter, *Arch. Microbiol.* **145**, 56–61 (1986).
25. W. Zillig, I. Holz, D. Janekovic, W. Schäfer, and W.D. Reiter, *Syst. Appl. Microbiol.* **4**, 88–94 (1983).
26. R.J. Pledger and J.A. Baross, *J. Gen. Microbiol.* **137**, 203–211 (1991).
27. R. Huber, P. Stoffers, J.L. Cheminee, H.H. Richnow, and K.O. Stetter, *Nature* **345**, 179–181 (1990).
28. G. Zellner, E. Stackebrandt, H. Kneifel, P. Messner, U.B. Sleytr, E.C. De Macario, H.-P. Zabel, and K.O. Stetter, *Syst. Appl. Microbiol.* **11**, 151–160 (1989).
29. W. Zillig, K.O. Stetter, D. Prangishvilli, W. Schäfer, S. Wunderl, D. Janekovic, I. Holz, and P. Palm, *Zbl. Bakt. Hyg., I. Abt. Orig.* **C3**, 304–317 (1982).
30. W.J. Jones, J.A. Leigh, F. Mayer, C.R. Woese, and R.S. Wolfe, *Arch. Microbiol.* **136**, 254–261 (1983).
31. K.O. Stetter, *Syst. Appl. Microbiol.* **10**, 172–173 (1988).
32. K.U. Klages and H.W. Morgan, *Arch. Microbiol.* **162**, 261–266 (1994).
33. A. Segerer, A. Neuner, J.K. Kristjansson, and K.O. Stetter, *Int. J. Syst. Bacteriol.* **36**, 559–564, (1986).
34. W. Zillig, A. Gierl, G. Schreiber, S. Wunderl, D. Janekovic, K.O. Stetter, and H.P. Klenk, *Syst. Appl. Microbiol.* **4**, 79–87 (1983).
35. M. Kurr, R. Huber, H. Konig, H.W. Jannasch, H. Fricke, A. Trincone, J.K. Kristjansson, and K.O. Stetter, *Arch. Microbiol.* **156**, 239–247 (1991).
36. A. Neuner, H.W. Jannasch, S. Belkin, and K.O. Stetter, *Arch. Microbiol.* **153**, 205–207 (1990).
37. R.J. Pledger and J.A. Baross, *Syst. Appl. Microbiol.* **12**, 249–256 (1989).
38. G.S. Grassia, K.M. McLean, P. Glenat, J. Bauld, and A.J. Sheehy, *FEMS Microbiol. Ecol.* **21**, 47–58 (1996).
39. H.W. Jannasch, C.O. Wirsen, and T. Hoaki, in F.T. Robb, A.R. Place, K.R. Sowers, H.J. Schreier, S. DasSarma, and E.M. Fleischmann eds., *Archaea: A Laboratory Manual*,

Cold Spring Harbor Laboratory Press, New York, 1995, pp. 9–13.

40. A. Kletzin, S. Mukund, T.L. Kelley-Crouse, M.K. Chan, D.C. Rees, and M.W.W. Adams, *J. Bacteriol.* **177**, 4817–4819 (1995).
41. R.N. Schicho, L.J. Snowden, S. Mukund, J.B. Park, M.W.W. Adams, and R.M. Kelly, *Arch. Microbiol.* **159**, 380–385 (1993).
42. K.D. Rinker and R.M. Kelly, *Appl. Environ. Microbiol.* **62**, 4478–4485 (1996).
43. M. Krahe, G. Antranikian, and H. Markl, *FEMS Microbiol. Rev.* **18**, 271–285 (1996).
44. A.K. Parameswaran, C.N. Provan, F.J. Sturm, and R.M. Kelly, *Appl. Environ. Microbiol.* **53**, 1690–1693 (1987).
45. H. Connaris, D. Cowan, M. Ruffett, and R.J. Sharp, *Lett. Appl. Microbiol.* **13**, 25–27 (1991).
46. J. Carlsson, G.P.D. Granberg, G.K. Nyberg, and M.B.K. Edlund, *Appl. Environ. Microbiol.* **37**, 383–390 (1979).
47. S.E. Childers, M. Vargas, and K.M. Noll, *Appl. Environ. Microbiol.* **58**, 3949–3953 (1992).
48. O.T. Harriott, R. Huber, K.O. Stetter, P.W. Betts, and K.M. Noll, *J. Bacteriol.* **176**, 2759–2762 (1994).
49. R. Huber, T. Wilharm, D. Huber, A. Trincone, S. Burggraf, H. Konig, R. Rachel, I. Rockinger, H. Fricke, and K.O. Stetter, *Syst. Appl. Microbiol.* **15**, 340–351 (1992).
50. N.D.H. Raven and R.J. Sharp, *FEMS Microbiol. Lett.* **146**, 135–141 (1997).
51. D.W. Grogan, *J. Bacteriol.* **171**, 6710–6719 (1989).
52. L. Watrin and D. Prieur, *Curr. Microbiol.* **33**, 377–382 (1996).
53. R. Pley, J. Schipka, A. Gambacorta, H.W. Jannasch, H.F.R. Rachel, and K.O. Stetter, *Syst. Appl. Microbiol.* **14**, 245–253 (1991).
54. K.O. Stetter, H. Konig, and E. Stackebrandt, *Syst. Appl. Microbiol.* **4**, 535–551 (1983).
55. W. Zillig, I. Holz, D. Janekovic, H.-P. Klenk, E. Imsel, J. Trent, S. Wunderl, V.H. Forjaz, R. Coutinho, and T. Ferreira, *J. Bacteriol.* **172**, 3959–3965 (1990).
56. K.O. Stetter, in T.D. Brock ed., *Thermophiles: General, Molecular and Applied Microbiology*, Wiley, New York, 1986, pp. 40–74.
57. T. Fuchs, H. Huber, S. Burggraf, and K.O. Stetter, *Syst. Appl. Microbiol.* **19**, 56–60 (1996).
58. W. Zillig, K.O. Stetter, S. Wunderl, W. Schulz, H. Priess, and I. Scholz, *Arch. Microbiol.* **125**, 259–269 (1980).
59. W. Zillig, A. Kletzin, C. Schleper, I. Holz, D. Janekovic, J. Hain, M. Lanzendorfer, and J.K. Kristjansson, *Syst. Appl. Microbiol.* **16**, 609–628 (1994).
60. D. Grogan, P. Palm, and W. Zillig, *Arch. Microbiol.* **154**, 594–599 (1990).
61. G. Huber, C. Spinnler, A. Gambacorta, and K.O. Stetter, *Syst. Appl. Microbiol.* **12**, 38–47 (1989).
62. T. Fuchs, H. Huber, K. Teiner, S. Burggraf, and K.O. Stetter, *Syst. Appl. Microbiol.* **18**, 560–566 (1995).
63. A.H. Segerer, A. Trincone, M. Gahrtz, and K.O. Stetter, *J. Bacteriol.* **41**, 495–501 (1991).
64. W. Zillig, S. Yeats, I. Holz, A. Bock, M. Rettenberger, F. Gropp, and G. Simon, *Syst. Appl. Microbiol.* **8**, 197–203 (1986).
65. E.A. Bonch-Osmolovskaya, A.I. Slesarev, M.L. Miroshnichenko, T.P. Svetlichnaya, and V.A. Alexeyev, *Microbiologyia* **57**, 78–85 (1985).
66. H.W. Jannasch, C.O. Wirsen, S.J. Molyneaux, and T.A. Langworthy, *Appl. Environ. Microbiol.* **54**, 1203–1209 (1988).
67. G. Fiala, K.O. Stetter, H.W. Jannasch, T.A. Langworthy, and J. Madon, *Syst. Appl. Microbiol.* **8**, 106–113 (1986).
68. P. Volkl, R. Huber, E. Drobner, R. Rachel, S. Burggraf, A. Trincone, and K.O. Stetter, *Appl. Environ. Microbiol.* **59**, 2918–2926 (1993).
69. R. Huber, J.K. Kristjansson, and K.O. Stetter, *Arch. Microbiol.* **149**, 95–101 (1987).
70. W. Zillig, K.O. Stetter, W. Schäfer, D. Janekovic, S. Wunderl, I. Holz, and P. Palm, *Zbl. Bakt. Hyg., I. Abt. Orig.* **C2**, 205–227 (1981).
71. E.A. Bonch-Osmolovskaya, M.L. Miroshnichenko, N.A. Kostrikina, N.A. Chernych, and G.A. Zavarzin, *Arch. Microbiol.* **154**, 556–559 (1990).
72. S. Schäfer, C. Barkowski, and G. Fuchs, *Arch. Microbiol.* **146**, 301–308 (1986).
73. F. Canganella and W.J. Jones, *Curr. Microbiol.* **28**, 299–306 (1994).
74. H.W. Jannasch, C.O. Wirsen, S.J. Molyneaux, and T.A. Langworthy, *Appl. Environ. Microbiol.* **58**, 3472–3481 (1992).
75. M. Taya, H. Hinoki, T. Yagi, and T. Kobayashi, *Appl. Microbiol. Biotechnol.* **29**, 474–479 (1988).
76. V.A. Svetlichnyi, A.I. Slesarev, T.P. Svetlichnaya, and G.A. Zavarzin, *Mikrobiologiya* **56**, 831–838.
77. Y. Sako, N. Nomura, A. Uchida, Y. Ishida, H. Morii, Y. Koga, T. Hoaki, and T. Maruyama, *Int. J. Syst. Bacteriol.* **46**, 1070–1077 (1996).
78. A.-L. Reysenbach and J.W. Deming, *Appl. Environ. Microbiol.* **57**, 1271–1274 (1991).
79. W. Zillig, I. Holz, H.-P. Klenk, J. Trent, S. Wunderl, D. Janekovic, E. Imsel, and B. Haas, *Syst. Appl. Microbiol.* **9**, 62–70 (1987).
80. G. Erauso, A.-L. Reysenbach, A. Godfroy, J.-R. Meunier, B. Crump, F. Partensky, J.A. Baross, V. Marteinsson, G. Barbier, N.R. Pace, and D. Prieur, *Arch. Microbiol.* **160**, 338–349 (1993).
81. M.L. Miroshnichenko, E.A. Bonch-Osmolovskaya, A. Neuner, N.A. Kostrikina, N.A. Chernych, and V.A. Alekseev, *Syst. Appl. Microbiol.* **12**, 257–262 (1989).
82. S. Belkin and H.W. Jannasch, *Arch. Microbiol.* **141**, 181–186 (1985).
83. G.D. Duffaud, O.B. d'Hennezel, A.S. Peek, A.-L. Reysenbach, and R.M. Kelly, *Syst. Appl. Microbiol.* **21**, 40–49 (1998).
84. R. Huber, J. Stohr, S. Hohenhaus, R. Rachel, S. Burggraf, H.W. Jannasch, and K.O. Stetter, *Arch. Microbiol.* **164**, 255–264 (1995).
85. M. Gonzalez, C. Kato, and K. Horikoshi, *Arch. Microbiol.* **164**, 159–164 (1995).
86. T. Kobayashi, Y.S. Kwak, T. Akiba, T. Kudo, and K. Horikoshi (1994).
87. A. Godfroy, J.-R. Meunier, J. Guezennec, F. Lesongeur, G. Raguenes, A. Rimbault, and G. Barbier, *Int. J. Syst. Bact.* **46**, 1113–1119 (1996).
88. M. Keller, F.-J. Braun, R. Dirmeier, D. Hafenbradl, S. Burggraf, R. Rachel, and K.O. Stetter, *Arch. Microbiol.* **164**, 390–395 (1995).
89. R.J. Sharp and N.D.H. Raven, in P.M. Rhodes and P.F. Stanbury eds., *Applied Microbial Physiology: A Practical Approach*, IRL Press, Oxford, 1997, pp. 23–51.
90. Y.S. Kwak, T. Kobayashi, T. Akiba, K. Horikoshi, and Y.B. Kim, *Biosci. Biotech. Biochem.* **59**, 1666–1669 (1995).

91. S. Burggraf, H.W. Jannasch, B. Nicolaus, and K.O. Stetter, *Syst. Appl. Microbiol.* **13**, 24–28 (1990).
92. W.J. Jones, C.E. Stugard, and H.W. Jannasch, *Arch. Microbiol.* **151**, 314–318 (1989).
93. H. Zhao, A.G. Wood, F. Widdel, and M.P. Bryant, *Arch. Microbiol.* **150**, 178–183 (1988).
94. S. Burggraf, H. Fricke, A. Neuner, J. Kristjansson, P. Rouvier, L. Mandelco, C.R. Woese, and K.O. Stetter, *Syst. Appl. Microbiol.* **13**, 263–269 (1990).
95. G. Lauerer, J.K. Kristjansson, T.A. Langworthy, H. Konig, and K.O. Stetter, *Syst. Appl. Microbiol.* **8** (1986).
96. K.O. Stetter, M. Thomm, J. Winter, G. Wildgruber, H. Huber, W. Zillig, D. Janecovic, H. Konig, P. Palm, and S. Wunderl, *Zbl. Bakt. Hyg., I. Abt. Orig.* **C2**, 166–178 (1981).
97. P.J.L. Derikx, G.A.H. de Jong, H.J.M. Op den Camp, C. van der Drift, L.J.L.D. van Griensven, and G.D. Vogels, *FEMS Microbiol. Ecol.* **62**, 251–258 (1989).
98. R. Huber, T.A. Langworthy, H. Konig, M. Thomm, C.R. Woese, U.B. Sleytr, and K.O. Stetter, *Arch. Microbiol.* **144**, 324–333 (1986).
99. S. Belkin, C.O. Wirsen, and H.W. Jannasch, *Appl. Environ. Microbiol.* **51**, 1180–1185 (1986).
100. G. Zellner and H. Kneifel, *Arch. Microbiol.* **159**, 472–476 (1993).
101. R. Huber, C.R. Woese, T.A. Langworthy, H. Fricke, and K.O. Stetter, *Syst. Appl. Microbiol.* **12**, 32–37 (1989).
102. F.B. Perler, M.W. Southworth, D.G. Wilbur, and D. Wallace, in F.T. Robb, A.R. Place, K.R. Sowers, H.J. Schreier, S. DasSarma, and E.M. Fleischmann eds., *Archaea: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York, 1995, pp. 139–147.
103. J. DiRuggiero, J.H. Tuttle, and F.T. Robb, *Mol. Mar. Biol. Biotechnol.* **4**, 123–127 (1995).
104. S.W.M. Kengen and J.M. Stams, *Arch. Microbiol.* **161**, 168–175 (1994).
105. I.I. Blumentals, M. Itoh, G.J. Olson, and R.M. Kelly, *Appl. Environ. Microbiol.* **56**, 1255–1262 (1990).
106. S.H. Brown, H.R. Costantino, and R.M. Kelly, *Appl. Environ. Microbiol.* **56**, 1985–1991 (1990).
107. T. Schäfer and P. Schonheit, *Arch. Microbiol.* **155**, 366–377 (1991).
108. I.A. Usenko, L.O. Severina, and V.K. Plakunov, *Microbiology* **62**, 272–277 (1993).
109. R.N. Schicho, K. Ma, M.W.W. Adams, and R.M. Kelly, *J. Bacteriol.* **175**, 1823–1830 (1993).
110. K. Ma, R.N. Schicho, R.M. Kelly, and M.W.W. Adams, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5341–5344 (1993).
111. S. Mukund and M.W.W. Adams, *J. Biol. Chem.* **270**, 8389–8405 (1995).
112. T. Schaefer, M. Selig, and P. Schoenheit, *Arch. Microbiol.* **159**, 72–83 (1993).
113. S. Aono, F.O. Bryant, and M.W.W. Adams, *J. Bacteriol.* **171**, 3433–3439 (1989).
114. H.R. Costantino, S.H. Brown, and R.M. Kelly, *J. Bacteriol.* **172**, 3654–3660 (1990).
115. I.I. Blumentals, A.S. Robinson, and R.M. Kelly, *Appl. Environ. Microbiol.* **56**, 1992–1998 (1990).
116. S.B. Halio, I.I. Blumentals, S.A. Short, B.M. Merrill, and R.M. Kelly, *J. Bacteriol.* **178**, 2605–2612 (1996).
117. S.B. Halio, M.W. Bauer, S. Mukund, M.W.W. Adams, and R.M. Kelly, *Appl. Environ. Microbiol.* **63**, 289–295 (1997).
118. K.A. Laderman, B.R. Davis, H.C. Krutzsch, M.S. Lewis, Y.V. Griko, P.L. Privalov, and C.B. Anfinsen, *J. Biol. Chem.* **268**, 24394–24401 (1993).
119. S.W.M. Kengen, E.J. Luesink, A.J.M. Stams, and A.J.B. Zehnder, *Eur. J. Biochem.* **213**, 305–312 (1993).
120. M.W. Bauer, E.J. Bylina, R.V. Swanson, and R.M. Kelly, *J. Biol. Chem.* **271**, 23749–23755, (1996).
121. K.A. Robinson, D.A. Bartley, F.T. Robb, and H.J. Schreier, *Gene* **152**, 103–106 (1995).
122. M.W. Bauer, S.H. Bauer, and R.M. Kelly, *Appl. Environ. Microbiol.* **63**, 1160–1164 (1997).
123. R. Eggen, A. Geerling, J. Watts, and W.M. de Vos, *FEMS Microbiol. Lett.* **71**, 17–20 (1990).
124. F.O. Bryant and M.W.W. Adams, *J. Biol. Chem.* **264**, 5070–5079 (1989).
125. P.R. Blake, J.B. Park, F.O. Bryant, S. Aono, J.K. Magnuson, E. Eccleston, J.B. Howard, M.F. Summers, and M.W.W. Adams, *Biochemistry* **30**, 10885–10895 (1991).
126. A. Guagliardi, D. de Pascale, R. Cannio, V. Nobile, S. Bartolucci, and M. Rossi, *J. Biol. Chem.* **270**, 5748–5755 (1995).
127. C.A. Raia, S. D'Auria, and M. Rossi, *Biosens. Bioelectron.* **10**, 135, (1995).
128. S. Mukund and M.W.W. Adams, *J. Biol. Chem.* **266**, 14208–14216 (1991).
129. X. Mai and M.W.W. Adams, *J. Biol. Chem.* **269**, 16726–16732 (1994).
130. E.T. Smith, J.M. Blamey, and M.W.W. Adams, *Biochemistry* **33**, 1008–1016 (1994).
131. J. Heider, X. Mai, and M.W. Adams, *J. Bacteriol.* **178**, 780–787 (1996).
132. K. Ma and M.W.W. Adams, *J. Bacteriol.* **176**, 6509–6517 (1994).
133. F.T. Robb, J.B. Park, and M.W.W. Adams, *Biochim. Biophys. Acta* **1120**, 267–272 (1992).
134. H.R. Badr, K.A. Sims, and M.W.W. Adams, *Syst. Appl. Microbiol.* **17**, 1–6 (1994).
135. K.S. Lundberg, D.D. Shoemaker, M.W.W. Adams, J.M. Short, J.A. Sorge, and E.J. Mathur, *Gene* **108**, 1–6 (1991).
136. X. Mai and M.W.W. Adams, *J. Bacteriol.* **178**, 5897–5903 (1996).
137. J.M. Muir, R.J. Russell, D.W. Hough, and M.J. Danson, *Protein Eng.* **8**, 583–592 (1995).
138. R.M. Kelly, S.H. Brown, I.I. Blumentals, and M.W.W. Adams, in M.W.W. Adams and R.M. Kelly eds., *Biocatalysis at Extreme Temperatures*, American Chemical Society, Washington, D.C., 1992, pp. 23–41.
139. M.J. Peak, J.G. Peak, F.J. Stevens, J. Blamey, X. Mai, Z.H. Zhou, and M.W. Adams, *Arch. Biochem. Biophys.* **313**, 280–286 (1994).
140. G. Andreotti, M.V. Cubellis, G. Nitti, G. Sannia, X. Mai, M.W. Adams, and G. Marino, *Biochim. Biophys. Acta* **1247**, 90–96 (1995).
141. J.P. Huelsenbeck and B. Rannala, *Science* **276**, 227–232 (1997).
142. C. Winterhalter and W. Liebl, *Appl. Environ. Microbiol.* **61**, 1810–1815 (1995).
143. G.D. Duffaud, C.M. McCutchen, P. Leduc, K.N. Parker, and R.M. Kelly, *Appl. Environ. Microbiol.* **63**, 169–177 (1997).
144. C. Schröder, M. Selig, and P. Schönheit, *Arch. Microbiol.* **161**, 460–470 (1994).

145. M.W.W. Adams, *FEMS Microbiol. Rev.* **15**, 261–277 (1994).

146. G. Ravot, B. Ollivier, M.-L. Fardeau, B.K.C. Patel, K.T. Andrews, M. Magot, and J.-L. Garcia, *Appl. Environ. Microbiol.* **62**, 2657–2659 (1996).

147. M.Y. Galperin, K.M. Noll, A.H. Romano, *Appl. Environ. Microbiol.* **62**, 2915–2918 (1996).

148. V. Muralidharan, K.D. Rinker, I.S. Hirsh, E.J. Bouwer, and R.M. Kelly, *Biotechnol. Bioeng.* **56**, 268–278 (1997).

149. M.F. Gulch, D. Typke, and W. Baumesiter, *J. Bacteriol.* **177**, 5473–5479 (1995).

150. K. Bronnenmeier, A. Kern, W. Liebl, and W.L. Staudenbauer, *Appl. Environ. Microbiol.* **61**, 1399–1407 (1995).

151. J.D. Bok, S.K. Goers, and D.E. Eveleigh, *ACS Symp. Ser.* **566**, 54–65 (1994).

152. J. Gabelsberger, W. Liebl, and K.H. Schleifer, *FEMS Microbiol. Lett.* **109**, 131–138 (1993).

153. L.D. Ruttersmith and R.M. Daniel, *Biochim. Biophys. Acta* **1156**, 167–172 (1993).

154. L.D. Ruttersmith, R.M. Daniel, and H.D. Simpson, *Ann. N.Y. Acad. Sci.* **672**, 137–141 (1992).

155. L.D. Ruttersmith and R.M. Daniel, *J. Biochem.* **277**, 887–890 (1991).

156. A. Sunna, J. Puls, and G. Antranikian, *Biotechnol. Appl. Biochem.* **24**, 177, (1996).

157. V. Zverlov, K. Piotukh, O. Dakhova, G. Velikodvorskaya, and R. Borris, *Microbiol. Biotechnol.* **45**, 245–247 (1996).

158. H.D. Simpson, U.R. Haufler, and R.M. Daniel, *J. Biochem.* **277**, 413–417 (1991).

159. H. Schurig, K. Rutkat, R. Rachel, and R. Jaenicke, *Protein Sci.* **4735**, 228–236 (1995).

160. A. Juszczak, S. Aono, and M.W. Adams, *J. Biol. Chem.* **266**, 13834–13841 (1991).

161. A. Tomschy, R. Glockshuber, and R. Jaenicke, *Eur. J. Biochem.* **214**, 43–50 (1993).

162. R. Ostendorp, W. Liebl, H. Schurig, and R. Jaenicke, *Eur. J. Biochem.* **216**, 709–715 (1993).

163. J.M. Blamey and M.W. Adams, *Biochemistry* **33**, 1000–1007 (1994).

164. J.M. Blamey, S. Mukund, and M.W. Adams, *FEMS Microbiol. Lett.* **121**, 165–169 (1994).

165. P. Palm, C. Schleper, I. Arnold-Ammer, I. Holz, T. Meier, F. Lottspeich, and W. Zillig, *Nucleic Acids Res.* **21**, 4904–4908 (1993).

166. H.D. Simpson, T. Coolbear, and R.M. Daniel, *Ann. N.Y. Acad. Sci.* **613**, 426–428 (1990).

167. J.W. Brown, E.S. Haas, and N.R. Pace, *Nucleic Acids Res.* **21**, 671–679 (1993).

168. J. Tong and J.G. Wetmur, *J. Bacteriol.* **178**, 2695–2700 (1996).

169. J.G. Wetmur, D.M. Wong, B. Ortiz, J. Tong, F. Reichert, and D.H. Gelfand, *J. Biol. Chem.* **269**, 25928–25935 (1994).

170. H. Schurig, N. Beaucamp, R. Ostendorp, R. Jaenicke, E. Adler, and J.R. Knowles, *EMBO J.* **14**, 442–451 (1995).

171. S.H. Brown, C. Sjøholm, and R.M. Kelly, *Biotechnol. Bioeng.* **41**, 878–886 (1993).

172. C. Vieille, J.M. Hess, R.M. Kelly, and J.G. Zeikus, *Appl. Environ. Microbiol.* **61**, 1867–1875 (1995).

173. A.M. Engel, Z. Cejka, A. Lupas, F. Lottspeich, and W. Baumeister, *EMBO J.* **11**, 4369–4378 (1992).

174. R.V. Swanson, M.G. Sanna, and M.I. Simon, *J. Bacteriol.* **178**, 484–489 (1996).

175. W. Liebl, R. Feil, J. Gabelsberger, J. Kellermann, and K.H. Schleifer, *Eur. J. Biochem.* **207**, 81–88 (1992).

176. J. Schumann, A. Wrba, R. Jaenicke, and K.O. Stetter, *FEBS Lett* **282**, 122–126 (1991).

177. J.F. Miller, N.N. Shah, C.M. Nelson, J.M. Ludlow, and D.S. Clark, *Appl. Environ. Microbiol.* **54**, 3039–3042 (1988).

178. E.A. Bonch-Osmolovskaya and K.O. Stetter, *Syst. Appl. Microbiol.* **14**, 205–208 (1991).

179. C. Schleper, G. Puehler, I. Holz, A. Gambacorta, D. Janekovic, U. Santarius, H.-P. Klenk, and W. Zillig, *J. Bacteriol.* **177**, 7050–7059 (1995).

180. G.J. Olson and R.M. Kelly, *Biotechnol. Prog.* **2**, 1–15 (1986).

181. J.D. Trent, J. Osipiuk, and T. Pinkau, *J. Bacteriol.* **172**, 1478–1484 (1990).

182. C.J. Han, S.H. Park, and R.M. Kelly, *Appl. Environ. Microbiol.* **63**, 2391–2396 (1997).

183. T.L. Peeples and R.M. Kelly, *Appl. Environ. Microbiol.* **61**, 2314–2321 (1995).

184. M.W.W. Adams, in F.T. Robb, A.R. Place, K.R. Sowers, H.J. Schreier, S. DasSarma, and E.M. Fleischmann eds., *Archaea: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York, 1995, pp. 47–49.

185. G. Bernhardt, A. Disteche, J. Rainer, B. Koch, H.D. Ludemann, and K.O. Stetter, *Appl. Microbiol. Biotechnol.* **28**, 176–181 (1988).

186. J. Konisky, P.C. Michels, and D.S. Clark, *Appl. Environ. Microbiol.* **61**, 2762–2764 (1995).

187. J.F. Miller, E.L. Almond, N.N. Shah, J.M. Ludlow, J.A. Zollweg, W.B. Streett, S.H. Zinder, and D.S. Clark, *Biotechnol. Bioeng.* **31**, 407–413 (1988).

188. K.D. Rinker, Ph.D. Thesis, North Carolina State University, Raleigh, N.C., 1998.

189. N. Raven, N. Ladwa, D. Cossar, and R. Sharp, *Appl. Microbiol. Biotechnol.* **38**, 263–267 (1992).

190. A. Ashkin, J.M. Dziedzic, and T. Yamane, *Nature* **330**, 769–771 (1987).

191. A. Ashkin and J.M. Dziedzic, *Science* **235**, 1517–1520 (1987).

192. R. Huber, S. Burggraf, T. Mayer, S.M. Barns, P. Rossnagel, and K.O. Stetter, *Nature* **376**, 57–58 (1995).

193. C.J. Bult, O. White, G.J. Olsen, L. Zhou, R.D. Fleischmann, G.G. Sutton, J.A. Blake, L.M. FitzGerald, R.A. Clayton, J.D. Gocayne, A.R. Kerlavage, B.A. Dougherty, J.F. Tomb, M.D. Adams, C.I. Reich, R. Overbeek, E.F. Kirkness, K.G. Weinstock, J.M. Merrick, A. Glodek, J.L. Scott, N.S.M. Geoghagen, and J.C. Venter, *Science* **273**, 1058–1073 (1996).

See also ENZYMES, EXTREMELY THERMOSTABLE.

THERMOPHILIC/THERMOSTABLE ENZYMES.

See ENZYMES, EXTREMELY THERMOSTABLE; THERMOPHILIC MICROORGANISMS.

THREONINE. See AMINO ACIDS, PRODUCTION PROCESSES.

TRANSFER PHENOMENA IN MULTIPHASE SYSTEMS IN MIXING VESSELS

RODICA-VIORICA ROMAN
Chemical Pharmaceutical Research Institute
Iaşi, Romania

KEY WORDS

Biosynthesis of antibiotics
Hydrodynamics
Mass Transfer
Modified agitators

OUTLINE

INTRODUCTION

The dispersion of gases and solids in liquids is one of the most commonly applied processes of chemical engineering, and the fully baffled stirred vessel is one of many types of equipment used for this purpose. The hydrodynamic regime of the impeller controls the mixing performance of the whole reactor, as is reflected in phases of homogenization, the intensity of local energy dissipation distribution in the region of severe shear conditions, the retained gas fraction, bubble size, solid concentration distribution, mass transfer, and heat transfer (1–3). Good understanding of the dispersion mechanism of the impeller is indispensable for a reliable prediction of the behavior of the multiphase system flow in the reactor as a whole.

In multiphase processes carried out in stirred vessels, interphase mass transfer is often an important step and may be rate determining for the overall process. In aerobic fermentation systems, the oxygen transfer from the gas phase to the surface of the microbial cells is of primary importance, especially at high cell densities when microbial growth is likely to be limited by the availability of oxygen in the liquid phase (4–6). For many years, a set of radial flow Rushton disk turbine impellers of roughly one-third of the standard reactor diameter has been considered the optimum design for the multiphase system mixing, including fermentation processes. Mass transfer cannot be improved by increasing the specific power consumption or the aeration rate above certain limits, both of which lead to increased overall costs, but only by a modification of the blades of the impellers or by the development of impellers that allow a better bulk mixing with a minimum power consumption.

It has been reported (7) that a simple modification of the blades of a Rushton turbine through increase in the blade height simultaneously with perforation of the blade surface could significantly diminish power consumption and improve oxygen transfer efficiency in either a non-biological system or in an antibiotic biosynthesis process.

DESCRIPTION OF THE MODIFIED RUSHTON TURBINE AGITATORS

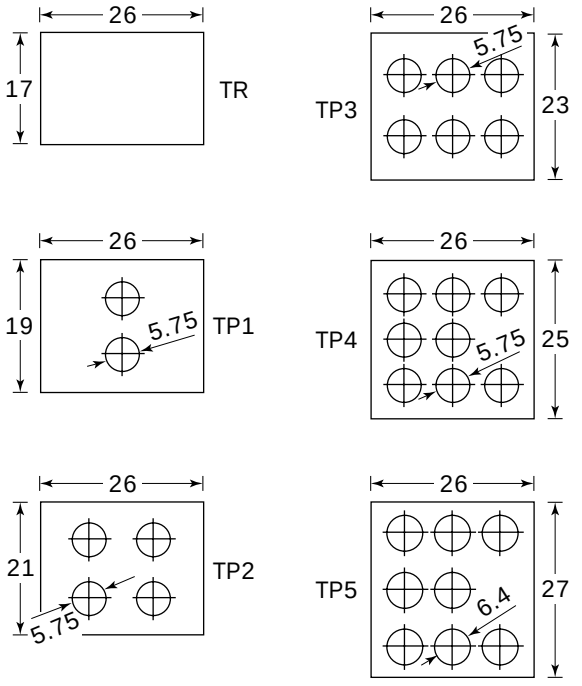
The modified blades were obtained by increase in the conventional (standard) blade height simultaneously with perforation of the blade surface. Five modified blade turbine types (denoted TP1–TP5) were investigated; these are characterized by the fact that the number of perforations and their size varied and that the filled surface of the modified blades is equal to the blade surface of the standard Rushton turbine (TR) (Fig. 1) (7,8).

The application of perforations to the modified blades decreases the pathway between successive shearing actions on the liquid streams, increasing the number of times that the fluid is sheared. This effect has direct implications on local bubble size distribution in a gas–liquid mixture, which develops as a result of the balance between the complementary processes of dispersion and coalescence. The surface fraction of the perforations, defined as the ratio between the area S_G of the perforations on the surface and the full surface area S_C of the blade is in the range of 0 to 0.588. The modified blade turbines do not require any modification in the electrical motor and drive assemblies and are simple to manufacture. The experiments were carried out in a transparent vessel of inside diameter $D = 0.25$ m, with a flat bottom and four symmetrical wall baffles of width $w = 0.1\,D$. Fluid mixing was carried out by conventional and modified blade turbines, positioned singly (1TR, 1TP1–1TP5) or doubly (2TR, 2TP1–2TP5) on the same shaft.

MULTIPHASE SYSTEM HYDRODYNAMICS

Gas–Liquid Dispersing Characteristics

The hydrodynamic regimes were defined by reference to the distribution of the gas phase in the vessel. For an air–water system, the flow regimes were detected through visual observation of the zone close to the outer edge of the TR and TP1–TP5 turbines (8). In the single-turbine configurations, the three classical regimes (9,10) were observed: flooding (F), loading (L), and complete dispersion (CD). For



Agitator	TR	TP1	TP2	TP3	TP4	TP5
S_G/S_C	0	0.117	0.235	0.353	0.470	0.588

Figure 1. Schematic representation of the standard (TR) and modified (TP1–TP5) blades of the Rushton turbine. (*Source*: From Ref. 8, with permission.)

a double-turbine system, two hydrodynamic configurations were defined in the upper zone: ineffective dispersion (DI) and effective dispersion (DE), respectively; the transition between the hydrodynamic regimes corresponds to the rotational speed at which the bubbles have a radial velocity sufficient to reach the vessel wall. The gassed power consumption, P_g, is usually determined by varying the operational parameters: the volumetric flow rate, Q_g, and the impeller speed, N.

The curve representing the ratio P_g/P as related to the gas flow number presents maximum and minimum points when the volumetric gas flow rate is kept constant. Each of these extreme points corresponds to a gas–liquid hydrodynamic regime in the impeller region. The power curves change notably when the power consumption is expressed as a power number. Interesting results were obtained by plotting the gassed power number $(Ne)_g$ against the impeller speed N, when the aeration rate Q'_g, expressed as volumetric flow rate of gas per minute/volume of liquid (vvm) was kept constant. These results are presented in Figures 2 and 3 for mixing systems having one or two turbines, respectively. Table 1 explains the symbols used in Figures 2 and 3.

In the range of low N values, the power number for a single turbine decreased with increasing impeller speed until a certain point: subsequently, it increased suddenly in the case of the standard blade turbine and slowly in the case of the modified blade turbines. The impeller speed at which the transition between these two zones occurred was very close to the transition impeller speed between flooding and loading of the turbine. For two turbines on the same shaft, clear delimitation between the two zones of the impeller speed was not apparent (Fig. 3). The $(Ne)_g$ values decreased slowly with increasing impeller speed until the agitation system exceed the flooding state of the bottom turbine. In the range of large N values, when the gas dispersed uniformly through the entire vessel section, the gassed power number became independent of the impeller speed, as it was a function of aeration rate and geometrical configuration of the mixing system only. The variation in the power number depending on the surface fraction of the perforations, which indicates the geometry of the mixing system, is presented in Figure 4.

The range of power number for an individual Rushton turbine reported in the literature (11) is Ne = 4.8 − 6.3 for $Q'_g = 0$. In our case, the value Ne = 5.5 for a single Rushton turbine is in good agreement with these data. The Ne values for Rushton turbines are dependent on a scale (12), and for a mixing system in which the geometrical symplexes deviate from the standard values, the introduction of a correction factor of the power number is recommended (13,14). The Ne and $(Ne)_g$ values for all the modified blade turbines are smaller than those for the Rushton turbine, as a result of the change in fluid flow around and over the blades and of the minimization of the drag forces associated with the motion of blades such that the energy losses due to form drag are very low. These results imply that it is possible to rotate the modified blade turbines much faster than the standard Rushton turbine at an equal specific power consumption. The TP3 modified blade turbine, with the surface fraction S_G/S_C of the perforations equal to 0.353, has the smallest gassed and ungassed power number values, these being reduced by approximately 50% in comparison with the standard Rushton turbine. The increase in the surface fraction of the perforations over $S_G/S_C = 0.353$ is not justified on the basis of the power consumption.

For double turbines, the power number values are approximately twice those obtained for the single turbines, irrespective of the turbine type, at constant aeration rate Q'_g. This shows that, in the case of two identical turbines, the power consumption P_g is twice that obtained using a single turbine: this is possible because, if Q'_g is kept constant, the volumetric gas flow rate, Q_g, ($m^3\ s^{-1}$) is 60% higher when the liquid volume is higher (Table 1). Also, the power consumption of two TP3 modified turbines is reduced by 50% in comparison with the double standard Rushton turbines. Because the surface of each modified blade increases simultaneously with the S_G/S_C ratio, it is to be expected that a larger gas volume is retained in the ventilated cavities behind the modified blades and dispersed afterward. Therefore, the modified blades have a greatly increased gas-handling capacity: the TP3 turbines handle more than 35% more gas than do the standard Rushton turbines at the same specific ungassed power consumption, before flooding or ineffective dispersion.

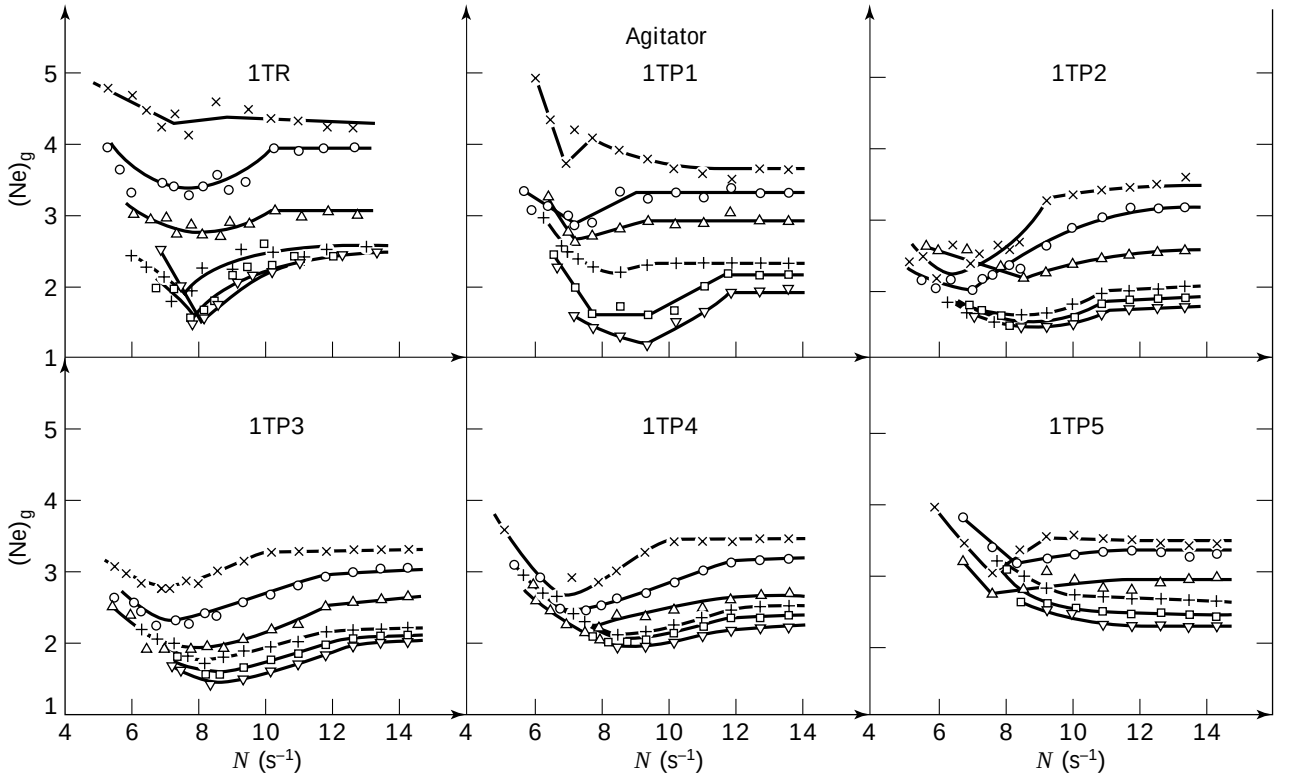


Figure 2. Variation in the gassed power number $(Ne)_g$, depending on the impeller speed (N) at various gas flow rates for single standard and modified Rushton turbines (symbols are defined in Table 1). (*Source:* From Ref. 8, with permission.)

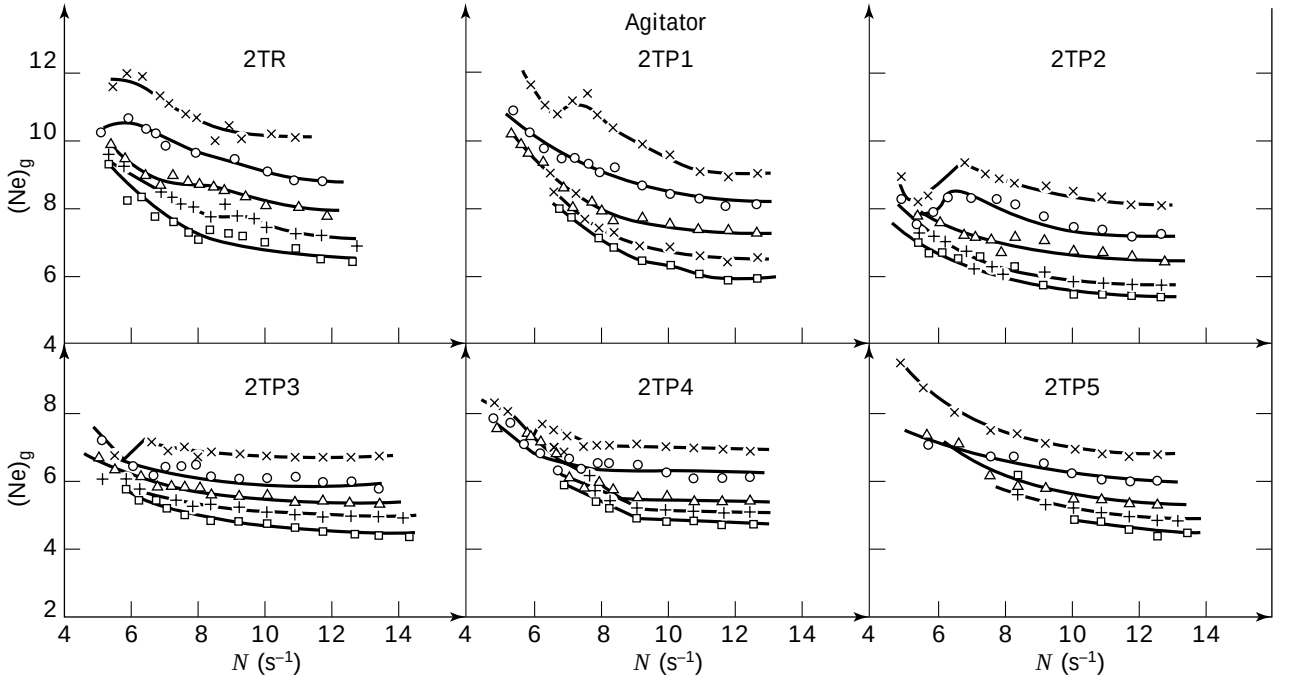


Figure 3. Variation in the gassed power number $(Ne)_g$ depending on the impeller speed (N) at various gas flow rates for double standard and modified Rushton turbines (symbols are defined in Table 1). (*Source*: From Ref. 8, with permission.)

Table 1. Symbols used in Figures 2 and 3

Symbol	Aeration rate, Q'_g (vvm)	Volumetric gas flow rate, $Q_g\ 10^4$ (m^3/s): One turbine ($V = 0.0125\ m^3$)	Two turbines ($V = 0.020\ m^3$)
×	0.30	0.625	1.0
○	0.50	1.042	1.666
△	0.75	1.555	2.500
*	1.00	2.083	3.33
□	1.25	2.611	4.166
▽	1.50	3.125	—

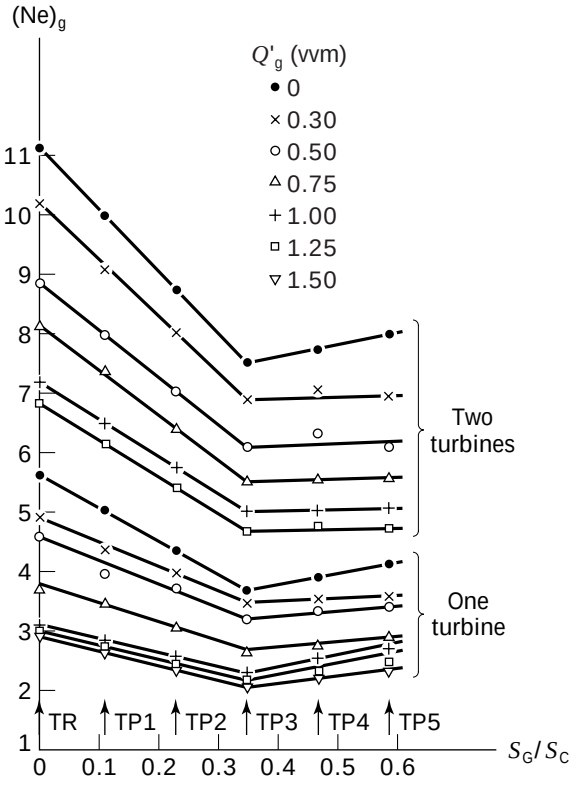


Figure 4. Variation in the power number depending on the surface fraction of the perforations (S_G/S_C) at various aeration rates (Q'_g). (*Source*: From Ref. 8, with permission.)

Suspension of Solid Particles in Liquid

Three suspension states can be defined, as follows (15):

1. Incomplete suspension, in which a part of the solid phase is deposited on the bottom of the vessels
2. Complete suspension, in which all particles are in suspension
3. Homogeneous suspension, in which the particle concentration is uniform throughout the vessel

Many earlier investigations concentrated on the impeller speed for complete suspension, when the maximum surface area available for processing is achieved (15–18). The most widely used criterion to describe solids suspension in mechanically agitated vessels is the just-off-the-bottom suspension condition, first described by Zwietering (19).

In this section, an attempt is made to understand the effects of the physical and rheological properties of the suspensions and number and turbine types on the complete suspension speed and on the power dissipation at complete and homogeneous suspension states of the solid particles (20). The solid particles used are $CaCO_3$, a cation exchange resin (Amberlite IR-120 type), and an anion exchange resin (Amberlite IRA-410 type). A range of density differences $\Delta\rho = (\rho_s - \rho_L)$ between 150 and 1,200 kg/m^3 was examined. The ion exchange resin particle shape is spherical, the particle size distribution being measured by a microscopic method. The change in particle size during the experimental run was negligible. The surface-to-volume (Sauter) mean diameter of the particles, d_p, was calculated for the ion exchange resin particles. In the $CaCO_3$ suspension, the mean diameter of the particles was 15 μm (21). The particle concentration, X, was between 20 and 150 kg/m^3 for each particle type used. The physical properties are given in Table 2. The suspensions used are pseudoplastic in behavior. The rheological properties of suspensions depend on the particle size and the concentration of solids. Rheological data of the suspensions were obtained with a rotational cylinder viscosimeter (Rheotest-2.1) and were fitted according to the power law model (22). The rheological parameters, consistency index, K, and flow behavior index, n, are also presented in Table 2. The definition of complete suspension speed, N_{js}, was taken as the speed at which no particles were visually observed to remain at rest on the vessel bottom for more than 1 or 2 s (19). The last particles to be suspended were located near the center of the vessel base and also behind the baffles. The experimental relationships between complete suspension speed and solid concentrations are shown in Figure 5 for each of the five particle types and for modified and standard turbines positioned singly or doubly on the same shaft.

For the same operation conditions, the N_{js} values for the TP3 modified turbine are slightly higher than those for the standard turbine. Also, Figure 5 shows that, independent of the agitator type, N_{js} actually is increased when two turbines are used, as was also observed by Armenante (23). Although the phenomenon of solid suspension off the vessel bottom is largely dominated by the lower impeller, the presence of two turbines modifies the flow pattern in the vessel.

In all geometrical systems, the N_{js} values are found to increase with solid concentration. This increase of the N_{js} value may result from an increase of probability to find particles at rest on the bottom as the number of particles in suspension increases. It was observed that at N_{js} the concentration of the solid particles in the zone near the bottom is considerably higher than the average value in the vessel. The physical background of the concentration effect is the increased power dissipation by presence of solid particles in the turbulent field (24). The particles can interfere with the turbulent field via two routes. (*1*) Because of their inertia, particles do not follow the motion of the liquid. Consequently, the kinetic energy of a particle

Table 2. Physical and Rheological Properties of Solid Particles

		Physical parameters				Rheological parameters	
Particle type	Symbol	d_p (μm)	ρ_s (kg/m^3)	X (kg/m^3)	ρ_{sp} (kg/m^3)	K (Pa s^n)	n (−)
$CaCO_3$	S1	15	2,200	20	1,026	0.105	0.834
				50	1,065	0.118	0.815
				100	1,130	0.133	0.812
				150	1,195	0.186	0.751
Amberlite IR-120	S2	680	1,350	20	1,008	0.135	0.760
				50	1,020	0.149	0.741
				100	1,040	0.164	0.729
				150	1,060	0.204	0.721
Amberlite IR-120	S3	860	1,350	20	1,008	0.140	0.811
				50	1,020	0.150	0.789
				100	1,040	0.220	0.710
				150	1,060	0.382	0.609
Amberlite IRA-410	S4	600	1,150	20	1,002	0.194	0.730
				50	1,005	0.244	0.686
				100	1,010	0.355	0.611
				150	1,015	0.530	0.570
Amberlite IRA-410	S5	1,000	1,150	20	1,002	0.235	0.662
				50	1,005	0.292	0.651
				100	1,010	0.428	0.568
				150	1,015	0.753	0.553

Source: Ref. 20.

changes continuously. Acceleration of a particle occurs at the expense of energy from the turbulent eddies, whereas deceleration leads to energy dissipation in the form of heat. (*2*) At high solids concentration, the particles will also interfere with each other (collison and/or hindrance). In a similar way, this will lead to increased dissipation of the energy of the turbulent eddies. Thus, for the point of complete suspension to be maintained, a higher solid concentration requires a higher power input to the systems.

Figure 6, where the power consumption expressed as power number for complete suspension condition $(Ne)_{js}$, is plotted against the Reynolds number Re, demonstrates that in the vessel equipped with the modified blade turbines, single or double, the $(Ne)_{js}$ values are approximately 30% lower than those in the standard vessel. In these plots, $(Ne)_{js}$ and Re numbers are calculated using the properties of the suspensions. The $(Ne)_{js}$–Re representation based solely on liquid properties is inadequate (25), because at the highest solids concentration laminar or transition range appear to persist up to very high Reynolds numbers. By using a suitable apparent viscosity for the solid–liquid suspension, it is possible to condense all the data in the adequate range. In our case, the Reynolds number values for complete suspension condition are in the transition range ($Re < 1000$); the $(Ne)_{js}$–Re dependency has a negative slope.

The main reason for the smaller power number of modified blade turbines is the change of suspension flow around and over the blades. The fluid flow around the blades of the standard Rushton turbine was studied by Vant' Riet (26). Dead zones behind each turbine blade are formed. From these dead zones with the stagnation lines, the trailing vortex pair originates as a combined effect of the vortex motion at the rear of the vertical inner edge of the blade and the wrapping of the vortex sheets behind both horizontal edges. In the case of the modified blade turbines, the fluid jets penetrating the perforations, eliminate the stagnation lines and change the conditions of the trailing vortex formation.

When all the observation are combined, the following correlation is evident between complete suspension speed and the physical and rheological properties of the solid particles:

$$N_{js} = c d_p^{0.2} \Delta\rho^{0.45} X^{0.13} K^{0.1} \quad (1)$$

The exponents of d_p, $\Delta\rho$, and X are essentially the same as those reported by Zwietering. The correlation, equation 1, includes the consistency index K as a pseudo-viscosity of the suspensions. The value of the dimensionless constant, c, is specific for each of the geometrical configurations studied. The influence of the turbine type and number of the turbines on N_{js} can be expressed by incorporation of two dimensionless groups, namely $(1 + S_G/S_C)$ and $n_p/6$ (S_G/S_C is the surface fraction of the perforation and n_p is the number of blades of the mixing system). Using nonlinear least squares regression, the following correlation was established:

$$N_{js} = 1.1 d_p^{0.2} \Delta\rho^{0.45} X^{0.13} K^{0.1} (1 + S_G/S_C)^{0.2} (n_p/6)^{0.25} \quad (2)$$

In the turbulent range ($Re > 1000$), when $N > N_{js}$, the particle concentration is uniform throughout the vessel, and power number in solid–liquid system, $(Ne)_s$, is constant.

In the vessel equipped with a standard turbine, the maximum value of the Reynolds number is 1,800; for the modified blade turbine, the maximum Reynolds number is

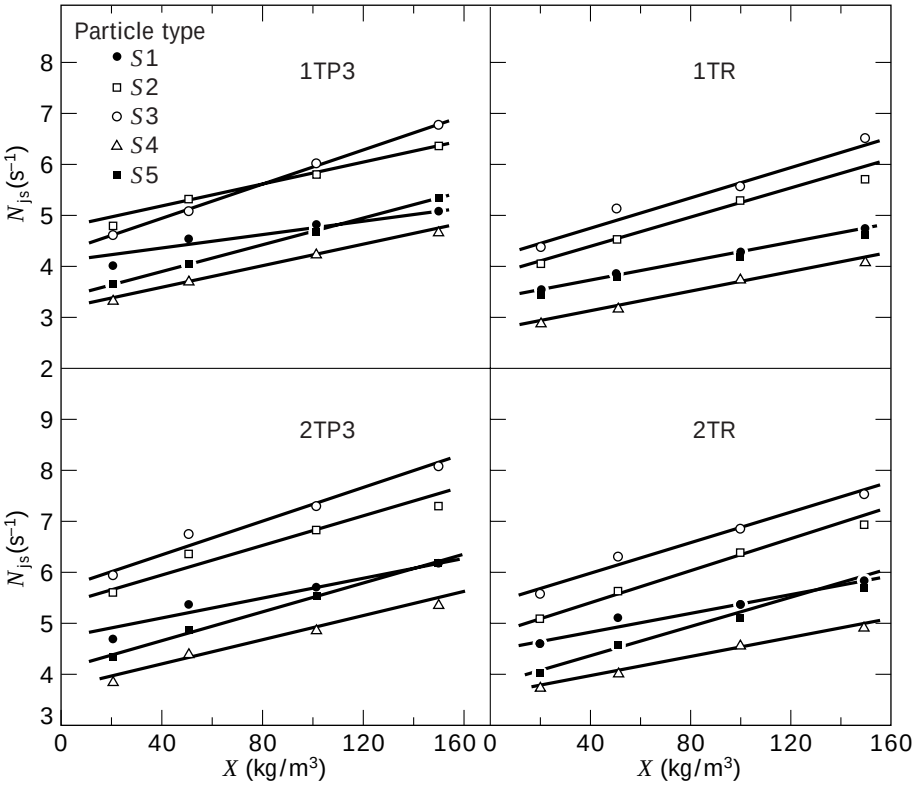


Figure 5. Effect of solid concentration on complete suspension speed for standard and modified Rushton turbines, positioned singly or doubly on the same shaft. S1–S5, particle types (see Table 2). (*Source*: From Ref. 20, with permission.)

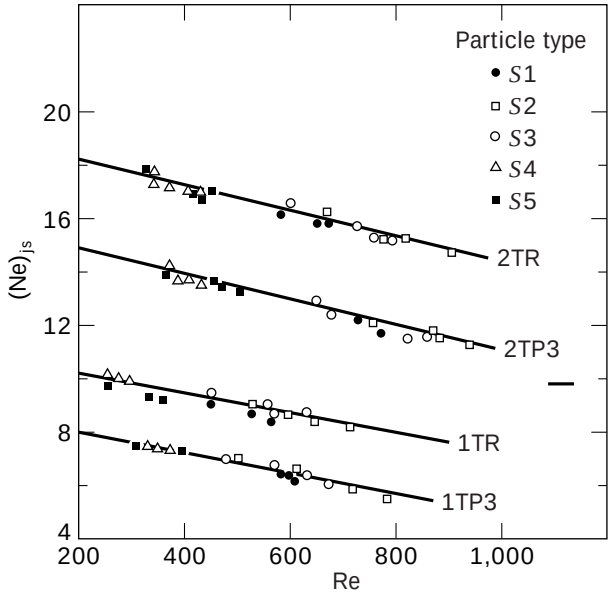


Figure 6. Relation between the power number and Reynolds number (Re) at the complete suspension state, for standard and modified Rushton turbines, single or double. (*Source*: From Ref. 20, with permission.)

2,400. This is possible because the modified blade turbine rotates much faster than the standard turbine at an equal specific power input (11). The relation between the mean values of the power number and the physical properties of the solid particle, in the turbulent range, is

$$(\mathrm{Ne})_3 = 5.6175(n_p/6)(1 + S_G/S_C)^{-1.3867} \cdot (1 + 10d_p^{0.2}\Delta\rho^{-0.45}) \quad (3)$$

The estimated error using equations 2 and 3 lies within ±10%.

Effect of Particles on Gas–Liquid Hydrodynamics

Establishment of the major interactions between gas dispersion and particle suspension mechanisms is a basis for examining in detail the effect of the major variables on the minimum agitator speed for particle suspension under gassed conditions and on the minimum agitator speed for gas dispersion to all parts of the vessel in the three-phase systems.

The manner in which the presence of the particles affects the gas dispersion is inferred from the variation of the relative power consumption P_{gs}/P_s (P_{gs} and P_s are power consumption in gas–liquid–solid and solid–liquid systems, respectively), with gas flow rate, Fl, which is associated with different stages of cavity growth behind the impeller blades. In the curves presented in Figure 7, for various volumetric gas flow rates and concentrations of the S5 type solid particles for 1TR and 1TP3 turbines, the transition to complete gas dispersion is obtained when a minimum occurs in P_{gs}/P_s and the cavity sizes are maximum (27). For the 1TR turbine, the maximum values of P_{gs}/P_s are between 0.5 and 0.3 for all solid concentrations and depend on the volumetric gas flow rate. The 1TP3 turbine shows a behavior similar to that of 1TR turbine but only at $X = 20$ kg/m^3 and $X = 50$ kg/m^3. At $X = 100$ kg/m^3 and $X = 150$ kg/m^3, the minimum values of the P_{gs}/P_s for

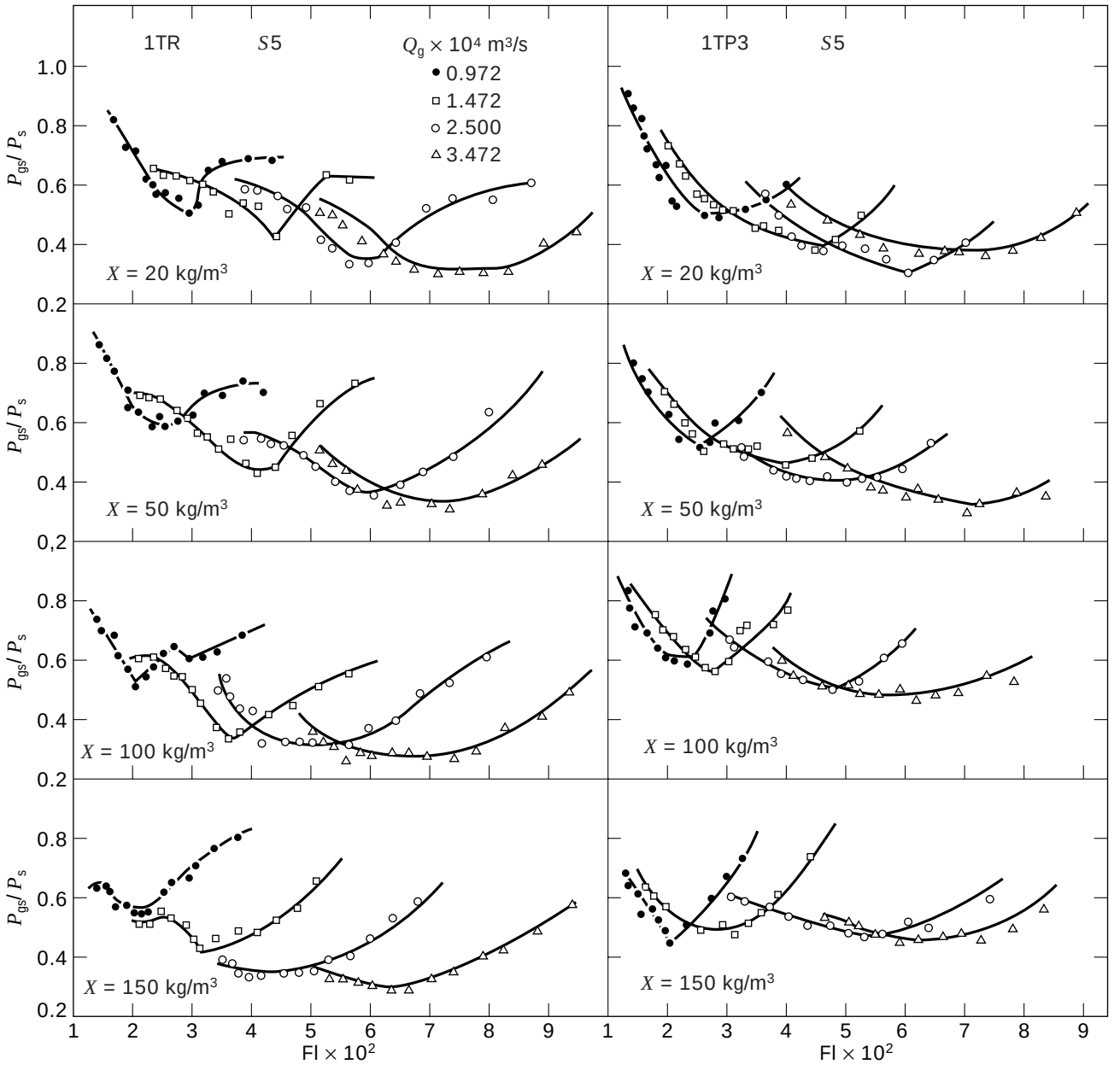


Figure 7. Relation between relative power consumption (P_{gs}/P_s) and gas flow number (Fl) for the S5 particles using single standard and modified turbines. (*Source*: From Ref. 27, with permission.)

1TP3 are approximately 0.5, irrespective of the volumetric gas flow rate. This confirms previous observations (8) as regards the fact that the modified blades have a greatly increased gas-handling capacity.

Effect of particle concentration on power number in a three-phase system, $(Ne)_{gs}$, at constant superficial gas velocity is presented in Figure 8 for the S2 type ion exchange resin, in modified turbines, single or double. A decrease in power number at $N < N_{jsg}$ (jsg is the moment of complete solid suspension) is due to the "false bottom" formed by the particles effectively reducing impeller clearance at low speeds. As the agitator speed N tends to N_{jsg} (as indicated in Fig. 8), the particles were suspended, the false bottom removed, and the power number tended to the common value. The power number generally increased as particle concentration increased before complete particle suspension had been obtained.

In the case of the 1TP3 turbine, the moment of the complete gas dispersion N_{CD} shifts into the range of smaller gas flow rates when the particle concentration increases. For two turbines (2TP3), the variation $(Ne)_{gs} = f$ (Fl) showed the following transition points: the transition to effective dispersion of the gas at the upper impeller (minimum point of the each curve) and also the flooding–loading transition of the bottom impeller (maximum point of each curve). The transition at complete gas dispersion

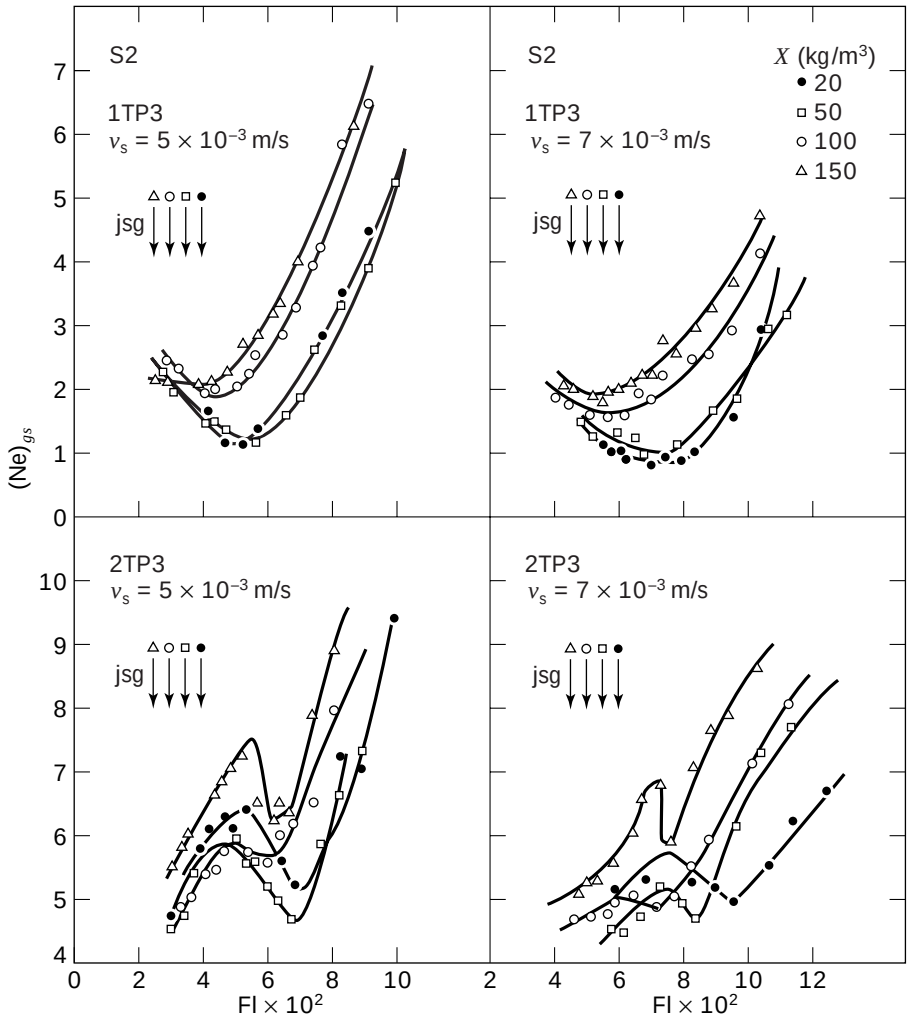


Figure 8. Variation of the power number with gas flow number for the S2 particles using modified turbines, single and double. (*Source*: From Ref. 27, with permission.)

does not show in Figure 8, probably because, once the particles are suspended, the increasing suspension density in the impeller region (and of the power consumption) compensates the decrease in power consumption because of large cavity formation.

The examination of the experimental data in correlation with the Reynolds number shows the following (27):

- For Re $<$ 1,000, mixing is in the transition range, in which the $(Ne)_{gs}$–Re dependence has a negative slope.
- In the range 1,000–1,400 of the Reynolds number, the $(Ne)_{gs}$ values for 1TR attain a minimum point more pronounced for the lower solids concentration. In the case of the 1TP3 turbine, the variation of $(Ne)_{gs}$ with Re is negligible, as a result of the change in fluid flow around and over the perforated blade. The formation of the ventilated cavities behind the modified blades was influenced by the impact of the jets that come through the blade perforations.
- For Re $>$ 1,400, $(Ne)_{gs}$ becomes independent of the Reynolds number, but depends on particle and agitator types. The mixing is in the turbulent range and the phases are uniformly dispersed within the entire vessel section.

The relation between the power number and the physical and rheological properties of solid particles is presented in Figure 9 for the 1TR turbine at various aeration conditions and for 1TP3, 2TP3, and 2TR turbines at constant aeration rate. For the same operation conditions, the modified blade turbines are more efficient, the power number $(Ne)_{gs}$ being reduced by 50% in comparison with the standard Rushton turbine.

Effect of Gas on Solid–Liquid Hydrodynamics

The effect of gas on solid–liquid hydrodynamics in mixing vessels was studied to determine the agitation speed required to just completely suspend all the particles under gassed conditions, N_{jsg} (28). The introduction of a small amount of gas into a solid–liquid system, for the just-suspended condition, causes slight sedimentation such that, for the same impeller speed, the particles were no longer just suspended (i.e., $N < N_{jsg}$). Further increases in the volumetric gas flow rate led to more sedimentation because the pumping capacity and ability of the impeller to circulate fluid and the power input decrease. Any decrease in pumping capacity and power input will decrease all the parameters that cause particle suspension: drag forces, energy dissipation, and associated turbulent eddies. An in-

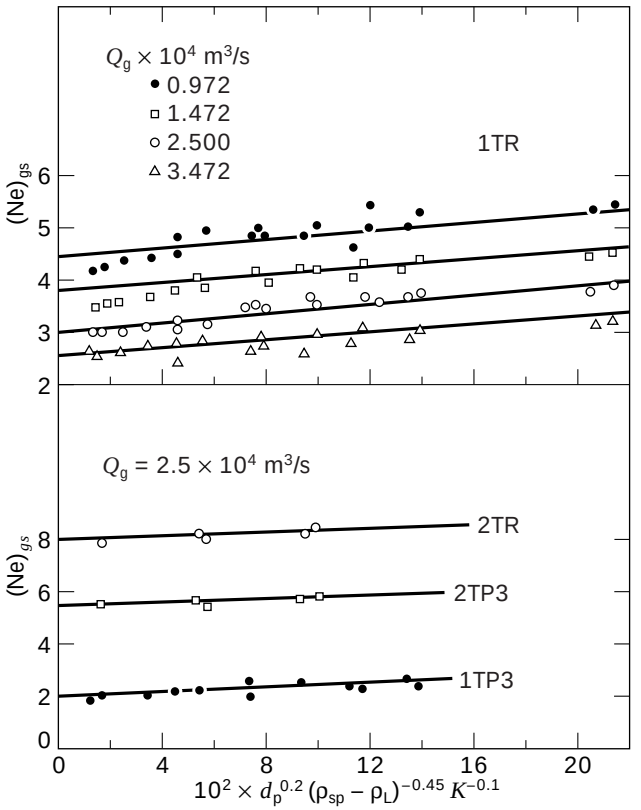


Figure 9. Relation between the power number and the physical and rheological properties of the suspension agitated by various agitator geometries. (*Source*: From Ref. 27, with permission.)

crease of the agitator speed to N_{jsg} ensured complete resuspension. Similarly, under gassed conditions, if $N < N_{jsg}$, a reduction in volumetric gas flow rate gave complete resuspension. In fact, there is a unique combination of gas rate and impeller speed at which the just-suspended condition is achieved and which depends on the agitator type used.

If N_{jsg} is plotted against superficial gas velocity, v_s, the form of the curves is slowly parabolical, for any particle concentration (Fig. 10). At superficial gas velocity of approximately 4×10^{-3} m/s, the slope of the dependence $N_{jsg} = f(v_s)$ decreases, demonstrating that the mechanism of the solid suspension changes. Indeed, this value of the superficial gas velocity is found around a gas flow number Fl = 0.03, when in gas–liquid systems, transition occurs between the clinging cavity regime and the large cavity regime (29,30). The presence of particles has no measurable effect on the formation of the gas-filled cavities.

As showed previously, for complete suspension to be maintained, a higher solid concentration requires a higher power input to the systems. The presence of the gas amplifies the effects mentioned, such that with increasing aeration rate, associated with increasing particle concentration, the agitator speed N_{jsg} must be larger (Fig. 11). The relationship for evaluation of the agitation speed required to just completely suspend all the particles under gassed conditions, N_{jsg}, takes into account the physical and rheological properties of suspensions, the aeration rate, and the equipment configuration (28).

MASS TRANSFER

The mass transfer process from the gas phase to the liquid phase is usually expressed in terms of the volumetric mass transfer coefficient, Kla. The determination methods of the volumetric oxygen transfer coefficient are divided by Rainer (31) into indirect methods, applied without biological system (gassing-out, electrode momentum, sulfite oxidation, carbon dioxide, and glucose oxidase methods), and direct methods (method of gas balance, dynamic method, continuous culture method). The Kla values in mechanically agitated reactors depend on various parameters such as system properties, vessel and stirrer dimensions and arrangement, and operating conditions (32). To examine the behavior of the modified turbines regarding mass transfer, the Kla coefficients were determined in the presence and absence of the solid particles by the sulfite oxidation method.

Gas–Liquid Mass Transfer Characteristics

Figure 12 shows the measured Kla values as a function of gas flow number, Fl, for single and double turbine agitator

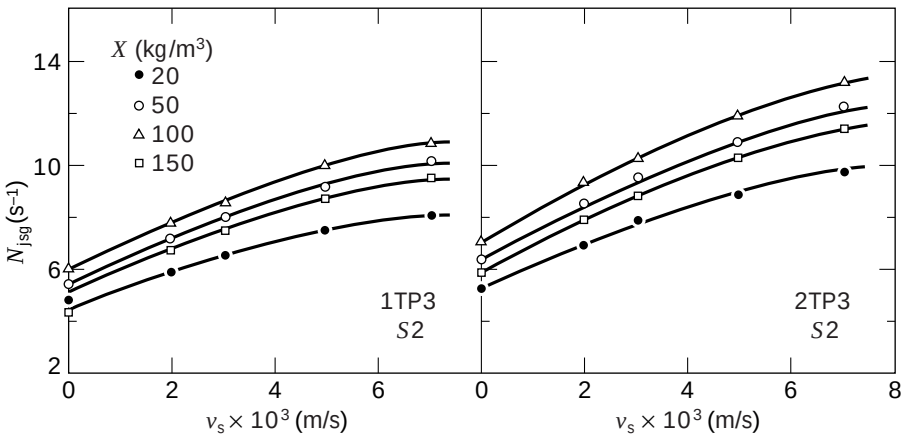


Figure 10. Relation between N_{jsg} and the superficial gas velocity (V_s) for the S2 suspension mixing with modified turbines. (*Source*: From Ref. 28, with permission.)

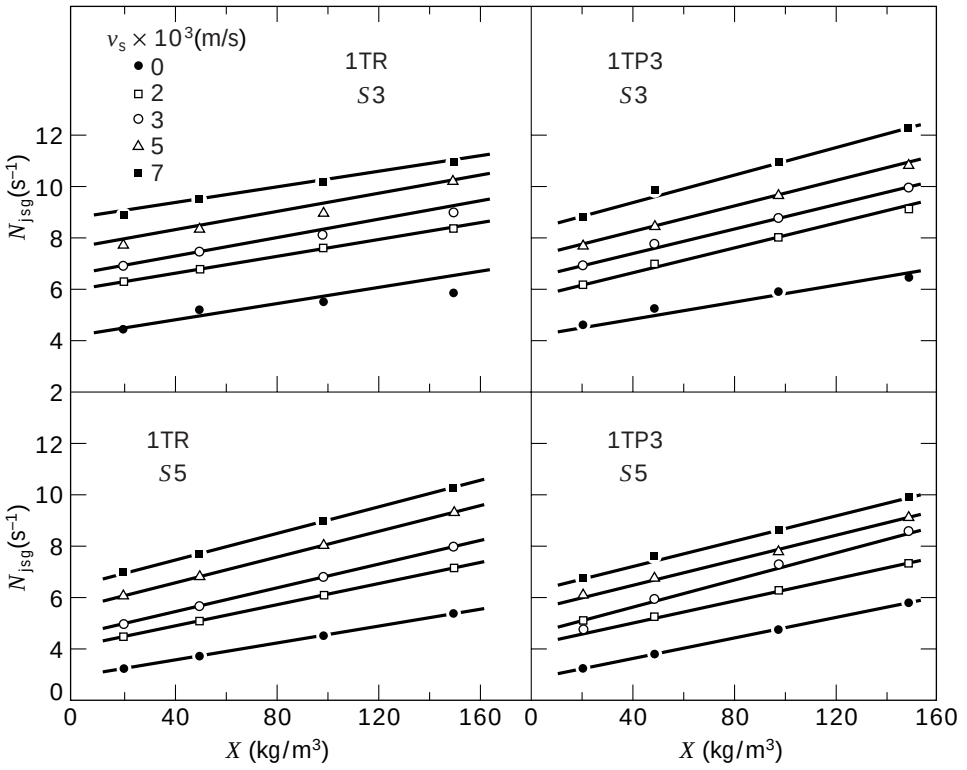


Figure 11. Relation between N_{jsg} (jsg, moment of complete solid suspension) and the solid concentration for various values of the superficial gas velocity. (*Source*: From Ref. 28, with permission.)

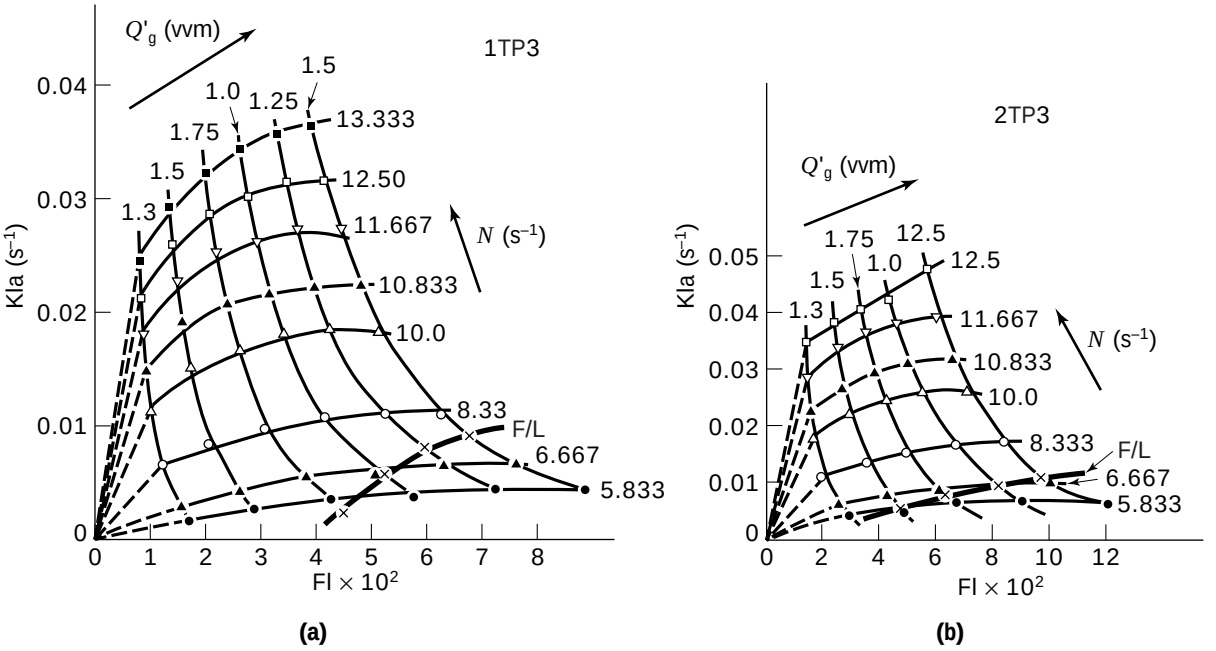


Figure 12. Variation of the volumetric mass transfer coefficient (Kla) depending on the gas flow number: (**a**) single TP3 modified Rushton turbine; (**b**) double TP3 modified Rushton turbines. (*Source*: From Ref. 33, with permission.)

of TP3 type, having as parameter the agitator speed (curves indicated in Fig. 12 by the continual line) or the aeration rate (curves indicated by the dotted line). The same variation, Kla = f(Fl), was also observed for the other modified turbines and the Rushton turbine (33). The dependence of Kla on Fl, when N is constant, is greater at lower gas flow rates (lower gas flow numbers). At higher gas flow numbers, the curves tend to constant Kla values for all agitator speed values and single turbine types (Fig. 12a). In the case of double turbines (Fig. 12b), this tendency of the curves Kla = f(Fl) appears only at values $N < 10\ s^{-1}$. This dependence of Kla on Fl can be explained by the prism of various hydrodynamic regimes, considering the value Fl = 0.03 as the transition point between the vortex-clinging cavity regime and large cavity regime (33). When the Kla values are represented depending on the gas flow number at a constant aeration rate (curves indicated by the dotted line in Fig. 12), these increase slowly once with increasing the rotational speed when the agitator works in flooding regime (below line noted F/L in Fig. 12). When the agitator works in the loading regime (over line F/L), the volumetric oxygen mass transfer coefficient increases considerably with the decreasing gas flow number.

One procedure for evaluating the volumetric oxygen mass transfer coefficient is based on knowledge of the specific power consumption (P_g/V) and the superficial gas velocity (v_s) (34):

$$\mathrm{Kla} = c(P_g/V)^{0.95}(v_s)^{0.67} \tag{4}$$

The relationship between Kla and the product of the variables $(P_g/V)^{0.95}$ $(v_s)^{0.67}$ for all turbine types and their geometrical disposition is linear, the average value of the dimensional constant c being specific for each of the geometrical configurations studied.

The influence of the mixing system geometry on gas–liquid mass transfer is quantified by including, in equation 4, the geometrical elements $(n_p/6)$ and $(1 + S_G/S_C)$, which indicate the number of turbines and their type. The following correlations were obtained for the evaluations of the volumetric mass transfer coefficient:

$$\mathrm{Kla} = 7.2202 \times 10^{-4}(1 + S_G/S_C)^{1.753}(n_p/6)^{-0.126} \cdot (P_g/V)^{0.95})(v_s)^{0.67} \text{ for } S_G/S_C \leq 0.353;\ r = 0.9843 \tag{5}$$

$$\mathrm{Kla} = 2.721 \times 10^{-3}(1 + S_G/S_C)^{-2.73}(n_p/6)^{-0.141} \cdot (P_g/V)^{0.95}(v_s)^{0.67} \text{ for } S_G/S_C > 0.353;\ r = 0.9629 \tag{6}$$

To obtain the same value of the volumetric mass transfer coefficient, the power consumption in the vessel equipped with the TR agitator is higher than that in the vessel equipped with the TP3 agitator. Therefore, the oxygen transfer efficiency E (m^3O_2/W h) (the oxygen volume transferred into the liquid phase for 1 W h of energy consumption for mixing) for one and two TP3 modified Rushton turbines, for any value of the aeration rate (Fig. 13), is more than 30% higher than those obtained with the TR turbines. The TR and TP5 turbines are similar as regards mass transfer efficiency.

Effect of Particles on Gas–Liquid Mass Transfer

It has been observed that the addition of solids has an influence on gas–liquid mass transfer characteristics, even if these particles are not reactive and do not show increased adsorption of dissolved gases. The inert solid particles used in our work are $CaCO_3$ (S1) and cation exchange resin Amberlite IR-120 type (S3) (35). The experiments were carried out in the range of three-phase homogeneous mixing. The presence of the solid phase, with the concentration between 20 and 50 kg/m^3, influences the gas–liquid mass transfer as shown in Figure 14 for the standard (TR) and modified (TP3) turbines, single or double, at various values of superficial gas velocity. All curves show the same general trend.

For $CaCO_3$ particles (S1 solid), a very slight increase of the Kla coefficient is observed until $X = 50$ kg/m^3, while the addition of the cation exchange resin in a proportion of 20 kg/m^3 leads to a pronounced increase of the Kla value, particularly at $N = 11.667\ s^{-1}$ and at higher values of the superficial gas velocity. This trend of the Kla coefficient to increase in the range of small solid concentration was explained by Brehm (36) that at the moderate agitator speed ($N < 15\ s^{-1}$) the solid particles can move independently of fluid elements and break up the gas–liquid interface, which leads to decreasing thickness of the boundary layer. When increasing the solid concentration beyond the values mentioned, the Kla coefficient gradually decreases toward values smaller than those obtained in gas–liquid systems.

Two main factors are responsible for this reduction of the Kla values (37): one factor is the change of the rheological properties of the suspensions and the other is enhanced bubble coalescence, which is caused by turbulence intensity damping. The higher Kla values for mixing with two turbines (Fig. 14) are due to the their higher power consumption for the same operation conditions.

In the three-phase system, the oxygen transfer efficiency, E, decreases with increase of solid concentration, irrespective of the agitator type (Fig. 15), but for mixing with the TP3 modified turbines, the E values are more than 45% higher than those obtained using TR turbines. The experimental data regarding the mass transfer in gas–liquid suspensions demonstrate that the Kla values depend on the physical and rheological properties. The effect of the solid particles on Kla value is given by the group $K^{-0.15}X^{-0.1}$, which includes the consistency index K as a pseudo-viscosity of the suspensions. Therefore:

$$\mathrm{Kla} = 7.2202 \times 10^{-4}(1 + S_G/S_C)^{1.753}(n_p/6)^{-0.126} \cdot (P_g/V)^{0.95}(v_s)^{0.67}K^{-0.15}X^{-0.1} \tag{7}$$

Equation 7 describes the experimental data with an average error of 20%.

PERFORMANCE OF THE MODIFIED AGITATORS IN BIOSYNTHESIS OF ANTIBIOTICS

The transport phenomena in submerged aerobic fermentation involve mixing and mass and heat transfer in three-phase systems consisting of gas bubbles and suspended microorganisms in a liquid phase and are strongly influenced by the rheological properties of the medium used. The rhe-

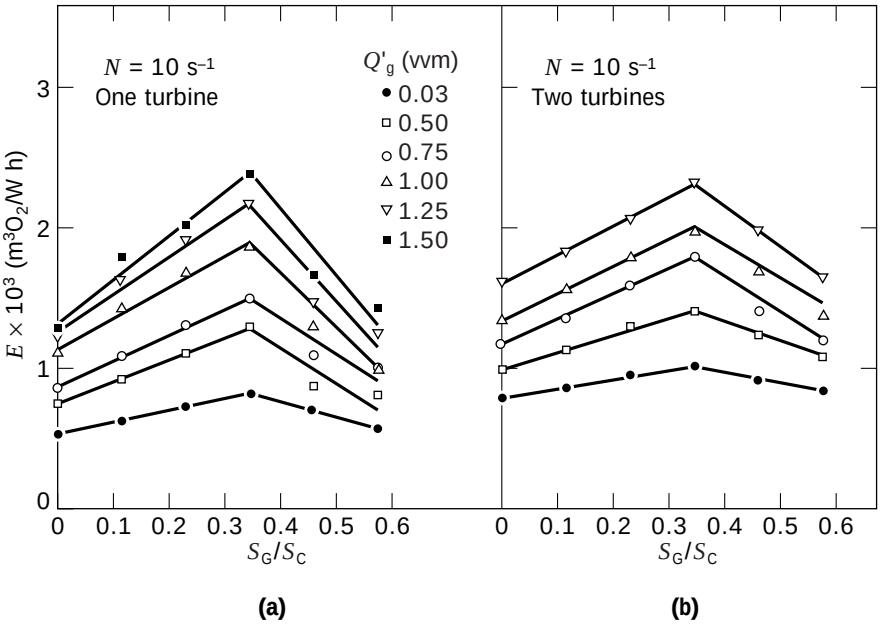


Figure 13. Variation of the oxygen transfer efficiency (E) depending on the surface fraction of the perforation at various values of the aeration rate: (**a**) single turbine; (**b**) double turbines. (*Source*: From Ref. 33, with permission.)

ology of the broth is affected by the composition of the original medium and its modification by the growing culture, the concentration and morphology of the biomass, and the concentration of microbial products. Complex interaction between transport phenomena and reaction kinetics characterizes bioreactors and determines their performance. The modified turbine TP3 with the surface fraction S_G/S_C of perforations equal to 0.353, considered an optimum in dispersion of the gas and solids in liquid, was investigated in antibiotic biosynthesis liquids in semiindustrial and industrial bioreactors.

Biosynthesis of Antibiotics in Semiindustrial-Scale Bioreactors

The experiments were carried out in two 20-m^3 stainless steel stirred tank bioreactors, equipped with three modified Rushton turbines (TP3) (blade width × blade length = 0.306 × 0.225 m), and three witness Rushton turbines (TR) (blade width × blade length = 0.225 × 0.225 m), respectively, having the same diameter and rotating with the same agitator speed (38). The experimental data in the biological systems were obtained using four different microorganisms: *Streptomyces erithreus*, *S. griseus*, *S. noursei*, and *Nocardia mediterranei*, producing erythromycin, streptomycin, nystatin, and rifamycin, respectively. During the four parallel fermentations, the specific power consumption, the volumetric oxygen mass transfer coefficient, the oxygen transfer efficiency, and the antibiotic concentrations were investigated. For each fermentation type, the power consumption for the TP3 agitator is diminished by a factor of 1.18 (streptomycin) to 1.6 (erythromycin) compared to that for the TR agitator (Fig. 16). Therefore, in the viscous mycelial fermentation, the TP3 agitator minimizes the drag forces associated with the motion of the blades, confirming the observations presented for a nonbiological system.

The Kla values, as measured by the dynamic method, decrease during the fermentation processes, depending on the microorganism growing in the liquid (Fig. 17). The fermentation media for *S. erithreus* and *S. noursei*, producing erythromycin and nystatin, respectively, contain starch as a degradable component (22,39), which confers a high viscosity on the initial medium; thus, Kla is smaller at the beginning of biosynthesis, compared to the initial medium for *S. noursei* and *N. mediterranei* without this component. The Kla values obtained under mixing conditions with the TP3 agitator are 25% (streptomycin) to 50% (nystatin) higher than those obtained by using the TR agitator. This effect can be explained by the improvement of the dispersion of the gas bubbles as a result of the impact of the jets that come through the perforations in the TP3 blades onto the relatively surface in the back of large cavities.

Because the TP3 blades are higher by 36% in comparison with the witness blades, the size of the well-mixed, low-viscosity, and high-Kla region is effectively increased. In the micromixing region, excellent bubble breakup and shear thinning of the viscous medium occurs, which results in oxygen transfer improving in accordance with the concepts of Bajpai and Reuss (40). Also, the macromixing region, away from the impeller where the oxygen transfer is relatively poor, is much smaller.

In the investigated biological systems, for 1 W h of energy consumption (for agitation and air compression), the oxygen volume transferred into the liquid phase is approximately 30% higher in the bioreactor equipped with a TP3 agitator (Fig. 18). The oxygen transferred into the medium is consumed mainly in the combustion reaction that takes place within the biomass, on which the antibiotic production depends. The relative antibiotic concentrations measured over the entire course of the fermentations are plotted in Figure 19.

The efficiency of the TP3 agitator was confirmed in that antibiotic production increased up to 35% in nystatin biosynthesis, as much as 9% in rifamycin biosynthesis, and up to 20% in streptomycin biosynthesis as well. No difference in erythromycin production was observed between

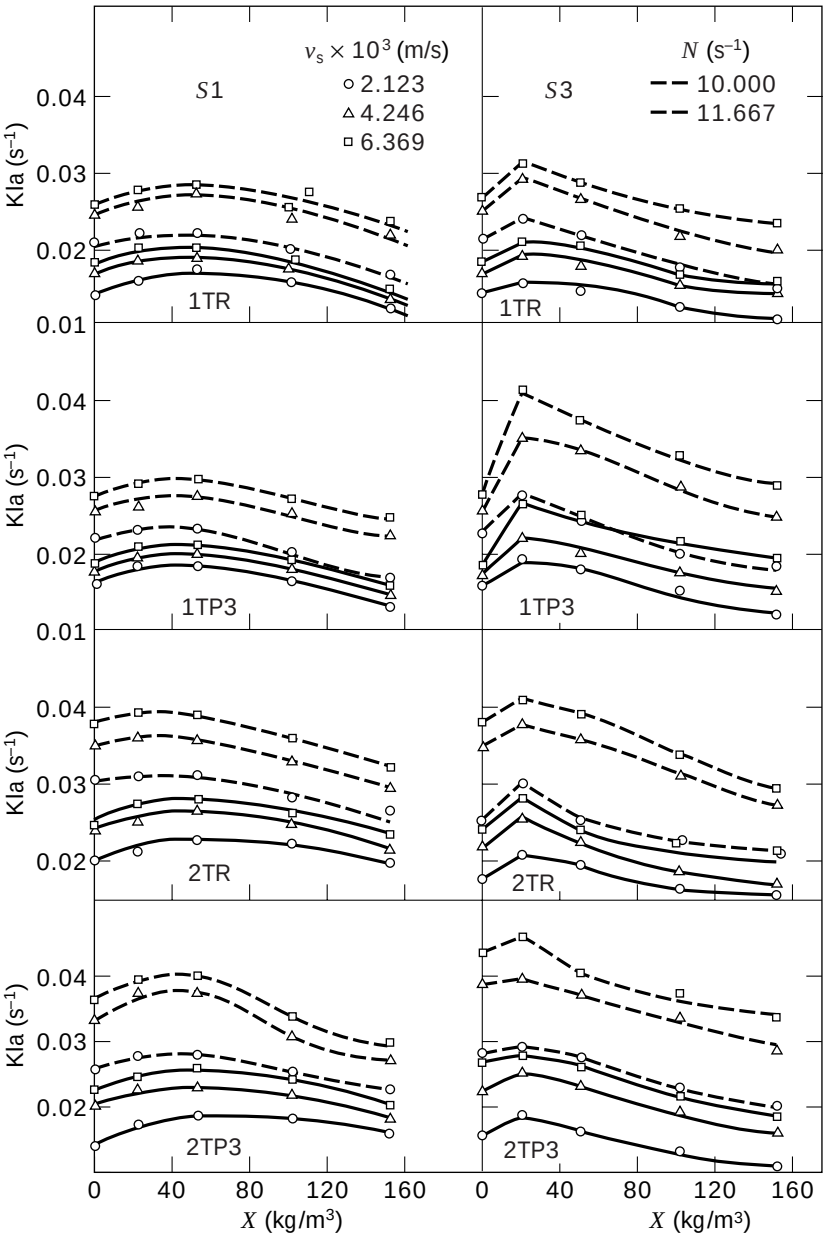


Figure 14. Effect of solid concentration on volumetric mass transfer coefficient for standard and modified Rushton turbine, positioned singly or doubly on the same shaft. (*Source*: From Ref. 35, with permission.)

the two parallel fermentations, but efficiency occurs from the energy savings. Therefore, using the TP3 agitator in the biosynthesis of antibiotics, it is possible to minimize the power consumption and to increase the antibiotic production for the same operating conditions as well as the TR agitator.

Biosynthesis of Antibiotics in Industrial-Scale Bioreactors

In the biosynthesis of antibiotics in semiindustrial-scale bioreactors, it was demonstrated that when the modified agitator has the same diameter and rotates with the same agitator speed as the standard agitator, the power consumption is reduced by a factor of 1.18 to 1.6 (38). Therefore, the use of a modified agitator in biosynthesis of the antibiotics may improve the productivity for a given power input by increasing the TP3 impeller diameter and keeping the agitator speed as high as for the TR agitator, and by increasing the TP3 agitator speed and keeping the impeller diameter as large as for the TR agitator.

The improved agitation performance by modified Rushton turbine agitators corresponding to the two aforementioned cases (if the specific power consumption is constant) was verified experimentally in industrial-scale bioreactors, as follows (41):

- In parallel fermentations of *S. aureofaciens* and *S. rimosus*, producing tetracycline and oxytetracycline, respectively, in two 100-m^3 stainless steel stirred tank bioreactors equipped with four modified impellers (TP3) and four standard impellers (TR), respectively (first case)
- In parallel fermentations of *Penicillium chrysogenum* producing penicillin, in two 63-m^3 stainless steel

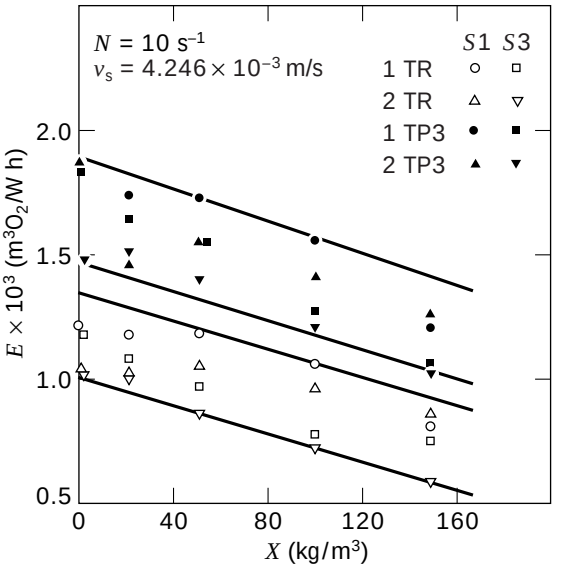


Figure 15. Relation between oxygen transfer efficiency and solid concentration for various agitators. (*Source*: From Ref. 35, with permission.)

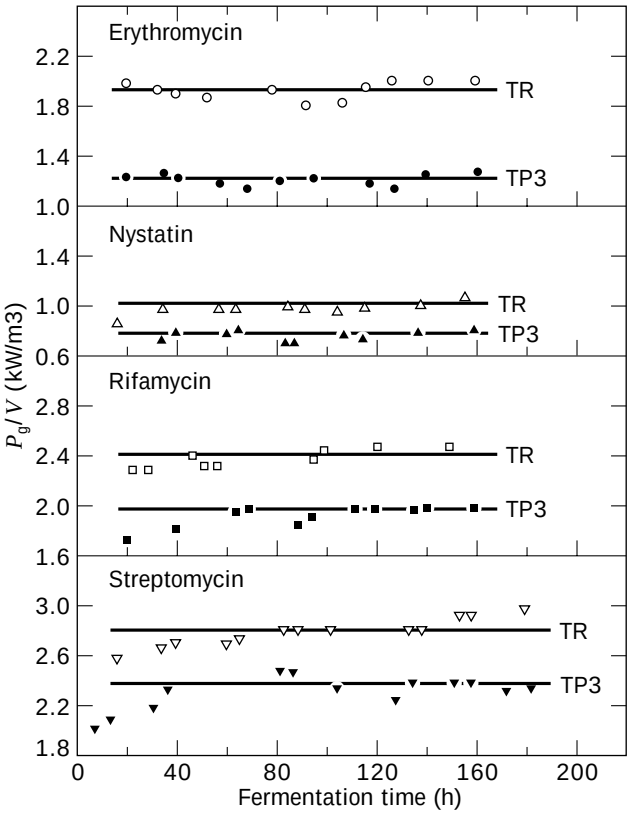


Figure 16. Comparison of TP3 and TR agitators with respect to specific power consumption in the course of processes in the biosynthesis of antibiotics. (*Source*: From Ref. 38, with permission.)

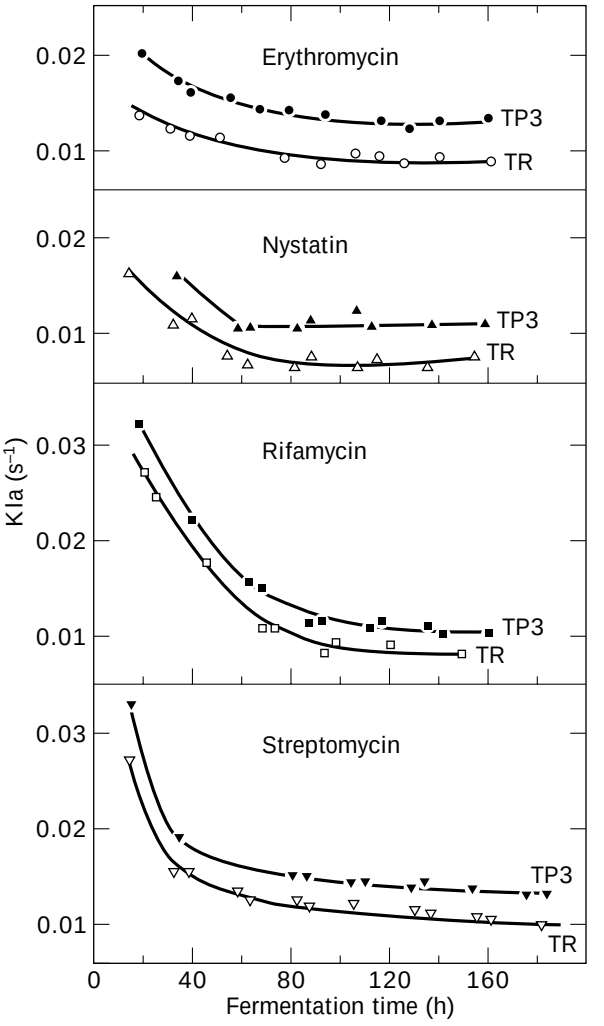


Figure 17. Comparison of TP3 and TR agitators with respect to the volumetric oxygen mass transfer coefficient in the course of processes in the biosynthesis of antibiotics. (*Source*: From Ref. 38, with permission.)

stirred tank bioreactors equipped with three modified impellers (TP3) and three standard impellers (TR), respectively (second case)

The morphology of molds is a function of shear stresses in the bioreactor. The growth rate and partial pressure of oxygen, and the following elements, play a role: the growth of hyphae, the branching of hyphae, the dispersion (breakup) of hyphae caused by the shear stress in the vessel, and the tensile strength of hyphae. The breakup frequency (λ/k_o) is proportional to the frequency that a particle will have in the so-called dispersion zone (V_{disp}), that is, the volume at which the turbulent shear forces are greater than the forces necessary to break the hyphal cell wall (42). The λ/k_o and V_{disp} values for the agitators investigated are presented in Table 3.

Therefore, by using the TP3 agitators, the size of the well-mixed region is effectively increased up to 27% and the breakup frequency of the mycelial hyphae is increased up to 75%. The relative antibiotic production and variation

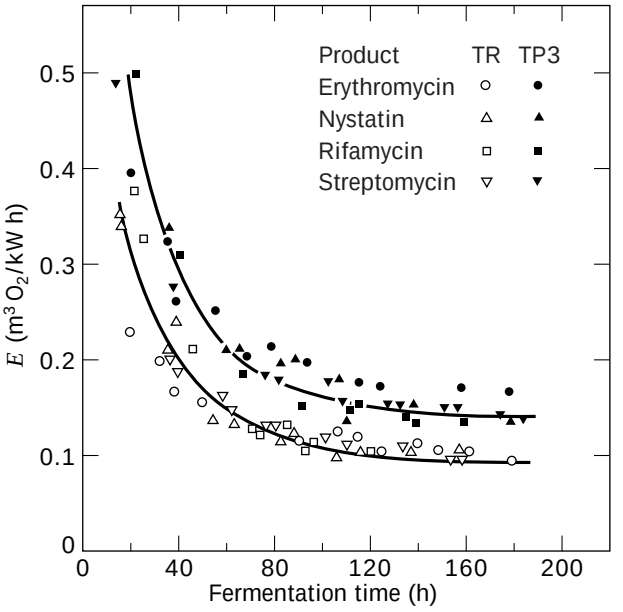


Figure 18. Comparison of TP3 and TR agitators with respect to oxygen transfer efficiency in the course of processes in the biosynthesis of antibiotics. (*Source*: From Ref. 38, with permission.)

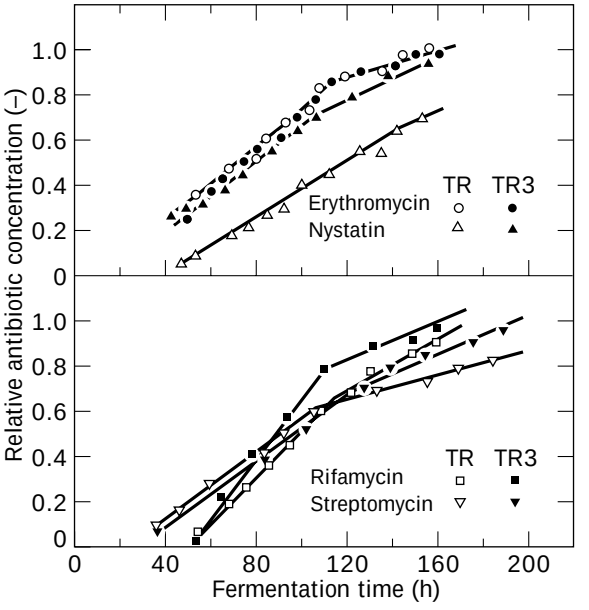


Figure 19. Relative antibiotic concentration profiles during the biosynthesis process using the TP3 and TR agitators in semiindustrial-scale bioreactors. (*Source*: From Ref. 38, with permission.)

of the specific antibiotic production rate, Q_p, during the cultivation time, are plotted in Figures 20 and 21, respectively. The improvement in antibiotic production by more than 30% by using modified agitators can be explained by creating a better hydrodynamic microclimate, and effectively increasing the well-mixed region, oxygen mass transfer, discharge efficiency, and the breakup frequency of mycelial hyphae.

Table 3. Breakup Frequency of the Mycelial Hyphae and the Dispersion Zone of TP3 and TR Agitators

Product	Agitator	V_{disp} (m^3)	$\lambda/k_0 10^2$ (s^{-1})
Tetracycline and oxytetracycline	TR	1.846	1.688
	TP3	3.522	4.808
Penicillin	TR	1.948	2.255
	TF3	2.474	3.937

Source: Ref. 41.

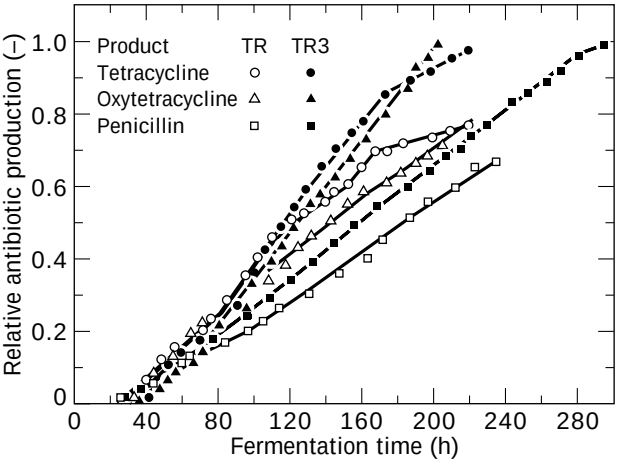


Figure 20. Relative antibiotic concentration profiles during the biosynthesis process using the TP3 and TR agitators in industrial-scale bioreactors. (*Source*: From Ref. 41, with permission.)

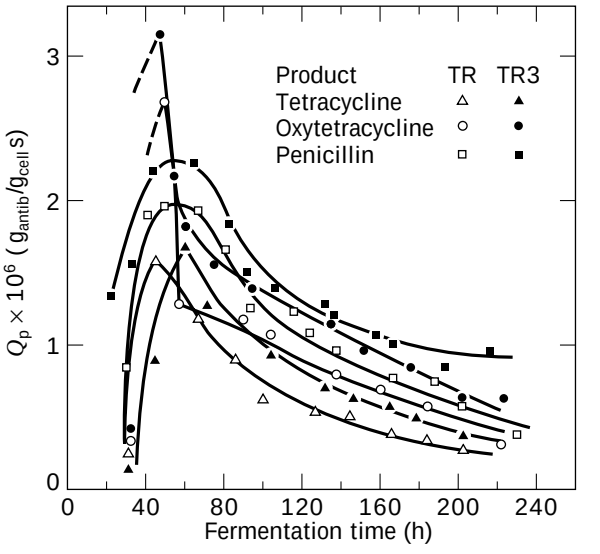


Figure 21. Specific antibiotic production rate profiles during the biosynthesis process using TP3 and TR agitators. (*Source*: From Ref. 41, with permission.)

The Q_p values reach a maximum in the early exponential phase of batch growth, although the cell mass concentration is small. The maximum values of the specific antibiotic production rate and the fermentation time at which these maximum points are reached depend on the type of producing microorganism and the type of agitator. Therefore, the standard and modified agitators create various types of environmental conditions under which the cells appear to adapt by structural and functional changes.

BIBLIOGRAPHY

1. M.M.C.G. Warmoeskerken, Ph.D. Thesis, University of Delft, The Netherlands, 1985.
2. J.M. Smith, in J.J. Ulbrecht and G.K. Patterson eds., *Mixing of Liquid by Mechanical Agitation*, Gordon and Breach, New York, 1985, pp. 139–202.
3. A.W. Nienow and J.J. Ulbrecht, in J.J. Ulbrecht and G.K. Patterson eds., *Mixing of Liquid by Mechanical Agitation*, Gordon and Breach, New York, 1985, pp. 203–237.
4. F. Kargi and M. Moo-Young, in M. Moo-Young ed., *Comprehensive Biotechnology*, vol. 2, Pergamon Press, Oxford, 1985, pp. 5–56.
5. R.V. Roman and R.Z. Tudose, *Rev. Chim.* **46**, 547–564 (1995).
6. R.V. Roman and R.Z. Tudose, *Roum. Chem. Q. Rev.* **4**, 269–293 (1996).
7. R.V. Roman, Ph.D. Thesis, Technical University, Iasi, Roumania, 1997.
8. R.V. Roman and R.Z. Tudose, *Chem. Eng. J.* **61**, 83–93 (1996).
9. R.V. Roman and M. Gavrilescu, *Hung. J. Ind. Chem.* **22**, 87–93 (1994).
10. V. Abrardi, G. Rovero, G. Baldi, S. Sicardi, and R. Conti, *Trans. Inst. Chem. Eng. A* **68**, 516–521 (1990).
11. R.V. Roman and R.Z. Tudose, *Rev. Roum. Chim.* **41**, 309–319 (1996).
12. W. Bujalski, A.W. Nienow, S. Chatwin, and M. Cooke, *Chem. Eng. Sci.* **42**, 317–326 (1987).
13. E.A. Bratu, *Operatii Unitare în Industria Chimică*, vol. 2, Editura Technică, Bucharest, 1985, p. 20.
14. R.V. Roman, M. Gavrilescu, and V. Efimov, *Hung. J. Ind. Chem.* **20**, 155–160 (1992).
15. R.K. Geisler, R.K. Buurman, and A.B. Mersmann, *Chem. Eng. J.* **51**, 29–39 (1993).
16. K. Takahashi, H. Fujita, and T. Yokota, *Chem. Eng. Jpn.* **26**, 98–100 (1993).
17. A. Barresi and G. Baldi, *Chem. Eng. Sci.* **42**, 2949–2956 (1987).
18. K.J. Myers, *Can. J. Chem. Eng.* **72**, 745–748 (1994).
19. T.N. Zwietering, *Chem. Eng. Sci.* **8**, 244–253 (1958).
20. R.V. Roman and R.Z. Tudose, *Bioprocess Eng.* **15**, 221–229 (1996).
21. T. Koloini, I. Plazi, and M. Zumer, *Chem. Eng. Res. Des.* **67**, 526–536 (1989).
22. M. Gavrilescu, R.V. Roman, and V. Efimov, *Acta Biotechnol.* **12**, 383–396 (1992).
23. P.M. Armenante and T. Li, *AIChE Symp. Ser.* **293**, 105–111 (1993).
24. C. Buurman, *I. Chem. E. Symp. Ser.* **121**, 343–350 (1990).
25. M. Greaves and V.Y. Loh, *I. Chem. E. Symp. Ser.* **89**, 69–96 (1984).
26. K. Vant'Riet, Ph.D. Thesis, University of Delft, The Netherlands, 1975.
27. R.V. Roman and R.Z. Tudose, *Bioprocess Eng.* **16**, 135–144 (1997).
28. R.V. Roman and R.Z. Tudose, *Bioprocess Eng.* **17**, 55–60 (1997).
29. R.V. Roman and R.Z. Tudose, *Hung. J. Ind. Chem.* **24**, 161–169 (1996).
30. R.V. Roman and R.Z. Tudose, *Bioprocess Eng.* **17**, 307–316 (1997).
31. B.W. Rainer, *Chem. Biochem. Eng.* **Q4**, 185–196 (1990).
32. H. Oguz, A. Brehm, and W.D. Deckwer, *Chem. Eng. Sci.* **42**, 1815–1822 (1987).
33. R.V. Roman and R.Z. Tudose, *Hung. J. Ind. Chem.* **24**, 185–192 (1996).
34. N.M.G. Oosterhuis, Ph.D. Thesis, University of Delft, The Netherlands, 1984.
35. R.V. Roman and R.Z. Tudose, *Bioprocess Eng.* **17**, 361–365 (1997).
36. A. Brehm, S. Ledakowicz, and R. Kokuun, *Proc. 8th Int. Cong. CHISA*, Praha, 1984.
37. W.M. Lu, R.C. Hsu, and H.S. Chou, *J. Chim. I. Ch. E.* **24**, 31–39 (1993).
38. R.V. Roman and M. Gavrilescu, *Acta Biotechnol.* **14**, 181–192 (1994).
39. M. Gavrilescu, R.V. Roman, and V. Efimov, *Acta Biotechnol.* **13**, 59–70 (1993).
40. R.K. Bajpai and M. Reuss, *Can. J. Chem. Eng.* **60**, 384–392 (1982).
41. R.V. Roman, R.Z. Tudose, M. Gavrilescu, M. Cojocaru, and S. Luca, *Acta Biotechnol.* **16**, 43–56 (1996).
42. J.C. Van Suijdam and B. Metz, *Biotechnol. Bioeng.* **23**, 111–148 (1981).

See also MASS TRANSFER.

TRANSGLUTAMINASE

KATSUYA SEGURO
Food Research and Development Lab
Kawaski, Japan

KEY WORDS

Acyl acceptor substrate
Acyl donor
Acyl transfer reaction
Calcium
Cross-linked proteins
Enzymatic activity
Food proteins
MTG (microbial enzyme)

OUTLINE

INTRODUCTION

Transglutaminase (glutaminyl-peptide γ-glutamyltransferase/R-glutaminyl-peptide: amine γ-glutamyl transferase, EC 2.3.2.13) is an enzyme that catalyzes the acyl-transfer reaction between the γ-carboxyamide group of peptide- and/or protein-bound glutamine as the acyl donor and a variety of primary amines as the acyl acceptors (Fig. 1a) (1). If the side chain of protein-bound lysine acts as the acceptor, transglutaminase cross-links protein molecule(s) inter- and/or intramolecularly through the formation of ϵ-(γ-glutamyl)lysine isopeptide bonds between the γ-carboxyamide groups of glutamine and ϵ-amino groups of lysine in the protein molecule(s) (Fig. 1b). If no amines are present in the reaction system, the water molecule acts the acceptor, and the reaction results in deamidation of the γ-carboxyamide groups of glutamine in the proteins (Fig. 1c). Transglutaminase therefore is responsible for the post-translational modification of proteins as a biocatalyst.

In 1950, Borsook et al. observed the incorporation of radioactive lysine into protein components in the crude extract of guinea pig liver (2). Since then, research on the enzyme has continuously progressed. At present, the activity of most transglutaminases known so far is expressed in a calcium-concentration-dependent manner (3–8). Under physiological conditions in which the calcium catch-release mechanism is well regulated, most transglutaminases are controlled by such a mechanism. Transglutaminase is thus said to be calcium dependent. A typical example of transglutaminase-catalyzed protein cross-linking is readily observed in the final stabilization step of the blood coagulation or wound-healing process in warm-blooded species. In this case, plasma, platelet, and/or placental factor XIIIa (the activated form of blood transglutaminase) plays an important role in fibrin cross-linking, when the thrombin-mediated proteolytic cascade reaction along with elevation of endogenous calcium ion levels is induced by injury. The resulting cross-links are covalent, stable, and resistant to proteolysis, thereby increasing the resistance of fibrin clots to chemical, enzymatic, and physical degradation. The transglutaminase reaction with polyamines results in protein modifications, possibly affecting the biological activity or turnover of the target protein, but not polymer formation. Transglutaminase has a broad specificity for the acyl acceptor substrates (a variety of primary amines, peptide-bound lysine, or polyamines), but a relatively narrow specificity for the acyl donors (peptide-bound glutamine residues) is confirmed (amino acid sequence around the susceptible glutamine residue influences the specificity). Denaturation or proteolysis of a nonreactive protein can convert it to a transglutaminase substrate.

Transglutaminase thus can be used for intentional alteration of food protein functionalities (9). The starting point for the present industrial production of transglutaminase is the discovery of Ando et al. (10); calcium-independent transglutaminase is obtained by fermentation of a nonpathogenic and nontoxicogenic variant of *Streptoverticillium mobaraense.*

DISTRIBUTION OF TRANSGLUTAMINASE IN NATURE

Distribution of transglutaminases is ubiquitous in nature. Most mammalian organs, epidermis, plasma, and extracellular fluids contain transglutaminases in multiformity (6–8). Mammalian transglutaminase exists in five types: keratinocyte transglutaminase (TG_K, or type I), tissue transglutaminase (TG_C, or type II), epidermal transglutaminase (TG_E, or type III), prostate transglutaminase (TG_P, or type IV), and plasma and platelet transglutaminases (factor XIIIa) (Table 1). Each enzyme may have distinct physiological functions in the body; that is, the well-known blood transglutaminase, referred to as factor XIIIa, stabilizes fibrin entanglement by cross-linking; the guinea pig liver transglutaminase is classified as TG_C. Although this guinea pig liver transglutaminase was the first enzyme to be found and has been well characterized, its physiological role was the last to be solved.

Transglutaminases are found in not only in mammals, but also in fish, shellfish, crustaceans, amphibians, plants, and microorganisms. For instance, white croaker, carp, sardine, Alaska pollock, chum salmon, rainbow trout (11), horse mackerel (12), hoki (13), *Crassostrea gigas* (oyster) (14), *Tachypleus tridentatus* (horseshoe crab) (15), tadpoles of *Rana catesbeiana* (bullfrog) (16), *Beta vulgaris* L. (Silver beet), broccoli, spinach, (17), *Medicago sativa* L. (alfalfa) (18), *Helianthus tuberosus* (19), *Pisum sativum* seed (20), *Physarium* (21), *Bacillus* (22), and *Streptoverticillium* (10,23) have been reported to contain transglutaminase ac-

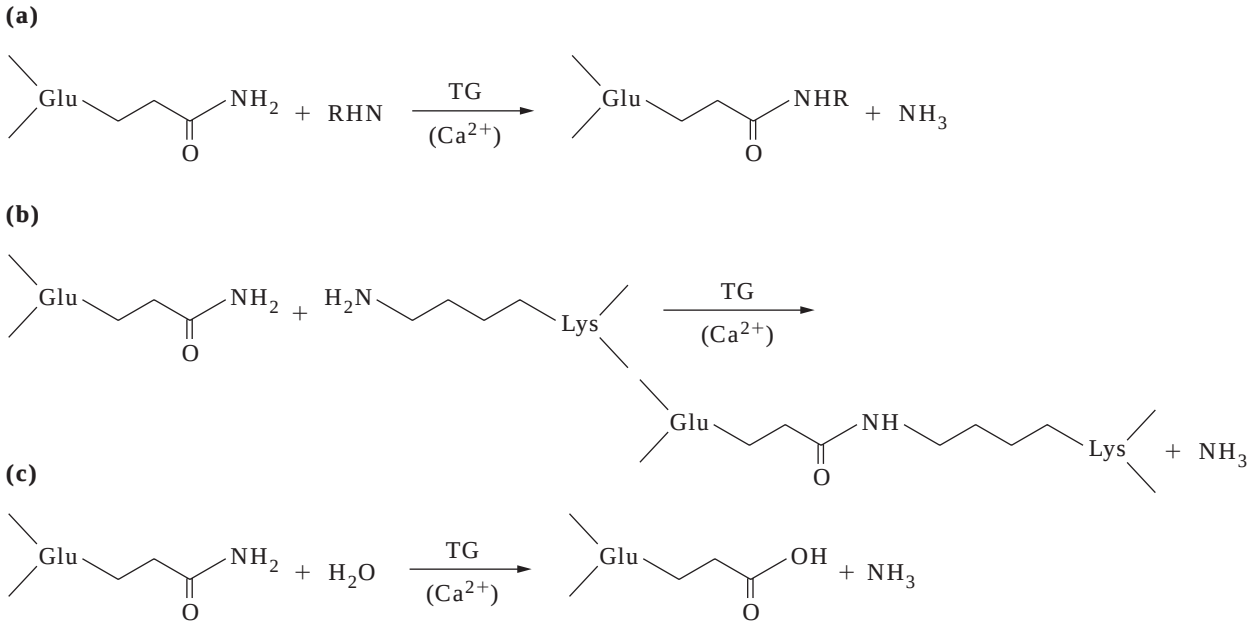


Figure 1. Reaction scheme of transglutaminase activity. (**a**) incorporation of a primary amine into γ-carboxyamide group of peptide-bound glutaminyl residue; (**b**) cross-linking through formation of ϵ-(γ-glutamyl)lysine bond between the γ-carboxyamide group and ϵ-amino group of peptide-bound lysyl residue; and (**c**) deamidation of the γ-carboxyamide group.

Table 1. Physicochemical Properties of Mammalian Transglutaminases

Type	Molecular weight	Subunit structure		Protease activation[a]	GTP effect[b]
Type I, TGK	≈90,000	Monomer	+ fatty acylation	No	None
Type II, TGC	≈8,000	Monomer	+ none	No	Inhibitory
Type III, TGE	≈80,000	Monomer	+ none	Yes[c]	?
Type IV, TGP	≈150,000 (≈71,000 × 2)	Homodimer	+ ?	?	None
Factor XIII					
Plasma	≈350,000	a2b2	+ none	Yes (thrombin)	None
Platelet	≈160,000	a2	+ none	Yes (thrombin)	None

[a]Protease responsible for in vivo activation of transglutaminase is indicated in parentheses.
[b]Effect of guanosine triphophate on transglutaminase activity.
[c]Proteases physiologically involved have not been specified.

tivity, for example, in organs, muscles, or tissue matrices. Although the classification of many enzymes is left undetermined, it is obvious that the enzymatic activity of all the transglutaminases, except microbial enzymes, is expressed only in the presence of calcium ions. The activity of transglutaminases from some microorganisms, such as *Bacillus* and *Streptoverticillium*, is expressed in both the absence and presence of calcium ions. This calcium independence is advantageous for food protein modification because many food proteins are affected by calcium ions. Further, recent investigations revealed that the activity of transglutaminases from oyster and scallop was quite interestingly regulated by both calcium and sodium ions.

ASSAYS FOR TRANSGLUTAMINASE ACTIVITY

There are both quantitative and qualitative assay methods for transglutaminase activity. Quantitative determination can be accomplished by (*1*) measuring the incorporation of any of several radioisotopically labeled primary amines (e.g., [^{14}C]- or [^{15}N]-labeled putrescence, cadaverine, methylamine, and glycine ethyl ester) into caseins, dimethylated caseins, the acetylated B-chain of oxidized insulin, or the glutaminyl peptide derivative, carbobenzoxy(*Z*)-glutaminylglycine; (*2*) measuring the incorporation of fluorescently-labeled primary amines (in most cases, dansylcadaverine) into caseins, dimethylated caseins, the acetylated B-chain of oxidized insulin, or *Z*-glutaminylglycine; (*3*) colorimetric measurement of an enzymatically produced colored γ-glutamylhydroxamate, between hydroxylamine and *Z*-glutaminylglycine, with acidic conditions (Fig. 2); or (*4*) measuring the ammonium released in the acyl transfer reaction (24).

For qualitative determination of the transglutaminases, the detection of either cross-linked proteins by

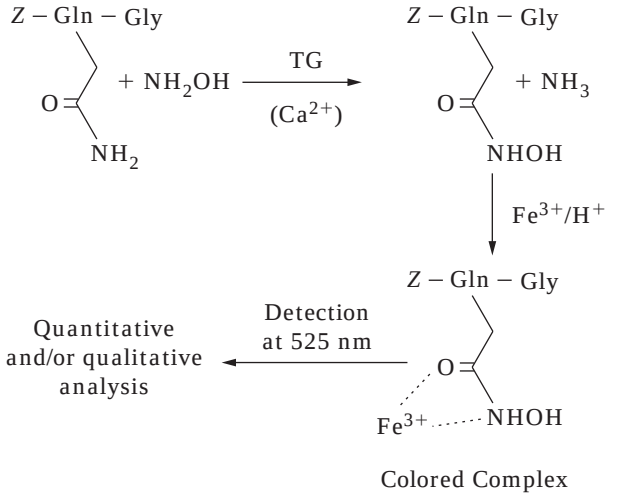


Figure 2. Reaction scheme of hydroxamate assay. Carbobenzoxy(*Z*)-glutaminylglycine and hydroxylamine act as the acyl donor and acceptor substrates. On the addition of ferric chloride/trichloroacetic solution in the acidic pH, the γ-glutamylhydroxamate derivative of *Z*-glutaminylglycine gives a red-colored complex.

(*1*) sodium dodecysulfate polyacrylamide gel electrophoresis (SDS-PAGE) with a reducing agent (e.g., 2-mercaptoethanol or dithiothreitol), or (*2*) test-tube inversion is effective. However, concentrated protein solution (a protein concentration of more than 50 to 100 mg/mL) must be used in the test-tube inversion assay because transglutaminase cross-links the protein substrate in both dilute and concentrate solutions, but a gel that remains at the bottom of the tube can be visually detected only in the concentrated protein solution. The detection of the ϵ-(γ-glutamyl)lysine bond, the enzymatic product, also indicates the presence of transglutaminase (25,26). Further, enzyme-linked immunosorbent assay (ELISA) using either a monoclonal or polyclonal antibody specifical to certain transglutaminases (27) indicates the presence of the enzyme both qualitatively and quantitatively, depending on the conditions of the assay. This method is effective for the detection of denatured transglutaminases in transglutaminase-utilized products.

To detect transglutaminase activity, appropriate assay methods should be carefully chosen. For instance, the sensitivity of the method, the quantity, and the specific activity of transglutaminase may alter the result. Also, substrate specificity, as well as the availability of the detective apparatus and appropriate facilities should be considered. It is known that transglutaminase from the hair follicles of guinea pigs did not recognize hydroxylamine and/or *Z*-glutaminylglycine as a substrate (28). Therefore, this method is not applied for the detection of hair follicle transglutaminases, although the hydroxamate method (hydroxylamine and *Z*-glutaminylglycine) is the most convenient for quantitative and qualitative detection of most transglutaminases.

In the all assays except that for ϵ-(γ-glutamyl)lysine detection, one has to bear in the mind that false results are always a possibility. For instance, some amidases and peptidoglutaminase (EC 3.5.1.43; an enzyme that catalyzes the incorporation of primary amines into the γ-carboxyamide group of glutamines in only short peptides, not in proteins [29]) give a positive result in the hydroxamate assay. Further, various amine oxidases, including lysyl oxidase and peroxidase, may introduce various primary amines in the protein substrates, or cross-link the protein substrates through Schiff base formation by oxidation of amines or the ϵ-amino group of lysine residues (6). Therefore, for final confirmation that the enzyme is truly a transglutaminase, one should check the presence or increment of the ϵ-(γ-glutamyl)lysine bond.

MAMMALIAN TRANSGLUTAMINASE

Transglutaminases in mammalian organs and tissues exit in multiformity depending on their physiological functions and significance. Ordinarily, for all the transglutaminases, including the well-characterized guinea pig liver enzyme and plasma factor XIIIa, the calcium ion requirement for expression of enzymatic activity is absolute (3–8). The primary structure of the five mammalian transglutaminases, namely, keratinocyte transglutaminase (TG_K, or type I), epidermal transglutaminase (TG_E, or type III), tissue transglutaminase (TG_C or type II), and prostate transglutaminase (TG_P or type IV), including plasma and platelet transglutaminase (factor XIIIa), have been established by either cDNA cloning or protein sequencing.

A major physiological role of all the distinct transglutaminases, except tissue transglutaminase, is to cross-link certain protein substrates in accordance with their existence in the organs, plasma, and extracellular fluids. Distinct transglutaminases may recognize the same protein as substrate, but different glutamine residues within the protein act as the acyl donors with different affinity and/or specificity. Their possible physiological functions are reviewed in detail elsewhere (6–8).

Physicochemical Properties

Transglutaminases can be purified by a set of various modes of chromatography (i.e., anion and cation exchange, heparin, blue Sepharose, or antibody-conjugated affinity, hydrophobic, and size-exclusion chromatography). The properties of the purified transglutaminases are summarized in Table 1. The monomeric forms of the four typical mammalian transglutaminases quite resemble one another, and their primary structures differ only slightly, thereby resulting in slightly different molecular weights (Table 1). The primary structures as revealed by cDNA cloning and/or protein sequencing have approximately 670–820 amino acid residues, and approximate molecular weights of the monomeric form of the four transglutaminases stay within the range of 71–92 kDa. TG_K has extra N- and C-terminal regions (about 60 amino acid residues) and fatty acylation through thioester-linked palmitic and myristic acid; hence, TG_K has the largest molecular weight. Other transglutaminases, such as TG_C and TG_E, have a few possible N-linked glycosylation sites and more

than 10 cysteine residues, but no glycosylation and disulfide bonding have been observed.

The structural similarity, as expressed by degrees of sequence identity among the four transglutaminases, is 92% between human and rat TG_K's, 80% between human and guinea pig TG_C's, and 75% between human and mouse TG_E's. On the other hand, the structural similarity between each of the four typical transglutaminases ranges from 31% to 37%. This indicates that the degree of sequence identity between species is rather more conserved than that between the types, thereby suggesting that the ancestral genes for each transglutaminase are limited to a few types (8).

Plasma and platelet factor XIII molecules have molecular weights of approximately 360,000 and 150,000, respectively (3,4,30). They are the zymogen of factor XIIIa, an activated form. Plasma factor XIII exists as a tetramer comprising two noncovalently associated a-subunits and two filamentous b-subunits, whereas platelet factor XIII exits as a dimer comprising two noncovalently associated a-subunits. Both plasma and platelet factor XIII are activated by α-thrombin cleavage of the a-subunit into factor XIIIa. In plasma factor XIIIa, b-subunits dissociate from a-thrombin-treated plasma factor XIII (or factor XIIIa) in the presence of calcium ions. In the continued presence of calcium ions, both plasma and platelet factor XIIIa become catalytically active. The a-subunit resembles four typical transglutaminases, with a sequence identity of 43% with TG_K, 36% with TG_C, 34% with TG_E, and 27% with TG_P. The three-dimensional structure of an activated form of factor XIII has been determined (29). Thorough observation of the structure revealed that the catalytic triad residues (cysteine-314, histidine-373, and aspartate-396) are essential for the expression of transglutaminase activity. Such structural information is critical not only for protein chemistry, but also for protein engineering and genetic engineering. The differences therefore are seen mainly in the physicochemical properties of the six transglutaminases from various sources and origins; amino acid sequences around the active site (an important region for expression of the catalytic activity) are highly conserved among all six transglutaminases (Fig. 3).

Figure 3. Sequence alignment of the active site region of six types of human transglutaminase. TG_K, keratinocyte enzyme; TG_C, tissue enzyme; TG_E, epidermal enzyme, TG_P, prostate enzyme; XIII, factor XIIIa subunit; MTG, microbial enzyme from *Streptoverticillium mobaraense* var. Underlines indicate residues conserved through all the transglutaminases, except MTG.

Enzymatic Properties

Enzymatic properties, such as optimum pH and temperature for the enzymatic activity of the guinea pig liver transglutaminase (TG_C) and factor XIII, have been extensively characterized and reviewed (3–5), whereas those of the remaining transglutaminases have been summarized briefly (7,8). For instance, TG_C can bind guanosine triphosphate (GTP) (32). At the same time, TG_C hydrolyzes GTP, and the TG_C activity is inhibited by GTP. In this context, the possible physiological function of TG_C has been postulated as a G-protein (Gh) in receptor-mediated signal transduction (33), and such a property is unique among the six transglutaminases.

MICROBIAL TRANSGLUTAMINASE

To achieve commercialization of transglutaminase, a constant mass supply is an absolute prerequisite. Among microbial transglutaminases from three microorganisms, *Bacillus*, *Physarium*, and *Streptoverticillium*, transglutaminase from *Physarium* has a molecular weight of 70,000 (21) (the largest of the three), whereas very few characteristics have been determined for *Bacillus* transglutaminase (22). On the other hand, transglutaminases from *Streptoverticillium mobaraense* var. (MTG) have been well characterized (10,23,34), and their industrial production has been established (see later). MTG is secreted into the culture broth so that cell disruption is unnecessary, and subsequently, easy purification has accelerated commercialization. *Streptoverticillium* transglutaminase was first named "bacterial transglutaminase" (BTG, or BTGase), however, the final taxonomical classification into *S. mobaraense* var. (23) urged the renaming of the enzyme as "microbial" transglutaminase (MTG, or MTGase).

Manufacturing Process

MTG is produced through the fermentation process using *S. mobaraense* var. (Fig. 4). The inoculum for industrial fermenters is built up in three stages: started in a shake flask, then moved via aseptic transfer to a large fermenter, and finally to a larger fermenter. Food-grade raw materials are sterilized and are either batched or continuously

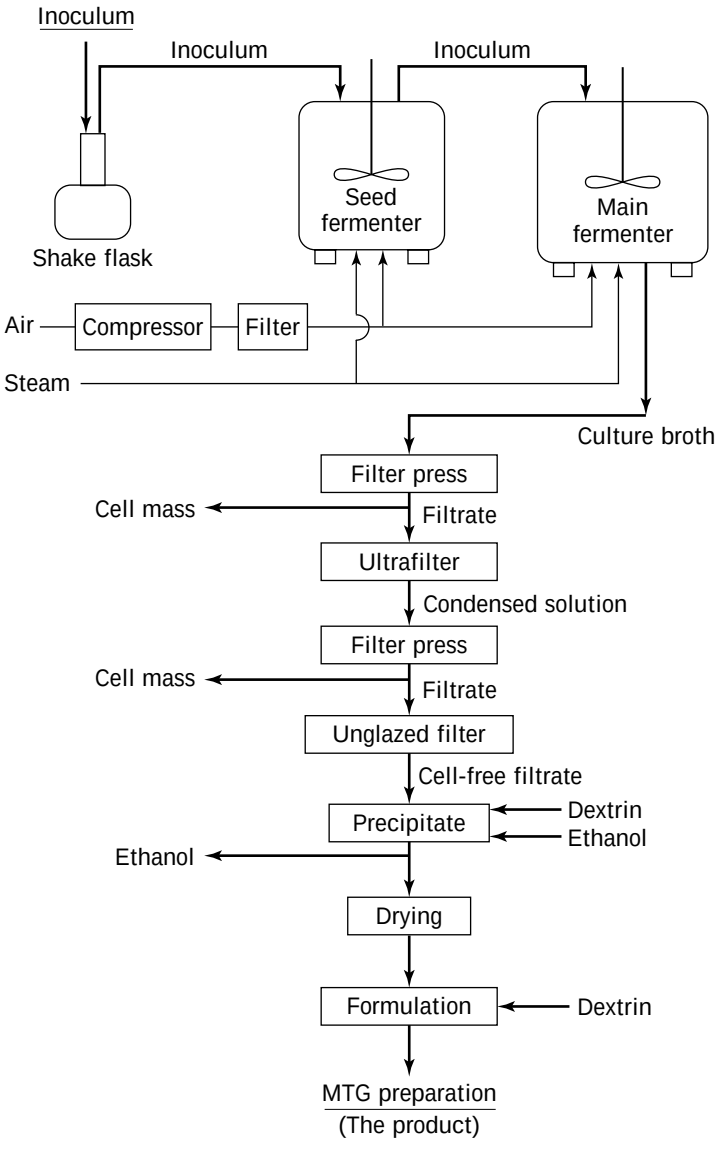


Figure 4. Manufacturing flow of transglutaminase from *Streptoverticillium mobaraense* var.

pumped to the fermenters. The build-up of the inoculum is started when the media are properly prepared and adjusted to suitable growth conditions, and a small inoculum of the pure culture is added. When proper growth is confirmed, the broth is pumped to the next phase. After these serial transfers, the inoculum is then transferred to the main fermenter. Under aseptic conditions, with pH, temperature, aeration and nutrients controlled, the organism begins to produce and excrete the enzyme. When MTG is accumulated to the proper level, the aerobic process is stopped. Thereafter, the whole broth is considered suitably treated, and the fermentation phase is considered completed. Next come the recovery and purification processes. The recovery of MTG from the cultural broth is initiated by the process that removes the organism and spent media. The liquid (broth) that contains the enzyme is subjected to ethanol precipitation, and the fraction containing the enzyme is concentrated by drying in vacuo before it is formulated for the product. All substances added during the recovery and formulating processes are of food grade. The product is then set aside until quality control can assure that it passes all the specifications.

There are various industrial technologies for enzyme recovery, most of which are currently employed. These technologies include the salting-out method, column chromatography, liquid–liquid extraction, ultrafiltration, diafiltration, and so forth. The processes employed in the recovery of MTG are ultrafiltration and solvent precipitation. Following fermentation of the producing organism, the organism is eliminated from the broth by a filter press and then a fine filter apparatus (made of unglazed filter tube). Food-grade dextrin is added to the concentrated enzyme solution for protection of the enzyme, and the enzyme is precipitated by ethanol. Then, the ethanol is removed by centrifugation, and the enzyme is dried in vacuo. The dried enzyme is formulated into the product, whose enzymatic activity is adjusted to 1,000 to 1,150 U/g. The product is kept until the quality control studies are completed.

Fermentation Flow. The pure culture of *S. mobaraense* var is lyophilized; the medium for storage consists of 10% skim milk and 1% monosodium L-glutamate. A lyophilized culture in an ampoule is thawed and suspended in sterile water. After suitable dilution with sterile water, a 0.15-mL portion of the suspension is aseptically inoculated in Venette agar medium containing glucose, peptone, meat extract, yeast extract, and agar. The inoculated media are incubated at 30 ± 1 °C for 10 days. After confirming sufficient mycelium growth and spore formation, the test tubes are examined macroscopically and microscopically. The strain is checked for the standard requirement of MTG activity (hydroxamate assay), protease activity, and pH value. The best strain selected by the above purity test is stored as stock spores at 5 ± 1 °C.

The preculture medium has the following ingredients, in decreasing order of predominance: starch, glycerol, meat extract, yeast extract, potassium dihydrogen phosphate, magnesium sulfate, and antifoam agent. These ingredients are dissolved in water, and the solution is neutralized with sodium hydroxide solution and sterilized at 121 °C for 20 min. Five 2-L shake flasks containing 500 mL of sterile preculture medium are inoculated with a loopful of the spores. The flasks are incubated on a reciprocal shaker at 30 ± 1 °C for 48 h. All of them are inoculated to seed fermenters including sterile preculture medium. The seed fermenter is aerobically incubated at 30 ± 1 °C for about 26 h under a pressure of 0.5 kg/cm^2 with agitation. The whole seed culture broth is transferred to the main fermenter containing a medium described later. The fermenter is conducted aerobically at 30 ± 1 °C for 85–90 h under a pressure of 0.5 kg/cm^2 with agitation. During the fermentation, the pH of broth and the amount of CO_2 in the waste gas are continuously measured; the mycelial growth is also checked. A microscope check for foreign microorganisms is done at 24-h intervals; if any are found, the entire batch is discarded.

The main culture medium has the following ingredients, in decreasing order of predominance: starch, peptone, meat extract, yeast extract, corn steep liquor, potassium dihydrogen phosphate, dipotassium phosphate, magnesium sulfate, and antifoam agent. These ingredients are dissolved in water, and the solution is neutralized with sodium hydroxide solution and sterilized at 121 °C for 30 min.

Purification Flow. The culture broth is passed through a filter press using diatomaceous earth to remove viable cells of *S. mobaraense* var. The obtained filtrate is concentrated to approximately 1/12 volume of the original culture broth and desalted by ultrafiltration. The concentration is filtered using a filter press and a fine-filter apparatus, which is made of an unglazed filter tube. The cell-free filtrate is mixed with ethanol after addition of food-grade dextrin, then stood at 10 ± 2 °C. Ethanol is removed with decantation. The precipitate is washed with ethanol twice. The precipitates are recovered by centrifugation, and the cake is dried in vacuo at 39 ± 1 °C for 22–24 h. The dry powder is ground out and passed through a 60-mesh screen. If necessary, the TG activity can be adjusted to an appropriate uniformity by mixing with a diluent such as dextrin.

Physicochemical Properties

MTG has a molecular weight of approximately 38,000 as determined by SDS-PAGE and gel-permeation chromatography (10). Mass spectrometric analysis indicates that MTG has a molecular weight of 37,869.2 ± 8.8. Amino acid (34) and cDNA (23) sequences reveal that MTG comprises 331 amino acid residues and has a single cysteine residue as a potential active site (Fig. 5). Although two potential glycosylation sites (-Thr-Xxx-Asn-) exist in the primary structure, MTG is a monomeric and simple protein (not a glycoprotein or lipoprotein, etc.). The cDNA structure reveals that MTG has a signal peptide of 18 amino acid residues in its N-terminal in the precursor. In fact, MTG is secreted into the cultural broth during the fermentation. A high isoelectric point (pI 8.9) suggests that MTG is a basic protein.

Enzymatic Properties

MTG has a wide optimum pH (from 5 to 8) for activity. Even at pH 4 or 9, MTG still shows some enzymatic activity. MTG is also stable at wide pH ranges. An optimum temperature for MTG activity was 45–55 °C, and MTG fully sustained its enzymatic activity even after heating at 50 °C for 10 min, both in the presence and absence of reducing agents (Fig. 6). On the other hand, MTG losses its activity within a few minutes on heating to 70°C. MTG expresses its activity even at 10 °C, and it exerts some activity even at near freezing (0 °C).

Expression of enzymatic activity for MTG is totally independent of calcium ions, and in this aspect MTG is unique as compared with other mammalian enzymes. Such calcium independence is useful for the modification of the functional properties of food proteins. Many food proteins, (e.g., caseins, soybean globulins, and myofibrillar proteins) are susceptible to calcium ions and, subsequently, are precipitated easily by calcium ions. The sensitivity toward other cations has also been investigated, and copper, zinc, and lead ions notably inhibited MTG. The results are reasonable because these heavy metals bind to and block the thiol group of the single cysteine residue; meanwhile, MTG was inhibited by thiol-modifying agents (10). For substrate specificity of MTG, to date, most food proteins, such as legume globulins, wheat glutens, egg yolk and white proteins, ovalbumin, actins, myosins, fibrins, milk caseins, α-lactalbumin, and β-lactoglobulin, as well as many other albumins, are cross-linked (35–42).

ALTERATION OF FUNCTIONALITIES OF FOOD PROTEINS BY TRANSGLUTAMINASE

In the formation and maintenance of tertiary and quaternary structures of protein molecules, hydrogen bonds, disulfide bonds, and electrostatic and van der Waal's hydrophobic interactions are essential. Especially in food, not only these bonds, but also other bonds and interactions, particularly those found in connective tissues, are very

```
1                                                                  50
DSDDRVTPPAEPLDRMPDPYRPSYGRAETVVNNYIRKWQQVYSHRDGRKQ

51                                                                 100
QMTEEQREWLSYGC*VGVTWVNSGQUPTNRLAFASFDEDRFKNELKNGRPR

101                                                                150
SGETRAEFEGRVAKESFDEEKGGQRAREVASVMNRALENAHDESAYLDNL

151                                                                200
KKELANGNDALRNEDARSPFYSALRNTPSFKERNGGNHDPSRMKAVIYSK

201                                                                250
HFWSGQDRSSSADKRKYGDPDAFRPAPGTGLVDMSRDNIPRSPTSPGEG

251                                                                300
FVNFDYGWFGAQTEADADKTVWTHGNHTHAPNGSLGAMHVYESKFRNWSE

301                                                331
GYSDFDRGAYVITFIPKSWNTAPDKVKQGWP
```

Figure 5. Amino acid sequence of microbial transglutaminase from *Streptoverticillium mobaraense* var. All amino acid residues are represented by one-letter. An asterisk indicates the active center.

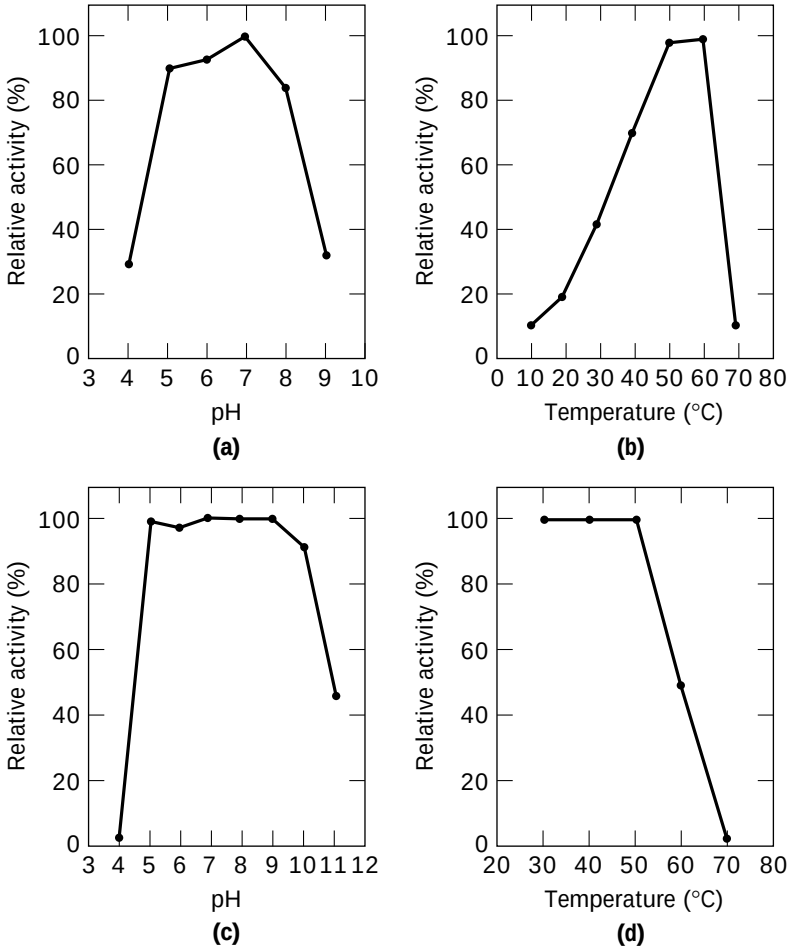


Figure 6. Enzymatic properties of microbial transglutaminase. (**a**) pH-activity profile; (**b**) temperature-activity profile; (**c**) pH stability profile; and (**d**) thermal stability profile.

important in the expression of textures or mouth-feel. Among them, covalent ϵ-(γ-glutamyl)lysine cross-linking is considered noteworthy. An example of such a cross-linking formation is found in the setting phenomenon or "suwari" of salted, ground fish protein pastes in Japanese kamaboko manufacturing. Contribution of such endogenous transglutaminase activity of fish in the suwari process has been well documented (43–45). As demonstrated in the these cases, transglutaminase has been believed to modify functional properties of protein substrates (9). In an attempt to develop new foods with unique protein utilization and processing methodology, we have been interested in the modification of the functional properties of food proteins through formation of ϵ-(γ-glutamyl)lysine bonds by transglutaminase.

Feasibility Study on Protein Modification by Transglutaminases

In the early 1980s, the possibility of modification of functional properties in milk caseins, soybean globulins, whey proteins, and beef, pork, chicken, and fish actomyosins was demonstrated using transglutaminase derived from guinea pig liver (46–59) or bovine plasma factor XIIIa (60,61). When protein solutions are cast on a flat surface, a transparent, water-resistant, and slowly digestible protein film can be prepared. A protein emulsion of the oil-in-water type can also serve as the substrate, and it is gelled by transglutaminase. In these studies, cross-linking between various proteins of different origins, gelation of various proteins by the transglutaminase reaction, as well as incorporation of amino acids or peptides to improve nutritive deficiency was shown. These results indicate that transglutaminase was potentially useful in creating proteins with new, unique functional properties, as previously postulated (9). However, limited supply and the unfamiliarity of guinea pig liver transglutaminase for food use hindered its commercialization.

Protein Cross-Linking

Discovery of MTG drastically accelerated its applicative studies on the alteration of food protein functionalities. MTG gels concentrate solutions of such proteins as soybean protein, milk proteins, gelatin, and beef, pork, chicken, and fish myosins, as does the guinea pig liver enzyme. Subsequently, improvement in solubility, water-holding capacity, and thermal stability was demonstrated. The gelled soybean globulin further hardened by heating at 100 °C, for 15 min, resulting in a protein with new gel properties. Caseins, non-heat-setting proteins, are also gelled by MTG without heating. Gelatin, a cold-setting protein, is also gelled, and in this case, the gelled gelatin no longer melts on heating at 100 °C. In practical application, soybean curd can be prepared by the ordinary process with the aid of MTG. This MTG-treated soybean curd becomes retort resistant in a way that the soybean curd retains its shape and is nonporous even after the retort treatment(s). More than two different proteins can be covalently conjugated by MTG to produce new proteins with novel functionalities, as with guinea pig transglutaminase. For instance, conjugation of caseins or soybean globulins with ovomucin, one of the egg white glycoproteins, improved emulsifying activity (44). Also conjugation of casein and gelatin by transglutaminase yielded novel proteins that were highly soluble in acidic pH regions (41). Practical applications of transglutaminases, including MTG, in food processing and manufacturing are summarized in Fig. 7, and detailed explanations of transglutaminase applications have been given in elsewhere (62–64).

Amino Acids and/or Peptides Incorporation into Proteins

Amino acids and/or peptides can be covalently introduced into substrate proteins by MTG. This improves nutritive values of caseins, soybean proteins, wheat gluten, and so on, whose methionine and lysine, are limiting amino acids. For practical application of amino acid incorporation, all the amino acids, except lysine, must have their α-carboxyl group amidated, esterified, or decarboxylated in order to be substrates of the transglutaminases. On the other hand, lysine with an ϵ-amino group on the side chain is a good substrate for transglutaminase. In the case of peptides, most should contain at least one lysine residue to be a suitable substrate. For instance, lysylmethionine or methionyllysine can be incorporated into caseins to improve the methionine deficiency. Likewise, lysylarginine or arginyllysine can be incorporated into caseins to improve the arginine deficiency in infant prawns. On the other hand, carnosine and β-alanylhistidine can serve as the substrate (65). Carnosine and anserine, a 1-methylated histidine derivative of carnosine, are thus used as inhibitors of protein cross-linking, because both peptides competitively act as the substrates in the transglutaminase cross-linking reaction.

BIOAVAILABILITY OF CROSS-LINKED PROTEINS

Although the scope of application of MTG-catalyzed modification of food proteins is vast, here attention is focused on the nutritional efficiency of the cross-linked proteins. MTG-modified proteins differ from native proteins only in the quantity of the ϵ-(γ-glutamyl)lysine cross-links, all else is the same. However, the digestibility or bioavailability of the ϵ-(γ-glutamyl)lysine moiety in the cross-linked proteins has been intensively investigated.

Distribution of Naturally Occurring ϵ-(γ-Glutamyl)lysine

Distribution of naturally occurring ϵ-(γ-glutamyl)lysine cross-links in raw materials such as animal organs and tissues and in commonly consumed processed foods have been determined by measuring the ϵ-(γ-glutamyl)lysine dipeptide after exhaustive digestion by proteases and peptidases (25,26). Raw food materials, e.g., meats, fish, shellfish, and most processed foods (e.g., kamaboko, ham, stewed beef, fried chicken, grilled pork, and hamburger) contained more or less certain amounts of ϵ-(γ-glutamyl)lysine cross-links (Table 2). However, dairy products (e.g., cheese, milk, and yogurt) scarcely contained these cross-links. The level of ϵ-(γ-glutamyl)lysine cross-links was higher in processed, especially cooked, foods than in raw materials. The ϵ-(γ-glutamyl)lysine cross-link

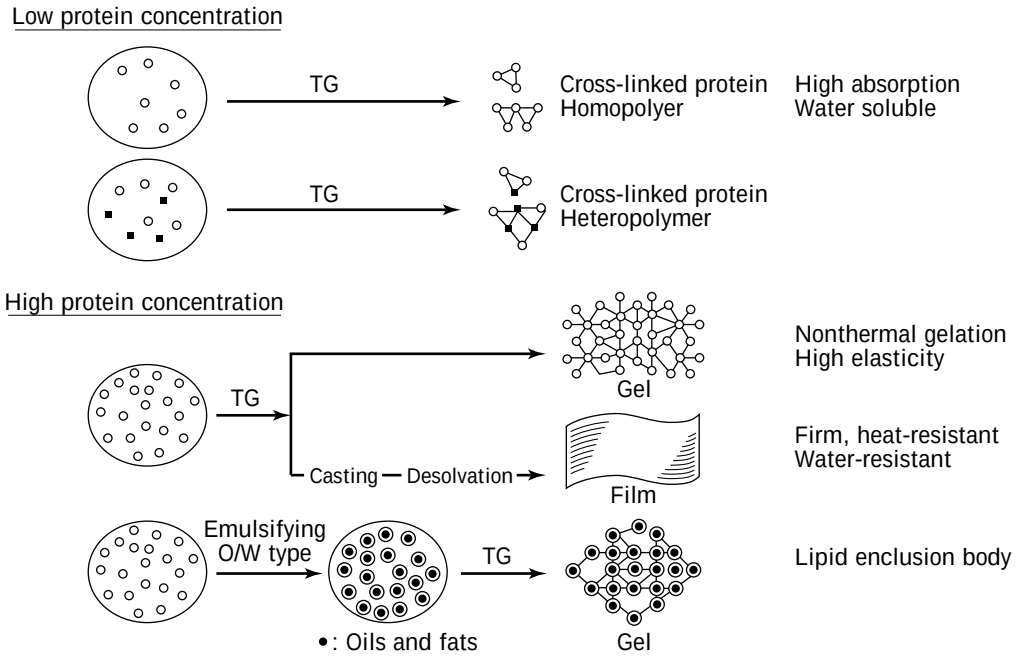


Figure 7. Application of transglutaminase on alteration of food functionalities. Open small circles represent substrate protein molecules. Closed small squares represent other substrate protein molecules. Closed small circles represent oils and/or fats.

Table 2. Distribution of Naturally Occurring ϵ-(γ-Glutamy)lysine Cross-Links

Items		Amount of ϵ-(γ-Glutamyl)lysine[a]
Raw materials	Beef tongue	78.4
	Crab	34.6
	Beef liver	14.5
	Prawn	3.77
	Milk	nd
	Chicken egg	nd
	Tomato	nd
	Apple	nd
	Potato	nd
Processed food	Stewed beef	68.7
	Hamberger	40.9
	Ham	39.4
	Peking duck	31.5
	Bread	15.7
	Cooked rice	11.1
	Popcorn	6.80
	Fried chicken	6.76
	Cheese	nd
	Yogurt	nd

[a]Micromoles of ϵ-(γ-Glutamyl)lysine in 100 g protein of each item.

occurs naturally in eggs of red salmon, lumpfish, herring, sardine, and Alaska pollock (66), all of which have been popularly consumed as caviar in both Western and Eastern countries.

Mechanism of ϵ-(γ-Glutamyl)lysine Formation

It is highly likely that endogenous transglutaminases are responsible in the formation of ϵ-(γ-glutamyl)lysine cross-links in protein during the cooking process, because a variety of living organisms contain intrinsic, endogenous transglutaminase in their tissues and organs. During cooking, the temperature elevation in food materials is in fact very slow, such that the endogenous transglutaminases have some time to exert their enzymatic activity before heat inactivation takes place. Thus cooked or processed foods may contain larger amounts of ϵ-(γ-glutamyl)lysine cross-links than raw materials due to transglutaminase activity. That dairy products contained no ϵ-(γ-glutamyl)lysine cross-links may support an idea that milk itself does not contain any endogenous transglutaminase activities.

Besides the catalysis of the transglutaminases, heating has been reported to induce the formation of ϵ-(γ-glutamyl)lysine cross-links in chemical dehydration between the γ-carboxyl group of glutamate and ϵ-amino group of lysine (67–69). In this respect, humans have been ingesting the ϵ-(γ-glutamyl)lysine moiety since the discovery of cooking. Consequently, the safety of the ϵ-(γ-glutamyl)lysine moiety has been unintentionally and automatically demonstrated.

Degradation In Vitro of ϵ-(γ-Glutamyl)lysine

Ordinarily mammalian gastrointestinal digestive enzymes cleave proteins into amino acids, but leave the ϵ-(γ-glutamyl)lysine cross-link uncleaved after ingestion of cross-linked proteins. The uncleaved ϵ-(γ-glutamyl)lysine dipeptide may be absorbed through intestinal brush-border membranes to the kidney. As a factor responsible for the degradation of the ϵ-(γ-glutamyl)lysine cross-link, a kidney enzyme, γ-glutamylamine cyclotransferase, has been reported to cleave the ϵ-(γ-glutamyl)lysine dipeptide

form. This reaction yields a free lysine and 5-*oxo*-proline (synonym: pyroglutamate) from the dipeptide (70). Because lysine is an essential amino acid for humans and other mammals, the presence of such a pathway would help satisfy nutritional requirements. On the other hand, 5-*oxo*-proline should be converted to glutamate by an ATP-dependent enzyme, 5-*oxo*-prolinase. If a large amount of the cross-linked protein was ingested, a large amount of ATP would be consumed in the kidney. Such a large ATP consumption might overload kidney function, however, recently, γ-glutamyltransferase (γ-glutamyltranspeptidase; EC 2.3.2.2, (71)), present mainly in the intestinal brush-border membranes, kidney, and blood, has been found to cleave the ϵ-(γ-glutamyl)lysine dipeptide into a lysine and glutamate (72). This reaction uses no ATP to generate glutamate. Thus, the ϵ-(γ-glutamyl)lysine moiety, could be effectively metabolized into lysine by the two enzymes.

Degradation In Vivo of ϵ-(γ-Glutamyl)lysine

Many researchers have demonstrated that the ϵ-(γ-glutamyl)lysine dipeptide could be metabolized in rats, and that lysine was integrated in rat tissues (73–76). On the other hand, the bioavailability of the ϵ-(γ-glutamyl)lysine cross-linked moiety in the cross-linked proteins has not yet been experimentally demonstrated. This is because it would be unrealistic and impractical to prepare large-scale cross-linked proteins for animal feeding by using the guinea pig liver transglutaminase. With abundantly available MTG, cross-linked caseins in kilogram scale were fed to rats to evaluate the nutritive value of lysine in the ϵ-(γ-glutamyl)lysine moiety. Lysine generated in the rats has been proved to be nutritionally beneficial in the rats using MTG-catalyzed cross-linked wheat glutens (77). It is thus suggested that the ϵ-(γ-glutamyl)lysine moiety in the cross-linked caseins is cleaved, and the lysine in the moiety metabolically utilized in the body.

CONCLUSION

As stated earlier (9), transglutaminases have been found to drastically alter protein functionalities through the formation of ϵ-(γ-glutamyl)lysine cross-links in many food proteins. Vast application of the transglutaminase modification on food proteins can bring forth such promising results as novel foods and innovative processing techniques. Further, the discovery of transglutaminase from *S. mobaraense* var. (MTG) has accelerated applicative studies, and

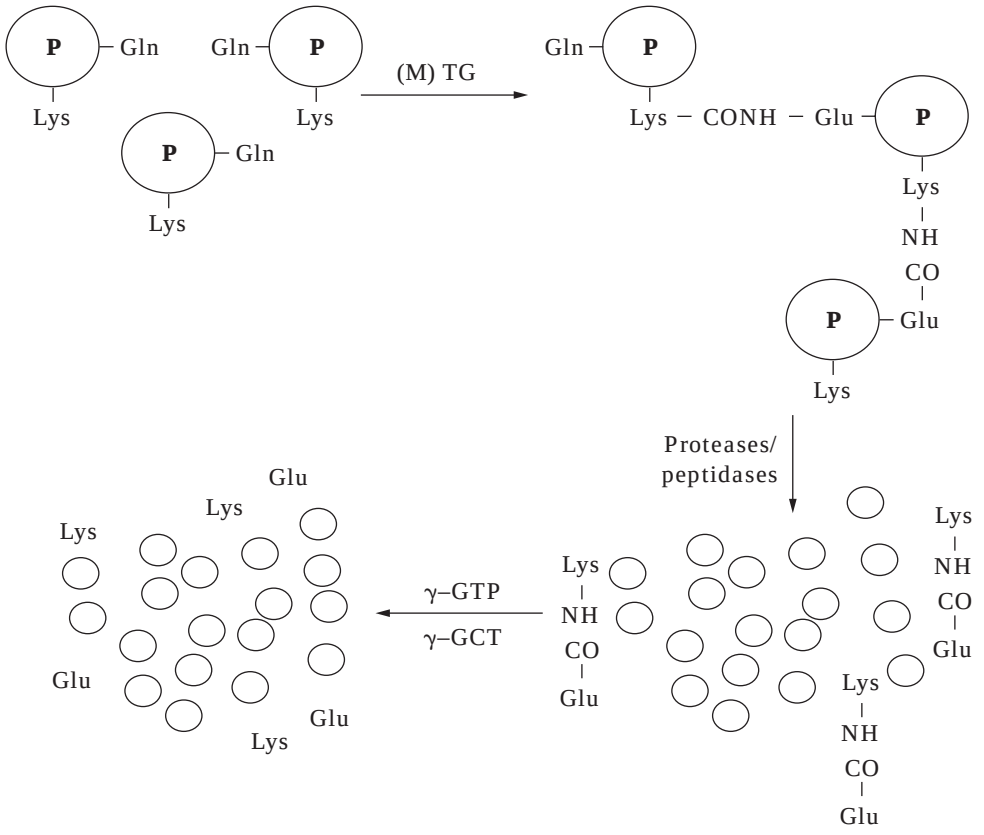


Figure 8. Reaction scheme of the digestion of transglutaminase-catalyzed cross-linked proteins. A large circle with *P* represents a protein or peptide; a small circle represents a single amino acid produced by digestion of proteases and/or peptidases; Glu-CONH-Lys, ϵ-(γ-glutamyl)lysine dipeptide; (M)TG, (microbial) transglutaminase; γ-GTP, γ-glutamyltransferase; γ-GCT, γ-glutamylamine cyclotransferase.

consequently, MTG can be used in surimi and sausage manufacturing in Japan.

With respect to the bioavailability of the ϵ-(γ-glutamyl)lysine moiety in the cross-linked proteins, many studies have been done, and it has been demonstrated that the ϵ-(γ-glutamyl)lysine dipeptide that usually can be produced from the cross-linked protein after exhaustive digestion by gastrointestinal digestive proteases and peptidases can be cleaved by both γ-glutamylamine cyclotransferase and γ-glutamyltransferase, yielding a free lysine. The lysine in the cross-linked proteins has been shown to be metabolizable in the body. The nutritive value of the lysine of the cross-linked protein has been proven using wheat gluten as a lysine-deficient model protein in rats (Fig. 8). Finally, with the evidence that the safety of ϵ-(γ-glutamyl)lysine moiety has been proved by long-established, habitual intakes of cooked foods by humans, the overall effectiveness and safety of protein modification by transglutaminase have been shown.

BIBLIOGRAPHY

1. Nomenclature Committee of the International Union of Biochemistry and Molecular Biology, *Enzyme Nomenclature*, Academic Press, San Diego, Calif., 1992, p. 201.
2. H. Borsook, C.L. Deasy, A.J. Haagen-Smit, G. Keighley, and P.H. Lowy, *J. Biol. Chem.* **184**, 529–543 (1950).
3. J.E. Folk and S.I. Chung, *Adv. Enzymol.* **38**, 109–191 (1973).
4. J.E. Folk, J.S. Finlayson, *Adv. Protein Chem.* **31**, 1–133 (1977).
5. J.E. Folk, *Adv. Enzymol. Relat. Areas Mol. Biol.* **54**, 1–56 (1983).
6. L. Lorand and S.M. Conrad, *Mol. Cell Biochem.* **58**, 9–35 (1984).
7. C.S. Greenberg, P.J. Birckbichler, and R.H. Rice, *FASEB J.* **5**, 3071–3077 (1991).
8. D. Aeschlimann and M. Paulsson, *Thromb. Haemost.* **71**, 402–415 (1994).
9. J.R. Whitaker, *Food Proteins—Improvement through Chemical and Enzymatic Modification*, American Chemical Society; Washington, D.C., 1977, pp. 95–105.
10. H. Ando, M. Adachi, K. Umeda, A. Matsuura, M. Nonaka, R. Uchio, H. Tanaka, and M. Motoki, *Agric. Biol. Chem.* **53**, 2613–2617 (1989).
11. H. Araki and N. Seki, *Nippon Suisan Gakkaishi*, **59**, 711–719 (1993).
12. Y. Kumazawa, K. Seguro, M. Takamura, and M. Motoki, *J. Food Sci.* **58**, 1062–1064 (1993).
13. I. Kimura, M. Sugimoto, K. Toyoda, N. Seki, K. Arai, and T. Fujita, *Nippon Suisan Gakkaishi* **57**, 1389–1396 (1991).
14. K. Sano, Y. Kumazawa, H. Yasueda, K. Seguro, and M. Motoki, *Fifth International Conference on Transglutaminase and Protein Crosslinking Reactions: Book of Abstracts*, 1996, p. 65.
15. F. Tokunaga, M. Yamada, T. Miyata, Y.L. Ding, M. Hiranaga-Kawabata, T. Muta, and S. Iwanaga, *J. Biol. Chem.* **268**, 252–261 (1993).
16. M. Ando, Y. Nagata, and R. Hammerschlag, *Zool. Sci.* **8**, 721–727 (1991).
17. M. Signorini, S. Beninati, and C.M. Bergamini, *J. Plant Physiol.* **137**, 547–552 (1991).
18. S.A. Margosiak, M.E. Alavarets, D. Louie, and G.D. Kuehn, *Plant Physiol.* **92**, 88–96 (1990).
19. D. Serafini-Fracassini, S.D. Duca, and P. Torrigiani, *Plant Physiol. Biochem.* **27**, 659–669 (1989).
20. I. Icekson and A. Apelbaum, *Plant Physiol.* **84**, 972–974 (1987).
21. J.D. Klein, E. Guzman, and G.D. Kuehn, *J. Biocteriol.* **174**, 2599–2605 (1992).
22. M.V. Ramanujam and J.H. Hageman, *FASEB J.* **168**, A3630 (1990).
23. K. Washizu, K. Ando, S. Koikeda, S. Hirose, A. Matsuura, H. Takagi, M. Motoki, and K. Takeuchi, *Biosci. Biotechnol. Biochem.* **58**, 82–87 (1994).
24. H. Waelsch and M.J. Mycek, *Methods Enzymol.* **5**, 833–838 (1962).
25. M. Griffin, J. Wilson, and A. Lorand, *Anal. Biochem.* **124**, 406–413 (1982).
26. H. Sakamoto, Y. Kumazawa, H. Kawajiri, and M. Motoki, *J. Food Sci.* **60**, 416–419 (1995).
27. T. Ohtsuka, K. Seguro, and M. Motoki, *J. Food Sci.* **61**, 81–84 (1996).
28. S.I. Chung and J.E. Folk, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 303–307 (1972).
29. Nomenclature Committee of the International Union of Biochemistry and Molecular Biology, *Enzyme Nomenclature*, Academic Press, San Diego, Calif., 1992, p. 427.
30. T.J. Hornyak, P.D. Bishop, and J.A. Shafer, *Biochemistry* **28**, 7326–7332 (1989).
31. V.C. Yee, L.C. Pedersen, I.L. Trong, P.D. Bishop, R.E. Stenkamp, and D.C. Teller, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7296–7300 (1994).
32. Y. Takeuchi, P.J. Birckbichler, M.K. Patterson, Jr., and K.N. Lee, *FEBS*, **307**, 177–180 (1992).
33. H. Nakaoka, D.M. Perez, K.J. Beak, T. Das, A. Husain, K. Misono, M.J. Im, and R.M. Graham, *Science* **264**, 1593–1596 (1994).
34. T. Kanaji, H. Ozaki, T. Takao, H. Kawajiri, H. Ide, M. Motoki, and Y. Shimonishi, *J. Biol. Chem.* **268**, 11565–11572 (1993).
35. M. Nonaka, H. Tanaka, A. Okiyama, M. Motoki, H. Ando, K. Umeda, and A. Matsuura, *Agric. Biol. Chem.* **53**, 2619–2623 (1989).
36. M. Nonaka, H. Sakamoto, S. Toiguchi, H. Kawajiri, T. Soeda, and M. Motoki, *J. Food Sci.* **57**, 1214–1218 (1992).
37. Y.P. Huang, K. Seguro, M. Motoki, and K. Tawada, *J. Biochem.* **112**, 229–234 (1992).
38. H. Sakamoto, Y. Kumazawa, S. Toiguchi, K. Seguro, T. Soeda, and M. Motoki, *J. Food Sci.* **60**, 300–304 (1995).
39. I.J. Kang, Y. Matsumura, K. Ikura, M. Motoki, H. Sakamoto, and T. Mori, *J. Agric. Food Chem.* **42**, 159–165 (1994).
40. E. Kim, M. Motoki, K. Seguro, A. Muhlrad, and E. Reisler, *Biophys. J.* **69**, 2024–2032 (1995).
41. M. Nonaka, Y. Matsuura, and M. Motoki, *Biosci. Biotechnol. Biochem.* **60**, 131–133 (1996).
42. A. Kato, T. Wada, K. Kobayashi, K. Seguro, and M. Motoki, *Agric. Biol. Chem.* **55**, 1027–1031 (1991).
43. G.G. Kamath, T.C. Lanier, E.A. Foegeding, and D.D. Hamann, *J. Food Biochem.* **16**, 151–172 (1992).
44. K. Sato, Y. Tsukamasa, C. Imai, K. Ohtsuki, Y. Shimizu, and M. Kawabata, *J. Agric. Food Chem.* **40**, 806–810 (1992).
45. J. Wan, I. Kimura, M. Satake, and N. Seki, *Fisheries Sci.* **60**, 107–113 (1994).

46. K. Ikura, T. Kometani, M. Yoshikawa, R. Sasaki, and H. Chiba, *Agric. Biol. Chem.* **44**, 1567–1573 (1980).
47. K. Ikura, T. Kometani, R. Sasaki, and H. Chiba, *Agric. Biol. Chem.* **44**, 2979–2984 (1980).
48. K. Ikura, M. Yoshikawa, R. Sasaki, and H. Chiba, *Agric. Biol. Chem.* **45**, 2587–2592 (1981).
49. K. Ikura, M. Goto, M. Yoshikawa, R. Sasaki, and H. Chiba, *Agric. Biol. Chem.* **48**, 2347–2354 (1984).
50. K. Ikura, K. Okumura, M. Yoshikawa, R. Sasaki, and H. Chiba, *Agric. Biol. Chem.* **49**, 1877–1878 (1985).
51. N. Nio and M. Motoki, *J. Food Sci.* **48**, 561–566 (1983).
52. N. Nio, M. Motoki, and K. Takinami, *Agric. Biol. Chem.* **48**, 1257–1261 (1984).
53. N. Nio, M. Motoki, and K. Takinami, *Agric. Biol. Chem.* **49**, 2483–2286 (1985).
54. N. Nio, M. Motoki, and K. Takinami, *Agric. Biol. Chem.* **50**, 851–855 (1986).
55. N. Nio, M. Motoki, and K. Takinami, *Agric. Biol. Chem.* **50**, 1409–1412 (1986).
56. M. Motoki, K. Seguro, N. Nio, and K. Takinami, *Agric. Biol. Chem.* **50**, 3025–3030 (1986).
57. N. Nio, M. Motoki, and K. Takinami, *Agric. Biol. Chem.* **51**, 237–239 (1987).
58. M. Motoki, H. Aso, K. Seguro, and N. Nio, *Agric. Biol. Chem.* **51**, 993–996 (1987).
59. M. Motoki, H. Aso, K. Seguro, and N. Nio, *Agric. Biol. Chem.* **51**, 997–1002 (1987).
60. L. Kurth and P.J. Rogers, *J. Food Sci.* **49**, 573–576, 589 (1984).
61. C.D. Backer-Royer, F. Traore, and J.C. Meunier, *J. Agric. Food Chem.* **40**, 2052–2058 (1992).
62. P.M. Neilsen, *Food Biotechnol.* **9**, 119–156 (1995).
63. Y. Zhu, A. Rinzema, J. Tramper, and J. Bol, *Appl. Microbiol. Biotechnol.* **44**, 277–282 (1995).
64. H. Singh, *Trends Food Sci. Technol.* **8**, 196–200 (1991).
65. J. Wan, I. Kimura, and N. Seki, *Fisheries Sci.* **61**, 968–972 (1995).
66. Y. Kumazawa, H. Sakamoto, H. Kawajiri, K. Seguro, and M. Motoki, *Fisheries Sci.* **62**, 331–332 (1996).
67. R.S. Asquith, M.S. Otterburn, and W.J. Sinclair, *Angew. Chem. Int. Ed.* **13**, 514–520 (1974).
68. R.F. Hurrell, K.J. Carpenter, W.J. Sinclair, M.S. Otterburn, and R.S. Asquith, *Br. J. Nutr.* **35**, 383–395 (1976).
69. M. Otterburn, M. Healy, and M. Sinclair, *Adv. Exp. Med. Biol.* **86B**, 239–262 (1977).
70. M.L. Fink, S.I. Chung, and J.E. Folk, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 4564–4568 (1980).
71. A. Meister, S.S. Tate, and O.W. Griffith, *Methods Enzymol.* **77**, 237–253 (1981).
72. K. Seguro, Y. Kumazawa, T. Ohtsuka, H. Ide, N. Nio, M. Motoki, and K. Kubota, *J. Agric. Food Chem.* **43**, 1977–1981 (1995).
73. G. Raczynski, M. Snochowski, and S. Buraczewski, *Br. J. Nutr.* **34**, 291–296 (1975).
74. P.A. Finot, F. Mottu, E. Bujard, and J. Mauron, *Nutritional Improvement of Foods and Food Proteins*, Plenum, London, 1978, pp. 549–570.
75. K. Iwami and K. Yasumoto, *J. Sci. Food Agric.* **37**, 495–503 (1986).
76. M. Friedman and P.A. Finot, *J. Agric. Food Chem.* **38**, 2011–2020 (1990).
77. K. Seguro, Y. Kumazawa, C. Kuraishi, H. Sakamoto, and M. Motoki, *J. Nutr.*, in press.

See also BIOCATALYSIS DATABASES.

TRANSIENT EXPRESSION SYSTEMS

A.R. BERNARD
H.D. BLASEY
Serono Pharmaceutical Research Institute
Geneva, Switzerland

KEY WORDS

Baculovirus
BHK
COS
DNA transfection
HEK293EBNA
Insect cell culture
Large-scale transient expression
Mammalian cell culture
Recombinant protein
Semliki Forest virus
Sf9
Virus expression vectors

OUTLINE

OVERVIEW

Historically, the search for new genes and new gene functions has been the major driving force for the development of several powerful transient expression systems such as direct DNA transfection into mammalian cells, recombinant virus infection, injection of *Xenopus* oocytes with mRNA, or complete in vitro–coupled transcription–translation. In particular, functional expression of large cDNA libraries has allowed the cloning of many new genes. The analysis of the gene products, however, has lagged behind because of the lack of efficient, quick, and easy ways of producing sufficient amounts of the gene product to be studied. On another front, methods for stable expression of genes in mammalian cells have been available for many years and presently form the basis for the development of biopharmaceutical production in cell culture. One of the major drawbacks of stable expression lies in the time required for the construction of an optimal, stable recombinant cell line. Transient expression of human or mammalian genes has received considerable interest in recent years for its use in producing recombinant proteins, mainly for research but potentially also for clinical purposes. The ability to rapidly and reproducibly express one or several recombinant genes in a eucaryotic cell environment and to produce large amounts of the gene product now complements the conventional set of biotechnological tools. This article reviews the transient expression methods that can be used on a large scale for the production of recombinant proteins of pharmaceutical interest for in vitro biological, biochemical, animal, and human studies. The various methods presently available can be classified according to the vector used: plasmid DNA, insect viruses, or mammalian viruses. Several of the vectors used in transient expression are also used in gene therapy applications, but we will restrict the scope of this review to the production of recombinant proteins in cell cultures.

There are several key features of transient expression systems:

- Simplicity, in particular in the construction of the expression vector
- Intrinsic genetic stability, because the gene is constantly maintained in the expression host
- Speed, that is, the time required to obtain the first quantities of the gene product
- Application to a wide range of host cells (except in the case of the baculovirus system)
- Suitability to parallel processing, that is, the ability to study several genes or mutants at the same time

All of the transient expression methods have historically been designed for small-scale operations (Petri dish or tissue culture flasks). However, the emphasis of recent developments has been on the scale-up and application of these laboratory methods to large-scale production systems, and all methods can now be operated at pilot bioreactor scale.

DIRECT DNA TRANSFECTION

Introduction

Direct DNA transfection is a widely used system to efficiently produce a desired gene product at levels of 1–3 mg/L within a few weeks. Both genomic DNA and cDNA can be expressed using the classical COS host cell line. This allows the generation of large amounts of protein material with mammalian posttranslational modifications (glycosylation, disulfide bridges, carboxylation, amidation). The material produced can then be utilized for studying the structure and function of the gene product. In parallel, COS cells are routinely used for studying gene expression regulation and for expression cloning. The cells are readily available from public collections (ATCC, ECACC) and are efficiently transfected. Cultivation of these adherent cells does not pose any major problem, even for large-scale cultures. These properties have contributed to the wide popularity of this expression system.

COS cells are derived from green monkey kidney cells (CV-1) transformed with an origin-defective SV40 virus. Historically, the cell line was developed by Yakov-Gluzman (1) to allow the generation of SV40 virus mutants. He developed three cell lines: COS-1, COS-3, COS-7. The cells express high levels of the wild-type SV40 large tumor (T) antigen, an early gene product of 100 bp, which binds to the vector DNA at the origin of replication and allows the host-cell polymerase to run repetitive cycles of DNA replication. As a consequence, plasmids carrying the SV40 origin are maintained in the cell at high copy number (10^4–10^5 copies per cell), giving rise to abundant mRNA yielding high-level recombinant gene expression. COS cells are known to express for up to 1 week posttransfection, with peak protein production between 48 and 72 h posttransfection.

Lusky and Botchan in 1981 (2) were the first to use the large T antigen expressing COS in combination with an SV40 origin–containing plasmid for the identification of DNA transcription sequences. In 1982, the COS cell expression system was used for the first time by Rose and Bergmann to express cell surface and secreted proteins such as vesicular stomatitis virus glycoprotein (3). Subsequent applications by numerous other investigators demonstrated that the COS cell expression system generally yields biologically active proteins. The main factors influencing the expression level in this system are the expression plasmid construct and the transfection efficiency.

Today, many SV40 T antigen–dependent vectors (e.g., ATCC 37192, 37193, 53100) are available for gene expression with COS cells. Typical plasmids for COS transient expression carry the following major components (Fig. 1):

1. The SV40 origin for plasmid replication
2. Appropriate promoter, enhancer, and polyadenylation sequences
3. A prokaryotic origin of replication
4. A prokaryotic selectable marker
5. Unique restriction sites for target gene subcloning

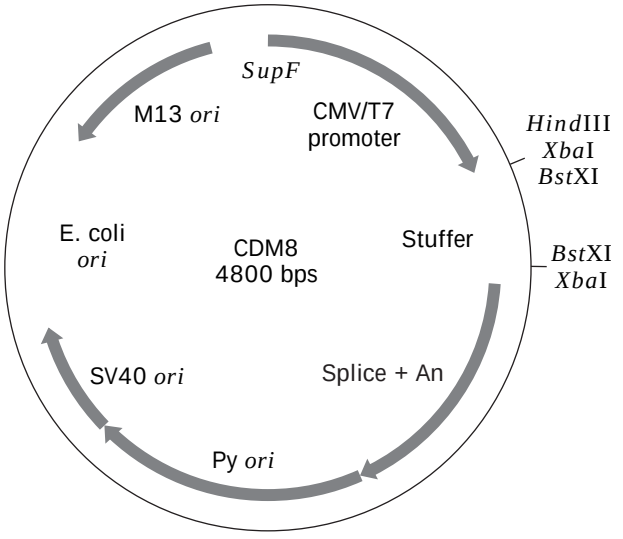


Figure 1. Typical COS expression vector. CDM8, a typical plasmid for transient expression in COS cells. The plasmid consists of CMV/T7 RNA polymerase promoter for high-level gene expression; (SV40 *ori*) SV40 and (Py) polyoma origin of replication; M13 origin of replication; a stuffer with unique restriction sites for subcloning; *E. coli* origin of replication; prokaryotic genetic marker (*SupF*); and splice and An, eukaryotic transcription regulatory elements. *Source*: Reprinted from Ref. 4, with permission of *Nature*.

The Epstein–Barr virus (EBV) also gave rise to a popular expression system based on vectors carrying the replication origin region (OriP) from EBV. EBV–transformed cell lines expressing the EBV nuclear antigen 1 (EBNA-1) replicate transfected plasmids and express the subcloned gene in a similar way to the COS cell–based system. Both genomic and cDNA sequences can be expressed in this system. EBV vectors replicate episomally in the nucleus and, under certain conditions, can be maintained by the cells upon prolonged passage, thus providing a way of obtaining either transient or stable expression of the recombinant protein. In addition, incorporation of the coding sequence for EBNA-1 on the expression vector yields a self-replicating expression plasmid that is autonomously maintained under selection pressure.

By maintaining transfected cells under selective pressure with hygromycin, 293 cells expressing EBNA-1 can be passaged for up to six months while still expressing recombinant proteins at a high level. For example, human placental alkaline phosphatase, tissue plasminogen activator, vascular endothelial cell growth factor, hepatocyte growth factor, prolactin, neurotrophins, and soluble tumor necrosis factor receptors were obtained at levels of 0.2–10 μg/mL (5).

Basic Principles and Protocols

COS Cells and Transfection. COS cell lines are available from the ATCC (COS-1:ATCC CRL 1650, and COS-7: ATCC CRL 1651). They are normally grown adherent onto tissue culture treated matrices in DMEM with 10% FCS and are split by trypsinization before reaching confluence to avoid aggregate formation. ATCC recommends handling the cells in accordance with Biosafety Level 2. Various transfection methods are available to introduce plasmids into mammalian cells. Calcium phosphate precipitation, DEAE dextran, lipofection, and electroporation are commonly used transfection procedures that are well established in cell culture laboratories. In addition, other methods have been described, such as receptor-mediated endocytosis, protoplast fusion, polylysine, and polybrene-mediated transfection. Direct DNA injection into cells with the gene gun (bombardment with DNA-coated gold particles) is used in special applications where cells are difficult to transfect or for certain electrophysiology studies.

Calcium Phosphate. DNA uptake by mammalian cells can be achieved by exposing cells to a calcium phosphate–DNA coprecipitate. The precipitate is formed within 20–30 min by diluting a $CaCl_2$/DNA mix into 2$\times$ HBS. It is then added to the cells in culture, and some protocols describe replacing the medium after about 10 min with fresh medium. In all cases, the transfection mixture is removed from the culture within 1 day after precipitate addition and replaced by fresh culture medium, allowing 2–3 days for gene expression.

Calcium phosphate transfection is also applicable to cells grown in suspension. After a wash with PBS, the cells are resuspended and incubated in a solution containing the calcium phosphate–DNA coprecipitate. After fresh medium top-up and a 1-day incubation, the cells are then pelleted and transferred into fresh medium. Several parameters for transfection with calcium phosphate are critical, including the time allowed for formation of the calcium phosphate–DNA coprecipitate and the pH at which this operation is performed. Scaling-up this method requires that the standard protocols be adapted to reduce or avoid the incubation, washing, and separation steps, because these steps are not compatible with large volumes.

Lipofection. DNA can be delivered to cells by liposome-mediated transfection. Positively charged liposomes are formed with various cationic lipids, such as for *N*-[1(2,3-diolexyloxy)propyl]-*N*,*N*,*N*-trimethylammonium chloride (DOTMA), and interact with DNA to produce DNA–liposome complexes. These complexes are adsorbed to and subsequently fuse with the lipid bilayer of cell membranes, delivering DNA into the cytoplasm.

The experimental protocols are relatively simple: plasmid DNA is diluted into serum free medium and mixed with preformed cationic liposomes. For the actual transfection, cells are overlaid with this mixture for several hours. Finally, fresh medium is added, and the culture is continued for a few days during which expression of the target gene is normally observed. No developments in large-scale transfections have been reported to date, probably because of the high costs of the liposome starting material required for lipofection. However, lipofection has been successfully carried out in spinner systems (Blasey et al., unpublished observation), which indicates the potential for scale-up.

DEAE-Dextran. One of the most widely utilized methods for the transfection of COS cells is the DEAE-dextran

transfection method. The method is based on the formation of a complex between DNA and high molecular weight DEAE-dextran, followed by the natural uptake of this complex by the cells.

COS cells are first grown up to 50–70% confluence, and medium is exchanged for the transfection medium, which is based on DMEM medium supplemented with 10% NuSerum (Collaborative Research, USA). Normal serum is substituted by NuSerum because of the formation of toxic precipitates. The transfection medium also contains DNA at about 2 μg/mL, DEAE-dextran and chloroquine. Chloroquine is believed to prevent DNA degradation induced by the acidic pH of the endosomes during transport. A DMSO shock is also normally applied to the cells to increase the transfection efficiency. Cells are finally plated out and 24 h after transfection are subjected to either a medium exchange, or a split. The cells or the secreted product are then harvested 24–96 h posttransfection. The major drawback of this method with respect to scale-up consists of the repeated medium changes and washing steps. In addition, the exact timing for the DMSO shock needs to be empirically determined for each transfection.

Electroporation. Transfecting cells by electroporation is achieved by imposing a short, powerful electrical field, (i.e., a pulse) onto the cells. The pulse is generated by discharging a capacitor, and as a consequence, small pores are created into the cell membrane, allowing DNA to enter the cell by diffusion.

Practically, cells are grown to submaximal density, pelleted by centrifugation, and suspended into a conductive medium (e.g., normal culture medium or PBS) together with the plasmid DNA. The mixture is transferred to a sterile electroporation cuvette built with two electrodes embedded in two opposing vertical walls. The electrical pulse is applied, and the cells are then diluted back into fresh medium. Following are the parameters that determine the transfection efficiency:

- The field strength (V/cm), defined by the voltage applied and the distance between the cuvette electrodes
- The pulse shape, which can be square or with a logarithmic decay
- The time constant, which is directly related to the type of the electroporation cuvette
- The type of conductive medium and the cell density
- The concentration of DNA
- The number of electrical pulses applied

The optimal settings of these parameters (Table 1) vary between different cell types and other factors such as the pre- and posttransfection incubation times with the DNA, the temperature at which the electroporation is performed, and the physiological state of the cells. Electroporation has been used to transfect COS cells at the liter scale, however, the process was carried out in batch mode (i.e., by electroporating a large number of cuvettes), and this operation mode clearly limits further scale-up. Commercially available continuous electroporation systems are available, but results using these systems have not been reported.

Results

Transient Expression Scale-Up with COS Cells. Large-scale transient expression with plasmids carrying the SV40 origin includes several steps, shown in Figure 2: production of large DNA quantities, scale-up of cell growth, transfection, and protein production.

Production of Large Quantities of DNA. After transformation of competent *Escherichia coli* cells, colony expansion, and propagation in fermentors, the required gram amounts of plasmid DNA can be obtained. Plasmid DNA is first extracted by conventional alkaline lysis and then purified by selective DNA adsorption onto commercial resins (Qiagen, Germany). One of the main goals of plasmid purification is to achieve a low final endotoxin level, which is essential for a high transfection efficiency.

Culture Scale-Up. COS cells, can be grown up to large quantities by any of several appropriate methods, such as tissue culture flasks, roller bottles, microcarrier cultures, or Cellcube, but clearly, only the last two provide an efficient means of large-scale cell production.

Large-Scale Transfection. COS cells are collected, suspended at high density (10^8 cells/mL) in 0.5 mL of electroporation buffer (see earlier) containing 80 μg plasmid DNA/mL, transferred into electroporation cuvettes (interplate distance of 4 mm), and electroporated (Biorad Gene Pulser) with 2 pulses at 260 V, 960 μF, and infinite resistance.

Protein Production. The transfected culture is allowed to recover for 1 day in FCS-containing medium, which is then exchanged for low-protein, serum-free medium for the protein production phase. Daily medium exchanges allow one to continue the culture for up to 10 days posttransfection. A typical run yields several milligrams of pure protein (6).

Alternatively, COS cells can be grown up in rollerbottles, and the culture is overlaid with a DEAE transfection mix. After applying a DMSO shock, medium is added, and the production is carried out for several days. This method was reported for producing, at 800-mL scale, 1–2 mg/L of the scFv::IgC-kappa fusion protein (7).

Scale-Up of the Calcium Phosphate Transfection with HEK Cells. HEK cells grown in stirred systems can be transfected by the calcium phosphate method. HEK cells previously adapted to suspension growth are propagated in a stirred bioreactor. Once the cell density reaches 3×10^5 cells/mL, the calcium concentration in the medium is elevated to 10 mM. After 30 min transfection is initiated by addition of 20 mL/L of the calcium phosphate/DNA precipitate (at 500 μg DNA/L); 4 h later, calcium phosphate is diluted by simple addition of fresh culture medium, thus avoiding a separation step. In a typical example, 0.5 mg/L recombinant tPA was produced during a 5-day production period (8).

BACULOVIRUS–INSECT CELL EXPRESSION SYSTEM

Introduction

Baculovirus: Overview. The baculovirus–insect cell expression system is based on the infection of insect cell lines

Table 1. Typical Electroporation Conditions (used in our laboratory)

Cell line	CHO Ta	HEK293EBNA	COS M6
Voltage (V)	230	230	260
Capacitance (μF)	960	960	960
Volume (mL)	0.25	0.25	0.5
Resistance (ω)	∞	∞	∞
Cuvette: plate distance (mm)	4	4	4
Plasmid (μg/cuvette)	40	40	30
Electroporation buffer	Basal medium with 10% FCS	Basal medium with 10% FCS	Hanks buffered salt solution
Number of pulses	2	1	1–2
Typical results (% positive)	35–40%	35–40%	40–70%

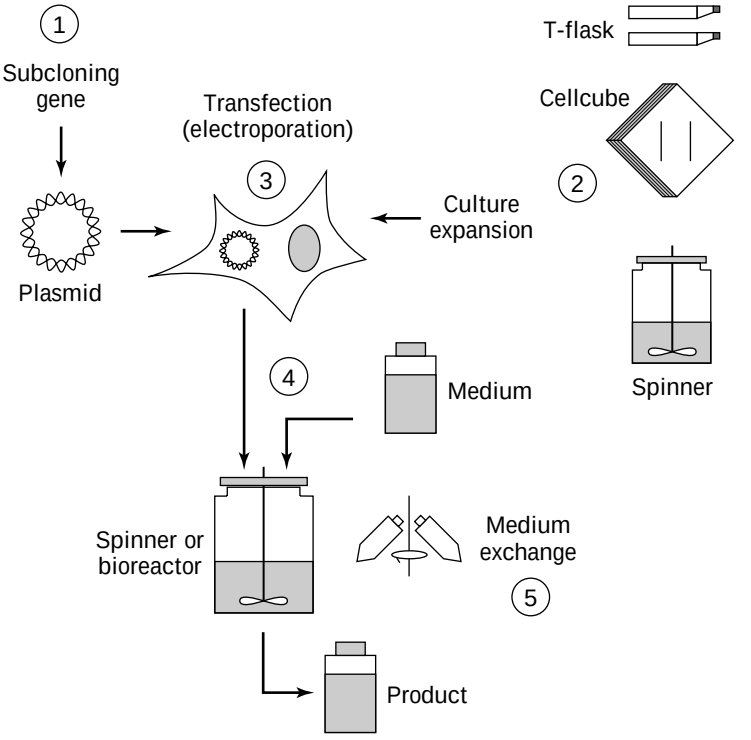


Figure 2. Large-scale COS expression. (1) The gene to be expressed is inserted into the multiple cloning site of the vector. COS cells are grown up in large amounts (2) and then transfected with the plasmid (3) by electroporation. After transfer into a spinner or bioreactor (4) the transfected culture resumes growth and produces the recombinant protein (4). Medium changes (5) extend culture life and allow long-term production (up to 10 days).

with a eucaryotic DNA virus: *Autographa californica* nuclear polyhedrosis virus (AcNPV). A powerful expression vector was originally derived from AcNPV by Summers (9) and Miller (10). It is helper independent and leads to very high expression of heterologous proteins. The expression system exploits strong viral promoters of nonessential genes such as the polyhedrin gene. It is also particularly well suited for coexpression of multisubunit proteins or complex, multicomponent assemblies such as viral capsids. The system is extremely safe because the virus does not replicate in mammalian cells and, in its nonoccluded form, cannot survive in the environment.

Virus Structure and Natural Life Cycle. AcNPV, a member of the Baculoviridae family is a large virus that infects mostly insects. The capsids are 40–50 nm in diameter and 200–300 nm in length. This large size explains in part why large insertions can be made into the viral DNA, a double-stranded circle of approximately 130 kb. The natural life cycle starts from polyhedra or occlusion bodies where virion particles are embedded into a crystalline protein matrix of polyhedrin (29 kDa). These polyhedra are ingested as part of their normal diet by insect larvae. The alkaline pH of the larva's mid-gut dissolves the polyhedrin matrix, releasing up to 100 virion particles per occlusion body. These particles fuse with the mid-gut epithelial cell plasma membranes, migrate to the nucleus where they replicate, and generate nonoccluded (NOV) or budded virus (BV). These budded virus particles are transported by the larva hemolymph to other tissues where they initiate secondary infection sites. In the secondary infection, virus penetrates the cells by receptor-mediated adsorptive endocytosis, a process that is also used during in vitro infection of cells in culture. The virus then infects other susceptible cells, is transported to the nucleus, and enters into a second round of replication that produces NOV again but also, later, occlusion bodies (OB), which close the cycle (Fig. 3). The polyhedrin present in the occlusion bodies can represent up to 50% of the total cellular protein at the peak of production.

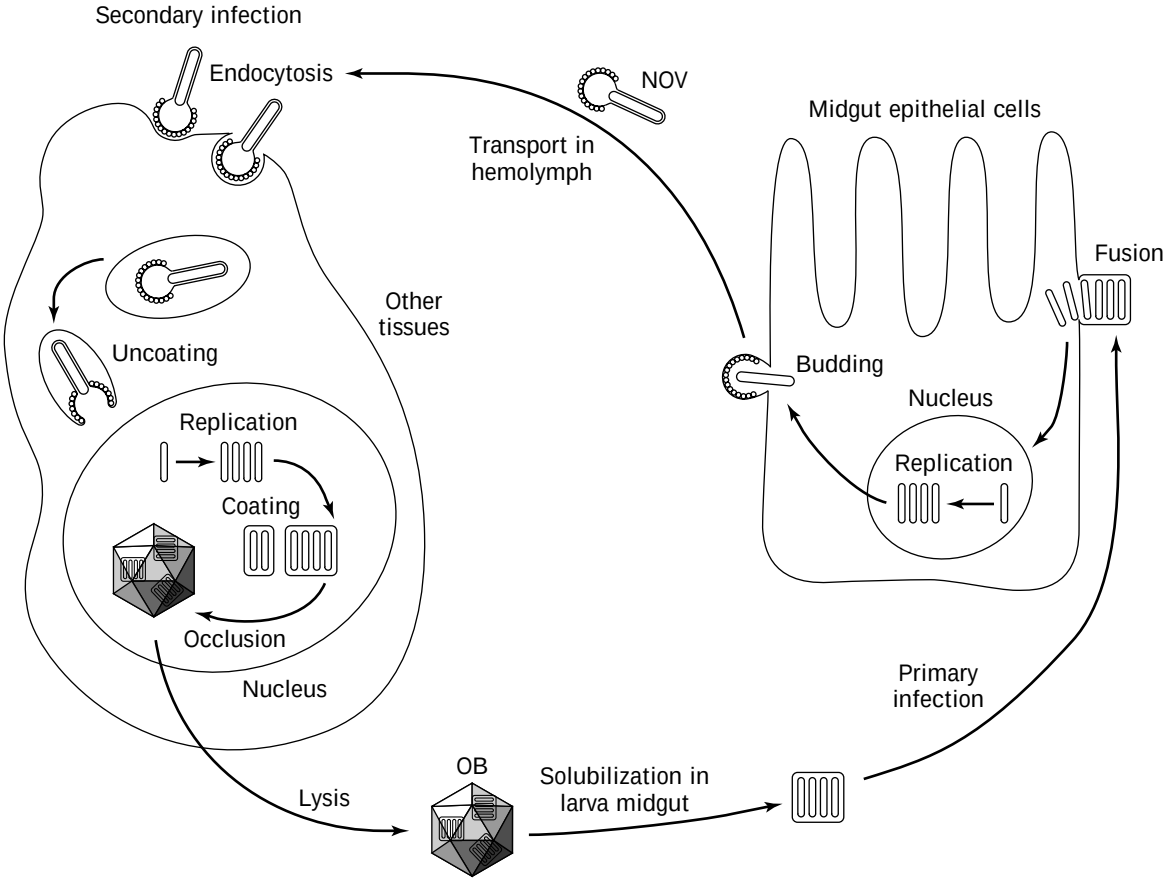


Figure 3. Baculovirus life cycle. The natural life cycle begins with primary infection: the insect larva ingests polyhedra or occlusion body (OB) particles as part of its normal diet. The alkaline pH of the larva's midgut dissolves the polyhedrin matrix. The released virion particles fuse with the midgut epithelial cell plasma membranes, migrate to the nucleus where they start replicating, and generate budded or nonoccluded (NOV) virus particles. NOVs are transported by the larva hemolymph to other tissues where they initiate secondary infection sites. In the secondary infection, virus penetrates the cells by endocytosis, uncoats, is transported to the nucleus, and enters into a second round of replication, which again produces OBs and closes the cycle.

Basic Principles and Protocols

Cell Lines and Media. Several lepidopteran cell lines are highly susceptible to AcNPV, but, two are most commonly used. One is derived from the fall armyworm, *Spodoptera frugiperda* (Sf9 clone), and is available from ATCC (CRL1711), the other is isolated from the cabbage looper, *Trichoplusia ni*, and commonly designated Tn5 (available as High Five cells from Invitrogen Corp.) (11). There are in fact, almost as many variants of these cell lines as laboratories working with them, and extreme caution should be taken when reporting or comparing results without a detailed knowledge of the particular cell line employed. All cell lines have very limited requirements for growth in tissue culture flasks (27 °C, no CO_2, adaptable to serum-free media), and most of them can be readily adapted to suspension culture.

Both cell lines can be used for the expression and production of recombinant proteins, but Sf9 is better suited for generating virus stocks and performing plaque assays because of its natural ability to stick to plastic, its higher susceptibility to the virus, and its ability to produce more virus particles per cell (Fig. 4).

Several media have been developed extensively for efficient growth and expression with the two cell lines described. Serum-containing media such as TNM-FH, IPL-41, and TC-100 are presently used less and less at the production stage, whereas serum-free and even protein-free media are available from commercial sources (JRH Biosciences, Gibco-BRL, Bio-Whittaker). Serum-free media offer clear advantages for large-scale manufacturing, particularly in the case of secreted proteins. When using serum-free media, cells should not be passaged for long periods of time (not more than 40–50 passages) to avoid gradual loss of susceptibility to the virus. Frozen stocks should thus be established to allow rapid and frequent expansion.

Transfer Vectors and Recombinant Virus Construction. In order to produce a recombinant baculovirus, and because the baculovirus genome is too large to be manipulated di-

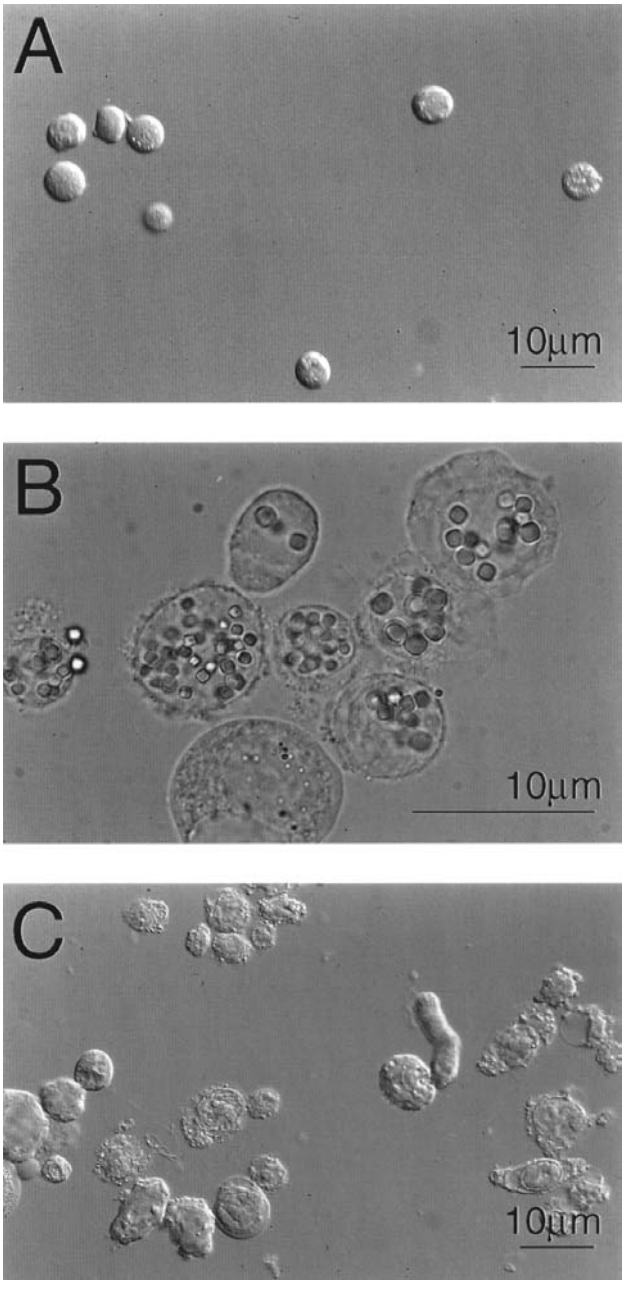


Figure 4. Noninfected Sf9 cells (**a**) and Sf9 cells infected with either WT (**b**) or with a recombinant virus (**c**). Noninfected cells appear round with smooth membranes; WT virus–infected cells contain many occlusion bodies in the nucleus, and recombinant (polyhedrin negative) virus–infected cells display an irregular, sometimes "sausage" shape with rough membranes.

rectly by restriction enzymes, the gene of interest needs to be inserted into a transfer vector (Figs. 5 and 6). Two different types of transfer vector are routinely used, depending on the recombination mechanism chosen:

Transfer Vectors for Recombination in Cells. Once assembled, the transfer vector generally contains one of the strong viral promoters, such as polyhedrin or p10, followed by the recombinant gene coding sequence and transcription termination signals and is flanked on both sides by AcNPV DNA sequences. These sequences allow homologous recombination with wild-type, linearized DNA. A wide variety of transfer vectors of this type have been constructed over recent years (12,13) and offer the possibilities of varying promoters, of multiple gene inserts, and of insertion of tags, fusion, partners, secretion signals, and other helpful features aimed at facilitating detection, expression, and purification.

The transfer vector and wild-type baculovirus DNA are then cotransfected into cells, and the foreign gene coding cassette is inserted into the viral DNA by homologous recombination. Because homologous recombination occurs at a low frequency (1–5%), the virus preparation needs to be purified by plaque assay in order to isolate a pure recombinant virus. It is possible to dramatically increase the proportion of recombinant over wild-type viruses (up to 90%) by using linearized and partially deleted viral DNA, as in the BaculoGold™ system (Pharmingen) (13), or in the MaxBac™ kit (Invitrogen Corp.), but this approach still does not eliminate the requirement for virus plaque purification.

Transfer Vectors for Recombination in E. coli. A newly developed system, BAC-TO-BAC, permits the rapid engineering of the recombinant virus in *E. coli* (14). It is based on site-specific transposition of the DNA cassette into the full viral DNA maintained in *E. coli* as a baculovirus shuttle vector (bacmid). This bacmid contains replication and kanamycin resistance functions and behaves like any common small plasmid in *E. coli*. In addition, a *lacZ*α region with a target site for transposition (mini *att*Tn7) was inserted into the bacmid, which allows transposition of the recombinant cassette from the transfer vector. The transposition machinery is provided in trans by a helper plasmid (tet resistant). This strategy allows fast, easy, and reliable selection of recombinant bacmid colonies on chromogenic substrate–containing plates. Most importantly, it eliminates the need for plaque purification of the recombinant virus. The recombinant bacmid DNA is prepared from the positive *E. coli* colonies and used to infect insect cells in culture in order to regenerate a pure recombinant virus stock.

Amplification and Titration. Titration of the virus is commonly performed by a standard plaque assay technique with Sf9 cell monolayers in presence of serum-containing media. This technique is laborious but gives the most accurate estimate of the number of infectious particles per unit volume (or plaque forming units) in the virus stock.

Table 2. Insect Cell Growth Parameters

Cell line	Medium	Td (h^{-1})	Max density (cells/mL)	Ref
Sf9	IPL41 + FCS	0.02	$5\ 10^6$	8
Sf9	SF900II	0.035	$1.5\ 10^7$	19
Sf9	TNMFH + FCS	0.032	$4.7\ 10^6$	1
Sf9	Excell 400	0.034	$2.9\ 10^6$	2
Tn5	Excel 401	0.030	$4\ 10^6$	3
Tn5	Excel 405	0.038	$6\ 10^6$	20
Tn5	SF-1	0.035	$5\ 10^6$	5

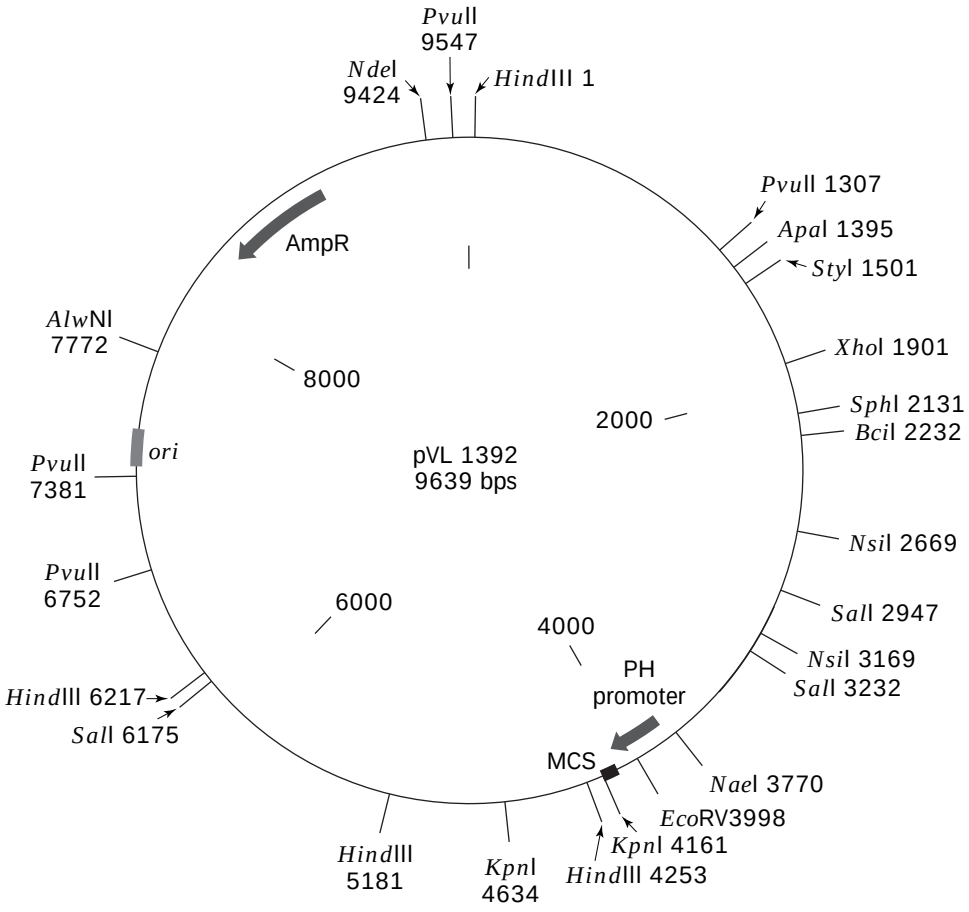


Figure 5. pVL1392 transfer vector. The pVL1392 baculovirus transfer vector was built from pUC8 with its origin of replication (*ori*) and ampicillin resistance gene (Amp^r). It contains the complete polyhedrin gene promoter sequence (PH promoter) but lacks part of the polyhedrin coding sequence. A multiple cloning site (MCS) region has been inserted 37 nucleotides downstream of the ATH polyhedrin start codon, which has been changed into an ATT. In a variant of this vector, pVL1393, the MCS contains the same restriction sites in the opposite orientation. *Source*: Reprinted from Ref. 12, with permission of Pharmingen, San Diego.

Once obtained in a pure form, the virus isolate must be expanded to larger volumes if required for manufacturing on a large scale. This can be accomplished by infection of a culture at a low multiplicity of infection (MOI = 0.01–0.1) so as to reduce the risk of producing defective interfering particles (DIPs) (15,16) and by harvesting the supernatant of the culture at the end of the infection phase (4 days). Titers between 10^7 and 10^8 PFU/mL are commonly obtained at this stage. Virus stocks can be stored at 4 °C, in the dark, for several months, although titers may be significantly reduced upon prolonged storage. A small aliquot can also be kept frozen at −70 °C for long-term storage.

Growth and Infection. Small-scale growth of the host insect cells and infection are easily obtained in stationary tissue culture flask cultures at 27 °C (15,17). Large-scale cultures can be performed with cells in suspension in spinner flasks or in bioreactors equipped with appropriate controls for temperature, dissolved oxygen, agitation, and a pH around 6.2 (18). Growth and product yields are fairly insensitive to bioreactor design but highly sensitive to reactor operation mode and other process parameters such as dissolved oxygen, cell density, and viability at onset of infection, MOI, and nutrient concentrations.

Insect cells have been known for a long time to be particularly sensitive to mechanical stress, in particular at the gas–liquid interface formed during the bubble-bursting process. This explains many of the initial failures in trying to grow insect cells in airlift bioreactors. This problem can be alleviated by supplementing the growth medium with appropriate surfactants such as Pluronic F-68 at a concentration of 0.02%, which prevents cells from being adsorbed on air bubbles.

Normally growth occurs with a doubling time of less than 24 h, and cell density reaches a plateau at 10^7 cells/mL in conventional media (Table 2). Infection is started by simple addition of a virus stock to the culture in full exponential growth ($1–2 \times 10^6$ cells/mL) and at high viability as determined by the Trypan blue exclusion method. A

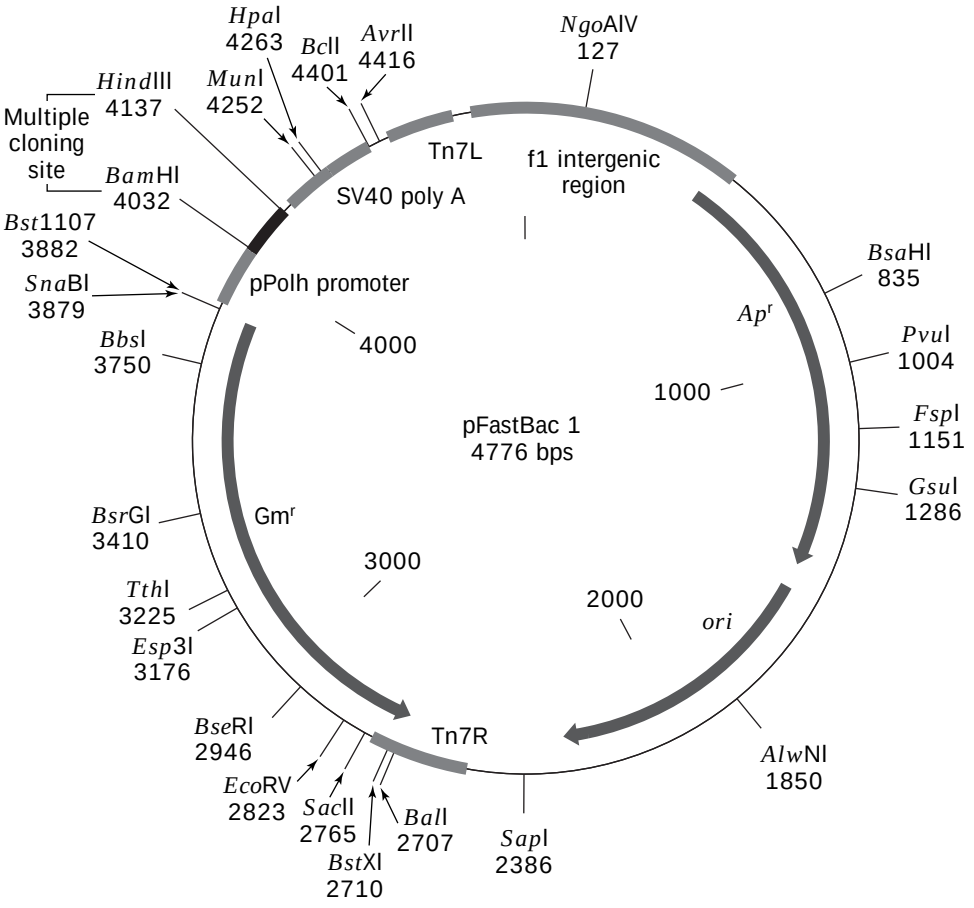


Figure 6. pFast-Bac1 transfer vector. The pFastBac1 vector contains the pUC origin of replication (*ori*), the ampicillin resistance gene (Ap^r), and transposon sequences (Tn7L, Tn7R) to produce recombinant bacmid DNA in *E. coli*. For selection and expression in insect cells, a gentamicin resistance gene (Gm^r), the polyhedrin promoter, SV40 polyadenylation signal, and a multiple cloning site are present. *Source*: reproduced from the GIBCO-BRL 1997–1998 product catalog, with permission from Life Technologies, Inc.

high MOI (5–10) guarantees a fast and uniform infection of all cells and maximizes the final product yield, but the amount of virus should not exceed 10% of the culture volume to reduce the addition of cell debris and spent medium. (Fig. 5).

The level of nutrients present in the medium at the time of infection is probably one of the most important parameters of the process. The nutritional state of the cells conditions the success of the infection and the final product yield, and the lack of a proper control of this parameter explains most observations relating the final expression level to the cell density. It has been clearly demonstrated by several groups that infection could actually be carried out at high cell densities (above 10^7 cells/mL), provided that the nutrient level was kept stable by either one or a combination of strategies, such as medium replacement, concentrated nutrient addition, and perfusion. (Fig. 6).

Another key element in large-scale insect cell processes is oxygen transfer, due to the high natural specific consumption rate (around 5×10^{-17} M/cell/s) by the insect cells. In the early infection process oxygen transfer becomes even more problematic because the cell's requirement can double in the first 24 h of infection. Thus, sparging pure oxygen, increasing agitation rates, and gas flow rates are the most efficient means of maintaining a stable dissolved oxygen concentration (set point between 20 and 50% of air saturation) throughout the whole process.

Results

Timing of Expression. The timing of expression is directly related to the promoter used in the construction of the recombinant virus. The kinetics of accumulation of the recombinant protein in the culture is a result of two inverse phenomena: translation of the recombinant gene from the viral DNA inside the infected cells, and progressive cell death due to viral induced cell lysis. Peak expression is usually observed between 48 and 72 h postinfection when the target gene is under control of very late promoters such as polyhedrin or p10. It may be advantageous, however, to drive expression from earlier promoters, such as the basic protein or *pcor* promoters, which leads to max-

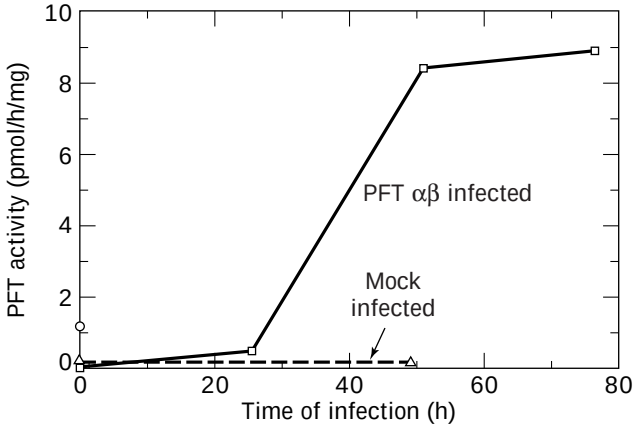


Figure 7. Time course of PFT production by baculovirus–infected Sf9 cells. Multiple cultures of Sf9 cells (4×10^6) were either infected with no virus (i.e., mock infected) (○), infected with wild-type virus (□), or infected with recombinant baculovirus containing the coding regions for both the a- and b-subunits of the rat brain PFT (△). Following infection, cells were incubated for the indicated times, whereupon they were harvested by centrifugation and soluble extracts prepared. The extracts were assayed for PFT activity as previously described (21). *Source*: Reprinted from Ref. 21, with permission of *Journal of Biological Chemistry*.

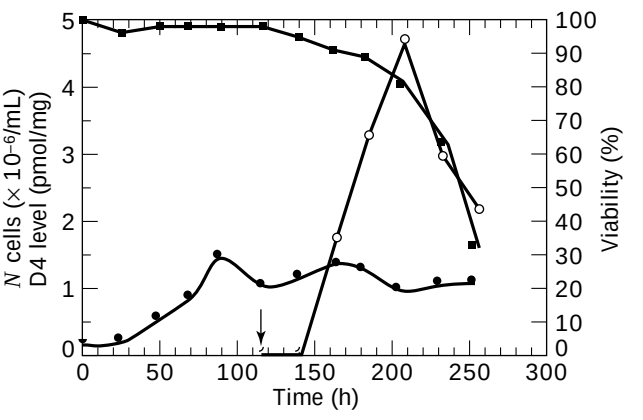


Figure 9. Expression of human dopamine D4 receptors in Sf9 cells. Sf9 cells were propagated in 15-L bioreactors at 27 °C, in TC100 medium (Gibco-BRL) supplemented with 10% FCS and 0.1% Pluronic F68 (BASF, Germany). Dissolved oxygen was controlled at 50% air saturation by pure oxygen sparging on demand. Once the culture reached a cell density of 1.5×10^6/mL (•) with a viability above 95% (■), fresh medium was added together with an aliquot of the virus stock so as to obtain a cell density of 1.0×10^6/mL and an MOI of 10 (arrow). Every 24 h postinfection, a sample was withdrawn for receptor binding activity measurements (○) with [^{3}H] spiperone, as previously described (24).

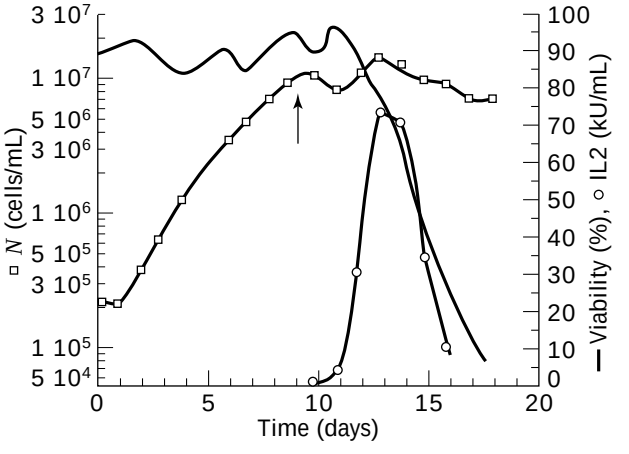


Figure 8. High-density perfusion culture of Sf-21 cells in serum-free Ex-Cell 401 medium. After 9 days, cells were infected with 1.8×10^8 AcNPV/pVL30 corresponding to a MOI of 0.015 (arrow). Maximum productivity was reached 4 days postinfection at a viability of 80%. *Source*: Adapted from Ref. 50, with permission.

imal expression earlier when cell viability is still above 80%.

Expression Level. A high expression level of recombinant proteins is commonly achieved with the baculovirus system; however, it is heavily dependent upon the nature and specific characteristics of each protein and each construct. Clearly, the expression level is also affected by the type of promoter used to drive expression of the foreign gene. The highest levels reported are in the range of several hundred micrograms of active protein per milliliter of culture, as in the case of expression of *β*-galactosidase in monolayer cultures (10). Some cytoplasmic or secreted proteins can also be produced in the 100-mg/L range in large-scale cultures (21,22) (Figs. 7 and 8), but more complex proteins such as multisubunit complexes and single or multiple transmembrane proteins are usually produced at lower levels, typically 0.1–5 mg/L (Fig. 9) (24).

Posttranslational Events. One of the strong arguments for the baculovirus is the ability of insect cells to process nascent polypeptides in a way that is similar to that mammalian cells, and most mammalian proteins produced by this system are virtually identical to their native equivalent in terms of their biochemical or biological activity. It is well established, however, that N-glycosylation performed by the most commonly used cell lines, Sf9 or Tn5, is significantly different from mammalian N-glycosylation, with most proteins being processed to the high-mannose structure steps. However, glycosylation by other specialized insect cell lines such as *Estigmene acrea* (25) appears to be more closely related to that of mammalian cell lines. This area is of course of considerable interest and is still the subject of intensive research.

Due to the availability of several promoters, and because large sequences of DNA can be inserted in the viral genome, the baculovirus is particularly well suited to the production of multisubunit complexes. Mammalian viral capsids or viruslike particles can also be reconstituted inside the cell, and this property is being exploited for the development of novel cattle and human vaccines.

Conclusion and Perspectives

The baculovirus system has shown to be a very useful tool for the production of numerous recombinant mammalian

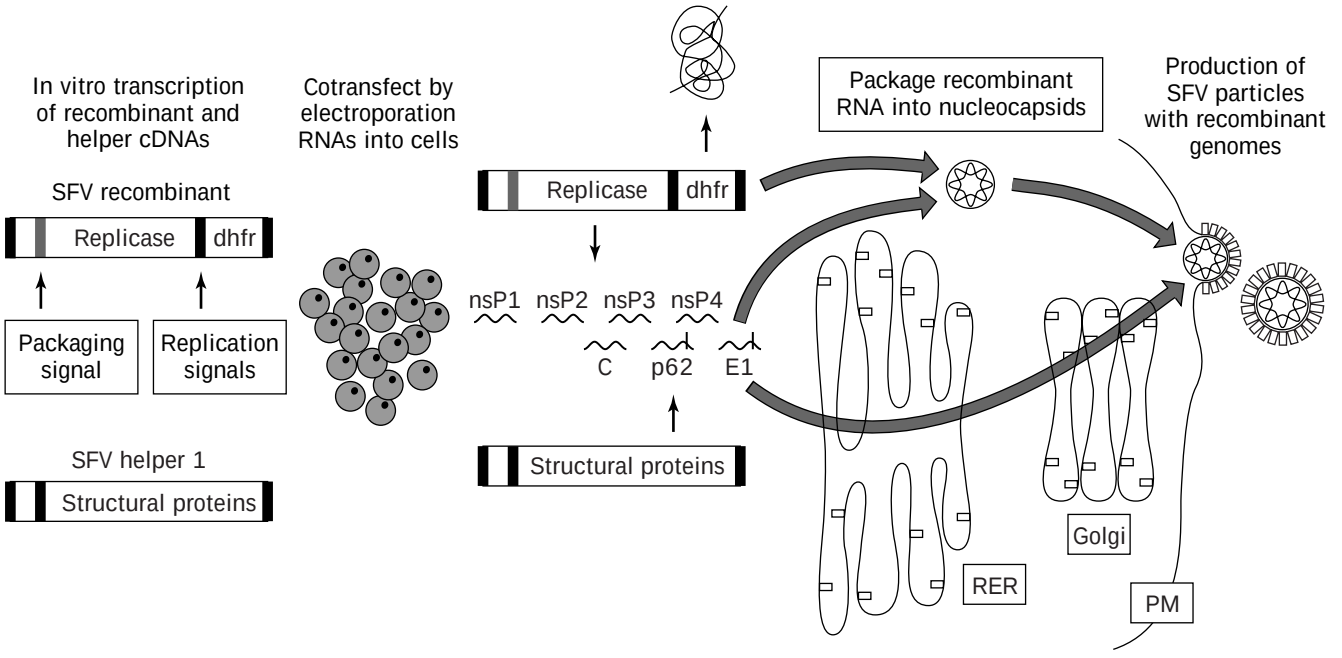


Figure 10. In vivo packaging of recombinant SFV RNA transcripts into infectious particles. After cotransfection of RNA transcripts from the expression and the helper vector, the replicase, coded on the recombinant vector, drives RNA replication and transcription for both plasmids. Structural genes from the helper are translated into proteins. While the capsid protein stays cytosolic and binds RNA to form the nucleocapsid, the membrane proteins undergo posttranslational modifications. E1 and p62 are dimerized in the endoplasmic reticulum and are transported to the plasma membrane where virus budding occurs. Because the viral packaging signal is found only on the recombinant RNA, selective packaging of the pSFVdhfr occurs and as a result yields nonreplicative infectious SFV particles. *Source*: Reprinted from Ref. 26, with permission of Nature-Biotechnology, N.Y.

proteins for research, animal vaccination, and clinical applications. Several biological products produced with this system are being tested in therapies or preventive vaccination, and it seems likely that at least some of them will soon obtain approval for marketing. The major advantages of the system are high-level expression, safety, speed, coexpression of several proteins, easy scale-up, and low cost, which explain its widespread use in particular for the production of research material. Future work is needed to improve control over posttranslational events, which is the only area of difference with more conventional mammalian cell–based expression systems.

SEMLIKI FOREST VIRUS

Introduction: Alpha Virus–Based Expression Systems

Alphaviruses such as the Sindbis virus (SIN) or the Semliki Forest virus (SFV) are very small enveloped viruses containing single-stranded RNA of positive polarity. The 12-kb RNA is capped at the 5′ terminus and polyadenylated at the 3′ terminus. SIN and SFV are the most-studied alphaviruses, and both have been used for the expression of various heterologous genes. Upon infection, the 5′ two-thirds of the RNA is translated into polypeptides, which are processed into four nonstructural proteins (nsP1–nsP4). The nonstructural proteins include transcriptase, which produces a negative-strand copy of the RNA genome, serving as a template for the generation of positive-strand RNA copies. The 3′ third of the genome codes for the structural proteins (i.e., for the capsid C and two membrane proteins, p62 and E1). After translation, the capsid protein C binds to genomic RNA within the cytoplasm. The membrane proteins are translocated into the endoplasmic reticulum, where they form heterodimers and allow virus budding from the cell surface. The released virus can then be taken up by receptor-mediated endocytosis and infect a wide range of mammalian cells.

Semliki Forest Virus Expression System

The broad host range of the SFV expression system has made it a particularly attractive system for functional studies and investigation of the biological role of proteins. It was originally developed by Liljeström and Garoff (26), who split the SFV genome into two plasmids: the cloning vector, pSFV1, which carries the replicase genes; and the pSFV-helper, which encodes the structural proteins of the virus (Figs. 10 and 11). In order to increase the level of biological safety, they later also introduced a mutation into the spike protein–encoding gene, generating pSFV-helper II (27). The mutation prevents the natural proteolytic processing of the spike protein into an active protein, which

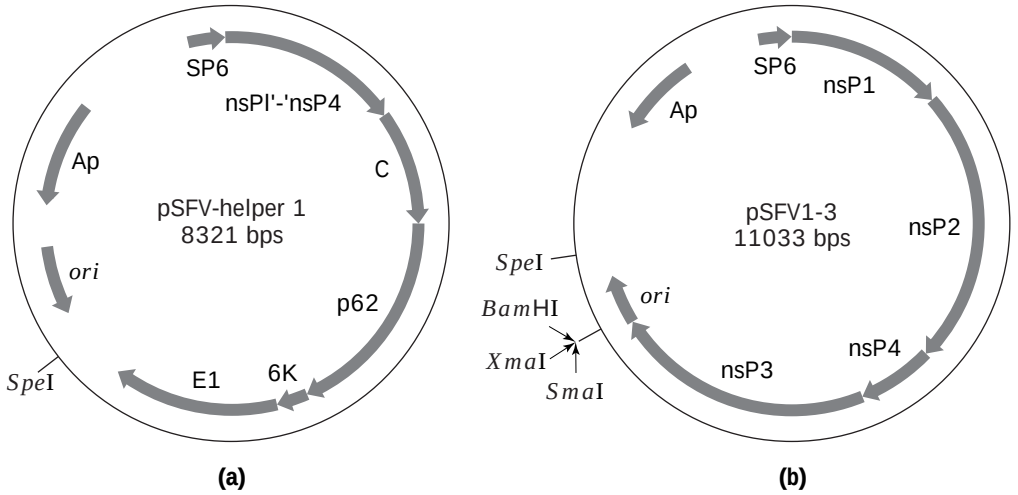


Figure 11. pSFV Helper1 (**a**) and pSFV1-3 plasmids (**b**). Construction of the SFV expression plasmids was done by deleting most of the coding region for the structural proteins (C, p62, 6K, E1) from the cDNA clone, which spans the complete SFV genome. The vectors retain the promoter region of the 26S subgenomic RNA (at the end of the nsP4 region) and the last 49 amino acids of the E1 protein as well as the complete noncoding 3′ end of the SFV genome. Downstream of the 26S promoter, a polylinker (*Bam*III, *Xma*I, *Sma*I) was introduced. The SP6 promoter, which is followed by the replicase (nsP1-nsP4), allows in vitro generation of full-length SFV RNA transcripts. The eukaryotic origin of replication (*ori*) and antibiotic resistance (Ap) allow for plasmid production in *E. coli*. A unique *Spe*I restriction site was inserted for plasmid linearization. The pSFV-helper1 represents the SFV genome cDNA being deleted of most of the nonstructural region but carrying all the structural genes (C, p62, 6K, E1). The pSFV-helper2 codes for a cleavage-deficient variant of the spike protein. *Source*: Reprinted from Ref. 26, with permission of Nature-Biotechnology, N.Y.

mediates entry into the cell. However, the processing can be obtained by chymotrypsin cleavage, thus restoring full infectivity of the virus stock (27). High expression levels are obtained due to the extremely efficient SFV 26S promoter used and also to the high number of copies of RNA (200,000) present in each infected cell. For example, levels up to 25% of total cell protein have been reported for β-galactosidase (26), and several million neurokinin 1 receptors per cell (28). Because the virus can be used to infect almost any mammalian cell, recombinant proteins can be produced with posttranslational modifications based on the specific processing capacities of the host cell selected (Table 3).

Infection and protein expression properties. SFV can infect a broad range of host cells; however, the level of expression differs between cell lines: BHK, CHO, and COS cells yield higher levels of expression than other cells such as NSO, RPMI 8226, and HeLa. This was clearly demonstrated for different types of proteins: a multitransmembrane receptor (neurokinin 2 receptor), a secreted protein (soluble placental alkaline phosphatase), and a cytosolic protein (β-gal).

At least in BHK cells it is known that the optimal pH is different for SFV infection (pH 6.9) and protein production (pH 7.3). The MOI also clearly modulates the expression level (30): a higher MOI results in elevated expression levels.

Transfecting cells with RNA generated in vitro from the cloning vector also leads to RNA replication and subsequent expression of the subcloned gene, a method that should be applicable to the expression of oncogenes and other hazardous genes. However, infection with recombinant packaged SFV particles is still a more efficient way to deliver recombinant RNA inside the cells.

Basic Principles and Protocols

Virus Generation. After subcloning the gene of interest into the multiple cloning site of pSFV, the plasmid is linearized to prepare for in vitro transcription and RNA production. BHK cells are then cotransfected with the recombinant and helper RNAs (Fig. 12), which initiates the in vivo packaging process. It is important to note that the packaging signal is present only on pSFV-1 and is absent from the pSFV-helper vector. As a consequence, infectious but nonreplicative recombinant virus particles are produced, and a high titer (approximately 10^9 virus particles/mL) is harvested 24 h after transfection.

Protein Expression. As a first step, BHK cells are grown to confluence in T-flasks. The virus is activated by incubation with chymotrypsin, which is inactivated after 20 min by addition of aprotinin. Virus is transferred into the culture to give a MOI of 30 for infection. Product can be detected by pulse labeling after 4–6 h postinfection while the host cell protein synthesis is dramatically reduced or shut down and cell growth is arrested. Maximum product concentration is usually reached 1 day postinfection but

Table 3. SFV-Expressed Recombinant Proteins

Protein group	Protein	Ref
Enzymes	β-Gal (up to 25% of cell protein)	26
	Mouse dehydrofolate reductase	26
	Chicken lysozyme	26
	Terminal deoxynucleotide transferase	29
	Human cyclooxygenase 2	30
	Rat and human catechol *O*-methyltransferase	31
	Wild-type and mutant thyroid peroxidase (TPO)	32
Receptors	Dopamine receptor 3 ($>1 \times 10^6$ receptors/cell)	33
	Human neurokinin receptor 1 receptor (2×10^6 receptors/cell)	28
	Human neurokinin receptor 2 receptor	29
	Human transferin receptor	33
	5HT-3 receptor	26
	5HT-1D alpha receptor	34
	Adenosin A3 receptor	35
	Human prostaglandin receptor E4	35
	Dopamin 3 receptor mutant	36
	Human P2×1, rat P2×2	37
	Human P2×1, rat P2×2, rat 2×3, rat 2×4	38
	GABA subunits	39
		40
Virus	Human papilloma type 16 capsid protein	41
	Hamster measles virus protein HNT-PI F and Ph26 H	42
	Influenza virus nucleoprotein	43
	HIV gp120, gp160	44
	Moloney murine leukemia virus spike protein	45
	Vesicular stomatitus virus glycoprotein G	46
	GB virus-C E2 envelope protein	47
	GP4 structural glycoprotein of Lelystad virus	48
	Respiratory-cynytial virus glycoprotein-G	49
	Murine leukemia virus envelopes	50
Others	Rab 8	51
	Rab 12, 22, 24	52
	VIP-21caveolin	53
	Yeast syntaxin Sso2p	54
	Rabbit pIgR	55
	Human APP	56
	Plasminogen activator inhibitor type 2	57
	Mitochondrial precurser protein	58
	APOE isoforms	59
	P35/p40 IL-12 heterodimer	60

may vary depending on the proteins expressed. The culture is harvested by centrifuging the cells after detachment.

Semliki Forest Virus Scale-Up Results

BHK cells are grown in single cell suspension to 10^6 cells/mL using a stirred-tank reactor equipped with a Vibromix stirring system and oxygenation by silicone tube gassing (Chemap/MBR, Switzerland). This agitation method proves to be optimal for minimizing cell aggregate formation, and the oxygenation method is preferred over direct gas sparging into the bulk of the liquid so as to eliminate any aerosol formation.

Before infection, 90% of the medium is exchanged by external cross-flow filtration. The virus is activated separately and the culture infected by addition of the activated virus stock. At 16 h postinfection the culture is harvested by centrifugation and the recombinant protein purified.

Recent results of large-scale protein production with SFV have recently been described:

- Cyclooxygenase-2: 16 mg microsomal protein per 10^9 cells, activity: 3,942 pg PGE_2/μg microsomal protein (30)
- 5HT3 receptor: 1–2 mg 5HT3 receptor protein per liter culture, or 25–50 pmol binding sites per cell (19)

Commentary

The SFV expression system has quickly found wide acceptance due to the high level of expression achievable combined with the short time scale for production. It is preferentially used to express receptors and intracellular products.

Virus generation for large-scale infection is still laborious, in particular the RNA generation and transfection of cells with the two RNAs. This constitutes the main difference with other viral-mediated expression systems.

Biological safety of the SFV expression system has been given significant attention. The virus is nonreplicative and only conditionally infective. However, because the culture

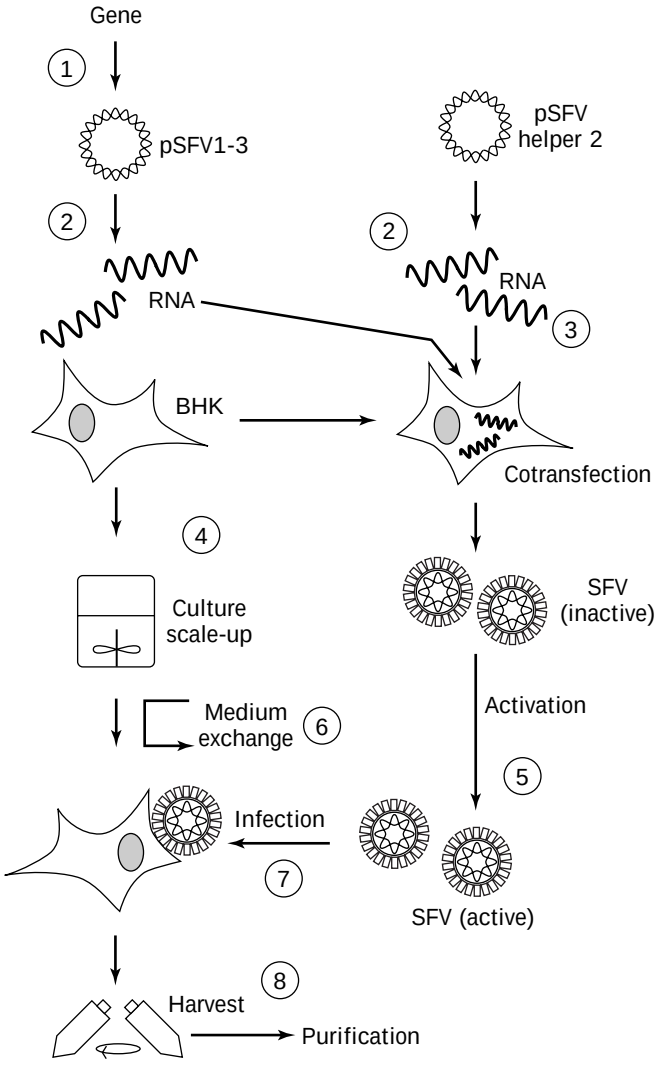


Figure 12. SFV infection process. Large-scale recombinant protein production with the Semliki Forest virus system. The gene to be expressed is subcloned (1) into the recombinant vector pSFV1-3. In vitro transcribed RNA (2) from the recombinant plasmid pSFV1-3 and the pSFV-helper-2 are cotransfected (3) into BHK cells; virus is harvested 24 h later. BHK cells are grown up to 10^6 cell/mL as suspension culture in a mechanically stirred reactor (4). Virus is activated (5), the culture medium is exchanged (6), and the culture infected (7). Cells are harvested (8) 1 day after infection by centrifugation, and the protein is purified.

at the time of harvest potentially contains active virus, it is recommended that this expression system and the handling of products obtained by this technology be operated under safety category two or higher.

OTHER VIRAL-MEDIATED EXPRESSION SYSTEMS

Adenovirus

Introduction. Isolated in the early 1950s, adenoviruses have recently received a renewed interest for their use as vectors for DNA delivery to cells in vivo in gene therapy applications, as potential vaccine vectors, or as expression vectors for heterologous genes in vitro. In this section, the latter application of adenoviruses is reviewed.

Adenovirus Life Cycle. The adenovirus life cycle starts with an early phase in which regions E1–E4 and L1 of the linear double-stranded DNA (36 kb total length) are actively transcribed from the early (E) and major late promoters (MLP), respectively (61). Viral replication begins 8 h after infection, and host transcription as well as protein synthesis are gradually shut off as the cycle enters the late phase of infection. In this phase, late transcripts from the MLP, L2–L5 appear together with structural proteins. The MLP promoter drives high levels of expression of these late proteins, which can account for up to 50% cellular protein. The cycle also leads to cell death and to the release of high virus titers (10^8–10^9 PFU/mL).

Expression Vectors. The extensive knowledge accumulated over the years on adenovirus transcription regulation allowed the engineering of adenovirus vectors (Ad) modified for the expression of heterologous genes (61). The first modification consists in deleting the early regions E1 and E3. Whereas E3 is not essential to viral replication in cell cultures, E1 deletion renders the virus replication incompetent, and this function must then be provided in trans by the host cell, as is the case for 293 cells, which constitutively express the E1 protein. The second modification consists in introducing an expression cassette in place of the deleted E1 region. In the cassette, the recombinant gene is placed under control of an additional (or ectopic) MLP and follows the tripartite leader and splicing signal sequences, thus taking advantage of the strong natural positive regulation of the promoter. Alternatively, the recombinant can be placed under control of an exogenous promoter such as CMV (62) or the natural MLP, but in this latter case, the recombinant virus must then be treated as a helper-dependent virus. Very recently, a new series of vectors, pAdBM5, was constructed (63) with significant improvements in the ectopic MLP region by insertion of enhancer sequences that stimulate the promoter's transcriptional activity. The fact that the virus can be propagated in suspension cell cultures is also of considerable advantage over other similar processes.

Results. Several heterologous proteins have been expressed from Ad vectors at high levels in 293 cells. Most constructions have been done in the Ad5 serotype background, and with a few exceptions, such as DHFR, CAT, factor IX or the vitamin D3 receptor, all were made for expressing viral proteins. SV40 large T antigen was expressed at high levels and could be purified to homogeneity, Herpes simplex virus type 2 ribonucleotide reductase subunit R2 was expressed in 293 cells at 4–5% of cellular proteins, and the polyomavirus middle T antigen and the hepatitis B surface antigen were also expressed at high levels in the same cells. The pAdBM5 vector has been reported to yield even higher levels of expression (10–20% of cellular

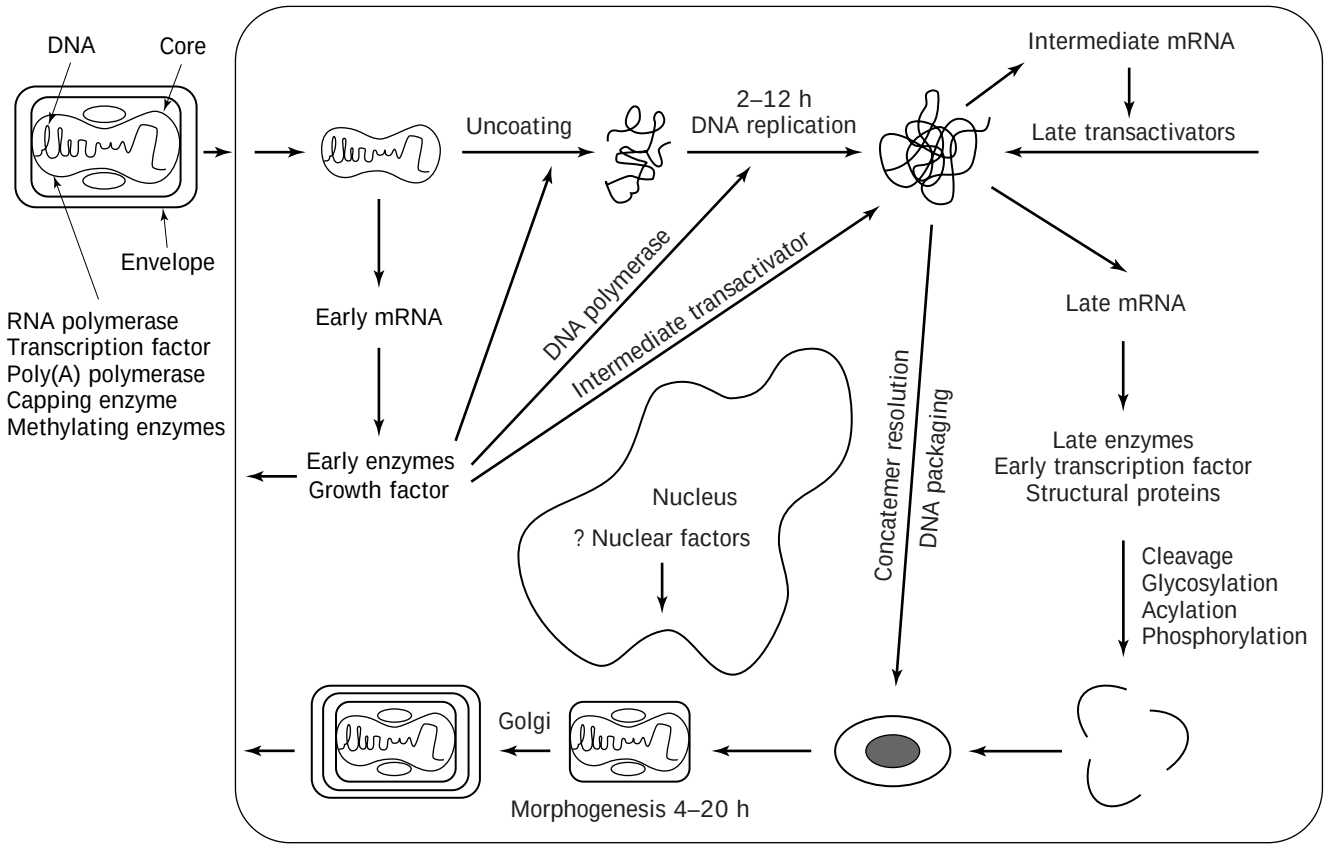


Figure 13. Replication cycle of vaccinia virus. After entry of vaccinia virus into cells, early genes are expressed, leading to secretion of several proteins (including a growth factor), uncoating of the virus core, and synthesis of DNA polymerase (and other replication proteins), RNA polymerase subunits, and transcriptional transactivators of the intermediate class of genes. After DNA replication, intermediate mRNAs are made, some of which encode late transcriptional transactivators, leading to expression of late genes. The latter encode structural proteins, enzymes, and early transcription factors, which are packaged in the assembling virus particles. Some of the mature virions are wrapped in Golgi-derived membranes and are released from the cell. The bold arrows indicate the products that exit the cell. *Source*: Reprinted from Ref. 66, with permission of Raven Press, N.Y.

Table 4. Vaccinia Large-Scale Production

Product	System	Cells	Company	Ref.
HIV-1 rgp160	150-L microcarrier	BHK	Pasteur Merieux	67
HIV-1 gp160	1200-L microcarrier	VERO	Immuno AG	68
Human protrombin	1200-L microcarrier	VERO	Immuno AG	69

proteins), with a tyrosine phosphatase (PTP1C) used as a model in bioreactor-scale cultures of 293 cells (64,65).

As is normally expected, posttranslational processing of Ad-derived recombinant proteins has been shown to be identical to the native protein.

Vaccinia Virus

Vaccinia Virus Characteristics. The vaccinia virus, an *Orthopoxivirus*, belongs to the family of Poxviridae. It has a double-stranded DNA of nearly 200,000 bp and replicates in the cytoplasm of the host cell (Fig. 13). The lipoprotein envelope contains the core structure, including an RNA polymerase, a transcription factor, capping and methylating enzymes, and a poly(A) polymerase.

Infection of cells with the vaccinia virus leads to the suppression of host cell protein synthesis, and cytopathic effects occur. Up to 5,000 virus particles are produced per cell, which, depending on the virus strain, will stay cell associated or can be found in the culture medium.

Table 5. Transient Expression Systems

	Time scale	Advantages	Disadvantages	Applications	Yields
			DNA vectors		
Large T–expressing cell lines (e.g., COS)	3–4 weeks	Established technology Simple to execute Scale up to liter scale demonstrated	Large DNA and cell quantities required for transfection Large-scale transfection Short expression time after transfection (1–2 weeks), overgrowth by non producers Safety category 2–3	Functional and structural studies	0.1–3 mg/L Up to 10 mg protein
HEK293 EBNA	4–8 weeks	Simple to execute Culture scale-up under selective pressure Transfection requires only small quantities of DNA and cells Stable under selection Potential selection of clones Establishing seed: bank possible	Safety category 2	Functional and structural studies, target validation, preclinical studies	0.1–10 mg/L 100 mg protein
			Viral vectors		
Baculovirus	6–8 weeks	Established technology Easy scale-up Safe viral system	Lytic system Nonmammalian glycosylation	Functional and structural studies, target validation, preclinical studies	0.5–30 mg/L gram quantities
Semliki Forest virus	2–6 weeks	Exceptionally short time from gene to product High yields Very broad host range	Safety category 2 for activated virus Limited virus scale-up Lytic system	Functional and structural studies, target validation	Up to 8×10^6 cell Up to 100 mg
Vaccinia	2–8 weeks	Easy scale-up to very large scale Virus simple to scale-up Broad viral host range	Pathogenic: safety category 2 Lytic system	Functional and structural studies, target validation, preclinical studies; vaccines	Around 0.1–5 mg Gram quantities

Vaccinia Virus for Heterologous Expression. For recombinant protein expression, the gene of interest is placed under the control of a vaccinia promoter. Late promoters are preferred for yielding high-level protein expression. Integration of the foreign gene can be accomplished by one of three methods: homologous recombination, direct cloning, or construction of a hybrid of vaccinia with the bacteriophage T7 RNA polymerase.

Homologous recombination is performed by (*1*) subcloning cDNA into a plasmid next to a vaccinia promoter, flanked by nonessential vaccinia gene sequences; (*2*) recombining into the vaccinia genome (cell infection with wild-type vaccinia virus and transfection of recombinant plasmid); and (*3*) screening for recombinant virus.

Direct cloning is performed by inserting the cDNA into a unique restriction site of the vaccinia virus genome, which is packaged after transfecting helper cells (fowlpox virus–infected cells). High-level expression can also be achieved with the vaccinia virus–T7 RNA polymerase hybrid system. This system relies on the cytoplasmic expression of the bacteriophage T7 RNA polymerase (e.g., by a stable transfected cell line). The recombinant gene is inserted into a vaccinia virus under the control of the T7 RNA polymerase promoter. To provide the T7 polymerase, either constitutively expressing cells are infected or cells are cotranfected with a polymerase expressing vaccinia virus.

Advantages of the vaccinia virus system are (*1*) the wide host range of mammalian cells that can be infected, (*2*) the mammalian posttranslational modification machinery, (*3*) the high expression levels, and (*4*) easy production of the virus. A disadvantage of the system is the safety level required when working with vaccinia.

The vaccinia system has been successfully scaled up to very large scale (Table 4).

CONCLUSION

Transient expression technologies have been available for many years and have been applied to many different objectives. Recent novel applications cover the use of these technologies for producing recombinant proteins in amounts sufficient for biochemical and biological studies (100–1,000 mg). Table 5 summarizes the essential features of each technology. A first step common to all technologies presented here is the large-scale preparation of the expression vector (plasmid DNA, mammalian or insect recombinant virus). Speed is the principal motivation (starting from cDNA, the protein is produced within 10–15 days) in contrast to more classical technologies such as stable cell lines, which require selection and amplification of the recombinant gene. In addition, novel routes to pharmaceutical discovery such as human genome sequencing create a serious need for high throughput, medium-scale (10–100 mg of protein) expression systems in order to quickly characterize and evaluate the function of novel gene products. The eukaryotic environment in which most of these transient expression systems operate ensures that elaborate posttranslational modifications are properly carried out and that the protein product obtained provides an authentic representation of its native human counterpart.

ACKNOWLEDGMENTS

We are thankful to Dr. Tom Kost for suggestions and corrections of the manuscript and to Chris Hebert for graphic work.

BIBLIOGRAPHY

1. Y. Glutzman, *Cell* **23**, 175–182 (1981).
2. M. Lusky and M. Botchan, *Nature* **27**, 279–288 (1981).
3. J.K. Rose and J.E. Bergmann, *Cell* **30**, 753–762 (1982).
4. B. Seed, *Nature* **329**, 840–842 (1997).
5. G. Cachianes, C. Ho, R.F. Weber, S.R. Williams, D.V. Goeddel, and D.W. Leung, *Biotechniques* **15**, 255–259 (1993).
6. H.D. Blasey, J.P. Aubry, G.J. Mazzei, and A.R. Bernard, *Cytotechnology* **18**, 183–192 (1995).
7. R. Ridder, S. Geisse, B. Kleuser, P. Kawalleck, and H. Gram, *Gene* **166**, 273–276 (1995).
8. M. Jordan, C. Koehne, and F. Wurm, *Cytotechnology* **26**, 39–47 (1998).
9. G.E. Smith, M.D. Summers, and M.D. Fraser, *Mol. Cell. Biol.* **3**, 2156–2165 (1983).
10. G.D. Pennock, C. Shoemaker, and L.K. Miller, *Mol. Cell. Biol.* **4**, 399–406 (1984).
11. T.J. Wickham, T. Davis, R.R. Granados, M.L. Shuler, and H.A. Wood, *Biotechnol. Prog.* **8**, 391–396 (1992).
12. V.A. Luckow and M.D. Summers, *Biotechnology (N.Y.)* **6**, 47–55 (1988).
13. M.D. Gruenwald and M.S. Heitz, in *Baculovirus Expression Vector System: Procedures and Methods Manual*, Pharmingen, San Diego, 1996, pp. 1–73.
14. V.A. Luckow, S.C. Lee, G.F. Barry, and P.O. Olins, *J. Virol.* **67**, 4566–4579 (1993).
15. D.R. O'Reilly, L.K. Miller, and V.A. Luckow, in *Baculovirus Expression Vectors: A Laboratory Manual*, Oxford University Press, Oxford, 1994, pp. 1–347.
16. M. Kool, J.W. Voncken, F.L. van Lier, J. Tramper, and J.M. Vlak, *Virology* **183**, 739–746 (1991).
17. M.D. Summers and G.E. Smith, in *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*, Texas A&M University, College Station, Texas, 1988, pp. 1–56.
18. A.E. Proudfoot, L. Goffin, M.A. Payton, T.N.C. Wells, and A.R. Bernard, *Biochem. J.* **318**, 437–442 (1996).
19. R. Hovius, A.-P. Tairi, H. Blasey, A.R. Bernard, K. Lundström, and H. Vogel, *Protein Sci.* (1997).
20. F.L. Graham and A.J. Van der Eb, *Virology* 456–467 (1973).
21. W.J. Chen, J.F. Moomaw, L. Overton, T.A. Kost, and P.J. Casey, *J. Biol. Chem.* **268**, 9675–9680 (1993).
22. V. Jäger and A. Kobold, *Biotechnol. Tech.* **9**, 435–440 (1995).
23. V. Jäger, A. Kobold, C. Köhne, S.M. Deutschmann, E. Grabenhorst, E. Karger, and H.S. Conradt, in R.E. Spier, J.B. Griffiths, and W. Berthold, eds., *Animal Cell Technology: Products for Today, Prospects for Tomorrow*, Butterworth-Heinemann, Oxford, 1994, pp. 207–211.
24. A. Mills, B. Allet, A.R. Bernard, C. Chabert, E. Brandt, C. Cavegn, A. Chollet, and E. Kawashima, *FEBS Lett.* **320**, 130–134 (1993).

25. R. Wagner, H. Geyer, R. Geyer, and H.D. Klenk, *J. Virol.* **70**, 4103–4109 (1996).
26. P. Liljeström and H. Garoff, *Biotechnology (N.Y.)* **9**, 1356–1361 (1991).
27. P. Berglund, M. Sjoberg, H. Garoff, G.J. Atkins, B.J. Sheahan, and P. Liljeström, *Biotechnology (N.Y.)* **11**, 916–920 (1993).
28. K. Lundström, A. Mills, G. Buell, E. Allet, N. Adami, and P. Liljeström, *Eur. J Biochem.* **224**, 917–921 (1994).
29. V. Ciccarone, J. Jessee, P. Berglund, and P. Liljeström, *Focus* 103–105 (1993).
30. H.D. Blasey, K. Lundström, S. Tate, and A.R. Bernard, *Cytotechnology* **24**, 65–72 (1997).
31. K. Lundström, J. Tenhunen, C. Tilgmann, T. Karhunen, P. Panula, and I. Ulmanen, *Biochim. Biophys. Acta* **1251**, 1–10 (1995).
32. H. Bikker, F. Baas, and J.-J.M. De-Vijlder, *J. Clin. Endocrinol. Metab.* **82**, 649–653 (1997).
33. K. Lundström, A. Mills, E. Allet, K. Ceszkowski, G. Agudo, A. Chollet, and P. Liljeström, *J. Receptor Signal Transduction Res.* **15**, 23–32 (1995).
34. P. Werner, E. Kawashima, J. Reid, N. Hussy, K. Lundström, G. Buell, Y. Humbert, and K.A. Jones, *Mol. Brain Res.* **26**, 233–241 (1994).
35. M. Patel, C. Harris, and K. Lundström, *Drug Dev. Res.* **40**, 35–40 (1997).
36. F.H. Marshall, K. Patel, K. Lundström, J. Camacho, S.M. Foord, and M.G. Lee, *Br. J. Pharmacol.* **121**, 1673–1678 (1997).
37. K. Lundström and M.P. Turpin, *Biochem. Biophys. Res. Commun.* **225**, 1068–1072 (1996).
38. A.D. Michel, K. Lundström, G.N. Buell, A. Surprenant, S. Velara, and P.-P.A. Humphrey, *Br. J. Pharmacol.* **118**, 1806–1812 (1996).
39. K. Lundström, A. Michel, H. Blasey, A.R. Bernard, R. Hovius, H. Vogel, and A. Surprenant, *J. Receptor Signal Transduction Res.* 115–126 (1997).
40. G.H. Gorrie, Y. Vallis, A. Stephenson, J. Whitfield, B. Browning, T.G. Smart, and S.J. Moss, *J. Neurosci.* **17**, 6587–6596 (1997).
41. P. Heino, J. Dillner, and S. Schwartz, *Virology* **214**, 349–359 (1995).
42. K.B. Hummel, J.A. Vanchiere, and W.J. Bellini, *Virology* **202**, 665–672 (1994).
43. X. Zhou, P. Berglund, H. Zhao, P. Liljeström, and M. Jondal, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3009–3013 (1995).
44. N.L. Paul, M. Marsh, J.A. McKeating, T.F. Schulz, P. Liljeström, H. Garoff, and R.A. Weiss, *AIDS Res. Hum. Retroviruses* **9**, 963–970 (1993).
45. M. Suomalainen and H. Garoff, *J. Virol.* **68**, 4879–4889 (1994).
46. M.M. Rolls, K. Haglund, and J.K. Rose, *Virology* **218**, 406–411 (1996).
47. T.J. Pilot-Matias, R.J. Carrick, P.F. Coleman, T.P. Leary, T.K. Surowy, J.N. Simons, A.S. Muerhoff, S.L. Buijk, M.L. Chalmers, G.J. Dawson, S.M. Desai, and I.K. Mushahwar, *Virology* **225**, 282–292 (1996).
48. J.-J.M. Meulenberg, A.P. Van-Nieustadt, Z.A. Van-Essen, and J.-P.M. Langeveld, *J. Virol.* **71**, 6061–6067 (1997).
49. J. Meanger, I. Peroulis, and J. Mills, *Biotechniques* **23**, 432–434, 436 (1997).
50. F.-K.N.-A.S.-J. Lebedeva-I, *J. Virol.* **9**, (1998).
51. M.J. de Hoop, V.M. Olkkonen, E. Ikonen, E. Williamson, C. von Poser, L. Meyn, and C.G. Dotti, *Gene Therapy* **1** (Suppl 1), S28–31 (1994).
52. V.M. Olkkonen, P. Dupree, I. Killisch, A. Lutcke, M. Zerial, and K. Simons, *J. Cell. Sci.* **106**, 1249–1261 (1993).
53. A.M. Fra, E. Williamson, K. Simons, and R.G. Parton, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8655–8659 (1995).
54. J. Jantti, S. Keranen, J. Toikkanen, E. Kuismanen, C. Ehnholm, H. Soderlund, and V.M. Olkkonen, *J. Cell. Sci.* **107**, 3623–3633 (1994).
55. M. de Hoop, C. von Poser, C. Lange, E. Ikonen, W. Hunziker, and C.G. Dotti, *J. Cell. Biol.* **130**, 1447–1459 (1995).
56. B. De Strooper, M. Simons, G. Multhaup, F. Van Leuven, K. Beyreuther, and C.G. Dotti, *EMBO J.* **14**, 4932–4938 (1995).
57. P. Mikus, T. Urano, P. Liljeström, and T. Ny, *Eur. J. Biochem.* **218**, 1071–1082 (1993).
58. A. Huckriede, A. Heikema, J. Wilschut, and E. Agsteribbe, *Eur. J. Biochem.* **237**, 288–294 (1996).
59. S. Aleshkov, C.R. Abraham, and V.I. Zannis, *Biochemistry* **36**, 10571–10580 (1997).
60. J. Zhang, P.C. Asselin, F. Bex, J. Bernrd, J. Chehimi, F. Willems, A. Caignard, P. Berglund, P. Liljeström, A. Burny, and S. Chouaib, *Gene Therapy* **4**, 367–374 (1997).
61. F.L. Graham and L. Prevec, *Mol. Biotechnol.* **3**, 207–220 (1995).
62. A.R. Fooks, E. Schadeck, U.G. Liebert, A.B. Dowsett, B.K. Rima, M. Steward, J.R. Stephenson, and G.W.G. Wilkinson, *Virology* **210**, 456–465 (1995).
63. B. Massie, J. Dionne, N. Lamarche, J. Fleurent, and Y. Langelier, *Biotechnology (N.Y.)* **13**, 602–608 (1995).
64. I. Nadeau, A. Garnier, J. Cote, B. Massie, C. Chavarie, and A. Kamen, *Biotechnol. Bioeng.* **51**, 613–623 (1996).
65. A. Garnier, J. Cote, I. Nadeau, A. Kamen, and B. Massie, *Cytotechnology* **15**, 145–155 (1994).
66. B. Moss, in B.M. Fields, D.M. Knipe, R.M. Chanock, M.S. Hirsch, J.L. Melnick, T.P. Monath, and B. Roizman eds., *Virology* Raven Press, New York, 1990, pp. 2079–2111.
67. G. Pialoux, J.L. Excler, Y. Riviere, G. Gonzalez-Canali, V. Feuillie, P. Coulaud, J.C. Gluckman, T.J. Matthews, B. Meignier, M.P. Kieny, P. Gonnet, I. Diaz, C. Meric, E. Paoletti, J. Tartaglia, H. Salomon, S. Plotkin, The Agis Group, and L'Agence Nationale de Recherche sur le Sida, *AIDS Res. Hum. Retroviruses* **11**, 373–381 (1995).
68. N. Barrett, A. Mitterer, W. Mundt, J. Eibl, M. Eibl, R.C. Gallo, B. Moss, and F. Dorner, *AIDS Res. Hum. Retroviruses* **5**, 159–171 (1989).
69. F.G. Falkner, P.L. Turecek, R.T. MacGillivray, W. Bodemer, F. Scheiflinger, S. Kandels, A. Mitterer, O. Kistner, N. Barrett, J. Eibl, and F. Dorner, *Thromb. Haemost.* **68**, 119–124 (1992).

See also CHINESE HAMSTER OVARY CELLS, RECOMBINANT PROTEIN PRODUCTION; EXPRESSION SYSTEMS, *E. COLI*; EXPRESSION SYSTEMS, MAMMALIAN CELLS; INSECT CELL CULTURE, PROTEIN EXPRESSION; INSECT CELLS AND LARVAE, GENE EXPRESSION SYSTEMS; *PICHIA*, OPTIMIZATION OF PROTEIN EXPRESSION.

TRANSPORT, MICROBIAL SOLUTE UPTAKE

R.E. MARQUIS
University of Rochester
Rochester, New York

KEY WORDS

Antiporters
ATPases
Biofilms
Energy recycling transport
Permeability
Phosphoenolpyruvate
Proton-motive force
Sugar phosphotransferase
Symporters
Transmembrane solute transport

OUTLINE

INTRODUCTION

This article covers transport processes in microorganism of industrial importance. General aspects of transport and uptake of solutes are considered, but as far as possible, examples involving industrial organisms are presented. Figure 1 presents a summary diagram reviewing the major types of transport systems in bacterial cells. Recent reviews from a less applied perspective are cited in the bibliography.

Cell Structure and Solute Movements

Solute transport across membranes plays major roles in the physiology of industrial organisms and may determine which solutes are transformed into products or if the organisms can function under the rigors of industrial process environments. Gram-positive bacteria generally have a single cytoplasmic or protoplast membrane surrounded by a thick cell wall with an elastic peptidoglycan network containing various other polymers and forming a heteroporous molecular sieve that restricts entry of large molecules into the cell. Fungal and plant cells also have thick, heteroporous walls but do not contain the peptidoglycans that are peculiar to eubacteria. The wall is a coarse sieve, whereas the membrane is a finer sieve. In fact, generally only very small solutes pass through the cytoplasmic membrane unless the membrane has specific transport proteins that catalyze movements into or out of the cell. Walled cells generally have what is called osmotrophic nutrition and only small nutrients, for which there are transport systems in the cytoplasmic membrane, can be transformed through the actions of cytoplasmic enzymes. However, bacteria and fungi are noted for secreting hydrolytic enzymes that catalyze breakdown of polymers in the environment. The released fragments then pass through the cell wall to be transported across the cytoplasmic membrane into the cell. Many of these exoenzymes are of major industrial importance, although their primary function for the producing cell is to provide nutrients.

Gram-negative eubacteria generally have two membranes, an inner or cytoplasmic membrane bounding the protoplast, and an outer, cell wall membrane. Both membranes are barriers to solute uptake. The outer membrane is a less-restrictive barrier but does largely prevent movement of larger molecules into the cell unless there are specific receptors for the molecules. Movement even of smaller molecules commonly involves protein porins, which form channels across the membranes. The inner membrane is the equivalent of the cytoplasmic membrane of Gram-positive bacteria and contains transport catalysts for movement of solutes into and out of the cytoplasm. The space between the two membranes contains a thin layer of peptidoglycan, lipoprotein molecules that bridge the peptidoglycan layer and the outer membrane and a variety of enzymes and binding proteins, some of which are involved in solute transport. Small hydrophilic solutes may pass through porins in the outer membrane and be bound by specific binding proteins in the periplasm between the membranes. Binding protein–solute complexes may then interact with specific transport proteins in the cytoplasmic membrane to catalyze transport of the solute into the cell and release of the binding proteins back into the periplasm.

Archeae, which are becoming more and more prominent in biotechnology, have cytoplasmic membranes with ether-linked lipids rather than the ester-linked lipids of other organisms and contain phytanols rather than the more usual types of fatty acids of eubacterial and eukaryotic cells. Various protective envelope structures external to the membrane have been described, but none contains the peptidoglycans characteristic of most eubacterial cells. The membrane transport systems of the *Archeae* are generally similar to systems in eubacterials cells but with differences in the molecular details of the components. Since many *Archeae* are adapted to function in what are considered to be extreme environments, their transport systems are also adapted to function at, for example, high temperatures,

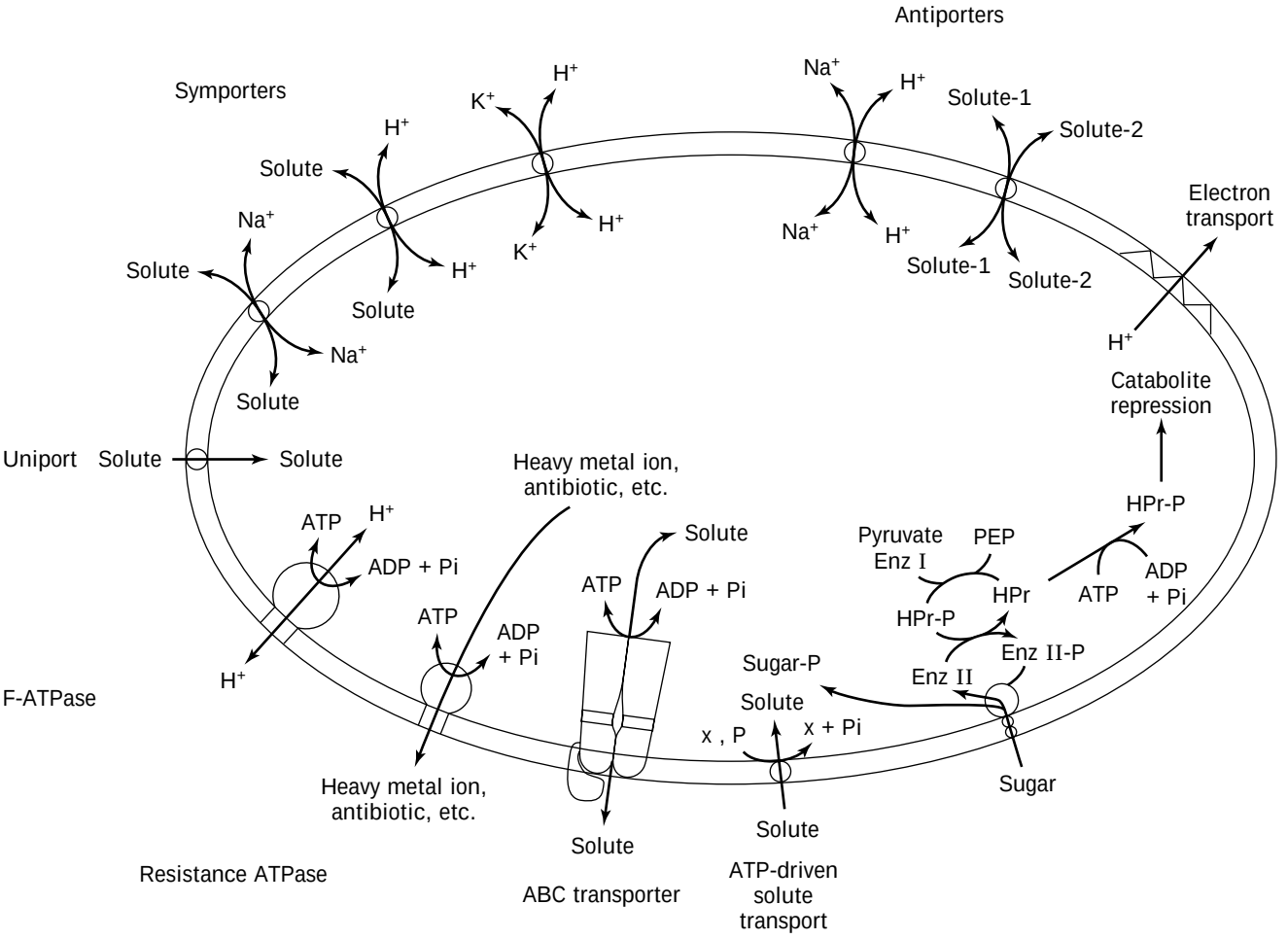


Figure 1. Solute transport systems.

low pH values, very high salt concentrations or low oxidation-reduction potentials.

Eukaryotic organisms used in industrial microbiology include yeasts and other fungi, but also animal cells used, say, for production of monoclonal antibodies, and increasingly, plant cells. Most eukaryotes have multiple membranes that regulate solute transport into and out of the cells and also into and out of the various organelles within in the cells. Indeed, these organisms nearly all have a mix of prokaryotic and eukaryotic characteristics. For example, their major energy-transferring organelles, mitochondria and chloroplasts, are derived from prokaryotic endosymbionts and so have modified prokaryotic membranes and transport systems.

Transport in Biofilms

Special problems of solute transport may be encountered in biofilms, which are important in many biotechnological applications ranging from food fermentations to sewage treatment. Entry of solutes into and out of these consortia of organisms associated with a surface is generally controlled by diffusion across the surface of the biofilm. The situation is often one in which there is flow of a fluid over the biofilm and passage of solues from the fluid across the surface of the film into the polysaccharide matrix and thence into the cells. The passage of solute generally follows the Fick diffusion equation, $dS/dt = PA(\Delta C)$, where dS/dt is the change in the amount of solute in the biofilm per unit of time, P is the permeability coefficient (which has units of distance divided by time), A is the surface area for diffusion, and ΔC is the difference in solute concentration between the interior and exterior of the biofilm surface. Theoretical aspects of antibiotic diffusion into biofilms have been reviewed by Stewart (1), and his analyses can be applied to other solutes.

The biofilm may be mainly exposed to air or other gases. However, the film is still covered by a liquid, and passage, say of O_2, is from the gas phase into the liquid phase and thence into the biofilm. O_2 movement into the film is limited, and biofilms generally contain a diversity of aerobic, facultative, and anaerobic organisms with stratification of the anaerobes away from the surface. Many or most biofilms are penetrated by aqueous channels that may help to deliver solutes to lower parts of the consortium. Nutrients may be leached from solid substrata by biofilms, but again, the limitation on leaching is commonly surface-limited diffusion.

Biofilms have long played key roles in industrial processes ranging from vinegar production to sewage treatment. Our appreciation of them is growing, and they are likely to play even more prominent roles in the future. As mentioned, they allow for wide biodiversity in populations, and this diversity is key in many types of biocatalysis involving cometabolism. They also allow for coexistence of genetically engineered organisms that otherwise would not be competitive. They serve also to mute adverse environmental influences so that members of the community have time to undergo phenotypic or even genotypic adaptations in response to environmental stresses. Certainly, it is well know that cells in biofilms are far more resistant to biocides than are cells in open cultures. However, biofilms do present special challenges in terms of supply of O_2 and other solutes and diffusion of products away from the consortium.

PRIMARY TRANSPORT SYSTEMS

In cells, the primary mechanisms for energy transfer from fuel to function include substrate-level phosphorylations, such as those of glycolysis and other fermentative pathways, and so-called chemiosmotic processes first elucidated by Peter Mitchell. The hypothesis of chemiosmotic energy coupling was based on two fundamental insights. The first involved realization that biological membranes allow for charge separation across a hydrophobic barrier and a conversion of chemical bond energy to electrical energy. During respiration or photosynthesis, protons are moved out of the cell or organelle, such as a mitocondrion or chloroplast, to create a gradient of pH and charge across the membrane. The potential energy of the charge separation was defined by Mitchell in terms of the protonmotive force (Δp) expressed in volts or millivolts

$$\Delta p = \Delta\psi + z\Delta\text{pH}$$

where $\Delta\psi$ is the electrical potential across the membrane (measured experimentally or calculated by use of the Nernst equation), ΔpH is the pH difference across the membrane, and z is equal to 2.3 RT/F with R being the gas constant, T the Kelvin temperature, F the Faraday constant, and the multiplier 2.3 is for conversion of natural log to base 10 log.

The Δp potential energy can then be transferred to synthesis of ATP, to solute movements across the membrane, or to other endergonic processes such as motility. Mitchell's other major insight was that there was associated with the membrane an ATPase that could catabolize synthesis of ATP from ADP and inorganic phosphate in association with proton movements back into the cell. Thus, the electrical energy associated with charge separation across the membrane could be converted back to chemical bond energy. These enzymes are generally called F-ATPases and have the peculiarities that they can operate reversibly and do not become phosphorylated during their catalytic cycles. They are the major agents for chemiosmotic ATP synthesis in respiring and photosynthetic organisms. For example, in the average human, they catalyze synthesis of some 70 kg of ATP per day. The equivalent enzymes of *Archeae* are generally called A-ATPases, and they are able to couple charge separation across the cell membrane to ATP synthesis.

F-ATPases serve major functions even in nonrespiring, fermentative organisms. For example, fermentation commonly results in acid production, and the F-ATPases act to move protons from dissociated acids out of the cytoplasm to develop ΔpH across the cell membrane with the interior alkaline relative to the exterior.

It is now apparent that chemiosmotic energy transfer can involve sodium ions (ΔpNa) and sodium-transporting ATP synthases (2). The ATPases responsible for this type of transfer are distinct from the enzymes responsible for proton-based energy transfer. Proton-based systems are more common than Na^+-based systems, but the latter are widely distributed among prokaryotes.

SECONDARY TRANSPORT SYSTEMS

ATP-Dependent Transport

F-ATPases play a key role in acid-base physiology by transporting protons out of the cytoplasm to reduce internal acidification. This role is clearly important in many industrial fermentations in which acid is produced (e.g., in vinegar production) or in which there is an acidification phase before production of another metabolite (such as in the butanol fermentation by *Clostridium acetobutylicum*). For the bacteria involved in a wide range of processes from yogurt production to acid leaching of copper ores, F-ATPases are needed for the organisms to operate under acidic conditions. Brewing or wine making also are examples in which acidification precedes ethanol production. Yeasts are eukaryotes, and their plasma membrane ATPases, which are not of the reversible F-type, are important for moving protons out of the cell. Also, the vacuole-ATPase (V-ATPase) of the organism appears to be important for acid tolerance. The V-ATPases are multisubunit enzymes with hydrophilic and hydrophobic membrane-embedded components involved in proton pumping or other ion movements. The actions of V-ATPases are not reversible, and so they do not serve for chemiosmotic ATP synthesis. The yeast must then depend on glycolysis for ATP under anaerobic conditions of ethanol production or on the F-ATPases of their mitochondria under aerobic growth conditions.

In bacteria, ATPases other than F-ATPases also are involved in ion movements across the cell membrane. These ATPases are often what are called P-ATPases because they become phosphorylated during their catalytic cycles. Among the best known are heavy-metal-transporting ATPases (3) that play major roles in the resistances of many bacteria to heavy metals such as cadmium. Bacteria also have multisubunit ATPases for the transport of mineral ions such as K^+, Na^+, and CA^{2+}.

ATP can also serve as the energy source for transport of macromolecules across cell membranes, either into the cell or out of the cell, as reviewed by Palmen et al. (4). An example in the dairy industry concerns metabolism of proteins such as casein in manufacture of cheese and related

products. As described by Poolman et al. (5) for *Lactococcus lactis*, the proteolytic pathway involves initial cleavage of proteins by proteinases outside of the cell membrane, in this case mainly the so-called Prt-P proteinase. The cleavage may release free amino acids, which are then transported into the cell by standard amino acid transport systems, mainly energized by Δp or by antiport systems. The products of Prt-P hydrolysis are a wide variety of peptides. Di- and tripeptides can be transported into the cell by means of specific transport systems with relatively broad specificities energized by Δp (Dtp-P transport system) or by ATP (Dtp-T system). The peptides are then broken down in the cytoplasm to amino acids. Oligopeptides can be transported into the cell by the oligopeptide transport system, which is an example of ATP binding cassette (ABC) system consisting of two membrane-spanning proteins, two ATP-hydrolyzing subunits, and a binding subunit associated with the outer face of the cytoplasmic membrane. Oligopeptides are then hydrolyzed in the cytoplasm to amino acids by a variety of peptidases. Oligopeptides are the major source of nitrogen for the organism when it is growing in milk. ABC transporters commonly work in the opposite direction to catalyze movement of polymers out of the cell and so are important for many industrial processes such as production of polysaccharides or enzymes and various other proteins.

Phosphoenolpyruvate-Dependent Transport

The phosphotransferase system (PTS), or as it is now more commonly called, the phosphoenolpyruvate:sugar phosphotransferase system, plays major roles in sugar transport in facultative bacteria. However, it appears that the system may be more widespread among organisms and may play major regulatory roles, especially in catabolite repression (6). The biochemistry of catabolite repression has been elucidated for Gram-positive bacteria such as *Bacillus subtilis*. Repression involves an HPr(Ser)-kinase, which can be activated by fructose-1,6-bisphosphate or gluconate 6-phosphate. The kinase can phosphorylate the HPr protein of the PTS at serine-46. The HPr-P can then combine with CcpA (the repressor-activator protein mediating catabolite repression and glucose activation) and fructose-1,6-bisphosphate to form a complex that binds to the cis-acting DNA sequence called the catabolite repression element (CRE) to inhibit transcription. Thus, many Gram-positive bacteria are able to regulate metabolism in response to rises and falls in cytoplasmic levels of glycolytic intermediates. In Gram-negative organisms, catabolite repression involves cyclic AMP and cyclic-AMP-receptor protein, as described in standard textbooks. Catabolite repression is important in many industrial microbiology processes and allows for sequential utilization of mixed substrates, often to allow mainly growth initially and product formation subsequently, for example, in many antibiotic fermentations.

A major function of the PTS is uptake of sugars. The system has common elements, specifically HPr protein and enzyme I, which catalyze transfer of phosphate from phosphoenolpyruvate to a histidine moiety on HPr. The PTS has elements designed for specific sugars, notably IIC proteins or peptides, which are the actual permeases embedded in the cytoplasmic membrane for catalysis of sugar transport across the cell membrane and phosphoryl transfer from enzyme IIB to the incoming sugar. For many of the industrial lactic acid bacteria, the PTS is the major agent for sugar uptake, although nearly all PTS-positive bacteria also have alternative systems for sugar uptake, especially if sugar fermentation plays a major role in their physiology. Thus, for example, dairy lactic acid bacteria may take up milk sugar lactose by a PTS or by transport systems energized either by Δp or ATP.

Proton-Motive Force–Dependent Transport

Coupling of Δp to transport is a very direct means for energy transfer. This coupling is dependent on what is often referred to as *the energized membrane*. The advantage to the organism in direct coupling is obvious, and direct coupling occurs not only for solute transport but also for functions such as motility. However, there can also be disadvantages because the energized state of the membrane is very sensitive to environmental changes. Energizers such as ATP or PEP generally are not released from cells unless there is major damage to membranes. However, Δp can be discharged by a variety of mild insults including exposure to weak acids, which can enter the cell in the protonated state and dissociate within the cytoplasm to diminish the ΔpH component of Δp. This sort of discharge can occur to a degree when weak acids such as formate, acetate, or lactate are produced metabolically. The discharge then puts additional energy demands on the organism, requiring additional catabolism for growth and maintenance of cell functions. The uncoupling may actually be an advantage in many fermentation processes, permitting high production of metabolic products with less production of cells. The discharge of ΔpH can also be of advantage in applications such as the use of fluoride to inhibit acid production by the plaque bacteria that cause dental caries. Fluoride acts in the protonated form, HF, to bring extruded protons back into the cell and thereby to reduce acid tolerance and the capacities of cariogenic bacteria to produce acid at low pH values (7).

Δp-dependent systems occur widely for uptake of small solutes. For example, uptake of many sugars is coupled to proton reentry. Thus entry of galactose, lactose, rhamnose, fucose, and so forth can be coupled through symport proteins to proton movements back into the cell. Many organisms also have PTS systems for the same sugars. We have found for oral streptococci that use of a system for galactose uptake coupled to proton movements, rather than the PTS, actually reduces acid tolerance of the organisms (8).

Δp-coupled uptake and excretion of a wide variety of organic solutes is known to involve proton symport or antiport. For example, lysine is taken up by *L. lactis* through a proton symport mechanism. In the same organism, the transport systems for Leu/Ile/Val, Ala/Gly and Ser/Thr are also dependent on Δp (9). Citrate uptake is commonly coupled to movements of protons back into the cell. Antibiotic

resistance in many organisms may have its base in transport systems that move antibiotics out of the cell. Perhaps the best example is that of tetracycline efflux systems in a number of organisms involving antiport of tetracycline out and protons or K^+ in. Multidrug efflux pumps occur also in a variety of organisms and have recently been reviewed by Nikaido (10).

Now that the specifics of Na^+-based energetics are well established (2), it is clear that organisms such as *Propionigenium modestum* depend on ΔpNa, rather than ΔpH, for coupling transport of various solutes to movements of Na^+. In any organism, there can be a mix of Δp and ΔpNa coupling, but in general in the biological world, Δp coupling is predominant.

In general, the biopolymers within cells have a net negative charge and a need for cations to neutralize the excess charge. The potential across the cell membrane ($\Delta\psi$) reflects a nonequilibrium state in metabolizing cells. However, it is a component of Δp that can serve to energize transport. Generally, it is considered that $\Delta\psi$ can be converted readily to ΔpH, and so it is often difficult to separate $\Delta\psi$-driven processes from ΔpH-driven processes. It should be appreciated that a large fraction of the cations within cells, predominantly K^+, are retained simply for charge neutralization. If the cell membrane is damaged in a walled organism, such as a eubacterial cell, there is an efflux of solutes, but there is also a retention of cations as counterion for polymers that cannot exit readily through the porous wall. Of course, when the membrane is damaged, the major agent for selectivity of solute entry is damaged, and cation exchange can occur readily. Movements of mineral ions across membranes by symport and antiport play major roles in cell physiology. A good example is in the capacities of many organisms, predominantly bacteria, to function at alkaline pH values. The most common basis for this ability is H^+/Na^+ antiporters that move protons into the cell to reduce the cytoplasmic pH value in association with movement of Na^+ out.

Energy Recycling Systems

Microorganisms often are able to use transport systems for energy conservation. One of the early proposals for this sort of conservation involved glycolysis by dairy streptococci (11). Glucose is catabolized by these organisms to yield lactate. At higher pH values, lactate excretion is electrogenic because more than one proton is excreted per lactate anion. This excretion then results in Δp across the cell membrane, which can be converted chemiosmotically to chemical energy in the form of ATP through the action of F-ATPases.

Another example pertinent to industrial fermentations is that of the malolactic fermentation carried out by *Oenococcus oenos* and related organisms and important in production of wine, cider, and a variety of fermented beverages. Malate is transported into the cell where it is decarboxylated by the malolactate enzyme to yield CO_2 and lactate, which are then moved out across the membrane in association with protons to create Δp (12). The use of the conserved energy in the Δp can then allow for some 25 to 50% increase in growth yields of the organisms.

Symporters and Antiporters Not Coupled to Δp or ΔpNa

As already reviewed, the energetic coupling of Δp or ΔpNa to solute movements into and out of cells commonly involves mechanisms of symport and antiport. Thus, the potential energy of a gradient of protons or of sodium ions is translated into kinetic energy of movement of other solutes, or the electrical potential across the cell membrane may be the basis for movement of counterions into the cell to reduce the potential. Other types of symporters and antiporters may not be so directly coupled to the energized state of the membrane. A good example is the ornithine/arginine antiport system of many organisms that use the arginine deiminase system (ADS) for ATP synthesis or for neutralization of cytoplasmic protons through production of ammonia (13). The ADS involves three enzymes, arginine deiminase, ornithine carbamyl transferase, and carbamate kinase, plus an antiporter. ATP is synthesized in association with the carbamate kinase–catalyzed reaction. Ornithine is produced in the reaction catalyzed by ornithine carbamyl transferase. It exchanges with arginine so that substrate is brought into the cell, and product is excreted without the need to make use of the ATP produced by the system. This exchange allows for operation with a net gain in ATP, and many organisms can grow with arginine as fuel without need for respiration. In fact, oxygen often is repressive for the ADS. CO_2 is produced by the system, and it can diffuse out of the cell. The other main product, NH_3, reacts with protons within the cytoplasm and serves to maintain acid-base balance in many organisms.

TRANSPORT ASSAYS

Various methods used to assay transport of molecules into cells have been reviewed recently (14), with distinction made between assays primarily for determining permeability of a cell or cell component to a particular compound, and those for assessing the activities of transport systems for active movement of solutes across the cell membrane. Also, methods are reviewed for assessing Δp from its two components, ΔpH and $\Delta\psi$, and for assessing osmotic responses of cells.

In practical situations, cell permeability often is key for entry of solutes into cells. For example, gases are hydrophobic and generally pass readily through hydrophobic membranes. In industrial fermentation in which O_2 must be supplied to aerobic or facultative organisms, the major barrier to be overcome in supplying the gas is at the boundary between the gas phase and the liquid phase. O_2 is relatively insoluble in aqueous solution, and so a high concentration cannot be built up in the aqueous phase of a culture. For a rapidly respiring microbial population, supplying sufficient O_2 can be nearly impossible. Efforts are made then to increase the surface area for diffusion of O_2 into the culture by sparging. Also, pure oxygen may be used to increase supplies. Still, it is common for the cells to be starved for O_2, which passes readily from the medium into the cell as it becomes available. Metabolically produced gases, such as CO_2 and H_2, also pass readily across the cell membrane so that cells do not build up stores of

the gases. In fact, gas bubbles generally do not form within cells unless special structures are produced by the cell, such as the gas vesicles used for flotation by certain aquatic microorganisms (15).

Perhaps the most important solute that enters cells without specific transport mechanisms is water. All cells are highly permeable to water, even bacterial spores; and when cells are mixed with water labeled with tritium or deuterium, equilibration of the labeled water generally occurs in seconds. Indeed, osmotic equilibration of water can occur in milliseconds. Other small solutes may also penetrate rapidly. In fact, labeled urea or glycerol is commonly used to assess the total water space of cells because of their ready movement into the total water volume of cells. Recently, the roles of ubiquitous major intrinsic proteins (MIP) in movements of water and neutral solutes across membranes have been studied (16). Hydrophobic solutes may also readily enter biological membranes, especially if they are of low molecular weight. Higher molecular weight hydrophobic solutes also may enter the membrane readily but may remain there and cause membrane dysfunction (17). Mention was made earlier of the effects of organic weak acids on membranes related to movements of the protonated forms into and through membranes (18).

For measurements of passive permeability of cells to solutes, thick suspensions of cells can be used, and the disappearance of solute from the suspending medium assayed. The uptake process is generally not saturable at high solute concentrations, although cells may passively concentrate solutes, especially at low concentration, because of passive binding, for example, binding of cations to negatively charged polymers in the cell wall, membrane, or cytoplasm. For assays of active transport, much less dense suspensions of cells are generally used. Cells and suspending medium may then be separated by centrifugation or by membrane filtration. In some cases, there is no need to make the separation. For example, in assays of proton movements into cells, the pH electrode detects protons outside the cell but not those within the insulating membrane. Thus, as protons move into the cell, the pH value measured by the electrode rises. From the level of pH rise and knowledge of the volumes of cells and suspending medium, the flow of protons across cell membranes can be calculated. When transport depends on metabolism, provision must be made for such metabolism, which may supply ATP, serve to energize the membrane, or produce solutes that may exchange with other solutes in the medium, for example, in arginine/ornithine antiport. In many cases, it is difficult to allow for adequate metabolism, and for example, measurements of movements of protons or other solutes into cells may be affected very much by the experimental conditions. Maintenance of Δp is particularly sensitive to changes in experimental conditions. The cell membrane is at least moderately permeable to protons, and cells must work to maintain ΔpH. If the proton-extruding functions of the F-ATPase are upset, or if proton permeability of the membrane is increased, say, by weak acids, then ΔpH declines and the energized state of the membrane declines. Similarly, $\Delta\psi$ may also be affected adversely, say, by agents such as gramicidin or lantibiotics that enhance membrane permeability to ions.

REGULATION OF TRANSPORT

Transport systems may be regulated in association with the catabolic systems they serve. The best-known example is that of the *lac* operon. The operon contains the gene *lacY* for lactose permease. The operon in *Escherichia coli* is derepressible by β-galactosides and also controlled by catabolite repression involving cyclic AMP and cyclic-AMP-receptor protein. Thus, the permease system is synthesized only when needed for transport of β-galactosides into the cell and maintained at a very low level when β-galactosides are not present. However, genes for transport proteins need not be in the same operon, or even the same regulon, as the genes for metabolism of the transported substrates. An example pertinent to industrial microbiology involves the inducible transport systems for citrate in organisms such as *Lactobacillus rhamnosus* (19). This organism and related organisms are important for flavor development in fermented products through production of diacetyl and acetoin. These products are produced in association with increased consumption of citrate or pyruvate from the environment. The transport system for citrate is repressed in the presence of glucose but becomes derepressed when glucose is exhausted. Thus, flavor enhancement follows derepression of citrate transport after the main phases of growth and acid production.

As might be expected, key enzymes such as the F-ATPases, are mainly constitutive. There is adaptive regulation of F-ATPase synthesis, but the excursion from minimum to maximum levels is small. Thus, for example, for *Streptococcus mutans* and *Enterococcus hirae*, the range is only three to four times the minimal level (20). These organisms depend on their F-ATPases for maintaining acid-base balance, and so appear not to be able to tolerate repression of synthesis of the enzymes to low values. However, upregulation of synthesis of the enzymes does result in increased acid tolerance, and this sort of adaptive tolerance occurs widely among bacteria. It is of importance in many industrial processes in allowing a variety of organisms to operate under adverse conditions.

One of the best examples pertinent to a wide variety of industrial organisms involves the adapation to adverse osmotic conditions of reduced water activity. Adaptation involves pooling of compatible solutes such as K^+, glycine betaine, proline, or polyols and generally involves upregulation of transport systems and enhanced synthetic capacities (21).

INDUSTRIAL EXAMPLES AND APPLICATIONS

Current knowledge of the molecular details of many of the transport systems of industrially important microorganisms is sufficiently advanced to allow for genetic engineering of improved strains, although to date the major work has been on engineering organisms so that they have systems for export of proteins coded by transferred genes. The range of possibilities for advantageous manipulation of

transport systems is wide. It is generally considered that organisms that are adapted to oligotrophic environments have transport systems geared for obtaining nutrients from very dilute solutions. Part of their strategy for survival involves formation of biofilms to take advantage of the associations of molecules with interfaces. However, they also need to have transport systems with low K_m (Michaelis–Menten coefficient) values. Thus, they are a source for genes for high-affinity transport systems. The possibilities for developing improved strains include extending the range of substrates that can be used by any one organism. For example, most strains of *Saccharomyces* yeasts used for ethanol production are not very versatile in the range of sugars they can utilize. *Kluyveromyces fragilis* is able to metabolize the sugar lactose, and so there is a possibility of developing hybrids that will allow *Saccharomyces* organisms to use milk sugar for alcohol production. The need to extend the metabolic range of an organism is evident also in development of organisms to transform various xenobiotics. Generally, the xenobiotic must be transported before it can be metabolized. Fortunately, many organisms have two separate transport systems for, say, an amino acid: one highly specific and one with lower specificity. Often xenobiotics can be taken up by the system with lower specificity. However, there are still possibilities for engineering high-affinity systems to the level that they favor the xenobiotic over a related natural compound.

Other examples include work on a variety of organisms for antibiotic production. Generally, producing organisms need to be resistant to the antibiotics they produce, and for example, resistance to tetracyclines can depend on efflux systems. Resistance to heavy metals also commonly depends on efflux systems, often involving efflux ATPases. Mention has already been made of bacterial adaptation to acid conditions and the role of proton-transport systems in the adaptation. However, adaptation to alkali environments in certain foods or in industrial processing also depends very much on transport systems, often K^+/proton antiport or Na^+/proton antiport (22).

Another area in which there has been a great deal of activity concerns excretion of proteins by various genetically engineered organisms. Often simply transferring the genes for synthesis of a particular protein from one organism to another does not result in a useful producing strain. In fact, the transfer can be lethal if the protein cannot be excreted.

ACKNOWLEDGMENT

The work of the author has been supported by grants (P01-DE11549 and R01-DE06127) from the National Institute of Dental Research and by the Center for Aseptic Processing and Packaging Studies, which is a U.S. National Science Foundation Industry/University Cooperative Research Center.

BIBLIOGRAPHY

1. P.S. Stewart, *Antimicrob. Agents Chemother.* **40**, 2517–2522 (1996).
2. P. Dimroth, *Phil. Trans. R. Soc. Lond. B* **326**, 465–477 (1990).
3. G. Ji and S. Silver, *J. Indust. Microbiol.* **14**, 61–75 (1995).
4. R. Palmen, A.J.M. Driessen, and K.J. Hellingwerf, *Biochim. Biophys. Acta* **1183**, 417–451 (1994).
5. B. Poolman, E.R.S. Kunji, A. Hagting, V. Juillard, and W.N. Konings, *J. Appl. Bacteriol.* **79**, 65S–75S (1995).
6. M.H. Saier, Jr., S. Chauvaux, J. Deutscher, J. Reizer, and J.-J. Ye, *Trends Biochem. Sci.* **20**, 267–271 (1995).
7. R.E. Marquis, *Can. J. Microbiol.* **41**, 955–964 (1995).
8. W.A. Belli and R.E. Marquis, *Oral Microbiol. Immunol.* **9**, 29–34 (1994).
9. B. Poolman, A.J.M. Driessen, and W.N. Konings, *Microbiol. Rev.* **51**, 498–508 (1987).
10. H. Nikaido, *J. Bacteriol.* **178**, 5853–5859 (1996).
11. R. Otto, A.S.M. Sonnenberg, H. Veldkamp, and W.N. Konings, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5502–5506 (1980).
12. T. Henick-Kling, *J. Appl. Bacteriol.* **79**, 29S–37S (1995).
13. A.J.M. Driessen, D. Molenaar, and W.N. Konings, *J. Biol. Chem.* **264**, 10361–10370 (1989).
14. R.E. Marquis, in P. Gerhardt ed., *Methods for General and Molecular Bacteriology*, American Society for Microbiology, Washington, D.C., 1994, pp. 587–599.
15. A.E. Walsby, *Microbiol. Rev.* **58**, 94–144 (1994).
16. J.H. Park and M.H. Saier, Jr., *J. Membrane Biol.* **153**, 171–180 (1996).
17. J.J. Sikkema, A.M. de Bont, and B. Poolman, *J. Biol. Chem.* **269**, 8022–8028 (1994).
18. W.A. Belli, D.H. Buckley, and R.E. Marquis, *Can. J. Microbiol.* **41**, 785–791 (1995).
19. R.M. de Figueroa, I.L. Benito de Cárdenas, F. Sesma, F. Alvarez, A.P. de Ruiz Holgado, and G. Oliver, *J. Appl. Bacteriol.* **81**, 348–354 (1966).
20. W.A. Belli and R.E. Marquis, *Appl. Environ. Microbiol.* **57**, 1134–1138 (1991).
21. L.N. Csonka and A.D. Hanson, *Ann. Rev. Microbiol.* **45**, 569–606 (1991).
22. R.J. Rowbury, Z. Lazim, and M. Goodson, *J. Appl. Bacteriol.* **22**, 429–432 (1996).

ADDITIONAL READING

W.N. Konings, B. Poolman, and H.W. van Deen, *Antonie van Leeuwenhoek* **65**, 369–380 (1994).
P. Mitchell, *Chemiosmotic Coupling and Energy Transduction*, Glynn Research, Ltd., Bodmin, England, 1968.
B. Poolman and W.N. Konings, *Biochim. Biophys. Acta* **1183**, 5–39 (1993).

See also MASS TRANSFER; TRANSFER PHENOMENA IN MULTIPHASE SYSTEMS IN MIXING VESSELS.

TRYPTOPHAN. See AMINO ACIDS, PRODUCTION PROCESSES.

TYROSINE PHENOL-LYASE

HIDEHIKO KUMAGAI
Kyoto University
Kyoto, Japan

KEY WORDS

Enzymes
Gene cloning
L-DOPA
β-Tyrosinase
L-Tyrosine

OUTLINE

INTRODUCTION

Tyrosine phenol lyase (TPL) is a pyridoxal 5′-phosphate (PLP)–dependent multifunctional enzyme catalyzing a degradation reaction of L-tyrosine into phenol, pyruvate, and ammonia. TPL is produced by enterobacteria grown in the medium containing L-tyrosine as an inducer. A crystalline preparation of the enzyme was obtained from two bacterial strains, *Escherichia intermedia* and *Erwinia herbicola*, and their properties were elucidated. The genes of TPL (*tpl*) were cloned, and their nucleotide sequences were determined. The molecule of TPL consists of four identical subunits, each with a molecular mass about 50 kDa, and binds four molecules of PLP per tetramer. The three-dimensional structure of apoenzyme form of TPL (apo TPL) has been refined at 2.3-Å resolution. TPL biosynthesis is regulated transcriptionally, and the DNA binding regulator proteins, TyrR and CRP, are working in this regulation system.

The enzyme catalyzes the synthesis reaction of 3,4-dihydroxyphenyl-L-alanine (L-DOPA) from pyruvate, ammonia, and pyrocatechol in a one-step reaction. L-DOPA is an effective agent in the treatment of Parkinson's disease.

Kakihara and Ichihara (1) reported that phenol is produced from L-tyrosine by an intestinal bacteria, not through stepwise degradation, but through primary fission of the side chain. The tyrosine-inducible enzyme responsible for this conversion was named *β*-tyrosinase and elucidated to be PLP dependent. The recommended name of this enzyme is tyrosine phenol lyase (deaminating) (EC 4.1.99.2).

Apparently homogeneous preparations of TPL were prepared from cells of *Escherichia intermedia* and *Erwinia herbicola* grown on media supplemented with L-tyrosine. The crystalline preparations of the enzyme catalyze a series of *α,β*-elimination (equations 1 and 2), *β*-replacement (equations 3–5), and racemization (equation 6) reactions. The reverse of the *α,β*-elimination reaction to synthesize L-tyrosine or L-DOPA (equations 7 and 8) was also found to be catalyzed by the crystalline preparations of the enzyme (2).

$$\text{L-Tyrosine} + H_2O \rightarrow \text{pyruvic acid} + NH_3 + \text{phenol} \quad (1)$$

$$\text{L-Serine} \rightarrow \text{pyruvic acid} + NH_3 \quad (2)$$

$$\text{L-Tyrosine} + \text{pyrocatechol} \rightarrow \text{L-DOPA} + \text{phenol} \quad (3)$$

$$\text{L-Serine} + \text{phenol} \rightarrow \text{L-tyrosine} + H_2O \quad (4)$$

$$\text{L-Serine} + \text{pyrocatechol} \rightarrow \text{L-DOPA} + H_2O \quad (5)$$

$$\text{L-Alanine} \rightarrow \text{DL-alanine} \quad (6)$$

$$\text{Pyruvic acid} + NH_3 + \text{phenol} \rightarrow \text{L-tyrosine} + H_2O \quad (7)$$

$$\text{Pyruvic acid} + NH_3 + \text{pyrocatechol} \rightarrow \text{L-DOPA} + H_2O \quad (8)$$

The bacterial cells containing this enzyme catalyze the synthesis of L-DOPA from pyruvic acid, ammonia, and pyrocatechol in significantly high yields, and this reaction is used in industrial production of L-DOPA (3).

PROPERTIES OF TPL

TPLs have been purified and crystallized in the laboratory (4,5,6) from cells of *Escherichia intermedia* and *Erwinia herbicola* grown in a medium supplemented with L-tyrosine by a procedure including cell disruption by ultrasonic oscillation, ammonium sulfate fractionation, protamine sulfate treatment, DEAE-Sephadex column chromatography, hydroxyapatite column chromatography, Sephadex G-150 gel filtration, and crystallization. Photomicrographs of the crystalline holoenzyme prepared from *Erwinia herbicola* are shown in Figure 1.

Crude and partially purified preparations of the TPL were preserved at 5 °C in 0.01 M potassium phosphate buffer, pH 6.0, containing 5 mM mercaptoethanol. Variations in pH in the range of pH 6.0 to 7.0 had little effect on enzyme stability. Outside this range, however, the rate of inactivation increased rapidly. Solutions of highly purified enzyme were less stable than those of partially purified preparations. For example, a sample of the enzyme from *Escherichia intermedia*, with a specific activity of 1.94, lost to 6 to 10% of its initial activity after standing for 2 days at 5 °C in 0.01 M potassium phosphate buffer, pH 6.0, containing 5 mM mercaptoethanol.

The crystalline enzyme preparations could be stored at 0 to 5 °C as suspensions in 10 mM potassium phosphate buffer, pH 6.0, containing 5 mM mercaptoethanol and 60% saturated ammonium sulfate. The specific and total activity of the enzyme remained constant for periods of more than 1 month. Preservation in this fashion constituted the best storage procedure.

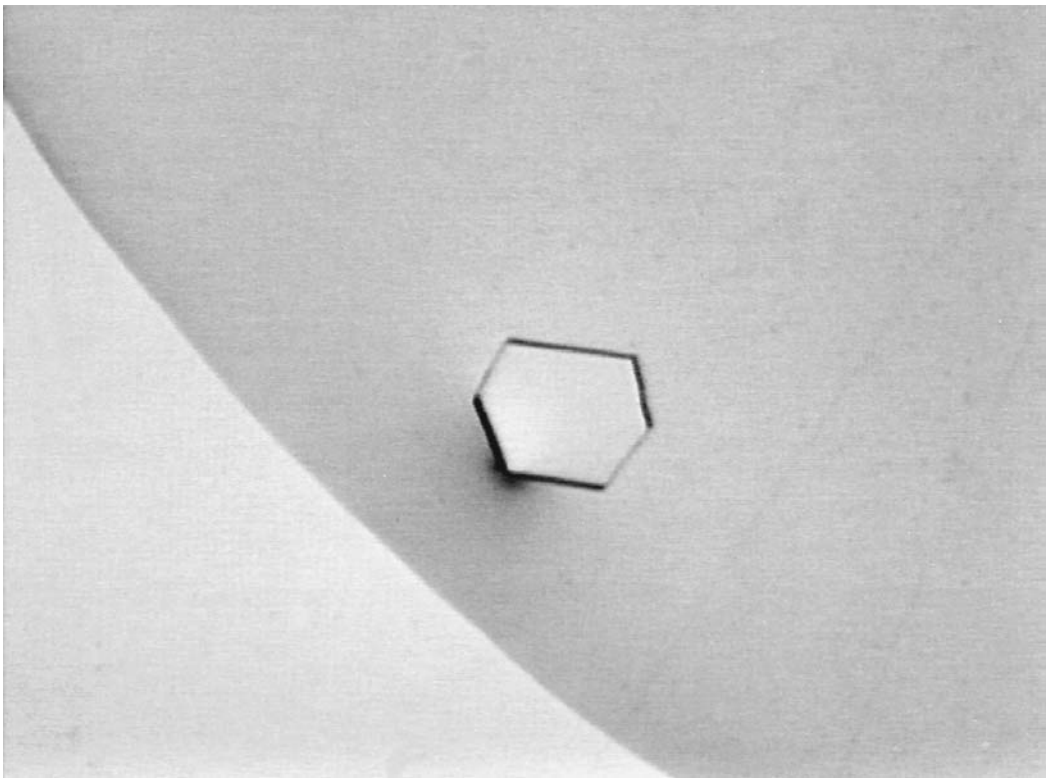


Figure 1. A photomicrograph of the crystalline holo TPL prepared from *Erwinia herbicola*.

On heating, the crystalline enzyme was rather stable up to 40 °C, but above 50 °C, it was rapidly inactivated.

Recrystallized enzyme preparations gave a single band on acrylamide gel electrophoresis. The molecular weight of the enzymes was estimated about 200,000, and that of the subunit is about 50,000. The gene cloning and subsequent nucleotide sequencing gave the molecular mass of the subunit 51,439 Da for *Escherichia intermedia* TPL and 51,364 Da for *Erwinia herbicola* TPL (7,8). These data indicate that TPL has homotetramer structure.

The enzyme has been shown to have cofactor requirements (5). Crystalline TPLs prepared by the above procedure showed negligible activity when assayed in the absence of PLP; they are purified as the apoenzyme. Half the maximum enzymic activity was obtained at the respective PLP concentrations of 1.3×10^{-6} M and 1.5×10^{-6} M for the enzymes from *Escherichia intermedia* and *Erwinia herbicola*. On association with PLP, pronounced absorption maxima appeared at 340 and 430 nm. Absorption in the 410 to 440 nm region is characteristic of many PLP enzymes and has been attributed to a hydrogen bound azomethine of PLP. The absorption maxima of holo TPLs were not appreciably shifted by variations in pH between 6.0 and 9.0.

The crystalline enzymes from *Escherichia intermedia* and *Erwinia herbicola* showed a similar requirement for K^+ and NH_4^+ for their maximum activities. However, Na^+ was inactive.

The amount of PLP bound by the apoenzyme was determined by chemical and spectrophotometric titration methods. The results obtained from these two procedures indicated that there are two PLP binding sites per molecule of enzyme. The crystallographic studies showed the existence of sulfate residue at the binding site of phosphate residue of PLP, suggesting the prevention of PLP binding to the binding site. The proper number of PLP molecules bound is estimated to be four per tetramer.

The catalytic properties of the enzyme have been determined with the crystalline TPL from *Escherichia intermedia*. The enzyme catalyzed a series of α,β-elimination, β-replacement, recemization, and reverse reaction of α,β-elimination (equations 1 to 8).

α,β-Elimination Reaction

TPLs catalyze the conversion of L-tyrosine into phenol, pyruvate, and ammonia in the presence of PLP (3,4). Pyruvate formation was also observed with β-chloro-L-alanine, *S*-methyl-L-cysteine, L- and D-serine, and L- and D-cysteine.

L-Alanine competitively inhibits pyruvate formation from L-tyrosine, *S*-methyl-L-cysteine, and L-serine. Phenol, pyrocatechol, and resorcinol also inhibit pyruvate formation, but the inhibition was a mixed type. The α,β-elimination reactions proceed optimally around pH 8.2.

β-Replacement Reaction

TPL catalyzed the synthesis of L-tyrosine from L- or D-serine and phenol in the presence of PLP. β-Chloro-L-alanine, *S*-methyl-L-cysteine, and L-cysteine replaced L- or

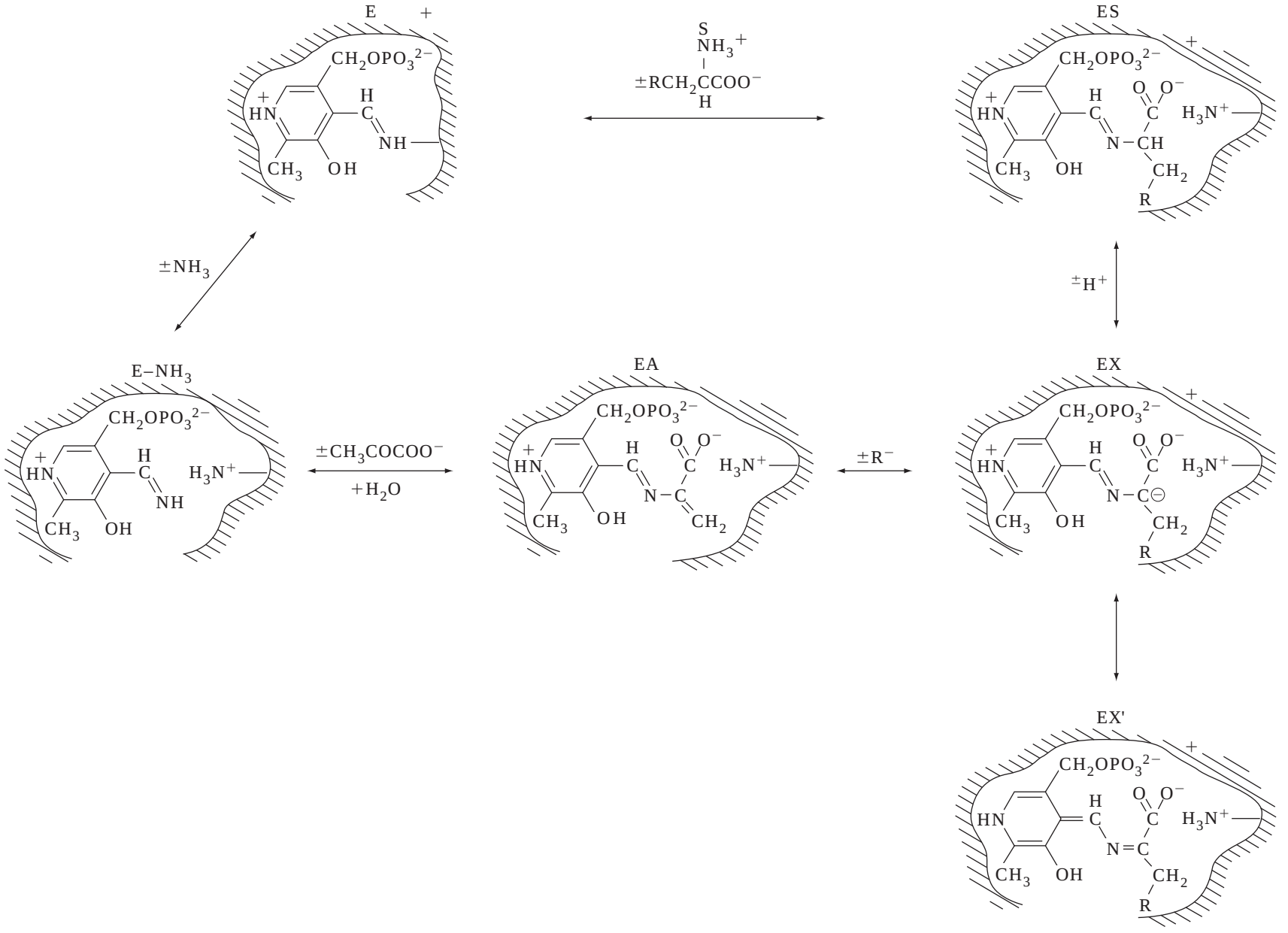


Figure 2. Schematic representation of the mechanism for the reactions catalyzed by TPL. Different species of the enzyme and enzyme-substrate complexes are represented.

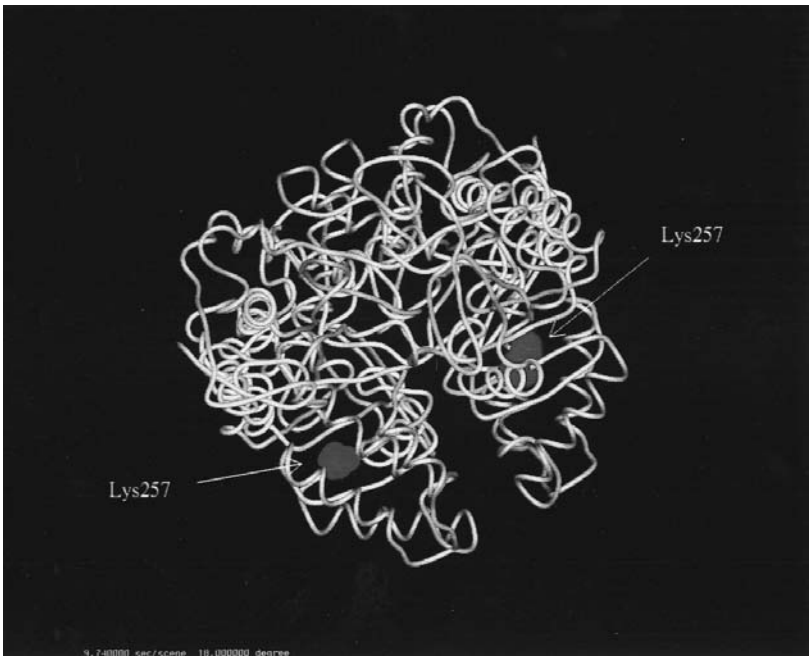


Figure 3. Three-dimensional structure of TPL from *Erwinia herbicola* estimated by computer modeling. TPL has a homotetramer structure, and only dimer structure is shown in this figure. Lysine 257 residues where coenzyme PLP bind are shown by arrows.

D-serine in the synthesis of L-tyrosine (3). When pyrocatechol, resorcinol, pyrogallol, and hydroxyhydroquinone were added to the reaction mixture in place of phenol, L-DOPA, 2,4-dihydroxyphenyl-L-alanine (2,4-L-DOPA), 2,3,4-trihydroxyphenyl-L-alanine (2,3,4-L-TOPA), and 2,4,5-trihydroxyphenyl-L-alanine (2,3,5-L-TOPA) were, respectively, synthesized. These β-replacement reactions proceed optimally at pH 8.0 to 8.5.

Racemization Reaction

The conversion of D- or L-alanine into the racemate is catalyzed by TPL in the presence of PLP (3). The K_m value for L-alanine was 2.6 mM and the maximum velocity of the racemization reaction was 0.05 μmol/min per milligram of protein. The optimum pH for this racemization reaction is between 7.2 and 7.5.

Synthetic Reaction

The synthesis of L-tyrosine from phenol, pyruvate, and ammonia is catalyzed by TPL in the presence of PLP (2,3). This synthesis proceeds as a function of the concentration of phenol, pyruvate, and ammonia. *K* values were obtained by using three substrate kinetics (3). The K_m values for phenol, pyruvate, and ammonia calculated were 1.1 mM, 12 mM, and 20 mM, respectively, and the maximum velocity was 3.3 μmol/min per mg of protein, which is about 1.5 times higher than that of L-tyrosine degradation by the reaction of equation 1.

When pyrocatechol and resorcinol were, respectively, added to the reaction mixture in place of phenol, L-DOPA and 2,4-L-DOPA were synthesized. Cresols (*m*- and *O*-) and chlorophenols (*m*- and *O*-) also replaced phenol to synthesize, respectively, methyl-L-tyrosine (2- and 3-) and chloro-L-tyrosine (2- and 3-) (3).

REACTION MECHANISM

The crystalline preparations of TPL from *Escherichia intermedia* and *Erwinia herbicola* catalyze a variety of α,β-elimination, β-replacement, racemization, and synthetic reactions (3). These reactions are partially explained by adopting the general mechanism for pyridoxal-dependent reactions proposed by Braunstein and Shemyakin (9) and by Metzler et al. (10). This mechanism is shown in Figure 2, with a modification in which PLP in the enzyme is shown as an internal azomethine. The enzyme-bound α-aminoacrylate (EA) is the key intermediate in this proposed mechanism. The α,β-elimination reaction is considered to proceed via steps

$$E + S \rightarrow ES \xrightarrow{-H^+} EX \xrightarrow{-R^-} EA \xrightarrow{+H_2O} E + \text{pyruvate} + NH_3$$

The β-replacement reaction proceeds via steps

$$E + S \rightarrow ES \xrightarrow{-H^+} EX \xrightarrow{-R^-} EA \xrightarrow{+R'^-} EX' \xrightarrow{+H^+} ES' \rightarrow E + S'$$

and the racemization reaction proceeds via steps

$$E + L - S \rightarrow E \cdot L - S \xrightarrow{-H^+} EX \xrightarrow{+H^+} E \cdot DL - S \rightarrow E + DL - S$$

The synthetic reaction of L-tyrosine seems to proceed via the reversal steps of α,β-elimination. The stereochemistry and reaction mechanism of TPL were investigated and reported (11).

GENE CLONING AND STRUCTURE ANALYSIS

TPL genes from *Escherichia intermedia* (12), *Cirobacter (Escherichia) freundii* (7), and *Erwinia herbicola* (8,13) were cloned and their nucleotide sequences were deter-

mined. The amino acid sequences deduced from the nucleotide sequences showed the following homology: 100% identical between *Citrobacter freundii* and *Escherichia intermedia*; 90.6% between *Escherichia intermedia* and *Erwinia herbicola*. The PLP binding site of TPL is determined as Lys 257 (8). The three-dimensional structure of apo TPL of *Escherichia intermedia* was analyzed by X-ray crystallography at 2.3 Å resolution and reported after refinement with primary amino acid sequence of the *tpl* from *C. freundii*. The three-dimensional structure of apo TPL was estimated from that of *Escherichia intermedia* by computer analysis (Fig. 3).

REGULATION OF TPL BIOSYNTHESIS

Induction and repression of TPL of *Erwinia herbicola* were regulated on the transcriptional level; mRNA of TPL was increased by the addition of tyrosine and decreased by the addition of glucose in the medium. The 5′ flanking region of its gene, *tpl*, was analyzed and its transcriptional start point was determined. TyrR box and operator-like region were also found in this region (14). TyrR box is a typical binding site of DNA where a regulator protein TyrR binds and regulates transcription of structure genes of enzymes or transporters responsible for aromatic amino acid biosynthesis or transportation.

BIBLIOGRAPHY

1. Y. Kakihara and K. Ichihara, *Med. J. Osaka Univ.* **3**, 497–507 (1953).
2. H. Yamada, H. Kumagai, N. Kashima and H. Torii, *Biochem. Biophys. Res. Commun.* **46**, 370–374 (1972).
3. H. Yamada and H. Kumagai, in D. Perlman ed., *Advances in Applied Microbiology*, vol. 19, Academic Press, New York, 1975, pp. 249–288.
4. H. Kumagai, H. Yamada, H. Matsui, H. Ohkishi and K. Ogata, *J. Biol. Chem.* **245**, 1767–1772 (1970).
5. H. Kumagai, H. Yamada, H. Matsui, H. Ohkishi and K. Ogata, *J. Biol. Chem.* **245**, 1773–1777 (1970).
6. H. Kumagai, K. Kashima, H. Torii, H. Yamada, H. Enei, and S. Okumura, *Agric. and Biol. Chem.* **36**, 472–482 (1972).
7. Y. Kurusu, M. Fukusima, K. Kohama, M. Kobayashi, M. Terasawa, H. Kumagai and H. Yukawa, *Biotechnol. Lett.* **13**, 762–772 (1991).
8. H. Suzuki, K. Nishihara, N. Usui, H. Matsui and H. Kumagai, *J. Ferment. Bioeng.* **75**, 145–148 (1993).
9. A.E. Braunstein and M.M. Shemyakin, *Biokhimiya* **18**, 393–411 (1953).
10. D.E. Metzler, M. Ikawa, and E.E. Snell, *J. Am. Chem. Soc.* **76**, 648–652 (1954).
11. M.M. Palcic, S.-J. Shen, E. Schleicher, H. Kumagai, S. Sawada, H. Yamada and H.G. Floss, *Z. Naturforsch.* **42C**, 307–318 (1987).
12. S. Iwamori, S. Yoshino, K. Ishiwata, and N. Makiguchi, *J. Ferment. Bioengineer.* **72**, 147–151 (1991).
13. S. Iwamori, T. Oikawa, K. Ishiwata, and N. Makiguchi, *Biotechnol. Appl. Biochem.* **16**, 77–85 (1992).
14. H. Suzuki, T. Katayama, K. Yamamoto, and H. Kumagai, *Biosci. Biotech. Biochem.* **59**, 2028–2032 (1995).

V

VACCINE TECHNOLOGY

MICHAEL L. DEKLEVA
Merck & Co., Inc.
West Point, Pennsylvania

KEY WORDS

Adjuvant
Immunization
Vaccine
Validation

OUTLINE

INTRODUCTION

Disease Applications for Vaccines

Vaccines are suspensions of live, killed, or attenuated microorganisms such as bacteria or viruses or components of them. They are designed for human or veterinary use to prime the immune system to produce antibodies or cellular responses that ward off future challenge by pathogenic organisms or their toxins (1–4). *Antibodies* are a group of related proteins that bind specifically to the molecules that induce their formation, *antigens*, and in doing so help facilitate their destruction and clearance from the body. Some molecules, such as carbohydrates, lipids, and small molecules, are themselves poorly immunogenic and must be coupled to proteins to more effectively stimulate an immune response. Although numerous examples exist of vaccines against diseases of bacterial origin (*Haemophilus influenzae, Corynebacteria diphtheria, Streptococcus pneumoniae*, etc.), they are particularly useful prophylactic treatments for viral pathogens, against which antibiotics are ineffective (measles, mumps, rubella, hepatitis B, hepatitis A, etc.). In addition, vaccines have been produced to effectively counteract bacterial toxins such as diphtheria (*Corynebacterium diptheria*) (3) and tetanus (*Clostridium tetanii*) (3) by producing a benign form of the toxin, referred to as a toxoid, by heat or chemical treatment (phenol, formaldehyde, etc.). Vaccines may also offer protection against certain cancers. Hepatocellular carcinoma resulting from chronic hepatitis B infection can be effectively eliminated by preventing the initial infection through vaccination (5–8). Likewise, evidence suggests that certain stomach cancers may be preventable through vaccination with *Helicobacter pylori* preparations (9), or some cervical cancers by vaccination with human papillomavirus vaccines (10,11).

Overview of Preparation

The success of vaccines is most significantly influenced by five factors: (*1*) the nature and dose of the immunogen, (*2*) the adjuvant or excipients used in the formulation, (*3*) the immunization schedule, (*4*) the route of administration, and (*5*) the immune status of the person being vaccinated, influenced by age and disease (2). The vaccine preparations themselves may consist of killed whole bacterial cells or live attenuated or inactivated viral preparations. Alternatively, subunit vaccines are those that rely on recognition of specific components of pathogens by the immune system, instead of the whole microorganism, to elicit a response. These components may include capsular polysaccharides (e.g., vaccines against *Haemophilus influenzae* type b or *Streptococcus pneumoniae*) (12–14), viral coat proteins (15), or other structural components of the organism that can be recognized and responded to by the immune system. Also, recombinant DNA technology provides a means to express subunits of pathogenic bacteria or viruses. For example, a highly effective vaccine against hepatitis B infection has been made by expressing the hepatitis B viral coat protein surface antigen in yeast (15), thereby eliminating the need to the handle infectious blood from which the vaccine was originally isolated. Modern techniques for cell culture, fermentation, isolation, and formulation allow the manufacturer of vaccines of unparalleled purity, specificity, and safety. The high levels of purity offer the advantage of low reactigenicity, with reduced pain, swelling, redness at the site of injection, or systemic effects such as fever. Higher purity presents new challenges, however, in that the highly pure components are typically poorer immunogens than their whole cell counterparts (16–19), driving active research into other agents that can be coadministered (adjuvants) to boost their effectiveness.

Routes of Administration

Vaccines are typically administered intramuscularly or subcutaneously but may be designed, as appropriate, for administration by other routes (e.g., oral polio vaccine). Active research programs exist to study alternate methods for vaccine administration, because most infections are acquired naturally by mucosal routes, either orally, nasally, or genitally, and parenteral vaccines are often poor inducers of mucosal immunity (2). Mucosal vaccines, however, must be designed to overcome the normal rapid clearance, and oral vaccine candidates must be designed to survive the enzymes and pH extremes associated with the gastrointestinal tract (2).

Immunization Strategies

Vaccination strategies include inoculation of individuals who are, because of lifestyle or circumstance, susceptible to a particular pathogen. Alternatively, widespread immunization of a demographic group may be desirable to protect entire populations. This is accomplished in the United States and many other countries through compulsory vaccination of schoolchildren, health care workers, and those serving in the armed forces. This approach has been particularly effective in eliminating the threat of endemic disease in children, such as polio and measles (20,21).

Schedules for vaccine administration vary from product to product and often from one country to the next (4,22). For full and lasting immunity to be achieved, repeated exposure to an antigen may be required (23–25). Therefore, most vaccines are administered through regimens of two to four doses. Even after serum antibody titers drop below detectable levels, immune memory results in a rapid and substantial *anamnestic* response, whereby antibody titers rise rapidly upon revaccination or subsequent challenge with the targeted disease entity (23,25). Table 1 lists a number of currently available vaccines and their administration schedules (2).

It is now widely accepted that the prophylactic prevention of disease through vaccination is an effective and cost-effective means to disease management (4). Smallpox has been eradicated through rigorous worldwide public vaccine programs, and many other diseases have been brought under control, fueling tremendous public optimism in our ability to alleviate morbidity and mortality associated with the many bacterial, viral, and parasitic infections that continue to plague us (20,21). Further applications toward the prevention of reproductive fertility and cancer treatment are also being explored. It is critical at this juncture that continued research into vaccine development and production provide the means to further reduce vaccine costs and make compliance to vaccine schedules easier through the advent of combination vaccines (to reduce required injections) or alternate delivery systems.

EARLY VACCINE HISTORY

Vaccine technology enjoys a rich history with brilliant contributors (26). The first successful vaccine was demonstrated by Jenner in the late 1700s. Jenner had observed that milk maids, who often harbored lesions from benign cowpox infections, were immune to smallpox. He found that a person inoculated with the fluid from a cowpox lesion would be protected from disease if later challenged with fluid from a person with smallpox. The word vaccine is derived from the Latin word *vaccinus*, which means "of or pertaining to cows."

Koch was responsible for advancing the concept that each infectious disease is caused by a single, unique pathogenic species, and whose work with anthrax, cholera, and tuberculosis shook the medical community of his time. Likewise, Pasteur's flamboyant style and flare for publicity popularized the notion, through work with cholera, anthrax, and rabies, that many diseases of animals and people could be prevented or cured. And so early successes with smallpox, cholera, anthrax, and rabies fueled an enthusiastic search for cures and preventative measures for such scourges as diphtheria, tetanus, and polio. Early vaccines for the first time offered hope that the infectious diseases responsible for widespread morbidity and mortality could be prevented.

METHODS FOR VACCINE MANUFACTURE

The earliest vaccine preparations were isolated from tissue fluids obtained from infected people or animals, beginning with Jenner's cowpox vaccine against human smallpox. Through the middle of the 20th century all viral vaccines originated from live or killed whole virus, derived from infected organs or tissues of animals or embryonated hens' eggs (15). Modern methods required for the preparation of vaccines are as varied as the vaccines themselves, tailored to the unique chemical and physical properties of the antigen being cultured or isolated (27–29). Because the course of vaccine treatment is confined to just a few highly potent doses, administered in low volumes (typically a milliliter or less), manufacturing scales for vaccines are small relative to many therapeutic products requiring higher doses or longer courses of treatment.

Propagation of Live, Attenuated Viral Vaccines

The basis for live, attenuated viral vaccines is that by repeatedly subculturing the virus through appropriate host cell lines (e.g., chicken embryo or monkey kidney cell), its virulence can be sufficiently attenuated to avoid a fulminant infection upon inoculation, while still preserving the antigens required to elicit protective immunity. Because viruses are obligate intracellular parasites, they require viable eukaryotic cells for propagation (30–32). The following paragraphs will provide further details on cell culture lines and cell culture techniques that are used in the preparation of vaccines.

Primary Cell Lines. Vaccines have been manufactured by culturing viruses in a variety of animal organs or tissues in vivo or primary cells in vitro. These include embryonated chicken or duck eggs and primary cells obtained from a variety of different animals. Primary cells refer to cells obtained directly from tissue sources and transferred

Table 1. Pharmacokinetic Parameters of Vaccines

Vaccine	Recommended age for first vaccination[a]	Interval for vaccination schedule (mo)	Route	Onset (wk)	Duration (y)	FDA-defined risk factors in pregnancy[b]
			Short-acting vaccines (<2 y)			
Cholera	5 y	0–0.5; every 6 mo	SC or IM (PO) available	2	0.5	C
Influenza	6 mo	Annually	SC	2	1	C
Japanese encephalitis	3 y	0–0.25–0.5; 4 y interval	SC	8	0.5–1	C
Pertussis, accellular[c]	6 wk	0–1–2	IM	4	>1.5	ND
			Moderately short-acting vaccines (2–5 y)			
Haemophillus influenzae b	6 wk	0–2–4–10	IM	16[d]	>4	C
Meningococcus	2–6 y	Single dose	SC	2	3	C
Polio, inactivated	6 wk	0–1–6	SC	6–24	5	C
Typhoid	3 y	4 doses over 1 wk	PO, fasting	1	5	C
			Moderately long-acting vaccines (5–10)			
Diphtheria	6 wk	0–1–2–8; 6 y	Deep SC or IM	8	>10	C
Hepatitis A	2 y	0–1–7	IM	3	5–10	ND
Hepatitis B	2 mo	0–1–6	IM	4–8	5–7	C
Pneumococcus	2 y	Single dose	SC or IM	2–3	5–8	C/X
Varicella	12 mo	0–1	SC	6	5–10	C (D in first trimester)
			Long-acting vaccines (>10 y)			
Measles	9 mo	Single dose; 6 y interval	SC	ND	>21	X
Mumps	15 mo	Single dose; 6 y interval	SC	4	15	X
Polio, live	2 mo	0–2–16	PO	1	>12	C/X
Rubella	12 mo	Single dose	SC	2–4	Lifelong	X
Smallpox	12 mo	Single dose	ID with bifurcated needle	ND	30	X
Tetanus	6 wk	0–1–2–8; 6 y interval	SC or IM	4	12	C
Yellow fever	9 mo	10 y interval	SC	2	30–35	D
			Duration not clear			
Rabies	12 mo	0–0.25–1	ID or IM; postexposure IM	1–2	ND	C
Tuberculosis (BCG)	3 mo	Single dose	ID	2–3	ND	C

Source: Ref. 2.
Note: BCG = bacille Calmette-Guerin; FDA = U.S. Food and Drug Administration; ID = intradermal; IM = intramuscular; mo = month(s); ND = no data available; PO = oral; SC = subcutaneous; wk = week(s); y = year(s).
[a]Younger children may receive some of the vaccines.
[b]C = animal studies have revealed adverse effects on the fetus (use only if potential benefit justifies the potential risk to the fetus); D = positive evidence of human fetal risk; X = human beings have demonstrated fetal abnormalities or evidence of fetal risk, based on human experience.
[c]Duration is still under investigation.
[d]Also dependent on age.

for the first time into an in vitro growth environment (32). Examples include chick fibroblasts cultured from chicken embryos for the production of measles and mumps vaccines. Cell lines for viral propagation grow either as attachment-dependent or, occasionally, suspension cultures (32).

To produce measles and mumps vaccines, those viruses are grown and attenuated by passage through cultures of chicken embryo cells. Chicken embryos are harvested from eggs, obtained from special isolated flocks, and then minced and treated with disaggregating enzymes such as trypsin, collagenase, hyaluronidase, and pronase. The chick fibroblast cells released by this process are attachment dependent, requiring solid surfaces for growth. The successful commercial manufacture of viral vaccines in attachment-dependent cell systems rely on the establishment and maintenance of healthy cell monolayers. Selection of the appropriate growth and maintenance media are critical, as are careful attention to nutrient depletion, waste accumulation, and changes in pH over time (33).

Diploid and Continuous Cell Lines. Serially passaged diploid strains of cultured human cells were first described in the 1960s by researchers at the Wistar Institute (34). The concept of using human diploid cell lines for vaccine preparation was hotly debated in the 1960s, for fear that

cancer-causing DNA or human viral agents might be unknowingly coadministered with the vaccine. Extensive karyological (i.e., chromosomal) characterization and thorough searches for viral contaminants were necessary to ensure that cultures were free of exogenous infectious agents before the first diploid cell product, poliomyelitis vaccine, was licensed in the United States (11). Human diploid cell lines (WI-38 or MRC-5) have since been used to produce a number of licensed vaccine products against poliovirus, adenovirus types 4 and 7, rubella (German measles) virus, rubeola (measles) virus, rabies, hepatitis A, and varicella virus (chickenpox) (30–32). A distinguishing feature of the human diploid cells such as WI-38 or MRC-5 is that they have a finite life span, reaching senescence after 40 to 60 population doublings (32). In contrast, continuous cell lines exhibit no such constraints and divide indefinitely. An example of a nonhuman continuous cell line that has been used successfully for the manufacturing of vaccines such as polio, rabies, and influenza is Vero (32,35). Vero cells are a continuous monkey kidney cell line (35).

Serially cultured or continuous cell lines are advantageous in that each new production batch is derived from a uniform master cell bank, characterized to be free of contaminating infectious agents, contaminating proteins, or nucleic acids. Proper maintenance of a master cell bank is critically important for the consistency of cell culture products. They are usually stored as small aliquots in liquid nitrogen and tested at intervals for chromosomal (karyology) and cell integrity to establish long-term stability (32).

Cell Culture Reactors. Cylindrically shaped roller bottles have been used successfully to establish monolayers of chick fibroblasts on their inner surfaces, which can be monitored microscopically through the clear plastic. The cell sheets are continuously bathed by a growth medium contained within the bottle during slow axial rotation on special roller racks. After the cell sheets are established, they are infected by introduction of the specific virus. After incubation, the cells and virus may be harvested (e.g., varicella). In the case of rubella, viral fluids may be harvested at approximately 2- to 3-day intervals, through 10 to 12 harvest cycles. Although millions of doses have been successfully manufactured using roller bottles, capacity is limited by the space that a large number of roller bottles requires and by the time-consuming and often labor-intensive manipulations for harvesting and pooling the viral fluids from them. Robotics may reduce the labor burden associated with the latter and have been successfully used by Merck for the manufacture of their chickenpox vaccine, Varivax, in roller bottles (33).

Multidisk reactors offer an alternative to roller bottles for attachment-dependent cell lines (36). They consist of 10-L stainless steel reaction vessels containing approximately 100 parallel titanium disks that slowly rotate through growth media and provide solid surfaces for cell attachment and monolayer formation. In addition to their improved capacity, the stainless steel reactors are jacketed for temperature control. Monolayers cannot be microscopically monitored, however. Another alternative for cultivating attachment-dependent cells was described in 1979 (37) and consists of stacked polystyrene unit trays (NUNC single-tray unit). Again, cell monolayers can be established, infected, and harvested more efficiently than with roller bottles (33). Other solid supports for attachment-dependent cell lines include ceramic or membrane (hollow fiber) solid supports, and microcarriers. Microcarriers may be composed of a variety of materials such as gelatin, polystyrene, cellulose, polyacrylamide, dextran, or glass, but usually are approximately 200 μm in diameter. They offer advantages of improved process control and scaleability because, like suspension cultures, they can be cultured in large-scale bioreactors. They are, however, susceptible to shear stress and physical deterioration and so must be suspended with low-shear agitators (38). For further reading on cell culture techniques for attachment-dependent cell lines and suspension cultures, please refer to *Large-Scale Mammalian Cell Culture Technology* (39,40).

Purified and Killed Whole Viral Vaccines

Examples of purified whole viral vaccines includes those to protect against polio, influenza, and hepatitis A viruses.

For the large-scale preparation of a killed polio vaccine, viral infected Vero cells have been grown on microcarriers in fermenters (35). Purification of the whole virus vaccine is as follows (36). The crude virus suspension is clarified by filtration, then concentrated 250-fold by hollow-fiber ultrafiltration (100,000 Da MW cutoff), which retains >99% of the virus. The virus is further purified by DEAE-Sephadex chromatography and then sterile filtered through a 0.22-μm filter before formaldehyde inactivation.

Traditionally, methods for influenza vaccine manufacture have always involved growth of the virus in embryonated chicken eggs. Recently, a manufacturing process for a formaldehyde-inactivated whole virus vaccine against influenza virus was developed using Vero cells in a microcarrier-based fermentation system (35). Supernatants from microcarrier cultures are clarified before benzonase treatment to digest Vero cell DNA. After formaldehyde inactivation, the product is concentrated by ultrafiltration (200,000 Da MW cutoff) and further clarified by protamine sulphate precipitation. Further purification is accomplished by continuous-flow zonal centrifugation of a 0 to 50% sucrose gradient, followed by ultrafiltration and sterile filtration.

Development of the hepatitis A vaccine began with demonstration that formaldehyde-inactivated virus extracted from the livers of infected marmosets would induce protective antibodies in susceptible marmosets (41). During the years that followed, various cell culture and purification schemes were developed for hepatitis A vaccine preparations (42,43). VAQTA® is one example (44,45), and is a highly purified, formaldehyde-inactivated hepatitis A viral vaccine. The virus is harvested from MRC-5 cells grown in Costar Cubes (46,47) after lysis with Triton X-100, followed by nuclease treatment of the lysate to digest nucleic acids. Concentration of the product is accomplished by adsorption on an anion exchange column and then PEG precipitation. Solvent extraction of the resuspended PEG pellet contributed significantly to the purification, as did the anion exchange and SEC steps that followed. The product is

subsequently formaldehyde inactivated and adsorbed onto aluminum hydroxide.

Purified Subunit Vaccines

The high levels of purity achieved by subunit vaccines have the potential to increase the specificity and significantly reduce the number or degree of adverse reactions associated with less-pure preparations. Their goal is to present the patients' immune systems with the specific feature, be it capsular polysaccharide in the case of *Haemophilus influenzae* or *Streptococcus pneumonia*, or viral coat proteins in the case of hepatitis B vaccines, conserved in all variants of a given species such that specific antibodies produced against it will neutralize the whole organism upon subsequent invasion. These components may be isolated from the organism itself, or expressed by a suitable host through recombinant DNA technology.

Manufacturing processes for subunit vaccines typically have six steps in common (1,48,49), which include (*1*) cell multiplication, either through fermentation or cell culture; (*2*) phase separation to remove cells, if the product is extracellular, or to harvest and wash cells if the product is intracellular; (*3*) product recovery; (*4*) purification; (*5*) polishing; and (*6*) adjuvanting and filling.

Cell Multiplication. The source culture for a subunit vaccine varies, depending on the nature of the product. Examples where components of the pathogens themselves are purified to produce vaccines include *Streptococcus pneumoniae, Haemophilus influenzae* type b, *Neisseria meningitidis, Clostridium tetanii, Corynebacteria diptheria*, or *Bordatella pertussis*. These bacteria are usually grown in batch culture on complex nutrients, such as yeast extract and soy peptone, with addition of specific growth factor supplements, such as hemin, that may be required for growth (50). As a general rule, defined or minimally complex media are sought to minimize purification demands. The antigens of interest, from the above examples, may be capsular polysaccharides (*Streptococcus pneumonia, Haemophilus influenzae*), or protein toxins that are chemically or thermally inactivated (tetanus, diphtheria, or pertussis). Capsular polysaccharide vaccines are sometimes adjuvanted by chemically conjugating them with protein components of other bacteria. Examples are Hib conjugate vaccines in which the capsular polysaccharide of *Haemophilus influenzae* type b is chemically coupled to the outer membrane protein complex of *Neisseria meningitidis*, diptheria toxoid, or tetanus toxoid (12,14). And although the current pneumococcal vaccine is formulated using 23 types of *Streptococcus pneumoniae* (1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19F, 19A, 20, 22F, 23F, and 33F), new pneumococcal polysaccharide–protein conjugate vaccines are being developed to improve T-cell-dependent antibody responses, improve long-term immunologic memory, and otherwise enhance immune responses in young children, the elderly, and other groups at risk for pneumococcal infection (13).

Alternatively, proteins representing key antigenic determinants from a target pathogenic microorganism can be recombinantly expressed in bacterial, yeast, or mammalian cell cultures. Important factors influencing the choice of a host expression system are freedom from adventitious agents (51,52), the scientific and organizational experience with a given system, efficiency of product expression, the need for specific glycosylation patterns for full efficacy, or the desirable purification characteristics. Proteins expressed in *Escherichia coli* are not glycosylated and accumulate intracellularly in inclusion bodies that require denaturation, then refolding to return the protein to its native conformation. Proteins expressed in yeast may be either excreted or accumulated intracellularly as viruslike particles, depending on the construct. Mammalian host systems express proteins extracellularly (17). Yeast and mammalian cell host systems, because they are eukaryotic, are often considered when consistent patterns of glycosylation are important (53). Expression systems are also chosen for their scaleability, so that conditions studied in the laboratory can be more easily applied to full-scale manufacturing. Growth media vary depending on the specific host system and on the recombinant expression system. For example, stable maintenance of a plasmid-encoded protein in yeast or bacteria may depend on a culture medium deficient in a particular amino acid, where the plasmid compensates or a specific host auxotrophy (54,55). Alternatively, the coding region for a particular protein product may be regulated by an inducible promoter, such that product expression is triggered by addition of the appropriate chemical or nutritional component (54,55).

Frozen master cell banks (-70 °C), consisting of clonal isolates tested to ensure purity and stability of expression, are subpassaged to generate a working cell bank. In this way, batch-to-batch consistency and purity can be ensured.

Stirred tank reactors, in volumes of hundreds to thousands of liters, are commonly used to cultivate bacterial or recombinant cultures. The bioreactors are jacketed for control of temperature and sparged to provide appropriate levels of oxygenation. Monitoring may be on-line or off-line, by periodic sampling and analysis, for such parameters as biomass, carbon-source (e.g., glucose) consumption, pH, temperature, dissolved oxygen, etc. Valuable information can also be attained on culture physiology by monitoring head space gases for carbon dioxide evolution or oxygen consumption or volatile organic by-products (e.g., ethanol, in the case of yeast cultures). In addition to proper maintenance of master and working cell banks, equipment maintenance and careful attention to fermentation raw materials are essential to ensure reproducible fermentation results (50,56).

Phase Separation. After the final production phase of cell culture, it is usually necessary to separate cells from the remainder of the fermentation broth. This may be accomplished either by continuous centrifugation or microfiltration. If the product is intracellular, the removal of residual culture media components or antifoaming agents from the whole cells is a useful method to reduce the number of contaminants that must be removed by subsequent downstream steps. Regardless of whether products are intracellular or extracellular, care must be taken to avoid cell disruption by excessive shear or osmotic stress to prevent

cells from spilling their contents. In the case of intracellular product, this would result in yield losses, and for extracellular products, it would result in the cells' intracellular contents adding to the total number of impurities that must be removed by subsequent downstream steps.

Product Recovery. Figure 1 depicts the typical flow of downstream processing from recovery through purification. Intracellular products are liberated by cell disruption, which is typically accomplished using either bead mills or continuous homogenizers. Bead mills rely on mechanical shear, whereas continuous homogenizers create fluid shear by forcing the cell slurry through a small opening at high pressure (50). Bacterial inclusion bodies require solubilization with chaotropes such as urea or guanidine HCl, followed by renaturation in appropriate buffers (1,27). Virus-like particles in yeast may be tightly associated with intracellular membranes, such as endoplasmic reticulum, requiring detergents for liberation (54). Removal of cell debris is often accomplished by filtration or centrifugation. Where concentration of the desired product is necessary, membrane ultrafiltration is common.

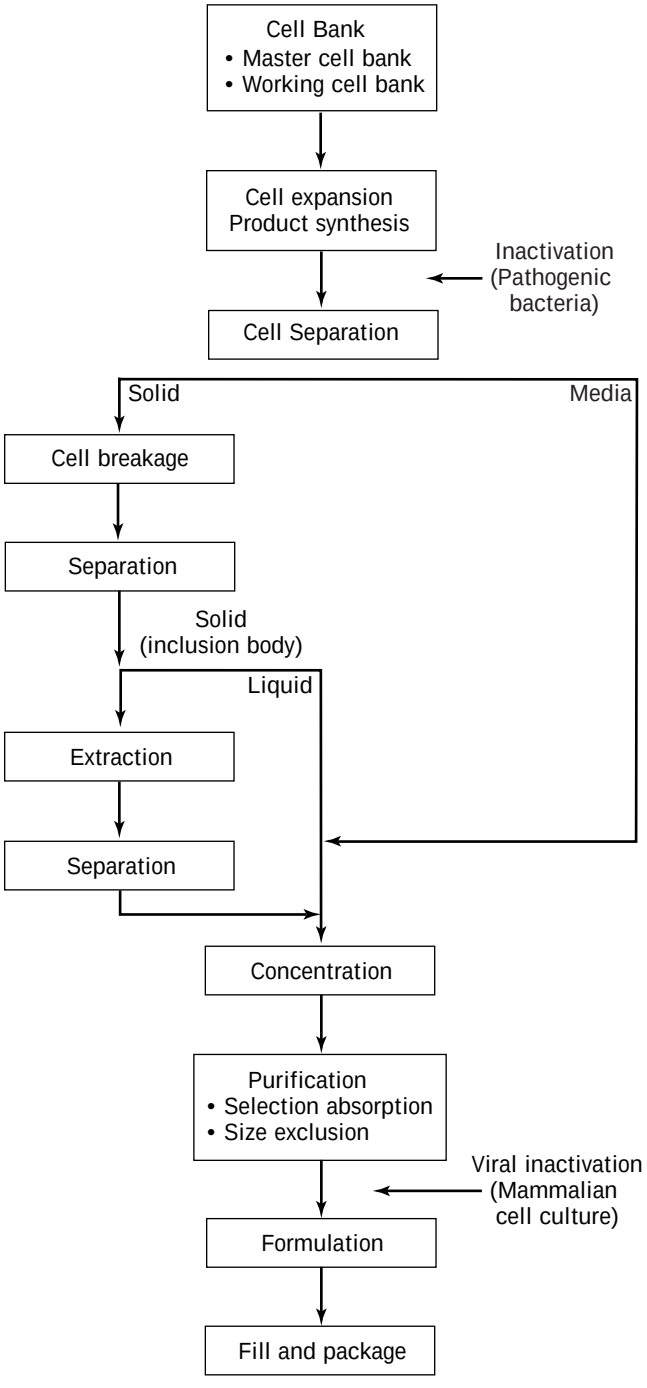


Figure 1. Generic process flow: microbial fermentation or cell culture products.

Purification. Numerous choices exist to accomplish purification, and selection of the most appropriate technologies must be grounded in an understanding of the physical, chemical, and biochemical properties of the starting material and key impurities (1,48,49). Some of the options are outlined in Table 2 (27–29,57). Because yield losses occur at each purification step, the number of steps should be kept to a minimum. This will improve overall recoveries and reduce costs by keeping resource requirements, including raw materials, testing costs, and labor, to a minimum (27–29,48,49,58).

Manufacture of Hepatitis B (Subunit) Vaccines: The original hepatitis B vaccine (HEPTAVAX B®), licensed in 1981, was a subunit vaccine isolated from the blood of hepatitis B infected donors (15). Although derived from infected blood, the final product consisted of a highly purified preparation of 22-nm particles containing HBV surface antigen, subunits of the HBV virus. These particles are expressed in copious amounts by patients suffering from acute infection or in hepatitis B carriers. Although these

Table 2. Downstream Operations for Large-Scale Recovery and Purification of Vaccines

Operation	Physicochemical Property
Cell breakage	
Bead Milling	Liquid/solid shear
Homogenization	Pressure gradient
Cell separation and product recovery	
Filtration	Particle size
Microfiltration	Particle size
Centrifugation	Sedimentation velocity
Precipitation	Solubility
Two-phase extraction	Partition coefficient
Purification	
Adsorption	Hydrogen bonds, polarity, dipole moments, etc.
Ion exchange	Charge
Affinity chromatography	Biological affinity
Hydrophobic interaction	Hydrophobicity
Ultrafiltration	Molecular size
Gel filtration	Molecular size
Reverse-phase liquid chromatography	Hydrophobic and hydrophilic interactions

subunits themselves are noninfectious, they are effective in stimulating a strong immune response and provide lasting protection against hepatitis B infection.

The isolation process consisted of (*1*) defibrination of the plasma with calcium; (*2*) ammonium sulfate precipitation, rate zonal centrifugation (sucrose gradient), and isopycnic banding (sodium bromide); (*3*) pepsin digestion; (*4*) urea denaturation/renaturation; (*5*) SEC; and (*6*) formaldehyde treatment. The process not only resulted in greater than 99% purity, but also ensured safety from extraneous infectious viruses through the sequence of viral inactivation steps—pepsin digestion, urea denaturation, and formaldehyde treatment (15).

Problems associated with the manufacture of plasma-derived hepatitis B vaccines included limitations in the supply of source plasma and a manufacturing process that was state-of-the-art at that time, but still laborious and requiring a year or more to complete each batch. In addition, although HEPTAVAX B® had an excellent safety and efficacy profile in millions of people around the world (59), it was not quickly accepted because of growing concerns over AIDS that might be present in the same infectious plasma that served as the source for the vaccine (59). These concerns led to strong public and government pressure to develop a recombinant hepatitis B vaccine for human use, of which RECOMBIVAX HB® was the first. This recombinant hepatitis B vaccine is prepared for human use by expression of the hepatitis B surface antigen (HBsAg) in a recombinant yeast strain that is batch cultured in stirred-tank fermenters. Cells are separated from the fermentation broth, concentrated, and washed with buffer (to remove media components and antifoam) by microfiltration. Cells are lysed to release the HBsAg, which is accumulated intracellularly by the yeast cells, by passage through a high-pressure homogenizer. Phenylmethylsulfonyl fluoride, a protease inhibitor, is added immediately before lysis to prevent proteolytic degradation. Triton X-100 detergent is added to liberate the HBsAg from tightly associated membrane components, and the product is recovered using a microfiltration system. Cell debris and unbroken cells are retained by the filter and eventually discarded, and the HBsAg is recovered in the filtrate. The product stream is ultrafiltered against a 100-kDa molecular weight hollow-fiber filter cartridge to clear small molecular weight contaminants and concentrate the product. Residual detergent is removed by recirculation through polystyrene beads (XAD-4) followed by adsorption of the antigen on colloidal silica (Aerosil). The HBsAg is eluted with warm borate buffer, which interacts with the silica surface to facilitate its release. Final polishing is accomplished by hydrophobic interaction chromatography, using butyl agarose, followed by treatment with thiocyanate to promote disulfide cross-linking of the protein components of the HBsAg particle. Afterward, the product is sterile filtered through a 0.22-mm membrane. The product is adjuvanted by coprecipitation with aluminum hydroxide (60).

Process Ruggedness. Once promising vaccine candidates are identified through research, and their safety and efficacy is proved through clinical studies, strong medical and financial drivers motivate a rapid introduction of the product into manufacturing. Once the manufacturing facility is constructed and qualified, and the final process is established, a consistency series of three to five lots serves to confirm that product profiles are unchanged through the final process and facility scale-up. Data from these full-scale runs support product licensure. Therefore, at the time of licensure, a considerable body of information will have been amassed at laboratory and pilot scale, but full-scale manufacturing data are limited. Because of the speed through which final process scale-up and manufacturing introduction are accomplished, abundant opportunities generally exist to improve manufacturing efficiencies and overall productivity after product introduction. Such changes must be approved by regulatory agencies before incorporation, supported by data confirming that the product's safety, purity, or potency are not diminished.

It is important to be aware of common pitfalls in new process introductions that might later interfere with a consistent product output. Process ruggedness is a key consideration. For example, a single lot of a key reagent, such as a fermentation raw material or chromatographic resin, may suffice to prepare numerous test lots at laboratory scale, and there may be a strong temptation to show process consistency by using a single lot of a key reagent during the manufacturing demonstration. It is absolutely crucial, however, to identify those critical raw materials early in process development and challenge the process with different vendor lots and from multiple vendors, if possible, to better understand the impact that those changes will have later on in manufacturing when much larger quantities will be consumed and when lot-to-lot variations will stress process consistency. Critical process parameters must also be identified early and the process challenged around them to better understand failure limits. In that way, unplanned process deviations associated with equipment failures or operator error can be more easily addressed, and rapid decisions can be made on the quality of a particular lot. Also, the more thorough the understanding of which steps are truly critical to achieving product quality goals, the less likely that a deviation at a noncritical step will result in product discard (52,61,62).

ADJUVANTS

The term *adjuvant* is derived from the Latin word *adjuvare,* which means to help. Adjuvants are used in conjunction with vaccines to aid in eliciting a rapid, high-level, and long lasting humoral or cellular immune response to the target antigen. Their significance is increasing because of the development of more highly purified subunit and synthetic vaccines that are more specific but inherently are less immunogenic. In addition, such vaccines tend to be very expensive to produce, so an effective adjuvant should also allow for the use of less antigen to stimulate the desired level of immunity. Ideally, an adjuvant should (*1*) enhance the ability of an antigen to stimulate a strong and persistent immune response in all age groups targeted for vaccination, whether newborns or older adults; (*2*) allow lower or fewer doses of antigen to be used; (*3*) be biode-

gradable; (*4*) not itself be immunogenic; (*5*) not interact with the antigen to cause its deterioration; and (*6*) be nontoxic and without side effects (63). An important issue for the development of vaccines today for human use is striking a balance between stimulation of a strong immune response and toxicity or other undesirable side effects. Those side effects may be localized or systemic. Local reactions, such as pain, swelling, and redness, can be a result of depot formation (localization of an antigen), which is also a primary mechanism of action of some adjuvants, such as mineral compounds, oil–water emulsions, liposomes, or biodegradable polymeric microspheres. These geltype or emulsion-based-adjuvants increase the half-life of the antigens associated with them, which might otherwise be cleared too rapidly from the body for an effective immune response to be mounted. Alternatively, systemic reactions may develop, depending on the type and amount of cytokines that might be elicited after stimulation of the immune system.

Adjuvant History

In 1916, LeMoignic and Pinoy first described the use of mineral oil and lanolin as an adjuvant emulsion in a killed *Salmonella* vaccine (64). In the 1920s, Ramon showed that immune responses to tetanus and diphtheria toxoids were enhanced when used in conjunction with other compounds such as agar, tapioca, lecithin, starch oil, saponin, or even breadcrumbs (65,66). Later, in 1926, Glenny described the use of aluminum compounds as adjuvants (19,67). In the early 1930s and 1940s, Freund developed his water-in-oil emulsion containing killed mycobacteria that became known as Freund's complete adjuvant. The reactigenicity was reduced significantly by elimination of the mycobacterial component. This modified formulation became known as Freund's incomplete adjuvant and was successfully used in several vaccine formulations (19). In 1956, Johnson et al. reported that lipolysaccharide endotoxins (LPS) from Gram-negative bacteria had adjuvant activity (19,68). This work provided the foundation for further studies into the use of detoxified LPS, or subcomponents such as lipid A (19,69). In 1974, Lederer et al. identified muramyl dipeptide (MDP) as the smallest adjuvant-active component of mycobacteria (19,70). Since then, more than 300 compounds have been investigated in search of those retaining the potent adjuvant activity but without unwanted side effects such as pain, redness, swelling, or fever.

Types of Adjuvants and Mechanisms of Action

Adjuvants are a chemically diverse group of substances that can be categorized based on chemical properties and physical or chemical properties (19). Adjuvants have also been characterized by mechanism of action (19,63,71), including (*1*) causing depot formation at the site of injection (e.g., mineral compounds, oil-based adjuvants, liposomes [72], oil adjuvants, biodegradable polymeric microspheres of >10 mm, etc.); (*2*) delivery vehicles for antigens to target specific immune-competent cells such as macrophages or dendritic cells (e.g., liposomes, oil adjuvants, biodegradable polymer microspheres <10 mm, nonionic block polymer surfactants); or (*3*) acting as immunostimulators (e.g., Freund's complete adjuvant, muramyl depeptide, lipopolysaccharide, lipid A, monophosphoryl lipid A, pertussis toxin, etc.) to facilitate production of relevant lymphokines. Antigens can also be modified to enhance immunogenicity (adjuvant effect) through (*1*) polymerization of the protein constituents using aldehydes (tetanus and diphtheria toxoids); (*2*) cationization of proteins (diphtheria toxoid); or (*3*) chemical conjugation of polysaccharide antigens to carrier proteins (e.g., *Haemophilus influenzae* vaccines), which may provide T-cell epitopes to the coupled antigen or simply increase its effective molecular weight.

The only adjuvants used routinely in human vaccines are aluminum salts (16,19), either aluminum hydroxide or aluminum phosphate (19). These have a long history of safety and efficacy, since it was discovered in the 1930s that suspensions of alum-precipitated tetanus and diphtheria toxoids had a much higher immunogenicity than the soluble toxoids, and their use became common with diphtheria-tetanus-pertussis (DTP) vaccines. Billions of doses have been safely administered. Despite their reputation for safety, alum adjuvants do have some drawbacks. They have occasionally been associated with subcutaneous nodules or allergic reactions. They cannot be frozen or lyophilized without collapse of the gel structure. Aluminum-based adjuvants are primarily effective in eliciting a humoral as opposed to cell-mediated response (16). This is a significant limitation when considering their use with vaccines targeting a T-cell response, such as against intracellular viral or parasitic agents such as AIDS or malaria, respectively (19,63,71).

Alum has several forms, including aluminum hydroxide and oxyhydroxides, aluminum phosphate gels (73–76) and precipitated alum [$KAl(SO_4)_2 . 12H_2O$]. Preparation of vaccines or toxoids with aluminum adjuvants is either by coprecipitation of the antigen with the aluminum-based compound ($KAI(SO_4)_2 . 12H_2O$ added to the product (76) or adsorption onto preformed aluminum compounds. Currently, most vaccines are manufactured by the adsorption methods. Their mechanism of adjuvant activity includes (*1*) depot formation at the site of injection, from which antigen is slowly released; (*2*) stimulation of immune-competent cells through complement activation and induction of eosinophils and macrophages; and (*3*) efficient uptake of aluminum-adsorbed antigens due to their particulate nature and optimal size (<10 mm) (21).

Research continues to contribute to a better understanding of adjuvant molecular mechanisms, the biochemistry of the immune response to various diseases, and how to selectively invoke classes of cytokines or immune-competent cells to best suit an antigen's specific purpose (i.e., humoral or cellular response) (19,71). In addition, controlled release of vaccines through the use of biodegradable microspheres may provide a mechanism for single-dose vaccines by either slow or intermittent antigen release. This is particularly important in some developing countries where contact between health care workers and the vaccine recipients for booster doses may be difficult to achieve (77,78).

REGULATORY CONSIDERATIONS FOR VACCINES

Published Federal Guidelines

Vaccine manufacture in the United States is regulated by the arm of the Food and Drug Administration (FDA) known as the Center for Biologics Evaluation and Research (CBER). In addition, the FDA's *Code of Federal Regulations* is a codification of the general and permanant rules published in the *Federal Register* by the executive departments and agencies of the federal government. Title 21, parts 600 to 699, was written with certain vaccines in mind (pertussis, typhoid, anthrax, cholera, polio, measles, mumps, and rubella), but it has evolved to include guidance for the manufacture and testing of all vaccines and many other biological products. Included are sections on personnel responsibilities and training, facilities, equipment, and animal care, record maintenance and retention, labeling, packaging, and shipping.

Parts 210 and 211 of Title 21 provide manufacturers of human drugs, veterinary drugs, and medical devices with guidance regarding current good manufacturing practices (cGMPs). The cGMPs are comprehensive and encompass manufacturing methods, facilities, controls, packaging, storage, and installations for drug operations. Currently, licensure of a vaccine product entails licensure of both the product and procedures by which it was made and tested, packaged, labeled, and distributed (product license) as well as the facility and equipment in which it was made (establishment license). The latter includes air handling systems; water and air supply and testing; personnel, product, and equipment flows; materials of construction for equipment; interior and exterior materials of construction; and cleaning and decontamination procedures. Government inspections are held before product approval and at a minimum of every 2 years thereafter to ensure compliance with existing regulations. Minor adverse inspection observations are presented by the inspectors, at the conclusion of an inspection, to the responsible head of the facility, on form 483. Other mechanisms exist for the FDA to prevent continued distribution or manufacture of products if more severe compliance violations are observed. In any event, the observations from FDA inspections are published and provide manufacturers with information on emerging trends within the FDA (79,80).

Process Validation

Regulations governing vaccine manufacture have evolved and continue to do so. In recent years, that evolution has led to an ever-increasing emphasis on validation, which is a formal method to demonstrate that a given process can accomplish its intended goals consistently and reproducibility. The need for validation of sterilization and cleaning procedures has gained widespread acceptance to ensure that equipment in direct contact with manufactured products is free from extraneous chemical or microbial contamination and from product-to-product or batch-to-batch carryover. In addition, testing procedures used to assess potency, and purity must also be validated.

Increasing emphasis is being placed by the FDA on process validation (81). The goals of process validation are to ensure the consistency of consecutive batches of vaccine using a particular process (82). Each process consists of a series of steps, referred to as unit operations, executed in sequence by a prescribed "recipe." Each of those operations has a specific goal. In the case of fermentation or cell culture, the goal is to increase cell density and expression of the particular antigen of interest. Downstream steps may include release of intracellular antigens through cell disruption, filtration or centrifugation steps, salt or alcohol precipitations, filtration steps to remove small molecular weight contaminants or exchange buffers, or chromatographies to remove contaminants such as unwanted proteins or nucleic acids. Additionally, chemical steps may be required achieve an optimal conformation for enhanced immunogenicity or adjuvant binding. Regardless of the steps, each serves a particular purpose to improve the purity or maintain or enhance the potency of the particular antigen of interest.

To accomplish process validation, a formal document or protocol is prepared delineating those steps critical to the overall success of the process, the intended goals of each of those steps, and the analytical methods required to monitor the consistency of the process to achieve its intended goal. Expectations are established early through laboratory- and pilot-scale runs, and then confirmed upon scale-up to the final manufacturing scale. Those equipment settings necessary to achieve a consistent outcome are referred to as critical process parameters. Analytical measurements of process performance are referred to as critical quality attributes. Prerequisites to process validation include appropriate controls over raw materials; qualification and maintenance of utilities and equipment; calibration and preventative maintenance of scales, gauges, and other mechanical and electronic monitoring devices; comprehensive quality control programs for analytical testing and established procedures for batch review and release; unambiguous manufacturing instructions approved by manufacturing and quality control personnel; and programs for comprehensive training. In addition, validation of equipment cleaning and sterilization are necessary.

Filter validation is another emerging arm of process validation. Goals include insurance that filter membranes, housings, and support components do not leach objectionable chemicals into the product stream during manufacturing and that the filters are sufficiently rugged to stand up to process conditions without integrity failure. Critical process filters are tested before and after use for integrity to provide greater assurance that they performed properly toward their intended goal.

Included among the overall benefits of process validation are the greater understanding of process variables, allowing the manufacturer to more reliably supply customers with the highest quality product with fewer discarded batches.

FUTURE DIRECTIONS FOR VACCINE TECHNOLOGY

Research into vaccine technology continues on many fronts to provide safe, efficacious, and affordable vaccine prod-

ucts. Many at-risk individuals living in Third World and endemic areas do not have ready access to health care, so compliance to the multidose regimens required to achieve full protection may be difficult. Also, cold chains may not exist for those vaccine preparations that must be kept refrigerated to maintain proper levels of potency. Improvements in the thermal stability of vaccine preparations is advantageous (83).

One approach to improve coverage that is being adopted by many vaccine manufacturers is to combine antigens into single preparation so that immunity to two or more diseases can be achieved with a single injection (84). Time-release polymeric preparations are also being explored even for monovalent preparations to achieve repeated antigenic stimulation of the immune system without the need for multiple injections. Other research initiatives seek to eliminate the need for needles for delivery of vaccines through the use of oral vaccines (85,86), aerosolizers for inhalation (87), or needleless devices that use compressed gases at high pressure to deliver transdermal doses of vaccine preparations (88).

A fairly recent development in vaccine technology may offer several advantages over more conventional preparations in terms of manufacturing ease, potency, and flexibility in terms of delivery systems and disease targets. Observations in the early 1990s that plasmid DNA-encoding antigenic proteins can elicit an immune response when injected directly into animals has sparked intensive research and tremendous excitement as well as concern over long-term safety (89). With DNA vaccines, the coding region for a specific antigenic protein is incorporated into plasmids that are amplified in a host bacterial strain, such as *Escherichia coli*, then purified. The bacterial hosts for the coding plasmids are easily scaleable, and generic purification schemes may be possible because the product would always be plasmid DNA. After injection, the naked plasmid DNA yields protein expression in its native conformation, to which both humoral and cell-mediated antibody responses are mounted (90). The DNA constructs can be customized to coexpress both the coding region for the antigen as well as an immune modulator to enhance the immune response or increase its specificity (91,92). The tremendous amount of research in the past few years has led to progress in the development of vaccines against disease targets associated with bacterial, viral, and parasitic infections and provides hope for combating autoimmune and inflammatory diseases (93) as well as cancer.

BIBLIOGRAPHY

1. M. Canales, A. Enriquez, E. Ramos, D. Cabrera, H. Dandie, A. Soto, V. Falcon, M. Rodriguez, and J. De la Fuente, *Vaccine* **15**, 414–422 (1997).
2. S. Gizurarson, *Clin. Pharmacokinet.* **30**, 1–15 (1996).
3. E.H. Relyveld, B. Bizzini, and R.K. Gupta, *Vaccine* **16**, 1016–1023 (1998).
4. R.E. Spier, *Vaccine* **15**, 1163–1164 (1997).
5. J.E. Maynard, M.A. Kane, and S.C. Hadler, *Rev. Infect. Dis.* **2**, S574–S578 (1989).
6. C.N. Shapiro and H.S. Margolis, *Epidemiol. Rev.* **12**, 221–227 (1990).
7. P. Tiollais and M.-A. Buendia, *Sci. Am.*, 116–123 (April 1991).
8. A.J. Zuckerman, *Baillieres Clin. Gastroenterol.* **4**, 775–788 1990.
9. M. Marchetti, M. Rossi, V. Giannelli, M.M. Giuliani, M. Pizza, S. Censini, A. Covacci, P. Massari, C. Pagliaccia, R. Manetti, J.L. Telford, G. Douce, G. Dougan, R. Rappuoli, and P. Ghiara, *Vaccine* **16**, 33–37 (1998).
10. D. Bonn and J. Bradbury, *Lancet* **351**, 810 (1998).
11. R.W. Tindle, *Immunol. Res.* **16**, 387–400 (1997).
12. J. Eskola and H. Kayhty, *Vaccine* **16**, 1433–1438 (1998).
13. C-J. Lee, S.D. Banks, and J.P. Li, *Crit. Rev. Microbiol.* **18**, 89–114 (1991).
14. J.B. Robbins and R. Schneerson, *J. Infect. Dis.* **161**, 821–832 (1990).
15. M. Hilleman, in W.W. Ellis ed., *Hepatitis B Vaccines in Clinical Practice*, Dekker, New York, 1993, pp. 17–39.
16. C.R. Alving, B. Detrick, R.L. Richards, M.G. Lewis, A. Shafferman, and G.A. Eddy, *Ann. N.Y. Acad. Sci.* **690**, 265–275 (1993).
17. R. Elelman, in M.M. Leving, G.C. Woodrow, J.B. Kaper, and G.S. Cobon eds., *New Generation Vaccines*, Dekker, New York, 1997, pp. 173–192.
18. E.H. Relyveld, *Dev. Biol. Stand.* **65**, 131–136 (1986).
19. F.R. Vogel, *Ann. N.Y. Acad. Sci.* **754**, 153–160 (1995).
20. J.-L. Excler, *Vaccine* **16**(14/15), 1439–1443 (1998).
21. A.R. Hinman, *Vaccine* **16**(11/12), 1116–1121 (1998).
22. *Weekly Epidemiological Record: WHO* **70**, 221–228 (1995).
23. L.S. Marsano, D.J. West, I. Chan, T.M. Hesley, J. Cox, V. Hackworth, and R.N Greenberg, *Vaccine* **16**(6), 624–629 (1998).
24. *Morbidity and Mortality Weekly Report* **38**(13), 205–227 (1989).
25. D.J. West and G.B. Calandra, *Vaccine* **14**(11), 1019–1027 (1996).
26. P. De Kruif, *Microbe Hunters*, Harcourt Brace Jovanovich, New York, 1954.
27. S.V. Ho, in M.R. Ladisch, R.C. Willson, C.-d.C. Painton, S.E. Builder eds., *Protein Purification from Molecular Mechanisms to Large-Scale Processes*, American Chemical Society Press, 1989, pp. 14–34.
28. E.N. Lightfoot, in M.R. Ladisch, R.C. Willson, C.-d.C. Painton, S.E. Builder eds., *Protein Purification from Molecular Mechanisms to Large-Scale Processes*, American Chemical Society Press, 1989, pp. 35–51.
29. R.C. Willson and M.R. Ladisch, in M.R. Ladisch, R.C. Willson, C-d.C. Painton, S.E. Builder eds. *Protein Purification From Molecular Mechanisms to Large-Scale Processes*, AChS Press, 1989, pp. 1–13.
30. L. Hayflick, *Dev. Biol. Stand.* **70**, 11–26 (1989).
31. L. Hayflick, S.A. Plotkin, and R.E. Stevenson, *Dev. Biol. Stand.* **68**, 9–17 (1987).
32. M.E. Wiebe and L.H. May, in A.S. Lubiniecki ed., *Large-Scale Mammalian Cell Culture Technology*, Dekker, New York, 1990, pp. 147–160.
33. A.Y. Elliott, in A.S. Lubiniecki ed., *Large-Scale Mammalian Cell Culture Technology*, Dekker, New York, 1990, pp. 207–216.
34. L. Hayflick and P.S. Moorhead, *Exp. Cell Res* **25**, 585–621 (1961).
35. O. Kistner, P.N. Barrett, W. Mundt, M. Reiter, S. Schober-Bendixen, and F. Dorner, *Vaccine* **16**, 960–968 (1998).

36. A.L. Van Wenzel, G. Van Stienis, C. Hannik, and A. Cohen, *Dev. Biol. Stand.* **41**, 159–168 (1978).
37. O. Molin and C.G. Heden, *Progress in Immunological Standardization*, vol. 4, Karger, New York, 1969, pp. 106–110.
38. Skoda, A. Hormann, O. Spath, and A. Johansson, *Dev. Biol. Stand.* **42**, 121–126 (1979).
39. S. Reuveny, in A.S. Lubiniecki ed., *Large-Scale Mammalian Cell Culture Technology*, Dekker, New York, 1990, pp. 271–323.
40. A.S. Lubiniecki ed., *Large-Scale Mammalian Cell Culture Technology*, Dekker, New York, 1990.
41. P.J. Provost and M.R. Hilleman, *Proc. Soc. Exp. Biol. Med.* **159**, 201–203 (1978).
42. D.B. Strader and L.B. Seeff, *Drugs* **51**, 359–366 (1996).
43. A.J. Zuckerman, *Baillieres Clin. Gastroenterol.* **4**, 775–788 (1990).
44. A.J. Hagen, C.N. Oliver, and R.D. Sitrin, *Biotechnol. Prog.* **12**, 406–412 (1996).
45. Eur. Pat. 0302692B (1992), J.A. Lewis, E.A. Emini, and M.A. Armstrong.
46. Pat. Appl. WO 94/03589 (February 17, 1994), R.A. Aboud, J.G. Aunins, B.B. Buckland, P.A. DePhillips, A.J. Hagen, J.P. Hennessey Jr., B. Junker, J.A. Lewis, C.N. Oliver, C.J. Orella, and R.D. Sitrin.
47. M.E. Armstrong, P.A. Giesa, J.P. Davide, F. Redner, J.A. Waterbury, A.E. Rhoad, R.D. Keys, P.J. Provost, and J.A. Lewis, *J. Hepatol.* **18**, S20–S26 (1993).
48. J.A. Asenjo, in J.A. Asenjo ed., *Separation Processes in Biotechnology, Bioprocess Technology*, vol. 9, Marcel Dekker, New York, 1990, pp. 3–16.
49. G.J. Prokopakis and J.A. Asenjo, in J.A. Asenjo ed., *Separation Processes in Biotechnology, Bioprocess Technology*, vol. 9, Marcel Dekker New York, 1990, pp. 571–600.
50. P.F. Stanbury and A. Whitaker, *Principles of Fermentation Technology*, Pergamon Press, New York, 1983.
51. A.J. Darling and J.J. Spaltro, *BioPharm* October, **9**, 42–50 (1996).
52. A.S. Lubiniecki, M.E. Wiebe, and S.E. Builder, in A.S. Lubiniecki ed., *Large-Scale Mammalian Cell Culture Technology*, Marcel Dekker, New York, 1990, pp. 515–541.
53. C.F. Goochee, M.J. Gramer, D.C. Andersen, J.B. Bahr, and J.R. Rasmusen, *Biotechnology* **9**, 1347–1355 (1991).
54. J. Fu, W.J. VanDusen, D.G. Kolodin, D.O. O'Keefe, W.K. Herber, and H.A. George, *Biotechnol. Bioeng.* **49**, 578–586 (1996).
55. S.M. Kingsman, A.J. Kingsman, M.J. Dobson, J. Mellor, and N.A. Roberts, *Biotechnol. Genet. Eng. Rev.* **3**, 377–416 (1985).
56. J.D. Kittle and B.J. Pimentel, *BioPharm*, 48–51 (October 1997).
57. R. Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag, New York, 1982.
58. S.M. Wheelwright, *Protein Purification: Design and Scale up of Downstream Processing*, Carl Hanser Verlag, Munich, 1991, pp. 10–25.
59. D.P. Francis, P.M. Feorino, S. McDougal, D. Wrfield, J. Getchell, C. Cabradilla, M. Tong, W.J. Miller, L.D. Schultz, F.J. Bailey, W.J. McAleer, E.M. Scolnick, and R.W. Ellis, *JAMA* **256**, 869–872 (1986).
60. R.D. Sitrin, E.D. Wampler, and R.W. Ellis, in R.W. Ellis ed., *Hepatitis B Vaccines in Clinical Practice*, Dekker, New York, 1993, pp. 83–101.
61. R. Kieffer and L. Torbeck, *Pharmaceutical Technol.*, 66–76 (June 1998).
62. N.M. Lugo, *BioPharm* **11**, 18–26 (1998).
63. R.K. Gupta and G.R. Siber, *Vaccine* **13**, 1263–1276 (1995).
64. LeMoignic and Pinoy, *Comp. Rend. Soc. Biol.* **79**, 352–354 (1916).
65. G. Ramon, *Bull. Soc. Centr. Med. Vet.* **101**, 227–234 (1925).
66. G. Ramon, *Ann. Institut Pasteur* **40**, 1–10 (1926).
67. A.T. Glenny, C.G. Pope, H. Waddington, and U. Wallace, *J. Pathol. Bacteriol.* **29**, 38–45 (1926).
68. A.G. Johnson, S. Gaines, and M. Landy, *J. Exp. Med.* **103**, 225–246 (1956).
69. E. Ribi, *J. Biol. Res. Mod.* **3**, 1–9 (1984).
70. F. Ellouz, A. Adam, R. Ciorbaur, and E. Lederer, *Biochem. Biophys. Res. Commun.* **59**, 1317–1325 (1974).
71. L. Rong, E. Tarr, and T.C. Jones, *Clin. Infect. Dis.* **21**, 1439–1449 (1995).
72. G. Gregoriadis and A.T. Florence, *Drugs* **45**, 15–28 (1993).
73. R.P. Weissburg, P.W. Berman, J.L. Cleland, D. Eastman, F. Farina, S. Frie, A. Lim, J. Mordenti, M.R. Peterson, K. Yim, and M.F. Powell, *Pharm. Res.* **12**, 1439–1446 (1995).
74. H.S. Warren, F.R. Vogel, and L.A. Chedid, *Annu. Rev. Immunol.* **4**, 369–388 (1986).
75. H.S. Warren and L.A. Chedid, *CRC Crit. Rev. Immunol.* **8**, 83–101 (1988).
76. S. Shirodkar, R.L. Hutchinson, D.L. Perry, J.L. White, and S.L. Hem, *Pharm. Res.* **7**, 1282–1288 (1990).
77. R.K. Gupta, B.E. Rost, E. Relyveld, and G.R. Siber, in M.F. Powell and M.J. Newman eds., *Vacine Design: The Subunit and Adjuvant Approach*, Plenum Publishing, New York, 1995, pp. 229–248.
78. R.K. Gupta, P. Griffin, Jr., An-Cheng Chang, R. Rivera, R. Anderson, B. Rost, D. Cecchini, M. Nicholson, and George Siber, in S. Cohen and A. Shafferman eds., *Novel Strategies in the Design and Production of Vaccines*, Adv. Exp. Med. Biol., vol. 397, Plenum Press, New York, 1996, pp. 105–113.
79. J.G. Dickenson ed., *Dickinson's FDA Review*, Ferdic, Camp Hill Pa., 1998, pp. 1–24.
80. *BioPharm* **11**(8), 8–51 (1998).
81. W.K. Hubbard, *Current Good Manufacturing Practice: Proposed Amendment of Certain Requirements for Finished Pharmaceuticals, Food and Drug Administration* 21 CRF Parts 210 and 211, Docket No. 95N-0362, RIN 0910-AA45, (March 29, 1996).
82. N.M. Lugo, *BioPharm* **11**, 18–26 (1998).
83. A. Robbins and P. Freeman, *Sci. Am.* November, 126–133 (1988).
84. F.E. Andre, *Biologicals* **22**, 317–321 (1994).
85. H.S. Mason, T.A. Haq, J.D. Clements, and C.J. Arntzen, *Vaccine* **16**(13), 1336–1343 (1998).
86. C.O. Tacket, H.S. Mason, G. Losonsky, J.D. Clements, M.M. Levine, and C.J. Arntzen. *Nat. Med.* **4**, 607–609 (1998).
87. A.R. Brown, D.W. George, and D.K. Matteson, *Vaccine* **15**, 1165–1173 (1997).
88. L. Jodar, T. Aguado, J. Lloyd, and P.-H. Lamber, *Gen. Eng. News* **18**, 6, 40, 45 (1998).
89. D.M. Klinman, M. Takeno, M. Ichino, M. Gu, G. Yamshchikov, G. Mor, and J. Conover, *Springer Semin. Immunopathol.* **19**, 245–256 (1997).
90. J.J. Donnelly, J.B. Ulmer, and M.A. Liu, *Life Sci.* **60**, 163–172 (1997).

91. S. Langermann, *Nat. Med.* **4**, 547–548 (1998).
92. H.L. Robinson and C.A.T. Torres, *Semin. Immunol.* **9**, 271–283 (1997).
93. I.A. Ramshaw, S.A. Fordham, C.C.A. Bernard, D. Maguire, W.B. Cowden, and D.O. Willenborg, *Immunol. Cell Biol.* **75**, 409–413 (1997).

VALIDATION. See CLEANING, CLEANING VALIDATION; GOOD MANUFACTURING PRACTICE (GMP) AND GOOD INDUSTRIAL LARGE SCALE PRACTICE (GLSP); PROCESS VALIDATION; VACCINE TECHNOLOGY.

VENT GAS ANALYSIS

DAVID POLLARD
PETER SALMON
Merck and Co., Inc.
Rahway, New Jersey

KEY WORDS

Fermentation
Growth kinetics
Magnetic sector
Mass spectrometer
Online monitoring
Quadrupole
Respiratory quotient
Vent gas

OUTLINE

INTRODUCTION

The monitoring of exhaust gas is an ideal universal method for fermentation characterization because most biological processes involve exchange or consumption of gases and volatile compounds. Its noninvasive application does not compromise the sterility of the fermentor, providing a clear advantage over in situ probes, which contact the broth. The monitoring of gas concentrations, such as carbon dioxide and oxygen, in the entry inlet and outlet gas streams and incorporated into liquid and gas balances makes it possible to quantify rates of gas utilization on production and respiratory quotient. The use of these data in conjunction with the fast response of the analyzers allows on-line real-time monitoring of the physiological state of the fermentation, including growth kinetics, and substrate consumption and product formation information. This utilization has made the technology invaluable in research and manufacturing for applications such as fermentation optimization, batch-to-batch reproducibility, on-line troubleshooting, process decision making, and inoculation criteria (1,2). The aqueous concentration of organic volatile compounds, such as ethanol, butanol, and methanol, can also be determined from vent gas analysis and have been used to optimize productivity in industrial processes (3,4).

METHODS OF VENT GAS ANALYSIS

A variety of methods are available for monitoring gases. However, only the methods most applicable for monitoring gases of interest to bioprocesses are described here.

Thermal Conductivity

The sensor for thermal conductivity is typically a resistance temperature detector whereby gas is passed over a heated filament and the magnitude of cooling is measured by the change of resistance of the filament as measured by a Wheatstone bridge (5). For a single flowthrough unit, the gas flow rate must be kept constant. Units with a bypass have two flows; one path houses the sensor with two heated coils. The heated gas is transported from upstream to downstream coils, changing their resistance. The bridge circuit with constant current input gives an output voltage in direct proportion to the difference in resistance, with heat input and specific heat assumed constant. For detectors used in gas chromatography, the heat transfer of the sample gas is compared with the transfer of the carrier gas. The rate of transfer depends on the thermal conductivity

of the gas, which is proportional to the square root of molecular weight.

The Wheatstone bridge balances the current through the sample element to restore its temperature to that of the reference. Tungsten resistance wires are used for ambient and low temperature; metallic oxides are used for high temperature (5). The reference and sample detectors are housed in a single metal block with a high heat capacity to minimize errors from transient heating and cooling (6). Instruments are calibrated for the gas to be used and can have a fast response time of 1 s, detection threshold of 30 ppm, and accuracy of 2%. Carbon dioxide is often measured with this technique as the thermal conductivity of CO_2 is about two-thirds that of oxygen and nitrogen. The bypass units where the bulk of the flow bypasses the sensor are less sensitive to gas flow rate fluctuations but have larger response times, 20 s compared with 1 s. Although this technology has been used for many years it does not provide accuracy or long-term reproducibility. It is sensitive to gas flow rate, temperature, and mechanical vibration. Contamination of the sample cell can lead to drift, and frequent calibration is required of these inexpensive devices. Sample pretreatment is often required to remove contaminants by condensation and drying.

Flame Ionization

In this method the sample gas is burned in a hydrogen/air-induced flame. An electrical voltage is applied across two electrodes to remove the resulting ions from the flame. The current created across the two electrodes is a measure of the ion concentration (5). The detector has high linearity and sensitivity (ppb) with a short response time. It is efficient for a variety of organic components, but no response occurs for O_2, N_2, H_2, CO_2, Ar, H_2O, NH_3, or simple organics such as formic acid and formaldehyde. This characteristic makes the method universal for organic analysis by gas chromatography, especially because contaminants, water, air, and ammonia are not detected. However, its application for vent gas analysis is thus limited to ethanol and organic volatiles.

Gas Chromatography

Gas chromatography is the phase separation of gas component by adsorption equilibrium differences between stationary and mobile phases (6). The stationary phase is coated on small solid particles (typically silica or polystyrene) that are packed or coated onto the walls of highly efficient fused silica columns (several meters in length). The particles have an affinity for the chosen gas analysis to achieve separation of gas components based on equilibrium differences. The system requires a carrier gas to transport the gas sample through the stationary phase (typically oxygen-free dry helium), a sample injection system, column temperature controller, and a detector. Thermal conductivity and flame ionization detectors are most commonly used, as previously described. An analysis time of about 5–30 min is typical for each sample, which is much longer than other simpler, faster, and less expensive methods. However, gas chromatography has been used to determine H_2, CO_2, O_2, N_2, and organic volatiles.

Electrochemical Sensors

Amperometric methods can measure oxygen concentration in the vent gas using technology similar to the probes used for dissolved oxygen tension. The Clark principle incorporates the separation of electrodes from the broth using a semipermeable membrane (silicone or polytetrafluroethylene, PTFE) (7,8). A polarization voltage (600–800 mV) reduces oxygen at the cathode (Pt or Au) to peroxide and hydroxyl ions. The resulting current between cathode and anode (Ag/AgCl) is proportional to the oxygen partial pressure. Cl ions are consumed and supplied by the electrolyte (KCl). The probe relies on diffusion between gas, membrane, and electrolyte, which results in a slow response time delay, usually about 10–100 s to reach 63% of total change in dissolved oxygen level. The oxygen diffusion across the membrane results in a temperature-sensitive process. Equally, electrolyte depletion and water loss into the gas stream and membrane fouling make the system prone to drift. Thus, the technology is rarely used for vent gas monitoring.

Zirconia ceramic sensors have been developed for oxygen detection. Thimble-shaped, nonporous, oxygen ion-conducting electrolyte (yttrium-stabilized zirconia) separates two platinum electrodes and the reference from the measurement gas (9,10). The sample gas passes through the center of the sensor tube, contacting the anode platinum electrode, while the reference gas passes between the cathode electrode and the cylinder outer wall. The electrolyte, which separates the two gas phases, is a good conductor of oxygen ions because the ceramic lattice has vacancies for oxygen from the substitution of tetravalent zirconia ions by trivalent yttrium ions. When the oxygen sensor is exposed to a gas mixture, oxygen molecules absorb onto the electrode surface. This changes the molecule concentration on sensor electrode, which in turn changes the charged oxygen vacancy concentrations inside the solid zirconia electrolyte. An electrostatic field is generated from the differences in charge density between the charged oxygen vacancies inside the zirconia electrolyte and electrons in the electrode. The electromotive force (emf) generated across the electrochemical zirconia cell is proportional to the concentration of the adsorbed oxygen. The electrochemical reaction is temperature sensitive and takes place between 600 and 800 °C, requiring good temperature control. Temperature elevation can catalyze a reaction between platinum anode and combustible gases in the sample, resulting in measurement of net oxygen concentration rather than total oxygen concentration (6). For cost-effectiveness as opposed to accuracy, this sensor may be applicable as it is inexpensive, robust, and simple to operate, with accuracy about 1%, suitable for fermentation monitoring (11).

Alternative sensors have been developed for volatile organic species based on solid-state technology. The technology is mostly proprietary, but the sensors are often tin (IV) oxide (SnO_2) and have been demonstrated for ethanol monitoring of fermentation vent gas (12). The gas of interest is detected using two chemically inert electrical conductors of the same composition interfacing with a solid electrolyte. The gas diffuses into the electrolyte and alters

its impedance by a change in the charge density. These inexpensive sensors offer good sensitivity but have slow response times.

Infrared Absorption

The bonds of nonsymmetrical molecules vibrate and stretch from bombardment with infrared radiation, resulting in energy absorption. The specific vibration or rotation of a particular bond only takes place at a given frequency. For infrared adsorption, the molecule must be exposed to radiation of the precise frequency at which the molecular motion occurs. Thus, the fundamental frequencies of vibration and rotation for a particular compound can be used to identify it from a mixture. For fermentation analysis, carbon dioxide is the only gas of interest that can be detected by infrared (IR) because oxygen and nitrogen contain only one element. Most commonly used analyzers are nondispersive, which excludes most wavelengths apart from a few narrow bands of radiation. They are less selective than dispersive analyzers but more sensitive, simple, and reliable (13). Sources of radiation include quartz iodide lamps, tungsten filaments, and nichrome wires (6). A specific interference filter assembly, a monochromator, is chosen to filter out all wavelengths except for a desired wavelength band. The resulting radiation is a narrow band based around the mean of the desired wavelength (4,200–4,400 nm for CO_2). Alternatively, a cell containing the gas of interest is used to isolate the desired wavelength source.

In a typical optical arrangement, such as the Luft analyzer, an optical beam splitter is used to split the single radiation source to two beams, that pass into two separate chambers (Fig. 1). One chamber houses the sample cell and the other the reference cell. Both chambers contain a gas that absorbs at the desired wavelength, and the detector gas is usually the same as the desired gas for analysis. The sample cell is a flowthrough device with inlet–outlet ports designed to have zero dead legs. Its material depends on the corrosivity of the sample, usually stainless steel or Teflon, and the sample window is usually quartz or sapphire (6). Differences of IR absorption occur between the two chambers, reference and sample, generating pressure imbalances. This imbalance distorts a diaphragm that separates two chambers in the detector and forms part of a condenser microphone. A chopper wheel, a perforated or segmented disk, is used to block the radiation from one chamber and to allow radiation to reach the detector through the other chamber. The alternate blocking of the reference and sample chambers allows each radiation to be sampled at the appropriate side of the detector. The alternating flow of radiation produces an oscillating effect on the diaphragm. The amplitude of the oscillations is measured by the capacitance of the condenser microphone, which is correlated to the concentration of absorbed gas. A variation of this analyzer is to use a microflow sensor to replace the capacitance detector. Fluctuating radiation flow, generated by the chopper, is detected by a flow cell. The cell measures the temperature difference between two heated elements in the flow path. The temperature difference produces a corresponding difference of electrical resistance using a Wheatstone bridge that is correlated to the sample gas concentration.

Other alternative detectors include a thermocouple detector made of blackened gold foil to which two wire leads made of dissimilar metals are attached (6). The metals are selected to provide large thermoelectric emf. As the gold foil absorbs radiation, an emf forms between the gold foil and the metal. The strength of the emf is a function of the radiant power absorbed. The entire device is enclosed to minimize heat loss. Pyroelectric detectors absorb IR radiation, which raises the temperature of the ferroelectric crystal sensing element, typically Pd-$LiTaO_3$ (14), or $Pb(Zr_{52}Ti_{48})O_3$ (15,16). Spacing between the crystalline lattice changes as heat is absorbed, which affects the electric polarization of the element (14). Two electrodes are used, one of which is transparent to radiation, and the crystal is mounted between the electrodes. The sample gas concentration is correlated from the voltage required to balance the polarization charge of the crystal.

IR analyzers have been used to monitor CO_2 and organic volatiles such as ethanol for vent gas analysis (17), but they are sensitive to interference from other components of the gas such as water. The accuracy of IR analyzers is reasonable for fermentation processes, typically about 1% of the range, but are prone to drift (as much as 1% full scale per day for both the zero and span). The cal-

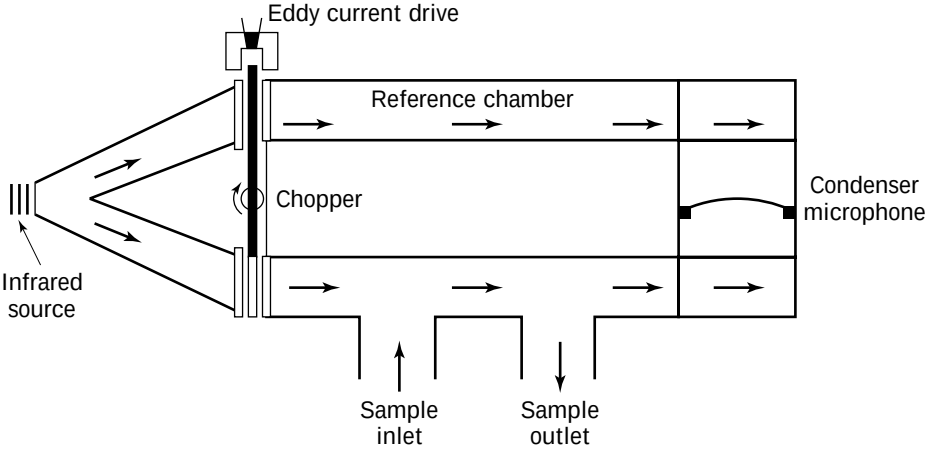


Figure 1. Luft nondispersive infrared analyzer with microphone sensor.

ibration curve is typically nonlinear, which gives added difficulty to frequent recalibration required to counteract drift. However, the 0–90% response time of 0.5–5 s is reasonably fast compared with other technologies already discussed.

Paramagnetism

Gases with unpaired electrons, such as oxygen, generally have high magnetic susceptibility. The unpaired electrons spin around the nuclei in a unbalanced manner. When a magnetic field is applied, the atoms orientate themselves in the orientation of the direction of the field, which has enabled specific instruments to be made for oxygen analysis. The principal instrument, a thermomagnetic analyzer, uses the principle that the magnetic susceptibility is proportional to the difference between sample temperature and the Curie point of 80 °C for oxygen (6). In a typical analyzer, gas flows through a heated magnetic field so that the oxygen becomes less magnetic. Cooler oxygen, which is more magnetic, pushes the hot oxygen, creating a magnetic wind. The magnitude of the flow determines the temperature of the coils, which changes the coil resistance as measured by a Wheatstone bridge. These instruments are sensitive to temperature variations and the thermal properties of the sample gas. Severe calibration shift and slow response times have also been reported (18).

An alternative analyzer is the magnetodynamic analyzer, which measures the mechanical deflection of a rigid test body (dumbbell) caused by the paramagnetic properties of oxygen. The dumbbell consists of two diamagnetic fused quartz spheres (2 mm in diameter) filled with nitrogen and mounted on a rod that is suspended by stiff quartz or platinum–iridium fiber in the cell chamber. A nonuniform magnetic field is applied to the gas sample in the chamber. The oxygen of the sample is attracted toward the most intense magnetic field, and the spheres are displaced until the force is balanced by the torque on the suspended fiber. The resulting torque is proportional to the difference of magnetic susceptibility between test body and surrounding gas. A light ray reflected by a mirror, mounted on the test body, transmits the angular deflection, which is directly proportional to the oxygen concentration in the chamber. Some instruments employ a feedback winding on the dumbbell with an electromagnet to maintain zero deflection, and the oxygen concentration is calculated from the current required. Unlike the thermomagnetic unit, the magnetic analyzer is not affected by temperature or thermal conductivity of the sample. However, care must be taken in orientation of the unit as it is sensitive to vibration. The unit is also sensitive to pressure as the amount of O_2 present in the sensing chamber, rather than the oxygen concentration, determines the balance of magnetic forces (6).

Another type of instrument is the Quincke-type analyzer, which measures the pressure difference when two gases with different oxygen concentrations come together in the magnetic field. The reference gas (N_2, O_2, air) flows through two restrictions so the gas entering either side of the sample chamber has equal flows (Fig. 2). One section of the reference gas is subjected to a pulsating magnetic field, and if oxygen is present then a pressure difference is created between the two reference flows. The Luft capacitance sensor has been used as a detector. Alternatively, Siemens AG developed an analyzer with a microflow detector in a connecting channel that detects the pulsing flow, producing a current proportional to the oxygen concentration (Fig. 2). The separation of the sample from the detector reduces the possibilities of drift and instrument failure from contamination or corrosion of parts.

These paramagnetic analyzers have found to be susceptible to vibration, pressure, and sample conditions with respect to moisture content and flow rate, resulting in frequent calibration to compensate for drift (2). Nevertheless, Quincke-type analyzers have been found to be more robust and have been widely used, including at the Merck & Co., Inc., production facilities. The units are priced similarly to IR analyzers but are more expensive than zirconia sensors. However, they provide only single-component analysis, with susceptibility to moisture and high sample flow requirement, and have poor accuracy (1–2% of range), with slow response times (5–100 s) (19). The units require calibration from drift, the frequency of which depends on care on calibration, maintenance of the machine, and sample treatment (20).

Mass Spectrometry

Mass spectrometry allows simultaneous measurement of both air gases (Ar, O_2, N_2, and CO_2) and volatile substances (including ethanol, acetone, and acetic acid) (21,22) present in the fermentor headspace. A mass spectrometer with multipoint analysis and fast response times (seconds) allows many fermentors and different processes to be monitored on the same instrument. The analysis involves admitting a gas stream to an ion source and then ionizing the molecules by electron bombardment under a high vacuum. The ions produced are subsequently separated, according to their mass-to-charge ratio, by either a magnetic sector or a quadrupole mass filter and detected sequentially. The systems are expensive but offer long-term accuracy and precision, with flexible analysis type and with sub-ppm detection limits and minimum operator intervention.

Vacuum System

The analyzer chamber of a vacuum system is kept under high vacuum (between 10^{-6} and 10^{-8} mmHg) to minimize collisions between particles. At this pressure, the mean free path of particles will be long enough, 50 m as opposed to <0.01 m at atmospheric pressure, and the number of collisions small enough for good analysis (6). Particle collision reduces instrument resolution and sensitivity by increasing kinetic energy distribution of the ions, inducing fragmentation or preventing ions from reaching the detector (23). The vacuum is created in two stages: a roughing pump, which reduces the vacuum to 10^{-3} mmHg, then a diffusion or turbomolecular pump, to obtain pressure as low as 10^{-9} mmHg. The roughing pump is usually a reciprocating mechanical pump lubricated with low vapor pressure oil to prevent oil contamination of the mass spectrometer (6). The diffusion pump uses an oil, often polyphenol

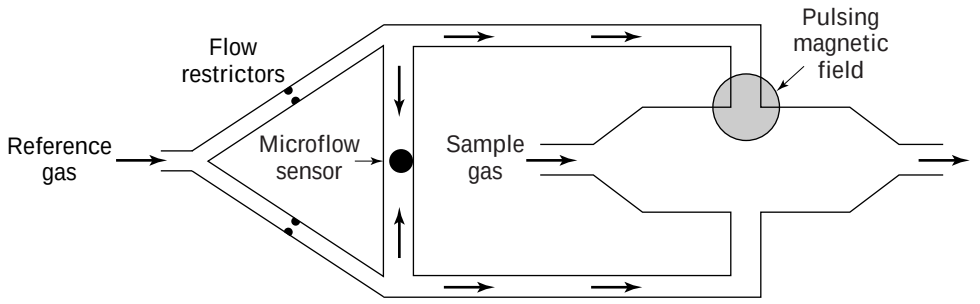


Figure 2. Quincke-type oxygen analyzer with capacitance sensor developed by Siemens AG.

ether, to force molecules out of the vacuum chamber. Oil is heated at the bottom of the pump, causing it to vaporize, and water-cooled coils at the top of pump then allow the oil to cool and condense. Large oil molecules force the smaller molecules down into the pump via momentum exchange, and smaller gaseous molecules exit through the mechanical pump at about 200 L/s. The disadvantage is that diffusion pumps are selective because noble gases are pumped at reduced rates (24). Turbomolecular pumping is mechanical; molecules are forced into the fins of the turbomolecular pump by momentum exchange. Fins revolve at about 10,000 rpm and produce an effective pumping capacity of about 400 L/s. Turbomolecular pumps are often preferred to diffusion types because they cause very low contamination by hydrocarbons and high vacuum is obtained very quickly on start-up. However, turbomolecular pumps are noisy, although they can be muffled to some degree, and highly mechanical, requiring regular servicing.

Sample Inlet

The capillary inlet design is most often used for mass spectrometer sampling, although some systems use a membrane inlet. The membrane material is typically silicone rubber or Teflon, which protects the vacuum integrity of the instrument and separates the analyzer from the atmosphere and contaminates such as foam or water. The membrane inlet has mostly been used for on-line in situ liquid analysis of organic volatiles in the fermentor (21,24–26) and has limited use for vent analysis because the membrane increases the response time to about 30 s. The most popular design is the capillary inlet on which a small representative sample flows from the main sample gas stream, which is continually venting, and "leaked" into the instrument, typically at 20 mL/min (18,27). The constant flow of gas stream from sample line to vent and to the leak of the analyzer means that line flushing can be ignored. The capillary is heated at 80–120 °C to preserve the vapor state of the sample. The rotary pump draws the sample by creating a pressure drop through the capillary. At the outlet, before the ionization chamber, the molecules cross a fritted glass or metal filter to further reduce pressure (13). The response times of such an inlet system can be about 0.5 s, achieved by low dead volumes and maintaining the whole system at 80 °C (28). Volatile components in the gas stream can contaminate the stainless steel capillary by adsorption and desorption, causing larger response times and memory effects from long transition times through the capillary (19,29). These factors can be reduced by using a capillary (diameter, 3 mm; length, 1–3 cm) lined with borosilicate glass that is heated resistively through a stainless steel outer sheath to 150 °C.

Ionization

The most widely used technique is electron impact ionization, where gas passes into an electron beam, resulting in electron ejection and a degree of fragmentation (2). Electrons ejected from a heated tungsten filament are accelerated through an electric field at 70 V to form an electron beam. The vacuum system draws molecules through the sample leak into the ionization chamber. Neutral molecules pass freely through the negatively charged ion repeller in front of the leak and are thermally desorbed into the electron beam, transferring kinetic energy to the sample molecule. The ionization (electron ejection) results in parent ions with the same mass to charge to MW with 6 eV of excess energy (Fig. 3). The excess energy in the molecule leads to some degree of fragmentation to molecular ions, fragment ions, and neutral fragments. A negative ionization step (electron capture) also occurs but is 100 times less efficient than electron ejection. Electron bombardment is most stable and reliable but is complicated by excessive fragmentation. Other soft ionization techniques have been developed: fast atom or ion bombardment, matrix-assisted laser desorption–ionization, and electrospray ionization, which produce only a few fragments. These methods give greater sensitivity and mass range but have primarily been developed for large molecules and solids analysis (23).

Ion Separation

Several methods are available to separate the ions after bombardment: quadrupole filter, magnetic analyzer, time-of-flight detectors, and Fourier transform detectors. The aim is to separate the ions of mass m from the ions of m', which is most commonly done by quadrupole filter or magnetic analyzer because these are easy to operate and compact (30).

The quadrupole filter is a set of four cylindrical rods, typically 125–300 mm long and 6–16 mm in diameter, held in a square array to which an RF and DC potential are applied (24,31). Two rods are connected and two are kept

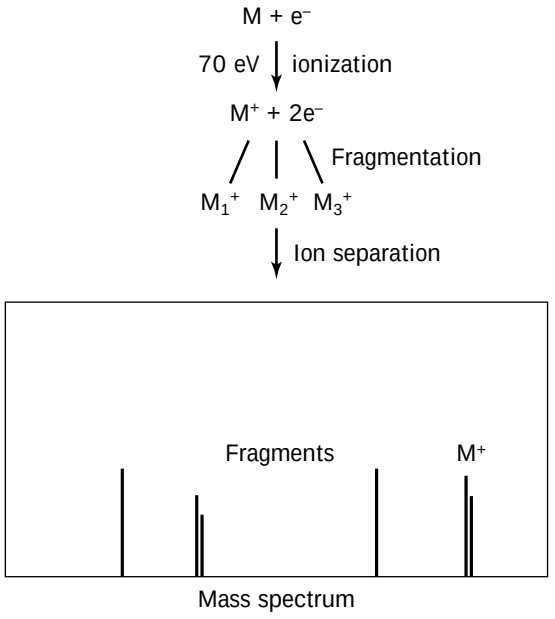


Figure 3. Ionization of molecules (M) producing fragment ions (M_1^+, M_2^+, and M_3^+).

180 degrees out of phase (Fig. 4). This forms crossed, oscillating electric fields perpendicular to the line of travel of the ions between the source and the detector. If the oscillating motion of the ion is out of resonance, the ions strike the walls and are not detected. Applying a certain voltage and RF can control the trajectory of a known mass-to-charge ratio to selectively strike the detector (30). Scanning RF and DC voltages from zero allows ions to be transmitted through the quad filter sequentially in order of their m/z values with constant resolution allowing multiple ion analysis. Resolution attained is determined by the accuracy of control of the parameters U, V, f, and ω radius within the quad rods (30).

The magnetic analyzer uses a magnetic field to separate the different ions. A moving charge passing through a magnetic field will experience a force and travel in a circular motion with a radius of curvature depending on the m/z of the ion. A magnetic analyzer separates ions according to their radii of curvature, and therefore only ions of given m/z will be able to reach a point detector at any given magnetic field (Fig. 5) (30). In a single detector system, an electromagnet is used to adjust the magnetic field to detect any m/z peak of interest (2). Alternatively, a multicollector system is set up whereby each ion fragment is measured by a small number of Faraday bucket collector assemblies, located along the focal plane according to which particular signals are of interest. Single detector systems are more common because they allow the combined use of the secondary electron multiplier that is too large to be used in conjunction with multiple detector systems.

Quadrupole systems are known to be faster, less expensive, and smaller than magnetic sector instruments. Yet the resolution is lower, easily contaminated, and less stable over time, making the magnetic sector a superior choice for industrial on-line monitoring (19,30). The response times of both systems are comparable because the magnetic hysteresis associated with magnetic sectors has been eliminated using laminated magnets. The precision from magnetic sectors comes from the flat-topped peak profiles of ion current versus magnetic field strength that they exhibit (Fig. 5). Thus, peak center and peak height can be precisely determined. Any minor drifting that may occur from temperature fluctuations has no effect, which results in high stability, accurate peak jumping, and long periods between recalibration (up to several weeks for fermentation application). Absolute accuracy is less than 0.05% for O_2 and less than 0.1 for N_2. Quadrupole filters exhibit a Gaussian-type peak profile and are more susceptible to drift (see Fig. 4) and temperature variation. Reduced accuracy and precision result in more frequent calibration. Insufficient resolution leads to an abundance sensitivity problem whereby peak discrimination is difficult at low concentrations with adjacent peaks. Experience with mass spectrometers at Merck during the past two decades favors the magnetic sector as opposed to quadrupoles. VG Gas Analysis Systems manufacture both quadrupoles and magnetic sectors and have shown improved accuracy and drift by a factor of 2 for magnetic sectors over quadrupole systems (27,29). Nevertheless, a considerable number of useful applications for fermentation vent gas analysis have been demonstrated using quadrupole systems (32,33).

Ion separation is still the limiting factor for mass spectrometry, however, because soft ionization produces larger molecular weight ions. Quads with electron ionization have been used since the 1950s but have found new utility when coupled with electron spray systems for solution analysis; however, this method has not been applied to vent gas analysis (23).

Detectors

Faraday cup and secondary electron multiplier are the two common detectors used for mass spectrometry. In the Faraday cup, ions pass the beam-defining slit of the cup and strike a metal dynode plate of a secondary emitting material, such as BeO, GaP, or CsSb, that induces several secondary electrons to be ejected and temporarily displaced. This temporary emission of electrons induces a current in the cup and provides a small amplification of the signal when an ion strikes the cup (23). The signal is measured by a electrometer and is directly proportional to the ion concentration. For the single-collector system, with a single amplifier and auto gain switching, high precision can be maintained over the widest range. In this arrangement any such drift is equal for all components and can therefore be compensated for by using any internal standard (2). The secondary electron multiplier (SEM) is a series of dynodes (about 10) maintained at ever-increasing potentials. Ions strike the dynode surface, resulting in emission of secondary electrons that are attracted to the next dynode where again secondary electrons are generated. A cascade of electrons is produced, transferring from plate to plate in a cycloidal flow path between two glass plates.

The signal amplification or current gain of an electron multiplier is 10^6 with a lifetime of 1–2 years, limited by

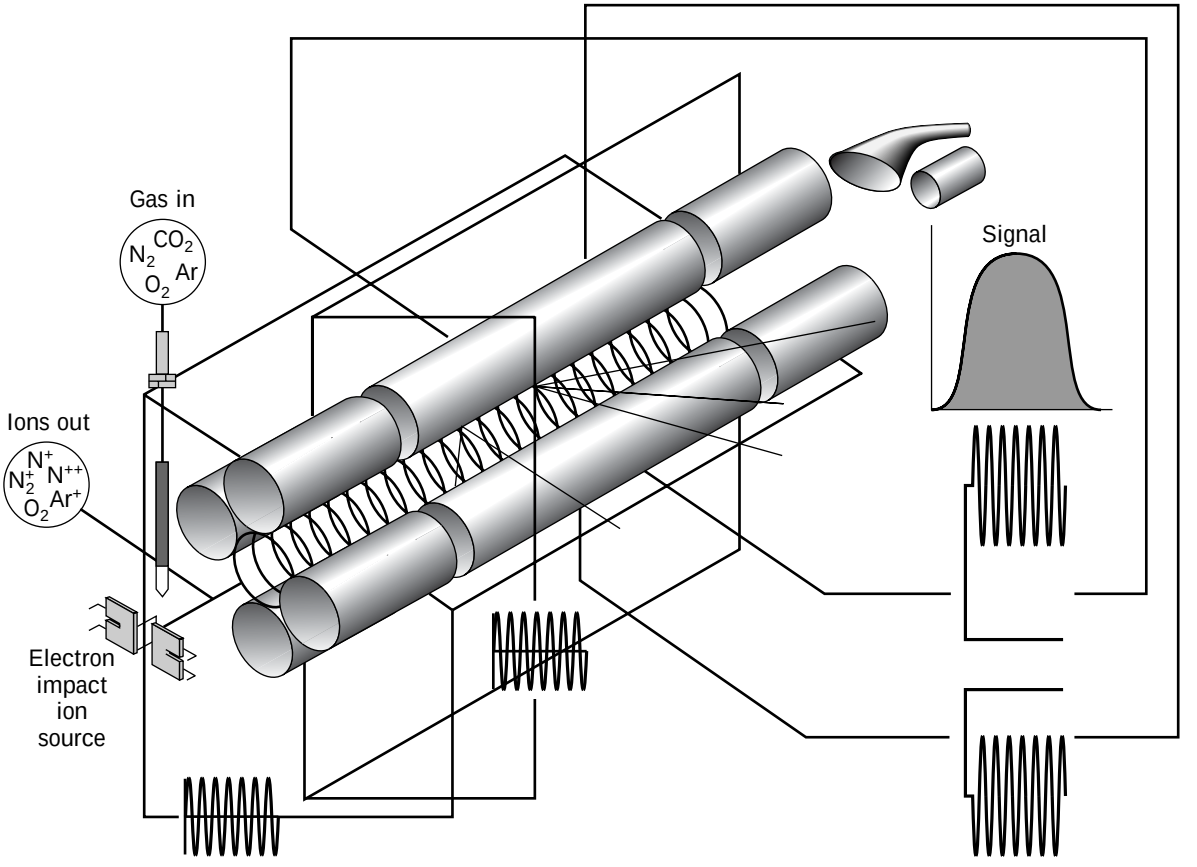


Figure 4. Quadrupole mass filter configuration and Gaussian-type profile curve for the signal, ion current versus electric field strength (VG Gas Analysis Systems Ltd.).

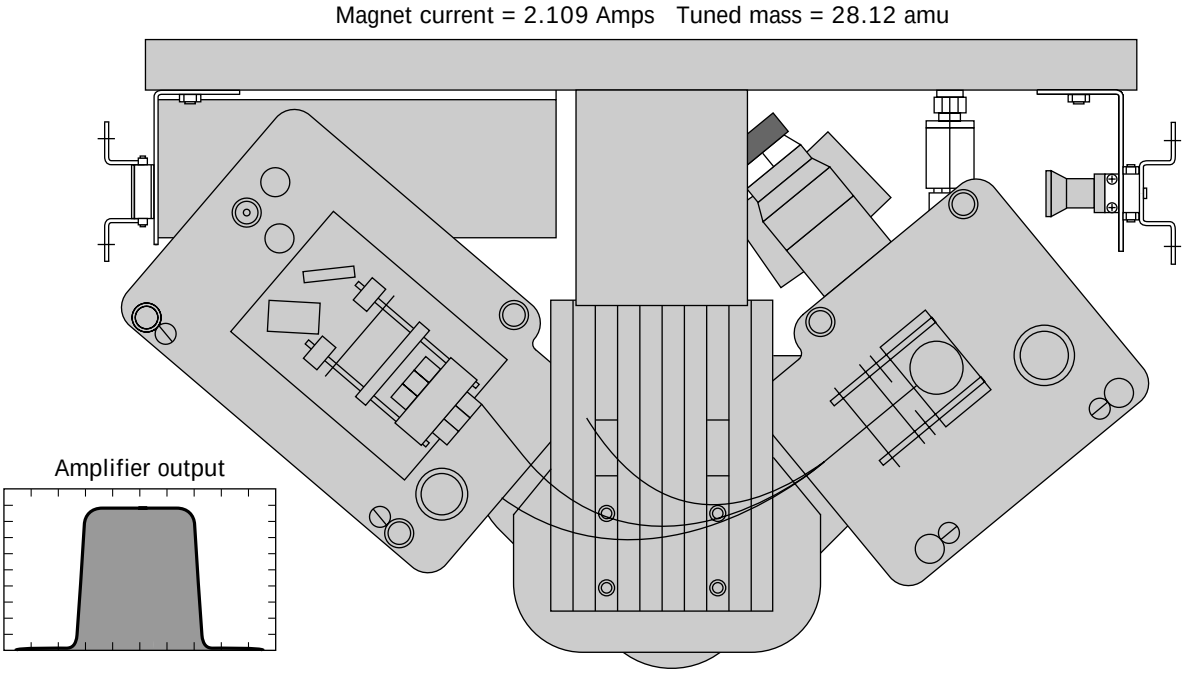


Figure 5. Magnetic sector mass spectrometer showing ion trajectory through magnetic field and the flat-top type profile of the ion current versus magnetic field strength (VG Gas Analysis Systems Ltd.).

surface contamination from incident ions or poor vacuum (19). The Faraday cup is considered rugged, reliable, and typically capable of measurements to 1–10 ppm concentration levels (28). The SEM detector offers greater sensitivity (sub-ppm) with a faster response but has poorer long term stability and is generally less accurate (2). SEM can be sensitive to repeated air exposure; background noise increases when the gain decreases, and it is far less rugged than the Faraday cup (13). In a single-collector arrangement with either a magnetic sector or quadrupole system, it is common to have remotely switchable Faraday cup–SEM assemblies. The appropriate detector is selected by the software, allowing all species to be monitored from 100% to parts per million (24). For example, a Faraday cup is used for air components and a secondary system for volatiles such as ethanol, acetone, or butandiol (22).

Spectral Analysis

To analyze a mixture of ion fragment intensities, with n components, a system of linear equations must be solved to calculate individual component concentrations $c_1, c_2 \ldots c_n$. The ion current signals, S_1, for each mass/charge is considered to be the sum of the contributions of all components, which are proportional to the concentration of each component:

$$\begin{aligned} s_1 &= c_1 m_{11} + c_2 m_{12} + \ \ldots \ + c_q m_{1q} \\ &\cdot \quad\ \cdot \quad\ \cdot \quad\ \cdot \\ &\cdot \quad\ \cdot \quad\ \cdot \quad\ \cdot \\ &\cdot \quad\ \cdot \quad\ \cdot \quad\ \cdot \\ s_p &= c_1 m_{p1} + c_2 m_{p2} + \ \ldots \ + c_q m_{pq} \end{aligned}$$

or

$$\vec{s} = \mathbf{M}\vec{c}$$

where s is the signal, M is the fragmentation pattern or sensitivity factor matrix determined by calibration with individual pure components, and c is the desired concentration (28). Most species produce several peaks due to fragmentation or double ionization. For instance, nitrogen has peaks at nominal m/z values 28, 14, 29, and 15. Thus, a set of linear equations is overdetermined and solved using an iterative route of multidimensional least squares technique (34). The best fit is then made between the observed spectrum and that which can be synthesized following small adjustments of the species concentration. These adjustments use the known fragmentation patterns of various species. Therefore, all the mass spectral information is made available by the single detector instrument, which is most effectively used in achieving the concentration analysis.

Suitability of Mass Spectrometry

A typical magnetic sector mass spectrometer instrument has excellent repeatability of $\pm 0.1\%$ of range at calibration, zero and gain drift less than 1% of range per month, and a response time of 1 s (1,19,28). The limiting step of the response time becomes the passage through the sampling system rather than the mass spectrometer itself. The fast analysis times (seconds) mean that many fermentors and indeed several different processes can be monitored simultaneously. This method provides a cost-effective device for multifermentor sampling despite the high purchase cost of an individual unit. Maintenance is usually associated with contamination of the inlet device and filament and ion pump replacement. However, this is offset by the rapid analysis for polar organic species, excellent long-term accuracy and precision, plus flexibility of analysis. These features make mass spectrometry the most flexible and versatile instrument for vent gas analysis.

INSTALLATION AND OPERATION

Careful consideration must be made in designing a sampling system for a vent gas system, which usually involves the multiplexing of one analyzer to a variety of fermentors. A sample handling system incorporates gas transport from the fermentor, followed by pretreatment before passing through the multipoint sampling manifold into the analyzer. The design of such a system must minimize the response time of sample analysis. The gas must take the shortest route from the fermentor to the pretreatment steps through to the analyzer. The amount of sample pretreatment required can be process specific. Most applications require removal of moisture and contaminating species from the vent gas, which involves steps such as filtering and heating before the analyzer. The materials of construction for transport lines, line purging, sample flow rate, and minimizing tubing lengths must all be considered.

Fermentor Sampling Location

The position of a gas sampling point from the fermentor requires care and to some degree depends on the species for analysis. If fermentor overpressure is to be used for gas transport to the mass spectrometer, then it is pertinent to have the sample port between the headplate and the control valve. The vent lines of fermentors often contain a condenser, double-piped steam heater, or vent filters that are steam dried from a jacket. These techniques can remove volatile compounds from the exhaust gas, so the sampling point should be upstream of these devices if organic volatile analysis is desired. The removal of water via the condenser may also affect carbon dioxide because the solubility of this species is high. Nevertheless, moisture must be removed before the analyzer, so sampling after the condenser may be advantageous.

If fermentor operation requires the use of a vent filter for containment, then sampling must be completed after the filter unless a separate vent gas filter is installed. Foaming and liquid entrainment can occur in the vent line. Some prevention of entrainment into the sampling system can be made by facing the sample port downstream of the gas flow (34). Gas has to make a 180-degree turn to the sample port, while tending to exclude liquid as its inertia prevents it from turning. Foam detection systems can also prevent sample line contamination. The foam sensor, using capacitance or conductivity between two elements, usually probe and tank wall or two elements in the probe, can be

mounted on the fermentor headplate; this can be part of an automatic defoamer addition system on foam detection. A second probe can be mounted in the vent gas line to close the off-gas sample valve on foam detection.

Sample Delivery and Pretreatment

Commonly, the fermentor headplate overpressure is used to transport gas to the analyzer. However, for laboratory-scale vessels a gas membrane or piston pump may be required. A 0.1-bar pressure is sufficient for a 100-m length of standard 6-mm nylon tubing (or ¼-in tubing) (28). The tube can be stainless steel, nylon, or preferably PTFE for volatile samples. The tubing is heated to between 60 and 160 °C to ensure compounds in the vapor phase, which avoids removal via condensation; this can be done electrically, by steam, or by jacketed hot water (29). For volatile analysis, heating the tube to about 150 °C and the use of Teflon are essential for fast responses times (35). Adsorption–desorption effects of volatile components with stainless steel tubing walls can hinder response times (>5 min) and cause memory effects.

The vent gas is often passed through a (0.2-μm) hydrophobic filter (PTFE) as a final particulate protector before the analyzer. For unheated lines, moisture is removed by condensation. Knockout pots are used in combination with a coalescing filter whereby large drops form and settle by gravity (34). A graded porosity hydrophilic matrix is used with a drain layer to direct coalescing liquids into a sump at the base of the housing. Automatic draining of the pot is ideal via solenoid actuation periodically or on liquid detection. If further moisture removal is required, various technologies are available, including small concentric tubes or coils in shell heat exchangers with refrigerated coolant. Other systems include the Peltier effect cylindrical gas cooler with a sealed tube containing a pure fluid. A wick lines the inner walls, which ensures fluid return by capillary action. One end of the tube is heated, liquid evaporates, and the gas flow displaces the steam toward the end of the tube. The steam expands and condenses back to the heated end (13). Desiccants such as anhydrous calcium chloride, alumina, or silica gel can be used for further sample drying, but this is rarely necessary.

Sample-Handling Manifolds

Most industrial applications utilize one mass spectrometer for vent gas analysis of an array of fermentors (>10). This requires a fast manifold system free of cross-contamination and leaks to automatically switch between fermentor samples and calibration gases. A typical system has the sampling streams from all tanks under permanent flow to guarantee instant sampling availability and avoid line purging (Fig. 6). Three-way valves can eliminate dead legs, and use of two valves per sample stream allows gas to flow to drain rather than contaminate other streams if both valves fail. Solenoid valves are often used as they are inexpensive and compact compared to air-actuated ball valves. Rotameters can be used to verify flow at the vent side of the three-way valve. The manifold is mounted in a cabinet with heat control to a temperature above the process temperature or to 110 °C required for organic volatiles. Multistream rotary valves have been designed that are actuated by an electric stepper motor for cross-contamination sampling (27,28). The unit (designed by VG Gas Analysis Systems Ltd.) can be configured for 16, 32, 48, or 64 sample streams, connected to the periphery of a circular flat distribution plate, that vent into a cylindrical housing with a common exhaust (Fig. 7). The electric stepper motor drives the rotary arm to position at a sample stream, and sample flows directly into the sample transfer tube. Cross-contamination of other gases is inhibited by a spring-loaded face sealing sleeve (2).

The response time of the manifold system can be reduced by minimizing tubing lengths. The time required to flush the lines should be four times the gas residence time (2). This is estimated using plug flow, but experimental measurements are preferred because of backmixing through valves and regulators.

Sample Flow Control

The signal from a gas analyzer is proportional to the volumetric concentration of the component of interest so that sample pressure will directly influence the results. This can be remedied by measuring internal standards (typically N_2, or Ar), or by measuring the concentrations of all other components, and converting to mole fractions. Alternatively the pressure can be measured and the data corrected using the ideal gas law or by controlling the pressure at the fixed value (34). Pressure control can be overcome by simply venting the instrument outlet to atmosphere. An upstream needle valve can be used to reduce the pressure and to control sample flow rate through instrument. Locating the valve as far upstream as possible reduces the lag time, as the gas velocity is increased after the pressure reduction. However, this leaves a 2–3% variation from atmospheric pressure. If this is an unacceptable error, then an absolute backpressure regulator must be employed on the outlet (34). Temperature is controlled within the analyzer but it is useful to control temperature of the manifold cabinet by an electric heater. A temperature greater than 60 °C will prevent condensation, unless moisture has been removed, and minimize adsorption of compounds.

System Monitoring

Water can be detected in a manifold system by a U-shaped liquid trap with a sensor to detect water in the sample lines. The low electrical conductivity of water in the sample means that the sensor must monitor either the dielectric constant or refractive index of the stream or even thermal dissipation from a self-heated thermistor or electric coil. The sample flow rate (>1 mL/s) can be controlled by a simple needle valve and rotameter. A thermal sensor and auto control valve will ensure sample flow rate is adequate and stable. Leak testing by pressuring the system from instrument air and closing the valves. Alternatively, the pressure can be monitored while the sample system is evacuated if this step is part of the sampling sequence. Diagnostic packages supplied with most analyzers will alarm for problems associated with the capillary, filament, vacuum, and power supply.

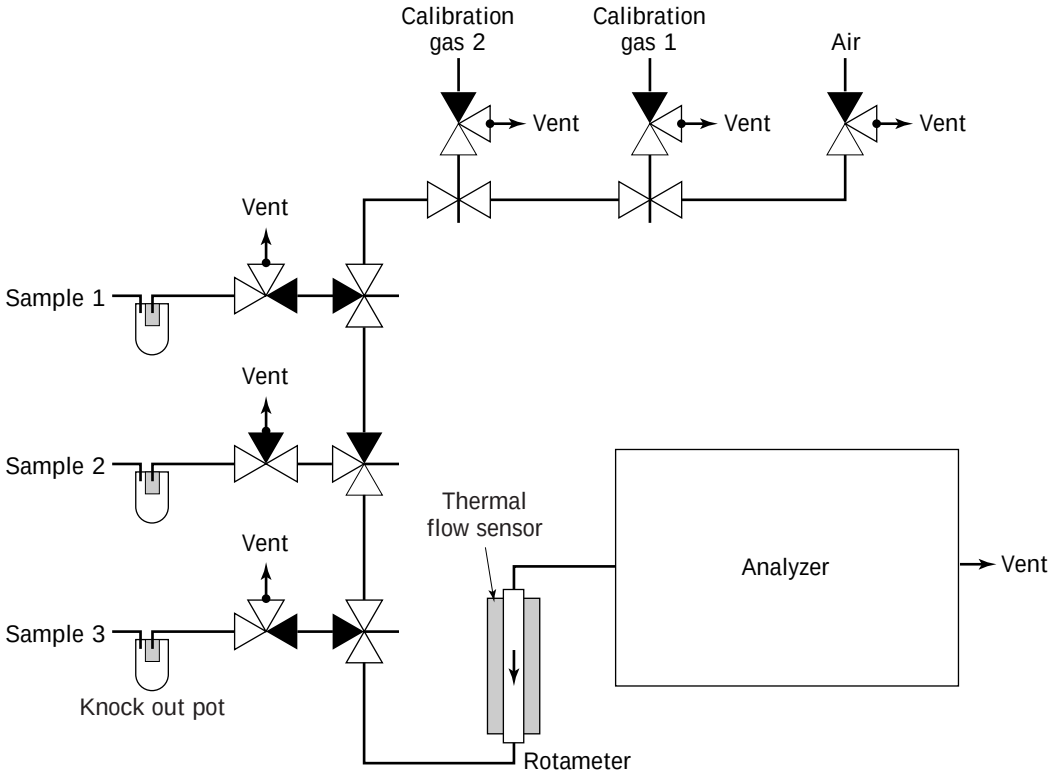


Figure 6. Multistream sample switching system. Solid segments indicate valve-closed position.

Process Measurements Required for Variable Calculation. Measurements of pressure, volume, air flow, and dissolved oxygen are required. To calculate the vent gas variables that are useful to fermentation (oxygen uptake rate [OUR], carbon evolution rate [CER], and respiratory quotient [RQ]). Broth volume can be determined by load cell or differential pressure transducers, air flow rate by a mass flow meter, and dissolved oxygen concentration by a polarographic probe whose response should be checked systematically by varying the headspace pressure or degassing with N_2. Strain gauge, piezoresistive, capacitance, or resonant wire transducer can be used for pressure measurement.

Data Acquisition. For a number of fermentors connected to a multistream manifold, such as a RMS, fast turnover of samples is required although measurements must be accurate. Therefore data acquisition must occur after a suitable line purge time. Multiple measurements can subsequently be made and filtered or smoothed to reduce noise (2). A typical approach is to use an algorithm to produce a stable reading (34). A linear exponential filter is used where each successive measured value c_t is used to update a smoothed estimate as $\hat{c}_t = \kappa c_t + (1 - \kappa)\hat{c}_{t-1}$ where k is a filtering parameter. The trend in the smoothed data is estimated as $\hat{S}_t = \lambda(\hat{c}_t - \hat{c}_{t-1}) + (1 - \lambda)\hat{S}_{t-1}$, where λ is another parameter. A typical algorithm acquires at least $1/\kappa$ values and continues until the absolute value of the trend $\hat{S}_t$ is less than a target tolerance, δ, indicating that a stable result has been obtained. The values of κ, λ, and δ depend on the rate of data acquisition and on instrument performance (34).

Vent gas analysis processes gases on a dry gas basis whereas the actual gases from the fermenter are saturated with water vapor. Standard handbooks can be used to correct these data (36). Most vent gas analysis is expressed as a mole fraction basis. If this is not the case, then data should be adjusted using the ideal gas law to standard conditions of temperature and pressure to allow mass balances to be performed.

Good precision is achieved if signals are noise free with no drift. However, the combination of vent gas data with online data such as air flow rate and dissolved oxygen for mass balancing introduces noise. These errors of measurement can dramatically enlarge errors in the data. Low-pass filters can be used to smooth the data, and many acquisition systems have built-in analog filters for this purpose.

PARAMETER CALCULATION

Vent gas analysis data are only useful when configured to quantify consumption rates and metabolic factors, which requires equation balances for the gas and liquid phases. When completed, the vent gas variables enable characterization of the microbial physiology of the fermentation broth.

Well-Mixed Liquid- and Gas-Phase State

The following balance equations for species i in the gas and liquid phases are written assuming well-mixed phases in

Figure 7. Magnetic sector mass spectrometer unit from VG Gas Analysis Systems Ltd. fitted with the rotary motor sample device.

the fermentor; this assumes that concentrations are uniform throughout the liquid phase, and that mole fractions are uniform throughout the gas phase. The assumption is accepted for small-scale stirred reactors but is less valid at large scale, >100 L. Axial volumetric concentration gradients of species in the gas phase as a consequence of the hydrostatic pressure gradient have been observed at large scale, which leads to alternatives for large-scale vessels, discussed in the next section. For low pressure and dilute concentrations of interest, the ideal gas law and the simplest form of Henry's law (relating equilibrium gas and liquid-phase concentrations through a simple proportionality constant) can be used. The equations have been given by Heinzle and Dunn (2) and modified by Salmon and Buckland (34). It is assumed that the fermentor is isothermal and has a uniform cross-sectional area.

$$\frac{dV_G C_{G_{e,i}}}{dt} = G_o C_{G_{o,i}} - G_e C_{G_{e,i}} - k_{L_i} a V_L (C^*_{L_{e,i}} - C_{L_{e,i}}) \quad (1)$$

$$\frac{dV_L C_{L_{e,i}}}{dt} = L_o C_{L_{o,i}} - L_e C_{L_{e,i}} + k_{L_i} a V_L (C^*_{L_{e,i}} - C_{L_{e,i}}) - V_L \Sigma r_i \quad (2)$$

where $C_{G_{o,i}}$, $C_{G_{e,i}}$, $C_{L_{o,i}}$ and $C_{L_{e,i}}$ are the gas- and liquid-phase concentrations, subscripts *o* and *e* refer to the influent and effluent, and subscripts *G* and *L* refer to the gas and liquid phases, respectively, V_G and V_L are the gas- and liquid-phase volumes, G_o, G_e, L_o, and L_e are the influent and effluent gas and liquid flow rates, $k_{L_i}a$ is a mass transfer coefficient, $C^*_{L_{e,i}}$ is the liquid-phase concentration that would be in equilibrium with $C_{G_{e,i}}$, and Σr_i is the total consumption rate of species *i* in the liquid phase. Standard conditions of pressure and temperature (typically $P_{ref} = 1$ atmosphere, $T_{ref} = 25$ °C) have been used for gas-phase concentrations, volumes, and flow rates.

An expression can be written for $C^*_{L_{e,i}}$, liquid-phase concentration, using Henry's law and integrating over the fermentation broth volume (34):

$$C^*_{L_{e,i}} = (RT_{ref}/H_i P_{ref})(P_h + V_L \rho_L g/2A) C_{G_{e,i}} \quad (3)$$

where H_i is the Henry coefficient, P_h is the headspace pressure, ρ_L is the liquid-phase density, and A is the cross-sectional area of the fermentor. It is usual to assume that the liquid-phase supply, removal, and accumulation terms in equation 1 are negligible for gases such as oxygen that have low aqueous solubility. The gas-phase accumulation term in equation 1 is also ignored because gas holdup is small and changes in oxygen utilization occur slowly. However, this simplification may not be valid for rapidly growing cultures (37) or when sudden metabolic changes occur (38), in which case the calculated results will lag actual events in the fermentor. Using the discussed assumptions, equations 1 and 2 can be added and rearranged to obtain an expression for $OUR = r_{O_2}$, the net oxygen utilization rate:

$$OUR = G_o(C_{G_{o,O_2}} - G_e/G_o C_{G_{e,O_2}})/V_L \quad (4)$$

An expression for the mass transfer coefficient, in a dimensionless form, can also be made from equation 1:

$$\alpha_{O_2} = \frac{k_{L_{O_2}} a V_L}{G_o} = \frac{C_{G_{o,O_2}} - G_e/G_o C_{G_{e,O_2}}}{\beta_{O_2}(1 + \gamma/2) C_{G_{e,O_2}} - C_{L_{e,O_2}}} \quad (5)$$

where $\beta_{O_2} = R\ T_{ref} P_h / H_{O_2} P_{ref}$, and $\gamma = V_L \rho_L\ g/AP_h$. Values of the Henry coefficient, H_{O_2}, can be obtained from the literature (39). The aqueous oxygen concentration, $C_{L_{e,O_2}}$, found in equation 5, can be measured on line using a polarographic probe calibrated to read $\beta_{O_2}(1 + \gamma/2) C_{G_{o,O_2}}$ before inoculation. The constant, 1, assumes that the difference of oxygen solubility in the fermentation medium compared to water is negligible. This assumption is reasonable as the constant can be estimated for most fermentation broths (39,40) and a value of 0.945 is achieved for a typical broth containing, salts, yeast extract, and glucose.

Inert gas species (such as Ar or N_2) are monitored to compensate for gas flow changes caused by the presence of moisture and volatiles in the exhaust gas. The net utilization for the inert gas is zero and consequently $G_e/G_o = C_{G_{o,inert}}/C_{G_{e,inert}}$ by using the simplifications for oxygen.

For carbon dioxide the situation is more complex because its high solubility (H_{CO_2} of 34 L bar/mol compared with H_{O_2} of 856.9 L bar/mol) means the liquid-phase accumulation cannot be ignored. Also, carbon dioxide reacts with water to give biocarbonate at pH > 6.5.

$$CO_2 + H_2O \underset{k_b}{\overset{k_f}{\rightleftarrows}} H_2CO_3 \overset{K_a}{\leftrightarrow} H^+ + HCO_3^- \leftrightarrow 2H^+ + CO_3^{2-} \quad (6)$$

The hydration of CO_2 to H_2CO_3 is a slow rate constant; however, its concentration is low (<0.3% of CO_2 present) (41) because it is almost instantaneously dissociated to HCO_3^-. At the pH range 4–8 for fermentation processes, the concentration of carbonate ions is negligible, complexing of CO_2 with amine groups can be ignored, and the reaction of dissolved CO_2 with OH^- is negligible (42). Bicarbonate accumulation increases with pH, and at pH 6.3 the concentration can be equal to the dissolved CO_2. Thus, the basic liquid-phase equation 2 must be modified to include accumulation of bicarbonate:

$$\frac{dV_L C_{L_{e,CO_2}}}{dt} = L_o C_{L_{e,CO_2}} - L_e C_{L_{e,CO_2}} + k_{L_{CO_2}} a V_L (C^*_{L_{e,CO_2}} - C_{L_{e,CO_2}}) + V_L(CER - r_{HCO_3^-}) \quad (7)$$

$$\frac{dV_L C_{L_{e,HCO_3^-}}}{dt} = L_o C_{L_{o,HCO_3^-}} - L_e C_{L_{e,HCO_3^-}} + V_L r_{HCO_3^-} \quad (8)$$

where $CER = -r_{CO_2}$ is the net carbon dioxide evolution rate from the biomass and $r_{HCO^-}3$ is the net rate of bicarbonate formation given by

$$r_{HCO_3^-} = k_f C_{L_{e,CO_2}} - k_b/K_a \, C_{L_{e,H^+}} \, C_{L_{e,HCO_3^-}} \quad (9)$$

Values for the constants k_f, k_b, and k_a can be obtained from the literature (43). Rearrangement of equation 1 gas-phase balance, and neglecting accumulation terms as before can produce the following balance:

$$k_{L_{CO_2}} a V_L (C^*_{L_{e,CO_2}} - C_{L_{e,CO_2}}) = G_o C_{G_{o,CO_2}} - G_e C_{G_{e,CO_2}} \quad (10)$$

and solved for the liquid-phase CO_2 concentration,

$$C_{L_{e,CO_2}} = \left\{\frac{1}{\alpha_{CO_2}} \frac{G_e}{G_o} + \beta_{CO_2}(1 + \gamma/2)\right\} C_{G_{e,CO_2}} - \frac{1}{\alpha_{CO_2}} C_{G_{o,CO_2}} \quad (11)$$

where $\alpha_{CO_2} = k_{L_{CO_2}} a V_L / G_0$, and $\beta_{CO_2} = RT_{ref} P_h / H_{CO_2} P_{ref}$. Henry coefficient H_{CO_2} estimates are found in the literature (39). Combining and rearranging equations 7, 8, and 10 gives

$$CER = \frac{1}{V_L}\left\{\frac{dV_L(C_{L_{e,CO_3}} + C_{L_{e,HCO_3^-}})}{dt} - L_o(C_{L_{e,CO_2}} + C_{L_{e,HCO_3^-}}) + L_e(C_{L_{e,CO_2}} + C_{L_{e,HCO_3^-}}) - G_o(C_{G_{o,CO_2}} - G_e/G_o C_{G_{e,CO_2}})\right\} \quad (12)$$

The evaluation of $C_{L_{e,HCO_3^-}}$ can be considered in three algorithms of complexity (34). Equations 8 and 12 can be solved progressively as the fermentation proceeds using current and past data. This practice enables monitoring of CO_2 evolution under dynamic conditions, so long as the rate of data acquisition by the analyzer is sufficiently rapid. This may be a problem when acid is introduced intermittently for pH control because the CO_2 concentration in the vent gas will increase for a short period after each addition (44). A second approach can be applied if the time constant for process changes is greater than 5 min. In this case, the hydration reaction can be assumed to be at equilibrium, yielding $C_{L_{e,HCO_3^-}} \approx (k_f/k_b) K_a 10_{pH} C_{L_{e,CO_2}}$. Substituting this expression into equation 12 allows the value of CER to be calculated directly if the rate of change of the dissolved CO_2 concentration is estimated using current and past data (34). Finally, it is common practice to ignore the transient terms so long as changes of pH and CER are slow. Heinzle et al. (45) demonstrated that including dynamic changes of CO_2 and HCO_3^- in a simulation only produced relative errors <2% in the CER.

Large-Scale Mixing Models

The assumption of ideal mixed gas and liquid phases is reasonable for tanks less than 100 L. However, upon scale up to production scale (100– 100,00 L), axial gradients and large mixing times are observed and so previous equations for oxygen transfer rate, OTR, and OUR based on the well-mixed system may not be valid. Hydrostatic pressure gradients in tall fermentors will cause large differences in oxygen solubility C_L from liquid surface to tank bottom. The dissolved oxygen concentration in a stirred tank can vary linearly from 5.4 mg/L at the surface to 12.1 mg/L at a depth of 10 m. At the bottom of the tank the oxygen solubility for a given gas composition is twice as a consequence of the doubled hydrostatic pressure, as shown by Henry's law. This change will double the maximum oxygen transfer rate, and, hence, dissolved oxygen concentration, oxygen solubility, gas holdup, and transfer parameters may vary with depth in a tall reactor.

Limited experimental and theoretical data are available on concentration in homogeneities or gradients within large fermentors, but radial and axial gradients of pH, conductivity, and dissolved oxygen have been observed (46,47). Residence time distribution data using liquid and gas transfers have characterized the mixing to produce a variety of models (48–51). Various models have been determined for large-scale mixing, where axial backmixing has been described by dispersion theory or a tank in series with recirculation based on the measurement of residence time distribution. Other models include compartmentalization of the reactor into zones, such as two well-mixed impeller compartments and the remainder for the rest of the tank, with transport between compartments based on the pumping power of impellers (46).

The model behavior is described through simulations of the coupled material balance equations for oxygen transfer rate, mixing, and kinetics. The importance of backmixing and hydrostatic pressure on scale-up was demonstrated by using a three-phase multimixed compartment model (50). The flow conditions of these models fall between extremes of well-mixed and plug flow. Simplification for large-scale reactors is to assume that the liquid phase is well mixed but that the gas phase moves in plug flow (no backmixing) from inlet to outlet (51). The gas-phase concentration of species *i* was described by Salmon and Buckland (34) using this equation:

$$\frac{dC_{G_i}}{dx} - \alpha_i\{\beta_i[1 + \gamma(1 - x)]C_{G_i} - C_{L_{e,i}}\} = 0 \qquad (13)$$

where α_i, β_i, and γ are defined in the previous section. For the gas-phase concentration of species *i*, a closed-form solution to this equation was used to calculate $C_{G_{e,i}} = C_{G_i}(1)$ from the inlet condition of $C_{G_i}(0) = C_{G_{o,i}}$. Salmon and Buckland (34) used this solution to produce a comparison between the estimates of equilibrium between liquid and gas phases at two scales, 0.15 and 60 m^3 (where differences are solely due to hydrostatic pressure), and between the two models of well-mixed phases and a gas phase with plug flow. For a given mass transfer coefficient, the predicted liquid-phase concentrations were closer to equilibrium with the effluent gas for the plug flow model than with the well-mixed system. This difference was accentuated with the large-scale tank (where hydrostatic pressure is significant). This conclusion seems reasonable because the mass transfer coefficient required to achieve a desired liquid-phase concentration is lower for a production vessel than for a small laboratory fermentor. Thus, for a given mass transfer coefficient the gas-phase and liquid-phase concentrations in a large vessel will be closer to equilibrium than those in a small vessel. This agrees with the observation that the dissolved carbon dioxide in a large tank can be calculated by presuming equilibrium with the vent gas, whereas this assumption is not appropriate for small fermentors.

Aqueous Concentration Estimation for Volatile Organic Compounds

The liquid-phase concentration for an organic species can be calculated as for $C_{L_{e,CO_2}}$ (equation 11) assuming that the system is well mixed with respect to the gas and liquid phases and that the inlet gas-phase concentration is zero (34):

$$C_{L_{e,i}} = \beta_i\left[\frac{1}{\alpha_i\beta_i}\frac{G_e}{G_o} + (1 + \gamma/2)\right]C_{G_{e,i}} \qquad (14)$$

As $\alpha_i\beta_i \to \infty$, the liquid-phase concentration approaches $\beta_i(1 + \gamma/2)\, C_{G_{e,i}}$, which is the concentration in equilibrium with the gas phase at the average fermentor pressure. Equilibrium constants can be attained for any volatile compound under most conditions (22) because most compounds have a high affinity for the liquid phase and hence a low Henry's coefficient (ethanol 10^{-2} bar/L/mol) compared to gas. Camelbeeck et al. (29) have experimentally confirmed this conclusion by demonstrating that changes in the rates of agitation and aeration have no influence on the vent gas measurements.

The ability to monitor a specific organic compound depends on concentration and volatility and analyzer sensitivity. Oeggerli and Heinzle (22) compared partial pressures above an aqueous solution at a concentration of 0.1 g/L at 30 °C. Compounds were considered analyzable if the partial pressure was greater than 1 μbar; hence, ethanol at 22.24 μbar and acetone at 181 μbar could be monitored (oxygen, $>10^6$ μbar). However 2,3-mesobutanediol was considered undetectable as the partial pressure was 0.0174 μbar, and its limit of detection was found to be 5 g/L with a quadrupole mass spectrometer.

Calibration is usually achieved by adding a known amount of compound to the fermentor before inoculation and measuring concentrations from liquid samples off line by HPLC or GC. Most volatiles require a SEM for accurate detection as volatiles are in the ppm level in the headspace (22,29).

Growth Kinetics

Growth rate kinetics including yield of biomass from oxygen consumed, Y_{x/O_2}, from carbon dioxide evolution, Y_{x/CO_2}, and the specific growth rate (52,53) can be simply determined using vent gas analysis:

$$Y_{x/O_2} = \frac{d(\mathrm{TOC})/dt}{dx/dt} \qquad (15)$$

$$X = Y_{x/O_2}(\mathrm{TOC}) \qquad (16)$$

where TOC is the total oxygen consumed estimated from the summation of the OUR versus time and X is the biomass concentration from oxygen. Linear regression from a curve of biomass concentration (determined off line) versus TOC and total carbon dioxide evolved (TCE) can then lead to the respective yields factors for biomass from oxygen and carbon dioxide. The specific growth rate, μ, can be determined by

$$\mu = \frac{d(\mathrm{TOC})/dt}{\mathrm{TOC}} = \frac{\mathrm{OUR}}{\mathrm{TOC}} \qquad (17)$$

where the Y_{x/O_2} is assumed to be constant. Good agreement can be made between specific growth rates as calculated on line using TOC data and off-line data from experimental dry cell weight (53). This calculation provides a useful on-line tool for analyzing growth rate during process and development. This analysis is especially important for secondary metabolites in which the understanding of growth and production rates are critical to process optimization, often involving substrate feeding. For example, the control of specific growth is very important for the synthesis of avermectin by *Streptomyces avermitilis* (53). The impact of decrease of temperature by a few degrees at 12 h into the fermentation clearly showed a dramatic decrease in specific growth rate compared to the control, as measured by the on-line method (Fig. 8).

Mass Balancing and RQ

Vent gas analysis data can be used to describe the substrate utilization and production kinetics of a fermentation via mass balancing (2,34). A typical fermentation could involve $CH_{h_S}N_{n_S}O_{o_S} + aNH_3 + bO_2 \rightarrow Y_{X/S}CH_{h_X}N_{n_X}O_{o_X} + Y_{P/S}CH_{h_P}N_{n_P}O_{o_P} + cCO_2 + dH_2O$, where biomass $CH_{h_X}N_{n_X}O_{o_X}$ and product $CH_{h_P}N_{n_P}O_{o_P}$ are formed from substrate $CH_{h_S}N_{n_S}O_{o_S}$. The coefficients a, b, c, and d can be solved from elemental balances for carbon, hydrogen, nitrogen, and oxygen. Coefficients for oxygen and carbon dioxide are $b = (\gamma'_s - Y_{X/S}\gamma'_x - Y_{P/S}\gamma'_p)/4$ and $c = 1 - Y_{X/S} - Y_{P/S}$, where $\gamma'_i = 4 + h_i - 3n_i - 2o_i$ is the so-called generalized degree of reduction for compound

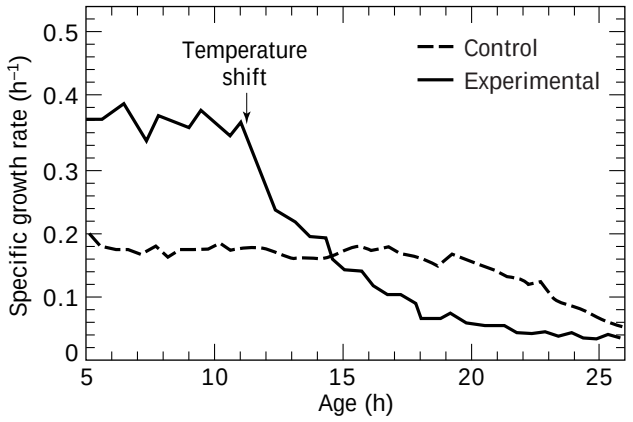


Figure 8. Impact of temperature shift on estimated specific growth rate (on line) of *Streptomyces avermitilis* at 800 L (53).

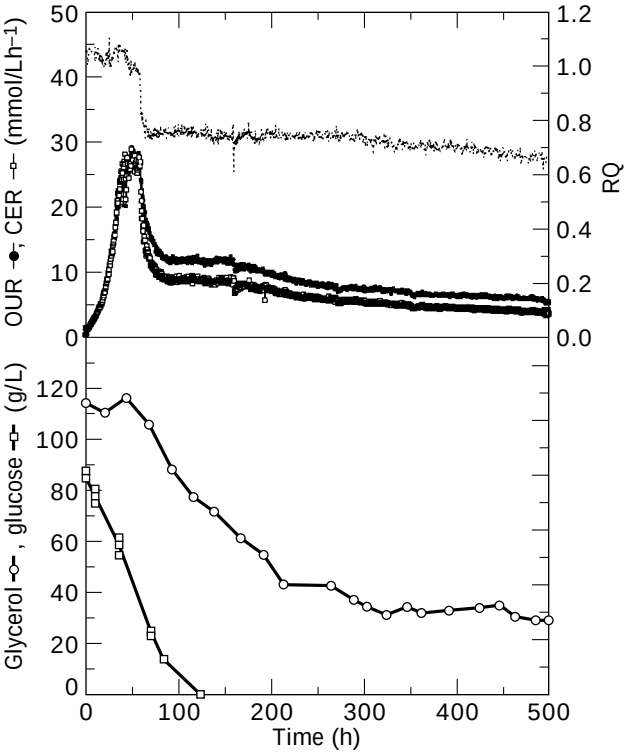


Figure 9. Comparison of vent gas analysis to off-line substrate data for a fungal fermentation at 1,600 L.

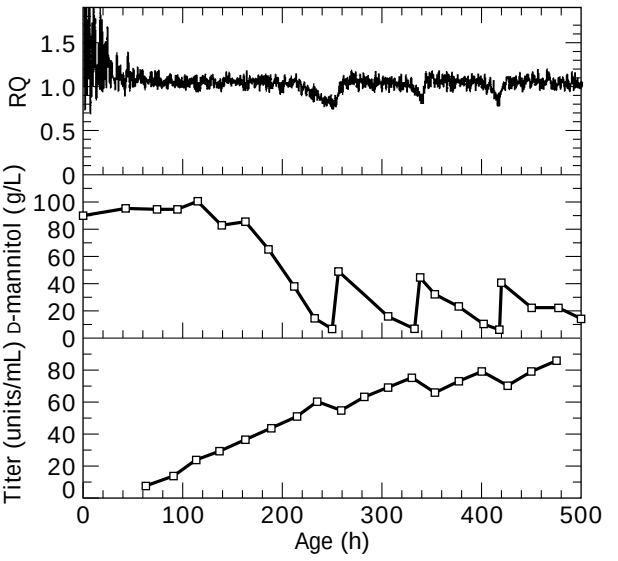


Figure 10. Use of on-line RQ measurements to trigger nutrient additions for secondary metabolite production.

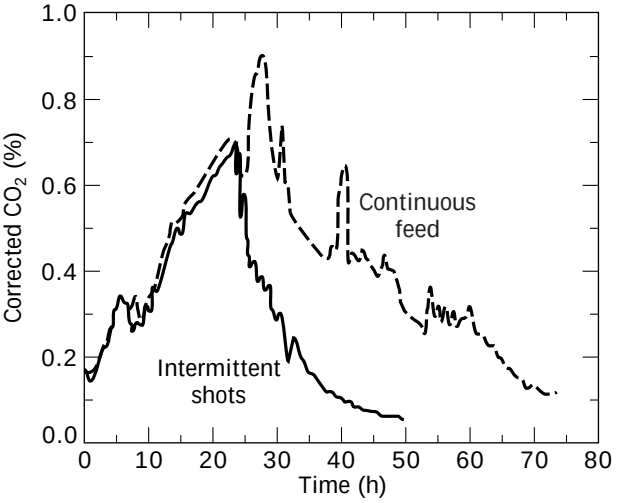


Figure 11. Comparison of the vent gas CO_2 concentration for *Nocardia* sp. between continuous and intermittent shot process for addition (54).

$CH_{h_i}N_{n_i}O_{o_i}$. The value of μ' for biomass is consistently about 4.2, and the values for some common substrates and products can be found in the literature (8). An expression can then be made for the respiratory quotient, RQ, the ratio between the rates of CO_2 evolution and oxygen utilization, as $RQ = 4(1 - Y_{X/S} - Y_{P/S})/(\gamma's - Y_{X/S}\,\gamma'_p)$. Thus, for a given fermentation the RQ, defined from vent gas analysis, CER/OUR, can be used with off-line analysis of biomass, carbon, nitrogen, and production concentrations to define the yield coefficients for substrate and production (32). Elemental balancing requires high-precision gas analysis with 0.01% accuracy because a 1% error relative offset calibration error for oxygen can result in an error for product estimation >50% (2).

For industrial applications, the unusual mutant strains have complex metabolism, of which little is understood. It is thus difficult to define stoichiometric equations for a process, limiting the use of elemental analysis. Nevertheless, vent gas analysis is a valuable tool for studying process kinetics and metabolism. For example, during a fungal fermentation the RQ was found to switch from unity to 0.7, indicating the change in metabolism from glucose to glycerol (Fig. 9). The reduction of growth rate, as shown by the OUR profile, and commencement of glycerol utilization indicated the start of secondary metabolite production. In

another fungal fermentation, the depletion of D-mannitol and subsequent reduction in volumetric productivity corresponded with the RQ reduction from 1 to 0.8 (Fig. 10). Shot additions of D-mannitol were made in response to the change in RQ, which prolonged the fermentation and increased production.

During the microbial hydroxylation of simvastatin (a cholesterol-lowering drug) by a *Nocardia* species, high concentrations of the substrate inhibit production (54). Maintaining a low concentration of the substrate using a controlled continuous feed strategy as opposed to a shot process enabled yield improvements of the biotransformation. The carbon dioxide profile was used to monitor process performance (Fig. 11) where the continuous substrate feed prolonged the decrease of carbon dioxide after 30 h compared to the rapid decrease observed with the shot processes. Therefore, combining vent gas data, including RQ, with off-line analysis of substrate utilization and productivity rates, along with an understanding of biochemical pathways, can enable process monitoring and feedback control.

BIBLIOGRAPHY

1. B. Buckland, T. Brix, H. Fastert, K. Gbewonyo, G. Hunt, and D. Jain, *Biotechnology* **3**, 982–988 (1985).
2. E. Heinzle and I.J. Dunn, in H.-J. Rehm, G. Reed, and K. Schugerl eds., *Biotechnology: Measuring, Modelling, and Control*, VCH Verlagsgesellschaft, Weinheim, 1991, pp. 27–74.
3. S. Aiba, S. Nagai, and Y. Nishizawa, *Biotechnol. Bioeng.* **18**, 1001–1006 (1976).
4. C. Fieschki, K.M. Egan, T. Ritch, R.A. Koski, and M. Jones, *Biotechnol. Bioeng.* **29**, 1113–1118 (1987).
5. J. Sevcik, *Detectors in Gas Chromatography*, Elsevier, New York, 1976, pp. 39–101.
6. G.D. Nichols, *On-Line Process Analyzers*, J. Wiley, New York, 1988, pp. 3–136.
7. Y.H. Lee and G.T. Tsao, *Adv. Biochem. Eng.* **9**, 30–42 (1979).
8. B. Atkinson and F. Mavituna, *Biochemical Engineering and Biotechnology Handbook*, Stockton Press, New York, 1991, pp. 1032–1038.
9. A.D. Brailsford, M. Yussouff, and E.M. Logothetis, *Sens. Actuators B* **44**, 321–326 (1997).
10. N. Hara and D.D. MacDonald, *J. Electrochem. Soc.* **144**, 4158–4160 (1997).
11. K.-D. Anders, W. Muller, R. Kammeyer, and T. Scheper, *Biotechnol. Tech.* **6**, 97–104 (1992).
12. H.M. Muhlemann and H.R. Bungay, *Biotechnol. Tech.* **7**, 575–578 (1993).
13. M.-N. Pons, *Bioprocess Monitoring and Control*, Carl Hanser Verlag, Munich, 1992, pp. 1–63.
14. A. Mandelis and C. Christofides, *Physics, Chemistry, and Technology of Solid State Gas Sensor Devices*, Wiley, New York, 1993, pp. 254–308.
15. Y. Taniguchi and K. Murakami, *Electron. Commun. Jpn.* **79**, 86–96 (1996).
16. C.C. Chang and C.S. Tang, *Sens. Actuators A* **65**, 171–174 (1998).
17. L. Rottenbacher, L. Behlau, and W. Bauer, *J. Biotechnol.* **2**, 137–147 (1985).
18. K. Schaefer and M. Schultis, in E. Henizle and M. Reuss eds., *Mass Spectrometry in Biotechnological Process Analysis and Control*, Plenum, New York, 1987, pp. 39–48.
19. A.T. Dadd, *Am. Biotechnol. Lab.* **8**, 35–38 (1990).
20. J.R. Swartz and C.L. Cooney, *Process Biochem.* **13**, 3–7 (1978).
21. S. Chauvatcharin, T. Seki, K. Fujiuyama, and T. Yoshida, *J. Ferment. Bioeng.* **79**, 264–269 (1995).
22. A. Oeggerli and E. Heinzle, *Biotechnol. Prog.* **10**, 284–290 (1994).
23. G. Suizdak, *Mass Spectrometry for Biotechnology*, Academic Press, New York, 1996, pp. 4–55.
24. H. Degn, R.P. Cox, and D. Lloyd, *Methods Biochem. Anal.* **31**, 165–194 (1985).
25. S. Bohatka, G. Langer, J. Szilagyi, and I. Berecz, *Int. J. Mass Spectrom. Ion Phys.* **48**, 277–280 (1983).
26. R.P. Cox, in E. Henizle and M. Reuss eds., *Mass Spectrometry in Biotechnological Process Analysis and Control*, Plenum, New York, 1987, pp. 63–74.
27. J.P. Camelbeeck, D.M. Comberbach, J. Goossens, and P. Roelants, *Biotechnol. Tech.* **2**, 183–188 (1988).
28. M.J. Winter, in E. Henizle and M. Reuss eds., *Mass Spectrometry in Biotechnological Process Analysis and Control*, Plenum, New York, 1987, pp. 17–38.
29. J.P. Camelbeeck, D.M. Comberbach, M. Orval, J.O. Petre, and P. Roelants, *Biotechnol. Tech.* **5**, 443–452 (1991).
30. E.R. Schmid, in E. Henizle and M. Reuss eds., *Mass Spectrometry in Biotechnological Process Analysis and Control*, Plenum, New York, 1987, pp. 7–16.
31. C.D. Bartman, in E. Henizle and M. Reuss eds., *Mass Spectrometry in Biotechnological Process Analysis and Control*, Plenum, New York, 1987, pp. 49–62.
32. E. Heinzle, *J. Biotechnol.* **25**, 81–114 (1992).
33. S.J. Coppella and P. Dhurjati, *Biotechnol. Bioeng.* **29**, 679–689 (1987).
34. P. Salmon and B.C. Buckland, in P.M. Rhodes and P.F. Stanbury eds., *Applied Microbial Physiology*, Oxford University Press, New York, 1997, pp. 131–163.
35. H. Heinzle, *Adv. Biochem. Eng. Biotechnol.* **35**, 1–46 (1987).
36. R.H. Perry, D.W. Green, and J.O. Maloney, *Chemical Engineers Handbook*, McGraw-Hill, New York, 1984.
37. S. Aiba and H. Furuse, *Biotechnol. Bioeng.* **36**, 534–538 (1990).
38. R.M. Dekkers, *Proceedings of the First IFAC Workshop on Modelling and Control of Biotechnical Processes*, Pergamon, Oxford, 1982.
39. A. Schumpe, G. Quicker, and W.D. Deckwer, *Adv. Biochem. Eng.* **24**, 1–38 (1982).
40. M. Popovic, H. Niebelschutz, and M. Reuss, *Eur. J. Appl. Microbiol.* **8**, 1–15 (1979).
41. J. Smith, S.W. Davison, and G.F. Payne, *Biotechnol. Bioeng.* **35**, 1088–1101 (1990).
42. P.N.C. Royce and N.F. Thornhill, *AIChE J.* **37**, 1680–1686 (1991).
43. H.S. Harned and B.B. Owen, *The Physical Chemistry of Electrolyte Solutions*, Reinhold, New York, 1958.
44. P.K. Yegneswaran, M.R. Gray, and B.G. Thompson, *Biotechnol. Bioeng.* **36**, 92–96 (1990).
45. E. Henizle, A. Oeggerli, and B. Dettwiler, *Anal. Chim. Acta* **238**, 101–115 (1990).
46. N.M.G. Oosterhuis and N.W.F. Kossen, *Biotechnol. Bioeng.* **26**, 546–550 (1984).

47. G. Larsson, M. Tornkvist, E. Stahl Wernersson, C. Tragardh, H. Noorman, and S.-O. Enfors, *Bioprocess Eng.* **14**, 281–289 (1996).
48. H.M. Ruffer, A. Petho, K. Schugerl, A. Lubbert, A. Ross, and W.D. Dekwer, *Bioprocess Eng.* **11**, 145–152 (1994).
49. A. Lubbert, in H.-J. Rehm, G. Reed, and K. Schugerl eds., *Biotechnology: Measuring, Modelling, and Control*, VCH Verlagsgesellschaft, Weinheim, 1991, pp. 107–148.
50. M. Reuss and R. Bajpai, in H.-J. Rehm, G. Reed, and K. Schugerl eds., *Biotechnology: Measuring, Modelling, and Control*, VCH Verlagsgesellschaft, Weinheim, 1991, pp. 299–348.
51. S. George, G. Larsson, K. Olsson, and S.-O. Enfors, *Bioprocess Eng.* **18**, 135–143 (1998).
52. D.W. Zabriskie and A.E. Humphrey, *AIChE J.* **24**, 138 (1978).
53. K. Gbewonyo, D. Jain, G. Hunt, S.W. Drew, and B.C. Buckland, *Biotechnol. Bioeng.* **34**, 234–241 (1989).
54. K. Gbewonyo, B.C. Buckland, and M.D. Lilly, *Biotechnol. Bioeng.* **37**, 1101–1107 (1991).

VINEGAR, ACETIC ACID PRODUCTION

HEINRICH EBNER
Linz, Austria
SYLVIA SELLMER-WILSBERG
Bad Honnef, Germany

KEY WORDS

Acetic acid bacteria
Bacteriophage
Cell recycling
Classification
Continuous submerged fermentation
Frings acetator
Semicontinuous submerged fermentation
Surface fermentation
Vinegar production

OUTLINE

INTRODUCTION

Worldwide, acetic acid is called vinegar, if obtained by oxidative fermentation of ethanol-containing solutions by acetic acid bacteria. Although vinegar does not entirely exclude diluted chemically produced acetic acid in every country of the world, this term is used here to define acetic acid that is produced from ethanol by a primary microbial metabolism, the so-called acetic acid fermentation or vinegar fermentation.

BASES OF ACETIC ACID FERMENTATION

Acetic acid fermentation (acetous fermentation) is an oxidative fermentation by which solutions of ethanol are oxidized to acetic acid and water by acetic acid bacteria using atmospheric oxygen. The oxidation proceeds according to the basic equation

$$C_2H_5OH + O_2 \rightarrow CH_3COOH + H_2O$$
$$G^\circ = -455 \text{ kJ mol}^{-1}$$

The alcohol-containing solution is called mash. Its alcohol concentration is given in percent by volume. Usually it also contains some acetic acid expressed in grams of acetic acid per 100 mL (% w/v). The sum of ethanol (vol%) and acetic acid (g 100 mL^{-1}) is called total concentration because the sum of these rather incommensurable values gives the maximal concentration of acetic acid that can be obtained by complete fermentation (1). The quotient of the total vinegar concentration produced over the total mash concentration indicates the concentration yield.

RAW MATERIALS FOR ACETIC ACID FERMENTATION

All mashes must contain ethanol and water and provide the acetic acid bacteria with nutrients.

Alcohol

By far the largest percentage of vinegar is alcohol vinegar, which is produced from diluted purified ethanol. Other common names for the same product are white vinegar or spirit vinegar. It is customary in almost all countries to denature the ethanol serving as raw material for the vinegar industry. In the United States, this is mostly done with ethyl acetate, and in most European Countries, with alcohol vinegar.

Mashes obtained by alcoholic fermentation of natural-sugar-containing liquids also serve as raw materials. The designation of vinegar is according to the particular raw material used: wine vinegar is produced by acetic acid fermentation of grape wine; cider vinegar is produced from fermented apple juice; malt vinegar is the undistilled product of alcoholic and subsequent acetous fermentation of an infusion of barley or other cereals, the starch of which was converted by malting; rice vinegar is made from saccharified rice starch, followed by alcoholic and acetous fermentation.

Water for Processing

The water used for the preparation of mashes must be clear, colorless, odorless, bacteriologically safe, and without sediments or suspended particles. It must also be free of chlorine, ozone, other chemical compounds, and damaging bacteria.

Nutrients

Most natural raw materials do not require the addition of extra nutrients. However, apple cider is usually low in nitrogenous compounds, but this can be corrected by the addition of ammonium phosphate. Some grape wines also require the addition of ammonium phosphate for an optimal fermentation. In rare cases, all nutrients discussed next must be added for the production of wine vinegar.

For the production of alcohol vinegar, a mixture of the required nutrients was developed. When producing vinegar with up to 15% acetic acid, the acetic acid bacteria definitely require glucose and potassium, sodium, magnesium, calcium, ammonium as ammonium phosphate, sulfate, and chloride. The following trace minerals are needed: iron, manganese, cobalt, copper, molybdenum, vanadium, and zinc. These additions are sufficient for an optimal acetic acid fermentation.

Commercially available mixtures contain supplements such as dried yeast extract in order to restart a fermentation more quickly if stopped by a disturbance such as a power failure. Principally, nutrients should be added sparingly in order to exert a selection pressure directing to a low requirement for nutrients.

PRODUCTION BY ACETIC ACID FERMENTATION

Microorganisms

Overview and Basic Problems of Classification. The microorganisms oxidizing ethanol to acetic acid are commonly called acetic acid bacteria. This special primary microbial metabolism at low pH of the surrounding medium differentiates them from all other bacteria. Acetic acid bacteria are polymorphous. Cells are Gram-negative, ellipsoidal to rod shaped, straight or slightly curved, 0.6 to 0.8 μm long, occurring singly, in pairs or in chains. There are nonmotile and motile forms with polar or peritrichous flagella. They are obligately aerobic. Some produce pigments, some cellulose.

Numerous attempts to classify acetic acid bacteria were carried out as summarized in the work by Ebner and Follman (1). A major contribution to the exact taxonomy of acetic acid bacteria were the DNA-rRNA hybridization studies of Gillis and De Ley (2). A first conclusion from this work was that the two groups *Gluconobacter* and *Acetobacter*, as classified earlier, were closely related groups, justifying their union in the family *Acetobacteraceae*. Although in 1984 a set of 32 features was given to differentiate the genera *Gluconobacter*, *Acetobacter*, and *Frateuria* (3), not all of them are necessary for a satisfactory identification.

The feature of acetate overoxidation identifies the strains used in high-acid fermentations as members of the genus *Acetobacter*, and thus separates them from the genera *Gluconobacter* and *Frateuria*.

Required Properties of Industrially Used Strains. Considering the interest of many microbiologists in acetic acid bacteria, it is surprising that the available information about these bacteria has had so little effect on vinegar production. Working strains were isolated, but lost again, demonstrating the difficulties to isolate and grow *Acetobacter* strains from industrial vinegar fermentations on solid media. Considerable progress was achieved by propagating bacteria from fermenters in Japan on a special double-layer agar. The bacteria isolated on this medium have been described as *Acetobacter polyoxogenes* (4). However, *A. polyoxogenes* was not available from the Japan Collection of Microorganisms because of problems of propagation and preservation.

A new species in the genus *Acetobacter*, for which the name *Acetobacter europaeus* was proposed (5), has been isolated and characterized in pure culture from vinegar fermentations at high acidity in Germany and Switzerland. All investigated strains isolated from submerged fermenters and trickle generators had very low (0 to 22%) DNA–DNA similarities with the traditional strains of *Acetobacter* and *Gluconobacter*. The phenotypical differentiation on the species level of the genus *Acetobacter* is rather difficult, because different strains of a single species do not necessarily utilize the same carbon source. A useful and significant criterion for the distinction of *Acetobacter europaeus* from the other *Acetobacter* species is the strong tolerance to acetic acid of 4 to 8% in AE-Agar and the absolute requirement of acetic acid for growth.

Independent of all these problems and difficulties, the first and foremost interest of the vinegar industry is to use a strain of acetic acid bacteria that tolerates high concentrations of acetic acid and total concentration, requires small amounts of nutrients, does not overoxidize the acetic acid formed, and yields high production rates. At present, it also seems to be necessary to make this strain phage resistant (6).

Industrially Used Strains. The vinegar industry has always worked with acetic acid bacteria that, in most cases, were not derived from pure cultures. The striking fact that common microorganisms used on a large scale for industrial vinegar production were not properly described and characterized in taxonomic terms is caused by the difficulty to cultivate them on semisolid media.

Industrial submerged vinegar fermentations are started by inoculation with inoculation vinegar, microbiologically undefined remains from previous fermentations. Acetic acid fermentation can be carried out for years without interruption or decrease in efficiency or yield, if suitable conditions are chosen to allow a permanent selection of strains tolerating high acidity at a minimum of nutrient concentration. Such cultures are delivered to vinegar factories in all parts of the world by Heinrich Frings GmbH & Co KG, Bonn, Germany.

Plasmid Profile Analysis for Characterization of Acetobacter Strains. Plasmids are circular, extrachromosomal DNA molecules that can be extracted from bacteria by a sophis-

ticated technique. The plasmid profile of a strain is very specific. Plasmids can determine certain properties of bacteria, such as phage resistance or citrate metabolism. The potential of this technique for vinegar fermentations was investigated (7–9). The majority of the acetic acid bacteria contain 1 to 8 plasmids from 1 to more than 17 MDa. Plasmid profile analysis has been applied to characterize the *Acetobacter europaeus* strains and the microflora of industrial vinegar production plants (10). It was shown that the microflora in submerged fermentations consists of one predominant strain, whereas in trickling generators the microflora is quite heterogenous, containing a mixture of several *Acetobacter europaeus* strains with quite different plasmid profiles. This is probably the reason for the susceptability of submerged fermentations to bacteriophage attack. In conclusion, the determination of plasmid profiles from submerged vinegar fermenters represents a new and powerful technique to characterize the microflora of such fermentations and to answer questions that could not have been approached until now, such as the stability, origin, and identity of the prevailing microflora.

Recent investigations have shown that the plasmid profiles of these strains are able to rest stable over a period of at least 5 years and are not changing their phenotypic features during this time. Complete sequencing of the 16S rDNA genes have shown the phylogenetic position of *Acetobacter europaeus* to be close to *Acetobacter xylinum* and well separated from seven other *Acetobacter* species. Acetic acid bacteria together with *Acidiphilium* species and *Rodopila globiformis* represent a distinct acidophilic branch in the α-subclass of *Proteobacteria*.

Electrophoretic protein patterns correlated with the percentage of DNA similarity of DNA–DNA hybridization proved to be very useful for the delineation of the strains from industrial vinegar fermentations at the species/subspecies level (11,12).

Genetic Improvement of *Acetobacter* strains. To improve strains of *Acetobacter* and *Gluconobacter* genetically, recombinant DNA techniques are considered to be useful. Host–vector systems and an efficient transformation method have been developed. The preparation of spheroplasts and the development of a spheroplast fusion method have been described (13). A strain growing at 37 °C and 4% acetic acid and a strain unable to grow at 35 °C and 5% acetic acid were spheroplasted and fused. Fusion products were received growing at 37 °C on agar plates containing 5% acetic acid, on which both parent strains were unable to grow. Some more genetic transformation systems have been described (14–16).

When summarizing the present technique, Fukaya (17) states that recombinant DNA technology can contribute to advances in the basic knowledge of acetic acid bacteria and additionally provide new tools for improving them. But to apply recombinant DNA technology to improving acetic acid bacteria even more, basic research concerning gene expression, protein engineering, and other fundamental studies are required.

Bacteriophages. *Acetobacter*-specific bacteriophages demonstrated by electron microscopy in submerged and trickling fermenters are said to be responsible for fermentation problems in industrial vinegar fermenters (18–21). In trickling generators, fermentation is sometimes slowed down, but hardly stops completely. In submerged fermentations, complete loss of productivity has been observed. Only three morphologically different *Acetobacter* phage types were described until 1992, when a great variety in phage types, regarding the dimensions of the phage heads and tails was reported (20). All phages isolated from vinegar fermentations in Germany, Austria, and Denmark showed long contractile tails and isometric heads belonging to Bradley's group A (22) and the family of *Myoviridae* (23), respectively. Although not all phages described have actually been shown to be infective for defined *Acetobacter* strains because of a lack of a suitable technique, their occurrence in high numbers (10^6 to 10^9 mL^{-1}) in disturbed fermentations suggests that bacteriophages may actually be the source of fermentation problems.

Biochemistry

During acetic acid fermentation, ethanol is almost quantitatively oxidized to acetic acid. Concentration yields between 95 and 98% are common, and the remainder is mainly lost in the effluent gas. At the same time, a suitable C_6-carbon source (preferably glucose) is oxidized. End products of this oxidation are CO_2 and H_2O.

Ethanol. Ethanol oxidation is performed by two sequential reactions catalyzed by alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH), located at the outer surface of the cytoplasmic membrane. Their function is linked to the respiratory chain of the organism (24,25).

Alcohol dehydrogenase has been purified from several acetic acid bacteria and shown to contain pyrroloquinoline quinone (PQQ) and heme C as prosthetic groups (26–28). Extensive purification and characterization of the enzymes involved in ethanol oxidation of *Acetobacter aceti*, *A. rancens*, and *G. suboxydans* have been reported (26–32). It has been shown that ADH of acetic acid bacteria donates electrons directly to ubichinone in the cytoplasmic membranes. Thus, the ethanol oxidase respiratory chain of acetic acid bacteria is made of only three membranous respiratory components, alcohol dehydrogenase, ubichinone, and terminal oxidase (33).

It has been confirmed that ADH and ALDH of acetic acid bacteria had the same prosthetic group, PQQ, as their glucose dehydrogenase (30). These quinoproteins, glucose, ADH, and ALDH are involved in sugar oxidation that links them to an electron transport system in the membrane of oxidative bacteria. ADH and ALDH from *A. polyoxogenes* sp. nov. have been purified (34,35). This strain originated from a vinegar factory showing higher acid productivity and higher tolerance to acetic acid than strains isolated from natural sources. In contrast to the ADH and ALDH obtained from *A. aceti* and *G. suboxydans*, the enzymes of *A. polyoxogenes* were quite stable, and this high stability seems at least to be partially involved in the ability of *A. polyoxogenes* to produce high concentrations of acetic acid.

Sugar. For the breakdown of sugars, *Acetobacter* is equipped with the hexose monophosphate (HMP) pathway

and the TCA cycle (36), whereas glycolysis is either absent or very weak. The enzymes of the Entner-Doudoroff pathway are present in *Gluconobacter* and in *Acetobacter xylinum*. *Gluconobacter* and *Acetobacter* are known for their direct oxidative capacity on sugars, alcohols, and steroids (37). For further details see the work by Ebner and Follman (1). The biochemistry of the ketogenic activities of acetic acid bacteria have been reviewed (38).

Acetate. Strains of *Acetobacter* are able to oxidize acetate as well as lactate either in the presence or in the absence of ethanol by using enzymes of the tricarboxylic cycle.

Carbon Dioxide. *Acetobacter* needs CO_2 for growth (39). CO_2 is incorporated into cell substances, where approximately 0.1% of the cell carbon is derived from CO_2. A very small but measurable portion of acetic acid too is derived from CO_2 (40).

Nitrogen. If an organic carbon source is present, a number of acetic acid bacteria can use ammonium as sole nitrogen source. Some strains need the presence of amino acids; others need cofactors such as vitamins or purines (41). No essential amino acids are known.

Growth Factors. Growth factor requirement depends on the carbon source supplied (42). Some strains require *p*-aminobenzoic acid, niacin, thiamin, or pantothenic acid as growth factors (43). Some *Acetobacter* strains do not need vitamins in the presence of glucose (44). But the combined addition of glutathione with Na glutamate has a cumulative effect on the growth of *A. aceti* (45).

PQQ, the novel prosthetic group in acetic acid bacteria, also shows an interesting growth-stimulating effect for various microorganisms (45–47). One of two PQQ effects was the stimulation of both the growth rate and the total cell yield with only a trace amount of PQQ as an essential growth factor. The second effect of PQQ observed universally was a marked reduction of the lag phase in microbial growth, but no increase of the growth rate in the subsequent exponential phase or in the total cell yield at the stationary phase.

Physiology

Oxygen Demand and Total Concentration. During commercial use of many aerobic microorganisms it is common practice to permit the inoculum to remain in a fermenter for some time up to several days without aeration before inoculation into the main fermenter. However, application of this procedure is impossible in the case of acetic acid fermentation.

Extensive tests have been conducted (48,49) concerning the damage to acetic acid bacteria caused by an interruption of aeration as a function of total concentration, acetic acid concentration, rate of fermentation, and length of interruption of aeration. The longer the interruption of aeration and the higher the total concentration of the mash, the more marked is the effect. At a total concentration of 5%, an interruption of aeration for 2 to 8 min leads to the same damage as an interruption for 15 to 60 s at total concentrations of 11 to 12%. At constant total concentrations, damage increases with increasing concentration of acetic acid and with increasing fermentation rate. The specific production of cell material is extremely low (1).

Experiments conducted later (50) confirmed these results. It was found that there was a sharp drop in the oxidation rate of ethanol as well as in the activity of enzymes involved (ADH and ALDH) after an interruption of oxygen supply, with an acetic acid concentration of 6%. ADH activity in the membrane fraction was highly stable in the culture broth containing 4% or less acetic acid. But the purified enzyme was very unstable even in the culture broth of 2% total concentration. It was suggested from this fact that the ADH has a membrane component contributing to its own stability at acetic acid concentrations of less than 4%, which becomes gradually liable to damage with further increasing acidity. But, the relationship between the ADH system and oxygen has not been elucidated at all.

Under conditions of industrial acetic acid production, the microorganisms undergo considerable stress caused by a high concentration of acetic acid and limited oxygen supply. A high utilization of oxygen up to 70% can be achieved in industrial production. A sparing use of air is very important because of the volatility of ethanol and acetic acid. However, the use of pure oxygen or highly oxygen-enriched air may easily damage the bacteria (48).

In 1985, it was found that at a total concentration of 13%, the ATP pool and the growth rate show a reverse behavior (51). The energy charge (average value of intracellular ATP, ADP, and AMP) showed a rather high value of 0.84 at the beginning of the fermentation. During fermentation, although the acetic acid concentration increased from 9 to 12%, the energy charge decreased to 0.7. After interruption of aeration for 45 s, the energy charge dropped to 0.58, and after several weeks of storage the charge of the inoculum was only 0.50.

Lack of Ethanol. Acetic acid bacteria are also damaged if a vinegar fermentation is carried out to the point where all the ethanol has been oxidized and the addition of fresh ethanol-containing mash is delayed. This is analogous to the damage resulting from an interruption of oxygen supply and also depends primarily on the total concentration and the duration of ethanol lack.

Specific Growth Rate. In 1953, the specific growth rate of acetic acid bacteria was calculated for the first time (52). In semicontinuous fermentations there was no decrease of the specific growth rate at increasing acetic acid concentrations, but at increasing total concentrations the specific growth rate decreased from 0.49 h^{-1} at 5% to 0.16 h^{-1} at 11%.

In continuous culture at a total concentration of 12%, the specific growth rate decreased from 0.027 h^{-1} at 4.5% ethanol to 0.006 h^{-1} at 1% ethanol. It should be mentioned that the specific growth rate obtained in continuous fermentations with constant concentrations of acetic acid and ethanol are less favorable than those obtained under semicontinuous conditions. Batch fermentations carried out with *Acetobacter suboxydans* (53) showed at a total con-

centration of 7.5% a specific growth rate of 0.30 h^{-1}, a result that is in good agreement with other studies (51).

Specific Product Formation. In 1949, a specific product formation was found of 21 g of acetic acid per gram of bacterial dry substance per hour in semicontinuous tests at a total concentration of 10% (54). This very high value was independent of the acetic acid concentration of 4.5 to 7.2%. In 1972, a strong dependence between specific product formation and acetic acid concentration in continuous fermentations at a total concentration of 7% was found (55). The maximum specific product formation was 15 $g\,g^{-1}\,h^{-1}$. The specific product formation follows a growth-associated mechanism as long as the acetic acid concentration is below 3%. At higher acetic acid concentrations up to 7%, an additional nongrowth-associated term must be considered. The highest value obtained with cell recycle in continuous culture of *A. suboxydans* was 11.5 $g\,g^{-1}\,h^{-1}$ (53).

Changes in Concentration and Temperature. Charging of a fermentation must be done with mash at the same total concentration and under rapid mixing because a locally formed strong concentration gradient damages the bacteria. The next cycle starts without any lag phase. At the same time, temperature has to be kept constant because a repeated change in temperature causes a similar result (1).

Overoxidation. Overoxidation is the undesirable oxidation of acetic acid to CO_2 and H_2O that has to be prevented by avoiding lack of ethanol and keeping the total concentration at a high level. In trickle fermenters, this is much more difficult than in submerged fermenters (1).

Industrial Processes

Submerged Vinegar Fermentation ***Semicontinuous and Continuous Fermentations.*** For the production of vinegar with more than 12% and up to 15% acetic acid (56), the process has to be carried out in a semicontinuous manner; each fermentation cycle taking about the same time as the preceding and following cycles. The starting concentration of each cycle is 7 to 10% acetic acid and about 5% ethanol. When an alcohol concentration between 0.05 and 0.3% has been achieved in the fermentation liquid, a quantity of vinegar is discharged from the fermenter. Refilling with new mash of 0 to 2% acetic acid and 12 to 15% alcohol leads to the starting concentrations for the new cycle mentioned earlier. Discharging must be carried out quickly in order to avoid complete alcohol depletion. Charging must be done slowly under constant fermentation temperature and rapid mixing.

Since the middle of 1994, it has been possible to produce vinegar of up to 19% acetic acid in a modified single-stage process. Starting with similar concentrations as mentioned earlier, alcohol is slowly added at a constant alcohol level of 2 to 3% in the fermentation liquid, creating a corresponding increase of the total concentration. Addition of alcohol is stopped when the desired total concentration has been reached. When alcohol approaches zero, a part of the fermentation liquid is discharged and replaced by mash with lower total concentration to bring the acetic acid concentration and total concentration back to starting conditions, thereby enabling the bacteria to multiply faster. Later, total concentration is increased again by adding alcohol.

Continuous fermentation is only possible up to a maximum of 9 to 10% acetic acid because the specific growth rate of the bacteria decreases with decreasing ethanol concentration. To obtain high yields, the fermentation must be carried out at a low alcohol concentration.

Two-stage Processes. In the canning industry, vinegar with a very high percentage of acetic acid is in demand. The vinegar industry is interested in producing vinegar of high acidity in order to save storage and transport costs. Therefore, a two-stage process for the production of vinegar with more than 15% acetic acid was developed (57). During a submerged fermentation with a total concentration below 15% in a first fermenter, alcohol is added slowly to increase the total concentration up to about 18.5%. After the acetic acid concentration has reached 15%, about 30% of the fermenting liquid is transferred into a second fermenter. The first fermenter is resupplied with new mash of lower total concentration. In the second fermenter, the fermentation continues until the alcohol is almost depleted. The whole quantity of the finished vinegar is discharged. The fermentation liquid in the first fermenter is supplied with alcohol at the appropriate time and later divided again.

Since the end of 1993, vinegar of more than 20% acetic acid has been produced using this process. An automatic control for this process has been described (58).

In 1981, a similar two-stage process for the production of vinegar with more than 20% acetic acid was described (59,60). The second fermentation stage is carried out at a reduced temperature of only 18 to 24 °C, whereas in the first fermentation stage, the normal temperature range of 27 to 32 °C is used.

In 1984, another two-stage process was claimed (61) with semicontinuous fermentation in at least two fermenters, membrane filtration of the mash from the semicontinuous stage, using lower temperatures at higher acetic acid concentrations, and with the membrane-filtered vinegar to be used as a second end product.

The Frings Acetator®. Frings Acetator® is the most common equipment for the production of all kinds of vinegar. At the end of 1996, 642 Acetators, continuously improved, enlarged and automated, were in operation all over the world with a total production of $1{,}524 \times 10^6$ L/year of vinegar with 10% acetic acid. Energy consumption is about 400 W L^{-1} of fermented ethanol with decreasing tendency; yield is up to 98%.

Commercial sizes of Acetators are for the fermentation of up to 3,600 L of pure ethanol in 24 h.

The different processes demand excellent performance of the fermenter, especially of its aeration system. The Frings aerator (62), as shown in Figure 1, is self-aspirating; no compressed air is needed. The hollow rotor is installed on the shaft of a motor mounted under the fermenter, connected to an air suction pipe and surrounded by a stator. It pumps liquid that enters the rotor from above outward through the channels of the stator that are formed by the wedges, thereby sucking air through the

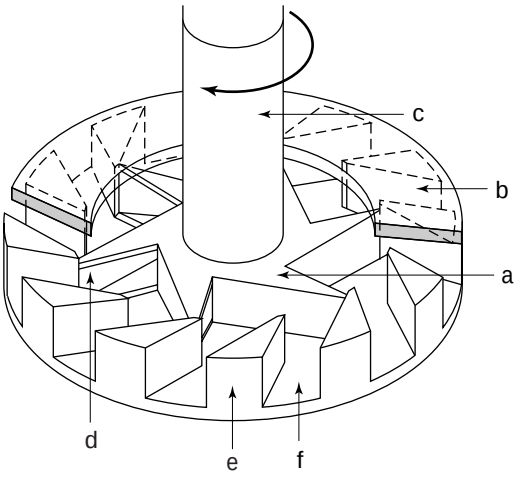


Figure 1. The Frings Aerator. a = hollow rotor; b = stator; c = air suction pipe; d = openings for air exit; e = wedges to form the channels; f = channels to form the beams of air–liquid emulsion.

openings of the rotor and creating an air–liquid emulsion that is ejected outward at a given speed. This speed must be chosen adequately so that the turbulence of the stream causes a uniform distribution of the air over the whole cross-*section of the fermenter.

Figure 2 shows this situation in a fermenter containing for demonstration a small amount of water immediately after the start of the aeration. Air is aspirated through the air suction pipe from the top, and the formed emulsion is spread out in beams at the bottom of the fermenter where they reach the wall of the fermenter somewhat outside of the cooling coils. Seconds later, the uniform aeration of the liquid takes place.

Because foam formation cannot always be avoided, the Acetator is equipped with a mechanical defoamer. From 1966 to 1991, the following equipment (63) was used. The rotor, turning in a spiral housing, separates liquid and gas and returns the liquid into the lower part of the fermenter. Only during discharging of the finished vinegar a part of this returned liquid is pumped out of the fermenter. Because this liquid causes foam formation, the defoamer must operate most of the time of a fermentation cycle. In an improved version (64), the separated liquid portion is not pumped back into the fermenter, but into a small collection vessel to be mixed with the vinegar end product later. Lysis of harmed vinegar bacteria releases foam-causing surface-active substances. By means of the total elimination of the separated liquid, further foam formation is avoided. Minimized foam formation with a shorter operating time of the defoamer and consequently a lower power uptake and a negligibly small volume of the separated liquid are the result.

The Frings Alkograph® (65), an automatic instrument for measuring the amount of ethanol in the fermenting liquid, has been in use since 1965. Small amounts of liquid flow through the analyzer continuously, at first through a heating vessel and then through three boiling vessels. The boiling temperature of the incoming liquid is measured in the first boiling vessel. While alcohol is distilled off continuously from the second and the third boiling vessel, the higher boiling point of the liquid from which ethanol has been removed is measured in the third boiling vessel. The difference in temperature corresponds to the ethanol content and is recorded automatically. Because the flow through the vessels takes some time, there is a delay of about 15 min between the beginning of the inflow of a sample and the appearance of the correct value on the recorder. Because alcohol concentration is decreasing slowly during fermentation, this delay has no disadvantage on fermentation control.

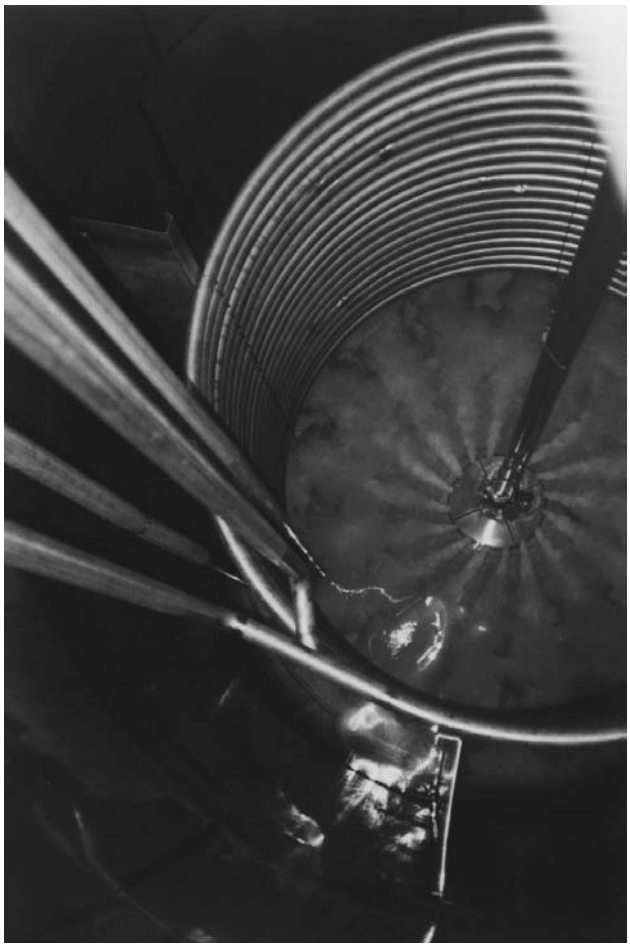


Figure 2. How the Frings Aeration System works. Acetator filled for demonstration with a small amount of water immediately after the start of the aeration. Air is aspirated through the suction pipe from the top and the formed emulsion is spread out in beams at the bottom of the fermenter, reaching the wall of the fermenter somewhat outside of the cooling coil.

To carry out the single-stage semicontinuous process at a defined alcohol content, a contact in the Alkograph activates the vinegar discharge pump. As soon as a preset level has been reached, the mash pump starts adding fresh mash. This pump is controlled by the fermentation temperature in order to refill under constant temperature. The pump is stopped when the desired level is reached and an automatic cooling system is activated. A fermentation cycle takes 24 to 48 h.

The Frings Alkograph® has worked with high reliability for about 30 years in 231 Acetators around the world. New biosensors together with modern membranes permitted the development of a device to measure alcohol without any time delay; the Frings Alkosens®. A device for measuring volatile constituents of the fermentation medium was described (66). A membrane serves to filter these volatile constituents from the fermentation liquid into a carrier medium, and they are recorded by an electronic sensor. Because the measurement relies on diffusion, ambient temperature has to be kept always at the same level. Thus, the alcohol probe is situated in the fermentation broth. The alcohol value displayed at the analyzer is influenced by the concentration of acetic acid, and, therefore, new calibrations are needed at different total concentration levels. This disadvantage has been overcome recently by a new system, which incorporates a multilayer membrane. New sensor technologies and more sophisticated computer equipment permit a fully automated process control of continuous, semicontinuous as well as single-stage or two-stage high-strength fermentations—preparation of mash included.

Further Trends in Submerged Processes. A computerized experimental laboratory system was applied 1992 (67) to submerged acetic acid fermentation with industrial *Acetobacter* strains at about 10% of final acidity using automated semicontinuous cycles during a period of one year. Based on these data, a mathematical model was developed to predict the time course of a cycle, even in the case when at the beginning of the cycle a lag phase occurred because of a lack of substrate at the end of the foregoing cycle. The parameters substrate, product, and biomass, however, were not sufficient to characterize the physiological status of the culture, since the behavior of the microorganisms also depends on their past. RNA concentration per unit of cell mass was found to be an internal key parameter that showed a reasonable agreement of the experimental data and the mathematical model.

In 1989, the possibilities of on-line measurement of alcohol concentration in vinegar production were summarized (68). In addition to the Frings Alkograph® the Alcolyzer®, a Japanese instrument based on the same principle, immobilized microbial sensors, and semiconductor alcohol sensors were described.

Experiments for the continuous production of acetic acid by submerged fermentation with continuous removal of acetic acid by electrodialysis, continuous supply of ethanol and concentrated nutrient solution, and control of ethanol and pH were reported (69–71). The reason for the continuous removal of acetic acid from the reaction mixture was to alleviate the inhibitory effect of the produced acetic acid. Ethanol concentration was kept constant at 1%, acetic acid concentration at approximately 2%. Without dialysis and pH control, acetic acid production already stopped after 2 days at approximately 4%, whereas with dialysis it lasted 30 days.

Other Submerged Fermentation Processes. Other processes of little practical significance and abandoned processes have been described (1).

Continuous Culture and Cell Recycle. Much work was invested (53,72) to increase the productivity of the acetic acid fermentation by continuous culture and cell recycle.

It must indeed be admitted that at first sight the productivity in an Acetator does not seem to be very high. The Acetator 3600® produces during 24 h 4,000 kg acetic acid with a concentration of up to 15% in a fermenting volume of 100 m^3, that is, 1.6 g L^{-1} h^{-1} acetic acid in average or about 2.4 g L^{-1} h^{-1} at the maximum fermentation rate. The aeration rate is 0.1 vvm only; the nutrient consumption is the same as described in Physiology.

Kinetic studies (72) clarified the dependence of the bacterial growth rate on ethanol concentration, on acetic acid concentration, and on concentration of dissolved oxygen. The optimum ethanol concentration is 23 g L^{-1}, the optimum acetic acid concentration is 36 g L^{-1}, and the optimum oxygen concentration is 2.5 − 3 mg L^{-1}. The productivity of acetic acid formation mainly depends on the growth rate but also on the cell concentration. Therefore, the maximum productivity of 6 to 6.6 g L^{-1} h^{-1} acetic acid can be obtained at 32 g L^{-1} ethanol, and 60 to 66 g L^{-1} acetic acid.

The synergistic effects of acetic acid and ethanol on the growth of *Acetobacter* sp. at low concentrations showed similar results (73).

At first sight, the potential productivity is about 4 times higher than presently achievable and could even be further increased by recycling of the bacteria. However, this high degree of productivity can only be achieved at a narrow concentration range. A slight change of ethanol, acetic acid, or oxygen concentration results in an immediate return to lower values.

For technical applications, a second fermenter in series to the fermenter, operating at optimal fermentation conditions, would be necessary to complete the fermentation. To keep a high oxygen concentration in the fermenting liquid at the higher fermentation rate, a high aeration rate would be necessary with recovery of ethanol and acetic acid from the exhaust gas to obtain high yields, and probably addition of oxygen would be necessary. Higher nutrient consumption as well as increasing foam formation would result. The cell recycle system would need a continuous microfiltration avoiding a lack of oxygen and require an automatic control of cell concentration. But all the expenditure in process control only allows the production of the same amount of vinegar, containing not more than 10% acetic acid, in a smaller fermenter. Apparently, the demand for high productivity does not necessarily indicate the optimal developments for future production.

An attempt was made to enhance continuous acetic acid production in a laboratory fermenter equipped with a hollow fiber filter module by increasing the concentration of dissolved oxygen (74,75). The maximum acetic acid concentration achieved was 5%. The highest productivity obtained with yeast extract as nutrient was 107 g L^{-1} h^{-1}. The concentration of dissolved oxygen was already lower than the critical concentration of 1 ppm, although pure oxygen was supplied at an aeration rate of 1 vvm. Furthermore, it was difficult to maintain a high concentration of viable cells in the cell recycle culture, because the cells were inactivated by the lethal conditions of acetic acid concentration, oxygen deficiency, and nutritional limitation.

The effect of dissolved oxygen and acetic acid concentration on acetic acid production in continuous culture has

extensively been investigated (75). The optimum concentration of dissolved oxygen was found to be about 2 ppm for acetic acid production at low acidity. The specific acetic acid production rate, however, was diminished to complete inhibition when the dissolved oxygen concentration was higher than 2 ppm at an acetic acid concentration of more than 4.5%. This is in good agreement with older findings (48), stating that the use of pure oxygen during fermentation can completely inhibit the acetic acid formation.

Because viability of the bacteria is important for high productivity, the factors affecting viability have been investigated (76). To study the effect of cell bleeding (withdrawing at constant flow rate), serial experiments were performed in a bioreactor with cell recycle and were supported by mathematical analysis. Acetic acid concentrations were up to 5.6%. The maximum value of productivity was 123 g L^{-1} h^{-1}. The bleed ratio and dilution rate at which concentration of acetic acid and viable cells are relatively high can be predicted now.

The authors expect these results to be useful for designing a two-stage culture for acetic acid production using the active cells in the bleed from the first stage in a second-stage vessel beyond a lethal acetic acid concentration.

How far these certainly interesting results will influence the construction of industrial equipment in Japan remains to be seen. The costs of such complicated fermenters may slow down or even prevent a fast conversion to industrial scale, although in Japan the demand for vinegar with 5% acetic acid is relatively high. However, these results will certainly have no influence on the production of vinegar with more than 10% acetic acid. A review of the Japanese development work (77) states that the problem of scale-up remains for further investigation.

Semicontinuous Culture with Cell Recycle. During investigation of semicontinuous cultures with cell recycle, an increase was found in the productivity compared to semicontinuous culture without cell recycle at an acidity of up to 6% (78). Although the concentration of total cells increased in every cycle, the viable cells reached a much lower quasi steady-state value, which was explained by product inhibition. The productivity was 2.9 g L^{-1} h^{-1} without and 5.0 g L^{-1} h^{-1} with cell recycle at 6% final acetic acid concentration, but 14 g L^{-1} h^{-1} at 3% final acetic acid concentration.

In another study (79), semicontinuous fermentations were reported in which the ethanol concentration was kept for a certain time between 20 and 30 g L^{-1} by adding ethanol-enriched medium. At the end of the cycle, ethanol-poor medium was added. Cells were recycled by membrane filtration. The goal was to obtain high concentrations of acetic acid; the highest amount reached was 9%. It was found that the *Acetobacter* cells lost their viability at about 6% acetic acid, but these nonviable cells were still effective for ethanol oxidation. This interesting feature of *Acetobacter* cells had already been reported much earlier (80).

Surface Fermentation. ***Old Surface Fermentation Processes.*** Detailed descriptions of the history of vinegar fermentation were published in 1976 (81). Today, the equipment used in the famous Orleans process for the production of wine vinegar can only be seen in museums.

New Experiments with Surface Fermentation. A new approach to surface fermentation has been tried (82). The authors used a bioreactor with a shallow, horizontal flow of the medium under a bacterial film with a surface of a few hundred cm^2. Liquid depth was up to 10 mm, and the acetic acid concentration of the effluent was 5.76%. The oxygen absorption rate through the microbial film was found to be very high; a kinetic study of acetic acid production was published (83).

A surface fermentation process in a system of serial intercommunicating fermentation vessels, with flow-regulating devices that enforce the flow to approach plug flow, was described (84). The liquid depth under the layer of acetic acid bacteria was about 50 mm. The inflowing liquid had 2% acidity and contained 3.5% alcohol. The acidity of the effluent stabilized at 5% acetic acid. The residence time was about 21 h.

Investigations on continuous surface fermentations in a single vessel with a working volume of 16 L were reported (84). Under optimum conditions with a liquid depth of 100 mm, 2 to 3% acidity and 20 to 30 g L^{-1} h^{-1} alcohol concentration of the inflowing liquid, vinegar with 4.5% acetic acid was produced during 80 days. The flavor of this vinegar was claimed to be better than that of vinegar produced by other processes.

As a consequence of the efforts to produce rice vinegar at low acidity with better quality, surface fermentation is developing in Japan into an automatic fermentation process at lower costs.

Trickling Processes. The old semicontinuous generator process used beech wood shavings, birch twigs, or corn cobs as carrier material for the vinegar bacteria. Large wooden tanks with up to 100 m^3 shavings are still in use. For more details refer to the work by Ebner and Follman (1). Despite the disadvantages of the trickling process, the use of a silicate carrier material with a large specific inner surface and an open pore structure was proposed recently (86).

Experiments with Immobilized Acetic Acid Bacteria. Many Japanese scientists investigated the fermentation of rice or fruit vinegar with the usual concentration of about 5% acetic acid in continuous fermentation of high productivity with immobilized acetic acid bacteria.

A 460-day run in an airlift bioreactor with acetic acid bacteria immobilized on κ-carrageenan gel beads, where living cells were newly released into the reactor, was reported (68). These newly released bacteria showed extremely high growth rates and respiratory activity that they retained for a few generations after leakage.

Polypropylene fibers as carrier material for *Acetobacter* producing rice vinegar of 3% acetic acid were proposed (87,88). With ceramic honeycomb-monolith as carrier in a fixed-bed reactor a maximum of 4% acetic acid was achieved (89). Calcium alginate beads in a fluidized bed reactor were investigated and 3.5% acetic acid obtained (90). The same author (91,92) used this reactor in combination with an ethanol fermenter to produce rice vinegar of up to 6.5% acetic acid from a saccharified rice medium.

A two-column bioreactor of porous ceramics was tested to produce vinegar of an orange wine with up to 8.4% acetic

acid (93). If the cells were fixed on woven cotton fabric, vinegar of kiwi fruit and persimmon of 4.5% acetic acid could be produced (94). A new ceramic carrier called Aphrocell with a continuous pore structure was tried (95), but the oxygen supply to this structure was difficult. None of the tested carriers proved to be ideal.

Experimental and theoretical studies on the continuous acetic acid production by immobilized cells in a three-phase fluidized-bed bioreactor at very low acetic acid concentrations were carried out (96). The influence of gel-entrapped and suspended cells, gel size, $k_L a$, and solid hold-up on productivity was studied by theoretical calculations and experiments.

Mori et al (97–99) reported that immobilized acetic acid bacteria are highly oxygen dependent. Therefore, it is advantageous to supply oxygen to a closed and pressurized culture. Because acetic acid inhibits the bacteria, it is considered that multiple reactors in series would be the best solution to obtain an increasing acidity, thereby keeping the dilution rate higher and creating the prolonged residence time necessary. An output acidity of 5.5% was obtained in two fluidized-bed type tabletop reactors in series. In scaling up to a pilot reactor, however, the microbial layer peeled off from the carrier too easily independent of its material. More operational know-how is needed for the practical application of this technology.

Environmental Protection. Today it is possible to produce vinegar without any environmental pollution. The exhaust gas contains alcohol, acetic acid, and ethyl acetate according to their vapor tension at a fermentation temperature of about 30 °C. To obtain high yields, the aeration rate is chosen as low as possible. It is furthermore common practice today to cool the exhaust gas down with cooling water as far as possible, and in order to absorb the rest of the volatiles it can be scrubbed with water that is recycled and used for mash preparation later.

The raw vinegar contains bacteria that have to be filtered off. Today, most modern factories use cross-flow filtration without adding any filter aid. The very small quantity of remaining solids of about 100 g $\times$ 1,000 L^{-1} of vinegar formed by acetic acid bacteria are degraded during wastewater treatment.

Production Volumes

United States. The U.S. vinegar industry ceased years ago to control and to publish vinegar production figures. Therefore, older figures must be quoted here. Production volumes in 1987 were 581 $\times$ 10^6 L of distilled vinegar, 54 $\times$ 10^6 L of cider vinegar and 126 $\times$ 10^6 L of other types of vinegar, all of different strength.

European Union. In 1995, the European Union produced 349 $\times$ 10^6 L alcohol vinegar, 160 $\times$ 10^6 L wine vinegar and 57 $\times$ 10^6 L other kinds of vinegar (all of 10% acetic acid), together 566 $\times$ 10^6 L.

Japan. In 1994, Japan produced 56 $\times$ 10^6 L of rice vinegar, 146 $\times$ 10^6 L of other grain vinegar, 22 $\times$ 10^6 L of fruit vinegar, and 166 $\times$ 10^6 L of other brewed vinegar, all together 390 $\times$ 10^6 L at a concentration of 4 to 5% acetic acid.

Worldwide. The world production (excluding the USSR and China) of vinegar of 10% acetic acid is assumed to be around 2,000 $\times$ 10^6 L per year, or 200, 000 tons of pure acetic acid.

BIBLIOGRAPHY

This article is a shortened and revised republication of H. Ebner, S. Sellmer, and H. Follmann, *Acetic Acid in Biotechnology*, 2nd ed., vol. 6, H.J. Rehm, G. Reed, A. Pühler, P. Stadler eds., Weinheim, VCH, 281–401, 1996, by permission of VCH Publishing Group of May 30, 1996.

1. H. Ebner and H. Follmann, in H.J. Rehm and G. Reed eds. *Biotechnology*. vol. 3, Verlag Chemie, Weinheim, Germany, 1983, pp. 387–407.
2. M. Gillis and J. De Ley, *Int. J. Syst. Bacteriol.* **30**, 7–27 (1980).
3. J. De Ley, J. Swings, and F. Gosele, in N.R. Kreig, G.J. Holt eds., *Bergey's Manual of Systematic Bacteriology*, Williams & Wilkins, Baltimore, 1984, pp. 268–274.
4. E. Entani, S. Ohmori, H. Masai, and K.J. Suzuki, *J. Gen. Appl. Microbiol.* **31**, 475–490 (1985).
5. M. Sievers, S. Sellmer, and M. Teuber, *Syst. Appl. Microbiol.* **15**, 386–392 (1992).
6. A.J. Schocher, H. Kuhn, B. Schindler, N.J. Palleroni, C.W. Despreaux, M. Boublik, and P.A. Miller, *Arch. Microbiol.* **121**, 193–197 (1979).
7. T. Inoue, M. Fukuda, and K. Yano, *J. Ferment. Technol.* **63**, 1–4 (1985).
8. M. Fukaya, T. Iwata, E. Entani, H. Masai, T. Uozumi, and T. Beppu, *Agric. Biol. Chem.* **49**, 1349–1355 (1985).
9. M. Teuber, M. Sievers, and A. Andresen, *Biotechnol. Lett.* **9**, 265–268 (1987).
10. M. Sievers, A. Andresen, and M. Teuber, *Food Biotechnol.* **4**, 555 (1990).
11. M. Sievers, M. Teuber, *J. Appl. Bacteriol. Symp. Suppl.* **79**, 84S–95S (1995).
12. M. Sievers, W. Ludwig, and M. Teuber, *System. Appl. Microbiol.* **17**, 189–196 (1994).
13. M. Fukaya, H. Tagami, K. Tayama, H. Okumura, Y. Kawamura, and T. Beppu, *Agric. Biol. Chem.* **53**, 2435–2440 (1989).
14. H. Okumura T. Uozumi, and T. Beppu, *Agric. Biol. Chem.* **49**, 1011–1017 (1985).
15. M. Fukaya, H. Okumura, H. Masai, T. Uozumi, and T. Beppu, *Agric. Biol. Chem.* **49**, 2083–2090 (1985).
16. M. Fukaya, H. Okumura, H. Masai, T. Uozumi, and T. Beppu, *Agric. Biol. Chem.* **49**, 2407–2411 (1985).
17. M. Fukaya, in Y. Murooka and T. Imanaka eds., *Recombinant Microbes for Industrial and Agricultural Applications*, Decker, New York, 1994, pp. 529–542.
18. M. Teuber, A. Andresen, and M. Sievers, *Biotechnol. Lett.* **9**, 37–38 (1987).
19. W.W. Stamm, M. Kittelmann, H. Follmann, and H.G. Trüper, *Appl. Microbiol. Biotechnol.* **30**, 41–46 (1989).
20. S. Sellmer, M. Sievers, M. Teuber, *System. Appl. Microbiol.* **15**, 610–616 (1992).
21. C. Defives, D. Ochin, J.P. Hornez, and M. Werquin, *Microbiol. Aliments Nutr.* **8**, 77–79 (1990).
22. O.E. Bradley, *Bacteriol. Rev.* **31**, 230–314 (1967).
23. H.W. Ackermann, *Microbiol. Sci.* **4**, 214–218 (1987).
24. M. Ameyama and O. Adachi, *Methods Enzymol.*, **89**, 450–457 (1982).

25. M. Ameyama and O. Adachi, *Methods Enzymol.*, **89**, 491–497 (1982).
26. O. Adachi, K. Tayama, E. Shinagawa, K. Matsushita, and M. Ameyama, *Agric. Biol. Chem.* **42**, 2045–2056 (1978).
27. O. Adachi, E. Miyagawa, E. Shinagawa, K. Matsushita, and M. Ameyama, *Agric. Biol. Chem.* **42**, 2331–2340, (1978).
28. H. Muraoka, Y. Watabe, N. Ogasawara, and H. Takahashi, *J. Ferment. Technol.* **59**, 247–235 (1981).
29. O. Adachi, K. Tayama, E. Shinagawa, K. Matsushita, and M. Ameyama, *Agric. Biol. Chem.* **44**, 503–515 (1980).
30. M. Ameyama, K. Matsushita, Y. Ohno, E. Shinagawa, and O. Adachi, *FEBS Lett.* **130**, 179–183 (1981).
31. R. Hommel and H.P. Kleber, *Proc. 3rd Eur. Congr. Biotechnol.* **1**, 133–137 (1984).
32. H. Muraoka, Y. Watabe, and N. Osagawara, *J. Ferment. Technol.* **60**, 171–180 (1982).
33. K. Matsushita, Y. Takaki, E. Shinagawa, M. Ameyama, and O. Adachi, *Biosci. Biotechnol. Biochem.* **56**, 303–310 (1992).
34. M. Fukaya, K. Tayama, T. Tamaki, H. Tagami, H. Okumura, Y. Kawamura, and T. Beppu, *Appl. Environ. Microbiol.* **55**, 171–176 (1989).
35. K. Tayama, M. Fukaya, H. Okumura, Y. Kawamura, and T. Beppu, *Appl. Microbiol. Biotechnol.* **32**, 181–185 (1989).
36. T. Asai, *Acetic Acid Bacteria*, Univ. of Tokyo Press, Tokyo, Univ. Park Press, Baltimore, 1968.
37. J. De Ley, and K. Kersters, *Bacteriol. Rev.* **28**, 83–95 (1964).
38. M. Kulhanek, in V. Krumpanzl and Z. Rehacek eds., *Modern Biotechnology*, vol. 2, Institute of Microbiology, Czechoslovak Academy of Sciences, Prague, 1984, pp. 614–676.
39. Z.G. Razumovskaya and T.Z. Belousova, *Mikrobiologiya* **21**, 403–407 (1952).
40. O. Hromatka and H. Gsur, *Enzymologia* **25**, 81–86 (1962).
41. G.D. Brown and C. Rainbow, *J. Gen. Microbiol.* **15**, 61 (1956).
42. M.R. Raghavendra Rao and J.L. Stokes, *J. Bacteriol.* **66**, 634–638 (1953).
43. F. Gossele, J. Swings, and J. De Ley, *Zentralbl. Bakteriol. Parasitenkd. Infektionskrankh. Hyg. I Abt. Orig. Cl.* 348–350 (1980).
44. M. Ameyama and K. Kondo, *Agric. Biol. Chem.* **30**, 203–211 (1966).
45. O. Adachi, K. Okamoto, K. Matsushita, E. Shinagawa, and M. Ameyama, *Agric. Biol. Chem.* **54**, 2751–2752 (1990).
46. M. Ameyama, E. Shinagawa, K. Matsushita, and O. Adachi, *Agric. Biol. Chem.* **48**, 2909–2911 (1984).
47. M. Ameyama, E. Shinagawa, K. Matsushita, and O. Adachi, *Agric. Biol. Chem.* **49**, 699–709 (1985).
48. O. Hromatka and H. Ebner, *Enzymologia* **15**, 57–69 (1951).
49. O. Hromatka and W. Exner, *Enzymologia* **25**, 37–51 (1962).
50. H. Muraoka, Y. Watabe, N. Ogasawara, and H. Takahashi, *J. Ferment. Technol.* **61**, 89–93 (1983).
51. A. Hitschmann and H. Stockinger, *Appl. Microbiol. Biotechnol.* **22**, 46–49 (1985).
52. O. Hromatka, G. Kastner, and H. Ebner, *Enzymologia* **15**, 337–350 (1953).
53. F.M. Vera and D.I.C. Wang, *174th Am. Chem. Soc. Mfg.*, Chicago, 1977.
54. O. Hromatka and H. Ebner, *Enzymologia* **13**, 369–387 (1949).
55. A. Mori and G. Terui, *J. Ferment. Technol.* **50**, 70–78 (1972).
56. U.S. Pat. 4,503,078 (Mar. 5, 1985), H. Ebner (to Heinrich Frings GmbH & Co. KG, Bonn, Fed. Rep. of Germany).
57. U.S. Pat. 4,076,844 (Feb. 28, 1978), H. Ebner and A. Enenkel (to Firma Heinrich Frings, Bonn, Germany).
58. U.S. Pat. 4,773,315 (Sept. 27, 1988), A. Enenkel (to Heinrich Frings GmbH & Co. KG, Bonn, Fed. Rep. of Germany).
59. U.S. Pat. 4,282,257 (Aug. 4, 1981), Y. Kunimatsu, H. Okumura, H. Masai, K. Yamada, and M. Yamada (to Nakano Vinegar Co., Ltd., Japan).
60. U.S. Pat. 4,364,960 (Dec. 21, 1982), Y. Kunimatsu, H. Okumura, H. Masai, K. Yamada, and M. Yamada (to Nakano Vinegar Co., Ltd., Tokyo, Japan).
61. U.S. Pat. 4,456,622 (Jan. 26, 1984), J.A. Maselli, and R.O. Horwath (to Nabisco Brands, Inc., Parsippany, N.J.).
62. U.S. Pat. 3,813,086 (May 28, 1974), H. Ebner and A. Enenkel (to Firma Heinrich Frings, Bonn, Germany).
63. U.S. Pat. 3,262,252 (July 26, 1966), H. Ebner.
64. U.S. Pat. 4,997,660 (Mar. 5, 1991), R. Wittler (to Heinrich Frings GmbH & Co. KG, Bonn, Fed. Rep. of Germany).
65. U.S. Pat. 3,290,924 (Dec. 13, 1966) H. Ebner and A. Enenkel (to Heinrich Frings Maschinen-und Apparatebau, Bonn, Germany).
66. U.S. Pat. 4,404,284 (Sept. 13, 1983), M. Heider, J. Hofschneider, W. Schallenberger, F. Nodes, and E. Kempe (to Vogelbusch Gesellschaft m.b.H., Vienna, Austria).
67. D. Hekmat and D. Vortmeyer, *J. Ferment Bioeng.* **73**, 26–30 (1992).
68. A. Mori, *Instrum. Control. Engin. Jpn., Kogyo Gijutsu-sha* 18–24 (1989).
69. Y. Nomura, M. Iwahara, and M. Hongo, *Appl. Environ. Microbiol.* **54**, 137–142 (1988).
70. Y. Nomura, M. Iwahara, and M. Hongo, *J. Biotechnol.* **12**, 317–326 (1989).
71. Y. Nomura, *Hakkokogaku* **70**, 205–216 (1992).
72. J.J. Bänsch and W. Bauer, *BioEngineering* **7**, 26–36 (1991).
73. A. Nanba, A. Tamura, and S. Nagai, *J. Ferment. Technol.* **62**, 501–505 (1984).
74. Y.S. Park, H. Ohtake, K. Toda, M. Fukaya, H. Okumura, and Y. Kawamura, *Biotechnol. Bioeng.* **33**, 918–923 (1989).
75. Y.S. Park, H. Ohtake, M. Fukaya, H. Okumura, Y. Kawamura, and K. Toda, *J. Ferment. Bioeng.* **68**, 315–319 (1989).
76. Y.S. Park and K. Toda, *J. Gen. Appl. Microbiol.* **36**, 221–233 (1990).
77. M. Fukaya, Y.S. Park, and K. Toda, *J. Appl. Bacteriol.* **73**, 447–454 (1992).
78. Y.S. Park, M. Fukaya, H. Okumura, Y. Kawamura, and K. Toda, *Biotechnol. Lett.* **13**, 271–276 (1991).
79. Y.S. Park, K. Toda, M. Fukaya, H. Okumura, and Y. Kawamura, *Appl. Microbiol. Biotechnol.* **35**, 149–153 (1991).
80. H. Ebner, in *Ullmanns Enzyklopädie der Technischen Chemie, 4.* Aufl. Bd. 11, Verlag Chemie Weinheim, 1976, pp. 41–55.
81. H.A. Conner and R.J. Allgeier, *Adv. Appl. Microbiol.* **20**, 81–133 (1976).
82. K. Toda, Y.S. Park, T. Asakura, C.Y. Cheng, and H. Ohtake, *Appl. Microbiol. Biotechnol.* **30**, 559–563 (1989).
83. Y.S. Park, H. Ohtake, and K. Toda, *Appl. Microbiol. Biotechnol.* **33**, 259–263 (1990).
84. Brit. Pat. 1 305 868 (Feb. 7, 1973), Kewpie Jozo Kabushiki Kaisa, Tokyo, Japan.
85. T. Higashide, C. Shimada, Y. Kawamura, M. Hisamatsu, and T. Yamada, *Bull. Fac. Biores. Mie Univ. Tsu* **8**, 43–49 (1992).
86. Eur. Pat. 0523 524 A1 (Jan. 20, 1993), Carl Kühne KG, Hamburg, Deutschland.

87. A. Okuhara, *J. Ferment. Technol.* **63**, 57–60 (1985).
88. A. Nanba, K. Kimura, and S. Nagai, *J. Ferment. Technol.* **63**, 175–179 (1985).
89. M. Kondo, Y. Suzuki, and H. Kato, *Hakkokogaku* **66**, 393–399 (1988).
90. A. Saeki, *Nippon Shokuhin Kogyo Gakkaishi* **37**, 191–198 (1990).
91. A. Saeki, *Nippon Shokuhin Kogyo Gakkaishi* **37**, 722–725 (1990).
92. A. Saeki, *Nippon Shokuhin Kogyo Gakkaishi* **38**, 891–896 (1991).
93. M. Takada and T. Hiramitsu, *Nippon Shokuhin Kogyo Gakkaishi* **38**, 967–971 (1991).
94. S. Yamashita, S. Ohta, and H. Suenaga, *Nippon Shokuhin Kogyo Gakkaishi* **38**, 608–613 (1991).
95. M. Sueki, N. Kobayashi, and A. Suzuki, *Biotechnol. Lett.* **13**, 185–190 (1991).
96. Y. Sun, S. Furusaki, *J. Ferment. Bioeng.* **69**, 102–110 (1990).
97. A. Mori, *Process Biochem.* **20**, 67–74 (1985).
98. A. Mori, in A. Tanaka, T. Tosa, and T. Kobayashi eds., *Industrial Application of Immobilized Biocatalysts*, Dekker, New York, 1992, pp. 291–313.
99. A. Mori, S. Tanaka, N. Matsumoto, and C. Imai, in S. Furusaki, I. Endo, and R. Matsuno eds., *Biochemical Engineering for 2001*, Springer-Verlag, New York, 1992, pp. 441–443.

WASTE GAS CLEANING, BIOLOGICAL

CORNELIUS G. FRIEDRICH
UDO WERNER
University of Dortmund
Dortmund, Germany

KEY WORDS

Biofilms
Biofilters
Bioscrubber
Flame ionization detection
Mass transfer
Microbial degradation
Plant Design
Trickle-bed bioreactor
Volatile Organic Carbon
Volatile organic emissions

OUTLINE

INTRODUCTION

Waste air or waste gases of industrial processes may contain volatile organic or inorganic odorous or toxic compounds. These gases are not to be released to the environment, according to national or European legislation, which defines the allowances of the individual compound to be released. Environmental biotechnology including waste gas treatment was first developed in countries using the German language and in The Netherlands. In these countries, waste gas treatment was applied by the industry for elimination of odor emitted from wastewater treatment plants, slaughterhouses, and such facilities. For more than two decades, plants have been operated with increasing efficiency, and their application has been extended to the chemical, pharmaceutical, and food industries. Initially, most information has evolved from these European countries; experience has been compiled and merged into guidelines for design and application of industrial biofilters, trickle-bed reactors, and bioscrubbers. The principles of these plants are shown in Figure 1. These guidelines have been published by the Verein Deutscher Ingenieure (VDI, Association of German Engineers) in German and in English (1,2). The developing technology was subsequently recognized on other continents. Previous articles on biological waste gas cleaning have covered different processes, mass transfer, biological, chemical, and olfactory aspects, and costs, and the reader is referred to comprehensive reviews (3–5) with detailed references therein.

High concentrations of organic or inorganic pollutants are currently eliminated from waste gases by sorptive processes or thermal of catalytic oxidation. Low pollutant concentrations and odorous compounds are also being traditionally eliminated by biological processes. Plant investment and operating costs of biological waste gas cleaning processes are significantly less than alternative physicochemical processes (6). This fact is the major driving force for academic and industrial research to develop biological processes and to extend these for elimination of high concentrations of pollutants from waste gases.

Presently, the three types of plants mentioned (Fig. 1) are currently used for biological treatment of different qualities of waste gases, which later are treated in detail. Biofilters require a substantial area and are filled to a low height with organic matter such as compost or chopped wood, which serves in part as a source for minerals for the attached microorganisms. Biofilters are suited for cleaning low concentrations of pollutants. Trickle-bed reactors require less area, are operated as towers, and are filled with incompressible porous minerals as air–liquid contactor and as support for microbial growth. Bioscrubbers are generally composed of two units—one for scrubbing the pollutants from the gas phase and the other as fermentor for microbial degradation of the pollutants comparable to activated sludge.

Microorganisms, bacteria or fungi, use the organic or inorganic compounds for energy transformation and for biomass formation. Biological waste gas cleaning is an aerobic process and is brought about by aerobic microorganisms either suspended in water or attached to some support material as a biofilm. Process development is required to establish and to maintain optimum conditions for these microorganisms to achieve an optimal elimination of pollutants from air. Microorganisms require different elements to build up biomass, which is composed of proteins, sugar polymers, fatty acids, nucleic acids, and so forth. The

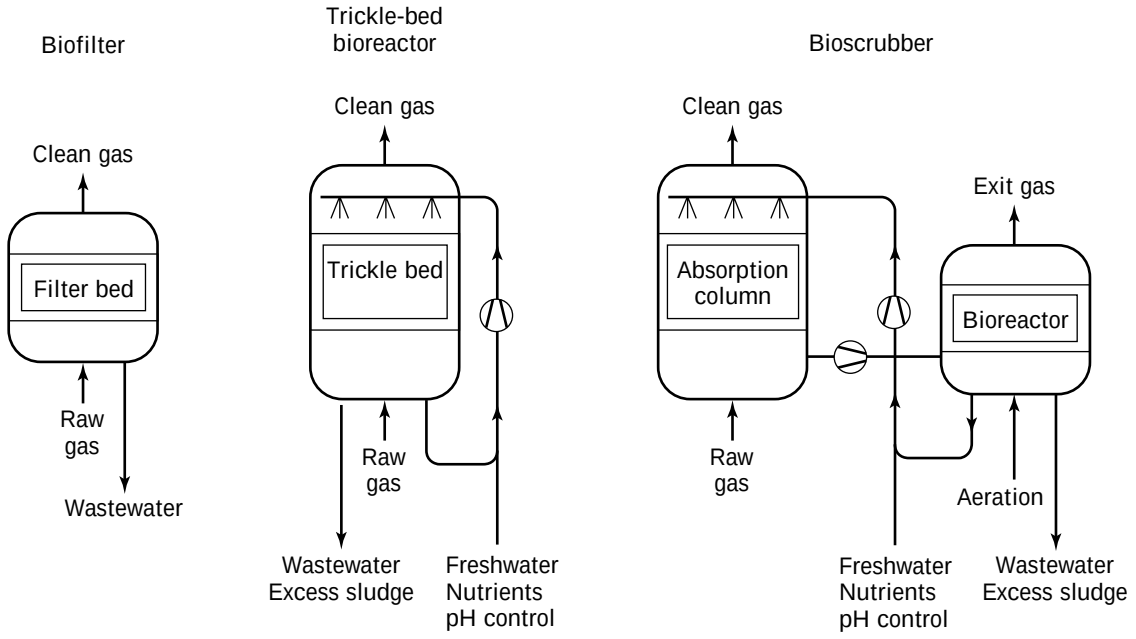


Figure 1. Schematic representation of different types of plants currently used for waste gas treatment.

energy to build up the respective monomers and to polymerize these is derived from the biochemical oxidation of volatile organic matter to carbon dioxide or from gaseous inorganic reduced compounds. Because the microorganisms are present in the aqueous phase, the organic or inorganic substrates must be transported from the gas to the liquid phase. Because biological waste gas cleaning is also an aerobic process, oxygen has to be transported to the liquid film and products have to be transported either to the gas or to the liquid phase.

The degree of elimination of a pollutant and the selection of the process is determined by different factors (Fig. 2): (*1*) the quality of the waste gas determined by the chemical nature of the pollutant, composition of different pollutants, pollutant concentrations, pollutant metabolic accessibility, and the volume of the gas emitted per unit of time; (*2*) the biomass and its physiological state determined by the rate of conversion of the substrates; (*3*) the availability of the substrates determined by the mass transfer of the pollutants and of oxygen from the gas phase to the biofilm and the export of degradation products therefrom (mass transfer depends on the process selected [e.g., biofilter, trickle-bed reactor, or bioscrubber] and on the physicochemical nature of the pollutants to be eliminated); and (*4*) costs of investment and of operation for waste gas cleaning determine the selection of each of these factors by the impact on overall productivity, reliability, and thus profitability.

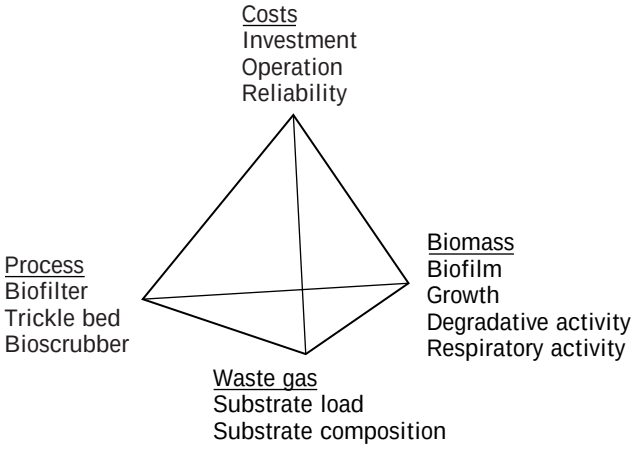


Figure 2. Interrelationships of different parameters crucial for costs and efficiency of waste gas treatment plants.

MICROORGANISMS AND THE BIOFILM

Industrial waste gas cleaning systems are currently operated under nonaxenic conditions and, as a result, contain different populations of bacteria, fungi, and protozoa. Such populations may establish randomly by means of the content of bacteria and bacterial and fungal spores in the surrounding air. To rapidly establish a microbial population in the reactors, these are often inoculated with suspensions of wastewater treatment plants. Because the composition of the waste gas determines the development of microorganisms best fitted to respective conditions, initial efficiency of a plant may be suboptimal. Also, because microorganisms are the crucial phase in biological treatment, they are emitted by the stream of the clean gas that is the subject of hygienic concern.

Establishment of the Microbial Population

Initial colonization of the microbes is the result of adsorption onto the surface of the support material, which occurs

almost instantly. Secondary colonization is characterized by the formation of filaments and of extracellular polysaccharides, which together form the biofilm. The kinetics of biofilm formation, its composition, changes, properties, and impact on biocorrosion have been reviewed previously (7). The composition of the species changes with duration of operation. Moreover, under steady-state conditions of the plant, the biofilm represents an ecosystem with grazing protozoa digesting free-living bacteria. Protozoan predation has resulted in a low and constant level of of bacteria, which, however, may not be detrimental to the efficiency of a plant but increase stability of trickle-bed reactors, for example. Enhancement of carbon mineralization up to 22 to 80% and decrease in pressure drop by preventing the extent of clogging have been reported from laboratory waste gas treatment systems. Moreover, most protozoa are not attached to surfaces and could be purged out by flushing the different systems (summarized in Ref. 8).

The use of pure bacterial cultures with defined catabolic properties with the ability to degrade xenobiotic or toxic compounds at known rates has been suggested to be advantageous with respect to waste gas cleaning. Several bacterial species specialized to degrade different solvents have been adsorbed on activated carbon. With suitable bacterial strains, 90% of the solvents could be eliminated from the gas phase, and mass transfer to the biofilm appeared to be limiting the efficiency of the laboratory trickle-bed reactor (9). No reports are presently available for industrial application of monocultures.

Health Considerations

Health hazards for personnel or the surrounding population emerging from waste gas treatment plants have not been reported. However, different bacterial species so far identified from various waste gas treatment plants have been mainly related to the degradative capability of the respective facility. Among these, species of the genus *Pseudomonas* have been identified that are highly versatile in utilizing aromatics and other volatiles. One concern is the release of potentially pathogenic microbes to the environment with the exit gas stream. Some studies have been devoted to this fact. The transport of microbial germs, mainly bacteria and to a smaller extent spores of molds, in the outlet gas of the different full-scale biofilters varied between 10^3 and 10^4 cells/m^3, which is of the same order of magnitude as indoor air (10). Search for fungal spores, the potential causative agent of mycogenic allergies, revealed a minor reduction of spores emitted from clean gas as compared to raw gas from moist biofilters, whereas emission from a dry biofilter temporarily exceeded that determined from raw gas by more than three orders of magnitude (11).

Identification of microorganisms is laborious, time consuming, and costly, and most reports consider the absorption of pollutants rather than the emission of microbes. The technology to detect and to identify specific bacteria within multispecies biofilms, however, is presently available and is used to describe microbial ecosystems. Nucleic acid hybridization techniques and comparative sequence analyses are increasingly being applied to environmental microbiology. More recently, fluorescent dye-labeled oligonucleotide probes have been used for identification of single cells (12).

MICROBIAL GROWTH AND DEGRADATIVE ACTIVITY

Organic or inorganic pollutants of waste gases are substrates for microorganisms. Organic carbon sources have a dual function as substrate for the living cell. Dissimilation of the carbon source characterizes its quality as an energy source to donate electrons for aerobic respiration, which yields the energy required for growth and maintenance of the cell. Carbon is also assimilated by the cell for growth to build up cell constituents. Carbon dioxide and water are the products of the dissimilatory aerobic respiration, and biomass is the product of the resulting assimilation. The quality of a given carbon source as an energy source depends on its degree of reduction and on the type of metabolism, as is shown for dissimilation of one-carbon compounds such as methane, methanol, and formaldehyde (Table 1).

These compounds are oxidized to carbon dioxide, water, and the reductant. Moreover, substitutions of hydrogen by halogens such as methylenechloride cause the release of hydrochloric acid. Degradation of methylmercaptane results in the release of hydrogen sulfide and that of methylamine results in the release of ammonia. The by-product hydrogen sulfide is subsequently oxidized by thiobacteria to sulfuric acid, and ammonia is oxidized by nitrifying bacteria to nitric acid, both resulting in acidification of the environment (Table 1). Therefore, the nature of the pollutant to be mineralized by the microorganisms has a significant impact on the pH of the microenvironment within the plant and on pH-dependent microbial degradative activity.

Table 1. Respiration and Mineralization of Differently Reduced and Substituted One-Carbon Compounds, Nitrification, and Sulfur Oxidation

Respiration

$$4[H]^a + O_2 \rightarrow H_2O$$

Mineralization

$$CH_4 + O_2 \rightarrow CO_2 + 4[H]$$
$$CH_3\text{-}OH + H_2O \rightarrow CO_2 + 6[H]$$
$$CH_3\text{-}NH_2 + 2H_2O \rightarrow CO_2 + NH_3 + 6[H]$$
$$CH_3\text{-}SH + 2H_2O \rightarrow CO_2 + H_2S + 6[H]$$
$$CH_2\text{-}Cl_2 + 2H_2O \rightarrow CO_2 + 2HCl + 4[H]$$
$$CH_2O + H_2O \rightarrow CO_2 + 4[H]$$

Nitrification

$$NH_3 + 2H_2O \rightarrow HNO_2 + 6[H]$$
$$HNO_2 + H_2O \rightarrow HNO_3 + 2[H]$$

Sulfur oxidation

$$H_2S + 4H_2O \rightarrow H_2SO_4 + 8[H]$$

[a][H] is equivalent to ($H^+ + e^-$) and represents the reductant hydrogen carrying one electron.

Acidification as a result of the elemental composition of the pollutants as well as the composition and distribution of the different microbial species may require mixing, washing, and pH control of the biofilm attached to the support material or being suspended in the aqueous phase. The different principles of waste gas treatment processes such as biofilter, trickle-bed reactor, or bioscrubber differ in efficiency in mass transfer, capacity to harbor biomass, and costs. Thus, knowledge of the quality of the waste gas with respect to its composition or load is essential for selection of the process.

One crucial criterion in microbial transformations is the overall rate of degradation. The mean residence time of a gas passing certain industrial waste gas treatment plants is occasionally less than a second (2). In this period of time, mass transfer and degradation of the pollutant must occur—and it does.

CELLULAR YIELDS AND MAINTENANCE REQUIREMENTS

The composition of the waste gas usually contains few elements—carbon, hydrogen, or oxygen as in the case of acetone—and other elements (N, P, S, Mg, K, Fe, etc.) are also required to build up biomass. Formation of biomass with constant elemental composition is a function of the organic substrate when each element required for growth is supplied by the environment and expressed as the cellular yield ($Y_{X/S}$) (gram dry cell weight per gram of substrate). In biofilters, some minerals may originate from the support material such as wood, heather, or compost, whereas in other waste gas cleaning plants minerals may be supplied from the added water. Mostly, there will be a severe imbalance in supply with elements required for microbial growth, whether the major pollutant was a hydrocarbon, solvent, hydrogen sulfide, ammonia, or methylenchloride. Imbalance in nutrient supply may originate from the surplus of the carbon source or from the limited supply of minerals, both resulting in two different consequences. First, imbalance of the elemental composition of the biomass may result from storage of different compounds. Excess phosphate supply to certain bacteria results in accumulation of intracellular polyphosphate; excess carbon and energy source may result in either intracellular accumulation of polyhydroxy fatty acids or the formation of extracellular polysaccharides and biofilms. Second, the growth rate may be affected and the requirements for maintenance energy may increase significantly. Maintenance energy is defined as the amount of substrate required by the cell to maintain vital cell functions without growth (equation 1). The concept of maintenance requirements was first described by Pirt (13) and was later extended by Neijssel and Tempest with the definition of overflow metabolism (14).

$$1/Y = m/\mu + 1/Y_{max} \quad (1)$$

where Y is the cellular yield, Y_{max} the maximum cellular yield, μ the specific growth rate (1/h), and m the maintenance coefficient (g substrate/g dry cell weight h^{-1}).

Overflow metabolism may result in two different phenomena: *(1)* the uncoupling of energy transformation of respiration leading to a low cellular yield and higher rate of oxidation of the substrate to carbon dioxide and *(2)* the formation of polymeric or monomeric by-products. The applicability of these academic findings to waste gas cleaning systems has been subsequently proven successful in laboratory plants (15). However, defined conditions of specific nutrient limitations are susceptible to changing concentrations of volatile organic carbon in the raw gas, which changes the balance in nutrient supply, and thus conditions for optimal mineralization to carbon dioxide and water are difficult to maintain.

ORIGIN, COMPOSITION, AND SENSING OF WASTE GAS COMPONENTS

Waste gas originates from very different industries, from air exchange of buildings, and from various processes. More than 500 biofilters and more than 200 bioscrubbers are currently operated in Germany and in The Netherlands, although outside Europe this technology is less often used. Different industries that apply biological waste gas cleaning as an air pollution control technology are listed in Table 2. The efficiencies of some of these to remove volatile organic carbon or odor are compiled in Table 3. A more comprehensive compilation is given elsewhere (1,2). Most of these facilities emit gases at high flow rates of 1,000 to 120,000 m^3/h but with low concentrations of volatile organic and inorganic compounds. Odor removal is quantified by olfactometry, which is costly and requires reliable calibration as evident from an interlaboratory comparison of odor measurements of *n*-butanol (25). Recent developments aim at on-line sensors for the quantification of odors. Odor threshold concentrations to be recognized by humans depend on the respective compound and composition of the volatiles. Some threshold concentrations for sensing the odor of chemicals such as dimethylethylamine, phenol, or volatiles of defined composition have been correlated to total organic carbon determined by flame

Table 2. Compilation of Some Industries Applying Biological Waste Gas Cleaning

Biofilter	Trickle-bed reactor	Bioscrubber
Animal rendering	Chemical industry	Aluminum foundry
Bone processing	Concrete slap production	Animal rendering
Chemical industry	Fibrous skin production	Blood processing
Coal mine exit gas	Grease melting plant	Chipboard production
Dump gas	Plastic production	Egg farms
Feces drying	Rayon production	Grease melting plant
Fish smoking	Sewage treatment plants	Grinding wheel production
Fish processing		Printing plant
Foundries		Tobacco industry
Fragrance production		Polyester production
Food production		
Pig farms		
Slaughterhouse		
Tobacco processing		
Varnish production		

Table 3. Compilation of Selected Efficiencies of Different Industrial Waste Gas Cleaning Facilities

Waste gas cleaning plant	Type of reactor	Eliminated pollutant	Total gas flow rate	Raw gas concentration	Clean gas concentration	Refs.
Rendering plant	Tree bark biofilter	Odor/VOC	160 $m^3\ m^{-3}\ h^{-1}$	256 mg organic C/m^3	38 mg organic C/m^3	
				1,700 OU	160 OU	1
Domestic refuse compost	Compost biofilter	Odor/VOC	60 $m^3\ m^{-3}\ h^{-1}$	230 mg organic C/m^3	8–9 mg organic C/m^3	1
Piggery	Peat/twigs biofilter	Odor	141 $m^3\ m^{-3}\ h^{-1}$	50–200 OU	Not detectable	1
Leather processing facility	Compost/chaff biofilter	Solvents	128 $m^3\ m^{-3}\ h^{-1}$	61 g $m^{-3}\ h^{-1}$	41 g $m^{-3}\ h^{-1}$	16
Canning industry	Compost biofilter	Solvents	40 $m^3\ m^{-3}\ h^{-1}$	470 mg organic C/m^3	55 mg organic C/m^3	23
Lacquerworks	Pilot biofilter	Solvents	30–71 $m^3\ m^{-3}\ h^{-1}$	642–1,415 mg/m^3	60–79 mg/m^3	17
Cacao roast house	Biofilter	VOC	75,000 m^3/h	7–615 mg organic C/m^3	1–6 mg organic C/m^3	18
		Odor		49,927–446,000 OU	35–773 OU	18
Printer's color production	Trickle-bed reactor	Toluene/ ethanol	214 $m^3\ m^{-3}\ h^{-1}$	450 mg organic C/m^3	<70 mg organic C/m^3	19
Resin-laden waste gas	Pilot trickle-bed reactor	Formaldehyde	760 $m^3\ m^{-2}\ h^{-1}$	10–30 mg/m^3	0.9–2.7 mg/m^3	20
		Phenol		10–30 mg/m^3	0.1–0.3 mg/m^3	20
		Ammonia		20–90 mg m^3	1.6–7.2 mg/m^3	20
Plastic production	Trickle-bed reactor	Styrene	51,000 m^3/h	0.6 g/m^3	Data not given	21
Lacquer production	Bioscrubber/ biofilter	Solvents	16,000 m^3/h	≤800 mg/m^3	50 mg/m^3	21
Car tunnel exit gas	Pilot trickle-bed reactor	Carbon monoxide	8–23 m^3/h	10–30 ppm	0.8–2.4 ppm	22
		NO_x		0.5 ppm	0.1 ppm	22
		C_xH_y		2 ppm	0.8 ppm	22
Wastewater treatment	Bioscrubber	Hydrogen sulfide	6.000 m^3/h	20 ppm	0.2 ppm	23
Rendering plant	Bioscrubber	Odor	40.000 m^3/h	10,000–70,000 OU	<100 OU	2
Plastic production facility	Bioscrubber	Hydrophobic VOC	10.000 m^3/h	≤200 mg organic C/m^3	10–40 mg organic C/m^3	24

Note: VOC, volatile organic compounds; OU, olfactometric units (cannot be compared from different references).

ionization detection, which is a reliable and available technology for continuous exit gas analysis. Threshold concentrations for odor sensing vary from 0.1 to 4.9 mg organic carbon/m^3 for defined chemicals and from 0.1 to 0.5 mg organic carbon/m^3 for waste gas from food processing and other industries. Usually, the degree of reduction of odor intensity is more effective than the elimination of volatile organic carbon.

Flame ionization detection of volatiles allows quantification of organic carbon but does not differentiate between different components of the waste gas. Thus, the degree of elimination of individual compounds cannot be differentiated. To follow the elimination of different components, identification by gas chromatography is required. This analysis has been used to follow degradation of different solvents in waste gases.

Corrosion of the detecting devices may be observed on on-line quantification of haloorganic compounds such as chloromethanes, chlorobenzenes, chloroethanes, or fluorinated or bromated compounds by flame ionization detection because of formation of the respective hydrohalic acids. Therefore, sampling, gas chromatography, and different detection systems may be advisable.

MASS TRANSFER IN WASTE GAS TREATMENT PLANTS

The principles of mass transfer were first elaborated to understand and to design chemical reaction processes and have been described elsewhere (26,27). However, these principles can be transduced also to biochemical and biological reactions where substrates must be delivered to the microbial cell and metabolic products have to be removed therefrom (3,28,29). This is especially important when a substrate is supplied from the gas phase to the aqueous phase, which is the environment for microbial life and activity.

The principles of mass transfer are applicable to the different processes treated here. However, the differences in design of these plants are based on different principles of gas–liquid contactors for transfer of volatile chemicals of different properties such as hydrophilic or hydrophobic and soluble or insoluble compounds. Therefore, the principles of the transfer of substrates of different physicochemical properties from the gas phase to the biofilm are discussed first, and subsequently mass transfer in biofilters, trickle-bed reactors, and bioscrubbers.

Mass Transfer from the Gas Phase to the Biofilm

Total mass transfer rate (J) from the gas to the liquid phase is driven by the difference in concentration of a given component in equilibrium with the gas phase (c_G) and the actual concentration in liquid phase or the biofilm (c_B), the specific interfacial area (a), and a correlation factor designated mass transfer coefficient (β) as described by equation 2:

$$J = \beta a(c_G - c_B) \quad (2)$$

The specific interfacial area is generally defined as the quotient of the total area for mass transfer and the total reactor volume.

The mass transfer coefficient is influenced by the different steps for mass transfer from the gas phase to the liquid phase and to the biofilm. These steps are the superpositions of convection and diffusion in both fluid phases and the absorptive transition of the component to the liquid phase and to the biomass, as shown schematically in Figure 3a. The equilibrium in absorption at the transition phase is expressed by Henry's law, equation 3:

$$c_G = Hc_F \quad (3)$$

where c_G is the partial pressure of the component in the gas and c_F of that in the fluid. A similar correlation holds for the phase-interface liquid or biofilm.

During steady-state operation of a plant, mass transfer rate (J) is in equilibrium with the substrate consumption rate (J_R) of the biomass as expressed in equation 4:

$$J = J_R = K_R a c_B^r \quad (4)$$

where K_R is the reaction rate of the biomass, a the specific area, c_B the concentration of the component in the biofilm, and r the order of the reaction. Mostly, biological reactions are of zero order when these are independent of the substrate concentration and of first order when the substrate concentration determines the reaction rate. The overall reaction rate is determined by the different steps just mentioned. However, in special cases one step may dominate the process, for example, when the rate of absorption of a pollutant is low as given for chemicals with Henry coefficients >1 MPa or low biological degradation rates. Parameters affecting the capacity of waste gas cleaning facilities are discussed next.

Limitations of Reaction Rates

Several examples have been described that demonstrate the significance of single steps on the overall process (27,28,30). The degradative activity of a given biofilm may be limited by suboptimal physiological conditions, which are discussed here, and the mass transfer rates. These conditions apply to the organic or inorganic electron donor for energy transformation, in our case the volatile pollutants, mass transfer rate of the electron acceptor oxygen, the low rate of pollutant degradation by the biofilm, or a combination of these caused by local differences within the reactor.

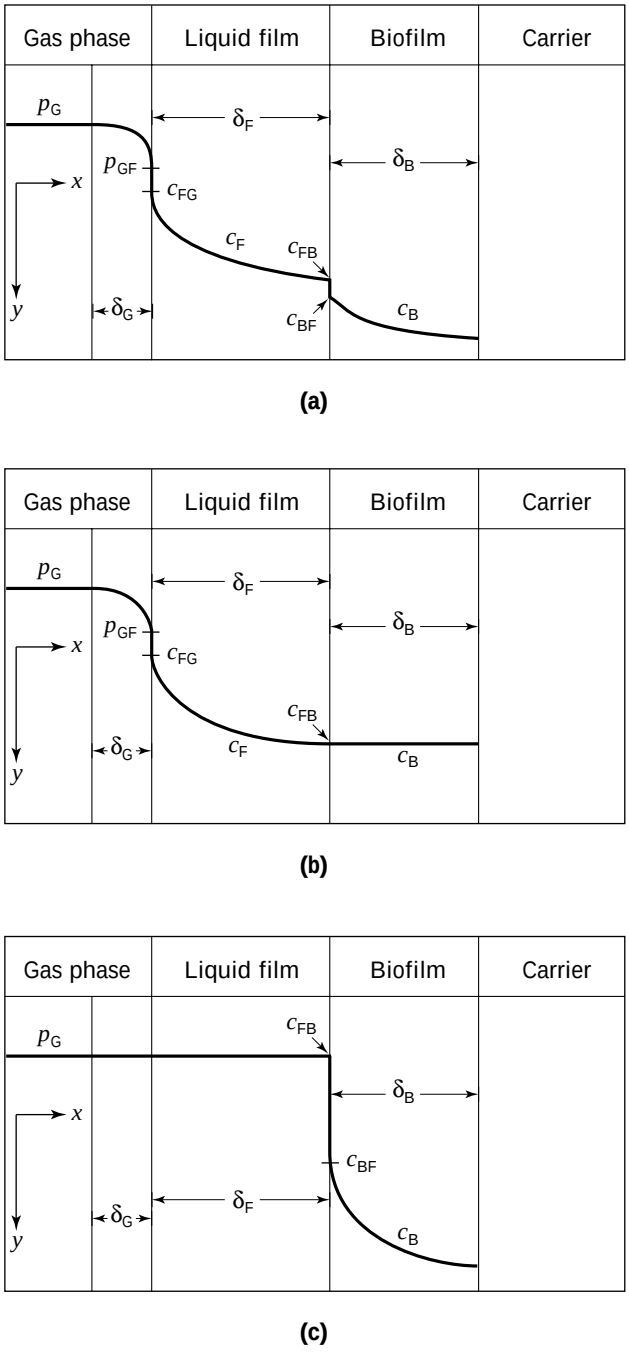


Figure 3. Substrate concentration profiles within the gas phase, the liquid phase, and the biofilm. (**a**) Decrease in concentration caused by biofilm degradative activity and substrate transport by diffusion and convection; (**b**) limitation by mass transfer from the gas phase to the liquid film; (**c**) limitation by degradative activity of the biofilm. p_G, p_{GF}, pollutant partial pressure in the gas phase and at the phase-interface gas–fluid; c_{FG}, c_F, c_{FB}, pollutant concentration at the phase-interface fluid–gas, fluid, and phase-interface fluid–biofilm, respectively; c_{BF}, c_B, pollutant concentration at the phase-interface biofilm–fluid and in the biofilm; δ_G, δ_F, δ_B, boundary layer thickness of the gas phase, fluid, and biofilm, respectively.

- Limitation by mass transfer of oxygen from the gas to the liquid phase may be given for a pollutant highly soluble in water, easily transferred from the gas into the biofilm, and rapidly degraded therein (Fig. 3a). Although the oxygen partial pressure of air is 21 kPa, it is barely soluble in water (7.5 mg of oxygen/L at 30 °C), and thus its mass transfer rate competes with that of low concentrations of the pollutants.
- Limitation by mass transfer of the pollutant from the gas to the liquid phase may result from rapid degradation of a pollutant by the biomass. In this case, mass transfer can be enhanced by increase in superficial gas velocity or by reduction of the irrigation rate. Both operations decrease the thickness of the boundary layers δ_G and δ_F (Fig. 3b).
- Limitation by microbial degradation may result either from low degradative activity or from low biomass. Both affect the efficiency of waste gas cleaning similarly. Low degradative activity is given for toxic or xenobiotic pollutants and other persistent chemicals not accessible to the common microbial metabolism (Fig. 3c). Low biomass may result from insufficient supply of mineral nutrients, nonphysiological pH, and support material inadequate to allow sufficient biomass attachment.
- Different limitations, as have been described, may be present in one system at different locations. In a system with a high load of a pollutant, insufficient degradation may take place at the entrance of the reactor where the mass transfer rate is higher than the degradation rate. As the pollutant is removed from the gas stream while passing through the reactor, the concentration in the gas phase decreases. Finally, a limiting mass transfer rate may result as compared to the potential microbial degradation rate.

The efficiency of elimination of a pollutant is clearly dependent on its solubility and mass transfer characteristics. Consequently elimination of a pollutant from the gas phase is linked to the Henry coefficient of the respective compound whether it is eliminated by physicochemical processes or microbial degradation. Such a link is evident for the theoretical and actual efficiency of elimination of different pollutants shown in Figure 4.

The differences in the processes of waste gas cleaning beneficial or deterious to the aspects as stated are discussed in the respective sections.

PARAMETERS FOR PLANT DESIGN

For a given waste gas to be treated, the appropriate type of plant has to be selected, and its size has to be estimated. Therefore, basic characteristics are crucial for plant design. Available area and costs are criteria for choosing between biofilter and trickle-bed reactor, while the hydropathic nature of a pollutant is crucial to decide for or against a bioscrubber. For elimination of pollutants with Henry coefficients >1 MPa, bioscrubbers are usually not chosen.

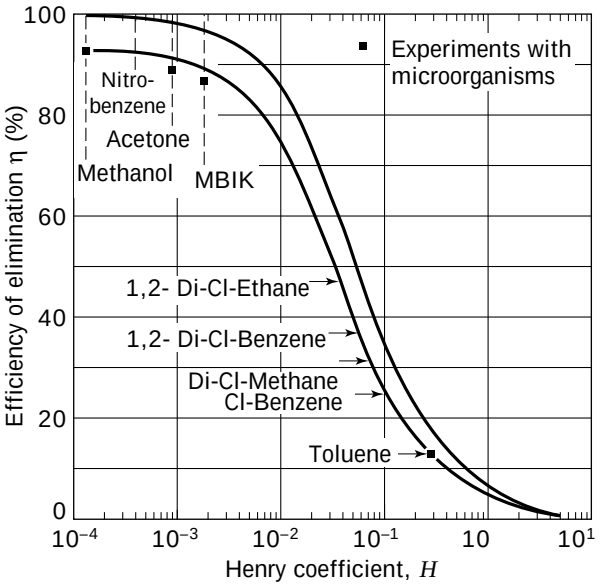


Figure 4. Dependency of the efficiency of elimination of various components on their Henry coefficients. The data are based on a single plate bioscrubber. The upper line represents theoretical efficiency; the lower line gives data of a single-plate bioscrubber. Liquid height on plate = 70 cm, T = 20 °C, u_0 = 0.55 m/s. The dimension of H (MPa) is equivalent to 135.4 ($kg/m^3_G/kg/m^3_L$). *Source*: Redrawn from Ref. 33, with permission.

In general, the mean residence time of the gas and the fluid is crucial for the three types of plants. The residence time is a measure to define the time given for contact of the gas, the liquid, and the biomass to allow mass transfer. The mean residence time (t_m) of a gas or fluid in a reactor with a given volume (V_R) at a given flow rate (V^*) is expressed in equation 5:

$$t_m = V_R/V^* \quad (5)$$

The inverse relation of this mean residence time of the gas defines the volumetric flow rate (τ_V), as expressed in equation 6:

$$\tau_V = V^*/V_R \quad (6)$$

The relation of the flow rate of a gas or fluid to the area of the cross section of the plant (A_R) is defined as area load (τ_A) as expressed in equation 7:

$$\tau_A = V^*/A_R = u_0 \quad (7)$$

where u_0 is the superficial velocity of a gas or fluid, which is the velocity with which they pass the free cross section of the reactor.

In waste gas cleaning, two criteria are important with respect to a pollutant: its concentration and its amount. The amount of a pollutant fed to a plant of a defined volume per unit of time is given by its concentration per volume of gas (c_{in}) and by the gas flow rate, and is expressed as volumetric load (V_L) as shown in equation 8:

$$V_L = (V^*/V_R)c_{in} \quad (8)$$

The pollutant is metabolized in the plant by the biomass and eliminated from the gas phase. Generally, the elimination is equivalent to the degradation of the pollutant, and this term is subsequently used. The overall degradation is defined as volumetric degradation rate (V_D), as expressed in equation 9:

$$V_D = (V^*/V_R)(c_{in} - c_{ex}) \quad (9)$$

The difference in concentration of a pollutant in the feed and the exit gas (c_{ex}) related to the feed concentration characterizes the efficiency of a plant, as expressed in equation 10:

$$\eta = (c_{in} - c_{ex})/c_{in} \quad (10)$$

The efficiency of a plant is determined by various parameters such as the chemical nature of the pollutant, biomass degradative activity, and gas or fluid flow rate. The latter is determined by the resistances within biofilters or trickle-bed reactors given by the type of filling, its density, porosity, and height of the package resulting in pressure loss (26), and may be a significant cost factor for operation. The power (P) required to achieve a certain gas flow rate (V^*) will compensate for the overall pressure loss (Δp) across the reaction zone of a plant, as expressed in equation 11:

$$P = \Delta p V^* \quad (11)$$

Other pressure losses resulting from such components as pipings, nozzles, and orifices are not specific to waste gas cleaning facilities and are not treated here.

The porosity (ε) is a crucial parameter of pressure loss for biofilters or trickle-bed reactors, and is defined by the interstitial volume of a rector. This factor is given by the difference in volume of the empty reactor (V_R) and the volume of the particles (V_P) related to the total reactor volume, as expressed in equation 12:

$$\varepsilon = (V_R - V_P)/V_R \quad (12)$$

TYPES OF WASTE GAS TREATMENT PLANTS

As mentioned, three types of plants are currently used for treatment of industrial waste gases of different origin and composition: biofilter, trickle-bed reactor, and bioscrubber (Fig. 1). Their principles, industrial sizes, capacities, and advantages or disadvantages are outlined next. Certain characteristics increase with the plant used. The applied volumetric flow rate (see equation 6) increases roughly 10-fold as the order of biofilters, trickle-bed reactors, and bioscrubbers. This order reflects not only an inverse relationship of efficiency and size of the plants but also their increasing complexities.

Biofilters

Different designs are currently applied for the construction of biofilters: single bed, multiple bed, and closed bed, either requiring land or installed on top of a roof. Biofilters require an area of 10 to more than 1,000 m^2 filled to a height of 0.5–1.6 m with different materials as support for microbes and for fine distribution of the gas. The filters are operated at rates of 100–120,000 m^3/h. The gas flow rate related to the filter volume (see equation 6) varies from 10 to 580 $m^3\ m^{-3}\ h^{-1}$. The superficial gas velocity of biofilters varies from 0.1 to 8 cm/s. The mean residence time within a filter bed of 1-m height is rather long and amounts to about 10 to 700 s (1). The filling material may be compost, peat, heather, chopped wood, bark, or disrupted root wood to which microorganisms are attached. Biofilters are suited to eliminate odors and volatile organic or inorganic compounds at rather low concentration, as detailed next. The area of the biofilter contains a wall of wood or concrete and a system for distribution of the main stream of the gas. It is practical for the biofilter to be vented from the bottom, that the orifices or slits allow constant gas flow rates over the filter area, and—for larger plants—to allow machines to enter the filter for exchange of the filling.

A waste gas may contain dust particles that are removed by a deduster before feeding into the biofilter. The gas passed through the biofilter should either be humidified to about 95% or the filter should be irrigated to ensure a moist environment as required for microbial activity. The concentration or the composition of the pollutants may vary with the operation of the emitting source, resulting in changes of load. Moderately higher concentrations can be temporarily reduced and a portion of the compounds stored in the filling material according to its adsorptive properties. The compounds are subsequently released therefrom for mineralization by the biomass. Thus, the efficiency of the biofilter can be maintained.

The filling material supplies the attached microorganisms with the minerals necessary for growth. The degradative capacity of the biofilters is limited by the low supply of minerals, which in turn causes a low amount of active biomass present on the support material, its limited buffering capacity for protons or volatiles, and the lack of control of the microbial environment with respect to pH. The filling material is of natural organic matter and subject to microbial degradation at different rates according to the structure, surface, and density. Moreover, environmental conditions may add to the rate of decomposition and loss of mechanical properties of the filling material. Such conditions may be acidification by organic or mineral acids as detailed in Table 1. Decomposition of the filling material results in settling, reduction in pore size, and increase in pressure loss and consequently in increased energy requirements for air compression. Moreover, settling of the filling may result in channeling wherein the gas stream circumvents but does not penetrate portions of lower porosity within the filter bed. This in turn causes reduction of the residence time of the gas and reduction in efficiency of pollutant elimination. Reduction in efficiency of the plant requires exchange of the decomposed filling material, which usually occurs after 1–3 years of continuous operation depending on filling texture, mechanical properties, and degradability and the load of the waste gas to be treated.

Trickle-Bed Reactor

Pollutants with low solubility in water may be eliminated from waste gases by trickle-bed reactors. The advantage of operation of a trickle-bed reactor over biofilters is not only in the increased mass transfer but also in the ease of controlling physiological conditions such as pH and the addition of minerals in the trickling fluid in the bottom of the column. Moreover, trickle-bed reactors require less land than biofilters and are operated in single or multiple stages. They are built as columns filled with incompressible and mostly inert support material for gas–liquid contact and attachment of biomass. The bed heights vary from 1 to 3 m, the reactor volumes from 2 to 160 m^3, and the flow rates from about 3,000 to 85,000 m^3/h. The mean residence time of the gas within the columns is generally less than 4 s, which is considerably shorter than reported for biofilters (see Table 3), and the superficial gas velocities vary from 0.1 to 1.6 m/s. The volumetric gas flow rates are rather divergent and vary from 250 to 4,700 $m^3\ m^{-3}\ h^{-1}$ (2).

The inert support material may be slag, natural porous stones, coke, charcoal, granular clay, ceramics, sintered glass, or plastics and other material proven to be suitable in absorption technology. The filling is trickled from top, and the surface area of the filling provides a major basis for mass transfer from the gas to the fluid phase. Sufficient empirical data are available for choosing geometry, size, and filling for various industrial applications. Moreover, based on the geometry of the fillings, the total surface area of the trickle bed can be calculated. The thickness of the fluid film is proportional to the irrigation rate and inversely related to the total surface area. This in turn allows calculating the mass transfer rate for a given pollutant (27). Also, total biomass can be deduced from the mean thickness of the biofilm. Trickle-bed reactors may be seeded with microorganisms by spraying activated sludge from wastewater treatment plants. The microbes will be selected and the population of the biofilm will be established according to the given pollutants.

A major resistance for mass transfer from the gas phase to the biofilm is the thickness of the fluid film, determined by the rate of irrigation and rate and direction of the gas flow. Therefore, the irrigation rate should be reduced to generate thin films. Depending on the waste gas to be treated, the rates vary from 0.04 to 50 $m^3\ m^{-2}\ h^{-1}$ as reported from laboratory and industrial reactors. However, industrial plants are preferentially operated at <0.5 $m^3\ m^{-2}\ h^{-1}$ (2). To enhance mass transfer and to reduce diffusional resistance by the liquid film, intermittent irrigation has been applied for elimination of pollutants by a laboratory trickle-bed reactor, which still ensured sufficient moisture for the biofilm (31–33).

Mass transfer may also be increased by increasing the surface area as given with support material of high specific surface. Application of such material results in lower interstitial volume, lower bed volume, and a smaller plant size. However, the thickness of the biofilm will increase with time of operation of the trickle-bed reactor, decreasing or abolishing the small interstitial volumes, which results in increase in pressure loss, clogging, and channel formation. The mean residence time is reduced and the efficiency of the plant declines in eliminating pollutants from the gas stream. The otherwise favorable characteristics of trickle-bed reactors with respect to mass transfer and control of circulation have led to the concept of removing excess biomass by various means: increased irrigation rate to wash off the biofilm (2), periodic pulses of air pressure after flooding, and movement of the trickle bed itself to shear off the biofilm (32). In laboratory systems this procedure of biomass removal has led to continuous operation for more than 2 years with constant degrees of elimination of pollutants. Research data now suggest that the degradative capacities of trickle-bed reactors can be better described mathematically than biofilters and allow conclusions for growth of biomass and of operation of trickle-bed reactors in industrial scale (31,32). Reduction of biomass as a result of a biological equilibrium of biomass formation and biomass degradation has been reported from a pilot trickle-bed reactor after long operation. In this system, elimination of the acidogenic pollutant dichloromethane was reported, involving conversion into biomass, whereas clogging as a result of biomass degradation was not observed (34). Dichloromethane is a substrate with low yield in energy transformation (Table 1) and maintenance requirements (see earlier discussion).

Bioscrubbers

The decision to select bioscrubbers for waste gas cleaning is based on the physicochemical properties of the pollutants and the availability of space for placing the plant. Bioscrubbers are usually smaller than biofilters and consist mostly, but not exclusively, of two units, one for absorption of the pollutant from the gas phase and the other as a fermentor for their mineralization and growth of microorganisms. Bioscrubbers are usually inoculated with activated sludge, and the microbial population will develop according to the volatiles supplied with the gas stream. Waste gas is passed through a spray tower, bubble column, absorber column, or other gas–liquid contactors for absorption of the pollutants. Industrial bioscrubbers presently treat 9,000 to 150,000 m^3 waste gas per hour. Scrubber units are built in heights of 2–4 m and volumes of 10–100 m^3 and exhibit volumetric gas flow rates of 1,000–3,600 $m^3\ m^{-3}\ h^{-1}$. Bioscrubbers are usually operated with superficial gas velocities of 0.8–2.0 m/s and with mean residence times of <3 s. The specific flow rate of the liquid phase to be circulated per area of the absorber unit ranges from 12 to 40 $m^3\ m^{-2}\ h^{-1}$. Some data of industrial plants are compiled in Table 3, and more have been given elsewhere (2). Elimination of hydrophobic pollutants from the gas phase in single-unit bioscrubbers is usually not advisable because of the high volumes of water to be moved. However, single-unit bioscrubbers have been successfully applied to clean industrial volatile pollutants with Henry coefficients of more than 1 MPa (24). Moreover, new developments in part operating in an industrial scale have been reported for application of bioscrubbers that combine both units in one apparatus. These systems consist of special devices for gas or liquid dispersion, special orifices at the bottom of bubble columns that contain the dispersed

microorganisms (33), or gassed closed-loop air-lift reactors (35), which, however, await industrial application.

The physiological conditions for the microorganisms such as pH, salinity, and oxygen partial pressure can be well controlled in bioscrubbers, as similarly described for trickle-bed reactors. Also, mineral salts nutrients can be added to the medium to allow microbial biomass to grow. The advantage of bioscrubbers over trickle-bed reactors is in the ease of controlling the environment for microorganisms and mechanically improving mass transfer rates by movement of the liquid according to the need to absorb pollutants from the gas phase. On the other hand, elimination of hydrophobic compounds with low solubilities in water requires movement and dispersion of large volumes of liquids. Bioscrubbers are an interesting alternative to other technologies that have been outlined with respect to the ease of operation, environmental control, and adjustment to changing conditions in the quality of the feed waste gas. Thus, bioscrubbers can be applied for elimination of hydrophobic pollutants (24). A variation of the process uses silicon oil for absorption of hydrophobic volatiles. Problems, however, may arise from disposal of silicon oil–containing sludge (2).

COSTS

Application of biological treatment has been proven useful to economically clean waste gases of different sources. Presently, waste gases with concentrations of less than 1 g of organic carbon/m^3 or of inorganic compounds are treated by biological processes. At these concentrations, biological processes are economically advantageous over alternative processes such as condensation, thermal or catalytic oxidation, or adsorptive and desorptive technologies. These processes are economically feasible at higher concentrations, as detailed in Table 4. Separation of organics from the gas phase by membrane technologies listed in Table 4 is just in its infancy (6,37). Performance of membrane bioreactors for waste gas treatment is presently being examined at a laboratory scale. Reactor performance, however, is hampered by unstable biofilms or by clogging of the liquid channels from excess biomass formation. On solution of these problems, membrane bioreactors are expected to be useful tools for treatment of poorly water-soluble pollutants such as highly chlorinated hydrocarbons (reviewed in Ref. 21).

In a comparative study, costs for investment and for operation were related to treatment by biofilters. Costs of investment include those for the plant and for measuring and controlling devices. Costs of operation include consumables, energy requirements, and costs for personnel (6) (Table 5). From this study it is evident that bioscrubbing is more costly with respect to investment and operation. At present, data on the economy of trickle-bed reactors are not available. However, biofiltration and bioscrubbing are significantly less costly than catalytic oxidation or adsorption (Table 5).

Adsorptive processes aim at the recovery of solvents, for example, whereas thermal or catalytic oxidation aims at the elimination of the pollutants. To make use of the low costs of biological processes for elimination of a pollutant, it is anticipated to extend this technology to waste gases with organic loads significantly larger than 1 g organic carbon/m^3. Two alternatives are presently considered. First, classical engineering processes as specified are applied to eliminate high concentrations or peak concentrations toxic to microorganisms and to combine these with conventional biological processes. This concept was successfully applied recently for elimination of solvents from waste gas of the lacquer industry (38). Second, the inexpensive biological processes are to be extended to treat waste gases with high organic load. The conversion of this concept requires addition of mineral nutrients, which results in formation of high amounts of biomass. Removal of the excess biomass is required to prevent undesired consequences such as clogging in trickle-bed reactors or increase in salt concentration in bioscrubbers or both. However, at laboratory scale high concentrations of biomass allowed the complete (100%) elimination of as much as 4.2 g toluene/m^3 or 8.7 g *n*-butanol/m^3 from the gas stream (39).

GENERAL CONSIDERATIONS

Biological waste gas treatment is less costly and requires less maintenance than conventional physicochemical, thermal, or catalytic processes. However, problems may arise from the chemical nature of the compounds to be

Table 4. Processes and Ranges of Concentrations Commonly Used to Eliminate Hydrocarbons from Waste Gases

Process	Concentration (g organic carbon/m^3 waste gas)
Condensation	>250
Membrane separation	>20
Thermal oxidation	6–20
Catalytic oxidation	1–10
Absorption	1–50
Adsorption	1–25
Biological waste gas treatment	<1
Biological treatment with biomass removal[a]	1–10

Source: Data compiled according to Ref. 36, extended. Pilot and laboratory scale.

Table 5. Relative Costs of Investment and of Operation for Different Waste Gas Treatment Plants

Type of process	Investment costs (%)	Annual costs of operation (%)
Biofilters	100	100
Bioscrubbers	158	196
Catalytic oxidation	182	288
Adsorption	236	365

Source: Altered from Ref. 6, with permission.
The data compare calculated costs of plants for treatment of waste gas with a flow rate of 30,000 m^3 and 6,000 h of operating time.

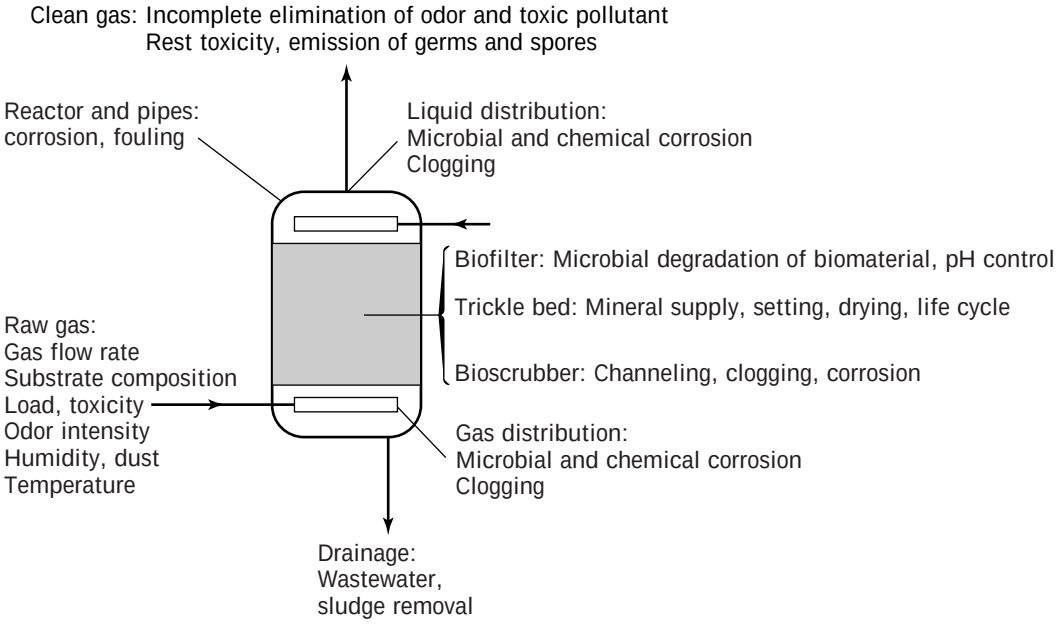


Figure 5. Summary of general problems of biological waste gas treatment plants.

eliminated from waste gases and from the material of which parts of the plant are constructed, such as the walls of the reactor and such parts as piping, orifices, and nozzles (Fig. 5). The waste gases to be treated may contain compounds that are hardly or very slowly biodegradable, such as benzene and xenobiotics, and other compounds toxic to different cell functions, such as solvents, cyanic acid, and hydrogen sulfide. However, successful operation of laboratory-scale waste gas treatment plants to eliminate such compounds point to the applicability of biological treatment to this category of pollutant also.

Different compounds may be converted to organic acids by incomplete oxidation when the mass transfer of oxygen is insufficient. Inorganic acids may be released on microbial degradation of acidogenic compounds, as detailed in Table 1. Acidification, however, not only results in decrease of pH and an unfavorable change of the physiological conditions for the microorganisms but also in corrosion of the material of the plant. The waste gas of rendering plants contains volatile organic sulfur compounds converted to sulfuric acid; this causes severe corrosion of concrete supports of the gas distribution devices of the filter bed. Therefore, such devices have to be exchanged for such components as wooden beams. The same applies for other acidogenic compounds of waste gases. Therefore, the appropriate material for piping and such should be selected according to the expected composition of the waste gas for the three types of plants.

Trickle-bed reactors may especially suffer from clogging and channeling by excess formation of biomass or slime with high loads of the waste gas or extended operation, as was outlined. Therefore, altered processes are presently being examined to extend the application of biological treatment of waste gases loaded with compounds that are slowly degraded or that contain high loads of organic and inorganic pollutants. This application in turn will require devices for the control of the environment for the microorganisms and the removal of excess biomass, and recent developments have been devoted to this need.

BIBLIOGRAPHY

1. VDI-Richtlinie 3477, *VDI/DIN-Handbuch Reinhaltung der Luft*, vol. 6, Beuth Verlag, Berlin, 1991.
2. VDI-Richtlinie 3478, *VDI/DIN-Handbuch Reinhaltung der Luft*, vol. 6, Beuth Verlag, Berlin, 1996.
3. S.P.P. Ottengraf, in H.J. Rehm and G. Reed eds., *Biotechnology*, vol. 8, VCH Verlagsgesellschaft, Weinheim, 1986, pp. 425–452.
4. G. Leson and A.M. Winer, *J. Air Waste Manage. Assoc.* **41**, 1045–1054 (1991).
5. P.L. Bishop and N. Kinner, in H.J. Rehm and G. Reed eds., *Biotechnology*, vol. 8, VCH Verlagsgesellschaft, Weinheim, 1986, pp. 114–176.
6. H. Menig, H. Krill, and T. Jäschke, in W.L. Prins and J. van Ham eds., *Biological Waste Gas Cleaning*, VDI Verlag, Düsseldorf, 1997, pp. 27–58.
7. H.-C. Flemming, *Wasser Abwasser* **132**, 197–207 (1991).
8. H.H.J. Cox and M.A. Deshusses, in W.L. Prins and J. van Ham eds., *Biological Waste Gas Cleaning*, VDI Verlag, Düsseldorf, 1997, pp. 233–240.
9. K. Kirchner, S. Wagner, and H.-J. Rehm, *Appl. Microbiol. Biotechnol.* **37**, 579–587 (1992).
10. S.P.P. Ottengraf and J.H.G. Konings, *Bioprocess Eng.* **7**, 89–96 (1991).
11. C. Huwe and H. Pöhle, *Biofilter* **7**, 47–49 (1997).
12. R.I. Amann, J. Stromley, R. Devereux, R. Key, and D.A. Stahl, *Appl. Environ. Microbiol.* **58**, 614–623 (1992).
13. S.J. Pirt, *Proc. R. Soc. London* **163**, 224–231 (1965).
14. O.M. Neijssel and D.W. Tempest, *Arch. Microbiol.* **107**, 215–221 (1976).

15. S.-M. Wübker and C.G. Friedrich, *Appl. Microbiol. Biotechnol.* **46**, 475–480 (1996).

16. K. Stefan, A. Windsperger, and R. Buchner, *Verfahrenstechnik* **24**, 12–15 (1990).

17. D. Eitner, *Wasser, Luft und Boden* **7–8**, 44–47 (1996).

18. D. Eitner, in K. Fischer ed., *Biologische Abluftreinigung*, Expert, Böblingen, 1990, pp. 55–73.

19. C. Kellner and E. Vitzthum, in W.L. Prins and J. van Ham eds., *Biological Waste Gas Cleaning*, VDI, Düsseldorf, 1997, pp. 61–66.

20. M. Arnold, J. Lehtomäki, and J. Roine, in W.L. Prins and J. van Ham eds., *Biological Waste Gas Cleaning*, VDI, Düsseldorf, 1997, pp. 75–82.

21. H. Herzog and N. Thißen, in W.L. Prins and J. van Ham eds., *Biological Waste Gas Cleaning*, VDI, Düsseldorf, 1997, pp. 83–90.

22. K.-H. Robra, M. Wellacher, F. Kirchmeir, R. Leistentritt, and K. Pucher, in W.L. Prins and J. van Ham eds., *Biological Waste Gas Cleaning*, VDI, Düsseldorf, 1997, pp. 131–140.

23. D. Eitner, in W.L. Prins and J. van Ham eds., *Biological Waste Gas Cleaning*, VDI, Düsseldorf, 1997, pp. 385–395.

24. S. Stockhammer, W. Schäfer-Treffenfeldt, and K. Zetzmann, *Chem. Ing. Tech.* **64**, 148–155 (1992).

25. P. Heeres and H. Harssema, *Gefahrst.-Reinhalt. Luft* **56**, 55–60 (1996).

26. R.H. Perry and C.H. Chilton, *Chemical Engineer's Handbook*, McGraw-Hill, New York, 1987.

27. M. Baerns, H. Hoffmann, and A. Renken, *Chemische Reaktionstechnik*, Thieme, Stuttgart, 1992, pp. 82–97.

28. K. Kirchner, in R. Margesin, M. Schneider, and F. Schinner eds., *Praxis der Biotechnologischen Abluftreinigung*, Springer, Berlin, 1996, pp. 1–16.

29. A. Windsperger and M. Sotoudeh, in R. Margesin, M. Schneider, and F. Schinner eds., *Praxis der Biotechnologischen Abluftreinigung*, Springer, Berlin, 1996, pp. 17–53.

30. R.M.M. Dicks and S.P.P. Ottengraf, in *Biologische Abgasreinigung*, VDI-Berichte, 1104, Düsseldorf, 1994.

31. H. Heits, A. Laurenzis, and U. Werner, in W.L. Prins and J. van Ham eds., *Biological Waste Gas Cleaning*, VDI, Düsseldorf, 1997, pp. 313–320.

32. A. Laurenzis, H. Heits, S.-M. Wübker, U. Heinze, C. Friedrich, and U. Werner, *Biotechnol. Bioeng.* **57**, 497–503 (1998).

33. F. Wolff, *Biologische Abluftreinigung mit Suspendierten und Immobilisierten Mikroorganismen*, Fortschrittsberichte VDI 94, Düsseldorf, 1992, pp. 71–106.

34. R.M.M. Diks, S.P.P. Ottengraf, and S. Vrijland, *Biotechnol. Bioeng.* **44**, 1279–1287 (1994).

35. A. Laurenzis and U. Werner, in W.L. Prins and J. van Ham eds., *Biological Waste Gas Cleaning*, VDI, Düsseldorf, 1997, pp. 115–122.

36. G. Börger, *Chem. Ing. Tech.* **64**, 905–914 (1992).

37. M. Schultes, *Abgasreinigung*, Springer, Berlin, 1996, pp. 209–238.

38. M.W. Reij, J.T.F. Keurentjes, and S. Hartmans, *J. Biotechnol.* **59**, 155–167 (1998).

39. S.-M. Wübker, U. Heinze, and C.G. Friedrich, in W.L. Prins and J. van Ham eds., *Biological Waste Gas Cleaning*, VDI, Düsseldorf, 1997, pp. 241–248.

WASTE TREATMENT, ACTIVATED SLUDGE, CONTROL STRATEGIES

GUSTAF OLSSON
Lund Institute of Technology
Lund, Sweden

KEY WORDS

Activated sludge
Control
Instrumentation
Operation

OUTLINE

TREATMENT PLANT PROCESSES

The treatment of municipal and some industrial wastewater generally uses a combination of primary, secondary, and tertiary treatment. Primary treatment simply screens and settles large particulates and skims off floating greases and oils. Secondary treatment biologically removes organic carbon, and in newer plants, soluble nitrogen or phosphates. Tertiary treatment attempts to limit the microorganisms and other pathogens in the treated water by membrane filtration or deep-bed filters and some form of disinfection using chlorine, ozone, or ultraviolet light.

The secondary or biological treatment of wastewater uses a number of key processes, unit operations in the process engineering terminology. These include primarily processes for biological reactions and processes for settling and clarification. Chemical precipitation to remove phosphorous is sometimes used just before the settling and clarification. There are many flowsheet configurations, often of a proprietary nature, that attempt to obtain the optimal combination of anaerobic, anoxic, and aerobic con-

ditions for the microorganisms. Commonly used flowsheets are shown in Figure 1. A detailed account of instrumentation, control and automation in wastewater treatment plants has been made in the textbook by Olsson and Newell (1).

GOALS AND INCENTIVES FOR CONTROL

The goal of the operation of a wastewater treatment plant is to satisfy the effluent requirements at all times at minimum cost. The effluent quality and a good performance has to be guaranteed *consistently*, which will have a great impact on instrumentation, control, and automation. Effluent requirements are usually formulated in terms of daily, weekly, or annual averages of suspended solids, biological oxygen demand (BOD), total nitrogen, and total phosphorus. To comply with the limits of the effluent concentrations, certain safety margins are required. The goal of automatic control is to keep the effluent quality consistently good, despite disturbances. This allows the margins to be smaller, thus saving costs.

Criteria

The capital costs are usually the dominant part of the total cost. Traditionally, many plants have been designed with large tank volumes to meet the varying loads to the system. The total cost can be decreased, if the balance between design and operational costs is systematically considered. There are many components of the operational costs, such as consumption of electric energy (for pumping, aeration, etc.), chemicals for phosphorus removal, or polymers for sludge. Obviously, labor costs are important. The balancing of costs is, of course, an important reason for control. Considering the increasing demand of consistent operation and effluent water quality around the clock, however, will probably be the main driving force for plant control and automation.

Driving forces for instrumentation, control, and automation (ICA) in a large number of countries were discussed at an IAWQ symposium in 1993 (2). The major incentives were categorized as

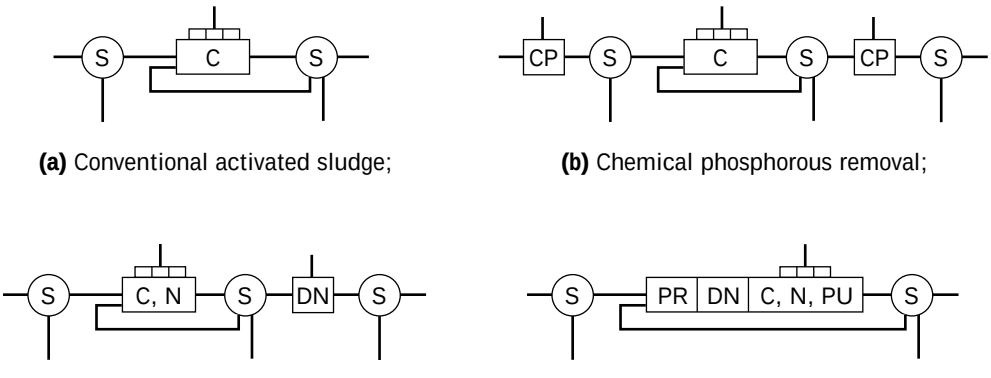
- Effluent quality standards
- Economy
- Increasing plant complexity
- Available tools (sensors, computers, models)

In some countries, there is a water shortage, which in itself is a driving force for improved wastewater treatment, such as for recycling.

In many industrial countries there is today an elaborate requirement for nutrient removal. This has led to the need for quality control based on advanced on-line instrumentation. With more complex processes, the needs for energy conservation and consistent operation are apparent. Ammonium and nitrate on-line instrumentation in the activated sludge was developed and documented in the late 1970s. At that period, the on-line measurements were still limiting the applicability. Very large efforts have been spent on the development of N and P analysers, and there has been a lot of progress in the instrumentation technology. In 1990s, the control is extended to include biological P removal in many plants. Today, commercial applications of nutrient removal control based on advanced instrumentation are common (3). A typical representative layout of nutrient removal plants is the University of Cape Town (UCT) activated sludge plant (Fig. 2).

Disturbances

The essence of the control problem in wastewater treatment is related to disturbances. They are critical to the approach to control and have some specific features:



(a) Conventional activated sludge; **(b)** Chemical phosphorous removal; **(c)** Postdenitrification process; **(d)** Typical BNR process.

C carbon removal, N nitrification, D denitrification, CP chemical precipitation, PR phosphorous release, PU phosphorous uptake, S setting and clarification

Figure 1. Typical wastewater treatment process flowsheets. (**a**) Conventional activated sludge; (**b**) Chemical phosphorous removal; (**c**) Postdenitrification process; (**d**) Typical BNR process. C = carbon removal; N = nitrification; D = denitrification; CP = chemical precipitation; PR = phosphorous release; PU = phosphorous uptake; S = settling and clarification.

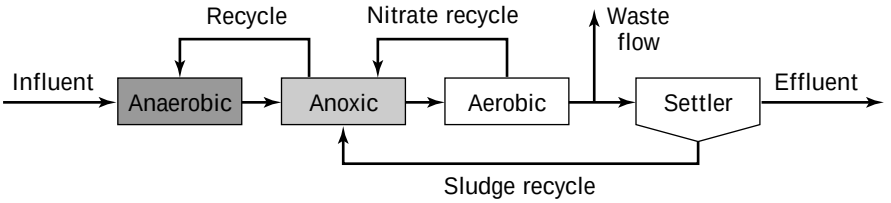


Figure 2. The UCT process, designed to remove organic material, ammonium nitrogen, and phosphorus.

- The daily volume of wastewater treated can be huge.
- The disturbances in the influent are enormous compared to most industries.
- The influent must be accepted and treated; no returning it to the supplier.

Because the disturbances to a wastewater treatment system are significant, the plant is for all practical reasons never in steady state. The disturbances are related to flow rate changes as well as concentration and composition variations. Figure 3 illustrates a typical influent situation in two municipal treatment plants.

The disturbances are not only related to the influent flow characteristics, but may be generated internally in the plant because of operational procedures. For example, poor pumping operations contribute to hydraulic disturbances (4). Clarifier upsets are often caused by inadequate pump control. A sudden flow increase caused by a pump start will influence the flow propagation along the plant within a fraction of an hour. A proper settler operation requires that the influent pumping has to be smooth. Variable speed pumping is a proved technology today and is a crucial component to create a smooth hydraulic operation of the plant. If primary pumps or other recycle pumps are operated on/off in too-large steps, the resulting solids–liquid separation will suffer, as will the effluent water quality.

Concentration disturbances are also generated within the plant itself. For example, sludge thickener overflow is usually returned to the plant inlet. This overflow is usually operated with on/off pumping without considering the influent load. Because the recirculated flow may have a large concentration, it will make a significant contribution to the overall load, both in terms of hydraulics and substrate loading.

It is usually impossible to remove the source of the disturbance, even if this is an important option, especially for industrial effluents, where production control in the industry may be improved. Sometimes the magnitude of the disturbance can be reduced *before* it reaches the plant. An integrated sewer-treatment plant control can attenuate hydraulic disturbances, and some waters can be pretreated in order to avoid harmful effects on the plant, such as neutralization of wastes with high or low pH. The control can also compensate for effects *within* the process. A common example is dissolved oxygen control.

All these disturbances will more or less influence the effluent water quality. An ultimate operational goal is to minimize the influence of disturbances and to maintain consistently a high effluent quality.

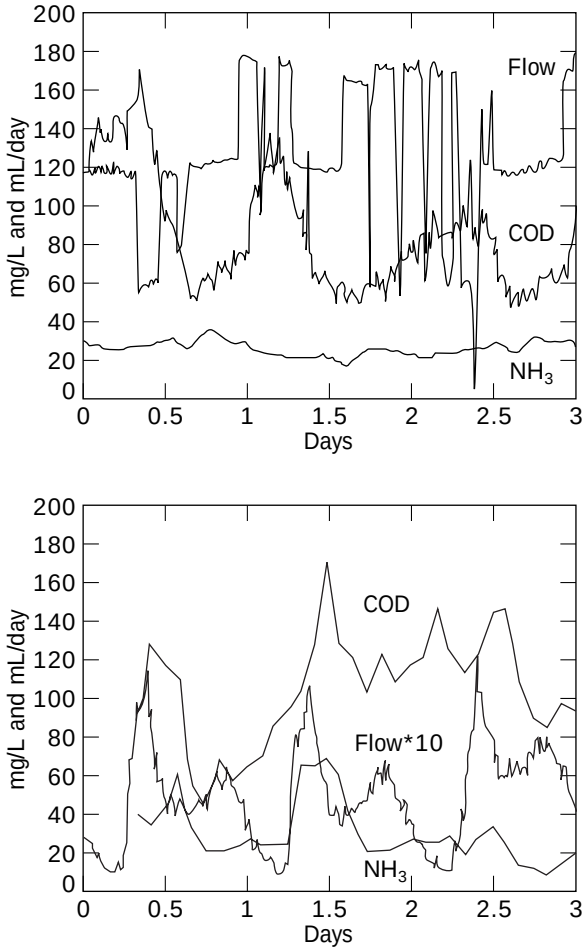


Figure 3. Typical influent disturbances from a large plant in Sweden (**a**) and a small plant in Australia (**b**). The flow rate as well as influent concentrations of chemical oxygen demand (COD) (related to organic content) and total nitrogen have significant variations.

CONTROL AND OPERATIONAL GOALS

It is far from straightforward to implement control in an activated sludge system; some of the reasons are listed here:

- The process has significant nonlinearities limiting the usefulness of simple controllers.
- There is a very wide range of response times, both a problem and a blessing.
- There are many interactions within the process with major recycles.

- There are many external couplings from the sewer to the receiving waters.
- Designers often leave too few manipulated variables and too little control authority.
- The concentrations of nutrients (pollutants) are very small, even challenging sensors.
- Sensors do not exist for many states and are expensive for others.
- The value of the product in the marketplace is remarkably low.

Other attributes why control is far from straightforward include:

- The microorganisms change their behavior and their population distributions.
- The separation of the effluent from the biomass is challenging and easily disturbed.
- Future effluent standards will be tight and based on spot checks.

Disturbance rejection is a major goal of activated sludge control. Large disturbances must be attenuated, and we must do it as early in the process as we can. This will ease the task of controllers later in the process.

The formulation of the goal of a wastewater treatment operation is much more demanding than to define the effluent standards. When a goal or a performance criterion is defined, then the necessary control actions can be analyzed. There are several goals of wastewater treatment, and one has to recognize the different levels and interest groups. The specific goals are set so that they contribute to meeting the more general corporate and community goals. They include at least

- To meet effluent discharge requirements.
- To achieve good disturbance rejection.
- To optimize operation to minimize the operating cost.

The first goal of course includes keeping the plant running. This is the major task for many operations, at least in terms of manpower. It is not a trivial task to maintain advanced instrumentation and equipment, to make sure it runs smoothly, and to detect any faults in its operation. Much of this is traditional automation and process control and is not specific to running wastewater treatment systems (5,6). The environment may be very specific and harsh for much equipment, but that is also the case in other industries as well, so from a strict control point of view the problems are traditional.

However, keeping the plant running and simply recording effluent quality is too often taken as the only goal. This has been and is still largely because of the lack of process knowledge on the part of plant operators. Too many operators of plants rely on the consultant for expertise in both design and operation. To satisfy the effluent quality requirements, it is crucial to ask: is this task solved primarily by design or do instrumentation, control, and automation play significant roles in this part? The traditional answer among designers used to be *no*. When only carbonaceous removal was considered, the plant was considered self-regulating and could take care of load disturbances and still produce a satisfactory effluent. The traditional method was large volumes, excess aeration, and chemical dosage. Consequently, a lot of plants have been oversized. No incentive was given for improved operation, because operational costs were not compared with investment costs. Often they were met from different financial sources.

Because of the significant load variations and disturbances typical for wastewater treatment systems, the plant is hardly ever in steady state. Despite the wide swings of load, the plant still has to *consistently* produce a satisfactory effluent. If the volumes of the tanks are not oversized, this means that some control system has to minimize the influence of disturbances and ensure that the operation is satisfactory. Pressures to fully utilize volumes and the complexities of nutrient removal are making process control essential. This of course includes disturbance attenuation.

Minimizing the operating cost and maximizing the use of available capacity is becoming necessary. Here, control and optimization are the key tools. Costs to be minimized include electrical energy to produce air, chemicals, pumping, and carbon addition.

MANIPULATED VARIABLES

There are quite a few variables to manipulate in an activated sludge process. Still, the possibilities to control the plant in a flexible way are quite limited. In many plants, however, there may be a potential to make a better use of the manipulated variables that are available. We categorize them in the following groups:

- Hydraulic, including sludge inventory variables and recirculations.
- Additions of chemicals or carbon sources.
- Air or oxygen supply.
- Pretreatment of influent wastewater.

There are several other manipulated variables in a plant that are related to the equipment and to basic control loops in the process, such as local flow controllers, level controllers, and so on. They are not included in this discussion. Further details are found in the work by Olsson and Jeppsson (7).

A majority of the manipulated variables are meant to change the hydraulic pattern in the plant. The different flow rates will influence the retention times in the different units. Moreover, the clarification and thickening processes are particularly sensitive to the *rate* of change. The hydraulic flows also determine the interaction between different unit processes. We classify the hydraulic manipulated variables into four major groups:

- Variables controlling the influent flow rate.
- Variables controlling the sludge inventory and its distribution, including return sludge flow rate, waste sludge flow rate and step feed control.
- Internal recirculations within the activated sludge system.

- External recycle streams, influencing the interactions between sludge and liquid processes.

Sequential batch reactors (SBR) offer other control possibilities because of the flexibility associated with working in time rather than in space. Some of the features of SBRs are described in some detail in the work by Irvine-Ketchum (8). The Danish Bio-Denitro and Bio-Denipho processes are not the traditional fill-and-draw process but are alternating processes with continuous discharge (9,10).

Some of the control variables are sometimes called *selection mechanisms* for the viable bacteria in the system (11). They are classified in the groups:

- Electron acceptors (e.g. oxygen or nitrate).
- Substrate.
- Thickening or clarification properties.
- Temperature.
- Growth rate.
- Free swimming organisms.

Not all of these can be manipulated. Only the electron acceptors (mostly oxygen) and (to some extent) the thickening and clarification properties can be controlled (by chemical precipitation). Moreover, substrate can be added (e.g. carbon sources).

IMPORTANT CONTROL LOOPS

Out of the many manipulated variables, we will consider a few. Probably the most important is the control of dissolved oxygen (DO) by using the air supply. Other important measures are the sludge inventory, controlled by the return sludge flow rate, and the waste sludge flow rate.

DO Control

The DO supply is a key control variable for the aerator system, both for carbonaceous and nitrogenous organisms. It influences the operating costs significantly, so there is an incentive to minimize the air consumption. From a microbiology point of view the choice of a proper DO set point is crucial, and too little DO will limit the growth of organisms.

From a process control point of view, DO control is known and proved technology. As long as the air supply is variable, the control of DO is quite straightforward. The tuning of the controller has to be different at different loads because of the inherent nonlinearities, and this can be compensated for by self-tuning regulators, prove in full scale for a long time (12). Another way is demonstrated in pilot scale, where a nonlinear controller can be used to operated over a wide range of loads (13).

The oxygen transfer rate determines how fast gaseous oxygen is transferred into DO. This in turn decides that the DO concentration can be manipulated by air flow changes within fractions of an hour. Changes in the oxygen demand, however, may influence the DO concentration within minutes. Not only the total amount of oxygen but its distribution in space and time are essential for the operation. Some of the considerations are:

- *Total air supply*. The total demand of air is basically determined by the total mass of viable organisms in the system multiplied with the biodegradable substrate concentration. The air supply control has an essential impact on the energy cost for the plant.
- *DO set points*. It is important to be able to select the DO set point. However, little is known about the proper choice of it. Generally, under closed loop DO control, it is made constant. There is, however, some evidence that a time variable set point may be advantageous (12).
- *DO spatial distribution (profile)*. It is a great advantage to be able to individually change the air flow rates along the aerator. This makes it possible to obtain an accurate air flow spatial distribution.
- *Pure oxygen addition*. Some aeration systems are supplied with pure oxygen instead of air. Pure oxygen is now and then used in conventional air systems to add extra DO to the system under extreme load situations.
- *Anoxic/aerobic zones*. Because loading conditions may change significantly on a weekly or longer basis, it is of interest to vary the size of the anoxic reactor volume and change the relative retention times between nitrification and denitrification. Such a control is, of course, easy to realize in SBR systems.

Given an adequate controller, it is straightforward to change the DO set point. However, to know which set point is the best from a microbiological point of view is a much more intricate question, and much remains to be studied before the ideal DO control system has been found. Consider an example of a compromise between competing goals within the same plant. The anoxic unit in the activated sludge system, Figure 2, is used to denitrify the nitrate, that is, to reduce the nitrate to free nitrogen by the oxidation of organic carbon. The dominating organisms in the process are the heterotrophs, a proportion of which reduce nitrate in an (ideally) oxygen-free environment. Therefore, the DO level has to be kept at a minimum. If not, the organisms will use oxygen instead of nitrate as the electron acceptor, and the zone will work at least partly like an aerobic reactor.

There are several disturbances that may violate the anoxic condition. The nitrate recycle stream may have a high DO concentration that will rapidly influence the DO concentration of the anoxic zone, thus the denitrification (DN) rate. It is apparent that from a DN point of view that the DO concentration at the outlet of the aerator has to be kept as low as possible, so that no more oxygen than necessary is recirculated. However, for the nitrification taking place in the aerator, a high DO concentration is required. Consequently, the DO concentration at the aerator outlet has to be a compromise between the needs of nitrification and denitrification.

Sludge Inventory Control

The sludge inventory can be controlled primarily by three manipulated variables: the waste-sludge flow rate, the

return-sludge flow rate, and the step-feed flow rates. Manipulation of the waste-sludge flow rate is used to control the total inventory of sludge in the process. Because the total inventory is a function of the total growth rate of organisms, it is used to control the sludge retention time, or the sludge age. It will influence the system in a time scale of several days or weeks.

Manipulation of the return-sludge flow rate is used to distribute the sludge between the aeration basins and the settler units or between the acidogenic and methanogenic reactors in two-stage anaerobic systems. Recycle from the settling stage is an important variable for obtaining the right operating point in the reactors, but seldom useful for the control on an hour-to-hour basis (14). Some systems are supplied with several feeding points for the return sludge. This has a potential for sludge redistribution for certain loads, such as toxic loading.

A *combination* of different recycle may be important. In systems with chemical precipitation, sludge from the secondary settler may be combined with chemical sludge from a postprecipitation settler unit. In that way, the flocculant properties may be influenced, and the chemicals better utilized for phosphorus removal.

By controlling the step feed in an activated sludge plant, the sludge within the aeration basin can be redistributed, given the proper amount of time (15). As a special case of step-feed control, one will obtain a *contact stabilization* structure.

MEASUREMENT AND DIAGNOSIS

Measurement technology is crucial for adequate operation. In particular, on-line analysis is of key interest to control the effluent water quality. It is of great importance to extract maximum information from available instrumentation. Therefore, detection and diagnosis, based on combination of measurements, are models that make up a basis for successful plant operation.

Sensor Technology

Sensor technology is crucial for the success of control and operation, and great progress has been made over the last few years (3). Measurement for control is not necessarily the same as measurements for quality check or for regulatory agencies. The sampling rates have to be consistent with the plant dynamics, and the location of some of the sensors has to be chosen with care, for example, DO probes. Important types of measurements include:

- Flow rates in different plant units.
- Air flow rates and air pressure.
- Sludge flow rates.
- Temperature in water basins and in digesters.
- Gas (e.g., methane) flow rates and temperatures.
- Alkalinity.
- pH.
- DO in different locations.
- Sludge levels.
- Suspended solids in different places.
- BOD, COD, TOC.
- Phosphorus fractions.
- Nitrogen fractions.
- Respiration rate.

There are many instruments that satisfactorily measure flows, levels, pressures, dissolved oxygen in liquids and suspended solids, and dry solids concentrations in sludge. Even if an instrument is deemed to be reliable, proper maintenance is still required. Maintenance requirements will vary greatly between different applications. The development of respirometry has been significant. Within IAWO, task group prepared a scientific and technical report during 1997 (16). Also, advanced N and P analyzers are becoming common in the advanced operation and control of nutrient removal plants (3).

Diagnosis and Estimation

It is clear that more information than the primary sensor signals is required in order to get good operation. The idea of estimation is to mix sensor information with mathematical models. In this way, variables that are not measurable may be reconstructed by combining a mathematical model with sensor information about some other variable. Consider the respiration rate (or the oxygen uptake rate) in an aerator. It can be measured by using respirometers. Alternatively, it can be obtained from measuring the DO concentration and calculation of the DO mass balance. Consider the DO mass balance in a complete mix aerator:

$$\frac{dc}{dt} = d_i \cdot c_i - d \cdot c + k_L a \cdot (c^s - c) - R \qquad (1)$$

where c is the DO concentration; c_i the influent and c^s the saturation DO concentrations, respectively; $k_L a$ the oxygen transfer rate, and R the oxygen uptake rate. The parameters d and d_i are dilution terms. If the oxygen transfer rate $k_L a$ is known, then R can be easily calculated from the DO measurements and the mass balance (equation 1). Usually, however, $k_L a$ must be determined, because it varies slowly with time. Therefore, both the oxygen transfer rate and the respiration rate R to be determined simultaneously from measurements of the DO, preferably under closed loop DO control. This is not trivial, because there is an identifiability problem. A positive change in R can be explained by a similar positive change in $k_L a$ and no unique solution is found (17). However, certain measures can be taken in order to find the parameters uniquely, so that the respiration is found under DO control (13,18,19).

INTEGRATED OPERATION

Because of the combined development of instrumentation, computer technology, operating technology, and the understanding of plant dynamics, the goal is today a plant-wide control. Recognizing the interaction between various unit processes, the overall control of a plant tries to make maximum use of volumes, energy, and other resources.

Because biological nutrient removal (BNR) will become more common, plant complexity will grow. This may be a major driving force for more instrumentation, control, and automation. In a BNR process, there is a delicate balance between carbon, nitrogen, and phosphorus removal. Some important couplings (see Fig. 2) are:

- The phosphorus release has to take place in an anaerobic environment—no oxygen or nitrate can be present. If oxygen-rich water is recirculated, it may impede the P release. If nitrate is allowed to enter the anaerobic zone, the bio-P process is also inhibited.
- Denitrification takes place in an anoxic environment. The various couplings to other unit processes were discussed earlier.
- Both the phosphorus release and the denitrification require a certain carbon concentration. At low loads this may not be sufficient, and an extra carbon source has to be added.

It will become more and more important to understand the impact of load variations. Furthermore, the increased quantities of sludge makes it necessary to develop better control methods for sludge treatment and disposal processes. The drive for a sustainable environment is an incentive to make better use of automation in order to reach such goals.

An efficient operation of a plant has to consider not only the couplings between various unit processes within the plant, but couplings to external processes. Realizing that sewer operation must be coordinated with plant control, there has been a strong development of combined sewer and plant control (20). Disturbances can be detected and controlled as early as possible.

BIBLIOGRAPHY

1. G. Olsson, and B. Newell, *Control of Biological Wastewater Treatment Plants*, 1997, in press.
2. G. Olsson, *Water Sci. Tech.* **28**, 1–7 (1993).
3. IAWO, *Water Sci. Technol.* **33** (1995).
4. G. Olsson, B. Andersson, B.G. Hellström, H. Holmström, L.G. Reinius, and P. Vopatek, *Water Sci. Technol.* **21**, 1333–1345 (1989).
5. G. Olsson, and G. Piani, *Computer Systems for Automation and Control*, Prentice Hall, Englewood Cliffs, N.J., 1992; also in German: *Steuern, Regeln und Automatisteren*, Carl Hanser, München, 1993 (revised and extended revision of the English version).
6. D.E. Seborg, T.F. Edgar, and D.A. Mellichamp, *Process Dynamics and Control*, Wiley, New York, 1989.
7. G. Olsson, and U. Jeppsson, *Forum for Applied Biotechnology (FAB)*, Bruges, Belgium, Sep. 28–30, 1994.
8. R.L. Irvine, and L.H. Ketchum Jr. *CRC Crit. Rev. Environ. Control* **18**, 255–294 (1989).
9. E. Bundgaard, *Proc. of the Int. Workshop on Wastewater Treatment Technology*, Copenhagen, 1988.
10. J. Einfeldt, *Water Sci. Technol.* **25**, 161–168 (1992).
11. W. Gujer, and M. Henze, *Water Sci. Technol.* **23**, 1011–1023 (1990).
12. G. Olsson, L. Rundqwist, L. Eriksson, and L. Hall, in R.A.R. Drake ed., *Advances in Water Pollution Control*, IAWPRC, Pergamon Press, New York, 1985, pp. 473–480.
13. C.F. Lindberg, B. Carlsson, *Water Sci. Technol.* **34**, 173–180 (1996).
14. G. Olsson, *Comprehensive Biotechnology*, Pergamon Press, New York, 1985, pp. 1107–1119.
15. J.F. Andrews, *Water Res.* **8**, 261–289 (1974).
16. H. Spanjers, P. Vanrolleghem, G. Olsson, and P. Dold, *Water Sci. Technol.* **34**, 117–126 (1996).
17. U. Holmberg, in R. Briggs ed., *Advances in Water Pollution Control*, IAWPRC, 1990, pp. 113–120.
18. U. Holmberg, G. Olsson, and B. Andersson *Water Sci. Technol.* **21**, 1185–1195 (1988).
19. B. Carlsson, C.F. Lindberg, S. Hasselblad, and S. Xu, *IAWO 17th Biennial Int. Conf.*, Budapest, Hungary, July 1994.
20. U. Nyberg, B. Andersson, and H. Aspegren, *Water Sci. Technol.* **34**, 127–134 (1996).

WASTEWATER TREATMENT, IMMOBILIZED CELLS

TAKAMITSU IIDA
Niigata University
Niigata, Japan

KEY WORDS

Activated sludge
Adsorption
Biological oxygen demand (BOD)
Entrapment
Immobilization
Immobilized cells
Nitrogen removal
Photo-crosslinked gels
Polyethyleneglycol
Wastewater

OUTLINE

INTRODUCTION

Activated sludge for biological wastewater treatment has been broadly available for a long time. But, water pollution and eutrophication in rivers lakes, and inland seas have resulted from insufficiently treated water. To protect and control these occurrences, it is critical to consider biological oxygen demand (BOD) and nitrogen-phosphorus (N-P) removal in the development of anaerobic–aerobic wastewater treatment systems.

Increasing volumes of wastewater, combined with limited space availability and progressively tightening standards and quality control, require the development of new intensive biotechnologies for wastewater treatment. Recently, there has been a significant increase in the use of immobilization processes for municipal and industrial wastewater treatment. Immobilization processes offer several advantages because of higher volumetric load, increased process stability, compactness, and so on, as shown in Table 1.

The successful application of an immobilized system as biocatalyst relies on the proper choice of the immobilizing materials and reactor type. The most important requirements for wastewater treatment are as follows:

1. High cell mass loading capacity.
2. Simple nontoxic immobilization procedure.
3. High surface-area-to-volume ratio.
4. Mechanical stability.
5. Chemical stability.
6. Reusability.
7. Suitable for conventional reactor systems.
8. Economic feasibility.

An ideal system that possesses all these optimal properties does not exist, and a suitable compromise has to be found.

For wastewater treatment, bioreactors with immobilized microbes have been examined as a combination of activated sludge process for BOD and nitrogen removal and as an anaerobic process for methane production. But, the attempted immobilization to an anaerobic process is scarcely used in a practical scale because the handling of microbes is difficult at the immobilization and the effects of immobilization are not clear (1). The immobilized microbe treatment process clearly is an effective method of aerobic wastewater treatment.

On a laboratory scale, the immobilization of microorganisms has also been applied to the removal of specific compounds such as phenol and its derivatives (2–5), coloring matter (6), and other inorganic compounds such as cadmium (7,8), uranium (9), and hydrogen sulfide (10). Consequently, they can maintain higher biomass concentrations, and a high degree of organic carbon or nitrogen removal, or both, is achieved by these compact systems. Moreover, minimized production of excess sludge is expected by maintaining a low food-to-microorganism ratio.

CLASSIFICATION OF IMMOBILIZATION TECHNIQUES

Whole-cell immobilization is always used in wastewater treatment and may be defined as the physical confinement or localization of intact cells to a certain defined region of space with the preservation of some decomposition activity (11–13). Accordingly, the immobilization technique using biocatalysts is developed to reuse the catalytic activity and for continuous operation in a bioreactor.

The immobilization of microbes for wastewater treatment is broadly classified into four major categories: adsorption of attachment to a surface, entrapment within a porous matrix, containment behind a barrier (semipermeable membrane and microcapsule), and self-aggregation. These techniques can be applied to all wastewater treatments that use a viable cell system.

Cells Attached to a Surface

Attaching cells to a surface is a very popular immobilization technique because it is simple and fast. A suitable adsorbent for attachment should display a high affinity toward microbial cells and yet cause minimal denaturation. A surface-attached process is the immobilization method in which cells are bound to a surface, regardless of the type of binding. Microbial cell walls carry a charge that depends on the pH of wastewater and other environmental conditions. Biofilms resulting from the growth of cells on the surfaces of inorganics such sand or rock are present in conventional industrial processes (14–21), and the self-immobilization of cells (22,23) is a widespread process. But in recent years, intensive treating processes have indicated a need for the development of immobilization for a given process.

Table 1. Advantages and Disadvantages of Immobilized Living Cell Systems

Advantages	Disadvantages
1. Self-proliferation or self-regeneration of immobilized cell systems.	1. The mechanical and chemical stability of some supports materials limits its use (shear, dissolution, and attack by product).
2. Higher productivities because of higher biomass density.	2. Cell growth can destroy the matrix.
3. Tailoring to the most efficient reactor is possible.	3. Diffusion limitations may limit the biodegradation rates.
4. During continuous fermentations, washout can be avoided at high dilution rates.	4. Biocompatibility of immobilization matrices is not sufficient for cells.
5. Better protection from surroundings.	5. Immobilization techniques and materials can be too expensive be used in large-scale processing.

The strength with which the cells are bonded to the support varies, with cell type, support type (24,25), and the variation of operation condition. The industrial use of surface-attached cell systems (biofilm) has been extensively examined in the field of wastewater treatment (26). For strengthening the attachment, porous supports such as porous glass are used (27).

However, in the fluidized-bed type bioreactors for wastewater treatment, inorganic materials are generally not used because of the difficulty in controlling their specific gravity, brittleness, and shortcoming of abrasion-resistance.

Cells Entrapped in a Porous Matrix

The second immobilization technique is entrapment within a porous matrix that is synthesized in situ around the cells. This method is rapidly coming into widespread use (28). The cells are entrapped in a matrix that protects them from the shear field outside the particles. A high degree of cell viability is retained in most entrapment methods. These systems are therefore useful for a variety of growth-related processes. An advantage of the preformed support particles is that they are more resistant to the compression and disintegration in fluidized-bed bioreactors. Cell growth is controlled by the continuous flow of wastewater past the particles (29).

A variety of compounds can be used for cell entrapment, with minimal loss of viability (30). Natural polymers used for immobilization include polysaccharides, such as alginate (31), carrageenan (32), cellulose, and agar, and polypeptides, such as gelatin and collagen. These natural polymers are compatible with microorganisms, but they are not suitable for wastewater treatment because of their assimilation by microorganisms in activated sludge.

Accordingly, only synthetic polymers, such as polyacrylamide (33), poly(vinyl alcohol) (34,35), poly(ethylene glycol), and polypropylene, and hydrophilic photo-cross-linkable resin can be applied to the immobilization of biomass in terms of mechanical strength and durability that are required for industrial usage. For successful immobilization, the hydrogel must be effective in cell viability and function (biocompatible) (36). The method of entrapment can reduce the culture viability. The entrapment in a polyacrylamide gel destroys viability in 50 to 90% of the cells by the toxic acrylamide monomers (37).

Several reviews (28,38,39) offer a thorough treatment of methods and application in immobilized cells. Examples of the various types of polymers are shown in Table 2.

Cells Contained Behind a Barrier

The third major category of cell immobilization is containment behind a barrier. When cell separation from the effluent is required, these systems are highly useful (56).

The barrier that immobilizes cells can be as simple as the liquid–liquid phase interface between two miscible fluids. Mohan and Li (57) immobilized the denitrifying bacteria, *Micrococcus denitrificans*, by emulsifying the cell suspension with surfactant into a hydrocarbon solvent. Cells in the surfactant-enclosed aqueous droplets did not leak from aqueous phase through the organic phase to the cell-free aqueous phase surrounding the droplets. Because the cells are retained behind a semipermeable wall or membrane, this type of system is very similar to microcapsulation in an engineering sense. Nutrients (mainly BOD components) are supplied to, and products are removed from, the cell mass by diffusion. In wastewater treatment, the barrier loses its separation activity during long operations.

Self-Aggregating Cells

Cells naturally aggregate or flocculate and produce extracellular polysaccharides. The large size of the aggregates makes their use possible in reactors designed for immobilized cells, such as packed-bed or fluidized-bed reactors. Artificial flocculating agents or cross-linking agents may be added to enhance the process of aggregation for cells that do not naturally flocculate. Similar mechanisms govern both artificially and naturally induced flocculation (58). The durability of the aggregate will depend on its hydrodynamic interactions with the environment. Many cells produce extracellular polysaccharides, some of which can act to promote the cohesion of a biofilm. These substances might be important determinants of the diffusive resistance in these films.

INDUSTRIAL APPLICATIONS OF IMMOBILIZED CELLS

Much research has focused on industrial applications using immobilized cell technology for wastewater treatment. In an industry setting, entrapment and adsorption are the predominant techniques used for wastewater treatment. The nitrifying process is described in this section, examples include entrapment using PBG gels and adsorption using photo-cross-linked resins.

Gel Entrapment of Microbial Cells

The entrapment of microbes is the most popular immobilization technique for living cells in wastewater treatment. The BOD sensor was developed to abbreviate the long measurement time by entrapping the activated sludge (59,60). In wastewater treatment, immobilization processes should be mild to minimize the damage to the living cells. Entrapping materials must also have proper permeability to allow sufficient diffusion and transport of oxygen and essential nutrients, metabolic waste, and secretory products across the hydrogel network at supplemental culture.

Considering the use of synthetic polymers serving as the synthetic extracellular matrixes for cells, it is important to understand the transport properties through these gels to predict whether nutrients can diffuse into the matrix. Supplemented culture media provide the cells with essential nutrients for growth and development as well as a means of oxygen and metabolic product transport while in vitro (61,62). Traditionally, mass transfer behavior in the gel matrix has often been recognized as a rate-limiting factor in an entrapping cell system. Some of the immobilizing matrixes are highly hydrated, allowing almost unhindered diffusion for products and substrates with molec-

Table 2. Environmental Application of Gel-Entrapped Microorganisms

Substrate	Microorganism	Entrapment method	Reference
Ammonia	*Nitrosomonas europaea*	Ca-alginate	40
		κ-carrageenan	41
	Nitrosomonas europaea	Polyelectrolyte	42
	Paracoccus denitrificans	Complex-stabilized Ca-alginate	
	Thiosphaera pantotropha	Agarose	43
	Scenedesmus quadricauda	κ-Carrageenan	44
Dairy effluents	*Candida pseudotropicalis*	Alginate	45
Nitrate	*Anasistis nidulans*	Collagen	46
	Denitrifying sludge	PVA	34
	Pseudomonas denitrificans	Ca-alginate	47,48
	Thiobacilus denitrificans	Agar/agarose	49
Nitrite	*Nitrobacter agilis*	Ca-alginate	50
		Carrageenan	51
	Nitrobacter sp.	PVA	35
Nitrogen	*Anabaena* CH3	Ca-alginate	52
	Nitrifiers	Polyethyleneglycol resin	53
Phenol	*Candida tropicalis*	Alginate	54
	Pseudomonas sp.	Ca-alginate, polyacrylamide hydrazide PVA-boric acid-Ca-alginate	55,36

ular weights less than 10,000 Da (36). The effective diffusivity of oxygen, glucose, and various ions in gel has been studied by several groups (63–65).

For industrial applications, microbes must be cultivated for entrapment and storage to keep the activity of immobilized microbes after immobilization.

Entrapping Materials. Mass transport through gels is governed by the size and shape of the pores and solute molecules as well as by the nature of the interaction between solute molecule and gel material. The value of the equilibrium coefficient reflects the relative affinity of the solute for the gel relative to its affinity for the surrounding medium. The species concentration in the gel at equilibrium is obtained by mass balance using the known values for the reservoir and gel volumes. The value of K, the partition coefficient, is calculated by the following equation:

$$K = C^{\text{gel}}/C^{\text{bath}} \quad (1)$$

The average pore diameters of gels were estimated from the diffusion of solutes and the hydrodynamic diameter of solute (60).

Biomass decay is an important parameter in the dynamics of the immobilized microbe system. *Nitrosomonas europaea* cells entrapped in a gel can lose their viability during cultivation; they die first in the center of a gel bead but also will die in the outer layers at low substrate loadings. If the substrate concentration is raised, the number of viable microbes near the surface of the gel increases again. The fluorescein diacetate and lissamine-green staining techniques were useful in distinguishing between viable and dead cells in gel beads (66).

Successive nitrification and denitrification are effective biological processes for the removal of nitrogen in the form of ammonia, nitrite, and nitrate from aqueous waste. In this biological nitrogen-removal system, nitrification is generally the rate-limiting step because of the extremely low growth rate of obligate autotrophic nitrifying bacteria. The use of immobilized cells in nitrification overcomes this disadvantage and shortens greatly the hydraulic retention time (HRT) in continuous nitrification.

In a synthetic resin gel such as polyacrylamide, the gel strength decreased about 50 to 60% in the presence of microbes. But the density of microbes in the gel did not influence its strength (67). The immobilization of nitrifying bacteria by using urethane prepolymers of the hydrophilic type exhibited a drastic improvement in the activity yield (68). These immobilized cells were prepared by entrapment using alginate (69–71), polyelectrolite complex (72), polyethyleneimine (73), photo-cross-linkable resin (74), poly(vinyl alcohol) (75), polyethylene (76), and poly(ethylene glycol) (PEG) (77).

Characteristics of Microbes Entrapped in Polyethyleneglycol. The PEG gel was prepared by radical reaction under existing microbes to exhibit about 3 kg/cm^2 of gel strength (78). These gels are nontoxic for microbial cells and show viability. Furthermore, this method is easy industrialized for a large-scale production.

Microorganisms such as nitrifying bacteria grow near the surface of gels (about 60-μm depth). After about a month's cultivation, they have constant activity. The rate of oxygen consumption the parameter of activity, is 60 mg O_2/L gel/h at initially and 800 mg O_2/L gel/h at the end of precultivation. At the steady state, the concentration of microbes in the gel carrier is increased to contain about 10^{10} cells per milliliter of carrier. On the other hand, the concentration of nitrifying microbes in the activated sludge is usually 10^3 to 10^5 cells/mL. Figure 1 depicts nitrifying pellets (about 3 mm^3) and a cross-section of pellets and bacteria in colony.

Nitrification and denitrification of entrapped microbes were scarcely influenced by the operation temperature, and the activity of nitrification was about three times higher than that of the activated sludge, as shown in Figure 2. In Figure 2, the reaction rates are exhibited on the basis of suspended solid for comparing the activity of im-

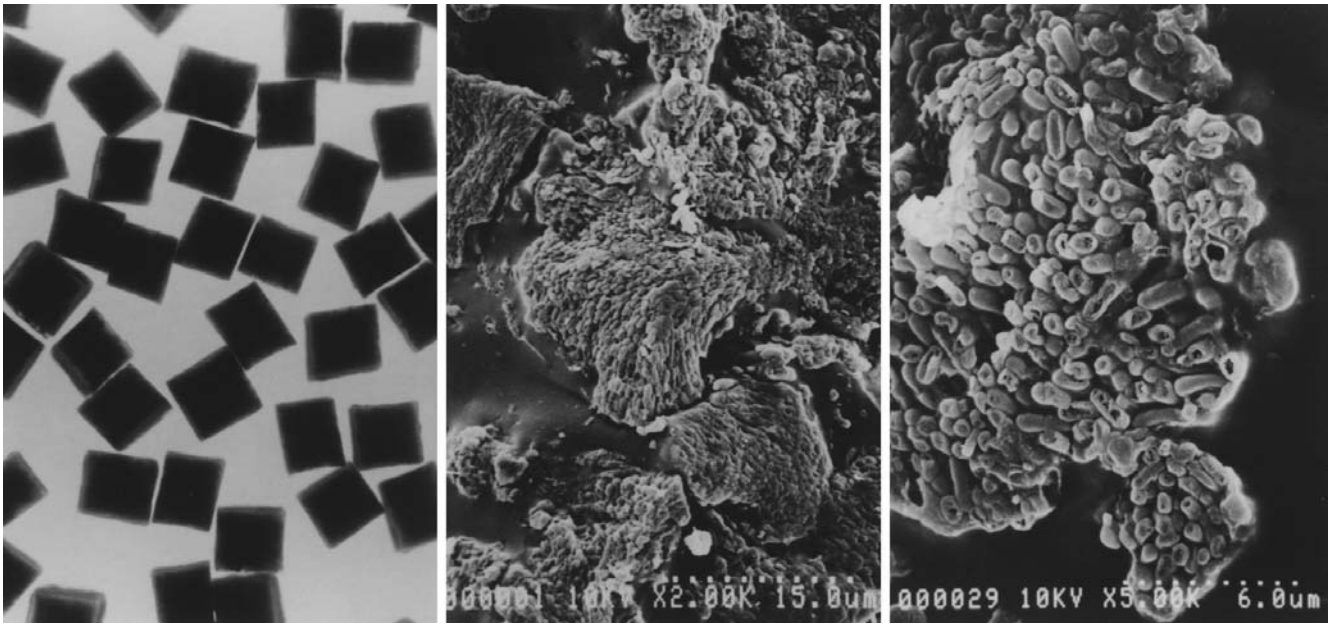


Figure 1. Photographs of nitrifying pellets. Cross-section of pellets and bacteria in colony. (a) Nitrifying pellets of 3-mm cubic gel; (b) Cross-section of pellet by electron scanning micrograph; (c) Bacteria in colony under electron scanning micrograph.

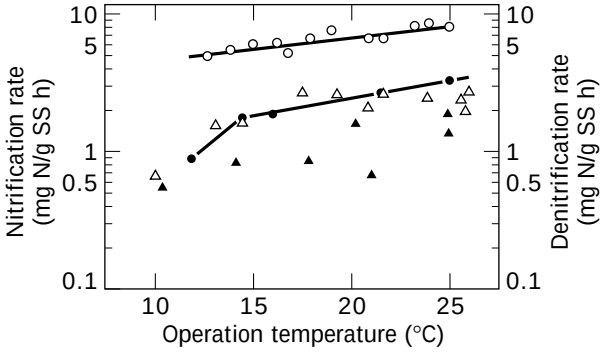


Figure 2. Effect of operation temperature on nitrification rate and denitrification rate of entrapping method against the activated sludge method. Nitrification: ○ = immobilized: △ = activated sludge. Denitrification: ● = immobilized: ▲ = activated sludge. Range of BOD loading: immobilized = 0.1–0.15 (kg/kg SS day). Activated sludge = 0.02–0.08.

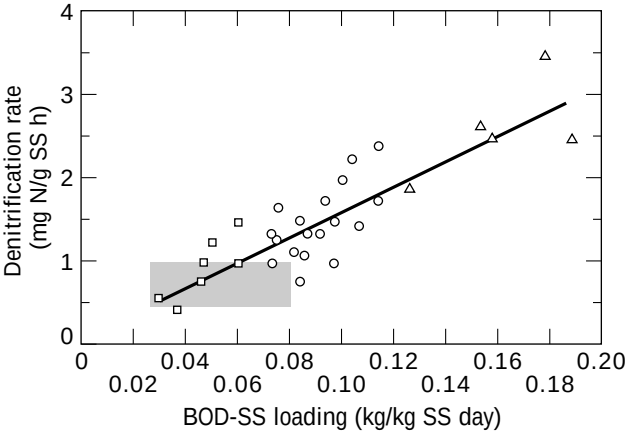


Figure 3. Relationship between denitrification rate and BOD-SS loading in the entrapping method. BOD loading (kg/kg SS day) at 15 °C: □ = 0.02–0.07; ○ = 0.07–0.12; △ = 0.12–0.19. Shaded part indicates the range of the activated sludge method.

mobilization with that of activated sludge. The nitrification rate of precultured immobilized gels was about 100 mg N/L gel/h at 13 °C. Accordingly, the system using entrapped microbes showed a high denitrification rate owing to high BOD loading resulting from short retention time in the nitrification tank (Fig. 3).

Continuous Wastewater Treatment by Entrapped Microbes (79). Figure 4 shows a schematic diagram of a pilot plant using PEG-entrapped gels. It is composed of an initial and final sedimentation tank, a denitrification tank, and a nitrification tank. The volume ratio of denitrification to nitrification is 1.5/1.0, and the gel separator (the net of 1.5-mm mesh) in the nitrification tank is installed for preventing gel bead washout. The mesh size of the gel separator should be half the gel diameter for sufficient separation. Wastewater treatment was recycled between the denitrification tank and the nitrification tank. Retention times are 4.8 and 3.2 h for denitrification and nitrification, respectively. The average total nitrogen removal is represented as

$$\text{Total N removal (\%)} = R/(R + 1) \times 100 \qquad (2)$$

with R as the recycling ratio (−). The total N removal is

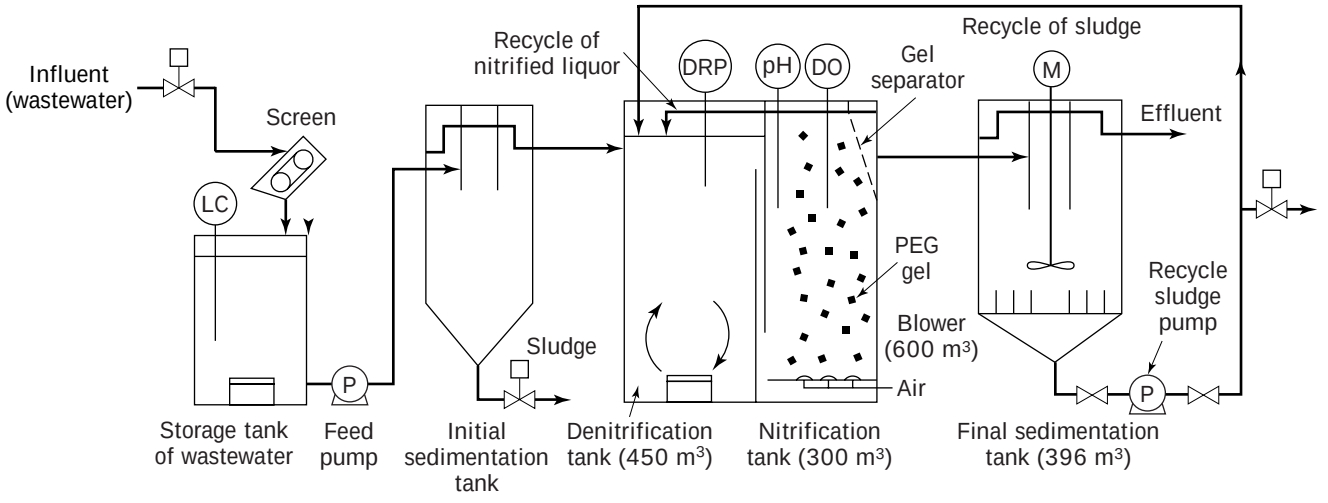


Figure 4. Schematic diagram of wastewater treatment plant with a capacity of 2,250 m^3/day (max, 3,000 m^3/day). DO = dissolved oxygen electrode; ORP = oxidation–reduction potential; pH = pH electrode; LC = level control; P = pump; M = mixer.

increased according to the enlargement of the recycling ratio, but the dissolved oxygen in the denitrification tank is also increased. Consequently, recycling is generally operated at a 1:2 recycling ratio to maintain the anaerobic condition of denitrification and compensate for the increased power cost in nitrification. Gel beads are fluidized by the air blowing at the bottom of nitrification tank.

The packing ratio of entrapped gels in the reactor is usually 5:15 volumetric percent depending on the treatment. In the entrapment method, gels are added at 7.5 volumetric percent to the nitrification tank. The floating carriers are distributed homogeneously. In the entrapment method, specialized equipment and sufficient time are needed to prepare an adequate number of gels for large-scale wastewater treatment.

The result of continuous wastewater treatment in the PEG entrapping method is shown in Figure 5, and the operation conditions of the plant are shown in Table 3. Even in winter nitrification was thoroughly carried out, and BOD as well as suspended solids in the effluent were lower than 10 mg/L. This method needs about 8 h of retention time to eliminate 70% of nitrogen in wastewater. After continuous operation for 2 y, it was concluded that gel beads must have above 1.5 kg/cm^2 compressive strength; therefore, the initial compressive strength of gel absent from microbial cells needs to be over 2.7 kg/cm^2.

For the nitrification, photo-cross-linkable resin was used to obtain an activity of 1.5 kg of N/m^3 gel/day for continuous wastewater treatment (80). On the other hand, the drop in denitrifying activity at the immobilization was restored in the medium containing nitrite or nitrate with sodium acetate as a hydrogen donor.

Industrial Application of Adsorption Method

Systems in which cells are immobilized by adsorption are popular because of the ease of this type of immobilization. The strength with which the cells are bonded to the support varies with cell type and support type (81,82).

The adsorption method differs from the entrapment method, advantages include easy gel formation (no consideration of microbes at cross-linking reaction), mass transfer of gel barrier, good storage stability before usage, and low transportation cost of gels at the industrial scale. A disadvantage is the long induction period for steady wastewater treatment.

The growth rate of nitrifying bacteria is about 10 times less than that of general microbes. The attachment of biomass on the gel surface occurs during the start-up phase. After this phase, the reactor is operated continuously at different hydraulic and surface loading rates. Accordingly, the gels must have high strength and good cell attachment on their surfaces.

A positively charged macroporous cellulose carrier is used for immobilization of nitrifying bacteria by the adsorption method (83). This carrier promotes stable adhesion of bacteria onto the carrier walls and allows transfer of nutrients into the carrier. A porous cellulose carrier was produced from cellulose with a polyethyleneimine coating to form a stable cross-linked structure. Similarly, polyurethane foam (84) and other carriers (85,86) have been used for wastewater treatment.

Advantages of Adsorption Using Photo-Cross-Linked Gels. To use immobilized microbes in wastewater treatment, the preparative method, especially the support material used for the preparation, must be selected from the standpoint of its economy and feasibility on a large-scale operation. For these reasons, photo-cross-linkable resins are promising support materials for the immobilization of microbes.

Although this immobilization process may have various conceivable features, its features are summarized as follows (87):

1. The immobilization procedure (gel formation) is simple because of ultraviolet irradiation.

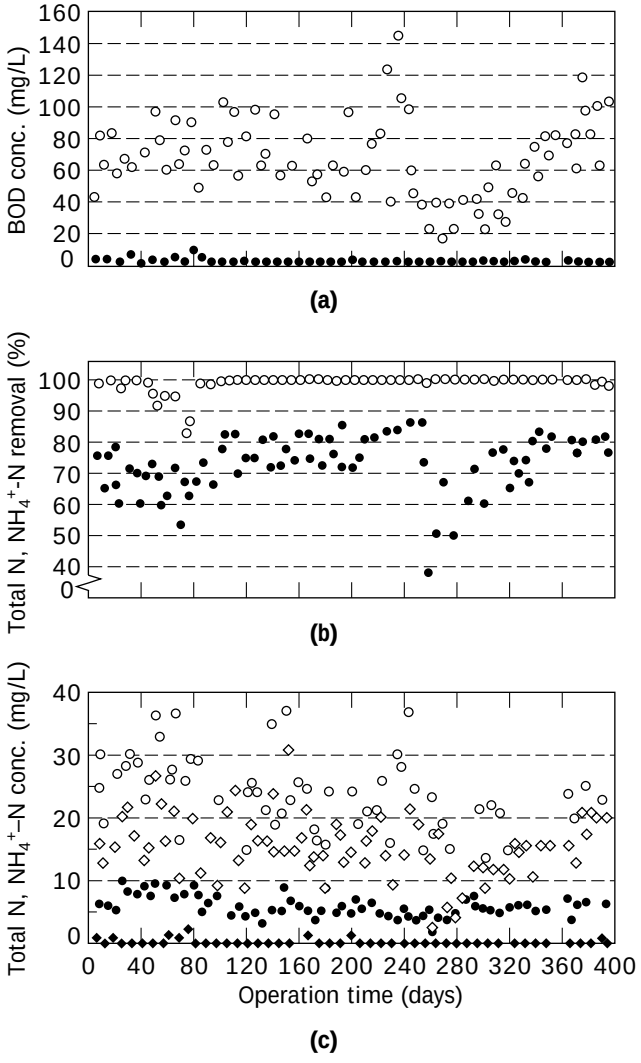


Figure 5. Continuous wastewater treatment using entrapped microbes. (**a**) Influent (mg/L): ○ = BOD; effluent (mg/L) ● = BOD. (**b**) Removal (%): ● = total N removal (%): ○ = NH_4^+-N. (**c**) Influent (mg/L): ˆ = total N, ◇ = NH_4^+-N; effluent (mg/L): ● = total N, ◆ = NH_4^+-N.

2. The gel materials have great mechanical strength.
3. The surface of gel beads can be regulated by the properties of photo-cross-linkable resin.
4. The manufacture of photo-cross-linkable resins and the process of gel formation can be easily industrialized.

The photo-cross-linkable resins usually have a main chain of PEG, a mixture of PEG and poly(propylene glycol) (PPG), or poly(vinyl alcohol) (PVA) with photo-cross-linkable ethylenic double bonds. By irradiation of ethylenic double bonds with UV eight, polymerization is induced, forming three-dimensionally-cross-linked gels.

Surface Structure of Gels for Good Cell Adsorption. The sludge used is taken from the recycle line of an activated sludge treatment plant treating wastewater, because organic nitrogen and ammonia nitrogen are present in wastewater. Acclimatization of the sludge is carried out in a batch reactor. The carrier elements are made of hydrophilic photo-cross-linkable resin with a 4-mm bead diameter.

Because the nitrifying bacteria do not flocculate differently from the activated sludge, it is difficult to attach the nitrifying bacteria on the surface of the carrier. The adsorption rate of nitrifying bacteria is maximum at pH 4.8, and nonbiological factors are important at the initial stage of attachment (88). Fluorescence studies of the adsorbing surface clarified that the affinity to water and ζ potentials of the surfaces are important factors for adsorption (89). Among the synthetic polymers, photo-cross-linkable resin has largely resolved the problems with inorganic carriers described in the Introduction. Its surface structure may be suitable for attachment of microbes, controlling the properties of photo-cross-linkable resins.

Several factors affect cell affinity and behavior on hydrogel: the general chemistry of the monomers and the cross-links, hydrophilic and hydrophobic properties, and surface properties involving charged functional chemical groups. Physical properties of the hydrogel also govern the adhesion affinity; therefore, altering pore size and network structure can convert cell adhesion as well as morphology and function.

The attachment of biomass on the charged gel surface occurs during the start-up phase. Designing gel carriers using photo-cross-linkable resin has resulted in good cell growth on their surface with good mechanical strength (90). The surface of PEG-PPG-type hydrophilic resin is most hydrophobic and has an excellent affinity to microbes. This phenomenon is practically confirmed in continuous denitrification, as shown in Table 4.

When observing the surface of PEG-PPG-type photo-cross-linked resin and the results of contact angle, the phase separation of the gel in which hydrophobic domains of about 1-μm diameter separated on the hydrophilic phase of PEG was seen on SEM of the surface (Fig. 6) and with the optical microscope. From these results, the surface of hydrophobic domains formed on hydrophilic phase, the so-called macrophase separation of hydrophilic gel, may be best for the attachment and growth of nitrifying microbes.

The structure of the gel bead is controlled with a combination of PEG and PPG, PVA type, and PEG type. Beads with a diameter of about 4 mm and compressive strength of more than 20 kg/cm^2 can be produced by immobilization (91).

Continuous Wastewater Treatment. Continuous wastewater treatment can be carried out by a process using gel for microbial adsorption, as shown in Table 5 (92–94). Removal performance of the pilot plant with a retention time of 8 to 9 h in the bioreactor is obtained as follows. The average total N of the influent is 31.8 mg/L, and that of the effluent total N is 8.3 mg/L. The average P of the influent is 2.57 mg/L, and that of the effluent total P is 0.60 mg/L.

The nitrification rate of immobilized gels is of zero-order

Table 3. Operation Condition of Plant

Item	Run 1	Run 2	Run 3
Operation period	18	11	18
Infuluent (m^3/day)	2,720	2,280	1,350–3,150
Retention time (h)	6.6	7.9	5.7–13.3
Aeration (Nm^3/h)	610	670	800
Recycling ratio (−)	2.8	2.7	2.0–7.0
MLSS (mg/L)	2,800	1,840	1,880
BOD-SS loading (kg/kg day)	0.070	0.149	0.0965–0.302
BOD loading (kg/m^3 day)	0.196	0.275	0.182–0.570
Total N and SS loading (kg/kg day)	0.0285	0.0535	0.0226–0.0772
Total N loading (kg/m^3 day)	0.0793	0.0980	0.0426–0.146
Recycling ratio of sludge (%)	64	64	47–109
Temp. of water (°C)	23.5	14.0	14.2
Total SRT (day)	14.6	9.0	8.8
Aerobic SRT (day)	5.8	3.6	3.5

Notes: Run 3 = Variation test of loading. MLSS = Mixed liquor suspended solid; BOD = Biological oxygen demand; SS = suspended solid; SRT = sludge retention time.

reaction in which the produced NO_z-N increases directly with the reaction time, and the nitrification rate is about 2.5 kg NO_k/m^2 gel/day, as calculated in Figure 7. In the relationship between the dissolved oxygen (DO) and the nitrification rate coefficient (K_M) of gels, K_M largely depended on the DO. It is considered that the oxygen diffusion through the biofilm on the surface is the rate-determining step for nitrification. The phosphorus uptake increased in proportion to the denitrified nitrogen, and the weight ratio of phosphorus to nitrogen is about 0.75. From these facts, nitrifying microbes are suggested to be phosphorus-accumulating microorganisms (95,96).

Water qualities and flow diagram for metal-finishing wastewater are shown in Table 5 and Figure 8, respectively. Since 1997, a denitrification plant using these gels (nitrifying tank volume, 10,000 m^2) has been operating in Tokyo (Fig. 9).

EVALUATION OF IMMOBILIZATION

The development of processes using immobilized microorganisms has made it possible to eliminate BOD and nitrogen in half the time (6 to 8 h), compared with conventional activated sludge. In addition, these processes maintained constantly high activity during winter. The treatment cost containing the depreciation is 80 to 90% of the activated sludge.

Table 4. Comparison of PEG-PPG/PVA Type Gels and PVA Type Gels

Species of gels	PEG-PPG/PVA type	PVA type
DO (mg/L)	2.8	6.0
NH_3-N value added (kg/m^3 day)	0.29	0.21
Nitrification velocity (kg/m^3 day)	0.21	0.15
Cells attached (mg/L)	6,920	1,620

To gain the advantages of immobilization, reductions are needed in the immobilized gel costs, particularly for large-scale operations, which include raw materials costs, immobilization facility, transportation cost, storage cost, and so on.

Wastewater treatment processes using immobilized

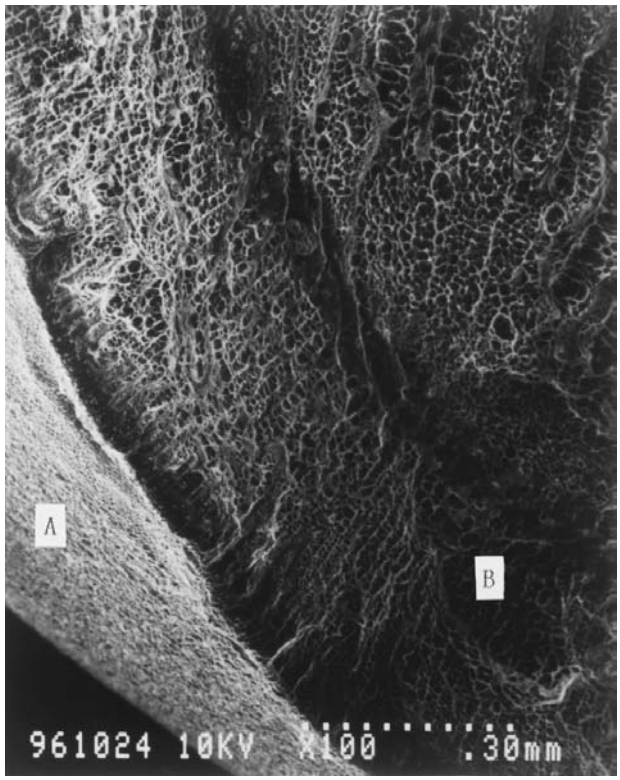


Figure 6. Scanning electron micrograph of photo-cross-linked resin gel. A = Gel surface. Convex is formed by phase separation of hydrophilic and hydrophobic structure. B = Cross-section of gel.

Table 5. Operation Conditions of the Pilot Plant

Item	Step 1	Step 2	Step 3
Influent flow rate (m^3/d)	25.0	25.0	28.6
HRT (h)	8.7	9.0	8.0
Anaerobic tank (m^3)	1.0	1.5	1.5
SRT (day)	16	11–14	8–11
Aerobic SRT (day)	5.7	3.7–4.7	2.7–3.7
Gel diameter (mm)	5.0	5.0	4.2
Charge ratio (%)	20	20	16 -n
DO (mg/L)	1.8–5.0	2.0–5.8	2.8–7.8
Influent BOD (mg/L)	49–151	45–110	48–118
Total N (mg/L)	14–34	13–47	16–45
Total P (mg/L)	1.2–2.9	2.0–5.8	2.8–7.8

HRT = Hydraulic retention time; SRT = sludge retention time; DO = dissolved oxygen.

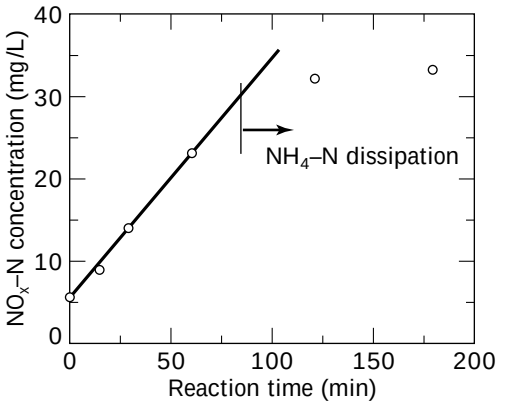


Figure 7. Nitrification rate of adsorpted gels.

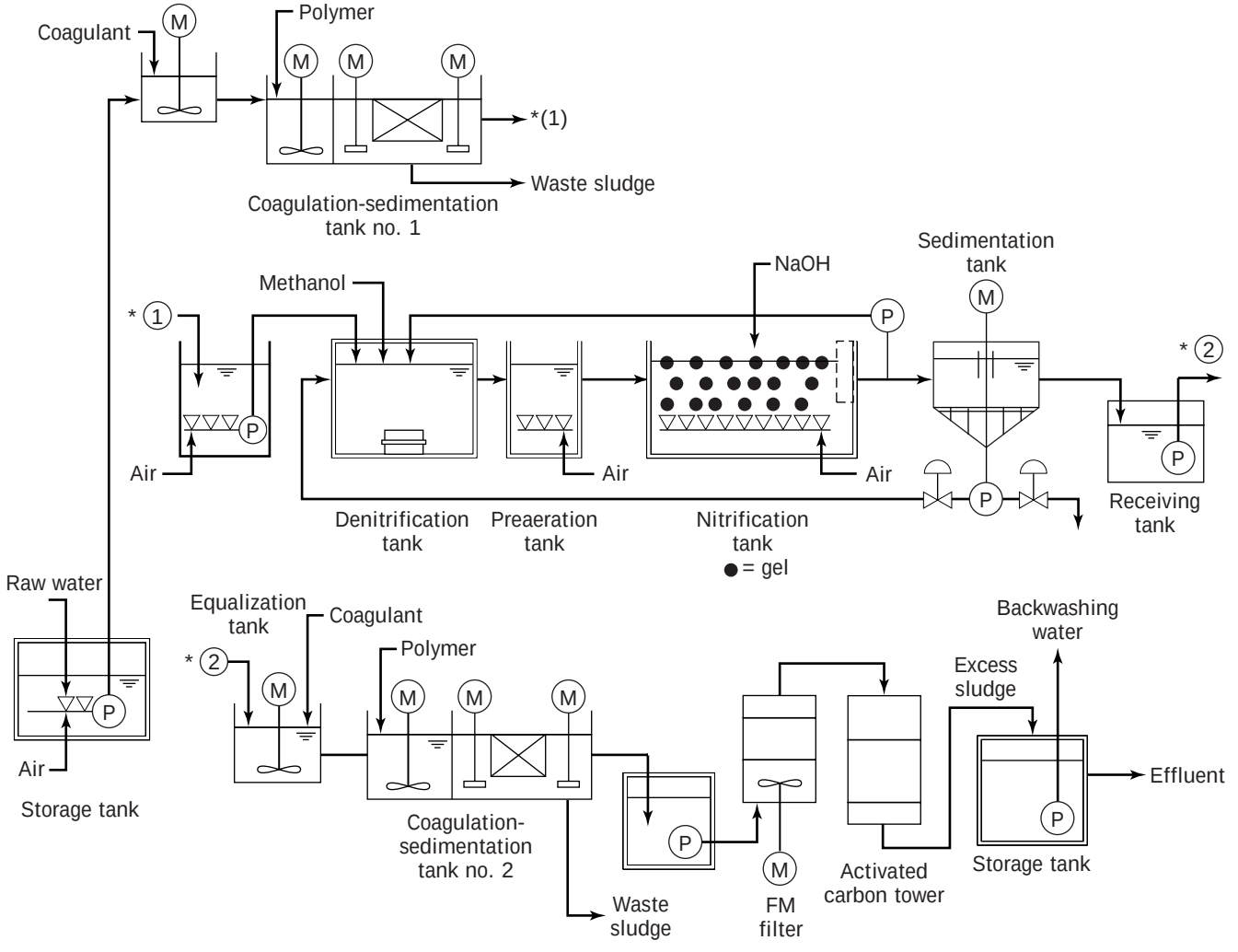


Figure 8. Flow diagram of metal finishing wastewater treatment process.

Figure 9. Photograph of denitrification plant using charged photo-cross-linked resin gels in Tokyo. Nitrifying tank volume, 10,000 m^3, Packing ratio of gel, 10%.

cells are still under development and will be further improved in the future. Particularly, it is expected that carrier gels will regulate their structure and form a particular reaction site in gels or on the surface of gels independent of the surroundings.

BIBLIOGRAPHY

1. J. Poels, P. van Assche, and W. Verstraete, *Biotechnol. Lett.* **6**, 747–752 (1984).
2. H. Bettman and H.J. Rehm, *Appl. Microbiol. Biotechnol.* **22**, 389–393 (1985).
3. A.M. Anselmo, M. Mateus, J.M.S. Cabral, and J.M. Novais, *Biotechnol. Lett.* **7**, 889–894 (1985).
4. H. Keweloh and H. Gruetzmacher, *Appl. Microbiol. Biotechnol.* **31**, 383–389 (1989).
5. J.E. Lin and H.Y. Wang, *J. Ferment. Bioeng.* **72**, 311–314 (1991).
6. T. Ogawa and C. Yatome, *Bull. Environ. Contam. Toxicol.* **44**, 561–566 (1990).
7. L.E. Macaskie and A.C.R. Dean, *Environ. Technol. Lett.* **5**, 177–186 (1984).
8. L.J. Michel, L.E. Macaskie, and A.C. Dean, *Biotechnol. Bioeng.* **28**, 1358–1365 (1986).
9. L.E. Macaskie and A.C.R. Dean, *Enzyme Microbiol. Technol.* **9**, 2–4 (1987).
10. K.L. Sublette, *Appl. Biochem. Biotechnol.* **20/21**, 675–686 (1989).
11. F.K. Steven, B.L. Shari, and R.R. Channing, *Chem. Eng. Sci.* **40**, 1321–1354 (1985).
12. I. Chibata and T. Tosa, *Appl. Biochem. Bioeng.* **4**, 1–9 (1983).
13. J. Klein and F. Wagner, *Appl. Biochem. Bioeng.* **4**, 11–51 (1983).
14. B. Atkinson and H.W. Fowler, *Adv. Biochem. Eng.* **3**, 221–277 (1974).
15. W.G. Characklis and K.E. Cooksey, *Adv. Appl. Microbiol.* **29**, 93–138 (1983).
16. M.S. Switzenbaum, *Enzyme Microb. Technol.* **5**, 242–250 (1983).
17. M. Henze and P. Harremoes, *Water Sci. Technol.* **15**, 1–101 (1983).
18. G. Bitton and K.C. Marshall, *Adsorption of Microorganisms to Surfaces*, Wiley, New York, 1980.
19. K. Elliot, M. O'Connor, and J. Whelan, *Adhesion and Microorganism Pathogenicity*, Pitman Medical, London, 1981.
20. R.C.W. Berkeley, J.M. Lynch, L. Melling, P.R. Rutter, and B. Vincent eds., *Microbial Adhesion to Surfaces*, Ellis Horwood, Chichester, U.K., 1980.
21. D.C. Ellwood, J. Melling, and P. Rutter, *Adhesion of Microorganisms to Surfaces*, Academic Press, London, 1979.
22. B. Atkinson and I.S. Daoud, *Adv. Biochem. Eng.* **4**, 41–124 (1976).
23. B. Metz and N.W.F. Kossen, *Biotechnol. Bioeng.* **19**, 781–799 (1977).
24. M. Fletcher, M.J. Latham, J.M. Lynch, and P.R. Rutter eds., *Microbial Adhesion to Surfaces*, Ellis Horwood, Chichester, U.K., 1980.
25. S.L. Daniels, *Dev. Ind. Microbiol.* **13**, 211–253 (1972).
26. B. Atkinson and I.S. Daoud, *Adv. Biochem. Eng.* **4**, 41–124 (1976).
27. R.A. Messing, C. Ghommidh, and J.M. Navarro, *Biotechnol. Bioeng.* **24**, 1991–1999 (1982).
28. S. Birnbaum, P.O. Larsson, and K. Mosbach, in W.H. Scouten ed., *Solid Phase Biochemistry: Analytical and Synthetic Aspects*, Wiley, New York, 1983.
29. G.M. Black, C. Webb, T.M. Matthews, and B. Atkinson, *Biotechnol. Bioeng.* **26**, 134–141 (1984).
30. I. Chibata and T. Tosa, *Annu. Rev. Biophys. Bioeng.* **10**, 197–216 (1981).
31. J. Tramper and A.W.A. de Man, *Enzyme Microb. Technol.* **8**, 472–476 (1986).
32. J. Tramper and D.R.J. Grootjen, *Enzyme Microb. Technol.* **8**, 477–480 (1986).
33. T. Sumino, H. Nakamura, and N. Mori, in A. Tanaka, T. Tosa, and T. Kobayashi eds., *Industrial Application of Immobilized Biocatalysts*, Dekker, New York, 1992, pp. 377–391.
34. W.R. Vieth and K. Venkatasubramanian, in *Immobilized Microbial Cells*, ACS Symposium Series 106, American Chemical Society, Washington D.C., 1979, p. 1.
35. H. Bettmann and H.J. Rehm, *Appl. Microbiol. Biotechnol.* **22**, 389–393 (1985).
36. H. Tanaka, M. Matsumura, and A.I. Veliky, *Biotechnol. Bioeng.* **26**, 53–58 (1984).
37. N.G. Starostina, K.A. Lusta, and B.A. Fikhte, *Eur. J. Appl. Microbiol. Biotechnol.* **18**, 264–270 (1983).
38. I. Chibata and L.B. Wingard, *Appl. Biochem. Bioeng.*, vol. 4, *Immobilized Microbial Cells*, Academic Press, New York, 1983, p. 355.
39. S. Fukui and A. Tanaka, *Annu. Rev. Microbiol.* **36**, 145–172 (1982).
40. C.G. van Ginkel, J. Tramper, K.Ch.A.M. Luyben, and A. Klapwijk, *Enzyme Microb. Technol.* **5**, 297–303 (1983).
41. R.H. Wijffels and J. Tramper, *Appl. Microbiol. Biotechnol.* **32**, 108–112 (1989).

42. E. Kokufuta, M. Shimohashi, and I. Nakamura, *J. Ferment. Technol.* **65**, 359–361 (1987).
43. C.M. Hooijmans, S.G.M. Geraats, E.W.J. van Neil, L.A. Roberttson, J.J. Heijnen, and K.Ch.A.M. Luyb, *Biotechnol. Bioeng.* **36**, 931–939 (1990).
44. P. Chevalier and J. de la Noue, *Enzyme Microb. Technol.* **7**, 621–624 (1985).
45. P. Chevalier and J. de la Noue, *Biotechnol. Lett.* **7**, 395–400 (1985).
46. S.S. Marwaha, J.F. Kennedy, P.K. Khanna, H.K. Tewari, and A. Redhu, in J.A.M. de Bont, J. Visser, B. Mattiasson, and J. Tramper eds., *Physiology of Immobilized Cells*, Elsevier, Amsterdam, 1990, p. 265.
47. K.C. Chen and Y.F. Lin, *Water Sci. Technol.* **28**, 159–164 (1993).
48. I. Nilsson, S. Ohlson, L. Haggstrom, N. Molin, and K. Mosbach, *Eur. J. Appl. Microbiol. Biotechnol.* **10**, 261–274 (1980).
49. A.B. Salleh and M. Aminuddin, *Pertanika* **4**, 21–24 (1981).
50. J. Tramper and A.W.A. de Man, *Enzyme Microb. Technol.* **8**, 472–476 (1986).
51. T. Willke and K.D. Vorlop, *Symposium Immobilised Cells: Basic and Applications, The Working Party Applied Biocatalysis of the European Federation of Biotechnology*, Noordwijkerhout, The Netherlands, November 26–29, 1995.
52. C.M. Lee, C. Lu, Y.H. Yin, and P.C. Chen, *Symposium Immobilised Cells: Basic and Applications, The Working Party Applied Biocatalysis of the European Federation of Biotechnology*, Noordwijkerhout, The Netherlands, November 26–29, 1995.
53. K. Tanaka, T. Sumino, H. Nakamura, and H. Emori, *Symposium Immobilised Cells: Basic and Applications, The Working Party Applied Biocatalysis of the European Federation of Biotechnology*, Noordwijkerhout, The Netherlands, November 26–29, 1995, pp. 622–632.
54. J. Klein and H. Eng, *Biotechnol. Lett.* **1**, 171–176 (1979).
55. K.Y.A. Wu and K.D. Wisecarver, *Biotechnol. Bioeng.* **39**, 447–449 (1992).
56. A.P. Jarvis and T.A. Grdina, *BioTechniques* **1**, 22/24–27 (1983).
57. R.R. Mohan and N.N. Li, *Biotechnol. Bioeng.* **17**, 1137–1156 (1975).
58. M.W. Tenny and F.H. Verhoff, *Biotechnol. Bioeng.* **15**, 1045–1073 (1973).
59. T. Iida, H. Izumida, Y. Akagi, and M. Sakamoto, *J. Ferment. Bioeng.* **75**, 28–31 (1993).
60. H.L. Rebecca, D.H. Altreuter, and F.T. Gentile, *Biotechnol. Bioeng.* **50**, 365–373 (1996).
61. J. Kein and G. Manecke, *Enzyme Eng.* **6**, 181–189 (1982).
62. D.A. Sirotti and A. Emery, *Biotechnol. Bioeng.* **25**, 1773–1779 (1983).
63. K. Williamson and P.L. McCarty, *J. Water Pollut. Control. Fed.* **48**, 281–296 (1976).
64. J.V. Matson and W.G. Characklis, *Water Res.* **10**, 877–885 (1976).
65. M. Onuma and T. Omura, *Water Sci. Technol.* **14**, 553–568 (1982).
66. E.J.T.M. Leenen, A.A. Boogert, A.A.M. van Lammeren, J. Tramper, and R.H. Wijffels, *Biotechnol. Bioeng.* **55**, 630–641 (1997).
67. T. Sumino, H. Nakamura, and N. Mori, *J. Ferment. Bioeng.* **72**, 141–143 (1991).
68. T. Sumino, H. Nakamura, N. Mori, Y. Kawaguchi, and M. Toda, *Appl. Microbiol. Biotechnol.* **36**, 556–560 (1992).
69. I. Nilsson, S. Ohlson, L. Haggstrom, N. Molin, and K. Mosbach, *Eur. J. Appl. Microbiol. Biotechnol.* **10**, 262–274 (1980).
70. I. Nilsson and S. Ohlson, *Eur. J. Appl. Microbiol. Biotechnol.* **14**, 86–90 (1982).
71. J. Tramper and A.W.A. de Man, *Enzyme Microb. Technol.* **8**, 472–476 (1986).
72. E. Kokufuta, M. Shimohashi, and I. Nakamura, *J. Ferment. Technol.* **64**, 533–538 (1986).
73. S. Cizinska, V. Vojtisek, J. Maixner, J. Barta, and V. Krumphanzl, *Biotechnol. Lett.* **7**, 737–742 (1985).
74. M. Taya, H. Miura, K. Uchiyama, S. Iijima, and T. Kobayashi, *Hakkoukougaku* **66**, 235–240 (1988).
75. Y.F. Lin and K.C. Chen, *Water Res.* **29**, 35–43 (1995).
76. F. Cecen and E. Orak, *J. Chem. Tech. Biotechnol.* **65**, 229–238 (1996).
77. Japan Sewerage Division, Evaluation Report, *Pegasus*, 1994.
78. T. Sumino, K. Noto, T. Ogasawara, N. Hashimoto, and Y. Suwa, *Proceedings of the Water Environment Federation 70th Annual Conference Exposition*, Chicago, 1997, pp. 165–172.
79. T. Sumino, H. Nakamura, N. Mori, and Y. Kawaguchi, *J. Ferment. Technol.* **73**, 37–42 (1992).
80. T. Takatera, H. Izumida, T. Iida, and T. Kobayashi, *Ferment. Bioeng. Conf. Proc.* 1987 p. 161.
81. M. Fletcher, M.J. Latham, J.M. Lynch, and P.R. Rutter, in R.C.W. Berkeley, J.M. Lynch, J. Melling, P.R. Rutter, and B. Vincent eds., *Microbial Adhesion to Surfaces*, Ellis Horwood, Chichester, 1980.
82. S.L. Daniels, *Dev. Ind. Microbiol.* **13**, 211–253 (1972).
83. M. Matsumura, T. Yamamoto, K. Shinabe, and K. Yasuda, *Water Res.* **31**, 1027–1034 (1997).
84. H. Deguchi and M. Kashiwaya, *Water Sci. Technol.* **30**, 143–149 (1994).
85. G. Claus and H.J. Kutzner, *Appl. Microbiol. Biotechnol.* **22**, 289–296 (1985).
86. T.H. Lessel, *Water Sci. Technol.* **23**, 825–834 (1991).
87. T. Iida, *Industrial Application of Immobilized Biocatalysts*, Dekker, New York, 1992, pp. 163–182.
88. Y. Lue, *Colloids Surface* **5**, 213–219 (1995).
89. P.G. Rouxhet and N. Mozes, *Water Sci. Technol.* **22**, 1–16 (1990).
90. M. Kikuta and T. Iida, *RadTech Asia '97 Radiation Curing Conference Proceedings*, Yokohama, 1997, pp. 820–823.
91. U.S. Pat. 4,605,622 (August 12, 1986), E. Hasegawa, T. Iida, M. Sakamoto (to Kansai Paint Co., Ltd.).
92. G.A. Stanley, T.J. Hobley, and N.B. Pamment, *Biotechnol. Bioeng.* **53**, 88–99 (1997).
93. K. Mishima, T. Nishimura, M. Goi, and N. Katsukura, *Water Sci. Technol.* **34**, 137–143 (1996).
94. K. Mishima, T. Yamada, and H. Itoh, *Ebara Time Signal* **170**, 81–85 (1996).
95. T. Kuba, E. Murnleitner, M.C.M. van Loosdrecht, and J.J. Heijnen, *Biotechnol. Bioeng.* **52**, 685–695 (1996).
96. E. Murnleitner, T. Kuba, M.C.M. van Loosdrecht, and J.J. Heijnen, *Biotechnol. Bioeng.* **54**, 434–450 (1997).

WINE PRODUCTION

O. COLAGRANDE
Università Cattolica Sacro Cuore
Piacenza, Italy

KEY WORDS

Alcoholic fermentation
Colloidal dispersion
Enzyme treatments
Glucan
Malolactic fermentation
Organoleptic
Polyphenols
Saccharomyces cerevisiae

OUTLINE

INTRODUCTION

During the past 20 years there have been considerable developments in winemaking techniques, with the introduction of processes and machinery borrowed from other sectors of the food industry. These developments have affected all phases of wine production, wine being the end product of a technological chain that includes the preparation and treatment of the must, fermentation, aging, and bottling. Thus, technological progress has improved the quality of the wine produced. Wine quality is measured by finesse, intensity, and originality in taste and smell and by microbiological and physicochemical stability (1–4).

The main phases of production of white and red wines are illustrated schematically in Figures 1 and 2. Figure 3 shows those technological innovations that may be of practical use. Among these, during the past few years particular attention has been paid to enzyme treatments and to the use of selected fermentation starters (both free and immobilized). Some of these processes have already been applied to production, and others are still in the experimental stage.

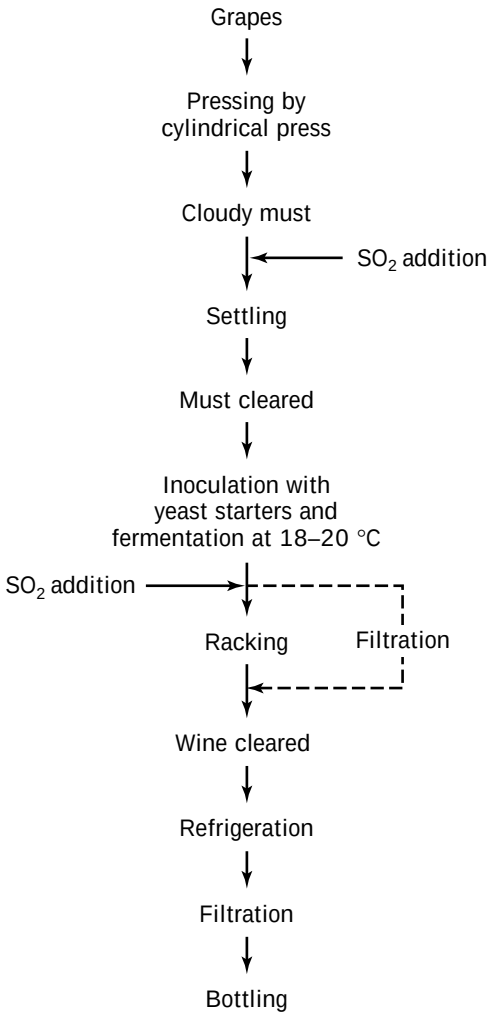


Figure 1. Scheme of a process for white wine preparation (fermentation with skins off). Terms are defined as follows. Pressing by cylindrical press: The operation of applying pressure to grapes to extract the liquid, which allows the juice to separate from the solids (skins, seeds, and stems). Settling: The operation of clarifying must at T = 5 to 10 °C (must can also be clarified by centrifugation or filtration). Refrigeration: The operation of cooling to 0 °C or below, used to hasten the precipitation of potassium bitartrate and to improve the stability of the wine.

ENZYME TREATMENTS

Winemaking is a biotechnological process in which enzymes play a fundamental role and in which the use of exogenous enzyme preparations makes it possible to overcome the problem of insufficient activity of endogenous enzymes in the grapes. The enzymic treatments have a multiple purpose in that the changes made to the grapes, the

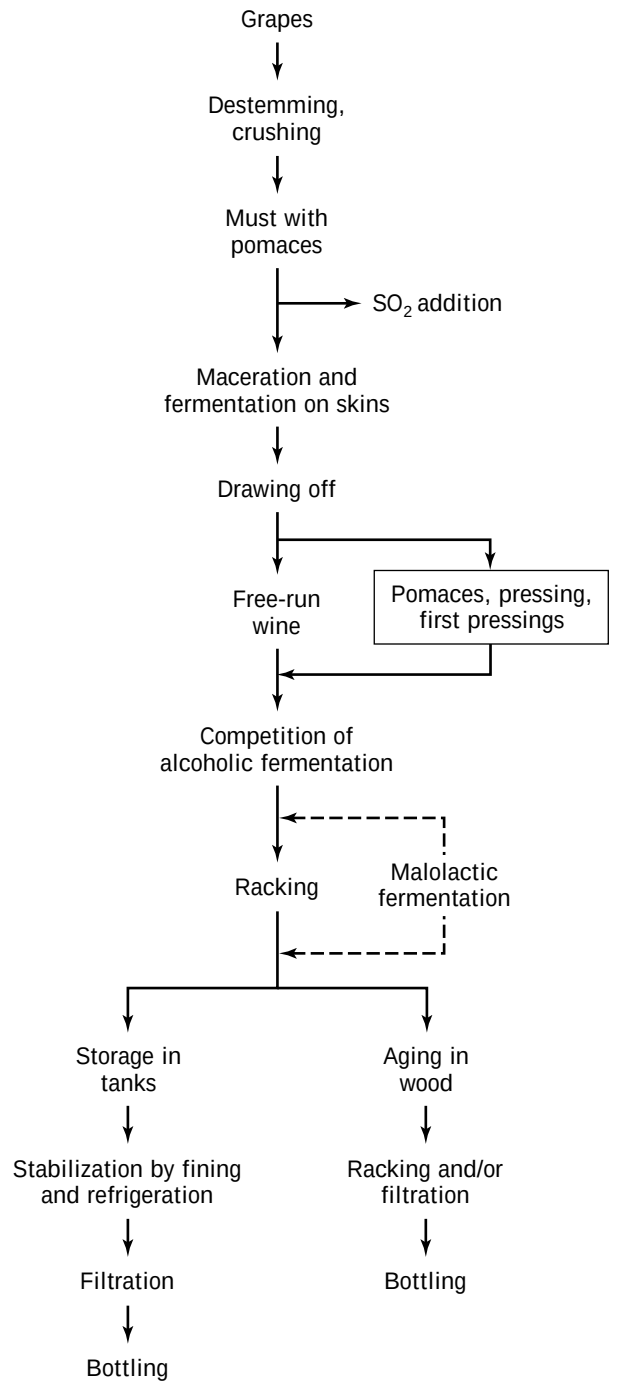


Figure 2. Scheme of a process for red wine preparation (fermentation with skins on). Terms are defined as follows. Destemming-crushing: The operation that mechanically removes the stems from the grapes and liberates the juice. Maceration: Contact of the must with the solid parts (skins and seeds) of the berries to enrich the juice with skin constituents. The maceration conditions (temperature, duration, pumping oversystem) change in relation to the grape variety and to the organoleptic characteristics of the wine produced. Alcoholic fermentation: With natural or selected yeasts. Malolactic fermentation: The biological decomposition of malic acid can be carried out soon after alcoholic fermentation or during storage (aging).

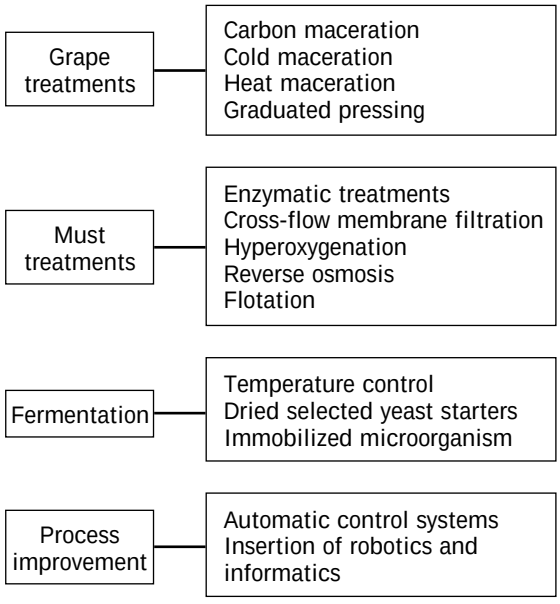


Figure 3. Technological improvement in the enological industry (5).

must, and the wine have an influence on the processes of clarification and filtration, the processes of extraction of color and aroma, and the stability of the wine (6–10). It must, however, be remembered that the use of enzymes in the wine industry, apart from pectolytic enzymes, remains extremely limited for a series of reasons that can be summarized as follows:

- The traditionalism of the wine industry, which is still anchored to classical methods
- Limitations of enzymic activity connected with the composition of the must and the wine (low pH, ethanol concentration, presence of tannins, and a fairly low working temperature compared to the optimum for enzyme preparations)
- Legal restrictions

However, during the past few years, research has suggested the use of enzymes that catalyze the transformation of glucan (glucanase), proteins (peptidase), and polyphenols (laccase, peroxidase, phenolase, and tannase), all of which are compounds that condition the stability and organoleptic characteristics of wines. Particular attention has been paid to glycosidases that are able to free odorous compounds by hydrolyzing terpene glycosides.

Pectolytic Enzymes

Pectolytic enzymes have been widely used for many years, because they act on the pectins. Pectins are complexes of acids (homogalacturonan and rhamnogalacturonan) and neutral polysaccharides (arabinan, galactan, and arabinogalactan) that are found in the must in colloid form. They act as protective colloids, causing problems in the clarification and stabilization (11–13).

Enzymes that can act on the pectic substances are usually classified into two groups, as shown in Figure 4 (14,15): de-esterifiers (pectinmethylesterase) and depolymerases that catalyze both hydrolysis (polygalacturonase) and β-elimination reactions (pectin lyase and pectate lyase). Purified enzymic preparations from cultures of microorganisms normally contain polygalacturonase and lyases associated with hemicellulase and cellulase. However, the presence of pectin methylesterase must not be too great in order to limit the liberation of methanol, which is found esterified with the carboxylic groups of the galacturonic acid in the pectic molecule (16). Pectolytic enzymes are used at vintage time in the preparation of both white and red wines. Italian law permits the use of pectolytic enzymes in the processing of pressed grapes, the must to be used for producing white wine, musts and wines obtained using the thermovinification system, and wines after they have been separated from the pomaces (Ministerial Decree of 8 September 1976, G.U.I.). In the context of numerous experiments that have been carried out during the past few years as well as the development of new commercial products, various authors have evaluated the influence of the use of these enzymes on the extraction of juice, color, and aroma and on the clarification and filtration processes. Ough and Crowell (17) investigated the effect of the use of some exogenous pectolytic enzymes at pressing on the must yields, turbidity, and solids, using different varieties of grapes. The enzyme-treated samples gave better juice yields than the untreated samples. The amount of sediment after overnight settling was less for the controls. After racking and centrifuging, the enzyme-treated juice samples were clearer and contained fewer insoluble solids.

The addition of pectolytic enzymes has a favorable influence on the clarification of musts and, consequently, also has a positive effect on such parameters as viscosity and filterability (18,19). The treatment of crushed grapes with pectolytic enzymes in the production of red wine leads to a more rapid and complete breakdown of the cell wall of the skin. This leads to greater extraction of anthocyanins and, thus, to an increase in color, although the results may vary according the cultivar of grapes used (20).

Rational use of pectolytic enzymes, especially in the case of grapes particularly rich in pectic substances (e.g., muscat) makes it possible to improve the economics of the process (e.g., through reduction in time required, better use of equipment, and reduction in loss of must and wine) without negatively altering or modifying the quality characteristics of the final product (21,22).

β-Glucanase

Glucan is a polysaccharide with a high molecular weight (M_r 10^5 – 10^6), consisting entirely of glucose subunits. It is a (1–3:1–6)-β-D-glucan. It is found in must and wines from grapes infected by *Botrytis cinerea* (23). In these wines, the glucan is of great technological importance because it tends to clog the filters if a quantity is present by forming a glutinous layer on them. This causes considerable problems in clarification and filtration. In order to solve this problem of clogging, an enzyme treatment was set up in the 1980s, based on β-glucanase, which hydrolyzes the glucan excreted by the *Botrytis cinerea* (25).

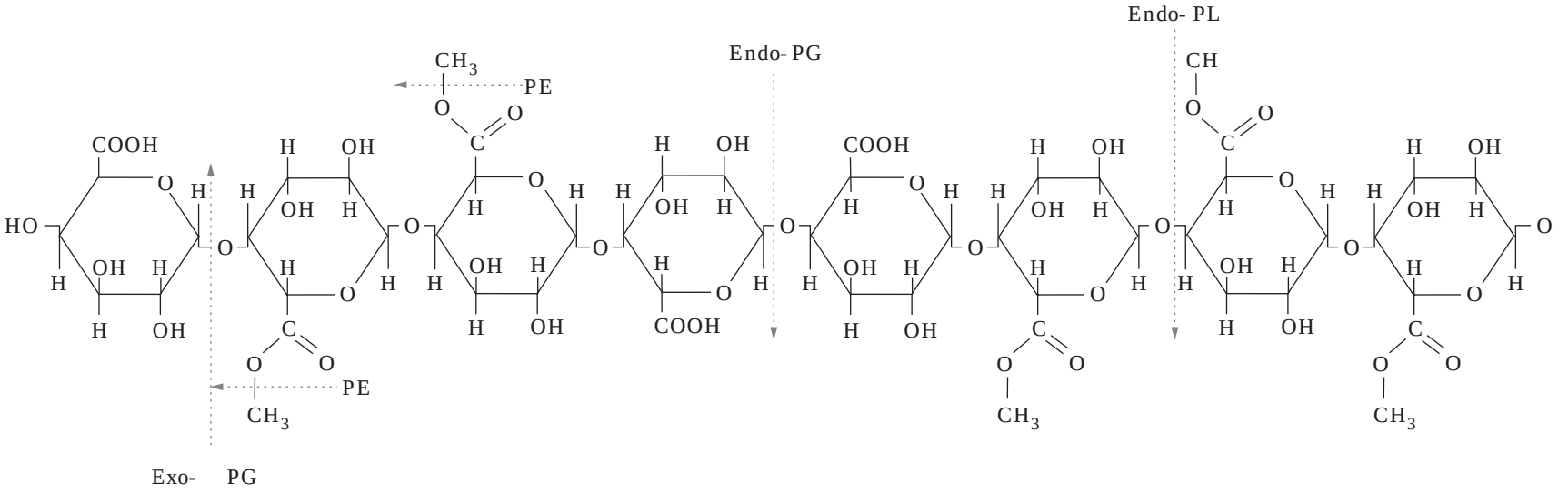
Experiments carried out by Villettaz (9,26,27) and Wucherpfenning and Dietrich (28) to evaluate the efficacy of β-glucanase have demonstrated the technological and economic importance of this enzyme in improving the filterability of wines produced from grapes affected by *Botrytis*.

β-glucanase has been included in the practices and treatments permitted by the EEC, although the conditions for use have yet to be defined (29,30).

Proteolytic enzymes

In winemaking, treatments are carried out to lower the protein content, because in must and wine, proteins are found in colloidal dispersion, which can become insoluble, forming haze and precipitates (31,32). When this occurs during processing it causes delays, but if it occurs when the wine is already on sale it causes economic loss. This phenomenon occurs predominantly in white wines and only very rarely in red wines. Red wines contain tannins, which are believed to react with the proteins to form insoluble tannoprotein compounds. During fermentation and aging, this leads to a drastic reduction in the protein levels (33). Although the proteins responsible for the instability have not yet been well characterized, it seems that the protein fractions of lower MW (12,600 and 20,000 to 30,000) and lower P_i (4.1 to 5.8) and glycoproteins are the major and most important fractions contributing to protein instability in wines (34). Treatments with proteolytic enzymes have been tried as an alternative to using traditional finings such as bentonite. These enzymes cause hydrolysis of the proteins into amino acids or short-chain peptides that are completely soluble in the wine (35–37); however, intervention with exogenous enzymes is not free from problems. Many of the traditional proteases (papain, pepsin, and trypsin) are not suitable for use in wine; they require high temperatures that are deleterious to wine quality, and they do not have their optimal pH in the range of 3 to 4, which is ideal for wines. For these reasons, studies have mostly been directed toward the use of acid proteases of fungal origin in free and immobilized forms; however, so far the results have not always been positive. Modra and coworkers (38) assessed the action of commercial peptidases on proteins isolated from muscat grapes. This treatment modified the HPLC elution profile of a high relative molecular weight fraction containing protein. The most generally discernible change was a decrease in the M_r 23.000 peak, with a concomitant increase in the 10.000 M_r peak, and a loss of resolution in both. The change in protein profile varied among the peptidases, but none of the enzymes gave complete proteolysis after 7 days. The tests carried out on muscat must confirmed that the peptidases modified the proteins; however, their efficacy was not sufficient to make the wine stable, and treatment with bentonite was required.

Woiwodow et al. (37) tested the action of an acid protease from *Aspergillus niger* immobilized on Sylochrom C-80 on the proteins of white table wines. Experiments were carried out with different quantities of enzyme (2 to



Polygalacturonases (PG)

Name	Elective substrate	Mode of action	End products
Pectinesterase (PE)	Exo-PG	Endo-PG	Endopectinlyase (PL)
Pectin	Pectate	Pectate	Pectine
Hydrolysis of ester bond	Hydrolysis of α-1,4 bonds	Random hydrolysis of α-1,4 bonds	
Pectate and methanol	D-galacturonate	Oligogalacturonates	Methyl esters of unsaturated oligogalacturonates

Figure 4. Enzymic degradation of pectins and some enzyme characteristics.

10 g/L) and with different times of contact (2 to 48 h) at a temperature of 23 to 25°C. Complete stability of the wine was reached after prolonged treatment (48 h) with very high quantities of immobilized enzyme (10 g/L). In other tests, electrophoresis of the proteins isolated from the wine before and after treatment showed a diminution in both number and intensity of protein fractions, but the wine was not completely stable. Bakalinsky and Boulton (35) studied the use of food-grade acid protease (from *A. niger*) covalently linked to agarose spheres of 100 μm. Four types of wine were treated at temperatures of 30 and 37 °C in a recirculating flow packed-bed reactor containing 226 mg of dry weight conjugate. At the end of the treatment, the wines were subjected to heat-stability testing. The results showed that the enzyme, although active in all of the wines tested, achieved protein stability only in Riesling wine. The Riesling was the only wine whose pH was adjusted to 3.0 for treatment, whereas the other wines were treated at their natural pH levels (Sauvignon blanc, 3.65; chardonnay, 3.57; Gewürztraminer 3.59). The pH adjustment may have been the cause of the successful stabilization.

That immobilized protease was active in the wine but failed to have any significant effect on the protein stability may indicate that the proteases do not hydrolyze the protein fractions responsible for the instability or that the hydrolysis occurs along with the formation of products that are still unstable. If the potentially unstable wine proteins were complexed with polyphenolic compounds, one would expect them to be more resistant to enzymic hydrolysis. Currently, the use of free or immobilized proteases to stabilize wine requires more information on the nature of the proteins, their interactions with the other components of wine, and the conditions that cause them to become insoluble. Furthermore, sensory evaluation of the treated wine is recommended to assess possible detrimental alterations in flavor caused by the probable increase in peptide content. Small hydrophobic peptides, in particular, are recognized as having a bitter flavor (39).

Enzymes with Specific Action for Polyphenols

Phenols in wine predominantly arise from the grapes and are primarily flavonoid. They are crucial components in wines because of their basic importance to the color and flavor. Flavonoids (catechins, proanthocyanidins) are also responsible for discoloration, turbidity, and flavor changes in wines. Review of the importance of phenolic compounds in grapes and wine has recently been published (40,41). Chemical and enzymic oxidation of phenolic compounds leads to the formation of highly reactive intermediates (quinones) that interact with amino acids to reduce sugars and alcohol and give rise to chromophores and volatile substances (40,42,43). This results in a loss of freshness, anywhere from yellowing to browning of the wine, and turbidity (42,44). A summary of these reactions can be found in Figure 5.

The removal of polyphenols is a common practice in winemaking processes to prevent discoloration and to obtain stabilization against oxidation, especially in white and rose wines (45). Wine technology tends toward the adoption of systems for extraction of the must that limit the solubilization of oxidizable substances (46) and toward the use of finings (gelatin and casein), adsorbants (active carbon), and polyvinylpolypyrrolidone (PVPP) (47,48). The use of these additives in wine production is aimed at reduction of the level of polyphenols and compounds that catalyze their transformation, particularly enzymes and heavy metals. To limit the amount of treatment with chemical-physical agents, it has been suggested that enzymes active against polyphenols (oxydases, hydrolases, and transferases) be used in the prefermentative phase. The aim of this is to set off a process of enzymic oxidation of polyphenols in the must under controlled conditions, so as to destabilize them and accelerate the process of polymerization and flocculation. Enzymes and reactions are shown in Table 1. The possibility of eliminating or achieving reactivity loss of the polyphenols responsible for wine instability has been tested on the bench and on a pilot scale using enzymic preparations containing, respectively, tannase (from *Aspergillus* sp.), phenolase (prepared as an extract or acetoin powder from *Agaricus campestris*), and laccase (from *Polyporus versicolor* in a liquid culture, partially purified) (47,49).

These different enzyme treatments were compared with each other, and some were then compared with conventional systems of stabilization. The polyphenol concentration in wines prepared from the same, but differently processed, pinot grapes showed that laccase was more effective than the other enzymes tested and produced a wine with a more stable color (Table 2). The possibility of using laccase was checked on a pilot scale with the must from Trebbiano grapes. Laccase treatment was compared with clarification carried out with casein, active carbon, and bentonite, with or without sulphur dioxide. The use of the enzyme was shown to be highly effective, preferable or practically identical to the results from traditional processing. The wines treated with laccase had different and better organoleptic maderization-resistant features (47). In the case of muscat and Riesling, the action of laccase (8,300 units/L) was compared with that of fining agents (gelatin/SiO_2) (50). The results obtained with muscat must without SO_2 and Riesling must with SO_2 (150 ppm) (Table 3) shows that the enzymic treatment is highly effective, preferable, or practically identical to traditional processing, especially when using must without sulphur dioxide (Table 3).

These interesting results allow us to conclude that laccase prefermentative treatments make it possible to eliminate instability in white wines caused by oxidizable polyphenols. Obviously, the process of winemaking must include physical (e.g., ultrafiltration) or chemical (e.g., addition of sulphur dioxide, deproteinization) systems that make it possible to eliminate the enzyme. The enzymic treatment coupled with a filtration system (filtration by diatomaceous earth, PVPP, or silica of ultrafiltration) is probably a technique that gives sufficient stability for white wines.

Glycosidases Acting on Precursors of the Aroma

The aroma of grapes includes volatile free odorous substances, especially terpenes (linalool, geraniol, nerol, citro-

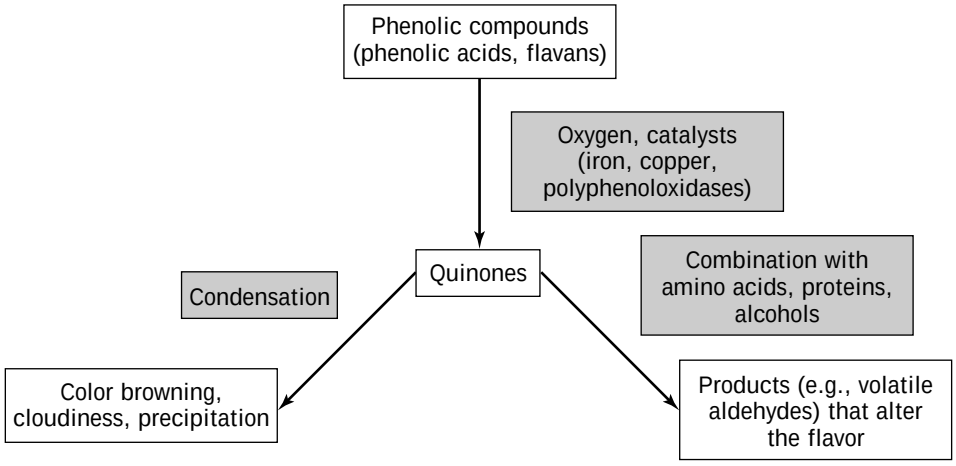


Figure 5. Possible reactions of maderization.

Table 1. Phenolic Enzymes

Oxidases
Peroxidase (EC 1.11.1.7)
Donor (catecol) + H_2O_2, oxidized donor (quinone) + $2H_2O$
o-Diphenoloxidase (catecholoxidase, polyphenoloxidase, phenolase, tyrosinase, cresolase[a]) (EC 1.10.3.1)
2 *o*-Diphenol + O_2; 2 *o*-quinone + $2H_2O$
p-Diphenoloxidase (laccase) (EC 1.10.3.2)
2 *p*-Diphenol + O_2; 2 *p*-quinone + 2 H_2O
Hydrolases
Tannin acylhydrolase (tannase) (EC 3.1.1.20)
Galloylglucose, gallic acid + glucose
Anthocyanase (EC 3.1.1)
Anthocyanidin 3-monoglycoside, anthocianidin + hexose

[a]Monophenol + O_2 + $2e^-$, *o*-diphenol + O^{2-}.

Table 2. Polyphenol Content and Color Increases Found in Maderization Test of Pinot Wines Obtained by Various Enzyme Treatments

	No enzyme	Laccase	Phenolase	Tannase
	Polyphenol (mg/L)			
Total	339	310	319	333
Nontannic	267	261	284	284
	Color			
E_{420mm} (×1,000)	92	82	78	64
Increase due to maderization test	121	95	160	156

Source: From Ref. 47.

Table 3. Comparison of Untreated and Use of Fining or Laccase Enzyme on Total Phenolics and Catechins of Muscat and Riesling Must

	Untreated	Fining	Laccase
	Muscat (raw)		
Total phenolics	232	184	90
Catechins	43	56	30
Color E_{420mm} (×1,000)	226	99	89
	Riesling (150 ppm SO_2 added)		
Total phenolics	303	282	286
Catechins	29	28	29
Color E_{420mm} (×1,000)	83	53	60

Note: Total phenolics expressed as mg/L gallic acid, and catechins expressed as mg/L *d*-catechin.
Source: From Ref. 48.

nellol, α-terpineol, linalool oxide) that play a fundamental role in giving character to certain cultivars, for example, muscat (51–53). In addition, grapes also include compounds called aroma precursors, which are glycosides mainly of linalool, nerol, and geraniol (for example 6-*O*-α-L-arabinofuranosyl-β-D-glucopyranosides, 6-*O*-α-L-rhamnopyranosyl-β-D-glucopyranosides and β-D-apiofuranosyl-β-D-glucopyranosides (54–57). The ratio between bound and free monoterpenols ranges from 1 to 5 in juice grape cultivars of muscat and Riesling and can be as high as 15 in the Gewürztraminer variety (58,59). From a technological point of view, the existence of such bound forms is interesting because these glycosides, nonvolatile precursors, can give flavoring aglycones when hydrolytic reactions take place. High-temperature acid hydrolysis can be used, but it may lead to extensive rearrangements of the terpenols. A study of the transformations of the terpenic compounds has in fact shown that, as the pH of the aqueous medium diminishes, the hydration and cyclization of linalool increase (60).

Enzymic methods are attractive for increasing the concentration of free flavorants with minimal changes in the natural monoterpene composition. Recent studies have shown that enzymic hydrolysis occurs in two stages, following a sequential mechanism illustrated in Figure 6. In the first stage, the intersugar linkage is cleaved by arabinofuranosidase, rhamnopyranosidase, or apiofuranosidase (depending on the structure of aglycon moiety), and the corresponding monoterpenyl-β-D-glucosides are released. In the second stage, the liberation of monoterpenols takes place after action of a β-glucosidase on the previous monoterpenyl β-D-glucosides (62,63).

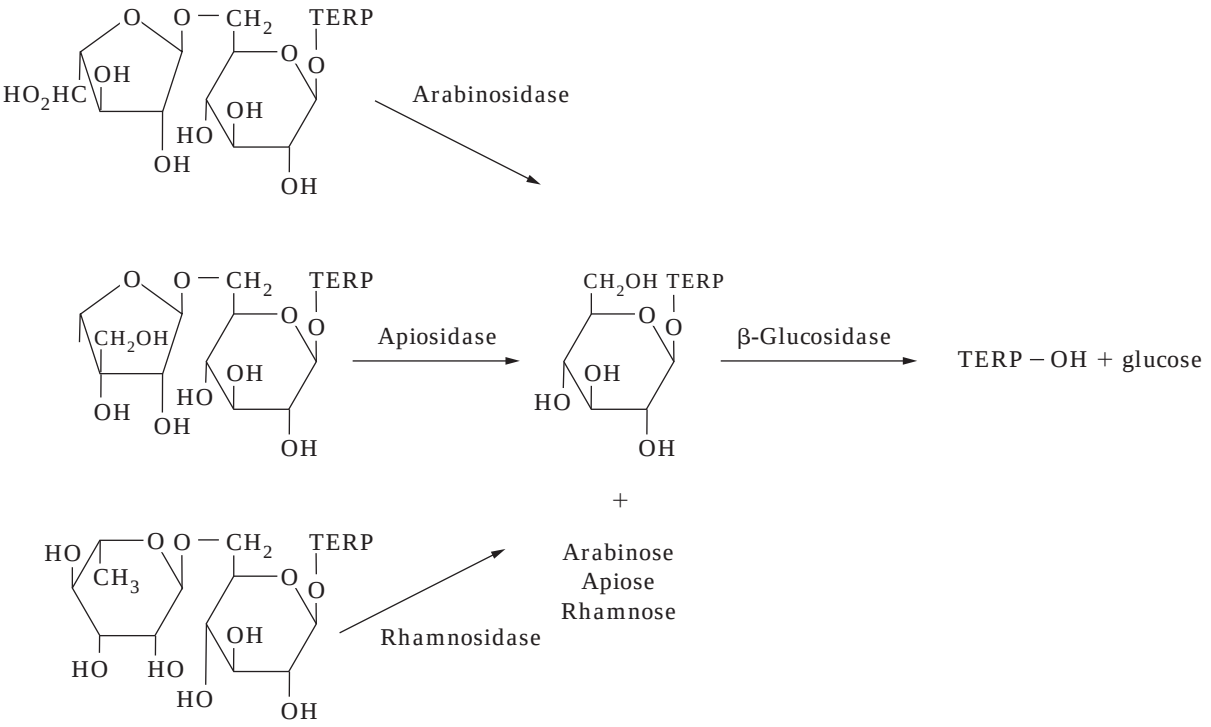


Figure 6. Sequential mechanism of terpene glycosides enzymic hydrolysis (61).

In the search for an enzymatic compound suitable for the hydrolysis of these precursors, different commercial products (pectinase, cellulase, hemicellulase, lypase, and protease) have been tested. Some of the preparations tested have contemporaneously good α-arabinofuranosidase, α-rhamnopyranosidase, and β-glucosidase action, but their efficacy in the must and wine is a function of numerous factors (64). The efficiency of hydrolysis of monoterpenyl β-D-glucosides by β-glucosidases was found to be dependent on the structure of the aglycon and the origin of the enzyme. Moreover, the efficiency of β-glucosidase is influenced by the concentration of glucose and gluconolactone, and in sweet wines, the hydrolysis is interrupted at the stage of the monoglucoside without freeing the odorous monomers. In the case of dry wines, on the other hand, there is a considerable increase of most of the free terpenes (65). Considering that the glycosidases used so far have not given the results hoped for, and that this is principally because of the properties of β-glucosidase, the hydrolytic activities of various plant and fungal β ;-glucosidases toward terpene glycosides are being studied. The goal of these studies is to find an enzyme with large aglycon specificity and good stability to low pH as well as tolerance to glucose and ethanol.

USE OF SELECTED CULTURES AND INNOVATION IN THE FERMENTATION PROCESS

Of the two fermentation processes, alcoholic fermentation causes the transformation of sugars into ethyl alcohol and CO_2, whereas malolactic fermentation causes the transformation of malic acid into lactic acid and CO_2. However, both processes produce compounds (organic acid, alcohols, ester, and carbonilic compounds) that influence or determine the organoleptic characteristics of wines and so play a fundamental role in fixing wine quality (3,66–68). For this reason, the microorganisms responsible for these transformations, yeast and bacteria, have for some time been the object of great interest. The work of Pasteur (1886) was followed by numerous studies of the microflora of grapes and of the factors that regulate the equilibria between the various species and their metabolism. In fact, a great variety of microorganisms are present in grapes and musts, and the yeasts responsible for alcoholic fermentation are but a small part of them (69–71). The fermentative yeasts (particularly *Saccharomyces cerevisiae* and related species) generally predominate because the chemical composition of the must is more favorable for their development than for that of other eumycetes and schizomycetes (72,73). However, it is difficult to control alcoholic fermentation that develops naturally, because certain parameters (such as temperature, degree of oxygenation, degree of clarity of the must, and phenomena of microbial antagonism) can modify the natural selection and therefore condition the development and completion of the process. This alters the ratio between the principal (ethanol and CO_2) and secondary (acetic acid, succinic acid, acetoin, 2,3-butanediol, acetaldehyde, glycerol) products.

Because it is impossible to cover in this review even all the main aspects of microbial biochemistry involved in winemaking and the influence of certain environmental conditions on grape must fermentation, textbooks or re-

views should be consulted for more detailed information on these topics (66,72,74–80).

Selected *Saccharomyces cerevisiae*

To overcome the problems mentioned earlier, for many years it has been standard practice, especially in the production of white wine, to use dry pure cultures of enologically suitable species (in particular *Saccharomyces cerevisiae* and all its physiological variants) with known properties. This ensures a quicker start to alcoholic fermentation (thanks to the massive addition of yeast), prevents the development of atypical microflora, and therefore eliminates the uncertainty and variation that occur in spontaneous alcoholic fermentation (50,66,76). Selected culture has therefore been the subject of many investigations that have considered the desirable characteristics of wine yeasts and their influence on the composition of wine and its organoleptic characteristics. Numerous properties of yeasts have been studied: some are always desirable, and others are always undesirable (Table 4) (81).

During the past few years, the wine industry has looked with interest at yeast genetics and strain development programs in order to obtain modified yeasts suitable for winemaking technologies that do not impair the flavor and bouquet of the final product. Yeast geneticists have carried out programs of genetic improvement to obtain new yeast strains that possess a high fermentation rate and high ethanol production as well as other relevant winemaking characteristics (82,83). This program has used both classical genetics (e.g., hybridization through conjugation between spores or aploid cells) and more recent biotechnological methods, such as somatic hybridization through fusion of protoplasts induced chemically or physically and transformation through the introduction of fragment of extraneous DNA using molecular vectors. Some reviews on this topic have recently been published by Falcone and Frontali (84) and by Pretorius and van der Westhuizen (85).

Thus far, it has been possible to eliminate foaming capacity of strains thought to be highly suitable for inducing alcoholic fermentation (86,87). Certain wine yeast strains produce proteins that are responsible for the foaming associated with fermentation. Excessive foaming during wine fermentation is an undesirable characteristic of wine yeast strains. The formation of a froth head can result in the loss of grape juice or reduce the capacity of plant equipment, because part of the fermentation vessel may have to be reserved to prevent the froth from spilling over. Foaminess also detracts from the visual appeal of the product (88).

Furthermore, it has been possible to effect genetic marking (89) to differentiate the selected strains from those of the indigenous microflora (always present in musts) and to give the killer phenotype to a strain (90–92) to increase its competitiveness with indigenous flora. The killer phenomenon is related to the ability of some yeasts to secrete toxic proteins that are lethal for so-called sensitive strains (93).

Particular attention is being given to the following possibilities: (*1*) introduction of the flocculent characteristic into yeast strains to facilitate the separation of cells from wine at the end of fermentation (94–97); (*2*) introducing malolactic enzymic activity into yeasts with the object of effecting alcoholic fermentation and malolactic fermentation with a single microorganism (yeast), a modification that would yield obvious technological advantages (98); (*3*) obtaining modified strains for the production of the aroma components (esters, fatty acids, and alcohols) (99); and (*4*) obtaining strains that do not produce H_2S, a component responsible for organoleptic alterations in the wine (100).

Achievements are still modest in the field of genetic improvement of enogical strains because of various factors: (*1*) the complex genetic background of industrial strains (ploidy, homothallism, poor sporulation, etc.) (101–104); (*2*) the impossibility of describing the genetic basis of most enological characteristics and, thus, of elaborating simple screening tests for them (85,105); and (*3*) the presence of indigenous microflora in the must that, unlike other substrata, cannot be subjected to sterilization without altering the wine characteristics.

Table 4. Some Characteristics of *Sacchromyces cerevisiae* Affecting the Winemaking Process

Desirable
Rapid initiation of fermentation
Efficient conversion of grape sugar to alcohol
Ability to conduct even fermentation
Ethanol tolerance
Ability to settle rapidly at the end of fermentation to aid clarification
Fermentation at low temperatures such as 10–14°C
Retention of viability during storage
Sulfur dioxide resistance (sulfur dioxide is normally used in winemaking)
Production of desirable fermentation bouquet and reproducible production of the correct levels of flavor and aroma compounds
Killer factor or resistance to killer toxins
Film-forming capacity
Undesirable
Production of sulfite, hydrogen sulfide, or mercaptan
High production of acetaldehyde, acetic acid (volatile acidity), ethyl acetate, and higher alcohols
foaming capacity
Production of urea, which can result in the formation of ethyl carbamate

Process Innovation in Alcoholic Fermentation

Another line of research that has developed in the field of enology is the search for and optimization of new processes that are able to meet the technological and organizational needs of the industry (greater productivity, reduction in the time needed for the process, reduction in capital and running costs). In this field, we can mention the use of yeasts immobilized in natural nontoxic polymers or contained in a microporous membrane and of the cell-recycle batch fermentation process (CRBF) for the preparation of

white wines and sparkling wines. The advantages that are sought are summarized in this section.

Bottle-Fermented Sparkling Wines (Champenois Method). It would be desirable to eliminate the phase of remuage, with a considerable savings in labor and time, thanks to instant sedimentation of the immobilized yeasts by gravity. Sparkling wine produced by the champenois method is characterized by fermentation of wine in the bottle, to which sugars, yeasts and other fermenting agents have been added (Fig. 7). The bottles thus prepared are stacked horizontally in thermoregulated environment (12 to 18 °C) where the secondary fermentation is carried out. Once fermentation has terminated and the wine has been in contact with the yeast for at least a year, the bottles are turned frequently and inclined (an operation referred to as remuage), so that the yeast sediment slides toward the bottle neck and settles on the stopper. The sediment is then eliminated by degorgement. Remuage is both long and arduous in that it requires a large amount of labor over a long period of time (1 to 4 months) and the occupation of large wine cellars. By using immobilized yeasts, it is possible to move directly from the stacks to degorgement with a saving in production costs of about 80% when compared with the traditional method (106).

Production of Sparkling Wine in Tanks (Charmat Process). The possibility of using a high cellular concentration and

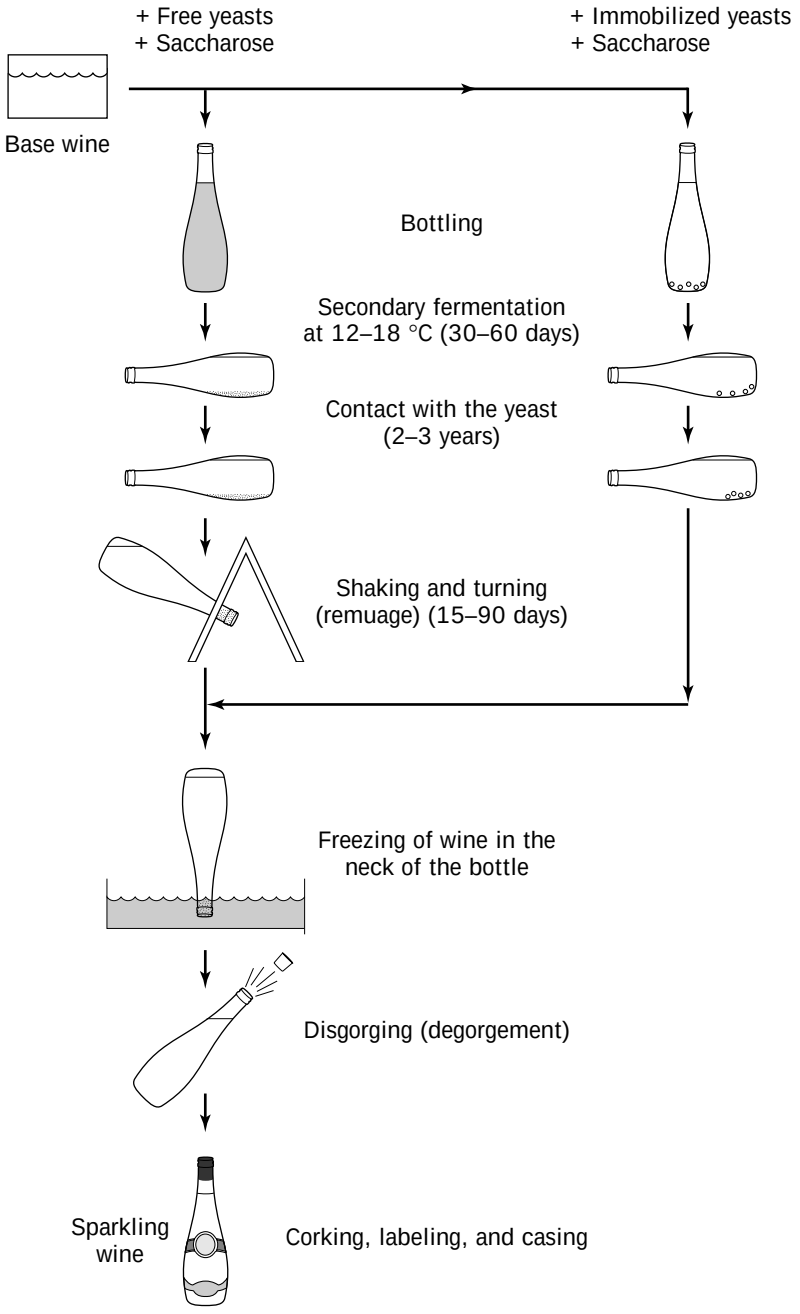


Figure 7. Scheme of the process for bottled sparkling wine preparation with traditional methods and with immobilized yeast.

thus of obtaining a very high rate of reaction so as to achieve a continuous process (107) is another desiderable innovation. This would allow a reduction in the dimension and number of pressure tanks, with a diminution in the costs of running and equipment. In the bulk process for sparkling wine production, the Charmat method (Fig. 8), sugars, other nutrients, and yeasts are added to the base wine, and the second fermentation occurs in large stainless steel sealed tanks. The sparkling wine is then filtered, fined, and bottled. This process is the cheapest for producing sparkling wines, but it still requires a considerable number of expensive tanks.

Use of Yeast Entrapped in Wine Production. The possibility of improving the rate and efficiency of transformation by increasing the concentration of yeast would be an advantage in general wine production. In this way, a linear development of the fermentative process is achieved, excluding the yeast induction phase and exponential growth. This reduces the time required by the fermentative process and the final level of the secondary products of fermentation (acetaldehyde, acetic acid, products of catabolism of the amino acids, etc.) that accumulate in the first phase of fermentation when the biomass is formed (50).

The use of yeasts entrapped in calcium alginate and in other materials has been investigated by Cantarelli (50); Duteurtre et al. (108); Lemonnier and Duteurtre (109); Loureiro (110); Zamorani et al. (111); and researchers at our Institute (106,112–115). This research has made it possible to observe the following:

- Of the various compounds used for entrapment—agar, agarose, calcium alginate, barium alginate, κ-carrageenan, gelatin, pumice stone, porous porcelain, controlled porous glass, and dyalysis membranes—agar and calcium alginate gel are the most suitable for fermentation (110).
- The retention power of the immobilizing matrix is inversely linked to the fermentative capacity of the yeast and to the alginate/yeast ratio (112).
- The presence of structural constituents of alginate (uronic acids) in the wine may be ignored up to bead content equal to 14 g/bottle (106).
- The phosphate and tartaric, malic, and citric acids present in wine do not have a disintegrating effect on the structure of the alginate (106).
- The enrichment of cells that proliferate in nucleic acids, reserve polysaccharides (trehalose and glycogen), and structural polysaccharides (mannan and glucan) is not reflected in a variation in the content of the wine (106).

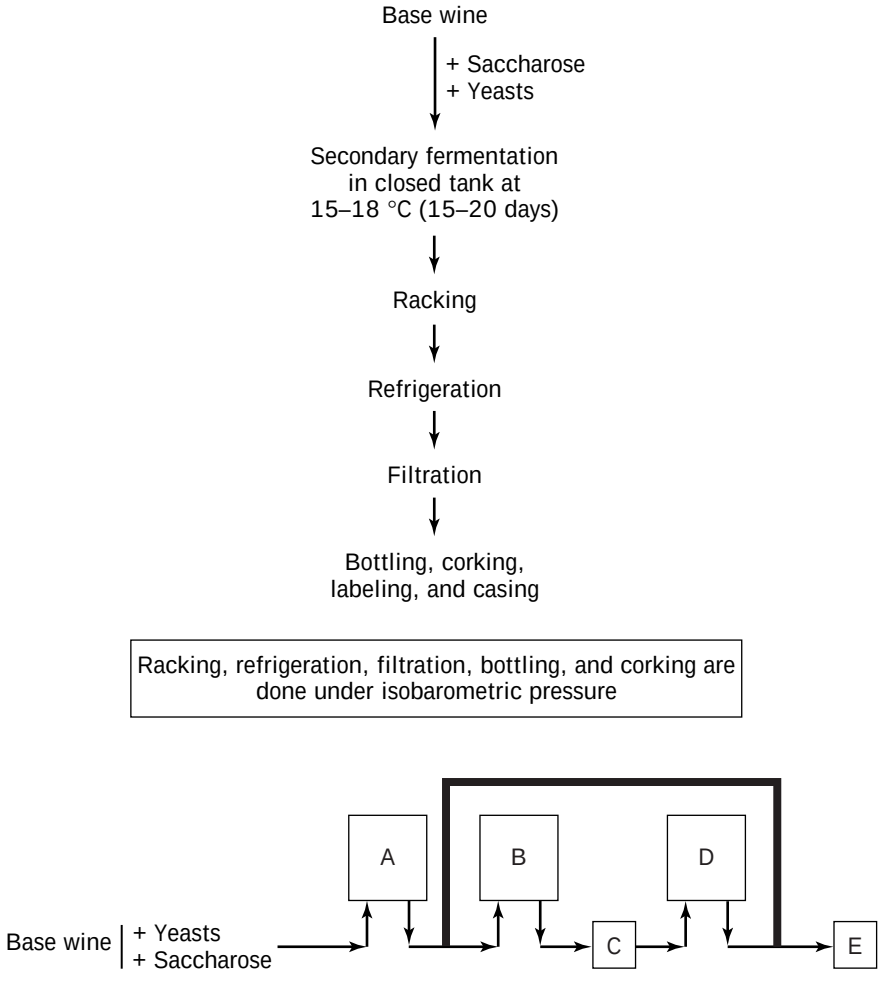


Figure 8. Scheme of a process for sparkling wine preparation by the close tank or Charmat methods. A, closed tank for secondary fermentation; B, tank for refrigeration; C, fieter; D, tank for ready to be bottled; E, isobarometric bottling.

- When the concentrations of alginate and cells in beads are equal, then the smaller the inoculum at draught, the greater the cellular release; however, this release can be avoided by using nonproliferating yeasts (106).

As for the evaluation of the influence of the alginate immobilization technique on the fermentative process, the composition, and the organoleptic characteristics of the bottled sparkling wine, Fumi et al. (114) observed that the use of immobilized cells does not cause any significant change in the second fermentation process. The main components (ethanol, organic acids, total nitrogen, and higher alcohols) are not appreciably different in sparkling wines obtained by immobilized yeast when compared with traditional sparkling wines (Table 5), but some differences are found as regards certain amino acids and some aroma components (Table 6). As shown by tasting, however, these differences do not affect the organoleptic characteristics. Sensory evaluation of sparkling wines obtained with immobilized yeasts versus a control was performed using a duo-trio test and black glasses.

Divies (116) and Duteurtre (108) conducted comparative tests with free and calcium alginate immobilized yeasts for bottled sparkling wine production, and they also noted that there are no variations in the principal chemical or physical characteristics.

Experiments to avoid cell release were performed by coating alginate/yeast beads with a protective cell-free alginate gel layer. These studies were carried out by Duteurtre (108), Fumi et al., (106), and Loureiro (110) for bottled sparkling wine production. A complete bottle fermentation was achieved without cell release, independent of the different conditions adopted—concentration of the alginate, time of gelification, and $CaCl_2$ concentration. However, Ors et al. (117) noted that the double layer must have a minimum critical thickness, equal to 0.2 mm, in order to avoid cellular release.

In an assay where agar cylinders covered with a protective layer of agar without cells were used, equally satisfactory results were not achieved (110).

Although the problem of cell release from the immobilization matrix during fermentation may be solved, there

Table 5. Analytical Characteristics of Sparkling Wine Produced by Free Yeast Cells and Yeast Cells Immobilized in 3% Alginate

	Free yeast	Immobilized yeast
Ethanol (% vol)	12.29	12.29
Sugar (g/L)	2.3	1.7
Ashes (g/L)	1.48	1.48
Ashes alkaline (mequiv/L)	11.0	11.0
pH	3.25	3.21
Titratable acidity (g/L)	6.79	6.93
Volatile acidity (g/L)	0.36	0.34
Tartaric acid (g/L)	1.73	1.82
Malic acid (g/L)	1.47	1.47
Lactic acid (g/L)	2.10	2.06
Citric acid (g/L)	0.20	0.24
Succinic acid (g/L)	0.40	0.44
Pyruvic acid (mg/L)	47.5	49.7
α-Ketoglutaric acid (mg/L)	58	63
Acetaldehyde (mg/L)	71	68
Glycerol (g/L)	5.90	5.60
Color (410 nm)	0.073	0.072
Gauge pressure CO_2 at 20°C (atm)	8.0	8.3
Calcium (mg/L)	66.5	65.5

Source: From Ref. 114.

Table 6. Free Amino Acids and Aroma Components of Sparkling Wines Produced by Free and Immobilized Yeast in 3% Alginate

	Free yeast	Immobilized yeast
Amino Acids		
Histidine (mg/L)	38.60	36.80
Asparagine (mg/L)	167.76	253.62
Arginine (mg/L)	542.11	612.24
Serine (mg/L)	67.83	65.88
Aspartic + glutamic acid (mg/L)	152.39	147.50
Glycine + threonine (mg/L)	88.17	84.16
Alanine (mg/L)	362.0	359.5
γ-Aminobutyric acid (mg/L)	151.2	145.8
Proline (mg/L)	227.03	220.97
Valine (mg/L)	22.75	23.08
Methionine (mg/L)	trace	trace
Tryptophan (mg/L)	trace	trace
Phenylalanine (mg/L)	17.70	18.38
Isoleucine (mg/L)	31.52	31.67
Ornithine (mg/L)	18.58	19.54
Lysine (mg/L)	12.53	14.26
Tyrosine (mg/L)	trace	trace
Total nitrogen (mg/L)	534	528
Esters		
Ethyl butyrate (mg/L)	0.38	0.27
Ethyl caproate (mg/L)	0.53	0.52
Ethyl pyruvate (mg/L)	0.56	0.43
Ethyl lactate (mg/L)	143	113
Diethyl succinate (mg/L)	0.63	0.60
Isobutyl acetate (mg/L)	0.13	0.08
Isoamyl acetate (mg/L)	1.15	0.82
Alcohols		
Methyl alcohol (mL/100 mL of EtOH)	0.03	0.03
1-Propanol (mg/L)	29	31
1-Butanol (mg/L)	1.62	1.23
2-Methyl-1-propanol (mg/L)	55	52
2-Methyl-1-butanol (mg/L)	24	19
3-Methyl-1-butanol (mg/L)	100	94
Hexanol (mg/L)	1.46	1.33
Benzyl alcohol (mg/L)	trace	trace
β-Phenylethanol (mg/L)	17.1	16.7
Acids		
3-Methylbutanois acid (mg/L)	0.55	0.57
Caproic acid (mg/L)	3.70	0.57
Caprylic acid (mg/L)	5.63	4.98

Note: Test with immobilized yeast: yeast dry weight, 20 mg/L gel; 4 g of beads/bottle. Chemicals were analyzed at the end of the fermentation.
Source: From Ref. 114.

are other difficulties to overcome for the industrial application of calcium alginate immobilization. The possibility of supplying wineries with ready-to-use immobilized cells would be a significant step for the implementation of this technique. For this, Loureiro (110) has performed freeze-drying experiments of cells immobilized by gel occlusion. These tests with different polysaccharide gels showed that it was possible to freeze-dry without changing bead shape using either κ-carrageenan or κ-carrageenan plus calcium alginate. However, cell release always occurred, and further studies will be necessary to optimize the thickness and rigidity of the protective bead layer for controlling cell release.

Lemmonier and Duteurtre (109) developed a device consisting of a small cartridge containing two microporous membranes, one hydrophilic and the other hydrophobic, which is filled with yeast and placed in the neck of the bottle. The two membranes hold the yeasts captive, enabling a free and total exchange between the yeasts and the wine. When the cap is opened, the cartridge is automatically ejected. The first results showed that production of sparkling wine with immobilized yeasts occurs in a manner similar to that using traditional methods with free cells. Some differences have been noted regarding protein and polysaccharides (118).

The use of a continuous fermentative system with a calcium alginate immobilized yeast bioreactor was tested by Fumi et al. (115) for the production of sparkling wines in tanks. The experimental system is shown in Figure 9. The general idea was to run the second fermentation outside the principal stainless steel tank using equipment that could easily be moved from one tank to another. The tests were carried out by placing 2 L of 1.5% alginate and 8% yeast beads (yeast inoculum was 250 g dry weight) or 3 L of 1.5% alginate and 10% yeast (yeast inoculum was 500 g dry weight) distributed or not on five distanced supports of pierced stainless steel, in a 10-L bioreactor. The base wine (9.95% v/v ethanol, 5.61 g/L titratable acidity, pH 3.04) was continuously supplied at a rate of 5 L/h and sucrose syrup at 0.14 L/h. The results showed that a high concentration of immobilized yeasts and a distribution of immobilized yeast mass in the bioreactor greatly affect the fermentation rate, making the process faster than the traditional one (78 h, as compared to 10 to 15 days needed for the traditional process). The sparkling wines have a composition (Table 7) comparable to that of traditional sparkling wines, and sensory evaluation (expert panel of eight tasters) of immobilized test versus control did not show significant differences at a level of $p < 0.05$.

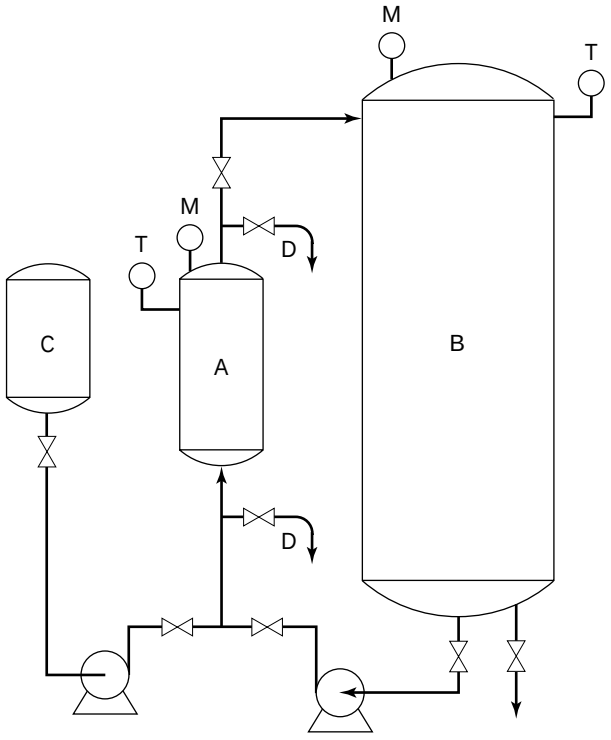


Figure 9. Scheme of the system used in the production of sparkling wine in tanks with immobilized yeast (117); A, bioreactor with immobilized yeast; B, sealed tank; C, syrup tank; D, sampling taps; M, manometers; T, thermometers.

The influence of cell-recycle batch fermentation process on sparkling wines produced in pressure tanks was studied by Ciani et al. (119). They noted a reduction in the time of fermentation as well as an improvement in ethanol productivity and yield. This was linked to the increased inoculum and also to the disappearance of the lag phase, which lasted approximately 150 h in batch. The increase in ethanol yield is probably correlated with a resting cell state highlighted by a lower percentage of new cells produced in steady-state conditions (120). The analytical characteristics of the sparkling wine obtained were not significantly different from those of sparkling wine produced by traditional batch fermentation.

As far as the fermentative process of must is concerned, Cantarelli (50) conducted comparative tests with free and calcium alginate immobilized yeasts. A fluidized or

Table 7. Chemical Composition of Sparkling Wines Obtained with Charmat Methods by Free and Immobilized Yeasts

	Free yeast	Immobilized yeast
Ethanol (v/v %)	10.60	10.62
Sugar (g/L)	4.5	4.7
pH	3.07	3.07
Total acidity (g/L)	5.70	5.78
Volatile acidity (g/L)	0.25	0.28
Tartaric acid (g/L)	2.37	2.28
Malic acid (g/L)	0.58	0.53
Lactic acid (g/L)	1.81	1.85
Succinic acid (g/L)	0.41	0.47
Glycerol (g/L)	4.79	5.08
Acetaldehyde (mg/L)	49	53
Total nitrogen (mg/L)	205	199
Calcium (mg/L)	81	87
n-Propanol (mg/L)	30	30
Isobutyl alcohol (mg/L)	55	62
Amyl alcohols (mg/L)	125	125
Higher alcohols (mg/L)	210	217
OD_{660nm}	>1	0.003

Source: From Ref. 115.

packed-bed reactor consisting of a 200-L tubular fermenter with immobilized yeast beads at various concentrations (maximum 8 g/L) was used for wine production. The results showed that immobilized yeast, compared with free yeast, had almost linear fermentation kinetics without a lag or exponential phase. In addition, organoleptic and chemiometric evaluation of analytical data both showed that there were no significant differences in by-products. Therefore, the final by-product content of wine produced by a fluid-bed fermenter was lower than by a packed one.

Batch fermentation tests on Riesling grape juice were carried out by Fumi et al. (113) using two 50-L working volume fermenters with calcium alginate immobilized biocatalysts (beads with alginate/yeast ratio of 1.5 wt/dry wt, reintroduced into 1% alginate and supported by stainless steel pierced plates vertically disposed) with surface area to volume ratios 1.33 and 0.66, respectively. The immobilized-system kinetic parameters had similar values, and they were not influenced by the different surface area to volume ratios of the biocatalysts. Analytical data showed no detectable differences between the wines obtained by the immobilized system and wines produced traditionally. During the first phases of fermentation with immobilized cells, the must was kept perfectly clean in both bioreactors. Nevertheless, at the end of the fermentation process, a free-cell deposit, caused by yeast leakage from the beads, was found in the bioreactors.

Selected Lactic Acid Bacteria and Innovation in Malolactic Fermentation

Malolactic fermentation is an important factor for many red wines or at least for those high acid wines produced from grapes grown in cool climate (121,122).

This process is effected by a strain of lactic bacteria of the genera *Leuconostoc*, *Lactobacillus*, and *Pediococcus* (67,122,123) and essentially consists of the degradation of malic acid into lactic acid and carbon dioxide, the consequence of which is a reduction in total acidity (deacidification). In addition, the transformation of malic acid is always accompanied by the fermentation of sugars, which although on a modest scale has the result of increasing aroma components such as acetaldehyde, acetic acid, acetoin, diacetyl, 2,3-butandiol, ethyl lactate, and high alcohols (67,68,122). As a whole, the wine gains in mellowness, roundness, and fullness and becomes more pleasing to the palate (68,123).

The difficulties that hinder growth of lactic bacteria in wine have causes known only in part (pH, ethanol concentration, SO_2 concentration, fermentation temperature, competition between lactic bacteria and other microorganisms, lack of nutrients). Thus, spontaneous biological deacidification of wines is not always possible. To help in the deacidification of wines, it is a common practice to make use of an inoculum of active indigenous bacteria effected through the addition of wines where malolactic fermentation is in process or has just finished. However, this system does not always give good results and is unsuitable for those wines that must maintain their peculiar characteristics unaltered (122).

For quite some time, attempts have been made to induce malolactic fermentation by inoculating the wine with selected lactic acid bacteria (121). For this purpose, lactic bacteria have been isolated from musts and wines and have been selected for their tolerance of ethanol (13 to 14 vol.%) and sulphur dioxide, their capacity to develop in the pH of wine, their capacity to produce small quantities of acetic acid, and their capacity to not alter organoleptic characteristics (121,124,125). Dried or frozen cultures of selected bacteria suitable for malolactic fermentation have been on the market for some years. Research carried out thus far on the efficacy of these preparations has shown that malic acid is broken down completely if the strain is previously reactivated in suitable substrata (67,126–130). Lonvaud-Funel (128) observed that complete breakdown of malic acid is achieved by using an inoculum containing up to 10^7 cells/mL that has been allowed to develop previously for 24 h in either a synthetic substrate (40 g/L glucose, 3 g/L malic acid, 5 g/L of neopeptone, 50 mg/L of Mg^{2+}, 20 mg/L of Mn^{2+}, 50 g/L of yeast extract, pH 4.5) or in a grape juice base substrate.

Some investigators (131–136) have suggested the use of immobilized lactic bacteria or enzymes to convert malic acid into lactic acid.

Spettoli et al. (136) immobilized *Leuconostoc oenos* ML34 in 1.67% calcium alginate gels and tested their ability to convert L-malic acid into L-lactic in a red wine having 10.20% v/v alcohol; 7.90 g/L of titratable acidity; pH 3.15; 21 mg/L of total SO_2, and 1.50 g/L of L-malic acid. The cells, in batch reaction, maintained 56% conversion of L-malic acid after 16 h.

Cuenat and Villettaz (132) immobilized a strain of *Leuconostoc oenos* on alginate (1 vol of 15% dry weight cells in 5 to 12 vol of alginate, 5% (w/v)). A complete degradation of malic acid (from 6.1 to 0.07 g/L) was obtained in a continuous system where 25 L of wines (pH 3.3; 11.3% (v/v) ethanol, 139 mg/L total SO_2) was passed through the columns in 5.5 h.

Naouri et al. (134) compared the malic-acid-degrading ability of a strain of *Leuconostoc oenos* with that of a strain of *Lactobacillus* sp. immobilized in alginate in a high compacting multiphase reactor (HCMR) containing 170 g of alginate beads and 1.41 L of red wines. After bioconversion in the reactor, malic acid was almost completely (82%) degraded by the immobilized *Leuconostoc* strain, whereas incomplete (45%) degradation was obtained with the *Lactobacillus* strain.

Crapisi et al. (131) reported on the efficiency of a continuous-flow bioreactor filled with *Lactobacillus brevis* immobilized on four types of alginate and κ-carrageenan for the control of malolactic fermentation in wine. Bioreactor efficiency seems to be dependent on gel properties ranging from 34 to 45%, depending on the immobilization matrix used.

Attempts to deacidify wine using malolactic enzyme immobilized on gels have not been successful (137) probably because of the inactivity of the enzyme at wine pH and because the required cofactor, NAD, is particularly unstable in wine.

CONCLUSION

Factors limiting the application of standardized biotechnological processes to wine production include the following:

- The seasonality and high variability in grape composition
- The limited annual season of vintage
- The length of the winemaking process for some wines, which delays experimental results
- The variations in the amount of wine production (ranging from a few hundred hectolitres to more than 200.000)

In fact, under the generic term wine there is a diversity and hierarchy of quality that is quite unique among the products of the soil and that is connected to the cultivar, the area of cultivation, and the transformation technologies applied to the grapes. The latter includes the transformation of the must by yeast fermentation, which is certainly the oldest of biotechnologies. The principles of winemaking have been established empirically over the course of several centuries of rigorous and methodical observation. The production of quality wines came before all scientific knowledge.

However, although its origins were empirical, the production of quality wine has, in the last quarter of a century, benefited most from research developments. In this area, biotechnological knowledge may well become more and more important, not only as regards the microorganisms involved in the fermentation process but also because of the use of enzymes in pressing grapes, prefermentation treatments, and refinement of the wines.

Since the development of the use of pectolytic enzymes, the tendency has been to search for enzymic preparations that have precise effects and that allow us to reduce the use of chemical additives. In this field, studies have made it possible to show that there are numerous limitations to the use of enzymes in the field of enology. For example, experiments with enzymes to improve the stabilization of white wines and to increase the aroma of some wines have demonstrated that the composition of the medium and the working conditions are not very favorable to the action of enzymes such as proteases and glycosidases. At present, the use of extracellular yeast enzymes, such as proteases, able to act on the constituents of the must are being studied (138,139). This is obviously of technological interest, because it would make it possible to take advantage of the enzyme activity of the yeast, not only for fermentation but also for stabilization of wines. In addition, the energy liberated during the fermentation process would give temperatures better suited to the activity of the enzymes.

These days a selected yeast is required not only to ensure that undesirable characteristics do not arise but also in order to have a positive effect on the characteristics of the wine. Methods of genetic improvement offer a means of programming and constructing new strains of yeasts. However, certain problems remain when one attempts genetic improvement of the strains, not the last among which is the lack of objective methods of evaluation of the quality of the yeast. The contribution of yeast to the quality of wine is complex because it is the result of interactions between grapes, yeast strains, and technology. In spite of this, the application of recombinant DNA technologies to the improvement of wine strains has undergone rapid development, and although practical applications have not yet been achieved, the basis for future advances has been established.

In the field of bottled sparkling wine production, studies have led to the creation of technologies that permit the use of alginate-immobilized yeasts at a semi-industrial level (108). However, the technology of entrapped yeast for must fermentation requires further optimization of immobilized systems, in the areas of mass transfer, reactor design, and yeast physiology. The most logical strategy for providing the means of increasing productivity without changing the flavor characteristics of fermented products must be found.

Reactors with immobilized lactic acid bacteria have also been shown to be potentially very interesting, but putting them into operation on an industrial scale still requires critical evaluation, not only of the efficiency of the breakdown of malic acid but also of the results on the organoleptic characteristics of the wines.

Every biotechnological advance that is to be accepted in winemaking must allow one to maintain the quality of the product, must take into account the hygienic and legal aspects as well, and must not impair the flavor and bouquet of the final product.

BIBLIOGRAPHY

1. D. Dubourdieu, *Experientia* **42**, 914–921 (1986).
2. A.C. Noble, in H.F. Linskens and J.f. Jackson eds., *Wine Analysis: Modern Methods of Plant Analysis*, vol. 6, Springer-Verlag, Berlin, 1988, pp. 9–28.
3. A. Rapp and H. Mandery, *Experientia* **42**, 873–884, (1986).
4. P. Schreier, *Crit. Rev. Food Sci. Nutr.* **12**, 59–111, (1979).
5. O. Colagrande, *Ann. Fac. Agraria U.C.S.C.* **28**, 99–126 (1988).
6. R.M. Canal-Llauberes, *Rev. Oenol.* **53**, 17–22 (1989).
7. R.M. Canal-Llauberes, *Rev. Franc. Oenol.* **30**, 28–33 (1990).
8. G. Montedoro, *Riv. Sci. Tech. Alim. Nutr. Um.* **6**, 133–144, (1976).
9. J.C. Villettaz, In E. Lemperle and H. Rasenberger, eds., *Proc. of 7th Int. Oenological Symposium, Rome, Italy*, International Association for Modern Winery Technology and Management, Breisach, Germany, 1984, pp. 99–119.
10. A. Zamorani, in C. Cantarelli, G. Lanzarini, eds., *Biotechnology Application in Beverage Production*, Elsevier Applied Science Publishers, New York, 1989, pp. 223–246.
11. M. Feuillat, *Rev. Oenol.* **45**, 7–17 (1987).
12. L. Saulnier, J.M. Brillouet, and M. Montounet, *Conn. Vigne Vin* **22**, 135–158, (1988).
13. L. Usseglio-Tomasset, *Conn. Vigne Vin* **10**, 193–226, (1976).
14. A. Baron, *Rev. Franc. Oenol.* **30**, 21–27 (1990).
15. J.M. Brillouet, L. Saulnier, and M. Montounet, *Rev. Franc. Oenol.* **30**, 43–54 (1990).
16. A. Silva, O. Colagrande, and P. Fontana. In G. Domenichini ed., *Proc. of the 4th Symposium La Difesa Antiparassitaria Nelle Industrie Alimentari e la Protezione Degli Alimenti* Camera di Commercio Industria Artigianato e Agricoltura, Piacenza, Italy, 1989, pp. 389–407.
17. C.S. Ough and E.A. Crowell, *Am. J. Enol. Vitic.* **30**, 22–27, (1979).
18. M.R. Brown and C.S. Ough *Am. J. Enol. Vitic.* **32**, 272–276 (1981).

19. P. Martinière, J.C. Sapis, G. Guimberteau, and J. Ribéreau-Gayon, *Compt. Rendus Acad. Agric. (France)* **59**, 267–273 (1973).
20. M. Castino and P. Bella, *Riv. Vitic. Enol.* **34**, 179–197 (1981).
21. P. Berta, *Vigne Vini* **18**, 37–42 (1991).
22. M. Castino, A. Bosso, and G. Marescalco, *Vini Ital.* **32**, 7–20 (1990).
23. D. Dubourdieu, P. Ribéreau-Gayon, and B. Fournet, *Carbohydr. Res.* **93**, 294–299 (1981).
24. J. Ribéreau-Gayon, P. Ribéreau-Gayon, and G. Seguin, in J.R. Coley-Smith, K. Verhoff, W.R. Jarvis eds., *The Biology of Botrylis*, Academic Press, London, 1980, pp. 251–274.
25. D. Dubourdieu, J.C. Villettaz, C. Desplanques, and P. Ribéreau-Gayon, *Conn. Vigne Vin* **15**, 161–177 (1981).
26. J.C. Villettaz, *Rev. Franc. Oenol.* **27**, 65–72, (1987).
27. J.C. Villettaz, *Rev. Franc. Oenol.* **30**, 59–63, (1990).
28. K. Wucherpfenning and H. Dietrich. *Bull. O.I.V.* **56**, 733–745, (1983).
29. *European Community Official Gazette* 1987, **30**, L86.
30. *European Community Official Gazette* 1988, **31**, L198.
31. M. Feuillat, Ph.D. Thesis, Université de Dijon, France, May 30, 1974.
32. P.R. Ngaba-Mbiahop, Ph.D. Thesis, Oregon State University, 1981.
33. J.C. Hsu and D.A. Heatherbell, *Am. J. Enol. Vitic.* **38**, 6–10 (1987).
34. J.C. Hsu and D.A. Heatherbell, *Am. J. Enol. Vitic.* **38**, 11–16 (1987).
35. A.T. Bakalinsky and R. Boulton, *Am. J. Enol. Vitic.* **36**, 23–29 (1985).
36. B.S. Gaina, N.M. Pavlenko, E.N. Datunashvili, Y.I. Krylova, L.V. Kozlov, V.K. Antonov, *Appl. Biochem. Micro.* **12**, 167–173 (1976).
37. K. Woiwodov, B. Galunsky, S. Djankov, N. Gorinova, and D. Tzakov, *Mitt. Klosterneuburg* **32**, 117–121, (1982).
38. E.J. Modra, P.J. Williams, T.H. Lee, and W. Wallace, in P. Ribéreau-Gayon and A. Lonvaud eds., *Actualités Oenologiques 89, Proc. of the 4th Symposium International d'Oenologie, Bordeaux 1989*, Dunod, Paris, 1990, pp. 217–221.
39. S. Arai, G. Charalambous ed., *The Analysis and Control of Less Desirable Flavors in Foods and Beverages*, Academic Press, New York, 1980, pp. 133–147.
40. J.J. Macheix, J.C. Sapis, and A. Fleuriet, *Crit. Rev. Food Sci. Nutr.* **30**, 441–486 (1991).
41. A.C. Noble, in R.L. Rouseff ed., *Bitterness in Foods and Beverages: Developments in Food Science*, vol. 25, Elsevier, Amsterdam, 1990, pp. 145–158.
42. V.L. Singleton, *Am. J. Enol. Vitic.* **38**, 69–77 (1987).
43. H.L. Wildenradt and V.L. Singleton *Am. J. Enol. Vitic.* **25**, 119–127 (1974).
44. G. Bonaga, U. Pallotta, and K. Syrghi, *Vini Ital.* **32**, 13–30, **32**, 31–38 (1990).
45. U. Pallotta and C. Cantarelli, *Vigne Vini* **6**, 19–46 (1979).
46. O. Colagrande, R. Ottina, and M.D. Fumi, *Vini Ital.* **28**, 43–56 (1986).
47. C. Cantarelli, *Vini Ital.* **28**, 87–98 (1986).
48. C. Cantarelli, G. Giovannelli, and S. Gallizia, in Atti Simposio *Criteri razionali di controllo e di impiego dei Coadiuvanti Enologici*, Università Cattolica del S. Cuore, Istituto di Enologia, Piacenza, Italy, 1989, pp. 55–67.
49. C. Cantarelli, O. Brenna, G. Giovannelli, and M. Rossi, *Food Biotechnol.* **3**, 203–213 (1989).
50. C. Cantarelli, C. Cantarelli and G. Lanzarini eds., in *Biotechnology Application in Beverage Production*, Elsevier Applied Science, New York, 1989, pp. 127–151.
51. P. Ribéreau-Gayon, J.N. Boidron, and A. Terrier, *J. Agric. Food Chem.* **23**, 1042–1047 (1975).
52. P.J. Williams, C.R. Strauss, and B. Wilson *Am. J. Enol. Vitic.* **32**, 230–235 (1981).
53. R. Di Stefano, *Vini Ital.* **130**, 29–43 (1981).
54. R. Di Stefano, *Vigne Vini* **9**, 45–47 (1982).
55. Z. Gunata, C. Bayonove, R. Baumes, and R. Cordonnier, *J. Sci. Food Agric.* **36**, 857–862 (1985).
56. P.J. Williams, C.R. Strauss, and B. Wilson *Phytochemistry* **19**, 1137–1139 (1980).
57. S. Voirin, R. Baumes, S. Bitteur, Z. Gunata, and C. Bayonove, *J. Agric. Food Chem.* **38**, 1373–1378 (1990).
58. P.J. Williams, C.R. Strauss, B. Wilson, R.A. Maay-Westropp, *Phytochemistry* **21**, 2013–2020 (1982).
59. Z. Gunata, C. Bayonove, R. Baumes, and R. Cordonnier, *J. Chromatogr.* **331**, 83–90 (1985).
60. R. Di Stefano, *Riv. Vitic. Enol.* **42**, 11–23 (1989).
61. C. Bayonove, Y.Z. Gunata, J.C. Sapis, and R.L. Baumes, *Rev. Oenol.* **18**, 15–18 (1992).
62. Z. Gunata, S. Bitteur, J.M. Brillouet, C. Bayonove, and R. Cordonnier, *Carbohydr. Res.* **184**, 139–149 (1988).
63. Z. Gunata, S. Bitteur, R. Baumes, J.C. Sapis, and C. Bayonove, *Rev. Franc. Oenol.* **30**, 37–41 (1990).
64. R.E. Cordonnier, Z. Gunata, R.L. Baumes, and C.L. Bayonove, *Conn. Vigne Vin* **23**, 7–23 (1989).
65. Z. Gunata, I. Dugelay, J.C. Sapis, R. Baumes, and C. Bayonove, *J. Int. Sci. Vigne Vin* **24**, 133–144 (1990).
66. R.E. Kunkee and M.A. Amerine, in A.H. Rose and J.S. Harrison eds., *The Yeasts* Academic Press, New York, 1970, pp. 5–71.
67. S. Lafon-Lafourcade, in H.-J. Rehm and G. Reed eds., *Biotechnology*, Verlag Chemie, Weinheim, 1983, pp. 81–116.
68. F. Radler, *Experientia* **42**, 884–893 (1986).
69. A. Martini, F. Federici, and G. Rosini, *Can. J . Microbiol.* **26**, 856–859 (1980).
70. H.J. Phaff, M.W. Miller, and E.M. Mrak, *The Life of Yeasts*, Harvard University Press: Cambridge, Mass., 1978.
71. G. Rosini, F. Federici, and A. Martini, *Microb. Ecol.* **8**, 93–97 (1982).
72. R.E. Kunkee, *Food Microbiol.* **1**, 315–332 (1984).
73. C. Zambonelli, P. Romano, and G. Suzzi in C. Cantarelli, G. Lanzarini, eds., *Biotechnology Application in Beverage Production*, Elsevier Applied Science Publishers, New York, 1989, pp. 17–30.
74. H. Bussey, *Adv. Microb. Physiol.* **22**, 93–122 (1981).
75. C. Geneix, S. Lafon-Lafourcade, and P. Ribereau-Gayon, *C.R. Acad. Sci. Paris* **296**, 943–945 (1983).
76. S. Lafon-Lafourcade, *Experientia* **42**, 904–914 (1986).
77. S. Lafon-Lafourcade, D. Dubourdieu, D. Hadjinicolaou, and P. Ribéreau-Gayon, *Conn. Vigne Vin.* **14**, 97–109 (1980).
78. S. Lafon-Lafourcade, C. Geneix, and P. Ribereau-Gayon, *Appl. Envir. Micribiol.* **47**, 1246–1249 (1984).
79. P. Pfeiffer and F. Radler *J. Gen. Microbiol.* **128**, 2699–2706 (1983).

80. R.B. Wickner, in E.W.J. Strathern and J.R. Broach eds., *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance*, Cold Spring Harbor Laboratory, New York, 1981, pp. 415–444.
81. C. Zambonelli, *Accademia Italiana Vite e Vino* **29**, 51–64 (1977).
82. L.F. Bisson, *World Biotech. Rep. Food Process.* **2**, 103–116 (1986).
83. C.V. Rous, R. Snow, and R.E. Kunkee, a leucine *J. Inst. Brew.* **89**, 274–278, (1983).
84. C. Falcone and L. Frontali, C. Cantarelli, and G. Lanzarini eds., in *Biotechnology Application in Beverages Production*, Elsevier Applied Science, London, 1989, pp. 31–39.
85. I.S. Pretorius and T.J. van der Westhuizen, *S. Afr. J. Enol. Vitic.* **12**, 3–31 (1991).
86. R. Eschenbruch, K.J. Cresswell, B.M. Fisher, and R.J. Thornton, *Europ. J. Appl. Microbiol. Biotechnol.* **14**, 155–159 (1982).
87. F. Vezinhet, *Sci. Aliments* **9**, 253–265, (1989).
88. P.C. Molan, M. Edward, and R. Eschenbruch, *Eur. J. Appl. Microbiol. Biotechnol.* **16**, 110–113, (1982).
89. F. Vezinhet, *Rev. Franc. Oenol.* **25**, 45–51, (1985).
90. G.A. Farris, F. Fatichenti, L. Bifulco, L. Berardi, P. Delena, and T. Satta *Biotechnol. Lett.* **14**, 219–222 (1992).
91. S. Hara, Y. Iimura, H. Oyama, T. Kozeki, K. Kitano, K. Otsuka, *Agric. Biol. Chem.* **45**, 1327–1330 (1981).
92. T. Seki, E.H. Choi, and D. Ryu *Appl. Environ. Microbiol.* **49**, 1211–1215, (1985).
93. T.W. Young, in A.H. Rose and J.S. Harrison eds., *The Yeasts*, vol. 2, Academic Press, New York, 1987, pp. 131–164.
94. P. Romano, M.G. Soli, G. Suzzi, and L. Grazia *Appl. Environ. Microbiol.* **50**, 1064–1067 (1985).
95. R.J. Thornton, *Am. J. Enol. Vitic.* **36**, 47–49, (1985).
96. F. Vezinhet and P. Barre, in G. Durand, L. Bobichon, J. Florent, eds., *8th International Biotechnology Symposium Paris 1988*, Abstract A43, p. 101–103.
97. J. Watari, Y. Takata, M. Ogawa, N. Nishikawa, and M. Kamimura *Agr. Biol. Chem.* **53**, 901–903, (1989).
98. S.A. Williams, R.A. Hodges, T.L. Strike, R. Snow, and R.E. Kunkee, *Appl. Environ. Microbiol.* **47**, 288–293, (1984).
99. R. Snow, in J.F.T. Spencer, D.M. Spencer, A.R.W. Smith, eds., *Yeast Genetics: Fundamental and Applied Aspects*, Springer Verlag, New York, 1983, pp. 439–459.
100. C. Zambonelli, M.G. Soli, and D. Guerra *Ann. Microbiol.* **34**, 7–15, (1984).
101. C. Gjermansen and P. Sigsgaard, *Carlsberg Res. Commun.* **46**, 1–11 (1981).
102. I. Herskowitz, *Microbiol. Rev.* **52**, 536–553 (1988).
103. G.C. Stewart, *Can. J. Microbiol.* **27**, 973–990 (1981).
104. R.J. Thornton and R. Eschenbruch, *Antonie van Leeuwenhoek* **42**, 503–509 (1976).
105. M. Polsinelli, in O. Colagrande, ed., *Sviluppi della Biotechnologia Nella Produzione dello Spumante Classico*, Proc. of the 4th Simposio *Internazionale sul Vino, Pavia* Chiriotti Publishers, Pinerolo, Italy, 1990, pp. 19–22.
106. M.D. Fumi, G. Trioli, and O. Colagrande, *Produzione dello Spumante Classico, Proc. of the Simposio Internazionale sul Vino, Pavia*, Chiriotti Publishers, Pinerolo, Italy, 1987, pp. 94–102.
107. J. Tampion and M.D. Tampion, *Immobilized Cells: Principle and Applications*, Cambridge University Press, Cambridge, U.K., 1987.
108. B. Duteurtre, P. Ors, and D. Hennequin, in O. Colagrande, ed., *Sviluppi della Biotechnologia Nella Produzione dello Spumante Classico, Proc. of the 4th Simposio Internazionale sul Vino*, Pavia Chiriotti Publishers, Pinerolo, Italy, 1990, pp. 71–73.
109. J. Lemonnier and B. Duteurtre, *Rev. Franc. Oenol.* **29**, 15–26 (1989).
110. V. Loureiro, in O. Colagrande ed., *Sviluppi della Biotechnologia Nella Produzione dello Spumante Classico Proc. of the 4th Simposio Internazionale sul Vino, Pavia*, Chiriotti Publishers, Pinerolo, Italy, 1990, pp. 74–77.
111. A. Zamorani, A. Crapisi, G. Borin, M. Dell'Eva, G. Versini, and P. Spettoli, *Yeast* **5**, 63–66 (1989).
112. M.D. Fumi, G. Trioli, and O. Colagrande, *Biotechnol. Lett.* **9**, 339–342 (1987).
113. M.D. Fumi, G. Trioli, and O. Colagrande *Proc. 4th Europ. Congr. Biotechnol.*, Amsterdam, June 14–19, Elsevier, 1987, pp. 81–84.
114. M.D. Fumi, G. Trioli, M.G. Colombi, and O. Colagrande *Am. J. Enol. Vitio.* **39**, 267–272 (1988).
115. M.D. Fumi, M. Bufo, G. Trioli, and O. Colagrande *Biotechnol. Lett.* **11** 821–824 (1989).
116. C. Divies, in C. Cantarelli, and G. Lanzarini, eds., *Biotechnology Application in Beverages Production*. Elsevier Applied Science, London, 1989, pp. 153–167.
117. P. Ors, B. Duteurtre, D. Hannequin, in P. Ribéreau-Gayon, A. Lonvaud, eds., *Actualités Oenologiques 89, Proc. of the 4th Symposium International d'Oenologie, Bordeaux 1989*, Dunod, Paris, 1990, pp. 270–274.
118. M. Feuillat, in *Progressi Scientifici e Tecnici Nella Moderna Spumantistica Italiana* Rovato, 1991, pp. 25–40.
119. M. Ciani and G. Rosini, *Biotechnol. Lett.* **13**, 533–536 (1991).
120. G. Rosini, *Appl. Microbiol. Biotechnol.* **24**, 140–143, (1986).
121. C.R. Davis, D. Wibowo, R. Eschenbruch, T.H. Lee, and G.H. Fleet, *Am. J. Enol. Vitic.* **36**, 290–301 (1985).
122. R.E. Kunkee, in A.D. Webb ed., *Chemistry of Wine*, American Chemical Society, Washington D.C., 1974, pp. 137, 151–170.
123. D. Wibowo, R. Eschenbruch, C.R. Davis, G.H. Fleet, and T.H. Lee, *Am. J. Enol. Vitic* **36**, 302–313, (1985).
124. P. Chagnaud, P. Nauri, A. Arnaud, and P. Galzy, *Belgian J. Food Chem. Biotechnol.* **45**, 12–20 (1990).
125. S. Lafon-Lafourcade, in J.G. Carr and C.V. Cutting eds., *Lactic Acid Bacteria in Beverages and Food*, Academic Press, London, 1975, pp. 43–53.
126. A. Joyeux and A. Lonvaud-Funel, *Conn. Vigne Vin* **19**, 149–159 (1985).
127. D.C. Hayman and P.R. Monk, *Food Technol. Australia* **34**, 14–18 (1982).
128. A. Lonvaud-Funel, *Application à l'Oenologie des Progrès Récents en Microbiologie et en Fermentation*, Paris 1988, Programme COMETT, O.I.V., Paris, pp. 91–103.
129. P. Naouri, P. Chagnaud, A. Arnaud, P. Galzy, and J. Mathieu, *J. Biotechnol.* **10**, 135–150, (1989).
130. M. Valade, in *Application à l'oenologie des Progrès Récents en Microbiologie et en Fermentation*, Programme COMETT, Paris, 1988, OIV, Paris, pp. 71–80.
131. A. Crapisi, A. Lante, M.P. Nuti, A. Zamorani, and P. Spettoli, in P. Ribéreau-Gayon, and A. Lonvaud, eds., *Actualités Oenologiques 89, Proc. of the 4th Symposium International d'Oenologie Bordeaux 1989*, Dunod, Paris, 1990, pp. 349–352.
132. Ph. Cuenat and J.CL. Villettaz, *Rev. Suisse Vitic. Arboric, Hortic.* **16**, 145–151 (1984).

133. C. Divies and M.H. Siess *Ann. Inst. Pasteur. Microbiologie Paris* **127**, 525–539 (1976).

134. P. Naouri, P. Chagnaud, A. Arnaud, P. Galzy, and J. Matthieu *J. Wine Res.* **2**, 5–20, (1991).

135. J. Rossi and F. Clementi, *Am. J. Enol. Vitic.* **35**, 100–102, (1984).

136. P. Spettoli, A. Bottacin, M.P. Nuti, and A. Zamorani, *Am. J. Enol. Vitic.* **33**, 1–5, (1982).

137. S. Gestrelius, *Enzyme Eng.* **6**, 245–250 (1982).

138. I. Rosi, L. Costamagna, and M. Bertuccioli, *J. Inst. Brew.* **93**, 322–324, (1987).

139. I. Rosi and M. Bertuccioli, in O. Colagrande ed., *Sviluppi della Biotecnologia Nella Produzione Dello Spumante Classico, Proc. of the 4th Simposio Internazionale sul vino*, Chiriotti Publishers, Pinerolo, Italy, 1990, pp. 43–48.

ADDITIONAL READING

C. Biron, R. Cordonnier, O. Glory, Y.Z. Gunata, and J.C. Sapis *Conn. Vigne Vin* **22**, 125–134, (1988).

C. Bucke, *Methods Enzymol.* **135**, 175–189, (1987).

G.B. Calleja, in A.H. Rose and J.S. Harrison eds., *The Yeasts*, vol. 2, Academic Press, New York, 1987, pp. 165–238.

P. Chatonnet, C. Barbe, R. Canal-LLauberes, D. Dubordieu and J.N. Boidron *J. Int. Sci. Vigne Vin* **26**, 253–269, (1992).

D. Dobourdieu, P. Darriet, P. Chatonnet, and J.N. Boidron in P. Riéreau-Gayon and A. Lonvaud eds., *Actualités oenologiques 89, Proc. of the 4th Symposium International d'Oenologie*, Dunod, Paris, 1990, pp. 151–159.

Y.Z. Gunata, Ph.D. Thesis, Academie de Montpellier-Université des Sciences et Techniques du Languedoc, April 19, 1984.

R. Felix and J.C. Villettaz, in T. Godfrey and J. Reichelt eds., *Industrial Enzymology*, Stockton Press, New York, 1983, pp. 410–421.

Th. Henick-Kling, in H.F. Linskens and J.F. Jackson eds., *Wine Analysis: Modern Methods of Plant Analysis* , vol. 6, Springer-Verlag, Berlin, 1988, pp. 276–316.

S. Lafon-Lafourcade, E. Carre, A. Lonvaud-Funel, and P. Ribereau-Gayon, *Conn. Vigne Vin.* **17**, 55–71 (1983).

G. Lanzarini and P.G. Pifferi, in C. Cantarelli and G. Lanzarini eds., *Biotechnology Application in Beverage Production*, Elsevier Applied Science Publishers, New York, 1989, pp. 189–222.

H.J. Peppler and G. Reed, in H-J. Rehm and G. Reed eds., *Biotechnology*, vol. 7, Verlag Chemie, Weinheim, 1987, pp. 584–586.

V.L. Singleton, in H.F. Linskens and J.F. Jackson eds., *Wine Analysis: Modern Methods of Plant Analysis*, vol. 6, Springer-Verlag, Berlin, 1988, pp. 173–257.

H.J.J. Vuuren and L.M.T. Dicks, *Am. J. Enol. Vitic.* **44**, 99–112, (1993).

See also ANAEROBES; CULTURE COLLECTIONS.

XANTHAN GUM

J.-L. FLORES CANDIA
W.-D. DECKWER
GBF-Gesellschaft für Biotechnologische Forschung GmbH
Braunschweig, Germany

KEY WORDS

Drug release control
Emulsifier
Enhanced oil recovery (EOR)
Food additives
Microbial exopolysaccharide
Non-Newtonian fluid
Oil drilling
Thickening agent
Xanthan
Xanthomonas campestris

OUTLINE

INTRODUCTION

A variety of polysaccharides are generated by microorganisms. Examples include alginate, curdlane, dextran, gellan, glucan, pullulan, and xanthan. Also cellulose can be produced by microbes. Among the microbial polysaccharides, xanthan plays a dominant role due to the relative easiness to produce it and as a result of its outstanding properties. Indeed, it has found widespread applications ranging from food and cosmetics additives to enhanced oil recovery. This polymer represents the fastest growing segment of the polysaccharide industry (Table 1). The largest growth is expected to be in food applications (8.3%/year), where demand for natural gums is decreasing. Xanthan gum is an exoheteropolysaccharide produced by the bacterium *Xanthomonas campestris*.

Structure and Properties

The xanthan polymer consists of repeated units of five sugars, namely, two D-glucose, two D-mannose, and one D-glucoronate, and varying amounts of acetate and pyruvate. Its primary structure is shown in Figure 1, based on the results of Jannson et al. (5) and Melton et al. (6), revised according to Hassler and Doherty (7), and confirmed by McIntyre et al. (8).

The polysaccharide backbone, made up of β-[1–4]-linked D-glucose units, is identical to that of cellulose. To alternate D-glucose units at the O-3 position, a trisaccharide side chain containing a D-glucoronosyl unit between two D-mannosyl units is attached. The terminal β-D-mannosyl unit is glycosidically linked to the O-4 position of the β-D-glucoronosyl unit, which in turn is glycosidically linked to the O-2 position of an α-D-mannosyl unit. The terminal D-mannosyl unit of the side chain contains a pyruvic acid moiety as a 4,6-cyclic acetal in variable degree. Finally, the nonterminal D-mannosyl unit is stoichiometrically substituted at O-6 with an acetyl group. Xanthan gum also contains monovalent cations, which can be Na, K, or ½ Ca.

The secondary structure is known to undergo an order–disorder conformational transition at particular salinity

Table 1. Annual Average Growth Rates of the Main Commercial Polysaccharides: Market Analysis and Forecast for the United States

	Annual average growth rate (%)		
Polysaccharide	Food	Petroleum	Cosmetics and toiletries
Starches	2.5	1.5	–
Natural gums			
Guar	2.1	–	–
Algins	2.7	–	3.7
Arabic	2.6	–	3.7
Carrageenan	4.5	–	3.7
Locust bean	1.9	–	3.1
Cellulosics			
CMC	3.5	0.5	3.0
HPG	–	1.5	–
HEC	–	0.5	–
MC	3.5	–	3.7
Xanthan	8.3	2.0	5.9
Pectin	5.2	–	–

Source: Business Communications Co.
Note: CMC, carboxymethylcellulose; HPG, hydroxypropyl guar; HEC, hydroxyethylcellulose; MC, methylcellulose.

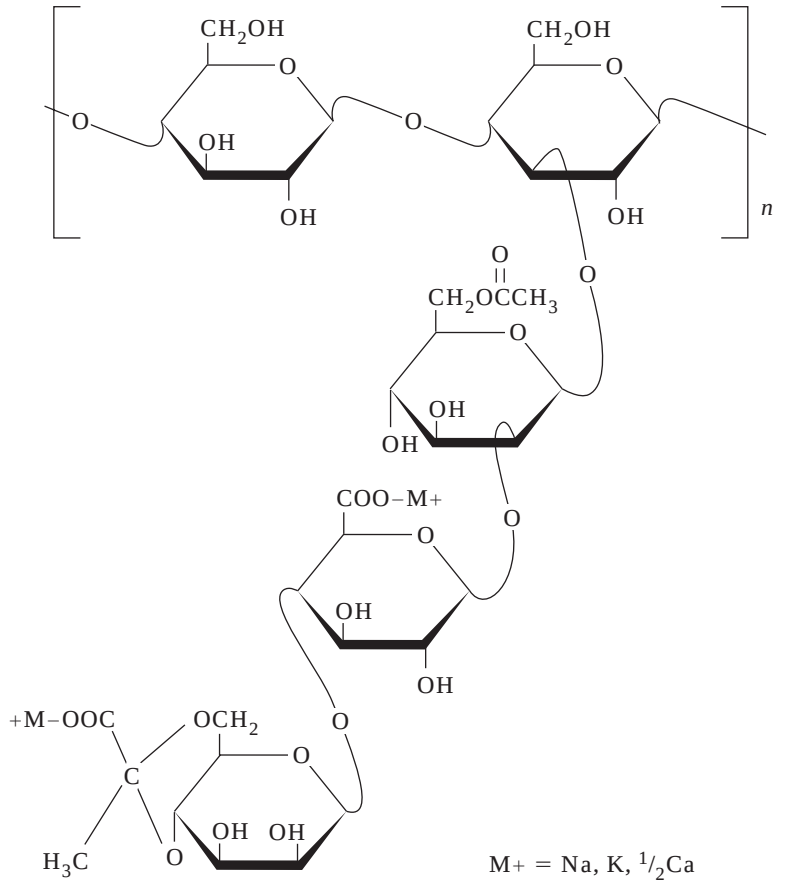


Figure 1. Chemical structure of the xanthan repeating unit.

and temperature levels, depending on the extent of pyruvate substitution (9,10). Most authors explain the stabilization of the ordered form by a noncovalent packing of the side chains along the backbone by H-bonding. X-ray fiber studies have indicated that the polymer adopts a right-handed helical conformation with fivefold symmetry and a helix pitch of 4.7 nm^3 (11). However, the question of whether this polymer dissolves in water as a single (11,12) or a double strand (13,14) has been a matter of controversial discussion. It now seems to be anticipated that xanthan may effectively be a single or a double strand, depending on the fermentation process, the acyl content, the thermal history, and the downstream process (12,15,16).

The most important property of xanthan is its ability to form solutions of high viscosity at rest or at low shear forces, which are highly pseudoplastic, and may display also a viscosity yield value. The viscosity of xanthan solutions is stable over a wide range of salt concentrations (up to 150 g/L NaCl), temperature (up to 90 °C) and pH (2–11). These solution characteristics of xanthan give rise to the functional properties such as thickening and stabilizing ability, which are used in a diverse range of applications. In addition, xanthan shows a synergistic interaction with various galactomannans (17,18), chitosan (19), and β-lactoglobulin (20), which results in enhanced solution viscosity or formation of gels, depending upon the type of mixture.

The rheological behavior of xanthan solution is most conveniently described by the two-parameter model of Ostwald and de Waele (power law)

$$\tau = k\dot{\gamma}^n \tag{1}$$

where τ is the shear stress and $\dot{\gamma}$ the shear rate. From equation 1, the effective viscosity, μ_{eff}, follows as

$$\mu_{\text{eff}} = k\dot{\gamma}^{n-1} \tag{2}$$

The fluid consistency factor, k, and the flow behavior index, n ($\leq$1 in case of the pseudoplastic xanthan solutions), are dependent on xanthan concentration and quality. Xanthan quality is predominantly determined by its molecular mass, but the pyruvate content has also a significant effect, though the influence of the pyruvate content on the viscosity yield is not yet clearly recognized. Flores-Candia (21) observed a sharp increase in viscosity (about a hundredfold) when the pyruvate content increased above ca. 3% wt. This is in accordance with Cadmus et al. (22) who stated that many industrial applications require a xanthan with a pyruvate content above 3.3% wt. The effect of xanthan molecular mass (pyruvate content $\geq$3.5 wt) on viscosity was studied by Herbst et al. (23) and Peters et al. (24). Figure 2 demonstrates the strong dependency of the vis-

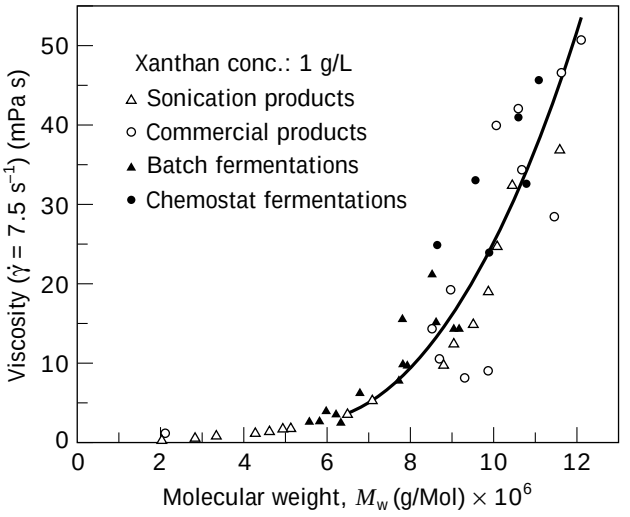


Figure 2. Effect of average molar mass on solution viscosity at a shear rate of 7.5 s^{-1} and a xanthan concentration of 0.1 % wt. *Source*: From Ref. 23.

cosity on the molecular mass of xanthan products of different origin.

Peters et al. (24) developed the following correlations for predicting the rheological parameters of equation 1 from the molar mass, M, (kg $kmol^{-1}$) and polymer concentration, c_p (kg m^{-3})

$$k = 10^{-3} + 2.58 \cdot 10^{-6}\left(\frac{M}{10^6}\right)^{3.69} c_p^{2.72} \quad (3)$$

$$\frac{1 - n}{n - 0.071} = 0.0227\left(\frac{M}{10^6}\right)^{2.08} (k - 10^{-3})^{0.47} \quad (4)$$

The thickening ability is related to the viscosity of a gum solution and can be defined as the resistance to flow of a liquid. Because of its high solution viscosity at low concentration and shear thinning behavior, xanthan is a particularly useful thickener. Xanthan is also used for the stabilization of suspensions, emulsions, and foams. This functional property is enhanced by a high solution viscosity and the existence of a viscosity yield value. Thus, if the suspended particles exert a force less than the yield value, they will be dispersed effectively (25). Again, the shear thinning behavior of xanthan is important because some stabilized products require a low viscosity during their application. Although xanthan does not possess the functional property of gelation, it will form thermoreversible gels with certain galactomannans. Gelation arises from intermolecular associations to form a three-dimensional continuous network that traps and immobilizes the water within it to form a rigid structure that shows resistance to flow under pressure.

BIOSYNTHESIS

Xanthan is produced by the bacterium *Xanthomonas campestris*, isolated originally from cabbage in which it causes black rot disease (26). *Xanthomonas* species are Gram-negative bacteria taxonomically placed in the Pseudomonadaceae family (27). According to the genetic basis of the classification, the average genome size of *Xanthomonas* is 2.5×10^9 Da. The percent G + C (guanine + cytosine) content of the DNA ranges from 63 to 71% (28). The cells have a rodlike form, slightly curved with rounded ends, measuring 0.2–0.6 by 0.8–2.9 μm. They occur mostly alone or in pairs, but chains and filaments are also observed. *X. campestris* forms smooth yellow mucoid colonies on solid media, and cells are surrounded by xanthan gum (extracellular polysaccharide). The chemical structure of the yellow pigment was determined to be mono- or dibromoaryl polyene and is probably bound to the cellular membrane. Genetic instabilities that affect quality and yield of xanthan have been observed in *Xanthomonas* cultivation (29). Genetic mutations are associated with the formation of L (large) and S (small) colony types. Whereas L-type colonies produce high xanthan yields with the desired rheological characteristics, the S-type colonies give low product yields with poor quality.

The natural function of xanthan is not fully recognized, but there is strong evidence indicating that the polysaccharide layer, surrounding the microbial cell, protects the microorganisms from various environmental factors. The water-retaining capacity of xanthan, for example, provides the microorganism with a protective layer under dry environmental conditions, thus providing some degree of desiccation resistance (30). Furthermore, recent reports of xanthan yield improvements in the presence of detergents (31) and bacterial-toxic acids (32) reveal that xanthan biosynthesis is induced under stress conditions.

In *X. campestris*, the Entner–Doudoroff pathway in conjunction with the tricarboxylic acid cycle pathway is the predominant mechanism for glucose catabolism (28,33,34). A small portion of glucose is routed via the pentose phosphate pathway (28,35). For glucose uptake, two discrete systems seem to exist (35). The biosynthesis of xanthan, as in most polysaccharide-producing bacteria, utilizes various activated carbohydrate donors to form the polymer on an acceptor molecule. The oligosaccharide repeated units of xanthan are constructed by the sequential addition of monosaccharides from sugar nucleotide diphosphates to an isoprenoid lipid acceptor molecule.

At the same time, acyl substitutes are added from appropriate activated donors. As shown in Figure 3, the xanthan backbone is formed by the successive addition of D-glucose-1-phosphate and D-glucose from two moles of UDP-D-glucose. Thereafter, D-mannose and D-glucoronic acid are added from GDP-mannose and UDP-glucoronic acid, respectively. O-Acetyl groups are transferred from acetyl-CoA to the internal mannose residue, and pyruvate from phosphoenolpyruvate is added to the terminal mannose. This pattern of reactions was elegantly demonstrated by Ielpi and his colleagues (36,37) using permeabilized cells and radioactively labeled precursors. Further, it was also shown that each of these steps, as just described, requires specific substrates and a specific enzyme for completion. If either the substrate or the enzyme is absent, the step is blocked (39–41).

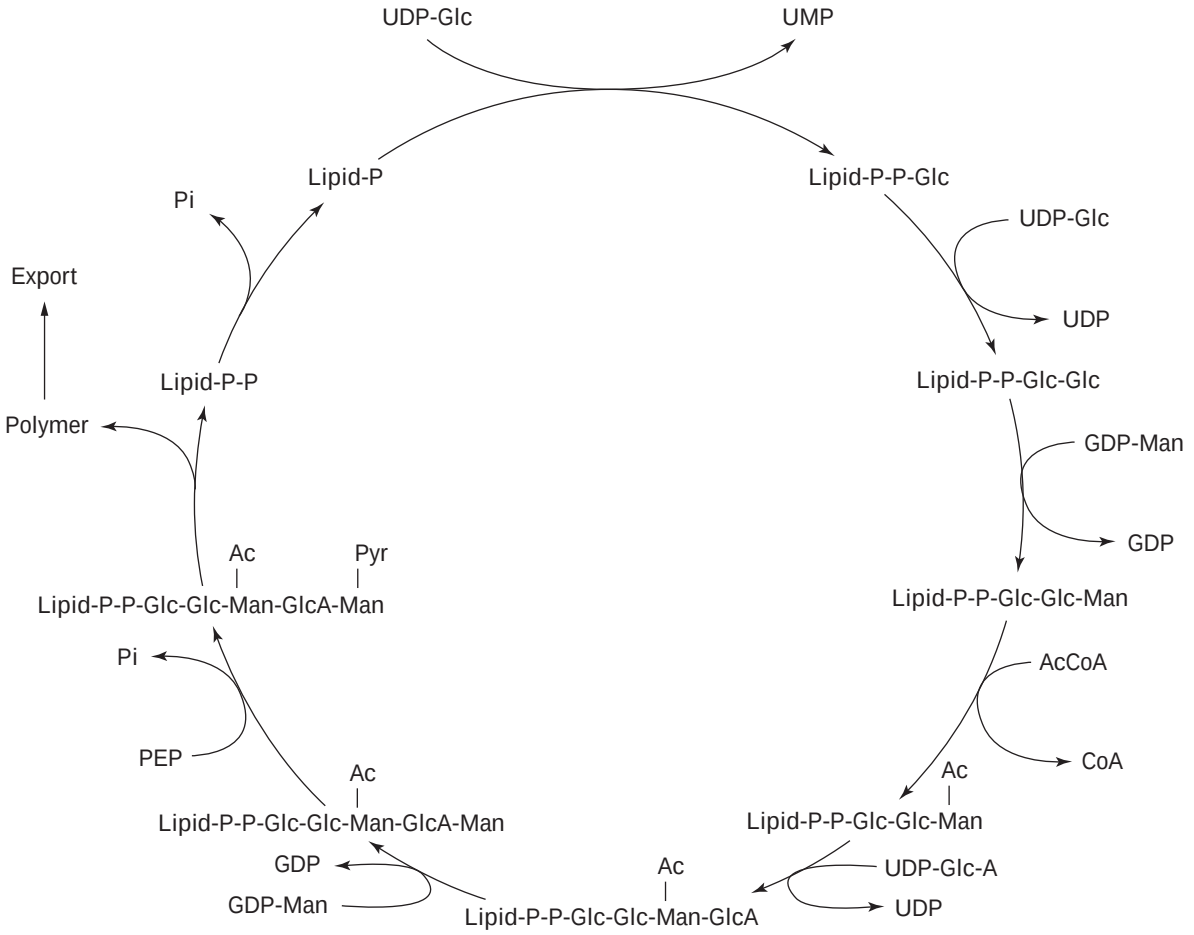


Figure 3. Biosynthesis of xanthan gum from activated precursors of carbohydrates and acyl groups. *Source*: From Refs. 36–38.

It has been suggested that the construction of the exopolysaccharide follows a "tail-to-head" polymerization (36–38,41). After the pentasaccharide repeated unit is formed, oligomers are formed by transfer to other lipid intermediates, gradually increasing the size of the carbohydrate chain. As noticed by Sutherland (38), oligomer construction normally involves the addition of the longer oligosaccharide sequence to the nonreducing terminus of a single repeat unit attached to the isoprenoid lipid diphosphate. The inactive lipid carrier is dephosphorylated to yield isoprenyl phosphate, which can then reenter the biosynthetic sequence.

Although the structure of the repeating units is determined by the sequential transfer of the different monosaccharides and acyl groups from their respective donors by highly specific sugar transferases, the polymerase enzyme responsible for polymerization of the pentasaccharides into a macromolecule of $M > 10^6$ has been shown to be less specific (38). The final stages of exopolysaccharide secretion from the cytoplasmic membrane, passing across the periplasm and outer membrane and finally excreted into the extracellular environment, are much less well specified than the previous biosynthetic events. This mechanism must exist in all polysaccharide-producing bacteria for releasing polymer from the isoprenoid lipid prior to transport to its final destination. The process obviously requires an energy source and may be analogous to the export of lipopolysaccharide to the outer membrane (42), in which ATP is the energy supplier.

Significant advances have been made in genetic research on *X. campestris*. Using transposon mutagenesis, Harding et al. (43) found a cluster of linked genes that encode enzymes for the assembly of the pentasaccharide. This same cluster was also isolated and analyzed by Thorne et al. (44). Vanderslice et al. (45) sequenced the DNA in this region and showed that it contains 12 genes (Fig. 4), which included seven gene products needed for monosaccharide transfer and for acylated repeating units. The genes responsible for xanthan acylation were further corroborated by Marzocca et al. (46). In addition, a recent analysis of the transcriptional organization of the gene cluster indicated that the gum region was mainly expressed as an operon from a promoter located upstream of the first gene, *gumB* (47). Another important contribution to the genetics of *Xanthomonas* is the identification of *xanA* and *xanB* genes (48). These genes were reported to be involved in UDP-glucose and GDP-mannose biosynthesis. Recently, it was also demonstrated that the gene encoding

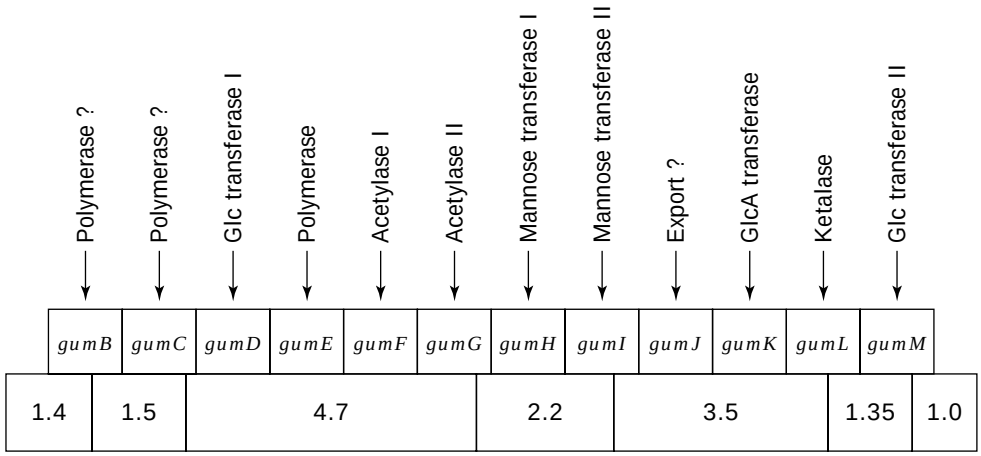


Figure 4. Genetic map of the genome segment of *X. campestris* involved in xanthan production. The *Bam*HI restriction map (bottom) indicates the order and the approximate size in kilobases (kb) of the fragments that comprise the xanthan gene cluster. This cluster has a total DNA length of 16 kb. The genetic map (center) indicates the 12 separate xanthan genes, designated *gumB* to *gumM*. The biochemical functions (top) indicate the enzymatic activity identified for each gene; (?) refers to possible involvement with polymerization or export as indicated; Glc, glucose, GlcA, glucoronic acid. *Source*: Adapted from Vanderslice et al. (45) and Sutherland (38).

UDP-glucose pyrophosphorylase is essential for xanthan synthesis (49). In fact, over 20 relevant genes seem to be identified by genetic analysis and DNA sequencing (50).

BIOPROCESS TECHNOLOGY

Raw Materials

The most efficient xanthan gum producer is *X. campestris*, however other species such as *Xanthomonas phaseolis, Xanthomonas malvacearum, Xanthomonas carotae, Xanthomonas manihotis, and Xanthomonas jugandlis* are known to be able to produce the exopolysaccharide as well. The nutritional versatility of *X. campestris* allows the use of industrial, complex, and synthetic media formulations. In industrial processes, commercial-grade glucose or hydrolized starch is used as the carbon substrate. Sucrose and acid whey are also used. Glucose concentrations up to 5% have been found to give the best product yield; at higher glucose concentrations, xanthan yields decrease. In fact, the maximum level of glucose is limited by the rheology of the fermentation broth. Nitrogen sources, such as corn steep liquor, casein hydrolysates, distiller's dried solubles, yeast, and soybean extracts, can be used, but because of their relatively undefined and variable composition they have been replaced by ammonium salts. Xanthan can indeed be produced in good yield when *X. campestris* is grown in a simple synthetic medium composed of ammonium chloride, glucose, and salts (51). The benefits of using synthetic media are twofold: control of genetic instabilities and improvement of product quality (52). Although the structure of xanthan synthesis is essentially independent of media formulation, there are some differences in acyl substituents and the rheological properties of the polysaccharide, as well as in the product yields (21).

Basic Structure of the Production Process

The design of a xanthan production process depends on the economic requirements to achieve optimum performance with regard to productivity, product concentration, and yield and, at the same time, to fulfill the desired product quality and application specifications. However, this is not a straightforward task because, during the production process, viscosity increases drastically. Therefore, product concentration are usually low and product quality also varies. A compromise between the commercial targets and the limitations inherent to the process must be reached at the industrial level by establishing the relationships between the bacteria's performance, environmental conditions, and equipment design and operation.

Xanthan is produced commercially in a conventional batch process using mechanically agitated vessels (Fig. 5) and a culture of *X. campestris* in a suitable fermentation medium. It is an aerobic process that runs for about 100 h under the following operating conditions: 28–30 °C, pH ~ 7, aeration rate higher than 0.3 vvm, and specific power input for agitation higher than 1 kW/m^3. A product yield approaching 50% is typical in the standard process. The inoculum preparation includes several stages requiring different set of reactors ranging from 10 L for the initial seed culture up to 100 m^3 for production purposes, the need volume being usually enlarged by a factor of 10.

Typical profiles of batch culture of *X. campestris* using synthetic media similar to that used by Davison (53) and the operating conditions mentioned above are shown in Figure 6. The data are plotted as the time courses for cell dried mass (CDM), xanthan, glucose, the rheological parameters k (consistency factor) and n (flow index), the oxygen transfer rate (OTR), and the oxygen partial pressure (pO_2). As the fermentation evolves, cells grow exponentially, resulting in a rapid consumption of the nitrogen

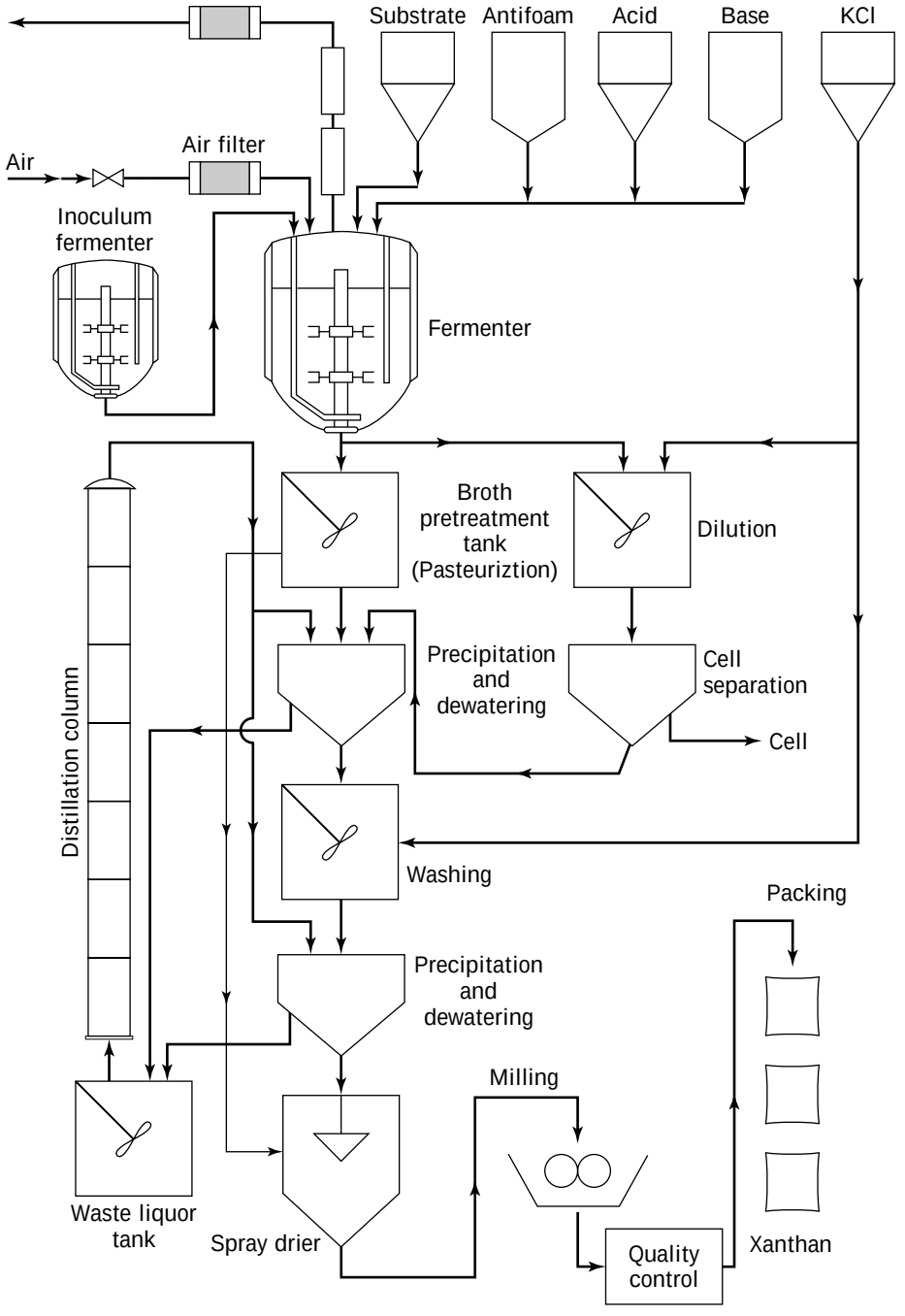


Figure 5. Flowsheet for production of xanthan.

source (not shown). Under N-limitation, cells enter the stationary phase. Xanthan is synthetized in both the growth and the stationary phases, reaching a final concentration of about 20 g L^{-1}. Xanthan biosynthesis ceases as glucose is completely consumed. The consistency factor (k) as a parameter of the rheological properties of xanthan increases from ~15 mPa sn to 30,000 mPa sn. The pO_2 dropped over the initial periods to a minimum value corresponding in time to peak oxygen uptakes rates during the growth phase. Thereafter, both OTR and pO_2 decrease continuously toward the end of the fermentation, reflecting the effect of increasing broth viscosity on the oxygen transfer rate.

After the fermentation stage, xanthan is recovered through a multistep downstream process (Fig. 5). The characteristics of this process are defined by the end product application and the difficulty of removing bacterial cells from viscous culture broths. As pointed out by Smith and Pace (54), the objectives of the recovering process are (*1*) concentration of the fermentation broth to a solid form, which has to be microbiologically stable and easy to handle, transport, and store; (*2*) purification to reduce the degree of nonpolymer solids, such as cells and/or salts, in order to improve the functional performance (odor, color, and taste) of the product; (*3*) deactivation of undesirable contaminating enzymes, such as pectinases and cellulases;

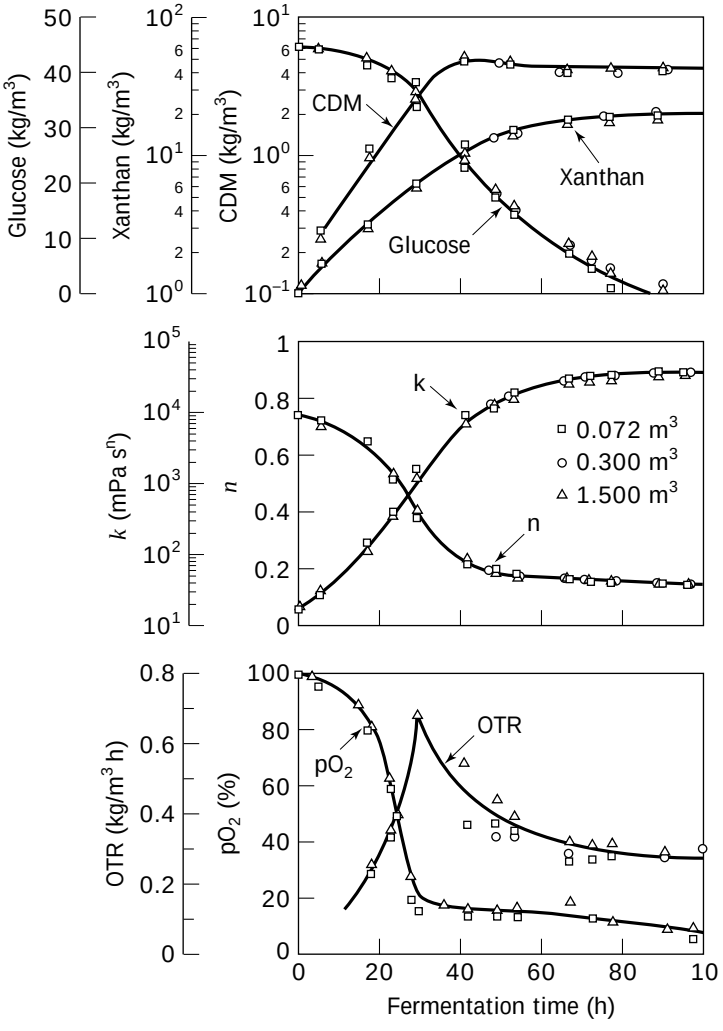


Figure 6. Time profiles of xanthan fermentation in reactors of 72 L, 0.3, and 1.5 m^3 applying the same power input (3.3 kW/m^3) and gas flow rate (7.2 mm s^{-1}). *Source*: Adapted from Herbst et al. (23).

and (*4*) modification of the chemical properties of the product to alter either the application performance or the dispersion and solution rate characteristics of the polymer.

When industrial grade Xanthan is required, the postfermentation process starts with the pasteurization of the fermentation broth to kill the bacterial cells (Fig. 5). Because *Xanthomonas* are sensitive to temperature and are nonspore formers, complete pasteurization can be readily accomplished. However, caution has to be taken with the temperature of the transition midpoint (T_m) for xanthan in order to avoid thermal degradation of the polymer. The T_m for xanthan is about 100 °C but varies with the total ionic strength and the ratio of the contents of pyruvate and acetate groups (55–57). For xanthan applications where maximum cell removal is required (such as in the food, pharmaceutical, and cosmetics industries), cell separation by centrifugation is facilitated by dilution of the fermentation broth to reduce viscosity. Alternatively, the bacterial cells can also be subjected to chemical or enzymatic digestion (58,59); however, the resultant product may contain cell fragments, and solutions are optically not clear. Such a product may be restricted to applications in the petroleum industry.

After the prerecovery treatment just mentioned, the isolation of xanthan is accomplished by lowering its solubility in order to obtain either a precipitate or a polymer-concentrated phase by either the addition of lower alcohols, or salt or acid, respectively. The more favored method is the addition of ethanol or isopropanol. The added alcohol is subsequently recovered from the spent liquor by distillation (Fig. 5). The addition of salts improves the precipitation by creating a charge reversal effect in which all the available anionic groups are bound with cations of the added salt. The addition of KCl to assist precipitation reduces, for example, the amount of isopropanol required by about 30% (60). Consequently, increasing both the product concentration and the ionic strength prior to precipitation will decrease the amount of alcohol required per mass of product. This has a positive inpact on the economics of the process because the cost of alcohol recovery and the inevitable losses of alcohol during xanthan precipitation represent an important portion of the production cost.

Once the polymer is obtained in a wet solid form, its final purity (quality) and the economics of the next operation units can be determined by dewatering and washing. The xanthan precipitate may be centrifuged or pressed in

order to remove free liquid, thereby improving the purity of the product and reducing drying costs. At this stage, the additional washing (Fig. 5) may be carried out to improve the quality of the product by removing as much as possible the particulate matter (cell debris, microgels, organics residues, and pigments). To this end the concentrated form of xanthan is redissolved, washed with water/KCl, precipitated, and dewatered again.

The precipitated cake is then dried in batch or continuous driers under vacuum or with forced air, or with inert gas if the solids contain inflammable solvents. To avoid undesirable effects, such as polymer deterioration, discoloration, and impairment of solubility, the drying temperature must be carefully controlled.

The dried xanthan is then milled to the desired mesh size to control dispersability and dissolution rate as well as to make solids handling easier. The mesh size can range from 40 to 200; the most common mesh sizes are 80 (177 μm) and 200 (74 μm). A range of products with these particle sizes is marketed.

Because an important part of the production costs represents the recovery and purification of xanthan, preparations that have been concentrated directly from the fermentation broth are also commercially available at lower prices. These xanthan concentrates, prepared by tangential flow filtration, contain between 8 and 12% dry polysaccharide to which antibacterial agents are added to prevent deterioration. This form of xanthan, which is specially useful in drilling operations in the oil industry, is available in different industrial grades and can be readily diluted to the required concentrations.

Reactor Designs and Operation Modes

During xanthan fermentation the hydrodynamic behavior of the broth changes substantially as the polymer is produced. The rheology exhibited by the xanthan fermentation culture is complex and includes pseudoplastic flow behavior, yield stress, and viscoelasticity. The effective viscosity increases by about three powers of 10. Under such conditions adequate aeration and mixing have to be provided if high specific xanthan productivity is to be maintained and stagnant zones and hence oxygen limitation avoided. Fermenter and impeller design, therefore, plays a critical role in ensuring good mass transfer and homogeneity and, hence, good product quality. Because the turbulence and shear rate in stirred tank-reactors (STRs) is higher near the impeller than in bulk mixing, optimal design requires the provision of high turbulence to give good gas dispersion and bulk mixing. The conventional Rushton turbine (the most widely used impeller in the fermentation industry) has shown to be disadvantageous because of its high unaerated power number (61) and compartimentalization effects (62). At high xanthan concentration the sorption efficiency (mass of oxygen absorbed at maximum driving force per energy supplied) of an STR with a Rushton turbine drops significantly below that of other reactors (Fig. 7). Consequently, alternative impeller types such as the Intermig, Prochem Maxflo T, and Scaba 6SRGT have been tested (23,64). Among these, the Intermig has shown a four times higher sorption efficiency at high xanthan concentrations, compared with Rushton turbines (23).

Reactor design has also been targeted as a means of improving xanthan production and quality. The plunging jet reactor (65), airlift reactor (66), centrifugal fibrous-bed bioreactor (67), STR with draft tube (68), and bubble column (66), have all been tested for that purpose. In spite of the lower energy requirements and higher oxygen sorption efficiency, the poor mixing in the wall region of the plunging jet reactor is a problem that needs to be solved. The limitation of oxygen supply in the downcomer of airlift reactors has negative effects on the xanthan production rate and product yield (66). Although the design is novel, the recently proposed centrifugal fibrous-bed bioreactor (67) does not show obvious advantages. An STR with a draft tube shows good performance as evaluated by its sorption efficiency (Fig. 7) (68). At low xanthan concentration the bubble column has significantly higher efficiency than the other reactors (Fig. 7). Even at high xanthan concentrations the bubble column exhibited excellent mixing performance; slug flow prevails and inhomogeneities (stagnant zones) are avoided (34,66).

The possibility of running a xanthan bioprocess continuously has been also examined at laboratory scale (69–71). The main features of this mode of operation are (*1*) higher specific productivity, (*2*) no definite limit (in theory) to the length of run because bacterial degeneration could be avoided using synthetic media with reduced levels of sulfur, (*3*) avoidance of time losses in restarting batch runs, and (*4*) near maximum process productivity and a constant product quality, owing to steady-state conditions. One of the most interesting applications of continuous mode operation may be found in the oil industry, where in situ production plants may provide xanthan in different industrial grades that can readily be diluted to the required solution concentration. However, a serious weakness of the process is the low product accumulation, which is around 5–7 g/L at the conditions where maximum productivity and high product quality are favored (dilution rate $\sim$0.075 h^{-1}). This constraint and limitations in avoiding culture contamination are important reasons why industry has not yet adopted this cultivation method.

Unfortunately, fed-batch strategies with glucose feeding (70,72–77) have not brought significant improvement to the process; some resulted in a performance that was even worse than that of batch culture. A common characteristic of these processes is that an important portion of the glucose supplied (up to 60%) remains in the reactor without being utilized. These limitations in fed-batch culture make this process strategy not economically attractive.

Recently, a new fed-batch operation mode was proposed, based on the observation that the N-source inhibits growth of *X. campestris* (21). Therefore, the novel feature of this process is the use of ammonia feeding to control growth. This allows major product accumulation with higher molecular weight during the growth phase (21). The benefits to be expected are as follows: (*1*) larger product yield, (*2*) higher volumetric productivity, (*3*) improvement of product quality and uniformity, (*4*) reduction in capital investment and unit production costs, and (*5*) increase in plant production capacity as compared with the standard batch process (Table 2). The application of this process at the production level may require combined strategies in-

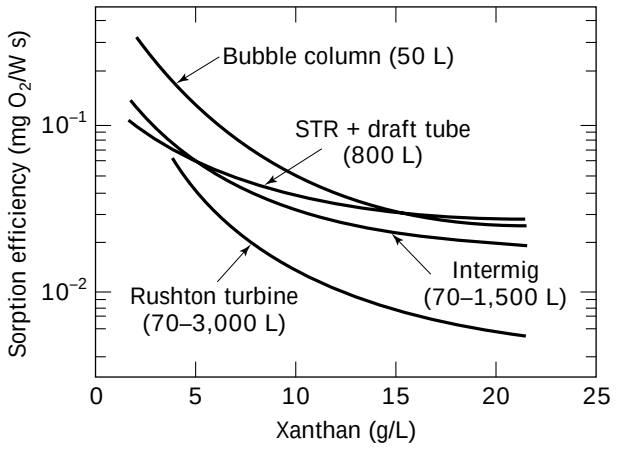


Figure 7. Comparison of the energetic efficiencies of oxygen transfer during xanthan production in various bioreactors and different impeller design. *Source*: From Herbst (63).

Table 2. Comparison of Performance and Costs of Xanthan Cultivation in Batch and Fed-Batch Operation Mode

Parameter	Standard batch	Fed-batch
X_O (g L^{-1})	0.2	0.5
X_f (g L^{-1})	4	5.2
$Product_{(harvest)}$ (g L^{-1})	25.38	30.8
Run time (h)	84	49
Productivity (g L^{-1} h^{-1})	0.3	0.63
$Y_{P/S}$	0.507	0.65
Capital investment cost ($ $\times$ 10^3)[a]	5,493	2,930
Production cost per kg ($)[a]	7.28	5.38

Cost analysis has been made taking a production plant capacity of 1,500 tons/year (21).

cluding the use of the Intermig impeller in STRs with a multiple feeding device, or the use of a bubble-column reactor.

Oxygen Transfer and Scale-Up

Transport processes are dependent on scale because the time constants for transport processes increase with scale. Scale-up problems are pronounced in xanthan fermentation because of the highly viscous pseudoplastic flow behavior of the culture broth, which deteriorates the mass transfer, mixing, and heat transfer capabilities of the fermenter. The actual oxygen concentration and the kinetic behavior of the bacteria are dependent on scale owing to the fact that oxygen and other nutrients involved in the bacterial metabolic activity are consumed and have to be supplied by transport processes. Sufficient oxygen supply to the obligately aerobic bacterium is of particular importance. If dissolved oxygen becomes depleted in the culture medium, the specific xanthan productivity of the cells decreases linearly with the specific oxygen uptake rate (78) (Fig. 8). Furthermore, and as shown in Figure 9, a lower molar mass product is formed under oxygen limitation (79). The information in Figures 8 and 9 shows that, among other scale-up criteria, scale-up at constant specific oxygen transfer rate is the most appropriate choice to achieve constant productivity and high product quality. Maintaining

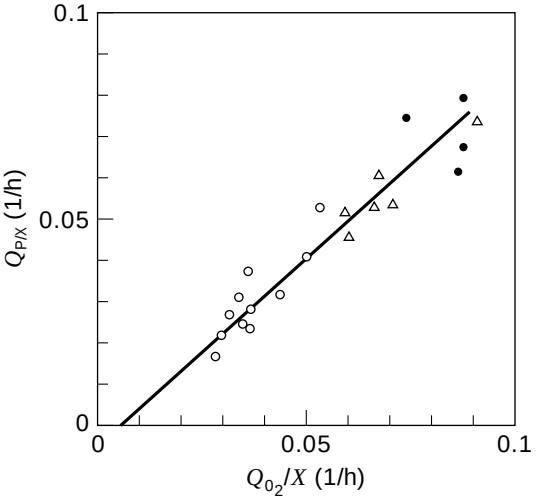


Figure 8. Dependence of the specific xanthan production rate on the specific oxygen uptake rate. Circles refer to runs in an STR at 15 L and 1,500 L, triangle correspond to bubble-column reactor; filled symbols denote no oxygen limitation.

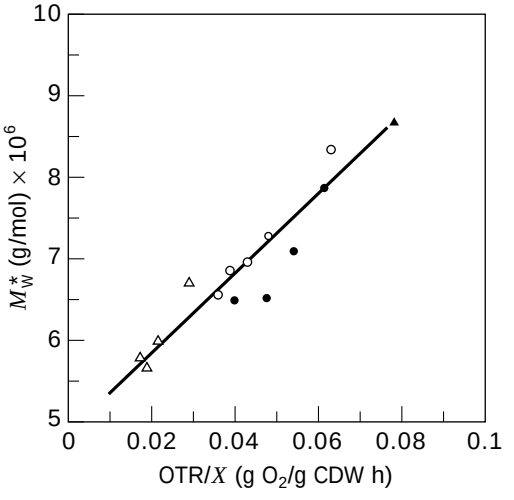


Figure 9. Influence of the specific oxygen transfer rate on the weight-mean molar mass of the xanthan produced.

a constant volumetric mass transfer coefficient (k_La) in process scale-up is approximately the appropriate condition to guarantee the same oxygen supply rate to satisfy the oxygen demand of the cell population.

Aiming at high product concentration for economic downstream processing, oxygen transfer limitation is likely to occur at the final stage of the fermentation. This applies particularly to industrial fermentations where the specific power input is bound to be low. Any error at the design stage when estimating the oxygen transfer capacity of the fermenter results in a proportional error in the maximum xanthan production rate. Consequently, the volumetric mass transfer coefficient, k_La, must be accurately estimated.

Various approaches toward the estimation of k_La in stirred-tank reactors with viscous pseudoplastic liquid have been proposed in the literature. The flow behavior is commonly described according to the Ostwald-de Waele relationship (equation 1). Because the shear rate in the STR is locally nonuniform, the effective mean shear rate, $\dot{\gamma}_{eff}$, caused by the impeller is described, for example, in terms used by Metzner & Otto (80):

$$\dot{\gamma}_{eff} = AN \tag{5}$$

where A is a constant characteristic of the impeller and N is the sirrer speed. Various expressions for $\dot{\gamma}_{eff}$ in the STR, employed to calculate the effective viscosity, η_{eff}, from equation 2, are listed in Table 3. Equation 6 of Table 3 allows one to predict shear rates at turbine impeller blades, whereas equation 7 applies in the "stirrer zone." Equation 8, is a modification of the Metzner–Otto relationship including a function of the flow behavior index (n) and differs only marginally from the original one (23). Herbst et al. (23) showed that the predicted $\dot{\gamma}_{eff}$, using Calderbank and Moo-young's relationship, is almost two orders of magnitude lower than the shear rate at the stirrer tip proposed by Henzler & Kauling (84). The last predicts a different trend because increasing viscosity is assumed to lower the shear rate. It is obvious that any correlation of the volumetric mass transfer coefficient will be affected by the choice of relationship for the effective shear rate $\dot{\gamma}_{eff}$.

A number of mass transfer correlations exist in the literature, some of them are given in Table 4. The relationships in Table 4 may be divided into those that reflect the effect of the stirring speed directly (equations 10, 11, and 13) and those that take into account the resultant power input (equation 12). The predictive capability of these equations, however, differs significantly (Fig. 10). Based on the results shown in Figure 10, it can be concluded that, the dimensional relationship of Kawase and Moo-Young (87), equation 12, gives the best representation of the experimental data. Herbst et al. (23), working with reactor volumes ranging from 0.015 to 1.5 m^3, demonstrated further that the power input required to achieve a desired mass transfer rate calculated from equation 12 was reduced considerably when Intermig impellers are applied. Figure 11 shows that the experimental data are well described by the relationship of Kawase and Moo-Young (equation 12) when the power input (P/V) was modified by a factor of three (23).

The scale-up criteria of a constant specific oxygen transfer rate was experimentally confirmed to be valid in xanthan bioprocesses by Herbst et al. (23). In order to keep the volume-specific oxygen transfer rate constant, equal power input ($\sim$3.3 kW/m^3) and superficial gas velocity (7.2 mm s^{-1}) in reactors ranging from 0.072 to 1.5 m^3 were employed. Under these operating conditions all the main fermentation performance parameters show an excellent reproducibility independent of the scale. This can be recognized in the results in Figure 6 where all concentrations, the rheological properties (K, n), and the oxygen transfer rate and the dissolved oxygen concentration are reproducible. Moreover, the product quality as evaluated by its molecular weight, the polydispersity index, and the pyruvat content were constant irrespective of the fermenter scale. These results highlight that constant specific oxygen transfer rate is the suitable scale-up criterion for constant xanthan productivity and product quality.

Effects of the Process Design on Product Quality

The high viscosity yield of xanthan at low shear rates and its shear thinning behavior are particularly important during and after processing of many industrial products. Thus, a highly pseudoplastic xanthan is required. This means that, according to the rheology (equation 2) for a determined shear rate, $\dot{\gamma}$, high values of the rheological parameters k and n are desired to achieve a maximum effective viscosity (η_{eff}). However, in real terms, there is no way to vary k and n independently because both appear to be interdependent (see equation 4). The solution of equation 4 is given by equation 14, which exhibits an inverse functional relationship between the flow index, n, and both the xanthan molar mass, M, and the consistency factor, k.

Table 3. Selection of Relationships for the Effective Shear Rate in Stirred Tanks

Reference	Approach	Correlation	Eq.
Wichterle et al. (81)	Semiempirical	$\dot{\gamma} = N(1 + 5.3n)^{1/n}(N^{2-n}d^2\rho_L/k)^{1/(1+n)}$	(6)
Obernosterer and Henzler (82)	Semiempirical	$\dot{\gamma} = \left[0.038 \frac{P}{V_L} \frac{1}{k} \left(\frac{D}{d}\right)^3 \frac{d}{b}\right]^{1/(1+n)}$	(7)
Calderbank and Moo-Young (83)	Semiempirical	$\dot{\gamma}_{eff} = AN\left(\frac{4n}{3n + 1}\right)^{1/(1-n)}$	(8)
Henzler and Kauling (84)	Dimensional analysis	$\dot{\gamma}_{eff} = \left(\frac{P/V}{\mu_{eff}}\right)^{0.5}$	(9)

Table 4. Selection of Correlations for the Volumetric Mass Transfer Coefficient ($K_L a$) in Stirrer Tanks with Viscous (Non-Newtonian) Media

Stirrer	Flow behavior	Shear rate	Correlation	Eq.
Rushton turbine $d/D = 0.4$	$k \leq 13.2$ Pa s^n $n \geq 0.42$	(Eq. 5)	$\frac{k_L a d^2}{D_L} = 0.06\left(\frac{d^2 N \rho_L}{\mu_{eff}}\right)^{1.5}\left(\frac{\mu_{eff}}{\rho_L D_L}\right)^{0.5}\left(\frac{\mu_{eff} u_G}{\sigma}\right)^{0.6}\left(\frac{dN^2}{g}\right)^{0.19}\left(\frac{Nd}{u_G}\right)^{0.32} \cdot [1 + 2(\lambda N)^{0.5}]^{-0.67}$	(10)
Rushton turbine $d/D = 0.27$–0.47	$k \leq 0.005$ Pa s $n = 1$	10^4 s^{-1}	$\frac{k_L a d^2}{D_L} = 2.5 \times 10^{-4}\left(\frac{d^2 N \rho_L}{\mu_{eff}}\right)^{1.8}\left(\frac{\mu_{eff}}{\rho_L D_L}\right)^{0.5}\left(\frac{\mu_{eff}}{\mu_w}\right)^{1.39} \cdot [1 + 100(\lambda N)]^{-0.67}$	(11)
Semitheoretical approach		(Eq. 8)	$k_L a = 0.675 D_L^{0.5} \frac{\rho_L^{3/5}(P/\rho_L V)^{(9+4n)/(10+n)}}{(k/\rho_L)^{(1+n)/2\sigma^{3/5}}}\left(\frac{u_G}{u_b}\right)^{0.5}\left(\frac{\mu_{eff}}{\mu_w}\right)^{-0.25}$ with $u_b = 0.265$ m/s	(12)
6-Blade paddle $d/D = 0.5$	$k \leq 33.2$ Pa s^n $n \geq 0.68$	(Eq. 5)	$\frac{k_L a}{N} = 0.0786\left(\frac{d^2 N \rho_L}{\mu_{eff}}\right)^{0.41}\left(\frac{\rho_L N^2 d^3}{\sigma}\right)^{0.13}\left(\frac{Nd}{u_G}\right)^{0.69}$	(13)

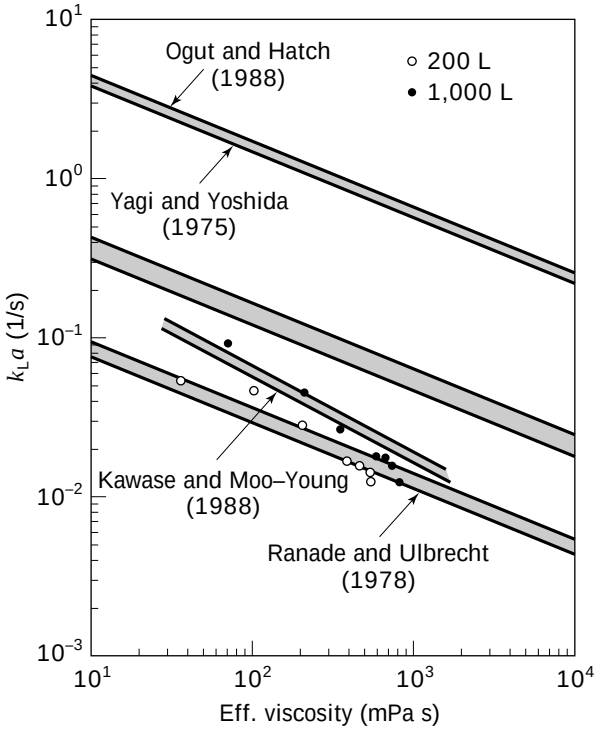


Figure 10. Comparison of experimental data with estimated values of the volumetric mass transfer coefficients using equations 10–13 of Table 4.

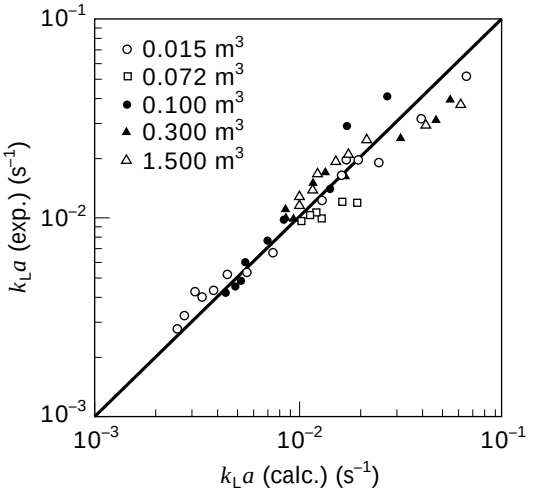


Figure 11. Parity plot for the relationships of Kawase and Moo-Young (82) (equation 12), with experimental data from Herbst et al. (23). Power input in equation 12 was modified by a factor of 3 for Intermig impeller.

$$n = \frac{1 + 0.071 \cdot B}{1 + B}, \text{ with}$$
$$B = 0.0227\left(\frac{M}{10^6}\right)^{2.08}(k - 10^{-3})^{0.47} \qquad (14)$$

As observed in the results in Figure 6, k rises by three orders of magnitude toward the end of the fermentation, while n declines by less than one order of magnitude. This indicates that the consistency factor, k, is a predominant factor in determining the rheological properties of xanthan. The functional dependence of the parameter k on xanthan molar mass and product concentration is well established (see equation 3). Therefore, to achieve high values of k for a given product concentration, conditions of high xanthan molar mass have to be achieved. To this end the empirical correlation of Peters et al. (24) may be considered (equation 15):

$$M = 4.9 \cdot 10^6 + 3.8 \cdot 10^6 \cdot \mathrm{OTR}/Q^*_{O_2} + 2.6 \cdot 10^7 \mu \qquad (15)$$

Equation 15 implies that conditions of no oxygen limitation and high growth rates have to be ensured when xanthan of high viscosity yield to be produced. This rationale was proven to be right by the scale-up results of Herbst et al. (23) mentioned earlier. Furthermore, the media formulation has shown to affect indirectly the mean molar mass of xanthan (21,66). On one hand, when complex media are

used, oxygen limitation is encountered at an early stage of the fermentation run, and hence, the mean molar mass of xanthan is impaired (66). On the other hand, high levels of the N-source in synthetic media formulations inhibit the bacterial growth, with the consequent result of low molar mass xanthan synthesis (21). It was further found that the N-source level also affects the pyruvate content of xanthan (21). The same author demonstrated further that the rheological properties of xanthan improve as the degree of pyruvate substitution increases (Fig. 12). This is in accordance with the initial observations of Sandford (86) and Cadmus et al. (52). Thus, the pyruvate content of the polymer should be regarded as an important parameter in order to design, predict, and control the rheological behavior of the fermentation broth, and as regards the product application.

COMMERCIAL APPLICATIONS

5.1 Food uses

Several applications that illustrate xanthan use in the food industry are given in Table 5. First of all xanthan gum is incorporated into foods to control the rheology of the final product. In performing this functional task, the polymer has a major effect on various sensory properties such as texture, flavor release, and appearance, which contribute to the acceptability of the product for consumer use. Because of its solution properties, xanthan can be used to provide a desired functionality at a lower concentration than most gums and, therefore, may be more cost-effective. Additionally, because of its pseudoplastic properties in solution, xanthan has a less "gummy" mouth-feel than gums that exhibit Newtonian behavior. An additional advantage in food formulations is the antioxidative action of xanthan as compared with other polysaccharides and additives (93).

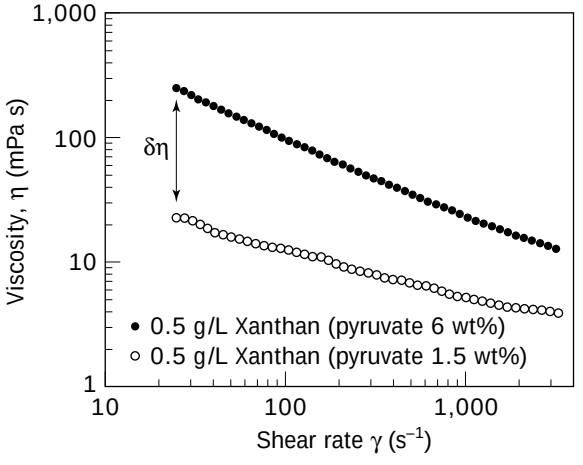


Figure 12. Comparative viscosity curves of 0.5 g/L xanthan in 0.12 M KCl with varying degree of pyruvate substitution. Filled symbols correspond to xanthan with about 6 wt% pyruvate content and M_w of 7×10^6 g/mol; open symbols relate to xanthan with about 1.5 wt% pyruvilation degree and M_w of 6.5×10^6 g/mol (run F7). *Source*: Results from Flores-Candia (21).

Table 5. Main Practical Uses of Xanthan for the Food Industry

Function	Application
Adhesive	Icings and glazes
Binding agent	Pet foods
Coating	Confectionery
Emulsifying agent	Salad dressing
Encapsulation	Powdered flavors
Film formation	Protective coatings, sausage casings
Foam stabilizer	Beer
Gluten substitution and dough processing	Baked goods, pasta
Stabilizer	Ice cream, salad dressings, juice drinks, margarine
Swelling agent	Processed meat products
Syneresis inhibitor, freeze–thaw stability	Cheese, frozen foods
Thickening agent	Jams, sauces, syrups, and pie fillings
Pumping and filling improvement	Processing of canned products

Source: From Refs. 87–92.

Xanthan is also being used in the formulation of reduced-calorie products, representing the second and third generation of fat replacers and stabilizers (87).

Cosmetics and Pharmaceutical Uses

Xanthan gum is used in cosmetics and toiletries as a bodying agent and emulsifier. Personal care products such as shampoos, creams, lotions, make-up, hair care products, and toothpastes can be formulated with xanthan (88,94,95). Additionally, xanthan provides creams and lotions with a good "skin-feel" during and subsequent to application (88). Recently, a new xanthan formulation (Ticaxan) was introduced into the market for that purpose (96).

In the pharmaceutical industry, xanthan is used to suspend pharmaceuticals such as antibiotics and to provide formulations of uniform dosage (88). It can also be used to stabilize emulsified cream formulations containing pharmaceuticals. Presently, xanthan is being evaluated in controlled-release applications, such as in gastrointestinal pharmaceutical applications, as rate limiting membranes for sustained delivery of oral formulations (97,98) and as dissolution and binding agents in tablets and rectal preparations (99,100).

Agricultural and Other Industrial Applications

Xanthan is used in a diverse range of industrial applications, particularly as a suspending or thickening agent. The polymer improves the flowability of fungicides, herbicides, and insecticides by uniformly suspending the solid components of the formulation in an aqueous system or by stabilizing emulsions and multiphase liquid systems (88). The unique rheological properties of xanthan solutions also improve sprayability, reduce drift, and increase pesticide cling and permanence (89,101,102). Recently, various "tolerance exemptions" were issued by the U.S. Envi-

ronmental Protection Agency for the use of xanthan as a surfactant in pesticide formulations (103). Because of its ability to disperse and hydrate rapidly, xanthan is used in jet injection printing. It is nonpolluting and gives a good color yield. It is also used in several mining processes for removing dissolved metals in a highly efficient process (104). Lately, in the formulation of a new generation of thermoset coatings, xanthan was introduced by Monsanto with the aim of addressing customers' environmental challenges (105).

Petroleum Industry

In the petroleum industry, xanthan gum finds a wide range of applications. Among these are oil drilling, fracturing, pipeline cleaning, workover, and completion. Xanthan's unique solution rheology, its excellent compatibility with salts, and its resistance to thermal degradation have made it useful as an additive to drilling fluids (88,106). The pseudoplasticity of its solutions provides low viscosity at the drill bit, where shear is high, and high viscosity in the annulus, where shear is low. Therefore, xanthan serves a dual purpose in allowing faster penetration at the bit because of the low solution viscosity in that area and, at the same time, suspending cuttings in the annulus because of the high solution viscosity in that area. Hydraulic fracturing is used to effect deep-penetrating reservoir fractures that improve productivity of the well. Xanthan's high solution viscosity at low shear is important for proppant suspension in this operation (107). The pseudoplastic rheology, shear stability, and friction reduction impaired by xanthan provide for maximum delivery of pressure to the formation and minimal friction loss in the tubing and reservoir. A workover and completion fluid is essential for achieving and maintaining maximum well productivity. Xanthan contributes to excellent hole cleaning, suspension, friction reduction, and stability to shear and high temperatures. An important speciality application employs a xanthan gel to remove rust, welding rods, wet slag, and other debris from gas pipelines. This pipeline cleaning method is considered safe, environmentally acceptable for open sea disposal, and cost-effective (88).

Enhanced Oil Recovery

For every barrel of oil actually produced, about another two remain in the ground (108). Therefore, enhanced oil recovery (EOR) may become important in the next decades to meet the increasing energy demand of the future. Xanthan gum can be used in two types of enhanced recovery systems. The principal use is for an improved form of water flooding in secondary oil recovery, a phase in which polymers are used to increase the efficiency of water contact and to displace reservoir oil. The second application is in micellar-polymer flooding as a tertiary oil recovery operation. Polymer-thickened brine is used to drive a slug of surfactant through the porous reservoir rock to mobilize residual oil; the polymer prevents fingering of the drive water through the surfactant ban and ensures good area sweep. In both applications, the function of the polymer is to reduce the mobility of injected water by increasing its viscosity. Evaluation of hundreds of biopolymers has found xanthan to be one of the best candidates for use in EOR operations. The reasons for this are its (*1*) non-Newtonian behavior, (*2*) high viscosity yield even at low concentrations (600–2,000 ppm), (*3*) low sensitivity of viscosity to salinity changes (unlike the polyacrylamides), (*4*) resistance to mechanical degradation, (*5*) stability with respect to higher temperature (up to almost 90 °C).

One of the major limiting factors in the application of xanthan in EOR is that, by the nature of its production method, it is expensive, and the current general market price is in excess of $11/kg (109). In addition, some drawbacks during field operation are associated with poor injectability behavior, adsorption into the rock strata, and plugging problems (110,111). These problems have been attributed to the quality and composition of the polymer (110,112), to water quality (111), and to well configuration (113,114).

MARKET SIZE AND FUTURE DEVELOPMENT

Xanthan gum currently dominates the microbial gum market, which is growing at between 6 and 7% per year (115). It has been suggested that global production exceeds 50,000 tons per year (116). The worldwide xanthan market is valued between $600 and $800 million per year (109,116,117). The European Union market accounted for about $158 million in 1994, and the U.S. market was reported to be about $120 million in 1995. Table 6 shows the U.S. market size and forecast of xanthan consumption by major end uses. Food accounts for about 70% of total consumption; the remainder is spread among petroleum production, cosmetics, and other uses. Overall growth is forecasted at about 7.0% yearly, with food applications providing most of the impetus at 8.3% annually (1). Consumption of xanthan in petroleum is forecasted at a very modest rate (2%), which is due mainly to recovery operations rather than oil well drilling (2). Because of its unique properties, the application of xanthan in cosmetics has been growing successfully at 5.9% per year (82). It should be noted that use in textile production is very nominal because the textile industry in the United States is conservative and only growing with the national economy (3).

An analysis of Table 6 shows that xanthan is positioned well in U.S. markets. Further analysis reveals that xan-

Table 6. Market Analysis and Forecast: Xanthan Consumption in The United States

End use	1993 ($ M)	1995 ($ M)	1998 ($ M)	AAGR (%) 1993–1998
Food	74.4	89.5	111.0	8.3
Petroleum	16.5	17.1	18.2	2.0
Cosmetics and toiletries	9.0	10.2	12.0	5.9
Textiles	1.4	1.4	1.4	0
Other	1.1	1.1	1.7	8.5
Total	102.4	119.1	144.3	7.1

Source: Business Communications Company (1–4,94).
Note: Major xanthan demand according to end use is listed in millions of dollars (M); AAGR, annual average growth rate.

than gum represents the fastest growing segment of the polysaccharide industry (1–4). Most growth is in foods, where demand for natural gums is falling as a result of market share erosion by xanthan (118).

The high growth rate of the xanthan market may explain not only why the major producers have all expanded in recent years (119) but also why Archer Daniels Midland Co. has decided to enter the xanthan market with an estimated 10,000 tons of annual capacity (119,120). Between 1992 and 1994, the global xanthan production capacity doubled in size (119,121). Recently, Nutrasweet Kelco Co., the world's leading producer of xanthan, has announced further expansion (122). In fact, an overcapacity has been created in the xanthan market, which may have two important consequences: (*1*) an accelerated penetration of xanthan into new markets, and (*2*) stiff competition for new polysaccharides. Success may be limited only to those polymers that have a clear advantage in performance or customer perception. Less-expensive polymers that do not perform better may be disadvantaged due to the high costs of incorporating new gum into the end product. An estimate from 1982 suggests that the testing required by the U.S. Food and Drug Administration to gain food clearance approval for a new additive can cost about $2 million and may take up to 8 years (123).

According to a market forecast, the two main factors shaping the xanthan market in the next century are growing xanthan demand in developing countries and worldwide energy demand, which will stimulate EOR operations as reserves decline (124,125). Indeed, the large strategic value of xanthan still remains in EOR operations, whose worldwide production at the beginning of 1996 was estimated to be around 2.2 million barrels/day (126).

The large economic value of xanthan in the EOR market can be assessed, as suggested by Moses (127), by looking at the amount of oil that remains in the earth's crust (Table 7). Thus, it becomes clear that the EOR markets are very large indeed. They are, however, low-price markets that could be developed only through the use of much improved technologies with low production costs. To improve the competitive advantage of xanthan in EOR, more efficient fermentation processes and better control of the polymer quality are required.

Table 7. Approximate Value of Oil Not Economically Recoverable by Present Technology

Source region	Estimated value of oil unrecoverable by present technologies ($ × 10^{12})
Asia-Pacific	0.6
Western Europe	0.8
Middle East	12.4
Africa	1.9
Western Hemisphere	4.1
Former East-Block states	2.9
Total for the whole world	22.7

Source: From Ref. 25.
Note: The estimate was based on the following assumptions: (*1*) a crude oil price of $18.00 per barrel, (*2*) that proven reserves (i.e., reserves regarded as presently recoverable) represent 35% of the original oil-in-place.

Relatively high xanthan costs, some technical drawbacks and the (relatively) low oil price have prevented market entry in recent years. Xanthan usage in EOR is typically 0.4% w/w (about 0.5 kg per barrel of oil), which represents about 25% of the estimated equilibrium oil price ($20 per barrel)—a value that is evidently too high for the oil industry.

A recent survey of EOR shows that polymer flooding is achieving commercial success (126). In this operation, up to 35% of original oil in place has been recovered in the Junggar basin (China). Polymer flooding is being used in several projects in France, Germany, China, India, and Romania. For example, polymer flooding accounts for about 50% of the total EOR operations in China. Even when it is not exactly clear what percentage xanthan is responsible for, the fact that the largest xanthan gum production plant has been brought onstream for oil applications is an indication that xanthan use in EOR is not far away. Analysts in marketing research are convinced that the use of xanthan as well as other microbial gums in EOR will become more prevalent as increasing worldwide energy demand results in rising oil prices.

NOMENCLATURE

a (m^{-1})	Interfacial area per unit volume of liquid
A	Impeller-related factor
b (m)	Height of stirrer blade
C_P ($kg\ m^{-3}$)	Product concentration
d (m)	Stirrer diameter
D (m)	Reactor diameter
g ($m\ s^{-2}$)	Gravitational acceleration
k (mPa sn)	Consistency index
k_L ($m\ s^{-1}$)	Liquid-side mass transfer coefficient
M ($kg\ kmol^{-1}$)	Weight-mean molecular weight
N (s^{-1})	Stirring speed
n	Flow behavior index
OTR ($g\ L^{-1}\ h^{-1}$)	Oxygen transfer rate
P (W)	Total power input
pO_2 (%)	Partial oxygen pressure
Q_{O_2} ($g\ L^{-1}\ h^{-1}$)	Microbial oxygen demand
u_b ($m\ s^{-1}$)	Bubble rise velocity
u_G ($m\ s^{-1}$)	Volumetric gas flow rate
V (m^3)	Dispersion volume
V_G ($k\ mol^{-1}$)	Mole volume of ideal gas
x	Molar fraction

Greek Symbols

$\gamma\cdot$ (s^{-1})	Shear rate
η (mPa s)	Viscosity
τ (Pa)	Shear stress
ρ ($kg\ m^3$)	Density
λ (s)	Relaxation time
σ ($N\ m^{-1}$)	Surface tension

Subscripts or Superscripts

L	Liquid phase
eff	Effective
w	Water

BIBLIOGRAPHY

1. Business Communications Company, July, 1994, 00039081.
2. Business Communications Company, July, 1994, 00039094.
3. Business Communications Company, July, 1994, 00039098.
4. Business Communications Company, July, 1994, 00039068.
5. P.E. Jansson, L. Keene and B. Lindberg, *Carbohydr. Res.* **45**, 275–282 (1975).
6. L.D. Melton, L. Mindt, D.A. Rees, and G.R. Sanderson, *Carbohydr. Res.* **46**, 245–257 (1976).
7. R.A. Hassler and D.H. Doherty, *Biotechnol. Prog.* **6**, 182–187 (1990).
8. H. McIntyre, H.J. Ceri, and Vogel, *Starch. Starke* **48** (7–8), 285–291 (1996).
9. A.J. Clarke-Sturman, J.B. Pedley, and P.L. Sturla, *Int. J. Biol, Macromol.* **8**, 355–360 (1986).
10. G. Muller, J. Lecourtier, G. Chauveteau, and C. Allain, *Makromol. Chem. Rapid Commun.* **5**, 203–208 (1984).
11. R. Moorhouse, M.D. Walkinshaw, and S. Arnott, *ACS Symp. Ser.* **45**, 90–102 (1976).
12. G. Muller, M. Anrhourrache, J. Lecourtier, and G. Chauveteau, *Int. J. Biol. Macromol.* **8**, 167–172 (1986).
13. G. Holzward and E.B. Prestridge, *Science* **197**, 757–759 (1977).
14. G. Holzward, *Carbohydr. Res.* **66**, 173 (1978).
15. K.P. Shatwell, I.W. Sutherland, and S.B. Ross-Murphy, *Int. J. Biol. Macromol.* **12**, 71–78 (1990).
16. M. Milas, W.F. Reed, and S. Printz, *Int. J. Biol Macromol.* **18**, 211–221 (1996).
17. K.P. Shatwell, I.W. Sutherland, S.B. Ross-Murphy, and I.C.M. Dea, *Carbohydr. Polymers* **14**, 115–130 (1991).
18. K.P. Shatwell, I.W. Sutherland, S.B. Ross-Murphy, and I.C. M. Dea, *Carbohydr. Polymers* **14**, 131–147 (1991).
19. H. Kumagai, C.H. Chu, T. Sakiyama, S. Ikeda, and K. Nakamura, *Biosci. Biotechnol. Biochem.* **60**, 1623–1626 (1996).
20. D.V. Zasypkin, E. Dumay, and J. C. Cheftel, *Food Hydrocolloids.* **10**(2), 203–211 (1996).
21. J.-L. Flores-Candia, VDI Verlag GmbH, Düsseldorf, (1998).
22. U.S. Pat. 4,394,447 (1983), M. Cadmus and C.A. Knutson.
23. H. Herbst, A. Schumpe, and W.-D. Deckwer, *Chem. Eng. Technol.* **15**, 495–498 (1992).
24. H.U. Peters, I.-S. Suh, A. Schumpe, and W.-D. Deckwer, *Can. J. Chem. Eng.* **70**, 742–750 (1992).
25. A. Parker, P.A. Gunning, K. Ng, and M.M. Robins, *Food Hydrocolloids* **9**(4), 333–342 (1995).
26. J.C. Walker, in *Diseases of Vegetable Crops*, McGraw-Hill, New York, 1952, pp. 128–131.
27. J.F. Bradbury, in *Guide to Plant Pathogenic Bacteria*, CAB International Mycological Institute, Slough, 1986, pp. 198–260.
28. J. Swings, L. Vauterin, and K. Kersters, in Xanthomonas, Chapman & Hall, London, 1993, pp. 121–156.
29. M. Cadmus, C.A. Knutson, A.A. Lagoda, J.E. Pittsley, and K.A. Burton, *Biotechnol. Bioeng.* **20**, 1003 (1978).
30. J.F. Kennedy and I.J. Bradshaw, in *Progress in Industrial Microbiology*, vol. 19, Elsevier, Amsterdam, 1984, pp. 319–365.
31. E. Galindo and G. Salcedo, *Enzyme Microb. Technol.* **19**, 145–149 (1996).
32. J.C. Roseiro, M.E. Esgalhado, A.N. Emery, and M.T. Amaral-Collaco, *J. Chem. Tech. Biotechnol.* **65**, 258–264 (1996).
33. T.R. Jarman and G.W. Pace, *Arch. Microbiol.* **137**, 231–235 (1984).
34. A. Pons, C.G. Dussap, and J.B. Gros, *Biotechnol. Bioeng.* **33**, 394–405 (1988).
35. C. Whitfield, I.W. Sutherland, R.E. Cripps, *J. Gen. Microbiol.* **128**, 981–985 (1982).
36. L. Ielpi, R.O. Couso, and M.A. Dankert, *Biochem. Biophys. Res. Commun.* **102**, 1400–1408 (1981).
37. L. Ielpi, R.O. Couso, M.A. Dankert, *Biochem. Int.* **6**, 323–333 (1983).
38. I.W. Sutherland, in Xanthomonas, Chapman & Hall, 1993, pp. 363–388.
39. J.-L. Tang, C.L. Gough, and M.J. Daniels, *Mol. Gen. Genet.* **222**, 157–160 (1990).
40. U.S. Pat. 4,713,449 (1988), R.W. Vanderslice and P. Shannon.
41. M.R. Betlach, M.A. Capage, D.H. Doherty, R.A. Hassler, N.M. Henderson, R.W. Vanderlice, J.D. Marrelli, and M.B. Wood, in M. Yalpani ed., *Industrial Polysaccharides: Genetic Engineering, Structure/Property Relations and Application*. Elsevier, Amsterdam, 1987, pp. 35–50.
42. G.J. Boulnois and K. Jann, *Mol. Microbiol.* **3**, 1819–1823 (1989).
43. N.E. Harding, J.M. Cleary, D.K. Cabanas, I.G. Rosen, and K.S. Kang, *J. Bacteriol.* **169**, 2854–2861 (1987).
44. L. Thorne, L. Tansey, and T.J. Pollock, *J. Bacteriol.* **169**, 3593–3600 (1987).
45. R.W. Vanderslice, D.H. Doherty, and M. Capage, in *Biomedical and Biotechnological Advances in Industrial Polysaccharides*, Gordon & Breach, New York, 1989, pp. 145–156.
46. M.P. Marzocca, N.E. Harding, E.A. Petroni, J.M. Cleary, and L. Ielpi, *J. Bacteriol.* **173**, 7519–7524 (1991).
47. F. Katzen, A. Becker, A. Zorreguieta, A. Pühler, and L. Ielpi, *J. Bacteriol.* **178**, 4313–4318 (1996).
48. R. Köplin, W. Arnold, B. Hötte, R. Simon, G. Wang, and A. Pühler, *J. Bacteriol.* **174**, 191–199 (1992).
49. C.L. Wei, N.T. Lin, S.F. Weng, and Y.H. Tseng, *Biochem. Biophys. Res. Commun.* **226**, 607–612 (1996).
50. M.O. Martin, *Res. Microbiol.* **145**, 93–97 (1994).
51. M. Cadmus, C.A. Knutson, A.A. Lagoda, J.E. Pittsley, and K.A. Burton, *Biotechnol. Bioeng.* **20**, 1003 (1978).
52. U.S. Pat. 4,394,447 (1983), M. Cadmus and C.A. Knutson.
53. I.W. Davidson, *FEMS Microbiol. Lett.* **3**, 347–349 (1978).
54. I.H. Smith and G.W. Pace, *J. Chem. Tech. Biotechnol.* **32**, 119–129 (1982).
55. K. Hatakenaka, W. Liu, and T. Norisuye, *Int. J. Biol. Macromol.* **9**, 346–348 (1987).
56. P. Foss, B.T. Stokke, and O. Smidsrod, *Carbohydr. Polymers* **7**, 421–433 (1987).
57. F. Callet, M. Milas, and M. Rinaudo, *Int. J. Biol. Macromol.* **9**, 291–293 (1987).
58. U.S. Pat. 3,966,618 (1976), G.T. Colegrove.

59. Br. Pat. 2,065,689 (1981), T.J. Holding and G.W. Pace.
60. A. Margaritis and G.W. Pace, in M. Moo-Young ed., *Comprehensive Biotechnology*, Pergamon Press, Toronto, 1985.
61. A.-W. Nienow and T.P. Elson, *Chem. Eng. Res. Des.* **66**, 5–15 (1988).
62. R. Kuboi and A.W. Nienow, *Chem. Eng. Sci.* **41**, 123–134 (1986).
63. H. Herbst, Ph.D. Thesis, Technical University of Braunschweig, Germany, 1990.
64. A. Amanullah, L. Serrano-Carreon, B. Castro, E. Galindo, and A.W. Nienow, *Biotechnol. Bioeng.* **57**, 94–108 (1998).
65. A. Zaidi, P. Ghosh, A. Schumpe, and W.-D. Deckwer, *Appl. Microbiol. Biotechnol.* **35**, 330–333 (1991).
66. I.-S. Suh, A. Schumpe, and W.-D. Deckwer, *Biotechnol. Bioeng.* **39**, 85–94 (1992).
67. S.-T. Yang, Y.-M. Lo, and D.-B. Min, *Biotechnol. Prog.* **12**, 630–637 (1996).
68. H. Herbst, I.-S. Suh, H.-U. Peters, A. Schumpe, and W.-D. Deckwer, in *DECHEMA Biotechnology Conferences 3*, VCH Verlagsgesellschaft, 1989.
69. H.U. Peters, A. Zaidi, P. Ghosh, A. Schumpe, and W.-D. Deckwer, *DECHEMA Biotechnology Conferences, Annual Meeting of Biotechnologists*, Frankfurt am Main, Germany, May 28–30, 1993.
70. J.C. Roseiro, A.N. Emery, P. Simoes, F. Estevao, and M.T. Amaral-Collao, *Appl. Microbiol. Biotechnol.* **38**, 709–713 (1993).
71. A.-K. Jana and P. Ghosh, *J. Ferment. Bioeng.* **80**, 485–491 (1995).
72. Eur. Pat. 0112661 (1984), G.T. Banks and P.D. Browning.
73. U.S. Pat. 4,828,321 (1981), W.C. Wernau.
74. L. De Vuyst, J. Van Loo, and E.J. Vandamme, *J. Chem. Tech. Biotechnol.* **39**, 263–273 (1987).
75. H. Funahashi, M. Maehara, H. Taguchi, and T. Yoshida, *J. Chem. Eng. Jpn.* **20**, 16–20 (1987).
76. J.C. Roseiro, D.C. Costa, and M.T. Amaral-Collao, *Lebensm. Wiss. Technol.* **25**, 289–293 (1992).
77. K.V. Rajeshwari, G. Prakash, and P. Ghosh, *Lett. Appl. Microbiol.* **21**, 173–175 (1995).
78. H.-U. Peters, H. Herbst, I.-S. Suh, A. Schumpe, and W.-D. Deckwer, in *Proc. 3rd Int. Workshop Recent Dev. Bacterial Polysaccharides: Biomed. Biotechnol. Adv.* pp. 275–281.
79. I.-S. Suh, H. Herbst, A. Schumpe, W.-D. Deckwer, *Biotechnol. Lett.* **12**, 201–206 (1990).
80. A.B. Metzner and R.E. Otto, Agitation of non-Newtonian fluids. *AICHE J.* **3**, 3–10 (1957).
81. K. Wichterle, M. Kadlec, L. Zak, and P. Mitschka, *Chem. Eng. Commun.* **26**, 25–32 (1984).
82. G. Obernosterer and H.-J. Henzler, in *Rheologie und mechanische Beanspruchung biologischer Systeme*, GVC-VDI Tagung, Tübingen, 1987.
83. P.H. Calderbank and M.B. Moo-Young, *Trans. Inst. Chem. Eng.* **37**, 26 (1959).
84. H.-J. Henzler and J. Kauling, *Proc. 5th Eur. Conf. Mixing*, Paper 30, BHRA Cranfield, 1985.
85. Y. Kawase and M. Moo-Young, *Chem. Eng. Res. Dev.* **66**, 284–288 (1988).
86. P.A. Sandford, J.E. Pittsley, C.A. Knudson, P.R. Watson, M.C. Cadmus, and A. Jeanes, *ACS Symp. Ser.* **45**, 192–210 (1977).
87. L.C. Fryer, F.M. Aramouni, and E. Chambers, *J. Food Sci.* **61**(1), 245 (1996).
88. K.S. Kang and D.J. Pettitt, in *Industrial Gums: Polysaccharides and Their Derivatives*, Academic Press, 1993, pp. 341–371.
89. U.S. Pat. 3,717,452 (1973), K.F. Gibsen and A.W. Saddington.
90. M. Satin, *New Sci.* April, 56–59 (1988).
91. I.W. Sutherland, *Food Biotechnol.* **6**, 75–86 (1992).
92. N.M. Edwards, C.G. Biliaderis, and J.E. Dexter, *J. Food Sci.* **60**(6), 1321–1324 (1995).
93. K. Shimada, H. Okada, K. Matsuo, and S. Yoshioka, *Biosci. Biotechnol. Biochem.* **60**, 125–127 (1996).
94. Business Communications Company, July, 1994, 00039104.
95. *Happy Household & Personal Products Industry*, Oct., 33(10), 152 (1996).
96. *Happy Household & Personal Products Industry*, April, 33(4), 136 (1996).
97. J. Sujjaareevath, D.L. Munday, P.J. Cox, and K.A. Khan, *Int. J. Pharm.* **139**(1–2), 53–62 (1997).
98. D.S. Roy and B.D. Rohera, *Pharm. Res.* **13**(9), 320 (1996).
99. J.D. Ntawukulliyayo, S.C. DeSemedt, J. Demeester, and J.P. Remon, *Int. J. Pharm.* **128**(1–2), 73–79 (1996).
100. A. Kondo, T. Irie, and K. Uekama, *Biol. Pharm. Bull.* **19**(2), 280–286 (1996).
101. U.S. Pat. 3,659,026 (1972), H.R. Schuppner.
102. Business Communications Company, July, 1994, 00039109.
103. *Food Chemical News* 12 Feb. **37**(51), 15 (1996).
104. Business Communications Company, November, 1996, 00065453.
105. The Royal Society of Chemistry, Press release, Oct. 23, 1996.
106. Eur. Pat. 611824 (1994), French Institute for Petroleum.
107. U.S. Pat. 5350528 (1994), Weyerhauser.
108. A. Gabriel, in *Microbial Polysaccharides and Polysaccharases*, Academic Press, 1979, pp. 191–204.
109. M. Lerner, *Chemical Market Reporter*, June 24, **26**, (1996), 12–14 (1996).
110. E.I. Sandvik and J.M. Maerker, in *ACS Symp. Series No. 45, Extracellular Microbial Polysaccharides*, 1977, pp. 242–264.
111. U. Grove and D. Kessel, *Chem. Ing. Tech.* **68**, 4196, 413–417 (1996).
112. J. Klein and A. Schumpe, *Chem. Ing. Tech.* **66**(2), 234–238 (1994).
113. D.G. Kessel, *Status Report*, German Institute for Petroleum Research, 1988, p. 41.
114. M. Hassan-Wakkas, Ph.D. Thesis, Technical University of Clausthal, 1995.
115. *Chemical Week*, International Edition, May 31, **156**(21), 47–48 (1995).
116. M. Lerner, *Chemical Market Reporter*, December 11, **248**, 7–14 (1995).
117. *Chemical Marketing Reporter*, June 27, **245**(26 suppl.), SR16–SR18 (1994).
118. *Chemical Marketing Reporter*, Jan. 3, **245**, 17 (1994).
119. B. Banner, *Chemical Marketing Reporter*, **244**(21), 21 (1993).
120. I. Cintron, *Chemical Marketing Reporter*, **243**(15), 7, 16 (1993).
121. *Informations Chimie Hebdo*, Nov. 18, **1192**, 31 (1994).

122. *Chemical Marketing Reporter*, Dec. 11, **284**(24) 7, 14 (1995).
123. M. Gliksman, G.O. Phillips, D.J. Wedlock, and P.A. Williams, in *Progress in Food and Nutrition Science*, vol. 6, *Gums and Stabilisers for the Food Industry*, Pergamon Press, 1982, pp. 299–321.
124. *Bioprocessing Technology*, July, 4–5 (1990).
125. Business Communications Company, November, 1995, 00054955.
126. G. Moritis, *Oil Gas J*. April, 39–61 (1996).
127. V. Moses, in *Biotechnology, the Science and the Business*, Harwood Academic, 1991, pp. 537–565.

See also CULTURE COLLECTIONS; RHEOLOGY OF FILAMENTOUS MICROORGANISMS, SUBMERGED CULTURE.

YEAST, BAKER'S

TADASHI KOMETANI
Toyama National College of Technology
Toyama, Japan

RYUICHI MATSUNO
Kyoto University
Kyoto, Japan

KEY WORDS

Asymmetric reduction
Bubble-column reactor
Enantioselective oxidation
Ethanol
NAD^+
NADPH regeneration
Optically active compound
Optical resolution

OUTLINE

INTRODUCTION

Baker's yeast consists of microorganisms selected for the production of bread. The history of the baker's yeast industry for breadmaking was briefly described in *Encyclopedia of Food Science and Technology* (1). In the modern baker's yeast industry, specially selected strains of *Saccharomyces cerevisiae* are used. Baker's yeast is able to metabolize the commonly available sugars, such as D-glucose, D-mannose, sucrose, and maltose, very readily under both aerobic and anaerobic conditions.

In 1874, the reducing action of fermenting *Saccharomyces cerevisiae* was observed in which hydrogen sulfide was liberated on addition of finely powdered sulfur to a suspension of fresh baker's yeast in sugar solution. The reduction of furfural to furfuryl alcohol, the first transformation of an organic molecule, was then reported. Since the beginning of the 20th century, numerous biotransformations, biodegradations, and fermentations using other microorganisms (or microbial enzymes) in addition to baker's yeast have been studied by biotechnologists and organic chemists (2–6). Organic synthetic chemists have applied the microbial (or enzymatic) transformation to the synthesis of chiral molecules in enantiomerically pure form, and they have used baker's yeast among microorganisms and lipases among enzymes by preference. The reasons for these choices are that baker's yeast is versatile and inexpensive, and its growth does not require the assistance of a microbiologist. In addition, hydrolytic enzymes do not require cofactor regeneration and are easy to use. Thus, a large number of references to scientific studies in this area have been reported to date.

Microbial transformation of unnatural molecules has been used in industry to prepare optically pure compounds. For example, optical resolution by enantioselective hydrolysis with lipase is used in the synthetic process of a medicine such as diltiazem hydrochloride. Conversely, the baker's yeast-mediated transformation is not industrially applied. However, metabolisms including cofactor regeneration in baker's yeast have been thoroughly investigated in the 1990s. Therefore, the baker's yeast-mediated reduction now has the potential for industrial use.

TYPE OF TRANSFORMATIONS

The scientific studies devoted to baker's yeast-mediated transformation have gradually grown since the 1950s. The transformations of various classes of synthetic organic molecules were examined using baker's yeast as a reagent in organic synthesis, and the number of the papers is now more than 500. Those transformations were reviewed from the viewpoint of chemical reaction in three excellent articles published in 1990, 1991, and 1992, respectively (7–9). In these reviews, the uses of baker's yeast in synthetic organic chemistry are fully covered.

The baker's yeast-mediated transformations are mainly classified into four types as follows.

1. Carbon–carbon bond formation
2. Reduction of ketones and carbon–carbon double bonds
3. Oxidation
4. Hydrolysis of esters

The general goal of these transformations mediated by baker's yeast is considered to be the preparation of chiral molecules in enantiomerically pure form, because the catalytic synthesis of optically pure compounds is one of the most exciting chapters in organic chemistry.

Among the reports on baker's yeast-mediated transformations, more than half of the studies are devoted to the asymmetric reduction of prochiral carbonyl compounds, such as cycloalkanones, aliphatic alkanones, sulfur-containing alkanones, cyclic diketones, acyclic diketones, keto esters, and so on. Chiral alcohols are easily obtained using baker's yeast under mild conditions on a laboratory scale. Asymmetric C–C bond-forming reactions have been investigated by many chemists because of their importance as a fundamental reaction and as an attractive method from a synthetic point in organic chemistry.

Organic chemists sometimes use active dry yeast (certain strains of *S. cerevisiae* having higher resistance to drying with good baking activity [10]) instead of baker's yeast (compressed). Biotransformation mediated by active dry yeast also proceeds in the same manner as that by baker's yeast (compressed); therefore, both biocatalysts are not distinguished in this article.

C–C BOND FORMATION

In 1921, the process for the stereospecific preparation of D-ephedrine was developed using baker's yeast (11). This was the first example of the successful combination of chemical and enzymatic reactions. The key intermediate, L-1-hydroxy-1-phenyl-2-propanone, was produced by treating benzaldehyde under fermenting conditions of baker's yeast in the presence of glucose (Fig. 1). In this reaction, benzaldehyde is stereospecifically condensed with a CH_3CO unit to give the α-hydroxy ketone. The unit comes from decarboxylation of pyruvate through the EMP pathway of glucose and reacts with aldehyde. Hitherto, this is only the industrial use of the baker's yeast-mediated transformation in synthetic organic chemistry.

This acyloin-type condensation with baker's yeast has been applied to other aldehydes and α,β-unsaturated aldehydes. Furthermore, unnatural α-keto acids instead of pyruvic acid were used in this condensation reaction with aldehyde. The α-hydroxy ketones obtained in these reactions were further reduced to afford the *erythro*-diols in more than 95% enantiomeric excess (ee).

Another C–C bond formation mediated by baker's yeast is a Michael-type addition. The reaction of α,β-unsaturated ketones with 2,2,2-trifluoroethanol under fermentative conditions afforded the corresponding trifluoromethyl methyl carbinols with more than 90% ee.

ASYMMETRIC REDUCTION OF PROCHIRAL KETONES

The baker's yeast-mediated reduction of prochiral ketones with varying substituents (Me, Et, *n*-Pr, *n*-Bu, Bz) was systematically investigated in the 1960s. The prochiral ketone, with R_L representing a large substituent and R_S a small substituent adjacent to the carbonyl group, was transformed to the secondary alcohol with an *S* configuration. Therefore, it was judged that a hydrogen transfer to the *re*-face would preferably occur in the baker's yeast cells (Fig. 2). However, it was pointed out that this rule, the Prelog rule (12), should be cautiously applied to baker's yeast-mediated reduction, because exceptions to this rule were found.

In asymmetric reduction, its stereoselectivity is practically most important. Unsatisfactory stereoselectivity would be the consequence of two situations (7). The first case is that a single enzyme, catalyzing the reduction of each substrate, cannot recognize its structural features. The second case is that two or more oxidoreductases for each substrate act simultaneously with opposite stereochemical preferences. Although each of the oxidoreductases catalyzes the enantioselective reduction with high enantioselectivity, stereoselectivity of the transformation and the optical purity of the product are the consequence of the activities of those enzymes. This second case was recently found in baker's yeast-mediated reduction. By characterizing oxidoreductases in baker's yeast followed by modifying the structures or procedures, improvement in the enantiomeric excess of the product was accomplished in some reactions. In this article, the successful result for production of (*S*)-ethyl 3-hydroxybutanoate with high ee is described.

Method

Compared with the chemistry of the baker's yeast-mediated reduction, the method, especially for large-scale production, has hardly been investigated. In the 1990s, regeneration systems for nicotinamide coenzymes in the baker's yeast-mediated reduction were extensively investigated, and consequently, an effective method for large-scale production has been developed (13).

The details of the use of baker's yeast in an asymmetric reduction of prochiral ketones were described on a laboratory scale in *Organic Syntheses*. Reductions of acetol (AC) to (*R*)-propylene glycol (PG) (14), of ethyl acetoacetate (EA) to (*S*)-ethyl 3-hydroxybutanoate (E3-HB) (15), and of 2,2-dimethylcyclohexanone-1,3-dione to (*S*)-3-hydroxy-2,2-dimethylcyclohexanone (16) were fully reported (Fig. 3). Thus, asymmetric reduction mediated by baker's yeast was recognized as a useful technique by synthetic chemists.

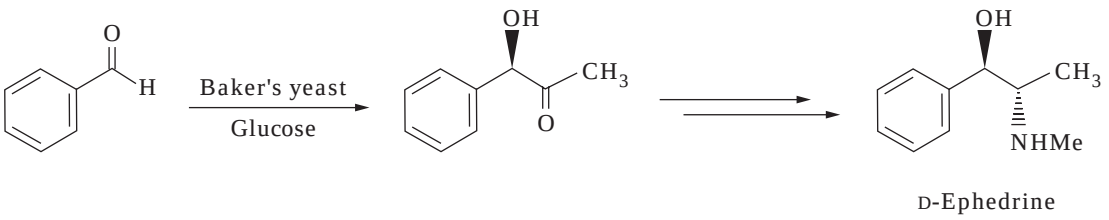


Figure 1. Production of L-ephedrine through baker's yeast-mediated reduction.

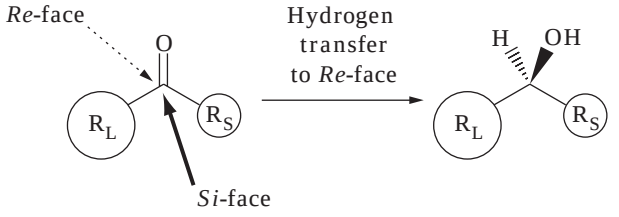


Figure 2. Asymmetric reduction of a prochiral ketone by hydrogen transfer to the *Re*-face.

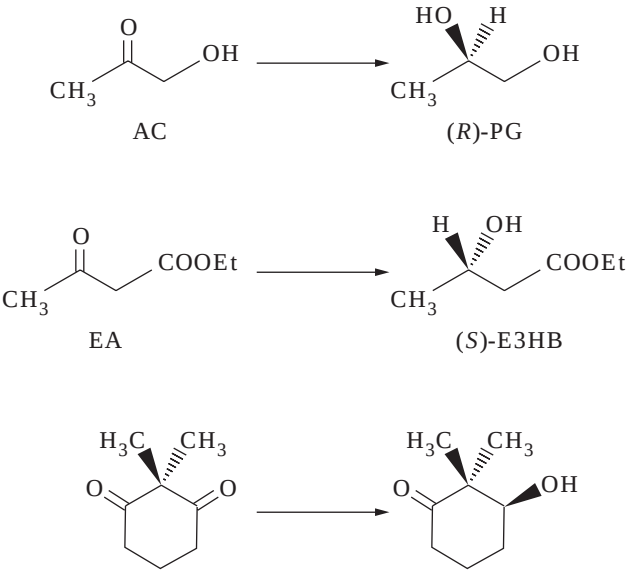


Figure 3. Baker's yeast-mediated reduction fully described in *Organic Syntheses*.

These three bioreductions, and other many similar ones, were done in an aqueous solution containing carbohydrate with stirring, as chemical reactions. Under these conditions, reductions successively proceed consuming carbohydrate, the energy source for coenzyme regeneration. Reduction by carbonyl reductases is coupled with regeneration systems of the reduced form of nicotinamide coenzymes NADPH from the oxidized form $NADP^+$ through glucose oxidation, as shown in Figure 4. However, the use of carbohydrate as the energy source for NADPH regeneration was thought to be inadequate, because the amount of carbohydrate is quite excessive compared with that of the substrate, and the metabolism of the carbohydrate is accompanied by the production of a vigorous foam of carbon dioxide, ethanol, and smelly by-products.

Recently, an efficient method for a baker's yeast-mediated reduction was developed (13). It uses ethanol and oxygen instead of carbohydrate as the energy source for NADPH regeneration. NADPH is regenerated from $NADP^+$ through ethanol oxidation to carbon dioxide, as shown in Figure 5. The asymmetric reduction of several carbonyl compounds by this new method afforded similar results in the chemical yield and optical purity of the product. The method, operating in about 100 mM EtOH under aerobic conditions, is clean and efficient, because no by-product except carbon dioxide was detected, and the energy source can be efficiently used for coenzyme regeneration.

NADPH Regeneration System

In the baker's yeast-mediated reduction, NADPH, which is the hydrogen donor in the reduction, must be successively regenerated from the corresponding oxidized form in the cells. The system responsible for the coenzyme regeneration must be then coupled with reduction of a substrate. Recently, Yamada et al. characterized the enzymes responsible for the NADPH regeneration system using carbohydrate as the energy source with some microorganisms (17). With *S. cerevisiae*, the enzymes were shown to be D-

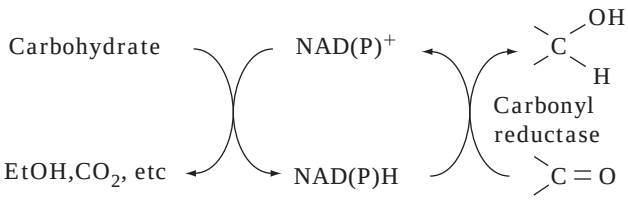


Figure 4. Outline of the reaction mechanism using carbohydrate as the energy source.

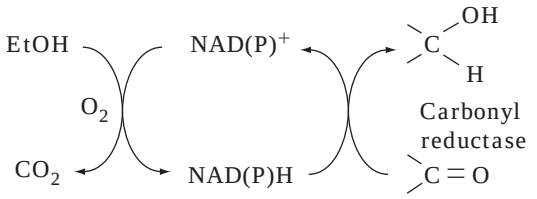


Figure 5. Outline of the reaction mechanism using EtOH as the energy source.

glucose-6-phosphate dehydrogenase and 6-phospho-D-gluconate dehydrogenase in the hexose monophosphate pathway.

In the method using ethanol as the energy source, NADPH is regenerated from $NADP^+$ through the oxidative pathway of ethanol to carbon dioxide, and the system is also efficiently coupled with reduction of a ketone. The enzyme system responsible for NADPH regeneration was investigated by studying the reduction of AC and EA, catalyzed by NADH- and NADPH-dependent carbonyl reductases, respectively. NAD^+ formed during the reduction of AC is reduced to NADH through the oxidative pathway of ethanol to acetate, catalyzed by alcohol dehydrogenase and aldehyde dehydrogenase. On the other hand, $NADP^+$ formed in the reduction of EA is reduced to NADPH through the oxidative pathway of acetate to carbon dioxide. The enzyme responsible for NADPH regeneration must be malate dehydrogenation (decarboxylating) in leaving the TCA cycle into pyruvate.

Asymmetric Reduction of Ethyl Acetoacetate

Reduction of EA by baker's yeast is the most extensively studied transformation, leading to an important chiral building block, (*S*)-ethyl 3-hydroxybutanoate (E3-HB), for the syntheses of the carbapenem thienamycin, (*S*)-(+)-sulcatol, nystatin, and so on. However, the methods described in the literature have been considered to be unacceptable for large-scale production of (*S*)-E3-HB. The reasons are as follows: (*1*) a relatively low concentration, (*2*) tedious isolation and purification processes, and (*3*) a relatively low optical purity. Fortunately, these disadvantages have recently been overcome in developing an efficient system for the baker's yeast-mediated reduction of EA suitable for actual large-scale production.

The first and second disadvantages were simply solved by applying a method using ethanol as the energy source for NADPH regeneration. Because the amount of byproducts inactivating baker's yeast cells was considerably diminished as described, they were active longer and, as a result, the product could accumulate in high concentration.

Regarding the third disadvantage, the reported ee values of the product vary from 70 to more than 95%. To realize those results and improve the ee values, purification and characterization of EA reductases in the baker's yeast cells were very helpful. In 1988, two NADPH-dependent oxidoreductases responsible for the reduction of EA in bakers' yeast were purified and characterized (18). One enzyme, with a K_m value for EA of 0.9 mM, catalyzes the reduction to give (*S*)-E3-HB, and the other, with a K_m value for EA of 17.0 mM, to give (*R*)-E3-HB. Therefore, it was supposed that the ee values might easily vary because of the condition of the initial concentration of EA. The former enzyme would mainly participate in the reduction at a low EA concentration, affording high optical purity. Actually, the optical purity was dramatically improved by applying a fed-batch operation of EA to afford more than 99% ee (*S*)-E3-HB (13).

Bubble-Column Reactor

The baker's yeast-mediated reduction of EA and AC were effectively achieved on a large scale using a bubble-column reactor, which had advantages such as decreasing the operating costs due to low energy requirements. Because the aeration rate is the most important factor in this reactor, its effect on the reduction rate was examined. In the case of EA, the specific reduction rate increased according to the increase in the aeration rate, then became constant. Conversely, the specific reduction rate of AC was almost constant regardless of the aeration rate. These results were consistent with coenzyme requirements and their described regeneration mechanisms.

The large-scale production of (*S*)-E-3HB and (*R*)-PG was performed by applying the fed-batch operation for the addition of the substrates and ethanol in a bubble-column reactor containing 10 L of medium (Table 1) (13). Compared with the results by the conventional procedure described in *Organic Syntheses*, the yield and the optical purity were markedly improved, and the procedure could be applied on an industrial scale (14,15). (*R*)-PG is actually used as a chiral building block for preparation of (*S*)-ofloxacin, although it is now industrially produced by catalytic hydrogenation using a chiral *Ru* complex.

Diltiazem Synthesis

Diltiazem hydrochloride, (2*S*,3*S*)-3-acetoxy-5-[2-(dimethylamino)ethyl]-2,3-dihydro-2-(4-methoxyphenyl)-1,5-benzothiazepin-4(5*H*)-one hydrochloride, is a representative calcium channel blocker that has been clinically used as an antianginal and antihypertensive agent in more than 100 countries. In the synthetic process, the key step is the

Table 1. Comparison between EtOH Procedure and the Conventional Procedure Described in *Organic Synethese*

	EtOH procedure (EA reduction)	Conventional (Acetol reductin)
Water	1,000 mL	1,000 mL
Baker's yeast	112 g (77 g)	112 g (100 g)
Energy source	104 g of EtOH (192 g of sucrose)	18 g of EtOH (100 g of sucrose)
Substrate	60 g (15.3 g)	40 g (10 g)
Air	25 L/min	3 L/min
Reaction time	62 h (74 h)	38 h (72 h)
Temperature	30 ° (r.t.)	30 °C (32 °C)
Recovery	40.5 g (10.6 g)	22 g (5.5 g)
Optical rotation	+44.3° (+37.2°)	−17.3° (−15.0°)
ee	99.3% (85%)	98.2% (90%)

introduction of two chiral centers. Optical resolution of a racemic intermediate, lipase-catalyzed enantioselective hydrolysis of methyl 3-phenylglycidate, is now industrially used. Recently, a new efficient synthetic process involving asymmetric reduction by baker's yeast as a key step was reported. The transformation of a racemic ketone, (*RS*)-2-(4-methoxyphenyl)-1,5-benzothiazepin-3,4(2H,5H)-dione afforded only the desired (2*S*,3*S*)-2,3-dihydro-3-hydroxy-2-(4-methoxyphenyl)-1,5-benzothiazepin-4(5*H*)-one (Fig. 6). The difficulty in the industrial application because of its low productivity, caused by low solubility of the substrate in water, was also overcome by changing its physical property. The crystalline substrate was first converted to an amorphous-like solid aggregate by pouring a DMF solution of the substrate into water. After separation, the aggregate was used for the reduction under the usual conditions. Thus, 100 g of the substrate in 1 L of a reaction medium was reduced to produce the chiral intermediate with more than 99% ee in 80% yield (19).

ENANTIOSELECTIVE OXIDATION

Baker's yeast is widely used as a reducing agent in organic synthesis for preparing enantiomerically pure chiral alcohols. One enantiomer can be conveniently prepared by this method, but the other is essentially impossible to prepare with the same system. Several methods for controlling the enantioselectivity of the baker's yeast-mediated reduction have been reported; modifying the structure of the substrate, the reaction conditions, and adding a third reagent. These methods are very interesting but not convenient for large-scale production. Therefore, baker's yeast-mediated oxidation was developed to prepare the antipode of the reduction product on a large scale.

(*R*)-1,2-Alkanediols were found to be easily oxidized to 1-hydroxy-2-alkanones with baker's yeast under aerobic conditions, and (*S*)-1,2-alkanediols, conversely, could not be oxidized under the same conditions. Thus, treatment of racemic 1,2-alkanediols with baker's yeast to oxidize (*R*)-1,2-alkanediols to 1-hydroxy-2-alkanones followed by separation of the 1-hydroxy-2-alkanones afforded the (*S*)-1,2-alkanediols (Fig. 7) (13).

NAD^+ Regeneration

In this biooxidation, NADH-dependent oxidoreductases participating in the asymmetric reduction of 1-hydroxy-2-alkanone catalyze the contrary enantioselective oxidation, and NAD^+ is regenerated from NADH in the yeast cells for the successive oxidation. A preliminary experiment showed that the oxidation did not proceed under anaerobic conditions. Thus, an outline of the oxidation, especially in relation to the condition of oxygen supply, was examined by analyzing the transformation of (*R*)-PG to AC (13). The amounts of oxygen consumed and AC produced increased with increasing $k_L a$, and the ratio was about 1.7 (AC mmol/O_2 mmol). Because 2 mol (*R*)-PG would be theoretically oxidized using 1 mol of oxygen, as shown in Figure 8, this result suggests that oxygen is effectively used during the biooxidation (85% efficiency).

Optical Resolution of 1,2-Alkanediols

Treatment of racemic 1,2-alkanediols with baker's yeast affords (*S*)-1,2-alkanediols and 1-hydroxy-2-alkanones that are easily separated because of a considerable difference in their volatilities. The asymmetric reduction of 1-hydroxy-2-alkanones with baker's yeast affords (*R*)-1,2-alkanediols (13). Thus, the optical resolution of racemic 1,2-alkanediols was conveniently accomplished using only baker's yeast, as summarized in Figure 9.

The oxidative resolution and the asymmetric reduction have been achieved using the bubble-column reactor on a large scale. The specific oxidation rate is sufficiently fast and reaches equilibrium within 24 h, affording the (*S*)-1,2-alkanediols with 79% ee. The aqueous layers after removal of (*S*)-1,2-alkanediols are directly used for baker's yeast-mediated reduction, affording the (*R*)-1,2-alkanediols with over 98% ee.

CHARACTERIZATION OF OXIDOREDUCTASES

There have been several reports on purification and characterization of enzymes from baker's yeast that catalyze the reductions of unnatural substrates. The purposes are to understand the mechanism of the stereochemistry of baker's yeast-mediated transformation and, as a result, to improve the ee of the reduction product by some modifications. Furthermore, some purified reductases would be useful for synthetic purposes.

First, Sih et al. isolated three oxidoreductases that catalyze the reduction of alkyl 4-chloroacetoacetate (20). They reported that stereochemical control by modification of the substrate stems from the difference in relative activities of the D- and L-enzymes depending on the length of the carbon chain in the substrate. Furuichi et al. purified 2-

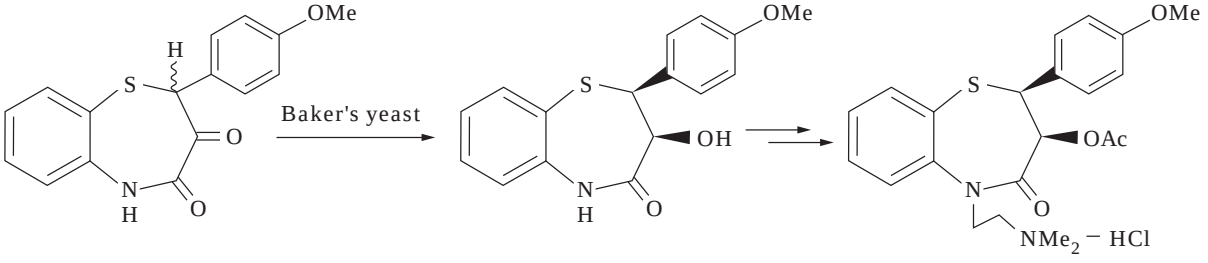


Figure 6. Production of diltiazem hydrochloride through baker's yeast-mediated reduction.

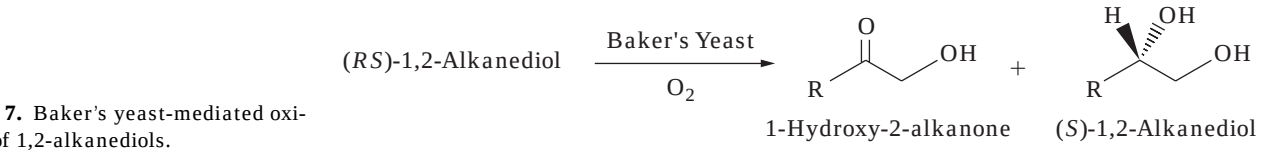


Figure 7. Baker's yeast-mediated oxidation of 1,2-alkanediols.

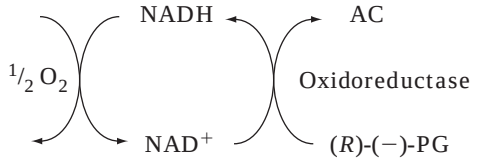


Figure 8. Outline of the reaction mechanism of the yeast-mediated oxidation.

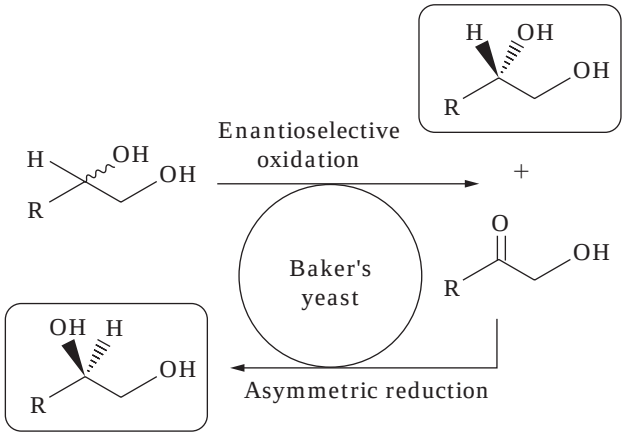


Figure 9. Oxidative resolution of racemic, 1,2-alkanediols with baker's yeast.

methyl-3-oxobutanoate reductase (21) and Heidlas et al. purified two oxidoreductases that catalyze β-, γ-, and δ-keto esters (18). The latter report was very useful for improvement of the ee of the product by baker's yeast-mediated reduction of EA as described. Nakamura et al. isolated four enzymes that catalyze β-keto- and α-alkyl-β-ketoesters and three α-ketoester reductases (22).

IMMOBILIZATION

Immobilized baker's yeast, using polyurethane, calcium alginate, or κ-carrageenan, has been applied to change the reaction conditions (8). Some marked differences have been reported in the stereochemistry, mostly due to a change in the concentration of the substrates. In some cases, immobilization of baker's yeast improved the ee of the products. The immobilized baker's yeast was shown to be active for several weeks, longer than free baker's yeast. However, reductions using immobilized baker's yeast gave lower yields as compared with the analogous reduction with free baker's yeast. Thus, the immobilization did not improve the efficiency of the baker's yeast-mediated transformation for large-scale production.

BIBLIOGRAPHY

1. Y.H. Hui, *Encyclopedia of Food Science and Technology*, Wiley, New York, 1993, pp. 2885–2887.
2. J.B. Jones, C.J. Sih, and D. Perlman, *Application of Biochemical Systems in Organic Chemistry*, Parts 1 and 2, Wiley, New York, 1976.
3. K. Kieslich, *Biotransformations*, Verlag Chemie, Weinheim, Germany, 1984.
4. H. Yamada and S. Shimizu, *Angew. Chem. Int. Ed. Engl.* **27**, 622–642 (1988).
5. C.-H. Wong and G.M. Whitesides, *Enzymes in Synthetic Organic Chemistry*, Pergamon Press, Oxford, 1989.
6. J. Halgas, *Biocatalysts in Organic Synthesis*, Elsevier, Amsterdam, 1992.
7. S. Servi, *Synthesis*, 1990, 1–25, (1990).
8. R. Csuk and B.I. Gläzer, *Chem. Rev.* **91**, 49–97 (1991).
9. E. Santaniello, P. Ferraboschi, P. Grisenti, and A. Manzocchi, *Chem. Rev.* **92**, 1071–1140 (1992).
10. Y.H. Hui, *Encyclopedia of Food Science and Technology*, Wiley, New York, 1993, pp. 2887–2888.
11. C. Neuberg and J. Hirsch, *J. Biochem. Z.* **115**, 282 (1921).
12. V. Prelog, *Pure Appl. Chem.* **9**, 119–130 (1964).
13. T. Kometani, H. Yoshii, and R. Matsumo, *J. Mol. Catal.* **1**, 45–52 (1996).
14. P.A. Levene and A. Walti, *Org. Synth. Coll.* **2**, 545–547 (1943).
15. D. Seebach, M.A. Sutter, R.H. Weber, and M.F. Zuer, *Org. Synth.* **63**, 1–7 (1983).
16. K. Mori and H. Mori, *Org. Synth.* **68**, 56–61 (1989).
17. M. Kataoka, Y. Nomura, S. Shimizu, and H. Yamada, *Biosci. Biotech. Biochem.* **56**, 820–821 (1992).
18. J. Heidlas, K.-E. Engel, and R. Tresal, *Eur. J. Biochem.* **172**, 633–639 (1988).
19. T. Kometani, Y. Sakai, H. Matsumae, T. Shibatani, and R. Matsuno, *J. Ferment. Bioeng.* **84**, 195–199 (1997).
20. W.-R. Shieh, A.S. Gopalan, and C.J. Sih, *J. Am. Chem. Soc.* **107**, 2993–2994.
21. A. Furuichi, H. Akita, H. Matsukira, T. Oishi, and K. Horikoshi, *Agric. Biol. Chem.* **49**, 2563–2570 (1985).
22. K. Nakamura, S. Kondo, Y. Kawai, N. Nakajima, and A. Ohno, *Biosci. Biotech. Biochem.* **58**, 2236–2240 (1994).

See also ANAEROBES; ENERGY METABOLISM, MICROBIAL AND ANIMAL CELLS; FOOD PROCESS ENGINEERING; PROCESS CONTROL, STRATEGY AND OPTIMIZATION; PROCESS MONITORING.

CONVERSION FACTORS, ABBREVIATIONS, AND UNIT SYMBOLS

SI UNITS (Adopted 1960)

The International System of Units (abbreviated SI), is being implemented throughout the world. This measurement system is a modernized version of the MKSA (meter, kilogram, second, ampere) system, and its details are published and controlled by an international treaty organization (The International Bureau of Weights and Measures).

SI units are divided into three classes:

BASE UNITS		SUPPLEMENTARY UNITS	
length	meter† (m)	plane angle	radian (rad)
mass	solid angle	steradian (sr)	kilogram (kg)
time	second (s)		
electric current	ampere (A)		
thermodynamic temperature‡	kelvin (K)		
amount of substance	mole (mol)		
luminous intensity	candela (cd)		

Quantity	Unit	Symbol	Acceptable equivalent
volume	cubic meter	m^3	
	cubic diameter	dm^3	L (liter) (5)
	cubic centimeter	cm^3	mL
wave number	1 per meter	m^{-1}	
	1 per centimeter	cm^{-1}	

In addition, there are 16 prefixes used to indicate order of magnitude, as follows:

Multiplication factor	Prefix	Symbol
10^{18}	exa	E
10^{15}	peta	P
10^{12}	tera	T
10^{9}	giga	G
10^{6}	mega	M
10^{3}	kilo	k
10^{2}	hecto	h[a]
10	deka	da[a]
10^{-1}	deci	d[a]
10^{-2}	centi	c[a]
10^{-3}	milli	m
10^{-6}	micro	€
10^{-9}	nano	n
10^{-12}	pico	p
10^{-15}	femto	f
10^{-18}	atto	a

[a]Although hecto, deka, deci, and centi are SI prefixes, their use should be avoided except for SI unit-multiples for area and volume and nontechnical use of centimeter, as for body and clothing measurement.

For a complete description of SI and its use the reader is referred to ASTM E380.

A representative list of conversion factors from non-SI to SI units is presented herewith. Factors are given to four significant figures. Exact relationships are followed by a dagger. A more complete list is given in the latest editions of ASTM E380 and ANSI Z210.1.

†The spellings "metre" and "litre" are preferred by ASTM; however, "-er" is used in the *Encyclopedia*.

‡Wide use is made of Celsius temperature (t) defined by

$$t = T - T_0$$

where T is the thermodynamic temperature, expressed in kelvin, and $T_0 = 273.15$ K by definition. A temperature interval may be expressed in degrees Celsius as well as in kelvin.

CONVERSION FACTORS TO SI UNITS

To convert from	To	Multiply by
acre	square meter (m^2)	4.047×10^{3}
angstrom	meter (m)	1.0×10^{-10}†
are	square meter (m^2)	1.0×10^{2}†
astronomical unit	meter (m)	1.496×10^{11}
atmosphere, standard	pascal (Pa)	1.013×10^{5}
bar	pascal (Pa)	1.0×10^{5}†
barn	square meter (m^2)	1.0×10^{-28}†
barrel (42 U.S. liquid gallons)	cubic meter (m^3)	0.1590
Bohr magneton ($€_B$)	J/T	9.274×10^{-24}
Btu (International Table)	joule (J)	1.055×10^{3}
Btu (mean)	joule (J)	1.056×10^{3}
Btu (thermochemical)	joule (J)	1.054×10^{3}
bushel	cubic meter (m^3)	3.524×10^{-2}
calorie (International Table)	joule (J)	4.187
calorie (mean)	joule (J)	4.190
calorie (thermochemical)	joule (J)	4.184†
centipoise	pascal second (Pa s)	1.0×10^{-3}†
centistokes	square millimeter per second (mm^2/s)	1.0†
cfm (cubic foot per minute)	cubic meter per second (m^3/s)	4.72×10^{-4}
cubic inch	cubic meter (m^3)	1.639×10^{-5}
cubic foot	cubic meter (m^3)	2.832×10^{-2}
cubic yard	cubic meter (m^3)	0.7646
curie	becquerel (Bq)	3.70×10^{10}†
debye	coulomb meter (C m)	3.336×10^{-30}
degree (angle)	radian (rad)	1.745×10^{-2}
denier (international)	kilogram per meter (kg/m)	1.111×10^{-7}
	tex‡	0.1111
dram (apothecaries')	kilogram (kg)	3.888×10^{-3}
dram (avoirdupois)	kilogram (kg)	1.772×10^{-3}
dram (U.S. fluid)	cubic meter (m^3)	3.697×10^{-6}
dyne	newton (N)	1.0×10^{-5}†
dyne/cm	newton per meter (N/m)	1.0×10^{-3}†
electronvolt	joule (J)	1.602×10^{-19}
erg	joule (J)	1.0×10^{-7}†
fathom	meter (m)	1.829
fluid ounce (U.S.)	cubic meter (m^3)	2.957×10^{-5}
foot	meter (m)	0.3048†
footcandle	lux (lx)	10.76
furlong	meter (m)	2.012×10^{-2}
gal	meter per second squared (m/s^2)	1.0×10^{-2}†
gallon (U.S. dry)	cubic meter (m^3)	4.405×10^{-3}
gallon (U.S. liquid)	cubic meter (m^3)	3.785×10^{-3}
gallon per minute (gpm)	cubic meter per second (m^3/s)	6.309×10^{-5}
	cubic meter per hour (m^3/h)	0.2271
gauss	tesla (T)	1.0×10^{-4}
gilbert	ampere (A)	0.7958
gill (U.S.)	cubic meter (m^3)	1.183×10^{-4}
grade	radian	1.571×10^{-2}
grain	kilogram (kg)	6.480×10^{-5}
gram force per denier	newton per tex (N/tex)	8.826×10^{-2}
hectare	square meter (m^2)	1.0×10^{4}†
horsepower (550 ft · lbf/s)	watt (W)	7.457×10^{2}
horsepower (boiler)	watt (W)	9.810×10^{3}
horsepower (electric)	watt (W)	7.46×10^{2}†
hundredweight (long)	kilogram (kg)	50.80
hundredweight (short)	kilogram (kg)	45.36
inch	meter (m)	2.54×10^{-2}†
inch of mercury (32°F)	pascal (Pa)	3.386×10^{3}
inch of water (39.2°F)	pascal (Pa)	2.491×10^{2}

CONVERSION FACTORS TO SI UNITS

To convert from	To	Multiply by
kilogram-force	newton (N)	9.807
kilowatt hour	megajoule (MJ)	3.6†
kip	newton (N)	4.448×10^{3}
knot (international)	meter per second (m/S)	0.5144
lambert	candela per square meter (cd/m^3)	3.183×10^{3}
league (British nautical)	meter (m)	5.559×10^{3}
league (statute)	meter (m)	4.828×10^{3}
light year	meter (m)	9.461×10^{15}
liter (for fluids only)	cubic meter (m^3)	1.0×10^{-3}†
maxwell	weber (Wb)	1.0×10^{-8}†
micron	meter (m)	1.0×10^{-6}†
mil	meter (m)	2.54×10^{-5}†
mile (statute)	meter (m)	1.609×10^{3}
mile (U.S. nautical)	meter (m)	1.852×10^{3}†
mile per hour	meter per second (m/s)	0.4470
millibar	pascal (Pa)	1.0×10^{2}
millimeter of mercury (0°C)	pascal (Pa)	1.333×10^{2}†
minute (angular)	radian	2.909×10^{-4}
myriagram	kilogram (kg)	10
myriameter	kilometer (km)	10
oersted	ampere per meter (A/m)	79.58
ounce (avoirdupois)	kilogram (kg)	2.835×10^{-2}
ounce (troy)	kilogram (kg)	3.110×10^{-2}
ounce (U.S. fluid)	cubic meter (m^3)	2.957×10^{-5}
ounce-force	newton (N)	0.2780
peck (U.S.)	cubic meter (m^3)	8.810×10^{-3}
pennyweight	kilogram (kg)	1.555×10^{-3}
pint (U.S. dry)	cubic meter (m^3)	5.506×10^{-4}
pint (U.S. liquid)	cubic meter (m^3)	4.732×10^{-4}
poise (absolute viscosity)	pascal second ($Pa \cdot s$)	0.10†
pound (avoirdupois)	kilogram (kg)	0.4536
pound (troy)	kilogram (kg)	0.3732
poundal	newton (N)	0.1383
pound-force	newton (N)	4.448
pound force per square inch (psi)	pascal (Pa)	6.895×10^{3}
quart (U.S. dry)	cubic meter (m^3)	1.101×10^{-3}
quart (U.S. liquid)	cubic meter (m^3)	9.464×10^{-4}
quintal	kilogram (kg)	1.0×10^{2}†
rad	gray (Gy)	1.0×10^{-2}†
rod	meter (m)	5.029
roentgen	coulomb per kilogram (C/kg)	2.58×10^{-4}
second (angle)	radian (rad)	4.848×10^{-6}†
section	square meter (m^2)	2.590×10^{6}
slug	kilogram (kg)	14.59
spherical candle power	lumen (lm)	12.57
square inch	square meter (m^2)	6.452×10^{-4}
square foot	square meter (m^2)	9.290×10^{-2}
square mile	square meter (m^2)	2.590×10^{6}
square yard	square meter (m^2)	0.8361
stere	cubic meter (m^3)	1.0†
stokes (kinematic viscosity)	square meter per second (m^2/s)	1.0×10^{-4}†
tex	kilogram per meter (kg/m)	1.0×10^{-6}†
ton (long, 2240 pounds)	kilogram (kg)	1.016×10^{3}
ton (metric) (tonne)	kilogram (kg)	1.0×10^{3}†
ton (short, 2000 pounds)	kilogram (kg)	9.072×10^{2}
torr	pascal (Pa)	1.333×10^{2}
unit pole	weber (Wb)	1.257×10^{-7}
yard	meter (m)	0.9144†

†Exact.